

令和 4 年 5 月 7 日現在

機関番号：72698

研究種目：基盤研究(B) (一般)

研究期間：2019～2021

課題番号：19H04070

研究課題名(和文)種-集団-個体レベルの多様性の遺伝的背景を総合理解する統計モデルの開発

研究課題名(英文)Statistical modeling of genetic background behind species-population-individual level diversity

研究代表者

岸野 洋久(Kishino, Hirohisa)

(財)進化生物学研究所・その他部局等・客員研究員

研究者番号：00141987

交付決定額(研究期間全体)：(直接経費) 13,200,000円

研究成果の概要(和文)：本研究は、表現型と遺伝子型の種内・種間の多様性を総合的に理解することを目的とした。種間の分析では、分子進化速度を予測因子として生活史の諸形質の間の共進化を推定した。種内の遺伝的多様度が長期にわたる分子進化速度に依存することが明らかになった。分子進化速度による予測からの偏差から、種特異的な選択圧を検出する方法を開発した。集団史に照らして多形質共適応を理解する探索的解析を提案した。多形質の方向性淘汰の歴史を表現する行列に主成分分析を適応し、主たる共適応を検出し、これを集団分化・混合グラフ上にマッピングした。

研究成果の学術的意義や社会的意義

本研究で提案する手法は、集団にかかる選択圧と適応を種間比較の中で相対化するものである。種の形態的進化には、結合組織よりもむしろ上皮細胞の進化が主に寄与しており、ヒトにおいては、免疫系の進化と感染症関連遺伝子への機能的な制約の強化が適応進化の代表例であった。共適応の探索的データ分析は、対称を絞った分析・実験では気づかない適応にも光を投げかける。アメリカ北西部の野生ボブラの解析では、光合成関連形質と成長関連形質の気温と日長への応答、地上部環境の選択への制限因子としての土壌条件が示唆された。

研究成果の概要(英文)：The aim of this study was to gain a comprehensive understanding of intra- and interspecific diversity of phenotypes and genotypes. In the interspecies analysis, coevolution among life history traits was estimated using the rate of molecular evolution as a predictor. We found that intraspecies genetic diversity depends on the rate of molecular evolution. Based on this, we developed a method to detect species-specific selection pressure by using deviations from predictions based on molecular evolutionary rates. We proposed an exploratory analysis of multi-trait coadaptations in the light of population history. By defining a matrix representing the history of adaptation of polymorphisms, we detected sets of major co-adaptations by principal component analysis and mapped onto an admixture graph. The newly defined factor loadings of environmental variables enabled to understand the background of co-adaptation.

研究分野：統計科学、分子進化学、集団遺伝学

キーワード：分子進化 集団ゲノム 高次多変量解析 階層分散分析

## 1. 研究開始当初の背景

ゲノムのデータが充実するのに伴い、遺伝子型と表現型を関連付ける統計遺伝学の方法も研ぎ澄まされてきている。が、種内の遺伝的多様性は近現代に生じた突然変異に起因するのに対して、種間の遺伝的多様性は長期にわたる桁違いに多くの変異を反映している。このため、集団の関連解析の方法を種間比較に適用しても、高い検出力は望めない。申請者は、多遺伝子系統樹を分散分析して突然変異率と遺伝子への淘汰圧を分離し、予測因子として分子進化速度を導入することにより表現型の種間の違いの遺伝的背景を偏りなく推測する方法を提案した。

哺乳類の生活史の進化を推測したところ、図1に見られるように、恐竜が絶滅した K-Pg 境界(およそ 6,500 万年前)以前はほとんどが食虫性であったものが、この時期を境に、肉食、草食、雑食へと食性を転換したことを示唆する、化石記録と符合する結果が得られた。回帰係数からは、群れを作る系統では脳および神経系に関わる遺伝子の機能的な制約を強め、夜行性から昼行性になると、聴覚や嗅覚、触覚に関わる遺伝子への機能的な制約が弱まったことが示唆された。

