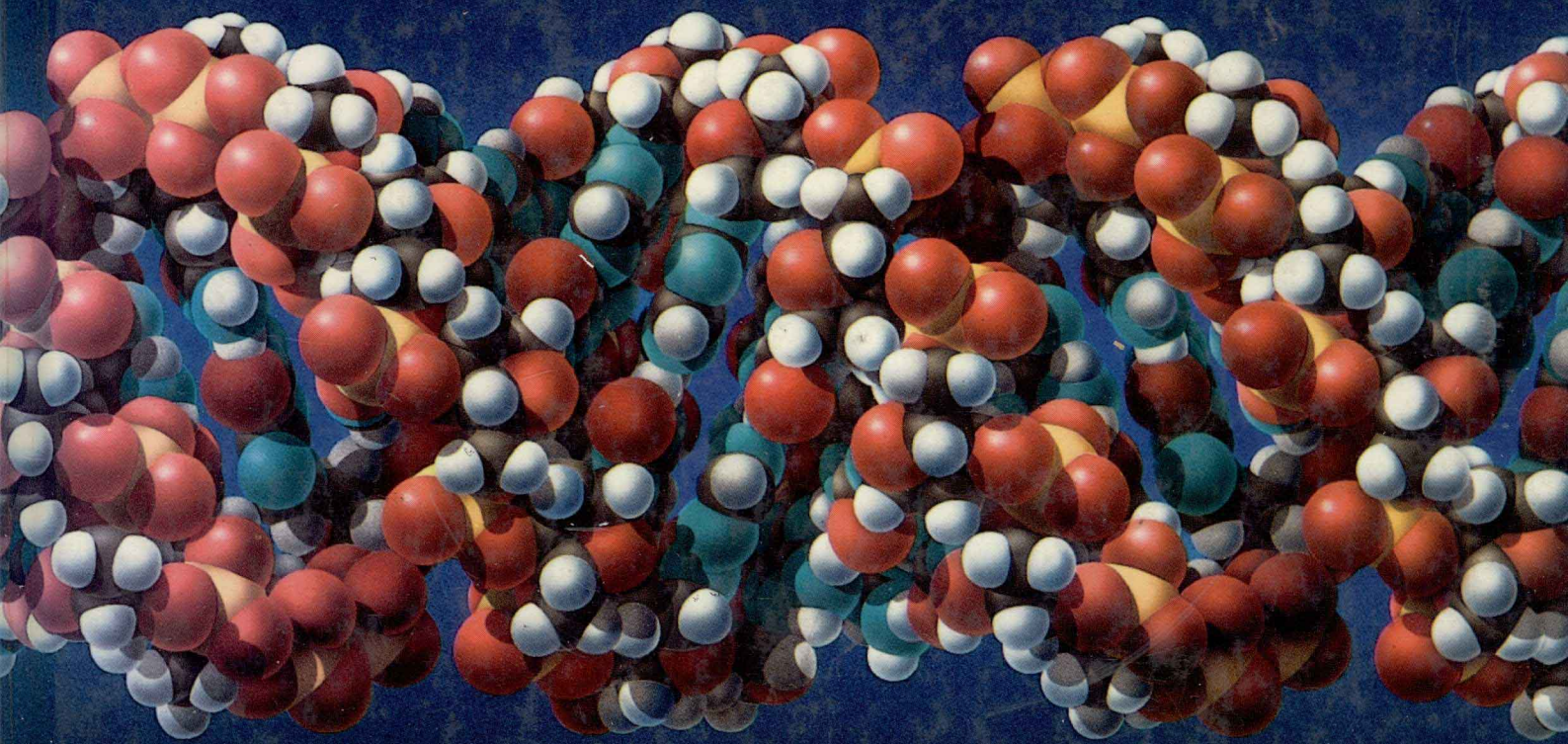
AN INTRODUCTION TO

GENETIC ANALYSIS



FOURTH EDITION

SUZUKI • GRIFFITHS • MILLER • LEWONTIN

An Introduction to Genetic Analysis

FOURTH EDITION

David T. Suzuki
University of British Columbia

Anthony J. F. Griffiths
University of British Columbia

Jeffrey H. Miller
University of California, Los Angeles

Richard C. Lewontin
Harvard University

W. H. FREEMAN AND COMPANY / NEW YORK

Cover illustration by Tomo Narashima

Library of Congress Cataloging-in-Publication Data

An introduction to genetic analysis, 4th ed. / David T. Suzuki [et al.]
p. cm.
Bibliography, p.
Includes index.
ISBN 0-7167-1956-8. ISBN 0-7167-1996-7 (pbk.)
1. Genetics. 2. Genetics—Methodology. I. Suzuki, David T., 1936–
QH430.I62 1989 88-34252
575.1—dc19 CIP

