

# Editorial

Welcome to a new issue of EMBnet.journal the international, Open Access, peer-reviewed bioinformatics journal.

In this issue, amongst other articles, you will read about SeqAhead, a new COST Action initiative that aims to tackle many of the complex informatics issues generated by the deluge of data flowing from Next Generation Sequencing (NGS) technologies.

Genome sequencing studies are showing promise of both diagnostic and therapeutic benefits in clinical settings; in the agricultural area, the complete genome sequences of important plants, like the tomato and the potato, are out in the public domain - such projects facilitate the creation of new plant varieties that are resistant to drought and harsh conditions. Clearly, a true biotechnological revolution has arrived, promising tangible biomedical and agri-food impacts. The need to train more researchers in the bioinformatics techniques necessary to deliver those impacts is therefore imperative; projects like SeqAhead are poised to help make this happen.

EMBnet.journal plays an integral part in the new data-driven landscape of the life sciences, not only providing updates on the latest tools, databases and educational materials relevant to the genomic revolution, but also offering general bioinformatics information to a growing community of researchers, from disciplines ranging from medicine to plant and farm animal sciences.

*EMBnet.journal Editorial Board*



Protein Spotlight (ISSN 1424-4721) is a periodical electronic review from the SWISS-PROT group of the Swiss Institute of Bioinformatics (SIB). It is published on a monthly basis and consists of articles focused on particular proteins of interest. Each issue is available, free of charge, in HTML or PDF format at <http://www.expasy.org/spotlight>.

We provide the EMBnet community with a printed version of issue 128. Please let us know if you like this inclusion.

Cover picture: Crowned Cranes, *Baelearica regulorum*, Masai Mara, Kenya, 2010. [© Martin Norling]

# Contents

Editorial .....	2
<b>Letters to the Editor</b>	
Mastering Data-Intensive Collaboration and Decision Making through a Cloud Infrastructure: The Dicode EU project .....	3
<b>News</b>	
Introduction into Systems Biology: Basics of Proteomics, Bioinformatics, Biostatistics & Integration of Data Generated by these Fields, Athens, Greece, 26-30 September 2011 .....	4
The Bioinformatics Roadshow, 1-3 November 2011, Athens, Greece .....	5
ISCB and EMBnet to collaborate on bioinformatics education and training .....	6
<b>Reports</b>	
SEQAHEAD - COST Action BM1006: Next Generation Sequencing Data Analysis Network .....	7
<b>Technical Notes</b>	
Cutadapt removes adapter sequences from high-throughput sequencing reads .....	10
Command line analysis of ChIP-seq results .....	13
GRID distribution supports clustering validation of large mixed microarray data sets .....	18
The future of HOPE: what can and cannot be predicted about the molecular effects of a disease causing point mutation in a protein? .....	25
<b>Research Papers</b>	
SEQscoring: a tool to facilitate the interpretation of data generated with next generation sequencing technologies .....	38
e-RGA: enhanced Reference Guided Assembly of Complex Genomes .....	46
Protein Spotlight 128 .....	53
<b>Node information</b> .....	55

## Editorial Board:

**Erik Bongcam-Rudloff**, The Linnaeus Centre for Bioinformatics, SLU/UU, SE, [erik.bongcam@bmc.uu.se](mailto:erik.bongcam@bmc.uu.se)

**Teresa K. Attwood**, Faculty of Life Sciences and School of Computer Sciences, University of Manchester, UK, [teresa.k.attwood@manchester.ac.uk](mailto:teresa.k.attwood@manchester.ac.uk)

**Domenica D'Elia**, Institute for Biomedical Technologies, CNR, Bari, IT, [domenica.delia@ba.itb.cnr.it](mailto:domenica.delia@ba.itb.cnr.it)

**Andreas Gisel**, Institute for Biomedical Technologies, CNR, Bari, IT, [andreas.gisel@ba.itb.cnr.it](mailto:andreas.gisel@ba.itb.cnr.it)

**Laurent Falquet**, Swiss Institute of Bioinformatics, Génopode, Lausanne, CH, [Laurent.Falquet@isb-sib.ch](mailto:Laurent.Falquet@isb-sib.ch)

**Pedro Fernandes**, Instituto Gulbenkian, PT, [pfern@igc.gulbenkian.pt](mailto:pfern@igc.gulbenkian.pt)

**Lubos Klucar**, Institute of Molecular Biology, SAS Bratislava, SK, [klucar@EMBnet.sk](mailto:klucar@EMBnet.sk)

**Martin Norling**, Swedish University of Agriculture, SLU, Uppsala, SE, [martin.norling@slu.se](mailto:martin.norling@slu.se)