

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 21 日現在

機関番号：12601

研究種目：挑戦的萌芽研究

研究期間：2012～2013

課題番号：24651226

研究課題名(和文) 新型シーケンサによるゲノムワイド・インビボ・フットプリンティング

研究課題名(英文) Genome-wide in vivo footprinting by next-generation sequencing

研究代表者

伊藤 隆司 (Ito, Takashi)

東京大学・理学(系)研究科(研究院)・教授

研究者番号：90201326

交付決定額(研究期間全体)：(直接経費) 3,100,000円、(間接経費) 930,000円

研究成果の概要(和文)：ゲノムの働きは、ゲノムDNAに働きかけるタンパク質によって制御されています。したがって、生きた細胞の中でゲノムDNAとタンパク質が相互作用する様子を網羅的に解き明かすことが出来れば、ゲノムの働き方の理解が格段に深まります。そのための新しい方法として、ジメチル硫酸(DMS)処理と次世代シーケンシングを用いるDMS-Seqと名付けた方法の開発を試みました。条件を最適化した上でこの方法をモデルとしてパン酵母に適用したところ、遺伝子発現を制御することが知られている領域にシグナルを得ることが出来ました。以上の結果は、この方法の有効性を示唆するものと思われまます。

研究成果の概要(英文)：Proteins that act on genomic DNA functionally regulate the genome. Accordingly, if we can comprehensively reveal the interactions between genomic DNA and proteins in living cells, we will substantially deepen our understanding of how the genome works. As a novel method toward this goal, we tried to develop a method termed DMS-Seq that uses dimethylsulfate (DMS) and next-generation sequencing. Following the optimization of the method, we applied it to the baker's yeast as a model system and successfully detected signals on regions known to regulate gene expression. These results would suggest the feasibility of this approach.

研究分野：複合新領域

科研費の分科・細目：システムゲノム科学

キーワード：ゲノム解析技術

1. 研究開始当初の背景

複製・組換え・修復・転写などゲノムにまつわる様々な事象の研究とは、その大半がゲノム DNA と転写因子等のトランス作用因子の相互作用の理解である言っても過言ではない。したがって、その包括的・システム的理解には、両者間の相互作用の全体像を把握する必要がある。

ゲノム DNA とトランス作用因子の相互作用解析において主流となっている方法論は、クロマチン免疫沈降シーケンス法 (ChIP-Seq) である。しかし、その成否は良質の抗体の有無に依存する。良質な抗体がない場合には、トランス作用因子側にエピトープタグを付加する等の方策を取らざるを得ず、生理的発現レベルの維持など別の問題が生じる。しかも、ChIP-Seq は個々の因子に対する解析法であるため、トランス作用因子の数だけ実験しない限り全体像の把握はできない。

これらの制約がない方法として、単離核の DNaseI 処理に基づくデジタルゲノムフットプリント法が報告されている (Nat Methods 6, 283, 2009)。この方法はシーケンス深度を上げて、高感受性部位のみならずフットプリントまでも検出するものであるが、技術的難易度が高く、核の単離が必須であるため厳密には *in vivo* 解析法ではない。

したがって、最も望まれる技術とは、細胞核の単離を必要とせず *in vivo* をより正確に反映できるゲノムワイドフットプリント法である。

2. 研究の目的

ゲノム機能発現の包括的理解には、生細胞におけるゲノム DNA とトランス作用因子の相互作用の全体像を把握する必要がある。そのために用いられる主な技術はクロマチン免疫沈降シーケンス法であるが、その成否は良質の抗体の有無に依存する上に、個別因子ごとの解析しかできない。これらの制約がない手法としては、単離核の DNaseI 処理に基づくデジタルゲノムフットプリント法があるが、核の単離を伴う点で厳密には *in vivo* 解析とは言えず、技術的にも難易度が高い。

そこで本申請は、核を単離することなく生細胞をそのまま処理できる方法として、ジメチル硫酸 *in vivo* フットプリント法に着目し、これを次世代シーケンシングと組み合わせたゲノムワイド *in vivo* フットプリント法の開発に挑戦する。

3. 研究の方法

出芽酵母をモデルに用いて、いくつかの遺伝子座位を標的とした Linker-Mediated (LM)-PCR 法で DMS *in vivo* フットプリンティングの条件を検討し、処理条件の最適化を図る。