Printed in the United States of America

3 4 5 6 7 8 9 0 KP 7 6 5 4 3 2 1 0 8 9

Preface

■ The great power of modern genetics and its position of prominence in biological research have grown from a blend of classical and molecular techniques. Each analytical approach has its unique strengths. Classical genetics is unparalleled in its ability to explore uncharted biological terrain; molecular genetics is equally unparalleled in its ability to unravel cellular mechanisms. It would be unthinkable to teach one without the other. By giving both due prominence in this book, we have attempted to present a balanced view of genetics as practiced today.

The partnership of classical and molecular genetics has always presented a teaching problem: the order and manner in which the two partners should be introduced to the student. A common solution is to introduce both classical and molecular sides of each genetic principle concurrently. This approach makes use of the wisdom of hindsight — the teacher or writer addresses the student from the lofty vantage point of current understanding of the diverse parts of genetics and how they interrelate. We have chosen a different approach. It seems to us that students begin much as biologists did at the turn of the century, asking general questions about the laws governing inheritance. Thus the first half of the book is a more or less historical treatment of classical eukaryotic genetics, modified somewhat by the need to organize the material under specific genetic concepts. Some molecular information is provided where necessary, but this aspect is not emphasized. Armed with classical principles, the students can proceed to the second half of the book, which integrates molecular techniques and information into the classical framework. Progression from the general perspective to the specific seems to be a natural one, and makes sense not only in research but in teaching about research.

Geneticists today can be divided into two broad groups: those concerned with the mechanisms of heredity and those using techniques of genetics to probe other fundamental biological processes or phenomena. Both of these approaches are emphasized and recur as themes throughout the text. However, as the title of this book suggests, the main theme is genetic analysis. This emphasis reflects our belief that the best way to understand

genetics is by understanding how genetic inferences are made. In any science, findings are important in themselves, but equally important are the modes of inference and the techniques of analysis, because these are the keys to future exploration. Quantitative analysis is particularly important. Many of the abstract ideas in genetics, from independent assortment to the existence of repetitive DNA, have been based on the analysis of quantitative data. The problems at the end of each chapter provide the student with the opportunity to apply these analytical methods to experimental situations.

Once again, this edition contains extensive revisions. These changes were made partly from our continuing attempts to update the contents and improve pedagogy in light of our own teaching experiences and partly from the suggestions of users and reviewes. In recent years, the balance between classical and molecular genetics in research has changed considerably; we have therefore given molecular genetics more emphasis in this edition than in previous ones, in recognition of its increasing application in all areas of the life sciences.

New features have been added and old ones improved. One prominent innovation is the list of key concepts found on the title page of each chapter. These are brief summary statements of the major themes to be found in the chapter, phrased to avoid the use of terms that have not yet been defined. They are intended to act as pedagogical signposts for the material that follows. A second major new feature is the addition of solved problems at the end of each chapter. Though these are often, but not always, representative of the problem sets, they are meant to show how the principles of the chapter can be applied to subsequent analysis. Approximately sixty new problems have been added to the problem sets. Most of these are simpler problems inserted early in the sets, and are designed to give practice and build confidence. As before, the problems are arranged roughly in order of difficulty, with particularly challenging problems marked with asterisks. The selected problem solutions at the back of the book have been assembled and modified by Diane Lavett (SUNY College at Cortland), who is also author of the *Companion,* which contains the complete set of problem solutions and a number of helpful study hints.

Also prominent is the newly added color insert. Although genetics is a somewhat cerebral subject, and is properly taught this way, it is nevertheless true that the subject material of genetics — the organisms themselves — are very often colorful and interesting. Furthermore this richness is a major source of inspiration in research as well as in teaching. We have therefore added a selection of colorful specimens that illustrate a variety of the principles covered in the book. Of course, some topic areas lend themselves naturally to color illustration and these comprise the bulk of the illustrations.

