# 1. Introduction

The Genome Comparison browser is developed to compare genomes with following results' visualization. The application has the following abilities:

- To compare genomes (aligning genomes with use of the gsbl program).
- To visualize genomes and similarity regions.
- To visualize the similarity matrix.
- To search for motifs in genomes.

2. General system description

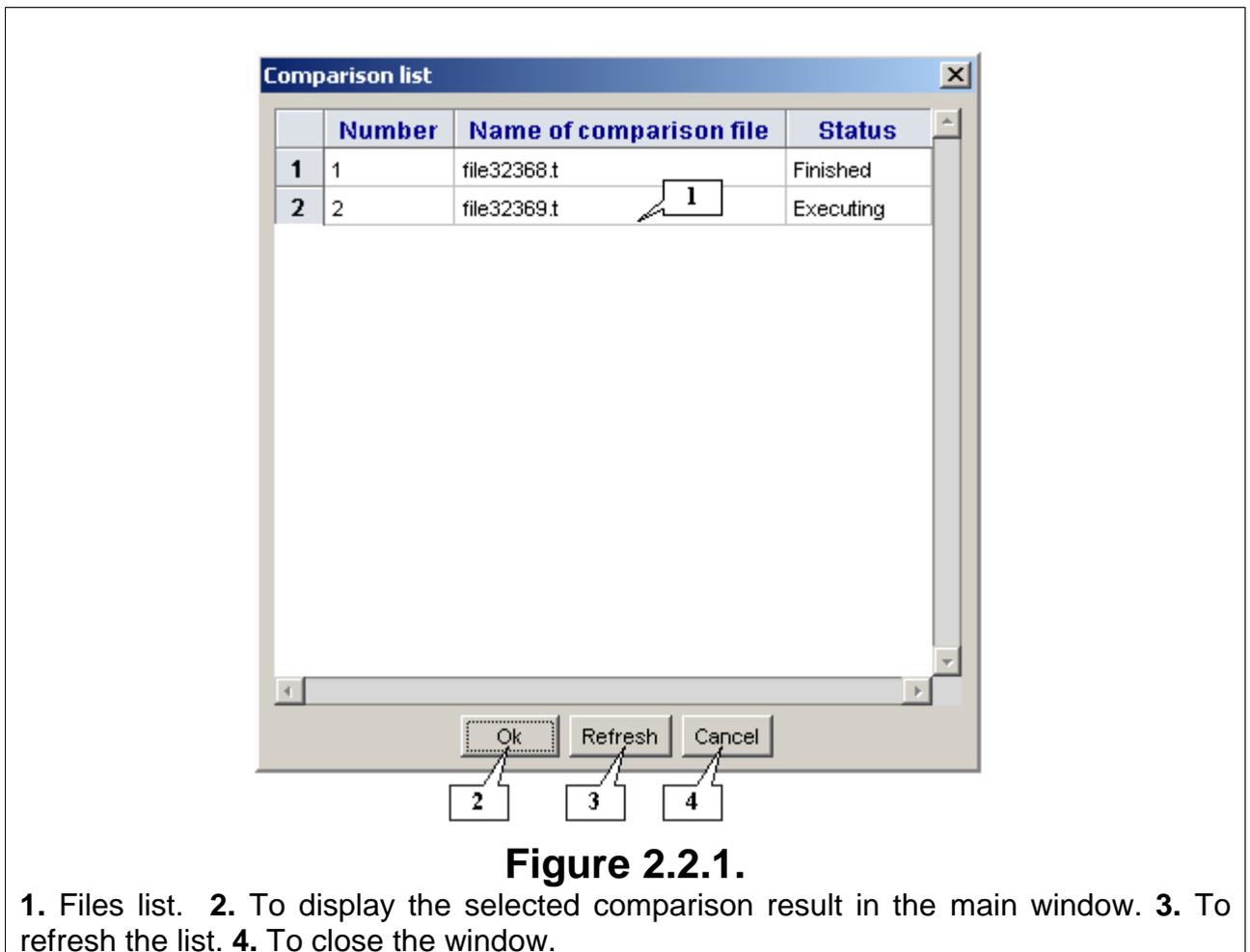## 2.1. Main window.
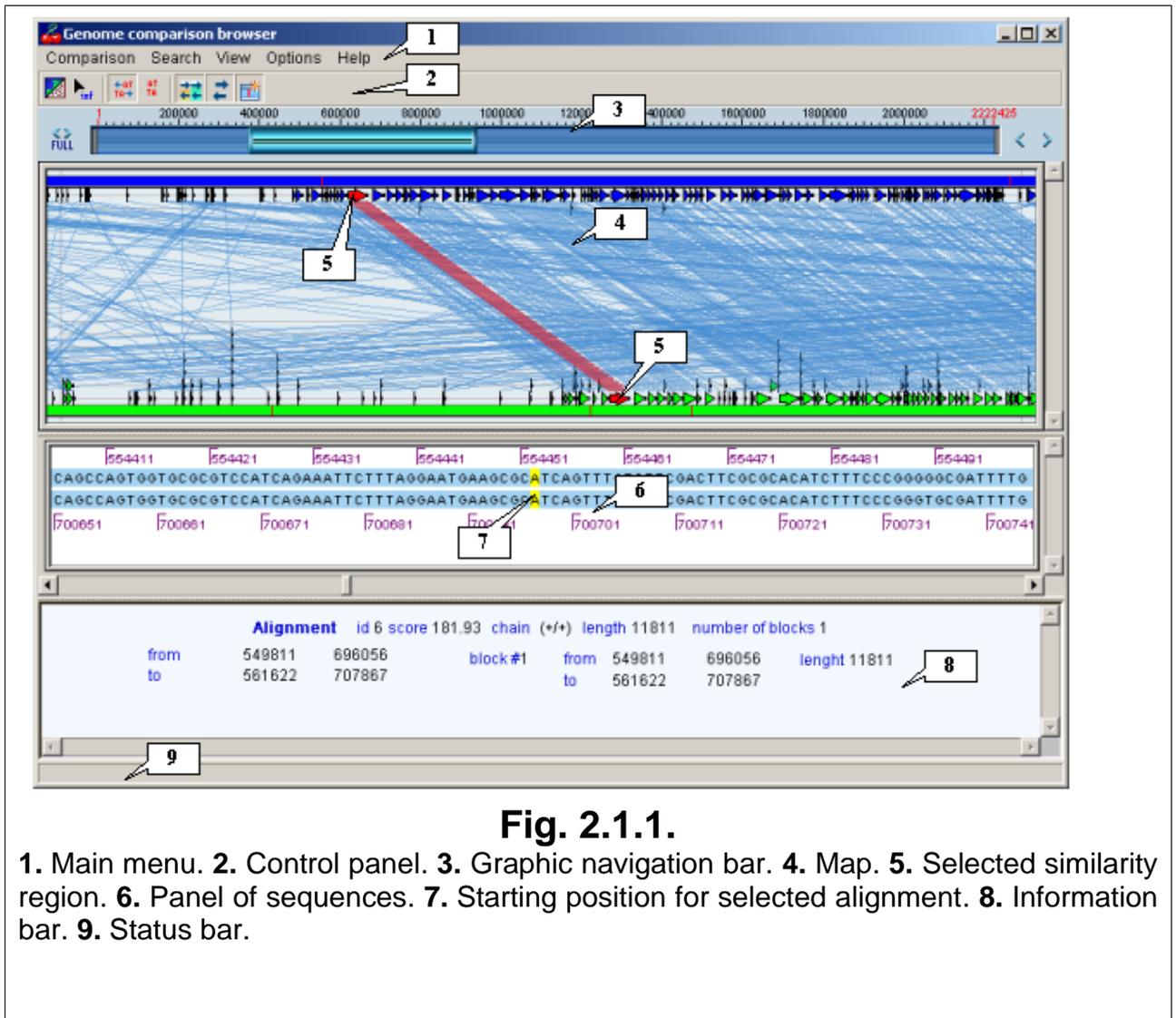
Fig. 2.1.1.

