

Aligning *C. elegans* and *C. briggsae* orthologues

Sockeye How-to: **Using Sockeye to Perform Multiple Alignments between *C. elegans* and *C. briggsae* orthologues.**

Author: Steven Jones

Date: January 19, 2004.

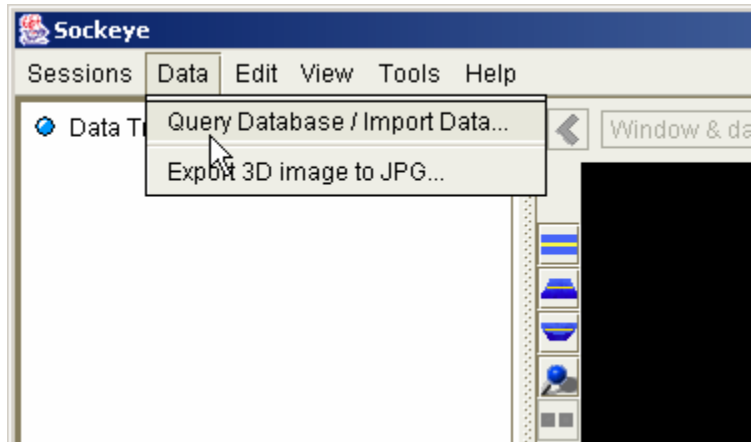
Information about how to install Sockeye:

<http://www.bcgsc.ca/gc/bomge/sockeye>

Description: This How-To describes the steps required to align orthologous genes between *C. elegans* and *C. briggsae* and subsequently view the aligned sequences. Alignment algorithms currently available to achieve this include LAGAN and CLUSTALW.

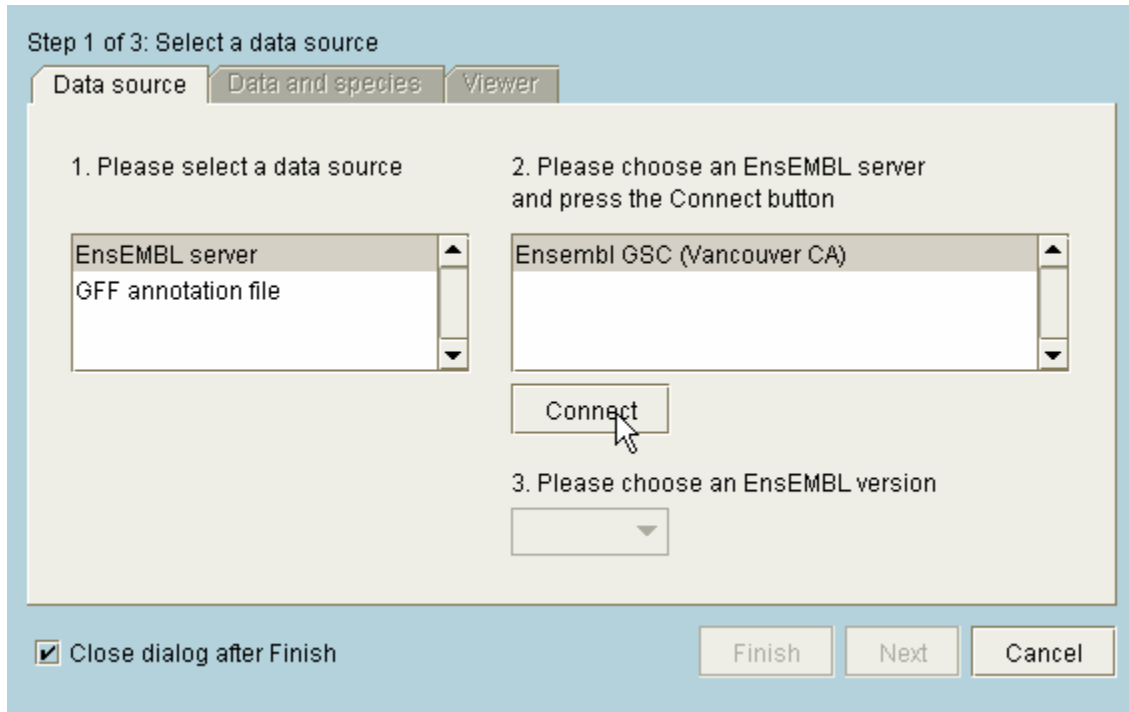
## Aligning *C. elegans* and *C. briggsae* orthologues

- 1) Connect to an Ensembl database by selecting the Data->Query Database/Import Data.



