

## jORCA new Features (updated May 2010)

### Pipeline execution

In the last jORCA release (November 2009), a user could create a pipeline from the execution LOG just dragging and dropping it into a Canvas. From this created pipeline, the user could export it into a SCUFL file and execute it with external programs such as Taverna Workbench ([www.taverna.org](http://www.taverna.org)).

Now, jORCA is able to create and execute pipelines in a single step. There are two ways to create a pipeline in jORCA:

- From the execution LOG
- Using Magallanes (<http://www.biomedcentral.com/content/pdf/1471-2105-10-334.pdf>) workflow creation tool.

Once created the pipeline, jORCA can execute and save it in the favorite list.

### *How to create pipelines using the execution LOG*

Imagine that you have executed a set of tools, which inputs and output are compatible. Just dragging and dropping them to the LOG canvas in the same execution order produces a pipeline.

Let's illustrate it with an example, execute the following tools (use INB testing repository-[chirimoyo.ac.uma.es](http://chirimoyo.ac.uma.es)) :

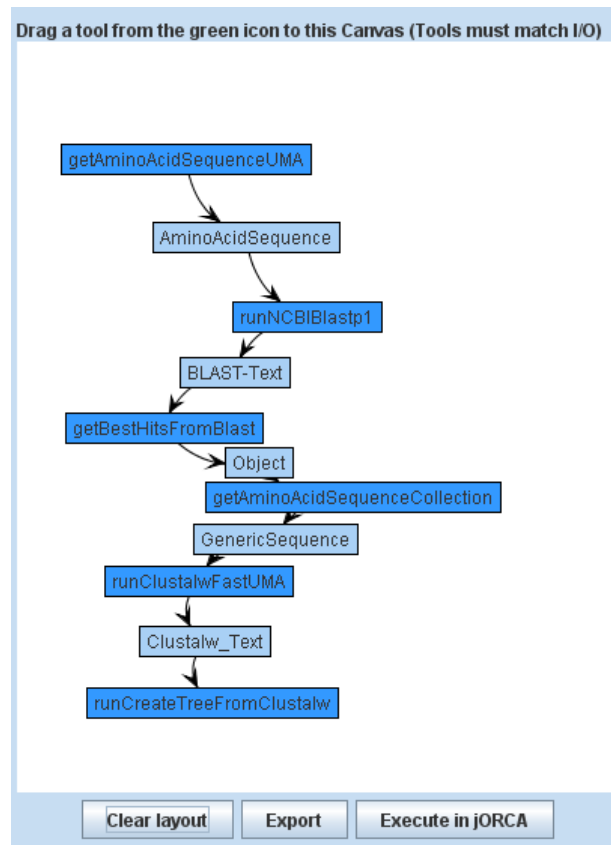
- Execute **getAminoacidSequenceUMA** with: DQA1\_HUMAN.
- Use the output to execute **runNCBIBlastp1** : (right click on the file system area and click on Exact compatible tools > runNCBIBlastp1)
- Use the output to execute **getBestHitsFromBlast**: (right click on the file system area and click on Exact compatible tools > getBestHitsFromBlast)
- Click on **getAminoacidSequenceCollection** on the favorites list and drag the result file from the getBesthitsFromBlast tool.
- Click on **runClustalwFastUMA** on the favorites list and drag the result file from the getAminoacidSequenceCollection tool.
- Use the output to execute **runCreateTreeFromClustalw**: (right click on the file system area and click on Exact compatible tools > runCreateTreeFromClustalw)



Now click on the log icon and open it, the executed tasks will be shown:

