

平成 30 年 6 月 12 日現在

機関番号：14401

研究種目：若手研究(B)

研究期間：2016～2017

課題番号：16K21624

研究課題名(和文)機械学習による上皮成長因子シグナル伝達の応答不均一性の推定

研究課題名(英文)Predicting heterogeneity in EGF signaling by machine learning

研究代表者

岩本 一成 (Iwamoto, Kazunari)

大阪大学・たんぱく質研究所・助教

研究者番号：70619866

交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)：細胞ごとに外部刺激に対する応答が異なる。応答不均一性は、細胞のがん化や薬剤耐性等に關与しているが、この機序に関しては未だ不明な点が多い。本課題では、上皮成長因子(EGF)シグナル伝達経路の新規数理モデル解析およびサポートベクターマシンによるERK応答予測を行い、EGF受容体およびRas発現量の変化がERK核内移行レベルを大きく変化させる事、また、これらの発現量から精度良くERK応答を予測できる事を明らかにした。これらの結果から、EGF刺激時のERK核内移行応答は、EGFシグナル伝達に關与するタンパク質、とくにEGF受容体およびRas発現量により制御されている可能性が示唆された。

研究成果の概要(英文)：Cellular heterogeneity, which means different response to same external stimuli in individual cells, is involved in malignant transformation and drug resistance, however, the detailed mechanism has been poorly understood. In this study, we developed a method integrating mathematical modeling with machine learning to elucidate the mechanism that generates cellular heterogeneity in nuclear translocation of ERK observed in epidermal growth factor (EGF) signaling pathway. Our mathematical analysis of EGF signaling pathway model and prediction of ERK response by support vector machine demonstrated that nuclear translocation of ERK after EGF stimulation at a single cell level could be predicted from expression levels of EGF receptor and Ras protein in individual cells. These results suggest that EGF receptor and Ras would determine the nuclear ERK translocation response under EGF treatment.

研究分野：システム生物学

キーワード：数値シミュレーション EGFシグナル伝達 応答不均一性

### 1. 研究開始当初の背景

細胞のシグナル伝達応答や様々な遺伝子の発現は、遺伝的に同一な細胞集団で、かつ均一な条件下においても、細胞ごとに異なる。これは、“応答不均一性”と呼ばれる現象で、例えば、大腸菌の増殖時の遺伝子発現 (Elowitz et al., Science, 2002) や細胞死誘導因子によるアポトーシス応答 (Spencer et al., Nature, 2009) また、免疫応答で中心的役割を果たす NF- $\kappa$ B の応答 (Shinohara et al., Science, 2014) 等、これまでに様々な系で観察されている。近年、このような応答不均一性は、細胞のがん化や分化、また薬剤耐性に関係する重要な現象であると考えられ始めた。特に、がん細胞や病原菌の薬剤耐性の克服は医学的に重要な課題の一つである。ところが、応答不均一性がどのように生み出され制御されているのか？その発生機序には未だ不明な点が多い。そのため、応答不均一性を生み出す細胞間の“ばらつき”を推定することが出来れば、上述のような医学・生物学的な重要な問題の解決につながる事が期待されている。

### 2. 研究の目的

本研究課題では、上皮成長因子 (EGF) シグナル伝達経路を対象に応答不均一性の制御機序の研究を進めた。EGF シグナル伝達経路は、細胞増殖や分化、アポトーシス等を制御する重要なシグナル伝達経路の一つである。EGF リガンドが細胞膜上の EGF 受容体に結合し、Ras、Raf、MEK、ERK タンパク質を順次活性化する。活性化した ERK タンパク質は最終的に細胞質から核内へと移行し、一連の転写因子の活性を制御する。それらは、細胞増殖や分化等に関与する多くの遺伝子の発現を制御する。これまでに、EGF シグナル伝達経路における ERK の核移行応答が応答不均一性を示す事が明らかにされてきた (Shindo et al., Nat. Commun. 2017)。そこで、この核内移行の応答不均一性の発生機序解明を目的に、数理モデリングと近年発展してきた人工知能関連技術、特に機械学習を組み合わせた手法を開発した。

### 3. 研究の方法

#### (1) EGF シグナル伝達経路の新規数理モデルの構築

これまで構築された EGF シグナル伝達経路の数理モデルでは、ERK 応答不均一性を再現できなかったため、新規に EGF シグナル伝達経路モデルを構築した。Hornberg らが構築した EGF 数理モデル (Hornberg et al., 2006) をベースに、ERK タンパク質の細胞質・核内移行プロセスおよび ERK による自身の移行制御プロセスを追加した。最終的に、数理モデルは、78 個の化学種および 152 個の反応プロセスを含む連立微分方程式によ

