

Contents

Part I Genetic Mapping 1

- 1 Mapping Populations and Principles of Genetic Mapping 3**
Katharina Schneider
Overview 3
Abstract 3
 - 1.1 Introduction 4
 - 1.2 Mapping Populations 6
 - 1.2.1 Mapping Populations Suitable for Self-fertilizing Plants 7
 - 1.2.1.1 F_2 Populations 7
 - 1.2.1.2 Recombinant Inbred Lines 8
 - 1.2.1.3 Backcross Populations 9
 - 1.2.1.4 Introgression Lines: Exotic Libraries 10
 - 1.2.1.5 Doubled Haploid Lines 11
 - 1.2.2 Mapping Populations for Cross-pollinating Species 12
 - 1.2.3 Two-step Strategies for Mapping Mutants and DNA Fragments 12
 - 1.2.4 Chromosome-specific Tools for Mapping 13
 - 1.2.5 Mapping in Natural Populations/Breeding Pools 14
 - 1.2.6 Mapping Genes and Mutants to Physically Aligned DNA 15
 - 1.2.7 Specific Mapping Problems 15
 - 1.3 Discussion 17

Acknowledgments 18
References 19
- 2 Molecular Marker Systems for Genetic Mapping 23**
Henry T. Nguyen and Xiaolei Wu
Abstract 23
 - 2.1 Introduction 23
 - 2.2 DNA-based Markers Popularly Used in Genetic Mapping 25
 - 2.2.1 RFLP 25
 - 2.2.1.1 Conventional RFLP Analysis 26
 - 2.2.1.2 PCR-RFLP 26
 - 2.2.1.3 Mismatch PCR-RFLP 27

2.2.2	RAPD	28
2.2.3	SSR Markers	29
2.2.3.1	Conventional SSR Analysis	30
2.2.3.2	ISSR	34
2.2.3.3	STMP	36
2.2.4	AFLP	36
2.2.4.1	Conventional AFLP Analysis	36
2.2.4.2	f-AFLP	39
2.2.4.3	cDNA-AFLP and HiCEP	40
2.2.4.4	TE-AFLP	42
2.2.4.5	MEGA-AFLP	42
2.2.4.6	MITE-AFLPs	44
2.2.4.7	AFLP Conversion	45
2.2.5	REMAP and IRAP	47
2.2.5.1	IRAP	48
2.2.5.2	REMAP	48
2.2.6	SRAP	48
2.3	Discussion	49
	References	50
3	Methods and Software for Genetic Mapping	53
	<i>James C. Nelson</i>	
	Overview	53
	Abstract	53
3.1	Introduction	54
3.1.1	Methods and Tools for Genetic Linkage Mapping in Plants	54
3.1.1.1	Statement of the Problem	54
3.1.2	Locus Grouping	55
3.1.3	Locus Ordering	55
3.1.4	Multilocus Distance Estimation	57
3.1.5	Using Variant and Mixed Cross Designs	57
3.1.5.1	Outbreeding Species	57
3.1.5.2	Autopolyploid Species	58
3.1.5.3	Combining Datasets	58
3.1.6	Linkage-mapping Software Availability, Interfaces, and Features	58
3.2	Methods and Tools for QTL Mapping in Plants	59
3.2.1	Statement of the Problem	59
3.2.2	Single-marker Association	60
3.2.2.1	Metric Traits	60
3.2.2.2	Categorical Traits	60
3.2.3	Interval Mapping: Simple (SIM)	61
3.2.3.1	ML Methods	61
3.2.3.2	Least-squares (Regression) and Nonparametric Methods	61
3.2.4	Interval Mapping: Composite (CIM)	63
3.2.5	Significance Testing	63