1. Main menu. **2.** Control panel. **3.** Graphic navigation bar. **4.** Map. **5.** Selected similarity region. **6.** Panel of sequences. **7.** Starting position for selected alignment. **8.** Information bar. **9.** Status bar.

2.2. Main menu.

The main menu contains the following sections:

- The "Comparison" menu includes the commands:
  - ○ "Open comparison file" - To open the file with results of gsbl comparison program work.
  - ○ "Compare sequences" – To call out the "Comparison parameters" dialog window purposed to set the values for comparison parameters (see ch. 3.1.2).

o "Comparison list" – To call out the "Comparison list" window (fig. 2.2.1) purposed to look through the tasks in background. The list contains the tasks' IDs and status. To view the result, select the file in the list and click the "OK" button.

	Number	Name of comparison file		Status
1	1	file32368.t		Finished
2	2	file32369.t	1	Executing

[Ok] [Refresh] [Cancel]
2   3   4

Figure 2.2.1.
1. Files list. **2.** To display the selected comparison result in the main window. **3.** To refresh the list. **4.** To close the window.

o "Set sequences" – To call out the "Set sequences" dialog window purposed to view sequences (see ch. 3.1.1).

- The "Search" menu includes the commands:
  o "Find motif" – To call out the "Find motif" dialog window purposed to search for motif in sequences (see ch. 3.2).

- The "View" menu includes the commands:
  o "Dot matrix diagram" – To switch the map visualization mode from sequences to similarity matrix (see ch. 2.4.2).

- The " Options" menu includes the commands:
  o "Transparency" – If the option is on, the lines connecting the sequences' similarity regions becomes transparent. Otherwise, the lines overlaying each other (see ch. 2.4.1).
  o "Alignments" – If the option is on, the lines connecting the sequences' similarity regions are shown. Otherwise, the lines are not (see ch. 2.4.1).

o "Details selection mode" – If the option is on, the only one block in similarity region is selected. Otherwise, all blocks are selected (see ch. 2.4.1).

- The "Help" menu includes the commands:
  o "Content" – N/A.
  o "About application" – To call out the "About" information window containing information on the application.

2.3. Control Panel

Functions of the control panel's buttons:

	"Show dot matrix diagram" – To switch the map visualization mode from sequences to similarity matrix. The same as the «View->Dot matrix diagram» command of the main menu.
	"Quick view mode" – While the button is pressed, the information bar shows the data on similarity region pointed out by mouse. When the button is released, the data is shown only after clicking on the region.
	"Enable positioning" – While the button is pressed, clicking on the similarity region selects the starting nucleotides of the region in the panel of sequences.
	"Disable positioning" – To switch off the effects of the "Enable positioning" button.
	"Opacity" - While the button is pressed, the lines connecting the sequences' similarity regions becomes transparent. The same as the «Options->Transparency» command of the main menu.
	"Draw alignment" - While the button is pressed, the lines connecting the sequences' similarity regions are shown. The same as the «Options->Alignments» command of the main menu.
	"Block selection" – While the button is pressed, the only one block in similarity region is selected. Otherwise, all blocks are selected. The same as the «Options->Details selection mode» command of the main menu.
	"Next zoom" – To switch to the following map scale. The button is available in similarity matrix visualization mode only.
	"Defaut zoom" – To restore the initial map scale. The button is available in similarity matrix visualization mode only.
	"Preiew zoom" – To restore the previous map scale. The button is available in similarity matrix visualization mode only.
	"Move diagram" – While the button is pressed, the similarity matrix can be scrolled using the mouse. Otherwise, the mouse is used to change the matrix scale. The button is available in similarity matrix visualization mode only.

