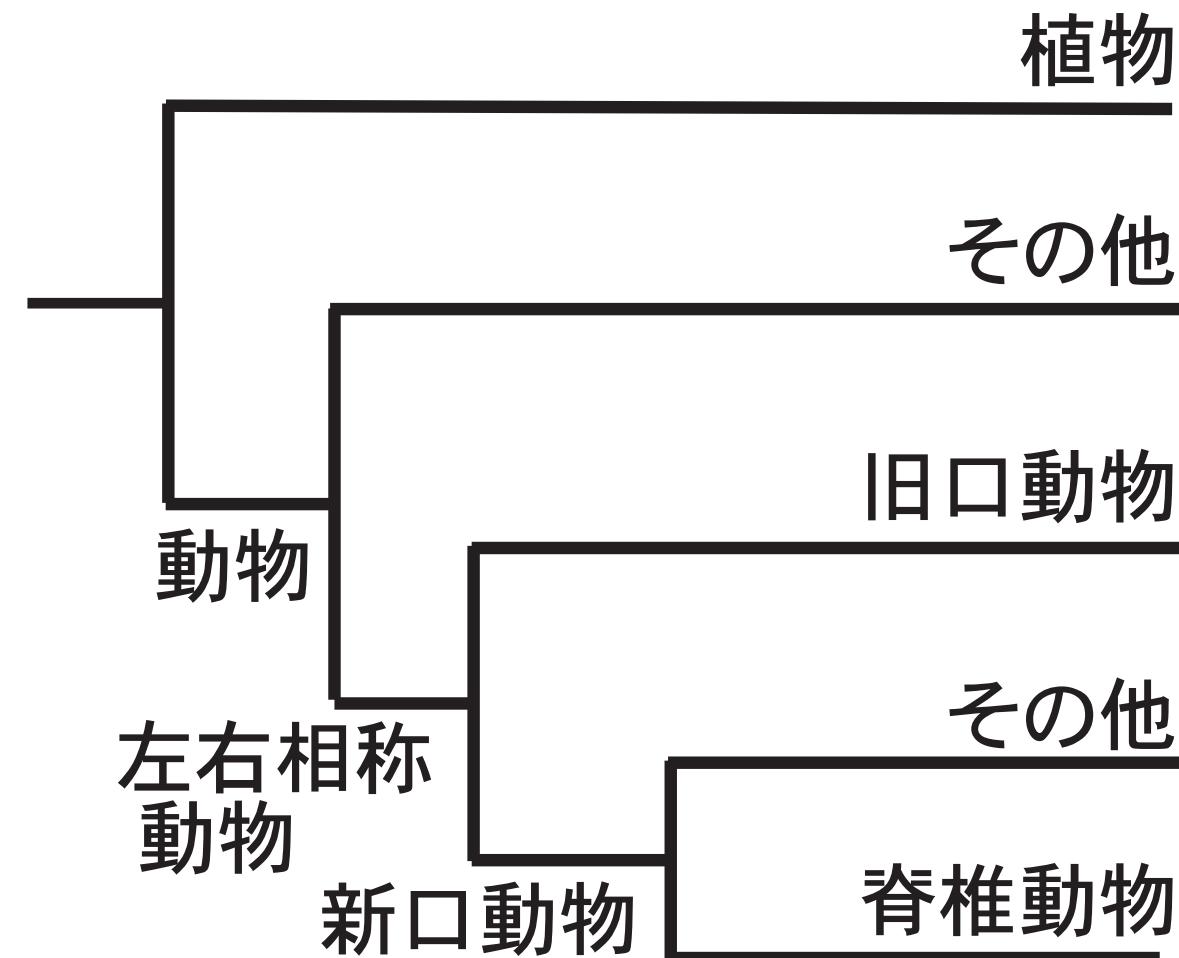


ORTHOSCOPE*： ゲノムに存在する全遺伝子の 歴史を推定する系統解析パイプライン

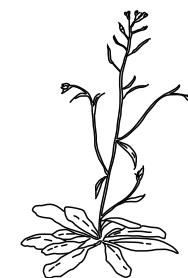
井上 潤

2022年9月8日
動物学会

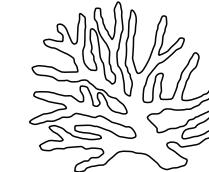
ゲノムデータが蓄積している



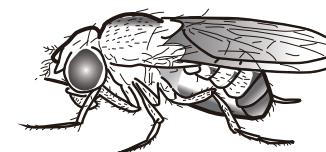
核ゲノムデータ(種)



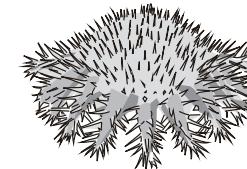
152



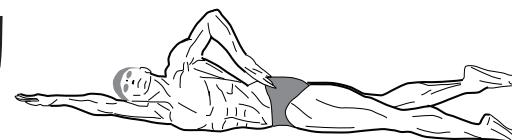
13



269



11



470

(RefSeq 2022年7月)

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

ORTHO SCOPE 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, [ORTHO SCOPE*](#), has bee 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): ファイルを選択 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
 3 5 10 20

Number of BLAST hits to report per genome 0.0 0.2 0.4 0.55

Alignment

Aligned site rate threshold within unambiguously aligned sites Amino acid DNA (Exclude 3rd) DNA (Include 3rd)
 60% 70% 80%

Tree search

Dataset Nephrozoa or Bilateria

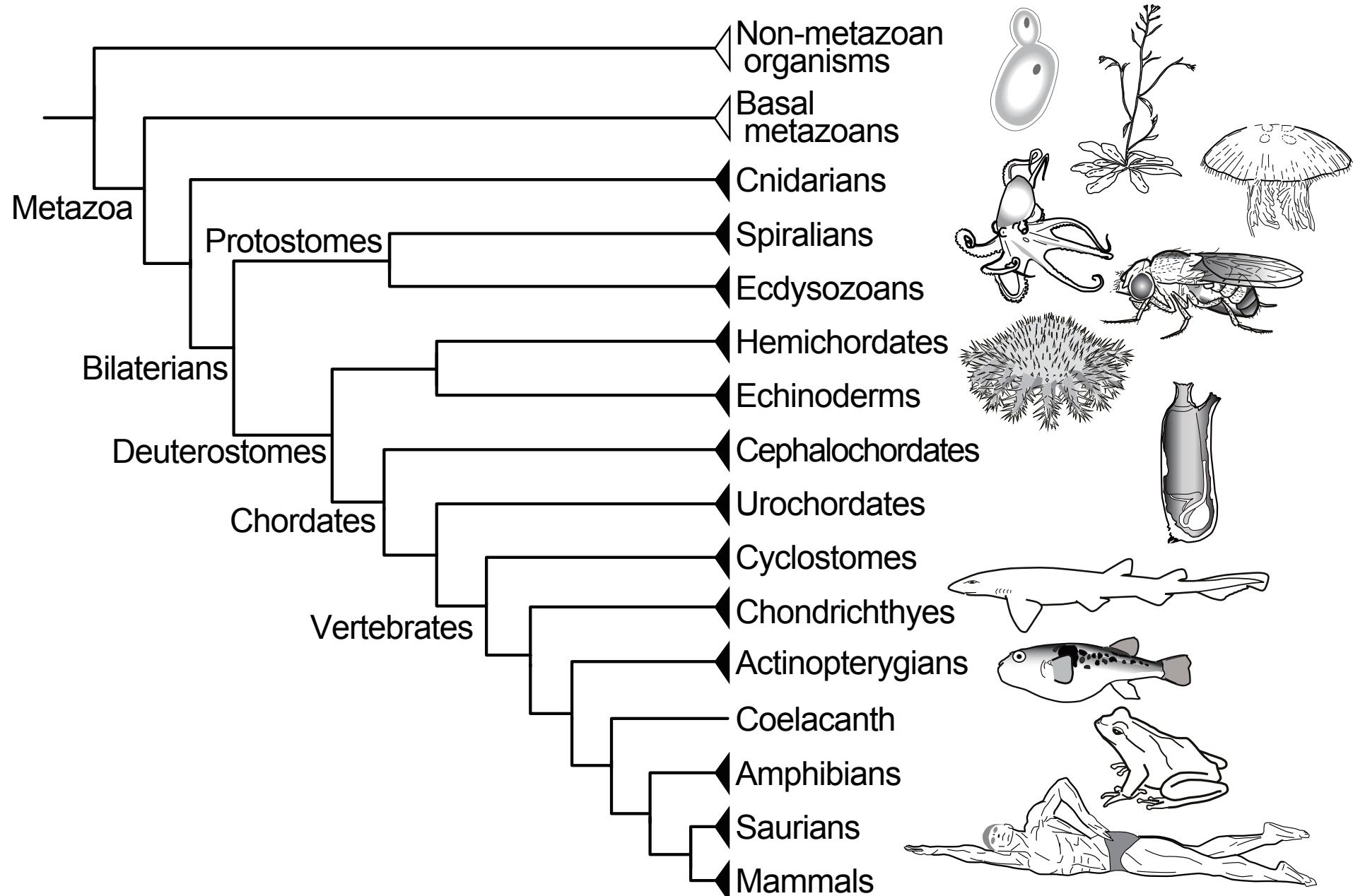
Rearrangement BS value threshold

Key node

BLAST
+
遺伝子系統樹

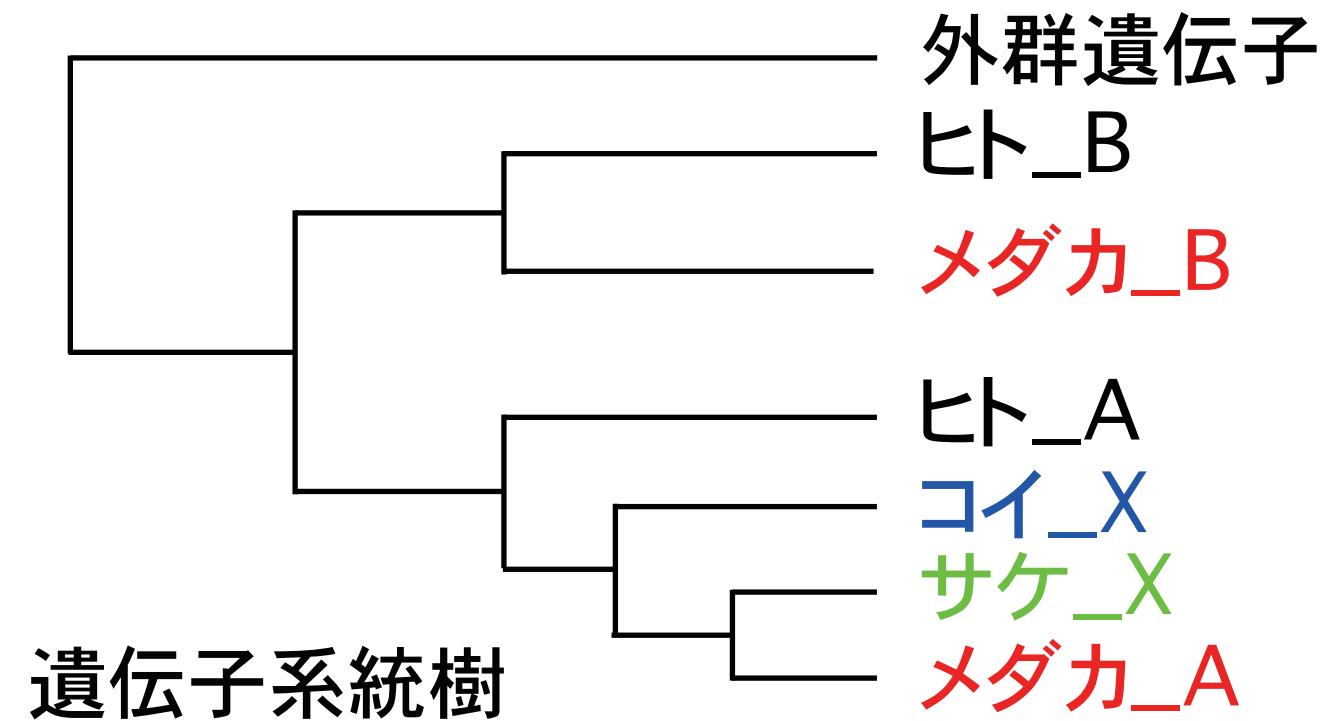
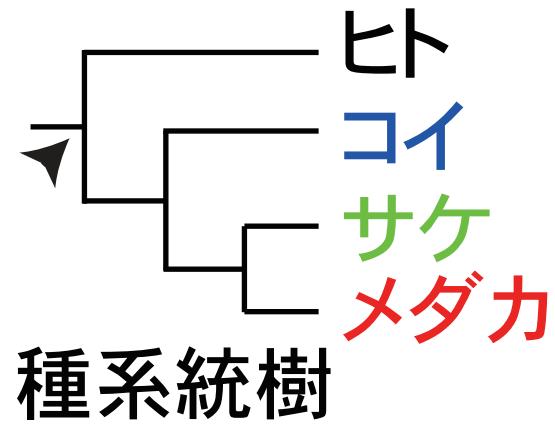
Inoue and Satoh (2019) MBE

