

## 科学研究費助成事業 研究成果報告書

平成 28 年 5 月 24 日現在

機関番号：12601

研究種目：基盤研究(B) (一般)

研究期間：2013～2015

課題番号：25280006

研究課題名(和文) ネットワークの許容領域による生物・生態系の健全性の評価手法の開発

研究課題名(英文) Development of a graphical model and a tolerance region for the assessment of physiological process and ecological diversity

研究代表者

岸野 洋久 (Kishino, Hirohisa)

東京大学・農学生命科学研究科・教授

研究者番号：00141987

交付決定額(研究期間全体)：(直接経費) 13,100,000円

研究成果の概要(和文)：本研究では第一に、遺伝子発現プロファイルのデータからエントロピー極大連結グラフを構築する方法を開発した。タイセイヨウサケの初期発生過程の発現プロファイルを解析したところ、鍵遺伝子が発生の段階とともに発現していく様子が明らかになった。第二に、被覆確率の上限と下限を保証する相補的な許容区間を導入し、対象とする集団の分布が参照母集団における異常値を含む確率を統計的に見積もるリスク評価の手法を開発した。第三に、分子系統樹から得られる分岐年代の分布に基づき、生物群集の種構成の系統的歪みを測る指標を開発した。シミュレーション、台湾周辺海域の魚群集、世界諸海域のサメ群集の解析により有効性を確認した。

研究成果の概要(英文)：First, we developed a method that constructs the entropy-based maximal core connected subgraph of gene expression profiles. The analysis of the expression profiles at the early developmental stage of Atlantic salmon showed that the resultant graph makes it easy to understand the key players and their interactions. Second, we defined two complementary tolerance intervals. The coverage probability of the standard tolerance interval is at least the pre-specified level, whereas that of the complementary tolerance interval is at most the specified level. With the two complementary tolerance intervals, we developed a statistical procedure to test a target population includes a pre-specified amount of outliers with respect to the reference population. Third, we developed a new index of community diversity. By comparing the species sampling proportion with the value estimated assuming random sampling, it measures phylogenetic skew of member species.

研究分野：統計科学、分子進化学

キーワード：医薬生物・ゲノム解析 生体と生態系の健全性 遺伝子ネットワーク 分子系統樹 リスク評価

1. 研究開始当初の背景

私たちは都市化、医療・育種とバイオテクノロジーなど、何らかの形で生物と生態系に選択圧を加え、かつてないスピードで彼らを改変している。その結果、病原菌の脅威、新薬・新品種の副作用、安全性を定量的に評価し、生態系の多様性を感度よくモニタリングすることが急務となってきている。生物と群集の健全性を偏りなく、定量的に評価し、生物学者が解釈可能な形で比較分析の結果を提示する枠組みが、強く求められている。

2. 研究の目的

本研究では、生体内の複雑な生理現象や生態系における複雑な種間相互作用を踏まえて、リスク評価の手法を開発することを目的とした。核となるネットワーク構造を検出し、その変化をモニタリングすることが鍵となる。第一に、遺伝子発現ネットワークにおいて鍵表現型と遺伝子発現ネットワークから注目する形質に関連する情報量を最大化させる連結成分を抽出し、処置の効果を総合評価する統計量を開発することを目指した。そして第二に、許容領域の実質同等性を検定する方法を構築することを目指した。第三に、生物群集の遺伝的多様性とその時空間変動を定量的に評価するために、サンプリング方式を考慮に入れた基準化多様度指数を開発し、季節変動パターンに揺らぎを導入したベイズ分散分析モデルを開発することを目指した。

