XCout - Guided Exercises

Version v1: 27th November 2018.

Developed by:

Sergio Diaz-Del-Pino

Pablo Rodriguez

Ricardo Holthausen

Oswaldo Trelles

Report incidences to:
ots@ac.uma.es

Contents:

Exercise 1: Selecting interesting comparisons from the multiple pairwise comparison view

1.1 Introduction

In this exercise we will use the multiple pairwise genome comparison mode, in which we begin by loading full genome comparisons between multiple species from the server. Then we proceed to utilize the features provided by the web application that allows us to detect comparisons with an interesting signal with a quick glance.

1.1.a. Load multiple pairwise comparisons

We will start by loading the genomes comparisons between *Homo Sapiens* (HOMSA) against the *Mus Musculus* (MUSMU) and against *Rattus Norvegicus* (RATNO). Moreover, we will add the full genome comparison between *Mus Musculus* (MUSMU) and *Rattus Norvegicus* (RATNO); visualize one of these results with a clearly conserved signal; and lastly proceed to demonstrate how to remove a set of comparisons (in this case MUSMU *vs.* RATNO) from the current view.

1.1.b Conserved signal detection

Now we can load the clean and repetition frag files stored in the server. After the loading process, the application displays the different views and layers. A Horizontal and Vertical view per file is generated, and a map with the active layers is shown. At this point, the data analyst can interact with the comparisons by zooming, filtering, searching for annotations, etc.

1.2 Exercise Development

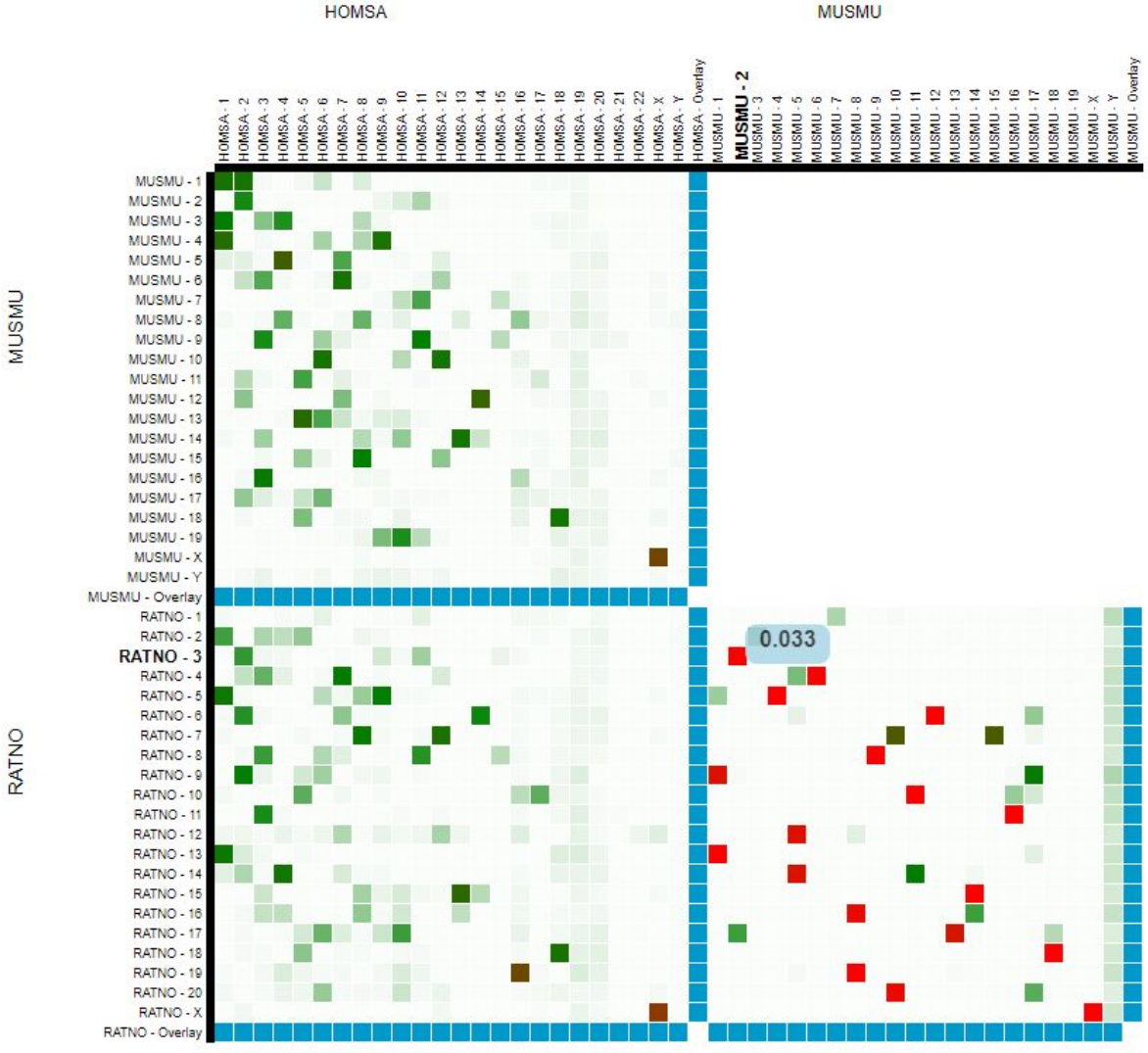
1. Enter to http://pistacho.ac.uma.es/xcout

2. Once the application is loaded, we proceed to select the comparisons of interest from the side menu located to the left. In this case we will begin by selecting HOMSA and MUSMU. To visualize them we click on the 'Add' button (+).