3.2.6	Interval Mapping: Multiple-QTL Model Building	64
3.2.6.1	Stepwise and Exhaustive-search Methods for Building Multiple-QTL Models	64
3.2.6.2	Markov Chain Monte Carlo (MCMC) Methods	65
3.2.6.3	Genetic Algorithms	65
3.2.7	Multiple-trait (MT) QTL Mapping	66
3.2.8	Multiple-cross (MC) QTL Mapping	66
3.2.9	Computational Optimization Methods	66
3.3	Future Directions in Mapping Methods and Tools	67
3.3.1	Future of Linkage and QTL Mapping	67
3.3.2	Adequacy of Software Tools for Plant Mapping	68
3.3.2.1	Software Merit Criteria	68
3.3.2.2	Analytical Scope	68
3.3.2.3	Ease of Learning and Use	68
3.3.2.4	Accessibility and Extensibility	69
3.3.3	A Development Model for Public Genetic Mapping Software	69
	References	70
4	Single nucleotide Polymorphisms: Detection Techniques and Their Potential for Genotyping and Genome Mapping	75
	<i>Günter Kahl, Andrea Mast, Nigel Tooke, Richard Shen, and Dirk van den Boom</i>	
4.1	Introduction	75
4.2	Selected Techniques	80
4.2.1	SNP Analysis I: The Invader Technology	80
4.2.1.1	Introduction	80
4.2.1.2	Cleavage Enzymes	80
4.2.1.3	Oligonucleotides and Structure	80
4.2.1.4	Probe Cycling and Signal Amplification	81
4.2.1.5	DNA Format	81
4.2.1.6	RNA Format	81
4.2.1.7	Alternate Detection Formats	82
4.2.1.8	Specificity	83
4.2.1.9	Robustness	85
4.2.1.10	Invader Applications	85
4.2.1.11	Conclusion	85
4.2.2	SNP Analysis II: Pyrosequencing	85
4.2.2.1	Introduction	85
4.2.2.2	SNP Genotyping Using Pyrosequencing Technology	86
4.2.2.3	Allele Frequency Quantification	88
4.2.2.4	Haplotyping	90
4.2.3	SNP Analysis III: A Scalable High-multiplex SNP Genotyping Platform	91
4.2.3.1	Introduction	91
4.2.3.2	The Illumina Genotyping Platform	91
4.2.3.3	Genotyping Data	95

3.2.6	Interval Mapping: Multiple-QTL Model Building	64
3.2.6.1	Stepwise and Exhaustive-search Methods for Building Multiple-QTL Models	64
3.2.6.2	Markov Chain Monte Carlo (MCMC) Methods	65
3.2.6.3	Genetic Algorithms	65
3.2.7	Multiple-trait (MT) QTL Mapping	66
3.2.8	Multiple-cross (MC) QTL Mapping	66
3.2.9	Computational Optimization Methods	66
3.3	Future Directions in Mapping Methods and Tools	67
3.3.1	Future of Linkage and QTL Mapping	67
3.3.2	Adequacy of Software Tools for Plant Mapping	68
3.3.2.1	Software Merit Criteria	68
3.3.2.2	Analytical Scope	68
3.3.2.3	Ease of Learning and Use	68
3.3.2.4	Accessibility and Extensibility	69
3.3.3	A Development Model for Public Genetic Mapping Software	69
	References	70
4	Single nucleotide Polymorphisms: Detection Techniques and Their Potential for Genotyping and Genome Mapping	75
	<i>Günter Kahl, Andrea Mast, Nigel Tooke, Richard Shen, and Dirk van den Boom</i>	
4.1	Introduction	75
4.2	Selected Techniques	80
4.2.1	SNP Analysis I: The Invader Technology	80
4.2.1.1	Introduction	80
4.2.1.2	Cleavage Enzymes	80
4.2.1.3	Oligonucleotides and Structure	80
4.2.1.4	Probe Cycling and Signal Amplification	81
4.2.1.5	DNA Format	81
4.2.1.6	RNA Format	81
4.2.1.7	Alternate Detection Formats	82
4.2.1.8	Specificity	83
4.2.1.9	Robustness	85
4.2.1.10	Invader Applications	85
4.2.1.11	Conclusion	85
4.2.2	SNP Analysis II: Pyrosequencing	85
4.2.2.1	Introduction	85
4.2.2.2	SNP Genotyping Using Pyrosequencing Technology	86
4.2.2.3	Allele Frequency Quantification	88
4.2.2.4	Haplotyping	90
4.2.3	SNP Analysis III: A Scalable High-multiplex SNP Genotyping Platform	91
4.2.3.1	Introduction	91
4.2.3.2	The Illumina Genotyping Platform	91
4.2.3.3	Genotyping Data	95

