BioJava 2009: an Open-Source Framework for Bioinformatics

Andreas Prlić (andreas.prlic@gmail.com), University of California San Diego, San Diego, CA, USA
Thomas Down, Wellcome Trust/Cancer Research UK Gurdon Institute UK
Andreas Dräger, Eberhard Karls University Tübingen, Center for Bioinformatics Tübingen (ZBIT) Germany
Sylvain Foisy, Laboratory in Genetics and Genomic Medicine of Inflammation, Montreal Heart Institute, Montreal, Québec, Canada
David Huen, University of Cambridge, UK
Jules Jacobsen, European Bioinformatics Institute, Cambridge, UK
Keith James, Wellcome Trust Sanger Institute, Cambridge, UK
Michael Heuer, Harbinger Partners, Inc., USA
Matthew Pocock, University Newcastle Upon Tyne, UK
Mark Schreiber,Novartis Institute for Tropical Diseases, Singapore
George Waldon, GeneInfinity, USA
Andy Yates, European Bioinformatics Institute, Cambridge, UK
Richard Holland, Eagle Genomics Ltd., Cambridge, UK

The BioJava website is http://biojava.org/.

The version 1.7 release can be downloaded from http://biojava.org/wiki/BioJava:Download

BioJava is a mature open-source project that provides a framework for processing of biological data. BioJava contains powerful analysis and statistical routines, tools for parsing common file formats, and packages for manipulating sequences and 3D structures. It enables rapid bioinformatics application development in the Java programming language. Here we present the latest BioJava release (version 1.7, released in Apr 2009).

Besides numerous bug fixes and stability improvements, a lot of development has been going on in the protein structure modules. BioJava now provides a framework for parsing mmCif files. The parsing of PDB header information has been improved and a new tool to read the Chemical component dictionary is in place.

An ongoing project is the open sourcing of the RCSB protein structure viewers. These have been made available as open source as a stand-alone project. Work is currently under way to provide a more close integration with BioJava.