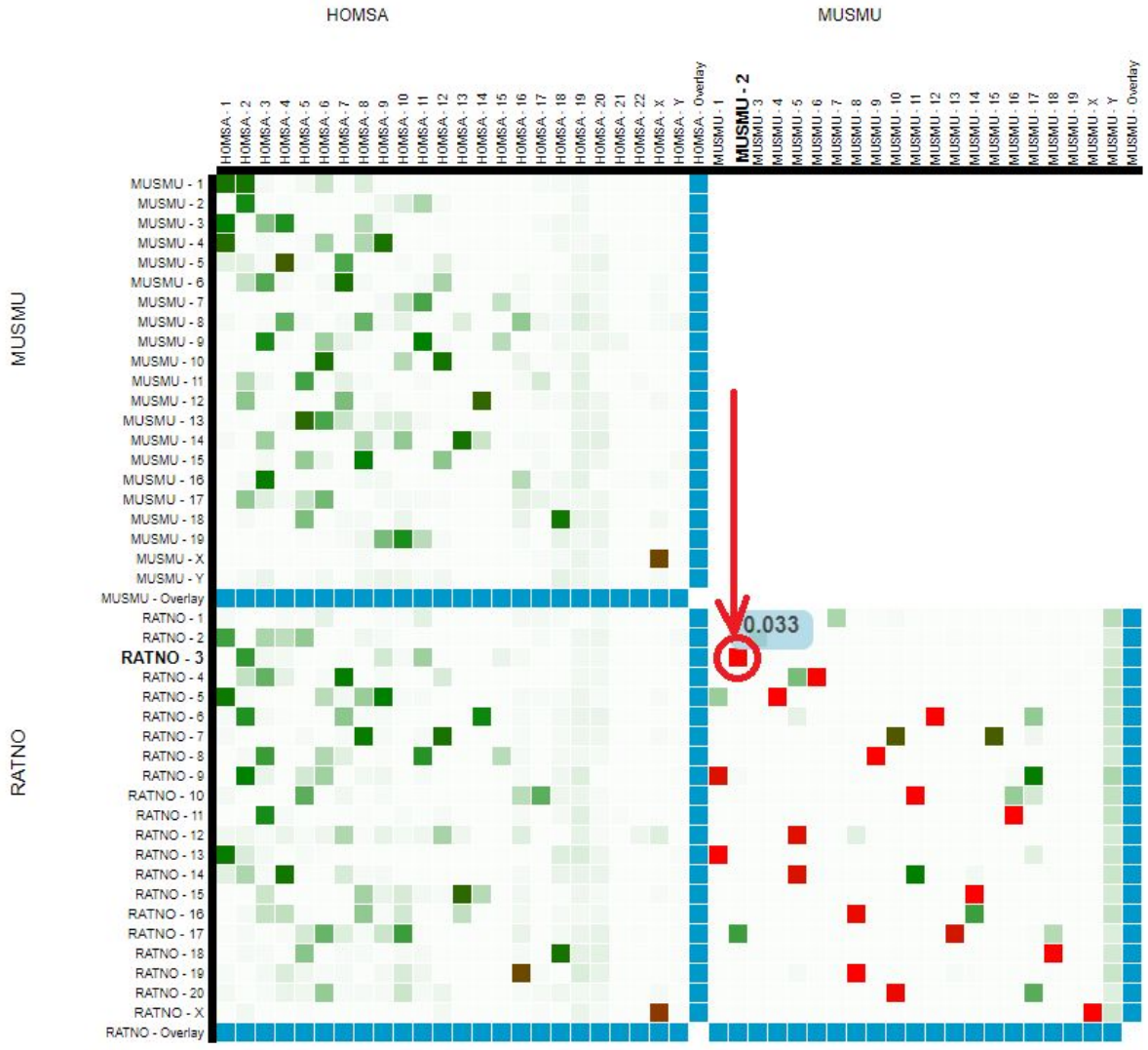
3. We continue to load the HOMSA *vs.* RATNO and the MUSMU *vs.* RATNO

genome comparisons. The loaded species table should like the following:

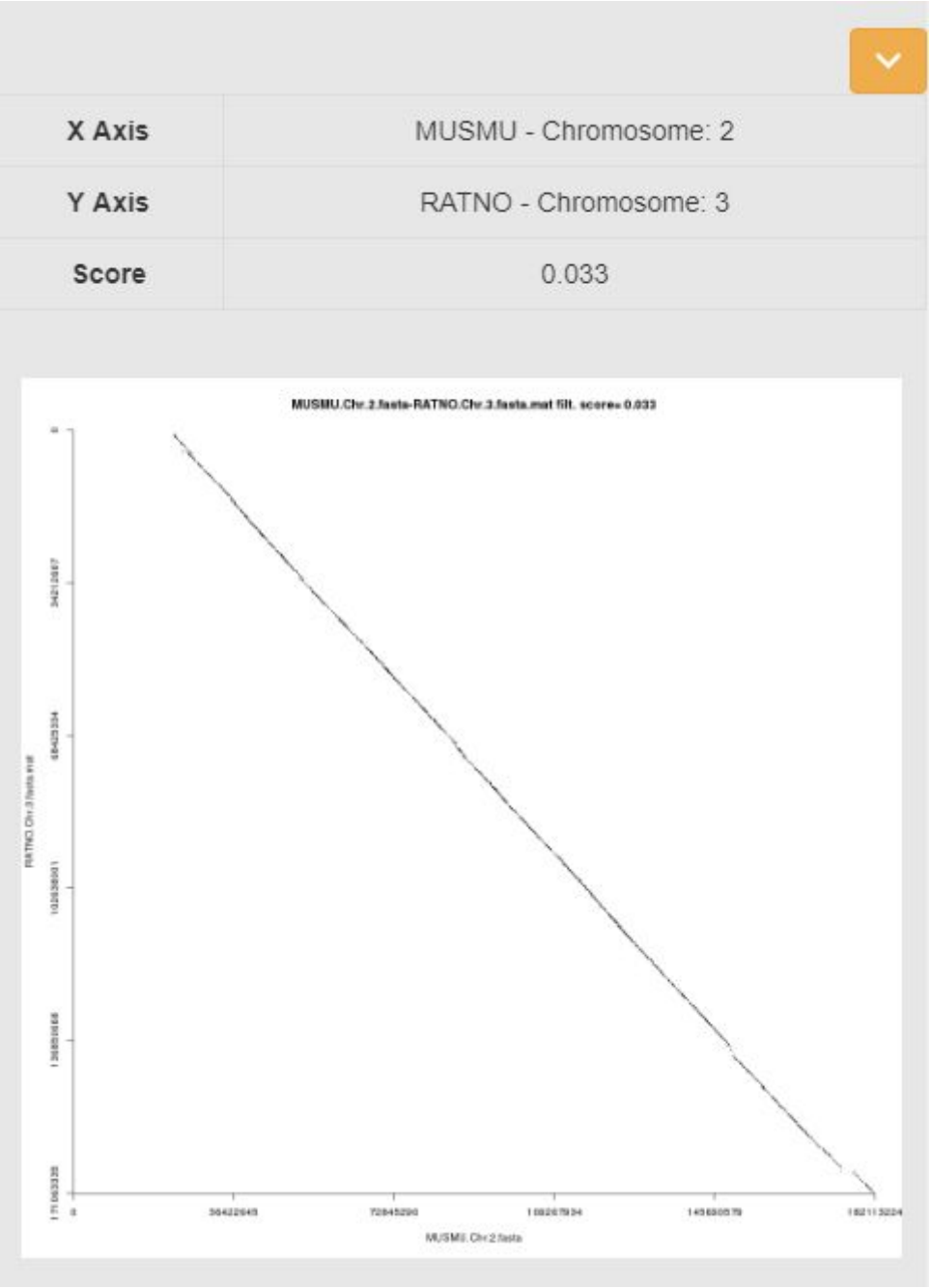
4. Once the visualization matrix is loaded, we can hover the mouse through the boxes and obtain the distance score calculated from Chromeister. The color is directly related to the metric of chromeister, enabling a quick detection of conserved signal comparisons. The score can go from 0 to 1, transitioning from light red to green to white as the score gets higher.