Aligning *C. elegans* and *C. briggsae* orthologues

2) Then connect to an available Ensembl database server.



- 3) Then select the “Data and Species” tag and then select *C. elegans*. Then choose “Get annotation by Ensembl ID”. This will therefore allow you to simply enter the Wormbase name for the gene in the adjacent box. In this example we will use the *myo-2* gene T18D3.4 (for more information on *C. elegans* genes see the wormbase database [www.wormbase.org](http://www.wormbase.org)). This box also allows you to choose how much DNA sequence should be taken from either side of the transcript. Since in this example, we want to align the upstream region of the gene change the 5' value to 5Kb.

Retrieve Data

Step 2 of 3: Select data set and species

Data source | Data and species | Viewer

1. Please select

- Get annotations by range
- Get annotations by Ensembl ID
- Search Ensembl

2. Please enter stable gene id (eg, B0513.6)

C.elegans (v.16.102) - 1 matches

T18D3.4

T18D3.4

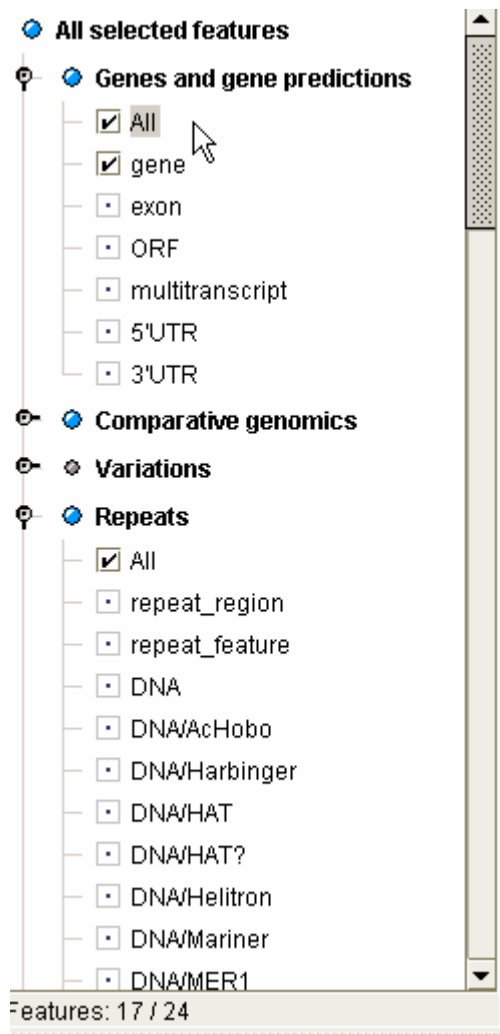
Plus 5' 5 000

Plus 3' 1000

Close dialog after Finish

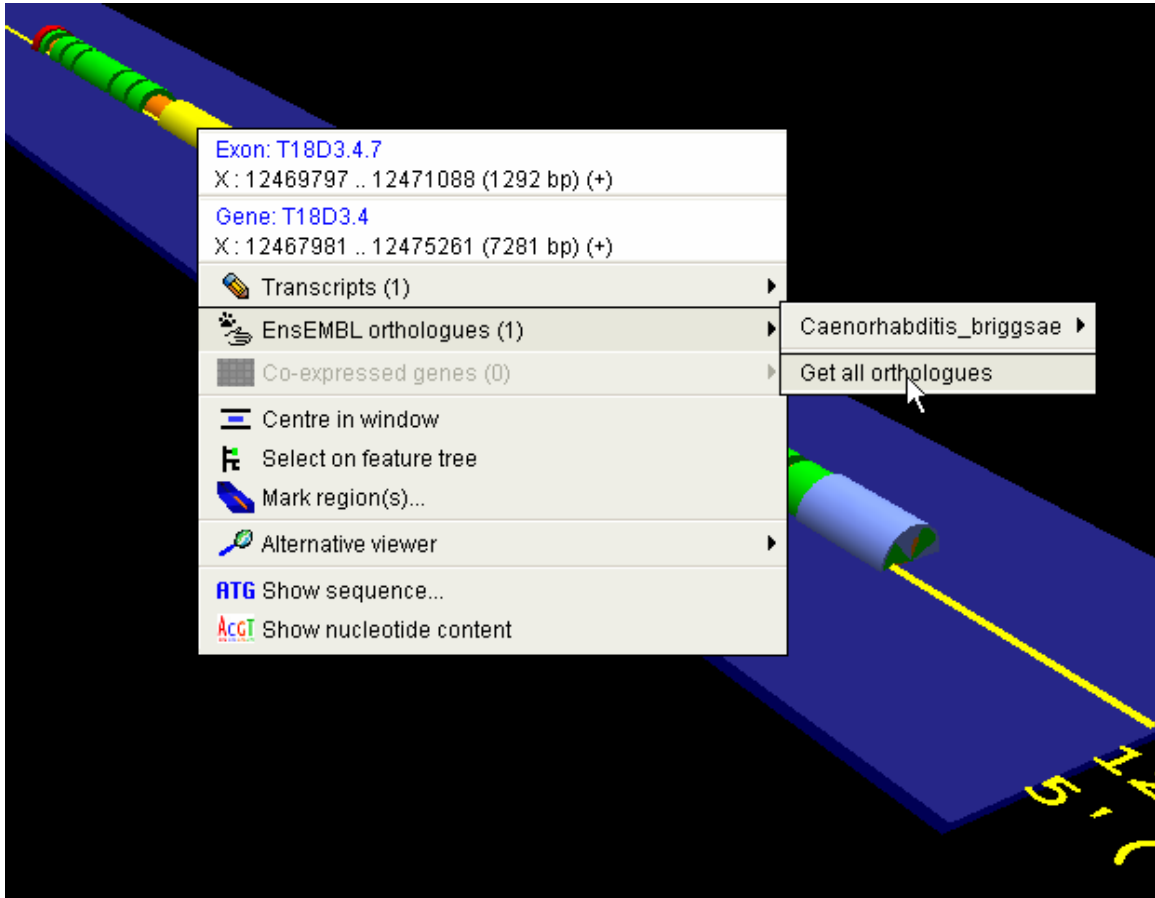
Finish Next Cancel

- 4) This will now show the gene T18D3.4 in the graphics plane. We can now ask for more genomic features to be displayed. In this case we want to see all the repetitive elements and the gene associated features. This is achieved by selecting “All” from the different expandable feature trees to the left of the genomic display. Click on the small “magnifying glass” icon to expand. Click on the blue button to turn on or off all the features in a tree. To navigate the 3D environment, the left button will rotate the image, the middle button will allow you to zoom in and out and the right button will allow you to grab and move the image.



