

## Welcome To PhyloEasy User Tutorial:

Thank you for deciding to conduct your research with **PhyloEasy**. This manual contains detailed instructions for PhyloEasy. Our PhyloEasy clients may want to read this manual which tries to cover every significant aspect of PhyloEasy that might be useful for clients.

PhyloEasy is a web-based program that has been designed to ensure the scientific study of evolution. The purpose of the program is to provide an easy-to-use web application for exploring, discovering, and assessing biological data from an evolutionary perspective. By employing distance-based approaches for molecular phylogenetic analysis, PhyloEasy expands the scope of its predecessor's investigations from single genes to the whole genome. It employs a basic methodology that is easily accessible to the scientific community for their research work. It is a user interface that allows users to select criteria that are appropriate for understanding ancestry, evolution and adaptability. This tutorial provides an overview of the PhyloEasy features.



### Welcome to PhyloEasy.

PhyloEasy a client-server freely available online web application that allows users to interactively build their own choice phylogenetic trees and visualize it in a graphical interface. PhyloEasy is a tool that allows the user to build different distance-based phylogenetic trees in one easy-to-use web application. PhyloEasy builds two types of trees, NJ and UPGMA trees.

### Build your own PhyloEasy Tree

Progressive sequence alignment by CLUSTALW

Upload your Fasta sequence file: Support Format: (FASTA pearson)

Choose file that contain your Query:  No file chosen

Display box shows the uploaded File:

```
Example:
>YP_009742608.1_SARS-CoV 2
MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQHLKDGTGCLVEVEKGVLPQLE
QPYYVFIKRSRDARTAPHGHVMVELVAELEGIQYGRSGETLGVLPVHVGEPVAYRKVLLRKNNGK
AGGHSYGADLKSFDLGDGLTDPYEDFQENWNKTHSSGVTRLMRELNGG
```

User won't be able to paste your own sequences into the display box.

Parameters for Phylogenetic Tree:

Complete-Deletion ☐ Yes ☐ No

Substitution Model : Pairwise Distance Model

Tree Clustering Type : ☐ NJ ☐ UPGMA

Bootstrap Range: 100

**PhyloEasy** is an interactive, user friendly and novel web program for comparative phylogenetic analysis. **Homepage** is highly compact, with tab buttons linking to the Tutorial, about, and Contact sections. If you have any questions or suggestions, please contact us.

PhyloEasy

TutorialAboutContact

---

Build your own PhyloEasy Tree

Progressive sequence alignment by CLUSTALW

Upload your Fasta sequence file: Support Format: (FASTA pearson)

Choose file that contain your Query:  No file chosen

Display box what's inside in your uploaded File:

Example:  
>YP\_009742608.1\_Severe\_acute\_respiratory\_syndrome\_coronavirus\_2  
MESLVPGFNEKTHVQLSLPVLQVRDVLVRGFGDSVEEVLSEARQHLKDGTGGLVEVEKGVLPQLE  
QPYVFIKRS DARTAPHGHVMVELVAELEGIQYGRSGETLGVLPVHVGEIPVAYRKVLLRKNGNKG  
AGGHSYGADLKSFDELGDLPYEDFQENWNTKHSSGVTRELMRELNGG

Clear

The Home-page contains three sections, the first section contain introduction, in the second section user upload their FASTA sequence file and remember in display box user won't be able to paste their own sequences. PhyloEasy supports sequence alignment using **CLUSTALW** program. The display box shows what's inside in your uploaded (**.FASTA**) file. The clear button clear the all text from display box.

### Parameters for Phylogenetic Tree:

Complete-Deletion ☐ Yes ☐ No

Substitution Model : Pairwise Distance Model

Tree Clustering Type : ☐ NJ ☐ UPGMA

Bootstrap Range: 100

Execute

After uploading sequence file user select Parameters according to their interest. If user want complete deletion (Gaps/Missing data treatment) from sequence file after alignment then select yes option otherwise select no. Substitution model contain P-Distance Model. In Clustering user select NJ and UPGMA option according to their requirement. To see if the tree's phylogeny is correct, PhyloEasy employs the bootstrap approach, with the number of bootstrap replications set at 100. Click on **Execute button** for phylogenetic tree.

## Evolutionary history of GLI gene family is explained using NJ and UPGMA tree.

**Result:** Through PhyloEasy the evolutionary relationships of taxa was inferred using Neighbor-Joining method and unweighted pair group method with arithmetic mean. The evolutionary distances were evaluated using the p-distance method. The finest tree based on the sum of branch lengths. The total percentage of duplicate trees with related taxa grouped together is shown above the branches.

