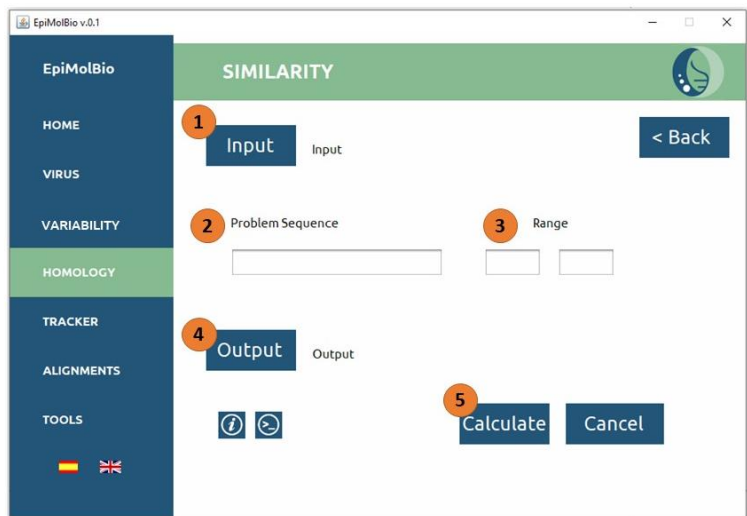


Homology Similarity

It allows searching for a user-introduced target sequence among the sequences in the input file, obtaining the proportion of sequences per file that contain the target sequence.



- 1 Select the input folder. It should only contain .fasta files of aligned NT or AA sequences to be analyzed (one or multiple files).
- 2 Enter the target sequence (e.g., KLKPGMDGPKVK).
- 3 Enter the positions of the input protein that encompass the region where you want to search.
- 4 Select the output folder and enter the file name in the following format: /NAME.html.
- 5 Click Calculate.

Homology Similarity Range 5 - 20		
Problem Sequence: WQRPLVT		
File	Frequency	Total Sequences
PR_01_AE.fasta	76.643%	26849
PR_02_AG.fasta	64.728%	9577
PR_03_A6B.fasta	69.355%	310
PR_04_cpx.fasta	53.333%	15
PR_05_DF.fasta	4.167%	24
PR_06_cpx.fasta	61.126%	746
PR_07_BC.fasta	76.695%	10916

The .html file is a table containing the target sequence, the names of the input files, the frequency of occurrence of the target sequence in each file, colored according to the color code, and the total number of sequences in the analyzed file.

More information about this function can be found in the User Manual, section:
III. HOMOLOGY III.1 SIMILARITY