

## 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](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

<b>type</b>	<b>Sequence type. Supported: upstream, downstream, ORF (unspliced open reading frame). Type = string</b>
<b>format</b>	<b>Sequence format. Supported: IG (Intelligenetics), WC (wconsensus), raw, FastA Type = string</b>
<b>all</b>	<b>Return sequences for all the genes of the organism. Type = boolean</b>
<b>lw</b>	<b>Line width (0 for whole sequence on one line). Type = int</b>
<b>label</b>	<b>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</b>
<b>label_sep</b>	<b>Separator between the label fields. Default:   (pipe character). Type = string</b>
<b>nocom</b>	<b>No comments. Only the identifier and the sequence are returned. By default, the comment indicates the ORF and upstream sequence coordinates. Type = boolean</b>
<b>repeat</b>	<b>Use the repeat masked version of the genome. Attention: repeated regions are annotated for some genomes only. Type = boolean</b>
<b>imp_pos</b>	<b>Admit imprecise positions. Type = boolean</b>

### **Output Parameters**

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

- 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

### oligo\_analysis

#### Description

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

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

### Output Parameters

<b>server</b>	<b>Location of the result file on the server. This can be used as input</b>
---------------	---

for a further request.

Type = string

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

**client** The result discovered pattern(s).  
Type = string

## Method

### gene\_info

#### Description

Get information about genes.

#### Parameters

#### Input Parameters

**organism (required)** Organism. Words need to be underscore separated (example: Escherichia\_coli\_K12).  
Type = string

**query (required)** List of gene(s) for which you want info on.  
Type = string

**full** Full match only (no substring matching).  
Type = boolean

**noquery** Do not print the query at the begining of each line.  
Type = boolean

**descr** Match query against the description, too, not just against gene ID and name.  
Type = boolean

**featype** Feature type (CDS, mRNA, tRNA, rRNA, scRNA).  
Type = string

#### Output Parameters

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

**client** The result gene(s) info(s).  
Type = string