

Genetic Mapping in Experimental Populations

Genetic linkage maps are an increasingly important tool in both fundamental and applied research, enabling the study and deployment of genes that determine important biological traits. This concise introduction to genetic mapping in species with disomic inheritance enables life science graduate students and researchers to use mapping software to produce more reliable results. After a brief refresher on meiosis and genetic recombination, the steps in the map construction procedure are described, with explanations of the computations involved. The emphasis throughout is on the practical application of the methods described; detailed mathematical formulae are avoided and exercises are included to help readers consolidate their understanding. A chapter on recognizing and solving problems provides valuable guidance for dealing with real-life situations. An extensive chapter dedicated to the more complex situation of outbreeding species offers a unique insight into the approach required for many economically important and model species, both plants and animals.

Dr J.W. (Johan) Van Ooijen has worked in the field of genetic linkage analysis for more than 20 years. He developed the QTL analysis software MapQTL[®] and was also involved in the development of JoinMap[®], Professor Piet Stam's genetic linkage mapping software. He continues the development of both software packages in his company, Kyazma B.V., and regularly teaches courses on genetic mapping and QTL analysis.

Dr J. (Hans) Jansen is a senior member of the Biometris group at Wageningen University & Research Centre. He is involved in research projects on various applications of molecular markers, including genetic diversity studies, the construction of genetic linkage maps and QTL analysis involving multiple pedigree-related populations.

Genetic Mapping in Experimental Populations

J.W. Van Ooijen

Kyazma B.V., Wageningen, The Netherlands

J. Jansen

*Biometris, Wageningen University &
Research Centre, Wageningen, The Netherlands*



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Published in the United States of America by Cambridge University Press, New York

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107013216

© J.W. Van Ooijen & J. Jansen 2013

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2013

Printed in the United Kingdom by Clays, St Ives plc

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Ooijen, J.W. Van, 1957–

Genetic mapping in experimental populations / J.W. Van Ooijen, J. Jansen.

p. ; cm.

Includes bibliographical references and index.

ISBN 978-1-107-01321-6 (hardback) – ISBN 978-1-107-60103-1 (pbk.)

I. Jansen, J. (Johannes), 1952– II. Title.

[DNLM: 1. Genetic mapping-methods, 2. genetic linkage, (3. diploid species)]
572.8'6 – dc23 2013005734

ISBN 978-1-107-01321-6 Hardback

ISBN 978-1-107-60103-1 Paperback

Additional resources for this publication at www.cambridge.org/9781107013216

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

To the memory of Piet Stam

Contents

<i>Preface</i>	<i>page xi</i>
1 Introduction	1
References	8
2 Meiosis and genetic recombination	9
2.1 The life cycle and the ploidy level	9
2.2 Two types of cell division	10
2.3 Two sexes or mating types	10
2.4 Meiosis	11
2.5 Following two loci in meiosis	13
Interchromosomal recombination	13
Intrachromosomal recombination	14
2.6 Recombination frequency	17
Crossover interference	17
Recombination frequency and physical distance	18
2.7 Polyploidy	18
2.8 Sex chromosomes	19
References	19
3 Estimation of recombination frequencies	21
3.1 What do we observe? From separate loci to pairs, from phenotype to genotype	21
Genotype and phenotype notation	23
3.2 The backcross	23
Estimation with the method of maximum likelihood	26
The likelihood ratio and the <i>LOD</i>	27
The variance of the estimator \hat{r} in the backcross	28
Doubled haploids	28
3.3 The F_2	29
The F_2 with dominance	33
One marker with a dominant allele	33
Coupling phase	34

Repulsion phase	35
Practical consequences of dominance	36
3.4 Advanced population types	37
Exercises	39
References	44
4 Determination of linkage groups	45
4.1 Why determine linkage groups and how?	45
4.2 Test for independent segregation	46
Example	48
Generalization	48
4.3 Test for the recombination frequency being equal to $\frac{1}{2}$	49
Example	50
Generalization	50
4.4 Some caution	51
4.5 Grouping algorithm	52
4.6 Determining the linkage groups in practice	53
Exercises	56
References	57
5 Estimation of a genetic map	59
5.1 Recombination frequencies and distances	59
5.2 An example experiment	63
5.3 Map estimation by linear regression	65
Missing genotype observations	68
Crossover interference	68
5.4 Map estimation by maximum likelihood	69
Missing genotype observations	71
Crossover interference	72
Exercises	72
References	75
6 Criteria for the evaluation of maps	77
6.1 Introduction	77
6.2 Log-likelihood	78
6.3 Sum of <i>LODs</i> of adjacent segments	79
6.4 Sum of recombination frequencies in adjacent segments	80
6.5 Residual sum of squares	81
6.6 χ^2 Goodness of fit	82
Exercises	83
7 How to find the best map order	87
7.1 Introduction	87