これらの結果は、表現型の適応進化の推定には分子進化速度を予測因子として利用することが有効なこと、ヒト集団における形質多型の遺伝的背景が、種をまたいだ長期にわたるそれと一致していることを意味している。それらの種内の多型性と種間の遺伝的差異を比較分析することにより、種をまたいで共通にかかる選択圧、種に特異的に働く選択圧を浮き彫りできる可能性があると思われた。

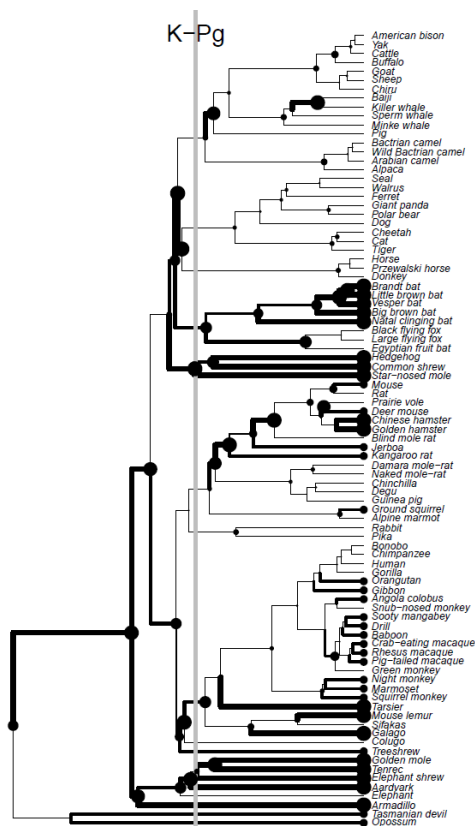


図1 分子進化速度を予測因子として復元された食虫性の祖先状態。丸の大きさは食虫性の予測確率を示す。

## 2. 研究の目的

本研究では表現型の種内の多様性と種間差を総合的に評価し、その遺伝的背景と環境への適応の歴史を調べる統計手法の開発を行うことを目的とした。種間の遺伝的多様性を説明する分子進化と、種内多型を形成する集団遺伝のアプローチを総合して表現型進化に関連付け、種内と種間の多様性を包括する階層的分散分析の方法を開発する。種間の比較では、遺伝子の変わりやすさと表現型の関連を分析する。種内の分集団間の比較では、遺伝子のアレルと環境の関連を分析する。個体間の比較では、分集団内の表現型と遺伝子のアレルの関連を分析する。生活史を構成する諸形質の変化と関連遺伝子を鳥瞰し、全体像を踏まえた上で詳細を掘り進める。種・集団・個体レベルの各階層の多様性の構図を比較分析して情報を補完しつつ、環境が生活史へ及ぼす淘汰圧の実体を探り、遺伝的背景について包括的に理解することを目的とした。

## 3. 研究の方法

### (1) 分子進化速度を予測因子とした祖先形質の復元

ゲノムを種間比較することにより導き出された分子進化の中立説<sup>1</sup>によると、有利な突然変異の割合は無視できるほどに小さい。このとき、分子進化速度は

$$r = (N \times v) \times p \times \frac{1}{N} = vp$$

と一次近似される。Nは集団中の遺伝子数、vは突然変異率、pは中立な変異の割合である。

分子系統樹の枝の長さbは分子進化速度rと進化時間tの積

<sup>1</sup> Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, **217**: 624–626

$$b = r \times t = (v \times t) \times p$$

である。数多くの遺伝子  $i = 1, \dots, N$  について分子系統樹が得られると、それらの枝の長さ  $b_j^{(i)}$  ( $j = 1, \dots, M$ ) は

$$b_j^{(i)} = r_{ij} \times t_j = \bar{p}^{(i)} \times (v_j \times t_j) \times \bar{p}_j^{(i)}$$

