Contents

List of Contributors page vii

Preface xi

1. An Introduction to Next-Generation Biological Platforms
   Virginia Mohlere, Wenting Wang, and Ganiraju Manyam 1

2. An Introduction to The Cancer Genome Atlas
   Bradley M. Broom and Rehan Akbani 31

3. DNA Variant Calling in Targeted Sequencing Data
   Wenyi Wang, Yu Fan, and Terence P. Speed 54

4. Statistical Analysis of Mapped Reads from mRNA-Seq Data
   Ernest Turro and Alex Lewin 77

5. Model-Based Methods for Transcript Expression-Level
   Quantification in RNA-Seq  Zhaonan Sun, Han Wu, Zhaohui Qin, and Yu Zhu 105

6. Bayesian Model-Based Approaches for Solexa Sequencing
   Data  Riten Mitra, Peter Mueller, and Yuan Ji 126


8. Bayesian Modeling of ChIP-Seq Data from Transcription
   Factor to Nucleosome Positioning  Raphael Gottardo and Sangsoon Woo 170

9. Multivariate Linear Models for GWAS  Chiara Sabatti 188

10. Bayesian Model Averaging for Genetic Association Studies
    Christine Peterson, Michael Swartz, Sanjay Shete, and Marina Vannucci 208

11. Whole-Genome Multi-SNP-Phenotype Association Analysis
    Yongtao Guan and Kai Wang 224
## Contents

12. Methods for the Analysis of Copy Number Data in Cancer Research  
   Bradley M. Broom, Kim-Anh Do, Melissa Bondy, Patricia Thompson, and Kevin Coombes  
   244

13. Bayesian Models for Integrative Genomics  
   Francesco C. Stingo and Marina Vannucci  
   272

14. Bayesian Graphical Models for Integrating Multiplatform Genomics Data  
   Wenting Wang, Veerabhadran Baladandayuthapani, Chris C. Holmes, and Kim-Anh Do  
   292

15. Genetical Genomics Data: Some Statistical Problems and Solutions  
   Hongzhe Li  
   312

16. A Bayesian Framework for Integrating Copy Number and Gene Expression Data  
   Yuan Ji, Filippo Trentini, and Peter Mueller  
   331

17. Application of Bayesian Sparse Factor Analysis Models in Bioinformatics  
   Haisu Ma and Hongyu Zhao  
   350

18. Predicting Cancer Subtypes Using Survival-Supervised Latent Dirichlet Allocation Models  
   Keegan Korthauer, John Dawson, and Christina Kendziorski  
   366

19. Regularization Techniques for Highly Correlated Gene Expression Data with Unknown Group Structure  
   Brent A. Johnson  
   382

20. Optimized Cross-Study Analysis of Microarray-Based Predictors  
   Xiaogang Zhong, Luigi Marchionni, Leslie Cope, Edwin S. Iversen, Elizabeth S. Garrett-Mayer, Edward Gabrielson, and Giovanni Parmigiani  
   398

   Laila M. Poisson and Debashis Ghosh  
   423

22. Discover Trend and Progression Underlying High-Dimensional Data  
   Peng Qiu  
   445

23. Bayesian Phylogenetics Adapts to Comprehensive Infectious Disease Sequence Data  
   Jennifer A. Tom, Janet S. Sinsheimer, and Marc A. Suchard  
   460

Index  
477

Color plates follow page 104.