
Preface

Most traits in nature and of importance to agriculture are quantitatively inherited. These traits are difficult to study due to the complex nature of their inheritance. However, recent developments of genomic technologies provide a revolutionary means for unraveling the secrets of genetic variation in quantitative traits. Genomic technologies allow the molecular characterization of polymorphic markers throughout the entire genome that are then used to identify and map the genes or quantitative trait loci (QTLs) underlying a quantitative trait based on linkage analysis.

Statistical analysis is a crucial tool for analyzing genome data, which are now becoming increasingly available for a variety of species, and for giving precise explanations regarding genetic variation in quantitative traits occurring among species, populations, families, and individuals. In 1989, Lander and Botstein published a hallmark methodological paper for interval mapping that enables geneticists to detect and estimate individual QTL that control the phenotype of a trait. Today, interval mapping is an important statistical tool for studying the genetics of quantitative traits at the molecular level, and has led to the discovery of thousands of QTLs responsible for a variety of traits in plants, animals, and humans. In a recent study published in *Science*, Li, Zhou, and Sang (2006, *311*, 1936–1939) were able to characterize the molecular basis of the reduction of grain shattering – a fundamental selection process for rice domestication – at the detected QTL by interval mapping. Among many other examples of the success of interval mapping are the positional cloning of QTLs responsible for fruit size and shape in tomato (Frary et al. 2000, *Science* *289*, 85–88) and for branch, floescence, and grain architecture in maize (Doebley et al. 1997, *Nature* *386*, 485–488; Gallavotti et al. 2004, *Nature* *432*, 630–635; Wang et al. 2005, *Nature* *436*, 714–719).

To make it suitable for various practical applications, interval mapping has been extensively modified and extended during the past 15 years. A host of useful statistical methods for QTL mapping have been produced through the collective efforts of statistical geneticists. However, these methods generally have various objectives and utilities and are sporadically distributed in a massive amount of literature. A single volume synthesizing statistical developments for genetic mapping may be helpful for many researchers, especially those with a keen interest in building a bridge between

genetics and statistics, to acquaint themselves with this expanding area as quickly as possible.

This book intends to provide geneticists with the tools needed to understand and model the genetic variation for quantitative traits based on genomic data collected in mapping research and equip statisticians with the uniqueness and ideas in relation to the exploration of genetic secrets using their computational skills. This book also intends to attract researchers toward multidisciplinary research and to introduce them to new paradigms in genomic science. In this book, the statistical and computational theories applied to genetic mapping are developed hand in hand and a number of examples displaying the implications of statistical genomics are introduced.

This book contains 14 chapters, broadly divided into three parts. Part 1, including Chapters 1 and 2, provides introductory genetics and statistics at the level appropriate for understanding general genetic concepts and statistical models for genetic mapping. Part 2, composed of Chapters 3–7, attempts to provide a thorough and comprehensive coverage of linkage analysis with molecular markers. Models and methods for linkage analysis and map construction are systematically introduced for different designs, such as the backcross/ F_2 (Chapter 3), outbred crosses (Chapter 4), recombinant inbred lines (Chapter 5) and structured pedigrees (Chapter 7), and for special marker types including distorted and misclassified markers (Chapter 6) and dominant markers (Chapters 4 and 7). Part 3, composed of Chapters 8–14, covers statistical models and algorithms of QTL mapping. The topics include simple marker-phenotype association analyses (Chapter 8), the statistical structure of interval mapping (Chapter 9), regression- (Chapter 10) and maximum likelihood-based analysis of interval mapping (Chapter 11), threshold and confidence interval determination (Chapter 12), composite interval mapping using multiple markers as cofactors (Chapter 13), and interval mapping for outbred mapping populations (Chapter 14). In the Appendices, we provide general statistical theories directly related to the genetic mapping approaches introduced and R programs for some of the examples used in the book. A webpage (<http://www.buffalo.edu/~cxma/book/>) was constructed for this book, which includes a complete list of programs and algorithms written in MatLab or R for all the examples.

Writing a book in such a rapidly developing and changing field is a pain but, more precisely speaking, full of excitement. In the summer of 1997, Wu delivered a series of lectures on statistical methods for QTL mapping to graduate students and faculty at Nanjing Forestry University, China. In the spring semester of 2002, Wu taught a statistical genetics course at the master's level at the University of Florida and then was joined for coteaching by Casella in the spring of 2003 and Ma in the spring of 2005. This course is now taught by Wu at the University of Florida and by Ma at the State University of New York at Buffalo on the regular basis. We all gave many lectures or short courses related to statistical genetics at other places and times. At each place and time, we were heavily impressed by the enthusiasm of students and other audiences to learn this fascinating area. All these encouraged us to write a book that can cover basic methods for statistical genetics research. The concepts, models and algorithms related to genetic mapping have been published in a variety of statistics and genetics journals by a large number of authors, but part of the material

contained in this book comes from our collaborative research program in the past five years. In particular, we apologize for those authors whose work was not mentioned in this book because of limited space.

During the writing of this book, many of our colleagues and friends both at the University of Florida and outside provided valuable help from different perspective. Wu is warmly grateful to his postdoctoral advisor, Dr. Zhao-Bang Zeng at North Carolina State University, for tremendous guidance and for leading him to the field of statistical genetics. Dr. Bruce Walsh at the University of Arizona provided insightful reviews of the book manuscript in different stages. Several anonymous reviewers gave constructive comments that significantly improve the presentation of the book. Students or postdocs who attended our lectures and classes or are working with us on statistical genetics in different places have provided many insightful suggestions to improve our presentation of the book. The following students or postdocs in our group, former or current, deserve special thanks: Yuehua Cui, Wei Hou, Hongying Li, Min Lin, Tian Liu, Fei Long, Xiang-Yang Lou, Qing Lu, Damaris Santana, Zhaojie Wang, Zuoheng Wang, Jiasheng Wu, Song Wu, Jie Yang, John Yap, Li Zhang, Wei Zhao, and Yun Zhu. The data used for examples in the book were kindly supplied by Dr. James Cheverud at Washington University (mouse), Dr. Junyi Gai at Nanjing Agricultural University (soybean), Rory Todhunter at Cornell University (dog), Drs. Stan Wulschleger and Tongming Yin at Oak Ridge National Laboratory (poplar), and Dr. Jun Zhu at Zhejiang University (rice).

We are grateful to the Department of Statistics and the Institute of Food and Agricultural Sciences at the University of Florida for writing this book and performing our research program. Finally, we are greatly indebted to our respective families for their continuous support of our research activities over the years. This work is partially supported by NSF grant 0540745.

Gainesville, FL
Buffalo, NY
December 2006

Rongling Wu
Chang-Xing Ma
George Casella

Statistical Genetics of Quantitative Traits

Linkage, Maps and QTL

Wu, R.; Ma, C.; Casella, G.

2007, XVI, 368 p., Hardcover

ISBN: 978-0-387-20334-8