

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)特許出願公開番号

特開2023-175516

(P2023-175516A)

(43)公開日 令和5年12月12日(2023.12.12)

(51)Int. Cl.	F I	テーマコード (参考)
<i>C 1 2 N 15/54 (2006.01)</i>	C 1 2 N 15/54	4 B 0 5 0
<i>C 1 2 N 9/10 (2006.01)</i>	C 1 2 N 9/10	Z N A 4 B 0 6 3
<i>C 1 2 N 15/63 (2006.01)</i>	C 1 2 N 15/63	Z 4 B 0 6 5
<i>C 1 2 N 1/15 (2006.01)</i>	C 1 2 N 1/15	
<i>C 1 2 N 1/19 (2006.01)</i>	C 1 2 N 1/19	

審査請求 未請求 請求項の数 8 O L (全 26 頁) 最終頁に続く

(21)出願番号 特願2022-87994(P2022-87994)

(22)出願日 令和4年5月30日(2022.5.30)

(71)出願人 504145342

国立大学法人九州大学  
福岡県福岡市西区元岡7 4 4

(74)代理人 100092783

弁理士 小林 浩

(74)代理人 100120134

弁理士 大森 規雄

(74)代理人 100153693

弁理士 岩田 耕一

(72)発明者 三浦 史仁

福岡県福岡市西区元岡7 4 4 国立大学法  
人九州大学内Fターム(参考) 4B050 CC04 DD01 LL03 LL05  
4B063 QA11 QQ42 QR06

最終頁に続く

(54)【発明の名称】 DNAメチル基転移酵素

(57)【要約】

【課題】内在性メチル化が起こるCGジヌクレオチド配列以外で且つ該配列と同程度の短いコンテキストを認識しメチル化することができる、DNAメチル基転移酵素等を提供する。

【解決手段】本発明は、M.CviQIX又はM.CviPIIのDNAメチルトランスフェラーゼのN末端を含む一部のアミノ酸残基を欠失させて、CCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を獲得したタンパク質等に係るものである。

【選択図】なし

## 【特許請求の範囲】

## 【請求項 1】

以下の(a)、(b)又は(c)のタンパク質。

(a) M.CviQIX又はM.CviPIIのDNAメチルトランスフェラーゼのN末端を含む一部のアミノ酸残基を欠失させて、CCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を獲得したタンパク質。

(b) 上記(a)のタンパク質のアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質。

(c) 上記(a)のタンパク質のアミノ酸配列と80%以上の同一性を有するアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質。

10

## 【請求項 2】

以下の(a)、(b)又は(c)のタンパク質。

(a) 配列番号6又は8に示されるアミノ酸配列を有するタンパク質。

(b) 上記(a)のタンパク質のアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質。

(c) 上記(a)のタンパク質のアミノ酸配列と80%以上の同一性を有するアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質。

20

## 【請求項 3】

請求項1又は2に記載のタンパク質をコードする遺伝子。

## 【請求項 4】

以下の(a)、(b)又は(c)のDNAを含む遺伝子。

(a) 配列番号5又は7に示される塩基配列を含むDNA。

(b) 上記(a)のDNAに対し相補的な塩基配列からなるDNAとストリンジェントな条件下でハイブリダイズするDNAであり、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質をコードするDNA。

(c) 上記(a)のDNAに対して80%以上の同一性を有するDNAであり、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質をコードするDNA。

30

## 【請求項 5】

請求項3又は4に記載の遺伝子を含む組換えベクター。

## 【請求項 6】

請求項5に記載の組換えベクターを含む形質転換体。

## 【請求項 7】

請求項6に記載の形質転換体を培養する工程と、

得られる培養物から、CCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質を採取する工程とを含む、当該タンパク質の製造方法。

40

## 【請求項 8】

請求項1若しくは2に記載のタンパク質、又は請求項7に記載の方法により製造されるタンパク質を用いることを含む、エピゲノム解析方法。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、新規なDNAメチル基転移酵素に関する。

## 【背景技術】

## 【0002】

50

DNAメチル基転移酵素（DNAメチルトランスフェラーゼ（MTase））は、DNAの代表的なエピジェネティック修飾の一つであるメチル化を誘導する酵素であり、エピゲノム解析にも有用なツールである。シトシンの5位のDNAメチル化（DNAシトシン5-メチル化（5mC））は、生体のゲノムDNAにおいて広く観察される。

しかしながら、DNAシトシン5-メチル化に関しては、これまでに同定されたMTase（非特許文献1～4）には、その認識配列が哺乳類細胞における固有のDNAメチル化（内在性メチル化）の認識配列であるCGジヌクレオチド配列と重なるという欠点や、また認識配列の塩基数が微細なエピジェネティックマッピングには長すぎるといった問題点がある。

【先行技術文献】

【非特許文献】

【0003】

【非特許文献1】Renbaum, P. et al., Cloning, characterization, and expression in *Escherichia coli* of the gene coding for the CpG DNA methylase from *Spiroplasma* sp. strain MQ1(M.Sssl). *Nucleic acids research*, 1990. 18(5): p. 1145-1152.

【非特許文献2】Xu, M. et al., Cloning, characterization and expression of the gene coding for a cytosine-5-DNA methyltransferase recognizing GpC. *Nucleic Acids Res*, 1998. 26(17): p. 3961-3966. (後述の参考文献21と同じ)

【非特許文献3】Zhang, B. et al., The M.ALuI DNA-(cytosine C5)-methyltransferase has an unusually large, partially dispensable, variable region. *Nucleic Acids Res*, 1993. 21(4): p. 905-911.

【非特許文献4】Vladimir S Dedkov et al., Cloning and Study of New DNA Methyltransferase M.FatI Modifying Cytosine in a Recognition Site CATG. *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, 2015. 6(6): p. 1341-1348.

【発明の概要】

【発明が解決しようとする課題】

【0004】

このような状況下において、内在性メチル化と区別可能なメチル化を導入することができる、すなわち、内在性メチル化が起こるCGジヌクレオチド配列以外で且つ該配列と同程度の短いコンテクストを認識しメチル化することができる、DNAメチル基転移酵素（MTase）等の開発が望まれていた。

【課題を解決するための手段】

【0005】

本発明は、上記状況を考慮してなされたもので、以下に示す、タンパク質（MTase）及びそれをコードする遺伝子等を提供するものである。

【0006】

[1] 以下の(a)、(b)又は(c)のタンパク質。

(a) M.CviQIX又はM.CviPIIのDNAメチルトランスフェラーゼのN末端を含む一部のアミノ酸残基を欠失させて、CCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を獲得したタンパク質。

(b) 上記(a)のタンパク質のアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質。

(c) 上記(a)のタンパク質のアミノ酸配列と80%以上の同一性を有するアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質。

【0007】

[2] 以下の(a)、(b)又は(c)のタンパク質。

(a) 配列番号6又は8に示されるアミノ酸配列を有するタンパク質。

(b) 上記(a)のタンパク質のアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識

10

20

30

40

50

し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質。

(c) 上記(a)のタンパク質のアミノ酸配列と80%以上の同一性を有するアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質。

【0008】

[3] 上記[1]又は[2]に記載のタンパク質をコードする遺伝子。

[4] 以下の(a)、(b)又は(c)のDNAを含む遺伝子。

(a) 配列番号5又は7に示される塩基配列を含むDNA。

(b) 上記(a)のDNAに対し相補的な塩基配列からなるDNAとストリンジェントな条件下でハイブリダイズするDNAであり、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質をコードするDNA。

10

(c) 上記(a)のDNAに対して80%以上の同一性を有するDNAであり、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質をコードするDNA。

【0009】

[5] 上記[3]又は[4]に記載の遺伝子を含む組換えベクター。

[6] 上記[5]に記載の組換えベクターを含む形質転換体。

[7] 上記[6]に記載の形質転換体を培養する工程と、

得られる培養物から、CCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質を採取する工程と

20

を含む、当該タンパク質の製造方法。

[8] 上記[1]若しくは[2]に記載のタンパク質、又は上記[7]に記載の方法により製造されるタンパク質を用いることを含む、エピゲノム解析方法。

【発明の効果】

【0010】

本発明によれば、哺乳類細胞における固有のDNAメチル化(内在性メチル化)と区別可能なメチル化を導入すること、より具体的には、「CCジヌクレオチド配列」を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化することができる、タンパク質(MTase)及びそれをコードする遺伝子等を提供することができる。本発明のタンパク質(MTase)は、内在性メチル化とそれ以外のエピゲノムを同時にプロファイリングするエピゲノム解析方法(エピゲノムマッピング)等が理論上可能となる点で有用なものである。

30

【図面の簡単な説明】

【0011】

【図1】 プラスミドベクター-pCZSのスキームを示す図である。

【図2】 候補MTaseを精製するために用いられるデュアルアフィニティー精製戦略を示す図である。A. 発現タンパク質の構造。B. HisTrap精製及びHisTrap及びStrepTrap精製と組み合わせたデュアルアフィニティー精製によって精製されたタンパク質。ここではMTaseの例としてM.CviPIを示している。C. 精製M.CviPIのDNAメチル化活性をメチル化感受性制限酵素HaeIIIを用いてアッセイした。基質DNAとして非メチル化DNA(Promega)を用いた。コントロールとして、NEBから購入したM.CviPIを使用したレーンを置いた。

40

【図3】 M.CviQIXのN末端伸長がDNAメチル化に対して阻害効果を有することを示す図である。A. M.CviQIXの組換えタンパク質構造を示す図。B. 制限酵素HaeIIIの認識配列。認識配列はCCジヌクレオチドを含む。最初のCが5-メチル化されていれば、HaeIIIは基質DNAを切断できない。C. Aに示した各種のM.CviQIXタンパク質を発現する細菌細胞から抽出したゲノムDNAをHaeIIIで消化し、E-Gel Exにて分析した結果を示す図。D. 各種のM.CviQIXタンパク質構築物の構造を示すスキーム。タグはベクター由来の配列を意味する。E. Dの構築物を用いた細菌細胞のゲノムDNA消化。

【図4】 MTase候補の認識配列を明らかにするための実験戦略を示す図である。A. 使用したスキーム。B. 同定された配列モチーフの例。C. アミノ酸配列の類似性に基づいて系統樹を描いた。決定された認識配列は遺伝子名に従って配置されている。各クレードの共通

50

認識配列はクレードの上部に注釈付けされている。

【図5】固定された酵母核のNOMe-Seqの結果を示す図である。A. CCMT処理核とGCMT処理核間のメチル化レベルの比較binの大きさは10bpである。B.酵母遺伝子(Track 1)、GC部位でのDNAメチル化レベル(Track 2、GCMTを用いたNOMe Seq)及びCC部位(Track 3、CCMTを用いたNOMe Seq)、ATAC-Seqのリードカバレッジ (Track 4、SRR12926697)(参照文献25)、及びMNase-Seqのリードカバレッジ(Track 5、SRR6729489)(参照文献26)。C-E. 酵母遺伝子の開始コドン周囲のアグリゲーションプロット。CC (C)部位とGC (D)部位でのDNAメチル化シグナル、ならびにMNase-Seq (SRR6729489)で描画した例を示す。

