

*Editorial*

## **Editorial: Special Issue on Algorithms for Sequence Analysis and Storage**

**Veli Mäkinen**

Helsinki Institute for Information Technology HIIT, Department of Computer Science, P. O. Box 68 (Gustaf Hällströmin katu 2b), University of Helsinki, Helsinki 00014, Finland;

E-Mail: vmakinen@cs.helsinki.fi

*Received: 14 March 2014; in revised form: 19 March 2014 / Accepted: 19 March 2014 /*

*Published: 25 March 2014*

---

**Abstract:** This special issue of *Algorithms* is dedicated to approaches to biological sequence analysis that have algorithmic novelty and potential for fundamental impact in methods used for genome research.

**Keywords:** high-throughput sequencing; compressed index structures; pattern discovery; population simulation; dynamic programming

---

### **1. Introduction**

Analysis of high-throughput sequencing data has become a crucial component in genome research. For example, methods based on latest developments in compressed data structures, namely index structures exploiting Burrows-Wheeler transform, are widely deployed in the discovery of disease causing mutations. The success of such approaches is due to being able to solve the dilemma of the indexing requiring more space than the data itself, where the data itself is enormous. With our already increased knowledge about the genomic structure of the whole human population, and with the development of sequencing techniques and their applications in studying RNAs, metapopulations, and epigenetics, the field seeks new innovative and universal algorithmic approaches that assess current and future needs in the analysis and storage of biological sequences.

### **2. Special Issue**

In response to the call for papers, ten articles were submitted to the special issue. Each article was reviewed by at least two anonymous reviewers, and after one to two review rounds, six articles

were selected for the special issue. The range of topics of accepted papers widely covers the analysis and storage theme: the authors contributed in compression of sequencing data [1], compressed index structures [2], pattern/motif discovery [3,4], population simulation [5], and with a novel framework for dynamic programming [6].

## Acknowledgments

As Guest Editor, I would like to thank the authors for submitting their work to the special issue, the reviewers for their constructive comments, and the publishers, Editor-in-Chief, Professor Kazuo Iwama, and Assistant Editors Chelly Cheng and Maple Lv, for their assistance in managing the review process.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Mazza, T.; Castellana, S. Multi-Sided Compression Performance Assessment of ABI SOLiD WES Data. *Algorithms* **2013**, *6*, 309–318.
2. Abeliuk, A.; Cánovas, R.; Navarro, G. Practical Compressed Suffix Trees. *Algorithms* **2013**, *6*, 319–351.
3. Comin, M.; Verzotto, D. Filtering Degenerate Patterns with Application to Protein Sequence Analysis. *Algorithms* **2013**, *6*, 352–370.
4. Fu, B.; Fu, Y.; Xue, Y. Sublinear Time Motif Discovery from Multiple Sequences. *Algorithms* **2013**, *6*, 636–677.
5. Haiminen, N.; Utro, F.; Lebreton, C.; Flament, P.; Karaman, Z.; Parida, L. Efficient *in silico* Chromosomal Representation of Populations via Indexing Ancestral Genomes. *Algorithms* **2013**, *6*, 430–441.
6. Giegerich, R.; Touzet, H. Modeling Dynamic Programming Problems over Sequences and Trees with Inverse Coupled Rewrite Systems. *Algorithms* **2014**, *7*, 62–144.

© 2014 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).