

# ゲノム転移因子と胚性幹細胞エピゲノム構築に関する研究

村本 玄紀

東京大学大学院 農学生命科学研究科

応用動物科学専攻

平成 20 年度 博士課程進学

指導教員 塩田邦郎

目次	頁
緒言	2
第一章	14
序論	15
材料および方法	17
結果	25
考察	31
図と表	34
第二章	51
序論	52
材料および方法	58
結果	62
考察	71
図と表	75
総合討論	108
引用文献	117
論文要旨	127
謝辞	130

## 緒言

エピジェネティクスとは、「塩基配列の変化を伴わず、細胞世代を超えて継承される遺伝子機能の変化、またはこの現象を研究する学問分野」と定義される。哺乳類の発生・分化過程においては、1つの受精卵から成体を構成する約200種類の細胞が生じる。この過程において、免疫系細胞などの一部の例外を除き、DNAの塩基配列に変化は起きない。同じゲノム情報を有しながら、分化系譜ごとに遺伝子の発現・抑制のパターン（トランスクリプトーム）が形成され、細胞分裂後も引き継がれるのは、エピジェネティックな遺伝子発現機構が存在するからである。哺乳類の発生・分化の本質は、受精卵から成体へ至る過程でのゲノム全体のエピジェネティック情報（エピゲノム）の変化および確立であるといえる。

哺乳類のエピジェネティック制御の主要な分子機構としてDNAメチル化とヒストン修飾がある（図1）。DNAメチル化は哺乳類ゲノムに、唯一広範に認められる化学修飾である。DNAメチル化は遺伝子発現抑制に加え、染色体の安定化、X染色体の不活化、ゲノムインプリンティング、ゲノム転移因子の抑制といった様々な生命現象に関わる（Bird, 2002; Shiota, 2004）。主にシトシンとグアニンが並び合った配列（CpG配列）のシトシン塩基の5位にメチル基が付加される（図2）。哺乳類におけるDNAメチル化はDNAメチル基転移酵素（DNA methyltransferase、DNMT）によってなされる（Li, 2002）。メチル化修飾の活性を持つ遺伝子として*Dnmt1*、*Dnmt3a*、*Dnmt3b*が存在する。*Dnmt1*はDNA複製時に親鎖上のメチル化修飾を認識し娘鎖にメチル化修飾を行う活性（維持メチル化活性）を、*Dnmt3a*と*Dnmt3b*は新規にメチル化修飾を行う活性（*de novo*メチル化活性）を担う。メチル化された遺伝子領域では一般的に転写が不活性化され、遺伝子発現は抑制される。これは、DNAがメチル化されることによって転写因子の結合が阻害される、あるいはメチル化されたDNAを認識するDNA結合因子が結合することで転写因子を阻害すること

による (Bird, 2002)。DNAメチル化は、後述するヒストン修飾と協調してエピジェネティックな遺伝子発現制御を行っている (Ikegami et al., 2009)。

ヒストン修飾とは、クロマチンを構成するコアヒストンH2A、H2B、H3、H4が翻訳後修飾を受ける現象であり、ヒストンテールと呼ばれるヒストンのN端が主に修飾される (Kouzarides, 2007)。これまでにアセチル化、メチル化、リン酸化、ユビキチン化など様々な修飾が知られている。これらのヒストン修飾の組み合わせは、遺伝子発現を正または負に制御すると考えられている。具体的には、転写を正に制御する修飾としてH3のアセチル化やH3の4番目、36番目、79番目のリジン（それぞれH3K4、H3K36、H3K79と表記する）のメチル化が、転写を負に制御する修飾としてH3の9番目、27番目のリジン（H3K9、H3K27）のメチル化が知られている (図1)。

ゲノム上には細胞の種類に依存してDNAメチル化状態が変化する領域 (tissue-dependent and differentially methylated region, T-DMR) が多数存在する (図3) (Imamura et al., 2001; Shiota et al., 2002; Khulan et al., 2006)。T-DMRのメチル化により、細胞・組織ごとのトランスクリプトームが制御される。例えば、分化多能性細胞に特異的に発現する *Pou5f1*、*Nanog* 遺伝子プロモーター領域には、分化多能性特異的に低メチル化状態になるT-DMRが存在する (Hattori et al., 2004; Hattori et al., 2007)。これらのT-DMRのメチル化は、ヒストン修飾と協調して遺伝子発現を制御する。転写開始点上流のエンハンサー領域に存在するT-DMRが遺伝子発現を制御する例もある (Tomikawa et al., 2006; Schmidl et al., 2009)。また、胚性幹 (ES) 細胞の解析から、ES細胞の転写制御に重要な転写因子の多くがT-DMRを持ち、さらに転写因子の標的遺伝子にもT-DMRを持つものが多く存在することが示されている (Sato et al., 2010)。つまりT-DMRのメチル化は、転写因子とその標的遺伝子という二つの階層を制御するこ



とで、直接的・間接的に膨大な数の遺伝子の発現を制御し、細胞・組織特異的なトランスクリプトームの形成に寄与する（図4）。

ゲノム上に存在する多数のT-DMRにおけるDNAメチル化・非メチル化のパターンのことを、DNAメチル化プロファイルと呼ぶ（Shiota, 2004）（図3）。DNAメチル化プロファイルは各細胞・組織に固有であり、それぞれに特有の機能が反映されている（Yagi et al., 2008）。例えば、肝臓で低メチル化であるT-DMRを持つ遺伝子には、脂質代謝などの肝臓の機能に関係する遺伝子が、脳で低メチル化であるT-DMRを持つ遺伝子については、シナプスなど脳特有の機能に関連する遺伝子が有意に多く含まれた。また、前述したように、細胞・組織固有の転写因子ネットワークが各細胞・組織で機能するためにも、DNAメチル化プロファイルによる制御が行われている（Yagi et al., 2008; Sato et al., 2010）（図4）。つまり、細胞・組織特異的な遺伝子機能を発揮するために、DNAメチル化プロファイルは重要である。

哺乳類のゲノムは細胞のなかで核（あるいは細胞核）という細胞小器官に収納されている。核内には内部を空間的に区切る膜構造は存在しない。そのため物質が自由に拡散し均一な構造をとりうると考えられるが、実際には核内は高度に区画化されている。最もよく知られた例としては、核内はDNAが凝集したヘテロクロマチン領域とDNAが緩んだユークロマチン領域に分かれることが挙げられる。転写の活発なクロマチンにみられるヒストンH3K4やH3K36のメチル化修飾はユークロマチン領域と、転写が抑制されたクロマチンに見られるH3K9やH3K27のトリメチル化はヘテロクロマチン領域と一致する（Kouzarides, 2007; Wang et al., 2009）。遺伝子の転写も核内で均一に行われるのではなく、転写ファクトリーとよばれる核内空間で行われる（Iborra et al., 1996）。核あたりの転写ファクトリーの数は細胞によって異なり、数百から数千と見積もられている。この数は、発現している遺伝子の数よりも少ないことから、複数の遺伝子が一つの染色体テリトリーで転写されるという考えが提唱されている（Cook, 2002）。

実際、同一染色体上の離れた遺伝子同士あるいは異なる染色体上の遺伝子同士が空間的に近接して転写されている例が報告されている (Mistli, 2010)。また、核内には特定の機能を担う分子群が集合体として存在する領域 (機能ドメイン) が存在する (Ferrai et al., 2010)。具体的にはリボソームの合成やRNAポリメラーゼIIによる遺伝子発現が行われる核小体、mRNAのスプライシングや核外輸送に関わると考えられる核スペckル、がん抑制遺伝子産物であるPML (promyelocytic leukemia) を主体とし転写、アポトーシス、細胞周期に関わるPMLボディなどが挙げられる。さらに、個々の染色体も間期核内において混ざり合って存在するのではなく、染色体テリトリー (chromosome territory) と呼ばれる各染色体に固有の核内空間を占めていることが明らかになっている (Cremer et al., 1982)。

哺乳類ゲノムの特徴として、反復配列が約半分をしめていることが挙げられる (Lander et al., 2001; Waterston et al., 2002) (図5)。反復配列は、縦列反復 (tandem repeat) 配列とゲノム全体に散在する反復配列に分類される。後者が反復配列の大部分を占めており、その実体は転移因子 (トランスポゾン) である。転移因子は、Barbara McClintockにより発見された、ゲノム中を移動するゲノム配列である。2000年以降様々な哺乳類のゲノム解読がなされることで、哺乳類ゲノム中における転移因子の全容が明らかになってきた。転移因子は、その転移様式によって、2つの種類に分類される。1つ目は、DNAの状態で行うもので、McClintockが発見したのはこの種類であった。転移因子自身にコードされた遺伝子を利用し、転移を行っている。原核生物においては主要なゲノムの構成要素であるが、真核生物においては数が少ない。2つ目の種類は、RNA中間体を介した転移 (レトロ転移、retrotransposition) を行うもので、特にレトロトランスポゾンと呼ばれる。レトロトランスポゾンは原核細胞においては現在までに見つかっておらず、真核細胞ゲノムの特徴である。真核生物のなかでも、特に哺乳類ではゲノムの約半数近くがレトロトランスポゾンにより占められる。レトロトランス

ポソンは、その配列からさらにLTR型、非LTR型に分けられる（図6）。LTR型は配列がレトロウイルスに類似しており、両末端に長い反復配列（long terminal repeat, LTR）を有する。非LTR型は両端にLTRを持たないレトロトランスポゾンであり、哺乳類ではLINE（long interspersed nuclear element）配列とSINE（short interspersed nuclear element）配列がある。LINE配列は、長さが約7～10 kbであり、LINE配列自身のレトロ転移に必要な逆転写酵素様の遺伝子を含む。SINE配列は、長さが150～300 bpであり、SINE配列自身にはタンパク質はコードされておらず、転写産物はノンコーディングRNAである。転写されたSINE配列のRNAは、LINE配列にコードされている逆転写酵素によって転移すると考えられている（Goodier & Kazazian, 2008）。LINE配列、SINE配列はそれぞれ哺乳類ゲノムの約10%を占める。

哺乳類ゲノム中のほとんどの転移因子はメチル化修飾がなされている。転移因子はDNAメチル化修飾により転写が阻害される（Liu & Schmid, 1993; Walsh et al., 1998）ことから、生体中においてもDNAのメチル化は、転移因子というゲノム中の「寄生因子」に対する防御システムであるという考えが提唱されている（Yoder et al., 1997）。

哺乳類ゲノムにとって、DNAメチル化による細胞・組織特異な遺伝子発現制御と転移因子の抑制は、どちらも必須なものである。ただし、前者が細胞・組織依存的にメチル化・脱メチル化されるのに対し、後者は基本的にどの細胞・組織に置いてもメチル化されるという性質の違いがある。同じゲノム中において、異なる性質のDNAメチルによる制御を受けるT-DMRと転移因子がゲノムワイドにどのような関係で存在しているのかは興味深い、これまでに解析はなされていない。

本研究は、ES細胞特異的なT-DMRに焦点を当て、染色体上でのT-DMRと転移因子の関係を明らかにすることを目的とする。第一章では、遺伝子プロモーター領域のT-DMRについて、周囲の転移因子の分布を解析した。第二章においては、第一章の結果をもとに染色体全域を対象

にして、T-DMRと転移因子の関係を調べた。さらにここで分かった関係をもとに、染色体構造とT-DMRのエピジェネティック制御の関係を明らかにした。

# エピジェネティック制御の分子機構

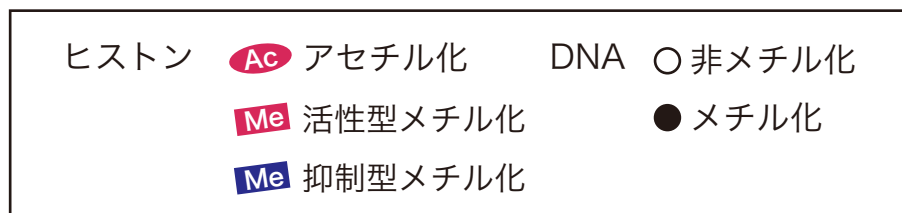
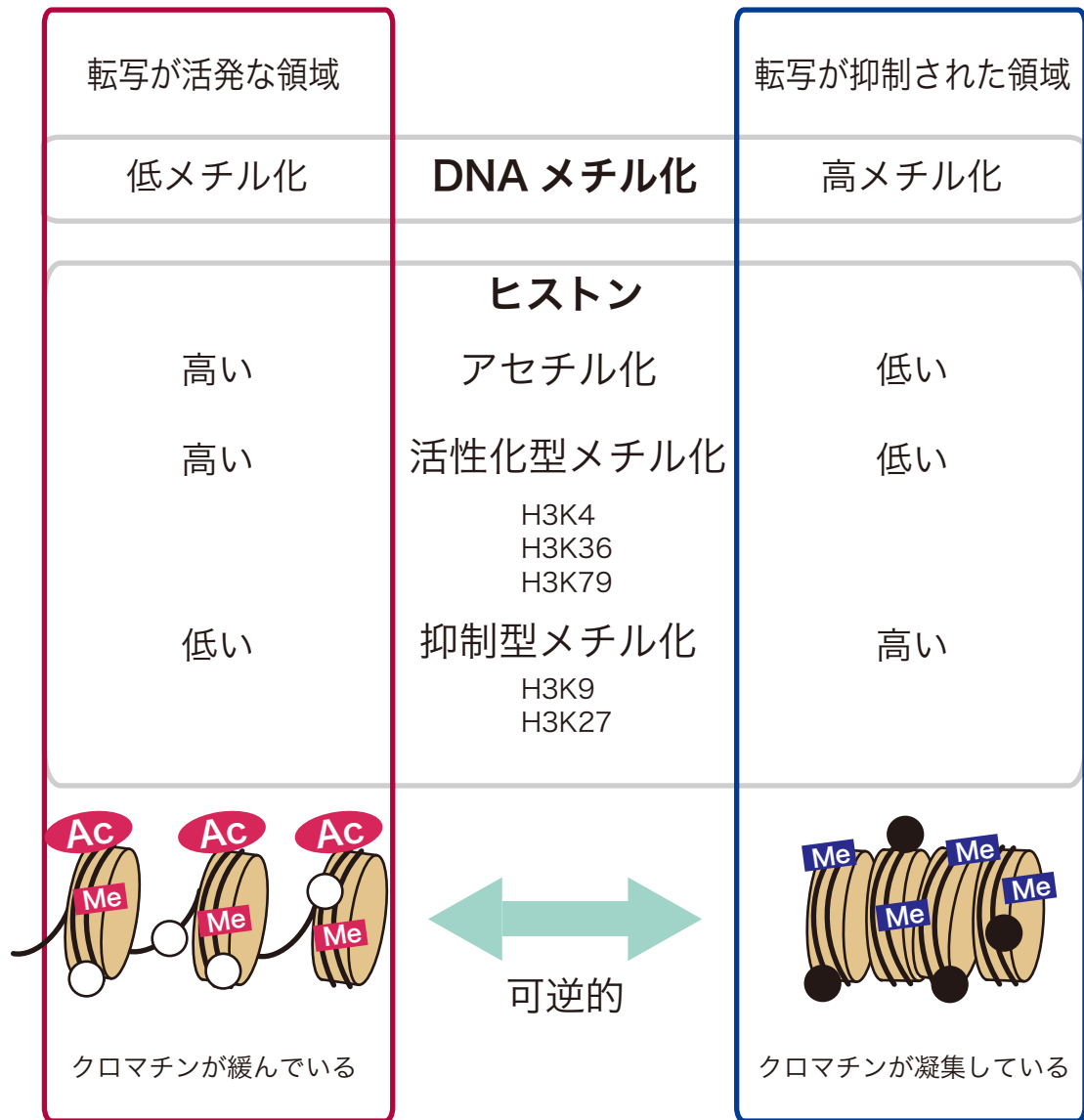


図1 DNAメチル化とヒストン修飾の模式図

哺乳類ゲノムの主要なエピジェネティック制御機構であるDNAメチル化・ヒストン修飾と転写活性との関係を模式的に示した。

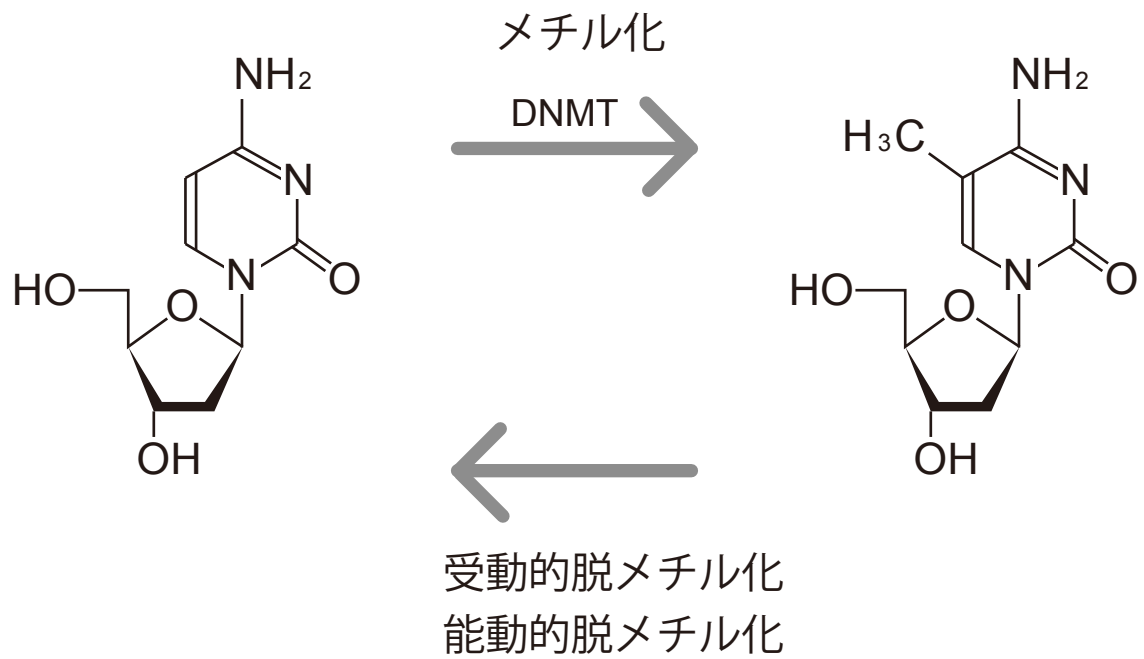


図2 DNAメチル化修飾

DNAメチル基転移酵素（DNMT）が、メチル基供与体であるSAM（S-adenosylmethionine）からシトシン塩基の5位へメチル基を転移することによりDNAのメチル化が起こる。DNAの脱メチル化機構は、DNA複製依存的に起こる受動的脱メチル化とDNA複製非依存的に起こる能動的脱メチル化に分けられる。能動的脱メチル化の分子機構は、現在のところよくわかっていない。

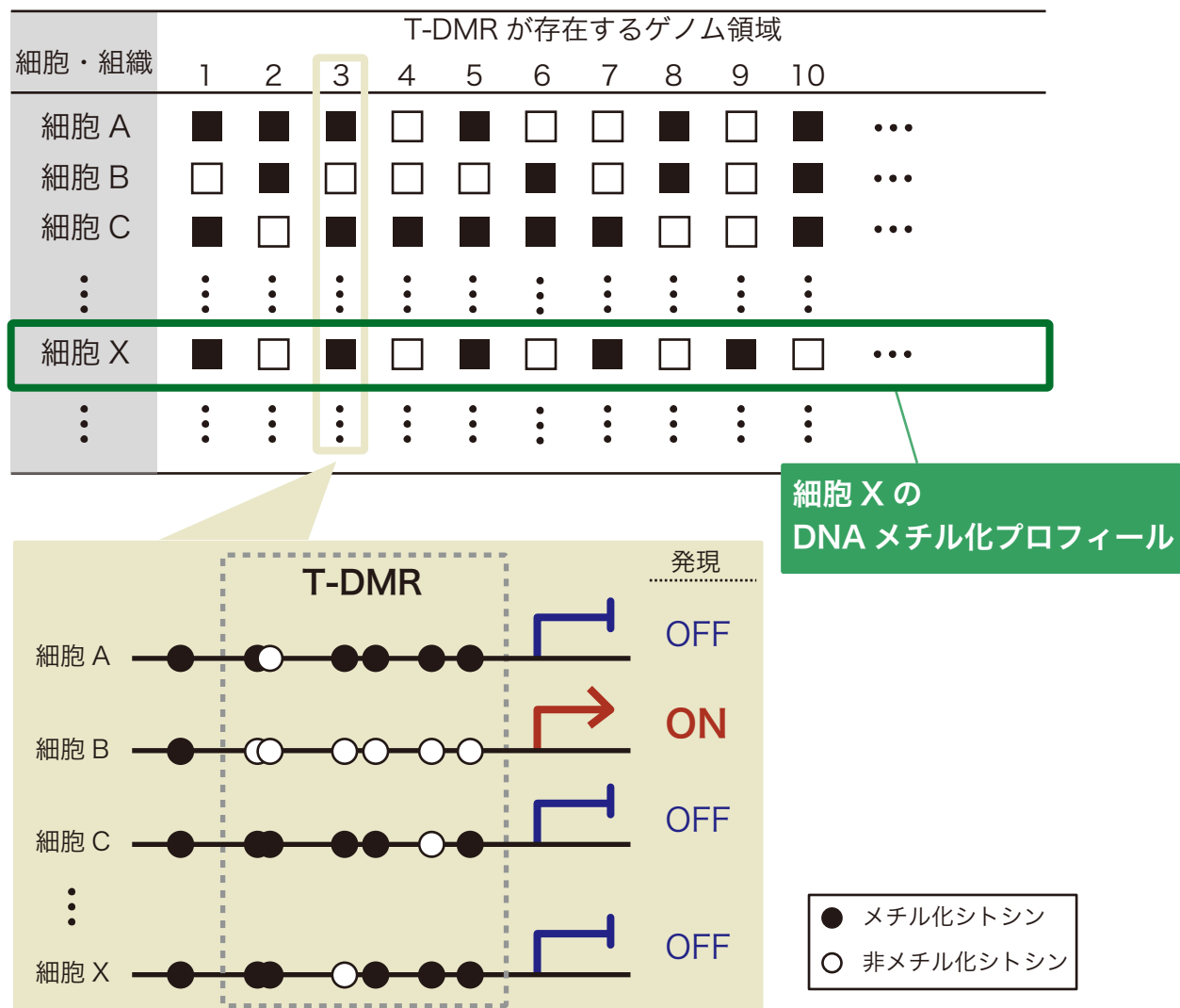


図3 DNAメチル化プロファイルの模式図

ゲノム上に存在する各T-DMRのメチル化状態の組み合わせが、細胞・組織に固有のDNAメチル化プロファイルを形成する。各T-DMRは細胞・組織特異的な遺伝子発現を制御している。従って、メチル化プロファイルは細胞・組織ごとのトランスクリプトーム形成に寄与する。

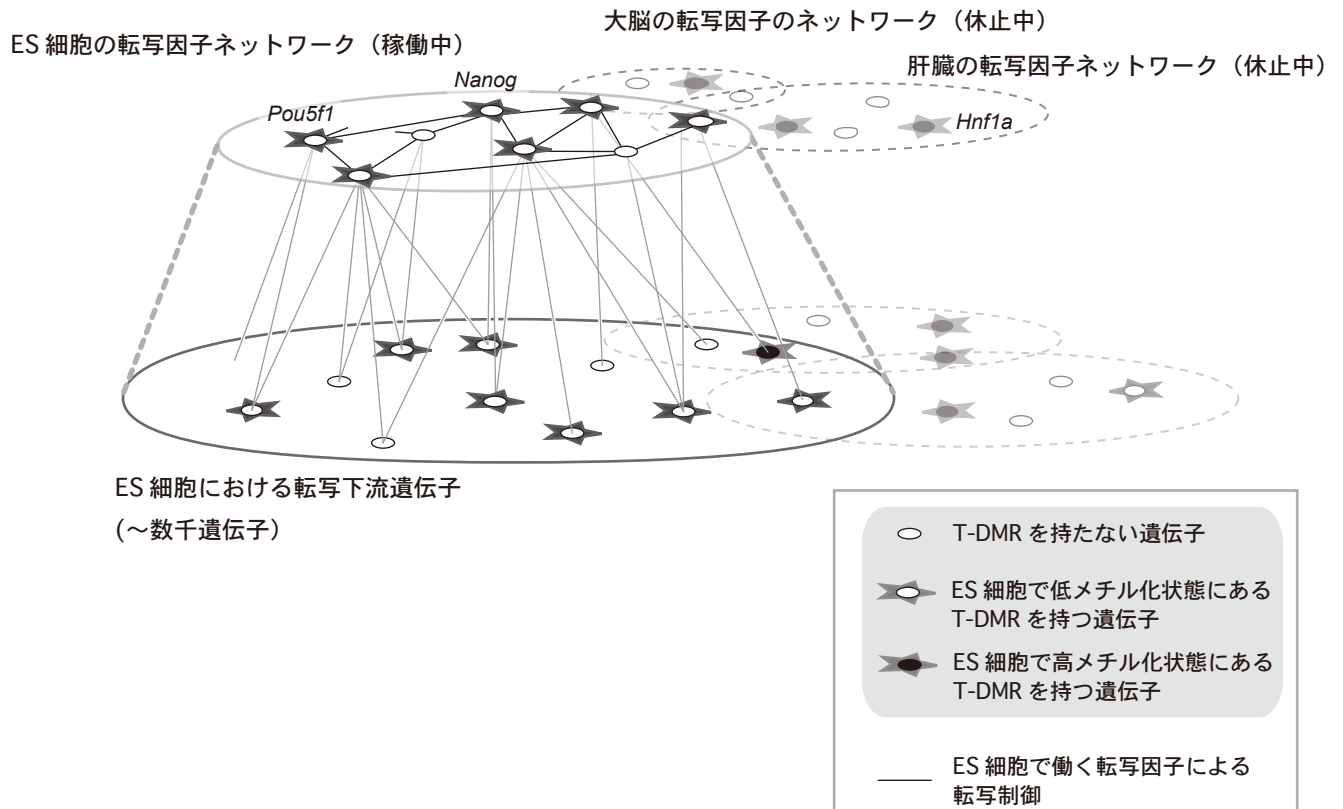


図4 T-DMRのメチル化による転写因子カスケードの階層的な制御

小楕円は遺伝子を、小楕円どうしを結ぶ直線は転写を制御する関係を示している。ES細胞を例にすると、ES細胞においては、*Pou5f1*や*Nanog*遺伝子などにより構成される転写因子ネットワークが機能し、その下流遺伝子（転写因子の標的遺伝子）を制御する。T-DMRはその両者に存在し、両階層で発現制御に関わる。結果的にメチル化により膨大な数の遺伝子が直接・間接的に制御を受け、ES細胞のトランスクリプトームが形成される。



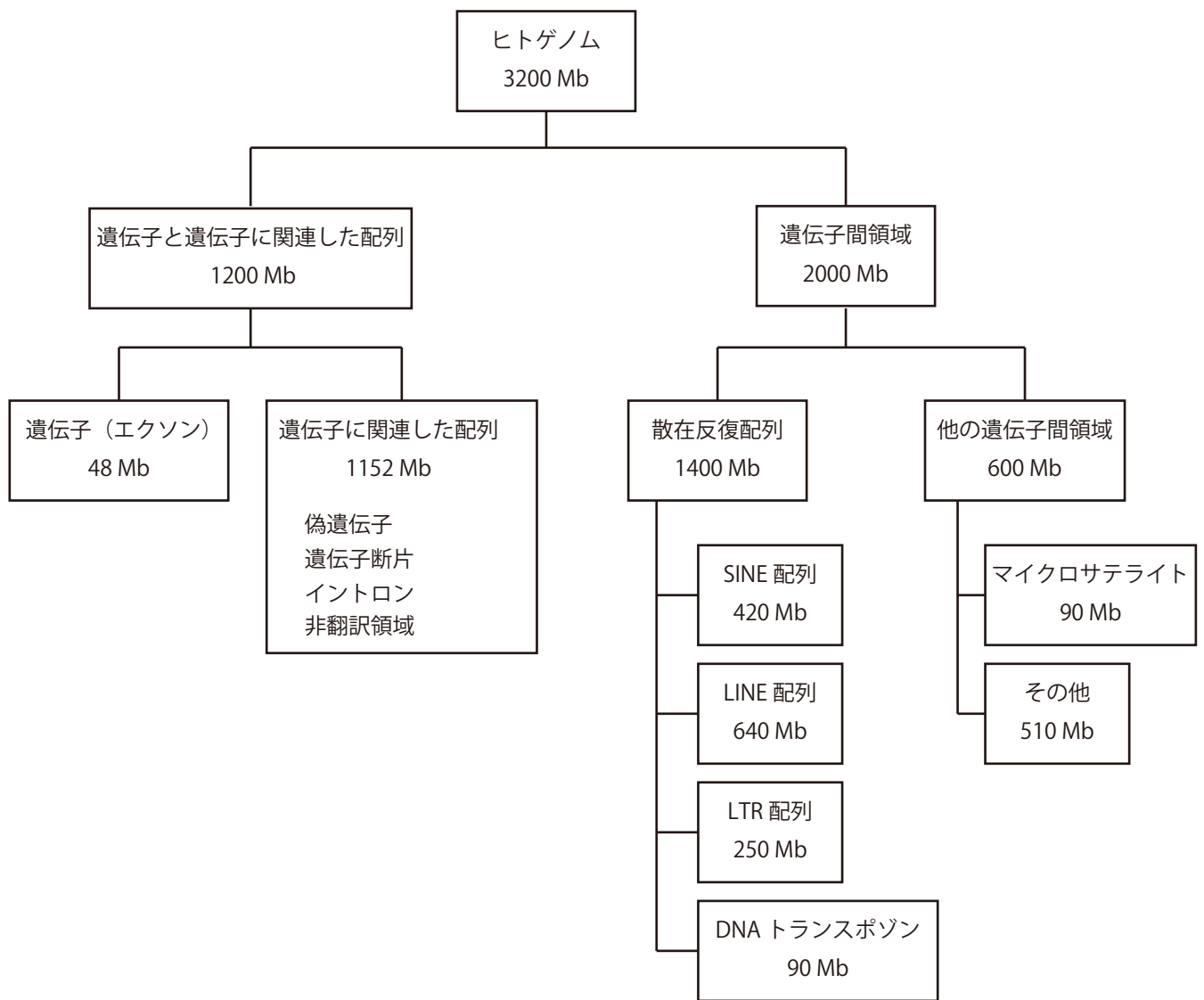


図5 ヒトゲノムの構成

哺乳類ゲノムの例として、ヒトゲノムの構成を示した。反復配列はヒトゲノムの約半分を占めている。Genomes (Brown, 2002) をもとに作成した。

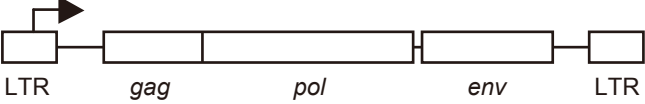
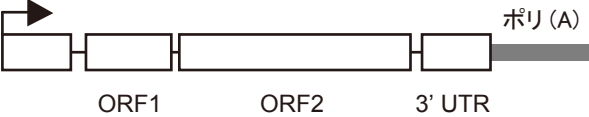

		平均的な長さ	マウスゲノム中の割合
LTR		4–10 kb	8.2%
LINE		5–8 kb	19.2%
SINE		100–500 bp	9.9%

図6 哺乳類ゲノム中のレトロトランスポゾン

哺乳類ゲノムに存在する主要な3つのタイプのレトロトランスポゾンについて、模式図を示した。

## 第一章

遺伝子領域に存在する ES 細胞特異的な組織特異的 DNA メチル化可変領域 (T-DMR) の配列解析

## 序論

内部細胞塊から樹立される胚性幹細胞（ES 細胞）は、生体を構成する全ての種類の細胞へと分化する能力である分化多能性を有する。分化多能性に必須な機能を持つ、様々な遺伝子がこれまでの研究から同定されている。例えば *Pou5f1* 遺伝子 (Niwa et al., 2000) や *Nanog* 遺伝子 (Chambers et al., 2003; Mitsui et al., 2003) は ES 細胞特異的に発現する転写因子であり、分化多能性の維持および獲得に必須な役割を持つ (Takahashi & Yamanaka, 2006; Chen et al., 2008)。

ゲノムワイドな DNA メチル化解析により、ゲノム中には多数の細胞・組織に特異的なメチル化修飾を受ける領域 (tissue-dependent and differentially methylated region, T-DMR) が存在することが明らかになった (Imamura et al., 2001; Shiota et al., 2002)。様々な遺伝子の細胞・組織特異的な遺伝子発現が、T-DMR のメチル化状態により制御されている (Hattori et al., 2004; Tomikawa et al., 2006; Hattori et al., 2007; Maeda et al., 2008)。ゲノム上に存在する各 T-DMR のメチル化状態の組み合わせを DNA メチル化プロフィールと呼ぶ (Shiota, 2004)。DNA メチル化プロフィールは ES 細胞や体細胞を含め、細胞ごとに固有である (Shiota et al., 2002; Sakamoto et al., 2007)。*Pou5f1* 遺伝子や *Nanog* 遺伝子は、ES 細胞特異的に低メチル化状態にある T-DMR をプロモーター領域に有する。これらの T-DMR のメチル化は、ヒストン修飾とともに、ES 細胞特異的な発現制御に関わる (Hattori et al., 2004; Hattori et al., 2007)。ES 細胞特異的な非メチル化状態は、ES-cell-associated transcript (ECAT) と呼ばれる ES 細胞に特異的な発現を示す遺伝子群にも認められる (Mitsui et al., 2003; Imamura et al., 2006)。さらに、マイクロアレイを用いた網羅的なメチル化プロフィールの解析から、ES 細胞特異的なメチル化制御を受ける T-DMR

が数千存在することが明らかになっている (Sato et al., 2010)。未分化状態の維持および分化過程において、DNA メチル化プロファイルが適切に維持・形成されることが ES 細胞の分化多能性に重要である。

DNA メチル化修飾は、細胞・組織特異的な遺伝子発現制御に加え、転移因子（トランスポゾン）の抑制にも働く。いくつかの哺乳類についてゲノム解読プロジェクトが終了し、ゲノムの約半数を占める反復配列は、大部分が転移因子に由来することが明らかになった (Waterston et al., 2002)。SINE 配列、LINE 配列、LTR 配列などの転移因子は、転写産物（RNA 中間体）が逆転写されるプロセスによりゲノム中で増幅する。転移因子の転写は DNA メチル化修飾により阻害される (Liu & Schmid, 1993; Walsh et al., 1998)。生体中においても DNA メチル化修飾はこれらの因子の転写制御に関わると考えられており、実際哺乳類ゲノム中においてはほとんどの転移因子がメチル化修飾を受けている。このことから、DNA のメチル化は、転移因子というゲノム中の「寄生因子」の増殖を抑制する防御システムであるという考えが提唱されている (Yoder et al., 1997)。

染色体上において、T-DMR と転移因子がゲノムワイドにどのような関係で存在しているのかは興味深い、これまでに解析はなされていない。第一章では、ゲノムワイドなメチル化解析法である D-REAM (Yagi et al., 2008) により ES 細胞特異的な T-DMR を有する遺伝子を同定し、それらの T-DMR を有する遺伝子周辺に存在する反復配列の分布状況を解析した。さらに、ES 細胞特異的な T-DMR を有する遺伝子周辺に存在する反復配列が持つ、配列の性質を明らかにした。

## 材料および方法

### 試薬

試薬はとくに断りがない限り和光純薬 (Osaka) より購入した。

### 実験動物

本実験に用いた C57BL/6NCrj マウスは日本チャールズリバー (Yokohama) より購入した。マウスは全て室温 22–25°C、湿度 40–60%、照明が明期 14 時間、暗期 10 時間に制御された環境下で飼育され、食餌および水を自由に摂取させた。膣栓が確認された日の正午を胎生 0.5 日とした。DNA 抽出用に成体マウス肝臓を摘出した。PBS (–) (1.37 M NaCl、27 mM KCl、100 mM  $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ 、18 mM  $\text{KH}_2\text{HPO}_4$ ) で洗浄後、液体窒素で凍結し、–80°C で保存した。

### 培養細胞

#### A) 培地

本実験に用いた培地の組成は以下に示した通りである。

#### EMFI 培地

DMEM pH 7.4 (Gibco BRL、Rockville) に、10% FBS (JRH、Lenexa)、100  $\mu\text{M}$  mercaptoethanol (Gibco BRL)、50  $\mu\text{g/ml}$  penicillin (Gibco BRL)、50  $\mu\text{g/ml}$  streptomycin (Gibco BRL) を添加した。

#### ES 細胞用培地

DMEM pH 7.4 (Gibco BRL) に、15% FBS (Biowet、Nuail)、100 U/ml penicillin (Gibco

BRL)、100  $\mu$ g/ml streptomycin (Gibco BRL)、100  $\mu$ M/ml MEM nonessential acids (Gibco BRL)、1 mM/ml sodium pyruvate (Gibco BRL)、100  $\mu$ M/ml 2-mercaptoethanol (Gibco BRL)、および 2 mM/ml L-glutamine (Gibco BRL) を添加した。

## B) Feeder 細胞

### Primary mouse embryonic fibroblast (EMFI) 細胞の調整

胎生 15.5 日目のマウス胎仔から頭部および全ての臓器、骨を取り除き、PBS (-) で洗浄した後に細断した。再び PBS (-) で洗浄後、175 x g で 3 分間遠心して沈殿物を回収した。0.05% Trypsin/1 mM EDTA を加えた PBS (-) 中で、室温において 30 分間穏やかに攪拌した後、再度トリプシンを加えて回転震盪した。メッシュを用いてろ過した後、175 x g、室温で 3 分間遠心して回収した細胞を EMFI 培地に懸濁し、新しいディッシュに移した。翌日培地を交換し、セミコンフルエントになるまで 3 日間培養した後、細胞を回収した。使用するまで液体窒素中に保存した。

### MitomycinC treated EMFI (MMC-EMFI) 細胞の調整

EMFI 細胞を EMFI 培地でセミコンフルエントまで培養した後、培地に 1 mg/ml Mitomycin C を加えて 37°C で 2 時間処理した。回収後液体窒素中に保存した。

## C) ES 細胞

### ES 細胞の培養

C57BL/6 マウス胚由来の ES 細胞 MS12 株は、明治製菓（現 京都大学）の末森博士より分与していただいた。ES 細胞は EMFI 細胞をフィーダー細胞として培養された。培地は ES 細胞用培地用い、細胞を未分化状態に保つため、1000 U/ml の白血病抑制因子 (LIF; leukemia

inhibitory factor、商品名：ESGR0、Chemicon Int.、Temecula）を加えた。培地は一日おきに交換した。

#### 未分化 ES 細胞の調整

フィーダー細胞を除くため、0.05% trypsin/1 mM EDTA で処理した後、ゼラチンコートされていない新しいディッシュにまき、37°C で 30 分培養した。その後上清を新たなゼラチンコートされていないディッシュに移し、30 分後 ES 細胞を含む上清を回収し、-80°C で保存した。

#### DNA メチル化解析

##### A) ゲノム DNA の抽出

サンプルを lysis buffer [10 mM Tris-HCl (pH8.0)、5 mM EDTA、200 mM NaCl、0.2% SDS、200 µg/ml proteinase K (Merck、Darmstadt) ] で懸濁し、55°C で 20 分間インキュベートした。次に等量の phenol/chloroform/isoamyl alcohol (PCI) (50:49:1) で抽出し、500 µg/ml RNase (Roche、Basel) を 0.5 µl 加え、37°C で 1 時間インキュベートした。再度 PCI で抽出した後、ゲノム DNA をエタノールで沈殿させ、TE (pH 8.0) に溶解した。使用するまで 4°C で保存した。

##### B) ゲノム DNA の制限酵素処理

抽出したゲノム DNA に *Pst*I (Takara, Kyoto) 、*Eco*RI (Takara)、または *Xba*I (Takara) を 50 U 加え、37°C で 16 時間反応させた。その後、制限酵素を除去するために、等量の PCI を加え、5 分間激しく振って混和した。20,400 x g、10 分間、4°C で遠心後、上清を新しいチューブに移し、再び等量の PCI を加え、振とう混和した。20,400 x g、10 分間、4°C で遠心後、上清



を回収し、3 M Na-acetate pH 5.2 を全量の 1/10 倍量加えた。2 倍量の 100% ethanol を加え転倒混和し、-80°C に 10 分間静置した。20,400 x g、15 分間、4°C で遠心した後、上清を除き、70% ethanol で洗浄し、ペレットを 31  $\mu$ l の滅菌蒸留水に溶解した。このうちの 1  $\mu$ l を用いて吸光光度計により濃度測定を行い、残りを -20°C で保存した。