と表される。 $\bar{p}^{(i)}$  は遺伝子  $i$  における突然変異のうち中立な変異の割合の系統樹にわたる期待値、 $v_j$  と  $t_j$  はその枝における突然変異率と時間、 $\bar{p}_j^{(i)}$  は遺伝子  $i$  における中立な変異の割合の枝の間の変動を表している<sup>2</sup>。ポアソン回帰により自然に遺伝子効果、枝効果、遺伝子 × 枝相互作用を分離することができる。相互作用は遺伝子にかかる機能的な制約の強さが系統間で変動する様を表現しており、祖先における表現型の推定における予測因子として有望である。

## (2) 種内の遺伝的多様度と分子進化速度

種内の遺伝的多型は、過去に生じた突然変異で、現在までに集団中に固定するまでに至らず、かつ集団から排除されずに残ったものにより生成される。ある部位が多型である確率  $q$  は

$$q = \int_{t=0}^{\infty} dt \int_{-\infty}^0 g(s) ds \Pr(T > t | s, N_e(u), 0 \leq u \leq t) N_e(t) v$$

と表される。ここで、 $v$  は突然変異率、 $s$  と  $g$  は選択強度とその分布、 $N_e$  は有効な集団の大きさである。 $T$  は突然変異が固定化されるか消失するまでの生存時間で、選択強度と集団の歴史に依存する。遺伝的多型が主として中立な突然変異によって生成されるとすると、 $q$  は

$$q = vpF$$

と書き表される。ただし、 $F = \int_{t=0}^{\infty} \Pr(T > t | s = 0, N_e(u), 0 \leq u \leq t) 2N_e(t) dt$  である。このことから、種内の遺伝的多型は分子進化速度に依存することが示唆される。集団遺伝の理論からはある領域の SNP 数の分散は

$$V(K) = E(K) + (E(K))^2 \frac{V(T_c)}{E(T_c)^2}$$

と整理される<sup>3</sup>。 $E(T_c)$  と  $V(T_c)$  は遺伝子の系図の合祖過程における待ち時間の平均と分散であるが、多くの場合データに基づきこれを求めることは難しい。が、この理論式は負の二項回帰に馴染んでおり、データに基づき過分散を評価することができる。

エクソンに働く選択圧は連鎖を通してその近傍にまで伝搬し、遺伝的多様度を押し下げる。選択強度が分子進化速度に依存することを考慮して Hudson-Kaplan モデル<sup>4</sup>を一般化し、ゲノム上のある部位の多様度  $\pi$  を

$$\pi = \pi_0 \prod_{i=1}^m \left[ 1 - \frac{U \Delta x \cdot s_i}{2(s_i + R(x_i))^2} \right] \approx \pi_0 \exp \sum_{i=1}^m \left[ 1 - \frac{U \cdot \Delta x \cdot \exp(\alpha + \beta r_i)}{2(\exp(\alpha + \beta r_i) + R(x_i))^2} \right]$$

と定式化する。和はその染色体上の窓をわたる。 $U$  は突然変異率、 $\Delta x$  は分析の窓幅、 $R(x_i)$  はその部位と窓の間の組換え価である。

## (3) 集団構造を進化の視点で理解する：集団対 $F_{ST}$ と集団固有 $F_{ST}$

集団構造を鳥瞰するために用いられる集団対  $F_{ST}$  は、2つの集団 ( $i, j$ ) に対して

$$pwF_{ST}^{ij} = \frac{H_T^{ij} - H_S^{ij}}{H_T^{ij}} = 1 - \frac{H_S^{ij}}{H_T^{ij}}$$

と定義される。ここで  $H_T^{ij}$  は全ヘテロ接合度、 $H_S^{ij}$  は集団内ヘテロ接合度である。これに対し、集団  $i$  の集団固有  $F_{ST}$  は

$$psF_{ST}^i = \frac{H_B - H_{Si}}{H_B} = 1 - \frac{H_{Si}}{H_B}$$

で定義される。 $H_B$  は2集団からサンプルされたアレルが異なる確率である。長い歴史を持つ祖先集団は多型性も大きく、集団固有  $F_{ST}$  は小さいのに対し、新しくできた集団は集団固有  $F_{ST}$  が