4.2.3.4	Conclusion	96
4.2.4	SNP Analysis IV: High-throughput SNP Analysis by MALDI-TOF MS	96
4.2.4.1	Introduction	96
4.2.4.2	SNP Analysis Using the Massarray Platform	97
4.3	Conclusions and Perspectives	103
	Acknowledgments	104
	References	104

5 **Breeding By Design: Exploiting Genetic Maps and Molecular Markers Through Marker-assisted Selection** 109

Johan D. Peleman, Anker P. Sørensen, and Jeroen Rouppe van der Voort

Abstract 109

5.1	Introduction	110
5.2	Marker-assisted Selection	111
5.2.1	Genetic Distance Analysis, Variety Identification, and Seed Purity Analysis	113
5.2.2	Indirect Selection	113
5.2.2.1	Monogenic Traits	113
5.2.2.2	Polygenic (Quantitative) Traits	114
5.2.2.3	Marker-assisted Backcrossing	116
5.3	The Creation of Novel Varieties (Marker-assisted Breeding)	116
5.3.1	Removal of Linkage Drag	117
5.3.2	Pyramiding Resistance Genes	117
5.3.3	Marker-assisted Breeding of Polygenic Traits	119
5.3.4	Introduction of Novel Characteristics	120
5.3.5	Effective Exploitation of (Exotic) Germ Plasm	120
5.4	Breeding by Design	121
5.4.1	Mapping Loci Involved in All Agronomically Relevant Traits	121
5.4.2	Assessment of the Allelic Variation at the Loci Associated With Agronomically Relevant Traits	124
5.4.3	Breeding by Design	125
5.4.4	Future of Breeding by Design	127
	Acknowledgments	127
	References	128

Part II **Physical Mapping** 131

6 **Physical Mapping of Plant Chromosomes** 133

Barbara Hass-Jacobus and Scott A. Jackson

6.1	Introduction	133
6.2	Classical Physical Mapping Techniques	134
6.2.1	Knob Mapping in Maize	134
6.2.2	Deletion/Aneuploid/Substitution Mapping	134
6.2.3	Polytene Chromosomes in <i>Drosophila</i>	136
6.2.4	Chromosome Banding	136

6.3	Molecular Physical Mapping Techniques	139
6.3.1	CHEF Gel Mapping	139
6.3.2	Radiation Hybrid Mapping	139
6.3.4	Large-insert Clone Libraries (YACs, BACs, Cosmids)	140
6.3.5	Fluorescent <i>in situ</i> Hybridization	142
6.3.6	Mapping Gene Space	144
6.4	Discussion	145
	References	146
7	Chromosome Flow Sorting and Physical Mapping	151
	<i>Jaroslav Doležel, Marie Kubaláková, Jan Bartoš, and Jiří Macas</i>	
	Overview	151
	Abstract	151
7.1	Introduction	152
7.2	Development of Flow Cytogenetics for Plants	153
7.2.1	The Uses of Flow Karyotyping	155
7.2.2	Applications of Flow-sorted Chromosomes	155
7.3	Methodologies and Techniques	157
7.3.1	Cell Cycle Synchronization and Metaphase Accumulation	157
7.3.2	Chromosome Isolation	158
7.3.3	Chromosome Analysis	158
7.3.3.1	Detection of Structural Chromosome Changes	160
7.3.3.2	Quantitative Detection of Numerical Chromosome Changes	160
7.3.3.3	Chromosome Sorting	160
7.3.3.4	Cytogenetic Mapping on Flow-sorted Chromosomes	162
7.3.3.5	Physical Mapping Using PCR	163
7.3.3.6	Chromosome and Chromosome Arm-specific DNA Libraries	165
7.3.3.7	Targeted Isolation of Molecular Markers	166
7.4	Discussion	167
	Acknowledgments	168
	References	168
8	Genomic DNA Libraries and Physical Mapping	173
	<i>Chengwei Ren, Zhanyou Xu, Shuku Sun, Mi-Kyung Lee, Chengcang Wu, Chantel Scheuring, and Hong-Bin Zhang</i>	
	Overview	173
	Abstract	173
8.1	Introduction	174
8.2	Methodologies and Techniques	176
8.2.1	Bacteria-based Large-insert DNA Clones	176
8.2.1.1	BAC	178
8.2.1.2	PAC	179
8.2.1.3	BIBACs	180
8.2.1.4	PBC	181
8.2.1.5	TAC	182