上記条件下での処理を行って得られたゲノム DNA を用いて Illumina 用のライブラリ

を作成して、シーケンスを行い、リードをゲノムにマッピングして、既存の DNaseI-Seq の結果と比較検討を通して評価を行う。そのために、接合フェロモン α 因子で処理した細胞としない細胞の比較検討を行う。

4. 研究成果

1) DMS 処理条件の最適化

常法に従って、DMS 処理を行い、単離したゲノム DNA を piperidine 処理に供して切断を行った上で、LM-PCR を試みた。標的遺伝子としては、接合フェロモン α 因子に応答する遺伝子として STE2 および STE3 を採り上げて、条件の検討を進めた。その結果、DMS *in vivo* フットプリンティングの条件の最適化が進んだ。

その過程で DMS 処理後に piperidine 処理を行わなくても十分に DNA が細断化されることが判明した。DMS 処理後に一級アミン存在下で加温するだけでも切断が起こることが報告されていたため、以降、piperidine 処理は行わないこととした。

2) DMS-Seq

まず、接合フェロモン α 因子で処理した酵母としなかった酵母を用意して、それぞれの半分について DMS 処理を行い、DNA を調製した。残り半分については、DMS 処理無の対照として利用するために、常法にしたがってゲノム DNA を調製した上で、DMS 処理細胞と同等の断片長になるように超音波処理を行って DNA の断片化を行った。

上記条件で調製したゲノム DNA について、末端修復および dA 付加を行い、Illumina MiSeq による配列決定を行った。配列決定は 36-nt のペアエンドモードで行った。公共の DNase-Seq との比較を行うために、36-nt のリードは 28-nt にトリミングして使用した。その結果、 α 因子-DMS+、 α 因子-DMS-、 α 因子+DMS+、 α 因子-DMS- の 4 種類について、それぞれ 37 万、51 万、45 万、57 万リードのデータを得ることに成功した。

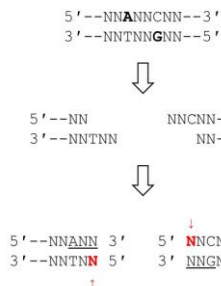
3) DMS-Seq データの解析

次にこれらのリードと DNaseI-Seq の結果の相関を調べるために、ゲノムを 100-nt のウィンドウに分割してそこにマッピングされた塩基数を計算した上で比較を行ったところ、全てのサンプル間で高い相関係数が得られた。その原因を探ったところ、コピー数の高い rRNA 遺伝子座位とミトコンドリア DNA に多数のリードがマップされており、それらの影響で見かけ上は高い相関係数が発生したことが予想された。そこで、これら 2 領域にマッピングされたリードを取り除いた上で再度計算を行った。その結果、超音波処理したものは、予想通りに他のものとほとんど相関を示さなくなった。一方で、 α 因子有無の DMS-Seq 同士および DMS-Seq と DNaseI-Seq では、それぞれ $R=0.74$ および

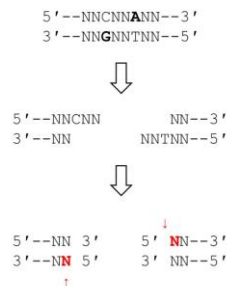
R=0.64 という相関を示すことが判明した。

続いて、読み出し塩基の偏りについての解析を行なった。DMS によって 2 本鎖 DNA が起こるのは、Watson 鎖上と Crick 鎖上で隣接したプリン塩基にメチル化が起こった場合である。そこに末端修復操作を行ってからアダプタを付加することになるので、読み出し塩基（下図の赤）の一塩基 5' 上流がプリン塩基となる筈である（下図参照）。

5'が突出する場合

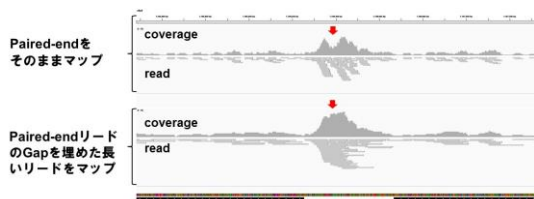


3'が突出する場合



実際にマッピングされたリードについて一塩基上流側の塩基を調べてみると、予想通りにプリン塩基となっており、DMS による切断で生じた末端を正しく捉えていることが示唆された。

