

世界初！ヒトゲノムのダイナミックな立体構造の計算に成功！

国立大学法人東海国立大学機構 名古屋大学大学院情報学研究科の藤城 新 研究員、笹井 理生 名誉教授（大学院情報学研究科 客員教授）らの研究グループは、ヒト細胞核の中でゲノムが立体構造を形成する過程について、世界最初の高精度計算シミュレーションを行い、構造形成の原理を明らかにしました。ゲノム立体構造は、DNA 情報の読み取りや複製に大きな影響を与えるため、計算シミュレーションによる解明が待たれていました。

本研究では、クロマチン^{注1)}の物理的性質が染色体の中で不均一であることに注目し、クロマチンの不均一な運動によって生じる相分離^{注2)}がゲノム立体構造を形成する駆動力であることを示しました。本研究で開発した計算モデルは、大規模生化学データや顕微鏡データを統合的に説明すると同時に、ゲノム立体構造のダイナミックな揺らぎの予測を可能にするものです。

このことは、ゲノム立体構造とその運動を定量的に解析する新しい方法として、細胞制御機構への理解に向けて重要な貢献をすると期待されます。

本研究成果は、2022 年 5 月 27 日（日本時間）付アメリカ科学アカデミー紀要「Proceedings of the National Academy of Sciences of the United States of America」オンライン版に掲載されました。

本研究は、科学技術振興機構・戦略的創造研究推進事業（CREST）（JPMJCR15G2）、RIKEN Pioneering Project、科研費新学術領域研究「クロマチン潜在能」（19H05258、21H00248）、科研費新学術領域研究「生命の情報物理学」（20H05530）、科研費基盤研究 B（19H01860）、科研費基盤研究 A（22H00406）Nagoya University Research Fund の支援のもとで行われたものです。

【ポイント】

- ・ 染色体の中でクロマチンの物理的性質が不均一に分布すること、そのため、クロマチンに不均一な斥力が生じることを示した。
- ・ 不均一な斥力により、細胞分裂後に新たな細胞核が形成される過程で、クロマチンに相分離が生じることを示し、ゲノム立体構造がクロマチンのダイナミックな運動の産物であることを明らかにした。
- ・ これまでヒト細胞について報告されてきた大規模生化学データ、顕微鏡データを定量的、統一的に説明できる計算モデルを開発し、ゲノム立体構造の計算に成功した。

【研究背景と内容】

私たちの体の設計図である DNA は、10 万分の 1 メートルほどの小さな細胞核の中に収納されていますが、この細胞核内の DNA すべてをまとめてゲノムと呼びます。DNA 情報が読み取られ、あるいは複製される様子は、ゲノム立体構造とその運動によって大きな影響を受けることが、近年、明らかになってきました。そのため、ゲノム立体構造が細胞核内でどのように形成されているかは、細胞を理解するための焦眉の問題として強い関心を集めてきました。

ゲノム立体構造の謎を解くため、様々な生化学測定法や顕微鏡法が開発され、大量の測定データが蓄積されて来ましたが、これらのデータから統一的な描像を得て新しい技術につなげるためには、ゲノム立体構造を計算機で再現する、計算シミュレーションの方法が欠かせないと考えられてきました。しかし、ゲノム立体構造の形成原理が未解明のため、有効な計算モデルが開発されず、大量の測定データに合うように計算機で立体構造の座標を調節するという、データ依存の方法による構造モデル作成がこれまでの計算法の主流でした。

第 10 染色体のクロマチンの物理的性質の配列

IMR90

BuAAAAAAAAAAuBBBBBBBBBBBBBBBBuuuAAAAAAAAAAAAAAAAuuuuuuAAAAAAAAAAuuBuBuu

GM12878

BBuuuuuuuuAAAAuBBBuBBBBBBBBBBBBuuuuuuBBuAAAAuBBBBBBBBBBuuAAAAAAAAAAuuuBBu

図 1: クロマチンは染色体の中で不均一な物理的性質を持つ。塩基 10 万個が連なる領域を 1 つのクロマチン領域として識別すると、密度が低いクロマチン領域を A (黄)、密度が高いクロマチン領域を B (青)、その中間を u (灰色) として、染色体を A, B, u 三文字の配列として表現することができる。例として、ヒト線維芽細胞 (IMR90)、ヒトリンパ芽球様細胞 (GM12878) の第 10 染色体の配列初端を示している。