All chapters have been revised, but some contain extensive changes, as follows. In the first chapter on linkage (Chapter 5), the section on mapping has been completely rewritten. In the second linkage chapter (Chapter 6) the section on tetrad analysis has likewise been completely rewritten. The chapter on bacterial and phage genetics (Chapter 10) has been relocated to its original position in the second edition, in order to better accommodate the ensuing discussion on the nature of the gene. The entire central part of the book (broadly, molecular genetics) has undergone extensive revisions. Noteworthy here are the major changes in Chapter 10 and the addition of a new section on DNA manipulation in eukaryotes to Chapter 15. Chapter 16 has been amended to include updated material on eukaryotic genetic control mechanisms. Models of recombination are described more clearly in Chapter 18 and new material concerning mechanisms of transposition is presented in Chapter 19. The material on developmental genetics (Chapter 21 in the Third Edition) has been extended and reorganized into two separate chapters (21 and 22) that reflect the exciting recent developments in research in that area. The two chapters on populations (23 and 24) have also been streamlined.

As before, several special features aid the student. Throughout the text, major conclusions are summarized as Messages. These provide convenient stopping points from which readers may orient themselves within each chapter and may also be used as a convenient way of reviewing the material. Each chapter ends with a concise summary. New terms are set in boldface type; boldface is also used for emphasis. Most boldface words are defined in the Glossary. Suggestions for Further Reading are also found at the end of the book.

Thanks are due to the following people at W. H. Freeman and Company for their considerable support throughout the preparation of this edition: Patrick Fitzgerald, acquisitions editor; Moira Lerner, development editor; Stephen Wagley, project editor; Mary George, copy editor; Mike Suh, designer; Bill Page, illustration coordinator; and Susan Stetzer, production coordinator.

We also extend our thanks to the following reviewers, whose insights and suggestions were most helpful in the revision process: Wyatt W. Anderson, University of Georgia; Anna Berkovitz, Purdue University; Bruce J. Cochrane, University of South Florida; Christopher A. Cullis, Case Western Reserve University; Jeffrey L. Doering, Loyola University of Chicago; James E. Haber, Brandeis University; Robert Holmgren, Northwestern University, Robert Ivarie, University of Georgia; Janet Kurjan, Columbia University; Diane K. Lavett, State University of New York, Cortland; Anthony J. Pelletier, University of Colorado, Boulder; Jeffrey Powell, Yale University; Mark F. Sanders, University of California, Davis; Katherine Spindler, University of

Georgia; Jill Tabor; Laurie Tompkins, Temple University; David A. West, Virginia Polytechnical Institute and State University; and John H. Williamson, Davidson College.

We hope this book will stimulate the reader to do some first-hand experimental genetics, whether as professional scientist, student, or amateur plant or animal breeder. Failing this, we hope to impart some lasting impression of the incisiveness, elegance, and power of genetic analysis.

David T. Suzuki
Anthony J. F. Griffiths
Jeffrey H. Miller
Richard C. Lewontin

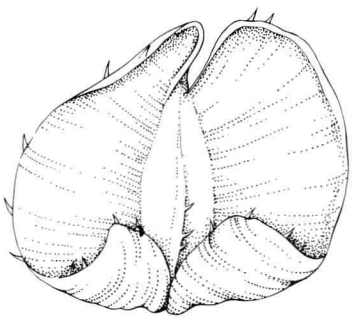
An Introduction to Genetic Analysis

Contents

CHAPTER

1

Genetics and the Organism



KEY CONCEPTS

Genetics has unified the biological sciences.

▪

Genetics may be defined as the study of genes through their variation.

▪

Gene variation contributes to variation in nature.

▪

The characteristics of an organism are determined by the interaction of its unique set of genes with its unique environment.

▪

Genetics is of direct relevance to human affairs.