Executed tasks in jORCA			
<b>Repository:</b> INB Testing repository chirimoyo.ac.uma.es <b>Parameters:</b> [null,DQA1_HUMAN] <b>Status:</b> FINISHED			
<b>Tool:</b> runNCBIBlastp1 <b>Date:</b> 26 Apr 2010 11:09:33 GMT <b>Where:</b> chirimoyo.ac.uma.es <b>Repository:</b> INB Testing repository chirimoyo.ac.uma.es <b>Parameters:</b> [null,D:\JorCA Files\getAminoAcidSequenceUMA.2010.04.26.13-08-1] <b>Status:</b> FINISHED			
<b>Tool:</b> getBestHitsFromBlast <b>Date:</b> 26 Apr 2010 11:10:39 GMT <b>Where:</b> chirimoyo.ac.uma.es <b>Repository:</b> INB Testing repository chirimoyo.ac.uma.es <b>Parameters:</b> [null,D:\JorCA Files\runNCBIBlastp1.2010.04.26.13-09-1],[null,Expected_value],[null,0.02],[null,5] <b>Status:</b> FINISHED			
<b>Tool:</b> getAminoAcidSequenceCollection <b>Date:</b> 26 Apr 2010 11:11:06 GMT <b>Where:</b> mango.ac.uma.es <b>Repository:</b> INB Testing repository chirimoyo.ac.uma.es <b>Parameters:</b> [null,D:\JorCA Files\getBestHitsFromBlast.2010.04.26.13-10-4] <b>Status:</b> FINISHED			
<b>Tool:</b> runClustalwFastUMA <b>Date:</b> 26 Apr 2010 11:15:45 GMT <b>Where:</b> chirimoyo.ac.uma.es <b>Repository:</b> INB Testing repository chirimoyo.ac.uma.es <b>Parameters:</b> [null,D:\JorCA Files\getAminoAcidSequenceCollection.2010.04.26.13-10-1],[null,0.5],[null,4],[null,10.0],[null,5],[null,BLOSUM],[null,1], <b>Status:</b> FINISHED			
<b>Tool:</b> runCreateTreeFromClustalw <b>Date:</b> 26 Apr 2010 11:15:58 GMT <b>Where:</b> mango.ac.uma.es <b>Repository:</b> INB Testing repository chirimoyo.ac.uma.es <b>Parameters:</b> [null,D:\JorCA Files\runClustalwFastUMA.2010.04.26.13-15-11] <b>Status:</b> FINISHED			

Drag and drop the tools in order into the white canvas, you will get this graph:



Click on export if you want to get a SCUFL file, click on execute in jORCA to launch the execution panel.

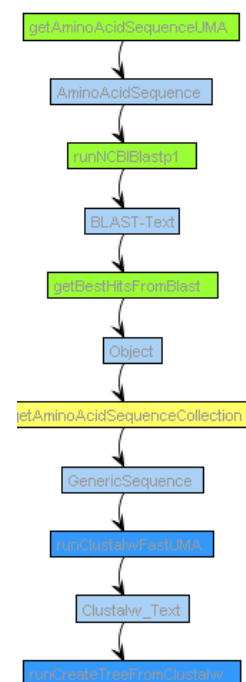
On the left of the execution panel, the pipeline will appear, on the right the parameters to be completed.

Only the first tool parameters and the outputs are shown to be completed by the final user. All the drag and drop and data-conversion features from jORCA are also available in this panel. If the user wants to configure any parameter such as the Clustal gap distance, clicking on the arrow these options will be displayed.

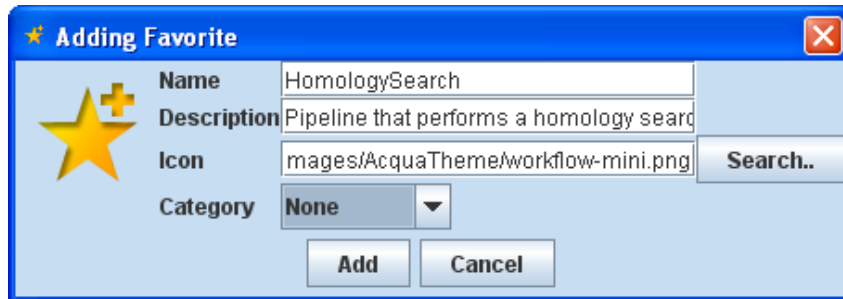
Some tools have more than one mandatory input, if this happens the tool name will be color in red. Also, a tool output could match with more than one input of the next tool, if jORCA detects this; it will ask the user about which parameter use.

The execution is quite simple, imagine that you don't want to change any default parameter, and there's no other mandatory input that the first one. Complete the string parameter with DQA1\_HUMAN. Just click on the "Launch" button and the execution will start.

jORCA monitors the execution coloring the different tool boxes in the graph, and also allows saving the intermediate tools' results.




The user can also save the pipeline in the favorites list clicking the button with a star in the panel or export it to a SCUFL file.



