

## Bioimage informatics: a new category in *Bioinformatics*

Hanchuan Peng<sup>1</sup>, Alex Bateman<sup>2</sup>, Alfonso Valencia<sup>3</sup> and Jonathan D. Wren<sup>4</sup>

<sup>1</sup>Janelia Farm Research Campus, Howard Hughes Medical Institute, VA 20147, USA, <sup>2</sup>Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK, <sup>3</sup>Structural Biology and BioComputing Programme, Spanish National Cancer Research Centre, E-28029 Madrid, Spain and <sup>4</sup>Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA

The last two decades have witnessed great advances in biological tissue labeling and automated microscopic imaging that, in turn, have revolutionized how biologists visualize molecular, sub-cellular, cellular, and super-cellular structures and study their respective functions. Tremendous volumes of multi-dimensional bioimaging data are now being generated in almost every branch of biology. How to interpret such image datasets in a quantitative, objective, automatic and efficient way has become a major challenge in current computational biology. Bioimage informatics methods have begun to turn image data into useful biological knowledge (Peng, 2008; Swedlow, *et al.*, 2009; Shamir, *et al.*, 2010; Danuser, 2011). The essential methods of bioimage informatics involve large-scale bioimage generation, visualization, analysis and management. Bioimage informatics also encompasses both hypothesis- and data-driven exploratory approaches, with an emphasis on how to generate biological knowledge and/or gain new insights that would otherwise be hard to achieve.

Early work in bioimage informatics began in the late 1990s. Increasingly, computer vision, image analysis, data mining, machine learning and pattern recognition methods have been applied to microscopic images to extract biological information and to generate ontology databases. The growing amount of bioimage data are quickly imposing additional demands on how to store, manage and retrieve such image datasets as well as the associated secondary meta-data. Data analysis, fusion and reconstruction techniques have also been developed to facilitate better image acquisition and formation. Joint analysis of image data in combination with other biological datasets, such as genomes and gene expression profiles, is also becoming more and more commonplace.

To meet the need of this growing field, the first international workshop on Bioimage Informatics was organized at Stanford University in 2005. It grew to be an annual event in this field. Other meetings on similar topics and related applications have also emerged since then. In 2010, the annual conference on Intelligent Systems for Molecular Biology (ISMB) established a paper-submission track on bioimaging data analysis and visualization.

While there is a noticeable need to publish high quality papers on bioimage informatics, so far no high-impact journal explicitly accepts this category of papers. We believe it is an appropriate time to create this new category in *Bioinformatics*. As of February 2012,

*Bioinformatics* now includes a new paper submission category in the scope described by the journal at its website as follows:

'Informatics methods for the acquisition, analysis, mining and visualization of images produced by modern microscopy, with an emphasis on the application of novel computing techniques to solve challenging and significant biological and medical problems at the molecular, sub-cellular, cellular, and super-cellular (organ, organism, and population) levels. This category also encourages large-scale image informatics methods/applications/software, various enabling techniques (e.g. cyber-infrastructures, quantitative validation experiments, pattern recognition, etc.) for such large-scale studies, and joint analysis of multiple heterogeneous datasets that include images as a component. Bioimage related ontology and databases studies, image-oriented large-scale machine learning, data mining, and other analytics techniques are also encouraged.

We will not consider image analysis and pattern recognition methods that are solely based on tuning parameters or swapping computational sub-steps, without an in-depth description or demonstration of why such changes are significantly superior for one or more biological problems.'

We would like to thank a number of colleagues and practitioners who have contributed to the creation of this new category. Especially, we thank Eugene Myers, B.S.Manjunath, Badri Roysam, Manfred Auer, Michael Hawrylycz, Jean-Christophe Olivo-Marin, Anne Carpenter and Vebjorn Ljosa in helping to define this new category of paper submissions. We hope the journal *Bioinformatics* becomes a valuable venue for bioimage informatics researchers to publish their most important work.

**Contact:** pengh@janelia.hhmi.org

### REFERENCES

- Danuser,G. (2011) Computer vision in cell biology. *Cell*, **147**, 973–978.
- Peng,H. (2008) Bioimage informatics: a new area of engineering biology. *Bioinformatics*, **24**, 1827–1836.
- Shamir,L. *et al.* (2010) Pattern recognition software and techniques for biological image analysis. *PLoS Comput. Biol.*, **6**, e1000974.
- Swedlow,J.R. *et al.*; OME Consortium. (2009) Bioimage informatics for experimental biology. *Annu. Rev. Biophys.*, **38**, 327–346.