

## 染色体の立体構造を1細胞解像度で解析 ～非線形時系列解析手法でより速く精密に～

### 1. 発表者：

平田 祥人（東京大学生産技術研究所 特任准教授）  
小田 有沙（東京大学大学院総合文化研究科 広域科学専攻 特任助教）  
太田 邦史（東京大学大学院総合文化研究科 広域科学専攻 教授）  
合原 一幸（東京大学生産技術研究所 教授）

### 2. 発表のポイント：

- ◆これまで困難であった、1細胞内の染色体の立体構造に関する新しい再構成法を提案した。
- ◆提案した手法は、非線形時系列解析手法を応用するもので、従来法とは異なり個々の細胞の染色体構造が一つに絞り込まれ、高速に再構成することができる。
- ◆本手法により細胞ごとに高解像度の染色体構造を簡便に明らかにすることができ、個々の細胞の染色体変化が重要な意味を持つがん細胞の特性の分析や治療手段の確立が可能になることが、今後期待される。

### 3. 発表概要：

DNA 配列解析技術向上に伴い、染色体(注 1)のどの部分とどの部分が空間的に近いかを調べる Hi-C(注 2)という実験で、1細胞ごとの局所的な染色体相互作用の検出が可能となりつつある。しかし、この局所的な情報の寄せ集めから、長大な染色体の立体構造の全体像を正確に復元することは、非常に困難であった。

そこで、東京大学の平田祥人特任准教授、小田有沙特任助教、太田邦史教授、合原一幸教授らの研究チームは、非線形時系列解析手法の一つである「リカレンスプロット(注 3)」の概念を応用することで、染色体の立体構造を解明する解決策を考案した。Hi-C 実験のデータはリカレンスプロットに数学的な構造が良く似ている(注 4)。また、近年、リカレンスプロットから元の時系列(注 5)を復元する手法が提案され、その手法を裏付ける数学的な証明がなされた。研究チームは、この手法を応用して、1細胞の染色体立体構造を、高解像度で高速に復元することに成功した(図 1)。

染色体の物理的な構造は、さまざまな生理的な現象に影響を与えるため、今後、細胞一つ一つの染色体構造の違いやそれに由来する遺伝子発現のばらつきを解析する事で、iPS 細胞の分化過程、がん細胞の進行度などが1細胞レベルで明らかになると期待される。

この成果は、2016年10月11日にネイチャー・パブリッシング・グループの総合科学雑誌「Scientific Reports (サイエンティフィック・リポーツ誌)」(オンライン版)に掲載される予定である。掲載論文は下記 URL からどなたでも無料で閲覧することができる。

[www.nature.com/articles/srep34982](http://www.nature.com/articles/srep34982)

### 4. 発表内容：

#### ■背景

近年の DNA 配列解析技術の発展に伴い、核内で染色体のどの部分とどの部分の DNA 配列が空間的に近いかどうかという情報が実験的に検出できるようになってきた(Hi-C)。しか

し、特に、少ないサンプル量由来のデータや、細胞1つ1つについて個別の状態を解析する事を目的とした実験から得られるデータは、情報の欠落やノイズも多く、このDNA配列の局所的な隣接情報のみから長大な染色体の元の立体構造を推定するのは難しかった。

## ■内 容

本研究チームは、この問題に数学的な構造が良く似ている非線形時系列解析手法である「リカレンスプロット」の手法を用いて、染色体の立体構造の復元を行うことを提案した。リカレンスプロット自体は、約30年ほど前に提案された、時系列を2次元平面内に視覚化するための手法である。リカレンスプロットは、時系列データを対象として、縦軸、横軸ともに時間軸の2次元平面上において、時系列データの2つの時刻の組に対応する状態同士が近いかどうかを示す。2つの時刻に対応する状態が近ければ、その座標に点を打ち、そうでなければ点を打たないことで、観察対象の時間内で繰り返し起こる再起的なイベントなどの情報が特徴的に可視化出来る。2008年にリカレンスプロットから元の時系列の概形を復元する手法が提案されていて、近年、その手法を数学的に保証する証明が複数得られている。そこで本論文では、リカレンスプロットの手法を染色体の立体構造の復元の問題に利用することを提案した。この手法はノイズやデータの欠落に対し頑強であるため、特にこれまで解析の難しかった1細胞のHi-Cデータにこの手法を応用する事で、速く、確実に、高解像度な染色体の全体像を再構成する事が可能となった。