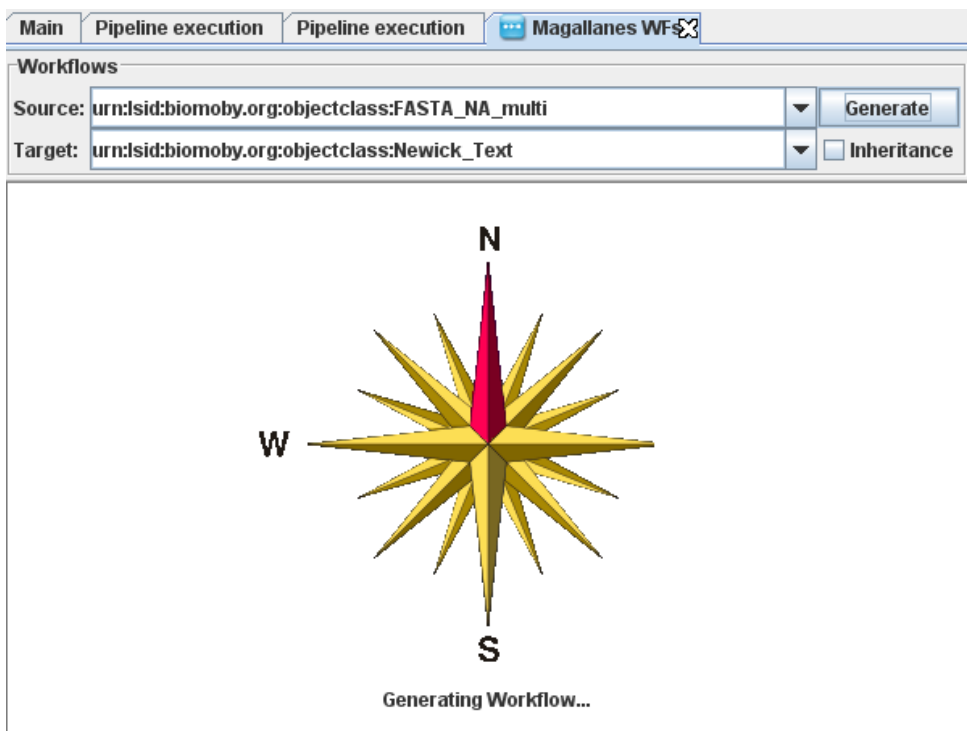
### *How to create pipelines using Magallanes*

Any workflow found using Magallanes is executable with jORCA's pipeline module.

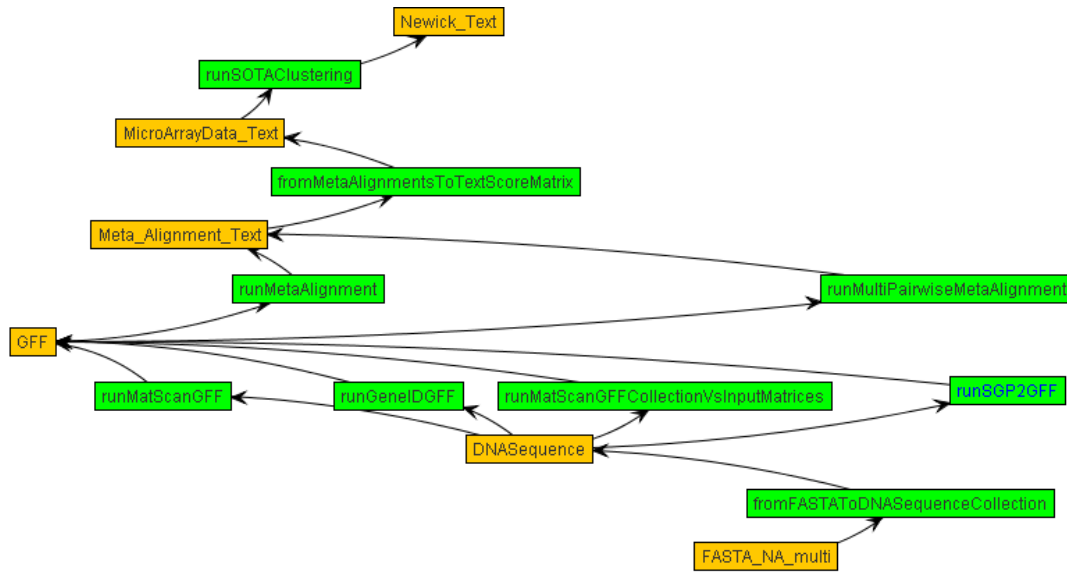
We will use the proposed workflow by Arnaud Kerhornou and Roderic Guigó in <http://bioinformatics.oxfordjournals.org/cgi/content/full/23/14/1831> as an example to

illustrate this feature, click on Magallanes workflow composition icon  and write fasta\_NA in the first text box (Source) and press the arrow down keyboard button to select FASTA\_NA\_Multi, in the target text box write Newick and press the arrow down keyboard button to select Newick\_Text. Click on Inheritance to de-select this option.

