

自ら触れる系統樹

～aLeavesを使った分子系統樹の推定～

工樂 樹洋
(国立遺伝学研究所)

全国生物教育研究会 2023年8月

自己紹介

分子進化・分子系統学（学部 / 修士）



発生生物学（博士）



ゲノム情報学

2007～2012年 ドイツ・コンスタンツ大で教員



ボーデン湖畔の街 コンスタンツ

学部・大学院の各種講義・実習を担当(英語)

シーラカンス、カメ、ウミヤツメの
国際ゲノムコンソーシアムに参加

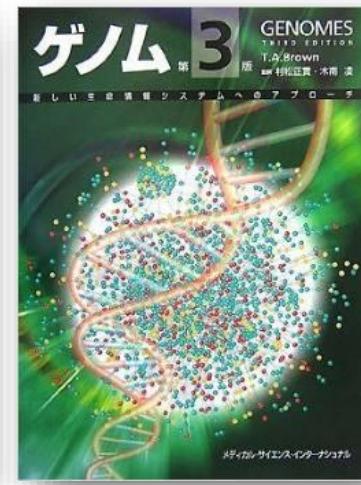
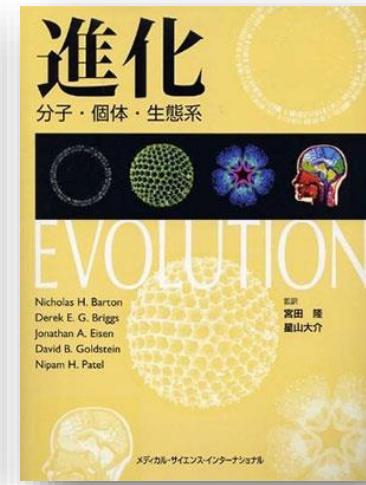
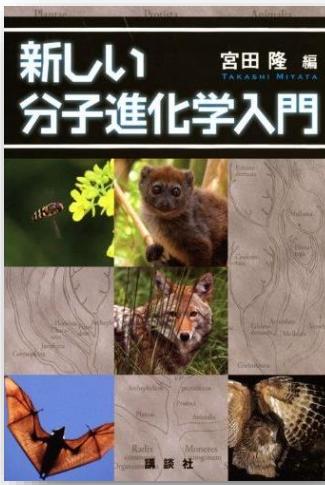
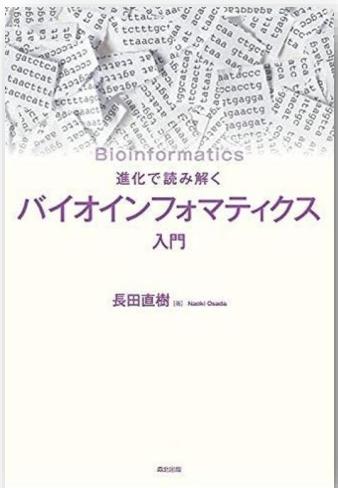
2012年～ 神戸理研でDNA解析施設・研究室を指揮



2021年～ 独自の研究は静岡県三島市の国立遺伝研へ移行



参考書籍



✓ いいね！ 130

ツイート



2018年10月9日

理化研究所

沖縄美ら島財団

海遊館

大阪市立大学

東京大学

サメのゲノムを解読

–サメの進化・生態の解明への新たな手がかり–

要旨

理化研究所（理研）生命機能科学研究センター分子配列比較解析ユニットの工楽樹洋ユニットリーダー、沖縄美ら島財団総合研究センターの佐藤圭一上席研究員（沖縄美ら海水族館副館長）、海遊館の西田清徳館長、大阪市立大学大学院理学研究科の小柳光正教授、東京大学大気海洋研究所の兵藤晋教授らの共同研究グループ※は、「イヌザメ」「トラザメ」「ジンベエザメ」の全ゲノム配列を解析しました。

本研究成果から得られた高精度のゲノム情報は、脊椎動物の進化および海の生態系において独特の位置を占める「軟骨魚類」の発生や生態、また水族館などでの飼育・繁殖に関する多様な研究に貢献すると期待できます。

軟骨魚類のうち、サメ類とエイ類が属する「板鰓（ばんさい）類^[1]」は、ゲノムサイズ^[2]が大きいことなどが障壁となり、これまで、高精度な全ゲノム配列情報が得られていませんでした。今回、共同研究グループは、約1,200種知られる板鰓類のうち、卵生^[3]で繁殖力の高いイヌザメとトラザメに着目し、全ゲノム配列を解読しました。また、現生で最大の魚類ジンベエザメについては、既に公開されていたゲノム配列をより高精度の情報に整えました。これらのゲノム情報に加えて、生体から採取した細胞のRNAデータを活用することで、形態形成や視覚などをつかさどる遺伝子群の進化を詳しく解析し、サメ類とヒトを含む他の脊椎動物との共通点と相違点を多数発見しました。

本研究は、英国のオンライン科学雑誌『Nature Ecology and Evolution』（10月8日付け：日本時間10月9日）に掲載されます。

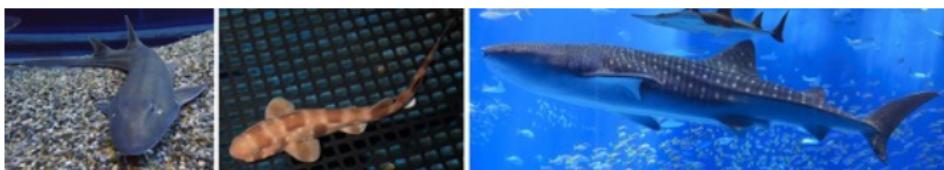


図 今回全ゲノム解析したイヌザメ成魚、トラザメ幼魚、ジンベエザメ成魚（左より）

2023年3月23日

プレスリリース

本論文はすでに公開されています
本情報はすぐにご利用いただけます

ジンベエザメだけに起きた視覚の進化 ～深海生活への適応か？～

■ 概要

動物は、生息環境中の光を活用するしくみをもっています。なかでも、深海や地中など微弱な光環境で暮らす生物種は、その限られた光を活用する特別なしくみを備えています。最大の魚類であるジンベエザメは、海面近くで摂餌をする一方で、光の届きにくい深海にも潜ることが知られています。

今回、情報・システム研究機構 国立遺伝学研究所の工楽樹洋教授（理化学研究所 チームリーダー）、大阪公立大学の小柳光正教授が率いる研究チームは、ジンベエザメが光を活用するしくみの解明に挑みました。

本研究チームは、眼の網膜にある光受容タンパク質（オプシン⁽¹⁾）のうち、微弱光下での視覚をつかさどるロドプシンについて、DNA情報と吸収する光の波長を測る分光測定を組み合わせることで、ジンベエザメと他のサメ類を比較しました（図1）。その結果、ジンベエザメのロドプシンは、従来の常識を覆すアミノ酸置換によって、深海の中で最も届きやすい青色の光を効率的に受け取ることができることを明らかにしました。また、ジンベエザメのロドプシンは熱に弱く、低温となる深海での機能に適していることもわかりました。

本結果から、ジンベエザメの視覚は、水温が低下した深海において、微弱な光を活用できるような進化を遂げたことが示唆されます。この進化の引き金になったアミノ酸置換部位はヒトの夜盲症⁽²⁾の原因となる置換部位でもあり、ジンベエザメの海面付近から水深2000メートル付近の深海まで潜水するというその独特的のライフスタイルの表れともいえるかもしれません。

さらに、本研究は、様々な生物の暮らしをその生体を犠牲にすることなく解き明かすために、DNAの情報を活用したタンパク質の人工合成技術が有用であることを示しました。

本研究成果は、米国科学雑誌『Proceedings of the National Academy of Sciences of the United States of America (PNAS)』に2023年3月22日午前1時（日本時間）に短報論文として掲載されました。



海洋博公園・沖縄美ら海水族館提供

2018年10月9日

理化学研究所

沖縄美ら島財団

海遊館

大阪市立大学

東京大学

サメのゲノムを解読

–サメの進化・生態の解明への新たな手がかり–

要旨

理化学研究所（理研）生命機能科学研究センター分子配列比較解析合研究センターの佐藤圭一上席研究員（沖縄美ら海水族館副館長）の小柳光正准教授、東京大学大気海洋研究所の兵藤晋教授らの共同「ジンベエザメ」の全ゲノム配列を解析しました。

本研究成果から得られた高精度のゲノム情報は、脊椎動物の進化おの発生や生態、また水族館などでの飼育・繁殖に関する多様な研究

軟骨魚類のうち、サメ類とエイ類が属する「板鰓（ばんさい）類^[1]これまで、高精度な全ゲノム配列情報が得られていませんでした。うち、卵生^[3]で繁殖力の高いイヌザメとトラザメに着目し、全ゲノムについて、既に公開されていたゲノム配列をより高精度のら採取した細胞のRNAデータを活用することで、形態形成や視覚などヒトを含む他の脊椎動物との共通点と相違点を多数発見しました。

本研究は、英国のオンライン科学雑誌『Nature Ecology and Evolution』に掲載されました。

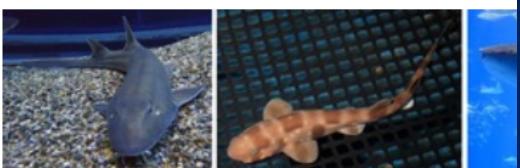


図 今回全ゲノム解析したイヌザメ成魚、トラザメ



海遊館ホームページより

2023年3月23日

エザメだけに起きた視覚の進化
・深海生活への適応か？～

しくみをもっています。なかでも、深海や地中など微弱な光環境で暮らる特別なしくみを備えています。最大の魚類であるジンベエザメは、海にいく深海にも潜ることが知られています。

遺伝学研究所の工樂樹洋教授（理化学研究所 チームリーダー）、大阪先チームは、ジンベエザメが光を活用するしくみの解明に挑みました。

受容タンパク質（オプシン⁽¹⁾）のうち、微弱光下での視覚をつかさどる口の波長を測る分光測定を組み合わせることで、ジンベエザメと他果、ジンベエザメのロドブシンは、従来の常識を覆すアミノ酸置換によつ光を効率的に受け取ることができることを明らかにしました。また、ジン

となる深海での機能に適していることもわかりました。

水温が低下した深海において、微弱な光を活用できるような進化を遂げき金になったアミノ酸置換部位はヒトの夜盲症⁽²⁾の原因となる置換部位から水深 2000 メートル付近の深海まで潜水するというその独特的のライフ

ん。
しをその生体を犠牲にすることなく解き明かすために、DNA の情報を活用であることも示しました。

edings of the National Academy of Sciences of the United States of America, 2023, 120(13), e2311353, pp. 1-10.
午前 1 時（日本時間）に短報論文として掲載されました。



海洋博公園・沖縄美ら海水族館提供

Whale shark ジンベエザメ *Rhincodon typus*
Ningaloo Coast World Heritage Area

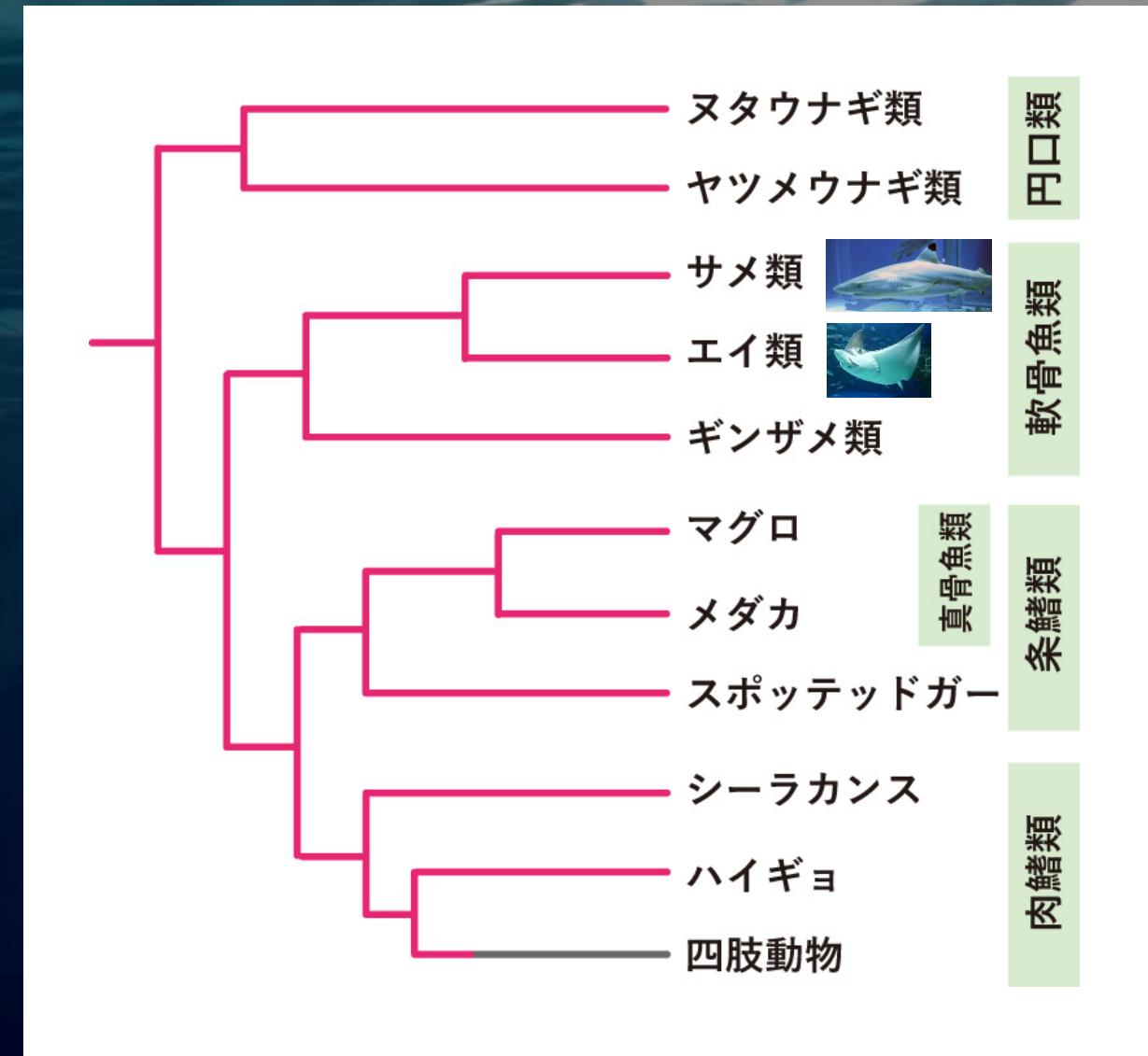


“魚”類研究 ≈ 脊椎動物研究

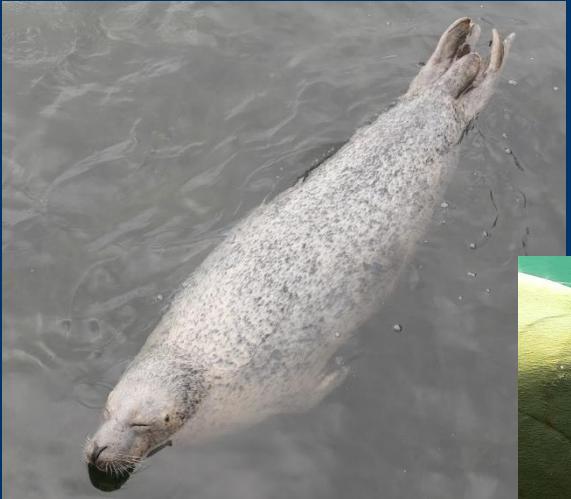
魚類大系統間の関係は決着
例、Kikugawa et al., *BMC Biol*, 2004

核の全ゲノム情報も得られ、
各魚類系統「らしさ」の理解を
目指す分子レベルの研究の
土台が整った

とくに遅れていたのは：
ヌタウナギ
板鰓類
ハイギョ

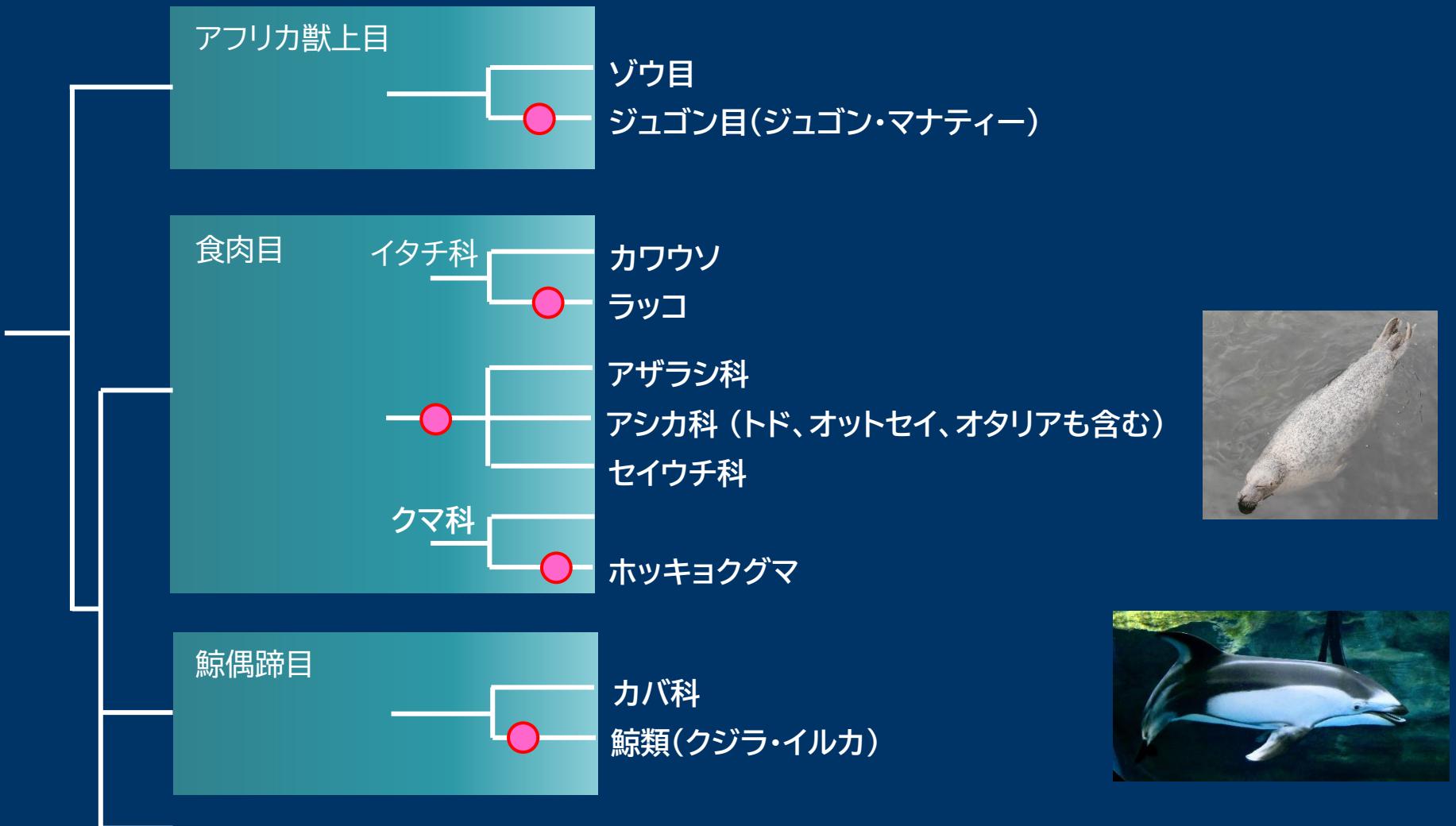


例) 海棲哺乳類

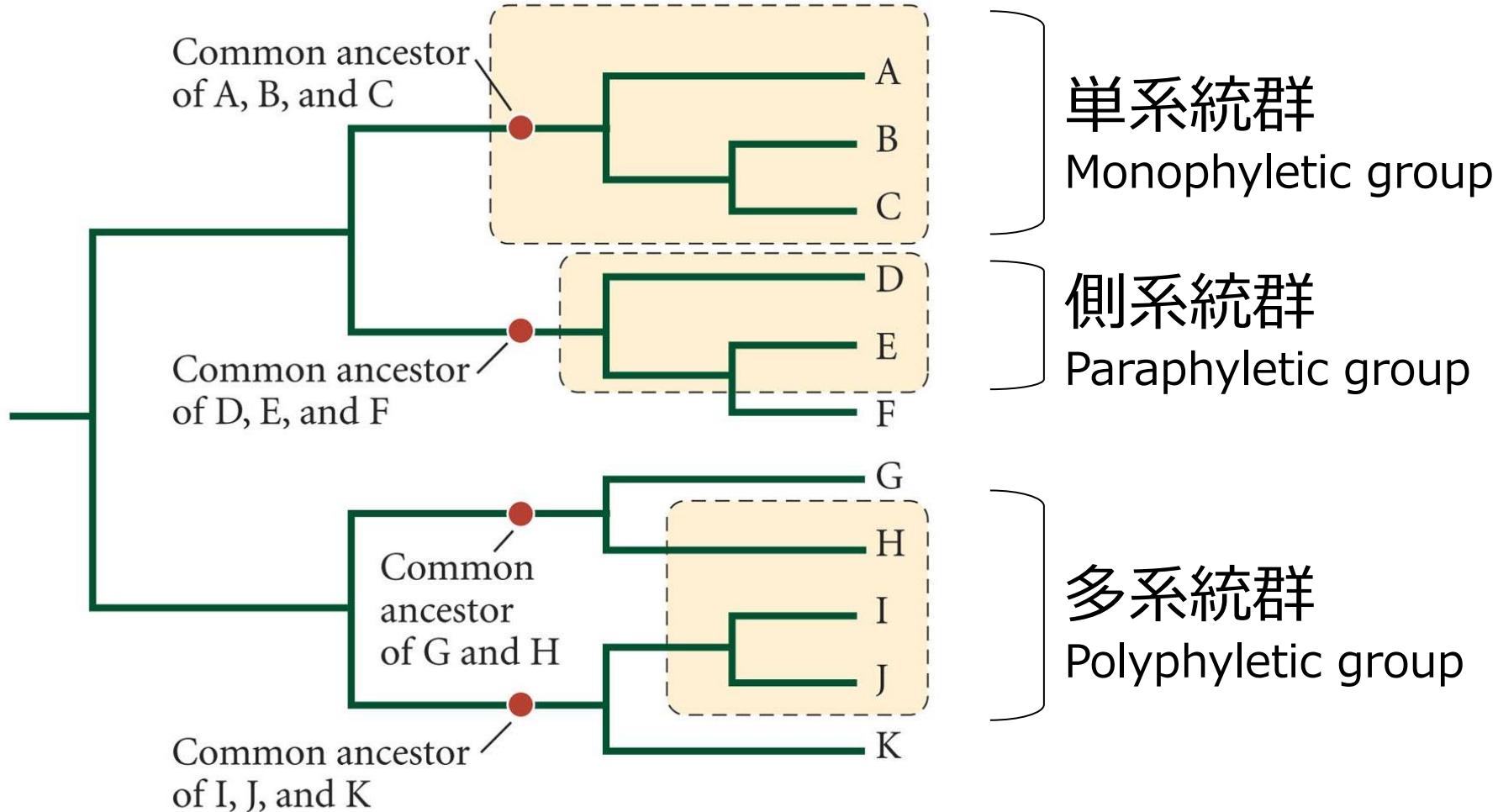


哺乳類の海生適応は複数回起きた

● 海への進出



系統分類学の用語



分子系統樹推定の作業の流れ

多重アラインメント
Multiple alignment



アルゴリズム
Algorithm



系統樹

Phylogenetic tree

UPGMA法
近隣結合 (Neighbor-joining) 法
最節約 (Maximum parsimony) 法
最尤 (Maximum-likelihood) 法
ベイズ法 (Bayesian inference)
など

