

Cyclic Sequence Aligner

User Manual

(Version 1.0)

Introduction

The Cyclic Sequence Aligner tool allows to find the best possible rotation for a set of cyclic sequences that will be later aligned using a linear multiple sequence alignment algorithm. The best rotation found for each sequence will effectively improve the alignment of all the sequences.

This user manual will guide the user through all the steps from uploading the sequences to understanding the output, covering all the options and results in the process.

Main Page

The tool is available through an intuitive and easy to use web interface where the user can upload the cyclic sequences and set the visualization parameters.

HOME SUBMIT DOCUMENTATION DOWNLOADS CONTACT US

Cyclic DNA Sequence Aligner

DESCRIPTION

This tool finds the correct rotation for a set of cyclic sequences that are to be aligned. It is very well suitable for example to apply to mitochondrial genome alignments.

The best rotation is calculated based on the largest non-repeated block that belongs to all the sequences. These maximum common blocks are obtained with the help of a generalized cyclic suffix tree.

As it is developed, this tool does not perform the multiple sequence alignment itself. The user is presented with some suggestions of other highly efficient multiple sequence alignment tools to perform this task.

SUBMISSION

The sequences can be submitted in the Multi-FASTA format by uploading a file in this format or by pasting the sequences in the text window. The length of the blocks displayed in the output can be specified by the user. By default only blocks with length greater than 15 are displayed.

Sequences

Options ☒ Upload File Minimum block size (≥0) ☐ Upload Text Maximum block size (0=∞)

MULTIPLE SEQUENCE ALIGNMENT TOOLS

[ClustalW](#)

[Mavid](#)

[Mauve](#)

[SincView](#)

SEQUENCE GENOME VIEWER

[DNAPlotter](#)

DOCUMENTS

[User Manual](#)

DOWNLOADS

[Source code](#)

CONTACT

Please send your comments, suggestions, bug reports or questions to:
csatool@kdh'o.inesc-id.pt

Copyright © 2009 Privacy Policy Terms of Use Sitemap CSS
Design by FreeTemplates, revised by FreeFlashTemplates

Uploading the Sequences

The submission form can be found in the “*Submission*” section at the middle of the main page. The sequences must be submitted all at once by storing the set in the Multi-FASTA format. There are two ways to upload the data:

- **File upload:** specify a file containing the data in this format.

Sequences	C:\CSA\Examples\Primates.fasta		Browse...
Options	<input checked="" type="radio"/> Upload File	15	Minimum block size (≥ 8)
	<input type="radio"/> Upload Text	0	Maximum block size ($0=\infty$)
Submit			

- **Text upload:** directly copy/paste the data into the text field.

Sequences	<pre>>gi 5835121 ref NC_001643.1 Pan troglodytes mitochondrion, complete genome GTTTATGTAGCTTACCCCTCAAAGCAATACACTGAAAATGTTTC GACGGGTTTACATCACCCATAAACAAACAGGTTTGGTCCTAGCC TTTCTATTAGCTCTTAGTAAGATTACACATGCAAGCATCCCCGCC</pre>		
Options	<input type="radio"/> Upload File	15	Minimum block size (≥ 8)
	<input checked="" type="radio"/> Upload Text	0	Maximum block size ($0=\infty$)
Submit			

In current version, the maximum number of sequences that can be processed simultaneously is limited to 32. After inputting the sequences and setting the optional visualization parameters, pressing the “*Submit*” button will automatically upload the data and redirect the user to the results page.