#### C) バイサルファイト反応

制限酵素処理したゲノム DNA が 3-10  $\mu$ g ずつ含まれるように、それぞれのサンプルを 32.3  $\mu$ l まで滅菌蒸留水でメスアップした。これに 10N NaOH を 1  $\mu$ l 加え、37°C で 15 分間変性処理した。その後、10 mM hydroquinone を 14.4  $\mu$ l (終濃度 2.0 M)、2.5 mM sodium metabisulfite (pH 5.0) を 229.7  $\mu$ l (終濃度 0.5 M)、滅菌蒸留水を 9.8  $\mu$ l 加えて全体を 287  $\mu$ l とした。この溶液を遮光し、55°C で 16 時間反応させた。反応後、Wizard DNA Clean-Up System (Promega、Madison) を用いて DNA を精製した。精製後、10 N NaOH 3  $\mu$ l を加え 37°C で 15 分間インキュベートし、バイサルファイト反応を停止させた。等量 (99.6  $\mu$ l) の 6 M  $\text{NH}_4\text{OAc}$  (pH 7.0) を加えて溶液を中和し、1  $\mu$ l の ethachinmate (NIPPON GENE、Tokyo)、2 倍量の 100% ethanol を加えて DNA を沈殿し、70% ethanol で洗浄した。得られたペレットを 20  $\mu$ l の滅菌蒸留水に溶解した。さらにそれぞれのサンプルのゲノム DNA 溶液の濃度が 2-10 ng/ $\mu$ l になるように希釈した。

#### D) バイサルファイト PCR

バイサルファイト反応後、Polymerase Chain Reaction (PCR) を行なった。Immolase Taq DNA polymerase (BIOLINE、Randolph) を用いて、95°C 10 分; 94°C 30 秒、60°C 30 秒、72°C 1 分  $\times$  43 サイクル; 72°C 10 分 の条件で増幅した。プライマーの配列は表 1-3 にまとめて記載した。

PCR 産物はアガロースゲル中で電気泳動を行い、ethidium bromide 染色によって可視化した。得られた DNA 断片は Wizard SV Gel and PCR Clean-Up System (Promega) を用いて精製した。

#### E) バイサルファイト Restriction Mapping

バイサルファイト PCR 産物では、バイサルファイト反応を行ったことにより非メチル化シトシンがチミンに置換されている。そのため制限酵素の認識配列に CpG 配列が含まれる場合、この置換を利用することで切断された PCR 産物の割合から制限酵素の認識配列に含まれるメチル化シトシンの割合を解析することが出来る。バイサルファイト PCR 産物を *TaqI* (Takara) により制限酵素処理した。切断された PCR 産物の割合は、マイクロチップ電気泳動装置 Multina (Shimadzu, Kyoto) を用いて定量した。

#### F) バイサルファイトシーケンス

精製したバイサルファイト PCR 産物を pGEM-T Easy Vector (Promega) にクローニングし、塩基配列を決定した。

#### D-REAM

5  $\mu$ g ゲノム DNA にメチル化感受性酵素である *HpyCH4IV* (New England BioLabs, Ipswich) で処理し、一晚 37 °C で反応させた。アガロースゲル中で電気泳動を行い、反応が行われたことを確かめた。切断された DNA をエタノールで沈殿させ PCI で抽出し、TE (pH 8.0) に溶解した。50 ng の DNA に、T4 DNA リガーゼ (New England BioLabs) によって R アダプターペア (表 1-3) を連結した。クレノウフラグメント (Takara) 処理した後、*TaqI* (Takara) で消化

し、Micfospin S-300 HR column (GE healthcare、Buckinghamshire) により精製した。DNA に N アダプターペア (表 1-3) を連結し、Wizard SV Gel and PCR Clean-Up System (Promega) を用いて精製した。dUTP 存在下で、Immolase Taq DNA polymerase (BIOLINE) を用いて、プライマーセット R18 と N18 (表 1-3) により PCR 反応を行った。PCR サイクルは、95°C 7 分; 94°C 30 秒、62°C 30 秒、72°C 2 分 × 20 サイクル; 72°C 2 分 とした。10 µg の増幅産物をマイクロアレイ解析に用いた。

マイクロアレイ解析は GeneChip System (Affimetrix, Santa Clara) を用い、全ての手順を Affimetrix chromatin immunoprecipitation assay protocol に基づいて行った。簡潔に説明すると、GeneChip WT Double-Strand DNA Terminal Labeling Kit (Affimetrix) によりサンプル DNA をビオチン標識し、GeneChip Mouse Promoter 1.0R Array とハイブリダイゼーションさせた。反応後のアレイは Fluidics Station 450 (Affimetrix) でプロトコール FS490\_0001 において洗浄、抗体反応、シグナル増幅反応を行った。抗体液、染色液は GeneChip Hybridization, Wash, and Stain Kit (Affimetrix) の Stain Cocktail を用いた。Affimetrix GeneChip Scanner 3000 (Affimetrix) でシグナルを読み取り、GeneChip Operating Software (GCOS ver1.4; Affimetrix) によって、個々のプローブにおけるシグナル強度を記述した CEL ファイルを得た。

#### D-REAM データ解析

D-REAM のデータ解析は以前の報告に従った (Yagi et al., 2008)。マイクロアレイにより得られた“.cel”ファイルを MAT プログラム (Johnson et al., 2006) を用いて処理し、ES 細胞と肝臓の間で有意にシグナル強度に差がある領域を検出し ( $p\text{-value} < 10^{-7}$ )、BED ファイルとして得た。MAT による計算における window サイズは 300 bp に設定した。

## T-DMRtag の同定

本研究では、D-REAM によって同定したメチル化状態に差がある *HpyCH4IV* 認識配列部位を “T-DMRtag” と呼ぶことにする。肝臓に比べ ES 細胞の方が低メチル化状態にあるものを ES hypo T-DMRtag、高メチル化状態にあるものを ES hyper T-DMRtag とした。T-DMRtag は、Galaxy (Giardine et al., 2005) <<http://g2.bx.psu.edu>> を利用して、BED ファイルの領域情報と *HpyCH4IV* 認識部位の位置情報を照合することにより得た。

## データの可視化

T-DMRtag と遺伝子の位置、およびゲノム上の反復配列の割合は、Integrated Genome Browser <[http://www.affymetrix.com/support/developer/tools/download\\_igb.affx](http://www.affymetrix.com/support/developer/tools/download_igb.affx)> または UCSC ゲノムブラウザー <<http://genome.ucsc.edu>> を用いて可視化した。

## ゲノム解析に使用したデータ

ゲノム配列情報、CpG アイランド情報および Gap 情報は UCSC ゲノムブラウザー (mouse, mm8; NCBI build 36) より取得した。反復配列情報は UCSC ゲノムブラウザーの RepeatMasker track より取得した。RepeatMasker track は RepBase library <<http://www.repeatmasker.org>> の反復配列データベースに基づいて計算されている。遺伝子情報として、Ensembl Gene ID (Ensembl 46, August 2007) <<http://aug2007.archive.ensembl.org/index.html>> を使用した。

## ゲノム配列に関する計算

ゲノム上の反復配列の割合は Galaxy および自作した Perl スクリプトにより解析した。CpG 配列の数、GC 含有量、CpG 配列の観測値／期待値 (Obs/Exp)、1 kb あたりの CpG 配列頻度

(CpG/kb) および GC 含有量 (GC/kb) は自作した Perl スクリプトにより計算した。全ての配列計算において、10%より大きい割合が Gap により占められている領域は解析から除外した。Gap が 10%以下の領域については、Gap をのぞいた部分について計算を行った。

## 統計解析

統計解析には統計解析ソフト R <<http://www.R-project.org>> を用いた。

## 結果

### ES 細胞において低メチル化状態にある T-DMR を持つ遺伝子周辺領域に認められたゲノム

#### 配列の特徴

ES 細胞と体細胞でメチル化状態の異なる T-DMR を同定するために、ES 細胞と肝臓を対象とし、マウスプロモーターアレイを用いた D-REAM 解析を行った。メチル化感受性酵素としては *HpyCH4IV* を使用し、ES 細胞と肝臓でメチル化状態の異なる *HpyCH4IV* 認識部位 (Hpy サイト) を T-DMRtag として検出した。常染色体上の 27194 遺伝子のうち、1923 遺伝子が ES 細胞において低メチル化状態にある T-DMRtag (ES hypo T-DMRtag) を有し、717 遺伝子が ES 細胞において高メチル化状態にある T-DMRtag (ES hyper T-DMRtag) を有した。以下、それぞれの遺伝子を ES hypo 遺伝子、ES hyper 遺伝子と呼ぶ。

遺伝子の転写開始点-150 kb から+150 kb に渡り、5 kb おきに 10 kb の領域に占める SINE 配列、LINE 配列、LTR 配列の割合を解析した。遺伝子周辺に各反復配列の占める割合 (図 1-1A 上段)、および全常染色体遺伝子周辺に対する ES hypo 遺伝子・ES hyper 遺伝子周辺の各反復配列の占める割合の相対値 (図 1-1A 下段) を図 1-1 に示した。その結果 ES hypo 遺伝子の周辺領域には、SINE 配列と LINE 配列の占める割合に偏りがあることが明らかになった。全常染色体遺伝子の周辺領域に比べ ES hypo 遺伝子の周辺領域では、SINE 配列の占める割合の平均値は約 1.4 倍であり、LINE 配列の占める割合の平均値は約半分であった。興味深いことに、ES hypo 遺伝子周辺領域では、SINE 配列は遺伝子転写開始点に近いほど高い割合を占め、転写開始点を中心とする分布を示した (図 1-1A 上段)。次に、転写開始点周囲 100 kb 領域に占める反復配列の割合を計算した (図 1-1B)。反復配列ではない領域 (非反復配列) の割合は、全常遺伝子、ES hypo 遺伝子、ES hyper 遺伝子のいずれの周辺領域においても共通して 60%以上を占めた。一方

で、SINE 配列と LINE 配列の割合は、ES hypo 遺伝子周辺領域でのみ偏りが認められた。

以上で明らかになった偏った反復配列の分布は、ウィルコクソン順位和検定 (Wilcoxon rank sum test) を用いた統計解析によって有意な傾向であることが確認された。ES hypo 遺伝子周辺 100 kb 領域では、SINE 配列、LINE 配列、LTR 配列の分布に有意な偏りが認められた (図 1-1C)。各常染色体ごとに解析したところ、SINE 配列と LINE 配列に関してはほぼ全ての常染色体上で有意な偏りが認められたのに対し、LTR 配列については 4 つの常染色体にのみ偏りが認められた (表 1-1、図 1-1D)。一方、ES hyper 遺伝子周辺領域については、SINE 配列が分布に有意な差を示したものの (図 1-1C、表 1-1)、染色体ごとに偏りを調べると有意な差が認められたのは 1 つの染色体のみであった (図 1-1D)。これらの結果は、ES hypo 遺伝子周辺において SINE 配列と LINE 配列がゲノムワイドに偏った分布を示すことを意味している。

### **SINE 配列が多く LINE 配列が少ない領域への ES hypo 遺伝子の偏り**

次に SINE 配列と LINE 配列に焦点をあて、ES hypo 遺伝子周辺における分布を解析した (図 1-2A)。ES hypo 遺伝子のうち、SINE 配列が多く LINE 配列が少ない領域に存在する遺伝子の割合は 69%であり、これは常染色体遺伝子の割合と比べ 29%高い。一方で、SINE 配列が少なく LINE 配列が多い領域に存在する遺伝子の割合は 12%であり、これは常染色体遺伝子の割合と比べ 28%低い。一方、ES hyper 遺伝子は常染色体遺伝子と同様の割合であった。SINE 配列が多く LINE 配列が少ない領域に ES hypo 遺伝子が存在する傾向は、染色体ごとに調べても認められた。典型的な例として、第 10 染色体における散布図を図 1-2B に示した。また、ES 細胞において特異的に発現することが知られている ECAT 遺伝子群 (Mitsui et al., 2003) について解析したところ、ECAT 遺伝子群の中で ES hypo T-DMR を有するものは SINE 配列が多く LINE 配列が少ない領域に偏っていた (図 1-3)。さらに、ES 細胞特異的な転写因子の標的となっている遺伝

子を解析した（図 1-4）。標的遺伝子は全体的に SINE 配列が多く LINE 配列が少ない領域に存在する傾向があったが、そのなかでも ES hypo T-DMR を有するものはさらに顕著に SINE 配列の多い領域に存在した。各ゲノム領域に着目した解析により同定された、SINE 配列が多く LINE 配列が少ない領域に存在する ES hypo 遺伝子の具体例を図 1-2C に示した。第 6、10 染色体の 2 つのゲノム座位では、1 Mb 以上にわたって SINE 配列が多く LINE 配列が少ない領域が存在し、ES hypo 遺伝子はその領域内に偏って存在していた。以上の結果より、SINE 配列が多く LINE 配列が少ないことは、ES hypo 遺伝子が存在しているゲノム領域が持つ配列の特徴であることが分かった。

#### **ES hypo 遺伝子周辺領域の偏った CpG 配列頻度**

CpG 配列の数、GC 含有量、CpG 配列の観測値/期待値（Obs/Exp）を、転写開始点前後 100 kb 領域について解析した。これらの値は ES hypo 遺伝子周囲において全常染色体遺伝子に比べ有意に高かった（表 1-2、図 1-5 A、B）。また、この傾向は染色体ごとに解析した場合でも、ほぼ全ての常染色体において認められた（図 1-5C）。一方、ES hyper 遺伝子周囲においては、このようなゲノムワイドな傾向は認められなかった（表 1-2、図 1-5A、B、C）。常染色体遺伝子と ES hypo 遺伝子の違いをより詳しく解析したところ、ES hypo 遺伝子周辺では CpG 配列の数、GC 含有量、Obs/Exp が少なくとも -150 kb から 150 kb にわたって値が高いだけでなく、- 20 kb から 20 kb で顕著に高い値を示すことが分かった（図 1-5A）。このことは、プロモーター領域（- 7kb から 3 kb）に CpG アイランドを有する割合が ES hyper 遺伝子と常染色体遺伝子は 50%以下であるのに対し、ES hypo 遺伝子は 84%であることと一致する（図 1-5D）。興味深いことに、CpG アイランドの有無で ES hypo 遺伝子分類した場合、どちらのグループも -150 kb から 150 kb の領域での CpG 配列の数、GC 含有量、Obs/Exp の平均値が常染色体遺伝子より高かった（図 1-5A 下段）。これらの結果から、ES hypo 遺伝子周辺のゲノム配列には、300 kb 以上にわたって CpG



配列頻度、GC 含有量が有為に高い傾向があることが示された。

次に、これらの傾向が全ての種類の配列（SINE 配列、LINE 配列、LTR 配列、その他の反復配列、非反復配列）に認められるのかを調べるために、各種類の配列について 1 kb あたりの CpG 配列の数および GC 含有量を解析した（CpG/kb および GC/kb と呼ぶ）（図 1-5E）。SINE 配列とその他の反復配列の GC/kb を除き、どの種類の配列においても CpG/kb、GC/kb は ES hypo 遺伝子周辺で常染色体遺伝子周辺に比べ高かった。さらに、ES hypo 遺伝子周辺領域における SINE 配列の CpG/kb は、常染色体遺伝子周辺領域における値よりも有意に高い傾向があった（図 1-5F）。このことは ES hypo 遺伝子が CpG 配列を多く含む SINE 配列に囲まれていることを示している。以上の結果より、ES hypo 遺伝子周辺における高い CpG 配列頻度は、ユニーク配列および SINE 配列を含めた反復配列の両方に認められる配列の特徴であることが明らかになった。

#### **ES hypo 遺伝子周辺領域には特定の SINE サブファミリーが多い**

マウスゲノム中に存在する SINE 配列は、B1（Alu と呼ばれる）、B2、B4、ID、MIR の 5 つのサブファミリーに分類される。B1、B2、B4、ID はげっ歯類に特有な SINE サブファミリーであり、MIR は哺乳類全体に認められる SINE サブファミリーである。サブファミリーごとの解析をしたところ、B1、B2、B4、ID は常染色体遺伝子に比べ ES hypo 遺伝子の周囲に占める割合が有意に高かった（ $P < 2.2 \times 10^{-16}$ ）（表 1-1、図 1-6A）。この傾向はほぼ全ての常染色体において認められたことから（図 1-6B）、B1、B2、B4、ID はゲノムワイドに ES hypo 遺伝子周囲に多く存在することが示された。一方 MIR は、常染色体全体では ES hypo 遺伝子周囲において常染色体遺伝子周囲に比べ有意に多く存在していた（ $P = 2.0 \times 10^{-6}$ ）（表 1-1、図 1-6A）が、ES hypo 遺伝子と常染色体遺伝子の周囲に占める割合の中央値の差は他のサブファミリーに比べると小さかった。また染色体ごとに解析すると、MIR が ES hypo 遺伝子周囲に有意に多い傾向が

認められたのは1つの染色体のみであった（図 1-6B）。ES hyper 遺伝子周囲では、各 SINE 配列サブファミリーの分布には弱い偏りしかないもしくは偏りが認められなかった（表 1-1、図 1-6）。以上より、SINE 配列のなかでも特定のサブファミリーが、ゲノムワイドに ES hypo 遺伝子周辺に有意に多く存在することが分かった。

### ES hypo 遺伝子周辺領域に存在する SINE 配列のメチル化解析

ES 細胞における SINE 配列のメチル化状態を解析するために、まず SINE 配列全体のメチル化状態を調べた。COBRA 法により、SINE B1 のコンセンサス配列に存在する *TaqI* 認識部位のメチル化状態を解析した（図 1-7A）。*TaqI* により切断されなかった PCR 産物の割合は ES 細胞と肝臓の間で有意な差はなかった（ $P > 0.05$ ）。一方、胚体外幹細胞（TS 細胞）、*Dnmt1* 欠損 ES 細胞、*Dnmt3a/3b* 両欠損 ES 細胞は野生型の ES 細胞に比べ、切断されなかった割合が有意に高かった。この結果、これらの細胞のゲノムが全体的に低メチル化状態にあることと一致している（Chapman et al., 1984; Li et al., 1992; Okano et al., 1999）。以上より、ES 細胞と肝臓では全体的な SINE B1 のメチル化状態は同様であることが示唆された。

次に、ES hypo 遺伝子の1つである *Dppa3* 遺伝子周辺に存在する SINE 配列のメチル化状態を解析した。*Dppa3* 遺伝子は分化多能性細胞および始原生殖細胞のマーカー遺伝子であり（Sato et al., 2002）、SINE 配列が多く LINE 配列が少ない領域に存在する（図 1-2C）。*Dppa3* 遺伝子が ES 細胞において肝臓よりも低メチル化な T-DMR を有することは、バイサルファイトシーケンスにより確認された（図 1-7B）。転写開始点上流（領域 I、II）および遺伝子内領域（領域 III、IV）の SINE 配列のメチル化解析を行ったところ、5つの SINE 配列が、ES 細胞において肝臓よりも 30%以上低メチル化状態にあることが分かった（図 1-7C）。以上の結果より、SINE 配列は ES 細胞と体細胞で共通して全体的に高メチル化状態にあるものの、*Dppa3* 遺伝子領域にお

いては一部の SINE 配列は ES 細胞で特異的な低メチル化を示す T-DMR であることが分かった。

## 考察

本章では、ES 細胞で低メチル化状態にある T-DMR (ES hypo T-DMR) を有する遺伝子 (ES hypo 遺伝子) 周辺には、ゲノム配列に偏りがあることを明らかにした。これまでに、エピジェネティック制御を受けるゲノム領域に反復配列の分布に偏りがみられる例がいくつか報告されている。例えば SINE 配列は、ヒトにおいてインプリント遺伝子領域、およびヒト・マウスにおいて単一对立遺伝子性発現する遺伝子周辺に数が少ないことが知られている (Greally, 2002; Allen et al., 2003)。また、ES 細胞でヒストン H3 の 27 番目のリジンがメチル化修飾を受けている領域では、SINE 配列を含む反復配列の密度が低い (Bernstein et al., 2006)。これらに対し、本章の研究は、反復配列のなかでも SINE 配列にのみ認められる傾向として、ES hypo 遺伝子周辺領域において密度が高いことを明らかにした。ECAT に着目した解析からも、ECAT 遺伝子のなかで ES hypo T-DMR を持つものは SINE 配列密度の高い領域に存在することが示された。さらに、ES 細胞に必須な転写因子の標的遺伝子のなかで、ES hypo T-DMR を有する遺伝子は有意に周囲の SINE 配列密度が高かった。以上の結果より、周囲の SINE 配列密度が高いことは ES hypo 遺伝子に認められる特有の性質であることが分かった。

これまでに、ゲノムプロットニング解析により体細胞においては Alu 配列が完全にメチル化されていることが示されていた (Hellmann-Blumberg et al., 1993)。ヒトの脳において Alu 配列を対象としたハイスループットなメチル化解析によっても、SINE 配列がメチル化修飾を受けていることが明らかにされている (Xie et al., 2009)。さらにバイサルファイトシーケンスによる大規模なメチル化解析により、ES 細胞、神経細胞、初代培養細胞において、ほとんどの SINE 配列がメチル化修飾をうけていることが示された (Meissner et al., 2008)。これらの報告は哺乳類ゲノム中では反復配列が高度にメチル化されているという考えに一致する。本

研究においても、SINE B1 配列全体は ES 細胞においても肝臓においてもよくメチル化されていた。しかし ES hypo 遺伝子の 1 つである *Dppa3* 遺伝子に焦点を当てた解析から、ES 細胞において低メチル化、肝臓において高メチル化を示す T-DMR である SINE 配列が存在することが分かった。このことは、他の ES hypo 遺伝子周辺に存在する SINE 配列のなかにも、ES 細胞において低メチル化状態を示す T-DMR であるものが存在する可能性を示唆している。これらの SINE 配列は T-DMR を構成するユニーク配列と同じように、ES 細胞の未分化状態においてはメチル化修飾から守られ、分化過程においてメチル化修飾がなされる必要がある。従って、ES hypo 遺伝子領域における SINE 配列のエピジェネティック状態の制御が、ES 細胞の未分化状態および分化過程における DNA メチル化プロファイルの形成および、分化多能性に重要である可能性が考えられる。

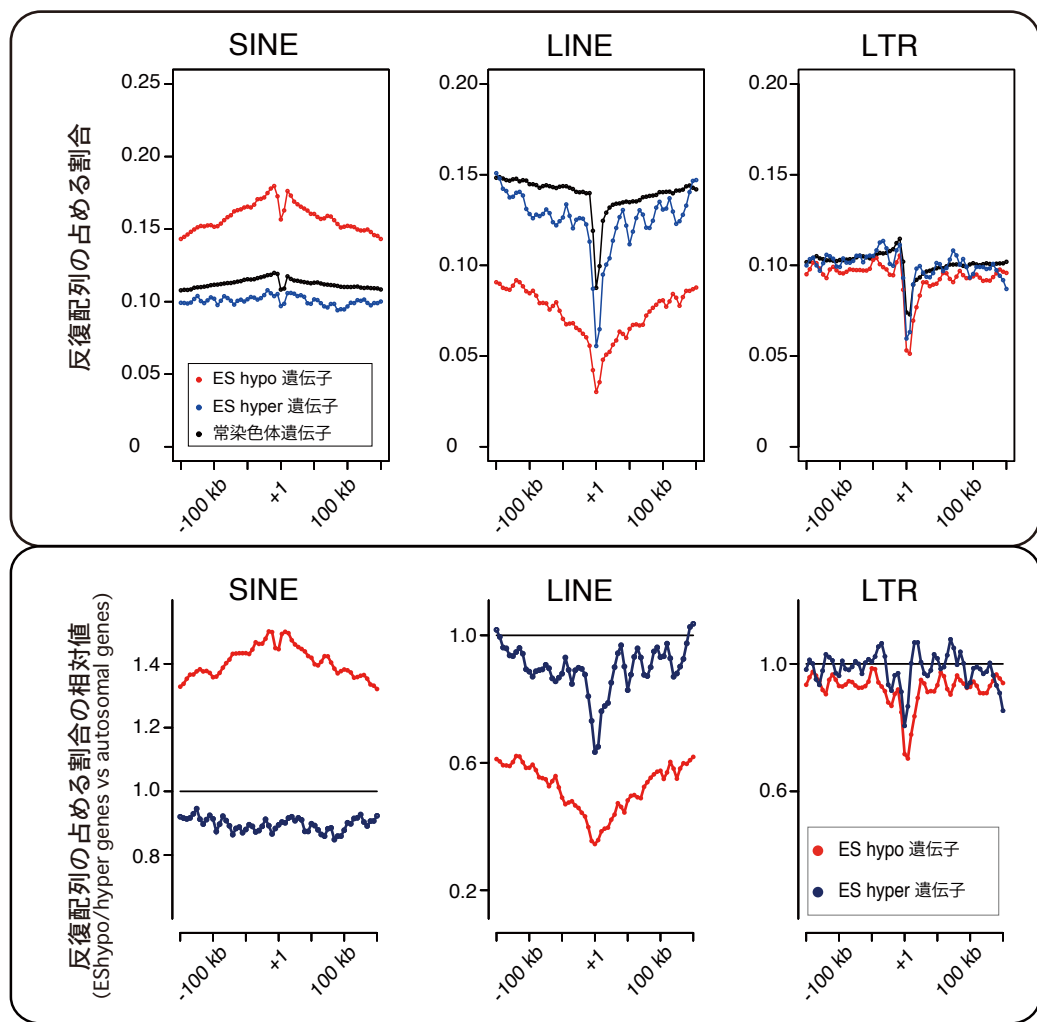
反復配列は哺乳類ゲノムの約半分を占めることから、ゲノムにとって反復配列の存在の悪い面、悪い面についてさまざま議論がなされてきた (Kramarov & Vassetzky, 2005; Ohshima & Okada, 2005; Goodier & Kazazian, 2008)。以前の研究から、SINE 配列が新規メチル化修飾の標的となりうることを示されている (Yates et al., 1999; Arnaud et al., 2000)。ES hypo 遺伝子周囲に存在する SINE 配列の一部は、分化多能性細胞の分化過程において、分化多能性関連遺伝子領域におけるメチル化修飾の標的として機能している可能性がある。その一方で、ES hypo 遺伝子が存在するゲノム領域は SINE 配列が組み込まれ、継承されやすいようなクロマチン環境になっているだけで、SINE 配列自身はメチル化状態の制御には関わらない可能性も考えられる。SINE 配列が T-DMR のメチル化制御に積極的な役割を持つのかについては、今後検討されるべき重要な問題である。

塩基配列の解析から、ES hypo 遺伝子周辺には広い範囲で CpG 配列頻度に特徴があることが分かった。常染色体遺伝子周囲の反復配列は、非反復配列よりも CpG 配列頻度が低かった。これは以前のヒトゲノムにおける報告と一致しており (Rollins et al., 2006)、ヒトとマ

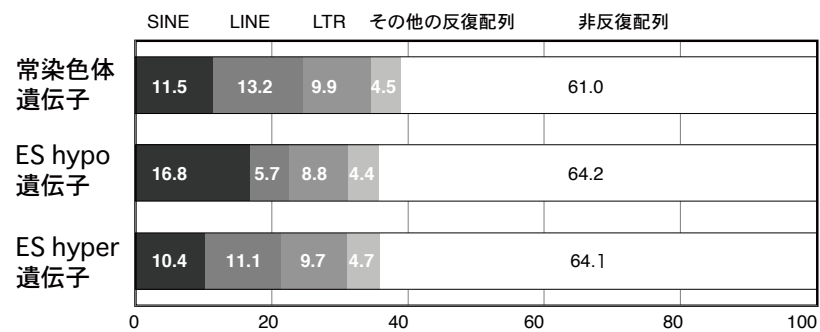
ウスに共通する性質であることが分かった。本章においてはさらに、ES hypo 遺伝子は全常染色体遺伝子に比べ、転写開始点周囲の CpG 配列頻度が 100 kb 以上にわたって高く、さらにこの CpG 配列頻度が高い傾向は反復配列・非反復配列ともに認められることを明らかにした。哺乳類ゲノム上において、CpG 配列頻度に関してもっとも偏りがあるゲノム領域は CpG アイランドである (Fazzari & Greally, 2004)。興味深いことに、ES hypo 遺伝子周辺での高い CpG 配列頻度は、CpG アイランドの有無とは独立した傾向であることが示唆された。これらの結果から、ES hypo 遺伝子周辺領域は、CpG アイランドの有無によらず、反復配列・非反復配列の両方に、100 kb 以上にわたって CpG 配列頻度が高い性質を持つことが分かった。近年のゲノムワイドな DNA メチル化解析においては、転写開始点近傍領域における CpG 配列頻度のみが注目されていることから、本章において明らかになったより広い領域にみられる CpG 配列頻度の特徴は今後のエピゲノム解析に有用な情報となると期待される。

以上、本章の結果より、SINE 配列密度が高いことは ES hypo T-DMR を持つ遺伝子が存在するゲノム領域の配列上の特徴であることが示された。

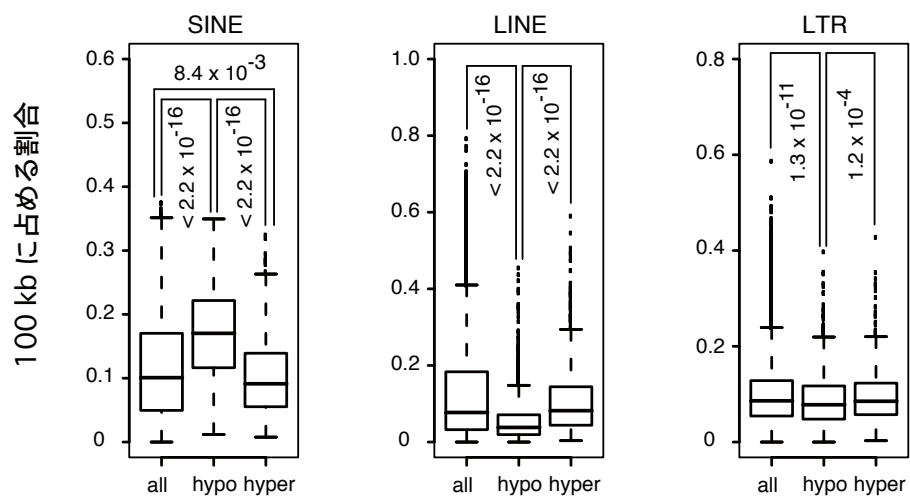
**A**



**B**



**C**



D

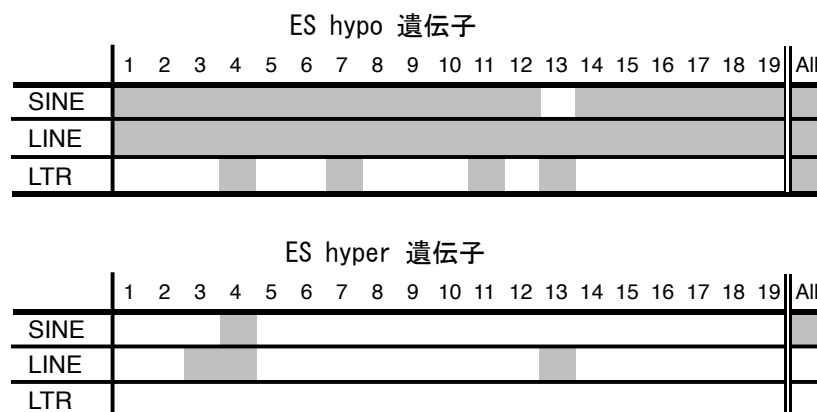


図 1-1 ES hypo 遺伝子周辺における偏った配列上の性質

(A) ES hypo 遺伝子、ES hyper 遺伝子と常染色体遺伝子の周囲における反復配列密度の比較。転写開始点の上流-150 kb から下流 150 kb まで、各 10 kb 領域における反復配列密度の平均値（上段）および相対値（下段）を 5 kb おきに計算した。相対値は、常染色体遺伝子に対する ES hypo 遺伝子、ES hyper 遺伝子の値をそれぞれ赤、青で示す。相対値 1.0 を黒横線で示す。+1 は転写開始点を意味する。

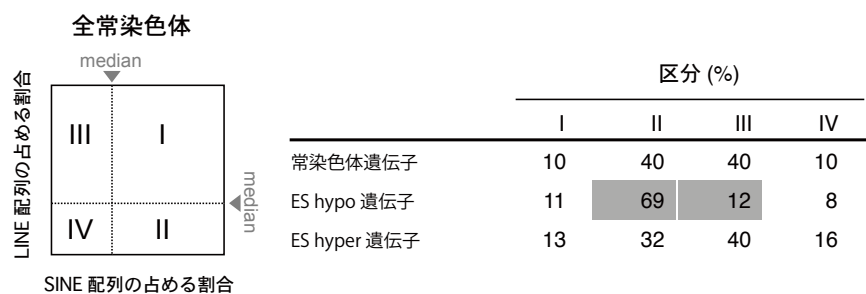
(B) 反復配列の種類による転写開始点周囲 100 kb 領域の分類。ゲノム配列を、UCSC ゲノムブラウザーの RepeatMasker の情報に基づいて、ゲノムを 5 つのカテゴリーに分類した。RepeatMasker により反復配列と判断されなかった領域は非反復配列領域（nonrepeat masked region）と呼ぶことにする。棒グラフは、全 ES hypo 遺伝子、全 ES hyper 遺伝子、全常染色体遺伝子の転写開始点周囲 100 kb 領域における 5 つのカテゴリーが占める割合を示す。具体的な数値はグラフ中に示した。

(C) 転写開始点周囲 100 kb 領域における SINE 配列、LINE 配列、LTR 配列が占める割合の箱ひげ図。箱の下部は第 1 四分位点であり上部は第 3 四分位点を表す。箱中の横線は中央値である。箱からのびた点線の先の横棒は、第 1 四分位点から IQR（interquartile range、第 3 四分位点 - 第 1 四分位点）の 1.5 倍の範囲で最も値の低いデータ、および第 3 四分位点から IQR の 1.5 倍の範囲で最も値の高いデータを表す。この範囲外のデータは、小円で示す。有意な差のある比較においては P 値を明記した（ $P < 0.01$ 、ウィルコクソン順位和検定）。all、全常染色体遺伝子；hypo、ES hypo 遺伝子；hyper、ES hyper 遺伝子

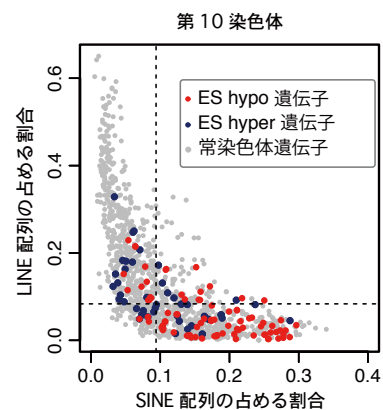
(D) 各常染色体における ES hypo 遺伝子、ES hyper 遺伝子周辺領域の反復配列の偏り。各染色体上の遺伝子周辺領域と比較し、有意に反復配列密度に差があった染色体は棒を灰色で示した（ $P < 0.01$ 、ウィルコクソン順位和検定）。



A



B



C

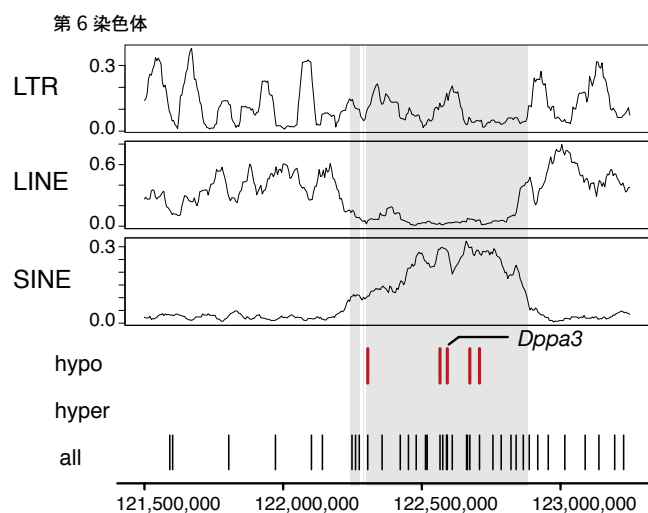
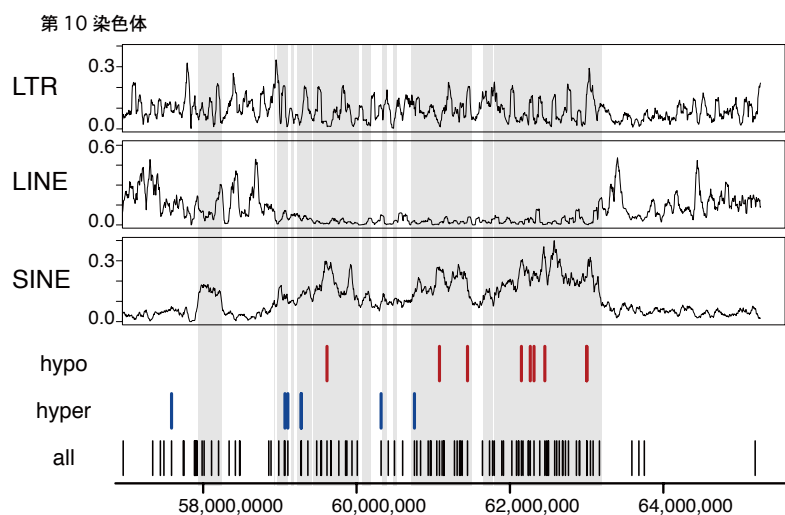


図 1-2 ES hypo 遺伝子周辺領域は SINE 配列密度が高く、LINE 配列密度が低い

(A) ES hypo 遺伝子の SINE 配列密度が高く LINE 配列密度が低い領域への偏り。常染色体遺伝子の転写開始点周辺 100 kb 領域における SINE 配列密度の中央値（縦線）、LINE 配列密度の中央値（横線）により遺伝子を 4 つの区分に分類した模式図を左に示す。右の表はそれぞれの区分に属する遺伝子の割合を示す。常染色体に比べ、値が 20%以上違う区画は灰色の四角で示した。

(B) X 軸が SINE 配列の占める割合、Y 軸が LINE 配列の占める割合を表した散布図。転写開始点周囲 100 kb 領域に占める反復配列の割合を計算した。常染色体遺伝子、ES hypo 遺伝子、ES hyper 遺伝子をそれぞれ灰色、赤、青の点で示した。

(C) ES hypo 遺伝子、ES hyper 遺伝子周辺領域における反復配列の分布の具体例。左が第 10 染色体上、右が第 6 染色体上のゲノム領域である。反復配列の占める割合は、5 kb おきに前後 50 kb 領域について計算した。赤縦線、青縦線、黒縦線はそれぞれ、ES hypo 遺伝子、ES hyper 遺伝子、常染色体遺伝子を表す。グラフの下段には染色体上の場所情報を示す。SINE 配列密度が 0.1（1 割）以上の領域は、灰色で示してある。第 6 染色体上の *Dppa3* 遺伝子は場所を明記した。

