

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4350800号  
(P4350800)

(45) 発行日 平成21年10月21日(2009.10.21)

(24) 登録日 平成21年7月31日(2009.7.31)

(51) Int.Cl.

F I

C 1 2 N 5/10 (2006.01)

C 1 2 N 5/00 B

C O 7 K 14/075 (2006.01)

C O 7 K 14/075

C 1 2 N 15/09 (2006.01)

C 1 2 N 15/00 Z N A A

A 6 1 K 31/711 (2006.01)

A 6 1 K 31/711

A 6 1 K 35/76 (2006.01)

A 6 1 K 35/76

請求項の数 7 (全 81 頁) 最終頁に続く

(21) 出願番号 特願平10-515273  
 (86) (22) 出願日 平成9年9月24日(1997.9.24)  
 (65) 公表番号 特表2001-505047(P2001-505047A)  
 (43) 公表日 平成13年4月17日(2001.4.17)  
 (86) 国際出願番号 PCT/EP1997/005251  
 (87) 国際公開番号 W01998/013499  
 (87) 国際公開日 平成10年4月2日(1998.4.2)  
 審査請求日 平成16年9月22日(2004.9.22)  
 (31) 優先権主張番号 08/719,806  
 (32) 優先日 平成8年9月25日(1996.9.25)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 597011463  
 ノバルティス アクチエンゲゼルシャフト  
 スイス国、4056 バーゼル、リヒトシ  
 ユトラーセ 35  
 (73) 特許権者 593052785  
 ザ スクリップス リサーチ インスティ  
 テュート  
 アメリカ合衆国 カリフォルニア州 92  
 037 ラ ジョラ ノース トーリー  
 パインス ロード 10550  
 (74) 代理人 100081422  
 弁理士 田中 光雄  
 (74) 代理人 100062144  
 弁理士 青山 葆

最終頁に続く

(54) 【発明の名称】 高キャパシティーアデノウイルスベクターの開発の促進に使用するためのパッケージング細胞系

## (57) 【特許請求の範囲】

## 【請求項 1】

アデノウイルスファイバータンパク質を発現するアデノウイルスパッケージング細胞系であって、アデノウイルスファイバータンパク質をコードし、第1の異種プロモーターに操作可能に結合したヌクレオチド配列を含む安定に統合された核酸配列(ここで、当該第1の異種プロモーターは導入アデノウイルス(A d) 2またはA d 5の3つにわかれたリーダー(T P L)をコードするヌクレオチド配列を含む)およびアデノウイルスE 4遺伝子をコードする核酸に操作可能に結合した第2の安定に統合された異種プロモーターを含む、パッケージング細胞系:

ここで、当該パッケージング細胞系は、32.5の許容温度での293細胞によるウイルス産生よりも2~3倍のレベルに、39.5でウイルス産生を支持する。

## 【請求項 2】

アデノウイルスE 4遺伝子をコードするヌクレオチド配列が、A d 5ゲノムのヌクレオチド32667~35780によってコードされる、請求項1記載のパッケージング細胞系。

## 【請求項 3】

アデノウイルスファイバータンパク質をコードするヌクレオチド配列が、A d 5ゲノムのヌクレオチド31040~32791を含む、請求項1記載のパッケージング細胞系。

## 【請求項 4】

さらにウイルスベクターを含む、請求項2または3記載のパッケージング細胞系であって

10

20

、当該細胞系において産生された前記ウイルスベクターが、下記制御タンパク質またはポリペプチド：E 2 A、E 2 BもしくはE 3の1つ以上をコードするDNA配列の欠失または変異を有する核酸配列を含み、当該欠失または変異が初期領域ポリペプチドの発現を阻止し、そして当該パッケージング細胞系が欠失または変異制御タンパク質またはポリペプチドの機能を補っている、パッケージング細胞系。

【請求項5】

外来タンパク質、ポリペプチドまたはそのフラグメントの1つ以上をコードする外来DNA配列が任意の欠失部位に挿入されている、請求項4記載のパッケージング細胞系。

【請求項6】

外来DNAが、腫瘍抑制タンパク質またはその生物学的に活性なフラグメントをコードする、請求項5記載のパッケージング細胞系。

【請求項7】

外来DNAが、自殺タンパク質またはその生物学的に活性なフラグメントをコードする、請求項5記載のパッケージング細胞系。

【発明の詳細な説明】

本発明は、NIH Grant No. HL 54352の下、米国政府の援助によりなされた。政府は本発明に一定の権利を有する。

本発明は、遺伝子治療、特にアデノウイルス-ベースの遺伝子治療に関する。特に、高キャパシティーベクターの開発の促進に使用するための、新規パッケージング細胞系を記載する。関連する組成物、キットとしての高キャパシティーアデノウイルスベクター、および記載のベクター、細胞系およびキットの製造法および使用がまた記載される。

細胞内へのDNA接合体の移入の促進が、上皮細胞に容易に感染するヒトDNAウイルスであるアデノウイルスで達成されている(Horwitz, "Adenoviridae and their replication", Virology, Fields and Knipe, eds., Raven Press, NY (1990) pp. 1679-1740)。

アデノウイルス介在遺伝子治療が、細胞内へのDNA移入の改善された方法を代表するが、この試みの可能性のある限界は、アデノウイルス複製の結果宿主細胞の破壊に至ることである。加えて、アデノウイルスはまた、そのタンパク質の一つが腫瘍抑制遺伝子生産物と結合することを含む発癌性特性を有する。アデノウイルスのいわゆる複製欠失株(これは、典型的に、宿主細胞内でウイルスが複製できなくなる、E 1 Aおよび/またはE 1 B欠失を保持する)の使用は、インビボ治療に原則としてより適している；しかしながら、組換えウイルスのトランス活性化をもたらすウイルスの野性型株での上皮細胞の共感染の可能性は、インビボ適用に関する著しい安全性懸念を提示し得る。更に、どの組換えアデノウイルスが、それらが運搬し得る外来遺伝子の長期間の安定発現を可能にするように、そのゲノムを宿主細胞DNAに統合をできるかまだ知られていない。

そのままのまたは複製コンピテントアデノウイルスの遺伝子移入手段としての使用の他の望ましくない態様は、その遺伝子生産物が宿主細胞腫瘍抑制タンパク質および主要組織適合遺伝子複合体(MHC)のような免疫認識分子を妨害することが知られている発癌ウイルスであるということである。加えて、アデノウイルスへの先在循環抗体がインビボ遺伝子運搬の効果を有意に減少し得る。最後に、6キロベース(kb)またはそれ以下の外来遺伝子しか、遺伝子移入実験ではそのままのアデノウイルスゲノム内に包含できないが、本発明の方法を使用して、15 kbより大きいDNAセグメントが移入できる。

Adベクターをより複製インコンピテントにするために、ある研究者は、ウイルス粒子のパッケージングに必要であることが知られている部分以外、そのゲノムのほとんど総てが欠失した組換えAd由来ベクターの構築を試みている。例えば、全ウイルスORFを欠失するが、本質的cis要素(逆方向末端反復---ITRs---および隣接パッケージング配列)を含むヘルパー依存的ベクターが構築されているが、ウイルス粒子はヘルパーよりパッケージの効率が劣り、時としてマルチマーとしてパッケージする。これはウイルスが完全DNA補体のパッケージを“望み”得ることを示す(例えば、Fisher, et al., Virology 217:11-22 (1996) 参照)。Mitani et al. (PNAS USA 92:3854-3858 (1995))はまた完全に複製欠失ではないように見えるヘルパー依存的Adベクターも記載する。

10

20

30

40

50

Amalfitano, et al. (PNAS USA 93:3352-3356 (1996)) は、インビボでの A d ベクターの複製の阻害の成果として、E 1 - およびポリメラーゼ 欠失 A d ベクターの生育を支持する A d パッケージング細胞系の構築を記載する。同様に、Armentano, et al. (Hum. Gene Ther. 6:1343-53 (1995)) は、ほとんど---ではあるが総てではない---の E 4 配列がそこから欠失された A d ベクターを記載する。しかしながら、このような少量の遺伝物質がベクターから欠失されたため、それらの治療的配列の運搬能力はむしろ制限される。

上記問題の観点から、本明細書に記載のパッケージング細胞系およびシステムの設計および構築は、更に本明細書に記載のように、新規でエレガントな解決を提供する。インビトロおよびインビボ両方での外来遺伝子の受容細胞への移入における本発明の組換え配列およびベクターの使用は、上記遺伝子移入システムの制限に打ち勝つ。本発明は、そのま

10

まのアデノウイルスビリオンの細胞受容体結合および DNA 運搬特性を二倍にし、従って遺伝子治療およびアンチセンス - ベースの抗ウイルス治療の改善された方法を示す。そのまの Aデノウイルス、ヘルパープラスミドまたはウイルスを必要とする修飾 Aデノウイルス、またはいわゆる複製欠失 Aデノウイルスの使用の欠点と比較して、本発明の組換え Aデノウイルス由来ベクターは遺伝子運搬にある利点を提供する。最初に、本発明の A d 由来ベクターは、上皮細胞受容体への結合およびエンドサイトーシス液胞の浸透を含む、遺伝子治療に必要な機能的特性の総てを含む。本発明の治療的ウイルスベクターは、非上皮細胞の受容体を標的とし、浸透を達成するように操作し得る；ウイルスベクターをこれらの目的のために操作する方法は下記に詳述する。

第 2 に、本発明のベクターは、A d ゲノムの実質的な部分を欠失し、これは A d 由来ベクターが他の宿主細胞または組織に“広がる”能力を制限するだけでなく、“外来”(または非天然)核酸の著しい量を、ウイルスゲノムの再生およびパッケージング無しでウイルスゲノムに包含させることを可能にする。従って、本発明のベクターは、広範囲の治療的適用への使用に理想的である。

20

第 3 に、種々のヒトの病気の処置の治療剤としての本明細書に記載のベクターの使用が安全であり、一方、主として、それらが再生される新規細胞系のために伝播およびパッケージングのための“ヘルパー”の存在を必要としない。このような細胞系---本明細書ではパッケージング細胞系と呼ぶ---は、本発明の更に別の態様を構成する。

野性型 Aデノウイルスによる汚染の頻度を減少するために、組換えの可能性を減少させるためのウイルスベクターまたは細胞系のいずれかを改善することが望まれる。例えば、グループ C ウイルスに対する相同性が小さいグループ由来の Aデノウイルスを使用して、293 細胞の A d 5 配列との組換えの傾向が小さい組換えウイルスを操作し得る。同様に、A d 3 由来の Aデノウイルスタンパク質またはポリペプチドおよび/または他の非グループ C またはグループ C 血清型由来のタンパク質またはポリペプチドを安定に発現する上皮細胞系---293 または他---を、本明細書に記載の方法に従って、製造し得る；このような細胞系は、このようなベクターが由来する血清型と無関係に、調節性および/または構造遺伝子の欠失を担持する Aデノウイルス由来ベクターの支持に有用である。

30

本発明の構築物および方法は、2 個またはそれ以上の A d 血清型由来のアミノ酸残基配列を発現するキメラウイルスベクターの設計および操作を支持することも意図される。従って、本発明の記載の出現より前に利用可能であった方法および構築物とは異なり、本発明はその構築および伝播を支持するウイルスベクターおよび細胞系の設計および製造の順応性の可能性を最大にする---総て、有害な組換え体を製造する可能性のある野性型 A d との組換えの危険性の減少を伴う。

40

一部において、本発明は、ウイルスと細胞性配列の間の組換えを、当分野で検討されているよりも減少させる、単純な、別の方法を記載する。一つのこのような手段は、組換えウイルスのサイズの増加であり、それによりウイルスとパッケージング細胞系に存在する A d 遺伝子、例えば、293 細胞系内の A d 5 遺伝子、または本発明の新規細胞系中の種々の A d 遺伝子との間の共有配列の程度を減少する。

Aデノウイルスの構造遺伝子の総てのまたは一部の欠失は、このような欠失によるウイルス再生およびパッケージングにおける予測される有害な作用のため、望ましくないと考え

50

られている。実際、“ヘルパー”ウイルスまたはプラスミドの使用は、構造タンパク質配列に大きな欠失を含むA d由来ベクターを使用する時、まさにこの理由のために、しばしば推奨される。

当分野で示しているものとは反対に、本発明は、A d構造タンパク質をコードする種々の遺伝子配列のすべてまたは一部を欠失した組換えA d由来ベクターの製造、伝播および使用を、ウイルス製造における野性型アデノウイルス汚染の危険性の減少の手段および、このようなベクターへの外来DNAの種々の診断的および治療的適用のためのパッケージを可能にする手段の両方として、記載および請求する。

従って、本発明の一つの態様において、一個またはそれ以上のアデノウイルス調節性ポリペプチドをコードするDNA配列および一個またはそれ以上のアデノウイルス構造ポリペプチドをコードするDNA配列が、細胞性ゲノムに安定に統合されているパッケージング細胞系を記載する。

10

従って、本発明の他の態様において、一個またはそれ以上のアデノウイルス構造タンパク質、ポリペプチド、またはそのフラグメントを発現し、該構造タンパク質が：

- a. ペントン塩基；
- b. ヘキソン；
- c. ファイバー；
- d. ポリペプチドIIIa；
- e. ポリペプチドV；
- f. ポリペプチドVI；
- g. ポリペプチドVII；
- h. ポリペプチドVIII；そして
- i. それらの生理学的に活性なフラグメント

20

からなる群から選択されるパッケージング細胞系を記載する。

一つの変法において、配列は構造に発現される；他には、一個またはそれ以上の配列が調節可能プロモーターの制御下にある。好ましい態様において、発現は構造である。種々の好ましい態様において、DNA配列により発現されるポリペプチドは生理学的に活性である。

別のそして好ましい態様において、本発明のパッケージング細胞系はウイルスベクターの製造を支持する。好ましい態様において、ウイルスベクターは治療的ベクターである。

30

本発明の一つの態様において、各DNA配列を本明細書に記載の細胞系のゲノム内に、別の相補的プラスミドを介して挿入する。一つの変法において、相補的プラスミドは、アデノウイルスファイバータンパク質、ポリペプチドまたはそのフラグメントをコードするDNA配列を含む。本発明の有用な相補的プラスミドの例は、pCLFの特徴を有するプラスミドである（寄託の詳細に関しては実施例3参照）。

本発明の他の態様において、本発明の細胞系を形質転換するのに使用する相補的プラスミドは、更に、アデノウイルス調節タンパク質、ポリペプチドまたはそのフラグメントをコードするDNA配列を含む。一つの変法において、調節タンパク質はE1A、E1B、E2A、E2B、E3、E4およびL4（“100Kタンパク質”とも呼ばれる）からなる群から選択される；例示的相補的プラスミドは、pE4/Hygro??（寄託の詳細に関しては実施例3参照）の特徴を有する。他の態様において、本発明の細胞系の形質転換に使用する相補的プラスミドは、更に、2個またはそれ以上の上記のアデノウイルス調節タンパク質、ポリペプチドまたはそのフラグメントをコードするDNA配列を含む。

40

一つの変法において、2個またはそれ以上の調節タンパク質、ポリペプチドまたはそのフラグメントは、E1A、E1B、E2A、E2B、E3、E4およびL4（“100Kタンパク質”とも呼ばれる）からなる群から選択される。他の変法において、構造タンパク質はペントン塩基；ヘキソン；ファイバー；ポリペプチドIIIa；ポリペプチドV；ポリペプチドVI；ポリペプチドVII；ポリペプチドVIII；およびこれらの生理学的に活性なフラグメントからなる群から選択される。

本発明の一つの変法において、パッケージング細胞系はファイバータンパク質を示す。一

50

つの態様において、ファイバータンパク質は、特異的受容体を標的とするが、3量体形成または核内へのファイバーの運搬を妨害しない、非天然アミノ酸残基配列を含むように調節する。他の変法において、非天然アミノ酸残基は、ファイバーの標的細胞型への結合特異性を変える。更に別の態様において、構造タンパク質は、一つ以上のアデノウイルス血清型由来のアミノ酸残基配列を含むファイバーである。本明細書に記載のように、ファイバータンパク質またはポリペプチドをコードするヌクレオチド配列は、1端または両端を単に修飾する必要はない；ファイバータンパク質---および実際、本明細書で教示の任意のアデノウイルス構造タンパク質---を、“内部的に”そして末端で修飾し得る。

本発明はまた、その細胞系内で製造されるウイルスベクターが、アデノウイルス構造タンパク質、ポリペプチド、またはそのフラグメントをコードするDNA配列の欠失または変異を有する核酸配列を含む、パッケージング細胞系も記載する。一つの変法において、ウイルスベクターは、更に、調節ポリペプチドE1AおよびE1BをコードするDNA配列の欠失または変異を有する核酸配列を含む。他の変法において、ウイルスベクターは、更に、一個またはそれ以上の以下の調節タンパク質またはポリペプチド；E2A、E2B、E3、E4、L4またはそれらのフラグメントをコードするDNA配列の欠失または変異を有する核酸配列を含む。

更に別の変法は、一個またはそれ以上の外来タンパク質、ポリペプチドまたはそのフラグメントをコードする外来DNA配列が、治療的ウイルスベクター内の欠失の代わりに挿入されていることを記載する。一つの態様において、外来DNAは腫瘍抑制タンパク質または生理学的に活性なそのフラグメントをコードする。他の態様において、外来DNAは自殺タンパク質または生理学的に活性なそのフラグメントをコードする。前記のように、本明細書に記載の細胞系は、起源が原核または真核であり得、哺乳類細胞系がしばしば好ましい。上皮および非上皮細胞系が前記の変法に有用である；特に有用な細胞系は293、A549、W162、HeLa、Vero、211および211A細胞系を含む。

本発明は、更に、一個またはそれ以上の外来タンパク質、ポリペプチドまたはそれらのフラグメントをコードする外来DNA配列を含むウイルスベクターの製造を支持する前記細胞系が、欠失した任意の構造および/または調節タンパク質（またはそれらの一部）の代わりに挿入されることを意図する。従って、一つの態様において、外来DNAは腫瘍抑制タンパク質；自殺タンパク質；嚢胞性線維症膜コンダクタンス制御因子（CFTR）タンパク質；またはこれらの生理学的に活性なフラグメントをコードする。

本明細書に記載の細胞系は、ゲノム内に安定に統合したファイバータンパク質---修飾またはキメラタンパク質を含む---の総てまたは一部をコードするDNAを有し得る。従って、一つの変法において、ファイバータンパク質を、特異的受容体を標的とするが、3量体形成またはファイバータンパク質の核内への運搬を妨害しない非天然アミノ酸残基配列を含むように修飾する。一つの変法において、非天然アミノ酸残基配列は、ファイバーのカルボキシル末端に結合する。更に別には、非天然アミノ酸残基配列は、更にリンカー配列を含む。あるいは、ファイバータンパク質は更にリンカーに結合したリガンドを含む。適当なリガンドは、細胞表面受容体に選択的に結合するリガンドおよび他のタンパク質または核酸分子との結合に使用できるリンカーからなる群から選択され得る。一つの変法において、リガンドは細胞表面受容体に選択的に結合するリガンドおよび他のタンパク質または核酸分子との結合に使用できるリンカーからなる群から選択される。

更に別の態様において、非天然アミノ酸残基配列を、一つのファイバー末端以外の位置でファイバーアミノ酸残基配列に挿入させる。あるいは、非天然アミノ酸残基配列は、ファイバーの結合特性を他の標的細胞型に変える。発現したファイバーは、種々の態様において、通常アデノウイルスの標的とならない特異的標的細胞型に結合し得るかおよび/または一つ以上のアデノウイルス血清型に由来するアミノ酸残基を含み得る。

本発明の種々の態様において、本発明のパッケージング細胞系は、原核細胞系由来である；他には、それは真核細胞系由来である。種々の態様が哺乳類細胞、およびより特には上皮細胞系の使用を示唆するが、種々の他の、非上皮細胞系を種々の態様で使用する。従って、種々の態様が293、A549、W162、HeLa、Vero、211および211A細胞系か

10

20

30

40

50

らなる群から選択される細胞系の使用を記載するが、種々の他の細胞系が同様に本明細書に記載の使用に意図される。

本発明は、更に、広範囲の種々の核酸配列およびウイルスベクターを記載する。従って、一つの態様において、本発明は上記のアデノウイルスファイバータンパク質、ポリペプチドまたはそれらのフラグメント---限定しないが、欠失または変異を含むもの；キメラのもの；およびリンカー、外来アミノ酸残基、または本明細書に記載の種々の目的で結合した他の分子を有するものを含む---の一つをコードする核酸配列を記載する。種々の他のアデノウイルス構造および/または調節タンパク質またはポリペプチドをコードする核酸配列はまた本発明の範囲内である。

広範囲の治療的ウイルスベクターは、また本発明の態様である。一つの態様において、ファイバータンパク質をコードするDNA配列またはその一部を欠く治療的ウイルスベクターを記載する。他の変法において、治療的ウイルスベクターは、更に、またはそれとは別に、一個またはそれ以上の調節タンパク質、ポリペプチドまたはそれらのフラグメントを欠失し得る。種々の態様において、外来DNA配列は、本発明のウイルスベクターのファイバータンパク質コードDNA配列に代わって挿入される。他の態様において、治療的ウイルスベクターは、更に、一個またはそれ以上の調節タンパク質、ポリペプチドまたはそれらのフラグメントおよび/または一個またはそれ以上の構造タンパク質、ポリペプチドまたはそれらのフラグメントをコードするDNA配列に代わって外来DNA配列を含む。本発明は多くのウイルスベクターを記載する。一つの変法において、ウイルスベクターは、アデノウイルス構造タンパク質、ポリペプチドまたはそれらのフラグメントをコードするDNA配列の欠失または変異を含む。ベクターは、更に、調節ポリペプチドE1AおよびE1BをコードするDNA配列の欠失または変異を含む；そして、それは更に1個またはそれ以上の以下の調節タンパク質またはポリペプチド：E2A、E2B、E3、E4、L4またはそれらのフラグメントをコードするDNA配列の欠失または変異も含み得る。他の変法において、本発明のウイルスベクターにおいて、構造タンパク質はファイバーを含み得る。前記の任意の組み合わせがまた本発明で意図される。本発明のウイルスベクターは、本明細書に記載の治療的ウイルスベクター---それらの組み合わせを含む---を含む医薬組成物の製造に適していることもまた記載する。本発明のウイルスベクターの更なる使用は、異なる細胞型を含む細胞集団中の特異的細胞の標的化である。

本発明は、更に、相補的プラスミドおよびその製造法も意図する。一つの態様において、相補的プラスミドは、アデノウイルス構造ポリペプチドをコードするヌクレオチド配列に操作可能に結合したプロモーターヌクレオチド配列を含む。一つの変法において、相補的プラスミドはpCLFを含む。他の変法において、相補的プラスミドは、更に、第1のアデノウイルス調節ポリペプチドをコードするヌクレオチド配列、第2の調節ポリペプチドをコードするヌクレオチド配列、第3の調節ポリペプチドをコードするヌクレオチド配列；または前記の組み合わせを含む。更に別の態様において、アデノウイルス構造ポリペプチドは、ペントン塩基；ヘキソン；ファイバー；ポリペプチドIIIa；ポリペプチドV；ポリペプチドVI；ポリペプチドVII；ポリペプチドVIII；およびそれらの生理学的に活性なフラグメントからなる群から選択される。

本発明はまた、アデノウイルス構造タンパク質、ポリペプチドまたはそれらのフラグメントをコードするヌクレオチド配列およびアデノウイルス調節タンパク質、ポリペプチドまたはそれらのフラグメントをコードするヌクレオチド配列に操作可能に結合したプロモーターヌクレオチド配列を含む相補的プラスミドも記載する。また他の変法において、初期領域ポリペプチドはE4である；他において、プラスミドはpE4/Hygroを含む。更に別の変法において、初期領域ポリペプチドはE1およびE4である。更に、プロモーターヌクレオチド配列がMMTV、CMVおよびE4プロモーターヌクレオチド配列からなる群から選択されるプラスミドとして、アデノウイルス構造タンパク質、ポリペプチドまたはそれらのフラグメントを含む相補的プラスミドがまた意図される。

パッケージングシグナルおよび外来タンパク質またはポリペプチドをコードするヌクレオチド配列を含み、アデノウイルス構造タンパク質をコードするヌクレオチド配列は欠失し

10

20

30

40

50

ているウイルスベクターがまた記載される。一つの変法において、外来タンパク質またはポリペプチドをコードするヌクレオチド配列は約3 kb長までである；他に、外来タンパク質またはポリペプチドをコードするヌクレオチド配列は約9.5 kb長までである；更に別には、外来タンパク質またはポリペプチドをコードするヌクレオチド配列は約12.5 kb長までである。中間の長さのヌクレオチド配列はまた本発明により意図され、配列は12.5 kbを超える。

本発明は、また外来タンパク質またはポリペプチドが抗腫瘍剤、腫瘍抑制タンパク質、自殺タンパク質、またはそらのフラグメントまたは機能的相同物をコードする配列であるウイルスベクターも記載する。一つの変法において、一個またはそれ以上の調節タンパク質をコードするヌクレオチド配列はまたベクターから欠失される。他には、調節タンパク質は、E1A、E1B、E2A、E2B、E3、E4およびL4(100Kタンパク質)からなる群から選択される。

種々の態様において、アデノウイルスは、血清型1、2、5または6から選択されるグループCアデノウイルスである；他の態様において、他の血清型から選択されるアデノウイルスが本明細書で記載のように有用である。本発明はまた前記具体例に従ったウイルスベクター、および製薬学的に有用な担体または賦形剤を含む有用なワクチンも記載する。

種々の有用な組成物もまた本明細書に記載する。一つの態様は、ファイバーをコードするヌクレオチド配列を欠失するアデノウイルスゲノムを含む運搬プラスミドを含む細胞を含む組換えアデノウイルスベクターの製造に有用な組成物を記載する。一つの変法において、細胞は更にファイバーをコードするヌクレオチド配列を含む完全プラスミドを含み、プラスミドは細胞の細胞性ゲノムに安定に統合される。他の変法において、運搬プラスミドは、更に、外来ポリペプチドをコードするヌクレオチド配列を含む。一つの変法において、運搬プラスミドはpDV44、pE1BgalまたはpE1sp1Bである。

他の態様において、ポリペプチドは治療的分子である。更に別の態様において、ポリペプチドは治療的分子である。他の変法は、運搬プラスミドが更に外来ポリペプチドをコードするヌクレオチド配列を含むことを提供する。

組換えアデノウイルスウイルスベクターの製造に有用な組成物をまた記載する。一つの態様において、組成物は、ファイバーをコードするヌクレオチド配列を欠き、第2の運搬プラスミドの非存在下では新規ウイルス粒子のパッケージングの指示ができないアデノウイルスゲノムを含む第1運搬プラスミド；そして第1運搬プラスミドの存在下で新規ウイルス粒子のパッケージングを指示できるアデノウイルスゲノムを含む第2運搬プラスミドを含む細胞を含む。

他の変法において、第1および第2運搬プラスミドは、細胞内で相互作用し、治療的ウイルスベクターを製造する。更に他には、細胞は更にファイバーをコードするヌクレオチド配列を含む相補的プラスミドを含み、プラスミドは細胞の細胞性ゲノム内に安定に統合される。更に別には、第1または第2運搬プラスミドが更に外来ポリペプチドをコードするヌクレオチド配列を含む。種々の態様において、ポリペプチドは治療的分子である。

他の態様は、第1運搬プラスミドがアデノウイルスパッケージングシグナル配列を欠く組成物を前記のように記載する。他の態様において、第2運搬プラスミドはLacZレポーター構築物を含む。他の変法は、第2運搬プラスミドが、アデノウイルス調節タンパク質をコードするヌクレオチド配列を更に欠くものを記載する。一つの変法において、調節タンパク質はE1である。上記組成物の一つの態様において、相補的プラスミドはpCLFの性質を有する。

他の態様において、第1運搬プラスミドがアデノウイルス構造タンパク質をコードするヌクレオチド配列を欠き、そして第2運搬プラスミドがアデノウイルスE1タンパク質をコードするヌクレオチド配列を欠く、組成物を記載する。他に、第1運搬プラスミドがアデノウイルスE4タンパク質をコードするヌクレオチド配列を欠き、そして第2運搬プラスミドがアデノウイルスE1タンパク質をコードするヌクレオチド配列を欠く。更に別には、細胞は、アデノウイルス調節タンパク質および構造タンパク質をコードする少なくとも一つの相補的プラスミドを含む。

10

20

30

40

50

別の態様において、調節タンパク質は4であり、構造タンパク質はファイバーである；または調節タンパク質がE1であり、構造タンパク質がファイバーである。更に別の態様において、アデノウイルス調節タンパク質および構造タンパク質が別の相補的プラスミドによりコードされる。

他の変法は、細胞が293、A549、W162、HeLa、Vero、211および211A細胞系からなる群から選択される組成物を記載する。他の態様において、運搬プラスミドはDV1またはpE1BgalまたはpE1sp1Bである。

本発明のベクター、プラスミド、細胞系および他の組成物および構築物を製造し、使用する種々の方法を本明細書に記載する。以下の方法は説明のためであり、限定するものではない。

10

従って、一つの変法において、本発明は、運搬プラスミドをAdファイバー-発現相補的細胞系に挿入することを含む、治療的ウイルスベクターの構築法を記載し、Adファイバータンパク質をコードするDNA配列は運搬プラスミドから除去される。一つの変法において、運搬プラスミドは、更に、外来タンパク質、ポリペプチドまたはそれらのフラグメントをコードするDNA配列を含む。他の態様において、運搬プラスミドはDV1、pE1Bgal、pE1sp1Bまたは類似の構築物である。

本発明は、また細胞を本明細書に記載の任意のベクターと接触させることを含む、病理性過増殖性哺乳類細胞の形質転換法も記載する。他の態様において、哺乳類標的細胞を、予め選択した外来ヌクレオチド配列を含むウイルスベクターで感染させる方法を記載する。一つのこのような変法は、以下の段階を含む：(a) 標的細胞を本発明のウイルスベクターで感染させ、ウイルスベクターは予め選択した外来ヌクレオチド配列を含む；そして(b) 標的細胞内で外来ヌクレオチド配列を発現させる。

20

本発明はまた、本明細書に記載の方法で製造した予め選択した外来ヌクレオチド配列に感染した哺乳類細胞もまた包含する。一つの変法において、標的細胞は複製、遅延複製および非複製ヒト細胞からなる群から選択される。

獲得または遺伝疾患を標的とする方法もまた記載する。一つの方法は、(a) ベクターが予め選択した治療的ヌクレオチド配列を含むものである、標的細胞への製薬学的に許容可能な量のウイルスベクターの投与；および(b) 細胞内の獲得または遺伝疾患の改善に十分な期間標的細胞内に治療的配列を発現させることを含む。本発明のパッケージング細胞系により製造した治療的ウイルスベクターの十分な量を患者に投与することを含む遺伝子治療法もまた記載する。

30

本発明によりまた意図されるのは、本発明の治療的ウイルスベクターの有効量を、適当な条件下で患者に投与することを含む、患者の腫瘍の増殖の阻害法である。一つの変法において、遺伝子は抗腫瘍剤をコードする。他の変法において、薬剤は腫瘍抑制因子である。更に別の態様において、薬剤は自殺遺伝子またはその機能的相同物である。他の変法において、ベクターを腫瘍内注射により投与する。

本発明はまた前記方法で使用するシステムまたはキットを記載する。システムまたはキットは、本明細書に記載のベクター、プラスミド、細胞系および付加的治療剤の適当な組み合わせを記載のように含み得る。好ましくはこのようなキットまたはシステムはそれぞれ少なくとも1回の投与に十分な量の一定量の適当な治療的物質または配列、および投与および使用のための説明書を含む。従って、一つのシステムは、更に、治療的ウイルスベクター含有組成物の治療的効果を促進する治療剤の有効な量を更に含む。

40

他の変法は、組成物および治療剤がそれぞれ別の貯蔵所または容器に包含されているものを記載する。

前記要素の任意の組み合わせがまた本明細書に記載のように有効であり得、総ての関連法がまた本発明の範囲内であることは認識される。

#### 【図面の簡単な説明】

図1は、プロモーターおよびターミネーターは配列と共に、ブロックの断片で示すオープンリーディングフレーム(ORF)を有する完全アデノウイルスE4転写単位の模式図である。E4の特異的部分を増幅するためのプライマーの位置もさらに実施例1Aに記載す

50



る。

図 2 は、更に実施例 1 B に記載するようなプラスミド pE4/Hygro の模式的地図である。

図 3 は、更に実施例 1 B に記載するようなプラスミド pCDNA3/ファイバーの模式的地図である。

図 4 は、更に実施例 1 B に記載するようなプラスミド pCLF の模式的地図である。

図 5 は、更に実施例 1 C に記載するような、211 細胞系内のそのままのアデノウイルス E 4 3.1 キロベース (kb) 挿入物の存在を示すサザンブロットの写真である。

図 6 は、実施例 1 C に記載の天然および変性電気泳動条件下で検出した標識ファイバータンパク質を示すウエスタンブロットの写真である。293 細胞はファイバーを欠くが、211A、211B および 211R は、官能的 3 量体形および変性 1 量体形で検出可能なファイバータンパク質を含む。

図 7 は、更に実施例 1 D に記載するようなプラスミド pDEX/E1 の模式的地図である。

図 8 は、更に実施例 1 F 1 に記載するようなプラスミド pE1/ファイバーの模式的地図である。

図 9 は、更に実施例 1 F 2 ) に記載するようなプラスミド pE4/ファイバーの模式的地図である。

図 10 は、多重アデノウイルス遺伝子欠失を有する組換えアデノウイルスベクターを形成するための共形質転換および組換えに使用するための直線化 pE1Bgal 運搬プラスミドの模式図である。プラスミドおよび組換え事象は、実施例 2 A により完全に記載する。

図 11 は、プラスミド p8.2 と共に、pDV44 運搬プラスミドを構築するために使用する、実施例 2 A に更に記載のようなプラスミド p11.3 の模式図である。

図 12 は、プラスミド p11.3 と共に、pDV44 運搬プラスミドを構築するために使用する、実施例 2 A に更に記載のようなプラスミド p8.2 の模式図である。

図 13. 組換えファイバーの 3 量体構造：示すような 293、211A、211B または 211R 細胞は、[<sup>35</sup>S]メチオニンで代謝的に標識し、可溶性タンパク質抽出物を調整し、ファイバーを免疫沈降させた。沈殿タンパク質の一部を 8 %SDS-PAGE ゲルで半天然または変性条件下で電気泳動させた。3 量体 (T) および 1 量体 (M) ファイバーの位置を示す。電気泳動条件のコントロールとして、バキュロウイルス感染細胞内で産生した組換え Ad2 ファイバーを同様の条件下で流し、クーマシーブルーで染色した。

図 14. ファイバー製造細胞によるファイバー変異アデノウイルスの相補性。示す細胞系 (2 × 10<sup>6</sup> 細胞 / サンプル) を温度感受性ファイバー変異アデノウイルス H5ts143 で、10 PFU/細胞で感染させ、複製 (32.5、点画棒) または制限 (39.5、黒棒) 温度でインキュベートした。感染 48 時間後、ウイルスを凍結 - 融解溶菌で単離し、収率を SW480 細胞での蛍光焦点アッセイにより測定した。各値は、二つのサンプルの平均を示し、示すデータは複数回実験の代表値である。