得られた系統樹の統計的評価
例) ブートストラップ法

分子系統樹推定：概論

- 配列セットの準備とアライメントも非常に重要
情報の欠落 & 誤ったInputは誤った結論を導きやすい
- 万能の系統樹推定法は存在しない
複数の方法を独立に試すことが重要 — 結果の相違を検証
'Tree space' — 答えはひとつとは限らない
- 系統樹の視覚化
「トポロジー」と「枝長」
系統樹は部分的にswappable
再度、計算時間も重要なファクター

分子系統樹推定：実践

- ・「根-root」の扱い
‘有根系統樹 (rooted tree)’ or ‘無根系統樹 (unrooted tree)’
- ・塩基配列かアミノ酸配列か？核ゲノムかミトコンドリアか？
 答える問題による—古い分岐はアミノ酸配列で
 ミトコンドリアDNAは進化速度が大きい—より新しい分岐に向く
- ・系統樹推定法
 距離行列法：UPGMA法、近隣結合法(NJ法)
 形質状態法：最尤法(ML法)、最節約法(MP法)
 ↓
 完全探索か発見的探索か？
- ・推定結果の統計的評価法 例) ブートストラップ法

分子系統樹推定：実践

- ・検討を要する配列

 - 部分（Partial）配列

 - 未決定領域（N/X）を含む配列

 - 冗長な（複数の同一な）配列

- ・アウトグループの選び方

 - 複数選ぶのがよい

 - イングループにできるだけ近い方が良い

 - イングループに内包される配列であってはならない

- ・どの範囲を含めるのか？

 - 最適な範囲は答えた問い合わせによる

 - 広げれば広げるほど、使える情報（座位数）は減る傾向

2013年日本動物学会岡山大会にて口頭発表後、情報を追加



aLeaves (エイリーヴズ) を用いて 分子系統樹推定におけるデータセット生成の ストレスを減らす

~ tackling a bottleneck of modern molecular phylogenetics ~

工樂 樹洋 (理研CDB・ゲノム資源解析ユニット)
当時

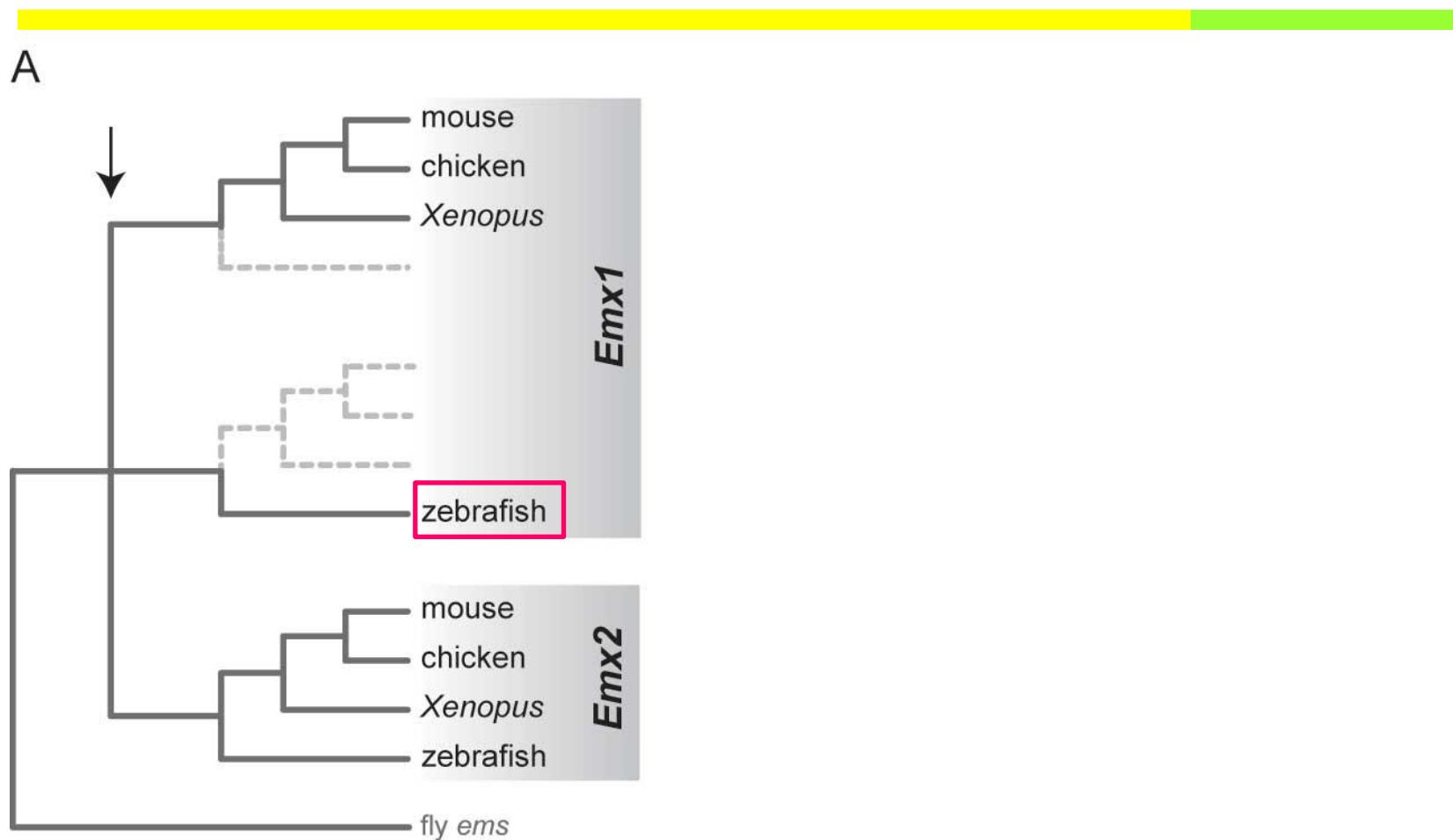
Shigehiro Kuraku (Genome Resource & Analysis Unit, RIKEN CDB)

コンセプト

系統樹推定法が成熟し、多数のゲノムプロジェクトにより配列情報が非常に豊かになったにもかかわらず、実験生物学者にとって分子系統樹推定の作業は身近になるどころか、より手の届きにくいものになったのではないか？

情報過多になりがちな昨今の配列リソースを網羅しつつも、生物学者が頼りにしやすい手掛けりを用いて配列数を絞り込むことができるオンラインツールをつくる

Hidden paralogy とは?: 例、zebrafish *Emx3*

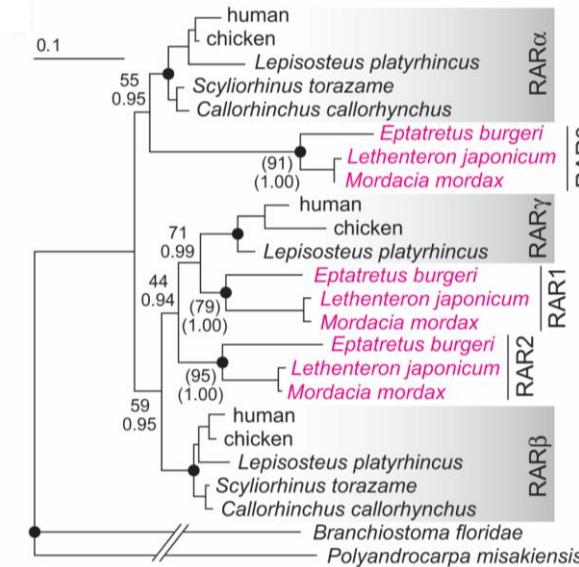
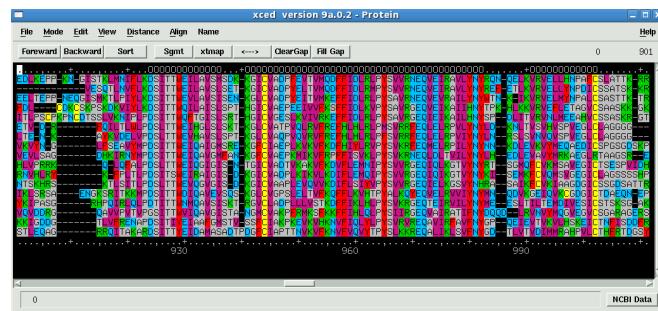


Morita et al., 1995

Reviewed in Kuraku, 2010. *Integ. Comp. Biol.*

ホモログ配列セットをどうやって用意すればよいか？

A) Exhaustive search
of homologs



B) Heuristic collection

Nucleotide ▾ AB292622 × Search

類似性検索にはBLASTが有用、ただし

BLAST® Basic Local Alignment Search Tool My NCBI [Sign In] [Register]

NCBI/ BLAST/ blastp suite Standard Protein BLAST

blastn blastp blastx tblastn tblastx

BLASTP programs search protein databases using a protein query. [more...](#) Reset pa Bookmar

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) Clear Query subrange [?](#)

From To

Or, upload file [参照...](#) ファイルが選択されていません. [?](#)

Job Title
Enter a descriptive title for your BLAST search [?](#)

Align two or more sequences [?](#)

Choose Search Set

Database Non-redundant protein sequences (nr) [?](#)

Organism [Optional](#) Enter organism name or id—completions will be suggested Exclude [+](#)
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown. [?](#)

Exclude [Optional](#) Models (XM/XP) Uncultured/environmental sample sequences

An Essential Guide to the Basic Local Alignment Search Tool

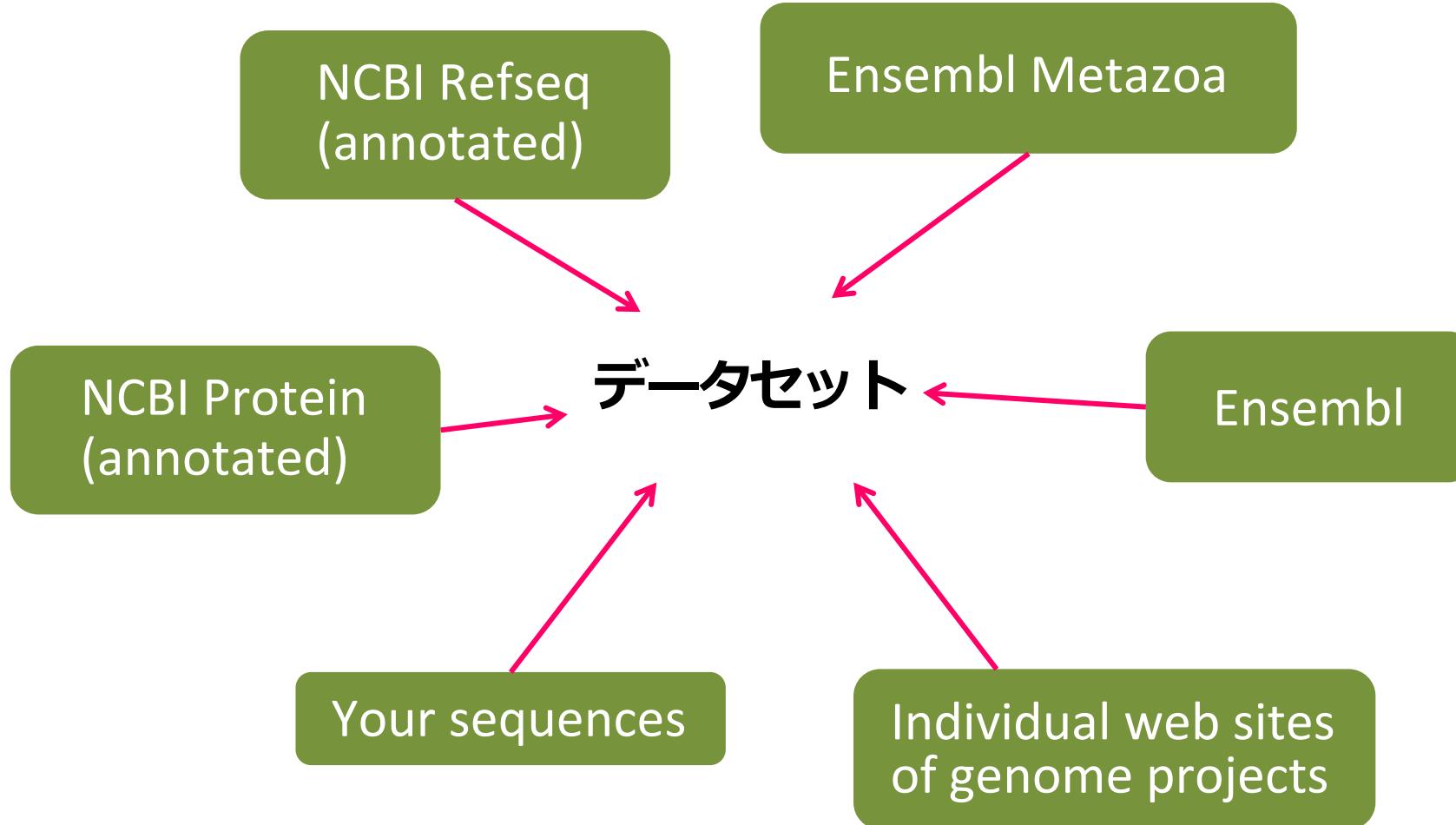
BLAST



OREILLY™ Ann Koef, Mark Yandell & Joseph Russell

"Every BLAST search is an experiment" by

配列情報が散在していることがネック



コンセプト

系統樹推定法が成熟し、多数のゲノムプロジェクトにより配列情報が非常に豊かになったにもかかわらず、実験生物学者にとって分子系統樹推定の作業は身近になるどころか、より手の届きにくいものになったのではないか？

情報過多になりがちな昨今の配列リソースを網羅しつつも、
生物学者が頼りにしやすい手掛けりを用いて配列数を絞り込む
ことができるオンラインツールをつくる

Collaborators

Osamu Nishimura 西村 理



GRAS, RIKEN CDB

Kazutaka Katoh 加藤 和貴



CBRC, AIST
&
iFReC, Osaka Univ.

Christian M. Zmasek



Sanford-Burnham
Medical Research Institute
USA

For citation

*W22-W28 Nucleic Acids Research, 2013, Vol. 41, Web Server issue
doi:10.1093/nar/gkt389*

Published online 15 May 2013

aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity

Shigehiro Kuraku^{1,*}, Christian M. Zmasek², Osamu Nishimura¹ and Kazutaka Katoh^{3,4,*}

¹Genome Resource and Analysis Unit, RIKEN Center for Developmental Biology, Kobe, Hyogo 650-0047, Japan, ²Program in Bioinformatics and Systems Biology, Sanford-Burnham Medical Research Institute, La Jolla, CA 92037, USA, ³Immunology Frontier Research Center, Osaka University, Osaka 565-0871, Japan and ⁴Computational Biology Research Center, The National Institute of Advanced Industrial Science and Technology (AIST), Tokyo 135-0064, Japan

Received February 17, 2013; Revised April 11, 2013; Accepted April 18, 2013

ABSTRACT

We report a new web server, aLeaves (<http://aleaves.cdb.riken.jp/>), for homologue collection from diverse animal genomes. In molecular comparative studies involving multiple species, orthology identification is the basis on which most subsequent biological analyses rely. It can be achieved most accurately by explicit phylogenetic inference. More and more species are subjected to large-scale sequencing, but the resultant resources are scattered in independent project-based, and multi-species, but separate, web sites. This compli-

INTRODUCTION

In any cross-species comparison at the molecular level, identification of orthology and paralogy is the basis on which most subsequent analyses rely (1). The most reliable approach for distinguishing orthologues from paralogues is by explicit phylogenetic inference. Some databases host genome-wide sets of molecular phylogenetic trees for individual gene families ('phylomes') (2–4). However, those existing databases provide only phylomes for a limited number of species with genome-wide sequence resources, and cannot fully accommodate biologists' daily demands for custom data sets, sometimes including organisms without genome-wide information or sequences identified on their own. To achieve smooth

aLeaves – <http://aleaves.cdb.riken.jp>

aLeaves - [èlívz]

first step to build zoologically informative phylogenetic trees

Top Database Species History Help About Links

What is 'aLeaves'?

aLeaves allows you to collect homologs of the sequence of your interest. Those sequences can be passed on to the web server for sophisticated multiple alignment, and then to tree building. This initial process one can perform here is just like collecting leaves (termini of trees) from diverse animals. This is why we designated this tool 'aLeaves'.

Start the search (powered by NCBI Blast)

Enter your "query" peptide sequence in the fasta format ([example](#)):

Or upload a file: Browse...

Search Reset Select/Unselect all

Select database (one or more) (number of sequences in parentheses):

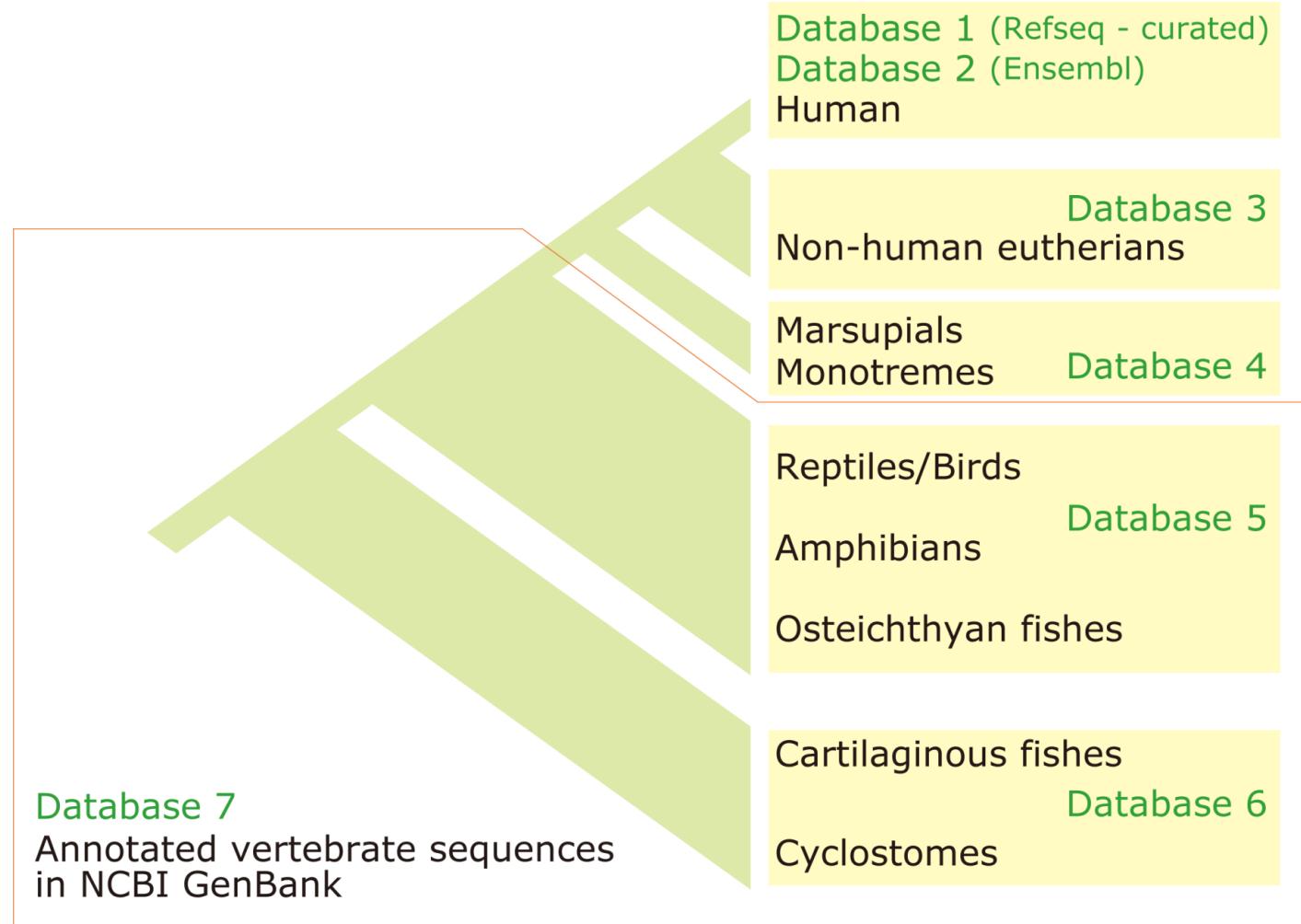
- 1. Human - Refseq (26,077) [[detail](#)]
- 2. Human - Ensembl (104,785) [[detail](#)]
- 3*. Non-human eutherians - Ensembl (743,219) [[detail](#)]
- 4. Non-human non-eutherians - Ensembl (1,111,111) [[detail](#)]