3. 研究の方法

(1) エントロピー極大連結有向グラフの推定

本研究では遺伝子発現プロファイルのデータに基づき、注目する形質に直接かかわる遺伝子群を統計的に抽出し、その間の関連を推定する。遺伝子間の相互作用やそれらと形質との関連は節と枝からなるグラフを用いて表現される。時系列的前後関係や転写因子データベースなどの補助情報を利用し、有向非巡回グラフを考える。一般にグラフ  $(V, E)$  の尤度は

$$L(V, E) = \prod_{v \in V} P(X_v | \{X_{v'} : v' \in pa(v)\})$$

と表すことができる。ここで  $pa(v)$  は  $v$  の親節全体である。形質節を含む連結部分グラフで AIC を最小化するグラフ  $(V_0, E_0)$  を遺伝アルゴリズムで探索する。遺伝子間の相互作用を重回帰で表現した場合は、対数尤度は

$$\log L = -\frac{n}{2} \sum_{v \in V} \log(1 - R_v^2)$$

と書ける ( $R_v^2$  は重相関係数)。グラフを構成する節とこれに関わる節の間の相互情報量

を足し合わせた形をしている。従って、AIC 最小化により選択された部分グラフは、推定の不確実性を加味して有意な相互情報量で結ばれた遺伝子を集めたものになっている。こうして抽出された部分グラフを通して、表現型の背後にある生理現象について、直接的な解釈を得ることが容易となる。

(2) 相補的許容区間とリスク評価

新品種、新規食品の安全性評価の問題を、分布の許容区間に関する統計的検定問題として定式化する。たとえば 95% 許容区間は集団中の 95% が含まれる範囲として定義される。標本に基づく従来型  $(1 - \gamma, 1 - \alpha)$  許容区間  $TI_0$  は、 $1 - \alpha$  の確率で集団中の  $1 - \gamma$  以上を包含する区間で、たとえば標本が遺伝・環境を変量効果に持つ正規混合モデルで記述される場合、

$$TI_0 = \left( \hat{\mu} - \hat{\sigma} \sqrt{\frac{m \cdot \chi_{1-\gamma}^2(\text{ncp} = 1/R_0)}{\chi_{m;\alpha}^2(\text{ncp} = 0)}}, \hat{\mu} + \hat{\sigma} \sqrt{\frac{m \cdot \chi_{1-\gamma}^2(\text{ncp} = 1/R_0)}{\chi_{m;\alpha}^2(\text{ncp} = 0)}} \right)$$

と定義する。ここで、 $R_0$  は誤差分散  $\sigma^2$  と  $\hat{\mu}$  の分散の比、 $m$  は  $2(\hat{\sigma}^2)^2$  と  $\sigma^2$  の推定量の分散の比である。これに対して本研究では、新たにこれと相補的な許容区間

$$TI_1 = \left( \hat{\mu} - \hat{\sigma} \sqrt{\frac{m \cdot \chi_{1-\gamma}^2(\text{ncp} = 1/R_0)}{\chi_{m;1-\alpha}^2(\text{ncp} = 0)}}, \hat{\mu} + \hat{\sigma} \sqrt{\frac{m \cdot \chi_{1-\gamma}^2(\text{ncp} = 1/R_0)}{\chi_{m;1-\alpha}^2(\text{ncp} = 0)}} \right)$$

を定義し、目標集団  $POP_{tar}$  と対照集団  $POP_{ref}$  について、対立仮説

$$H_1: TI(\gamma_{tar}, POP_{tar}) \subset TI(\gamma_{ref}, POP_{ref})$$

に対して帰無仮説

$$H_0: TI(\gamma_{tar}, POP_{tar}) \not\subset TI(\gamma_{ref}, POP_{ref})$$

を検定する枠組みを提示する。検定統計量を

$$\alpha_{\pi 01} = \alpha_{\pi 01}(X_{tar}, X_{ref}) = \arg \min \{ \alpha \mid TI_0(\alpha/2, \gamma_{tar}, X_{tar}) \subset TI_1(\alpha/2, \gamma_{ref}, X_{ref}) \}$$

で定義する。これにより、目標集団の中に、対照集団の分布に照らし合わせると異常値となるものが無視できない程度含まれないか、統計的に検定することが可能となる。

(3) 群集の種組成の系統的歪み

さまざまな種が生物群集を形作る。この群

集の多様性を測る尺度として、種数などの種豊度と均等度がある。また系統的な広がりや加味する。本研究では、対象となる群集の構成種の分岐年代の分布を、メタ群集の種プールの分岐年代の分布と比較する。対象となる群集はメタ群集から種を間引くことにより得られた形になっていることから、群集の分岐年代はメタ群集の分岐年代に比して相対的に古くなる傾向がある。そこで、この傾向の大きさを対象となる群集がメタ群集からの無作為標本である場合の期待値と比較することにより、種組成の分子系統的歪みの大きさを表す尺度が得られる。

