

ORTHOSCOPE: instruction

This page is an instruction for [ORTHOSCOPE](#).

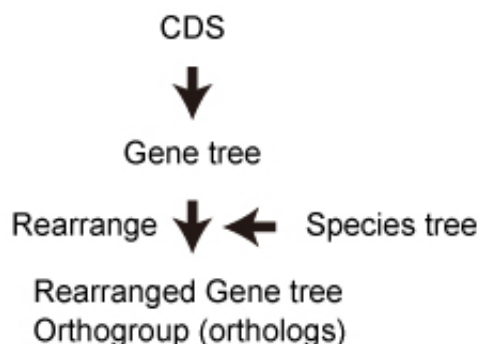
Modified: 17 Jul. 2018

Two analyses (mode) of ORTHOSCOPE

Tree search only



Search & rearrangement



Query sequences: Inoue et al.

Query sequences from genes with known function

Actinopterygii	Vertebrata	Deuterostomia	Protostomia
PLCB1*	ALDH1A	Bra	Bra
Queries	Queries	Queries	Queries
Result	Result	Result	Result

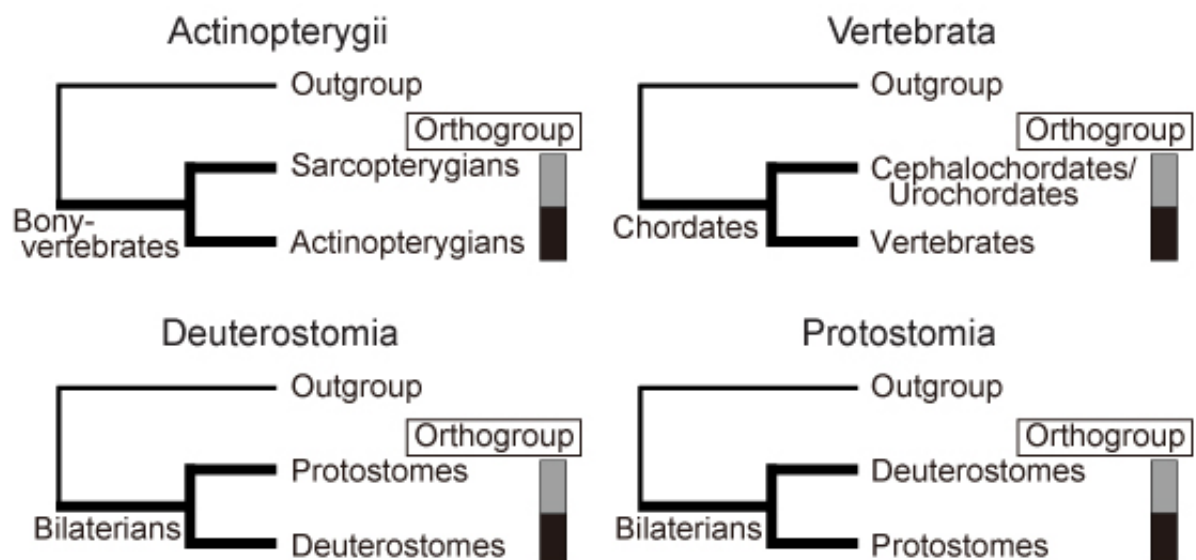
Query sequences from assemble database*

1. Download *Coregonus lavaretus* TSA file ([GFIG000000000.1](#)) from NCBI.
2. Translate raw sequences into amino acid and cDNA sequences using [TrandDecoder](#).
`./TransDecoder.LongOrfs -t GFIG01.1.fsa_nt`
3. Make blast databases using [BLAST+](#).
`makeblastdb -in longest_orfs.pep -dbtype prot -parse_seqids`
`makeblastdb -in longest_orfs.cds -dbtype nucl -parse_seqids`
4. BLASTP seaech against amino acid database.
`blastp -query query.txt -db longest_orfs.pep -num_alignments 10 -`
`evaluate 1e-12 -out 010_out.txt`
5. Retrieve blast top hit sequences from cDNA file using seq id.
`blastdbcmd -db longest_orfs.cds -dbtype nucl -entry_batch`
`queryIDs.txt -out 020_out.txt`

