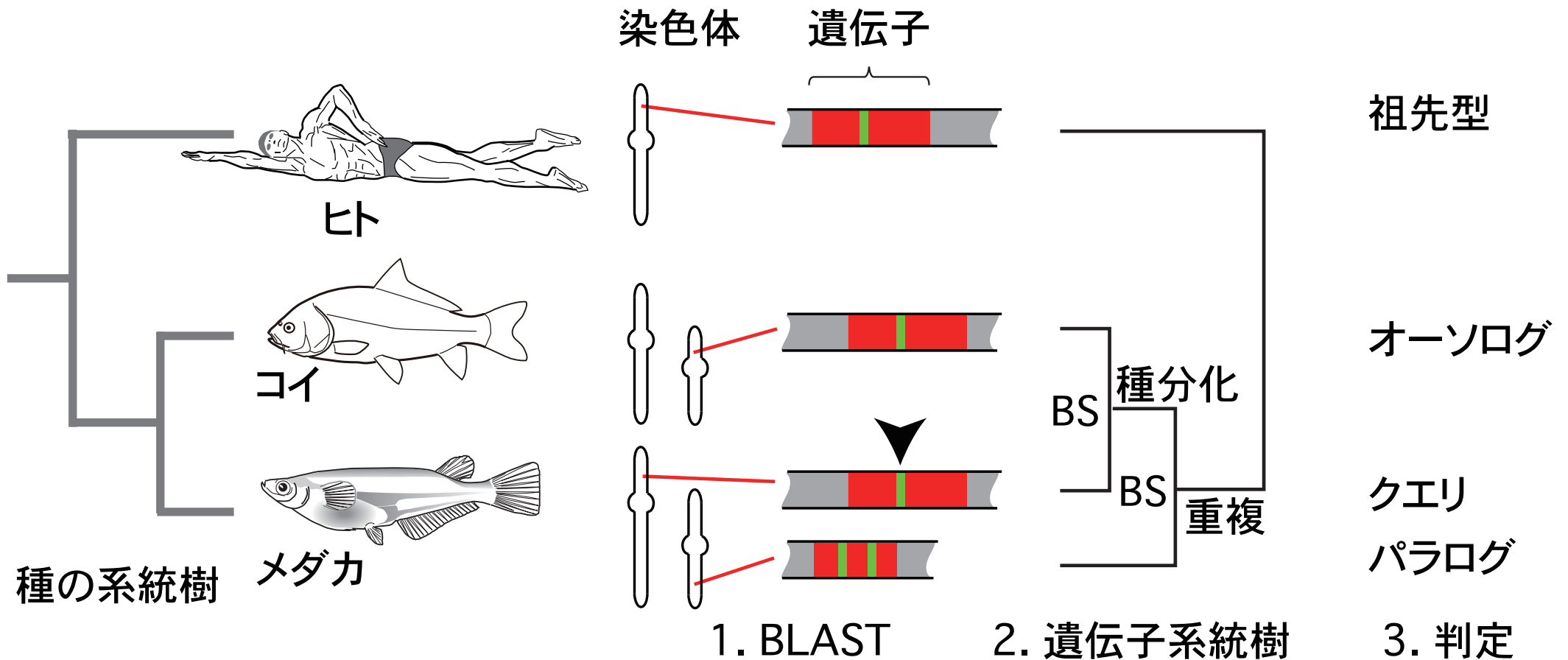


ORTHOSCOPE\* :  
ゲノムに存在する全遺伝子の  
歴史を推定する系統解析パイプライン

井上 潤

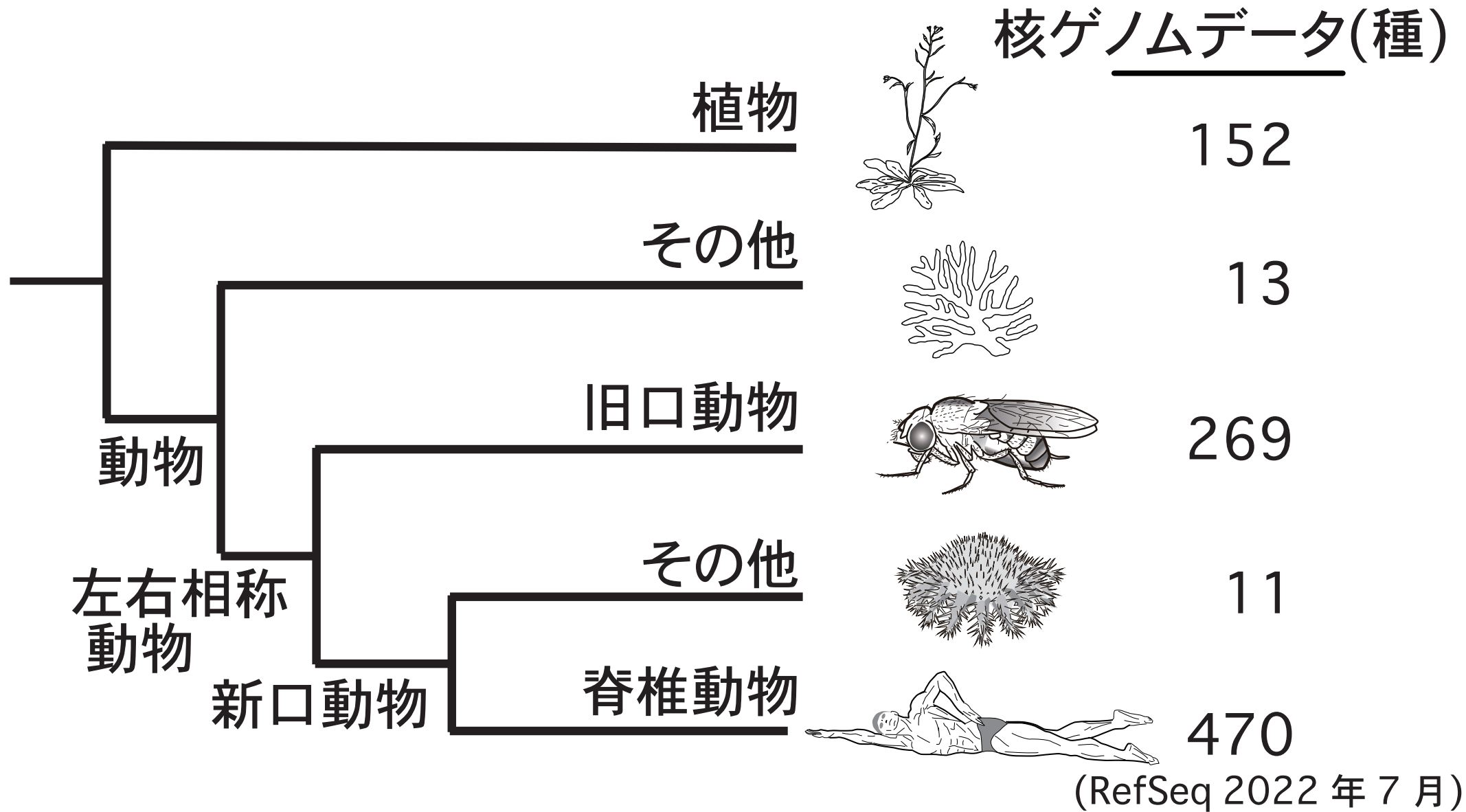
2022年9月8日  
動物学会

# ORTHOSCOPE\* は何ができる？



- この作業を全遺伝子について行う。
- 動植物 600 種のデータを使える。

# ゲノムデータが蓄積している



## ゲノムを比較すれば進化の謎に迫れる

# Web ツール ORTHOSCOPE を開発

## ORTHOSCOPE v1.5.2 - AORI: yurai

Gene tree and orthogroup estimation using a species tree (< 5 min as default). 2 July 2022  
[Instructions](#), [Species tree](#) Support: Safari(except for v14), Firefox, Chrome

---

**NEWS** A script for genome wide data, [ORTHOSCOPE\\*](#), has been published.

---

**Focal group**  
[Actinopterygii](#) [Mammalia](#) [Vertebrata](#) **Deuterostomia** [Protostomia](#) [Acropora](#) [Plants](#)

---

**Status** Ready.

---

**Mode**  Estimating gene tree  Comparing gene and species trees

---

**Upload file** Sequences (fasta):  DNA  Amino acid Example fasta file: [DNA](#), [Amino acid](#)

Species tree (newick):  ファイル未選択 If not selected, [this tree](#) is used.

---

**Sequence collection**  
BLAST E-value threshold for reported sequences  1e-5  1e-4  1e-3  1e-2  1e-1  1  
Number of BLAST hits to report per genome  3  5  10  20

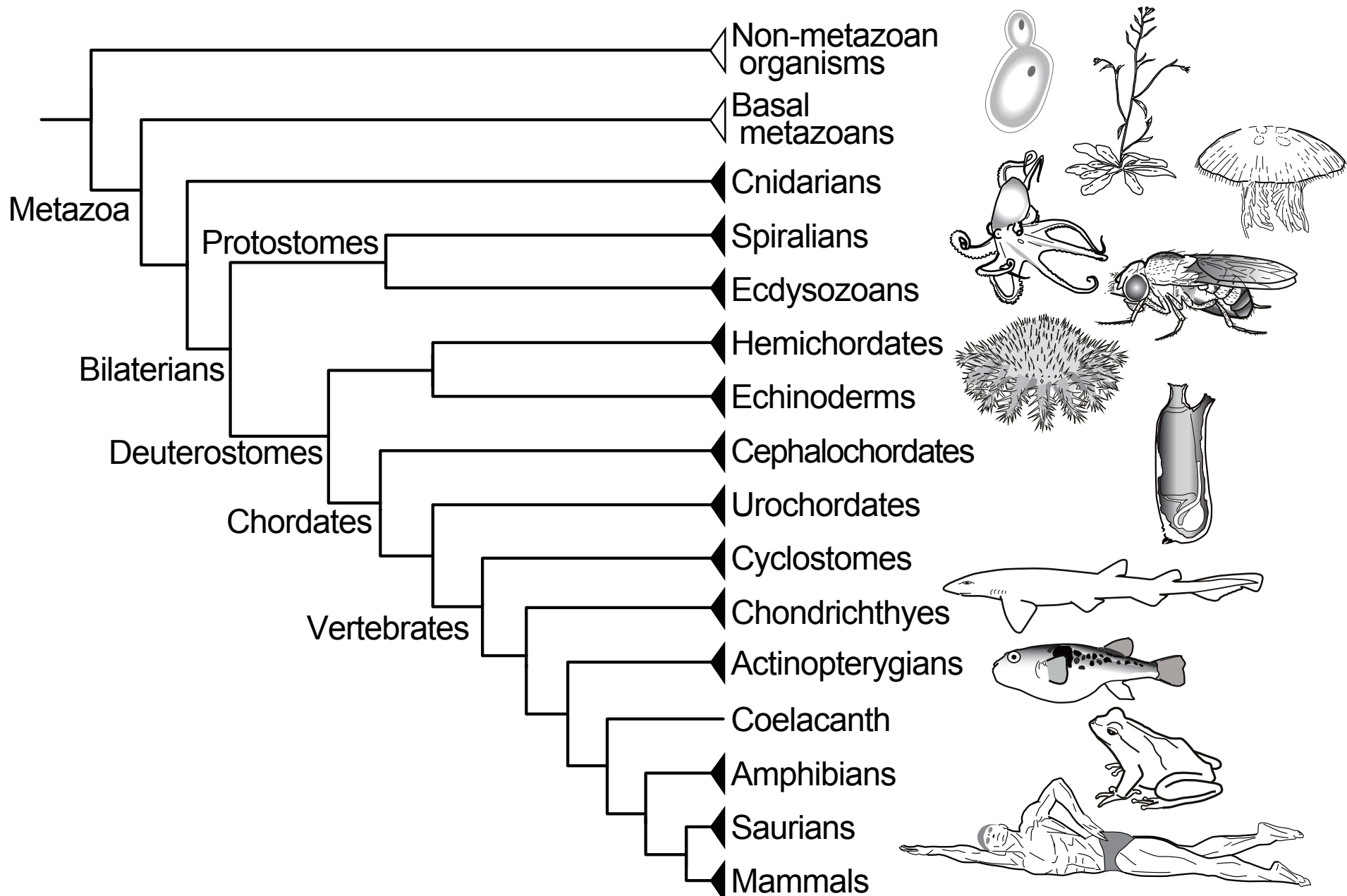
**Alignment**  
Aligned site rate threshold within unambiguously aligned sites  0  0.2  0.4  0.55

**Tree search**  
Dataset  Amino acid  DNA (Exclude 3rd)  DNA (Include 3rd)  
Rearrangement BS value threshold  60%  70%  80%  
Key node Nephrozoa or Bilateria