分岐年代  $\mathbf{t} = (t_1, \dots, t_{s-1})$ ,  $t_1 > \dots > t_{s-1}$  の尤度は種分化率  $\lambda$ , 種絶滅率  $\mu$ , 種抽出率  $\rho$  を用いて

$$L_0(\mathbf{t} | \lambda, \mu, \rho) = (s-2)! \prod_{j=2}^{s-1} \frac{\lambda \rho_1(t_j)}{v_{t_j}}$$

と表される。ここで

$$\rho_1(t) = \frac{1}{\rho} P(0, t)^2 e^{(\mu-\lambda)t}$$

$$P(0, t) = \frac{\rho(\lambda - \mu)}{\rho\lambda + (\lambda(1-\rho) - \mu)e^{(\mu-\lambda)t}}$$

$$v_t = 1 - \frac{1}{\rho} P(0, t)e^{(\mu-\lambda)t}$$

である。3つのパラメータ  $\lambda$ ,  $\mu$ ,  $\rho$  は識別不可能で、上記尤度は純多様化率  $\theta_1 \equiv \lambda - \mu$ , および見かけの種分化率  $\theta_2 \equiv \lambda\rho$  の関数  $L_0(\mathbf{t} | \theta_1, \theta_2)$  として表現できる。群集の構成種はそれらを包括するメタ群集の種プールからランダムサンプリングされたものであるとすると、各群集の見かけの種分化率はメタ群集の表面種分化率に種抽出率を乗じたものとなる。従って、 $K$  個の群集とそれらを包括するメタ群集について、構成種の分岐年代  $\mathbf{t}_{\text{meta-comm}}, \mathbf{t}_{\text{comm}}^{(1)}, \dots, \mathbf{t}_{\text{comm}}^{(K)}$  の複合尤度

$$L(\mathbf{t}_{\text{meta-comm}}, \mathbf{t}_{\text{comm}}^{(1)}, \dots, \mathbf{t}_{\text{comm}}^{(K)} | \theta_1, \theta_2, \rho_E^{(1)}, \dots, \rho_E^{(K)}) \\ = L_0(\mathbf{t}_{\text{meta-comm}} | \theta_1, \theta_2) \prod_{k=1}^K L_0(\mathbf{t}_{\text{comm}}^{(k)} | \theta_1, \theta_2 \times \rho_E^{(k)})$$

を得る。種抽出率の最尤推定量は構成種が種プールからの無作為標本であることを仮定していることに注意すると、これを観察値に対する比の形で相対化した統計量は種組成の系統的偏りを表現することになる。種組成が分子系統樹上いくつかのクラスターに偏る群集は大きな値をもち、逆に種々のクラスターにわたって万遍なく存在する群集は小さな値を持つ。従って、この尺度は系統的歪み(phylogenetic skew)を表す指数とみなすこ

とができる。そこでこの尺度を PS と名付ける。

#### 4. 研究成果

##### (1) 大西洋サケの発生初期における発現プロフィール

タイセイヨウサケの受精直後から孵化し稚魚が浮上するまでの発現プロフィールのデータを解析し、エントロピー極大連結グラフを推定したところ、147 遺伝子が抽出された。最初期には転写制御とエネルギー代謝、細胞成長の遺伝子および腫瘍抑制の免疫遺伝子が発現しており、体組織の形成に合わせて筋肉・骨格形成遺伝子が発現し、血液合成の活発化を経て、鱗が形成され、消化管の活動が開始し、さらに孵化して外界に接すると同時に免疫遺伝子が活発に働き始める様子が読み取れた。

