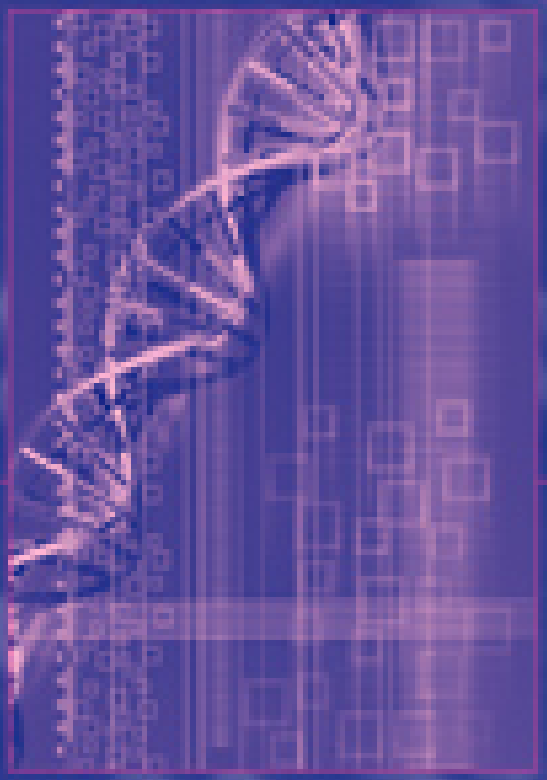


JAMES N. THOMPSON, JR.
JENNA J. HELLACK
GERALD BRAVER
DAVID S. DURICA

THIRD EDITION

Primer of GENETIC ANALYSIS

A Problems Approach

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521842686

PRIMER OF GENETIC ANALYSIS, THIRD EDITION

An invaluable student-tested study aid, this primer provides guided instruction for the analysis and interpretation of genetic principles and practice in problem solving. Each section is introduced with a summary of useful hints for problem solving and an overview of the topic with key terms. A series of problems, generally progressing from simple to more complex, then allows students to test their understanding of the material. Each question and answer pair is provided with a detailed explanation.

This new edition includes additional problems in basic areas that often challenge students, extended coverage in molecular biology and development, an expanded glossary of terms, and updated historical landmarks.

Students at all levels, from beginning biologists and premedical students to graduates seeking a review of basic genetics, will find this book to be a valuable aid. It will complement the formal presentation in any genetics textbook or can stand alone as a self-paced review manual.

James N. Thompson, Jr., is a David Ross Boyd Professor in the Department of Zoology at the University of Oklahoma. He has taught genetics for almost thirty years. His research interests are in genetic responses and adaptation to stress, genotype and environment interactions in quantitative traits, and variation in mutation rate.

Jenna J. Hellack is Professor and Chairperson of the Department of Biology at the University of Central Oklahoma. She currently teaches introductory genetics, evolution, and molecular and population genetics courses. Her research interests are in the area of population genetics.

Gerald Braver was Professor of Zoology at the University of Oklahoma from 1958 until his retirement in 1985. In addition to teaching genetics, he served as Assistant Dean of the College of Arts and Sciences and was a member of the steering committee that originally designed the university's Honors Program.

David S. Durica has been a member of the Department of Zoology at the University of Oklahoma since 1988, where he teaches both undergraduate and graduate courses in genetics. His research focuses on developmental genetics and the organization and expression of multigene families.