Query sequences: example

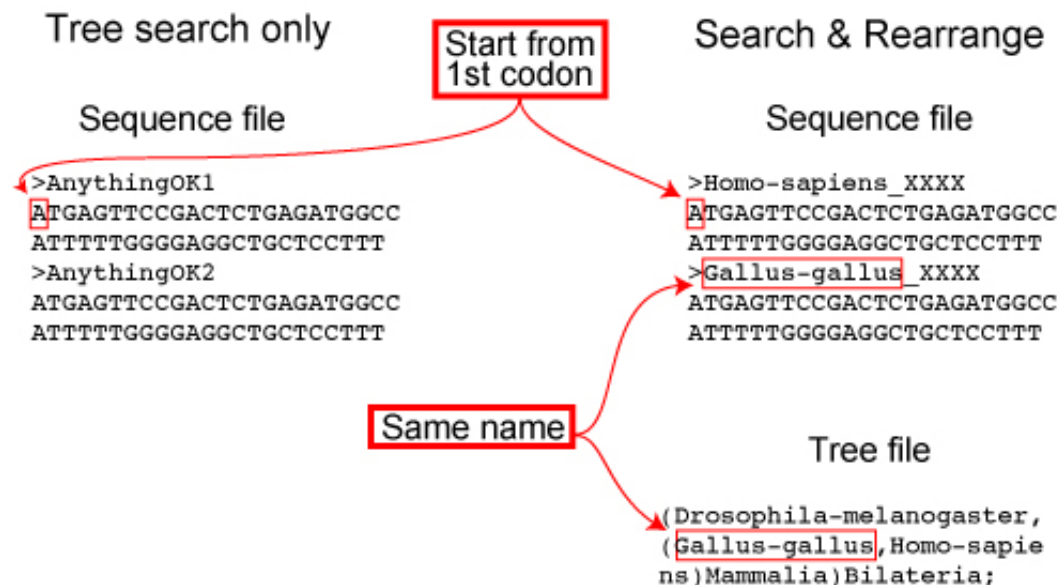
	Actinopterygii	Vertebrata	Deuterostomia	Protostomia
Bra	3queries.txt	6queries.txt	6queries.txt	3queries.txt
Actin	4queries.txt	4querist.txt	7queries.txt	9queries.txt
MHC	4queries.txt	10queries.txt	6queries.txt	2queries.txt

Analysis group



Upload files

Contents



Case 1: Query sequence is present in ORTHOSCOPE database

Coding sequence (fasta.txt)

Upload file

ファイルを選択

Homo_sapiens...uence.fa.txt

Species tree (newick.tre)

ファイルを選択

Metazoanのコピー.tre

Examples

Sequence file

```

1 >Halocynthia-roretzi_XXXXX
2 ATGCAGCGAAGCGTCACCTCTATACCACTTTACTG
3 GCAAGTTTCCCAAGTAACATCAATATAGGAGGATTGTC
4 GAGGTGTTCCGTTCCCTCTCGACCAACAGGACATA
5 GATATGGTCAAAATGGGAACAGCTTCTCCATGACATAC
6 AAAGGTGTGTACCCATCATCGGACTCTACGACCGAAG
7 TCTGTGGAGCGCTGCACGTCTGCTTTGTCACCCATCC
8 CAGTTCGTATCCAGTTAAGGCCGAGCTCCAGGACGCA
          
```

