

# DNA ポリメラーゼ動態と突然変異生成機構

## DNA polymerase dynamics and mutagenesis

(日本遺伝学会推薦)

代表研究者 公益財団法人がん研究会 大学 保一 Japanese Foundation for Cancer Research Yasukazu Daigaku  
協同研究者 公益財団法人がん研究会 ベインブリッジ ルイス Japanese Foundation for Cancer Research Bainbridge Lewis

Eukaryotic genome replication requires the coordinated function of multiple DNA polymerases, each differing in fidelity and catalytic efficiency. While replication mechanisms have been elucidated in yeast, mammalian cells utilize a broader array of polymerases, suggesting a more flexible and redundant replication system. This study investigates how error-prone polymerases, particularly Pol $\eta$  and Pol $\kappa$ , contribute to genome replication under physiological conditions. Using CRISPR-Cas9, we generated human cell lines with active-site mutations in Pol $\eta$  and Pol $\kappa$  that promote ribonucleotide incorporation. By detecting genomic regions with elevated levels of embedded rNMPs, we identified the genome-wide activity of these polymerases. This approach, known as polymerase usage sequencing (Pu-seq), revealed that although Pol $\eta$  and Pol $\kappa$  exhibited only modest synthesis profiles across the genome, their activity was specifically detectable in replication initiation zones, particularly on the lagging strand. Moreover, mutation signature analysis based on cancer genome data suggested that Pol $\eta$ - and Pol $\kappa$ -mediated synthesis may underlie strand-biased mutagenesis near IZs. We also mapped replication initiation using Pol $\epsilon$  and Pol $\alpha$  profiles and found that IZs correlated with intergenic regions marked by H2A.Z and H3K27me3. In contrast, heterochromatic regions lacked distinct IZs, likely due to dispersed initiation events. These findings highlight the complexity of genome replication control, where both DNA polymerase dynamics and chromatin context collectively shape the spatial organization and fidelity of replication initiation.

### 研究目的

DNA 合成を担う酵素である DNA ポリメラーゼは、ヒトにおいて 10 種類以上存在し、それぞれ DNA 合成の効率や正確性が異なる (1)。これら多様な DNA ポリメラーゼが、ゲノム複製においてどのように分業・協調して機能するかのメカニズムは、遺伝情報の安定性を理解する上で解明すべき重要な課題である。

これまで、DNA 複製のメカニズムは主に出芽酵母や分裂酵母を用いて研究してきた。その結果、DNA 複製フォークの進行に伴い、リーディング鎖は Pol $\epsilon$  (イプシロン) が、ラギング鎖は Pol $\alpha$  (アルファ) および Pol $\delta$  (デルタ) が主に合成を担うことが明らかとなっている (2,3)。一方で、はるかに大きなゲノムを有する哺乳類細胞では、より多くの DNA ポリメラーゼがゲノム複製に関与し、柔軟かつ冗長性

のある複製メカニズムを構築していることが示唆されている。

本研究では、哺乳類細胞においてゲノム複製の主役となる DNA ポリメラーゼに加え、エラーが起こりやすいとされる DNA ポリメラーゼ (error-prone polymerases) が、複製のバックアップ機構としてどのように機能するかを明らかにする。また、これらの error-prone polymerases による DNA 合成が自然突然変異の原因となるかを検証し、ゲノム上の突然変異生成における DNA ポリメラーゼ動態の役割を解明することを目的とする。

さらに、ヒト細胞のような大規模ゲノムを持つ細胞においては、1 つの複製装置が長大な DNA 領域を効率的に合成する必要があることから、複製開始領域のゲノム上での分布は、全体の複製を限られた時間内に完了させるための重要な要素となる。また、

その機構の制御はゲノム安定性維持という点でも欠かせない仕組みである。これまでの知見から、複製開始反応の位置は、クロマチン構造、核内高次構造、遺伝子との相対的位置といった多様なゲノム因子に依存することが示唆されているが、その決定機構の全容は明らかになっていない。本研究では、Pu-seq (Polymerase usage sequencing) により特定される複製開始領域の分布および活性に関するデータを基盤として、複製開始領域の決定に寄与するゲノム上の要因を包括的に解析し、そのメカニズムの解明を目指す。

