

令和元年6月18日現在

機関番号：38005

研究種目：基盤研究(C) (一般)

研究期間：2015～2018

課題番号：15K07172

研究課題名(和文)脊椎動物の全ゲノム比較解析を可能にする統合オーソログ判定パイプラインの構築

研究課題名(英文) Development of an integrated pipeline of ortholog identification for comparative genome analyses of vertebrates

研究代表者

井上 潤 (Inoue, Jun)

沖縄科学技術大学院大学・マリンゲノミクスユニット・研究員

研究者番号：10596779

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：遺伝子の機能を推定する web tool, ORTHOSCOPE を開発した。ORTHOSCOPE は、ユーザーが注目する左右相称動物に含まれる数種のゲノムデータから、同じ機能を持つと考えられる遺伝子を判定する。

ORTHOSCOPE を用いて、脊椎動物祖先の出現を可能にした 2 つの器官の由来に迫った：(1) 脊椎動物のすばやい動きを可能にする骨格筋は、骨格筋アクチンとトロポニン C を新たに用いることで獲得された。(2) 脊索動物を特徴付ける脊索は、Brachyury 遺伝子のコピーを新たに作成したのではなく、既存の遺伝子に新たな機能を加えることで獲得された。

研究成果の学術的意義や社会的意義

作成した ORTHOSCOPE はウェブツールであるため、広く一般の利用が可能である。このツールを用いれば、分子系統解析を専門としない研究者であっても、自身が注目する遺伝子がある分類群のゲノムに存在するのか判定し、その由来を推定できる。

このため学術的意義は、ORTHOSCOPE を用いれば、左右相称動物のある動物群(脊椎動物など)の出現を可能にした器官と遺伝子の起源に迫る研究を、多くの研究者が独自に展開できるようになった点である。社会的な意義としては、医療で注目される遺伝子の進化的起源にも迫れるため、医学的な知見にも進化的解釈を導入しやすくなった点が挙げられる。

研究成果の概要(英文)：We developed a web tool, ORTHOSCOPE, to infer gene functions. ORTHOSCOPE identifies candidate genes sharing the same function from genome data of focal bilaterians.

By using ORTHOSCOPE, we estimated origins of two organs, those enabled the emergence of chordates: (1) the skeletal muscle, which sustains fast movements of vertebrates, was acquired by using skeletal actin and troponin C. (2) the notochord, which supports fish-like larvae sharing among chordates, was acquired by using a single copy of the Brachyury gene with two expression domains, not by multiple copies of the gene produced by gene duplication events.

研究分野：分子系統分析を用いたゲノム比較解析。主に脊索動物を対象とする。

キーワード：解析パイプライン Web tool ゲノム比較 左右相称動物 脊索動物の起源

1. 研究開始当初の背景

(1) 背景

本課題の申請書を作成した 2014 年秋の時点であっても、すでにゲノムデータが多くの脊椎動物で解読されていた。そして、解読されたデータを簡潔に比較しただけではあったが、脊椎動物は 3 回の太古の全ゲノム重複を経験したことは、疑いようのない事実となっていた (図 1)。このことは、脊椎動物のゲノムを比較するには、全ゲノム重複の痕跡を考慮に入れる必要があることを意味する。

全ゲノム重複とは、遺伝子をコードする領域と非コード領域にかかわらず、遺伝情報が全て重複するイベントである。このため当然、イベントの後にコピー遺伝子が残り、種間で遺伝子の対応関係が複雑になっている。種間で遺伝子を比較する場合、その関係は、遺伝子重複によってコピーされたのか、あるいは、種分化によって分かれたのか、どちらかである。前者をパラログス、後者をオーソログスな関係という。

例えば種の系統関係を推定するには、オーソログだけを集めて解析しなければならないのは、想像に難くない。オーソログ/パラログ判定は、あらゆるゲノム比較解析の根幹をなす問題である。