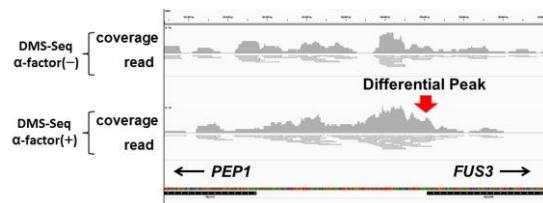
ゲノムにマッピングされたリードを用いて、まずはピークの検出を行った。ペアエンドリードで挟まれた領域を繋いだ形でリードとしてマッピングを行った。この場合にはトランス作用因子が結合した部分は、末端となることはなくて、各リードに共通する内部領域となるために結合部分がピークとして検出される筈である。それに対してペアエンドリードをそれぞれ別個にマッピングして切断点の位置を表現するとトランス作用因子の結合部位は切断点のフットプリントとして表現されると予想される。そこでその双方を試みたところ、下図のように予想通りのパターンが得られた。



まずはピークとしての検出を行うことにして、得られたパターンをゲノムブラウザで可視化したところ、上図にも示すように基本的にピークはプロモータ領域に相当する遺伝子間領域に集中することが判った。このことは DMS による修飾がクロマチンの状態がオープンな領域に集中していることを示している。そのパターンは、概ね DNaseI-Seq と一致する傾向にもあることから、クロマチ

ンアクセシビリティを反映するものと考えられた。

これらのピークの中にはシャープなものとブロードなものが観察された。後者はヌクレオソームによるものである可能性が高く、本研究の目的にはシャープなもの、すなわち切断点間の距離が短いものが重要と考えられたので、総リード数は減少するものの 125-nt 以下になるリードのみを選択して更に解析を進めることとした。これらのリードがマップされた数について、再びウィンドウベースでの解析を行い、接合フェロモン α 因子の有無でマップされるリード数が有意に変化するものを選択した (FDR=0.05)。そのようにして選択したピークについて、前後の遺伝子を抽出して、GO 解析を行ったところ、Conjugation の濃縮が示された。実際に接合に関わる FUS3 遺伝子の 5' 上流のピークが、接合フェロモン α 因子によって誘導されていることも確認できた（下図）。



以上の結果は、DMS-Seq が想定した通りに機能すること、生物学的に意義のある変化を捕捉できることを示唆するものであった。今回の解析ではリード数が少ないこともあって、フットプリントの検出はまだ現実的ではないが、DNaseI-Seq に比較すると格段に簡便に in vivo でのクロマチンアクセシビリティを評価する手法としては有望な手法であることが示されたと考える。

なお上記の解析では、細胞から得られたリードをマッピングしてそのまま用いているが、DNA 配列による DMS 感受性の違いの影響をサブトラクトする必要がある。しかしながら、精製した DNA については現在と同様の条件の DMS 処理では断片化がほとんど起こらない。検討を重ねた結果、1 級アミンの共存が重要であることが判り、putresine 存在下であれば精製 DNA も DMS 処理によって十分に断片化できることも判明した。したがって、これを対照試料として用いることによって、トランス作用因子の結合による DMS のアクセシビリティの変化を評価することが可能になると考えられる。

また、ライブラリの調製についても今回はゲノム DNA 断片をそのまま利用したが、フットプリントを得る場合には主に短い断片を対象に実験を行った方がよいと考えられた。ゲノム DNA の段階か、あるいはアダプタを付加した後の段階において、サイズ選択を行うことでこれは実現できると思われた。

以上の結果は、DMS-Seq が当初の期待通りに新しい in vivo のクロマチン解析法にな

り得ること、更に in vivo におけるタンパク質と DNA との相互作用の総体を把握できる新しい手法ともなり得ることを示唆するものである。

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 0 件)

[学会発表] (計 0 件)

[図書] (計 0 件)

[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

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 : 9 0 2 0 1 3 2 6

(2)研究分担者

なし ()

研究者番号 :

(3)連携研究者

なし ()

研究者番号 :