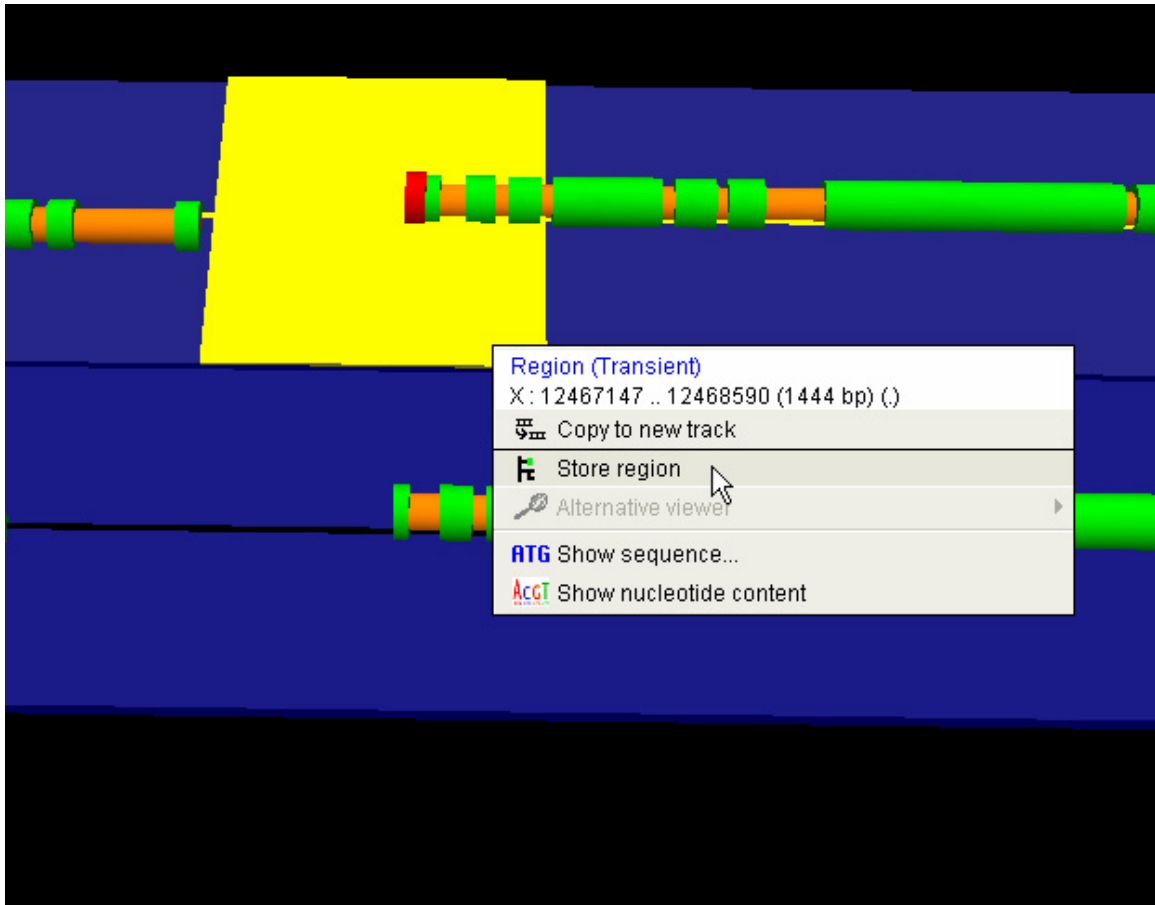
Aligning *C. elegans* and *C. briggsae* orthologues

- 5) We now need to get the orthologous gene from *C. briggsae*. This will be the orthologue as determined within the