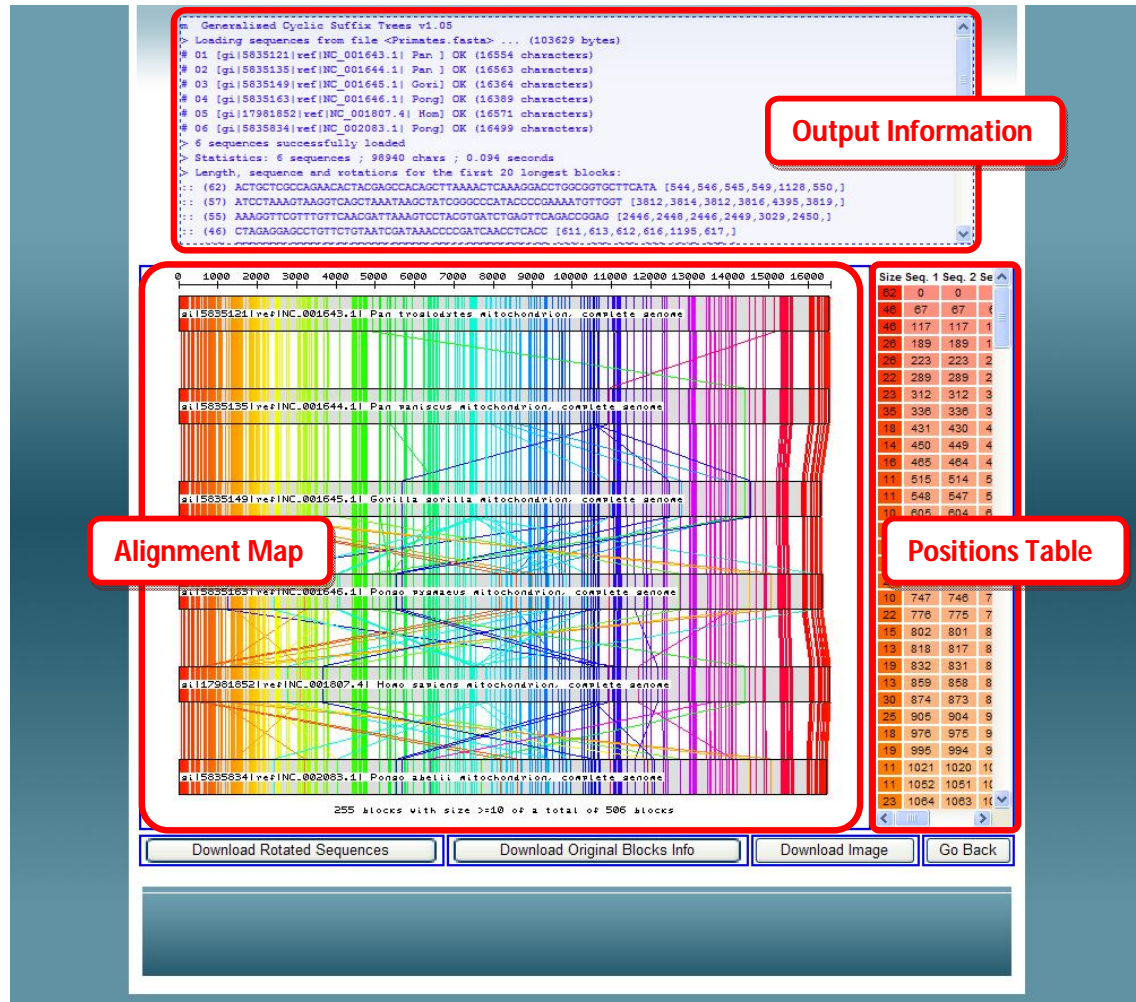
Visualization Parameters

The user also has the possibility to specify the minimum and maximum lengths of the blocks that will be shown in the alignment map later in the results page.

Sequences			Browse...
Options	<input checked="" type="radio"/> Upload File	15	Minimum block size (≥ 8)
	<input type="radio"/> Upload Text	0	Maximum block size ($0=\infty$)
Submit			

The minimum allowed length is 8. If no maximum length is required, the corresponding value should be set to 0. By default, the values are set to 15 and 0 respectively.

This page shows the general output information and an alignment map of the blocks in all the sequences after their optimal rotation, as well as a table with all their lengths and positions, following the color code shown in the map.



By right-clicking over a specific section of one of the sequences in the alignment map, the positions table automatically sorts its rows to reflect the correct order of the blocks inside that sequence and it automatically scrolls itself to show the information of the blocks from the selected region/sequence.

This page also gives the possibility to download the rotated sequences and information about the blocks:

- **Download Rotated Sequences:** returns a file in the Multi-FASTA format containing all the resulting sequences after the optimal rotation is applied to each one.
- **Download Original Blocks Info:** gives a plain text file where each line shows the size, nucleotide sequence and position on each original sequence (before the rotation) of each block.
- **Download Image:** transfers the image file of the displayed alignment map.

The "Go Back" button takes the user to the initial main page.

Understanding the Results

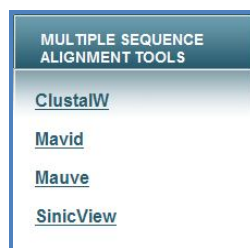
One important remark is that only the maximum blocks which occur in all the sequences and only once per sequence are considered for the results. The longest such block is used to set the optimal rotation of each sequence, so this block appears in the beginning of each rotated sequence. Beside this, the limits set in the visualization parameters are also taken into account when drawing the blocks in the alignment map.

The results page is divided into the following sections:

- **Output Information:** shows the general status and output messages resulting from the running of the tool, including size and description of each input sequence, running time, information about the 20 first longest blocks, and eventually any processing errors.
- **Alignment Map:** this image represents the relative location of the blocks inside each sequence. The blocks are marked using a gradient coloring to allow their easy detection between distinct sequences when translocations occur. All the represented sequences are already rotated. Note that only the blocks with lengths within the limits specified in the visualization parameters are shown. As already mentioned, this map is sensible to clicks.
- **Positions Table:** shows the size and position in each rotated sequence of all the blocks represented in the alignment map, following the same color scheme for each block in order to facilitate its identification in the image on the left.

Multiple Sequence Alignment

Although this feature is planned to be incorporated in a near future, this tool does not yet perform the multiple sequence alignment itself. In the meanwhile, some alternatives are presented.



After retrieving the file containing all the optimally rotated sequences using the "*Download Rotated Sequences*" button, they can be aligned using any regular linear multiple sequence alignment tool. For this purpose, some links of very efficient such tools are presented in the main page.

Additional Information

Any additional questions, comments, bug reports or suggestions can be sent to the following e-mail address: csatool@kdbio.inesc-id.pt