

武田 洋幸

東京大学大学院理学系研究科
教授

DNA3 次元クロマチン動態の理解と予測

§ 1. 研究実施体制

(1) 武田グループ

- ① 研究代表者: 武田 洋幸 (東京大学大学院理学系研究科、教授)
- ② 研究項目
 - (1) メダカ胚からの epigenetic code、発現情報の収集
 - (2) Hi-C 法による DNA 接触位置情報取得の技術の確立と実施
 - (3) 再解読が完了した近交系 HNI ゲノムを用いてエピジェネティックコードの系統差を同定
 - (4) DNaseI hypersensitive sites をゲノムワイドに収集

(2) 森下グループ(研究機関別)

- ① 主たる共同研究者: 森下 真一 (東京大学大学院新領域創成科学研究科、教授)
- ② 研究項目
 - (1) 接触位置情報の背後にある3次元数理モデルを最尤推定するアルゴリズムを構築
 - (2) モデルの妥当性を公開データにより検証し、epigenetic code の集積性を分析
 - (3) 最良のモデルと見落とさないようにアルゴリズムを高速化

§ 2. 研究実施の概要

平成 26 年度では、メダカ胚細胞からの包括的エピゲノム情報、接触位置情報の取得が順調に進み、森下グループでの3Dクロマチンモデルの構築、検討が始まった。またエピジェネティックコードの成立に必要な cis 配列の探索を機械学習アルゴリズム(Support Vector Machine)を用いて行った。具体的には以下の 3 項目について研究を実施した。

I. クロマチン構造を予測する数理モデル (主に森下グループ)

- ・ 武田グループで観測する多様な細胞系譜での DNA 接触位置頻度から、クロマチン 3 次元構造を高速に計算するプログラムを構築し高速化を達成した。3 次元構造が揺らいでいる様子を、固有ベクトルを変更することで表現できるか否かについて考察を開始した。
- ・ epigenetic code が DNA 上に連続領域として存在する性質を説明するために、巨大なタンパク質複合体が DNA3次元構造にどのように相互作用するかについてモデル化を試み、推測された DNA 3 次元構造が存在しない部分(3 次元的空洞)に、巨大なタンパク質複合体が存在すると考えるアイデアの妥当性を検証中である。

II. 発生・再生過程における包括的エピゲノム情報、接触位置情報、遺伝子発現プロファイルの抽出 (主に武田グループ)

- ・ DNA の 3 次元構造推定の基となる全ゲノム接触位置情報を取得する Hi-C 実験を行った。メダカ胚において良好な実験条件を決定でき、トポロジカルドメイン、コンパートメントといった構造を検出することに成功した。ヒト、ショウジョウバエの先行研究データについても再解析を行い、メダカとの比較解析を進めている。
- ・ 包括的エピゲノム情報の収集については、H3K4me2、H3K27me3、H3K9me3 の ChIP-seq を行った。これらの修飾パターンと高次構造の関連について今後解析する予定である。

III. Epigenetic code の確立とクロマチン構造に寄与する DNA 領域(配列)の同定 (武田グループおよび森下グループ)

- ・ クロマチン構造推定精度の障害となるゲノムの未解読領域を減らすため、平均長 1 万塩基の DNA 断片を解読できる PacBio RS II を利用しメダカゲノムを再解読した。ギャップ数を約 1000 個に減らせる見通しが立った。これにより、以下の多型近交系間での比較が可能となった。
- ・ Epigenetic code に重要な cis DNA 配列を同定するために、メダカの2つの多型近交系(Hd-rR と HNI)において Epigenetic code に違いのある領域を同定した。
- ・ 転写因子などが結合する機能的なゲノム領域を同定するために、メダカ胞胚期において、DNase hypersensitive site の網羅的同定を行った。
- ・ 機械学習アルゴリズム(Support Vector Machine)を用いることで Epigenetic code を DNA の一次配列から予測できるかどうか検証し、各種の修飾状態に特徴的な配列を同定した。

代表的な論文

●1

Ryohei Nakamura, Tatsuya Tsukahara, Wei Qu, Kazuki Ichikawa, Takayoshi Otsuka, Katsumi Ogoshi, Taro L Saito, Kouji Matsushima, Sumio Sugano, Shinichi Hashimoto, Yutaka Suzuki, Shinichi Morishita and Hiroyuki Takeda “Large hypomethylated domains serve as strong repressive machinery for key developmental genes in vertebrates”, *Development*, vol. 141, pp.2568-2580, 2014 (DOI: dev.108548 [pii]10.1242/dev.108548)

●2

Kazuki Ichikawa and Shinichi Morishita. “A linear time algorithm for detecting long genomic regions enriched with a specific combination of epigenetic states”, *BMC Genomics*, Vol.16 (Suppl. 2), pp. S8, 2015 (DOI:10.1186/1471-2164-16-S2-S8)

●3

Kazuki Ichikawa and Shinichi Morishita. “A simple but powerful heuristic method for accelerating k-means clustering of large-scale data in life science,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 11, Issue 4, pp. 681-692, July/August 2014 (DOI: 10.1109/TCBB.2014.2306200)