

Gecko-MGV

Document: User guide

The main screen

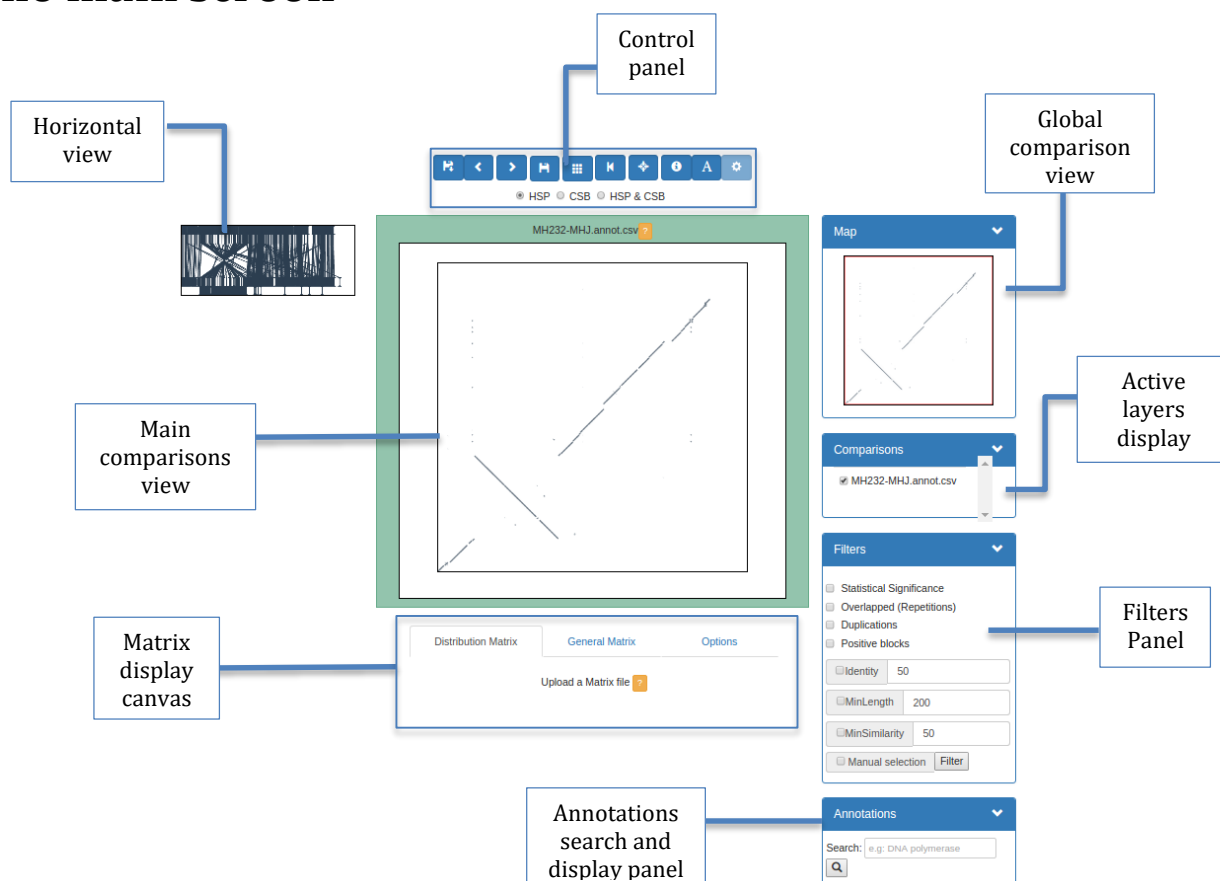


Figure 1: Gecko-MGV main screen with a central canvas and in the right the control panel. In the upper right corner a small map displays the position of zoomed area –during zoom-out operations