本研究ではデータ依存の方法ではなく、物理的原理に基づいてゲノム立体構造をシミュレートする新しい計算法を提案しました。そのためにまず、染色体の局所領域ごとに、クロマチンは不均一な物理的性質を持つことを示しました。DNA の塩基対 10 万個が連なる領域を 1 つのクロマチン領域として識別すると、それぞれの染色体は数 100 個から数 1000 個のクロマチン領域がつながった鎖と考えられ、ヒトゲノム^{注3)} は、それらを合わせた約 6 万個のクロマチン領域の集まりとなります。本研究では、これらのクロマチン領域をその物理的性質に応じて A、B、u の 3 種類に分類しました。A 領域は密度が低く遺伝子活性が高い傾向にあり、B 領域は密度が高くて遺伝子活性は低く、u 領域はその中間です。こうして、それぞれの染色体を A、B、u の三文字が連なる配列として表現することができました (図 1)。

各クロマチン領域はそれぞれ密度が異なるため、その間の相互作用も異なると考えられます。密度の低い A 領域どうしは重なることが許されるため、A 領域の相互作用は弱い斥力であると考えられ、より密度の高い B 領域どうしは重なることによって反発を生じ、より強い斥力を示すと考えられます。本研究では、この仮説を積分方程式の理論^{注4)} を用いて確認しました。こうして各染色体は、様々な強さの斥力が不均一に混ざった鎖と考えることができます。

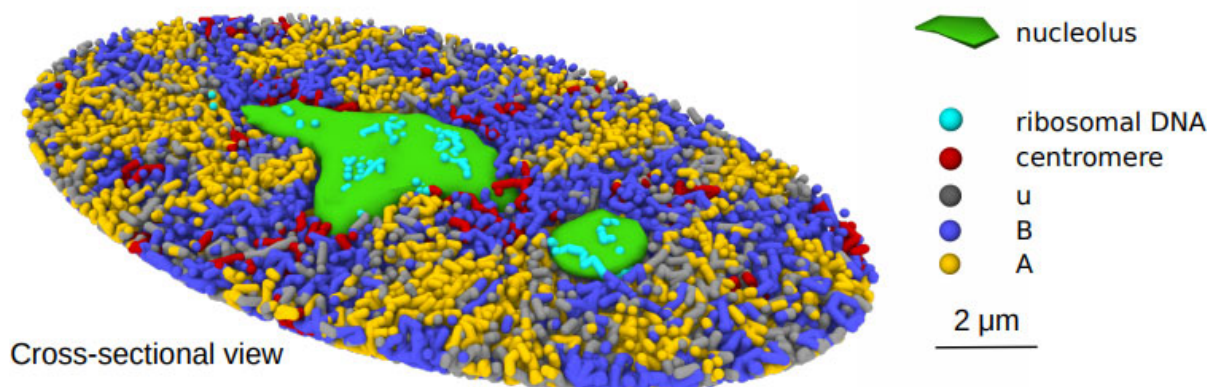


図 2: 計算シミュレーションで得られたゲノム立体構造。A 領域 (黄)、B 領域 (青) は相分離により A コンパートメントおよび B コンパートメントをつくり、u 領域 (灰色) は A/B コンパートメント境界付近に局在する。B コンパートメントは核小体 (緑) や核膜付近に集まる傾向がある。リボソーム RNA 遺伝子をコードしているリボソーム DNA (水色) は核小体に取り込まれる傾向があり、染色体の中ほどにあるセントロメア (赤) は B コンパートメントに埋もれている。ヒト線維芽細胞 (IMR90) の細胞核断面図を例として表している。

本研究では、ヒトゲノムが細胞分裂後の凝縮した染色体の集まりとして存在している状態からシミュレーションを開始しました。細胞分裂が終わると、上記の斥力により染色体は膨張し始めるとともに、核を覆う膜 (核膜) や核小体^{注5)} がつくられ、新しい細胞核が形成されます。この過程で、斥力の不均一さにより不均一な運動が生じ、A 領域、B 領域はそれぞれ同じ領域どうしが集まる相分離を起こし、より大きな領域 (A コンパートメントと B コンパートメント) を形成します。こうして、クロマチン領域の