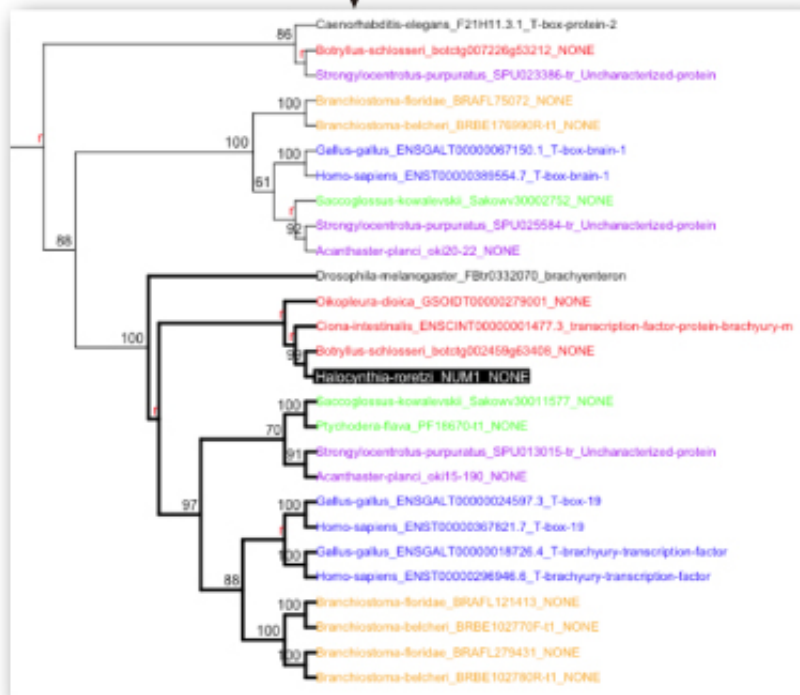
Tree file

```

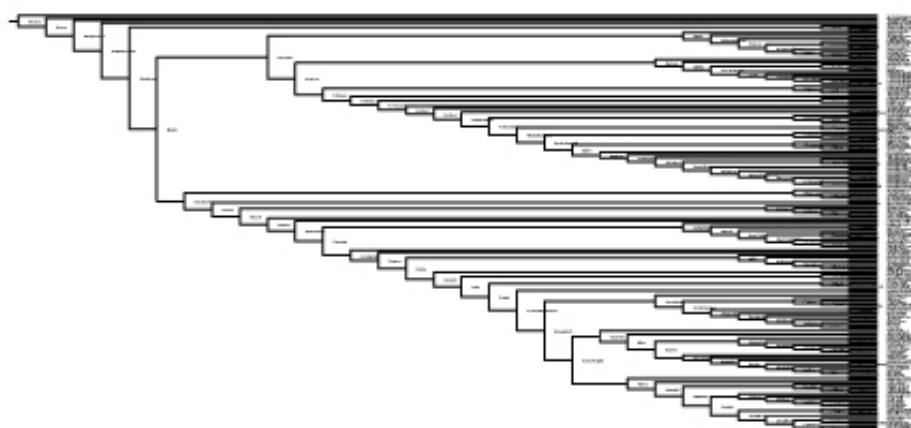
1. trotus-purpuratus, (Acanthaster-planci, Acanthaster-
   - Ambulacraria, ((Branchiostoma-floridae, Branchiostoma-
   - eura-dioica, ((Botryllus-schlosseri, Halocynthia-roretzi),
   - myi, (Ciona-intestinalis, Ciona-intestinalis), intestinalis),
   - troyzon-narinus, ((Callorhinchus-
   - atus, ((Scleropages-formosus, Paramyrion-
          
```