アミノ酸配列クエリー 1 本
から検索スタート

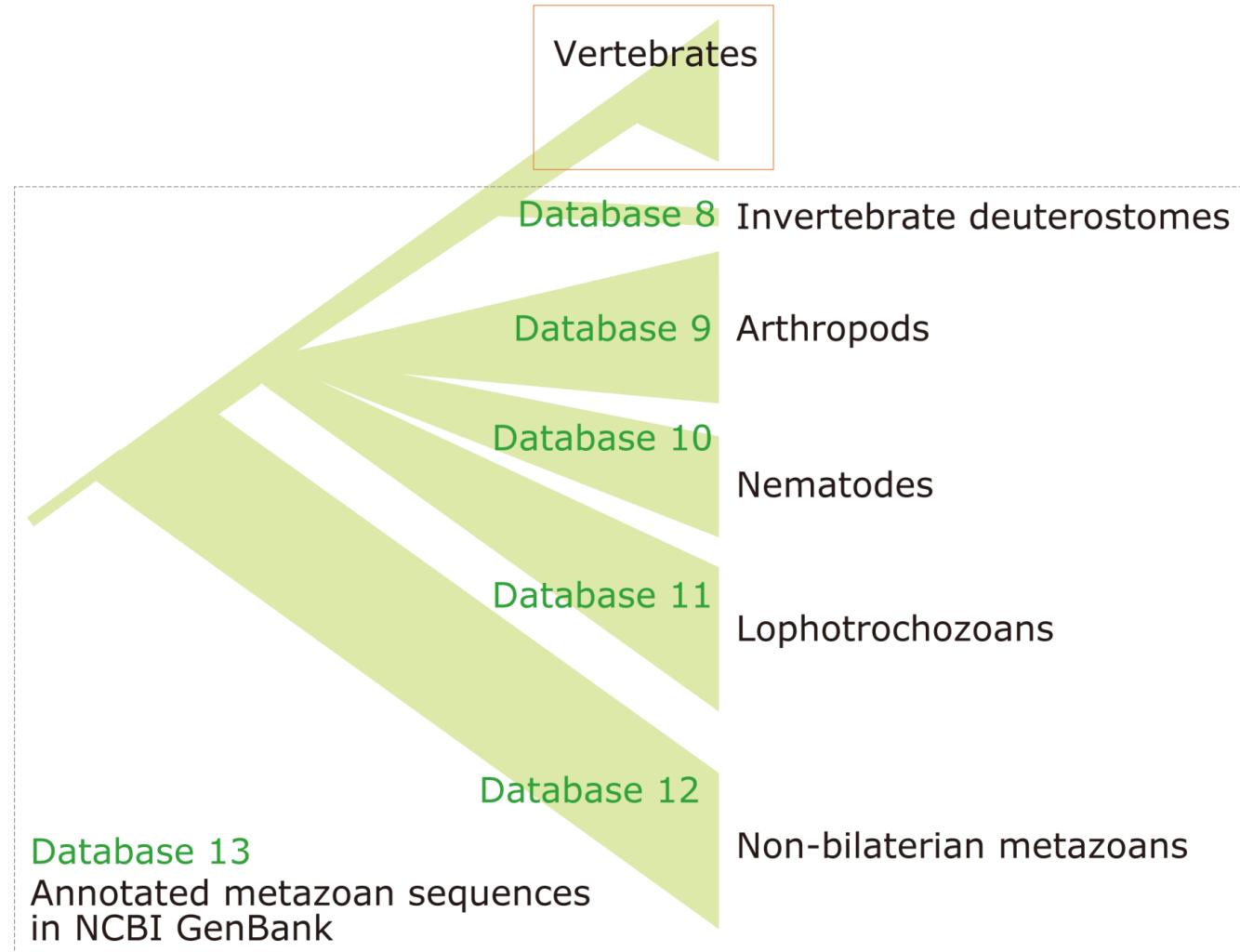
EnsemblやNCBIに加え、
それらに含まれないナメクジウオなどの
配列も含めてワンサーチで

5 分以内に 対象データベース
から配列を収集し multifasta
形式で出力

検索対象となるデータベースの構成 (1)



検索対象となるデータベースの構成（2）



検索をスタート

NCBI Proteinなどで、起点となる1本のアミノ酸配列を取得

The screenshot illustrates the process of retrieving a protein sequence from the NCBI Protein database.

Left Panel (Search Results):

- NIH National Library of Medicine logo.
- Protein search bar: Protein dropdown set to Protein, search term: Homo Rhodopsin.
- Species filter: Animals (8,346).
- Summary: 20 per page, Sort by Default order.
- Result count: Items: 1 to 20 of 9013.
- Link: See [rhodopsin rhodopsin-like](#) in the Gene database.
- Link: [rhodopsin reference sequences](#) Transcript (1).
- Link: [See the results of this search \(21 items\) in our](#) database.
- Link: [rhodopsin \[Homo sapiens\]](#)
- Details for the first result:
 - NCBI Reference Sequence: NP_000530.1
 - Identical Proteins, FASTA, Graphics links.
 - Go to: dropdown menu.
 - DEFINITION: rhodopsin [Homo sapiens].
 - ACCESSION: NP_000530.
 - VERSION: NP_000530.1
 - DBSOURCE: REFSEQ: accession NM_000539.3
 - KEYWORDS: RefSeq; MANE Select.
 - SOURCE: Homo sapiens (human)
 - ORGANISM: Homo sapiens
- Links at the bottom: GenPept, Identical Proteins, FASTA, Graphics.

Middle Panel (Sequence View):

rhodopsin [Homo sapiens]

NCBI Reference Sequence: NP_000530.1

Identical Proteins, **FASTA**, Graphics

Sequence details:
NP_000530.1 rhodopsin [Homo sapiens]
MNGTEGPNFVYPFSNATGVVRSPFEYPQQYLAEPWQFSMLAAYMFLILVLGFPINFLTLVYTVQHKKLRT
PLNYILLNLAVADLFMVLGGFTSTLYTSLHGKFVFGPTGCNLEGFFATLGGEIALWSLVLAIERYVV
KPMNSNRFGENHAIMGVAFTWMALACAAPPLAGWSRYIPEGLOQSCGIDYTLKPEVNNEFVIYMFMVV
HFTIPMIIIFCYGQLVFTVKEAAQQQESATTQKAKEEVTRMVIMVIAFLICWVPYASAFYIFTHQG
SNFGP1FM1T1PAFFAKSAAIYNPVIIYIMMNQFRNCMLTT1CCGKNPLGDDEASATVSKTETSQVAPA

Right Panel (aLeaves Tool):

aLeaves - [èlÍ:vz]

first step to build zoologically informative phylogenetic trees

Top Database Species History Help About Links

What is 'aLeaves'?

aLeaves allows you to collect homologs of the sequence of your interest. Those sequences can be passed on to the web server for sophisticated multiple alignment, and then to tree building tools. This initial process one can perform here is just like collecting leaves (termini of trees) from diverse animals. This is why we designated this tool 'aLeaves'.

Start the search (powered by NCBI Blast)

Enter your "query" peptide sequence in the fasta format (example):

```
>NP_000530.1 rhodopsin [Homo sapiens]
MNGTEGPNFVYPFSNATGVVRSPFEYPQQYLAEPWQFSMLAAYMFLILVLGFPINFLTLVYTVQHKKLRT
PLNYILLNLAVADLFMVLGGFTSTLYTSLHGKFVFGPTGCNLEGFFATLGGEIALWSLVLAIERYVV
KPMNSNRFGENHAIMGVAFTWMALACAAPPLAGWSRYIPEGLOQSCGIDYTLKPEVNNEFVIYMFMVV
HFTIPMIIIFCYGQLVFTVKEAAQQQESATTQKAKEEVTRMVIMVIAFLICWVPYASAFYIFTHQG
SNFGP1FM1T1PAFFAKSAAIYNPVIIYIMMNQFRNCMLTT1CCGKNPLGDDEASATVSKTETSQVAPA
```

検索をスタート

目的（調べたい範囲）に合わせてデータベースを選定

Select database (one or more) (number of sequences in parentheses):

- 1. Human - Refseq (62,342) [detail]
- 2. Human - Ensembl (116,646) [detail]
- 3*. Non-human eutherians - Ensembl (2,260,241) [detail]
- 4. Non-eutherian mammals - Ensembl (122,040) [detail]
- 5. Non-mammalian vertebrates - Ensembl (2,766,415) [detail]
- 6. Cartilaginous fishes 初期状態では、さまざまな動物門を含め
動物を広くカバーするチョイスとなっている (7,537,596)
- 7*. All vertebrate entries except mammalian (7,537,596)
- 8. Invertebrate deuterostomes (250,212) [detail]
- 9. Arthropods (2,106,748) [detail]
- 10. Nematodes (779,022) [detail]
- 11. Other protostomes (644,150) [detail]
- 12. Non-bilaterian metazoans (cnidarians, placozoan & poriferan) (644,150)
- 13*. All metazoan entries except vertebrates in NCBI Protein (11,024,596)

(Caution: selecting the database marked with * will largely slow down the search.)

Number of homologs to collect: (This number should not exceed 1000000000)

Threshold E-value:

Set low complexity filter or not [help of NCBI Blast]: no yes

昆虫の遺伝子だけ調べる場合。
全ゲノム情報に基づくデータベース9に加えて、
無脊椎動物の個別登録に基づくデータベース13も選定すると
探し漏らしがない。

Select database (one or more) (number of sequences in parentheses):

- 1. Human - Refseq (62,342) [detail]
- 2. Human - Ensembl (116,646) [detail]
- 3*. Non-human eutherians - Ensembl (2,260,241) [detail]
- 4. Non-eutherian mammals - Ensembl (122,040) [detail]
- 5. Non-mammalian vertebrates - Ensembl (2,766,415) [detail]
- 6. Cartilaginous fishes and cyclostomes (36,660) [detail]
- 7*. All vertebrate entries except mammalian (7,537,596) [detail]
- 8. Invertebrate deuterostomes (250,212) [detail]
- 9. Arthropods (2,106,748) [detail]
- 10. Nematodes (779,022) [detail]
- 11.
- 12.
- 13. 哺乳類には拘泥せず、脊椎動物を広くカバーしたい
場合にはこういうチョイスも有力。
データベース8は、アウトグループ（外群）として選定。

集める配列の保存性が低すぎると思われる場合には、この数値を大きくする（例、0.1や1）

Select database (one or more) (number of sequences in parentheses):

- 1. Human - Refseq (62,342) [detail]
- 2. Human - Ensembl (116,646) [detail]
- 3*. Non-human eutherians - Ensembl (2,260,241) [detail]
- 4. Non-eutherian mammals - Ensembl (122,040) [detail]
- 5. Non-mammalian vertebrates - Ensembl (2,766,415) [detail]
- 6. Cartilaginous fishes and cyclostomes (36,660) [detail]
- 7*. All vertebrate entries except mammalian (7,537,596) [detail]
- 8. Invertebrate deuterostomes (250,212) [detail]
- 9. Arthropods (2,106,748) [detail]
- 10. Nematodes (779,022) [detail]
- 11.
- 12.
- 13*. All metazoan entries except vertebrates in NCBI Protein (11,024,596) [detail]

配列検索後のステップはMAFFTサーバ上で

様々な基準に従い、配列の体系的な取捨選択が可能

- ・配列長
- ・酷似する配列を削除 (CD-HIT)
- ・Gap-freeな座位数を制限している配列を削除(Max-Align)
- ・特定の生物種に絞る
- ・Guide-tree(案内木)上の位置

From aLeaves to Tree Building
aleaves.cdb.riken.jp/aleaves/tmp/dir444_13021712

aLeaves - [èllí:vz]

first step to build zoologically

Top Database Species

Rough clustering of unaligned sequences

Input from aLeaves:

```
>query sequence
MNCTEGDNFVYVPSNKTGLARSPYEYQYYLAEI
VTIQHKKLRTPLNYILLNLAMANLFMVLFGTV
GEVALWSLVLAIERYIVICKPMGNFRGNTA
EGMQCSCGPDYTTLPNPNFNESYVYMFVVFHFL
ASTQKAKEVTRMVVLMVIGFLVCWVVPYASVAF
YNPVIYILMNQFRNCMITTLCCGKNPGLDDES
>PELSI|ENSPSIP0000008756 pep:nova
MNCTEGPNFYVPMNSNKTGIVRNPFYEPQYYLAEI
VTIQHKKLRTPLNYILLNLAFANLFMVGFFTT
GEITALWSLVLAIERYVYVCKPMNSFRSESHEA
EGMQCSCGIDYYTLKPEVHNESFYIMVFVHF
ATTQKAEREVTRMVVIMVISFLVCWVVPYASVAF
YNPVIYIVMNQFRNCMITTLCCGKNNLLGDDA
>TAEGU|ENSTGUP00000010860 pep:know
MNCTECODEVYDNMSNKTGIVRNPFYEPQYYLAEI
Submit Reset
```

Settings (help)

Distance measure:

- Rough (based on frequencies of amino acids)
- Accurate (based on scores of pairs)
- Accurate (based on scores of pairs)

Tools View Font Size Options Type Help

Phylogram
 Dyna Hide
 Rollover
 Show Internal Data
 Taxonomy Colorize
 Colorize Branches
 Use Branch-Widths

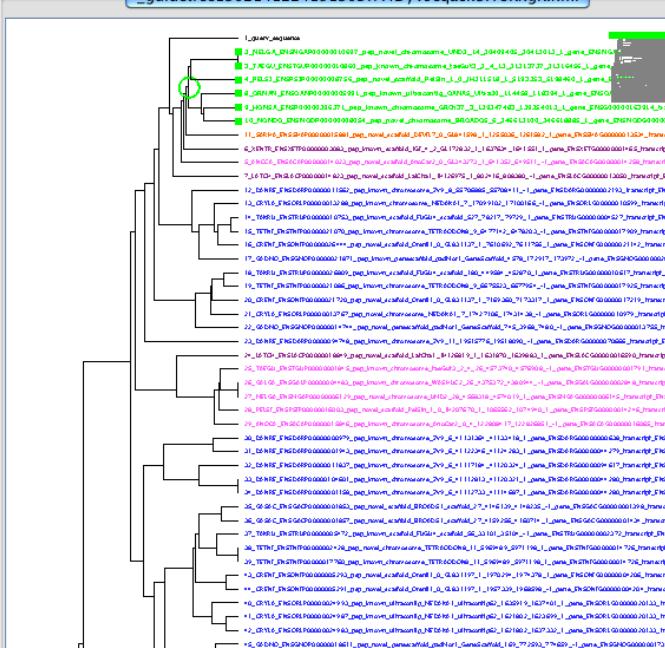
Display Data:
 Node Name
 Taxonomy Code
 Seq Name
 Confidence Values

Click on Node to:
Select/Unselect

Zoom:
Y+ Y- F +

Back to Super Tree
Order Subtrees
Uncollapse All

Search:



配列の絞り込み方法（1）—配列長

Unselected 16 sequence(s) ×

Select / Unselect all Invert selection Fill between top and bottom selections

CD-HIT removes redundant sequences. [help](#)

MaxAlign increases the number of gap-free sites. [help](#)

Show / Hide lengths Limit with length : 341 - 398

Highlight selected or unselected sequences on Phylo.io.

9 selected / 25 loaded / 25 input

0 / 25 Jump to bottom ↓

	アミノ酸残基数	蛋白名	属種
<input type="checkbox"/>	489	1 AAH50072.1 Forkhead box G1	[Homo sapiens]
<input type="checkbox"/>	451	2 NP_990524.1 forkhead box protein G1	[Gallus gallus]
<input type="checkbox"/>	482	3 XP_003214519.1 PREDICTED: forkhead box protein G1	[Anolis carolinensis]
<input type="checkbox"/>	432	4 AAI59145.1 foxg1 protein	[Xenopus tropicalis]
<input type="checkbox"/>	409	5 XP_006001628.1 PREDICTED: forkhead box protein G1	[Latimeria chalumnae]
<input type="checkbox"/>	409	6 XP_015205581.1 PREDICTED: forkhead box protein G1	[Lepisosteus oculatus]
<input type="checkbox"/>	420	7 NP_571142.1 forkhead box protein G1	[Danio rerio]
<input type="checkbox"/>	411	8 XP_007891603.1 forkhead box protein G1	[Callorhinus milii]
<input type="checkbox"/>	409	9 BPG57206.1 forkhead box protein FoxG1	[Scyliorhinus torazame]
<input checked="" type="checkbox"/>	409	10 Chihu0006553.p1 Chiloscyllium punctatum protein FoxG1 seq on C punctatum_v1.0	
<input type="checkbox"/>	333	11 XP_016853233.1 PREDICTED: forkhead box protein G1-like (FoxG2)	[Anolis carolinensis]
<input type="checkbox"/>	371	12 XP_005991682.2 PREDICTED: forkhead box protein G1-like (FoxG2)	[Latimeria chalumnae]
<input type="checkbox"/>	403	13 XP_015196379.1 PREDICTED: forkhead box protein G1-like (FoxG2)	[Lepisosteus oculatus]
<input checked="" type="checkbox"/>	379	14 NP_001038680.1 forkhead box protein G1c (foxg1c)	[Danio rerio]
<input type="checkbox"/>	316	15 GCC18449.1 hypothetical protein chiPu_0020798 (FoxG2), partial	[Chiloscyllium punctatum]
<input type="checkbox"/>	350	16 BBG57207.1 forkhead box protein FoxG2	[Scyliorhinus torazame]
<input type="checkbox"/>	330	17 XP_014342331.1 PREDICTED: forkhead box protein G1-like (FoxG3)	[Latimeria chalumnae]
<input checked="" type="checkbox"/>	356	18 XP_006638532.1 PREDICTED: forkhead box protein G1-like (FoxG3)	[Lepisosteus oculatus]
<input checked="" type="checkbox"/>	341	19 NP_998079.2 forkhead box protein G1b (FoxG3)	[Danio rerio]

数字を入力し、
保持する配列の長さの
範囲を指定する
Limit with lengthボタンで
合致する配列のみ残す

アミノ酸残基数の大小に
併せて長方形の背景色の
濃淡が変わる。他と比べ
長い・短い配列を見つけやすい

配列の絞り込み方法（2）—CD-HITで冗長性を除く

左側の操作パネルに「CD-HIT remove」ボタンがあります。これを押すと、確認ダイアログが表示されます。

確認ダイアログの内容は「mafft.cbrc.jp」と「Apply CD-HIT to the 25 loaded sequences?」です。

確認ダイアログの「OK」ボタンを押すと、右側の「Run CD-HIT」ダイアログが表示されます。

「Input data:」欄には「25 sequences from .2308111344140340」が表示されています。

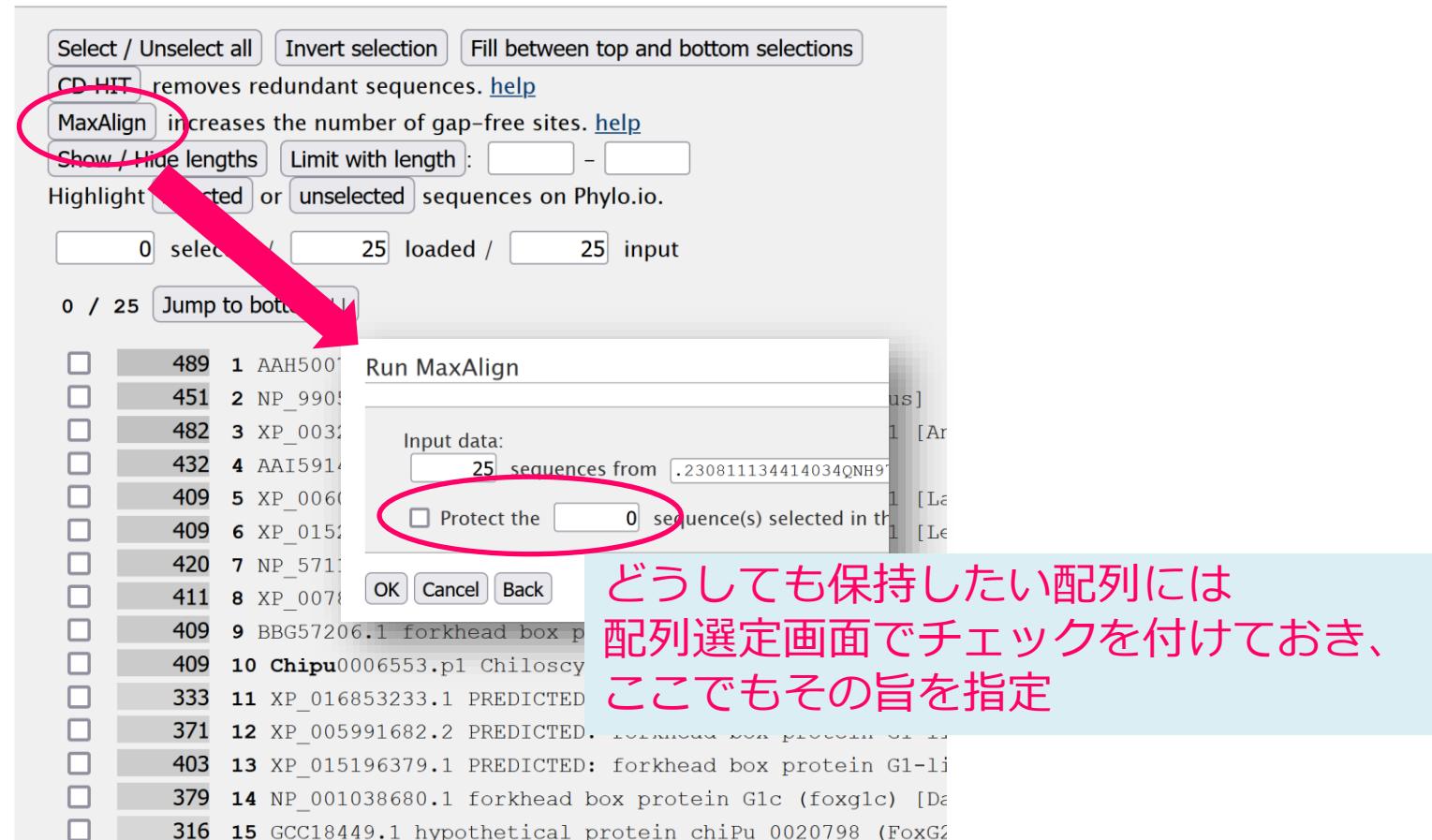
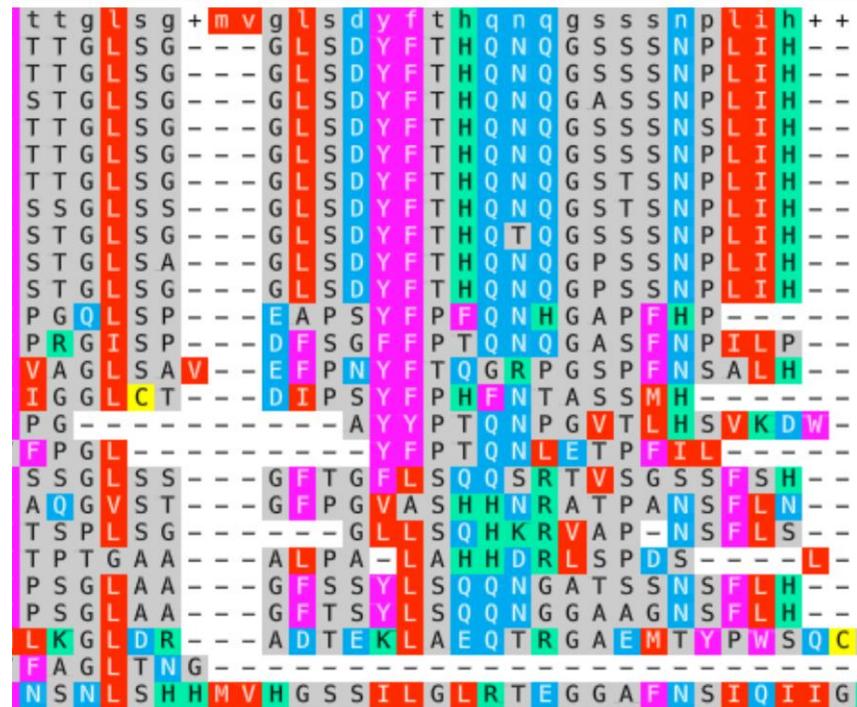
「Threshold to exclude:」欄には「100 % identical」が設定されています。この値を変更する場合は、赤い丸で囲まれた部分を修正できます。