■ Why study genetics? The answers to this question constitute the major part of this book, but at the outset a summary answer can be given. Although a relatively young discipline, genetics has assumed a position of central importance in the biological sciences, because a knowledge of the structure and function of the genetic material has been found to be essential to an understanding of most aspects of a living organism. In addition to this powerful unifying role, genetics has gained a position of great importance in human affairs. The findings of genetic research have had considerable impact not only in the applied areas of biology, medicine, and agriculture but also in such areas as philosophy, law, and religion. It is a rare newspaper issue nowadays that does not address some aspect of genetics.

The Scope of Genetics

Why has genetics become so important? To answer this question we must first define genetics. The science of **genetics** attempts to understand the properties of the genetic material, *deoxyribonucleic acid,* best known by its abbreviation DNA. Geneticists study the properties of DNA at many levels, ranging from cells to populations. The cells of all organisms, from bacteria to humans, contain one or more sets of a basic DNA complement that is unique to the species. This fundamental complement of DNA is called a **genome.** The genome may be subdivided into **chromosomes,** each of which is a very long single continuous DNA molecule. In its turn, a chromosome can be demarcated along its length into thousands of functional regions, called **genes,** and also into regions of less well understood function. For example, each of the trillions of cells that comprise a human being has 46 chromosomes in two equivalent sets, or genomes, of 23. Each of the 23 chromosomes in a genome is unique; it is matched only by its equivalent partner in the other set and by all other chromosomes of that type in the rest of the members of the species. An average human chromosome represents about a 50-millimeter length of DNA; therefore, the human genome is the equivalent of about one meter of DNA. This amount of DNA can only be packed into a nucleus by very efficient coiling and folding.

The functional units of heredity, the genes, have quite naturally become the focus of geneticists as they try to understand the laws of heredity. However, because genes are relevant to many different biological processes, many other kinds of biologists — from physiologists to ecologists — are also interested in genes, but these scientists would probably not consider themselves to be geneticists. Although it is true that many subdisciplines of biology are separated by "gray areas" today, it is nevertheless an interesting excercise to ask what, if anything, sets genetics apart. To this question, two kinds of answer can be given. First, geneticists tend to concentrate more on the basic properties of the genes themselves. Second, genetics works in a unique way that makes use of naturally occurring or induced gene variation in a population of organisms. In fact, variation is the raw material for genetic studies: if all members of a population were identical, genetic analysis as we know it could not be done. Therefore, it is useful to define genetics as *the study of genes through their variation.* Many examples of the way in which genes are studied through their variation will be found throughout this book, and the topic will be addressed again later in this chapter. The principle is, however, simple: only when a gene is different in some way can the geneticists follow its inheritance and its effects on the organism.

Geneticists study all aspects of genes. The study of the modes of gene transmission from generation to generation is broadly called **transmission genetics,** the study of gene structure and function is called **molecular genetics,** and the study of gene behavior in populations is called **population genetics.** These form the three major subdivisions of the field of genetics, although, as with all categories invented by humans, the subdivisions are to a certain extent arbitrary and there is considerable overlap. It is the knowledge of how genes act and how they are transmitted down through the generations that has unified biology; previously, specific sets of biological phenomena had each been relegated to separate disciplines. An understanding of how genes act is now an essential prerequisite for such biological fields of study as development, cytology, physiology, and morphology. An understanding of gene transmission is a fundamental aspect of areas such as ecology, evolution, and taxonomy. Further unification has resulted from the discovery that the basic chemistry of gene structure and function is very similar across the entire spectrum of life on the earth. These points may seem trite to those who have grown up in the light of current knowledge, but it is important to realize that our modern view of biology and its inter-related parts is a relatively recent phenomenon. Not so long ago, biology was fragmented into many camps that rarely communicated with each other. Today, however, every biologist must be a bit of a geneticist, because the findings and techniques of genetics are being applied and used in all fields. Take physiology, for example. Most aspects of the physiology of a cell, from photosynthesis to microtubule function, are strongly influenced by genes. In medicine, many human diseases, including many cancers, are influenced at the genetic level. Genetics, in fact, provides the modern paradigm for all of biology.

Message Genetics has provided a unifying thread for the previously disparate fields of biology.

