

PRESS RELEASE

2019 年 2 月 22 日

理化学研究所
三重大学

ゲノム DNA 複製の真の姿を捉えた

ーゲノム DNA が倍加していく様子を 1 細胞レベルで明らかにー

理化学研究所（理研）生命機能科学研究センター発生エピジェネティクス研究チームの平谷伊智朗チームリーダー、高橋沙央里基礎科学特別研究員、三浦尚研究員、三重大学大学院医学系研究科の竹林慎一郎講師、柴田隆豊大学院生（研究当時）らの共同研究グループ*は、増殖中の細胞でゲノム DNA が倍加していく過程を網羅的に解析する「scRepli-seq 法」を開発し、その様子を 1 細胞レベルで詳しく観察することに成功しました。

真核生物の「ゲノム DNA 複製^[1]」の真の姿を捉えた本成果は、細胞核内でゲノム DNA 複製が制御される仕組みや、クロマチン^[2]高次構造とゲノム機能の関係性の理解に大きく貢献すると期待できます。

増殖中の細胞におけるゲノム DNA 複製のプロセスは、これまでは技術的な限界から、数万個に及ぶ細胞の平均像としてしか捉えられていませんでした。

今回、共同研究グループは、1 細胞での全ゲノム解析^[3]を実現する scRepli-seq 法を開発し、さらに「一塩基多型^[4]」の情報を用いることで、一つ一つの細胞で父方と母方由来の染色体 1 本 1 本が複製されていく様子を捉えることに成功しました。染色体のどの領域がいつ複製されたかを集団中の個々の細胞間で比較すると、そのゆらぎ（ばらつき）は予想以上に小さく、ゲノム DNA 複製の時間的制御は、細胞間で高度に保存されていました。つまり、これまで見えていた数万個の細胞の平均像は、個々の細胞におけるゲノム DNA 複製の姿に驚くほど近かったことが明らかになりました。