BLAST  
+  
遺伝子系統樹

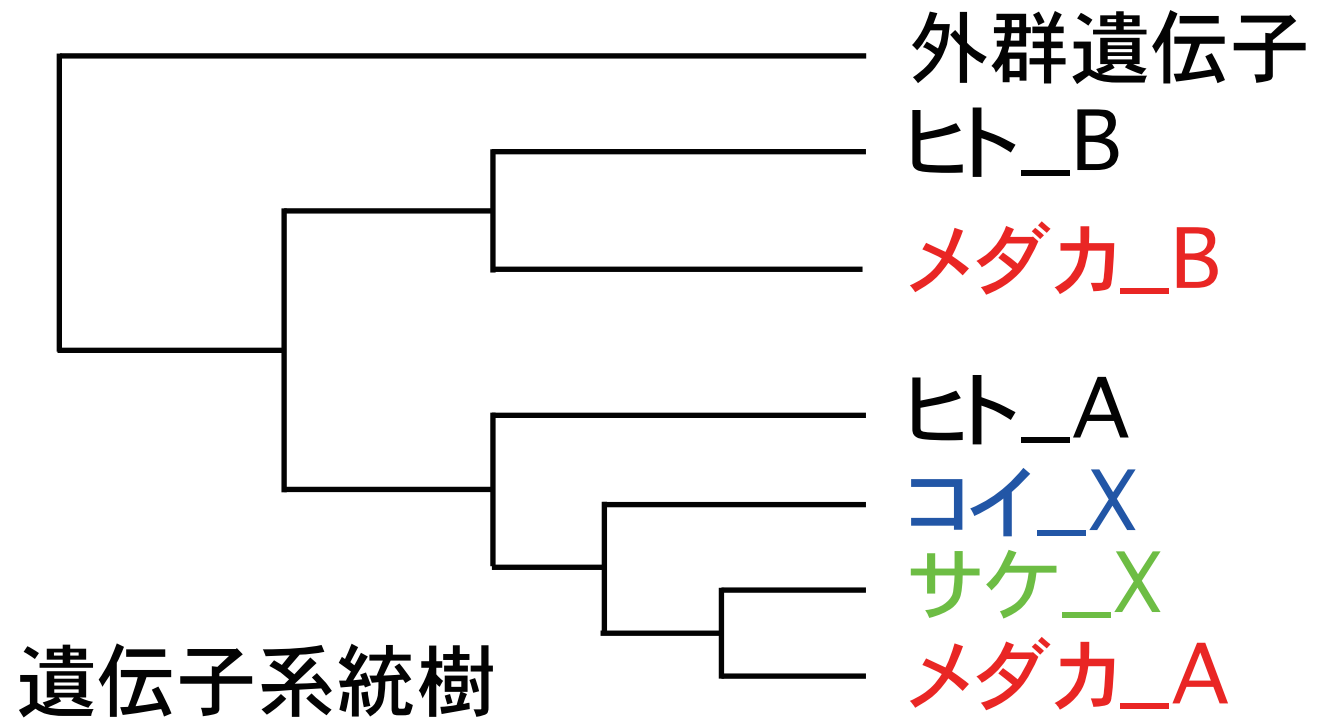
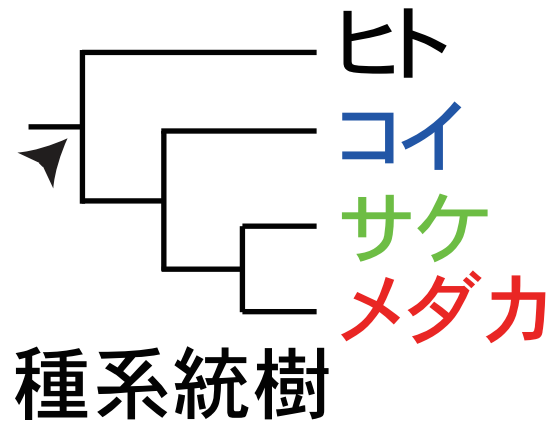
*Inoue and Satoh (2019) MBE*

# ORTHOSCOPE のデータベース

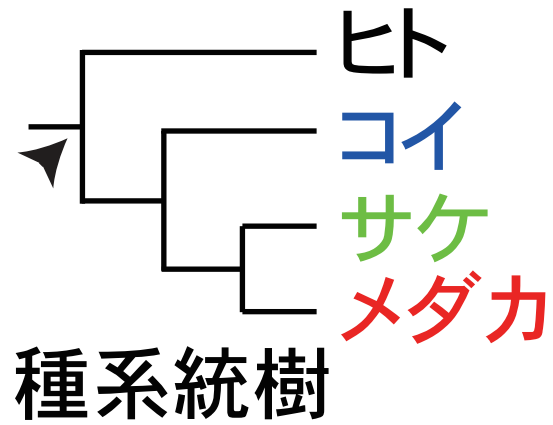


動物 550 種、植物 50 種の遺伝子モデル

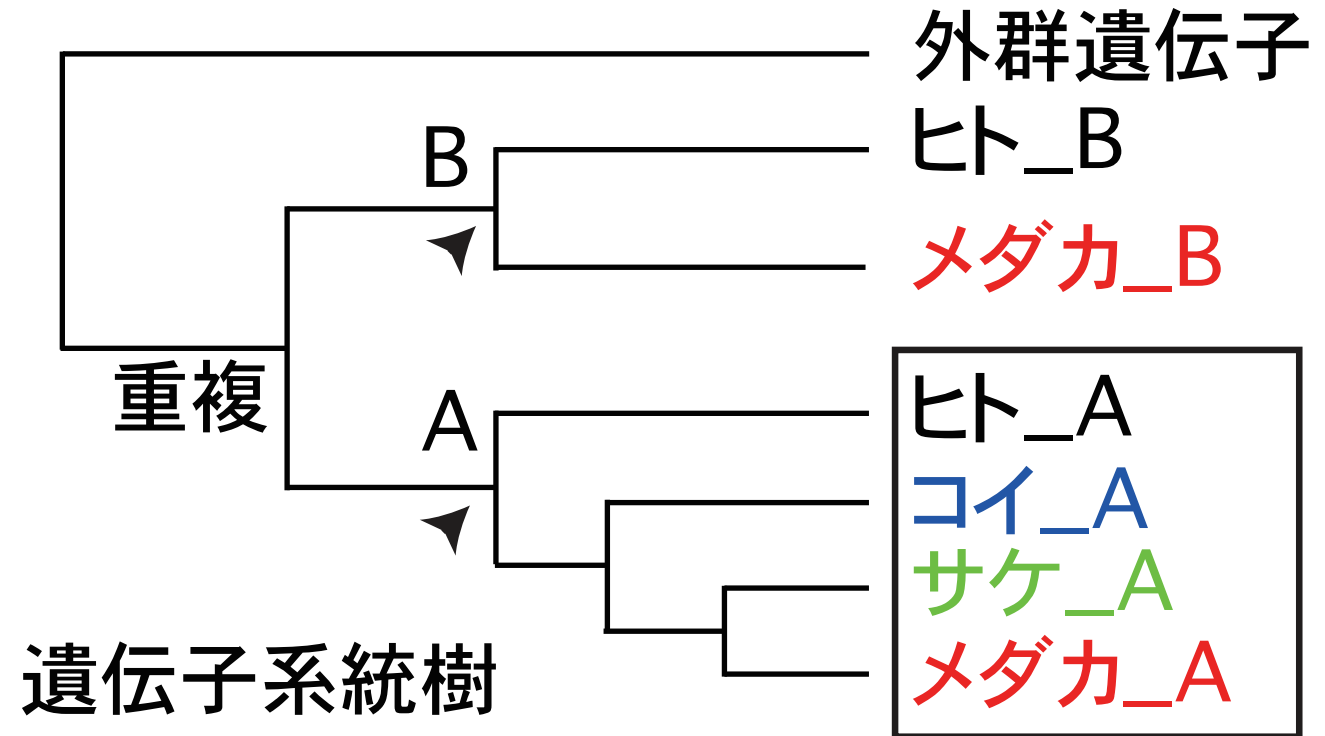
# ORTHOSCOPE は何ができる？



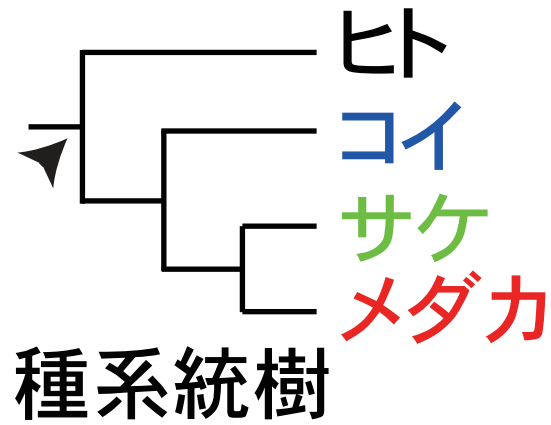
# ORTHOSCOPE は何ができる？



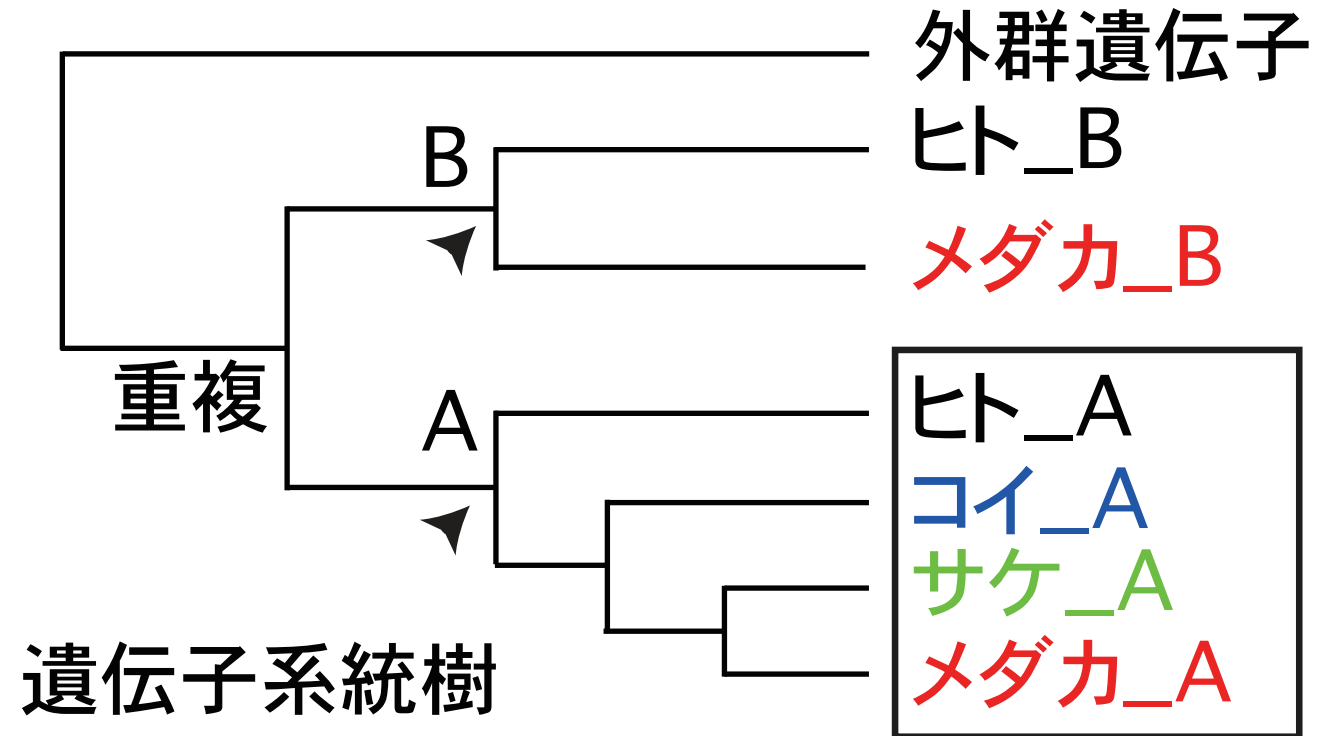
オーソグループを見つけて、



# ORTHOSCOPE は何ができる？



オーソグループを見つけて、



- 遺伝子の機能が推定できる
- 遺伝子（遺伝子座）の数がカウントできる



# ORTHOSCOPE demonstration

## ORTHOSCOPE v1.5.2 - AORI: yurai

Gene tree and orthogroup estimation using a species tree (< 5 min as default).

2 July 2022

[Instructions](#), [Species tree](#)

Support: Safari(except for v14), Firefox, Chrome

**NEWS** A script for genome wide data, [ORTHOSCOPE\\*](#), has been published.

### Focal group

[Actinopterygii](#)

[Mammalia](#)

[Vertebrata](#)

[Deuterostomia](#)

[Protostomia](#)

[Acropora](#)

[Plants](#)

### Status

Ready.

Execute

### Mode

Estimating gene tree

Comparing gene and species trees

### Upload file

Sequences (fasta):  DNA  Amino acid

Example fasta file: [DNA](#), [Amino acid](#)

Species tree (newick):  ファイル未選択

If not selected, [this tree](#) is used.

