Foundations of Biology

Samuel M. Scheiner and Kayla I. Scheiner

Copyright (c) 2014 by S. M. Scheiner and K. I. Scheiner. This work is made available under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 license (http://creativecommons.org/licenses/by-nc-nd/3.0/)

Front cover
American egret, Gainsville, Florida.
Photograph by Samuel Scheiner
Dedication

To Judy Scheiner, wife and mother.
Brief Contents

Chapter 1: Life ................................................................. 11
Chapter 2: Science .......................................................... 45
Chapter 3: Genetics ......................................................... 72
Chapter 4: Evolution ....................................................... 110
Chapter 5: Cells .............................................................. 144
Chapter 6: Organisms ..................................................... 179
Chapter 7: Ecology ......................................................... 230
Contents

Chapter 1: Life ................................................................................................ 11
  The Study of Life ......................................................................................... 11
    The persistence of life
    Box 1A - The Chemistry of Life ................................................................. 16
  What is Life? ............................................................................................ 14
    Life as an open system
    Life as bonded units
    Information
    Variation
    Complex systems
    Emergent properties
    Contingency
    Persistence and change
    New cells from old
    Box 1B - Louis Pasteur .............................................................................. 26
  The Origin of Life ...................................................................................... 29
    The problem
    Complex molecules
    Replication first
    Metabolism first
    Reconciliation
    Box 1C - Critical Experiment: The Precursors of Life ............................... 30
    Box 1D - Antonio Lazcano ......................................................................... 34
  Life, the Universe, and Everything .............................................................. 39
    Is it alive?
    Nonterrestrial life
    Box 1E  - Carl Edward Sagan ..................................................................... 41

Chapter 2: Science ........................................................................................... 45
  Ways of Knowing ...................................................................................... 45
    Alternatives to science
    Biology’s special assumptions
  How We Know .......................................................................................... 48
    The philosophy of science
    Theories and models
    Testing theories
    Science and other ways of knowing, revisited
    Multicausality
    Explaining the characteristics of life
    Box 2A - Aristotle .................................................................................... 50
    Box 2B - The Nature of Theory ................................................................. 53
    Methodologies .......................................................................................... 56
    Box 2C - Francis Bacon ............................................................................. 58
    Box 2D - Ronald Aylmer Fisher ................................................................. 61
      Experiments
      Creativity, objectivity, and subjectivity
  Science as a Social Activity ........................................................................ 67
    Peer review
Ethics

When is it not misconduct?
Responsibilities for ethical conduct

Chapter 3: Genetics ........................................................... 72

The Basics of Genetic Structure and Duplication
Box 3A - Critical Experiment: The Structure of DNA
The Theory of Genetics
Resemblance of Relatives ........................................................ 78
Discrete traits
Continuous traits
Fidelity of information transmittal
Box 3B - Gregor Johann Mendel ........................................ 83
Box 3C - A Simple Genetic System and the Resemblance of Relatives .......... 88
Box 3D - Barbara McClintock .............................................. 91
New Information ................................................................. 89
   Capability for new information
   Mutation
   Exchange and recombination
   Random processes
Box 3E - Thomas Hunt Morgan ........................................... 98
Information Usage ........................................................... 102
   Transcription and translation
   Direction of information flow
   Context dependency
Evolution of the Information System ................................... 108
   What is a Gene?

Chapter 4: Evolution .......................................................... 110

The Process of Evolution .................................................... 110
The Theory of Evolution .................................................... 111
Box 4A - Charles Darwin ................................................... 113
Patterns of Evolution ......................................................... 116
   Change over generations
   Speciation
   Common descent
What is a Species? ........................................................... 120
Mechanisms of Evolution .................................................. 122
   Gradual processes
   Variation
Box 4B - Critical Experiment: Predicting Transitional Forms ................. 124
Natural selection ............................................................. 128
   Contingency
Box 4C - Models and Theories ............................................ 131
Box 4D - Theodosius Grygorovych Dobzhansky .......................... 133
Box 4E - Motoo Kimura ..................................................... 138
Creationism ................................................................. 140

Chapter 5: Cells ............................................................... 144
Cells as Foundational Units ............................................. 145
Chapter 7: Ecology ................................................................. 230
  The Theory of Ecology .......................................................... 231
  Distributions and Interactions ................................................. 232
    Uneven distributions
    Biotic and abiotic interactions
    Contingency
  Box 7A - Fredric Edward Clements ........................................ 235
  Box 7B - Robert Helmer MacArthur ....................................... 242
  Variation ................................................................................. 247
    Individuals
    Environmental conditions
    Finite resources
  Box 7C - Critical Experiment: Large-scale Experimental Manipulations ...... 251
  Box 7D - Rachel Louise Carson ................................................ 255
  Fitness .................................................................................... 258
    Birth and death
    Evolution
  What is a Community? ........................................................... 261
Preface

This book is a different kind of introductory biology textbook that is suitable for all classes, majors and non-majors alike. It takes a different approach to teaching biology by deliberately eschewing the typical bewildering array of detail and focusing on core concepts. The central conceit of this book is the slogan of the Bauhaus architectural movement: less is more. This book is emphatically NOT meant to be encyclopedic.

If you are a biology major, this book will equip you with a conceptual framework that you can use all of your other courses. If you are a non-major, you will gain a fundamental understanding of biology that you can relate to your own field of study. Towards these ends, the book is theory oriented. The theories prioritize the big ideas of biology, so that you will know where to focus, what you need to remember, and help you see the connections among the parts of biology. A theoretical allows you to see how knowledge comes about; it makes biological knowledge into something dynamic, and not received wisdom. To determine the list of fundamental concepts, we took as our guiding philosophy: “If I were teaching an upper-level course in topic X, what would I want my students to know about all of the other topics that I am not teaching.”

The book avoids the very human-centric approach of nearly all introductory textbooks, an approach which greatly distorts your appreciation for the diversity of life. Our approach is to take a topic and discuss how it applies to all species. For example, we discuss the structure of cells of bacteria, animals and plants together. While we use human examples throughout the book in order to engage the student, we make it clear that humans are just one of many species. We emphasize that to properly understand human biology, you must appreciate biological science in its entirety. The book avoids, as much as possible, the specialized vocabulary of biology, instead presenting the concepts using plain English. We limit the use of specialized vocabulary to concepts that are used repeatedly in the book or are otherwise ones that a student is likely to need to understand topics in more than one advanced class.

Each chapter is presented within the framework of a unifying theory, focusing on the fundamental principles of those theories. The concepts are placed within a historical context which shows science as a dynamic process and the result of the activities of real people, rather than as a set of facts handed down from on high. To this end, each chapter features brief biographies of three important scientists. To relate the importance of the scientific process in gaining understanding, each chapter features a Critical Experiment.

Using the book

Throughout this book, we explore the broad properties that make up living systems, the characteristics of all or very many of those systems, rather than focusing on the numerous details and variation. There is certainly an important need to understand those details. However, you must first understand the overarching framework of life and the science of biology so that you have a context within which to place the details. In this chapter we will examine the fundamental characteristics of living systems. What separates life from non-life? How would we recognize life elsewhere in the universe if we came across it?

Those fundamental characteristics form the groundwork upon which are built theories that form our understanding of specific aspects of living systems. In the second chapter we look at the structure of theories and how different pieces fit together to create the science of biology. For now, we will just say that the fundamental characteristics are those that both are necessary for the specific theories, and are properties that arise
in each of those theories. They are also characteristics that supercede the details of the carbon-based life found on the Earth and, thus, can be used as a way of identifying non-terrestrial living systems.

Once we lay the groundwork for the characteristics of life and how we study it, the rest of the book looks at life from five points of view: genetics, evolution, cells, organisms, ecology. Each point of view examines an aspect of the structure of living systems and the processes responsible for that structure. Each point of view represents a discipline within the science of biology. We emphasize that science is ultimately unitary. All parts connect to all other parts. Thus, you will find themes that cut across all of the disciplines, some of the fundamental characteristics of life. Some of the most exciting new research happens at the boundaries of disciplines when ideas from different disciplines are given a chance to mingle. By focusing this book on the broad properties of living systems, we hope that you will be better able to see how those parts connect.

The chapters in this book can be approach in many different orders. The material in each chapter is largely independent of the others, with a few caveat. Chapter 2 (Science) can be read at any point. Chapter 1 (Life) should be read before chapters 3-7. Chapter 3 (Genetics) should be read before Chapter 4 (Evolution). Chapter 5 (Cells) should be read before Chapter 6 (Organisms).

You will note a lack of citations throughout the book. This absence is not meant to convey the impression that these concepts are simply given truths. Quite the contrary, they are the result of the work of very many people over decades. To give proper credit to all of the ideas and facts would require a thicket of citations that would obscure the text. Instead, we leave it to follow-up courses to provide those details, along with all of the others that we are leaving out.

We are releasing this book under a Creative Commons license. All of the text is copyrighted by the authors. The images are either the property of the authors, given to them for use in this book, or in the public domain. You are free to copy and share this book as much as you wish. Our one request is that you provide feedback on the positive and negative aspects of the book so that we may improve future editions.

Many people commented on various drafts and versions of this book including: Michael Barfield, Judith Bronstein, Robin Bush, James Collins, Richard Kliman, Gordon Fox, George Gilchrist, Gabriel Herrick, Norman Johnson, Milton Muldrow, Adam Porter, Andrew Sinauer, Alan Tessier, Saran Twombly, Michael Willig, Grace Wyngaard, William Zamer, as well as several anonymous reviewers. All helped improve its content and for that we thank them.
To the reader:

This book is an ongoing project. Any help that you can give to improve the book would be appreciated. First, please let us know if you find an errors of fact, spelling errors, typos, or format problems. Second, we are aware that many of the figures are not ideal. We would appreciate it if you have, or can point us to, a better alternative; keep in mind that the alternative must be available under a Creative Commons license. Finally, any general feedback about the content of the book is welcome. If you are a teacher, we would greatly like to know how you are using it in your classes.
Chapter 1
Life

Introduction
The fundamental concepts of biology presented in this book are a continually evolving set of ideas, and only skim the surface of biological knowledge. Biology can be an intimidating subject to begin to study. It can seem like the entire body of human scientific knowledge is being thrown at you at once - beyond even the basics of biology, you’re learning chemistry, physics, mathematics, and statistics as well. It can be very difficult to really make sense of the deluge of information – to fit all of those seemingly disparate facts into a cohesive framework. With this book, our aim is not to stuff you with every iota of knowledge in the discipline of biology, but to give you that basic framework with which you can build a deeper, more comprehensive understanding of biology, life, the scientific process. We present the basic concepts that you will need as you begin your studies of the biological sciences, without going into so much detail at once that the underlying ideas are lost. It is our hope that you will find this book to be a useful tool in laying the groundwork for your future studies, or even simply building a basic understanding of biology if your academic path leads elsewhere. All biological knowledge is related, and to understand any one piece requires knowing its relationship to the whole.

The Study of Life
It might seem incredible, given the amazing diversity, complexity, and pervasiveness of life, but the very fact that life exists at all is both remarkable and unexpected. After all, from a simple consideration of the laws of physics, it seems as if life should not exist. The reason that it does, as we will see in this chapter, is that the universe contains the capacity for complexity, and it is from this complexity that life has arisen.
Biology is the study of life in all of its diversity and complexity. Why should anyone want to learn about this complexity? The typical answer is that knowledge of biology will help to cure human diseases and feed the hungry. Those are certainly worthy goals, but should not be the entire answer. Humans are just one of millions of species on this planet (Figure 1.1), and a somewhat unusual one at that. To understand living systems means widening one’s perspective to all species, and such a widened perspective has several advantages.

First, all of life is connected through a complex web of direct and indirect interactions, and even to solve human-centric problems requires knowledge of the entire living world. This need is most obvious for problems like global warming that involve changes in our environment. It can also be seen, however, by realizing that problems of human health are often caused by environmental factors such as pollution, or by infectious diseases that exist primarily in non-human animals (e.g., avian influenza).
Second, all of life is related through a chain of descent to a single, common ancestor. Because of this, there are some aspects of human biology that are only fully understood by learning about both similar and differing aspects of other species. Because of the close relationships between and interconnectedness of all life on earth, to understand any one piece of biological knowledge requires knowing its relationship to the whole.
For example, understanding why human natal development from a single cell to a fully-formed individual happens the way it does requires understanding the structure of the most primitive animals.
Third, all of life is fascinating. Biology offers a chance to explore a myriad of exciting questions. The pace of scientific investigation of living systems has been accelerating over the past half century, with new technologies giving us the ability to answer long-standing questions as well as raising new ones. What is the basis of consciousness and are other animals conscious? Why are there more species in equatorial regions of the Earth and what will happen to them as global climates change? Why do salamanders
have much more genetic material in each cell than humans have? How do all of the parts of a single cell manage to work together? How many different kinds of single-celled organisms are there? What might life elsewhere in the universe look like?

The science of biology is how we will answer those questions. Science is a method for gaining understanding about the world; it explains the vast and complex systems of the natural world. In this book you will learn about how science operates, especially in studies of living systems. You will learn about the common properties of all living systems, the fundamental concepts of biology. By the end of this book, you will have a broad understanding of living systems that will allow you to place further learning into an overarching framework consisting of a series of general theories.

We start with a theory for all of life. In its most basic sense, life can be described as a variety of systems, and the easiest system to consider is a living organism such as yourself. A living organism has a cohesiveness and a wholeness that is easy to understand, although it is not always so easy to define what, exactly, an individual is. Nearly everything in life can be described and studied as systems – for example, in addition to an individual as a whole, the cells and organs that make up that individual are systems, and populations of individuals act as systems as well. Throughout this book, especially in this chapter, we will talk about the characteristics of living systems, which apply to everything from a single cell to the collection of the entirety of life on Earth.

The general theories of biology, and the chapters in this book, center around six broad questions (Table 1.1). The processes responsible for the persistence of life are those that form the answers to all of these questions. All of these questions are intertwined, to understand the answer to one you need to understand the others. This book will start you on the road to that understanding.

<table>
<thead>
<tr>
<th>Table 1.1 The six fundamental questions in biology.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Why does life manage to persist? (Chapter 1)</td>
</tr>
<tr>
<td>2. Why do offspring resemble their parents? (Chapter 3)</td>
</tr>
<tr>
<td>3. What is the cause of organismal change and diversity? (Chapter 4)</td>
</tr>
<tr>
<td>4. How does a cell maintain its structure and function? (Chapter 5)</td>
</tr>
<tr>
<td>5. How does an individual maintain its integrity? (Chapter 6)</td>
</tr>
<tr>
<td>6. What explains the distribution and abundance of organisms? (Chapter 7)</td>
</tr>
</tbody>
</table>

The persistence of life

The first, and possibly central, characteristic of life that we need to explain is its remarkable persistence. Life originated on the Earth possibly as long ago as four billion years, about a half billion years after the Earth itself coalesced from rocks, dust, and gases, and has been around ever since. At a much shorter time scale, individuals persist from moment to moment. Yet this persistence seems at odds with some of the basic laws of physics; the second law of thermodynamics says that in a closed system order must break down into disorder. Life not only persists as ordered systems, but has led to even more complex and diverse life, seemingly creating order out of disorder. The solution to this apparent contradiction is to realize what physicists mean when they speak of order and disorder.

What the laws of thermodynamics really mean is that the universe is proceeding from the Big Bang – the entire universe tied up in one very high energy point – to the final heat death when all of that energy has spread out across infinite space and exists in the lowest possible energy state. The universe as a whole must run down from high
energy to low because it is a single, closed system. A living system, on the other hand, does not decay this way because it is not a closed system. A living organism takes in matter and energy, uses it, and then expels matter and energy again. Ultimately the matter taken in must balance the matter going out, while the energy that comes in goes from a higher state (more ordered) to a lower state (less ordered). In thermodynamic terms, life persists by creating local pockets of order at a cost of creating more disorder somewhere else. In this way it can be said that life, by perpetuating itself, is (infinitesimally) speeding up the heat death of the universe.

**What is Life?**

As persistence is one of the fundamental characteristics of living systems, it is important to understand what the term means as applied to biology. Persistence is not just that living systems simply endure. A rock can endure without changing – living systems continuously change. While change does not necessarily guarantee persistence, a lack of change guarantees death or extinction. However, persistence is more than just simple change. Living systems also must be robust, which means that they are able to persist and to thrive for very long times in the face of a changing environment. It is this robust persistence that the science of biology attempts to explain.

The theory of biology (Table 1.2) attempts to explain why living systems are persistent. The first two fundamental principles of that theory describe the basic characteristics of living systems. The next six principles describe the causes and consequences of living systems. The last two principles explain where living systems come from. In this chapter, we examine the broad nature of these aspects of living systems, and in Chapters 3 through 7 we see how they play out in more specific properties of living systems. Chapter 2 will look at the process of science, including the nature of scientific theories like this one.

**Life as an open system**

Life consists of open, non-equilibrial systems (Table 1.2, principle 1). Open simply means that living systems take in matter and energy and eventually release them again, although the form of that matter and energy is changed. In living systems, matter (in the form of either simple atoms or complex molecules) has two basic roles: it is used for making the structures of living systems, and it is the carrier for energy in the form of chemical bonds (Box 1A). Living systems use energy to move matter around and build its structures. In doing so, the energy gets changed from higher states to lower states. Keep in mind, though, that this change from high to low is the net change of all of the energy that comes into the system. Some of the energy may be retained in a high-energy state (e.g., the ordered structures built by the organism), at a cost of transforming even more energy to a low state that passes out of the system. This explains why life must be an open system in order to exist; the openness of living systems is what allows them to maintain themselves.

Non-equilibrial means that living systems consist of ordered structures in a universe that otherwise tends toward disorder. A living system stands out from the nonliving universe around it, in that life persists despite all of the stress and trauma that an individual experiences. For life to persist this way, that order must be actively maintained, and thus living organisms are continually building and repairing themselves. This active maintenance of whole organisms is the result of all of the activities that go on within the systems in the organism. In turn, the active maintenance of individuals results in the persistence of the systems of which those individuals are a part.

By active maintenance, we do not mean that the maintenance is consciously
directed. While we consciously maintain our automobiles in running order by continually checking if they need more oil and by having them repaired when parts start to fail, living systems maintain themselves because they have autonomic feedback mechanisms. For example, many plants have leaves that change orientation so that light will more fully fall upon their surfaces. The light triggers some of the cells in the leaf's stem to take in water and expand, while others expel water and contract, thus bending or twisting the stem. This process stops when the maximal amount of light is hitting the leaf, thereby maximizing the leaf's ability to turn sunlight into stored energy (see Chapter 5), all without conscious direction.

Table 1.2 The fundamental principles of the theory of biology.

1. Life consists of open, non-equilibrial systems that are persistent.
2. Life consists of bounded units that are capable of metabolism and reproduction.
3. Life requires a system to store, use, and transmit information.
4. Living systems vary in their composition and structure at all levels.
5. Living systems consist of complex sets of interacting parts.
6. The complexity of living systems leads to emergent properties.
7. The complexity of living systems creates a role for contingency.
8. The persistence of living systems requires that they are capable of change over time.
9. Living systems come from other living systems.
10. Life originated from non-life.
Box 1A
The Chemistry of Life

Living systems are built from a diverse set of chemical structures. These molecules, sometimes referred to as biomolecules, are built from atoms of carbon (C), hydrogen (H), oxygen (O), and nitrogen (N), along with smaller amounts of phosphorus (P) and sulfur (S) and occasionally other trace elements. These molecules can be classified into four broad types: proteins, sugars, lipids, and nucleic acids. All are built by taking smaller units and combining them into large, complex structures.

The basic unit of proteins is the amino acid (Fig. 1A.A). There are 20 common amino acids, with the unit labeled R in the diagram to the left representing a variety of possible structures that give the amino acid many of its chemical properties. They are strung together in various ways and then folded into complex 3-dimensional shapes. Proteins can act as vital structural elements, especially flexible ones such as muscles and skin. Their other important role is to direct and enhance chemical reactions; when they act in this capacity, they are called enzymes.

Sugars can come as single units, such as the molecule glucose, as pairs of units, such as the molecule sucrose, and as long chains of units. When they occur in single or double units, sugars are important sources of energy. In living systems, energy is stored as chemical bonds, with different types of bonds containing different amounts of energy. The bonds between single and double sugars are easy to break, which releases their energy, making it readily available to living organisms. As long chains of units, sugars are important structural units, especially when those structures need to be rigid. Wood, for example, is composed of long chains of sugar molecules.

Like sugars, lipids can act as both energy storage units and structural units. Because lipids are very stable, they are often used for long-term energy storage. As structural units, lipids form the basis of the envelope that surround cells and the parts of cells.
Nucleic acids are the basic unit of information storage. They exist in units called nucleotides that consist of the nucleic acid, a sugar molecule, and one or more phosphate groups. The chemical bonds between the phosphate groups contain a lot of energy, so nucleic acids also are important short-term energy storage units. However, these bonds are not very stable, so their role is that of a shuttle of energy from one part of the cell to another, rather than long-term storage.
Life as bounded units

Life maintains itself as pockets of order in a disordered universe. The foundational unit of those pockets are cells (Table 1.2, principle 2). The cell holds together the complex machinery of life along with the energy needed to power that machinery. In its most fundamental sense, a cell is an envelope or a bubble that acts to hold in matter and energy. Each cell contains high concentrations of the ions and molecules that are essential for life to exist, arranged in very specific, highly ordered structures. The envelope itself has a specific chemical make-up (Figure 1.2) that allows ions and molecules to pass in and out. That passage in and out of the cell is actively regulated, part of the way that life maintains itself.

While all organisms are composed of cells, the nature of that cell varies. Cells can be simple or complex (see Figure 5.1). Some organisms consist of a single cell, while others consist of multiple cells (Figure 1.1). For some multicellular organisms, the divisions between individual cells can break down so that the organism itself acts like a single, giant cell (see Figure 6.3).

The recognition that cells form the fundamental unit of life slowly developed during the 17th, 18th, and 19th centuries and is an example of how technical innovation helps to drive scientific progress. Much of the early understanding of cells as fundamental units was dependent on the development of better and better microscopes and microscopic techniques such as stains that allowed parts of cells to be more easily distinguished. The term “cell” comes from the English scientist Robert Hooke; in 1665, he was peering through one of the earliest microscopes at a sample of cork, and the regular divisions he saw there reminded him strongly of the rooms, called cells, that monks lived in.

The first living cells were described in 1674 by the Dutch scientist Antonie van Leeuwenhoek who did much to refine early microscopes. van Leeuwenhoek primarily observed single-celled organisms. It would take another century and a half to clearly establish that much larger organisms, such as plants and animals were also made up of cells. In the early 19th century, the science of life and cellular structure advanced quickly, beginning with the French scientist Henri Milne-Edwards making careful observations of muscles of frogs and other animals and clearly describing their cellular make up.

Ludolph Christian Treviranus, a German scientist, proposed that plant tissues consisted of individual, separable cells, and all of this led the French scientist Henri Dutrochet to perform experiments, published in 1824, in which he broke apart plant tissues into cells that could be clearly distinguished and described. In that publication, Dutrochet declared that the cell is the fundamental unit of all life. This dictum makes up one of the two principles of the classic cell theory, the other being that all cells come from other cells (see below).

Information

In order for the ordered complexity of life to exist and persist, it has to contain and perpetuate information. To understand why this is, consider a random collection of letters [tbtesornoThottunoobsehtiqateei]. In such a format, it conveys nothing. However, put those letters into a specific order [To be or not to be, that is the question], and suddenly you have one of the most recognizable quotes in the history of English theater. The same principle holds true for living organisms; to maintain themselves, they must have a way to capture - and use - the information contained in their ordered complexity (Table 1.2, principle 3). We explore the details of the information system in Chapter 3. Here we just present its overall characteristics, especially those characteristics that would likely be found in the information system of any form of life.
In organisms on the planet Earth, the primary means of storing information is as long molecules of deoxyribonucleic acid (DNA). As a language is built up from individual letters to words, sentences and whole paragraphs, so too is the information of life built up from individual nucleotides to strings of DNA code. Just as only 26 letters in the English language can be combined to form tens of thousands of English words, a key feature of information systems is the ability to build up complex formations from just a few basic elements, in this case just four types of nucleotides (see Chapter 3).

The information in those DNA molecules constitutes the instruction manual for how an organism operates. But like any instruction manual, it is useless if it is not read and the information turned into actions. In order to put the information contained in DNA into action, organisms contain a large number of other molecules that read the information and turn it into the central machinery of the cell that in turn builds and maintains the organism.

Besides being used as an initial construction guide or blueprint, the information must be maintained internally as an organism grows, one cell dividing into two, and passed along as an individual reproduces the next generation. This transmittal of information, whether as part of growth or reproduction, starts with making a copy of the information. Again, a large set of molecules does the copying, producing a new DNA molecule that matches the original. Thus, the machinery that the information codes for in turn is responsible for reproducing the machinery. Such interdependence of information and the structures that come from it, however, immediately raises the question of how the whole system got started.

An essential piece of the answer is found in one aspect of the transmittal system. Elaborate mechanisms exist to try to ensure that the information is passed along correctly. However, changes in the information can occur. The transmittal system is not perfect, and the changes caused by that imperfect copying of information can sometimes be disastrous, resulting in the death of an offspring or cells that turn cancerous. Sometimes, however, those changes prove beneficial and ultimately lead to new characteristics of living systems.
Variation

We can look at life from many different perspectives. One common framework is referred to as levels of organization (Figure 1.3), in which molecules are first put together to make cells, groups of cells become tissues and organs, and tissues and organs form individuals. A group of individuals make up a population or species, and a set of species co-existing in an area constitute a community or ecosystem. Multiple communities spread across a landscape become a biome, and the sum total of the biomes on the Earth is our biosphere.

![Figure 1.3](image)

One way to organize our view of life is as a hierarchy beginning with molecules within cells, going up to whole organisms, and then organisms within larger units until it reaches the level of all of life on the Earth. (Created by Mikala14, Source: Wikipedia)

One fundamental property of living systems is that there is a lot of variation at all of these levels (Table 1.2, principle 4). The basic units of a DNA molecule vary along its length and from one individual to another, a cell consists of many different kinds of molecules, and the individuals within a population differ from each other. Some communities are dominated by trees and others by grasses, and contain a wide array of animal life. Thus, differences in composition create variation from one place to another, whether that is from one side of a cell to the other or from one continent to another.

Just as important as its variation, life is dynamic, changing from one moment to the next. A molecule within a cell may, from second to second, alter how it is folded in three dimensions or in its interactions with other molecules or cellular structures. A multicellular organism begins life as single cells; as it grows and develops it divides into additional cells which differ in structure and function. This dynamic sets the stage for life's complexity, which in turn leads to emergent properties and creates a role of contingency. Life's dynamism is also critical for the persistence of living systems. All of this variation creates both the challenge and excitement of the science of biology.

Complex systems

Compared to the largest structures in the universe, the stars and galaxies, the simplest, microscop-ic single-celled organism is vastly more complex with many more types of components that are arranged in a much more complicated fashion. A hallmark of living systems is that they are complex (Table 1.2, principle 5), meaning they are made of up many different kinds of parts, which interact with each other in many different ways. This complexity can be found at any level you look, from an individual molecule, to a single cell, to a whole organism, to the entire Earth. Such complexity
is one way that living systems differ from much of the nonliving world.

It seems obvious that an individual is made up of many different kinds of parts. A human body has arms and legs, eyes and ears, a heart and lungs. A closer look at one of those parts shows that it is also made up of different pieces, which in turn are made up of cells that contain many subsystems. Even single molecules are often large and built up from many pieces (Box 1A). While it is necessary to know the many different kinds of building blocks and how they are combined in order to understand how organisms work, that is a level of detail that is unnecessary for now. For now it is only necessary to understand that from a few simple pieces, many different kinds of molecules can be built because of the many different ways those pieces can be combined.

While the parts of living systems interact with each other, not everything interacts with everything else. You do not interact with every other person in the world, or even in your neighborhood. You have a small circle of friends with whom you communicate on a regular basis, as well as a set of family members you may see every day, but with whom you communicate very differently than your friends. Then there are other people who you may only see once a month or once a year. This type of variation in how you interact can also be found in other types of organisms and in other aspects of living systems - such as the molecules within a cell - and is one aspect of the complexity of living systems. This variation in amounts of interaction creates modules, parts of a system within which components interact a lot with each other, and much less than other parts of the system.

**Emergent properties**

The concept of emergent properties can be summarized by an old adage: “The whole is more than the sum of its parts.” An emergent property is one that is found at a certain level of organization due to properties, structures, and processes at that level of organization based on the interactions among the more basic parts. The complex relationships that characterize biological systems give rise to emergent properties, including the very existence of life – and intelligent life – itself (Table 1.2, principle 6).

Emergent properties can be contrasted with properties that are merely aggregates of properties at a lower hierarchical level. A simple example is that the properties of a water molecule are not simply the properties of hydrogen and oxygen atoms taken together; the molecule has its own unique properties that result from the way the hydrogen and oxygen atoms interact. For a biological example, consider the locomotion of a human. For a human to walk down the street requires the existence of many different parts—bones, muscles, connective tissues, and so forth—arranged in a particular three-dimensional configuration, and all operating together in a certain way. The separate parts cannot move on their own, thus movement is an emergent property of the whole organism.

Such emergent properties can be found at all levels of living systems. The function of a protein depends on the sequence of amino acids and how that chain is folded together into a precise three-dimensional shape (Box 1A). Cells function by separating and concentrating molecules into particular subsections (see Chapter 5). Emergent properties mean that we cannot simply measure the properties of individual parts and add them up to get the properties of the whole. Consider the amount of CO2 taken up by all of the trees in a forest; we cannot measure the uptake rate of an entire forest just by measuring the uptake rates of the individual trees under average conditions. The total uptake rate depends on how individuals interact in several ways, including shading one another, interfering with wind, and competing for the available CO2. The total uptake rate is an emergent property of the community of trees.
ing shading one another, interfering with wind, and competing for the available CO2. The total uptake rate is an emergent property of the community of trees. Emergent properties are the result of feedbacks within living systems so that the behavior or properties of one part affect another which in turn affects the first part. An example of feedback is what happens when you touch a very hot object. The hot object begins to damage your finger, and a pain signal is sent to your central nervous system. It then signals the muscles in your arm and hand to stop touching the object. Finally, your hand moves away from the hot object and further damage is halted. On a larger scale, take the case of the arctic hare and its predator, the Canada lynx (Figure 1.4). As the number of hares goes up, the number of lynx goes up a year or two later. Eventually, the lynx are so numerous that they are eating hares faster than the hares can reproduce. This causes the number of hares to go down, which in turn causes the lynx to begin to starve and have fewer offspring, until the number of hares being eaten is fewer than the number being born. Then the cycle starts all over again. This cycling of numbers is an emergent property of the feedbacks between the two species.

Contingency

Contingency is the combined effects of two processes—randomness and a sensitivity to initial conditions. Contingency is the occurrence of chance events – you happen to meet another person at a party – and the changes that can happen from that – you end up getting married. The complexity of living systems means that contingency can play an important role in the properties of those systems (Table 1.2, principle 7). Many processes in living systems are subject to random events; for example, how information
gets passed from one generation to the next. Molecules within cells move around, in part, due to random jostling by other molecules. The dynamic nature of living systems is one factor that allows randomness to play a role.

What is meant by randomness, though, is not always obvious and is a concept that physicists and philosophers debate. First, random does not mean equally likely. Consider a 6-sided die; an “honest” die has an equal chance of landing on any of the six sides because the mass of the die is evenly distributed. A “loaded” die changes that likelihood by subtly altering the mass distribution. In a loaded die, out of 600 rolls, the number 1 might come up 120 times, rather than 100 times. Yet, in a single roll any of the six numbers might come up, and which number does appear is guided by random chance.

What determines that random chance? When you flip a coin it might seem as if coming up heads or tails were random. Yet it is possible to build a machine that can flip a coin exactly the same way every single time, giving the coin exactly the same force so that it spins exactly the same number of times and always land heads up. Much harder would be building a machine that spins a coin on a flat surface so that when it stops spinning it lands on the same side every time. It is much more likely that the coin would land heads up half the time and tails up the other half. The difference between flipping and spinning the coin is that the coin will spin around many, many more times than the number of flips it makes. A flipping machine can be built that always has the coin make 10 rotations in the air before it hits the ground. But a spun coin will go around tens or hundreds of times. Which way the coin happens to be tilting at the exact moment it starts to slow down can vary due to tiny differences in the surface or the tiny currents in the air around it. If such factors could be precisely controlled every time, then the outcome would be predictable. Some philosophers argue that such a system is not truly random because precise control is possible in theory. From a real world perspective, however, such causes of variation have the same effect as if they were random. Even very tiny differences in the amount of spin that the machine gives the coin will change the outcome. Such differences in initial spin is an example of sensitivity to initial conditions, which is sometimes referred to as the butterfly effect, as in “a butterfly flapping its wings in Tokyo can change the weather in New York City.” Living systems exhibit such sensitivity in many ways. For example, the amount of nutrition that a human fetus receives or its exposure to certain chemicals when it is in the mother’s womb can change the course of development, resulting in a baby being born with deformed limbs that can never grow properly, or defects in its internal organs. At the other end of the scale in space and time, 65 million years ago a huge asteroid slammed into the Earth, leading to the extinction of the last of the dinosaurs. If, by chance, that asteroid had missed the Earth, humans might never have evolved. Even our existence today is the result of contingency.

Emergent properties are one cause of sensitivity to initial conditions. The way that molecules function is due to the precise way that they are folded into three-dimensional structures. It is so easy for us to die from accident, disease, or violence because very small changes to our bodies (e.g., a small hole in a vital organ) can disrupt its intricate relationships. Changes in the amount of carbon dioxide released into the atmosphere in the United States through the burning of coal is leading to changes in the amount of rainfall in Central America, which can change the rainforests into deserts and result in the extinction of thousands of species. Thus, the ways that systems interact can magnify effects, making those systems very sensitive to changes.

The importance of contingency differs among the various aspects of living systems. Random events are continually at play when it comes to how the information system passes along information from one generation to another (see Chapter 3). On
the other hand, contingency usually has little effect on how cells function (see Chapter 5). The reason why contingency plays a big role in one instance and not in another is the effect of large numbers. Consider what happens when you start flipping a coin. If you flip a coin just two times, the odds that it will come up heads once and tails once is only 50%. Half the time, you will get either all heads or all tails. If you flip that same coin 20 times, the chance of getting all heads is approximately 1 in 1,000,000, while there is a 95% chance of getting between 28% and 72% heads, or about 6-14 heads. If you flip the coin 2000 times, there is a 95% chance of getting between 48% and 52% heads, or 960-1040. The larger the number of flips, the more likely you will get close to half heads and half tails. The functioning of cells is like flipping a coin very many times. Cells contain both a large number of molecules of a single type as well as many different types, all interacting with each other. Any single molecule moves randomly, but the movements of all of those molecules average out so that the total movement is very predictable. In contrast, an individual may only have a few offspring, so random differences in how information is passed along become very important. There are other instances where contingency plays important roles in living systems, as we will see in Chapters 4 and 7.

 Persistence and change

A hallmark of living systems is that they are dynamic and constantly changing. This change occurs over many different time scales, from moment to moment, within the life of an individual, and from one generation to the next. The ability to change is vital to the existence of living organisms; life itself could not persist if living systems were incapable of change (Table 1.2, principle 8).

An individual has to continually change to survive. In the simplest sense, you have to continuously take in new matter by breathing, eating, and drinking, to replace the energy that you are continually using and parts that get worn out. Perhaps counterintuitively, change in one part of the system creates stability in other parts. For example, mammals and birds tend to maintain a constant body temperature by having an active system for converting the energy locked up in food into heat. That energy conversion means that the cells are constantly creating and breaking chemical bonds – constantly changing those parts of the organism – which results in the temperature of the entire body remaining unchanged.

Individual change can be a response to a changing environment. In regions of the world that experience freezing temperatures for part of the year, many organisms enter some sort of period of low metabolic activity such as bears hibernating through the winter. A living organism might derive a myriad of beneficial results from such a period of reduced activity; it could make it much easier for a mammal to pass through a period when food is scarce (a hibernating bear doesn’t need to eat very much!). It can also be a way for an individual to avoid being frozen, or it can be a way to prepare an individual, such as a tree, to survive exposure to freezing temperatures. For example, many trees change the chemical make up of their cells so as to lower the temperature at which ice crystals form, thereby avoiding damage.

Change is also important for the survival of a lineage over much longer periods of time, such as from one generation to the next. While most offspring contain the same information as their parents, sometimes mutations occur, changes in information. Some of those changes in information lead to changes in the characteristics of organisms, such as a child being a little smaller or larger than its parent, an offspring using food a little more efficiently, or the flowers on a new plant being a little more attractive to bees and less attractive to butterflies. If the individuals that differ from their parents end up leaving more offspring than the others, the collective characteristics of that group of individuals will have changed. It is this change that is called evolution.
Because the world keeps changing, if organisms did not evolve, they would eventually cease to exist. Since life arose some 3 billion years ago, the sun has waxed and waned in intensity, the length of the day has increased from 15 hours to 24 hours, and the concentrations of oxygen and carbon dioxide in the atmosphere have increased, decreased, and increased again. Glacial epochs have come and gone. And all that is just the nonliving world! The existence of life brings its own changes, with the advent of new species that will eat others, or compete with them for scarce resources.

All these constant changes mean that while an organism may be doing perfectly fine one day, it might find survival more and more difficult as conditions change, until the conditions become so different that the organism cannot survive at all. But if individuals in a group differ from each other, some may do better than others in those new conditions. Their offspring will survive better and have more offspring of their own. Thus, organisms with the capacity to change over generations, to evolve, are the ones that persist. Change does not guarantee persistence, but a lack of change guarantees death or extinction.

New cells from old

For change to occur from one generation to the next, there has to be continuity of living systems. That continuity embodies two principles, that living systems come from other living systems (Table 1.2, principle 9) and that, on the whole, those new living systems are extremely similar to the ones from which they come.

While it may seem obvious today that life comes from other life, that principle was not enshrined in biology until the middle of the 19th century. While it was certainly very clear that at least some life came from other life – birds laid eggs that hatched out new birds, humans gave birth to other humans – the life cycles of very many organisms had not been recorded, and many were not easy to observe. For example, it was thought that maggots, the larval stage of flies, would arise out of meat that was left out to decay. The notion that life could arise from nonlife was called spontaneous generation. As early as the 17th century, scientists set about testing this idea. The Italian scientist Francesco Redi performed the first experimental test of spontaneous generation, specifically the idea that maggots would form out of the material of rotting meat. He set up eight jars with raw meat placed inside. One set of four jars was simply left open, while the tops of the other four were covered with a fine gauze that let in air, but not insects. After several days he found maggots only in those jars that were left open. He let the maggots mature and saw that they became flies. Finally, when he placed the flies in jars with a piece of meat, new maggots were found. This experiment and others like it gave rise to the saying “Omne vivum ex ovo”, which translates to “all life comes from other life” (literally, “from an egg”). The discovery of cells as the fundamental unit of life during this same period was an important further step. Based on the many studies of the cellular basis of life that were made in the first half of the 19th century, the French scientist François-Vincent Raspail put forth the dictum “Omnis cellula e cellula” (“every cell originates from another existing cell”). This idea was later broadly promoted by the German doctor and scientist Rudolf Virchow and became the other cornerstone of cell theory. However, these studies still left open the possibility that microscopic, single-celled organisms might arise by spontaneous generation. The final nail in the coffin came about through the experiments of Louis Pasteur who showed that even microorganisms had to arise from other individuals (Box 1B).

As is often the case in science, solving one conundrum created a new one. If all living systems come from other living systems, where did life itself come from? We answer that question in the next section.
Louis Pasteur is most widely known for lending his name to the process of heating liquids such as milk to eradicate the bacteria and mold spores they contain, an invention which every dairy farmer and schoolchild with a milk carton can be thankful for. Born on December 27, 1822 in Dole in the Jura region of France, Pasteur spent his childhood in Arbois, where he later had his house and laboratory (this structure still stands today as a museum). The bulk of his work was dedicated to research into the causes of diseases and their prevention, his experiments supporting germ theory and his breakthroughs vastly reducing the mortality of rabies, puerperal fever (also called childbed), and cholera. Together with Ferdinand Cohn and Robert Koch, Pasteur was one of the founders of the science of microbiology. Young Pasteur’s aptitude was recognized by his college headmaster, who recommended he apply to the École Normale Supérieure. He served briefly as the professor of physics at Dijon Lycee in 1848, then later became a professor of chemistry at Strasbourg University, where he met and courted the daughter of the university rector. He and Marie Laurent were married on May 29, 1849.

In addition to his work in the biological sciences, Pasteur also made discoveries in chemistry, most notably solving a problem concerning the nature of a compound called tartaric acid. While a solution of the compound derived from biological material rotated the plane of polarization of light passing through it, the same solution derived via chemical synthesis let light pass through unchanged. This presented a very bizarre problem; the solutions themselves are identical, down to their elemental composition and chemical reactions, and thus should not display any variance in structure. Pasteur’s examinations of crystals of sodium ammonium tartrate revealed that the crystals come in two asymmetric forms, mirror images of one another. Sorting the crystals by hand, he found two forms of the compound in question; solutions of one rotated polarized light counterclockwise, while solutions of the other rotated it clockwise, and an equal mix of the two types had no effect on light at all. From this, he deduced that the molecule was asymmetric and could exist in two different forms that are mirror images of each other; while the synthesized version had both structures, the natural version had only one. This was the first demonstration of the existence of such molecules.

By far, Pasteur’s most well-known contributions fall into the realm of germ theory. While Pasteur wasn’t the originator, he performed experiments that helped support and develop it, and helped deal the death blow to the notion of spontaneous generation. He demonstrated that the fermentation processes is caused by the growth of microorganisms, which were not spontaneously generated out of food, but from previous individuals. In his most famous experiment, he placed boiled broths in glass vessels, some with long, curved necks that were open to the air but would prevent dust from falling into the broth, some with filters that had the same effect, and some with neither, which were open and exposed to the outside world. Nothing grew in the broths that had the long necks and filters, while the exposed broths quickly spoiled. This meant that the microorganisms growing in the spoiled broths came in from the outside, a fact which both supported germ theory and disproved spontaneous generation. Some of those sealed vessels still exist and nothing has grown in them in nearly 150 years. His investigations also led to the development in 1862 with his colleague Claude Bernard of a process for
removing the microorganisms that live in fermenting beverages by heating the liquid, a process eventually named pasteurization in his honor.

The results of his experiments with broths led Pasteur to conclude that microorganisms also infect animals and humans, causing illness. He proposed that by preventing these microorganisms from entering the body, illness could be avoided. These conclusions would greatly influence Joseph Lister to develop his antiseptic methods in surgery, which have saved countless lives. Pasteur’s own work with diseases lead to the discovery of a new and more effective method of producing vaccines; while a smallpox vaccination had been in use since 1796, using a weak form of cowpox to inoculate humans, it was necessary with that method to locate a naturally weak form of the disease to use as an inoculum. Pasteur, while working with chicken cholera, exposed a group of experimental chicks to a batch of the microbes that had gone bad and lost its potency. The chicks developed only mild symptoms, and when Pasteur tried to re-infect them with a fresh batch of cholera, none of them got sick. Despite being weakened, the original bacteria had inoculated the chickens. In the 1880s, Pasteur applied this method of inoculation to cattle, exposing an anthrax bacillus to an oxidizing agent, potassium dichromate, effectively creating a vaccine. Because that technique had been developed by rival Jean-Joseph-Henri Toussaint, Pasteur claimed he had exposed the bacteria to oxygen; while he did develop a technique using oxygen, it wasn’t until after he had secured a patent on an anthrax vaccine. Pasteur’s artificially produced vaccines had a huge impact on the medical field, and today humanity is nearly free from many deadly diseases, including measles, scarlet fever, and polio.

Pasteur also produced a rabies vaccine by growing the virus in rabbits, then weakening it by drying the affected nerve tissue. While a vaccine for rabies was first created by Emile Roux, a doctor and colleague of Pasteur, Pasteur himself first used it on 9-year-old Joseph Meister on July 6, 1885, after the boy had been badly mauled by a rabid dog. This was done at considerable personal risk for Pasteur, who, as he wasn’t a licensed doctor, could have faced prosecution if anything went wrong. However, the vaccine worked; Meister didn’t develop the disease, and Pasteur was lauded as a hero. This success led to the manufacture of many other vaccines, and the first of the Pasteur Institutes were founded on the basis of his achievement.

Pasteur’s achievements are the result of his careful attention to detail in his experimental work. He famously said, “Dans les champs de l’observation le hasard ne favorise que les esprits préparés” (In the fields of observation chance favors only the prepared mind). He was recognized in 1895 with the award of the Leeuwenhoek Medal, microbiology’s highest achievement, but died that same year near Paris from complications arising from a series of strokes that started in 1868. He was originally buried in Notre Dame, but was later reinterred in a crypt at the Institut Pasteur in Paris.
The Origin of Life

The problem

If life always comes from life, if new cells always come from old cells, how did it all get started? At some point, life had to arise out of nonlife (Table 1.2, principle 10). Although it might seem that way, this statement about the origin of life does not contradict the statement that living systems come from other living systems. We are speaking about two different aspects of life, its origin and its continuity. The key to understanding the coexistence of these two ideas is to realize that the conditions that gave rise to life are no longer found on the Earth, and life itself would prevent them from re-occurring. The origin of life created an interesting juxtaposition: life first arose from nonlife, thereby ensuring that from that point on, life had to come from already existing life.

The key difficulty in explaining how life originated is that even the simplest, single-celled organism consists of a complex system consisting of many different parts. We can divide those parts into three broad categories: the information storage system (DNA and its associated molecules), the information usage system (the rest of the cellular machinery), and a wrapper to keep all of the pieces together. Replication of a cell requires an intricate set of interactions and feedbacks between all of these pieces. The information storage system codes for the cellular machinery, which in turn is responsible for duplicating the information system.

Answering this basic question of the origin of life has become a very active area of research in the past decade or so. While definitive answers do not yet exist, several theories are gaining in strength. Those theories are gathered under two broad headings: replication first and metabolism first. As implied from the names, each set of theories champions a start of life from a different part of the cellular complex: the information storage system and the cellular machinery.

Complex molecules

Life on Earth is mostly built from just a few common elements—carbon (C), hydrogen (H), oxygen (O), and nitrogen (N)—along with some phosphorus (P) and sulfur (S), and a smattering of many other elements. If we look around the rest of the solar system, we find these elements mostly in the form of very simple molecules, such as methane (CH4), water (H2O), carbon dioxide (CO2), ammonia (NH3), cyanide (CN), and hydrogen sulfide (H2S). But the molecules of living organisms are often very large, consisting of hundreds or thousands of atoms. How did those complex molecules come about?

The answer lies in how these large molecules are structured – they are generally composed of smaller, simpler repeating subunits. For example, DNA is made up of subunits called nucleotides, while proteins have subunits called amino acids (Box 1A). A number of experiments have shown that simple chemical reactions that might have gone on early in Earth’s history are all that are necessary for the formation of those subunits (Box 1C).

While these experiments have shown a number of plausible sources for the molecules of life, there are still many disagreements among scientists. First, we still do not know that much about the actual conditions of the early Earth, such as the exact chemical composition, the average temperatures, the amount of solar irradiance, how much volcanism was occurring, and so forth. All of the experiments may be giving us some hints about the sources of life’s building blocks, but they still do not tell us which were the main sources.

Even more contentious is the concentration problem. How could enough of the building blocks get together in one place long enough for them to be assembled into larger molecules, and what kept those molecules from simply breaking down again?
Recall that one of the hallmarks of life is that it consists of pockets of order – cells – that use energy to keep themselves in that condition.

Various solutions to the concentration problem have been proposed. One solution is various sorts of “sticky” surfaces that act to accumulate organic molecules, such as mineral particles or metal ions that have an electric charge (and so act like a magnet). Another possibility is that certain types of organic molecules formed natural spheres or bubbles, with the other molecules concentrated on the inside of the bubble. Yet another solution is that the molecules formed in pockets in rocks, so that the first wrapper was not organic but just part of the environment. Each of these solutions suggest different possible environmental conditions – high temperatures or low temperatures – or different locations – deep sea vents or along the shores of the oceans – where life might have formed. We will need to know more about the environmental conditions of the early Earth before scientists can say with any greater certainty which of those possibilities was more likely correct.
Box 1C  
Critical Experiment: The Precursors of Life

The science of biology deals with a variety of types of questions. Some of these, such as the origin of life, deal with events that happened in the deep past. Unfortunately for human curiosity, these events were often unique, never to be repeated. How can the scientific process, which is built on predictive understanding and repeatable experiments (see Chapter 2), answer those sorts of questions? Much like a forensics team investigating the scene of a crime, scientists use multiple lines of evidence, including experiments designed to test various explanations and separate the plausible from the implausible. An important tool in this arsenal is the use of experiments that attempt to duplicate conditions from the distant past and recreate those past events. Such an experiment was performed by Stanley Miller when he was a graduate student at the University of Chicago.

Figure 1C. Diagram of the Miller-Urey experiment.  
(Created by Yassine Mrabet,  
Source: Wikipedia)

As a young graduate student in 1950, Stanley Miller heard a lecture by the Nobel-prize winning chemist Harold Urey on the likely chemical composition of the early Earth as it might relate to the origin of life. Miller decided that he wanted to tackle the problem of how the basic building blocks of life (Box 1A) might have come about and approached Urey the following year about doing a series of experiments that would form the basis for his graduate dissertation.

The experiments were very simple and although they differed in detail, they all used a similar approach. They started with the assumption that the atmosphere of the early Earth was primarily made up of water (H2O), methane (CH4), hydrogen (H2), and ammonia (NH3). The other important ingredient was energy. He chose to focus on lightning as a likely source. Thus, the experiment consisted of a gaseous mixture of the four molecules, plus an electrical spark. After letting this mixture “cook” for two days, he looked at the chemical composition of the liquid that had condensed out of the atmosphere. What he found were several kinds of amino acids, the building blocks of proteins, along with other small organic molecules.

From these results, Miller speculated that over the course of millions of years, as these molecules continued to rain down from the sky, the Earth’s oceans would eventually become a rich prebiotic “soup” of all of the building blocks of life. Eventually, the mixture would become concentrated enough for life to form. The validity of these speculations were enhanced when, a few years later, other scientists found conditions suitable for the formation of other building blocks, such as sugars and amino acids. These led, in the 1960s, to the proposition of the RNA World hypothesis (see main text).
Thus, this experiment addressed one of the key issues about the origin of life, where the precursors of life might have come from. For the first time, a scientist was able to demonstrate that it was possible for simple chemical reactions to produce the building blocks of life. However, we now know that the conclusions of this study were wrong. Later calculations suggested that the early atmosphere had much less hydrogen and much more carbon dioxide ($CO_2$).

Such an atmosphere was much less likely to result in complex molecules. In addition, doubts were cast as to whether the rate of formation of the molecules would be great enough to overcome the rate at which they would spontaneously break down. However, the study was critical because it demonstrated for the first time that it was possible to use a modern laboratory set-up to mimic events in the distant past. Newer experiments have superseded its results and addressed its limitations, but it established the principle that the distant past was open to modern experimental tests.

Beyond the effects of the results themselves, this particular study provides an interesting view into the inner workings of the scientific community. Once the experiment was complete and the results ready for publication, there came the question of whose name the results would be published under - after all, getting credit for a scientific discovery is one of the incentives that motivate scientists (see Chapter 2). Although the experiments described here were done in the laboratory of Harold Urey and were inspired by his lecture, the idea and motivation for the experiment came from Stanley Miller. Urey insisted that the paper reporting the results have only the name of Miller on it; Urey realized that as a graduate student, Miller would not receive proper credit for his ideas if the name of a Nobel laureate was also on the paper. Urey also insisted that the paper be published in one of the leading journals, Science, and worked very hard to make sure that was where it would appear. In the scientific community, the student-mentor relationship is often extremely important, as proven by the efforts of Urey on Miller’s behalf. The paper was published in Science on July 10, 1953, just a few weeks after the paper on the structure of DNA was published by Watson and Crick in the other leading journal Nature (see Box 4A). That other paper was a further critical piece in the puzzle of the origin of life.
Replication first

A major breakthrough in explaining life’s origin was the discovery that a single type of molecule, ribonucleic acid (RNA), could act both as an information storage system and as an information usage system. RNA is very similar to DNA, but its subtle differences are important. Unlike DNA, certain RNA molecules, when put together with individual nucleotides, can cause those individual building blocks to assemble into larger molecules, similar to the way that enzymes function (see Chapter 5), thereby duplicating the structure of the original RNA molecule. RNA molecules are capable of self-replication thereby avoiding the problem of how there can be replication of a system that requires many different types of molecules at once. We look at the details of the replication process of DNA and RNA in Chapter 3.

Besides directing their own reproduction, RNA molecules are also capable of linking amino acids together to form proteins. How RNA became associated with amino acids is another part of the puzzle. It turns out that sometimes the ability of RNA molecules to replicate and perform other functions can be enhanced when the RNA molecule has amino acids associated with it via hydrogen bonds. In such instances, forms of RNA that were better at acquiring those amino acids would be favored. A further enhancement would be the development of RNA molecules that were capable of synthesizing amino acids. Plausible scenarios for how that might have come about have been proposed. One piece of evidence for this history is that today, RNA is still involved in all aspects of the basic machinery of cells (see Chapters 3 and 5).

There is still the problem of how the RNA World, as this model is known, got started. One discovery that has provided a possible first step is that single RNA nucleotides, when attracted to the surface of certain clay particles, will spontaneously join together to form larger molecules. While the process would initially be entirely random, eventually combinations would form that would enhance that linkage process. Once that happened, the process would naturally accelerate, since those molecules would continue to add other nucleotides more quickly than other single nucleotides, eventually far outstripping those that had to rely purely on random chance. Eventually, combinations would arise that could continue the building processes independently of the clay particles. Even in this world, the existence of variation and the emergence of new properties from the building of more complex structures were key factors in the origins of life.

All of these chemical reactions to build molecules require energy. In this nonliving world, the energy would have to come either directly from some environmental source, or from other molecules that were formed from those sources. The most likely sources are some combination of very high temperatures from molten rock within the Earth and chemical energy (it is also quite plausible that the origin of life was fueled by a combination of the two). It is the latter source, however, that is hypothesized to have been the primary source for building RNA molecules; specifically, other molecules containing phosphorus. This idea is supported by the fact that today, the chemical binding of phosphorus with nucleic acids is a central form of energy storage and transfer in living systems.

A big stumbling block to the Replication First theory is explaining where the nucleic acid building blocks came from. As yet there are no convincing ways to build those molecules from the simpler forms that could clearly be synthesized from the chemical mix of the early Earth (Box 1C). One suggestion is that a different type of nucleic acid, not RNA, was the actual first replicator, possibly one that combined a nucleic acid with an amino acid. Another major question is how those molecules could last long enough to get to high enough concentrations for the rate of replication to be faster than the rate that the molecules spontaneously broke back down. One solution to these problems is to consider a different start to the entire system, metabolism.
**Metabolism first**

An alternative to the Replication First theory, the Metabolism First theory was put forward in 1988 by a German biochemist and patent attorney, Günther Wächtershäuser. He proposed what is now known as the Iron-Sulfur World hypothesis by noting that iron disulfide, also called pyrite or fool’s gold, can speed up several types of chemical reactions, most notably chemical reactions that are today at the heart of the metabolism of all living organisms (see Chapters 5 and 6). Through this series of chemical reactions, simple organic molecules can use simple, common molecules such as CO2, H2O, CN, and so forth, to make new copies of themselves. If such reactions were occurring, they likely happened at deep-sea vents where very hot water was coming up out of the Earth’s crust or even inside the hot crustal rocks themselves, with geothermal heat providing the needed energy.

After an initial core metabolic cycle was established, more and more additional cycles could then be built and added to the original. The molecules in the central cycle (see Figure 5.CC) are the precursors for many other building blocks, including amino acids and nucleic acids. Recent experiments have shown that even single amino acids are capable of enhancing chemical reactions, so that the multiple, interlocking cycles could have built themselves up.

As with the Replication First theory, the Metabolism First theory also has its limitations. Again, there is the concentration problem. If the metabolic cycles depend on several different types of enzymes all working together, how do they stay associated with each other at high enough concentrations? Similar solutions have been proposed: attachment to mineral particles such as pyrite, concentration within organic bubbles, or concentration within rock pores. The other problem is the lack of an explanation for how these metabolic cycles became associated with replicator molecules such as RNA. The environmental conditions that favor the metabolic cycles – high temperatures and high pressures deep in the ocean – are just the opposite of those favoring the sustainability of RNA.
Science is an international activity, as can be seen by the life and career of Antonio Lazcano, one of the world’s authorities on life’s origin. His life also exemplifies the ways in which one’s culture combined with chance events can shape a person’s outlook and career choices. Lazcano’s international experiences began when he was just a boy. Although born in Mexico, he spent the better part of his childhood in the United States, while his father worked in San Francisco, Santa Fe, and El Paso. However, his family always had strong nationalistic roots, with Spanish being spoken at home and Lazcano spending as much time as possible in Mexico City with his grandmother. Although she helped emphasize his Mexican roots, she had a very cosmopolitan attitude, and like many of her generation was a pronounced Francophile.

Lazcano attended elementary school in the United States during the time of Sputnik and the American-Soviet Space Race, which lead to a huge push in science education and a much more enthusiastic social attitude toward science. This new movement even had an effect on Mexican education, via textbooks imported into Latin American countries. This environment, combined with his natural curiosity, lead him into the sciences; he discovered not only his own love of science, but that it was possible to pursue science as a career. This desire was bolstered by his family, and even his culture - it is a deeply ingrained attitude in Mexican families that education is the best gift one can give to a child, and Lazcano’s family nurtured his budding curiosity.

Two chance events of his childhood helped to shape his life-long interests. The first occurred when Lazcano was around the age of ten; he received Paul de Kruif’s 1926 book Microbe Hunters, a collection of biographies of physicians from the “Heroic Age” of medicine. He became deeply fascinated by the biography of Lazzaro Spallanzani, an Italian biologist and physiologist whose work on biogenesis laid the foundation for Louis Pasteur’s refutation of spontaneous generation (Box 1B). The second event occurred when his grandmother took him to the opera in Mexico City to see The Tales of Hoffmann by Jacques Offenbach, where he was entranced by the character of Spalanzani, an inventor and magician who had been inspired by the actual Lazzaro Spallanzani. Other relatives gave him works by Darwin (see Box 4A) and other scientists; when he moved back to Mexico after his parents’ divorce, his mother and grandmother continued to supply him books as an indispensable part of life – a tradition he now carries on with his own family. His family also gave him scientific equipment; Christmas meant the arrival of items such as microscopes, telescopes, science encyclopedias, and chemistry sets.

Lazcano’s own family has old and illustrious roots in science; his grandaunt was a chemist whose teacher was a famous Mexican chemistry professor who had been taught by Louis Pasteur (Box 1B). This aunt supplied young Lazcano with books on chemistry,
some quite advanced for his age. Her influence meant that he seriously considered becoming a chemist himself. However, upon finishing secondary school and visiting the School of Chemistry at the Universidad Nacional Autónoma de México (UNAM), Mexico’s leading university, he found the chemistry department too politically, socially, and intellectually conservative, and so went into biology instead.

Both Lazcano’s undergraduate and graduate work were done at UNAM, where he later joined the faculty rising to the rank of Professor. Mexican universities are strong promoters of new ideas in science, art, humanities, and politics, and are deeply involved in public affairs, to a greater degree than even many American universities. This involvement has its roots in the idea of the public figure as an intellectual, a major social and cultural point of reference in Mexican and Latin American society. This atmosphere introduced Lazcano to a wide-ranging and very intense cultural atmosphere.

An important influence on Lazcano’s choice of research was the opportunities that being at a leading university gave him to interact with an array of leading scientists. While a student, he met and befriended other geneticists and evolutionary biologists from across the globe, such as the Russian Aleksandr Oparin who wrote one of the first major treatises on the origin of life, the American Stanley Miller a chemist and biologist who did some of the first experiments on the origin of life (Box 1C), and the Spaniard Joan Oró, a biochemist who was working at the Astronomy Institute of UNAM on extraterrestrial organic compounds and later worked with NASA on various missions to Mars. These friendships played a major role in shaping Lazcano’s understanding of the field of the origin and early evolution of life. Although Mexican society is much more accepting of evolutionary theory than the United States, science itself plays a much smaller role. Mexico has a much smaller percentage of scientists than in countries such as the United State, the United Kingdom, France, or Germany, and the resources devoted to science’s promotion and development fall far short of the standards recommended by UNESCO and other international organizations. However, Lazcano’s own research efforts have always been well-received and supported by the government. In addition, his books written for non-scientists have been well-received. His most popular book The Origin of Life, has sold over 600,000 copies since its publication in 1984. He continuously receives an overwhelming number of invitations to speak to both scientific and general audiences, and has written numerous articles for newspapers and often spoken on radio programs; he is currently involved with a radio program run by one of his friends.

Lazcano believes strongly in promoting the study of Darwinian evolution and opening new frontiers to students as part of his duties as a mentor – although some of his motivation comes from the challenge of sparking new debates. Surprisingly, at least to the average American, the strongest opposition Lazcano has faced in his scientific career comes from American Creationists visiting Mexico – “creationism” as a concept (see Chapter 4) does not exist in Mexican society. Although the Catholic Church plays an important social and cultural role, as an institution it has no problem with evolution or studies on the origin of life. Often, the invitations he receives to lecture on evolution or life’s origin are from private schools run by nuns or priests, a fact that might seem very strange to an American student. Part of the reason for this openness comes from the strength of the Catholic church in Mexican society, which does not preach Biblical literalism. This stands in contrast to the strongly Protestant strain of Christianity more common in America.

Even the Mexican government has served to spread evolutionary theory, via textbooks freely available for schoolchildren published by the Mexican secretary of Public Education. In the 1980s, Lazcano proposed to teach a course on the origins of life at UNAM; because of the groundwork laid by Alfonso L. Herrera, a Mexican naturalist in the
early 20th century, who was one of the first promoters of evolutionary ideas in Mexico, the idea was well-received.

Attempting to explain the origin of life, an event not directly observable, has been challenging Lazcano for most of his career. As with any other relatively young field of study, Lazcano has seen the attitude of the scientific community towards this question change drastically over the past several decades. Although it was once seen as little more than useless speculation, the spreading awareness of genes and proteins as rich repositories of evolutionary information has radically changed the acceptance and gravitas of the discipline. The work of Carl Woese (see Box 5A) in subunit ribosomal DNA did much to lend support to the theories, and the discovery that RNA could be both an information carrier and a catalyst (see main text) lead to new perspectives in the study of the origin and evolution of life. NASA has even lent a great deal of support to it, which has greatly contributed to the development of astrobiology, and multidisciplinary studies of the origin and existence of life in the wider Universe.

Lazcano currently studies the mechanisms that might have led to the synthesis and accumulation of organic compounds in the primitive earth. By comparing the genes of Bacteria, Archea, and Eukaryotes, he hopes to gain insights into very early biological evolution, the stepwise evolution of replication and translational apparatus, and the development of the genetic code (see Chapter 3). Although it is extremely unlikely that we will ever observe direct evidence of prebiotic or early-life stages, the types of inferential work exemplified by Lazcano is likely to provide a convincing explanation of life’s origin. Throughout these endeavors, Lazcano’s international connections have served to strengthen his research. He has been professor-in-residence or visiting scientist in France, Spain, Cuba, Switzerland, Russia, and the United States. He has served on many international advisory and review boards, including one run by NASA on astrobiology. His greatest honor was his election (and re-election) as president of the International Society for the Study of the Origin of Life, the first Latin American scientist to occupy that position. He continues his research and teaching at UNAM to this day.
Reconciliation

One answer to the limitations of the Replication First and Metabolism First theories is to try to wed them. There is no reason to assume that the origin of life had to occur from a purely replication or metabolic starting point. After all, the two theories are not mutually exclusive, and each could have been happening simultaneously. This scenario provides a possible solution to the problem of how the replication and metabolic systems came to be associated with one another, despite the disparate conditions that would favor each type of molecular reaction. If both types got started simultaneously but under very different conditions, it is possible that they each could have begun to transition to more and more similar conditions through the common use of organic bubbles as concentration systems. Such bubbles, chemically similar to but much simpler than the molecules that make up cell membranes, will spontaneously merge when put together. Given enough time – hundreds of millions of years – through random chance many different combinations of “replicator” bubbles and “metabolism” bubbles would be formed. Eventually, some combinations would form in which the replicators would enhance the enzymes in the metabolic cycles and the metabolic cycles would produce the nucleotides needed for the replicators. Those combinations would be favored because their chemical reactions would occur faster than others through the feedbacks between the different types of systems. At this point, the system would have become complex enough to foster the emergent property that we call “life.” The transition from nonliving to living would have occurred as information and metabolism merged together into these discrete packages, the first cells.

The first cells would still have looked very different from even the simplest single-celled organism currently alive. A modern microbiologist investigating their structures and lifecycles would be struck by how much simpler they would appear, with few metabolic cycles and comparatively small RNA molecules. From these first tiny sparks of life, there would still be many steps needed before RNA replication began to depend completely on enzymes and the building of the enzymes began to depend completely on RNA. This switch likely occurred because proteins, relative to RNA, have the potential for greater complexity in their three-dimensional structure, allowing for greater efficiency and more types of enzymatic functions (see Chapter 5). There is also the question of how the information “language” came into being. Specific RNA “words” are translated into specific amino acid building blocks (see Chapter 3), but how that code came to be is not known. We also don’t yet know why the switch in information storage from RNA to DNA occurred. The most likely answer is that because DNA is more a stable molecule than RNA, meaning that it acts as a more effective storage system. These transitions have left a very visible legacy in today’s living systems; although DNA is the primary information storage system, RNA plays a central role in the cell’s machinery (see Chapter 5). Those transitions have not yet been explored. So, there is still much research ahead before we fully understand how life originated.

Although the origin of life is usually talked about in terms of a single, explosive event, it is equally likely that life actually emerged multiple times. Although all of the available evidence points to a single origin, that may simply mean that one form of life proved to be more efficient than the others, so that today’s living systems are all descended from that winner. It is also possible that life originated multiple times, each time quickly going extinct, before it finally took hold. Because these multiple origins would have left no traces, any evidence is lost in the depths of time. As frustrating as it is, sometimes science raises questions that are impossible to answer.
Is it alive?

Perhaps the most basic question of biology is whether or not something is alive, and thus falls within the purview of the biological sciences. It might seem self-evident, but sometimes there are things that seem to fall into a grey area; this is when we must take a step back and take a systematic look at the object in question, using the list of the characteristics of life (Table 1.2) to really determine if various types of objects are alive or not. No single property distinguishes living systems from non-living systems; rather, it is all of those properties together that allow us to identify life. Consider a crystal, such as a diamond or a lump of table salt. It is a highly ordered system that came about through the use of energy, and given the proper chemical conditions, you can grow your own salt crystals. But we would not say that those crystals represented life because they are unable to maintain their own structure. In addition, although they are ordered, they are not complex. Rather than consisting of many different parts, they are uniform in structure.

What about fire? A fire actively maintains itself by consuming fuel, much the same way that you “burn” calories. But a fire works in only one direction, moving from order to disorder. It does not have the complexity to create order.

What about computers and robots? Are they alive? Certainly, as currently constituted, none of them are alive, although the various aspects of living systems can be found scattered among them. There are computer programs that mimic the evolutionary process (see Chapter 4); the software is designed to generate random changes that are then selected in ways that permanently change the software code. Computers certainly include systems for the storage, use and transmission of information, and both computers and robots are complex systems that consist of interacting parts with emergent properties. What they are unable to do, though, is actively maintain themselves. While software code is capable of change and evolution, that is not true of the computers or robots running that code. Computers and robots also do not directly reproduce themselves without direction from and materials provided by humans. Currently, all software, computers and robots require outside intervention, us, to create new versions. It is certainly possible, though, to envisage a time when there is a self-guiding robot that is capable of gathering materials to repair itself and to create new, similar robots. At that point we would have to say that the robot was alive. Even though we would consider it to be a form of life, it would still be fundamentally different from any life from currently known today - one of the key differences of such robotic life from ours is that reproduction by the robot would be occurring externally. A robot would reproduce by building another robot, as opposed to having the new robot grow directly from the body of the first robot. For life on Earth, all reproduction occurs with at least some aspect occurring as a process internal to a currently living progenitor (see Chapters 5 and 6), but such internal reproduction is not a necessary condition for life itself.

