

科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年 6月 3日現在

機関番号：12601  
 研究種目：挑戦的萌芽研究  
 研究期間：2011～2012  
 課題番号：23651199  
 研究課題名（和文） 塩基修飾と次世代シーケンサを用いたクロマチン構造のゲノムワイド単一分子解像度解析  
 研究課題名（英文） Genome-wide analysis of chromatin structure at single molecule resolution by base modification and next-generation sequencing  
 研究代表者  
 伊藤 隆司 (ITO TAKASHI)  
 東京大学・大学院理学系研究科・教授  
 研究者番号：90201326

研究成果の概要（和文）：クロマチン動態の理解を深めるには、ヌクレオソームや転写因子とDNAの相互作用イベント間の連鎖・相関を把握することが重要である。そのために、我々はDNA切断ではなくて塩基修飾の導入と次世代シーケンサによるその網羅的検出を基盤とするゲノムワイドクロマチン解析法の開発を試みた。同様の手法 NOME-Seq が報告されたが、我々独自のバイサルファイトシーケンス技術 PBAT の併用によって、実用的で高感度な解析を実現できる目途が立った。

研究成果の概要（英文）：To improve our understanding of chromatin dynamics, we have to reveal linkage/association among nucleosome/transcription factor-DNA interactions. Toward this goal, we tried to develop a novel method that utilizes base modifications, but not DNA cleavage, to reveal chromatin structure in a genome-wide scale. While a similar method termed NOME-Seq was reported, its combined use with PBAT, a bisulfite sequencing method of our own, would enable a practical and highly-sensitive analysis.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
交付決定額	3,100,000	930,000	4,030,000

研究分野：エピゲノミクス

科研費の分科・細目：ゲノム科学・システムゲノム科学

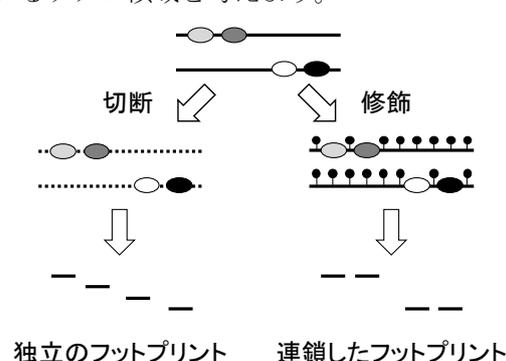
キーワード：クロマチン構造、次世代シーケンサ

1. 研究開始当初の背景

クロマチン構造の動的変化は様々な生命現象の根幹にあり、その理解にはゲノム上のタンパク質結合配置の解明が不可欠である。次世代シーケンサの登場によって、マイクロコッカルヌクレアーゼ消化によるヌクレオソーム配置の決定や DNaseI フットプリント法による転写因子結合部位の検出もゲノムワイド化された。

次に理解すべきは、これらの結合イベント間の関連性(独立・協同・依存・排他 etc)である。しかし、現行法ではその情報が得られない。極端な例として、図の様な状態で4つのヌクレオソーム・転写因子が配置・結合して

いるゲノム領域を考えよう。



現行法で4つの結合部位が同定されたと

しても(図左)、2つの異なるクロマチン状態の存在を見抜くことはできず、4因子が同一分子に同時に結合した状態と理解しがちである。なぜだろうか？それは、現行法がそれぞれの結合部位を「独立」の結合断片またはフットプリントとして検出し、結合イベント間の「連鎖」情報を完全に捨ててしまうからである。つまり、ヌクレアーゼによる DNA の切断は、結合部位間の関連性情報をも切断してしまう。したがって、クロマチン状況の正しい理解には、DNA 切断反応を用いない解析法が必要である。

## 2. 研究の目的

クロマチン動態の理解を深めるには、ヌクレオソームや転写因子の結合部位を同定するのみならず、各結合イベント間の関連性(相関)を把握することが重要である。しかし、ヌクレアーゼに依拠する現行の解析法では、各結合部位が独立の結合断片またはフットプリントとして検出されるために、結合イベント間の関連性に関する情報が失われ、集団全体の平均像しか得られない。そこで我々は、出芽酵母をモデル系に、DNA 切断ではなくて、塩基修飾の導入と次世代シーケンサによるその網羅的検出を基盤とするゲノムワイドクロマチン解析法を開発する。こうして得られるデータの単一分子解像度を活用して、クロマチン構造に関する集団構造や結合イベント間の関連性を把握し、クロマチン動態の本質の理解に迫る新しい研究スタイルの開拓を目指す。