り記述された。数値シミュレーションは、CVODE を用いて行い、実験データ (Shindo et al., Nat comm. 2017) を再現するようにマニュアルでパラメータを調整した。実験データは、PC12 細胞を EGF で刺激した際の ERK タンパク質の核内移行をライブセルイメージングで計測したもので、EGF 刺激後のタイムコースと EGF の刺激濃度を変更したときの ERK 核内移行応答 (用量応答) のデータを使用した。

#### (2) ERK 応答不均一性の数値シミュレーション

応答不均一性のシミュレーションを行うにあたり、本研究では 2 種類のばらつきを考慮した (反応ノイズおよび細胞間でのタンパク質量の違い)。反応ノイズは、化学反応に関わる分子数に起因するものであり、分子数が少ないほど化学反応がより確率的になる。これは、確率シミュレーションにより再現可能で、その 1 つである Gillespie 法を使用した。一方、細胞ごとの各タンパク質の量の違いは、数理モデルの各化学種の初期値をシミュレーションごとに変更することで再現した。これらのばらつきの大きさを様々変更し、実験で観察された ERK 核内移行の応答不均一性と一致する適切なばらつきの大きさを推定した。ERK 核内移行の不均一性は、ERK の核内移行レベルの変動係数 (CV) を用いて評価し、実験データとシミュレーション結果の比較は、変動係数および相互情報量を用いて行った。

#### (3) 機械学習による応答不均一性の予測

上記数値シミュレーションデータを用いて、ERK 核内移行応答が予測可能かをサポートベクターマシン (SVM)、決定木、線形判別、主成分分析等の手法で検討した。分類精度は交差検証により評価し、どのような指標を入力にすることで、ERK 核内移行応答が予測できるかを調べた。

### 4. 研究成果

#### (1) 数理モデル解析による ERK 応答不均一性の発生機序解明

まず、反応ノイズおよびタンパク質量の細胞間での変化により、ERK 核内移行応答が不均一になるかを確認した。両ノイズのばらつきを変え、数値シミュレーションを行ったところ、反応ノイズは ERK 核内移行にほぼ影響をおよぼさなかった。詳細にモデルの解析を行った結果、数理モデルに含まれる各化学種の分子数が 1,000~1,000,000 と非常に多く、反応ノイズはほとんど発生していなかった。この分子数は実験値から推定したものであるため、EGF シグナル伝達経路では分子数に起因する反応ノイズはほとんど ERK 核内移行の応答不均一性に寄与していないことが考えられた。一方、各タンパク質

量の変化は、ERK 核内移行レベルを大きく変化させた。これらの結果から、EGF 刺激後の ERK 核内移行の応答不均一性は、細胞間での各タンパク質量の違いに起因する現象であることが示唆された。さらに、タンパク質量のばらつきを大きく変えてシミュレーションしたところ、CV25%程度のばらつきを与えた場合に、最も実験データとよく一致した。CV25%では、細胞間でのタンパク質量の最大/最小の比が、およそ4になる。また、これまでに実験的に計測された ERK や MEK 等一部のタンパク質の細胞間での発現量のばらつきは、およそ CV25%程度であった。本シミュレーションによる予測は、その値とよく一致しており、構築した数理モデルは高精度である事が示された。次に、どの化学種の発現量のばらつきが、ERK 核内移行の応答不均一性に影響をおよぼすかを調べた。それぞれの化学種にばらつきを発生させた数値シミュレーションを行い、ERK 核内移行のばらつきを評価したところ、EGF 受容体、Ras、Raf、MEK タンパク質の変化が比較的大きな ERK 核内移行レベル変化を引き起こした。この結果は、ERK 核内移行がこれらのタンパク質発現量に敏感であることを示唆する。実際、EGFR 受容体や Ras、Raf タンパク質の過剰発現は、肺がんや悪性黒色腫（メラノーマ）で観察されており、細胞のがん化の一因である。したがって、これらの過剰発現は ERK の核内移行レベルを大きく引き上げ、その結果として細胞のがん化を誘導したと考えられる。まとめると、EGF 刺激後の ERK 核内移行の応答不均一性は、EGF シグナル伝達に参与するタンパク質、特に EGF 受容体および Ras、Raf、MEK タンパク質の発現量に起因する可能性が示唆された。

#### (2) 機械学習による応答不均一性の推定

上記の数値シミュレーションから得られたデータを使用して、タンパク質発現量から ERK 核内移行応答が予測可能かどうかを検証した。本研究では、SVM、決定木、線形判別、主成分分析の4手法をテストした。まず、数理モデルに含まれる全ての化学種の初期値を入力にして、ERK 核内移行応答が予測できるかを交差検証により評価した。その結果、4手法の中で SVM が最も高い予測精度を示し、その予測精度は99%と非常に高精度であった。本シミュレーションでは、確率的影響は化学種の初期値のみに反映されているため、そのような高い精度を示したと考えられた。しかしながら、実験的な検証を考えた場合、細胞に刺激を与える前に全ての化学種の濃度を知ることは不可能であるため、より少ない種類の化学種でかつ精度よく ERK 応答を予測できないかを調査した。数値シミュレーションで予測された ERK 核内移行に重要な化学種、EGF 受容体、Ras、Raf、MEK の発現量を入力に ERK 応答を SVM で予測