Which brings us to an always contentious question, are viruses alive? Viruses consist of either DNA or RNA (i.e., an information storage system), that is usually, but not always, contained in some sort of protein wrapping. Viruses store and use information, they vary in their composition, they consist of complex parts with emergent properties, and they are capable of change. What they fail to do, however, is to actively maintain themselves. We cannot even categorize viruses as cellular – viruses do not have their own metabolic system. They are not self-sustaining in the sense of taking in matter and energy that is then used for maintenance, growth, and reproduction. Instead, they insert themselves into a cell and redirect the cell’s machinery for their own purposes. Like today’s robots and computers, they require an external agent for their reproduction.
That dependence on an external agent is different from our requirement that we need to eat other living organisms to survive. In our case, we take in that matter, but break it down into small pieces. Rather than using the machinery itself, we break down the total organism and use the constituent parts; viruses, on the other hand, are directly dependent on the machinery of the cell. All of this means that viruses are non-living parasitic molecules.

However, this does not mean that viruses have no connection to life at all. Life was originally cellular in nature, and viruses may have evolved later as “rogue” bits of DNA or RNA that became independent of the cells within which they originated. It is unlikely that viruses evolved independently of cellular life; in the RNA World, the chains of nucleic acids that formed the basis of life were capable of self-replication; in contrast, all known viruses require cellular proteins to replicate. It is likely, in fact, that new viruses have arisen throughout Earth’s history.

**Nonterrestrial life**

If you were traveling the universe, how would you recognize life if you saw it? If it comes up and says “Hello” or tries to eat you, it is pretty obvious. As amusing as it may be – and as good a basis for a science fiction story – that scenario, is not very likely. The non-Earthbound life that we are most likely to encounter is on Mars, and certainly not complex enough to interact with humanity. We currently have robots on Mars looking for life, or at least signs that it was once around. What should we look for?

The answer is quite simple: life must have the characteristics listed in Table 1.2. It doesn’t need to be based on the same molecules that we find here on Earth (Box 1A), nor does the information storage system need to be based on DNA or other nucleic acids. It doesn’t even need to be carbon based. If, for example, we succeed in creating robots that meet the criteria for being alive, they will be based on iron and silicon, or carbon-based plastics.

The likelihood that life elsewhere in the universe will resemble life on Earth very much depends on the conditions that led to the origin of life here. Part of debate over origin of life is the question of how many of the characteristics of Earthbound life are inevitable or necessarily universal because the origin of life on Earth was the most likely or only route to life, and how much is a “frozen accident” or contingent. The only way to answer this question is by finding life on other planets and comparing it to life on Earth (Box 1E). The Metabolism First theory predicts that the core metabolism of life elsewhere should be very similar or identical to that of life on Earth. The Replication First hypothesis predicts that life elsewhere should also use DNA and RNA as its core information storage and translational molecules. Some versions of these theories even predict that the words of the DNA language will be the same. This is because these theories claim that the information and metabolic systems here on Earth are the way they are because they are the most chemically stable and therefore the mostly likely to occur. If they are correct, then life elsewhere might consist of very similar organic molecules to those on Earth.

All of those predictions depend on conditions elsewhere in the universe being similar to those here on Earth. One reason that the search for life on Mars is of such interest is that early in its history conditions were likely similar to those on the early Earth. There has also been speculation about other possible places in our solar system where life might be found, with the leading candidate being Titan, a moon of Saturn. While Titan has a chemical composition similar to that postulated for the early Earth, a key difference is that Titan is much further from the Sun and so is much, much colder. Life there would likely depend on thermal energy from within Titan itself.
What about outside our solar system? Now we are free to speculate all we want because there are few data to constrain us. Every year astronomers are discovering more and more extrasolar planets, although none so far look like the Earth. That is not because Earth-like planets do not exist, just that it is very hard to see something that small at those distances. That lack is likely to be corrected in coming decades as astronomical observing capabilities are improved. Our first hint of life elsewhere in the universe is likely to come from the detection of substantial amounts of free oxygen (O2) in the atmosphere of other planets. This is because oxygen is highly reactive and tends to combine with other molecules (e.g., CO2 and H2O); thus, free oxygen is almost certainly a sign that a planet has a living system.

The various theories of the origin of life suggest that life is likely to be common elsewhere in the universe, because Earth-like planets are likely common and life is almost inevitable given those conditions. This still does not mean that a human visitor would be able to talk to these lifeforms; such life would not necessarily be sentient, or even multicellular. On the Earth, life consisted of single-celled organisms for about 3 billion years. Multicellular forms arose only about 1 billion years ago, and complex organisms only about 600 million years ago. The odds of finding another sentient race are slim, especially given the vast size of the universe. Mathematically, the odds are based on the Drake equation, which calculates the probable number of civilizations in the universe with which we might be able to communicate by taking into account factors of how many life-bearing planets exist, how many of those planets actually support life, and how many of those life forms create civilizations capable of interstellar signals. Any solid proof of extra-terrestrial life would represent an incredible discovery, however, and one that would mark a watershed moment for both the biological sciences and humanity as a whole.
Are we alone in the universe, the sole intelligence among billions and billions of stars? While we still don’t know the answer to that question, the search for that answer found an early and vigorous advocate in Carl Sagan. An astronomer, astrochemist, and prolific author, he popularized astrophysics and other natural sciences, and was especially well-known for the PBS series “Cosmos: A Personal Voyage,” which he narrated and co-wrote. The series, which as been seen by over 600 million people in 60 countries, is the most-watched PBS series of all time. He also published more than 600 scientific papers and popular articles and over 20 books, in which he advocated skeptical inquiry, secular humanism, and the scientific method.

Born in Brooklyn, New York on November 9, 1934 into a Russian Jewish family, Sagan’s father was an immigrant garment worker. He attended high school in Rahway, New Jersey, and graduated in 1951, afterwards attending the University of Chicago for both his undergraduate and graduate degrees in physics, astronomy, and astrophysics. In 1960 he accepted the Miller Fellowship at the University of California-Berkeley, and two years later began working at the Smithsonian Astrophysical Observatory in Cambridge, Massachusetts. While there, he lectured at Harvard until 1968, when he moved to Cornell, where he attained full professorship in 1971 and directed the Laboratory for Planetary Studies. He became ever more deeply involved in astronomical research, becoming the Associate Director of the Center For Radio Physics and Space Research from 1972 until 1981, and was a foremost figure in the space program.

Figure 1E.A
Plaques carried by the Pioneer 10 and 11 spacecraft. The plaques show the nude figures of a human male and female along with several symbols that are designed to provide information about the origin of the spacecraft.
(Designed by Carl Sagan & Frank Drake; artwork by Linda Salzman Sagan; Source: Wikipedia)
From the 1950s on he was an adviser to NASA, and would brief Apollo astronauts before their flights to the Moon. He contributed heavily to most robotic spacecraft missions, arranging experiments on many of the expeditions around the solar system. It was Sagan who first conceived the now-iconic idea of including an unalterable and universal message, understandable to extraterrestrial life, on spacecraft destined to leave the solar system. He assembled the first one, a gold-anodized plaque on the Pioneer 10 probe launched in 1972, and another on Pioneer 11, launched a year later. He continued to refine the design throughout his life; the most elaborate message was the Voyager Golden Record, sent on the Voyager probes in 1977.

Some of Sagan’s major contributions to planetary science came in his predictions of Venusian surface conditions. While its inhospitable nature is well-known today, at the time most imagined a balmy paradise underneath Venus’ thick cloud layer. Sagan analyzed radio emissions from the planet, and concluded that its surface temperature would be around 500°C, and he depicted his conclusions in a report that would later be the basis for the Time-Life book Planets. His predictions were proven true in 1962 when the Mariner 2 probe landed on Venus and provided data on its inhospitable surface conditions. Sagan himself worked on the Mariner project while he was a visiting scientist at NASA’s Jet Propulsion Lab. He also established Venus’ atmosphere as extremely hot and dense with steadily rising pressure down to the surface. These observations led him to the possibility of throwing Earth into an artificial version of the process that made Venus a barren, life-hostile planet through a vastly accelerated greenhouse effect, a prospect that we are now facing.

Sagan was among the first to hypothesize that Saturn’s moon Titan might possess oceans of liquid compounds and that Jupiter’s moon Europa might have subsurface oceans of water, which would provide the potential for life; his predictions about Europa were indirectly confirmed by the spacecraft Galileo. He helped solve the mystery of the red haze that spears around Titan, which has proven to be composed of complex organic molecules constantly raining down to the surface. He proved that color changes on Mars, which has been previously attributed to seasonal or vegetative shifts, were actually the product of massive shifts in the surface dust due to windstorms sweeping across the planet’s surface.

Sagan is best known, however, for his contributions to the search for extraterrestrial life. Firmly believing that other worlds supported intelligent life, he urged the scientific community to listen with radio telescopes for signals that would indicate the existence of extraterrestrial civilizations. By 1982 he had succeeded in getting a petition signed by 70 scientists (including seven Nobel Prize winners) advocating a formal Search for Extra-Terrestrial Intelligence (SETI), which was published in the journal Science. It marked a huge turnaround in the respectability of the field, which had always been considered more science fiction than real science. Not content to simply search for evidence of other life in the universe, Sagan helped Dr. Frank Drake write the Arecibo Message, a radio message broadcast into space from the Arecibo radio telescope on November 16, 1974, aimed at informing extraterrestrial intelligences about Earth and the existence of human life. He became a member of the SETI Institute’s Board of Trustees, and co-founded the Planetary Society, the largest space-interest group in the world, with over 1 million members in more than 149 countries. His involvement in the astronomical community was extensive; he became the Chairman of the Division for Planetary Science of the American Astronomical Society, the President of the Planetary Section of the American Geophysical Union, and the Chairman of the Astronomy Section of the American Association for the Advancement of Science.
In addition to being a widely-recognized figure in his scientific fields, Sagan also dedicated his life to bringing science into the realm of public knowledge. His most recognized public offering was the 13-part PBS series *Cosmos*, which he co-wrote with author Ann Druyan (who was also his third wife). In it, Sagan covered a diversity of topics, from the origin of life to a perspective on our place in the universe. When *Cosmos* was first broadcast in 1980, it won an Emmy and a Peabody Award. Sagan also wrote a book by the same name, a reflection on and expansion of the series, which became the best-selling science book in the English language. In all his writings, Sagan elaborated a skeptical, naturalistic view of the world writing about science and its role in human society, the difference between pseudoscience, modern superstition, and viable scientific inquiry, and examined the principles of religion and the relationship religion has to science, expressing skepticism about conceptualizations of God. He promoted critical thinking and the scientific method as tools to examine the world, hoping to equip humanity with the ability to better advance society. In addition to broadcast awards for *Cosmos*, Sagan also won a Pulitzer for *The Dragons of Eden: Speculations on the Evolution of Human Intelligence*. Straying a bit from nonfiction, he also wrote *Contact*, a best-selling science fiction novel, but with a central message of the story is one of rational skepticism and scientific inquiry.

At the height of the Cold War, when the threat of a nuclear war seemed to be just over the horizon, Sagan became involved in public awareness efforts for the environmental effects of nuclear war. He was alarmed after a mathematical climate model suggested that a worldwide conflict involving the launch of multiple nuclear warheads could disrupt the balance of the Earth’s climate. He co-authored a paper hypothesizing global nuclear winter, a condition in which so much dust gets thrown into the upper atmosphere by nuclear explosions that the Earth’s climate would experience a devastating drop in temperature, essentially creating winterlike conditions all year. Later, he spoke against President Reagan’s Strategic Defense Initiative. He believed the level of precision necessary to deflect missiles in that way would be unattainable, and the defensive missiles themselves easy to deflect with decoys; the very existence of such a program, he believed, would destabilize the nuclear balance between the United States and Russia.

All these concerns might have come from the Fermi Paradox: While logic suggests that a high number of extraterrestrial civilizations would form throughout the universe, we have not seen any evidence for their existence. Sagan suggested that their absence was because such civilizations tend to destroy themselves quickly. He hoped that by identifying and publicizing these huge risks to humanity, we might avoid that same fate. While Sagan had some interest in the phenomenon of UFOs, he was quite skeptical of any extraordinary answer to the mystery of strange lights and shapes in the night sky. Because of the public interest in the matter, he thought it merited at least some study, but, as with the rest of his work, he wrote frequently on the local and empirical fallacies regarding UFOs and stories of alien abductions. When the Air Force commissioned a study on UFOs, named Project Blue Book, Sagan was on the committee that reviewed it in 1966. They found that Project Blue Book was lacking as science, and recommended a university-based project. This lead to the formation of the Condon Committee, which ran from 1966 to 1968, led by physicist Edward Condon. Unfortunately for the fervor of UFO enthusiasts everywhere, the Committee found no evidence of extraterrestrial visits to Earth, concluding only that the phenomenon was not a threat to national security. Sagan himself examined UFOs in books and one episode of *Cosmos*, but did not spend much time researching them due to the lack of empirical evidence, and said that attributions of UFOs to extraterrestrial activity were emotional, not rational in nature, arguing that the chances of an extraterrestrial visit to Earth were extremely minute. With regard to
such claims and the lack of evidence, he famously stated that “Extraordinary claims require extraordinary evidence.”

Despite Sagan’s numerous scientific achievements, he was never made a member of U.S. science’s most prestigious body, the National Academy of Sciences. His colleagues felt that his work presenting science to the general public was unbecoming for a serious scientist; ironically, his very popularity with the general public worked against his popularity with his scientific peers. It is also possible that jealousy of that popularity played a role. Today attitudes about the value of explaining science to the general public are changing, especially as issues like global warming require an informed electorate. Carl Sagan died on December 20, 1996 after a long fight with a rare type of bone cancer. Six months later, the unmanned Mars Pathfinder landing site was renamed the Carl Sagan Memorial Station.
Ways of Knowing

One of the most important aspects of the human mind is its ability to organize and process information about the world around us. It is this ability that has allowed humanity to grow into what we are today, and it is what will determine what becomes of us in the future. Science is a way of organizing and processing information – understanding the world – that systematically combines logical thinking with observations of the world. It provides explanations for how the universe works and where it came from. The methodologies of science – the procedures scientists use for coming up with those explanations – have developed over thousands of years, with a particularly swift advance in the past 400. Science as we know it today is a child of the Enlightenment, the movement in western Europe that claimed that rational thought could provide guidance in humanity’s endeavors to understand and manipulate the world around us. In this chapter we will explore what science is and how it operates, both as a rational enterprise and as a social enterprise; that is, how it helps individuals understand the information they acquire, and how scientists themselves interact with each other (Table 2.1).

At its most basic level, science is built on a tripod of pattern, process, and theory. Patterns consist of the relationships between the phenomena or entities of the natural world, processes are the causes of those patterns, and theories are the explanations of those causes. Theories – what ultimately allows us to understand the world – are built with logic and tested against those patterns and processes. The work of a scientist is to document patterns, investigate and understand processes, and ultimately to put together theories that explain what they have found out. Although a single scientific study may be focused on only one of those three aspects, science seeks to understand all the parts of a system and how that system functions as a whole. This is particularly true of the science of biology; life is amazingly complex with properties that are more than the sum of the constituent parts, so biologists are always faced with this duality of taking things apart and then putting them back together.

Alternatives to science

The easiest way to understand the nature of science is to contrast it with other ways of knowing and understanding the world. Mathematics, for example, is a way of knowing separate from science. Science tells us about the empirical world, while mathematics tells us how logical symbols can be manipulated to provide knowledge of patterns of time, space, and numbers. Science can not apply value judgments. Although science can tell us about the physics of light and color and about how our eyes interact with our brain to affect our perception, it cannot create standards to judge whether one painting is better than another – for that we need aesthetics. Aesthetics is a separate way of knowing about the world with a goal of reflecting on art, culture, and nature so as to render judgments on sentiments or taste. Science tells us that we are capable of eating both animals and plants, but cannot tell us whether we should live as strict vegetarians. To answer questions of proper behavior, we turn to ethics; ethics is a separate way of knowing about the world with a goal of helping us decide how we should live our lives. What science can do is provide information about the consequences of those decisions, for example, how to achieve a healthy diet as a vegetarian and the alternative effects on the environment of eating or not eating meat.
Table 2.1. The structure of science

A. Premises
   1. One’s senses provide reliable information about an external reality.
   2. All scientific explanations must rely on phenomena that are the result of natural processes.

B. Methodologies
   1. Science provides explanations through theories and models.
   2. Scientific theories should be consistent with each other.
   3. Theories are built on observations.
   4. Theories are tested by building models that generate hypotheses.
   5. Hypotheses are falsified through experiments.
   6. Experiments can be of various types: manipulative, natural, or observational.
   7. Theories are revised based on experiments and additional observations.

C. Science as a human endeavor
   1. Research is affected by its social environment.
   2. Science is self-correcting.
   3. Peer review provides feedback between a scientist and the rest of the scientific community.
   4. Science proceeds according to ethical rules and norms.

The most contentious border between these different ways of knowing lies between science and religion. To understand how they differ, consider the bases of each: science as a way of knowing rests on two basic premises. The first is a statement about the world, while the second is a statement about the domain of science.

The premise about the world is science’s single, fundamental assumption, that one’s senses provide reliable information about an external reality. The alternative is that one is a disembodied intelligence and that the universe is all an illusion or a dream. It is impossible to prove that the world is not an illusion, since the only evidence one has is the evidence of one’s senses. This assumption is the nearest thing that science has to an article of faith or belief. Although specific theories rely on a host of assumptions, all of those assumptions share a common property: they are potentially refutable by observations. It is only the assumption that those observations are actually real that cannot itself be tested.

The premise about the domain of science is that all scientific explanations must rely on phenomena that are the result of natural processes, the “naturalism” premise. Science does not deny the existence of supernatural phenomena, this premise simply states that any such phenomena lie outside the realm of science. The reason for this premise is that explanations that rely on natural processes can be refuted by observations of the natural world, while supernatural processes are inherently beyond the world and thus cannot be so refuted. Thus, the naturalism premise is not a claim about the world, but a statement about the boundaries of science.

Some assert that science is just another form of religion, but that assertion is based on a misunderstanding of the two basic premises. Religion also has two basic premises that are mirror images of those of science. First, religion makes a basic claim about the world, that supernatural phenomena exist. Second, religion relies on faith,
addition to observation, to understand those supernatural phenomena. Notice that science makes no claim about the existence of supernatural phenomena, just a statement about the admissibility of those phenomena as scientific explanations. In contrast, religion makes no claims about the reliability of one’s senses, instead indicating that knowledge can come to one from something beyond the senses.

Interestingly, science and religion are each based on two premises – claims about the world and what counts as evidence – that are complementary. Science is not a form of religion; rather, both are differing systems for trying to understand the world. The two systems clash only when each tries to make assertions about the premises of the other: when scientists claim that supernatural phenomena do not exist and when proponents of religion claim that data about the world can come from sources other than one’s senses. While either assertion is consistent within the worldview of science or religion, respectively, neither are consistent with the other’s worldview. Science can no more refute religion than religion can make claims about scientific theories. Because of this, sometimes conflicts occur between religion and science, especially when religion tries to make claims that affect the realm of science (see Chapter 4).

**Biology’s special assumptions**

The science of biology has two assumptions that are unique to itself. First, related to the naturalism assumption, biologists reject vitalism, the notion that living systems are imbued with some sort of nonmaterial life force. The discoveries of the chemical bases of life in the past few centuries, along with our understanding of the properties of living systems, has led to the understanding that life is something that can emerge from nonlife through purely material processes without the necessity of invoking any nonmaterial force (see Chapter 1).

Second, life as a whole is not goal directed, although individuals can certainly show goal directed behaviors. When you are hungry, you eat in order to stop that hunger. But wolves did not evolve sharper teeth because they wanted to be better hunters and acquire food more easily. A cell does not produce more of a particular protein because it wants to perform a specific chemical reaction. These are both examples of feedbacks that are simply part of living systems. We would no more say that a boulder wanted to roll down a hillside after the ground underneath it eroded away. Goal directedness is called teleology, and in biology it is very easy to fall into teleological explanations because the many feedback systems appear to be goal directed. Biologists are often guilty of using teleological sounding explanations because it is a convenient shorthand. However, you should always be suspicious of any such explanation and should always remind yourself that there is always a nonteleological, albeit more complicated, explanation.

**How We Know**

**The philosophy of science**

One of the concerns of the discipline of philosophy is trying to understand how it is that we know things. The study of how we know is called **epistemology**. In this book we are concerned with just one branch of epistemology, the philosophy of science. The philosophy of science has a very long and complex history (Box 2A) and, like all of philosophy, contains many different schools of thought, some with very subtle distinctions. For our purposes, we will collapse that complexity into three broad classes: empiricism, social constructivism, and realism. Understanding these three main approaches tells us how science lets us know about the world.

**Empiricism** approaches the world with the goal of producing theories that can make useful predictions through logical deduction from basic assumptions combined with
data gathered from previous experiments. For example, we can use Newton’s laws of motion to tell us how to aim a spaceship so that it will travel from the Earth to the Moon. Fundamental to this approach is the notion that even when a theory makes correct predictions, we have no reason to believe that the theory tells us anything about how the world really works. Although we can use Newton’s laws to guide our spaceship because the theory works under some circumstances, physicists have replaced Newton’s theories with those of Einstein. At some time in the future, Einstein’s theories may, in turn, be discarded.

Like empiricism, social constructivism also posits that we have no reason to believe that theories tell us how the world really works. Theories are seen as consensus agreements by communities of scientists; they are consistent with the data because the community of scientists concur on which data are relevant to those theories. When data disagrees with a theory, the theory is tweaked while retaining its core structure. Only when disagreements between data and theory become overwhelming is a theory replaced, again through a social consensus. Social constructivism differs from empiricism with regard to which aspects of the scientific process are the primary drivers: logic and data (empiricism) or social conventions (constructivism).

In contrast, realism posits that theories do tell us about how the world works. Theories make accurate predictions because they capture some true aspect of the world. Scientists do not merely discard one set of theories for another, but as theories are refined or replaced, we get closer and closer to the underlying truth. Going from a theory that the Sun goes around the Earth in a perfect circle to one in which the Earth goes around the Sun in a perfect circle moves us closer to reality. Later, that theory was refined from movement in a circle to movement in an ellipse and combined with Newton’s laws of motion to make predictions about the movement of all of the planets. Later, it was discovered that those laws could not account for the motion of the planet Mercury around the Sun. That discrepancy was one of the factors that resulted in the replacement of Newton’s laws by Einstein’s theory of relativity which could account for movement close to a large gravitational object. In contrast to both empiricism and social constructivism, realism posits that as scientific theories get refined they get closer to the truth and thus are less likely to be overturned.

The science of biology has had similar changes in its basic theories, although most of them have not been nearly as dramatic as those in the field of physics. The major foundations of biology were built in the 19th century with the establishment of the chemical and cellular bases of life (see Chapter 5), the theory of evolution (see Chapter 4), and the roots of genetics (see Chapter 3). Studies of the functioning of organisms go back farther (see Chapter 6), while the discipline of ecology emerged from natural history near the end of the 19th century (see Chapter 7). The 20th century can be seen as a period of refining those broad theories, including some notable debates and the refuting of some widely held theories. By the 1960s, the broad outlines of the theories presented in the rest of this book were in place (see Tables 1.2, 3.1, 4.1, 5.1, 6.1, 7.1). While we are still refining the details, most biologists from that decade would not quarrel with most of those fundamental concepts.

Since the science of biology has not seen the sorts of major upheavals that have roiled the history of other disciplines, biologists now believe that a realist position is warranted. As we keep refining our theories and making better predictions we are able to produce better crops, cure more diseases, and predict the effects of overfishing on population sizes in the ocean. One reason for this realist confidence is that biology very often deals with things that we can directly see and feel. Biologists, using that assumption of the reliability of our senses, can directly observe the agreement between theory and the
Most philosophers of science can be classified as either empiricists or social constructivists, while the vast majority of scientists are realist, perhaps because scientists are the ones who are constantly confronted with observations. Given this difference in outlook, many scientists question the value of the philosophy of science. However, the philosophy of science is valuable because it provides scientists with tools for gaining a deeper understanding of themselves and their chosen vocation, in particular a basis for the building and testing of scientific theories, which we explore next.
No discussion of the history and foundations of Western thought is complete without Aristotle. One of the founders of Western philosophy itself, he was the first to fully create a comprehensive system of thought systematically encompassing everything from morality, politics, and aesthetics to scientific inquiry, mathematics, and metaphysics. His views on physical science stood as unquestioned fact well into the Renaissance and formed the basis for Medieval Scholasticism. His writings had a profound influence on the philosophical and theological system of all three major monotheistic religions. His system of formal logic, the first of its kind to be created, is still used today. This influence exists despite the fact that most of his writings have been lost; it is thought that only one-third of his works are still extant.

Aristotle was born in Strageira, Chalcidice in 384 BCE, the son of Nicomachus, who was the personal physician to King Amyntas of Macedon, and as a result received the training and education appropriate to the aristocracy. Around the age of 18 he traveled to Athens to attend Plato’s Academy and stayed there about 20 years until Plato’s death in 347 BCE. Afterwards, he traveled with Xenocrates to the court of Hermias of Atarneus in Asia Minor, then with Theophrastus to the Isle of Lesbos, where they engaged in a major botanical and zoological survey of the local plant and wildlife.

He recorded his observations and dissections in three major works, History of Animals, Generation of Animals, and Parts of Animals. He recorded incredibly detailed observations on the marine life around Lesbos, making entries on catfish, electric fish, angler fish, octopus, cuttlefish, and the paper nautilus. He created an early form of systematics for the animals he documented, separating aquatic mammals from fish and naming sharks and rays as a group called Selache, or selachians.

After Hermias’ death, he was invited by Philip of Macedon to be the tutor of his son Alexander. After his tutoring post he returned to Athens where he established his own school, called the Lyceum, in 335 BCE. This was his most prolific era in his writings, producing many dialogues, few of which survive today. Most of his extant writings are in treatise form, and were probably intended as lecture aids for his students. Aristotle’s writings were separated into two categories, esoteric and exoteric; his treatises, which were meant for his students and other philosophers, were the former, while his dialogues, meant for the general public, were the latter. However, what history would come to regard as his most
important works, his writings on Physics, Metaphysics, Nicomachean Ethics, Politics, De Anima (On the Soul), and Poetics, would fall into his category of esoterics.

Aristotle’s style of philosophy differed from his teacher’s in some striking ways. Plato studied the “forms” of things as a universal truth existing in a spaceseparate from the physical world, then moved into the realm of the physical particulars (considering them pale imitations of their pure forms). In contrast, Aristotle’s natural philosophy looked for the higher “essence” in the physical phenomena he studied, aiming to uncover the universal truth via the things themselves. Aristotle referred to the entirety of human inquiry as “science,” dividing disciplines into “practical science,” which would encompass disciplines such as politics and ethics, “poetical science,” encompassing poetry and the fine arts, and “theoretical science,” which would be most of what we call science today, as well as mathematics and metaphysics.

His investigations blazed new ground in nearly every discipline known at the time, although it was qualitative rather than quantitative. He lacked concepts like mass, velocity, force, and temperature, and was severely limited by the simple technological lack of instruments like clocks and thermometers. Perhaps because of this (and a lack of human dissection), his scientific laws are a mix of ideas centuries ahead of his time and complete errors. He claimed, for example, that heavier objects fall faster than lighter ones (which Galileo would famously disprove) but refuted Democritus’ claim that the Milky Way was brighter because the stars there were shielded from the sun’s light by the earth, citing the relative size and distance of the sun to the earth and the rest of the stars to assert that the sun’s light would strike all stars equally.

The Aristotelian universe was geocentric, and based on far too little empirical observation when it came to astronomy and physics. However, he performed an immense amount of firsthand research in the life sciences, from his cataloguing of the life of Lesbos, describing ruminants (e.g., cows) four-chambered forestomachs and making an extensive study of embryonic chick development, breaking open fertilized chicken eggs at intervals during incubation to track the generation of visible organ systems. He classified organisms in a hierarchical Ladder of Life, placing them according to their complexity of structure and function. His systematic classifications would encompass vertebrates (as animals “with blood”) and invertebrates (animals “without blood”), further divided into live-bearing (mammals) and egg-bearing (birds and fish), and insects, crustaceans (with non-shelled and shelled cephalopods) and testacea (mollusks), respectively. With his studies of biology, he introduced the idea that nature is composed of things that change, and by studying the changes, one can discover the underlying constants. Despite his focus on observations of the natural world, the concept of the experiment, the manipulation of nature to reveal causal relationships, never occurred to him. Aristotle’s scientific systems would stand unchallenged for centuries, forming the basis of Western knowledge and thought through the Renaissance, until Sir Francis Bacon (Box 2C) initiated the creation of the modern scientific method.

Unfortunately for Aristotle, after Alexander the Great’s death in 323 BCE, anti-Macedonian sentiment bloomed in Athens, and Eurymedon the hierophant denounced Aristotle for not honoring the gods. Unwilling to be executed as Socrates had been, he fled to his mother’s family estate in Chalcis, dying of natural causes within the year.
Theories and models

Science provides explanations about the world through theories and models. The word “theory” has a very different meaning in science than it does in common usage: a scientific theory is a broad, comprehensive explanation of a large body of information that, over time, must be supported and ultimately confirmed (or rejected) by the accumulation of a wide range of different kinds of evidence (Box 2B).

In popular usage, the word theory usually refers to a limited, specific conjecture or supposition, or even a guess or hunch. Equating the meaning of a scientific theory with a guess has caused no end of mischief in the popular press and in public debates on politically charged issues. A well-known example is the theory of evolution. While sometimes portrayed as “just a theory” by creationists and advocates of intelligent design, it is actually a comprehensive explanation of a large number of observed patterns in nature – in fact, it is one of the best-tested theories in biology (see Chapter 4).

When a theory is buttressed over many years with strong evidence, with new findings consistently supporting and amplifying the theory while producing no serious contradictory evidence, it becomes an accepted framework or pattern of scientific thought. This is what has occurred with Newton’s theory of gravity, Darwin’s theory of evolution, and Einstein’s theory of relativity. Scientists use such overarching theories to organize their thinking and derive additional predictions about nature.

Theories are intertwined, just as all life on Earth is interconnected. The five theories of biology presented in chapters 3-7 are not independent of each other; rather, each of them relies on aspects of all of the others so that causes in one show up as consequences in all of the others. For example, the process of evolution (see Chapter 4), influences the properties of the genetic system (see Chapter 3), cells (see Chapter 5), organisms (see Chapter 6), and ecological systems (see Chapter 7). Conversely, genetics, organismal structure, and ecological interactions all influence the process of evolution. This interdependence of scientific theories requires that they be consistent with each other, which is termed consilience. Consilience extends beyond biology, so that biology must be consistent with other areas of science such as chemistry, physics, geology, and astronomy. For example, understanding the origin of life (see Chapter 1) draws on aspects of all of those disciplines.

Theories are, in turn, used to generate and organize models. A model is an abstraction or simplification that expresses structures or relationships. Models are one of the ways in which the human mind attempts to understand complex structures and relationships, whether in science or in everyday life. Building a model airplane from a kit can tell you a lot about the basic form of an airplane; civil engineers often build small models of structures such as bridges or buildings (either physical models or three-dimensional images on a computer) before construction is begun.

Models can be abstract or tangible, made of words or plastic (e.g., Watson and Crick’s ball and wire model of a DNA molecule; see Box 3A), diagrams on paper, sets of equations, or a complex computer program. In science, models are used to define patterns, summarize processes, and generate hypotheses. One of the most valuable uses of models is to make predictions. All models are necessarily based on simplifications and rest on a set of assumptions. Those (implicit and explicit) simplifications and assumptions are critical to recognize because they can alert you to the limitations of the model, and because faulty assumptions and unjustified simplifications can sink even the most widely accepted or elegant model.
Box 2B
The Nature of Theory

Theories provide generalizations of data and concepts, playing many roles and coming in many different forms depending on their use and on the breadth of the phenomena they attempt to explain. This book deals with foundational, general theories consisting of a framework of broad principles (see Tables 1.1, 3.1, 4.1, 5.1, 6.1 and 7.1). Conversely, a theory can also be a highly specific model that makes precise predictions about a particular circumstance. Between these two ends of the continuum are theories that generalize beyond a single model, but have a somewhat limited scope. A good example of this intermediate form is the theory of evolution by natural selection (see Box 4C). While theories are purely intellectual constructs, models are where the theoretical rubber meets the empirical road. Models generate hypotheses that are then tested with experiments (Figure 2.1). The results of those tests along with additional observations are then used to further refine theories and generate new models and hypotheses.

Each theory defines a domain, the scope of the theory. For example, the domain of the theory of genetics (see Chapter 3) is the form and process of information storage, transmission, and usage. Theories may have overlapping domains, and such intersections between theoretical domains are where some of the most exciting science can take place. Often, these places are where new and seemingly disparate elements are being brought together. Currently there is great interest in joining our understanding of how an individual changes during its lifetime (development; see Chapter 6) with how change occurs over generations (evolution; see Chapter 4). It is important to remember that theories and their domains are not real properties of the natural world, but human constructs, how we put structure on the world so that we can understand it.

At the heart of a theory is a set of broad statements (fundamental principles) about empirical patterns and the processes responsible for them. Fundamental principles are meant to be broad in scope, often encompassing multiple, interrelated patterns and mechanisms. They define the guiding questions of a domain. For example, the central question for evolution is why organisms change over generations in the way that they do (see Chapter 4). For genetics it is why offspring resemble their parents (see Chapter 3), and for organismal biology it is how organisms maintain themselves (see Chapter 6). The principles of the general theories of these disciplines are the statements necessary to answer those questions. Of course, those theories are not the entire story; for that, we have to examine all the details of the more specific theories and models within each general theory, a level of detail beyond the scope of this book.

Scientists are often looking for laws, statements of relationship or causation. Laws describe how particular processes result in specific outcomes or patterns. For example, Mendel’s Laws of Genetics (see Chapter 3), describe the pattern of offspring appearances, given the appearances of their parents. In biology, laws reside within specific theories but not at a more general level, because within biology there is no law that holds true across the entire domain of a given general theory.

Another role for theories is to provide a framework for guiding and evaluating research. They tell scientists where to look and how to do the looking. For example, theories about how organisms are distributed across the globe (see Chapter 7) tell us that we are likely to find many new types of organisms in forests near Earth’s equator. Scientists would then use other theories about how organisms are related to each other (see Chapter 4) to recognize a new type of organism. By providing guidance, theories help scientists be much more efficient in their search for knowledge.
Testing theories

The testing of scientific theories, especially biological ones, is a more subtle, nuanced, and complicated endeavor than nonscientists often realize. The popular image of the scientific method portrays it as a process of testing and falsifying hypotheses. This approach was codified by the German philosopher of science Karl Popper and is a form of empiricism. In this framework, we are taught that we can never prove a scientific hypothesis or theory. Rather, we propose a hypothesis and test it; the outcome of the test either falsifies or fails to falsify the hypothesis. While hypothesis testing and falsification is an important part of theory testing, it is not the whole story, for two reasons.

First, the falsification approach fails to recognize knowledge accumulation. In a strict Popperian framework, all theories are held to be potentially false. We never prove anything to be true; we merely disprove ideas that are false. This assumption goes against our own experience and the history of the accumulation of scientific understanding. Today we know that the Earth revolves around the sun, even though this was once just a hypothesis. We know that the universe is approximately 14 billion years old and began with the Big Bang, even if we still do not know the details of that event. We know that life began and assumed its present shape through the process of evolution. We know that many diseases are caused by microbial infections, not by the imbalance of “humours,” and that hereditary traits are conveyed by DNA (or in a few viruses, by RNA), not by blood itself. While we may acknowledge that all of this knowledge has not, in a strictly philosophical sense, been proven true, but has only failed thus far to be falsified, we also recognize that some knowledge is so firmly established and bolstered by so many facts that the chance that we are wrong is so small as to be nearly nonexistent. In contrast to empiricism, realism recognizes this progressive accumulation of knowledge.

Second, and more important, the Popperian framework fails to account for a second type of question that we very commonly ask in biology. Often the issue is not one of falsifying a hypothesis, but of the relative importance of different processes. When we examine the structure of a community (see Chapter 7), we do not ask, ”Is it true or false that competition is occurring?” Instead, we ask, ”How much, and in what ways, do the processes of competition and predation each contribute to shaping this community?” When we are building our theories about community structure, our activities are more akin to estimating the necessary quantities and assembling a complex model than to falsifying a set of propositions.

Falsification does play a role in science, but a more limited one than Popper envisaged. Theory construction is like assembling a jigsaw puzzle from a pile of pieces from more than one box – we can ask whether a particular piece belongs in this spot by erecting a hypothesis and falsifying it. We may even conclude that this particular piece does not belong in this puzzle. Occasionally we are attempting to completely throw the piece away, saying that it does not belong to any puzzle.

Science and other ways of knowing, revisited

Science demands internal and external consistency: ultimately, theories must be consistent with one another, and data must be consistent with theories. Other ways of interpreting the world do not share this characteristic; systems of morality or religion may or may not include obvious contradictions, but none demand consistency with data, in any sense of the term. Making science internally and externally consistent is a constant effort. Theories—even successful ones—contradict one another in places. Some experimental results seem to contradict theory at times, and even well-designed studies can contradict one another because chance events result in different outcomes.

The fact that we find contradictions simply means that we are still learning, and it is often through discovering inconsistencies that scientists are able to expand our body
of knowledge about the world. One reason for inconsistencies, and one that requires closer examination, is multicausality.

**Multicausality**

Living systems have a critical property that affects the structure and evaluation of biological theories: *multicausality*. We distinguish two types of multicausality: first are instances in which a single outcome can arise as a consequence of a number of different components, even though they are all isolated from each other. This aspect of multicausality is important for the structure of many models. If a model includes all of the multiple causes, it will provide accurate predictions or explanations.

If a model does not include all causes, the utility of the model depends on how those causes interact. For example, weight gain and loss is caused by both the amount of food eaten—energy taken in—and the amount of exercise an individual performs—energy expended. A model that attempted to predict weight loss based just on the amount of food eaten would accurately predict that people eating less food would lose more weight. However, it would fail to predict the actual amount of weight loss without accounting for the amount of exercise. This shows that the excluded factors in a model may affect absolute predictions of a model, but not relative ones. This aspect of multicausality is important for the evaluation of models and theories. The fact that diet and exercise are not the only factors that affect weight loss, for example the different rates at which people absorb the energy from food, just makes the job that much more difficult. In practice, biologists must often use multiple lines of evidence to discern the relative role of different biological processes in producing patterns.

The second type of multicausality are instances in which multiple causes act together to create an outcome. For some types of interactions, conclusions about the relative magnitudes of the processes included in the model are still accurate. For example, exercise might also affect the amount of food that someone ate. If exercising caused all people to eat less, a model might still accurately predict that more exercise led to lower weight, even if the exact amount of weight loss was not accurately predicted. But what if the effect of exercise on the amount of food eaten depended on the mass of the person, so that a thin person ate more food while a fat person ate less? In that case, a model of the effects of exercise on weight loss would also need to account for the weight of the person to predict weight gain or loss. At minimum, it is necessary to acknowledge that any particular model may not include all possible important causes and that those other causes might have unexpected effects.

Another aspect of multicausality is that some effects come about through the short-term circumstances of an individual, and others through long-term effects such as an organism’s evolutionary history. Consider the question: Why are male lions larger than female lions? A short-term explanation involves an individual’s development and food intake during growth, while a more long-term explanation involves how male lions compete with each other for mates, thus affecting how many offspring a larger lion produces. Beyond that may be effects that are common to all types of cats or carnivores. These alternative explanations often derive from different domains, so theories need to either draw on those multiple domains or to acknowledge the limitations of their explanatory scope.

**Explaining the characteristics of life**

The characteristics of living systems can be studied from four perspectives: functional, developmental, historical, and adaptational. These perspectives are alternative explanations for why those characteristics exist in the form that they do; these perspectives are often intertwined in biological studies, and it is necessary to take all of them into account in understanding living systems.
The functional perspective asks what the constraints or limitations are on how a living system can be put together and operates. For example, if you compare the legs of a deer and an elephant, the latter has much thicker legs. Those legs are not just thicker, though, they are also proportionally thicker. If you simply expanded a deer to the size of an elephant its legs would snap under the much greater weight. A functional perspective is especially important for molecular, cellular, and organismal biology (see Chapters 5 and 6).

The developmental perspective asks how the process of building an organism constrains or limits its form. For example, some types of ants can develop an amazing array of forms among genetically identical individuals (see Figure 6.ANT), while most other types of insect have only a single form. Ants are able to produce such diverse forms because of key differences in their growth processes. A developmental perspective is central to much of organismal biology (see Chapter 6).

Delving further into the past is the historical perspective, which asks how the past history of a system determines and constrains its current and future characteristics. For example, ants have six legs while elephants have four because, by chance, the ancestors of each had six and four legs, respectively. There is no necessary reason why ants could not function with four legs or with eight. A historical perspective is particularly important for studies in evolution and ecology (see Chapters 4 and 7) and understanding organismal function (see Box 6D).

Lastly, the adaptational perspective asks how the ways in which a living system interacts with its environment mold its characteristics over many generations. For example, although an elephant has difficulty standing on just its hind legs, its trunk allows it to reach parts of trees it could not otherwise. The length and flexibility of its trunk came about because of the ways in which those properties of the trunk affected feeding and other activities which, in turn, enhanced the survival and reproduction of the elephant’s ancestors. The adaptational perspective is explored in detail in Chapter 4.

**Methodologies**

Scientists gain knowledge by using the **scientific method**. They carry out a particular series of steps designed to structure their questions in as controlled and well-designed a method as possible, although not always in a fixed order (Figure 2.1). These steps can be summarized as follows: observation, description, quantification, posing hypotheses, testing those hypotheses using experiments (in a broad sense of the word, as discussed below), and verification, rejection, or revision of the hypotheses, followed by retesting of the new or modified hypotheses.

Throughout this process, scientists gather various kinds of information, look for patterns or regularities in their data, and propose processes that might be responsible for those patterns. They often put together some sort of model to help in advancing their understanding. Eventually, they construct theories, using assumptions, data, models, and the results of many tests of hypotheses, among other things. The building of comprehensive scientific theories proceeds simultaneously from multiple directions by numerous people, sometimes working together and sometimes at cross-purposes. Science in operation can be a messy and chaotic process, but out of this chaos comes our understanding of nature.

A scientific hypothesis is a possible explanation for a particular observation or set of observations. A **hypothesis** is smaller in scope than a fully developed theory, and must be testable: they must contain a prediction or statement that can be verified or rejected using scientific evidence. Experiments are the heart of science, and we discuss their design and use in more detail in the next section. A crucial characteristic of science
is the need to revise or reject a hypothesis if the evidence does not support it. Science does not accept hypotheses on faith and scientists must at all times retain a healthy skepticism. A scientist needs to remain open to the possibility that any hypothesis, model or theory may be wrong or incomplete, but recognizes that some models and theories are so well established and so well buttressed by evidence that it would be unreasonable to act as if they might be wrong.
Most people know Sir Francis Bacon as a scholar, the man whose writing formed the basis of what we know today as the scientific method. Along with a student of everything from biology and systematics to philosophy and metaphysics and a prolific writer, he was also a lifelong courtier and politician, constantly striving to raise his position in English society. His twin lives were interconnected as he sought royal patronage for his works while standing as a political adviser and minister for Elizabeth I and James I.

His political ambitions began with his father, Sir Nicholas Bacon, who was a lawyer (in a time when that profession was closely connected with Parliament and the peerage), a statesman, a privy councilor and Lord Keeper of the Great Seal, England’s highest judicial position, for Elizabeth I for over 20 years. Francis was a younger son by his second wife. He entered Cambridge University when he was 12 and studied there for two and a half years before beginning his study of law.

His father arranged a posting to France, which was cut short when his father died in 1579. Bacon was consequently forced to return to England. Left out of his father’s will (apparently by accident), Bacon had to work for his living, returning to his study of law, and beginning his work in 1582. His ambition, however, was to obtain a position in service to the crown, which was partially fulfilled in 1581. He was elected to the House of Commons, an event which marked the commencement of his occasionally rocky public career. In his political ambitions, he didn’t scruple to use people to gain advantage; in his private writings he spoke of others as merely means he could exploit to his own ends, and of using flattery and craft to aggrandize himself. He made plans to ingratiate himself with his cousin, Salisbury, who was the Lord Treasurer to Henry Prince of Wales, James I’s principal minister, while simultaneously currying royal favor with an eye to replacing his cousin in his position. Unfortunately for Bacon, his attempts at subterfuge did not work; while he did rise in position, he was ultimately used as a scapegoat in a public legal scandal.

Through his politicking, Bacon became solicitor general, later attorney general, and finally Lord Chancellor. He also was promoted to the peerage with the title of Lord Verulam and then Viscount St. Alban. He was the presiding judge of the Court of Chancery, the nominal head of England’s judiciary. Perhaps because of his further ambitions, he granted repeated requests from Buckingham for his favor in particular disputes, and was lax enough to accept gifts from suitors to court. Because of this, in 1621 when the House of Commons attacked royal grants of monopoly, financial peculation and corrupt officials, Bacon was accused of bribery. The charge against Bacon was headed by his longtime enemy, Sir Edward Coke, who he had gotten dismissed from a judgeship in 1616. When the case was sent to the House of Lords, the king let Bacon become the scapegoat for the whole mess. Bacon, knowing appearances were against him (although he maintained impartiality and could point to many cases in which he had actually ruled against those who had given him gifts) submitted to trial and confessed guilt.
He received a fine of 40,000 pounds, imprisonment at royal discretion, and was barred from holding state office, sitting in Parliament, or coming within verge of the court. Essentially, he was banished from London. However, the king showed leniency, imprisoning Bacon for three or four days, essentially remitting the fine and awarded him a partial pardon. Afterward Bacon threw himself into his writings while continually striving in vain to regain political office.

Despite his mostly thwarted ambitions in the political sphere, Bacon was a prolific writer and philosopher on the nature and role of science for mankind. In Bacon’s time, the number of disciplines and subdisciplines of knowledge was much smaller; in addition, the boundaries between the disciplines were much vaguer, or drawn differently. Great thinkers still practiced an encyclopedic approach to knowledge, although Bacon himself was unique in the sheer number of disciplines he investigated, both in what is now considered the life and physical sciences, and medicine, psychology, music, social science and crafts. He lived during the Protestant Revolution, and was necessarily concerned with religion and the ethics thereof; he urged mercy for Catholics and Puritans, although he looked to former ministers for their value in the fight against popery, urging that Catholics be discouraged but not driven to desperation.

His primary intellectual goal was to restructure what he knew as philosophy – what we today would consider science – as a form of inquiry, enabling a systematic, continual progress of knowledge. In an astonishingly modern attitude, Bacon believed that the purpose of science was for the betterment of humanity, and hoped that by establishing a more systematic approach, many improvements and benefits to human life would emerge as a result. He knew that in order to bring this about, he would have to completely reform and restructure the methods of acquiring and testing scientific knowledge. Bacon wanted to do away with the old modes of natural philosophy, with all energy aimed at preserving the knowledge of the ancient world as current.

Interestingly enough, Bacon’s motives for creating his new method of investigating the world were based in religion; he desired to use science to restore humanity to the pristine condition it had before the Fall - being cast out of the Garden of Eden. However, his philosophy had a consistent naturalism, being concerned with empiricism and the reformation of practical investigative procedure. The mix of religious or mystical inspirations for eminently practical methods was the hallmark of Bacon’s new system of philosophy; he owned his most basic inspirations – the understanding and manipulation of the natural world for the benefit of humanity – to the medieval practice of alchemy. According to Bacon, true understanding of a phenomenon and the ability to create said phenomenon were interchangeable; to know a cause of nature (verum) is to be able to manipulate nature to your own designs (factum).

Bacon’s greatest break from previous thought was the idea of the experiment as a way of manipulating nature to reveal its secrets. Previous thinkers had made a sharp distinction between the natural and the artificial. He also conceived of the use of observations for the generation of theories and hypotheses that would be tested by new observations and experiments. Previously, knowledge was conceived as a much simpler process of inducing generalizations from a few observations. Bacon also saw science as a collective enterprise based on communication among its practitioners. At the same time, he also firmly believed in secrecy where the transmission of knowledge was concerned. He wrote about publishing knowledge in such as way that either held back part of what was being taught, or obfuscated it so much as to render it incomprehensible to everyone but the most erudite and determined students. Bacon was a child of his time, an elitist and aristocrat who had a strongly paternalistic, albeit benevolent, view of most of humanity. Bacon was a key influence to the coming explosion of scientific thought in the
seventeenth century, and without Bacon’s contributions, modern science would be drastically different. Bacon died on April 9, 1626 of pneumonia, possibly contracted while stuffing a chicken with snow as an experiment in using cold to preserve meat.
A brilliant statistician and geneticist, Ronald Fisher was a rarity in modern science, making enormous contributions to both fields. Coming from rather unlikely origins, he was born in East Finchley, London, on February 17, 1890. His father was a successful fine arts dealer, but lost his business in a series of bad transactions eight months after Fisher’s mother died when he was 14. Fisher showed an early aptitude in mathematics; because of his poor eyesight, he was tutored in math without the use of pen and paper, with allowed him to develop the ability to visualize complex problems in geometric models instead of step-by-step algebraic calculations. This later became his legendary ability to produce mathematical results to problems without having to go through any intervening steps. He won early accolades for his scholarship, including the Neeld Medal (a competitive math essay) at the Harrow School when he was 16, and later winning a scholarship to Gonville and Caius College, Cambridge. It was there he discovered his passion for genetics and evolutionary science, and saw the growing range of statistical methods as a way to reconcile the discontinuous nature of the traits studied by Mendel (see Chapter 3) with the continuous variation and gradual evolution of Darwin (see Chapter 4).

He also had a keen interest in eugenics, the scientific improvement of the human condition, at a time that its reputation was still unblemished by the atrocities and gross human rights violations that would occur in Europe and the United States. Fisher saw eugenics as both a social and scientific issue in statistics and genetics, dedicating much of his professional time and energy to their study. He co-founded the Cambridge University Eugenics Society in 1911 and became the Professor of Eugenics at University College London in 1933.

Fisher graduated in 1913, at the height of World War I, and was eager to join the war effort. He failed the physical due to his eyesight, and instead began working as a statistician for the City of London, and teaching physics and math at public schools, including Bradfield College in Berkshire and aboard the H.M. Training Ship Worcester. During the war, he met and married his wife, Eileen Guinness. The couple set up a subsistence farming operation on the Bradfield Estate, raising animals and a large garden. At this time he did some of his rare empirical work by carrying on selective breeding experiments. Fisher also started writing book reviews for the Eugenic Review, and was hired in a part time position by Major Leonard Darwin, one of Darwin’s sons, with whom Fisher claimed a close friendship. He began publishing articles on biostatistics, including the groundbreaking “The Correlation Between Relatives on the Supposition of Mendelian Inheritance,” written in 1916 and published in 1918. In it, he laid the foundation for what came to be known as biometrical genetics and introduced the methodology of analysis of variance, a huge improvement over the correlation methods used previously. He showed that inheritance of traits could be measured by real values (values of continuous variables) and that it is consistent with Mendelian principles. Remarkably, in a single paper he established two entire fields, one in genetics and one in statistics, and laid the foundation stone for the evolutionary Modern Synthesis (see Chapter 4).

After the war, he was offered a job at Galton Laboratory by Karl Pearson, but
because of a developing rivalry between the two men, instead he accepted a post at Rothamsted Experimental station, a small agricultural research station in Harpenden, Hertfordshire. This decision was very fortuitous because it put him in close contact with empirical researchers who needed advice on experimental design and data analysis. There he began a major statistical study of the data the station had collected over many years, resulting in a series of reports titled “Studies in Crop Variation.” Over the next seven years, he pioneered the principles of the design of experiments and elaborated studies of analysis of variance, a systematic approach of analysis of real data as a springboard for the development of new statistical methods. He developed practical and rigorous methods for the labor necessary in statistical computations, in 1925 publishing Statistical Methods for Research Workers and ten years later publishing The Design of Experiments, which became standard reference works for scientists in many different disciplines. It was during this time he made huge strides in the field of statistics, analyzing the technique of maximum likelihood, which fits a statistical model to data, extrapolating parameters for a large population based on actual data from a smaller representative sample. He introduced the “randomization test” beginning the field of non-parametric statistics.

Along with his major contributions to statistics, Fisher was also one of the founders of the evolutionary Modern Synthesis. Fisher’s work on the mathematical underpinnings of the theory of evolution culminated in his major biological work, The Genetical Theory of Natural Selection, which he began writing in 1928 and published in 1930. In this book he summarized his work showing how the evolution of continuous traits could be explained by changes in the frequency of Mendelian genes, as well as developing ideas on the evolution of mating, mimicry and genetic dominance. Despite the enormous influence of the ideas presented in that book, it is notorious for being nearly impossible to read. Fisher used his own, very obscure system of notation, and it was mostly through the translation of his work by others that his ideas became widely understood.

In keeping with his interest in eugenics, a third of the book was concerned with the application of these concepts to humanity. He attributed the decline and fall of civilizations to their arrival at a state wherein the fertility of the upper classes is forced down. Using 1911 British census data, he showed an inverse relationship between fertility and social classes, and believed it partly due to the rise in social status of families not capable of producing a large number of children, who rose because of the financial advantage of a smaller family. To rectify this, he proposed subsidies to families with a large number of children, proportional to the wages of the father.

With the onset of World War II in 1939, University College tried to dissolve its eugenics department; although Fisher fought his decision, he was exiled back to Rothamsted with a significantly reduced staff and resources. WWII was a bad time for Fisher; he was unable to find war work, his marriage dissolved, and his oldest son George, a pilot, was killed in action. In 1943 he was offered the Balfour Chair of Genetics at Cambridge (his alma mater), which had also been nearly destroyed by the war. Fisher accepted the position with promises from the University that he would be able to rebuild it, but he was given very few resources, and it grew very slowly.

Fisher’s evolutionary theories were built around the notion that the traits of organisms were determined by many genes, each having a small effect that added up to the final phenotype. The primary mode of evolutionary change was through selection on those genes in large populations. As a result, his theory saw natural selection as the sole determinant of evolutionary change, denying any role for contingency. His collaborations with the field biologist E. B. Ford were aimed at demonstrating these effects in natural populations. He clashed frequently with one of the other founders of the Modern Synthesis, Sewall Wright, whose theories emphasized evolution in small populations and
importance of genetic drift.

Fisher was marked by an intense loyalty, both to his friends and his country. He was conservative politically and a scientific naturalist as well as member of the Anglican Church, writing articles for church magazines. Despite this, he held a firmly rational view of the world; in a 1955 broadcast on science and Christianity, he said, “The custom of making abstract dogmatic assertions if not, certainly, derived from the teachings of Jesus, but has been a widespread weakness among religious teachers in subsequent centuries. I do not think that the word for the Christian virtue of faith should be prostituted to mean the credulous acceptance of all such piously intended assertions. Much self-deception in the young believer is needed to convince himself that he knows that of which in reality he knows himself to be ignorant. That surely is hypocrisy, against which we have been most conspicuously warned.” In later years, a certain reputation for lassness in manners and dress blossomed into an archetypal nature as an absentminded professor; despite this, he was often sought after as a conversationalist. He was well recognized and well-traveled for his work, being inducted into the Royal Society in 1929 and two years later spending six weeks at the Statistical Laboratory at Iowa State College. In 1937 he visited the Indian Statistical Institute in Calcutta, which at the time consisted of a single part-time employee, P. C. Mahalanobis. Fisher returned often, encouraging their growth and development, and attended as the guest of honor in their 1957 twenty-fifth anniversary, when it had grown to 2000 members.

Fisher was dubbed a Knight-Bachelor by Queen Elizabeth II in 1952, and retired from Cambridge five years later in 1957. After Cambridge, he spent some time as a senior research fellow in Australia, and died there of colon cancer on July 19, 1962. He was awarded the Linnean Society of London’s Darwin-Wallace Medal the next year.
Experiments

A cornerstone of the scientific process is the experiment. We use the term “experiment” here in its broadest sense: a test of an idea. Experiments can be classified into three broad types: manipulative, natural, and observational. Manipulative, or controlled, experiments are what most of us think of as experiments: A person manipulates the world in some way and looks for a pattern in the response. For example, a biologist might be interested in the effects of different amounts of nutrients on the growth of a particular type of plant. She can grow several groups of plants, giving each of them a different nutrient treatment, and measure such things as their time to maturity and their final size. This experiment could be done in a controlled environment such as a growth chamber, in a greenhouse, in an experimental garden, or in a natural community in a field setting.

This range of potential settings for the experiment comes with a set of trade-offs. If the experiment is conducted in a laboratory or growth chamber, the scientist is able to control most of the possible sources of variation so that the differences among treatments can be clearly attributed to the factors being studied in the experiment. These sorts of controlled experiments exemplify the scientific method as it was first laid out by Francis Bacon in the seventeenth century (Box 2C). Baconian experiments are the mainstay of most of molecular, cellular and organismal biology (see Chapters 5 and 6) as well as the physical sciences. By working in a controlled environment, however, the biologist sacrifices something. The controlled environment is highly artificial, which compromises realism, and it is also narrow in scope, sacrificing generality because the results apply only to a limited range of conditions.

If an experiment is conducted in a field setting, it is more realistic or more natural, but now many factors may vary in an uncontrolled fashion. In a field experiment, the only factors that are controlled are the ones being studied. Instead of attempting to control all variation, variation due to factors other than the experimental ones is randomized among replicates, and conclusions are based on the use of statistics to separate effects due to the factors being manipulated from other, uncontrolled factors. Such experiments can be carried out in many settings and are not restricted to the field. The design and analysis of these sorts of experiment relies heavily on the pioneering work of Ronald A. Fisher in the early twentieth century (Box 2D). Fisherian experiments are a mainstay of ecology and evolutionary biology as well as the social sciences. They are typically less narrowly defined than Baconian experiments, and thus their results may be more readily generalized. A biologist must decide where along this continuum of control versus realism she needs to carry out her experiment based on her scientific goals as well as practical considerations.

Experiments are usually designed as tests of hypotheses. If the hypothesis is partially or wholly falsified, the scientist goes back, revises his ideas, and tries again. If the hypothesis is not falsified by the outcome of the experiment, the scientist gains confidence that his hypothesis might be correct. Sometimes, however, scientists design experiments to “poke at it and see what happens.” Even here, one or more hypotheses are being tested (though they are sometimes not stated as such): by creating a difference between groups (such as feeding some animals more than others) and then measuring some quantity (like the time until adulthood), a biologist implicitly generates hypotheses about the relationship between the manipulation and the things measured. Such experiments are common throughout the biological sciences. Scientists have studied in detail only a few hundred of the millions of species; of these, only a few such as humans, mice, rats, fruit flies, and corn approach being well studied. A biologist beginning the study of a new species or other component of living systems must do many of these
general types of experiments. Of course, she is guided by her knowledge of other similar species and ecosystems. Each species is unique, however, which is why each study expands our scientific knowledge.

Manipulative experiments are powerful tools for two major reasons: first, because the scientist can control which parts of the natural world will be altered to study their effects, and second, because she can separate factors that typically occur together to test them individually. So, manipulative experiments are well poised to separate out some types of multicausality. Such experiments have limitations, however. If multicausality is due to complex interactions of many factors, it may not be practical to create an experiment that considers all possible combinations of those factors, even if you know what they are. Sometimes manipulative experiments are plagued by artifacts—outcomes caused by a side effect of the experimental manipulation itself rather than being a response to the experimental treatment being tested. Good experiments avoid artifacts or take them into account in evaluating the results.

Another limitation is that of scale. Evolution (see Chapter 4) and ecology (see Chapter 7) are often concerned with patterns and processes that occur across large scales of space and time—for example, the causes of differences in the numbers of species on different continents, or the responses of populations to climate change over the next two centuries. We cannot do manipulative experiments at these great scales of time and space, and in many cases no true replicates might exist (continents, for example) even if we could work at these scales. Biologists are, however, increasingly making use of longer-term and larger-scale manipulative experiments in ways that mimic natural processes (see Box 7C).

Some types of manipulative experiments would be unethical to carry out. For example, we would not cause the extinction of a species just to study the effects of such an event. In such cases, biologists must rely on two other types of studies – these are natural and observational studies, which may be thought of as two different kinds of experiments.

Natural experiments are “manipulations” caused by some natural occurrence. For example, a species may go extinct in a region, a volcanic eruption may denude an area, or a flash flood may scour a streambed. Natural and manipulative experiments represent a trade-off between realism and precision, similar to the trade-off between laboratory and field experiments. Just as with a manipulative experiment, the biologist compares the altered system either with the same system before the change or with a similar, unchanged system.

The major limitation of natural experiments is that there is never just a single difference before and after a change or between systems being compared. There are no guarantees, for instance, that the altered and unaltered systems were identical prior to the event. For example, if we are comparing areas burned in a major fire with others that remained unburned, the unburned areas might have been wetter, might have had a different site history or different vegetation before the fire, and so on. Therefore, it can be difficult to determine the cause of any one particular change. The best natural experiments are ones that repeat themselves in space or time. If a biologist finds similar changes each time, then she gains confidence about the causes of those changes. Another approach is to combine natural experiments with manipulative experiments.

Observational experiments consist of the systematic study of natural variation. Such observations or measurements are experiments if an biologist starts with one or more hypotheses (predictions) to test. For example, one could measure patterns of species diversity across a continent to test hypotheses about the relationship between the
number of species and rainfall. Or, one could compare the process of growth of different species of fish, with some that live in the ocean and some that live in freshwater, to test hypotheses about how organisms cope with differences in water salinity. An advantage of natural and observational experiments is that biologists can let nature tell them what the multiple causes are. Statistical procedures developed in recent decades are able to tease apart complex, causal relationships, adding a third alternative to the approaches of Bacon and Fisher.

An important limitation of this type of experiment is the need for the scientist to know which of the many possible factors should be measured. Again, as with natural experiments, there is the potential for multiple factors to vary together. If several factors are tightly correlated, it becomes difficult to determine which factor is the cause of the observed pattern. This circumstance is when multiple or repeated studies become extremely beneficial; as with natural experiments, observational experiments repeated in space or time or among different groups of species add confidence to our conclusions. Other sciences, notably geology and astronomy, also rely largely or exclusively on observational experiments because of the spatial or temporal scales of their studies or because direct manipulation is impossible.

**Creativity, objectivity, and subjectivity**

When you read a typical scientific paper, it may at first seem esoteric and dull. The format follows a rigid protocol, designed for efficiently conveying essential information to other scientists. Ideas are tightly packaged, with a clear logical line running from start to finish. It may seem as if the researchers knew exactly what they would find even before they began. We will let you in on an open secret: That is not usually how real science works. The justifications for the research presented in a paper’s introduction may have been thought up or discovered long after the research project began, or even after the work was finished. Because of serendipitous discoveries, laboratory or field disasters, or unusual natural occurrences, the original purpose of a research project is sometimes modified or, occasionally, entirely discarded and replaced with something else.

Ideas in science, especially in biology, come from a variety of sources. While everyone knows that science is objective and rational, that is only half the story. In order to reach a genuinely new understanding, subjectivity and creativity must also come into play. While one must be objective in, for example, examining the weight of evidence in support of a hypothesis, subjectivity plays a subtle but important role throughout all of scientific research. What one chooses to study is a subjective decision. Given that choice, there is usually a range of possible places to look for answers – another subjective decision. All of these choices largely depend on the questions one asks, and while determining the answers must be objective, choosing what questions to ask, and how to ask them, is largely subjective.

Many scientific endeavors are highly creative as well. Coming up with a good experiment, looking at a seemingly intractable problem from a new perspective, switching gears after a disastrous laboratory failure to extract a successful outcome from the jaws of catastrophe, and pulling a large number of disparate facts together to build a comprehensive theory all require a high level of creativity. A scientist must be able to come up with multiple courses of action when faced with problems, and cannot become too rigid in his work, or risk failure or stagnation.

As with other creative endeavors, the inspiration behind an experiment can come from almost anywhere. Many discoveries start with casual observations, such as Newton’s apocryphal apple. An idea might also arise as a “what if” thought: What if the world works in a particular way? A scientist may draw on precious discoveries or earlier experiments of her own which have raised new questions. What makes a scientist
successful is the ability to recognize the worth of these casual observations, what-if thoughts, and new questions. From these sources, a scientist constructs hypotheses and designs experiments to test them.

There is a distinction between the kind of research a scientist does and the kind of research done for a term paper, or by any member of the public trying to gather information about a topic using textbooks (such as this one), other library books, or material posted on websites. Although there are exceptions, research carried out by students or the general public is usually secondary research: gathering data or confirming facts that are already known. This sort of research is not only useful, it is essential: every scientific study must begin by assessing what is already known. But the heart of what research scientists do is primary research: gathering information that no one has ever known before, or coming up with new, testable ideas about how nature works. In recent decades, with the explosion of scientific publications and the advent of electronic databases, synthetic analyses of existing data has burgeoned as a new form of primary research; these efforts differ from secondary research in that they aim at assembling data in new ways and testing hypotheses using those data. These experiences of discovery are what makes doing science so incredibly exciting and fun.

Science as a Social Activity

Obviously, scientific research is not carried out by a cadre of mindless robots. Scientists are human, which means that human concerns and aspirations affect the scientific process. While that process is designed to maximize our ability to understand the world and decipher reality, sometimes aspects of human behavior can intrude and hamper the process. In some cases, however, those aspects can enhance the process.

Today, the publication of results is central to the scientific process; as the goal of science is to provide understanding, a scientific study is incomplete until its new piece has been added to the overall puzzle. It is not sufficient to simply announce one’s conclusions, either; the scientist must describe how she arrived at those conclusions and provide the data used to reach them so that others can independently decide if the conclusions are justified and can replicate the study if necessary. However, science was not always done in this fashion. In the Middle Ages, the practice was to hoard one’s knowledge, or to provide conclusions without justification. This began to change in the 17th century when the British Royal Society and the French Royal Academy established the first scientific journals. Scientists were encouraged to put their studies, including complete descriptions of methods and results, in those publications for others to read. A social contract was forged. In return for publishing one’s research, a scientist received recognition in the form of citations to that research in later papers. In this key respect, science differs from nearly all other professions in that the primary reward is neither money nor power, but recognition and status.

There are other, sometimes more tangible rewards for scientists as well. Scientists may be driven by a desire to alleviate a societal problem, such as curing a disease or preventing global warming. Such goals create enormous satisfaction when achieved. If the scientist is also an academic, he may enjoy the lifestyle. For the most part, you are your own boss and able to pursue projects of your own choice, although the continual scramble for research funding can interfere with that ability. Finally, there is simply the desire to know. Scientists spend long hours in the laboratory or slogging through miles of swamp because they have a question that they want answered. In this regard, scientists are like the private detective in a crime novel who keeps on looking for the murderer despite being beaten up and threatened with death.

Sometimes the questions that a scientist tries to answer are driven by her own
curiosity, and sometimes they are determined by others. Frequently, those others are the ones who hold the purse strings, who can be members of a government, a not-for-profit foundation, or a for-profit company. Such control is not necessarily bad as it can push scientists towards questions that will have immediate benefits to society. Often the benefits of research may not be immediately obvious; the payoff may not come for decades. That is why foundations and governments fund curiosity-driven research, because they hope for that long-term benefit. Even then, some standards need to be applied to decide which research is more promising than others; that is one role for peer review.

**Peer review**

Despite the popular image of the lonely scientist toiling away for years, science is a highly social activity, albeit carried out by people who may lack some social skills. Central to the scientific process is peer review, the notion that one’s scientific peers are the best judges of the worth of a scientific study. Peer review happens most notably as part of the publication process. Before a scientific paper or book gets published, it is evaluated by multiple people. A manuscript may be revised several times, and sometimes additional experiments may be necessary, before a study is deemed ready for publication. Research grant money is given out based on the results of peer review. In academia, tenure and promotion decisions frequently rely on evaluations of scientists from others outside the institution who can judge the quality of the science that has been done. The bestowing of academic degrees, especially the all-important Doctor of Philosophy (Ph.D.), happens through a review of a panel of people who themselves have Ph.D.s. The scientific community decides what counts as good science and who gets to call themselves scientists.

Such self-regulation by the scientific community is not just insularity for its own sake. Peer review makes science more efficient, and helps the scientific community reach consensus on the most important questions to pursue and the most appropriate methods for achieving the answers. In the biological sciences, there are millions of species that could be studied, each one with its own unique natural history. However, some species are better at illuminating a general truth than others; for example, Charles Darwin’s formulation of the theory of evolution came about, in part, because of his studies of a particular set of finches on the Galapagos Islands (see Box 4A). While he or someone else would undoubtedly have eventually formulated a similar theory through other observations, those Galapagos finches had the virtue of clearly laying forth critical pieces of the puzzle. Today, the methods for determining the information contained in a molecule of DNA is changing continuously and, importantly, the costs keep dropping with each advance. Peer review of funding requests helps ensure that the most up-to-date and cost-effective methods are used. Peer review of manuscripts helps to clarify the writing, if not always as much as it should, and to make sure that conclusions match the data.