本研究は、英国の科学雑誌『Nature Genetics』オンライン版（2 月 25 日付け：日本時間 2 月 26 日）に掲載されます。

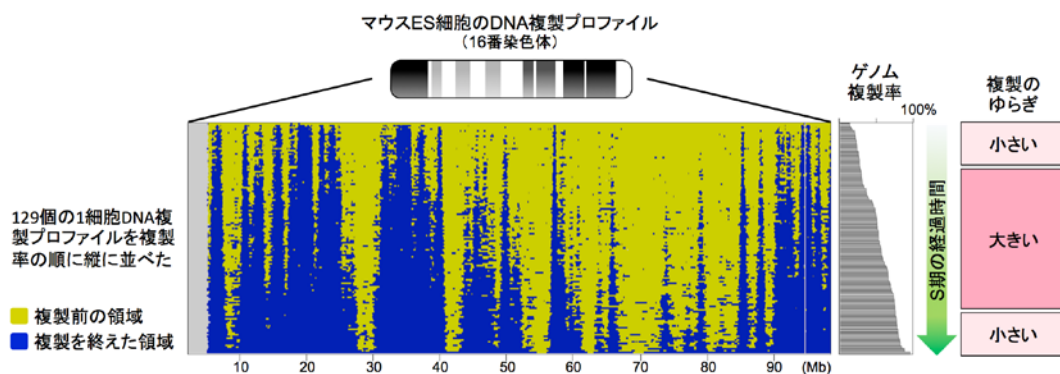


図 マウス ES 細胞のゲノム DNA が複製されていく様子と細胞間の複製のゆらぎ

※共同研究グループ

理化学研究所 生命機能科学研究センター 発生エピジェネティクス研究チーム

チームリーダー 平谷 伊智朗（ひらたに いちろう）

基礎科学特別研究員 高橋 沙央里（たかはし さおり）

研究員 三浦 尚（みうら ひさし）

（以上、旧所属は多細胞システム形成研究センター 発生エピジェネティクス研究チーム）

三重大学大学院 医学系研究科 基礎医学系講座 機能プロテオミクス分野

教授 緒方 正人（おがた まさと）

講師 竹林 慎一郎（たけばやし しんいちろう）

三重大学大学院 生物資源学研究科 生命機能化学講座 分子細胞生物学分野

教授 奥村 克純（おくむら かつずみ）

大学院生（研究当時） 柴田 隆豊（しばた たかひろ）

（現 滋賀県農業技術振興センター）

大阪大学大学院 理学研究科 生物科学専攻 染色体構造機能学研究室

教授 小布施 力史（おぶせ ちかし）

准教授 長尾 恒治（ながお こうじ）

※研究支援

本研究は、日本学術振興会（JSPS）科学研究費補助金若手研究「1 細胞 DNA 複製タイミング解析による発生分化過程の核内コンパートメント動態予測（研究代表者：高橋沙央里）」、同新学術領域研究（研究領域提案型）「染色体機能ドメインの可塑性とその意義（研究代表者：竹林慎一郎）」、同新学術領域研究（研究領域提案型）「細胞分化にともなうクロマチンポテンシャルの変化とその分子基盤（研究分担者：平谷伊智朗）」、同新学術領域研究（研究領域提案型）「ヘテロクロマチンの凝縮構造を作り出すノンコーディング RNA 群の解明（研究代表者：長尾恒治）」、同新学術領域研究（研究領域提案型）「性染色体ヘテロクロマチンが規定する性スペクトラム（研究代表者：長尾恒治）」による支援を受けて行われました。

1. 背景

生物が持つ基本的な性質の一つは、「増える」ことです。あらゆる生物の細胞は、生命の設計図であるゲノムを正確に倍加させる「ゲノム DNA 複製」というプロセスを経て初めて、二つの娘細胞に分かれ、増殖することができます。

真核生物では、細胞周期^[5]の S 期^[5]にゲノム DNA が複製される際に、どの領域から順に複製が進行していくかという順序（複製タイミング^[6]）と、それらの領域におけるクロマチン高次構造との間に強い相関が見られることが報告されています^{注 1)}。すなわち、DNA が緩く折り畳まれて遺伝子がよく転写^[7]されている領域ほど複製タイミングが早く、密に折り畳まれて遺伝子の転写が抑えられている領域ほど複製タイミングが遅いことが知られています。

複製の過程を正確に理解することは、細胞分裂を通してゲノム DNA が安定に維持され、クロマチン高次構造が適切に制御される仕組みの理解という観点から非常に重要です。しかし、多細胞生物を構成する一つ一つの細胞について、複製タイミングとクロマチン高次構造がどの程度関連しているかは分かっています。

せんでした。複製の過程を網羅的に調べるには、少なくとも数万個の細胞が必要であり、そのような解析から得られた複製に関する記述は、実は数万個の細胞から得られた平均像にすぎなかったからです。個々の細胞で実際にどのように複製が進行するのかは、分子生物学における未解決の問題でした。

注 1) Rhind, N. and Gilbert, D.M. DNA replication timing. *Cold Spring Harb Perspect Biol.* 5, 1-26 (2013).

2. 研究手法と成果

1 細胞レベルで解析を行うためには、従来法とは根本的に違う方法論が必要です。共同研究グループは、S 期の細胞 1 個からゲノム DNA を取り出し、各領域間のコピー数の差を検出できれば、複製前（1 コピー）と複製後（2 コピー）の領域を識別できる可能性があると考えました。この考えをもとに新しい方法論の開発に取り組み、ゲノム DNA が複製される様子を 1 細胞レベルの全ゲノム解析によって捉える「scRepli-seq (single-cell DNA Replication sequencing) 法」の開発に成功しました。この手法で S 期中期のヒト細胞を解析した結果、全ての染色体について、どの領域がどこまで複製されているかを詳細に表す「DNA 複製プロファイル」が得られました (図 1)。