5. Now we will click on the box belonging to MUSMU's chromosome 2 and RATNO's chromosome 3.

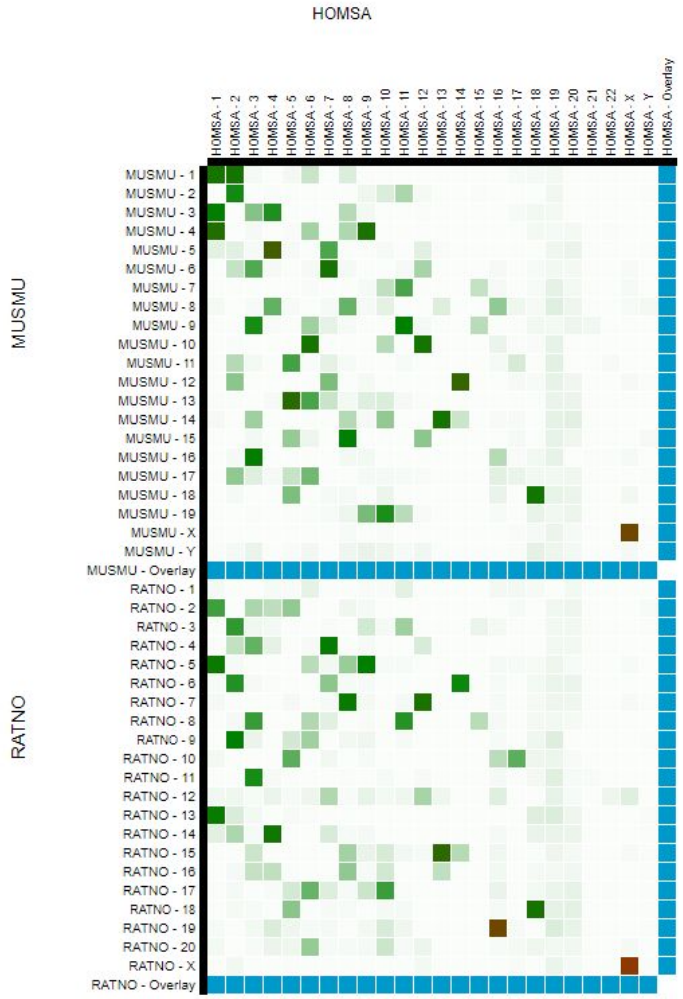


As a result, the pairwise comparison will be visualized in the side menu. The down arrow button ( ) will hide the currently visualized comparison from the side menu.

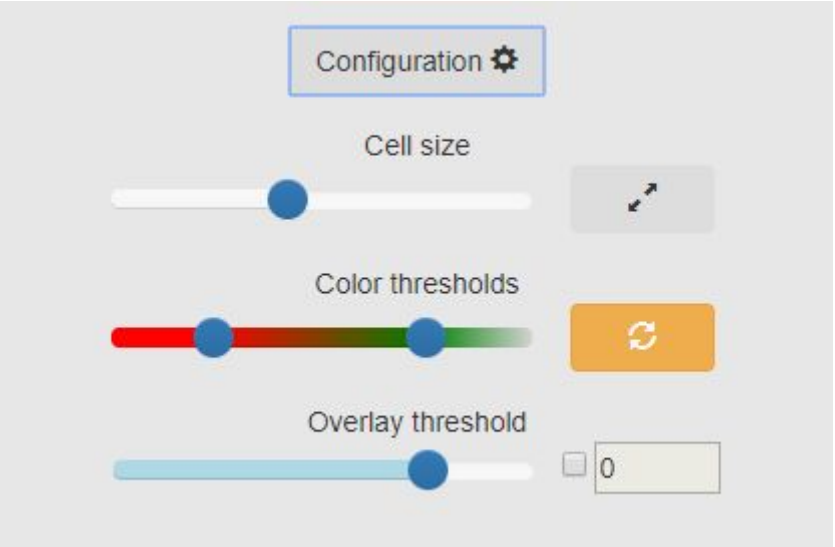


6. To demonstrate how to remove a full genome comparison from the visualization, we click in the 'X' button ( ) in the **side menu** table of currently loaded comparisons.

For this exercise we will remove the RATNO *vs.* MUSMU. As result, the visualization grid show will be the following:



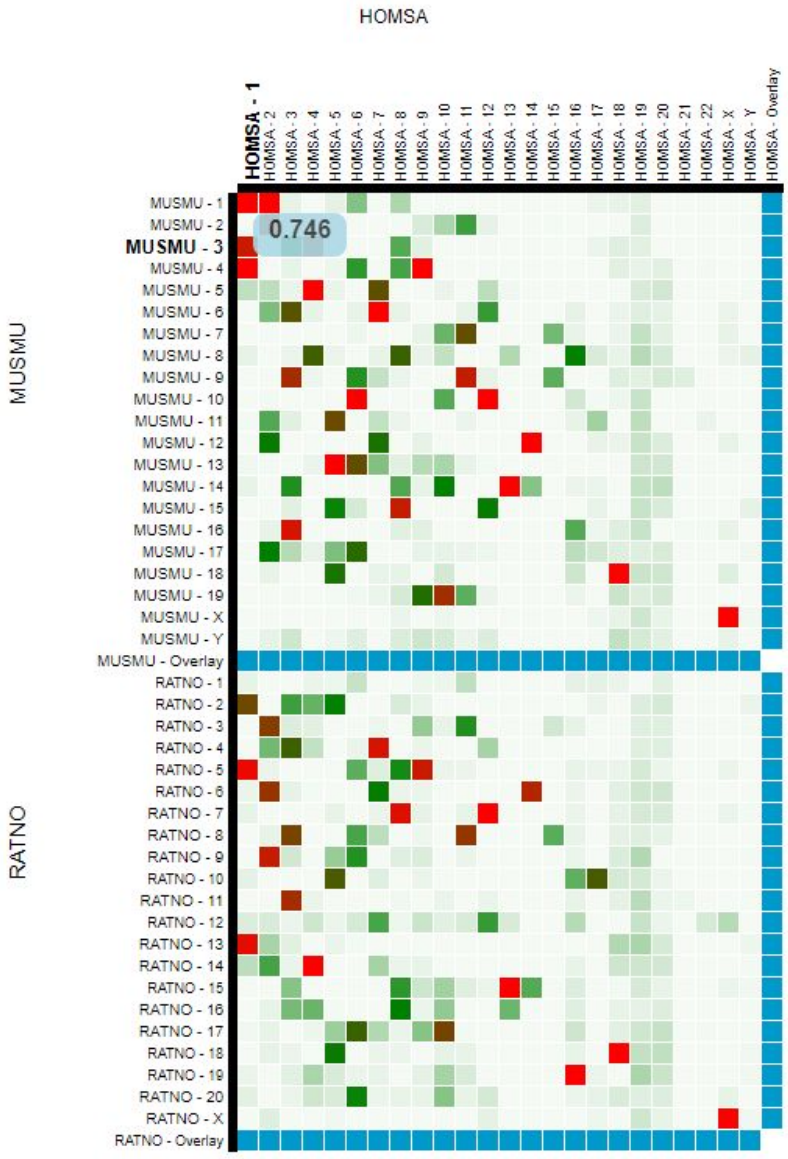
7. in order to take advantage of the additional features provided by XCout, we must open the Configuration menu by clicking on the corresponding () button. This will show the following additional information in the menu:



- **Cell Size:** The slider bar controls the size of each of the comparison boxes. The 'Fit to Screen' button () will adapt the visualization size to the current window size, and if clicked again it will return to the previous (original) size.
- **Color thresholds:** The slider bar allows the user to manually color gradient in the visualization. The 'Suggested Thresholds' button () will send automatically change it towards the most appropriate values to detect interesting comparisons with a quick glance. For this exercise we will click on it and the bar should adjust as the following:

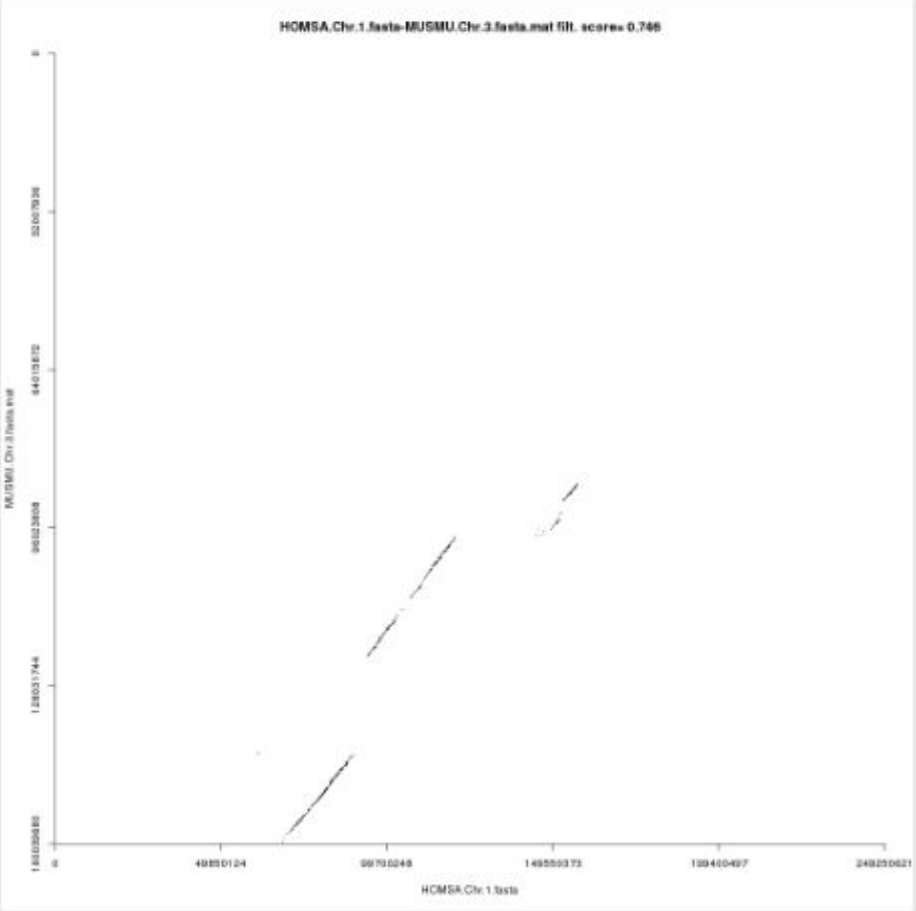
- **Overlay threshold:** This controls the number of comparisons to be overlaid when visualizing the Overlay of detected Collinear Synteny Blocks (CSB). We will explore this feature in the next exercise.

8. Having pressed on the 'Suggested Thresholds' button, the grid should now look like this:



9. Having changed from green to red, we can now quickly select a comparison of interest, in this case HOMSA chromosome 1 against MUSMU chromosome 3, and visualize it by clicking on it.

X Axis	HOMSA - Chromosome: 1
Y Axis	MUSMU - Chromosome: 3
Score	0.746



Exercise 2: Overlay multiple comparison

2.1 Introduction

In this exercise we will make use of the Overlay feature of XCout. This allows the user to overlay various pairwise comparisons of multiple chromosomes of a selected species to an specific chromosome of another organism. We will begin the exercise having loading a full genome comparison between two species and proceed to overlay multiple comparisons of one of the chromosomes. Additionally, we will explore additional characteristics of this feature.

2.1.a. Overlay comparisons

We will begin by loading the genomes comparisons between *Homo Sapiens* (HOMSA) against the *Mus Musculus* (MUSMU); apply the suggest thresholds to quickly detect a chromosome that has a visible signal with multiple chromosomes of another specie and proceed to overlay it.

2.1.b Overlay features

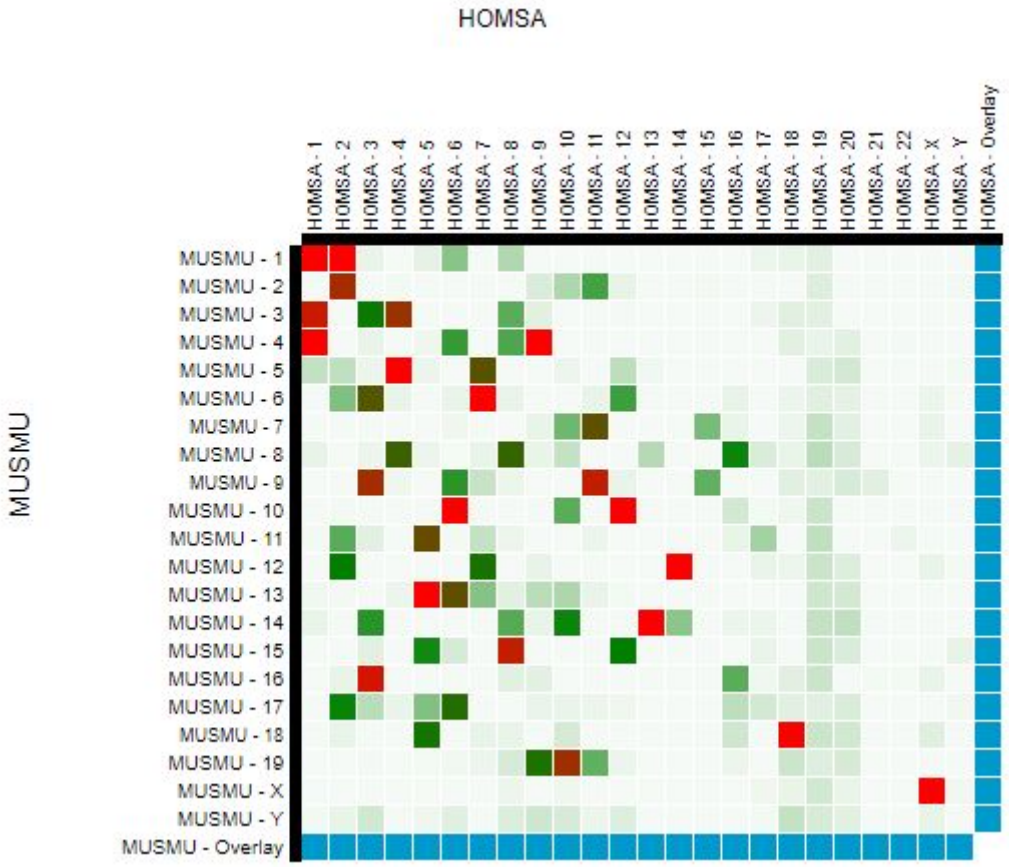
Afterwards we will take advantage of the overlay features within the configuration menu. Once we obtain an adequate result, we will proceed to export it to the 'Block Tracer' feature.

