

機関番号：82110

研究種目：基盤研究 (B)

研究期間：2008～2010

課題番号：20300103

研究課題名 (和文) 転写因子の結合とヌクレオソーム構造にもとづく遺伝子転写制御機構の推定

研究課題名 (英文) Regulation of Gene Transcription Inferred From Transcription Factor Binding and Nucleosome Positioning

研究代表者

河野 秀俊 (KONO HIDETOSHI)

独立行政法人日本原子力研究開発機構・量子ビーム応用研究部門・研究主幹

研究者番号：40291918

研究成果の概要 (和文) : 遺伝子の発現は、転写因子の特定DNA配列への結合とヌクレオソーム構造によって制御されている。本研究では、研究が進んでいないヌクレオソームの位置を予測する方法を開発し、酵母ゲノムに対して適用した。結果、予測されたヌクレオソーム位置の転写開始点付近での分布から、酵母遺伝子の転写パターンは大きく5つに分類されることがわかった。GC配列の含有量が多いものほどヌクレオソーム構造を形成しやすい傾向にあり、この含有量と特定の2塩基配列の周期性がヌクレオソーム形成に深く関わっていることが示唆された。

研究成果の概要 (英文) : Gene Regulation is controlled by binding of regulator proteins to DNA and positioning of nucleosome structures. We have developed a method that can predict nucleosome positions based on the physico-chemical properties of DNA sequences. The predicted distributions of nucleosome positions around transcription start sites of yeast genes indicate that transcription patterns of the genes can be classified into five groups. In addition, GC content was highly correlated with the distribution of nucleosomes.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	3,400,000	1,020,000	4,420,000
2009年度	5,100,000	1,530,000	6,630,000
2010年度	5,000,000	1,500,000	6,500,000
年度			
年度			
総計	13,500,000	4,050,000	17,550,000

研究分野：総合領域

科研費の分科・細目：生体生命情報学

キーワード：バイオインフォマティクス、ヌクレオソーム、転写、遺伝子発現、制御

1. 研究開始当初の背景

今日、500種類以上の生物のゲノムが読まれ、我々はDNA配列として大量のデジタル情報を手にした。生命現象を解き明かすためには、このゲノムに書き込まれたデジタル情報を読み解き、遺伝子ネットワークを明らかにする必要がある。遺伝子ネットワークを知るためには、最も根幹的で重要な生物機能のひとつ、遺伝子の転写制御関係を理解する必要がある。現在、さまざまなハイスループット実験が開発、実施され、酵母に関してはクロマチン免疫沈降法によって約200の転写因子と遺伝子の関係が解き明かされるまでになっている。遺伝子の転写は、特異的なDNA配列に結合するタンパク質（転写因子）によって、巧妙に制御されている。従って、転写因子のターゲット配列を特定することができれば、どの転写因子がどの遺伝子の発現を制御しているか知ることができる。しかし、その認識塩基配列パターンは、ゲノム上に非常に多く存在し、実際に転写因子が結合するのはその一部である。

転写制御のもうひとつの重要な要素は、核内でのDNAの構造、すなわち、クロマチン構造である。クロマチンは、ヒストンのまわりに約150塩基対のDNAが2回巻きついたヌクレオソームを構造単位とし、多数のヌクレオソームが凝集した構造である。転写因子がゲノム上のターゲット領域に結合するためには、その領域がヌクレオソームからほどけてタンパク質が結合できる状態になければならない。2005年には、酵母の第3染色体および約500の遺伝子の上流1kbの領域において、ヌクレオソームの位置が初めて大規模に調べられ、多くの遺伝子で転写制御領域のDNAはヌクレオソームからほどけていることが明らかにされた。

最近、申請者らはDNA配列に依存した構造変形能を調べるために、4塩基対すべての配列パターンについて分子動力学計算を行い、配列ごとに構造変形のしやすさや平均構造を定量的に特徴づけた。このDNA構造を特徴づけるパラメータとヌクレオソーム位置の実験データ、ヌクレオソームの立体構造を用いれば、ゲノム中でのヌクレオソーム位置を予測する方法を開発できると考える。また、ヌクレオソームの位置と転写因子の結合領域を統合的に見ることによって、転写機構の理解を深めることができると考える。