Primer of Genetic Analysis

A Problems Approach

Third Edition

JAMES N. THOMPSON, JR.
University of Oklahoma

JENNA J. HELLACK
University of Central Oklahoma

GERALD BRAVER
Formerly of the University of Oklahoma

DAVID S. DURICA
University of Oklahoma

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK
Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521842686

First published in print format 2007

ISBN-13 978-0-511-35569-1 eBook (NetLibrary)
ISBN-10 0-511-35569-6 eBook (NetLibrary)

ISBN-13 978-0-521-84268-6 hardback
ISBN-10 0-521-84268-9 hardback

ISBN-13 978-0-521-60365-2 paperback
ISBN-10 0-521-60365-X paperback

Contents

Preface

Many beginning students find some aspects of genetics much more difficult to grasp than others. Not very surprisingly, this is especially true of the problem solving and calculations involved in Mendelian genetics, pedigree analysis, linkage and mapping, quantitative analysis, and population genetics. We have found that texts often lack the space to provide much feedback to students regarding problem sets. Even if the correct answer is given, a student who arrives at a wrong one may not know whether the incorrect answer resulted from a minor mathematical error or a more fundamental misunderstanding of the material. This primer of genetic analysis is designed to address some of these difficulties.

In order to help you build confidence in your grasp of the subject and in your ability to apply that knowledge, we start with an array of problems that test basic ideas. More complex problems then follow naturally. Most importantly, we have supplied all questions with detailed explanations of how to work them. You can therefore work at your own pace through the logical analysis. With this practice and some guidance on how to avoid common pitfalls, you will gain experience and confidence in your ability to solve problems in the major areas of genetics.

Each chapter begins with Study Hints, in which we suggest ways of organizing your approach to a topic or take you step-by-step through a problem-solving exercise. Some areas, such as linkage and mapping, are given special attention, since they are particularly well suited to practicing data analysis and interpretation. It is important to remember, however, that these Study Hints are not comprehensive. They are intended to complement the material in a general textbook. The hints and problems should also be useful in reviewing for examinations like the Medical Candidacy Aptitude Test (MCAT) or Graduate Record Examination (GRE). We strongly recommend that you attempt the problems before looking at the answers. It is one thing to agree that the answers we provide make sense, but quite another to arrive at them on your own.

In preparing this and the earlier editions, we have benefited from the helpful suggestions made by students in our courses. In addition, Fred B. Schnee, William E. Spivey, and Robert R. Tucker made many constructive recommendations during the writing of the first edition. Eric M. Weaver, Stanton B. Gray, and K. April Sholl helped us see the material again through the eyes of students for the second edition. Their suggestions led to the addition of several "overview" chapters that we have now updated. Coral McCallister Cashion, Clayton N. Hallman, D. Jeremy Madrid, and Christopher H. May also contributed significant suggestions and perspectives for this third edition.

Sadly our colleague Gerald Braver died before the third edition was completed, and it is dedicated to his memory. While we learned from him the excitement of genetics, he taught each of us a great deal more. One of his hobbies was writing limericks. Our effort does not approach his wit or creativity, but we thought it would be a fitting tribute to end with this:

There once was a teacher named Braver
Whose DNA problems did favor
Those students who caught
The ideas that they taught
On exams this was sure a life saver.

Note on Genetic Symbols

It is very important to learn to recognize the meaning of genetic symbols and to use them consistently. Some symbols can be used interchangeably, and we have chosen to use several common versions in this text. Care in defining symbols will help prevent errors in interpreting and solving genetic problems.

The various forms in which a gene can exist (its alleles) are all given the same letter symbol. Uppercase letters (for example, *A* or *Cy*) denote the dominant alleles, and the lowercase letters (*a* or *cy*) denote the recessive alleles. When more than two alleles are known for a particular genetic locus, numerical superscripts are usually used to identify them. For example, a^1, a^2, and *A* could represent three different alleles at the *A* locus. Allele symbols, you will note, are italicized.

Another common convention is to use a plus (+) sign to denote the normal (or "wild-type") allele. If several wild-type alleles are being considered in the same problem or breeding plan, it is often useful to distinguish among them by adding the plus sign, as a superscript, to the mutant symbol. For example, *ri* would be the recessive mutant allele for the wild-type locus ri^+.

In this text, the complete genetic makeup of an individual may not necessarily be known. In those instances, a short dash will be used to indicate that an allele is present but unknown. For example, in the genotype *AaB* –, we know that the individual is heterozygous for the *A* locus and carries at least one *B* allele. We do not, however, know whether the individual is *Bb* or *BB*.