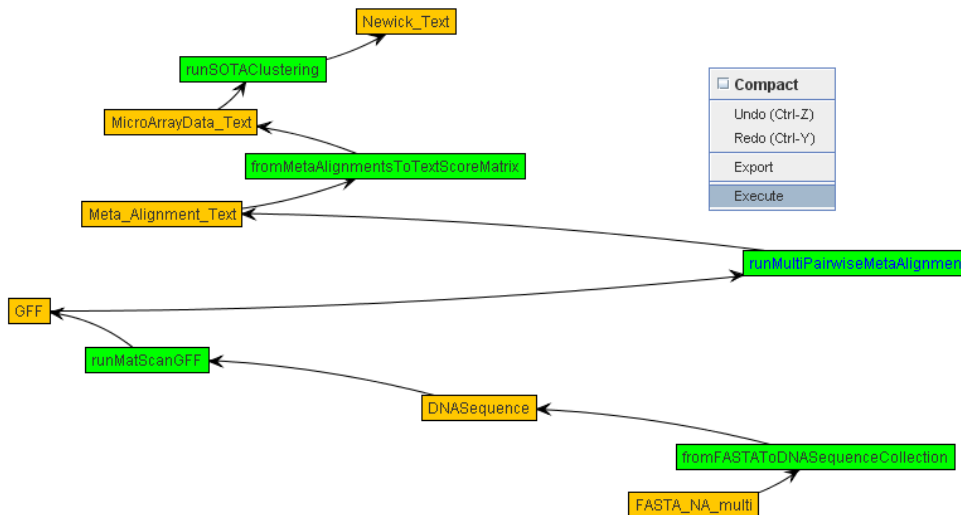
Now click on generate workflow.



The minimal path from both data-types will be displayed, if there is more than one option all the possibilities will be displayed.



Double click the tools that you want to use in the pipeline (runMatScanGFF-runMultiPairwiseMetaAlignment) and once there's a pipeline, right click on the graph and select the "Execute" option.



jORCA pipeline's execution panel will open. You can test this pipeline creating a FASTA\_NA\_Multi object with this content:

```
>PDEF_ENSG00000124664_PDEF:NCBI36:6:34632070:34632569:-1:TSS_predicted
GGCCCAAGCCTCGGGATCCCTCCCCAGGGTCTCTGAAGCTCTCTCCATCCCTGGCCCT
GAGTAGCCAGACAGCACCTCCTCCAGGAAGCCCTCACTGATTTCCCTAGTTGGTGCCAC
CCTCAGTGCCCCCTCAGTCCTCCATCTGGGCATGGGTGGTTCTGGATCTCCACTGCTGCT
CACTTGCTCTGCTGGCCCTCAGTGATCCATCTTAGAACCCAGCCCTGGACCCACTC
GACGTATCTCTGGCGCCTTGACGTAATATGAGCTGAGTGGCTATGCAGCAACCAATGAA
CGAGTAATGAGCGAGTGAATGAATGAGTCCCCTAGCTGTGAGGCATGGATCCCCAGC
AAGGAGGGGAGACCTGCAAGGGTTAATCAGGAGCCTGCCTGTGGTCTGAGGTAAGCAAG
GAGTCTATTTGTTTCAGCTAAATAAGGAAGGATTACTTATAATGGAAATCAGGCCCTGGC
CAACTCTTCATCTCGCGGCT
>H2AFO_ENSG00000203812_H2AFO:NCBI36:1:148088752:148089251:1:TSS_predicted
GAAGATGCTGTGAGGAAGGAGTTCATGATGCCCATGGCCCTGCACCAGATGCCGGTCTC
GGGTGGACCCGCTTCAGCACCTTGTACACGTAGATGGAGTAGCTCTCCTTGCGGCTGCG
CTTGCGCTTCTGCGCTTCTTCTTGGGCTTGGTGACGGCTTCTTGGAGCCCTTCTT
GGGAGCCGCGCAACTTGCAGGCTCAGGCATGGCCAGACCCAGACCCGACACCGACCC
CCGAGAAGCAAGCAGACGGTAGGCTCGGGTCTACCCGAAACGACTGTGTACTTACAG
```