図 15. Ad3 粒子への組換え Ad5 ファイバーの挿入：A) 数個の異なる血清型由来のファイバータンパク質の N - 末端 (ペントン塩基結合) ドメインの配置。B) タイプ 3 アデノウイルスを 293、211B または 211R 細胞内で、示すように伝播させ、二回の連続 CsCl 遠心で精製した。精製ウイルス粒子 10 mg を次いで変性条件下で電気泳動し、PVDF 膜に移した。Ad5 ファイバーを、組換え Ad2 ファイバーに対して惹起したポリクローナルウサギ抗体で検出した。検出の陽性コントロールとして、野性型 Ad2 400 ng を 'Ad2' と記したレーンに流した、これらの条件下で、Ad2 および Ad5 ファイバーの移動性は区別がつかず、抗体は両方のタンパク質と反応する。

図 16. 三つのパッケージング細胞系における組換えファイバータンパク質の核局在化：細胞を 8 ウェルチャンバースライドで生育させ、ウサギ抗ファイバーポリクローナル抗体で染色し、FITC-結合ヤギ抗ウサギ抗体で可視化した。A) 211A 系 B) 211B 系 C) 211R 系 D) 293 細胞 (陰性コントロール) E) 1 pfu/w 細胞の Ad.RSVbgal で感染させ、感染後 24 時間に染色した 293 細胞 (陽性コントロール) F) (E) のように製造したが、一次抗体で染色していない感染細胞。

詳細な説明

10

20

30

40

50

野性型アデノウイルスによる汚染の頻度を減少させるために、組換えの可能性を減少させるようにウイルスベクターまたは細胞系のいずれかを改善する必要性が考えられている。例えば、グループCと相同性が少ないグループ由来のアデノウイルスを使用して、293細胞内でAd5配列と組換えの傾向が少ない組換えウイルスを操作する。同様に、上皮細胞系---例えば、293として既知の細胞---を使用し得るかまたは本明細書に記載の方法に従って、Ad3由来のアデノウイルスタンパク質またはポリペプチドおよび/または他の非グループCまたはグループC血清型由来のタンパク質またはポリペプチドを安定に発現するように修飾する；このような細胞系は、調節および/または構造遺伝子の欠失を担持するアデノウイルス由来ウイルスベクターを指示するために、このようなベクターが由来した血清型に無関係に有用である。

10

本発明の構築物および方法が二個またはそれ以上のAd血清型由来のアミノ酸残基配列を発現するキメラウイルスベクターの設計および操作を支持することもまた意図される。従って、本発明の記載の出現より前に利用可能であった方法および構成とは異なり、本発明はその構築および伝播を支持するウイルスベクターおよび細胞系の設計および製造の順応性の可能性を最大にする---総て、有害な組換え体を製造する可能性のある野性型Adとの組換えの危険性の減少を伴う。

一部において、本発明は、ウイルスと細胞性配列の間の組換えを、当分野で検討されているよりも減少させる、単純な、別の方法を記載する。一つのこのような手段は、組換えウイルスのサイズの増加であり、それによりウイルスとパッケージング細胞系に存在するAd遺伝子---例えば、293細胞系内のAd5遺伝子、または本発明の新規細胞系中の種々のAd遺伝子との間の共有配列の程度を減少する。

20

“実質的に相同”なる用語は、少なくとも80%、好ましくは少なくとも90%、最も好ましくは少なくとも95%の相同性を有することを意味する。

本明細書に記載のアミノ酸残基は好ましくは“L”異性体形である。しかしながら、所望の機能特性がポリペプチドにより保持されている限り、“D”異性体形の残基を、L-アミノ酸残基と置換できる。NH<sub>2</sub>はポリペプチドのアミノ末端に存在する遊離アミノ基を意味する。

DNA相同：予め選択した保存的ヌクレオチド配列を有し、本発明に従った好ましいポリペプチドをコードする配列を有する核酸。

外来遺伝子：本用語は、野性型アデノウイルスに見られる対応DNA分子として正確な配向および位置で存在しないDNAを同定するために使用する。また、他のAd血清型由来または完全に異なる種---例えば、ヒトDNA配列由来のDNA分子も意味し得る。

30

ペントン：“ペントン”または“ペントン複合体”なる用語は、本明細書で主にペントン塩基とファイバーの複合体を意味するために使用される。“ペントン”なる用語はまたペントン塩基およびペントン複合体を示すように使用し得る。“ペントン”なる用語単独の意味は、それが使用される前後関係から明らかである。

ポリペプチドおよびペプチド：これらの用語は、本明細書で、一方が他方に、隣接残基の-アミノとカルボキシ基の間のペプチド結合により結合した約50アミノ酸残基を超えない配列を意味するために、交換可能で使用される。

受容体：受容体は、本明細書で、特異的に他の分子に（または、と）結合する生理学的活性分子を意味するために使用される。“受容体タンパク質”なる用語は、特異的受容体のタンパク質様性質をより特異的に示すために使用し得る。

40

トランスジーンまたは治療的ヌクレオチド配列：本明細書に記載され、請求のように、このような配列はRNAまたはポリペプチドをコードするDNAおよびRNA配列を含む。このような配列は“天然”であるか、または天然由来配列であり得る；それらはまた“非天然”または天然または組換え由来の“外来”配列でもあり得る。本明細書では“治療的ヌクレオチド配列”と交換可能に使用し得る“トランスジーン”なる用語は、しばしば、ウイルスベクターにより運搬され、宿主細胞内に形質導入される同種または外来（異種）遺伝子を記載するために使用する。

従って、治療的ヌクレオチド配列は、アンチセンス配列またはアンチセンス配列に転写さ

50

れ得るヌクレオチド配列を含む。治療的ヌクレオチド配列（またはトランスジーン）は、更に、該治療的配列が運搬される細胞または細胞核内に所望の作用を産生するために機能する配列を含む。例えば、治療的ヌクレオチド配列は、このような機能的タンパク質を製造できない細胞内に運搬することを意図した機能的タンパク質をコードし得る。

発現または運搬ベクター：外来DNAを適当な宿主細胞内に発現するために挿入し得るプラスミドまたはウイルス---即ち、DNAによりコードされるタンパク質またはポリペプチドを宿主細胞システム内で合成する。一個またはそれ以上のタンパク質をコードするDNAセグメント（遺伝子）の発現を指示できるベクターは、本明細書で“発現ベクター”と呼ぶ。また、逆転写酵素を使用して、製造したmRNAからのcDNA（相補的DNA）のクローニングを可能とするベクターも含む。

10

アデノウイルスベクターまたはAd由来ベクター：外来DNAを挿入または発現し得る任意のアデノウイルス由来プラスミドまたはウイルス。本用語は、また“ウイルスベクター”と交換可能に使用し得る。この“型”のベクターは、更に下記のように、治療的タンパク質またはポリペプチドをコードするヌクレオチド配列の、処置を必要とする患者の特異的細胞または細胞型への運搬に使用し得る。

相補的プラスミド：本用語は、一般に、本明細書でパッケージング細胞系内に特定のヌクレオチド配列を、該配列が細胞性ゲノム内に安定に統合される目的を有して、運搬するために使用されるプラスミドベクターを記載するために使用する。

運搬プラスミド：本用語は、一般的に、本発明の治療的ウイルスベクターの伝播の目的で、細胞系（例えば、パッケージング細胞系）内にまたは中に運搬または運搬するプラスミドベクターを記載するために本明細書内で使用する。

20

アデノウイルス（Ad）粒子は、相対的に複雑であり、種々の基礎構造に分解し得る。外部殻は、形において明白に二十面体であり、一見して、25の数の三角形分割を有するように見える。5つに折りたたまれた所（“ペントン”）の構造は、残り（“ヘキソン”）と異なるが、ヘキソンは化学的に6量体よりむしろ3量体である。従って、構造は単純な、下位で三角形分割された二十面体設計には本当は対応しない（例えば、Fields, et al., Virology, Vol. 1, Raven Press, NY, pp.54-56 (1990) 参照）。

ファイバーは、ウイルスを細胞表面の特異的受容体に結合させることにより、アデノウイルス感染で重大な役割を担う。ファイバーは3つのドメインから成る：ペントン塩基と相互作用するN-末端尾；薄層および屈曲部を形成する15アミノ酸セグメントの22反復を構成するシャフト；および型特異的抗原を含み、細胞表面受容体への結合を担うC-末端のノブ。ファイバータンパク質はまたウイルス核酸の核内への運搬も担う。Ad2由来のファイバータンパク質をコードする遺伝子はヒト細胞内に発現され、3量体に正確に集合し、糖付加され、核内に運搬されることが示されている。（例えば、HongおよびEngler, Virology 185:758-761 (1991) 参照）。従って、組換えアデノウイルスベクターにより介在される、特異的細胞タイプへの遺伝子運搬の変化は、種々の遺伝子治療適用に大きな実用性を有し、従って、本発明の目的の一つである。

30

ヘキソンおよびペントンキャプソメアは、ビリオンの表面の腫瘍性分である。番号II、IIIおよびIVであるそれらの構成ポリペプチドは、ビリオンの表面に暴露され---例えば、そのままの粒子のヨウ素化により標識できるチロシン残基を含む。

40

ファイバーは、582アミノ酸長の3つの同一のポリペプチドの3量体（ポリペプチドIV）として存在する延長タンパク質である。ファイバーのN-末端は、一般的にペントンキャプソメアと呼ばれているものを形成するためのペントン塩基の結合を介在する。ファイバーのC-末端はウイルスの細胞性受容体への最初の結合に関与する。

アデノウイルスタイプ2の35,000塩基対（bp）は配列決定され、腫瘍コートタンパク質（ヘキソン、ファイバーおよびペントン塩基）の予測されたアミノ酸配列は記載されている。（例えば、Neumann et al., Gene 69:153-157 (1988); Herisse et al., Nuc. Acids Res. 9:4023-4041 (1981); Roberts et al., J. Biol. Chem. 259:13968-13975 (1984); Kinloch et al., J. Biol. Chem. 259:6431-6436; および Chroboczek et al., Virol. 161:549-554 (1987) 参考）。

50

A d 5 DNAの配列はより最近に完全になった；その配列は全35,935bpを含む。多くの他のアデノウイルスゲノムの部分もまた配列決定されている。近年、アデノウイルスのパッケージングの上限が野性型ゲノム長の約105%であることが理解されてきている。（例えば、Bett, et al., J. Virol. 67(10):5911-21(1993)参照）。従って、A d 2およびA d 5に関して、パッケージングの上限はDNA約38kbである。

アデノウイルスDNAはまた血清型に依存して約100から150bpのサイズの範囲である逆方向末端反復配列(ITRs)も含む。逆方向反復は、一本鎖ウイルスDNAがその末端配列での塩基対形成により環化することを可能とし、得られる塩基対“フライパンの柄状”構造は、ウイルスDNAの複製に重要であると考えられる。

効率的パッケージングのために、ITRsおよびパッケージングシグナル(数百bp長)は、“最小要求”を含むように見える全ウイルスPRFを欠くが、これらの本質的cis要素(ITRsおよび隣接パッケージングシグナル)を含むヘルパー依存的ベクターが構築されているが、ピリオンパッケージは、マルチマーとしてのヘルパーとパッケージより時間に関して効果がなく、これはウイルスが完全DNA補体のパッケージを“望み”得ることを示す(例えば、Fisher, et al., Virology 217:11-22(1996)参照)。

複製コンピテントベクターが安全の争点を発生させることがないように、複製欠失A d ウイルスベクターの使用がある場合好ましいが、本発明のウイルスベクターは、その中にパッケージされたゲノムを発現させる能力を保持し得る---即ち、“感染性”を保持し得る---それらはそれらが治療目的で投与された患者において疾病の原因となる程度までは感染剤として作用しない。

本発明のウイルスベクターは、当分野で記載のアデノウイルスおよびA d 由来ベクターよりもいくつかの明白な利点を有する。例えば、このようなベクターの組換えは稀である；アデノウイルスの一般的ヒト感染にもかかわらずヒト悪性腫瘍とアデノウイルス感染の関連は知られていない；ゲノムは適切に実際のサイズの外来遺伝子の適用のために操作し得る；そして宿主増殖がアデノウイルスタンパク質の発現に必要なではない。

本発明の拡大は、本明細書に記載のA d 由来ウイルスベクターが、組換えDNA法によりファイバータンパク質上に他の受容体(CD4のような)のための結合配列を挿入させることにより、遺伝子の特異的細胞に標的化および運搬するために、従って、キメラ分子を製造するために使用できる。これは、遺伝子を、宿主免疫系による認識をかいくぐる利点を有して、広範囲の細胞型に標的化および運搬できる能力をもたす。本明細書に記載の運搬系は、従って、増殖および非増殖細胞型への安定な統合を可能にするための遺伝子設計の増加した柔軟性を提供する。

例えば、その記載を本明細書に引用して包含させる公開国際出願WO 95/26412およびKrasnykh, et al. (J. Virol. 70:6839-46(1996))は、アデノウイルスファイバータンパク質を製造するためであり得る修飾を記載する。このような修飾は、標的化機構およびアデノウイルスの特異性の変更に有用であり、本発明の構築物と関連して、本明細書に記載の新規ウイルスベクターを異なる受容体および異なる細胞に標的化するのに容易に使用できる。更に、その親和性を変えるファイバータンパク質の修飾は、治療的適用におけるウイルスベクターの局在化の非常な制御を可能にし得る。

同様に、種々の構造タンパク質の本発明への細胞系への挿入は、これらのタンパク質が修飾されていてもいなくても、本発明によりまた意図される。従って、例えば、Wickham, et al. (J. Virol. 70:6831-8(1996))に記載のような修飾ペントン塩基ポリペプチドは、本明細書に記載の方法に従って使用した場合、治療的利用性を有し得る。

#### C. パッケージング細胞系

現在利用可能な組換えアデノウイルスベクターの最初の世代は、一般的にE 1と呼ばれる、E 1 aおよびE 1 b領域を含む第1ウイルス初期遺伝子領域に欠損を有する傾向がある。(これらの領域は、典型的に遺伝子地図単位を1.30から9.24に伸ばした)。Fields Virology, 3d Ed. (Fields et al. (編), Lippincott-Raven Publ., Philadelphia, (1996), p. 2116)のチャプター67の図3は、役立つ例であるアデノウイルスタイプ2(A d 2)の転写および翻訳地図を説明する。

種々の刊行された報告に従って、ウイルス E 1 領域の欠失は、組換えアデノウイルスを複製に関して不完全とし、続く感染標的細胞内での感染ウイルス粒子の製造を不可能とする。従って、E 1 欠失アデノウイルスを産生する能力は、しばしば 293 と呼ばれるヒト初期腎臓パッケージング細胞系の能力に基づく。この細胞系は、アデノウイルスの E 1 領域を含み、これは細胞系内で E 1 欠失ウイルスが生育するのを“支持する”ための E 1 遺伝子領域生産物を提供する（例えば、Graham et al., J. Gen. Virol. 36:59-71 (1977) 参照）。

それにもかかわらず、現在の第 1 世代組換えアデノウイルスの本質的な問題は、患者へのその使用に関する懸念を増加させている。例えば、いくつかの近年の実験は、E 1 欠失アデノウイルスが完全に複製 - インコンピテントでないことを示している（Rich, Hum. Gen. Ther. 4:461-476 (1993); Engelhardt, et al., Nature Genet. 4:27-34 (1993) 参照）。

3 つの一般的限定がアデノウイルスベクター法に関連する。第 1 に、感染の高い多重度（“MOI”）でのアデノウイルスのインピボおよびインピトロ両方の感染は、それ自体哺乳類細胞に毒性であるペントンタンパク質の蓄積のため、標的細胞に細胞毒性をもたらす（Kay, Cell Biochem. 17E:207 (1993)）。第 2 に、ペントンタンパク質を含むアデノウイルス後期遺伝子生産物に対する免疫応答が、ベクターを受けた感染組織の炎症性反応および破壊をもたらす（Yang, et al., PNAS USA 92:4407-4411 (1994)）。最後に、宿主免疫応答および細胞特性作用は、共に、トランスジーンの長期発現を妨げ、アデノウイルスベクターの続く投与後の遺伝子発現の減少したレベルをもたらす（Mittal, et al., Virus Res. 28:67-90 (1993)）。

本明細書に記載のパッケージング細胞系は、ヘルパーウイルスの必要性無しに、ウイルスゲノムの主要部分の欠失を有するウイルスベクターを支持する。

#### D. 治療的ウイルスベクターおよび関連システム

##### 1. 核酸セグメント

本発明の治療的ウイルスベクターまたは組成物は、本明細書に記載のような、治療的適用に使用し得るタンパク質またはポリペプチド分子---またはそれらの生物学的に活性なフラグメント---をコードするヌクレオチド配列を含む。治療的ウイルスベクターまたは組成物は、更に、このような治療的ヌクレオチド配列または遺伝子の発現のためおよび制御のための、5' に位置するエンハンサー要素またはプロモーターを含み得る。

一般に、プロモーターは、3' またはプロモーターの下流に位置する遺伝子の発現を制御する DNA 配列を含む DNA セグメントである。プロモーターは、RNA ポリメラーゼが特異的に結合し、その典型的にプロモーターの 3' に位置する遺伝子の RNA 合成（転写）を開始させる DNA 配列である。一個以上の特定のポリペプチドまたはタンパク質をコードする核酸配列が治療的ウイルスベクターまたはヌクレオチド配列に包含される場合、特に、それが発現の効率を促進する場合、一個以上のプロモーターまたはエンハンサーを含み得る。本発明の目的のために、調節可能（誘導可能）ならびに構造プロモーターを、別々のベクターであれ同じベクターであれ使用し得る。

主題の治療的ヌクレオチド組成物またはベクターは、少なくとも 2 つの異なる操作可能に結合した DNA セグメントを含む核酸を含む。DNA は、本明細書に記載のように PCR により、および Molecular Cloning: A Laboratory Manual, 2nd Ed., Sambrook et al., 編, Cold Spring Harbor, New York (1989) に記載のような標準法を使用して操作および増幅できる。典型的に、本発明の治療的ウイルスベクターを製造するために、選択的治療組成物およびプロモーターまたはエンハンサーをコードする配列は、インピボまたはインピトロで細胞内の自発的複製を可能にする DNA 分子に操作可能に結合する。エンハンサー要素またはプロモーターと治療的ヌクレオチド組成物をコードするヌクレオチド配列を、ベクターに操作可能に結合させることにより、結合セグメントはベクター配列と一緒に複製する。

従って、本発明の組換え DNA 分子（rDNA）は、通常天然では一緒には見られない少なくとも 2 つのヌクレオチドを含むハイブリッド DNA 分子である。種々の好ましい態様

10

20

30

40

50

において、配列の一つはA d由来ポリペプチド、タンパク質またはそのフラグメントである。他の方法で記載すると、本発明の治療的ヌクレオチド配列は、発現可能タンパク質、ポリペプチドまたはそのフラグメントをコードするものであり、更に活発な構造または調節可能（例えば、誘導可能）プロモーター配列を含み得る。

本発明の治療的ウイルスベクターまたは組成物は、最適には約20塩基対から約40,000塩基対長である。好ましくは、核酸分子は、約50bpから約38,000bp長である。種々の態様において、核酸分子は、一個またはそれ以上のアデノウイルスタンパク質またはその機能的ポリペプチド部分をコードするのに十分な長さである。個々のA dポリペプチドが約19アミノ酸残基から約967アミノ酸残基まで長さが代わるため、対応するヌクレオチド配列は、本発明の治療的ヌクレオチド配列によりウイルスベクター内で“置換”される個々のポリペプチドコード配列のサイズに依存して、約50bpから3000bpまでである。

10

種々のA dタンパク質は、一個以上のポリペプチド配列を含む。従って、本明細書で教示のようなA dベクター由来の対応する遺伝子の欠失は、このようにベクターがより大きい“外来”DNAセグメントでさえ適応することを可能にする。従って、一個またはそれ以上のアデノウイルスポリペプチドまたはタンパク質をコードする配列が本発明の組換えヌクレオチドに取って代わる場合、組換え配列の長さは、関連するアデノウイルス由来ベクターのパッケージングの限界近くまで伸長できることが考えられる。

本明細書に記載の好ましい態様が、ヘルパー非依存的A d由来ベクターであるという事実の観点から、全野性型A dゲノムは、このようなベクターをある種の“助け”を必要とするベクターへの形質転換無しに、組換え核酸分子により完全に取って代わられる。しかしながら、本発明のA d由来ベクターは、ヘルパーウイルスに依存しない；代わりに、本発明のベクターは、“外来”DNAのベクター内への付加を可能にするために該ベクターから除去されているタンパク質またはポリペプチドを安定に発現する細胞系内で増殖する。種々の記載の態様において、特異的初期領域および構造ポリペプチドが本発明のベクターから欠失し、それにより、種々の長さの組換え核酸配列（またはカセット）がベクターに適合するのを可能にする。例えば、本発明のA d由来ベクターは、容易に12kbまたはそれ以上の外来（または“治療的”）DNA配列を含み得る。

20

治療的（または外来）ヌクレオチド配列は、1-アンチトリプシンまたは嚢胞性線維症膜コンダクタンス制御因子タンパク質（CFTR）などの置換のような所望の治療的効果を提供するタンパク質またはポリペプチド---または生理学的に活性なそのフラグメント---をコードする遺伝子または遺伝子フラグメントであり得る。（例えば、その内容を本明細書に引用して包含させるCrystal, et al., Nature Genetics 8:42-51 (1994); Zabner, et al., Cell 75:207-216 (1993); Knowles, et al., NEJM 333 (13):823-831 (1995); およびRosenfeld, et al., Cell 68:143-155 (1992) 参照）。

30

本発明のA d由来ベクターは、また---p53、Rbまたはマイトシンのような---細胞サイクルの調節に有効なまたは、チミジンキナーゼのような細胞死を誘導するのに有効なタンパク質、ポリペプチドまたはそのフラグメントも含み得る。（例えば、その内容を引用して本明細書に包含させる国際出願WO 95 / 1984 参照）。本発明の治療的ウイルスベクターにより発現される治療的タンパク質またはポリペプチドは、適当な場合、より効果的にするために他の治療剤と組み合わせて---例えば、チミジンキナーゼ代謝産物は、チミジンキナーゼをコードする遺伝子およびその生産物と組み合わせて使用し得る---使用し得ることが更に意図される。

40

あるいは、治療的ウイルスベクターは、治療的効果を発揮する前にポリペプチド生産物に翻訳される必要性が無く、酵素的治療活性を示すDNAまたはRNAオリゴヌクレオチド配列を含み得る。後者の例は、有害遺伝子の転写を阻害するアンチセンスオリゴヌクレオチドまたは選択した変異遺伝子配列を開裂するための部位特異的リボヌクレアーゼとして作用するリボザイムを含む。他の変法において、本発明の治療的ヌクレオチド配列は、公開PCT出願WO 92 / 06693（その内容を本明細書に引用して包含させる）に記載のような、リボザイムおよびアンチセンスDNAを含む、標的細胞内に高コピー数で、治

50

療的ヌクレオチド分子の産生ができるDNA構築物を含み得る。他の好ましい本発明の治療的ヌクレオチド配列は、HIVアンチセンスヌクレオチドを、CD4を介して、潜伏感染T細胞に運搬することができる。同様に、エプスタイン・バーウイルス(EBV)EBNa-1アンチセンスヌクレオチドの、CR2を介したB細胞への運搬は、治療的結果をもたらす。

本明細書の他の場所にも記載のように、本発明のAd由来ベクターは、またプロモーター配列も含み得る。構造および調節(しばしば“誘導可能”と呼ばれる)プロモーターが本発明の構築物および方法で有用である。例えば、ある有用な調節可能プロモーターは、CREB-調節遺伝子ファミリーのものであり、インヒビン、ゴナドトロピン、シトクロームc、グルカゴンなどを含む。(例えば、その内容を引用して本明細書に包含させる公開国際出願WO96/14061参照)。

調節可能または誘導可能プロモーターは、RNAポリメラーゼ結合および開始の速度が、外部刺激により調節されるプロモーターとして記し得る。このような刺激は、種々の化合物または組成物、光、熱、負荷、化学的エネルギー源などを含む。誘導可能、抑制可能および抑圧可能プロモーターが調節可能プロモーターと見なされる。

調節可能プロモーターは、また組織特異的プロモーターも含み得る。組織特異的プロモーターは、それらが特異的細胞型に操作可能に結合している遺伝子の発現を指向する。組織特異的プロモーターは、プロモーターがその内在遺伝子を発現する場所の特異的細胞内において、その3'に位置する遺伝子を、優勢的にではあるが、排他的で無く発現させる。典型的に、組織特異的プロモーターが、その3'に位置する遺伝子を総て発現した場合その発現は、正確な細胞型で適切に発現されていると考える(例えば、Palmiter et al., Ann. Rev. Genet. 20:465-499 (1986) 参照)。

組織特異的プロモーターが遺伝子の発現を制御する場合、遺伝子は、実質的に総ての組織および細胞型よりむしろ少数の組織または細胞型で発現される。組織特異的プロモーターの例は、Brinster et al., Nature 306:332-336 (1983) およびStorb et al., Nature 310:238-231 (1984) により記載の免疫グロブリンプロモーター; Swift et al., Cell 38:639-646 (1984) により記載のエラストーゼ-Iプロモーター; Townes et al., Mol. Cell Biol. 5:1977-1983 (1985) およびMagram et al., Mol. Cell Biol. 9:4581-4584 (1989) により記載のグロビンプロモーター; Bucchini et al., PNAS USA, 83:2511-2515 (1986) およびEdwards et al., Cell 58:161 (1989) により記載のインシュリンプロモーター; Rusconi et al., Nature 314:330-334 (1985) およびGrosschedl et al., Cell 38:647-658 (1984) により記載の免疫グロブリンプロモーター; Shani, Mol. Cell Biol. 6:2624-2631 (1986) により記載のアルファ・アクチンプロモーター; Overbeek et al., PNAS USA 82:7815-7819 (1985) により記載のアルファ・クリスタリンプロモーター; Crenshaw et al., Genes and Development 3:959-972 (1989) により記載のプロラクチンプロモーター; Tremblay et al., PNAS USA 85:8890-8894 (1988) により記載のプロオピメラノコルチンプロモーター; Tatsumi et al., Nippon Rinsho 47:2213-2220 (1989) により記載のベータ-チロイド刺激ホルモン(BTSH)プロモーター; Muller et al., Cell 54:105 (1988) により記載のマウス乳腫瘍ウイルス(MMTV)プロモーター; Palmiter et al., Ann. Rev. Genet. 20:465-499 (1986) により記載のアルブミンプロモーター; Vasser et al., PNAS USA 86:11,111-11,116 (1988) により記載のケラチンプロモーター; McVey et al., J. Biol. Chem. 263:11,111-11,116 (1988) に記載のオステオネクチンプロモーター; Allison et al., Mol. Cell Biol. 9:2254-2257 (1989) により記載の前立腺特異的プロモーター; Nathans et al., PNAS USA 81:4815-4855 (1984) により記載のオブシンプロモーター; Danciger et al., PNAS USA 86:8565-8569 (1989) により記載の嗅覚マーカータンパク質プロモーター; Forss-Peltter et al., J. Neurosci. Res. 16:141-151 (1986) により記載の神経特異的エノラーゼ(NSE)プロモーター; Sutcliffe, Trends in Genetics 3:73-76 (1987) により記載のL-7プロモーターおよびPeschon et al., Ann. New York Acad. Sci. 564:186-197 (1989) およびBraun et al., Genes and Development 3:793-802 (1989) により記載のプロタミン1プロモーターを含む。(総

10

20

30

40

50

ての引用文献の記載を、本明細書に引用して包含させる。) )

## 2. 組成物

本発明の種々の別の態様において、本明細書に記載の治療法を実施するために有用な治療的配列および組成物が意図される。本発明の治療的組成物は、生理学的に耐受性の担体と、活性成分としてその中に溶解または分散した一個またはそれ以上の本発明の治療的ヌクレオチド配列を含み得る。好ましい態様において、本組成物は、治療目的で患者に投与した時に、免疫原性ではなく、またはそうでなければ望ましくない副作用の原因とならない。

本明細書での使用において、“製薬学的に許容される”、“生理学的に耐受性”およびそれらの語尾変化した形は、組成物、担体、希釈剤および試薬に関する限り、交換可能に使用され、患者---例えば、哺乳類---に、嘔吐、めまい、胃の不調などのような望ましくない生理学的作用の発生無しに投与できる物質を意味する。

例えば、発明は、上皮細胞または非上皮細胞の特異的標的化に、およびそれらの細胞への治療的ヌクレオチド配列の運搬に有効な治療的組成物を含む。上皮細胞を主に標的とするように設計した治療的組成物は、治療的ヌクレオチド配列を含むアデノウイルス由来ベクターを含み得る。本明細書に記載のように、多くのアデノウイルス由来部分が、本明細書に記載の治療的組成物および方法に有用である。

下記の実施例のいくつかは、ファイバタンパク質、ポリペプチドおよびそのフラグメントを特に記載しているが、他の構造および非構造 A d タンパク質およびポリペプチド（例えば、調節タンパク質およびポリペプチド）が種々の記載のベクターおよび細胞系の成分として使用し得ることは、ここに特に記載する。更に、異なる A d 血清型由来のタンパク質、ポリペプチドおよび/またはそれらのフラグメントを含むキメラ分子も、本明細書に記載の方法、構築物および組成物で使用し得る。同様に、本発明の組換え D N A 配列を、本明細書に記載され、請求のように、広範囲の適用を有する有用な構築物を設計するために、異なる A d 血清型由来の核酸配列を使用して製造し得る。

グループ C アデノウイルス---即ち、A d 血清型 1、2、5 および 6---が本明細書の種々の実施例で記載されているが、本発明はこれらの血清型単独に限定されないことは認められる。アデノウイルス血清型が、構造および機能的に密接に関連しているという観点から、本発明の治療的ウイルスベクター、パッケージング細胞系およびプラスミドは、任意のそして総ての A d 血清型の成分から構築し得る---そして、本明細書に記載の、本発明の種々の構築物および細胞系の製造および使用の方法は、総ての該血清型に適用される。

その中に溶解または分散した活性成分を含む医薬組成物の製造法は、当分野では良く理解されている。典型的に、このような組成物は、注射可能物---液体溶液または懸濁液のいずれか---として製造されるが、使用前に液体中の溶液または懸濁液とするのに適した固体形もまた製造できる。製剤はまた乳化でき、または坐薬、軟膏、クリーム、経皮パッチなどに、所望の投与経路に依存して調剤できる。

活性成分は、製薬学的に許容され、活性成分と適合性である賦形剤と、本明細書に記載の治療法への使用に適した量で混合できる。適当な賦形剤は、例えば、水、食塩水、デキストロース、グリセロール、エタノールなどおよびこれらの組み合わせであり、植物油、プロピレングリコール、ポリエチレングリコールおよびベンジルアルコール（注射または液体製剤のため）；およびペトロラタム（例えば、ワセリン）、植物油、動物脂肪およびポリエチレングリコール（外用製剤のため）を含む。加えて、望ましい場合、組成物は湿潤または乳化剤、等張剤、溶解促進剤、安定化剤、色素、防腐剤、緩和剤（soothing agents）および同様の添加剤（医薬製剤の通常の補助添加剤）、p H 緩衝化剤および活性剤の効果を促進する類似のものを含み得る。

本発明の治療的組成物は、その中に組成物の製薬学的に許容される塩を含み得る。製薬学的に許容される塩は、例えば、塩酸またはリン酸のような無機酸、または酢酸、酒石酸、マンデル酸などのような有機酸と形成される酸付加塩（ポリペプチドの遊離アミノ基と形成）を含む。遊離カルボキシル基と形成される塩はまた、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、水酸化カルシウムまたは水酸化鉄（III）のような

10

20

30

40

50



無機塩基から、およびイソプロピルアミン、トリメチルアミン、2-エチルアミノエタノール、ヒスチジン、プロカインなどのような有機塩基にも由来し得る。

生理学的に耐容性の担体は、当分野で既知である。液体担体の例は、活性成分と水以外にいかなる物質も含まない、または生理学的pH値のリン酸ナトリウムのような緩衝剤または生理食塩水もしくはリン酸緩衝食塩水のような両方を含む、滅菌水溶液である。更に、水性担体は、一つ以上の緩衝塩、ならびに塩化ナトリウムおよび塩化カリウム、デキストロース、ポリエチレングリコールおよび他の溶質を含み得る。

液体組成物は、また、水に加えておよび水以外に液体相を含み得る。このような付加的液体相の例は、グリセリン、綿実油のような植物油および水-油エマルジョンである。

治療的組成物は、典型的に、標的組織に治療的有效量を運搬するのに十分な本発明の治療的ヌクレオチド配列の一定量、典型的に、全治療組成物当たり、少なくとも0.1重量%から約90重量%の治療的ヌクレオチド配列を含む。重量%は、治療的ヌクレオチド配列対全組成物の重量比である。従って、例えば、0.1重量%は、全組成物100g当たり、DNAセグメント0.1gである。

本発明の合成オリゴヌクレオチドを含む治療的ヌクレオチドは、ホスホトリエステルまたはホスホジエステル法のような適当な方法を使用して製造できる。その内容を引用して本明細書に包含させる、Narang et al., Meth. Enzymol. 68:90 (1979); 米国特許第4,356,270号; およびBrown et al., Meth. Enzymol. 68:109, (1979) 参照。

変異体のファミリーが好ましい、治療的オリゴヌクレオチド組成物のために、ファミリーメンバーの合成を、一反応容器で同時に行い得、または独立して合成し、予め選択したモル比を混合し得る。同時の合成に関して、オリゴヌクレオチドの位置が、オリゴヌクレオチドポリマーの伸長のために化学的に添加される場合、一つの予め選択したヌクレオチド前駆体の固相オリゴヌクレオチド反応混合物への添加により、ファミリーメンバーの配列の予め選択した位置で保存されるヌクレオチド残基を、化学合成プロトコールで同時に変異体に挿入できる。変化する配列内のこれらの位置へのヌクレオチド残基の付加は、一定量の、好ましくは等モル量の多重の予め選択したヌクレオチド前駆体の固相オリゴヌクレオチド反応混合物への化学合成中の添加により、同時に、挿入できる。例えば、4つの可能性のある天然ヌクレオチド(A、T、GおよびC)を、予め選択した位置に添加する場合、その前駆体をその段階でオリゴヌクレオチドに添加し、4つの変異体を同時に形成する。

同時に関連オリゴヌクレオチドのファミリーを形成する方法は、Ausubel et al. (Current Protocols in Molecular Biology, Suppl. 8, p. 2.11.7, John Wiley & Sons, Inc., New York (1991)) により“宿重オリゴヌクレオチド”の製造として先に記載されており、本明細書に記載の治療的オリゴヌクレオチド組成物の製造に容易に適用できる。

