Application of VAMSAS enabled tools for the investigation of protein evolution

J. B. Procter, I. Milne, F. Wright, P. Marguerite, A. M. Waterhouse, D. Lindner, D.M.A. Martin, T. Oldfield, D. Marshall, G. J. Barton.

Analysis of protein evolution involves the application of a combination of methods: nucleic acid and protein alignment, phylogenetic modelling, tree building, and functional annotation analysis. Whilst there are many interactive applications capable of applying one or more of these approaches, very few efficiently enable the user to perform all aspects of such a study. We present an example that demonstrates how this kind of analysis can be performed using applications that have been modified to dynamically exchange data; *via* the 'Visualization and Analysis of Molecular Sequences, Alignments, and Structures' (VAMSAS) framework.

The VAMSAS framework is a data exchange and interoperation framework for bioinformatics applications. It enables them to share sequences, alignments, phylogenetic trees and annotation, and rapidly exchange messages so that each one's visualization of the shared data can be synchronised. For example, messages are exchanged to indicate current display state such as the position of the mouse pointer within an amino acid sequence, or a newly selected area in an alignment. Shared data may also be exported, to provide a complete record of the current state of an analysis that can be exchanged with other researchers.

An LGPL licensed implementation of the VAMSAS framework was developed in Java to enable interoperation between three popular graphical applications: Jalview[1] is a widely used workbench for multiple sequence alignment, visualization, editing and analysis developed at the University of Dundee. TOPALi[2] is a program for evolutionary analysis of protein and nucleotide sequence alignments and phylogenetic tree construction, developed by Biomathematics and Statistics Scotland, and the Scottish Crop Research Institute. AV@MSD-EBI[3, 4] enables the visualization and analysis of protein and nucleic acid structures, and interfaces to web services provided by the Macromolecular Structure Database (MSD) at the European Bioinformatics Institute (EBI). The VAMSAS web site (www.vamsas.ac.uk) provides more information about the library, and links to the different programs.

1. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ: **Jalview Version 2 - a multiple sequence alignment editor and analysis workbench**. *Bioinformatics* 2009, doi:10.1093/bioinformatics/btp033.
2. Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, Marshall DF, Wright F: **TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops**. *Bioinformatics* 2009, **25**(1):126-127.
3. Hartshorn MJ: **AstexViewer: a visualisation aid for structure-based drug design**. *J Comput Aided Mol Des* 2002, **16**(12):871-881.
4. Oldfield TJ: **A Java applet for multiple linked visualization of protein structure and sequence**. *J Comput Aided Mol Des* 2004, **18**(4):225-234.