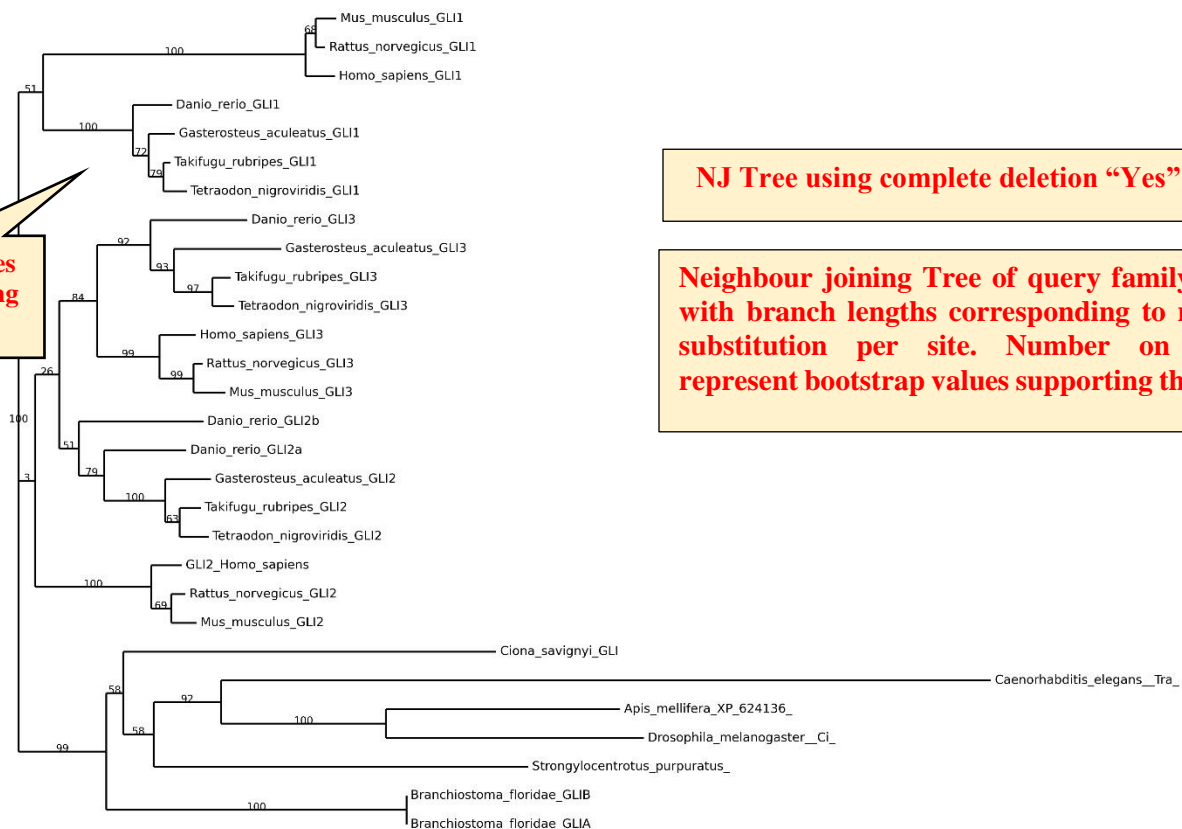
### PhyloEasy Results

[Alignment File](#)

[Pairwise Distance File](#)

[Tree Files](#)

[Tree Images](#)

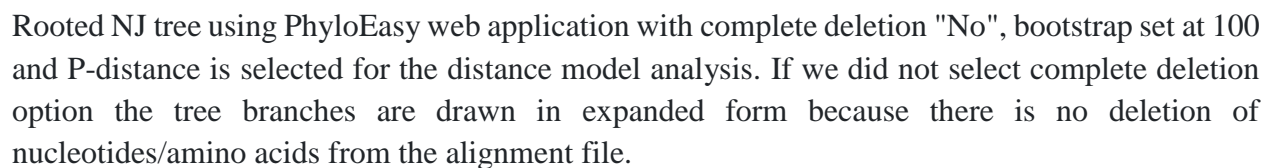


**Figure: PhyloEasy built NJ tree based on pairwise distances with complete deletion for 29 GLI sequences.**

---

[PhyloEasy](#)
[Tutorial](#)
[About](#)
[Contact](#)

