

# Visual Analytics for Biological Data

**Raghu Machiraju** ■ *The Ohio State University*

**Carsten Görg** ■ *University of Colorado Medical School*

**Arthur Olson** ■ *The Scripps Research Institute*

**T**he preeminence of a systems approach in the biological and other life sciences creates enormous challenges for computational visualization techniques to enable researchers to gain insight from their large, highly complex, and multiple datasets.

For this special issue, we sought contributions that suitably describe applications of visual analytics (VA) techniques to yield viable results that are biologically relevant and significant. In particular, we solicited case studies that were greatly facilitated by VA or couldn't have been completed without it. Existing, modified, or new analytical and visualization techniques, when applied to verify prevailing or novel hypotheses, often lead to new insights and hypotheses as manifested in the data. Ideally, the submitted case studies would be conducted by teams of researchers from the visualization and bioinformatics or biology communities.

The submissions we received included solutions for data on the organization, form, and function of phenotypes. They covered scales ranging from the molecular and subcellular to cells, tissue, organisms, and populations thereof. At least two experts reviewed each paper: one from the visualization community and one from the bioinformatics or biology community. After considering the reviews, the magazine accepted the following three articles.

"DIVE: A Graph-Based Visual-Analytics Framework for Big Data," by Steven Rysavy, Dennis Bromley, and Valerie Daggett, presents software that lets users manipulate structured data models. DIVE (Data-Intensive Visualization Engine) is

data-agnostic and ontologically expressive, and allows the streaming of large datasets at interactive speeds. Besides presenting the platform's technical details, the authors discuss several usage examples and a case study from a molecular-dynamics project.

In "Characterizing Cancer Subtypes Using Dual Analysis in Caleydo StratomeX," Cagatay Turkay, Alexander Lex, Marc Streit, Hanspeter Pfister, and Helwig Hauser describe how they comprehensively analyze and characterize cancer subtypes. Their approach uses sufficient statistics to characterize a high-dimensional dataset in StratomeX, a Caleydo view for analyzing cancer subtypes. Because their approach uses robust significant-difference plots to show the elements of a subtype that significantly distinguish it from other subtypes, analysts can easily characterize subtypes. Furthermore, analysts can investigate how samples relate to the assigned subtypes and to the other groups of samples. This ability is critical to developing biomarkers and patient stratification. Three case studies demonstrate the approach's utility.

In "Visual Exploration of Parameter Influence on Phylogenetic Trees," Martin Hess, Sebastian Bremm, Stephanie Weissgraeber, Kay Hamacher, Michael Goesele, Josef Wiemeyer, and Tatiana von Landesberger explain how they created and used VA tools to explore phylogenetic trees inferred from multiple sequence alignments. The ensuing parametric space is potentially very large, which renders the task of determining the most reliable evolutionary relationships highly intractable. The authors' approach depicts the alignment param-

eters' impact on the trees. It hierarchically clusters input trees from different parameter settings and visualizes similar clusters together with their construction parameters. The authors applied the approach to analysis of ribosomal RNA and protein sequences of ion channels in bacteria.

**W**e hope these articles inspire new perspectives on, and spur efforts in, visualizing large, heterogeneous biological data. This special issue complements and supplements other forums in biological data analysis and visualization. Many authors who submitted articles to this issue are members-at-large of the research community frequenting the Symposium on Biological Data Visualization (BioVis; [www.biovis.net](http://www.biovis.net)). From 2011 to 2013, BioVis was colocated with the flagship visualization conference, IEEE VIS (formerly IEEE VisWeek). In 2014, BioVis will be colocated with the Conference on Intelligent Systems for Molecular Biology (ISMB), in July in Boston. To further grow the young field of biological data visualization, we urge you to submit pertinent work to BioVis and other related forums.

We thank Associate Editor in Chief Pak Chung Wong and the CG&A staff for their help with editing this special issue. We especially thank the reviewers for their due diligence in completing the reviews. Finally, we thank all the authors who submitted articles. ■■

**Raghu Machiraju** is a professor in the Ohio State University's Departments of Computer Science and Engineering

and Biomedical Informatics. His research interests include visualization, image analysis, and computational biology. Machiraju received a PhD in computer science from the Ohio State University. He was a founding cochair of the Symposium on Biological Data Visualization and served as the general and paper chairs of the IEEE Visualization Conference. He's currently completing a sabbatical at Stanford University. Contact him at [raghu.machiraju@gmail.com](mailto:raghu.machiraju@gmail.com).

**Carsten Görg** is an instructor in the University of Colorado Medical School's Computational Bioscience Program. His research interests include visual analytics and information visualization, especially designing, developing, and evaluating visual-analytics tools for biological and biomedical datasets. Görg received a PhD in computer science from Saarland University. He served on the program and organizing committees of the Symposium on Biological Data Visualization from 2011 to 2013. Contact him at [carsten.goerg@ucdenver.edu](mailto:carsten.goerg@ucdenver.edu).

**Arthur Olson** is the Anderson Research Chair Professor in the Department of Integrative Structural and Computational Biology at the Scripps Research Institute. His research focuses on developing and applying biomolecular modeling and visualization software. He's the director of the HIV Interactions and Viral Evolution (HIVE) Center, which studies mechanisms of viral drug resistance. His recent visualization research employs 3D printing and augmented reality to develop tangible user interfaces for molecular biology for research and education. Olson received a PhD in physical chemistry from the University of California, Berkeley. Contact him at [olson@scripps.edu](mailto:olson@scripps.edu).

**cn** Selected CS articles and columns are also available for free at <http://ComputingNow.computer.org>.



## Call for Articles

**IEEE Software** seeks practical, readable articles that will appeal to experts and nonexperts alike. The magazine aims to deliver reliable information to software developers and managers to help them stay on top of rapid technology change.

**IEEE  
Software**

**Author guidelines:**  
[www.computer.org/software/author.htm](http://www.computer.org/software/author.htm)  
Further details: [software@computer.org](mailto:software@computer.org)  
**[www.computer.org/software](http://www.computer.org/software)**