Ensembl compara database. We call up the orthologues by right mouse clicking on our gene of interest, T18D3.4



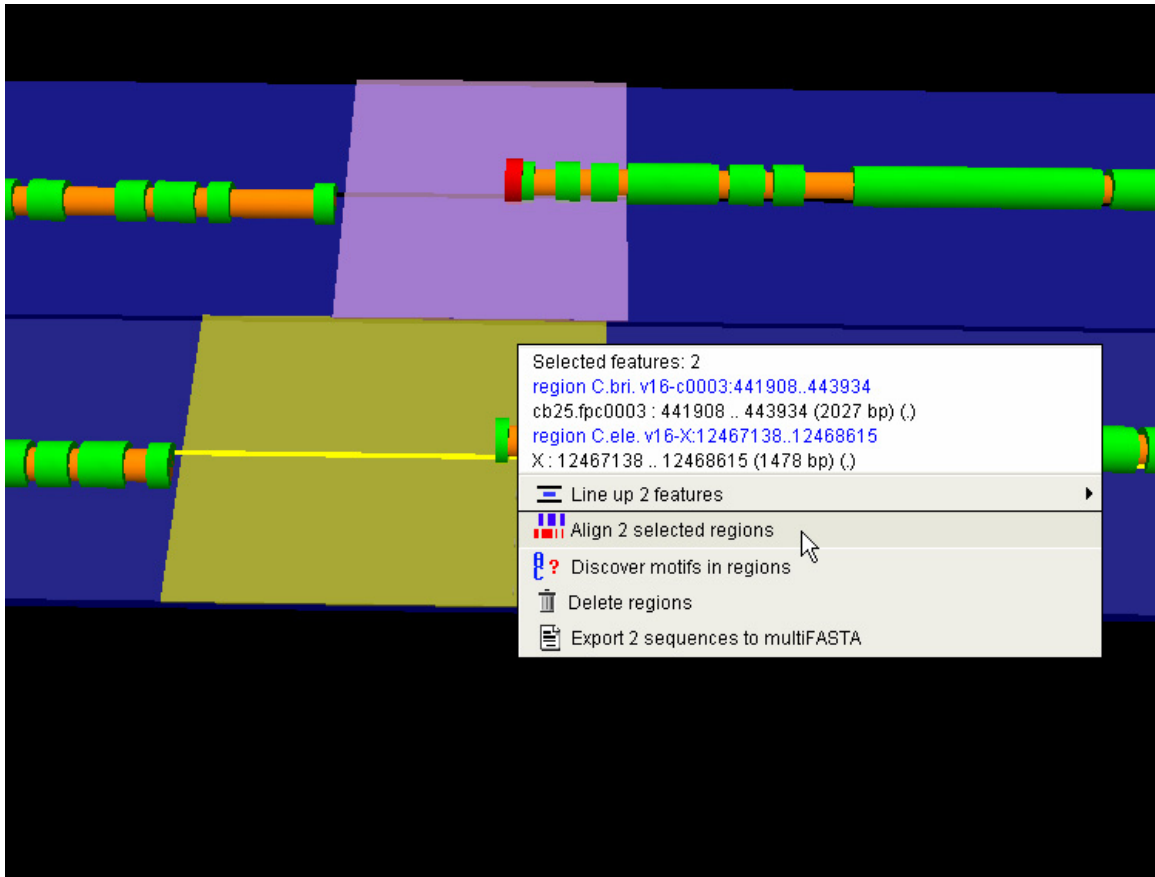
## Aligning *C. elegans* and *C. briggsae* orthologues