ORTHOSCOPE のデータベース

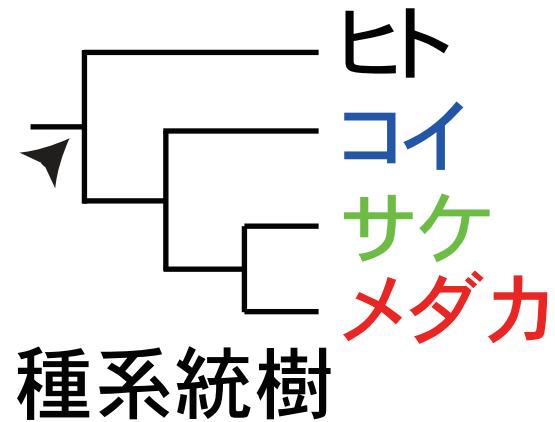


動物 550 種、植物 50 種の遺伝子モデル

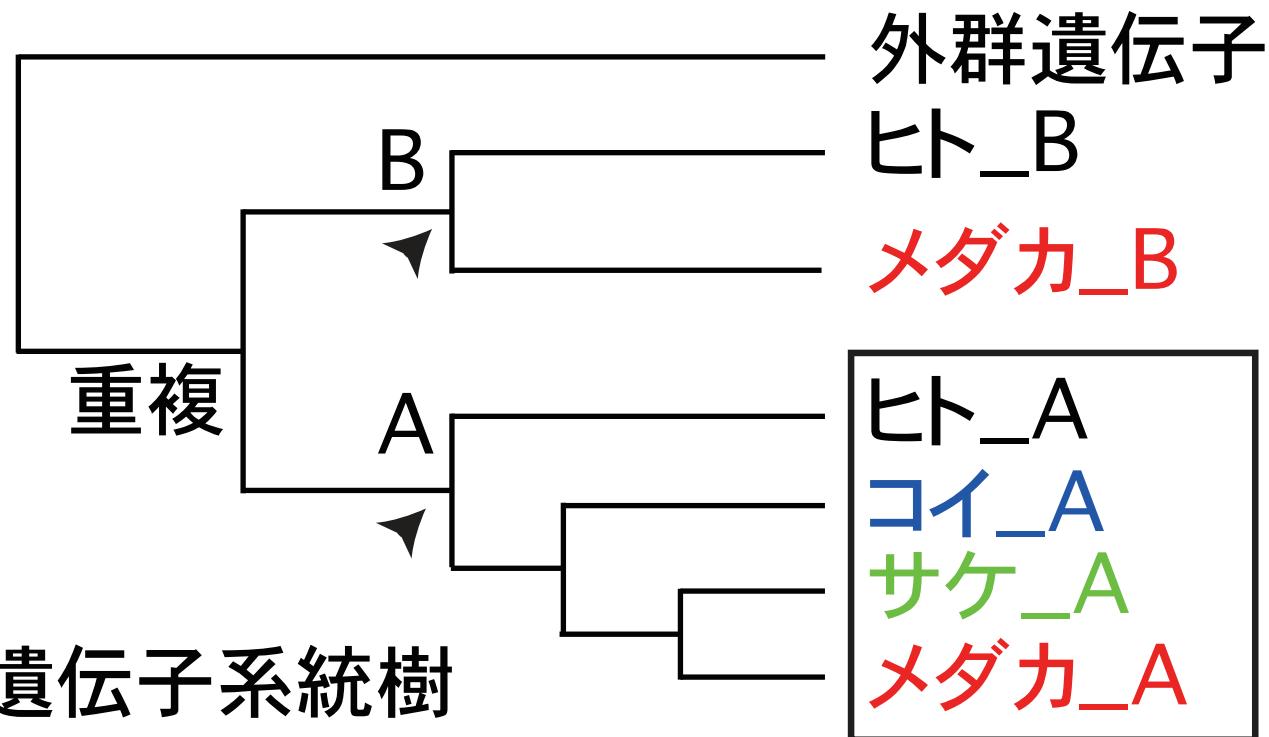
ORTHO SCOPE は何ができる？



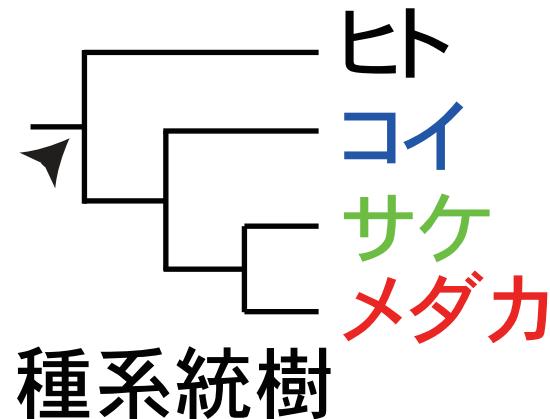
ORTHO SCOPE は何ができる？



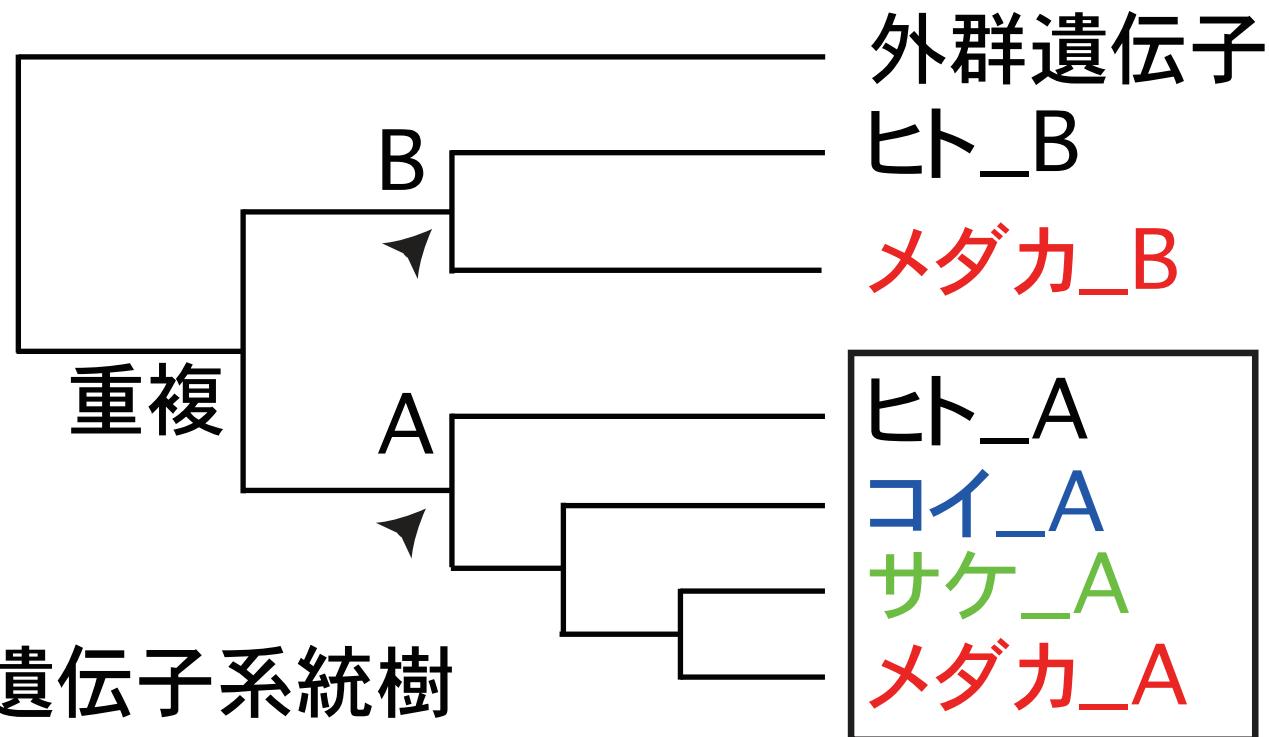
オーソグループを見つけて，



ORTHO SCOPE 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 bee 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): ファイルを選択 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