赤い矢印が示す通り、チェックマークが付いている配列を対象として、基準以上に似ている配列を除く機能です。

ここで基準を設定（100%なら、完全一致の配列が除かれ、うち1本のみが保持される）

配列の絞り込み方法（3）—Max-Alignで部分配列を除く

この下から2つ目のような配列が含まれていると、
系統樹推定に利用できる（ギャップ = 欠失部分のない）
領域が減ってしまうため、除きたい。



どうしても保持したい配列には
配列選定画面でチェックを付けておき、
ここでもその旨を指定

配列の絞り込み方法（4）一生物種（群）で絞る

Select / Unselect all Invert selection Fill between top and bottom selections

CD-HIT removes redundant sequences. [help](#)

MaxAlign increases the number of gap-free sites. [help](#)

Show / Hide lengths Limit with length : 311 - 459

Highlight selected or unselected sequences on Phylo.io.

Select / Unselect representative organisms:

CALMI **HOMSA** **ORYLA** **LEPOC**
NOS **SAR** **SAU** **ACT**

555 selected / 633 loaded / 633

0 / 633 Jump to bottom ↓

<input checked="" type="checkbox"/> 354 1 Rhincodon NOS QQP64014.1 rhodopsin [Rhincodon typus]
<input checked="" type="checkbox"/> 311 2 NOS XP_020378075.1 rhodopsin [Rhincodon typus]
<input checked="" type="checkbox"/> 354 3 NOS QGW08841.1 rhodopsin [Chiloscyllium punctatum]
<input checked="" type="checkbox"/> 345 5 NOS AFS63881.1 rhodopsin, partial [Orectolobus maculatus]
<input checked="" type="checkbox"/> 331 292 SAU ATY46668.1 rhodopsin 1, partial [Oxyrhopuss guibei]
<input checked="" type="checkbox"/> 353 193 ORNAN ENSDANP0000000005991 2 non primary assembly:mOrnAnal.p.v1:X1:5
<input checked="" type="checkbox"/> 348 194 MONDO EN
<input checked="" type="checkbox"/> 348 195 HOMSA NP
<input checked="" type="checkbox"/> 348 196 SARHA EN
<input checked="" type="checkbox"/> 352 371 SPHPU EN
<input checked="" type="checkbox"/> 354 167 ACT XP_0
<input checked="" type="checkbox"/> 354 168 ACT BBH55931.1 rhodopsin [Polypterus senegalus]
<input checked="" type="checkbox"/> 354 372 ACT XP_036386254.1 rhodopsin, like [Megalops cyprinoides]

集められた配列全体に含まれるIDがここに羅列されており、クリックすることで、特定のIDに紐づいた配列のみを選択・除外することが可能

aLeaves - [əlī:vz]

first step to build zoologically informative phylogenetic trees

Top Database Species History Help About Links

Species included

A. Species ID

The table below lists all species based on genome-wide sequence resources (not NCBI Protein or GenBank) covered at aLeaves. Please see the 'Database' page for a list of the sequences of individual species. Species newly added in the latest aLeaves release are indicated with *. If you proceed to the MAFFT server for data set refinement and inference, these Species IDs will be automatically recognized, and the sequences colored accordingly.

Species ID	Latin species name	English common name	Category
HOMSA	<i>Homo sapiens</i>	human	Mammals
AILME	<i>Ailuropoda melanoleuca</i>	giant panda	Mammals

配列の絞り込み方法（5）—Guide treeによる選定

History

Step 1: 1001 unaligned sequences → Aligned, 15:12:35

Step 2: 633 sequences × 228

Refine dataset

Open guide tree on Phylo.io and click slow.

Apply the sequence selection below

- Realign the selected sequences. Reverse complementary sequences can be generated.
- Clip the selected sequences and rebuild a tree (alignment→kept; numbering→kept)
- Clip the selected sequences (alignment→kept; numbering→kept)

Select / Unselect all Invert selection Fill between top and bottom selections

CD-HIT removes redundant sequences. [help](#)

MaxAlign increases the number of gap-free sites. [help](#)

Show / Hide lengths Limit with length : 311 - 459

Highlight selected or unselected sequences on Phylo.io.

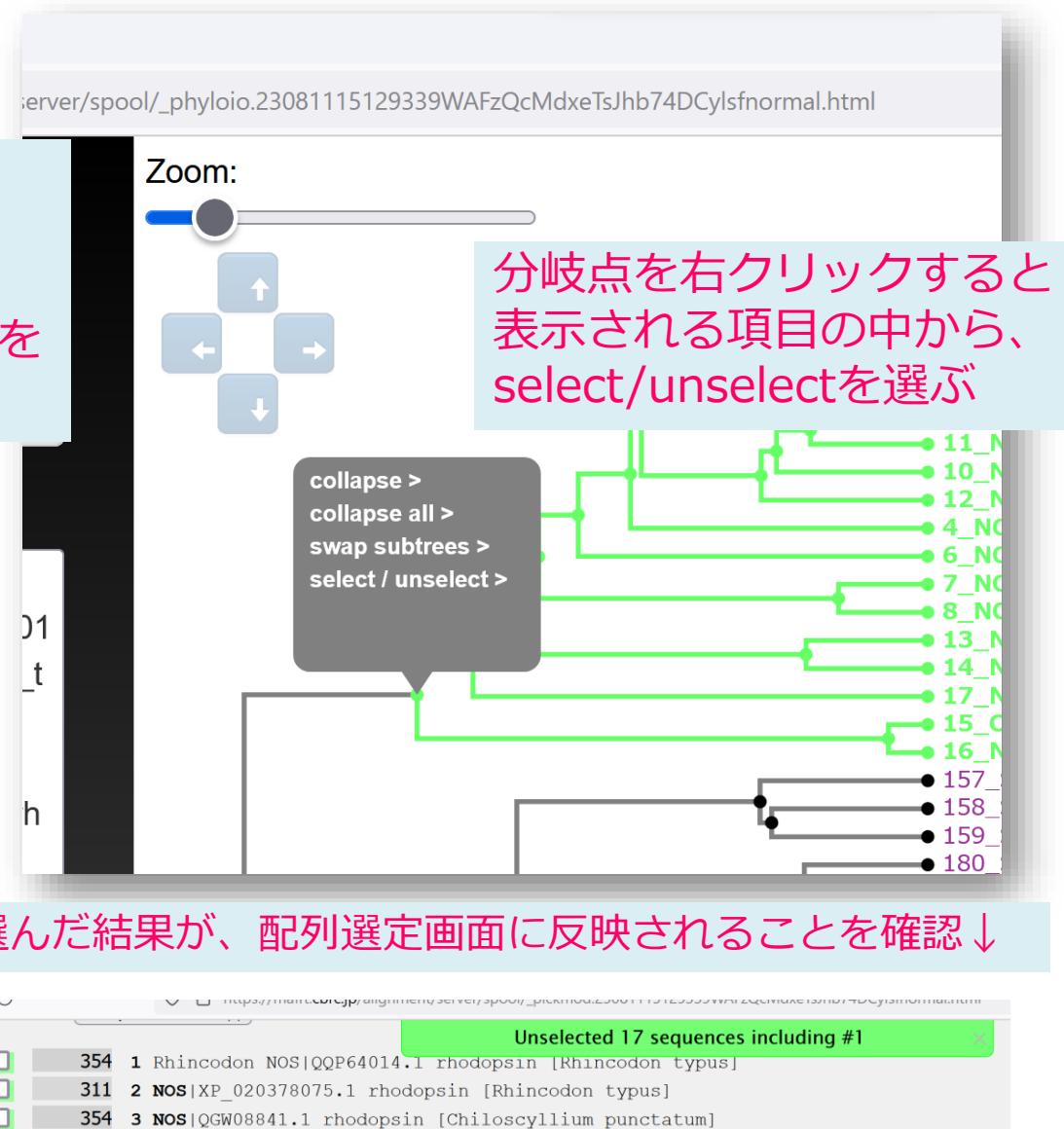
Select / Unselect representative organisms:

[CALMI](#) [HOMSA](#) [ORYLA](#) [LEPOC](#)

[NOS](#) [SAR](#) [SAU](#) [ACT](#)

555 selected / 633 loaded / 633 input

配列選定画面のこのボタンを押すことで、簡易系統樹を表示できる。樹形の全体像を見ながら、特定の分岐点の内側の配列を選定・除外することが可能。



配列の絞り込み過程で役立つ機能（1）

Select / Unselect all Invert selection Fill between top and bottom selections

CD-HIT removes redundant sequences. [help](#)

MaxAlign increases the number of gap-free sites. [help](#)

Show / Hide lengths Limit with length : 311 - 459

Highlight selected or unselected sequences on Phylo.io.

Select / Unselect representative organisms:

CALMI HOMSA ORYLA LEPOC
NOS SAR SAU ACT

555 selected / 633 loaded / 633 input

0 / 633 Jump to bottom ↓

<input checked="" type="checkbox"/>	354	1 Rhincodon NOS QQP64014.1 rhodopsin [Rhincodon typus]
<input checked="" type="checkbox"/>	311	2 NOS XP_020378075.1 rhodopsin [Rhincodon typus]
<input checked="" type="checkbox"/>	354	3 NOS QGW08841.1 rhodopsin [Chiloscyllium punctatum]
<input checked="" type="checkbox"/>	345	5 NOS AFS63881.1 rhodopsin, partial [Orectolobus maculatus]
<input checked="" type="checkbox"/>	354	9 NOS QGW08840.1 rhodopsin [Neotrygon kuhlii]
<input checked="" type="checkbox"/>	354	11 NOS QTX16876.1 rhodopsin-like-1 rod opsin [Potamotrygon motoro]
<input checked="" type="checkbox"/>	354	10 NOS QGW08839.1 rhodopsin [Aptychotrema vincentiana]
<input checked="" type="checkbox"/>	354	12 NOS QGW08843.1 rhodopsin [Glaucostegus typus]
<input checked="" type="checkbox"/>	354	4 NOS CAA76707.1 opsin [Gymnophionae sp.]

上段の3つのボタン、左から

- ・全部の配列を選択・削除
- ・選択・削除を逆にする
- ・選んだ2つの間の配列を選択
(別途右に図示)

Before		After
<input type="checkbox"/>	354	1 Rhincodon
<input type="checkbox"/>	311	2 NOS XP_0
<input type="checkbox"/>	354	3 NOS QGW0
<input checked="" type="checkbox"/>	345	5 NOS AFS6
<input type="checkbox"/>	354	9 NOS QGW0
<input type="checkbox"/>	354	11 NOS QTX
<input type="checkbox"/>	354	10 NOS QGW
<input type="checkbox"/>	354	12 NOS QGW
<input type="checkbox"/>	354	4 NOS CAA7
これらも 選ばれる		1 NOS XP_0 1 NOS QGW0
<input type="checkbox"/>	354	8 NOS QGW0
<input type="checkbox"/>	354	13 NOS XP_0
<input type="checkbox"/>	354	14 NOS AAC
<input type="checkbox"/>	354	17 NOS CAA
<input type="checkbox"/>	354	15 CALMI E
<input checked="" type="checkbox"/>	354	16 NOS ABU
<input type="checkbox"/>	354	157 SAR XP
<input type="checkbox"/>	354	158 SAR XP

配列の絞り込み過程で役立つ機能（2）

The screenshot shows a web-based sequence alignment tool interface. At the top, there's a navigation bar with links like 'Clustal format', 'Fasta format', 'MAFFT result', 'View', 'Tree', 'Refine dataset', and 'Return to home'. Below this is a 'History' section listing four steps of sequence clipping:

- Step 1: 1001 unaligned sequences → Aligned, 15:12:35
- Step 2: 633 sequences × 228 gap-free sites → Clipped, 09:20:43
- Step 3: 536 sequences × 305 gap-free sites → Clipped, 09:21:23
- Step 4: 424 sequences × 315 gap-free sites →

Below the history is a 'Refine dataset' section with instructions to open a guide tree on Phylo.io and apply sequence selection. It includes three radio button options:

- Realign the selected sequences. Reverse complementary sequences can be generated in the case
- Clip the selected sequences and rebuild a tree (alignment→kept; numbering→reset)
- Clip the selected sequences (alignment→kept; numbering→kept)

At the bottom, there are various filtering and selection tools:

- Select / Unselect all
- Invert selection
- Fill between top and bottom selections
- CD-HIT removes redundant sequences. [help](#)
- MaxAlign increases the number of gap-free sites. [help](#)
- Show / Hide lengths
- Limit with length : -
- Highlight selected or unselected sequences on Phylo.io.
- Select / Unselect representative organisms:
CALMI HOMSA ORYLA
NOS ACT

Status indicators at the bottom show: 0 selected / 424 loaded / 424 input.

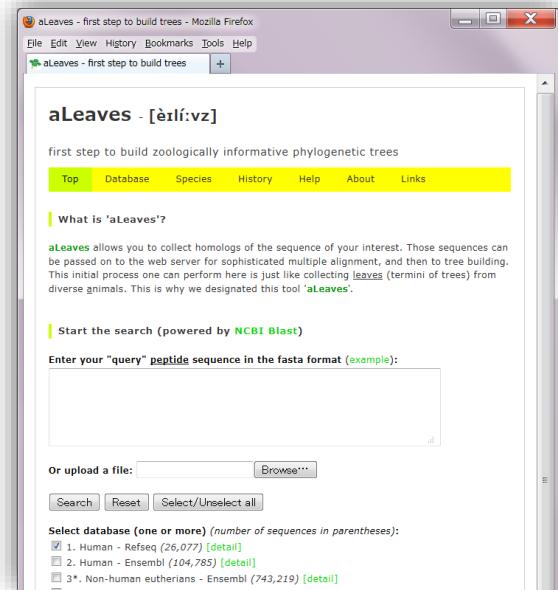
左から

- 配列群をClustalフォーマットのファイルとして取得
- 配列群をFASTA形式のファイルとして取得
- アラインメント結果のテキスト表示
- アラインメント結果のグラフィカル表示
- 系統樹推定（近隣結合法など）へ

配列選定画面の上段のヒストリから、
クリックすると、前の段階に戻ることが可能。
除きたくない配列を除いてしまったことに
気づいた際などに便利。

aLeaves-MAFFTで実現するワークフロー

Exhaustive collection of homologs

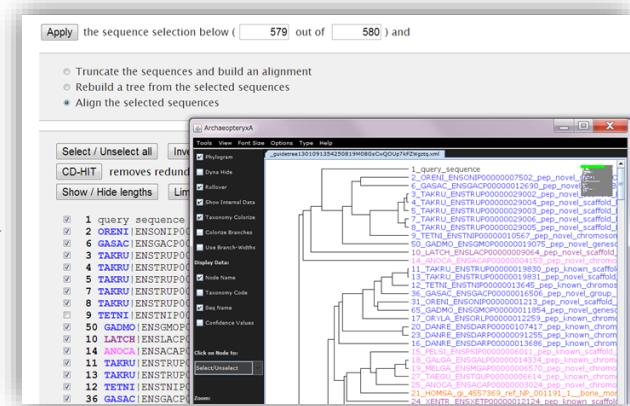


(on this aLeaves server at CDB, RIKEN)

athostomes, we identified three paralogs in some species, which were designated as *RAK1* (AB292622–AB292624) in the genomes *RAK1* and *RAK2* clustered with *Ry* and *RAR α* , respectively (fig. 5C). A

Heuristic identification of homologs
(in publications, etc.)

Careful refinement of dataset
by deleting unnecessary sequences



(on MAFFT server at CBRC, AIST)

- 取り漏らしの起きにくい配列収集プロセス
- ソフトウェアのインストールは不要

Phylogenetic tree inference

A search interface with a dropdown menu set to "Nucleotide", a search input field containing "AB292622", and a "Search" button.

Retrieval of limited number of sequences

注意点

- ・ **aLeaves** は他所すでに公開されているリソースだけに基づいたツールです（オリジナルサイトへリンク）
- ・ **aLeaves**プロジェクトとしては、独自にタンパクコード遺伝子の推定およびその吟味を行わず、オリジナルリソースの情報をそのまま踏襲しています
- ・ **aLeaves**-MAFFTサーバを利用し予備的な分子系統解析が可能だが、本格的な解析は、作成した配列セットをダウンロードし、やはりローカルで行ってください

