

GENETICS
AND PROBABILITY
IN ANIMAL BREEDING
EXPERIMENTS

GENETICS AND PROBABILITY IN ANIMAL BREEDING EXPERIMENTS

A primer and reference book on probability, segregation, assortment, linkage and mating systems for biomedical scientists who breed and use genetically defined laboratory animals for research

EARL L. GREEN

M

© Earl L. Green 1981

Softcover reprint of the hardcover 1st edition 1981

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission

First published 1981 by

The Scientific and Medical Division

MACMILLAN PUBLISHERS LTD

London and Basingstoke

Associated companies in Delhi Dublin

Hong Kong Johannesburg Lagos Melbourne

New York Singapore and Tokyo

Typeset by

Reproduction Drawings Ltd, Sutton, Surrey

ISBN 978-1-349-04906-6 ISBN 978-1-349-04904-2 (eBook)

DOI 10.1007/978-1-349-04904-2

This book is sold subject to the standard conditions of
the Net Book Agreement

Dedicated to

THE JACKSON LABORATORY,
founded in 1929;
to its founder,
a man of vision;
and to its people,
past, present, and future;
on the occasion of its fiftieth anniversary in 1979

Preface

For nearly 20 years, I was Director of the Jackson Laboratory in Bar Harbor, Maine, one of the world's centres for research in mammalian genetics. During most of those years, I gave a series of lectures on genetics and mouse breeding to each year's new crop of staff members and postdoctoral fellows. This book is an expanded version of those lectures. It is intended to be useful to research workers who are just starting to work with genetically defined animals, such as mice of various inbred strains, and who need an introduction to the probability aspects of transmission genetics in order to enrich their understanding of the kinds of animals already available and to guide them in the design and analysis of their own breeding experiments.

The new staff members and postdoctoral fellows I dealt with had many different formal educational backgrounds. They included physiologists, biochemists, embryologists, psychologists, pathologists, immunologists, microbiologists and physical chemists. The yearly group also included geneticists, but contemporary geneticists are of many kinds: cytogeneticists, behavioural geneticists, biochemical geneticists, physiological geneticists, immunogeneticists, developmental geneticists, radiation geneticists and quantitative geneticists. All of these kinds of scientists usually have a working knowledge of the classical principles of heredity and of the classical methods of statistical inference. Yet, except for the quantitative geneticists, they usually feel uneasy with the probability aspects of transmission genetics and with the methods of statistical inference for discrete variables.

Scientists of this sort will soon encounter the folkways and jargon of the geneticists who breed laboratory animals. They will want to know, and will need to know: how new genes are discovered, how one establishes that a new gene is not like any already known gene, how one finds the position of a new genetic locus on the chromosome map, how alike are the animals of an inbred strain, how many generations of breeding does one have to arrange to ensure a given probability of likeness, how many matings should one make up and how many progeny per mating should one observe in each generation when perpetuating a recessive mutation? Answers to questions of this sort will be found in the following pages.

This book is intended, in two ways, to meet the needs of research workers who use laboratory animals: first, as a coherent self-contained account of the probability

and statistical aspects of laboratory animal breeding and, second, as a ready reference book for terminology and formulae. The organisation of the book reflects these two objectives. To provide an account, the reader is conducted through five model breeding experiments of the sort that he, himself, may soon want to carry out in his own research. To provide a ready reference, the more useful formulae are assembled in tables or, otherwise, are identified by serial numbers within chapters.

I am assuming that everyone who reads this book knows that the units of heredity are called *genes* and that genes are located on *chromosomes*. I am also assuming that everyone knows about the elementary principles of transmission genetics: *segregation of alleles* and *assortment* or *recombination of non-alleles*; and that the genetic phenomena of segregation and assortment exactly parallel the chromosomal phenomena of *disjunction*, *assortment* and *crossing-over* during germ-cell formation. The book deals with various modifications and complexities of these elementary principles of transmission genetics. It does not deal with the molecular nature of the gene, nor with how genes replicate or how they make proteins, for these aspects of physical, physiological, developmental and biochemical genetics are not relevant to its purpose.