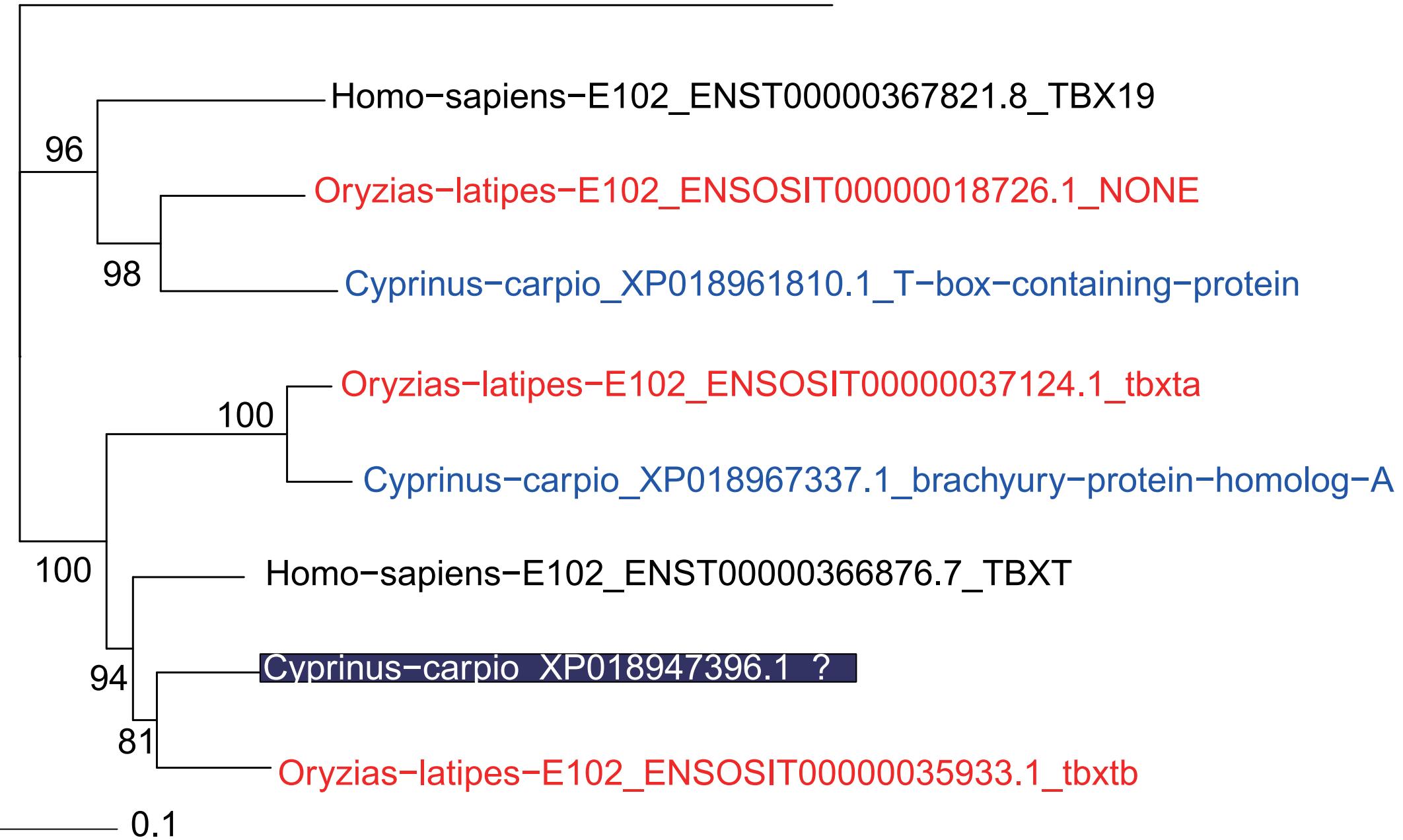
Nephrozoa or Bilateria

Deuterostomia

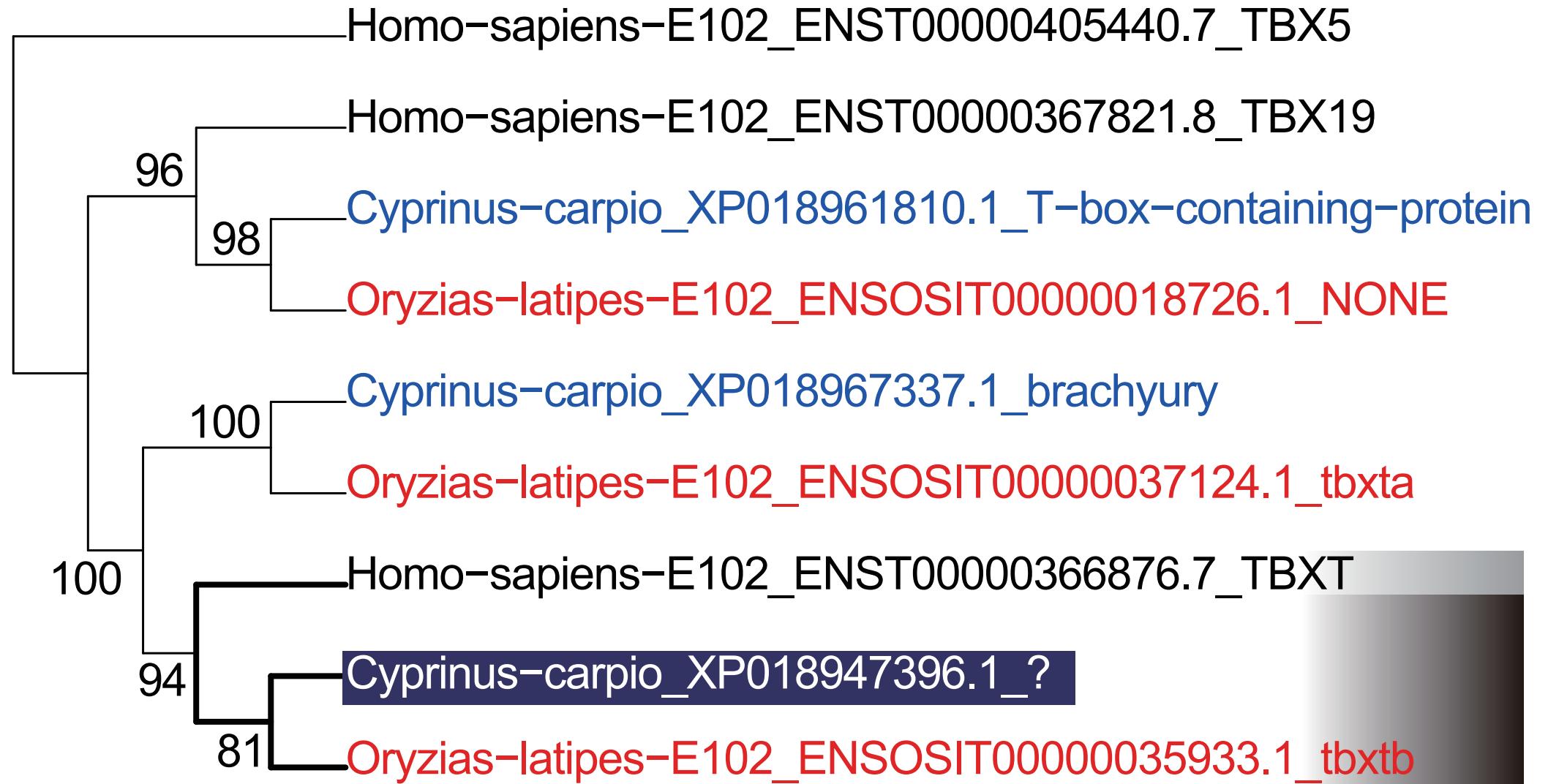
Hemichordata	<input checked="" type="checkbox"/> <i>Saccoglossus kowalevskii</i>	Helical acorn worm	OIST-R	34239
	<input type="checkbox"/> <i>Saccoglossus kowalevskii</i> 1	Helical acorn worm	RefSeq89	20922
Echinodermata	<input checked="" type="checkbox"/> <i>Ptychoderma flava</i>	Yellow acorn worm	OIST-S	34647
	<input checked="" type="checkbox"/> <i>Strongylocentrotus purpuratus</i>	Purple urchin	EnsMet38	28842
	<input type="checkbox"/> <i>Strongylocentrotus purpuratus</i> 1	Purple urchin	RefSeq89	27728
	<input checked="" type="checkbox"/> <i>Acanthaster planci</i>	Crown-of-thorns starfish (Okinawa)	OIST-S	24323
	<input type="checkbox"/> <i>Acanthaster planci</i> 1	Crown-of-thorns starfish (Okinawa)	RefSeq89	16468
	<input type="checkbox"/> <i>Acanthaster planci GBR</i>	Crown-of-thorns starfish (Great Barrier Reef)	OIST-S	24747
Cephalochordata	<input type="checkbox"/> <i>Asymmetron lucayanum</i>	Lancelet	TSA	104311
	<input checked="" type="checkbox"/> <i>Branchiostoma belcheri</i>	Belcher's lancelet	LanceletDB	30392
	<input type="checkbox"/> <i>Branchiostoma belcheri</i> 1	Belcher's lancelet	RefSeq89	23855
	<input type="checkbox"/> <i>Branchiostoma lanceolatum</i>	Mediterranean amphioxus	Amphiencode	23586
	<input checked="" type="checkbox"/> <i>Branchiostoma floridae</i>	Florida lancelet	JGI	21954
	<input type="checkbox"/> <i>Branchiostoma floridae</i> 1	Florida lancelet	RefSeq89	28621
Urochordata				
Appendicularia	<input checked="" type="checkbox"/> <i>Oikopleura dioica</i>	Pelagic tunicate	OikoBase	17217
Thaliacea	<input type="checkbox"/> <i>Salpa thompsoni</i>	Salpa	TSA	76798
Enterogona	<input checked="" type="checkbox"/> <i>Ciona savignyi</i>	Pacific transparent sea squirt	Ens91	11616
	<input type="checkbox"/> <i>Ciona savignyi</i> 1	Pacific transparent sea squirt	aNISEED	20115
	<input checked="" type="checkbox"/> <i>Ciona intestinalis</i>	Transparent sea squirt	Ens91	16671
	<input type="checkbox"/> <i>Ciona intestinalis</i> 1	Transparent sea squirt	RefSeq89	13520
Molgulidae	<input type="checkbox"/> <i>Molgula occidentalis</i>	—	aNISEED	33023
	<input type="checkbox"/> <i>Molgula oculata</i>	Sea grape	aNISEED	16616
Stvelidae	<input type="checkbox"/> <i>Botrylloides leachii</i>	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

result3300

- 010_candidates_nucl.txt
- 020_treeRearranged.pdf
- 020_tree.pdf
- 100_2ndTree.zip
- 000_summary.txt
- 010_candidates_prot.txt
- 010_a1n.aln.html

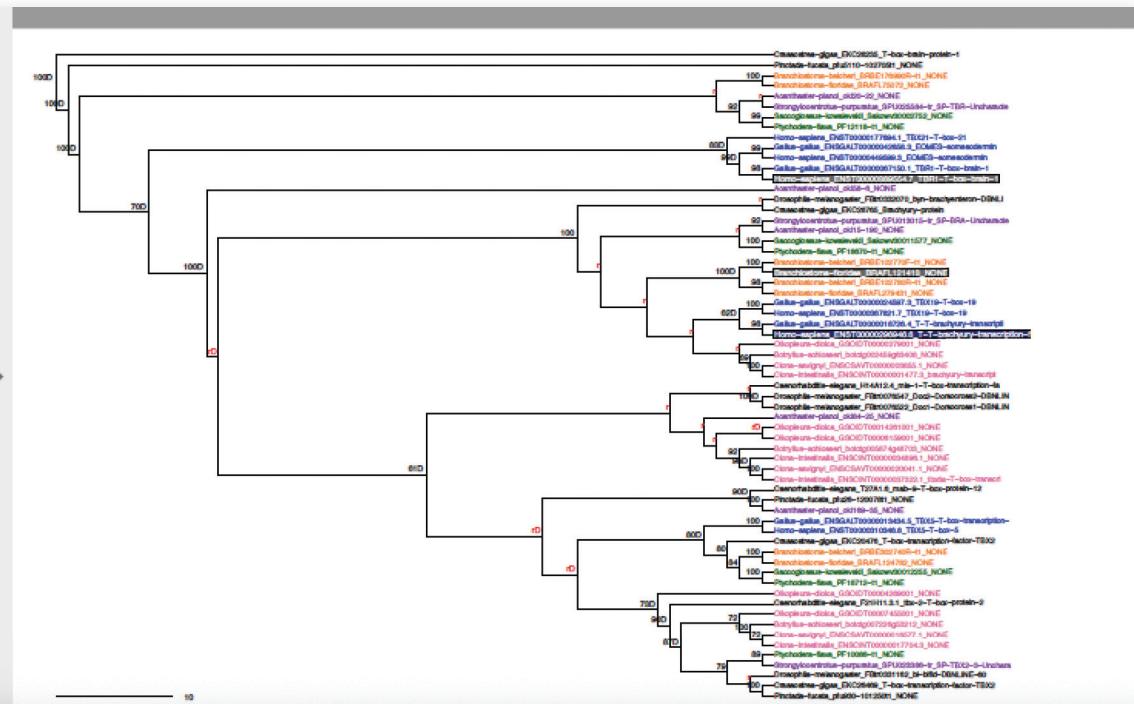
000_summary.txt

```
#####
# Results #####
#####

>Queries_used_in_the_analysis
Homo-sapiens_ENST00000296946.6_T-T-brachyury-transcrip
Branchiostoma-floridae_BRAFL121413_NONE
Homo-sapiens_ENST0000389554.7_TBR1-T-box-brain-1

>BootstrapValue_OrthogroupMonophyly
The orthogroup monophyly is supported by 100% bootstrap

>Orthogroup
Acanthaster-planci_oki15-190_NONE
Botryllus-schlosseri_botctg002459g63408_NONE
Branchiostoma-belcheri_BRBE102770F-t1_NONE
Branchiostoma-belcheri_BRBE102780R-t1_NONE
Branchiostoma-floridae_BRAFL121413_NONE
Branchiostoma-floridae_BRAFL279431_NONE
Ciona-intestinalis_ENSCINT00000001477.3_brachyury-tr
Ciona-savignyi_ENSCSAVT00000003855.1_NONE
Crassostrea-gigas_EKC28765_Brachyury-protein
Drosophila-melanogaster_FBtr0332070_byn-brachyentero
Gallus-gallus_ENSGALT00000018726.4_T-T-brachyury-tr
Gallus-gallus_ENSGALT00000024597.3_TBX19-T-box-19
Homo-sapiens_ENST00000296946.6_T-T-brachyury-transcrip
```



These results can be downloaded.

