

[ScienceWatch Home](#)[Inside This Month...](#)[Interviews](#)[Featured Interviews](#)[Author Commentaries](#)[Institutional Interviews](#)[Journal Interviews](#)[Podcasts](#)[Analyses](#)[Featured Analyses](#)[What's Hot In...](#)[Special Topics](#)[Data & Rankings](#)[Sci-Bytes](#)[Fast Breaking Papers](#)[New Hot Papers](#)[Emerging Research Fronts](#)[Fast Moving Fronts](#)[Corporate Research Fronts](#)[Research Front Maps](#)[Current Classics](#)[Top Topics](#)[Rising Stars](#)[New Entrants](#)[Country Profiles](#)[About Science Watch](#)[Methodology](#)[Archives](#)[Contact Us](#)[RSS Feeds](#)

scienceWATCH[®].com

TRACKING TRENDS & PERFORMANCE IN BASIC RESEARCH

[Interviews](#)[Analyses](#)[Data & Rankings](#)

2010 : January 2010 - New Hot Papers : Geoffrey J. Barton Describes the Latest Version of Jalview

new hot papers - 2010

January 2010



Geoffrey J. Barton talks with *ScienceWatch.com* and answers a few questions about this month's New Hot Paper in the field of Computer Science.



Article Title: Jalview Version 2-a multiple sequence alignment editor and analysis workbench

Authors: Waterhouse, AM;Procter, JB;Martin, DMA;Clamp, M;Barton, GJ

Journal: BIOINFORMATICS, Volume: 25, Issue: 9, Page: 1189-1191
Year: MAY 1 2009

* Univ Dundee, Sch Life Sci Res, Coll Life Sci, Dow St, Dundee DD1 5EH, Scotland.

* Univ Dundee, Sch Life Sci Res, Coll Life Sci, Dundee DD1 5EH, Scotland.

* Broad Inst, Cambridge, MA 02142 USA.

SW: Why do you think your paper is highly cited?

The paper describes the latest version of the **Jalview** multiple sequence alignment editor and analysis workbench. Jalview is one of the most powerful tools available for manipulating sequence alignments and integrating annotations from biological databases around the world. As a consequence, many thousands of scientists make use of Jalview in their daily work.

The Jalview software is installed on over 20,000 computers worldwide and is also available as an applet that is installed on over 100,000 web pages including those run by major international databases of sequence alignments such as **Pfam**. Analysis of alignments by Jalview is often of key importance in a scientific publication and so this leads to citations of the paper describing Jalview.

Sequences of DNA, RNA, and proteins are the fundamental currency of modern biological and medical research. Sequences link the different levels of the biological hierarchy, from gene to three-dimensional structure.

Multiple Sequence Alignments (MSAs) arrange sequences that are similar as a table that highlights which amino acids or nucleotides are common across all sequences. For proteins, MSAs permit the identification of common features between species or identify functionally important amino acids. MSAs provide the

basis for a spectrum of computational methods, including the prediction of protein secondary structure and solvent accessibility, functional sites, and interaction sites. MSAs are also the essential first step in studying molecular evolution and are core to the identification of genomic rearrangements.

In journal publications, MSAs provide a convenient way to display common features and complex annotations relating to sequences and their functions. Although there are many programs that generate multiple alignments from unaligned sequences, none give a perfect result in all circumstances. Jalview provides a convenient way to generate alignments by a variety of methods and then to edit them to correct errors or choose the most informative subsets of sequences.

SW: Does it describe a new discovery, methodology, or synthesis of knowledge?

The paper describes significant updates to the Jalview system. Updates include more sophisticated editing functions and new visualization methods, the ability to recall computer-intensive alignment and analysis methods on remote servers from within the program and provide access to over 50 different types of annotation provided by computer servers worldwide. These new enhancements, together with updates to core file management and format conversions, have made the program useful to a larger number of potential users.

SW: Would you summarize the significance of your paper in layman's terms?

Medical and biological research is producing enormous quantities of data about the DNA and protein molecules that make up all living organisms and are central to understanding function and disease. While generating data is getting easier and cheaper, the volume of data presents big problems for scientists to visualize, edit, and analyze.

This paper describes a powerful tool for visualizing, aligning, and analyzing very large sets of DNA and protein sequences. It is effectively a specialized word processor, web-browser and desktop publishing package for sequences rolled into one.

The Jalview software described in this paper makes it easier to carry out common analyses on biological sequences, but most importantly makes possible analyses that would otherwise be too difficult or impossible to do.

SW: How did you become involved in this research, and were there any problems along the way?

By the mid-1990s we had worked for more than 10 years on the generation and analysis of protein multiple sequence alignments ([view](#)). Visualization of alignments was always a problem, so we first developed a program ALSCRIPT (1) that allowed flexible annotation of a static alignment: ([view](#)).

However, to work more efficiently, we required an interactive tool to edit alignments and display the results of some of our other techniques such as AMAS (2) sub-family analyses ([view](#)), and JPred (3) secondary structure predictions ([view](#)).

Jalview was developed as a successor to ALSCRIPT, although ALSCRIPT still has a strong following. The main problem we encountered was the continuation of funding for Jalview. However, recent

"Sequence analysis is central to all modern biological research, whether in agriculture, biotechnology, or the study and treatment of human disease. Jalview is in use daily by scientists working in all these fields and, since it makes it possible for them to work more efficiently, has direct impact on the many social and political issues that their research influences."

initiatives by the UK Biotechnology and Biological Sciences Research Council (BBSRC) have been friendly to the project. Thanks to BBSRC, core funding is now secure until 2014.

SW: Where do you see your research leading in the future?

Next-generation sequencing technology has become available over the last three years and has led to an explosion in the volume of sequence data available. This volume of data presents significant challenges for alignment visualization and analysis. Already, some protein families have over 100,000 members and this will be the norm for most families of proteins within the next five years.

Accordingly, we will be developing Jalview to work efficiently with such large sequence families as well as longer sequences such as complete genomes. We will also be making it easier for other scientists to add new features to the program that are specific to their needs.

SW: Do you foresee any social or political implications for your research?

Sequence analysis is central to all modern biological research, whether in agriculture, biotechnology, or the study and treatment of human disease. Jalview is in use daily by scientists working in all these fields and, since it makes it possible for them to work more efficiently, has direct impact on the many social and political issues that their research influences.

Geoff Barton, Ph.D.

The Barton Group

Professor of Bioinformatics

College of Life Sciences

University of Dundee

Scotland, UK

Web

References:

1. Barton, GJ, "ALSCRIPT - A Tool to Format Multiple Sequence Alignments," *Prot. Eng.* 6, 37-40, 1993.
2. Livingstone, CD and Barton, G, "Protein Sequence Alignments: A Strategy for the Hierarchical Analysis of Residue Conservation", *Comp. Appl. Bio. Sci.* 9, 745-56, 1993.
3. Cole, C., Barber, JD and Barton, GJ, "The JPRED 3 Secondary structure prediction server," *Nucleic Acids Research*, doi: 10.1093/nar/gkn238, 2008.

KEYWORDS: SECONDARY STRUCTURE; STRUCTURAL BIOLOGY; PROTEIN SEQUENCES; ACCURACY; TOOLS; SYSTEM.

 PDF

2010 : January 2010 - New Hot Papers : Geoffrey J. Barton Describes the Latest Version of Jalview