What have been some of the success stories of genetics within basic biology? For just about any feature of biological structure or function—such as size, shape, number of parts, color, pattern, behavioral pattern, or biochemical function—that has been looked at in experimental organisms, genes have been found to be involved. Determining the location of these genes on their respective chromosomes is relatively easy. It has been found that the major way in which genes exert their effect is by controlling the myriad chemical reactions that go on inside cells, and thousands of specific genes have each been associated with a specific chemical reaction. The control of gene action (how genes are "turned on" and "turned off") has been intensely studied in many organisms, and the mechanisms in these cases are well understood. Many different genes have been isolated in the test tube, and their particular chemical structures have been determined. Such studies have provided important clues about how genes perform their functions. A gene can be removed from one organism and introduced into another, either for the convenience of propagating large amounts of the gene for later study or to examine its effects in another biological system. Genes can be modified at will to study the effects of these changes on biological processes; such changes can involve a large part of the genes or very localized regions of interest to the experimenter. Most genes have been found to reside rather stably at specific chromosomal locations, but other pieces of DNA have been found to be capable of sudden relocation to new areas. Last but not least, most of these findings have tremendous relevance for evolutionary processes, which of course are concerned with changes in the structure and function of the gene set.

Clearly, the advances in genetics have been truly astounding, particularly over the last three decades. Many recent accomplishments—such as the isolation and characterization of individual genes, which researchers in the 1950s believed could never happen in their lifetimes—have already come to be regarded as routine procedures in current work in genetics.

This chapter presents an overview of the subject of genetics, by way of an orientation to the rest of the book. We shall deal first with generalities about genes, their inheritance, and the ways in which they interact with the environment. Then we shall discover the unique ways in which geneticists identify specific genes and examine the use of these techniques in studying biological phenomena. Finally, we shall consider some of the ways in which genetics has interacted with human society.

Gene Transmission

Genetics embraces two contradictory aspects of nature: offspring resemble their parents, yet they are not identical to their parents. The offspring of lions are lions and never lambs, yet no two lions are identical, even if they come from the same litter. We have no trouble recognizing the differences between sisters, for example, and even "identical" twins are recognized as distinctive individuals by their parents and close friends. But we also can notice subtle similarities between parents and children. As we shall see, **heredity** (the similarity of offspring to parents) and **variation** (the difference between parents and offspring and between the offspring themselves) turn out to be two aspects of the same fundamental mechanism.

Humans necessarily became involved with both heredity and variation when they began to domesticate plants and animals (around 10,000 years ago), because humans had to choose the organisms with advantageous characteristics from among all the organisms at their disposal and then seek to propagate these traits in future generations. References in Egyptian tomb inscriptions and in the Bible to sound breeding practices convince us that conscious concern with genetic phenomena is at least as old as civilization. The farmers and shepherds involved with such concerns could quite deservedly have claimed to be called geneticists. But the formal study of genetics as a coherent and unified theory of heredity and variation is little more than a century old.

Modern genetics as a set of principles and analytic rules began with the work of an Augustinian monk, Gregor Mendel, who worked in a monastery in the middle of the nineteenth century in what is now Brno in Czechoslovakia. Mendel was taken into the monastery by its director, Abbot Knapp, with the express purpose of trying to discover a firm mathematical and physical foundation underlying the practice of plant breeding. Knapp and others in Brno were interested in fruit breeding. They believed that recent advances in mathematics in the physical sciences could establish a model for building a science of variation. Mendel was recommended to Knapp as a good scholar of mathematics and physics, although a rather mediocre student of biology!

Mendel's methods, which he developed in the monastery garden, are still used today (in an extended form) and form an integral part of genetic analysis. (Mendel's work is considered in detail in Chapter 2.) Mendel realized that both the similarities and the differences among parents and their offspring can be explained by a me-

chanical transmission of discrete hereditary units, which we now call genes, from parent to offspring. We now know that in all organisms — whether bacteria, fungi, animals, or plants — there is a regular passage of hereditary information from parent to offspring by means of the genes. The regularities we observe in heredity and variation are consequences of the regularities of the mechanical lanes of transmission and activity of these genes.

Recall that each gene is a portion of a DNA molecule. In more than one sense, then, DNA is truly the thread of life: not only is a DNA molecule itself a thread-like string of genes, but the DNA handed down from parent to offspring represents a narrow connecting thread between the generations. When we say that a woman has her mother's hair or a man has his father's nose, what we really mean is that the parent has handed on, in egg or sperm, the instructions necessary to direct the synthesis of that specific feature.