### Sequence collection

BLAST E-value threshold for reported sequences

1e-5  1e-4  1e-3  1e-2  1e-1  1

Number of BLAST hits to report per genome

3  5  10  20

### Alignment

Aligned site rate threshold  
within unambiguously aligned sites

0  0.2  0.4  0.55

### Tree search

Dataset

Amino acid  DNA (Exclude 3rd)  DNA (Include 3rd)

Rearrangement BS value threshold

60%  70%  80%

Key node

Nephrozoa or Bilateria

## Deuterostomia

### Hemichordata

- Saccoglossus kowalevskii* Helical acorn worm [OIST-R](#) 34239
- Saccoglossus kowalevskii* 1 Helical acorn worm [RefSeq89](#) 20922
- Ptychodera flava* Yellow acorn worm OIST-S 34647

### Echinodermata

- Strongylocentrotus purpuratus* Purple urchin EnsMet38 28842
- Strongylocentrotus purpuratus* 1 Purple urchin [RefSeq89](#) 27728
- Acanthaster planci* Crown-of-thorns starfish (Okinawa) OIST-S 24323
- Acanthaster planci* 1 Crown-of-thorns starfish (Okinawa) [RefSeq89](#) 16468
- Acanthaster planci* GBR Crown-of-thorns starfish (Great Barrier Reef) OIST-S 24747

### Cephalochordata

- Asymmetron lucayanum* Lancelet [TSA](#) 104311
- Branchiostoma belcheri* Belcher's lancelet [LanceletDB](#) 30392
- Branchiostoma belcheri* 1 Belcher's lancelet [RefSeq89](#) 23855
- Branchiostoma lanceolatum* Mediterranean amphioxus [Amphicode](#) 23586
- Branchiostoma floridae* Florida lancelet [JGI](#) 21954
- Branchiostoma floridae* 1 Florida lancelet [RefSeq89](#) 28621

### Urochordata

#### Appendicularia

- Oikopleura dioica* Pelagic tunicate [OikoBase](#) 17217

#### Thaliacea

- Salpa thompsoni* Salpa [TSA](#) 76798

#### Enterogona

- Ciona savignyi* Pacific transparent sea squirt [Ens91](#) 11616
- Ciona savignyi* 1 Pacific transparent sea squirt aNISEED 20115
- Ciona intestinalis* Transparent sea squirt Ens91 16671
- Ciona intestinalis* 1 Transparent sea squirt [RefSeq89](#) 13520

#### Molgulidae

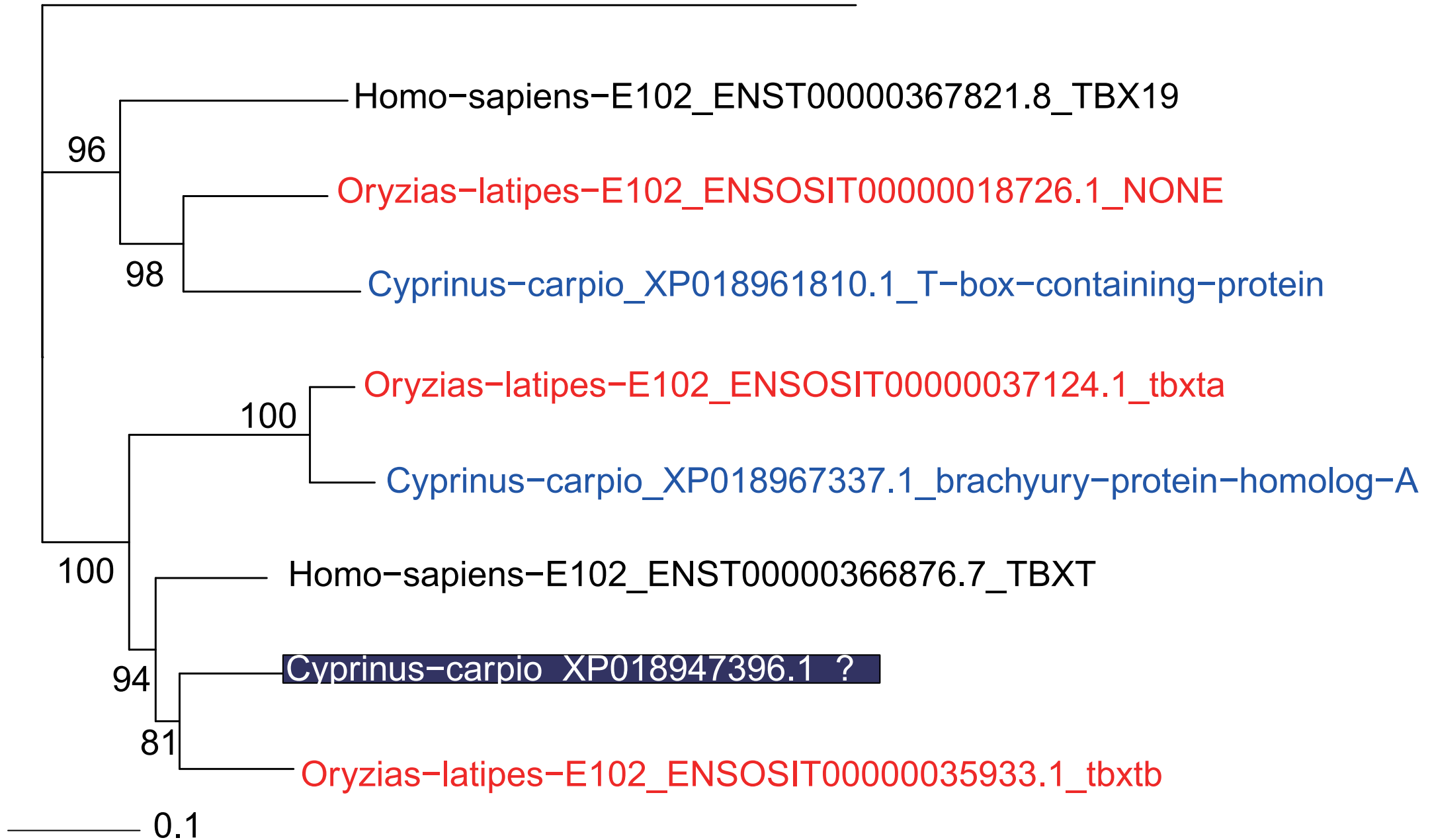
- Molgula occidentalis* — [aNISEED](#) 33023
- Molgula oculata* Sea grape aNISEED 16616

#### Stvelidae

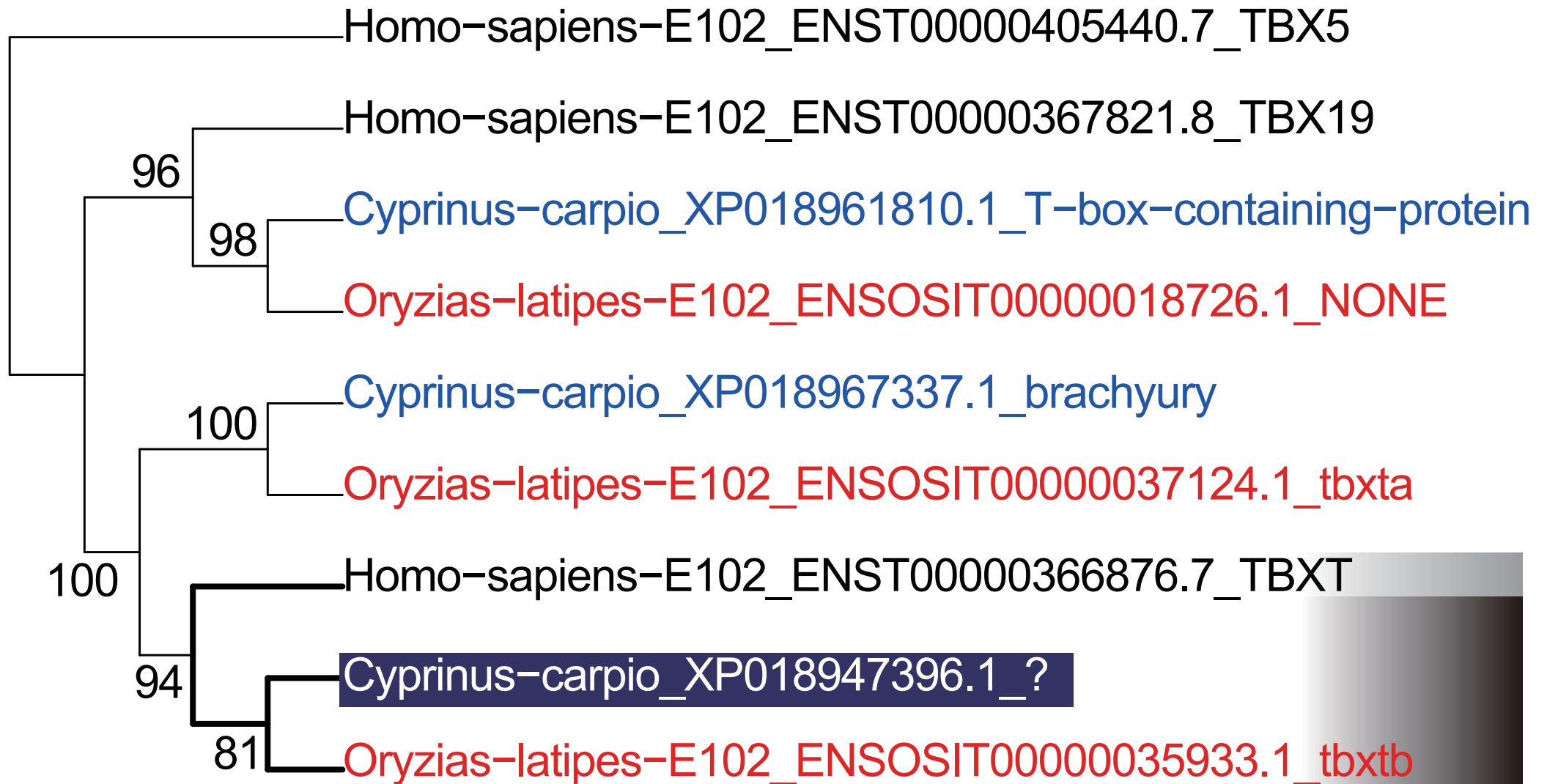
- Botrylloides leachii* Leache's ascidian aNISEED 15839