Perspectives

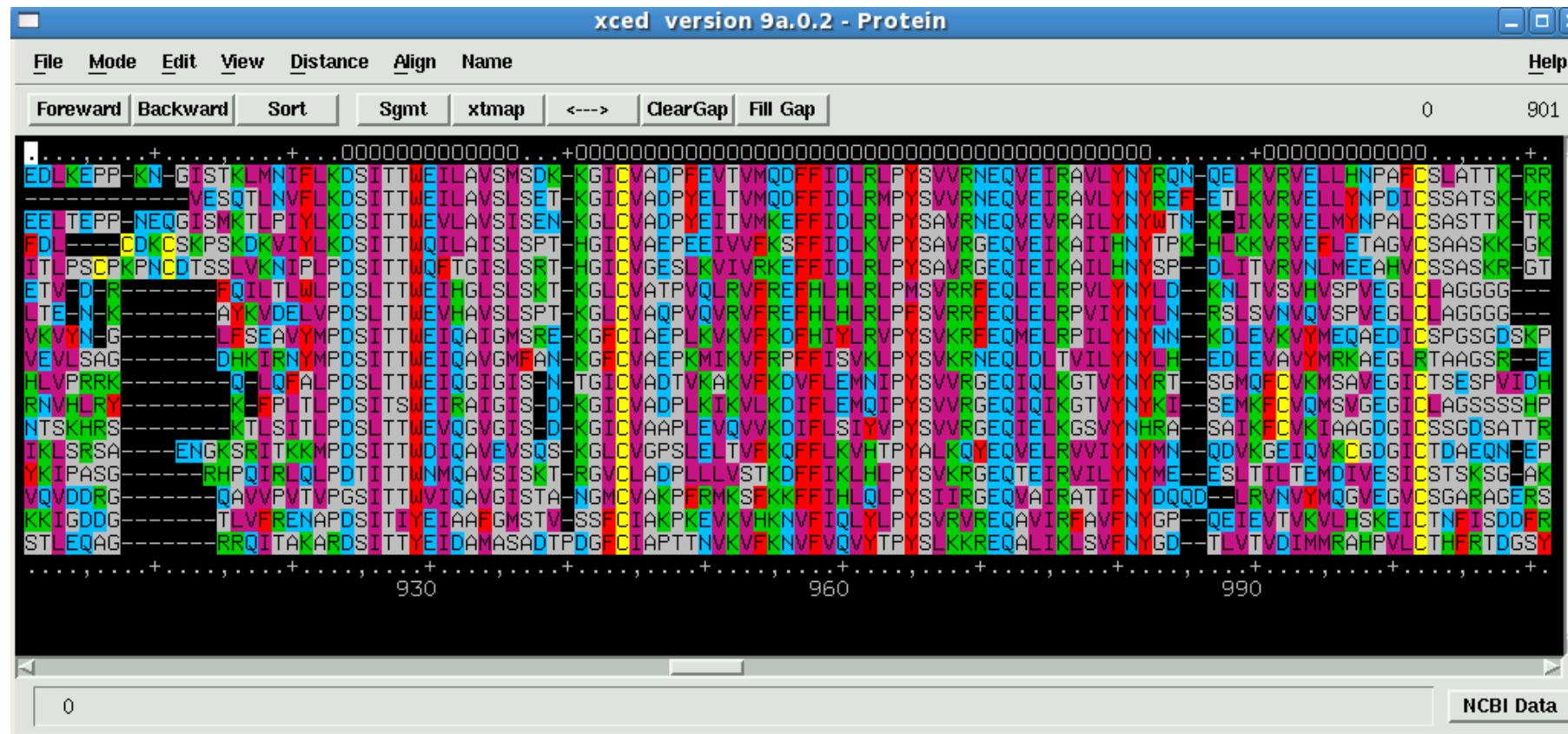
機能拡張の可能性

- ・塩基配列への対応
- ・遺伝子配列だけでなくゲノムアセンブリに対応
- ・動物以外の生物群に対応

サポート

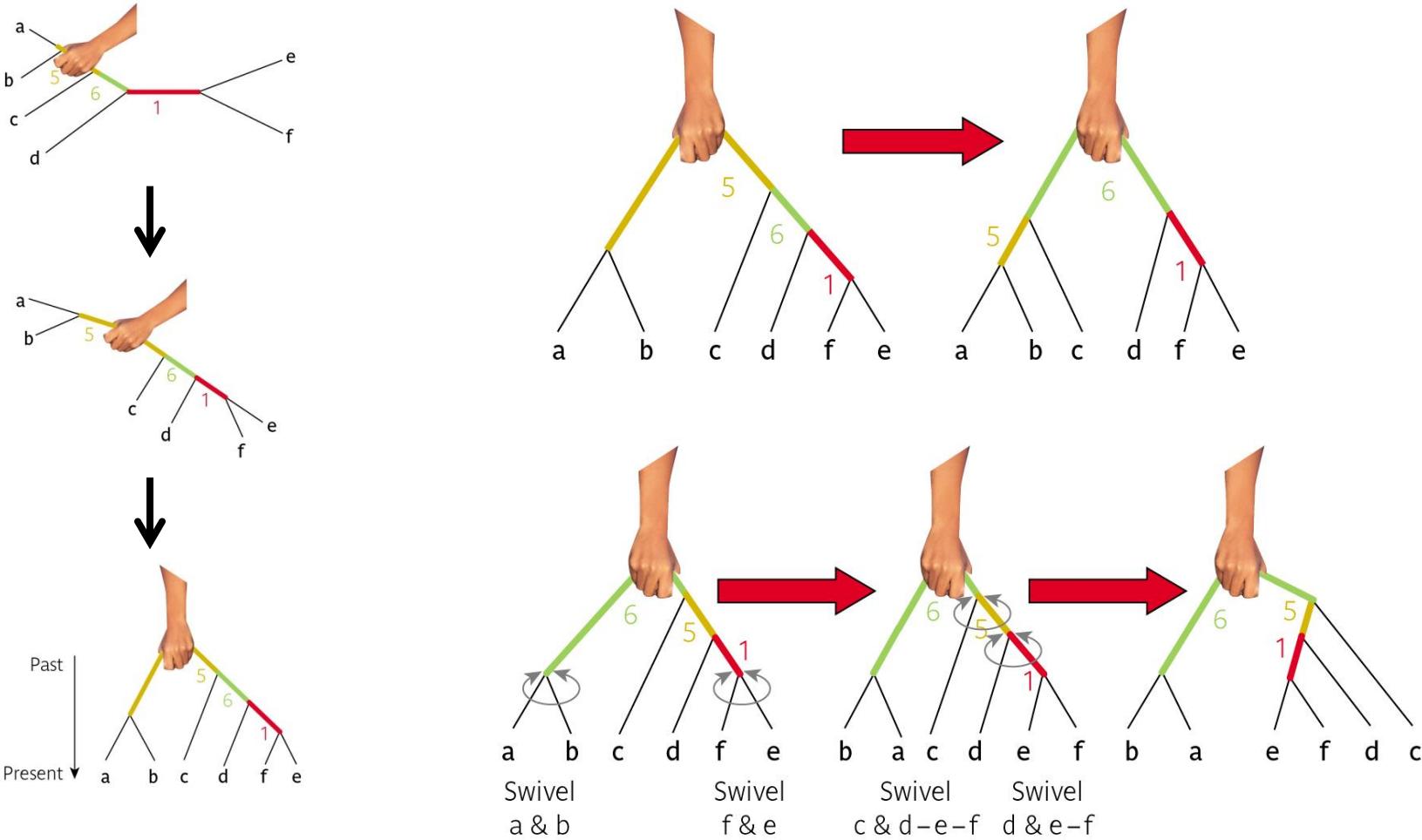
- ・特定の生物種の追加などのリクエストにも可能な限り対応
- ・質問等は skuraku@nig.ac.jp へ
- ・分子系統解析についての一般的なご相談もぜひどうぞ

配列の多重アライメント (multiple alignment)



「相同な」塩基・あるいはアミノ酸座位を見つける作業

Rerooting (根の付け替え) とは



こういった変形を加えても、実質的な樹形（トポロジー）は変わらない

- ・オーソログ（直系遺伝子）=種分岐で分かれた遺伝子
 - ・パラログ（側系遺伝子）=遺伝子重複で分かれた遺伝子
-

注意：「オーソログ=機能が同じ遺伝子」ではない

- 理由 1 遺伝子の機能も進化の過程で変化するから
したがって、オーソロガスな関係にある遺伝子も異なる機能を持ちうるから
- 理由 2 特定の分子機能が、共通の起源をもたない（ホモログでない）
2つ以上の遺伝子によって担われることがあり、そもそも
ホモログでない遺伝子間の関係は、オーソログやパラログ
という言葉で形容されないため

(何を「機能」と呼ぶのか、ということも深くて複雑な問題である)

派生語

形容詞	オーソロガス orthologous = 「オーソログの関係にある」という 遺伝子の関係を形容する語	パラロガス paralogous = 「オーソログの関係にある」という 遺伝子の関係を形容する語
名詞	オーソログ ortholog = オーソロガスな <u>遺伝子</u>	パラログ paralog = パラロガスな <u>遺伝子</u>
名詞	オーソロジー orthology = オーソロガスな <u>関係のこと</u>	パラロジー paralogy = パラロガスな <u>関係のこと</u>