相分離により、A および B コンパートメントが生じることが明らかになりました（図 2）。また、B コンパートメントは核膜や核小体付近に集まることが示され、ゲノム立体構造の形成原理が明らかにされました。

本研究では、こうして計算されたゲノム立体構造と実験データを比較し、計算結果が多様な実験データを統一的に説明することを示しています。例えば、染色体の各部分のペアが互いに近づく頻度を、Hi-C 法と呼ばれる生化学的な方法で測定したデータ（ゲノム全体にわたる多数のペアを測定した大規模なデータ）と計算の結果は良く一致しています。さらに、染色体の各部分が核膜の近くに来る頻度、各部分が核小体の近くに来る頻度についても、測定データと計算結果は良い一致を示しています（図 3）。その他にも、各染色体の拡がりの程度や、各染色体の細胞核内における位置、生成された核小体の個数についても、計算結果は顕微鏡による測定データを良く説明しています。

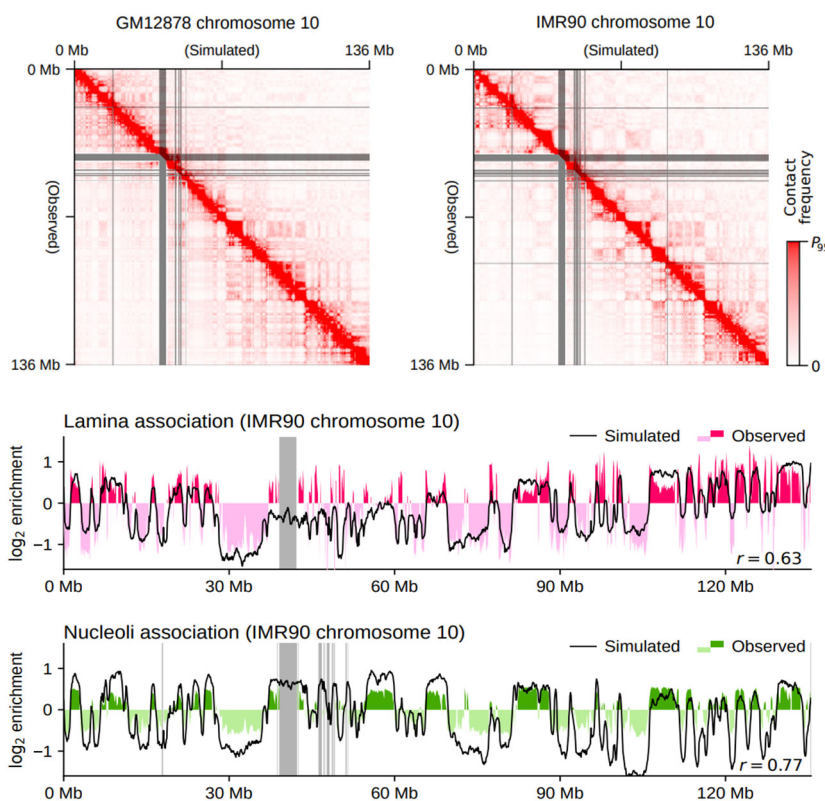


図 3: 測定データと計算結果の比較。(上段) 下左の三角は染色体の各位置どうしの接近頻度についての測定データ (Rao et al. 2014 Cell)。上右の三角は計算結果。(中段) 染色体の各位置が核膜の裏打ち構造であるラミナに接近する頻度。測定データ (Sadaie et al. 2013 Genes Dev.) は接近頻度が高い位置 (赤) と低い位置 (ピンク) を表す。黒い線は計算結果。(下段) 染色体の各位置が核小体に接近する頻度。測定データ (Dillinger et al. 2017 PLOS One) は接近頻度が高い位置 (濃緑) と低い位置 (薄緑) を表す。黒い線は計算結果。ヒト線維芽細胞 (IMR90)、ヒトリンパ芽球様細胞 (GM12878) の第 10 染色体を例として表している。

このように、物理的原理に基づいた計算により、クロマチンのダイナミックな運動によって生じた相分離がゲノム立体構造をつくることが明らかになりました。本研究で

は、細胞分裂後の凝縮した染色体が膨張して、細胞核を形成する過程で、不均一な斥力によって相分離が生じるという新しい考え方を提案しています。この新しい考え方により、これまで解釈の難しかった多様な実験データを統一的に説明することが可能になりました。