2.4. Map

The map is purposed to schematically display the loaded data and to retrieve the information on similarity region.
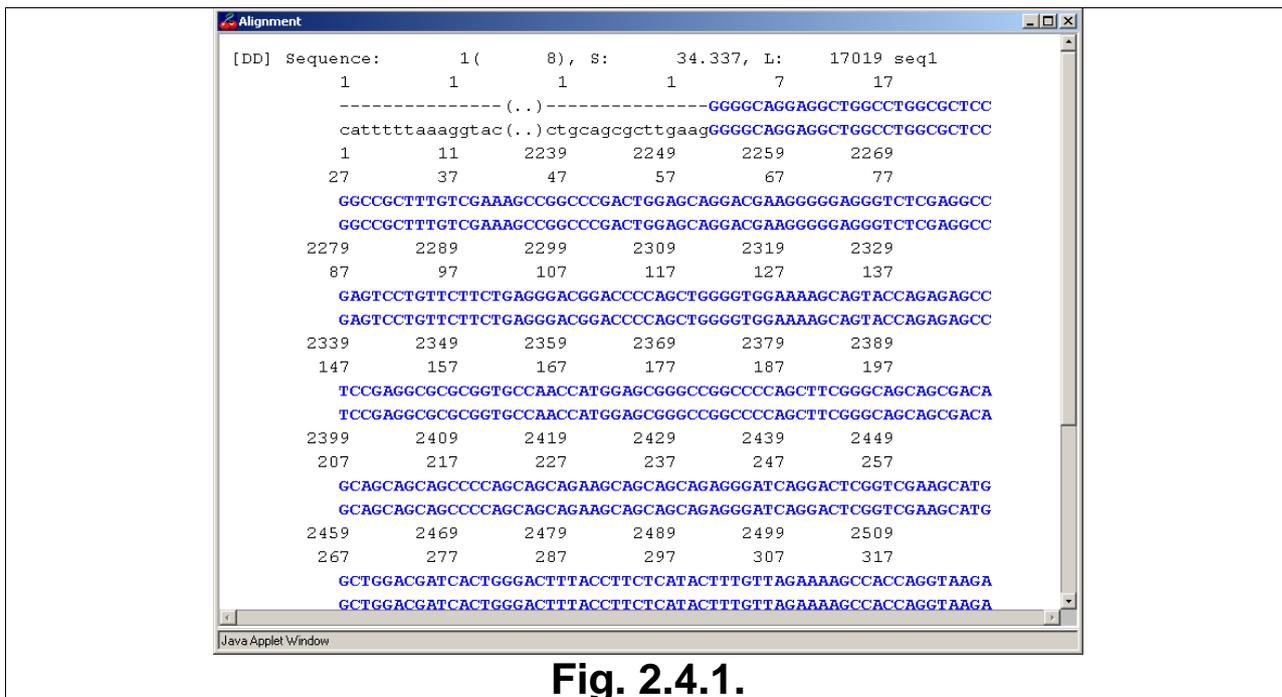Map visualization modes:
- The mode for graphic representation of sequences' similarity (fig. 2.4.1.1).
- The mode for visualization of genomes' similarity matrix (fig. 2.4.2.1.).

To switch between the modes, use either the button of the control panel, or the «View->Dot matrix diagram» of the main menu.

In any mode, the mouse right click over the similarity region calls out the window with nucleotide sequences of regions, between which the similarity is found (fig. 2.4.1).