Modify

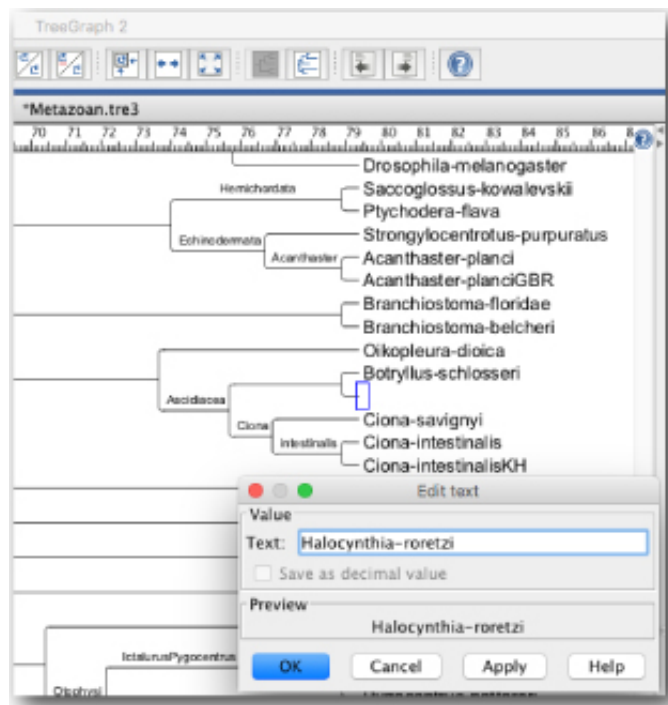
Add species name



Species tree hypothesis



The tree file can be modified using [TreeGraph2](#).



Sequence collection

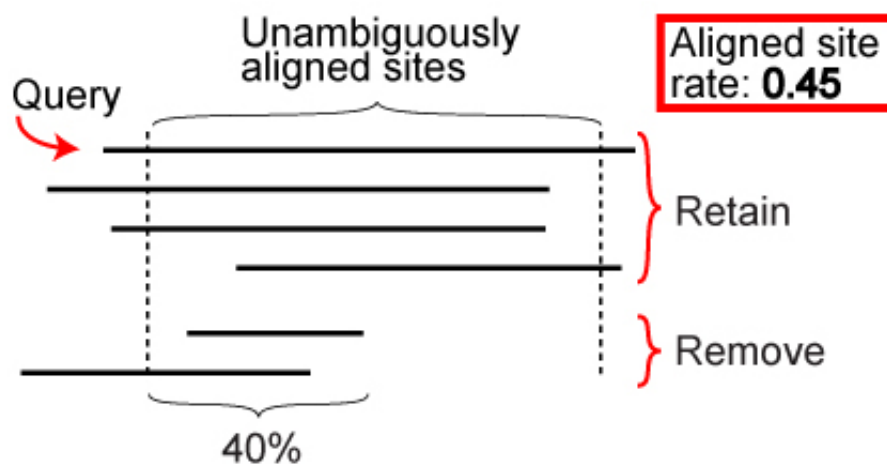
BlastP results against D.melanogaster

Hit sequences	E value
Drosophila-melanogaster_FBpp01_X	0.0
Drosophila-melanogaster_FBpp14_X	1e-7
Drosophila-melanogaster_FBpp34_X	1e-6
Drosophila-melanogaster_FBpp34_X	1e-5
Drosophila-melanogaster_FBpp67_X	1e-4
Drosophila-melanogaster_FBpp14_X	1e-3
Drosophila-melanogaster_FBpp24_X	1e-2
Drosophila-melanogaster_FBpp53_X	1e-1