2.2 Exercise Development

1. Enter to http://pistacho.ac.uma.es/xcout

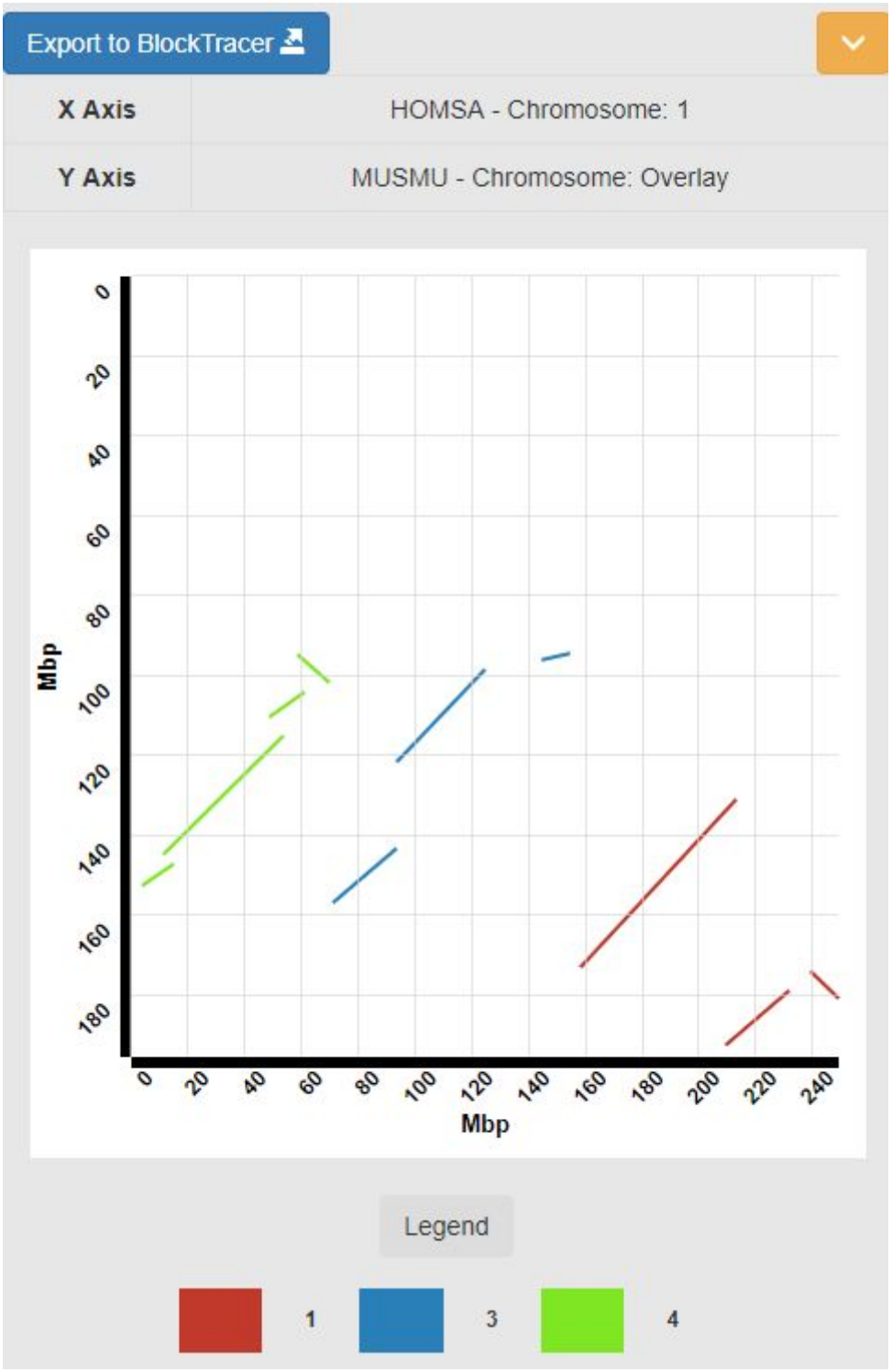
2. Once the webpage is loaded, we proceed to select the comparisons of interest from the side menu located to the left. In this case we will begin by selecting HOMSA and MUSMU and proceed to load it by clicking the 'Add' button (+).

3. Once loaded, we click on the configuration button (Configuration) and press on the 'Suggested Thresholds' button ( ) to enable an easier detection of comparisons with an interesting comparison. The resulting grid view should be the following:

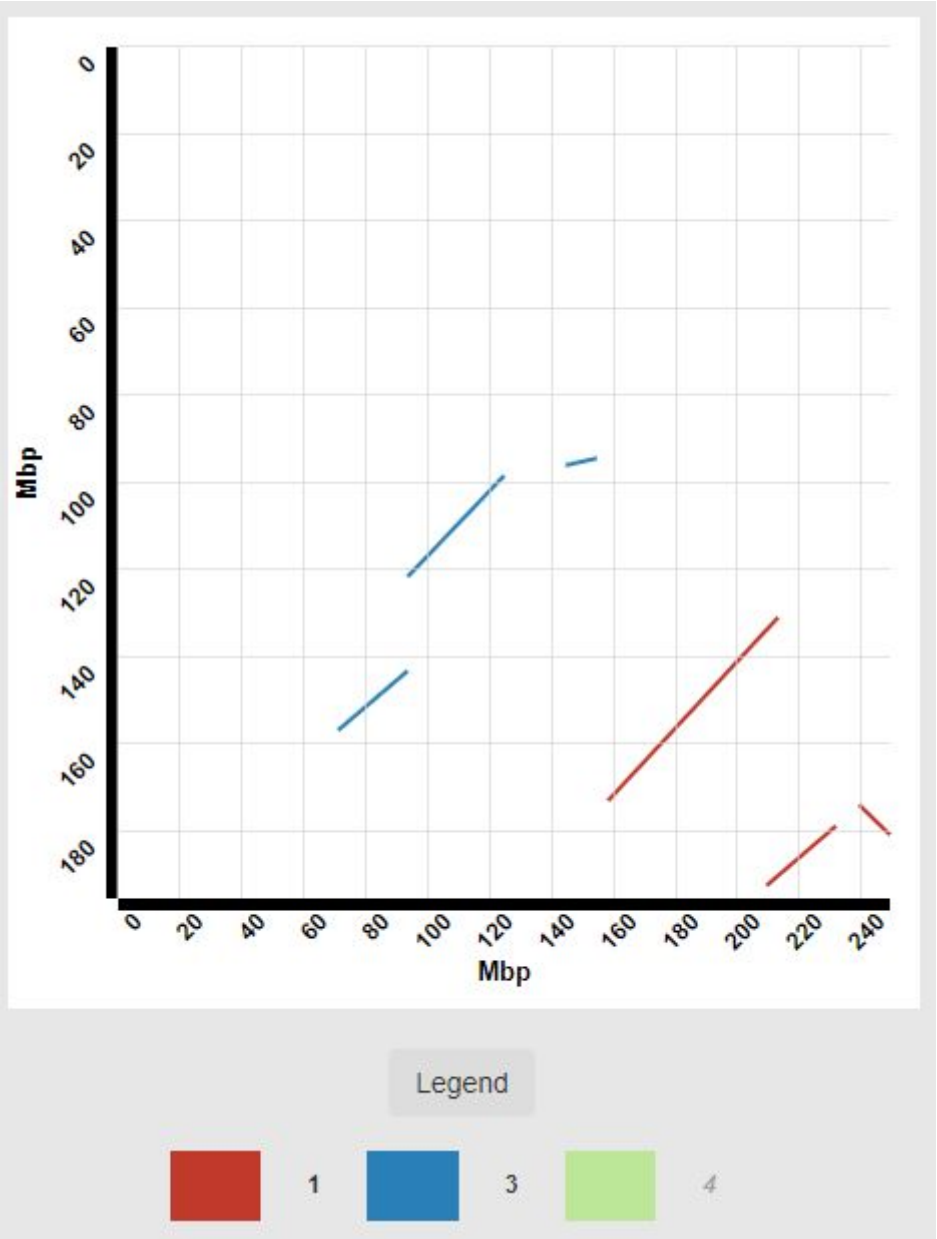


4. From this canvas we can observe that the HOMSA chromosome 1 has a significant signal (red boxes) in various chromosomes of MUSMU (1, 3 and 4). For this exercise we will overlay the MUSMU chromosomes against the HOMSA chromosome 1 by click the 'MUSMU - Overlay' blue box on the 'HOMSA - 1' column.

