

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

A simple modular architecture research tool for the identification of signaling domains

ArticleInfo		
ArticleID	:	3608
ArticleDOI	:	10.1186/gb-2000-1-1-reports234
ArticleCitationID	:	reports234
ArticleSequenceNumber	:	99
ArticleCategory	:	Web report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 2000-1-10 Received : 2000-1-10 OnlineDate : 2000-3-17
ArticleCopyright	:	BioMed Central Ltd2000
ArticleGrants	:	
ArticleContext	:	130591111

Todd Richmond

Abstract

SMART compares query sequences with its databases of domain sequences and multiple alignments while concurrently identifying compositionally biased regions such as signal peptide, transmembrane and coiled-coil segments.

Content

The abstract describing SMART says it best: "SMART compares query sequences with its databases of domain sequences and multiple alignments while concurrently identifying compositionally biased regions such as signal peptide, transmembrane and coiled-coil segments". Annotated and unannotated regions of the sequence can be used as queries in searches of sequence databases. The SMART alignment collection represents more than 250 signaling and extracellular domains. Each alignment is curated to assign appropriate domain boundaries and to ensure its quality. The website contains more than this, however. Each of the domains is carefully annotated, containing information about the proposed role of the domain, a taxonomic breakdown of where the domain is found, links to sequences, three-dimensional structures, primary literature and other databases. In addition, once you have determined the structural organization of the domains in your protein, you can use the SMART website to search for other proteins with the same domains, or use the 'SMART alert' to notify you when similar proteins are deposited in the database.

Navigation

As with most domain searches, the navigation is straightforward. Paste your sequence into the webform, or supply a sequence identifier (ID)/accession number (ACC), and press 'Search'. The only choices are which databases to search and an option to look for signal sequences. The output includes a graphic showing the putative domains and their location in the protein, as well as a table containing the domain name, sequence positions and the expected value for each prediction. The domains in both the graphic and the table are linked to another page with the full description of the domains. This second page also contains additional options to align your sequence in the region to the SMART alignment, view the complete family alignment or submit a BLAST search with the domain sequence. There are also two additional links on the first page: one to search immediately for proteins with a similar domain

organization, and the other to search for proteins with a similar domain composition. For an architecture analysis query, you type in the domains you want to look for (for example, LRR and S_TKc, to find all proteins with leucine-rich repeats and a serine/threonine protein kinase domain) and press 'Search'. To narrow searches, you can limit the search to certain species or taxa. The output consists of a list of proteins, with an accession number (linked to [TrEMBL](#), a computer-annotated supplement to SWISS-PROT), a description and the species. By clicking on a checkbox to the left of each sequence, you can choose a graphic view of the domain structures of proteins that you select. You can also enter the successful search into the SMART alert system to notify you when new proteins that match the same conditions are entered in the database.

Reporter's comments

Timeliness

There is no indication of when the site was last updated.

Best feature

The graphical output, with the underlying HTML links, is a very simple yet powerful method for representing what can be a tremendous amount of information. All the information necessary to determine what each domain does is at the user's fingertips.

Worst feature

It is limited to signaling and extracellular domains and will not find other important domains, including many domains associated with enzymatic functions.

Wish list

It would be more useful if the number of domains represented in the SMART database could be expanded to include more than just signaling and extracellular domains. An option for Postscript output would also be helpful. The GIF output is fine for web pages but it would not take too much effort to make a Postscript version, which would be useful for making figures.

Related websites

Sites for determining protein domains and motifs in proteins include [The protein domain database ProDom](#), [The Pfam database of protein domains and HMMs](#), [EMOTIF](#), [Blocks](#) and [The protein motif fingerprint database PRINTS](#).

Table of links

[SMART: simple modular architecture research tool](#)

[The protein domain database ProDom](#)

[The Pfam database of protein domains and HMMs](#)

[EMOTIF](#)

[Blocks](#)

[The protein motif fingerprint database PRINTS](#)

[TrEMBL](#)

References

1. [SMART: simple modular architecture research tool](#).