

論文内容の要旨

論文題目：

In silico Analyses of Nucleosome Positioning and Dynamics

(ヌクレオソームの配置とダイナミクスの *in silico* 解析)

氏名 田中 義章

1. 序論

真核生物はクロマチン構造を形成して巨大なゲノム DNA を核内に収納している。このクロマチンの基本単位であるヌクレオソームは多くの制御因子（転写因子など）の DNA へのアクセスを制限するため、転写や複製などの細胞内プロセスの制御に重要である。またヌクレオソーム配置は細胞分化や外部環境等によって変化することも知られており、プロモーターなどの機能的な領域上のヌクレオソームリモデリングは遺伝子発現誘導調節の一端を担っている。そのためヌクレオソームの配置やダイナミクスを理解することは転写制御機構の解析に必要不可欠である。そこで本論文ではヌクレオソームの配置とダイナミクスに着目し、大量ヌクレオソーム配列データとコンピューターを用いた3つの研究を行った。

2. 研究内容

2.1. 既存のヌクレオソーム配置予測モデルの比較と生物間でのヌクレオソーム配置の DNA 配列依存性の比較

<背景>

ヌクレオソーム配置には DNA 配列依存性があることが知られている。例えば AA/TT や GC が 10bp 周期で現れる配列はヌクレオソームを形成しやすいことが知られている。近年ではこのような配列依存性を利用したコンピューターによるヌクレオソーム配置予測モデルが数多く提案されている。しかしこれらのモデルのうち、どの方法が最も実験データとよく一

致するのか（精度が高いのか）、またこれらのモデルがヒトから酵母までどのようなゲノムにも適応できるのか十分な議論がされていなかった。そこで本研究ではヒト、メダカ、ショウジョウバエ、線虫、酵母のヌクレオソーム分布のデータを用い、代表的な既存の3つのモデル（Segal *et al.* 2006, Gupta *et al.* 2008, Miele *et al.* 2008）について精度の評価を行った。

<結果と考察>

実験的に同定された 1000 個のヌクレオソーム・非ヌクレオソーム配列をそれぞれ陽性・陰性の正解データとして用意し、各モデルの予測精度を Receiver Operating Characteristic (ROC)カーブで評価した（図1）。その結果、ほとんどのゲノムで Segal らの最新のモデルと Gupta らのモデルが高い精度を示したが、メダカでは Gupta らのモデルのみが精度が高かった。また生物間でのヌクレオソーム配置の配列依存性の違いを見るために、各生物のヌクレオソーム配列・非ヌクレオソーム配列で富むオリゴヌクレオチドを比較解析した。その結果、全生物のヌクレオソーム配列で共通に観察されたのは CA/TG と AC/GT でだけであった。また非ヌクレオソーム配列においてはメダカ以外の生物では polyA 等の AT-rich な配列が観察されたが、メダカでは GC-rich なオリゴヌクレオチドが観察された。このことからヌクレオソームの配列依存性の生物間で相違が示唆された。

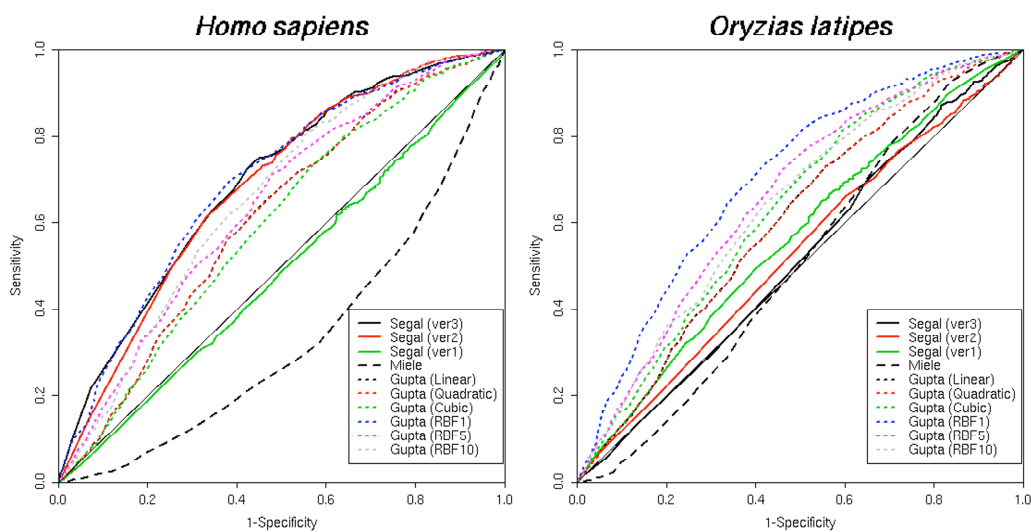


図1. ヒトとメダカにおける各モデルのヌクレオソーム配置予測精度の比較。

(Y. Tanaka and K. Nakai, *Genome Informatics*, 2009 より引用)