But does such constant peer review just result in conformity rather than achieving better understanding? While the social constructivist viewpoint would say that it does, there are two ways that a scientist can gain recognition. On the one hand, she can follow the crowd, accept the reigning consensus on some problem, and add one more brick to the scientific edifice. Or she could be an iconoclast and challenge the consensus (see Boxes 3C, 4E and 5D). Challenging a consensus involves more than simply claiming that you are right and everyone else is wrong. In science, ultimately the evidence prevails. You might think that the process of peer review would work against such iconoclasts. It does to the extent that scientists are skeptical of any attempt to throw out a well established consensus involving either a fundamental principle or a methodological approach. If peer review is about achieving efficiency, then you need a good reason to
throw out something that is working. But scientists are well aware that well established ideas and good methods have been tossed in the past. Ideas prove to be wrong, methods are shown to be inaccurate, and scientists are always willing to give alternatives an opportunity to prove their worth. There are many examples of vigorous debates in the scientific literature where the participants have not just allowed, but encouraged the airing of the views of the other side. Sometimes the scientific community can be pigheaded and a scientist must be willing to spend years convincing her peers that she is correct. In the long run, however, it is these rebels who get remembered and cited.

**Ethics**

Scientific ethics involves three central components: maintaining the integrity of the conclusions reached by the research, taking responsibility for the consequences of how a study is conducted, and taking responsibility for the way that the results are used. Scientific misconduct consists of data fabrication or falsification, plagiarism, or other serious deviations from accepted scientific practices. Data fabrication or falsification means that the results of the research were either made up or were manipulated in some way so as to alter their meaning. Plagiarism means using someone else’s ideas, results or words without giving proper credit. Both types of activities are treated as serious breaches of community norms.

Unlike what you hear about the treatment of misconduct in other professions, the scientific community tightly polices itself and deals with such breaches harshly. The reason is that unlike other professions, those most harmed by scientific misconduct are other scientists. If a doctor or lawyer commits misconduct, the victim is the patient or client, not other doctors or lawyers. But if a scientist publishes false data, other scientists may spend years of effort and hundreds of thousands of dollars pursuing a phantom. Because citation is a primary currency of the scientific community and scientists strive for recognition, stealing another’s ideas violates the heart of the scientific enterprise. Thus it is no surprise that misconduct receives thorough, if not always swift, punishment. It also explains why misconduct by scientists is much rarer than in other professions. False or misleading data, if important enough to be worth publishing, will eventually be discovered when used or checked by others. And plagiarism is almost impossible to hide in today’s world of electronic publications.

When it comes to any particular subdiscipline or question, the number of scientists working in that area is typically rather small, no more than would be found in a small village or town. Just like in a small town, everyone knows everybody else and gossip travels fast. A scientist must trust the data published by others. While in theory any experiment must be repeatable by others, in practice very few experiments are exactly repeated because to do so is inefficient. Why repeat an experiment when there are many more questions to be answered and experiments to be conducted? So, scientists must trust the other inhabitants of their small town and once you lose that trust it is hard to regain it. Scientists learn to be honest and careful, or they quickly stop being scientists.

**When is it not misconduct?**

Scientific misconduct does not include honest error or differences of opinion. Controversy actually plays an important part in science. During the process of amassing evidence regarding the validity of a theory, different interpretations of experimental data and different weight given to various pieces of evidence will lead scientists to differing opinions. These opinions may be passionately held and forcefully argued, and discussion can sometimes become heated. As the evidence supporting a theory accumulates, some scientists will be willing to accept it sooner, while others will wait until a larger bulk of the evidence is accumulated.

It is not misconduct to assert a particular conclusion even if others think that the
data do not warrant that conclusion, as long as the data themselves are available for examination. Honest errors occur, either in conducting experiments or in preparing manuscripts for publication. Sometimes the rush to publish and establish priority can result in sloppiness. As long as you clearly correct the record as soon as you discover your error, it is not considered misconduct. However, continued sloppiness may lead others to not completely trust your results.

There is one important nuance that scientists must consider when preparing their results for publication. Since data falsification includes omitting data so that the research is not accurately represented, it raises the question of whether you must always publish every single bit of data collected. Part of the purpose of a scientific paper is to provide understanding of the world in the most efficient way possible. To that end, it would be counterproductive to obscure a result with all of the false starts and missteps along the way. Instruments malfunction and humans make mistakes. Experiments may be done to refine a method or to gather preliminary data that will inform the design of a much bigger study, and while these studies are useful for a scientist in getting to the important data, the greater bulk of the information may not be of interest to anyone else. It is legitimate for a scientist to decide which components of a study merit publication and which data are erroneous. These are judgment calls, and, as with any other opinion scientists may disagree on the correct course of action. The scientist is responsible, however, for keeping a record of all that went on during a study in case questions ever come up about how it was conducted and so that other scientists, who may disagree on these issues, can make their own judgments.

**Responsibilities for ethical conduct**

Scientists must take responsibility for the conduct of their research and the use of their conclusions. For example, ecologists study the effects of interactions among species (see Chapter 7). One method for measuring such effects is to exclude one species from an area, say a fence that keeps foxes out while letting in the mice that they feed on. One way to conduct such an experiment would be to make a species go extinct, the ultimate exclusion. Clearly such an experiment would be unethical. Other situations are more complex, however, involving instances where one must weigh the ethical values on either side of an issue. For example, what amount of pain or harm to animals should be allowed if the experiment will lead to the prevention of human disease? These sorts of questions can arise even when the subjects are humans; since humans must be allowed to give informed consent about their participation, how do you obtain informed consent if the individual is not conscious, for example in a study of the best treatment for an accident victim? These and other such issues are continually being wrestled with by scientists in conjunction with ethicists.

When data are collected in an unethical manner, they are shunned. The Nazis performed horrible experiments on people to see how they would respond to traumatic conditions, and while those data are unique and might provide life-saving information about the treatment of severely injured people, by consensus of the scientific community, those data are never used. Today, scientific funding agencies and journals require that studies of humans and many other animals be reviewed by special boards to ensure ethical treatment before they will be funded or published.

Scientists must see that their conclusions are used in an accurate manner. Scientists are not policy makers, although the results of science are often, or should often, be used to inform policy decisions. For example, global climate change is a critical issue today, and one approach to ameliorate the problem is to replace fossil fuels with fuels made from plants growing today, such as ethanol made from corn or sugar cane. Another approach would be to replace those fossil fuels with electricity generated by wind
turbines. Scientists cannot make decisions about which alternative to pursue. However, scientists have a responsibility to see that all of the positive and negative effects of the alternatives are presented fairly. The use of various crop species have different effects on both the human food supply and the conversion of forests or prairies into agricultural fields. Wind turbines can cause mortality of birds and bats. While a scientist might have his or her own preferences, for example favoring saving tropical rainforests over temperate songbirds, that scientist must present all of the information needed by the policy makers for informed decisions.

Because scientists need to trust each other, it is important that scientists report on any possible biases that they might have. Some sorts of biases are inevitable because scientists are human. If a scientist has presented a new or daring theory that, if proven true, would earn a Nobel Prize, then that scientist will be looking for results that support the theory. Those sorts of biases are generally obvious and other scientists can be suitably skeptical. More pernicious are hidden biases and conflicts of interest, such as when a scientist has a monetary stake in the outcome of a study. Many journals require authors to provide information on such conflicts when submitting a manuscript for review.

Peer reviews are usually done blind so that the person being reviewed does not know who the reviewer is. Such anonymous reviews are felt to be more honest because the reviewer does not need to worry about a negative review leading to retaliation or a friendship destroyed. In such cases it is the responsibility of the referee, the editor of the journal or the grant-giving agency, to ensure that the reviewer does not use that anonymity to promulgate a biased opinion. Scientists are members of their cultures and not immune to prejudices based on race, sex or other conditions. Unfortunately, the conduct of science has been affected by such prejudices (see Boxes 4D and 5B). Today, several journals use a double-blind review process where the identity of the author(s) is not known to the reviewer. One comforting result is that in at least some areas of biology, it has been shown that the acceptance rate of manuscripts by women is no different using single-blind and double-blind review. While scientists are not angels, ultimately scientists are interested in understanding the world and are willing to accept information from all types of people to gain that understanding.
Chapter 3
Genetics

Attend any family reunion or holiday gathering and certain similarities shared among the attendees quickly become obvious. The same is true across the entire spectrum of life. Relatives resemble each other in both appearance and behavior; offspring look like their parents and siblings act like each other. That resemblance is determined by information, and the way information is transferred. This information is at the heart of life itself, and those similarities form the basis of how the great diversity of life on earth arose from the simple envelopes of self-replicating molecules proposed by the Replication First and Metabolism First hypotheses (see Chapter 1). This holds true for the differences as well as the similarities; after all, the resemblances we see are rarely perfect, and sometimes relatives may look and act nothing like each other. As explained in Chapter 1, information exists in the ordered complexity of living systems, particularly in the order of the base-pairs in a DNA molecule, just like information exists in the specific ordering of the letters in a sentence. In this chapter, we explore how the information systems of organisms work, how that information is transmitted from one generation to another, and how that information is used.

Living systems come from other living systems, resulting in a tree of life and connections among all species (see Figure 4.2), and it is the transmission of information that is responsible for that continuity. Evolution requires variation (see Chapter 4), and in this chapter we will look at where that variation comes from. The concepts presented in this chapter are framed in general terms of information, rather than in specific terms of DNA, RNA, and so forth. The reasons are two-fold. First, the theory of genetics presented here could apply to any living system, even ones based on other biomolecules, as might exist on other planets. Second, even for life on Earth, there are many variants in the details of how the genetic system works. The theory is written to encompass all of those variants, and in this chapter we will explore the major variants.

The first question we must address is a seemingly very basic one: What is information? One way to think about information is that it is a sort of symbol that represents something else, such as the way the letter “m” represents a particular sound. The more types of symbols, the more possible information. But information is more than each individual symbol – it is also how those pieces are combined to create new information. This creation of new information through new combinations is an example of an emergent property (see Chapter 1).

The Basics of Genetic Structure and Duplication

The gene is the fundamental unit of information. While genes are composed of parts that themselves act as information, it is the gene itself that turns information into a biological molecule that determines the characteristics of organisms. Our understanding of what a gene is and how it functions in information storage, transmission, and usage has grown considerably in the past 150 years.

Genes are, in turn, assembled into larger structures called chromosomes. Those chromosomes come in two forms: closed circles and open strings. Of the three domains of life (see Figure 4.2), the Bacteria and Archaea have closed circles and the Eukaryota have open strings. Before we explore how the information contained in these genes and chromosomes are transmitted and used, we must first examine the basics of their structure.
The basic building-block of the chromosome is deoxyribonucleic acid, DNA. While all DNA units contain one part that is the same for all types, the sugar-phosphate portion is the same for all types. They differ in the structure of the base. The paired bases are held together by hydrogen bonds, which are weaker than the chemical bonds that hold the rest of the molecule together. (Source: Wikipedia) The DNA molecule is twisted, forming a double helix. (Created by Richard Wheeler, Source: Wikipedia)

The basic building-block of the chromosome is deoxyribonucleic acid, DNA. While all DNA units contain one part that is the same for all types, the sugar-phosphate portion (Figure 3.1), they differ in their bases. There are four types of bases that result in four types of DNA units: adenine (A), guanine (G), thymine (T), and cytosine (C). These bases are strung together in long strands that link up the sugar-phosphate portions, with the bases sticking out the side. A DNA chromosome consists of a series of two of these strands arranged so that the bases of one strand match up with the bases of the other. The form of the bases is such that A and T fit together, while G fits with C. The structure of the chromosome is somewhat like a ladder with the sugar-phosphate portion forming the side poles and the base pairs forming the rungs. The ladder is twisted, resulting in a structure known as a double-helix (Figure 3.1B). The double-helix strand is then further wound around a protein forming a much larger structure, the chromosome (Figure 3.2). The packing is so efficient that if all of the DNA contained in one human cell were stretched out end-to-end, it would be 1.8 meters long.

The structure of DNA is essential to transmission of information from parent to offspring, including from a parental cell when it divides to create two daughter cells. During that process, the chromosomes can first be seen to duplicate, then to be split up so that each daughter cell receives one copy of each chromosome. The double-helix structure is physically unzipped, or split down the middle, opening up the A-T and G-C pairs. The
new DNA units – previously synthesized by the cell – are then lined up with each strand so that each of the two strands becomes the template for a new, matching strand. Because the bases are complementary, A with T and G with C, the end result are two chromosomes, each a duplicate of the original (Figure 3.3). The process of chromosome duplication and sorting into the daughter cells is known as **mitosis** (Figure 3.4).

Living systems can carry their information in multiple places. All cells have at least one chromosome that is the primary information carrier. In Bacteria and Archaea, besides the main chromosome, a cell might additionally contain one or more small, circular pieces of DNA called **plasmids** containing just a few genes which will also get replicated and passed along to daughter cells. Eukaryotes have multiple chromosomes. In addition, nearly all eukaryotic cells have specialized substructures that are responsible for processing energy called **mitochondria**, and plants have specialized structures for capturing light energy called **chloroplasts**. Both mitochondria and chloroplasts have circular chromosomes similar to those of Bacteria. We look at both in detail in Chapter 5.
The duplication of chromosomes can be seen during the process of cell division when a parental cell gives rise to daughter cells. During that process, the chromosomes duplicate and split up so that each daughter cell received one copy of each chromosome. (A) Chromosome duplication in Eukaryotes which have linear chromosomes contained in a specialized structure called the nucleus (see Chapter 5). (From The Science Primer, NCBI, NIH, Source: Wikipedia) (B) Chromosome duplication in Bacteria and Archaea which have circular chromosomes that are attached to the cell membrane. (Created by Ecoddington14, Source: Wikipedia)
Box 3A
Critical Experiment: The Structure of DNA

The discovery and elucidation of the structure and duplication process of DNA came about through performing a series of experiments and building various models. Early in the 20th century it was recognized that chromosomes were the structures that stored and transmitted genetic information. Analyses revealed that chromosomes contained both DNA and proteins (Figure 3.2), but nothing was known about its structure. Initially it was thought that the proteins were the information carriers and the DNA was just there for structural integrity. This was because it was known that DNA contained just four types of bases (Figure 3.1), while proteins contained twenty different kinds of amino acids (see Box 1A). This greater variability of the basic units of proteins was thought to permit it to hold more complex information.

The first critical breakthrough came with experiments designed to explore the method by which information is transmitted, and ultimately demonstrated that DNA was the information carrier. In one of those experiments, Alfred D. Hershey and Martha Chase of the Carnegie Laboratory of Genetics used a very simple system, a bacterium and a virus that infected it. Hershey and Chase knew about previous experiments that seemed to show that DNA was the information carrier, but which had not yet convinced other scientists. They thought that they could demonstrate conclusively that it was the DNA, rather than amino acids, that serves as the main carrier of information in living organisms. In their experiment, the virus they used consisted of a small piece of DNA packaged inside of a protein, and had the advantage that the DNA and the protein were separate structures, not tangled up together like most chromosomes.

They exploited the fact that DNA and proteins have a key difference in their chemical make-up: DNA contains phosphorus, but no sulfur, while some amino acids – and thus proteins – contain sulfur, but no amino acids contain phosphorus. They grew batches of bacteria and viruses in broth that contained either radioactive phosphorus or radioactive sulfur, which would be absorbed by the bacteria and viruses, effectively labeling their proteins and DNA.

Using a centrifuge, they could then separate the viruses from the bacteria so that they had pure samples of the virus that were labeled with either the radioactive phosphorus or radioactive sulfur. These radioactively-labeled viruses were placed in new batches of bacteria that had no radioactivity. After just a short...
time, the solution was shaken up and spun in a centrifuge so that any parts of the virus remaining outside the bacteria would be washed away, and only the parts of the virus that entered the bacterial cells were left. They then measured the radioactivity of those cells and found radioactivity only in the cells exposed to the phosphorus-labeled viruses. The viral protein did not enter the cell. Further, if the viruses were allowed to reproduce, the new viruses contained radioactive phosphorus. These studies, published in 1952, convinced scientists that DNA was the information carrier.

The British scientist Francis Crick and the American scientist James D. Watson are credited with determining the structure of DNA. They met each other while both were working at Cambridge University in the early 1950s. Although credit is usually given exclusively to them, two clues from other scientists were critical in finally determining the structure of DNA. The first clue came from Erwin Chargaff of Columbia University who published an analysis of the DNA of many species in 1950, in which he found that the amount of adenine always matched the amount of thymine and the amount of guanine always matched the amount of cytosine. The second clue came from a process of photographing the DNA molecules using X-rays. Rosalind Franklin had taken such photographs using samples prepared by Maurice Wilkins. Unlike ordinary photographs, this type of X-ray photograph results in a series of darker and lighter spots that require skilled knowledge to interpret.

There is still controversy over how Watson and Crick obtained copies of those photographs, whether the photographs were made publicly available, whether Wilkins had shown them to Watson on a visit to Wilkin’s laboratory, or whether Watson had simply seen them during that visit. What is not in dispute is that Watson and Crick realized that the photographs indicated that the structure of DNA was helical, something that Franklin and Wilkins did not. Armed with all of this information and using wire models, Watson and Crick were able to deduce the structure of DNA. For this discovery, they and Wilkins
The Theory of Genetics

The fundamental principles of the theory of genetics (Table 3.1) trace back to the pioneering work of Gregor Mendel in the 19th century (Box 3B). However, unlike the theory of evolution (see Chapter 4), no single person set forth this theory; rather, the principles were developed by many people over the course of the 20th century. One such critical development was the determination of the structure of DNA in the middle of that century (Box 3A). Since that time the focus of most genetics studies has been the molecular details of the structure of genetic material, how it is replicated, and what regulates the usage of that information. As with other aspects of biology, discoveries in genetics have led to further developments in other fields as well; for example, the discoveries in genetics during the first half of the 20th century were particularly important for our understanding of the process of evolution. It was bringing together that theory with the developing theory of genetics that lead to the development of the today’s understanding of the principles of the theory of evolution as the Modern Synthesis (see Chapter 4).

Today, new technological advances are driving our deepening understanding of genetic mechanisms. For example, primary to the understanding of genetic systems is knowledge of the sequence of base pairs in DNA molecules. The cost and speed of determining those sequences continues to drop by an order of magnitude about every five years. The original effort to determine the sequence of the DNA of a human - which consists of 3 billion base pairs - was begun in 1990, took 13 years, and cost about 3 billion dollars. There is now talk that by 2015, you will be able to have your personal DNA sequence for only $10,000. However, knowing what that sequence means and how it influences your life will take a bit longer, and will come through the application of the principles of the theory of genetics.

Table 3.1. The fundamental principles of the theory of genetics

<table>
<thead>
<tr>
<th>Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Offspring resemble their parents.</td>
</tr>
<tr>
<td>2. The information system requires an error correction system.</td>
</tr>
<tr>
<td>3. The information system must be capable of producing new information.</td>
</tr>
<tr>
<td>4. The imperfections of error correction create new information.</td>
</tr>
<tr>
<td>5. The exchange and recombination of information create new information.</td>
</tr>
<tr>
<td>6. Random processes play an important role in the information system.</td>
</tr>
<tr>
<td>7. Information usage must be robust to changes.</td>
</tr>
<tr>
<td>8. Information usage is context dependent.</td>
</tr>
<tr>
<td>9. The properties of information systems are the result of evolution.</td>
</tr>
</tbody>
</table>

Resemblance of Relatives

Understanding the information in your DNA sequence begins with an understanding of why you look like your relatives (Table 3.1, principle 1). The first steps were the working out of the rules by which the characteristics of parents were manifest in their offspring, and were first explained in multicellular organisms such as plants, mammals, insects, and fungi. This is not surprising as those are the most visible species and the easiest to manipulate and measure. Most of these early studies focused on characteristics such as form, color, and size. Technological advances made the measurement of molecular characteristics easier and easier, and by the middle of the 20th century much of the focus shifted to single-celled organisms, especially Bacteria and their viruses. Those
organisms had much smaller chromosomes than those of Eukaryotes and so were easier to sequence and study. Now, with modern sequencing technologies, the focus has shifted back to Eukaryotes.

The earliest geneticists were confronted with trying to explain two types of characteristics. One type consisted of discrete categories, such as the color of the eyes of fruit flies which could be red, brown, white and so forth (Figure 3.9). The other type was continuous, such as height or weight. We will explore each in turn, and then show how the two apparently discordant types can be explained by a single theory.

**Discrete traits**

In order to explore discrete traits, we start with a simple example, flower color, and consider the case of a species that shows three different colors: red, pink and white. We then perform an experiment designed to model how this trait is passed from parents to offspring. First, we mate individuals with red flowers with others that also have red flowers. We then plant the seeds and grow up the offspring and find that they all have red flowers. We do the same with white-flowered individuals and get a similar result, with all of their offspring having white flowers. Then we mate an individual with red flowers with one with white flowers, and find that their offspring have pink flowers, as if they had taken the colors of the two parents and averaged them together. We then take two pink-flowered individuals and mate them, and their offspring are a mixture of red, pink and white in proportions 1/4, 1/2 and 1/4 respectively. How can we explain these results?

First, we start by postulating that flower color is determined by a gene (C). We assume that there are “red” versions of this gene (that we will indicate as C\(_R\)), and “white” versions of this gene (C\(_W\)). Different versions of the same gene are called **alleles**. Red-flowered individuals have the C\(_R\) allele which they pass on to their offspring who then also have red flowers. The same is true for white flowers. But where do the pink flowers come from?

We can explain both the pink flowers and the variation in their offspring if we postulate that each individual has two copies of the flower color gene and one copy is passed along to a given offspring. A red-flowered individual has a genetic make-up of C\(_R\)C\(_R\). When it mates it produces reproductive cells, called **gametes**, that contain one copy of that gene (C\(_R\)). Another red-flowered individual does the same. Those gametes combine to form a new individual, again with two copies of the gene (C\(_R\)C\(_R\)). This system of segregation of the two copies of the gene into separate gametes explains both why it takes two individuals to reproduce and how the amount of information or genetic material stays the same from one generation to the next.

Pink-flowered individuals are now easy to explain. They receive a C\(_R\) allele from one parent and a C\(_W\) allele from the other, resulting in a genetic makeup of C\(_R\)C\(_W\). If each allele contributes to half of the color, the resulting individual has a color that is
intermediate between the two parents. Next, when we cross two pink-flowered individuals, we have offspring with four possible sets of alleles (Figure 3.5). In this example, parent 1 produces gametes containing either the $C_R$ or $C_W$ allele. To keep track of them we will designate them as $C_{R1}$ and $C_{W1}$. Parent 2 does the same: $C_{R2}$ and $C_{W2}$. So we end up with the following offspring: $C_{R1}C_{R2}$, $C_{R1}C_{W2}$, $C_{W1}C_{R2}$ and $C_{W1}C_{W2}$. If the alleles of a given type are identical no matter which parent they come from, we end up with 1/4 of the offspring being $C_R C_R$, 1/2 being $C_R C_W$, and 1/4 being $C_W C_W$, thus explaining the observed proportions of flower colors in the offspring. These observations form the basis of Mendel’s Law of Segregation (Table 3.2).

**Table 3.2. The propositions that result in Mendel’s Laws of Segregation and Assortment**

1. There are alternate forms of genes.
2. An individual carries two copies of each gene. Those copies may be the same form or alternate forms.
3. An individual passes one copy of a gene to each gamete and an offspring receives one copy from each parent.
4. Pairs of genes segregate (separate) during gamete formation and the fusion of gametes at fertilization recreates gene pairs. (Law of Segregation)
5. Different pairs of genes segregate independently during gamete formation. (Law of Assortment)

Figure 3.6

Similar to the previous example, here flower color is determined by one gene with two variants $C_R$ and $C_W$. Individuals with two copies of the $C_R$ allele have red flowers and those with two copies of the $C_W$ allele have white flowers. However, because the $C_R$ allele is dominant over the $C_W$ allele because individuals with one copy of each have red flowers. (Created by Benutzer: Magnus Manske, Source: Wikipedia)

One of the experiments performed by Mendel involved flower color. Unlike the example shown in Figure 3.5, here there are only two colors, red and white. When a red-flowered individual is mated with a white-flowered individual, all of the offspring have red flowers. Rather than look like an intermediate of the two parents, the offspring look like just one of them. If two of those individuals are mated with each other, 3/4 of the offspring have red flowers and 1/4 have white flowers (Figure 3.6). These results can be explained by Mendel’s Law of Segregation (see main text) with one critical difference, the way in which individuals with mixed alleles appeared. The red-flowered parent had alleles $C_R C_R$, the white-flowered parent had alleles $C_W C_W$ and their red-flowered offspring had alleles $C_R C_W$. In this case, the red allele is **dominant** over the white allele. Dominance is a type of emergent property since the flower color is not a simple averaging of the effects of each separate allele.

The Law of Segregation matches what we know about chromosomes. In most
Eukaryotes, the chromosomes come in pairs. In this case, we refer to such individuals as being diploid, which comes from the Greek word “diplos” meaning double. Individuals with only one copy of each chromosome are called haploid, while those with more than two copies are call polyploid. For example, humans are diploid with 23 pairs of chromosomes that are distinguishable by their length (Figure 3.7).

A given pair contains the same sets of genes, but may have different alleles. When gametes are formed, one member of the pair goes to one gamete and the other to the other gamete (we look at the process in more detail later in this chapter). Thus, most Eukaryotes have two types of doubles in their DNA. A single chromosome consists of a double-helix of two complementary strands, and those chromosomes come in pairs. We emphasize, though, that this is not a universal pattern. Some viruses consist of only a single strand of DNA, and many species have only a single chromosome or may have chromosomes that come in three, four or more copies. We will come back to this issue later in the chapter.

The second great insight of Mendel has to do with the inheritance of two traits at a time. Consider the following example: peapods can come in a variety of colors and shapes. Assume that there are three pea colors: green, yellow-green, and yellow. Similarly, there are three pea shapes: smooth, partially wrinkled, and very wrinkled. We start with two types of parents, one with smooth yellow seeds and one with very wrinkled green seeds. These are mated and the offspring all have partially wrinkled, yellow-green seeds. Two of those individuals are mated. Among their offspring are all nine possible combinations of seed colors and shapes in the following proportions: 1/16 smooth and yellow, 2/16 smooth and yellow-green, 1/16 smooth and green, 2/16 partially wrinkled and yellow, 4/16 partially wrinkled and yellow-green, 2/16 partially wrinkled and green, 1/16 very wrinkled and yellow, 2/16 very wrinkled and yellow-green, and 1/16 very wrinkled and green (Figure 3.8). These results can be explained if there are two genes each with two alleles, one gene determining seed color (C), with yellow (CY) and green (CG) alleles, and one gene determining seed shape (S) with smooth (SS) and wrinkled (SW) alleles. The parents have genotypes SSSCYCY and SWSCGCG, and their offspring have genotype SSSWCYCG (genotype refers to the genetic makeup of an individual). Those individuals can produce four types of gametes, each carrying one copy of the color-determining gene and the shape-determining gene: SSCY, SSCG, SWCY and SWCG. Their offspring each get one gamete from each parent, resulting in 16 possible combinations of genotypes and 9 possible phenotypes.

Figure 3.7
The 23 pairs of chromosomes of a human male. Maleness is determined by the one copy of the X chromosome and one copy of the Y chromosome that can be seen in the lower right corner of the figure. The chromosomes are shown following duplication, just before they would be split between the daughter cells. (Photo from the National Human Genome Research Institute, NIH, Source: Wikipedia)
Figure 3.8
The process of independent assortment of two genes, one that determines seed pod color and one that determines seed pod shape. There are 16 possible genotypes that give rise to 4 possible phenotypes. (Source: K. Scheiner)
Born Johann Mendel – taking the name Gregor upon entering the monastery – Mendel’s pioneering work in genetics became the basis of our modern understanding of the process of inheritance and was a critical precursor to the development of evolutionary theory in the 20th century (see Chapter 4). A modest man, he was unlike most modern scientists in not publicly advocating his ideas. Yet his theories of inheritance set the stage for the green revolution in agriculture in the 20th century and today’s revolution in genetic technologies.

Mendel was born July 20, 1822 to an ethnic German family in Heinzendorf bei Odrav, Austrian Silesia in the Austrian Empire (now called Hynčice in the Czech Republic). He grew up with his parents, Anton and Rosine Mendel, and two sisters – one older and one younger. They lived on a farm that had been in the Mendel family for 130 years, which might have sparked his interest in botany. Growing up he worked as a gardener and studied beekeeping, and attended the Philosophical Institute in Olomouc from 1840 to 1843. It was the recommendation of his physics teacher, Friedrich Franz that lead him to enter the Augustinian Abbey of St. Thomas in Brno in 1843, where he took the name Gregor.

His scientific career flourished in the monastery. While a monk he attended the University of Vienna from 1851 to 1853, where he studied zoology, botany, chemistry and physics, and began his experiments with pea plants. Inspired by his university professors and monastic colleges, he began researching variation in plants, cultivating and testing 29,000 plants between 1856 and 1863. With the encouragement of his abbot, Cyrill Franz Napp, he began his experiments in the monastery garden; Napp himself had a lively interest in the sciences and plant cultivation. These experiments resulted in the Law of Segregation and Law of Independent Assortment, later called Mendel’s Laws of Inheritance (Table 3.2).

Despite the enormous impact Mendel’s experiments would have, at the time he presented his work he was met with overwhelming ambivalence. He read his preliminary paper, Experiments on Plant Hybridization, at two meetings of the Natural History Society of Brünn in Moravia in 1865, publishing them the following year in the Natural History Society’s Proceedings. Over the following 35 years, Mendel’s work would be cited only 3 times. Contemporary writings say that audiences received Mendel’s lectures courteously, but with blank incomprehension. No one had previously attempted to use mathematical and statistical analyses to interpret the results of biological inquiry, and at the time, most biologists held with the idea of blending inheritance, such as Darwin’s attempts to explain inheritance via pangenesis (which were, ultimately, unsuccessful).

Mendel himself was a relatively shy person, and might not have presented his groundbreaking results with the necessary emphasis and stress. Although he sent out forty special reprints of his papers to various botanists and biologists known to be interested in the hybridization of plants, most seem to have ignored it or paid scant attention. One of those recipients was Charles Darwin, although it appears that Darwin never read the work as the pages of the copy in his library are uncut. If he had read them, Darwin might have been able to reconcile Mendel’s ideas with his own theory of evolution, something which did not happen for another 60 years. However, as that reconciliation required the work of several brilliant mathematicians, skills not evidenced by Darwin, it is likely that he would have failed to make the connections. It is impossible to know. The only recipient to correspond with Mendel was Carl Wilhelm von Nägeli, one
of the most highly acclaimed botanists of the time. It appears, however, that he only glanced at the work. Although it dealt with 355 cross-bred strains and 12,980 resultant hybrids, Nägeli described it as “incomplete” and urged Mendel to continue his investigations. Unfortunately, Nägeli advised Mendel to begin using hawkweed, a member of the sunflower family that reproduces asexually, a fact not appreciated at the time. This meant any results would be uninformative and contrary to the predictions of Mendel’s theories, since hawkweed’s genetic information is transferred exclusively through the maternal line. When the experiments failed, the disappointing results combined with the previous audience reactions led Mendel to become frustrated with his investigations.

Mendel tried expanding his scientific inquiries into the animal kingdom using honeybees, but was stymied there as well. It was difficult to control the mating behaviors of the queens, and thus impossible to generate a clear picture of their heredity. He also managed to create a hybrid strain of bees that was so vicious it had to be destroyed. When he was elevated to abbot in 1868, his scientific work largely ended. Although he regretted the loss of its pursuit, his time became consumed by his new administrative responsibilities, such as a dispute with the civil government over their attempt to impose taxes on religious institutions. The truncation of a scientific career in the morass of administrative duties stymies academic scientists even today. Despite the lack of appreciation for his work in his lifetime, in 1883, a matter of months before his death Mendel remarked “My scientific studies have afforded me great gratification, and I am convinced that it will not be long before the whole world acknowledges the results of my work.” In less than 20 years he was proven correct in what is now called the rediscovery of Mendel. At about 1900, several scientists working on problems of inheritance and (unknowingly) performing similar experiments, independently came upon his work, including Hugo de Vries, Carl Correns, Erich von Tschermak and William Bateson. In May of 1900, while on the train going to present a paper on heredity to the Royal Horticultural Society, Bateson read Mendel’s actual paper for the first time. He immediately incorporated Mendel’s laws into his lecture.

Until then the predominant view on patterns of heredity was that the traits of offspring were a blend of those of its parents, a view influenced by Darwin’s theories of slow, continuous evolutionary change. Those who held this view were termed biometricians. In the 1890s the biggest opposition to that school of thought came from Bateson, who pushed the idea of discontinuous variation, an idea that fit well with Mendel’s notions of discrete genes that could segregate independently. (Bateson also coined the term “genetics”.) The dichotomy between the established biometric school of thought and Mendel’s newly rediscovered laws raged in the first few decades of the 20th century, with biometricians claiming the weight of statistic and mathematical rigor, and Mendelians claiming a better and deeper understanding of biology. The conflict was finally resolved by R. A. Fisher (see Box 2D) who developed a mathematical model from which many discrete genes of small effect could result in continuous variation in a trait.

One controversy has dogged Mendel’s legacy, the claim that he fudged his data to make them better fit his theories. This issue was first raised in 1936 by Fisher who published a re-analysis of some of Mendel’s data. Because of the way that Mendel assessed the genotype of his pea plants, such as the experiment shown in Figure 3.6, chance events should have resulted in ratios that were slightly less than the expected ones). While Fisher thought that the data were falsified, he attributed the act to a supposed assistant, rather than Mendel himself. There the issue lay until 1965, the centenary of Mendel’s original publication. A series of papers that re-examined the issue resulted in a vigorous debate that continued for the next 40 years. We now conclude that this is probably not a case of scientific malpractice or fraud. Mendel’s results may have been
the result of confirmation bias; Mendel might have detected an approximate 3:1 ratio in a small sample size early on in his experiments, then collected more data until his results conformed to a more exact ratio. Some suggest that he may have censored his results. The seven traits he studied occur on separate chromosomes, which is extremely unlikely if they were chosen at random. It is also possible that the way he scored some of the traits was more accurate than that assumed by Fisher. Finally, additional analyses of some of the data examined by Fisher show that the results closely match the ratios that Fisher predicts. While Mendel carefully reported these somewhat deviant results, he still interpreted them as supporting his theories. Given the absence in Mendel’s day of the kinds of statistical analyses developed in the 20th century, such interpretations were justified. So, there is no evidence that Mendel was anything less than scrupulous and honest in his scientific work.
Continuous traits

While some characteristics of organisms come in discrete types, many characteristics show continuous variation - traits that are measured on a spectrum, such as height or weight. Consider a trait such as height in a plant at the end of the growing season. In an experiment, we would choose pairs of plants, pollinate one individual with pollen from the other in each pair, and cover the flowers to prevent any other plant from pollinating that individual. Thus, when we collect the seeds at the end of the growing season, we know exactly who the parents were. Before the parental plants die, we measure their heights. Then we plant the seeds, let them grow, and measure the heights of the new plants at the end of the next growing season.

We then plot the heights of the offspring against the heights of their parents (Figure 3.9). In this case, we find that taller parents tend to produce taller offspring. Because the trait being measured does not sort into discrete categories, we must measure this tendency using a statistical technique called correlation. By plotting the data on a graph comparing the parents’ and offsprings’ height, we can draw a line through the middle of those points; once the mathematic slope of the line has been calculated, we determine how closely the offsprings’ height matches that of their parents. If offspring always exactly matched their parents, the correlation between parental height and offspring height would be 1.0, and if there were no relationship, the correlation would be 0.0. In our example, the correlation is 0.41, and the slope of the line is 0.78; there is a resemblance, but some offspring are taller than their parents, while others are shorter. This correlation is one measure of heritability, the amount of resemblance among relatives that is due to shared genes. Offspring tend to resemble their parents and their siblings because the phenotype (physical characteristics) of an individual is determined in part by its genotype and an individual receives its genes from its parents and shares those genes with its siblings.

However, comparing offspring trait values against parental values is only one measure of the heritability of a trait. In the example above, because we used information from both parents, the slope is exactly equal to the heritability. If we had measured the trait in only one parent, we would have information about only half of the genes being contributed to the offspring, and the slope would be one-half the heritability.

The other common way of measuring heritability is to measure the correlation among siblings. Suppose we took two seeds from each of many plants. We could germinate the seeds and grow the pairs of siblings, measure their heights, and construct a graph much like Figure 3.8, except that now the axes would be the heights of the two
siblings, and each point would represent a sibling pair. Again, the slope would measure heritability, with the exact relationship depending on whether the individuals shared both parents or just one parent. We can do such an analysis with cousins or with any individuals that are related as long as we know their relationships. Nor are we restricted to using pairs of individuals. Various statistical techniques can be used to measure heritability in groups of related individuals with different degrees of relatedness.

During the first decades of the 20th century there were heated disagreements between geneticists that focused on discrete traits and those that focused on continuous traits. At the time it was not clear that continuous traits were controlled by the same sorts of genetic factors that controlled discrete traits. These differences related to arguments about whether evolution occurred through gradual change or by large, discrete changes. These disagreements were resolved by R. A. Fisher (see Box 2D), who showed that continuous variation and its heritability could be explained by assuming that a continuous trait was determined by very many genes, each of which had just a small effect on the phenotype.

There is a critical distinction between the heritability of a trait - how closely offspring resemble their parents - and whether that trait has a genetic basis. Heritability requires that phenotypic differences among individuals be due, at least in part, to the genetic information those individuals carry. In Box 3C we describe a case in which height is genetically determined. In that example, some individuals have a genotype of AA, some Aa, and some aa - height, therefore, is at least partially related to the information passed down from an individual’s parents. Instead, imagine that all individuals in the population have the same genotype, AA. Assume, however, that height also depends on the amount of nitrogen in the soil. If the population is growing in a field that varies in soil nitrogen from spot to spot, then individuals will differ in height. However, none of those phenotypic differences will be due to genotypic differences. Consider what would happen if we were to measure these plants, collect their seeds, and grow the offspring in that same field. By chance the seeds from a tall individual that grew in nitrogen-rich soil would end up in a variety of soils, some nitrogen rich and some nitrogen poor, and thus would have both tall and short offspring. The resulting correlation between parental height and offspring height would be 0 and the heritability of height in that population would be 0. Yet there is still a gene in that population that determines height. In this case, the heritability of height is zero because phenotypic variation in height is due to variation in an environmental factor, not to variation in the gene for height.

This example also demonstrates that the heritability of a trait depends on the frequencies of its alleles in the population. When the frequency of A is 1.0 – all individuals have the AA genotype – the heritability of the trait is 0. Thus, heritability estimates for the same trait can differ among populations, or in the same population measured at different times. Heritability estimates are always specific to the population and environment in which they are measured.

Heritability values tell us whether there is genetic variation for a trait in a population, and if so, whether there is just a little variation or a lot of variation. In terms of evolution, the amount of genetic variation may impose a constraint on evolution (see Chapter 4). If there is no genetic variation, the constraint is strong. No matter how much natural selection there is on a trait, there will be no genetic response; the potential for evolution requires variation in a trait before change can occur. If there is a little bit of genetic variation, the constraint is weak; there will be a genetic response, but it will be small, and evolution will proceed slowly. If there is a lot of genetic variation, there is almost no constraint on evolution.
Box 3C
A Simple Genetic System and the Resemblance of Relatives

This example shows how heritability as measured as a correlation between the phenotype of parents and offspring can be related to the effects of individual genes. Although this example is based on just a single gene, the same principles hold no matter how many genes affect a trait. Consider a simple genetic system in which plant height is determined by a single diploid gene. We assume that individuals with genotype AA are tall (100 cm) and those with genotype aa are short (20 cm).

**Case 1: Strict Additivity**

If individuals with genotype Aa have a phenotype that is exactly intermediate between AA and aa individuals, then genetic variation is strictly additive. In this case, Aa individuals would be intermediate in height (60 cm tall). Because the effects of the alleles are strictly additive, we can predict the phenotypes of the offspring of a cross. If both parents are tall, the cross will be AA × AA, and all offspring will be tall. If both parents are short, the cross will be aa × aa, and all offspring will be short. If one parent is tall and the other short (AA × aa), all offspring will be 60 cm tall (Aa). That is also the height that we get by averaging the parental phenotypes; the mean offspring phenotype equals the mean value of the parents’ phenotypes. If one parent is 100 cm tall and the other is 60 cm tall (AA × Aa), half the offspring will be 100 cm tall and half will be 60 cm tall. Again, the mean value of the parents’ phenotypes, 80 cm, exactly equals the mean value of the offspring phenotypes. Note that, for this cross, no parent or offspring is actually at the mean height; the mean is a descriptor of the group, not a property of any particular individual. A graph of mean parental phenotype against mean offspring phenotype (part A of the accompanying figure) has a slope of 1.0. That is, the heritability of this trait is 1.0, because we can perfectly predict the mean offspring phenotype from our knowledge of the parental phenotypes.

**Case 2: Dominance**

Now assume that A is dominant to a, such that Aa individuals are also 100 cm tall. In this case, predicting offspring phenotypes becomes more difficult. If both parents are short, all offspring will be short. But if both parents are tall, their genotypes could be both Aa, both AA, or one AA and the other Aa. In the latter two instances, all offspring will be 100 cm tall. Again, the mean value of the parents’ phenotypes, 80 cm, exactly equals the mean value of the offspring phenotypes. Now assume that A is dominant to a, such that Aa individuals are also 100 cm tall. In this case, predicting offspring phenotypes becomes more difficult. If both parents are short, all offspring will be short. But if both parents are tall, their genotypes could be both Aa, both AA, or one AA and the other Aa. In the latter two instances, all offspring will be 100 cm tall. But if both parents are Aa, then 1/4 of the offspring will be aa and will be short. The mean offspring phenotype will be 80 cm (3/4 x 100 + 1/4 x 20).
x 20), even though the mean phenotype of the parents was 100 cm. If we assume that the two alleles exist at equal frequencies in our population, then a graph of mean parental phenotype against mean offspring phenotype will have a slope of 0.67 (part B of the figure). The heritability of the trait is less than 1.0 because some of the genetic variation is nonadditive due to the dominance relationship. In other words, some offspring differ phenotypically from their parents because of the effects of dominance; if they do not inherit the dominant allele from either of their parents, they do not resemble their parents. Thus, the exact heritability of a trait in a population depends on both the degree of dominance and allele frequencies in the population.

**Fidelity of information transmittal**

The resemblance of parents and offspring is explained in the double-stranded nature of DNA and its replication process. The double strands accomplish two tasks at once, efficient replication and keeping the chromosome true to its template (Table 3.1, principle 2). During cell replication, each daughter cell must end up with all of the information contained in the parental cell. The DNA duplication process (Figure 3.3) results in two new chromosomes in a single operation. Those chromosomes are easy to separate; a fact which becomes important when it comes time to pass the new chromosome to the daughter cell. If the DNA molecule consisted of just a single strand, the two halves would tend to stick together through the same forces that hold the base pairs together. Also, a DNA molecule is very long and if unpaired, all of those “sticky” nucleic acids would tend to match up with bases elsewhere on the same strand, creating a tangled mess. Double-strandedness solves the templating problem while also keeping the information in a form that is easy to replicate and store.

Double-strandedness is also one way in which the information can be kept from changing. Occasionally, a chemical reaction will happen to the base pairs on one of the two strands changing its composition. When that occurs, specialized proteins will come along and repair the strand using the information from the complementary strand. Thus, the two strands act as “back up” copies of each other. We discuss these processes in more detail in the follow section.

**New Information**

**Capability for new information**

Living systems are the product of natural selection, which requires genotypic and phenotypic variation (see Chapter 4). If species could not evolve, they would go extinct; the constantly changing environment would eventually reach a point that survival and reproduction of that species would not longer be possible. Thus, the persistence of life requires the generation of variation in the form of new information (Table 3.1, principle 3). Of course, all species eventually go extinct; their evolution allows them to persist for far longer than they would otherwise.

The genetic system creates new information through a variety of processes which can be collected under two broad headings: mutation and recombination (Table 3.1, principles 4 and 5). **Mutation** is a change in the sequence of the DNA molecule. In the next section we explore the many ways that this can happen. **Recombination** is, as the name implies, the bringing together of DNA sequences in new combinations.

In general, mutations create new information thereby increasing the amount of genetic variation, although not all genetic variation results in phenotypic variation (as we will see later in this chapter). There are a few types of mutations that involve the loss of DNA and can even act to decrease variation. Recombination is different in that it is equally likely to increase as decrease variation. However, it produces new variation at a faster rate than mutation. Thus, both processes create new information in different ways.
and at different rates.

The rates at which mutation and recombination happen are determined by evolution and natural selection. The basic chemistry of DNA molecules is such that mutations constantly occur. But the rate at which those mutations get passed along to the next generation is controlled, in part, by the efficiency and accuracy of genetic repair mechanisms. Similarly, the rate of recombination is under the control of the organism through a variety of mechanisms that are described below. If the mutation and recombination rates are very low, then extinction eventually occurs. But if those rates are too high, offspring tend not to resemble their parents. Although the environment is always changing, it is not changing that quickly. For the most part, offspring that resemble their parents tend to do better at surviving and reproducing. Thus, natural selection will hone the mutation and recombination rates.

Natural selection has resulted in organisms that have both some background mutation rate, that is, a tendency for mutations to appear under normal circumstances, and in at least some species an increase in mutation rate under stressful conditions. If organisms never mutated, genetic variation would eventually disappear, evolution would cease, and life would likely go extinct as the environment changed and species did not. Only those forms of life that continue to mutate avoid extinction.

**Mutation**

The concept of mutation encompasses a wide variety of types of changes in DNA sequence. The most obvious is a simple change in the identity of one of the base pairs, for example if one of the strands changed from a sequence of ATTCCG to ATACCG, with a corresponding change in the matching strand. Some changes swap the order of base pairs, so that the previous sequence could change to ACCTTG, the middle four bases flipping position. Such flips can include stretches of DNA as long as tens of thousands of base pairs. Some changes involve deletions, for example changing ATTCCG to ATCG (again, the deleted piece can be very long). Some changes involve moving pieces of DNA, for example from one chromosome to another (Box 3D).

Some changes involve creating duplications; the previous sequence could be changed to ATTCCTTCCG. Such duplications could be just a few base pairs, consist of very long stretches of DNA, or even involve the duplication of whole chromosomes. In some cases the entire set of chromosomes is duplicated, so that an organism that had 8 chromosomes would now have 16, and so on. The duplication of whole genes and chromosomes can be an important source of new gene functions. Once there are two copies of a gene, one can continue to perform its original function while the other can evolve a new function. We describe how genes function later in this chapter.
While science can sometimes seem like an all-boys’ club, there have been many women who have made substantial contributions to the scientific edifice. Born on June 16, 1902, Barbara McClintock stands as one of the world’s most distinguished geneticists, discovering the phenomenon of the movement of genes from one chromosomal location to another, and demonstrating the roles of chromosome structures in the conservation of genetic information. She is a Nobel Laureate, having received the Nobel Prize for Physiology and Medicine in 1983.

Born the third of four children in Hartford, Connecticut, from the time she was three until she started school she lived with her aunt and uncle in Massachusetts, in order to alleviate the financial burden on her parents. Her father, Thomas Henry McClintock, was a physician, and it took him several years to successfully establish a practice so that he could comfortably support his family. When he was finally able to do so in 1908, the McClintocks moved to semi-rural Brooklyn, New York. Barbara was solitary, independent and a tomboy from early on; she had a good relationship with her father, but her relationship with her mother was plagued with strife. During high school she discovered her love of science and wanted to attend Cornell University, but her mother resisted the idea of higher education for her daughters, believing it would make them unmarriageable. Fortunately, her father intervened on her behalf, and she entered Cornell in 1919.

McClintock enrolled in the College of Agriculture where she studied botany, and received a BS degree in 1923. As an undergraduate, her interest in genetics was sparked when she took a course in the subject taught by C. B. Hutchinson, a plant breeder and geneticist. Hutchinson was impressed by McClintock’s keen interest in genetics, and invited her to participate in his graduate genetics course, even though she was still an undergraduate. McClintock would later point to that invitation as the reason she continued in genetics.

McClintock remained at Cornell as a graduate student from 1923 to 1927, and then as an instructor until 1933 McClintock. She was instrumental in assembling a group that studied maize (corn) and the then-new field of cytogenetics, which focused on the structure and function of chromosomes, bringing together plant breeders and cytologists. McClintock’s work focused on developing ways to visualize and characterize maize chromosomes, techniques which are still taught today. By studying chromosomes’ morphology, McClintock was able to link specific chromosomes to groups of traits that are inherited together, triggering a surge of interest in maize cytogenetics.

In 1930 she became the first person to describe the interaction of chromosomes during meiosis (Figure 3.4), and the next year she and a graduate student, Harriet Creighton, proved a link between the process during meiosis which results in the physical reassortment of chromosomal pairs and the recombination of genetic traits. She observed that when chromosomes were recombined, the resulting phenotype would result in the inheritance of new traits; until then, it was only hypothesized that genetic recombination was possible during meiosis, although there was genetic evidence that it occurred.

During the summers of those years, she worked with geneticist Lewis Stadler at the University of Missouri - Columbia, who introduced her to the use of X-rays to gener-
ate mutations. Using mutagenized maize, McClintock discovered ring chromosomes, which form when the ends of a single chromosome fuse together after radiation damage. From this, she hypothesized the existence of a structure on the chromosome tip that normally ensures the stability of the genetic material.

Unfortunately, McClintock was unhappy at the University of Missouri because of how she was treated by the other faculty, for example being excluded from faculty meetings. Early in 1941 she was invited by the Director of the Department of Genetics at Cold Spring Harbor to spend the summer there, and used the opportunity to take a leave of absence from Missouri. In December of the following year, she was offered a research position at the Carnegie Institute of Washington’s Department of Genetics Cold Spring Harbor Laboratory, where she continued her work with the breakage-fusion-bridge cycle, using it as a substitute for X-rays as a tool for mapping new genes.

By this time she was receiving wide recognition for her groundbreaking work. In 1939, at the age of 37, she was elected vice-president of the Genetics Society of America, and in 1945 she became its first woman president. In 1944 she became the third woman in history to be elected to the National Academy of Sciences. By 1951 she had received the Achievement Award of the Association of University Women, and been awarded two honorary degrees. She was undoubtedly one of the best-respected cytogenetics at that time; however, in that same year this respect was to be tested with her theory about the nature of the gene.

In the summer of 1944 at Cold Spring Harbor McClintock began a series of systematic studies on the mechanisms of mosaic color patterns of maize seed and the unstable inheritance of mosaicism. She identified two new dominant and interacting genes she named Dissociator (Ds) and Activator (Ac), finding that Ds did not only cause the chromosome to break, but had a variety of effects on neighboring genes when Ac was also present. By 1948 McClintock had discovered that Ds and Ac could move on the chromosome; the effects of this transposition could be observed via the changing patterns of coloration in maize kernels over generations of controlled crosses. McClintock described the relationship between the two loci through microscopic analysis, concluding that Ac controls transposition of Ds, and the movement of Ds is accompanied by breakage of the chromosome. The transposition of Ds in different cells is random, moving in some but not others, thus producing color mosaicism. The size of the colored spot on the seeds is determined by the stage of seed development reached during dissociation, and the transposition of Ds is determined by the number of Ac copies in the cell.

Over the next two years McClintock developed a theory by which these mobile elements regulated genes by inhibiting and modulating action. She referred to Ds and Ac as “controlling units” (later “controlling elements”) to distinguish them from genes. Hypothesizing that gene regulation explained how complex multicellular organisms made of cells with identical genomes have cells with different functions, this theory challenged
the concept of the genome as a static set of instructions passed between generations. It would not be until a few years later, after the structure of DNA was discovered (see Box 3A) that the more complex nature of genes began to be revealed.

McClintock reported her findings in a 1950 paper, and the next year at Cold Spring Harbor’s annual genetics symposium. Her paper addressed a critical missing piece in the understanding of genes, how their expression was regulated. Unfortunately for McClintock, she conflated the action of Ac as a controller of transposition with its potential role as a controller of gene expression. The combination of a conceptually difficult idea that challenged accepted wisdom along with an extremely long and dense verbal (and later written) presentation, meant that the reaction of her peers was mostly to ignore her. While McClintock was greatly admired as an experimentalist, her theory seemed to go to far beyond her data. It did not help that she aligned herself with an even more extreme iconoclast, Richard Goldschmidt, who gave the keynote address at that same symposium.

Ironically, McClintock is now primarily remembered for what to her was a side issue, the transposition of genes. Her major point was that one had to considered the context of a gene to understand how it was regulated. While the details of her theory were wrong, she was correct that more than the expressed part was critical. However, because she presented these ideas in review articles and non-peer-reviewed venues and because it went against the current theory it was not appreciated at the time. In hindsight, however, we can see that she was a pioneer in understanding gene regulation. Even in the 1960s, when the first models of gene regulation were worked out in bacteria, McClintock was not credited with the fundamental concept of regulation of expression in one part of the chromosome by elements in other parts (because the mechanism of that regulation was so different from the one she had earlier described). McClintock took a philosophical view on it, however, writing in 1973 on her decision to cease publishing by 1953 that “One must await the right time for conceptual change.”

McClintock officially retired from her position at the Carnegie Institution in 1967, but continued working with graduate students and colleagues in Cold Spring Harbor Laboratory as scientist emerita. Although she published little after the early 1950s, her previous achievements continued to gain her honors. She was awarded the National Medal of Science by Richard Nixon in 1971, and Cold Spring Harbor named a building after her in 1973. In 1981 she was the one of the first recipients of a MacArthur Foundation Grant (called “genius grants”) and awarded the Albert Laster Award for Basic Medical Research. She received the Wolf Prize in Medicine and the Thomas Hunt Morgan medal by the Genetics Society of America the same year. In 1982 she was awarded the Louisa Gross Horwitz Prize from Columbia University. All of this was capped in the following year when she received the Nobel Prize for discovering mobile genetic elements.

She remained a regular presence in Cold Spring Harbor community, giving talks on mobile genetic elements and the history of genetics research for the benefit of younger scientists. She died in Huntington, New York on September 2nd, 1992 at the age of 90. Today McClintock is widely held up as a role model for girls in children’s literature. On May 4, 2005 the United States Postal Service issued their “American Scientists” commemorative postage stamp series, which featured McClintock.

It is common to think of mutations as having a dramatic effect on the phenotype (Figure 3.10A and 3.10B). But mutations can also have very subtle effects (Figure 3.10C), so small that they are only apparent in very close contrast to the species as a whole. Some mutations may even have no effect at all on the phenotype. Other mutations, such as deletions, may destroy the function of a gene, which can result in extreme or even fatal effects on the organism. Duplications of whole genes may have no
immediate effect on the phenotype, but may allow the eventual evolution of a new function in one of the two copies. Further, in order to produce a permanent change in a population rather than a single individual, a mutation must occur in gametes or gamete-producing cells.

Figure 3.10
Mutations can have a variety of effects on organisms from large and striking to small and subtle. (A) Typically *Drosophila melanogaster* (fruit flies) have red eyes and yellow bodies (lower left in photo), but various mutations that affect pigment formation results in flies with a variety of eye and body colors. This is an example of a qualitative difference. (Source: Wikipedia) (B) A single mutation also can have a large effect on a quantitative (continuous) trait. The mouse on the left has a mutation that prevents the production of a hormone that controls the amount of fat in its body. (Source: Wikipedia) (C) The plant *Arabidopsis thaliana* typically has small oval leaves. Mutations in some genes can create small changes in size or shape, while others create wildly misshapen leaves. (Photo courtesy of Hirokazu Tsukaya)
Mutations are relatively rare events. Mutations of single base pairs occur at a frequency of $10^{-8}$ to $10^{-10}$ per base pair per generation. For the average-sized gene, mutations occur at a frequency of $10^{-5}$ to $10^{-7}$ per gene per generation. On average, in a given population for most species, one individual in ten has a new mutation somewhere in its genome each generation. New technologies that allow the measurement of variation in DNA are quickly adding to our understanding of mutation rates. At any one time there are likely to be many variants of a given gene even though mutations of a per gene or per individual basis are rare, since most species consist of millions of individuals. It is the accumulation of a lot of small mutations that provides much of the variation upon which evolution depends. Usually mutations with large effects on the phenotype are harmful, while mutations with very small effects can be either harmful or beneficial. Despite nearly 90 years of trying to determine the frequency of small and large, or harmful and beneficial mutations, this is still an intense area of research.

Mutation is never purposeful. Mutations do not occur because they can make individuals or species better adapted to their environment. Rather, both beneficial and harmful mutations occur all the time and are then winnowed out by natural selection. We often observe mutations only after many of the harmful mutations have been eliminated because the individual with that mutation dies or produces few offspring. Thus, it can seem as if all mutations are beneficial. During the 1980s some scientists claimed to demonstrate that Bacteria placed in conditions where they required specific mutations to survive—the ability to use unusual food sources—purposely mutated to better survive under those conditions. Others showed, though, that the results could be explained entirely because only bacteria with the correct mutations were ever seen.

Organisms under stressful conditions have an increased rate of mutation, and such an increased mutation rate can increase the chances that beneficial mutations will appear. But there is a distinct difference between a mechanism that simply increases mutation rate and one that directs those mutations to be best suited for the individual’s environment.

Although mutation is not purposeful or directed, it does not mean that all mutations are equally likely. Random does not mean that something occurs with equal frequencies (see Chapter 1). This non-equality of the occurrence of mutations happens in several ways. For example, because of their chemical similarity (Figure 3.1) it is much more likely that an A will change to a G and vice versa, than either will change to a C or T. Similarly swaps between C and T are more likely. Nor is the likelihood of mutation the same at all base pairs. There are mutational “hotspots” on chromosomes where base pair changes can be 10 times more likely than at other places. Similarly, some types or locations of deletions or duplications are more likely than others.

![Figure 3.11](image)
Frog with an extra hind leg due to a mutation caused by pollution.

Certain types of phenotypic changes are more likely to happen. For the most part, the development of an organism from fertilization to fully formed individual is highly regulated (see Chapter 6) so that small changes in that process are more likely than large changes and some large changes are more likely than others. For example, terrestrial vertebrate animals have four limbs. Mutations that change the length of those limbs, making them longer or shorter, happen frequently. Just consider the relatively long legs and arms of a
professional basketball player as compared to that of an average person. Eliminating limbs entirely is also possible (e.g., the hind limbs of whales, see Figure 4.1). But, there are no six-limbed terrestrial vertebrates, indicating that such a change in development is either extremely difficult or impossible. Although there have been instances of vertebrates born with extra limbs due to mutations (Figure 3.11), these limbs are often misshapen and not functional, showing that the developmental limitations on the kinds of mutations that are possible can provide important directionality to the course of evolution.

**Exchange and recombination**

**Recombination** is the process by which the genetic material of one individual is mixed with that of another. In Bacteria and Archaea this process occurs in a haphazard fashion. In some species, when a cell dies and ruptures, its DNA can get taken up by other cells. Most of the time, the other cell simply uses the individual nucleotides to build new copies of its own DNA molecules. Occasionally, however, a stretch of DNA from one individual can get incorporated into the chromosome of another. Sometimes this exchange occurs when a virus picks up a piece of its host’s DNA during replication (Box 3A). This sort of exchange of genetic material can occur among very distantly related species, including Eukaryotes, Bacteria and Archaea. How often this occurs among Bacteria and Archaea is not known and currently a hotly debated issue. This type of genetic exchange can act like an extreme form of mutation, bringing genes with entirely new functions to a species. However, this sort of movement is limited to cellular functions (see Chapter 5). A cat is not going to suddenly grow a dog’s tail. In some cases, the exchange process among Bacteria is somewhat less haphazard. Some species have specialized structures and special enzymes for transferring DNA from one individual to another (Figure 3.12). Such transfers, however, still occur only occasionally.

The process of DNA exchange and recombination became regularized with the evolution of Eukaryotes. Recombination occurs during the process of meiosis during which a diploid individual produces haploid gametes (Figure 3.13). Even those Eukaryotes that are haploid for nearly their entire life will mate and produce diploid cells that have at least a brief existence. Recombination occurs during the process of meiosis during which a diploid individual produces haploid gametes (Figure 3.13).

The process of meiosis creates new variation in several ways. For example, if an entire population consists of just red-flowered (CRCR) and white-flowered (CW CW) individuals as in Figure 3.5, the process of gamete formation and mating creates new pink-flowered (CRCW) individuals.

Recombination also allows genes linked in some way to separate, creating greater variation. Consider Mendel’s Law of Independent Assortment (Figure 3.7). Since independent assortment occurs only if genes are located on different chromosomes, if two genes are on the same chromosome, they will very likely end up in the same gamete. The process of recombination unlinks those genes by breaking the DNA chain of each chromosome and creating new links with the other chain. For example, if there were only genes for seed shape and seed color (Figure 3.8) and there was no recombination, then the only types of gametes that would exist would have either the SSCY combination of genes or the SWCG combination. The only types of offspring would be either SSSC CYC, SWSWCYG or SSSWCYCG. Tracing patterns of recombination has even helped scientists uncover the structure and function of genetic material itself. Only if there was recombination could there be combinations such as SSSWCYCY or SWSWCYCG. One of the confirmations that genes were carried on chromosomes came from the work of Thomas Hunt Morgan while studying patterns of recombination in fruit flies (Box 3E).
From the perspective of evolution, recombination is important because it creates new combinations of genes which can increase the rate of evolution. Consider two different mutations (A and B) that each increase the fitness of an organism. An individual has mutation A. Since mutations are rare, it is highly unlikely that its offspring would also have mutation B. But, if that individual mated with an individual that carried mutation B, then many of their offspring would carry both mutations and be more fit. Hybridization, matings between individuals of different species, is an extreme example of this process of bringing together very different genes to form new combinations.

Figure 3.13
The process of meiosis. After chromosomes are duplicated, the matching pairs line up. At that point, pieces of the chromosomes can be swapped creating new combinations of genetic materials among the various gametes. (From the Science Primer, NCBI, NIH, Source: Wikipedia)
Box 3E
Thomas Hunt Morgan

Thomas Hunt Morgan helped create the modern science of genetics, both because of the ideas that he fostered and because of the methods that he developed to test those ideas. Yet, he was initially skeptical of the entire field and later largely abandoned it to return to his first love, the development of marine organisms.

Morgan was born on September 25, 1866, in Lexington, Kentucky, the eldest son of an old southern aristocratic family. Thomas could trace his family deep into the history of the region; he was the nephew of Confederate General John Hunt Morgan, and the great-grandson of the first millionaire west of the Allegheny Mountains, John Wesley Hunt.

As with many biologists, Morgan showed an interest in the study of life before the age of 10, collecting birds and their eggs as well as fossils. He continued to pursue these interests at a time in America when a career in the sciences was unusual, receiving a B.S. degree in biology in 1886 from the University of Kentucky, the only person in his class to do so. That summer he visited the seashore laboratory of Alpheus Hyatt at Annisquam, Massachusetts, starting him on his life-long interest in the biology of marine organisms. When he began his graduate studies at Johns Hopkins University, he brought with him a newfound interest, beginning a project on the development of sea spiders. Much of this work was done at the Marine Biological Laboratory at Woods Hole, Massachusetts, a place that inspired many biologists (see Boxes 5B and 7D). After receiving his Ph.D. in 1890, he was hired as a professor of zoology at Bryn Mawr College, Johns Hopkins’ sister school (the schools were segregated by gender, like nearly all colleges and universities of the time). He stayed until 1904. For the next 24 years, he took up a position as professor of Experimental Zoology at Columbia University, then became a professor of biology at the California Institute of Technology, Pasadena, where he was also the director of the William G. Kerckhoff Laboratory.

During the year he received his Ph.D. he received a fellowship to visit Europe, spending a great deal of time in Naples, Italy. There he met Hans Driesch, a German biologist noted for his work in embryology, with whom he later collaborated. Perhaps as a result of this association, he turned to experimental embryology after returning from his trip. Although he would later become deeply involved with genetic research using the fruit fly, *Drosophila*, throughout his life he worked on the problem of developmental stability using the sea squirt, *Ciona*, as a model organism. Sea squirts are closely related to the group that gave rise to vertebrates and, thus, much studied as a glimpse into the origins of vertebrate traits. He was associated with Woods Hole continuously from 1902 onward, actively taking part in biological expeditions to the Bahamas and Jamaica.

As a student Morgan was boldly critical, skeptical and of a very independent judgment. His work in experimental embryology and regeneration won him a high reputation during his early career, becoming the president of the American Morphological Society in 1900. He was an intent worker, impatient of unnecessary interruptions, and a prolific and sometimes hasty writer. Although he did no genetics work prior to 1905, his early papers show considerable distrust of Mendelian laws. Even up to 1909, after working himself with mice and others in his department began working on *Drosophila*, Morgan
did not clearly distinguish problems of heredity from those of development. However, it was the work he did with *Drosophila* that eventually brought him around.

Morgan’s groundbreaking work was built upon a long line of scientists interested in genetics. *Drosophila* were first bred in quantity by Charles W. Woodworth, who studied them at Harvard during the winter of 1900-1901. There he suggested to William E. Castle that the flies would be useful for genetic study; Castle and his students used it for studies on the effect of inbreeding. From their activities, Frank E. Lutz became interested and suggested them to Morgan as an experimental animal.

Throughout his career, Morgan always insisted that his work was that of a team. He assembled a group headed by Alfred H. Sturtevant, Hermann J. Muller and Calvin B. Bridges who carried out experiments in a much larger scale than anyone had attempted before. His laboratory at Columbia University became known as the Fly Room and was the hub of some of the most important work in genetics in the first part of the 20th century. Morgan’s early *Drosophila* papers focused attention on the demonstration that the gene for white eyes was associated with those that determined gender. From these early studies, Morgan’s skepticism about Mendelian genetic laws quickly faded. In a 1911 letter in *Science*, he put forward the theory that genes were linked to each other on chromosomes in a linear arrangement. He also discovered that males and females had differences in the chromosomes that determined sex, with females having two large chromosomes, dubbed X and so being XX, while males had one large and one small chromosome, dubbed Y and so being XY. In 1915 Morgan, Sturtevant, Bridges and Muller wrote *The Mechanism of Mendelian Heredity*, a text that would stand as a seminal work of genetics, and would form the basis of geneticists’ efforts for decades to come.

Because of Morgan’s success with *Drosophila*, it was picked up for genetic studies all over the country and the world, with Columbia at the center of an exchange network of promising mutant strains. When he established the biology division at the California Institute of Technology, he wanted to distinguish it by focusing on genetics and evolution, experimental embryology, physiology, biophysics and biochemistry; he succeeded not only in advancing and expanding biology at CIT, but in spreading his impact throughout the entire field of biology via his efforts with *Drosophila*.

In keeping with the personality that lead to sometimes hasty research publication, Morgan did not keep organized notes on experiments, but would pull envelopes and scrap paper out of his pockets when examining ongoing and completed experiments. His colleague Tyler kept an eye on Morgan’s experiments, and probably helped to keep the chaos at bay. Morgan, for all his apparent absent-mindedness, was extremely passionate about experimentation and distasteful of speculation, believing only what could be proven. He was known for his sardonic wit; in 1909 during his speech at the American Breeders Association, he was critical of genetics, saying, “In the modern interpretation of Mendelism, facts are being transformed into factors at a rapid rate. If one factor will not explain the facts, then two are involved; if two prove insufficient, three will sometimes work out....” As fate would have it, a year later he discovered the gene that controls white eye pigmentation in *Drosophila*. This is an excellent example of the fact that real scientists change course when confronted by facts.

In 1904 Morgan married Lilian Vaughn Sampson, a former research student of his at Bryn Mawr and a frequent associate in his lab work. They eventually had a son and three daughters. While at Cal Tech, he lived in a comfortable ranch house on the north side of Kerckhoff Labs, and would give a General Biology Seminar at 7:30 PM on Tuesdays. He would open the seminar by commenting on stories in the New York Times with a scientific bent, making fun of human gullibility. He would introduce the speaker for the night, then sit in the front row next to his wife, and usually fell asleep within two
sentences. He wife would nudge him awake, and he would be refreshed and usually had acute questions for the speaker at the end.