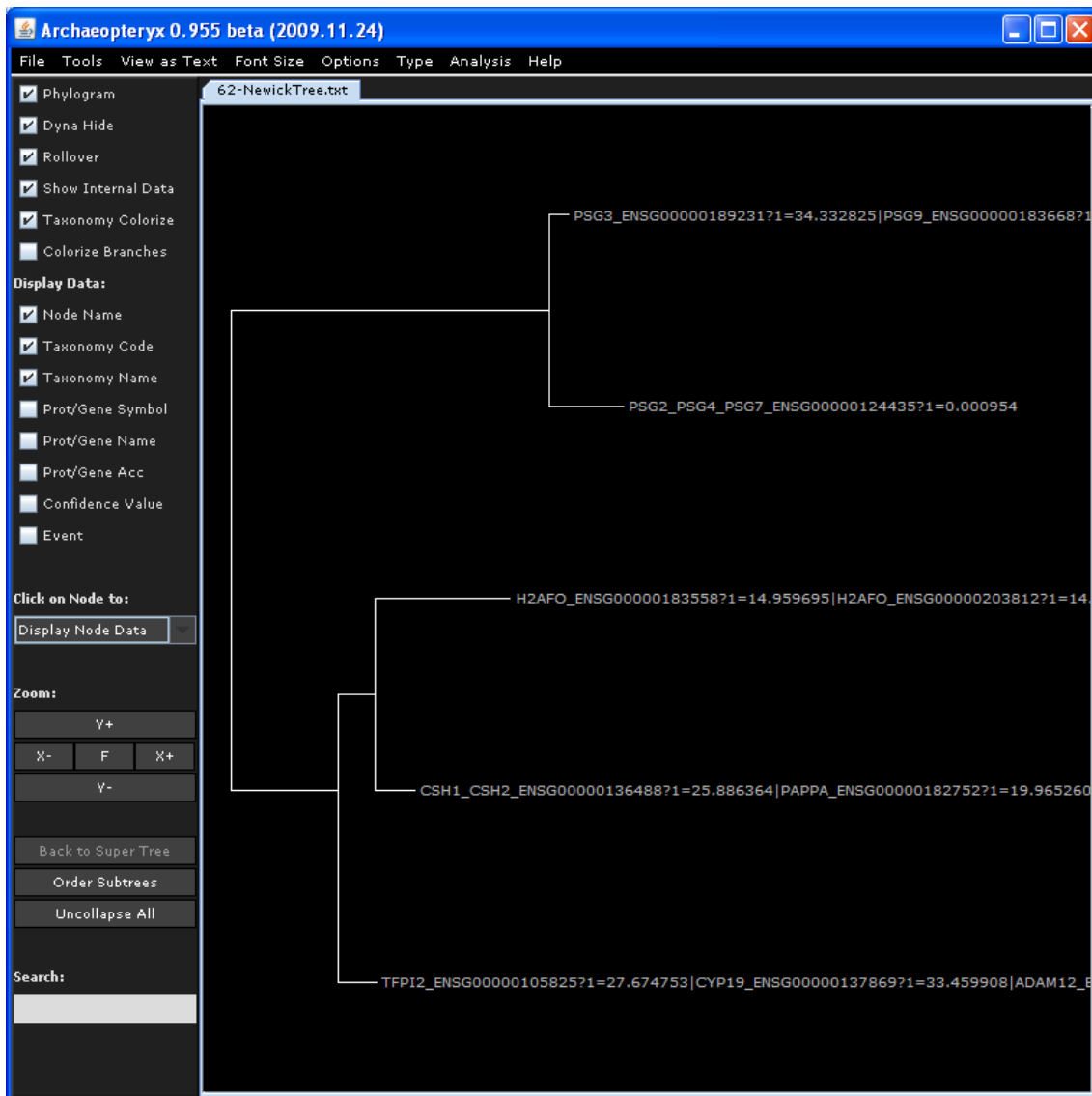
AGGCTGTGCGCATGACGCTGCGTTATGGTTCCGCGAGTTTTCCGCGGCGCGCAATGCGAGG  
GAGACGAGATTATGTAATGAGTGGATTCTGGCTGAGCTATCCTATTGGCTATCGGGACA  
AAATTTGCTTGAGCCAATCAAAGTCTCCGTGGACAATCGCCGTTCTGTCTATAAAAAGG  
TGAAGCAGCGGGCTTTCCG  
>H2AFO\_ENSG00000183558 H2AFO:NCBI36:1:148081103:148081602:-1:TSS predicted  
GAGCTGCGCTACTTGGTACCCTGGTGGCTCCGACACCGCGTGGTGGCCAGCTCG  
CCGGGACAGCAGCGGCGCAGCGCTCTCCGCGATCGCGTCAAGATGTCTGTGAGGAAAG  
TGTAGTCCGCGAGGGGAGCCCTCTCCGCGATCGCGTCAAGATGTCTGTGAGGAAAG  
AGTTATGATGCCCCATGGCCTTGCACAGATGCGCGTCTGGGGTGGACCCGCTTCAGCA  
CCTTGTACAGTATGAGTGGAGTCTCTCTGCGGCTGCGCTTCCGCTTCTGCGCTCTT  
TCTTCTGGGCTTTGGTACGGCTTTCTTGGAGCCCTTCTTGGAGCCGCGCGAACTTTG  
CAGGCTCAGGCATGGCAGCCAGACCCAGACCCGACCCCGGAGAACGCAAGCAGAGCGG  
GTAGGCTCGGGTCTACCGAAACACTGTGTACTTACAGAGGCTGTGCGCATGACGCTG  
CGTTATGGTTCGCGAGTTTT  
>PSMD7\_ENSG00000103035 PSMD7:NCBI36:16:72887682:72888181:1:TSS predicted  
TTTAGCCCTAGTTACCGAGTAATAACGACAGGTGGAAACCTGCAGGTTGCCAGTTC  
ATTGCAAAAGTGTCCACAGTAAAGTAAAGGAGAAATCCCCCGCAATGAAAACAGGAAT  
GAGATAAATTACCAAAATACATTAAGAGCGTGCCTTATGTTAGAGATTAAAGTTCCTC  
TCAGAAATCGCGGAGGATTTGACCTGAAGCAAAAGTGAATCTTCCGGGTGGTAGAAA  
AGTAACTGTCAACTATTACTTTGGGGGGTGGCGGGGGAGGTTGCTGGGAAGGAATT  
GAGGCGAAAAGCCTCTGAGGTGACGGTACAGGCCAATACCCAGAGGTCAGTCCCGGT  
GTCCCTTCGGCCCTTAAGTCCAGGAAACGCACCGCGGAGGACGAGGGCGATTCCCGCC  
TTTCCAGCCCTTTCTCCGCCCAAAATGACGGCGTGGCGAAGATCCGGGCGCGCG  
GCGAGGCACTCTGGGAGCGG  
>NUCKS\_ENSG00000069275 NUCKS:NCBI36:1:203986028:203986527:-1:TSS predicted  
GATATTCCTGATATGATCCTGCCTCTACTCCAGCAGCCCTATTGGTCCCTGAGGG  
TTAATGCCGAGCCAGGGAAGGGCTTGTAGTTAAAGTCAAAATTTTGAACCTAAG  
GAATTTCCCTCAATCTGTAGCAGCCCGCAGTGAAGTCTGTCTGCTCTCCACGGG  