【図6】IMR-90細胞のNOMe-Seqの結果を示す図である。A-D. CCMT処理核と未処理核の間のDNAメチル化レベルの比較CG部位(AとB)の内在性DNAメチル化、及びin vitroでCCMTを介して導入されたCC部位(CとD)でのメチル化を示した。E.ヒト遺伝子(Track 1)、固有のメチル化レベル(未処理、CCMT処理、GCMT処理のそれぞれTrack 2、3、4)、MTase誘導メチル化(未処理、CCMT処理、GCMT処理のそれぞれTrack 5、6、7)、ATAC-Seqのリードカバレッジ(SRR7765313)を示すゲノム・ブレイサー・スクリーンショットを示す(Track 8)。右側の図は、左側の図の指示された位置の拡大図である。MTase誘発性のメチル化パターンを示すトラック(Track 5、6、7)については、各トラックの下部にmacs2で検出されるピークを示す。

【図7】いくつかの発現ベクターからM.CviQIX及びM.CviPIIの発現を試みた図である。A. M.CviQIXの発現のために使用される構築物のスキーム。B. ゲノムDNAのメチル化状態を調べるために用いられる制限酵素の認識配列。下線を引いたCがメチル化されると、酵素の活性が阻害される。C. Aに示されるM.CviQIXの発現を誘導した大腸菌細胞から抽出したゲノムDNAのメチル化状態の解析D. M.CviQIXのN-末端欠失突然変異体のスキーム。コールドショックプロモーターに基づく構築物を用いた。E. Dに示された構築物を有する大腸菌細胞から抽出したゲノムDNAの制限消化アッセイ。F. M.CviPIIの欠失突然変異体を比較した。G. Fの構築物を含む大腸菌細胞から抽出したゲノムDNAの制限消化アッセイ。

【図8】MTaseの精製及びその活性を示す図である。A. タンパク質精製を分析する代表的なSDS-PAGEゲル画像。 W: 全細胞抽出物; 3及び4: それぞれHisTrapカラム精製の3番目及び4番目のフラクション。 I: StrepTrap精製の入力。 S: 精製タンパク質。B. 精製タンパク質のMTase活性の解析。メチル化されていない DNAを精製したMTaseで処理し、メチル化感受性制限酵素HpaIIで消化し、E-Gel Exで分析した。

【図9】CCMTの至適塩濃度の検討結果を示す図である。A. 解析に用いた戦略。B. 用いた塩条件で観察されたメチル化レベル。

【図10】固定出芽酵母核のGCMT-及びCCMTベースのNOMe-Seqによる分析結果を示す図である。A. GCMT-、CCMTベースのNOMe-Seq、及びMNase-seqのメチル化レベルに関する凝集プロット。B-C. GCMTベース、CCMTベースのNOMe-Seq、ATAC-seq、及びMNase-seqで同定された重複ピークを示すベン図。

【発明を実施するための形態】

【0012】

以下、本発明を詳細に説明する。本発明の範囲はこれらの説明に拘束されることはなく、以下の例示以外についても、本発明の趣旨を損なわない範囲で適宜変更し実施することができる。

本明細書において引用された全ての刊行物、例えば先行技術文献、及び公開公報、特許公報その他の特許文献は、参照として本明細書に組み込まれる。

本明細書において、「DNAメチル基転移酵素」と、「DNAメチルトランスフェラーゼ」及び「MTase」とは同義である。

なお、以降の記載中に掲載する各々の参考文献の詳細は、本明細書の末尾に列挙する。

【0013】

## 1. 本発明の概要、及び技術的背景

DNAシトシン5-メチル化(5mC)は、生体のゲノムDNAに広く観察され、主要なエピジェネ

10

20

30

40

50

ティック修飾の1つを形成する。古典的には、哺乳類細胞では、5mCレベルの増加は遺伝子サイレンシングにつながると考えられていたが、遺伝子本体のメチル化レベルは遺伝子の発現レベルと正の相関を示すことが示されている(参照文献1-4)。したがって、5mCの分布パターンは細胞型特異的であり、5mCまたはメチロームの全ゲノム分布は、常に多くの研究者の関心を引きつけている。

#### 【0014】

哺乳類細胞における、シトシンの5-メチル化は、DNMT1、DNMT3A、DNMT3Bの3種類のDNAメチルトランスフェラーゼ(MTase)に依存している(参照文献5-7)。DNMT1は、ヘミメチル化されたCGジヌクレオチドにメチル化を導入する維持DNAメチルトランスフェラーゼである。一方、DNMT3AとDNMT3Bは、非メチル化CG部位にDNAメチル化を導入するde novoメチルトランスフェラーゼである。多能性幹細胞、卵母細胞、ニューロン、グリア細胞にはいくつかの例外が知られているが、正常な体細胞組織ではCG以外の部位へのDNAメチル化が起こることはまれである(参照文献8,9)。したがって、5mCはほとんどCGジヌクレオチドのみに導入され、CGジヌクレオチド以外の状況におけるシトシンはほとんどの哺乳動物細胞においてメチル化されていないことが一般的に認められている。

#### 【0015】

5mCのゲノム中の配置は他のエピジェネティック修飾と密接に関係している。例えば、遺伝子プロモーターでは5mCが枯渇しており、ヒストンH3(H3K4me3)のリジン4のトリメチル化が濃縮されている(参照文献10)。別の例は、ヒストンH3リジン36(H3K36me2及びH3K36me3)のジメチル化及びトリメチル化であり、そのゲノム中の配置は転写活性と正に関係している。H3K36me2とH3K36me3は、遺伝子体に富み、DNMT3A/Bを動員して5mCを導入する。したがって、これらの修飾は遺伝子体メチル化(参照文献11,12)を支持する。エピジェネティックな修飾の多くは相互に関連しているため、同一DNA分子上の他のエピジェネティックマーカーと同時にDNAメチル化を検出することは測定意欲を刺激するものとなり、このような測定の実現はエピジェネティックな測定を次の段階にシフトさせるであると考えられる。しかし、エピゲノムを検出するための現在のほとんどの方法は、同時検出には適していない。

#### 【0016】

例えば、エピゲノムを検出するために最もよく用いられる技術である、クロマチン免疫沈降と、それに続く塩基配列決定法(ChIP-Seq)(参照文献13,14)で検討すると、ChIP-Seqでは、まずクロマチンが断片化され、次に断片化されたクロマチン上の標的エピゲノムが免疫沈降され、エピゲノムに結合したDNAが配列決定される。ChIP-Seqは、クロマチン断片化に基づいているので、異なるヌクレオソーム上の近位に位置するエピゲノムは、この最初の断片化段階で分離されるはずである。したがって、単一DNA分子上の2つのエピジェネティック修飾間の関係を測定することは、ChIP-Seqでは実用的ではない。したがって、断片化に基づく手順は、単一DNA分子上で複数のエピジェネティック修飾を同時に測定するのに適していない。

#### 【0017】

しかし、エピゲノム検出には断片化のない手順がいくつかある。たとえば、DNAメチルトランスフェラーゼ(MTase)はクロマチンの接近しやすさ(クロマチンアクセシビリティ)を測定するプローブとして用いることができる。

ヌクレオソーム占有率及びメチロームシーケンシング(NOMe-Seq)は、MTaseによるクロマチンアクセシビリティを測定する最も初期の方法の1つであり(参照文献15)、最近ではナノポアシーケンサーと組み合わせて行われている(参照文献16)。非特異的アデニンN-6 MTase M.EcoGIIとナノポアシーケンサーの組み合わせであるFiber-seq(参照文献17)についても、最近同じ原理が適用された。クロマチンアクセシビリティは、特定されていないタンパク質のDNA-タンパク質相互作用の測定であるが、DNAと特定のタンパク質との相互作用もMTaseで検出することができる。代表的な技術の1つにDamID(参照文献18)がある。DamIDでは、目的タンパク質と融合したdamメチルトランスフェラーゼを標的細胞に発現させ、融合タンパク質とDNAとの特異的相互作用は、アデノシンのN-6位に導

10

20

30

40

50

入されたDNAメチル化を検出することによって測定することができる。当初のDamID報告(参照文献18)でメチル化シグナルがメチル化感受性制限酵素を用いて検出されたが、これはDNAメチル化シグナルを検出するための絶対要件ではない。実際のところ、DamIDを導入したDNAメチル化が、単分子シーケンサーを用いて検出された例もある。また、特定のエピゲノム修飾近傍にアデニンのN-6メチル化を導入するDiMeLo-Seqも実現されており、MTaseを用いる(エピゲノム)検出技術の需要はさらに高まっている。

#### 【0018】

したがって、MTaseは1つのDNA分子上の複数のエピゲノムの断片化を伴わない同時検出を可能にする。

MTaseに基づくエピジェネティック解析では、分解能は主にゲノム中の認識された配列の頻度によって決定され、認識された配列の長さが配列の頻度を決定する。最も基本的なエピゲノム単位ヌクレオソームの1つは、ゲノムDNAの約150塩基対(bp)に結合しているため、エピゲノムマッピングに必要な解像度は150bpよりも高い必要がある。したがって、エピジェネティックな測定に十分な分解能を得るためには、認識配列は3bp(64bp分解能)以下でなければならない。多くのMTaseが細菌及びウイルスから同定されており、M. EcoG IIなどの非特異的アデニンN-6 MTaseが、クロマチンアクセシビリティのファインマッピングを達成している(参照文献17)。しかしながら、DNAシトシンC-5 MTaseに焦点を当てると、エピゲノムマッピングのプロープとして利用可能なMTaseは、3つのみに限定される：M. SssI (CG), M. CviPI (GC), M. CviPII (CCD) (標的のシトシンには下線を付した)(参照文献19,20)。さらに、哺乳類細胞にはCGジヌクレオチドの本来のDNAメチル化(内在性メチル化)があるため、CGジヌクレオチドと重複する配列が、完全に(M. SssI)または部分的に(M. CviPI)存在することが問題となる。そこで、CGジヌクレオチドと決して重複せず且つ短い認識配列をもつシトシンC-5 MTaseを開発し同定することが望まれていた。

#### 【0019】

本発明者は、短い認識配列中のシトシンC-5 MTaseを同定するための系統的戦略を見出し、6つの新規認識配列のMTaseを同定することを可能にした。そして、これらのMTaseのうち、CGジヌクレオチドとの非重複特異性が哺乳類細胞のエピゲノムマッピングに有益であると考えられる、CCジヌクレオチドを認識し得るMTaseに着目し、新規なMTaseを開発した。この新規なMTaseは、より具体的には、CCジヌクレオチド配列を特異的に認識し、当該配列中の5'側のシトシン残基を5-メチル化することができるMTaseである。CCジヌクレオチドという僅か2塩基であり、且つ内在性メチル化が起こるCGとは重複しないコンテキストを認識し、メチル化することができる当該新規なMTaseは、これまでに報告も実現もされていない。