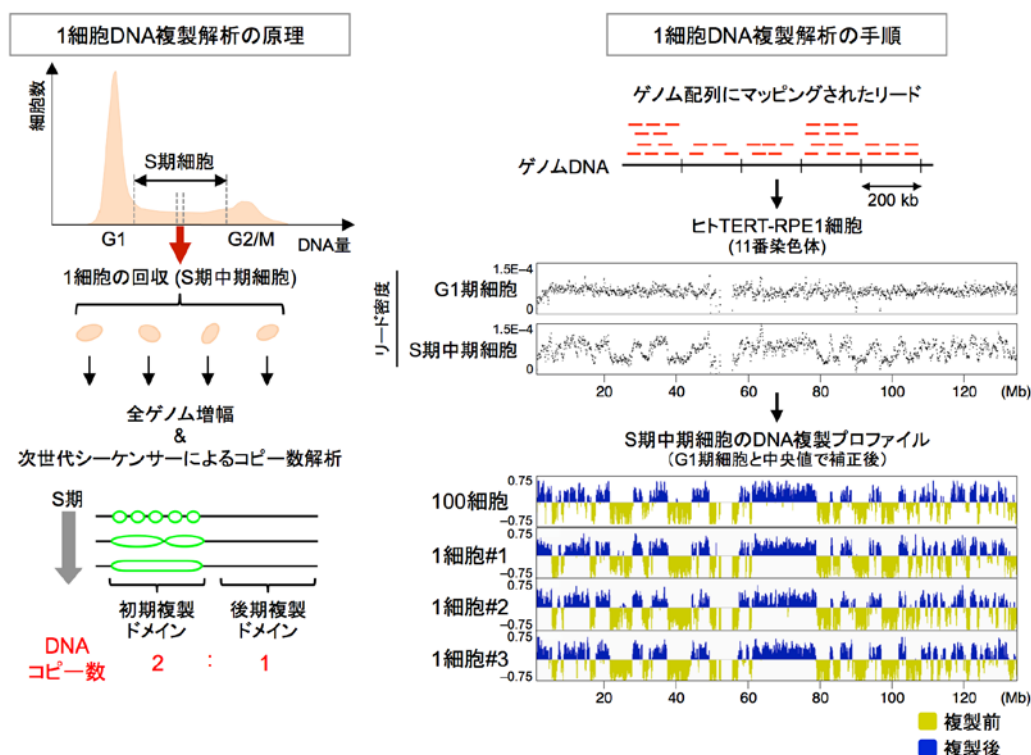


図 1 本研究で開発した 1 細胞ゲノム DNA 複製解析法「scRepli-seq 法」の原理と解析の手順

左： scRepli-seq 法の原理 (図は S 期中期細胞の場合)。フローサイトメーターという細胞を分ける機器を用いて、S 期の細胞を一つ一つ個別に回収する。各々の細胞から抽出したゲノム DNA の全ゲノム増幅を行い、次世代シーケンサーを用いてこのゲノム DNA を解読し、ゲノム DNA 上の初期複製ドメイン（複製後）と後期複製ドメイン（複製前）を DNA コピー数によって識別する。

右：解析の手順。（上段）次世代シーケンサーで得られた配列（リード）をマッピング（配列が一致するゲノム領域にリードを対応させる作業）し、解析を行った。（中下段）ヒト TERT-RPE1 細胞（網膜色素上皮細胞）の 11 番染色体のデータで、横軸に染色体上の位置と距離を示している（Mb：メガベース、100 万塩基対）。染色体ごとに見ると、複製前の G1 期の細胞では、染色体全域でリード密度がほぼ一定なのに対し、S 期中期の細胞ではリード密度が領域によって異なることが分かる（中段）。S 期中期細胞のリード密度を、G1 期細胞データ及び S 期中期細胞の中央値で補正すると、複製前（黄）と複製後（青）の領域が存在すること、また、1 細胞での結果が 100 細胞からの結果がほぼ同じであることが分かる（下段）。

さらに共同研究グループは、scRepli-seq 法を用いた全ゲノム解析に、もう一つ工夫を加えました。父方由来と母方由来の相同染色体^[8]はほぼ同じ DNA 配列を持ちますが、父母間で異なる「一塩基多型」の情報を利用することで、1 細胞中の各染色体について父方由来と母方由来の染色体を識別することができます。scRepli-seq 法にこの識別を組み合わせることで、細胞中の染色体 1 本 1 本がまさに複製されていく様子を捉えることにも成功しました（図 2）。

