

令和元年5月7日現在

機関番号：12601

研究種目：基盤研究(B) (一般)

研究期間：2016～2018

課題番号：16H02788

研究課題名(和文)次世代型分子データによる詳細分子生物地理の推定法の開発

研究課題名(英文)Statistical modeling of highly-resolved phylogeography based on molecular data of new generation

研究代表者

岸野 洋久(KISHINO, Hirohisa)

東京大学・大学院農学生命科学研究科(農学部)・教授

研究者番号：00141987

交付決定額(研究期間全体)：(直接経費) 12,200,000円

研究成果の概要(和文)：研究成果の概要(200字程度)
次世代型分子データは、種内の多様性と種間の多様性をシームレスに、高解像度で表現することを可能とする。私たちは、進化学、集団遺伝学、系統地理学の古典理論を拡張して要約統計量をモデリングする、というアプローチを提案した。アジア・オセアニアに広域に棲息する広食性昆虫ハスモンヨトウの集団ゲノム解析、哺乳類の多遺伝子系統樹と生活史の関連解析を通して手法の有効性を実証した。遺伝子流動の高い海洋生物などを想定し、高解像度に集団構造を検出する方法を提示し、プログラムパッケージを公開した。

研究成果の学術的意義や社会的意義

本研究は、人の構築した集団遺伝と分子進化の理論を拡張し、要約統計量を統計的にモデリングするというアプローチを開発した。多遺伝子分子進化速度の分散分析により分子進化と表現型進化を橋渡しするアプローチは、ゲノムと表現型の種間比較の新たな方法を提供する。ハスモンヨトウの分析で観察された広範囲に及ぶ遺伝子流動は、害虫管理方策の策定に有用な情報を提供する。また、哺乳類の生活史の分析から得られた、社会性や一夫一婦制では全般に機能的な制約が緩むこと、社会性については脳や神経系に携わる遺伝子が予測因子として抽出されるなどの結果は、行動の遺伝的背景を理解する上でも示唆を与えるものと思われる。

研究成果の概要(英文)：New-generation molecular data enables to represent the intra-species polymorphism and inter-species diversity seamlessly, and with high resolution. By extending classical theory of population genetics and molecular evolution, we developed a statistical model to express the summary statistics such as allele frequencies spectra and multi-gene phylogenetic trees. The population genomic study of *Spodoptera litura*, a polyphagous insect widely distributed in East Asia, identified a high gene flow along a South India-South China-Japan axis. We developed a novel method of ancestral state reconstruction of life history traits that associates the variation of trait values with the variation of molecular evolutionary rates. It inferred, consistent with fossils, that placentals were insectivores before Cretaceous-Paleogene boundary (65mya) and only nocturnal animals survived over the Eocene-Oligocene transition (EOT, 33.9 mya). These studies show clearly the validity of our approach.

研究分野：統計科学, 分子進化学, 集団遺伝学

キーワード：ゲノムビッグデータ 要約統計量 分子進化の中立説 アレル頻度スペクトル 多遺伝子系統樹 分散分析

様式 C-19、F-19-1、Z-19、CK-19 (共通)

1. 研究開始当初の背景

SNP (single nucleotide polymorphism, 1塩基多型)などのゲノムデータのコストが急速に縮小しており、今後この傾向はますます強くなっていく。これに伴い、集団遺伝の方法論も進化しているが、データの性格の質的な変化に十分に対応できていない。

集団を構成する個体の間には祖先を共有することから来る相関関係がある。このため、限られた遺伝子座を分析するときは、調査する個体数を増やしても推定の精度の向上は限定的となる。これに対して、ゲノム上で大きく離れた部位では連鎖不平衡が及ばず、祖先への合体のパターンが独立である。従って、大量のSNPsをタイピングすることは、時として多数の個体をサンプリングすること以上の効果を持つ。このことは個体サンプリングの自由度が格段に広がることを意味する。現在の集団遺伝学の分析では、通常各地から数十個体をサンプリングするのに対し、標本の総数を変えることなく、各地からの標本を極限的に抑えて時空間的に密なサンプリングを行い、詳細な分子生物地理学的知見を得ることへの可能性が開かれる。

2. 研究の目的

本研究の着想に至った背景には、岡崎登志夫博士、三田和英教授との出会いがある。岡崎博士は国立水産研究所に在職中30年にわたり、日本と韓国の数多くの河川から淡水魚ほぼ全種にわたる約80魚種をサンプリングし、ミトコンドリア分子系統樹に基づき、移入・拡散・塗り替わりについて仮説を練り上げて来られた。三田教授は1990年代からカイコ(*Bombyx mori*)のゲノムプロジェクトをリードして来られたが、近年同じ鱗翅目昆虫であるハスモンヨトウ(*Spodoptera litura*)のゲノムプロジェクトを推し進めている。前者は空間的に密なサンプリングにより得られたデータの代表例といてよく、後者は広域にわたるサンプリングにより得られたゲノムデータである。これらのデータを集中的に統計的モデル分析することを通して、次世代型分子データに基づく詳細分子生物地理の可能性を開く、強靱な方法論を開発することを目的とした。