- 6) Selecting the regions to align. This is simply done by dragging the mouse with left button depressed over the regions to be aligned and storing the region with the right mouse button. Repeat this process for both the *C. elegans* and *C. briggsae* genes. For a sanity check, it is always a good idea when aligning upstream regions to include a few exons from each gene to indicate that the alignment behaves as expected [Problems often occur since the 5' and 3' UTRs are often well defined in *C. elegans* but not in *C. briggsae*].



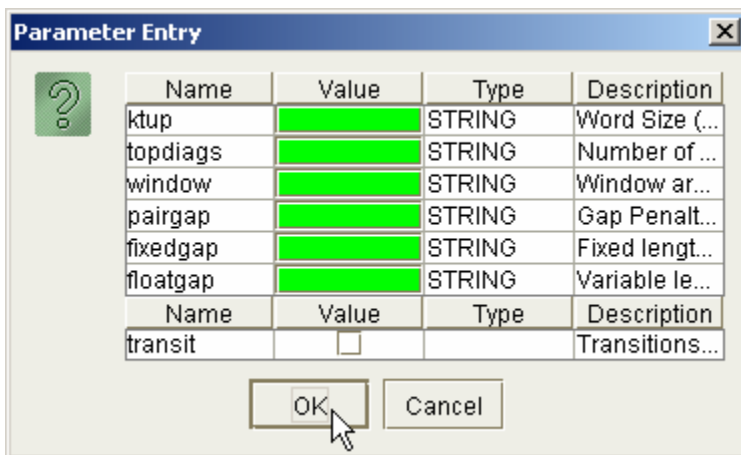
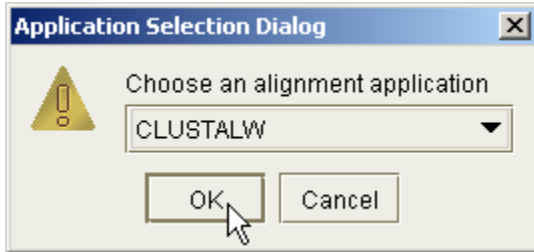
## Aligning *C. elegans* and *C. briggsae* orthologues

- 7) Once the regions to be aligned have been stored, we select them both by clicking with the left mouse button and **holding down the CTRL-key** allowing multiple regions to be selected. Once the regions are selected we can pull down a menu using the right mouse button and choose to align the sequences.