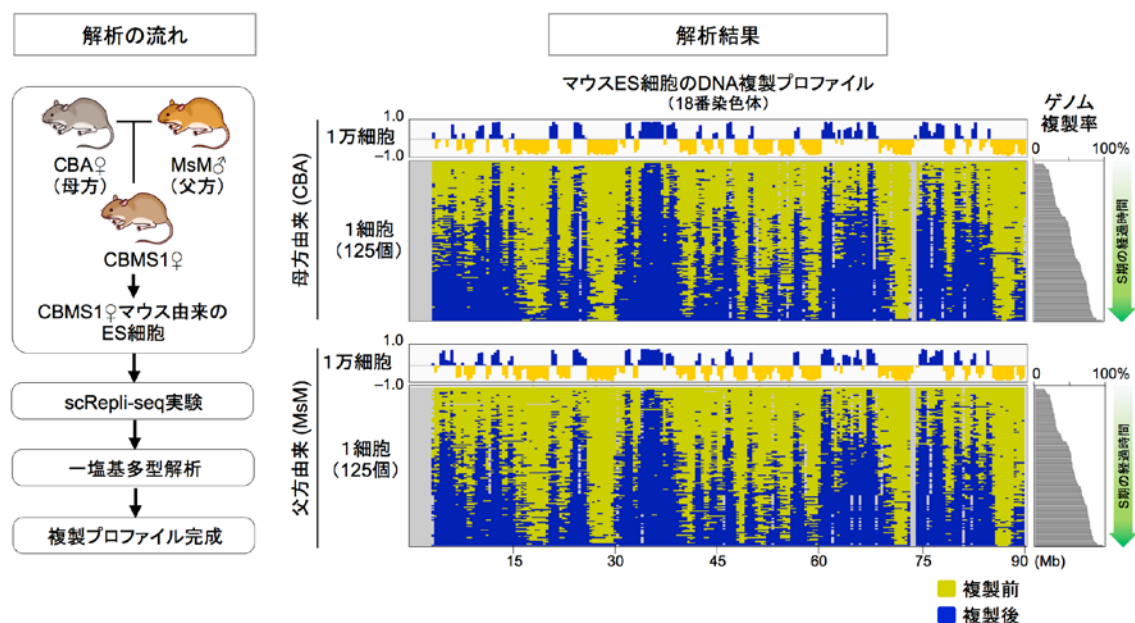


図 2 父方由来と母方由来の染色体を識別した scRepli-seq データ

左：解析の流れ。母＝CBA 系統（西ヨーロッパ産マウス亜種由来）、父＝MsM 系統（日本産マウス亜種由来）の掛け合わせで生まれた雑種第一代（CBMS1）雌マウスに由来する ES 細胞を用い、通常の scRepli-seq 実験を行った後、一塩基多型の解析を行って、母方(CBA)由来と父方(MsM)由来の染色体を識別し、最終的にそれぞれの染色体の DNA 複製プロファイルを完成させる。

右：解析結果。CBMS1 マウス ES 細胞について、合計 125 個の 1 細胞データからの母方(CBA)由来と父方(MsM)由来の染色体ごとの 1 細胞 DNA 複製プロファイルを示す。細胞は、S 期全域から回収し、ゲノム DNA 複製率（ほとんど未複製なものから、複製完了に近いものまで）に従って並べた。1 万細胞の平均像として見えていたゲノム DNA 複製のプロセスが、個々の細胞のそれになりに近いことが分かる。

さらに、マウス ES 細胞^[9]を解析した結果、1 細胞ゲノム DNA 複製プロファイル間のゆらぎ（ばらつき）は、予想よりもはるかに小さく、複製のプロセスは細胞間で高度に保存されていました（図 3）。また、マウス ES 細胞を分化^[10]させ

たところ、分化後に DNA 複製プロファイルは大きく変わるものの、分化後も細胞集団中のゆらぎは依然小さいことも分かりました（図 3）。これらの細胞において、相同染色体同士はよく似たプロファイルを示したことから、特定の分化状態の均一な細胞集団において、複製プロセスの時間的制御は、細胞間のみならず、相同染色体間でもかなり保存されていることが明らかになりました（図 2）。