TTTTGCTATTAGAGGATTTCAAGTGATCAAGCAGACGAGATGAACACCAATTAAGT  
AACACTGAGCAACGCTGAACCTACATAAAATCTTTTTTCTAGTCTTTTAAGATACGAC  
CCCAGAGAACTGTCAACTCTGGAGTGGAAACCGGGGGGGGTAAGCAGGACCACTTT  
TTGTTTTGGCGCGTGGCTCTCAGAGCATCTTACTAACAGGACCCCGCAGTGAAGT  
GGAGCCTATCAGAGCAGGGCGGGGCTGGAAGACGTGATGATCAAAAGGACTAATGA  
TCTGGGCGCATCTGGCAGAA  
>FLJ10111\_ENSG00000092098 FLJ10111:NCBI36:14:23685436:23685935:1:TSS predicted  
AGCAGGAATCTGGAAATCCCGAGAACTGAAGCACTAGCAGATTGGGGAGTTCTCT  
GGCACTGCTGGTCCCTGAACACCCAGCTTGGCTCCACCAAGAGGGTCTATGACC  
CCTTCTGGCCTCCGATTCCTTACCTGTTGGGGCTTCCCGCTCAGGGCGCACCC  
ACACGGCTTGGCCATGCTGCTTCAAGTGTAGATCTCTGGTCTCCCGCTAGTCCGCGG  
GCTTTCGCTTTCACCTCCAGGTCAGGCGCGCCCGCTCAACCCCGCCCTTTCTTCT  
TTCCGCGCCCTCTCCCGCTGGCAGGCCGCAAGCAGTAGGTAGCCCTCAGCATTCC  
AGCCAGTTGCCACTAGGAGGGCGGTGCTGGTGGTCCAGTCCAGGGGCTCCCGGAG  
GAGGGCTAGGGTAAAGTGGACCGGAGACTGGAGCGGGGCTGCTCCCTCGAGAGGGG  
CGGGCTTATAGCTAGGCC  
>SQSTM1\_ENSG00000161011 SQSTM1:NCBI36:5:179180003:179180502:1:TSS predicted  
CTGCCACCGCGCAGGCTGACCATGGCAGGTCGGAAACAGCAACCTGAGAGCCAGCTGT  
TCTGGCCAGCAGTGGCCACTGGGCGCACTAGCAGCCCTCTGATATGGGGGCTGTGCTCC  
CTCTCCCTGCACTGGGTACCCCAACTGAGGATATTTGCTGAGTCAATGGCCAGGCCAAGC  
CTGGGAGGGGCGAGGGGCTGGACCCCGCCAGTACCCTGATCCAGGTCAGAGGCTGGA  
GCCAGCCCTGTATGAGTCCAGGGCGGTTCTTGGGGTCTTGGTGCACCGGGCAAT  
GAAGAGAGGGGTGCAATAGGGGGCGGAGACCTGGAGCGAGGGGTAGCGGGGAAAG  
GGAGACTGTAAGGGCCCTCTCAGGGCGGCTCTCCGCGCCGACGACGCTGCGGGG  
CGGGAGGGGCGCAGAGACTCCGCCCTCTCAGAGGGGGGGGGGCTCCGCGTTCCGTA  
CAAAAGCCCGCGGGGGCTG  
>PSG9\_ENSG00000183668 PSG9:NCBI36:19:48465521:48466020:-1:TSS predicted  
GGGTCCTCTCACCAGGCAACACAGCTCACCCACAGCAATGGGGTGTACTGTGAC  
AGGACCTGGTGCATCTCCAGCCTCTGCACTGAAGGGGAGGCCAGATGGGGATGAAAG  
AGGGTAAAGGCAATGATGTGACCCCAACCCAGAGCCATGGGGACAGCAGGAGGCTGAG  