Morgan died on December 4, 1945, leaving behind a legacy writ large in genes. He was highly decorated during his lifetime, becoming a foreign member of the British Royal Society in 1919, receiving the Darwin Medal in 1924, the Copley Medal in 1939, and finally the Nobel Prize in 1933. When he established the biology division at Cal Tech, he wanted to distinguish it by focusing on genetics and evolution, experimental embryology, physiology, biophysics and biochemistry; he succeeded not only in advancing and expanding biology at Cal Tech, but in spreading his impact throughout the entire field of biology via his efforts with Drosophila. Morgan’s work laid the foundation of the science of genetics and the theoretical foundation for the mechanism of evolution via natural selection, fields that are still seeing advances today.
Random processes

Mutation and recombination are an important source of contingent effects in living systems (Table 3.1, principle 6). As discussed previously, mutations occur randomly; so, while they are an important part of new variation for the evolutionary process, when and where they occur is unpredictable. Cougars are found in North America and tigers in Asia because the mutations that led to each of these species appeared on one continent and not the other.

The random nature of mutation has direct implications for human health. For example, humans and the influenza virus are in a continual arms race. The human immune system produces molecules that are able to attack the virus when it enters your body (see Chapter 6). Your body recognizes that a particle is a flu virus because of the proteins in the virus’s outer envelop, and your body has the ability to continually produce new molecules to match those proteins. Once a particular molecule is produced your body continues to produce them for several years and they remain circulating in your blood. This is important because those molecules can attack an entering virus particle immediately; you only get sick when a virus enters your body that is not recognized by your immune system. In that case, the virus has time to multiply and attack your body for several days before your body can mount a counterattack. Vaccination keeps you from getting sick by prompting your body to make the antivirus molecules before the virus enters your body.

It takes about six months for flu vaccine production to go from the start to the point of having the tens of millions of doses necessary to protect the U.S. population. So, it is important to know well ahead of time which influenza strains will be circulating in the population. The problem is that the influenza virus is continually mutating, changing the composition of the proteins in its outer envelop. Those mutations are unpredictable, to some extent. Most years the mutations create small changes in the proteins and so the chance of the virus making you very sick is small. Based on past patterns of mutation, those small changes are somewhat predictable. Scientists can use mathematical models to decide which form of the vaccine should be produced – even if the vaccine is not an exact match, being close to a match is enough to give your body the head start it needs to mount its defense. In some years, though, the mutations create much larger changes. While scientists know that these mutations with larger effects have a small chance of happening each year, exactly when they will occur and what the nature of the mutation will be is much harder to predict. In those years the vaccine will not match the strains and many people will end up sick with the flu.

The independent assortment of genes on different chromosomes is the best example of truly random processes in biology. To see why this must be so, consider the following thought experiment. In some species, not all of the products of meiosis go on to produce gametes. For example, in human females only one of the four gametes goes on to become an egg. Imagine that there was a gene that had two alleles and one was able to increase its chances of ending up in that egg. Even a very small advantage in getting passed along to the next generation would mean that this allele would quickly become the only type in the population. Alleles of any other genes that are linked to that allele would also be quickly fixed, and alleles linked with the unfavored allele would quickly become much rarer in a population. In fact, the entire meiotic machinery is geared to suppressing the possibility of such skewing of gene segregation during meiosis. The end result is that unlinked genes end up in gametes in the proportions expected by random chance.
Information Usage

In the previous sections we described how the genetic system accounts for resemblances among relatives in Eukaryotes, especially complex multicellular organisms. Although the link between genotype and phenotype is much simpler for single-celled organisms, the problem of the resemblance of relatives was first considered and solved for those more complex organisms. This is another example of the advancement of technology leading to further biological discovery, albeit in a somewhat ironically backwards fashion. While the greater number and duplication of chromosomes in Eukaryotes results in a much more complex information system, Bacteria and Archaea are much simpler to study, because much of the phenotype involves the direct expression of the information stored in the DNA; there is only a single chromosome, and thus one copy of each gene.

Transcription and translation

Transcription and translation is the two-stage process (Figure 3.14) by which the information encoded in the DNA molecule is used. First, the information contained in the DNA molecule is copied to a RNA molecule (transcription). One of the two DNA strands is used as a template and through a process similar to DNA replication, a complementary strand of RNA is produced. The RNA strand consists of A, G and C bases just like DNA. The one difference is that thymine (T) is replaced with a closely related base, uracil (U).

![Figure 3.14](https://example.com/diagram.png)

The process of transcription copies the information contained in the DNA molecule to an RNA molecule. That information is then used to build a protein in the process of translation using other types of RNA molecules. (Created by Madeleine Price Ball, Source: Wikipedia)

Next, the information in that RNA molecule is used to create a protein (translation). The correspondence of the sequences of bases in the RNA molecule and the amino acids in the protein molecule is called the genetic code. Thus, the genetic code is how the information in the genotype is translated into the phenotype.

The genetic code is built on triplets of bases called codons. That is, a set of three bases corresponds to one amino acid. With four different bases (A, G, C, and U), there are 64 different possible three base combinations. However, because there are only 20 different amino acids, nearly all amino acids are coded for by more than one combination of bases (Figure 3.15). For example, tyrosine is coded for by two sequences: UAU and UAC. In this case, the first two bases are the same and only the third base differs. Thus, the genetic code is redundant as a given amino acid is coded for by more than one possible combination of bases. The use of only 4 bases to code for 20 amino acids by triplet codons is an example of an emergent property because the grouping of the individual bases into codons creates new information in the same way that words are combinations of letters.
Figure 3.15
The genetic code. Each triplet codes for either one of 20 amino acids, or indicates the starting or stopping point for protein synthesis. Slight variants in the code are found in some organisms.

The redundancy of the genetic code is part of the robustness of the information system (Table 3.1, principle 7). Earlier we mentioned that mutations between A and G and between C and T are much more likely than the other possible types. When an amino acid is coded for by just two different codons, the two codons have in their third position bases that are either A and G or C and T. So, the most likely mutations in third-position bases will not change the resulting protein and the phenotype of the organism. For example, consider arginine. A mutation from AGA to AGG still codes for arginine. Because of this lack of change, we speak of silent mutations. Whether different triplets that code for the same amino acid are truly equivalent and silent is an active area of research.

Diploidy further increases the error tolerance of the system. A mutation that simply changes one base into another will likely have a very small change in the properties of the resulting protein. But many types of mutation can completely destroy the functioning of the protein, such as deletion, changes in the order of the sequence, or a change in the base sequence that causes translation to halt before the entire protein is
produced. However, a diploid organism has two copies of each gene, so even if one copy does not produce a functioning protein, the other will. Some forms of dominance will result in the dominant form of the gene producing the functioning protein, thus preserving its function in the phenotype.

The phenotype of an organism is not just determined by whether a protein is produced or not, or by the form of the protein; it is also determined by the relative amount of protein produced. This can be regulated at many different steps in the process: the rate of transcription or translation can be sped up, slowed down, or shut off entirely. The rate of transcription is controlled by the DNA sequence adjacent to the region that is transcribed. These parts of the sequence are referred to as noncoding, regulatory regions. In translation, typically many copies of a protein will be made from a single RNA molecule, so the amount of protein can also be regulated by how long that RNA molecule lasts before it is broken down. Cells contain a complex machinery for this regulation (see Chapter 5).

In humans, only about 2% of the information in the DNA ends up as proteins. Typically in Eukaryotes much of the DNA that is transcribed into RNA is not translated. Some of this RNA is snipped out of the strand before translation. Other stands of RNA are part of the translation machinery, or combine with proteins to regulate their function. Yet other RNA strands directly interact with DNA or the proteins that provide structure to the chromosomes and help regulate the rate of transcription. Until recently, scientists generally considered all of these untranslated portions of the genome to be “junk.” The more we learn about the regulation of information usage, the more we have come to understand that even parts of the DNA not directly coding for proteins help determine an organism’s phenotype.

The relationship between the genotype and phenotype is far more complex than it might first appear. Mendel’s original experiments generally involved characteristics that were controlled by one or maybe two genes, making it easy to discover their effects on a particular aspect of the phenotype. However, as the biochemistry of the cell was worked out in the early decades of the 20th century, it was realized that genes coded for enzymes, which in turn were responsible for an organism’s characteristics. In the 1940s, the American scientists George Beadle and Edward Tatum put forward the “one gene-one enzyme” model, for which they were later awarded the Nobel Prize. Genes were no longer seen as having a direct effect on the outward appearance of the organism, but a direct role in the production of the cellular machinery that eventually was responsible for the structure and function of the organism (see Chapters 5 and 6).

**Direction of information flow**

In the process of transcription and translation, information flows in one direction - from DNA to RNA to proteins. The idea that information flow is only in one direction was termed the central dogma by Francis Crick (Box 3A). For several decades it was thought that this unidirectional flow was universal. We know now of several important exceptions.

The first exception is a flow of information from RNA to DNA. The information in viruses can be stored and transferred as either DNA or RNA. When some RNA viruses enter a cell, they will get translated back into DNA. In that form they can be incorporated in the chromosome of the host cell. If that cell is one that produces gametes, the virus can even be passed from parent to offspring. The most infamous of these viruses is the agent that causes AIDS, the human immunodeficiency virus (HIV). The incorporation of HIV into human chromosomes is why it is so difficult to cure an infected person and eradicate the disease. To eliminate the virus, a person’s own cells must be killed. After initial infection, the virus can lay dormant in the chromosome for several years,
multiplying as the cells divide, before it eventually emerges and makes the infected person sick.

The second exception involves changes in the chemical structure of the DNA molecule (other than base pair sequence changes). Cytosine (in Eukaryotes) and adenine (in Bacteria) will sometimes have a hydrogen atom replaced with a CH3 group. Because CH4 is methane, this process is called **methyla**tion. Methylation affects the rate at which genes get transcribed, thus affecting how information is used by the organism, which can change its phenotype. When the cell divides, the methylation changes get passed along to the daughter cells, preserving the information changes. If the methylation occurs in cells that produce gametes, those changes in the DNA chemistry can get passed along to the offspring, sometimes for several generations. Because methylation is controlled by enzymes, this is an instance of information going from proteins to DNA. A heritable change in the information content of a cell is occurring, even if this is not a change in the DNA sequence. It is important to note, however, that such changes are not permanent and likely to be reversed.

Such changes sound like an old idea. It was once believed that organisms could pass along changes that occurred in their phenotype to their offspring, a process termed the inheritance of acquired characteristics. However, this idea was discarded by the middle of the 20th century as the genetic basis of inheritance was worked out. The inheritance of acquired characteristics posited that changes in the traits of an organism through use or disuse (such as bulking up a muscle) would get passed on to its offspring. This is quite different from the idea that a virus can be incorporated into a chromosome and be inherited, or that DNA methylation patterns might be passed on. However, there are many ways in which the environment of an individual, especially as determined by its parents, can affect its phenotype, which is the topic of the next section.

**Context dependency**

How the information contained in a stretch of DNA gets used depends on its context (Table 3.1, principle 8). That context includes the other DNA sequences surrounding it on the chromosome, the sequence of the DNA in the matching gene on the paired chromosome, other genes, and most notably, the rest of the environment outside the cell. That environment can include other cells within the same organism, other organisms, and the individual’s physical surroundings. Thus, the phenotype of an organism is the result not just of that individual’s genes, but also the environment within which those genes reside and are expressed.

Previously we discussed how flower color might depend on the combined effects of the two paired genes, either as an average of the two (Figure 3.5) or with one gene being completely dominant to the other (Box 3B). Other patterns are also possible. Dominance may be incomplete - for example, the color of the flower might be mostly red (as opposed to a strictly additive genotype in which the color of the offspring is the average of its parents). Sometimes the phenotype of the offspring falls outside the range of the parents entirely. In other cases, both genes may be expressed, such as in the blood type of humans. Humans have blood types O, A, B and AB, a condition that is due to three alleles. The A and B alleles each codes for an enzyme that attaches a different sugar molecule to a protein on the surface of red blood cells, while the O allele produces an enzyme that lacks activity. Because of this, an individual that has either the AA or AO genotype has blood type A. Individuals with blood type O all have an OO genotype, and individuals with an AB genotype have blood type AB.
The products of different genes can interact with each other. The hair color of wild mice is a grayish color (described as agouti) due to bands on individual hairs (Figure 3.16). One of the genes that controls a mouse’s hair color comes in two forms, the dominant form of which is agouti in color. The recessive form lacks the banding pattern, so a mouse with both forms of that gene is black. Another gene also affects hair color, and individuals with both recessive forms have white hair. The effects of the gene causing the white color gene completely override the effects of the agouti gene. Such interacting genes may be on different chromosomes, or even in different parts of the cell. In Figure 3.16, Wild color and white mice (Source: Wikipedia) Eukaryotes, for example, specialized structures called mitochondria and chloroplasts (see Chapter 5) also contain chromosomes and the enzymes coded for by the genes on those chromosomes can interact with the enzymes coded for by the genes in the nucleus.

Some genes code for enzymes that are responsible for regulating the transcription processes; these enzymes can bind to the DNA and prevent the transcription enzymes from doing so. The binding sites are adjacent to the parts of the DNA sequence that are transcribed, thus defining the limits of the information being read, similar to punctuation in a sentence. Changes in the sequence to those stretches of DNA can change the binding ability of both the transcription enzymes and the regulatory enzymes. Changes in DNA sequences both adjacent to a transcribed gene and at other genes affects the rate of transcription and the usage of the information - similar to replacing a period with a semicolon, or vice versa. Transcription rates may be determined by chemical signals that come from other cells in the organism (see Chapter 6).

The phenotype of an organism depends not just on all of its genes, but also the environment. One aspect of that environment is the parents of the organism. Previously we described how changes in methylation patterns on chromosomes can be passed to offspring. Much more common, however, are direct nutritional effects. The size of an egg or a seed will also affect the size of the organism at birth or germination and many types of animals provide food for their offspring after birth. Those behaviors are controlled, in part, by the genes of the parents, so you can think of such parent-offspring effects as interactions between the parents’ genes and the offspring’s genes.

How a gene is expressed can depend on where it is within an organism. All of the cells in your body have the same genes, and the differentiation of those cells into specialized function (e.g., skin, muscles, nerves) is due to differences in gene expression (see Chapters 5 and 6), which can be determined by the external environment. Siamese cats have light color hair on most of their body, but dark colored hair on their extremities because the temperature is cooler. Kittens are all white because the mother’s womb is warm, and their extremities darken with age. You can affect this darkening process by raising the cats at warmer or cooler temperatures.

The disease sickle cell anemia provides an example of two types of context dependency, the identity of the paired gene and the environment. Anemia is a condition of having a deficiency of red blood cells leading to breathlessness and weakness due to a lack of oxygen, which is carried in red blood cells by the protein hemoglobin. In humans, a variant of that molecule causes the red blood cell to form a characteristic bent or sickle shape (Figure 3.17). For simplicity, we will label the allele that gives rise to the
normal variant A, and the allele that gives rise to the sickle variant B. If a person has two copies of the sickle allele (BB), a severe loss of red blood cells in that individual can lead to an increased likelihood of dying in childhood. A person with just one copy (AB), however, will be anemic but will be able to live into adulthood.

Another context is equally important, the environment. Malaria is caused by a single-celled parasite that is transmitted from the blood of one person to another by mosquitoes. If a person with the AA genotype gets infected, they can become very ill. But a person with the AB genotype is resistant to infection by the malaria parasite. Malaria is not found in temperate or polar regions of the world, but is very common throughout the tropics. So, for someone living in the far north who never gets infected, the hemoglobin characteristic of “malaria resistance” is never expressed. However, if that same individual were to travel to the tropics and become infected, she would be resistant to the disease. When living in the far north, individuals with AA and AB genotypes have identical phenotypes with respect to avoiding becoming ill with malaria because the parasite is absent. But in the tropics, AB individuals are much less likely to become ill with malaria, so in that environment the information content of “malaria resistance” is expressed or used.

The environment can also influence the heritability of continuous traits. Our previous discussion of heritability assumed that differences among individuals with different genotypes do not depend on their environment. That is, we assume that if an individual is 10% taller than another when growing in one environment, it will still be 10% taller in a different environment. But what happens when this assumption does not hold? Suppose, for example, that when a certain plant species is grown under shady conditions, all individuals are short and about the same size, but when it is grown in a sunny spot, some of those individuals are much taller than the others due to genetic differences. In other words, the genetic differences are apparent in some environments, but not in others. These kinds of differences in genetic expression as a function of the environment are referred to as **genotype-environment interactions** (Figure 3.18).

The presence of variation resulting from genotype-environment interactions can have large effects on heritability. Heritability is not simply a result of the genetic differences among individuals: Those genetic differences must result in phenotypic differences. Some kinds of genetic differences among individuals never result in phenotypic differences; for example, some types of variation in regions of the DNA that are not translated into proteins. In other cases, whether genetic differences result in phenotypic differences depends on the environment. When variation in genotype-environment interactions is present, the amount of expressed genetic variation may differ among environments. In the example given above, the plants grown in the shade were all of similar height; in other words, phenotypic differences were minimized in that environment. If the heritability of height were measured only in the shade, we would conclude that it

---

**Figure 3.17**
Some of the red blood cells of individuals with sickle cell disease have a distinctive flattened or sickle-shaped appearance. Those cells are rigid and unable to pass through narrow capillaries restricting the blood supply, and are rapidly broken down in the spleen leading to anemia. (Photo from OpenStax College, Source: Wikipedia)
was very low because the amount of phenotypic variation would be low. On the other hand, if heritability were measured only in the sunny environment, it would be larger. Evolution would be constrained in the shady environment because of a lack of heritable variation.

Variation within an individual’s genetic makeup leads to variation among individuals. Some of that variation is deterministic, e.g., individuals that differ in their genetic makeup. Some of that variation is context dependent, e.g., individuals that differ because they reside in different environments. Finally, some of that variation is due to contingent events, coming about through random events in metabolic processes or development. Individuals with different forms of a gene may be phenotypically identical in one environment and different in another, and those environmental triggers may have a random component. Thus, the usage of information in living systems depends on complex interactions that lead to emergent properties and a role for contingency.

**Evolution of the Information System**

The genetic system came about through a process of evolution (Table 3.1, principle 9). Several key events occurred during that process. The first event was the appearance of a molecule that could act as a carrier of information. As described in Chapter 1, it is likely that the first information carrier was RNA, rather than DNA. It is likely that the origins of the genetic code trace to this period as RNA molecules evolved relationships with amino acids.

The next key event was the replacement of RNA by DNA as the information carrier. This replacement likely occurred because DNA is a more stable molecule that can exist as very long double strands. The double-strand structure for the first time allowed for error corrections using the information in the complementary strand, making the information system more resistant to mutations. The legacy of RNA as the information
carrier can still be seen in our cells. RNA acts as an intermediary so that information in DNA is first transferred to RNA before being translated into proteins and other forms of RNA play various roles in the translation process (Figure 3.14).

At some point the genetic code was set in place. It is possible that the original code consisted of two-base codons. Such a code has only 16 possible codon combinations and lacks the redundancy of the current code (Figure 3.15). The evolution of a three-base code, therefore, both increased the number of possible amino acids being coded for and, again, added robustness to the information system.

Why is the genetic code the way it is? Are the particular associations of codons and amino acids inevitable because of the biochemistry of those molecules? Or, is the code merely a frozen accident – if life were to evolve all over again, would a different code form? This issue is one of current debate among scientists with no clear resolution at the moment. It is clear that the genetic code is not simply random, providing clues to its evolutionary origin. Amino acids with similar chemical properties tend to have similar codons, so that mutations that changed the amino acid composition of a protein would tend to have a small change in its biochemical properties and be less likely to completely disrupt its function. In addition, amino acids made through the same synthetic pathway tend to have the same first base in their codons, which is what you might predict if relationships between specific RNA sequences and amino acids came about through gradual evolution from an original code consisting of two-base codons to one consisting of three-base codons. One way to resolve this debate would be if we were to find life elsewhere in the universe and could examine its genetic code.

The genetic code is not quite universal. While the code shown in Figure 3.15 is found in nearly all organisms, exceptions exist. For example, in most species the codon AUG serves two purposes, coding for methionine and serving as a marker for the start position for translation. In most proteins the methionine is then clipped off the protein. However, in Bacteria and Archaea, UUG and GUG can also serve as start codons. Bacteria in the genus Mycoplasm translate the codon UGA as tryptophan, rather than as a stop signal. Several differences from the standard code are found in mitochondria and chloroplasts, especially a reduction in the code because several codons are no longer used.

In general, we would expect the code to evolve extremely slowly. Once the code is set in place, a change effects not just a single protein, but possibly every protein. Thus, any changes are likely to be highly deleterious. Many of the known variants are replacements of start or stop signals with amino acids, rather than replacements of amino acids with stop signals, or switches of one amino acid for another. One can imagine that the first sorts of changes are less likely to be deleterious than the others. For example, it would be very hard to create a functional protein if a change created a stop codon in the middle of the gene. That some of the most numerous changes are found in mitochondria and chloroplasts can be explained by the fact that they are the result of an ancient symbiosis, so that many of their proteins are now produced by other genes (see Chapter 5).

What is a Gene?

At the beginning of this chapter, we stated that the gene was the fundamental unit of information in living systems. But what is a gene? The concept of the gene has undergone a continual evolution since it was first proposed by Mendel. Initially, a gene was associated with a specific characteristic of an organism such as the color of a pea. With the discovery of linkage, genes on chromosomes were treated like beads on a string. The concept of a gene became associated with a discrete unit that could be localized on a chromosome.
The association of genes with the molecular properties of cells produced the “one gene-one protein” concept. A gene was still thought of as a discrete unit, but now the association between genes and organismal properties was moved back a step. Along with this move was the growing appreciation that organismal characteristics like size, shape and color were often determined by more than a single gene.

The concept of the gene became much more complex in the middle of the 20th century as the structure of DNA and the genetic code was worked out. The gene was no longer a discrete unit, but a long strand. Recombination could occur within a gene as well as among genes. Still, the gene was seen as a single stretch of DNA. That idea had to be modified beginning in the late 1970s with the discovery that a protein was not the straightforward product of a single stretch of DNA. For many genes, it was discovered that after transcription some of the RNA was snipped out of the sequence so that the eventual protein was coded for by several disconnected stretches of DNA. Then scientists discovered that RNAs were sometimes assembled from pieces coming from very different places along the chromosome, not even from a single stretch. In some cases, a single stretch of DNA would result in a RNA molecule that would be associated with many other RNAs in different combinations, so that a single stretch of DNA could be translated as parts of more than one protein. In other cases, it was discovered that a single stretch of DNA could be translated into more than one protein by the simple process of starting the translation offset by one base pair, or by translating both strands of the DNA molecule. Finally, a stretch of DNA could also include parts that were never transcribed, but were essential in regulating the transcription of the adjacent parts. Thus, we have gone from a concept of a gene as a discrete unit associated in a simple fashion with a characteristic of an organism, to a complex, construct that is multipartite with a multifaceted relationship between the DNA and the resulting proteins and on to the characteristics of the organism.

Today scientists no longer have a single concept of the gene. Instead, they use different concepts depending on the circumstances. While such conceptual flexibility can be useful, it also has its pitfalls. For the purposes of modeling the evolutionary process, the concept of the gene as a discrete unit is often convenient (see Box 4C). But such models fail to account for the complex ways that mutation and recombination can occur. More complex models can be built, but are often much less tractable; one result is that computer simulations take much longer to run. The key is deciding when the simpler model is sufficient to provide reasonable predictions.

Technology is now capable of sequencing your genome in just a few days, but that’s only a small part of how your genetic information is expressed. We need more than just the sequence of base pairs, we also want to associate that sequence with functional consequences, the proteins that they produce and the organismal characteristics that they affect. Unfortunately, it would be impossible to work out those associations separately for every stretch of DNA. Instead, sequences of DNA are matched against other, known sequences from other species. Here, the gene concept that is used becomes critical. Do we look for matches against a stretch of DNA that includes the control regions, just the transcribed regions, or just the translated regions? How do we account for cases where RNAs are assembled from pieces transcribed from widely separated stretches of DNA? How do we ascribe a function to a stretch of DNA whose RNA is combined to create more than one protein? All of these issues are combined with the fact that the sequences of base pairs in a stretch of DNA varies among individuals within a species as well as between species. Sometimes those sequence differences result in no change in the protein product, but sometimes they do and those changes in the protein may or may not alter the way it functions (see Box 6D). How different can they be and
still function in a similar manner? Would we be able to recognize these two genes as being related to each other?

Recognizing the complexity of the concept of the gene is important for how biologists pursue their science. It is also a lesson in how all of the concepts presented in this book need to be considered as possibly having multiple meanings that can change over time and depend on the context in which they are used.
Chapter 4
Evolution

As discussed in Chapter 1, one of the most important and pervasive hallmarks of life is change. Individuals begin changing the moment they are born and continue doing so until they die. Change also happens from one generation to the next; children differ from their parents, and these changes can accumulate over long stretches of time. Such inter-generational change, when it is due to differences in the genetic makeup of those individuals, is called evolution. Recognizing that evolution happens and putting forth a clear theory of how it occurs was one of the central achievements of biology in the 19th century. It was in the 20th century that the theory of evolution was fully fleshed out and united with the rest of the biological sciences. Evolution is now seen as a cornerstone to understanding why life on Earth takes the forms that it does, and why organisms function the way that they do. The centrality of evolution is summed up in a famous statement by Theodosius Dobzhansky, “Nothing in biology makes sense except in light of evolution.”

Evolutionary biologists try to answer two broad questions: where did living organisms in all their diversity come from, and what is the reason for that variety of form? These questions lead to studies of the evolution of species and their characteristics from four perspectives: functional, developmental, historical, and adaptational. These perspectives each provide alternative explanations for why a trait exists in the form that it does. The functional perspective asks how limitations on the way an organism can be put together and operate constrain trait evolution. The developmental perspective asks how the developmental system generates forms and how that process constrains the forms that are possible. The historical perspective asks how the characteristics of the ancestors of a species determine the current form and constrain future evolution. Finally, the adaptational perspective asks how the ecology of a species molds its traits. In this chapter we consider several historical effects and the process of adaptation. Functional, developmental, and ecological processes are considered in Chapters 5, 6, and 7.

The Process of Evolution

Take a walk in a garden or the woods. Look closely at the flowers and the birds. At first it seems as if all of life has been well designed so that organisms are all perfectly suited to their environments. Some flowers have broad petals that provide a landing platform for bees. The bees have hairs on their heads and on their hind legs that carry the pollen they collect from one flower to the next for pollination. Yet as we look closer, what seems like a perfectly designed machine begins to look much more piecemeal. Most organisms are more like a collection of parts than a streamlined machine, some of which do a job well while other parts are simply adequate to the task, and some parts even have no function at all. While some flowers provide nectar for the bees, others that look similar provide no nectar. Some bees will chew holes in the side of a flower and drink the nectar, while never carrying away pollen to the next flower. Other aspects of organisms seem odd or useless. For example, human males have nipples even though they will never nurse a child. Many species that live in caves have rudimentary eyes, even though they live in complete darkness and thus the eyes serve no function. Because the production of eyes takes energy and resources, eyes in cave-dwellers could even be considered to be a detriment.

This mixture of organisms and characteristics, some well fit for their environments and some not, occurs because of the process of evolution. As a process of nature,
evolution often leads to organisms that perform well, but also to ones that fail to do so. Understanding the process of evolution, the goal of this chapter, will provide an understanding of why and how the living world exists in all of its variety, especially why life sometimes is not as well designed as it might seem.

The process of evolution often works by changing the function of something that already exists, such as the feathers of birds adapting to aid flight. When they originally arose on dinosaurs, feathers most likely functioned as insulation, just like hair on mammals. Later, in the lineage that lead to birds, evolution shaped those feathers to aid in running (like a spoiler on a sports car) or in gliding (Box 4E). Only then did the process of natural selection lead to feathers that functioned best as aids to powered flight.

Because the process of evolution often has to start with something that already exists, the result is often a compromise among functions. For example, in human women the width of the pelvis is a compromise between being wide enough to permit childbirth, without being so wide that walking becomes difficult. Often, the outcome of evolution is something that functions well for some purposes, but at a cost. For example, humans stand upright unlike our nearest relatives, chimpanzees and gorillas. That upright posture has many advantages, among them the freeing of our hands to better manipulate our environment, but we pay for that upright posture with backaches and spinal problems.

Evolution is not a process of increasing perfection and complexity. Despite popular images that show a straight line progress from fish to reptiles to mammals to apes to humans, the process of evolution does not progress toward some ultimate goal. Rather than sitting at the pinnacle, humans share the same 3.8 billion year history as all living organisms on earth. While evolution can lead to more complex organisms, the opposite is also true; blind cave-dwellers, for example, have evolved to be less complex. As described in this chapter, many processes contribute to evolutionary change; some of those processes increasing complexity, and some decreasing complexity.

Despite appearances, the process of evolution is not the result of organisms or nature trying to achieve some goal. Birds did not evolve wings because they needed to fly; rather, birds are able to fly because they have wings. Evolution is always backward looking, meaning that organisms are adapted to the conditions that existed in the last generation. Only because those conditions – which favored their parents and allowed them to survive and have more offspring than their competitors – are likely to occur again do organisms appear to have adapted to function at their peak in the conditions in which they live. In general, the environment that an organism experiences is much the same from year to year, generation to generation. If change occurs, it occurs slowly enough that while an organism may not be perfectly adapted to its environment, it is adapted well enough to survive and continue to propagate. Only during times of extremely rapid environmental change, as is happening now with global climate change, will many species find themselves in environmental conditions beyond their abilities to survive and reproduce, and well beyond the capacity of the species to adapt over future generations.

The Theory of Evolution
The theory of evolution consists of seven fundamental principles (Table 4.1). The first three principles (change over generations, speciation, and single origin) are
statements about the fact of evolution itself, while the other principles are statements about the mechanisms responsible for evolutionary change. These principles make up the core of Charles Darwin’s book On the Origin of Species by Natural Selection published in 1859. What Darwin was missing, however, was an understanding of genetics (see Chapter 3). Beginning in the 1920s, the unifying of the theory of genetics with that of evolution, and the enfolding of the fields of systematics, paleontology, and ecology through the 1950s, came to be known as the Modern Synthesis. The Modern Synthesis refined those principles, adding especially the knowledge of genetic variation (principle 5) and its pattern and source in mutation (see Chapter 3).

The shape of the theory of evolution, like all scientific theories, continually changes. From Darwin’s original theory, many controversies and new developments in technology and related sciences have constantly refined and bolstered the concept of evolution. For example, contingency, which was included in Darwin’s original theory, was considered a minor factor until the rise of molecular genetics in the 1960s. This development sparked a two-decade long debate about the importance of contingency, especially whether it should supplant natural selection as the primary agent of evolution. Today the primacy of natural selection (principle 6) remains as first put forth by Charles Darwin 150 years ago, but the importance of contingencies, both large and small, are also recognized (principle 7). The theory of evolution may itself continue to change with the explosion of new biological knowledge in the 21st century.

<table>
<thead>
<tr>
<th>Table 4.1. The fundamental principles of the theory of evolution.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The characteristics of organisms change over generations.</td>
</tr>
<tr>
<td>2. Species give rise to other species.</td>
</tr>
<tr>
<td>3. All organisms are linked through common descent.</td>
</tr>
<tr>
<td>4. Evolution occurs through gradual processes.</td>
</tr>
<tr>
<td>5. Organisms within species vary in genotype and phenotype.</td>
</tr>
<tr>
<td>6. Evolutionary change is caused primarily by natural selection.</td>
</tr>
<tr>
<td>7. Evolution depends on contingencies.</td>
</tr>
</tbody>
</table>
Because of the centrality of evolution to understanding living organisms, Charles Darwin is arguably the most important biologist to have ever lived. However, his importance could not have been foretold by his early life. Darwin was born February 12, 1809 in Shrewsbury, Shropshire, England. He attended a nearby Anglican boarding school in his youth, then entered the University of Edinburgh in Scotland in 1825 to study medicine at the behest of his father. Soon finding that medicine was not to his taste he switched his focus to natural history, joining the Plinian Society in his second year, and began studying under Robert Edmund Grant, a proponent of Lamarck’s theory of evolution by acquired characteristics.

Darwin’s father soon enrolled him in Christ’s College, Cambridge in a Bachelor of Arts course to enter the clergy, but Darwin selectively neglected his studies as he had done at Edinburgh, preferring to go riding and collecting beetles with his cousin. It was this interest in natural history that led him to his great discoveries. He was introduced to Reverend John Stevens Henslow, who taught botany at Cambridge, becoming his favorite pupil. Enrolling in Henslow’s botany course and staying at Cambridge, he planned to travel to the Madeira Islands with classmates to study natural history, but decided to forgo the trip when Henslow recommended him to fulfill the role of naturalist aboard the H.M.S. Beagle on its planned two-year voyage. That voyage stretched from two years to five and became the turning point in his life.

Before the voyage began, the Beagle’s captain, Robert Fitzroy, gave Darwin the first volume of Charles Lyell’s Principles of Geology, which explained that landforms are made through gradual processes over immense periods of time. During the first stop of the voyage Darwin observed geological and fossil formations which bore out Lyell’s writings, including a white band high on a cliff made of coral and seashells, which pointed to the vertical movement of landmasses. Through the entirety of the voyage, Darwin would observe a wide variety of geological formations, and collect numerous fossils and living animals, many of which had not yet been catalogued. In South America, he uncovered fossils of giant extinct mammals, including huge armored plates like those found on modern armadillos.

While in the Galapagos Islands, Darwin made his now famous observations of birds, at first believing them to be a mix of finches, blackbirds and grosbeaks, each bearing characteristics that would allow them to fill a certain niche in their particular ecosystem. Darwin sent his samples and observations back to England throughout the voyage, establishing his reputation as a naturalist so that by the time he returned in 1836, he immediately began collaborating, discussing his observations with other prominent scientists of the day. In addition to his own research into natural selection, he was writing reports on his collections from the years aboard the Beagle as well as a multi-volume series entitled Zoology of the Voyage of H.M.S. Beagle and a book on the geology of South America supporting Lyell’s theories. Studying the fossils that Darwin had collected, including giant sloths, a hippopotamus-like rodent called Toxodon, and the
giant armadillo-like Glyptodon, the anatomist Richard Owen realized that they were unrelated to any European animal, but instead were closely related to modern South American animals. Ornithologist John Gould determined that the Galapagos birds were, in fact, different species of finches rather than the many different types that Darwin had first identified. All of these revelations lead Darwin to begin hypothesizing that the animals he had observed had altered in response to environmental pressures unique to each individual island.

Soon after returning to England, Darwin’s health began to suffer from stress and overwork. He also likely contracted Chagas disease while exploring South America, an infectious condition that persisted for the rest of his life. He developed heart palpitations and, under the advice of a doctor, stayed with his Wedgwood relatives for a month to rest. There he met his cousin Emma Wedgwood, whom he would later marry in 1839. Unfortunately for his health, Darwin continued work on his burgeoning evolutionary theories unabated, questioning everyone from naturalists to farmers and pigeon fanciers, including mankind in his explorations after observing an ape in a zoo. His health problems continued throughout his life, periodically incapacitating him, with doctors at the time unable to arrive at a diagnosis.

Through his marriage Darwin inherited sufficient funds to become that fixture of 19th century British science, one that no longer exists, the gentleman scientist. Despite never earning even a college degree, he became one of the greatest scientists of his generation. He did this by funding all of his research using his own money, most of which could be carried out by examining specimens at home. Thus, he could devote himself to thinking about evolution and formulating his theories. This advantage is rarely available today, as most scientists are employed as university professors who must devote a large part of their time to teaching or other activities (such as endless faculty meetings) or are in industry where they must focus their research on products for market.

Darwin was spurred by his readings of Thomas Malthus’s human population theories, which stated that the human population would double every 25 years, but was held in check by disease and famine, as well as by A. P. de Candolle’s observations of the struggle for resources among wildlife. He postulated that species always breed beyond available resources, thus creating pressures that might favor individuals with certain variations, allowing those individuals to obtain more resources and produce more offspring, thus increasing the frequency of those variations in a population, eventually forming a new species. On the Origin of Species was finally published in 1859, and despite avoiding the even-then loaded word “evolution,” it immediately sparked controversy.

The book became something of a rallying point and firebrand for the conflict between the older generation of aristocracy- and clergy-based amateurs and the new generation of professional scientists, with the conflict being played out on a very public stage of debates, articles, satire and newspaper caricatures. Botanist Joseph Hooker and comparative anatomist Thomas Huxley emerged as the foremost proponents of Darwin’s
work, Huxley’s fierce defense of Darwin earning him the nickname “Darwin’s bulldog.” Richard Owen, meanwhile, lead the opposition. Darwin followed the controversy closely, and although his continuing ill health prevented him from appearing at public debates, he gathered support through correspondence. Public interest was quickly sparked by On the Origin of Species, and people flocked to Huxley’s lectures.

Despite illness, Darwin continued his work through the rest of his life, taking up the question of human evolution in *The Descent of Man*, published in 1871. In it, he addressed both the concept of sexual selection as it applied to human evolution and human culture, and as it applied in the natural world, in structures such as the showy but seemingly useless plumage of birds. He would continue to publish in the years before he died in Downe, Kent, England on April 19, 1882, publishing a total of 17 books over his lifetime. He was given a state funeral and buried in Westminster Abbey, near Isaac Newton.
Patterns of Evolution
Change over generations

The recognition that the characteristics of organisms change over generations, which Darwin called “descent with modification” (Table 4.1, principle 1), began in the late 18th century with the rapid discovery of fossils of creatures that had no known living counterparts, including the first dinosaur fossils. By the early 19th century it was obvious that these fossils were not just a few oddball creatures that could be ignored, and it was becoming clear that the Earth was much older than previously believed, long enough for species to have come and gone. We now know that the Earth is approximately 4.5 billion years old, and with that long age has come a long and varied history of life.

Figure 4.2
Whales are descended from animals that originally lived on land. Their ancestors were much like hippopotamuses, spending much of their time in the water eating plants. Later, they evolved to being fully aquatic and meat eaters, a process laid out in the fossil record. Our understanding of the origin of whales is still growing, much of it discovered in just the past 20 years. The first fossils of *Ambulocetus* were found in the 1990s. (Original illustrations by Nobu Tamura, compiled by Niusereset, Source: Wikipedia)

This change through time is exemplified by the evolution of whales from land-dwelling animals (Figure 4.2). Whales arose about 53 million years ago from hoofed animals and are most closely related to the hippopotamus. The ancestors of whales, while originally fully land-living, evolved to a life partly on land and partly in the water like today’s hippopotamuses. Later they adapted to a life mostly in the water and switched from being plant eaters to being carnivores, similar to modern seals. Eventually, their descendants became fully aquatic with their hind legs completely disappearing.

While we often think about evolution in terms of these sweeping changes over very long periods of time, all such changes result from the accumulation of many small changes in each generation. Nor are such large changes the only types that count as evolution. Any heritable change from one generation to the next is a form of evolution. A mutation (see Chapter 3) that appears in an offspring is evolution because that offspring has a different genetic makeup from its parent. That mutation in a single individual may eventually end up in all of the individuals in a species, or it may be quickly lost.

One such set of small changes that is having an increasing impact on humanity are the ones taking place in various populations of disease-causing micro-organisms. For example, in the past few decades many bacterial species have become much more resistant to the effects of antibiotics. When antibiotics are used, a few individuals that are more resistant than others may survive. The patient may become healthy, but the offspring of those resistant bacteria may be transmitted to others with the process repeating itself. The result is that today Bacteria that cause sexually transmitted diseases, tuberculosis, staph infections, and other illnesses have become serious health problems.
**Speciation**

One outcome of evolution and natural selection is **speciation**, the production of new species (Table 4.1, principle 2). During the process of speciation two or more populations of the same species become isolated from each other and may adapt to different environmental conditions over a long period of time. The most common process of speciation – ecological speciation – occurs when populations begin to respond to differences in natural selection in their disparate environments, slowly changing the genetics of the population as a whole. Eventually, the populations’ genetic makeup become so different that they are reproductively isolated, meaning they can no longer interbreed. Most biologists consider reproductive isolation to be the hallmark of two populations that have become different species.

We still know surprisingly little about the process of speciation. Besides divergent natural selection, speciation can also happen through a variety of other processes such as mutation and random genetic drift. As discussed later in this chapter, there has been vigorous debate over the past century about the relative importance of natural selection versus these other processes, with the current consensus being that natural selection plays a predominant role. The genetic mechanisms responsible for reproductive isolation are just now being unraveled and are an exciting area of current research that is driven by the quickly falling costs of sequencing DNA. The geography of speciation has also been a topic of contention. Typically differentiation happens in populations that are geographically distant from each other. However, it also can occur in adjacent populations, or even within a single population. The relative importance of speciation in distant populations versus adjacent populations has been debated, or even if speciation in adjacent populations could even occur. Today we know that the latter is unlikely, but possible. How to decide which individuals should be grouped together into a single species is a task that stretches back well before the concept of evolution entered the picture, at least as far back as the ancient Greeks 2500 years ago. Our modern system derives from that proposed in the 18th century by the Swedish naturalist Carl Linnaeus (Table 4.2). He gave names to over 12,000 plant and animal species; today there are approximately 2,000,000 named species and current estimates of the total number of species range from 5,000,000 to 30,000,000.

---

**Table 4.2. The systematic hierarchy with an example, humans.**

<table>
<thead>
<tr>
<th>Domain</th>
<th>Eukaryota</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kingdom</td>
<td>Animalia</td>
</tr>
<tr>
<td>Phylum</td>
<td>Chordata</td>
</tr>
<tr>
<td>Class</td>
<td>Mammalia</td>
</tr>
<tr>
<td>Order</td>
<td>Primates</td>
</tr>
<tr>
<td>Family</td>
<td>Hominidae</td>
</tr>
<tr>
<td>Genus</td>
<td>Homo</td>
</tr>
<tr>
<td>Species</td>
<td>Homo sapiens</td>
</tr>
</tbody>
</table>

Scientists now recognize that the levels in the hierarchy are arbitrary, since species come about through a process of continual branching. Organizing those species into a hierarchy, however, makes it easier to describe their relationships and discuss their common features.
While the Linnaean system is hierarchical, Carl Linnaeus did not intend it to indicate actual relationships; that is, his original system did not include the concept of descent from a common ancestor or species relatedness. However, Charles Darwin realized that his theory of evolution led naturally to a hierarchical set of relationships among all species with implications for the relatedness of all life.

**Common descent**

All life on Earth is linked through common descent from a single origin of life (Table 4.1, principle 3; see Chapter 1). As discussed previously, this single origin was not inevitable. It is quite possible to imagine an Earth in which life originated multiple times, each with its own set of descendant species. The most important evidence for a single, common origin can be found in our DNA. The genetic code that directs how the information in our genes determines our phenotype (see Chapter 3) is the same for nearly all species on this planet. The code is complex, and alternative versions are possible. However, the chance of exactly the same code arising more than once is so small, the simplest explanation for this shared code is that it arose once in the common ancestor of all lifeforms.

We can use the information in those genes to infer relationships in a science called **systematics**, the process of determining the relationships among organisms. Two species are inferred to be related to each other if they share common features, either DNA sequences or traits like similarly-structured skeletal forms. If two species share a trait that no other species share, then we can infer that they are more likely to be more closely related. When we look at the millions of species in the world, inferring relationships is not simple. One trait may be unique to one pair of species, while one of those species may share a unique trait with yet a different species. It is only by looking at very many traits and very many species that we can determine likely relationships. For example, all mammals have hair, one piece of evidence that they are all related to each other. The inverse is also true; although birds and bats both have wings, because of many other features not shared by those groups such as hair on bats and feathers on birds, we know that they are not closely related. Even with all of the information that we now have about the millions of species worldwide, our best understanding of their relationships may change as more information becomes available.

![Figure 4.3](http://example.com/chimpanzee.jpg) A. chimpanzee, (Photo by: Thomas Lersch, Source: Wikipedia) B. human. (Photo by A. Cohen)
The past 25 years has seen a great rise in the use of genetic information to infer relationships among species. Sometimes this information has resolved long-standing disputes about relationships, such as resolving that humans are more closely related to chimpanzees than to gorillas (Figure 4.3). In other cases, this information has lead to surprising new understandings about relationships. For example, a recent analysis of flowering plants showed that many species living in North America and Eurasia were more closely related to tropical species from South America and Africa, respectively, than to each other. The latter example demonstrates how morphology and genetics can sometimes lead to different conclusions about relationships.

Figure 4.4

To further muddy the waters, two species may appear alike because both have been shaped by natural selection to live in similar environments. For example, dolphins, which are mammals, and Ichthyosaurs, extinct reptiles that lived at same time as dinosaurs and only very distantly related to mammals, had similar body shapes because both lived in the ocean and ate fish (Figure 4.4). When trying to infer relationships among species, we need to account for how evolutionary processes such as natural selection for certain types of traits such as body shape can make them different or similar. As our understanding of the evolutionary process improves, we can better infer relationships. Despite these difficulties, scientists are confident that they understand the broad pattern of relationships among all of life (Figure 4.5). A striking feature of that pattern is that it is tree-like, which is exactly what is predicted if all of life is linked through common descent.

While the overall structure is tree-like, if we look closely at some groups we find a more complex pattern. A tree-like pattern occurs because a new species always branches off from an existing species (Table 4.1, principle 2). However, some species even after they branch off will occasionally form successful matings with other species, a process known as hybridization. While many such matings produce no offspring or sterile offspring (e.g., the production of a mule through the mating of a horse and a donkey), in some cases the offspring are fertile. If these matings occur often enough, a new species might arise which is derived from two other species, rather than branching off from a single species. So among some groups of species, rather than a tree-like pattern there would be a web-like pattern (Figure 4.6). Usually, hybridization occurs only among species that are closely related. Thus, while at the very tips of the tree in a few places a web-like pattern may be found, the larger pattern remains tree-like.
As with all of biology, there are always exceptions to the rule. Unlike more complex organisms, in which offspring receive half their genes from the mother and half from the father, among Bacteria and Archea the sharing of genes often occurs piecemeal (see Figure 3.12). One individual may pick up just a piece of a gene from another, or a whole gene, or a large number of genes. This sort of gene sharing between different species is called **horizontal gene transfer** and can occur among species that are not closely related. The extent of horizontal gene transfer is unknown, and is still an active area of research. If it is very extensive, then the structure of relationships among Bacteria and Archea would be much better described as a web than a tree.

**What is a Species?**

The concept of “species” is central to our understanding of the process of evolution. Species are a fundamental unit of evolutionary change. The process of speciation is central to the divergence of organisms leading to the great diversity that we see today (see Chapter 7). Yet, despite the centrality of this concept, there are extensive debates among scientists about what is meant by “species.”

There are two broad ways in which species can be defined, based either on the history of past relationships among individuals or on the future potential for independent evolution. The **phylogenetic species concept** defines a species as a group of similar
organisms that share a common ancestor so that species are based on past relationships. The problem is that the true relationships are not known but must be inferred based on the characteristics of the individuals. Species are generally defined based on shared traits, which indicates that they share a recent ancestor. Because individuals vary, scientists are faced with the problem of deciding when two individuals are different enough to warrant being classified as members of different species. For this reason, the German taxonomist Ernst Mayr proposed the biological species concept in 1942.

The biological species concept defines a species as a group of actually or potentially interbreeding organisms that are reproductively isolated from other such groups. This delineates populations that are linked by their genetic changes through time. When groups of individuals in a species become adapted to different conditions, natural selection can reinforce any tendencies to be reproductively isolated, thus creating two species that can then become more distinct. Of this occurs because reproductive isolation can allow two groups to adapt further and further to different environments. Note that this definition is based not on the past history of the species, but about what might happen in the future. If two populations are separated by a large distance, such as on the two sides of a large continent, we cannot know for certain that their descendants will never interbreed. If individuals of that species were found across the entire continent, it is more likely that the populations will be part of a single evolutionary unit,
one species. In most cases, the two broad approaches – phylogenetic and biological – result in classifying individuals into the same sets of species, which is to be expected if past evolutionary processes are continuing today and into the future.

The types of information used to decide which individuals belong to the same species can sometimes lead to differences in how they are classified. For example, using morphological data can fail to distinguish individuals that appear to belong to the same species and yet are reproductively isolated from each other. More common are individuals that appear to be different, especially if they come from distant locations, yet are still able to interbreed. If those individuals are connected by others that are morphologically and geographically intermediate, they would probably be classified as members of the same species. Using genetic information – DNA sequences – does not necessarily solve these problems. While genetic differences will always exist, those differences may be in only a few genes out of tens of thousands and hard to find. More common is the opposite problem, when individuals within a species show extensive variation among individuals. It is then hard to decide how different individuals must be to belong to different species.

All of these issues have resulted in decades of debate over how species should be defined. Systematists, whose job consists of naming species and determining their relationships, almost always use some sort of phylogenetic species concept. Within that perspective, though, are dozens of subtle distinctions in exactly how one decides if two individuals are different enough to belong to different species and a variety of methods – mathematical and otherwise – in how relationships are determined. Other evolutionary biologists tend to use the biological species concept. This difference in perspective has resulted in a lot of unnecessary misunderstandings, especially claims that the concepts are incompatible. Once it is clear that one concept is about what did occur in the past and the other is about what might occur in the future, the two concepts can co-exist in harmony.

Figure 4.7
The results of an experiment in which there was selection on body weight in chickens. Individuals were selected at 56 days of age for either increased or decreased weight. After 54 generations, males in the line selected for increased weight were twelve times heavier than those selected for decreased weight. (from Jacobsson, L., H-B Park, R. Fredriksson, M. Perez-Enciso, P. B. Siegel, and L. Andersson. 2005. *Genet. Res.* 86:115-125.; Data courtesy of Paul Siegel; Photo by Christa F. Honaker)
Mechanisms of Evolution

Gradual processes

It is sometimes difficult to fully understand the sliding time scales involved in evolutionary change. Evolution works via processes that take place over many generations (Table 4.1, principle 4); however, that length of time varies by species, and many generations can mean several months for a bacterium or several thousand years for a giant redwood tree. Extreme differences in phenotype and genotype only happen after tens or hundreds of generations (Figure 4.7), since offspring tend to resemble their parents. This is why laboratory experiments use bacteria or other fast-breeding species, and it is so difficult to race the evolutionary path of species such as human beings.

Even though changes are gradual, this does not mean that evolution happens at a constant rate. A species may undergo a very dramatic change quickly and then might remain that way (with only small changes) for millions of years. This can be seen in creatures termed “living fossils”, species that look just like fossils that are tens, even hundreds of millions of years old. Of course, not all aspects of an individual can be fossilized and while the living fossil may possess the same physical structure as its many times over great-grandparent, it might differ in its behavior or physiology. Gradualism, therefore, encompasses a large variety of patterns. Some lineages may change rapidly, while others change slowly or not at all; some parts of an organism may go through many changes, while other parts are static.

As mentioned before, the idea of “gradual” is a relative concept. On a human timescale, tens and even hundreds of generations can seem gradual, while on a geological scale, a hundred generations is the blink of an eye. This difference in perspective of what is meant by gradual as opposed to rapid has lead to substantial arguments among scientists about this fundamental principle. Darwin’s original argument for evolution being gradual was related to his putting forth natural selection as the process responsible for that change. He recognized that because natural selection typically depended on small differences among individuals each generation, any changes would occur slowly.

Later generations of scientists argued about the evidence for gradual change, the pattern shown in the fossil record. In some instances, when looking at the fossils of a particular species, they saw evidence for extremely rapid changes happening in one period followed by very little change. This led them to argue that natural selection was not the primary cause of evolutionary change (Table 4.1, principle 6), but rather that mutation is the primary cause. However, we know now that much of the variation in the rate of evolution is likely due to variation in natural selection. While mutation can be one factor leading to change, it is not sufficient and is rarely the sole or leading cause. Natural selection can favor changes in organisms, but once a species has become well adapted to its environment, natural selection can act to prevent change. On the other hand, a lack of variation can prevent change as well, so the lack of mutations for favorable forms could also contribute to stasis. In addition, though, an apparent lack of change could simply be due to the incompleteness of the fossil record.

Although evolution takes place primarily through gradual change, in very rare instances large changes can occur. Both hybridization and horizontal gene transfer can result in very extensive changes in the characteristics of an organism in a single event. When two different but still genetically compatible individuals produce offspring or when entire genes are added piecemeal into an individual’s genotype, genes are brought together in new combinations. Such events may have been responsible for the origin of some of the fundamental characteristics of Bacteria such as their metabolic processes (see Chapter 6).
Box 4B
Critical Experiment: Predicting Transitional Forms

One of the key tests of the theory of evolution is the existence of species that are intermediate in their form and appearance between two other species. Predictions of the occurrence of intermediate forms come from the combination of the second, third and fourth fundamental principles (Table 4.1). These predictions can be made concerning living species, or two species found in the fossil record. A particularly striking and historically important example of the fulfillment of this prediction occurred just after the 1859 publication of Darwin’s On the Origin of Species. In 1861, in a limestone quarry near Langenaltheim, Germany, a fossil was discovered that had feathers, but also had many characteristics of dinosaurs, such as small teeth and a long, bony tail. This fossil was named Archaeopteryx, meaning ancient feather or wing. In later editions of his book, Darwin used this fossil as an example of a link between reptiles and birds.

This fossil fulfills the predictions of the theory of evolution in two ways. First, it contains a mix of features that belong to two groups of organisms, representing a transition in form from one type of organism to another. Second, Archaeopteryx lived at the same time as dinosaurs, around 150-155 million years ago during the Jurassic Period, thus making it reasonable to surmise that Archaeopteryx is the product of changes in a previous species of dinosaur. If the fossil had been only 50 million years old, long after the dinosaurs had gone extinct, or 300 million years old, long before the dinosaurs had existed, serious doubt would have been cast on the theory of evolution. Two strong predictions are therefore borne out: organisms with forms that are intermediate between other already known forms will be found, and those organisms will have lived at a time after one of those species but before the other.

Since the discovery of Archaeopteryx, additional fossils have provided more knowledge of the origins of birds, while at the same time yielding more tests of the theory of evolution. While Archaeopteryx has some birdlike features, it is still far from being a bird. Evolutionary theory therefore predicts that there must be additional fossils that bridge the gap between dinosaurs and birds. Since the early 1990s, various fossils have been found in China dating from the Early Cretaceous Period (approximately 140-100 million years ago). These fossils have feathers that are much more like those of modern birds than the feathers of Archaeopteryx, as well as other features that make them much more bird-like than dinosaur-like.

In addition to the gap between Archaeopteryx and birds, other fossil discoveries filled the gap between it and other dinosaurs. In 1969, a dinosaur named Deinonychus was found whose wrist bones were similar to those of Archaeopteryx. There have also been a number of claims for the discovery of very primitive feather-like structures on animals that otherwise are dinosaurs. Since then, many more fossils have been found, including Sinornithosaurus, which is a member of the group of dinosaurs thought to be the most closely related to birds. The question of exactly how birds are related to dinosaurs and the exact sequence of the appearance of various traits remains a very active, and sometimes contentious, area of research.

Reconstructed Sinornithosaurus, Carnegie Museum of Natural History (Photo by S. Scheiner)
This example of the evolution of birds, as well as that of whales (Figure 4.1), demonstrates how we can test scientific theories through observational experiments (see Chapter 2). In situations like this, where the theory relates to events that happened long ago and cannot be practically repeated in either a laboratory setting or in the field, we can only look for natural evidence of what might have occurred to shape the world as it is today. Yet, we have met the conditions necessary for a test of the theory—making a prediction that is then tested with data. It is important to recognize that a failure to find a transitional form is not, by itself, a repudiation of the theory of evolution. The fossil record is very incomplete because it is difficult for fossils to form, and some parts like skin and feathers are not preserved as easily as other parts like bones. The fossil has to survive for millions of years and then be found. So, while having a fossil is very strong, positive evidence, the importance of negative evidence is harder to judge. The continual discovery of new fossils that meet the two criteria of being of intermediate form and being the correct age continually reinforces evidence for the theory, making a lack of intermediate forms in some instances much more likely to be the result of a lack of fossils rather than a problem with the theory of evolution.

The theory of evolution does not rest just on observational experiments; other parts of the theory have been tested by manipulative experiments, such as artificial selection experiments (Figure 4.4). The broad acceptance of the theory of evolution as the explanation for the forms of life on Earth comes from 150 years of these repeated tests.
Variation

Within a species – even within families – individuals show remarkable variation, to the point that no two individuals of the same species are ever identical. Individuals of some species vary so wildly that they might not seem to be related at all - take the wide range of characteristics shown by domestic cats (Figure 4.8). Although domestic cats are an extreme example, such types of variation are found in all species. Human biases being what they are, we easily recognize such variation in our fellow humans. In some species the variation can be subtle, at least to human eyes, but is revealed by morphological characteristics such as measurements of body length and weight, behaviors, or physiological characteristics such as how fast an individual processes its food.

Figure 4.8
Domestic cats (*Felis silvestris*) are an extremely variable species, helped along by human breeding and artificial selection. Domestic cats are descended from wild cats that today inhabit the Near East. From this stock come individuals with long hair, short hair or no hair; a multitude of colors arranged in solids, stripes or patches; sleek and stout; and even lacking tails. (Source: *The New International Encyclopædia*, v. 4, 1905, facing p. 312.)

This variation exists in two ways, as variation in the appearance of individuals, the phenotype, and in the genetic makeup of individuals, the genotype (see Chapter 3). Phenotypic characteristics include outward appearance (such as height, having wings, or flower color), whether individuals tend to have a long or short lifespan, the behaviors of individuals, or an individual’s physiology and biochemistry (such as protein composition). Phenotypic variation can be extensive for traits such as seed size, which can vary 20-fold within a single population of a wildflower species, or mass, which can vary over several orders of magnitude in trees. Other kinds of traits, such as body size in insects, tend to vary much less among individuals in a population. Still other traits, such as the number of petals on a flower, may not vary at all within some species.

The same genetic individual may have a very different phenotype when grown under different environmental conditions. In addition, individuals with different genotypes may respond differently to environmental conditions (see Figure 3.16). Even two individuals with identical genotypes (twins) growing in identical environments, however, do not necessarily look alike or function identically. Small, random differences in when and how genes are expressed can lead to measurable differences in individuals. Those two types of variation, while often linked, need not be. Identical twins have the same genes, but differences in how they are raised, or even just random differences in how they develop in the womb, can lead to differences in how they look or act. And two individuals with a similar feature, such as identical height, may have different genotypes. In some cases, differences in genes may not even be expressed. Since evolution is change in genotypes.
over generations, a change in phenotype that is not due to a genetic change is not considered evolution. For example, the average American today is much heavier than a few generations ago due to a change in our diet, not our genetic makeup.

Variation, which is a primary requirement for natural selection (Table 4.1, principle 5), comes from genetic mutation and recombination (see Chapter 3). While natural selection acts to eliminate genetic variation as alleles that lead to increased fitness become fixed in a population and other alleles are lost, genetic mutation and recombination serve to provide new variation. Equally important are processes that move that variation among populations, since natural selection occurs within a population. Several processes can bring variation into a population from elsewhere. Migration is the movement of organisms or propagules (e.g., seeds, pollen, or gametes) and is a very important way in which mutations that appear in one population move to others.

### Natural selection

The idea that evolution came about through the process of natural selection was conceived separately by both Charles Darwin and Alfred Russell Wallace, and presented to the Royal Society of London in 1858. The following year, Charles Darwin published his ideas in his book, On the Origin of Species by Natural Selection, which brought them to the wider scientific community. That book was important for both presenting the broad outlines of the theory of evolution, as well as providing a simple and logical mechanism to explain evolutionary change. While the idea of evolution had been around for over 50 years, it was only after a clear and simple mechanism was proposed that the scientific community embraced the idea.

Central to Charles Darwin’s theory was that the process of natural selection was the most important process in causing evolutionary change (Table 4.1, principle 6). Since then, there have been various challenges to the primacy of natural selection, with natural selection retaining its leading role each time. The first challenges came because of a lack of knowledge about genetics and the source and maintenance of variation (see Chapter 3). Others disagreed that the process of selection was gradual, and felt that evolution was primarily driven by mutations which effect large changes in the phenotype. The refutation of these challenges led to the Modern Synthesis, the 30 year process from the 1920s through the 1950s during which the theory of evolution was placed into its current framework, with its relations to genetics and biology as a whole. The final challenge to the primacy of natural selection came over arguments about the importance of contingency, which led to the addition of the seventh fundamental principle. Today scientists agree that natural selection is the only mechanism that can produce sustained and large directional changes in species characteristics. However, it is important to keep in mind that those changes are influenced and channeled by other processes, such as contingency and limitations on the sources and types of phenotypic variation (see Chapter 6).

Natural selection occurs when individuals with differences in their traits leave different numbers of descendants because of that variation. Natural selection requires three conditions in order to lead to evolutionary change (Table 4.3): phenotypic variation among individuals in a particular trait; fitness differences (some individuals must leave more descendants than others as a result of the phenotypic differences), and genetic variation (the phenotypic differences must have a genetic basis). If these three components are present for a trait within a population, then the frequency of that trait will change in that population from one generation to the next: evolution by natural selection will occur.
The fitness of an individual is the result of the many parts of its life that affect its survival and reproduction, including the chance of surviving to adulthood, its mating ability, its fertilizing ability, and the number of offspring produced. There are many ways of measuring fitness, with one of the most comprehensive being to count the number of grandchildren that an individual has. This measure accounts for both the production of offspring, and for how well those offspring are provisioned or raised so that they also survive and reproduce. For evolution to occur, the other part – the genetic response – is equally important. However, differences in fitness among individuals could be due to chance and, thus, will not be associated with trait differences in a consistent fashion. A useful way to study the process of natural selection is to subdivide it into two parts: that which occurs within a single generation and that which occurs from one generation to the next. Within a single generation individuals vary in their phenotype and may have differences in fitness – survival and reproduction. This part is what we often think of as natural selection. The genetic response is the change in the genetic makeup of the population that occurs from one generation to the next, and depends on the heritability of a trait (see Chapter 3). If a trait is heritable and favored by phenotypic selection in one generation, the next generation will have a greater proportion of individuals with that trait. From one generation to the next, the change in the population may be small. But if the process of natural selection goes on for many generations, the population may become very different from the ancestral population. Much of the striking variation among species is due to such long-term evolutionary responses to natural selection.

The fitness of an individual is always relative to other individuals in a population. For example, consider a population consisting of brown mice and grey mice (Figure 4.9). On average, the grey mice have 10 offspring, and the brown mice have 11 offspring. Since the brown mice are more successful – more fit – than the grey mice, over time the population will evolve to be mostly or all brown. On the other hand, if the brown mice have 9 offspring on average, the population will evolve to be mostly or all grey. It is even conceivable that an individual could increase its relative fitness by harming itself, as long as it harmed the individuals around it even more. If that process continued, a species could evolve itself to extinction.

---

Table 4.3. The necessary and sufficient conditions for evolution by natural selection.

1. Individuals must vary in their characteristics (phenotypic variation).
2. That phenotypic variation must lead to differences in fitness (fitness differences).
3. That phenotypic variation must have, in part, a genetic basis (genetic variation).

---

The fitness of an individual is the result of the many parts of its life that affect its survival and reproduction, including the chance of surviving to adulthood, its mating ability, its fertilizing ability, and the number of offspring produced. There are many ways of measuring fitness, with one of the most comprehensive being to count the number of grandchildren that an individual has. This measure accounts for both the production of offspring, and for how well those offspring are provisioned or raised so that they also survive and reproduce. For evolution to occur, the other part – the genetic response – is equally important. However, differences in fitness among individuals could be due to chance and, thus, will not be associated with trait differences in a consistent fashion. A useful way to study the process of natural selection is to subdivide it into two parts: that which occurs within a single generation and that which occurs from one generation to the next. Within a single generation individuals vary in their phenotype and may have differences in fitness – survival and reproduction. This part is what we often think of as natural selection. The genetic response is the change in the genetic makeup of the population that occurs from one generation to the next, and depends on the heritability of a trait (see Chapter 3). If a trait is heritable and favored by phenotypic selection in one generation, the next generation will have a greater proportion of individuals with that trait. From one generation to the next, the change in the population may be small. But if the process of natural selection goes on for many generations, the population may become very different from the ancestral population. Much of the striking variation among species is due to such long-term evolutionary responses to natural selection.

The fitness of an individual is always relative to other individuals in a population. For example, consider a population consisting of brown mice and grey mice (Figure 4.9). On average, the grey mice have 10 offspring, and the brown mice have 11 offspring. Since the brown mice are more successful – more fit – than the grey mice, over time the population will evolve to be mostly or all brown. On the other hand, if the brown mice have 9 offspring on average, the population will evolve to be mostly or all grey. It is even conceivable that an individual could increase its relative fitness by harming itself, as long as it harmed the individuals around it even more. If that process continued, a species could evolve itself to extinction.

---

Figure 4.9
Grey and brown mice.
(Photo credits: Roger McLassus, George Shuklin, Source: Wikipedia)
How do we know that a particular trait was shaped by natural selection? Answering this question is not simple. There is no single method for doing so, and all methods require assembling multiple kinds of information. Some scientists have been guilty of assuming that characteristics of organisms were optimal, and were always adaptations produced by natural selection, without carefully documenting whether that was, in fact, the case. These assumptions were termed “Panglossian” after the character Dr. Pangloss in the novel Candide by Voltaire. In that story, the naive Dr. Pangloss goes through life confidently declaring that “All is for the best in this, the best of all possible worlds.”

Sometimes scientists have constructed scenarios to explain what processes may have led to particular adaptations, and then simply accepted those scenarios without substantiation. For example, some plant species produce seeds with a sticky coating. Some scientists originally claimed that this coating was an adaptation for dispersal; seeds getting stuck on the feet of ducks was one particular scenario they proposed. Such scenarios are termed “just-so stories,” after the title of the book in which Rudyard Kipling recounts fanciful tales about the origins of various animals. Today we understand that in many species, the coating has more to do with water retention by the seed than with dispersal. The problem with “just-so stories” is not that they are necessarily wrong; it is that they are based on scant evidence or on unfounded extrapolation from what is known.

Although speculations are an indispensable first step in posing hypotheses to test whether a trait has come about by natural selection, we must be cautious in accepting or dismissing such speculations. As scientists are usually knowledgeable about the organisms and systems they study, such speculations often prove to be correct. But until adequate evidence has been assembled that firmly supports these suppositions, they remain speculations.

For example, through a variety of types of evidence we now understand how natural selection has shaped the size of guppies in the Aripo River on the island of Trinidad (Figure 4.10). Guppies that live in the lower part of the river grow quickly, become adults at a small size, and have small offspring. Those in the upper part of the river, separated from the lower part by waterfalls, grow slower, become larger, and have larger offspring. A key difference in the environment is that the lower part of the river has large fish that eat both adult and juvenile guppies. Since the adults may get eaten, natural selection favors adults that reproduce quickly. In the upper part of the river are smaller fish that only eat the juveniles. In this environment, natural selection favors adults that can produce many offspring. So, the first piece of evidence that the differences in body size are due to natural selection is that the characteristics appear to match expectations for which would be most fit for each environment. The second piece of evidence is obtained by looking at survival of individuals of different sizes in each environment. As expected, when individual fish are marked and followed over several months, larger fish survive longer in the upper part of the river than in the lower part. The third piece of evidence comes from an experiment in which small guppies were placed in a stream that had neither guppies nor fish that would eat them. Within 20 generations the guppies had evolved to grow slower and become adults at a larger size, exactly as predicted. Evolution by natural selection was directly observed. While such complete evidence of natural selection exists for only
a few species, there is sufficient partial evidence for scientists to have confidence that they understand how natural selection has shaped the characteristics of most species. It is important to differentiate between “selection of” and “selection for.” Natural selection mostly works through differences in the fitness of individuals, selection of individuals, because it is the individual that lives or dies or has a certain number of offspring. But those fitness differences are due to differences in some trait, as with our example of brown and grey mice, selection for color. The result of the selection of individuals are changes in the relative numbers of brown and grey mice in a population, and the frequencies of the genes for coat color, and thus evolution of the population in which they exist.

Evolution is always ultimately about changes in the genetic composition of a species, but it is not always about selection of or for particular alleles. Sometimes there can be selection of species. For example, when the dinosaurs went extinct 65 million years ago and the mammals survived, there was selection of mammal species relative to dinosaur species. What was selected for is unclear and currently subject to debate. Mammals had some combination of traits – possibly warm-bloodedness, small size, and rapid reproduction – that increased their fitness relative to dinosaurs at that time. Selection often is for a combination of traits, rather than just for a single trait.
“Essentially, all models are wrong, but some are useful.” – attributed to statistician George E. P. Box

Although it can seem as though the ultimate goal of scientists is to develop general theories about the world, it is equally important to the advancement of the scientific method to develop specific models to help apply general theories to specific phenomena (see Box 2B). Since a general theory is only a guide for thinking about a scientific question, it requires a specific model that can be applied to a particular situation. Models can be used to more fully explore the implication of a theory and provide quantitative predictions and tests.