共通の4つのヌクレオチド(A、T、GおよびC)またはRNA相同ヌクレオチドウラシル(U)以外のヌクレオチド塩基もまた本発明で使用できる。例えば、イノシン(I)がA、TおよびGとハイブリダイズできるが、Cとはできないことは既知である。他の有用なヌクレオチドアナログの例は、当分野で既知である; 多くが37 C.F.R. § 1.822のリストに見られ得る。

従って、4つの共通のヌクレオチドがオリゴヌクレオチドのファミリーの一つの位置を占める場合、即ち、予め選択された治療的ヌクレオチド組成物が、一つの位置で変化する4つの配列とハイブリダイズできるように設計された場合、数個の異なるオリゴヌクレオチド配列が意図される。組成物は、4つのメンバーを含み得、選択された場所はA、T、GまたはCを含む。あるいは、組成物は、二つのメンバーを含み得、予め選択された位置がIまたはCを含み、その位置で全4つの共通ヌクレオチドとハイブリダイズする能力を有する。最後に、非不安定化法で、一個以上の共通ヌクレオチドと、イノシンと同様の方法でハイブリダイズする能力を有する他のヌクレオチドを、予め選択した位置に含み得る。

### 3. 発現ベクター系

外因性DNAを真核細胞内に挿入する方法は、分子生物学者の最も強力な道具の一つとなっている。“外因性”なる用語は、本発明の治療法で投与される、本発明の治療的組成物

10

20

30

40

50

を包含する。従って、“外因性”は、また“外来”、“非天然”等とまた本明細書では記載し得る。本発明の方法は、好ましくは、DNAの受容細胞への効率的な運搬および続く外来DNAを発現する細胞の同定を必要とする。

広く使用されているプラスミドはpBR322であり、そのヌクレオチド配列およびエンドヌクレアーゼ開裂部位が既知であるベクターである。種々の他の有用なプラスミドベクターを下記実施例で記載する。

本発明のベクターは、細胞内で自発的複製ができ、それにDNAセグメント、例えば、遺伝子またはポリヌクレオチドが、結合セグメントの複製をもたらすように、操作可能に結合できる核酸（好ましくはDNA）分子を含む。本発明において、ベクター配列に操作可能に結合するヌクレオチドセグメントの一つは、治療的核酸分子の少なくとも一部をコードする---事実上、一個またはそれ以上の治療的タンパク質またはポリペプチド、またはそのフラグメントをコードする核酸。

種々の態様において、治療的遺伝子の完全なペプチドコード配列をベクターに挿入し、発現させる；しかしながら、ある非コード配列をまた同様に含むベクターの構築もまた可能である。好ましくは、しかしながら、非コード配列を排除する。あるいは、ポリペプチドの可溶性形のヌクレオチド配列を使用し得る。他の好ましい治療的ウイルスベクターは、発現のためのベクターに操作可能に結合した治療的ヌクレオチド配列の少なくとも一部をコードするヌクレオチド配列を含む。DNA配列またはセグメントに関して本明細書で使用する限り、語句“操作可能に結合する”は、一般に、配列またはセグメントが、DNAの断片に共有結合的に、一重結合または二重結合形であれ、結合していることを意味する。

本発明の治療的ヌクレオチド配列を操作可能に結合させるウイルスベクターの選択は直接、当分野で既知のように、所望の機能的特性、例えば、ベクター複製およびタンパク質発現および形質転換する宿主細胞に依存する---これらは、組換えDNA分子構築の分野で固有に限定されている。あるアデノウイルス血清型を、具体的実施例の形で本明細書で記載しているが、本発明は、ハイブリッドおよびその誘導体を含む、任意のアデノウイルス血清型の使用を意図することは理解される。観察されるように、2個またはそれ以上の血清型のヌクレオチドおよび/またはアミノ酸残基配列を本発明の構築物、組成物および方法に使用することは、珍しいことではなく、本発明の範囲外ではない。

当業者が注目するように、本発明の種々の態様において、異なる“タイプ”のベクターを記載する。例えば、ベクターの一つの“型”は、特定のヌクレオチド配列をパッケージング細胞系に運搬するために使用し、該配列が細胞性ゲノムに安定に統合される目的を有する；これらの“型”のベクターは、一般に、本明細書では完全プラスミドと呼ぶ。本明細書に記載の更なる“型”のベクターは、本発明の治療的ウイルスベクターを増殖する目的で、ヌクレオチド配列を細胞系にまたは細胞系内（例えば、パッケージング細胞系）に運搬または輸送する；従って、これらのベクターは、本明細書で運搬プラスミドと一般的に呼ぶ。本明細書に記載の第3の“型”のベクターは、治療的タンパク質またはポリペプチドをコードするヌクレオチド配列を、処置を必要とする患者の特異的細胞または細胞型に運搬するために使用される；これらのベクターは、本明細書では、治療的ウイルスベクターまたはAd由来ベクターと同定する。

一つの態様において、指向性ライゲーション手段が、上流ヌクレオチド配列、下流ヌクレオチド配列または両方に存在するヌクレオチドにより提供される。他の態様において、指向性ライゲーションに適したヌクレオチドの配列は、多重指向性クローニング手段を停止するヌクレオチドの配列を含む。指向性ライゲーションに適したヌクレオチドの配列が、多くの制限部位を定義する場合、それは多重クローニング部位と呼ぶ。

翻訳可能ヌクレオチド配列は、一つのリーディングフレーム内のポリペプチドをコードする中断されていない一連の少なくとも8つのコドンを提供するヌクレオチドの連続直線である。好ましくは、ヌクレオチド配列はDNA配列である。ベクターそれ自体、ウイルスベクター（RNAまたはDNA）、剥き出しの直鎖または環状DNA、または核酸材料および細胞に挿入されるポリペプチドを含む小胞または覆いのような適当な形であり得る。

本発明の治療的ヌクレオチド組成物が存在する好ましいウイルスベクターは、アデノウイルス (A d) 由来である。ベクターがプロモーター配列を含むことも望ましい。本明細書で教示のように、本発明のウイルスベクターは、処置をしようとする治療の性質に依存して、それらが特異的に予め選択した受容細胞型を標的とするような方法で設計および構築している。特異的細胞を標的とする治療的ウイルスベクターの製造および使用法は、更に、下記の実施例に記載する。

新規ベクターおよび組成物はまたそうでなければ、野性型アデノウイルスビリオンの標的とならない細胞を選択的に標的とするように設計および製造もし得る。例えば、非上皮細胞を標的とするために、本発明の教示に従って、非上皮細胞または特定のアデノウイルスにより一般的に標的とされるのと異なる受容体を使用使用する外来タンパク質、ポリペプチドまたは他のリガンドをコードするヌクレオチド配列を含む治療的ベクターも製造し得る。特異的受容体 (括弧内に同定) を指向する有用なリガンドの例は HIVgp120 の V 3 ループ (C D 4) ; トランスフェリン (トランスフェリン受容体) ; L D L (L D L 受容体) ; および脱グリコシルタンパク質 (アシアロ糖タンパク質受容体) を含む。アデノウイルスファイバーに添加し得る有用なリガンド---およびそれを製造し結合する方法---は、その内容を引用して本明細書に包含させる公開国際出願 W O 9 5 / 2 6 4 1 2 に明示されている。

本発明のウイルスベクター内に含まれる外来ヌクレオチド配列によりコードされ得るか、または該ベクターが患者に投与された後にそれにより発現されるタンパク質またはポリペプチドに結合し得る有用なリガンドは、また抗体および結合配列、ならびに受容体それ自体を含む。例えば、M H C クラス I およびクラス II、アシアロ糖タンパク質受容体、トランスフェリン受容体、L D L 受容体、C D 4 および C R 2 のような、インテグリンなどへの細胞受容体分子への抗体は、本発明に有用な数個である。典型的に受容体に結合するリガンドおよびこれらのリガンドのアナログは、本明細書に記載のように、細胞性標的化剤として使用し得ることも理解される。

#### E. 治療法

本発明のベクターは、特に遺伝子治療に適している。従って、種々の治療法が本発明により意図される。

例えば、A d 由来ウイルスベクターが、治療的ヌクレオチド配列を特異的細胞または組織に運搬でき、それにより、より伝統的な方法が効果の限界のある多くの状態に利用可能な処置選択を広げ、促進することが発見された。従って、これらのベクターを使用する遺伝子治療法は、本発明の範囲内である。ベクターは、典型的に精製し、次いで、有効な量をインビボまたはエキソビボで患者に投与する。

例えば、本組成物は、H I V 感染を破壊するためにインビボで予防的にまたは治療的に使用し得、作用機構はアンチセンス H I V 配列またはリボザイムの T 細胞または単核細胞への運搬を介した遺伝子発現または活性化の阻害による。本発明の方法を使用して、本発明に記載の治療的ウイルスベクターを、このような細胞の感染は、本発明のウイルスベクターが容易に標的化し得る明確なインテグリンにより介在されるように思えるため、造血細胞を含む特異的細胞および組織に標的化し得る。(例えば、Huang, et al., J. Virol. 70:4502-8 (1996) 参照)。

他の有用な治療的ヌクレオチド配列は、EBV EBNa-1 遺伝子に相補的なアンチセンスヌクレオチド配列を含む。このような治療的配列の使用は、E B V への B 細胞の潜伏感染を治療するか、予防する。本明細書および下記実施例に記載のように、標的化および運搬は、種々のリガンド、受容体、および他の適当な標的化剤の使用を介して達成し得る。

従って、一つの態様において、本発明の治療法は、E B V または H I V に感染している患者の細胞を、本発明の治療的ヌクレオチドを含む製薬学的に許容可能な組成物の治療的有效量と接触させることを含む。関連する態様において、接触は、治療的ヌクレオチド配列組成物を、E B V または H I V - 介在感染をしている細胞に挿入することに関与する。

遺伝子治療法は、当分野で既知である (例えば、Larrick および Burck, Gene Therapy: Application of Molecular Biology, Elsevier Science Publ. Co., Inc. New York, NY (19

10

20

30

40

50

91) ;Kriegler, Gene Transfer and Expression: A Laboratory Manual, W. H. Freeman and Company, New York (1990) 参照)。“患者”なる用語は、マウス、ラット、ウシ、ブタ、イヌ、ネコ、ウマ、クマまたはヒト患者のような任意の動物---特に哺乳類---患者を含むと理解される。

ベクターに担持される外来遺伝子が、腫瘍抑制遺伝子または他の抗-腫瘍タンパク質をコードする場合、ベクターは患者の細胞の過増殖の処置または減少、患者の腫瘍増殖の阻害または特定の、関連する病状の軽減に有用である。病理学的過増殖細胞は、甲状腺の肥厚、乾癬症、湿疹、良性前立腺肥大症、胸部癌を含み、フラウメニ症候群、肉腫および他の新生物、膀胱癌、大腸癌、肺癌、種々の白血病及びリンパ腫のような種々の病気の状態を特徴とする。

10

非病理的過増殖細胞は、例えば、傷回復に関連する細胞で見られる。しかしながら、病理的過増殖細胞は、特徴的に、接触阻害の欠失を示し、細胞の表面特性の変化を含む選択的粘着の能力が減少し、更に、細胞内伝達が破壊している。これらの変化は、分割の刺激およびタンパク質分解酵素の分泌の能力を含む。

本発明はまた、本発明のベクターを使用した、細胞集団内への野性型腫瘍抑制因子の挿入を介した、病理的哺乳類過増殖細胞に汚染された骨髓再組成中の造血前駆体の適当なサンプルの枯渇の方法も意図する。本明細書で使用する限り、適当なサンプルは、患者から得られた異種細胞集団、例えば、表現型的に正常なおよび病的な細胞の両方を含む細胞の混合集団として定義する。

投与は、本発明の治療的薬剤の細胞または患者への、直接注射、静脈内、腹腔内を含む種々の手段、腫瘍内注射、エアゾールによるまたは局所的な挿入を含む---がこれらに限定されない---。本明細書に記載の治療剤は、本明細書に記載のように、有効量の治療剤と製薬学的に許容される担体の投与のための組み合わせであり得る。

20

本明細書で使用する限り、“有効量”は、一般的に、ベクターを投与した患者において良性の結果を達成するベクター（またはそれにより産生/放出されたタンパク質）の量を意味する。投与する全量は、必然的に、関連の技術者が認識するように、投与形態に依存して変化し、投与量も同様に変化し得る。

生理学的ベクターの量は、幾分複雑であり、濃度（ミリリットル当たりのプラーク形成単位（pfu/ml）として）、全量（pfusとして）、そして細胞当たりの投与したベクターの概算数（感染の概算多重度またはMOI）の用語で記載し得る。従って、ベクターを一定量で、注入---例えば、鼻上皮を介して---により投与した場合、それぞれの濃度は、例えば、下記のように記載し得る：

30

| 濃度<br>(pfu/ml) | 用量<br>(ml) | 投与量<br>(pfu)       | 概算MOI |
|----------------|------------|--------------------|-------|
| $10^7$         | 2          | $2 \times 10^7$    | 1     |
| $10^8$         | 2          | $2 \times 10^8$    | 10    |
| $10^9$         | 2          | $2 \times 10^9$    | 100   |
| $10^{10}$      | 2          | $2 \times 10^{10}$ | 1000  |

一般に、アデノウイルスベクターを鼻上皮を介した注入により投与した場合、約10またはそれより大きい概算MOIとなる投与量が、それより少ない用量より有効である。（例えば、Knowles, et al., New Eng. J. Med. 333:823-831 (1995) 参照）。同様に、直接注入---例えば、腫瘍へのウイルスベクターの直接注射---が、好ましい処置様相である場合、 $1 \times 10^9$  pfuまたはそれ以上の投与量が一般に好ましい。（例えば、公開国際出願WO 95 / 11984 参照）。

40

従って、投与形態に依存して、一回量で投与する有効な量は、好ましくは約 $10^6$ から $10^{15}$ 感染単位を含む。典型的な処置の経過は、5日間にわたるこのような一日当たりの投与量である。当業者が認識するように、有効量は、（1）処置する病状または他の状態、（2）患者の状態および感受性および（3）先に適用していてもよい他の処置経歴に対する患者の耐容性のような、当業者に既知の種々の他の因子に依存し得る。従って、当業者

50

は、このような因子の理解および評価に基づいて、特定の患者に投与し得る、本発明の薬剤／ベクターの量を、容易に、そして正確に測定し得る。

本発明は、また患者における過増殖性細胞または遺伝的欠失により特徴付けられる病状の軽減法も意図する。このようなベクターは、好ましくは、適当な条件下で病状を軽減することができる遺伝子生産物（例えば、ポリペプチドまたはタンパク質）をコードする外来遺伝子を含む。本明細書での使用に関して、“遺伝的欠失”は、遺伝された因子によりもたらされる病気、状態または異常、例えば、ハンティントン病、テイ－サックス病または鎌状赤血球病を意味する。

本発明は、更に、腫瘍抑制遺伝子以外の抗腫瘍遺伝子を含むアデノウイルス発現ベクターの有効量を、腫瘍塊に挿入することによる、患者の腫瘍細胞の増殖の減少法を提供する。抗腫瘍遺伝子は、例えば、チミジンキナーゼ（TK）をコードできる。有効量の治療剤を、次いで、患者に投与する；治療剤は、抗腫瘍遺伝子存在下で、細胞に毒性である。

例示のようにチミジンキナーゼを使用して、治療剤はガンシクロビル（GCV）、6－メトキシプリンアラビノヌクレオシド（araM）またはそれらの機能的相同物のようなチミジンキナーゼ代謝産物である。チミジンキナーゼ遺伝子およびチミジンキナーゼ遺伝子代謝産物の両方を、宿主細胞に毒性を作用されるために同時に使用しなければならない。TK遺伝子存在下で、GCVはリン酸化され、DNA合成の強力なイニシエーターとなり、一方araMは細胞毒性同化生成物araATPに変換する。従って、厳密な作用法または相乗作用は、治療効果に意味は無い；適当な遺伝子および治療剤の同時の使用が、特異的疾患状態の軽減に有効であり得ることが意味がある。

他の有用な例は、酵素シトシンデアミナーゼを発現する、本発明のベクターの使用を意図する。このようなベクターは、医薬5－フルオロウラシル（AustinおよびHuber, Mol. Pharm. 43:380-387（1993））または近年記載のイー・コリDeo遺伝子と組み合わせた6－メチル－プリン－2’－デオキシリボヌクレオシド（Sorscher et al., Gene Therapy 1:233-238（1994））の投与に関連して使用する。

先に記載の腫瘍抑制遺伝子の使用に関連して、他の抗腫瘍遺伝子の使用が、単独であれ、適当な治療剤との組み合わせであれ、腫瘍および悪性腫瘍の特徴である非制御細胞生育または増殖特性の処置を提供する。従って、本発明は、患者の非制御細胞性生育の停止の治療を提供し、それにより患者に存在する症状または疾患または悪液質を軽減する。本処置の効果は、患者の生存期間の延長、腫瘍塊または重量の減少、または循環腫瘍細胞の数の減少を含むが、これらに限定されない。本治療の優れた作用の定量手段は、当業者は既知である。

本発明は、機能的な外来ポリペプチド、タンパク質または生理学的に活性なそのフラグメントをコードする治療的外来核酸配列を収容し、ファイバーをコードする遺伝子のような、一個またはそれ以上のアデノウイルス構造タンパク質遺伝子の部分的または全欠失により特徴付けられる組換えアデノウイルス発現を提供する。例えば、このような機能的ポリペプチド部分は、抗癌遺伝子TKが一つの例であるが、自殺遺伝子またはその機能的相当物であり得る。TK遺伝子は、発現した時、特にGCV存在下で、細胞に致死である遺伝子生産物を産生する。TK遺伝子の一つの源は単純ヘルペスウイルス（HSV）であるが、他の源が同様に既知であり、本明細書に教示のように使用し得る。TK遺伝子は、当業者に既知の方法で、HSVから容易に得られ得る。例えば、イー・コリHB101（ATCC#39369から）のプラスミドpMLBTKは、HSV-1 TK遺伝子の源であり、本明細書に記載のように使用し得る。（例えば、その内容を引用して本明細書に包含させる公開国際出願WO 95 / 11984参照）。

治療的遺伝子配列は、アデノウイルス発現ベクターと製薬学的に許容される担体の組み合わせにより、腫瘍塊に挿入し得る。挿入は、例えば、組換えAdベクターの腫瘍塊への直接注射により達成できる。肝細胞癌（HCC）のような癌の特別な例の場合、肝動脈への直接注射が、ほとんどのHCCが本動脈からのその循環に由来するため、運搬に使用し得る。同様の投与方法は、当業者に既知のように、他の特異的型の腫瘍および悪性腫瘍に使用し得る。

腫瘍抑制遺伝子の腫瘍特異的運搬は、患者の標的組織を、有効量の本発明の組換え A d 由来ベクターと接触させることにより達成される。抗腫瘍治療において、遺伝子は、機能的腫瘍抑制遺伝子生産物または自殺遺伝子生産物のような抗腫瘍剤をコードすることを意図される。“接触”なる用語は、腫瘍内注射のような、ベクターの効率的な運搬のための運搬法を含むことを意図する。

他の実施例において、本発明のアデノウイルスは、インビボで中枢神経系 (CNS) 腫瘍への遺伝子の運搬に使用できる。定位的運搬を使用して、A d 由来ベクターは腫瘍治療を意図した CNS への遺伝子の運搬ができる。例えば、Badie, et al. (引用して本明細書に包含させる Neurosurgery 35 (5) :910-916 (1994)) は、約  $10^7$  および  $10^8$  プラーク形成単位 / ml のベクター価での 50% および 90% の形質導入が、インビトロ実験で観察されたと報告している。適当な動物脳腫瘍モデルを使用した彼らのインビボ実験において、 $10^7$  を超える価が、細胞変性効果を有することが観察された；腫瘍細胞生育の 50% を超える減少が、 $10^8$  pfu/ml で示された； $10^{10}$  pfu/ml 程高い価を患者動物の脳組織に注射した場合、毒性効果は見られなかった (Id)。従って、 $10^7$  pfu/ml より大きい価の使用が、CNS 腫瘍を攻撃するときに適当であると考えられる。

本発明はまた本明細書に記載の治療的組成物および方法の効果の測定法も意図する。効果を確認するための方法の一つは、EBV - アンチセンス治療的ヌクレオチド配列が腫瘍形成を遮断するかの評価のための、EBV - 誘発 LPD (リンパ組織増殖性疾患) のヒト / SCID (重症複合免疫不全症) マウスモデルの使用である。(例えば、その内容を引用して本明細書に包含させる Pisa, et al., Blood 79:173-179 (1992); Rowe, et al., Curr. Top. Microbiol. Immunol. 166:325 (1990); および Cannon, et al., J Clin. Invest. 85:1333-1337 (1990 参照)。

最後に、種々の疾病の処置、治療および / または診断のための医薬の製造における本発明の A d ベクターの使用がまた本発明により意図される。更に、他の抗腫瘍遺伝子に対応する治療剤と組み合わせて使用して、腫瘍細胞の増殖を減少し得る。このような他の遺伝子 - および - 治療剤組み合わせは当業者に既知であり、本明細書に教示のように使用し得る。

#### F. 遺伝子運搬のための治療的ウイルスベクターの構築

インビトロ遺伝子運搬のために、投与は、しばしば、最初に肺上皮細胞、リンパ球などのような患者の細胞集団から選択細胞を単離し、続いて本発明の治療的組成物のインビトロ遺伝子運搬をし、そして患者に細胞を戻すことにより、達成される。インビトロ治療は、また例えば、本発明の治療的組成物の種々の運搬手段による投与を介して、意図される。例えば、エアゾール投与および皮下、静脈内、腹腔内、筋肉内、眼内等を介した投与はまた本発明の範囲内である。

他の遺伝子運搬法はまた本発明の方法、組成物および構築物と関連して有用である；例えば、内容を引用して本明細書に包含させる公開国際出願 WO 95 / 11984 参照。

同様に、種々の非ヒト動物が、本発明のベクターまたは形質転換細胞を挿入されている。これらの“トランスジェニック”動物を、当業者に既知の方法を使用して作る。例えば、米国特許第 5,175,384 号参照 (その内容を引用して本明細書に包含させる)。

本発明はまた特異的細胞 --- 例えば、診断および / または処置を必要とする患者の細胞 --- の標的化のための種々の方法も意図する。本明細書に記載のように、本発明は、本発明のウイルスベクターおよび組成物が、これらのベクターおよび組成物を特異的細胞または細胞型に運搬するという最終的な目的のために、特異的受容体または細胞を指向し得る。本発明のウイルスベクターおよび構築物は、この点で特に有用である。

一般に、細胞へのアデノウイルスの結合および取りこみは、別のウイルスコートタンパク質と、内在化のための結合用受容体および V インテグリン受容体の相互作用によりもたらされる、別々ではあるが協調した事象である。ファイバーコートタンパク質を介したアデノウイルスの細胞表面への結合は、内在化の続く段階と区別でき、異なることが記載されており、本発明は、これらの受容体の区別と関連して、利点を有し得る。

#### G. 他の適用

本発明の細胞系、ウイルスベクターおよび方法はまた治療的ヌクレオチド配列の直接投与以外の目的にも有用であり得る。このような適用の一つは、本発明のウイルスベクターで形質転換した細胞内での大量の生理学的活性タンパク質またはポリペプチドの製造が意図される。例えば、ヒトリンパ芽球状細胞は、エリスロポエチン（EPO）の遺伝子のようなヒト造血生育因子を担持する本発明の完全ウイルスベクターで形質転換し得る；このように形質転換した細胞は、生理学的に活性なEPOを製造できる。（例えば、Lopez et al., Gene 148:285-91（1994）参照）。

記載した本発明の方法、細胞系、プラスミド、ベクターおよび組成物のその他様々な応用および使用は、下記の実施例のより詳しい研究により明らかになるであろう。

#### 実施例

下記の実施例は、本発明の例示説明を意図するものであって、限定するものではない。それ自体、下記の説明は、本発明の特定の実施態様を製造および使用できるような詳細な手段を提供するものである。この説明は、本発明の例示説明であるが、発明を具体的に限定するものと解釈されるものではない。現在既知の、あるいは後に開発される変形物および均等物は、当業者の理解および技術的能力の範囲内であり、本発明の範囲内にあると考えられる。

#### 実施例 1

##### アデノウイルスパッケージング細胞系の製造

アデノウイルスを増殖させるために通常使われる細胞系は、宿主細胞としてアデノウイルスパッケージング細胞系の製造にも有用である。好ましい細胞には、293、即ちATCCから入手した受託番号CRL 1573のアデノウイルス形質転換ヒト胎児腎臓細胞系；ヒーラー、即ちヒト上皮癌細胞系（ATCC受託番号CCL 2）；A549、即ちヒト肺癌細胞系（ATCC受託番号CCL 1889）；および類似の上皮由来細胞系がある。アデノウイルス形質転換の結果として、293細胞は、E1初期領域調節遺伝子を含む。全ての細胞は、特記しない限り、完全DMEM + 10%ウシ胎児血清中に維持した。本発明の細胞系は、アデノウイルス遺伝子の細胞相補性により、前もって選択した遺伝子領域に欠失のある新規アデノウイルスベクターの生産および増殖を可能にする。本発明の新規ウイルスベクターを作る目的で、このような欠失アデノウイルスゲノムの望ましい相補性を提供するため、前もって選択した機能性単位を含むプラスミドベクターを本明細書に記載のとおり設計した。このような単位には、E1初期領域、?????、およびウイルスファイバー遺伝子があるが、これらに限定されない。このような相補性を与えるプラスミドを調製することによって宿主細胞染色体内に安定に挿入される“相補性プラスミドまたは構築物”とすることは、下記に説明する。

##### A. E4 遺伝子欠失アデノウイルスの相補性用のE4 - 発現プラスミドの製造

ウイルスE4調節領域は、互い違いにスプライシングされて数種のmRNAを産生する、単一転写単位を含む。本明細書に記載のとおり製造され、かつ293細胞系をトランスフェクションするために使用される、E4発現プラスミドは、図1に示したように、E4転写単位全体を含む。Chroboczek et al. (Virology, 186: 280-285 (1992), GenBank受託番号M73260)に記載の5型アデノウイルス（以後、Ad5と呼ぶ）のヌクレオチド32667 - 35780に相当する、天然E4プロモーターを含むE4転写開始部位の上流175ヌクレオチドから、天然E4ターミネーターシグナルを含むE4ポリアデニル化シグナルの下流153ヌクレオチドへと伸びるDNAフラグメントは、ATCCから入手したAd5ゲノムDNAからポリメラーゼ連鎖反応（PCR）を介して増幅させた。使用したプライマーの配列は、5' CGGTACACAGAAATTCAGGAGACACAACTCC3'（E4Lとして表す前方または5'プライマー）（配列番号1）および5' GCCCTGGATCCGGGAAGTTACGTAACTGGGAAAC3'（E4Rとして表す後方または3'プライマー）であった。PCRフラグメントのクローニングを助長するため、これらのオリゴヌクレオチドを、下線のヌクレオチドで示したように、制限酵素EcoRIおよびBamHIそれぞれに対する新規部位を持つように設計した。DNAは、PCRにより92 で1分間、50 で1分間、72 で3分間のサイクル

10

20

30

40

50

を30回実施して増幅させ、増幅全長E4遺伝子産物を得た。

その後、増幅DNA E4産物は、p Bluescript/SK+の融和性部位にクローニングするため、プラスミドpBS/E4を作る標準技術によりEcoRIおよびBamHIで消化した。単純ヘルペスウイルスチミジンキナーゼプロモーター、ヒグロマイシン耐性遺伝子およびチミジンキナーゼポリアデニル化シグナルを含む2603塩基対(bp)カセットは、FspIで消化し、次いでBamHIでの二次消化用にBamHIリンカー(5'CGCGGATCCGCG3')(配列番号3)を添加して、ヒグロマイシン含有フラグメントを単離することによって、プラスミドpMEP4(Invitrogen, San Diego, CA)から切り取った。

次いで、8710bp含有プラスミドpE4/Hygroを作るために、単離したBamHI-修飾フラグメントをE4領域を含有するpBS/E4のBamHI部位にクローニングした(図2)。pE4/Hygroプラスミドは実施例3に記載のように、ATCCに寄託されている。pE4/Hygroの完全ヌクレオチド配列は、配列番号4に挙げる。線状ベクターの位置番号1は、図2に、ヒグロマイシン挿入物の3'BamHI部位とE4挿入物の3'EcoRI部位の間の細い線で示したように、pBS/SK+主鎖のほぼ中間位置に相当する。E4遺伝子の5'および3'末端は、配列番号4のそれぞれのヌクレオチド位置3820および707に位置するが、ヒグロマイシン挿入物の5'および3'末端は、それぞれのヌクレオチド位置3830および6470に位置する。使用のために選択したクローンでは、E4およびヒドロマイシン耐性遺伝子を分岐的に転写した。

#### B. ファイバー遺伝子欠失アデノウイルスの相補性のためのファイバー発現プラスミドの製造

ファイバーコード構築物を製造するために、プライマーを消化し、フラグメントの5'および3'末端にそれぞれユニークBamHIおよびNotI部位を付加してAd5ゲノムDNAからファイバーコード領域を増幅させた。Ad5ヌクレオチド配列は、GenBank受託番号M18369で入手できる。5'および3'プライマーは、それぞれのヌクレオチド配列、5'ATGGGATCCAGATGAAGCGCGCAAGACCG3'(配列番号5)および5'CATAAACGCCGCCCTTCTTTATTCCTTGCGC3'(配列番号6)を持ち、ここで、挿入したBamHIおよびNotI部位は下線で示す。5'プライマーもまた、開始部位である第2ATGコドン(CないしA)の5'の3つのヌクレオチドにヌクレオチド置換を含有した。このヌクレオチド置換はファイバータンパク質翻訳を開始するためのコンセンサスを向上するために含められた。

増幅DNAフラグメントは、pcDNA3(Invitrogen)のBamHIおよびNotI部位に挿入して、7148bpのプラスミド消化pCDNA3/ファイバーを作った。そのプラスミド地図は、図3に示す。親のプラスミドは、CMVプロモーター、ウシ成長ホルモン(BHG)ターミネーター、およびネオマイシン耐性を与える遺伝子を含有した。この構築物に含まれるウイルス配列は、Ad5ゲノムのヌクレオチド31040-32791に相当する。

pCDNA3/ファイバーの完全ヌクレオチド配列は、配列番号7に挙げる。ここで、ヌクレオチド位置1は、pcDNA3ベクター配列のほぼ中間に相当する。ファイバー遺伝子5'および3'末端は、ATGではヌクレオチド位置916に、TAAではヌクレオチド位置2661にそれぞれ位置する。

pcDNA3ベクターによって提供される構成CMVプロモーターによるファイバータンパク質の発現を強化するために、2型アデノウイルスの3つにわかれたリーダー(tripartite leader)(TPL)を含有するBglIIフラグメントをpRD112aから切り取り(Sheay et al., *BioTechniques*, 15: 856-862 (1993))、pCDNA3/ファイバーのBamHI部位に挿入して、7649bpを持つプラスミドpCLFを作った。そのプラスミド地図は、図4に示す。Logan et al., *Proc. Natl. Acad. Sci., USA*, 81: 3655-3659 (1984)およびBerkner, *BioTechniques*, 6: 616-629 (1988)に記載のように、全ての主要な後期アデノウイルスmRNAの5'末端に存在するアデノウイルスの3つに分かれたリーダー配列は、2型アデノウイルスゲノムのヌクレオチド位置6071-6079(第

10

20

30

40

50



1 リーダーセグメントの3'末端)、7101-7172(第2リーダーセグメント全体)および9634-9721(第3リーダーセグメント)に相当する3つの空間的に離れたエクソンによりコードされている。しかしながら、3つに分かれた配列はまた、ヌクレオチド位置6081-6089(第1リーダーセグメントの3'末端)、7111-7182(第2リーダーセグメント全体)および9644-9845(第3リーダーセグメントおよびそのセグメントの下流にある配列)に相当する3つの空間的に離れたエクソンを持つAd5リーダー配列との一致を示す。pCLFに存在する3つに分かれたリーダー配列の対応cDNA配列は、ヌクレオチド位置907-912ないし1228-1233それぞれにBamHI/BglII 5'および3'部位が接している配列番号8に列挙する。

pCLFプラスミドは、実施例3に記載のようにATCCに寄託されている。pCLFの完全ヌクレオチド配列は、配列番号8に列挙するが、ここではヌクレオチド位置1は、pcDNA3親ベクター配列のほぼ中間に相当する。Ad5ファイバー遺伝子の5'および3'末端は、それぞれATGではヌクレオチド位置1237-1239に、TAAでは2980-2982に位置する。残りのベクター構築物は、既に上記した。

#### C. 機能的E4およびファイバータンパク質をコードするプラスミドを持つアデノウイルスパッケージング細胞系の製造

Graham et al., J. Gen. Virol., 36: 59-74 (1977)により製造され、更にSpector, Virol., 130: 533-538 (1983)に特性化されているとおり、293細胞系は既にE1遺伝子を含むので、第1アデノウイルスパッケージング系を製造する場合もこの細胞系を選択した。電気穿孔の前に、293細胞をRPMI培地+10%ウシ胎児血清中で増殖させた。BioRad Gene Pulserを用い、300V、25μFに設定して、 $4 \times 10^6$ 細胞をpE4/Hygro DNAおよびpCLF DNAそれぞれ20μgと共に電気穿孔した。電気穿孔用のDNAは、Qiagenシステム(Bio-Rad, Richmond, CA)を用いて製造元指導書に従い製造した。

電気穿孔の後、細胞を、200μg/mlヒグロマイシンB(Sigma, St. Louis, MO)を含む新鮮な完全DMEM+10%ウシ胎児血清中に分配した。

展開されたコロニーから、“MICROTURBOGEN”システム(Invitrogen)を用いて製造元指導書に従い、ゲノムDNAを単離した。組込まれたE4 DNAの存在は、PCRによりプライマー対E4RおよびORF6L(5'TGCTTAAGCGGCCGCGAAGGAGAGAGTCC3') (配列番号9)を用いて評価した。ここで、後者は、5'前方プライマー近辺アデノウイルス5オープンリーディングフレーム6である。E4遺伝子に関連するプライマーの位置については図1参照。

親細胞系293でみられた増殖特性とは違う増殖特性を示した1クローンを選択し、これを211と呼んだ。この211クローンは、予想生成物を含むし、pE4/Hygroプラスミドに含まれる、全てではないが、殆どのE4フラグメントに相当する挿入DNAの存在を示した。211細胞系は、実施例3に記載のようにATCCに寄託されている。この系は、更に、上記プライマー対E4L/E4Rを用いる増幅により評価し、全長E4挿入物に相当する生成物を検出した。ゲノムサザンブロッティングをEcoRIおよびBamHIで制限したDNAにて実施した。次いで、製造元指導書に従いGeniumシステム(Boehringer Mannheim, Indianapolis, IN)での標識化プローブとして使用するためpBS/E4にクローニングしたEcoRI/BamHI E4フラグメントを持つ標準と比較して、およそ1コピー/ゲノムでE4フラグメントを検出した。211細胞系由来のDNAでは、予想標識化内部フラグメントpE4/Hygroを、単離したE4配列とハイブリダイズさせた。更に、プローブを第2の挿入事象の結果であり得るより大きいフラグメントにハイブリダイズさせた(図5)。