# 結果：遺伝子系統樹

Homo-sapiens-E102\_ENST00000405440.7\_TBX5



# 結果: 種の系統樹と比較



モデル生物の情報を参照



# 全遺伝子を解析したい

## ORTHOSCOPE

### 1 遺伝子

### Web ツール

*Inoue and Satoh (2019) MBE*

# 全遺伝子を解析したい

ORTHOSCOPE

1 遺伝子

Web ツール

*Inoue and Satoh (2019) MBE*

ORTHOSCOPE\*

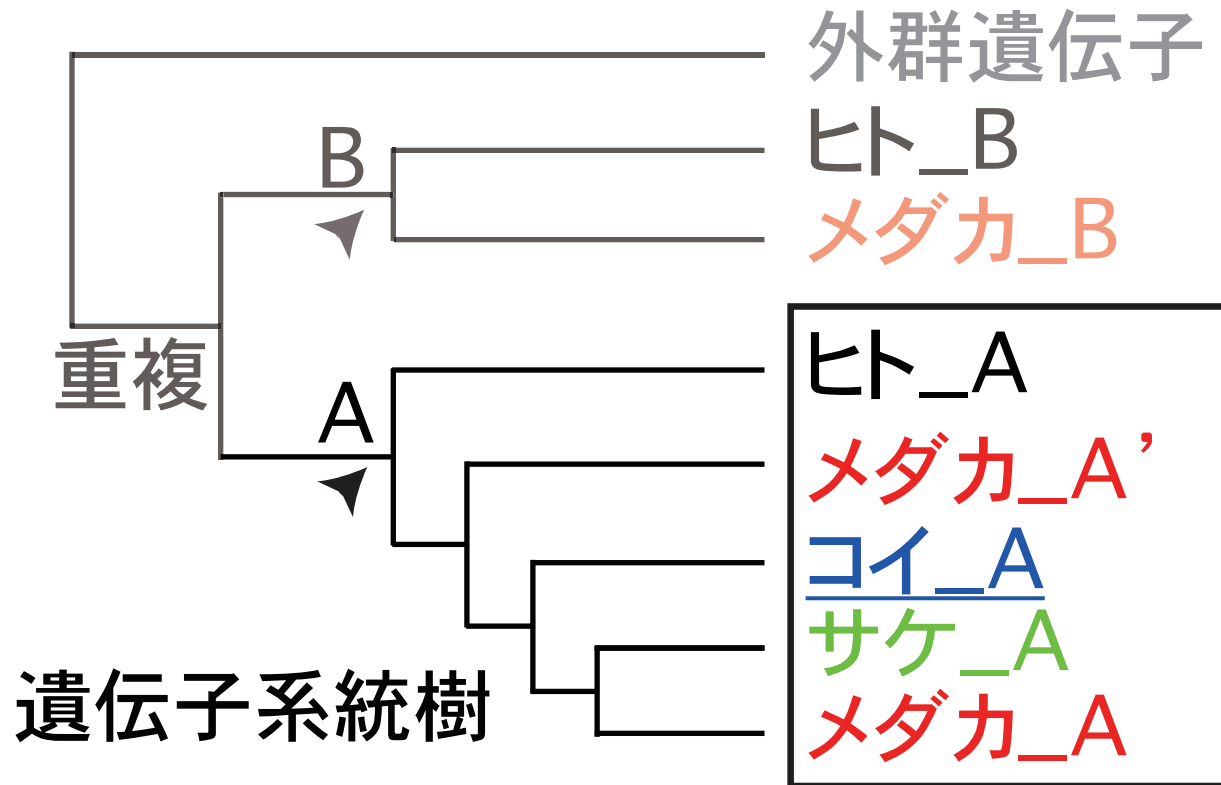
全遺伝子

スクリプト

*Inoue (2022) MBE*

# ORTHOSCOPE\*

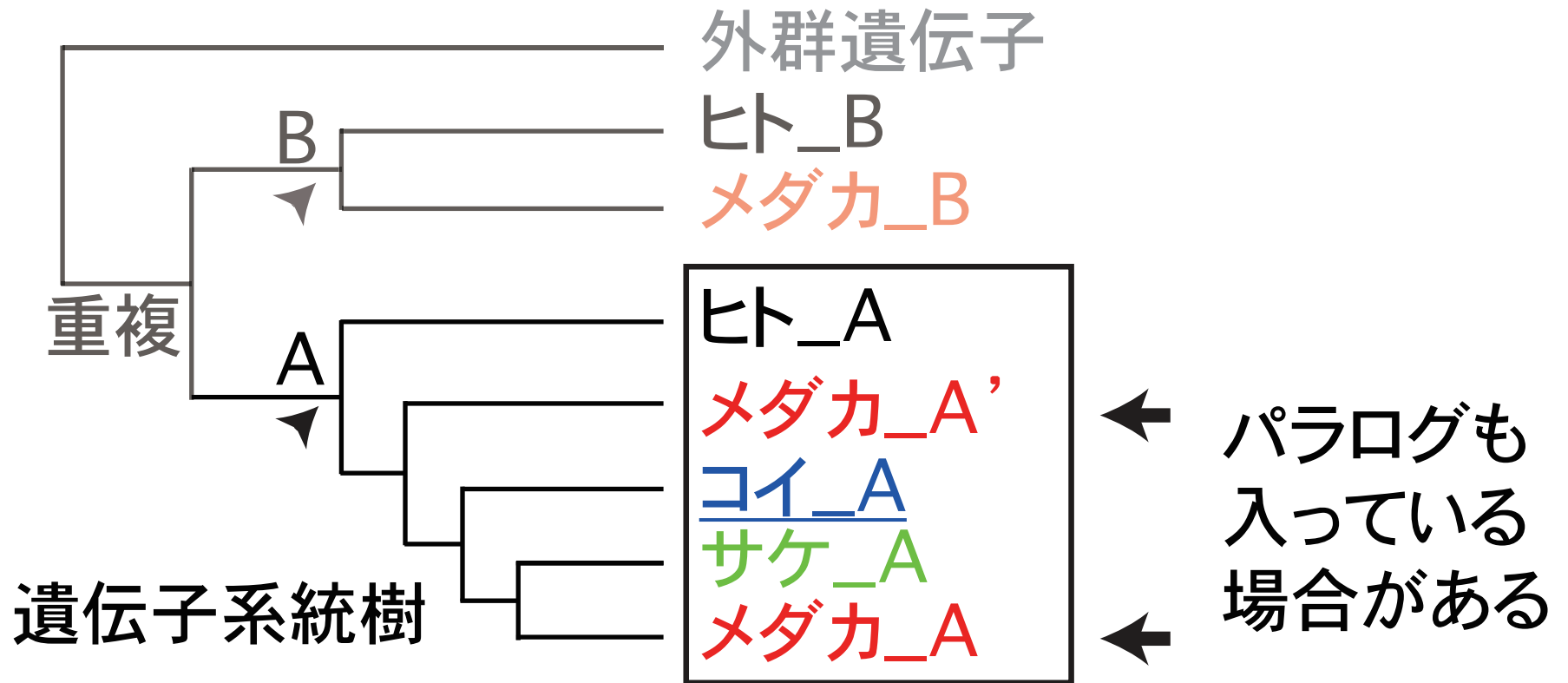
## オーソグループだけの系統樹も推定





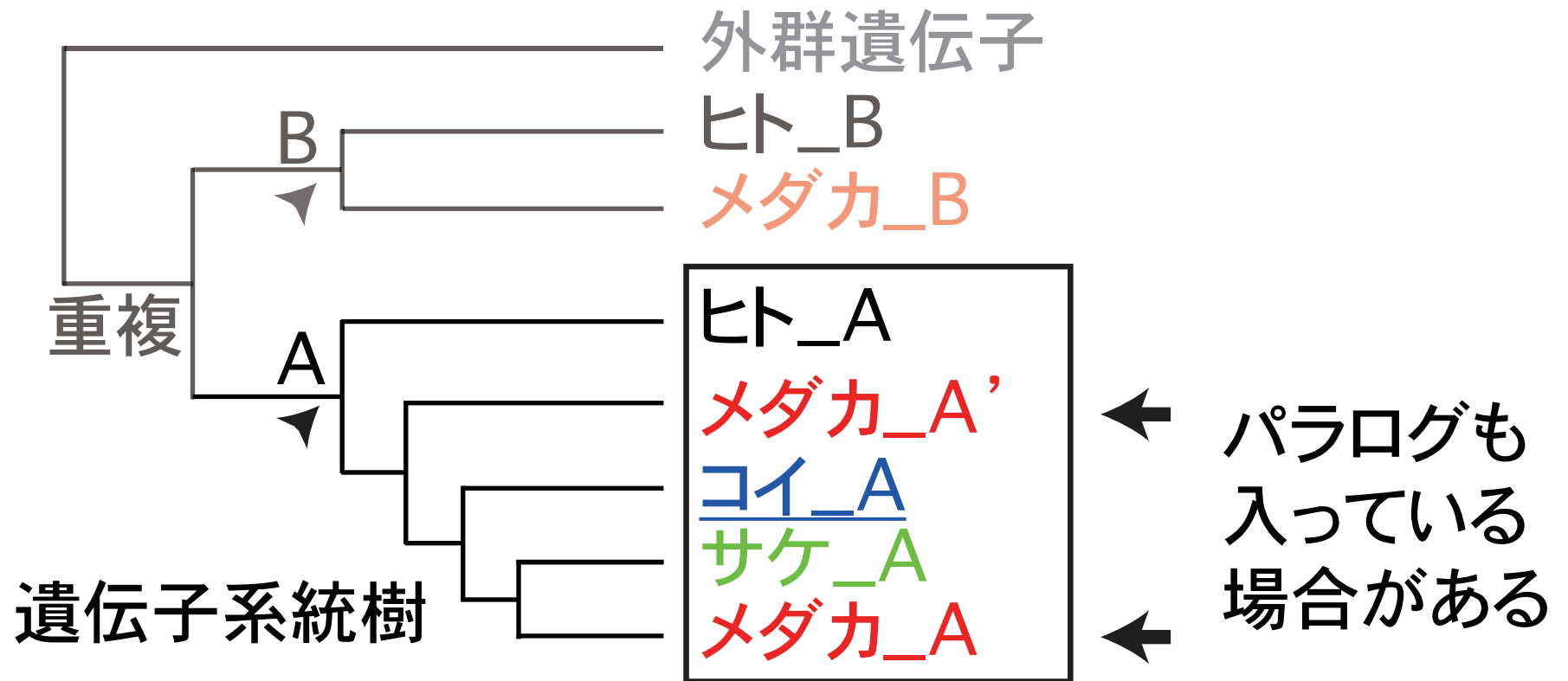
# ORTHOSCOPE\*

## オーソグループだけの系統樹も推定



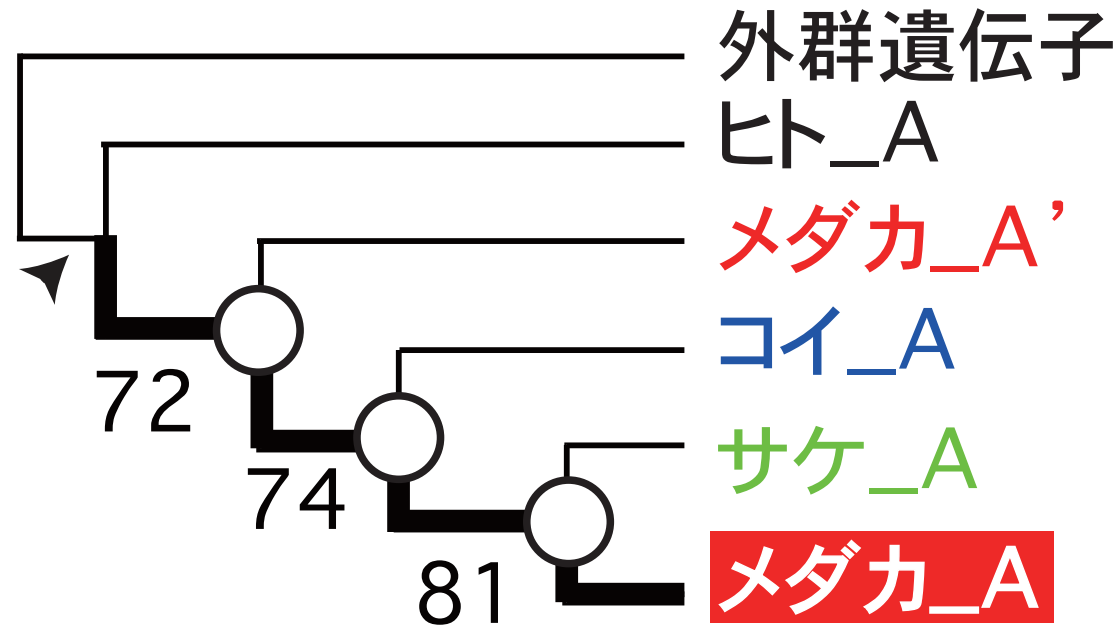
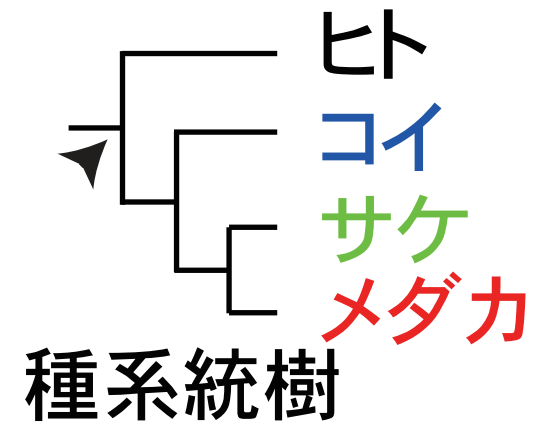
# ORTHOSCOPE\*

## オーソグループだけの系統樹も推定

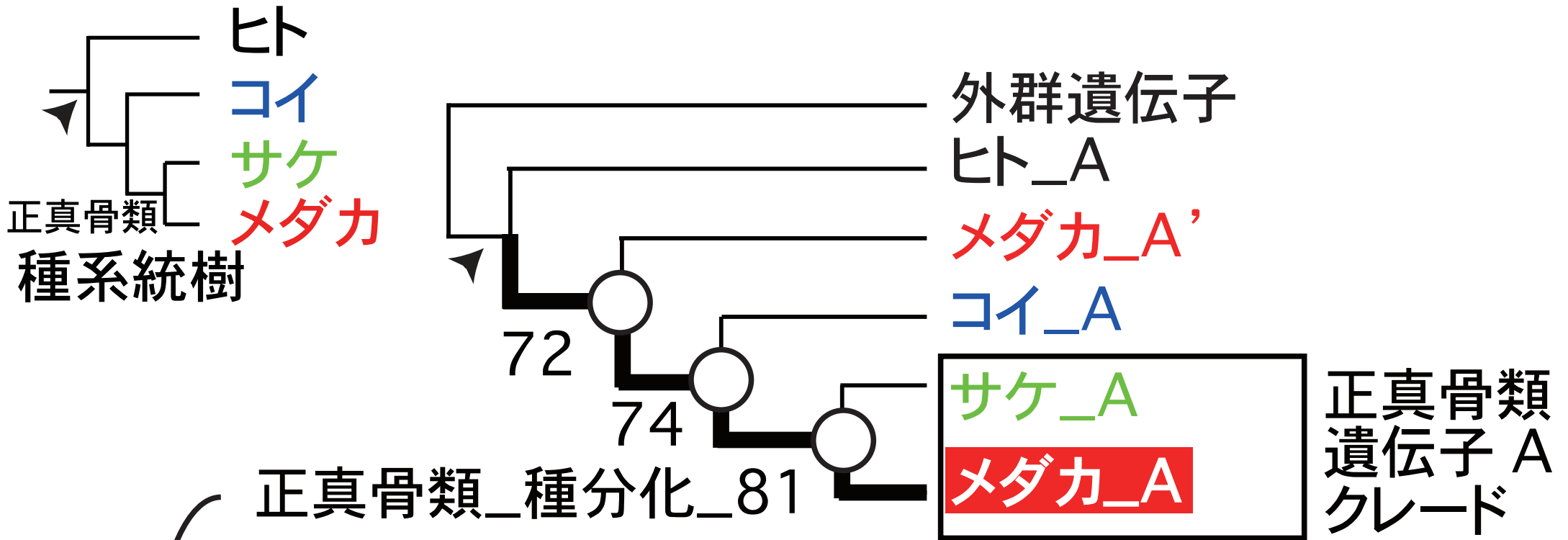


問題点: どれを Concatenate する?