7.2	The travelling salesman problem	88
7.3	Optimization algorithms	90
7.4	Greedy algorithms	91
7.5	Simulated annealing	92
7.6	Tabu search	93
7.7	Evolutionary algorithms	94
7.8	Ant colony optimization	95
7.9	Distance geometry	96
7.10	Final remarks	96
	Exercise	96
	References	98
8	Outbreeding species	101
8.1	Introduction	101
8.2	Segregation type and linkage phases	102
8.3	Two-way pseudo-testcross strategy	106
	Linkage phase determination in the pseudo-testcross	108
8.4	Estimation of the pairwise recombination frequency in the full-sib family	110
	Linkage phase determination in the full-sib family	111
8.5	Simultaneous estimation of the separate parental recombination frequencies	112
	Example	114
	Multipoint estimation	116
8.6	Map estimation	117
	Exercises	119
	References	120
9	Mapping in practice	123
9.1	Introduction	123
9.2	Model assumptions	124
9.3	Observation errors	125
9.4	Phenotype and genotype	127
9.5	Missing observations	128
9.6	Pre-mapping diagnostics	129
	Phenotype frequencies	129
	Segregation ratio	129
	Identical loci and individuals	131
	Pairwise recombination frequencies	131
9.7	Post-mapping diagnostics	132
	Segregation distortion	132
	Number of recombinations per individual	132
	Phenotype probabilities and graphical genotypes	133

x Contents

Stress	134
Model fit	135
Plausible map orders	136
References	137
<i>Answers to exercises</i>	139
<i>Index</i>	155



Preface

Genetic linkage maps play an important role in many branches of biological research. The specific calculations involved in the construction of genetic maps ask for dedicated computer software. When using this software, the construction of genetic maps looks simple: just press a few buttons and you get a map. However, things are not as simple as they look. The various computations are based on a fair amount of theory from genetics, probability, statistics and optimization. We felt the need to explain this theory and write a special textbook dedicated to the construction of genetic maps. We think it is essential reading for those who want to construct reliable genetic maps.

Ever since the late 1980s, when the first RFLP projects started in Wageningen, we have been involved in linkage analysis with molecular genetic markers. By the mid-1990s, we started teaching short courses on the construction of genetic maps. Currently, our course covers 3 days of lectures and practicals. This book is a complete rewrite of our course reader. The book is aimed at the biologist with an interest in the construction of genetic maps. For a proper understanding, knowledge of genetics, probability and statistics at the undergraduate level is required.

Writing this book would not have been possible without the cooperation of many colleagues, especially the geneticists and plant breeders of Wageningen University and Research Centre. We are very grateful to Dr Hans de Jong, Professor of Cytogenetics at Wageningen University, for his comments and suggestions on the chapter about meiosis.

We dedicate this book to the memory of the late Professor Piet Stam. For a long time, he was a great inspiration to both of us. For J.W.V.O., this started as early as 1976 when he attended Piet's course on population genetics at the Agricultural University in Wageningen. For J.J., it really started in 1990 when Piet joined the group of geneticists and statisticians at the Centre for Plant Breeding and Reproduction Research (now part of Wageningen University & Research Centre). Piet Stam was a very enthusiastic scientist who contributed significantly to genetics. We name just two of his papers: 'Interference in genetic crossing over and chromosome mapping' (*Genetics*, 1979, 92, 573–94), and 'Construction of

integrated genetic linkage maps by means of a new computer package: JoinMap' (*The Plant Journal*, 1993, 3, 739–44). We owe much to him for his inspiration and knowledge.

*J.W. Van Ooijen & J. Jansen
Wageningen, The Netherlands
7 November 2012*