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*

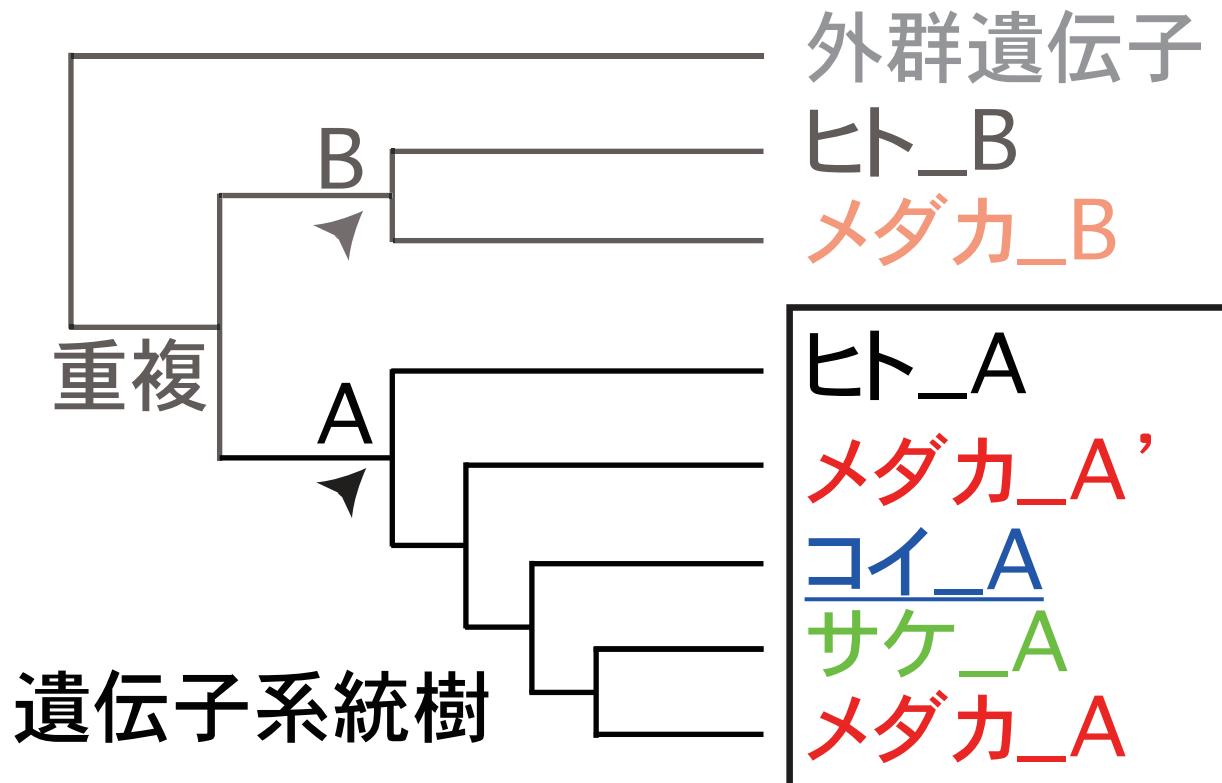
全遺伝子

スクリプト

Inoue (2022) MBE

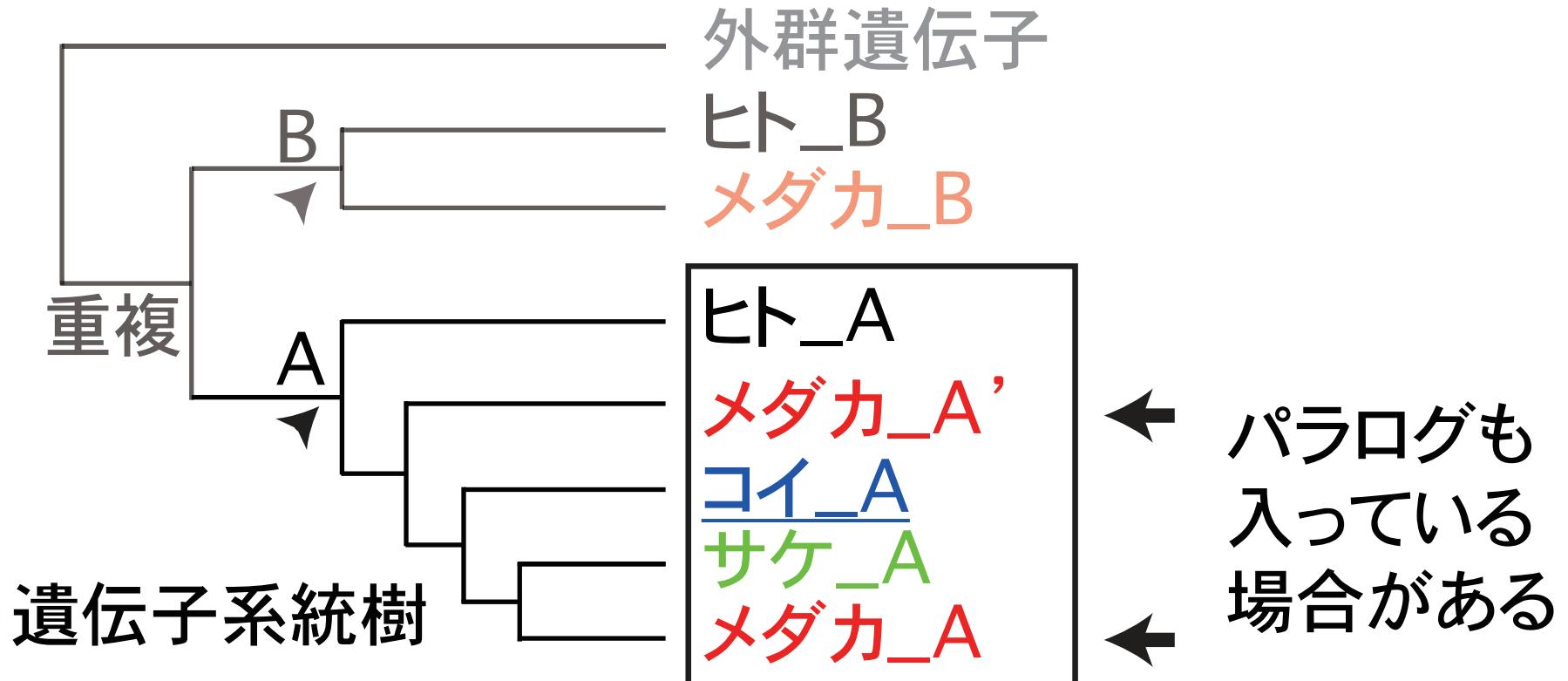
ORTHO SCOPE*

オーソグループだけの系統樹も推定



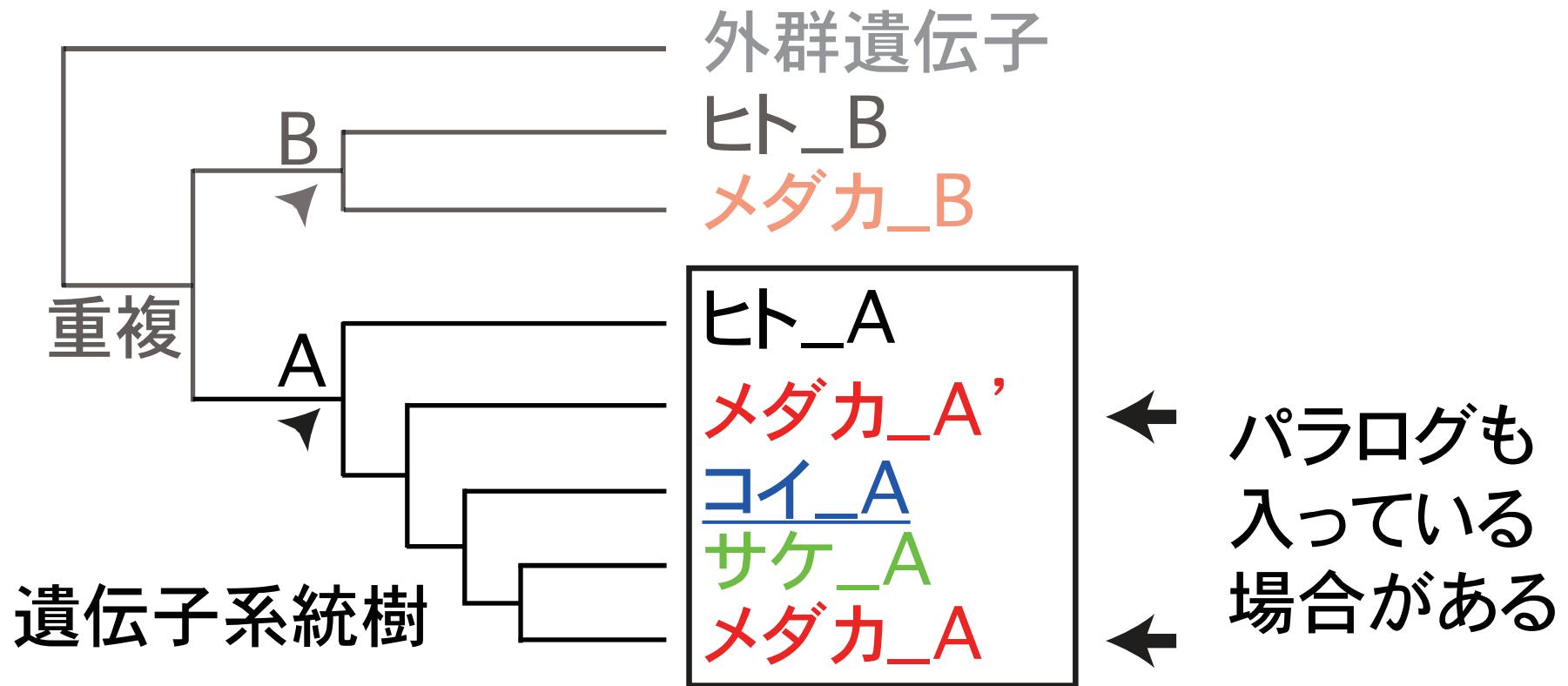
ORTHO SCOPE*

オーソグループだけの系統樹も推定



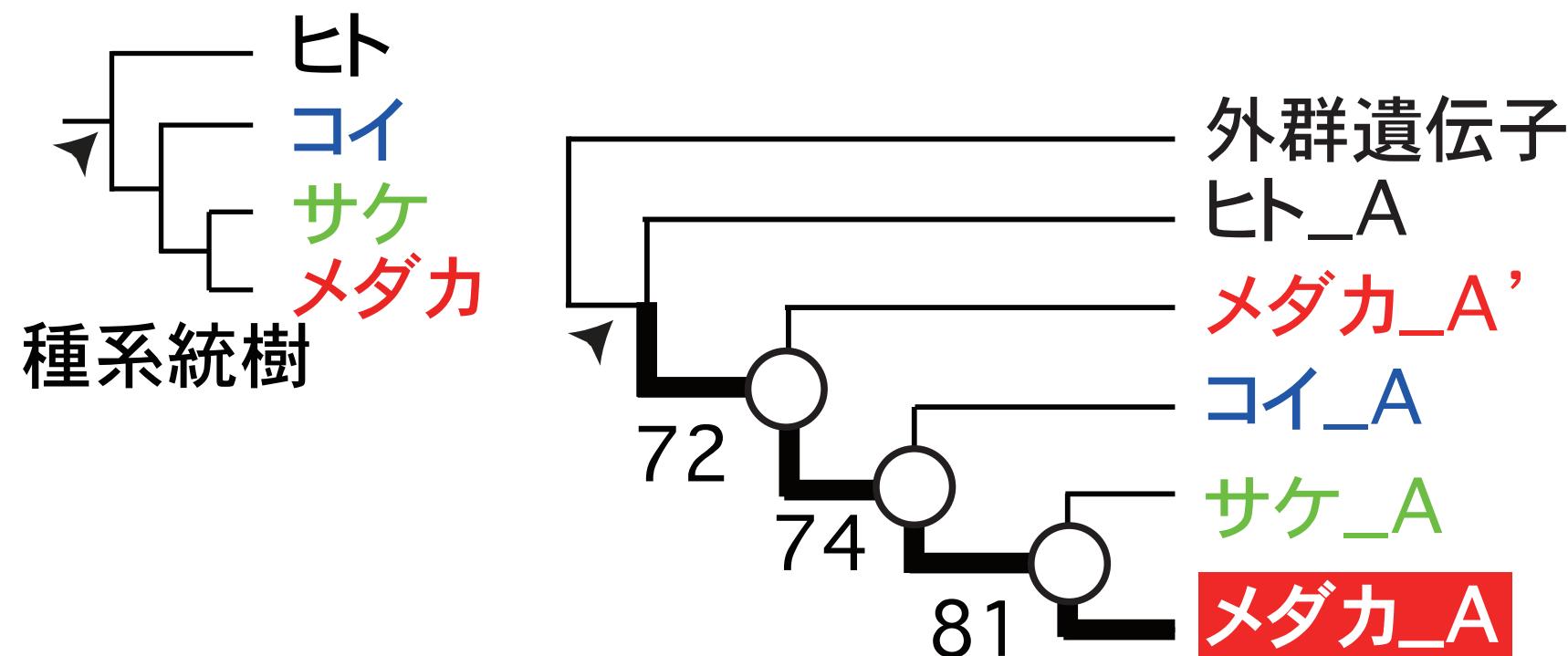
ORTHO SCOPE*

オーソグループだけの系統樹も推定

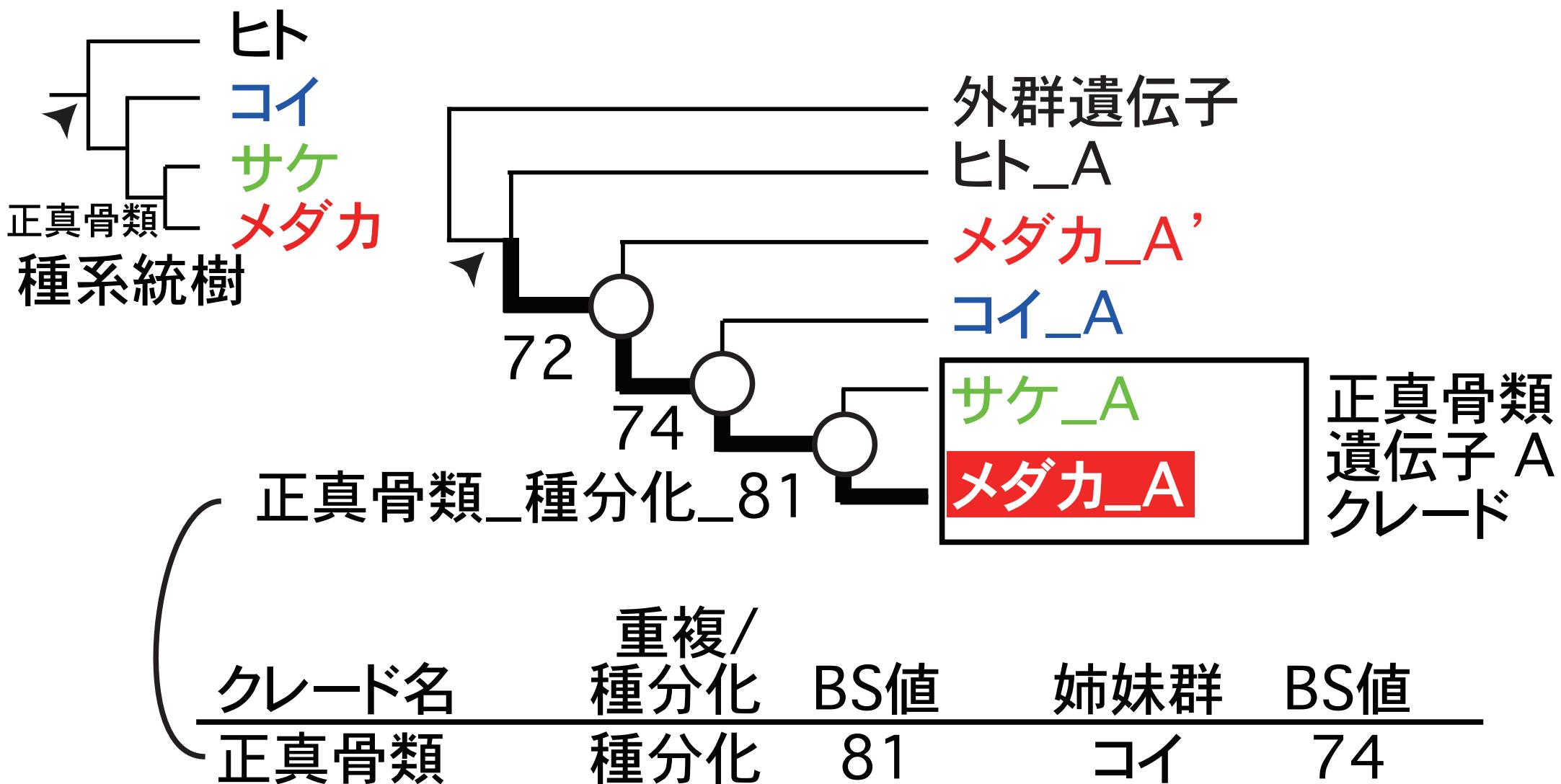


問題点: どれを Concatenate する?