#### 【0020】

## 2. タンパク質

本発明に係るタンパク質は、DNAメチル基転移酵素(MTase)としての機能を有するタンパク質である。具体的には、CCジヌクレオチド配列を特異的に認識し、当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質である。より具体的には、M. CviQIX又はM. CviPIIのMTaseのN末端を含む一部のアミノ酸残基を欠失させることにより、上記活性(CCジヌクレオチド配列中の5'側のシトシン残基を5-メチル化する活性)を有するものとなった(当該活性を獲得した)タンパク質である。

#### 【0021】

ここで、M. CviQIXは、Paramecium bursaria Chlorella virus NY2A由来であるとされてきたMTaseであり、そのアミノ酸配列は、配列番号2に示されるアミノ酸配列である。なお、当該アミノ酸配列をコードする塩基配列としては、限定はされないが、例えば、配列番号1に示される塩基配列が挙げられる。

M. CviPIIは、Chlorella virus由来のMTaseであり、そのアミノ酸配列は、配列番号4に示されるアミノ酸配列である。なお、当該アミノ酸配列をコードする塩基配列としては、限定はされないが、例えば、配列番号3に示される塩基配列が挙げられる。

## 【0022】

なお、M.CviQIX及びM.CviPIIのMTaseは、いずれも、従来から知られている多数のMTaseのうちの一つであり、「CCDトリヌクレオチド配列」（ここで、DはA、G又はC）を特異的に認識してその5'側のシトシン残基を5-メチル化し得るものとして知られているものである。このような、CCDトリヌクレオチドという比較的短い配列を認識してメチル化し得るMTaseを、さらに改変しようとする合理的な理由は、通常、当業者においては存在しない。ましてや、仮に更に短いジヌクレオチド配列を認識しメチル化し得るMTaseに改変させようとした場合であっても、そのN末端を含むアミノ酸残基を欠失させるという手段に至る合理的な理由はない。

## 【0023】

M.CviQIXであるMTaseのN末端を含む一部のアミノ酸残基を欠失させたものとは、限定はされず、例えば、当該MTaseのN末端から、11アミノ酸残基を欠失させたものや、13アミノ酸残基を欠失させたものや、15アミノ酸残基を欠失させたものであってもよいが、なかでも、15アミノ酸残基を欠失させたものが好ましく挙げられる。このM.CviQIXのN末端から15アミノ酸残基を欠失させたタンパク質のアミノ酸配列は、配列番号6に示されるアミノ酸配列である（なお、当該アミノ酸配列をコードする塩基配列は、配列番号5に示される塩基配列である。）。そのため、本発明に係るタンパク質としては、配列番号6に示されるアミノ酸配列を有するタンパク質も、好ましく挙げられる。

## 【0024】

M.CviPIIであるMTaseのN末端を含む一部のアミノ酸残基を欠失させたものとは、限定はされず、例えば、当該MTaseのN末端から、11アミノ酸残基を欠失させたものや、13アミノ酸残基を欠失させたものや、15アミノ酸残基を欠失させたものであってもよいが、なかでも、15アミノ酸残基を欠失させたものが好ましく挙げられる。このM.CviPIIのN末端から15アミノ酸残基を欠失させたタンパク質のアミノ酸配列は、配列番号8に示されるアミノ酸配列である（なお、当該アミノ酸配列をコードする塩基配列は、配列番号7に示される塩基配列である。）。そのため、本発明に係るタンパク質としては、配列番号8に示されるアミノ酸配列を有するタンパク質も、好ましく挙げられる。

## 【0025】

本発明に係るタンパク質としては、限定はされず、例えば、本項において上述した各種タンパク質のアミノ酸配列において1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有する、タンパク質（いわゆる変異型のタンパク質）も、好ましく挙げられる。

ここで、上記「1若しくは数個のアミノ酸が欠失、置換若しくは付加されたアミノ酸配列」としては、例えば、1個～10個程度、好ましくは1～数個、1～5個、1～4個、1～3個、1～2個、又は1個のアミノ酸が欠失、置換又は付加されたアミノ酸配列であることが好ましい。当該欠失、置換又は付加等の変異の導入は、部位特異的突然変異誘発法を利用した変異導入用キット、例えば、GeneTailor™ Site-Directed Mutagenesis System（インビトロジェン社）、及びTaKaRa Site-Directed Mutagenesis System（Prime STAR（登録商標）Mutagenesis Basal kit、Mutan（登録商標）-Super Express Km等：タカラバイオ社製）等を用いて行うことができる。また、上記欠失、置換又は付加の変異が導入されているかどうかは、各種アミノ酸配列決定法、並びにX線及びNMR等による構造解析法などを用いて確認することができる。

## 【0026】

また、本発明に係るタンパク質としては、限定はされず、例えば、本項において上述した各種タンパク質のアミノ酸配列と80%以上の同一性（相同性）を有するアミノ酸配列を含み、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有する、タンパク質（いわゆる変異型のタンパク質）も、好ましく挙げられる。

ここで、上記同一性は、85%以上、90%以上、95%以上、96%以上、97%以上、98%以

10

20

30

40

50

上、又は99%以上であることがより好ましい。

上述した変異型のタンパク質は、該タンパク質のアミノ酸配列をコードする遺伝子を用いて遺伝子工学的に作製することもできる。

#### 【0027】

本発明においては、「CCジヌクレオチド配列を特異的に認識する」とは、シトシン(C)が2塩基連なったジヌクレオチドのみを認識することができることを意味する。「CCジヌクレオチド配列中の5'側のシトシン残基を5-メチル化する」とは、前記認識したCCのうちの5'側の方のシトシン残基に対し、該残基の5位の炭素をメチル化することを意味し、そのメチル化により、5-メチルシトシン又は5-メチルシチジン(5mC)となる。このようなメチル化の活性は、例えば、5mCの存在により切断活性が阻害されることが知られているHaeIII(GGCCを認識)やHpaII(CCGGを認識)によるDNA切断の有無を調べる方法や、

10

バイサルファイトシーケンシング法等により、検出、測定、評価をすることができる。

#### 【0028】

本発明に係るタンパク質は、天然物由来のペプチドを組み合わせて連結したものであってもよいし、人工的に化学合成して得られたものであってもよく、限定はされない。

天然物由来のタンパク質は、天然物から公知の回収法及び精製法により直接得てもよいし、又は、公知の遺伝子組換え技術により、当該タンパク質をコードする遺伝子を各種発現ベクター等に組込んで細胞に導入し、発現させた後、公知の回収法及び精製法により得てもよい。あるいは、市販のキット、例えば、試薬キットPROTEIOS<sup>TM</sup>(東洋紡)、TNT<sup>TM</sup> System(プロメガ)、合成装置のPG-Mate<sup>TM</sup>(東洋紡)及びRTS(ロシュ・ダイアグノステイクス)等を用いた無細胞タンパク質合成系により当該タンパク質を産生し、公知の回収法及び精製法により得てもよく、限定はされない。

20

#### 【0029】

また、化学合成タンパク質は、公知のタンパク質合成方法を用いて得ることができる。合成方法としては、例えば、アジド法、酸クロライド法、酸無水物法、混合酸無水物法、DCC法、活性エステル法、カルボイミダゾール法及び酸化還元法等が挙げられる。また、その合成は、固相合成法及び液相合成法のいずれをも適用することができる。市販のタンパク質合成装置を使用してもよい。合成反応後は、クロマトグラフィー等の公知の精製法を組み合わせてタンパク質を精製することができる。

本発明においては、前述した本発明に係る各種タンパク質とともに、又はそれに代えて、当該タンパク質の誘導体を含むことができる。当該誘導体とは、当該タンパク質に由来して調製され得るものをすべて含む意味であり、例えば、構成アミノ酸の一部が非天然のアミノ酸に置換されたものや、構成アミノ酸(主にその側鎖)の一部に化学修飾が施されたもの等が挙げられる。

30

#### 【0030】

また本発明においては、前述した本発明に係る各種タンパク質、及び/又は、当該タンパク質の誘導体とともに、あるいはそれに代えて、当該タンパク質及び/又は当該誘導体の塩を含むことができる。当該塩としては、生理学的に許容される酸付加塩又は塩基性塩が好ましい。酸付加塩としては、例えば、塩酸、リン酸、臭化水素酸、硫酸などの無機酸との塩、あるいは酢酸、ギ酸、プロピオン酸、フマル酸、マレイン酸、コハク酸、酒石酸、クエン酸、リンゴ酸、蔞酸、安息香酸、メタンスルホン酸、ベンゼンスルホン酸などの有機酸との塩が挙げられる。塩基性塩としては、例えば、水酸化ナトリウム、水酸化カリウム、水酸化アンモニウム、水酸化マグネシウムなどの無機塩基との塩、あるいはカフェイン、ペペリジン、トリメチルアミン、ピリジンなどの有機塩基との塩が挙げられる。

40

#### 【0031】

当該塩は、塩酸などの適切な酸、又は水酸化ナトリウムなどの適切な塩基を用いて調製することができる。例えば、水中、又はメタノール、エタノール若しくはジオキサンなどの不活性な水混和性有機溶媒を含む液体中で、標準的なプロトコルを用いて処理することにより調製することができる。

#### 【0032】

50

## 2. 組換え遺伝子

上述した本発明のタンパク質をコードする遺伝子としては、限定はされないが、例えば、以下の(a)、(b)又は(c)のDNAを含む遺伝子が挙げられる。これらDNAを含む遺伝子としては、これらDNAのみからなるものであってもよいし、これらDNAの一部を含み、その他に遺伝子発現に必要な公知の塩基配列（転写プロモーター、SD配列、Kozak配列、ターミネーター等）や各種制限酵素サイトの認識配列を含むものであってもよく、限定はされない。

### 【0033】

(a) 配列番号5又は7に示される塩基配列を含むDNA。

(b) 上記(a)のDNAに対し相補的な塩基配列からなるDNAとストリンジェントな条件下でハイブリダイズするDNAであり、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質をコードするDNA。

(c) 上記(a)のDNAに対して80%以上の同一性を有するDNAであり、かつCCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質をコードするDNA。

### 【0034】

上記(a)のDNAにおいて、配列番号5に示される塩基配列は、前述したとおり、M.CviQIXのMTaseのアミノ酸配列（配列番号2）におけるN末端から15残基を欠失させたアミノ酸配列を、コードする塩基配列である。また、配列番号7に示される塩基配列は、M.CviPIIのMTaseのアミノ酸配列（配列番号4）におけるN末端から15残基を欠失させたアミノ酸配列を、コードする塩基配列である。

### 【0035】

上記(b)のDNAは、上記(a)のDNA若しくはそれと相補的な塩基配列からなるDNA、又はこれらを断片化したものをプローブとして用い、コロニーハイブリダイゼーション、プラークハイブリダイゼーション、及びサザンプロット等の公知のハイブリダイゼーション法を実施し、cDNAライブラリーやゲノムライブラリーから得ることができる。ライブラリーは、公知の方法で作製されたものを利用してよいし、市販のcDNAライブラリーやゲノムライブラリーを利用してよく、限定はされない。