2.2. Alu エLEMENTのヌクレオソーム配置に与える影響

<背景>

ゲノムの約 70~80%はヌクレオソームによって占められていることが知られている。そこで次に我々はゲノム全体でヌクレオソームの配置がどれだけヌクレオチドの周期性による影響を受けているのかということに着目した。本研究ではまず16種類のゲノム配列に対しフーリエ解析を行うことでゲノムワイドなヌクレオチドの周期性を観察した。

<結果と考察>

フーリエ解析の結果、哺乳類以上の生物ではヌクレオソーム形成に大きく関わるものが知

られている 10bp の周期性が観察されず、霊長類ゲノムでは別に 84bp と 167bp の周期性が特異的に観察された。この 167bp はコアヒストン(147bp)とリンカーヒストン(20bp)が結合する領域の長さに一致する。さらにこれらの周期性は霊長類ゲノム特異的な反復配列の Alu エlementを除くと観察されなくなることも明らかになった。そこで次にモノヌクレオソームの paired-end シーケンシングデータを使って、Alu エlementとその周辺のヌクレオソーム分布を調べた。その結果、Alu とその周辺のヌクレオソーム配置が整列化されていることが明らかとなった (図 2)。また別のヌクレオソーム分布のタイリングアレイデータを使って同様の解析を行うと、周辺領域におけるヌクレオソーム配置の整列化が観察された。よって Alu 配列にはその中だけでなく周辺のヌクレオソーム配置にも影響を与えることが示唆された。

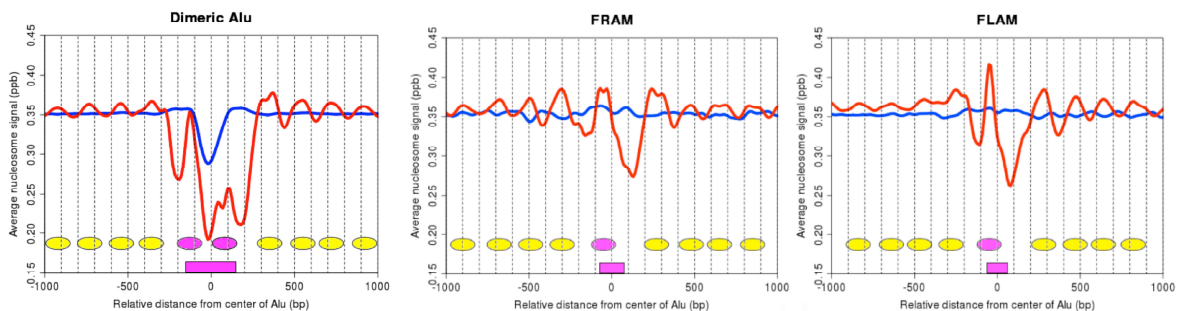


図 2. Alu エlement周辺のヌクレオソームの分布。赤線はユニークにマッピングされたヌクレオソーム DNA タグ、青線はランダムに抽出したゲノム断片から作成。円は赤線より推定したヌクレオソーム位置を示す。X 軸は Alu の中心位置からの距離、Y 軸はヌクレオソーム DNA タグの頻度の平均値を表す。

(Y. Tanaka, R. Yamashita, Y. Suzuki and K. Nakai, *BMC Genomics*, 2010 より引用)

2.3. ゲノムワイドなヌクレオソームリモデリングの解析

<背景>

ヌクレオソームリモデリングはこれまで一部のプロモーター上でのみで解析されてきたが、ゲノム全体でどのヌクレオソームが安定または不安定に配置されているのか十分にわかっていなかった。そこで我々は異なる条件下で得られたヌクレオソーム分布のデータを比較解析することで、ゲノムワイドなヌクレオソームリモデリングのプロファイルを得る方法を提案し、検証・解析を行った。