A single theory can give rise to more than one model. In the case of the theory of natural selection, two very different models are possible, one for traits that vary in a continuous manner and are due to variation in many genes, and a second for traits that consist of a few different types and are due to variation in a single gene. Creating these models can take several steps, depending on how complex the general theory is, and what, precisely, it is trying to predict. Take the, verbal description given in Table 4.2 as an example. The first step is to translate the description into a very basic mathematical statement: Evolutionary Response = Phenotypic Variation × Fitness Differences × Genetic Variation.

The next step is to make this a formal, mathematical model. One version is known as the breeder’s equation – $\Delta \bar{z} = s h^2$ – because it was used in agriculture to predict the effects of artificial selection on domestic animals and plants. Keep in mind that the process of evolution is not about what happens to individuals, but about what happens to groups of individuals – populations and species – over generations. This equation is a prediction about the change in the average characteristic of a population from one generation to the next, when that characteristic is caused by multiple genes.

This equation says that for some trait $z$, say the number of kernels on an ear of corn, we can predict the change in average number of kernels, represented in the model by the term (the Greek letter delta ($\Delta$) is often used to symbolize change; the line over the $z$ indicates that it is an average). This change is a function of three factors. First is the strength of selection ($s$) or fitness differences. In this instance, the farmer may select the 10% of corn plants that have the most kernels, or a strength of selection of 0.1. The second term in the equation is ($h^2$), the heritability of that trait (the amount of resemblance among relatives that is due to shared genes; Box 3C). If the farmer knows the heritability of the number of kernels produced by each ear of corn by plants in his field, he can predict how many more he will get on average in the next generation. Figure 4.4 shows the results of just such a selection experiment.

Since a single trait can be controlled by several different genes, models like the breeder’s equation don’t necessarily predict selection effects on any single specific gene. However, we can build a different type of model that will allow us to predict the change in the frequency of a single gene. Take the case of a gene that has has two forms, with the frequency of one of those forms symbolized as $p$. The model is expressed as: $\Delta \bar{z}$. Although this equation looks much more complicated than the previous one, it still uses all the same concepts as the breeder’s equation, just put together in a slightly different way. Again, $s$ is the strength of selection, and the rest of the equation is just a way of symbolizing the amount of genetic variation in the population for the gene we are tracking. In these examples, both equations are more formal representations (models) of the
same theory. Both contain a long list of overlapping but different background assumptions, which can be linked if additional assumptions are made. The more general theory serves to unify what at first may seem to be very different types of selection, a trait determined by many genes and one determined by a single gene. In turn, the theory of natural selection is one component of the more general theory of evolution (Table 4.1), which involves other processes as well.
Throughout his life, his passion for collecting never abated; a friend of his, Victor Luchnik, convinced him to switch to collecting ladybird beetles in 1915, the same year he read Darwin’s Origin of Species. These early experiences in etymology would shape the course of his research through the rest of his life as he studied populations of Drosophila in the wild as well as in the laboratory. When he entered the University of Kiev in 1917, he pursued a more traditional course of life science studies while continuing his work with ladybird beetles. Upon completing his studies (although he never received a formal diploma) he took a position as an Assistant with the Faculty of Agriculture at the Polytechnic Institute of Kiev, the equivalent of an assistant professorship in the United States.

Dobzhansky was introduced to the then-new field of genetics by a botany professor, Gregory Levitsky (with whom he would share an apartment in Kiev), and began to apply the principles of this recent science to his passion for ladybird beetles. When another professor, Nicolai Vanilov, brought a number of books and pamphlets on genetics into Petrograd (now Saint Petersburg) from Germany in 1921, Levitsky studied the collection and brought back a wealth of information to Dobzhansky. This was an immensely valuable resource, as at that time Russia had closed its borders to much of Western European culture, including scientific advancement. Using his newly-acquired information and his own body of ladybird beetle research, he published a paper in Germany in 1924 that came to two significant conclusions in regards to the genetics of natural populations: (1) He found no essential differences in variation between groups culled from wild populations and individuals within those groups. (2) While variation within a population is controlled by differences of single genes, variation among populations from different geographical regions was due to entire complexes of genes. This was to be the end of

Figure 4D.B
Ladybird beetle
(Photo by Gilles San Martin, Source: Wikipedia)
his significant work with ladybird beetles, however, and he began to use *Drosophila melanogaster* as a lab animal in 1922. Since he was focusing more and more on genetics, *Drosophila* represented a much easier subject to work with, with *D. melanogaster* having one-quarter the number of genes as the beetles. It was from his first steps in studying the shape of mutant female *Drosophila*'s sperm storage sacks that Dobzhansky concluded that every gene acts upon all parts of the body.

In 1927 Dobzhansky received a research grant from the International Education Board, and used it to take advantage of an opportunity to go to America and work with Thomas Hunt Morgan at Columbia University (see Box 3E). The pairing was a natural one, as Morgan had begun the use of *Drosophila* in genetics experiments, making use of its prolific and swift breeding process. Despite Dobzhansky’s initial handicap in his volume of genetics knowledge – due to the general dearth of new research in Russia – during the ensuing decade, Dobzhansky would accompany Morgan to the California Institute of Technology and begin his field experiments with *Drosophila pseudoobscura*, surveying populations in California and throughout the American West as well as South America. Because *D. melanogaster* was strictly a laboratory animal, Dobzhansky’s studies of *D. pseudoobscura* represented an important step towards uniting ecological and genetic studies within an evolutionary context.

During this time, he forged relationships with Alfred H. Sturtevant and Sewell Wright. In 1932, he briefly returned to his work with ladybeetles, publishing his first paper in English in which he argued that there was no essential difference between geographical and local variability, and that varieties of species are, in fact, incipient species themselves, as Darwin advocated. He found significant genetic similarities among regional populations, even, surprisingly, that populations of different species living in the same area would bear more genetic similarities than populations of the same species living in different regions.

Ten years after he arrived in America, Dobzhansky published *Genetics and the Origin of Species*, one of the most significant works of the Modern Synthesis. In it, he put forth a definition of evolution as “a change in the frequency of an allele within a gene pool,” an explanation in keeping with his field observations. Writing this book helped crystallize his view of evolution and the relation between Wright’s theories and his own experimental study of natural populations (Wright’s own papers being so thick with math, even his fellow geneticists such as Dobzhansky had to extrapolate the information contained within and assume his calculations were correct!)

Dobzhansky continued his fieldwork, collecting *Drosophila* from mountains in the Death Valley region of California. The data, which further supported Wright’s theories of semi-isolated populations and genetic drift, were the beginning of a series of papers on the genetics of natural populations that would continue for nearly 40 years and 43 publications. These studies were designed to test the hypotheses being generated by the new evolutionary theories. For example, to test the assumption that, freed from events such as a winter die-off, tropical climates would support populations of Drosophila large enough to almost completely eliminate genetic drift. He even made a trip to the Brazilian rainforest - an environment free from major seasonal variation – to test this hypothesis.

Dobzhansky returned to Columbia University in 1940 after a falling out with Sturtevant, and remained there until 1962, when he joined the Rockefeller Institute, staying until he retired in 1971. Afterward he remained active in academia, taking a position as an emeritus professor at the University of California-Davis and continuing his field work in North and South America. Suffering with leukemia, which had worsened by the summer of 1975, he passed away from heart failure on November 11, during a trip to San Jacinto, California.
Contingency

Although natural selection is a primary process in shaping the form and function of organisms, other processes may also be responsible for the forms of life we see today. Contingency, or chance events, can be important in several ways (Table 4.1, principle 7). First, chance plays a role in where particular species happen to live. For example, Australia and neighboring islands are the only place in the world that we find living kangaroos (Figure 4.11) and their close relatives, as well as all kangaroo-like fossils. This limited distribution happened because, by chance, the ancestor of kangaroos arose in Australia at a time that the continent was isolated from the rest of the world. Despite the fact that hopping can be a very energy efficient way to move at high speeds, nowhere else in the world is there a very large animal that moves by hopping. By chance, only in Australia did mutations appear in large mammals that led to a hopping gait, nor were there any other large terrestrial mammals that ran on four legs that might have out competed them. Thus, the particular mutations and patterns of selection that lead to these large hopping animals existing in one place and not others are the result of a long series of chance events, and show the importance of contingencies both large and small.

Second, chance plays a role in which traits some species happen to have. Mammals have backbones not only because a backbone is a handy thing to have, but because they inherited this trait from vertebrate ancestors (Figure 4.2B); all vertebrates share this trait for that reason. All warm-blooded animals with hair also happen to have backbones, because those traits happened to appear in the same lineage. When vertebrates first arose, backbones were likely favored by natural selection. But that past selection does not necessarily mean that there is currently selection for backbones in all vertebrate species, although admittedly it is hard to imagine selection against the presence of a backbone given the central role it plays in mammal form and function.

Third, chance plays a role in which genes happen to get passed along to the next generation, a process known as genetic drift. Genetic drift refers to changes in gene frequencies due to random sampling effects, the way in which by chance alone only some alleles get passed to the next generation (Figure 4.12). Take the example of a population that consists of yellow-flowered and white-flowered plants. In this population, flower color has a genetic basis. In one generation, by chance alone, more yellow-flowered plants happen to be growing in richer soil. Because of this, the yellow-flowered plants, on average, produce slightly more seeds than the white-flowered plants. In the next generation, the frequencies of the two flower colors in the population will have changed, but not because of natural selection. Having yellow flowers instead of white did not play a direct role in those plants being able to leave more offspring; rather, the change in trait frequency is due to random seed establishment. If in the next generation white-flowered plants happen to end up on the richer soil, then the opposite change will occur.

In most populations, changes due to genetic drift will be small. Generally, genetic drift is important only in populations smaller than 100 individuals; in populations greater than 1000 individuals, its effects are usually much smaller than those of other evolutionary processes. Thus, population size is a critical parameter in determining how evolution
Figure 4.12
Changes over 20 generations in gene frequencies due to random sampling. (A) In this small population there are 9 diploid individuals. In the first generation there are 9 copies of the A allele and 9 copies of the a allele, making the frequency of the A allele is 1/2 (p = 0.5). Over generations, random sampling from such a small pool means that the frequency of the allele can go through large changes in just a single generation. Different populations, represented by each line in the figure, can quickly end up with different gene frequencies. (B) In this larger population there are 50 diploid individuals. Again, in the first generation the frequency of the A allele is 1/2 (p = 0.5). Because of the larger population size, sampling effects are much smaller and the populations take many more generations to differ in their gene frequencies. As population size increases, the effects of genetic drift decrease.

New variation through migration will not be available because the few populations will all contain the same alleles, and new variation through mutation will be rare because the number of new mutations depends on the number of individuals. In such populations, if the environment changes such that new alleles would increase fitness, those alleles may not be present, and the population may be unable to adapt to the new conditions. Such concerns are even greater today because of human-caused changes to the environment; for example, global warming may cause large changes in local environments in just a few decades. Small populations, unable to adapt, will be faced with extinction.

The importance of contingency as a factor in evolution, and its elevation from a minor phenomenon to a fundamental principle, occurred over several decades. As with many ideas in evolution, the role of contingency in evolution was first proposed by Charles Darwin. However, the idea that evolution might depend on random chance quickly fell into disfavor and by the beginning of the 20th century was cited as a reason why Darwin’s theory of evolution by natural selection should be replaced by some other mechanism. Then, in the 1930s, Sewall Wright proposed the first models of genetic drift. However, it was Motoo Kimura (Box 4D) in the 1950s who first proposed that genetic drift could be a major evolutionary process. Even then, the importance of genetic
drift was not actively debated until the 1960s when the first data on variation in proteins became available. Those data suggested that there was much more genetic variation at the molecular level than seemed possible, given how much importance natural selection was assumed to have. That debate intensified into the 1970s and 1980s as data on variation in DNA became available. It was proposed that evolution at the molecular level was fundamentally different than evolution at the level of visible traits, with genetic drift being the primary process determining the evolution of molecules. Today, as we have obtained more and more information on DNA variation in many species along with methods for assessing selection on such traits, it has become clear that much of this variation can be explained as the result of natural selection.

Along with that debate was a second assault on the primacy of natural selection. This debate concerned the cause and rate of evolution (gradual vs rapid). A theory termed punctuated equilibrium was proposed that stated (1) all evolutionary change occurred at the moment of speciation, and (2) chance, contingent events like mutation were the primary process responsible for those changes. While evolutionary change within a species may be rapid at some times and slow or non-existent at others, there is no evidence that the change happens only at speciation. Still, the debate over punctuated equilibrium helped to highlight the importance of contingency. Since then controversies over the general importance of natural selection vs contingency have subsided. Instead, biologists ask how both processes have shaped the evolution of particular sets of species.
The most important change in the theory of evolution following the Modern Synthesis came from Motoo Kimura, a Japanese biologist and mathematician. Born November 13, 1924 in Okazaki, Aichi Prefecture, Kimura had a keen interest in botany and mathematics from a young age, teaching himself geometry and other disciplines during a long recovery from food poisoning. While attending high school in Nagoya, he studied plant morphology and cytology, and worked in M. Kumazawa’s laboratory studying the genetic composition of lilies. He soon found that he could pursue his interest in biology and mathematics in biometry, the application of statistics to biology (a discipline widely used in fields such as agriculture and medicine).

Entering Kyoto Imperial University in 1944, Kimura joined the botany program under the advice of geneticist Hitoshi Kihara, whose laboratory Kimura eventually joined. His choice was heavily influenced by his desire not to fight in World War II; by going into the botany rather than the cytology program, he was able to avoid military duty. His early life might have helped to spark this interest as well; his father, a businessman, loved flowers and raised ornamentals. Kimura’s curiosity in their development spurred his father to buy him a microscope. Later in life he became an avid orchid breeder, creating prize-winning varieties. He also liked to paint pictures of favorite flowers, especially on chinaware.

Working with Kihara, Kimura was introduced to the study of population genetics and worked closely with the chromosomal structure of plants. Joining the National Institute of Genetics in Mishima in 1949, he published his first paper on population genetics in 1953, in which he expanded upon models of the effects of population structure on evolution put forward by the geneticist Sewall Wright. He developed his understanding of these models based almost entirely on self-study, his formal mathematical training being very limited and there being almost no colleagues in Japan at the time. Kimura entered graduate school at Iowa State College that summer under a Fulbright Fellowship, but soon transferred to the University of Wisconsin. There he worked with James F. Crow, Sewall Wright, and other prominent geneticists, and developed a general model for genetic drift which took into account multiple alleles, selection, migration and mutations. In 1955, he presented a paper at the Cold Spring Harbor Symposium which drew great praise from Wright and J. B. S. Haldane, who, along with Wright and R. A. Fisher (see Box 2D), laid the theoretical groundwork for the Modern Synthesis. In just two years as a graduate student at the University of Wisconsin he managed to produce numerous papers that established his reputation.

After receiving his Ph.D. in 1956, Kimura returned to Japan and once again began work at the National Institute of Genetics, where he would remain for the rest of his working life. During the following decade, he would collaborate on much of his work with Takeo Maruyama, and introduced new models for the study of genetic drift. These were used increasingly widely as the field of molecular evolution gained prominence and as information about protein composition and genetic sequences grew. He worked on non-genetic molecular studies as well, developing methods for distinguishing different forms of the same protein. This work contributed to the first glimpses into variation at the molecular level. The large amounts of variation discovered, much more than could be accounted for by the theories at that time that relied on natural selection, helped spur interest in alternative theories. In 1968, drawing heavily from his previous work with
mathematical models and evolutionary genetics, Kimura introduced the neutral theory of molecular evolution, proposing that, at the molecular level, the majority of mutations cause no differences in the fitness of individuals. Rather than being selected for by environmental pressures, mutations are entirely random, making genetic drift, rather than natural selection, a primary factor in evolution.

Kimura’s theory was immediately controversial. As molecular biology and subsequently molecular evolution gained prominence, tensions grew between molecular biologists and organismal biologists, the traditional domain of evolution. These tensions crystallized around Kimura’s theory, with many molecular biologists supporting it and many traditional evolutionary biologists rejecting it, although opinions were by no means that starkly divided. The neutral theory would represent the majority of Kimura’s work for the rest of his life; as techniques and knowledge of genetic study expanded, so did the scope and refinement of the neutral theory. He developed new mathematical methods for testing the theory, and worked to defend the neutral theory through scientific and popular writing. In 1973, a student of Kimura’s, Tomoko Ohta, developed a more general version of the neutral theory, the “nearly neutral theory” that would account for the high volumes of harmful mutations found in populations. Despite difficulties in testing it against alternative theories, the neutral theory has become part of the modern approach to molecular evolution.

As with many people, Kimura was complex. He had broad interests in both Eastern and Western culture, including readings that ranged from the Greek philosopher Sophocles to the science fiction of Arthur C. Clarke. While he could be generous and helpful to his friends, he could also be self-centered, demanding and dogmatic. Although he treated scientific disagreements in a cordial fashion, later in life he became increasingly concerned with his place in scientific history despite his growing recognition. He began treating scientific disagreements as personal attacks and would attack others in turn. Kimura received the Darwin Medal from the British Royal Society for the Improvement of Natural Knowledge 1992, given for “work of acknowledged distinction in the broad area of biology in which Charles Darwin worked” and he was made a Foreign Member of the Royal Society in 1993. He was married to Hiroko Kimura with whom he had one son. Kimura died in 1994 on his seventieth birthday; although he had been suffering from amyotrophic lateral sclerosis, the cause of death was from hitting his head during a fall.
Creationism

To properly understand the theory of evolution it is useful to examine the primary
disparate collection of views that go under the heading of
creationism. Creationism is first and foremost not a scientific theory. Rather, it is a re-
ligious viewpoint that is sometimes dressed up in scientific clothing; it is not a single,
ated viewpoint, but a collection of views that are sometimes at odds with each other.
The one common thread is that all versions of creationism assume some form of su-
pernatural intervention in the world (see Chapter 2). While we will describe the broad
outlines of these different viewpoints, it is beyond the scope of this book to go into the
details of the theological, sociological, historical, cultural and psychological reasons for
the various creationism beliefs or their legal history.

Creationist views run a gamut from those that deny nearly all scientific evidence
for evolution (and other scientific theories), to those that accept all of the evidence but
claim that supernatural intervention has occurred in certain instances (Table 4.4). Young
Earth Creationism begins with the premise that the origin story in the Book of Genesis is
literally true and the Earth is no more than several thousand years old; the usual figure
being 10,000. According to this position, creation occurred by supernatural means over
six short days. This position denies all of the evidence of evolution from biology, as well
as all of the evidence from geology, physics and astronomy that the Earth is 4.5 billion
years old and the universe is 13.7 billion years old. With regard to the theory of evolu-
tion, Young Earth Creationism denies the first and most crucial of the fundamental prin-
ciples, that organisms have changed over time (Table 4.1). It also explicitly denies prin-
ciples 2 and 3. The other principles are simply ignored as not relevant, once these three
are denied.

Table 4.4. Types of creationism.

<table>
<thead>
<tr>
<th>A. Types that deny any evolution, all species created as they now appear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Earth Creationism – a single, special creation that occurred only several thousand years ago; also referred to as Special Creationism</td>
</tr>
<tr>
<td>Old Earth Creationism – creation took place millions of years ago; there are several variants of this viewpoint</td>
</tr>
<tr>
<td>Gap Creationism – creation took place twice, once many years ago which was then replaced by a version of Special Creationism</td>
</tr>
<tr>
<td>Day-Age Creationism – each “day” in the creation story was millions of years long</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Types that allow evolution of some sort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Progressive Creationism – multiple creations over a very long time, with some evolution permitted within “kinds” but not between “kinds”</td>
</tr>
<tr>
<td>Theistic Evolution – evolution proceeds mostly through natural processes, but supernatural intervention occurs to guide the evolutionary process; the extent and types of intervention vary depending on viewpoint</td>
</tr>
<tr>
<td>Intelligent Design – a form of theistic evolution that requires supernatural intervention to accomplish evolutionary changes not possible through natural processes</td>
</tr>
</tbody>
</table>

Old Earth Creationism is a collection of various attempts to reconcile creationism
with the evidence of an old Earth and universe and the fossil record of changes in the
composition of life on Earth. Going by various names—Day-Age Creationism, Gap
Creationism—Old Earth Creationism starts with the assumption that a “day” as described in the creation story was actually a very long time, typically millions of years. Creation through supernatural means then happened repeatedly over those long periods. These viewpoints deny the same fundamental principles as Young Earth Creationism, but admit that the evidence seems to indicate that change was occurring. That change, though, is ascribed to strictly supernatural processes.

Next are various forms of creationism that allow for at least some evolution through natural processes. In effect, these forms of creationism accept all of the fundamental principles, with one important caveat—some of the observed evolution was due to supernatural intervention. They may also deny the details of some of the principles. Progressive Creationism, sometimes classified as a type of Old Earth Creationism, allows for evolution within a “kind” but not beyond. For example, Progressive Creationism allows that all cats (e.g., house cats, lions, tigers) are descended from a common ancestor, but denies that all carnivores (e.g., cats, dogs, bears) are related through common descent. In this viewpoint, each kind came about from a separate creation. Thus, they greatly limit the scope of principle 2.

All of the previously described types of creationism are views held by a minority of people in the United States. The most commonly held viewpoint is a form of Theistic Evolution that allows for all of evolutionary change to have occurred by natural and gradual processes. Theistic Evolution accepts all of the fundamental principles and all of the scientific evidence for an ancient world and the pattern of evolution revealed in the fossil record and the comparison of genes of living organisms. However, it asserts that at various times a supernatural process was necessary to either guide that evolution or cause changes that natural processes are incapable of causing. So, this viewpoint restricts the scope of principle 8 (contingency), by asserting that some of these contingent events were not truly random. How much supernatural guidance and intervention is supposed to have occurred differs among various proponents of this position. Because these occurrences are supernatural, they are untestable. The untestability is compounded by the fact that some of this supernatural guidance occurs through apparently natural processes.

Today the most vociferously promulgated of these positions is Intelligent Design. Intelligent Design (ID) accepts all of the evidence for evolution and all of the processes responsible for that change. However, it claims that certain changes are “irreducibly complex,” that is they could not have occurred through a gradual process or any natural process. So, it places restrictions on the scope of principle 4. Typically, these irreducibly
complex traits involve molecular processes or the origin of life itself. The problem with ID is that “irreducible complexity” is not defined except as “I personally cannot understand how this could have evolved by small steps.” However, every example of supposed irreducible complexity that has been put forward has been shown to be explainable by gradual evolution. The tendency of evolutionary change to occur through the co-option of something previously evolved for another function plays a large role in these explanations.

For example, Bacteria often use a hair-like flagellum for locomotion, the movement of which is driven by an almost machine-like structure made of several kinds of proteins. While it may seem improbable that such a structure would have evolved on its own, and thus may give the appearance of having been created from whole cloth, it is, in fact, a more complex version of another structure found in Yersinia pestis, the Bacteria which causes bubonic plague. Yersinia pestis has, instead of a flexible flagellum, a rigid, needle-like structure it uses to inject host cells with proteins. Movement of this structure is driven by some of the same proteins that drive flagella. The complex flagellum came about by altering a structure already present that served a very different function, negating the need for any supernatural explanations.

ID shows itself to not be a scientific theory in several ways. First, ID never defines irreducible complexity in a concrete enough way to be able to be tested. Second, the proponents of ID usually ignore explanations of change through gradual processes put forward by others, often by simply raising a new example. So, the theory itself is unscientific by relying on arbitrary, and thus untestable, supernatural intervention, while the proponents deny the existence of evidence that does not suit their viewpoint. Because much of the theory of evolution is about events that occurred in the distant past, nearly all forms of creationism rely on the argument that what we cannot directly observe, we cannot know. However, much of science, and even every day life, is based on coming to conclusions about things that we cannot directly observe. Someone who is at work or school all day will come home to find mail in the mailbox. Since no one sees how the mail gets there, it is possible that it is spontaneously generated at the doorstep every day. A simple observational experiment proves, however, that it is delivered by a human being. In some cases we rely on instruments to extend our ability to observe that which is not directly observable (e.g., the far side of the Moon or the microscopic).

Just as important, though, is logic and rationality that uses evidence and reconciles it with what else we know about the world (see Chapter 2). For example, when someone is murdered and there are no witnesses, the police must use indirect evidence to deduce the identity of the murderer. If the person was strangled, they would conclude that the murderer was a large or strong individual. If the body was found outside after a rainstorm, but the ground was dry underneath the body, they would conclude that the body was left there before the rain fell. Similarly, scientists are like detectives putting together the clues of the history of life on Earth.

Proponents of creationism have tried to claim that disagreements among scientists about some of the fundamental principles, or even the details of the evidence for evolution, means that all of the principles are wrong or that the entire theory is false. But scientific theories are robust. It is a strength of the scientific process that changes in the evidence can result in changes in details of a theory. While all theories are potentially
open for refutation, at some point the evidence for either a theory or parts of a theory are so strong that we treat those parts as settled. It would be highly inefficient if we had to continually reprove something that is well established – even something such as gravity or that planets orbit around the sun. James Oberg, NASA engineer, is quoted as saying “You must keep an open mind, but not so open that your brains fall out.” A scientific theory is both facts and the explanations for those facts; when told that “evolution is only a theory,” substitute the phrase “evolution is only an explanation.” When speaking about theories, saying “is only” makes no sense. The only question is whether the explanation – for whatever phenomenon is being explained – is sufficient and well substantiated.
Chapter 5
Cells

Living organisms, for all their size and complexity, are all formed of tiny discrete units called cells. These cells are the foundation of life; life itself exists only because it is possible to maintain highly ordered systems against the decay of entropy, and it is the cell which provides the wall between order and disorder. Cells come in a myriad of forms (Figure 5.1). Some of this variety is due to specialization of cells within an organism; for examples, humans contain about 210 different types of cells. Some is due to diversity among organisms; the cells of plants differ from those of animals which differ from those of fungi. The enormous array of forms among single-celled organisms accounts for this diversity as well; there are more than 250,000 species of single-celled Eukaryotes that represent some of the widest array of forms on Earth.

Figure 5.1

This chapter will focus primarily on what is common across all of these different types of cells, although the enormous range of cellular variation means that we cannot possibly explore all of its byways. Instead we will focus on what is going on inside the cell, the molecular basis of life. Because all of life on Earth most likely arose from a common ancestor (see Chapter 1), the basic molecular machinery is similar in all cells. Yet there are also important differences. Evolution has led to many variations on the basic theme and new functions have arisen in some lineages. Examining the cellular machinery is a good way of understanding biology’s unity in diversity, how much of life is variation on a central theme.

Scientists who study the biology of cells and molecules have been very successful over the past century in figuring out the pieces that make up a typical cell. This understanding has been driven primarily by improved instrumentation; microscopes have gotten better and better, first using light and then electrons, so that we have been able to observe and study smaller and smaller objects. Following World War II, the development of radioactive chemicals allowed scientists another breakthrough, the ability to trace biochemical pathways. However, this reductionist strategy can take us only so far. We now understand that the functional properties of cells cannot be explained simply by adding up all of the pieces. The cell has emergent properties that come about from how
those pieces are connected to each other, both physically and functionally. The past de-
decade or so has seen the rise of what is now termed systems biology, which is attempting
to put the pieces back together to understand the whole. Systems biology has created a
renaissance and a new, exciting dynamic in the studies of cells.

Cells as Foundational Units

As the foundation of living systems, the cell (Figure 5.2) encompasses life’s key at-
tributes: information, variation, complexity, and emergence (Table 5.1). One view of life
pertinent to cellular biology is that it consists of a set of controlled chemical reactions.
As we will see in this chapter, cells are made up of large, complex molecules, which
sustain life because they are able to control chemical reactions that would be impossible
outside of a cell. Because of this, the cell exemplifies life as a property that emerges
from non-living building blocks.

The theory of cells consists of ten fundamental principles (Table 5.1). The first
three principles are about the internal structure and function of cells, the next three are
about how cells interact with their external environment and one principle is about en-
ergy use and efficiency. The final three are about where cells and their properties come
from and provide links with the theories of genetics and evolution (see Chapters 3 and
4).

Table 5.1 The fundamental principles of the theory of cells

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Cells are highly ordered, bounded systems.</td>
</tr>
<tr>
<td>2.</td>
<td>Cells are composed of heterogeneous parts consisting of subsystems that act to localize resources and processes.</td>
</tr>
<tr>
<td>3.</td>
<td>Cells are regulated by a network of biochemical and supermolecular interactions.</td>
</tr>
<tr>
<td>4.</td>
<td>Cells interact with their external environment, including other cells.</td>
</tr>
<tr>
<td>5.</td>
<td>Cells exchange matter through boundaries consisting of semipermeable membranes.</td>
</tr>
<tr>
<td>6.</td>
<td>Cells require an external energy source, either chemical or electromagnetic.</td>
</tr>
<tr>
<td>7.</td>
<td>Cells use energy to create concentration gradients of ions and molecules.</td>
</tr>
<tr>
<td>8.</td>
<td>New cells are formed from other existing cells.</td>
</tr>
<tr>
<td>9.</td>
<td>Cells contain all of the information necessary for their own construction, operation and replication.</td>
</tr>
<tr>
<td>10.</td>
<td>The properties of cells are the result of evolution.</td>
</tr>
</tbody>
</table>

There are two ways that scientists have gained understanding of cell function. One
way is studying those functions within a living cell, commonly referred to by the Latin
term in vivo. The other way is isolating a function outside of the cell, commonly referred
to by the term in vitro, which comes from the Latin word for glass, because these stud-
ies are typically carried out in glass test tubes or beakers. Scientists have come to real-
ize that in vitro studies are insufficient for providing understanding of how cells function,
one reason being that the chemical reactions that go on in vivo are much more efficient
than those same reactions carried out in vitro. This increased efficiency occurs because
within the natural environment of the cell, the molecules are in a precise orientation to
enhance the reactions, conditions that are difficult or impossible to recreate artificially.

How is it that cells maintain themselves? That question can be answered in several
different ways. Much of this chapter will focus on the machinery that makes up a cell,
what the parts are in a broad sense, and how they operate. To maintain itself, a cell has
to maintain its structure and take in energy and nutrients. Among Earth’s many species,
these tasks are accomplish in a wide variety of ways. We can understand this variety by considering a few simple key differences among the many kinds of cells. First, is the cell an entire organism that is fully self-contained, or is it just one unit of a larger organism and not able to exist independently? The latter type live as part of a multicellular organism, while the former type may exist in an aggregation of cells or as a separate individual. Single-celled individuals must be able to accomplish all of a cell’s necessary tasks, while a cell that is part of a multicellular individual may only perform some of those tasks.

The Earth’s species are divided into three broad domains, Bacteria, Archaea and Eukaryota (see Figure 4.2A), a system first proposed by Carl Woese (Box 5A). Bacteria and Archaea have much simpler cells than those of Eukaryotes (Figure 5.2). They are usually described as single-celled organisms, although as discussed in Chapter 6, some species can be considered to be multicellular. Eukaryotes can be either single-celled or multicellular.

Figure 5.2
Idealized examples of a bacterial cell, a plant cell, and an animal cell. The cells of Eukaryotes have many different types of structures, most notably a nucleus. The cells of Bacteria and Archaea are much simpler. (Created by Mariana Ruiz, Source: Wikipedia)

When we describe cells and their structure and function in this chapter, keep in mind that we are talking about a mythical “average” cell and that no single type of cell does everything. In Chapter 6, when we investigate whole organisms, we will return to many of these same issues of maintenance, but focusing much more on the relationship of the organism with its external world. This chapter looks mostly inward.
Box 5A
Carl Richard Woese

A key component to the science of biology is systematics: the organization of the “tree of life” showing relationships between and the paths of descent of every living creature, from the earliest and simplest single-celled organisms to the most complex variations of life, including humans (see Figure 4.2). Efforts in systematics go back as early as Aristotle (Box 2A), although the modern structure is generally credited to Carolus Linneus, who set the framework that we use today. However, in 1976, it was all turned on its head by a single man: Carl Woese. From the nineteenth century on, the very base of the tree of life was divided into two sections: Prokaryotes, or organisms whose cells do not have nuclei, and Eukaryotes, organisms who do carry nuclei. Woese discovered a third basic classification of life, Archaea.

Born in Syracuse, New York, on July 15, 1928, Woese grew up in the Depression and World War II. After graduating from Amherst College with a degree in physics, he earned a Ph.D. in biophysics (the term first used to describe molecular biology) from Yale University. He stumbled into the microbial world as a post-doctoral scholar at Yale in the 1950s, where he investigated the development of ribosomes, the cell’s protein-synthesis machines. From there, he became interested in the origin of DNA, an interest which would eventually lead to a decades long study of the bacterial genetic code. This study would eventually redefine the way we look at the tiniest and strangest forms of life on Earth.

Because Woese wanted to unravel the complex evolutionary history of DNA and RNA, he needed a comprehensive phylogeny of all organisms to exist on Earth. At the time, however, most microbes were lumped into the broad category of Prokaryotes (most multicellular life, of course, had been studied since ancient Greece and Rome and by then was well-documented). By the 1930s, leading microbiologist C. B. van Niel of Stanford University’s Hopkins Marine Station had cited the classification of Bacteria as the most important unresolved issue in microbiology, but at the time it was impossible to obtain enough information on the characteristics of prokaryotes to determine their relationships. In this instance, as with so many other advances in microbiology, it was the advance of technology that drove the next wave of discovery. Since scientists had, by Woese’s time, spent decades trying unsuccessfully to classify Bacteria by size, shape and metabolism (the same way the huge variety of plants and animals had originally been classified) there was a deep-seated bias against classification of microbes; it was considered an impossible task to phylogenetically order the tiny life forms.

Woese, coming from a background in physics instead of biology, believed that “the world has deep and simple principles, and that if you look at it the right way,” you can find them. He was convinced the then-new molecular science held the key to microbial phylogeny, and used ribosomal RNAs (rRNAs), nucleic acid sequences found in the ribosomes, to attempt to forge relationship ties between Prokaryotes. The advantage of this method was in the fact that rRNA is one of the most conserved elements in all organisms; ribosomes are abundant in cells, and their RNA serves as a comprehensive record of life’s evolutionary history. In order to map out this history, Woese used a tedious, labor-intensive technique known as oligo-nucleotide cataloguing, wherein an rRNA molecule is cut at every guanine (G) residue, and the pieces are then broken into...
subfragments with enzymes that sliced at different residues, allowing the original sequence to be reconstructed. These tiny fragments appeared as fuzzy spots on film, and each one had to be physically examined for similarities.

Woese worked this way for a decade, and eventually completed the sequences of 60 types of Bacteria, arranging them by genetic similarity. The work was incredibly tedious, both determining the sequences and comparing them. Today, using modern sequencing technology and computers, the entire process could be accomplished in less than a day. Woese was willing to work alone, staring for hours at strips of film hanging the length of his office.

His work brought him modest $50,000 grants from NASA while he taught molecular biology at the University of Illinois, where he had been hired in 1964. He published phylogenies of chloroplasts and mitochondria, and made the surprising discovery that anaerobic bacteroids (Bacteria that don’t process oxygen the way most life on Earth does) and one group of aerobic Bacteria were closely related. Up until that time, a primary criteria for grouping Bacteria was their form of metabolism, so aerobic and anaerobic Bacteria were placed in completely separate groups.

His biggest breakthrough came in 1976 when Ralph Wolfe, a close colleague of Woese, suggested that he study a strange group of Bacteria that produced methane as a byproduct. At the time, Wolfe was the world expert on these odd Bacteria, known as methanogens. Although these methanogens showed a diverse range of physical forms, they all shared the same metabolism. When he began mapping their genetic sequences, Woese made a startling discovery: genetically, methanogens did not register as Bacteria at all. They bore no relation to either Prokaryotes or Eukaryotes, but stood apart as a third branch of life, which Woese called archaebacteria – the name was later shortened to simply Archaea. His findings were published in the Proceedings of the National Academy of Sciences (PNAS), with Ralph Wolfe as co-author. This was done to help add weight to a radical discovery; although Woese was considered something of a recluse and eccentric, shunning scientific conferences as places for politicking and social advancement, Wolfe was already a well-regarded figure in his field and a member of the U.S. National Academy of Sciences.

Despite Wolfe’s advocacy of Woese’s findings, most scientists were skeptical. The tiny snippets of rRNA Woese worked with were considered too fragmentary to serve as useful sources of information, especially to overturn a major concept in science like the basic division of life into Prokaryotes and Eukaryotes. For the most part, these doubts were not done in scientific publications, which disheartened Woese. Instead, other scientists spoke directly to Wolfe, telling him that he would ruin his career by associating himself with such a crazy idea. The situation was not helped by the fact that Woese had a retiring nature and a great dislike for professional meetings. As a result, he had few opportunities to try to persuade his colleagues that his hypothesis was correct.

Slowly, however, the tide of scientific opinion began to change. The first major turning point came when an influential microbiologist in Germany, Otto Kandler, supporting Woese’s work with his own studies of the structure of cell walls. His analyses also pointed to basic differences between archaebacteria and single-celled Prokaryotes and Eukaryotes. By 1980, the American scientific community had begun to follow suit, with other types of Archaea being discovered, such as salt-loving halophiles and the strange and fascinating thermoacidophiles, sulfur-metabolizing methanogens found in the most seemingly life-in hospitable places on Earth. Even so, many resisted his ideas. As late as 1986, Bergey’s Manual, the definitive compendium of microbiology, showed Archaea as a subgroup within the kingdom Prokayotae. Today, however, the notion of three major domains is firmly established and providing new insights into the origins of life and the
rise of Eukaryotes (see main text).

Woese remained bitter over this resistance and the lack of recognition at the time by the leading microbiologists in the United States. Throughout the 1980s, he continued to publish major studies on the aspects of the three domains of life, with his work appearing in leading journals such as Science and PNAS. Eventually, he was recognized with the John D. and Catherine T. MacArthur Award, the Waksman Award from the National Academy of Sciences of the United States, the Crafoord Prize from the Swedish Royal Academy, and in 1990 won the Leeuwenhoek Medal, microbiology’s highest honor. And, remaining true to form, when he received invitations to give lectures at universities and conferences, Woese declined most of them. Although he was familiar with the work by philosophers on the important role of social activities in the process of science, he continued to shun such activities. Woese died on December 30, 2012.
Cells as Bounded Systems

A cell is defined by its boundary, a lipid membrane, which allows a cell to maintain an internal environment that differs from its external environment, concentrating and organizing chemical processes (Figure 1.2). If, again, we regard life as a series of controlled chemical reactions, then that control happens because the chemical reactions are kept together by that boundary (Table 5.1, principle 1).

A membrane sets the stage for two types of feedback systems, those internal to the cell and those between the cell and its exterior. The internal feedback system keeps the chemical reactions in balance. A cell is a dynamic entity, in which parts are continually being assembled, disassembled, and re-assembled for other functions. It is critical that the assembly of new structures keeps pace with the breaking down of old, or the cell will be left without parts necessary for life.

A cell also needs to maintain itself in the face of a changing external environment. Cells do so because their membranes are selectively permeable; they allow some materials in and exclude others. Because of this, the cell is not at equilibrium with its surroundings, and energy is required to continually move molecules into and out of the cell so as to maintain it in that state. If a cell is part of a multicellular organism, the membrane is a critical component of the signal system among the cells. All of these membrane properties will be described in more detail later in this chapter.

Cells as Heterogeneous Systems

Physical structure

Cells are highly structured in ways that increase the efficiency of cellular functions (Table 5.1, principle 2). That structure exists in two senses. First is the physical structure, the ways in which the molecules are arranged. Second is the interaction structure, the ways in which those molecules react with each other. While we can think of these aspects of structure as separate, they are mutually reinforcing.

If life is a set of organized chemical reactions, then a central function of cells is to enable those reactions. Chemical reactions consist of the breaking of chemical bonds and the formation of new bonds, a process which requires energy. This is because the chemical bonds of biological molecules are usually stable – the bonds tend to remain in place when they exist and tend not to form spontaneously – and organisms must be able to overcome this stability.

Consider the following example. A molecule of table sugar (sucrose) consists of two subunits, a molecule of glucose and a molecule of fructose (Figure 5.4).

It takes 6.6 kilocalories (kcal) of energy to break the chemical bond joining those two subunits. If you dissolve a tablespoon of sucrose in a glass of water, nearly all of it remains as sucrose. Only a few molecules manage to break the bond joining the subunits. Some of the sucrose molecules break down because they have a certain amount of energy from the heat of the solution, that comes about from the molecules always bumping around and creating friction. Those molecules also vary in the amount of energy they have, as some will be moving faster than others.

Figure 5.4
A molecule of sucrose consists of two subunits, glucose and fructose.
By chance alone, if some of those molecules are moving fast enough and happen to bump into each other in the right way, the bond on one of them will break. But the odds of this happening are very small. At room temperature, few of the sucrose molecules will break down in this way.

**Enzymes**

All animal species have the capacity to use sucrose, break it down into smaller pieces and use the resulting molecules and energy. To do this, our cells can enhance the chemical reaction that breaks the bond between the component glucose and fructose molecules. That enhancement is done by helper molecules, proteins called enzymes. Enzymes speed up chemical reactions by reducing the amount of energy needed to make or break a chemical bond (Figure 5.5). This happens because of the properties of the enzymes’ structure. Enzymes are usually much larger than the molecules that they are helping and the enzyme has a particular niche or pocket within which the molecule sits. Once in that niche, the enzyme can enhance the chemical reaction in four ways. (1) The molecule might be bent in a precise way, thereby straining its shape and breaking the bond. (2) Two molecules might be brought together in a way that makes bond formation highly likely. (3) The chemical environment immediately around the molecule(s) will be different making bond formation or breakage more likely. (4) The enzyme might provide an alternative chemical pathway for the reaction that requires less energy. For example, the formation of a bond between two molecules might occur by first creating a bond between one of those molecules and the enzyme, then by breaking that bond and using the energy released to form a bond with the other molecule. The enzyme itself is unchanged, thus being available to repeat the reaction many times.

Enzymes are known to enhance over 4000 different chemical reactions, each one due to a different enzyme. They come in many different shapes and sizes since it is the precise three-dimensional structure that gives an enzyme its specificity. An enzyme must be flexible because it is often the bending of the enzyme that makes the chemical reaction happen. Sometimes the enzyme may be made of two or more subunits, separate proteins that are bound together, or it may require other smaller molecules for proper functioning. All of this complexity of enzyme structure allows for many ways that the chemical reactions can be regulated. Finally, the chemical reaction may require additional energy that comes from another molecule acting as an energy carrier. We will discuss both those energy carriers and cellular regulation later in this chapter.

**Spatial structures**

Although enzymes enhance chemical reactions, there is still the necessity that all of the pieces of that reaction be assembled in one place. If a chemical bond in a molecule is being broken, the molecule must be brought together with the enzyme in the right way; the same is true when chemical bonds are formed between two separate molecules. If a cell was like a beaker filled with a watery solution, enzymatic reactions would
still be very slow. Even worse, in most cells the number of any one kind of enzyme and the molecules that they are reacting with are usually small, so the chances of them randomly encountering each other are correspondingly small. In addition, a single chemical reaction is almost always insufficient to achieve a meaningful result. Almost all reactions are part of a chain in which the final result is either the creation of a complex structure or the reduction of a complex structure to smaller molecules. Yet, cells manage to carry out these chemical reactions.

The answer to this conundrum is that the cell consists of a complex system of membranes that organize the chemical reactions. The membranes are made of lipids arranged to form sheets, tubes and spheres (Figure 5.3). These lipid membranes are similar in form to those that make up the surface of all cells. Consider the formation of sucrose by the creation of a bond between glucose and fructose. Unlike the breaking of that bond, the creation of the bond happens through a series of six different chemical reactions, and it requires even more energy (21.9 kcal) than to break the bond. The formation of sucrose happens a lot in some species, notably the species of grass called sugar cane, from which we get most of our table sugar. The cells in sugar cane are able to carry out the synthesis of sucrose in an efficient manner because the enzymes involved in this process are lined up on membranes in such a way that the glucose and fructose molecules are passed along so that they have a high probability of being brought together properly, almost like an assembly line in a factory. Unlike a human-built assembly line, which is consciously designed to maximize efficiency, cellular ones are the product of evolution. Because they were not designed at the outset, they can be less efficient than they might be, and inevitably some of the molecules do not make it all the way through. However, in general these are very efficient processes that create localized concentrations of molecules that greatly enhance the reactions.

All types of cells contain a variety of large molecules, often called macromolecules, to perform many different functions. Even the simpler cells have such molecules arranged as structured systems, but this is even more true of Eukaryotic cells that are highly structured (Figure 5.3). At first they can seem like a jumble of disparate parts, each with its own name and specialized function, but despite their number, the parts carry out just a few types of functions, and it is really a just single, highly organized system.

Manufacturing processes, for example, involve many different systems. DNA and RNA are made in the nucleus. Proteins are made on ribosomes that typically are attached to a membrane called the endoplasmic reticulum (ER), which gives it a rough appearance (Bacteria and Archaea also have ribosomes, although not ER). Lipids are made on smooth-looking ER. The Golgi apparatus receives molecules made by the ER, and may make further modifications to them as well. Golgi apparatus may also package materials into small membrane spheres, vesicles, that move to the surface of the cell, merge with the surface membrane, and export them from the cell (Figure 5.6).

The breakdown of molecules in a cell is often isolated into specialized compartments. This occurs for two reasons: first, to ensure that the enzymes performing these functions do not also break down cellular structures, and second because some of the breakdown products are harmful. [One particularly harmful breakdown product is hydrogen peroxide (H2O2) and reactions that create this product are isolated in peroxisomes.] This breakdown occurs in lysosomes in animals and vacuoles in plants (which are also used for storage, especially of water).

In most organisms, energy is processed in two specialized structures. Chloroplasts convert light energy into chemical energy. Mitochondria convert chemical energy in food into other forms that the cell can use. Some Bacteria have neither structure,
Figure 5.6
Linkages among the membrane systems in the Secretory pathway diagram, including nucleus, endoplasmic reticulum and Golgi apparatus. Movement of materials into and out of the cell.
1. Nuclear membrane
2. Nuclear pore
3. Rough endoplasmic reticulum (rER)
4. Smooth endoplasmic reticulum (sER)
5. Ribosome attached to rER
6. Macromolecules
7. Transport vesicles
8. Golgi apparatus
9. Cis face of Golgi apparatus
10. Trans face of Golgi apparatus
11. Cisternae of Golgi apparatus
(Created by Magnus Manske, Source: Wikipedia)

instead using specialized membrane systems for energy metabolism. We discuss both chloroplasts and mitochondria in detail later in this chapter.

Similar to mammals and their skeletal and muscular systems, cells contain structures that help provide support and give it form, that allow it to move, and that enhance communication among cells. The **cytoskeleton** consists of a meshwork of fine fibers that extends through the entire cell, providing structural support as well as performing other functions. These fibers, which can be thick or thin and are composed of globular and fibrous proteins, provide a place for the other cellular structures to anchor and act as tracking mechanisms for their movements. Recent research suggests that some fibers may mechanically transmit signals from the cell surface to its interior. They are responsible for cell division, particularly guiding the movement of chromosomes (see Figure 3.4).

Cells actively move two ways: crawling and swimming. Crawling is done by building up fibers in the direction of movement and breaking them down at the opposite end. The same process is responsible for changes in cell shape.

Swimming is done by the beating of fibers that stick out from the surface of cells. A cell may have many short fibers, **cilia**, or just a few or even one long fiber, **flagella** (Figure 5.7). Although they differ in length, both have a similar structure, and are found in Bacteria and Archaea as well as Eukaryotes.

In multicellular organisms, cilia are used for the opposite function, moving materials outside a cell while the cell remains stationary. Some of the cells that line your windpipe have cilia that push out foreign substances that you have inhaled.

The outside of cells can be rigid or flexible, although all cells are surrounded by a lipid membrane, which itself is flexible. Rigidity is achieved by a cell wall, which is composed of various types of polysaccharides. All plant cells, for example, have walls made of cellulose (which is what wood is made of). Fungi have cell walls made of chitin, the same material that makes up the shells of lobsters and insects’ exoskeletons. Bacteria and Archaea have cell walls made of a variety of types of polysaccharides. Many animal cells excrete a sticky layer composed of molecules that combine sugars and proteins. This layer can provide support and protection, and also acts to hold cells together. Cells may also be held together by specialized junctions that include cytoskeleton fibers that extend into the cell. In multicellular organisms, adjacent cells may have channels that allow for the flow of materials and for communication.

The result of all of this is that cells are highly structured. Rather than watery bags of chemicals, they are viscous solutions, more Jell-o™ than Kool-aid™. The efficient functioning of a cell is, thus, an emergent property of this viscosity and structure. Although the existence of most of these structures has been known for over a century,
scientists are just beginning to understand how they work. Understanding the chemistry of a low-concentration, viscous medium such as the inside of a cell is one of the current

**Cells as Regulated Networks**

The complex physical structures just described exist to support the intricate network of chemical reactions that form the basis for living systems. Inside every cell is a vast network of molecular reactions, each consisting of many steps carried out by a multitude of enzymes (Table 5.1, principle 3). Previously, we described just a single reaction, the linking of fructose and glucose to form sucrose starting with separate sugar molecules, each of which goes through a different chemical reaction before the final linking reaction. An important feature of these networks is that a few highly connected enzymes act as key nodes in the entire network. Scientists are rapidly gaining a lot of knowledge about the structure of the networks that exist in various organisms, mapping out the thousands of connections. However, they are still very poorly understood. The network structure in yeast (Figure 5.8) is probably the best known of any Eukaryote because it is a single-celled organism that is easy to grow in the laboratory. But even for this network, we still know very little about its dynamic properties, how the interaction structure determines how the individual enzymes function together.

In order for life to exist, the component chemical reactions must be controlled. For example, many Bacteria can use different sugars as carbon sources. However, if the only sugar available is glucose, a bacterial cell would waste resources if it also produced the enzymes need to digest fructose. Thus, it is important that the enzymes it creates remain in balance with the molecules available to it as raw materials. Many different types of controls are used to manage this complex dance.

First, control can occur at many different levels. Chemical reactions are controlled by enzymes that are produced through the process of transcription and translation (see Figure 3.14). The amount of enzyme can be controlled by regulating the amount of RNA produced, by regulating how quickly those transcripts are translated into proteins, or by regulating how quickly those enzymes are broken back down into amino acids. Besides regulating enzyme availability, control can also be achieved by regulating the efficiency of the enzyme. Enzymes speed up chemical reactions, but that effect is not necessarily all or nothing. As mentioned previously, enzymes sometimes exist as complexes of several molecules, with others acting in ways to make the enzymes more efficient.
By adding or removing those helper molecules, the rate of the chemical reaction can be varied.

Finally, networks can be controlled by regulating the amount of substrate, or raw material, available. In the same way that a chain is only as strong as its weakest link, a series of chemical reactions that depend on the product of the previous step can proceed only as fast as the slowest step. But what happens to all of the substrate that is piling up behind that rate limiting step, like the shoppers in line at a checkout counter? Often there are negative feedback controls; as the substrate builds up, it acts as a signal to slow down production. This signal can act at any of the control points previously mentioned. In a similar fashion, a substrate can act in a positive fashion. Presenting the bacterium E. coli with glucose triggers the production of the enzymes necessary to break down that molecule.

Just like people in a store finding another counter with a shorter line, chemical reaction networks consist of many divergence points where substrate can be shunted from one series of reactions to another. The tendency of the substrate to be used in one series or another can be controlled through either a push or a pull system; halting or slowing down the movement of a substrate along one pathway will push the remaining substrate to another pathway. Conversely, speeding up the processing of a substrate along a pathway will tend to pull more substrate along. At one of those divergence points, the enzymes responsible for movement towards one or another pathway can be thought of as competing for the substrate. The enzymes may differ in their affinities for the substrates, and that affinity can be regulated by helper molecules as mentioned previously.

The membrane system itself can also be considered part of the interaction network – it is dynamic, with parts continually budding off and merging. For example, the endoplasmic reticulum transfers materials to the Golgi apparatus by creating small membrane spheres. The rate at which that occurs, the rate at which substrate gets moved around inside the cell, or the rate that RNA molecules get moved from the DNA to the ribosomes can also act as network regulators.

### External Interactions

The network of interactions can also extend beyond the bounds of the cell itself (Table 5.1, principle 4) to other cells (for multicellular organisms) and to the external environment (for all organisms). In some instances those interactions are direct: cells that are next to each other, cells in multicellular organisms that are in direct contact with the environment, or for all single-celled organisms when interacting with their environment. Those same interactions, however, may be indirect as well.
In multicellular organisms a cell may produce a signal which must get passed along, for example through the blood, before it gets acted upon by another cell. Those cell-cell interactions act to regulate processes at the level of the whole organism, in the same way that interactions among components regulate those functions within cells. One outcome of those interactions is the specialization of cells for particular functions. We examine organism function and cell specialization in detail in Chapter 6. Because such signals are often molecules at very low concentrations, understanding their functions and actions required the development of sensitive detection techniques (Box 5B).
Box 5B
Rosalyn Sussman Yalow

Rosalyn Yalow made her career at the cutting edge – as a woman and a nuclear physicist, she carved inroads into early nuclear medicine and served in the vanguard of her gender’s entry into the atomic age. As a female scientist in the middle of the 20th century, she faced blatant discrimination, yet went on to receive the Nobel Prize. Even in her own family she forged an unprecedented path; neither her father, a packaging materials wholesaler, nor her mother, a German immigrant and a homemaker, attended high school or college. Her research demonstrates how scientific breakthroughs can come about through the invention of new tools, which often come from someone working across disciplinary boundaries.

Yalow, born on July 19, 1921 in the South Bronx as Rosalyn Sussman, showed an early interest in science. As her parents owned no books, she and her brother Alexander made weekly trips to the public library, and by the age of eight Rosalyn knew she wanted to become a scientist. She gravitated to the logic of science and loved its capacity to explain the natural world.

Yalow was the product of New York City schools, attending Walton High School in the Bronx (one of two Walton graduates to receive the Nobel Prize), where a teacher sparked an interest in chemistry. She went on to attend Hunter College. After reading a biography of Marie Curie and attending Enrico Fermi’s 1939 Columbia University colloquium on nuclear fission, she shifted her studies to nuclear physics. To her, the field seemed to be the most exciting one in the world, with each major experiment garnering a Nobel Prize. At the age of 19 she graduated magna cum laude as Hunter College’s first physics major.

Despite her stellar academic record, Yalow immediately ran into obstacles upon embarking on her graduate studies. She applied for an assistantship at Perdue University in Illinois and was rejected, the University writing to her professor that “She is from New York. She is Jewish. She is a woman. If you can guarantee her a job afterward, we’ll give her an assistantship.” Since a guarantee was not possible, the rejection stood. Yalow ended up having to take a position as a secretary at the College of Physicians and Surgeons at Columbia University. Undeterred despite assertions that a woman would be unable to get into graduate school in physics, she attended classes at Columbia and continued to pursue a graduate education.

Her opportunity came in the form of an offer for a teaching assistantship at the College of Engineering at the University of Illinois at Champaign-Urbana. While at Illinois she met and married Aaron Yalow, a fellow graduate student, eventually having two children, Benjamin and Elanna,. The U.S. had entered World War II, and the draft meant that universities were faced with filling their graduate programs with women or shutting them down entirely. Yalow was able to seize this entry into science and academia. She was the first female graduate student in engineering in 24 years, and as the only woman among 400 peers, she faced an incredible amount of pressure to perform. This was compounded by the fact that Hunter’s late adoption of a physics curriculum meant that she was behind her classmates academically; to make up for it, she took two undergraduate courses as well as her three graduate courses. She excelled in her studies, earning
was behind her classmates academically; to make up for it, she took two undergraduate courses as well as her three graduate courses. She excelled in her studies, earning straight A’s – with the exception of the laboratory component of her Optics course, in which she received an A-minus. The chairman of the physics department pointed to her grade as proof that women couldn’t excel in lab work.

Despite the physics chairman’s assertions, Yalow became proficient in both the construction and operation of apparatus for measuring radioactive substances while studying for her Ph.D. under Maurice Goldhaber. She received her doctorate in nuclear physics in 1945 and returned to New York, where she worked as a researcher in the Federal Telecommunications Laboratory – again the only woman among her peers – until the research group left New York the following year. Unable to secure another research position, she ended up teaching physics at Hunter College. Still wanting to pursue a research career, she volunteered to work in Columbia University’s medical laboratory, where she was introduced to the then-brand-new field of radiotherapy. She continued to pursue nuclear medical research over the next three years, first in a part-time position at the Bronx Veterans Administration Hospital, then later becoming a full-time researcher there.

In 1950 she began her 22-year-long collaboration with Dr. Solomon Berson. At the time, Berson was accorded more gravitas than Yalow; due to his medical degree and gender, he had more contacts with scientific journals, professional societies and within academia. Yalow, however, was absolutely single-minded in the pursuit of her research, living in a modest house just a mile from the hospital and eschewing hobbies and leisure travel. At the beginning of their partnership, Yalow and Berson focused on using radioactive isotopes to more accurately estimate blood volume, their first major contribution the development of a technique to measure iodine filtered by the thyroid and kidneys in a set amount of time. From this beginning, they widened their scope to include research into other peptide hormones, in particular insulin. This thrust had personal significance to Yalow, whose husband was diabetic.

Peptide hormones are present in much lower concentrations than other types of hormones, and correspondingly more difficult to track in the body. Although chemical methods of analysis were available for those other hormones, there were no such techniques available for peptide hormones in the 1950’s. Yalow and Berson’s research changed that. Key was the development of a new technique, radioimmunoassay (RIA) that came from Yallow’s training as a nuclear physicist.

Yalow discovered that people who had received insulin injections developed an antibody response to the insulin. This discovery, coming in 1956, challenged the current scientific understanding of insulin – few believed that so small a molecule could trigger an antibody response – but was key to Yalow’s research. She and Berson used a radioactive isotope to tag a known quantity of insulin, then mixed the newly-tagged insulin and its antibody with a blood sample in which the levels of the hormone were unknown. The antibodies would abandon the tagged molecules to attack the naturally-occurring molecules in the blood sample, and by counting the number of tagged molecules without antibodies, they were able to ascertain the concentration of insulin in the blood sample. Radioimmunoassay revolutionized the field of endocrinology, and decades later would garner Yalow a Nobel Prize.

Using RIA, Yalow delved into hormonal research, publishing a series of papers with Berson beginning in 1956 elucidating not only the basic nature of peptide hormones such as insulin, but also the mechanics of diseases caused by hormonal abnormalities. She demonstrated that it was the binding of insulin antibodies leading to abnormal degradation of the hormone that was at the root cause of diabetes, rather than the pancreas
secreting too little insulin as was previously thought. Today we know that such autoimmunity reactions are the source of more than 80 human illnesses, including type I diabetes, rheumatoid arthritis, and lupus.

RIA also offered a life-altering solution to newborn babies afflicted with underactive thyroid. This condition can cause mental retardation, but symptoms only show up at three months of age after the damage has been done. By allowing a newborn’s thyroid levels to be checked immediately, the afflicted baby can be treated immediately and the damage prevented. In addition, they adapted RIA to track vitamins such as B12 as well as viruses, which allowed donated blood to be tested for hepatitis B.

Despite her achievements, Yalow still faced obstacles. As her work challenged the current scientific edifice, she faced disbelief and outright rejection in trying to publish her results. She had to delete her explanation of insulin antibodies before the Journal of Clinical Investigation would accept her paper on RIA – Yalow saved the rejection letter and later included it in her Nobel lecture. Despite the huge commercial potential for the technique, Yalow and Berson refused to patent it, instead making every effort to get RIA into common use in hospitals and laboratories.

Yalow went on to become the acting chief of the RIA Reference Laboratory at the Bronx Veterans Administration Hospital in 1968, and then chief from 1969 to 1992. She was elected to the National Academy of Sciences in 1975 and received the Albert Lasker Medical Research Award in 1976. The following year she was awarded the Nobel Prize in Physiology or Medicine, only the second woman to receive the award. In 1982, when speaking to a group of schoolchildren about the challenges and opportunities of a life dedicated to science, she said, “Initially, new ideas are rejected. Later they become dogma, if you’re right. And if you’re really lucky, you can publish your rejections as part of your Nobel presentation.”

Because they are not awarded posthumously, Berson, who died in 1972, did not share in the 1977 Nobel Prize. However, at his death Yalow insisted the VA rename the lab where she and Berson had worked in his honor, so that his name would continue to appear on her research papers. Yalow died on May 30, 2011, at the time the senior medical investigator emeritus at the Bronx Veterans Administration Hospital and (quite appropriately) the Solomon A. Berson distinguished professor-at-large at the Mount Sinai School of Medicine.
The cell membrane is the gateway that regulates cellular interactions, that gateway can act in several ways (Figure 5.9). Key to these interactions are the complex structure of the membrane. A cell membrane is not a uniform surface. Rather, it is full of holes, called pores, that are made of specialized proteins embedded in the membrane. The first type of interaction is passive diffusion. Although the pores in the lipid membrane will block the passage of large molecules, single ions can pass through. In passive diffusion the cell does not regulate this movement; instead, the rate is determined by the relative concentrations inside and outside the cell. For example, if the concentration of sodium ions (Na+) is higher inside than outside, on average more ions will be passing through the membrane from the inside out than from the outside in.

Although some movement of sodium ions occurs by passive diffusion, much happens through active transport, the second, very important method of regulation of external interactions. For these interactions, the pores actively move ions or molecules from one side of the membrane to another, using energy to do so. Each pore is specialized for the movement of one or a few ions or molecules, which acts as a further regulatory control.

Sometimes what is moved is not material, but information. Again, this information movement is carried out by specialized proteins that extend through the cell membrane. Those proteins act as receptors for specific molecules that provide the signals. When the signal molecule reacts with the protein on the outside of the cell, the protein changes its shape. Depending on how it does this, it will either release another signal molecule on the inside of the cell, or if the protein is an enzyme, the signal molecule may change its activity.

Pores are sufficient to move single molecules, but sometimes a cell must move large amounts of material in or out. Movement of large objects into a cell is how many single-celled Eukaryotes eat. For example, an amoeba eats bacteria. It does so by surrounding the bacterial cell with its membrane. That membrane then pinches off, forming a sphere inside the cell (Figure 5.AE-A). Enzymes that are able to break down the bacterial cell are then moved into the bubble. The reverse process occurs in cells of multicellular organisms that are responsible for manufacturing substances used by other cells (Figure 5.6).

The cell membrane is dynamic not only on small scales, but also on very large scales across the entire cell. In animals, when an egg is fertilized by a sperm cell, a change occurs in the membrane that prevents any other sperm from fertilizing that egg. This process occurs in just fractions of a second, spreading outward from the point of fertilization. The cell membranes of both the egg and the sperm perform another important function in this process, containing proteins that bind to each other. Those proteins
are specific to a given species preventing fertilization of an egg by a sperm of a different species. Many people had a hand in fully researching this process, the first of whom was E. E. Just (Box 5C).
It is a sad reality that social prejudices often thwart the growth and advancement of all subjects, from the sciences to the arts. This is, unfortunately, what happened to Ernest Everett Just, a brilliant biologist and embryologist who published over 70 articles and two very well-received books, and nevertheless was unable to secure a research position in America because of his race. Born on August 14, 1883 in Charleston, South Carolina, Just was African-American, the grandson of a slave who inherited the name Just from his master. His father was an alcoholic and a womanizer, who kept a mistress despite the fact he didn’t make enough money to support his wife and child. He died when E. E. Just was four, and after his father’s death, his mother sold their house in Charleston, and moved to James Island, off the South Carolina coast.

Just’s mother took a job in a phosphate factory, an unusual occupation for a woman, but one that paid much better than traditional “women’s work.” She eventually made enough money to invest in real estate, and founded the first school on the island. She pushed her son to succeed, and at 13 Just enrolled in South Carolina State College (also known as The Colored Normal, Industrial, Agricultural and Mechanical College) where he completed the usual four-year course of work in only three. He then applied for entrance at Kimball Union Academy, a private secondary school in Meriden, New Hampshire, and promptly left on a boat bound for New York without waiting to hear if he had been accepted. He worked aboard the boat to pay for passage, then worked odd jobs in the city until he had made enough money to pay for the rest of the trip to New Hampshire, where he found that he had been accepted and awarded a scholarship reserved for “deserving” students.

At Kimball he studied the classics, and excelled at oratory and journalism, winning an oratory competition and running the student paper. At that point he planned on becoming a classical scholar, and was encouraged by his teachers, who recognized the brilliance of the young Just. After Kimball, he entered Dartmouth in the fall of 1903 at the age of 20 with the intention of continuing to study the classics. While there, he displayed a talent for writing, some of his short stories and poetry being published in various Dartmouth publications. It was here that his interests started gravitating toward biology.

Just was drawn to William Patten, a distinguished paleontologist with a strong effect on Dartmouth’s curriculum, who eventually organized a course on evolution that was required for freshmen. Just was also mentored by J. H. Gerould, who was known for his genetic studies of butterflies.

He graduated in 1907, earning his diploma magna cum laude. It was then that he ran into the first big obstacle of his career: because of his race, he only had two employment opportunities; Howard or Moorehouse College. He went with Howard College, where he was initially assigned to the English department. He was popular with students, and recognized for his teaching skills; he received the Spingarn Medal of the NAACP in 1915.

When he began teaching biology in 1909, his interests again shifted from the classical to the scientific, and he sought Patten’s advice on doing graduate work in zoology. Although Pattern advised him that medicine was a better direction, he nevertheless
recommended Just to Frank R. Lillie, the head of the zoology department at Chicago University. Lillie accepted Just as his assistant at Woods Hole in the summer of 1909, where Just quickly earned a reputation as an excellent scientist, working closely with Lillie, sharing a great mutual respect. Just quickly became an expert collector of sea invertebrates as well as a skilled microscopist, and almost immediately began publishing highly regarded papers, his first article reporting that the first cleavage plane of the *Nereis* sea worm was determined by where the sperm enters the egg. He was soon sought-after for advice.

At Woods Hole Just soon developed several close friendships; he and A. H. Sturtevant would often eat together, and he spent time with geneticists Donald and Rebecca Lancefield, cytologists Franz Schrader and Sally Hughes (Schrader). With Sally he found an outlet for his literary passions as well as his scientific ones, and the two would often discuss poetry, literature and music, especially the work of D. H. Lawrence, who was considered quite scandalous at the time. Socially, Just was quite popular at Woods Hole; he was handsome, intelligent and personable, with a wide range of interests.

After his first paper, he continued work on *Nereis*, charting the life history, especially the breeding habits, of the worm. He developed a technique to mark the entry point of the sperm into the egg using tiny particles of India ink, and used his experience with *Nereis* to study the fertilization process of other marine organisms, including *Platyneres*, *Echinarchnius*, and *Arbacia*. All of his work was characterized by careful observations, care taken in the living conditions of the animals, and meticulous attention to experimental details. His *Echinarchnius* studies verified that eggs become impermeable once a sperm enters; Just mapped two events triggered by its entry - first the nucleus releases a substance making the egg fertilizable, then the sperm enters. In addition to his research work, Just also published articles on techniques of collecting and experimental methods, culminating in a book based on his work at Woods Hole.

In the fall of 1911 he enrolled at the University of Chicago as Lillie’s graduate student, and received his Ph. D. on June 6, 1916. Despite this, and despite his glowing reputation, the only two places he could find employment remained Howard and Moore-house. Just had by that time proved himself to be a superb technician as well as an extremely careful worker; he set high experimental standards for himself, and was often critical of the methods of others when they fell short of his own expectations. He was open with his criticism and freely disagreed with others when their observations were different from his own. This would prove a serious detriment to his career.

The most notable disagreement took place with Jacque Loeb, who, ironically enough, had recommended Just for the Springarn Award. Just thought Loeb’s work was flawed, and openly said as much. Loeb had argued that the development of an egg was initiated by two steps; cytolysis (which could be induced in the lab by butyric acid), then a quenching with hypertonic seawater. Just showed that the seawater alone was sufficient, and thought that Loeb was simply careless with the details of his experiments. This, along with other criticisms, lead to a major controversy in the scientific community, with embryologists taking sides. Loeb’s public opinion of Just plummeted; unfortunately, Loeb was a liberal deeply interested in social causes, with a special interest in Howard College. When Just was being considered for a position at the Rockefeller Institute for Medical Research, naturally Loeb’s advice was sought. The response Loeb gave probably cost Just his only opportunity for a research position he was likely to encounter as an African American man: “...the man is limited in intelligence, ignorant, incompetent and conceited; in fact his research work is not only bad but a nuisance.” Just was thus forced to remain at Howard.

Although Just continued to spend his summers at Woods Hole and eventually got
a grant to spend half his time at Howard in research, he felt buried under his administra-
tive duties and seemingly endless committee responsibilities. Although he remained
productive, publishing around 50 articles by 1930, Howard began having major admin-
istration problems, and Just was caught up in them as his relationship with the college’s
president deteriorated. Even at Woods Hole, too much of his time was spent helping oth-
ers rather than working on his own research, and he ran into the ugly spectre of racism
when he brought his wife and children up one summer, who quickly left after encounter-
ing offensive remarks. After a visit to the Bay of Naples biological station and Europe,
Just began to desire a move across the Atlantic.