Out of these basic considerations emerge two vastly powerful and unique properties of DNA that make it the fundamental molecule of life. The first of these is its ability to serve as a model for the production of replicas of itself, termed **replication.** This property is the key to transmission and forms the basis of transmission genetics. A parental organism transmits a replica of its DNA to the progenitor cell of an individual of the next generation. As this progenitor cell goes through its rounds of division to produce a multicellular organism, each division is accompanied by the production of identical replicas of the DNA of the progenitor cell, which are apportioned into each new cell. Thus, replication is the mechanism through which life persists across the generations in a stable fashion.

The second property of DNA that makes it a fundamental molecule of life is its ability to act as a carrier of information. For example, embodied into the one meter of DNA that constitutes a human chromosome set is the information needed to build a specimen of *Homo sapiens.* The word information means literally "that which is necessary to give form"; this is precisely what the DNA of the genes does. The information is "written" into the sequence of DNA in the form of a molecular code.

Gene and Organism

Precisely how does information become form? At the level of molecules, the answer to this question embraces much of what was defined previously as molecular genetics. Basically, the phenomena and structures of life are produced by an interaction of DNA with the inanimate world, the nonliving environment. The universe naturally tends to disarray; order tends spontaneously to disorder; complex and orderly objects become piles of dust; the reverse does not occur unaided. Yet DNA causes an eddy in this river of chaos; through its interaction with the disorderly components of the universe, the most orderly system that we know about is born: the phenomenon of life. One of the unexpected discoveries arising from the study of DNA function is that the mechanism of converting information into form is virtually identical across all groups of organisms on this planet. We humans share a common genetic chemistry with the entire variety of life forms on the earth — a staggering spectrum, including some 286,000 species of flowering plants, 500,000 species of fungi, and 750,000 species of insects.

A general view of the interaction of DNA with the environment is a necessary prelude to the detailed analyses that are found in the chapters ahead. We must put the gene and the environment into perspective in order to provide a framework on which the details of genetic analysis can be hung.

It is a characteristic of living organisms that they mobilize the components of the world around themselves and convert these components into their own living material, or into artifacts that are extensions of themselves. An acorn becomes an oak tree, using in the process only water, oxygen, carbon dioxide, some inorganic materials from the soil, and light energy.

The seed of an oak tree develops into an oak, while the spore of a moss develops into a moss, although both are growing side by side in the same forest (Figure 1-1).

Figure 1-1. The genes of a moss direct environmental components to be shaped into a moss, whereas the genes of a tree cause a tree to be constructed from the same components. (From Grant Heilman.)

The two plants that result from these developmental processes resemble their parents and differ from each other, even though they have access to the same narrow range of inorganic materials from the environment. The specifications for building living protoplasm from the environmental materials are passed in the form of genes from parent to offspring through the physical materials of the fertilized egg. As a consequence of the information in the genes, the seed of the oak develops into an oak and the moss spore becomes a moss.

What is true for the oak and moss is also true within species. Consider plants of the species *Plectritis congesta,* the sea blush. Two forms of this species are found wherever the plants grow in nature: one form has wingless fruits, and the other has winged fruits (Figure 1-2). These plants will self-pollinate, and we can observe the offspring that result from such "selfs" when these are grown in a greenhouse under uniform conditions. It is commonly observed that the progeny of a winged-fruited plant are all winged-fruited and that the progeny from a wingless-fruited plant all have wingless fruits. Since all the progeny were grown in an identical environment, we can safely conclude that the difference between the original plants must result from the different genes they carry.

The *Plectritis* example involves two inherited forms that can both be considered perfectly normal. Yet the determinative power of genes is equally well demonstrated when a gene becomes abnormal. The human inherited disease sickle-cell anemia provides a good example. In this case, careful study has revealed the chain of events whereby the gene impacts on the organism from the submicroscopic molecular level, through the microscopic level, to the macroscopic anatomical level. The underlying cause of the disease is a variation in hemoglobin, the oxygen-transporting protein molecule found in red blood cells. Normal people have a type of hemoglobin called hemoglobin A, the information for which is encoded in a gene. A minute chemical change at the molecular level in the DNA of this gene results in the production of a slightly changed hemoglobin, termed hemoglobin S. In people possessing only hemoglobin S, the ultimate effect of this small change is severe ill health and usually death. The gene works its effect on the organism through a complex "cascade effect," as summarized in Figure 1-3.