## 3. 研究の方法

塩基修飾の導入と次世代シーケンサによるその網羅的検出を基盤に、ゲノムワイドクロマチン構造解析法を開発する。具体的には、出芽酵母をモデルに以下の3手法の開発に取り組む。

- (1) DNA-(シトシン-C5)-メチレートを *in vivo* で発現させるか、または単離核に作用させた上で、全ゲノムバイサルファイトシーケンサ法(WGBS)でメチル化部位を同定する方法
- (2) DNA-シチジンデアミナーゼを *in vivo* で発現させるか、または単離核に作用させた上で、全ゲノムショットガンシーケンサ法(WGS)で C→T 変異部位を同定する方法
- (3) 単離核・クロマチンに適切な方法でバイサルファイト処理を施した上で、WGS でシトシンの脱アミノ化部位(C→T 変異部位)を同定する方法

## 4. 研究成果

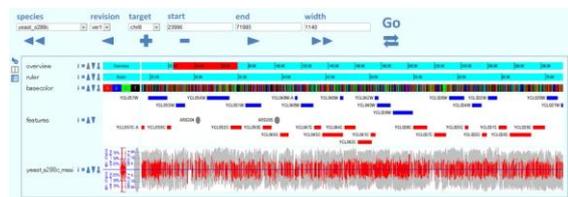
- (1) DNA-(シトシン-C5)-メチレートをによるメチル化

細菌由来の CpG メチレートを *M.SssI* を用いてメチル化の条件検討を進めた。酵母ゲノム DNA を対象にメーカー推奨の条件でメチル化を行い、新たに発売されたメチル化依存性制限酵素 *MspJI* による切断での評価を進めることとした。

しかしながら、結果が安定せず、その原因を追究したところ、*MspJI* の切断自体がしばしば不安定であることが判明した。この酵素にはアクティベーターオリゴヌクレオチドを添加することが当初はメーカーから推奨されており、それに従って使用をしていたが、対照の非メチル化 DNA も切断するなどのトラブルが起こった。やがて、アクティベーターオリゴを入れない反応が推奨されるようになったものの、今後は元来ゲノムがメチル化されている生物の DNA も切れづらくなるなど、その使用には困難が伴った。

そこで切断部位数は減少するが、確実性の高い *MspI* と *HpaII* の組み合わせでの評価を行った。また、*M.SssI* に関して、別の実験からメーカーによって品質が異なることも明らかになったため、当時、市販されていた2社のものをテストした。

更に当研究室に小型の次世代シーケンサ *Illumina MiSeq* が導入されたことを踏まえて、実際にメチル化処理したものをシーケンスして評価を行った(下図)。



上手のグレーの部分はリード数を表し、赤い部分はメチル化レベルを表している。この図からも明らかなように、ゲノム全体が概ね均一にメチル化されるものの、そのレベルは50%程度と不満足なものであった。その後も様々な条件検討を加えたが、満足のゆくレベルのメチル化を得るには至らなかった。

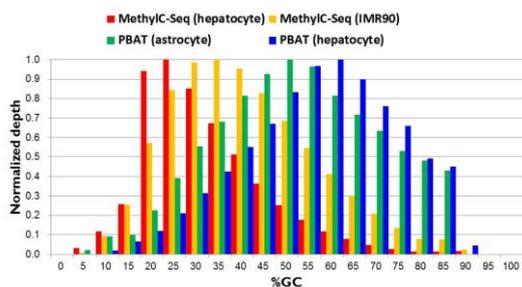
一方、Peter A Jones らは、我々と同様のアプローチに成功して、これを *NOME-Seq* として発表した。彼らは、哺乳類細胞を対象にしているため、内在性メチル化と識別ができない CpG メチル化ではなくて、GpC をメチル化する酵素 *M.CviPI* を利用している。更に *NOME-Seq* キットが *Active Motif* 社から販売されるようになった。したがって、この段階で敢えて我々自身で *M.SssI* によるメチル化を最適化することは断念することにした。

*NOME-Seq* では単離核を *M.CviPI* でメチル化処理した後に、DNA を抽出してから、通常の *MethylC-Seq* プロトコルに沿って全

ゲノムバイサルファイトシーケンシング (WGBS)を行っている。つまりメチル化アダプタ付加後にバイサルファイト変換を施してから、シーケンスを行っている。そのため、5 µg の DNA、つまり 10<sup>6</sup> 個程度の細胞が解析に必要となっている上に、バイサルファイトによる鋳型 DNA の破壊を補うべく、PCR 増幅のステップが含まれている。