Sadly, some of his impetus to move stemmed from the racism still rampant in the
United States; his opportunities were very limited at home, and moving would afford
him the ability to hold a research position. For a man deeply interested in literature and
music, the live opera and high-quality chamber music available in Paris and Italy were a
strong draw, especially when he met and began an affair with a European woman, Hed-
wig Schnetzler, whom he eventually married. Unfortunately, his financial situation pre-
vented a move; all his money went to his current wife and children.

This did not deter his interest, however. He became friends with Reinhard Dhorn,
director of the Statione Zoologica in Naples, and felt quite at home with European scien-
tsists. Dhorn encouraged Just, both in pursuits of science and philosophy, making him his
secretary; with long lab hours and concerts in the lobby of the Statione. Just’s natural
inclination toward philosophical studies and willingness to speculate was discouraged in
America, but nurtured by European scientists like Dhorn, and Just’s later work reflect-
ed this encouragement. He was strongly drawn to the Morgan school of genetics, and
thought that chromosome theory was a great modern breakthrough, but felt that it fell
short in explaining how genetic information translates to development and phenotype.

Just turned to cytoplasm, especially ectoplasm (which has more contact with outside
agents) for his explanations (although he was always careful to label his speculations as
such). In “A single theory fr the physiology of development and genetics,” a 1936 paper
in which he attempted a synthesis of genetics and embryology, he speculated that the
egg starts out with pluripotent cytoplasm, which the chromatogosome synthesis then uses
to copy itself in cell division. The cytoplasm then has restricted potencies, with somatic
cells having more and more restricted set of properties. He later began to speculate that
genes are nucleic acid, realizing that chemical studies of nucleoproteins were of greater
and greater importance. He hoped to do “…a more exact study of nucleo-protein syn-
thesis to embrace as many different types of eggs as possible” but unfortunately passed
away before he could investigate further.

Just later spent time at Kaiser-Wilhelm-Gellschaft in Dahlem in the laboratory of
Max Hartmann, where he interacted with Richard Goldschmidt, Otto Mangold, and Jo-
hannes Holtfretner. He had earlier become interested in ectoplasm, the outer layer of
the cytoplasm in cells, and here he was able to study Amoeba, taking advantage of the
creature’s huge cell size. When he began to look in earnest for funding to finance his
move to Europe, he immediately ran into problems, with foundations, millionaires, and
all other avenues turning him down.

Some of the reluctance to fund Just’s move was concern for the plight of African-
Americans – they believed that someone like Just could do more good for the community
at Howard. Eventually he received a little funding from the Carnegie Institute, and was
given a desk at the Sorbonne (although it was a great honor, it carried no monetary re-
compense). He managed to get a European divorce in 1939, and married Hedwig; without
a family to support, he was able to settle in a small biological station in Roscoff on the
French coast overlooking the Channel. The facilities were primitive, but the Channel
waters teemed with biodiversity.

Unfortunately, by 1940 the Germans had invaded Czechoslovakia and the siege of Paris had begun. Just’s colleagues at Woods Hole worried for his safety, and eventually Just decided that he and Hedwig must flee. Problems arose with Just’s passport, and he was interned by the Nazis for a time until his release was negotiated. He was able to book passage out of Spain to New York City, but all his research done at Roscoff was lost in the confusion. Just returned to Howard in failing health, weakened and in pain, and was diagnosed with pancreatic cancer. He died on October 27, 1941, and was commemorated on a United States stamp in 1996.

Just was a very early frontrunner of the field that has come to be known as eco-devo - ecological developmental biology (officially established in 2001 as a subcategory of evolutionary developmental biology). Concerned with studying organisms as a whole, in their natural (as opposed to laboratory) conditions, the field relies on organismic, or materialistic holism, a strong characteristic of Just’s work with cytoplasm and heredity, insisting that lab conditions match closely as possible the natural conditions of an embryo.
Metabolism: Unity and Diversity

The basic chemical reactions

Living systems maintain themselves in an ordered state by using energy (Table 5.1, principle 6; see Chapter 1). The set of co-coordinated chemical reactions that provide organisms with that energy is termed metabolism. Metabolism consists of two linked processes: anabolism, the synthesis of complex substances from simpler ones, and catabolism, the breaking down of those complex substances. Anabolism both creates the structures of living systems and stores energy in organic molecules. Catabolism breaks down structures so that the components can be used for other purposes and uses that stored energy. These twin processes are occurring constantly within each cell. They also occur at a larger scale with some organisms being net creators of organic molecules and stored energy and other organisms being net consumers.

Energy begins in the form of high-energy electrons which are then harnessed for chemical reactions and stored in chemical bonds. The transfer of electrons from one substance to another is called a reduction-oxidation reaction, or redox reaction for short. The gain of an electron is called reduction, and the loss is called oxidation. Cells need to balance the negative electrical charge in the electron (e-) with a positive charge from a proton, i.e., a hydrogen ion (H+) [these reactions can also be thought of as the loss (oxidation) or gain (reduction) of a hydrogen atom].

Figure 5.10
The two parts of a reduction-oxidation reaction. (Created by Cameron Garnham, Source: Wikipedia)

Living systems use a variety of electron donors and electron acceptors as the basis of their energy systems, but all energy systems have that same basic set of paired reactions. All of the energy systems we find today arose in Bacteria and Archaea early in the history of life on Earth, and some components even pre-date life's origin. Some of those systems are also found in Eukaryotes, where again the systems have a basic similarity across all species while showing diversity in particular aspects.

Organisms can be divided based on their source of energy: whether the energy comes from inorganic sources or from consuming other organisms. Among inorganic sources, light accounts for almost all of the energy fixed by living systems on Earth. Each year approximately 100 terawatts of energy is captured by this process, about seven times more than used by the entire human civilization, converting around 100 billion metric tonnes of carbon into biomass per year.

Light energy is used by nearly all plants, as well as numerous single-celled Eukaryotes and some Bacteria. When light is the energy source, photons are absorbed by specialized molecules that release high-energy electrons. The electron donor is most commonly water (H₂O) and the electron acceptor is carbon dioxide (CO₂) with molecular oxygen (O₂) being released in the process. We look at this process, called photosynthesis, in detail below. Some Bacteria use light energy with other electron donors including hydrogen sulfide (H₂S, which produces elemental sulfur, S) and various organic molecules.

Some Bacteria can also use the chemical energy contained in a variety of inorganic substances. These alternative electron donors include methane (CH₄), hydrogen sulfide (H₂S), ammonia (NH₃), hydrogen (H₂), iron (Fe₂+), and manganese (Mn₂+) and a
Photosynthesis is the process of converting light energy into chemical energy. Chemically, the process can be summarized as: $\text{H}_2\text{O} + \text{CO}_2 = \text{O}_2 + \text{organic compounds}$. That summary, however, hides a lot of detail. The entire process can be divided into two segments (Figure 5.11). The first segment is the oxidation step that uses light energy to break $\text{H}_2\text{O}$ into $\text{H}^+$ and $\text{O}_2$ and captures the energy in two compounds: ATP and NADPH.

ATP stands for adenine triphosphate, indicating that it contains three phosphate ($\text{PO}_3^-$) units (Figure 5.ATP). The bond between the third phosphate unit and the rest of the molecule contains a lot of energy, but it is also easily broken in a way that can transfer that energy to another chemical bond. This makes ATP one of the primary energy carriers in cells. NADP+, nicotinamide adenine dinucleotide phosphate, is an electron acceptor that can be converted to NADPH, again as a high-energy carrier.

Because this segment uses light energy, it is often referred to as the light reactions. The key molecule in this step is chlorophyll, which is capable of absorbing photons. Those photons change the chemical composition of the chlorophyll, shifting it into an excited state in which it can act as a reducing agent. The energy in that reaction gets passed along through a chain of molecules, in the process of which ATP and NADPH are produced (Figure 5.12). Although it may seem unnecessarily convoluted to have the electron passed along a chain, rather than having the chlorophyll react directly with the ADP and NADP+, the process is necessary to “cool off” the energetic particle.
Figure 5.12
The molecular structures of adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADPH), two of the major carriers of energy in cells. The two molecules have related structures which may indicate a common evolutionary origin. Since a common feature of evolution is to take a structure used for one function and alter it so that it performs a related function, the fact that they are both based on a nucleic acid may indicate an origin in the pre-biotic RNA World (see Chapter 1). (Created by NEUROtiker, Source: Wikipedia)

The light absorbing system is like an antenna that consists of many molecules. Chlorophyll is at the center of the antenna, but other related molecules are part of it, with different molecules best able to absorb different wavelengths of light. Land plants appear green because these molecules absorb light across most of the visible spectrum, and reflect back mostly the green wavelengths. Other groups of photosynthetic Eukaryotes and Bacteria look red, brown, yellow, orange, cyan and even purple because they have other forms of chlorophyll and associated molecules with slightly different absorbance characteristics. All of these photosynthesis systems are based on the same general principle – the capture of light energy as excited electrons – but they carry out that process in many different ways.

The second segment is the reduction step that starts with CO$_2$ and uses the energy captured in the ATP and NADPH molecules to produce organic compounds, notably glucose. Glucose is the end point of this process because it is much more stable than ATP or NADPH, and thus suitable for both long-term storage of energy and for transportation to other parts of a multicellular organism. Because this segment does not require light energy, it is sometimes referred to as the dark reactions. It is also known as the Calvin cycle or the Calvin-Benson cycle after the people who first described its details (Box 5D). In Eukaryotes, the process of photosynthesis takes place in specialized organelles within the cell called chloroplasts (Figure 5.13). Just as with their other metabolic processes, Bacteria that carry out photosynthesis have membranes that are used to organize the molecules responsible for this process.

Because photosynthesis is the way in which many species get their energy, you might think that those species would always be carrying out photosynthesis as quickly as possible. However, this rarely happens. Just as the amount of food you can consume at any one time is limited, the rate of photosynthesis is limited by many different factors. First, more light energy does not necessarily mean more light captured. A cell contains only so many chloroplasts, which contain only so many chlorophyll molecules. The exciting of the electrons and the passing of that energy to the next molecule in the chain takes a certain amount of time, even if that time can be measured in nanoseconds. The
process also requires $H_2O$ and $CO_2$, and even if both are available in unlimited quantities, it takes time to move them into the cell and into the chloroplasts. Very often light or $H_2O$ are available in limited amounts (a plant in the shade or in a desert), and even the availability of $CO_2$ can be limiting. Those are all short-term limitations; over a longer span of time, there may be limits on the materials needed to build chloroplasts.

Figure 5.13
A. Chloroplasts are specialized organelles in the cells of photosynthetic Eukaryotes where photosynthesis is carried out. (Created by Kelvinsong, Source: Wikipedia)
Box 5D
Critical Experiment: Describing Photosynthesis

The process of working out the dark reactions of photosynthesis were carried out by a group of scientists at the University of California-Berkeley in the years following World War II. They were led by Melvin Calvin – who was awarded the Nobel Prize in Chemistry in 1961 for this work – and included Andrew Benson and James Bassham. This set of experiments highlights the importance of how the development of new scientific tools can provide new ways to answer questions. In this instance, radioactive isotopes of carbon (14C) became available as a by-product of atomic power and atomic bomb research, at the same time that chemists developed paper partition chromatography, a process which can separate individual compounds from a solution. The researchers utilized these new advances along with autoradiography, a process which can locate radioactive compounds on a paper chromatogram. Although none of these techniques were developed for the problem of describing photosynthesis, as commonly occurs in the sciences, Calvin and his team were able to recognize how those methods could be used to answer their question.

The experiments were carried out with the eukaryotic, one-celled alga Chlorella which had the advantage of growing quickly in solution. They grew them in a flattened flask that was dubbed a “lollipop” because of its shape, one which exposed the alga to as much light as possible, thus maximizing the rate of photosynthesis (Fig. 5D.A). At the start of the experiment, liquid containing 14CO₂ was injected into the lollipop. Then, after a set amount of time, a sample was rapidly placed into boiling ethanol. The boiling ethanol quickly killed the alga, halting photosynthesis.

Critically, the ethanol also extracted the various molecules that were part of the dark reaction. The ethanol, along with the extracted molecules, was then placed on the paper chromatogram. Different molecules would move at different rates along the paper depending on their chemical structure, thereby separating them. The researchers could then use autoradiography to identify those that had incorporated the 14C, and perform further chemical analyses to identify the molecule and even to show which of the carbon atoms were radioactive.

The first thing that they discovered was that the reactions were very fast; after just 30 seconds many different types of molecules contained the radioactive label. Wanting to be as precise as possible – something very important in any scientific experiment – they kept repeating the experiment using shorter and shorter exposure times, until they had it down to just 2 seconds. Even then, a half dozen or more types of molecules were labeled. However one type, 3-phosphoglyceric acid (3PG) (Figure 5D.B), was always the most abundant. From this they concluded that it was the first molecule in the process.

An important clue to what was happening was based on their determination of which of the carbon atoms were radioactive. Most of the labeled atoms were part of the COO group on the end, lending support to the conclusion that this was the newly incorporated 14CO₂ molecule. Sometimes the other atoms were also radioactive, and from this they deduced that the process was cyclical, that is, the 3PG was made by adding...
CO₂ to a molecule that was previously derived from 3PG. But what was this molecule? The obvious answer was some sort of two-carbon molecule to which the third was added – however, none of the radioactive molecules that showed up on the paper chromatogram had this structure. When faced with this apparent obstacle, the researchers had to draw on disciplines outside biology itself – in this case, chemistry – to figure out what their results meant. This is a very common occurrence in all scientific fields, not just biology, and this experiment is a prime example of how science rarely happens in isolation, but is extremely interconnected. One molecule that they found was a reduced form of 3PG, a reduction that requires energy, which they hypothesized came from light energy captured in the form of ATP. To prove this, they performed the experiment in the light and in the dark. As predicted, in the dark amounts of 3PG increased and very little of the reduced form were found.

Now that they knew where the 3PG was going to, the next step was to determine where it was coming from. What, exactly, was the unknown precursor molecule – which they called molecule X – and how was it getting turned into 3PG? To figure out the identity of molecule X, the Berkeley group needed another bit of information, which was what would happen if the supply of CO₂ were cut off in the presence of light. In that case, the amount of 3PG should decrease as it gets reduced, but the amount of molecule X should remain the same, or possibly even increase, because there is no CO₂ for the next step of the cycle. When they sifted through the data from their experiments, they found a molecule that fit the bill perfectly, the five-carbon ribulose bisphosphate (RuBP). Instead of a two-carbon precursor, the CO₂ was added to a five-carbon molecule producing an unstable six-carbon molecule that immediately split into two molecules of 3PG.

Once they had worked out these key steps, the rest of the cycle was deduced using similar methods. They discovered that the dark reactions consisted of a number of steps involving four-, five-, six- and seven-carbon molecules. They also discovered the key enzyme in the process, the one responsible for the reaction between RUBP and CO₂, giving it the name ribulose bisphosphate carboxylase (usually called rubisco). Later workers showed that rubisco is found exclusively in the chloroplasts, as are the other molecules and enzymes postulated to be part of the dark reactions, just as would be predicted.

The experiments of Calvin and his team demonstrate one very fruitful way of gaining scientific knowledge. They performed a series of carefully controlled experiments where they would vary one factor at a time, keeping all other factors constant, the sort of experiment advocated by Francis Bacon (see Box 2C). The experiments were guided by a careful logic based on both their broad knowledge of chemistry in general and other similar discoveries. For example, they concluded that the process was cyclical based both on their own data and knowledge that other biochemical reaction systems were also cyclical. Finally, they were able to carry out the experiment when they did because of new tools and techniques that others had developed.
**Energy usage**

In many regards, the usage of the energy stored in organic molecules simply reverses the steps that occurred when those molecules were created. Many of the basic features are similar. All energy usage involves redox reactions, just in reverse. Some parts of the process are shared by all species – again pointing to descent from a common ancestor – yet there is much diversity involving the specific details. In Eukaryotes, parts of the process take place in specialized organelles, the **mitochondria**.

![Diagram of glucose metabolism](image)

**Figure 5.14**
Glucose, a six-carbon molecule, gets broken down into pyruvate, a three-carbon molecule, releasing energy. The pyruvate may get further oxidized completely into CO$_2$, if oxygen is available, a process that in Eukaryotes occurs in the mitochondria. If oxygen is unavailable, the pyruvate may be reduced into a variety of three-carbon or two-carbon molecules. (Created by JohnyAbb, Source: Wikipedia)

Because the most common product of photosynthesis is glucose, we will focus on its usage, although the same general principles hold for the breakdown of all organic molecules. One vital, if slightly counterintuitive, aspect of the process is that it takes energy to release energy. Glucose is a six-carbon molecule, and the first steps involve breaking it into two three-carbon molecules that each have a phosphate molecule added. The phosphate comes from ATP, necessitating the use of energy to transfer the phosphate. Those three-carbon molecules are then oxidized to pyruvate, also a three-carbon molecule, in the process producing four molecules of ATP, for a net gain of two ATP, plus two molecules of NADH (similar in structure and function to NADPH). At this point there are a variety of fates for those pyruvate molecules (Figure 5.14).

The first consideration is whether oxygen is available. If it is not, the pyruvate may be involved in a reduction reaction that produces a two-carbon molecule of ethanol plus one molecule of CO$_2$. This is the process responsible for creating the alcohol in wine and beer. It also uses two molecules of NADH, thus using energy. Although ethanol is a common end-product in these reactions, some organisms produce acetate instead.

Alternatively, the pyruvate may be reduced to a variety of three-carbon molecules, including alanine and lactate, again consuming two NADH molecules. Those molecules may be used to build other structures – alanine, for example, is an amino acid used in proteins – or they may be waste products and eventually expelled. These alternative
fates can take place in a wide variety of organisms, including Bacteria, Archaea, and Eukaryotes that live in places with little or no oxygen. Thus, **anaerobic metabolism** (metabolism in the absence of oxygen) is another example of both the unity and diversity of metabolic systems.

Anaerobic metabolism can also occur in organisms with aerobic metabolisms during short-term periods when oxygen becomes limiting. For example, when you perform extreme physical events like running a marathon, you can exceed the ability of your blood to supply oxygen to your muscles. When that happens, the sugar that your muscles are trying to use gets converted to lactate, and it is the presence of lactic acid that causes your muscles to feel sore.

If oxygen is available, the pyruvate is oxidized to produce three molecules of CO$_2$ while producing one molecule of ATP, four molecules of NADH and one molecule of FADH$_2$ (another, similar energy carrier). Most of this happens as part of a cyclical process involving a series of four-, five- and six-carbon intermediates (Figure 5.15). The cycle is variously known as the citric acid cycle, for the first stable intermediate, or the Krebs cycle, for the German-British scientist Hans Krebs, who described much more...
of the process in the 1930s. In Eukaryotes this process is carried out in the mitochn-
dria, an organelle found in all but a handful of species (Figure 5.16).

This oxygen-requiring process extracts much more energy from the glucose mol-
ecule than the first part of the process where glucose was split into two pieces. So, once oxygen-producing photosynthesis evolved and the levels of oxygen in the atmosphere became substantial, organisms could be much more complex. It is probably no coinci-
dence that it is not until after this happens that Eukaryotes evolved because their cells are generally larger much more complex than those of Bacteria and Archaea.

A common misconception is that plants do not use oxygen or give off CO₂. Plants, just like other Eukaryotes, have mitochondria and break down glucose. During the day, this process is far outstripped by photosynthesis, so that overall plants are taking in much more CO₂ than they are giving off. At night, however, they are giving off CO₂ just like animals.

Scientists have suggested as part of the Metabolism First theory for life’s origin (see Chapter 1) that a version of the citric acid cycle predated the origin of life. A ver-
sion of that cycle is found in all species and is linked to the synthesis of most biological molecules. Originally it would have run in the opposite direction as a synthetic cycle, possibly using hydrogen sulfide (H₂S) as the electron donor and iron sulfide (FeS) as the electron acceptor. These chemical reactions do not necessarily require enzymes to happen at rates high enough to create sufficient concentrations of organic molecules, a requirement for pre-living systems. Because the system is cyclic it could form the basis of a self-organizing living system. Only about a billion years later, after cells had evolved and oxygen became plentiful, was the cycle co-opted as part of oxidative metabolism.

**Nutrients**

Because cells are dynamic entities, they require a constant input to maintain themselves. Besides energy, organisms require materials to build their structures and their dynamic nature means that structures are constantly being rebuilt and changed. The building blocks for these structures can be obtained in two ways: a cell can build them from scratch or it can take them in as food. This movement of ions and molecules both into a cell and among the parts of the cell is one of the primary uses of a cell’s energy (Table 5.1, principle 7).

Some organisms can build all of their needed structures beginning with the sim-
plest inorganic molecules: hydrogen and oxygen from H₂O, or from CO₂ or CH₄ which also serve as a source of carbon, nitrogen from N₂, NH₄, HNO₂ or HNO₃, and phosphorus from HPO₄. Such organisms also have to be able to obtain energy from light or chemi-
cal sources. Species able to do this include a variety of Bacteria and Archaea, along with plants and other photosynthetic Eukaryotes.

For all other species, these materials are obtained, at least in part, by taking in organic materials built either by the species just mentioned, or from others that have eaten those species. Some material is broken down completely and the basic molecules reused. In other cases, material is broken down just partially, while other material is used as consumed. Consider what happens when you eat a piece of protein, as a burger made either from beef or tofu. The protein is first broken down into its amino acids, many of which may be further broken down into ammonium (NH₃⁺) and carboxylate (HCO₃⁻) ions and then reused to build new amino acids. However, humans are incapable of building the amino acid tryptophan so we must take it in as part of our diet. In this case, both beef and soybeans in the form of tofu can be sources of tryptophan. Because tryptophan must be obtained by consumption, it is called an essential amino acid.

The term **vitamin** is used for those molecules that are taken in and used as is without first being broken down, although they may be incorporated into larger
molecules or further modified. For example, vitamin B5 (pantothenic acid) is the coA part of acetyl-coA that participates in the citric acid cycle (Figure 5.CC). Vitamin C (ascorbic acid) is necessary for humans, great apes, bats, and a few other mammals, notably guinea pigs, and some birds. Other mammals and birds are capable of synthesizing this molecule. In some cases, vitamins are needed only under some circumstances. For the synthesis of vitamin D, humans require sunlight which reacts with the pigment melanin in our skin. People living in very northern or southern latitudes which have little sun for large parts of the year must consume vitamin D in their diet.

The network of chemical reactions in a cell (Figure 5.YP) is a linking of the synthesis and breakdown of organic molecules. RNA molecules are assembled as part of the process of transcription (see Chapter 3) and then broken back down into their nucleic acid components to be recycled into new RNA molecules. Often the chemical reactions involved in synthesis and breakdown are very similar, even identical, and in a few cases may even make use of the same enzyme. Depending on which way the reaction is run, it may use energy or it may generate energy. When you eat a bowl of ice cream that is full of sugar, fats and protein, the breakdown of all of those materials provides both energy and material that is then re-used to build new forms of the sugars, fats and proteins to meet the specific needs of your cells.

It is no surprise that the same chemical reactions running forward in plants when synthesizing materials run in the opposite direction in animals when they consume those materials. Most of the biochemical pathways in plants and animals first evolved in the bacterial and archaeal ancestors of Eukaryotes. Natural selection tends to be conservative, re-using and re-purposing functions and structures that may have originally evolved for other purposes (see Chapter 4). Evolution also explains why metabolic pathways sometimes seem overly complex. Evolution always starts with what is already present, then adding or subtracting as mutation and natural selection allow. The result is often a workable solution, rather than the best if it were designed from scratch.

**Cell Replication**

The processes of life exist for and are perpetuated through living organisms maintaining themselves, growing, and reproducing. In chapter 6, we will look at these processes from the perspective of the whole organism. Here, we examine them from the viewpoint of a single cell. We have already looked in detail at the process of maintenance through the intake of matter and energy and the use of those materials for building and regulating cellular structures. If an organism consists of only a single cell, then growth occurs through the intake of matter. At some point, if it has ingested enough material, it will divide into two cells. For multicellular organisms, cell division is the primary way that they grow, although some growth can occur by cell expansion. All cells come from the division of previously existing cells (Table 5.1, principle 8). This process of cell replication is called **mitosis**. In Chapter 3 we looked at one aspect of this process, the replication of chromosomes (see Figure 3.4). Elaborate cell machinery involving the cytoskeleton exists to ensure that the duplicated chromosomes are correctly sorted into the daughter cells. For the rest of the cellular components, no specific machinery exists. Instead, because all of the parts are spread out across the entire cytoplasm, when cells divide enough material of all types necessary to form the entire machinery of the cell ends up in each of the daughter cells. Mitochondria and chloroplasts go through their own processes of growth, chromosome duplication, and division within a cell, so that they exist in multiple copies. These copies subsequently end up being distributed among the daughter cells.

The information necessary for these processes is contained in the cell’s
chromosome(s) (Table 5.1, principle 9), which constitute the mechanism by which information is passed from one generation of cells to the next (Chapter 3). In Bacteria and Archaea there is typically one large circular chromosome, although there may also be additional small circular chromosomes. In Eukaryotes, the nucleus contains one or more large linear chromosomes, while the mitochondria and chloroplasts within the cells contain their own, smaller circular chromosomes. The chromosomal theory of inheritance was first proposed by T. H. Morgan (see Box 3E) in the early twentieth century, which eventually lead to the discovery of the structure of DNA (see Box 3A).

That new cells come from existing cells is one of the two central tenets of cell theory put forward in the 19th century (see Chapter 1), and the promulgation of cell theory was part of the general development of biology as a discipline. The other tenet – that all organisms are composed of cells – is now considered part of the theory of organisms (see Chapter 6).

The Evolution of Cells

Cell evolution (Table 5.1, principle 10) has gone through three critical stages: the origin of the first cells, the origin of Eukaryotes, and multicellularity. The first two stages occurred once; all of life on Earth traces to a single common ancestor and all Eukaryotes trace to a single common ancestor. That does not mean that the process happened in a single step; undoubtedly it was a multiphase process. The evolution of multicellular species, however, clearly happened multiple times. The first two stages are similar in that they involved the coming together of independent entities, each with different abilities, to form a new, single cell. The last stage reversed this process to some extent, taking a cell and dividing its abilities among multiple entities. Because evolution includes contingent events (see Table 4.1, principle 7), at least some of the features of cells are due to chance events, although we do not know which ones; for example, scientists debate whether the basic metabolic processes that are shared by all cells were inevitable because they were chemically favored, or if they were just an accident of life’s origin.

Stage 1: Life’s origin

The origin of life and the first cell is discussed in detail in Chapter 1. To summarize, if the reconciliation model is correct, then that cell originated when self-replicating nucleic acids and self-catalyzing redox reactions were brought together within a lipid bubble. Each of those components accounted for a different cellular function: information, metabolism, and structure. It was a type of symbiosis, an interdependent or mutually beneficial relationship between two or more entities.

Critical to this process is that the various entities develop a way to avoid “cheating,” which is to say, one entity improving its fitness at the expense of its partners. The first cells did this by linking the parts into a shared fate. The information and metabolic components became unable to reproduce on their own. The information contained in the structures of and relationships among the metabolic components became encoded within the information system. Conversely, the information system in moving from RNA to DNA became incapable of independent replication. The linking of all of the pieces of information (nucleic acids) into a single chromosome is another way of creating a shared fate, a strategy found in the single large, circular chromosome of all Bacteria and Archaea. Recently, scientists have tried to answer the question of what comprises the minimum structure and function necessary to be a living cell. In answering this question, they hope to both gain a better understanding of how cells function and insight into what the common ancestor of all life might have looked like, an entity sometimes called the Last Universal Common Ancestor (LUCA). Since that ancestor is now long extinct, its biochemical and genetic make-up has to be deduced by looking at all of the living
species and determining what functions all of them share. Depending on how that calculation is done, the LUCA may have had somewhere between 500 and 1500 genes, about half of which were involved in the maintenance and replication of the information system and the other half involved in cellular metabolism and structure. If this is correct, then clearly a lot of evolution had to have occurred between the time of the origin of the first cell and the ancestor that ultimately gave rise to today’s species. This raises the possibility that life arose multiple times independently and all but one of those origins went extinct. Of course, that is a hypothesis that is impossible to test, as there are no traces of those other possible lineages.

**Stage 2: Eukaryotic cells**

The next stage was the origin of the eukaryotic cell. While much is still unknown about this process, our understanding has improved markedly in the past 40 years since Lynn Margulis popularized the endosymbiotic theory that had been first proposed in 1905 by the Russian botanist Konstantin Mereschkowsk. To understand this theory, consider the mitochondrion, which has several curious features. First, mitochondria replicate through a process of binary fission like Bacteria; they are not built from scratch like other cellular structures. Second, it has not just one membrane but a double membrane. Third, it has a small circular chromosome similar to those in Bacteria that codes for some, but not all, of its machinery. Fourth, the ribosomes in mitochondria are much more like those in Bacteria rather than those in the rest of the cell.

Based on this evidence, Margulis and others suggested that the mitochondrion was originally a free-living Bacteria. Typically, Bacteria are consumed by larger single-celled organisms by being engulfed (Figure 5.AE). The hypothesis is that in one instance a bacterium that was particularly good at oxidative metabolism was not consumed, but remained alive inside the cell. The bacterium formed a symbiosis with the surrounding cell: in exchange for pyruvate and other materials, it supplied the cell with energy in the form of ATP, NADH and FADH. Over time, some of the genes on the bacterial chromosome were transferred to the nuclear chromosomes, tying the fate of the mitochondrion to the rest of the cell. The engulfment explains the double membrane; one would have originally been the outer envelope of the independent bacterium, and one formed when the original cell engulfed and attempted to consume the bacterium. It also explains why mitochondria have many features similar to Bacteria. Although this theory was controversial for some time, it is now the accepted explanation. The strongest evidence is based on more recent sequencing of the mitochondrial chromosome showing that its DNA sequence is more like that of some Bacteria than it is like nuclear genes. This process has also identified the group of Bacteria from which mitochondria most likely derived, the Proteobacteria. DNA sequencing has also been able to identify genes now located on nuclear chromosomes that appear to have originated in the mitochondria. As expected, these genes code for functions involved in oxidative metabolism and material that are transferred into the mitochondria.

What about the rest of the features of eukaryotic cells? Eukaryotes are distinguished from Bacteria and Archaea by having a flexible cell surface, linear chromosomes, a nuclear envelope, a complex membrane system including digestive vesicles, and a cytoskeleton. Some of these features can be explained, while others are still the subject of speculation and research. Based on DNA sequence evidence, Eukaryotes are more closely related to Archaea, particularly the genes that code for proteins involved in information system functions (translation, transcription, replication, and repair). On the other hand, some aspects of eukaryotic cells, such as the cell membrane are more like those of Bacteria. Interestingly, the genes that code for proteins involved with functional and structural aspects of the cell (metabolic enzymes, components of membranes, and
other cellular structures) appear to be more closely related to Bacteria. One possibility is that the original Eukaryote came about through a fusion of an archaeon and a bacterium with the former becoming the nucleus and the latter the cytoplasm. Suggestive of this is that the nuclear membrane, like that of the mitochondria, is also a double membrane. At some point, either before or after this fusion, the organism had to lose its cell wall leaving it with a flexible membrane, and infoldings of that membrane would have produced the complex membrane structure that we see today. The cytoskeleton would have been derived from what was previously strictly external cilia or flagella. What is currently completely unexplained is how chromosomes went from being circular to linear. Even the timing is still very uncertain. There are clear fossil single-celled Eukaryotes 1.6 billion years old, with some possibly as old as 2.1 billion years, and some much sketchier trace evidence as far back as 2.7 billion years.

Once the first Eukaryote appeared, it would have been quickly favored. The complex membrane system provided for greater spatial structuring and more efficient chemistry. Organelles such as mitochondria allow for compartmentalization of functions, especially those involving chemicals that could harm other parts of the cell. At some point, Eukaryotes evolved meiosis and a regularized system of recombination, greatly increasing the amount of genetic variation (Chapter 3) and the rate of evolution (Chapter 4).

The origin of chloroplasts is similar to that of mitochondria. Like mitochondria, chloroplasts also have their own circular chromosome and a double membrane. This symbiosis, though, occurred at least three different times, the first one about 1.5 billion years ago. The evolution of photosynthesis in single-celled Eukaryotes is quite complex, with evidence of secondary symbiosis: a photosynthetic Eukaryote being engulfed by another Eukaryote. Evidence for this comes from both DNA sequence data and the presence of a triple membrane around the chloroplast.

**Stage 3: Multicellularity**

The third stage in cellular evolution is multicellularity. Like the advent of eukaryotic cells, this development granted an important advantage for organisms that possessed it; the component cells of an organism specialize for different functions. Like the evolution of compartmentalization in eukaryotic cells, specialization increases the efficiency of cell functions. Not only can functions be separated that might interfere with each other, but cells can take on different forms suited to particular tasks. However, separating these functions requires that cells be able to communicate, both during their initial creation and throughout an organism’s lifespan. After all, it would not do for all of your cells to become heart cells, some must become lungs, nerves, skin, and so forth. This means that the communication systems described earlier are critical components; there must be ways to ensure that the cells cooperate. If this communication system breaks down, the results can be disastrous; one way to think about a cancerous cell is as a single cell that is reproducing itself at the expense of the entire individual. We take up the story of multicellularity in more detail in the next chapter.
Chapter 6
Organisms

Life on Earth consists of a bewildering array of organisms, from tiny microbes to giant blue whales (Figure 6.1, also see Figure 1.1); as varied in form as the organisms themselves are the strategies they employ to survive and flourish. Some survive on just sunlight and a few simple nutrients, while others have elaborate ways to capture and consume other living organisms. How can we make sense of this vast hodgepodge? One theme of this chapter is the wide diversity of forms, functions and life history patterns taken by life on Earth. Instead of trying to describe all the different ways that organisms exist, as with the rest of this book we will extract and describe the broad characteristics that unite all of life. In doing so we recognize that we are skipping over an enormous amount of detail that is important to explore in order to truly and deeply understand how organisms function.

One aspect of this great diversity is that except at the most fundamental level (e.g., the need for energy), there are no universal generalities about how organisms function. Although there are basic generalizations, everything has exceptions. The goal of this book is to provide you with an overview of those generalizations. We will give some sense of the frequency or rarity of the exceptions through the use of terms like “some”, “many”, and “most”. However, a full exploration of all the byways and detours of those exceptions would require a book many times the size of this one, and those details are best left to others.

The exact number of species on Earth is unknown. Currently about 1.6 million are named, with estimates of the actual number ranging from 3 million to 10 million; a few scientists estimate as many as 100 million. Part of the reason for this great range in estimates is due to definitions of “species” (see Chapter 4), especially with regard to Bacteria and Archaea. The vast majority of organisms are single-celled and they are less well studied than multicellular life forms because they are harder to observe and raise in the laboratory. However, even in groups that are extremely well studied, such as flowering plants and mammals, new species are being discovered each year. Even large species can be elusive – a new species of deer was discovered in Vietnam in the mid-1990s.

An organism is a bounded unit. Boundedness is a basic property of living systems that is necessary for their persistence (see Chapter 1), and that boundedness occurs at two levels, the cell and the whole organism. In Chapter 5 we explored the properties of cells that make them the foundational units of living systems. In this chapter we look at whole organisms, the units that integrate those cells into a complete system. For the multitude of single-celled organisms, the foundational and integrative units are one and the same. In this chapter we will touch briefly on aspects of single-celled organisms not addressed in the Chapter 5, notably the ways in which they interact with their external environment. Much of this chapter will focus on multicellular organisms, which have numerous additional properties simply because they are multicellular. We humans are included among those multicellular organisms, and so there is a natural tendency to focus on the properties of humans. However, it is important to understand that we are just one (somewhat peculiar) species, so this chapter will mostly discuss the rest of that vast diversity.

As mentioned in Chapter 1, the individual is biology’s integrative unit, where the information in the genome is manifest as a phenotypic whole. The expression of those genes is context dependent (see Chapter 3), depending on both the rest of the individual and its environment. The individual provides the filter through which the environment
interacts with that genome. All individuals are born and die, and in between there is maintenance, growth and reproduction. Each individual has a particular form that accomplishes all of these tasks. This chapter will explore all of these aspects of individuals.

**What is an Individual?**

Central to understanding organisms is to understand the concept of the individual. To us humans, what is an individual seems obvious. Each of us has a boundary layer, the skin, which separates us from other humans and the rest of the world. However, that simple appearance is deceiving. Contained inside of us are millions of other organisms, primarily Bacteria, and even the surface of our hair is home to numerous mites, tiny organisms related to spiders. To some extent we are actually a community of organisms, a topic that we discuss in more detail in Chapter 7. In this chapter we focus on the separation of entities that are members of the same species.

Most species are like us; each individual has an obvious boundary. For single-celled organisms that boundary is its lipid membrane (see Chapter 5). However, some single-celled organisms live in aggregations. In those instances, how can we decide if an entity is a collection of potentially independent units, or if it is a multicellular organism? The key is whether or not a part that is removed from an entity is able to exist on its own. That means that the part can perform active self-maintenance and can reproduce. In other words, it is able to persist both as itself and across generations, persistence be
ing the first fundamental principle of living systems (see Table 1.2). For example, we can take a cell from a human body and grow it in a petri dish. That cell can survive and divide, but only if it is continually supplied with nutrients that it would normally receive from the rest of the body. Without that support system, the cell would die; therefore, we would not consider such cells as individuals.

For most organisms, the criteria of self-maintenance and reproduction let us distinguish collections of individuals from integrated units. But what about organisms such as honey bees (see Figure 7.10B)? A honey bee hive is a collection of apparently independent individuals. A closer examination, however, shows that nearly all of those individuals are sterile. Although they are able to achieve self-maintenance, they cannot reproduce. The entire hive acts as the reproductive unit, with different organisms carrying out different tasks that are necessary for the self-maintenance and reproduction of the entire hive. For honey bees, is the hive an individual? The possibly unsatisfying and certain intriguing answer is sort of, in that the hive as a whole has many, but not all, of the properties of an individual.

What counts as an individual is not always so easy to discern and there is a very blurry line defining that border. We can evaluate how closely tied different components are by two measures – physiological integration and genetic homogeneity – that correspond to the self-maintenance and reproduction criteria. For example, many plants grow through the spread of roots, such as the grass in your yard. When a new shoot starts it gets all, or nearly all, of its nutrients from adjacent shoots through those roots. At some point, the shoot becomes big enough for its own leaves and roots to supply the needed nutrients. The shoot may also be capable of flowering and producing seeds. So the shoot has the properties of an independent individual. If it was still connected to the rest of the shoots, however, we would consider the entire collection of shoots as one individual. Even if the connections are severed, all of the shoots are genetically identical, and so are still very close to being a single individual.

Those shoots may not be genetically uniform, though, if mutations have occurred during growth. Even in a human, mutations can occur during growth. In some cases, those mutations cause a cell to act more like a separate individual than part of an integrative whole. The cell begins to grow and divide (i.e., reproduce) uncontrollably and possibly move through the body. This is what happens when a cell becomes cancerous. Cancer cells can be considered to be parasitic invaders inside your body, just like other disease-causing organisms. Even though they started out as part of your body, they have become genetically different and physiologically separate.

A hive of honey bees is at the far end of this continuum. Physiological integration is loose. Each bee is capable of maintaining its own body, although some bees are dependent on others to bring back nectar and pollen. All of the bees are related to each other as sisters, but just like any siblings they are not genetically identical. This genetic relatedness can get even more tenuous in other species of bees or wasps in which a hive may have multiple queens who may or may not be sisters. In those cases, the other bees may be cousins, or even unrelated, although they still rely on each other for functions such as foraging for food. Thus, a collection of organisms may be physiologically integrated but not genetically homogeneous, or may be genetically homogeneous but not physiologically integrated. Throughout this chapter we will present various examples of organisms that span this continuum.

How does persistence of an entity affect how we identify an individual? Your body is continually being renewed so that after about 7-10 years every single molecule has been replaced. Are you the same individual as seven years ago? Would the answer be different if you could instantaneously replicate your body while destroying the original?
What about the process by which a single-celled organism reproduces by dividing in two – are the two new cells new individuals or the old individual now grown and separated? Does the answer differ if the division is into a large and a small cell, rather than two of the same size? The answers to these questions are as much a matter of definition as anything and have been debated by philosophers for centuries. The key point is that the notion of an individual, while seemingly simple, can get quite complex when considered in the light of the diversity of life.

### The Theory of Organisms

The structure and function of individuals, especially their immense variety, can be explained by a theory of organisms (Table 6.1) which consists of ten fundamental principles. The first four principles deal with the internal structure and function of organisms. They provide links with the theories of genetics and cells (see Chapters 3 and 5). The next four principles deal with interactions with the external environment and provide links to the theories of cells and ecology (see Chapter 7). The last two principles are about the causes of organismal properties and link to the theory of evolution (see Chapter 4). That this theory links so closely with all of the other theories is indicative of the fact that biology is, at its heart, a science of organisms. Understanding the form and function of individuals is central to understanding life.

The principles put forth in Table 6.1 apply to all organisms. Multicellular organisms have additional properties that need to be accounted for in an additional subtheory (Table 6.2). To understand both the common properties of organisms and how those common properties vary, throughout this chapter we will explore both those general properties and those specific to multicellular individuals, referring back to both sets of principles. Just to make a complex case even messier, to some extent these principles also apply to groups of organisms that act like individuals.

---

**Table 6.1. The fundamental principles of the theory of organisms.**

<table>
<thead>
<tr>
<th>Principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. An individual organism actively maintains its structural and functional integrity.</td>
</tr>
<tr>
<td>2. All organisms are composed of cells at some point in their life cycle.</td>
</tr>
<tr>
<td>3. Maintenance of organismal integrity requires dynamic change.</td>
</tr>
<tr>
<td>4. Maintenance of organismal integrity is a function of interactions with the abiotic and biotic environment.</td>
</tr>
<tr>
<td>5. Organisms require external sources of materials and energy for maintenance, growth and reproduction.</td>
</tr>
<tr>
<td>6. Organismal functions trade off against each other.</td>
</tr>
<tr>
<td>7. Because organisms are changeable, external influences can force change.</td>
</tr>
<tr>
<td>8. Environmental heterogeneity in space and time leads to variation in life history patterns.</td>
</tr>
<tr>
<td>9. Organismal reproduction is both a cause and consequence of evolutionary processes.</td>
</tr>
<tr>
<td>10. The properties of organisms are the result of evolution.</td>
</tr>
</tbody>
</table>
Table 6.2. The fundamental principles of the subtheory of multicellularity.

1. Multicellularity allows for specialization of cells.
2. Cell-cell interactions are necessary for cell specialization.
3. Specialization of cells requires their spatial or temporal localization at some point in the life cycle.
4. Specialization of cells allows for modularity.
5. Specialization of cells leads to emergent organismal properties.
6. Development requires heterogeneity in cellular or organismal composition.

Structure and Function

Integrity

First and foremost, individuals must maintain themselves (Table 6.1, principle 1). This maintenance has two aspects: mechanisms that keep the individual intact and functioning, and mechanisms that keep one individual distinct from other individuals. Keeping individuals distinct seems to be a simple proposition – after all, you are not likely to suddenly merge with the person next to you. But this is not so easy for the simplest multicelled organisms. If an individual has a rigid covering – as do plants and fungi – then individuals can easily stay distinct. However, the simplest multicellular animals, the sponges (Figure 6.2) lack such a covering and are not much more than bundles of cells. What prevents two such bundles from merging with each other? They must have mechanisms that prevent that merging, which include special proteins on the cell surface that are recognition signals – they identify self from non-self. If the cells of a sponge are broken apart and two individuals mixed together, they will gradually sort themselves back into the two original individuals.

This recognition system plays another very important role, prevention of bodily invasion by other species. Such recognition systems exist in all species, from Bacteria and Archaea to animals, plants, and fungi. For example, all individuals, regardless of complexity, can be invaded by viruses. To enter an individual, a virus has to be able to chemically link with the proteins on the surface of the species. Nearly all multicellular species – animals, plants, and fungi – have some type of system that responds to non-self. In animals, this system consists of specialized cells that engulf and digest invaders. Vertebrates have two systems. One system is found in all animals and consists of generalized cells that recognize any invader. The other system produces specialized cells that are keyed to particular invaders. This second system is the one that protects you against viruses like influenza. Vaccinations provide protection from disease by stimulating that system to produce specialized cells before you get infected. The continual evolution of the influenza virus means that it is constantly producing new forms which will not be recognized by those cells, which is why you need to be revaccinated each year.

Plants have very different strategies for dealing with invaders. Instead of specialized cells, they respond with chemicals. Some chemicals directly attack the invader while...
others act as signals that wall off the invaded part of the plant. The plant may even simply shed the infected part; perhaps unfortunately, such shedding of body parts is generally not an option for animals.

Recognition systems also play important roles in reproduction. The maintenance of species as independent units requires that individuals mate with others of their own species (see Chapter 4). In many species, this recognition is managed through elaborate behaviors, but chemical recognition also plays an important role. Gametes have specific proteins on their surface that must match before fertilization can occur. For many species that produce both male and female gametes, like most flowering plants, the proteins also act to prevent two gametes from the same individual from merging. Thus, the system even acts to allow one cell to recognize a gradient of similarities, identifying others that are enough like it to be members of the same species, but not so alike that the particular cell in question is part of the same individual.

In animals that have live births, like humans, a different problem occurs: how to keep the mother from rejecting the growing embryo as a foreign invader. In mammals there is a complex membrane system that separates the blood and circulatory system of the embryo from that of the mother so that her immune system does not attack the embryo. The reaction of the mother’s immune system can be especially severe if the mother and embryo have different blood types (see Chapter 3). Recently, scientists found that the part of the human recognition system has undergone rapid evolution during the past few million years as we evolved from our ape-like ancestors. They speculated that this evolution was driven by the longer gestation period in humans resulting in prolonged interactions between the mother and the embryo.

**Cells**

All organisms are composed of cells at some point in their life cycle (Table 6.1, principle 2). Most organisms consist of cells all of the time (a single-celled organism, of course, always consists of a cell). The vast majority of multicellular organisms also consist of cells all of the time. However, some organisms do not have recognizable, separate cells. Fungi are made up of a netlike mass of filaments. In some species, the filaments do not have cell walls, instead consisting of a continuous cytoplasm with many nuclei. However, during reproduction distinct cells are formed. So, even in organisms that are acellular for much of their life cycle, at some point, usually as part of sexual reproduction, individual cells are produced.

Animals are typically always cellular. Insects, however, are a case just as strange and interesting as fungi, albeit in the inverse; after fertilization the chromosomes divide and produce multiple nuclei, but no cell division takes place. Eventually, cell membranes are formed and the new individual grows into a cellular organism. We discuss this process of development in more detail later in this chapter.

Slime molds are organisms that straddle the line between single-celled and multicelled, existing sometimes as individual cells and sometimes as a multinucleate, acellular organism (Figure 6.3). Slime molds start out as single-celled amoebae-like organisms living in the soil and eating bacteria. Those cells grow and divide. Under certain circumstances, often in response to low food levels, a signal is sent out that causes the cells to aggregate into a large, acellular structure. Because the aggregating cells are generally closely related but not genetically identical, this aggregation process requires cells to have a recognition system similar to that of a gamete in its sensitivity to differences and similarities. In most species, this acellular structure is just a few centimeters across, but the largest ever found was 30 meters across, making it the biggest acellular organism known. These large aggregates can move, and in some species two of them might merge to become an even larger individual. Eventually some of the nuclei will produce
spores and start the cycle all over.

The principle that organisms consist of cells is one of two from the cell theory that was first put forward in the nineteenth century. The other central tenet of that theory was that new cells came from old cells (see Chapter 5). In the new, more expanded versions of that theory, the two principles are separated into one that is a statement about organisms (discussed here) and one that is about cells per se (discussed in Chapter 5). Theories are dynamic entities, continually being modified and updated as new information becomes available. What began in the nineteenth century as a single theory consisting of just two principles, has become two theories each consisting of ten principles in the twenty-first century. The expansion and elaboration of the theory reflects a similar expansion and elaboration of our knowledge about cells and the nature of organisms. The original version of this principle stated that all organisms consist of cells, and as we have just discussed, some organisms are acellular for part of their life cycle; thus, the principle had to be modified. Such modifications are an example of the dynamic nature of theories.

Cell specialization

A striking feature of life on Earth is the multitude of multicellular organisms, including ourselves. Multicellular organisms are a rather late arrival, showing up about 1 billion years ago, approximately 3 billion years after the origin of life. When they finally did arrive, though, they quickly came to diversify and dominate the world. During the Cambrian period about 500 million years ago, the rate of diversification of animals was so great that scientists refer to this event as the Cambrian Explosion. What was it about multicellularity that lead to this diversification?

A key feature of multicellular organisms is that they consist of more than one cell type (Table 6.2, principle 1). It is the existence of such specialized cells that allows us to distinguish a multicellular individual from a mere collection of cells. Specialization leads to increased functionality and efficiency because it ameliorates the trade-offs (discussed later in this chapter) caused by having to perform multiple functions simultaneously or with the same set of structures.

Cell specialization is not the exclusive domain of Eukaryotes. Some Bacteria are also multicellular, and this increased functionality brought about by specialization can be seen in Anabaena. The species can perform two different metabolic functions; it is
The photosynthetic bacteria *Anabaena* grows as a chain of connected cells. While most of the cells carry out photosynthesis, some of the cells are specialized for the fixation of nitrogen. (Photo by Bdcarl, Source: Wikipedia)

A key feature of multicellularity is that it allows for more complex structures and functions. One way to view organismal evolution is by considering the number of cell types contained within a single individual (Figure 6.5). That number increased greatly (from about 10 to 120) during the period from 1.5 billion years ago to 500 million years ago, the period leading up to the Cambrian Explosion. Diversity of cell types, therefore, led to diversity of species.

Unlike the single origin of all modern living organisms, multicellularity has originated multiple times. Animals, plants, and fungi each represent a separate origin (see Figure 4.2). There are many other groups of organisms, though, that are multicellular, albeit consisting of just two or a few specialized cell types, for example, several groups of algae, the slime molds, and various Bacteria such as *Anabaena*.

---

Figure 6.4
The photosynthetic bacteria *Anabaena* grows as a chain of connected cells. While most of the cells carry out photosynthesis, some of the cells are specialized for the fixation of nitrogen. (Photo by Bdcarl, Source: Wikipedia)

Figure 6.5
The maximum number of cell types per individual has increased over the history of life on Earth. There was an increase when Eukaryotes first arose about 2 billion years ago, and then a second, larger increase during the origin of multicellular Eukaryotes about 1.5 billion years ago continuing through the Cambrian Explosion 500 million years ago. Life today (time 0) shows a wide variety of cell types per individual. (from: Hedges, S. B., M. L. Venturi, and J. L. Shoe. 2004. *BMC Evolutionary Biology* 4:2 doi:10.1186/1471-2148-4-2)
Cell specialization comes about through the differential expression of genes. All cells in an organism contain the same genes, but different parts of the genetic code are activated in each specialized type of cell. The process of differential gene expression leading to cell specialization is one component of development, a topic that we consider later in this chapter.

One type of cell of special interest is involved in reproduction, called a gamete. In multicellular Eukaryotes, reproduction involves the processes of meiosis, which produces haploid cells that have half the number of chromosomes as the original, and fusion of those cells to produce a diploid cell with the original number (see Chapter 3). In nearly all animals, the only haploid cells are the gametes; they always exist only as single cells, unable to grow or divide without fusion with another gamete. In contrast, in plants and fungi, the haploid cells can divide and grow into complete individuals. In some species, the individuals that you see are haploid while the diploid stage has been reduced to just one or a few cells. In flowering plants, although the haploid part of the life cycle has been greatly reduced, it still exists as a distinct stage (Figure 6.23A).

The sex of an individual is defined based on the type of gametes it produces. Fungi and some simple Eukaryotes produce gametes that are all the same size and so do not have distinct sexes. All animals and nearly all plants produce two types of gametes, one much larger than the other. Individuals that produce large gametes (i.e., eggs or ovules) are called females and individuals that produce small gametes (i.e., sperm or pollen) are called males. Some individuals produce both types and are called hermaphrodites. Diploid reproductive cells are typically called spores although spores can also be haploid.

Multicellularity itself drives further cell specialization. In Anabaena, products from one type of cell can be moved to the other type through simple diffusion. However, diffusion is a slow process and not very efficient. Although organisms that live in an aquatic environment can often depend on simple diffusion from the surrounding environment to provide needed materials, such as oxygen, they are still constrained by the surface-to-volume problem.

The surface of an object increases with the square of its size while the volume increases with the cube. A sphere with a diameter of 2 cm has a surface area of approximately 12.6 cm² and a volume of 4.2 cm³. If you double that diameter to 4 cm, the surface area increases to 50.3 cm², a four-fold increase, while the volume increases to 33.5 cm³, an eight-fold increase. Because of this, organisms that are more than a few cells thick (e.g., sponges, Figure 6.2) require active transport mechanisms to get materials to cells in their interior.

Animals contain systems for taking in solids and liquids (e.g., eating and drinking), breathing in oxygen, moving those materials around inside the body, and then expelling waste material or breathing out CO₂ (Figure 6.6A). Land-dwelling species need to have complex systems to accomplish these tasks, especially ways to avoid water loss. Plants have fewer cell types and systems than animals, but they also have specialized systems for performing photosynthesis, taking up water and nutrients from the soil, and moving materials among different parts (Figure 6.6B). The other group of multicellular organisms, fungi, have much simpler bodies with just a few specialized systems (Figure 6.6C). In all of these types of organisms, coordination of those systems requires mechanisms for communication. Communication systems are particularly important for maintaining organismal integrity, the topic of the next section.
Figure 6.6
The structure and various organ systems of (A) a human (Source: Wikipedia), (B) a plant (Created by Kelvinsong, Source: Wikipedia), and (C) a fungus (Created by Debivort, Source: Wikipedia).
Non-adjacent cells require more elaborate mechanisms for communication, especially in complex and large organisms. These communication mechanisms can be either chemical or electrical. All multicellular species use chemical signals, which can be specialized molecules that interact with specific receptors. In humans, for example, after consuming carbohydrates, the concentration of glucose in the blood rises. This increase stimulates a special organ, the pancreas, to release the molecule insulin. Insulin circulates throughout the body and acts as a signal for cells to take up glucose from the blood. If the concentration of glucose in the blood falls below a set level, other cells in the

Cell-cell interactions

In order for cells to specialize in an orderly fashion during development and to coordinate their functions, they must be able to interact and communicate (Table 6.2, principle 2). Those cell-cell interactions come about through a variety of mechanisms, although we can broadly categorize them into two types of interactions, those involving the exchange of information and those involving the exchange of matter (although information itself is transported as matter). The difference is that an information exchange involves a small signal (just one or a few molecules) that initiates a large response. In contrast, the exchange of matter (e.g., large amounts of carbohydrates) typically initiates a response that is proportional to the amount of matter. To some extent, however, this is a false dichotomy and interactions among cells involve a wide range of exchanges.

The mechanisms of information exchange depend, to some extent, on whether cells are adjacent or distant. If cells are adjacent to each other, they can communicate and interact through the exchange of molecules across their membranes. Often, adjacent cells have special pores to facilitate such exchanges. It is not even necessary for a molecule to actually pass into another cell to create an interaction. The molecule can attach to a protein that extends through the cell membrane, change the configuration of that molecule, and trigger a chemical reaction on the inside of the cell (see Figure 5.3).

Non-adjacent cells require more elaborate mechanisms for communication, especially in complex and large organisms. These communication mechanisms can be either chemical or electrical. All multicellular species use chemical signals, which can be specialized molecules that interact with specific receptors. In humans, for example, after consuming carbohydrates, the concentration of glucose in the blood rises. This increase stimulates a special organ, the pancreas, to release the molecule insulin. Insulin circulates throughout the body and acts as a signal for cells to take up glucose from the blood. If the concentration of glucose in the blood falls below a set level, other cells in the

Figure 6.7
The chemical structure of ethylene.
pancreas release the molecule glucagon. This molecule stimulates cells in the liver to release glucose into the bloodstream. Through this system of feedbacks between two different signals, glucose levels in the body are maintained in a dynamic equilibrium. The disease diabetes occurs when the body is unable to produce insulin or respond properly to the signal, thus interfering with this feedback system (see Box 5B). This example also shows how signals can be a combination of specialized molecules acting at low concentrations (e.g., insulin) and more common molecules (e.g., glucose).

Figure 6.8
A nerve cell showing the process of signaling. (Source: National Institute of Aging, U.S. NIH)

Signaling molecules can be very simple or quite complex. In plants, the chemical ethylene (Figure 6.7), which consists of only six atoms, causes fruit to ripen, among other effects. Human insulin, in contrast, is a protein made up of 51 amino acids.

Animals have a second communication system that is electrical, the nervous system. Chemical signals cause an electrical charge to travel along the extensions of specialized nerve cells (Figure 6.8). At the end of this extension, the electrical charge causes the release of a new chemical signal that can either stimulate another nerve cell to continue to pass the signal along, or can stimulate other types of cells, such as muscle cells. These other types of cells can also initiate signals and pass them to nerve cells, such when light enters the eye. This electrical communication system has several key differences from chemical systems. First is the speed of response. In humans, a signal (the sight of a rock in front of you) can pass from the eye to the brain triggering a change in the movement of your legs in just 0.02 seconds. Second is the complexity of the response. Catching a ball thrown to you requires coordinating the movement of all of your limbs as well as your visual system. Your eyes track the ball's movement sending a signal to your brain which then translates the movement into a trajectory. A further signal to your legs causes you to move towards where the ball is heading, while other signals move your arms and hands to catch the ball. Third is the flexibility of the response system. The same system can let you catch a ball, throw a ball, and run an obstacle course. Fourth is the potential for learning, a topic that we will explore later in this chapter when we consider behavior. One trend in the evolution of animals is the development of more complex nervous systems and the centralization of control (i.e., the brain; Box 6A).
Although she does not have the popular recognition of other women scientists such as Marie Curie or Rachel Carson, Rita Levi-Montalcini was a critical figure in the development of neurobiology in the latter half of the 20th century. At 103 upon her death, she was the only Nobel Prize Laureate to reach the age of 100. Her life also demonstrates a fierce determination to continue her science no matter what the obstacles.

Born on April 22, 1909 in Turin, Italy along with her twin sister Paola, Rita was part of a remarkable Sephardic Jewish family. Although she was raised in a traditionally Victorian setting, with all decisions being made by her father as head of the household, her childhood was filled with love and devotion. Her parents, both highly cultured themselves - her father, Adamo Levi, was an electrical engineer and mathematician, while her mother, Adele Montalcini, was a painter - instilled in all their children a deep appreciation of intellectual pursuits. Rita’s older brother Gino became one of the best-known architects in Italy and a professor at the University of Turin, while her older sister Anna carried a great enthusiasm for the Swedish writer and Nobel Laureate Selma Lagerlöf; as children, Anna infected Rita with her admiration so much that she wanted to be a writer and create an Italian saga “a la Lagerlöf.” Rita’s twin Paola inherited her mother’s artistic passions, showing great talent for painting early in life, and dedicated herself to artistic pursuits, becoming one of Italy’s most acclaimed female painters.

Despite his appreciation for intellectual pursuits, Adamo discouraged his daughters from pursuing professional careers, fearing it would make them unsuitable for marriage. Although he loved his daughters and respected women, he initially forbade them from enrolling in the university. However, at 20 Rita realized she could not adjust to a traditional feminine role as envisioned by her father, and asked for permission to pursue her own career. Adamo granted it and Rita quickly set to filling in the gaps in her education, finishing the necessary studies in Greek, Latin, and mathematics and graduating from high school in eight months, then enrolling in medical school in Turin.

At the university, she was a student of the famous Italian histologist Giuseppe Levi (no relationship), along with her friends Salvador Lucira and Renato Dulbecco, both of whom also were awarded the Nobel Prize. From Levi, all three learned a rigorous approach to scientific problems and received excellent biological training. Rita graduated summa cum laude from medical school in 1936 with a degree in Medicine and Surgery; she then enrolled in a three-year specialization in neurology and psychiatry, uncertain whether to devote herself fully to the practice of medicine or to pursue basic neurological research as well.

While she was pursuing knowledge and her budding scientific career, world events were conspiring to make her choices for her. In 1938, Mussolini issued “Manifesto per la Difesa della Razza,” the Manifesto of Race. Closely modeled on the Nuremberg Laws of Germany, the Manifesto declared Italians to be descendants of the Aryan race, stripping non-Aryans (in particular, Jews and people of African descent) of Italian citizenship and barring them from academic and professional careers. While this was happening at home, Rita was at a neurological institute in Brussels, Belgium, and returned to Turin on
the verge of the German invasion of Belgium in the spring of 1940. Reunited with her family, their options were limited; emigrate to the United States or pursue activity that was isolated from mainstream Italian society. They chose the second path, and Rita built a small research unit installed in her bedroom. She began a course of neurological research on chick embryos, taking inspiration from a 1943 article by experimental embryologist Viktor Hamburger reporting on the effects of limb extirpation on chick embryos. Soon her old teacher Giuseppe Levi, having escaped Nazi-occupied Belgium, returned to Turin and joined Rita as her first and only research assistant.

Turin came under heavy bombing by Allied forces in 1941, which forced the Levi-Montalcini family to abandon their home and retreat to a country cottage in Piedmont. There, Rita rebuilt her laboratory and continued her experiments, but again had to abandon her home and work when the Nazis invaded Italy in the fall of 1943. She and her family fled to Florence, where they made contact with friends in the Italian resistance, the Partito di Azione, and were able to go into hiding until the end of the war. In August 1944, the advance of Allied forces drove the Nazis out of Florence, and Rita was hired by the Anglo-American headquarters as a medical doctor and assigned to a refugee camp. The occupants came from northern Italy where the war still raged on, and although Rita did her best to stave off death, she acted alone against great suffering as both nurse and doctor. Infectious diseases ran rampant in the camp, and abdominal typhus decimated the refugees.

When the war ended in May 1945, Rita returned with her family to Turin, and resumed her academic positions at the university. Two years later in the fall of 1947, she received an invitation from Viktor Hamburger to come to Washington University in St. Louis and repeat her experiments with chick embryos. Although she only planned to be in America for a little less than a year, she postponed her return in the face of the results of her research. While there she focused on determining the molecules responsible for cell growth and differentiation. At the time, although scientists had been describing patterns of development of animals for more than a century, the mechanisms controlling those patterns were not understood.

In 1952, she and her colleague Stanley Cohen decided to focus on certain cancerous tissues that cause extremely rapid growth of nerve cells. They discovered nerve growth factor (NGF), a small protein essential for the development and differentiation of many types of neurons. NGF was one of the first growth factors described; we now know that there are hundreds arranged into dozens of related types. The discovery of NGF was an important milestone in the formation of neurobiology, of which Levi-Montalcini is recognized as a pioneer. Beyond that field, these discoveries, along with other scientific advances, transformed the field of embryology – descriptions of embryo development – into the modern science of developmental biology that is rapidly working out the genetic and biochemical mechanisms of those developmental patterns.

Four years later she was offered and accepted an Associate Professorship, and became a full professor in 1958, a position she held until her retirement in 1977. In 1962 she established a research unit in Rome, and began dividing her time between Rome and St. Louis. In 1968, she became the tenth woman elected to the United States National Academy of Sciences, and a year later became the Director of the Institute of Cell Biology of the Italian National Council of Research in Rome; although she retired from directorship in 1978, she remained at the Italian National Council of Research as a guest professor in 1979.

In 1986 she received the Nobel Prize in Physiology or Medicine for her 1952 discovery of NGF. In 1987, she was awarded the National Medal of Science, the highest scientific honor in the United States. In 2001 she became a Senator for Life in the Italian
Senate, and continued to actively participate in Upper House discussions when not busy with academics. Due to her support of the government of Romano Prodi, she was often criticized by some right-wing senators, who accused her of saving the government when its majority in the senate was at risk. She was also frequently insulted in public and on blogs since 2006 by center-right senators such as Francesco Storace, as well as far-right bloggers for her age and Jewish origins. Levi-Montalcini died on December 30, 2012.
Localization, modularity, and emergent properties

Except for the very simplest species, all multicellular organisms have bodies that have some sort of organized structure. The extent of organization depends on the number of cell types and can be as little as segregating the cells into an outer and inner layer (Figure 6.2), or as much as a system of tissues and organs, each with its own specific, specialized function (see Figure 1.3). What all of these species have in common is that the specialized cells are separated or localized (Table 6.2, principle 3). Usually we think of this separation in terms of space—existing in different places in the body—but it can also be separation in time—being produced during particular times in an individual’s life cycle (Figure 6.21). The body of a multicellular organism can be thought of as a system of systems. Interactions by means of cell-to-cell communication allows those systems to stay coordinated, resulting in the maintenance of organismal integrity.

An important result of localization is modularity (Table 6.2, principle 4). A module is a functional and developmental unit, such as a leaf or an arm. Modularity allows for more extensive differentiation and the separation of functions, and also moderates bodily integration. The necessity of this moderation is twofold. On the one hand, integration enhances coordination. On the other hand, at times independence of parts is beneficial. In plants, for example, each leaf is a separate module. The rate of photosynthesis in a leaf is a function of the local conditions surrounding it (e.g., amount of light, wind speed) and is independent of the rate in other leaves. This means that a plant that is partly shaded can still take in enough CO₂ to survive, as long as the other parts receive sufficient sunlight. In a system like this, the carbon fixed in one leaf can be shunted to others as needed. In your own body, an injury to one limb does not prevent the other limbs from functioning.

A second important outcome of modularity is that it allows for parts of an organism to evolve independently. As we discuss later in this chapter, the development of an organism involves the complex coordination of cell differentiation and movement. If each module can develop independently, then evolutionary changes in one part can happen without affecting other parts. For example, insects consist of a series of segments (modules), each of which is specialized for a different function (Figure 6.9). Insects evolved from a segmented organism in which each segment was originally identical, and the independence of those segments allowed for the evolution of the highly differentiated insect body we see today.

The existence of multiple cell types and their spatial or temporal localization allows for emergent properties (Table 6.2, principle 5). These emergent properties can be found in just part of an individual (the ability to see with an eye) or, they might be the
property of an entire individual (your ability to walk down the street). The mechanics of walking requires the coordination of your skeletal, muscular, and nervous systems, and the energy needed to create that movement comes about through a metabolic system that must take in and store energy until it is needed and then expel the waste products. The coordination of those systems from moment-to-moment, from day-to-day, and across an entire lifetime results in an individual. In this way, individuality is an emergent property of organismal systems.

**The Dynamic Individual**

An organism is not a static entity. In order to maintain its structure and function, an organism is highly dynamic (Table 6.1, principle 3), and that dynamism is expressed in many different aspects of an individual. Consider body temperature. Species use many different ways to either regulate body temperature or respond to changes in temperature. All organisms face the same challenge: metabolic processes limited to working within a range of temperatures. At the lower end is the freezing point of seawater (-1.8°C) as all enzymatic functions require liquid water. At the upper end are temperatures around 60°C, although one species of worm that lives near deep-sea hydrothermal vents can thrive at 80°C, and a few Archaea are able to flourish in water that is close to the boiling point (100°C).

An examination of the chemical activity of an enzyme shows that it has a characteristic temperature-response curve. For an individual, the sum of the response curves of all of its enzymes helps shape a similar looking curve for the functional response of the entire organism (Figure 6.10). When an organism is faced with temperatures at the fringes of or beyond the range within which its enzymes can function, several responses are possible. The organism might modify its biochemistry so that the existing enzymes can still function (e.g., adding chemicals that prevent freezing) or, in a few cases, by producing different enzymes with a different temperature profile. Otherwise, the organism must go into some form of stasis and wait out the unfavorable conditions.
An alternative to simply responding to changes in the surrounding temperature is to use mechanisms that maintain a constant body temperature, or at least moderate any temperature changes. One simple way to moderate temperatures is through behaviors by which an individual moves between warmer and cooler areas, out into the sun when it is cold and into the shade when it is hot. Aquatic organisms move up and down in the water seeking cooler or warmer spots. Even some plants, which cannot move their entire bodies, will change the orientation of their leaves, moving them so that they catch more sunshine or letting them droop to catch less.

Then there are metabolic mechanisms of temperature regulation. Birds and mammals maintain a nearly constant body temperature by converting chemical energy into heat energy, i.e., “burning” carbohydrates and fats. When you are cold, shivering warms you because you are generating heat from your muscles (even though your muscles are not moving one or more parts of your body) thereby using stored chemical energy. Animals, such as reptiles, that do not maintain a constant body temperature can also use metabolic energy to moderate body temperature. Even some plants generate heat (Figure 6.11). At the other end, under hot conditions, land-dwelling species (both plants and animals) can lower their body temperatures by sweating, letting heat be carried away from the body with the water vapor.