<sup>2</sup> Wu, J., Yonezawa, T., and Kishino, H. (2017). Rates of molecular evolution suggest natural history of life history traits and a post-K-Pg nocturnal bottleneck of placentals. *Current Biology*, **27**: 3025–3033

<sup>3</sup> Živković, D. and Wiehe, T. (2008). Second-order moments of segregating sites under variable population size. *Genetics*, **180**(1): 341–357

<sup>4</sup> Hudson, R.R. and Kaplan, N.L. (1995). Deleterious background selection with recombination. *Genetics*, **141**: 1605–1617

大きい。そこで、集団対  $F_{ST}$  から推定される集団構造に集団固有  $F_{ST}$  の情報を載せることにより、現在の集団構造と集団の範囲拡大の歴史を理解することが容易になる。

#### (4) 集団の履歴に照らした多形質共適応の探索的解析

中立 SNPs の対立遺伝子頻度スペクトラムを集団間で比較することにより、集団の分化・混合の履歴を推定し、グラフ表現することができる<sup>5</sup>。このグラフ上における形質関連 SNPs の対立遺伝子頻度の遺伝的浮動を超えた変動として、方向性淘汰を検出することができる<sup>6</sup>。ただし、個体群は相互に影響し合う形質を変化させることで、その地域の環境に対応する。そこで本研究では、上の手法に多変量解析を組み合わせることで多形質共適応を探索的に検出する方法を開発した。

### 4. 研究成果

#### (1) 分子進化速度を予測因子とした祖先形質の復元と多形質共進化の検出

保存性の高い 1204 の遺伝子の分子進化速度を予測因子として哺乳類 89 種の祖先形質を推定し、これに基づき進化の履歴の形質間の偏相関をグラフ表現した (図 1)。その結果、体重と寿命の間には相関関係があることが知られているが、両者が妊娠期間と共進化することから来る間接的なものであることがわかった。また、昼行性哺乳類は夜行性哺乳類よりも一夫多妻制をとる傾向があり、樹上性哺乳類は陸上性哺乳類よりも子サイズが小さい傾向があることが示された。食性と季節繁殖行動の間の共進化が示唆されたため、結合マルコフモデルを適用して因果関係を推測した。その結果、通年繁殖は雑食性への食性変化に先行し、季節繁殖は肉食性への食性変化に先行することが示された。

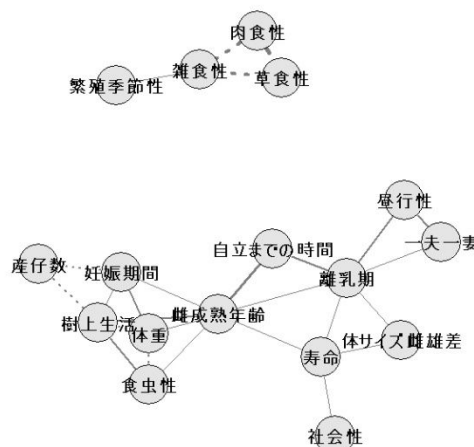


図 1 哺乳類における生活史と繁殖行動の共進化グラフ (Wu et al 2021).

#### (2) 種内の遺伝的多様度と分子進化速度

翻訳領域であるエクソンと 5'非翻訳領域、3'非翻訳領域、およびイントロンを対象に、1000bp 幅の窓の遺伝的多様度を網羅的に分析した。その結果、エクソンに働く選択圧が連鎖を通して、特に 5'非翻訳領域の遺伝的多様度を押し下げることがわかった。図 2 はヒトの 14,671 の遺伝子について、SNP の密度と分子進化速度の関係を調べたものである。青線は負の二項回帰による予測曲線で、色付きの箱ひげ図は、ヒトでは物質関連障害に関連する遺伝子の機能的制約が減少し、感染症に関連する遺伝子の機能的制約が強化されたことを示している。ヒト、チンパンジー、ゴリラ、マウス、イヌ・オオカミの集団ゲノムデータの比較分析から、種の形態的進化には、結合組織よりもむしろ上皮細胞の進化が主に寄与していることがわかった。