したところ、予測精度は80%であった。さらに、これら4つのタンパク質から2つの組み合わせで予測を行ったところ、EGF 受容体と Ras タンパク質の組み合わせが最も高い予測精度を示した(80%)。以上の結果から、EGF 刺激時の ERK 核内移行応答は EGF 受容体と Ras の発現量から SVM で精度良く予測できることが示された。本研究での数値シミュレーションおよび機械学習法による数理モデル解析結果を踏まえると、細胞のシグナル伝達における応答不均一性は、そのシグナル伝達に参与するタンパク質の量に起因する可能性が示唆された。さらに、全てのタンパク質が均等に応答不均一性に寄与するわけではなく、応答不均一性を決定づけるキーとなるタンパク質が存在することも示唆された。しかしながら、本研究では、EGF シグナル伝達経路のみしか解析していないため、他のシグナル伝達経路や遺伝子発現制御等も同様のシステムを検証していく必要がある。

今後、本研究で得られた予測に基づいた検証を行うことが出来れば、細胞シグナル伝達機構の詳細や細胞のがん化機序等が明らかになっていく事が期待される。

#### 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3件)

1. Shigeyuki Magi, Kazunari Iwamoto, Noriko Yumoto, Michio Hiroshima, Takeshi Nagashima, Rieko Ohki, Amaya Garcia-Munoz, Natalia Volinsky, Alexander Von Kriegsheim, Yasushi Sako, Koichi Takahashi, Shuhei Kimura, Boris N Kholodenko and Mariko Okada-Hatakeyama, Transcriptionally inducible Pleckstrin homology-like domain, family A, member 1, attenuates ErbB receptor activity by inhibiting receptor oligomerization, *Journal of Biological Chemistry*, **293**(6), 2206-2218, doi: 10.1074/jbc.M117.778399, 2018 (査読あり)

2. Shigeyuki Magi, Kazunari Iwamoto, Mariko Okada-Hatakeyama, Current Status of Mathematical Modeling of Cancer - From the Viewpoint of Cancer Hallmarks, *Current Opinion in Systems Biology*, **2**, 38-47, doi: 10.1016/j.coisb.2017.02.008, 2017 (査読あり)

3. Kazunari Iwamoto, Yuki Shindo and Koichi Takahashi, Modeling cellular noise underlying heterogeneous cell responses in

the epidermal growth factor signaling pathway, *PLoS Computational Biology*, **12**(11): e1005222. doi: 10.1371/journal.pcbi.1005222, 2016 (査読あり)

〔学会発表〕(計6件)

1. 岩本一成, シグナル依存的なクロマチン構造変化によるスーパーエンハンサーの制御機構, 日本生物工学会バイオインフォマティクス相談部会 第一回講演会, 12月26日, 2017.

2. Kazunari Iwamoto, Formation of NF-kappa B super enhancer regulated by chromatin structural changes, Australian National University (ANU) & IPR 2nd Joint Symposium 2017 "PROTEIN STRUCTURE AND FUNCTION", Osaka, 3-5 December, 2017.

3. 岩本一成, 遺伝子変異不均一性によるシグナル伝達破綻機序の解明に向けた数理モデリング, 新学術研究領域「数理シグナル」第1回若手ワークショップ, 静岡, 8月6-8日, 2017.

4. 岩本一成, スーパーエンハンサーを介した転写因子 NF- $\kappa$ B の遺伝子発現制御機構の解明, 第5回NGS現場の会 仙台, 5月22-24日, 2017.

5. Kazunari Iwamoto, Elucidation of NF-kappa B-regulated gene expression mechanism based on various sequence data, 11th International Symposium of The Institute Network, Tokushima, 26-27 January, 2017.

6. Kazunari Iwamoto, Dynamic gene expression regulated by NF-kappa B through super-enhancer, Astbury-IPR Young Scientists Symposium, Leeds, UK, 17-18 January, 2017.

〔図書〕(計2件)

1. 岩本一成, 岡田眞里子, シングルセルシーケンスデータを読み解くための情報解析, 実験医学別冊 シングルセル解析プロトコル, 326-331 (2017)

2. 岩本一成, 間木重行, 岡田眞里子, 数理モデル解析を用いたシグナル伝達機構の解明 ErbB 受容体の負のフィードバック制御の解析事例, 実験医学, vol.35 No.5 788-793 (2017)

## 6 . 研究組織

### (1)研究代表者

岩本 一成 (IWAMOTO, Kazunari)

大阪大学 蛋白質研究所・助教

研究者番号 : 70619866