# 遺伝子ごとの歴史を推定

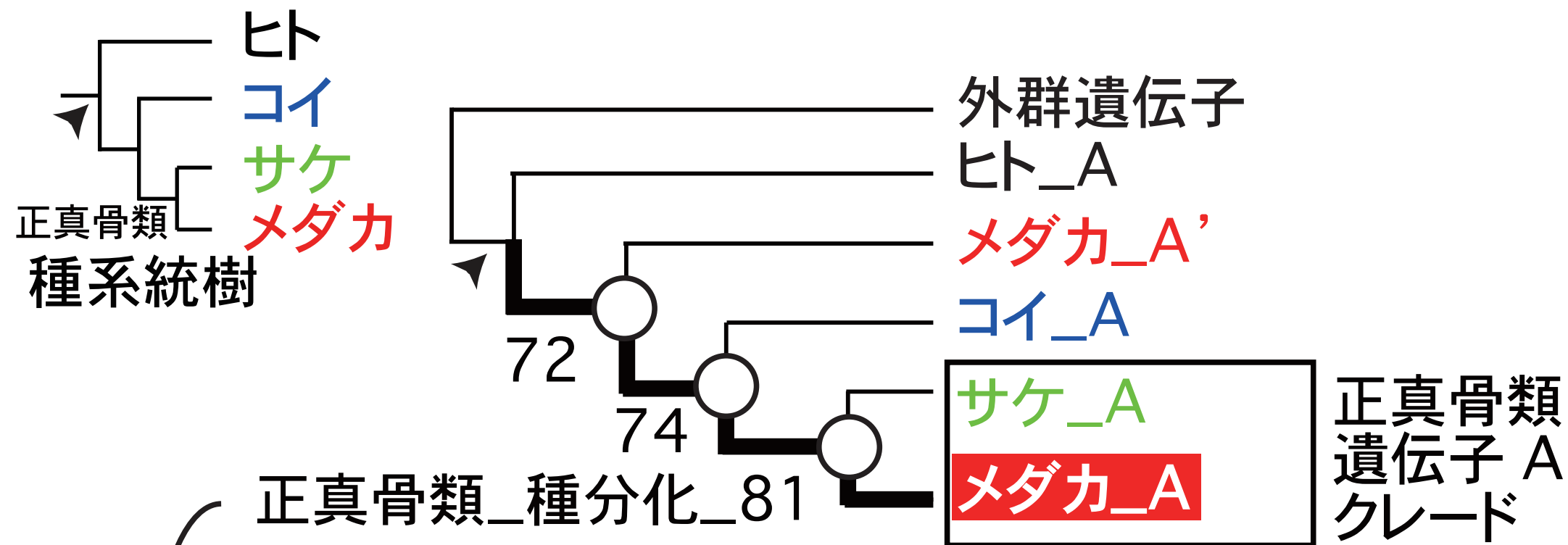


# 遺伝子ごとの歴史を推定



クレード名	重複/ 種分化	BS値	姉妹群	BS値
正真骨類	種分化	81	コイ	74

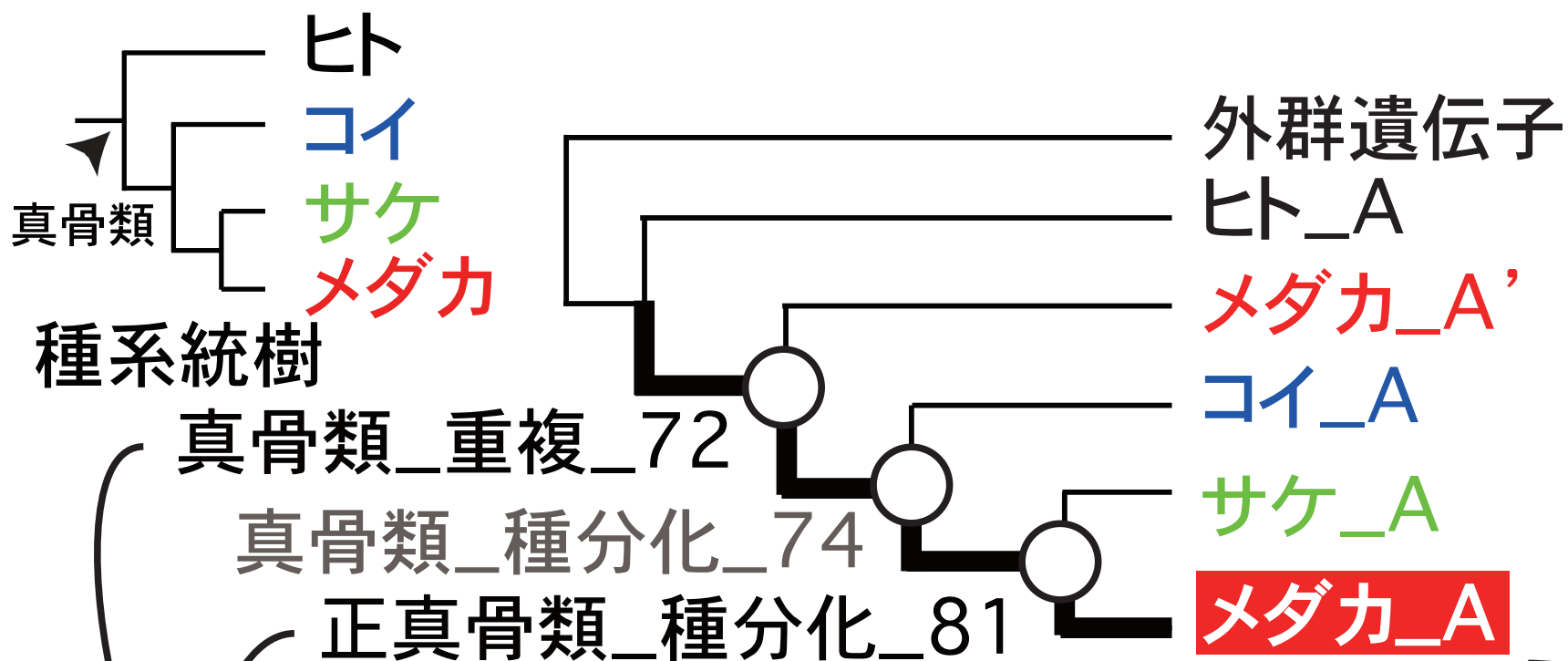
# 遺伝子ごとの歴史を推定



クレード名	重複/ 種分化	BS値	姉妹群	BS値
正真骨類	種分化	81	コイ	74

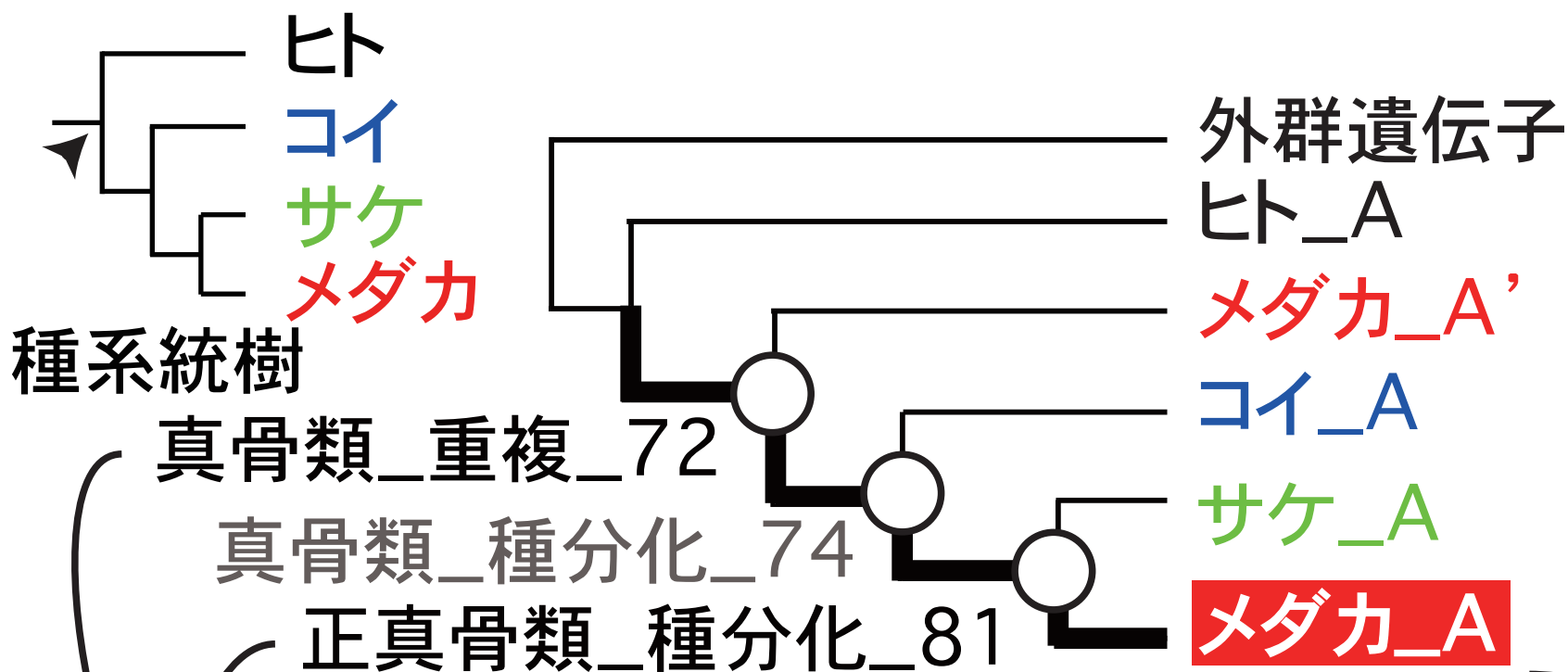
エクセルファイルとして出力

# 他のノードも記述



分岐名	重複/種分化	BS値	種の姉妹群	BS値
真骨類	重複	72	—	—
正真骨類	種分化	81	コイ	74
メダカ_A	—	—	サケ	81

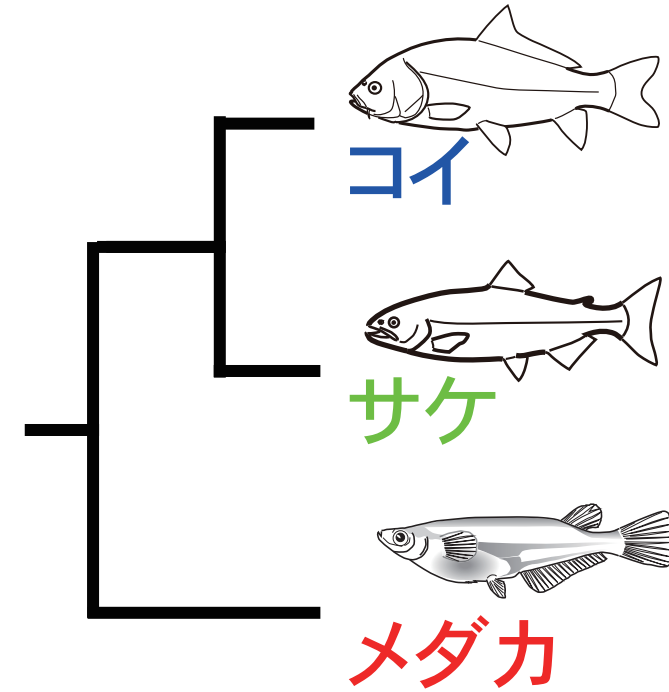
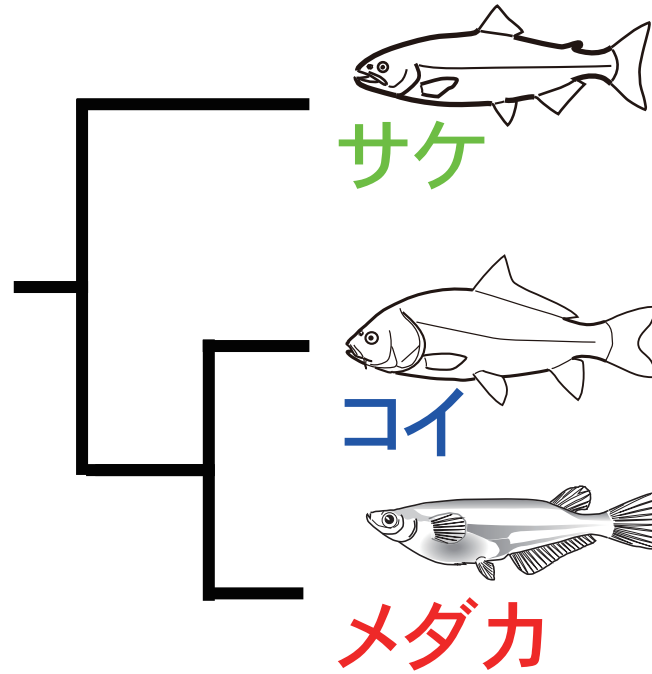
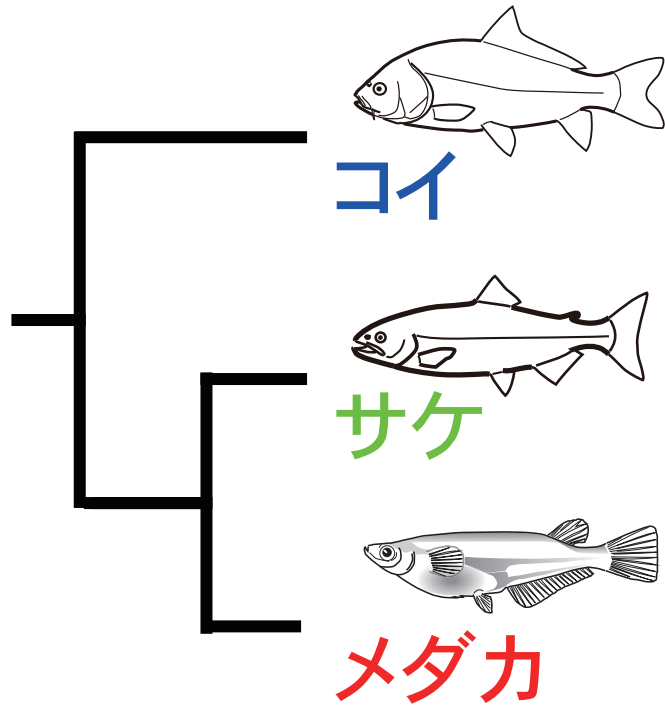
# 他のノードも記述