脊椎動物根幹で生じた 2 回の全ゲノム重複は (図 1)、脊椎動物のゲノム構造の成立にも大きな影響を及ぼした。ゲノム構造の比較では、数億年におよぶ染色体の融合と分離の結果、全ゲノム重複で生じたコピー染色体領域がパッチ状に存在することが問題となる。コピー染色体領域の由来判定に遺伝子のオーソログ/パラログ関係が利用できるが、以下の理由から解析が滞っている： 1) ゲノム重複と種分化で生じた遺伝子が入り混じっており、その対応関係が複雑である。2) 単一遺伝子のみの短い配列に基づく系統解析では、5 億年前のイベントを検証するのに信頼性の高い遺伝子系統樹が推定しにくい。そこで申請者と研究グループのメンバーは、本課題の前に採用された若手 (B) の課題 (2012~2015 年) で、より最近でしかも 1 回のイベントである、真骨魚類根幹で生じた全ゲノム重複に注目し、全ゲノム重複で生じたコピー遺伝子の欠失・保持過程を解析した。本課題 (基盤 C) に着手した 2015 年春には、その成果を論文としてまとめ (Inoue et al. 2015)、投稿先の雑誌とのやりとりを行っていた。

(2) 動機

Inoue et al. (2015) では、真骨魚類を中心とした硬骨魚類のゲノム比較に特化した解析パイプラインを作成した。全てのタンパク質遺伝子のオーソログ情報を推定できる解析パイプラインを手にしたあと、注目すべきは、染色体領域同士の比較であった。Inoue et al. (2015) では、解析パイプラインで推定した遺伝子のオーソログ情報が、種間でオーソログな染色体領域を判定するアンカーとして、もっとも適していることが判明した。オーソログの並び方 (シンテニーと呼ばれる) が種間で類似する領域を判定すれば、種間でゲノム構造を比較する際に、配列の類似した染色体領域のうち、どのペアが種分化で、どのペアが全ゲノム重複で別れたのか、明らかになる。つまり染色体領域の歴史的背景がわかるのだ。そこで、この申請課題では当初、ヒトから真骨魚類を含む硬骨魚類それぞれの種で保存されるであろう、このオーソログ染色体領域に注目した。

2. 研究の目的

上記の動機から、申請時における本課題の目的は、オーソログな遺伝子を判定し、これをもとにオーソログな染色体領域を脊椎動物のゲノムから選定する解析パイプラインを作成することであった。さらに、アミアやアロワナなど、古代魚のゲノムデータを解読し、作成した解析パイプラインを適用して、真骨魚類根幹で生じた全ゲノム重複直後に、どのようにコピーされた遺伝子が欠失してゆくのかを、より詳細に検討する予定であった。

しかし、研究環境が変わってしまった。この本課題が始まった 1 年後の 2016 年 4 月に、申

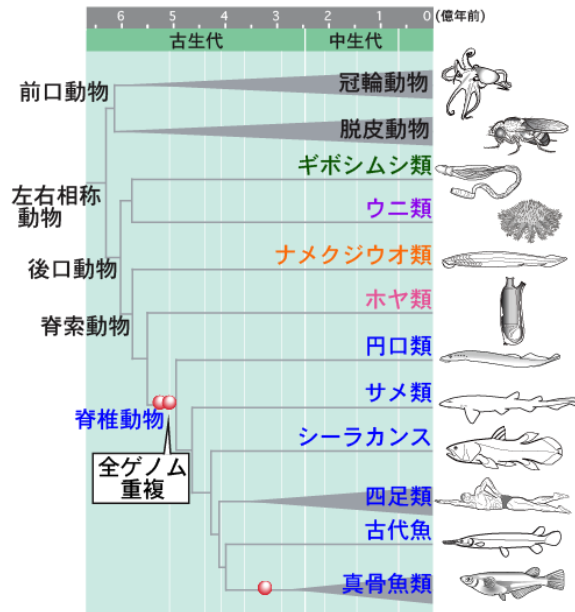


図 1 左右相称動物の時間軸付き系統樹。

請者は、所属研究室を変更せざるを得なくなったのだ。幸運なことに、変更先の研究室では、対象生物を脊椎動物から、脊索動物を中心とした後口動物全体へと変更すれば、ゲノム比較解析を続けるチャンスに恵まれたのである。所属先の研究室で与えられたミッションを進行するには、以下の課題があった：1) 後口動物の主要 5 系統を網羅するゲノムデータを扱う。2) 脊椎動物に比べ、多くの無脊椎動物では、解読されたゲノムデータの質が低い。3) ある形質の出現に主要な役割を演じた特定のタンパク質遺伝子を詳細に解析する。

そこで、申請者は、研究目標を、以下のよう改めた：1) 脊索動物を中心とした後口動物のゲノムから、オーソログな遺伝子を抽出する解析プログラムを作成する。2) 全タンパク質遺伝子を一度に解析するのではなく、脊索動物の出現を可能にした形質を司る遺伝子に焦点を当て、その由来を探る。