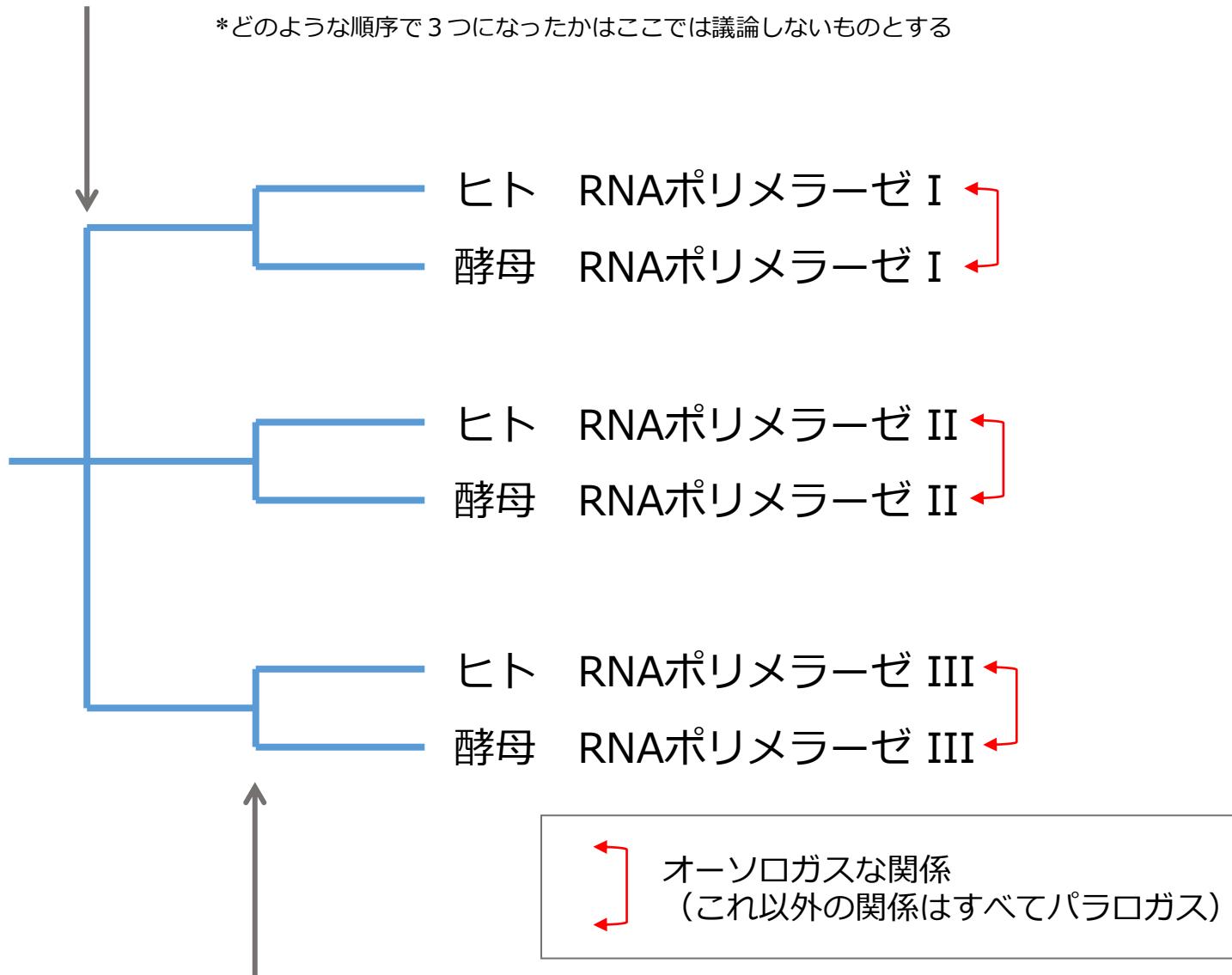
系統 lineage = 進化の結果としての生物の遺伝的なつながり

遺伝子重複 gene duplication = 進化の過程で遺伝子のコピーがつくられること

種分岐 speciation = 生物の系統が枝分かれすること

遺伝子重複（ひとつの祖先遺伝子が3つに増えた）

*どのような順序で3つになったかはここでは議論しないものとする



ヒト（動物）の祖先と酵母（菌類）の祖先の種分岐

オーソログ・パラログを判定する手がかり

遺伝子が「分かれる」イベントは「遺伝子重複」と「種分岐」のどちらかである

→2つの遺伝子同士の関係はオーソロガスかパラロガスの常にどちらかである

→「オーソロガスでない関係」はすべてパラロガスである

同じ生物のゲノム内の互いに相同な遺伝子同士の関係は常にパラロガス

→遺伝子重複による産物であるから

オーソログは常に1：1の数の関係にあるとは限らない

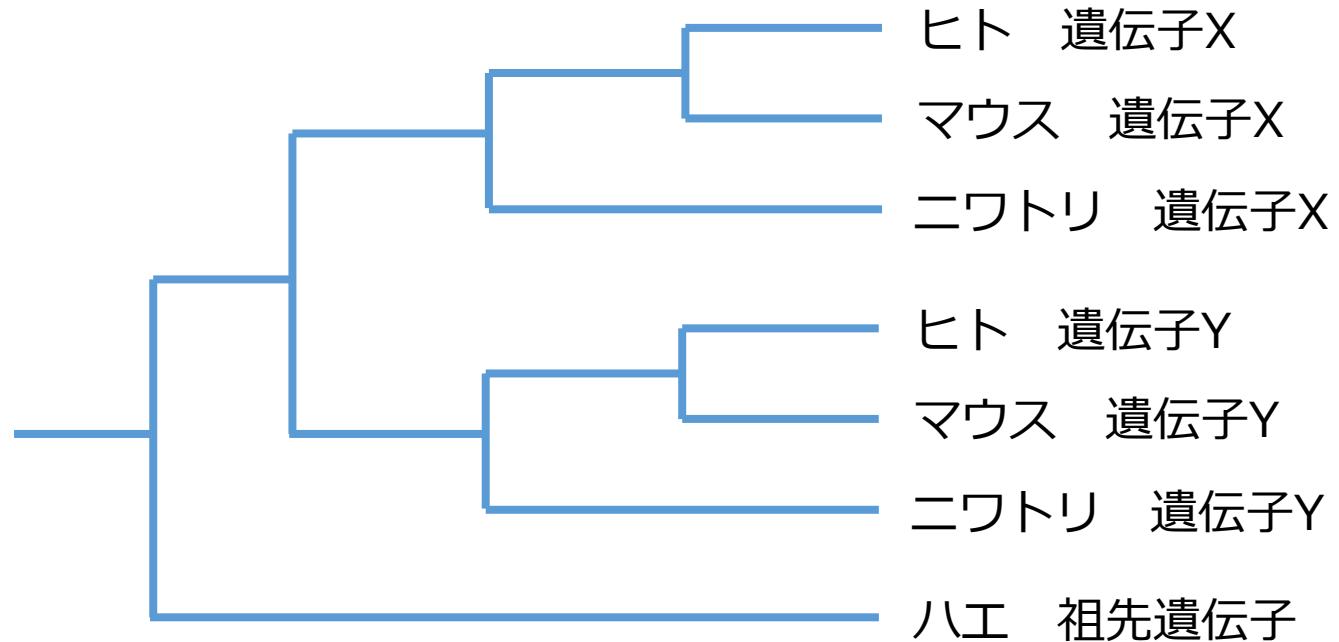
→系統特異的な遺伝子重複が起きることがあるから

注意：系統特異的重複が起きたとしても、オーソログが「増える」だけで
あって、オーソログでありつづけることに変わりはない

いくつか例を

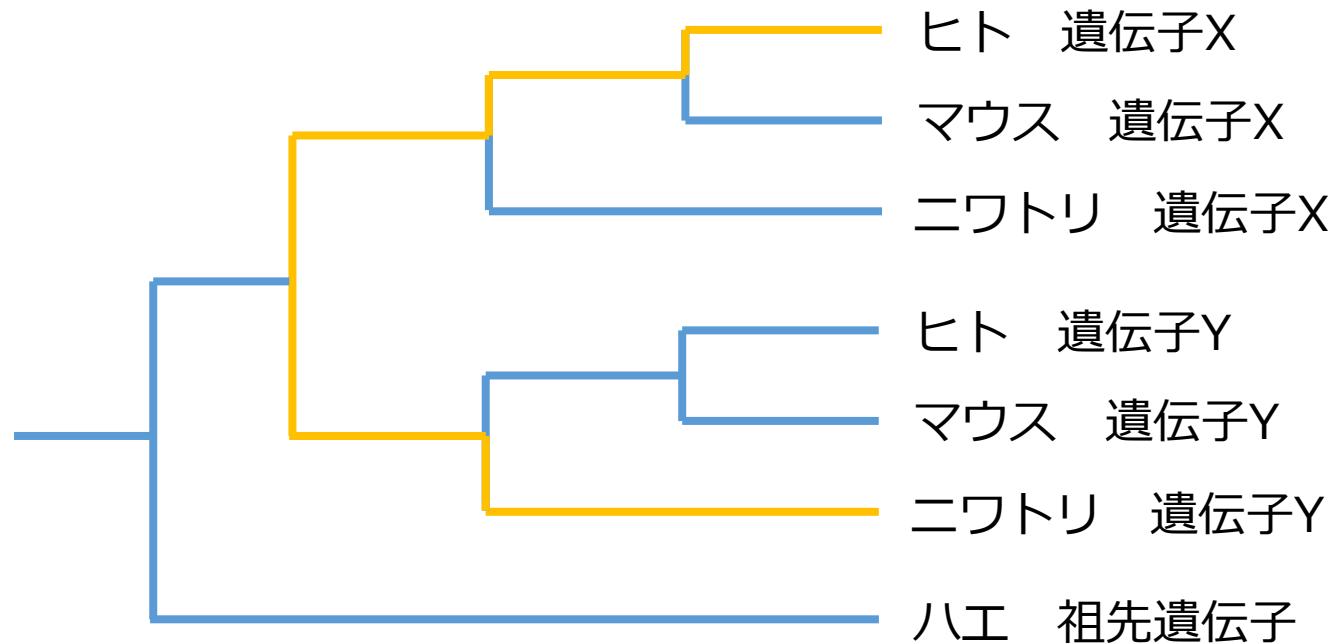
オーソログ・パラログを判定するための究極的な方法

例題：ヒト 遺伝子Xとニワトリ遺伝子Yの関係は？



オーソログ・パラログを判定するための究極的な方法

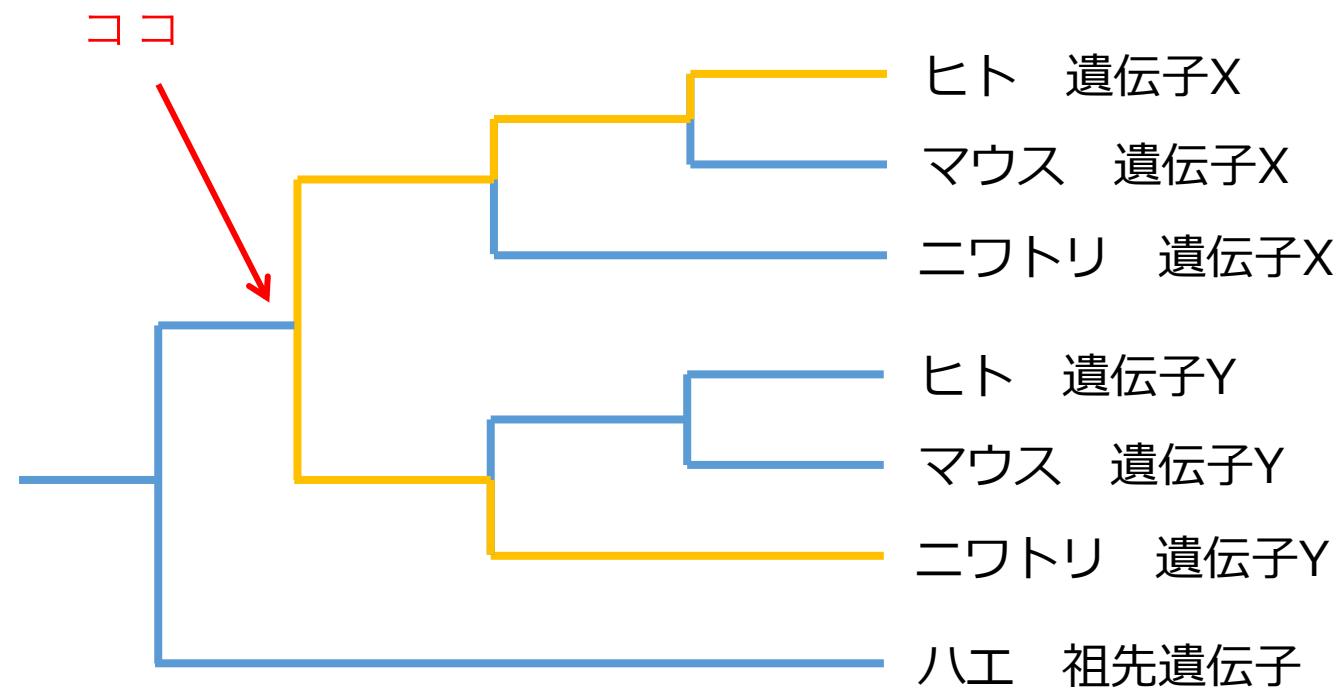
例題：ヒト 遺伝子Xとニワトリ遺伝子Yの関係は？



ステップ1 当該遺伝子の関係を線でたどる

オーソログ・パラログを判定するための究極的な方法

例題：ヒト 遺伝子Xとニワトリ遺伝子Yの関係は？

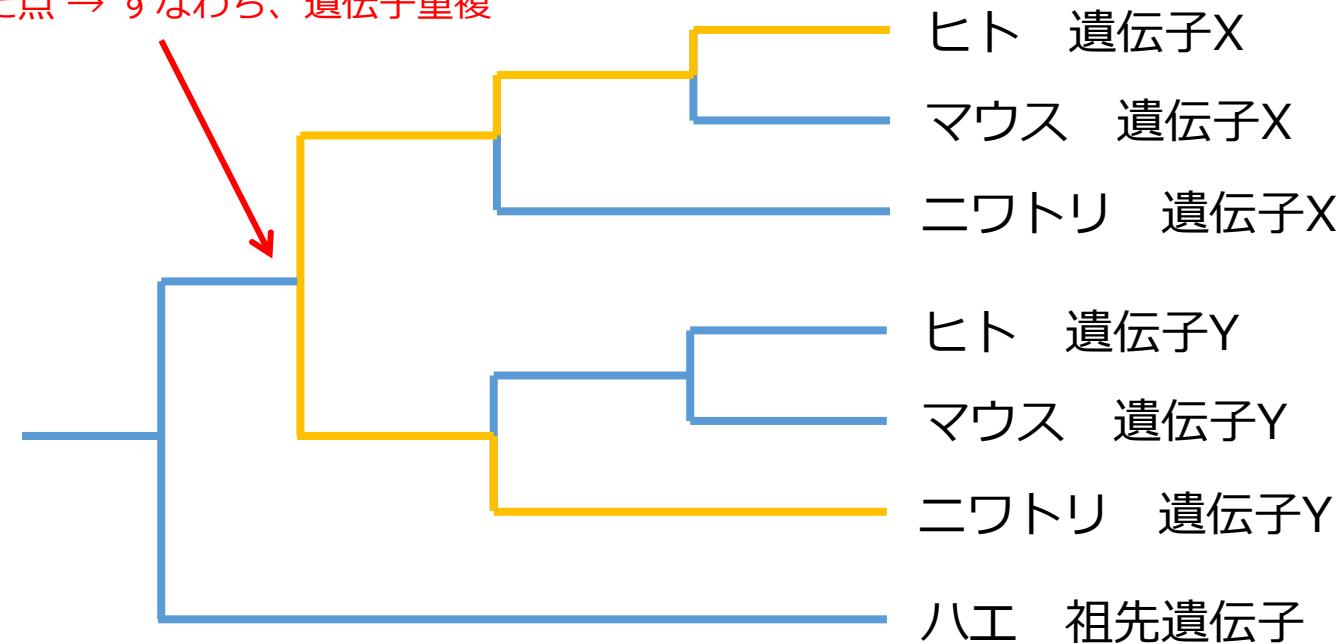


ステップ2 もっとも古い分岐点（通常はもっとも左側）が
「遺伝子重複」か「種分岐」かを判定する

オーソログ・パラログを判定するための究極的な方法

例題：ヒト 遺伝子Xとニワトリ遺伝子Yの関係は？

この場合、当分岐点が遺伝子Xと遺伝子Yが
分かれた点 → すなわち、遺伝子重複

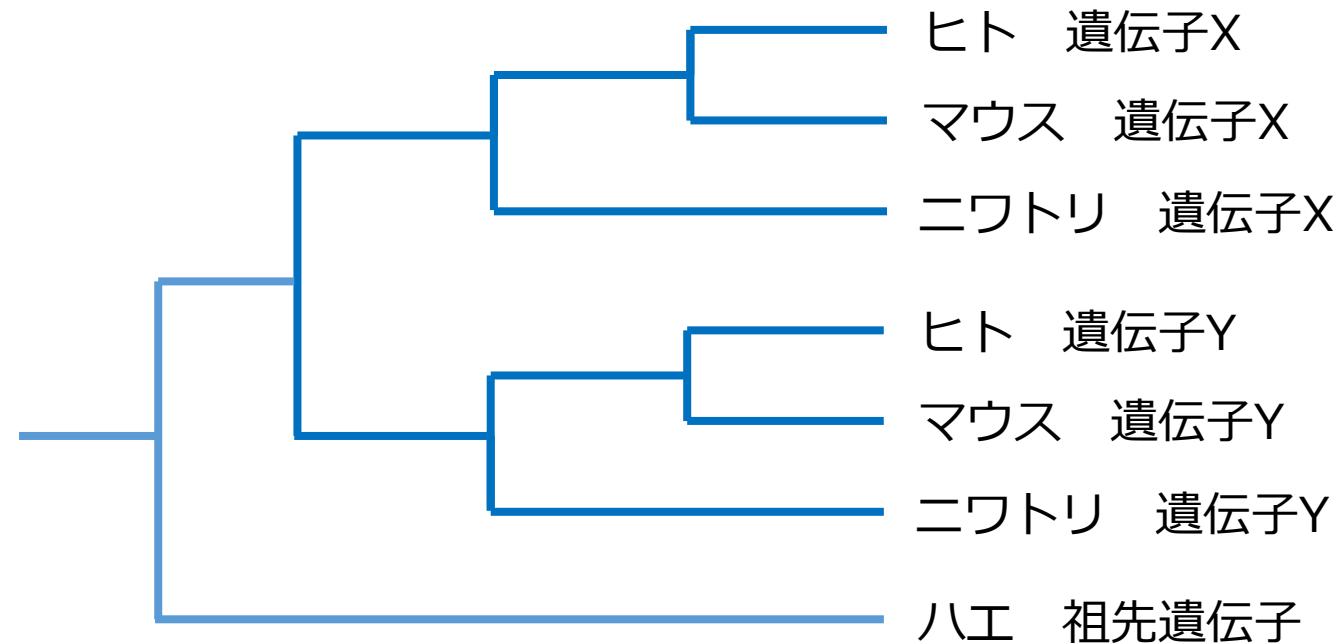


ステップ3 定義に従い、「種分岐」ならオーソロガス
「遺伝子重複」ならパラロガス と判定

(場合によっては、ここで生物の系統関係についての詳しい知識が必要)

練習問題 1

ヒト 遺伝子Xとニワトリ 遺伝子Xの関係は？

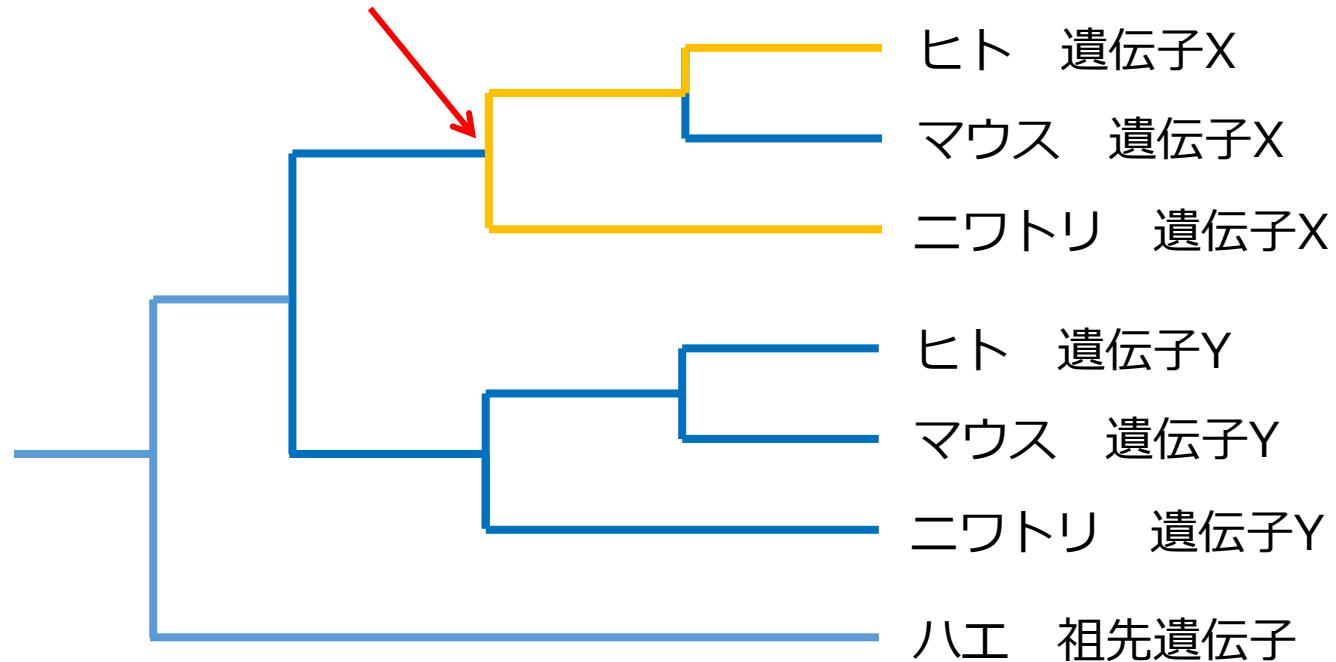


練習問題 1

ヒト 遺伝子Xとニワトリ遺伝子Xの関係は？

問題となるのはこの分岐点

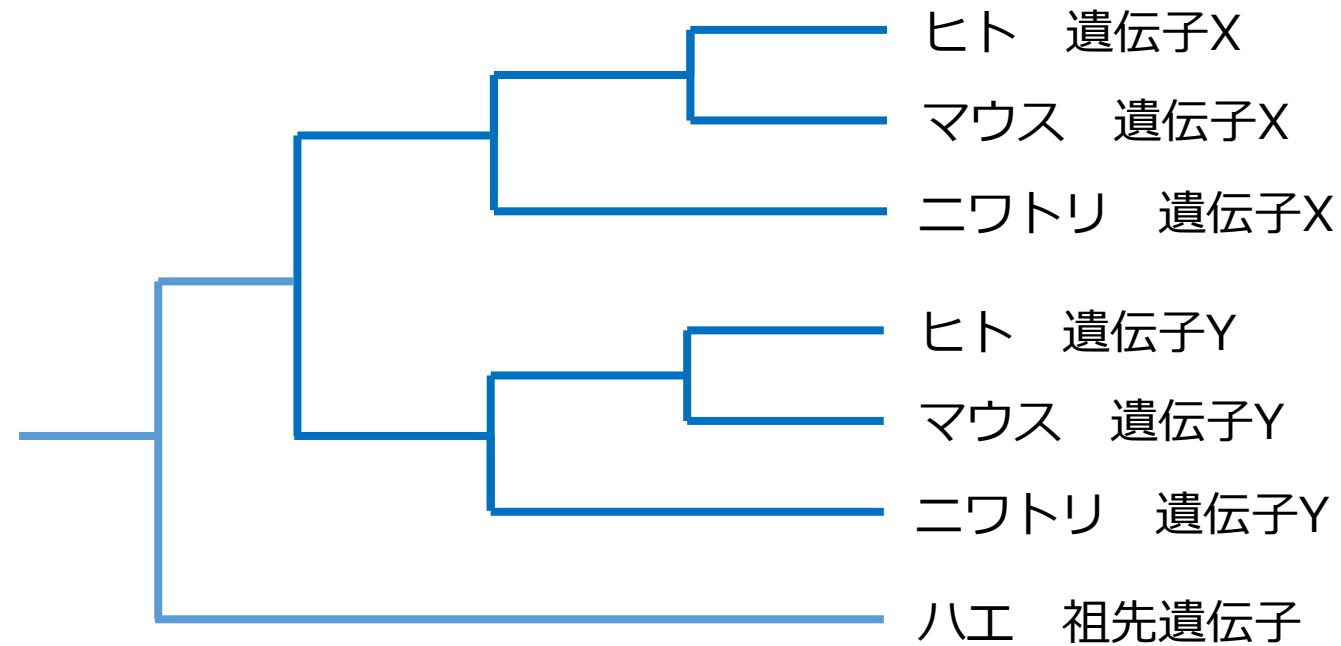
このグループ内は、ヒト、マウス、ニワトリ間で種の系統関係が成り立っており、
この分岐が遺伝子重複であると判断する理由はない、したがって、種分岐。



正解：オーソロガス

練習問題 2

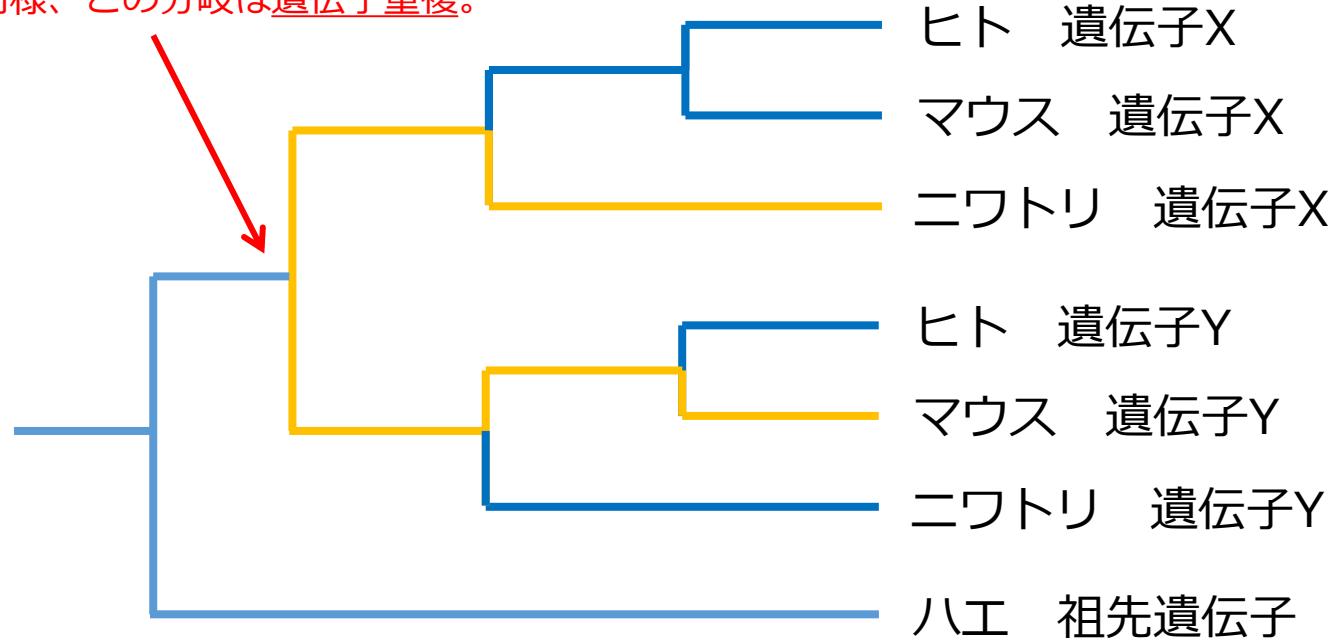
ニワトリ 遺伝子Xとマウス遺伝子Yの関係は？



練習問題 2

ニワトリ 遺伝子Xとマウス遺伝子Yの関係は？

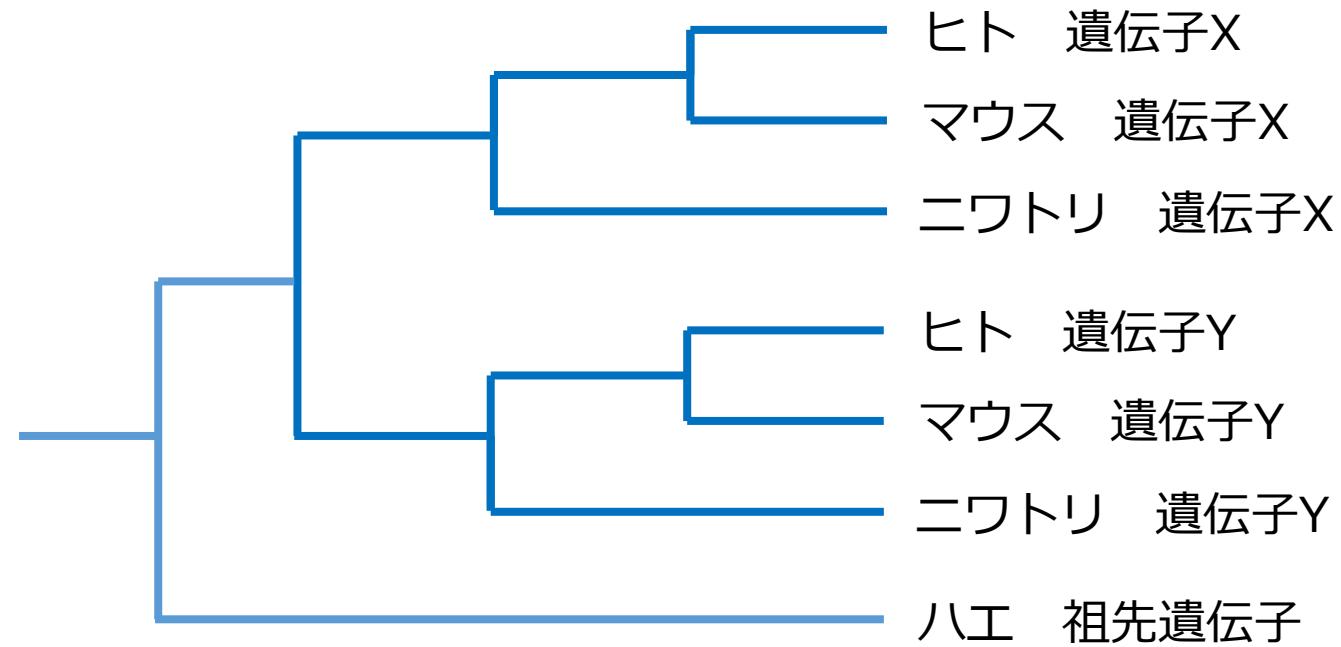
問題となるのは再びこの分岐点。
例題と同様、この分岐は遺伝子重複。



正解：パラ口ガス

練習問題 3

マウス 遺伝子Xと「ハエ祖先遺伝子」の関係は？

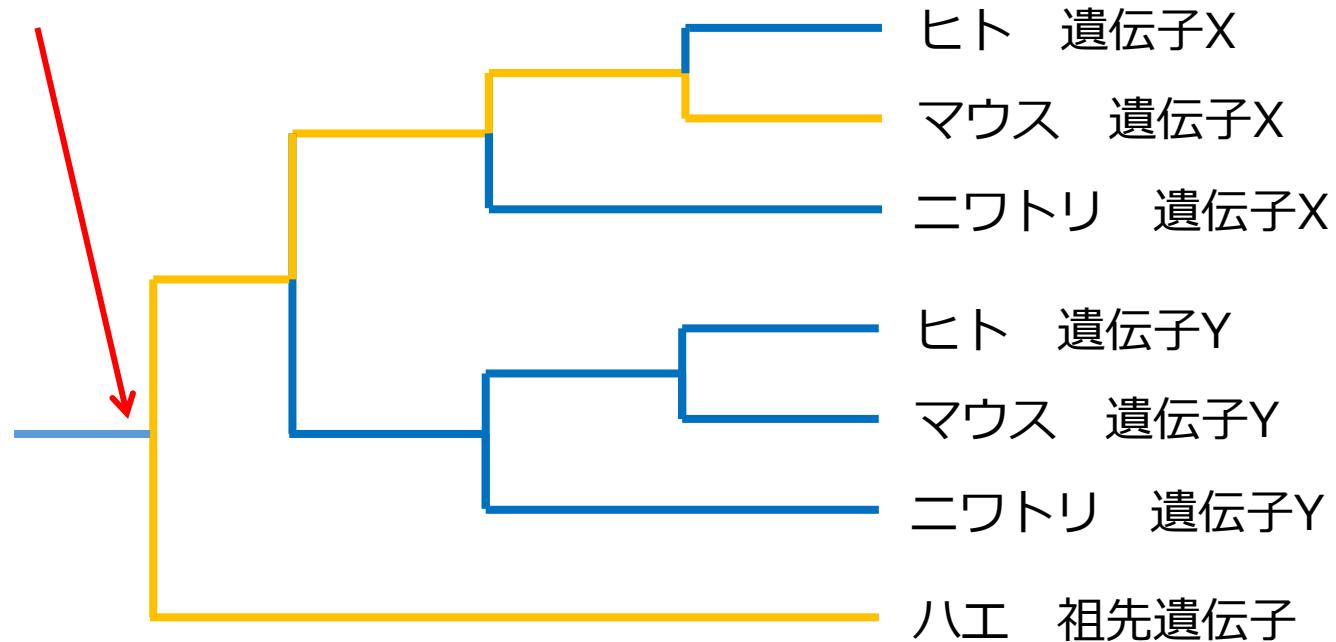


練習問題 3

マウス 遺伝子Xと「ハエ祖先遺伝子」の関係は？

問題となるのはこの分岐点。

ここは、脊椎動物と節足動物の系統が分かれた種分岐とみるのが妥当



正解：オーソロガス

まず、問題となる分岐点が「種分岐」なのか「遺伝子重複」なのかを判定するために、正しい生物種の系統関係を把握しておく必要がある。

最後に、軟骨魚類のロドプシンについての話題を

プレスリリース

2023年3月23日

本論文はすでに公開されています
本情報はすぐにご利用いただけます

ジンベエザメだけに起きた視覚の進化 ～深海生活への適応か？～

■ 概要

動物は、生息環境中の光を活用するしくみをもっています。なかでも、深海や地中など微弱な光環境で暮らす生物種は、その限られた光を活用する特別なしくみを備えています。最大の魚類であるジンベエザメは、海面近くで餌食をする一方で、光の届きにくい深海にも潜ることが知られています。

今回、情報・システム研究機構 国立遺伝学研究所の工楽樹洋教授(理化学研究所 チームリーダー)、大阪公立大学の小柳光正教授が率いる研究チームは、ジンベエザメが光を活用するしくみの解明に挑みました。

本研究チームは、眼の網膜にある光受容タンパク質(オプシン⁽¹⁾)のうち、微弱光下での視覚をつかさどるロドプシンについて、DNA情報と吸収する光の波長を測る分光測定を組み合わせることで、ジンベエザメと他のサメ類を比較しました(図1)。その結果、ジンベエザメのロドプシンは、従来の常識を覆すアミノ酸置換によって、深海の中で最も届きやすい青色の光を効率的に受け取ることができることを明らかにしました。また、ジンベエザメのロドプシンは熱に弱く、低温となる深海での機能に適していることもわかりました。

本結果から、ジンベエザメの視覚は、水温が低下した深海において、微弱な光を活用できるような進化を遂げたことが示唆されます。この進化の引き金になったアミノ酸置換部位はヒトの夜盲症⁽²⁾の原因となる置換部位でもあり、ジンベエザメの海水面付近から水深2000メートル付近の深海まで潜水するというその独特的なライフスタイルの表れともいえるかもしれません。

さらに、本研究は、様々な生物のくらしをその生体を犠牲にすることなく解明かすために、DNAの情報を活用したタンパク質の人工合成技術が有用であることを示しました。

本研究成果は、米国科学雑誌「Proceedings of the National Academy of Sciences of the United States of America (PNAS)」に2023年3月22日午前1時(日本時間)に短報論文として掲載されました。



