

TreeDomViewer: A tool for the visualization of phylogeny and protein domain structure

Blaise T.F. Alako^{1,2}, Daphne Rainey³, Harm Nijveen¹, and Jack A.M. Leunissen¹

1. *Laboratory of Bioinformatics, Wageningen University and Research Centre, PO Box 8128, 6700 ET Wageningen, The Netherlands*
2. *Centre for BioSystems Genomics, PO Box 98, 6700 AB Wageningen, The Netherlands*
3. *KEYGENE NV, PO Box 216 6700 AE Wageningen, The Netherlands*

Abstract

Phylogenetic analyses and examination of protein domains are critical to comparative genomic studies. These techniques allow precise genome annotation and are invaluable to studies of protein and protein complex evolution. It is worth noticing that two sequences may be homologous without sharing statistically significant amino acid or nucleotide identity. Recognizing this type of homology is a particular challenging bioinformatics problem.

We present here TreeDomViewer, a visualization tool available as a PERL web-based interface that analyzes the phylogenetic tree description, multiple sequence alignment, and InterProScan results of the sequences, and generates a phylogenetic tree with the corresponding protein domain information projected onto the multiple sequence alignment. TreeDomViewer produces an interactive scalar vector graphic (SVG) image; this allows the user to see the orthological relationships and domain content of the proteins of interest at one glance. Alternatively, the image can also be obtained in PDF, JPEG or PNG. In our description TreeDomViewer adopts an evolutionary perspective of how domain structure in two or more sequences can be aligned and compared to subsequently infer the function of an unknown homolog. This provides insight into the analysis of, in terms of amino acid substitution very divergent but yet closely related family members.

TreeDomViewer is a flexible phylogenetic tree and domain structure visualization tool that makes use of outstanding existing domain prediction tools such as InterProScan and EMBOSS package for protein or gene relatedness inference. This makes it a beneficial addition to the annotation pipeline of unknown gene or gene product. An important aspect of TreeDomViewer is that domain structures are aligned, thereby allowing visualization of structurally conserved domains throughout evolution. TreeDomViewer is freely available for academic use.

Application area: Sequence analysis

Platform: Independent, Web-based

Availability: <http://www.bioinformatics.nl/tools/treedom/>

Contact: jack.leunissen@wur.nl, blaise.alako@wur.nl