2. 研究の目的

(1) DNAの配列に依存した物理化学的な性質（アナログ情報）にもとづいて、ヌクレオソームの位置を予測する方法を開発する。

(2) ヌクレオソーム位置分布から遺伝子を分類し、その転写機構を探る。

(3) ヌクレオソーム位置とDNAのメチル化位置の関係、細胞種間の違いから特徴的な量を見つけることで転写機構を探る。

3. 研究の方法

(1) ヌクレオソーム位置の予測方法の開発

これまで、DNAは塩基配列ごとに異なる力学特性をもつことをシミュレーション計算や立体構造の解析により示してきた。そこで、塩基配列の力学特性にもとづいて、ゲノム配列の中でヌクレオソーム構造によく適合する塩基配列部位とそうでない配列部位を推定する方法の開発を行う。

(2) ヌクレオソーム位置予測方法の検証

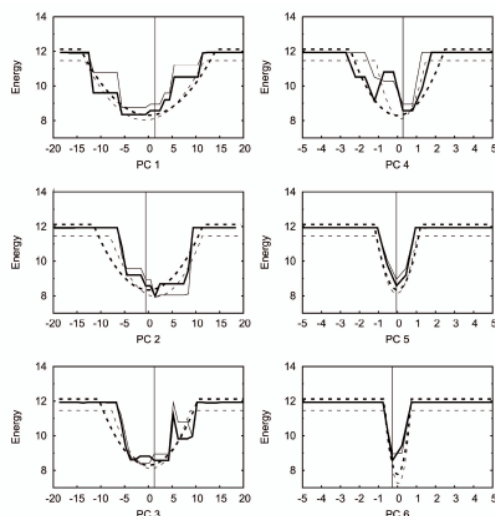
実験的にヌクレオソームの位置が明らかにされている酵母のゲノムの一部などを使い、推定した結果が正しいかどうか評価する。

(3) ヌクレオソーム位置分布による遺伝子のクラスター解析

酵母のゲノムに対して、転写開始部位の上流、下流それぞれ2kbの領域に対して、ヌクレオソームポジションの予測を行う。ヌクレオソームポジションの分布に従って、遺伝子をクラスター分類する。

(4) ヌクレオソーム部位とDNAのメチル化部位の比較

実験的にわかっているヌクレオソーム部



位とメチル化部位を同時にゲノム上にマッピングしたデータベースを作成し、両者の関係を調べる。

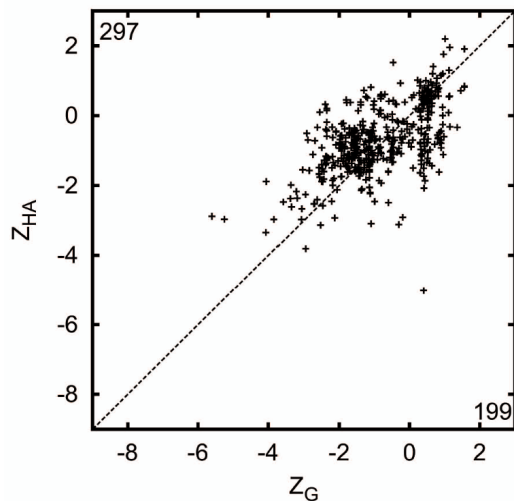
4. 研究の成果

(1) ヌクレオソーム位置の予測方法の開発

X線結晶構造解析など実験的に決定されているDNAの立体構造データは未だ数が少なく、いろいろな塩基配列パターンに対して十分な構造情報がない。そこで、すべての4塩基対を網羅するように分子動力学計算を行い、そのサンプリングした構造から力学特性を決定した。これまでの解析では、サンプルされた構造の分布をガウシアン分布を仮定して解析を行っていたが、必ずしも分布はガウシアン分布を満たさない。そこで、より正確に分布を表現するために、構造を表現する6自由度のパラメータ空間においてグリッドを切り、確率密度関数として分布を表現した。そして、その分布の対数をエネルギーとし、DNA配列がヌクレオソームの構造に適合するかどうか調べた。