I am further assuming that the reader's background in the probability aspects of genetics is skimpy at best, even though he or she is familiar with the 1:1 and 1:1:1:1 segregation ratios of backcrosses and the 3:1 and 9:3:3:1 segregation ratios of intercrosses for one and two loci. Such readers should find many familiar concepts of a first course in genetics recast in a more sophisticated symbolism of probability and thereby put into forms that lead to symbolic predictions and bases for statistical inferences about the outcomes of breeding procedures. Even so, the reader's level of mathematical competence need not exceed algebra. All manipulations requiring calculus have been relegated to appendices.

The book is intended to be read, to be studied and to be used as a reference. Yet it is not cast in the form of a textbook. I have tried to compensate for this deliberate omission by including numerous worked examples in the text. Beyond that I suggest that for such exercises the interested reader should procure elementary textbooks on probability and statistics and on experimental genetics, of which there are dozens of good ones now available.

As a convenience in exposition and as a reflection of my own first-hand familiarity, I have referred to *mice* throughout the text. Yet the concepts and methods described in the book are applicable to any bisexual organism whose reproductive pattern is similar to that of mice. This includes all the common laboratory mammals: rats, guinea pigs, hamsters, rabbits, gerbils, nutria, cats and dogs.

This book deals with the overlap of the domain of genetics and the domain of probability. To pursue that figure of speech just a bit further: each domain has some lofty peaks and some profound caverns. To ascend the peaks or to explore the caverns would require special hand tools and foot gear. Our explorations will, instead, be confined to the gentle paths among the foothills between the domains, with short excursions into a few nooks and crannies. Soft-soled shoes and hand-held calculators are all we shall need.

The first chapter is composed of extremely condensed statements about the

concepts of probability needed for genetic predictions and the methods of statistics needed in the analysis of genetic data. My emphasis is on the meaning and use of the numerous concepts and methods needed by the laboratory animal breeder. Biologists who feel repelled by statistical notation may prefer to skim this chapter at first reading and refer to it only when a particular idea in a later chapter depends upon a formula in chapter 1.

Chapters 2 and 3 are devoted to deriving Mendel's principles of heredity from the data of a few mouse breeding experiments, but with emphasis on full probability statements of these principles rather than upon the ratios of more elementary treatments.

Chapter 4, on linkage, is like chapters 2 and 3, on segregation and assortment, in putting the concepts in probability form. Chapter 5 does the same thing for various regular systems of mating in widespread use, but also uses the probability results to predict the genetic consequences of using the systems over many generations.

The appendices are essential for the serious breeder of laboratory animals, both for extensions of the theoretical aspects and the practical problems he will face. Appendices 1-7 contain technical details that I thought should be removed from the text because they require more mathematical preparation than I am assuming for most of the users of this book. They are, however, an invitation to the reader to delve a little deeper into the fascinating topics of linkage estimation and of analysis of mating systems. Appendix 7 deals with the practical problems of how many animals to raise and how many matings to make up for various mating systems described in chapter 5. Appendices 8, 9 and 10 deal with the intensely practical problems of nomenclature of genes and strains of mice, of a record-keeping system for breeding laboratory animals, and of mouserroom layout and operation.

The elaboration of the principles of transmission genetics is largely the work of Sewall Wright, R. A. Fisher and J. B. S. Haldane, along with those who followed in their wake. I have not given references to the origin of each probability and statistical statement in this elementary introduction. The formulae, or the fundamental relationships from which they are derived, can be found in the vast writings on genetics, probability and statistics in the middle two quarters of this century. I have, however, given references to various contemporary sources of general information about genetics, probability and statistics and to specific recent developments cited in the text.