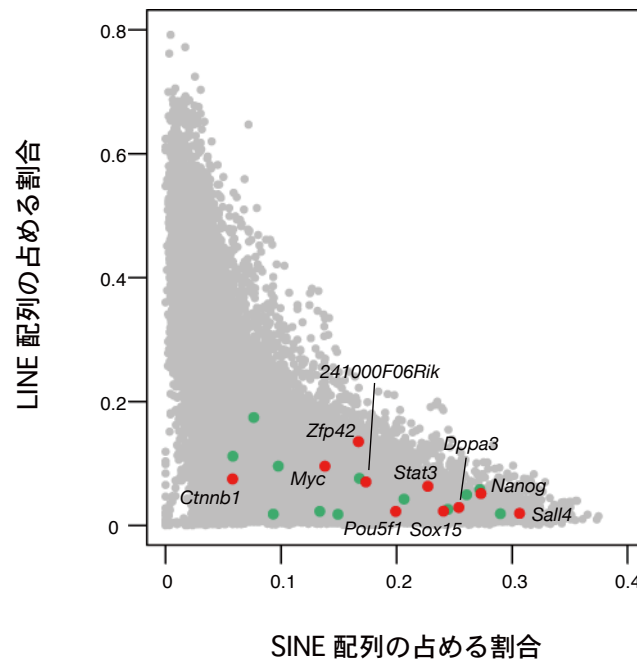
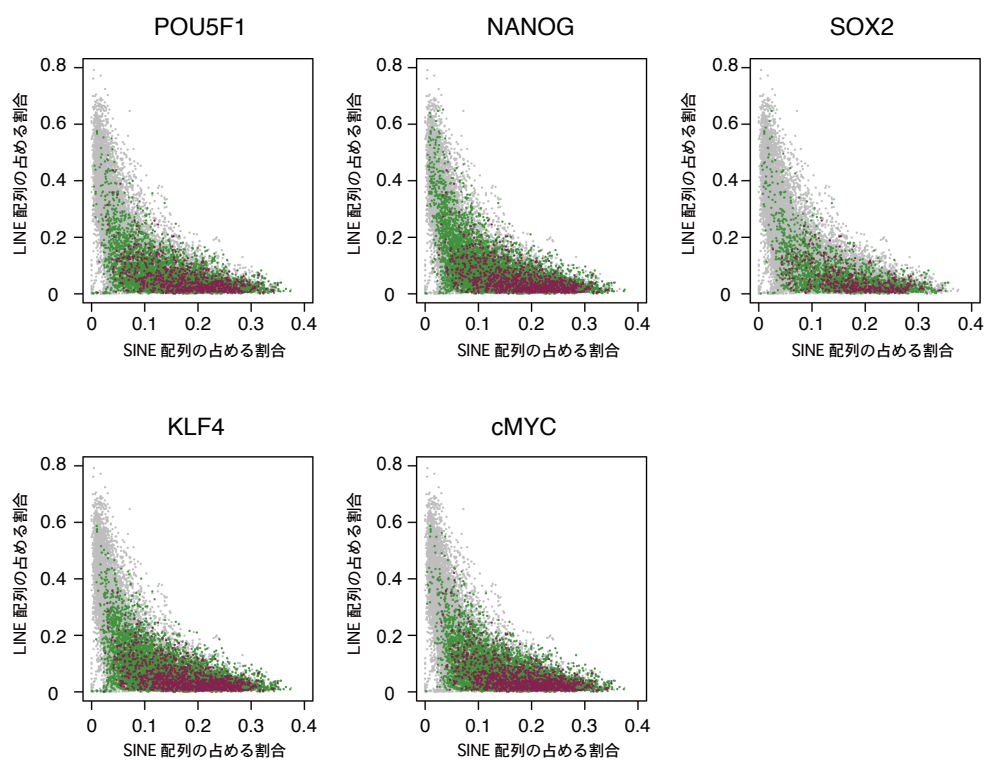


図 1-3

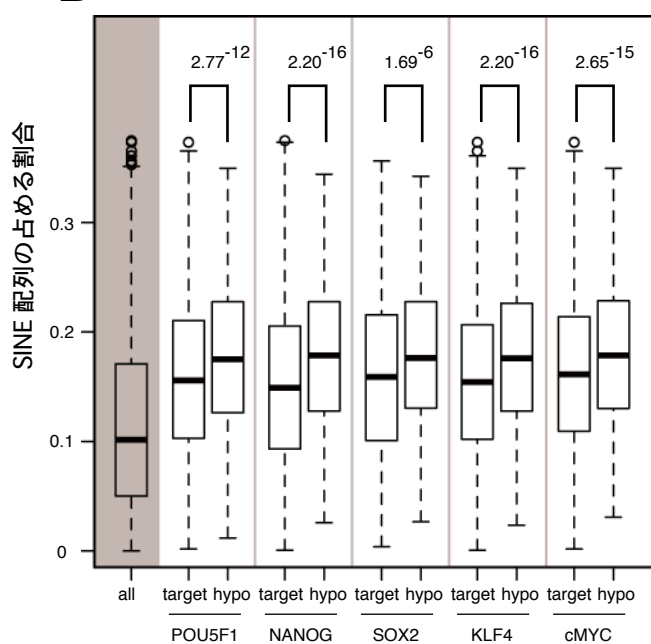
ES hypo T-DMR を有する ECAT 遺伝子周辺領域における SINE 配列、LINE 配列密度

X 軸が SINE 配列の占める割合、Y 軸が LINE 配列の占める割合を表した散布図。転写開始点周囲 100 kb 領域に占める反復配列の割合を計算した。常染色体遺伝子、ES hypo T-DMR を有する ECAT 遺伝子、ES hypo T-DMR を持たない ECAT 遺伝子をそれぞれ灰色、赤、緑の点で示した。ES hypo T-DMR を有する ECAT 遺伝子については遺伝子名を示した。

**A**



**B**



**C**

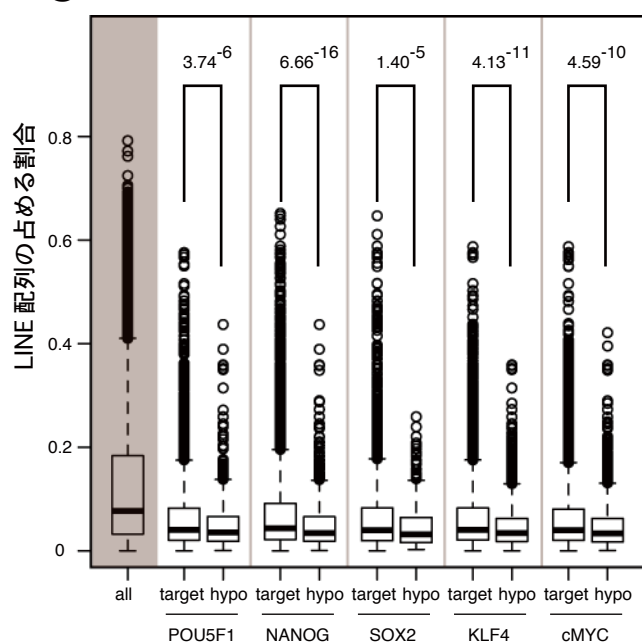


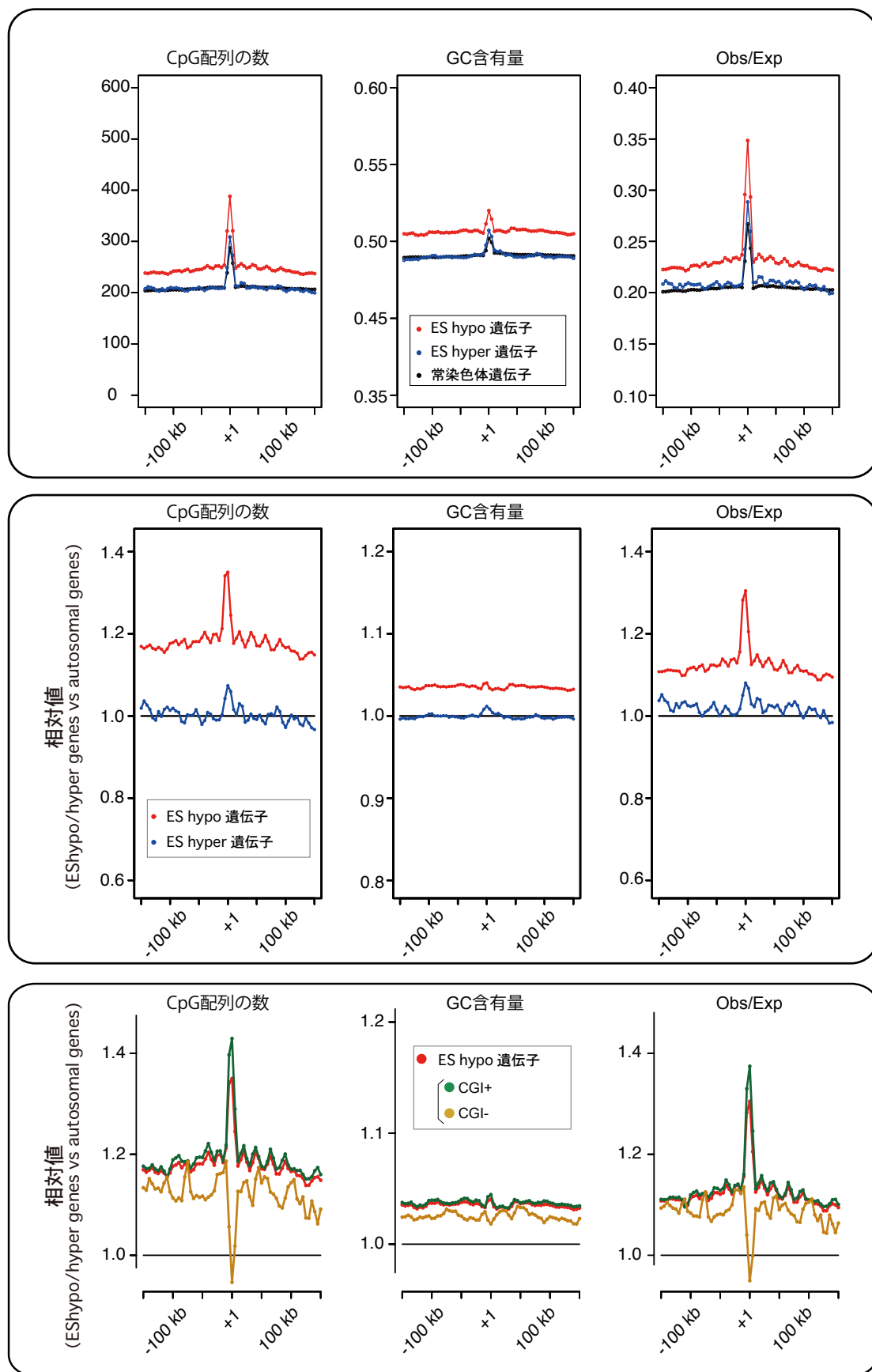
図 1-4 ES 細胞特異的転写因子の標的遺伝子周辺領域における SINE 配列・LINE 配列密度

(A) X 軸が SINE 配列の占める割合、Y 軸が LINE 配列の占める割合を表した散布図。転写開始点周囲 100 kb 領域に占める反復配列の割合を計算した。常染色体遺伝子を灰色の点で、ES 細胞特異的転写因子の標的遺伝子のなかで ES hypo T-DMR を有する遺伝子、持たない遺伝子をそれぞれ赤、緑の点で示した。ES 細胞特異的転写因子 (POU5F1、NANOG、SOX2、KLF4、cMYC) の標的遺伝子は、プロモーター領域 (転写開始点の上流 7 kb から下流 3 kb) に転写因子の結合が ChIP-Seq 解析 (Chen et al., 2008) により検出された遺伝子とした。

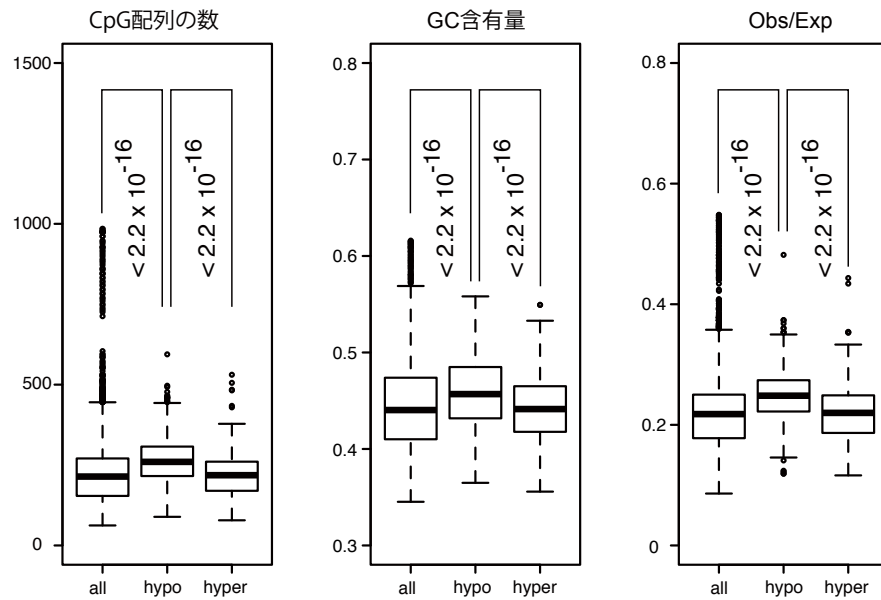
(B) ES 細胞特異的転写因子の標的遺伝子周辺領域における SINE 配列密度の箱ひげ図。転写開始点周囲 100 kb 領域に占める反復配列の割合を示した。箱ひげ図は図 1-1 と同様に描写した。有意な差のある比較においては P 値を明記した ( $P < 0.01$ 、ウィルコクソン順位和検定)。all、常染色体遺伝子; target、転写因子の標的遺伝子; hypo、転写因子の標的である ES hypo 遺伝子

(C) ES 細胞特異的転写因子の標的遺伝子周辺領域における LINE 配列密度の箱ひげ図。転写開始点周囲 100 kb 領域に占める反復配列の割合を示した。箱ひげ図は図 1-1 と同様に描写した。有意な差のある比較においては P 値を明記した ( $P < 0.01$ 、ウィルコクソン順位和検定)。all、常染色体遺伝子; target、転写因子の標的遺伝子; hypo、転写因子の標的である ES hypo 遺伝子

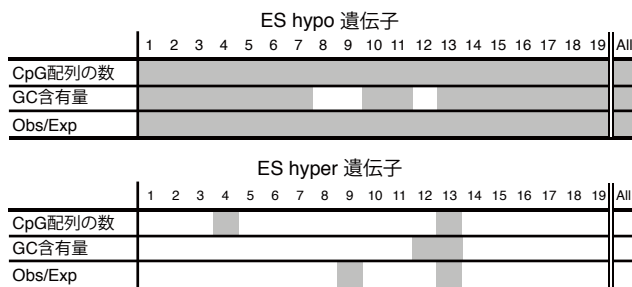
**A**



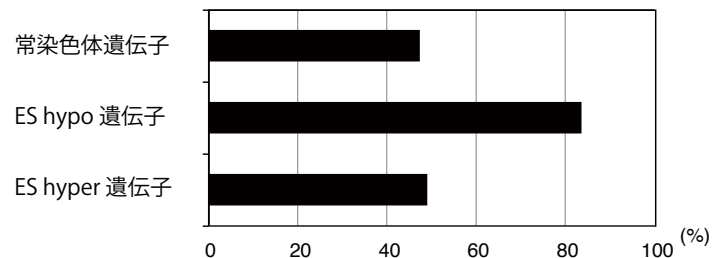
**B**



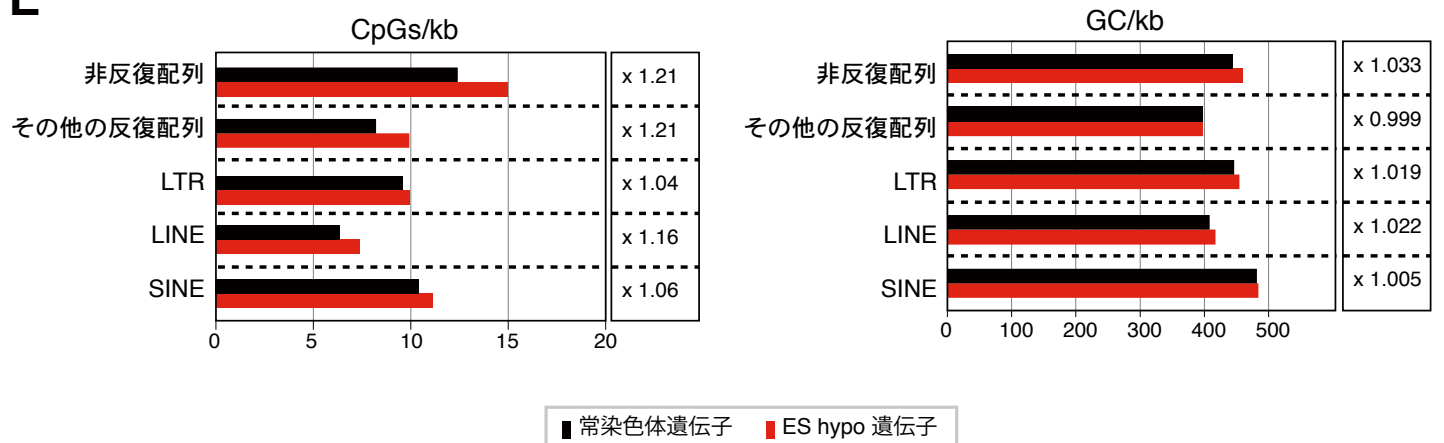
**C**



**D**



**E**



**F**

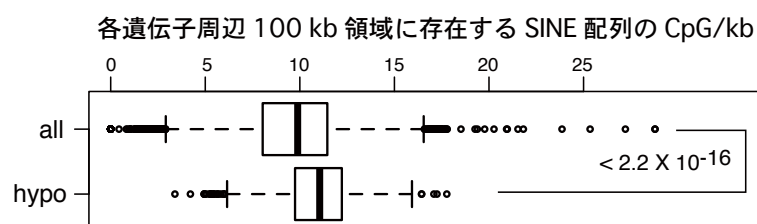


図 1-5 ES hypo 遺伝子周辺における CpG 配列の数、GC 含有量、Obs/Exp

(A) 転写開始点の上流-150 kb から下流 150 kb まで、各 10 kb 領域における CpG 配列の数、GC 含有量、Obs/Exp について平均値、およびその相対値を 5 kb おきに計算した。常染色体遺伝子に対する ES hypo 遺伝子（赤）、ES hyper 遺伝子（青）の相対値を中段に示す。下段には、常染色体遺伝子に対する ES hypo 遺伝子（赤）、CpG アイランドを持つ ES hypo 遺伝子（緑）、CpG アイランドを持たない ES hypo 遺伝子（黄色）を示す。黒横線は相対値 1.0 を表す。+1 は転写開始点を意味する。

(B) 転写開始点周囲 100 kb 領域における CpG 配列の数、GC 含有量、Obs/Exp の箱ひげ図。箱ひげ図は図 1-1 と同様に描写した。有意な差のある比較においては P 値を明記した ( $P < 0.01$ 、ウィルコクソン順位和検定)。all、全常染色体遺伝子；hypo、ES hypo 遺伝子；hyper、ES hyper 遺伝子

(C) 各常染色体における ES hypo 遺伝子、ES hyper 遺伝子周辺領域の CpG 配列の数、GC 含有量、Obs/Exp の偏り。各染色体上の遺伝子周辺領域と比較し、有意な差があった染色体は棒を灰色で示した ( $P < 0.01$ 、ウィルコクソン順位和検定)。

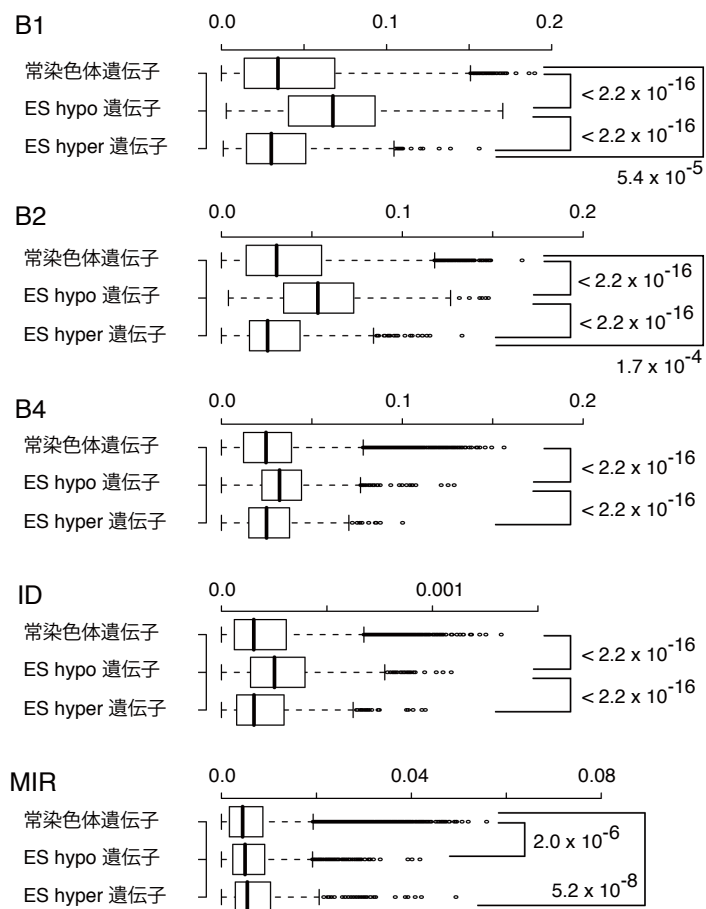
(D) 転写開始点上流 7 kb から下流 3 kb の領域に CpG アイランドを有する遺伝子の割合。

(E) 転写開始点周囲 100 kb 領域における、各ゲノムの区分（SINE、LINE、LTR、その他の反復配列、非反復配列）の 1 kb あたりの CpG 配列の数および GC 含有量（CpG/kb、GC/kb）。黒棒は全常染色体遺伝子周辺における値を、赤棒は全 ES hypo 遺伝子周辺領域における値を示す。常染色体遺伝子に対する ES hypo 遺伝子の相対値を棒グラフの右に示した。

(F) ES hypo 遺伝子周辺領域における各 ES hypo 遺伝子周辺の CpG/kb の箱ひげ図。箱ひげ図は図 1-1 と同様に描写した。有意な差のある比較においては P 値を明記した ( $P < 0.01$ 、ウィルコクソン順位和検定)。all、全常染色体遺伝子；hypo、ES hypo 遺伝子



**A**



**B**

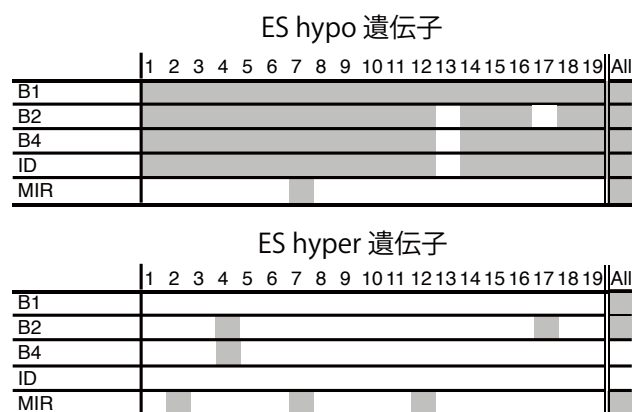
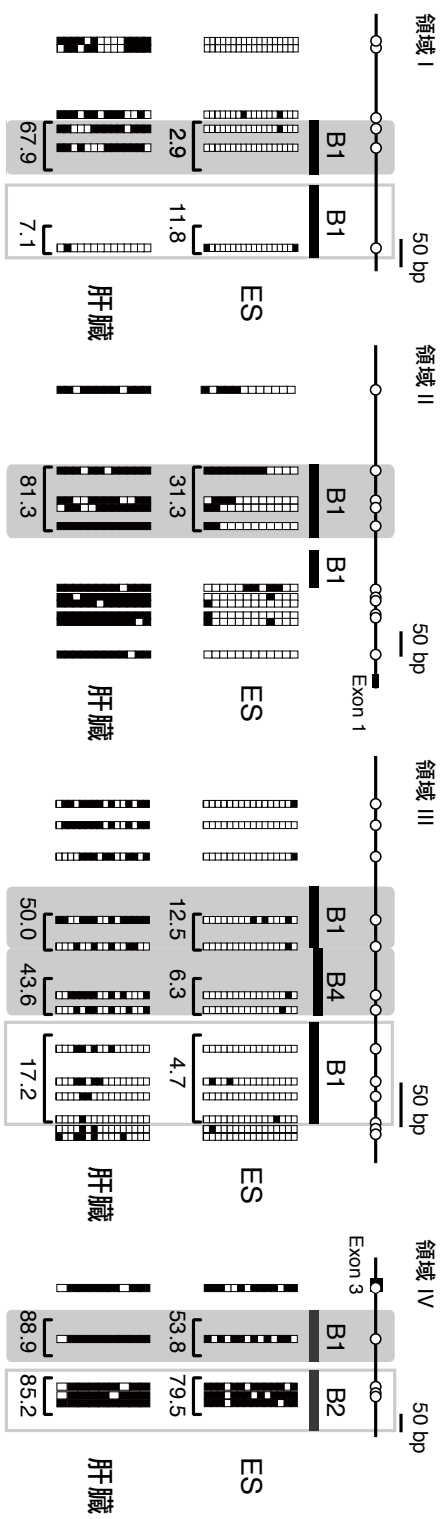
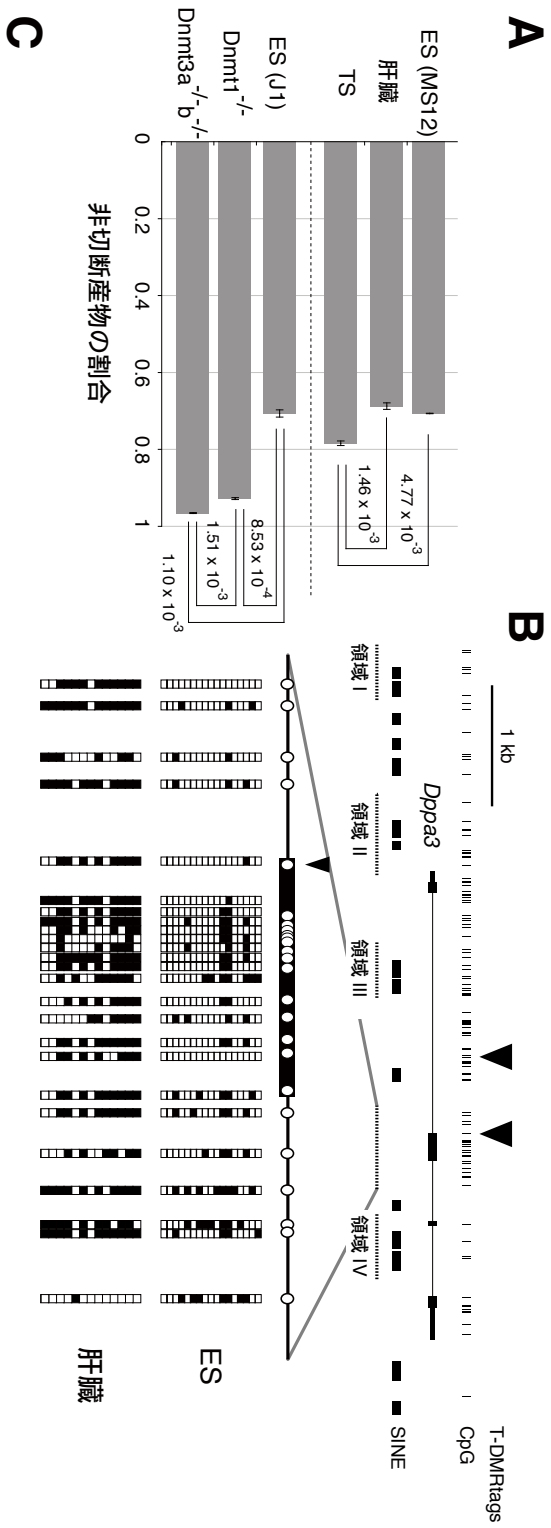


図 1-6 ES hypo 遺伝子周辺における SINE サブファミリーの占める割合

(A) 転写開始点周囲 100 kb 領域における各 SINE サブファミリーが占める割合の箱ひげ図。箱ひげ図は図 1-1 と同様に描写した。有意な差のある比較においては P 値を明記した ( $P < 0.01$ 、ウィルコクソン順位和検定)。

(B) 各常染色体における ES hypo 遺伝子、ES hyper 遺伝子周辺領域に SINE サブファミリーが占める割合の偏り。各染色体上の遺伝子周辺領域と比較し、ES hypo 遺伝子または ES hyper 遺伝子周辺領域の反復配列密度に有意な差があった染色体は枠を灰色で示した ( $P < 0.01$ 、ウィルコクソン順位和検定)。



#### 図 1-7 SINE 配列の DNA メチル化解析

(A) SINE B1 サブファミリーの COBRA 解析。 *TaqI* 制限酵素処理した PCR 産物中で、切断された産物と非切断産物の合計に対する非切断産物の割合を示した。有意な差のある比較においては P 値を明記した ( $P < 0.01$ , t 検定)。 *Dnmt1* 欠損 ES 細胞、 *Dnmt3a/3b* 両欠損 ES 細胞をそれぞれ *Dnmt1*<sup>-/-</sup>、 *Dnmt3a*<sup>-/-</sup>*3b*<sup>-/-</sup> と表記した。

(B) *Dppa3* 遺伝子領域の模式図と ES hypo T-DMRtag のバイサルファイトシーケンス解析。図中において上から、ES hypo T-DMRtag (黒い矢頭)、CpG 配列 (縦線)、*Dppa3* 遺伝子、SINE 配列 (黒い横実線)、バイサルファイトシーケンスによる解析対象領域 (黒い横点線) を示す。図の下部では、CpG 配列 (白い円) と *Dppa3* 遺伝子の位置、ES 細胞および肝臓におけるメチル化状態を示す。白い四角は非メチル化 CpG を、黒い四角はメチル化 CpG を表す。

(C) *Dppa3* 遺伝子領域に存在する SINE 配列のメチル化状態。図 4B 中の領域 I から IV を解析した。上から CpG 配列と *Dppa3* 遺伝子の位置、SINE 配列のサブファミリーの種類と位置、ES 細胞および肝臓におけるメチル化状態を示す。各 SINE 配列のメチル化率については数値を示した。ES 細胞の方が肝臓よりもメチル化率が 30%以上低い SINE 配列は灰色で塗って示してある。

表 1-1 ES hypo 遺伝子、ES hyper 遺伝子周辺領域の反復配列密度

		P value*	Ratio of the median †	The number of autosomes ‡
ES hypo 遺伝子	LTR	$2.3 \times 10^{-11}$	0.91	4
	LINE	$< 2.2 \times 10^{-16}$	0.50	19
	SINE	$< 2.2 \times 10^{-16}$	1.68	18
	B1	$< 2.2 \times 10^{-16}$	2.05	19
	B2	$< 2.2 \times 10^{-16}$	1.81	17
	B4	$< 2.2 \times 10^{-16}$	1.32	18
	ID	$< 2.2 \times 10^{-16}$	1.72	18
	MIR	$2.0 \times 10^{-6}$	1.09	1
ES hyper 遺伝子	LTR	n. s. §	0.99	0
	LINE	n. s. §	1.07	3
	SINE	$8.8 \times 10^{-2}$	0.91	1
	B1	$5.4 \times 10^{-5}$	0.86	0
	B2	$1.7 \times 10^{-4}$	0.83	2
	B4	n. s. §	1.00	1
	ID	n. s. §	1.01	0
	MIR	$5.2 \times 10^{-8}$	1.18	3

\*ウィルコクソン順位和検定

† 常染色体遺伝子に対する ES hypo 遺伝子または ES hyper 遺伝子の、各反復配列が転写開始点前後 100 kb 領域に占める割合の中央値の相対値

‡ 各染色体の解析から、常染色体遺伝子と ES hypo 遺伝子または ES hyper 遺伝子の間に、周辺における反復配列に有意な差が認められた常染色体の数

§ n. s., 有意な差は認められなかった (not significant)

表 1-2 ES hypo 遺伝子、ES hyper 遺伝子周辺領域の CpG 配列の数、GC 含有量、Obs/Exp

		P value*	Ratio of the median †	The number of autosomes ‡
ES hypo 遺伝子	CpG 配列の数	$< 2.2 \times 10^{-16}$	1.21	19
	GC 含有量	$< 2.2 \times 10^{-16}$	1.04	16
	Obs/Exp	$< 2.2 \times 10^{-16}$	1.14	19
ES hyper 遺伝子	CpG 配列の数	n. s. §	1.02	2
	GC 含有量	n. s. §	1.00	2
	Obs/Exp	n. s. §	1.01	2

\*ウィルコクソン順位和検定

† 常染色体遺伝子に対する ES hypo 遺伝子または ES hyper 遺伝子の、転写開始点前後 100 kb 領域における中央値の相対値

‡ 各染色体の解析から、常染色体遺伝子と ES hypo 遺伝子または ES hyper 遺伝子の間に、周辺における CpG 配列の数、GC 含有量、Obs/Exp に有意な差が認められた常染色体の数

§ n. s., 有意な差は認められなかった (not significant)

表 1-3 プライマーリスト

名前	配列	対象となるゲノム領域 (mm8)
SINE_B1_F	ttagtatgtggaggtagaggtag	図1-7A
SINE_B2_R	aaacaaaa ttctctatataaccctaa	
Dppa3_F	ggggagaa tatgtaggttttagaattt	chr6:122593970-122594654
Dppa3_R	caaacacccaataactaatactactacta	図1-7B
RegionI_F	gaagtattt tgggttatgtgaaat	chr6:122590187-122590644
RegionI_R	aaaataaaaaccaa aatcccata	図1-7C領域I
RegionII_F	gatttaagatgaagtagttaaaggaaat	chr6:122591421-122592077
RegionII_R	acctaaa accaaaacccaattacaa	図1-7C領域II
RegionIII_F	gaggtaatagtgttggaattgaattt	chr6:122592630-122593097
RegionIII_R	aacccaaattcaactaaacttaaaaa	図1-7C領域III
RegionIV_F	gtgtttaaagtaggtttgatagtttga	chr6:122594862-122595295
RegionIV_R	aaacactaaactactcttccaaaaatccta	図1-7C領域IV
N18	ggcaactgtgctatccga	D-REAM
R18	gcactctccagcctctca	
Rアダプター	agcactctccagcctctcaccgct cgagcgtga	D-REAM
Nアダプター	aggcaactgtgctatccgaggga cggtccctcg	D-REAM

## 第二章

ES 細胞に特有な組織特異的 DNA メチル化可変領域 (T-DMR) と染色体構造および核内構造との関係



## 序論

真核生物のゲノムは、核膜によって囲われた核という空間に収容されている。核は核膜以外に物理的な区切りを持たないため、核内では物質が自由に離合、集散できると考えられる。そのなかで核内は物質的、機能的に不均一であることが様々な観点から明らかになっている。例えば、核内は DNA が凝集し DAPI 染色によりよく染まるヘテロクロマチン領域とその他のユークロマチン領域に分かれる。転写の活発なクロマチンにみられるヒストン H3K4 や H3K36 のメチル化修飾はユークロマチン領域と、転写が抑制されたクロマチンに見られる H3K9 や H3K27 のトリメチル化はヘテロクロマチン領域と一致する (Kouzarides, 2007; Wang et al., 2009)。また転写に関しては、核内で RNA ポリメラーゼ II (Pol2) が転写ファクトリーと呼ばれる空間領域に集まって存在しており (Iborra et al., 1996)、この転写ファクトリーにクロマチンが接近することで遺伝子が転写されるというモデルが提唱されている (Cook, 2002)。さらに、染色体自身も核内にランダムに存在するのではなく、染色体ごとに核内の異なる区画 (染色体テリトリーと呼ばれる) に存在している (Cremer et al., 1982)。

真核生物のゲノムは間期核のなかで、物理的に階層性を持った構造をとって収容されている (Misteli, 2010)。最も基本的な構造は八量体のコアヒストンに DNA が巻き付いたヌクレオソームである。複数のヌクレオソームはリンカーヒストンによってつながれ、DNA という糸にヌクレオソームというビーズを通したような直径約 10 nm の繊維状構造をとる。この構造はさらに折り畳まれるが、生体内での詳しい構造は分かっていない。より高次の構造として、折り畳まれたクロマチンの核内空間での位置配置があり、その上位にクロモソームテリトリーがある。ゲノムの高次構造は、細胞・組織ごとに異なり、遺伝子発現制御に関わる (Parada & Misteli, 2002)。

間期核の高次構造はさまざまなクロマチン因子によって制御される。その一つがジンクフィンガー（zinc finger）タンパク質の CTCF である。CTCF は、ゲノム上における発現制御の区切り役として機能する、いわゆるインスレータータンパク質としてよく知られている（Xie et al., 2007）。近年の研究から CTCF が、同一染色体の離れた領域同士あるいは異なる染色体上の領域同士の相互作用に関わることが明らかになってきた（Phillips & Corces, 2009）。また、コヒーシン（Cohesin）も高次構造に関わるクロマチン因子である。コヒーシンとは間期の S 期で DNA が複製され生じた姉妹染色分体を束ねて接着させるのに必須なタンパク質複合体である（Peters et al., 2008）。この複合体は、SMC1、SMC3 を含む進化的に保存された 4 つの中心的なタンパク質と、染色体へのローディングに関わる NIPBL などの複数の関連因子からなる。近年、コヒーシンがゲノムの高次構造に関わることを示唆する結果が酵母、ショウジョウバエ、マウス、ヒトで得られている（Glynn et al., 2004; Lengronne et al., 2004; Misulovin et al., 2008; Wendt et al., 2008）。

ゲノムのエピジェネティック制御と空間配置の関係は興味深い。しかし、この関係を網羅的に解析するのは、現時点では技術的に困難である。従来核内におけるゲノム配置の解析には FISH 法が用いられてきたが、解像度は 100 kb 以上である上に、同時に解析できるゲノム領域の数には限りがあった。その中で、核膜からの距離は様々な生物・細胞でゲノムワイドな解析が可能な数少ないゲノムの空間配置に関するパラメーターの 1 つである。核膜はクロマチンとの相互作用を介して遺伝子発現に関与する。これまでに電子顕微鏡や光学顕微鏡を用いた解析により、ヘテロクロマチンが物理的に核膜と結合していること（Fawcett 1966; Paddy et al., 1990）、発現が抑制された遺伝子や遺伝子が少ないゲノム領域が高い頻度で核膜と相互作用すること（Boyle et al., 2001; Gremer et al., 2001）、人工的に核膜へ結合させたレポーター遺伝子は発現が抑制されること（Reddy et al., 2008）などが報告されており、核膜との相互作

用は転写を抑制すると考えられている。核膜内膜は、核ラミンを主成分とする繊維のネットワーク構造により裏打ちされている。この裏打ち構造が核ラミナである。近年、DNA adenine methyltransferase identification (DamID) 法を用いることで、核膜とゲノムの相互作用を網羅的に解析することが可能になった。核ラミンと大腸菌由来の DNA アデニンメチル基転移酵素 (Dam) との融合タンパク質を発現させ、アデニンがメチル化されたゲノム領域として核ラミナと相互作用するゲノム領域を検出する (Guelen et al., 2008)。これまでの核ラミナと相互作用するゲノム領域 (lamina-associated domain、LAD) の解析から、LAD には抑制性のヒストン修飾である H3K9 ジメチル化や H3K27 トリメチル化が有意に多く認められ、遺伝子発現も抑制されている傾向があることが分かっている (Guelen et al., 2008)。ES 細胞の *in vitro* 神経分化系を用いた解析から、ゲノムの大部分では核膜との相互作用は分化に伴い変化しないものの、分化過程で核膜方向へ移動した遺伝子は発現が分化後に低下する傾向にあり、一方核膜から離れた遺伝子は発現が分化後に上昇する傾向にあることが示された (Peric-Hupkes et al., 2010)。

ES 細胞は分化細胞とは異なる核内構造を持つ。核内動態に関しては、ES 細胞では HP1 などのクロマチン関連因子の置き換えりが分化細胞よりも早く、核内での物質の置き換えりが活発であると考えられている (Meshorer et al., 2006)。また、ヘテロクロマチンにも特徴がある。ES 細胞ではセントロメアのヘテロクロマチン領域どうしが集まり、数個のクロモセーターと呼ばれるヘテロクロマチン領域を形成している。クロモセーターの大部分は核膜とは接しておらず、核内部に存在する。クロモセーターは分化に伴い分裂し、核膜とより密に接するようになる (Yamazaki et al., 2003)。電子顕微鏡による解析では、核膜・核小体周辺のヘテロクロマチンが ES 細胞では薄く、クロマチンが緩んだ領域 (ユークロマチン) の占める割合が多いことが明らかにされている (Baharvand & Matthaei, 2003)。

ES 細胞のトランスクリプトームを制御する因子として、様々なクロマチン関連因子

が同定されてきた。*Pou5f1* 遺伝子、*Nanog* 遺伝子をはじめ、ES 細胞特異的な転写因子は相互に制御し合う転写因子ネットワークを構成して下流遺伝子の発現を制御する (Chen et al., 2008)。また Kagey らによって、メディエーター因子、コヒーシン関連因子が ES 細胞の分化多能性に必須であること、転写が活発な遺伝子にゲノムワイドに結合し転写を制御することが示された (Kagey et al., 2010)。

ES hypo T-DMR の ES 細胞特異的な DNA の低メチル化は、ヒストン修飾およびクロマチン関連因子の結合と協調して遺伝子発現を制御する (Ikegami et al., 2009)。ES 細胞特異的に発現し、分化多能性に必須な転写因子である *Pou5f1* 遺伝子、*Nanog* 遺伝子のプロモーター領域に存在する ES hypo T-DMR は、ES 細胞において活性型のヒストン修飾であるアセチル化および H3K4 のメチル化修飾を受ける (Hattori et al., 2004; Hattori et al., 2007)。ヒトにおいても、繊維芽細胞よりも ES 細胞で低メチル化状態にある領域のなかに、H3K4 のトリメチル化と H3K36 のトリメチル化修飾を受ける領域があり、その近傍の遺伝子は繊維芽細胞に比べ ES 細胞で発現が高いことが示されている (Lister et al., 2009)。また T-DMR と転写因子に関しては、遺伝子領域を対象にしたゲノムワイドなメチル化解析から、プロモーター領域に ES hypo T-DMR を有する遺伝子は有意に高い割合で ES 細胞特異的な転写因子ネットワークを構成する転写因子の標的となっていることが示されている (Sato et al., 2010)。

