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, cheminformatics, computational biology and clinical/medical informatics that have a molecular foundation to the study. Large-scale instrumentation and computerisation is reducing the time that needs to be spent in the laboratory. Instead, the rate-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 brief 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 focussed 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 We do not publish work on new methods that have not yet been described elsewhere.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 bottleneck problems in important bioinformatics approaches will be especially helpful to readers. It must be emphasised 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)
- Protocol 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 originator) (500–1000 words)
- Book reviews (500–2000 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, cheminformatics, computational biology and clinical/medical informatics that have a molecular foundation to the study. Large-scale instrumentation and computerisation is reducing the time that needs to be spent in the laboratory. Instead, the rate-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 brief 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 focussed 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 bottleneck problems in important bioinformatics approaches will be especially helpful to readers. It must be emphasised 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)
- Protocol 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 originator) (500–1000 words)
- Book reviews (500–2000 words)
# Contents

## Papers

1. **A systematic evaluation of normalization methods in quantitative label-free proteomics**  
   Tommi Välikangas, Tomi Suomi and Laura L. Elo

12. **Inter-kingdom prediction certainty evaluation of protein subcellular localization tools: microbial pathogenesis approach for deciphering host microbe interaction**  
   Abdu A. Khan, Zahir Khan, Mohd. A. Islam and Azmat Q. Khan

23. **The present and future of de novo whole-genome assembly**  
   Jang-il Sohn and Jin-Wu Nam

41. **A bioimage informatics platform for high-throughput embryo phenotyping**  

52. **Combinatorial epigenetic regulation of non-coding RNAs has profound effects on oncogenic pathways in breast cancer subtypes**  
   Juan Xu, Zihan Wang, Su-yong Li, Jian Chen, Jiaojiao Zhang, Changjie Yang, Jiaona Zhao, Jing Li, Yongzheng Li and Xin Li

65. **Synthetic data sets for the identification of key ingredients for RNA-seq differential analysis**  
   Guillem Rigaill, Sandrine Balzergue, Véronique Brunaud, Eddy Blondet, Andrea Rau, Oded Roger, José Costa, Cathy Magey-Robassinou, Ludwine Soubigou-Tessonnet, Sébastien Aubourg, Claire Latre, Marie-Laure Martin-Maugisse and Étienne Berlandoy

77. **Notions of similarity for systems biology models**  
   Ruo Wenkel, Robert Werthofer, Tim Korzeniewski, Christian Knappe, Wolfram Liebermeister and Dagmar Wohlfarth

89. **The framework for population epigenetic study**  
   Linna Zhao, Di Liu, Jing Xu, Zhaoyang Wang, Yang Chen, Changgui Li, Ying Li, Guoyu Liu and Yonghuan Jiang

101. **Genome-wide DNA methylation analysis identifies candidate epigenetic markers and drivers of hepatocellular carcinoma**  
   Yangchang Zheng, Qiangam Huang, Zijian Bing, Tingying Liu, Chenghao Xue, Xinrui Song and Jin Gu

109. **An empirical likelihood ratio test robust to individual heterogeneity for differential expression analysis of RNA-seq**  
   Maoqi Xu and Liang Chen

118. **Computational pan-genomics: status, promises and challenges**  
   The Computational Pan-Genomics Consortium

136. **Transferring entropy to the realm of GxG interactions**  
   Pascale G. Ferrario and Inke R. König

148. **Comprehensive assessment and performance improvement of effect protein predictors for bacterial secretion systems III, IV and VI**  
   Yi An, Jianwei Wang, Chen Li, André Lemos, Tatiana Marquez-Lope, Jonathan Wilks, Yang Zhang, Geoffrey I. Webb, Jianglei Song and Pierre Airiau

162. **A graphical model approach visualizes regulatory relationships between genome-wide transcription factor binding profiles**  
   Parasa P. I. Ng, David Rau, Lance Weisnag and Berthold Gätgens

174. **A review of validation strategies for computational drug repositioning**  
   Adam S. Brown and Dingy J. Patel