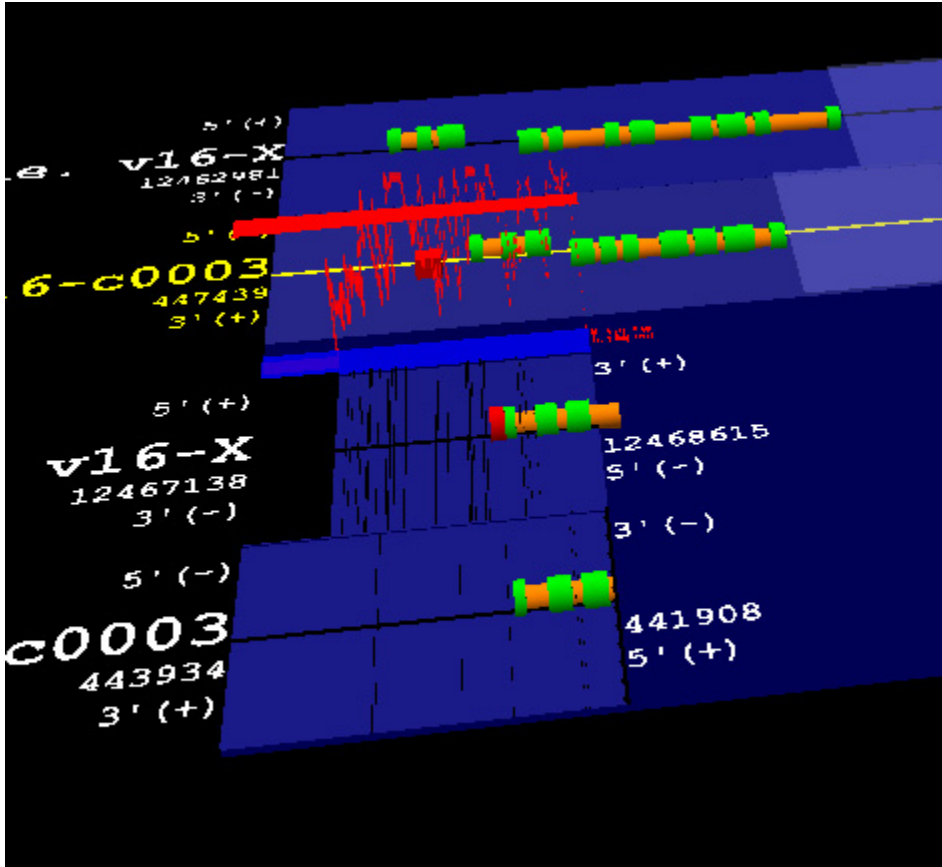


- 8) Choose the alignment algorithm. You will also be presented with a box to change any parameters that the alignment algorithm can accept. A green box indicates that a default value is being used.



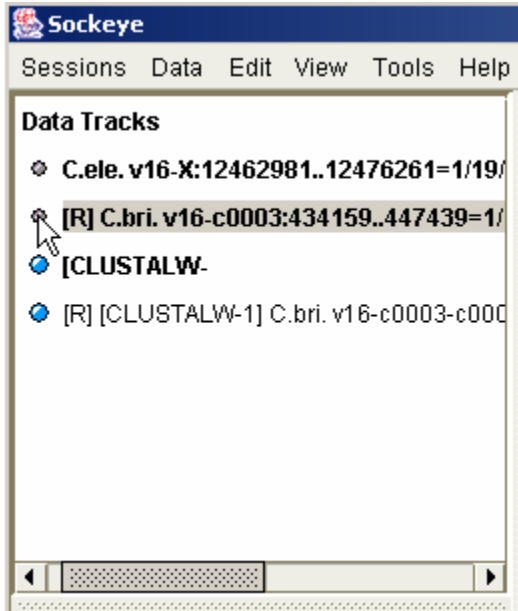
Aligning *C. elegans* and *C. briggsae* orthologues

- 9) Once the alignment is submitted it will take a few seconds for the alignment to proceed and the results returned. The alignment will be returned onto a new track.