Alignment File      Pairwise Distance File      Tree Files      Tree Images



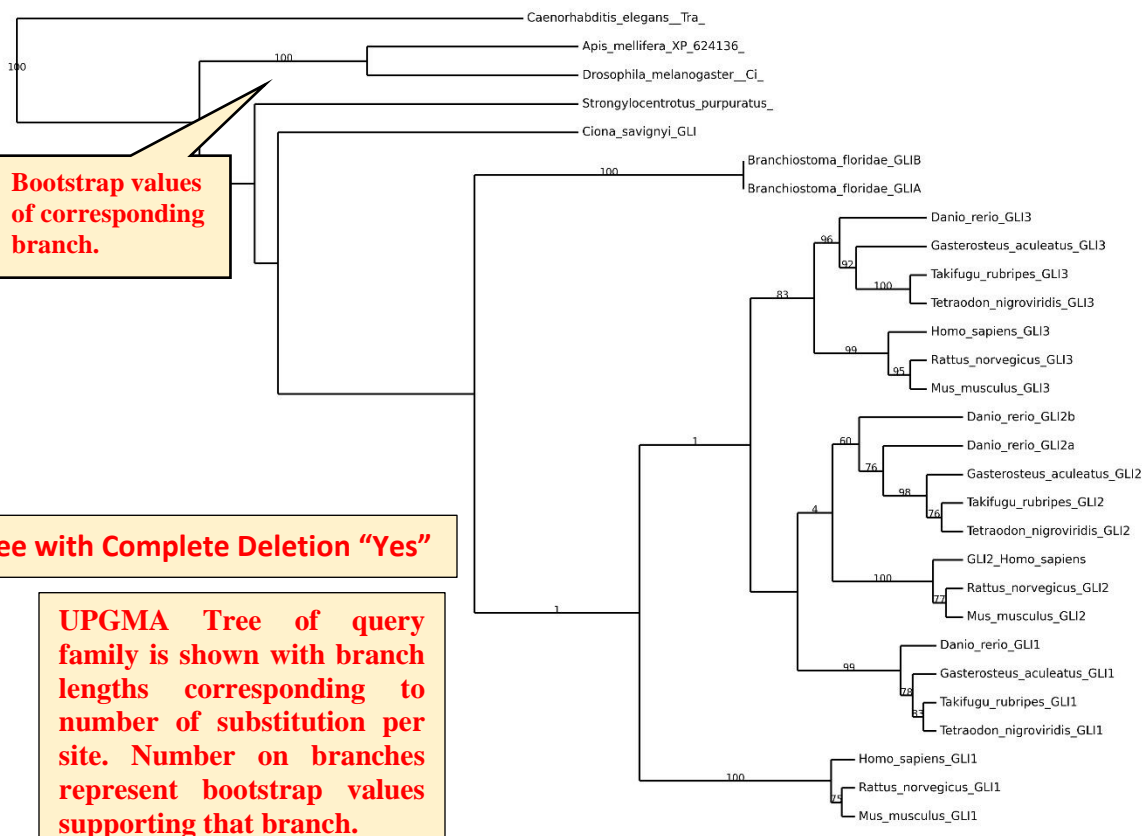
## PhyloEasy Results

Alignment File

Pairwise Distance File

Tree Files

Tree Images



**Figure: PhyloEasy construct UPGMA tree based on pairwise distances with complete deletion for 29 GLI sequences.**

PhyloEasy built UPGMA tree with complete deletion "Yes", bootstrap set at 100 and P-distance is selected for the distance model analysis. UPGMA tree explain the evolutionary history of GLI gene family by duplicating nodes. Each node represent recent and accent duplication.