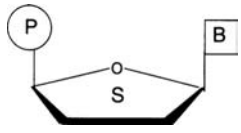
When the linkage relationships of a set of genes are important, the linked combinations will be separated by a slash. In the genotype *R Y U / r y u*, one chromosome carries the dominant alleles of all three genes, and the other chromosome carries the recessive alleles. Other special types of symbols should be clear from the context of the problem or will be defined for a specific situation.

CHAPTER ONE

Overview of Genetic Organization and Scale

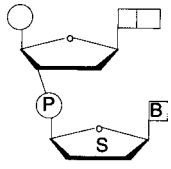
The genetic material is a molecule called **deoxyribonucleic acid (DNA).** Each **chromosome** contains a single long strand of DNA that encodes the information needed to produce hundreds or even thousands of different proteins. Each species has a characteristic array of chromosomes that carries all the **genes** needed to produce that organism from a single cell. The relationship between the genetic makeup of an organism (the **genotype**) and the developmental effects of these genes (the **phenotype**) can be complex. It is, therefore, useful to begin with a simple overview of these processes. Here we introduce some of the key concepts of genetics using an illustrated guide that begins at the smallest unit of genetic organization within a **nucleus** and ends at the level of the **population.** Some important terms are shown in boldface type, and definitions are given in the Glossary.

- DNA is made up of subunits called **nucleotides** composed of a sugar (S), a phosphate group (P), and a nitrogenous base (B). There are four nucleotides that differ by the nucleotide base they contain: **adenine (A), guanine (G), thymine** (T), and **cytosine (C)**. Genetic information is encoded in DNA by the sequence of these four bases.



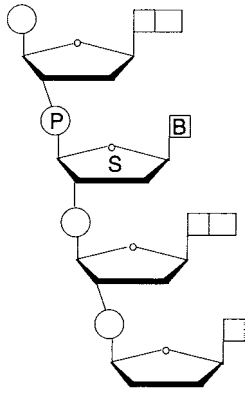
nucleotide

- **Nucleotides** are linked by a bond between the sugar of one nucleotide and the phosphate group of the next.



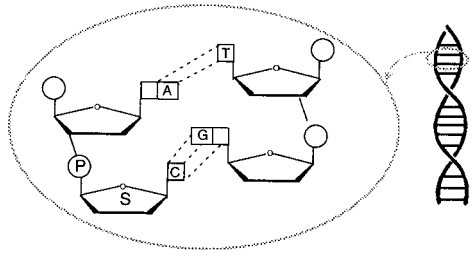
two nucleotides (schematic representation)

- This produces a long chain that can be literally millions of nucleotides long.



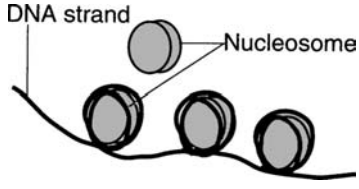
a portion of one strand

- **A single DNA molecule** is composed of two such strands that join together by bonds between the nucleotide bases (A paired with T, and C paired with G). This forms a DNA **double helix.**



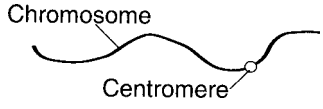
schematic of DNA and double helix

- The DNA is bound to structural proteins (**histones** that make up the **nucleosome**) that help pack the DNA in the nucleus and to regulatory proteins that turn genes on and off during development.

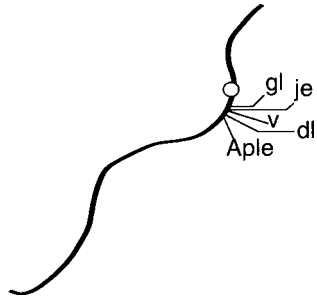


nucleosomes with DNA

- Each **chromosome** is made up of one long DNA molecule (one DNA double helix) and its associated proteins. The **centromere** is the attachment site for the spindle fiber that moves the chromosome during cell division.