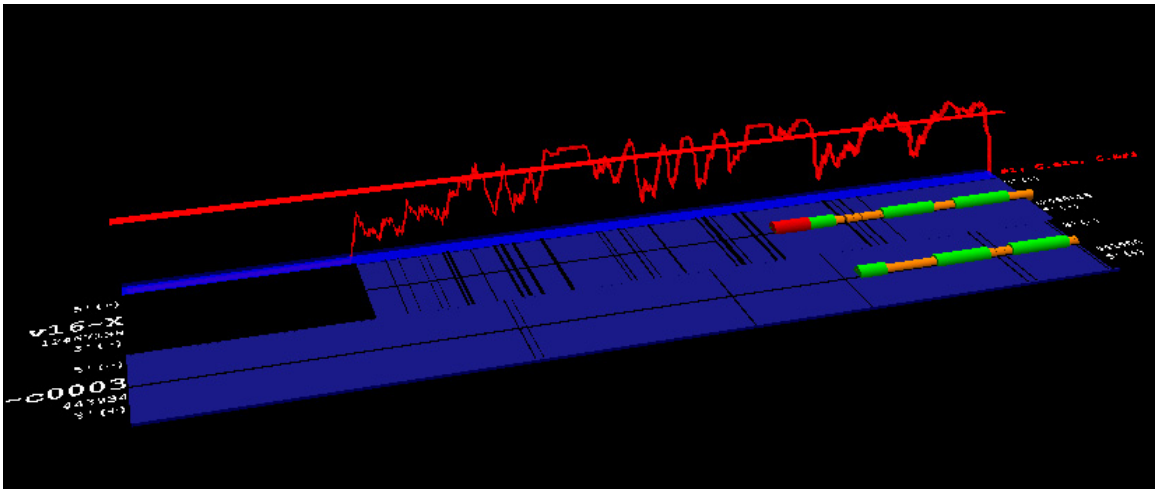


## Aligning *C. elegans* and *C. briggsae* orthologues

10) We can change the display so that only the aligned regions are displayed, simply by turning off the two parent tracks, by clicking on their blue buttons.

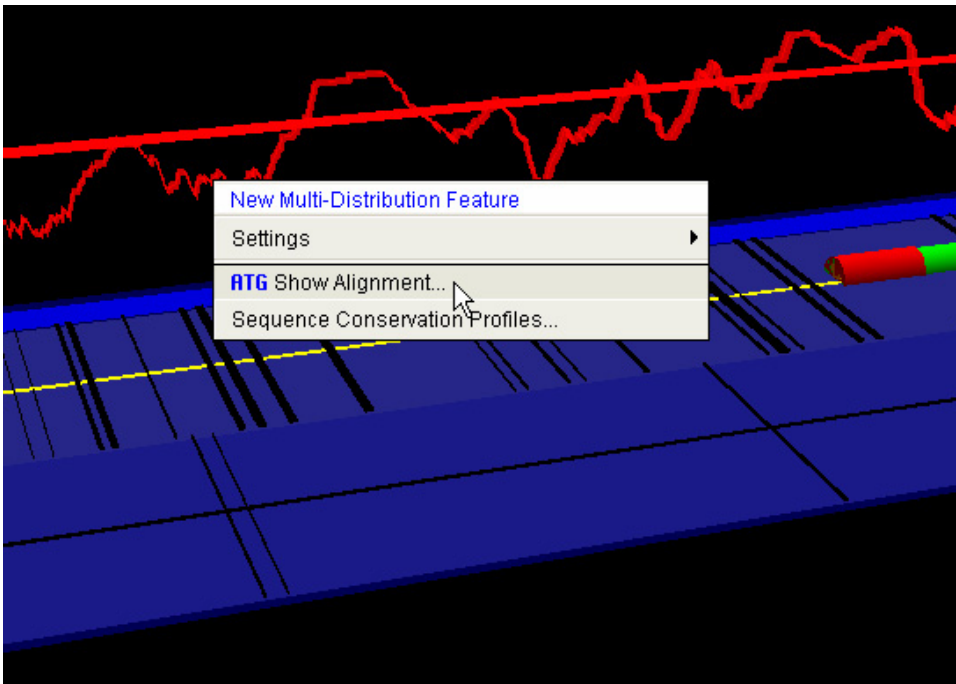


This now shows the alignment in more detail. This can be rotated and viewed using the standard 3D navigation tools provided by the mouse buttons. Gaps in the alignment are shown on the DNA plane.