Body temperature is just one aspect of an individual that is regulated through dynamic mechanisms. Individuals must ensure that the energy used for maintenance and growth is replaced by new energy. Food is consumed and waste products are expelled. This entire dynamic is regulated by elaborate feedback mechanisms, such as the two hormones that moderate blood glucose levels as we described previously.

Such a system is typical in that it consists of two processes, each of which moves the system in opposite directions, together maintaining the system in dynamic equilibrium. Such systems can be relatively simple or they can be very complex; they can be found everywhere from Bacteria and Archaea to the largest multicellular organism. Consider the simple, two-molecule system that regulates the use of lactose by the bacteria E. coli. Lactose is a sugar consisting of a molecule of glucose and a molecule of galactose (Figure 6.12). The first step in using lactose as a food source is to break down this molecule into its two parts, which is done by the enzyme β-galactosidase. It would be
energetically costly for an individual to produce this enzyme if lactose were not present. So, the gene for β-galactosidase is regulated by two molecules, one that represses production in the absence of lactose and another that stimulates production in the absence of glucose (which serves as an indicator that lactose needs to be broken down). The amount of the enzyme is then maintained as a dynamic equilibrium controlled by the balance between these two molecules.

These dynamic equilibrium systems also can be quite complex. For example, in human females the system that coordinates the monthly release of eggs involves five hormones produced by three different organs. The reaction times of such systems vary from a matter of milliseconds (when your eye is exposed to bright light the iris will contract immediately) or over many days or weeks (a plant will allocate more growth to roots when water is scarce). Some of this dynamic is driven by internal changes, such as during natal development in vertebrates when the cells that will become eyes induce adjacent cells to form the lens. Some of the dynamic is simply due to the growth of an individual. Much of it occurs, however, in response to the changing environment, a topic that we address later in this chapter.
Box 6B

Graham Lusk

Graham Lusk was one of the central figures in the development of our current understanding of mammalian, especially human, metabolism, work that led to our current understanding of metabolism as an integrated system. His work was anchored by a core assumption that basic metabolic rates were constant and that reports of fluctuations in rates were due to poor measurement techniques. This assumption was coupled with a methodology focused on rigid control of experimental conditions. However, he coupled this rigid methodological approach with a flexible intellectual appreciation, and was quite willing to accept new ideas as data warranted.

The son of a physician, Lusk was prevented from following in his father’s footsteps by a hearing impairment, instead focusing on research in the medically-related fields of nutrition and metabolism. Born in Bridgeport, Connecticut on February 15, 1866, Lusk focused his life’s study on the science of physiological chemistry, attending the Columbia University School of Mines, earning a Ph.B. degree in 1887, and immediately thereafter moving to Germany to study under the leading physiologists of the time. He spent four years abroad, first studying under Carl Ludwig at Leipzig, then in Munich under Carl Voit. After receiving a Ph.D. from the University of Munich, he returned to the USA to become an instructor of physiology at Yale Medical School. For his $30-a-month salary, he had to take on his own janitorial as well as instructional duties.

Lusk’s deep passions for his work would bring him an assistant professorship within a year, and make him a full professor by 1895. He accepted a position at the University and Bellevue Hospital Medical College in New York City in 1898, where his father had previously been a professor of obstetrics, then moving to the Cornell University Medical College in that same city in 1909, where he stayed for 21 years, retiring only a few weeks prior to his death. In 1899, he married Mary Woodbridge Tiffany, the eldest daughter of the famous designer Louis Tiffany, with whom he had three children. Their summer home on Long Island was adjacent to the Tiffany estate where groups of scientists meeting at their home would be treated to a tour of the beautiful collection of art glass.

Cornell is where the bulk of his research occurred. Early in his career there, he helped construct a device for measuring the heat of chemical reactions within the body (and thus metabolic processes), called a respiration calorimeter, designed for use on dogs and small babies. The instrument was designed by H. B. Williams, who also supervised its construction; he, Lusk and another colleague, J. A. Riche, operated it. This instrument exemplified Lusk’s approach of tightly controlling experimental conditions. The resulting collaboration led to a classic 42-paper series entitled “Animal Calorimetry.”

His most well-known and enduring work was The Science of Nutrition, first published in 1906. It went through four editions and was the standard text on metabolism for decades.

Most of their studies were in what was known at the time as specific dynamic action, which is referred to today as the thermic effect of food, or TEF. TEF is the cost of increase in energy expenditure above the body’s resting metabolic rate created by the body processing food for use and storage. Some of their studies looked at other factors related to energy metabolism as well. One of Lusk’s most significant findings was the fact that diet has no effect on the number of calories burned during physical work; when
he measured the caloric output of well-fed dogs who had received glucose (and were thus using food for energy) and the output of dogs who had been starved (using their stored body fat for energy), both running on treadmills, the output was the same for both groups of dogs. He also showed that basal metabolism, under normal conditions with a steady, uniform diet, is constant, using measurements taken from one of the lab dogs, Lady Astor, over a period of several years, as well as himself as his colleague Eugene DuBois.

He could be very forceful in defending his scientific views, often taking the position that contradictory studies were likely to be faulty in execution and data collection. For example, a study performed by colleague and later collaborator E. DuBois of the metabolism of a person with diabetes seemed to show variation in metabolic rate. Careful checking discovered that relatives had been bringing food to the hospital-bound patient; two sandwiches were found secreted in the bed, and two more were hidden in his armpits. Over the course of several decades, the high regard that others had for his experimental technique along with his strong questioning of contrary results led the rest of the community to firmly accept his view of the constancy of basic metabolic rates. However, he was always willing to let contrary opinions be aired. In one case, a junior scientist from Poland spent a year in his laboratory. Based on his study of fasting dogs done in Lusk’s laboratory, he came to an opposite conclusion. Rather than block the publication, as he could have as head of the lab, he was quite willing to let the colleague publish his findings with a footnote that Lusk disagreed with the interpretation of the data. Later he helped the scientist get a job in Poland. Lusk was highly regarded by his peers, considered very generous and kind, even going so far as to raise money to cover the travel expenses of his European colleagues to the International Physiological Congress in Boston in 1929 and hosting many of them at his home.

He also was known for the honesty in his convictions. He scrupulously avoided even the appearance of impropriety, at one point refusing to invest in Corn Products stock, despite a tip that it was poised for an enormous gain in value, so that his future work with glucose wouldn’t be questioned. That attitude could be a lesson to today’s community of biomedical scientists who have been wrestling with just such conflict-of-interest issues. Lusk was also years ahead of his time in calling for corporate accountability, advocating the creation of nutrition labels for packaged food products, so that “…the citizen should know how best to maintain [his body] at a maximum of efficiency. Not only that, but in times of trouble he should know where to turn to find nourishment in the form which is best and cheapest. Who will give him this information? Will the manufacturer? …No, not unless forced to do so…. By being compelled by law to label his can: ‘This contains X calories of which Y percent are in proteins of grade C....’” His public efforts included service during WWI as one of the two American representatives to the Interallied Scientific Food Commission.

In addition to experimental work, Lusk was incredibly active in the scientific community, helping to found and organize most of the early societies focus on metabolism and human health, including the Society for Experimental Biology and Medicine in 1903, the American Society of Pharmacology and Experimental Therapeutics in 1908, the American Institution of Nutrition in 1928. He helped found the Harvey Society in 1905 which was dedicated to the diffusion of scientific knowledge to the medical community. He joined the American Physiological Society at its fifth annual meeting in 1892, and was a charter member of the American Society of Biological Chemists, serving as its chairman in 1914. Recognition for his scientific achievements and service efforts included honorary degrees from Yale University and the University of Glasgow, fellowships in the U.S. National Academy of Sciences and the Royal Society of Edinburgh, and being made
a foreign member of the Royal Society of London. On July 18, 1932, Lusk died at the age of 66.
Organisms in Their Environment

External interactions

Organisms interact with their external environments in a myriad of ways that are both causes and consequences of their need to maintain bodily integrity (Table 6.1, principle 4). In the next section we look in detail at the flow of materials and energy in and out of the body. Here we will consider organismal-environment interactions in terms of the flow of information.

Organisms obtain information from the environment through a variety of types of sensors. Like internal communication systems, many of those sensors rely on chemical or electromagnetic energy. Human senses of taste and smell are chemical, while sight is electromagnetic (light energy). We can sense heat, another form of electromagnetic energy, and we also respond to pressure. Hearing occurs when variation in air pressure moves our eardrum, ultimately stimulating nerve cells (Figure 6.13), and touch is another form of pressure sensing.

All species respond to electromagnetic and chemical cues in some fashion, but these are not the only forms of information. Birds that migrate long distances use magnetic information for orientation, which has been found in organisms as wildly different as Bacteria, insects, fish and cows. A recent study using remote sensing images found that when grazing or resting, cows and deer tend to face in a north-south direction, but oriented towards magnetic rather than polar north. Organisms also respond to gravity in a variety of ways. Your sense of balance involves the pull of gravity on tiny hairs inside your ears, creating movement which then stimulates nerve cells. In the roots of plants, gravity pulls starch grains to the bottom of the cells causing the roots to grow downward.

Information can also come from signals among individuals of the same species. Chemical communication goes back to the earliest single-celled organisms, and is still found in the interactions of cells within our bodies. Even complex organisms communicate through chemicals, for example the scents used by many mammals to mark their territories. The release of chemicals by plants that have been partially eaten can cause nearby plants to produce chemical defenses. Many animals communicate through sound and sight. Some species produce light through chemical reactions (e.g., lightning bugs, Figure 6.14), but most sight-based communication simply uses ambient light.

This information is used by an organism in various ways. In addition to regulating body temperature and orientation, it allows individuals to find food. When put in a medium with a higher glucose concentration in one direction, Bacteria will move in that direction. Wolves will track their prey by following a scent (i.e., chemical) trail.

All organisms respond to environmental cues, reactions which are often called behaviors. Complex behaviors in organisms like mammals are easy to observe. For
Materials and energy

All organisms, including humans, must eat to live (Table 6.1, principle 5). You have to take in matter to grow and reproduce and need energy to process that matter. These functions are at the heart of the constantly-changing nature of living organisms, and there is a variety of ways in which they happen. As discussed in Chapter 5, photosynthetic organisms take in energy in the form of light. All other organisms take in energy in the form of chemical bonds, primarily as organic molecules. Even this single task can be accomplished in a variety of ways.

Eating consists of two actions, taking in matter and breaking that matter down to small pieces that can then be used for various functions. In humans and many other animals, the breaking down happens in several steps and locations, both inside and outside the body. We would not eat a cow or an orange whole; first we cut them up into small pieces before we put them in our mouths. Our teeth continue the break-up process, which is finished in our stomach and intestine (Figure 6.16). Snakes, which have no hands or claws and only two fangs, do consume prey whole, with all of the breakdown happening in the gut. In a similar fashion, many single-celled organisms, such as amoebae, eat by taking in smaller organisms whole and then breaking them down in specialized vacuoles. Fungi do the reverse, first breaking down large items by excreting digestive enzymes, then absorbing the resulting small molecules.

One common need in all of these situations is to keep digestive enzymes and other chemicals sequestered so that the organism does not start digesting itself. Inside cells, enzymes that break down molecules are kept in special vacuoles. In humans, the lining of the stomach is specialized to be able to contend with the very acidic fluids it contains. Ulcers occur when there is a break in that lining, allowing those acids to attack other cells. Fungi, which do not have specialized organs for digestion, move the task outside their bodies.

Organisms obtain energy and matter from a variety of kinds of diets which go by a long list of names. We can reduce that long list to a very short list of general categories, noting that a given organism or species may make use of more than one category. First, does the energy and matter come from strictly inorganic sources, or from organic...
molecules? Second, if it does come from organic molecules, does it come as part of a whole organism or in smaller forms that are already partially broken down? For example, many Bacteria, Archaea, and fungi take in organic molecules from organisms that have previously died and started to decompose.

Third, does the organism consume its food live, does it kill its food, or does it eat food that is already dead? This distinction is not always clear. For example, a seed is alive, but dormant, when it is consumed. Humans, for the most part, eat food that is already dead, but not yet breaking down. We would not find a decaying carcass very appetizing, although a vulture would. Disease-causing organisms eat their food from the inside out.

Fourth are distinctions based on what kind of food is being consumed. Is the organism in question one which only takes in inorganic matter (i.e., plants), or is it an organism that eats other organisms? Does it have a mixed diet of both inorganic and organic matter, and if the latter, what kinds of organic matter does it consume? These distinctions are especially important in defining the ways in which individuals of different species interact with each other (see Chapter 7).

Just as individuals must take in energy and matter, they must also give off energy and matter. All organisms give off energy in the form of heat. This is obvious in birds and mammals, such as humans, that maintain body temperatures greater than their surroundings. But even animals such as fish or insects, as well as plants, fungi and Bacteria, give off some heat energy generated by the chemical reactions in their cells. Organisms also give off waste products in the form of solids, liquids or gases, generally molecules that can no longer be used for metabolism. Some molecules, especially water, are expelled because they are the medium by which other molecules are carried. It is through the balance between intake and outflow that organisms maintain their structure and integrity.

**Trade-offs**

No organism is able simultaneously to do everything in an optimal fashion necessary for survival and reproduction; because of this, life is a series of trade-offs (Table 6.1, principle 6). Some types of trade-offs are inevitable, such as allocation of resources or time. When an individual consumes food, that food can be used to simply maintain its body, to grow, or to reproduce. Resources used for one purpose are unavailable for the other. Similarly, there are only 24 hours in the day. Time spent foraging for food is time...
not available for searching for mates or guarding against predators. While this sort of trade-off may seem obvious, why an individual divides up the day in a particular way is not always clear. Consider a kangaroo rat, a small desert rodent (Figure 6.17). It has to forage for seeds while avoiding predators like hawks, owls, and coyotes. It can dig for seeds underneath bushes where it is generally protected from predators. As it depletes the seeds under one bush, it must make a decision: should it keep digging or should it try to find another bush? It takes time to find another place to dig and to move to that location, time which is not spent digging for food. Moving exposes the animal to predators, creating the possibility that it might become food for another animal. So, there is a complex calculation involving the current rate at which it is finding seeds, the possibility of increasing that rate somewhere else, the time used to find that new place, and the chance of being eaten on the way there. Of course, kangaroo rats are not consciously weighing all of these factors. Rather, the process of natural selection has resulted in animals that use a series of cues that guide behavior: how much time has passed since the last seed was found, how hungry the animal is, whether the moon is out that night or the night is dark. Different individuals (or other species of small rodent) will move or stay depending on its particular circumstances and conditions.

Because there is almost never a single optimal solution in a given situation, different species have evolved to employ different solutions. For example, land plants are faced with a dilemma. Photosynthesis requires that they take in CO2 from the air (see Chapter 5). Most plants do this through special pores in their leaves. But opening those pores to let in CO2 also means that water can get out. If there is abundant water in the soil, the plant can replenish any water lost. If water is limiting, however, the plant has to trade off water loss with carbon gain. Different species deal with this trade-off in a variety of ways. Some plant species are more tolerant of water stress than others, but they achieve this tolerance by dedicating resources to special proteins and structures. Some species grow quickly and only when water is abundant, thereby avoiding the problem. Slower-growing species may shut their pores in the middle of the day when temperatures are highest and water loss greatest. Some species, notably cacti (see Figure 7.1), take this strategy to the extreme. They open their pores only at night, completing only the initial capture of CO2, and storing the molecules in special vacuoles. Then, during the day when light energy is available, they complete the photosynthesis process. While this solution avoids water loss, it means that cacti must grow very slowly, as they can take up only small amounts of CO2 in a given day.

Trade-offs can also be circumvented in a variety of ways. The vast majority of animals are bilaterally symmetric, meaning that the left and right sides of their bodies are almost mirror images of each other. In some species that symmetry is broken, leading to ways to avoid trade-offs. For example, lobsters (Figure 6.18) feed on a wide variety of other animals, including crabs, mussels, and clams. Eating these animals requires two steps; first the shell must be crushed, and then the animal inside must be cut into small pieces so that it can be consumed, tasks done with the front claws. Claws that are optimal for crushing would not do as well at cutting, and visa versa. In most species of lobsters, the claws are intermediate between the optimal design for either chore. But in the American lobster, a different solution has evolved. Their left and right front claws are not
Figure 6.18
An American lobster. Note the difference in size and shape of the left and right claws. In this individual, the right-hand claw is the crushing claw, others may have it on the left. (Photo by Roberto Rodríguez, Source: Wikipedia)

Figure 6.19
Seven leafcutter ant workers of various castes (left) and two Queens (right) These individuals are siblings. The extreme difference in size is due to the diet of the larva, not genetic differences. (Photo by Sarefo, modified by GameKeeper; Source: Wikipedia)

Mortality

Death is inevitable (Table 6.1, principle 7), even more so than taxes. Life is an intricate and complex system that requires constant change to maintain itself, and death is the disruption of that system.

But why is death inevitable? Although this question has been addressed in a very different way by philosophers and religious thinkers, biologically speaking death comes about through environmental interactions. Some causes of death are easy to explain. Because an individual requires matter and energy to maintain itself, any individual can be starved to death. Less obvious is why organisms are not invulnerable; any organism might die as the result of predation, stress, trauma, or starvation. The vulnerability of an organism is inherent in its being a dynamic system, which can be disrupted. That disruption can be internal – the ingestion of a substance that interferes with biochemical processes, substances that we conventionally call poisons – or it can be external (being eaten by another organism or having a rock fall and crush you). The list of possible ways to disrupt a living system is very long. For example, coniine, which has a chemical structure that is similar to nicotine, comes from the hemlock plant; it works by blocking the transmission of signals from nerve cells to muscles, causing paralysis and death from lack of breathing.

Only an inert object can potentially avoid external change. Even if an inert object like a rock can be eroded away, we can imagine an object that is so hard that only an extreme event, such as being dropped into the sun, would destroy it. Such an object
could never be alive. The vulnerability of organisms appears to be inherent in the carbon-based metabolism underlying all life found on the Earth. It may also be true of any living system – whatever its chemical or energy basis – if living systems must be dynamic, as any dynamic system can be disrupted.

This principle does not mean that all organisms senesce. **Senescence** is an increase in the probability of dying with age. Humans senesce. At about age 60 the probability of dying in the next year dramatically increases (Figure 6.20). Careful studies of mortality rates have found evidence for senescence in a variety of species, including other primates, mice, snakes, fish, bees, fruit flies, and roundworms, indicating that it may be a common feature of most animals. However, it is very hard to know for sure because it is not easy to study. To show an increase in the rate of death at old age, you have to follow a very large number of individuals, since at the oldest ages only a small percentage will remain in the study. We have these data for humans because of the extensive census and death records that are kept by many governments. For other species, experiments have been performed in which thousands (mice) or hundreds of thousands (fruit flies and roundworms) of individuals are followed from birth to death. These species have the advantage of being both small and short-lived making such experiments possible in a laboratory setting. There is evidence, though, that some simple animals do not senesce.

![Figure 6.20](image)

Actual and estimated human survival curves in England and Wales. Each curve is a group of persons born in 1851 to 2031 at 20-year intervals. The median life span for each group is the age at which 50% of individuals are still surviving. Survival probabilities increased dramatically between the middle of the 19th and 20th centuries and more slowly since then. Maximal lifespan has increased to a much lesser extent. (Data from: National Population Projections, 2010-based reference volume: Series PP2, UK Office for National Statistics)

The few studies done with long-lived plants have failed to find evidence of senescence. It may be that plants, which do not separate germ line cells from somatic cells (see below), work in different ways from animals. Even among plants, though, patterns of mortality vary. Many species grow for some set period of time, which may be less than a year, two years, or several years, then reproduce once and die. Some species of trees may live for centuries or millennia, but eventually succumb. However, there is no evidence that the probability of death changes with age. It is just that, for example, if your chance of dying in a given year was 1 in 100, eventually the odds would catch up with you. That all individuals eventually die does not imply senescence.

As for other species, few data exist. There have been no studies of senescence in multicellular fungi. As for single-celled organisms, the entire concept of senescence may not apply. One possible cause of senescence is that individuals have a limited number of cell divisions, effectively limiting their lifespan. Single-celled organisms, however, are continually dividing in order to produce new individuals. They cannot have a limited number of cell divisions, otherwise they would go extinct. Thus, if senescence in animals is related to limits on cell division, there must be something fundamentally different
about how it occurs, perhaps linked to the process of cell specialization. More research is needed to answer this question.

But why should any organism senesce? We do not yet have a good answer to this question, despite the extensive research that has gone on because of our vested interest in the answer. There is some evidence that our cells are able to divide only so many times, although why is unclear. Some research points to imperfections in chromosomal replication, but that has been called into question. Undoubtedly research into the mechanisms of aging and senescence will continue, given the possible practical applications of such knowledge.

There are two evolutionary theories to explain the existence of senescence. The fitness of an individual with regard to natural selection (see Chapter 4), has two broad components: survival and reproduction. To be favored by natural selection, an individual must survive long enough to reproduce, and the longer it survives the more reproduction must be possible. One explanation for senescence is that if a mutation occurs beyond the age of reproduction, it will not be purged by natural selection because it has no effect on fitness (i.e., the number of offspring). The result is an accumulation of late-life harmful mutations that cause senescence. The other explanation is one of trade-offs, that vigor and high fecundity early in life inevitably lead to decreased vigor late in life. There is some evidence for both mechanisms in different species. However, both of these theories are still unconnected to the mechanistic factors described above, so more research is needed.

**Change over time**

**Life history**

All organisms are born, grow, reproduce, and die. That series of events is called a **life history**. How those events unfold for any given species or individual varies in myriad ways and is determined by a combination of the intrinsic biology of the individual and how it interacts with its environment (Table 6.1, principle 8). Those environmental interactions play a role in two ways. For a given individual they determine how much food it eats, how factors such as temperature affect its metabolism, how successful it is in finding a mate, and so forth. For a given species they determine the process of natural selection (see Chapter 4), thereby shaping the broad outline of a species’ life history within which a given individual’s life history is played out.

Life histories come in many different lengths and patterns (Figure 6.21). For most single-celled organisms they are relatively simple. The process of cell division wraps together birth, growth, and reproduction. The rate of cell division is typically limited by how fast the individual can take in matter and energy and build new structures. Even for such species, though, life histories can be more complex. The environment changes over the course of a day or a year (see Chapter 7). For a single-celled organism, 24 hours can represent several generations. There may be a pattern to cell division, perhaps occurring only during the day or only during the night, depending on the particular circumstances of that species. Over the course of a year, if the species lives where there is a pronounced summer and winter or a wet season and dry season, the environment may be too harsh at certain times for an individual to survive with an active metabolism. In that case, an individual may enter some sort of state of arrested metabolism. Many single-celled organisms encase themselves in special structures as part of this period of quiescence.

Life histories get more complex for multicellular organisms. Life always starts with birth and ends with death, but in between growth and reproduction can combine in many different ways. Some species just grow until they reach a final size and then
simply maintain themselves while they reproduce. Reproduction may happen in a single bout quickly followed by death, or it may be spread across multiple bouts. Other species alternate between periods of growth and reproduction. Still others grow continuously while they are also reproducing. All of this may happen within a single year or be spread over multiple years. Additional complexities come about if an individual substantially changes its form over the course of its life (Figure 6.21B,C). Then there are all of the various ways that organisms actually reproduce, a topic that we explore in detail in the next section.

Figure 6.21
Idealized life cycles of a flowering plant (A), an insect (B), and a frog (C). (A) The sporophyte – the plant with which we are familiar – is the diploid multicellular part of the life cycle. It gives rise to a haploid, unicellular spore by meiosis, which grows into a haploid, multicellular gametophyte – either a pollen grain or an ovule – by mitosis. At fertilization, a diploid, unicellular zygote is formed, which by mitosis grows into a multicellular embryo. The endosperm is triploid, formed by the union of a sperm cell with a binucleate cell. Tissue in the seed coat comes from the maternal sporophyte. (Created by Mariana Ruiz, Source: Wikipedia) (B) The diploid adult butterfly produces haploid sperm and egg cells that unite to form diploid eggs. The egg hatches into a larva that grows through several stages by molting. The larva spins a cocoon within which it changes into its adult form. (Created by Nicholas Cafarilla, Source: Wikipedia) (C) A diploid adult frog produces haploid sperm and egg cells that unite to form diploid eggs. The egg hatches and the larva grows through several stages during which it gradually changes its form to that of the adult. (Created by Mariana Ruiz, Source: Wikipedia)
To see some of this variation in life histories, we consider a few specific cases. We start with humans, although in many ways humans have very unusual life history patterns. For the most part, contemporary humans grow to adulthood then reproduce. Humans are capable of reproducing well before growth is completed, once puberty is reached around the age of 12 or 13, but reproduction is typically delayed for several years. Nor is this delay just a feature of modern societies; evidence exists for such delays in pre-agricultural and early agricultural societies. Interestingly, although human women are capable of reproducing for many years as an adult, most reproduction is concentrated in just a few years. In most families, siblings tend to be fairly close in age. This pattern allows families to concentrate resources on the growth of current offspring, rather than on the production of new offspring, driven probably by the extremely long period in which offspring are dependent on their parents. While many species take a long time to reach the age of reproduction, humans are unusual in the continual reliance of offspring on their parents during that maturation period. Humans are also unusual in their long life after they are no longer capable of reproduction. In most human societies, those post-reproductive individuals help in raising their grandchildren, another manifestation of the long period of offspring reliance. Thus, one aspect of human biology, our long period of growth, serves to shape many other aspects of our social structures as well as biological processes.

As with single-called organisms, multicellular organisms need to avoid periods of harsh climate (very high temperatures, very low temperatures, or very dry conditions) drives many life history patterns. There are three general solutions to periods of harsh climate: avoidance, quiescence, or just being able to tough it out. Humans, like many other mammals and birds that can maintain body temperature metabolically, remain active during periods of low temperature (i.e., winter) and survive by both building up fat within the body during good times and by finding sufficient food during the winter to keep going.

Many animals enter a quiescent state, sometimes called hibernation (e.g., bears), in which metabolic rates are lowered. Many insects survive winter in temperate climes in this fashion, and the quiescent state may be during any stage of the life cycle (Figure 6.21B), depending on the species. Similarly, trees and other plants that live for more than a year enter such a quiescent state. A key to the survival of such species that do not regulate their body temperature (e.g., insects and plants) are biochemical changes that prevent damage below 0°C, the temperature at which water freezes. Because water expands upon freezing, the formation of ice inside cells can cause them to rupture. Biochemical changes can include the removal of most of the water inside a cell or the production of special chemicals that lower the temperature at which ice will form.

Avoidance can take two forms. Many animals simply move away from harsh conditions, such as the many birds and insects that migrate to tropical regions during winter.
Migrations can occur over much shorter distances in response to other types of environmental variation. For example, in east Africa, wildebeest migrate a vast circuit of more than a thousand kilometers in response to changes in rainfall, so that they will always be in areas with sufficient grass to eat (Figure 6.23).

Another form of avoidance is akin to hibernation, but involves having a life history stage that is metabolically inactive. Plants cannot move away from harsh conditions, so some species avoid harsh conditions as seeds in the soil. Sometimes this is combined with very rapid growth: some desert-dwelling plants, where water is available for a short period of time, can go from a germinating seed to the production of new seeds in as little as six weeks.

Another way to avoid harsh conditions is to have a life history that involves the ability to live in two very different kinds of habitats. Frogs and toads typify this type of life history (Figure 6.21B). Some live in temporary ponds that dry out quickly. In those species, they can go from egg to air-breathing adult in just a few weeks. As with many species, adult survival is increased in larger individuals, so it is advantageous to remain a growing larva as long as possible, exiting the pond only when conditions are no longer favorable (e.g., the onset of winter). Such species can also speed up the developmental process, for example in response to cues that the pond is getting close to drying out, thus ensuring survival albeit at a smaller size.

Harsh conditions need not just be due to climate. Predators can also take their toll, and in response many animals go through some sort of change in body form. For example, many sea-dwelling animals have a larval stage that drifts in the water before changing form and settling to the bottom (Figure 6.24). The larva are small and easily eaten and the timing of the change to the adult form is determined, in part, by predation rates. Some amphibian species avoid larval predation by going through the entire larval stage inside the egg. The evolution in mammals of bearing live young might have been driven, in part, as a way to avoid predation on eggs.

**Reproduction**

Reproduction is the other component of species fitness, and large parts of an individual’s biology is geared towards reproduction. All species reproduce at some point in their life cycle, a fact driven by the inevitability of an individual’s death. Because no organisms are invulnerable, natural selection has favored individuals whose biological processes devote a large portion of resources toward reproduction, so that the persistent lineages are those that reproduce (Table 6.1, principle 9).

Besides being necessary for persistence, reproduction is necessary for evolution to occur, and is itself also shaped by evolution. This feedback loop is caused by the fact
that evolution requires variation (see Table 4.1, principle 4), which is generated by a combination of mutation and recombination (see Chapter 3). If organisms did not reproduce, variation would not be generated, and the process of evolution and natural selection would grind to a halt. Because of this, the existence of evolution implies reproduction.

Reproduction can occur two ways: sexually (i.e., involving some process of meiosis and recombination) or asexually. Asexual reproduction is the older form, and this is what happens when single-celled organisms divide. But, as discussed in Chapters 4 and 5, even Bacteria and Archaea have processes by which recombination occurs, albeit in a haphazard fashion. Even without recombination, variation among individuals can still be generated through the accumulation of mutations, so that evolution can still happen.

Sexual recombination arose at some point early in the evolution of Eukaryotes and is found throughout Eukaryotic species. Some Eukaryotic groups, however, reproduce primarily asexually, and some do not have sexual reproduction at all. For example, there is a species of lizard (Figure 6.25A) which has repeatedly given rise to asexual lineages. Asexuality has arisen many times in flowering plants. A familiar example are the dandelions that infest many lawns (Figure 6.25B); all of the seeds that it produces are created asexually.

There is a long-standing debate over why some species are sexual and others asexual. Over short periods of time asexuality should be favored because it is more efficient. Consider mammals which have both males and females. If we humans reproduced asexually, we would not need to “waste” resources on the production of males; if all sons were replaced with daughters we could reproduce twice as fast. On the other hand, over long periods of time, sexual reproduction and recombination produces more variation than mutation alone. Changing environments mean that the optimal phenotype will change, favoring sexual lineages that can evolve in response. It may also be that once a lineage has evolved for sexual reproduction, especially the production of separate males and females, it is very difficult for the machinery of reproduction to evolve the means to reproduce asexually. As mentioned previously, asexuality is more common in plants, which may be due to their very different developmental process (see next section). The extent of asexuality varies among groups. As far as is known, there are no asexually reproducing mammals, and asexual reproduction is very rare in other vertebrates. It is more commonly found (although still rare) in other multicellular animals. Some species do a bit of both, reproducing asexually most of the time, but then occasionally
going through a round of sexual reproduction. One of the best examples are Daphnia, also called water fleas (Figure 6.27). Based on genetic evidence, at least one group of multicellular animals may have been strictly asexual for at least 80 million years (Figure 6.28). Such ancient asexuality would contradict the evolutionary explanation for the maintenance of sexual reproduction because they have manage to persist during vast changes in the environment. However, recent data has called into question whether these animals are strictly asexual, so this is very much an open question. It is unclear how common asexual reproduction is in fungi and single-celled Eukaryotes because it is difficult to raise many of those species in the laboratory. Cheap and easy DNA sequencing is allowing scientists to use genetic evidence instead, so we will know more about the extent of asexual reproduction in coming years.

Multicellular species that reproduce sexually have very elaborate procedures for obtaining mates or for ensuring that gametes can find one another. Some species simply release lots of gametes into the surrounding environment. This is a strategy found in corals in the ocean (Figure 6.29), which release both male and female gametes into the water, and plants and fungi on land, which release large amounts of pollen (male gametes) and spores into the air. In both situations the adults are unable to move, so the gametes have to do all of the moving.

Such solutions are very inefficient; many gametes have to be produced in order for a few find each other. Organisms have many different ways to make that process more efficient. Plants do so by utilizing the help of animals that can move: notably insects but also a few birds and bats (Figure 6.30; see Figure 7.8). Much of the variation in flowers has evolved to both attract these animals, and to ensure that the animals will move to other flowers of the same species, thereby getting the gametes to the correct location. For animals that can move, there are procedures that help individuals find others of their same species, for example the color patterns of birds (Figure 6.31), the calls of frogs, or chemicals released by moths.

Just finding each other, though, is not the whole story. In many species there is competition for mates. The number of eggs being laid by females may be limited
Patterns of reproduction add another layer of complexity to life histories. Most land animals and some plants have simple patterns, with individuals developing as either a male or a female and staying that way throughout their lives. But some animals switch from male to female, and this switch might happen once or multiple times. Other animals are hermaphrodites, and may produce both male and female gametes simultaneously or sequentially. Most plants are hermaphrodites although they can also alternate between the production of male and female gametes. Fungi are completely different from animals and plants; their gametes have a number of mating types, and a gamete can combine only with a type different from itself.

The timing and length of the reproductive period adds additional complexity. In general, birth is timed for periods of abundant resources and thus set by the environment. But there can be an extended time between when fertilization occurs and when an animal is born or a seed is germinated, thereby allowing mating to occur under one
set of conditions and birth or germination under a different set. Among animals, if fertilization is internal, the development of the embryo can be arrested for days, months or years. Once development begins, it can happen as quickly as two weeks in a mouse or as slow as 22 months in elephants. A key evolutionary innovation in flowering plants was the shortening of seed development from 1-2 years in pine trees and their relatives to just a few months or weeks. This quicker development both increased seed production and allowed for species that could complete their life cycle very quickly, and thus survive in environments that are only favorable for very short periods of time. The process of development is the subject of the next section.

**Development**

All organisms, both single-celled and multi-cellular, grow. Following cell division, single-celled organisms must acquire resources before they can divide again. Multicellular organisms go through an additional process, development, the process by which cells differentiate into their specialized function and the structure of the body is built. Development, unlike growth, involves cellular differentiation in addition to cell division. Development is often thought of as occurring during distinct periods in an organism’s life, such as between the time of fertilization and birth or until adulthood. However, some types of cell differentiation can occur throughout one’s lifetime, as with cells in human bone marrow that are continually producing specialized red and white blood cells. While we can separate growth from development as categories, the individual processes blend into each other.

We will explore two aspects of development. In this section we look at the process of development and the machinery behind that process. In the next section we consider the relationship between the processes of development and evolution: how differences in development have evolved and how development shapes and constrains evolutionary patterns.

Multicellular organisms have one of two developmental patterns, open and closed. Nearly all animals have closed developmental systems; plants, the simplest animals, and other multicellular species have open development. In differentiating these two patterns, a key difference between is whether, early in development, there is a segregation of the cells that will go on to produce the gametes from those of the rest of the body. This segregation occurs in animals but not in plants. In animals, the cells that will later produce gametes, the **germ line**, become isolated into specialized reproductive organs early in development. In plants, at the ends of stems or roots where active growth is occurring, the **meristem cells** remain unspecialized and able to develop into roots, stems, leaves or reproductive structures that then go on to produce gametes (Figure 6.6).

The concept of the segregation of the germ line was first put forward in the late nineteenth century by the German biologist August Weismann. Such segregation would preclude the characteristics an individual acquired during its lifetime from being passed along to its offspring. Weismann’s ideas were an important milestone in discrediting that notion; it was finally put to rest in the early part of the twentieth century with the development of the science of genetics. Today we understand that even in plants – where this segregation does not occur – the machinery responsible for the use of genetic information does not allow for the inheritance of acquired characteristics. See Chapter 3 for a more detailed discussion of this issue.

Plants and animals differ in another key respect. While animals generally have a fixed adult form, plants generally have a plastic form. That is, a plant not only grows throughout its life, but the form that the plant takes can be easily modified; it may put out more leaves or fewer, or grow an additional stem. In contrast, an animal has either a single, fixed form, or a form which can change only within very narrow parameters. For
example, many fish continue to grow during their lives, but they simply get larger with possibly minor changes in shape from narrow to wide.

Related to this difference in plasticity of form, regeneration is very common in plants. In most species, you can take a bit of root or stem and grow an entirely new plant. While some animals have the ability to regenerate parts of their limbs or tails if lost, only in very simple animals can you regrow an entire individual from a small piece. This difference in regeneration ability is a product of cell specialization.

In animals, nearly all cells in the body become specialized, and once they do they cannot become unspecialized. Regeneration occurs either from unspecialized cells, or from cells that can become unspecialized. The fact that regeneration happens in some animals from previously specialized cells shows that unspecialization is not impossible, just very rare. Today a fierce debate is going on around this issue, although it is not a scientific debate so much as a social debate. Unspecialized cells, often called stem cells, have the potential to be used to cure diseases caused by cell malfunction or death. For example, nerve cells repair or regrow very poorly, and diseases such as Parkinson’s are caused by nerve cell malfunction. Diabetes, as discussed previously, comes about because of malfunctioning cells in the pancreas. Those diseases could be abated if stem cells could be directed to replace the malfunctioning cells. The debate concerns the source of those cells. Currently, the only reliable source is extremely young embryos in which the cells have not yet specialized. People disagree on whether it is ethical to use embryonic stem cells. While some proponents are continuing to fight for the use of unspecialized embryonic cells, research is also being done to determine if it is possible to take specialized cells and make them unspecialized, turning them into stem cells that have the same properties as embryonic stem cells.

Figure 6.32
Plant forms can be grouped into just a few general categories, including trees (A), shrubs (B), and herbs (C) shown here for various species growing on the island of Tenerife, Spain. (Photos: S. Scheiner)

Development among both plants and animals progress in a myriad of different ways, resulting in a dizzying array of forms. Although plants are overall more varied in form, we can group those forms into just a few general categories (Figure 6.32). While the overall size and shape of an individual may be highly variable, even among individuals of the same species, the units that make up that individual are often fixed in their shape or size. A red oak tree may have just a few leaves or thousands of leaves, but all of those leaves have the same overall shape (Figure 6.33). The reproductive structures of plants, i.e., flowers or cones, are often the least plastic in general form and often are the basis by which plant families and orders are defined (Figure 6.34).

Figure 6.33
Leaf of a red oak. (Photo: Ninjatacoshell; Source: Wikipedia)
For animals, differences in form are generally limited to the shape of the entire body (Figure 6.9) rather than involving particular parts or units. Individuals tend to differ in size, shape, or color, rather than in the number of units possessed (Figure 6.35). All insects have six legs, for example. Most land-dwelling vertebrates have four limbs, and while a few have only two limbs or no limbs (e.g., snakes), no species has more than four limbs. Body plans are the basis by which animal phyla and classes are defined. For development to occur in an orderly fashion, some sort of process must exist that determines which cell will specialize into which particular type. This process requires the existence of some sort of variation or asymmetry within the organism (Table 6.2, principle 6). In animals, that asymmetry is established at the time the egg is created. For example, in insects such as the fruit fly, the egg has a front and a back end that are defined by gradients in protein concentration, some with high concentrations at one end and others at high concentration at the other end. Similar gradients occur from the top to the bottom. After fertilization, the chromosomes divide, and nuclei migrate to different parts of the cell without forming separate cells. Once there, the different proteins stimulate the expression of particular genes. It is this differential gene expression which causes cells to form in different parts of the egg that will specialize into different types. In other animals, similar initial differences in proteins within different parts of the egg result in differences in the contents of the first two cells following cell division, which in turn determine subsequent patterns of cell division. Plants go through a similar process of starting with an asymmetric egg leading to unequal cell division and cell specialization.
Further cell specialization is guided by cell-cell interactions which ensures that the formation of each part is coordinated with those of adjacent parts. All such specialization is due to differential gene expression that is determined by a small set of molecules. These molecules set up a cascade of effects, with the translation of the first gene initiating the expression of several others and so on. That way, a single trigger can alter the expression of hundreds of genes resulting in very different phenotypes. How that differential gene expression is altered by evolution to create differences forms among species is the topic of the next section.
Although Charles Darwin is the name that most often comes to mind when thinking of the theory of evolution, other scientists were just as instrumental in the spread of evolutionary theory. One of the most important was Ernst Haeckel, a contemporary of Darwin whose work was praised by the English naturalist. Haeckel was born February 16, 1834 in Potsdam, Germany. Unlike Darwin, however, Haeckel led a life filled with drama, a great example of 19th-century German Romanticism.

Like Francis Bacon (see Box 2C), Haeckel’s family had a background in law, his father Karl standing as a jurist and a privy counselor to the Prussian court. His parents made certain he had a well-rounded education, his mother Charlotte raising him on classic German poetry, especially the work of Friedrich Schiller, while his father taught him the nature-philosophy of Johann Wolfgang von Goethe and the religious writings of Friedrich Schleiermacher, who was a friend of Haeckel’s aunt Bertha. It is his father’s legal background, however, that might have lead to Haeckel’s attempts to bring the systematic clarity of the legal code to what he saw as the disorder of the biological sciences, trying to codify the laws by which nature operates.

In August of 1852, following the wishes of his father, Haeckel enrolled in the medical school at Würzburg, the best in Germany at the time. Haeckel came under the tutelage of leading luminaries of the field like Albert von Kölliker, who introduced Haeckel to microbiology as part of his courses in histology, and Rudolph Virchow, who championed the (at the time) controversial concept of the cellular basis of life. Haeckel quickly found that he didn’t care for the clinical practice of medicine, but had an abiding love of scientific inquiry, especially microbiology. He found that he had a talent for examining a specimen under the microscope with one eye, and making a highly detailed and accurate drawing of it with the other. It was the love of science that lead Haeckel to continue studying medicine, as a way of achieving a scientific vocation. Virchow’s lectures in particular impressed him and kept him at Würzburg.

Haeckel followed Alexander von Humboldt in the conception of the role of science and scientists in the world; Humbolt insisted that the principles of astronomy, chemistry, physics, geology, botany, and zoology could be tied together, and it was the work of the natural scientist to reveal what he considered to be the great harmony of life. He also fiercely believed that the wonder of nature should be captured with a certain amount of artistry. This melding of science and art continued in Haeckel’s adoption of Goethe’s conception of the complementary nature of science and art in metaphysics, with nature harboring certain archetypes that the naturalist had to uncover in order to correctly articulate theory, and the artist had to know in order to render true natural beauty in paintings and poetry. In addition to his childhood exposure to these ideas, Haeckel devoured Goethe while in medical school, leading to his own work being rich in both science and art. Haeckel’s own nature was that of the Romantic, finding affinities in Schiller, Goethe and Shakespeare, and often feeling as though he were possessed of two souls; the “loving man” who felt deeply and kindled passions with nature and poetry, and the “scientific man,” who douses emotions with cold reason to achieve objective understanding of
natural laws.

Haeckel conceived of a way to unite the two aspects of his nature in a Humboldtian vision of the researcher in exotic lands who occasionally attends to the medical needs to the natives. It was this vision that helped him stay on the path of a medical career, which he found increasingly distasteful. In the fall of 1855 he was required to participate in the clinical treatment of patients, where he encountered the effects of parasites, rickets, scrofula and horrible eye diseases, all in children. The experience served to discourage him further, especially after spending the summer collecting marine specimens with Johannes Müller, the most famous physiologist and zoologist of the time. However, he found salvation in Virchow’s encouragement to study pathological anatomy, discovering that he enjoyed autopsies and anatomical investigation. Unfortunately, when Haeckel became Virchow’s assistant the following summer, the two proved to be a poor match; Virchow was very cool and detached, in contrast to the volatile and passionate Haeckel.

Despite his extreme dislike of clinical practice, Haeckel received his medical degree in March of 1857, writing his dissertation on the histology of river crabs. Biological research was still his ruling passion, and he arranged to conduct his habilitation study (necessary at the time to hold an academic position) with Müller. Unfortunately, before Haeckel could travel to Berlin, Müller, suffering from extreme depression, committed suicide. In the wake of this, one of Müller’s protégés, Carl Gegenbaur, invited Haeckel to visit him in Jena, where he was a professor of anatomy. In May 1858, he offered to have Haeckel travel with him to Messina in Italy on a biological survey expedition, to which Haeckel quickly acceded. That trip was to alter the course of Haeckel’s scientific career.

Before leaving, Gegenbaur and another of Haeckel’s friends from Würzburg gave him some advice: don’t marry anyone if he wanted a fruitful career. This advice came two months after he had secretly become engaged to his cousin, Anna Sethe. Haeckel wrote about Anna as the lodestar of his life, an all-consuming love for her giving meaning to his work and the universe. This love helped keep the “loving man” and the “scientific man” in balance, drawing Haeckel back from the dark abyss of materialism necessitated by science and reason. The two announced their engagement, which had become an open secret by then, on September 14, 1858.

Gegenbaur canceled his Italian trip, but Haeckel planned a trip himself for his habilitation study and Bildung, a journey of intellectual and personal formation. He planned to travel to Florence and Rome in the spring of 1859 to study art, to spend the summer in Naples to conduct marine research, and travel to Palermo and Messina by winter. He arrived in Florence on February 6, but quickly grew weary of religious themes in the paintings – he had planned to study the masterworks – the continual depiction of the Virgin Mary wearing on his northern German Romantic sensibilities. Rome was even worse to Haeckel’s mind, the very day-to-day life of Italians heavily based on religion, daily throngs of religious processions clogging the streets and cardinals displaying rich robes

Ernst Haeckel and his assistant Nicholas Miklouho-Maclay, photographed in the Canary Islands in 1866. (Source: First Run/Icarus Films)
to the crowds of the poor. However, despite the current realities of Rome, Haeckel still found beauty in the classical Rome of Vergil, Horace and Cicero.

Haeckel left for Naples on March 28 to begin his biological research, but found himself stymied. After six months of fruitless research, unpleasant weather and unpleasant people (according to the letters he wrote to Anna) he left Naples on June 17. He traveled to the islands of Ischia, and found the wonder of nature he was looking for in the mountainous vistas, and a kindred spirit in poet and painter Hermann Allmers. Allmers and Haeckel became lifelong friends, spending a month wandering the Italian countryside. Haeckel was even tempted to give up an academic life for one of the bohemian painter, but recognized that his own artistic skills were not quite up to par, and, more importantly, he wouldn’t be able to support a wife unless he received a professor’s salary. The pair traveled to Capri, then Messina in Sicily. Allmers left in mid-October, and Haeckel began to concentrate on marine research again, and again despaired at the breadth of variety of oceanic life.

Fortunately for Haeckel’s research, he had brought Müller’s last monograph with him, in which Müller had researched a group of single-celled marine animals, radiolaria. That study became the jumping-off point for Haeckel’s own habilitation monograph, a much more extensive one than Müller’s. When he returned to Berlin in April 1860, he worked on his collection at the Berlin Zoological Museum, and described several thousand new species. This study increased by half the number of known species of radiolarians, and produced careful description of the distinguishing characteristics of both their skeletons and soft parts, including exact measurements. Until then radiolarian internal characteristics had never been described. Haeckel’s work remains even today as the starting point for interpreting images produced by electron microscopes. He specified the seas in which given species lived, and the depths at which they were found, and attempted to arrange them into a “natural system” based on homology – the relation of the skeleton to the central capsule and the form or absence of the skeleton – separating them into about 15 families.

His two-volume monograph, Die Radiolarien (Rizopoda Radiaria), was published in 1862 and earned him the gold Cothenius medal of the Leopold-Caroline Academy of German Scientists. Charles Darwin, upon receiving a copy from Haeckel, wrote back to him expressing his astonishment at the insightfulness of the work and began a long and fruitful intellectual exchange.

Haeckel found himself inspired by Darwin’s Origin of Species, becoming a zealous proponent of evolutionary theory. He thought he could establish evolutionary theory empirically, arguing that relatedness within radiolaria families pointed to genealogy, and transitional species joining families confirmed this. Darwin himself was so impressed by Haeckel’s work that in The Descent of Man he claimed that his ideas on human evolution were antecedently confirmed by Haeckel.

Haeckel’s background in art stood him in good stead in his work on radiolaria; he was brilliant at creating detailed illustrations, and made his own drawings for the copper
double plate illustration showing embryos of fish (F), salamander (A), turtle (T), chick (H), pig (S), cow (R), rabbit (K), and human (M), at “very early”, “somewhat later” and “still later” stages, from Haeckel’s Anthropogenie published in 1874

Today, the idea that he is best remembered for is the theory of recapitulation which posits that an individual’s biological development, its ontogeny, is similar to its species’ entire evolutionary development, or phylogeny. This theory is often summarized by the slogan “ontogeny recapitulates phylogeny.” Haeckel promoted this theory based on a series of drawings of embryos, in particular an illustration from his textbook Anthropogenie, that compared those of fish, salamander, turtle, chicken, pig, cow, rabbit and human. These illustrations were the focus of extensive debate for the next 100 years, especially whether Haeckel was accurate in his representations or whether he drew them so as to support his theory. Today the theory of recapitulation is generally discredited. Although some features of embryos mirror those of a species’ ancestor (e.g., gill slits in the early embryos of humans), embryological development cannot be taken as a map of a species’ evolutionary history. While Haeckel’s drawings are considered an accurate representation of what he observed, like any rendering they were done in the context of some theory that affects which features a scientist chooses to highlight. Such choice is not evidence of fraud, as has been claimed in this case. Rather it is a legitimate part of they way that each scientist interprets the data with which she is confronted.

Haeckel had a long and distinguished career. During his 47 years at the University
of Jena he published nearly 50 books ranging from systematic studies of various single-celled and simple multicelled marine organisms, anatomy and evolution textbooks, books on philosophy and travelogues. In 1908 he was awarded the Linnean Society of London’s prestigious Darwin-Wallace Medal. Haeckel died at the age of 85 on August 9, 1919.
Evolution

The process of change

As with other aspects of biological systems, the characteristics of organisms are the result of evolutionary processes (Table 6.1, principle 10). Similarly, the process of evolution depends on the characteristics of organisms. Evolution requires phenotypic variation (see Chapter 4), which comes from the entire set of processes that give rise to individuals. That interaction between the characteristics of individuals and the process of evolution is nowhere more apparent than in multicellular individuals and the incredible variety of developmental processes which have shaped and been shaped by the course of evolution.

For convenience we can think of evolutionary change of two kinds. First are simple changes such as in size (the species becoming bigger or smaller) (see Figure 4.4), or limited metabolic changes such as producing more or less of an enzyme. Second are complex changes involving multiple systems or major structural developments, such as evolving legs. Often the second type of change is framed in terms of the transition from one body type to another (Figure 6.36) or the appearance of phenotypic novelties, that is, new types of structures or functions.

A debate has existed for 150 years about these kinds of evolutionary change. Charles Darwin theorized that there was really only one kind of change, the slow accumulation of simple changes that eventually result in more complex changes. This view has held center sway among most evolutionary biologists, while others have felt that the two types of changes were of a different nature requiring a different evolutionary mechanism. The latter position came from two observations: there were gaps in the form of organisms and the rate of change recorded in the fossil record seemed to be too fast to explain these complex, coordinated changes in form.

For most of that time the debate could only draw on morphology due to the limitations of the fossil record and contemporary biological techniques. However, in recent decades, as we have learned more about the process of development – especially its genetic basis – the debate has focused more on genes and cells. As previously described, development is a complex process in which the expression of a single gene can initiate a cascade of effects. A single genetic change can initiate multiple changes in the phenotype, not just because of direct change in gene regulation, but also because cell-cell interactions cause changes in other cells, tissues and organs. Small changes early in that cascade can have large effects on the form of the organism. On the one hand, this cascade can explain how complex structures can evolve relatively quickly. On the other hand, this coordination poses a problem. In general, large changes will not be favored by selection because they are more likely to decrease than to increase the fitness of an individual. While developmental processes are coordinated, if initiated at random such coordination is more likely to result in a mishmash of parts than a functioning organism.

Part of the resolution of this quandary is to recognize that changes in developmental processes do not have to produce large changes in and of themselves to have a large impact on the final phenotype of the individual. This is due in large part to the complex intercellular signaling system that exists; in other words, cell-cell interactions allow changes in one part to coordinate and cause changes in other parts. Consider the evolution of the legs of terrestrial vertebrates from the fins of fishes (Figure 6.36). A leg consists of a complex set of bones, muscles and nerves that must all fit together to work. This coordination of parts occurs because as the parts develop they send signals to each other so that muscles attach in particular places and that nerves are able to find the muscles. Given the existence of such a signaling system, small evolutionary changes
in one part of the system, say the underlying skeletal structure, would create corresponding changes in the rest of the organism’s systems. Eventually the accumulation of these small changes results in large changes in the body plan. Such changes could occur in just a few hundred or thousand generations, a mere blink of the eye when viewed from the resolution of the fossil record that usually can only record changes over millions of years.

A current debate involves what types of genetic changes are mostly responsible for evolutionary changes, both large and small. We can divide the DNA on a chromosome into two categories: genes that are transcribed and give rise to proteins that carry out cellular functions, and genes that regulate the rate at which the first category of DNAs are transcribed. The proteins that are responsible for transcription first have to bind to the DNA strand, and changes in the nucleic acid sequence can alter how easily such binding occurs – and thus, the rate of transcription. Some genes even code for proteins that are binding factors; their primary role is to regulate the rate of transcription of other genes. See Chapter 3 for a more detailed description of the process of transcription.

One position in the debate is that changes in regulatory genes or DNA binding sites are the primary cause of evolutionary change. In this case, changes take place due to alterations in the rate of gene expression.

In particular, changes in body plans are due to the evolution of the regulatory genes that control the first part of the developmental cascade. The other position is that change takes place due to alterations in the structure of enzymes that carry out cellular functions, alterations that change the rate of enzyme reactions or the substrate to which the enzyme reacts. As with most such debates, the answer is somewhere in the middle. Both types of changes are responsible for evolution, we simply do not yet know the relative importance of each type of genetic change and whether they are responsible for small or large phenotypic changes.

**The pattern of change**

One result of the coordination of developmental processes is to make some directions of evolutionary change less probable or even impossible. Consider three mythical creatures – the unicorn, the Pegasus, and the centaur (Figure 6.37). If we were to start with a horse, is it possible or probable for any of these three creatures to evolve? (For simplicity, consider the man-like portion of the centaur not as human, but as a human-looking extension of the horse.) It is easy to see how a unicorn could evolve; horns have evolved many times in mammals. In this case it would involve a relatively simple change in the growth of bones in the skull of the horse, a change that could be initiated by just one or a few small genetic changes. The myth of the unicorn likely came from distorted descriptions of a rhinoceros, a relative of the horse, although its horn is at the end of its nose rather than on its forehead.

A Pegasus is very hard to imagine evolving for several reasons. First, all terrestrial vertebrates have no more than four limbs. A Pegasus has six limbs: four legs and two wings. Substantial changes in the developmental process would be necessary to cause the growth of a new set of limbs in the middle of the back, ones involving large changes
in the genetic controls of that development. In addition, a whole new set of muscles would be necessary to properly attach those wings to the back so that they could flap. (We will ignore that fact that even wings that work would never be big enough to get a horse off the ground.) Finally, the wings require feathers. Both feathers and hair evolved from the scales of reptiles, but along completely different developmental and evolutionary pathways. While mammals have evolved wings (e.g., bats), they do not have feathers. Following this evolutionary path, at best you would end up with a bat-winged horse.

What about a centaur? At first glance, it seems to have the same limitation as a Pegasus of having six limbs; however, the architecture of a centaur is different. Imagine starting with a horse and duplicating the forward part of the ribcage and shoulder girdle, so that you would have a horse with two chests. It would be a very weird looking horse, but it would not involve any new structures, just the duplication of existing structures. Presumably all of the existing developmental machinery that produced one set of ribs and legs could produce a second set. Such structural duplications are known in fruit flies; they occur through mutations of regulatory genes. Once such a duplication existed, assuming that it was viable, smaller evolutionary changes could be selected for that would cause the front limbs to shorten and the spine to hinge behind the first set of limbs to make the front of the creature upright. It is even possible to imagine the hoof turning into a hand. A horse’s hoof is an elongated middle toe, the other four toes having been lost during its evolution. However, the genetic and developmental machinery necessary for a five-fingered limb may still exist and just be turned off or possibly changed in relatively small ways. Such reappearance of ancient structures is even known to occur. The phrase “rare as hen’s teeth” has a biological basis: Birds, whose ancestors stopped producing teeth at least 80 million years ago (see Box 4B), can be induced to develop them. So, although a centaur is a highly improbable creature, it could evolve.

We can summarize what is known or hypothesized about the evolution of new phenotypic structures and functions in four basic principles. First, the process of development is controlled by a small set of similar molecules that regulate gene expression.
Second, new genetic functions come about through gene duplication and divergence. One constraint on changes in gene functions (e.g., enzymes that react with different substrates) is that the original function will be lost; a solution is for the gene to be duplicated allowing one copy to evolve a new function while preserving the function of the other copy.

Third, the developmental process occurs through discrete, interacting modules. Because those modules are semi-independent, evolution of one module can happen without affecting the development of another module. In addition, evolution can occur by the duplication of a module. Once duplicated, the modules can diverge from each other. Fourth, many developmental changes occur through changes in the location, time, amount, and kind of gene expression. These developmental changes can be initiated by simple mutations in the molecules responsible for gene regulation, and can also occur through changes in the environment which induce phenotypic changes (see Chapter 3).

Once new phenotypic structures and functions arise, they need to become fixed in a population in order to cause any significant evolutionary change. That is, any mutation that appears in just a single individual then has to spread through the rest of the population through natural selection. If the developmental change is caused by the environment (see Figure 3.16), the change needs to become a genetic change. Various processes can increase the probability of the change becoming fixed. Once the first mutation happens, mutations at other genes can occur that affect the coordination of other aspects of development. For example, the reduction of horses’ toes from five to one involved a change in not just the morphology, but also the muscle and nerve systems. to some extent, these changes occurred by pre-existing signals that existed between those systems. Any new mutations, however, that created the needed changes in the muscles and nerves without the signals would likely be favored because the new arrangement of parts could develop at a an earlier stage of the embryo. Eventually, more complex developmental changes become possible as new patterns of coordination evolve.

**Phylogeny**

We can also use developmental patterns to help discern evolutionary relationships. A central figure in this effort was the 19th century German scientist Ernst Haeckel (Box 6C) who famously stated that ontogeny recapitulates phylogeny. By this he meant that you could look at the developmental changes of an embryo (ontogeny) and see the forms of the ancestors of that individual (phylogeny). Today we know that Haeckel’s theory is only partially correct. For example, young embryos of all mammals have structures that look like the gill slits of fish, and young embryos of humans have tails. Both of these structures disappear later in development. However, mammalian embryos never really resemble their fish or reptilian ancestors. Still, the appearance of gill slits in those embryos provides additional evidence for common ancestry. In that sense, Haeckel was correct that the features of embryos could be used to help determine relationships, but not to the overwhelming extent that he thought. Of course, when he was working in the 19th century, morphological features were the only ones available for study, so a desire for a theory that would guide their use was understandable. Today we have many more types of traits to work with, especially DNA sequence data, that together with developmental information can help us trace the history of evolution on Earth.

Once we have those relationships, they become essential tools in helping us understand organismal structure and function. For example, organisms need to maintain just the right balance of ions in their blood relative to their cells. The ion concentration in our blood approximates that of seawater, another legacy of our fish ancestors. Some animals have evolved to live in fresh water while others live in places that are even
saltier than the ocean (e.g., the Great Salt Lake in Utah). A simple comparison of species in those different environments can tell us something of the physiological changes necessary for those evolutionary shifts. Such comparisons are even more informative, though, when the relationships among the species are known (Box 6D). While it may seem obvious that comparative studies which use phylogenetic relationships can provide important information, even today many studies of organismal structure and function are not comparative. Rather, they are based on one species at a time rather than making use of the diversity generated by evolution. The problem with this approach is that you cannot confirm that specific traits are actually adaptations to a given environment. For example, a study of a species living in the Great Salt Lake is likely to have very many differences from species living in freshwater. Only some of those differences will be due to adaptations to saline conditions; others will be due to chance (see Chapter 4). Despite this limitations, many scientists doing these studies tended to (erroneously) attribute all differences to adaptations. Even those studies that compare species do not necessarily use a phylogenetic framework. However, as phylogenetic relationships become easier to decipher with DNA sequencing technologies, they are becoming more and more common.
Box 6D
Critical experiment: comparative physiology

Studying the structure and function of organisms involves addressing two questions. First, how does a particular aspect of an organism work to help maintain integrity and to allow for growth and reproduction? Second, are differences among organisms (either individuals within a population, among different populations, or among different species) due to adaptive changes or to random chance? While the second question clearly requires comparative studies, often, the first question is best answered using a comparative framework as well. Those comparisons, however, must be done in a meaningful way or they can give misleading results. One experiment that exemplifies this fact was conducted on enzyme activity in fish.

Aquatic environments can experience a wide range of temperatures. The oceans vary in temperature from the warm waters of the tropics to the frigid polar seas and from sun-warmed surface waters to the cold abyssal depths. At middle latitudes, the temperatures of lakes and rivers track seasonal changes. Temperature affects enzyme function, so we expect organisms that live in these different environments to differ in the enzymes used for their metabolic functions. Such differences provide an opportunity for scientists who wish to discover how enzyme structure relates to its function at different temperatures.

One way to address that question would be to take a single enzyme and change each of its amino acids one at a time to see how the change affects enzyme activity. You would then have to test how the new enzyme interacts with other enzymes and processes, because of the profoundly interconnected nature of living creatures. That approach would obviously be incredibly time consuming, given that there are 20 common amino acids. For an enzyme consisting of 100 amino acids, there would be 1020 (more than a billion-billion) possible combinations to try! Instead, we can make use of the fact that the evolutionary process is continually exploring many possibilities and look at natural variation in enzyme structure.

Such studies start from the premise that natural selection should act to change the function of an enzyme so that it enhances the fitness of an individual. If individuals live in different environments, and if different enzyme structures have optimal functions in each environment, then natural selection should result in organisms whose metabolisms use different enzymes, each tailored to fit best with its environment. If we wished to know, for example, how temperature affects enzyme function in fish, we could compare enzymes from fish living in tropical areas and temperate areas and look at the differences in their amino acid sequences.

George Somero and his colleagues addressed these questions over several decades while at the University of California-San Diego, Oregon State University, and Stanford University. They focused their attention on an enzyme that is important in energy metabolism, especially in muscle tissue: lactate dehydrogenase (LDH). In an early study they compared the temperature stability of LDH among several species that lived in very different environments, including the bald rockcod which dwells in the Antarctic Ocean at temperatures as low as -1.8°C, and the desert iguana of the southwest U.S. which can experience temperatures as high as 47°C. As predicted, the enzyme from species living in hotter environments became destabilized at higher temperatures. However, a direct comparison of the enzymes of these species would find a great many differences because they are so distantly related.

Their next step was to look at a set of closely related species, barracuda living in the eastern Pacific Ocean along the coast from California to Ecuador. Barracudas are
voracious predators that lie in wait, relying on surprise and short bursts of speed (up to 45 km/h) to overran their prey. Muscle energy metabolism is therefore likely to be important for fitness. The four species that they studied are all closely related members of the same genus with similar life styles. However, they differ in the temperatures of the waters that they inhabit: 18°C for the two living off the coast of California and Baja, Mexico, 23°C for the species off the coast of Central America, and 26°C for the one off the coast of Peru and Equator. They found that LDH activity differed as a function of temperature in the expected direction, with higher activity at lower temperatures for the species from the cooler waters, and vice versa.

Critically, when you looked at the range of temperatures each species inhabited, the enzymes tended to have the same activity. That is, enzyme activity at 23°C for the Central American species was the same as at 18°C for the more northern species. The enzymes had evolved so as to converge to the same activity levels despite living in different environments; that is the peak of the temperature response (Figure 6.FR) shifted higher or lower so as to be at its maximum at the temperature of the local environment for that species.

The next step was to determine how the enzymes had changed. Based on the DNA sequences of three of the species, they found just four amino acid differences in an enzyme consisting of more than 300 amino acids. They then took the version of the enzyme found in the Baja species and one at a time made the four substitutions found in the tropical species. Just one of those amino acids accounted for the differences in temperature stability, and two other amino acids were responsible for the differences in enzyme activity. The fourth amino acid difference had no effect. Most surprising, none of those amino acids were found in the portion of the enzyme responsible for its chemical activity, rather in parts of the enzyme that controlled the overall structure. If the scientists had tried to manufacture changes in the enzyme, and had come up with a plan based on their preconceived ideas of which changes were most important, changes in these amino acids would not have been at the top of their list. So by studying a set of variants filtered by natural selection, they greatly increased their ability to discover the key changes.

This study was one of the first to show that adaptation to temperature could occur by very few changes. Another aspect of their comparison found that different properties of the enzymes – stability and activity – could evolve independently. It shows the value of making comparisons among closely related species. Importantly, not all amino acid changes were necessarily linked to functional and adaptive differences, showing the need to separate adaptive differences from those occurring by chance alone. Otherwise, future studies exploring functional changes would waste time creating and examining enzyme variants with no functional differences. Focusing on closely related species with few differences in amino acids made it easy to determine which differences were the critical ones.
Chapter 7
Ecology

When you go outside, take note of the organisms around you. At first they will seem to be all alike; every tree seems like every other, with a trunk, branches and leaves all fed by an underground root system. After a closer examination, differences soon become apparent, until it seems as though there are no similarities between individuals at all; each squirrel is a different size and behaving differently from all of the others. When taken as a whole, however, patterns will emerge in the behaviors, morphology, and geography of the organisms in the world; different kinds of trees are found on hilltops and others in valleys; the squirrels in one location might be bigger than those in another. Some birds tend to stick to the interior of the forest, while others are found mostly along the edges. The differences among species form patterns as well; fish are found only in the water, kangaroos live only in Australia. Some of this variation seems obvious and unsurprising, and other aspects are subtle and require detailed measurements to uncover.

As we have discussed throughout this book, living systems bear several major hallmarks, one of which is variation (Figure 7.1). While some variation is due to genetic differences among individuals (see Chapter 3), and some is due to differences in growth and development (see Chapter 6), there is a third cause of variation that must be taken into account. this is the external factors affecting individual organisms; in other words, the environment. Like many other aspects of biology, organisms and their environments form a complex set of interactions that feed back on each other. In this chapter we will

![Figure 7.1](https://example.com/figure71.jpg)

explore both how the distribution and abundance of organisms vary in space and time and how that variation is driven by interactions with environmental factors. The explanations for these issues fall under the purview of the subdiscipline of biology known as ecology.

The Theory of Ecology

The general theory of ecology consists of eight fundamental principles (Table 7.1). The roots of these principles can be traced to the origins of the science of ecology in the 19th century, the roots of which can be traced back through the study of natural history to ancient Greece (see Box 2A). The word “ecology” was coined by Ernst Haeckel (see Box 6C) in German as “oekologie.” It comes from the Greek word “oïkos,” meaning “house,” indicating that ecology is the study of an organism’s abode.

Table 7.1. The eight fundamental principles of the general theory of ecology

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Organisms are distributed unevenly in space and time.</td>
</tr>
<tr>
<td>2.</td>
<td>Organisms interact with their abiotic and biotic environments.</td>
</tr>
<tr>
<td>3.</td>
<td>The distributions of organisms and their interactions depend on contingencies.</td>
</tr>
<tr>
<td>4.</td>
<td>Variation in the characteristics of organisms results in variation of ecological patterns and processes.</td>
</tr>
<tr>
<td>5.</td>
<td>Environmental conditions as perceived by organisms are heterogeneous in space and time.</td>
</tr>
<tr>
<td>6.</td>
<td>Resources as perceived by organisms are finite and heterogeneous in space and time.</td>
</tr>
<tr>
<td>7.</td>
<td>Birth rates and death rates are a consequence of interactions with the abiotic and biotic environment.</td>
</tr>
<tr>
<td>8.</td>
<td>The ecological properties of species are the result of evolution.</td>
</tr>
</tbody>
</table>

Ecology emerged as its own discipline between 1880 and 1910, along with much of biology. The roots of many of today’s ecological theories can be found in this period, although the fundamental principles emerged only slowly over the next 50 years. While all of the fundamental principles had been put forth by the middle of the 20th century, there were still serious debates about their relative importance. For example, the process of evolution in shaping global patterns of species distributions was recognized by Charles Darwin (see Box 4A) in the nineteenth century, but it was not until after the Modern Synthesis shaped the theory of evolution in the middle of the 20th century (see Chapter 4) that there was extensive study of the interplay of ecological and evolutionary processes within local settings.

The domain of the science of ecology is the spatial and temporal patterns of the distribution and abundance of organisms, including causes and consequences. Strikingly, four of the eight fundamental principles are about variation of either the biotic or abiotic world. It is this focus on variation that makes ecological studies so challenging and exciting. Notably, the theory of ecology is the only theory of the six put forth in this book that includes fundamental principles about the abiotic world; in part, principles 5 and 6 are about aspects of the world like rainfall and temperature (see below). As a result, ecology is the aspect of biology that pays the most attention to the world outside of the organism.