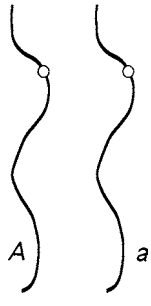


- Along the length of this DNA molecule are the regions that code for the production of proteins. These regions (or genetic **loci**) are the **genes.** Each gene can be up to a thousand or more nucleotides long, and every chromosome carries as many as several thousand different genes. All of the genes on a given type of chromosome are thus linked on a single DNA molecule. This linked group of genes is called a **linkage group.** The linear order of genes on a chromosome can be mapped to produce a linkage map.



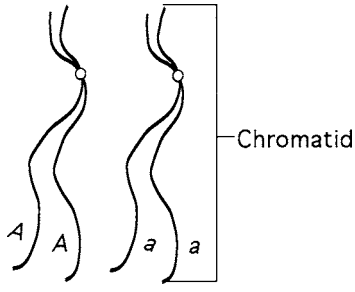
chromosome showing a map of five genes in linear order on its linkage group

- The body cells (**somatic** cells) of most organisms contain two copies of each type of chromosome (**diploid**). These are the homologous chromosomes. Since **homologous** chromosomes carry the same series of genes, they are members of the same linkage group. They can, however, differ in the form that a given gene takes (that is, normal or mutant). The different forms of a gene are called **alleles.**



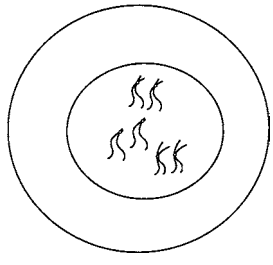
two homologous chromosomes carrying alleles *A* and *a*

- When the nucleus prepares to divide, each DNA molecule **replicates** except for the centromere. This yields two identical copies of the DNA molecule bound at the centromere. At this stage, the two copies are called sister **chromatids.** Since they have not yet divided at the centromere, however, each unit is still considered a single chromosome.



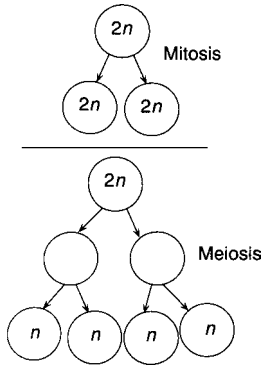
two homologous chromosomes, each with two sister chromatids

- Chromosomes that carry different sets of genes are called **nonhomologous** chromosomes. Every species has its own characteristic number of different chromosomes (***n***). **The total number of chromosomes in a somatic cell is, therefore, 2*n*.** All the genes needed to produce that organism will be found somewhere on one of these ***n*** linkage groups. The total **2*n*** genetic makeup is the **genome.** In the figure of the hypothetical cell, there are three pairs of nonhomologous chromosomes in the genome.



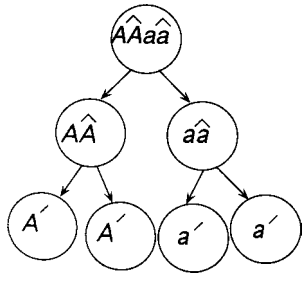
hypothetical cell with $2n = 6$, as seen during nuclear division

- **Mitosis** is a type of nuclear division that yields two identical diploid cells ($2n$). **Meiosis** is a special type of nuclear division found in reproductive **(germinal)** tissue that yields **gametes.** Each gamete carries only one copy of each linkage group and has a haploid (n) number of chromosomes. The diploid chromosome number is re-created at fertilization when the haploid maternal set and the haploid paternal set fuse.



mitosis and meiosis

- If we focus our attention on one gene, the alleles on the two homologous chromosomes can either be the same (*AA* or *aa* = **homozygous** genotypes) or be different (*Aa* = **heterozygous** genotype). These separate **(segregate)** during meiosis to produce haploid gametes.