8.2.2	Genomic DNA Library Quality and Genome Physical Mapping	182
8.2.2.1	Insert Sizes	183
8.2.2.2	Clone Genome Coverage	184
8.2.2.3	Representation for Genome	185
8.2.2.4	The Importance of Binary Vectors in Plant Physical Mapping	187
8.2.3	Construction of Bacteria-based Large-insert Genomic DNA Libraries	188
8.2.3.1	Megebase-size Nuclear DNA Preparation	188
8.2.3.2	Vector Preparation	191
8.2.3.3	Library Construction	196
8.2.4	Applications of Bacteria-based Large-insert Genomic DNA Libraries in Genome Physical Mapping	202
8.2.4.1	Physical Mapping by Fingerprint Analysis	203
8.2.4.2	Physical Mapping by Fluorescent <i>in situ</i> Hybridization	204
8.2.4.3	Physical Mapping by Optical Mapping	204
8.2.4.4	Physical Mapping by Iterative Hybridization	205
8.2.4.5	Physical Mapping by Other Techniques	206
8.3	Discussion	208
	Acknowledgments	209
	References	210
9	Integration of Physical and Genetic Maps	215
	<i>Khalid Meksem, Hirofumi Ishihara, and Tacco Jesse</i>	
9.1	Introduction	215
9.2	Colony Hybridization Techniques for the Integration of DNA Sequences Into a Physical Map	216
9.3	Integrating Gene-rich Sequences into the Physical Map Via Overgo Hybridization	218
9.3.1	Overlapping Oligo Labeling	218
9.3.2	Hybridization	219
9.3.3	Washing	220
9.3.4	Autoradiography	220
9.4	Pooling Strategy for the Integration of DNA Sequences Into a Physical Map	220
9.4.1	DNA Pooling	220
9.4.2	PCR Multiplexing	222
9.4.3	DNA Isolation from Pooled Large-insert Clones	222
9.5	Forward and Reverse Integrated Physical-Genetic Mapping: Targeted DNA Marker Mapping	223
9.5.1	Integrated AFLP Mapping	224
9.5.2	Targeted SSR Mapping	225
9.5.2.1	Subcloning of Large-insert DNA Clones	226
9.5.2.2	Colony Hybridization	226

9.6	Bioinformatic Tools for an Integrated Physical-Genetic Map: Genome Browser	227
9.6.1	GBrowse	227
9.6.2	iMap	229
9.7	Discussion	230
	References	231
10	Positional Cloning of Plant Developmental Genes	233
	<i>Peter M. Gresshoff</i>	
	Overview	233
	Abstract	234
10.1	Introduction	235
10.2	Gene Discovery Through Insertional Mutagenesis	238
10.3	Technology Requirements for Map-based Cloning	240
10.4	Positional Cloning Successes in Legume Nodulation Genes	245
10.4.1	Nodulation Biology	245
10.4.2	Non-nodulation Genes	247
10.4.3	Autoregulation of Nodulation Genes	249
10.5	Conclusions	252
	Acknowledgments	253
	References	253
11	Whole-genome Physical Mapping: An Overview on Methods for DNA Fingerprinting	257
	<i>Chengcang Wu, Shuku Sun, Mi-Kyung Lee, Zhanyou Xu, Chengwei Ren, Teofila S. Santos, and Hong-Bin Zhang</i>	
	Overview	257
	Abstract	257
11.1	Introduction	258
11.1.1	Inception and Development of DNA Fingerprinting	258
11.1.2	Clone-based Whole-genome Physical Mapping	259
11.1.3	Source DNA Libraries for Whole-genome Physical Mapping	260
11.2	Techniques and Methodologies	261
11.2.1	Preparation of DNA from Bacteria-based Large-insert Clones	263
11.2.1.1	Isolation of DNA from Bacteria-based Large-insert Clones Using the Clone-by-Clone Approach	263
11.2.1.2	Isolation of DNA from Bacteria-based Large-insert Clones in 96-well Format Using a Manual Approach	264
11.2.1.3	Isolation of DNA from Bacteria-based Large-insert Clones in 96-well Format Using the AutoGenprep 960 Robotic Workstation	265
11.2.2	Fingerprinting the Clone DNA Using Restriction Enzymes	265
11.2.2.1	Agarose Gel-based, Restriction Fingerprinting Method	265
11.2.2.2	Polyacrylamide (Sequencing) Gel-based, Restriction Fingerprinting Method	268