3. 研究の方法

本研究は種内の多様性と種間の多様性をシームレスに描き出す方法を開発した。当初は種内の多様性の分析を目的としていたが、この研究を進める中で、次世代型分子データは、種内の遺伝的多型性の分析にとどまらず、種間の遺伝的多型性と表現型を関連付けるための分析にも革新をもたらす可能性があることに注目した。全世界のヒト集団は、10-20万年で共通祖先へ行きつく。突然変異が起きる確率は1世代あたりサイト当たり 10^{-8} 程度であるため、ゲノム中の各サイトにおける遺伝的多型の多くは、高々1回の突然変異に由来する。一方、ヒトとその最近縁種であるチンパンジーが分岐したのは400万年から1300万年前である。ゲノムの種間比較においては、集団中の変異よりもはるかに多い突然変異の蓄積を相手にするため、ゲノムワイド関連解析をそのまま行っても、ノイズに埋もれたシグナルを掘り起こすことは期待できない。

ビッグデータを扱うため、今後膨大なデータの中から探索的に埋もれた情報をあぶり出していくアプローチが盛んになるであろう。これに対して本研究では、第一段で粗削りではあるが有効性が期待される情報を要約統計量という形で算出し、第二段では進化学、集団遺伝学、系統地理学の古典理論を拡張して要約統計量をモデリングする、というアプローチをとる。集団遺伝学は、集団中に生じた新たなアレルがときに次第に集団を埋め尽くし、ときに集団から消え去る様を、またアレル頻度が空間的に不均質である様子を確率過程で記述することにより定量評価する。20世紀初頭にS. Wright, R.A. Fisher, J.B. Haldaneにより得られた確率モデルは、1遺伝子中に起きた突然変異の運命とその不確実性を記述するものであった。SNPデータは、こうして導入された確率過程から生成されるサンプルパスの無数のスナップショットと見ることができる。

SNPの密度とアレル頻度は、集団の履歴や遺伝子にかかる淘汰圧に関する情報を提供する。たとえば稀なアレルによるSNPが卓越していると、集団中に起きた突然変異の多くは比較的最近に起きたものであることが想定される。これは、何らかの要因で集団が極端に細り、近年になって膨張して急速に突然変異の機会を広げた、と解釈できる。あるいはこの遺伝子が強い純化圧を受けており、突然変異が集団中に拡大することが抑えられている、と解釈できる。1遺伝子を観察するのみでは、こうした集団の履歴の影響と遺伝子にかかる淘汰圧の効果を識別できない。ところが、集団の履歴の影響はゲノム全体に一律に及ぶのに対して、淘汰圧の効果は遺伝子の機能と関係するため、SNPごとに個別に働く。したがって、ゲノム中のSNPを総合的に分析することにより両者を識別することができるのである。

分子進化は集団に固定された突然変異を集めたもので、種間の遺伝的多様性を醸成する。分子進化・分子系統学の中心をなす分子進化の中立説(Kimura (1969))は、ゲノムの多型性が適応進化からの予測をはるかに超えることを強く示唆する観察結果から得られたもので、いまでは定説となっている。集団を塗り替えてきた突然変異のうち有利な突然変異の割合は無視できる

ほど少なく、大部分は周りとは差がないか、あるいは弱有害のものであると唱える。相同な配列を比較して分子系統樹を得ることにより、分子進化速度の変化を推測することができる。中立説の下では、分子進化速度は突然変異率と中立な変異の割合の積で表現される。中立な変異の割合は遺伝子に働く機能的な制約の強さと逆相関している。これまでは分子進化速度の変化が突然変異率の変化によるものか、あるいは遺伝子に働く機能的な制約の強さの変化によるものか、識別不能で、補足情報に基づき推測するほかに術はなかった。ところが、突然変異率は世代の長さや変異源への暴露率に依存し、それによる影響はゲノム全体に及ぶのに対し、機能的な制約は遺伝子ごとに個別に働くことに注意すると、多遺伝子系統樹を総合することにより、両者が識別可能になることに気づくのである。