##### (2) 2つの許容区間によるリスク評価

南北アメリカの 47 環境における 9 年にわたる組換え・非組換えトウモロコシの圃場試験データについて、線形混合モデルで栄養素の遺伝効果、環境効果、交互作用の分散成分を求め、許容区間を比較した。また相補的な許容区間の特性を調べた。提案する検定手法が偏りを持たないことを示し、検出力を求めた。これを下に、日本のコメで作付を異にする 14 品種を遺伝効果とみなしてたんぱく質含量を九州と他地域で比較したところ、p 値は 0.195 であった。

##### (3) 確率変動する季節パターン

野生生物の個体群に見られる季節変動のパターンは必ずしも固定的ではなく、確率的に年変動する。たとえば沿岸魚類の生息数は海流の接岸パターンに依存するため、海流の蛇行パターンの変動に伴い、接岸時期が前後するのである。そこで、この揺れを事前分布として導入した階層ベイズ ANOVA モデルを開発した。台湾北部の主要魚種であるハリセンボン個体数の月次データを解析し、季節成分の年変動と年次効果を推定した。

##### (4) 系統的歪みの尺度 PS の性能評価

シミュレーションにより、既存の尺度とは異なり、新規に提案された PS は種数の大きさによらないこと、種組成の系統的歪みのパターンを感度よく表現することが確認された。台湾北部の原子力発電所において 30 年にわたり継時的に調査された魚群集標本からは、種数の直線的な減少が観察されたが、これに伴い分子系統的歪みが増大していることがわかった。夏季における優占種が近年見られなくなっていることから、温暖化の影響

響が示唆された。

#### (5) 世界各海域のサケ群集

新規に開発した PS に基づき世界中の海域におけるサメの種組成を納めたデータベースを解析したところ、赤道海域で特異的に大きな系統的歪みが観察された。そこで種組成を詳細に調べたところ、赤道海域では太平洋、大西洋、インド洋の広範にわたり共通して、メジロザメ科 (Carcharhinidae)、シュモクザメ科 (Sphyrnidae)、ジンベイザメ (Rhincodon typus) が生息していることがわかった。群集構造の特徴抽出のための探索的アプローチに PS が有効であることが示された。分子系統樹の統計的モデリングを通して群集構造を捉えるアプローチがさらに発展していくことが期待される。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 22 件)