ハイブリダイゼーション法の詳細な手順については、Molecular Cloning, A Laboratory Manual 4th ed., Cold Spring Harbor Laboratory Press (1989)等を適宜参照することができる。

### 【0036】

ハイブリダイゼーション法を実施における「ストリンジェントな条件」とは、ハイブリダイゼーション後の洗浄時の条件であって、例えば、バッファーの塩濃度が15~330mM、温度が25~65、好ましくは塩濃度が15~150mM、温度が45~55の条件を意味する。具体的には、例えば80mMで50等の条件を挙げることができる。さらに、このような塩濃度や温度等の条件に加えて、プローブ濃度、プローブの長さ、反応時間等の諸条件も考慮し、上記(b)のDNAを得るための条件を適宜設定することができる。

### 【0037】

上記(c)のDNAは、上記(a)のDNAに対して80%以上の同一性（相同性）を有するDNAであり、より好ましくは85%以上、90%以上、95%以上、96%以上、97%以上、98%以上、又は99%以上の同一性を有するDNAが挙げられる。当該DNAは、前記(b)のハイブリダイズするDNAとしても好ましい。

上記(c)のDNA等のような変異型のDNAは、例えば、Molecular Cloning, A Laboratory Manual 4th ed., Cold Spring Harbor Laboratory Press (1989)、Current Protocols in Molecular Biology, John Wiley & Sons (1987-1997)等に記載の部位特異的変位誘発法に準じて調製することができる。具体的には、Kunkel法や Gapped duplex法等の公知手法により、部位特異的突然変異誘発法を利用した変異導入用キットを用いて調製ことができ、当該キットとしては、例えば、QuickChange<sup>TM</sup> Site-Directed Mutagenesis Kit（ス

10

20

30

40

トラタジーン社製)、GeneTailor<sup>TM</sup> Site-Directed Mutagenesis System (インビトロジェン社製)、TaKaRa Site-Directed Mutagenesis System (Mutan-K、Mutan-Super Express Km等:タカラバイオ社製)等が好ましく挙げられる。

#### 【0038】

また、所望のアミノ酸のコドンを示す塩基となるようにミスセンス変異が導入されるように設計したPCRプライマーを用い、野生型MTGをコードする塩基配列を含むDNA等をテンプレートとして、適当な条件下でPCRを行うことにより調製することもできる。PCRに用いるDNAポリメラーゼは、限定はされないが、正確性の高いDNAポリメラーゼであることが好ましく、例えば、Pwo DNA (ポリメラーゼロシュ・ダイアグノスティックス)、Pfu DNAポリメラーゼ (プロメガ)、プラチナPfx DNAポリメラーゼ (インビトロジェン)、KOD DNAポリメラーゼ (東洋紡)、KOD-plus-ポリメラーゼ (東洋紡)等が好ましい。PCRの反応条件は、用いるDNAポリメラーゼの最適温度、合成するDNAの長さや種類等により適宜設定すればよいが、例えば、サイクル条件であれば「90~98 で5~30秒 (熱変性・解離) 50~65 で5~30秒 (アニーリング) 65~80 で30~1200秒 (合成・伸長)」を1サイクルとして合計20~200サイクル行う条件が好ましい。

#### 【0039】

上記(b)や(c)のDNAとしては、例えば、上記(a)のDNAと比較して、塩基配列については完全に同一ではないが、翻訳された後のアミノ酸配列については完全に同一となるような塩基配列からなるDNA (すなわち上記(a)のDNAにサイレント変異が施されたDNA)が、特に好ましい。

上記(b)や(c)のDNAにおける、「CCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性」に関する説明等については、前項における記載内容が適宜適用できる。

#### 【0040】

本発明に係るタンパク質をコードする遺伝子としては、翻訳後の個々のアミノ酸に対応するコドンは、特に限定はされないので、転写後、ヒト等の哺乳類において一般的に用いられているコドン (好ましくは使用頻度の高いコドン)を示すDNAを含むものであってもよいし、また、大腸菌や酵母等の微生物や、植物等において一般的に用いられているコドン (好ましくは使用頻度の高いコドン)を示すDNAを含むものであってもよい。

また、本発明に係るタンパク質をコードする遺伝子としては、当該タンパク質が前述したリンカー配列を含むものである場合は、そのリンカー配列のアミノ酸配列をコードするDNAを含むものであってもよい。

#### 【0041】

### 3. 組換えベクター及び形質転換体

本発明に係るタンパク質を発現させるためには、まず、上述した本発明に係る遺伝子を発現ベクターに組込んで組換えベクターを構築することが必要である。この際、発現ベクターに組込む遺伝子には、必要に応じて、予め、上流に転写プロモーター、SD配列 (宿主が原核細胞の場合)及びKozak配列 (宿主が真核細胞の場合)を連結しておいてもよいし、下流にターミネーターを連結しておいてもよく、その他、エンハンサー、スプライシングシグナル、ポリA付加シグナル、選択マーカー等を連結しておくこともできる。なお、上記転写プロモーター等の遺伝子発現に必要な各要素は、初めから当該遺伝子に含まれていてもよいし、もともと発現ベクターに含まれている場合はそれを利用してもよく、各要素の使用態様は特に限定されない。

#### 【0042】

発現ベクターに当該遺伝子を組込む方法としては、例えば、制限酵素を用いる方法や、トポイソメラーゼを用いる方法など、公知の遺伝子組換え技術を利用した各種方法が採用できる。また、発現ベクターとしては、例えば、プラスミドDNA、バクテリオファージDNA、レトロトランスポゾンDNA、レトロウイルスベクター、人工染色体DNAなど、本発明のタンパク質をコードする遺伝子を保持し得るものであれば、限定はされず、使用する宿主細

胞に適したベクターを適宜選択して使用することができる。

次いで、構築した上記組換えベクターを宿主に導入して形質転換体を得、これを培養することにより、本発明のタンパク質を発現させることができる。なお、本発明で言う「形質転換体」とは宿主に外来遺伝子が導入されたものを意味し、例えば、宿主にプラスミドDNA等を導入すること（形質転換）で外来遺伝子が導入されたもの、並びに、宿主に各種ウイルス及びファージを感染させること（形質導入）で外来遺伝子が導入されたものが含まれる。

#### 【0043】

宿主としては、上記組換えベクターが導入された後、本発明のタンパク質を発現し得るものであれば、限定はされず、適宜選択することができるが、例えば、ヒトやマウス等の各種動物細胞、各種植物細胞、細菌、酵母、植物細胞等の公知の宿主が挙げられる。

動物細胞を宿主とする場合は、例えば、ヒト繊維芽細胞、CHO細胞、サル細胞COS-7、Vero、マウスL細胞、ラットGH3、ヒトFL細胞等が用いられる。また、Sf9細胞、Sf21細胞等の昆虫細胞を用いることもできる。

細菌を宿主とする場合、例えば、大腸菌、枯草菌等が用いられる。

酵母を宿主とする場合は、例えば、サッカロミセス・セレビスエ (*Saccharomyces cerevisiae*)、シゾサッカロミセス・ポンベ (*Schizosaccharomyces pombe*) 等が用いられる。

植物細胞を宿主とする場合は、例えば、タバコBY-2細胞等が用いられる。

#### 【0044】

形質転換体を得る方法は、限定はされず、宿主と発現ベクターとの種類の組み合わせを考慮し、適宜選択することができるが、例えば、電気穿孔法、リポフェクション法、ヒートショック法、PEG法、リン酸カルシウム法、DEAEデキストラン法、並びに、DNAウイルスやRNAウイルス等の各種ウイルスを感染させる方法などが好ましく挙げられる。

得られる形質転換体においては、組換えベクターに含まれる遺伝子のコドン型は、実際に用いた宿主のコドン型と一致していてもよいし、異なってもよく、限定はされない。

#### 【0045】

### 4. タンパク質の製法

本発明に係るタンパク質の製造は、具体的には、前述した形質転換体を培養する工程と、得られる培養物から、CCジヌクレオチド配列を特異的に認識し当該配列中の5'側のシトシン残基を5-メチル化する活性を有するタンパク質を採取する工程とを含む方法により実施することができる。ここで、「培養物」とは、培養上清、培養細胞、培養菌体、又は細胞若しくは菌体の破砕物のいずれをも意味するものである。上記形質転換体の培養は、宿主の培養に用いられる通常の方法に従って行うことができる。目的のタンパク質は、上記培養物中に蓄積される。本発明においては、前記採取する工程は、タンパク質の精製工程を含んでいてもよい。

#### 【0046】

上記培養に用いる培地としては、宿主が資化し得る炭素源、窒素源、無機塩類などを含有し、形質転換体の培養を効率的に行うことができる培地であれば、公知の各種天然培地及び合成培地のいずれを用いてもよい。

培養中は、形質転換体に含まれる組換えベクターの脱落及び目的タンパク質をコードする遺伝子の脱落を防ぐために、選択圧をかけた状態で培養してもよい。すなわち、選択マーカーが薬剤耐性遺伝子である場合には、相当する薬剤を培地に添加することができ、選択マーカーが栄養要求性相補遺伝子である場合には、相当する栄養因子を培地から除くことができる。

#### 【0047】

プロモーターとして誘導性のプロモーターを用いた発現ベクターで形質転換した形質転換体等を培養する場合は、必要に応じて、好適なインデューサー（例えば、IPTG等）を培

10

20

30

40

50

地に添加してもよい。

形質転換体の培養条件は、目的タンパク質の生産性及び宿主の生育が妨げられない条件であれば特に限定はされず、通常、10 ~ 40、好ましくは20 ~ 37 で5~100時間行う。pHの調整は、無機又は有機酸、アルカリ溶液等を用いて行うことができる。培養方法としては、固体培養、静置培養、振盪培養、通気攪拌培養などが挙げられる。

#### 【0048】

培養後、目的タンパク質が菌体内又は細胞内に生産される場合には、菌体又は細胞を破碎することにより目的タンパク質を採取することができる。菌体又は細胞の破碎方法としては、フレンチプレス又はホモジナイザーによる高圧処理、超音波処理、ガラスビーズ等による磨砕処理、リゾチーム、セルラーゼ又はペクチナーゼ等を用いる酵素処理、凍結融解処理、低張液処理、ファージによる溶菌誘導処理等を利用することができる。破碎後、必要に応じて菌体又は細胞の破碎残渣（細胞抽出液不溶性画分を含む）を除くことができる。残渣を除去する方法としては、例えば、遠心分離やろ過などが挙げられ、必要に応じて、凝集剤やろ過助剤等を使用して残渣除去効率を上げることにもできる。残渣を除去した後得られた上清は、細胞抽出液可溶性画分であり、粗精製したタンパク質溶液とすることができる。