## ■効 果

従来法では多数の可能な染色体構造を確率的に推定することしかできなかったが、リカレンスプロットから元の時系列を復元する手法を染色体の立体構造の復元に応用する本提案手法では、高速かつ高解像度で各細胞の染色体構造を一つに絞り込んで再構成することができる。生体内の細胞では、染色体構造は、細胞周期や分化状態に依存して絶えず変化しており、この立体構造レベルの制御が細胞の遺伝子発現を調節していることが知られている。本研究チームは、細胞1つ1つについて、染色体の立体構造を解析することを可能とする本手法を提案して、実際のヒトの1細胞Hi-Cデータを用いた解析により、細胞ごとに染色体立体構造が部分的にばらつくことを見いだした。さらに、この解析手法は外部環境の変化によって生じる染色体状態のばらつきを明確化し、さまざまな生理現象や細胞状態を制御するメカニズムを1細胞レベルで紐解く鍵を与えるものと期待される。

## 5. 発表雑誌：

雑誌名：「Scientific Reports」2016年10月11日

論文タイトル：“Three-dimensional reconstruction of single-cell chromosome structure using recurrence plots”

著者：Yoshito Hirata\*, Arisa Oda\*, Kunihiro Ohta, and Kazuyuki Aihara

なお、本成果は本論文の投稿に先立ち、東京大学から特許申請されている（発明の名称：空間的な近さの概念を用いた生体分子データの3次元構造の再構成方法 出願番号：特願2016-023214）。

本研究の一部は、創薬等ライフサイエンス研究支援基盤事業(生命動態システム科学推進拠点事業)により、文部科学省と日本医療研究開発機構を通して助成されたものである。

## 6. 問い合わせ先 :

東京大学 生産技術研究所  
特任准教授 平田 祥人 (ひらた よしと)  
〒153-8505 東京都目黒区駒場 4-6-1  
Tel: 03-5452-6697 Fax: 03-5452-6692  
E-Mail : yoshito@sat.t.u-tokyo.ac.jp

東京大学 総合文化研究科  
特任助教 小田 有沙 (おだ ありさ)  
〒153-8505 東京都目黒区駒場 3-8-1  
Tel: 03-5454-6630 Fax: 03-5465-8834  
E-Mail : odar@bio.c.u-tokyo.ac.jp

東京大学 総合文化研究科  
教授 太田 邦史 (おおた くにひろ)  
〒153-8505 東京都目黒区駒場 3-8-1  
Tel: 03-5465-8834 Fax: 03-5465-8834  
E-Mail : kohta@bio.c.u-tokyo.ac.jp

東京大学 生産技術研究所  
教授 合原 一幸 (あいはら かずゆき)  
〒153-8505 東京都目黒区駒場 4-6-1  
Tel: 03-5452-6691 Fax: 03-5452-6692  
E-Mail : aihara@sat.t.u-tokyo.ac.jp

## 7. 用語解説 :

### (注1) 染色体

遺伝子情報を担う糸状の DNA が凝集して構造体を形成した生体物質。ヒトは 22 本の常染色体と 1 対の性染色体を持つ。

### (注2) Hi-C

細胞内で、染色体のどの部分とどの部分が空間的に近いかを実験的に検出する 3C(chromosome conformation capture) という実験を応用した技術。次世代シーケンサーを用いることで、染色体上の相互作用を網羅的に (high-throughput) 検出できる。近年、細胞ごとに、この染色体情報を解析する技術が進んでいる。