## 研究経過

本研究では、ヒト培養細胞 HCT116 を用いて、複数の DNA ポリメラーゼがどのように分業し、協調しながらゲノム複製を行っているのか、その分子機構を解明することを目指している。また、これらの機構がゲノム情報の安定性の維持にどのように寄与しているのかを明らかにするため、主にゲノム複製の大部分を担う複製ポリメラーゼ Pol $\epsilon$ , Pol $\alpha$  に加え (4)、誤りやすい合成を行う複数の DNA ポリメラーゼも対象とし、Pu-seq 法を用いた解析系の確立を進めてきた。Pu-seq 実験では、(i) 細胞内の特定の DNA ポリメラーゼの活性部位に変異を導入し、そのポリメラーゼによるゲノム DNA へのリボヌクレオチド (rNMP) の取り込みを誘導する。(ii) アルカリ処理により rNMP 部位で DNA を切断し、得られた断片を次世代シーケンサーで解析することで、取り込まれた rNMP のゲノム上の分布を明らかにし、対象ポリメラーゼが合成を行った領域を特定する。本年度は、誤りを起こしやすい DNA ポリメラーゼの中でも、特にゲノム複製に関与すると予想される Pol $\eta$  (エータ)、Pol $\kappa$  (カッパ)、に着目し、これらに対する実験系の確立を優先的に実施した。Pol $\eta$  および Pol $\kappa$  は、従来、DNA 損傷バイパスに関与するポリメラーゼとして認識してきた。しかし、本研究では DNA 損傷を誘導するような処理を一切行わず、通常のゲノム複製過程において、これらのポリメラーゼがどのような役割を果たしているのかを検証した。

まず、Pol $\eta$  および Pol $\kappa$  を対象とした Pu-seq を実施するにあたり、これらのポリメラーゼが rNMP を取り込みつつ DNA 合成を行う変異細胞の作製が必要であった。そのため、Pol $\eta$  および Pol $\kappa$  の変異導入

候補となるアミノ酸残基を、過去の遺伝学的あるいは生化学的研究に基づいて選定した。さらに、それらのアミノ酸が各ポリメラーゼの三次元構造において、活性中心に位置し rNMP の選択に関与する可能性のある部位であることを構造情報から確認した。続いて、選定したアミノ酸残基に変異を導入するため、CRISPR-Cas9 システムを用いたゲノム編集を行った。具体的には、変異配列を含む単鎖 DNA (ssDNA) をドナーテンプレートとして用いたノックイン (knock-in) 手法を適用した。その結果、Pol $\eta$  および Pol $\kappa$  のそれぞれにおいて、目的とする変異が両対立遺伝子座に導入された細胞株の単離に成功した。

これらの細胞株を使用して、Pol $\eta$  および Pol $\kappa$  の合成活性を、ゲノム DNA 中に取り込まれたリ rNMP の分布を指標として解析した。その結果、リーディング鎖およびラギング鎖の合成を担う主要ポリメラーゼである Pol $\epsilon$  や Pol $\alpha$  ほど、明瞭な合成プロファイルは得られなかった。しかし、複製開始が活発に起こるゲノム領域に限定して解析を行ったところ、Pol $\eta$  および Pol $\kappa$  がラギング鎖合成に関与していることを示唆するプロファイルが得られた。ラギング鎖合成は、リーディング鎖合成とは異なり、レプリソーム (ヘリカーゼを含む複製フォークの進行を担うタンパク質複合体) から一時的に乖離することがある。この過程で、通常ラギング鎖合成を担う Pol $\delta$  に代わって、Pol $\eta$  や Pol $\kappa$  が一部のラギング鎖伸長に関与している可能性が本研究から示唆された。

誤りがちな DNA ポリメラーゼが、リーディング鎖よりもラギング鎖合成に関与しやすいという観察結果から、これらのポリメラーゼによる合成が、複製フォーク周辺に非対称な変異分布を引き起こしている可能性が考えられる。過去の生化学的研究において、Pol $\eta$  は鋲型鎖にチミン (T) が存在する際、本来挿入されるべきアデニン (A) の代わりにグアニン (G) を誤挿入する傾向があり、また Pol $\kappa$  はアデニンの代わりにシトシン (C) を誤って挿入することが報告されている。そこで本研究では、これらの変異 (A→G 変異および A→C 変異) が、複製開始領域 (initiation zone: IZ) 周辺でどのように分布するかを解析した。IZ の位置および開始効率に関するデータは、本研究グループが独自に作成したものであり、そのうち開始効率が高い上位 25% の IZ を解析対象とした。変異のデータは、がんゲノムの大規模解析プロジェクト「Pancancer Analysis of Whole Genomes