GCCAGCACTGTGCTTCCCAACTGCAGGGTATGTGCTGACTGTGTGGTCTGTGTGT  
GTCTTCTGT  
GTTGAGGTTTGGTAAAGAACTCACTGTGAAAAGGCGAGGGCTCCACAATCCCGAGG  
ACCTGAAACACAGACAAAAGGAAAACAGAAAGGAGGCAAGAGGAGCAGGACTGAGAGAG  
AAGGGACAGAGAGGTTCTCTGGGACTGACCCCGCCATGAGCTGAGAAAGTGTCTCTG  
CCCGGAAAGGCTCAGCAC  
>PSG3\_ENSG00000189231 PSG3:NCBI36:19:47936467:47936966:-1:TSS predicted  
AAAAGGAGGTTCCCTCTCATCAGGCAACACAGCTCACCCACAGCAATGGGGTGT  
ACTGTGACAGGACCTGGTTCAGCTCCAGCCTCTGCACTGAAGGGGAGAGCCAGATGGG  
GATGAAAAGGCAAGGGCATGATGTGACCCCGACCCAGAGCCATGGGGACAGCAGG  
AGGCTGAGGGCCAGGACTCTGCTTCCCAACTGCAGGGTATGTTTCTGACTGTGTGGT  
CTGTGTGTCTCTCTGT  
AGGCTTGGTAAAGAACTCACTGCTGAAAAGGCGAGGGCTCCACAATCCCGAGGACCT  
GAAACACAGACAAAAGGAAAACAGGAGGACAGGAGGCGAGGACTGAGAGAGGAGGGGA  
CAGAGGCTGCTCTGGGCTGACCCCGCCATGAACCTGAGAAGTGTCTCTGCCCCGGGA  
AGAGGCTCAGCCGAGAGGA  
>PSG6\_ENSG00000170848 PSG6:NCBI36:19:48113884:48114383:-1:TSS predicted  
AGACCGATTAGAAAAGGAAAGGTCCTCTCACCAGGCAACACAGCTCACCCACAGC  
ACAATGGGACATCACTGTAAAGGGACCTGGTGCAGCTCCAGCCTCTGCACTGAAGGGG  
AGAGCCAGATGGGGATGAAAGAGGGCAAAATGGCATGAAATGTGACCCCGACCCAGAGCA  
TGGGGACAGCAGGAGGCTGAGGCGCCAGGACTGTCTCCGCAACCTGCAGGGTATGTGTG  
TCACTGTGTGGTCTGT  
GT  
GGACTGAAACACAGACAAAAGGAAAACAGAAAGGAGGCAAGGAGGCAAGGAGGCAAGG  
AGGAGGGACAGAGAGGTTCTGGGCTGACCCCGCCATGAGCTGAGAAGTGTCTCTGCCCCGGGA  
AGAGGCTCAGCCGAGAGGA  
>PAPPA\_ENSG00000182752 PAPPA:NCBI36:9:117955392:117955891:1:TSS predicted  
ATAAGTGGGGGAGTGGGGAGGAAAAAAGCAAGCAAAACAAAACAAAACAGGAATG  
TGTGTTTTGAGCAGGAATGAGAGCTGGGAGGACCTAAGGGGGTGTGACAGAAATTCAT  
TAAATTTGTTACTTTAAGATTGTCTCTTAAAAAATAAATAAATAAAGAGAGAGAGAG  
AGAGAAAAGGAGGGGGTGGTGGAGAAGCGGGAGCGAAGGAAAAGGAGGCAAAAGGCAAA  
GTGAAGAAAAGCTGGATAGCTCGGCTCTCAAACTGATTTAGTCTGATCCCGCA  
