

**THE MEASUREMENT
OF LINKAGE IN
HEREDITY**

K. MATHER

**METHUEN'S MONOGRAPHS
ON BIOLOGICAL SUBJECTS**

Methuen's Monographs on Biological Subjects

General Editors: G. R. de Beer and Michael Abercrombie

**THE MEASUREMENT OF LINKAGE
IN HEREDITY**

**THE
MEASUREMENT OF LINKAGE
IN HEREDITY**

by

K. MATHER, D.Sc., F.R.S.

PROFESSOR OF GENETICS

IN THE

UNIVERSITY OF BIRMINGHAM

LONDON: METHUEN & CO. LTD.

NEW YORK: JOHN WILEY & SONS, INC.

First published . . . April 28th, 1938
Second edition, revised and enlarged 1951

CATALOGUE NO. 4104/U (METHUEN)

PRINTED IN GREAT BRITAIN

PREFACE

GREGOR MENDEL did not use probability methods in the analysis of his classical experiments, but after the rediscovery of his work, at the beginning of this century, the application of such a technique was soon recognized as a necessary part of genetical analysis. The problems at first were chiefly involved with the testing of the significance of departures from expected ratios, but, with the discovery of linkage, questions of estimation became equally important. The absence of a suitable statistical technique is one of the prime causes of the failure to understand linkage and linkage groups during the years before the *Drosophila* technique of backcrossing was extensively practised, and the true relations of coupling and repulsion realized. Since that time statistical methods have become more and more favoured by genetical workers and are to-day more than ever necessary for genetical analysis.

Now statistics, like genetics, is a growing science and the crude and often arbitrary methods of yesterday have been superseded and rendered redundant by the development of more exact and more adaptable techniques. This is largely a result of the work of R. A. Fisher and his associates. In some branches of biology, as for example agronomical experimentation, such re-

finer statistical methods are now in common use for several reasons, not the least important of which is the existence of articles and textbooks giving details of their application. In genetics these methods have been less fully described and applied. It is hoped that the present work will serve to bring to the notice of geneticists the desirability of employing such methods and provide the necessary instruction for their use.

It is not claimed that this book is complete. In fact some statistical devices such as the analysis of variance and the use of regressions have been entirely omitted as they are, at present, of secondary importance to the geneticist. On the other hand, the uses of χ^2 and of maximum likelihood have been dealt with in considerable detail as they are of wide application in this branch of science. The methods are described in general and some specific applications are discussed in detail. No attempt has been made to prove all the general formulae used, though some have been considered in detail. Such a course is beyond the scope of this work and would, in any case, merely result in confusion for a non-mathematical reader. If desired, proofs may be found in the original literature cited. But numerical examples have been used wherever possible in the hope that the reader, by working through them, will familiarize himself with the methods and be able to apply them to other problems. The list of general formulae in the last chapter should make reference to any method, and its use for other analyses, a simple matter. It cannot be overemphasized that in order to

make full use of the book the actual examples used should be worked through in detail, as it is only by this means that they will be fully understood and appreciated.

For further details of the genetical and cytological principles, which have been assumed here in order to permit fuller development of the statistics, the reader is referred to *Mendelism and Evolution*, by E. B. Ford and to *The Chromosomes*, by M. J. D. White, both of which are in this series of Biological Monographs.

I am indebted to Professor R. A. Fisher, Head of the Galton Laboratory, for permitting the reproduction herein of a number of tables mainly from the *Annals of Eugenics*, and for continued encouragement during the writing of this book, and to Messrs. W. J. C. Lawrence and J. C. Cullen, of the John Innes Horticultural Institution, for reading and criticizing the manuscript. I also wish to thank Messrs. Oliver and Boyd for allowing me to reproduce Tables I and II, from R. A. Fisher's *Statistical Methods for Research Workers*.