- ① Chen, H., Nagai, S., and Kishino, H. (2016). Assessment of the network of protected areas for birds in Taiwan with regard to functional and phylogenetic diversity. *Pacific Conservation Biology*. **22**: 61-71. 査読有, DOI: <http://dx.doi.org/10.1071/PC15046>
- ② Lee, H-J., Kishino, H., Rodrigue, N., and Thorne, J. L. (2016). Grouping substitution types into different relaxed molecular clocks. *Philosophical Transactions of Royal Society B*. 査読有, accepted.
- ③ Chen, H. and Kishino, H. (2015). Hypothesis testing of inclusion of the tolerance interval for the assessment of food safety. *PLoS ONE*. **10(10)**: e0141117. 査読有, DOI: [10.1371/journal.pone.0141117](https://doi.org/10.1371/journal.pone.0141117)
- ④ Chen, H. and Kishino, H. (2015). Global pattern of phylogenetic species composition of shark and its conservation priority. *Ecology and Evolution*. **5(19)**: 4455-4465. 査読有, DOI: [10.1002/ece3.1724](https://doi.org/10.1002/ece3.1724)
- ⑤ Chen, H., Shao, K-T., and Kishino, H. (2015). A Bayesian hierarchical ANOVA model of stochastic seasonality for *Diodon Holocanthus* in northern Taiwan. *the Journal of Marine Science and Technology*. 査読有, DOI: [10.6119/JMST-015-0428-1](https://doi.org/10.6119/JMST-015-0428-1)
- ⑥ Miyashita, S., Ishibashi, K., Kishino, H., and Ishikawa, M. (2015). Viruses roll the dice: The stochastic behavior of viral genome molecules accelerates viral adaptation at the cell and tissue levels. *PLoS Biology*. **13(3)**: e1002094. 査読有, DOI: [10.1371/journal.pbio.1002094](https://doi.org/10.1371/journal.pbio.1002094)
- ⑦ Chen, H., Shao, K-T., and Kishino, H. (2015). Phylogenetic skew: an index of community diversity. *Molecular Ecology*. **24**: 759-770. 査読有, DOI: [10.1111/mec.13064](https://doi.org/10.1111/mec.13064)
- ⑧ Chen, H., Yamagishi, J., and Kishino, H. (2014). Bayesian inference of baseline fertility and treatment effects via a crop yield-fertility model. *PLoS ONE*. **9 (11)**: e112785. 査読有, DOI: [10.1371/journal.pone.0112785](https://doi.org/10.1371/journal.pone.0112785)
- ⑨ Hong, B., Fisher, T., Sult, T., Maxwell, C., Mickelson, J., Kishino, H., and Locke, M. (2014). Model-based tolerance intervals derived from cumulative historical composition data: Application for substantial equivalence assessment of a genetically modified crop. *Journal of Agricultural and Food Chemistry*. **62**: 9916-9926. 査読有, DOI: [10.1021/jf502158q](https://doi.org/10.1021/jf502158q)
- ⑩ Koyano, H., Tsubouchi, T., Kishino, H., and Akutsu, T. (2014). Archaeal  $\beta$  diversity patterns under the seafloor along geochemical gradients. *Journal of Geophysical Research G*. **119**: 1770-1788. 査読有, DOI: [10.1002/2014JG002676](https://doi.org/10.1002/2014JG002676)
- ⑪ Nakajima, K., Kitada, S., Habara, Y., Sano, S., Yokoyama, E., Sugaya, T., Iwamoto, W., Kishino, H., and Hamasaki, K. (2014). Genetic effects of marine stock enhancement: a case study based on the highly piscivorous Japanese Spanish mackerel. *Canadian Journal of Fisheries and Aquatic Sciences*. **71**: 1-14. 査読有, DOI: [10.1139/cjfas-2013-0418](https://doi.org/10.1139/cjfas-2013-0418)
- ⑫ Watabe, T. and Kishino, H. (2013). Spatial distribution of selection pressure on a protein based on the hierarchical Bayesian model. *Molecular Biology and Evolution*. **30**: 2714-2722. 査読有, DOI: [10.1093/molbev/mst151](https://doi.org/10.1093/molbev/mst151)
- ⑬ Kitada, S., Fujikake, C., Asakura, Y., Yuki, H., Nakajima, K., Vargas, K. M., Kawashima, S., Hamasaki, K., and Kishino, H. (2013). Molecular and morphological evidence of hybridization between native *Ruditapes philippinarum* and the introduced *Ruditapes* form in Japan. *Conservation Genetics*. **14**: 717-733. 査読有, DOI: [10.1007/s10592-013-0467-x](https://doi.org/10.1007/s10592-013-0467-x)

[学会発表] (計 34 件)

- ① Chen, H., Nagai, S., and Kishino, H. Assessment of protected area networks for bird biodiversity in Taiwan. 日本生態学会, 2016.3.20~2016.3.24, 仙台国際センター, 宮城県仙台市
- ② Kishino, H. On the adaptive mutations on the viral genomes and genetic diversity of