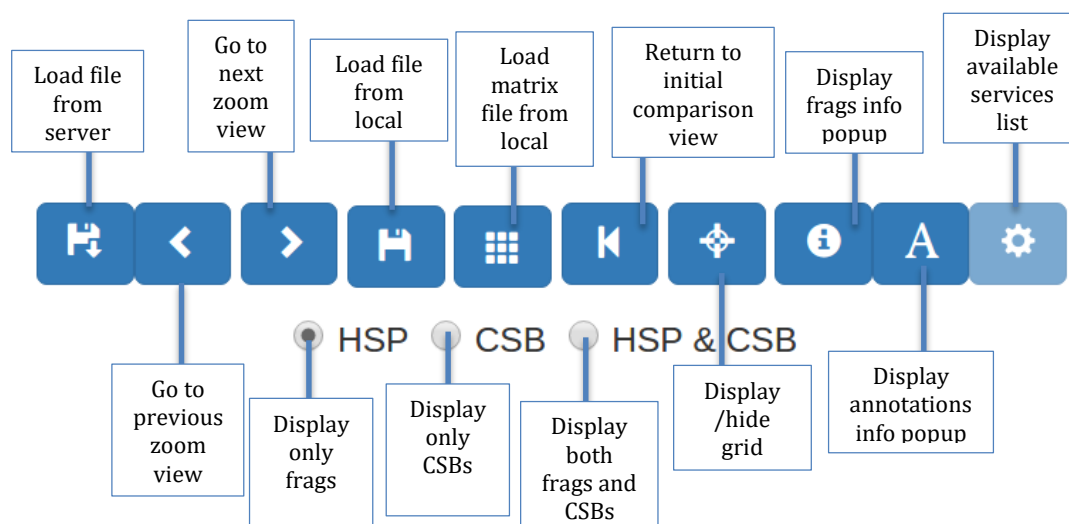


Figure 2: Gecko-MGV control panel

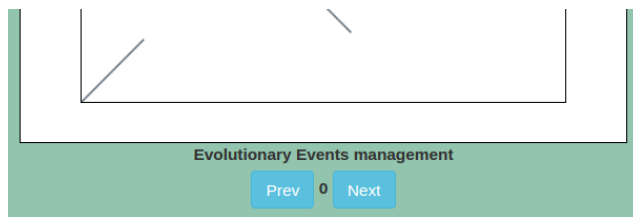


Figure 3: A small control panel showing the evolution index and two buttons to navigate through them is displayed under the main view when an evolutionary event is loaded.

a. Basic functionality

Load data files: Browsing the file system or upload a compatible file. The original output is a binary file, but in this case the processed CSV file must be used.

Saving options.-

- Save selected fragments. This option saves into a file the current selected fragments. This file can be used as an input.
- Save services execution results

Fragments distribution matrix: Traditional filters based on minimal length, number of identities or percentages, etc. produce an abrupt segmentation of fragments that could be avoided using dynamic thresholds based on the probability density function of fragments (p-Value). This distribution is computed –without additional cost- during fragments computation.

Type of display (views).

Display style

- (1) Traditional Dotplot-like
- (2) Horizontal (Mauve-like representation)

When filters has been applied, the system shows all the filtered fragments with a degraded (grey) colour.

Signal level

All options apply over the active “type of representation”

- Change different representations using HSP, CSB or both

Only HSP: displays HSP
Only CSB Displays computational syntenic blocks
HSP and CSB: Displays both