これまでのゲノムワイドなメチル化解析から様々な遺伝子が T-DMR を持つことが明らかにされてきたが、これらの解析では主にタンパク質をコードする遺伝子に焦点が当てられてきた。しかし、遺伝子にはタンパク質をコードしない遺伝子（非コーディング遺伝子）が存在する。例えば、古くから知られている転移 RNA 遺伝子やリボソーム遺伝子などに加え、近年存在が明らかになったマイクロ RNA (miRNA) 遺伝子や長鎖遺伝子間非コーディング RNA (large intergenic non-coding RNA、lincRNA) などが代表的な例である。これまでに様々な生命現象

に非コーディング遺伝子が関わる事が知られている。特に miRNA と lincRNA は転写制御やエピジェネティック制御に関わる事が近年明らかになってきており (Khalil et al., 2009; Wang et al., 2010)、これらの遺伝子を含め非コーディング遺伝子が T-DMR を持つかは重要な問いであるが、T-DMR の有無をゲノムワイドに調べた報告はない。また、遺伝子領域以外 (遺伝子間領域) の T-DMR ついても網羅的な解析は行われていない。遺伝子間領域に存在する T-DMR が存在することは明らかになっているが (Sakamoto et al., 2007; Schmidl et al., 2009)、解析対象が限定されており、全体像は不明である。これまでに数多くの T-DMR が同定されてきたが、これまでの解析対象は染色体全域のなかでは限られた範囲であったのが現状である。

第一章の解析では、少なくとも遺伝子のプロモーター領域に焦点を当てる限り、ゲノム上に ES hypo T-DMR が存在する場所と SINE 配列の分布様式に関係があることを明らかにした。この関係が染色体全域で成り立っているのかを検討するためには、プロモーター領域以外の染色体領域においても反復配列と T-DMR の関係を調べる必要がある。第一章で、第 6 染色体上の約 500 kb にわたって SINE 配列密度の高い領域に、ES hypo T-DMR を持つ遺伝子が複数存在することが明らかになった。この領域は *Dppa3* 遺伝子以外にも、*Nanog* 遺伝子、*Gdf3* 遺伝子、*Phc1* 遺伝子などの ES 細胞や初期胚で発現する遺伝子が存在することが知られており (Morgan et al., 2004)、ES hypo T-DMR が存在していたのはこれらの遺伝子のプロモーターであった。本章において染色体全域を対象に反復配列と T-DMR の関係を解析するにあたり、この領域は全体の傾向を知るための 1 つの典型例として利用することが出来る。従って、本章では第 6 染色体を解析対象に含むフルタイリングアレイを使用した。

本章では、染色体全域を解析対象として ES 細胞に特異的な低メチル化 T-DMR 周囲の反復配列密度の関係を解析した。さらにヒストン修飾、クロマチン関連因子の結合、さらに核内空間制御の観点から解析を行い、ES hypo T-DMR のエピジェネティック制御と SINE 配列密度

の関係を明らかにした。

## 材料および方法

### ES 細胞および肝臓

ES 細胞の培養および生体マウスからの肝臓のサンプリングは第一章に準ずる。

### NIH/3T3細胞

NIH/3T3用培地を用い、5% CO<sub>2</sub>、37°Cで培養した。培地は一日おきに交換した。

#### ・ NIH/3T3細胞用培地

DMEM pH 7.4 (Invitrogen、Carlsbad) に、10% FBS (JRH Biosciences、Lenexa)、100 U/ml penicillin (Invitrogen)、100 mg/ml streptomycin (Sigma-Aldrich、St. Louis) を添加した。

### 神経分化ES細胞

PA6細胞 (Kodama et al., 1986) をフィーダー細胞として用い、ES細胞を神経分化細胞培地中で10日間培養した。培地は4日目、6日目、8日目に交換した。ゼラチンコートされた10 cmディッシュ上でコンフルエントになるように神経分化細胞培地中で培養したPA6細胞を使用した。

#### ・ 神経分化培地

G-MEM medium (Invitrogen) に、10% KSR (Invitrogen)、2 mM L-glutamine (Invitrogen)、1 mM/ml sodium pyruvate (Invitrogen)、100  $\mu$ M/ml MEM nonessential acids (Invitrogen)、100  $\mu$ M/ml 2-mercaptoethanol (Invitrogen) を添加した。

## ゲノム DNA の抽出

操作は第一章に準ずる。

## D-REAM

D-REAM は第一章と同様に行った。アレイとして、第 6、8、16 染色体全域を解析対象とする GeneChip Mouse Promoter 2.0R (Affimetrix、Santa Clara) の F アレイを用いた。ES 細胞および肝臓を対象とした。各サンプル 2 回ずつ D-REAM を行った。

## マスアレイ解析

ES細胞、神経分化ES細胞、肝臓を用いたマスアレイ解析は、Sequenome (San Diego) の標準プロトコルに従って行った。メチル化状態を解析する原理は、バイサルファイト反応により生じるDNA配列から予測される質量と、MALDI-TOF質量分析により実際に計測されたサンプルの質量を比較することで、バイサルファイト反応産物中の各CpG配列のメチル化状態を明らかにするというものである。プライマーはEpidesigner web tool (<http://www.epidesigner.com/>) により設計した。この際、reverseプライマーには、T7プロモーター配列を5'末端に付加した。forwardプライマーには、融解温度をreverseプライマーにあわせるために、10塩基を5'末端に付加した。プライマーの配列は表2-3に示した。バイサルファイト反応はEZ DNA Methylation Kit (Zymo Research、Irvine) を用いた。反応条件は、95°Cを30分間、50°Cを15分間のサイクルを45サイクル行った。PCR反応のテンプレートDNA濃度は10 ng/μl 以上となるようにした。PCR反応溶液中においては、最終濃度が2 ng/μl 以上となる。プライマーの反応溶液中最終濃度は200 nMとした。PCR反応にはHot Star Taq (QIAGEN、Venlo) を用いた。反応条件は、94°C、15分間：94°C、20秒間、56°C、30秒間、72°C、1分間を45サイクル：



72°C、3分間とした。ただし、プライマーの溶解温度によって、アニーリング温度は調整した。

PCR反応産物は、Shrimp Alkaline Phosphatase (SAP) により脱リン酸化した。試験管内転写反応 (In vitro transcription) とRNase A処理はMassCLEAVE Kit (Sequenome) を用いて行った (Ehrich et al., 2005)。RNase処理により切断された試験内転写反応産物は、MALDI-TOF質量分析計 (MassARRAY system、Sequenome) により解析した。メチル化状態は、マスペクトル情報よりEpityper software v1.0 (Sequenome) を用いて、定量化した。

#### 核ラミンを用いたDamIDデータの取得と処理

ES細胞、神経前駆 (NPC) 細胞、アストロサイト (AC) 細胞、マウス胎仔繊維芽 (MEF) 細胞のDamIDのデータ (GSE17051) (Peric-Hupkes et al., 2010) はGene Expression Omnibus (GEO) <<http://www.ncbi.nlm.nih.gov/geo/>>より取得した。第6、8、16染色体上の全てのHpyCH4IV認識部位 (Hpyサイト) について、周囲6 kbに含まれるプローブのシグナル値の平均を計算した。シグナル値はPeric-Hupkesらによって正規化されたものを用いた。周囲6 kb領域にプローブが3個以下しか存在しないHpyサイトは、解析から除外した。各Hpyサイトについて、ES細胞と分化細胞間での周辺領域の平均シグナル値の差を計算し、 $\Delta$ Laminとした。

#### ヒストン修飾、クロマチン関連因子結合データの取得と処理

ES細胞のヒストン修飾データおよびクロマチン関連因子結合データは、UCSCおよびGEOより取得した (Mikkelsen et al., 2007; Meissner et al., 2008; Kagey et al., 2010)。結合あるいは修飾の領域情報 (BED ファイル) は、Galaxy (Giardine et al., 2005) <<http://g2.bx.psu.edu/>>のlift over機能によりmm9ゲノム上での位置情報へ変換した。

## 配列解析

反復配列の計算方法は第一章と同様に行った。染色体上5 kbごとに前後50 kb領域に含まれる反復配列の密度の計算は、Galaxyのcoverage機能により行った。遺伝子情報はEnsemblより取得した。この情報をもとに、遺伝子領域を、「転写開始点上流7 kbから転写終結点下流3 kb」と定義した。遺伝子領域に含まれるHpyサイトは、Galaxyのintersect機能により調べた。

## RNA発現解析

TRIzol reagent (Invitrogen)を用いてtotal RNAを抽出した。これをRNase-free DNase I (Takara, Kyoto)を用いて37°Cで15分間処理した。1 µgのtotal RNAを調整し、RT反応に用いた。RT反応ではランダムヘキサマーをプライマーとし、SuperScript III First Strand Synthesis for RT-PCR (Invitrogen)を逆転写酵素として利用した。

合成したcDNAを鋳型としてPCRを行なった。14倍希釈したcDNAを2 µl加えた20 µlの系で、反応条件は $\beta$ -actinのプライマーセットでは95°C 30秒; 94°C 30秒, 55°C 30秒, 72°C 1分 × 25サイクル; 72°C 10分とし、その他のプライマーセットでは95°C 30秒; 94°C 30秒, 60°C 30秒, 72°C 1分 × 35サイクル; 72°C 10分とした。使用したプライマーの配列は表2-4に示した。PCR産物は2% アガロースゲル中で電気泳動を行い、ethidium bromide染色によって可視化した。

## 結果

### 染色体全域を対象としたES細胞特異的なT-DMRの解析

マウスES細胞および肝臓を用いて、第6、8、16染色体全域を解析対象とするマウスフルタ  
イリングアレイによるD-REAM解析を行った。ES細胞において肝臓よりも低メチル化状態にある  
*HpyCH4IV*認識部位 (ES hypo T-DMRtag) を5027ヶ所、高メチル化にある *HpyCH4IV*認識部位 (ES hyper  
T-DMRtag) を4320ヶ所検出した。図2-1に第6染色体の結果を示した。

T-DMRtagが遺伝子領域と遺伝子間領域に存在する割合を調べた (図2-2)。本研究では、遺  
伝子領域を遺伝子転写開始点上流7 kbから転写終結点下流3 kbまでの領域とした。ES hyper  
T-DMRtagは、*HpyCH4IV*認識部位 (Hpyサイト) とほぼ同じ約5割が遺伝子領域に存在した。これ  
に対し、ES hypo T-DMRtagは83%が遺伝子領域に存在した。これらの結果からT-DMRは遺伝子領  
域・遺伝子間領域ともに存在すること、さらに、ES細胞に特異的な低メチル化T-DMR (ES hypo  
T-DMR) は遺伝子領域に偏っていることが明らかになった。

### ES細胞特異的なT-DMRは非コーディング遺伝子にも存在する

T-DMRtagが存在した遺伝子領域について、遺伝子の種類を調べた (図2-3)。4172のES hypo  
T-DMRtagは、1386の遺伝子領域に存在した。これらのES hypo T-DMRtagを持っていた遺伝子の  
うち、1174 (85%) はタンパク質をコードする遺伝子であり、212 (15%) はタンパク質をコード  
しない遺伝子 (非コーディング遺伝子) であった。ES hyper T-DMRtagを有する遺伝子においても、  
17%が非コーディング遺伝子であった。

miRNA遺伝子に着目すると、23個のmiRNA遺伝子領域にES hypo T-DMRtagが存在した。miRNA  
が産生される工程には、(1)ゲノムに複数がまとまって存在するmiRNA遺伝子クラスターが転写制

御機構を有し、RNAポリメラーゼⅡ（PolⅡ）によりmiRNA前駆体が転写される場合と、(2)miRNA遺伝子が他の遺伝子のイントロン内に存在しスプライシングにより前駆体が作られる場合の2通りが知られている（Tiscornia et al., 2010）。ES hypo T-DMRtagを有するmiRNA遺伝子には、前者（図2-4上段）と後者（図2-5下段）の両方があった。

長鎖ノンコーディングRNA遺伝子にもT-DMRを持つものが存在した（図2-5）。*RP23-103L13.1* 遺伝子はエクソン・イントロン構造を持つ非コーディング遺伝子である。D-REAM解析により第一イントロン領域に3つのES hypo T-DMRtagが検出された。また、ES hypo T-DMRtagの近傍（図2-5上段領域A、B）を対象にしたマスアレイによるメチル化解析から、領域Aの2つのCpG配列が神経分化細胞と肝臓よりもES細胞で低メチル化であることが分かった（図2-5下段）。以上の結果から、*RP23-103L13.1* 遺伝子がES hypo T-DMRを有することが示された。

以上より、T-DMRはタンパク質をコードする遺伝子だけでなく、非コーディング遺伝子にも存在することが明らかになった。

### SINE配列密度の高いゲノムドメインへのES hypo T-DMRの偏り

ES hypo T-DMRtag、ES hyper T-DMRtagの周辺におけるSINE配列、LINE配列、LTR配列の分布を解析するために、T-DMRtagおよびHpyサイトの周囲50 kb領域に反復配列が占める割合を計算した（図2-6上段）。HpyサイトとES hyper T-DMRtagの周辺では、反復配列密度は同様の傾向を示した。一方でES hypo T-DMRtagはHpyサイトに比べ、周囲のSINE配列密度が高く、LINE配列密度が低い傾向が認められた。さらに、SINE配列密度とLINE配列密度の散布図（図2-7）から、Hpyサイト、ES hyper T-DMRtagに比べ、ES hypo T-DMRtagは周囲のSINE配列密度が高くかつLINE配列密度が低い傾向を持つことが分かった。

次に、染色体上でのT-DMRと反復配列の分布の関係を明らかにするために、SINE配列密度の高さを指標に5つの区分に染色体全域を分類し、各区分にHpyサイトおよびT-DMRtagが存在する割合を調べた（図2-8）。Hpyサイトは各区画に約20%が属し、ゲノムの構成比と同様の分布を示した。このことは、*HpyCH4IV*を用いたD-RMEAは、ゲノム上をSINE配列密度に関してほぼ偏りなく解析できることを意味している。ES hyper T-DMRtagについて調べたところ、各区画への分布はHpyサイトと同様であった。それに対しES hypo T-DMRtagは、59%がhighest、25%がhighの区分に属したことから、ゲノムの中でも特にSINE密度の高い区画へ偏って存在することが示された。

さらに、ES hypo T-DMRtagを遺伝子領域・遺伝子間領域に存在するものに分類して周囲のSINE配列密度を解析した（図2-6中段、下段）。その結果、両方の領域でES hypo T-DMRtagはHpyサイトに比べ周囲のSINE密度が高かった。つまり、以上の傾向は遺伝子領域・遺伝子間領域に関わらず、ES hypo T-DMRに認められる配列密度の性質である。

以上より、染色体全域を対象とするメチル化解析から、ES hypo T-DMRは、遺伝子領域・遺伝子間領域に関わらずSINE密度の高いゲノムドメインに集中して存在することが明らかになった。

### **Nanog遺伝子が存在するSINE配列密度の高いゲノムドメインの解析**

染色体全域で、ES細胞におけるES hypo T-DMRのクロマチン関連因子の結合およびヒストン修飾を明らかにするために、GEO（Gene Expression Omnibus）等のデータベースからこれらの因子のChIP-seqデータ（表2-1、表2-2）を取得し、D-REAMのデータと比較した。

まず、第一章においても解析した、第6染色体上の約500 kbにわたってSINE配列密度が高いゲノムドメインに着目した（図2-9）。遺伝子領域と遺伝子間領域の両方に複数のES hypo T-DMRtag

が存在した。まず遺伝子領域に存在するT-DMRとして、*Nanog*遺伝子、*Phc1*遺伝子が持つES hypo T-DMRを解析した。*Nanog*遺伝子領域にES hypo T-DMRが存在することは、これまでの研究でも明らかにされている (Hattori et al., 2007; Imamura et al., 2007; Sato et al., 2010)。*Phc1*遺伝子領域にES hypo T-DMRが存在することはマスアレイ法によるメチル化解析によって確認された (図2-11下段)。これらの遺伝子領域には、ES細胞特異的な転写因子、メディエーター因子、転写関連因子、核内高次構造関連因子といった様々なクロマチン関連因子が結合することが既にKageyらによって報告されている (Kagey et al., 2010) (図2-10、2-11)。ヒストンの修飾を調べたところ、*Nanog*遺伝子 (図2-10)、*Phc1*遺伝子 (図2-11) の両方で転写の活発な領域にみられる修飾であるH3K79のジメチル化やH3K4のメチル化が認められた。

遺伝子間領域 (図2-9下段の遺伝子間領域1、2) のT-DMRについて解析した。これらの領域にES hypo T-DMRが存在することは、マスアレイによるメチル化解析によって確認した (図2-12下段、2-13下段)。興味深いことに、遺伝子間領域1ではユニーク配列だけでなく一部のSINE配列もES hypo T-DMRであった (図2-12の領域A、B、D、E)。クロマチン関連因子の結合およびヒストン修飾を調べたところ、遺伝子間領域1においてはES hypo T-DMRtagの周辺にPOU5F1、NANOGとメディエーター因子の結合およびH3K4のモノメチル化修飾が示唆された (図2-12上段)。遺伝子間領域2においても、ES hypo T-DMR周辺にクロマチン関連因子の結合と転写の活発な領域にみられるヒストン修飾が認められた (図2-13上段)。ここで、遺伝子間領域2は遺伝子間領域であるが、基本転写因子であるPol2やTBPを含め多くの種類のクロマチン関連因子の結合が示唆されたことから、転写産物の有無を解析したところ、ES細胞において転写産物が存在した (図2-13)。これはエンハンサー活性に伴う転写産物 (Kim et al., 2010)、あるいは新規の遺伝子である可能性が考えられた。

以上よりこのゲノムドメインでは、遺伝子領域・遺伝子間領域に関わらず、ES hypo T-DMRにはクロマチン関連因子の結合と転写の活発な領域にみられるヒストン修飾を伴うことが示唆された。一方で、転写の抑制された領域にみられるH3K27トリメチル化修飾は認められなかった。

次に、ES細胞と分化細胞におけるES hypo T-DMRの核内位置について解析した。T-DMRの核膜からの距離は、核ラミナの主成分である核ラミンを用いて行ったDamID (Peric-Hupkes et al., 2010) のシグナル値を指標とした。シグナル値が高いほど、核ラミナとの相互作用頻度が高いこと、つまり核膜により近い位置にあることを意味する。ES細胞、神経前駆 (NP) 細胞、マウス胎仔繊維芽 (MEF) 細胞のシグナル値を比較した。*Phc1*遺伝子領域に着目すると、ES hypo T-DMRtagは周辺のシグナル値がES細胞に比べNP細胞、MEF細胞で高い (図2-9下段左)。同様の傾向は*Nanog*遺伝子領域、遺伝子間領域1、2においても認められた (図2-9下段右)。つまり、ES hypo T-DMRtag周辺のシグナル値が、遺伝子領域・遺伝子間領域に関わらず上昇する傾向にあった。このゲノムドメインでは、ES hypo T-DMRの核内位置がES細胞に比べ分化細胞でより核膜に近いことが示唆された。

以上の結果より、*Nanog*遺伝子が存在するゲノムドメインのES hypo T-DMRには、ES細胞においてクロマチン関連因子の結合・活性型のヒストン修飾が認められ、分化に伴い核内位置が核膜に近づくことが示唆された。以下では、これらの傾向がゲノムワイドに成り立っているのかを検証した。

## **ES hypo T-DMRにおけるクロマチン関連因子の結合・ヒストン修飾と周辺領域のSINE配列密度の関係**

T-DMRへのクロマチン関連因子（表2-1）の結合をゲノムワイドに解析した。各T-DMRtagの近傍前後6 kb領域におけるクロマチン関連因子の結合を計算した結果、今回解析した全てのクロマチン関連因子がES hypo T-DMRtag近傍領域に有意に高い頻度で結合することが示唆された（図2-14上段）。遺伝子領域と遺伝子間領域に分けて解析した場合も、ES hypo T-DMRtag近傍領域ではクロマチン関連因子の結合頻度が有意に高かった（図2-14下段）。

次に、ES hypo T-DMRにおけるクロマチン関連因子の結合頻度と周辺領域のSINE配列密度に関係があるかを検討するために、クロマチン関連因子の結合をSINE配列密度で分けた区画ごとに解析した。ES hypo T-DMRtagへのクロマチン関連因子の結合の全体の様子は図2-15に示した。ES細胞特異的転写因子が結合するES hypo T-DMRtagのなかでhighest、highに属するものの割合はES hypo T-DMRtag全体とほぼ同じであった（図2-16）。また、ES hypo T-DMRtagにこれらの因子が結合する割合にも各区画の間で大きな差は認められなかった（図2-17）。つまり、ES細胞特異的転写因子の結合頻度と周辺領域のSINE配列密度との間に相関は認められなかった。一方、メディエーター因子および転写関連因子について解析したところ、これらの因子が結合するES hypo T-DMRtagは全体に比べ、highest、highに属するものの割合が5%から10%程度高かった（図2-16）。また、これらの因子はSINE配列密度の高い区画ほどES hypo T-DMRtagに高い割合で結合する傾向を示した（図2-17）。間期核ゲノムの高次構造に関する因子についても、メディエーター因子、転写関連因子と比較すると傾向は弱いものの、SINE密度の高い領域に存在するES hypo T-DMRへ結合の偏りを認めた（図2-17）。以上の傾向は、遺伝子領域・遺伝子間領域に存在するES hypo T-DMRとともに認められた（図2-15、2-18）。

ES hypo T-DMRtagのヒストン修飾について解析した。ヒストンH3K4のモノ・ジ・トリメチル化修飾はどれもES hypo T-DMRtagにはHpyサイトに比べ有意に高い頻度で起きており（図2-14）、さらにSINE配列密度の高い領域ほど修飾頻度が高い傾向があった（図2-17）。この傾向は遺伝子



領域・遺伝子間領域によらずに認められた（図2-15、2-18）。ヒストンH3K79のトリメチル化は、大部分が遺伝子領域のES hypo T-DMRに認められた（図2-15）。これは、H3K79のトリメチル化が転写の伸張反応が起きている領域にみられるという性質と一致している。修飾頻度とSINE配列密度の関係を調べたところ、ヒストンH3K79のジメチル化は解析したクロマチン関連因子・ヒストン修飾の中で最もhighestに属する割合が高く（図2-16）、修飾頻度がSINE配列密度の高い領域へ最も顕著に偏っていた（図2-17）。

一方、転写の不活性な領域に認められるヒストンH3K27のトリメチル化修飾は、ES hypo T-DMRtagに有意に高い頻度で認められたが、濃縮（enrichment）の程度が1.41倍と他の修飾に比べると低かった（図2-14上段）。ES hypo T-DMRのH3K27トリメチル化修飾の9割がH3K4のメチル化修飾とともに認められる、いわゆるbivalent修飾であった。bivalent修飾を受けていたES hypo T-DMRtagは、近傍の反復配列密度が非常に低い傾向にあった（図2-19上段）。この傾向は、bivalent修飾を含めH3K27トリメチル化修飾が起きている領域の近傍には反復配列が少ない（Bernstein et al., 2006）という報告と一致している。周囲50 kb領域のSINE配列密度を解析すると、bivalent修飾が認められたES hypo T-DMRtagはHpyサイトに比べ有意に密度が高いものの、ES hypo T-DMRtag全体の中では低い傾向にあった（図2-19下段）。また、H3K4、H3K79のメチル化修飾とは異なり、H3K27トリメチル化修飾にはSINE配列密度の高い領域に存在するES hypo T-DMRへの偏りはなかった（図2-17）。

以上より、ES hypo T-DMRにはクロマチン関連因子結合および活性型のヒストン修飾が有意に高い頻度で認められた。さらにES hypo T-DMRにおけるこれらの因子の結合および修飾は、ES細胞特異的な転写因子を除き、SINE配列の占める割合が高いゲノムドメインにおいてより高い頻度で起きていることが分かった。

### ES hypo T-DMRのES細胞分化に伴う核内位置の変化

ES細胞の分化に伴う、T-DMRの核内での位置の変化について解析した。核膜からの距離の指標として、核ラミンを用いたDamIDにより調べられた核ラミナとの相互作用を用いた。T-DMRtagおよびコントロールとしてHpyサイトの周辺前後6 kbにおけるDamIDシグナルの平均値を求め、ES細胞と分化細胞〔NP細胞、アストロサイト（AC細胞）、MEF細胞〕間の差を計算した。差の値を $\Delta$ Laminと呼ぶことにする。Hpyサイトの $\Delta$ Laminは0を中心とした分布を示した（図2-20）。これは、大部分のHpyサイト周辺では分化に伴う核内位置の変化が起きないことを意味しており、Peric-Hupkesらの結論と一致していることから今回のデータ計算の妥当性を示している。

ES細胞と分化細胞間でのT-DMRにおける核ラミナとの相互作用の差を解析した。まず、ES細胞とNPC細胞を比較した。Hpyサイトに比べES hypo T-DMRtagは、 $\Delta$ Lamin  $> 0$ 、ES hyper T-DMRtagは $\Delta$ Lamin  $< 0$ である傾向を示した（図2-21）。この傾向はES細胞とAC細胞、MEF細胞の比較でも認められた。これらの結果は、ES細胞に比べ分化細胞においては、ES hypo T-DMRはより核膜側に存在し、ES hyper T-DMRより核膜からはなれて存在することを示唆している。

次にES hypo T-DMRに着目し、周辺のSINE配列密度と核ラミナとの相互作用の変化の関係を調べた（図2-22）。周辺50 kb領域のSINE配列密度と $\Delta$ Laminの関係を散布図で示した。まずES細胞とNPC細胞を比較した結果、ES hypo T-DMRtagは、SINE配列密度が高い領域に存在するものほど $\Delta$ Laminが大きい傾向を示した。また、SINE配列密度で分けた区画ごとに統計解析をした結果、SINE配列密度が高い区画（highest、high）では、ES hypo T-DMRtagはHpyサイトに比べ $\Delta$ Laminが有意に大きいことが示された（図2-23）。この傾向は、MEF細胞との比較においても認められた（図2-23）。つまりSINE配列密度の高いゲノムドメインのなかにおいて、ES hypo T-DMRで特に $\Delta$ Lamin の上昇が大きいことが分かった。遺伝子領域と遺伝子間領域に存在するES hypo T-DMR

に分けて解析した結果、両方のグループでSINE配列密度が高いドメインに存在するものほど $\Delta$ Laminが大きい傾向があり（図2-22）、さらにHpyサイトとの比較で有意な差が認められた（図2-24）。従って、SINE配列密度と $\Delta$ Laminとの間に相関があるのは、遺伝子領域・遺伝子間領域によらず、SINE配列密度が高いドメインに存在するES hypo T-DMRの特徴である。

以上の結果に基づき、ES細胞の分化過程でのES hypo T-DMRの核内位置変化についての模式図を示す（図2-25）。ゲノム全体は分化にともない核膜との距離に大きな変化は示さない。しかしそのなかで、SINE配列密度の高いゲノムドメインに存在するES hypo T-DMRは、遺伝子領域・遺伝子間領域によらず分化に伴い核膜方向へ核内位置を移動することが示唆された。

## 考察

本章では、ES細胞特異的低メチル化T-DMR (ES hypo T-DMR) と染色体配列構造、核内構造との関係について新たな知見が得られた。染色体全領域におけるES細胞特異的メチル化修飾と反復配列の関係を明らかにすることを目的に、第6、8、16染色体全域を対象としたD-REAM解析を行った結果、ES hypo T-DMRには、周囲のSINE配列密度が高くLINE配列密度が低い傾向が認められた。染色体上での分布を調べたところ、ES hypo T-DMRの約6割がSINE配列密度の高い上位2割の領域に存在し、約8割がSINE配列密度の高い上位4割の領域に存在した。このES hypo T-DMRと周辺領域のSINE密度の関係は遺伝子領域・遺伝子間領域ともに認められたことから、ES hypo T-DMRは遺伝子領域・遺伝子間領域を問わず、SINE配列が豊富なゲノムドメインに集中していることが明らかになった。つまり、ゲノム上のES hypo T-DMRの分布は、SINE配列密度によって規定できる染色体の配列構造と相関することが示された。

SINE配列密度の高いドメインに着目した解析から、ES hypo T-DMRのなかでもSINE配列密度の高いドメインに存在するものほど転写に関連するクロマチン関連因子がよく結合し、転写が活発な領域にみられるヒストン修飾が認められた。つまりSINE配列が豊富な染色体領域にあるES hypo T-DMRは転写活性が高いクロマチン状態にあった。従来T-DMRの低メチル化は転写の活発なクロマチン状態と関係すると考えられていたが、本研究からこの関係には染色体配列構造も考慮する必要があることが示唆された。さらにSINE配列密度の高い領域のES hypo T-DMRは、分化に伴ってメチル化修飾を受けるだけでなく、核膜方向へと核内位置を移動することが示唆された。核膜は一般的に転写の不活性な環境であるので (Steensel & Dekker, 2010)、核膜方向への移動によってより強固な転写の不活性化が起きていると考えられる。ES hypo T-DMRのクロマチン高次構造の制御が染色体配列構造と相関することが示唆された。

以上、本章の研究により、ES hypo T-DMRのゲノム上の分布が染色体配列構造と関連すること、さらにES hypo T-DMRのクロモソームレベルおよび核内高次構造レベルのクロマチン構築(chromatin architecture)が染色体配列構造と関連することが明らかになった。ES細胞特異的な低メチル化は、クロマチン関連因子などの様々な因子や核内構造などの様々な要素と協調してゲノムワイドなエピジェネティック制御を行う。本章の研究は、ES細胞のエピゲノムを統一的に理解するために、染色体配列構造を考慮することが有効であることを示唆している。

これまでに、ゲノムワイドな細胞・組織ごとのメチル化修飾については、主に遺伝子の転写開始点近傍に焦点を当てた解析がなされ、数千の遺伝子について組織特異的なメチル化修飾と発現制御の関係が明らかにされてきた。一方で、T-DMRは遺伝子領域だけでなく、遺伝子領域外にも存在することが知られていたが、ゲノムワイドな解析はなされてこなかった。本章の研究では、遺伝子間領域に約850個のES hypo T-DMRtagを検出した。これらの遺伝子間領域ES hypo T-DMRには転写に関わるクロマチン関連因子の結合や、転写の活発な領域に認められるヒストン修飾が起きていたことから、エンハンサー活性を持ち転写の活性化に働いていることが示唆された。さらに、コヒーシン関連因子やCTCFという間期でクロマチンの高次構造に関わる因子も有意に高い頻度でES hypo T-DMRと共局在した。これまでに*Naong*遺伝子周辺域では、POU5F1 (Levasseur et al., 2008) やコヒーシン関連因子 (Kagey et al., 2010) を介して離れたゲノム領域どうしがES細胞において空間的な相互作用を行っていることが報告されている。核内高次構造を介したT-DMRの遺伝子発現制御により、より多くの遺伝子がT-DMRのメチル化修飾により制御を受ける可能性が考えられる。

本章の結果は、SINE配列密度の高いゲノムドメインの持つ新たな性質を明らかにした。ゲノム配列解読以降、トランスクリプトーム、エピゲノム解析から、染色体のゲノムドメインごとに異なるエピジェネティック制御がなされる傾向があることが分かってきた。ヒトのトランスクリプトーム解析から、SINE配列が多くGC含有量の高い染色体上のゲノムドメインには発現の高い遺伝子が多く、LINE配列が多くGC含有量の少ないゲノムドメインには発現が低い遺伝子が多いことが明らかにされた (Versteeg et al., 2003)。また、Thurmanらは主にHeLa細胞を対象とした解析から、Alu配列が多いゲノムドメイン中にはヒストンのアセチル化やH3K4のメチル化といった活性型のエピジェネティック修飾が多く、Alu配列が少ないゲノムドメインにはH3K27トリメチル化といった抑制型のエピジェネティック修飾が多いことを明らかにした (Thurman et al., 2007)。これらに対し本章では、ES細胞特異的に転写の活性化に働くES hypo T-DMRがSINE配列の多いゲノムドメインに集中して存在すること発見した。SINE配列が多いゲノムドメインには、様々な細胞で転写が活発な傾向がある一方で、ES細胞特異的に転写が活性化される遺伝子も多く存在するという特徴があることが明らかになった。

本章のメチル化解析から、タンパク質をコードする遺伝子以外に、miRNA や lincRNA を含むノンコーディング遺伝子にも T-DMR が存在することが明らかになった。これまでに、組織特異的なメチル化修飾が、転写調節領域のメチル化修飾を制御することで遺伝子の発現を直接制御する以外に、転写因子の発現を制御することで間接的に膨大な数の遺伝子の発現制御に関わることが示されていた (Sato et al., 2010)。miRNA や lincRNA は遺伝子発現制御やエピジェネティック制御に関わることが明らかになってきている (Khalil et al., 2009; Wang et al., 2010)。これらのノンコーディング RNA の発現が T-DMR のメチル化修飾により制御されることで、その標的遺伝子も間接的に制御されている可能性がある。本章の結果は、ノンコーディング RNA

を介したトランスクリプトーム制御という、細胞・組織特異的メチル化修飾の新たな機能を示唆している。

以上第二章では、染色体全域での T-DMR と SINE 配列の関係の解析から、ES 細胞特異的な低メチル化領域のゲノム上の分布およびそのエピジェネティック制御が、SINE 配列によって特徴付けられる染色体配列構造と相関することが明らかになった。

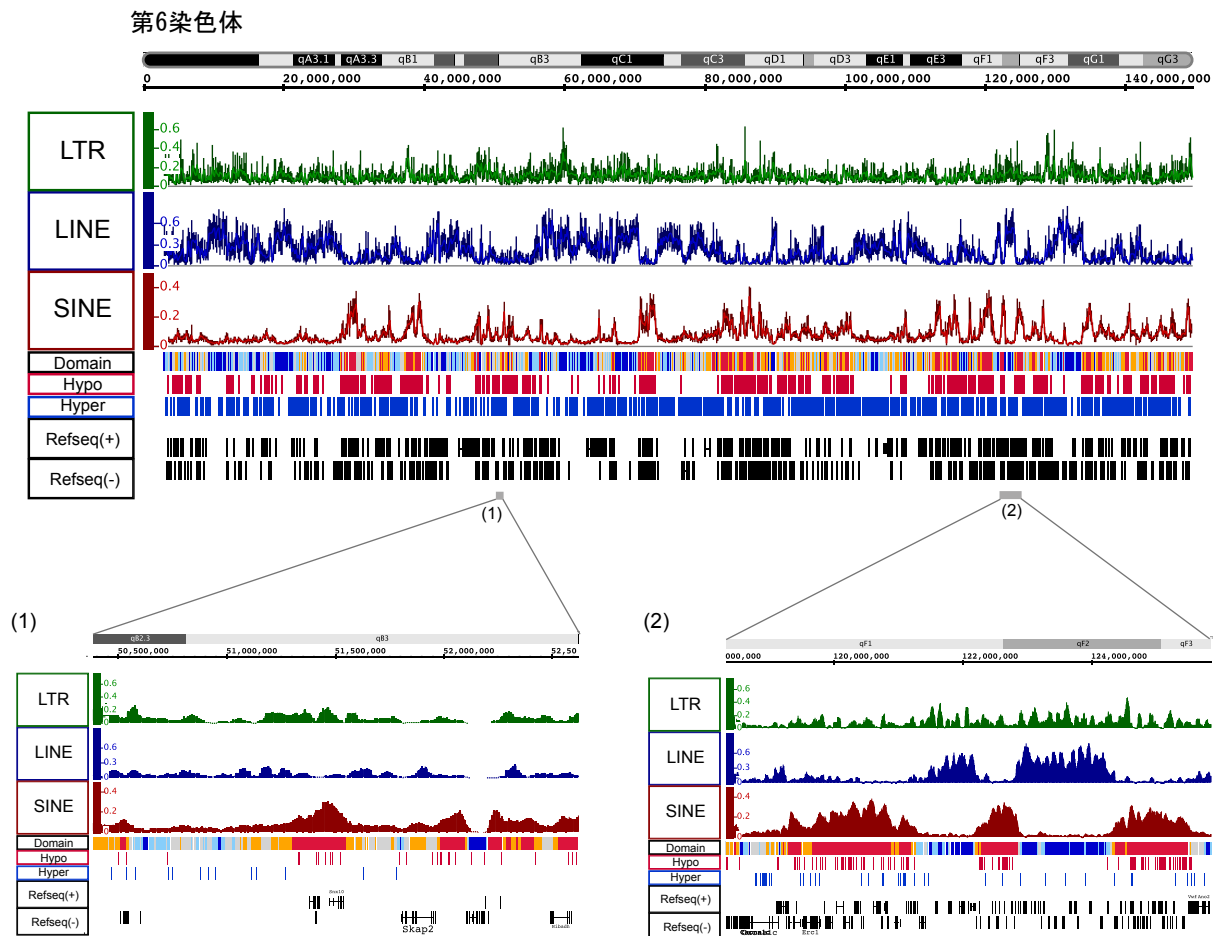


図 2-1 マウスフルタイリングアレイを用いた D-REAM 解析

第 6 染色体の D-REAM 解析結果を示した。上から、第 6 染色体の模式図、反復配列（LTR 配列、LINE 配列、SINE 配列）の密度、SINE 配列密度によって分類したドメインの種類 (Domain)、ES hypo T-DMRtag (Hypo)、ES hyper T-DMRtag (Hyper)、遺伝子 (Refseq) を示す。反復配列密度は 5 kb おきに前後 50 kb について計算した。ドメインごとの色は図 2-8 に基づく。遺伝子情報は Refseq データベースを参照した。上段の図の下に示した (1)、(2) の領域を、下段で拡大して表示した。ES hypo T-DMRtag および ES hyper T-DMRtag は遺伝子領域・遺伝子間領域ともに存在すること、ES hypo T-DMRtag は SINE 配列密度の高い領域に偏っていることが分かる。