以上の結果から、これまで数万個の細胞の平均像として見えていたゲノム DNA 複製のプロセスは、個々の細胞のそれにかかなり近いことが明らかになりました。

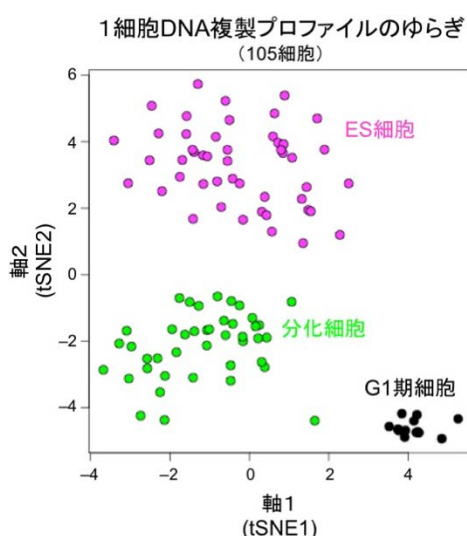


図 3 分化前の ES 細胞と分化後の細胞の scRepli-seq プロファイルのゆらぎ

tSNE 法というデータ解析手法を用いて、105 個の細胞の DNA 複製プロファイルのゆらぎ（ばらつき）を 2 次元のグラフで表現した。解析に用いたのは S 期中期にある ES 細胞（48 個）と、ES 細胞から分化させた細胞（42 個）で、G1 期細胞は対照群（ばらつきが少なくグラフ上でまとまった集団）。個々の細胞間に多少のばらつきはあるものの、ES 細胞、分化細胞、G1 期細胞はそれぞれ違う個性を持ち、同一細胞種内での 1 細胞 DNA 複製プロファイルのゆらぎは比較的小さいことが分かった。

次に、複製タイミングとクロマチン高次構造の相関について、1 細胞レベルで検証することを試みました。近年の Hi-C 法^[11]と呼ばれる全ゲノム解析を用いた染色体の三次元構造の研究から、間期^[5]の細胞核の中で各々の染色体は 100 万塩基対^[12]ほどの球状に折り畳まれたトポロジカルドメイン（TAD：Topologically Associating Domain）が数珠つながりになった構造をとると考えられています（図 4 上）。TAD はさらに折り畳まれ、遺伝子がよく転写されている TAD が集まって A コンパートメントと呼ばれる領域を、転写されていない TAD が集まって B コンパートメントと呼ばれる領域を核内空間に形成するとされています^{注 2)}。

scRepli-seq 法によって得られた 1 細胞 DNA 複製プロファイルは、この A/B コンパートメントの分布を非常によく反映していました（図 4 下）。このことは、A/B コンパートメントという構造の分布も、細胞間で高度に保存されている可能性を示唆しています。

注 2) Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (New York, N.Y.) 326, 289-93 (2009).