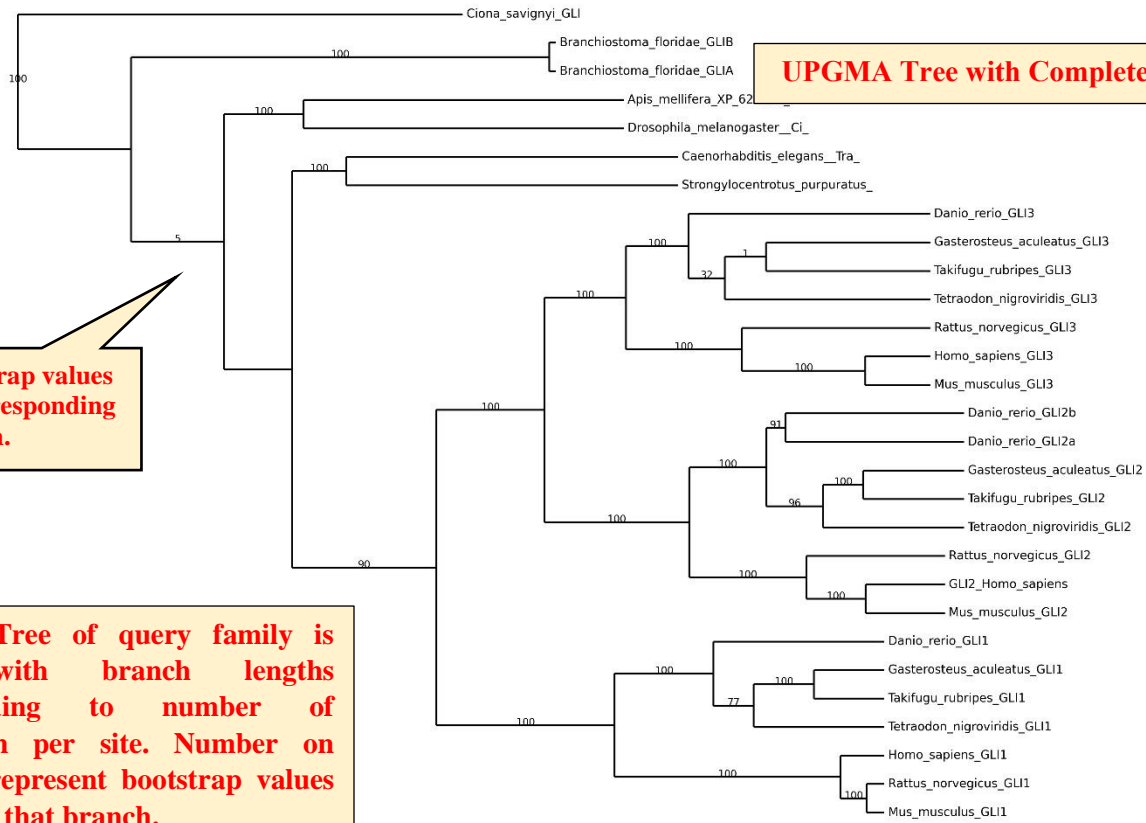
## PhyloEasy Results

Alignment File

Pairwise Distance File

Tree Files

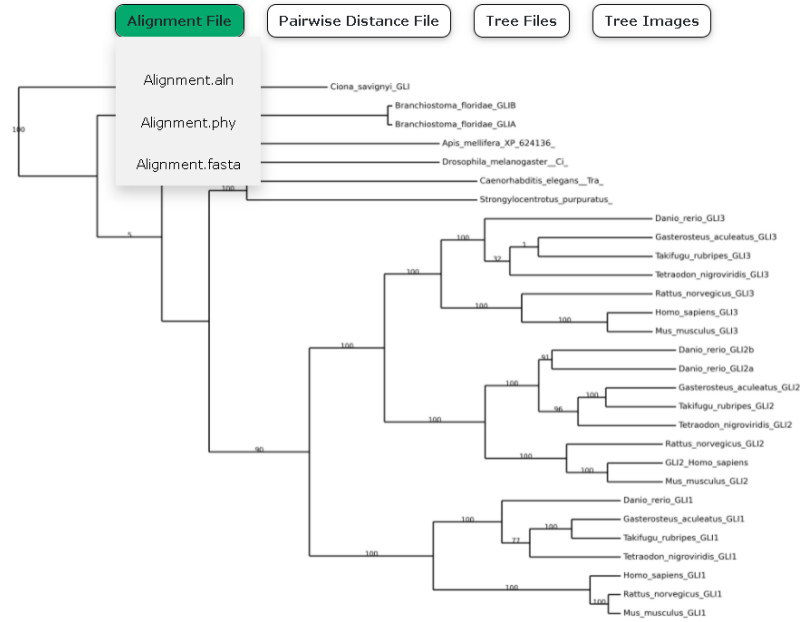
Tree Images



**Figure: PhyloEasy construct UPGMA tree based on pairwise distances without complete deletion for 29 GLI sequences.**

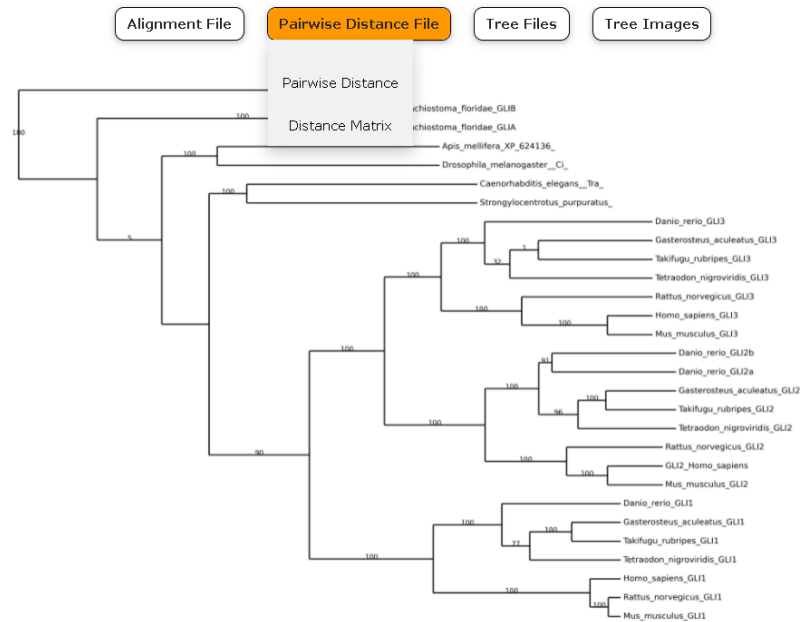
PhyloEasy built UPGMA tree with complete deletion "No", bootstrap set at 100 and P-distance is selected for the distance model analysis. UPGMA tree explain the evolutionary history of GLI gene family by duplicating nodes. Each node represent recent and accent duplication.