PREFACE TO THE SECOND EDITION

THE chief alterations made in preparing the second edition are the inclusion of a new section in Chapter VII on the scoring method devised by Professor R. A. Fisher for use in testing and estimating linkage, and a chapter on the estimation of gene frequencies, as is needed in the analysis of the distribution of, for example, blood groups in the various races of man. References to certain other developments of the past

viii MEASUREMENT OF LINKAGE IN HEREDITY

twelve years have also been added, and the occasion has been taken to correct a number of minor, though misleading, mistakes. I am indebted to several correspondents for drawing my attention to these errors.

CONTENTS

CHAP.	PAGE
PREFACE	v
I INTRODUCTION	1
II TWO CLASS SEGREGATIONS	13
III THE PLANNING OF EXPERIMENTS (I)	26
IV THE DETECTION OF LINKAGE	32
V THE ESTIMATION OF LINKAGE	44
VI INFORMATION AND THE PLANNING OF EXPERIMENTS (II)	56
VII COMBINED ESTIMATION AND TESTING HETEROGENEITY	69
VIII DISTURBED SEGREGATIONS	91
IX HUMAN GENETICS (I)	104
X HUMAN GENETICS (II)	114
XI THE ESTIMATION OF GENE FREQUENCIES	125
XII SYMBOLS AND FORMULAE	137
REFERENCES	145
INDEX	147

REFERENCES

- BAILEY, N. T. J. (1949.) The estimation of linkage with differential viability, II and III. *Heredity*, **3**, 220-8.
- BATESON, W. (1909.) *Mendel's Principles of Heredity*. University Press, Cambridge.
- FALCONER, D. S. (1947.) Sensory thresholds for solutions of phenyl-thio-carbamide. *Ann. Eugenics*, **13**, 211-22.
- FINNEY, D. J. (1942.) The detection of linkage. VI. The loss of information from incompleteness of parental records. *Ann. Eugenics*, **11**, 233-44.
- FISHER, R. A. (1921.) On the mathematical foundations of theoretical statistics. *Phil. Trans. Roy. Soc., A*, **122**, 309-68.
- (1928.) On a property connecting the χ^2 measure of discrepancy with the method of maximum likelihood. *Atti del Congresso Internazionale dei Matematici (Bologna)*, **6**, 95-100.
- (1934.) The effect of methods of ascertainment upon the estimation of frequencies. *Ann. Eugenics*, **6**, 13-25.
- (1935a.) The detection of linkage with 'Dominant' abnormalities. *Ann. Eugenics*, **6**, 187-201.
- (1935b.) The detection of linkage with 'Recessive' abnormalities. *Ann. Eugenics*, **6**, 339-51.
- (1936.) Tests of significance applied to Haldane's data on partial sex-linkage. *Ann. Eugenics*, **7**, 87-104.
- (1946.) A system of scoring linkage data, with special reference to the pied factors in mice. *Amer. Nat.*, **80**, 568-78.
- (1949a.) *Statistical Methods for Research Workers*. (11th edn.) Oliver and Boyd, Edinburgh.
- (1949b.) *The Design of Experiments*. (5th edn.) Oliver and Boyd, Edinburgh.
- and BALMAKUND, B. (1928.) The estimation of linkage from the offspring of selfed heterozygotes. *Jour. Genet.*, **20**, 79-92.
- and MATHER, K. (1936.) A linkage test with mice. *Ann. Eugenics*, **7**, 265-80.
- HALDANE, J. B. S. (1936.) A search for incomplete sex-linkage in man. *Ann. Eugenics*, **6**, 339-51.