11.2.2.3	Automatic Sequencing Gel- and Capillary Electrophoresis-based Restriction Fingerprinting Methods	271
11.2.3	Applications of DNA Fingerprinting for Genome Physical Mapping	275
11.2.3.1	Whole-genome Physical Maps Generated from Large-insert Bacterial Clones by Fingerprint Analysis	275
11.2.3.2	Comparison of Different Fingerprinting Methods	277
11.3	Discussion	280
	Acknowledgements	282
	References	282
12	Software for Restriction Fragment Physical Maps	285
	<i>William Nelson and Carol Soderlund</i>	
	Overview	285
12.1	Introduction	285
12.1.1	Review of Agarose Fingerprinting and FPC	285
12.2	HICF Techniques	288
12.2.1	Generalities	288
12.2.2	Approaches Using Type IIS Enzymes	289
12.2.3	SNaPshot HICF	290
12.3	Processing HICF Data	291
12.3.1	Peak Scoring	291
12.3.2	Sizing of Fragments	292
12.3.3	Quality and Contamination Screening	293
12.3.4	Vector and Repeat Screening	294
12.3.5	Packaging HICF Data for FPC	295
12.3.6	GenoProfiler	295
12.4	Building HICF Maps in FPC	296
12.4.1	Creating a New Project, Loading the Fingerprints, and Screening Vectors	297
12.4.2	Tolerance and Gel Length	298
12.4.3	Cutoff	299
12.4.4	Building	300
12.4.5	Build Quality, Q Clones, and the DQer	300
12.4.6	multiFPC	301
12.5	Theoretical Aspects	302
12.5.1	Simulations	302
12.5.2	Overlap Equations	303
	Acknowledgments	304
	References	304

13	Reduced Representation Strategies and Their Application to Plant Genomes	307
	<i>Daniel G. Peterson</i>	
	Overview	307
13.1	Introduction	308
13.2	Reduced Representation Techniques	311
13.2.1	EST Sequencing	311
13.2.2	Methylation Filtration	311
13.2.3	Cot-based Cloning and Sequencing	313
13.2.4	<i>De Novo</i> Polymorphism Discovery	316
13.2.4.1	Reduced Representation Shotgun (RRS) Sequencing	317
13.2.4.2	DOP-PCR	318
13.2.4.3	SSR Capture	318
13.3	Other Reduced Representation Techniques	320
13.4	Discussion	320
13.4.1	Repeat Sequence Enrichment	320
13.4.2	Sequencing Gene space	322
13.4.2.1	EST Sequencing	322
13.4.2.2	Methylation Filtration	323
13.4.2.3	Cot-based Cloning and Sequencing	325
13.4.2.4	Integration of Reduced Representation Strategies	327
13.4.3	Polymorphism Discovery	328
13.4.3.1	RRS Sequencing and DOP-PCR	328
13.4.3.2	Microsatellite Isolation	328
13.5	Conclusions	329
	Acknowledgments	330
	References	330
14	Large-scale DNA Sequencing	337
	<i>Christopher D. Town</i>	
	Abstract	337
14.1	Introduction	337
14.2	Motivation for Sequencing and Choice of Strategy	338
14.3	EST Sequencing	338
14.4	Large-scale Sequencing BAC-based Projects	339
14.5	Chromosome-based Sequencing	340
14.6	Whole-genome Shotgun Sequencing	341
14.7	Selective Genome Sequencing by Differential Cloning Strategies	341
14.8	Low-coverage Sequencing	341
14.9	BAC End Sequencing	343
14.10	The Sequencing Process Itself	343
14.10.1	Library Production	343
14.10.2	Template Production	344
14.10.3	Sequencing Reactions and Analysis	344
14.10.4	Sequencing Capacity and Costs	345

14.10.5	Post-sequencing Data Processing	345
14.10.5.1	Quality Trimming	345
14.10.5.2	Sequence Assembly	346
14.10.5.3	Scaffolding	346
14.10.5.4	Sequence Editing and Gap Closure	347
14.11	Conclusions	347
	Acknowledgments	348
	References	348

Glossary	353
-----------------	-----

Subject Index	371
----------------------	-----