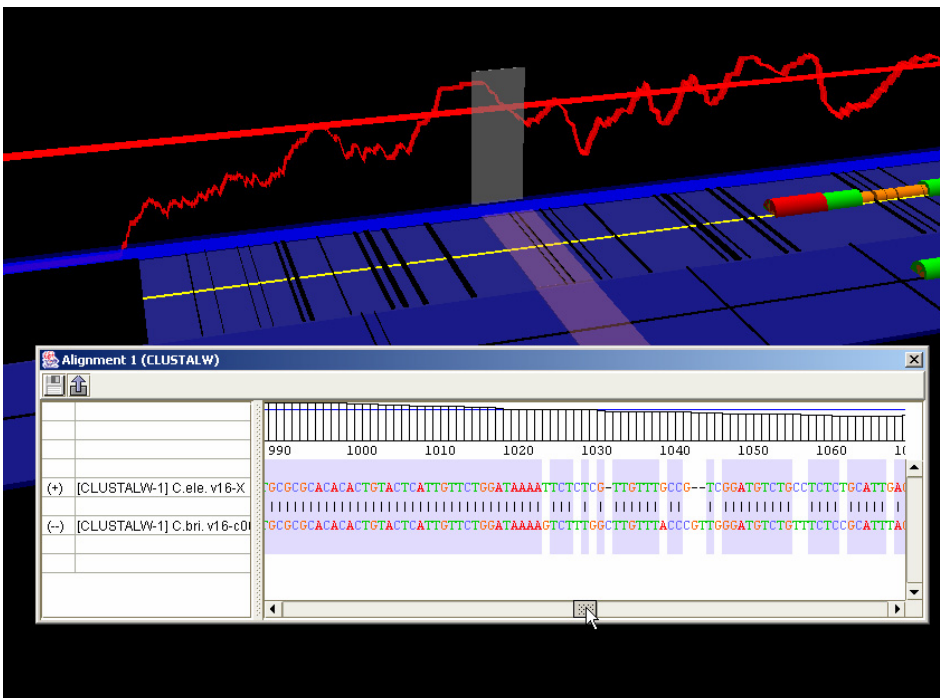


## Aligning *C. elegans* and *C. briggsae* orthologues

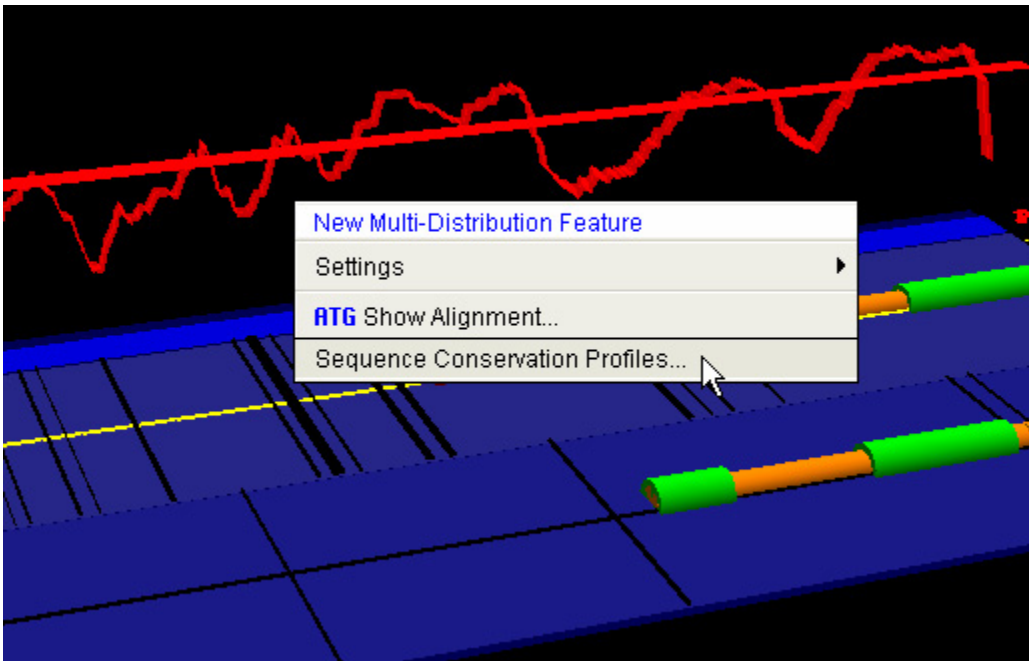
- 11) Exploring the actual alignment can be done by right-mouse clicking on the alignment graph. This will bring up a menu from which we select “ATG Show alignment”.



This will bring up the alignment tool viewer. The shaded region on the alignment graph indicates the region being shown in the sequence alignment viewer.



12) Changing the way that the alignment graph is plotted. This can be done by selecting the “Sequence Conservation Profile” menu item.



This box can, for instance, allows the adjustment of the window size, the way that the graph is scored and the height of the threshold line.