- communities. Invited seminar. 2016年3月10日. 重慶大学 (中国重慶市)
- ③ Kishino, H. Spatio-temporal modeling to measure viral adaptation and community diversity. Invited seminar. 2015年11月26日. University of Linz, Linz, Austria
- ④ Kishino, H. On the adaptive mutations on the viral genomes and genetic diversity of communities. Invited seminar. 2015年11月24日. Vienna Graduate School of Population Genetics, Vien, Austria
- ⑤ 岸野洋久. 日本進化学会学会賞・木村賞受賞講演. 日本進化学会. 2015年8月20日～2015年8月23日. 中央大学. 東京都文京区
- ⑥ Nakamichi, R., Kishino, H., and Kitada, S. The Core gene expression cascade during the early developmental process identifies genetic mechanisms for salmonids captive breeding. the American Fisheries Society, 2015.8.16～2015.8.20, Portland, U.S.A.
- ⑦ 陳虹諺・岸野洋久. Global pattern of phylogenetic species composition of shark and its conservation priority. 日本水産学会. 2015年3月27日～2015年3月31日. 東京海洋大 (東京都港区)
- ⑧ 中道礼一郎・岸野洋久・北田修一. ベニザケの産卵回帰における遺伝子発現ネットワークと上流/下流産卵群の環境適応の系統差. 日本水産学会. 2015年3月27日～2015年3月31日. 東京海洋大 (東京都港区)
- ⑨ 陳虹諺・長井敏・漆崎慎吾・山本圭吾・及川寛・安池元重・中村洋路・田邊晶史・本郷悠貴・藤原篤志・岸野洋久. Algal bloom and its phylogenetic aggregation. 日本計量生物学会. 2015年3月12日～2015年3月13日. 京都大学 (京都府京都市)
- ⑩ 陳虹諺・岸野洋久. Hypothesis testing of inclusion of the tolerance interval. 統計関連学会連合. 2014年9月13日～2014年9月16日. 東京大 (東京都文京区)
- ⑪ 渡部輝明・岸野洋久. インフルエンザヘマグルチニンタンパク質上における選択圧の分布と変遷の検出. 日本進化学会. 2014年8月21日～2014年8月24日. 高槻現代劇場 (大阪府高槻市)
- ⑫ 陳虹諺・岸野洋久. A high phylogenetic skew of sharks in the tropical sea. 日本進化学会. 2014年8月21日～2014年8月24日. 高槻現代劇場 (大阪府高槻市)
- ⑬ Chen, H-Y., Yamagishi, J., and Kishino, H. Baseline fertility connects crop yield and soil. International Biometric Society. 2014年7月6日～2014年7月11日. Florence, Italy
- ⑭ Chen, H-Y., Shao, K-T., and Kishino, H. Phylogenetic skew: A new index of community diversity. the joint annual meeting of the Society for the Study of Evolution (SSE), the Society of Systematic Biologists (SSB), and the American Society of Naturalists (ASN). 2014年6月20日～2014年6月24日. Raleigh, USA.
- ⑮ 中道礼一郎, 岸野洋久, 北田修一. タイセイヨウサケの初期発生における遺伝子発現ネットワークにみられる養殖の影響. 日本水産学会. 2014年3月27日～2014年3月31日. 北海道大学 (北海道函館市)
- ⑯ 陳虹諺, 岸野洋久. A new index of effective community size. 統計関連学会連合. 2013年9月8日～2013年9月11日. 大阪大学 (大阪府吹田市)
- ⑰ Kishino, H., Watabe, T., Nakamichi, R., and Kitada, S. Spatiotemporal modeling to measure the effects of mutations and selection pressures. International Statistical Institute (invited talk) 2013年8月25日～2013年8月30日. Hong Kong Convention and Exhibition Centre (Hong Kong, China)
- ⑱ 陳虹諺, 岸野洋久. Effective community size and the effect of weighted species sampling. 日本進化学会. 2013年8月28日～2013年8月31日. 筑波大学 (茨城県つくば市)
- ⑲ Watabe, T. and Kishino, H. Spatial distribution of selection pressure on a protein based on the hierarchical Bayesian model. Society for Molecular Biology and Evolution. 2013年7月7日～2013年7月11日. Hyatt Regency Chicago (Chicago, USA)
- ⑳ 中道礼一郎, 岸野洋久, 北田修一. サケ科魚類の生活史における遺伝子発現コアネットワーク: 最大エントロピー法による極大連結有向グラフの推定. 日本計量生物学会. 2013年5月23日～2013年5月24日. パルセイイざか (福島県福島市)

## 6. 研究組織

### (1) 研究代表者

岸野 洋久 (Kishino, Hirohisa)  
 東京大学・農学生命科学研究科・教授  
 研究者番号: 00141987

### (2) 研究分担者

北田 修一 (Kitada, Shuichi)  
 東京海洋大学・海洋科学技術研究科・教授  
 研究者番号: 10262338