Because the science of ecology deals with interactions between organisms and the rest of the world, it is often linked with political concerns about the state of the world,
which can be grouped under the heading of environmentalism. However, ecology is not environmental advocacy or political activism, although scientists are sometimes environmental activists in their personal lives, and environmental activists may, and should, rely on scientific research. Ecology itself is not about one’s feelings about nature, although ecologists may have strong feelings about what they study.

**Distributions and Interactions**

**Uneven distributions**

Organisms are not uniformly distributed across the face of the Earth (Table 7.1, principle 1; Figure 7.2). As far back as 2500 years ago, Greek natural historians were aware that different species were limited in the types of habitats in which they were found. However, scientists first became aware of large-scale patterns in species diversity as a result of the voyages of exploration and colonization undertaken by Europeans in the 18th and 19th centuries. Their ships often carried botanists and zoologists as part of the crew, the most famous example being the voyage of the Beagle (1831–1836), on which Charles Darwin served as the ship’s naturalist (see Box 4A). These naturalists made records of the plants and animals they found, brought back specimens, and added them to the growing catalog of described species.

From these records, it soon became apparent that tropical regions were typically very rich in species, polar regions were very species-poor, and temperate regions had intermediate numbers of species (Figure 7.2). Brazil alone has over 56,000 named plant species, while the United States has about 18,000, and Canada has about 4200. An explanation for this pattern of more species in the tropics may be the oldest major ecological hypothesis. Between 1799 and 1804, Baron Alexander von Humboldt traveled through Mexico, Central America, and northwestern South America. He subsequently published a series of essays under the title *Ansichten der Natur* (“Views of Nature”), in which he described this pattern and postulated that it was due to differences in climate, specifically winter temperatures and the effects of freezing.

Even within a single population of one species, individuals generally are not uniformly distributed (Figure 7.3). They may be concentrated in patches, as with prairie dogs in the western United States. Prairie dogs live in large colonies called towns; in northern Colorado each town contains about 850 individuals and the towns are separated by an average distance of 50 km. Individuals of a species may also be evenly spread throughout the landscape, as is seen with populations of creosote bushes in the Sonoran

![Figure 7.2](image-url)
The number of individuals may also vary through time, such as the cycles of the snowshoe hare and Canada lynx (see Figure 1.4). Prior to the development of vaccines, many disease-causing viruses showed similar boom-and-bust cycles (Figure 7.4). Not all change is cyclical, however. Following a major disturbance, such as a forest fire or a hurricane, the plants in a community will slowly change in composition (Figure 7.5). Immediately after the disturbance some plants will resprout from still-living root systems and others will germinate from seeds that sat dormant in the soil or that are newly arrived in the site. The result can be a very different mix of species than were in that site prior to the disturbance. Over time the first species may be displaced by later arriving ones, until over the course of decades or centuries the original community reappears.

Change also occurs over much longer time spans. For the past 2 million years the Earth has been in the midst of a glacial period. During that time glaciers have waxed and waned at least four times. In North America, at the height of the last advance, glaciers extended as far south as the region around the Great Lakes. At that time forests were found only in what is now the southern United States. For the past 20,000 years we have been in a period of glacial retreat and those forests have migrated to northern Canada. Over even longer time spans, however, some patterns are extremely stable. For example, the pattern of many species in the tropics and few at the poles is very old, going back at least to the time of the dinosaurs 100 million years ago.
Thus, the number of individuals in a population or the number of species in a community are dynamic. This dynamic aspect of ecological systems contrasts with an older view of the world existing in a “balance of nature,” an idea which goes back at least 2500 years to the Greek historian Herodotus, and was the prevailing view into the first half of the 20th century. While scientists recognized that communities changed following a disturbance, the notion was that communities tended to return to an equilibrium in which all plants and animals were at numbers sufficient to sustain the entire system. These ideas were most notably put forward in modern science by Fredric Clements (Box 7A), who formulated the first theory of change in plant communities. We now know that the idea that nature is in balance is wrong. Not only are systems dynamic, populations are often disappearing from a given site and replaced by other species or perhaps reappearing as new individuals move back into the system.
One of the seminal figures in ecology was Frederic E. Clements. He produced one of ecology’s first major theories, thereby shaping much of the research in plant ecology in the first half of the twentieth century. Despite that enormous influence, today his theory is largely discredited. How he came to create and lose such influence demonstrates how the social aspects of science can play a large role in shaping research agendas.

Clements was born on September 16, 1974 in Lincoln, Nebraska, where he lived until he attended the nearby University of Nebraska to study botany. He received his undergraduate degree in 1894 and remained to continue his graduate studies, earning his Ph.D. in 1898. He stayed at the university as a faculty member, rising to the rank of full professor in 1905 at the early age of 32. In part, because of the presence of Clements, the University of Nebraska was one of the two premier centers for the newly developing science of ecology at the turn of century, the other being the University of Chicago.

During this period the nascent science of ecology was coalescing. Up until then, most botanical studies were description of floras, and Clements was instrumental in moving the discipline beyond simple description. He and others noticed that a striking feature of plant communities was that they changed over time, in seemingly predictable ways. A field that had been used to raise crops, once abandoned, would first be colonized by fast growing herbs, and within a few years those plants would be replaced by longer-lived herbs as well as woody shrubs and trees. Among the tree species, some fast growing trees that need a lot of sunlight would come in first and later be displaced by more slowly growing and shade-tolerant ones. This process was eventually termed succession. The question facing ecologists at the turn of the 20th century was whether this process was predictable.

At that same time, much of biology was moving away from being a descriptive science with a focus on organisms in their natural settings, to being an experimental science done in laboratories. Field-based work was considered old-fashioned and substandard. Clements directly confronted that attitude by publishing Research Methods in Ecology in 1905. That book accomplished two goals; it presented a vision of a plant ecology as a unified discipline, and was the first attempt to define the conceptual boundaries of the new discipline of ecology and to distinguish it from natural history. It took the physiological perspective that pervaded biology at the time – the study of process and structure-function relationships – and applied them to natural communities. As important, it laid out a series of standardized techniques for measuring vegetation, which were deliberately designed to mimic the rigorous methodologies being developed for laboratory-based studies. This book was the first comprehensive set of ecological methods and became a standard reference for decades.

At about the same time, in 1904, he published The Development and Structure of Vegetation in which he presented an early form of his ideas about succession. Those ideas were further shaped by his field work in the western United States, work that could not have happened without his wife Edith Clements. The story of Frederic Clements is as much a story about Edith.

Edith G. Schwartz was also a student at the University of Nebraska, majoring in German. As an undergraduate she was Phi Beta Kappa and an accomplished athlete in...
basketball, fencing, and tennis. They married soon after graduation. With Frederic’s encouragement, Edith enrolled in the graduate program in botany; she would become the first woman to receive a Ph.D. from that institution. For the next four decades they traveled hundreds of thousands of miles studying vegetation and worked together growing plants in their greenhouses and experimental gardens.

This work happened only because of Edith. Frederic was decidedly not a person well suited to practical concerns. Besides being diabetic and of ill health, Frederic would forget to eat and sleep to the point of producing hallucinations. Edith had to remind him to care for himself. During their field trips, she was driver, car mechanic, cook, all around handy person, photographer, artist, and stenographer. She also used her language skills to translate his books and articles into several foreign languages. With Frederic she published a number of books including floras of the Rocky Mountains that she illustrated. At the age of 86 she published Adventures in Ecology where she described some harrowing tales of life in the field, including an incident when they were almost robbed and killed in the remote Arizona desert.

Those field surveys, along with a comprehensive summary of the literature of ecological studies, culminated in 1916 in the publication of Frederic Clements’ book *Plant Succession: An Analysis of the Development of Vegetation*. Here he put forth a sweeping theory that envisaged communities as equivalent to individuals, with successional change equivalent to the process of development and inevitably leading to a final climax community whose composition was determined by climate, especially water availability. That theory had both strengths and weaknesses. On the one hand, it was broadly conceived and complete, based on Clements’ comprehensive knowledge and insights. On the other hand, it was a very rigid structure. Clements started with the assumption that the community was an organism and, therefore, always had to be seen from the perspective of the final climax state. His focus on water relationships was very much influenced by the fact that all of his field work was done in the dry climate of the western U.S. where water availability is often a primary determinant of plant community composition. Ecologists who worked in eastern North America and Europe had very different experiences and were less inclined to concede the primacy of climate and water, further eroding support for his ideas.

Bolstering his significant scientific contributions was Clements’ powerful personality. He was variously described as a workaholic and puritanical in his personal habits; he abstained from tobacco and alcohol and did not care for those habits in others. He was aloof, abrasive, and arrogant. This arrogance gave him a penchant for sweeping generalizations that reduced complex problems to just one or a few causal factors. While allowing him to put forth a broad theory, it also made him intellectually inflexible and incapable of acknowledging error or modifying his ideas. Of course, as with all people, he was complex. While some ascribed to him a lack of humor, others saw someone who was able to laugh at himself.

Despite its flaws, his theory provided important intellectual direction to a fledgling discipline. It stressed dynamics and the importance of process – rather than just pattern – and unified early plant ecology towards a focus on succession. However, this unification was not because others agreed with his theory; on the contrary, few ecologists accepted his central premise equating a community with an organism. Rather, it inspired careful work aimed at proving the theory wrong. In science, it is much better to be wrong than ignored.

Clements’ inability to bring together other scientists to advance his theory came despite his leadership positions. In 1907 he left Nebraska to head the botany
department at the University of Minnesota, a position he held for 10 years. His hiring was strongly opposed by many members of that department, and during his time there he often clashed with the faculty, who found him difficult to work with. In 1917 he became a Research Associate at the Carnegie Institution of Washington, a private research foundation. There he was given a lavish budget to assemble a team of four or five scientists plus support personnel. He used his funds to continue his work at the Alpine Laboratory, a research station in Angel Canyon on the slopes of Pikes Peak, Colorado, that he and Edith had established years earlier. Winters were spent at the Carnegie Institute’s facility in Tucson, Arizona, and its Coastal Laboratory in Santa Barbara, California.

Clements was extremely ambitious. Realizing that promulgation of one’s ideas came through collaborators, he had many students and visiting faculty work at the alpine research station in the summers. The other method for spreading one’s ideas is through publication, and Clements was a prolific writer; in the 1920s he published seven books, including the first plant ecology textbook. In 1939 he co-authored *Bio-Ecology*, the first textbook that dealt with animals and plants equally as part of his attempt to widen the scope of his theory to include animal communities.

Initially he was tremendously respected and managed to attract many young scientists to work with him. The problem was that his domineering personality failed to let those young scientists develop independent research and ideas. As a result, his team never expanded on his original ideas. Clements always insisted on being the senior author on major publications. At the Carnegie Institute he attempted to organize all botanical research around his ideas, a position not well received by the other senior scientists. Eventually, when a reorganization took place in the 1930s, Clements was shut out of the process. In addition, his strong commitment to a version of evolution based on the inheritance of acquired characteristics, including unpublished claims of experimental proof, further eroded his prestige at the institute. He continued to work there even after his formal retirement in 1941. To a great extent his theory expired with his death on July 26, 1945. Yet, his physiological and functional view of communities set the groundwork for an approach to the study of ecosystems that began in the 1950s and continues to this day.
Biotic and abiotic interactions

All organisms live within a larger environment and, by necessity, interact with that environment (Table 7.1, principle 2); these interactions in turn determine the distribution and abundance of organisms. For simplicity we can divide those interactions into ones involving the abiotic (non-living) parts of the environment, and those involving the biotic (living) parts. The nature of those interactions vary depending on whether the organism is a plant or an animal, single-celled or multicellular, able to live exclusively on inorganic materials or requires organic materials. However, the basic nature of those interactions is the same for all species.

One common aspect is how the environment affects an individual’s body temperature. For example, when you jump into water that is cooler than you are, you feel chilly. You do not get cold as quickly as a smaller animal would, however, because your greater mass stores more heat energy. If you were to begin swimming vigorously, the metabolic energy produced by your muscles from burning stored calories would warm you up. Because you are a mammal, you maintain a constant body temperature using that metabolic energy (see Chapter 6). Most organisms, however, do not. Their body temperature depends on absorbing energy from the environment, energy which we label “heat” or “light.” Besides just absorbing energy, you also give it off in the form of heat; when you get out of the water, the evaporation of water from your skin makes you feel chilled because the water is carrying off heat. Even just standing still you give off heat energy, which is why your body feels warm to others.

All organisms must take in matter and energy to survive (see Chapter 6). For organisms that can perform photosynthesis (see Chapter 5), the energy and matter are somewhat decoupled; the energy comes from sunlight and the matter from CO2 in the air and water and various minerals in the soil. For all other organisms, the energy is tied up in organic molecules (e.g., sugars) that they must acquire by eating other organisms. Such consumption of one organism by another forms the vast majority of biotic interactions.

Many organisms that consume others for food do so by either consuming it while still alive, killing the other organism, or eating an organism that another animal has just killed or has recently died. This can include a group of lions chowing down on a zebra (Figure 7.6), or the vultures that feast on the carcass once the lions are done. Such consumption is called predation. Some organisms take in organic molecules from other organisms that are already dead, or from dead tissues such as the skin that you are constantly shedding. These latter organisms include most fungi and many bacteria (Figure 7.7). In these cases, the individual interactions can be less obvious. When you are infected with a disease-causing organism (e.g., a cold virus or a bacteria), that organism is, in effect, eating you from the inside.
When one organism eats another, the eater gets a benefit while the eaten suffers a cost. Many times that cost is death, but not always. When you harbor a cold virus, it will reproduce inside you and get passed on to others while only causing you a few days of illness. Most of the time when a plant gets eaten, only part of the plant gets taken, such as a caterpillar munching on a leaf. Thus, the amount of benefit and cost to each of the participants in the interaction can vary. Nor are interactions always of the type where one participant benefits while the other suffers a cost.

Many types of interactions benefit both participants and are termed mutualisms. An obvious example is when a hummingbird visits a flower (Figure 7.8). The hummingbird drinks the nectar from the flower, a solution very high in sugar. At the same time, the hummingbird gets pollen (see Chapter 6) deposited on its head. When the hummingbird visits the next flower, the pollen from the first gets left behind. Importantly, that next flower will generally be one of the same species, so that the pollen can fertilize that flower. Thus, the hummingbird gains energy and the flower gets its pollen moved to exactly the right place.

The pollination of a flower is one example of a large class of interactions involving mating and sexual reproduction. As with eat-or-be-eaten events, these interactions can be either positive or negative. Clearly, two individuals that produce an offspring benefit from that interaction. Often, though, males will compete with each other to gain access to females either through combat – which involves the risk of injury or death – or through choices made by the females (Figure 7.9). Organisms that live in groups (e.g., bees in a hive or meerkats in a family group; Figure 7.10) are constantly interacting in ways both positive and negative. For example, one male baboon will watch for lions while others are looking for food, yet those same males compete for the females in the group.
Positive interactions among species can be extremely intimate or very diffuse. For example, a **lichen** consists of an alga that is living inside of a fungus. The alga performs photosynthesis, providing both partners with energy, and the fungus absorbs other nutrients and provides the alga a place to live. On the other end of the spectrum, a given species of plant may be visited by many different pollinators (e.g., various species of bees and butterflies), while those pollinators may each visit many different species of plant.

In some cases, one participant may benefit from an interaction while having no effect on the other organism. For example, in many tropical forests small plants, rather than being rooted in the soil, may grow directly on tall trees (Figure 7.11). The small plant benefits by growing up in the sunlight rather than on the ground where it is very shaded. That plant is getting all of its nutrients from the air or from rainwater and just using the tree for support, and it is small enough that it has no effect on the tree. (At least, a single plant has no effect – if many such plants grew on the tree, they could interfere with the tree’s ability to photosynthesize or even cause branches to break.) Interactions among organisms can be very positive for both individuals, negative for one, and everything in between, and the nature of the interaction can depend on the context within which it happens.

All of the interactions just described are direct, one individual doing something to another. But interactions can also be indirect. For example, there is a finite number of zebras, and if one gets eaten by a lion, that leaves fewer zebras for other lions or hyenas to eat. A tree absorbs sunlight in its leaves for photosynthesis, which shades the plants below it. Sometimes such indirect effects can be positive. If a lion eats a zebra...
rather than a Cape Buffalo, the buffalo benefits from not being eaten. When the lion eats a zebra, the grass that would have eaten by the zebra is not, so the lion indirectly benefits the grass. On the other hand, if the buffalo eats more grass than the zebra would, then the net effect on the grass might be negative. This network of interactions within a given community based on which species are eating each other is called a food web (Figure 7.12).

![The Soil Food Web](image)

Figure 7.12
A diagram showing the interactions among some of the species in a freshwater pond and the nearby forest. These interactions make up a food web. (Source: USDA Natural Resources Conservation Service)

Indirect effects can also occur through changes in the abiotic environment. An especially dramatic example involves the actions of beavers. Beavers cut down trees and use the wood to build dams across streams, which causes the surrounding area to flood. Just like a man-made dam and attendant reservoir, this creates ponds within which the beavers build houses – creating a place protected from predators. In doing so, they change the environment. By flooding an area, they increase the amount of water in the soil around the pond as well as the amount of light on the area by removing the trees. Some plant species grow much better in areas with lots of light and very wet soil, so they indirectly benefit from the beavers’ activities, while other species are excluded by those changes.

![Beaver and beaver lodge](image)

Figure 7.13

The sum total of all of the interactions of a given species, both positive and negative, with the abiotic and the biotic environment, defines the niche of that species. In turn, the sum of the niches of all of the species in a community and the totality of their interactions make up an **ecosystem**. Much of the science of ecology consists of describing those interactions. Because the interactions, especially indirect ones, can be quite complex, it can take years of work to tease them apart. Additional complexity comes about because the interactions can involve time lags, such as the population cycles of snowshoe hares and Canada lynx (see Figure 1.E). And, as described above, the environment itself is changing over both short and long time spans. While ecologists have always worked to understand these interactions, such efforts were greatly enhanced with the movement towards descriptions through mathematical models. This movement took off in the middle of the 20th century based on the work of many ecologists, notably Robert MacArthur (Box 7B).
Despite his tragically short career, Robert H. MacArthur was one of the most influential – and controversial – ecologists of the 20th century. He was born on April 7, 1930 into an academic family. His father was a professor of genetics, first at the University of Toronto, then at Marlboro College in Vermont. It was there that MacArthur got his undergraduate degree in mathematics, going on for a master’s degree in the same subject at Brown University. This early training in mathematics was central to MacArthur’s influence on ecology. MacArthur received his Ph.D. in biology in 1957 at Yale University under the supervision of G. Evelyn Hutchinson followed by a year of study with David Lack at the University of Oxford. After returning from England in 1958 he joined the faculty of the University of Pennsylvania where he quickly rose through the ranks to full professor. In 1965 he moved to Princeton University where he remained for the rest of his career.

Hutchinson was a theoretical ecologist and Lack was a bird ecologist, and these two, along with MacArthur’s older brother (who was a physicist), shaped MacArthur’s blend of mathematical and ecological interests. He was both a creative mathematician and a skilled naturalist, especially when it came to birds. Not only was he willing to simply sit and observe them for extended periods of time, he was also skilled at taking those observations and putting them into a quantitative framework. This ability was manifest even as a graduate student. His first two publications included a new method for quantifying the structure of communities based on information theory and a theory to explain why communities tend to consist of a few very common species and many rare species. His dissertation examined the way that various species of warbler foraged for insects in different parts of a tree: along the trunk, out on the branches, or on the ground underneath. This study is still considered a classic and earned him the Mercer Award of the Ecological Society of America (ESA) for the best research paper by a young ecologist. Most notable about that work was how he did not simply describe those patterns, but quantified them by inventing new measures.

One reason that MacArthur was so influential was not just that he was a good mathematician, but that he was also a master at taking complex mathematics and presenting them as graphs that could be grasped by those without mathematical training, that is, nearly all other ecologists. The work done by him and his collaborators during the 1960s was a large reason that ecology went from being a mostly a descriptive to a predictive science with an emphasis on hypothesis testing. Ecology also became a more experimental science, although MacArthur’s own empirical work was observational, rather than experimental.

The controversy lay in the theories that he expounded. He gave a predominant role to the process of competition in shaping communities, an influence of his mentors Hutchinson and Lack. He tended to focus on the equilibrial predictions of his theories, rather than the trajectories that would lead to those equilibria, because those solutions were easy to derive mathematically. As a result, those theories focused on the nonrandom aspects of ecological processes, rather than those that depended on contingent processes. These characteristics of his theories – competition, equilibrium, determinism – were the focus of fierce debates among ecologists for the next several decades, long
after MacArthur’s death. Those debates would have pleased him, however. He once noted “There are worse sins for a scientist than to be wrong. One is to be trivial.”

Another important and controversial characteristic of his theories was their simplicity. MacArthur looked to mathematics to provide simplifications of a complex world. He felt that “[s]cience should be general in its principles” and so he typically formulated simple but general models. One of his goals was to replace models that were just descriptions of nature with ones that were built from biological mechanisms grounded in an organism’s biology and natural history. While many would complain that his models were oversimplifications, he felt that providing a basic framework of how the living world works was more important than capturing every nuance of nature’s complexity.

Perhaps his most enduring legacy came from the other characteristic of his work, his striving to unify disparate disciplines. He was working just after evolutionary theory was brought together in the Modern Synthesis (see Chapter 4). While that process showed how evolution could be a central organizing principle in biology, it had not produced models that could be used for that purpose. Thus, in July 1964 a small group of researchers from the universities of Chicago, Princeton, and Harvard met at McArthur’s lakeside home in Marlboro, Vermont. They had two goals, to unite the disciplines of ecology, evolutionary genetics, and biogeography, and to do so in the form of simple mathematical models. They spent two days working toward these goals, during which each person discussed their own work and how to produce such a theory. The result was a series of projects, both separately and, most importantly, in collaboration, that created the discipline of evolutionary ecology.

Among the participants was Edward O. Wilson with whom MacArthur had already been collaborating on a theory to explain the patterns of species numbers on islands. They first put forward this theory in a paper in 1963, but it gained prominence with the slim volume published in 1967, The Theory of Island Biogeography. In a simple graphical model, they showed how you could explain the numbers of species found on islands based on just two factors, the size of the island and the distance from other islands or the mainland. That number was a dynamic equilibrium between the processes of colonization and local extinction. They tied together ecological and evolutionary processes and took a subject, biogeography, that until then had been mostly the providence of descriptive natural history and made it predictive. That theory became one of the foundations of conservation biology because it could be used to predict the number of species that could be maintained in nature reserves of different sizes.

Despite its enormous importance and influence, that book also exposes the weaknesses of MacArthur’s approach. In much of his work, he would attempt to capture the essence of a problem through a simple dichotomy. In this case, the patterns of natural selection were divided into those experienced by newly arriving colonists (fast growth, high mortality rates, allocation of resources primarily to reproduction) and long-established residents (slow growth, low mortality rates, allocation of resources primarily to growth and maintenance); these were dubbed r-selection and K-selection based on terms from an equation for population growth rate. Unrecognized, because of the simplicity of the mathematics, were the many hidden assumptions behind that simple dichotomy. For many years afterward numerous ecologists attempted to put all species into those two boxes, rather than recognizing that you could find species with all possible combinations of those characteristics. Ecologists spent years uncovering the assumptions buried in this and many other of MacArthur’s theories.

Part of MacArthur’s success stemmed from his personality. He had a calm, understated manner, almost shy and reticent, although he could come across as arrogant to some. While he had no tolerance for sloppy thinking or mediocrity, he was quite happy
to take the arguments of others seriously. Notably, he gave personal attention and encouragement to many young scientists beyond his own students and direct collaborators. He ended up mentoring or influencing many people who went on to become leaders and important thinkers in ecology.

He was also very devoted to his family. He married Elizabeth B. Wittemore in 1952 and had four children, with whom he lived in Princeton in a lovely old house. A visitor, Jim Brown, described though that “the interior was, to put it bluntly, not exactly well maintained. I soon learned why. That evening, Robert..., the four MacArthur children, ranging in age from 17 downwards, and I played a ferocious game of ball hockey on the living room floor.”

MacArthur’s life ended on November 1, 1972, at the age of 42, from kidney cancer. During the mere 15 years of his professional career he helped found the discipline of population biology, reinvigorated the study of biogeography, and inspired decades of research. In recognition of his importance to the field of ecology, in 1983 the ESA created the MacArthur Award to honor excellence in research by a mid-career ecologist.
Contingency

Chance can play both a small and large role in the abundance and distribution of organisms (Table 7.1, principle 3). Consider a field with many different species of plants, including grasses and wildflowers, such as goldenrod (Figure 7.14). Each year the goldenrod individuals produce seeds that get blown about in the wind, landing in various places in the field. Some of those places are not near any other plants, and thus are good places to germinate and grow; other seeds will end up next to or directly underneath another plant, so the plants that germinate from those seeds may not grow very big or may simply die. Chance dictates how many seeds land in good places and which seeds those are. Once the seed starts to grow it must compete with the other plants around it for light and for the water and nutrients in the soil. Goldenrod is very different in form from the grasses in that field, and so does not compete very much with them. However, say that asters also grow in that field. Those plants are similar to goldenrods and so compete a bit more. And, of course, two goldenrods growing next to each will compete quite a bit. In addition to simple placement, which species an individual happens to be growing next to also depends a great deal on chance.

At much larger scales, the same sort of process can be seen. On the continent of Australia, the dominant mammals are marsupials like kangaroos and koalas. Placental mammals like humans and horses had not spread to Australia before that continent was separated from Africa, South America, and Antarctica. That chance event led to many differences in the animals of Australia. We can see those differences by comparing the animals to those in Africa which has a similar climate. For one thing, today the largest plant-eater on much of the continent is the kangaroo, which moves at top speeds by hopping on two legs rather than running on four like an African gazelle (Figure 7.15). Up until 40,000 years ago, the largest plant-eater was the Diprotodon (Figure 7.16). It grew to about 3 m in length and 2 m in height and weighed about 2500 kg. It mostly ate leaves of trees.
and shrubs and some grass. In many ways, it was similar to an elephant, but not as large.

Figure 7.16
A. Diprotodon optatum, a large plant-eating marsupial that used to live in Australia before going extinct 40,000 years ago. (Credit: Dmitry Bogdanov, Source: Wikipedia) B. An African elephant, the largest living terrestrial plant-eater. (Credit: Magister, Source Wikimedia Commons)

Figure 7.17

Figure 7.18
Varanus priscus, a giant monitor lizard that lived in Australia before going extinct between 30,000 and 40,000 years ago. (Credit: Arthur Weasley, Source: Wikipedia)

The largest predator in Australia was the marsupial lion which was about the size of a female African lion (Figure 7.17). It could not run as fast as African lions and likely hunted by either sneaking up on its prey and pouncing when close, or by dropping from trees. One important difference was that, unlike the African lion, the marsupial lion was not the undisputed top predator in Australia. It had to share that distinction with a 6 m long monitor lizard (Figure 7.18). No equivalent land-based reptile is found in Africa; the closest similar animal is the crocodile which is aquatic. All of these large animals went extinct about
40,000 year ago, some time after humans migrated to Australia. So, just as their presence in Australia was due to chance events tens of millions of years ago, their eventual extinction was due to yet another chance event, the advent of humans.

Ecologists have debated for many decades about the importance of chance. For the most part, in the first half of the 20th century, most ecologists thought that chance played a relatively small role, although there were always some dissenters. The most vigorous debates were about the role of chance in changes in population numbers, a debate that was renewed about once a decade. This debate intensified during the 1960s and spread to other aspects of individual and species numbers and distributions; as it spread, this debate spurred the rapid development of mathematical models through which the effects of chance could be tested. Finally, by the 1980s most ecologists agreed that chance events could play a significant role in most ecological processes. It is no coincidence that this increase in the appreciation for the importance of chance happened at the same time that computers were becoming faster and easier to use. Models that incorporate random variation require fast computers for implementation. This linkage of computer technology and a change in ecological theory is an example of how changes in technology can drive scientific advancement.

**Variation**

**Individuals**

No two individuals are ever identical. Even individuals that are genetically identical grow up under slightly different circumstances resulting in differences in form or behavior. Differences become more and more apparent the more genetic variation is involved: collections of individuals, populations, will also differ from each other, and most noticeable are the differences between species, many of which are caused by natural selection (see Chapter 4). All of these types of variation can affect ecological patterns and processes (Table 7.1, principle 4), and understanding those effects can have important practical implications when it comes to species conservation.

Because population numbers naturally cycle up and down, one cause of local extinction of that population is if the changes in the environment lead to the death of just a few more individuals during one of those down phases, causing the population numbers to drop to zero. For example, an especially cold winter may result in a few more individuals freezing to death, or a bad drought may mean more plants die from insufficient water. If all of the individuals in a population were identical, then the chance of every individual dying in especially adverse conditions will increase. However, if during that very cold winter, some individuals are a bit fatter than others, then their chance of survival will be greater and at least some individuals will survive. Then, when conditions improve the population numbers will again rise. This effect is why financial advisors tell you never to invest all of your money in a single company, but to always have a diversified set of investments; when the financial crash comes, at least some of those investments will survive.

Conserving biodiversity involves managing species’ populations so that they do not go locally extinct. An important tool in this effort is the use of models that predict the probability of local extinction under various management scenarios, such as how much hunting is allowed or how much land is put into development rather than maintained as species habitat. When conservation managers model extinction risk for different management strategies, they need to take survival variation into account. Depending on the model used, treating all individuals as identical can over- or underestimate the extinction risk.
Variation among species can have similar effects on the stability of entire communities. For example, plant species differ in their ability to grow given different amounts of water. Some species grow best under drier conditions, while others grow best under wetter conditions. Since rainfall can vary substantially from one year to the next, a plant community that contains a diverse array of species varies less in total growth from year to year than a community consisting of just one or two species.

In the diverse community, at least a few species can grow well under any given
Figure 7.19
Temperature (left-hand) and precipitation (right-hand) for the entire Earth (top) and North America (bottom). Used by permission of The Center for Sustainability and the Global Environment, Nelson Institute for Environmental Studies, University of Wisconsin-Madison
amount of rainfall, thus maintaining a minimum biomass. Species diversity alone, however, is not the only factor that affects the structure of communities. As food webs (Figure 7.12) increase in number of species, they can become more sensitive to small disturbances because a change in the number of individuals in one species quickly affects all of the other species. However, if the food web consists of some species that are tightly linked to each other (e.g., a predator that specializes on a single prey species) and others that are only loosely linked (e.g., a predator that eats lots of different species so that it does not rely on any single prey species), then changes in species numbers in one part of the food web do not affect those in other parts.

**Environmental conditions**

Variation in a species’ location is caused by many factors, one of which is variation in the environment (Table 7.1, principle 5). As shown in Figure 7.2, the greatest numbers of species are found in equatorial regions and the fewest numbers are found in polar regions and deserts. This variation in species numbers is associated with differences in annual temperatures and rainfall (Figure 7.19A,B). In general, more species are found in warmer and wetter environments, warmer than it is today (Figure 7.21B).

These sorts of global differences are also found across continents (Figure 7.19C,D) and even at distances of tens of kilometers (Figure 7.20). In addition to distance, time also creates variations; the environment varies in the same site from day to night, from week to week, and across seasons.

The Earth itself has been subject to much longer cycles, such as the advance and
retreat of glaciers over tens of thousands of years. The last warming period began about 12,000 years ago, and since that time, the average global temperature has risen, peaking about 7000 years ago, and then declining (Figure 7.21A). If not for human activities bringing on a new, rapid period of global warming, we would likely be seeing the growth of glaciers in the next few thousand years. Even these global cycles are just a small part of even longer cycles – 50 million years ago the average global temperature was 6°C.

All aspects of the environment vary, each having its own effects on the distribution of organisms. At the ocean’s shore, for example, are places that are always under water, others that are dry part of the day and wet at other times as the tides go up and down, and others that only occasionally get wet, leading to very different species in those various zones (Figure 7.22). Few plants can grow in soils with high amounts of salt, so the spray from the waves limits which species grow near the shore. Sometimes differences in soil conditions can create quite abrupt changes in which specific locations plant species are found (Figure 7.23).

Figure 7.21
A. Temperature changes during the past 12,000 years as recorded at 8 sites around the globe and their average. B. Average global temperatures for the past 540 million years. (Images created by Robert A. Rohde/Global Warming Art)
These large differences in the scale and pace of environmental variation pose challenges to scientists trying to understand the interaction between organisms and their environment. For small organisms, experiments can be done in a laboratory or greenhouse. For global-scale processes, manipulative experiments are not possible, so scientists must rely on mathematical models using correlations between environmental factors and species distributions to formulate these larger-scale theories. Such observational experiments have limitations in what they can tell us about causal mechanisms and are best when they can be paired with manipulative experiments (see Chapter 2). In the past few decades, scientists have been performing much larger and longer manipulative experiments that have provided new insights into ecological processes (Box 7C).

Environmental variation occurs at very short distances and across the entire globe, from moment to moment and from aeon to aeon. As dramatic as this variation can be, not every aspect of it affects all species equally. The relative importance depends on the organism; what a bacterium perceives is not what an elephant perceives. For a bacterium, differences in soil nutrients over distances of micrometers can be critical, a distance that would make no difference to an elephant. A squirrel may live and die within a year, while the oak tree that it climbs might live for hundreds of years. Changes in climate that take decades to happen are not experienced by the squirrel, but can be vital to the life of the oak tree.

Figure 7.22
A rocky coast at low tide along the Pacific Ocean in Washington. Different organisms can live at different heights depending on their tolerance for drying, with the most tolerant living at the highest distance from low tide. (Credit: Bcasterline, Source: Wikipedia)

Figure 7.23
Serpentine Barrens in Rock Springs Nature Preserve, southern Lancaster County, Pennsylvania, USA. This is a Savannah-like ecosystem in an area where the climate would normally support a forest; minerals in the Serpentine group create soils toxic to most plants. (Credit: Alex Zorach, Source: Wikipedia)
Scientists often study processes that occur over large areas or play out over many years. Much of their work is done using strictly observational studies, recording the changes that occur naturally to trace the patterns that define the ecology of an area. Although such observational studies are the only option for some questions (e.g., global patterns of species diversity, Figure 7.2), in recent years, scientists have been performing manipulative experiments that encompass larger areas and extend for longer periods. One example is the long-term study of prairie ecology at the Konza Prairie Research Natural Area in the Flint Hills of northeastern Kansas. It is 3500 hectares in size and one of the largest unplowed tallgrass prairies in North America. Some of the experimental manipulations were begun in 1972 and so have been going on for nearly 40 years. The reserve is divided into a series of large patches, which are subjected to different combinations of controlled burning at various time intervals and grazing by bison or cattle (Figure 7C.1). The intent of these manipulations is to understand the processes that were responsible for community composition prior to European settlement (fire and bison) and how to manage other grasslands that are used to graze cattle.

Figure 7C.1
Large-scale manipulative experiments are being carried out at the Konza Prairie Research Natural Area (A). Prescribed burns (B) are done at various intervals to investigate the effects of fire and fire frequency on prairie communities. In addition, areas grazed by bison (C) are studied and compared with ungrazed areas and with plots subjected to cattle grazing. The experimental patches (D), which are watershed units, vary in size from approximately 3 to 200 hectares. In this map, each patch is designated by a code indicating the fire treatment. (Credits: Alan Knapp, Konza Prairie Biological Station, Scott Collins, Konza Prairie Biological Station)
Figure 7C.2
Comparison of a patch that has been burned every year (top) with one that has been burned every fourth year (bottom). Both are ungrazed. When burning occurs every year the patch is dominated by grasses; less frequent burning allows invasion by woody plants. (Credit: Melinda Smith, John Briggs)
Laboratory, greenhouse and small-scale field experiments are often able to manipulate or control a wide array of factors. In contrast, large-scale manipulative experiments are often limited in the types of treatments that are possible. For example, at Konza Prairie, almost all of the controlled burning is carried out in the spring because later in the summer or fall the prairie would be so dry that any experimental burn would run a high risk of burning a much larger area than intended. Large scale experiments are also limited by historical information; because we do not have firm data on prairie fire regimes prior to European settlement, we do not know how this spring-burning treatment compares to “natural” fire regimes.

Being able to study and manipulate a system for decades has led to insights that would not have been possible in the 3 to 5 years of most ecological studies. One insight was the effect of bison on the pattern of spatial variation in the vegetation. The activities of bison, which were introduced to Konza Prairie in 1987, make the landscape much more variable, thereby increasing the total number of plant species that can co-exist. They are allowed to freely roam over part of the area (Figure 7C.1) so that their effects mimic as closely as possible those in pre-settlement days. Because bison eat mostly grass, leaving other types of plants untouched, those other plants flourished when otherwise they would be have been out-competed by the grasses. In addition, this grazing occurred more often in areas that were burned. By itself, burning increases the dominance of grasses over other species, but this effect is moderated by bison. Other behaviors of bison include the production of wallows, depressions of bare soil about 3-5 m in diameter created by pawing the ground and rolling in the exposed soil. These areas fill with water in the spring and later contain a unique set of plants. Bison create patches that are high in nitrogen through their urine deposits. When a bison dies, its carcass decays and creates an area that is high in nutrients. All of these effects of bison on environmental variation and the subsequent changes in the vegetation became apparent only after more than a decade of bison grazing.

These sorts of large-scale manipulative experiments are more important than ever today in our need to understand the potential effects of global warming. Scientists at Konza Prairie and elsewhere are manipulating the amount of rainfall that gets to the plants through the use of large, moveable canopies, and manipulating local temperatures through the use of heaters. At other places, scientists are also manipulating CO2 levels in large patches of forest and grassland. The goal of these large-scale experiments is to mimic the possible changes that might happen to the climate over the next century, and thus to understand how changes in these factors, separately and together, might affect the vegetation. Because global warming is a long-term process, large-scale manipulative experiments are a key component in understanding its effects.

Finite resources

An important distinction among types of environmental factors are conditions versus resources. Conditions are unaffected by organisms – for example, temperature is a condition in that it affects the organism, but no matter how warm or cold the organism gets the organism does not increase or decrease the amount of heat in the surrounding environment. In contrast resources are finite and potentially depleted by organisms. If foxes are eating rabbits, there are only so many rabbits to go around. Every rabbit eaten means one fewer available for another animal to eat. Resources, therefore, are subject to competition, unlike environmental conditions; both environmental conditions and resources, however, vary in space and time (Table 7.1, principle 6). The heterogeneity of resources has a number of causes.

Some resource heterogeneity is caused by the action of organisms, especially Bacteria and Archaea. For example, nitrogen is often a limiting nutrient for plant growth.
Most of the Earth’s nitrogen is found in the atmosphere in the form of N2, which plants cannot use; they need nitrogen to be in the form of NH4+, NO2− and NO3−. Nearly all of the conversion of atmospheric nitrogen into these forms is done by Bacteria (Figure 7.24). Bacteria and Archaea, as well as fungi, are primarily responsible for breaking down dead tissues, thereby making the nutrients that they contain – such as nitrogen compounds – available for other organisms.

For some species, the critical environmental variation is the presence of other species that serve as resources. Monarch butterfly caterpillars grow only on milkweeds (Figure 7.25). The measles virus infects only humans. Woodpeckers eat insects that live inside trees. Variation in environmental conditions can have indirect effects on the distribution and abundance of species through their affects on organisms that serve as resources.

The processes that cause variation in resources are often the same as those that cause variation in environmental conditions. For example, seasonal variation in light and temperature are caused by the movement of the Earth around the Sun. However, light is subject to competition (e.g., when one plant shades another) whereas temperature is a condition and not subject to competition.

Whether a particular environmental factor is a condition or a resource is context dependent. For example, water is sometimes a resource subject to competition (e.g., plants in a desert) and sometimes a condition (e.g., fish in the ocean). Some elements (e.g., manganese) can be limiting to plants at low levels, thereby acting as a resource. On the other hand, that same element can be toxic at high levels, in that case acting as a condition (Box 7D). Thus, while the distinction between conditions and resources is important, the two categories blur into each other.
Figure 7.24
The nitrogen cycle. (Source: U.S. Environmental Protection Agency)
Although the majority of modern scientific breakthroughs appear first as publications in journals, it is also sometimes possible to have an enormous influence on a scientific discipline while never publishing a single paper in a professional journal. This feat was accomplished by one of the most well-known figures in ecology, Rachel Carson. Born on May 27, 1907 in the small town of Springdale, Pennsylvania, Carson grew up in close proximity to nature, a relationship fostered by her mother. She spent much of her childhood exploring the pond, fields and forests of their 65-acre farm. She had a sense of wonder in her appreciation of the natural world that she retained into adulthood and made a centerpiece of her writing.

From her earliest childhood she knew that she wanted to be a writer, publishing her first story at age 11 in the children’s periodical, *St. Nicholas Magazine*. She tended to be a loner and did a lot of reading; at a young age she was particularly enamored with Beatrix Potter, and later, writers such as Robert Louis Stevenson and Herman Melville, both of whom wrote about life at sea.

After graduating first in her high school class in 1925, she attended the Pennsylvania College for Women (now Chatham University), in Pittsburgh. She began as an English major, but eventually graduated magna cum laude with a degree in biology in 1929. That summer she studied at the Marine Biological Laboratory in Woods Hole, Massachusetts. She found herself in a picture-book seaside village, initiating a lifelong passion for the seashore. From there she began graduate studies at Johns Hopkins University where she worked with Raymond Pearl, a human biologist. She was only able to afford to be a part-time student, working as a laboratory assistant and doing experiments with rats and fruit flies in order to earn money for tuition. After three years she completed a master’s degree working on the developmental patterns in fish. Although none of this work involved the science of ecology, from her advisor she gained a holistic view of biology as a way of understanding the human condition.

She planned on staying on and earning a Ph.D., but was forced to leave after a year to help support her family. The following year, in 1935, her father died suddenly, leaving Carson to earn a living in the midst of the Great Depression. Eventually she would take on responsibility for other family members: In 1937 her older sister died, and Carson was left as the sole breadwinner for her mother and two nieces. Twenty years later, when one of those nieces died at the age of 31, leaving a five-year-old orphan son, Carson adopted him while continuing to care for her aging mother. These actions were part of her overall reverence for life, which ran so deep she replaced a starfish in the spot from which she had collected it when she was finished studying it. She never married, and when the firestorm erupted over her book *Silent Spring* that fact was used as proof that she was a communist.

This need to support a family had an important effect on her career. Based on the advice of her undergraduate biology mentor, Mary Scott Skinker, in 1934 she took what was to have been a temporary position at the federal Bureau of Fisheries (later the Fish and Wildlife Service), where she wrote scripts for a weekly educational radio show, “Romance Under the Waters.” On her civil service exam, she outscored all other applicants and was only the second woman to be hired by the Bureau for a full-time, professional
position. She stayed there for the next fifteen years, eventually rising to Editor-in-Chief of all publications for the Fish and Wildlife Service. While in that position she published many natural history pieces in newspapers and leading magazines.

In the 1950s she published a trio of books – *Under the Sea Wind*, *The Sea Around Us*, and *The Edge of the Sea* – that established her reputation as a writer. The first book was based on a magazine piece that she had originally written for a Fisheries Bureau brochure, but that her supervisor had deemed too good for that purpose. A publisher was highly impressed and asked her to expand the piece into an entire book.

Carson’s strong background as a writer came through in her notes for the book, addressing technique and language as well as raw content. Among other questions, she specifically asked “What age child do editors prefer to attract?” This lead to the inclusion in the book of the story of Scomber the mackerel. Still, Carson made sure that the book reflected the best science known at the time.

Her first book sold poorly, and it wasn’t until she published her second book that she was able to quit her government job in 1952 and become a full-time writer. Chapters from that book were published in several leading magazines, and one piece won the American Association for the Advancement of Science’s George Westinghouse Science Writing Prize. The book was a bestseller, remaining on the New York Times list for 86 weeks, and was translated into over 25 languages. It won both the National Book Award and the Burroughs Medal, and resulted in Carson being awarded two honorary doctorates. Her first book was also republished, selling much better this time.

After the publication of her third book, Carson was looking for a new project and decided to focus on the issue of pesticide use. She first encountered the topic in 1945, but it was the extensive spraying of DDT in the 1950s by the Department of Agriculture to control the spread of the fire ant that prompted her concern. In 1957 the Audubon Society, which actively opposed such spraying programs, recruited Carson to publicize these practices. For the next four years, Carson would research this topic, gathering examples of environmental damage attributed to DDT and using her connections to government and academic scientists to collect information, some of it confidential.

The result of her efforts revealed a split that was to be repeated over and over. On the one side were various scientists and citizens raising concerns about possible environmental harms of industrial and governmental practices. On the other side were industry, governmental representatives, and scientists who would minimize those concerns and point out the important economic benefits of those same practices. The case for pesticides is particularly complex – the direct benefits include control of agricultural pests resulting in more and cheaper food, stopping the spread of invasive species like the fire ant, and improving health in developing countries by killing mosquitoes and preventing the transmission of malaria. What Carson focused on were the indirect effects, all of the other insects killed by indiscriminant spraying and the failure of eagles to reproduce because of the weakening of eggshells by DDT that they absorbed from the fish they ate.

The result of these efforts was *Silent Spring*, published in September of 1962, but first serialized in *The New Yorker* magazine beginning in June of that year. The title was originally just the heading of the chapter on birds, but based on a suggestion of her literary agent Marie Rodell, it became a symbol of a bleak future for the entire world. Although vehement attacks began even before publication, the book withstood it all; the years Carson had spent researching meant the science could not be contradicted. Along with her rigorous review of the scientific literature, Carson documented hundreds of individual incidents of pesticide exposure, human sickness, and ecological damage. The book was also timely; in 1959 the “Great Cranberry Scandal” broke; the 1957, 1958 and 1959 crops of U.S. cranberries were found to contain high levels of a cancer-causing
herbicide, resulting in the halt of the sale of all cranberry products. In addition, just before the book was published the story broke about the birth defect-causing drug thalidomide. The general awareness of these threats to public health were simply increased by the criticisms of the book.

Silent Spring has an enduring legacy. In its immediate aftermath, there was both a congressional review and a report by the President’s Science Advisory Committee on the effects of pesticides in the environment. The book helped galvanize the environmental movement, including the 1967 formation of the Environmental Defense Fund, and the eventual banning of DDT in 1972. This ban, with respect to agriculture, was extended worldwide by the Stockholm Convention in 2004. The concerns raised in the book led to the National Environmental Protection Act of 1969 and the creation of the Environmental Protection Agency in 1970. For professional ecologists, the book helped push their discipline from one focused on basic research of pristine systems to one that gives equal weight to applied questions in human-dominated systems such as urban areas.

The book demonstrates that scientific research can take many forms and that the synthesis of information is just as important and original in its way as the collection of new data. It shows that good science need not be divorced from good writing and can be found in a work intended for a popular audience. Despite all of the criticisms of her science, the eventual assessment was that while the book contained some factual errors, all were minor. The major conclusions remain unchallenged.

Following its publication, Carson received a flurry of awards and honors; the Audubon Medal from the National Audubon Society, the Cullum Medal from the American Geographical Society, and induction into the American Academy of Arts and Letters. In 1980, Carson was posthumously awarded the Presidential Medal of Freedom, the highest civilian honor in the United States, by President Carter. Numerous buildings, bridges, and natural areas have been named in her honor. Carson is also a frequent namesake for prizes awarded by philanthropic, educational and scholarly institutions.

She is still a controversial figure. On the one hundredth anniversary of her birth, a resolution presented in the U.S. Senate to honor her was blocked with the accusation that Silent Spring was responsible for banning pesticides that controlled disease-spreading mosquitoes. In fact, Carson never called for a ban on DDT, just that all pesticides should be used responsibly.

Carson did not live to see all of these results. While working on Silent Spring in 1960, she discovered cysts in her left breast, one of which necessitated a mastectomy. By December she discovered that the tumor was malignant and the cancer was spreading. Weakened by the cancer and treatment regimen, she became ill with a respiratory virus in January 1964. By March the cancer had reached her liver, and Carson died of a heart attack on April 14, 1964.
Fitness

Birth and death

The evolutionary fitness of an organism is the result of that individual’s life history and reproduction (see Chapter 4). Organisms are born through a complex process of cellular and individual replication and die due to trauma or starvation (see Chapters 5 and 6). Those birth and death processes are at least partially dependent on the environment of the organism (Table 7.1, principle 7), both environmental conditions and resource availability.

Reproduction requires the resources necessary to build a new cell or to birth an offspring and to raise that offspring to independence. The rate at which an organism reproduces – for example, how quickly a single-celled organism can divide or how many eggs a bird can lay – depends on the rate at which that individual can take up nutrients from the environment. It is more than simply how fast an animal can eat, however. Often, one particular nutrient will be the limiting factor. For example, nucleic acids contain a lot of phosphorus. Because gametes are mostly DNA, the rate of gamete production can be limited by the amount of available phosphorus. Some species of fruit flies live on rotting mushrooms, which have a low phosphorus content. Those fruit fly species reproduce much slower than others that live on rotting fruit, which has a higher phosphorus content.

Sometimes the limiting resource is other individuals of the same species. In general, sperm are much smaller than eggs (see Chapter 6), so often the males in a population have many more sperm than the females have available eggs. This leads to competition among the males for access to those females (Figure 7.9). In other cases, where males provide access to limited resources such as food-rich territories, females may compete for access to those males. In some circumstances where individuals are widely separated, such as tigers where each adult male or female will control a large territory, simply finding another individual to mate with can create limitations on the rate of reproduction.

Competition for limited resources determines the numbers of individuals that can live in an area. Such competition can occur between individuals within a species – Canada lynx competing for snowshoe hares (see Figure 1.4) – or among different species – arctic foxes eating those same hares (Figure 7.26). If the persistence of a population depends on it being above some minimum size, the amount of available food would limit not just the number of individuals but also the number of species. Consider, for example, if the minimum population size of both the lynx and fox was 100 individuals; there would have to be enough hares for 200 individuals for both species to be able to persist in an area. If there were only enough hares for 120 individuals, then one species would go locally extinct.
One way that the lynx and fox can co-exist, even if they both eat hares, is if at least one species also eats something else. For example, foxes might be better at catching small animals like mice, while the lynx might be better at catching the larger hares. Although the two species compete for some food resources, they differ enough in which foods they eat so that each species is competing more with individuals of its own species than of the other species. The result is co-existence.

Scientists have developed mathematical models to describe the dynamics of population numbers (Box 7D). These models provide insight into the role of contingency. Population numbers are often limited by a resource. As numbers increase, the rate of reproduction per individual decreases until, on average, each individual in the population is being replaced each generation. However, consider the following scenario in which there is enough food to support 100 foxes. When there are 100 animals, 50 females and 50 males, each pair produce exactly 2 offspring and population numbers stay the same from one generation to the next. But, when there are 50 foxes (25 pairs) each pair has enough food to produce 6 offspring. In the next generation there will be 150 individuals \((6 \times 25)\). Now there is only enough food for each of the 75 pairs to produce one offspring, so that in the next generation there will be 75 individuals. This cyclical pattern occurs because population numbers overshoot, either too high or too low, the point at which the number of individuals would be stable.

![Figure 7.27](image)

**Figure 7.27**
Models of population dynamics as a function of the base reproductive rate (number of offspring per adult). The base reproductive rate is measured under conditions when resources are not limiting, i.e., at low population numbers. This rate declines as numbers increase. (A) When the base rate is low, population numbers may smoothly approach the stable number, or may bounce between two values above and below that number. (B) At higher base rates, the numbers may bounce between several different values, or change in such a way that they never return to exactly the same value.

Depending on the details of how many offspring are produced for a given number of adults, population numbers can vary in a variety of ways. The size of the population may eventually converge to the stable number, or it may continue to bounce between sizes above and below that number in a regular fashion (Figure 7.27A). Or, if the reproductive rate (the number of offspring per adult) is very high at low population numbers, the numbers could change in a complex pattern (Figure 7.27B). In such circumstances, the pattern of population numbers in subsequent generations is very sensitive to the exact number in
Evolution

The process of change over generations (evolution) and environmental interactions (ecology) are intertwined. The biology of organisms creates the context within which evolution occurs and, in turn, evolution determines the properties of organisms (Table 7.1, principle 8). The environment creates context in two ways. First, the phenotype of the organism is partially determined by the environment within which genes are expressed (see Chapter 3). Second, the relationship between the phenotype of an organism and its fitness is determined by the organism’s ecology; thus, the direction of natural selection is a function of the environment (see Chapter 4).

It may seem obvious that the ecology of an organism has been shaped, at least in part, by its evolutionary history. But such evolutionary thinking was not always common among ecological scientists. The inclusion of evolution within ecological thinking was an important outcome of the Modern Synthesis (see Chapter 4). Although evolutionary thinking about ecological processes goes back at least to Darwin, evolutionary thinking had been infusing ecology in a significant way since the 1920s and its widespread acceptance occurred primarily in the latter half of the 20th century.

Even today, most ecological studies do not take into account evolutionary processes. In some instances, evolution can be safely ignored; for example, the cycles of population numbers of snowshoe hares and Canada lynx (see Figure 1.4), are not likely to be affected by the evolution of either species. On the other hand, cycles of measles outbreaks (Figure 7.4) may be affected by evolution, as even within a single person the measles virus is going through many generations allowing many opportunities for mutation and new genetic variation. This means that in the case of measles, evolutionary change is happening on the same time scale as ecological processes. Natural selection can lead to substantial change within tens of generations, so that even in longer-lived species like fish, mammals and trees evolutionary changes can interact with ecological dynamics.

Such interactions of ecology and evolution can be seen when two species interact and evolve in response to each other. Consider the case of the Australian wild flax plant and a fungus that can infect it, causing flax rust (Figure 7.28). A long-term study was undertaken in southeastern Australia beginning in 1986. In 1990, the population being monitored suddenly crashed, dropping from a density of 22.6 individuals/m$^2$ to 8.5 individuals/m$^2$ in just one year. A new strain of rust had appeared in the population, one to which most plants were susceptible. However, within a single year the percentage of the population that was resistant to that new strain increased from about 3% to over 20%, and as a result the numbers of individuals in the population began to climb. This change in the average resistance of the population is an example of very fast evolution in response to strong natural selection.

Accounting for evolutionary history can lead to new insights into how and why species sort themselves into communities. For example, approximately 150 species of hummingbird are found in western South America in the Andes Mountains and the adjacent lowlands. Closer examination
shows that different, related groups of species occupy different habitats. The groups called Brilliants and Coquettes are common at high elevations, and the groups called Hermits and Emeralds are most often found at low elevations (Figure 7.29). The reason for this difference is that at high elevations the amount of oxygen is lower and the air is generally thinner. Flying under those conditions requires larger wings and fewer wing flaps per second, and only the Brilliants and Coquettes had individuals with the right mutations to live under those conditions. Once the appropriate adaptations appeared, those groups diversified in those high elevation habitats. In this case, a contingent evolutionary event determined which groups would be most common in the different habitats.

Figure 7.29
A. A Collard Inca from the Andes Mountains (Credit: Tad Boniecki, Source: Wikipedia)
B. A Saw-billed Hermit from the lowland forests of Brazil. (Credit: Dario Sanchez, Source: Wikipedia)

What is a Community?

A community is a group of populations that coexist in space and time and interact with one another directly or indirectly. By “interact” we mean that they affect one another’s population dynamics. This definition of “community” includes all plants, animals, fungi, bacteria, and other organisms living in an area. Because a community encompasses all of the organisms and their interactions, it is often considered the basic ecological unit. Despite this central role of communities in ecological thinking, how one delimits a community has been the subject of vigorous debate by ecologists for over a century. In practice, the boundaries of communities are usually defined by differences in the abundances of the most common species. Because these abundances may change gradually and only involve some of the inhabitants, these boundaries are often hard to place. Only in special cases (e.g., islands, ponds, forest preserves surrounded by suburban development, vacant lots) are community boundaries defined easily. Even then, the movement of organisms and the transport of matter by wind and water make their boundaries fuzzy. Ecologists, therefore, are often of two minds when dealing with communities. On the one hand, they recognize that their boundaries are inexact; on the other hand, they often need to define discrete entities for convenience of analysis. Typically, ecologists define a community based on the relative uniformity of the resident species and use their knowledge of species biology to decide when they are moving from one community to another.

Ecologists based in different countries and educated in different historical traditions tend to view communities in somewhat different ways. In particular, ecologists in continental Europe were historically influenced by an approach that emphasizes the discreteness of communities. In contrast, most ecologists in English-speaking countries have been more strongly influenced by the history described in the following paragraphs; as a result, they tend to think of communities as blending continuously into one another. These distinct ways of thinking are becoming less prominent as a result of increased travel and communication among ecologists worldwide.

Within ecology, especially plant ecology, there has been, and still is, a range of views on the nature of communities. The extremes are sometimes labeled the Gleasonian and Clementsian views, named after Henry A. Gleason and Frederic E. Clements (Box 7A), their first major proponents in the English-speaking world. These two extreme
viewpoints differ in the importance they ascribe to biotic versus abiotic factors and predictable versus random processes in shaping community structure. Today, most ecologists hold a middle ground between these viewpoints, and to a large extent have moved beyond both of them.

The Clementsian view was the majority view among English-speaking plant ecologists during the first half of the 20th century. Clements saw plant communities as highly organized entities made up of mutually interdependent species. In his view, communities are **superorganisms** – the large-scale analogue of individual organisms – that are born, develop, grow, and eventually die. Two of the hallmarks of the superorganism concept were the presence of very tight linkages among species within communities and cooperation among these species in order to benefit the function of the entire community.

Even at the height of Clements’s influence, many ecologists held more moderate views. The moderate version of Clementsian ecology asserted only that species interactions such as competition, mutualism, and predation are important in determining community structure. Clements focused on the idealized nature of communities, and saw them as spatially distinct, with one superorganism complex giving way to another with a very different collection of species. His major focus was on the nature and development of the community as a superorganism, however, rather than on the boundaries between communities. The more moderate version of this view admits that communities are not entirely discrete, but still divides them into non-arbitrary groups with recognizable boundaries.

Although Clements himself acknowledged the effects of abiotic factors such as site history and soils in determining community composition, his conception of a community as being defined by absolute spatial boundaries stands in striking contrast to Gleason’s more holistic approach. Gleason posited that communities are the result of interactions between individual species and the environment (biotic and abiotic factors) in combination with chance historical events. Each species has its own environmental tolerances and thus responds in its own way to environmental conditions. The implication of this belief was that along an environmental gradient, different species would have their boundaries at different places. Not only were communities not tightly linked superorganisms, but defining the collection of species living together in a particular place as a community was an arbitrary human construct.

According to the Gleasonian view, within the range of environmental conditions a species can tolerate, chance events determine whether a species is actually found in a given location. At the local scale, chance dictates whether a seed happens to get to a particular spot. On larger scales, the chance events of history play a strong role. For example, species in the cactus family are found in the desert communities of the Americas because the family happened to originate in this region, while deserts in Asia, Africa, or Australia have no cacti (except where they have been recently introduced by humans). Furthermore, the mix of species changes from place to place as one moves across the landscape. The Gleasonian viewpoint posits gradual changes in community composition as opposed to abrupt boundaries between communities unless there are abrupt and large environmental boundaries. A more moderate viewpoint was that some identifiable community types exist, but that these tend to blend into other community types.

Today most ecologists take a middle position between Clements’s and Gleason’s views, and in many ways have diverged from both of their views. There is wide agreement that each species has its own pattern of distribution, and that community composition typically changes gradually along environmental gradients. Abrupt changes are most likely found where there are abrupt changes in the environment (Figure 7.SER).
However, abiotic boundaries and community boundaries do not always match. Because of processes such as dispersal from one habitat into another, a population may extend partway into an unfavorable environment. Abrupt changes may also reflect past events, such as the edge of a fire or a part of a forest that was plowed at some time in the past, which means that, current environmental boundaries do not always match past boundaries. Ecologists still disagree as to the relative importance of biotic and abiotic processes and chance events in determining community structure.

Thus, the extent to which communities are “real” has been a contentious issue among ecologists for much of the 20th century. The heart of the debate has been philosophical: What types of entities are real, and what types are just intellectual constructs? Are communities real entities, or are they merely convenient but arbitrary human inventions? In the past, these questions often focused on the problem of boundaries – of identifying where a community begins and ends. However, since many studies have shown that, except where there are abrupt physical discontinuities, communities tend not to have discrete boundaries.

One alternative to that debate is to ask a different question: Are community-level processes important in structuring the living world? We have already discussed several processes responsible for interactions among species: competition, herbivory, and mutualisms. These are all community-level processes that occur among the component parts of communities (e.g., populations). If such processes are significant factors in the structuring of a given set of populations, then we can regard that system as a community with unique properties. Such an outlook eliminates the need to worry as much about the existence of clear boundaries between communities. We can recognize the existence of communities if it is useful to do so, and ignore them if it is not.

The debate as to whether communities are real entities or imaginary constructs is more than just an academic exercise. For example, The Nature Conservancy (TNC), based in the United States, makes decisions about land acquisition and restrictions on land use based on the classification of the communities present. In collaboration with the Natural Heritage Network, TNC has devoted substantial effort to describing and classifying communities under the U.S. Natural Vegetation Classification (USNVC) system. Classifications of communities are even incorporated into law in a number of places. In southern California, for example, land development is regulated quite differently for “coastal sage scrub” than for “chaparral” communities, although by any definition the two include many of the same species, and different scientists may have used either name to categorize the same place.
Glossary

**abiotic** The nonliving parts of the environment.

**allele** Different versions of the same gene.

**anabolism** The synthesis of complex substances from simpler ones; one of the two linked processes that comprise metabolism.

**anaerobic metabolism** Metabolism in the absence of oxygen.

**biological species concept** A group of actually or potentially interbreeding organisms that are reproductively isolated from other such groups.

**biotic** The living parts of the environment.

**catabolism** The breaking down of complex substances into simpler ones; one of the two linked processes that comprise metabolism.

**chloroplasts** Specialized substructures in eukaryotic cells that capture energy from sunlight.

**cilia** Short fibers that extend from the surface of a cell and often used for propulsion.

**codon** Triplets of DNA bases that make up the genetic code.

**community** A group of populations that coexist in space and time and interact with one another directly or indirectly.

**consilience** The idea that the interdependence of scientific theories requires that they be consistent with each other.

**cytoskeleton** A meshwork of fine fibers that extends throughout a cell that gives them shape and coherence.

**degenerate** The characteristic of the genetic code in which a given amino acid is coded for by more than one possible combination of DNA bases.

**diploid** An individual with chromosomes in pairs so that it has two copies of each gene.

**dominant** An allele which is expressed preferentially in heterozygous individuals.

**ecosystem** All of the species in a community including their interactions with each other and the abiotic environment.

**egg** In species which produce two types of gametes, the larger type (see also ovule).

**empiricism** The school of thought in philosophy that understands the world by producing theories that can make useful predictions through logical deduction from basic assumptions combined with data gathered from previous experiments without assuming that the theories necessarily capture any basic truth about the world.

**endoplasmic reticulum (ER)** A membrane containing ribosomes that manufacture proteins.

**enzyme** A protein that directs or speeds up a chemical reaction.

**epistemology** The branch of philosophy that studies ways of knowing.

**essential amino acid** An amino acid which an organism cannot synthesize, and thus must take in via consumption of food.

**evolution** Changes in successive generations of a population of organisms due to changes in their genetic makeu.

**experiment** A test of an hypothesis.

**female** In species which produce two types of gametes, an individual that produces the larger type.

**flagella** Long fibers that extend from the surface of a cell used for propulsion.

**food web** The network of interactions within a community based on which species are eating each other.

**gamete** A reproductive cell.

**gene** The fundamental unit of biological information.

**genetic drift** Changes in gene frequencies due to random sampling effects.
genotype The genetic makeup of an individual.
genotype-environment interactions Differences in genetic expression as a function of the environment.

germ line In animals, the cells that will produce gametes.
Golgi apparatus A complex of vesicles and folded membranes within the cytoplasm of most eukaryotic cells, involved in secretion and intracellular transport.
haploid An individual with a single copy of each chromosome and each gene.
heritability The amount of resemblance among relatives that is due to shared genes.
hermaphrodite In species which produce two types of gametes, an individual that produces both types.
horizontal gene transfer The process of moving a gene, a whole gene, or a large number of genes from an individual of one species to an individual of another species, usually in reference to distantly related species.
hybridization The process by which individuals of different species mating and produce offspring.
hypothesis A proposed explanation for a series of observations.
learning Changes in behavior resulting from experience
lichen A symbiotic organism consisting of an alga living inside of a fungus
life history The series of events in which an organism is born, grows, reproduces and dies.
lysosome A structure in eukaryotic cells that contain degradative enzymes.
macromolecule A very large molecule.
male In species which produce two types of gametes, an individual that produces the smaller type.
manipulative experiment A test of an hypothesis based on a deliberate change in the physical world.
meiosis The process during which a diploid individual produces haploid gametes.
meristem cell In plants, an unspecialized cell at the tip of a stem or root where active growth is occurring.
metabolism The set of co-coordinated chemical reactions that provide organisms with energy.
Metabolism First theory The idea that life first arose on the Earth as simple organic molecules that could replicate through a series of coupled biochemical reactions.
methylation The replacement of a hydrogen atom with a CH3 group in the cytosine of Eukaryotes or the adenine of Bacteria.
migration The movement of organisms or propagules.
mitochondria Specialized substructures in eukaryotic cells that process energy.
mitosis The process of chromosome duplication and sorting into daughter cells.
model An abstraction or simplification that expresses structures or relationships.
module A part of a system within which components interact a lot with each other, and much less than with other parts of the system.
multicausality A single outcome that can arise as a consequence of a number of different components, or multiple causes act together.
mutation A change in information, possibly leading to a change in the characteristics of an organism; a change in the sequence of the DNA molecule.
natural experiment A test of an hypothesis based on a change in the physical world caused by some natural occurrence.
niche The ecological role and space that an organism fills in an ecosystem.
nucleotide A biochemical unit of information storage consisting of a nucleic acid, a sugar molecule, and one or more phosphorus atoms.
observational experiments  The systematic study of natural variation undertaken to test a hypothesis.

ontogeny  An individual’s biological development.

ovule  In species which produce two types of gametes, the larger type (see also egg).

peroxisome  A structure in eukaryotic cells that contains and isolates some enzymes and their breakdown products.

phenotype  The physical characteristics of an individual.

photosynthesis  The process by which organisms capture light energy and store it as chemical energy.

phylogenetic species  A smallest group of organisms that share an ancestor and can be distinguished from other such sets.

phylogeny  A species’ evolutionary history.

plasmid  A small, circular piece of DNA found in bacteria consisting of just a few genes.

pollen  In plant species which produce two types of gametes, the smaller type (see also sperm).

primary research  Gathering information or finding fact not previously known before, and the generation and testing of scientific hypotheses.

protein  A biochemical compound consisting of chains of amino acids.

realism  The school of thought in philosophy that holds that theories can tell us truths about the natural world; as more accurate data is gathered about the world, theories are refined and expanded to more accurately reflect how the world works.

recombination  The bringing together of DNA sequences in new combinations.

reduction-oxidation reaction  The transfer of electrons from one substance to another.

Replication First theory  The idea that life first arose on the Earth as self-replicating RNA molecules.

reproductive rate  The number of offspring per adult organism.

scientific method  Techniques for investigating and gathering measurable evidence about natural phenomena, using steps of systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses.

secondary research  Gathering data or confirming facts that are already known, usually undertaken by students or the general public.

senescence  An increase in the probability of dying with age.

social constructivism  The school of thought in philosophy that posits that theories are supported by the consensus opinion of the scientific community rather than drawn from facts and data.

speciation  The process by which one species gives rise to one or more new species.

sperm  In species which produce two types of gametes, the smaller type (see also pollen).

spontaneous generation  The idea that life can arise from non-life.

spore  A diploid reproductive cell.

stem cell  In animals, an unspecialized cell.

superorganism  A theory that an ecological community is an analogue of an individual organism.

symbiosis  An interdependent or mutually beneficial relationship between two or more entities.

systematics  The science of determining the relationships among organisms.

The Central Dogma  The idea that information flow in a cell is only in one direction, from DNA to RNA to proteins.
**transcription**  The process by which the information contained in a DNA molecule is copied to a RNA molecule.

**translation**  The process by which the information in a RNA molecule is used to create a protein.

**vacuole**  A vesicle in plant cells that contains fluids and organic molecules.

**vesicle**  A membrane sphere that stores and transports materials within eukaryotic cells.

**vitamin**  An organic molecule that is essential for normal growth and nutrition and are required in small quantities in the diet.

**zygote**  A diploid, unicellular entity that is formed at fertilization, which by mitosis grows into a multicellular embryo.