Finding Pattern Using Smith-Waterman Algorithm

Task Name: find-sw

Searches for a pattern in a nucleotide or protein sequence using the Smith-Waterman algorithm and saves the regions found as annotations.

Parameters:

in — input sequence file. [String, Required]

out — output file with the annotations. [String, Required]

name — name of the annotated regions. [String, Optional, Default: “misc_feature”]

ptrn — subsequence pattern to search for (e.g. *AGGCCT*). [String, Required]

score — percent identity between the pattern and a subsequence. [Number, Optional, Default: 90]

matrix — scoring matrix. [String, Optional, Default: “Auto”]

Among others the following values are available:

- blosum62
- dna
- rna
- dayhoff
- gonnet
- pam250
- etc.

The matrices available are stored in the $UGENE\data\weight_matrix directory.

filter — results filtering strategy. [String, Optional, Default: “filter-intersections”]

The following values are available:

- filter-intersections
- none

Example:

```
ugene find-sw --in=human_T1.fa --out=sw.gb --ptrn=TGCT --filter=none
```