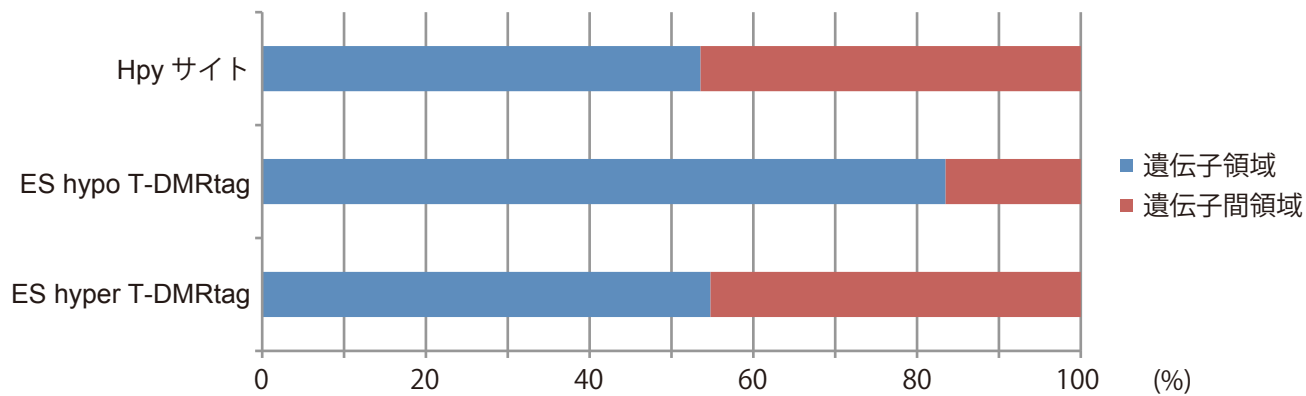


図 2-2 T-DMRtag の遺伝子領域・遺伝子間周辺への分布

Hpy サイト、ES hypo T-DMRtag、ES hyper T-DMRtag について、遺伝子周辺・遺伝子間領域に存在する割合を示した。

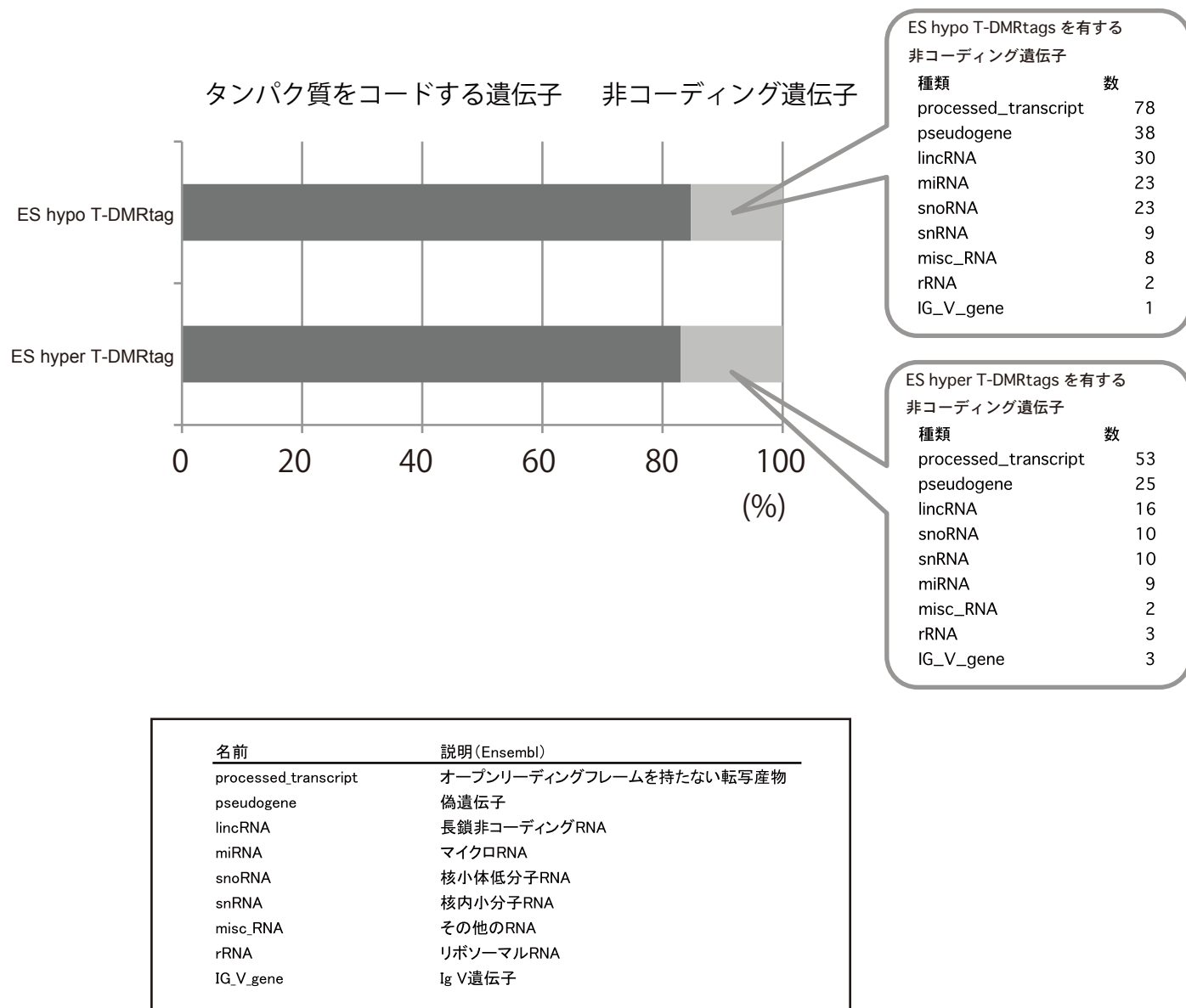
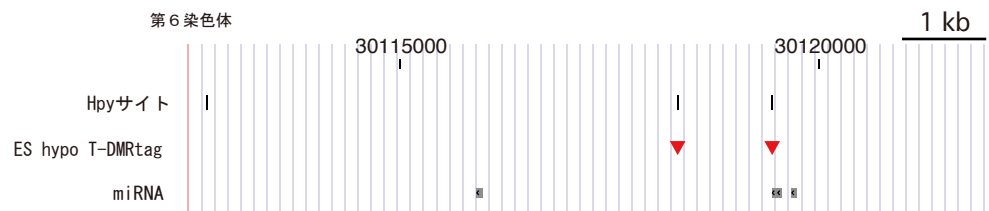


図 2-3 T-DMRtag が存在した遺伝子の種類

タンパク質をコードする遺伝子と非コーディング遺伝子の割合を示した。非コーディング遺伝子については、種類を示した。遺伝子の種類は Ensembl の遺伝子情報に基づいて分類した。下段に種類の説明を示した。

#### クラスターとして存在するmiRNA遺伝子の例



#### イントロンに存在するmiRNA遺伝子の例

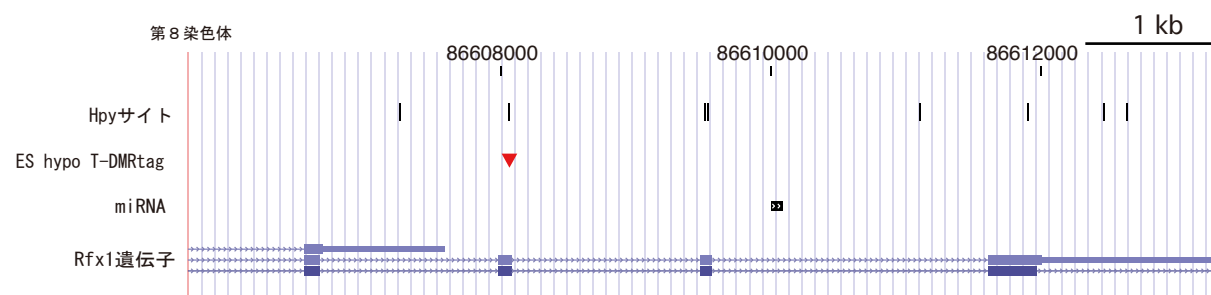
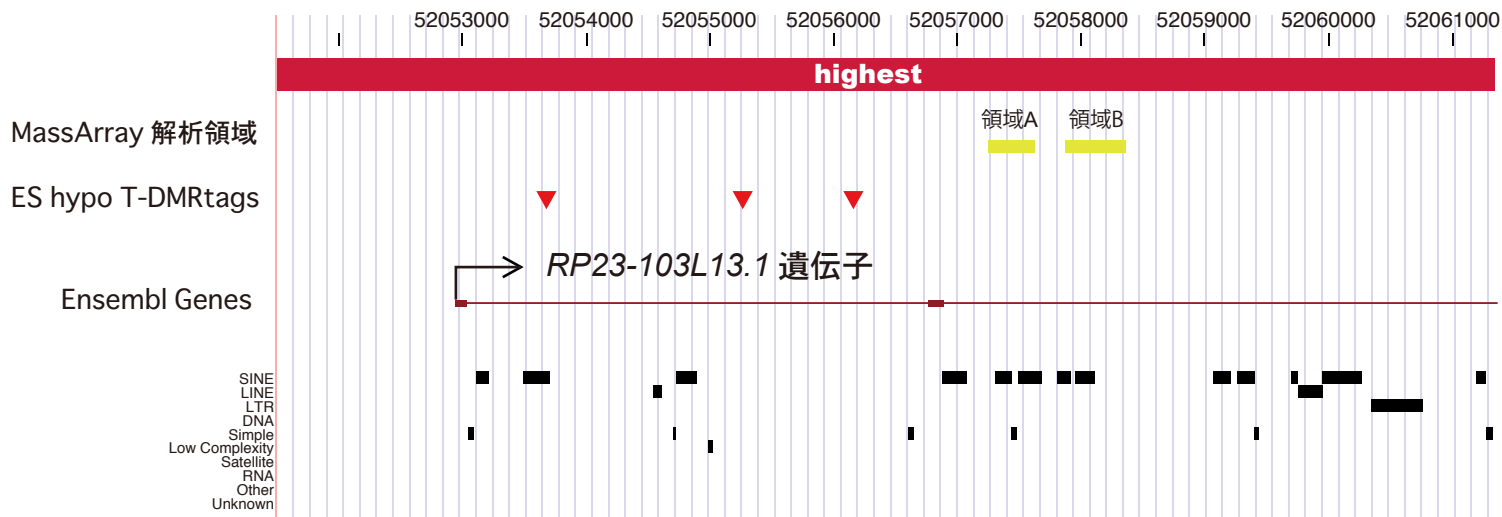


図 2-4 miRNA 遺伝子間領域に存在した ES hypo T-DMR の例

上から、染色体上の場所情報、Hpy サイト、ES hypo T-DMRtag、miRNA 遺伝子の場所を示した。

下段ではさらに *Rfx1* 遺伝子を示した。

非コーディング遺伝子の例



MassArray 解析

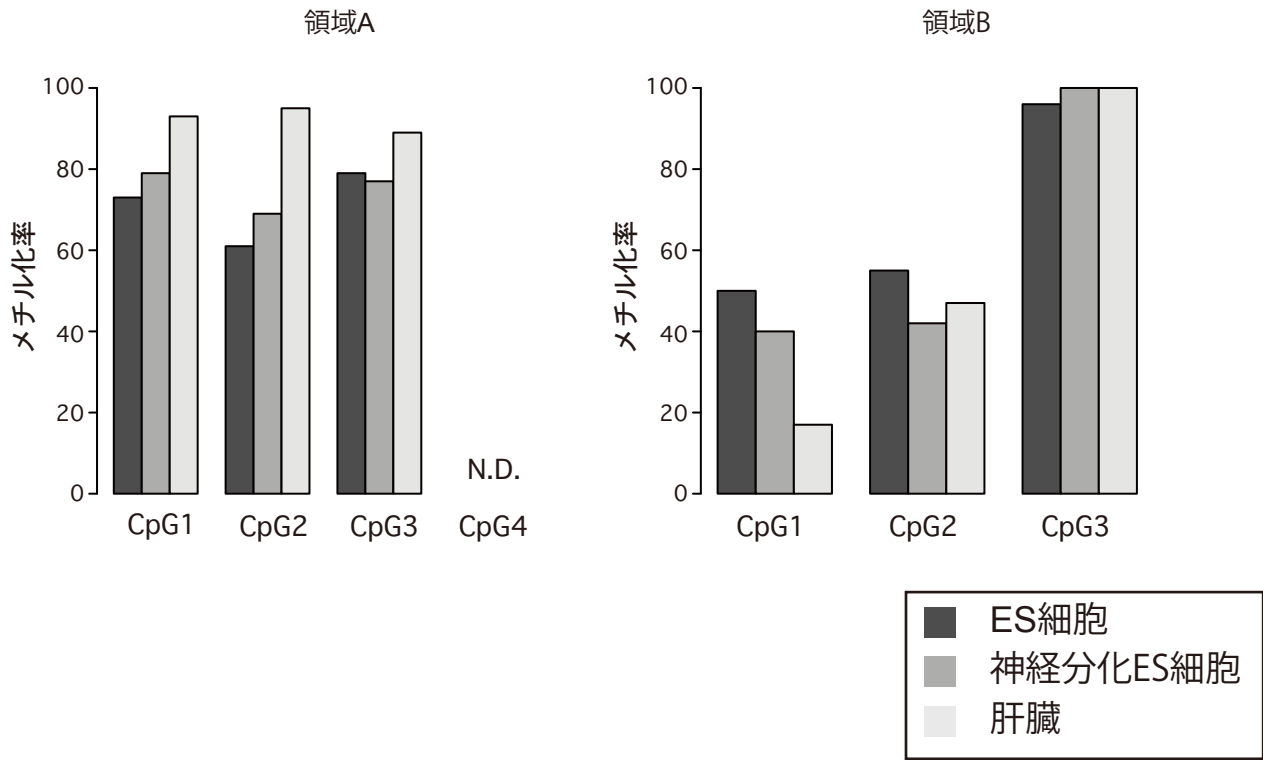
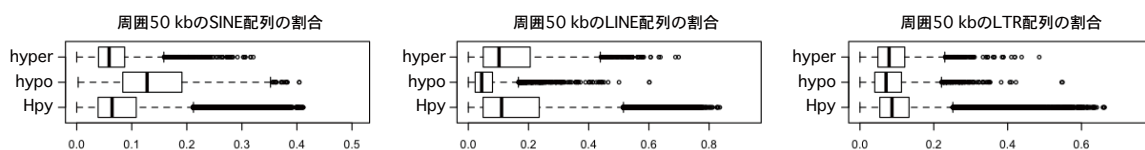


図 2-5 長鎖ノンコーディング RNA 遺伝子領域に存在した ES hypo T-DMR の例

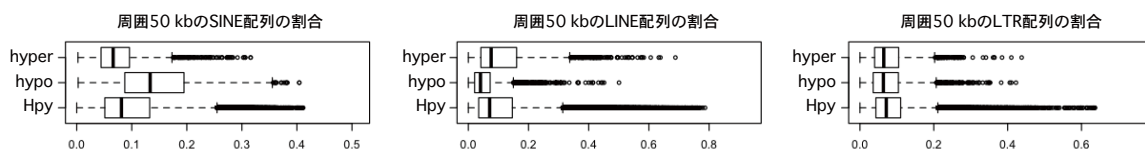
(上段) 上から、染色体上の場所情報、SINE 配列密度によって分類したドメインの種類、マスアレイの解析領域 (領域 A、B)、遺伝子の位置、反復配列の分布を示している。

(下段) マスアレイ解析の結果を示す。縦軸はメチル化率 (%) を示す。データが取得できなかった場合は、N. D. と示した。

## 全体



## 遺伝子領域



## 遺伝子間領域

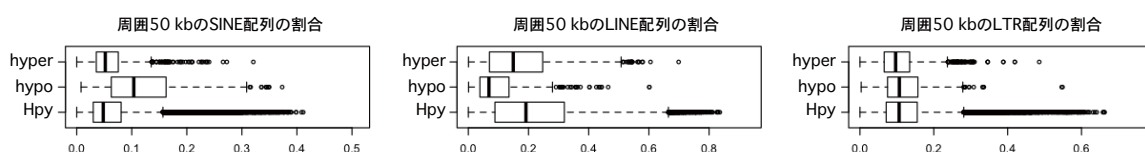
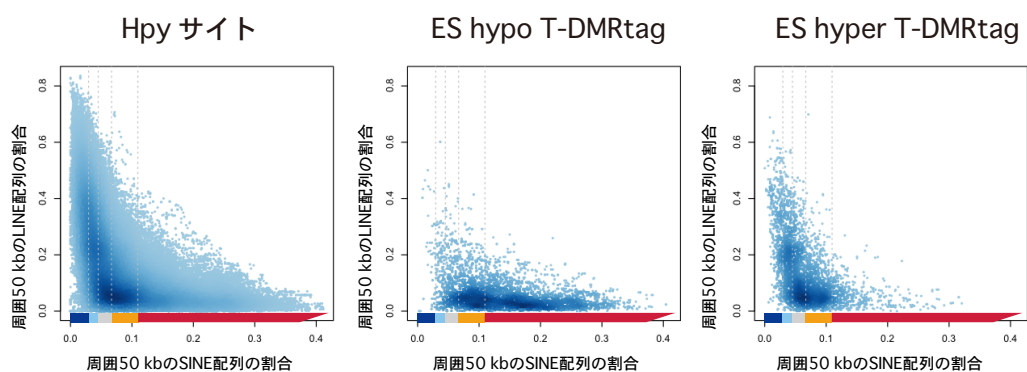


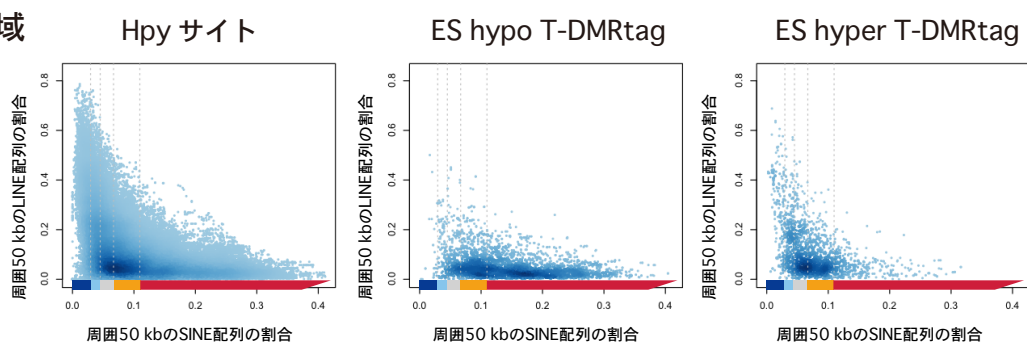
図 2-6 Hpy サイト、ES hypo T-DMRtag、ES hyper T-DMRtag 周囲 50 kb 領域に占める反復配列の割合の箱ひげ図

遺伝子領域・遺伝子間領域をあわせた結果を上段に、遺伝子領域と遺伝子間領域に分けて図示した結果を中段、下段に示している。箱の下部は第 1 四分位点であり上部は第 3 四分位点を表す。箱中の横線は中央値である。箱からのびた点線の先の横棒は、第 1 四分位点から IQR (interquartile range、第 3 四分位点 - 第 1 四分位点) の 1.5 倍の範囲で最も値の低いデータ、および第 3 四分位点から IQR の 1.5 倍の範囲で最も値の高いデータを表す。この範囲外のデータは、中空円で示す。Hpy、Hpy サイト; hypo ES hypo T-DMRtag; hyper、ES hyper T-DMRtag

全体



遺伝子領域



遺伝子間領域

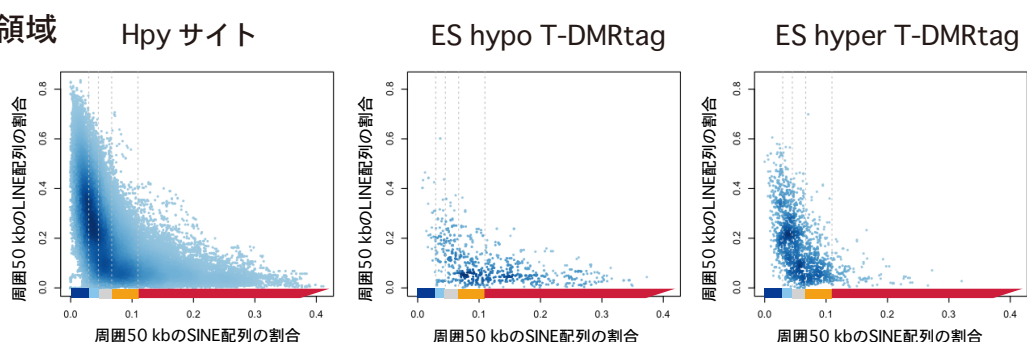


図 2-7 SINE 配列と LINE 配列が占める割合を表した散布図

Hpy サイト、ES hypo T-DMRtag、ES hyper T-DMRtag の周囲 50 kb 領域に占める反復配列の割合を計算した。X 軸に SINE 配列の占める割合、Y 軸に LINE 配列の占める割合を表した。X 軸には、SINE 配列密度による分類を色で示した。ドメインごとの色は図 2-8 に基づく。遺伝子領域、遺伝子間領域に分けて図示した結果を中段、下段に示している。

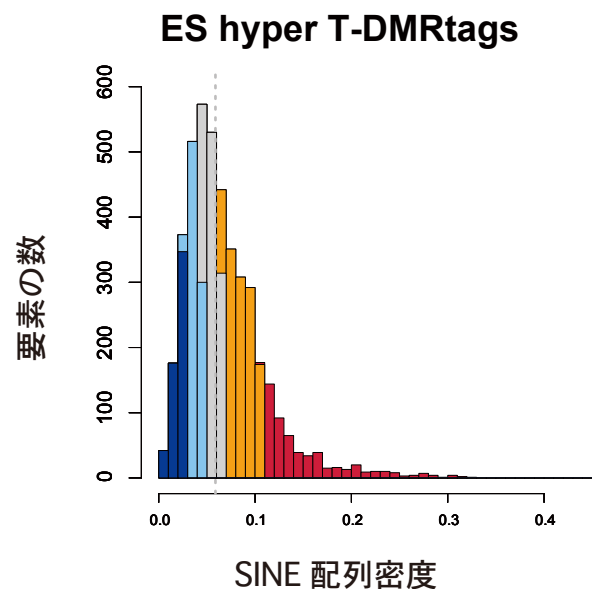
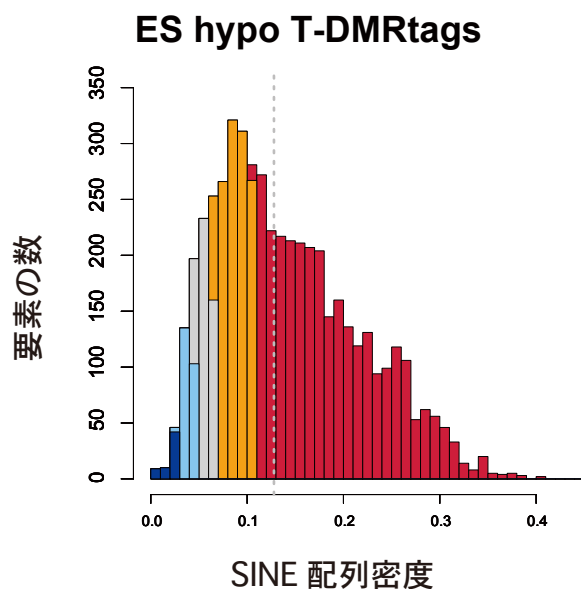
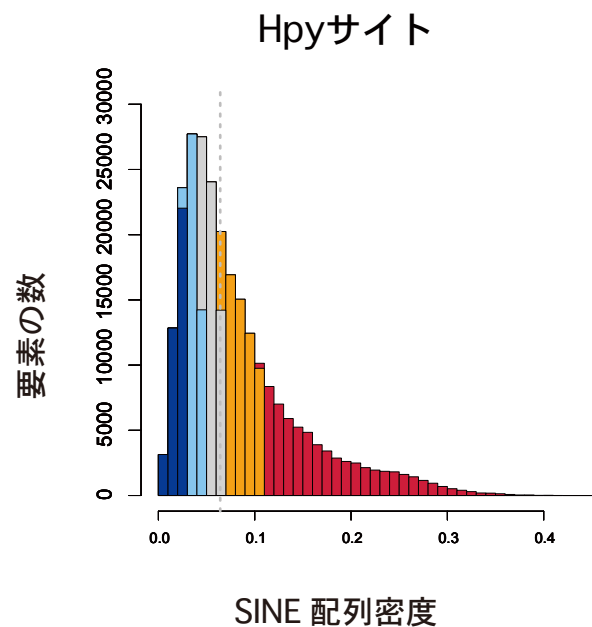
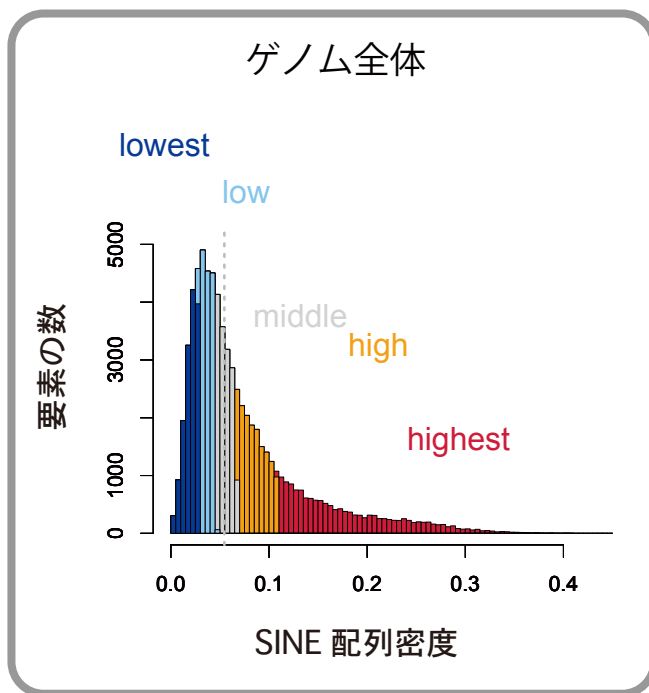


図 2-8 SINE 配列密度を指標としたゲノムの分類

各ヒストグラムは、X 軸が SINE 配列の占める割合、Y 軸が要素の数を表す。縦点線は SINE 配列の占める割合の中央値を示す。左上には、ゲノム上を 5 kb おきに前後 50 kb 領域に占める SINE 配列の割合を計算し、割合の高いものから 20%ずつ highest (赤)、high (オレンジ)、middle (灰色)、low (水色)、lowest (青) の 5 つの区分に分類した結果を示した。Hpy サイト、ES hypo T-DMRtag、ES hyper T-DMRtag についても、5 つの区分に分類した結果を色で示した。

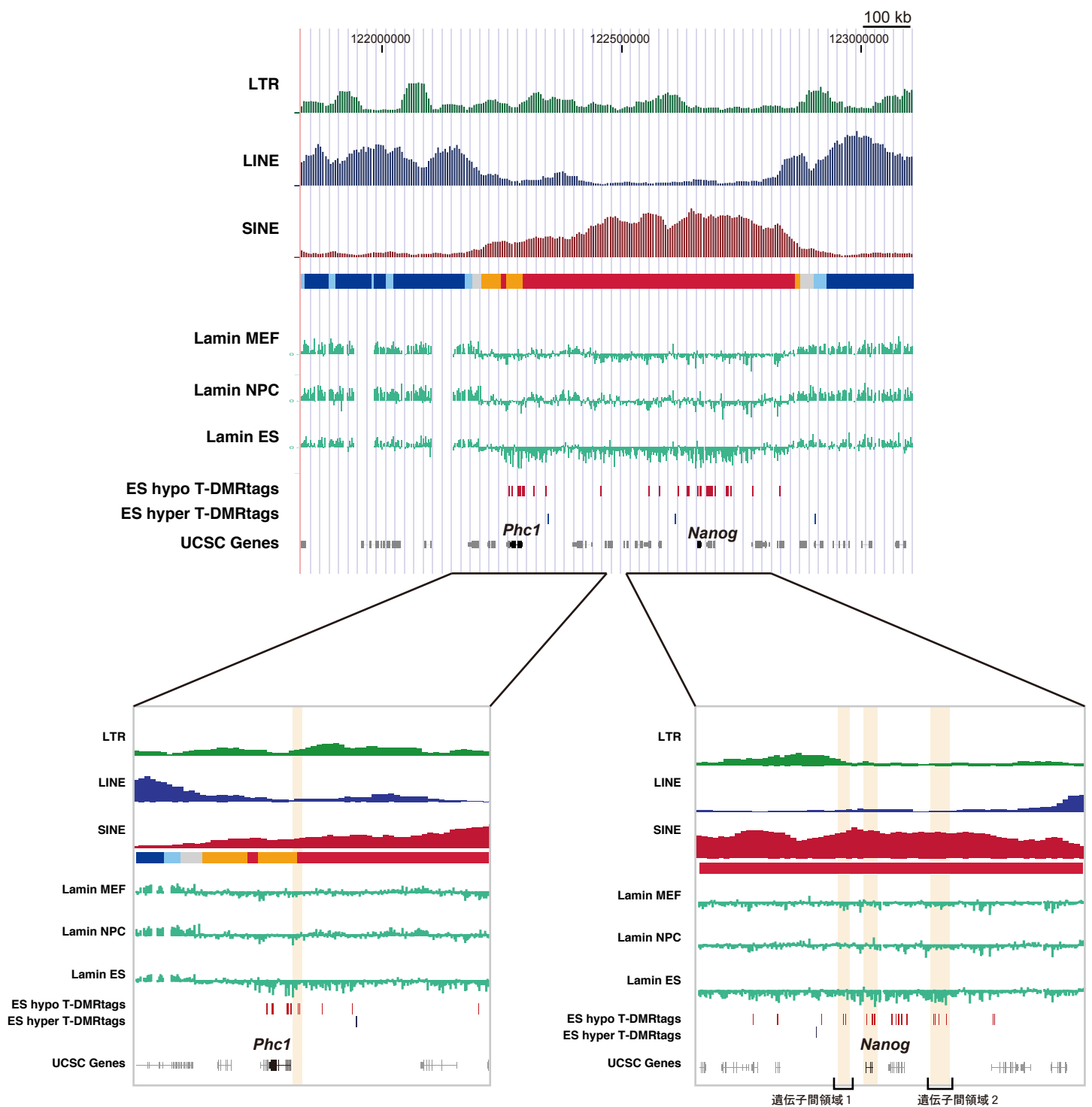


図 2-9 *Naong* 遺伝子が存在する SINE 配列密度の高いゲノムドメインの解析

(上段) 上から、反復配列の密度、SINE 配列密度によって分類したドメインの種類、Lamin を用いた DamID のシグナル (MEF 細胞、NPC 細胞、ES 細胞)、T-DMRtag の位置、遺伝子の位置を示す。ドメインごとの色は図 2-8 に基づく。DamID のシグナルは、値が高いほど核膜との相互作用頻度が高いことを意味する。

(下段) *Phc1* 遺伝子、*Nanog* 遺伝子を含む領域を拡大して示した。図 2-12、2-13 で解析した遺伝子間領域の場所を、遺伝子間領域 1、2 として示した。



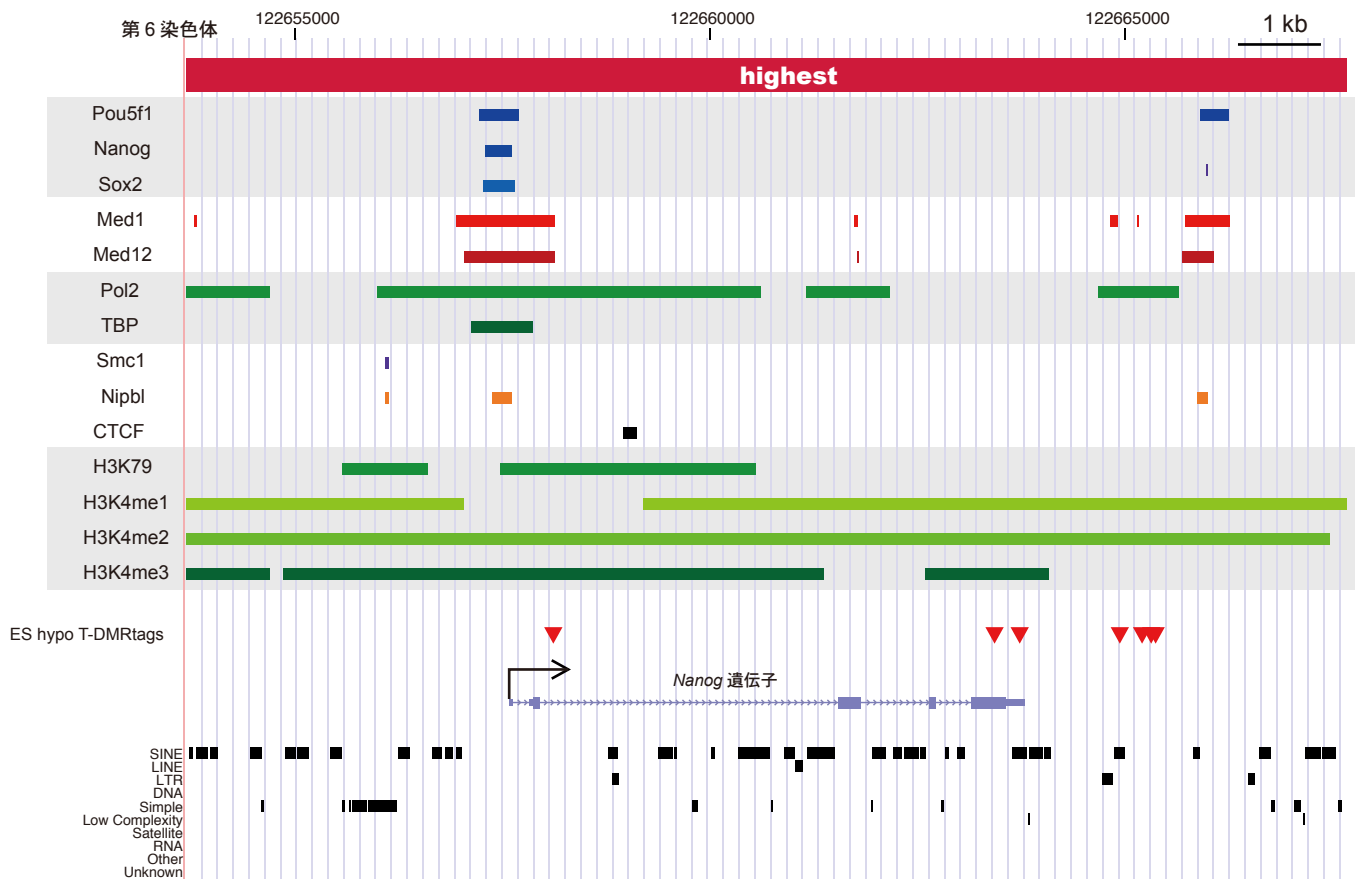
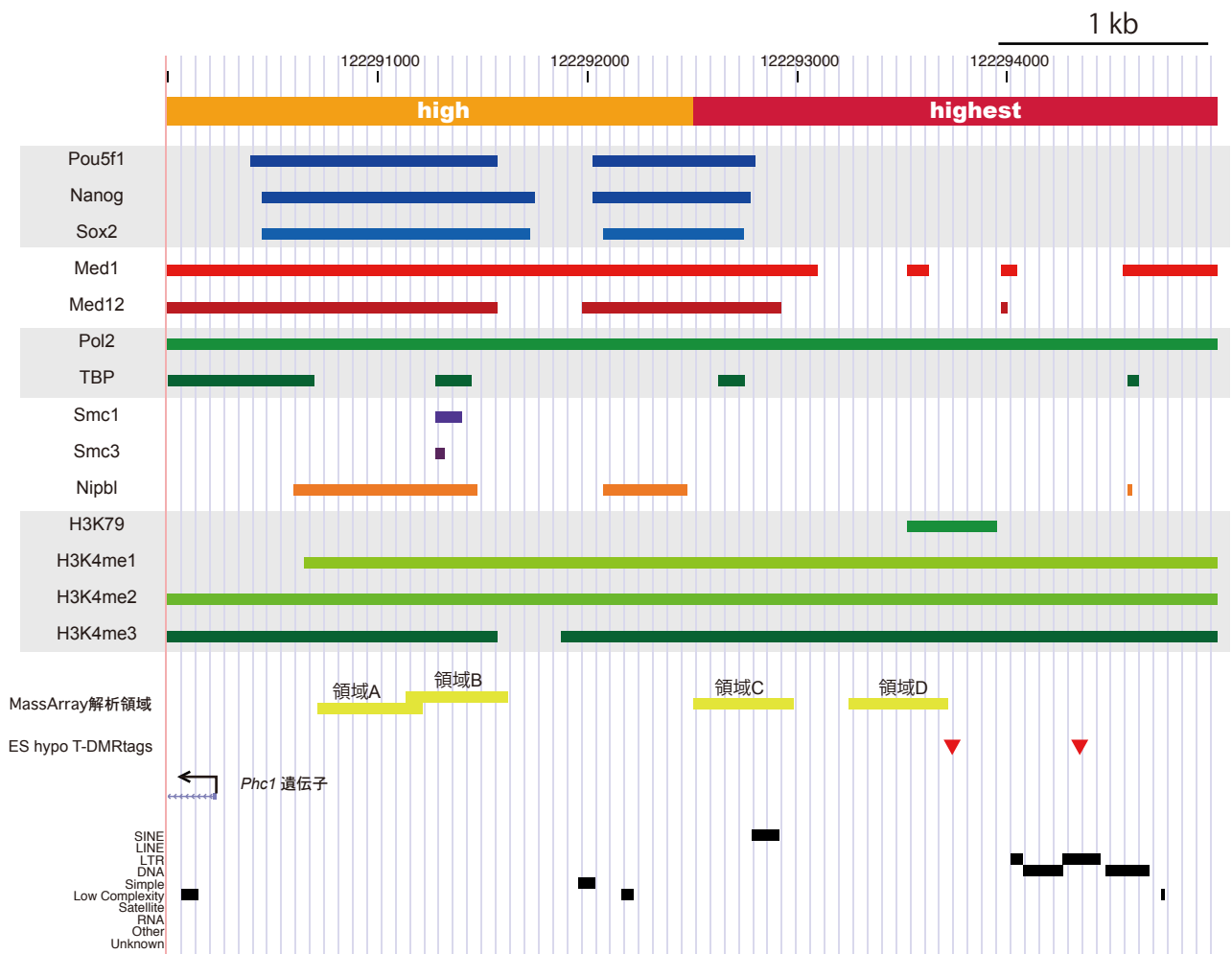


図 2-10 *Nanog* 遺伝子領域における ES hypo T-DMR とクロマチン関連因子の結合およびヒストン修飾の関係

上から、染色体上の場所情報、クロマチン関連因子の結合、ヒストン修飾、ES hypo T-DMRtag の位置、遺伝子の位置、反復配列の分布を示している。



## MassArray 解析

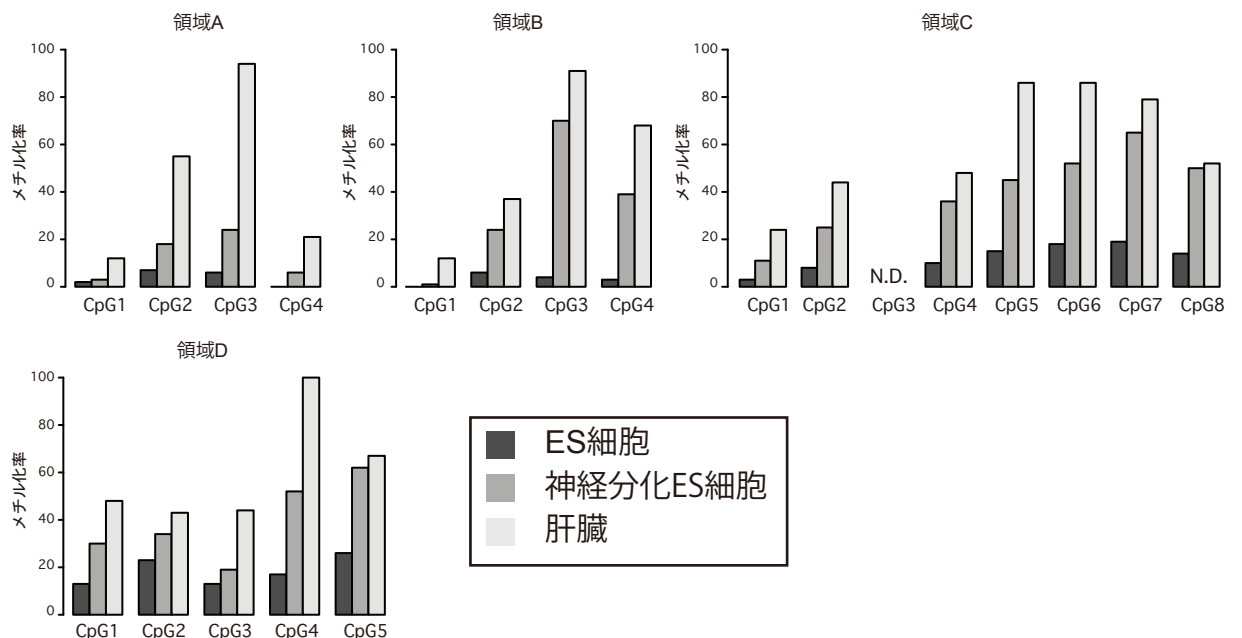
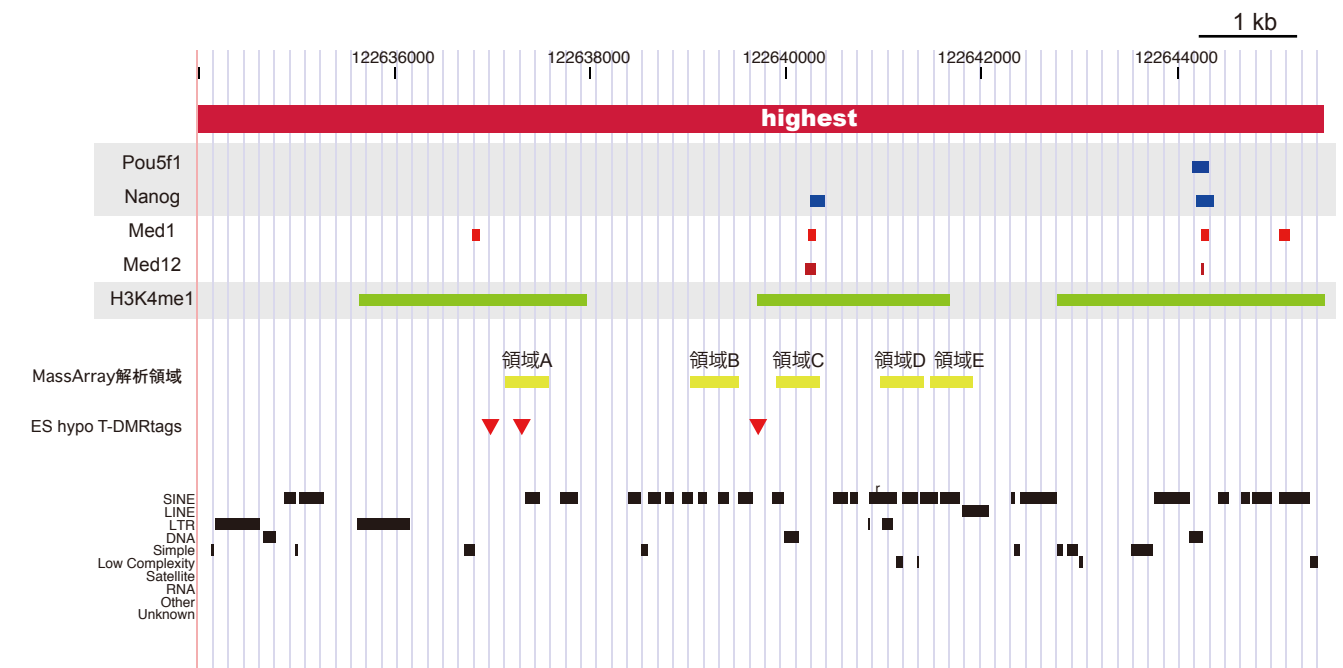