分岐名	重複/種分化	BS値	種の姉妹群	BS値
真骨類	重複	72	—	—
正真骨類	種分化	81	コイ	74
メダカ_A	—	—	サケ	81

精度の高い情報だけで、ゲノムの歴史を推定

# 1. 系統仮説を比較

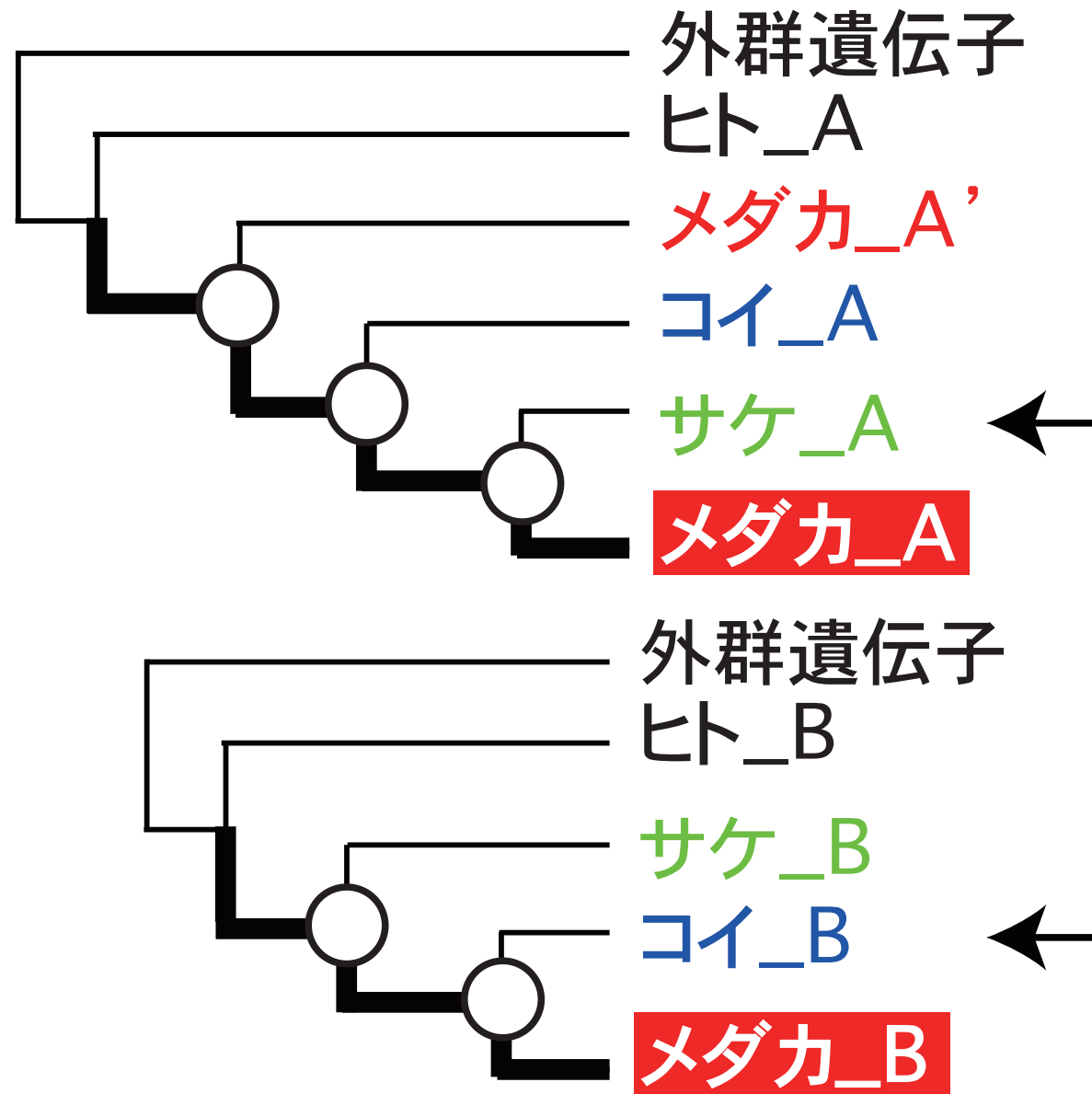


広く認められる関係

メダカの姉妹群は？

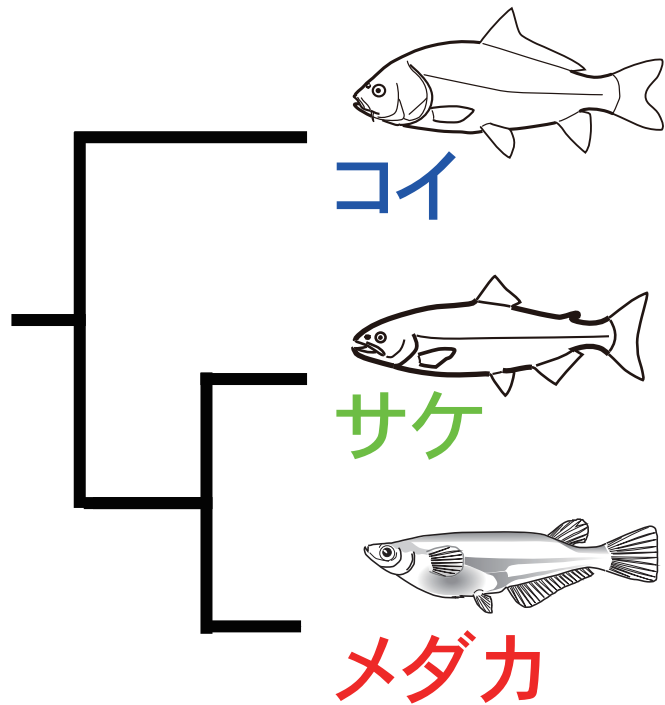


# 1. 系統仮説を比較

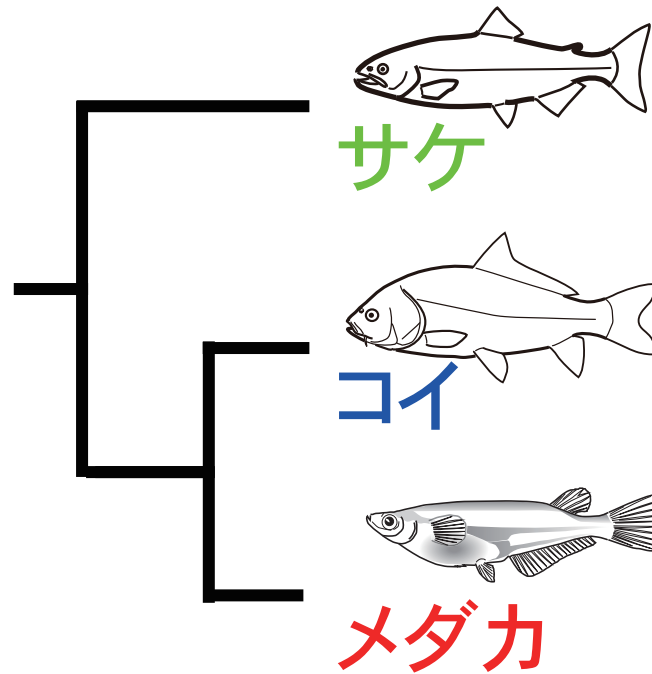


遺伝子ごとに姉妹群（種）を推定

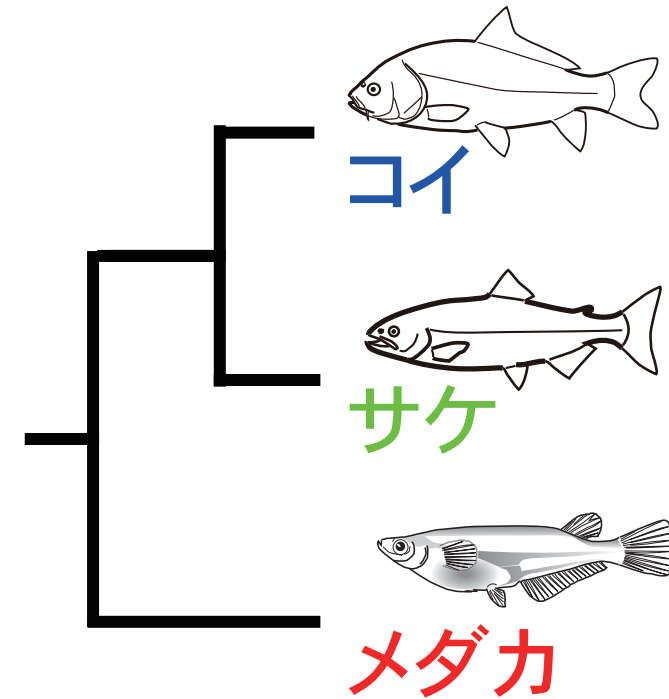
# 1. 系統仮説を比較



2520 遺伝子



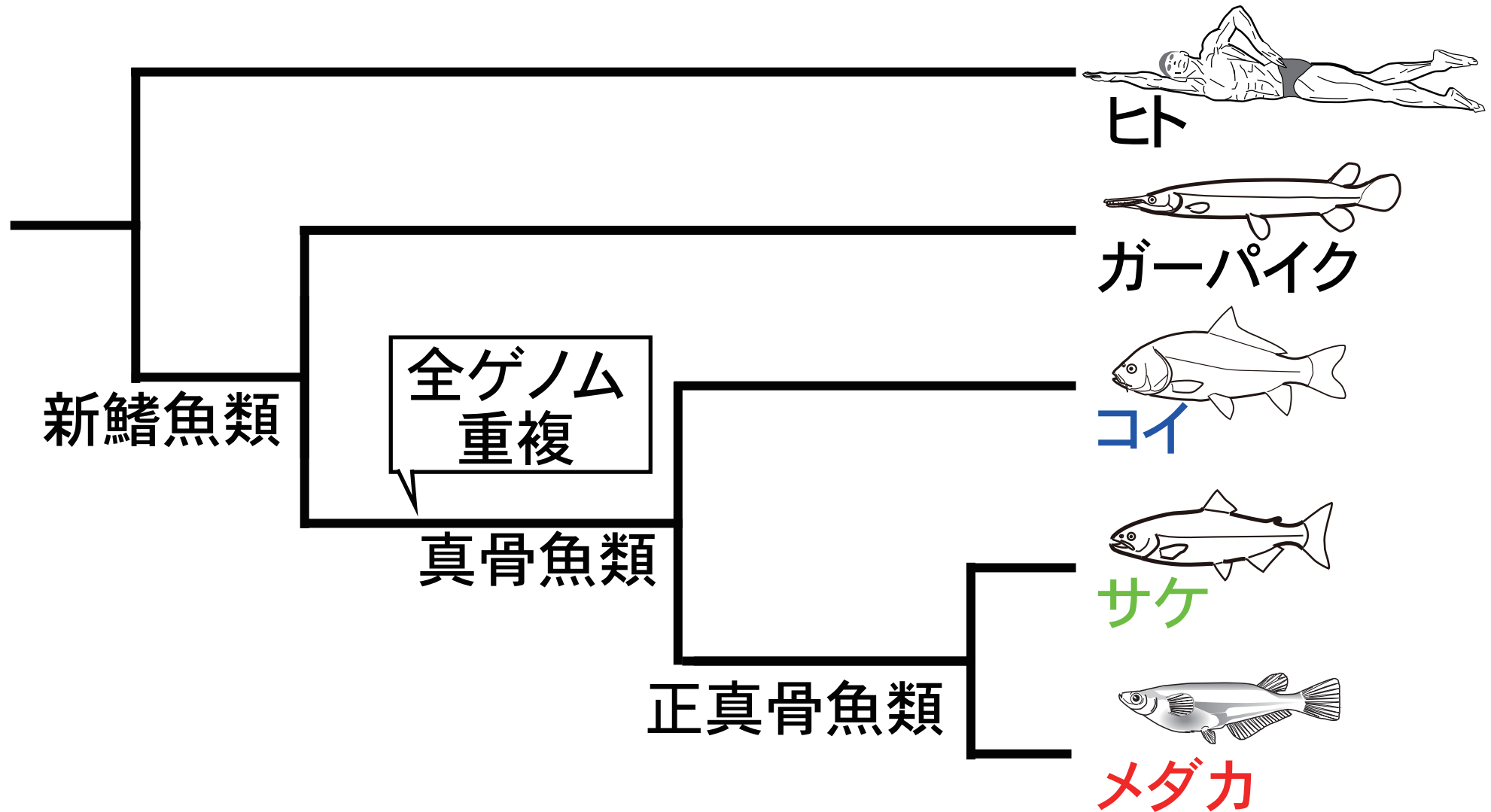
66 遺伝子



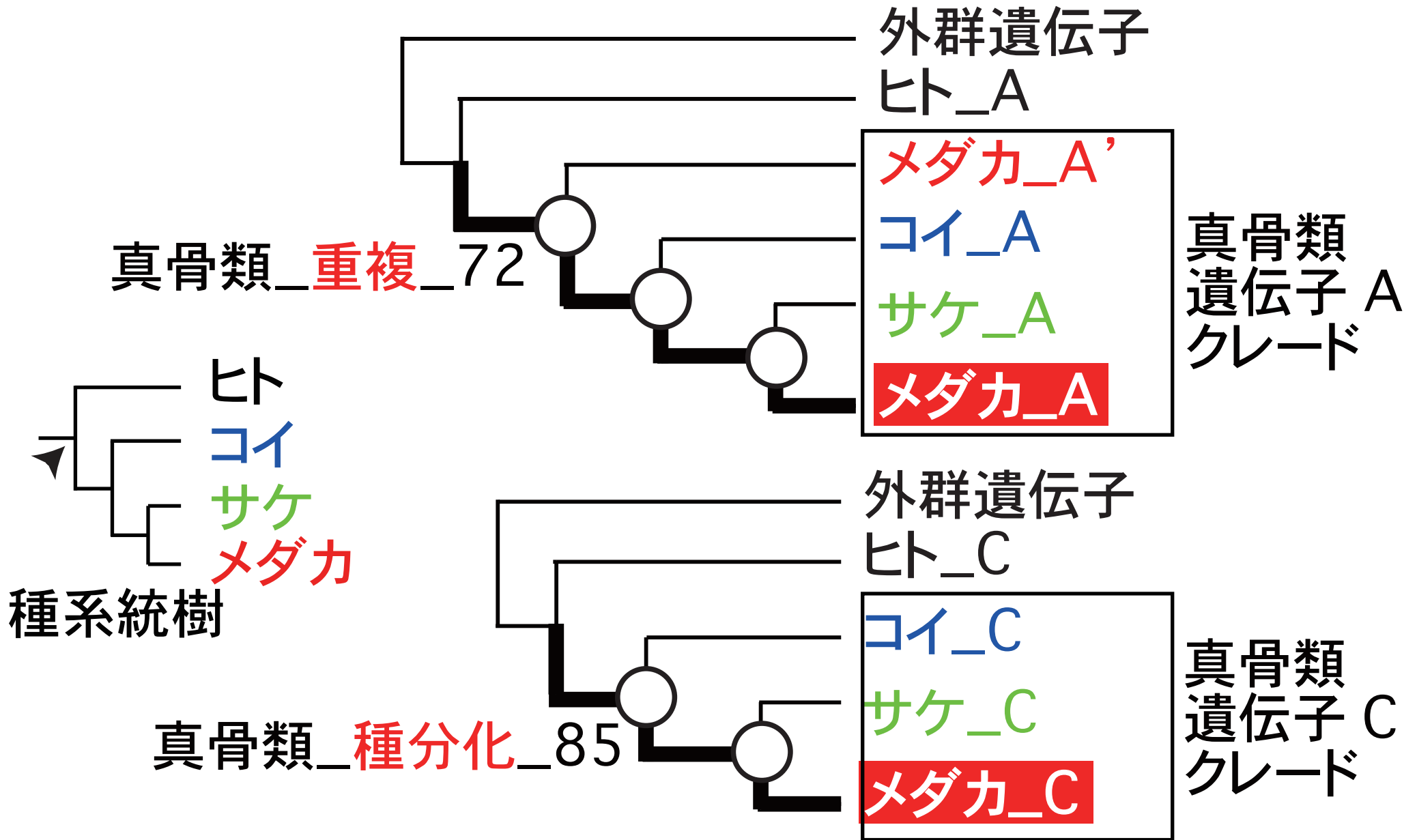
264 遺伝子

多くの遺伝子が再現に成功

## 2. 全ゲノム重複の検出

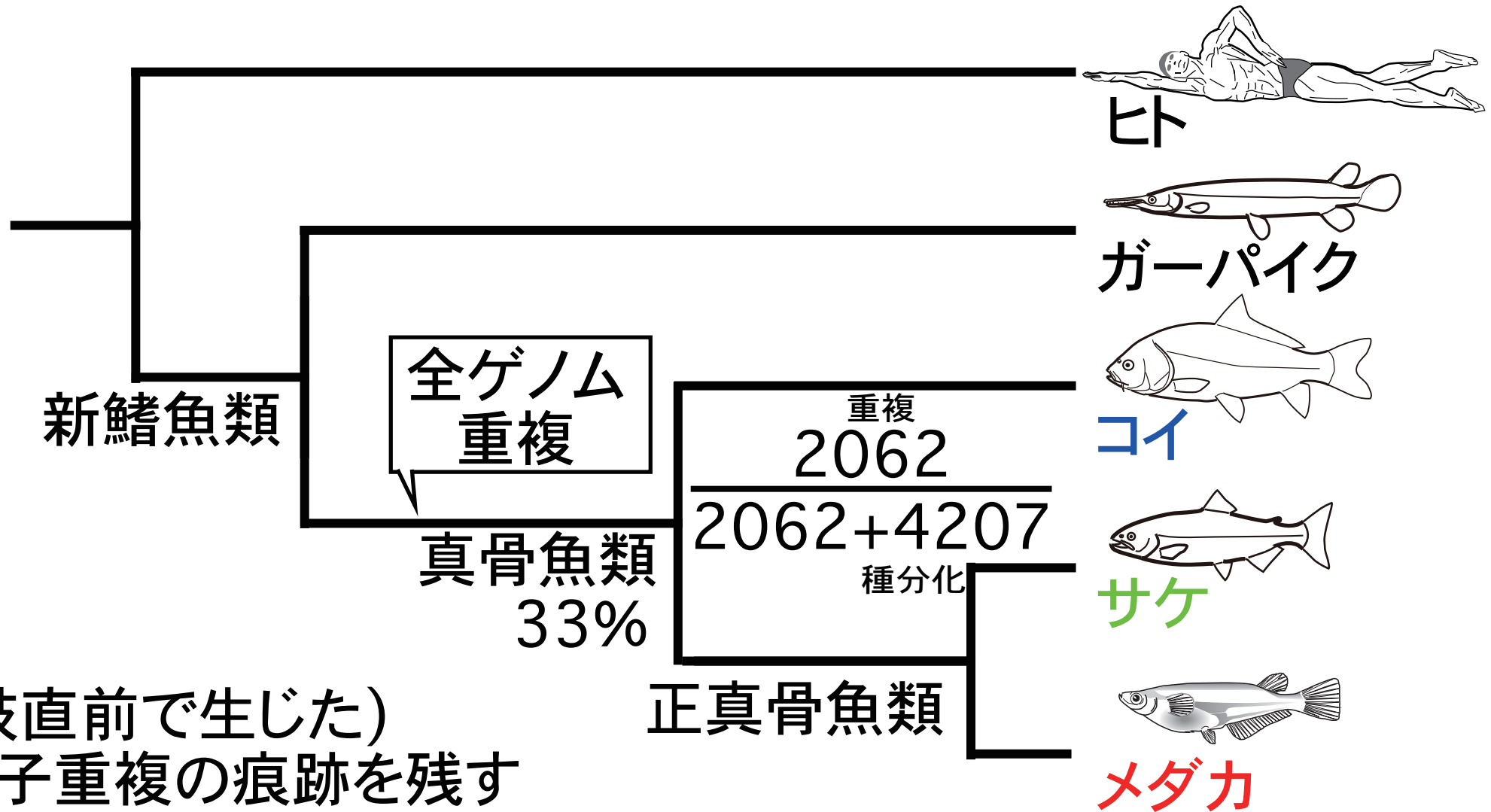


# 2. 全ゲノム重複の検出



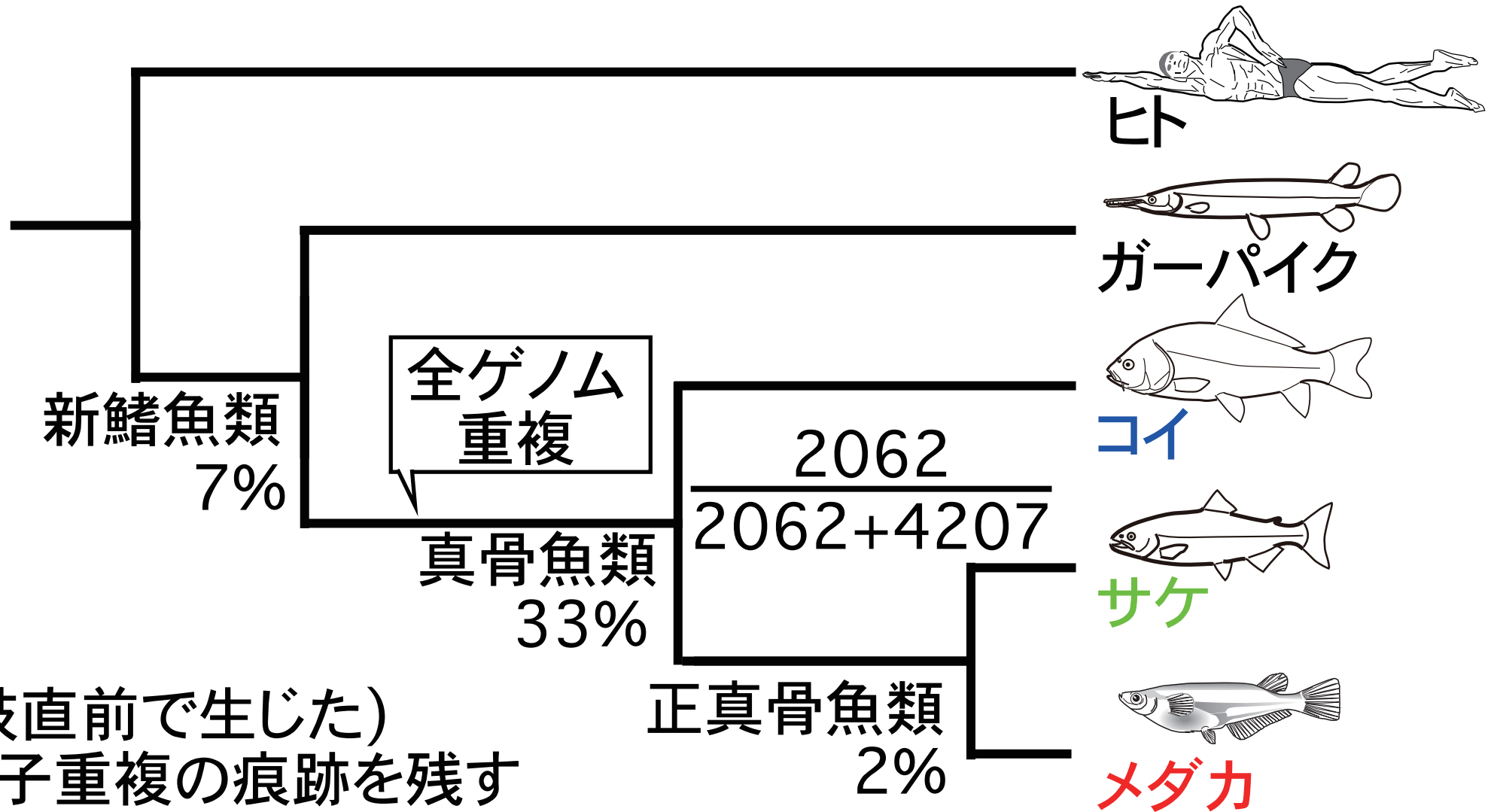
重複か種分化どちらか

# 2. 全ゲノム重複の検出



(分岐直前で生じた)  
遺伝子重複の痕跡を残す  
遺伝子の割合