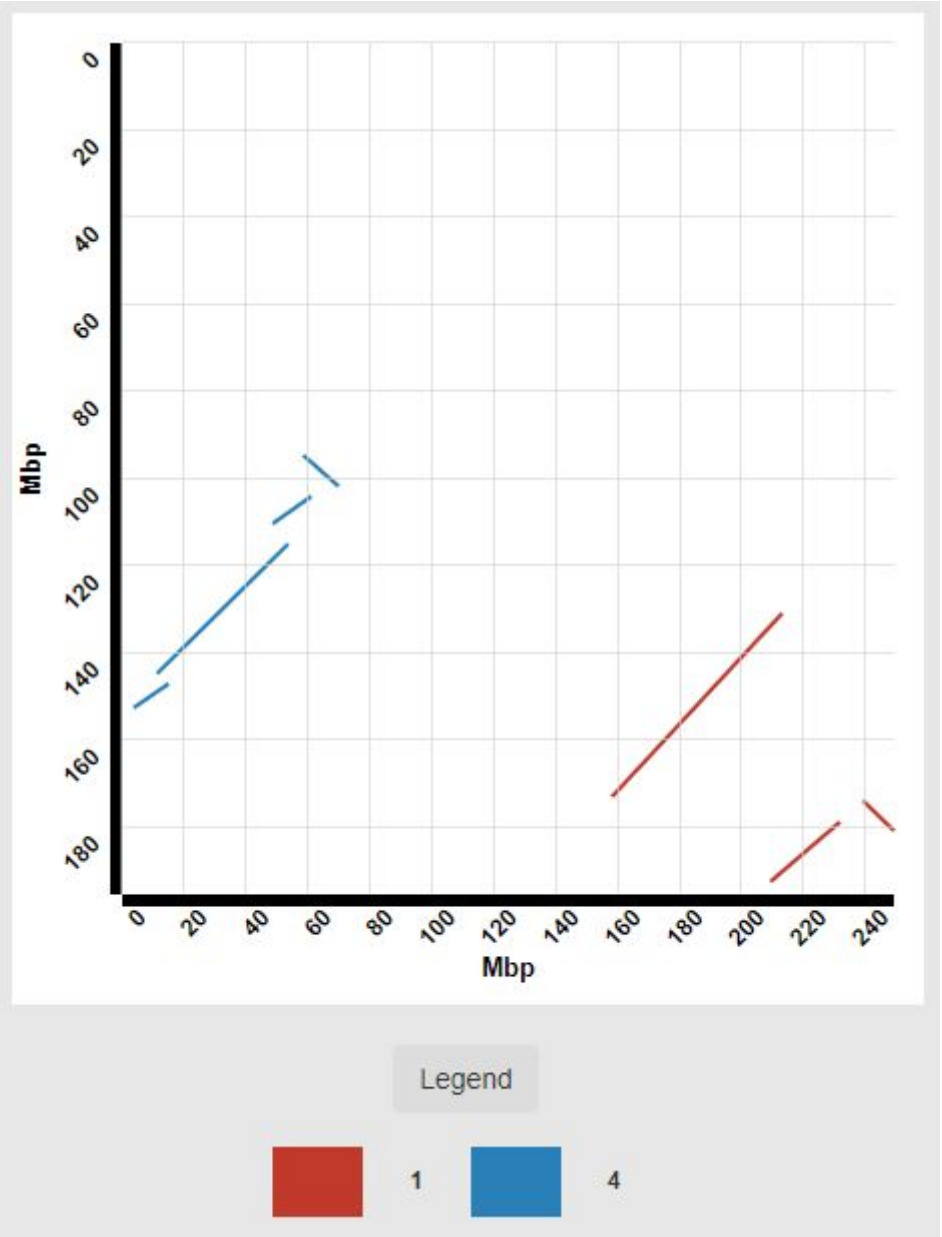
The resulting interactive canvas shall appear on the side menu. By now we are visualizing the common regions of DNA that are located in the chromosome 1 of the *Homo Sapiens* when compared to *Mus Musculus*. The resulting canvas is very interesting because it allows us to visualize different chunks of DNA (or groups of genes) that are common between species, but due to evolutionary events they have inherited independently into different chromosomes.



5. We can toggle the visibility of each of the overlaid comparisons by pressing on the corresponding chromosome in the 'Legend'. For instance, if we press on the green box (or the number 4), it will disappear from the canvas (as seen on the figure below) and if we press it again, it will reactive it.



6. We can observe that there are 3 overlaid pairwise comparisons against HOMSA's chromosome 1. By manipulating the configuration option of 'Overlay Threshold' slider, the comparisons overlaid will vary depending on the metric decided. As an alternative, the user may also activate the 'Overlay Threshold' checkbox in the configuration menu, and type the number of comparisons he wants to overlay, which will select the lowest score comparisons. For this exercise, we will check the box, select 2 comparisons to be overlaid and **press** the Enter key.



7. We can observe that the resulting overlay belongs to the two MUSMU comparisons with the lowest score: Chr 4 and Chr 1 with 0.683 and 0.705 distance metric respectively. Lastly we will proceed to click on the blue button on top of the overlay (Export to BlockTracer) to set up for the next exercise.

Exercise 3: Block Tracer

3.1 Introduction

In this exercise we will use the 'BlockTracer' feature of XCout that enables us to track Collinear Synteny Blocks (CSB) across multiple chromosomes amongst various species. This exercise is a continuation of the previous exercise (Exercise 2) in which the user finishes by exporting the overlaid comparison to the BlockTracer tab.

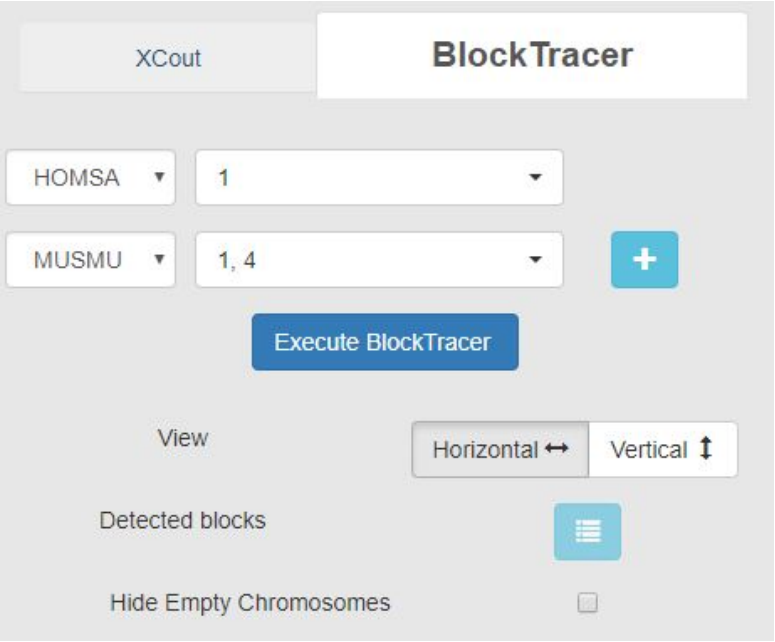
The user will execute Block Tracer and explore the features offered by XCout.