211細胞系はファイバー遺伝子の欠損を確認するネオマイシン耐性では選択されないが、ファイバー遺伝子の不在を示すため、この211細胞系を、抗ファイバーポリクローナル抗体と2次抗体としてFITC-標識化抗ウサギIgG(KPL)による間接的な免疫蛍光によりファイバータンパク質の発現について分析した。免疫反応性は、なんら検出されなかった。従って、組換えファイバー遺伝子を含む211クローンを作るために、

10

20

30

40

50

211 クローンを、RPMI 培地中で増殖させて展開させ、ファイバーをコードする上記の pCLF プラスミドで追加的に電気穿孔にかけた。

電気穿孔後、細胞を DMEM + 10% ウシ胎児血清中でプレートにまき、コロニーを 200 µg/ml G418 (Gibco, Gaithersburg, MD) で選択した。ポジティブ細胞系は、ヒグロマイシン耐性を維持した。次いで、これらの候補となる 211 のサブラインを、上記の間接的免疫蛍光によりファイバータンパク質発現についてスクリーニングした。スクリーニングした 3 つのサブライン、211 A、211 B および 211 R は、他の多くのサブラインと共に、全て、AdRSVgal (1 pfu/細胞) に感染させ、感染後 24 時間で染色されたポジティブコントロールの 293 細胞と比べ、性質的に核染色を示した。

次いで、このアッセイでの核染色にポジティブな系を、同じ抗体を用いる変性条件下でウェスタン・ブロット分析にかけた。抗体が予想分子量 (Ad5 ファイバータンパク質の場合 62 kd) のタンパク質を検出した、211 A、211 B および 211 R を含む幾つかの系を更に研究するために選択した。211 A 細胞系は、実施例 3 に記載のように ATCC に寄託されている。

これら 3 つの細胞系からの可溶性核抽出物を用いるウェスタンブロット分析と半天然電気泳動システムから、発現されたファイバータンパク質が図 6 に示したように、天然ファイバータンパク質の特徴から機能的な三量体形であることが示された。三量体化ファイバーの予想分子量は、186 kd である。293 と印したレーンはファイバーを欠くが、サブラインは検出可能なファイバーを含有する。図 6 に示すように、変性条件下で、三量体形は破壊されて、検出可能なファイバー単量体になった。内生 E1、新規に発現した組換え E4 およびファイバータンパク質を含有するクローンを選択し、実施例 2 に記載のように欠失した対応アデノウイルス遺伝子を持つアデノウイルス遺伝子運搬ベクターに相補させるのに使用した。

#### D. E1 - 遺伝子欠失アデノウイルスの相補性に対する E1 - 発現プラスミドの製造

上記実施例 1 C に記載の 293 細胞系を含有する E1 - 遺伝子に基づくもの以外のアデノウイルスパッケージング細胞系を製造するために、単独または E4 と様々に組み合わせた E1 とファイバー遺伝子を含有するプラスミドベクターを下記のとおり構築する。

E1 a および E1 b 遺伝子を含有するアデノウイルスゲノムの領域は、前記のとおりウイルスゲノム DNA から PCR により増幅させる。使用したプライマーは、E1 L、即ち、5' または前方プライマー、および E1 R、即ち、3' または後方プライマーであり、それぞれヌクレオチド配列 5' CCGAGCTAGCGACTGAAAAATGAG3' (配列番号 10) および 5' CCTCTCGAGAGACAGCAAGACAC3' (配列番号 11) を持つ。E1 L および E1 R プライマーは、下線で示したとおり、それぞれ制限部位 NheI および XhoI を含む。この部位は、増幅させた E1 遺伝子フラグメントを、Clontech (Palo Alto, CA) から市販されている pMAM の NheI / XhoI 部位内にクローニングし、11152 bp のプラスミド pDEX/E1 を形成するのに使用する。そのプラスミド地図は図 7 に示す。

pDEX/E1 の完全ヌクレオチド配列は、配列番号 12 に列挙するが、ここでヌクレオチド位置 1 は、pMAM 主鎖ベクター配列の 3' 末端からおよそ 1454 ヌクレオチドに相当する。pDEX/E1 プラスミドは、pMAM のグルココルチコイド誘発マウス乳癌ウイルス (MMTV) プロモーターの下流に位置するアデノウイルスゲノム (pDEX/E1 プラスミド中、ヌクレオチド位置 1460 から始まり、4998 で終わる) のヌクレオチド 552 から 4090 を含む。pMAM ベクターは、ヒポキサンチン/アミノプテリン/チミジン (HAT) 選択を用いて安定なトランスフェクション体を単離可能にする E. コリ gpt 遺伝子を含有する。pMAM 主鎖は、配列番号 12 のヌクレオチド位置 1 - 1454 および 5005 - 11152 を占める。

#### E. 機能的 E1 およびファイバータンパク質をコードするプラスミドを持つアデノウイルスパッケージング細胞系の作成

実施例 1 C に記載のように、211 サブライン、211 A、211 B および 211 R のものに匹敵する別個のアデノウイルスパッケージング細胞系を作るために、アデノウイルス

10

20

30

40

50

ゲノムを欠く別の細胞系を、下記のプラスミド構築物とのトランスフェクション用を選択した。許容できる宿主細胞には、A 5 4 9、ヒーラー、ペロ、および実施例 1 の記載と同様の細胞系がある。選択した細胞系を別のプラスミド p D E X / E 1 および p C L F でそれぞれトランスフェクションさせて、E 1 およびファイバー相補性タンパク質を発現させる。前記のトランスフェクション操作の後に、2つのプラスミドの安定な挿入物を含むクローンを、ネオマイシンおよび H A T での選択により単離した。E 1 遺伝子の全長コピーの組み込みは、上記のプライマーセット E 1 L / E 1 R を用いるゲノム D N A からの P C R 増幅により評価する。ファイバー遺伝子の機能的挿入は、前記の抗ファイバー抗体での染色によりアッセイする。

その後、生じた安定に組込まれた細胞系は、実施例 2 に記載のように、対応するアデノウイルス遺伝子欠失を持つアデノウイルス遺伝子運搬ベクターに相補するパッケージング細胞システムとして使用する。

#### F. 相補性遺伝子欠失アデノウイルス用の 2 以上のアデノウイルス遺伝子を含むプラスミドの製造

前述の実施例に記載の方法は、アデノウイルス細胞パッケージングシステム作成用の 2 つのプラスミド p E 4 / Hygro と p C L F、または p C L F と p D E X / E 1 の使用によるものである。本発明の方法での使用が企図される別の実施態様では、様々な組み合わせでコードされたタンパク質の発現について 2 以上のアデノウイルス遺伝子を含む相補性プラスミドもまた、下記のとおり製造される。この場合、得られたプラスミドは、対応するアデノウイルス遺伝子欠失を持つ運搬プラスミドと共に様々な細胞系で使用する。よって、本発明の組換えアデノウイルスのウイルス性ベクターを作るのに使用するためのパッケージング細胞の選択、運搬プラスミドの内容、相補性プラスミドの内容は、その他のアデノウイルス遺伝子がアデノウイルスファイバー遺伝子と共に欠失するかどうか、もしそうならば、どちらの方であるかによって代わる。

##### 1. ファイバーおよび E 1 アデノウイルス遺伝子を含む相補性プラスミドの製造

C M V プロモーター、アデノウイルスの 3 つに分かれたリーダー、ファイバー遺伝子およびウシ成長ホルモントミネーターに対する配列を含む D N A フラグメントは、p C L F 中の p C D N A 3 ベクター主鎖のヌクレオチド 1 - 1 9 にアニールする前方プライマー 5 ' G A C G G A T C G G G A G A T C T C C 3 ' (配列番号 1 3)、および p C D N A 3 ベクター主鎖のヌクレオチド 1 2 7 8 - 1 2 5 7 にアニールする後方プライマー 5 ' C C G C C T C A G A A G C C A T A G A G C C 3 ' (配列番号 1 4) を用いて実施例 1 B で製造した p C L F から増幅させる。前記のとおりフラグメントを増幅させ、次いで、実施例 1 D で製造した p D E X / E 1 プラスミド中にクローニングする。D N A フラグメントでクローニングするために、まず、p D E X / E 1 ベクターを Nde I で消化して、p D E X / E 1 中の p M A M ベクター主鎖のユニーク部位で切断し、次いで、末端をバクテリオファージ T 4 ポリメラーゼおよび d N T P での処理により修復する。

得られた p E 1 / ファイバーと呼ばれる、E 1 およびファイバー遺伝子含有プラスミドは、D E X / E 1 について記載したデキサメタゾン誘発 E 1 機能と上記の A d 5 ファイバータンパク質の発現の両方を提供する。1 4 4 5 5 b p の p E 1 / ファイバーのプラスミド概略地図は、図 8 に示す。

p E 1 / ファイバーの完全ヌクレオチド配列は、配列番号 1 5 に列挙するが、ここで、ヌクレオチド位置 1 は、親ベクター p M A M 配列の 3 ' 末端からおよそ 1 4 5 9 ヌクレオチドに相当する。A d 5 E 1 遺伝子の 5 ' および 3 末端は、p M A M 主鎖に続くヌクレオチド位置 1 4 6 0 および 4 9 9 8 にそれぞれ位置し、次いで、充填平滑末端化 Nde I 部位により p C L F 由来の A d 5 ファイバーから分離する。p C L F ファイバー遺伝子フラグメントの 5 ' および 3 ' 末端は、p C L F について前述したエレメントを含むヌクレオチド位置 1 0 9 2 2 - 1 4 2 2 3 にそれぞれ位置する。

次いで、得られた p E 1 / ファイバープラスミドを用いて、E 1 およびファイバーを発現する 1 以上の運搬プラスミドに相補させる。

それから、p E 1 / ファイバー構築物を用いて実施例 1 E に記載の選択した宿主細胞をト

10

20

30

40

50

ランスフェクションさせ、前述のとおり実施して安定な染色体挿入物を作り、その後 H A T 培地で選択する。その後、安定な細胞を実施例 2 に記載のパッケージング細胞として使用する。

## 2) E 4 およびファイバーアデノウイルス遺伝子を含有する相補性プラスミドの製造

実施例 1 B に記載のように製造した p C L F を、Bgl I I で部分的に消化し、p C D N A 3 主鎖のその部位のみで切断する。実施例 1 A で製造した p E 4 / Hygro プラスミドを Bam H I で消化し、E 4 含有フラグメントを作る。次いで、E 4 フラグメントを p C L F の Bam H I 部位に挿入し、プラスミド p E 4 / ファイバーを形成させる。得られたプラスミドにより、p C L F のところで記載したファイバー遺伝子と p E 4 / Hygro のところで記載した E 4 機能が発現される。

10

1 0 6 1 0 b p の p E 4 / ファイバーの概略プラスミド地図は図 9 に示す。p E 4 / ファイバーの完全ヌクレオチド配列は配列番号 1 6 に列挙するが、ここで、ヌクレオチド位置 1 は、親ベクター p C D N A 3 主鎖配列の 3 ' 末端からおよそ 1 4 b p に相当する。A d 5 E 4 遺伝子の 5 ' および 3 末端は、ヌクレオチド位置 2 1 および 3 1 4 9 それぞれに位置し、これらは Bgl I I / Bam H I 部位と融合し、そして C M V プロモーターを含む p C D N A 3 主鎖が再び Bgl I I / Bam H I 部位に続く。アデノウイルスリーダー配列は、ヌクレオチド位置 4 0 5 1 から始まり、4 3 6 6 まで伸びて Bam H I / Bgl I I 部位と融合し、そして、ファイバー遺伝子の 5 ' および 3 ' 末端は、ヌクレオチド位置 4 3 7 2 および 6 1 2 4 にそれぞれ位置する。

宿主細胞における p E 4 / ファイバーの安定な染色体挿入物は、上記のようにして得られる。

20

## 実施例 2

### アデノウイルスパッケージング細胞系を用いるアデノウイルス遺伝子運搬ベクターの製造

本発明のアデノウイルス運搬ベクターは、E 1 / ファイバーおよび E 4 / ファイバーの組み合わせを別々に欠くように製造する。このようなベクターは、複数のウイルス遺伝子がないため、前に使用したものよりも複製欠損性である。複製コンピテントであるが、非ファイバー手段を介するのみである、本発明の好ましいアデノウイルス運搬ベクターは、単にファイバー遺伝子を欠くだけのベクターであるが、但し、残りの機能的なアデノウイルス調節および構造遺伝子を含有する。更に、本発明のアデノウイルス運搬ベクターは、外来 D N A の挿入に対してより高い能力を持つ。

30

### A . 特定の遺伝子欠失を持つアデノウイルス遺伝子運搬ベクターの製造と使用法

Lac Z レポーター遺伝子構築物を含有する E 1 / ファイバー欠失ウイルスベクターを構築するために、2 つの新規プラスミドを構築した。プラスミド p E 1 B gal を下記のように構築した。p S V gal (ProMega Corp., Madison, WI) を Vsp I で消化することにより、S V 4 0 調節配列および E . コリ- - ガラクトシダーゼ遺伝子を含有する D N A フラグメントを単離した。得られたフラグメントは突出末端を持ち、次いで、これを、d N T P の存在下で D N A ポリメラーゼ 1 のクレノウフラグメントで充填し、Bam H I で消化した。得られたフラグメントを p E 1 sp1 B (Microbix Biosystems, Hamilton, Ontario) のポリリンカー中の Eco R V および Bam H I 部位にクローニングし、p E 1 B gal を形成させたため、これは、p S V gal の Lac Z カセット (ヌクレオチド 6 6 9 0 から 4 1 5 1) により置換された E 1 a 領域と共にアデノウイルスゲノムの左端を含有した。プラスミド D N A は、プラスミドの展開 (expand) に使用した形質転換細胞から Birnboim and Doly, Nuc. Acids Res., 7: 1513-1523 (1978) に記載のとおり、アルカリ性溶解法により製造した。次いで、D N A を CsCl - エチジウムブロマイド密度勾配遠心により精製した。

40

本明細書に記載のようにして製造した第 2 プラスミド (p D V 4 4) は、Bett et al., Proc. Natl. Acad. Sci., USA, 91: 8802-8806 (1994) に記載のように製造かつ Microbix から市販されているベクター p B H G 1 0 から得られ、このベクターは、欠失した左端にパッケージングシグナルと共に A d 5 ゲノムを、そして制限酵素 Pac I に対するユニーク部位と共にリンカーにより置換された E 3 領域 (ヌクレオチド 2 8 1 3 3 : 3 0 8 1 8) 50

を含有する。アデノウイルスゲノムの右端を含有する 11.3 kb の BamH I フラグメントを pBHG10 から単離し、pBS/SK(+) の BamH I 部位にクローニングして、およそ 14,658 bp のプラスミド p11.3 を作る。その概略プラスミド地図は図 13 に示す。次いで、p11.3 プラスミドを PacI および SalI で消化してファイバー、E4 および逆方向末端反復 (ITR) 配列を除去した。

このフラグメントを ITR セグメント含有フラグメントで置換し、次のオリゴヌクレオチド配列 (配列番号 17) (配列番号 18) を用いて pBHG10 からの PCR 増幅により E4 遺伝子を作る。これらのプライマーは、それぞれ PacI および BamH I に対する部位を組み込んでいる。このフラグメントを p11.3 主鎖の PacI と平滑末端化 SalI 部位にクローニングすると、ITR および E4 領域単独により、pBHG10 に存在する融合 ITR、E4 領域およびファイバー遺伝子の置換が起こった。

一般に、細胞培養でのプラスミドの組換えと、それに続く相補性によるウイルス生産法は、実施例 1 で製造したアデノウイルスパッケージング細胞系、即ち 211A、211B、211R、A549、ペロ細胞などのいずれか 1 つと、ウイルス遺伝子運搬ベクターに相当する配列を持つプラスミドとの同時トランスフェクションによる組換えウイルスの単離を含む。

選択した細胞系をペトリ皿で平板培養し、Bett et al., *Proc. Natl. Acad. Sci., USA*, 91: 8802-8806 (1994) に記載のリン酸カルシウム法を用いて、pDV44 および pE1Bgal と同時トランスフェクションする。2 つのプラスミド中の重複アデノウイルス配列間の組換えにより、pDV44 と pE1Bgal が再結合して、複数の欠失を持つ組換えアデノウイルスベクターを形成している全長ウイルス染色体が作られる。ウイルス染色体からの E1 およびファイバー遺伝子の欠失は、パッケージング細胞ゲノムに組み込まれた配列により補われ、感染性ウイルス粒子が作られる。よって、生じたプラークを単離し、組換えウイルスのストックを標準法で生成する。

本発明の好ましい実施態様では、ファイバー遺伝子欠失を持つ治療用ウイルスベクターを調整するための上記組換え事象を必要としない運搬プラスミドを製造する。パッケージングに必要なアデノウイルスゲノム全てを含有するが、ファイバー遺伝子は欠いている単一の運搬プラスミドは、Microbix から市販されている全長 Ad85 を含有するプラスミド pFG140 から製造する。次いで、pFG140-f と呼ぶ得られた運搬プラスミドを、上記のように、pCLF を安定に組み込んだ細胞と共に使用して、ファイバーを欠く治療用ウイルスベクターを製造する。本発明の好ましい態様では、ファイバー遺伝子を、治療用運搬アデノウイルスベクター製造の対象となる治療用遺伝子で置換する。

所望の治療用遺伝子を運搬するためのベクターは、上記 pE1Bgal 製造の際に実施したのと同様にして、対象となる遺伝子を市販の pE1sp1B (Microbix Biosystems) のポリリンカー中のマルチクローニングサイトにクローニングすることにより製造する。次いで、同じ同時トランスフェクションおよび組換え法を本明細書に記載のように実施して、ウイルス遺伝子運搬ベクターを得る。

こうして生産した組換えウイルスは、培養細胞およびインビボのいずれもで遺伝子運搬ツールとして使用される。多重欠失ベクターの効果および相対的免疫原性について研究するため、実施例 1 に記載のパッケージング系にて増殖させることにより、ウイルス粒子を生産し、CsCl 勾配遠心により精製する。力価測定 (titering) 後、ウイルス粒子を全身または局所注射により、または肺へのエアゾール送達によりマウスに投与する。LacZ リポーター遺伝子は、うまく形質導入された細胞の数および種類の評価を可能にする。導入遺伝子発現の持続期間を評価して、現在までに臨床試験に使用されてきた標準技術に対する多重欠失組換えアデノウイルス処置の長期間の効果を測定する。ここに記載した改良ベクターに対する免疫応答は、炎症、ベクターに対向する細胞毒性 T リンパ球の産生およびウイルスタンパク質に対向する抗体応答の性質および規模などのパラメーターを評価することにより測定する。

嚢胞性繊維症の処置のための CFT R または癌の処置のための腫瘍抑制遺伝子などの治療用遺伝子を含有する数種のベクターを、遺伝子導入および発現の安全性と効率について動

10

20

30

40

50

物系にて評価する。この評価の後、ヒト臨床治験において実験的治療剤としてこれらを使用する。

B. 異種または変化したファイバータンパク質を含有するウイルス粒子を生産することによる、アデノウイルス遺伝子運搬ベクターの再ターゲティング

標的細胞に結合するアデノウイルスの特異性は、ファイバータンパク質により広範囲に測定されるので、修飾ファイバータンパク質または異なるアデノウイルス血清型（偽型ベクター）由来のファイバータンパク質を組み込むウイルス粒子は、異なる特異性を持つ。よって、上記のアデノウイルスパッケージング細胞における天然 A d 5 ファイバータンパク質の発現もまた、異種ファイバータンパク質の生産に応用できる。

本発明の一態様では、キメラファイバータンパク質は、Stevenson et al., *J. Virol.*, 69: 2850-2857 (1995) の方法に従い、製造する。この著者は、ファイバー受容体結合活性の決定因子が、ファイバーのヘッドドメインに位置すること、および単離したヘッドドメインが三量体化して細胞受容体に結合する能力があることを示した。3 型アデノウイルス (A d 3) および A d 5 のヘッドドメインは、キメラファイバータンパク質を生産するために交換された。本発明の方法に使用するキメラファイバータンパク質をコードする同様の構築物も包含される。よって、アデノウイルスパッケージング細胞における相補性ウイルスベクターとして実施例 1 で製造した無傷の A d 5 ファイバー-コード構築物を使用する代わりに、本明細書に記載した構築物を用いて、E 4 および/または E 1 -コード構築物とともに、細胞をトランスフェクションする。

要するに、全長 A d 5 および A d 3 は、鋳型として精製アデノウイルスゲノム DNA から増幅させた。A d 5 および A d 3 ヌクレオチド配列は、それぞれ、GenBank 受託番号 M 1 8 3 6 9 と M 1 2 4 1 1 で入手できる。オリゴヌクレオチドプライマーは、開始コドン A T G から始まり、終止コドン T A A で終わる全長ファイバー遺伝子の全コード配列を増幅するように設計する。クローニングの目的では、5' および 3' プライマーは、実施例 1 A に記載のように、p c DNA プラスミドにクローニングするため、それぞれの制限部位 BamH I および Not I を含有する。P C R は上記のように実施する。

その後、得られた産物を用いて、Horton et al., *BioTechniques*, 8: 525-535 (1990) に記載の P C R 遺伝子重複伸長 (P C R gene overlap extension) によりキメラファイバー構築物を構築する。A d 5 ファイバーテイルおよびシャフト領域 (5 T S ; アミノ酸残基位置 1 から 4 0 3 をコードするヌクレオチド領域) を A d 3 ファイバーヘッド領域 (3 H ; アミノ酸残基位置 1 3 6 から 3 1 9 をコードするヌクレオチド領域) につなげて、5 T S 3 H ファイバーキメラを形成させる。逆に、A d 3 ファイバーテイルおよびシャフト領域 (3 T S ; アミノ酸残基位置 1 から 1 3 5 をコードするヌクレオチド領域) を A d 5 ファイバーヘッド領域 (5 H ; アミノ酸残基位置 4 0 4 から 5 8 1 をコードするヌクレオチド領域) につなげて、3 T S 5 H ファイバーキメラを形成させる。ファイバーシャフト-ヘッド接合部の保存 T L W T (配列番号 2 1) 配列で融合を行う。

次いで、得られたキメラファイバー P C R 産物を BamH I および Not I で消化し、同様に消化した p c DNA ベクター内へ別個に指向性ライゲーション (directional ligation) する。次いで、A d 2 リーダー配列を実施例 1 A に記載のように BamH I にサブクローニングし、上記 2 1 1 細胞への、または前述の別のパッケージング細胞系への 2 次トランスフェクション用の発現ベクターを製造する。次いで、得られたキメラファイバー構築物含有アデノウイルスパッケージング細胞系を用いて、前述のアデノウイルス運搬ベクターに相補させる。様々な既知のアデノウイルス血清型でも同様のアプローチを用いて、その他のファイバーキメラ構築物が得られる。

別の実施態様では、本発明の方法は、出典明示により本明細書の一部とする Michael et al., *Gene Therapy*, 2: 660-668 (1995) や国際公開 W O 9 5 / 2 6 4 1 2 に記載の新規エピトープを含む修飾タンパク質の使用を企図する。いずれの刊行物も、内生ウイルス結合特異性の破壊と共にウイルス内へ組み込まれる新しい結合特異性を持つ細胞型特異的治療用ウイルスベクターの構築を記載している。特に、その筆者らは、A d 5 ファイバー遺伝子のコード配列の 3' 末端でガストリン放出ペプチド (G R P) をコードするアデノウイ

10

20

30

40

50

ルスベクターの製造を記載していた。得られたファイバー-G R P 融合タンパク質は発現されて、合成後にヒラー細胞の核へと正確に運ばれる機能的ファイバー三量体をアセンブルすることを示した。

この文献と国際公開公報の教示に基づき、同様の構築物が、複製欠損性かつ低免疫原性の新規アデノウイルス遺伝子運搬ベクターを作る本発明の相補性アデノウイルスパッケージング細胞系に使用できることが企図される。ファイバー特異性を再指向するためにここで使用が意図される異種リガンドは、アミノ酸サイズ10程度から大きい球形構造までの範囲であり、その幾つかは、ファイバーによる異種リガンドの立体障害を低減または排除するか、またはファイバータンパク質の三量体化を防ぐためにスペーサー領域の付加を必要とする。リガンドは、末端に、またはリンカー領域内に挿入される。好ましいリガンドには、特異的細胞受容体を標的とするもの、またはビオチンおよびアビジンなどの他の部分に結合させるために使用されるものがある。リガンドが結合した結果として起こる細胞シグナルの種類は、そのリガンドの特異性、即ち、受容体インターナリゼーションまたはその欠乏によって変わる。

好ましいスペーサーには、各末端でプロリン残基とフランクした一連のセリンおよびアラニンから成る短い12アミノ酸ペプチドリンカーがある。当業者は、十分なタンパク質特性を達成するための、およびウイルスインターナリゼーションに続く細胞事象を害することなくファイバータンパク質結合特異性を変えるためのリンカーの製造に精通している。更に、本発明の状況では、本明細書に記載の方法により使用される、下記のアミノ末端およびカルボキシ末端でファイバータンパク質の内部に配置されるリガンドを持つ修飾ファイバーの製造も企図される。

異種結合リガンドを持つファイバーの製造は、実質的に上記引用文献に記載のとおり製造される。要約すると、選択したリガンドについて、部位指向性突然変異誘発法を使用して、リンカーのコード配列を実施例1で製造したp C L F 中のA d 5 ファイバー構築物の3'末端のNot I 部位に挿入する。3'またはアンチセンスオリゴヌクレオチドは、選択した異種リガンドに対するコード配列の挿入を可能にし、確実に正しく翻訳を終止する、P r o S e r A l a S e r A l a S e r A l a S e r A l a P r o G l y S e r ( 配列番号22 ) の好ましいリンカー配列、続く、ユニーク制限部位および2つの停止コドンそれぞれをコードする。アンチセンスオリゴヌクレオチドの3'末端は、オリゴヌクレオチドが部位指向性突然変異誘発法を介して挿入されているベクター配列と重複する配列を含む。リンカーおよび停止コドン配列を加えるp C L F 配列の突然変異誘発の後、予め選択したリガンドをコードするヌクレオチド配列を得、そのユニーク制限部位に相当するリンカーを結合させ、次いで、その配列に対応する線状化制限部位にクローニングする。

リンカーおよびリガンドをコードする3'ヌクレオチドと共にA d 5 ファイバー遺伝子配列を含有する得られたp C L F ベクターに、前述のようにA d 2 リーダー配列を挿入する。それから、得られたファイバー-リガンド構築物を用いて、前述の211または別の細胞パッケージング系をトランスフェクションし、本発明の方法に使用する相補性ウイルスベクターパッケージング系を産生する。

更なる実施態様では、異なるアデノウイルス血清型から単離したファイバー遺伝子によりコードされるファイバータンパク質を前述の211または別の細胞パッケージング系にトランスフェクションするために無傷で使用する。

まず、対象となるファイバータンパク質をコードする遺伝子をクローニングして、p C L F に類似のプラスミドを作り、ファイバータンパク質を産生する安定な細胞系をA d 5 ファイバーについて上記したようにして製造する。次いで、ファイバー遺伝子を欠いている記載のアデノウイルスベクターを、すぐ使う目的に適したファイバータンパク質を産生する細胞系にて増殖させる。存在する唯一のファイバー遺伝子のみがパッケージング細胞中のものであるため、産生されたアデノウイルスは、対象のファイバータンパク質のみを含有し、よって、相補性タンパク質により与えられる結合特異性を持つ。実験動物系におけるその特性を測定するために上記したような研究においてこのようなウイルス粒子を使用する。

10

20

30

40

50

### C. ファイバータンパク質を欠くウイルスベクター粒子を用いる標的化遺伝子運搬

造血細胞のアデノウイルス感染に入る別の様式は、Huang, et al., J. Virol., 69: 2257-2263 (1995) に記載されており、これはファイバータンパク質-宿主細胞受容体相互作用を必要としない。多数のその他の細胞型の感染は、ファイバータンパク質の存在を必要とするので、ファイバーを欠くベクター粒子は、単球またはマクロファージなどの造血細胞に優先的に感染する。

ファイバーなしのアデノウイルスベクター粒子を製造するためには、上記実施例 2 A に記載のように、ファイバー遺伝子は欠くが、運搬対象の遺伝子は含有するベクターを、ファイバータンパク質を産生しない細胞、例えば、実施例 1 で製造した 211 細胞中で増殖させ、それによって、多数のファイバータンパク質欠損粒子を産生させることにより増幅させる。次いで、アデノウイルスベクターの他の領域により提供されるターゲティング機構を介して、即ち、天然ペントン塩基を介して、本発明の方法に従い、回収したファイバーなしのウイルス粒子を対象の挿入遺伝子の送達に使用する。

#### 実施例 3

##### 材料の寄託

下記の細胞系およびプラスミドは、1996年9月25日、アメリカ合衆国、メリーランド州、ロックビル、パークラウンドドライブ1301番、アメリカン・タイプ・カルチャー・コレクション(ATCC)に寄託されている：

| 材料              | ATCC受託番号  |
|-----------------|-----------|
| プラスミド pE4/Hygro | 97739     |
| プラスミド pCLF      | 97737     |
| 211細胞系          | CRL-12193 |
| 211A細胞系         | CRL-12194 |

前記の詳述は、当業者が本発明を実施するのに十分であると考えられる。寄託された具体物は本発明の一態様の単なる例示であると意図されるので、本発明は、寄託された細胞系やプラスミドにより範囲限定されるものではなく、また機能的に均等ないかなる細胞系またはプラスミドベクターも本発明の範囲内にある。

特定の実施態様や実施例を含む前記詳述は、本発明の例示であり、限定と解釈されるものではない。その他多くの変形や修飾が、本発明の真の精神および範囲から逸脱することなく、実行できる。

#### 実施例 4

天然ファイバータンパク質は、ホモ三量体であり (Henry L. J. et al 1994 Characterization of the knob domain of the adenovirus Type 5 fiber protein expressed in *Escherichia coli*, J. Virol. 68: 5239-5246)、三量体化は、ペントン/ファイバー複合体のアセンブリーに必須である (Novelli A et al 1991 Assembly of adenovirus type 2 fiber synthesized in cell-free translation system. J. Biol. Chem. 266: 9299-9303)。細胞系により産生された組換えファイバータンパク質の多量体構造を評価するために、細胞を  $50 \mu\text{Ci/ml}$  [ $^{35}\text{S}$ ] Translabel (ICN) により 37 で 2 時間標識し、RIPA 緩衝液中で溶解させ、ファイバータンパク質を (Harlow E. et al., 1988, Antibodies. Cold Spring Harbour Laboratory, cold Spring Harbour) に記載のとおり免疫沈降させた。免疫複合体をプロテイン A - セファロースビーズ (Pierce) 上にて集め、RIPA 緩衝液で十分に洗浄し、0.1 M トリエチルアミン、pH 11.5 中、室温でインキュベーションして、結合ファイバータンパク質を解離させた。沈殿したファイバーの一部を、変性 (充填緩衝液中 1% SDS、試料を 5 分間煮沸する) または半天然 (充填緩衝液中 0.1% SDS、試料は加熱しない) 条件下、8% SDS-PAGE ゲルにて電気泳動した。図 13 から分かるように、対照の 293 細胞以外の系 211A、211B および 211R は、三量体 (186 kD) の場合は半天然条件下、単量体 (62 kD) の場合は変性条件



下、予想分子量で移動した免疫学的に反応性のタンパク質を発現した。沈殿ファイバーの挙動は、精製したバキュロウイルス産生組換え A d 2 ファイバー (Wickham T., et al., 1993, Cell, 73: 309-319) と見分けがつかなかった (58 kD A d 2 および 62 kD A d 5 ファイバーは、これらの条件で非常によく似た移動度を持つ)。

ファイバー発現系がファイバー欠損アデノウイルスの増殖を維持するかどうかを測定するため、本発明者らは、温度感受性ファイバー変異体 A d H 5 ts 1 4 2 (Harold Ginsberg から提供) を用いて 1 段階増殖実験を実施した。制限温度 (39.5) で、この変異体は、成熟ピリオンに組み込まれないグリコシル化を受けたファイバータンパク質を産生する (Chee-Sheung C. C. et al., 1982, J. Virol. 42: 932-950)。この結果、非感染性ウイルス粒子が蓄積される。本発明者らは、本発明の細胞系により発現される組換えファイバータンパク質が H 5 ts 1 4 2 欠損を補い、ウイルス増殖を救出し得るかどうかを問うた。

細胞系 293、211A、211B および 211R ( $2 \times 10^6$  細胞/試料) を H 5 ts 1 4 2 に 10 pfu/細胞で感染させた。48 時間後、細胞を 25 mM EDTA で分離し、ウイルスを 4 回の急速凍結解凍サイクルにより回収した。1500 x g で 10 分間回転させて異物を除去し、ウイルス力価を蛍光フォーカスアッセイ (Thiel J. F et al., 1967, Proc. Soc. Exp. Biol. Med. 125:892-895) により、ポリクローナル抗ペントン塩基 A b (Wickham T. et al., 1993, Cell, 73: 309-319) を持つ SW 480 細胞にて、測定した。図 14 に示すように、ファイバー変異体ウイルスは、293 細胞中、32.5 (許容温度) で高力価まで複製したが、39.5 の制限温度では、その程度はかなり低かった。ファイバー産生パッケージング系 211A、211B または 211R は、293 細胞において、39.5 では、許容温度で見られたよりも 2 または 3 倍以内のレベルまでウイルス産生を維持し、これは、これらの細胞がファイバー欠損の部分的相補を提供することを示している。

興味深いことに、ファイバー産生細胞系からのウイルス収量もまた、32.5 (“許容” 温度) では 293 細胞からの収量よりもやや高かった。このことは、ts 1 4 2 ウイルスにより産生されたファイバーは、許容温度でさえも部分的に欠けている場合があることを示唆している。あるいは、アデノウイルス力価における非特異的増大は、ウイルスを本発明のパッケージング細胞中で増殖させるとき、ファイバー相補性を含まない機構によって、結果的に起こってもよい。しかしながら、本発明者らは、野生型ファイバー遺伝子を持つウイルス (A d . R S V b g a l など) は、本発明のパッケージング系でも 293 細胞でも同一レベルまで複製することが分かった (データ示さず)。総括すると、これらの結果は、H 5 ts 1 4 2 増殖において観察された増大がファイバー変異の特異的相補性によることを表している。

ファイバー - 発現細胞系においてでさえ、ファイバー変異体は、39.5 よりも 32 でより高い力価まで増殖する。この不完全な相補性は、野生型感染でみられるよりもやや低いレベルでのファイバーのパッケージング系発現によるものであり得る (図 16)。最近の研究では、同時的にファイバータンパク質発現の衰えた E 4 - 欠失ベクターは、産生ウイルスの力価の大きな低下をもたらすことが報告された (Brough et al., 1996, J. Virol., 70: 6497-6501)。その他の可能性は、制限温度で産生された欠損 ts 1 4 2 ファイバータンパク質が、細胞により産生された野生型タンパク質のいくつかと複合体を形成し、粒子へのアッセンブリーを阻止するかもしれないというものである。

