

RSATWSPortType

Service Documentation

Web services for the Regulatory Sequence Analysis Tools (RSAT). Tools developed by Jacques van Helden (jvanheld@scmbb.ulb.ac.be), SOAP/WSDL interface developed by Olivier Sand (oly@scmbb.ulb.ac.be).

Server Address

http://rsat.scmbb.ulb.ac.be/rsat/web_services/RSATWS.cgi

Method

retrieve_seq

Description

Returns upstream, downstream or coding DNA sequences for list of query genes.

Parameters

Input Parameters

- | | |
|----------------------------|---|
| output | Return type. Accepted values: 'server' (result is stored on a file on the server), 'client' (result is directly transferred to the client), or 'both'. Default is 'both'
Type = string |
| organism (required) | Organism. Words need to be underscore separated (example: Escherichia_coli_K12).
Type = string |
| query (required) | A list of query genes
Type = string |
| noorf | Prevent overlap with upstream open reading frames (ORF).
Type = boolean |
| from | Inferior limit of the region to retrieve. Default is organism dependant (example: Saccharomyces cerevisiae = -800).
Type = int |
| to | Superior limit of the region to retrieve. Default is '-1'.
Type = int |
| featype | Type of genome features to load. Supported: CDS, mRNA, tRNA, rRNA.
Type = string |

type	Sequence type. Supported: upstream, downstream, ORF (unspliced open reading frame). Type = string
format	Sequence format. Supported: IG (Intelligenetics), WC (wconsensus), raw, FastA Type = string
all	Return sequences for all the genes of the organism. Type = boolean
lw	Line width (0 for whole sequence on one line). Type = int
label	Field(s) to be used in the sequence label. Multiple fields can be specified, separated by commas. Supported: id, name, organism_name, sequence_type, current_from, current_to, ctg, orf_strand, reg_left, reg_right. Default: name. Type = string
label_sep	Separator between the label fields. Default: (pipe character). Type = string
nocom	No comments. Only the identifier and the sequence are returned. By default, the comment indicates the ORF and upstream sequence coordinates. Type = boolean
repeat	Use the repeat masked version of the genome. Attention: repeated regions are annotated for some genomes only. Type = boolean
imp_pos	Admit imprecise positions. Type = boolean

Output Parameters

server	Location of the result file on the server. This can be used as input for a further request. Type = string
---------------	--

command The stand alone command executed on the server.
Type = string

client The result sequence(s).
Type = string

Method**purge_seq****Description**

Mask repeated fragments of an input sequence.

Parameters**Input Parameters**

output Return type. Accepted values: 'server' (result is stored on a file on the server), 'client' (result is directly transferred to the client), or 'both'. Default is 'both'.
Type = string

sequence Sequence to purge. You need to supply either this parameter or the next one (tmp_infile).
Type = string

tmp_infile Name of the file with input sequence on the server. You need to supply either this parameter or the previous one (sequence).
Type = string

format Sequence format. Supported: IG (Intelligenetics), WC (wconsensus), raw, FastA.
Type = string

match_length Minimal match length. Default is 40.
Type = int

mismatch Number of mismatches allowed. Default is 3.
Type = int

str Discard duplications on the direct strand only (1str) or on the reverse complement as well (2str). Default is 2str.
Type = string

- delete** Delete repeats instead of masking them.
Type = boolean
- mask_short** Mask (replace by N characters) sequences shorter than the specified length.
Type = int

Output Parameters

- server** Location of the result file on the server. This can be used as input for a further request.
Type = string
- command** The stand alone command executed on the server.
Type = string
- client** The result sequence(s).
Type = string

Method

Description

oligo_analysis

Analysis of the statistical significance of all the oligomers of a given size in a sequence. Commonly used to detect over-represented oligonucleotides in a set of promoter sequences.

Parameters

Input Parameters

- output** Return type. Accepted values: 'server' (result is stored on a file on the server), 'client' (result is directly transferred to the client), or 'both'. Default is 'both'.
Type = string
- sequence** Input sequence. You need to supply either this parameter or the next one (tmp_infile).
Type = string
- tmp_infile** Name of the file with input sequence on the server. You need to supply either this parameter or the previous one (sequence).
Type = string

format	Input sequence format. Supported: IG (Intelligenetics), WC (wconsensus), raw, fasta. Default is fasta Type = string
length (required)	Oligomer length. Type = int
organism (required)	Organism. Words need to be underscore separated (example: Escherichia_coli_K12). Type = string
background	Background model: Type of sequences used as background model for estimating expected oligonucleotide frequencies. Supported: upstream, upstreamL, upstream-noorf, intergenic, input Type = string
stats	List of statistics to return. Supported: occ, mseq, freq, proba, ratio, zscore, like, pos, rank Type = string
noov	No overlapping. Disable the detection of overlapping matches for self-overlapping patterns (ex TATATA, GATAGA). Type = boolean
str	Oligonucleotide occurrences found on both stands are summed (2str) or not (1str). Default is 2str. Type = string
sort	Sort oligomers according to overrepresentation. Type = boolean
lth	Lower threshold on some parameter. Format='param value' Type = string

Output Parameters

server	Location of the result file on the server. This can be used as input
---------------	---

	<p>for a further request. Type = string</p> <p>command The stand alone command executed on the server. Type = string</p> <p>client The result discovered pattern(s). Type = string</p>
Method	gene_info
Description	Get information about genes.
Parameters	<p>Input Parameters</p> <p>organism (required) Organism. Words need to be underscore separated (example: Escherichia_coli_K12). Type = string</p> <p>query (required) List of gene(s) for which you want info on. Type = string</p> <p>full Full match only (no substring matching). Type = boolean</p> <p>noquery Do not print the query at the begining of each line. Type = boolean</p> <p>descr Match query against the description, too, not just against gene ID and name. Type = boolean</p> <p>featype Feature type (CDS, mRNA, tRNA, rRNA, scRNA). Type = string</p> <p>Output Parameters</p> <p>command The stand alone command executed on the server. Type = string</p> <p>client The result gene(s) info(s). Type = string</p>