さらにこの計算モデルは、安定な細胞核が形成された後も、不均一な斥力による不均一な運動が残ることを予測しています。その結果、速く動くクロマチン領域と遅く動くクロマチン領域がゲノムに混在すること、クロマチンの運動方向のそろった動的なドメインが存在することなど、新しい現象を予測しています。こうしたクロマチンの運動は、様々な因子が DNA に接近する確率を変えて、遺伝子活性の制御に影響を与えることが予想されます。

【成果の意義】

細胞の活動の多くは、DNA から遺伝情報を読み取る過程から始まりますが、この読み取り過程はゲノム立体構造によって大きく影響を受けると考えられ、ゲノム立体構造と細胞の性質の関係は大きな注目を集めています。本研究で明らかになった構造形成の原理は、ゲノム立体構造とその運動を調べる新しい考え方と方法を提供します。

これまでは、ゲノム立体構造形成の原理が不明であったため、有効な計算法が提案されていませんでしたが、本研究による新しい考え方によって提案された計算法は、強力な方法としてその有効性が示されました。これは、ゲノム立体構造と細胞の性質の関係を解明する有力な手掛かりです。本研究の成果によって、細胞の機能が制御されている仕方への理解が進むことが期待されます。

【用語説明】

注 1) クロマチン :

細胞核内の DNA は、ヒストンを始めとした様々なタンパク質と相互作用をしている。このようなタンパク質と DNA の複合物質を、クロマチンと呼ぶ。

注 2) 相分離 :

常温常圧で水と油を混ぜても、水と油はそれぞれまとまって互いに分離するが、このように 2 種類以上の物質が溶け合わずに分離する現象を相分離と呼ぶ。本研究では、クロマチンの A 領域と B 領域が分離して、それぞれより大きな領域 (A コンパートメント、B コンパートメント) にまとまる現象を指す。

注 3) ヒトゲノム :

ヒトの細胞核は、22 種類の染色体それぞれ二つずつに加えて、X 染色体二つ(女性)、あるいは X 染色体と Y 染色体一つずつ(男性)という具合に、合計 46 本の染色体を持つ。これらをまとめてゲノムと呼ぶ。各染色体は短いもので約 5 千万個、長いもので約 2.5 億個の塩基対がつながった DNA 鎖を持ち、ゲノム全体は約 60 億個の塩基対

に相当する。本研究では、塩基対 10 万個のつながりをクロマチン領域の 1 つとみなすため、ゲノムは約 6 万個のクロマチン領域から成り立っていると考える。本研究の計算シミュレーションでは、この約 6 万個のクロマチン領域の細胞核内における空間分布と運動を計算した。

注 4) 積分方程式の理論 :

液体中の分子間相互作用が分かっているとき、積分方程式の方法を用いたオルンシュタイン-ツェルニケ (Ornstein-Zernike) 理論は、分子の空間分布を計算する理論として広く用いられる古典的理論である。この方法を高分子溶液の解析に使えるように拡張したのが、Polymer Reference Interaction Site Model (PRISM) 理論と呼ばれる方法である。これらの理論は、相互作用が与えられたときに分子の空間分布を計算するために用いられることが多いが、逆に、分子の空間分布が分かっているときに分子間の相互作用を求める方法として用いることができる。本研究では、クロマチンの計算シミュレーションにより、その空間分布を求めたのち、その結果から PRISM 理論を用いてクロマチン領域どうしの相互作用を計算し、A 領域間の相互作用、B 領域間の相互作用がそれぞれ強さの異なる斥力であることを導いた。

注 5) 核小体 :

リボソーム RNA およびリボソーム RNA に関連したタンパク質は、細胞核内で集合して液滴状のまとまりをつくる。この液滴のことを核小体 (nucleolus 複数の場合は nucleoli) と呼ぶ。

【論文情報】

雑誌名 : Proceedings of the National Academy of Sciences of the United States of America (アメリカ科学アカデミー紀要)

著者 : Shin Fujishiro and Masaki Sasai (藤城新、笹井理生)

論文投稿時の所属 : 名古屋大学大学院工学研究科応用物理学専攻

現所属 : 名古屋大学大学院情報学研究科複雑系科学専攻

DOI : 10.1073/pnas.2109838119

URL : <https://www.pnas.org/doi/full/10.1073/pnas.2109838119>