4. 研究成果

アジア・オセアニアに広域に棲息する広食性の昆虫である、ハスモンヨトウ(*Spodoptera litura*)について、広食を可能とするゲノムの背景を調査し、集団ゲノム解析を行った。集団構造と遺伝子流動を精密推定し、インド南部から中国南部、日本にまたがり遺伝的に均質であり、本種が恐らくモンスーンに流され広範囲に移動することを強く示唆する結果を得た。Nature Ecology and Evolution に掲載され、おすすめコンテンツに取り上げられた。また、広範囲に混じり合う海洋生物において、水温や塩分などの環境や増養殖に影響されて生成される遺伝的な不均質性の微かなシグナルを高感度に検出する方法を提示し、プログラムパッケージを Molecular Ecology Resources に公開した。

研究を進める中で、当初の計画を超えて、分子進化と表現型の進化を結びつける新たな方法に着想した。多数の遺伝子の分子系統樹について枝・遺伝子の積型二元配置分散分析を行うことにより、遺伝子に働く機能的な制約の強さとその変動を抽出し、これを生活史と関連付けることにより、その遺伝的背景を推測する方法を開発した。その結果、暁新世から始新世 - 漸新世への移行期 (EOT, 33.9 mya) にかけて胎盤哺乳類の祖先は群れを作らず、繁殖には季節性があったこと、昆虫を捕食していたこと、そして夜行性動物のみが強い寒冷期である EOT を生き延びたことが推定された。また社会性や一夫一婦制では全般に機能的な制約が緩むこと、社会性については脳や神経系に携わる遺伝子が予測因子として抽出されるなど、行動の遺伝的背景を理解する上でも示唆を与える結果が得られた。Current Biology に掲載され、DNA 情報により哺乳類の祖先形質を推測する全く新しい手法として Peer Community in Evolutionary Biology に紹介された (Galtier and Chang 2017)。

さらに、ビッグデータに基づき、ゲノムが複雑に進化するパターンを推測することを可能にするために、新規に提案された理論に支えられた高速アルゴリズムを Molecular Biology and Evolution に発表した。マルコフ連鎖モンテカルロ法(MCMC)はシミュレーションにより事後分布を描き出すもので、ベイズ推測理論のデータへの適用において核をなしているが、計算負荷が大きく、ゲノムデータに基づく分子系統樹の解析ではネックとなっていた。そこで事後分布をモデリングし、Kullback-Leibler ダイバージェンス最小化によりパラメータ推定を行う変分ベイズの手法を導入した。構造の制約や機能的な制約を受けて、タンパク質は複雑で多様に進化する。進化のパターンは遺伝子により異なり、遺伝子内においてもサイトにより異なる。CAT モデル (Galtier et al 2004)はこの複雑な進化をディリクレ過程混合分布で記述するが、変分ベイズが MCMC による事後分布を極めてよく近似し、計算負荷を大幅に軽減することが示された。

5. 主な発表論文等

[雑誌論文] (計 15 件)

- (1) Dang, T. T. and Kishino, H. (2019). Stochastic variational inference for Bayesian phylogenetics: A case of CAT model. *Molecular Biology and Evolution*. **36**: 825–833. doi:10.1093/molbev/msz020
- (2) Kitada, S., Nakajima, K., Hamasaki, K., Shishidou, H., Waples, R., and Kishino, H. (2019). Rigorous monitoring of a large-scale marine stock enhancement program demonstrates the need for comprehensive management of fisheries and nursery habitat. *Scientific Reports*. **9**: 5290. doi:10.1038/s41598-019-39050-3
- (3) Kitada, S., Nakamichi, R., and Kishino, H. (2017). The empirical Bayes estimators of fine-scale population structure in high gene flow species. *Molecular Ecology Resources*. **17**: 1210–1222. doi: 10.1111/1755-0998.12663
- (4) Cheng, T, Wu, J., Wu, Y., Chilukuri, R., Huang, L., Yamamoto, K., Feng, L., Li, W., Chen, Z., Guo, H., Liu, J., Li, S., Wang, X., Peng, L., Liu, D., Guo, Y., Fu, B., Li, Z., Liu, C., Chen, Y., Tomar, A., Hilliou, F., Montagne, N., Jacquin-Joly, E., d'Alencon, E., Seth, R., Bhatnagar, R., Jouraku, A., Shiotsuki, T., Kadono-Okuda, K., Promboon, A., Smaghe, G., Arunkumar, K., Kishino, H., Goldsmith, M., Feng, Q., Xia, Q., and Mita, K. (2017). Genomic adaptation to polyphagy and insecticides in a major East Asian noctuid pest. *Nature Ecology and Evolution*. **1**: 1747-1756. doi: 10.1038/s41559-017-0314-4