図1. エネルギー関数の比較。実線が確率密度分布に忠実なエネルギー関数。破線がガウス近似した関数。太線はGTTA配列、細線は、GTCG配列に対するエネルギー。

ヌクレオソームに適用する前に、蛋白質-DNA複合体のDNA構造をこのエネルギー関数(6次元)で評価し、その有効性を検

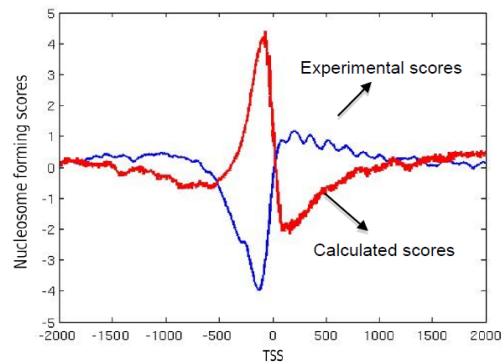


証した。図1に従来のエネルギー関数と改良したエネルギー関数を示す。ガウス関数で構造の分布を表現したエネルギーは、各自由度に対して2次曲線(点線)になるが、分子動力学計算によってサンプルされたアンサンブルに忠実に従った分布からのエネルギーは凸凹しているのがわかる。

構造既知の496の4塩基対に対してエネルギーを計算した結果を図2に示す。X線結晶構造は安定な構造と考えられるので、ランダムな塩基配列に比べて、結晶構造に見られる塩基配列のエネルギーは低いはずである。ランダムな塩基配列に対するエネルギー分布からのずれ(Zスコア)として結晶構造の塩基配列のエネルギーを評価すると、ガウス近似

のエネルギー関数、確率密度分布にもとづくエネルギー関数ともにZスコアがマイナスになっているのがわかる。さらに、この傾向は確率密度分布にもとづくエネルギー関数の方が強く、ガウス近似を仮定した関数よりもDNAの構造を評価するのに適切であることを示す。

図2. 496の4塩基対DNAの構造エネルギーのZスコア。 Z_G は確率密度関数にもとづくエネルギー関数、 Z_{HA} はガウス近似にもとづくエネルギー関数。

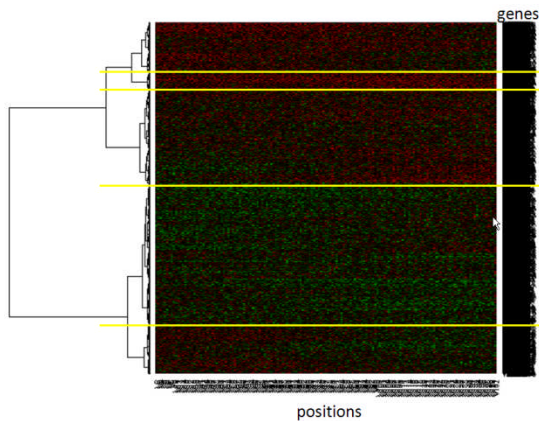


(2) ヌクレオソーム位置予測方法の検証

酵母ゲノムにおいてヌクレオソームを形成しているDNA配列の特徴解析を上記のポテンシャル関数を用いて行った。結果、ヌクレオソームを形成しているDNA配列のエネルギーは低いこと、つまり、このポテンシャルがヌクレオソーム位置予測に使えることを確認した。図3に酵母のゲノムに対するヌクレオソーム形成スコア(ヌクレオソーム構造に対するエネルギー分布を規格化してスコアにした)と実験から得られたヌクレオソーム形成傾向スコアを示す。これからわかるように、両スコアは見事に反比例し、ヌクレオソームを形成し難い塩基配列は、構造エネルギーが高く、ヌクレオソーム構造をとり難いことがわかる。また、その領域は転写開始点の50bp上流付近で強い。また、転写開始点直後はヌクレオソーム形成傾向が強く、徐々にその傾向が弱まっていくのも実験結果によく一致している。

