

2023年6月15日
一般財団法人電力中央研究所

発がんリスクに対する外的要因の影響を推定する 新しい数理モデルの構築

— 複製エラーと外的要因による DNA 損傷の影響を組み込んだ
数理モデルによって経時的な発がんリスクの説明に成功 —

ポイント

- ・ 変異の蓄積を複製エラーと外的要因に分けてモデル化
- ・ 統計データと比較し、複製エラーにより経時的な発がんリスクが説明できることを示した
- ・ 外的要因の影響を組み込むことで、より妥当なパラメータの推定値を得られることがわかった

概要

一般財団法人電力中央研究所（理事長：松浦昌則、本部：東京都千代田区）サステナブルシステム研究本部 生物・環境化学研究部門の内之宮光紀主任研究員、富田雅典上席研究員のグループは、代謝等に伴う内因性の複製エラー^{※1}による変異（以下、複製エラー）と外的要因による DNA 損傷による変異（以下、損傷）の蓄積を分けて調べることのできる数理モデルを構築し、複製エラーだけでも複数のがんの経時的な累積罹患リスク^{※2}を説明できることを示しました。さらに、外的要因を考慮することで、生物学的にも妥当なパラメータを得られることがわかりました。これは、発がんリスクに与える影響が非常に小さい低線量・低線量率放射線のような外的要因の影響評価につながる成果です。

本研究成果は、2023年6月14日付で米国科学誌 PLOS ONE に掲載されました。

1. 背景

生涯の発がんリスクは幹細胞の分裂回数によって説明できることが Tomasetti と Vogelstein (Science, 2015)によって報告されました。これは、内因性の複製エラーの蓄積によって細胞ががん化するためだと考えられます。一方で、放射線などの外的要因が発がんリスクに影響することもよく知られており、外的要因により DNA が傷つき、損傷が蓄積することでがんが生じる可能性があります。しかし、低線量・低線量率放射線の被ばくのように、影響があったとしても非常に小さい要因は、他の要因の影響に隠れて検出が難しく、発がんリスクに影響するかについて未だに議論が続いています。発がんについての数理モデルを構築して現実では排除できない要因を取り除くことで、そのような小さな影響でも評価できる可能性があります。我々は、生涯の発がんリスクが複製エラーの蓄積で説明できるのであれば、各年齢における発がんリスクも同様に説明できるとの発想のもと、そこに損傷の蓄積を加えた数理モデルによって、発がんリスクに対する外的要因の影響をより正確に評価できると考えました。

2. 研究手法・成果の特長

① **変異蓄積モデルの構築と Gillespie アルゴリズムを用いたシミュレーション**

初めに、複製エラーと損傷を分けて表現できる、変異蓄積モデルを構築しました。細胞が分裂して組織を

形成するようなイメージで、最初は細胞分裂により総細胞数が増加していきます (Growth Phase)。細胞数が一定数に到達すると細胞は分裂を停止し、それ以降を Stable phase とします。Stable phase では、細胞が死亡するなどして数が減少すると、それを補うように分裂します。複製エラーによる変異は細胞分裂の際に確率 λ_1 で生じ、損傷は時間的に一定の割合 λ_2 で生じるとします。このモデルを Gillespie アルゴリズム^{※3}という手法を使ってシミュレーションを行うことで、複製エラーと損傷がそれぞれ蓄積していく様子を表現することができました^{※図2}。

② 発がんモデルの構築と複製エラーのみを考慮した場合の統計データとの比較

次に、この変異蓄積モデルを用いた発がんモデルを構築しました。多段階発がん説^{※4}に基づいて、複製エラーもしくは損傷 1 個に対して確率 p でがん関連遺伝子に変異が生じ、 g 個のがん関連遺伝子に変異が生じた場合にがんが生じると想定します。この発がんモデルを、「がん情報サービス」(https://ganjoho.jp/reg_stat/index.html)で公開されているがんの累積罹患リスクの中から、9種類のがんのデータに適用しました (図1)。白血病を除き、経時的な発がんリスクも複製エラーのみでデータをよく説明できました。しかし、がん関連遺伝子の数は数個であるとされているにも関わらず、男性の肺がんにおけるがん関連遺伝子の数は10個と多く推定されました。喫煙は肺がんのリスク要因であることが知られているため、喫煙を想定して20歳以降に外的要因による損傷の蓄積が生じるとして解析を行うと、男性の肺がんにおけるがん関連遺伝子の数は7個と推定され、現実に知られている値に近づきました。