図 2-11 *Phc1* 遺伝子領域における ES hypo T-DMR とクロマチン関連因子の結合およびヒストン修飾の関係

上から、染色体上の場所情報、クロマチン関連因子の結合、ヒストン修飾、マスアレイの解析領域、ES hypo T-DMRtag の位置、遺伝子の位置、反復配列の分布を示している。その下にマスアレイ解析の結果を示す。縦軸はメチル化率（%）を示す。データが取得できなかった場合は、N. D. と示した。



## MassArray 解析

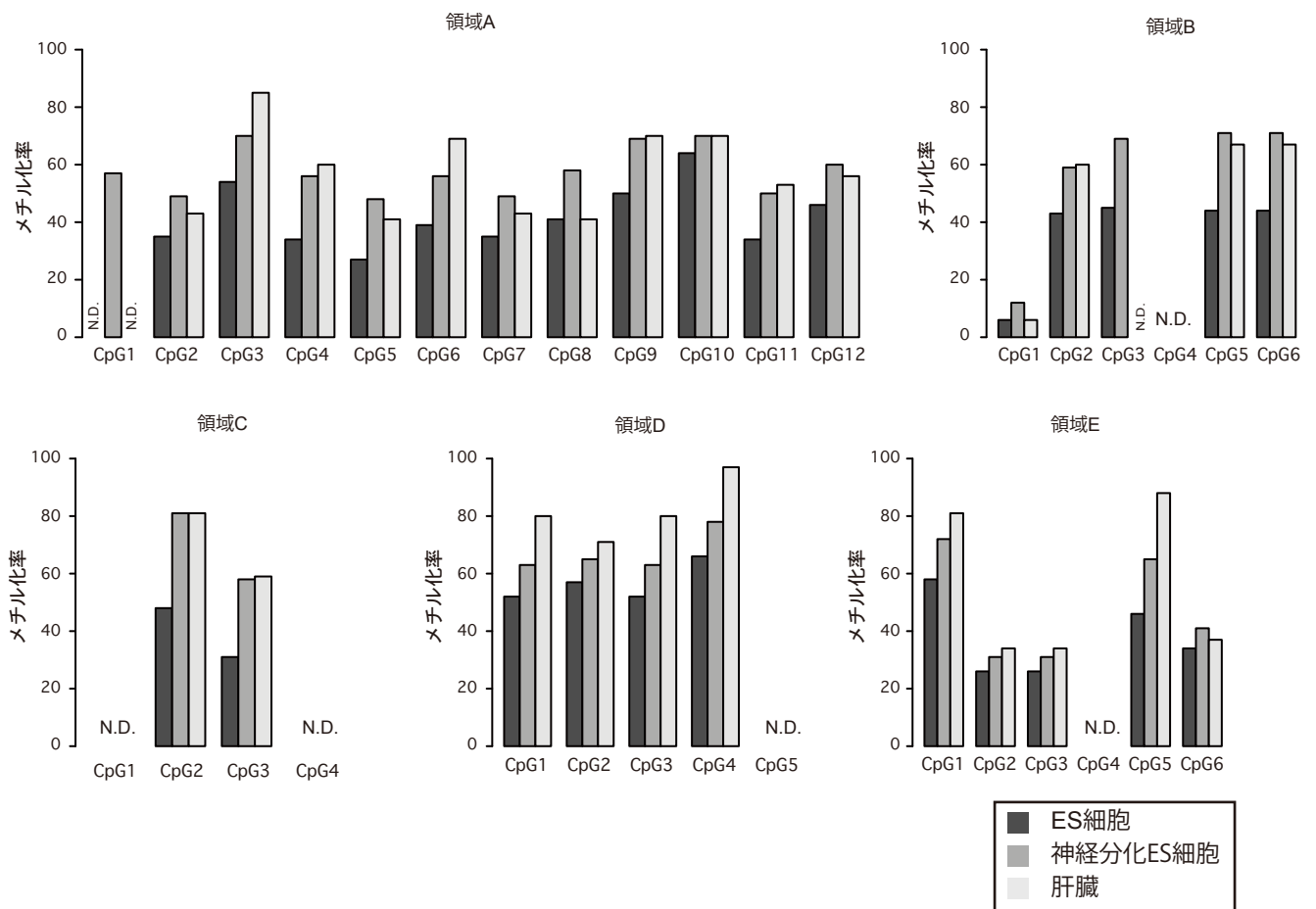
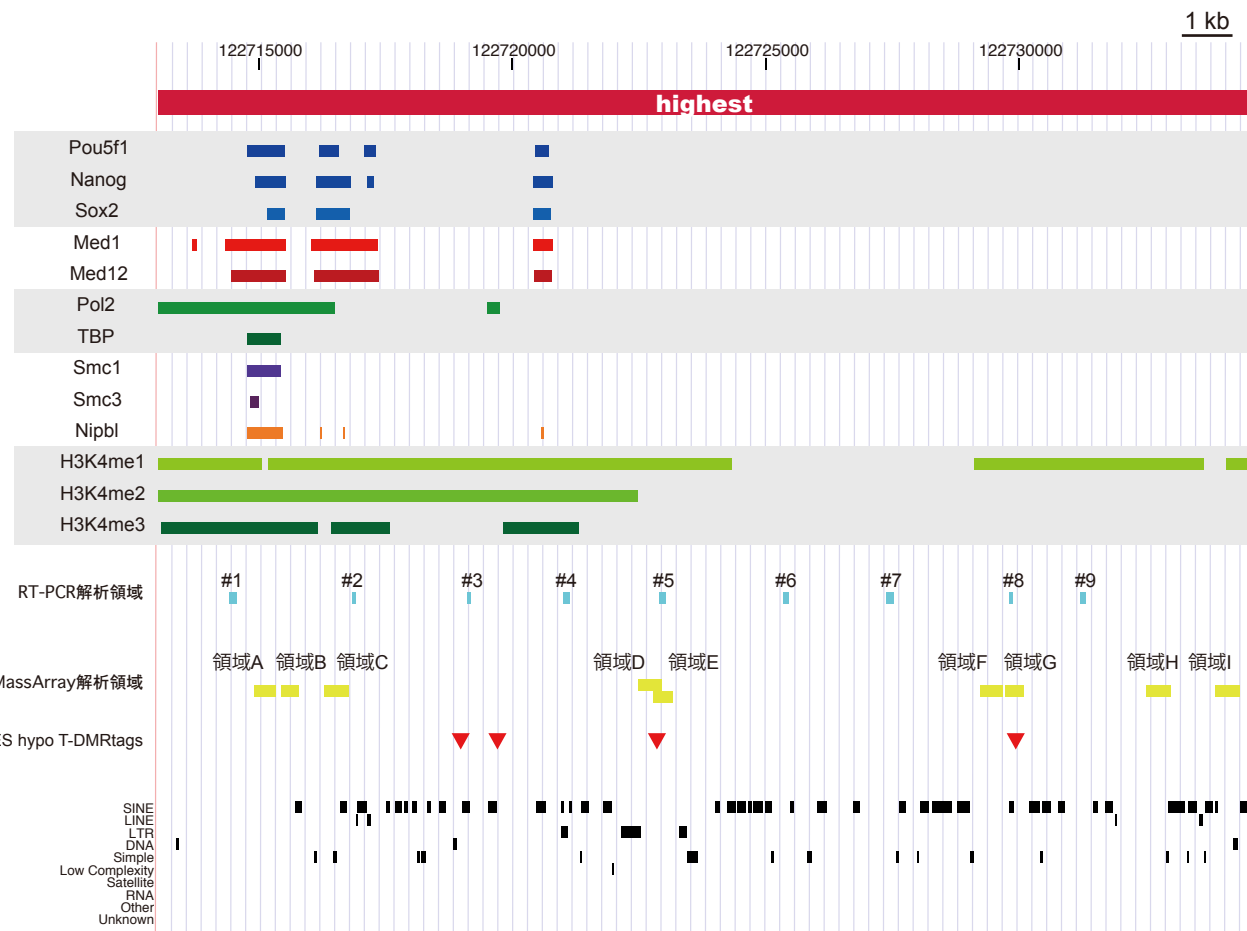


図 2-12 遺伝子間領域 1 における ES hypo T-DMR とクロマチン関連因子の結合およびヒストン修飾の関係

上から、染色体上の場所情報、クロマチン関連因子の結合、ヒストン修飾、マスアレイの解析領域、ES hypo T-DMRtag の位置、遺伝子の位置、反復配列の分布を示している。その下にマスアレイ解析の結果を示す。縦軸はメチル化率 (%) を示す。データが取得できなかった場合は、N. D. と示した。



## MassArray 解析

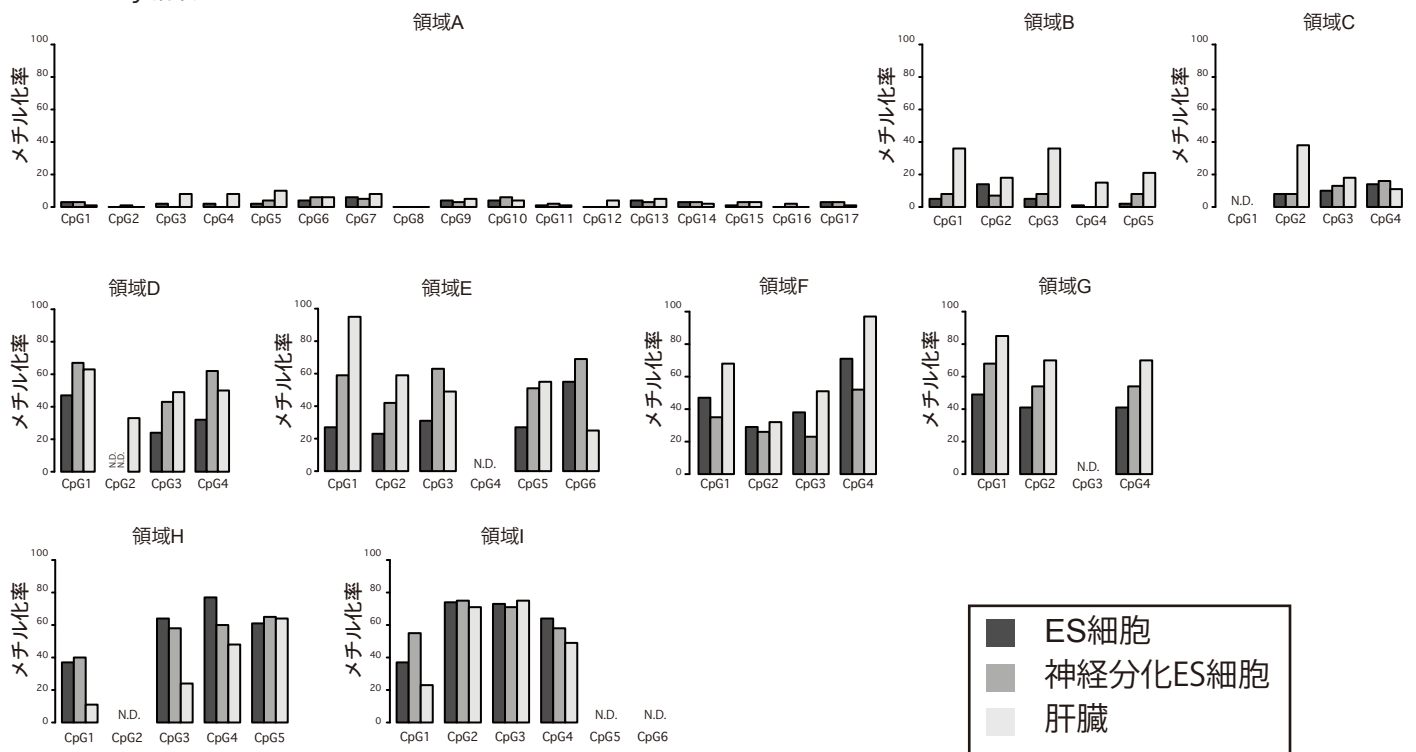




図 2-13 遺伝子間領域 2 における ES hypo T-DMR とクロマチン関連因子の結合およびヒストン修飾の関係

前ページの上から、染色体上の場所情報、クロマチン関連因子の結合、ヒストン修飾、RT-PCR の解析領域、マスアレイの解析領域、ES hypo T-DMRtag の位置、遺伝子の位置、反復配列の分布を示している。その下にマスアレイ解析の結果を示す。縦軸はメチル化率 (%) を示す。データが取得できなかった場合は、N. D. と示した。また本ページに RT-PCR の結果を示した。#1 から #9 は遺伝子間領域 2 における解析領域の番号を示す。(+) は逆転写反応を行った鋳型を、(-) は逆転写反応を行わなかった鋳型を用いて PCR を行ったことを意味する。ES 細胞は NIH/3T3 細胞と同じ cDNA 濃度の鋳型に加え、4 倍、16 倍希釈した cDNA を鋳型とした PCR を行った。B-actin に関しては、NIH/3T3 細胞も 4 倍、16 倍希釈した cDNA を鋳型とした PCR を行った。

		ES hypo T-DMRtags		ES hyper T-DMRtags	
		enrichment	P-value	enrichment	P-value
クロマチン関連因子	Pou5f1	3.92	< 2.2E-16	0.86	0.02976
	Nanog	3.62	< 2.2E-16	0.98	0.79
	Sox2	3.93	< 2.2E-16	0.91	0.1711
	Med1	4.19	< 2.2E-16	0.41	< 2.2E-16
	Med12	4.17	< 2.2E-16	0.53	< 2.2E-16
	Smc1	2.87	< 2.2E-16	0.52	< 2.2E-16
	Smc3	2.27	< 2.2E-16	0.49	< 2.2E-16
	Nipbl	4.38	< 2.2E-16	0.46	9.669E-16
	CTCF	2.26	< 2.2E-16	0.56	< 2.2E-16
	Pol2	3.85	< 2.2E-16	0.40	< 2.2E-16
	TBP	3.84	< 2.2E-16	0.44	4.415E-15
ヒストン修飾	H3K79	4.52	< 2.2E-16	0.61	1.447E-10
	H3K4me1	3.90	< 2.2E-16	0.60	< 2.2E-16
	H3K4me2	3.53	< 2.2E-16	0.57	< 2.2E-16
	H3K4me3	3.47	< 2.2E-16	0.49	< 2.2E-16
	H3K27me3	1.41	0.0000019	0.68	0.0008027

## 遺伝子領域

		ES hypo T-DMRtags		ES hyper T-DMRtags	
		enrichment	P-value	enrichment	P-value
クロマチン関連因子	Pou5f1	3.55	< 2.2E-16	0.74	0.001078
	Nanog	3.30	< 2.2E-16	0.81	0.02364
	Sox2	3.62	< 2.2E-16	0.77	0.00447
	Med1	3.74	< 2.2E-16	0.40	< 2.2E-16
	Med12	3.78	< 2.2E-16	0.44	< 2.2E-16
	Smc1	2.73	< 2.2E-16	0.46	< 2.2E-16
	Smc3	2.03	< 2.2E-16	0.46	< 2.2E-16
	Nipbl	3.59	< 2.2E-16	0.37	< 2.2E-16
	CTCF	2.12	< 2.2E-16	0.55	< 2.2E-16
	Pol2	2.80	< 2.2E-16	0.36	< 2.2E-16
	TBP	2.75	< 2.2E-16	0.40	2.666E-15
ヒストン修飾	H3K79	3.95	< 2.2E-16	0.56	2.883E-13
	H3K4me1	5.11	< 2.2E-16	0.53	< 2.2E-16
	H3K4me2	3.83	< 2.2E-16	0.50	< 2.2E-16
	H3K4me3	3.27	< 2.2E-16	0.44	< 2.2E-16
	H3K27me3	0.93	0.3376	0.67	0.001261

## 遺伝子間領域

		ES hypo T-DMRtags		ES hyper T-DMRtags	
		enrichment	P-value	enrichment	P-value
クロマチン関連因子	Pou5f1	9.38	< 2.2E-16	1.07	0.5071
	Nanog	8.60	< 2.2E-16	1.26	0.01752
	Sox2	9.74	< 2.2E-16	1.15	0.1695
	Med1	12.83	< 2.2E-16	0.42	8.016E-07
	Med12	11.80	< 2.2E-16	0.77	0.05295
	Smc1	7.09	< 2.2E-16	0.64	0.00001727
	Smc3	3.88	< 2.2E-16	0.52	5.59E-08
	Nipbl	11.65	< 2.2E-16	0.77	0.1104
	CTCF	3.81	< 2.2E-16	0.55	1.392E-08
	Pol2	12.24	< 2.2E-16	0.57	0.0286
	TBP	13.23	< 2.2E-16	0.60	0.04277
ヒストン修飾	H3K79	7.50	< 2.2E-16	1.81	0.0459
	H3K4me1	13.34	< 2.2E-16	0.76	0.003924
	H3K4me2	12.06	< 2.2E-16	0.73	0.00419
	H3K4me3	9.45	< 2.2E-16	0.64	0.002045
	H3K27me3	3.34	1.594E-08	0.63	0.1561

図 2-14 T-DMR 周辺領域におけるクロマチン関連因子の結合とヒストン修飾

(A) ES hypo T-DMRtag、ES hyper T-DMRtag 周辺 6 kb 領域におけるクロマチン関連因子の結合とヒストン修飾。enrichment は Hpy サイトに対する ES hypo T-DMRtag または ES hyper T-DMRtag の、クロマチン関連因子の結合およびヒストン修飾が周辺領域に認められた頻度の相対値。P-value はカイ自乗検定の結果を示す。

(B) 遺伝子領域（上段）、遺伝子間領域（下段）における ES hypo T-DMRtag、ES hyper T-DMRtag 周辺 6 kb 領域におけるクロマチン関連因子の結合とヒストン修飾。

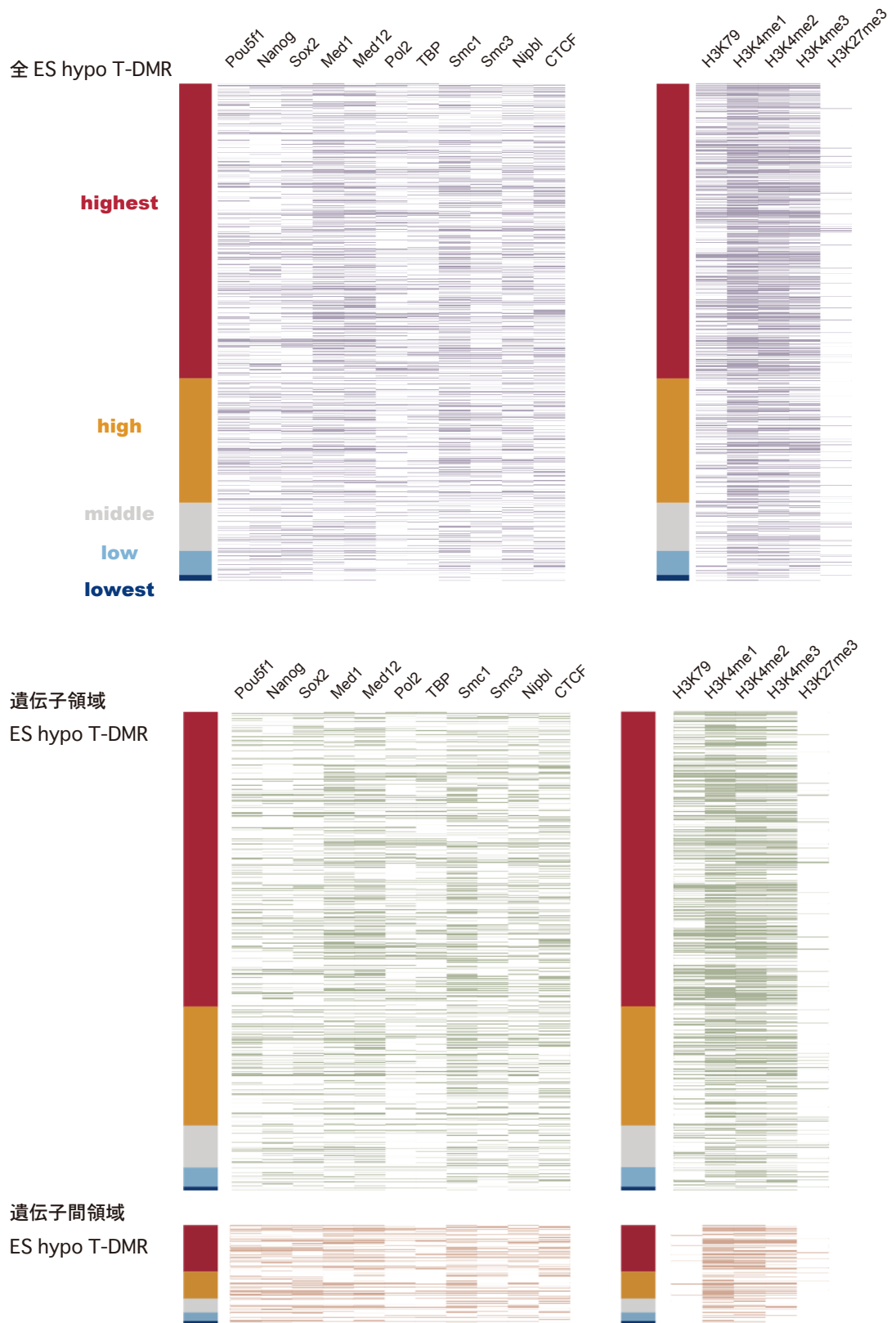
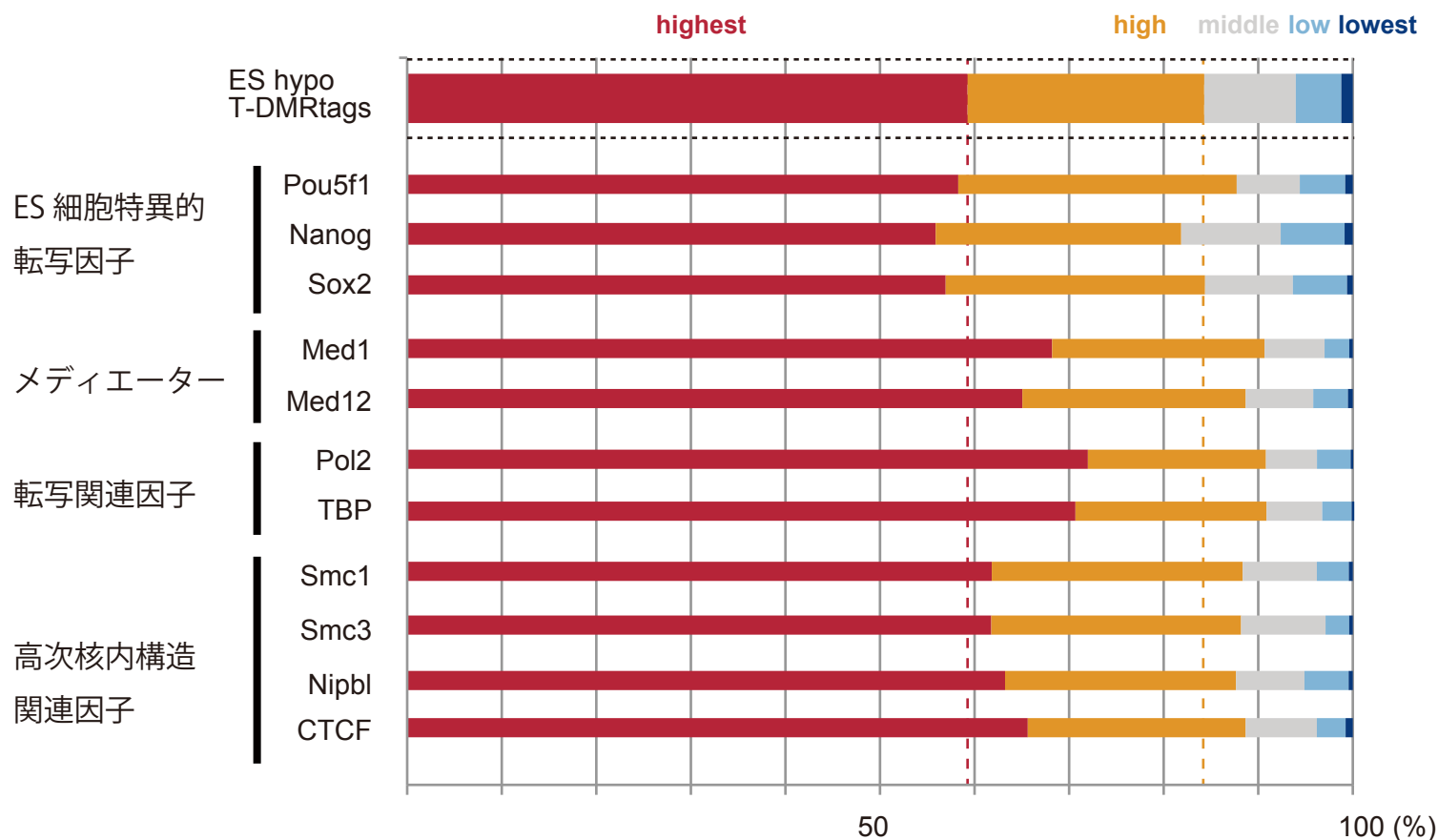


図 2-15 ES hypo T-DMR におけるクロマチン関連因子の結合・ヒストン修飾のヒートマップ  
 全ての ES hypo T-DMRtag (上段)、および遺伝子領域・遺伝子間領域に存在する ES hypo T-DMRtag  
 (中段、下段) について、クロマチン関連因子の結合・ヒストン修飾状況を示した。各 ES hypo  
 T-DMR が周囲の SINE 配列密度の高いものから縦に並べられており、クロマチン関連因子の結合・  
 ヒストン修飾が認められた場合は色付きの線で示してある。SINE 配列密度は周辺 50 kb 領域に  
 おける SINE 配列の占める割合を調べた。クロマチン関連因子の結合・ヒストン修飾は ES hypo  
 T-DMRtag の周囲 6 kb 領域を解析した。各ヒートマップの左側に示したドメインごとの色は図  
 2-8 に基づく。

## クロマチン関連因子



## ヒストン修飾

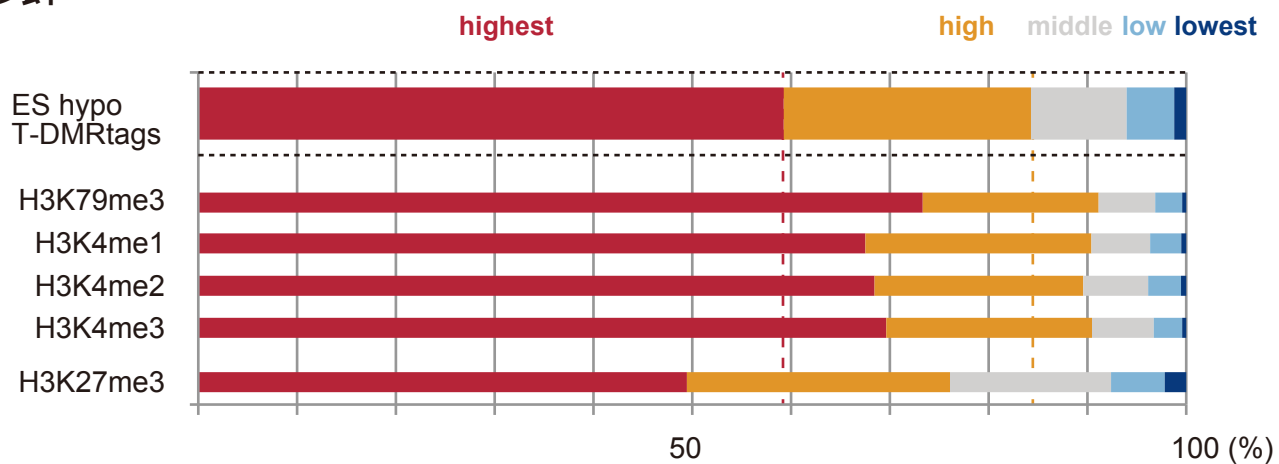
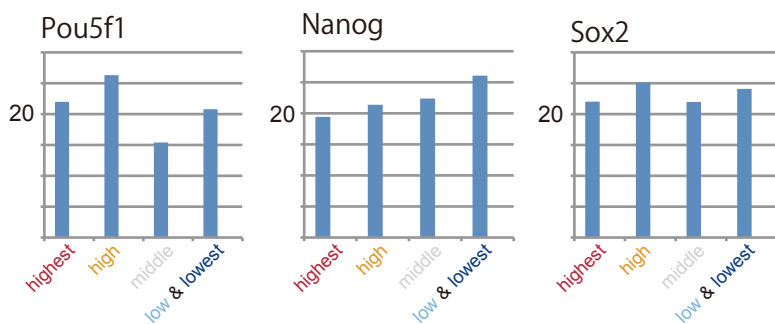


図 2-16 ES hypo T-DMR におけるクロマチン関連因子の結合およびヒストン修飾と周囲の SINE 配列密度の関係

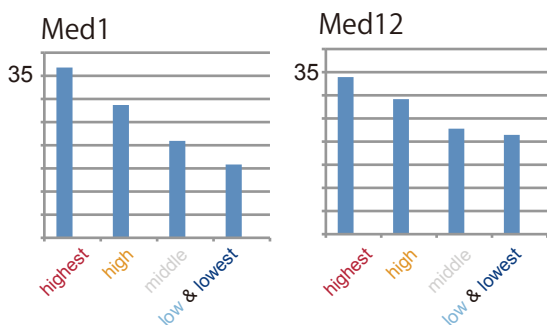
クロマチン関連因子の結合（上段）およびヒストン修飾（下段）の認められた ES hypo T-DMRtag が、SINE 配列密度によって分類した各ドメインに存在する割合を示した。比較対象として、ES hypo T-DMRtag のドメインごとの割合を示した。



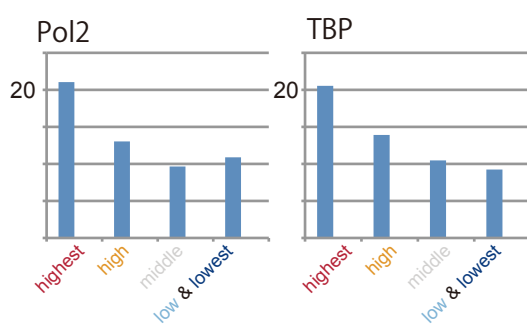
## ES 細胞特異的転写関連因子



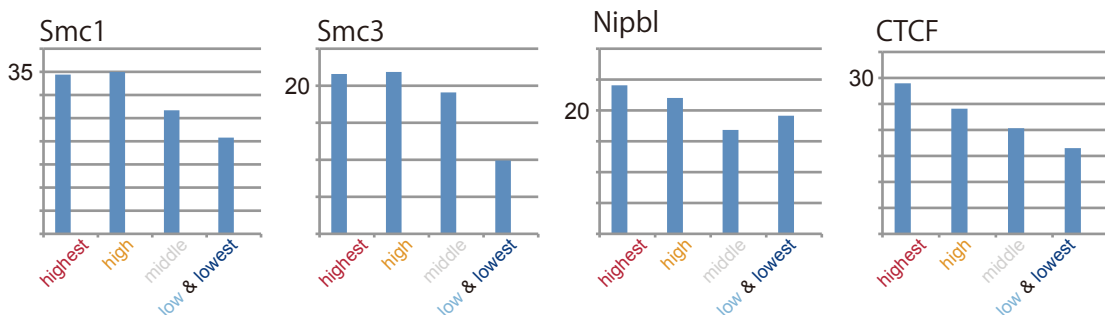
## メディエーター因子



## 転写関連因子



## 核内高次構造関連因子



## ヒストン修飾

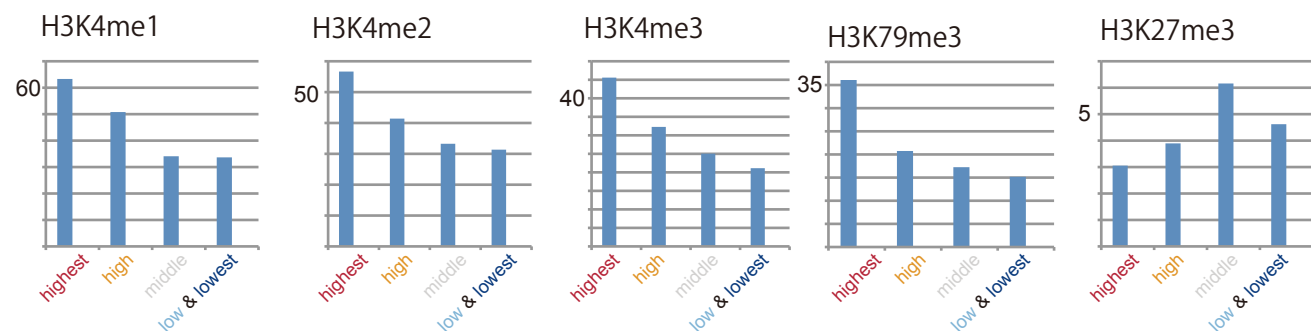
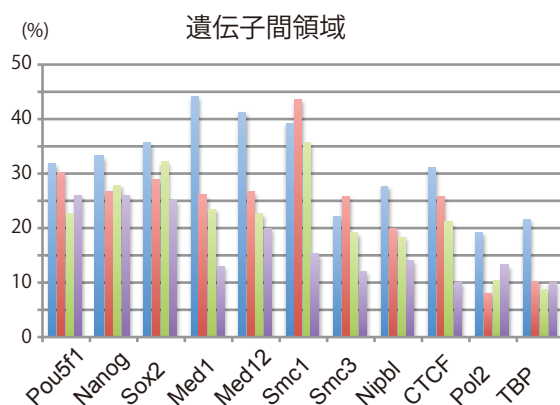
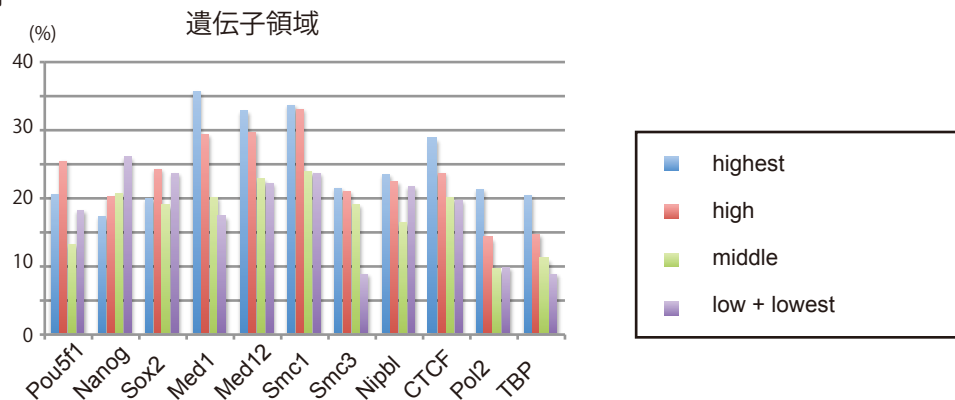


図 2-17 SINE 配列密度によって分類した区画ごとの、ES hypo T-DMR におけるクロマチン関連因子の結合とヒストン修飾の頻度

縦軸は各区画の ES hypo T-DMRtag にクロマチン関連因子の結合・ヒストン修飾が認められた割合 (%) を示す。low と lowest の区画はまとめて扱った。

## クロマチン関連因子



## ヒストン修飾

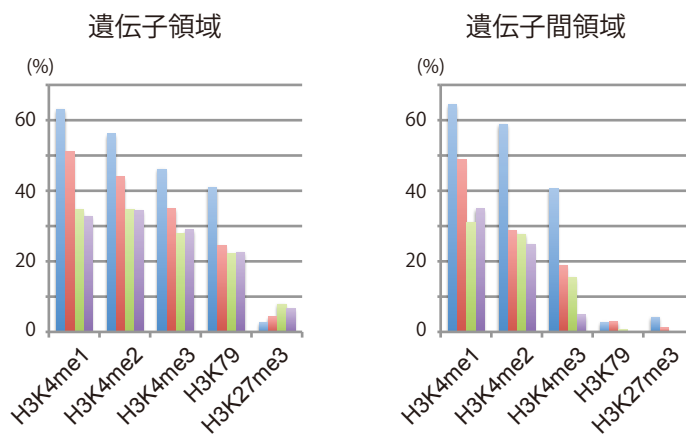


図 2-18 SINE 配列密度によって分類した区画ごとの、遺伝子領域・遺伝子間領域 ES hypo T-DMR におけるクロマチン関連因子の結合とヒストン修飾の頻度  
縦軸は各区画の ES hypo T-DMRtag にクロマチン関連因子の結合・ヒストン修飾が認められた割合 (%) を示す。low と lowest の区画はまとめて扱った。

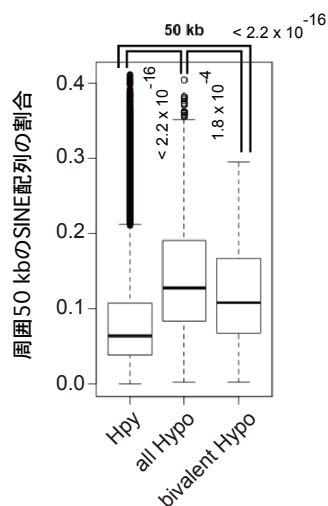
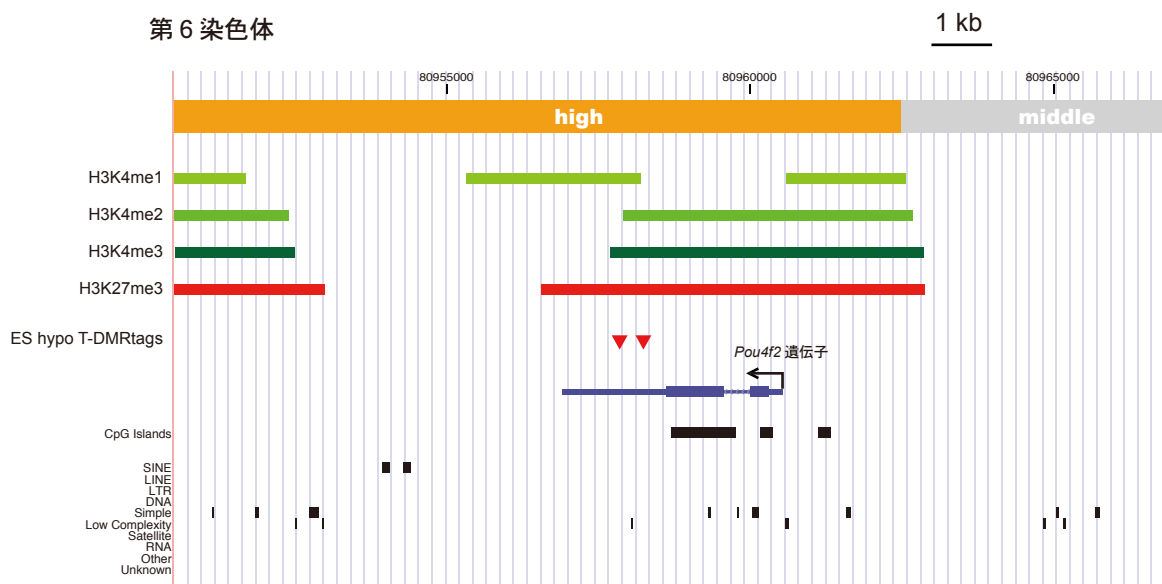