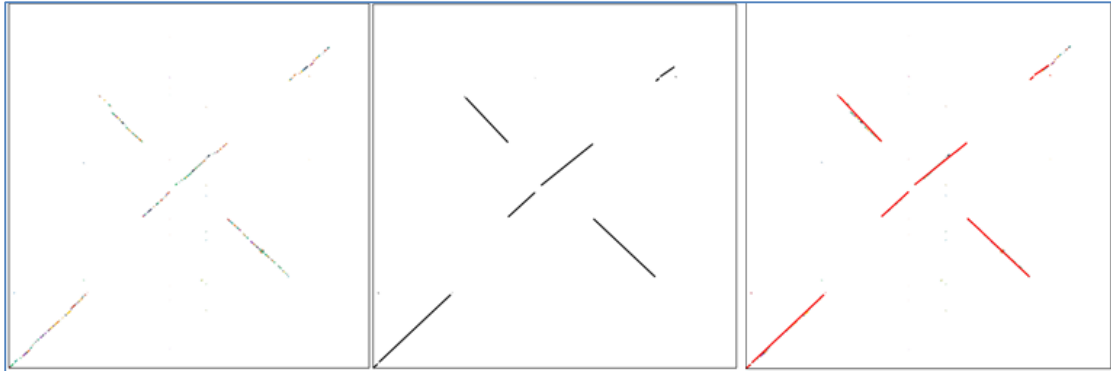


Figure 4: From left to right: HSP, CSB and both representations (* CSB are painted in red when both HSP and CSB are represented)

Additional functionality:

- Initial view: Restores the original view
- Grid: Displays a 5 x 5 grid as reference, including the numeric coordinates
- CSV Info: opens a floating window for text based info on the Fragments, the CSB or both
- Annotations: opens a floating window which shows the annotations referring to the fragments that are actually on the main canvas.

Filters.

There are several applicable filters which will mark those filtered, including a manual filter that will filter a selection and will unfilter those selected while holding 'f'

Filters

☐ Statistical Significance
☐ Overlapped (Repetitions)
☐ Duplications
☐ Positive blocks

☐ Identity
50

☐ MinLength
200

☐ MinSimilarity
50

☐ Manual selection
Filter

Figure 5: View of the filter panel

- By Statistical Significance
- Overlapped
- Duplications
- Positive blocks
- Minimal Length
- Minimal Similarity

Context information

SeqK filename	G1.fna
SeqY filename	G2.fna
SeqK name	gi 148377268 ref NC_009497.1
SeqY name	gi 313678134 ref NC_014760.1
SeqK length	877438
SeqY length	1003404
Min.fragment.length	50
Min.identity	50
Tot Hits (seeds)	3177
Tot Hits (seeds) used	685
Total fragments	523

Frag - 5

xStart: 400000
yStart: 400000
xEnd: 500000
yEnd: 500000
Strand: f
Block: 0
length: 0
Score(AVG): 0
Identity: 0
Similarity: 0

Execute service

CSB & Frag info

Save

Type	xStart	yStart	xEnd	yEnd	strand	block	length	score	ident	similarity	perc_ident	SeqX	SeqY	gene	locus_tag	product
Frag	0	8	2924	2932	f	1	2925	62.94	2383.00	81.47	62.94	7364	-	-		
GX	1	0	1401	0	f	1	1400	0	0	0	0	0	0	dnaA	MAG_0010	chromosomal replication initiation protein
GX	1535	0	2644	0	f	1	1109	0	0	0	0	0	0	dnaN	MAG_0020	DNA polymerase III

Figure 5: Contextual information, from left to right, for the global HSP identification process; at a fragment level and for a collection of fragments –in general, obtained by applying filters

Annotations

Functional, evolutionary or structural information are considered under this epigraph.

The way to exploit this information is by linking it with fragments or CSB, etc. This is to say:

- Display the annotations (for both or all the involved sequences) associated with selected fragments
- Text-based search over the annotations
- Identify annotations found

Results from searching is highlighted in the main canvas in the axis through small lines.