<結果と考察>

公共データベースに登録されている 7 種類の別々の実験条件から得られた酵母のヌクレオソーム分布のデータから、“entropy”と“linker ratio”の 2 つの指標を作成した (図 3 A)。前者はヌクレオソーム配置のデータセット間での差の度合いを、後者はオープンクロマチンの割合を示す。ヌクレオソーム配置が unstable か stable か実験的に証明されたデータ間で entropy 値を比較したところ、unstable なヌクレオソームは entropy 値が有意に高いことが示された(図 3B)。また転写因子結合サイト周辺について両値を調べたところ、クロマチンリモデリング活性を持つと報告されている 4 つの転写因子は“entropy”と“linker ratio”が他の因子に比べ高く、一方、リモデリングの阻害機能が報告されている Yhp1 は両方の値が全因子

の中で最も低い値を示した。以上の結果から我々が提唱した2つの指標はヌクレオソームリモデリングの度合いを反映していることが示唆された。

さらにプロモーター領域における本指標を利用し、遺伝子発現との関連を解析した。遺伝子発現量としては100以上のDNAマイクロアレイの標準偏差値を使った「発現変化の大きさ」とRNA-Seqのタグ数による「発現量の強さ」について解析した。その結果、プロモーター領域における”entropy”値は「発現変化の大きさ」と有意に相関するが、「発現量の強さ」とは相関しないことが明らかとなった。このことからプロモーターにおけるヌクレオソームリモデリングはmRNA量を制御せず、遺伝子発現のon/offのみを制御することが示唆された。また他のエピジェネティック因子との関連としてヒストンメチル化・アセチル化のChIP-chipデータとの比較解析も行った。その結果H3K4me1、H3K4me3、H3K14acがentropy値と有意な相関を示した。特にエンハンサーマーカであるH3K4me1が富む領域ではentropyの値が有意に高く、ヌクレオソームリモデリングが起りやすいことを示した。

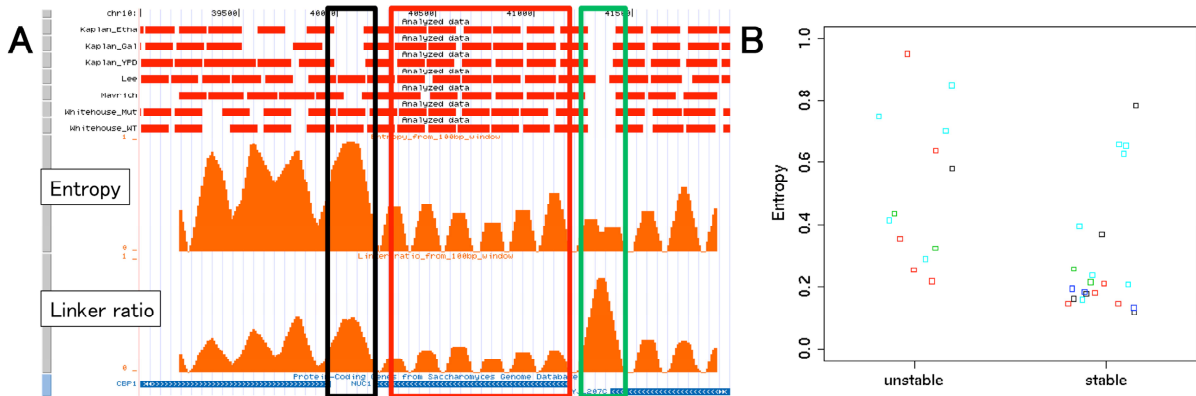


図3. A) ヌクレオソームリモデリングのプロファイル。赤のボックスは各データセットのヌクレオソーム位置を示す。黒枠や緑枠のようにヌクレオソームの欠失がある領域では”linker ratio”の値が高くなり、赤枠のようなヌクレオソームが密に存在する部分(closed chromatin)では低くなる。一方でヌクレオソームの欠失が一部(dynamic open chromatin、黒枠)か全データセット(stable open chromatin、緑枠)かは”entropy”の値で区別できる。B) 既知の unstable と stable ヌクレオソーム間での entropy 値の差。

(Y. Tanaka, I. Yoshimura and K. Nakai, *Chromosoma*, 2010 より引用)

3. 結論

ここ数年の間、ヌクレオソーム、DNAメチル化、ヒストン修飾といったゲノムの後天的な修飾(エピゲノム)情報が大量に得られるようになった。現在、これらの情報は遺伝子発現制御の基礎研究だけでなく再生医療や疾患などの解析にも応用されており、今後大規模なエピゲノム情報を利用した解析が様々な分野で必要不可欠になっていくと考えられる。そこで我々はヌクレオソームデータに重点を置き、そのDNA配列依存性の生物間における特徴やリモデリングの解析を示した。本研究で得られた結果は、今後ヌクレオソームの形成のメカニズムや遺伝子発現制御等への影響を理解していく上で重要な結果であると考えられる。