対象とする生物の変更に当初は戸惑いがあった。実際、脊椎動物に比べて無脊椎動物ゲノムデータは質が低いため、オーソログ染色体領域の判定は諦めざるを得なかった。しかし、状況の変化は、第一著者として合計 3 報の論文出版という、充実した成果を生み出す結果につながった。

3. 研究の方法

本課題で出版した 3 つの論文 (Inoue et al. 2017; Inoue and Satoh, 2018, 2019) を通じて、ゲノム比較研究に有用な以下 3 種類の手法を新た開発することとなった。

(1) ORTHOSCOPE の開発。

ウェブツール ORTHOSCOPE を開発した (図 2)。ORTHOSCOPE は、配列データをアップロードするだけで、遺伝子系統樹が構築され、これに基づきアップロードした遺伝子配列の機能を推定できるウェブツールである。

(2) オーソグループによる遺伝子機能の推定。

ORTHOSCOPE の中核となるのが、遺伝子の機能推定に、オーソグループを活用する点である (図 3)。これまでの研究では、ある遺伝子の機能を推定するのに、モデル生物のオーソログを探し、その機能を参照する、というものであった。しかし、一つだけオーソログを判定しただけでは不十分で、複数のモデル生物のオーソログを含むオーソグループの判定こそが、遺伝子機能の推定に重要であることが、ORTHOSCOPE の開発途上で明確になった。

(3) 遺伝子発現データに基づく祖先遺伝子の発現部位推定

この手法を用いれば、遺伝子系統樹から、注目した遺伝子の祖先となる遺伝子が、どのような部位で発現していたのか推定できる (Inoue and Satoh 2018)。ゲノム解読

