

平成 21年 6月 8日現在

研究種目：基盤研究(C)

研究期間：2007～2008

課題番号：19591241

研究課題名（和文） 小児白血病のメチル化 DNA のグローバル解析

研究課題名（英文） Global methylated DNA profile analyses of childhood leukemias

研究代表者

堤 修一 (Shuichi Tsutsumi)

東京大学・先端科学技術研究センター・助教

研究者番号 30345152

研究成果の概要：

研究代表者はメチル化 DNA 免疫沈降法(MeDIP 法)とタイリングアレイを組み合わせて、メチル化 DNA の領域を網羅的に取得する方法を開発した。今回はこれをより広範なプロモータアレイ、また高速大量シーケンス法によって全ゲノムレベルに拡張することが目的であった。細胞株のサンプルを使い、全ゲノムレベルのメチル化パターンを取得することに成功した。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2007 年度	2,200,000	660,000	2,860,000
2008 年度	1,300,000	390,000	1,690,000
年度			
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：医師薬学

科研費の分科・細目：内科系臨床医学・小児科学

キーワード：ゲノムワイド解析、DNA メチル化、高速大量シーケンス、エピジェネティクス、DNA 免疫沈降

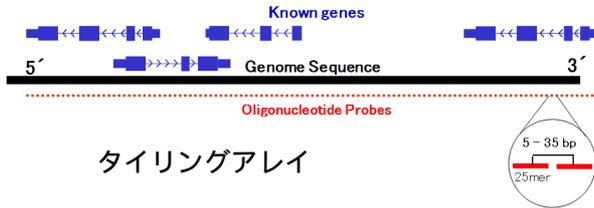
1. 研究開始当初の背景

(1)がん化に伴う遺伝子発現の異常に DNA のメチル化が関与していることが、多く報告されている。白血病においても、染色体転座による融合タンパクなどによって、発がん関連遺伝子の DNA メチル化に変動が認められ、多段階に発症していると考えられている。しかし、DNA メチル化状態を網羅的かつ高解像度で検出する強力な手法は存在しなかった。

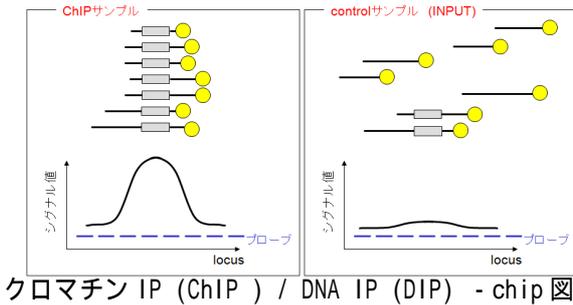
(2)一方で、科学技術の進歩によって詳細なマイクロアレイが開発されてきた。1.25cm X 1.25cm の領域に 600 万種類の配列を任意に作

製可能となったのである。ヒトゲノム配列の 35塩基対(bp)から 25bp を選択するとしても、わずか 7 枚でヒトゲノムをカバーする。この得られた連続したデータの処理によって、サンプル内の濃縮した領域を決定できる。さらに近年、高速大量シーケンス法が開発された。こちらは、1 ラン (= 8 レーン) で約 1 億配列 x 36bp のシーケンスを得ることができる。これをヒトゲノムのリファレンス配列にマップする。サンプル内で濃縮している領域には、マップした数が集中すると考えられる。

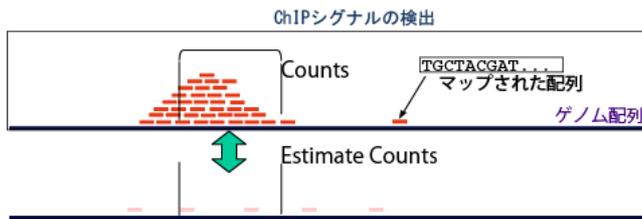
(3) 研究代表者はメチル化シトシン免疫沈降と高密度タイリングマイクロアレイ解析(以下 MeDIP-chip 解析)を用いて全ゲノムレベルで高解像度にメチル化 DNA 領域を検出する手法を開発した。(Hayashi *et.al.Human Genetics* 2006) MeDIP-chip法は前年より数個の報告があった(Weber M *et.al. Nature*