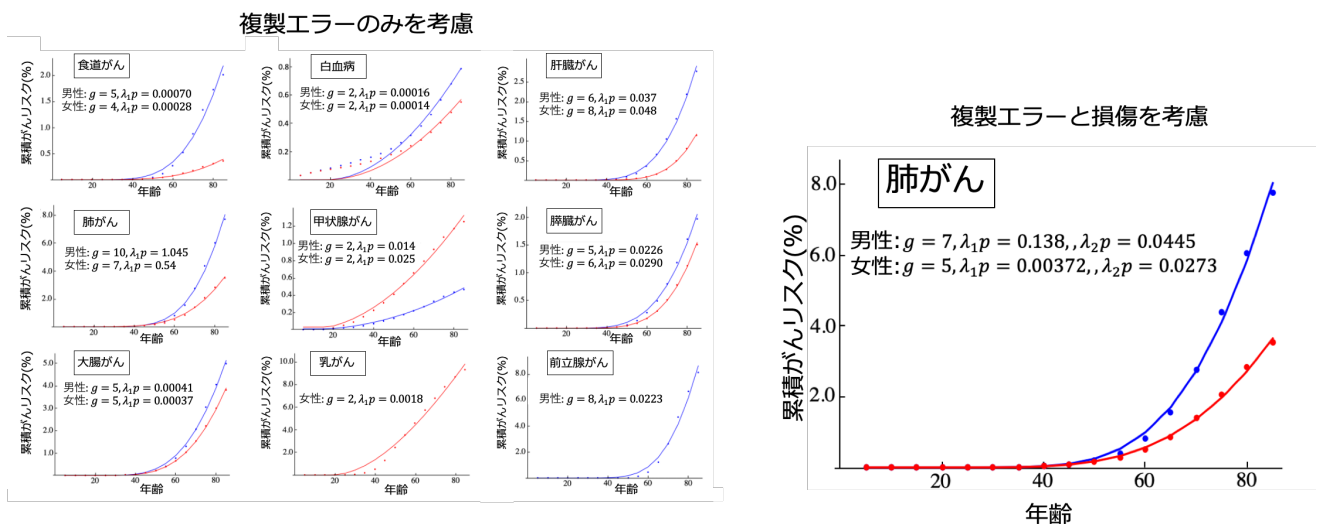


図1 モデルと統計データの比較

青のドットと赤のドットはそれぞれ、男性と女性の発がんリスクの統計データのプロット。曲線は、それぞれのデータにフィットするようにパラメータを推定した数理モデルから導き出された曲線。白血病以外は複製エラーのみを考慮してよく説明できている。肺がんの男性ではがん関連遺伝子の数が10個と推定されているが、これは実験的に知られている値に比べて多い。複製エラーに加えて、損傷を考慮した場合、男性で推定されたがん関連遺伝子の数は7個となり妥当な推定値となった。 $\lambda_1 p$ は分裂あたりに平均何個のがん関連遺伝子に変異が生じるかの推定値、 $\lambda_2 p$ は外的要因によって時間あたり平均何個のがん関連遺伝子に変異が生じるかの推定値、 g はがん関連遺伝子の数の推定値。

3. 本研究成果が社会に与える影響 (本研究成果の意義)

本研究で開発した数理モデルにより、発がんリスクにおける複製エラーの影響と損傷の影響を分けて調べることで、非常に小さな外的要因のリスクについても議論することができます。また、本モデルは細胞同士の相互作用は考慮していませんが、放射線発がんを抑制する可能性がある幹細胞競合^{※5}などを組み込むように改良することで、低線量・低線量率放射線被ばくの影響のような具体的な外的要因の影響の評価につながると期待できます。