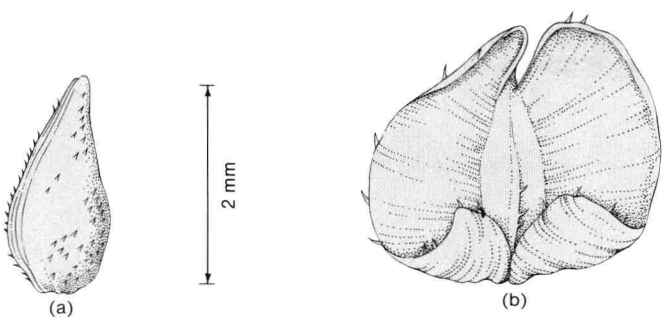


Figure 1-2. The fruits of two different forms of *Plectritis congesta,* the sea blush. (a) Wingless fruits. (b) Winged fruits. Any one plant has either all wingless or all winged fruits. In every other way the plants are identical. The striking difference in the appearance of the fruits is determined by a simple genetic difference.

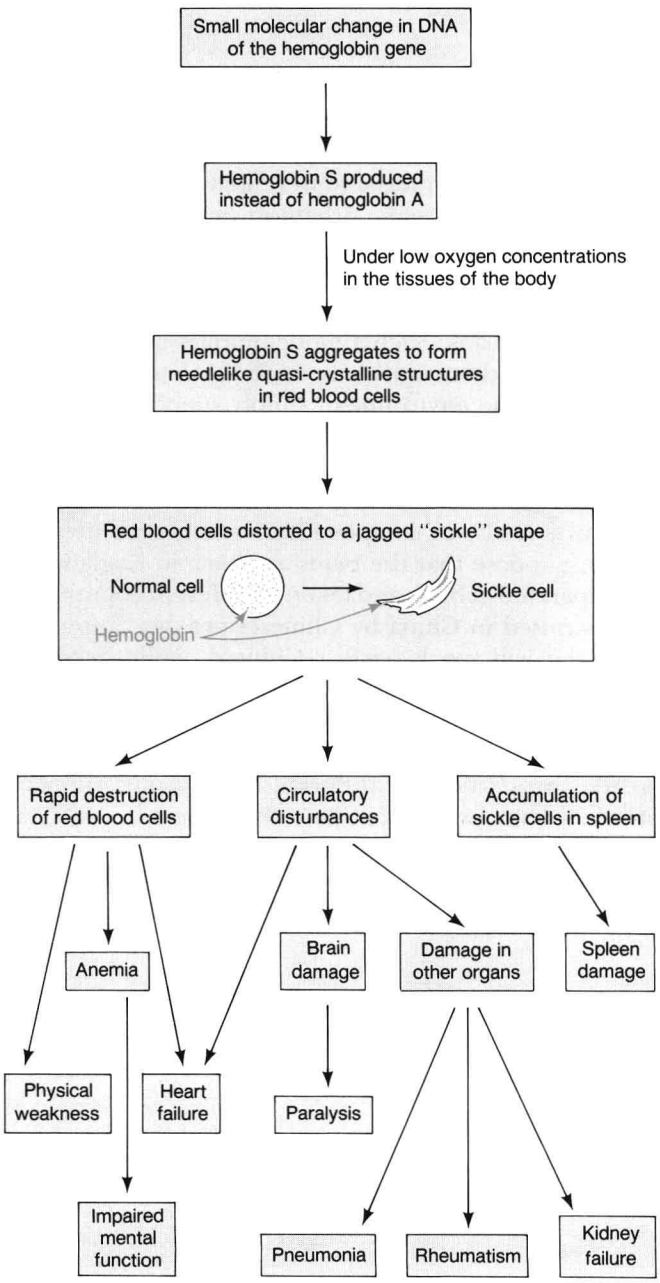


Figure 1-3. Chain of events resulting in sickle-cell anemia in humans.

