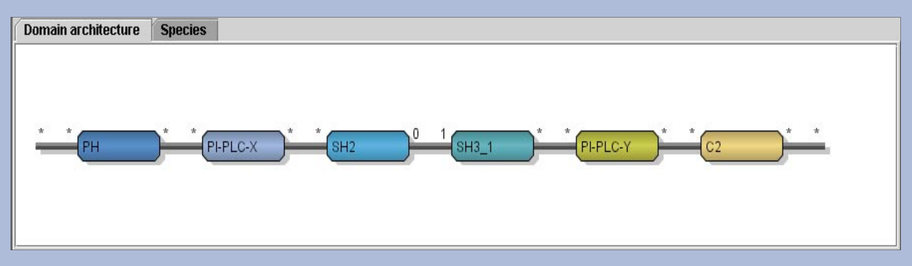
Pfam::Alyzer

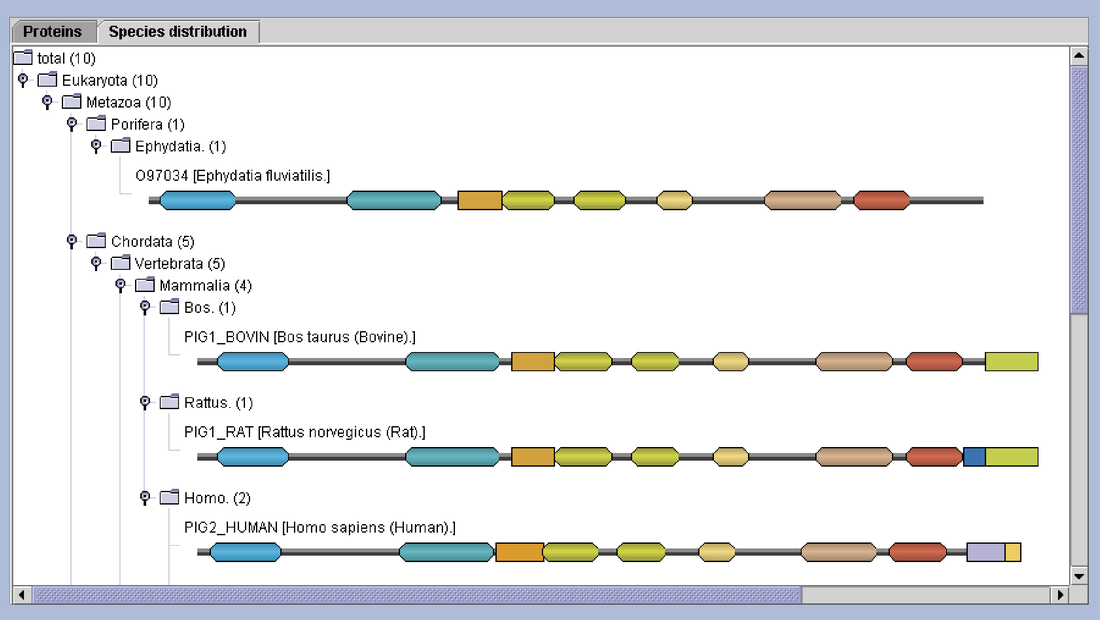
Volker Hollich and Erik L. L. Sonnhammer
Center for Genomics and Bioinformatics
Karolinska Institutet, Stockholm

PfamAlyzer paves way for an augmented understanding of protein domain architectures. It utilises the wealth of protein domain families present in the Pfam database and enables easy-to-use access to complex domain architectures as well as providing means for the study of their evolution.

Domain query

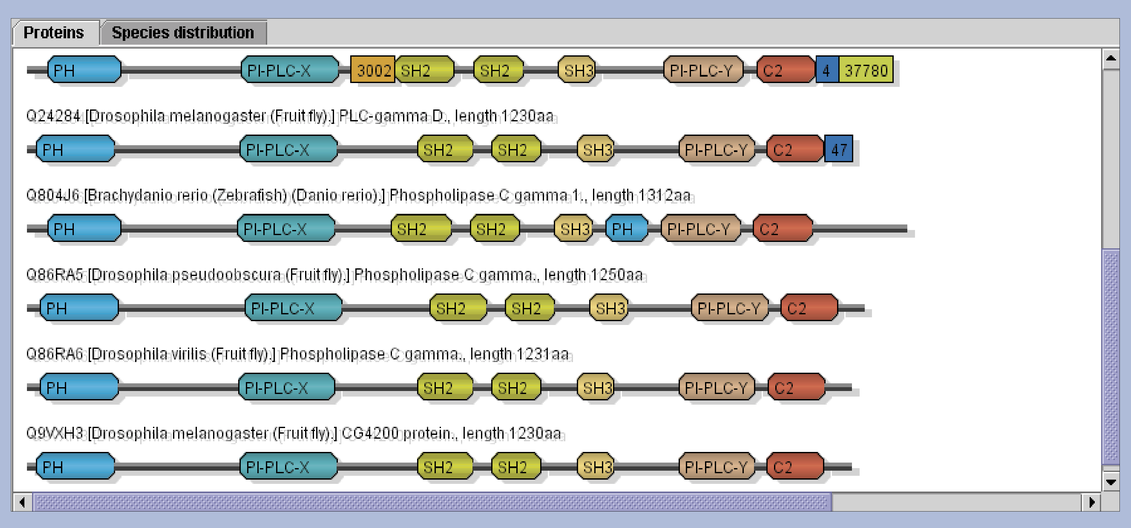
PfamAlyzer enables comprehensive study of domain architectures. Using a powerful, user-friendly graphical query language, arbitrary domains may be freely combined on the screen. PfamAlyzer uses Drag & Drop functionality as an intuitive approach for query formulation. Specific domains can be excluded from the search, in addition tolerated gaps between domains may be specified as domains or amino acids.