I am indebted to several people for direct encouragement with respect to the preparation of this book. My thanks are especially due to my wife, Margaret C. Green, and to my former colleagues at the Jackson Laboratory, Donald W. Bailey and Robert L. Collins, for reading the text critically during its preparation. My thanks are also due to Randall C. Adams for carrying out the computations used for the figures in chapter 5.

Contents

<i>Preface</i>	vii
Chapter 1 Probability and statistics	1
1.1 Variables and probabilities	2
1.2 Empirical probabilities	3
1.3 Theoretical probabilities	4
1.4 Points, probabilities and events	4
1.5 Complements and unions	5
1.6 Union of mutually exclusive events	6
1.7 Union of intersecting events	8
1.8 Conditional probabilities	9
1.9 Terminology	12
1.10 Summary about probability	12
1.11 Binomial distribution	13
1.12 Products of binomial distributions	17
1.13 Multinomial distributions	18
1.14 Mean and variance of binomial and multinomial distributions	19
1.15 Areas of the normal distribution	21
1.16 Maximum likelihood estimates of parameters	23
1.17 Amount of information	24
1.18 Terminology	24
1.19 Estimating an unknown probability	25
1.20 Test of a hypothesis about a probability	27
1.21 Errors in testing statistical hypotheses	29
1.22 Test of a difference between two proportions	30
1.23 Comparing class frequencies	32
1.24 Partitioning deviations	36
1.25 Orthogonal functions	37
1.26 Matrix algebra	39
1.27 Sources	39

Chapter 2	Segregation of alleles	41
2.1	Experiment 1: Matings of pale eared and normal eared mice	42
	Proportions and probabilities	44
	Inferences, symbols and assumptions	45
2.2	Definitions	47
2.3	Modified and disturbed segregations	50
2.4	Estimation of the differential viability parameter	54
2.5	Mating types	58
2.6	Nomenclature	59
2.7	Sources	59
Chapter 3	Assortment of non-alleles	61
3.1	Experiment 2: Matings of mice with pale ears and dilute coats	61
	Symbols	65
	Gametic output	66
	Mating types	67
	Inference	70
3.2	Other two-locus phenotypic ratios	70
	Codominance	71
	Epistasis	72
3.3	Experiment 3: Matings of albino and himalayan mice	73
	Inferences	75
3.4	Concept of the gene	75
3.5	Sources	76
Chapter 4	Linkage, recombination and mapping	77
4.1	Experiment 4: Matings of mice with pale ears and ruby eyes	77
	Definitions and symbols	80
4.2	Gametic output	81
4.3	Kinds of mice and mating types	82
4.4	Kinds of progeny and their probabilities	86
4.5	Maximum likelihood estimates of the linkage parameter	91
4.6	Extensions and complications	98
4.7	Use of a recombination fraction as a map distance	99
4.8	Summary on segregation, assortment and linkage	101
4.9	Experiment 5: A three-point cross involving extra toes, muted and pearl	101
4.10	Mapping functions	106
4.11	Linkage map of the mouse	108
4.12	Linkage testing stocks	109
4.13	Sources	112

Chapter 5	Mating systems	114
5.1	Experimental designs	115
	Comparison of treatments	115
	Correlation between variables	116
	Effects of mutated genes	116
	Discovery of new genetic traits	117
	Genetic dissection of complex traits	117
	Use of random-bred mice	117
5.2	Kinds of mice	117
5.3	Symbols and definitions	118
5.4	Method of analysis	120
5.5	Inbred strains	124
	Brother-sister inbreeding (system 1)	126
	Illustration	129
	Ratios of successive probabilities	131
	Nomenclature	131
	F ₁ hybrids	131
5.6	Recombinant inbred strains	133
	Mating plan	134
	Linkage estimates	137
	Strain distribution pattern	138
	Nomenclature	141
5.7	Congenic and coisogenic inbred strains	141
	Backcross system (system 2)	142
	Cross-intercross system (system 3)	147
	Cross-backcross-intercross system (system 4)	149
	Nomenclature	151
5.8	Segregating inbred strains	152
	Brother-sister inbreeding with heterozygosity forced by backcrosses (system 5)	153
	Brother-sister inbreeding with heterozygosity forced by intercrosses (system 6)	155
	Brother-sister inbreeding with heterozygosity forced by crosses and intercrosses or by backcrosses and intercrosses (systems 7 and 8)	157
	Crosses and intercrosses (system 7)	157
	Backcrosses and intercrosses (system 8)	159
	Nomenclature	161
5.9	Comparison of systems; strategies of use	161
5.10	Two-locus systems	167
	To propagate lethal, sterile or deleterious recessive mutations	167
	To propagate handicapped, but breedable, recessive mutations	172
	To identify recessive homozygotes early	173
	To identify mice with 0, 1 or 2 doses of recessive mutations	174