## PhyloEasy Results

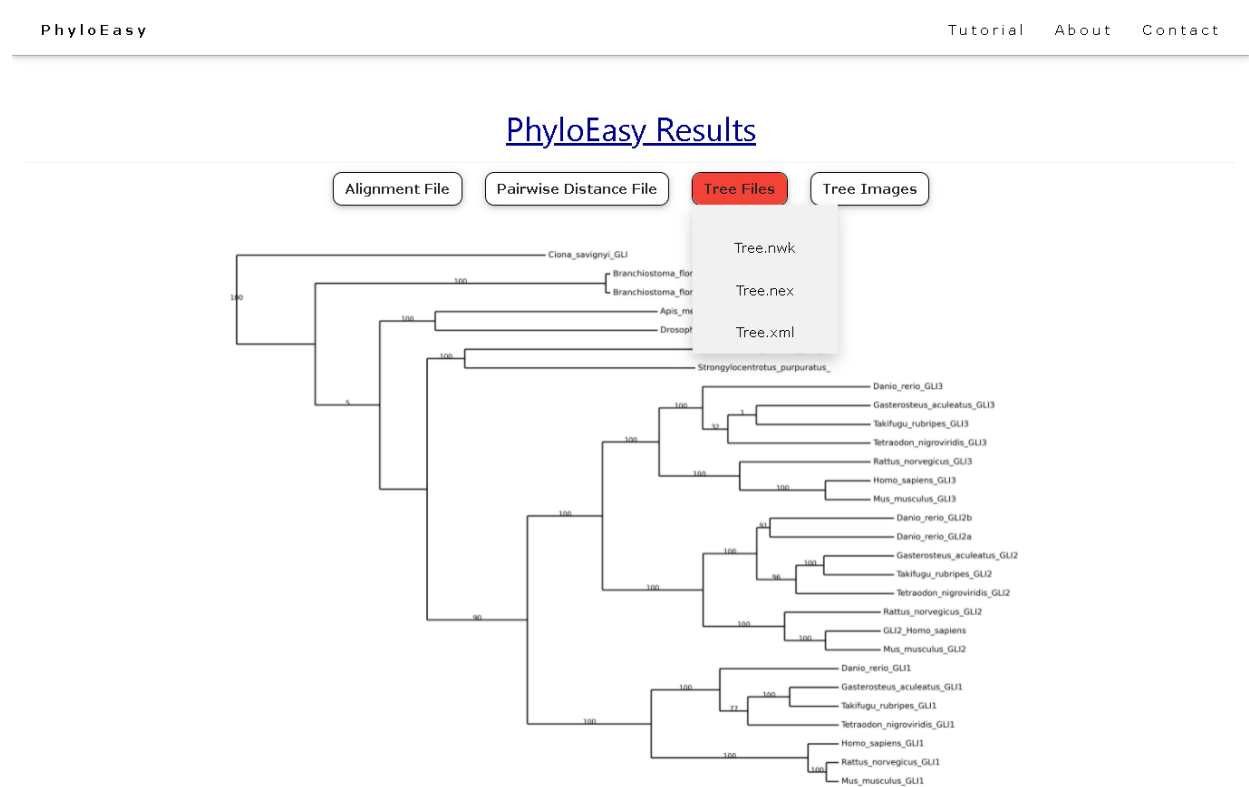


**Figure: From alignment file button user can download alignment files in three alignment formats (.aln, .fasta, .Phy)**

## PhyloEasy Results



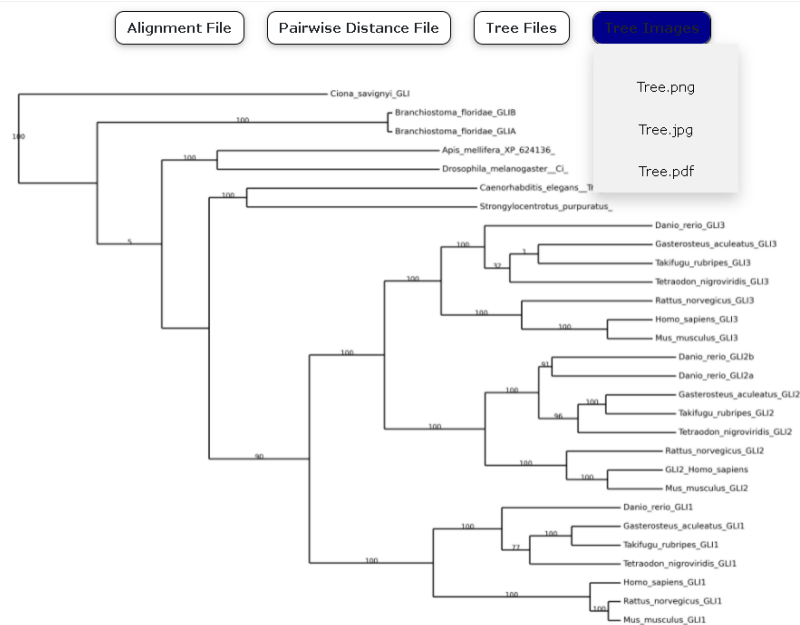
**Figure: From Pairwise distance file button user can access P-Distance files in .txt and .csv format. Both files contain pairwise distances between their query species. .csv file provide species distances in a matrix form.**



**Figure: From Tree file button user can download Tree files in different format according to their requirement (.nwk, .nex, .xml). Each Tree file contain species branch lengths according to their clade divisions.**



## PhyloEasy Results



**Figure: PhyloEasy users download Tree images in these format (.jpg, .png, .pdf). For Tree visualization in Mega software, User Can download Tree.nwk file from PhyloEasy site and open it with Mega software.**

## PhyloEasy Result Output Files

Mus_musculus_GLI1	-----EGCRKSYSRLENLTKHLRSHTGEKPYMCEQEGCSKAFSN
Rattus_norvegicus_GLI1	-----EGCRKSYSRLENLTKHLRSHTGEKPYMCEQEGCSKAFSN
Homo_sapiens_GLI1	-----EGCRKSYSRLENLTKHLRSHTGEKPYMCEHEGCSKAFSN
Tetraodon_nigroviridis_GLI1	-----EGCNKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Takifugu_rubripes_GLI1	-----EGCNKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Gasterosteus_aculeatus_GLI1	-----EGCNKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Danio_rerio_GLI1	-----EGCNKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Mus_musculus_GLI2	-----EGCSKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Rattus_norvegicus_GLI2	-----EGCSKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