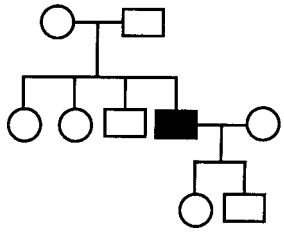
branching diagram to show haploid products

- The products of segregation and fertilization are highly predictable, giving rise to the basic rules of **genetic transmission.** Gregor **Mendel** set the foundation for this area of genetics.

♀ \ ♂	A	a
A	AA	Aa
a	Aa	aa

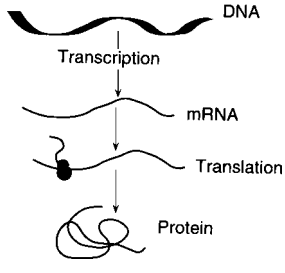
Mendelian cross using Punnett square for two heterozygous parents

- A **pedigree** diagram shows genetic relationships from a series of different Mendelian crosses. Circles indicate females and squares indicate males.



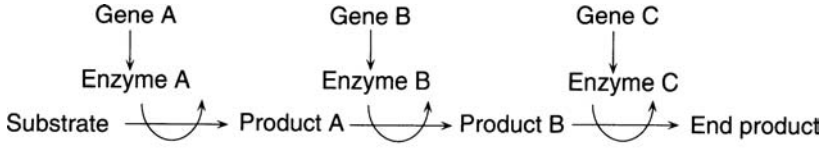
a simple pedigree

- Most genes code for the production of proteins. One of the two strands (the template strand) of a DNA molecule is "read" (through **transcription**) to yield a molecule of messenger **RNA** (mRNA). This then binds with ribosomes, where it defines the sequence of **amino acids** needed to produce the correct **polypeptide** (protein). This is **translation.**



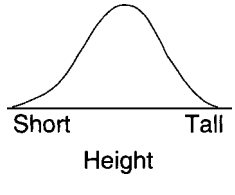
DNA → mRNA → protein

- Many proteins are **enzymes**, which catalyze specific biochemical steps. Thus, genes work by controlling the biochemical activities for growth and function of cells. In this way, the **genome** codes for all of the morphological, physiological, and behavioral characteristics (**phenotypes**) of an animal or plant.

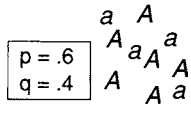


biochemical pathway

- Some characteristics are the result of several genes and environmental factors working together. Their expression is measured on an appropriate scale (such as height in meters). These are **quantitative traits (multifactorial** or **polygenic traits).**



- The genetic makeup of an individual is the **genome,** whereas the total genetic makeup of all individuals in the **population** is the **gene pool.** The gene pool is described in terms of allele frequencies, where ***p*** is the frequency of the *A* allele and ***q*** is the frequency of the *a* allele. By using appropriate assumptions, the genetic makeup of individuals in the population can be predicted.



alleles in a hypothetical gene pool

- **Hardy, Weinberg,** and **Castle** established the foundation for **population genetics** by showing that allele frequencies remain in **equilibrium** unless acted upon by **selection, migration, mutation,** sampling error in **small populations**, or deviations from **random mating.** Population genetics is the study of changes in allele and genotype frequencies that occur when these factors act on animals or plants.

	A (.6)	*a* (.4)
A (.6)	*AA* (.36)	*Aa* (.24)
a (.4)	*Aa* (.24)	*aa* (.16)

Punnett square with allele frequencies

Genetics is a dynamic and exciting field (but, of course, you would expect us to say something like that). But it can also be confusing, since there are so many levels at which you can look at inheritance and the use of genetic information. This introduction to genetic organization and scale is intended as a kind of outline to some key levels and processes. Although they overlap, we can readily see three main perspectives. First is the molecular level of DNA structure and coding (nucleotide—DNA—transcription—translation—protein product). Second are the rules of genetic transmission, which are based on probabilities of inheritance (segregation and independent assortment in meiosis—probabilities and genotypes—genes in families—genes in populations). Third is the way the genotype controls biochemical activities during development to produce the organism's phenotype (proteins—enzyme control of biochemical pathways—gene interactions—development). This primer will investigate these areas of genetics individually. But it is always important to keep in mind that they are really just different ways of looking at the same thing: the coding, transmission, and use of information by cells.

