

Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling

Submitted by Emmanuel Lemoine on Tue, 02/24/2015 - 15:24

Titre Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling

Type de publication Article de revue

Auteur Pele, J. [1], Becu, J. M [2], Abdi, H. [3], Chabbert, Marie [4]

Editeur BioMed Central

Type Article scientifique dans une revue   comit  de lecture

Ann e 2012

Langue Anglais

Date 2012

Volume 13

Titre de la revue BMC Bioinformatics

ISSN 1471-2105

Mots-cl s Algorithms [5], Animals [6], Cluster analysis [7], Humans [8], Multivariate Analysis [9], Phylogeny [10], Receptors, G-Protein-Coupled/chemistry/genetics [11], Sea Anemones/chemistry [12], Sequence Alignment/methods [13], Sequence Analysis, Protein/methods [14], Software [15]

R sum  en anglais
BACKGROUND: The distance matrix computed from multiple alignments of homologous sequences is widely used by distance-based phylogenetic methods to provide information on the evolution of protein families. This matrix can also be visualized in a low dimensional space by metric multidimensional scaling (MDS). Applied to protein families, MDS provides information complementary to the information derived from tree-based methods. Moreover, MDS gives a unique opportunity to compare orthologous sequence sets because it can add supplementary elements to a reference space. RESULTS: The R package bios2mds (from BIOlogical Sequences to MultiDimensional Scaling) has been designed to analyze multiple sequence alignments by MDS. Bios2mds starts with a sequence alignment, builds a matrix of distances between the aligned sequences, and represents this matrix by MDS to visualize a sequence space. This package also offers the possibility of performing K-means clustering in the MDS derived sequence space. Most importantly, bios2mds includes a function that projects supplementary elements (a.k.a. "out of sample" elements) onto the space defined by reference or "active" elements. Orthologous sequence sets can thus be compared in a straightforward way. The data analysis and visualization tools have been specifically designed for an easy monitoring of the evolutionary drift of protein sub-families. CONCLUSIONS: The bios2mds package provides the tools for a complete integrated pipeline aimed at the MDS analysis of multiple sets of orthologous sequences in the R statistical environment. In addition, as the analysis can be carried out from user provided matrices, the projection function can be widely used on any kind of data.

URL de la notice <http://okina.univ-angers.fr/publications/ua8181> [16]

DOI 10.1186/1471-2105-13-133 [17]
Lien vers le document <http://dx.doi.org/10.1186/1471-2105-13-133> [17]
Titre abrégé BMC Bioinformatics

Liens

- [1] [http://okina.univ-angers.fr/publications?f\[author\]=13612](http://okina.univ-angers.fr/publications?f[author]=13612)
- [2] [http://okina.univ-angers.fr/publications?f\[author\]=13613](http://okina.univ-angers.fr/publications?f[author]=13613)
- [3] [http://okina.univ-angers.fr/publications?f\[author\]=13614](http://okina.univ-angers.fr/publications?f[author]=13614)
- [4] <http://okina.univ-angers.fr/marie.chabbert/publications>
- [5] [http://okina.univ-angers.fr/publications?f\[keyword\]=7854](http://okina.univ-angers.fr/publications?f[keyword]=7854)
- [6] [http://okina.univ-angers.fr/publications?f\[keyword\]=964](http://okina.univ-angers.fr/publications?f[keyword]=964)
- [7] [http://okina.univ-angers.fr/publications?f\[keyword\]=11783](http://okina.univ-angers.fr/publications?f[keyword]=11783)
- [8] [http://okina.univ-angers.fr/publications?f\[keyword\]=991](http://okina.univ-angers.fr/publications?f[keyword]=991)
- [9] [http://okina.univ-angers.fr/publications?f\[keyword\]=8053](http://okina.univ-angers.fr/publications?f[keyword]=8053)
- [10] [http://okina.univ-angers.fr/publications?f\[keyword\]=8736](http://okina.univ-angers.fr/publications?f[keyword]=8736)
- [11] [http://okina.univ-angers.fr/publications?f\[keyword\]=12881](http://okina.univ-angers.fr/publications?f[keyword]=12881)
- [12] [http://okina.univ-angers.fr/publications?f\[keyword\]=12882](http://okina.univ-angers.fr/publications?f[keyword]=12882)
- [13] [http://okina.univ-angers.fr/publications?f\[keyword\]=12883](http://okina.univ-angers.fr/publications?f[keyword]=12883)
- [14] [http://okina.univ-angers.fr/publications?f\[keyword\]=12884](http://okina.univ-angers.fr/publications?f[keyword]=12884)
- [15] [http://okina.univ-angers.fr/publications?f\[keyword\]=311](http://okina.univ-angers.fr/publications?f[keyword]=311)
- [16] <http://okina.univ-angers.fr/publications/ua8181>
- [17] <http://dx.doi.org/10.1186/1471-2105-13-133>

Publié sur *Okina* (<http://okina.univ-angers.fr>)