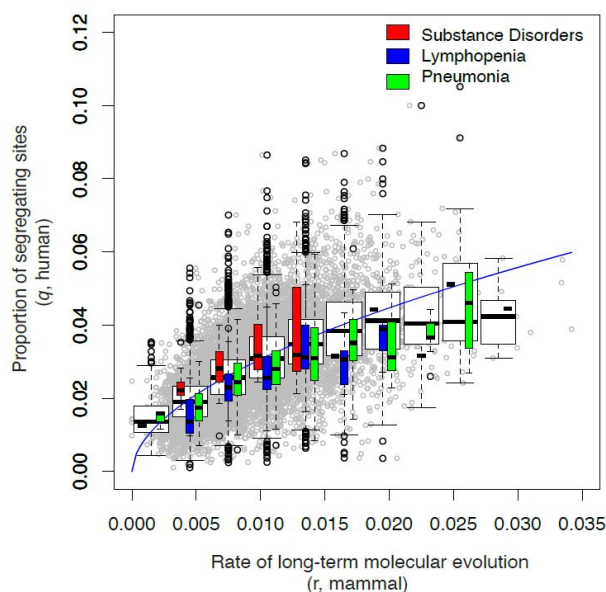


図 2 ヒトの遺伝的多様度と哺乳類分子進化速度 (Wu et al 2022).

#### (3) 集団構造と分布拡大の履歴

祖先集団からの集団植民を模したシミュレーションと、ヒト、大西洋タラ、野生ポプラの公開データを解析して、集団固有  $F_{ST}$  が祖先集団を特定し、そこから集団が派生する進化史を追跡で

<sup>5</sup> Pickrell, J. and Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, 8: e1002967

<sup>6</sup> Racimo, F., Berg, F. J., and Pickrel, J. K. (2018). Detecting polygenic adaptation in admixture graphs. *Genetics*, 208: 1565–1584.

きることが実証された。図3はヒト51集団について、集団対  $F_{ST}$  から得られる集団構造に集団固有  $F_{ST}$  の情報を付加している。アフリカから中央・南アジア、中東、ヨーロッパ、東アジア、オセアニア、アメリカへと分布拡大している様子が見て取れる。ちなみに階層ベイズモデルは、密に調査されているヨーロッパを祖先集団と推定してしまう。

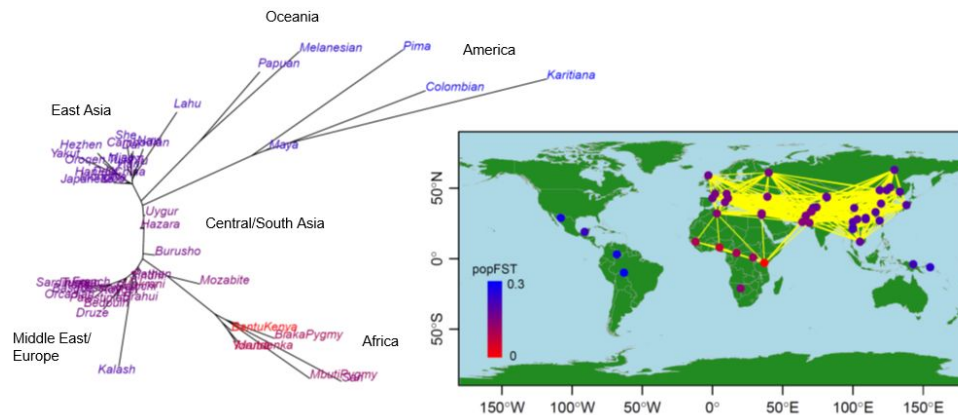


図3 ヒト51集団の集団構造と分布拡大の遺伝的痕跡 (Kitada *et al* 2021).