海洋博公園・沖縄美ら海水族館提供

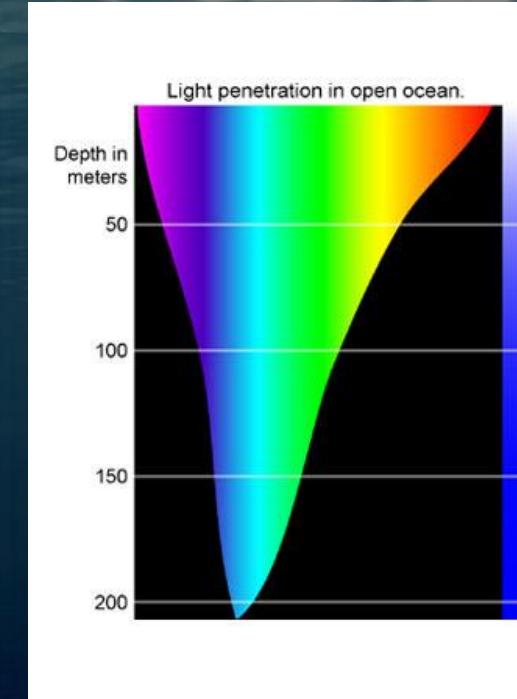
ジンベエザメの際立った特徴である
深海潜行に関する予備知識

深海性の硬骨魚の場合、
ロドプシン遺伝子を多数に増幅



Musilova, et al., 2019 Science

最深部に届くのは、約480nmの光

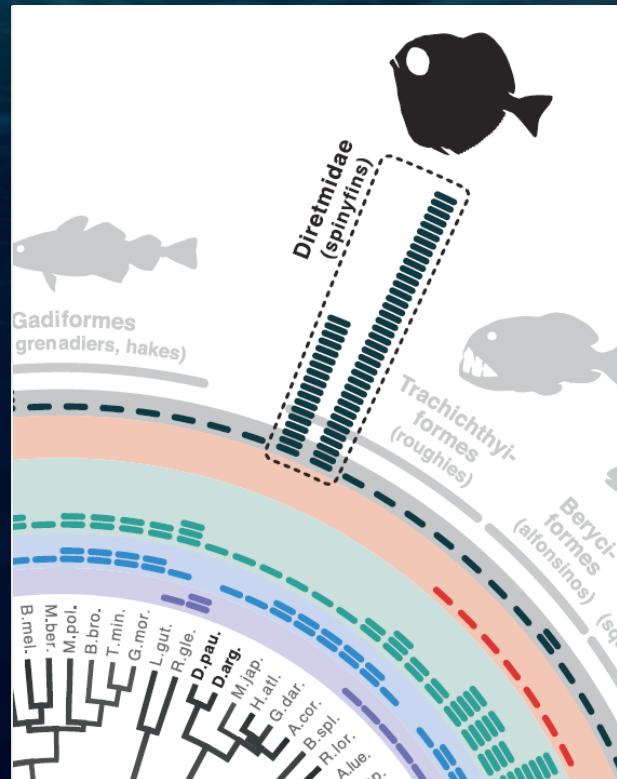


[http://oceanexplorer.noaa.gov/explorations/04deepscope/
background/deeplight/media/diagram3.html](http://oceanexplorer.noaa.gov/explorations/04deepscope/background/deeplight/media/diagram3.html)

全ゲノム情報に基づくサメ・エイ類のオプシン遺伝子群の解析は皆無であった

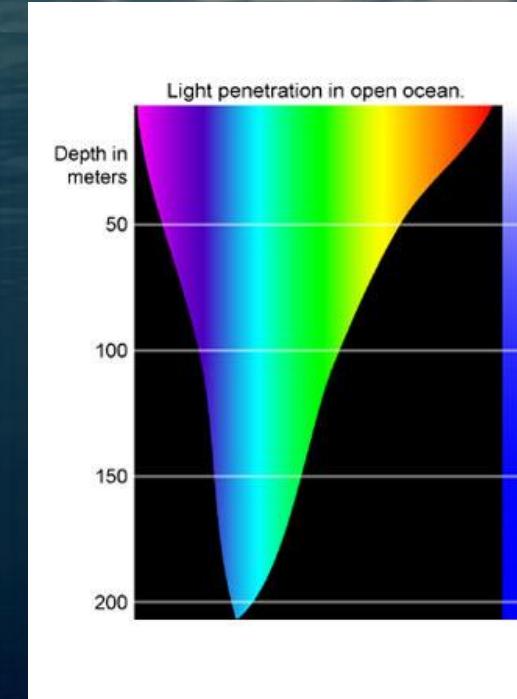
ジンベエザメの際立った特徴である 深海潜行に関する予備知識

深海性の硬骨魚の場合、
口ドプシン遺伝子を多数に増幅



Musilova, et al., 2019 Science

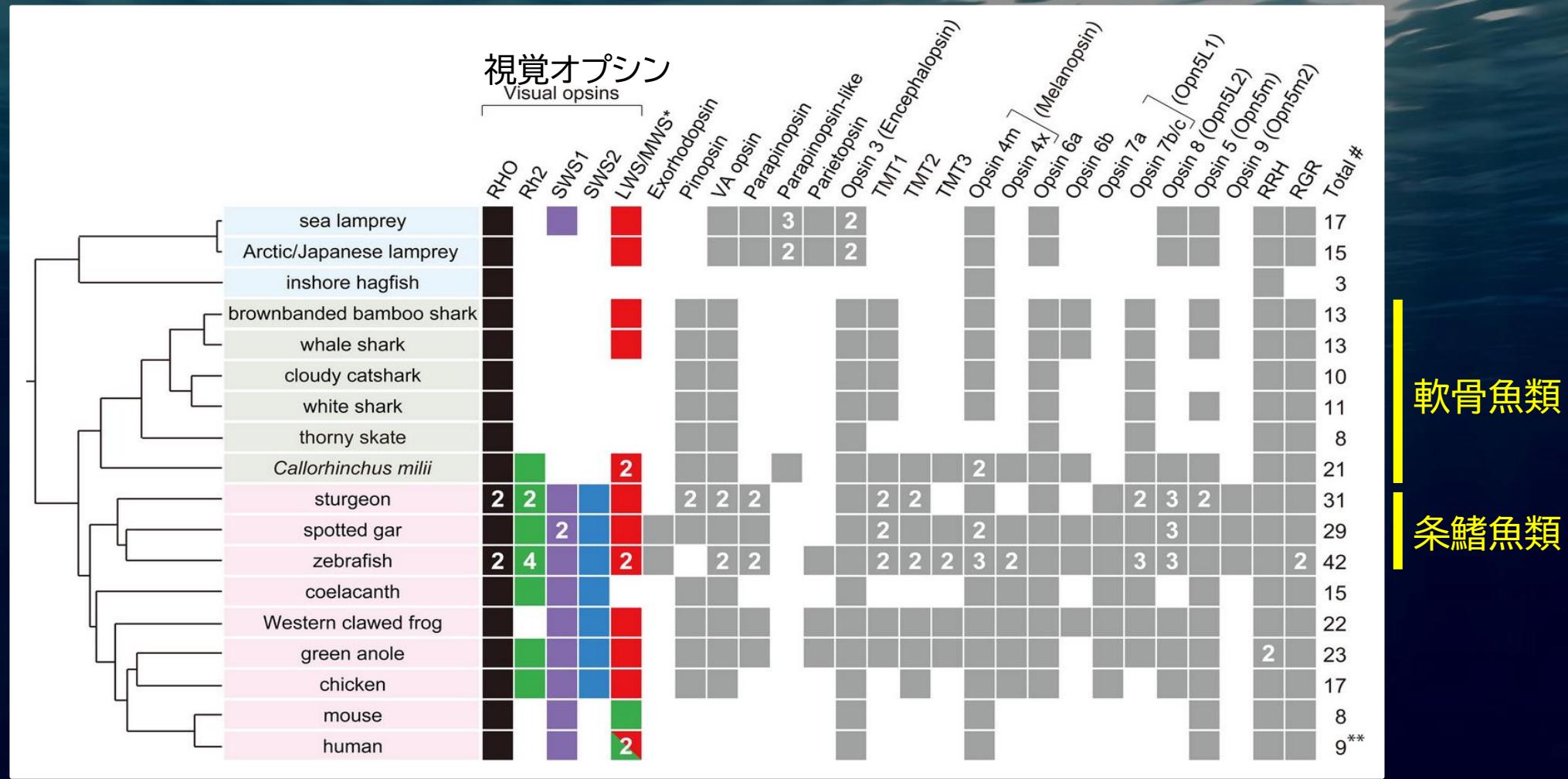
最深部に届くのは、約480nmの光



[http://oceanexplorer.noaa.gov/explorations/04deepscope/
background/deeplight/media/diagram3.html](http://oceanexplorer.noaa.gov/explorations/04deepscope/background/deeplight/media/diagram3.html)

全ゲノム情報に基づくサメ・エイ類のオプシン遺伝子群の解析は皆無であった

サメ類のオプシン遺伝子の探索



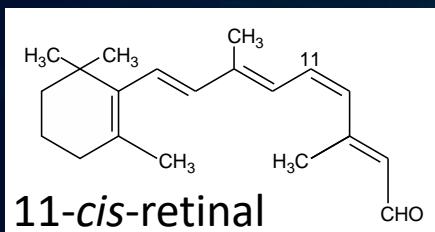
サメ類は、進化の過程で多くのオプシン遺伝子を失った
結果、ヒトの遺伝子レパートリと類似？

Hara, et al., *Nat. Ecol. Evol.* 2018
山口 & 工樂、比較生理生化学 37: 170-179. 2020

山口和晃さん



大阪公立大
寺北先生
小柳先生と共同



発色団レチナール

+

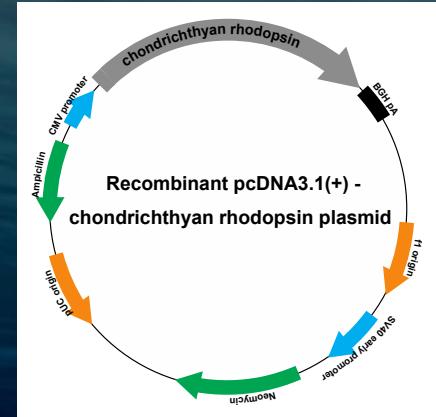
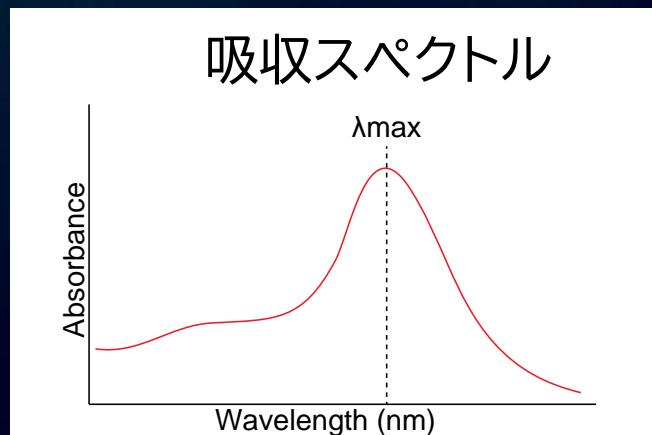


タンパク質を抽出し精製

↓
分光測定

サメの口ドプシンタンパク質を
発現させるためのプラスミドの作製

↓ HEK293 細胞を形質転換しタンパク質発現
口ドプシン



ロドプシンの吸収スペクトル

約5000万年前



Rhincodon typus
whale shark

ジンベエザメ



Stegostoma tigrinum
zebra shark

トラフザメ



Chiloscyllium punctatum
brownbanded bamboo shark

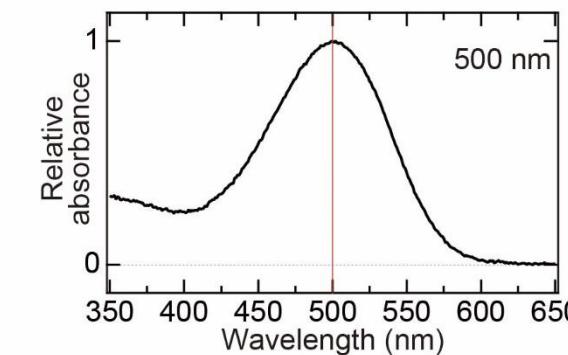
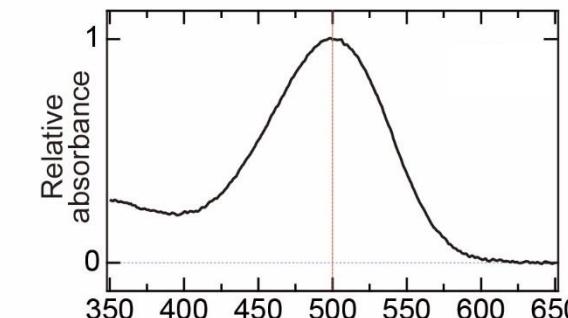
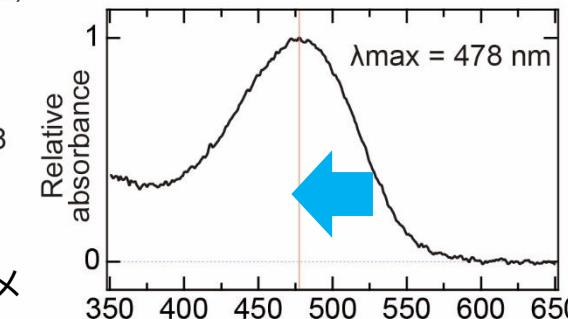
イヌザメ

Habitat depth (m)