While the inf button is pressed, the information bar shows the data on similarity region pointed out by mouse. When the button is released, the data is shown only after clicking on the region.
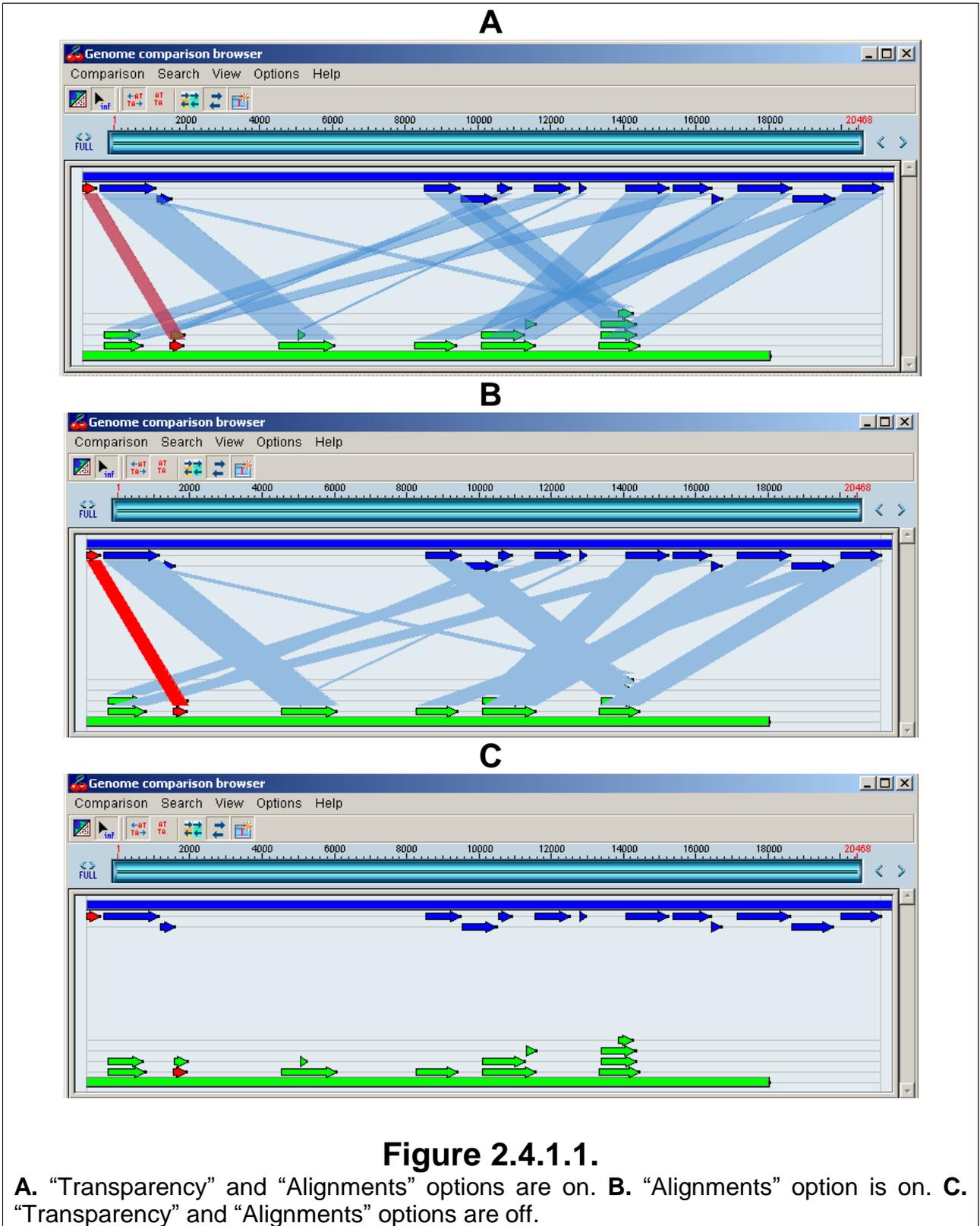
```
Alignment                                                                    _|□|×|

[DD] Sequence:      1(     8), S:      34.337, L:     17019 seq1
        1         1        1         1        7        17
        ---------------(..)---------------GGGGCAGGAGGCTGGCCTGGCGCTCC
        cattttttaaaggtac(..)ctgcagcgcttgaagGGGGCAGGAGGCTGGCCTGGCGCTCC
        1        11       2239      2249     2259     2269
        27        37        47        57        67        77
            GGCCGCTTTGTCGAAAGCCGGCCCGACTGGAGCAGGACGAAGGGGGAGGGTCTCGAGGCC
            GGCCGCTTTGTCGAAAGCCGGCCCGACTGGAGCAGGACGAAGGGGGAGGGTCTCGAGGCC
        2279      2289      2299      2309     2319     2329
            87        97       107       117      127      137
            GAGTCCTGTTCTTCTGAGGGACGGACCCCAGCTGGGGTGGAAAAGCAGTACCAGAGAGCC
            GAGTCCTGTTCTTCTGAGGGACGGACCCCAGCTGGGGTGGAAAAGCAGTACCAGAGAGCC
        2339      2349      2359      2369     2379     2389
            147       157       167       177      187      197
            TCCGAGGCGCGCGGTGCCAACCATGGAGCGGGCCGGCCCCAGCTTCGGGCAGCAGCGACA
            TCCGAGGCGCGCGGTGCCAACCATGGAGCGGGCCGGCCCCAGCTTCGGGCAGCAGCGACA
        2399      2409      2419      2429     2439     2449
            207       217       227       237      247      257
            GCAGCAGCAGCCCCAGCAGCAGAAGCAGCAGCAGAGGGATCAGGACTCGGTCGAAGCATG
            GCAGCAGCAGCCCCAGCAGCAGAAGCAGCAGCAGAGGGATCAGGACTCGGTCGAAGCATG
        2459      2469      2479      2489     2499     2509
            267       277       287       297      307      317
            GCTGGACGATCACTGGGACTTTACCTTCTCATACTTTGTTAGAAAAGCCACCAGGTAAGA
            GCTGGACGATCACTGGGACTTTACCTTCTCATACTTTGTTAGAAAAGCCACCAGGTAAGA

Java Applet Window
```

Fig. 2.4.1.

2.4.1. Mode for graphic representation of sequences' similarity

In this mode sequences and similarity regions are shown schematically. Regions, between which the similarity is found, are linked by lines (fig. 2.4.1.1). Only compared region is shown. Flanking regions, that are not included in comparison, are not shown.

Imaging on the map can be regulated using the following commands of the main menu or the buttons of the of the control panel (fig. 2.4.1.1):

- "Transparency" (![icon]) – If the option is on, the lines connecting the sequences' similarity regions becomes transparent. Otherwise, the lines overlaying each other.

- "Alignments" (![icon]) – If the option is on, the lines connecting the sequences' similarity regions are shown. Otherwise, the lines are not.

- "Details selection mode" (![icon]) – If the option is on, the only one block in similarity region is selected. Otherwise, all blocks are selected.

4

Figure 2.4.1.1.
A. "Transparency" and "Alignments" options are on. **B.** "Alignments" option is on. **C.** "Transparency" and "Alignments" options are off.

The map in this mode can be navigated either with use of mouse, or with use of graphic navigation bar (ch.2.4.1.1).
Navigation by mouse allows:
o To scroll the map to left or right without pressing any buttons.
o To change the map scale smoothly by moving the mouse while the "Shift" key is hold. The scaling occurs from the map center. If the mouse moves to right, the map widens, otherwise it shortens. Releasing and holding the "Shift" key

during the movement will cause switching from scaling mode to map scrolling and back again.

o To maximize the map scale by clicking on similarity region while the "Ctrl" key is hold.

o To select a certain map area drawing a rectangle by mouse while the "Ctrl" key is hold. Only left and right edges of rectangle define the area to be scaled out after releasing the mouse button.

