## Contents

### Papers

735 Measures for the degree of overlap of gene signatures and applications to TCGA  
Xingjie Shi, Huangdi Yi and Shuangge Ma

745 Environmental genes and genomes: understanding the differences and challenges in the approaches and software for their analyses  
Mads Lundegaard, Verena Meurkens, Thomas Schmidt-Nielsen and M. Thomas P. Gilbert

759 Detecting associations of rare variants with common diseases: collapsing or haplotyping?  
Aileng Wang and Shi Lin

769 Pathogenicity prediction of non-synonymous single nucleotide variants in dilated cardiomyopathy  
Sabiné C. Muñoz, Christina Rickes, Jan Moor, the RENAISSANCE Study Group, Hugo A. Katus, Benjamin Mehle, Eckart Meese and Andreas Keller

780 Comprehensive overview and assessment of computational prediction of microRNA targets in animals  
Kiaa Fan and Ekkart Adlung

795 Advantages of mixing bioinformatics and visualization approaches for analyzing sRNA-mediated regulatory bacterial networks  
Patricia Thebaud, Ximom Bouarah, William Brenchley, Christine Goupil, Pascal Standl-Pogon, Raluca Uricaru and Isabelle Dutour

806 LncTar: a tool for predicting the RNA targets of long noncoding RNAs  
Jianwei Li, Wen Mei, Pan Zeng, Jinyi Wang, Bin Geng, Jichuan Yang and Qinghua Cui

---

### Software Review

813 Sample size calculation in metabolic phenotyping studies  
Elise Billoir, Vincent Navratil and Benjamin J. Blaise

---

### Papers

820 Case study: using sequence homology to identify putative phosphorylation sites in an evolutionarily distant species (honeybee)  
Brett Trost, Scott Napper and Anthony Kusalik

830 A review of in silico approaches for analysis and prediction of HIV-1-human protein-protein interactions  
Sanghamitra Bandyopadhyay, Sumanta Ray, Anirban Mukhopadhyay and Ujjwal Maulik

852 Making the difference: integrating structural variation detection tools  
Ke Liu, Sandra Smil, Gaojun Bannerman, Gabriele Santoro-Pieroni and Dick de Ridder

865 The complexity, challenges and benefits of comparing two transporter classification systems in TCDB and Pfam  
Zachary Chiang, Ake Vastermark, Marco Pinto, Penelope C. Coogill, Jenny Nixtry, Robert D. Finn and Milton H. Saier, Jr

873 A selective review of robust variable selection with applications in bioinformatics  
Cen Xu and Shuangge Ma

884 Discriminative pattern mining and its applications in bioinformatics  
Jadorya Liu, Jun Wu, Feiyang Gu, Jie Wang and Zengyou He

---

### Opinion Note

901 Virtual Reference Environments: a simple way to make research reproducible  
Daniel A. Morley, David M. Budden and Edmund J. Crampin
The aim of Briefings in Bioinformatics is to provide an indispensable resource for the experimental practitioner seeking awareness of the disparate sources of data and analytical tools of contemporary biology, biotechnology and medicine based on the molecular level. This includes all areas of genomics, proteomics, metabolomics, interactomics and network biology, imaging, systems biology, chemoinformatics, computational biology and clinical/medical informatics that have a molecular foundation to the study. Large-scale instrumentation and computation is reducing the time that needs to be spent in the laboratory. Instead, the time-limiting step is the analysis and interpretation of data. The journal provides topical reviews of new methodologies as they become established.

The Editors, Martin Bishop and Russ Altman, welcome the submission of review papers and case studies for publication. Essential criteria for the publication of papers are that they provide practical guidance to the users of bioinformatics software and databases to supplement available manuals and tutorials, providing a broad conceptual review and overview of the problem addressed, its solution, dynamic range and limitations. We publish reviews of clearly defined subject areas for both experimental biologists and for bioinformatics specialists. Reviews may be broader or more narrowly focused but must cover a variety of approaches to a very well-specified biological problem or research area. We do not publish work on new methods that have not yet been described elsewhere. Please do not submit manuscripts concerning original research as they will be immediately rejected.