GLI2_Homo_sapiens	-----EGCSKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Tetraodon_nigroviridis_GLI2	-----GCSKAYSRLNLTHTLRSHTGEKPYLCEHEGCKNAFSN
Takifugu_rubripes_GLI2	-----EGCSKAYSRLNLTHTLRSHTGEKPYLCEHEGCKNAFSN
Gasterosteus_aculeatus_GLI2	-----EGCAKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Danio_rerio_GLI2a	-----EGCSKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Danio_rerio_GLI2b	-----EGCSKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Mus_musculus_GLI3	-----EGCTKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Rattus_norvegicus_GLI3	-----EGCTKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Homo_sapiens_GLI3	-----EGCTKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Tetraodon_nigroviridis_GLI3	-----EGCAKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Takifugu_rubripes_GLI3	-----EGCAKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Gasterosteus_aculeatus_GLI3	-----EGCAKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Danio_rerio_GLI3	-----EGCSKAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Branchiostoma_floridae_GLIA	ISLRTLPLVLPFEGCKNAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Branchiostoma_floridae_GLIB	ISLRTLPLVLPFEGCKNAYSRLNLTHTLRSHTGEKPYVCEHEGCKNAFSN
Strongylocentrotus_purpuratus_	-----EGCYKAYSRLNLTHTLRSHTGERPYVCEFGCTKAFSN
Drosophila_melanogaster_Ci_	-----EGCFKAYSRLNLTHTLRSHTGEKPYTCEYPGCSKAFSN
Apis_mellifera_XP_624136_	-----EGCFKAYSRLNLTHTLRSHTGEKPYTCEYPGCSKAFSN
Ciona_savignyi_GLI	-----EGCSKAYSRLNLTHTLRSHTGEKPYVCEYPSCTKAFSN
Caenorhabditis_elegans_Tra_	-----GCGKEYSRLENLTKHTRHTHTGEKPYKCEFADCEKAFSN
	*** ***** *:****.* ** .* *****