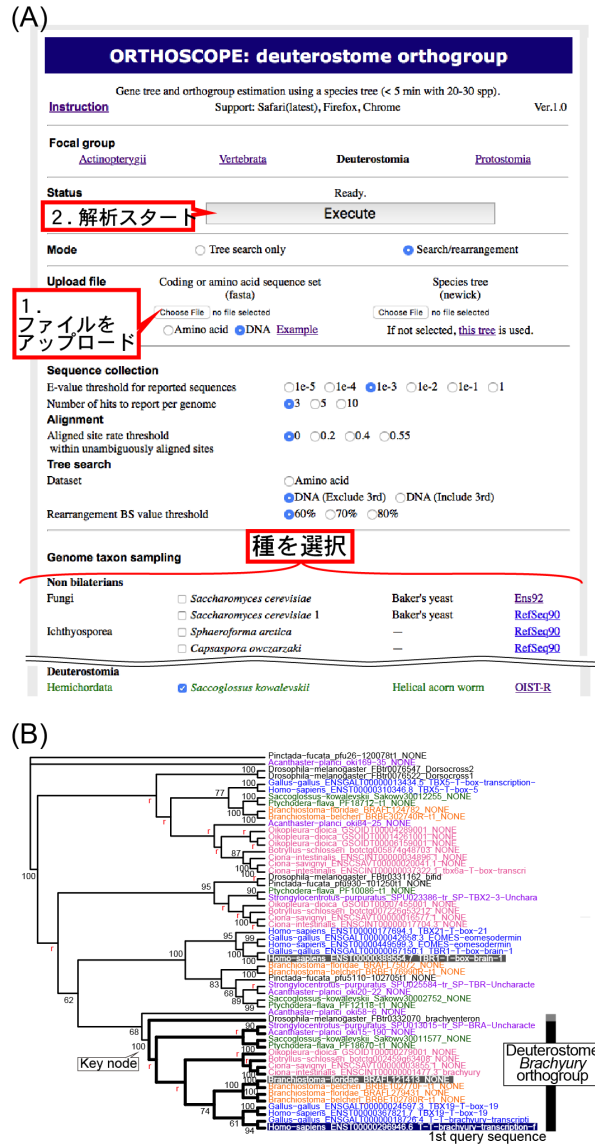


図 2 ORTHOSCOPE のインターフェース (A) と得られる遺伝子系統樹 (B)。

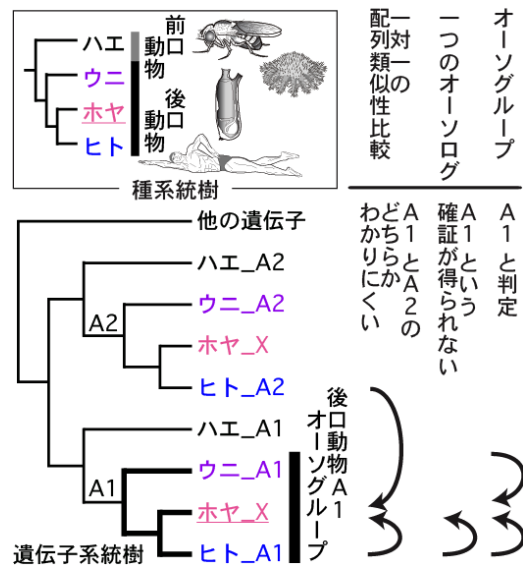


図 3 オーソグループによるホヤ遺伝子の機能推定。モデル生物を参考に遺伝子の機能を推定するには、オーソログの集まりとしてオーソグループを判定することが有効である。

で得られる遺伝子の一部は、過去の研究から実験的にその発現部位が明らかになっている。これらを遺伝子系統樹上にマッピングし、その遺伝子が含まれるクレードの祖先となる遺伝子の発現部位を推定する。

4. 研究成果

(1) ORTHOSCOPE の開発

① 主な成果

遺伝子の機能を自動的に推定するウェブツール ORTHOSCOPE を開発した (図 2)。

② 成果の位置付けとインパクト

本研究以前は、対象とする種とモデル生物の間で 1:1 対応だけで配列類似性を比較して、遺伝子の機能を推定することがほとんどであった (図 3)。しかし本研究によって、遺伝子系統樹からオーソグループを判定して、遺伝子の機能を推定するという簡潔で確実な方法が、ウェブツールとして広く一般に利用できるようになった。このウェブツールを利用して、遺伝子の機能推定を専門としない研究者から、すでに 2 つの論文 (Ishikawa et al. 2019; Shiraiishi et al. 2019) が、出版されている。

③ 今後の展望

ORTHOSCOPE をさらに発展させる。特に、1 つの遺伝子解析に特化した ORTHOSCOPE を改定し、全てのタンパク質遺伝子をスーパーコンピュータで同時に解析するスクリプトの開発が必要である。実際にそのスクリプトは、ほぼ完成している。

④ 予期しない結果から得た新知見

脊椎動物の起源など、理学的なトピックの解析を目標に ORTHOSCOPE を開発したが、医学的な利用にも対応可能であることがわかった。例えば Yasuoka et al. under review は、ORTHOSCOPE を用いて再生医療で注目される遺伝子の由来を推定した。彼らは、その遺伝子の祖先的な機能を推定するとともに、脊椎動物の共通祖先からヒトに至る進化のどの段階で、その遺伝子が現在ヒトで見られるような機能を持つに至ったか推定した。

(2) 脊椎動物骨格筋の起源

① 主な成果

脊索動物の出現を可能にした脊椎動物骨格筋の原型は、2 つの遺伝子を新たに利用することによって誕生したことを明らかにした (図 4)。

② 成果の位置付けとインパクト

骨格筋など筋肉の起源を調べる研究はこれまで多くなされてきたが、万人を納得させる成果に至らなかった。本研究では、遺伝子系統樹の正確な推定と他の研究で観察された発現部位を合わせた緻密な祖先状態の推定から、構成するタンパク質の由来を推定することが、器官全体の起源を推定につながることを示した。

③ 今後の展望

本研究で用いた遺伝子機能の推定方法は、ORTHOSCOPE に組み込まれている。ORTHOSCOPE を使えば、例えば心臓や内臓などを構成する筋肉など、脊椎動物以外の生物で相同器官が知られない筋肉でもその由来を推定することができる。さらに他の生物群にも目を向け、昆虫類の飛翔筋、貝類のキャッチ筋など、それぞれの分類群の繁栄をもたらした筋肉の解析にも適用可能である。ORTHOSCOPE は、脊索動物の解析を中心としているが、前口動物や条鰭魚類などの解析も可能である。もちろん、筋肉だけでなく他の器官の解析も期待できる。