Some areas that members of the Editorial Board have identified as important and presently under represented in the pages of the Journal include the complex fields of epigenetics and the construction of evolutionary models and the elucidation of evolutionary relationships. Methodological approaches of interest include software comparison and benchmarking, data cleaning and curation, accuracy of predicted and extracted information, ontologies and text-mining, solutions that allow for the large-scale analysis of biological data in reasonable time (high performance computing solutions and cloud systems), standards, training and change management activities, and the determination of causal relationships from data. There was specific mention of the use of ontologies for semantic-based analysis of molecular data and interaction networks, methods and tools for the automatic or semiautomatic annotation of biological data with terms extracted from ontologies, and methods and tools for enrichment analysis. Articles focusing on illuminating the bottlenecks problems in important bioinformatics approaches will be especially helpful to readers. It must be emphasized that these comments are suggestions and are not intended to be prescriptive. As science advances the details of what is important changes and the Editors, Editorial Board and Reviewers will be flexible in their policies.

Types of Submission
Submissions of the following types are accepted for review in the Journal:

- Review articles (2000–7000 words)
- Protocols for solving a specific problem using different sets of programs (2000–5000 words)
- Case studies in biological research applied to clinical practice (2000–5000 words)
- Opinion articles: topical or controversial areas that do not warrant a full review (500–1000 words)
- Letters to the Editors (relating to a topic previously published in the journal) (500–1000 words)
- Software and website reviews (by authors other than the originators) (500–1000 words)
- Book reviews (500–1000 words)

Aims
The aim of Briefings in Bioinformatics is to provide an indispensable resource for the experimental practitioner seeking awareness of the disparate sources of data and analytical tools of contemporary biology, biotechnology and medicine based on the molecular level. This includes all areas of genomics, proteomics, metabolomics, interactomics and network biology, imaging, systems biology, chemoinformatics, computational biology and clinical/medical informatics that have a molecular foundation to the study. Large-scale instrumentation and computation is reducing the time that needs to be spent in the laboratory. Instead, the time-limiting step is the analysis and interpretation of data. The journal provides topical reviews of new methodologies as they become established.

The Editors, Martin Bishop and Russ Altman, welcome the submission of review papers and case studies for publication. Essential criteria for the publication of papers are that they provide practical guidance to the users of bioinformatics software and databases to supplement available manuals and tutorials, providing a broad conceptual review and overview of the problem addressed, its solution, dynamic range and limitations. We publish reviews of clearly defined subject areas for both experimental biologists and for bioinformatics specialists. Reviews may be broader or more narrowly focused but must cover a variety of approaches to a very well-specified biological problem or research area. We do not publish work on new methods that have not yet been described elsewhere. Please do not submit manuscripts concerning original research as they will be immediately rejected.

Some areas that members of the Editorial Board have identified as important and presently under represented in the pages of the Journal include the complex fields of epigenetics and the construction of evolutionary models and the elucidation of evolutionary relationships. Methodological approaches of interest include software comparison and benchmarking, data cleaning and curation, accuracy of predicted and extracted information, ontologies and text-mining, solutions that allow for the large-scale analysis of biological data in reasonable time (high performance computing solutions and cloud systems), standards, training and change management activities, and the determination of causal relationships from data. There was specific mention of the use of ontologies for semantic-based analysis of molecular data and interaction networks, methods and tools for the automatic or semiautomatic annotation of biological data with terms extracted from ontologies, and methods and tools for enrichment analysis. Articles focusing on illuminating the bottlenecks problems in important bioinformatics approaches will be especially helpful to readers. It must be emphasized that these comments are suggestions and are not intended to be prescriptive. As science advances the details of what is important changes and the Editors, Editorial Board and Reviewers will be flexible in their policies.

Types of Submission
Submissions of the following types are accepted for review in the Journal:

- Review articles (2000–7000 words)
- Protocols for solving a specific problem using different sets of programs (2000–5000 words)
- Case studies in biological research applied to clinical practice (2000–5000 words)
- Opinion articles: topical or controversial areas that do not warrant a full review (500–1000 words)
- Letters to the Editors (relating to a topic previously published in the journal) (500–1000 words)
- Software and website reviews (by authors other than the originators) (500–1000 words)
- Book reviews (500–1000 words)