**Figure: Build amino acid multiple sequence alignment of GLI gene family.**

Species include for generating multiple sequence alignment are: *Homo sapiens* (Human), *Mus musculus* (Mouse), *Rattus norvegicus* (brown rat), *Tetraodon nigrovirds* (green spotted puffer), *Gasterosteus aculeatus* (Stickleback), *Danio rerio* (Zebrafish), *Takifugu rubripes* (Fugu), *Branchiostoma floridae* (Amphioxus), *Ciona savignyi* (Ascidian), *Strongulocentrotus purpuratus* (Sea urchin), *Drosophila melanogaster* (Fruit fly), *Apis mellifera* (Honey bee). An asterisk (\*) indicates position in which amino acid residues is 100% conserved in a column. A colon (:) indicates conservative substitutions. A dot (.) indicates less conservative substitutions. The hyphens (-) indicate the positions of insertions or deletions.

Column1	Column2	Column3	Column4	Column5	Column6	Column7	Column8
Mus_musculus_GLI1	0						
Rattus_norvegicus_GLI1	0.027105	0					
Homo_sapiens_GLI1	0.068807	0.0663052	0				
Tetraodon_nigroviridis_GLI1	0.388657	0.3890742	0.3853211	0			
Takifugu_rubripes_GLI1	0.43536	0.432026	0.431609	0.18015012	0		
Gasterosteus_aculeatus_GLI1	0.42994	0.4261884	0.4274395	0.2201834	0.108006	0	
Danio_rerio_GLI1	0.4211843	0.4174311	0.4157631	0.2944120	0.234361	0.2247706	0
	Mus_musculus_GLI1	Rattus_norvegicus_GLI1	Homo_sapiens_GLI1	Tetraodon_nigroviridis_GLI1	Takifugu_rubripes_GLI1	Gasterosteus_aculeatus_GLI1	Danio_rerio_GLI1

**Figure: UPGMA Distance Matrix for Pairwise distance between species.**

# Formats for describing Phylogenetic trees

```
#NEXUS
Begin Taxa;
  Dimensions NTax=29;
  TaxLabels Ciona_savignyi_GLI Branchiostoma_floridiae_GLIB
Branchiostoma_floridiae_GLIA_Apis_mellifera_XP_624136
Drosophila_melanogaster_Ci_Caenorhabditis_elegans_Tra
Strongylocentrotus_purpuratus_Danio_rerio_GLI3_Gasterosteus_aculeatus_GLI3
Takifugu_rubripes_GLI3_Tetraodon_nigroviridis_GLI3_Rattus_norvegicus_GLI3
Homo_sapiens_GLI3_Mus_musculus_GLI3_Danio_rerio_GLI2b_Danio_rerio_GLI2a
Gasterosteus_aculeatus_GLI2_Takifugu_rubripes_GLI2_Tetraodon_nigroviridis_GLI2
Rattus_norvegicus_GLI2_GLI2_Homo_sapiens_Mus_musculus_GLI2_Danio_rerio_GLI1
Gasterosteus_aculeatus_GLI1_Takifugu_rubripes_GLI1_Tetraodon_nigroviridis_GLI1
Homo_sapiens_GLI1_Rattus_norvegicus_GLI1_Mus_musculus_GLI1;
End;
Begin Trees;
  Tree
tree1=(Ciona_savignyi_GLI:0.34054,((Branchiostoma_floridiae_GLIB:0.00500,Branchios
toma_floridiae_GLIA:0.00500)100.00:0.32016,((Apis_mellifera_XP_624136:0.24500,Dro
sophila_melanogaster_Ci:0.24500)100.00:0.06108,((Caenorhabditis_elegans_Tra:0
.25396,Strongylocentrotus_purpuratus:0.25396)100.00:0.04145,(((Danio_rerio_GLI3
:0.18489,((Gasterosteus_aculeatus_GLI3:0.12573,Takifugu_rubripes_GLI3:0.12573):0.
03159,Tetraodon_nigroviridis_GLI3:0.15732)27.00:0.02758)100.00:0.04805,(Rattus_no
rvegicus_GLI3:0.14418,(Homo_sapiens_GLI3:0.04962,Mus_musculus_GLI3:0.04962)100.00
:0.09456)100.00:0.08877)100.00:0.06249,(((Danio_rerio_GLI2b:0.13657,Danio_rerio_G
LI2a:0.13657)79.00:0.01473,((Gasterosteus_aculeatus_GLI2:0.07736,Takifugu_rubripe
s_GLI2:0.07736)100.00:0.03065,Tetraodon_nigroviridis_GLI2:0.10801)89.00:0.04329)1
00.00:0.05928,(Rattus_norvegicus_GLI2:0.10592,(GLI2_Homo_sapiens:0.06068,Mus_musc
ulus_GLI2:0.06068)100.00:0.04525)100.00:0.08993)100.00:0.11081)100.00:0.08251,((D
anio_rerio_GLI1:0.13099,((Gasterosteus_aculeatus_GLI1:0.05400,Takifugu_rubripes_G
LI1:0.05400)100.00:0.04608,Tetraodon_nigroviridis_GLI1:0.10008)79.00:0.03091)100.
00:0.07555,(Homo_sapiens_GLI1:0.03378,(Rattus_norvegicus_GLI1:0.01355,Mus_musculu
s_GLI1:0.01355)100.00:0.02023)100.00:0.17277)100.00:0.13641)93.00:0.11052):0.0521
1)1.00:0.07120):0.08658)100.00:0.00000;
End;
```