異なる A d 血清型のファイバータンパク質は、そのシャフトドメインの長さおよびその受容体結合ノブドメインにおいて異なるが、ウイルスペントン塩基との相互作用を担う N - 末端領域は、高度に保存されている (Arnberg N. et al., 1997, Virology, 227: 239-244) (図 15 A)。このことは、異なる細胞-結合特異性を持つ多くのウイルス血清型由来のファイバーは、遺伝子運搬ベクターの生産に使用しやすいことを示唆している。

本発明者らは、パッケージング細胞により産生される組換え A d 5 ファイバーを、その他のアデノウイルス血清型の粒子に組み込んでもよいかどうかを問うた。3 型アデノウイルスは、ファイバー産生細胞系でも 293 細胞系でも増殖した。ウイルス粒子は、15 - 4

10

20

30

40

50

0 % CsCl 勾配を行い、2 連続遠心 ( 1 1 1 , 0 0 0 × g で 3 時間 ) により精製して、可溶性細胞性タンパク質を除去し、次いで、1 0 mM トリス-HCl、pH 8.1、1 5 0 mM NaCl、1 0 % グリセロールに対して十分に透析した。ポリクローナル抗ファイバー血清を用いる免疫プロットティング、次いで西洋ワサビペルオキシダーゼコンジュゲート化ヤギ抗ウサギ抗体 ( Kirkegaard and Perry Laboratories ) および ECL 化学発光基質 ( Amersham ) での検出により、Ad 5 ファイバータンパク質を検出した。精製した Ad 3 粒子は、ファイバー発現細胞系で 1 回継代した後に Ad 5 ファイバータンパク質を含有したが、2 9 3 細胞で継代した後では含有しなかった。前述の研究は、Ad 2 ファイバーがインビトロで Ad 3 ペントン塩基と相互作用できることを表しており ( Fender et al., 1997, *Nature Biotech.*, 15: 52-56 )、更に、これらの結果は、細胞により産生された 5 型ファイバータンパク質が完全な Ad 3 粒子内にアッセンブルする能力があることを表している。

10

Ad 5 ベースだが、Ad 7 ファイバータンパク質の遺伝子を含むベクターは、( Gall J. et al., 1996, *J. Virol.*, 70: 2116-2123 ) に記載されており、同じく、Ad 含有キメラファイバー遺伝子も ( Krasnykh V. N. et al., *J. Virol.*, 70: 6839-6846 ) に記載されている。異なる細胞タンパク質へ結合させるために短いペプチドリンカーをファイバーに付加することもまた報告されている ( 8188 )。本明細書に記載のようなパッケージング技術を用いることにより、異なるファイバータンパク質を備えた Ad ベクターを、各応用に対し、新規ベクターゲノムを作る時間のかかる工程を必要とせずとも、対照のファイバーを発現する細胞中で増殖させることにより簡単に製造できる。

20

ベクター染色体中のファイバー遺伝子を置換または修飾するにもまた、新規ファイバータンパク質は増殖する細胞の表面上の受容体に結合する必要がある。パッケージング細胞アプローチは、所望のファイバー遺伝子が発現する細胞中で 1 ラウンド増殖させることにより、もはやその宿主細胞には結合できないファイバー含有 Ad 粒子の製造を可能にする。これは、粒子内に取りこまれ得るファイバータンパク質のレパートリーを大きく広げるものであり、同時に、遺伝子運搬ベクターの再ターゲティングプロセスを簡略化するものである。

最後に、ファイバーに関係のない新規の感染経路が最近、造血細胞で報告され、その内容は、ペントン塩基が、インテグリン  $\alpha_5\beta_1$  に結合することにより初期ウイルス細胞相互作用を提供するというものであった ( Huang S. et al., 1996, *J. Virol.*, 70: 4502-4508 )。このことは、ファイバータンパク質を欠くウイルス粒子が、この経路を介した特定細胞型へのターゲティング遺伝子運搬に有用であり得ることを示唆している。

30

#### 配列表

( 1 ) 一般的情報

( i ) 出願人：ネメロー・グレン・アール

フォン・セジェーン・ダニエル・ジェー

( ii ) 発明の名称：高キャパシティーアデノウイルスベクターの開発の促進に使用するためのパッケージング細胞系

( iii ) 配列の数：2 0

( iv ) 連絡住所：

40

( A ) 宛名：ザ・スクリップス・リサーチ・インスティテュート

( B ) 通り：ノース・トリー・パインズ・ロード

( C ) 街：ラ・ホヤ

( D ) 州：カリフォルニア

( E ) 国：アメリカ合衆国

( F ) ZIP：9 2 0 3 7

( v ) コンピューター読解可能形式

( A ) 媒介型：フロッピーディスク

( B ) コンピューター：IBM PC 互換性

( C ) オペレーティング・システム：PC-DOS/MS-DOS 互換性

50

- ( D ) ソフトウェア : PatentIn Release #1.0, Version #1.25
- ( iv ) 現出願データ :
- ( A ) 出願番号 : U S
- ( B ) 出願日 : 1 9 9 6 年 9 月 2 5 日
- ( C ) 分類 :
- ( viii ) 代理人 / 会社情報 :
- ( A ) 氏名 : ローガン・エイプリル・シー
- ( B ) 登録番号 : 3 3 , 9 5 0
- ( C ) リファレンス / ドケット番号 : TSRI554.0
- ( ix ) 遠隔通信情報 : 10
- ( A ) 電話番号 : ( 6 1 9 ) 5 5 4 - 2 9 3 7
- ( B ) ファックス番号 : ( 6 1 9 ) 5 5 4 - 6 3 1 2
- ( 2 ) 配列番号 : 1 の情報 :
- ( i ) 配列の特性 :
- ( A ) 配列の長さ : 3 0 塩基対
- ( B ) 配列の型 : 核酸
- ( C ) 鎖の数 : 一本鎖
- ( D ) トポロジー : 直鎖状
- ( ii ) 配列の種類 : D N A ( genomic )
- ( iii ) ハイボセティカル : NO 20
- ( iii ) アンチセンス : NO
- ( xi ) 配列の記載 : 配列番号 : 1 :
- CGGTACACAG AATTCAGGAG ACACAACTCC 30
- ( 2 ) 配列番号 : 2 の情報 :
- ( i ) 配列の特性 :
- ( A ) 配列の長さ : 3 5 塩基対
- ( B ) 配列の型 : 核酸
- ( C ) 鎖の数 : 一本鎖
- ( D ) トポロジー : 直鎖状
- ( ii ) 配列の種類 : D N A ( genomic ) 30
- ( iii ) ハイボセティカル : NO
- ( iii ) アンチセンス : NO
- ( xi ) 配列の記載 : 配列番号 : 2 :
- GCCTGGATCC GGGAAGTTAC GTAACGTGGG AAAAC 35
- ( 2 ) 配列番号 : 3 の情報 :
- ( i ) 配列の特性 :
- ( A ) 配列の長さ : 1 2 塩基対
- ( B ) 配列の型 : 核酸
- ( C ) 鎖の数 : 一本鎖
- ( D ) トポロジー : 直鎖状 40
- ( ii ) 配列の種類 : D N A ( genomic )
- ( iii ) ハイボセティカル : NO
- ( iii ) アンチセンス : NO
- ( xi ) 配列の記載 : 配列番号 : 3 :
- CGCGGATCCG CG 12
- ( 2 ) 配列番号 : 4 の情報 :
- ( i ) 配列の特性 :
- ( A ) 配列の長さ : 8 7 1 0 塩基対
- ( B ) 配列の型 : 核酸
- ( C ) 鎖の数 : 二本鎖 50

( D ) トポロジー : 環状

( ii ) 配列の種類 : D N A ( genomic )

( iii ) ハイボセティカル : NO

( iii ) アンチセンス : NO

( xi ) 配列の記載 : 配列番号 : 4 :

|  |     |    |
|--|-----|----|
| CACCTAAATT GTAAGCGTTA ATATTTTGT AAAATTCGCG TTAAATTTTT GTTAAATCAG   | 60  |    |
| CTCATTTTTT AACCAATAGG CCGAAATCGG CAAAATCCCT TATAAATCAA AAGAATAGAC  | 120 |    |
| CGAGATAGGG TTGAGTGTTG TTCCAGTTTG GAACAAGAGT CCACTATTAA AGAACGTGGA  | 180 | 10 |
| CTCCAACGTC AAAGGGCGAA AAACCGTCTA TCAGGGCGAT GGCCCACTAC GTGAACCATC  | 240 |    |
| ACCCTAATCA AGTTTTTTTG GGTGAGGTG CCGTAAAGCA CTAAATCGGA ACCCTAAAGG   | 300 |    |
| GAGCCCCCGA TTTAGAGCTT GACGGGGAAA GCCGGCGAAC GTGGCGAGAA AGGAAGGGAA  | 360 |    |
| GAAAGCGAAA GGAGCGGGCG CTAGGGCGCT GGCAAGTGTA GCGGTCACGC TCGCGTAAC   | 420 |    |
| CACCACACCC GCCGCGCTTA ATGCGCCGCT ACAGGGCGCG TOCCATTGCG CATTGAGGCT  | 480 |    |
| GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCTCTTCG CTATTACGCC AGCTGGCGAA   | 540 | 20 |
| AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA GGGTTTTCCC AGTCACGACG  | 600 |    |
| TTGTAAAACG ACGGCCAGTG AATTGTAATA CGACTCACTA TAGGGCGAAT TGGGTACCGG  | 660 |    |
| GCCCCCCTC GAGGTCGACG GTATCGATAA GCTTGATATC GAATTCAGGA GACACAACCTC  | 720 |    |
| CAAGTGCATA CTCTATGTCA TTTTCATGGG ACTGGTCTGG CCACAACCTAC ATTAATGAAA | 780 |    |
| TATTTGCCAC ATCCTCTTAC ACTTTTTCAT ACATTGCCCA AGAATAAAGA ATCGTTTGTG  | 840 |    |

|   |      |    |
|---|------|----|
| TTATGTTTCA ACGTGTATTTAT TTTTCAATTG CAGAAAAATTT CAAGTCATTT TTCATTTCAGT | 900  |    |
| AGTATAGCCC CACCACCACA TAGCTTATAC AGATCACCGT ACCTTAATCA AACTCACAGA     | 960  |    |
| ACCCTAGTAT TCAACCTGCC ACCTCCCTCC CAACACACAG AGTACACAGT CCTTTCTCCC     | 1020 |    |
| CGGCTGGCCT TAAAAAGCAT CATATCATGG GTAACAGACA TATTCTTAGG TGTATATTC      | 1080 |    |
| CACACGGTTT CCTGTGAGC CAAACGCTCA TCAGTGATAT TAATAAACTC CCCGGGCAGC      | 1140 |    |
| TCACTTAAGT TCATGTGCT GTCCAGCTGC TGAGCCACAG GCTGCTGTCC AACTTGCGGT      | 1200 | 10 |
| TGCTTAACGG GCGGCGAAGG AGAAGTCCAC GCCTACATGG GGGTAGAGTC ATAATCGTGC     | 1260 |    |
| ATCAGGATAG GCGGGTGGTG CTGCAGCAGC GCGCGAATAA ACTGCTGCCG CCGCCGCTCC     | 1320 |    |
| GTCCTGCAGG AATACAACAT GGCAGTGGTC TCCTCAGCGA TGATTGCGAC CGCCCGCAGC     | 1380 |    |
| ATAAGGCGCC TTGTCCTCCG GGCACAGCAG CGCACCTGA TCTCACTTAA ATCAGCACAG      | 1440 |    |
| TAACTGCAGC ACAGCACCAC AATATTGTTT AAAATCCCAC AGTGCAAGGC GCTGTATCCA     | 1500 |    |
| AAGCTCATGG CGGGGACCAC AGAACCACG TGGCCATCAT ACCACAAGCG CAGGTAGATT      | 1560 | 20 |
| AAGTGGCGAC CCCTCATAAA CACGCTGGAC ATAAACATTA CCTCTTTTGG CATGTTGTAA     | 1620 |    |
| TTCACCACCT CCCGGTACCA TATAAACCTC TGATTAAACA TGGCGCCATC CACCACCATC     | 1680 |    |
| CTAAACCAGC TGGCCAAAAC CTGCCCGCCG GCTATACACT GCAGGGAACC GGGACTGGAA     | 1740 |    |
| CAATGACAGT GGAGAGCCCA GGAATCGTAA CCATGGATCA TCATGCTCGT CATGATATCA     | 1800 |    |
| ATGTTGGCAC AACACAGGCA CACGTGCATA CACTTCCTCA GGATTACAAG CTCCTCCCGC     | 1860 |    |
| GTTAGAACCA TATCCCAGGG AACAACCCAT TCCTGAATCA GCGTAAATCC CACTGTCAG      | 1920 |    |
| GGAAGACCTC GCACGTAAT CACGTTGTGC ATTGTCAAAG TGTACATTC GGGCAGCAGC       | 1980 | 30 |
| GGATGATCCT CCAGTATGGT AGCGCGGGTT TCTGTCTCAA AAGGAGGTAG ACGATCCCTA     | 2040 |    |
| CTGTACGGAG TGCGCCGAGA CAACCGAGAT CGTGTGGTC GTAGTGTCT GCGAAATGGA       | 2100 |    |
| ACGCCGGACG TAGTCATATT TCCTGAAGCA AAACCAGGTG CGGGCGTGAC AAACAGATCT     | 2160 |    |
| GCGTCTCCGG TCTCGCCGCT TAGATCGCTC TGTGTAGTAG TTGTAGTATA TCCACTCTCT     | 2220 |    |
| CAAAGCATCC AGGCGCCCCC TGGCTTCGGG TTCTATGTAA ACTCCTTCAT GCGCCGCTGC     | 2280 |    |
| CCTGATAACA TCCACCACCG CAGAATAAGC CACACCCAGC CAACCTACAC ATTCGTTCTG     | 2340 | 40 |
| CGAGTCACAC ACGGGAGGAG CGGGAAGAGC TGGAAGAACC ATGTTTTTTT TTTTATTCCA     | 2400 |    |
| AAAGATTATC CAAAACCTCA AAATGAAGAT CTATTAAGTG AACGCGCTCC CCTCCGGTGG     | 2460 |    |
| CGTGGTCAAA CTCTACAGCC AAAGAACAGA TAATGGCATT TGTAAGATGT TGCACAATGG     | 2520 |    |
| CTTCCAAAAG GCAAACGGCC CTCACGTCCA AGTGGACGTA AAGGCTAAAC CCTTCAGGGT     | 2580 |    |

|  |      |    |
|--|------|----|
| GAATCTCCTC TATAAACATT CCAGCACCTT CAACCATGCC CAAATAATTC TCATCTCGCC  | 2640 |    |
| ACCTTCTCAA TATATCTCTA AGCAAATCCC GAATATTAAG TCCGGCCATT GTAAAAATCT  | 2700 |    |
| GCTCCAGAGC GCCCTCCACC TTCAGCCTCA AGCAGCGAAT CATGATTGCA AAAATTTCAGG | 2760 |    |
| TTCTTCACAG ACCTGTATAA GATTCAAAAG CGGAACATTA AAAAAAATAC CGCGATCCCCG | 2820 |    |
| TAGGTCCCTT CGCAGGGCCA GCTGAACATA ATCGTGCAGG TCTGCACGGA CCAGCGCGGC  | 2880 |    |
| CACTTCCCCG CCAGGAACCT TGACAAAAGA ACCCACACTG ATTATGACAC GCATACTCGG  | 2940 | 10 |
| AGCTATGCTA ACCAGCGTAG CCCCGATGTA AGCTTTGTTG CATGGGCGGC GATATAAAAT  | 3000 |    |
| GCAAGGTGCT GCTCAAAAAA TCAGGCAAAG CCTCGCGCAA AAAAGAAAGC ACATCGTAGT  | 3060 |    |
| CATGCTCATG CAGATAAAGG CAGGTAAGCT CCGGAACCAC CACAGAAAAA GACACCATTT  | 3120 |    |
| TTCTCTCAAA CATGTCTGCG GGTTCCTGCA TAAACACAAA ATAAAATAAC AAAAAACAT   | 3180 |    |
| TTAAACATTA GAAGCCTGTC TTACAACAGG AAAAACAACC CTTATAAGCA TAAGACGGAC  | 3240 |    |
| TACGGCCATG CCGGCGTGAC CGTAAAAAAA CTGGTCACCG TGATTAAAAA GCACCACCGA  | 3300 | 20 |
| CAGTCCTCG GTCATGTCCG GAGTCATAAT GTAAGACTCG GTAAACACAT CAGGTTGATT   | 3360 |    |
| CATCGGTCAG TGCTAAAAAG CGACCGAAAT AGCCCGGGG AATACATACC CGCAGGCGTA   | 3420 |    |
| GAGACAACAT TACAGCCCCC ATAGGAGGTA TAACAAAATT AATAGGAGAG AAAAACACAT  | 3480 |    |
| AAACACCTGA AAAACCCTCC TGCCTAGGCA AAATAGCACC CTCCCGCTCC AGAACAACAT  | 3540 |    |
| ACAGCGCTTC ACAGCGGCAG CCTAACAGTC AGCCTTACCA GTAAAAAAGA AAACCTATTA  | 3600 |    |
| AAAAAACACC ACTCGACACG GCACCAGCTC AATCAGTCAC AGTGTA AAAA AGGGCCAAGT | 3660 |    |
| GCAGAGCGAG TATATATAGG ACTAAAAAAT GACGTAACGG TTAAAGTCCA CAAAAACAC   | 3720 | 30 |
| CCAGAAAACC GCACGCGAAC CTACGCCAG AAACGAAAGC CAAAAACCC ACAACTTCCT    | 3780 |    |
| CAAATCGTCA CTTCCGTTTT CCCACGTTAC GTAACCTCCC GGATCCGCGG CATTACAGT   | 3840 |    |
| TCTCCGCAAG AATTGATTGG CTCCAATTCT TGGAGTGGTG AATCCGTTAG CGAGGTGCCG  | 3900 |    |
| CCGGCTTCCA TTCAGGTGGA GGTGGCCCGG CTCCATGCAC CGCGACGCAA CGCGGGGAGG  | 3960 |    |
| CAGACAAGGT ATAGGGCGGC GCCTACAATC CATGCCAACC CGTTCCATGT GCTCGCCGAG  | 4020 |    |
| GCGGCATAAA TCGCCGTGAC GATCAGCGGT CCAGTGATCG AAGTTAGGCT GGTAAGAGCC  | 4080 | 40 |
| GCGAGCGATC CTTGAAGCTG TCCCTGATGG TCGTCATCTA CCTGCCTGGA CAGCATGGCC  | 4140 |    |
| TGCAACGCGG GCATCCCGAT GCCGCCGAA GCGAGAAGAA TCATAATGGG GAAGGCCATC   | 4200 |    |
| CAGCCTCGCG TCGCGAACGC CAGCAAGACG TAGCCCAGCG CGTCGGCCGC CATGCCCTGC  | 4260 |    |
| TTCATCCCCG TGGCCCGTTG CTCGCGTTTG CTGGCGGTGT CCCC GAAGA AATATATTTG  | 4320 |    |

|  |      |    |
|--|------|----|
| CATGTCTTTA GTTCTATGAT GACACAAACC CCGCCCAGCG TCTTGTCATT GGCGAATTTCG | 4380 |    |
| AACACGCAGA TGCAGTCGGG GCGGCGCGGT CCCAGGTCCA CTTCGCATAT TAAGGTGACG  | 4440 |    |
| CGTGTGGCCT CGAACACCGA GCGACCCTGC AGCGACCCGC TTAACAGCGT CAACAGCGTG  | 4500 |    |
| CCGCAGATCC CGGGCAATGA GATATGAAAA AGCCTGAACT CACCGCGACG TCTGTGAGAG  | 4560 |    |
| AGTTTCTGAT CGAAAAGTTC GACAGCGTCT CCGACCTGAT GCAGCTCTCG GAGGGCGAAG  | 4620 |    |
| AATCTCGTGC TTTCAGCTTC GATGTAGGAG GCGGTGGATA TGTCTGCGG GTAAATAGCT   | 4680 | 10 |
| GCGCCGATGG TTTCTACAAA GATCGTTATG TTTATCGGCA CTTTGCATCG GCCGCGCTCC  | 4740 |    |
| CGATTCCGGA AGTGCTTGAC ATTGGGGAAT TCAGCGAGAG CCTGACCTAT TGCATCTCCC  | 4800 |    |
| GCCGTGCACA GGGTGTACG TTGCAAGACC TGCCTGAAAC CGAACTGCCC GCTGTTCTGC   | 4860 |    |
| AGCCGGTCGC GGAGGCCATG GATGCGATCG CTGCGGCCGA TCTTAGCCAG ACGAGCGGGT  | 4920 |    |
| TCGGCCCATT CGGACCGCAA GGAATCGGTC AATACACTAC ATGGCGTGAT TTCATATGCG  | 4980 |    |
| CGATTGCTGA TCCCCATGTG TATCACTGGC AAAGTGTGAT GGACGACACC GTCAGTGCGT  | 5040 | 20 |
| CCGTGCGCA GGCTCTCGAT GAGCTGATGC TTTGGGCCGA GGAAGTCCGGC             | 5100 |    |
| ACCTCGTGCA CGCGGATTC GGCTCCAACA ATGTCCTGAC GGACAATGGC CGCATAACAG   | 5160 |    |
| CGGTCATTGA CTGGAGCGAG GCGATGTTG GGGATTCCCA ATACGAGGTC GCCAACATCT   | 5220 |    |
| TCTTCTGGAG GCCGTGGTTG GCTTGATGG AGCAGCAGAC GCGCTACTTC GAGCGGAGGC   | 5280 |    |
| ATCCGGAGCT TGCAGGATCG CCGCGGCTCC GGGCGTATAT GCTCCGCATT GGTCTTGACC  | 5340 |    |
| AACTCTATCA GAGCTTGGTT GACGGCAATT TCGATGATGC AGCTTGGGCG CAGGGTCGAT  | 5400 |    |
| GCGACGCAAT CGTCCGATCC GGAGCCGGGA CTGTGCGGCG TACACAAATC GCCCGCAGAA  | 5460 | 30 |
| GCGCGGCCGT CTGGACCGAT GGCTGTGTAG AAGTACTCGC CGATAGTGA AACCGACGCC   | 5520 |    |
| CCAGCACTCG TCCGAGGGCA AAGGAATAGG GGAGATGGGG GAGGCTAACT GAAACACGGA  | 5580 |    |
| AGGAGACAAT ACCGGAAGGA ACCCGCGCTA TGACGGCAAT AAAAAGACAG AATAAAACGC  | 5640 |    |
| ACGGGTGTTG GGTGTTTTGT TCATAAACGC GGGGTTCGGT CCCAGGGCTG GCACTCTGTC  | 5700 |    |
| GATACCCAC CGAGACCCCA TTGGGGCCAA TACGCCGCG TTTCTTCCTT TTCCCACCC     | 5760 |    |
| CACCCCCAA GTTCGGGTGA AGGCCAGGG CTCGCAGCCA ACGTCGGGGC GGCAGGCCCT    | 5820 | 40 |
| GCCATAGCCA CTGGCCCCGT GGGTTAGGGA CGGGGTCCCC CATGGGGAAT GGTATATGGT  | 5880 |    |
| TCGTGGGGT TATTATTTTG GCGTTGCGT GGGGTCTGGT CCACGACTGG ACTGAGCAGA    | 5940 |    |
| CAGACCCATG GTTTTTGGAT GGCCTGGGCA TGGACCGCAT GTACTGGCGC GACACGAACA  | 6000 |    |
| CCGGGCGTCT GTGGCTGCCA AACACCCCCG ACCCCCCAAA ACCACCGCGC GGATTTCTGG  | 6060 |    |

|   |      |    |
|---|------|----|
| CGCCCAGTGC CGTCGACCGG TCATGGCTGC GCCCCGACAC CCGCCAACAC CCGCTGACGC   | 6120 |    |
| GCCCTGACGG GCTTGTCTGC TCCCGGCATC CGCTTACAGA CAAGCTGTGA CCGTCTCCGG   | 6180 |    |
| GAGCTGCATG TGTCAGAGGT TTTCACCGTC ATCACCGAAA CGCGCGAGGC AGCCGGATCA   | 6240 |    |
| TAATCAGCCA TACCACATTT GTAGAGGTTT TACTTGCTTT AAAAAACCTC CCCACCTCCC   | 6300 |    |
| CCTGAACCTG AAACATAAAA TGAATGCAAT TGTTGTTGTT AACTTGTTTA TTGCAGCTTA   | 6360 |    |
| TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTTACACA AATAAAGCAT TTTTTCCTACT | 6420 | 10 |
| GCATTCTAGT TGTGGTTTGT CCAAACCTCAT CAATGTATCT TATCATGTCT GGATCCACTA  | 6480 |    |
| GTTCTAGAGC GGCCGCCACC GCGGTGGAGC TCCAGCTTTT GTTCCCTTTA GTGAGGGTTA   | 6540 |    |
| ATTTCCGAGCT TGGCGTAATC ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC  | 6600 |    |
| ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA   | 6660 |    |
| GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG   | 6720 |    |
| TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG   | 6780 | 20 |
| CGCTCTTCCG CTTCCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT GCGGCGAGCG  | 6840 |    |
| GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA   | 6900 |    |
| AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG   | 6960 |    |
| GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG   | 7020 |    |
| AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC   | 7080 |    |
| GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT TCTCCCTTCG   | 7140 |    |
| GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTCGTT   | 7200 | 30 |
| CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCGTTCAGC CCGACCGCTG CGCCTTATCC    | 7260 |    |
| GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC   | 7320 |    |
| ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG   | 7380 |    |
| TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA   | 7440 |    |
| GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC   | 7500 |    |
| GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA AAAAAGGATC TCAAGAAGAT   | 7560 | 40 |
| CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGGAACG AAAACTCAG TTAAGGGATT    | 7620 |    |
| TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC TTTTAAATTA AAAATGAAGT   | 7680 |    |
| TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC   | 7740 |    |
| AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTTCAT CCATAGTTGC CTGACTCCCC  | 7800 |    |



GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA 7860  
 CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG 7920  
 GCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC 7980  
 CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT 8040  
 ACAGGCATCG TGGTGTACAG CTCGTCGTTT GGTATGGCTT CATTTCAGCTC CGGTTCCCAA 8100  
 CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT 8160  
 CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT CACTCATGGT TATGGCAGCA 8220  
 CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC 8280  
 TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA 8340  
 ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT 8400  
 TCTTCGGGGC GAAAACTCTC AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC 8460  
 ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA 8520  
 AAAACAGGAA GGCAAAATGC CGCAAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA 8580  
 CTCATACTCT TCCTTTTTCA ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC 8640  
 GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCC 8700  
 CGAAAAGTGC 8710

10

20

- ( 2 ) 配列番号 : 5 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 3 0 塩基対  
 ( B ) 配列の型 : 核酸  
 ( C ) 鎖の数 : 一本鎖  
 ( D ) トポロジー : 直鎖状  
 ( ii ) 配列の種類 : D N A ( genomic )  
 ( iii ) ハイボセティカル : NO  
 ( iii ) アンチセンス : NO  
 ( xi ) 配列の記載 : 配列番号 : 5 :

ATGGGATCCA AGATGAAGCG CGCAAGACCG 30

- ( 2 ) 配列番号 : 6 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 3 0 塩基対  
 ( B ) 配列の型 : 核酸  
 ( C ) 鎖の数 : 一本鎖  
 ( D ) トポロジー : 直鎖状  
 ( ii ) 配列の種類 : D N A ( genomic )  
 ( iii ) ハイボセティカル : NO  
 ( iii ) アンチセンス : NO  
 ( xi ) 配列の記載 : 配列番号 : 6 :

CATAACGGG CCGCTTCTTT ATTCTTGGGC 30

- ( 2 ) 配列番号 : 7 の情報 :  
 ( i ) 配列の特性 :

30

40

50

- ( A ) 配列の長さ : 7 1 4 8 塩基対  
 ( B ) 配列の型 : 核酸  
 ( C ) 鎖の数 : 二本鎖  
 ( D ) トポロジ : 環状  
 ( ii ) 配列の種類 : D N A ( genomic )  
 ( iii ) ハイボセティカル : NO  
 ( iii ) アンチセンス : NO  
 ( xi ) 配列の記載 : 配列番号 : 7 :

|  |     |    |
|--|-----|----|
| GACGGATCGG GAGATCTCCC GATCCCCTAT GGTCGACTCT CAGTACAATC TGCTCTGATG  | 60  |    |
| CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG  | 120 | 10 |
| CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC  | 180 |    |
| TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT  | 240 |    |
| GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA  | 300 |    |
| TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC  | 360 |    |
| CCCGCCCAT T GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC | 420 |    |
| ATTGACGTCA ATGGGTGGAC TATTTACGGT AAAGTGGCCA CTTGGCAGTA CATCAAGTGT  | 480 | 20 |
| ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT  | 540 |    |
| ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA  | 600 |    |

|   |      |    |
|---|------|----|
| TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG     | 660  |    |
| ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC     | 720  |    |
| AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG     | 780  |    |
| GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA     | 840  |    |
| CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTTGGTACC     | 900  |    |
| GAGCTGGAT CCAAGATGAA GCGCGCAAGA CCGTCTGAAG ATACCTTCAA CCCCCTGTAT      | 960  | 10 |
| CCATATGACA CGGAAACCGG TCCTCCAACT GTGCCTTTTC TTA CTCTCCCTCC CTTTGTATCC | 1020 |    |
| CCCAATGGGT TTCAAGAGAG TCCCCCTGGG GTACTCTCTT TGC GCCTATC CGAACCTCTA    | 1080 |    |
| GTTACCTCCA ATGGCATGCT TGC GCTCAAA ATGGGCAACG GCCTCTCTCT GGACGAGGCC    | 1140 |    |
| GGCAACCTTA CCTCCCAAAA TGTAACCACT GTGAGCCCAC CTCTCAAAAA AACCAAGTCA     | 1200 |    |
| AACATAAACC TGGAAATATC TGCACCCCTC ACAGTTACCT CAGAAGCCCT AACTGTGGCT     | 1260 |    |
| GCCGCCGCAC CTCTAATGGT CGCGGGCAAC AACTCACCA TGCAATCACA GGCCCCGCTA      | 1320 | 20 |
| ACCGTGCACG ACTCCAACT TAGCATTGCC ACCCAAGGAC CCCTCACAGT GTCAGAAGGA      | 1380 |    |
| AAGCTAGCCC TGCAAACATC AGGCCCCCTC ACCACCACCG ATAGCAGTAC CCTTACTATC     | 1440 |    |
| ACTGCCTCAC CCCCTCTAAC TACTGCCACT GGTAGCTTGG GCATTGACTT GAAAGAGCCC     | 1500 |    |
| ATTTATACAC AAAATGGAAA ACTAGGACTA AAGTACGGGG CTCCTTTGCA TGTAACAGAC     | 1560 |    |
| GACCTAAACA CTTTGACCGT AGCAACTGGT CCAGGTGTGA CTATTAATAA TACTTCCTTG     | 1620 |    |
| CAAACTAAAG TTA CTGGAGC CTGGGGTTTT GATTCACAAG GCAATATGCA ACTTAATGTA    | 1680 |    |
| GCAGGAGGAC TAAGGATTGA TTCTCAAAAC AGACGCCTTA TACTTGATGT TAGTTATCCG     | 1740 | 30 |
| TTTGATGCTC AAAACCAACT AAATCTAAGA CTAGGACAGG GCCCTCTTTT TATAAACTCA     | 1800 |    |
| GCCCACAAC TGGATATTAA CTACAACAAA GGCCTTTACT TGTTTACAGC TTCAAACAAT      | 1860 |    |
| TCCAAAAAGC TTGAGGTTAA CCTAAGCACT GCCAAGGGGT TGATGTTTGA CGCTACAGCC     | 1920 |    |
| ATAGCCATTA ATGCAGGAGA TGGGCTTGAA TTTGGTTCAC CTAATGCACC AAACACAAAT     | 1980 |    |
| CCCCTCAAAA CAAAAATTGG CCATGGCCTA GAATTTGATT CAAACAAGGC TATGGTTCCT     | 2040 |    |
| AAACTAGGAA CTGGCCTTAG TTTTGACAGC ACAGGTGCCA TTACAGTAGG AAACAAAAAT     | 2100 | 40 |
| AATGATAAGC TAACTTTGTG GACCACACCA GCTCCATCTC CTA ACTGTAG ACTAAATGCA    | 2160 |    |
| GAGAAAGATG CTAAACTCAC TTTGGTCTTA AAAAAATGTG GCAGTCAAAT ACTTGCTACA     | 2220 |    |
| GTTTCAGTTT TGGCTGTTAA AGGCAGTTTG GCTCCAATAT CTGGAACAGT TCAAAGTGCT     | 2280 |    |
| CATCTTATTA TAAGATTTGA CGAAAAATGGA GTGCTACTAA ACAATTCCTT CCTGGACCCA    | 2340 |    |