タイリングアレイ



クロマチン IP (ChIP) / DNA IP (DIP) - chip 図



クロマチン IP (ChIP) / DNA IP (DIP) - Seq 図

Genetics 2005, Aug) もの、研究代表者の方法以外に DNA メチル化領域を高信頼性かつ高解像度、網羅的に検出する手法は報告されていない。

2. 研究の目的

代表者はタイリングアレイを用い、主に小児白血病の RefSeq 遺伝子領域における DNA メチル化パターンを取得して比較解析を行うことを目的とする。

3. 研究の方法

(1)ゲノム DNA 2ug よりメチルシトシン抗体により免疫沈降を行う。得られた 5-10ng の DNA には高メチル化 CpG を含む部分が濃縮していると考えられる。Bisulphite シークエンスで既に確認済みの領域が定量 PCR 法で少なくとも 10 倍の濃縮が得られれば、次のステップに移行する。

(2)この配列にアデニンを 10-20 塩基ほ

ど付加して、T7 配列を付加する。ここから T7 RNA ポリメラーゼによる RNA 増幅を行う。通常であれば、ここで数百~1000 倍の増幅が得られる。この RNA 産物を逆転写することにより、再び増幅工程を行う事により、10µg 以上の RNA が得られる。これをタイリングアレイにハイブリダイゼーションする。タイリングアレイは、RefSeq 遺伝子の転写開始点(TSS)付近を設計したプロモータアレイ(1枚)。および全ゲノムアレイ(7枚組)にて解析する。各実験ポイントでは、定量 PCR を行いメチル化領域が 10 倍以上の濃縮が行われている事を適宜確認した。

(3)高速大量シーケンス機(Solexa システム)が導入され、タイリングアレイより簡便かつ低コストで解析が可能となることが判明した。このため、同様に 5ng-10ng のサンプルにアダプターを付加して Solexa システムにて解析を行う。データは濃縮した配列から 50bp ほどの上流や下流の配列を読んでいることとなる。データはマッピングツールにより、ヒトゲノムに一意に決定される領域のカウントから、濃縮している領域を算出できる。

(4)得られている高メチル化領域を近傍の遺伝子との相関で、どのような部分に濃縮があるかを比較する。また、発現アレイを用いて得た遺伝子発現プロファイルと比較する。近傍遺伝子は 10kbp 以内に転写開始点または 3' 端が存在する遺伝子を近傍のものとして、発現スコアと比較した。さらに、メチル化位置と発現スコアの相関も比較した。

(5)タイリングアレイ(プロモータアレイ)とシーケンス(Solexa)をホメオボックスのクラスターで比較した。また、既報(Hayashi *et.al.Human Genetics* 2006)の確認された領域をどの程度検出できたかを比較した。

(6)また同様に代表者が以前行って、公共に得られている同細胞株のアセチル化ヒストンのデータとの重なりを比較した。

4. 研究成果

(1)細胞株 HCT116 よりの DNA からメチル化 DNA 免疫沈降を行い、プロモータアレイにより高メチル化候補領域を得た。論文で以前に報告した領域は、ヒト全ゲ

ノムの 1%の領域で、アレイプローブの密度は 22bp 中に 1 つの設計を行っている。実験はサンプル 3, コントロール 3 であった。

(2) これらメチル化を確認した中で、プロモータアレイにて同様に検出できた領域は 31 ヶ所中 19 ヶ所にすぎなかった。これはサンプル数が $n=2$ であることと、プローブが 35bp に 1 つとやや疎なことが原因と考えられた。全ゲノムデータでは、計 28 枚アレイを必要とするため、Solexa システム法で処理を行う方が良いと考えられた。