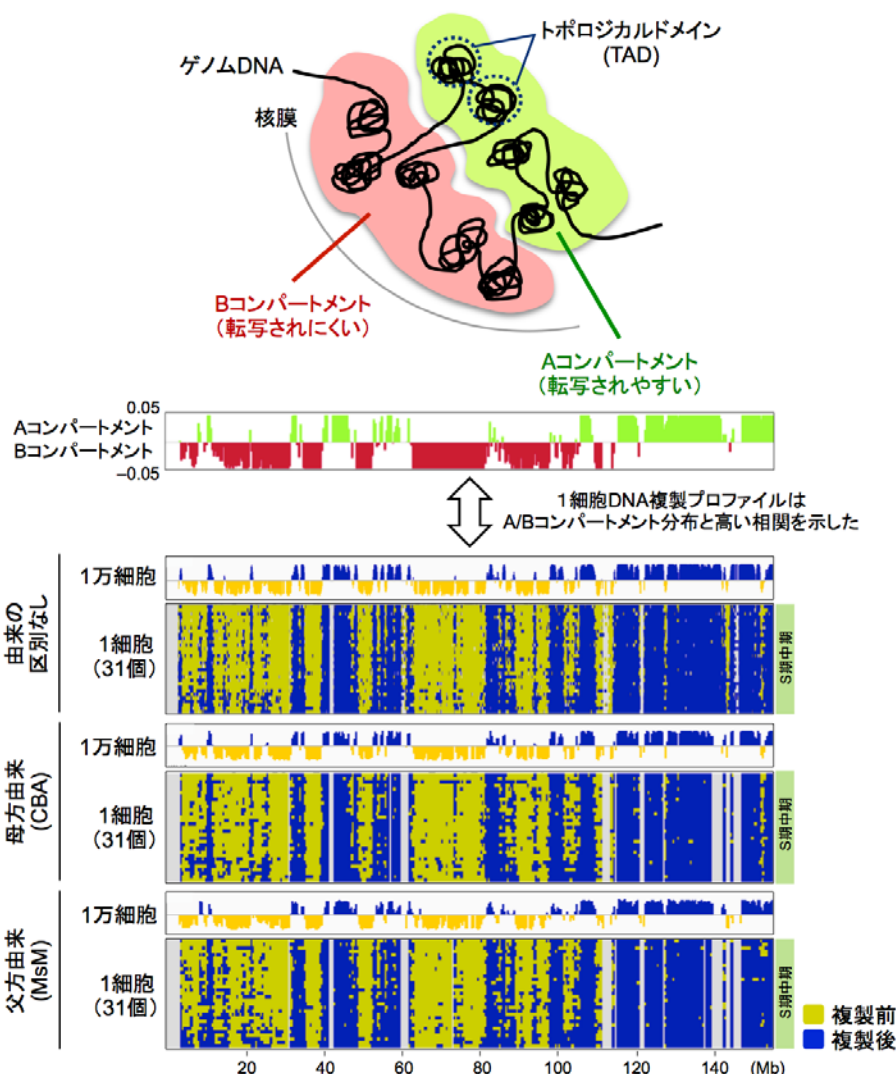


図4 A/B コンパートメント分布と1細胞レベルの複製ドメイン分布の相関

上段は、細胞核内の染色体三次元構造の模式図と、マウス ES 細胞の Hi-C 解析によって得られた A/B コンパートメント分布を示す。下段には S 期中期細胞から得られた scRepli-seq データを示す。「由来の区別なし」は、父方母方由来染色体識別前のデータ。「母方由来」と「父方由来」は識別後のデータ。1 細胞ごとのゲノム DNA 複製プロファイルは、細胞集団の解析から得られた A/B コンパートメント分布に非常によく似ており、上段の A コンパートメント（緑）と B コンパートメント（赤）は、それぞれ青色の複製後（S 期前半に複製される領域）と黄色の複製前（＝S 期後半に複製される領域：黄）のゲノム DNA 領域に対応していた。

以上の結果から、均一な細胞集団に属する個々の細胞は、よく似た DNA 複製プロファイルを持つという傾向が明らかとなりました。一方で、DNA 複製プロファイルの中で差異の大きい箇所、すなわち細胞間・相同染色体間で比較的ゆらぎの大きい領域を詳しく見ていくと、以下の特徴があることも分かりました。

- 1) 細胞間の複製時期のゆらぎは、S 期の初期と後期でかなり小さく、中期に複製される領域についてはゆらぎが比較的大きい（図 5）。
- 2) 細胞分化^[10]に伴い複製タイミングが変化する領域に注目すると、未分化 ES 細胞の段階で既に細胞間の複製時期のゆらぎが比較的大きい。