|   |      |    |
|---|------|----|
| GAATATTGGA ACTTTAGAAA TGGAGATCTT ACTGAAGGCA CAGCCTATAC AAACGCTGTT | 2400 |    |
| GGATTTATGC CTAACCTATC AGCTTATCCA AAATCTCACG GTAAAACTGC CAAAAGTAAC | 2460 |    |
| ATTGTCAGTC AAGTTTACTT AAACGGAGAC AAAACTAAAC CTGTAACACT AACCATTACA | 2520 |    |
| CTAAACGGTA CACAGGAAAC AGGAGACACA ACTCCAAGTG CATACTCTAT GTCATTTTCA | 2580 |    |
| TGGGACTGGT CTGGCCACAA CTACATTAAT GAAATATTTG CCACATCCTC TTACACTTTT | 2640 |    |
| TCATACATTG CCCAAGAATA AAGAAGCGGC CGCTCGAGCA TGCATCTAGA GGGCCCTATT | 2700 | 10 |
| CTATAGTGTC ACCTAAATGC TAGAGCTCGC TGATCAGCCT CGACTGTGCC TTCTAGTTGC | 2760 |    |
| CAGCCATCTG TTGTTTGCCC CTCCCCCGTG CCTTCCTTGA CCCTGGAAGG TGCCACTCCC | 2820 |    |
| ACTGTCCTTT CCTAATAAAA TGAGGAAATT GCATCGCATT GTCTGAGTAG GTGTCATTCT | 2880 |    |
| ATTCTGGGGG GTGGGGTGGG GCAGGACAGC AAGGGGGAGG ATTGGGAAGA CAATAGCAGG | 2940 |    |
| CATGCTGGGG ATGCGGTGGG CTCTATGGCT TCTGAGGCGG AAAGAACCAG CTGGGGCTCT | 3000 |    |
| AGGGGGTATC CCCACGCGCC CTGTAGCGGC GCATTAAGCG CGGCGGGTGT GGTGGTTACG | 3060 | 20 |
| CGCAGCGTGA CCGCTACACT TGCCAGCGCC CTAGCGCCCG CTCCTTTCGC TTTCTTCCCT | 3120 |    |
| TCCTTTCTCG CCACGTTGCG CGGCTTTCCC CGTCAAGCTC TAAATCGGGG CATCCCTTTA | 3180 |    |
| GGGTTCGAT TTAGTGCTTT ACGGCACCTC GACCCCAAAA AACTTGATTA GGGTGATGGT  | 3240 |    |
| TCACGTAGTG GGCCATCGCC CTGATAGACG GTTTTTCGCC CTTTGACGTT GGAGTCCACG | 3300 |    |
| TTCTTTAATA GTGGACTCTT GTTCCAAACT GGAACAACAC TCAACCCTAT CTCGGTCTAT | 3360 |    |
| TCTTTTGATT TATAAGGGAT TTTGGGGATT TCGGCCTATT GGTTAaaaaa TGAGCTGATT | 3420 |    |
| TAACAAAAAT TTAACGCGAA TTAATTCTGT GGAATGTGTG TCAGTTAGGG TGTGGAAAGT | 3480 | 30 |
| CCCCAGGCTC CCCAGGCAGG CAGAAGTATG CAAAGCATGC ATCTCAATTA GTCAGCAACC | 3540 |    |
| AGGTGTGGAA AGTCCCCAGG CTCCCCAGCA GGCAGAAGTA TGCAAAGCAT GCATCTCAAT | 3600 |    |
| TAGTCAGCAA CCATAGTCCC GCCCCTAACT CCGCCCATCC CGCCCCTAAC TCCGCCAGT  | 3660 |    |
| TCCGCCATT CTCCGCCCA TGGCTGACTA ATTTTTTTTA TTTATGCAGA GGCCGAGGCC   | 3720 |    |
| GCCTCTGCCT CTGAGCTATT CCAGAAGTAG TGAGGAGGCT TTTTGGAGG CCTAGGCTTT  | 3780 |    |
| TGCAAAAAGC TCCCGGGAGC TTGTATATCC ATTTTCGGAT CTGATCAAGA GACAGGATGA | 3840 | 40 |
| GGATCGTTTC GCATGATTGA ACAAGATGGA TTGCACGCAG GTTCTCCGGC CGCTTGGGTG | 3900 |    |
| GAGAGGCTAT TCGGCTATGA CTGGGCACAA CAGACAATCG GCTGCTCTGA TGCCGCCGTG | 3960 |    |
| TTCCGGCTGT CAGCGCAGGG GCGCCCGGTT CTTTTGTCA AGACCGACCT GTCCGGTGCC  | 4020 |    |
| CTGAATGAAC TGCAGGACGA GGCAGCGCGG CTATCGTGGC TGGCCACGAC GGGCGTTCCT | 4080 |    |

|  |      |    |
|--|------|----|
| TGCGCAGCTG TGCTCGACGT TGTCACTGAA GCGGGAAGGG ACTGGCTGCT ATTGGGCGAA  | 4140 |    |
| GTGCCGGGGC AGGATCTCCT GTCATCTCAC CTTGCTCCTG CCGAGAAAAGT ATCCATCATG | 4200 |    |
| GCTGATGCAA TGCGGCGGCT GCATAOGCTT GATCCGGCTA CCTGCCCATT CGACCACCAA  | 4260 |    |
| GCGAAACATC GCATCGAGCG AGCACGTACT CGGATGGAAG CCGGTCTTGT CGATCAGGAT  | 4320 |    |
| GATCTGGACG AAGAGCATCA GGGGCTCGCG CCAGCCGAAC TGTTCGCCAG GCTCAAGGCG  | 4380 |    |
| CGCATGCCCC ACGGCGAGGA TCTCGTCGTG ACCCATGGCG ATGCCTGCTT GCCGAATATC  | 4440 | 10 |
| ATGGTGAAAA ATGGCCGCTT TTCTGGATTG ATOGACTGTG GCCGGCTGGG TGTGGCGGAC  | 4500 |    |
| CGCTATCAGG ACATAGCGTT GGCTACCCGT GATATTGCTG AAGAGCTTGG CGGCGAATGG  | 4560 |    |
| GCTGACCGCT TCCTCGTGCT TTACGGTATC GCGCTCCCG ATTCGCAGCG CATCGCCTTC   | 4620 |    |
| TATCGCCTTC TTGACGAGTT CTTCTGAGCG GGACTCTGGG GTTCGAAATG ACCGACCAAG  | 4680 |    |
| CGACGCCCCA CCTGCCATCA CGAGATTTG ATTCACCGC CGCCTTCTAT GAAAGGTTGG    | 4740 |    |
| GCTTCGGAAT CGTTTTCCGG GACGCCGGCT GGATGATCCT CCAGCGCGGG GATCTCATGC  | 4800 | 20 |
| TGGAGTTCTT CGCCCACCCC AACTTGTTTA TTGCAGCTTA TAATGGTTAC AAATAAAGCA  | 4860 |    |
| ATAGCATCAC AAATTTTACA AATAAAGCAT TTTTTTCACT GCATTCTAGT TGTGGTTTGT  | 4920 |    |
| CCAAACTCAT CAATGTATCT TATCATGTCT GTATACCGTC GACCTCTAGC TAGAGCTTGG  | 4980 |    |
| CGTAATCATG GTCATAGCTG TTTCCTGTGT GAAATTGTTA TCCGCTCACA ATTCCACACA  | 5040 |    |
| ACATACGAGC CGGAAGCATA AAGTGTAAG CCTGGGGTGC CTAATGAGTG AGCTAACTCA   | 5100 |    |
| CATTAATTGC GTTGCGCTCA CTGCCGCTT TCCAGTCGGG AAACCTGTG TGCCAGCTGC    | 5160 |    |
| ATTAATGAAT CGGCCAACGC GCGGGGAGAG GCGGTTTGG TATTGGGCGC TCTTCGCTT    | 5220 | 30 |
| CCTCGCTCAC TGA CTGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT  | 5280 |    |
| CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG AACATGTGAG  | 5340 |    |
| CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG TTTTCCATA   | 5400 |    |
| GGCTCCGCCC CCCTGACGAG CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC  | 5460 |    |
| CGACAGGACT ATAAAGATAC CAGGCGTTTC CCCCTGGAAG CTCCTCGTG CGCTCTCCTG   | 5520 |    |
| TTCCGACCCT GCGGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC  | 5580 | 40 |
| TTTCTCAATG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTGTTTCG TCCAAGCTGG   | 5640 |    |
| GCTGTGTGCA CGAACCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT AACTATCGTC   | 5700 |    |
| TTGAGTCCAA CCCGGTAAGA CACGACTTAT CGCCACTGGC AGCAGCCACT GGTAACAGGA  | 5760 |    |
| TTAGCAGAGC GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG  | 5820 |    |

|  |      |    |
|--|------|----|
| GCTACACTAG AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA  | 5880 |    |
| AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT GGTTTTTTTT  | 5940 |    |
| TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA AGAAGATCCT TTGATCTTTT  | 6000 |    |
| CTACGGGGTC TGACGCTCAG TGGAACGAAA ACTCACGTTA AGGGATTTTG GTCATGAGAT  | 6060 |    |
| TATCAAAAAG GATCTTCACC TAGATCCTTT TAAATTAATA ATGAAGTTTT AAATCAATCT  | 6120 |    |
| AAAGTATATA TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA  | 6180 | 10 |
| TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC GTGTAGATAA  | 6240 |    |
| CTACGATACG GGAGGGGCTTA CCATCTGGCC CCAGTGCTGC AATGATACCG CGAGACCCAC | 6300 |    |
| GCTCACCGGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA  | 6360 |    |
| GTGGTCCTGC AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG  | 6420 |    |
| TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG  | 6480 |    |
| TGTCAOGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA TCAAGGCGAG  | 6540 |    |
| TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG  | 6600 | 20 |
| TCAGAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC  | 6660 |    |
| TTACTGTCAT GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT  | 6720 |    |
| TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GCGTCAATA CGGGATAATA   | 6780 |    |
| CCGCGCCACA TAGCAGAACT TAAAAAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA  | 6840 |    |
| AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA  | 6900 |    |
| ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC  | 6960 | 30 |
| AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC  | 7020 |    |
| TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG  | 7080 |    |
| AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC  | 7140 |    |
| CTGACGTC   | 7148 |    |

( 2 ) 配列番号 : 8 の情報 :

( i ) 配列の特性 :

( A ) 配列の長さ : 7 4 6 9 塩基対

( B ) 配列の型 : 核酸

( C ) 鎖の数 : 二本鎖

( D ) トポロジー : 環状

( ii ) 配列の種類 : D N A ( genomic )

( iii ) ハイボセティカル : NO

( iii ) アンチセンス : NO

( xi ) 配列の記載 : 配列番号 : 8 :

|  |      |    |
|--|------|----|
| GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGACTCT CAGTACAATC TGCTCTGATG     | 60   |    |
| CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG    | 120  |    |
| CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC    | 180  |    |
| TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT    | 240  |    |
| GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA    | 300  |    |
| TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC    | 360  |    |
| CCCGCCCATTT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC   | 420  | 10 |
| ATTGACGTCA ATGGGTGGAC TATTTACGGT AAAC TGCCCA CTTGGCAGTA CATCAAGTGT   | 480  |    |
| ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT    | 540  |    |
| ATGCCCAGTA CATGACCTTA TGGGACTTTC CTA CTGTTGGCA GTACATCTAC GTATTAGTCA | 600  |    |
| TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG    | 660  |    |
| ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC    | 720  |    |
| AAAATCAACG GGA CTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG   | 780  | 20 |
| GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA    | 840  |    |
| CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTTGGTACC    | 900  |    |
| GAGCTCGGAT CTGAATTGGA GCTCGCTGTT GGGCTCGCGG TTGAGGACAA ACTCTTCGCG    | 960  |    |
| GTCTTTCCAG TACTCTTGGA TCGGAAACCC GTCGGCCTCC GAACGGTACT CCGCCACCGA    | 1020 |    |
| GGGACCTGAG CGAGTCCGCA TCGACCGGAT CGGAAAACCT CTCGAGAAAAG GCGTCTAACC   | 1080 |    |
| AGTCACAGTC GCAAGGTAGG CTGAGCACCG TGGCGGGCGG CAGCGGGTGG CGGTCGGGGT    | 1140 | 30 |
| TGTTTCTGGC GGAGGTGCTG CTGATGATGT AATTAAAGTA GGCGGTCTTG AGACGGCGGA    | 1200 |    |
| TGGTCGAGGT GAGGTGTGGC AGGCTTGAGA TCCAAGATGA AGCGCGCAAG ACGTCTGAA     | 1260 |    |
| GATACCTTCA ACCCGTGTA TCCATATGAC ACGGAAACCG GTCCTCCAAC TGTGCCTTTT     | 1320 |    |
| CTTACTCCTC CCTTTGTATC CCCC AATGGG TTTCAAGAGA GTCCCCCTGG GGTACTCTCT   | 1380 |    |
| TTGCGCCTAT CCGAACCTCT AGTTACCTCC AATGGCATGC TTGCGCTCAA AATGGGCAAC    | 1440 |    |
| GGCCTCTCTC TGGACGAGGC CGGCAACCTT ACCTCCCAAA ATGTAACCAC TGTGAGCCCA    | 1500 | 40 |

|  |      |    |
|--|------|----|
| CCTCTCAAAA AAACCAAGTC AAACATAAAC CTGGAATAT CTGCACCCCT CACAGTTACC   | 1560 |    |
| TCAGAAGCCC TAACTGTGGC TGCCGCCGCA CCTCTAATGG TCGCGGGCAA CACACTCACC  | 1620 |    |
| ATGCAATCAC AGGCCCCGCT AACCGTGCAC GACTCCAAAC TTAGCATTGC CACCCAAGGA  | 1680 |    |
| CCCCCTCACAG TGTCAGAAGG AAAGCTAGCC CTGCAAACAT CAGGCCCCCT CACCACCACC | 1740 |    |
| GATAGCAGTA CCCTTACTAT CACTGCCTCA CCCCCTCTAA CTACTGCCAC TGGTAGCTTG  | 1800 |    |
| GGCATTGACT TGAAAGAGCC CATTTATACA CAAAATGGAA AACTAGGACT AAAGTACGGG  | 1860 | 10 |
| GCTCCTTTGC ATGTAACAGA CGACCTAAAC ACTTTGACCG TAGCAACTGG TCCAGGTGTG  | 1920 |    |
| ACTATTAATA ATACTTCCTT GCAAATAAA GTTACTGGAG CCTTGGGTTT TGATTCACAA   | 1980 |    |
| GGCAATATGC AACTTAATGT AGCAGGAGGA CTAAGGATTG ATTCTCAAAA CAGACGCCTT  | 2040 |    |
| ATACTTGATG TTAGTTATCC GTTTGATGCT CAAAACCAAC TAAATCTAAG ACTAGGACAG  | 2100 |    |
| GGCCCTCTTT TTATAAACTC AGCCCACAAC TTGGATATTA ACTACAACAA AGGCCTTTAC  | 2160 |    |
| TTGTTTACAG CTTCAAACAA TTCCAAAAAG CTGAGGTTA ACCTAAGCAC TGCCAAGGGG   | 2220 | 20 |
| TTGATGTTG ACGCTACAGC CATAGCCATT AATGCAGGAG ATGGGCTTGA ATTTGGTTCA   | 2280 |    |
| CCTAATGCAC CAAACACAAA TCCCCTCAA AAAAAATTG GCCATGGCCT AGAATTTGAT    | 2340 |    |
| TCAAACAAGG CTATGGTTCC TAACTAGGA ACTGGCCTTA GTTTTGACAG CACAGGTGCC   | 2400 |    |
| ATTACAGTAG GAAACAAAAA TAATGATAAG CTAACCTTGT GGACCACACC AGCTCCATCT  | 2460 |    |
| CCTAACTGTA GACTAAATGC AGAGAAAGAT GCTAACTCA CTTTGGTCTT AACAAAATGT   | 2520 |    |
| GGCAGTCAAA TACTTGCTAC AGTTTCAGTT TTGGCTGTTA AAGGCAGTTT GGCTCCAATA  | 2580 | 30 |
| TCTGGAACAG TTCAAAGTGC TCATCTTATT ATAAGATTG ACGAAAATGG AGTGCTACTA   | 2640 |    |
| AACAATTCCT TCCTGGACCC AGAATATTGG AACTTTAGAA ATGGAGATCT TACTGAAGGC  | 2700 |    |
| ACAGCCTATA CAAACGCTGT TGGATTTATG CCTAACCTAT CAGCTTATCC AAAATCTCAC  | 2760 |    |
| GGTAAACTG CAAAAGTAA CATTGTCAGT CAAGTTTACT TAAACGGAGA CAAACTAAA     | 2820 |    |
| CCTGTAACAC TAACCATTAC ACTAAACGGT ACACAGGAAA CAGGAGACAC AACTCCAAGT  | 2880 |    |
| GCATACTCTA TGTCATTTTC ATGGGACTGG TCTGGCCACA ACTACATTAA TGAAATATTT  | 2940 |    |
| GCCACATCCT CTTACACTTT TTCATACATT GCCCAAGAAT AAAGAAGCGG CCGCTCGAGC  | 3000 | 40 |
| ATGCATCTAG AGGGCCCTAT TCTATAGTGT CACCTAAATG CTAGAGCTCG CTGATCAGCC  | 3060 |    |
| TCGACTGTGC CTTCTAGTTG CCAGCCATCT GTTGTGTTGCC CCTCCCCCGT GCCTTCCTTG | 3120 |    |
| ACCCTGGAAG GTGCCACTCC CACTGTCTTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT  | 3180 |    |
| TGTCTGAGTA GGTGTCATTC TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG  | 3240 |    |



|  |      |    |
|--|------|----|
| GATTGGGAAG ACAATAGCAG GCATGCTGGG GATGCGGTGG GCTCTATGGC TTCTGAGGCG  | 3300 |    |
| GAAAGAACCA GCTGGGGCTC TAGGGGGTAT CCCCACGCGC CCTGTAGCGG CGCATTAAGC  | 3360 |    |
| GCGGCGGGTG TGGTGGTTAC GCGCAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGCGCCC  | 3420 |    |
| GCTCCTTTTCG CTTTCTTCCC TTCCTTTCTC GCCACGTTTG CCGGCTTTCC CCGTCAAGCT | 3480 |    |
| CTAAATCGGG GCATCCCTTT AGGGTTCCGA TTTAGTGCTT TACGGCACCT CGACCCCAAA  | 3540 |    |
| AAACTTGATT AGGGTGATGG TTCACGTAGT GGGCCATCGC CCTGATAGAC GGTTTTTCGC  | 3600 | 10 |
| CCTTTGACGT TGGAGTCCAC GTTCTTTAAT AGTGGACTCT TGTTCAAAC TGAACAACA    | 3660 |    |
| CTCAACCCTA TCTCGGTCTA TTCTTTTGAT TTATAAGGGA TTTTGGGGAT TTCGGCCTAT  | 3720 |    |
| TGGTTAAAAA ATGAGCTGAT TTAACAAAAA TTAAACGCGA ATTAATTCTG TGAATGTGT   | 3780 |    |
| GTCAGTTAGG GTGTGGAAG TCCCCAGGCT CCCCAGGCAG GCAGAAGTAT GCAAAGCATG   | 3840 |    |
| CATCTCAATT AGTCAGCAAC CAGGTGTGGA AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT  | 3900 |    |
| ATGCAAAGCA TGCATCTCAA TTAGTCAGCA ACCATAGTCC CGCCCCTAAC TCCGCCCATC  | 3960 | 20 |
| CCGCCCCTAA CTCCGCCAG TTCGCCCAT TCTCCGCCCC ATGGCTGACT AATTTTTTTT    | 4020 |    |
| ATTTATGCAG AGGCCGAGGC CGCCTCTGCC TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC  | 4080 |    |
| TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTCCCGGGAG CTTGTATATC CATTTTCGGA  | 4140 |    |
| TCTGATCAAG AGACAGGATG AGGATCGTTT CGCATGATTG AACAAGATGG ATTGCACGCA  | 4200 |    |
| GGTTCTCCGG CCGCTTGGGT GGAGAGGCTA TTCGGCTATG ACTGGGCACA ACAGACAATC  | 4260 |    |
| GGCTGCTCTG ATGCGCCGT GTTCCGGCTG TCAGCGCAGG GCGGCCCGGT TCTTTTTGTC   | 4320 |    |
| AAGACCGACC TGTCCGGTGC CCTGAATGAA CTGCAGGACG AGGCAGCGCG GCTATCGTGG  | 4380 | 30 |
| CTGGCCACGA CGGGCGTTCC TTGCGCAGCT GTGCTCGACG TTGTCACTGA AGCGGGAAGG  | 4440 |    |
| GACTGGCTGC TATTGGGCGA AGTGCCGGGG CAGGATCTCC TGTCATCTCA CCTTGCTCCT  | 4500 |    |
| GCCGAGAAAG TATCCATCAT GGCTGATGCA ATGCGGCGGC TGCATACGCT TGATCCGGCT  | 4560 |    |
| ACCTGCCCAT TCGACCACCA AGCGAAACAT CGCATCGAGC GAGCACGTAC TCGGATGGAA  | 4620 |    |
| GCCGGTCTTG TCGATCAGGA TGATCTGGAC GAAGAGCATC AGGGGCTCGC GCCAGCCGAA  | 4680 |    |
| CTGTTCCGCA GGCTCAAGGC GCGCATGCCC GACGGCGAGG ATCTCGTCGT GACCCATGGC  | 4740 | 40 |
| GATGCCTGCT TGCCGAATAT CATGGTGGA AATGGCCGCT TTTCTGGATT CATCGACTGT   | 4800 |    |
| GGCCGGCTGG GTGTGGCGGA CCGCTATCAG GACATAGCGT TGGCTACCG TGATATTGCT   | 4860 |    |
| GAAGAGCTTG GCGGCGAATG GGCTGACCGC TTCCTCGTGC TTTACGGTAT CGCCGCTCCC  | 4920 |    |
| GATTCGCAGC GCATCGCCTT CTATCGCCTT CTTGACGAGT TCTTCTGAGC GGGACTCTGG  | 4980 |    |

|  |      |    |
|--|------|----|
| GGTTCGAAAT GACCGACCAA GCGACGCCCA ACCTGCCATC ACGAGATTTC GATTCCACCG  | 5040 |    |
| CCGCCCTTCTA TGAAAGGTTG GGCTTCGGAA TCGTTTTCCG GGACGCCGGC TGGATGATCC | 5100 |    |
| TCCAGCGCGG GGATCTCATG CTGGAGTTCT TCGCCCACCC CAACTTGTTT ATTGCAGCTT  | 5160 |    |
| ATAATGGTTA CAAATAAAGC AATAGCATCA CAAATTTTAC AAATAAAGCA TTTTTTTTAC  | 5220 |    |
| TGCATTCTAG TTGTGGTTTG TCCAACTCA TCAATGTATC TTATCATGTC TGTATACCGT   | 5280 |    |
| CGACCTCTAG CTAGAGCTTG GCGTAATCAT GGTTCATAGCT GTTTCCTGTG TGAAATTGTT | 5340 | 10 |
| ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAG GCCTGGGGTG  | 5400 |    |
| CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG  | 5460 |    |
| GAAACCTGTC GTGCCAGCTG CATTAATGAA TCGGCCAAGC CGCGGGGAGA GGCGGTTTGC  | 5520 |    |
| GTATTGGGCG CTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC  | 5580 |    |
| GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA TCAGGGGATA  | 5640 |    |
| ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC CAGGAACCGT AAAAAGGCCG  | 5700 | 20 |
| CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC CCCCTGACGA GCATCACAAA AATCGACGCT  | 5760 |    |
| CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCTGGAA   | 5820 |    |
| GCTCCCTCGT GCGCTCTCCT GTTCGACCC TGCCGCTTAC CGGATACCTG TCCGCTTTC    | 5880 |    |
| TCCCTTCGGG AAGCGTGGCG CTTTCTCAAT GCTCACGCTG TAGGTATCTC AGTTCGGTGT  | 5940 |    |
| AGGTCGTTTC CTCCAAGCTG GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG  | 6000 |    |
| CCTTATCCGG TAACTATCGT CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG  | 6060 |    |
| CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT  | 6120 | 30 |
| TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TCGGCTCTGC  | 6180 |    |
| TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA CAAACCACCG  | 6240 |    |
| CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC  | 6300 |    |
| AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCACGTT  | 6360 |    |
| AAGGGATTTT GGTCATGAGA TTATCAAAAA GGATCTTAC CTAGATCCTT TTAAATTAAA   | 6420 |    |
| AATGAAGTTT TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT  | 6480 | 40 |
| GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC ATAGTTGCCT  | 6540 |    |
| GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT ACCATCTGGC CCCAGTGCTG  | 6600 |    |
| CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAGCCAG  | 6660 |    |
| CCGGAAGGGC CGAGCGCAGA AGTGGTCTG CAACTTTATC CGCTCCATC CAGTCTATTA    | 6720 |    |

ATTGTTGCCG GGAAGCTAGA GTAAGTAGTT CGCCAGTTAA TAGTTTGCGC AACGTTGTTG 6780  
 CCATTGCTAC AGGCATCGTG GTGTCACGCT CGTCGTTTGG TATGGCTTCA TTCAGCTCCG 6840  
 GTTCCCAACG ATCAAGGCGA GTTACATGAT CCCCCATGTT GTGCAAAAAA GCGGTTAGCT 6900  
 CCTTCGGTCC TCCGATCGTT GTCAGAAGTA AGTTGGCCGC AGTGTATCA CTCATGGTTA 6960  
 TGGCAGCACT GCATAATTCT CTTACTGTCA TGCCATCCGT AAGATGCTTT TCTGTGACTG 7020  
 GTGAGTACTC AACCAAGTCA TTCTGAGAAT AGTGTATGCG GCGACCGAGT TGCTCTTGCC 7080  
 CGGCGTCAAT ACGGGATAAT ACCGCGCCAC ATAGCAGAAC TTAAAAAGTG CTCATCATTG 7140  
 GAAAACGTTC TTCGGGGCGA AAACCTCTCA GGATCTTACC GCTGTTGAGA TCCAGTTCGA 7200  
 TGTAACCCAC TCGTGCACCC AACTGATCTT CAGCATCTTT TACTTTCACC AGCGTTTCTG 7260  
 GGTGAGCAAA AACAGGAAGG CAAAATGCCG CAAAAAAGGG AATAAGGGCG ACACGGAAAT 7320  
 GTTGAATACT CATACTCTTC CTTTTTCAAT ATTATTGAAG CATTATCAG GGTTATTGTC 7380  
 TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAAATAGGG GTTCCGCGCA 7440  
 CATTTCCCCG AAAAGTGCCA CCTGACGTC 7469

10

20

( 2 ) 配列番号 : 9 の情報 :

( i ) 配列の特性 :

( A ) 配列の長さ : 28 塩基対

( B ) 配列の型 : 核酸

( C ) 鎖の数 : 一本鎖

( D ) トポロジー : 直鎖状

( ii ) 配列の種類 : D N A ( genomic )

( iii ) ハイボセティカル : NO

( iii ) アンチセンス : NO

30

( xi ) 配列の記載 : 配列番号 : 9 :

TGCTTAAGCG GCCGGAAGG AGAAGTCC 28

( 2 ) 配列番号 : 10 の情報 :

( i ) 配列の特性 :

( A ) 配列の長さ : 23 塩基対

( B ) 配列の型 : 核酸

( C ) 鎖の数 : 一本鎖

( D ) トポロジー : 直鎖状

( ii ) 配列の種類 : D N A ( genomic )

( iii ) ハイボセティカル : NO

( iii ) アンチセンス : NO

40

( xi ) 配列の記載 : 配列番号 : 10 :

CCGAGCTAGC GACTGAAAT GAG 23

( 2 ) 配列番号 : 11 の情報 :

( i ) 配列の特性 :

( A ) 配列の長さ : 23 塩基対

( B ) 配列の型 : 核酸

( C ) 鎖の数 : 一本鎖

( D ) トポロジー : 直鎖状

( ii ) 配列の種類 : D N A ( genomic )

50

( iii ) ハイボセティカル : NO  
( iii ) アンチセンス : NO  
( xi ) 配列の記載 : 配列番号 : 1 1 :  
CCTCTCGAGA GACAGCAAGA CAC 23  
( 2 ) 配列番号 : 1 2 の情報 :  
( i ) 配列の特性 :  
( A ) 配列の長さ : 1 1 1 5 2 塩基対  
( B ) 配列の型 : 核酸  
( C ) 鎖の数 : 二本鎖  
( D ) トポロジー : 環状 10  
( ii ) 配列の種類 : D N A ( genomic )  
( iii ) ハイボセティカル : NO  
( iii ) アンチセンス : NO  
( xi ) 配列の記載 : 配列番号 : 1 2 :  
AAGCTTGGGC AGAAATGGTT GAACTCCCGA GAGTGTCCCTA CACCTAGGGG AGAAGCAGCC 60  
AAGGGGTTGT TTCCACCAA GGACGACCCG TCTGCCACA AACGGATGAG CCCATCAGAC 120

|  |      |    |
|--|------|----|
| AAAGACATAT TCATTCTCTG CTGCAAACTT GGCATAGCTC TGCTTTGCCT GGGGCTATTG  | 180  |    |
| GGGGAAGTTG CGGTTCTGTC TCGCAGGGCT CTCACCCTTG ACTCTTTTAA TAGCTCTTCT  | 240  |    |
| GTGCAAGATT ACAATCTAAA CAATTCGGAG AACTCGACCT TCCTCCTGAG GCAAGGACCA  | 300  |    |
| CAGCCAACTT CCTCTTACAA GCCGCATCGA TTTTGTCTT CAGAAATAGA AATAAGAATG   | 360  |    |
| CTTGCTAAAA ATTATATTTT TACCAATAAG ACCAATCCAA TAGGTAGATT ATTAGTTACT  | 420  |    |
| ATGTTAAGAA ATGAATCATT ATCTTTTAGT ACTATTTTAA CTCAAATTCA GAAGTTAGAA  | 480  | 10 |
| ATGGGAATAG AAAATAGAAA GAGACGCTCA ACCTCAATTG AAGAACAGGT GCAAGGACTA  | 540  |    |
| TTGACCACAG GCCTAGAAGT AAAAAAGGA AAAAAAGAGT TTTTGTCAA AATAGGAGAC    | 600  |    |
| AGGTGGTGGC AACCAGGGAC TTATAGGGGA CCTTACATCT ACAGACCAAC AGATGCCCCC  | 660  |    |
| TTACCATATA CAGGAAGATA TGA CTTAAAT TGGGATAGGT GGGTTACAGT CAATGGCTAT | 720  |    |
| AAAGTGTTAT ATAGATCCCT CCCTTTTCGT GAAAGACTCG CCAGAGCTAG ACCTCCTTGG  | 780  |    |
| TGTATGTTGT CTCAAGAAGA AAAAGACGAC ATGAAACAAC AGGTACATGA TTATATTTAT  | 840  | 20 |
| CTAGGAACAG GAATGCACTT TTGGGGAAAG ATTTTCCATA CCAAGGAGGG GACAGTGGCT  | 900  |    |
| GGACTAATAG AACATTATTC TGCAAAAAC CATGGCATGA GTTATTATGA ATAGCCTTTA   | 960  |    |
| TTGGCCCAAC CTTGCGGTTT CCAGGGCTTA AGTAAGTTT TGGTTACAAA CTGTTCTTAA   | 1020 |    |
| AACGAGGATG TGAGACAAGT GGTTCCTGA CTTGGTTTGG TATCAAAGGT TCTGATCTGA   | 1080 |    |
| GCTCTGAGTG TTCTATTTTC CTATGTTCTT TTGGAATTTA TCCAAATCTT ATGTAAATGC  | 1140 |    |
| TTATGTAAAC CAAGATATAA AAGAGTGCTG ATTTTTTGAG TAACTTGCA ACAGTCCTAA   | 1200 |    |
| CATTCACCTC TTGTGTGTTT GTGTCTGTTT GCCATCCCGT CTCGCTCGT CACTTATCCT   | 1260 | 30 |
| TCACTTTCCA GAGGGTCCCC CCGCAGACCC CGGCGACCCT CAGGTCCGCC GACTGCGGCA  | 1320 |    |
| GCTGGCGCCC GAACAGGGAC CCTCGGATAA GTGACCCTTG TCTCTATTTT TACTATTTGG  | 1380 |    |
| TGTTTGTCTT GTATTGTCTC TTTCTGTCT GGCTATCATC ACAAGAGCGG AACGGACTCA   | 1440 |    |
| CCATAGGGAC CAAGCTAGCG ACTGAAAATG AGACATATTA TCTGCCACGG AGGTGTTATT  | 1500 |    |
| ACCGAAGAAA TGGCCGCCAG TCTTTTGGAC CAGCTGATCG AAGAGGTA CTGGTATAAT    | 1560 |    |
| CTTCCACCTC CTAGCCATTT TGAACCACCT ACCCTTCACG AACTGTATGA TTTAGACGTG  | 1620 | 40 |
| ACGGCCCCCG AAGATCCCAA CGAGGAGGCG GTTTCGAGA TTTTCCCGA CTCTGTAATG    | 1680 |    |
| TTGGCGGTGC AGGAAGGGAT TGA CTTACTC ACTTTTCCGC CGGCGCCCG TTCTCCGGAG  | 1740 |    |
| CGCCTCACC TTTCCCGGCA GCGGAGCAG CCGGAGCAGA GAGCCTTGGG TCCGGTTTCT    | 1800 |    |
| ATGCCAAACC TTGTACGGA GGTGATCGAT CTTACCTGCC ACGAGGCTGG CTTTCCACCC   | 1860 |    |

|  |      |    |
|--|------|----|
| AGTGACGACG AGGATGAAGA GGGTGAGGAG TTTGTGTTAG ATTATGTGGA GCACCCCGGG  | 1920 |    |
| CACGGTTGCA GGTCTTGTC A TTATCACCGG AGGAATACGG GGGACCCAGA TATTATGTGT | 1980 |    |
| TCGCTTTGCT ATATGAGGAC CTGTGGCATG TTTGTCTACA GTAAGTGAAA ATTATGGGCA  | 2040 |    |
| GTGGGTGATA GAGTGGTGGG TTTGGTGTGG TAATTTTTTT TTTAATTTTT ACAGTTTTGT  | 2100 |    |
| GGTTTAAAGA ATTTTGTATT GTGATTTTTT TAAAAGGTCC TGTGTCTGAA CCTGAGCCTG  | 2160 |    |
| AGCCCGAGCC AGAACCGGAG CCTGCAAGAC CTACCCGCG TCCTAAAATG GCGCCTGCTA   | 2220 | 10 |
| TCCTGAGACG CCCGACATCA CCTGTGTCTA GAGAATGCAA TAGTAGTACG GATAGCTGTG  | 2280 |    |
| ACTCCGGTCC TTCTAACACA CCTCCTGAGA TACACCCGGT GGTCCCGCTG TGCCCCATTA  | 2340 |    |
| AACCAGTTGC CGTGAGAGTT GGTGGGCGTC GCCAGGCTGT GGAATGTATC GAGGACTTGC  | 2400 |    |
| TTAACGAGCC TGGGCAACCT TTGGA CTGTAACG CCCCAGGCCA TAAGGTGTAA         | 2460 |    |
| ACCTGTGATT GCGTGTGTGG TTAACGCCTT TGTTTGCTGA ATGAGTTGAT GTAAGTTTAA  | 2520 |    |
| TAAAGGGTGA GATAATGTTT AACTTGCATG GCGTGTTAAA TGGGGCGGGG CTTAAAGGGT  | 2580 | 20 |
| ATATAATGCG CCGTGGGCTA ATCTTGGTTA CATCTGACCT CATGGAGGCT TGGGAGTGTT  | 2640 |    |
| TGGAAGATTT TTCTGCTGTG CGTAACTTGC TGGAACAGAG CTCTAACAGT ACCTCTTGGT  | 2700 |    |
| TTTGGAGGTT TCTGTGGGGC TCATCCCAGG CAAAGTTAGT CTGCAGAATT AAGGAGGATT  | 2760 |    |
| ACAAGTGGGA ATTTGAAGAG CTTTGTAAAT CCTGTGGTGA GCTGTTTGAT TCTTTGAATC  | 2820 |    |
| TGGGTACCA GCGCTTTTC CAAGAGAAGG TCATCAAGAC TTTGGATTTT TCCACACCGG    | 2880 |    |
| GGCGCGCTGC GGCTGCTGTT GCTTTTTTGA GTTTTATAAA GGATAAATGG AGCGAAGAAA  | 2940 |    |
| CCCATCTGAG CGGGGGGTAC CTGCTGGATT TTCTGGCCAT GCATCTGTGG AGAGCGGTTG  | 3000 | 30 |
| TGAGACACAA GAATCGCCTG CTA CTGTTGT CTTCCGTCCG CCCGGCGATA ATACCGACGG | 3060 |    |
| AGGAGCAGCA GCAGCAGCAG GAGGAAGCCA GCGGCGGGC GCAGGAGCAG AGCCCATGGA   | 3120 |    |
| ACCCGAGAGC CGGCCTGGAC CCTCGGGAAT GAATGTTGTA CAGGTGGCTG AACTGTATCC  | 3180 |    |
| AGAACTGAGA CGCATTTTGA CAATTACAGA GGATGGGCAG GGGCTAAAGG GGGTAAAGAG  | 3240 |    |
| GGAGCGGGG GCTTGTGAGG CTACAGAGGA GGCTAGGAAT CTAGCTTTTA GCTTAATGAC   | 3300 |    |
| CAGACACCGT CCTGAGTGTA TTACTTTTCA ACAGATCAAG GATAATTGCG CTAATGAGCT  | 3360 | 40 |
| TGATCTGCTG GCGCAGAAGT ATTCCATAGA GCAGCTGACC ACTTACTGGC TGCAGCCAGG  | 3420 |    |
| GGATGATTTT GAGGAGGCTA TTAGGGTATA TGCAAAGGTG GCACTTAGGC CAGATTGCAA  | 3480 |    |
| GTACAAGATC AGCAA ACTTG TAAATATCAG GAATTGTTGC TACATTTCTG GGAACGGGGC | 3540 |    |
| CGAGGTGGAG ATAGATACGG AGGATAGGGT GGCCTTTAGA TGTAGCATGA TAAATATGTG  | 3600 |    |