図3. ヌクレオソーム形成傾向スコア。青：実験データ。赤：計算結果。

さらに、詳細にエネルギーの分布をみると、10塩基周期のエネルギーの振動が観測され、特に、CC, CTでそれが顕著であった。このことは、CC, CTが10塩基ごとに出現することを意味している。さらに、エネルギーとDNAのGC含有量に相関があることを見出した。酵母の場合、ヌクレオソームを形成しにくい領域はGC含有量が約0.33であるのに対し、形成しやすいところは約0.41とGC含有量が



大きくなっていった。

(3) ヌクレオソーム位置分布による遺伝子のクラスター解析

ヌクレオソーム位置予測方法と実験的にわかっているヌクレオソーム位置の分布は高い相関を示したので、酵母のすべての遺伝子の転写開始地点の上流、下流それぞれ 2 kbp についてヌクレオソーム位置分布を予測した。予測されたヌクレオソーム位置の転写開始点付近の分布パターンから遺伝子をクラスター分類した (図 4)。結果、大きく 5 つのパターンにわかれることがわかった。今後は、プロモータ配列の転写誘導の強弱とヌクレオソーム分布パターンの関係など詳細な解析を行っていく。

図 4. 酵母ゲノムのヌクレオソーム分布による遺伝子の分類。転写開始点の前後 150bp の領域を示す。遺伝子が 5 つに分類されることがわかる。赤：ヌクレオソーム形成傾向が高いところ。緑：ヌクレオソームを形成し難いところ。

(4) ヌクレオソーム部位と DNA のメチル化部位の比較

ヒトゲノムについて、ヌクレオソーム位置や細胞種ごとの DNA メチル化部位をゲノムにマッピングしたインハウスデータベースを構築した。両者の位置を比較すると、これまでに指摘されていたように、両者の位置

が重なっていることがわかった。また、メチル化の量が細胞種ごとに著しく異なる遺伝子があることがわかった (図 5)。しかし、これらの遺伝子間に明らかな関係を見出すまでには至らなかった。今後は、これらのデータを活用して、個々の遺伝子の転写制御領域とヌクレオソーム形成能の詳細な関係解析を継続して行い、ヌクレオソーム構造やその上の階層のクロマチン構造の観点から遺伝子発現制御の仕組みを解明していきたい。

図 5. ヒトゲノムのメチル化位置の例。細胞種ごとにメチル化位置が異なる。下図は細胞間のメチル化パターンの相関。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

① Yamasaki, S., Terada, T., Shimizu, K., Kono, H. & Sarai, A.

A Generalized Conformational Energy Function of DNA Derived from Molecular Dynamics Simulations (2009) *Nucleic Acids Res.* 37, e135. 査読有

②Fernandez, M., Fujii, S., Kono, H. & Sarai, A. (2009). Evaluation of DNA Intramolecular Interactions for Nucleosome Positioning in Yeast. *Genome Informatics* 23, 13-20. 査読有

③Yonetani, Y. & Kono, H. (2009). Sequence dependencies of DNA deformability and hydration in the minor groove. *Biophys J* 97, 1138-47. 査読有

[学会発表] (計 14 件)

①Kono, H., Kanaeda, N. & Ishida, H. (2011/03). Free energy profiles for nucleosome position changes. 55th Biophysical Society Annual Meeting, Baltimore, USA.

②河野秀俊. (2010/12). 構造バイオインフォマティクスと分子動力学計算による遺伝子発現制御メカニズムの理解. 第 33 回日本分子生物学会年会 第 83 回日本生科学会大会合同大会, 神戸. (招待講演)

③山崎智, 河野秀俊, 清水謙多郎, 皿井明倫 & 寺田透. (2010/06). 情報エントロピーを用いたタンパク質 - DNA 認識における indirect readout の寄与の評価. 第 10 回日本蛋白質科学会年会, 札幌.