Comment. The "Ctrl" key must be hold during all mouse operations, starting from pressing the left mouse button and until it has been released; otherwise the operation will not be completed.

2.4.1.1. Graphic navigation bar

The navigator (fig. 2.4.1.1.) provides the possibility to select the frames of region displayed on the navigation panel.

The region of runner movement corresponds to frames of sequence and the runner itself – to a sequence fragment displayed on the navigation panel.

Scale. Scale represents the length of sequence. Starting and ending positions are marked with red color.

Manipulations can be performed with use of mouse. Redrawing of navigation panel region occurs on the runner relocation. The frames and scaling of this region can be changed by several ways:

- Relocate the runner to left or right by mouse.
- Catch the right or left frame of runner with mouse and change it by dragging.
- Hold the "Shift" button and drag the runner's frame by mouse. Both frames will change – the runner will contract toward its center or stretch out.
- Click on the runner movement area to the right or left of the runner itself. It will cause the runner to displace in appropriate direction for its current length.
- Press the ⟨⟩ FULL «Full» button. It will cause the whole sequence to appear on the map. I this case the runner will take up the whole movement area.
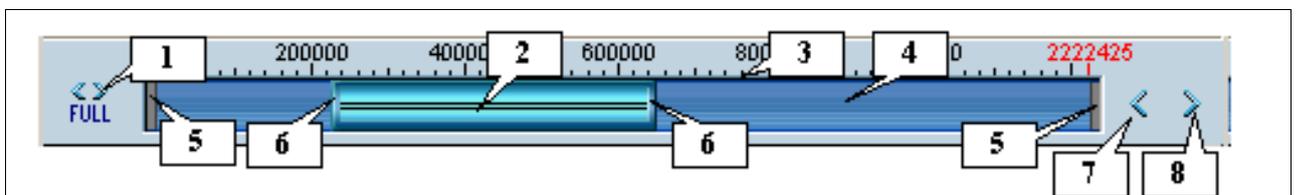
Fig. 2.4.1.1.1
1. To display the whole sequence. **2.** Runner. **3.** Scale. **4.** Runner movement area. **5.** Frames of the runner movement area. **6.** Runner frames. **7.** Left shift. **8.** Right shift.

- Press the ⟨ and ⟩ shift buttons. It will cause the runner to displace in appropriate direction for its current length.

2.4.2. Mode for visualization of genomes' similarity matrix

While in this mode, the map scaling and scrolling can be performed with use of the mouse and following buttons of the control panel:

- Hold the "Ctrl" key and draw a rectangle by the mouse to select a certain map area. Only left and right edges of rectangle define the area to be scaled out after releasing the mouse button.

- The ← button ("Next zoom") – To switch to the following map scale.
- The ↔ button ("Defaut zoom") – To restore the initial map scale.
- The → button ("Preiew zoom") – To restore the previous map scale.
- The 🖐 button ("Move diagram") – While the button is pressed, the diagram can be scrolled using the mouse. Otherwise, the mouse is used to change the matrix scale.

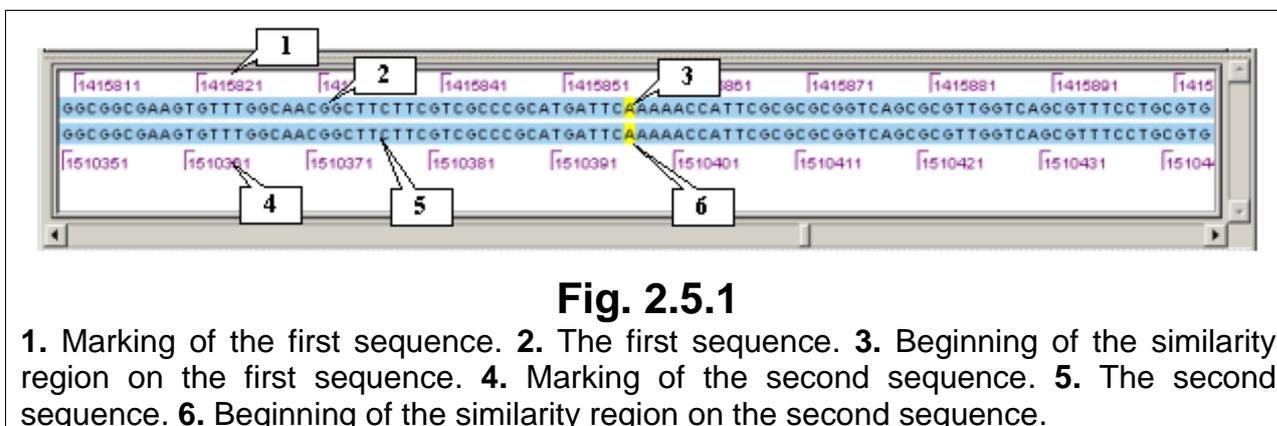The bottom and left rulers (fig. 2.4.2.1) represent coordinates of first and second sequences.