GTTTTAACAGGACTATTCAATGGGAAGGTGGAGCGTGGGGAGCAGATTAGCATACG  
CTTGTACTCATCTCTGAGGGATTTTTCCCGCTTTTCTTTTCAATTTGAGAAGAAAG  
GAGGGAGGGGAGGGGACTTGGGGGGGAGAAAGGGGGCTGTGGCTGTGTTATAAAGGA  
CGCAAAAAATAAATAAATTA  
>CSH1\_CSH2\_ENSG00000136488 CSH1:NCBI36:17:59327648:59328147:-1:TSS predicted  
TCAATGGTGTGCTCAGAACCCCACTATTTGGTGTGTGTGGCCCTTTTCCCAACACA  
CACATTTCTGTGGTGGTGAAGTAAACACCGGGGAGGAGAAAGGAATAGGATAGA  
GAGTGAATGGGGTGGTGGGCTCAAGGACTGGCTATCTGCAGCCTTCCCGCGT  
TCAGGTTGACCAACATGGCCTGCAGCCAGAGGGCACCCCTGACCTTAAAGAGAGGAC  
AAGTTGGTGGATCTGTGGTGCACCTGTGTGCACAACTTCAACAACCTGGTGTGGT  
GAGAAGGAAAAGCAGCAAGCCAGGGGCGATGATCCAGCATGTGTGGGAGGAGCTTCTA  
AATTAATCACTAGCACAAGCCGTCAGTGGCCCATGCAATAATGTACACAGAAACAGGT  
GGGTCAGCAGGAGAGAACTGGCAGGTTAAAAAGGGCCCAAGAGACCCGGT



Once the execution has finished, the user can visualize the results using one of jORCA's Newick tree viewer. (Han M.V. and Zmasek C.M. (2009). phyloXML: XML for evolutionary biology and comparative genomics. BMC Bioinformatics, 10:356. Zmasek C.M. and Eddy S.R. (2001) ATV: display and manipulation of annotated phylogenetic trees. Bioinformatics, 17, 383-384. )

Name	Type	
runCreateTreeFromClustalw.2010.04.26.13-...	Newick_Text	
runCreateTreeFromClustalw.2010.04.26.16-...	Newick_Text	
runNCBIblastp1.2010.04.26.13-09-1	BLAST-Text	XML viewer
runNCBIblastx.2010.04.26.10-51-1	BLAST-Text	HTML viewer
runSOTAClustering.2010.04.26.12-10-4	Newick_Text	<b>Convert to..</b> ▶
runSOTAClustering.2010.04.26.17-27-3	List_Text	Object editor
search.2010.04.26.16-07-8	SearchResult	<b>Hierarchy compatible Tool</b> ▶
		HyperTreeViewer
		Archeopteryx

(use Archeopteryx for this result)



## **Bugs fixed**

- The favorite list search feature was bugged.