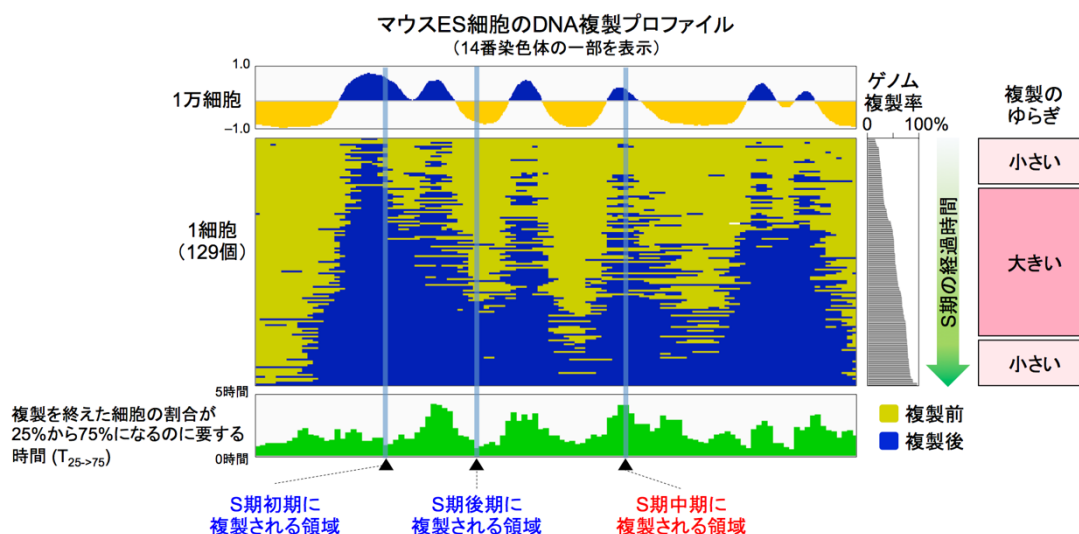


図5 S期初期・後期および中期におけるゲノムDNA複製時期のゆらぎの大きさ

S期全域から得られたマウスES細胞の1細胞DNA複製解析データのうち、14番染色体の一部を示す。下の緑の棒グラフは、ゲノムDNAのある領域を複製し終えた細胞の割合が25%から75%になるのに要する時間を $T_{25 \rightarrow 75}$ と定義し、各領域ごとの時間を示したもの。 $T_{25 \rightarrow 75}$ が大きいほど細胞間の複製時期のゆらぎは大きい。S期初期と後期に複製される領域は $T_{25 \rightarrow 75}$ が非常に短く、対照的に、S期中期に複製される領域は $T_{25 \rightarrow 75}$ が長い。

これらの観察結果は、①ゲノムDNAの複製がどの領域から開始してどの領域で完了するかは、細胞の種類ごとにあらかじめほぼ決定されているが、それ以外の領域がいつ複製されるかは細胞間で多少ゆらぐ、②細胞分化に伴う複製タイミングの変化は、その変化に先立つ複製時期のゆらぎ、すなわちコンパートメント構造の不安定性に起因している可能性がある、と解釈できます。これらの結果は、1細胞ゲノムDNA複製プロファイルを解析して初めて明らかになった現象であり、いまだにほぼ手付かずの状態にある複製タイミングおよびA/Bコンパートメント構造の制御メカニズムを考える上で重要な発見です。

3. 今後の期待

本研究で、共同研究グループは、ゲノムDNA複製の時間的制御が個々の細胞間でほぼ共通であることを明らかにしました。scRepli-seq法は非常にシンプルで汎用性が高く、あらゆる生物種に適用できるため、今後、DNA複製研究に大きく貢献することは確実です。一方、複製時期が細胞間で比較的大きくゆらぐ領域の存在も明らかになりました。このゆらぎは、ゲノムDNAの局所的なクロマチン構造のゆらぎと考えられ、クロマチン構造制御に関する新たな問いを生み出すものでもあります。

scRepli-seq法の確立により、ゲノムレベルのDNA複製研究、そしてクロマチン構造研究の一部については、細胞数のハードルはもはや完全に乗り越えられたこととなります。今後は本手法を活用して、個体や臓器中の個々の細胞における

ゲノム DNA の振る舞いや状態を理解する研究が進み、複製異常や染色体構造異常を伴うさまざまな疾患の理解にもつながることが期待できます。

4. 論文情報

<タイトル>

Genome-wide stability of the DNA replication program in single mammalian cells

<著者名>

Saori Takahashi, Hisashi Miura, Takahiro Shibata, Koji Nagao, Katsuzumi Okumura, Masato Ogata, Chikashi Obuse, Shin-ichiro Takebayashi, Ichiro Hiratani