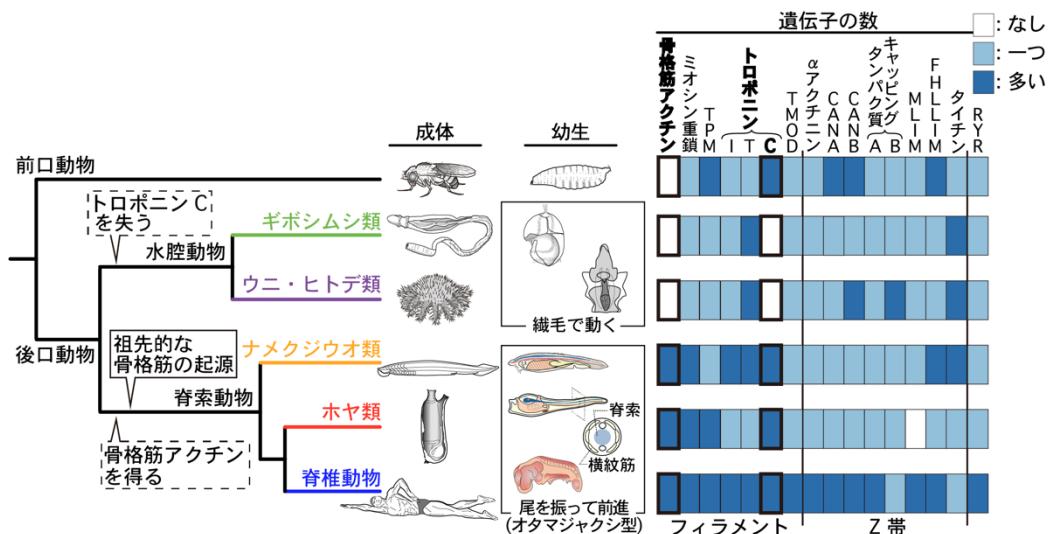


図 4

後口動物の種系統樹と、骨格筋タンパク質遺伝子の保持パターン。骨格筋アクチンとトロポニンCを利用して祖先的な骨格筋は誕生し、脊索動物が出現した。

④ 予期しない結果から得た新知見

これまで、筋肉など器官の進化を扱う研究で遺伝子系統樹は利用されていたが、推定された系統樹は曖昧で十分な考察がなされていなかった。本研究では、遺伝子系統樹の威力を遺憾なく発揮したと言えよう。遺伝子系統樹に発現データを合わせることで、祖先状態を推定した。この考えは、系統地理学で発展した祖先的生息域の推定を、発現器官の推定に適用したものである。タンパク質の祖先状態を推定するこの新たな手法によって、器官全体の起源を推定できるという新知見を本研究で得た。

(3) 脊索を形成する *Brachyury* の起源

① 主な成果

ナメクジウオ類以外は、後口動物の各系統では *Brachyury* 遺伝子が 1 つ存在することを明らかにした (図 5)。このことは、脊索を獲得した脊索動物の祖先は

Brachyury 遺伝子を 1 つしか持たなかったことを意味する。

② 成果の位置付けとインパクト

脊索の獲得という現象は、重要でありながらもゲノムデータが無脊椎動物で解読されていなかったため、分子分析の対象にされなかった。しかし 2015 年、ついにギボシムシ類のゲノムデータが公開され、後口動物主要 5 系統を代表するデータが出揃った。これらのデータを用いて本研究では、脊索動物の祖先が脊索を、新たなコピー *Brachyury* 遺伝子を作成して獲得したのではなく、既存の遺伝子に新たな機能を付加して獲得した、ことを明確にした。

③ 今後の展望

Brachyury 遺伝子を脊索にも発現させている調節配列を明らかにする必要がある。さら本研究では、ナメクジウオ類では *Brachyury* 遺伝子が 2 コピーあることが判明した (図 5)。このナメクジウオ類に存在する 2 つの *Brachyury* 遺伝子がそれぞれどのような役割を担うのか、発現解析によって検証する必要がある。

④ 予期しない結果から得た新知見

ORTHOSCOPE の開発段階で、最初に行った研究がこの脊索動物 *Brachyury* 起源の解明であった。この研究は、後口動物全 5 系統からゲノムデータが出揃って初めて可能になった。このため当初は、これらのゲノムデータの質が解析に十分なものであるか不明であった。しかし本研究のように、遺伝子モデルに注目すれば、ノイズの中から真実を反映するデータを抽出できることが判明した。