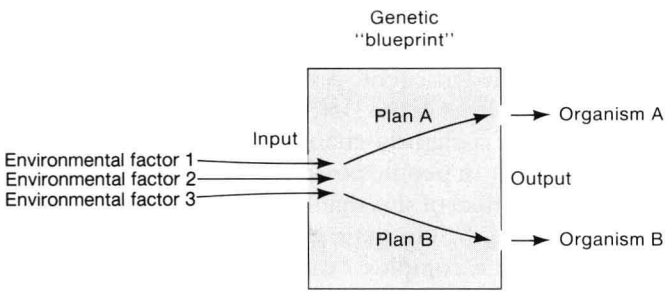


Figure 1-4. A model of determination that emphasizes the role of genes.

Observations like these lead to a model of the interaction of genes and environment like that shown in Figure 1-4. In this view, the genes act as a set of instructions for turning more or less undifferentiated environmental materials into a specific organism, much as blueprints specify what form of house is to be built from basic materials. The same bricks, mortar, wood, and nails can be made into an A-frame or a flat-roofed house, according to different plans. Such a model implies that the genes are really the dominant elements in the determination of organisms; the environment simply supplies the undifferentiated raw materials.

But now consider two monozygotic ("identical") twins, the product of a single fertilized egg that divided and produced two complete individuals with identical genes. Suppose that the twins are born in England but are separated at birth and taken to different countries. If one is raised in China by Chinese-speaking foster parents, she will speak perfect Chinese, while her sister raised in Budapest will speak fluent Hungarian. Each will absorb the cultural values and customs of her environment. Although the twins begin life with identical genetic properties, the different cultural environments in which they live will produce differences between the sisters (and differences from their parents). Obviously, the difference in this case is due to the environment, and genetic effects are of little importance.

This example suggests the model of Figure 1-5, which is the opposite of that shown in Figure 1-4. In the model in Figure 1-5, the genes impinge on the system, giving certain general signals for development, but the environment determines the actual course of change. Imagine a set of specifications for a house that simply calls for "a floor that will support 30 pounds per square foot" or "walls with an insulation factor of 15"; the actual appearance and nature of the structure would be determined by the available building materials.

Our different types of examples — of purely genetic effect versus that of the environment — lead to two very different models. Given a pair of seeds and a uniform growth environment, we would be unable to predict future growth patterns solely from a knowledge of the environment. In any environment we can imagine, if growth occurs at all, the acorn will become an oak and the spore will become a moss. On the other hand, considering the twins, no information about the set of genes they inherit could possibly enable us to predict their ultimate languages and cultures. Two individuals that are *genetically different* may develop differently in the *same environment,* but two *genetically identical* individuals may develop differently in *different environments.*

In general, of course, we deal with organisms that differ in both genes and environment. If we wish to understand and predict the outcome of the development of a living organism, we must first know the genetic constitution that it inherits from its parents. Then we must know the *historical sequence* of environments to which the developing organism is exposed. We emphasize the historical sequence of environments rather than simply the general environment. Every organism has a developmental history from birth to death. What an organism will become in the next moment depends critically both on the environment it encounters during that moment and on its present state. It makes a difference to an organism not only what environments it encounters but in what sequence it encounters them. A fruit fly *(Drosophila)* develops normally at 20°C. If the temperature is briefly raised to 37°C early in the pupal stage of

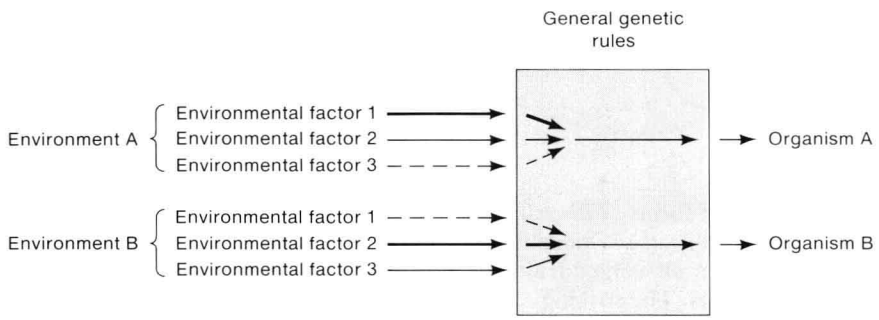


Figure 1-5. A model of determination that emphasizes the role of the environment.