CHAPTER TWO

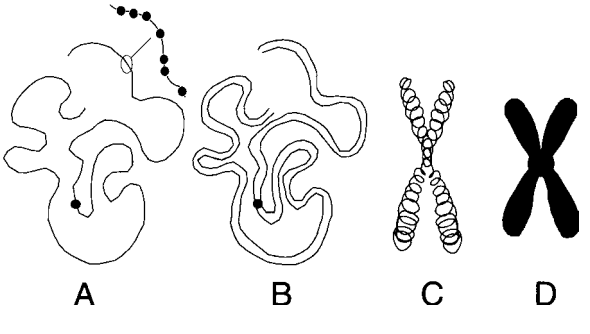
Mitosis and Meiosis

STUDY HINTS

Genes are located on chromosomes, and the stable manner in which chromosomes are first replicated and then distributed to daughter cells during cell division is the basis for genetic inheritance. Since much of genetic theory is based on the behavior of chromosomes and the genes they carry, it is very important to understand clearly how nuclear division occurs. In this way you can predict its consequences and understand the effect of errors that might occur in it. Yet the subject of cell division is complex, with many new terms to memorize and numerous things happening simultaneously. It is a continuous process that has been divided into stages somewhat artificially, so that we can describe it conveniently. All of this makes it rather hard to grasp at the beginning. Do not despair! It is really much simpler than it looks at first. The secret is to learn in stages. First one must understand the "strategies" of mitosis and meiosis, and the differences between them.

Mitosis has evolved as a mechanism to distribute accurately a copy of each chromosome present in the original cell to two new cells. The "goal" of *meiosis* is quite different. Meiosis passes alternate (homologous) copies of each type of chromosome to daughter cells and reduces the total chromosome number by half. These different objectives require slightly different chromosome behaviors. We shall briefly summarize these two processes, keeping in mind the different strategies they represent.

Both processes begin in essentially the same way. The chromosome (and the deoxyribonucleic acid [DNA] molecule it contains) duplicates, forming two identical chromosome strands attached to each other at the centromere. This is accompanied by a physical reorganization (coiling) that greatly reduces the chromosome's apparent length. The transition between these levels is illustrated in the following figure. One of the earliest signs of coiling is the formation of chromomeres, shown in this "magnified" insert to part A in the figure.



During late interphase, nuclear DNA and the histone proteins in chromosomes (A) duplicate (B), though remaining attached at the centromere. The centromere is really a constriction, but we draw it here as a dark spot so it can be seen easily. The two identical DNA copies are called sister chromatids. During prophase (C and D), these DNA strands coil into discretely identifiable chromosomes.

Each of the identical chromosome strands coils to form one of the two strands (sister chromatids) in a duplicated chromosome. The shape of the chromosome is determined by the position of the centromere. Without such coiling (or condensation), separating chromosomes would be a little like trying to separate a plate of spaghetti into two piles without breaking anything. People sometimes get confused about chromosome number at this point. Just remember that chromosome means "colored body." It is a single structural unit, no matter how much DNA it holds. So when counting chromosomes, count the centromeres, since whatever is attached to a centromere is a chromosomal unit.

In meiosis, the objective is to reduce the chromosome number so that there is only one copy of each kind in a gamete. The most direct way to do this is to pair the chromosomes carrying the same type of information or the same linear array of genes (homologous chromosomes). This is one purpose of synapsis. The first meiotic division, therefore, simply separates (segregates) homologous copies of each chromosome. The second meiotic division, like mitosis, distributes the identical copies (sister chromatids) that originated from chromosome replication.