3.2 Exercise development

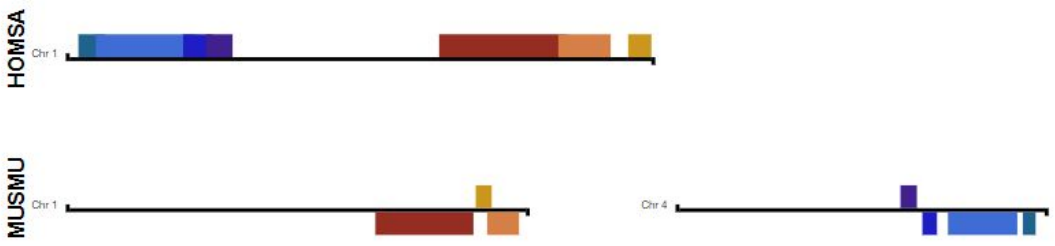
1. Two options:

a) Enter to http://pistacho.ac.uma.es. Press on the 'BlockTracer' tab, select HOMSA's chromosome 1 in the first row and MUSMU's chromosomes 1 and 4.

b) Finish the second exercise last step.



2. Press on the 'Execute BlockTracer' button. After loading, the result should look as the following:



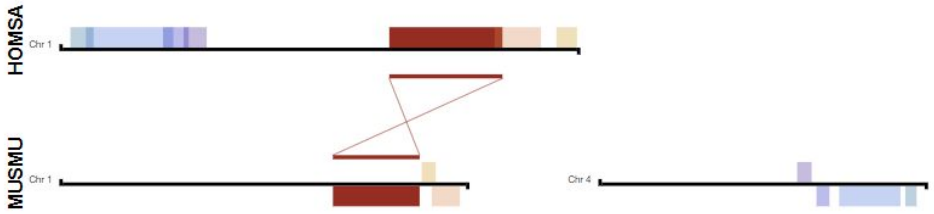
The general guidelines to interpret the visualization are the following:

- Each chromosome can have blocks above or below the chromosome line. Their position implies the blocks relationship to the previous organism: above means the strand is forward (5'-3') and below means it is reversed

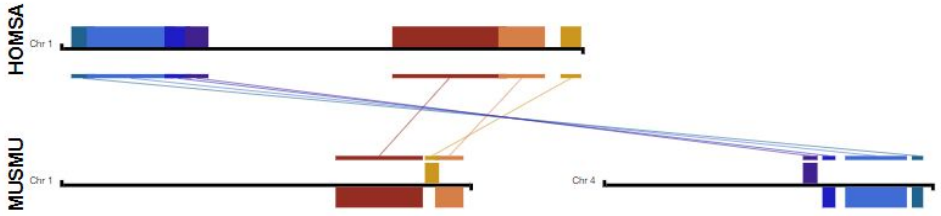
(3’-5’).

- All the blocks in the first species will always be forward. Hence, it will be our baseline reference genome.
- Analogous colors tend to belong to the same pairwise comparison

3. Once blocks have been traced, the user is capable of selecting each of the blocks to focus on them. Moreover, connection lines will appear, if they are crossed it means the block of the latter specie is reversed in reference to its previous one, just like in the following example:



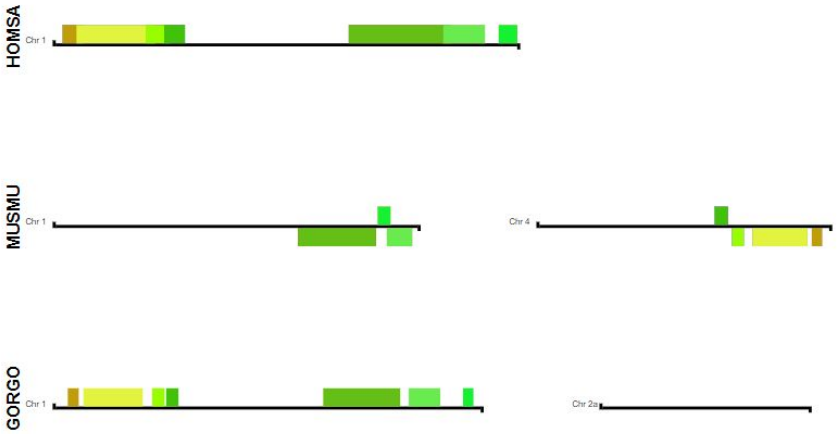
4. On the ‘Configuration’ options, users may click on the ‘Resize’ button ( ) to fit the visualization into the current window size, and can also click on ‘Show connection lines’ button ( ) to show all the connection lines, simplified (not doubled as a block selection line) as in the following image:



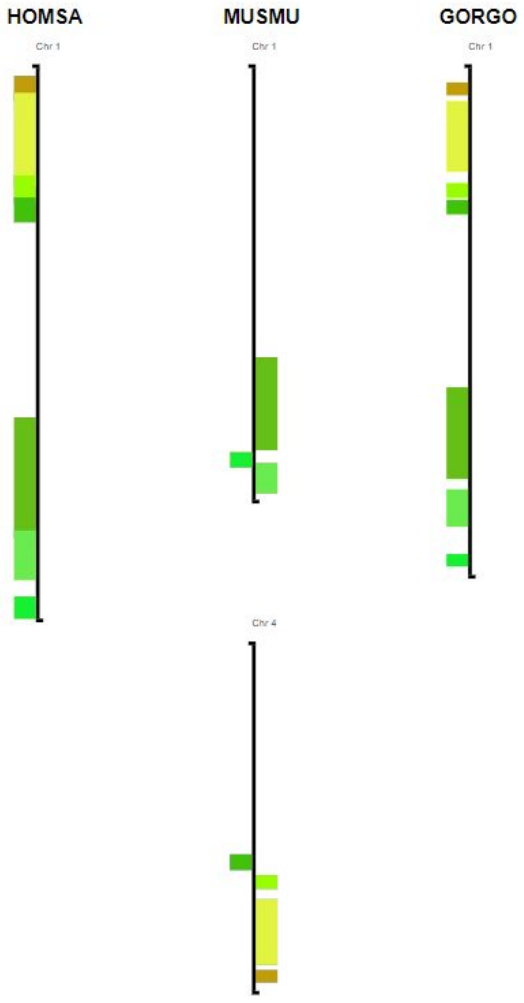
To turn off this option click on a block or on the previous button ( ) which now represents ‘Hide connection lines’.

5. Now we will load another specie by clicking on the ( ) button and select GORGO’s chromosome 1 and 2 for the third row.