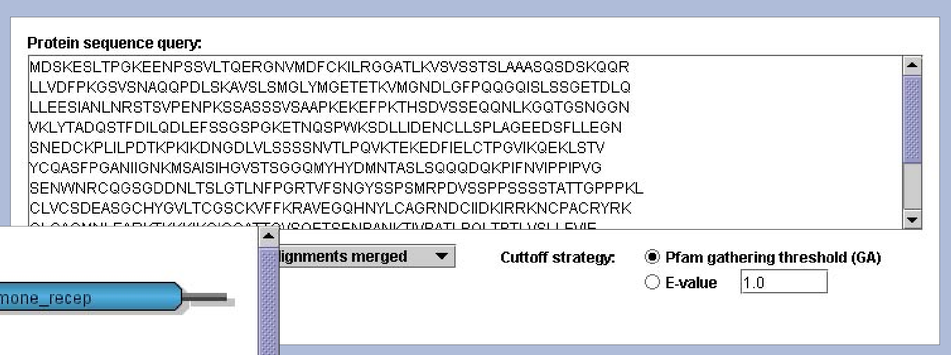
Optionally, the search can be limited to taxonomic groups, e.g., chordata. The proteins of SwissProt/TrEMBL with predicted domain architecture according to Pfam are queried for the selected architecture. The outcome is presented either in a list-fashion or as a species distribution. Here, the proteins are shown as leaves on the species tree. Subtrees may be in-/excluded from visualisation, which simplifies analysis of larges result sets.

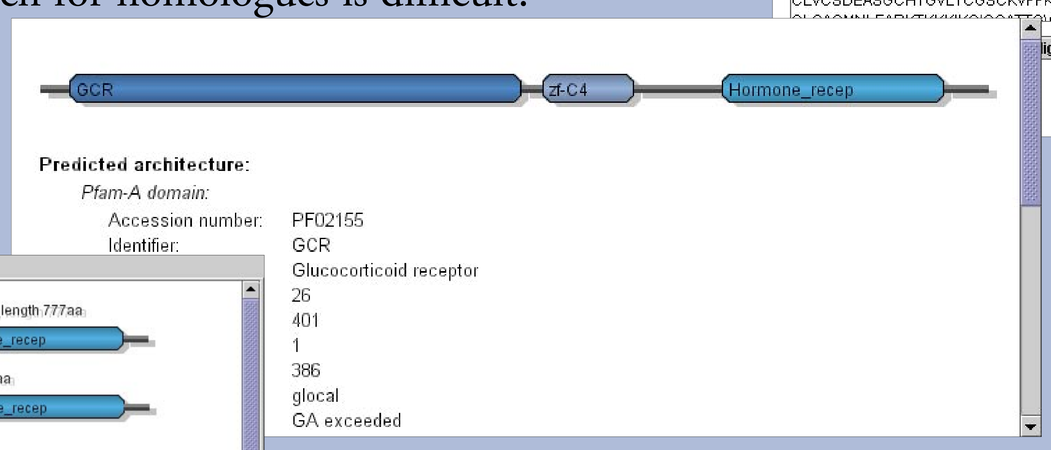
The manually curated Pfam-A domains are displayed as ovals, whereas automatically curated Pfam-B domains are shown as rectangles. Additional information for Pfam-A domains is available through context menus as well as links to other resources.

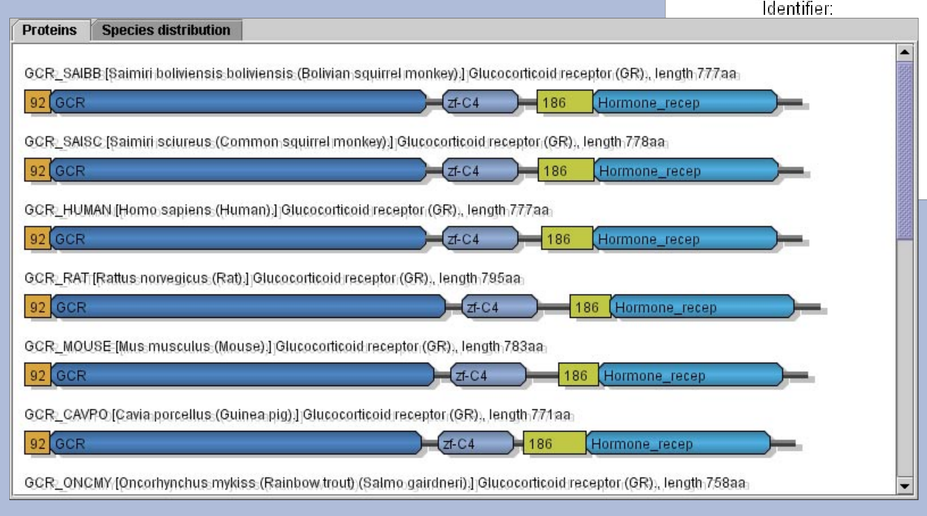


Homology seach

PfamAlyzer provides innovative homology searching. So far, the identification of homologues has been conducted by direct sequence comparison. This is appropriate if sequences are not very distant. However, for distantly related proteins the search for homologues is difficult. Domain querying allows a new tool for finding such homologues.







PfamAlyzer introduces a domain-centric homology search based on Pfam domains. In a first step, Pfam domains are predicted from the sequence. This is achieved using the hmmer package which is also applied for generating Pfam full alignments from curated seeds. Subsequently, SwissProt/TrEMBL is queried for the obtained domain architecture. This type of domain-centric homology analysis provides a means to identify relationships that are distant yet possess a conserved domain architecture.

http://pfam.cgb.ki.se/pfamalyzer