### (注3) リカレンスプロット

元々は、時系列(注5)を視覚的に観察するために作られた手法。2次元平面図で、縦軸、横軸ともに同じ時間軸をとり、2つの時刻の組で、時系列の対応する状態が似ていればその座標に点を打ち、そうでなければ点を打たないことによって得られる。リカレンスプロットは、いわば、時系列の情報を2値の行列情報として表現したものであるが、近年この2値の行列情報から元の時系列の概形を復元する手法が提案されていて、復元可能性に関する数学的証明も複数与えられている。

(注 4) 2つの問題の数学的な類似性

リカレンスプロットから元の時系列を復元する問題も、Hi-C のデータから染色体の立体構造を復元する問題も、突き詰めると、次のような数学的パズルの問題と等価と考えられる：

「AさんとBさんは近所です。BさんとCさんも近所です。でも、AさんとCさんは近所ではありません。では、この町の地図を書いてみましょう」。このようなご近所さんの情報が3次元空間上で多数得られれば、もとの立体的な地図を近似的に復元することが数学的に可能である。

(注 5)時系列

時間とともに何らかの観測によって得られた観測量の系列のこと。本研究では、3次元の時系列と、染色体を構成する DNA 配列の3次元空間内での位置情報の数学的な類似性に着目して研究を行った。

## 8. 添付資料：

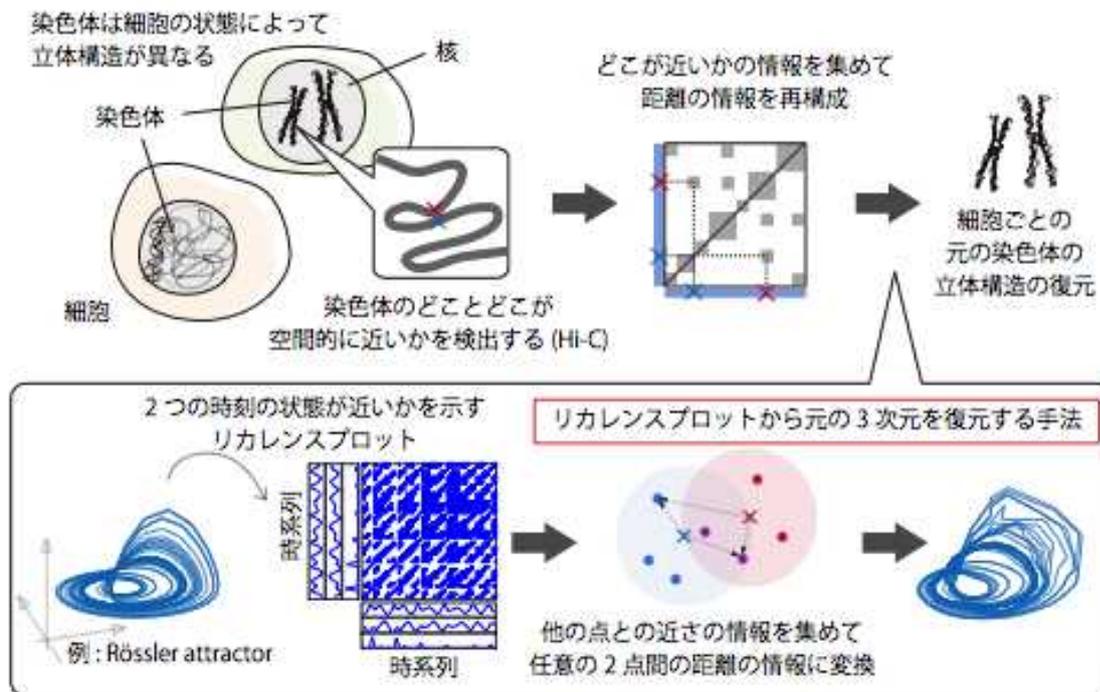


図 1：1細胞の Hi-C と、リカレンスプロットの関係