

Homology Partial Similarity

It allows comparing a user-introduced sequence with the input sequences to search for similar regions between them, defining the percentage of similarity between the sequences.

The screenshot shows the EpiMolBio v0.1 web interface. The sidebar on the left contains navigation links: HOME, VIRUS, VARIABILITY, HOMOLOGY (highlighted), TRACKER, ALIGNMENTS, and TOOLS. The main panel is titled 'PARTIAL SIMILARITY' and features a series of numbered steps (1-7) indicating the workflow. Step 1 is 'Input', Step 2 is 'Length', Step 3 is '% Similarity', Step 4 is 'Align' (with a checkbox for 'Not Align'), Step 5 is 'Sequence', Step 6 is 'Output', and Step 7 is 'Calculate'. A '< Back' button is located at the top right of the main panel.

- ① Select the input folder. It should only contain .fasta files of sequences to be analyzed in either nucleotide (NT) or amino acid (AA) format.
- ② Enter the length of fragments into which the sequences will be divided for comparison.
- ③ Enter the minimum percentage of similarity.
- ④ Select 'Align' for a comprehensive analysis or 'No Align' for a quicker one.
- ⑤ Enter the sequence for comparison with the input file without line breaks or spaces.
- ⑥ Select the output folder and enter the file name in the following format: /NAME.html.
- ⑦ Click Calculate.

| Homology Partial Similarity 50.0% | | | | |
|-----------------------------------|-------------------------------|----------------|----------------|------------|
| File | Sequence | Input Sequence | Found Sequence | Similarity |
| PR_D.fasta | >D.ET.2003.ETH_G_230.AB285830 | GHRATGTVLV | GTDTTITVNV | 50.000% |
| PR_D.fasta | >D.ET.2003.ETH_G_230.AB285830 | IGRNLLTQLG | NGMNGRTILG | 50.000% |
| PR_D.fasta | >D.ET.2003.ETH_G_230.AB285830 | GRNLLTQLGC | GMNGRTILGS | 50.000% |
| PR_D.fasta | >D.ET.2003.ETH_G_230.AB285830 | NLLTQLGCTL | NGRTILGSAL | 50.000% |
| PR_D.fasta | >D.JP.-patient_88.AB356098 | YDQIHVEICG | LTQDHVDILG | 50.000% |
| PR_D.fasta | >D.JP.-patient_88.AB356098 | DQIHVEICGH | TQDHVDILGP | 50.000% |

The .html file is a table containing the following information: input file name, sequence headers, a fragment of the introduced sequence, the corresponding found fragment, and the similarity percentage colored according to the color code.

More information about this function can be found in the User Manual, section:

III. HOMOLOGY III.2 PARTIAL SIMILARITY