Fig. 2.4.2.1
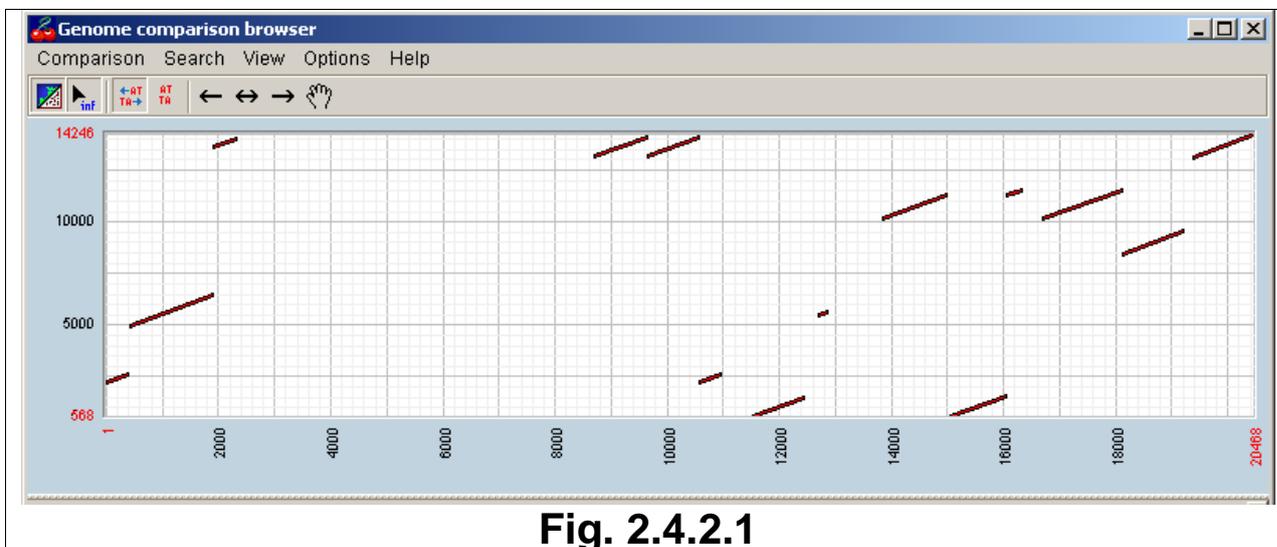
2.5. Panel of sequences

The panel displays the loaded sequences with numbering of positions (fig. 2.5.1).

If the 🔼 button is pressed, then clicking on similarity region causes the sequences to be replaced so that starting nucleotides for similarity region could be seen. The 🔼 button cancels the effect of the 🔼 button.

Sequences can be replaced using the runner.

Fig. 2.5.1

1. Marking of the first sequence. **2.** The first sequence. **3.** Beginning of the similarity region on the first sequence. **4.** Marking of the second sequence. **5.** The second sequence. **6.** Beginning of the similarity region on the second sequence.

2.6. Information bar

The bar displays the information on similarity region and on its block that is currently pointed out on the map (fig. 2.6.1).

7

While the button is pressed, the information bar shows the data on similarity region pointed out by mouse. When the button is released, the data is shown only after clicking on the region

Fig. 2.6.1

2.7. Status bar

The status bar displays the name or function of the control panel's button that is currently pointed out.

3. System operating.

3.1. Definition of sequences and comparison parameters

3.1.1. View sequences

The "Comparison>Set sequences" command of the main menu calls out the "Set sequences" dialog window purposed to load sequence into the program without execution of the comparison program (fig. 3.1.1.1). This operation is used to view the diagram of sequences on the map.

It is required to define the sequence in the window and press "OK".
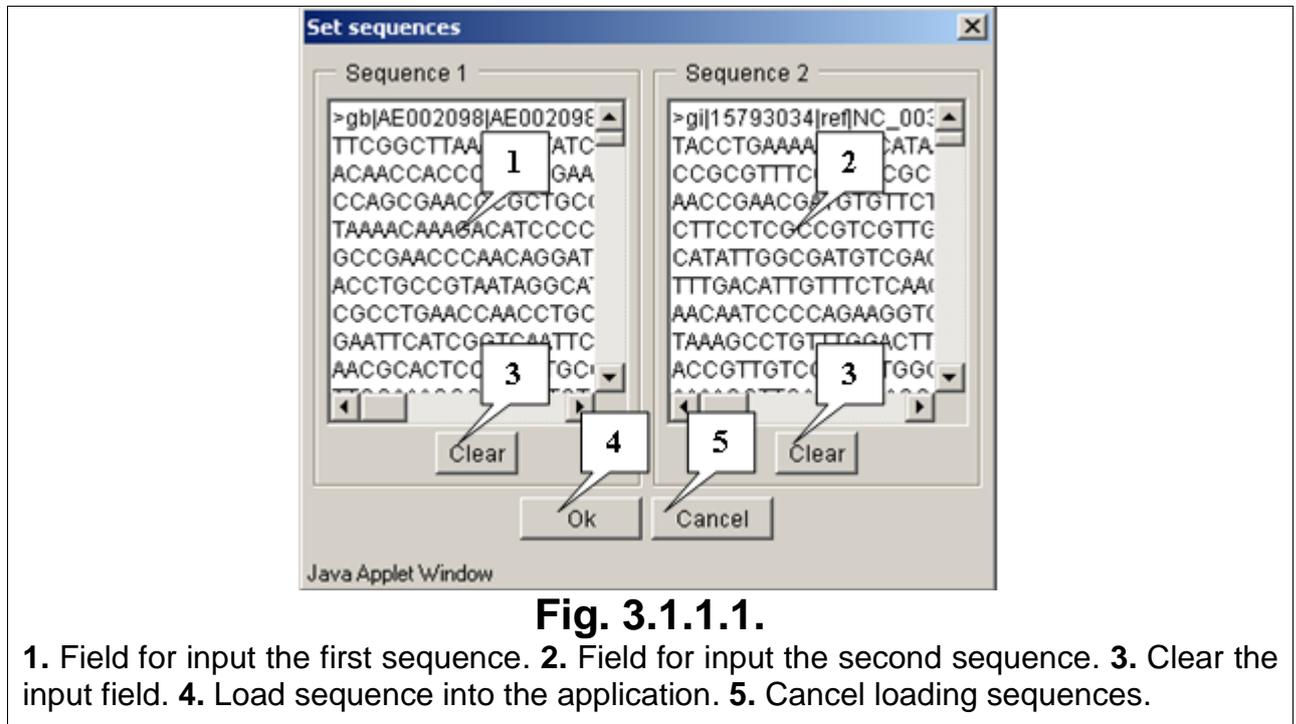
Fig. 3.1.1.1.
1. Field for input the first sequence. **2.** Field for input the second sequence. **3.** Clear the input field. **4.** Load sequence into the application. **5.** Cancel loading sequences.