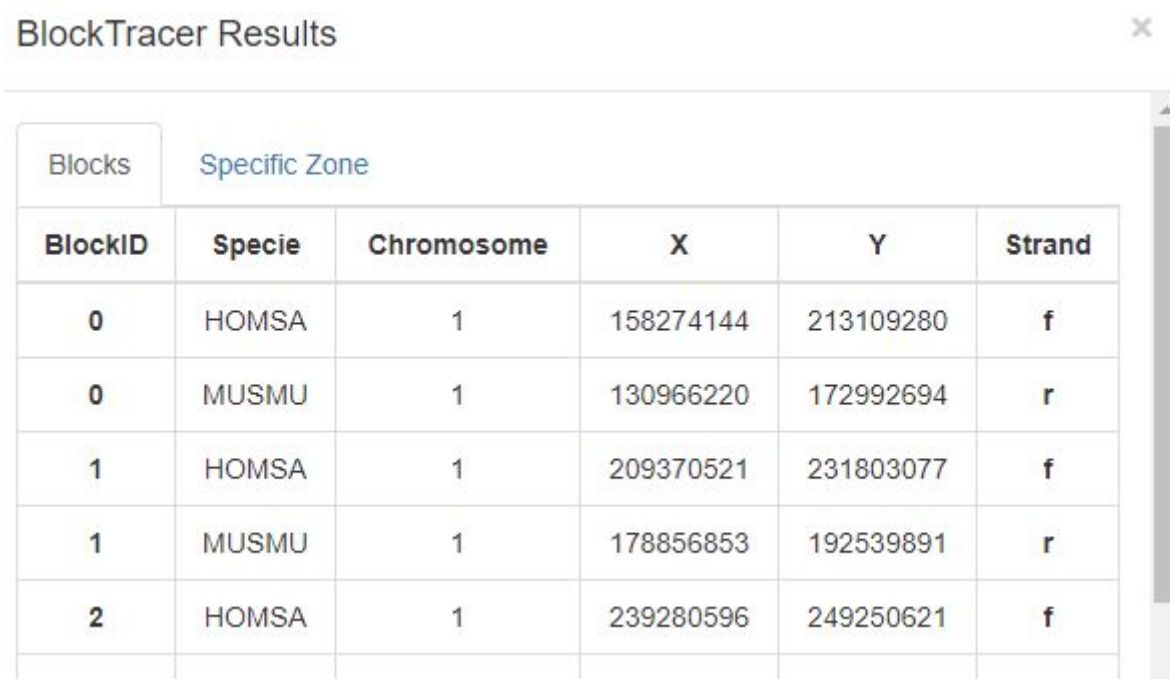
6. We proceed to ‘Execute BlockTracer’ and obtain the following results



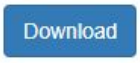
7. To explore additional functionalities of XCout's BlockTracer, we will click on the 'Vertical' button () from the View row in order to visualize the BlockTracer results from a vertical perspective. **Additionally**, we will check the 'Hide Empty Chromosomes' checkbox in the configuration menu to remove the chromosomes that do not have a block traced from the canvas.



8. Lastly, we will press on the ‘Detected blocks’ button () and a modal with two tables will appear, one with the information of the blocks that have been traced and another one which contains each chromosome’s specific regions (View Annex 1 for detailed information).



BlockID	Specie	Chromosome	X	Y	Strand
0	HOMSA	1	158274144	213109280	f
0	MUSMU	1	130966220	172992694	r
1	HOMSA	1	209370521	231803077	f
1	MUSMU	1	178856853	192539891	r
2	HOMSA	1	239280596	249250621	f

The user can download the currently active table by pressing the ‘Download’ button () to further analyze these results.

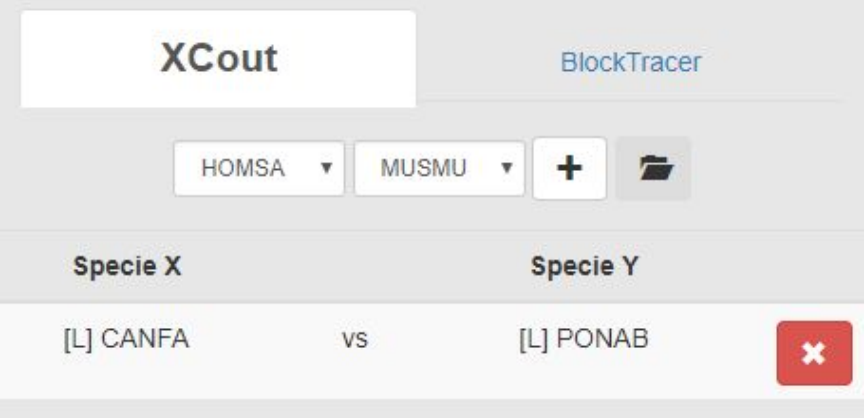
Exercise 4: Load local genome comparison

4.1 Introduction

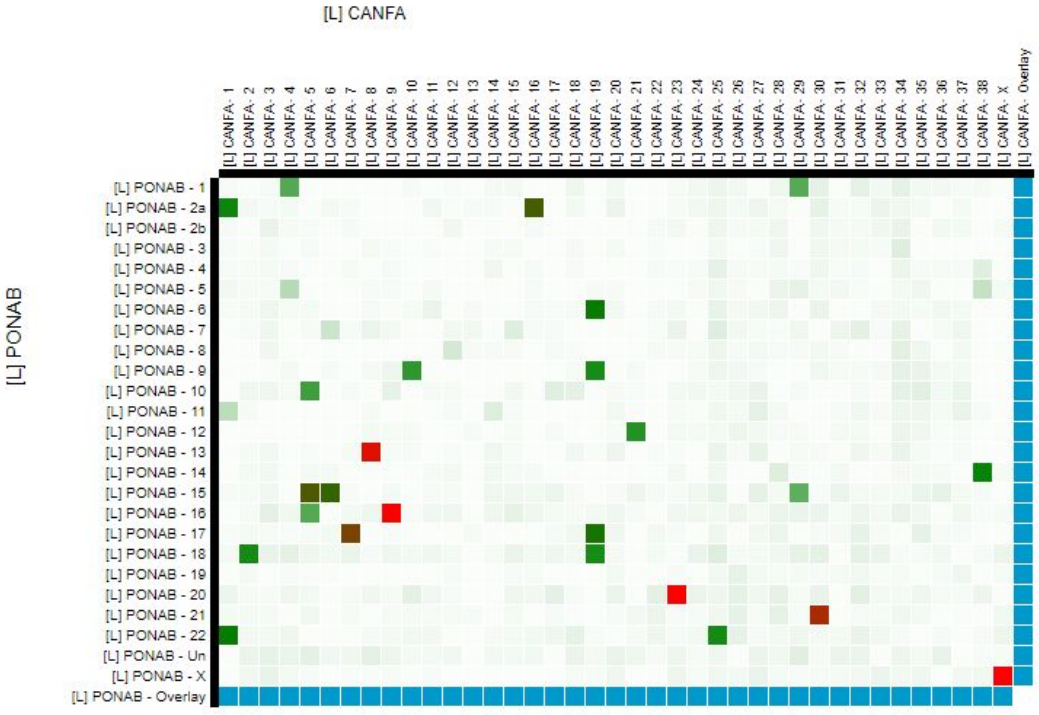
In this exercise we are going to learn how to load a local genome comparison performed with Chromeister. To perform this exercise you need to download the 'canfa_ponab.zip' file from http://mango.ac.uma.es/compartir/XCout/genome-comparisons/ and **unzip** it afterwards.

4.2 Exercise development

1. Being on the 'XCout' tab we will proceed to load our local genome comparisons from Chromeister by clicking on the 'Load local' button ( ) and selecting the folder which contains all the comparisons. The species loaded locally will have a '[L]' tag identifier, hence the loaded species table should look like this.



2. We perform an automatic threshold color threshold ( ) and visualize the results:



Annex 1:

Detected Blocks: Data Structures

- **Blocks:** CSB that have been tracked along the combinations of chromosomes from multiple species. Composed of:
    - **BlockID:** Block identifier
    - **Specie:** Block belongs to this specie
    - **Chromosome:** Block belongs to this chromosome
    - **X:** Nucleotide position where current zone begins
    - **Y:** Nucleotide where current zone ends
    - **Strand:** Direction of the block
- **Specific Zone:** Regions of the chromosome which do not belong to any block traced
    - **ZoneID:** Zone identifier
    - **Specie:** Zone belongs to this species
    - **Chromosome:** Zone belongs to this chromosome
    - **X:** Nucleotide position where current zone begins
    - **Y:** Nucleotide position where current zone ends