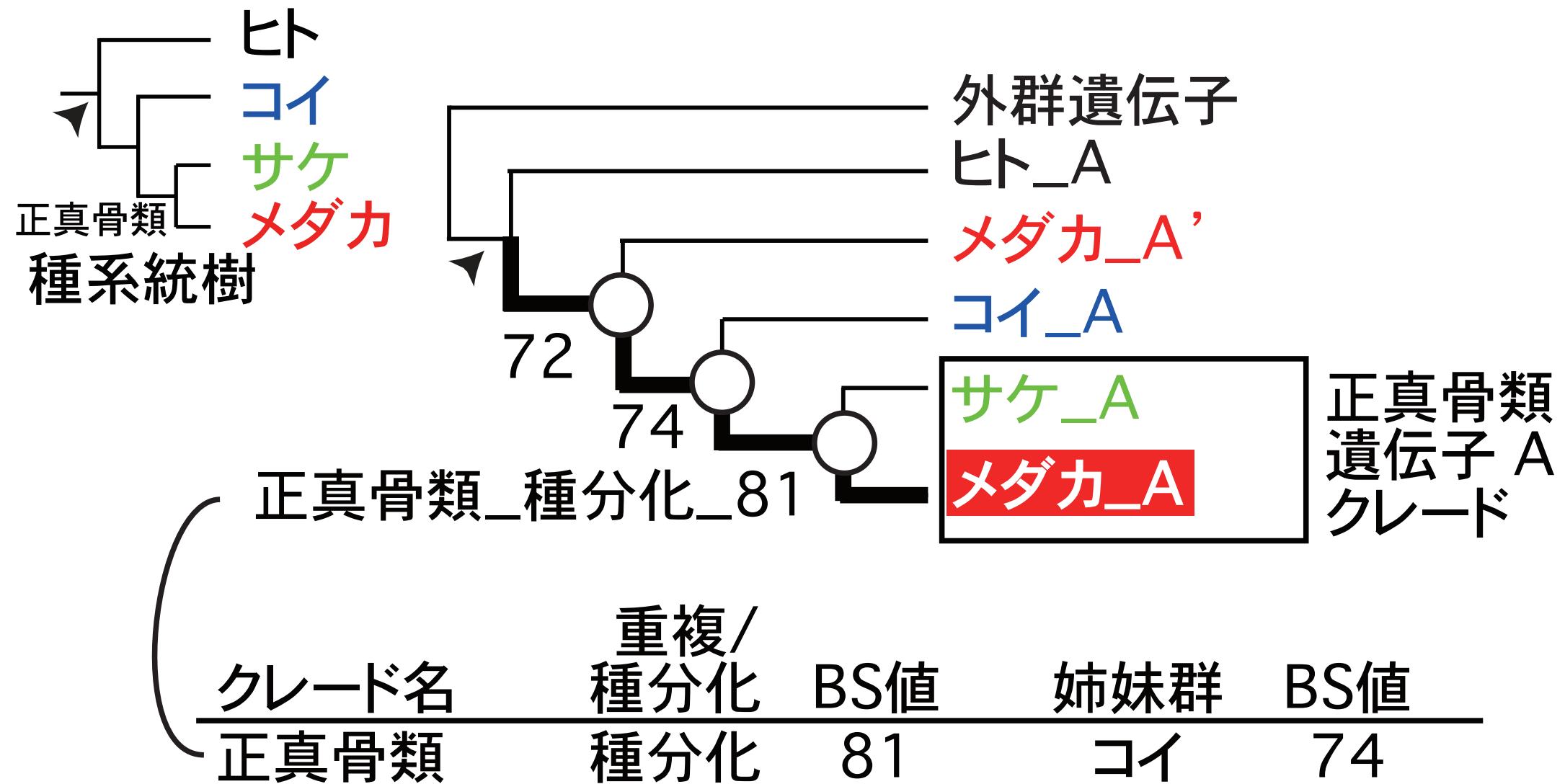
遺伝子ごとの歴史を推定



遺伝子ごとの歴史を推定

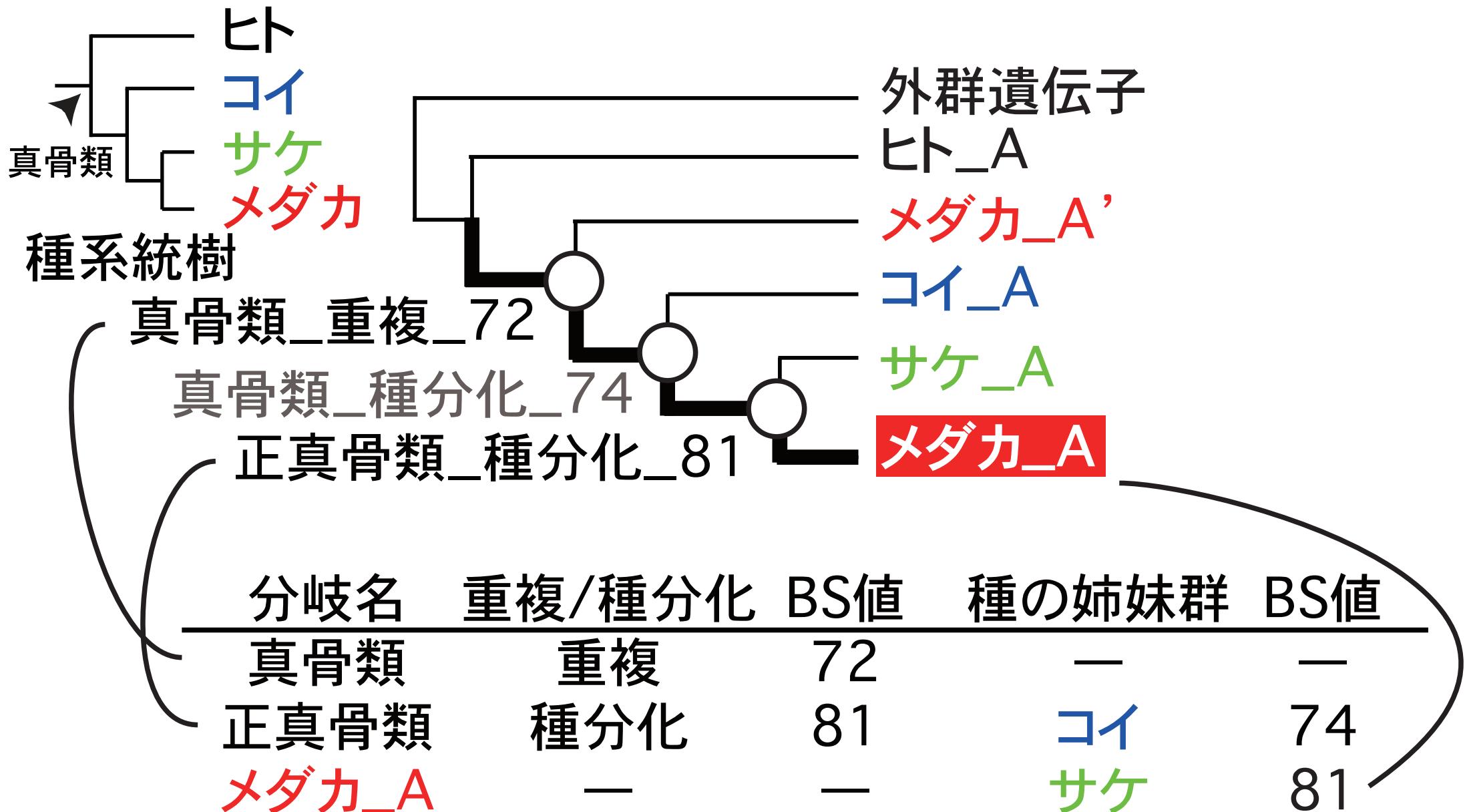


遺伝子ごとの歴史を推定

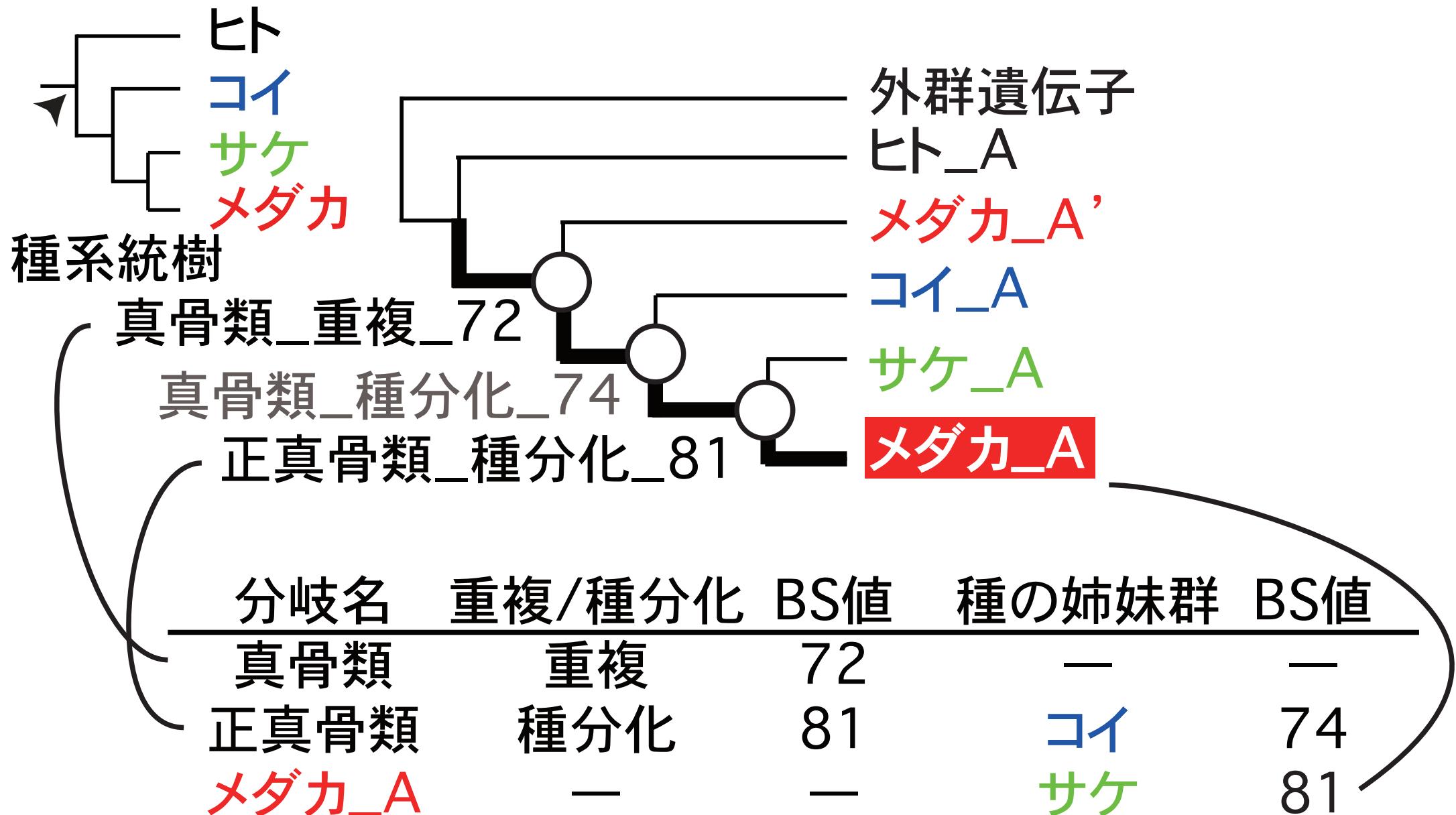


エクセルファイルとして出力

他のノードも記述

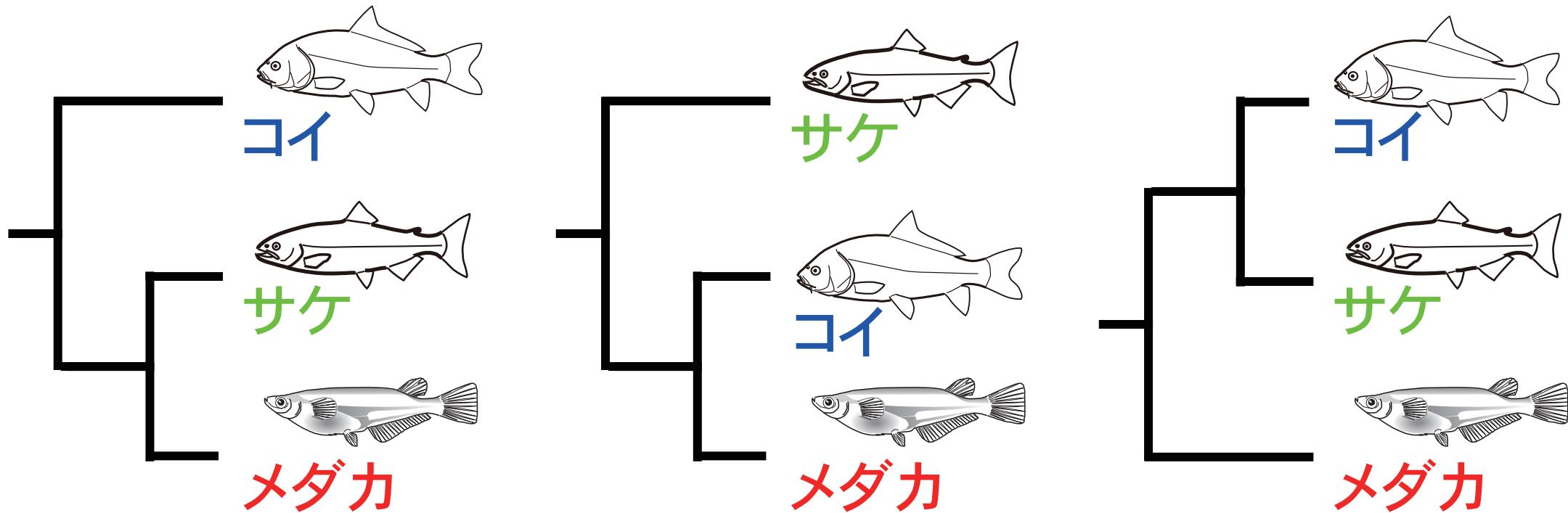


他のノードも記述



精度の高い情報だけで、ゲノムの歴史を推定

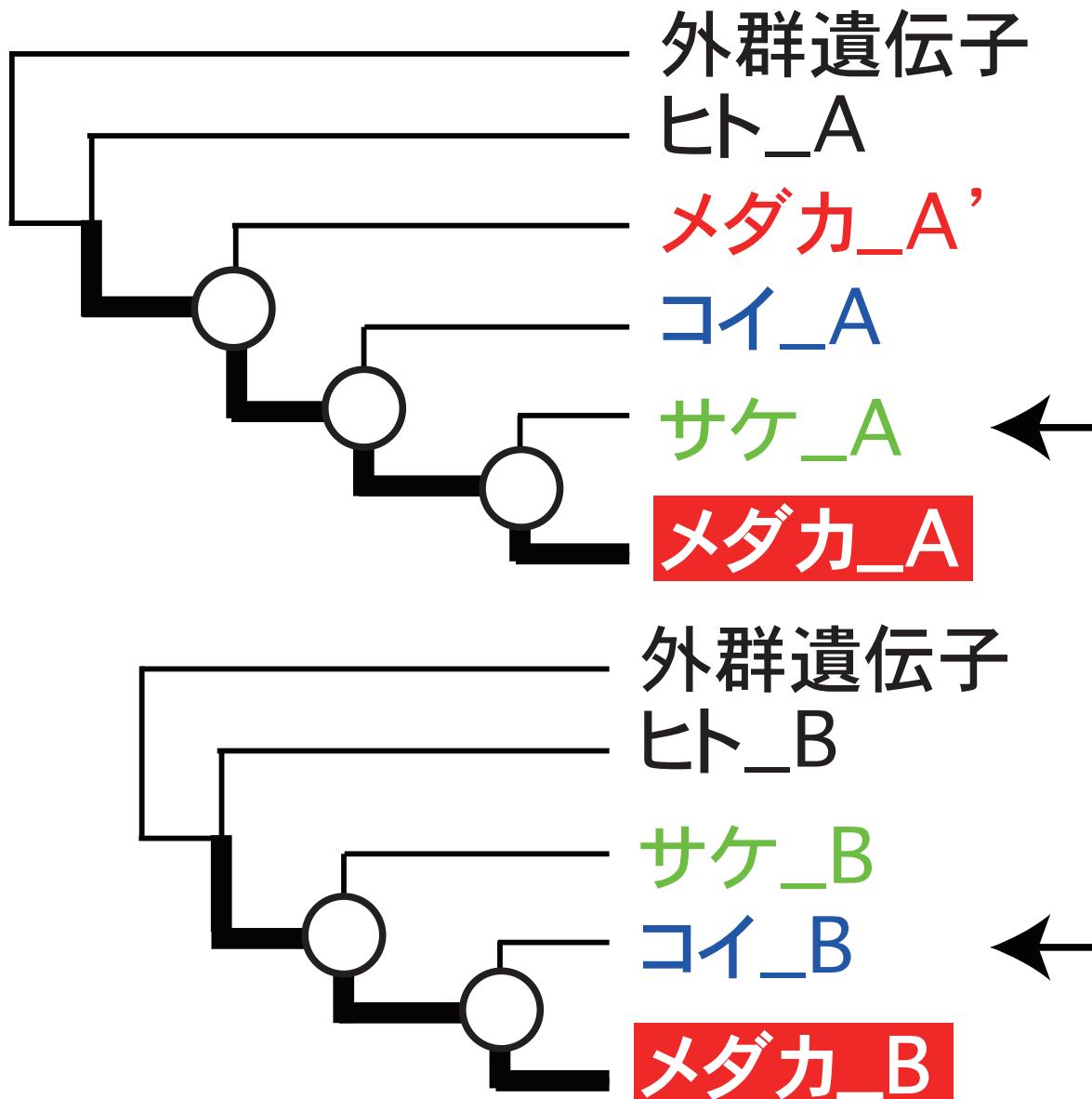
1. 系統仮説を比較



広く認められる関係

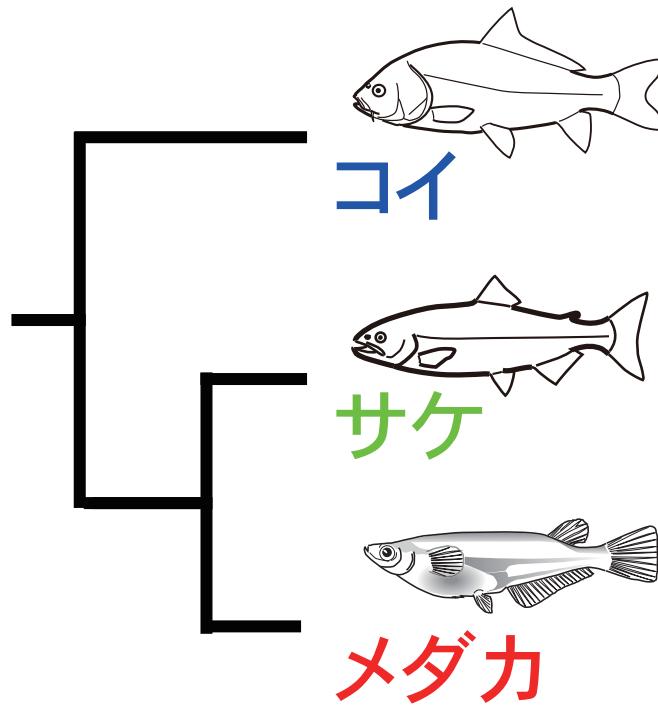
メダカの姉妹群は？

1. 系統仮説を比較

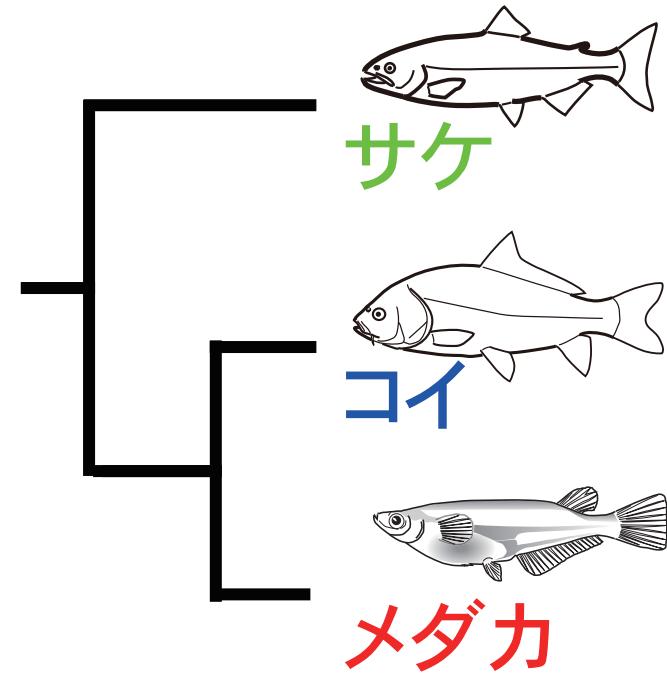


遺伝子ごとに姉妹群（種）を推定

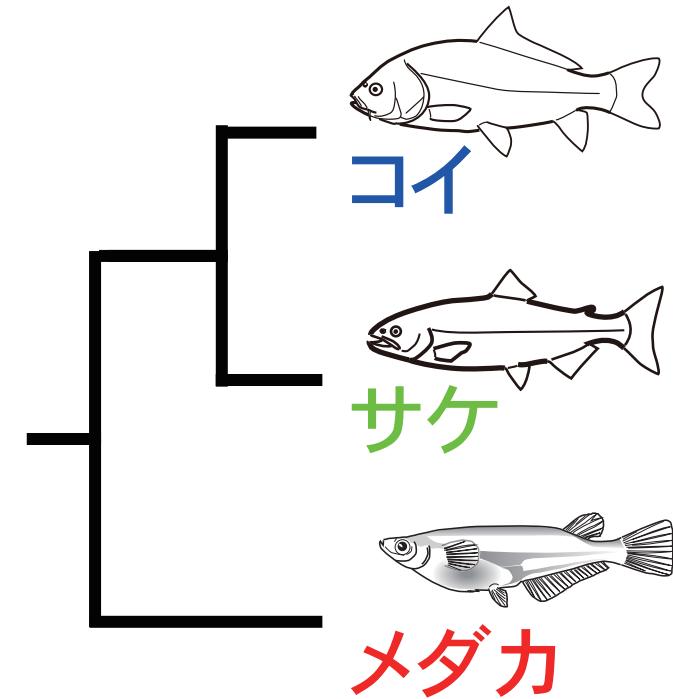
1. 系統仮説を比較



2520 遺伝子



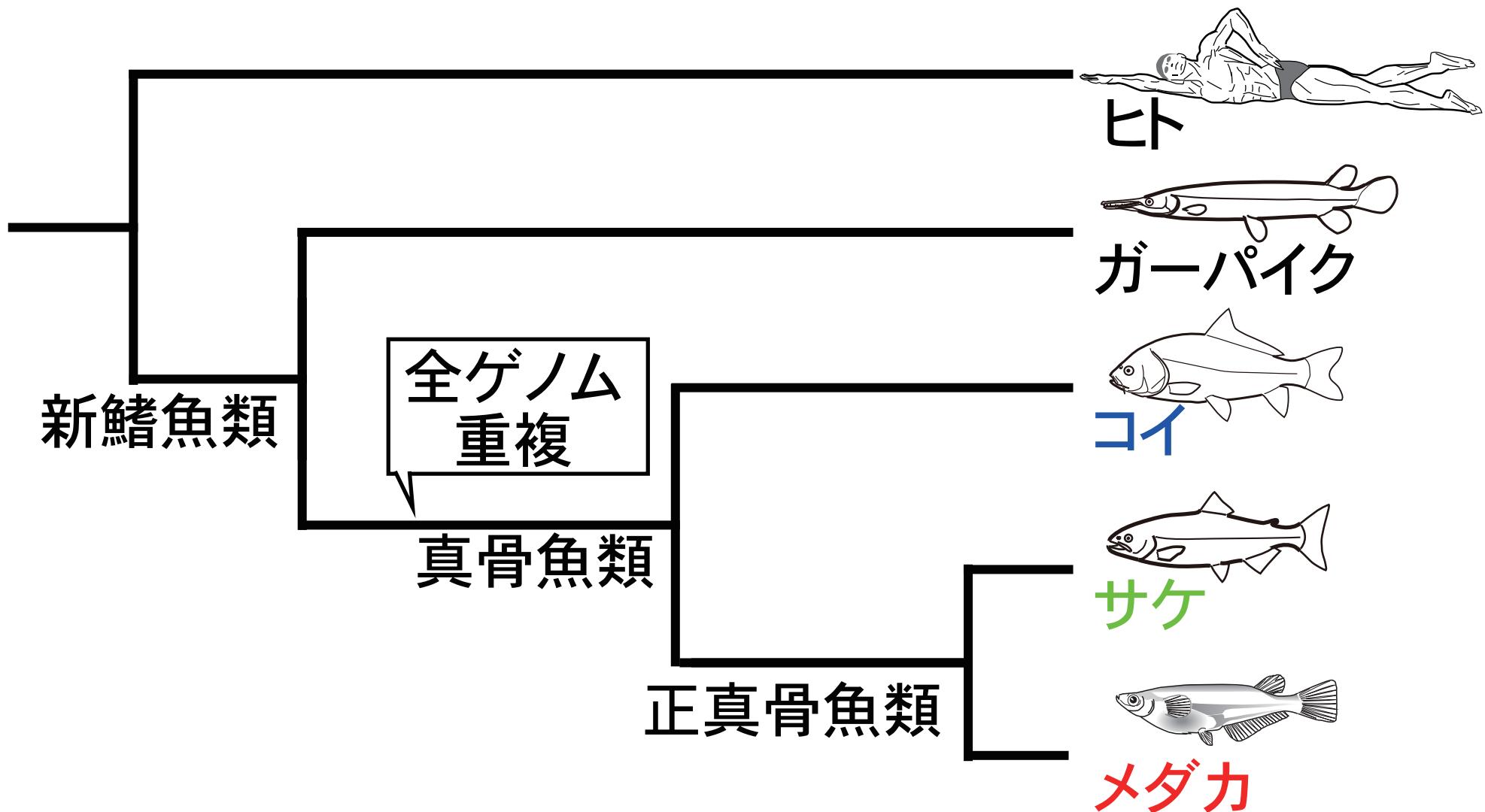
66 遺伝子



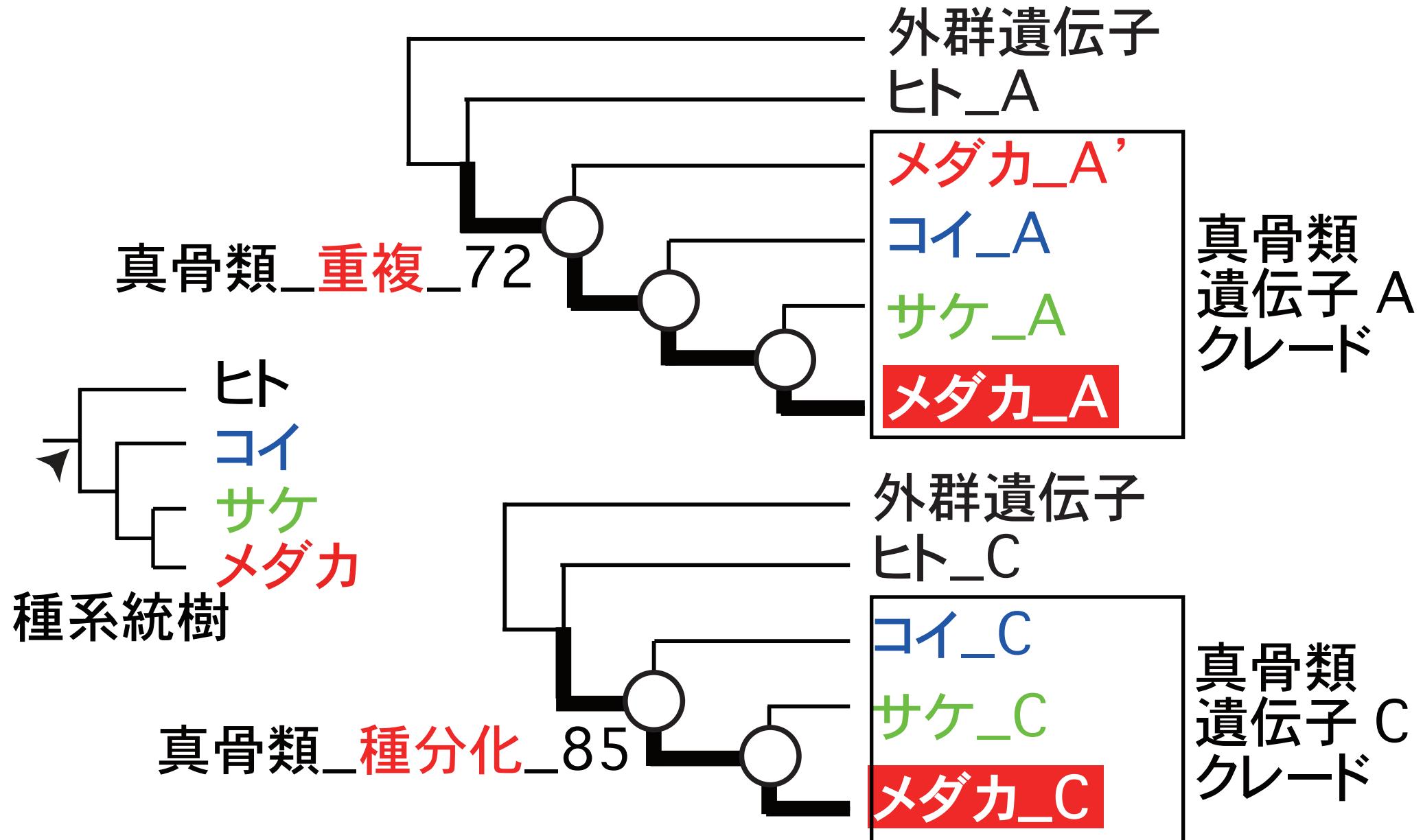
264 遺伝子

多くの遺伝子が再現に成功

2. 全ゲノム重複の検出

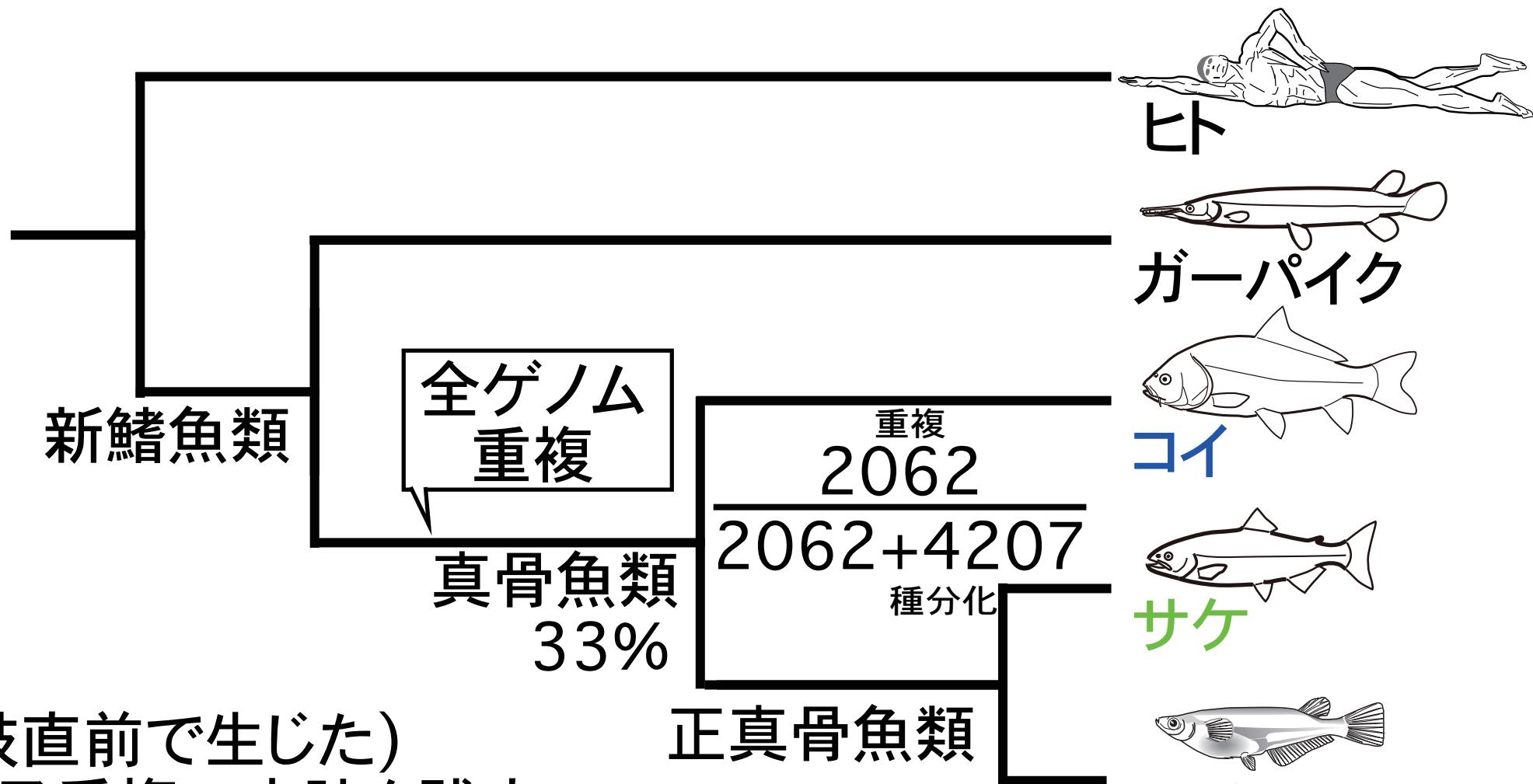


2. 全ゲノム重複の検出



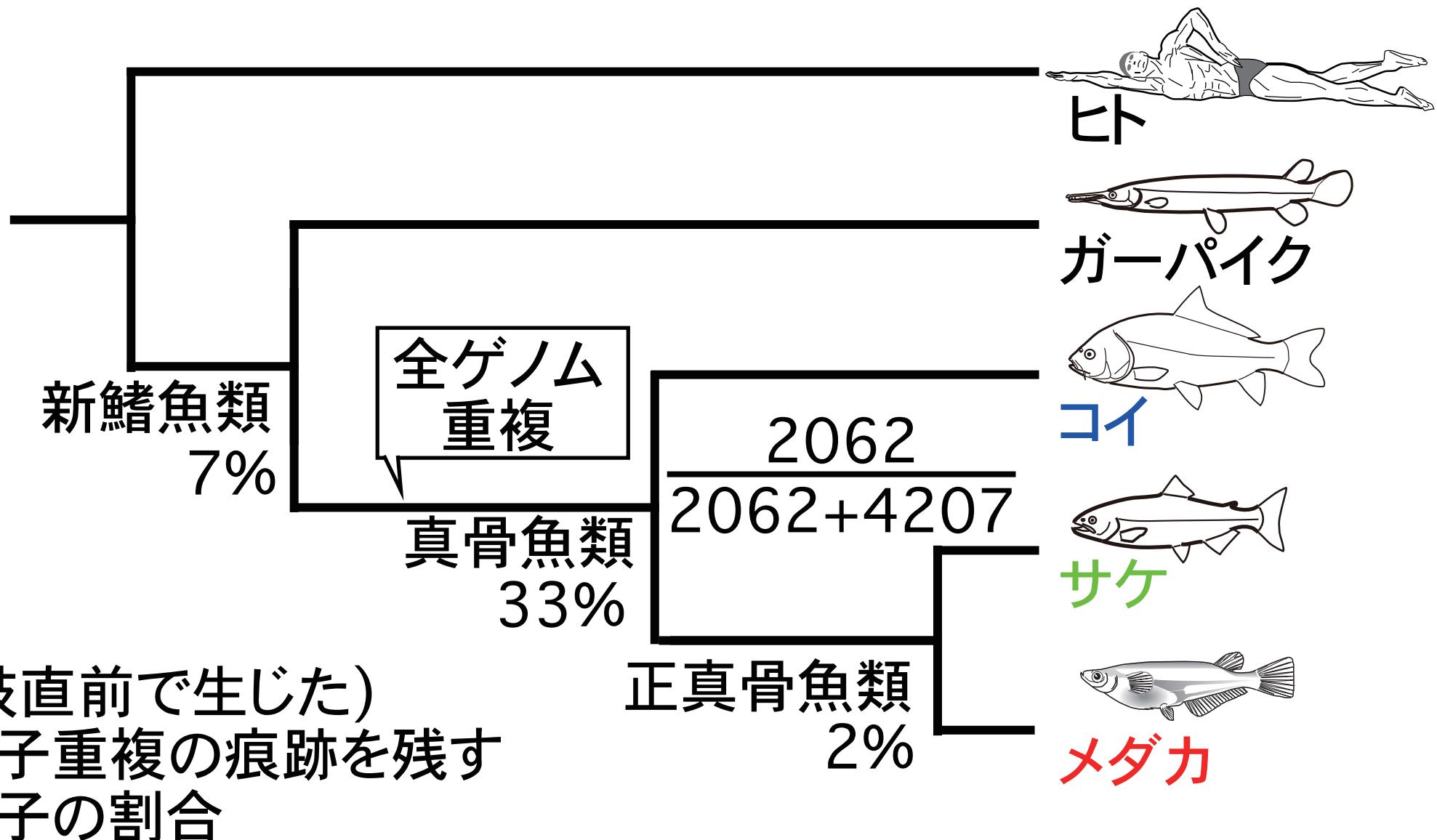
重複か種分化どちらか

2. 全ゲノム重複の検出



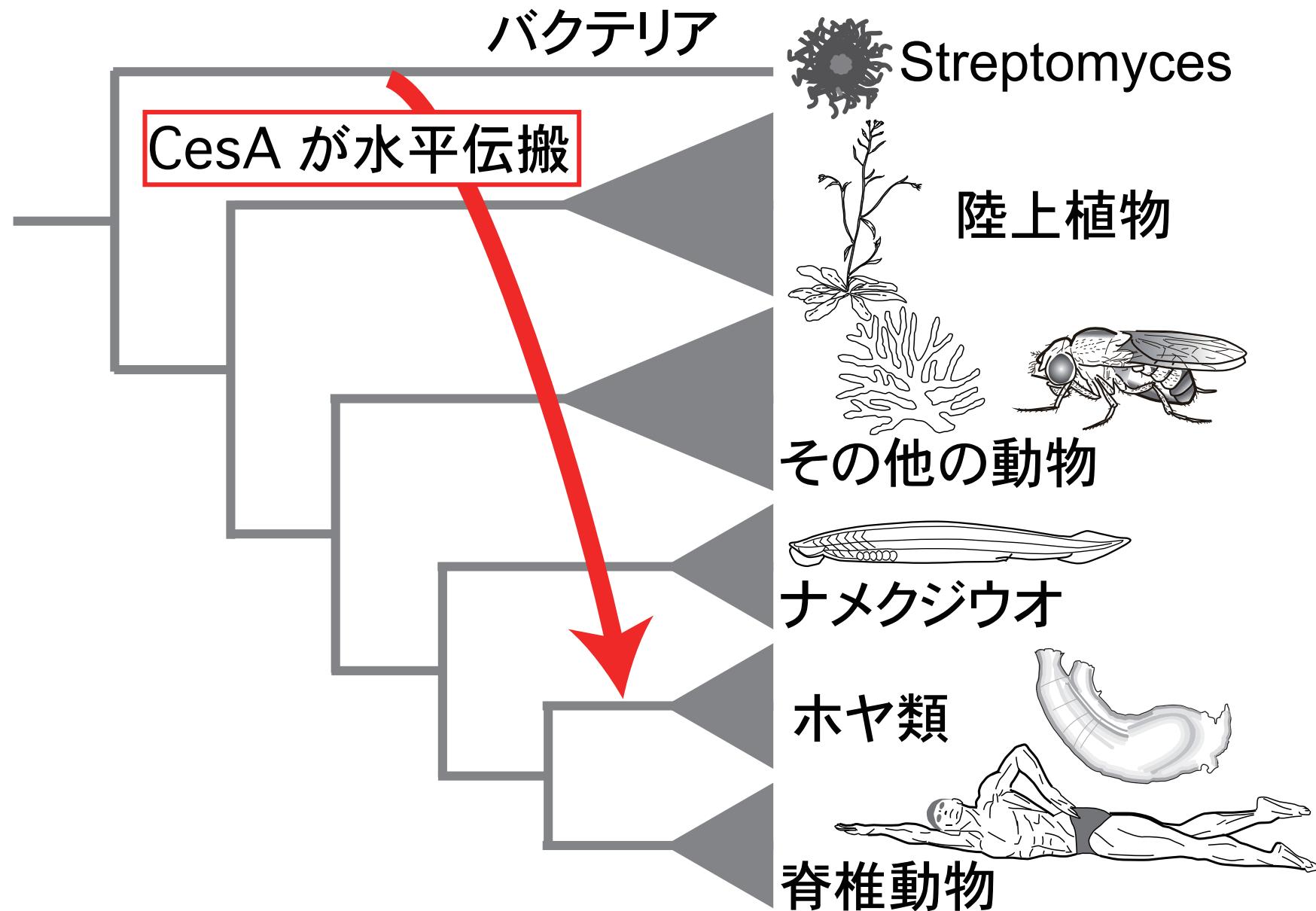
(分岐直前で生じた)
遺伝子重複の痕跡を残す
遺伝子の割合

2. 全ゲノム重複の検出

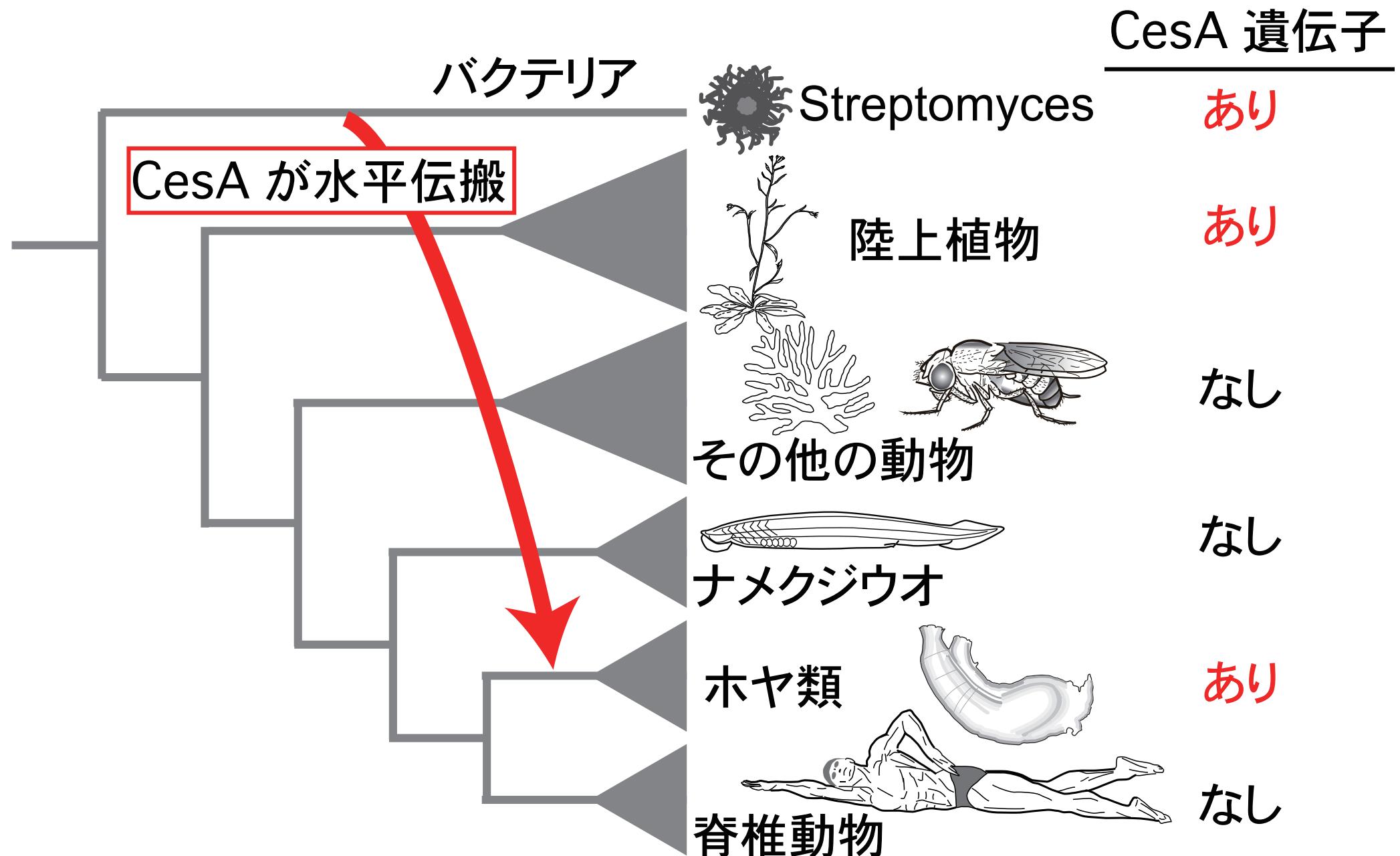


検出に成功。祖先ゲノムの状態を比較。

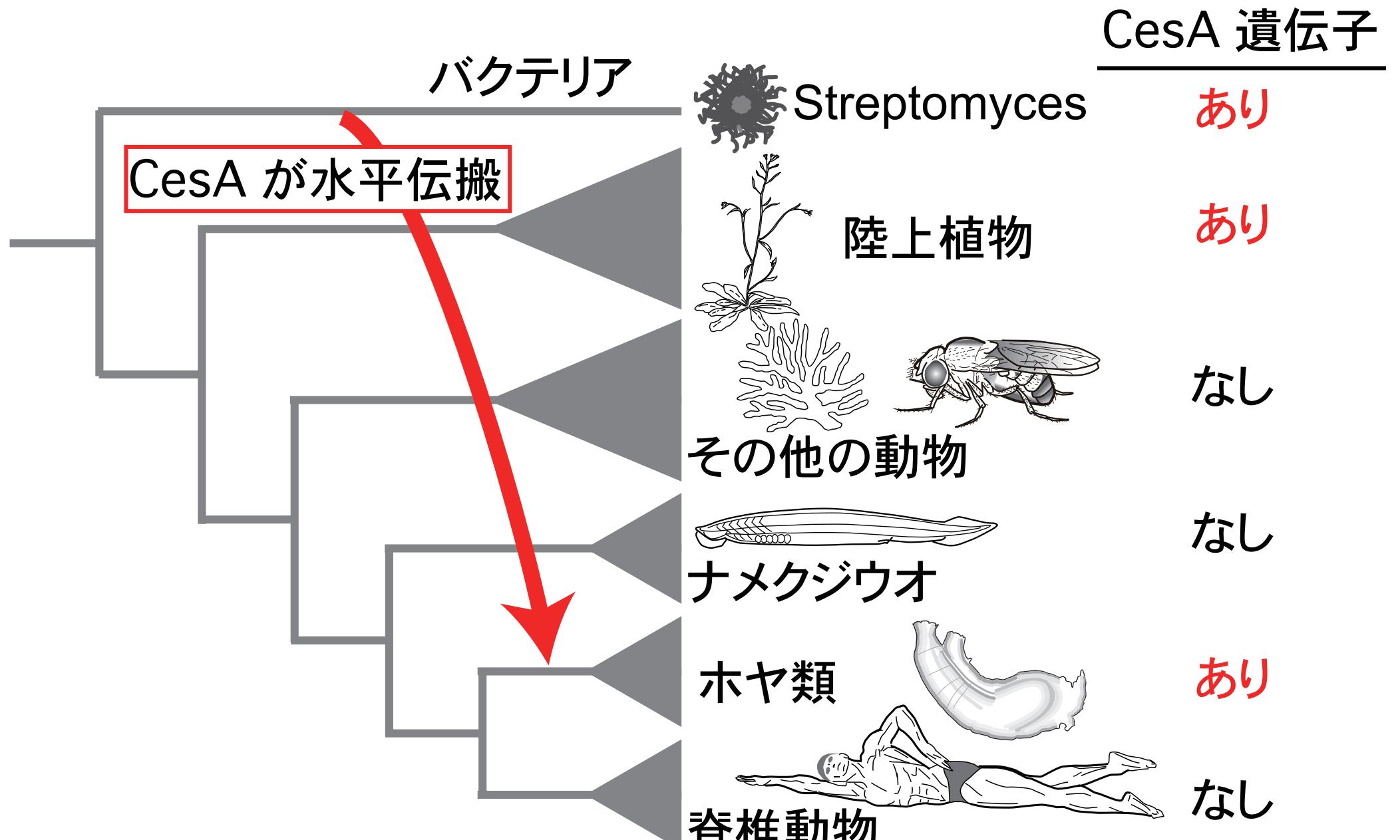