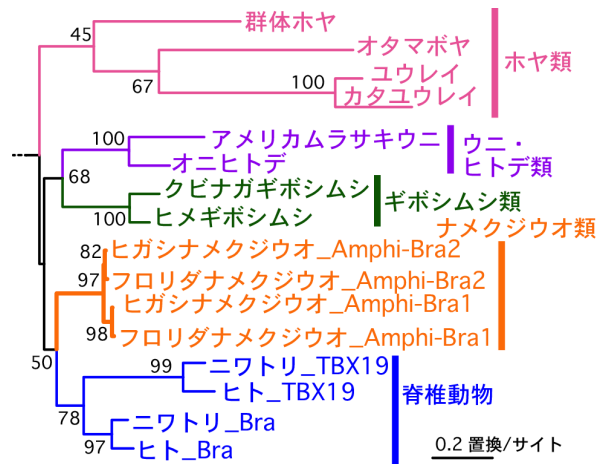


図 5 *Brachyury* の遺伝子系統樹。ナメクジウオ類以外、後口動物の各種は、*Brachyury* 遺伝子を 1 コピー持っていた。

5. 主な発表論文等

[雑誌論文] (計 6 件)

- (1) Inoue J, Satoh N. 2019. ORTHOSCOPE: an automatic web tool for phylogenetically inferring bilaterian orthogroups with user-selected taxa. *Molecular Biology and Evolution*. **36**: 621–631. 査読有. doi.org/10.1093/molbev/msy226
- (2) Ishikawa, A, Kabeya, N, Ikeya, K, Kakioka, R, Cech, JN, Osada, N, Leal, MC, Inoue J, Kume, M, Toyoda, A, Tezuka, A, Nagano, AJ, Yamasaki, YY, Suzuki, Y, Kokita, T, Takahashi, H, Lucek, K, Marques, D, Takehana, Y, Naruse, K, Mori, S, Monroig, O, Ladd, N, Schubert, C, Matthews, B, Peichel, CL, Seehausen, O, Yoshizaki, G, Kitano J. 2019. A key metabolic gene for recurrent freshwater colonization and radiation in fishes. *Science*. **364**: 886–889. 査読有. doi.org/10.1126/science.aau5656
- (3) Shiraishi A, Okuda T, Miyasaka N, Osugi T, Okuno Y, Inoue J, and Satake H. 2019. Repertoires of G protein-coupled receptors for Ciona-specific neuropeptides. *Proceedings of the National Academy of Sciences of the United States of America*. **116**: 7847–7856. 査読有. doi.org/10.1073/pnas.1816640116
- (4) Inoue J, Satoh N. 2018. Deuterostome genomics: Lineage-specific protein expansions that enabled chordate muscle evolution. *Molecular Biology and Evolution*. **35**:914–924. 査読有. doi.org/10.1093/molbev/msy002
- (5) Inoue J, Yasuoka Y, Takahashi H, Satoh N. 2017. The chordate ancestor possessed a single copy of the *Brachyury* gene for notochord acquisition. *Zoological Letters*. **3**: 4. 査読有. doi.org/10.1186/s40851-017-0064-9
- (6) Inoue J, Sato Y, Sinclair R, Tsukamoto K, Nishida M. 2015. Rapid genome reshaping by multiple-gene loss after whole-genome duplication in teleost fish suggested by mathematical modeling. *Proceedings of the National Academy of Sciences of the United States of America*. **112**: 1211–1216. 査読有. doi.org/10.1073/pnas.1412111112

〔学会発表〕 (計 3 件)

- (1) 井上 潤, 佐藤 矩行. 種の系統樹に基づいて左右相称動物のオーソログを判定するウェブツールの開発. 日本動物学会 第 89 回大会 (代替)・本郷 (2018 年 12 月).
- (2) Inoue, J, Satoh, N. A Genomic survey of muscle structural proteins that enabled the rise of chordates. Society for Molecular Biology and Evolution Conference 2016. Gold Coast, Australia (July, 2016).
- (3) 井上 潤, 佐藤 行人, シンクレア ロバート, 塚本 勝巳, 西田 睦. 全ゲノム重複と種の系統樹を考慮した系統解析マーカー遺伝子の選定. 日本進化学会年会「夏の学校」. 中央大学 (2015 年 8 月).

〔図書〕 (計 1 件)

- (1) 井上 潤. 2018. 古代魚の系統進化. 下位真骨類の系統進化. 魚類学の百科事典. 日本魚類学会 編. 丸善出版. p60-63.

〔産業財産権〕

○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕

ホームページ等

ORTHOSCOPE (<https://www.orthoscope.jp>)

6. 研究組織

(1) 研究分担者

なし.

(2) 研究協力者

佐藤 矩行 (SATO, Noriyuki)

沖縄科学技術大学院大学・マリンゲノミクス・教授

研究者番号: 30025481