3.1.2. Parameters of comparison

The "Comparison>Compare sequences" command calls out the "Comparison parameters" dialog window purposed to set the values for comparison parameters (fig. 3.1.2.1).

In this window the following parameters can be set:
- sequences;
- chains' direction;
- number of non-overlapping alignments;
- mode for visualization of similarity regions.
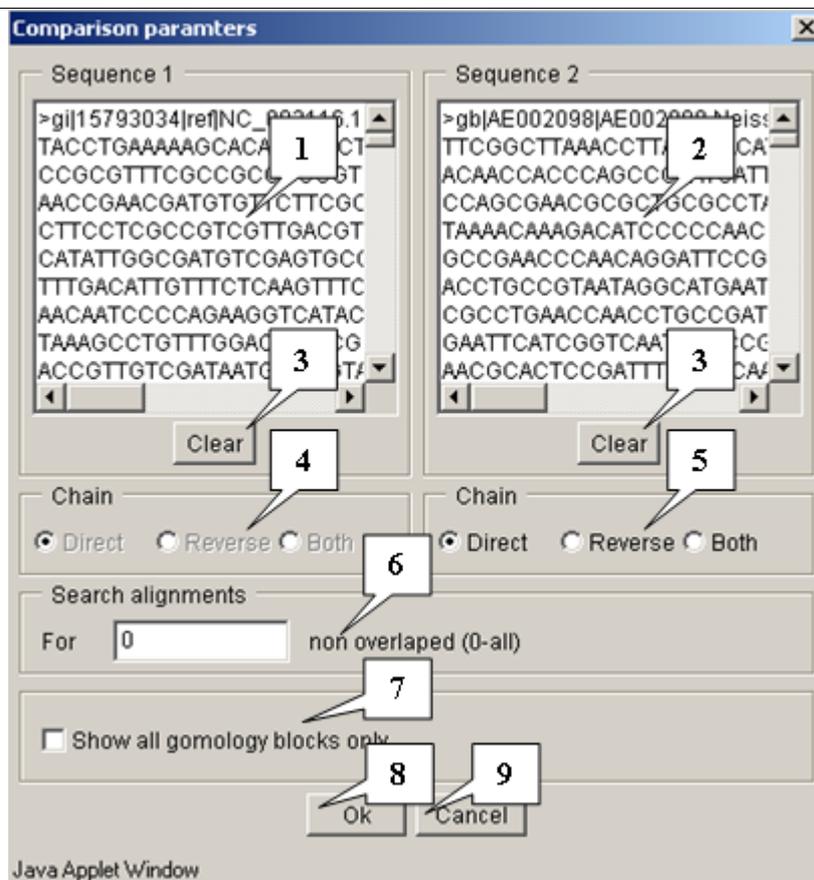
Fig. 3.1.2.1.

1. Field for input the first sequence. **2.** Field for input the second sequence. **3.** Clear the input field. **4**. Panel for choosing the first sequence's direction. **5.** Panel for choosing the second sequence's direction**. 6.** Field for setting up the number of non-overlapping alignments. **7.** 'Show all homology blocks only' checkbox. **8.** Button to start the search. **9.** Button to cancel the search.

3.2. Search for motif

This operation can be performed via the «Search->Find motif» command.

The "Find motif" dialog window (fig. 3.2.1) contains the IDs of previously made searches. On selecting a search, its parameters – ID, allowed number of mismatches, motif and search region – are shown in the window.

10
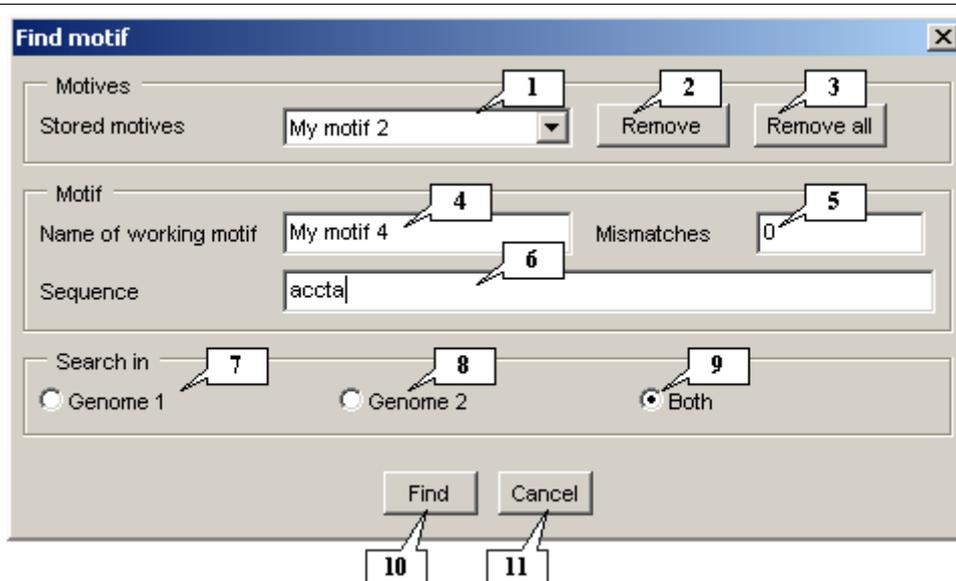
Fig. 3.2.1.