E-value for reported sequence: **1e-3**

Number of hits to report: **5**

Alignment



Tree search

Dataset

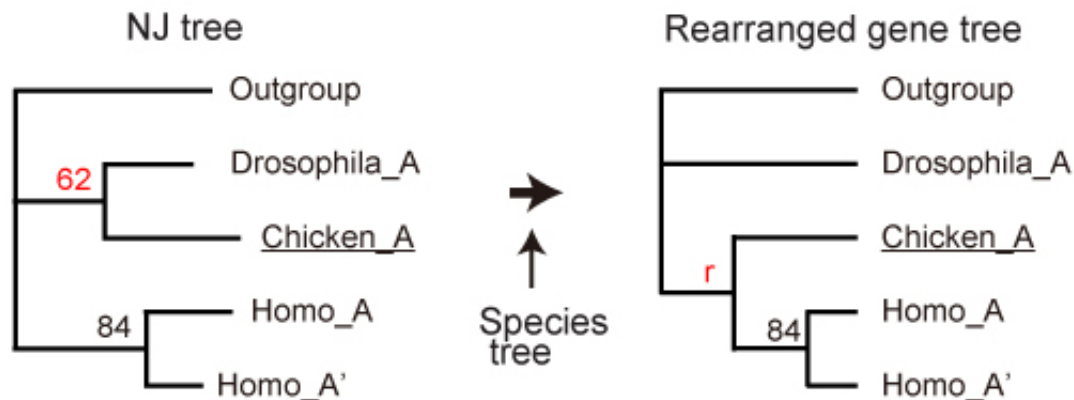
Dataset (3rd codon): Exclude

↑↑↑↑↑↑
 123123123
 Drosophila ATGAAAGTG
 Gallus ATGGGTAAG
 Homo ATGGGAACG
 ↓↓↓↓↓↓↓↓

Include

Rearrangement BS value threshold

Rearrangement BS value threshold: **70%**

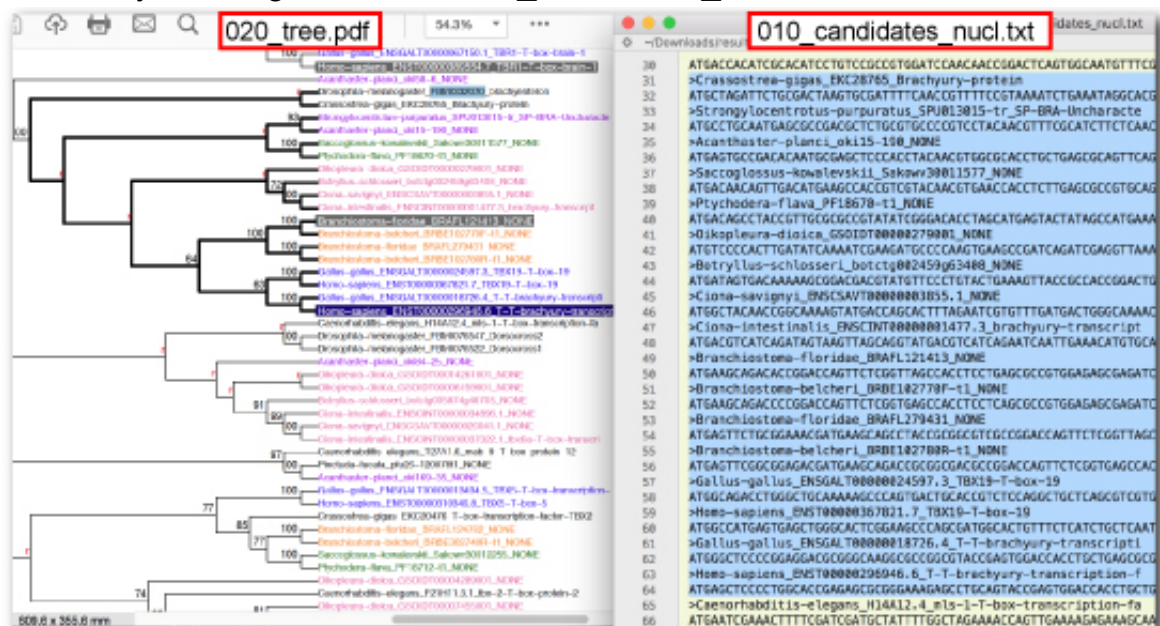


NJ analysis is conducted using the software package [Ape](#) in R (cDNA) and [FastME](#) (amino acid). Rearrangement analysis is done using a method implemented in [NOTUNG](#).

Tree estimation using identified orthologs

Mac only

1. Select only orthologs and save 010_candidates_nucl.txt file



2. Decompress 100_2ndTree.tar.gz file

3. cd into 100_2ndTree file

4. Run the pipeline

./100_estimate2ndTree.py

5. ML tree is saved is 200_RAxMLtree_Exc3rd.pdf.



Browser

Chrome	Firefox	Safari	IE
OK	OK	11.0 or later	Not supported

History

10 July 2018

Version 1.0.

Citation

Inoue J. and Satoh N. ORTHOSCOPE: an automatic web tool of analytical pipeline for ortholog identification using a species tree. in prep.

Previous versions

jun.inoue@toist.jp