146 MEASUREMENT OF LINKAGE IN HEREDITY

- HUTCHINSON, J. B. (1929.) The application of the 'Method of Maximum Likelihood' to the estimation of linkage. *Genetics*, 14, 519-37.
- IMAI, Y. (1931.) Linkage studies in *Pharbitis Nil* I. *Genetics*, 16, 26-41.
- IMMER, F. R. (1930.) Formulae and tables for calculating linkage intensities. *Genetics*, 15, 81-98.
- (1934.) Calculating linkage intensities from F₂ data. *Genetics*, 19, 119-36.
- JENKINS, M. T. (1927.) A factor for yellow-green chlorophyll colour in maize and its linkage relations. *Genetics*, 12, 498-518.
- MATHER, K. (1935.) The combination of data. *Ann. Eugenics*, 6, 399-410.
- (1936a.) Types of linkage data and their value. *Ann. Eugenics*, 7, 251-64.
- (1936b.) Segregation and linkage in autotetraploids. *Jour. Genet.*, 32, 287-314.
- (1937.) The analysis of single factor segregations. *Ann. Eugenics*, 8, 96-105.
- NABOURS, R. K., LARSON, I., and HARTWIG, N. (1933.) Inheritance of colour patterns in the grouse locust, *Acrydium arenosum*, Burmeister (Tettigidae). *Genetics*, 18, 159-72.
- PEARSON, K., NETTLESHIP, E., and USHER, C. H. (1913.) A monograph on albinism in man IV. *Drapers Company Research Memoirs, Biometrica*, IX.
- PENROSE, L. S. (1935.) The detection of autosomal linkage in data which consist of pairs of brothers and sisters of unspecified parentage. *Ann. Eugenics*, 6, 133-8.
- PHILP, J. (1934.) The genetics of *Papaver Rhoeas* and related forms. *Jour. Genet.*, 28, 175-204.
- RACE, R. R., SANGER, R., LAWLER, S., and KEETCH, D. V. (1948.) Blood groups of Latvians. *Ann. Eugenics*, 14, 134-8.
- SMITH, H. F. (1937.) A test of significance of Mendelian ratios when classification is uncertain. *Ann. Eugenics*, 8, 94-5.
- STEVENS, W. L. (1938.) Estimation of blood group gene frequencies. *Ann. Eugenics*, 8, 362-75.
- TAYLOR, G. L., and PRIOR, A. M. (1939.) The distribution of the M and N factors in random samples of different races. *Ann. Eugenics*, 9, 97-108.
- WINTON, D. DE, and HALDANE, J. B. S. (1935.) The genetics of *Primula sinensis*. III. Linkage in the diploid. *Jour. Genet.*, 31, 67-100.
- YATES, F. (1931.) Contingency tables involving small numbers and the χ^2 test. *Supp. Jour. Roy. Stat. Soc.*, 1, 217-35.

INDEX

- Achromatopsia, 117
Acridium arenosum, 97
 Albinism, 109
Antirrhinum majus, 14
 Ascertainment, 106, 117
 Autotetraploids, *see Simplex*
- Backcross, 3, 8, 43, 45, 79,
 116, 121
 — exact treatment, 96 *et seq.*
 — singles, 78, 116 *et seq.*
 Bailey, 96
 Balmakund, 54
 Bateson, 73, 87
 Binomial distribution, 8
 Blood groups, 125
 MN—, 127 *et seq.*
 ABO—, 131 *et seq.*
 Brandt, 22, 38, 128
- χ^2 , 10, 15 *et seq.*, 51, 53, 77,
 90, 93, 105, 111, 127, 129
 et seq.
 Classification, 63
 Comparisons, independent,
 34, 38
 Complementary factors, 3,
 64, 92
 Consistency, criterion of, 44
 Contingency table, 92, 122
 Continuity, correction for,
 123
 Continuous distribution, 10
 Coupling, 4, 61, 96, 117
 Covariance of estimates, 135
- Degrees of freedom, 16, 33
 — loss of, 22, 54
- Deviation, 69, 77
 Diploid, 2
 Dominance, 60, 114
Drosophila melanogaster, 2
 Duplicate factors, 41
- Efficiency, criterion of, 45
 — of statistics, 51 *et seq.*, 96,
 121
 Estimation, combined, 69 *et*
 seq.
 — criteria of, 44 *et seq.*
 — equation of, 46
 — inefficient, 54
 — method of, 7, 11, 44
 — scoring method of, 86 *et*
 seq.
 — simultaneous, 133 *et seq.*
 Eye colour, 123
- F_2 , 3, 36, 45, 57, 79, 88, 116,
 121
 — disturbed, 91, 96
 Falconer, 126
 Family size, 26 *et seq.*, 105 *et*
 seq.
 Finney, 124
 Fisher, 17, 24, 45, 51, 54, 86,
 96, 99 *et seq.*, 116 *et seq.*
 Frequency distribution, 8
- Gene frequencies, 125 *et seq.*
 Grouse locusts, *see Acridium*
- Haldane, 52, 114
 Heterogeneity, of data, 17 *et*
 seq., 37, 73 *et seq.*

