Guest Editorial for the 17th Asia Pacific Bioinformatics Conference

Louxin Zhang, Shaoliang Peng, Yi-Ping Phoebe Chen, David Sankoff, Guoliang Li, and Hong-Yu Zhang

THIS special section of *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* is a collection of eight papers presented at the 17th Asia Pacific Bioinformatics Conference (APBC), which was held in Wuhan, China, 14–16 January 2019. The conference aims to bring together researchers, academics, and industrial practitioners to interact. Its scientific program included three keynote, eight highlight, 51 contributed oral and 33 contributed poster presentations. The conference drew close to 150 attendees from all over the world.

The seven papers in this special section cover diverse topics, including gene tree-species tree reconciliation, genome evolution, data structures for genome assembly, protein binding-site inference and Boolean network modelling of protein-protein networks.

Gene tree-species tree reconciliation is an approach to inferring evolutionary history for gene families. It is also used to investigate the evolutionary histories of parasites and hosts as well as of species and their biogeographical habitats. In this approach, the discordance of a gene tree and a species tree is visualized by mapping the gene tree onto the species tree and is measured by the minimum number of postulated evolutionary events (e.g., gene duplication, gene loss, horizontal gene transfer) that are necessary for reconciling the trees, called the reconciliation cost of the two trees. In "The Unconstrained Diameters of the Duplication-Loss Cost and the Loss Cost," Gorecki, Eulenstein, and Tiuryn study a few mathematics and algorithmic issues of the maximum reconciliation cost between two arbitrary binary trees in the duplication-loss model and the gene loss model.

In "Reconciliation Reconsidered: In Search of a Most Representative Reconciliation in the Duplication-Transfer-Loss Model," Grueter *et al.* studies how well a single reconciliation can represent the entire space of optimal reconciliations. They answer an open problem of Huber *et al.* (Inform. Process. Letts, vol. 136, 2018) by developing an efficient algorithm for computing a median reconciliation under the path distance metric in the duplication-transfer-loss model. They also present an efficient algorithm for determining the maximum possible average support value that an optimal reconciliation can achieve.

In "Multiple Optimal Reconciliations under the Duplication-Loss-Coalescence Model," Du *et al.* presents new algorithms for computing the size of the space of the parsimony reconciliations under the duplication-loss-coalescence model and for uniformly sampling from this space. The proposed algorithms are efficient, with runtime polynomial in the sizes of the species tree and gene tree when the number of genes that map to any given species is fixed. De Bruijn graph is a data structure that is frequently used to design efficient computer assembler for de novo genome sequencing. It also plays an important role in metagenomics analysis. In "[D]eGSM: Memory Scalable Construction of Large Scale de Bruijn Graph," Guo *et al.* develops a lightweight parallel approach to constructing de Burijn graph of millions of reads. With its high scalability and efficiency, the construction method has great potential in large scale genomics study.

Boolean networks that are used to model different biological systems are usually large, dense and are of modular structure. In "Towards Optimal Decomposition of Boolean Networks," Su, Pang, and Paul propose a novel method for decomposition of BNs to balance the relation between structure and dynamics of a network.

Inverse tandem duplication random loss (iTDRL) event is proposed by researchers to model the gene order evolution of the mitochondrial genome, in which a continuous segment of a gene order is duplicated and inverted, followed by the random loss of one of the redundant copies for each gene. In "Sorting Signed Permutations by Inverse Tandem Duplication Random Losses," Hartmann, Bannach, and Middendorf study the algorithmic issues of genome rearrangement in iTDRL model. They show that a shortest rearrangement scenario that transforms a gene order into another can be obtained in quasilinear time and the length of such a scenario can be calculated in linear time.

Flavin mono-nucleotide (FMN) is a protein that is involved as a co-factor in carrying and transferring electrons in cellular respiration. In "Prediction of FMN Binding Sites in Electron Transport Chain based on 2-D CNN and PSSM Profiles," Nguyen and Le propose a deep learning model for FMN binding sites using a two-dimensional

- *Louxin Zhang is with the Department of Mathematics, National University of Singapore, Singapore 119076. E-mail: matzlx@nus.edu.sg.*
- *Shaoliang Peng is with the National Supercomputer Centre, Hunan University, Wuhan 410082, China. E-mail: slpeng@hnu.edu.cn.*
- *Yi-Ping Phoebe Chen is with the Department of Computer Science and Information Technology, La Trobe University, Bundoora VIC 3086, Australia. E-mail: phoebe.chen@latrobe.edu.au.*
- *David Sankoff is with the Department of Mathematics and Statistics, University of Ottawa, Ottawa, ON K1N 6N5, Canada. E-mail: sankoff@uottawa.ca.*
- *Guoliang Li and Hong-Yu Zhang are with the College of Information, Huazhong Agricultural University, Wuhan 430070, China. E-mail: {guoliang.li, zhy630}@mail.hzau.edu.cn.*

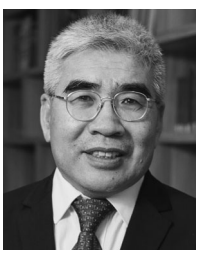
(Corresponding author: Louxin Zhang)

Digital Object Identifier no. 10.1109/TCBB.2021.3099948

convolutional neural network and position specific scoring matrices profiles.