<雑誌>

Nature Genetics

<DOI>

10.1038/s41588-019-0347-5

5. 補足説明

[1] ゲノム DNA 複製

細胞分裂における核分裂の前に、DNA ポリメラーゼという酵素によって、ゲノム DNA が過不足なく倍加される過程。ゲノム複製、DNA 複製と同義。ゲノム DNA とは、生物の細胞が持っている全ての DNA 配列情報のことを指し、真核生物の場合は核内にある染色体 DNA を指すことが多い。

[2] クロマチン

真核生物の核内で、ゲノム DNA が形成する高次構造体のこと。その主成分は、DNA とヒストンと呼ばれるタンパク質。

[3] 全ゲノム解析

ヒトやマウスなど、単一の生物種のゲノムを構成する全ての DNA 配列にわたって、特定の性質（RNA 転写量、特定の DNA 結合タンパク質の結合性など）を、次世代シーケンサーを用いて網羅的に調べる解析のこと。

[4] 一塩基多型

ある生物種集団のゲノム DNA 配列中に 1 塩基が変異した多様性が見られ、その変異が集団内で 1 %以上の頻度で見られるとき、これを一塩基多型 (SNP; Single Nucleotide Polymorphism) という。塩基とは核酸の構成要素のことで、DNA の場合はアデニン(A)、チミン(T)、グアニン(G)、シトシン(C)の 4 種類を指す。実際には、1 塩基に限らず、2 塩基から十数塩基程度の短い置換、挿入、欠失などの変異も含めて一塩基多型と呼ぶことも多く、本研究でもこの立場をとっている。

[5] 細胞周期、S 期、間期

増殖中の真核細胞が繰り返す時間的なサイクル（期間）のことを細胞周期と呼ぶ。間

期とは、細胞周期の中で細胞分裂を行う分裂期（M 期）以外の期間のこと。間期は、さらに G1 期、S 期、G2 期に分けられる。S 期は、DNA 複製が行われる期間。

[6] 複製タイミング

ゲノム DNA 複製の時間的制御のこと。S 期において、各ゲノム DNA 領域は、それぞれ特徴的な時期に複製されると考えられている。本研究以前は、その時間的制御が 1 細胞レベルで、どの程度ゆらぐのかは未解決の問題であった。

[7] 転写

染色体 DNA の遺伝子配列をもとに RNA が合成されること。その担い手は RNA ポリメラーゼと呼ばれる酵素。

[8] 相同染色体

二倍体の生物の細胞核内に存在する、父方および母方から由来した形態の等しい一対の染色体のことを指す。ヒトの細胞は、1 番から 22 番染色体まで 22 対の相同染色体を持ち、これに加えて 2 本の性染色体を持つ（オスは XY、メスは XX）。

[9] ES 細胞

胚性幹細胞（Embryonic Stem Cell; ES cell）とも呼ばれる。哺乳類生物の発生初期の胚盤胞期と呼ばれる時期の胚の一部、内部細胞塊と呼ばれる細胞群から作られる幹細胞細胞株のこと。性質は iPS 細胞（人工多能性幹細胞; induced Pluripotent Stem Cell）とよく似ているが、iPS 細胞は体細胞から作られる細胞株である。

[10] 分化、細胞分化

分化ともいう。多細胞生物の発生過程において、特殊化していない未分化の細胞が、より特殊化したタイプの細胞に変化するプロセスのことを指す。

[11] Hi-C 法

3C（Chromosome Conformation Capture）法を発展させた全ゲノム解析手法のことで、細胞核内三次元空間におけるあらゆるゲノム DNA 配列どうしの相対距離を測定して、染色体の三次元的な構造を推定できる画期的な手法。

[12] 塩基対

ゲノム中で 2 本のデオキシリボ核酸（DNA）分子が 2 重らせん構造をとる際に、アデニン(A)とチミン(T)、グアニン(G)とシトシン(C)、という決まった組を作り、水素結合でつながった状態をとるが、その決まった各々の組のことを塩基対と呼ぶ。DNA の長さを表現するときに「～塩基対(base pairs: bp)の DNA」、あるいは単に「～塩基(bases: b)」という表現が用いられることが多い。Mb はメガベース、100 万塩基（対）のこと。Kb はキロベース、1000 塩基（対）のこと。