148 MEASUREMENT OF LINKAGE IN HEREDITY

- Heterozygote, 3, 26
 Hierarchical tables, 20
 Homozygote, 26
 Human data, 104
 Hutchinson, 64

 Imai, 36, 48
 Immer, 54, 55
 Information, 53, 72, 77, 87,
 111, 134
 — calculation of, 57 *et seq.*
 Interpolation, solution by, 71

 Jenkins, 78, 92

 Landsteiner and Levine, 129
 Latvians, 132
 Likelihood, 46, 70
 — logarithm of, 46, 70
 Linear functions, 39
 Linkage, 4, 7, 44
 — with sex, 114

 Maize, 78, 92
 Manifestation, of characters,
 84, 95
 Mather, 17, 57 *et seq.*, 69 *et*
 seq.
 Matings, human, 105
 Maximization, 46, 70
 Maximum Likelihood, method
 of, 11, 45 *et seq.*, 77, 86,
 91, 95, 131
 Mean information, 73
 — weighted, 111
 Mice, 17, 22
 Misclassification, 27, 29
 Multinomial distribution, 11

 Nabours, 97
 Nettleship, 109
 Normal deviates, 14, 74
 — distribution, 10

 Orthogonal functions, 32 *et*
 seq., 38 *et seq.*, 129

 Orthogonal functions, deriva-
 tion of, 40
 Ovule sterility, 113

 Parameter, 7, 10, 44
 Pearson, 15, 109
 Penrose, 121
Pharbitis, 36, 48
 Phenylthiocarbamide, 125 *et*
 seq.
 Philp, 34, 47
Pisum, 113
 Precision, 12, 28
Primula sinensis, 2, 52
 Proband method, 106
 Product formula, 54, 94, 100
 — tables of, 55
 Poppy, 34, 47

 Race *et al.*, 132
 Random mating, 126
 — test of, 129
 Recessives, human, 116
 Recombination, 4, 5, 44
 Repulsion, 4, 61, 96, 117
 Retinitis pigmentosum, 114

 Sansome, 69
 Score, 87
 Segregation, 2, 29, 69
 Separation, index of, 69
 Sib method, 106 *et seq.*
 Significance, test of, 5, 9
 — meaning of, 6
 Simplex autotetraploid, 69
 Single factors, 2, 13
 — in humans, 104
 Smith, 84
 Snedecor, 22, 38, 128
 Standard error, 10, 13, 48, 50,
 73, 109, 119
 Statistic, 10, 11
 Sufficiency, criterion of, 45
 Sweet Pea, 73, 87

 Taste blindness, 125 *et seq.*
 Tetrad analysis, 3, 4
 Tomato, 70

INDEX

149

u statistics, 116
Usher, 109

Variance, 12, 39, 48, 50, 72,
96, 100 *et seq.*, 109, 119,
135

Viability, 95, 97

Weights, *see Mean*

Wheat, 84

Winton, De, 52

Yates, 123

**Printed in Great Britain by
Butler & Tanner Ltd.,
Frome and London**