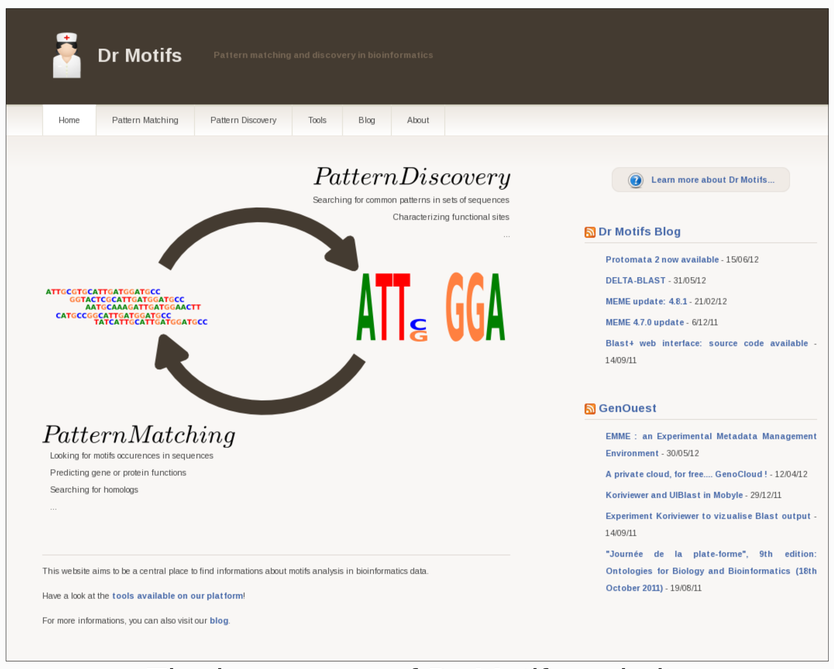
Dr Motifs: a web resource for pattern matching and discovery

Anthony Bretaudeau and Olivier Collin

Pattern discovery and pattern matching have become a very common task when analysing biological sequences. However, the diversity of algorithms and implementations to achieve these analyses make it difficult for users to choose the right tool. Dr Motifs is a web portal designed to ease analysis of motifs.

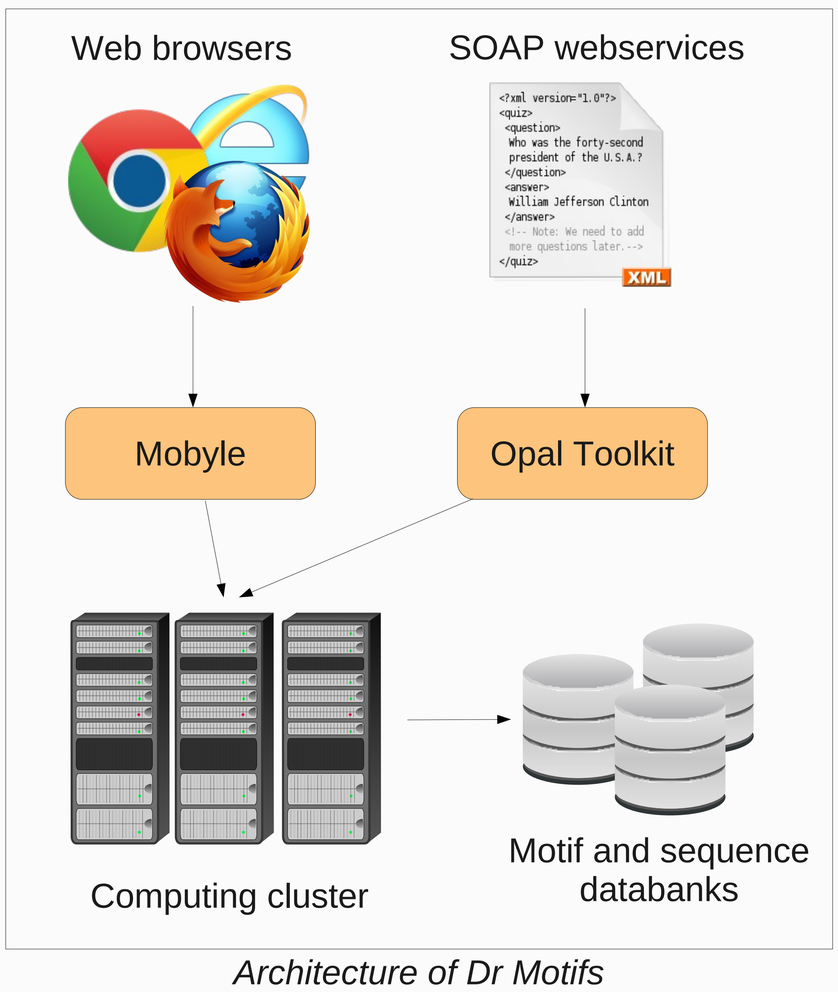


The homepage of Dr Motifs website

www.drmotifs.org

Architecture

Dr Motifs website provides a convenient access to motif analysis tools via web interfaces and web-services. Analyses are processed on a computing cluster hosted by the GenOuest platform, with access to a rich collection of motif and sequence databanks (e.g. Pfam, GenBank). These databanks are kept up-to-date thanks to BioMAJ (http://biomaj.genouest.org).