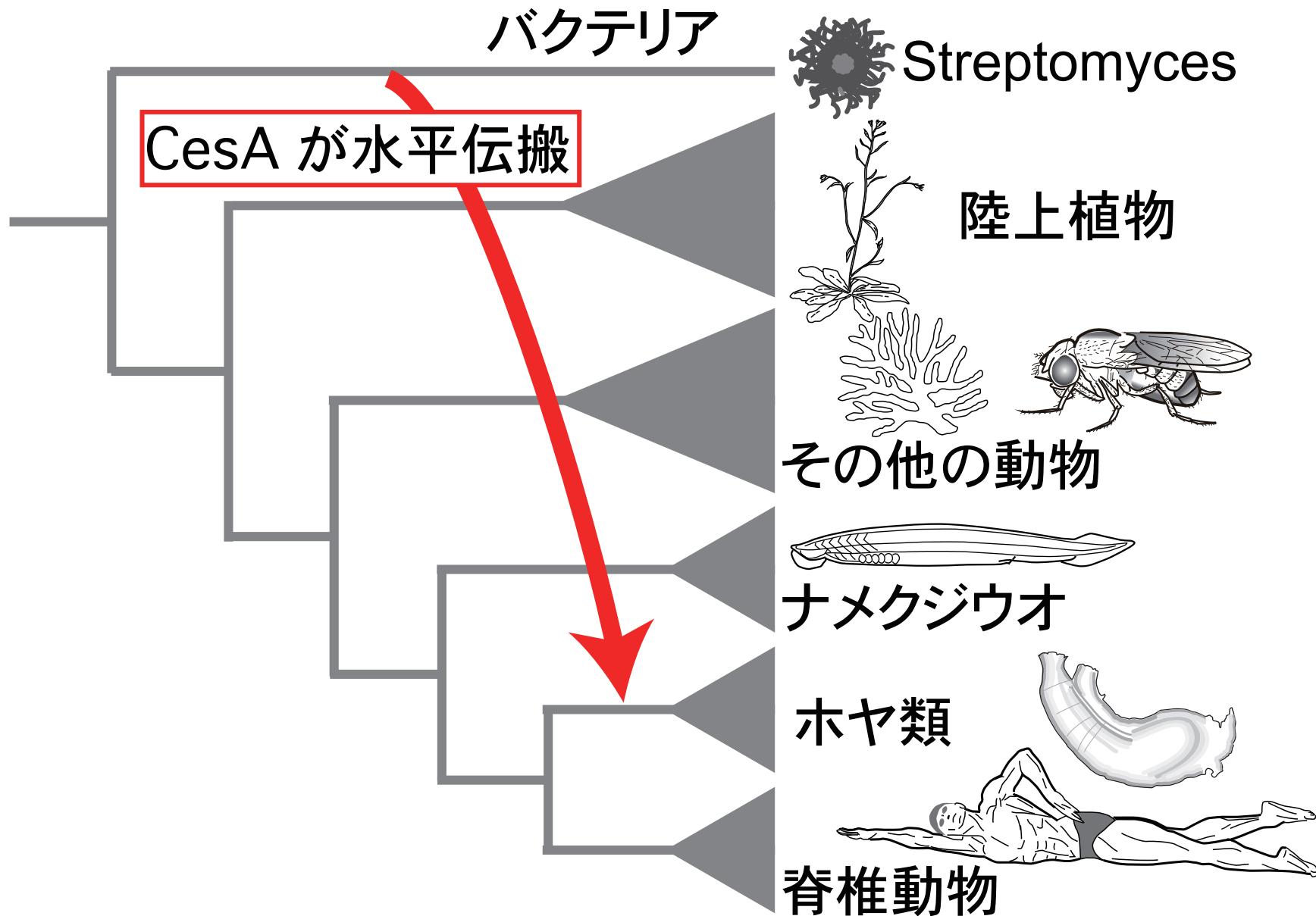
## 2. 全ゲノム重複の検出



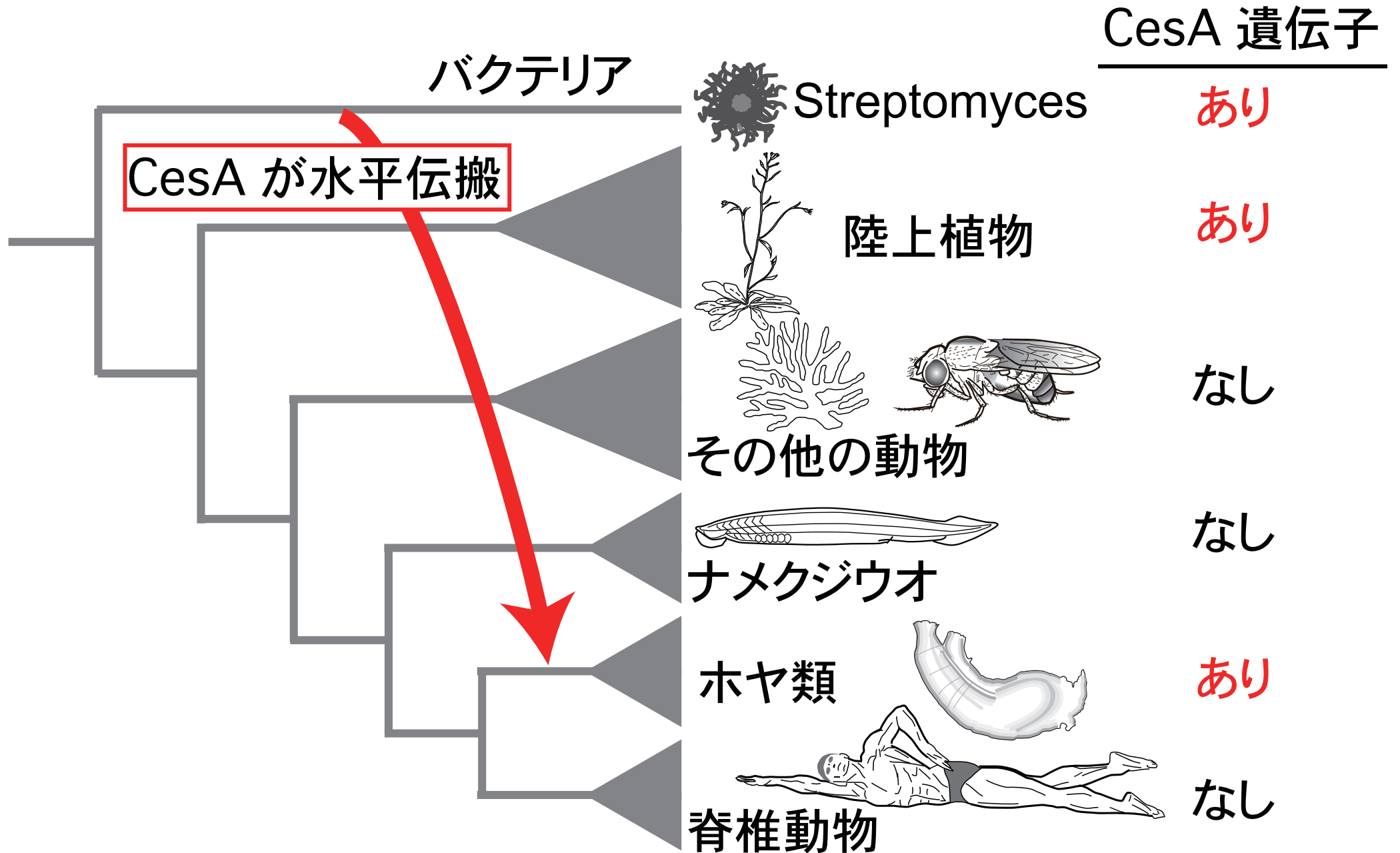
(分岐直前で生じた)  
遺伝子重複の痕跡を残す  
遺伝子の割合

検出に成功。祖先ゲノムの状態を比較。

# 3. 水平伝播の検出

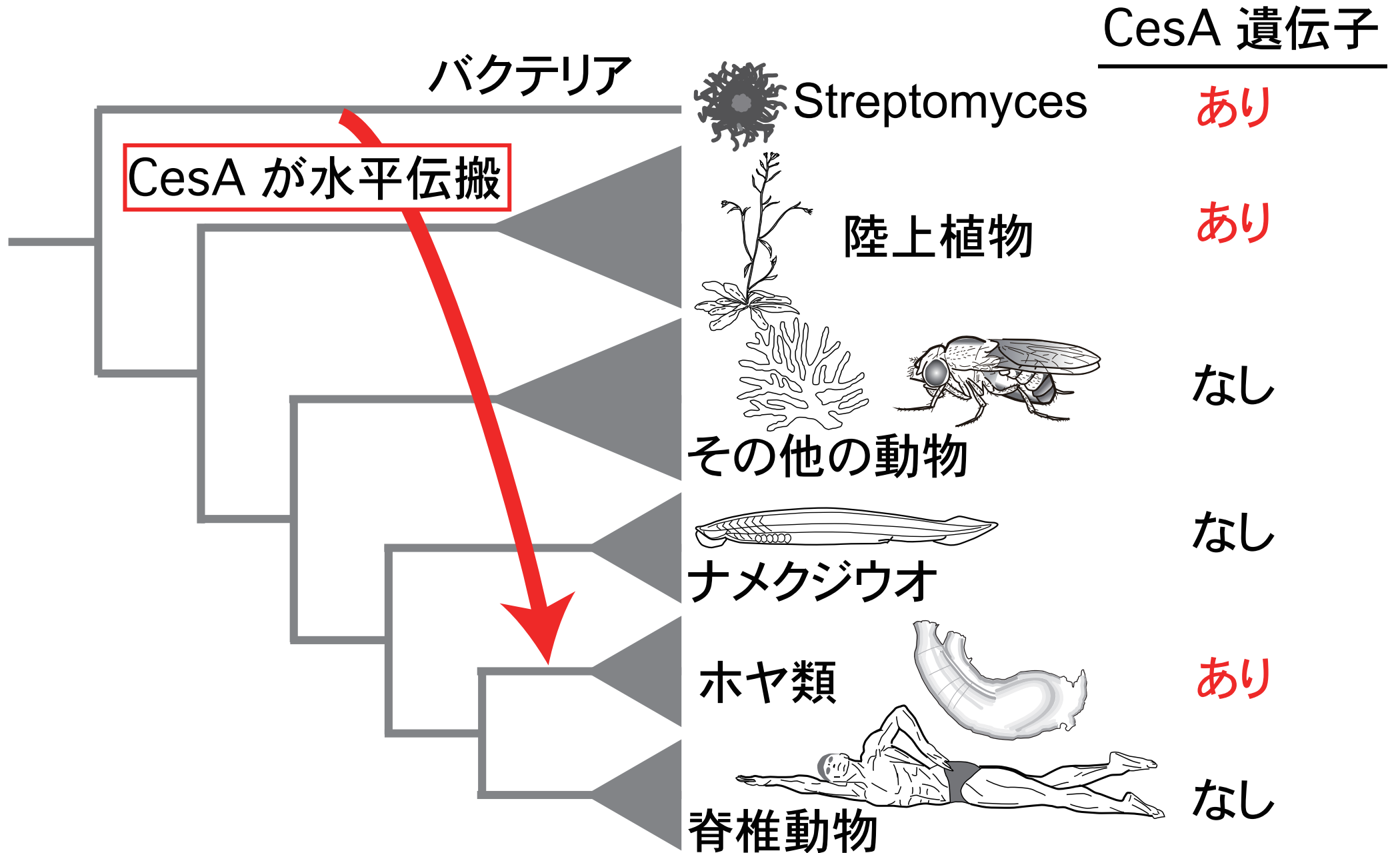


# 3. 水平伝播の検出





# 3. 水平伝播の検出



検出に成功。しかも、CesA 以外は伝搬していない。

# ORTHOSCOPE\* のまとめ

## 特徴

- ・ゲノムの歴史（祖先ゲノムの状態も）を推定できる。
- ・Web バージョンと併用できる。

# ORTHOSCOPE\* のまとめ

## 特徴

- ・ゲノムの歴史（祖先ゲノムの状態も）を推定できる。
- ・Webバージョンと併用できる。

## 問題点

- ・コマンドラインを使う程度の技術が必要。

Why GitHub? Enterprise Explore Marketplace Pricing Search Sign in Sign up

jun-inoue / orthoscope Watch 0 Star 3 Fork 2

Code Issues Pull requests Projects Wiki Security Insights

Join GitHub today  
Dismiss  