~1,928

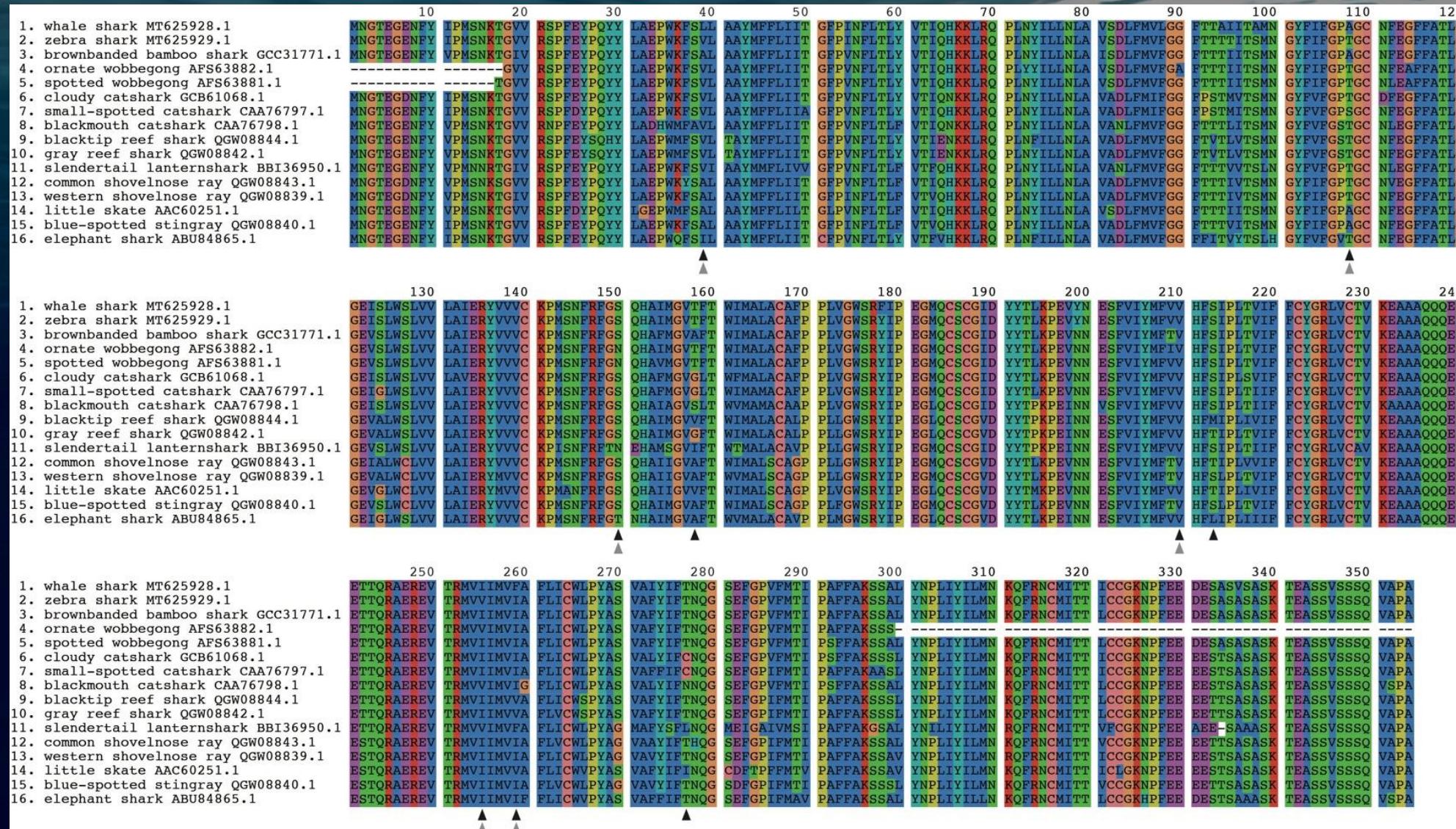
~62

~85

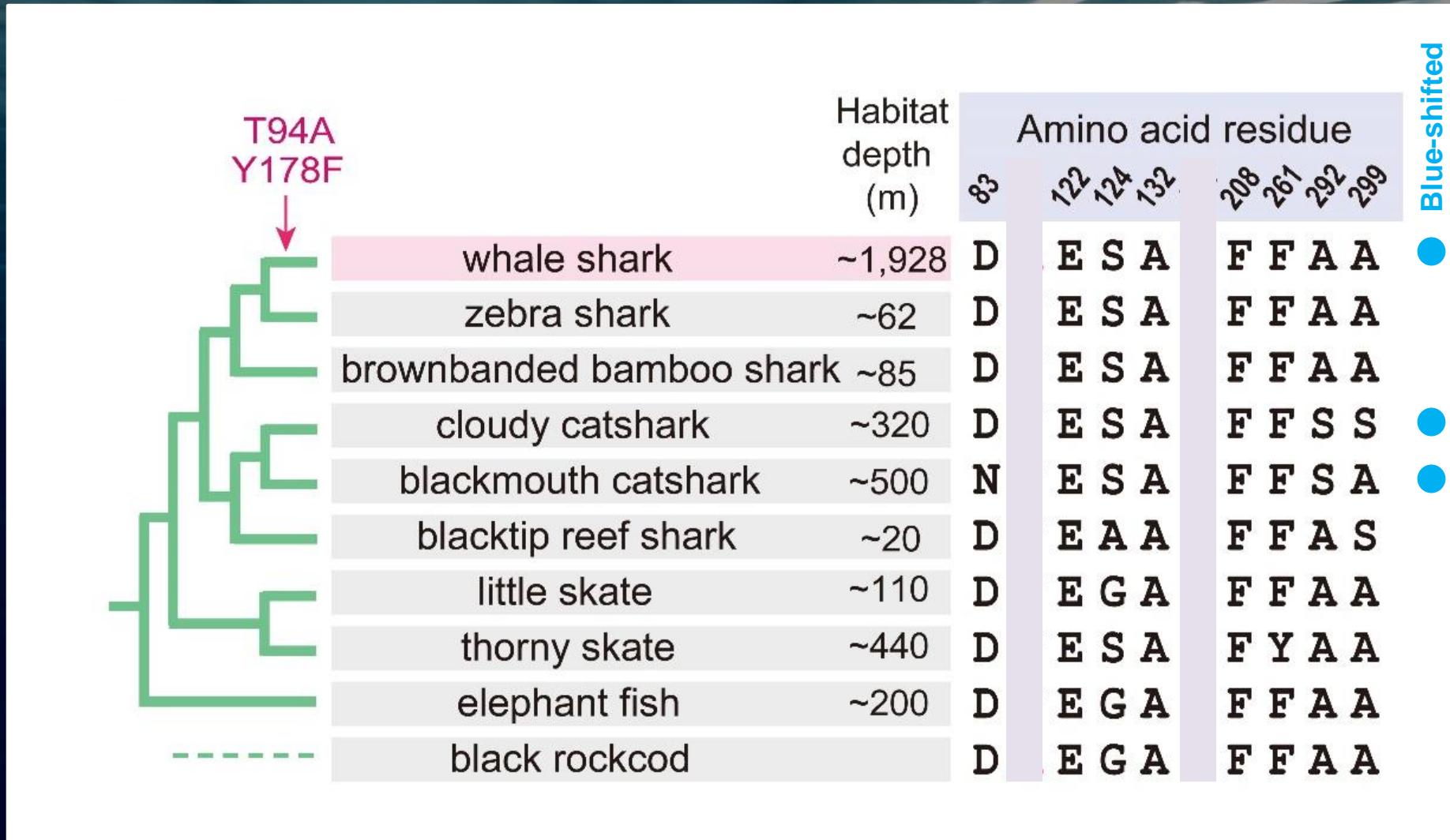


生体試料が
なくても…

軟骨魚類のロドプシンのアミノ酸配列の比較



既知の'Spectral tuning sites'は？

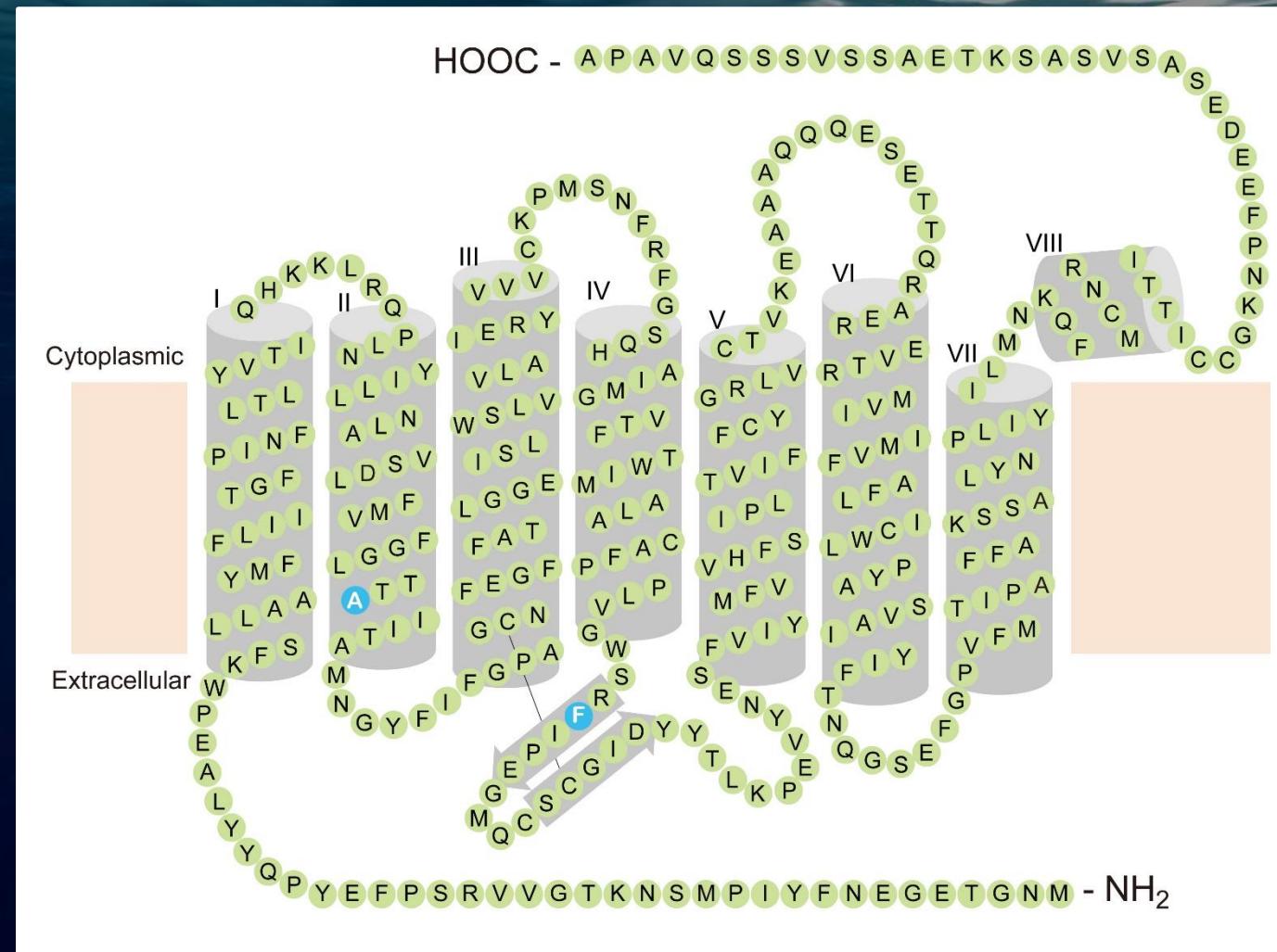


参照) プレスリリース「ジンベエザメだけに起きた視覚の進化」

<https://www.nig.ac.jp/nig/images/research highlights/PR20230322.pdf>

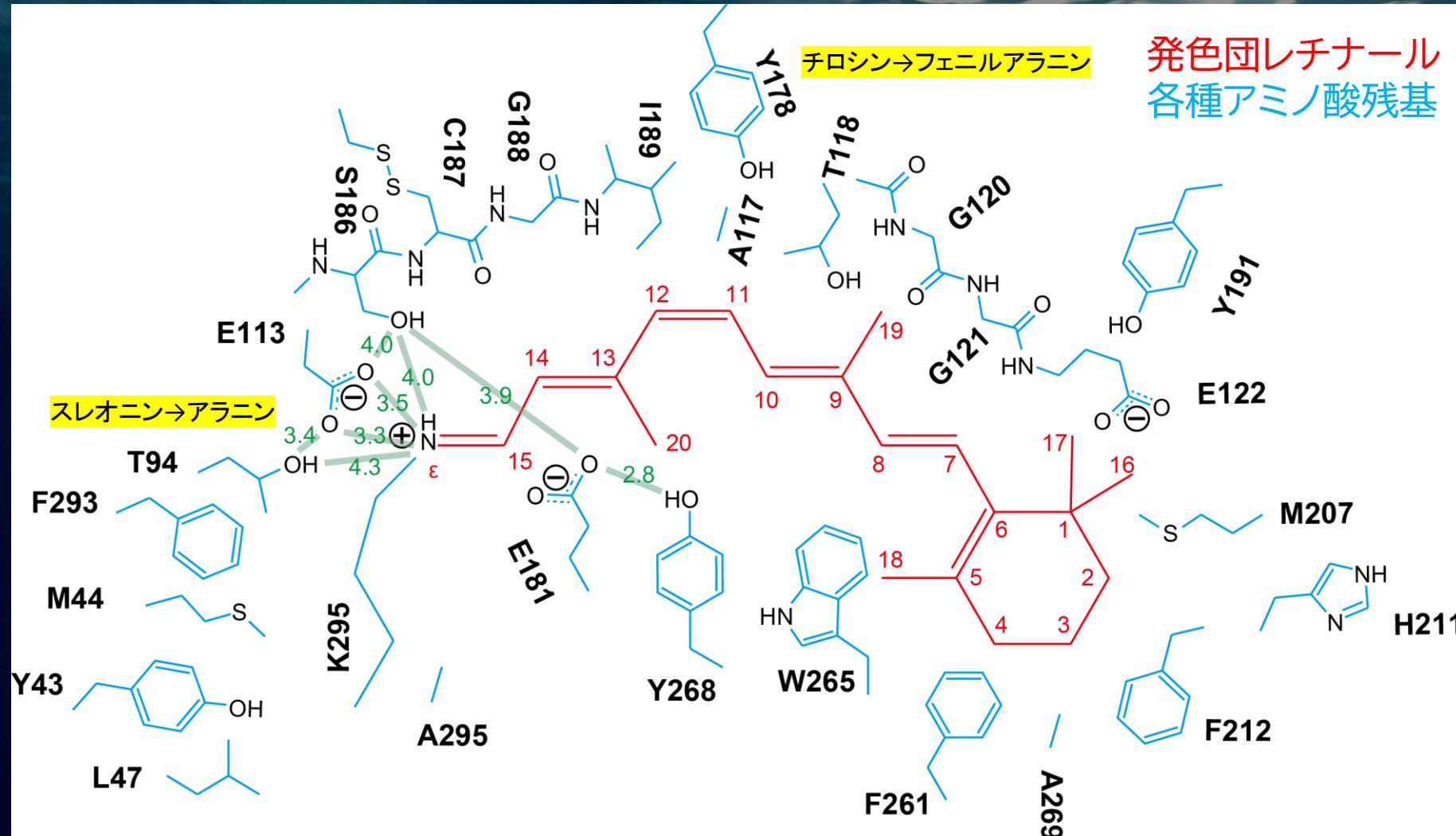
ロドプシンの構造

ジンベエザメのアミノ酸配列

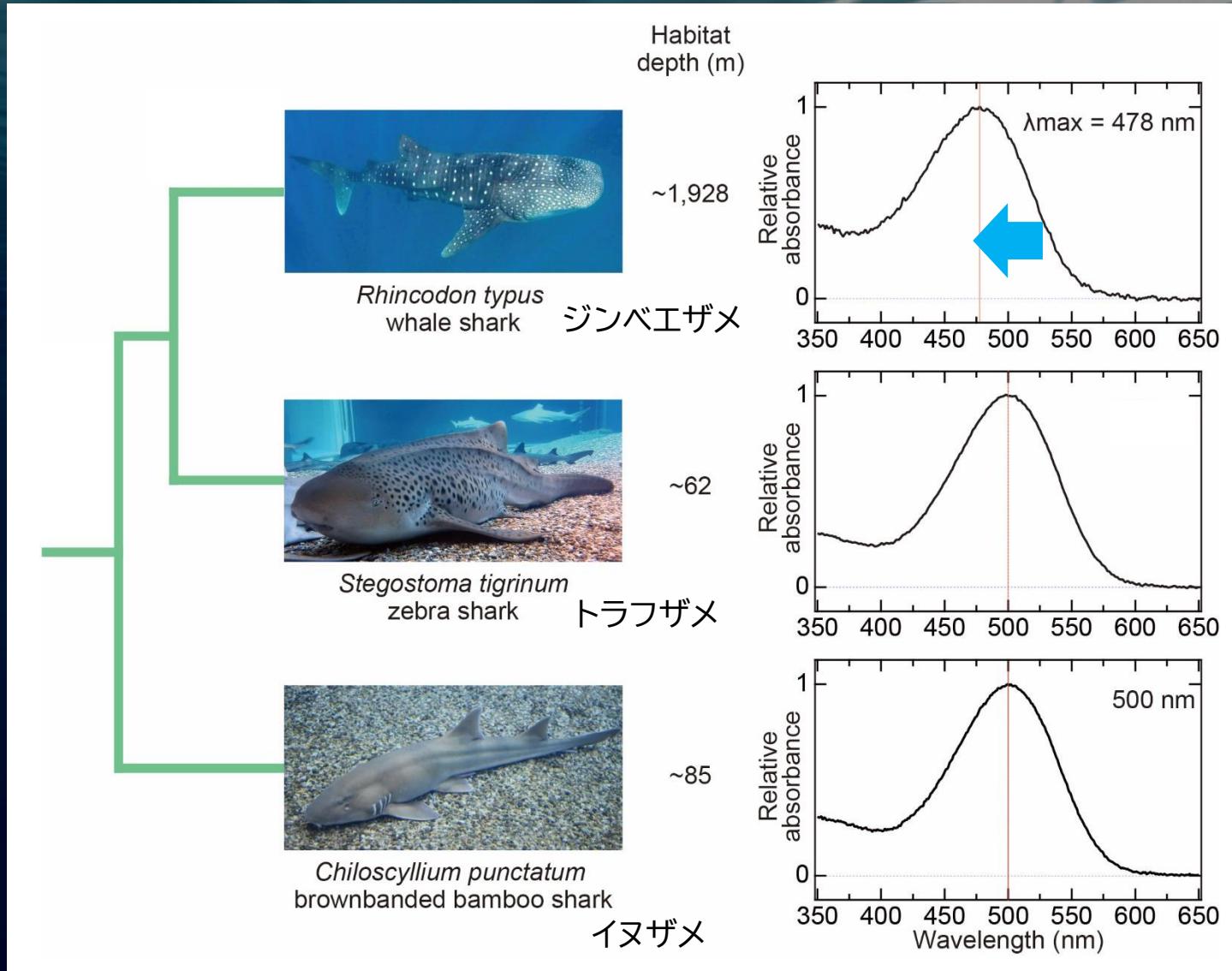


ロドプシンの構造

ジンベエザメ特有のアミノ酸置換はウシロドプシンならどこに？

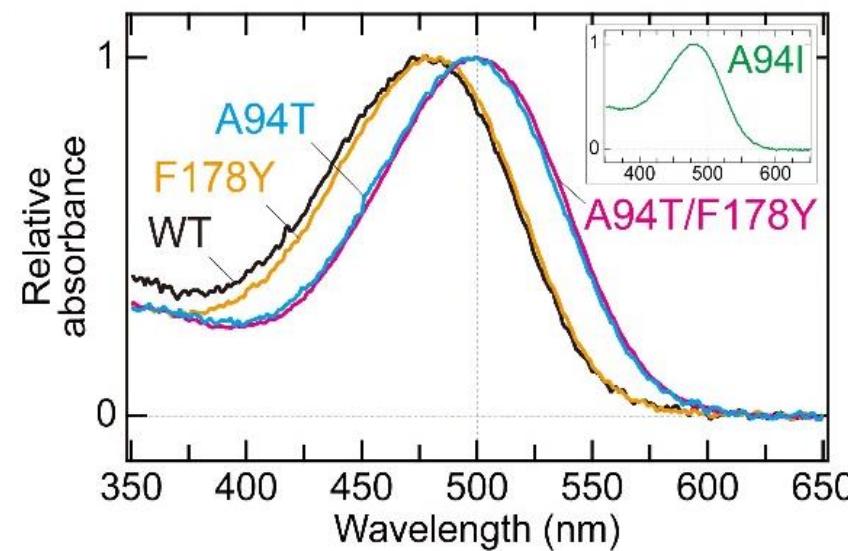


ブルーシフトの原因残基は？

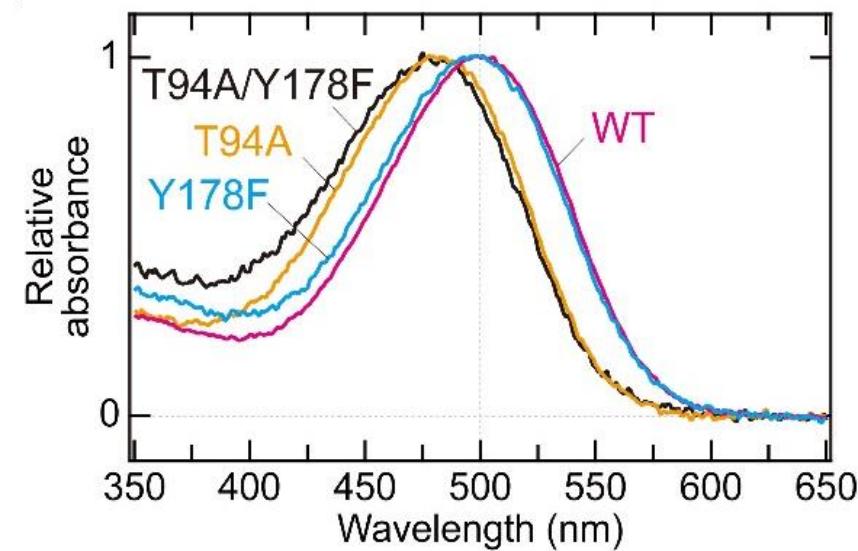


ブルーシフトの原因残基は？

ジンベエザメRHOを浅海型に

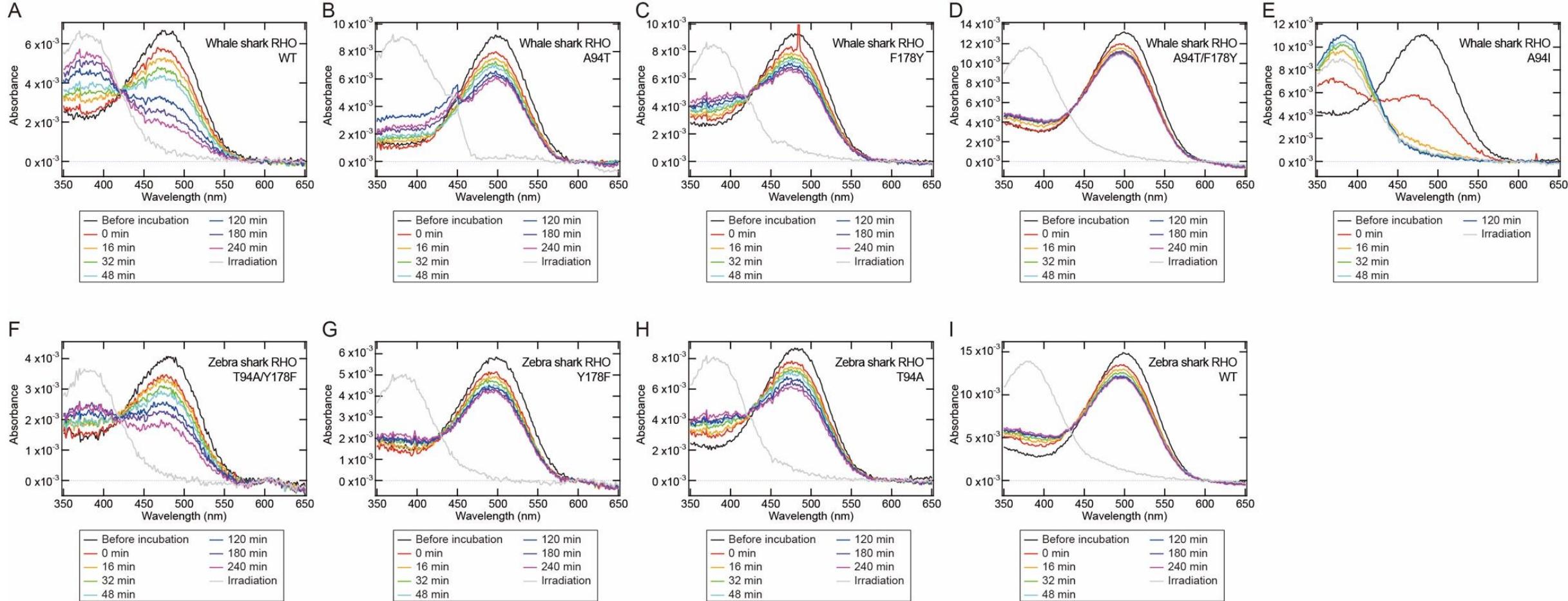


トラフザメRHOを深海型に



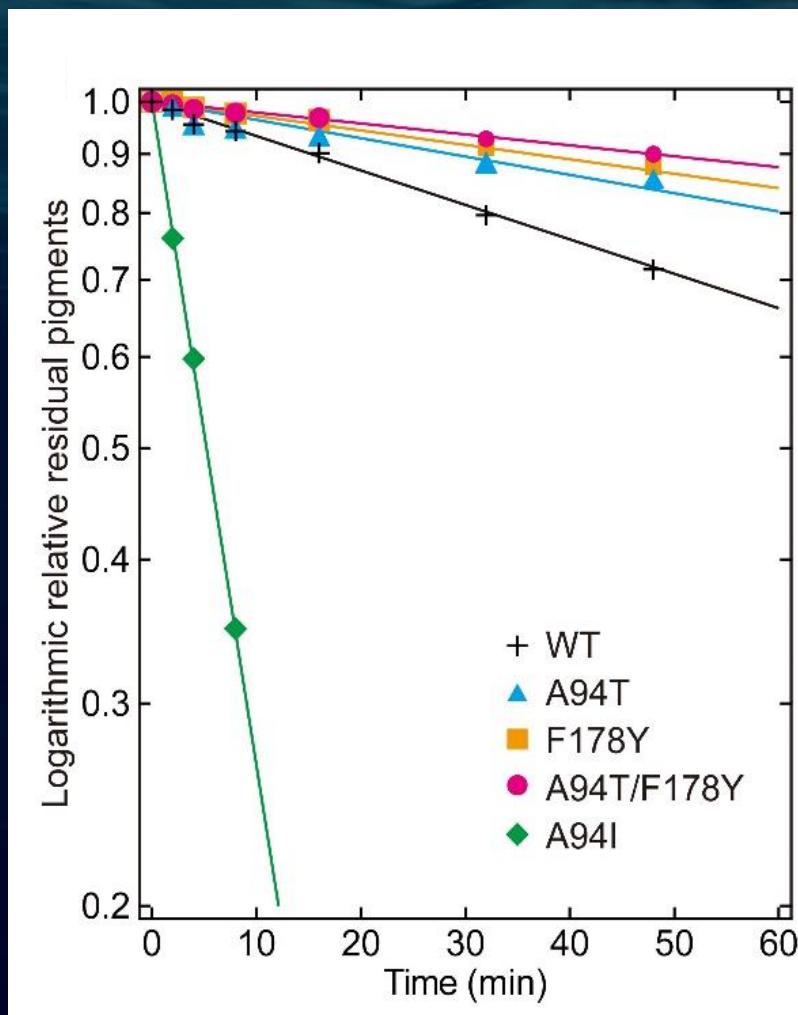
$$500 \text{ nm} - 478 \text{ nm} = 22 \text{ nm}$$

ジンベエザメ・トラフザメの野生型・変異型ロドプシンの熱安定性

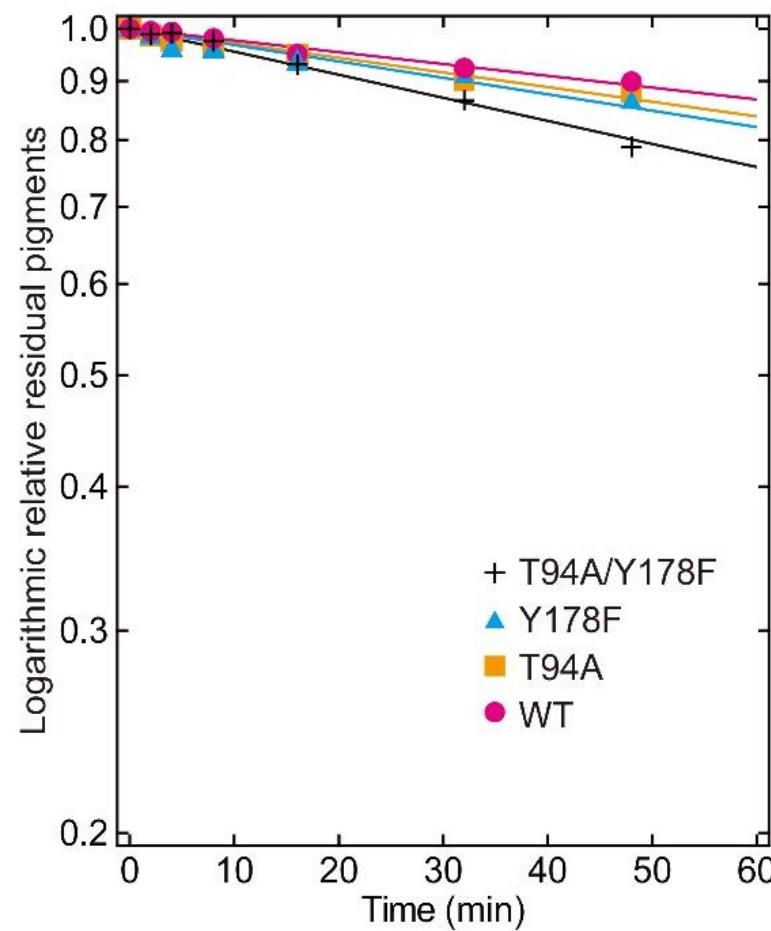


野生型・変異型ロドプシンの熱安定性

ジンベエザメ



トラフザメ



まとめ

深海性のサメ類ではいくつかの種がブルーシフト(ロドプシンの最大吸収波長の480nm付近へのシフト)を起こしているが、94番目と178番目の残基が原因となっているのはジンベエザメのみ。

ジンベエザメの生態を考えたとき、特徴的なのは、深海利用のいっぽうで海水面近くで活動する時間が長い。他の深海ザメと異なる仕組みが必要？

94番目の残基によるブルーシフトは夜盲症の原因のひとつでもあり、夜盲症と同様、ロドプシンの熱安定性の低下が観察される。94番目の残基の置換は南極に棲むノトテニア属硬骨魚にも見られ、その低温環境は、ジンベエザメが利用するとされる深海と共通性がある。

生体組織を用いずに、近縁種を含めたゲノム情報比較と人工合成による分子機能実験で示した成果。希少生物でも生態研究の広大な可能性。