**Figure: Nexus file of species that contain GLI gene.**

Nexus file starts with #NEXUS header. Nexus file contain multiple blocks like Taxa, Data, and Tree blocks. Taxa block of nexus file contain all information related to taxa or species. Data block contain sequence alignment information in a data matrix form. The Tree block contain information about species phylogenetic trees in a specified Newick format.

```
(Ciona_savignyi_GLI:0.34054,((Branchiostoma_floridae_GLIB:0.00500,Branchiostoma_floridae_GLIA:0.00500)100.00:0.32016,((Apis_mellifera_XP_624136_:0.24500,Drosophila_melanogaster_Ci_:0.24500)100.00:0.06108,((Caenorhabditis_elegans_Tra_:0.25396,Strongylocentrotus_purpuratus_:0.25396)100.00:0.04145,(((Danio_rerio_GLI3:0.18489,((Gasterosteus_aculeatus_GLI3:0.12573,Takifugu_rubripes_GLI3:0.12573):0.03159,Tetraodon_nigroviridis_GLI3:0.15732)27.00:0.02758)100.00:0.04805,(Rattus_norvegicus_GLI3:0.14418,(Homo_sapiens_GLI3:0.04962,Mus_musculus_GLI3:0.04962)100.00:0.09456)100.00:0.08877)100.00:0.06249,(((Danio_rerio_GLI2b:0.13657,Danio_rerio_GLI2a:0.13657)79.00:0.01473,((Gasterosteus_aculeatus_GLI2:0.07736,Takifugu_rubripes_GLI2:0.07736)100.00:0.03065,Tetraodon_nigroviridis_GLI2:0.10801)89.00:0.04329)100.00:0.05928,(Rattus_norvegicus_GLI2:0.10592,(GLI2_Homo_sapiens:0.06068,Mus_musculus_GLI2:0.06068)100.00:0.04525)100.00:0.08993)100.00:0.11081)100.00:0.08251,((Danio_rerio_GLI1:0.13099,((Gasterosteus_aculeatus_GLI1:0.05400,Takifugu_rubripes_GLI1:0.05400)100.00:0.04608,Tetraodon_nigroviridis_GLI1:0.10008)79.00:0.03091)100.00:0.07555,(Homo_sapiens_GLI1:0.03378,(Rattus_norvegicus_GLI1:0.01355,Mus_musculus_GLI1:0.01355)100.00:0.02023)100.00:0.17277)100.00:0.13641)93.00:0.11052):0.05211)1.00:0.07120):0.08658)100.00:0.00000;
```

**Figure: Newick tree file represent trees with edge lengths using commas and parentheses.**

Newick tree file contain distances and leaf name. Leaf name and distance length is separated by colon sign (:). Subtree are ended with semicolon sign (;).

```
<phyloxml xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="http://www.phyloxml.org" xsi:schemaLocation="http://www.phyloxml.org
http://www.phyloxml.org/1.10/phyloxml.xsd">

  <phylogeny rooted="false">

    <clade>

      <branch_length>0.0</branch_length>

      <confidence type="unknown">100.0</confidence>

    </clade>

    <name>Ciona_savignyi_GLI</name>

    <branch_length>0.34054</branch_length>

  </clade>

</phylogeny>

</phyloxml>
```

**Figure: Phylogenetic trees are stored in XML files.**

The Phyloxml file contains information such as clade branch lengths, species names, and support values for each clade. The Clade is the fundamental structural element of these phylogenetic trees, with each clade including a recursive succession of clades.