図 2-19 Bivalent 修飾を受ける ES hypo T-DMR の解析

(上段) Bivalent 修飾が認められた ES hypo T-DMR の例。上から、染色体上の場所情報、SINE 配列密度によって分類したドメインの種類、ヒストン修飾、ES hypo T-DMRtag の位置、遺伝子の位置、反復配列の分布を示している。

(下段) Bivalent 修飾が認められた ES hypo T-DMR 周辺 50 kb 領域における SINE 配列の占める割合。有意な差のある比較においては P 値を明記した ( $P < 0.01$ 、ウィルコクソン順位和検定)。箱ひげ図は図 2-6 と同様に描写した。Hpy、Hpy サイト; all Hypo、ES hypo T-DMRtag; bivalent Hypo、Bivalent 修飾が認められた ES hypo T-DMRtag

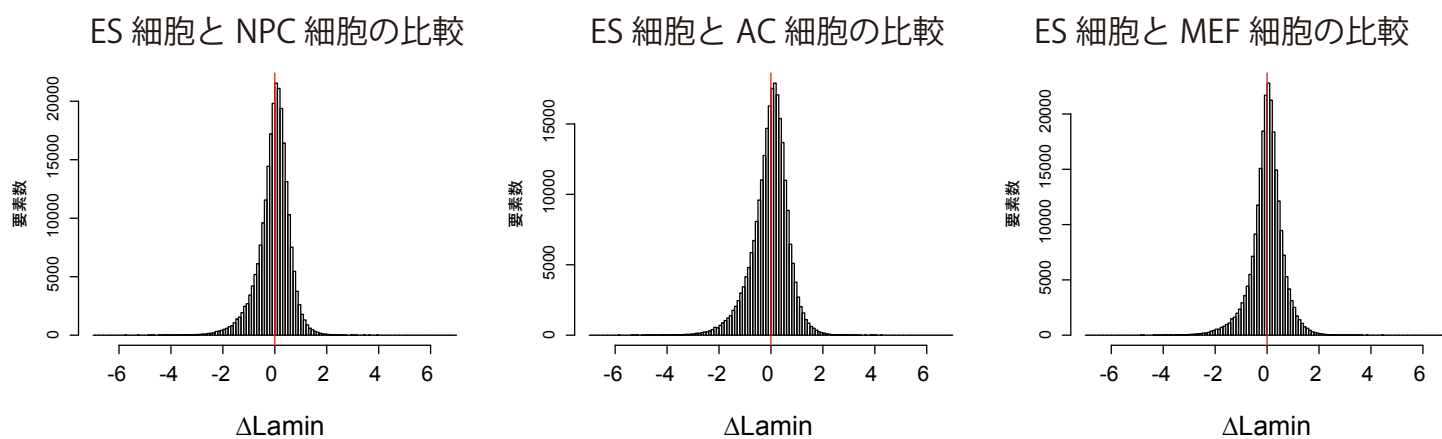
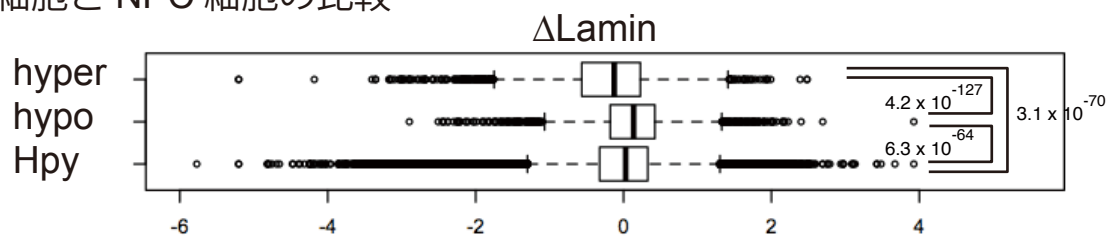


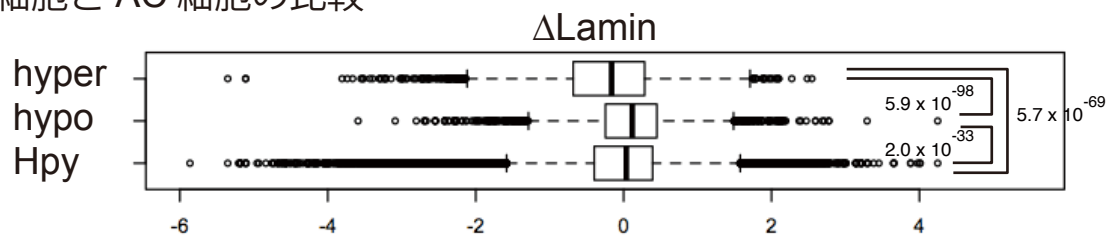
図 2-20 ES 細胞分化過程でのゲノム全体の核ラミナとの相互作用変化

Hpy サイトにおける、ES 細胞と分化細胞（NPC 細胞、AC 細胞、MEF 細胞）の核ラミナとの相互作用の差 ( $\Delta\text{Lamin}$ ) を、ヒストグラムで示した。X 軸は  $\Delta\text{Lamin}$ 、Y 軸は要素数を表す。赤縦線は  $\Delta\text{Lamin} = 0$  を示す。

### ES 細胞と NPC 細胞の比較



### ES 細胞と AC 細胞の比較



### ES 細胞と MEF 細胞の比較

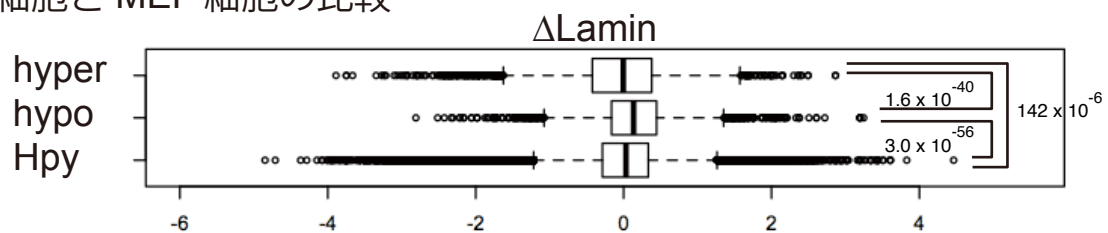
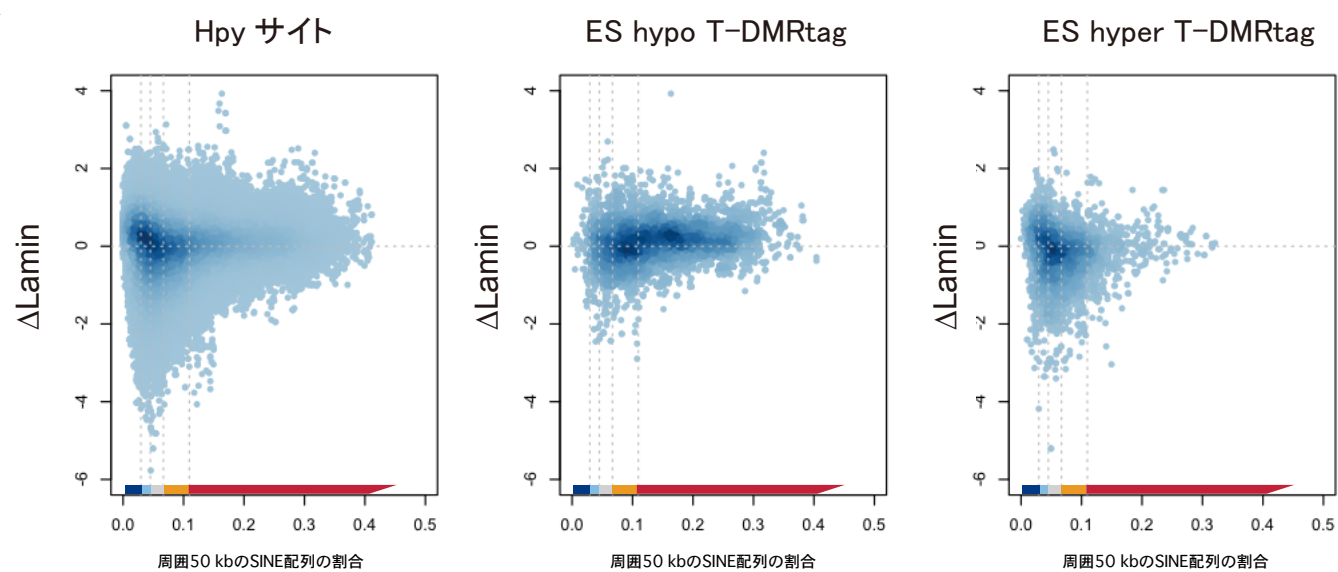


図 2-21 T-DMRtag 周辺領域の $\Delta$ Lamin

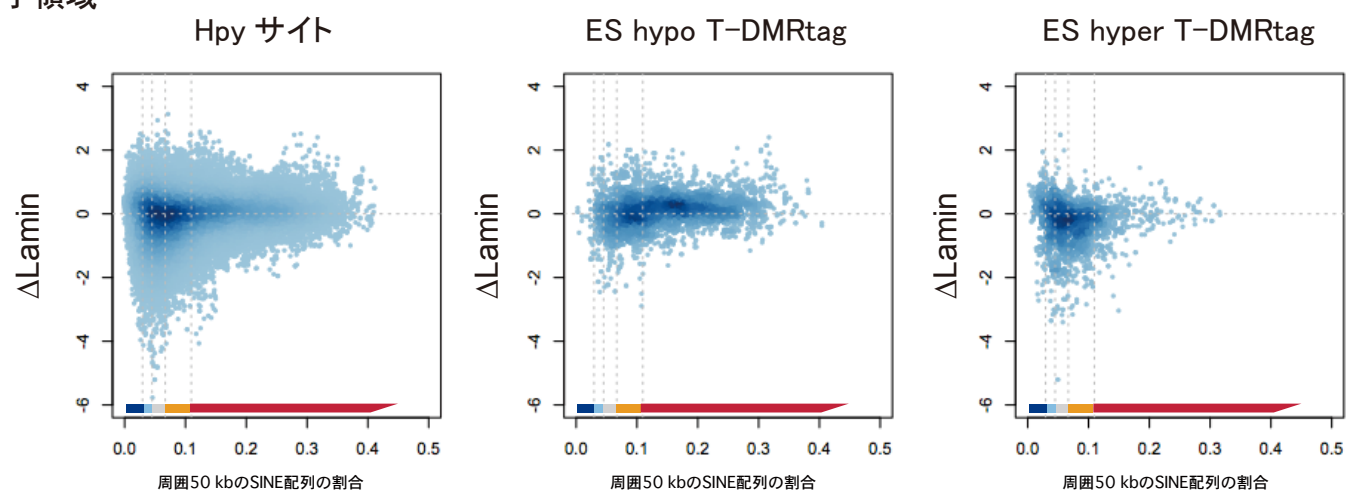
Hpy サイト (Hpy) および ES hypo T-DMRtag (hypo)、ES hyper T-DMRtag (hyper) 周辺 6 kb 領域における $\Delta$ Lamin の箱ひげ図。有意な差のある比較においては P 値を明記した (P < 0.01、ウィルコクソン順位和検定)。

## ES 細胞と NPC の比較

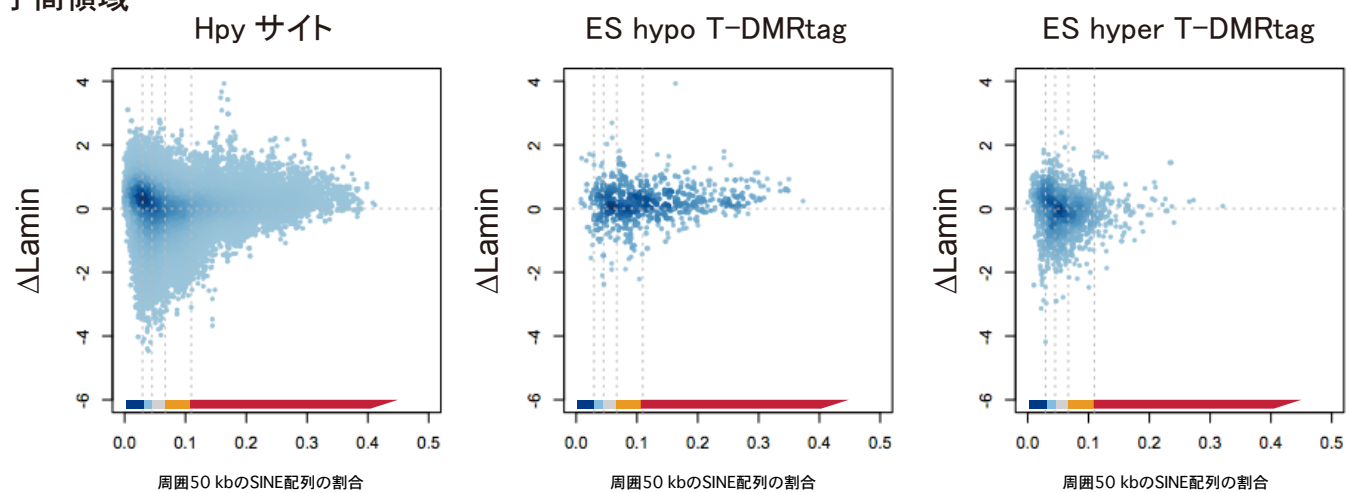
### 全体



### 遺伝子領域

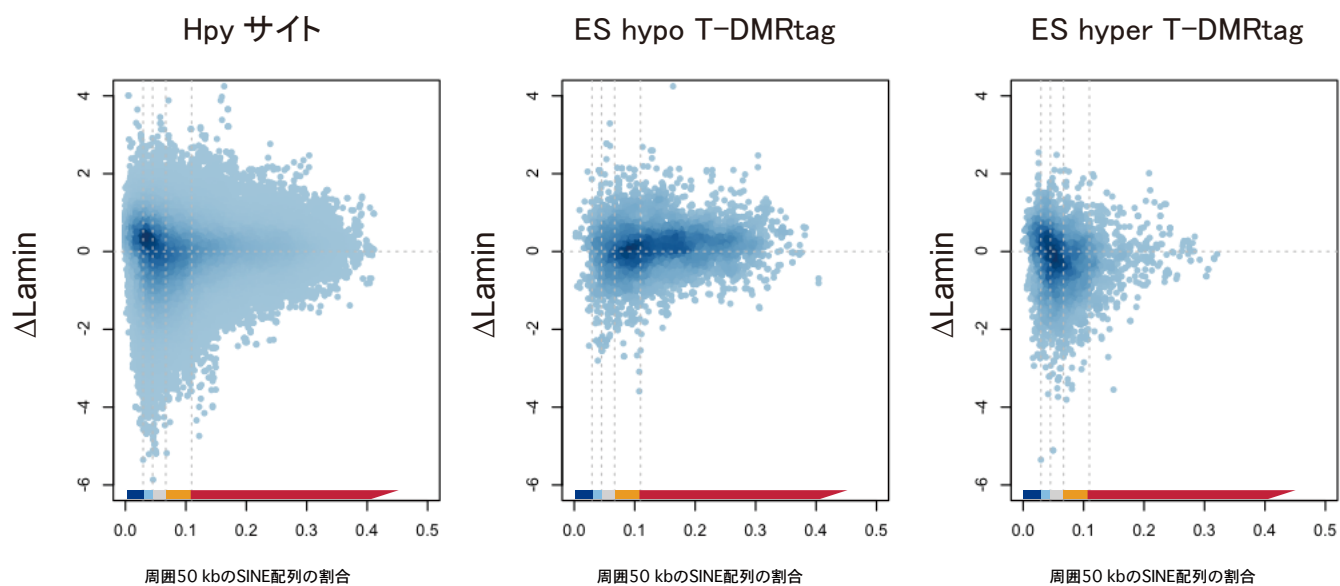


### 遺伝子間領域

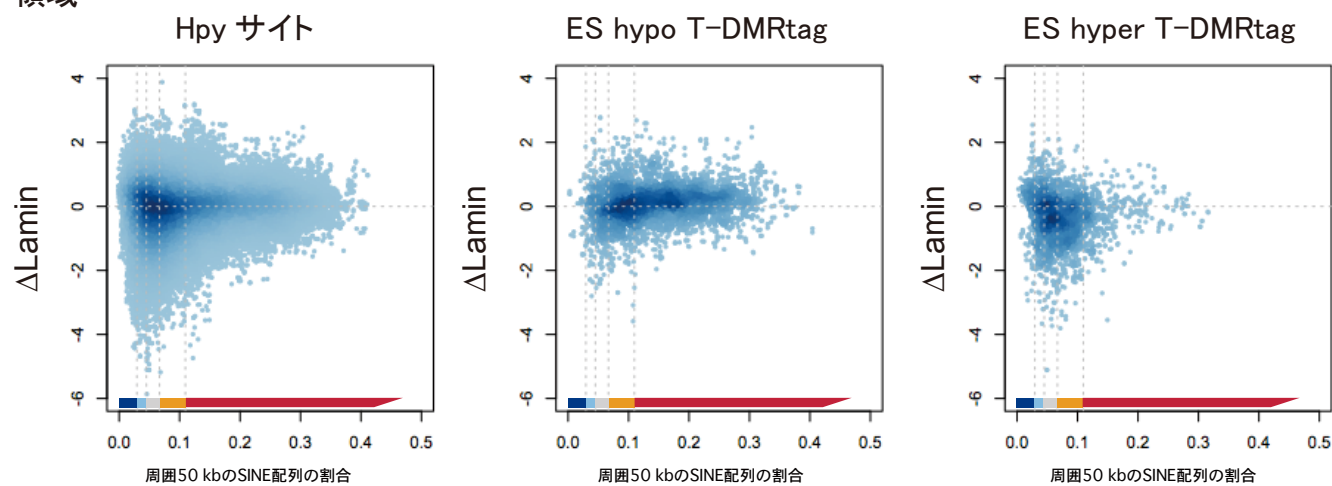


## ES 細胞と AC の比較

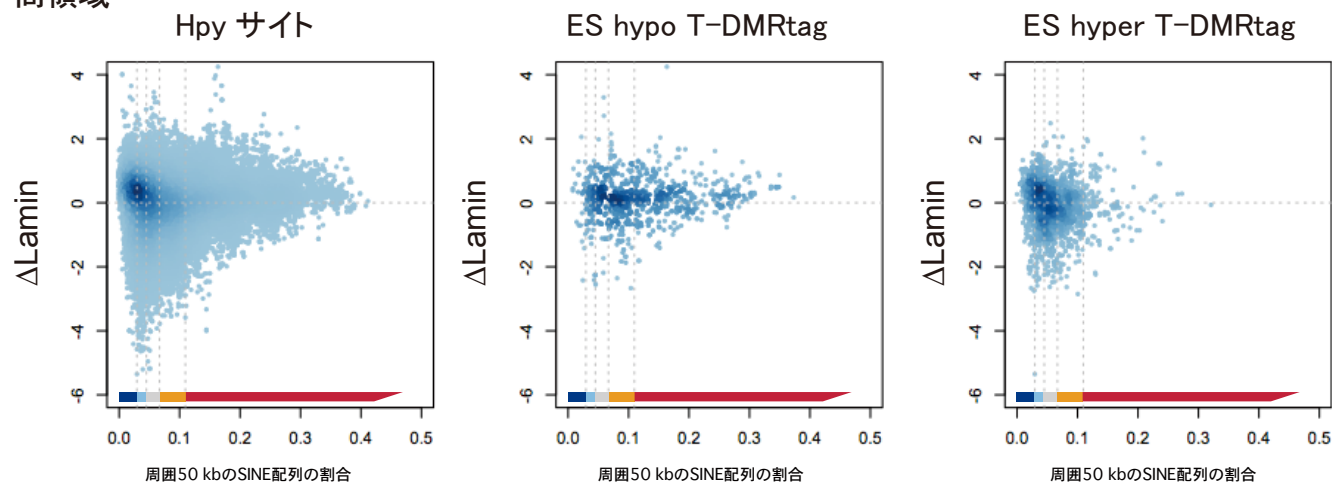
### 全体



### 遺伝子領域

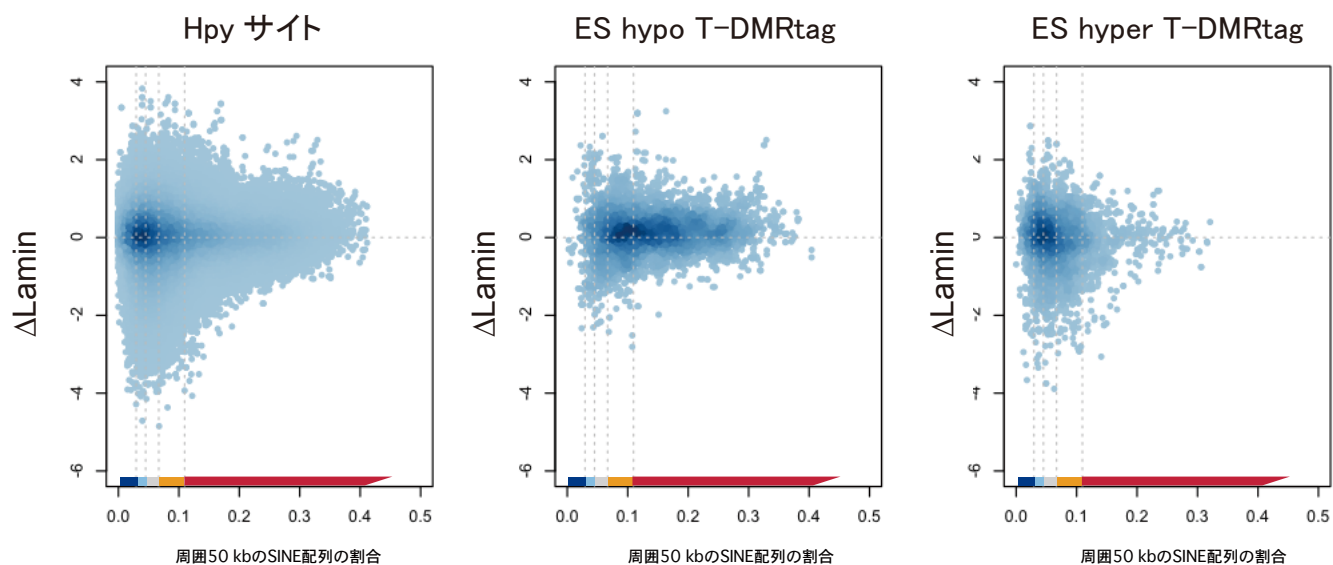


### 遺伝子間領域

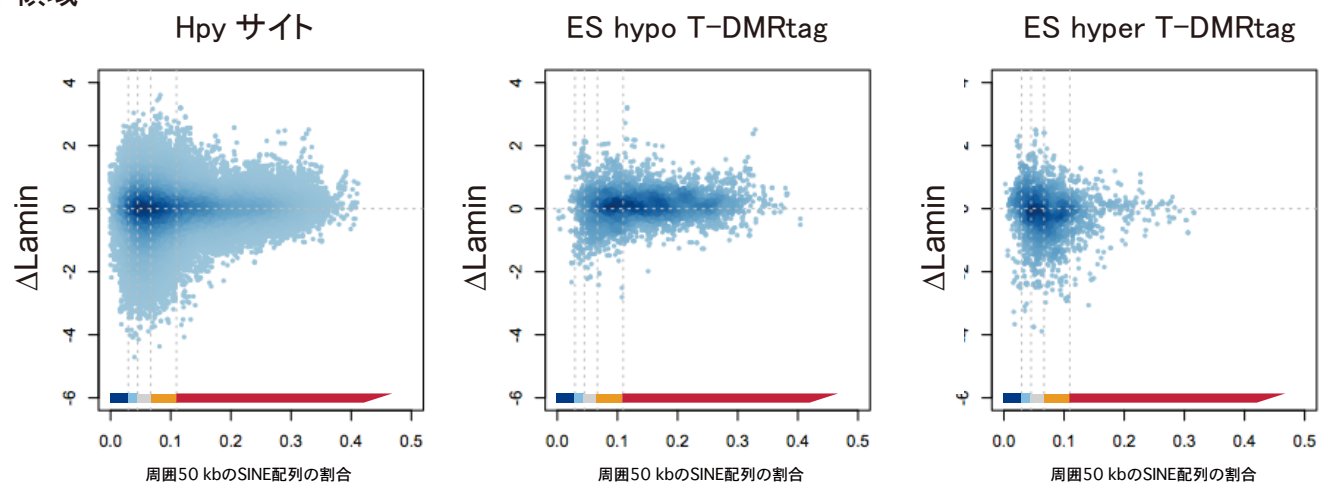


## ES 細胞と MEF の比較

### 全体



### 遺伝子領域



### 遺伝子間領域

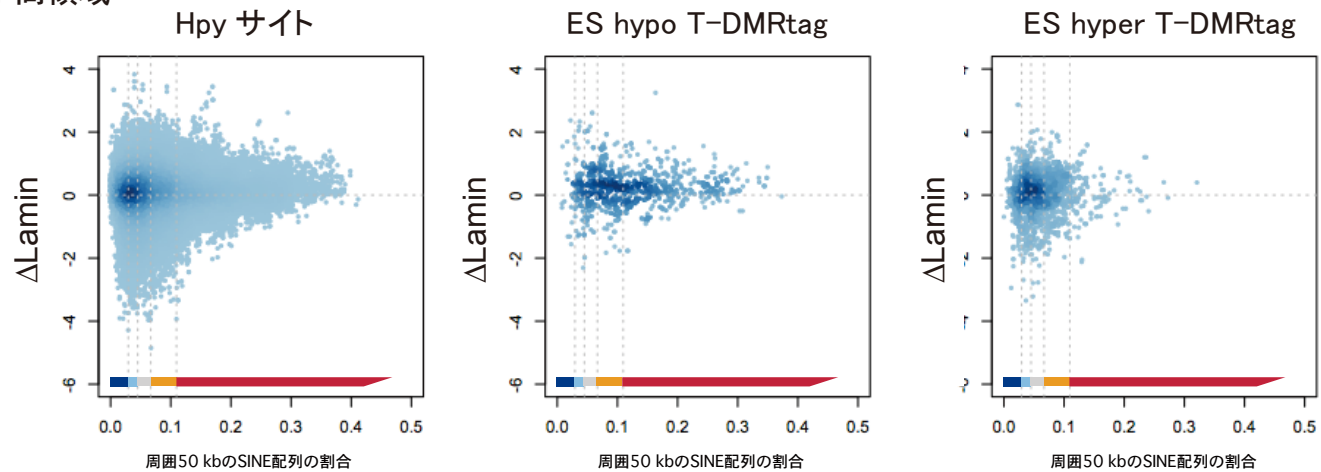


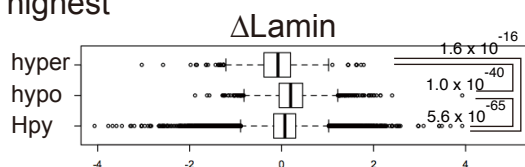


図 2-22  $\Delta$ Lamin と周辺領域の SINE 密度の関係

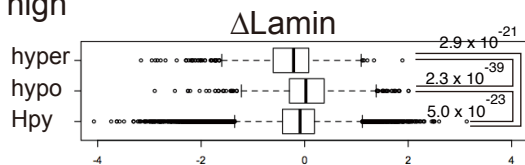
X 軸は周囲 50 kb 領域における SINE 配列の占める割合、Y 軸は $\Delta$ Lamin を表す。X 軸には SINE 配列密度によって分類したドメインの種類を示した。ドメインごとの色は図 2-8 に基づく。ES 細胞と分化細胞（NPC 細胞、AC 細胞、MEF 細胞）のそれぞれの比較において、全体での比較結果（上段）、遺伝子領域での比較結果（中段）、遺伝子間領域での比較結果（下段）を示した。

## ES 細胞と NPC 細胞の比較

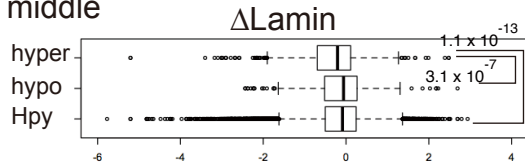
highest



high

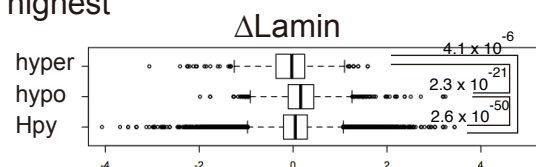


middle

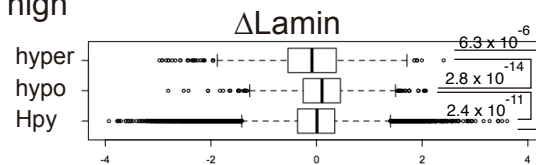


## ES 細胞と MEF 細胞の比較

highest



high



middle

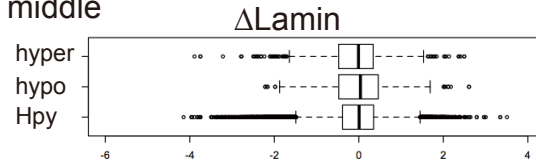


図 2-23 SINE 配列密度によって分類した区画ごとの  $\Delta$ Lamin の解析

ES 細胞と NPC 細胞(左側)および MEF 細胞(右側)の  $\Delta$ Lamin を、Hpy サイト(Hpy)、ES hypo T-DMRtag (hypo)、ES hyper T-DMRtag (hyper) について、SINE 配列密度によって分類した区画ごとに計算し、箱ひげ図を作成した。有意な差のある比較においては P 値を明記した ( $P < 0.01$ 、ウィルコクソン順位和検定)。箱ひげ図は図 2-6 と同様に描写した。上段が highest、中段が high、下段が middle の区画の箱ひげ図。

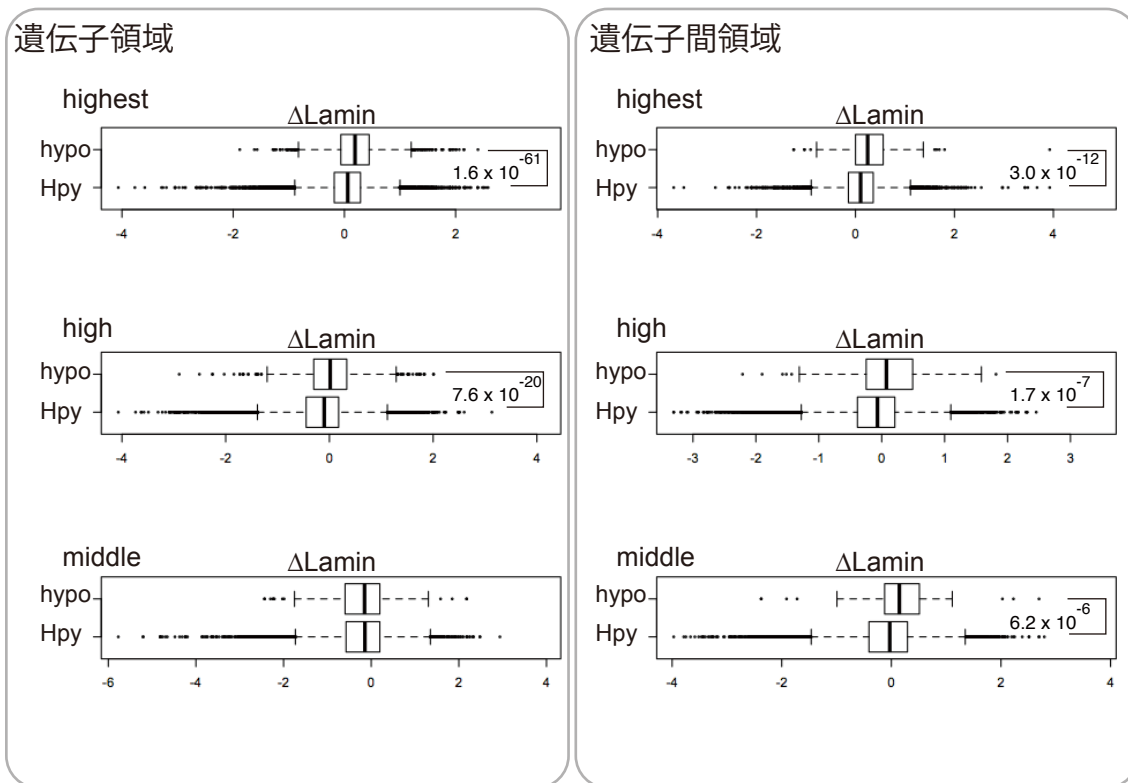


図 2-24 SINE 配列密度によって分類した区画ごとの遺伝子領域・遺伝子間領域における $\Delta\text{Lamin}$ の解析

ES 細胞と NPC 細胞の $\Delta\text{Lamin}$  を、遺伝子領域・遺伝子間領域の Hpy サイト (Hpy)、ES hypo T-DMRtag (hypo) について、SINE 配列密度によって分類した区画ごとに計算し、箱ひげ図を作成した。有意な差のある比較においては P 値を明記した ( $P < 0.01$ 、ウィルコクソン順位検定)。箱ひげ図は図 2-6 と同様に描写した。上段が highest、中段が high、下段が middle の区画の箱ひげ図。

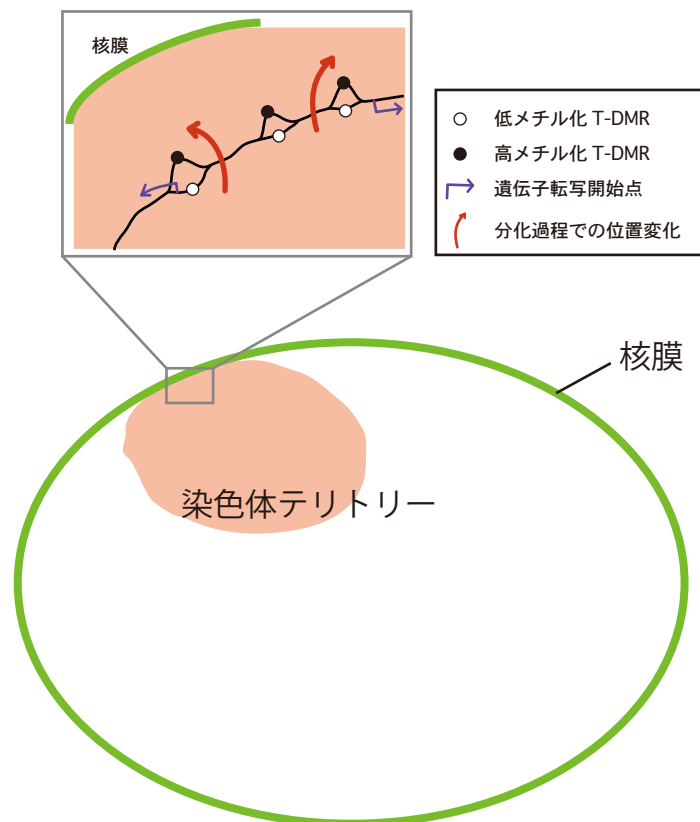


図 2-25 ES 細胞の分化に伴う T-DMR の核内位置変化の模式図

ゲノム全体は ES 細胞の分化過程で核膜との距離は変化しない。SINE 配列密度が高い領域においては、ES hypo T-DMR が特に核膜方向へと移動する。

表 2-1 本章で解析したクロマチン関連因子

名前	説明	その他の機能
Pou5f1	ES 細胞特異的転写因子	
Nanog	ES 細胞特異的転写因子	
Sox2	ES 細胞特異的転写因子	
Med1	メディエーター複合体	
Med12	メディエーター複合体	
Pol2	RNA ポリメラーゼ II	
TBP	TATA ボックス結合タンパク質	
Smc1	コヒーシン複合体	核内高次構造に関わる*
Smc3	コヒーシン複合体	核内高次構造に関わる*
Nipbl	コヒーシン関連因子 (cohesin loading factor)	核内高次構造に関わる*
CTCF	DNA 結合タンパク質	核内高次構造に関わる\$

#### 参考文献

\* Kagey et al., 2010

\$ Phillips & Corces, 2009

表 2-2 本章で解析したヒストン修飾

名前	修飾がみられる領域
H3K79me3	転写の活発な遺伝子のエクソン・イントロン (gene body)
H3K4me1	転写の活発な遺伝子のプロモーター、エンハンサー
H3K4me2	転写の活発な遺伝子のプロモーター
H3K4me3	転写の活発な遺伝子のプロモーター
H3K27me3	転写の抑制された遺伝子のプロモーター、特に CpG アイランド

表 2-3 マスアレイ解析で使したプライマー

解析領域			配列
図2-5	領域A	F	aggaagagagtttatggattttttaaagtaggaaa
		R	cagtaatacgactcactatagggagaaggctcactactcaaacaaacccccaaaaa
	領域B	F	aggaagagagggtagagggagagttgtgtttat
		R	cagtaatacgactcactatagggagaaggctaatacaaaaaccaacttaccaaaa
図2-11	領域A	F	aggaagagagtggttttgggaaggttgagtatatg
		R	cagtaatacgactcactatagggagaaggcttcctacctctccaatcctttaact
	領域B	F	aggaagagagtttttaggaggttttaggaatttg
		R	cagtaatacgactcactatagggagaaggctaataacctaataatctatacacacct
	領域C	F	aggaagagagaggggttttaagttgattgtgaa
		R	cagtaatacgactcactatagggagaaggctacccttttctactaaaaacaaaa
	領域D	F	aggaagagagatttttgggttatgttttttagg
		R	cagtaatacgactcactatagggagaaggctaataattcccaattcttacc
図2-12	領域A	F	aggaagagagaaaggttatgtaaagagtggtttt
		R	cagtaatacgactcactatagggagaaggctcaaatcacaaacaaactacacacc
	領域B	F	aggaagagaggggtattatatttttaggggtgga
		R	cagtaatacgactcactatagggagaaggctacaattaaaaaattccaacaacaca
	領域C	F	aggaagagagtggtgttttttagagggttttgag
		R	cagtaatacgactcactatagggagaaggcttcaaacaaaaataacacctttct
	領域D	F	aggaagagagaaattatgaatttttggtatgttggg
		R	cagtaatacgactcactatagggagaaggctcaaaaatcctttaaaaaacacacca
図2-13	領域A	F	aggaagagagaggtatttttgggagtagtagtta
		R	cagtaatacgactcactatagggagaaggctaatactctctcccaattatctctc
	領域B	F	aggaagagagggaggtgggtgtagattttt
		R	cagtaatacgactcactatagggagaaggctcctcctcctctctccaaaatac
	領域C	F	aggaagagagaaattaggtattttaagttttatggtgtg
		R	cagtaatacgactcactatagggagaaggctcctaatactttacaacaacattacaaca
	領域D	F	aggaagagagattgttttttggttatgaggtttg
		R	cagtaatacgactcactatagggagaaggctaatacaaatatcaccacctactcctc
	領域E	F	aggaagagaggttttttggtgatttttgattttg
		R	cagtaatacgactcactatagggagaaggctttctctcctctctataaaaaa
	領域F	F	aggaagagagtggaatatattttaatgtaggggg
		R	cagtaatacgactcactatagggagaaggctcacataaaaaacataaactcaaca
	領域G	F	aggaagagagagaaggttgaaaatttatttttgt
		R	cagtaatacgactcactatagggagaaggctataaaacttaacatccctctctccc
	領域H	F	aggaagagagtttgaatgttaaaagaattttgtgtatg
		R	cagtaatacgactcactatagggagaaggctcacaaaaaacctatctcaaaaaa
	領域I	F	aggaagagagagttggtggtggtatgtatgtttt
		R	cagtaatacgactcactatagggagaaggctaatttaacccaacaaaaacctct

表2-4 RT-PCRで使したプライマー

名前	配列	解析対象
#1_F	GCACGCCTAAATAGCAAGTAGCA	#1
#1_R	GGGAAGGATCAGATGTGTGTGAG	
#2_F	GCTGACCCATTACTCAACAGGTG	#2
#2_R	AAGGACTTCTGTGACTGGCCTTC	
#3_F	TTGGCTGAGAGCTTCTTTGTCTG	#3
#3_R	GGTCTACACAGTTCCAGGACAGC	
#4_F	TTGTTCCAGAATGATGTGTGTGG	#4
#4_R	TGCATTCCAGTACGCGAATTTAC	
#5_F	CACGAGGCTGTTTTAGAATGTGG	#5
#5_R	TCCAGGTCTCTCAGGCTCTACCT	
#6_F	AGGGCCACAACCTGTTTCCTTCTA	#6
#6_R	AAGGGAACCTGACACAGTGATGA	
#7_F	ACCTGTAGGC AAAAAGCCTGTTC	#7
#7_R	CAATCAGCAGCTCTGTGGTTAGC	
#9_F	ACAAAGGGGGTGTCTTCAAATA	#8
#9_R	GTGACCAGTGAGCTCTGGGTTTA	
#8_F	CCTTATCTCTCAGCCTCCTTTGC	#9
#8_R	AAACAGCTTTCTGGGTCTCAGG	