(PCAWG)」のデータベースから取得したものである。解析の結果、A→G 変異については、がん種によって変異の分布パターンが大きく異なることが明らかになった。しかし、皮膚がんや胃がんなど、複数のがん種においては、ラギング鎖に偏った変異分布が観察された。A→C 変異に関しては、極端なラギング鎖バイアスを示すがん種は少なかったが、全体的にはその傾向がみられた。これらの結果から、がん細胞における A→G 変異および A→C 変異の分布パターンが、 $\text{Pol}\eta$  や  $\text{Pol}\kappa$  による合成活動と関係している可能性が示唆される。しかし、本研究時点では因果関係を断定することはできず、今後さらに詳細な解析を行う予定である。

上記の DNA ポリメラーゼ機能の直接的な検証に加えて、DNA ポリメラーゼ  $\text{Pol}\epsilon$  (イプシロン) および  $\text{Pol}\alpha$  (アルファ) の合成プロファイルを利用し、複製開始反応が起きるゲノム領域を同定し、その領域の特徴を探索した。我々は本研究の開始以前に、DNA ポリメラーゼ  $\text{Pol}\epsilon$  および  $\text{Pol}\alpha$  の合成プロファイルをそれぞれ、リーディング鎖およびラギング鎖合成の指標として利用し、ヒトゲノム上の複製開始領域 (initiation zone: IZ) を、従来にない高精度で特定する解析手法を確立していた。よって、帆研究期間中には、上記の解析系、特に複製開始および停止反応の指標として導入した initiation index (複製開始度) を用いて、複製開始に必要な染色体 DNA 上の特徴を明らかにするための包括的な解析を行った。これまでの研究により、initiation index のピークは、遺伝子の前後、すなわち転写開始点 (TSS) および転写終結点 (TES) の近傍、約 15–20 kb の範囲に位置することが知られている。この知見に基づき、我々は複製開始が活発に起こると予想される発現レベルの高い遺伝子間領域 (10–100 kb) を選び出し、それらの領域における initiation index のパターンを詳細に解析した。その結果、initiation index が正值を示す、すなわち複製開始が高頻度で生じる領域は、遺伝子間領域と正確に一致することが確認された。興味深いことに、これらの領域には、特定のコンセンサス配列や明確な塩基配列のバイアスといった DNA 配列に基づく特徴は見られなかった。

一方で、エピジェネティックな特徴として、H2A.Z バリアントヒストンおよび H3K27me3 (ヒストン H3 の第 27 番目リジンのトリメチル化) が、複製開始が頻発する遺伝子間領域に顕著に集積していることが

判明した。これらの修飾は、H3K4me3 や H3K27ac など、転写活性に関する修飾とは異なる空間的分布を示し、IZ に特異的なクロマチン構造であることが明確となった。H2A.Z や H3K27me3 は、必要な因子が局在するための開いたクロマチン環境の形成や、直接的な因子のリクルート (誘導) に関与している可能性があるが、その役割については今後さらに詳細な検証が必要である。また、複製開始領域に関する各種クロマチン因子のゲノム分布データを用いて主成分分析 (PCA) を行った結果、複製開始の効率は H2A.Z および H3K27me3 の局在量と有意に相関することが明らかになった。しかし、それらの特徴だけでは IZ の構造全体を説明するには不十分であることも示された。この結果から、複製開始を制御するメカニズムには、当初の予想を超えて遺伝子配置やクロマチン構造など、多様な要素が複雑に関与していることが示唆された。

上述のように、IZ には特有のクロマチン構造が存在する一方で、H3K9me3 によって特徴づけられるヘテロクロマチン領域には IZ が全く存在しないことが、我々の研究によって明らかになった。ヘテロクロマチン領域は一般に S 期後半に複製される領域であるが、そこで複製開始が全く起きないというわけではない。これらの領域は染色体上に散在しており、そのサイズは数 Mbp におよぶ場合がある。しかし、1–2 kb/min で進行する複製フォークだけでは、そのような広範囲を S 期の限られた時間内にカバーすることは困難である。

このような状況においては、複製開始がランダムに分散して起こると考えられ、一定の領域に集中的に複製開始が起こる IZ としては検出されないことが予想される。すなわち、ヘテロクロマチンは複製開始反応を空間的に拡散させることで、IZ の形成を阻害していると考えられる。このように、染色体上のクロマチン構造は、複製開始反応に対して正または負の影響を与え、ゲノム複製機構の空間的制御において極めて重要な役割を果たしていることが示された。さらに、細胞の分化や増殖状態などに応じてクロマチン状態は刻々と変化するため、それに伴ってゲノム上の複製パターンも動的に変化していくことが予想される。この点については、今後、詳細な解析を進める予定である。