<用語説明>

※¹ **複製エラー**：細胞分裂を行う前に DNA を複製する必要があり、その際に元の DNA の塩基配列とは異なる塩基配列が複製されてしまうこと。

※² **累積罹患リスク**：ある年齢までにある病気であると診断される確率。

※³ **Gillespie アルゴリズム**：確率的なシミュレーションを行うためのアルゴリズムの一つ。本研究では、複製エラーが生じるなどの確率的なイベントであると考えているため採用した。

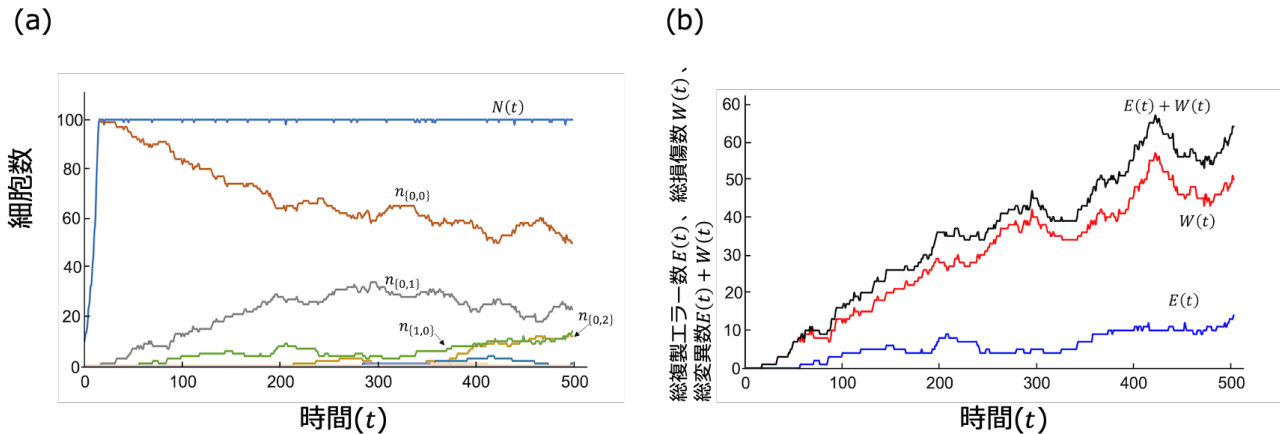


図 2 Gillespie アルゴリズムを用いたシミュレーションの一例

(a)細胞数を一定時間毎に記録した結果。 $n_{\{i,j\}}$ は複製エラーが*i*個、損傷が*j*個蓄積した細胞数。総細胞数 $N(t)$ は全ての $n_{\{i,j\}}$ の合計。Growth phaseの間は $N(t)$ は増加を続け、Stable phaseではほぼ一定を保つ。(b)細胞集団内の総複製エラー数 $E(t)$ 、総損傷数 $W(t)$ 及び、それらの合計である総変異数 $E(t) + W(t)$ の時間変化。確率的なモデルであるため、増減を繰り返しながら、全体としては時間と共に変異が増加していることが見て取れる。

※⁴ **多段階発がん説**：複数のがん関連遺伝子に徐々に変異が生じて正常な細胞ががん細胞になるとする学説

※⁵ **幹細胞競合**：幹細胞間で生じる細胞競合。細胞競合とは、組織内で適応度が異なる細胞が相互作用することで、適応度が低い細胞が排除される現象のこと。

<論文タイトルと著者>

論文タイトル：A mathematical model for cancer risk and accumulation of mutations caused by replication errors and external factors

著者：Kouki Uchinomiya, Masanori Tomita

掲載誌：PLOS ONE, 18(6): e0286499.

URL：<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0286499>

なお、本研究は JSPS 科研費 若手研究 20K19972 の助成を受けたものです。

以上

お問い合わせ：電力中央研究所 広報グループ 担当：林田、藤本 TEL：03-3201-5349（広報グループ直通）

※本件は、文部科学記者会、科学記者会、エネルギー記者会で資料配布致しております。