3. 水平伝播の検出



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

Code Issues 0 Pull requests 0 Projects 0 Wiki Security Insights

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

Dismiss

Sign up

An automatic web tool for phylogenetic i

640 commits

Branch: master New pull request

jun-inoue a

images tarfiles .DS_Store README.md

README.md

ORTHOSCOPE

Web servise: <https://orthoscope.org>.
Jan. 2019.
Mirror site: <http://www.fish-evolution.org>
Japanese instruction: <http://www.fish-evolution.org>

Our hypothetical species tree

Metazoa

Species Tree Hypothesis

Vertebrate species tree hypothesis

Basal metazoans
Cnidarians
Spiralians
Bilateria
Protostomes
Deuterostomes
Chordates
Vertebrates
Gnathostomata (Kikugawa et al. 2004)
Cyclostomata
Chondrichthyes
Chimaeriformes
Orectolobiformes
Actinopterygii
Coelacanthiformes
Gymnophiona
Amphibia
Batrachia
Anura
Gekkonidae
Lepidosauria (Townsend et al. 2004)
Unidentata
Serpentes
Colubroidea
Viperidae
Agamidae
Dactyloidae
Trionychidae
Testudines
Archelosauria (Meyer and Zardoya 2003)
Sauria
Monotremata
Marsupialia
Didelphimorphia
Dasyuromorphia/Diprotodontia
Diprotodontia
Cinquulata
Pilosa
Tubulidentata/Tenrecidae
Afrotheria
Xenarthra
Mammalia (Meyer and Zardoya 2003)

Eptatretus burgeri
Petromyzon marinus
Callorhinus milii
Rhinodon typus
Actinopterygii
Latimeria chalumnae
Rhinatremata bivittatum
Caecilia tentaculata
Typhlonectes compressicauda
Microcaecilia dermatophaga
Microcaecilia unicolor
Typhlonectes compressicauda
Microcaecilia dermatophaga
Microcaecilia unicolor
Nanorana parkeri
Xenopus tropicalis
Xenopus laevis
Gekko japonicus
Python bivittatus
Thamnophis sirtalis
Protobothrops mucrosquamatus
Protobothrops flavoviridis
Pogona vitticeps
Anolis carolinensis
Pelodiscus sinensis
Chelonia mydas
Terrapene mexicana
Chrysemys picta
Crocodylus porosus
Gavialis gangeticus
Alligator mississippiensis
Alligator sinensis
Aves
Ornithorhynchus anatinus
Monodelphis domestica