|   |      |    |
|---|------|----|
| GCCGGGGGTG CTGGCATGG ACGGGGTGGT TATTATGAAT GTAAGGTTTA CTGGCCCCAA  | 3660 |    |
| TTTtagCGGT ACGGTTTTCC TGGCCAATAC CAACCTTATC CTACACGGTG TAAGCTTCTA | 3720 |    |
| TGGGTTTAAC AATACCTGTG TGGAAGCCTG GACCGATGTA AGGGTTCGGG GCTGTGCCTT | 3780 |    |
| TTACTGCTGC TGGAAGGGGG TGGTGTGTG CCCCCAAAGC AGGGCTTCAA TTAAGAAATG  | 3840 |    |
| CCTCTTTGAA AGGTGTACCT TGGGTATCCT GTCTGAGGGT AACTCCAGGG TGCGCCACAA | 3900 |    |
| TGTGGCCTCC GACTGTGGTT GCTTCATGCT AGTGAAAAGC GTGGCTGTGA TTAAGCATAA | 3960 | 10 |
| CATGGTATGT GGCAACTGCG AGGACAGGGC CTCTCAGATG CTGACCTGCT CGGACGGCAA | 4020 |    |
| CTGTCACCTG CTGAAGACCA TTCACGTAGC CAGCCACTCT CGCAAGGCCT GGCCAGTGTT | 4080 |    |
| TGAGCATAAC ATACTGACCC GCTGTTCCCT GCATTTGGGT AACAGGAGGG GGGTGTTCCT | 4140 |    |
| ACCTTACCAA TGCAATTTGA GTCACACTAA GATATTGCTT GAGCCCGAGA GCATGTCCAA | 4200 |    |
| GGTGAACCTG AACGGGGTGT TTGACATGAC CATGAAGATC TGGAAGGTGC TGAGGTACGA | 4260 |    |
| TGAGACCCGC ACCAGGTGCA GACCCTGCGA GTGTGGCGGT AAACATATTA GGAACCAGCC | 4320 | 20 |
| TGTGATGCTG GATGTGACCG AGGAGCTGAG GCCCGATCAC TTGGTGCTGG CCTGCACCCG | 4380 |    |
| CGCTGAGTTT GGCTCTAGCG ATGAAGATAC AGATTGAGGT ACTGAAATGT GTGGGCGTGG | 4440 |    |
| CTTAAGGGTG GGAAAGAATA TATAAGGTGG GGGTCTTATG TAGTTTTGTA TCTGTTTTGC | 4500 |    |
| AGCAGCCGCC GCCGCCATGA GCACCAACTC GTTTGATGGA AGCATTGTGA GCTCATATTT | 4560 |    |
| GACAACGCGC ATGCCCCCAT GGGCCGGGGT GCGTCAGAAT GTGATGGGCT CCAGCATTGA | 4620 |    |
| TGGTCGCCCC GTCCTGCCCC CAAACTCTAC TACCTTGACC TACGAGACCG TGTCTGGAAC | 4680 |    |
| GCCGTTGGAG ACTGCAGCCT CCGCCGCCGC TTCAGCCGCT GCAGCCACCG CCCGCGGGAT | 4740 | 30 |
| TGTGACTGAC TTTGCTTTCC TGAGCCCGCT TGCAAGCAGT GCAGCTTCCC GTTCATCCGC | 4800 |    |
| CCGCGATGAC AAGTTGACGG CTCTTTTGGC ACAATTGGAT TCTTTGACCC GGGAACCTAA | 4860 |    |
| TGTCGTTTCT CAGCAGCTGT TGGATCTGCG CCAGCAGGTT TCTGCCCTGA AGGCTTCCTC | 4920 |    |
| CCCTCCCAAT GCGGTTTAAA ACATAAATAA AAAACCAGAC TCTGTTTGGA TTTGGATCAA | 4980 |    |
| GCAAGTGTCT TGCTGTCTCT CGAGGGATCT TTGTGAAGGA ACCTTACTTC TGTGGTGTGA | 5040 |    |
| CATAATTGGA CAAACTACCT ACAGAGATTT AAAGCTCTAA GGTAATATA AAATTTTTAA  | 5100 | 40 |
| GTGTATAATG TGTAAACTA CTGATTCTAA TTGTTTGTGT ATTTTAGATT CCAACCTATG  | 5160 |    |
| GAAGTATGA ATGGGAGCAG TGGTGAATG CCTTTAATGA GGAAAACCTG TTTTGCTCAG   | 5220 |    |
| AAGAAATGCC ATCTAGTGAT GATGAGGCTA CTGCTGACTC TCAACATTCT ACTCCTCCAA | 5280 |    |
| AAAAGAAGAG AAAGGTAGAA GACCCCAAGG ACTTTCCTTC AGAATTGCTA AGTTTTTTGA | 5340 |    |

|            |            |            |            |            |            |      |    |
|------------|------------|------------|------------|------------|------------|------|----|
| GTCATGCTGT | GTTTAGTAAT | AGAACTCTTG | CTTGCTTTGC | TATTTACACC | ACAAAGGAAA | 5400 |    |
| AAGCTGCACT | GCTATACAAG | AAAATTATGG | AAAAATATTC | TGTAACCTTT | ATAAGTAGGC | 5460 |    |
| ATAACAGTTA | TAATCATAAC | ATACTGTTTT | TTCTTACTCC | ACACAGGCAT | AGAGTGTCTG | 5520 |    |
| CTATTAATAA | CTATGCTCAA | AAATTGTGTA | CCTTTAGCTT | TTTAATTTGT | AAAGGGGTTA | 5580 |    |
| ATAAGGAATA | TTTGATGTAT | AGTGCCTTGA | CTAGAGATCA | TAATCAGCCA | TACCACATTT | 5640 |    |
| GTAGAGGTTT | TACTTGCTTT | AAAAAACCTC | CCACACCTCC | CCCTGAACCT | GAAACATAAA | 5700 | 10 |
| ATGAATGCAA | TTGTTGTTGT | TAACTGTGTT | ATTGCAGCTT | ATAATGGTTA | CAAATAAAGC | 5760 |    |
| AATAGCATCA | CAAATTTTAC | AAATAAAGCA | TTTTTTTAC  | TGCATTCTAG | TTGTGGTTTG | 5820 |    |
| TCCAAACTCA | TCAATGTATC | TTATCATGTC | TGGATCCGGC | TGTGGAATGT | GTGTCAGTTA | 5880 |    |
| GGGTGTGGAA | AGTCCCCAGG | CTCCCCAGCA | GGCAGAAGTA | TGCAAAGCAT | GCATCTCAAT | 5940 |    |
| TAGTCAGCAA | CCAGGTGTGG | AAAGTCCCCA | GGCTCCCCAG | CAGGCAGAAG | TATGCAAAGC | 6000 |    |
| ATGCATCTCA | ATTAGTCAGC | AACCATAGTC | CCGCCCCCTA | CTCGCCCAT  | CCCGCCCCTA | 6060 | 20 |
| ACTCCGCCCC | GTTCCGCCCC | TTCTCCGCCC | CATGGCTGAC | TAATTTTTTT | TATTTATGCA | 6120 |    |
| GAGGCCGAGG | CCGCCTCGGC | CTCTGAGCTA | TTCCAGAAGT | AGTGAGGAGG | CTTTTTTGA  | 6180 |    |
| GGCCTAGGCT | TTTGCAAAAA | GCTTGACAC  | AAGACAGGCT | TGCGAGATAT | GTTTGAGAAT | 6240 |    |
| ACCACTTTAT | CCCGCGTCAG | GGAGAGGCAG | TGCGTAAAAA | GACGCGGACT | CATGTGAAAT | 6300 |    |
| ACTGGTTTTT | AGTGCGCCAG | ATCTCTATAA | TCTCGCGCAA | CCTATTTTCC | CCTCGAACAC | 6360 |    |
| TTTTTAAGCC | GTAGATAAAC | AGGCTGGGAC | ACTTCACATG | AGCGAAAAAT | ACATCGTCAC | 6420 |    |
| CTGGGACATG | TTGCAGATCC | ATGCACGTAA | ACTCGCAAGC | CGACTGATGC | CTTCTGAACA | 6480 | 30 |
| ATGGAAGGC  | ATTATTGCCG | TAAGCCGTGG | CGGTCTGGTA | CCGGGTGCGT | TACTGGCGCG | 6540 |    |
| TGAACTGGGT | ATTCGTCATG | TCGATACCGT | TTGTATTTCC | AGCTACGATC | ACGACAACCA | 6600 |    |
| GCGCGAGCTT | AAAGTGCTGA | AACGCGCAGA | AGGCGATGGC | GAAGGCTTCA | TCGTTATTGA | 6660 |    |
| TGACCTGGTG | GATACCGGTG | GTACTGCGGT | TGCGATTTCG | GAAATGTATC | CAAAAGCGCA | 6720 |    |
| CTTTGTCACC | ATCTTCGCAA | AACCGGCTGG | TCGTCCGCTG | GTTGATGACT | ATGTTGTTGA | 6780 |    |
| TATCCGCAA  | GATACCTGGA | TTGAACAGCC | GTGGGATATG | GGCGTCGTAT | TCGTCCCGCC | 6840 | 40 |
| AATCTCCGGT | CGCTAATCTT | TTCAACGCCT | GGCACTGCCG | GGCGTTGTTC | TTTTAACTT  | 6900 |    |
| CAGGCGGGTT | ACAATAGTTT | CCAGTAAGTA | TTCTGGAGGC | TGCATCCATG | ACACAGGCAA | 6960 |    |
| ACCTGAGCGA | AACCCTGTTC | AAACCCCGCT | TAAACATCC  | TGAAACCTCG | ACGCTAGTCC | 7020 |    |
| GCCGCTTTAA | TCACGGCGCA | CAACCGCCTG | TGCAGTCGGC | CCTTGATGGT | AAAACCATCC | 7080 |    |



|  |      |    |
|--|------|----|
| CTCACTGGTA TCGCATGATT AACCGTCTGA TGTGGATCTG GCGCGGCATT GACCCACGCG  | 7140 |    |
| AAATCCTCGA CGTCCAGGCA CGTATTGTGA TGAGCGATGC CGAACGTACC GACGATGATT  | 7200 |    |
| TATACGATAC GGTGATTGGC TACCGTGGCG GCAACTGGAT TTATGAGTGG GCCCCGGATC  | 7260 |    |
| TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG ACAAACCTACC TACAGAGATT | 7320 |    |
| TAAAGCTCTA AGGTAAATAT AAAATTTTTA AGTGTATAAT GTGTAAACT ACTGATTCTA   | 7380 |    |
| ATTGTTTGTG TATTTTAGAT TCCAACCTAT GGAAGTATG AATGGGAGCA GTGGTGGAAT   | 7440 | 10 |
| GCCTTTAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC CATCTAGTGA TGATGAGGCT  | 7500 |    |
| ACTGCTGACT CTCAACATTC TACTCCTCCA AAAAAGAAGA GAAAGGTAGA AGACCCCAAG  | 7560 |    |
| GACTTTCCTT CAGAATTGCT AAGTTTTTTG AGTCATGCTG TGTTTAGTAA TAGAACTCTT  | 7620 |    |
| GCTTGCTTTG CTATTTACAC CACAAAGGAA AAAGCTGCAC TGCTATACAA GAAAATTATG  | 7680 |    |
| GAAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT ATAATCATAA CATACTGTTT  | 7740 |    |
| TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA ACTATGCTCA AAAATTGTGT  | 7800 | 20 |
| ACCTTTAGCT TTTTAATTTG TAAAGGGGTT AATAAGGAAT ATTTGATGTA TAGTGCCTTG  | 7860 |    |
| ACTAGAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT TTAAGTCTT TAAAAACCT    | 7920 |    |
| CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA ATTGTTGTTG TTAAGTCTT   | 7980 |    |
| TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA CAAATAAAGC  | 8040 |    |
| ATTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAACTC ATCAATGTAT CTTATCATGT   | 8100 |    |
| CTGGATCCCC AGGAAGCTCC TCTGTGTCCT CATAAACCTT AACCTCCTCT ACTTGAGAGG  | 8160 |    |
| ACATTCCAAT CATAGGCTGC CCATCCACCC TCTGTGTCCT CCTGTAAATT AGGTCACCTA  | 8220 | 30 |
| ACAAAAAGGA AATTGGGTAG GGGTTTTTCA CAGACCGCTT TCTAAGGGTA ATTTTAAAT   | 8280 |    |
| ATCTGGGAAG TCCCTTCCAC TGCTGTGTTT CAGAAAGTGT GGTAACAGC CCACAAATGT   | 8340 |    |
| CAACAGCAGA AACATACAAG CTGTCAGCTT TGCACAAGGG CCCAACACCC TGCTCATCAA  | 8400 |    |
| GAAGCACTGT GGTGCTGTG TTAGTAATGT GCAAAACAGG AGGCACATTT TCCCCACCTG   | 8460 |    |
| TGTAGGTTCC AAAATATCTA GTGTTTTTCA TTTTACTTGG ATCAGGAACC CAGCACTCCA  | 8520 |    |
| CTGGATAAGC ATTATCCTTA TCCAAAACAG CCTTGTGGTC AGTGTTTCATC TGCTGACTGT | 8580 | 40 |
| CAACTGTAGC ATTTTTTGGG GTTACAGTTT GAGCAGGATA TTTGGTCCTG TAGTTTGCTA  | 8640 |    |
| ACACACCCTG CAGCTCCAAA GGTTCACAC CAACAGCAAA AAAATGAAAA TTTGACCCTT   | 8700 |    |
| GAATGGGTTT TCCAGCACCA TTTTCATGAG TTTTTTGTGT CCTGAATGC AAGTTTAAAC   | 8760 |    |
| TAGCAGTTAC CCCAATAACC TCAGTTTTAA CAGTAACAGC TTCCACATC AAAATATTTT   | 8820 |    |

|            |            |            |             |            |            |       |
|------------|------------|------------|-------------|------------|------------|-------|
| CACAGGTAA  | GTCCTCATTT | AAATTAGGCA | AAGGAATTCT  | TGAAGACGAA | AGGGCCTCGT | 8880  |
| GATACGCCTA | TTTTTATAGG | TTAATGTCAT | GATAATAATG  | GTTTCTTAGA | CGTCAGGTGG | 8940  |
| CACTTTTCGG | GGAAATGTGC | GCGGAACCCC | TATTTGTTTA  | TTTTTCTAAA | TACATTCAAA | 9000  |
| TATGTATCCG | CTCATGAGAC | AATAACCCTG | ATAAATGCTT  | CAATAATATT | GAAAAAGGAA | 9060  |
| GAGTATGAGT | ATTCAACATT | TCCGTGTCGC | CCTTATTCCC  | TTTTTTGCGG | CATTTTGCCT | 9120  |
| TCCTGTTTTT | GCTCACCCAG | AAACGCTGGT | GAAAGTAAAA  | GATGCTGAAG | ATCAGTTGGG | 9180  |
| TGCACGAGTG | GGTTACATCG | AACTGGATCT | CAACAGCGGT  | AAGATCCTTG | AGAGTTTTCG | 9240  |
| CCCCGAAGAA | CGTTTTCCAA | TGATGAGCAC | TTTTAAAGTT  | CTGCTATGTG | GCGCGGTATT | 9300  |
| ATCCCGTGTT | GACGCCGGGC | AAGAGCAACT | CGGTCGCCGC  | ATACACTATT | CTCAGAATGA | 9360  |
| CTTGTTGAG  | TACTCACCAG | TCACAGAAAA | GCATCTTACG  | GATGGCATGA | CAGTAAGAGA | 9420  |
| ATTATGCAGT | GCTGCCATAA | CCATGAGTGA | TAACACTGCG  | GCCAACTTAC | TTCTGACAAC | 9480  |
| GATCGGAGGA | CCGAAGGAGC | TAACCGCTTT | TTTGACACAAC | ATGGGGGATC | ATGTAACCTG | 9540  |
| CCTTGATCGT | TGGGAACCGG | AGCTGAATGA | AGCCATACCA  | AACGACGAGC | GTGACACCAC | 9600  |
| GATGCCTGCA | GCAATGGCAA | CAACGTTGCG | CAAACATTA   | ACTGGCGAAC | TACTTACTCT | 9660  |
| AGCTTCCCGG | CAACAATTAA | TAGACTGGAT | GGAGGCGGAT  | AAAGTTGCAG | GACCACTTCT | 9720  |
| GCGCTCGGCC | CTTCGGGCTG | GCTGGTTTAT | TGCTGATAAA  | TCTGGAGCCG | GTGAGCGTGG | 9780  |
| GTCTCGCGGT | ATCATTGCAG | CACTGGGGCC | AGATGGTAAG  | CCCTCCCGTA | TCGTAGTTAT | 9840  |
| CTACACGACG | GGGAGTCAGG | CAACTATGGA | TGAACGAAAT  | AGACAGATCG | CTGAGATAGG | 9900  |
| TGCCTCACTG | ATTAAGCATT | GGTAACTGTC | AGACCAAGTT  | TACTCATATA | TACTTTAGAT | 9960  |
| TGATTTAAAA | CTTCATTTTT | AATTTAAAAG | GATCTAGGTG  | AAGATCCTTT | TTGATAATCT | 10020 |
| CATGACCAAA | ATCCCTTAAC | GTGAGTTTTT | GTTCCACTGA  | GCGTCAGACC | CCGTAGAAAA | 10080 |
| GATCAAAGGA | TCTTCTTGAG | ATCCTTTTTT | TCTGCGCGTA  | ATCTGCTGCT | TGCAAACAAA | 10140 |
| AAAACCACCG | CTACCAGCGG | TGGTTTGTTT | GCCGGATCAA  | GAGCTACCAA | CTCTTTTTCC | 10200 |
| GAAGGTAAct | GGCTTCAGCA | GAGCGCAGAT | ACCAAATACT  | GTCTTTCTAG | TGTAGCCGTA | 10260 |
| GTTAGGCCAC | CACTTCAAGA | ACTCTGTAGC | ACCGCCTACA  | TACCTCGCTC | TGCTAATCCT | 10320 |
| GTTACCAGTG | GCTGCTGCCA | GTGGCGATAA | GTCGTGTCTT  | ACCGGGTTGG | ACTCAAGACG | 10380 |
| ATAGTTACCG | GATAAGGCGC | AGCGGTCGGG | CTGAACGGGG  | GGTTCGTGCA | CACAGCCCAG | 10440 |
| CTTGGAGCGA | ACGACCTACA | CCGAACTGAG | ATACCTACAG  | CGTGAGCTAT | GAGAAAGCGC | 10500 |
| CACGCTTCCC | GAAGGGAGAA | AGGCGGACAG | GTATCCGGTA  | AGCGGCAGGG | TCGGAACAGG | 10560 |

10

20

30

40

AGAGCGCACG AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGGTT 10620  
 TCGCCACCTC TGA CTGAGC GTCGATTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG 10680  
 GAAAAACGCC AGCAACGCGG CCTTTTACG GTTCCTGGCC TTTTGCTGGC CTTTGCTCA 10740  
 CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG CCTTTGAGTG 10800  
 AGCTGATACC GCTCGCCGCA GCCGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC 10860  
 GGAAGAGCGC CTGATGCGGT ATTTTCTCCT TACGCATCTG TGCGGTATTT CACACCGCAT 10920  
 ATGGTGCACT CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGT ATACACTCCG 10980  
 CTATCGCTAC GTGACTGGGT CATGGCTGCG CCCCACACC CGCCAACACC CGCTGACGCG 11040  
 CCCTGACGGG CTTGTCTGCT CCGGCATCC GCTTACAGAC AAGCTGTGAC CGTCTCCGGG 11100  
 AGCTGCATGT GTCAGAGGTT TTCACCGTCA TCACCGAAAC GCGCGAGGCA GC 11152

10

- ( 2 ) 配列番号 : 1 3 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 1 9 塩基対  
 ( B ) 配列の型 : 核酸  
 ( C ) 鎖の数 : 一本鎖  
 ( D ) トポロジー : 直鎖状  
 ( ii ) 配列の種類 : D N A ( genomic )  
 ( iii ) ハイボセティカル : NO  
 ( iii ) アンチセンス : NO  
 ( xi ) 配列の記載 : 配列番号 : 1 3 :

GACGGATCGG GAGATCTCC

19

- ( 2 ) 配列番号 : 1 4 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 2 2 塩基対  
 ( B ) 配列の型 : 核酸  
 ( C ) 鎖の数 : 一本鎖  
 ( D ) トポロジー : 直鎖状  
 ( ii ) 配列の種類 : D N A ( genomic )  
 ( iii ) ハイボセティカル : NO  
 ( iii ) アンチセンス : NO  
 ( xi ) 配列の記載 : 配列番号 : 1 4 :

CCGCCTCAGA AGCCATAGAG CC

22

- ( 2 ) 配列番号 : 1 5 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 1 4 4 5 5 塩基対  
 ( B ) 配列の型 : 核酸  
 ( C ) 鎖の数 : 二本鎖  
 ( D ) トポロジー : 環状  
 ( ii ) 配列の種類 : D N A ( genomic )  
 ( iii ) ハイボセティカル : NO  
 ( iii ) アンチセンス : NO  
 ( xi ) 配列の記載 : 配列番号 : 1 5 :

40

|  |     |    |
|--|-----|----|
| AAGCTTGGGC AGAAATGGTT GAACTCCCGA GAGTGTCTTA CACCTAGGGG AGAAGCAGCC  | 60  |    |
| AAGGGGTGTG TTCCCACCAA GGACGACCCG TCTGCGCACA AACGGATGAG CCCATCAGAC  | 120 |    |
| AAAGACATAT TCATTCTCTG CTGCAAACTT GGCATAGCTC TGCTTTGCCT GGGGCTATTG  | 180 |    |
| GGGGAAGTTG CGGTTCTGTC TCGCAGGGCT CTCACCCTTG ACTCTTTTAA TAGCTCTTCT  | 240 |    |
| GTGCAAGATT ACAATCTAAA CAATTCGGAG AACTCGACCT TCCTCCTGAG GCAAGGACCA  | 300 |    |
| CAGCCAACTT CCTCTTACAA GCCGCATCGA TTTTGTCCCT CAGAAATAGA AATAAGAATG  | 360 |    |
| CTTGCTAAAA ATTATATTTT TACCAATAAG ACCAATCCAA TAGGTAGATT ATTAGTTACT  | 420 | 10 |
| ATGTTAAGAA ATGAATCATT ATCTTTTAGT ACTATTTTAA CTCAAATTCA GAAGTTAGAA  | 480 |    |
| ATGGGAATAG AAAATAGAAA GAGACGCTCA ACCTCAATTG AAGAACAGGT GCAAGGACTA  | 540 |    |
| TTGACCACAG GCCTAGAAGT AAAAAAGGGA AAAAAGAGTG TTTTGTCAA AATAGGAGAC   | 600 |    |
| AGGTGGTGGC AACCAGGGAC TTATAGGGGA CCTTACATCT ACAGACCAAC AGATGCCCCC  | 660 |    |
| TTACCATATA CAGGAAGATA TGACTTAAAT TGGGATAGGT GGGTTACAGT CAATGGCTAT  | 720 |    |
| AAAGTGTTAT ATAGATCCCT CCCTTTTCGT GAAAGACTCG CCAGAGCTAG ACCTCCTTGG  | 780 | 20 |
| TGTATGTTGT CTCAAGAAGA AAAAGACGAC ATGAAACAAC AGGTACATGA TTATATTTAT  | 840 |    |
| CTAGGAACAG GAATGCACTT TTGGGGAAAAG ATTTTCCATA CCAAGGAGGG GACAGTGGCT | 900 |    |
| GGACTAATAG AACATTATTC TGCAAAAACT CATGGCATGA GTTATTATGA ATAGCCTTTA  | 960 |    |

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| TTGGCCCAAC | CTTGCGGTTT | CCAGGGCTTA | AGTAAGTTTT | TGGTTACAAA | CTGTTCTTAA | 1020 |
| AACGAGGATG | TGAGACAAGT | GGTTTCCTGA | CTTGGTTTGG | TATCAAAGGT | TCTGATCTGA | 1080 |
| GCTCTGAGTG | TTCTATTTTC | CTATGTTCTT | TTGGAATTTA | TCCAAATCTT | ATGTAAATGC | 1140 |
| TTATGTAAAC | CAAGATATAA | AAGAGTGCTG | ATTTTTTGAG | TAAACTTGCA | ACAGTCCTAA | 1200 |
| CATTCACCTC | TTGTGTGTTT | GTGTCTGTTT | GCCATCCCGT | CTCGCTCGT  | CACTTATCCT | 1260 |
| TCACTTTCCA | GAGGGTCCCC | CCGCAGACCC | CGGCGACCCT | CAGGTCGGCC | GACTGCGGCA | 1320 |
| GCTGGOGCCC | GAACAGGGAC | CCTCGGATAA | GTGACCCTTG | TCTCTATTTT | TACTATTTGG | 1380 |
| TGTTTGTCTT | GTATTGTCTC | TTTCTGTCTT | GGCTATCATC | ACAAGAGCGG | AACGGACTCA | 1440 |
| CCATAGGGAC | CAAGCTAGCG | ACTGAAAATG | AGACATATTA | TCTGCCACGG | AGGTGTTATT | 1500 |
| ACCGAAGAAA | TGGCCGCCAG | TCTTTTGGAC | CAGCTGATCG | AAGAGGTACT | GGCTGATAAT | 1560 |
| CTTCCACCTC | CTAGCCATTT | TGAACCACCT | ACCCTTCACG | AACTGTATGA | TTTAGACGTG | 1620 |
| ACGGCCCCCG | AAGATCCCAA | CGAGGAGGCG | GTTTCGCAGA | TTTTTCCCGA | CTCTGTAATG | 1680 |
| TTGGCGGTGC | AGGAAGGGAT | TGACTTACTC | ACTTTTCCGC | CGGCGCCCGG | TTCTCCGGAG | 1740 |
| CCGCCTCACC | TTTCCCGGCA | GCCCGAGCAG | CCGGAGCAGA | GAGCCTTGGG | TCCGTTTCTT | 1800 |
| ATGCCAAACC | TTGTACCGGA | GGTGATCGAT | CTTACCTGCC | ACGAGGCTGG | CTTCCACCCC | 1860 |
| AGTGACGACG | AGGATGAAGA | GGGTGAGGAG | TTTGTGTTAG | ATTATGTGGA | GCACCCCGGG | 1920 |
| CACGGTTGCA | GGTCTTGTC  | TTATCACCGG | AGGAATACGG | GGGACCCAGA | TATTATGTGT | 1980 |
| TGCCTTTGCT | ATATGAGGAC | CTGTGGCATG | TTTGTCTACA | GTAAGTGAAA | ATTATGGGCA | 2040 |
| GTGGGTGATA | GAGTGGTGGG | TTTGGTGTGG | TAATTTTTTT | TTAATTTTTT | ACAGTTTTGT | 2100 |
| GGTTTAAAGA | ATTTTGTATT | GTGATTTTTT | TAAAAGGTCC | TGTGTCTGAA | CCTGAGCCTG | 2160 |
| AGCCCGAGCC | AGAACCGGAG | CCTGCAAGAC | CTACCCGCCG | TCCTAAAATG | GCGCCTGCTA | 2220 |
| TCCTGAGACG | CCCGACATCA | CCTGTGTCTA | GAGAATGCAA | TAGTAGTACG | GATAGCTGTG | 2280 |
| ACTCCGGTCC | TTCTAACACA | CCTCCTGAGA | TACACCCGGT | GGTCCCGCTG | TGCCCCATTA | 2340 |
| AACCAGTTGC | CGTGAGAGTT | GGTGGGCGTC | GCCAGGCTGT | GGAATGTATC | GAGGACTTGC | 2400 |
| TTAACGAGCC | TGGGCAACCT | TTGGACTTGA | GCTGTAAACG | CCCCAGGCCA | TAAGGTGTAA | 2460 |
| ACCTGTGATT | GCGTGTGTGG | TTAACGCCTT | TGTTTGCTGA | ATGAGTTGAT | GTAAGTTTAA | 2520 |
| TAAAGGGTGA | GATAATGTTT | AACTTGCATG | GCGTGTTAAA | TGGGGCGGGG | CTTAAAGGGT | 2580 |
| ATATAATGCG | CCGTGGGCTA | ATCTTGGTTA | CATCTGACCT | CATGGAGGCT | TGGGAGTGTT | 2640 |
| TGGAAGATTT | TTCTGCTGTG | CGTAACTTGC | TGGAACAGAG | CTCTAACAGT | ACCTCTTGGT | 2700 |

10

20

30

40

|  |      |    |
|--|------|----|
| TTTGGAGGTT TCTGTGGGGC TCATCCCAGG CAAAGTTAGT CTGCAGAATT AAGGAGGATT  | 2760 |    |
| ACAAGTGGGA ATTTGAAGAG CTTTTGAAAT CCTGTGGTGA GCTGTTTGAT TCTTTGAATC  | 2820 |    |
| TGGGTCACCA GCGCTTTTC CAAGAGAAGG TCATCAAGAC TTTGATTTT TCCACACCGG    | 2880 |    |
| GGCGCGCTGC GGCTGCTGTT GCTTTTTTGA GTTTTATAAA GGATAAATGG AGCGAAGAAA  | 2940 |    |
| CCCATCTGAG CGGGGGGTAC CTGCTGGATT TTCTGGCCAT GCATCTGTGG AGAGCGGTTG  | 3000 |    |
| TGAGACACAA GAATCGCCTG CTA CTGTTGT CTTCGTC CGCGCGATA ATACCGACGG     | 3060 | 10 |
| AGGAGCAGCA GCAGCAGCAG GAGGAAGCCA GCGCGCGGCG GCAGGAGCAG AGCCCATGGA  | 3120 |    |
| ACCCGAGAGC CGGCCTGGAC CCTCGGAAT GAATGTTGTA CAGGTGGCTG AACTGTATCC   | 3180 |    |
| AGAACTGAGA CGCATTTTGA CAATTACAGA GGATGGGCAG GGGCTAAAGG GGGTAAAGAG  | 3240 |    |
| GGAGCGGGGG GCTTGTGAGG CTACAGAGGA GGCTAGGAAT CTAGCTTTTA GCTTAATGAC  | 3300 |    |
| CAGACACCGT CCTGAGTGTA TTA CTTTCA ACAGATCAAG GATAATTGCG CTAATGAGCT  | 3360 |    |
| TGATCTGCTG GCGCAGAAGT ATTCCATAGA GCAGCTGACC ACTTACTGGC TGCAGCCAGG  | 3420 | 20 |
| GGATGATTTT GAGGAGGCTA TTAGGGTATA TGCAAAGGTG GCACTTAGGC CAGATTGCAA  | 3480 |    |
| GTACAAGATC AGCAA ACTTG TAAATATCAG GAATTGTTGC TACATTTCTG GGAACGGGGC | 3540 |    |
| CGAGGTGGAG ATAGATACGG AGGATAGGGT GGCCTTTAGA TG TAGCATGA TAAATATGTG | 3600 |    |
| GCCGGGGGTG CTTGGCATGG ACGGGGTGGT TATTATGAAT GTAAGGTTTA CTGGCCCCAA  | 3660 |    |
| TTTTAGCGGT ACGGTTTCC TGGCCAATAC CAACCTTATC CTACACGGTG TAAGCTTCTA   | 3720 |    |
| TGGGTTTAAC AATACCTGTG TGGAAGCCTG GACCGATGTA AGGGTTCGGG GCTGTGCCTT  | 3780 |    |
| TTACTGCTGC TGGAAGGGGG TGGTGTGTCG CCCCCAAGC AGGGCTTCAA TTAAGAAATG   | 3840 | 30 |
| CCTCTTTGAA AGGTGTACCT TGGGTATCCT GTCTGAGGGT AACTCCAGGG TGCGCCACAA  | 3900 |    |
| TGTGGCCTCC GACTGTGGTT GCTTCATGCT AGTGAAAAGC GTGGCTGTGA TTAAGCATAA  | 3960 |    |
| CATGGTATGT GGCAACTGCG AGGACAGGGC CTCTCAGATG CTGACCTGCT CGGACGGCAA  | 4020 |    |
| CTGTCACCTG CTGAAGACCA TTCACGTAGC CAGCCACTCT CGCAAGGCCT GGCCAGTGTT  | 4080 |    |
| TGAGCATAAC ATACTGACCC GCTGTTCCCT GCATTTGGGT AACAGGAGGG GGGTGTTCCT  | 4140 |    |
| ACCTTACCAA TGCAATTTGA GTCACACTAA GATATTGCTT GAGCCCGAGA GCATGTCCAA  | 4200 | 40 |
| GGTGAACCTG AACGGGGTGT TTGACATGAC CATGAAGATC TGGAAGGTGC TGAGGTACGA  | 4260 |    |
| TGAGACCCGC ACCAGGTGCA GACCCTGCGA GTGTGGCGGT AAACATATTA GGAACCAGCC  | 4320 |    |
| TGTGATGCTG GATGTGACCG AGGAGCTGAG GCCCGATCAC TTGGTGCTGG CCTGCACCCG  | 4380 |    |
| CGCTGAGTTT GGCTCTAGCG ATGAAGATAC AGATTGAGGT ACTGAAATGT GTGGGCGTGG  | 4440 |    |