We would like to thank the program committee members and external researchers who reviewed the selected papers: Tatsuya Akutsu, Paola Bonizzoni, Pawel Gorecki, Katharina Huber, Sun Kim, Manuel Lafond, Kui Lin, Minsik Oh, Laxmi Parida, Mingfu Shao, Jens Stoye, Glenn Tesler, Lusheng Wang, and Shuqin Zhang, Lastly, special thanks go to the editorial staff of the *IEEE/ACM TCBB* for assistance in publishing this special section of the APBC'2019.

Guest Editors
Louxin Zhang
Shaoliang Peng
Yi-Ping Phoebe Chen
David Sankoff
Guoliang Li
Hong-Yu Zhang



Louxin Zhang received the PhD degree in computer science from the University of Waterloo, Canada. He is currently a professor with the Department of Mathematics and the director of the Center for Data Science and Machine Learning, the National University of Singapore. He has authored or coauthored more than 100 papers in mathematics, computer science, and bioinformatics, and is the coauthor of a monograph on sequence comparison (with K. M. Chao) and a bioinformatics textbook (with B. Ma). His current research interests include bioinformatics for comparative genomics and cancer genomics, complex networks, big data in biomedicine, phylogeny, comparative genomics, and cancer genomics. He is on the Editorial Board of Quantitative Biology and other journals.

Shaoliang Peng is currently the executive director and a Changjiang Scholars professor with the National Supercomputing Center and College of Computer Science and Electronic Engineering, Hunan University, Changsha, China. He has authored or coauthored dozens of academic papers in several internationally influential journals, including Science, Nature Communications, Cell, The American Journal of Human Genetics, Nucleic Acids Research, Genome Research, Bioinformatics, ACM/IEEE Transactions, Bioinformatics, and Biomedicine. His research interests include high-performance computing, bioinformatics, big data, precision medicine, health informatics, and computer-aided drug design. He was the editor of several international journals, including executive editors of the International Journal of Biological Sciences and an executive editors-in-chief of Interdisciplinary Sciences: Computational Life Sciences International Journal. He was the chairman and sponsor of the Hunan Provincial Bioinformatics Society, China, the deputy director of Biological Information Professional Group in CCF, and invited speaker of the BIBM 2016 conference.

Yi-Ping Phoebe Chen (Senior Member, IEEE) received the BInfTech and PhD degrees in computer science from the University of Queensland. She is currently a professor and the chair with the Department of Computer Science and Computer Engineering, La Trobe University. She has coauthored more than 150 research papers in many top journals and conferences, such as the IEEE Transactions on Biomedical Engineering, Nucleic Acids Research, and ACM Transactions. She is currently the steering committee chair of the Asia Pacific Bioinformatics Conference (founder) and Multimedia Modelling. She has been on the program committees of more than 100 international conferences, including top ranking conferences, such as ICDE, ICPR, ISMB, and CIKM. She is a member of the ACM.

David Sankoff received the PhD degree in mathematics from McGill University. He is currently a member of the Centre de Recherches Mathematiques, Montreal, Canada. He is currently the Canada research chair in mathematical genomics with Mathematics and Statistics Department, University of Ottawa, and is cross appointed with the Biology and the Computer Science Departments. His research interests include comparative genomics, particularly probability models, statistics, and algorithms for genome rearrangements, with a focus on the genomes of flowering plants.

Guoliang Li received the PhD degree in computer science from the National University of Singapore. He is currently a professor and a vice dean of the College of Informatics, Huazhong Agricultural University, Wuhan, China. He is a senior member of the China Computer Federation and a member of the ACM. He has co-authored more than 80 research papers with many published in top journals such as Cell, Nature Genetics, Nature Communications, Genome Biology, Nucleic Acids Research, and IEEE TKDE. His research interests include bioinformatics and epigenomics, particularly 3D genomics.

Hong-Yu Zhang received the PhD degree in biophysics from the Institute of Biophysics, Chinese Academy of Sciences, in 1997. He is currently a professor and the dean of the College of Informatics, Huazhong Agricultural University. His research interests include systems evolution and evolutionary biomedicine. He has authored or coauthored more than 100 articles in internationally influential journals. He was the recipient of the Servier Young Investigator Prize from Chinese Pharmacological Society in 1999, Young Investigator Prize in Chemistry from Chinese Chemical Society in 2005, and Prize for Natural Science Research in 2011 from the Ministry of Education, China. He is the chairman of the Hubei Society of Bioinformatics and is on the Editorial Board of Frontiers in Genetics, Frontiers in Bioscience-Landmark, Protein & Peptide Letters, Current Computer-Aided Drug Design, Current Bioinformatics, and Current Chinese Science.