	To transfer a recessive mutation to a standard inbred background	175
	To transfer a not-easily-recognised recessive mutation to a standard inbred background	176
5.11	Linkage estimation	177
5.12	Number of matings and number of mice per mating	181
5.13	Sources	184
Appendix 1	Mean and variance of a binomial distribution	186
Appendix 2	Estimation of a parameter by the method of maximum likelihood	188
A2.1	Example 1: Binomial probabilities	190
A2.2	Example 2: Multinomial distribution with one parameter	192
Appendix 3	Extensions of the method of maximum likelihood	194
A3.1	An alternative form for the amount of information	194
A3.2	Propagation of variance	196
A3.3	Amount of information per observation	197
A3.4	Use of scores	198
A3.5	Completely worked example	199
Appendix 4	Comparative efficiency of matings for detecting and measuring linkage	202
Appendix 5	The Fibonacci sequence	210
Appendix 6	Systems of mating	212
Appendix 7	Numbers of matings and numbers of mice per mating	217
A7.1	General method	217
A7.2	Number of progeny of one-locus testcrosses	218
A7.3	Kinds of one-locus testcrosses	219
A7.4	Number of one-locus testcrosses	220
A7.5	Number of progeny of two-locus testcrosses	221
A7.6	Number of two-locus testcrosses with recessive mutations	223
A7.7	Number of two-locus testcrosses with a recessive mutation and a dominant marker	224
Appendix 8	Nomenclature	226
A8.1	Rules for gene nomenclature in mice	226
	Names of gene loci	226
	Symbols for gene loci	226
	Loci in series	227
	Allele symbols	227

	Phenotype symbols	228
	Gene complexes	228
	Viruses	230
	Antigenic variants	230
A8.2	Guidelines for nomenclature of biochemical variants	231
	Biochemical nomenclature	231
	Symbols for structural loci	231
	Allele symbols	231
	Series of loci	232
	Phenotype symbols	232
	Identification of loci	232
	Genetic variants affecting enzyme activity of mice	232
A8.3	Rules for nomenclature of chromosome anomalies	233
	Symbols for chromosome anomalies	233
	Nomenclature for variations in heterochromatin and chromosome banding	234
	Use of human chromosome nomenclature	234
A8.4	Rules for nomenclature of inbred strains of mice	235
	Definition of inbred strain	235
	Symbols for inbred strains	235
	Indication of inbreeding	235
	Priority in strain symbols	235
	Recombinant inbred strains	235
	Substrains	235
	Designation of substrains	236
	Sublines	236
	Designation of sublines	236
	Coisogenic, congenic and segregating inbred strains	237
	Inbred strains of mice preserved by freezing	238
Appendix 9	Record-keeping	240
A9.1	Forms	241
	Cage tags	241
	Mating cards	243
	Litter cards	243
	Mating record sheet	245
	Summary cards	246
	Other cards and hardware	247
A9.2	Acts	249
	Accessioning	249
	Mating	249
	Checking	249
	Recording	250
	Summarising	251
A9.3	Alternative system	251

CONTENTS

xv

Appendix 10	Mouseroom layout and procedures	253
Glossary of signs and symbols		256
Literature cited		258
<i>Index</i>		261