Architecture of Dr Motifs

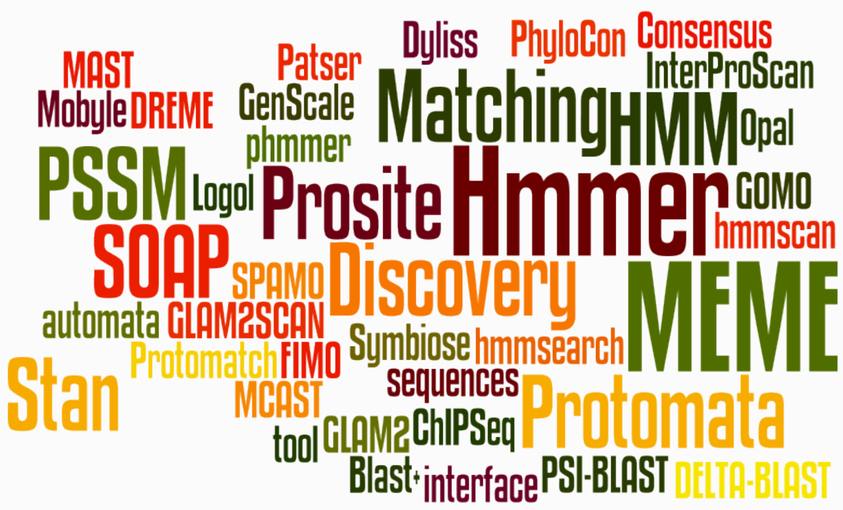
Tools access

HMMER	jackhmmer	Iteratively search protein sequence(s) against a protein database	Pattern matching (psi-blastp-like)	HMM		✓	3.0			Tutorial
	hmmconvert	Convert profile HMM file to a HMMER format	Other			✓				
	hmmalign	Align sequences to a profile HMM	Other			✓				
	hmmemit	Generate sequences from a profile HMM	Other			✓				
	hmmfetch	Retrieve an HMM from an HMM database	Other			✓				
	hmmsim	Collect profile HMM score distributions on random sequences	Other			✓				
	hmmstat	Display summary statistics for a profile file	Other			✓				
InterProScan		In a given sequence, searches for known motifs from databanks like Jaspar.	Pattern matching (for annotation)	HMM, PSSM, Prosite	✓	✓	4.7			Tutorial
Protomata	Protomata-learner	Build a motif (protomaton) representing a set of proteic sequences	Pattern discovery	Protomaton		✓	2.0			Tutorial
	Protomatch	Search new sequences containing a given protomata motif	Pattern matching			✓				
Stan		Search for a given pattern in sequences. The syntax allows insertion, deletions and string variables.	Pattern matching	SVG syntax (Improved Prosite)	✓	✓	-			
	Patser	Scan a DNA sequence with a position-specific scoring matrix (PSSM)	Pattern matching		✓	✓	3e			

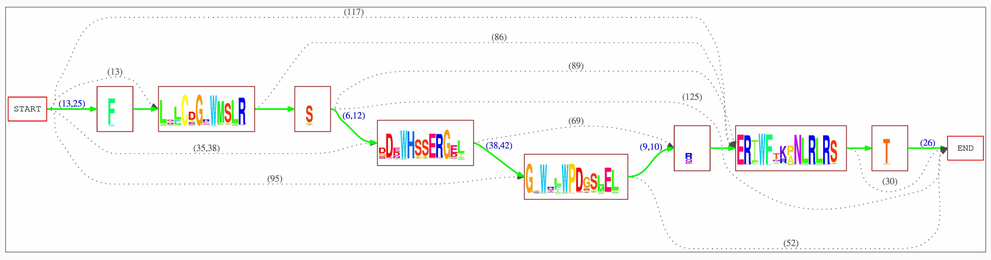
Synthetic list of available tools on Dr Motifs website

A synthetic listing of 37 tools is available, displaying the specificities of each one. One-click access to web interfaces, web-services and tutorials are provided on this page. Tutorials, documentation and use-cases are also available to guide users in their experiments.
Web interfaces and web-services can be used for both small and large scale analyses, with a tight integration of the different tools.

Available tools



Using Dr Motifs, you get access to the best known motif analysis software like Hmmer, MEME or InterProScan.
But more specialized tools like Protomata, Logol or Stan which are developed by the Inria/IRISA Dyliss and Genscale teams (Rennes, France) are also available. These tools allow access to other motifs representations with better expressivity.



An example of a protein motif produced by the Protomata suite

Plate-forme GenOuest
INRIA Rennes Bretagne Atlantique
Campus de Beaulieu
35042 Rennes Cedex
support@genouest.org