

Roded Sharan (Ed.)

Research in Computational Molecular Biology

18th Annual International Conference, RECOMB 2014
Pittsburgh, PA, USA, April 2-5, 2014
Proceedings

Table of Contents

| | |
|---|-----|
| Tractatus: An Exact and Subquadratic Algorithm for Inferring Identical-by-Descent Multi-shared Haplotype Tracts | 1 |
| <i>Derek Aguiar, Eric Morrow, and Sorin Istrail</i> | |
| HapTree: A Novel Bayesian Framework for Single Individual Polyplotyping Using NGS Data | 18 |
| <i>Emily Berger, Deniz Yorukoglu, Jian Peng, and Bonnie Berger</i> | |
| Changepoint Analysis for Efficient Variant Calling | 20 |
| <i>Adam Bloniarz, Ameet Talwalkar, Jonathan Terhorst, Michael I. Jordan, David Patterson, Bin Yu, and Yun S. Song</i> | |
| On the Representation of de Bruijn Graphs | 35 |
| <i>Rayan Chikhi, Antoine Limasset, Shaun Jackman, Jared T. Simpson, and Paul Medvedev</i> | |
| Exact Learning of RNA Energy Parameters from Structure | 56 |
| <i>Hamidreza Chitsaz and Mohammad Aminisharifabad</i> | |
| An Alignment-Free Regression Approach for Estimating Allele-Specific Expression Using RNA-Seq Data | 69 |
| <i>Chen-Ping Fu, Vladimir Jojic, and Leonard McMillan</i> | |
| The Generating Function Approach for Peptide Identification in Spectral Networks | 85 |
| <i>Adrian Guthals, Christina Boucher, and Nuno Bandeira</i> | |
| Decoding Coalescent Hidden Markov Models in Linear Time | 100 |
| <i>Kelley Harris, Sara Sheehan, John A. Kamm, and Yun S. Song</i> | |
| AptaCluster – A Method to Cluster HT-SELEX Aptamer Pools and Lessons from Its Application | 115 |
| <i>Jan Hoinka, Alexey Berezhnoy, Zuben E. Sauna, Eli Gilboa, and Teresa M. Przytycka</i> | |
| Learning Sequence Determinants of Protein: Protein Interaction Specificity with Sparse Graphical Models | 129 |
| <i>Hetunandan Kamisetty, Bornika Ghosh, Christopher James Langmead, and Chris Bailey-Kellogg</i> | |
| On Sufficient Statistics of Least-Squares Superposition of Vector Sets | 144 |
| <i>Arun S. Konagurthu, Parthan Kasarapu, Lloyd Allison, James H. Collier, and Arthur M. Lesk</i> | |

| | |
|---|-----|
| IDBA-MTP: A Hybrid MetaTranscriptomic Assembler Based on Protein Information | 160 |
| <i>Henry C.M. Leung, S.M. Yiu, and Francis Y.L. Chin</i> | |
| MRAlign: Protein Homology Detection through Alignment of Markov Random Fields | 173 |
| <i>Jianzhu Ma, Sheng Wang, Zhiyong Wang, and Jinbo Xu</i> | |
| An Integrated Model of Multiple-Condition ChIP-Seq Data Reveals Predeterminants of Cdx2 Binding | 175 |
| <i>Shaun Mahony, Matthew D. Edwards, Esteban O. Mazzone, Richard I. Sherwood, Akshay Kakumanu, Carolyn A. Morrison, Hynek Wichterle, and David K. Gifford</i> | |
| PASTA: Ultra-Large Multiple Sequence Alignment | 177 |
| <i>Siavash Mirarab, Nam Nguyen, and Tandy Warnow</i> | |
| Fast Flux Module Detection Using Matroid Theory | 192 |
| <i>Arne C. Müller, Frank J. Bruggeman, Brett G. Olivier, and Leen Stougie</i> | |
| Building a Pangenome Reference for a Population | 207 |
| <i>Ngan Nguyen, Glenn Hickey, Daniel R. Zerbino, Brian Raney, Dent Earl, Joel Armstrong, David Haussler, and Benedict Paten</i> | |
| CSAX: Characterizing Systematic Anomalies in eXpression Data | 222 |
| <i>Keith Noto, Carla Brodley, Saeed Majidi, Diana W. Bianchi, and Donna K. Slonim</i> | |
| WHATSHAP: Haplotype Assembly for Future-Generation Sequencing Reads | 237 |
| <i>Murray Patterson, Tobias Marschall, Nadia Pisanti, Leo van Iersel, Leen Stougie, Gunnar W. Klau, and Alexander Schönhuth</i> | |
| Simultaneous Inference of Cancer Pathways and Tumor Progression from Cross-Sectional Mutation Data | 250 |
| <i>Benjamin J. Raphael and Fabio Vandin</i> | |
| DIPSPADES: Assembler for Highly Polymorphic Diploid Genomes | 265 |
| <i>Yana Safonova, Anton Bankevich, and Pavel A. Pevzner</i> | |
| An Exact Algorithm to Compute the DCJ Distance for Genomes with Duplicate Genes | 280 |
| <i>Mingfu Shao, Yu Lin, and Bernard Moret</i> | |

