

Annotate with DAS Element

Finds similar protein sequence using remote BLAST. Using IDs of sequences found loads annotation for DAS sources. Nucleotide sequences are skipped if any supplied to input.

Parameters in GUI

Parameter	Description	Default value
Max result IDs	Use first IDs of similar sequences to load annotations.	5
Database	Database against which the search is performed: UniProtKB or clusters of sequences with 100%, 90% or 50% identity.	UniProtKB
Min identity	Minimum identity of a BLAST result and an input sequence.	90%
Threshold	The expectation value (E) threshold is a statistical measure of the number of expected matches in a random database. The lower the e-value, the more likely the match is to be significant.	10
Matrix	The matrix assigns a probability score for each position in an alignment.	Auto
Filtering	Low-complexity regions (e.g. stretches of cysteine in Q03751, or hydrophobic regions in membrane proteins) tend to produce spurious, insignificant matches with sequences in the database which have the same kind of low-complexity regions, but are unrelated biologically. If 'Filter low complexity regions' is selected, the query sequence will be run through the program SEG, and all amino acids in low-complexity regions will be replaced by X's.	None
Gapped	This will allow gaps to be introduced in the sequences when the comparison is done.	true
Hits	Limits the number of returned alignments.	250
Feature sources	The DAS sources to read features from.	InterPro-Matches-Overview, Pride DAS 1.6, UniProt, cbs_sort, signalp

Parameters in Workflow File

Type: dasannotation-search

Parameter	Parameter in the GUI	Type
idsnumber	Max result IDs	<i>numeric</i>
db	Database	<i>string</i>
identity	Min identity	<i>numeric</i>
threshold	Threshold	<i>string</i>
matrix	Matrix	<i>string</i>
filtering	Filtering	<i>string</i>
gapped	Gapped	<i>string</i>
maxres	Hits	<i>string</i>
fsources	Feature sources	<i>string</i>

Input/Output Ports

The element has 1 *input port*:

Name in GUI: Input sequences

Name in Workflow File: in-sequence

Slots:

Slot In GUI	Slot in Workflow File	Type
Sequence	sequence	<i>string</i>

The element has 1 *output port*:

Name in GUI: DAS annotations

Name in Workflow File: out-annotations

Slots:

Slot In GUI	Slot in Workflow File	Type
Set of annotations	annotations	<i>annotation-table-list</i>