10

#### 【0049】

また、目的のタンパク質が菌体内又は細胞内に生産される場合は、菌体や細胞そのものを遠心分離、膜分離等で回収して、未破碎のまま使用することも可能である。

一方、目的のタンパク質が菌体外又は細胞外に生産される場合には、培養液をそのまま使用するか、遠心分離やろ過等により菌体又は細胞を除去する。その後、必要に応じて硫酸沈澱による抽出等により、培養物中から目的タンパク質を採取し、さらに必要に応じて透析、各種クロマトグラフィー（ゲルろ過、イオン交換クロマトグラフィー、アフィニティクロマトグラフィー等）を用いて単離精製することもできる。

20

#### 【0050】

形質転換体等を培養して得られたタンパク質の生産収率は、例えば、培養液あたり、菌体湿重量又は乾燥重量あたり、粗酵素液タンパク質あたりなどの単位で、SDS-PAGE（ポリアクリルアミドゲル電気泳動）等により確認することができる。

また、目的タンパク質の製造は、上述したような形質転換体を用いたタンパク質合成系のほか、生細胞を全く使用しない無細胞タンパク質合成系を用いて行うこともできる。

30

#### 【0051】

無細胞タンパク質合成系とは、細胞抽出液を用いて試験管等の人工容器内で目的タンパク質を合成する系である。また、使用し得る無細胞タンパク質合成系としては、DNAを鋳型としてRNAを合成する無細胞転写系も含まれる。

この場合、使用する細胞抽出液の由来は、前述の宿主細胞であることが好ましい。細胞抽出液としては、例えば真核細胞由来又は原核細胞由来の抽出液、より具体的には、CHO細胞、ウサギ網状赤血球、マウスL-細胞、HeLa細胞、小麦胚芽、出芽酵母、大腸菌などの抽出液を使用することができる。なお、これらの細胞抽出液は、濃縮又は希釈して用いてもよいし、そのままでもよく、限定はされない。

#### 【0052】

細胞抽出液は、例えば限外濾過、透析、ポリエチレングリコール(PEG)沈殿等によって得ることができる。

40

このような無細胞タンパク質合成は、市販のキットを用いて行うこともできる。例えば、試薬キットPROTEIOS<sup>TM</sup>（東洋紡）、TNT<sup>TM</sup>System（プロメガ）、合成装置のPG-Mate<sup>TM</sup>（東洋紡）、RTS（ロシュ・ダイアグノスティクス）等が挙げられる。

無細胞タンパク質合成によって産生された目的のタンパク質は、前述したようにクロマトグラフィー等の手段を適宜選択して、精製することができる。

#### 【0053】

## 5. タンパク質の用途

50

本発明に係るタンパク質（前項の製法により得られるタンパク質も含む）を用いることを含む、エピゲノム解析方法（エピゲノムマッピング）も、本発明に含まれる。

【0054】

エピゲノムマッピングのツールとしてのMTaseの需要が増加しているにも関わらず、利用可能な酵素は限られており、特に、短い認識配列のシトシンC-5 MTaseについては限られている。本発明に係るタンパク質は、哺乳動物細胞において内在性メチル化が起こるCGジヌクレオチド配列とは異なる（CGジヌクレオチド配列とは重複しない）、CCジヌクレオチド配列を認識しメチル化することが可能なものである。そのため、本発明に係るタンパク質は、エピゲノム解析方法（エピゲノムマッピング）のための新しいプローブとして有用なものである。

【0055】

以下に、実施例を挙げて本発明をより具体的に説明するが、本発明はこれらに限定されるものではない。

【実施例1】

【0056】

#### 1. 材料及び方法

##### 配列解析

M. CviQIXのアミノ酸配列はREBASE (<http://rebase.neb.com/rebase/rebase.html>)から（参照文献19,20）、M.CviPIIのアミノ酸配列は参照文献21から、入手した。

【0057】

##### 人工的な遺伝子合成と発現ベクターへのサブクローニング

M.CviQIX及びM.CviPIIをコードする遺伝子はEurofin Genomics Inc.により、E. coliのコード最適化を行った上で合成された。遺伝子内部にBamHI及びEcoRIの認識配列を持たず、それぞれ5'及び3'末端にBamHIとEcoRIの認識配列を持つ遺伝子を合成した（配列番号9及び10（配列番号9及び10の塩基配列は、それぞれ、配列番号5及び7の塩基配列の5'及び3'末端に上記認識配列を付加したものである。）。この遺伝子断片をpCZS（図1）のBamHI-EcoRI部位にサブクローニングした。M.CviQIX及びM.CviPIIのN末端の欠損変異遺伝子の作成は、それぞれの完全長体を有する発現ベクターから、インバースPCRを実施した上でInfusionクローニング（タカラバイオ社）により環状化して実現した。

【0058】

##### MTaseの小規模誘導と細菌ゲノムDNA抽出

M.CviQIX及びM.CviPIIをコードする遺伝子を有する細菌発現ベクターpCZSをT7Express（ニューイングランド・バイオラボ、イプスウィッチ、MA）に導入した。この細胞をLB培地3 mL（1% [w/v]トリプトン、0.5% [w/v]酵母エキス、及び0.5% [w/v]塩化ナトリウム）に100 µg/mLのカルベニシリン（日本、京都、ナカライ・テスク）を接種し、激しく振り混ぜながら37 °Cで一晩増殖させた。各細胞懸濁液50マイクロリットルをカルベニシリン添加LB培地3 mLで希釈し、細胞を振盪しながら37 °Cで3時間培養した。この細胞懸濁液を氷冷水中で30分間冷却し、1Mイソプロピル β-D-1チオガラクトピラノシド（IPTG、ナカライテスク）3 µLを加えてタンパク質発現を誘導した。細胞は遠心分離により採取し、DNeasy Blood & Tissue Kit（Qiagen）によるゲノムDNA精製に使用した。ゲノムDNA濃度は、Qubit dsDNA BRアッセイキット及びQubit蛍光光度計（Thermo Fisher Scientific社、ワルタム、MA、米国）を用いて測定した。

【0059】

##### 制限酵素消化によるDNAメチル化の解析

1×CutSmart Bufferと、制限酵素HaeIII又はHpaIIを含む10 µLの反応にゲノムDNA（100 ng）を加えた。反応を37 °Cで1時間インキュベートし、酵素を70 °Cで10分間熱不活性化した。上記いずれかの制限酵素による処理の後、2% E-Gel Exゲル上のE-Gel Power SNAPシステム（Thermo Fisher Scientific社製）を用いてDNAを分析した。

【0060】

10

20

30

40

50

### 組換えMTaseの精製

MTaseをコードする遺伝子を発現するpCZSで形質転換した細菌細胞を、100 µg/mLのカルベニシリンを含む30mLのLB培地に接種し、37 °Cで一晩激しく振とうして培養した。種子培養は、カルベニシリン添加LB培地1Lを含むバツフルのついた2Lエルレンマイヤーフラスコに移し、さらに37 °Cで4時間振とう培養した。このフラスコを氷水中で30分間冷却し、次いで、230mgのIPTGを加えることによりタンパク質発現を誘導した。フラスコを16 °Cで一晩激しく振り混ぜながらインキュベートした。細胞を2500 × gで15分間の遠心分離を介して収集し、収集した細胞を、使用するまで-80 °Cで凍結した。

#### 【0061】

細菌細胞を、プロテアーゼ阻害剤カクテル(Nacalai Tesque)を含む20mLのHisTrap緩衝液A(20mMリン酸ナトリウム、pH 7.5、800mM NaCl、20mMイミダゾール、1mM DTT、及び10%グリセロール)中に再懸濁し、超音波処理を介して破壊した。溶解物は15,000 × gで15分間遠心分離により除去し、0.45 µmシリンジフィルターを用いてろ過した。HisTrap緩衝液A及びHisTrap緩衝液B(リン酸ナトリウム20mM、pH 7.5、NaCl 800mM、イミダゾール200mM、DTT 1mM、グリセロール10%)を用いて、5mLのHisTrap HPカラム(Cytiva、Marlborough、MA)を用いて、1回目の親和性精製を行った。標的タンパク質を含む画分は、StrepTrap緩衝液A(100mM Tris-HCl、pH8.0、150mM NaCl、1mM EDTA、1mM DTT)とStrepTrap緩衝液B(00mM Tris-HCl、pH8.0、150mM NaCl、1mM EDTA、1mM DTT、及び2.5mM d-デスチオビオチン)を用いて1mL StrepTrap HP (Cytiva)を用いてさらに精製した。クロマトグラフィー精製はAKTA Start System (Cytiva)を用いて行った。デュアルアフィニティー精製タンパク質をAmicon Ultra-4 30 K装置(Millipore)を用いて限外濾過により濃縮し、牛血清アルブミン(BSA)及びグリセロールを添加してそれぞれ100 µg/mL及び50% (v/v)の最終濃度を得た後、使用まで-20 °Cで保存した。

#### 【0062】

### 出芽酵母NOMe-Seq

S288C細胞(National Institute of Technology and Evaluation, NBRC1136の生物資源センター)を50mLのYPD培地(イーストエキス1%、2%のバクトペプトン、2%のデキストロース)で培養し、30 °Cで一晩の振とう培養を行った。細胞懸濁液を1LのYPDで希釈し、さらに30 °Cで4時間振とう培養した。細胞を2000 × gで15分間遠心分離により集め、水50mLで2回洗浄し、1Mソルビトール20mL中で再懸濁した。セル・サスペンションに1-マーカプトエタノール20 µL、Zymolyase 100T 20mg(セイカガク工業(株)、日本)を加え、15分間常温でインキュベートし、細胞壁を消化・除去してスフェロプラストを得た。次に、スフェロプラストを1Mソルビトール20mLで2回洗浄し、1Mソルビトール20mLに再懸濁した。再懸濁したスフェロプラストに38%ホルムアルデヒド(和光化学、日本、大阪)240 µLを加え、室温で5分間インキュベートして固定を行った。固定したスフェロプラストを1Mソルビトール20mLで2回洗浄し、5000 × gで5分間遠心分離によりペレット化した。次いで、スフェロプラストを10mLの核調製溶液(10mM HEPES KOH、pH7.5、10mM NaCl、3mM MgCl<sub>2</sub>、及び0.5% [v/v] Nonidet P 40代替物)に再懸濁し、室温で5分間インキュベートした。核を5000 × gで5分間遠心分離により集め、核洗浄液10mL(10mM HEPES-KOH, pH7.5,10mM NaCl,3mM MgCl<sub>2</sub>)で2回洗浄した。核を10mLの核洗浄液に再懸濁し、そのうち500 µLを20本のチューブに分注し、使用するまで-80 °Cで保存した。