GitHub is home to over 36 million developers working together to host and review code, manage projects, and build software together.  
Sign up

An automatic web tool for phylogenetic...

640 commits

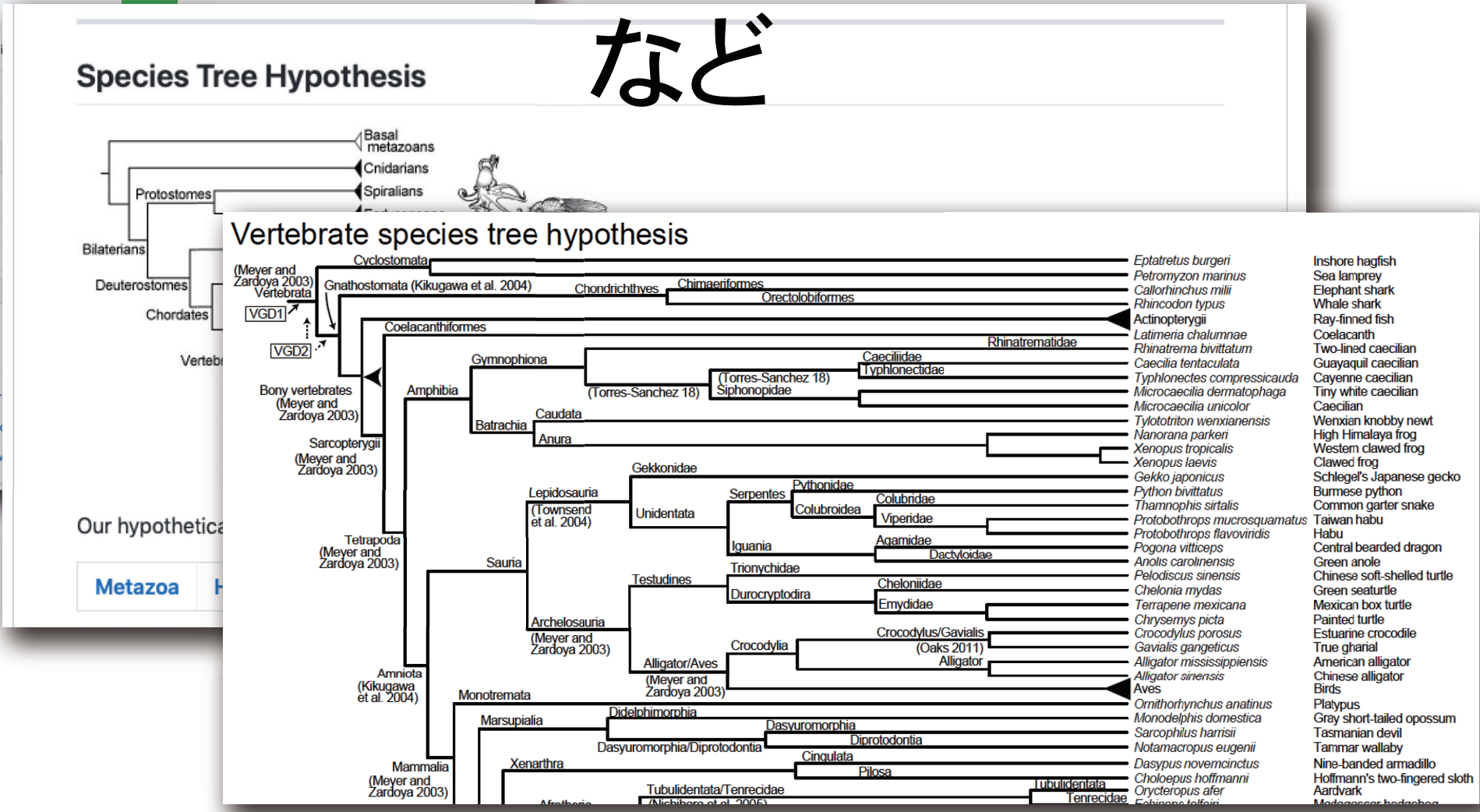
Branch: master New pull request

- jun-inoue a
- images
- tarfiles
- .DS\_Store
- README.md

### ORTHOSCOPE

Web service: <https://orthoscope.>  
Jan. 2019).  
Mirror site: <http://www.fish-evol.>  
Japanese instruction: <http://www.>

# ORTHOSCOPE, dbCNS, 種の系統樹 など



# ゲノムレベルで比較したい

ヒトゲノム

コード領域 (2%)

非コード領域 (98%)

## 1 遺伝子（配列）の解析

**ORTHOSCOPE**

Inoue & Satoh (2019)

**dbCNS**

Inoue & Saitou (2021)

## ゲノムレベルの解析

**ORTHOSCOPE\***

Inoue (2022)

—