## 総合討論

DNAのメチル化は哺乳類の発生・分化に必須なエピジェネティック制御である。発生・分化の過程での機能として、(1)細胞・組織特異的な遺伝子発現制御、および(2)ゲノム転移因子の抑制が挙げられる。前者はゲノム情報から細胞・組織に必要な情報の読み取りを、後者はゲノム情報の保護を担い、ともに重要である。哺乳類ゲノムには、細胞・組織特異的にメチル化される領域 (T-DMR) が多数存在する。一方で、ゲノムの半数は、どの細胞組織においてもメチル化修飾をうけるトランスポゾンによって占められている。本研究では、ES細胞特異なメチル化状態を示すT-DMRに焦点を当て、染色体上でT-DMRとトランスポゾンがどのような関係で存在しているのかを明らかにすることを目的に、ゲノムワイドなメチル化解析と配列解析を行った。第一章では遺伝子プロモーター領域に存在するES細胞特異的T-DMRに焦点を当てた解析を行い、ES細胞特異的低メチル化T-DMR (ES hypo T-DMR) をプロモーターに有する遺伝子周辺領域には、SINE配列が多くLINE配列が少ないことを明らかにした。SINE配列、LINE配列の偏った分布は、少なくとも転写開始点前後300 kbという広い領域にわたって認められた。さらに第二章では、対象を染色体全域に広げた解析から、遺伝子領域・遺伝子間領域に関わらず、ES hypo T-DMRはSINE配列が占める割合の高いゲノム領域に偏って存在することを明らかにした。このことは、ES細胞特異的な低メチル化制御と染色体配列構造が関係することを意味している。

第二章ではさらに、染色体構造とES hypo T-DMRのエピジェネティック制御の関係を明らかにした。周辺領域のSINE配列密度が高いES hypo T-DMRほど、遺伝子領域・遺伝子間領域ともにES細胞において転写の活性な領域にみられるヒストン修飾とクロマチン関連因子の結合を認めた。つまり、周辺のSINE配列密度が高い領域のES hypo T-DMRほどES細胞で転写活性が高いクロマチン状態にあることが示唆された。さらに核膜とES hypo T-DMRの相互作用の解析から、SINE

配列密度の高い領域のES hypo T-DMRほどES細胞の分化過程で転写が不活性な領域である核膜へ移動する傾向が認められた。以上より、SINE配列密度の高い染色体領域に存在するES hypo T-DMRほど、クロマチンレベルおよび核内高次構造レベルのES細胞特異的なエピジェネティック制御を受けることが示唆された。

以上本研究より、ES細胞特異的な低メチル化T-DMRは、染色体上に存在する場所およびそのエピジェネティック制御が染色体配列構造と相関を示すことが明らかになった（図1）。このことは、DNAメチル化制御による「ES細胞特異的なゲノム情報の取り出し」が、染色体の配列構造と関係することを示唆している。

本研究の結果をES細胞の核内高次構造および核内動態と考え合わせると、哺乳類ゲノムによる分化多能性制御の新たな一面が浮かび上がってくる。本研究における染色体全域を対象とした解析から、遺伝子領域・遺伝子間領域に関わらずES hypo T-DMRがES細胞のトランスクリプトーム形成に関わることが示唆された。ここで、核内で転写が行われるのは、転写ファクトリーとよばれる核内空間に限られることに注意する必要がある。ES hypo T-DMRはES細胞においては低メチル化状態にあり、転写に許容的な状態にあるが、転写ファクトリーに局在することが出来なければ転写は起こらない。ES細胞は核内動態が活発であることから、ES hypo T-DMRをいかに転写ファクトリーに局在させるかは分化多能性を維持するために重要な問題である。本研究から、ES hypo T-DMRはSINE配列密度が高い染色体領域に集中しており、SINE配列密度の高い領域に存在するES hypo T-DMRほど核内高次構造に関わる因子がよく結合することが明らかになった。このことから、ES hypo T-DMRは染色体上に偏って存在している上にお互いにゲノムの高次構造を介した相互作用することで、空間的にまとまって存在していることが示唆される（図2）。ES hypo T-DMRを空間的にまとめて制御することは、転写ファクトリーへ局在させることを容易

にしており、核内動態の活発なES細胞において上記の「局在問題」を解決する合理的な戦略である。SINE配列密度の高い領域に存在するES hypo T-DMRほど転写関連因子の結合および転写が活発であることを示すヒストン修飾が認められたことは、この考えを支持している。ES細胞では、活発な核内動態に対し、ES hypo T-DMRを空間的にまとめて制御することにより、効率的にES hypo T-DMRによるES細胞特異的なトランスクリプトーム制御を実現していることが示唆された。

ES細胞が分化多能性を発揮し分化を達成するためには、分化過程で分化多能性関連遺伝子を抑制する必要がある。本研究から、分化細胞においてES hypo T-DMRはメチル化修飾を受けるだけでなく、核膜方向へと移動することが明らかになった。SINE配列密度の高い領域には、どの細胞においても転写が活発な傾向があることが知られている (Versteeg et al., 2003; Thurman et al., 2007)。つまり、SINE配列密度の高い領域は分化細胞においても転写ファクトリーに局在する傾向があると考えられる。ES hypo T-DMRは核膜方向へ移動することにより、分化細胞において周囲の遺伝子が転写ファクトリーに局在しても、ES hypo T-DMRは転写ファクトリーに入らないように制御を受けていると考えられる (図2)。分化細胞は、ES hypo T-DMRに対して(1)メチル化修飾による転写関連因子の結合を防ぐ、(2)空間的に移動することで転写ファクトリーへの局在をさける、という二種類の制御をすることで、分化多能性関連遺伝子の抑制を行っていることが考えられた。

以上より哺乳類ゲノムは、「DNAメチル化およびクロマチン関連因子の結合とヒストン修飾によるクロマチンのヌクレオソームレベルの制御」に加え、「ES hypo T-DMRの染色体上および核内空間での存在様式と核内における転写機能の区画化に基づいたゲノムの高次構造レベルの制御」によって、分化多能性を制御しているというモデルが本研究により示唆された。

本研究により、遺伝子領域・遺伝子間領域を問わずES hypo T-DMR周辺には、少なくとも50 kbにわたってSINE配列が多くLINE配列が少ないという配列上の特徴があることが明らかになった。これまでのT-DMRと配列の関係は、T-DMRの近傍に焦点が当てた解析が行われてきた。T-DMR全般に見られる傾向として、T-DMRの近傍500 bp領域の解析から、T-DMRはCpGアイランドに較べてCpG配列が少ない領域に存在することが知られている (Sakamoto et al., 2007)。同様に、プロモーター領域に存在するT-DMRは、転写開始点近傍にCpGアイランドが存在する場合にはCpGアイランド内には少なくCpGアイランドの端 (CpG配列が多い領域と少ない領域の境界) およびその外側に存在することが知られている (Yagi et al., 2008; Sato et al., 2010)。またES細胞で低メチル化を示すT-DMRに特有の性質として、ES細胞で低メチル化を示すT-DMRはCpGアイランドを有する遺伝子に存在する割合が高い (Sato et al., 2010) ことが分かっており、これは第一章の結果とも一致している。これらに対し、本研究ではES hypo T-DMRが有する配列上の特徴が、少なくとも50 kbという広い範囲において認められることを示した。第一章においても議論したように、ES hypo T-DMRを持つ遺伝子周辺には広い範囲でCpG配列頻度が高い傾向があり、しかもその傾向は転写開始点におけるCpGアイランドの有無とは独立して認められた。以上より、本研究によりT-DMR周辺には広い範囲においても、近傍領域での傾向とは異なる配列上の傾向が存在することが示された。今後のエピゲノム解析においては、様々な範囲の配列傾向を解析することで、今まで見過ごされてきたエピゲノムとゲノムの関係が明らかになることが期待できる。

ES hypo T-DMRのメチル化解析から、ES細胞で特異的な低メチル化状態を示すSINE配列が遺伝子領域 (第一章)、遺伝子間領域 (第二章) において同定された。これまでも細胞・組織間のメチル化状態の比較解析により、反復配列のなかに細胞・組織特異的なメチル化状態を示す

ものが存在する例が報告されている (Sakamoto et al., 2007; Xie et al., 2009)。ES hypo T-DMR が周囲にSINE配列が多い傾向を持つことから、今回T-DMRであることを明らかにしたSINE配列以外にも、ユニーク配列と同様にES細胞特異的なメチル化制御を受けているSINE配列がゲノム上に多数存在する可能性がある。従来ゲノムワイドなT-DMRの探索はユニーク配列に焦点があたられてきたが、今後の網羅的なES細胞特異的なメチル化解析においてはユニーク配列、反復配列ともに解析する必要がある。

進化の過程で、なぜES hypo T-DMRの周りにSINE配列が多いという傾向が生じたのだろうか。SINE配列の挿入によってT-DMRが生じた可能性、あるいは、T-DMRが最初にあり、その周辺領域にSINE配列が挿入されていった可能性が考えられる。魚類のなかでフグ、ゼブラフィッシュは全ゲノムバイサルファイトシーケンスによるゲノムワイドなメチル化解析がなされている。CpG配列の9割以上がメチル化され、遺伝子転写開始点近傍では低メチル傾向にあるという哺乳類と同様な特徴を持つことが示されている (Zemach et al., 2010; Feng et al., 2010)。魚類ゲノムにおいても、細胞・組織特異的なDNAメチル化修飾が認められている。ゼブラフィッシュでは、初期発生に重要な遺伝子である *Sox2*、*Pou5f1*、*Klf4* 遺伝子は発生初期の胚では非メチル化状態にあるが、繊維芽細胞においてはメチル化修飾を受けることが明らかにされている (Lindeman et al., 2010)。これらのマウスオーソログ遺伝子はES hypo T-DMRを持つ。このことは、哺乳類ゲノムの進化の過程で挿入されたSINE配列はES hypo遺伝子の分化細胞でのメチル化修飾に必須ではないことを示唆しているだけでなく、さらには哺乳類と魚類が種分化する以前の段階でES hypo遺伝子の祖先遺伝子は「初期胚において低メチル化状態・分化細胞において高メチル化状態」であるT-DMRを既に持っていたことを示唆する。つまり、哺乳類と魚類が種分化する段階で、すでにDNAメチル化制御が形成されており、種分化後のゲノムでSINE配列が挿入

されてきたと考えられるのだ。ES hypo T-DMRの周辺に多いSINE配列は、マウスに特有のSINEサブファミリーであったこと（第一章）、マウス、ヒトでそれぞれのゲノムに特有のSINE配列サブファミリーの分布は非常に高い相関を示すこと（Waterston et al., 2002）は、この考えを支持する。ヒト・マウスが種分化する以前のゲノムの時点で、ES hypo T-DMRは存在しており、その周辺領域はSINE配列の挿入に許容的な性質を持っていたため、ES hypo T-DMRの周りにSINE配列が多いという配列傾向が生じたと考えられる。

では進化の過程でES hypo T-DMR周辺へのSINE配列の挿入が起きてきたのは、どのような良い点があったためなのだろうか。第一章で議論したように、ES hypo T-DMR周辺のSINE配列が分化過程におけるES hypo T-DMRのメチル化制御に関わるかは現在の時点で判断することは出来ない。しかしSINE配列が分化過程におけるメチル化修飾に許容的であることは確かである。このことから、分化過程でメチル化修飾に許容的な配列がES hypo T-DMRの周辺に増えることで、メチル化修飾がより確実になされるようになるという利点があったのではないか、という仮説が立てられる。もちろん、SINE配列の長さが約150 bpから300 bpと他の転移因子にくらべ短く、遺伝子領域に挿入された場合に遺伝子発現制御機構を破壊する確率が低い点も、ES hypo T-DMR近傍への転移が可能であった理由の一つであろう。

地球上の一部の生物は多細胞化することでより複雑な生命活動を可能にしたのと引き換えに、初期胚の有する分化多能性を分化細胞において失わせるために、分化過程において初期発生特異的な遺伝子を抑制する必要性が生じた。特に哺乳類は途中で胎生を獲得したが、これは卵生に比べ特に母親に対しより多くの投資を要求する。従って確実な発生プログラムの進行が進むような淘汰圧がかかったと考えられる。哺乳類ゲノム中のSINE配列はメチル化に許容的である性質によって、発生・分化過程における初期発生特異的な遺伝子の抑制に貢献することにより、ゲノム中で増殖することが出来たという可能性が考えられた。

## 結論

ES hypo T-DMRのDNAメチル化は、SINE配列密度により規定される染色体配列構造、ヌクレオソームレベルおよび核内高次構造レベルの核内クロマチン構築 (chromatin architecture) と協調し、ES細胞特異的なエピゲノムを構築する。

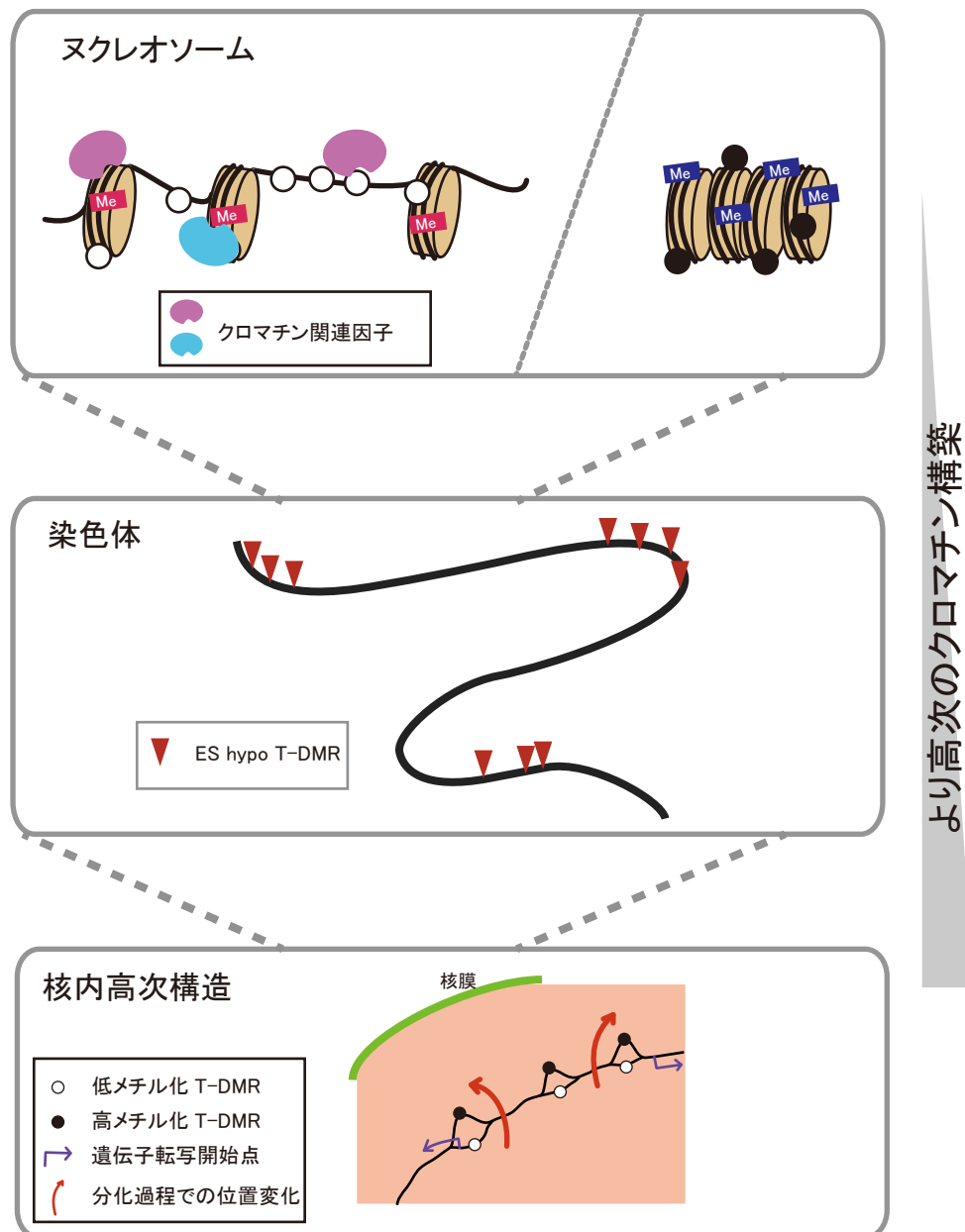


図1 ES hypo T-DMRと様々なレベルの核内構造の関係

ES hypo T-DMRは、ヌクレオソームレベルの制御から核内高次構造を介した制御まで、様々なクロマチン構築 (chromatin architecture) の制御と協調して、ES細胞特異的なエピゲノムを構築する。



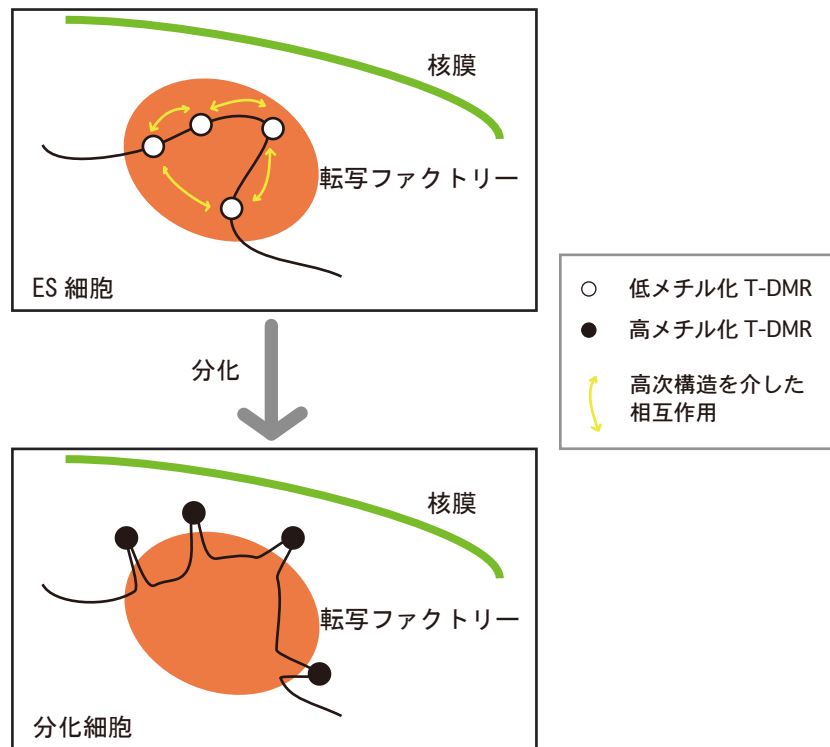


図2 ES hypo T-DMRと転写ファクトリー

核内高次構造レベルの制御による、ES hypo T-DMRのES細胞特異的な転写制御の模式図。

## 引用文献

Allen E, Horvath S, Tong F, Kraft P, Spiteri E, Riggs A, Marahrens Y. 2003. High concentrations of long interspersed nuclear element sequence distinguish monoallelically expressed genes. *Proc Natl Acad Sci U S A* **100**(17): 9940-9945.

Arnaud P, Goubely C, Pélissier T, Deragon J. 2000. SINE retroposons can be used in vivo as nucleation centers for de novo methylation. *Mol Cell Biol* **20**(10): 3434-3441.

Baharvand H, Matthaie KI. 2003. The ultrastructure of mouse embryonic stem cells. *Reprod Biomed Online* **7**(3): 330-335.

Bernstein B, Mikkelsen T, Xie X, Kamal M, Huebert D, Cuff J, Fry B, Meissner A, Wernig M, Plath K et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**(2): 315-326.

Bird A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev* **16**(1): 6-21.

Boyle S, Gilchrist S, Bridger JM, Mahy NL, Ellis JA, Bickmore WA. 2001. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum Mol Genet* **10**(3): 211-219.

Brown T. 2002. *Genomes*, 2<sup>nd</sup> edition. Bios Scientific Publishers, Oxford.

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**(5): 643-655.

Chapman V, Forrester L, Sanford J, Hastie N, Rossant J. 1984. Cell lineage-specific undermethylation of mouse repetitive DNA. *Nature* **307**(5948): 284-286.

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega V, Wong E, Orlov Y, Zhang W, Jiang J et al. 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**(6): 1106-1117.

Cook PR. 2002. Predicting three-dimensional genome structure from transcriptional activity. *Nat Genet* **32**(3): 347-352.

Cremer M, von Hase J, Volm T, Brero A, Kreth G, Walter J, Fischer C, Solovei I, Cremer C, Cremer T. 2001. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res* **9**(7): 541-567.

Cremer T, Cremer C, Baumann H, Luedtke EK, Sperling K, Teuber V, Zorn C. 1982. Rabl's model of the interphase chromosome arrangement tested in Chinese hamster cells by premature chromosome condensation and laser-UV-microbeam experiments. *Hum Genet* **60**(1): 46-56.

Eckhardt F, Lewin J, Cortese R, Rakyan V, Attwood J, Burger M, Burton J, Cox T, Davies R, Down T et al. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* **38**(12): 1378-1385.

Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, Cantor CR, Field JK, van den Boom D. 2005. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A* **102**(44): 15785-15790.

Fawcett DW. 1966. On the occurrence of a fibrous lamina on the inner aspect of the nuclear envelope in certain cells of vertebrates. *Am J Anat* **119**(1): 129-145.

Fazzari MJ, Grealley JM. 2004. Epigenomics: beyond CpG islands. *Nat Rev Genet* **5**(6): 446-455.

Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME et al. 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc Natl Acad Sci U S A* **107**(19): 8689-8694.

Ferrai C, de Castro IJ, Lavitas L, Chotalia M, Pombo A. 2010. Gene positioning. *Cold Spring Harb Perspect Biol* **2**(6): a000588.

Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J et al. 2005. Galaxy: a platform for interactive large-scale genome analysis. *Genome Res* **15**(10): 1451-1455.

Glynn EF, Megee PC, Yu HG, Mistrot C, Unal E, Koshland DE, DeRisi JL, Gerton JL. 2004. Genome-wide mapping of the cohesin complex in the yeast *Saccharomyces cerevisiae*. *PLoS Biol* **2**(9): E259.

Goodier J, Kazazian HJ. 2008. Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* **135**(1): 23-35.

Greally J. 2002. Short interspersed transposable elements (SINEs) are excluded from imprinted regions in the human genome. *Proc Natl Acad Sci U S A* **99**(1): 327-332.

Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**(7197): 948-951.

Hattori N, Imao Y, Nishino K, Ohgane J, Yagi S, Tanaka S, Shiota K. 2007. Epigenetic regulation of Nanog gene in embryonic stem and trophoblast stem cells. *Genes Cells* **12**(3): 387-396.

Hattori N, Nishino K, Ko Y, Ohgane J, Tanaka S, Shiota K. 2004. Epigenetic control of mouse Oct-4 gene expression in embryonic stem cells and trophoblast stem cells. *J Biol Chem* **279**(17): 17063-17069.

Hellmann-Blumberg U, Hintz M, Gatewood J, Schmid C. 1993. Developmental differences in methylation of human Alu repeats. *Mol Cell Biol* **13**(8): 4523-4530.

Iborra FJ, Pombo A, Jackson DA, Cook PR. 1996. Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. *J Cell Sci* **109** ( Pt 6): 1427-1436.

Ikegami K, Ohgane J, Tanaka S, Yagi S, Shiota K. 2009. Interplay between DNA methylation, histone modification and chromatin remodeling in stem cells and during development. *Int J Dev Biol* **53**(2-3): 203-214.

Imamura M, Miura K, Iwabuchi K, Ichisaka T, Nakagawa M, Lee J, Kanatsu-Shinohara M, Shinohara T, Yamanaka S. 2006. Transcriptional repression and DNA hypermethylation of a small set of ES cell marker genes in male germline stem cells. *BMC Dev Biol* **6**: 34.

Imamura T, Ohgane J, Ito S, Ogawa T, Hattori N, Tanaka S, Shiota K. 2001. CpG island of rat sphingosine kinase-1 gene: tissue-dependent DNA methylation status and multiple alternative first exons. *Genomics* **76**(1-3): 117-125.

Johnson W, Li W, Meyer C, Gottardo R, Carroll J, Brown M, Liu X. 2006. Model-based analysis of tiling-arrays for ChIP-chip. *Proc Natl Acad Sci U S A* **103**(33): 12457-12462.

Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* **467**(7314): 430-435.

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A et al. 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* **106**(28): 11667-11672.

Khulan B, Thompson R, Ye K, Fazzari M, Suzuki M, Stasiek E, Figueroa M, Glass J, Chen Q, Montagna C et al. 2006. Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res* **16**(8): 1046-1055.

Kodama H, Hagiwara H, Sudo H, Amagai Y, Yokota T, Arai N, Kitamura Y. 1986. MC3T3-G2/PA6 preadipocytes support in vitro proliferation of hemopoietic stem cells through a mechanism different from that of interleukin 3. *J Cell Physiol* **129**(1): 20-26.

Kouzarides T. 2007. Chromatin modifications and their function. *Cell* **128**(4): 693-705.

Kramerov DA, Vassetzky NS. 2005. Short retroposons in eukaryotic genomes. *Int Rev Cytol* **247**: 165-221.

Lander ES Linton LM Birren B Nusbaum C Zody MC Baldwin J Devon K Dewar K Doyle M FitzHugh W et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**(6822): 860-921.

Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, Itoh T, Watanabe Y, Shirahige K, Uhlmann F. 2004. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* **430**(6999): 573-578.

Levasseur DN, Wang J, Dorschner MO, Stamatoyannopoulos JA, Orkin SH. 2008. Oct4 dependence of chromatin structure within the extended Nanog locus in ES cells. *Genes Dev* **22**(5): 575-580.

Li E. 2002. Chromatin modification and epigenetic reprogramming in mammalian development. *Nat Rev Genet* **3**(9): 662-673.

Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**(6): 915-926.

Lindeman LC, Winata CL, Aanes H, Mathavan S, Alestrom P, Collas P. 2010. Chromatin states of developmentally-regulated genes revealed by DNA and histone methylation patterns in zebrafish embryos. *Int J Dev Biol* **54**(5): 803-813.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**(7271): 315-322.

Liu W, Schmid C. 1993. Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. *Nucleic Acids Res* **21**(6): 1351-1359.

Maeda C, Sato S, Hattori N, Tanaka S, Yagi S, Shiota K. 2008. DNA Hypomethylation Circuit of the Mouse Oocyte-Specific Histone H1foo Gene in Female Germ Cell Lineage. *Biol Reprod*.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**(7205): 766-770.

Meshorer E, Yellajoshula D, George E, Scambler PJ, Brown DT, Misteli T. 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* **10**(1): 105-116.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**(7153): 553-560.

Misteli T. 2010. Higher-order genome organization in human disease. *Cold Spring Harb Perspect Biol* **2**(8): a000794.

Misulovin Z, Schwartz YB, Li XY, Kahn TG, Gause M, MacArthur S, Fay JC, Eisen MB, Pirrotta V, Biggin MD et al. 2008. Association of cohesin and Nipped-B with transcriptionally active regions of the *Drosophila melanogaster* genome. *Chromosoma* **117**(1): 89-102.

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**(5): 631-642.

Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. 2004. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* **279**(50): 52353-52360.

Niwa H, Miyazaki J, Smith AG. 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**(4): 372-376.

Ohshima K, Okada N. 2005. SINEs and LINEs: symbionts of eukaryotic genomes with a common tail. *Cytogenet Genome Res* **110**(1-4): 475-490.

Okano M, Bell D, Haber D, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**(3): 247-257.

Paddy MR, Belmont AS, Saumweber H, Agard DA, Sedat JW. 1990. Interphase nuclear envelope lamins form a discontinuous network that interacts with only a fraction of the chromatin in the nuclear periphery. *Cell* **62**(1): 89-106.

Parada L, Misteli T. 2002. Chromosome positioning in the interphase nucleus. *Trends Cell Biol* **12**(9): 425-432.

Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M et al. 2010. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell* **38**(4): 603-613.

Peters JM, Tedeschi A, Schmitz J. 2008. The cohesin complex and its roles in chromosome biology. *Genes Dev* **22**(22): 3089-3114.

Phillips JE, Corces VG. 2009. CTCF: master weaver of the genome. *Cell* **137**(7): 1194-1211.

Reddy KL, Zullo JM, Bertolino E, Singh H. 2008. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452**(7184): 243-247.



Rollins R, Haghighi F, Edwards J, Das R, Zhang M, Ju J, Bestor T. 2006. Large-scale structure of genomic methylation patterns. *Genome Res* **16**(2): 157-163.

Sakamoto H, Suzuki M, Abe T, Hosoyama T, Himeno E, Tanaka S, Greally J, Hattori N, Yagi S, Shiota K. 2007. Cell type-specific methylation profiles occurring disproportionately in CpG-less regions that delineate developmental similarity. *Genes Cells* **12**(10): 1123-1132.

Sato M, Kimura T, Kurokawa K, Fujita Y, Abe K, Masuhara M, Yasunaga T, Ryo A, Yamamoto M, Nakano T. 2002. Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. *Mech Dev* **113**(1): 91-94.

Sato S, Yagi S, Arai Y, Hirabayashi K, Hattori N, Iwatani M, Okita K, Ohgane J, Tanaka S, Wakayama T et al. 2010. Genome-wide DNA methylation profile of tissue-dependent and differentially methylated regions (T-DMRs) residing in mouse pluripotent stem cells. *Genes Cells*.

Schmidl C, Klug M, Boeld TJ, Andreesen R, Hoffmann P, Edinger M, Rehli M. 2009. Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. *Genome Res* **19**(7): 1165-1174.

Shiota K. 2004. DNA methylation profiles of CpG islands for cellular differentiation and development in mammals. *Cytogenet Genome Res* **105**(2-4): 325-334.

Shiota K, Kogo Y, Ohgane J, Imamura T, Urano A, Nishino K, Tanaka S, Hattori N. 2002. Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. *Genes Cells* **7**(9): 961-969.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**(4): 663-676.

Thurman RE, Day N, Noble WS, Stamatoyannopoulos JA. 2007. Identification of higher-order functional domains in the human ENCODE regions. *Genome Res* **17**(6): 917-927.

- Tomikawa J, Fukatsu K, Tanaka S, Shiota K. 2006. DNA methylation-dependent epigenetic regulation of dimethylarginine dimethylaminohydrolase 2 gene in trophoblast cell lineage. *J Biol Chem* **281**(17): 12163-12169.
- van Steensel B, Dekker J. 2010. Genomics tools for unraveling chromosome architecture. *Nat Biotechnol* **28**(10): 1089-1095.
- Versteeg R, van Schaik BD, van Batenburg MF, Roos M, Monajemi R, Caron H, Bussemaker HJ, van Kampen AH. 2003. The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes. *Genome Res* **13**(9): 1998-2004.
- Walsh C, Chaillet J, Bestor T. 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet* **20**(2): 116-117.
- Wang X, Song X, Glass CK, Rosenfeld MG. 2010. The Long Arm of Long Noncoding RNAs: Roles as Sensors Regulating Gene Transcriptional Programs. *Cold Spring Harb Perspect Biol*.
- Wang Z, Schones DE, Zhao K. 2009. Characterization of human epigenomes. *Curr Opin Genet Dev* **19**(2): 127-134.
- Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P et al. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**(6915): 520-562.
- Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, Tsutsumi S, Nagae G, Ishihara K, Mishihiro T et al. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**(7180): 796-801.
- Xie H, Wang M, Bonaldo MF, Smith C, Rajaram V, Goldman S, Tomita T, Soares M. 2009. High-throughput sequence-based epigenomic analysis of Alu repeats in human cerebellum. *Nucleic Acids Res* **37**(13): 4331-4340.

Xie X, Mikkelsen TS, Gnirke A, Lindblad-Toh K, Kellis M, Lander ES. 2007. Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. *Proc Natl Acad Sci U S A* **104**(17): 7145-7150.

Yagi S, Hirabayashi K, Sato S, Li W, Takahashi Y, Hirakawa T, Wu G, Hattori N, Ohgane J, Tanaka S et al. 2008. DNA methylation profile of tissue-dependent and differentially methylated regions (T-DMRs) in mouse promoter regions demonstrating tissue-specific gene expression. *Genome Res* **18**(12): 1969-1978.

Yamazaki T, Kobayakawa S, Yamagata K, Abe K, Baba T. 2007. Molecular dynamics of heterochromatin protein 1beta, HP1beta, during mouse preimplantation development. *J Reprod Dev* **53**(5): 1035-1041.

Yates P, Burman R, Mummaneni P, Krussel S, Turker M. 1999. Tandem B1 elements located in a mouse methylation center provide a target for de novo DNA methylation. *J Biol Chem* **274**(51): 36357-36361.

Yoder JA, Walsh CP, Bestor TH. 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet* **13**(8): 335-340.

Zemach A, McDaniel IE, Silva P, Zilberman D. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **328**(5980): 916-919.

## 論文内容の要旨

応用動物科学専攻

平成20年度博士課程 進学

村本 玄紀

指導教員 塩田邦郎

## 論文題目

ゲノム転移因子と胚性幹細胞エピゲノム構築に関する研究

## 序論

エピジェネティクスとは、「塩基配列の変化を伴わず、細胞世代を超えて継承される遺伝子機能の変化、またはこの現象を研究する学問分野」として定義される。DNAメチル化およびヒストン修飾はエピジェネティック制御の主要な分子機構である。

ゲノム中には多数の細胞・組織に特異的なメチル化修飾を受ける領域 (tissue-dependent and differentially methylated region、T-DMR) が存在する。ゲノム全体に存在するT-DMRのメチル化状態の組み合わせをDNAメチル化プロフィールとよぶ。各T-DMRのメチル化状態は細胞・組織の種類に依存して異なっていることから、それらを組み合わせたDNAメチル化プロフィールは細胞・組織に固有である。DNAメチル化プロフィールは細胞・組織特有の細胞機能の制御に重要である。

ES細胞は、胚盤胞の内部細胞塊より樹立された培養細胞である。ES細胞は分化多能性を有し、胎仔を構成する全ての種類の細胞へ分化する能力を有する。未分化状態の維持・分化過程におけるDNAメチル化プロフィールの制御が、ES細胞の分化多能性に必要であることが示唆されている。

哺乳類ゲノムの特徴として、反復配列が約半分を占めていることが挙げられる。反復配列として最も多いものは転移因子である。哺乳類ゲノム中のほとんどの転移因子はメチル化修飾により抑制されていることから、DNAのメチル化は転移因子というゲノム中の「寄生因子」に対する防御システムであるという考えが提唱されている。

本研究は、ES細胞特異的なT-DMRに焦点を当て、染色体上におけるT-DMRとトランスポゾンとの関係を明らかにすることを目的とする。第一章において、遺伝子プロモーター領域のT-DMRについて、周囲のトランスポゾンの分布を解析した。第二章においては、第一章の結果をもとに

染色体全域を対象にして、T-DMRと転移因子の関係を調べた。さらにここで分かった関係をもとに、染色体構造とT-DMRのエピジェネティック制御の関係を明らかにした。

## 第一章

遺伝子プロモーター領域に焦点を当て、ES細胞特異的なメチル化制御を受けるT-DMR周辺の反復配列の分布を解析した。その結果、ES細胞特異的に低メチル化をしめすT-DMR (ES hypo T-DMR) を有する遺伝子の周辺には、SINE配列が多く、LINE配列が少ないという特徴があることが分かった。SINE配列が多い傾向は、ES hypo T-DMRを持つ遺伝子の転写開始点の前後少なくとも300 kbにわたって認められた。ゲノム中には複数のES hypo T-DMRが存在する、1 Mb以上にわたりSINE配列が多いゲノムドメインもあった。ES hypo T-DMR周辺のSINE配列の配列上の性質を調べたところ、CpG配列が豊富であり特定のSINEサブファミリーに属することが分かった。SINE配列自身は、ゲノム全体ではES細胞と分化細胞で高メチル化状態にあったが、*Dppa3*遺伝子領域では一部のSINE配列がES細胞特異的に低メチル化を示すT-DMRであった。以上の結果より、SINE配列密度が高いことは、ES hypo T-DMRを持つ遺伝子が存在するゲノム領域の配列上の特徴であることが示された。

## 第二章

第二章では第一章の結果をもとに、ES hypo T-DMRと反復配列の関係を染色体全領域において検証することを目的に、第6、8、16染色体全域を対象としたD-REAM解析を行った。その結果、ES hypo T-DMRには、遺伝子領域・遺伝子間領域に関わらず、周囲のSINE配列密度が高く、LINE配列密度が低い傾向が認められた。さらに、ES hypo T-DMRは染色体上でSINE配列が豊富なゲノムドメインに集中して存在することが明らかになった。ES hypo T-DMRのなかでもSINE配列密度の高いドメインに存在するものほど転写に関連するクロマチン関連因子がよく結合し、転写が活性な領域にみられるヒストン修飾が認められたことから、転写活性が高いことが示唆された。これらの領域は、分化後には転写の不活性な核内領域である核膜との相互作用頻度が上昇する傾向を示した。以上より、ES hypo T-DMRのゲノム上での分布はSINE配列密度という染色体構造の配列構造の特徴と関係し、さらにES hypo T-DMRのエピジェネティック制御も染色体配列構造と相関を示すことが明らかになった。

## 総合討論

本研究では、ES細胞特異的なメチル化状態を示すT-DMRに焦点を当て、染色体上でT-DMRとトランスポゾンがどのような関係で存在しているのかを明らかにすることを目的に、ゲノムワイドなメチル化解析と配列解析を行った。第一章では遺伝子プロモーター領域に存在するES細胞

特異的T-DMRに焦点を当てた解析を行い、ES細胞特異的低メチル化T-DMR (ES hypo T-DMR) をプロモーターに有する遺伝子周辺領域には、SINE配列が多くLINE配列が少ないことを明らかにした。さらに第二章では、対象を染色体全域に広げた解析から、遺伝子領域・遺伝子間領域に関わらず、ES hypo T-DMRはSINE配列が占める割合の高いゲノム領域に偏って存在することを明らかにした。このことは、ES細胞特異的な低メチル化制御と染色体の配列構造が関係することを意味している。第二章ではさらに、染色体配列構造とES hypo T-DMRのエピジェネティック制御の関係を明らかにした。周辺のSINE配列密度が高い領域のES hypo T-DMRほどES細胞における転写活性が高いクロマチン状態にあり、分化過程で核膜へ移動する傾向が認められた。つまり、SINE配列密度の高い染色体領域に存在するES hypo T-DMRほど、ES細胞特異的なエピジェネティック制御を受けることが示唆された。以上より、ES細胞特異的な低メチル化は、存在する場所およびそのエピジェネティック制御において染色体配列構造と相関を示すことが明らかになった。このことは、DNAメチル化制御による「ES細胞特異的なゲノム情報の取り出し」が、染色体の配列構造と関係することを示唆している。

本研究の結果をES細胞の核内高次構造および核内動態と考え合わせると、哺乳類ゲノムによる分化多能性制御の新たな一面が浮かび上がってくる。核内は高度に区画化されており、核内で転写が行われるのは、転写ファクトリーとよばれる核内空間に限られる。ES hypo T-DMRがSINE配列密度の高い染色体領域に集中し、さらにSINE配列密度の高い領域に存在するES hypo T-DMRほど核内高次構造に関わる因子がよく結合することは、ES hypo T-DMRを空間的にまとめて制御することにより、転写ファクトリーへの局在効率を高めることに寄与している可能性がある。また、SINE配列密度の高い領域のES hypo T-DMRは核膜方向へ移動することにより、分化細胞において近隣の遺伝子が転写ファクトリーに局在しても、ES hypo T-DMRは転写ファクトリーに入らないように制御をされていると考えられる。つまり、ES hypo T-DMRによるES細胞特異的な転写活性化を実現するために、哺乳類ゲノムはDNAメチル化修飾というクロマチンレベルの制御に加え、核内高次構造レベルの制御を行っていることが示唆された。

本研究により、ES hypo T-DMRのDNAメチル化は、反復配列密度により規定される染色体構造、ヌクレオソームレベルおよび核内高次構造レベルの核内クロマチン構築 (chromatin architecture) と協調し、ES細胞特異的なエピゲノムを構築することが明らかになった。

## 謝辞

研究の遂行、論文の作成にあたり御指導、御鞭撻いただきました塩田邦郎教授に深く感謝いたします。

本研究にあたり、的確な御助言、御指導をしていただいた八木慎太郎特任教授に感謝いたします。

日々の研究において、さまざまな助言をいただいた細胞生化学研究室の皆様に感謝いたします。

本研究の審査をしていただいた、塩田邦郎教授、千田和広教授、八木慎太郎特任教授、杉浦幸二准教授、武内ゆかり准教授、田中智准教授に感謝いたします。

私を見守り、支えてくださいました家族、友人に感謝いたします。