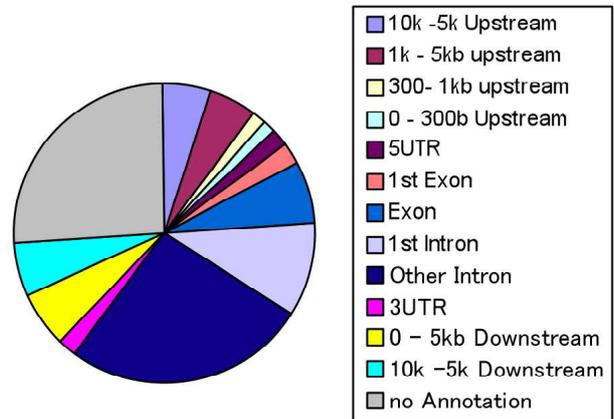
(3) このため、Solexa システムによって、DIP を行ったサンプルから、約 2000 万のシーケンスデータを得た。これらは、ヒトゲノム配列と比較し、マッピングを行った。得られたマッピング領域より約 36% の配列がヒトゲノム領域にマップ可能であった。これは、Input の DNA が 60% 以上のマッピング率から比べると低めである。しかし、テロメアやセントロメア付近のリピート配列はメチル化されていることが原因と考えられた。これらのマッピング領域はゲノム上でランダムにポワソン分布すると帰無仮説をたてて、マップされた数が多い領域を 500bp 幅のウィンドウで計算して確率を算出した。それら P 値の $1e^{-4}$ で有意な領域をメチル化候補領域として出力した。その結果、約 6 万ヶ所の高メチル化領域の候補を得た。

(4) これらの領域を確認すると IGF2 のインプリティング領域などでは、13 ヶ所中 11 ヶ所の領域を検出可能であった。これらの領域は 50% がメチル化されている部分であるが、十分な感度があると考えられる。さらに、Bisulfite シーケンス法によりメチル化している領域の確認を行った。10 ヶ所中 9 ヶ所で 50% 以上のメチル化を確認することができた。しかし、CpG の数が 500bp に数個しかないような低 CpG 領域のメチル化 DNA は感度が低く正しいデータを得る事ができなかった。これは数百塩基対から免疫沈降を行うのに、少ないメチル化シトシンの数では十分に免疫沈降を行うことが出来ないと考えられた。

しかし、この部分までをカバーするようなメチル化 DNA のグローバル解析法は、Bisulphite 処理後 DNA のダイレクトシーケンス法を全ゲノムレベルで行うしかなく、不可能ではないが現状ではまだ高コストである。このため、低 CpG 領

域のメチル化は除外した解析となっている。

(5) 次にメチル化候補領域を既知の遺伝子領域との位置関係で解析した。図のように、遺伝子近傍の領域がよりメチル化されている結果が得られた。



メチル化領域の遺伝子との相関分布

(6) 遺伝子発現との比較をアレイ解析による発現データと比較すると、遺伝子の転写開始点近傍(上流 600bp、下流 100bp)の領域のメチル化と低発現が相関することがわかった。この位置にメチル化が存在すると、近傍の遺伝子は発現が低い傾向にある。しかし、転写開始点から離れて、3' UTR 付近に高メチル化が認められる遺伝子などで、高発現が認められ、他の領域の DNA メチル化とは発現は単純に低発現とは相関しないことが解った。これらの領域におけるメチル化がどのような意味を持つかはさらなる解析が必要と考えられた。

次に、ヒストン活性化のマーカーである、アセチル化ヒストンの ChIP - chip データと比較した。このデータは以前の研究で、代表者が同細胞株で行っていたものである。これらをホメオボックス領域(HoxA)と比較した。すると、アセチル化ヒストン領域と DNA のメチル化は同時には存在せず、明確にクロマチンの状態がコントロールされていることが示唆された。

(7) 現在、他に MEF 細胞や KT 細胞株の全ゲノムデータを取得している。現在の Solexa システムは 1 レーン(1/8 ラン)で 1,000 万以上の配列を読む事ができるため、コスト効率が良い。インプリティング領域については十分な感度でシグナルを有している。しかし、データは現在取得済みではあるが、解析に時間を要している状態である。これらをまとめ

て報告する予定となっている。

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 1 件)

Shuichi Tsutsumi, Methylated DNA profile analysis using MeDIP - chip, Gordon Research Conference ; Cancer Genetics & Epigenetics 20 -25 May 2007, Lucca (Barga), Italy

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

なし

6. 研究組織

(1) 研究代表者

堤 修一 (Tsutsumi Shuichi)
東京大学・先端科学技術研究センター・助教
研究者番号 : 30345152

(2) 研究分担者

なし

(3) 連携研究者

なし