#### 【0063】

凍結核懸濁液を解凍し、500 µLの1 × メチル化緩衝液(50mM Tris HCl, pH 8.5,50mM NaCl,1mM DTT,1.6mM SAM)中に再懸濁した。次に、精製されたM.CviQIXのN末端15アミノ酸残基欠損タンパク質(CCMTと呼ぶ、後述の結果セクションを参照のこと)を適切な量加えた。37 °Cで1時間インキュベートした後、核を5000 × gで1分間遠心分離をしてペレット化し、10mM Tris-HCl,pH 8.0,40 µLに再懸濁した。この核懸濁液に20mg/mlプロテイナーゼK及び10% (w/v) SDSを5 µL添加し、50 °Cで1時間インキュベートした。次に、この反応に1M TrisHCl、pH 8.0の150 µLを添加し、白色鉱油で重層し、80 °Cで24時間インキュベートした。水相を新しい試験管に移し、Buffer AL及びisopropanol 200 µLと混合した。この混合

物をDNeasyカラムに10,000 × gで1分間遠心分離して負荷し、Buffer W1及びBuffer W2を500 µLずつ順次洗浄した。最後に、精製DNAを200 µLのBuffer AEで溶出し、tPBATプロトコルを用いてWGBS分析に用いた。Buffer AL、Buffer W1、Buffer W2、Buffer AEはDNeasy Blood & Tissue Kitに付属の試薬である。

【 0 0 6 4 】

#### 哺乳類細胞のNOMe-Seq

IMR-90細胞(JCRB細胞バンク#JCRB9054)を、10%ウシ胎児血清(Thermo Fisher Scientific社製)及び50 U/mLのペニシリン-ストレプトマイシン(Thermo Fisher Scientific社製)を添加したミニマムエッセンシャル培地(Thermo Fisher Scientific社製)30mLを含む4つのT75フラスコで、湿度100%、CO<sub>2</sub> 5%の雰囲気中、37 °Cで培養した。増殖した細胞をトリプシンで処理し、2000 × gで5分間遠心分離により集めた。D-PBS (Nacalai Tesque)20mLで細胞を2回洗浄し、CELLBANKER 1(Takara Bio)に1 × 10<sup>7</sup>細胞/mlで再懸濁した。細胞懸濁液は500 µLで分注し、使用まで-80 °Cで保存した。

【 0 0 6 5 】

使用前に凍結細胞懸濁液を解凍し、PBSで2回洗浄した。次に、細胞を核調製緩衝液(10mM Tris-HCl、pH 8.0、10mM NaCl、3mM MgCl<sub>2</sub>、0.1mM EDTA)、及び0.5% (w/v) NP-40に再懸濁し、10分間氷上でインキュベートした。次に、核を2000 × gで5分間遠心分離を介して集め、最適量のCCMTで1 × メチル化緩衝液500 µLで再懸濁し、37 °Cで30分間インキュベートした。その後、2000 × gで5分間遠心分離して核を採取し、上清を除去した。核を20mg/mlプロテイナーゼKの20 µLに溶解し、Buffer ATLを200 µL添加した。反応液は56 °Cで1時間インキュベートした後、Buffer AL 200 µLとイソプロパノール200 µLを添加した。この混合物をDNeasyカラムにロードし、Buffer W1またはBuffer W2 500 µLで洗浄した。精製したDNAは200 µLのBuffer AEで溶出し、WGBS解析に用いた。

【 0 0 6 6 】

## 2 . 結果

### N末端欠失はM.CviQIX及びM.CviPIIのMTase活性を増強する

本発明者は、CCDトリヌクレオチドの最初のCに対するMTase活性を有することがデータベース上に記載されていたM.CviQIXの活性型タンパク質の生産を試みた。しかし、この遺伝子がサブクロニングされたいくつかの発現ベクターを持つほとんどの細菌細胞において、DNAメチル化活性を検出することはできなかった(図3A-3C)。ただし、発現ベクターpC<sub>TF</sub>及びpC<sub>N5</sub>でサブクロニングしたM.CviQIXを発現する細胞では比較的強いMTase活性が認められた。つまり、これら二種類の発現ベクターを有する細胞のゲノムDNAは、HaeIII及びHpaII消化の両方に対して耐性を示した(図3C、及び図7A-7B)。pC<sub>TF</sub>及びpC<sub>N5</sub>は、それぞれM.CviQIXのN末端にトリガー因子(TF)及びSUMOと融合したタンパク質を産生するようにデザインされている(図3A)。TFとSUMOはタンパク質の可溶化やフォールディングに対する効果が期待される(参照文献22)。また、N末端がSUMO標識される発現ベクターを保有する細菌細胞から抽出したゲノムDNAは、C末端がSUMO標識される発現ベクターを保有する細菌細胞から抽出されたゲノムDNAに比べて、メチル化感受性制限酵素による処理に対してより強い性を示した(図3C)。そこで、これらのN末端TF及びSUMO標識コンストラクトの産物の精製を試みた。しかし、N末端His6タグを用いたアフィニティー精製後、両タンパク質はほとんどMTase活性を示さなかった。これらの結果は、タンパク質の可溶性向上やフォールディングとは異なる何らかの要因がN末端にあり、M.CviQIXのDNAメチル化活性を制御している可能性を示唆していた。そこで、N末端配列が精製M.CviQIXの低メチル化活性を引き起こす原因かもしれないという仮説を立てた。そこでM.CviQIXのN末端欠損タンパク質を有する発現ベクターを有する細菌細胞のゲノムDNAの分析を行うことにした。M.CviQIXのN末端には3つのメチオニン残基がある。そこで、この遺伝子を元々保有する生物の細胞内においては、これらの3残基のうちいずれかが翻訳開始に使用されるのではないかと考えた。そこで、それぞれのメチオニンの一残基手前のアミノ酸残基までが欠損した3種類の変異体を作成し、その宿主ゲノムのメチル化状態を分析した。その結果、M.CviQI

XのN末端11、13、または15アミノ酸の欠失は宿主ゲノムのメチル化を増強することが判明した(図3D-3E、及び図7C)。つまり、これら3つのメチオニン残基のいずれも、活性なM.CviQIXを産生するための開始コドンとして使用することができた。また、M.CviPIIにおいても同様の結果が得られた(図7F-7G)つまり、M.CviQIX及びM.CviPIIのN末端アミノ酸配列はDNAメチル化活性に対して阻害作用を示すことが明らかとなった。

【0067】

#### WGBSによるM.CviQIXとM.CviPIIの認識配列の決定

本発明者は、E. coli細胞においてM.CviQIX及びM.CviPIIのN末端15アミノ酸欠失(NDEL15)変異体の発現を誘導し、細菌のゲノムDNAを抽出し、全ゲノムバイサルファイトシーケンシング解析(WGBS)を実施した(図4A)。各シトシン残基についてメチル化レベルを測定し、メチル化レベルに寄与する周辺塩基配列の出現頻度をロゴ表示した結果、M.CviQIX及びM.CviPIIはCCジヌクレオチドの最初のCに対してメチル化を導入することが明らかになった。

10

【0068】

#### MTaseの精製と in vitroでの活性の確認

次に、本発明者は、M.CviQIX及びM.CviPIIのN末端15アミノ酸残基欠損タンパク質遺伝子を用いて活性タンパク質の精製を試みた。この際、精製条件が確立されているM.CviPIの活性タンパク質の精製もコントロールとして同時に試みた。この遺伝子がクローニングされた発現ベクターを有する細菌細胞を培養し、タンパク質発現を誘導した後、HisTag及びStrepII-tagを用いた二重アフィニティー精製法を用いてこれらのMTaseを精製した(図2)。

タンパク質を100 µg/ml BSA、10mM DTT、及び50%グリセロールの存在下で保存し、-20度の保存中に酵素活性を保存した。精製タンパク質は、SDS-PAGEでは高度に均質であると思われた。ファージDNAを基質として制限酵素及びWGBSでアッセイしたところ、これらのMTaseはDNAの切断阻害を引き起こしたことから、メチル化活性を示すことが確認された(図8B)。

20

【0069】

#### CC認識MTaseを有するNOMe-SeqはGCMTと同様のメチル化パターンを生成する

CCジヌクレオチドは、認識配列が短く、約10bp分解能の微細なエピゲノムマッピングを可能にする。そこで、本発明者はCCジヌクレオチドを認識するMTaseであるM.CviQIXあるいはM.CviPIIをエピゲノムマッピングのプロープとして確立することを試みた。発現ベクターと遺伝子の組み合わせを比較した結果、コールドショックプロモーターの制御下で発現させたM.CviQIXが最も高い収率と活性を示したため、このタンパク質をCCMTと命名した。CCMTのメチル化反応における塩濃度を検討した結果、50mM塩化ナトリウムが厳密なDNAメチル化活性の最良条件であることを見出した(図9)。

30

【0070】

次に、ホルムアルデヒド架橋した出芽酵母核のNOMe-Seq解析にCCMTを用いた。ホルマリン固定酵母スフェロプラストをNP-40が含まれる緩衝液内で透過処理することで細胞膜を除き、次いでS-アデノシルメチオニン存在下でCCMTによるメチル化反応を適用することでクロマチン上のDNAのメチル化反応を実施した。ゲノムDNAを精製した後にこのDNAのメチル化状態をWGBSで分析した。

40

【0071】

10-bpピンのメチル化レベルを計算し、GCジヌクレオチドを認識する市販のM.CviPI(GCMT)で処理した核とCCMTで処理した核のDNAのメチル化状態を比較した結果、これら2つのデータセットは同様の5mCのゲノム上の分布パターンを示した。なお、出芽酵母には内在性のDNAメチル化が無いため、検出されたD5mCは全てGCMTあるいはCCMTにより導入されたものであると解釈できる。図5A-5Bに示すように、同様の傾向がゲノムDNAメチル化パターンにおいて観察された。CCMTシグナルとGCMTシグナルの凝集プロットは、いくつかの類似したゲノムの特徴を中心にしていた(図5C-5D、及び図10A)。遺伝子のプロモーター領域に鋭いDNAメチル化ピークを観察した(図5B)。これらの強いDNAメチル化シグナルは、ATAC-Seqを用いて検出されたシグナルとよく共局在した。したがって、これらのシグナルは開い

50

たクロマチン領域を示すと考えられた。CCMT及びGCMTベースのNOMe-Seq及びATAC-Seqで検出されたピークはよく共局在した(図10B)。

【0072】

CC及びGCジヌクレオチドにおけるDNAメチル化レベルは、プロモーター近位領域でよく凝集した(図5C-5D)。これらのシグナルは、ヌクレオソームの結合が無い、もしくはその結合が弱い領域として遺伝子プロモーターで頻繁に観察されるヌクレオソームフリー領域であることは疑いようが無い。転写開始部位の-100から-250領域のプロモーター近位DNAメチル化シグナルは、MNase-Seqを介して検出される低シグナル領域と逆相関していた(図5E)。転写開始部位で観察されたDNAメチル化のシグナルは160~170bpの周期性を示し、MNase-Seqによって観察されたシグナルと逆相関した(図5B,5C-5E)。これらの観察は、既報のNOMe-SeqとMNase-seq(参照文献23)の特性を再現していた。

