# Guest Editorial for Special Section on BSB 2012

Marcilio C.P. de Souto and Maricel Kann

THIS special section includes a selection of papers presented at the Seventh Brazilian Symposium on Bioinformatics (BSB 2012) held 15-17 August 2012 in Campo Grande (Mato Grosso do Sul), Brazil. BSB is an international symposium that covers all aspects of bioinformatics and computational biology. The symposium is organized by the special interest group in Computational Biology of the Brazilian Computer Society (SBC).

BSB 2012 had an international program committee of 55 members. All papers submitted to the symposium were reviewed by at least three independent reviewers. After the review process, 16 papers were accepted. These papers appeared in the conference proceedings published by Springer-Verlag in the series *Lecture Notes in Computer Science/Lecture Notes in Bioinformatics* (volume 7409).

The authors of the best BSB papers were invited to submit an extended version to this special section. Following a rigorous review process, three papers were selected for publication. The selected papers cover a broad range of bioinformatics topics, including the genome rearrangement problem, RNA secondary structure prediction, and validation methodologies for cluster analysis of gene expression data.

In the first paper, "Extending the Algebraic Formalism for Genome Rearrangements to Include Linear Chromosomes," Pedro Feijão and João Meidanis extend the algebraic approach to genome rearrangement to the most general genome model consisting of circular and linear chromosomes. More specifically, they propose a new way of modeling genomes which can be seen as a mixture of algebraic rearrangement theory with the adjacency formalism. As a result, linear chromosomes can be modeled directly. The authors also fully discuss and put into perspective their framework with respect to those existing in the combinatorial field.

The paper "2D Meets 4G: G-Quadruplexes in RNA Secondary Structure Prediction," by Ronny Lorenz, Stephan H. Bernhart, Jing Qin, Christian Honer zu Siederdissen, Andrea Tanzer, Fabian Amman, Ivo L. Hofacker, and Peter F. Stadler, describes how G-quadruplex structures can be incorporated, using an energy function, into standard RNA structure prediction dynamic programming algorithms. G-quadruplexes are abundant locally stable structural elements in nucleic acids. In terms of results, the authors show that, with the preliminary energy parameters used, their algorithm could find that the majority of putative quadruplex-forming sequences in the human genome are likely to fold into canonical secondary structures.

Finally, in "Proximity Measures for Clustering Gene Expression Microarray Data: A Validation Methodology and a Comparative Analysis," Pablo A. Jaskowiak, Ricardo J.G.B. Campello, and Ivan G. Costa present a data-driven analysis of the choice of proximity measures for the clustering of microarray data. The authors investigate the performance of 16 proximity measures in 52 data sets from time-course and cancer experiments. The proximity measures, based on their distinct characteristics, are evaluated separately for cancer and time-course experiments. The authors show that measures rarely employed in the gene expression literature can provide better results than commonly employed ones, such as Pearson and Spearman correlation in addition to euclidean distance.

We would like to express our sincere thanks to the BSB 2012 Program Committee members for their invaluable help in making BSB 2012 such a success. Also, our special thanks to the external reviewers for volunteering their time to review the submissions to the symposium and the special section. We would also like to thank Marie-France Sagot and Ying Xu, as the former and current editor-in-chief of the *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (*TCBB*), for offering us this opportunity for wider dissemination of the research presented at BSB in *TCBB*. Last but not least, we would like to thank the authors for their time and effort in submitting their high quality work to BSB. Without them, the symposium would not have been possible.

Marcilio C. P. de Souto
Maricel Kann
*Guest Editors*

- *M.C.P. de Souto is with the Laboratoire d'Informatique Fondamentale d'Orléans (LIFO), Université d'Orléans, LIFO Bat. 3IA, Rue Léonard de Vinci, B.P. 6759, 45067, Orléans, France, and the Centro de Informática, Universidade Federal de Pernambuco, Av. Jornalista Anibal Fernandes, 50.740-560, Recife–PE, Brazil. E-mail: marcilio.desouto@univ-orleans.fr.*
- *M. Kann is with the Department of Biological Sciences, University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250. E-mail: mkann@umbc.edu.*

*For information on obtaining reprints of this article, please send e-mail to: tcbb@computer.org.*

**Marcilio C.P. de Souto** received the BSc degree in computer science from the Universidade Federal do Rio Grande do Norte (UFRN), Brazil, in 1991, the MSc degree in computer science from the Universidade Federal de Pernambuco (UFPE), Brazil, in 1995, and the PhD degree in electrical engineering (artificial intelligence) from Imperial College London, United Kingdom, in 1999. After receiving the PhD degree, he worked as a visiting professor at the Centro de Informãtica (CIn) of the UFPE. In 2004, he joined, as an associate professor, the Departamento de Informática e Matemática Aplicada of the UFRN. Currently, he is a full professor in the Laboratoire d'Informatique Fondamentale d'Orléans (LIFO) at the Université d'Orléans, France. He also collaborates with the postgraduate program in computing of the CIn/UFPE. Dr. de Souto is senior member of IEEE. He has consistently served on the program committees of the main Brazilian conferences on artificial intelligence, neural networks, and bioinformatics. He has also been guest editor of the *Journal of Intelligent and Fuzzy Systems*, the *International Journal of Neural Systems*, and *Neurocomputing*. His main research topic is machine learning, with focus on cluster analysis, hybrid intelligent systems, and bioinformatics.

**Maricel Kann** received the BSc and graduate degrees in chemistry from the Universidad de la Republica, Uruguay. In 2001, she received the doctoral degree from the University of Michigan under the guidance of Dr. Richard A. Goldstein. Her PhD thesis focused on the theory, statistics, and methods for protein sequence alignment. She is an associate professor at the University of Maryland, Baltimore County (UMBC). In 2002, Dr. Kann joined the National Center for Biotechnology Information (NIH) as a postdoctoral fellow, where she developed new algorithms for protein sequence comparison. Dr. Kann joined the Department of Biological Sciences at UMBC in 2007. Her research interests include classification of human variants, prediction of protein-protein interactions, and the study of protein domains and their associations with disease. Dr. Kann is one of the leading experts in the area of translational bioinformatics. She edited the first open access textbook on translational bioinformatics for *PLOS-Computational Biology* and organized the Translational Bioinformatics Conference. In addition, she has organized and chaired several international conference sessions at the Pacific Symposium on Biocomputing, Brazilian Symposium on Bioinformatics, the Intelligent Systems and Molecular Biology (ISMB), and the American Medical Informatics Association (AMIA) Summit in Bioinformatics. She is an associate editor for the journals *PLOS-Computational Biology* and the *Journal of Biomedical Informatics* and she is an advisory board member of the UniProt consortium and the NIH PubMedCentral National Committee.