| | |
|--|-----|
| HIT'nDRIVE: Multi-driver Gene Prioritization Based on Hitting Time | 293 |
| <i>Raunak Shrestha, Ermin Hodzic, Jake Yeung, Kendric Wang, Thomas Sauerwald, Phuong Dao, Shawn Anderson, Himisha Beltran, Mark A. Rubin, Colin C. Collins, Gholamreza Haffari, and S. Cenk Sahinalp</i> | |
| Modeling Mutual Exclusivity of Cancer Mutations | 307 |
| <i>Ewa Szczurek and Niko Beerenwinkel</i> | |
| Viral Quasispecies Assembly via Maximal Clique Enumeration | 309 |
| <i>Armin Töpfer, Tobias Marschall, Rowena A. Bull, Fabio Luciani, Alexander Schönhuth, and Niko Beerenwinkel</i> | |
| Correlated Protein Function Prediction via Maximization of Data-Knowledge Consistency | 311 |
| <i>Hua Wang, Heng Huang, and Chris Ding</i> | |
| Bayesian Multiple Protein Structure Alignment | 326 |
| <i>Rui Wang and Scott C. Schmidler</i> | |
| Gene-Gene Interactions Detection Using a Two-Stage Model | 340 |
| <i>Zhanyong Wang, Jae-Hoon Sul, Sagi Snir, Jose A. Lozano, and Eleazar Eskin</i> | |
| A Geometric Clustering Algorithm and Its Applications to Structural Data | 356 |
| <i>Shutan Xu, Shuxue Zou, and Lincong Wang</i> | |
| A Spatial-Aware Haplotype Copying Model with Applications to Genotype Imputation | 371 |
| <i>Wen-Yun Yang, Farhad Hormozdiari, Eleazar Eskin, and Bogdan Pasaniuc</i> | |
| Traversing the k -mer Landscape of NGS Read Datasets for Quality Score Sparsification | 385 |
| <i>Y. William Yu, Deniz Yorukoglu, and Bonnie Berger</i> | |
| Reconstructing Breakage Fusion Bridge Architectures Using Noisy Copy Numbers | 400 |
| <i>Shay Zakov and Vineet Bafna</i> | |
| Reconciliation with Non-binary Gene Trees Revisited | 418 |
| <i>Yu Zheng and Louxin Zhang</i> | |

| | |
|---|------------|
| Learning Protein-DNA Interaction Landscapes by Integrating Experimental Data through Computational Models..... | 433 |
| <i>Jianling Zhong, Todd Wasson, and Alexander J. Hartemink</i> | |
| Imputation of Quantitative Genetic Interactions in Epistatic MAPs by Interaction Propagation Matrix Completion | 448 |
| <i>Marinka Žitnik and Blaž Zupan</i> | |
| Author Index | 463 |