(4) 集団の履歴に照らした多形質共適応の探索的解析

一次元の飛び石型集団拡大の数値シミュレーションにより、個々の形質の分析では見逃されていた集団の環境適応のパターンが多形質の探索的分析により再現されることが示された。アメリカ北西部の野生ポプラの集団の公開データを解析した結果、多形質の適応の歴史を表現する SET (selection on edge-by-trait) 行列の主成分分析により、光合成関連形質と成長関連形質が共適応し、気温と日長の地理的なクラインにตอบสนองすることが明らかになった(図4)。また、体積関連形質の第2主成分の共適応から、土壌条件が地上部環境選択の制限因子であることが示唆された。

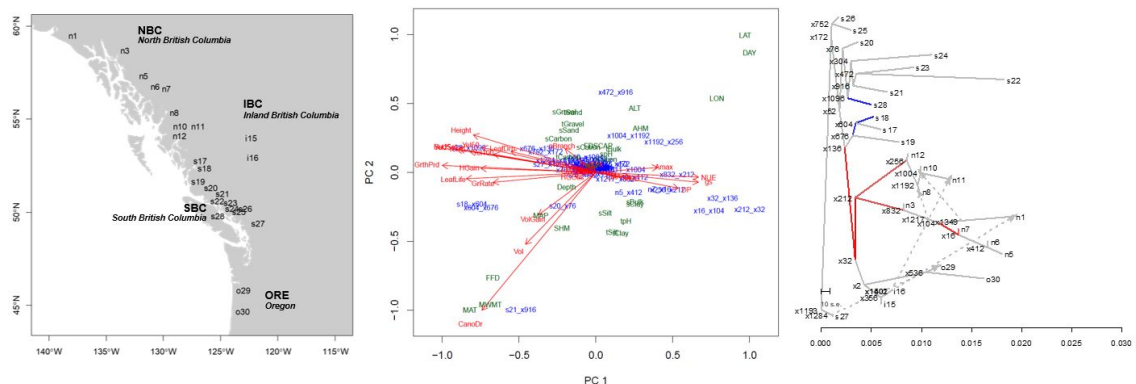


図4 アメリカ北西部に分布するポプラの分布拡大・混合と多形質共適応のマッピング (Nakamichi *et al* 2022).

## 5. 主な発表論文等

〔雑誌論文〕 計12件（うち査読付論文 12件/うち国際共著 12件/うちオープンアクセス 12件）

1. 著者名 Wu, J., Yonezawa, T., and Kishino, H.	4. 巻 13
2. 論文標題 Molecular evolutionary rate predicts intraspecific genetic polymorphism and species-specific selection	5. 発行年 2022年
3. 雑誌名 Genes	6. 最初と最後の頁 708
掲載論文のDOI（デジタルオブジェクト識別子） 10.3390/genes13040708	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Nakamichi, R., Kitada, S., and Kishino, H.	4. 巻 12
2. 論文標題 Exploratory analysis of multi-trait coadaptations in the light of population history	5. 発行年 2022年
3. 雑誌名 Ecology and Evolution	6. 最初と最後の頁 e8755
掲載論文のDOI（デジタルオブジェクト識別子） 10.1002/ece3.8755	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Dang, T. T. and Kishino, H.	4. 巻 6
2. 論文標題 Forward variable selection improves the power of random forest for high-dimensional microbiome data	5. 発行年 2022年
3. 雑誌名 Journal of Cancer Science and Clinical Therapeutics	6. 最初と最後の頁 87-105
掲載論文のDOI（デジタルオブジェクト識別子） 10.26502/jcsct.5079147	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する
1. 著者名 Kitada, S. and Kishino, H.	4. 巻 11
2. 論文標題 Population structure of chum salmon and selection on the markers collected for stock identification	5. 発行年 2021年
3. 雑誌名 Ecology and Evolution	6. 最初と最後の頁 13972-13985
掲載論文のDOI（デジタルオブジェクト識別子） 10.1002/ece3.8102	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