## 考察

本研究は、 $\text{Pol}\eta$  および  $\text{Pol}\kappa$  は DNA 合成の効率が高い酵素ではないが、複製フォークの進行過程においてラギング鎖合成に関与していることを示した。この結果は、ヒト細胞におけるゲノム複製には、主要な複製ポリメラーゼに加えて、誤りを起こしやすい複数の DNA ポリメラーゼも機能していることを示唆しており、これらが潜在的に変異生成の原因となり得ることを意味している。これまでのところ、個々の DNA ポリメラーゼの合成領域とゲノム上の変異分布との関係について、網羅的かつ体系的な解析は行われていない。今後は、本研究で一部示したように、取得可能ながんゲノム変異データと、我々が Pu-seq などを通じて得る各 DNA ポリメラーゼの合成プロファイルとを比較・検討することで、変異生成に強く関与する DNA ポリメラーゼの特定を目指す。また、がん種や症例ごとに亢進している DNA ポリメラーゼの同定も試みる予定である。さらに、複製開始反応を誘導するゲノム上の特徴を明らかにすることにより、複製フォークの動態と変異分布との関連性についても検証を進めていく。

## 参考文献

1. Lange S. S., Takata K., Wood R. D., DNA polymerases and cancer. *Nat Rev Cancer*, 11, 96–110, 2011.
2. Daigaku Y., Keszthelyi A., Müller C. A., Miyabe I., Brooks T., Retkute R., Hubank M., Nieduszyski C. A., Carr A. M.\*, A global profile of replicative polymerase usage. *Nat. Struct. Mol. Biol.*, 22, 192–8, 2015
3. Clausen A. R., Lujan S. A., Burkholder A. B., Orebaugh C. D., Williams J. S., Clausen M. F., Malc E. P., Mieczkowski P. A., Fargo D. C., Kunkel T. A., Tracking replication enzymology *in vivo* by genome-wide mapping of ribonucleotide incorporation. *Nat Struct Mol Biol.*, 22, 185–191, 2015.
4. Koyanagi E.#, Kakimoto Y.#, Minamisawa T.#, Yoshifuji F.#, Natsume T., Higashitani A., Ogi T.,

Carr A. M., Kanemaki M. T., Daigaku Y.\*; Global landscape of replicative DNA polymerase usage in the human genome. *Nat Commun.*, 13, 7221, 2022  
(\*筆頭著者, #同等貢献)

## 研究の発表

### 口頭発表

1. Yasukazu Daigaku, Genome Stability through the Lens of DNA Polymerase Function. 21st Birthday Symposium at the Genome Damage and Stability Centre; 2024.09, Brighton, ES, UK
2. Bainbridge, L., 増田雄司, 南澤宝美后, 高橋真美, 益谷央豪, 大學保一,  $\text{Pol}\eta$  および  $\text{Pol}\kappa$  のゲノム複製への関与とがんゲノム変異分布との関連性. 日本遺伝学会第 96 回大会; ;2024.09, 高知市.
3. 大學保一, Bainbridge, L., 高橋真美, 南澤宝美后 DNA polymerase dynamics and mutagenesis in cancer cells/DNA ポリメラーゼ動態とがん細胞における変異生成機構. 第 83 回日本癌学会学術総会; ;2024.09, 福岡市
4. Yasukazu Daigaku, Flexible usage of diverse DNA polymerases and its implications for mutagenesis. The 12th 3R+3C International Symposium; 2024.11, 福岡市
5. 大學保一, 多様な DNA ポリメラーゼによる染色体複製と変異生成.第 42 回染色体ワークショップ・第 23 回核ダイナミクス研究会; 2025.01, 大分県速見郡日出町

### 誌上発表

1. Bainbridge L.J., Daigaku Y.\*; Bulk synthesis and beyond: The roles of eukaryotic replicative DNA polymerases. *DNA repair*, 141, 103740, 2024 (総説)
2. Bainbridge L.J.\*; Daigaku Y.\*; Adaptive use of error-prone DNA polymerases provides flexibility in genome replication during tumorigenesis. *Cancer Science*, 115, 2125-2137, 2024 (総説)