10

【0073】

CCMTベースのNOMe-Seqは哺乳類エピゲノムの微細マッピングを達成する

次に、CCMTベースのNOMe-Seqをヒト由来培養細胞株であるIMR-90細胞に適用した。NP-40を含む緩衝液での処理で細胞膜を透過処理した後に、CCMT処理を実施し、ゲノムDNAを精製してこれにWGBS解析を適用した。また、CCMT処理を行わないサンプルも用意し、これをコントロールとした。CCMT処理の有無のサンプルでそれぞれ25.7倍及び24.5倍のカバー率を有する高品質のメチロームデータを得た(表1)。

【0074】

【表 1】

## IMR90 細胞を対象とした NOMe-Seq 解析におけるメチロームデータの基本統計値

			Control	CCMT-treated	GCMT-treated
Accession numbers at NCBI GEO			GSM5769478	GSM5769479	GSM6111856
Mapping	Read 1	Number of reads	462,135,990	445,678,391	447,579,538
		Unique map	73.00%	72.80%	69.5%
		Unmap	24.20%	24.70%	27.9%
	Read 2	Number of reads	462,135,988	445,678,391	447,579,537
		Unique map	65.10%	64.90%	63.6%
		Unmap	32.60%	32.90%	34.0%
Mapped read coverage		Top	12.9	12.3	11.4
		Bottom	12.8	12.2	11.3
		Both	25.7	24.5	22.7
		N	12.9	12.3	11.3
		C	14.2	13.4	12.4
		CpG	14.2	13.1	12.0
		CHG	14.4	13.6	12.8
		CHH	14.1	13.4	12.5
Methylation level		C	3.4	3.9	5.4
		CpG	56.2	56.2	58.0
		CHG	1	1.2	3.1
		CHH	0.8	1.5	2.8
Coverage vs read depth		>=1	86.2%	86.1%	85.7%
		>=3	84.0%	83.8%	82.2%
		>=5	80.9%	80.6%	76.9%
		>=10	64.8%	63.2%	55.5%

## 【 0 0 7 5 】

予想通り、CG部位における固有のDNAメチル化レベルは、CCMT処理細胞と未処理細胞の間でほとんど同一であった(図6A-6B)。その一方で、CC部位のDNAメチル化レベルはCCMT処理核でのみ増加した(図6C-6D)。これらの結果は、CCジヌクレオチドでの内因性CGメチル化とCCMT誘導メチル化が明示的に分離できることを示した。

40

## 【 0 0 7 6 】

次に、ゲノムブラウザを用いて2つのデータセットを視覚的に比較した。図6Eに示すように、両データセットで固有のDNAメチル化レベルが異なるいくつかの大きなドメインを観察した。これらのドメインは、疑いなく高度メチル化領域と中等度メチル化ドメインであり、その存在はIMR-90細胞について記載されている(参照文献24)。CCMTベースのNOMe間の比較Seqと無処理対照は、これら2つのデータセットがCGジヌクレオチドを対象とした哺乳類細胞の内在性メチル化状態は、ほぼ同じ状態で観察されることを示していた。その一方で、CCMTで処理したサンプルではCCメチル化の強い鋭いメチル化ピークが観察されたのに対して、無処理対照では観察されなかった(図6E)。また、これらのピークはATAC-Seqを用いて同定したピークと大部分が共局在していた。したがって、CCMT誘導メチル化シグナ

50

ルで観察されたピークは、高度にアクセシブルなクロマチン領域を示した。(図6E)これらの結果は、CCMTに基づくNOMe-Seqが哺乳類細胞の固有のメチロームとクロマチンアクセシビリティの両方を同時に測定できることを強く支持していた。

【 0 0 7 7 】

### 3 . 考察

エピゲノムマッピングのツールとしてMTaseの需要が増加しているにもかかわらず、特に短い認識配列のシトシンC-5MTaseについては利用可能な酵素が限られている。本実施例では、M.CviQIXとM.CviPIIのメチル化活性がそのN末端側数残基のアミノ酸配列によって阻害されていることを明らかにし、逆にそのN末端欠損株は高いDNAメチル化活性を有することを示した。これらの人工変異タンパク質は精製可能であり、in vitroでも活性が維持されていた。また、シトシンC-5MTaseとしては新規の認識配列であるCCジヌクレオチドを標的とすることも明らかにした。

このCCジヌクレオチドは、哺乳類内在性のMTaseの認識配列であるCGジヌクレオチドとは決して重なることはないことから、哺乳類の内在性MTaseに加えて追加で任意のエピゲノム状態を同時検出のためのエピジェネティックマッピングにおけるプローブとして大きな利点があると考えられた。本実施例ではNOMe-Seqの実現によりそのことを実証した。

【 0 0 7 8 】

以下に、本明細書中に掲載する各々の参考文献を列挙する。

1. Greenberg, M.V.C. and Bourc'his, D. (2019) The diverse roles of DNA methylation in mammalian development and disease. *Nature Reviews Molecular Cell Biology*, 20, 590 607.
2. Edwards, J.R., Yarychkivska, O., Boulard, M. and Bestor, T.H. (2017) DNA methylation and DNA methyltransferases. *Epigenetics & Chromatin*, 10, 23.
3. Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, 13, 484 492.
4. Petryk, N., Bultmann, S., Bartke, T. and Defossez, P. A. (2021) Staying true to yourself: mechanisms of DNA methylation maintenance in mammals. *Nucleic Acids Res.*, 49, 3020 3032.
5. Jones, P.A. and Liang, G. (2009) Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet*, 10, 805 811.

【 0 0 7 9 】

6. Edwards, J.R., Yarychkivska, O., Boulard, M. and Bestor, T.H. (2017) DNA methylation and DNA methyltransferases. *Epigenetics Chromatin*, 10, 23.
7. Tajima, S., Suetake, I., Takeshita, K., Nakagawa, A. and Kimura, H. (2016) Domain Structure of the Dnmt1, Dnmt3a, and Dnmt3b DNA Methyltransferases. *Adv. Exp. Med. Biol.*, 945, 63 86.
8. Jang, H.S., Shin, W.J., Lee, J.E. and Do, J.T. (2017) CpG and Non CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes*, 8, 148.
9. Patil, V., Ward, R.L. and Hesson, L.B. (2014) The evidence for functional non CpG methylation in mammalian cells. *Epigenetics*, 9, 823 828.
10. van de Lagemaat, L.N., Flenley, M., Lynch, M.D., Garrick, D., Tomlinson, S.R., Kranc, K.R. and Vernimmen, D. (2018) CpG binding protein (CFP1) occupies open chromatin regions of active genes, including enhancers and non CpG islands. *Epigenetics & Chromatin*, 11, 59.

【 0 0 8 0 】

11. Teissandier, A. and Bourc'his, D. (2017) Gene body DNA methylation conspires with H3K36me3 to preclude aberrant transcription. *The EMBO Journal*, 36, 1471 1473.

10

20

30

40

50

12. Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S. and Jeltsch, A. (2010) The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *The Journal of biological chemistry*, 285, 26114–26120.
13. Solomon, M.J., Larsen, P.L. and Varshavsky, A. (1988) Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell*, 53, 937–947.
14. Johnson, D.S., Mortazavi, A., Myers, R.M. and Wold, B. (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science*, 316, 1497–1502.
15. Kelly, T.K., Liu, Y., Lay, F.D., Liang, G., Berman, B.P. and Jones, P.A. (2012) Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res.*, 22, 2497–2506.

10

## 【 0 0 8 1 】

16. Lee, I., Razaghi, R., Gilpatrick, T., Molnar, M., Gershman, A., Sadowski, N., Sedlazeck, F.J., Hansen, K.D., Simpson, J.T. and Timp, W. (2020) Simultaneous profiling of chromatin accessibility and methylation on human cell lines with nanopore sequencing. *Nat Methods*, 17, 1191–1199.
17. Stergachis, A.B., Debo, B.M., Haugen, E., Churchman, L.S. and Stamatoyannopoulos, J.A. (2020) Single-molecule regulatory architectures captured by chromatin fiber sequencing. *Science*, 368, 1449–1454.
18. Steensel, B.v. and Henikoff, S. (2000) Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. *Nat. Biotechnol.*, 18, 424–428.
19. Roberts, R.J. and Macelis, D. (1997) REBASE: restriction enzymes and methylases. *Nucleic Acids Res.*, 25, 248–262.
20. Roberts, R.J., Vincze, T., Posfai, J. and Macelis, D. (2015) REBASE: a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.*, 43, D298–D299.

20

## 【 0 0 8 2 】

21. Xu, M., Klädde, M.P., Van Etten, J.L. and Simpson, R.T. (1998) Cloning, characterization and expression of the gene coding for a cytosine-5 DNA methyltransferase recognizing GpC. *Nucleic Acids Res.*, 26, 3961–3966.
22. Malakhov, M.P., Mattern, M.R., Malakhova, O.A., Drinker, M., Weeks, S.D. and Butt, T.R. (2004) SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *Journal of Structural and Functional Genomics*, 5, 75–86.
23. Klemm, S.L., Shipony, Z. and Greenleaf, W.J. (2019) Chromatin accessibility and the regulatory epigenome. *Nature Reviews Genetics*, 20, 207–220.
24. Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filipini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M. et al. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462, 315–322.
25. Braberg, H., Echeverria, I., Bohn, S., Cimermancic, P., Shiver, A., Alexander, R., Xu, J., Shales, M., Dronamraju, R., Jiang, S. et al. (2020) Genetic interaction mapping informs integrative structure determination of protein complexes. *Science*, 370.

30

## 【 0 0 8 3 】

26. Miura, F., Fujino, T., Kogashi, K., Shibata, Y., Miura, M., Isobe, H. and Ito, T. (2018) Triazole linking for preparation of a next-generation sequencing library from single-stranded DNA. *Nucleic Acids Res.*, 46, e95.