1. 著者名 Kitada, S., Nakamichi, R., and Kishino, H.	4. 巻 11
2. 論文標題 Understanding population structure in an evolutionary context: population-specific FST and pairwise FST	5. 発行年 2021年
3. 雑誌名 G3: Genes Genomes Genetics	6. 最初と最後の頁 jkab316
掲載論文のDOI (デジタルオブジェクト識別子) 10.1093/g3journal/jkab316	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Wu, J., Yonezawa, T., and Kishino, H.	4. 巻 12
2. 論文標題 Evolution of reproductive life history in mammals and the associated change of functional constraints	5. 発行年 2021年
3. 雑誌名 Genes	6. 最初と最後の頁 740
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/genes12050740	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Taniguchi, S., Bertl, J., Futschik, A., Kishino, H., and Okazaki, T.	4. 巻 12
2. 論文標題 Waves out of the Korean Peninsula and inter- and intra-species replacements in freshwater fishes in Japan	5. 発行年 2021年
3. 雑誌名 Genes	6. 最初と最後の頁 303
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/genes12020303	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Zhao, X-W. and Kishino, H.	4. 巻 11
2. 論文標題 Multiple isolated transcription factors act as switches and contribute to species uniqueness	5. 発行年 2020年
3. 雑誌名 Genes	6. 最初と最後の頁 1148
掲載論文のDOI (デジタルオブジェクト識別子) 10.3390/genes11101148	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Kitazoe Yasuhiro, Kishino Hirohisa, Tanisawa Kumpei, Udaka Keiko, Tanaka Masashi	4. 巻 18
2. 論文標題 Renormalized basal metabolic rate describes the human aging process and longevity	5. 発行年 2019年
3. 雑誌名 Aging Cell	6. 最初と最後の頁 1-11
掲載論文のDOI (デジタルオブジェクト識別子) 10.1111/ace1.12968	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Dang, T. and Kishino, H.	4. 巻 36
2. 論文標題 Stochastic variational inference for Bayesian phylogenetics: a case of CAT model	5. 発行年 2019年
3. 雑誌名 Molecular biology and evolution	6. 最初と最後の頁 825-833
掲載論文のDOI (デジタルオブジェクト識別子) 10.1093/molbev/msz020	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

〔学会発表〕 計7件 (うち招待講演 1件 / うち国際学会 1件)

1. 発表者名 Dang, T. and Kishino, H.
2. 発表標題 Phylogenetic variable selection for microbiome study
3. 学会等名 日本バイオインフォマティクス学会
4. 発表年 2020年

1. 発表者名 北田修一, 中道礼一郎, 岸野洋久
2. 発表標題 集団固有FSTと集団対FST: 歴史と環境圧.
3. 学会等名 日本水産学会
4. 発表年 2020年



1. 発表者名 中道礼一郎, 北田修一, 岸野洋久
2. 発表標題 集団の歴史と分化、環境適応の推測
3. 学会等名 日本水産学会
4. 発表年 2020年

1. 発表者名 吉田伸弥, 内山郁夫, 福世真樹, 矢原耕史, 桂有加子, 岸野洋久, 小林 一三
2. 発表標題 DNAメチル化を考慮した塩基置換モデルによる、メチロームの分子系統解析
3. 学会等名 日本ゲノム微生物学会
4. 発表年 2020年

1. 発表者名 Wu, J., Yonezawa, T., and Kishino, H.
2. 発表標題 Bridging molecular evolution and phenotypic evolution
3. 学会等名 日本哺乳類学会 (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 中道礼一郎, 北田修一, 岸野洋久
2. 発表標題 遺伝子・環境・形質の集団解析とグラフ表現
3. 学会等名 日本統計関連学会
4. 発表年 2019年

1. 発表者名 中道礼一郎, 北田修一, 岸野洋久
2. 発表標題 遺伝子・環境・形質の集団構造と適応進化のグラフ表現
3. 学会等名 日本進化学会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	北田 修一  (Kitada Shuichi)  (10262338)	東京海洋大学・学術研究院・名誉教授   (12614)	
研究分担者	中道 礼一郎  (Nakamichi Reiiichiro)  (70401255)	国立研究開発法人水産研究・教育機構・水研機構(横浜)・主任研究員   (82708)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------