我々はこれらの問題を回避すべく、バイサルファイト処理後の DNA にアダプタを付加する独自の Post-Bisulfite Adaptor Tagging (PBAT)法を開発してきた。PBAT を用いると 125 pg からでも PCR 増幅なしで鋳型を作成することが可能であり、30 ng (5 x 10<sup>3</sup> 細胞相当)の DNA があれば、非増幅で哺乳類ゲノムを 30X カバーできる。したがって、PBAT 法を NOME-Seq の後半部分にあたる WGBS のステップで利用すれば格段の高感度が達成される。

NOME-Seq では GC-rich 領域ほど情報が豊かになるが、PCR 増幅を含む従来法ではそのカバー率が低くなることが指摘されていた。そこで我々は同一サンプルを従来法と PBAT 法で WGBS を行って評価した。その結果、GC-rich 領域のカバー率は PBAT 法が圧倒的によいことが分かった (下図)。その意味でも NOME-Seq では PBAT を利用することが有効であろうと考えられた。



以上より、NOME-Seq については市販キットを踏襲しつつ、PBAT を導入することで現実的に有効な高感度解析を実現できる基盤が整えられた。

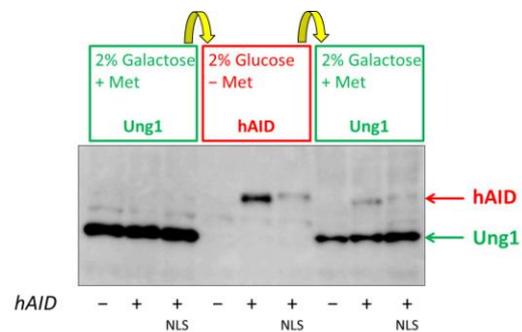
## (2)シチジンデアミナーゼによる脱アミノ化

シチジンデアミナーゼで C を脱アミノ化して U に変換することでフットプリントを得ることを目指した。用いる酵素としては酵母細胞内で発現実績があった AID を用いることにして、産総研 BIRC よりヒト全長 cDNA を入手してこれを利用した。

酵素として利用するには安定な供給源がなかったために、*vivo* での利用を検討した。M.SssI については *vivo* での発現が toxic になることを以前の研究で経験していたため、AID にそれがないことを期待したためでもある。但し、C が U に変換されるとこれを修復するウラシル DNA グリコシラーゼ UDG1

が働く。そこで出芽酵母を改変して、UDG1 非存在下で一過性に AID を誘導できる株を構築することとした。

具体的には、AID をメチオニン欠乏で誘導可能な MET17 プロモータで、UDG1 をグルコースで抑圧可能な GAL1 プロモータで、それぞれ制御する株を作成した。この株を、グルコースを含みメチオニンを含まない培地に移すと AID がアクセシブルなクロマチン領域に C の脱アミノ化を惹起するが、UDG1 が存在しないために修復を受けずに残る。そこでこれらの2つの酵素の発現をレシプロカルに制御可能であるかどうかをテストした。その結果、発現は期待通りにスイッチできることが分かった。AID に GFP を付加した実験では細胞質のシグナルが強かったために、核移行シグナルを付加することも試みたが、発現レベルが減少するという結果が得られた。



併せて、カナバニン耐性株の出現で変異導入率も評価を行った。その結果、GFP やエピトープタグを付加しない AID が最もよい成績を与えることが分かった。しかしながら、変異率が不十分であったので、ミスマッチ修復系の酵素を破壊したりなどの操作も加えた株も作成した。更に、AID による変異導入効果を増強することが報告されている THO 複合体の変異も試みた。

上記のカナバニンテストで評価した上で、次世代シーケンサによるゲノム配列決定を行い、変異の評価を進めた。しかしながら、期待された C→T のトランジション変異については、有意の上昇は見られなかった。

AID の活性が不十分である可能性や、高変異株が淘汰された可能性もあり、*vivo* での利用には一定の限界があることを思わせる結果となった。単離核に対して *in situ* で精製酵素を作用させるメチラーゼと同様のアプローチの方が有効であるかもしれない。

## (3)単離核・クロマチンの直接バイサルファイト処理

上記項目の遅れたために、十分な検討は行えなかった。しかしながら、市販キットのうちに、比較的精製度の低い DNA にも有効な

ものが見出され、今後の検討における有用な情報が得られた。

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1件)

Fumihito Miura, Yusuke Enomoto, Ryo Dairiki and Takashi Ito, Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* 40, e136, 2012 (査読有)

6. 研究組織

(1) 研究代表者

伊藤 隆司 (ITO TAKASHI)

東京大学・大学院理学系研究科・教授

研究者番号：90201326

(2) 研究分担者

なし

(3) 連携研究者

なし