④Fernandez, M., 藤井聡, 河野秀俊 & 皿井明倫. (2010/06). ヌクレオソームポジションにおける間接認識の役割. 第 10 回日本蛋白質科学会年会, 札幌.

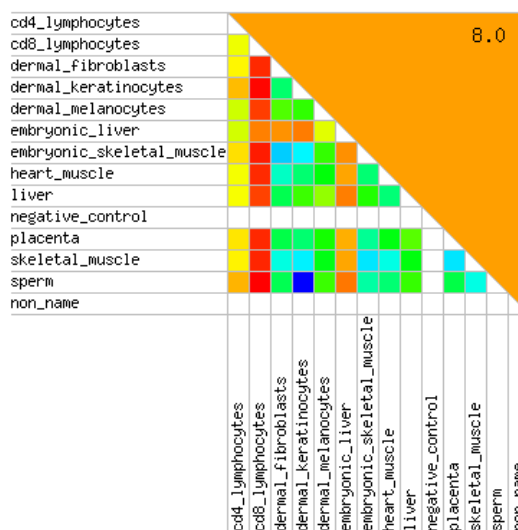
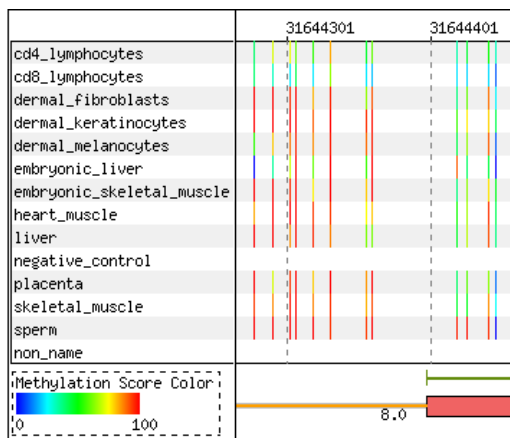
⑤Tanaka, A., Fujii, S., Kono, H. and Sarai, A. (2009/10) Structure-based Analysis of Cooperativity in Protein-DNA Recognition, 第 47 回日本生物物理学会年会, 徳島.

⑥ Yamasaki, S., Fukui, K., Kono, H., Shimizu, K., Sarai, A. & Terada, T. (2009/10). Role of indirect readout in protein-DNA recognition assessed by a Bayesian approach. 第 47 回日本生物物理学会年会, 徳島.

⑦徳久淳師 & 河野秀俊. (2008/6). タンパク質-DNA 相互作用の認識塩基配列と非認識塩基配列による違い. 第 8 回日本蛋白質科学会年会, 東京 (日本) .

[図書] (計 3 件)

①米谷佳晃 & 河野秀俊. (2010). DNA の水和水とダイナミクスの塩基配列依存症—DNA—タンパク質、DNA—低分子認識の理解に向けて. *メディカルバイオ 揺らぎと生体機能* 10 月別冊, 66-71.



②Kono, H.
Chapter 13 Rotamer Libraries for Molecular Modeling and Design of Proteins (Protein Design and Engineering) (2009) Eds. Park&Cochran, Taylor & Francis Group

③皿井明倫 & 河野秀俊. (2008). タンパク質と DNA の結合予測. *実験医学増刊号 バイオデータベースとソフトウェア最前線* 26, 133-139.

[産業財産権]

○出願状況 (計 0 件)

名称 :
発明者 :
権利者 :
種類 :
番号 :
出願年月日 :
国内外の別 :

○取得状況 (計 0 件)

名称 :

発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

河野 秀俊 (KONO HIDETOSHI)

日本原子力研究開発機構・量子ビーム応用
研究部門・生体分子シミュレーショングル
ープ・研究主幹

研究者番号：40291918

(2) 研究分担者

なし ()

(3) 連携研究者

松本 淳 (MATSUMOTO ATSUSHI)

日本原子力研究開発機構・量子ビーム応用
研究部門・生体分子シミュレーショングル
ープ・研究副主幹

研究者番号：10399420