- (5) 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. doi: 10.1016/j.cub.2017.08.043
- (6) Kitada, S., Yoshikai, R., Fujita, T., Hamasaki, K., Nakamichi, R., and Kishino, H. (2017). Population structure and persistence of Pacific herring following the Great Tohoku earthquake. *Conservation Genetics*. **18**: 423-437. doi: 10.1007/s10592-016-0918-2
- (7) Liao, H.-I., Chen, H. and Kishino, H., and Liao, C.-T. (2016). A reference-population-based conformance proportion. *Journal of Agricultural, Biological, and Environmental Statistics*. **21**: 684-697. doi: 10.1007/s13253-016-0268-z

[学会発表] (計 22 件)

- (1) Kishino, H. (2018). Bridging molecular evolution and phenotypic evolution. Invited seminar. University of Linz, Linz, Austria
- (2) Dang, T. and Kishino, H. (2018). Stochastic variational inference of mixture models in phylogenetics. 統計関連学会連合大会.
- (3) 中道礼一郎, 北田修一, 岸野洋久. (2018). 集団の遺伝的分化と環境適応のグラフィカルモデリング. 日本進化学会.
- (4) Wu, J., Yonezawa, T., and Kishino, H. (2018). Rates of molecular evolution suggest life history and a post-K-Pg nocturnal bottleneck of Placentals. Society of Molecular Biology and Evolution.
- (5) Dang, T. and Kishino, H. (2018). Fast approximate inference for phylogenetic reconstruction via stochastic variational inference in large data sets. Society of Molecular Biology and Evolution.
- (6) 岸野洋久. (2018). 分子進化速度が語るもの. 日本計量生物学会.
- (7) 中道礼一郎, 北田修一, 岸野洋久. (2018). 集団構造と環境因子のグラフィカルモデリング. 日本計量生物学会.
- (8) Wu, J., Yonezawa, T., and Kishino, H. (2017). Natural history of life history traits and a post-K-Pg nocturnal bottleneck of Placentals implicated by neutral theory. EMBL Conference Mammalian Genetics and Genomics.
- (9) 呉佳齊, 米澤隆弘, 岸野洋久. (2017). Natural history of life history traits and a post-KPg nocturnal bottleneck of Placentals: Implications from neutral theory. 統計関連学会連合大会.
- (10) 呉佳齊, 米澤隆弘, 岸野洋久. (2017). 分子進化速度が語る動物行動の進化. 日本行動計量学会.
- (11) 谷口昇志, 岡崎登志夫, Bertl, J., Futschik, A., 岸野洋久. (2017). 同一種内の淡水魚の個体群間の置き換わりで生じる地理的分断の検出. 日本進化学会.
- (12) 岡崎登志夫, 谷口昇志, 指田昌樹, 田祥麟, 岸野洋久. (2017). 分子系統と地誌から辿る日本の淡水魚類相の形成過程とそれに関与した大陸産魚種. 日本生態学会.
- (13) 谷口昇志, 指田昌樹, 岡崎登志夫, 田祥麟, Bertl, J., Futschik, A., 岸野洋久. (2017). 大陸からの侵入淡水魚による置き換りの推定: 分子系統地理的前向きシミュレーションによるアプローチ. 日本生態学会.
- (14) 呉佳齊, 岸野洋久. (2017). Gene-by-branch analysis of evolutionary rates predict times, traits and genes of ecological importance. 日本生態学会.
- (15) Kishino, H. (2017). Spatio-temporal modeling to measure viral adaptation and community diversity. Invited talk at International Conference on Bioinformatics and Biostatistics for Agriculture, Health and Environment. University of Rajshahi, Rajshahi, Bangladesh
- (16) 呉佳齊, 岸野洋久. (2016). Inter-specific genome wide association study and prediction of ecotype. 統計関連学会連合大会.
- (17) 岡崎登志夫, 谷口昇志, 田祥麟, 岸野洋久. (2016). 分子系統と地誌から見る淡水魚の大陸から日本への侵入パターンと国内での拡散・置き換り. 日本進化学会.
- (18) 谷口昇志, 岡崎登志夫, Jeon, S.-R., Bertl, J., Futschik, A., 岸野洋久. (2016). 淡水魚の侵入、拡散、置き換りの推定; ベイズ法及びシミュレーションをベースとしたアプローチ. 日本進化学会.

[図書] (計 1 件)

- (1) Mariadassou, M, Bar-Hen, A. and Kishino, H. (2019). Tree Evaluation and Robustness Testing. In *Encyclopedia of Bioinformatics and Computational Biology* (edited by Ranganathan, S., Nakai, K., Schönbach, C., and Gribskov, M.). Pp. 736-745. Academic Press.

6. 研究組織

(1)研究分担者

研究分担者氏名：北田修一

ローマ字氏名：KITADA, Shuichi

所属研究機関名：東京海洋大学

部局名：学術研究院

職名：名誉教授

研究者番号（8桁）：10262338

(2)研究協力者

研究協力者氏名：

ローマ字氏名：

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。