1. List of saved motifs. **2.** Remove the saved search from the list. **3.** Remove all saved searches from the list. **4**. ID of the current search. **5.** Maximal number of mismatches for the current search**. 6.** Motif's sequence. **7.** Search for motif in the first sequence. **8.** Search for motif in the second sequence. **9.** Search for motif in both sequences. **10.** Start search. **11.** Cancel search.

In order to perform the new search for motif, it is required:
1. In the «Name of working motif» field to input the ID of current motif. If such an ID is already in the list, the error message appears (fig. 3.2.2)

Fig. 3.2.2.

2. In the «Mismatches» field to set up a maximal number of allowed mismatches.
3. In the «Sequence» line to input the motif to be searched for (it must contain no gaps, otherwise motif will not be found).
4. With use of switches to define a region to be searched in:
   - «Genome 1» - in the first sequence;
   - «Genome 2» - in the second sequence;
   - «Both» - in both sequences.
5. To press the «Find» button.

For the time of search the information window (fig. 3.2.3) appears. Once the search is finished, the window disappears.
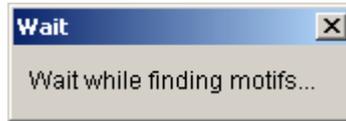
Fig. 3.2.3.

Once the search is finished, the "Motif match" window containing the list of found motifs (fig. 3.2.4) will appear. The list contains the name of genome, in which the motif was found, the starting and ending positions of the motif in genome, the chain direction, number of mismatches and number of blocks in the motif.

Comment. If there are mismatches, the motif may consist of two or more blocks.
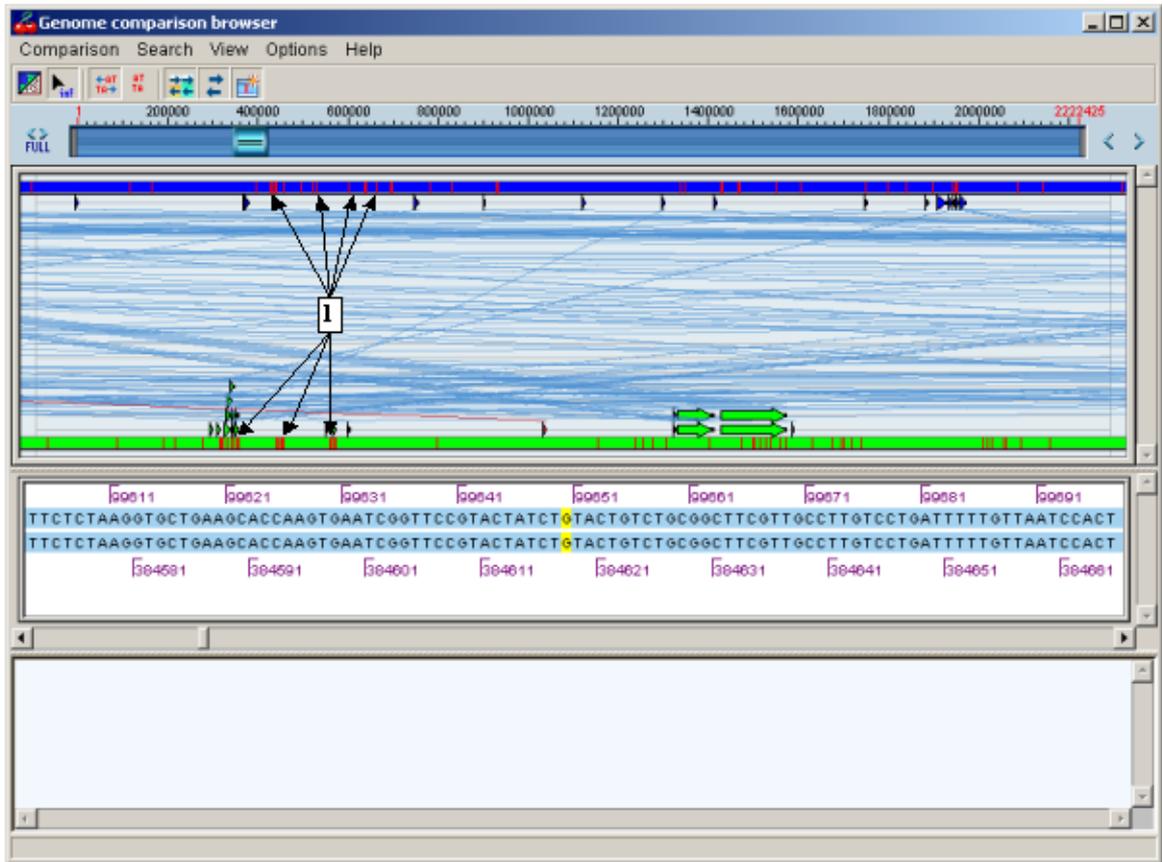

Fig. 3.2.4.
1. List of found motifs. **2.** Selected motifs. **3.** Contextual menu of the list. **4.** Close the window.

To select a motif in the list, click on it by the left mouse button. To unselect, make one more click.

Contextual menu of the list can be called out by the right mouse button click and contains the following commands:
- Show – To show selected motifs on the map (fig. 3.2.5).
- Hide – To hide selected motifs from the map.
- Select all – To select all motifs in the list.
- Deselect all – To unselect all motifs in the list.

Fig. 3.2.5.

1. Selected motifs.