40

50

【配列表フリーテキスト】

【0084】

配列番号5：組換えDNA

配列番号6：組換えタンパク質

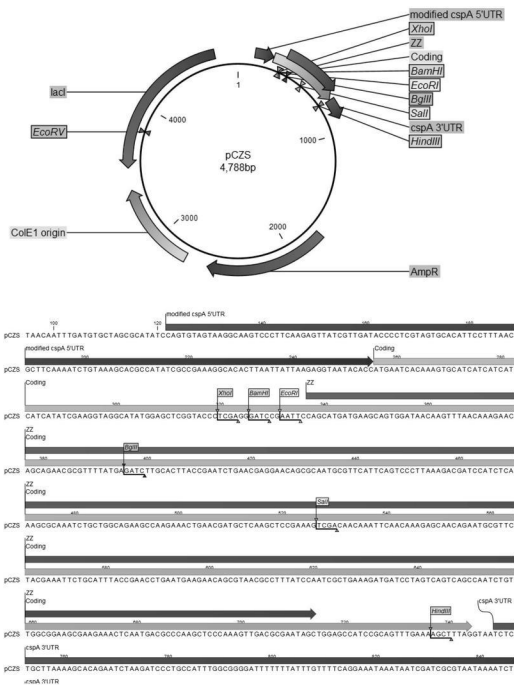
配列番号7：組換えDNA

配列番号8：組換えタンパク質

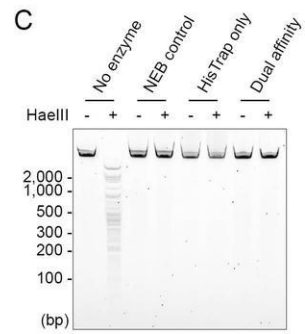
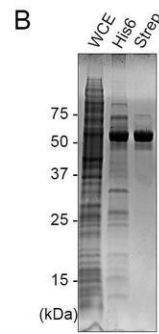
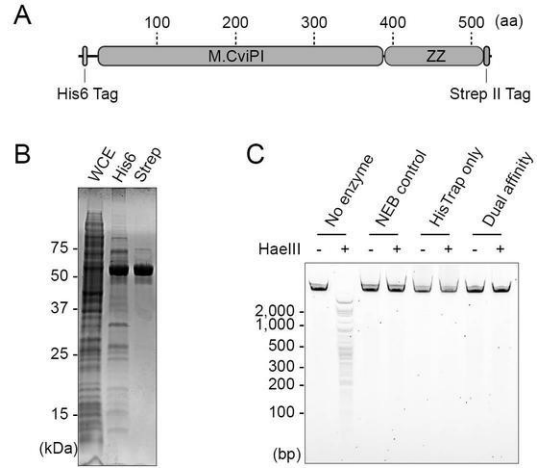
配列番号9：合成DNA

配列番号10：合成DNA

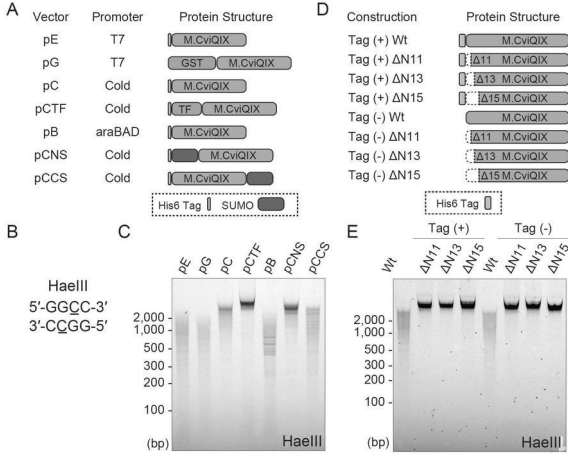
【図1】



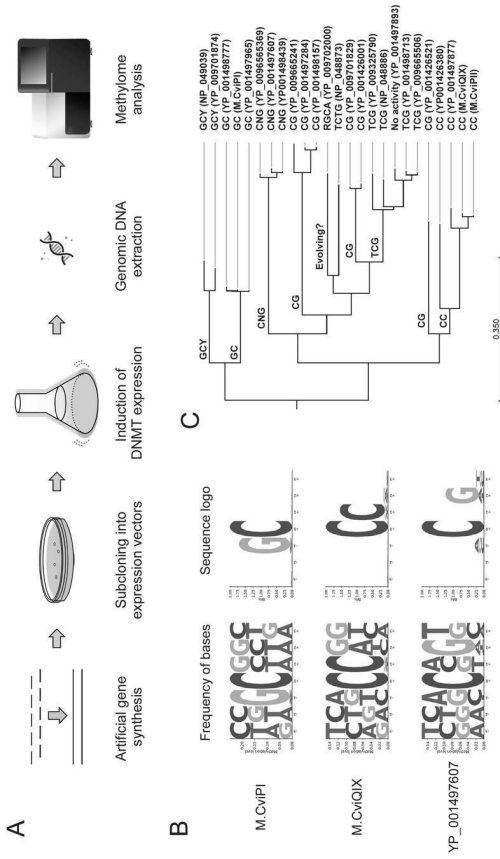
【図2】



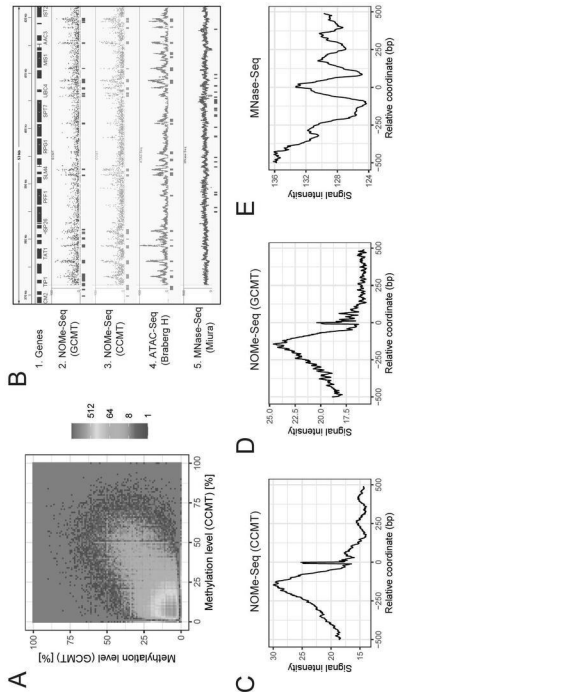
【 3 】



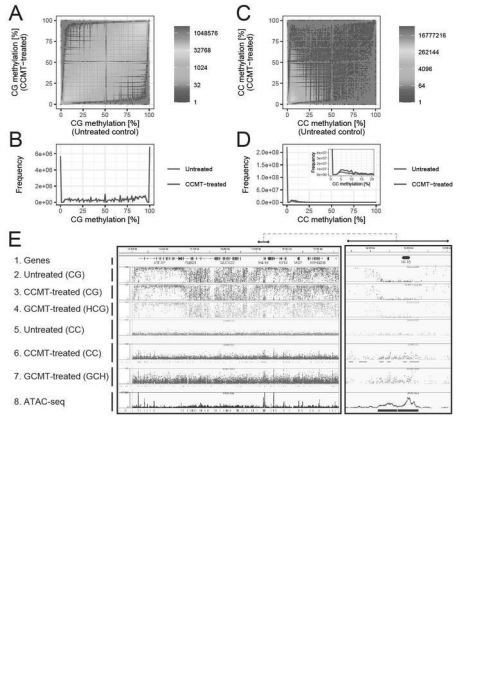
【 4 】



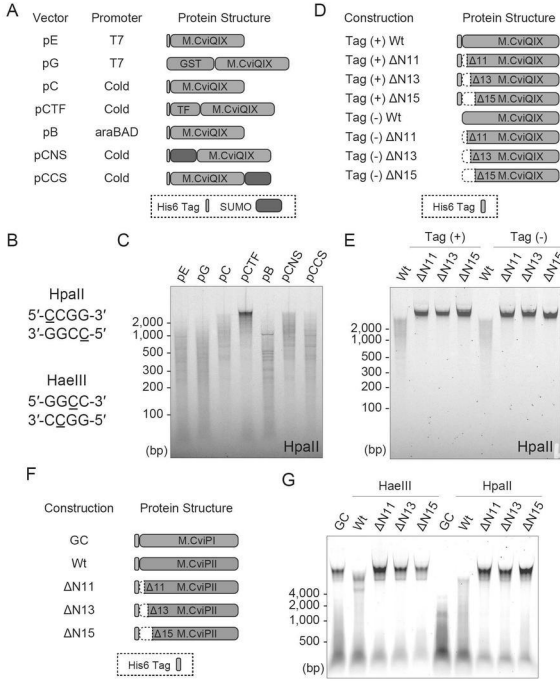
【 5 】



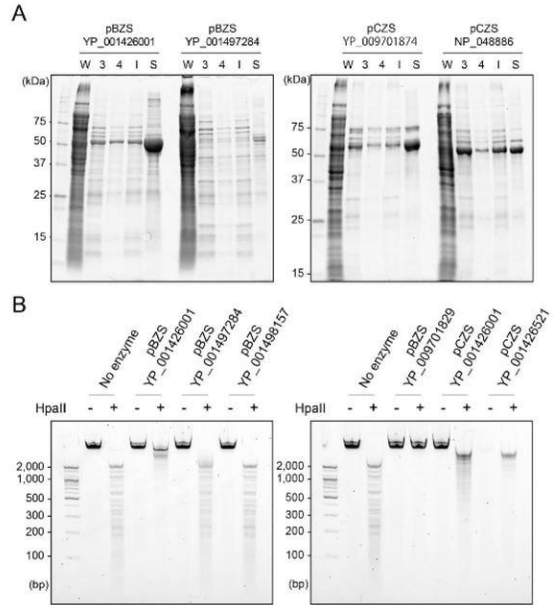
【 6 】



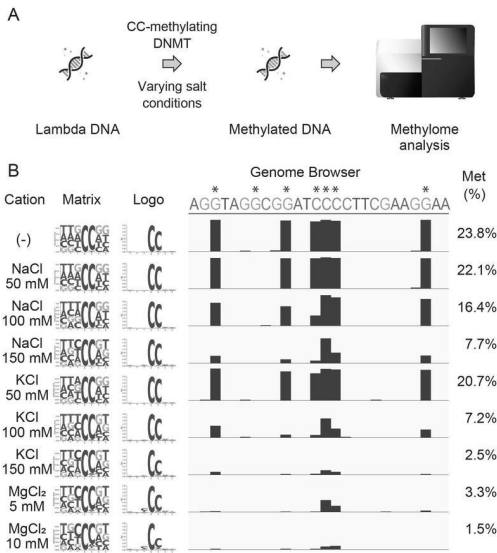
【 7 】



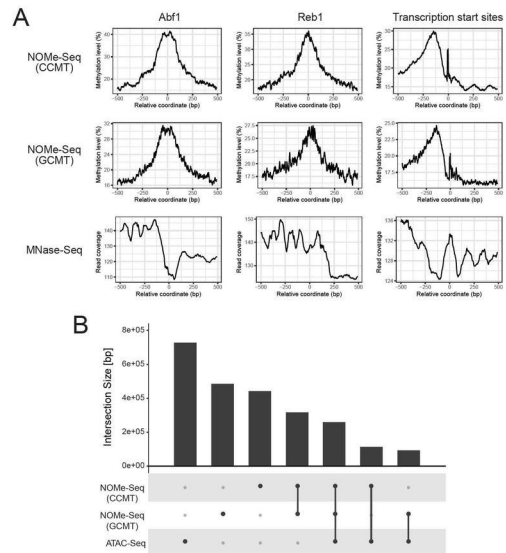
【 8 】



【 9 】



【 10 】



【配列表】

2023175516000001.app

---

フロントページの続き

(51)Int.Cl.			F I		テーマコード(参考)
C 1 2 N	1/21	(2006.01)	C 1 2 N	1/21	
C 1 2 N	5/10	(2006.01)	C 1 2 N	5/10	
C 1 2 Q	1/48	(2006.01)	C 1 2 Q	1/48	
C 1 2 Q	1/68	(2018.01)	C 1 2 Q	1/68	

Fターム(参考) 4B065 AA01X AA01Y AA57X AA57Y AA72X AA72Y AA83X AA83Y AA90X AA90Y AB01 AC14 BA02 CA29  
CA46 CA60