Sarcophilus harrisii
Notamacropus eugenii
Dasypus novemcinctus
Choloepus hoffmanni
Ocyteropus afer
Echinosorex boehmi

Inshore hagfish
Sea lamprey
Elephant shark
Whale shark
Ray-finned fish
Coelacanth
Two-lined caecilian
Guayaquil caecilian
Cayenne caecilian
Tiny white caecilian
Caeclian
Wenxian knobby newt
High Himalaya frog
Western clawed frog
Clawed frog
Schlegel's Japanese gecko
Burmeese python
Common garter snake
Taiwan habu
Habu
Central bearded dragon
Green anole
Chinese soft-shelled turtle
Green seaturtle
Mexican box turtle
Painted turtle
Estuarine crocodile
True gharial
American alligator
Chinese alligator
Birds
Platypus
Gray short-tailed opossum
Tasmanian devil
Tamar wallaby
Nine-banded armadillo
Hoffmann's two-fingered sloth
Aardvark
Madagascar hedgehog

ORTHOSCOPE, dbCNS, 種の系統樹 など

研究協力者

西田睦（琉球大学）

佐藤矩行（OIST）

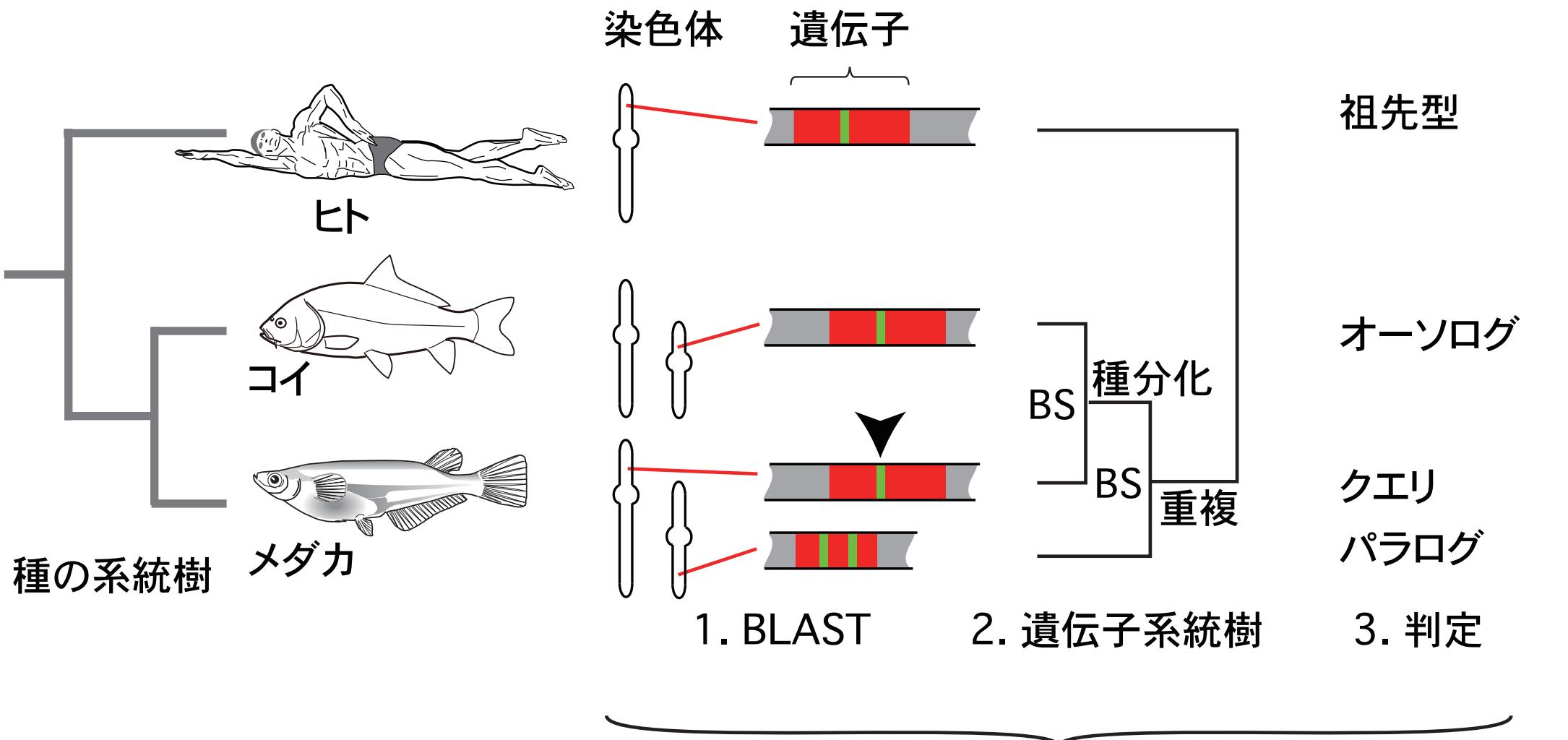
塚本勝巳（東大・大海研）

有本飛鳥（広島大学）

佐藤行人（琉球大学）

ユーザーの皆様

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)

—