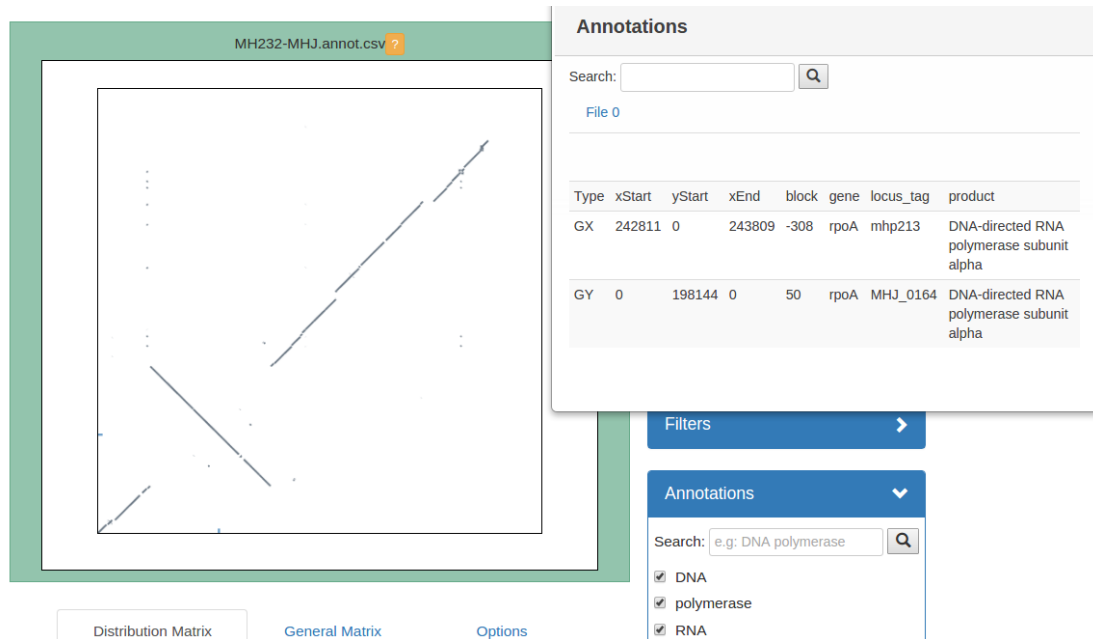


Figure 6: The annotations searched through the search boxes on the info panels or on the annotations panel

External services

The application includes a way to contact with external services to define new-processing tasks using a Web-Services interface. Results will be displayed in a new tab

There is a big collection of sequence analysis procedures available through web-services so it can be added in any moment.

As example we offer in our application the Clustal Omega web service integrated with the rest of the services.

Internal services

There is a set of internal services already given in the services tab, more services can be added through the admin side with a simple Python form and filling a form with the necessary information.

Services










Calculate Kmers Calculate Kmers from a given sequence	 Start
Extract sequence from frags Extract sequences from frags	 Start
Reverse Complement Write into the second parameter the reversed complement of the first parameter.	 Start
Extract overlapped All fragment partially or totally overlapped are separated into a new collection	 Start
Connect frags up Connects ungapped fragments that are likely to be part of a gapped fragment	 Start
Refine overlapped A fine-tuning separation is performed by extracting full overlapped regions from partially overlapped fragments	 Start
Classify repeats Separates overlapped fragments in three categories: duplications, tandem repeats and dispersed repeats	 Start
Gecko Program for genomes comparison handling long sequences.	 Start
Clustal Align sequences	 Start

Figure 7: In the services tab can be found all the currently available services.

The information about any service is provided through the creation form and can be found on the services tab.

Once in a service it will ask the user for the needed data, those files that will be as output files will be directly created on the user file system.

frags file:

X Fasta file:

Y Fasta file:

Y-reversed Fasta file:

Output FragFile:

Block:

Save results as:

Figure 8: The annotations searched through the search boxes on the info panels or on the annotations panel

1.4 Files system

Registered users can upload their own files, and also store the generated using the external services directly in their particularly user folder.

<input type="button" value="Upload file"/>		<input type="button" value="New"/>
Filename	Size	Options
AE017244.1-hypo-FloccurareATC27399.csv	17,0 KB	<input type="button" value="Delete"/> <input type="button" value="Open"/>
FloccurareATC27399.fasta	771,5 KB	<input type="button" value="Delete"/> <input type="button" value="Open"/>
AE017244.1-hypo.fasta	924,3 KB	<input type="button" value="Delete"/> <input type="button" value="Open"/>

Figure 9: View of the file management system