|   |      |    |
|---|------|----|
| CTTAAGGGTG GGAAGAATA TATAAGGTGG GGGTCTTATG TAGTTTTGTA TCTGTTTTGC  | 4500 |    |
| AGCAGCCGCC GCCGCCATGA GCACCAACTC GTTTGATGGA AGCATTGTGA GCTCATATTT | 4560 |    |
| GACAACGCGC ATGCCCCCAT GGGCCGGGGT GCGTCAGAAT GTGATGGGCT CCAGCATTGA | 4620 |    |
| TGGTCGCCCC GTCCTGCCCC CAACTCTAC TACCTTGACC TACGAGACCG TGTCTGGAAC  | 4680 |    |
| GCCGTTGGAG ACTGCAGCCT CCGCCGCCGC TTCAGCCGCT GCAGCCACCG CCCGCGGGAT | 4740 |    |
| TGTGACTGAC TTTGCTTTCC TGAGCCCGCT TGCAAGCAGT GCAGCTTCCC GTTCATCCGC | 4800 | 10 |
| CCGCGATGAC AAGTTGACGG CTCTTTTGGC ACAATTGGAT TCTTTGACCC GGGAACTTAA | 4860 |    |
| TGTCGTTTCT CAGCAGCTGT TGGATCTGCG CCAGCAGGTT TCTGCCCTGA AGGCTTCCTC | 4920 |    |
| CCCTCCCAAT GCGGTTTAAA ACATAAATAA AAAACCAGAC TCTGTTTGA TTTGGATCAA  | 4980 |    |
| GCAAGTGTCT TGCTGTCTCT CGAGGGATCT TTGTGAAGGA ACCTTACTTC TGTGGTGTGA | 5040 |    |
| CATAATTGGA CAACTACCT ACAGAGATTT AAAGCTCTAA GGTAATATA AAATTTTAA    | 5100 |    |
| GTGTATAATG TGTAAACTA CTGATTCTAA TTGTTTGTGT ATTTTAGATT CCAACCTATG  | 5160 | 20 |
| GAAGTATGA ATGGGAGCAG TGGTGAATG CCTTAAATGA GGAAAACCTG TTTTGCTCAG   | 5220 |    |
| AAGAAATGCC ATCTAGTGAT GATGAGGCTA CTGCTGACTC TCAACATTCT ACTCCTCAA  | 5280 |    |
| AAAAGAAGAG AAAGGTAGAA GACCCCAAGG ACTTTCCTTC AGAATTGCTA AGTTTTTTGA | 5340 |    |
| GTCATGCTGT GTTTAGTAAT AGAACTCTTG CTGCTTTTGC TATTTACACC ACAAAGGAAA | 5400 |    |
| AAGCTGCACT GCTATACAAG AAAATTATGG AAAAATATTC TGTAACCTTT ATAAGTAGGC | 5460 |    |
| ATAACAGTTA TAATCATAAC ATACTGTTTT TTCTTACTCC ACACAGGCAT AGAGTGTCTG | 5520 |    |
| CTATTAATAA CTATGCTCAA AAATTGTGTA CCTTTAGCTT TTAAATTTGT AAAGGGGTTA | 5580 | 30 |
| ATAAGGAATA TTTGATGTAT AGTGCCCTGA CTAGAGATCA TAATCAGCCA TACCACATTT | 5640 |    |
| GTAGAGGTTT TACTTGCTTT AAAAAACCTC CCACACCTCC CCCTGAACCT GAAACATAAA | 5700 |    |
| ATGAATGCAA TTGTTGTTGT TAACTTGTTT ATTGCAGCTT ATAATGGTTA CAAATAAAGC | 5760 |    |
| AATAGCATCA CAAATTTTAC AAATAAAGCA TTTTTTTCAC TGCATTCTAG TTGTGGTTTG | 5820 |    |
| TCCAAACTCA TCAATGTATC TTATCATGTC TGGATCCGGC TGTGGAATGT GTGTCAGTTA | 5880 |    |
| GGGTGTGGAA AGTCCCAGG CTCCCAGCA GGCAGAAGTA TGCAAAGCAT GCATCTCAAT   | 5940 | 40 |
| TAGTCAGCAA CCAGGTGTGG AAAGTCCCCA GGCTCCCCAG CAGGCAGAAG TATGCAAAGC | 6000 |    |
| ATGCATCTCA ATTAGTCAGC AACCATAGTC CCGCCCCTAA CTCGCCCCAT CCGCCCCCTA | 6060 |    |
| ACTCGCCCA GTTCGCCCCA TTCTCGCCC CATGGCTGAC TAATTTTTTT TATTTATGCA   | 6120 |    |
| GAGGCCGAGG CCGCCTCGGC CTCTGAGCTA TTCCAGAAGT AGTGAGGAGG CTTTTTTTGA | 6180 |    |

|   |      |    |
|---|------|----|
| GGCCTAGGET TTTGCAAAAA GCTTGGACAC AAGACAGGCT TGCAGATAT GTTTGAGAAT  | 6240 |    |
| ACCACTTTAT CCCGCGTCAG GGAGAGGCAG TCGTAAAAA GACGCGGACT CATGTGAAAT  | 6300 |    |
| ACTGGTTTTT AGTGCGCCAG ATCTCTATAA TCTCGCGCAA CCTATTTTCC CCTCGAACAC | 6360 |    |
| TTTTTAAGCC GTAGATAAAC AGGCTGGGAC ACTTCACATG AGCGAAAAAT ACATCGTCAC | 6420 |    |
| CTGGGACATG TTGCAGATCC ATGCACGTAA ACTCGCAAGC CGACTGATGC CTTCTGAACA | 6480 |    |
| ATGGAAGGC ATTATTGCCG TAAGCCGTGG CGGTCTGGTA CCGGTGCGT TACTGGCGCG   | 6540 | 10 |
| TGAACTGGGT ATTCGTCATG TCGATACCGT TTGTATTTCC AGCTACGATC ACGACAACCA | 6600 |    |
| GCGCGAGCTT AAAGTGCTGA AACGCGCAGA AGGCGATGGC GAAGGCTTCA TCGTTATTGA | 6660 |    |
| TGACCTGGTG GATACCGGTG GTACTGCGGT TCGATTTCGT GAAATGTATC CAAAAGCGCA | 6720 |    |
| CTTTGTCACC ATCTTCGCAA AACCGGCTGG TCGTCCGCTG GTTGATGACT ATGTTGTTGA | 6780 |    |
| TATCCCGCAA GATACCTGGA TTGAACAGCC GTGGGATATG GGCCTCGTAT TCGTCCCGCC | 6840 |    |
| AATCTCCGGT CGCTAATCTT TTCAACGCCT GGCCTGCGG GGCCTTGTTT TTTTAACTT   | 6900 | 20 |
| CAGGCGGGTT ACAATAGTTT CCAGTAAGTA TTCTGGAGGC TGCATCCATG ACACAGGCAA | 6960 |    |
| ACCTGAGCGA AACCCTGTTC AAACCCCGCT TTAAACATCC TGAAACCTCG ACGCTAGTCC | 7020 |    |
| GCCGCTTTAA TCACGGCGCA CAACCGCCTG TGCAGTCGGC CCTTGATGGT AAAACCATCC | 7080 |    |
| CTCACTGGTA TCGCATGATT AACCGTCTGA TGTGGATCTG GCGCGGCATT GACCCACGCG | 7140 |    |
| AAATCCTCGA CGTCCAGGCA CGTATTGTGA TGAGCGATGC CGAACGTACC GACGATGATT | 7200 |    |
| TATACGATAC GGTGATTGGC TACCGTGGCG GCAACTGGAT TTATGAGTGG GCCCCGGATC | 7260 |    |
| TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG ACAAATACC TACAGAGATT  | 7320 | 30 |
| TAAAGCTCTA AGGTAAATAT AAAATTTTAA AGTGTATAAT GTGTAAACT ACTGATTCTA  | 7380 |    |
| ATTGTTTGTG TATTTTAGAT TCCAACCTAT GGAAGTATG AATGGGAGCA GTGGTGGAAT  | 7440 |    |
| GCCTTTAATG AGGAAAACCT GTTTTGCTCA GAAGAAATGC CATCTAGTGA TGATGAGGCT | 7500 |    |
| ACTGCTGACT CTCAACATTC TACTCCTCCA AAAAAGAAGA GAAAGGTAGA AGACCCCAAG | 7560 |    |
| GACTTTCCTT CAGAATTGCT AAGTTTTTTG AGTCATGCTG TGTTTAGTAA TAGAACTCTT | 7620 |    |
| GCTTGCTTTG CTATTTACAC CACAAAGGAA AAAGCTGCAC TGCTATACAA GAAAATTATG | 7680 | 40 |
| GAAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT ATAATCATAA CATACTGTTT | 7740 |    |
| TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA ACTATGCTCA AAAATTGTGT | 7800 |    |
| ACCTTTAGCT TTTTAATTTG TAAAGGGGTT AATAAGGAAT ATTTGATGTA TAGTGCCTTG | 7860 |    |
| ACTAGAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT TTAATTGCTT TAAAAACCT  | 7920 |    |



|  |      |    |
|--|------|----|
| CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA ATTGTTGTTG TTAACCTTGT  | 7980 |    |
| TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA CAAATAAAGC  | 8040 |    |
| ATTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC ATCAATGTAT CTTATCATGT  | 8100 |    |
| CTGGATCCCC AGGAAGCTCC TCTGTGTCCT CATAAACCCCT AACCTCCTCT ACTTGAGAGG | 8160 |    |
| ACATTCCAAT CATAGGCTGC CCATCCACCC TCTGTGTCCT CCTGTTAATT AGGTCACCTA  | 8220 |    |
| ACAAAAAGGA AATTGGGTAG GGGTTTTTCA CAGACCGCTT TCTAAGGGTA ATTTTAAAT   | 8280 | 10 |
| ATCTGGGAAG TCCCTTCCAC TGCTGTGTTT CAGAAGTGTT GGTAACAGC CCACAAATGT   | 8340 |    |
| CAACAGCAGA AACATACAAG CTGTCAGCTT TGCACAAGGG CCCAACACCC TGCTCATCAA  | 8400 |    |
| GAAGCACTGT GGTGCTGTG TTAGTAATGT GCAAAACAGG AGGCACATTT TCCCCACCTG   | 8460 |    |
| TGTAGGTTCC AAAATATCTA GTGTTTTCAT TTTTACTTGG ATCAGGAACC CAGCACTCCA  | 8520 |    |
| CTGGATAAGC ATTATCCTTA TCCAAAACAG CCTTGTGGTC AGTGTTTCATC TGCTGACTGT | 8580 |    |
| CAACTGTAGC ATTTTTTGGG GTTACAGTTT GAGCAGGATA TTTGGTCCTG TAGTTTGCTA  | 8640 | 20 |
| ACACACCCTG CAGCTCCAAA GGTTCCTCCAC CAACAGCAAA AAAATGAAAA TTTGACCCTT | 8700 |    |
| GAATGGGTTT TCCAGCACCA TTTTCATGAG TTTTTGTGT CCCTGAATGC AAGTTTAACA   | 8760 |    |
| TAGCAGTTAC CCCAATAACC TCAGTTTTAA CAGTAACAGC TTCCCACATC AAAATATTTT  | 8820 |    |
| CACAGGTAA GTCCCTCATTT AAATTAGGCA AAGGAATTCT TGAAGACGAA AGGGCCTCGT  | 8880 |    |
| GATACGCTA TTTTATAGG TTAATGTCAT GATAATAATG GTTCTTAGA CGTCAGGTGG     | 8940 |    |
| CACTTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTTA TTTTCTAAA TACATTCAAA   | 9000 | 30 |
| TATGTATCCG CTCATGAGAC AATAACCCTG ATAAATGCTT CAATAATATT GAAAAAGGAA  | 9060 |    |
| GAGTATGAGT ATTCAACATT TCCGTGTCGC CCTTATTCCC TTTTTCGGG CATTTTGCCT   | 9120 |    |
| TCCTGTTTTT GCTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG  | 9180 |    |
| TGCACGAGTG GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTTCG  | 9240 |    |
| CCCCGAAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTATGTG GCGCGGTATT  | 9300 |    |
| ATCCCGTGTT GACGCCGGG AAGAGCAACT CGGTGCGCGC ATACACTATT CTCAGAATGA   | 9360 |    |
| CTTGGTTGAG TACTCACCAG TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA  | 9420 | 40 |
| ATTATGCAGT GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGACAAC  | 9480 |    |
| GATCGGAGGA CCGAAGGAGC TAACCGCTTT TTTGCACAAC ATGGGGGATC ATGTAACCTG  | 9540 |    |
| CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC  | 9600 |    |
| GATGCCTGCA GCAATGGCAA CAACGTTGCG CAAACTATTA ACTGGCGAAC TACTTACTCT  | 9660 |    |

AGCTTCCCGG CAACAATTAA TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCACTTCT 9720  
 GCGCTCGGCC CTTCGGCTG GCTGGTTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG 9780  
 GTCTCGCGGT ATCATTGCAG CACTGGGGCC AGATGGTAAG CCCTCCCGTA TCGTAGTTAT 9840  
 CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG 9900  
 TGCCTCACTG ATTAAGCATT GGTAAGTGTG AGACCAAGTT TACTCATATA TACTTTAGAT 9960  
 TGATTTAAAA CTTCATTTTT AATTTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT 10020  
 CATGACCAAA ATCCCTTAAC GTGAGTTTTT GTTCCACTGA GCGTCAGACC CCGTAGAAAA 10080  
 GATCAAAGGA TCTTCTTGAG ATCCTTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAACAAA 10140  
 AAAACCACCG CTACCAGCGG TGGTTTGTTT GCCGGATCAA GAGCTACCAA CTCTTTTTC 10200  
 GAAGGTAAGT GGCTTCAGCA GAGCGCAGAT ACCAAATACT GTCCTTCTAG TGTAGCCGTA 10260  
 GTTAGGCCAC CACTTCAAGA ACTCTGTAGC ACCGCCTACA TACCTCGCTC TGCTAATCCT 10320  
 GTTACCAGTG GCTGCTGCCA GTGGCGATAA GTCGTGTCTT ACCGGGTGG ACTCAAGACG 10380  
 ATAGTTACCG GATAAGGCGC AGCGGTCGGG CTGAACGGGG GGTTCGTGCA CACAGCCCAG 10440  
 CTTGGAGCGA ACGACCTACA CCGAACTGAG ATACCTACAG CGTGAGCTAT GAGAAAGCGC 10500  
 CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG 10560  
 AGAGCGCAG AGGGAGCTTC CAGGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTGGGTT 10620  
 TCGCCACCTC TGAATTGAGC GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG 10680  
 GAAAAACGCC AGCAACGCGG CCTTTTTACG GTTCCTGGCC TTTTGCTGGC CTTTGTCTCA 10740  
 CATGTTCTTT CCTGCGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG CCTTTGAGTG 10800  
 AGCTGATACC GCTCGCCGCA GCCGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC 10860  
 GGAAGAGCGC CTGATGCGGT ATTTTCTCCT TACGCATCTG TGCGGTATTT CACACCGCAT 10920  
 ACCGCCTCAG AAGCCATAGA GCCCACCACA TCCCCAGCAT GCCTGCTATT GTCTTCCCAA 10980  
 TCCTCCCCCT TGCTGTCTTG CCCACCCCCA CCCCCAGAA TAGAATGACA CCTACTCAGA 11040  
 CAATGCGATG CAATTTCTC ATTTTATTAG GAAAGGACAG TGGGAGTGGC ACCTTCCAGG 11100  
 GTCAAGGAAG GCACGGGGGA GGGGCAAACA ACAGATGGCT GGCAACTAGA AGGCACAGTC 11160  
 GAGGCTGATC AGCGAGCTCT AGCATTTAGG TGACACTATA GAATAGGGCC CTCTAGATGC 11220  
 ATGCTCGAGC GGCCGCTTCT TTATTCTTGG GCAATGTATG AAAAAGTGTA AGAGGATGTG 11280  
 GCAAATATTT CATTAATGTA GTTGTGGCCA GACCAGTCCC ATGAAAATGA CATAGAGTAT 11340  
 GCACTTGGAG TTGTGTCTCC TGTTTCCTGT GTACCGTTTA GTGTAATGGT TAGTGTTACA 11400

10

20

30

40

GGTTTAGTTT TGTCTCCGTT TAAGTAAACT TGACTGACAA TGTACTTTT GGCAGTTTFA 11460  
CCGTGAGATT TTGGATAAGC TGATAGGTTA GGCATAAATC CAACAGCGTT TGTATAGGCT 11520  
GTGCCTTCAG TAAGATCTCC ATTTCTAAAG TTCCAATATT CTGGGTCCAG GAAGGAATTG 11580  
TTTAGTAGCA CTCCATTTTC GTCAAATCTT ATAATAAGAT GAGCACTTTG AACTGTTCCA 11640  
GATATTGGAG CCAAAGTCC TTTAACAGCC AAAACTGAAA CTGTAGCAAG TATTTGACTG 11700  
CCACATTTTG TTAAGACCAA AGTGAGTTTA GCATCTTCT CTGCATTTAG TCTACAGTTA 11760  
GGAGATGGAG CTGGTGTGGT CCACAAAGTT AGCTTATCAT TATTTTGTGTT TCCTACTGTA 11820  
ATGGCACCTG TGCTGTCAAA ACTAAGGCCA GTTCCTAGTT TAGGAACCAT AGCCTTGTTT 11880  
GAATCAAATT CTAGGCCATG GCCAATTTTT GTTTTGAGGG GATTTGTGTT TGGTGCATTA 11940  
GGTGAACCAA ATTCAAGCCC ATCTCCTGCA TTAATGGCTA TGGCTGTAGC GTCAAACATC 12000  
AACCCTTGG CAGTGCTTAG GTTAACCTCA AGCTTTTGG AATTGTTTGA AGCTGTAAAC 12060  
AAGTAAAGGC CTTTGTGTGA GTTAATATCC AAGTTGTGGG CTGAGTTTAT AAAAAGAGGG 12120  
CCCTGTCCTA GTCTTAGATT TAGTTGGTTT TGAGCATCAA ACGGATAACT AACATCAAGT 12180  
ATAAGGCGTC TGTTTTGAGA ATCAATCCTT AGTCCTCTG CTACATTAAG TTGCATATTG 12240  
CCTTGTAAT CAAAACCCAA GGCTCCAGTA ACTTTAGTTT GCAAGGAAGT ATTATTAATA 12300  
GTCACACCTG GACCAGTTGC TACGGTCAAA GTGTTTAGGT CGTCTGTTAC ATGCAAAGGA 12360  
GCCCCGTA CTAGTCCTAG TTTTCCATTT TGTGTATAAA TGGGCTCTTT CAAGTCAATG 12420  
CCCAAGCTAC CAGTGGCAGT AGTTAGAGGG GGTGAGGCAG TGATAGTAAG GGTACTGCTA 12480  
TCGGTGGTGG TGAGGGGGCC TGATGTTTGC AGGGCTAGCT TTCCTTCTGA CACTGTGAGG 12540  
GGTCCTTGGG TGGCAATGCT AAGTTTGAG TCGTGCACGG TTAGCGGGGC CTGTGATTGC 12600  
ATGGTGAGTG TGTGCCCCG GACCATTAGA GGTGCGGCGG CAGCCACAGT TAGGGCTTCT 12660  
GAGGTAAGTG TGAGGGGTGC AGATATTTCC AGGTTTATGT TTGACTTGGT TTTTTTGAGA 12720  
GGTGGGCTCA CAGTGGTTAC ATTTTGGGAG GTAAGGTTGC CGGCCTCGTC CAGAGAGAGG 12780  
CCGTTGCCCA TTTTGAGCGC AAGCATGCCA TTGGAGGTAA CTAGAGGTTT GGATAGGCGC 12840  
AAAGAGAGTA CCCCAGGGGG ACTCTCTTGA AACCATTGG GGGATACAAA GGGAGGAGTA 12900  
AGAAAAGGCA CAGTTGGAGG ACCGGTTTCC GTGTCATATG GATACACGGG GTTGAAGGTA 12960  
TCTTCAGACG GTCTTGCGCG CTTTCATCTG GATCTCAAGC CTGCCACACC TCACCTCGAC 13020  
CATCCGCCGT CTCAAGACCG CCTACTTTAA TTACATCATC AGCAGCACCT CCGCCAGAAA 13080  
CAACCCCGAC CGCCACCCGC TGCCGCCCGC CACGGTGCTC AGCCTACCTT GCGACTGTGA 13140

10

20

30

40

CTGGTTAGAC GCCTTTCTCG AGAGGTTTTT CGATCCGGTC GATGCGGACT CGCTCAGGTC 13200  
 CCTCGGTGGC GGAGTACCGT TCGGAGGCCG ACGGGTTTCC GATCCAAGAG TACTGGAAAG 13260  
 ACCGCGAAGA GTTTGTCTC AACCGCGAGC CCAACAGCGA GCTCGAATTC AGATCCGAGC 13320  
 TCGGTACCAA GCTTGGGTCT CCCTATAGTG AGTCGTATTA ATTTGATAA GCCAGTAAGC 13380  
 AGTGGGTCT CTAGTTAGCC AGAGAGCTCT GCTTATATAG ACCTCCCACC GTACACGCCT 13440  
 ACCGCCCATT TCGTCAATG GGGCGGAGTT GTTACGACAT TTTGGAAAGT CCCGTTGATT 13500  
 TTGGTGCCAA AACAACTCC CATTGACGTC AATGGGGTGG AGACTTGAA ATCCCCGTGA 13560  
 GTCAAACCGC TATCCACGCC CATTGATGTA CTGCCAAAAC CGCATCACCA TGGTAATAGC 13620  
 GATGACTAAT ACGTAGATGT ACTGCCAAGT AGGAAAGTCC CATAAGGTCA TGTACTGGGC 13680  
 ATAATGCCAG GCGGGCCATT TACCGTCATT GACGTCAATA GGGGGCGTAC TTGGCATATG 13740  
 ATACACTTGA TGTACTGCCA AGTGGGCAGT TTACCGTAAA TAGTCCACCC ATTGACGTCA 13800  
 ATGGAAAGTC CCTATTGGCG TTAATATGGG AACATACGTC ATTATTGACG TCAATGGGCG 13860  
 GGGGTCGTTG GCGGTCAGC CAGGCGGGCC ATTTACCGTA AGTTATGTAA CGCGGAATC 13920  
 CATATATGGG CTATGAACTA ATGACCCCGT AATTGATTAC TATTAATAAC TAGTCAATAA 13980  
 TCAATGTCAA CGCGTATATC TGGCCCGTAC ATCGCGAAGC AGCGCAAAAC GCCTAACCCCT 14040  
 AAGCAGATTC TTCATGCAAT TGTCGGTCAA GCCTTGCCTT GTTGTAGCTT AAATTTTGCT 14100  
 CGCGCACTAC TCAGCGACCT CCAACACACA AGCAGGGAGC AGATACTGGC TTAATATGTC 14160  
 GGCATCAGAG CAGATTGTAC TGAGAGTCGA CCATAGGGGA TCGGGAGATC TCCCGATCCG 14220  
 TCTATGGTGC ACTCTCAGTA CAATCTGCTC TGATGCCGCA TAGTTAAGCC AGTATACACT 14280  
 CCGCTATCGC TACGTGACTG GGTGATGGCT GCGCCCCGAC ACCCGCCAAC ACCCGCTGAC 14340  
 GCGCCCTGAC GGGCTTGTCT GCTCCCGGCA TCCGCTTACA GACAAGCTGT GACCGTCTCC 14400  
 GGGAGCTGCA TGTGTCAGAG GTTTTCACCG TCATCACCGA AACGCGCGAG GCAGC 14455

10

20

30

( 2 ) 配列番号 : 1 6 の情報 :

( i ) 配列の特性 :

( A ) 配列の長さ : 1 0 6 1 0 塩基対

( B ) 配列の型 : 核酸

( C ) 鎖の数 : 二本鎖

( D ) トポロジー : 環状

( ii ) 配列の種類 : D N A ( genomic )

( iii ) ハイボセティカル : NO

( iii ) アンチセンス : NO

( xi ) 配列の記載 : 配列番号 : 1 6 :

40

|  |      |    |
|--|------|----|
| GACGGATCGG GAGATCCGCG CGGTACACAG AATTCAGGAG ACACAACTCC AAGTGCATAC  | 60   |    |
| TCTATGTCAT TTTCATGGGA CTGGTCTGGC CACAACTACA TTAATGAAAT ATTTGCCACA  | 120  |    |
| TCCTCTTACA CTTTTTCATA CATTGCCCAA GAATAAAGAA TCGTTTGTGT TATGTTTCAA  | 180  |    |
| CGTGTTTATT TTTCAATTGC AGAAAATTC AAGTCATTTT TCATTCAGTA GTATAGCCCC   | 240  |    |
| ACCACCACAT AGCTTATACA GATCACCGTA CCTTAATCAA ACTCACAGAA CCCTAGTATT  | 300  |    |
| CAACCTGCCA CCTCCCTCCC AACACACAGA GTACACAGTC CTTTCTCCCC GGCTGGCCTT  | 360  |    |
| AAAAAGCATC ATATCATGGG TAACAGACAT ATTCTTAGGT GTTATATTCC ACACGGTTTC  | 420  | 10 |
| CTGTCGAGCC AAACGCTCAT CAGTGATATT AATAAACTCC CCGGGCAGCT CACTTAAGTT  | 480  |    |
| CATGTCGCTG TCCAGCTGCT GAGCCACAGG CTGCTGTCCA ACTTGCGGTT GCTTAACGGG  | 540  |    |
| CGGCGAAGGA GAAGTCCACG CCTACATGGG GGTAGAGTCA TAATCGTGCA TCAGGATAGG  | 600  |    |
| GCGGTGGTGC TGCAGCAGCG CGCGAATAAA CTGCTGCCGC CGCCGCTCCG TCCTGCAGGA  | 660  |    |
| ATACAACATG GCAGTGGTCT CCTCAGCGAT GATTGCGACC GCCCGCAGCA TAAGGCGCCT  | 720  |    |
| TGTCTCCGG GCACAGCAGC GCACCCTGAT CTCACTTAAA TCAGCACAGT AACTGCAGCA   | 780  | 20 |
| CAGCACCACA ATATTGTTCA AAATCCCACA GTGCAAGGCG CTGTATCCAA AGCTCATGGC  | 840  |    |
| GGGGACCACA GAACCCACGT GGCCATCATA CCACAAGCGC AGGTAGATTA AGTGGCGACC  | 900  |    |
| CCTCATAAAC ACGCTGGACA TAAACATTAC CTCTTTTGGC ATGTTGTAAT TCACCACCTC  | 960  |    |
| CCGGTACCAT ATAAACCTCT GATTAAACAT GGCGCCATCC ACCACCATCC TAAACCAGCT  | 1020 |    |
| GGCCAAAACC TGCCCGCCGG CTATACACTG CAGGGAACCG GGAAGTGAAC AATGACAGTG  | 1080 |    |
| GAGAGCCCAG GACTCGTAAC CATGGATCAT CATGCTCGTC ATGATATCAA TGTTGGCACA  | 1140 | 30 |
| ACACAGGCAC ACGTGCATAC ACTTCCTCAG GATTACAAGC TCCTCCCGCG TTAGAACCAT  | 1200 |    |
| ATCCCAGGGA ACAACCCATT CCTGAATCAG CGTAAATCCC AACTGCAGG GAAGACCTCG   | 1260 |    |
| CACGTAATC ACGTTGTGCA TTGTCAAAGT GTTACATTCG GGCAGCAGCG GATGATCCTC   | 1320 |    |
| CAGTATGGTA GCGCGGGTTT CTGTCTCAAA AGGAGGTAGA CGATCCCTAC TGTACGGAGT  | 1380 |    |
| GCGCCGAGAC AACCGAGATC GTGTTGGTCG TAGTGTCTATG CCAAATGGAA CGCCGGACGT | 1440 |    |
| AGTCATATTT CCTGAAGCAA AACCAGGTGC GGGCGTGACA AACAGATCTG CGTCTCCGGT  | 1500 | 40 |
| CTGCGCGCTT AGATCGCTCT GTGTAGTAGT TGTAGTATAT CCACTCTCTC AAAGCATCCA  | 1560 |    |

|   |      |    |
|---|------|----|
| GGCGCCCCCT GGCTTCGGGT TCTATGTAAA CTCCTTCATG CGCGCTGCC CTGATAACAT  | 1620 |    |
| CCACCACCGC AGAATAAGCC ACACCCAGCC AACCTACACA TTCGTTCTGC GAGTCACACA | 1680 |    |
| CGGGAGGAGC GGAAGAGCT GGAAGAACCA TGTTTTTTTT TTTATTCCAA AAGATTATCC  | 1740 |    |
| AAAACCTCAA AATGAAGATC TATTAAGTGA ACGCGCTCCC CTCCGGTGGC GTGGTCAAAC | 1800 |    |
| TCTACAGCCA AAGAACAGAT AATGGCATT TTAAGATGTT GCACAATGGC TTCCAAAAGG  | 1860 |    |
| CAAACGGCCC TCACGTCCAA GTGGACGTAA AGGCTAAACC CTTCAGGGTG AATCTCCTCT | 1920 | 10 |
| ATAAACATTC CAGCACCTTC AACCATGCCC AAATAATTCT CATCTCGCCA CCTTCTCAAT | 1980 |    |
| ATATCTCTAA GCAAATCCCG AATATTAAGT CCGGCCATTG TAAAAATCTG CTCCAGAGCG | 2040 |    |
| CCCTCCACCT TCAGCCTCAA GCAGCGAATC ATGATTGCAA AAATTCAGGT TCCTCACAGA | 2100 |    |
| CCTGTATAAG ATTCAAAAGC GGAACATTAA CAAAATACC GCGATCCCGT AGGTCCCTTC  | 2160 |    |
| GCAGGGCCAG CTGAACATAA TCGTGCAGGT CTGCACGGAC CAGCGCGGCC ACTTCCCCGC | 2220 |    |
| CAGGAACCTT GACAAAAGAA CCCACACTGA TTATGACACG CATACTCGGA GCTATGCTAA | 2280 | 20 |
| CCAGCGTAGC CCCGATGTAA GCTTTGTTGC ATGGGCGGCG ATATAAAATG CAAGGTGCTG | 2340 |    |
| CTCAAAAAAT CAGGCAAAGC CTCGCGCAAA AAAGAAAGCA CATCGTAGTC ATGCTCATGC | 2400 |    |
| AGATAAAGGC AGGTAAGCTC CGGAACCACC ACAGAAAAAG ACACCATTTT TCTCTCAAAC | 2460 |    |
| ATGTCTGGG GTTTCTGCAT AAACACAAAA TAAAATAACA AAAAAACATT TAAACATTAG  | 2520 |    |
| AAGCCTGTCT TACAACAGGA AAAACAACCC TTATAAGCAT AAGACGGACT ACGGCCATGC | 2580 |    |
| CGGCGTGACC GTAAAAAAC TGGTCACCGT GATTAAAAAG CACCACCGAC AGCTCCTCGG  | 2640 |    |
| TCATGTCCGG AGTCATAATG TAAGACTCGG TAAACACATC AGGTTGATTC ATCGGTCAGT | 2700 | 30 |
| GCTAAAAAGC GACCGAAATA GCCCGGGGGA ATACATACCC GCAGGCGTAG AGACAACATT | 2760 |    |
| ACAGCCCCCA TAGGAGGTAT AACAAAATTA ATAGGAGAGA AAAACACATA AACACCTGAA | 2820 |    |
| AAACCTCCT GCCTAGGCAA AATAGCACCC TCCGCTCCA GAACAACATA CAGCGCTTCA   | 2880 |    |
| CAGCGGCAGC CTAACAGTCA GCCTTACCAG TAAAAAGAA AACCTATTAA AAAAACACCA  | 2940 |    |
| CTCGACACGG CACCAGCTCA ATCAGTCACA GTGTAAAAAA GGGCCAAGTG CAGAGCGAGT | 3000 |    |
| ATATATAGGA CTAAAAATG ACGTAACGGT TAAAGTCCAC AAAAAACACC CAGAAAACCG  | 3060 | 40 |
| CACGCGAACC TACGCCAGA AACGAAAGCC AAAAAACCA CAACTTCCTC AAATCGTCAC   | 3120 |    |
| TTCCGTTTTT CCACGTTACG TAACTTCCCG GATCCTCTCC CGATCCCCTA TGGTCGACTC | 3180 |    |
| TCAGTACAAT CTGCTCTGAT GCCGCATAGT TAAGCCAGTA TCTGCTCCCT GCTTGTGTGT | 3240 |    |
| TGGAGGTGCG TGAGTAGTGC GCGAGCAAAA TTAAAGCTAC AACAAGGCAA GGCTTGACCG | 3300 |    |

|   |      |    |
|---|------|----|
| ACAATTGCAT GAAGAATCTG CTTAGGGTTA GCGTTTTGC GCTGCTTCGC GATGTACGGG  | 3360 |    |
| CCAGATATAC GCGTTGACAT TGATTATTGA CTAGTTATTA ATAGTAATCA ATTACGGGGT | 3420 |    |
| CATTAGTTCA TAGCCCATAT ATGGAGTTCC GCGTTACATA ACTTACGGTA AATGGCCCCG | 3480 |    |
| CTGGCTGACC GCCCAACGAC CCCCGCCCAT TGACGTCAAT AATGACGTAT GTTCCCATAG | 3540 |    |
| TAACGCCAAT AGGGACTTTC CATTGACGTC AATGGGTGGA CTATTTACGG TAAACTGCCC | 3600 |    |
| ACTTGGCAGT ACATCAAGTG TATCATATGC CAAGTACGCC CCCTATTGAC GTCAATGACG | 3660 | 10 |
| GTAAATGGCC CGCCTGGCAT TATGCCCAGT ACATGACCTT ATGGGACTTT CCTACTTGGC | 3720 |    |
| AGTACATCTA CGTATTAGTC ATCGCTATTA CCATGGTGAT GCGGTTTTGG CAGTACATCA | 3780 |    |
| ATGGGCGTGG ATAGCGGTTT GACTCACGGG GATTTCCAAG TCTCCACCCC ATTGACGTCA | 3840 |    |
| ATGGGAGTTT GTTTTGGCAC CAAAATCAAC GGGACTTTCC AAAATGTCGT AACAACTCCG | 3900 |    |
| CCCCATTGAC GCAAATGGGC GGTAGGCGTG TACGGTGGGA GGTCTATATA AGCAGAGCTC | 3960 |    |
| TCTGGCTAAC TAGAGAACCC ACTGCTTACT GGCTTATCGA AATTAATACG ACTCACTATA | 4020 | 20 |
| GGGAGACCCA AGCTTGGTAC CGAGCTCGGA TCTGAATTCG AGCTCGCTGT TGGGCTCGCG | 4080 |    |
| GTTGAGGACA AACTCTTCGC GGTCTTTCCA GTACTCTTGG ATCGGAAACC CGTCGGCCTC | 4140 |    |
| CGAACGGTAC TCCGCCACCG AGGGACCTGA GCGAGTCCGC ATCGACCGGA TCGGAAAACC | 4200 |    |
| TCTCGAGAAA GCGTCTAAC CAGTCACAGT CGCAAGGTAG GCTGAGCACC GTGGCGGGCG  | 4260 |    |
| GCAGCGGGTG GCGGTCGGGG TTGTTTCTGG CGGAGGTGCT GCTGATGATG TAATTAAAGT | 4320 |    |
| AGGCGGTCTT GAGACGGCGG ATGGTCGAGG TGAGGTGTGG CAGGCTTGAG ATCCAAGATG | 4380 |    |
| AAGCGCGCAA GACCGTCTGA AGATACCTTC AACCCCGTGT ATCCATATGA CACGGAAACC | 4440 | 30 |
| GGTCCTCAA CTGTGCCTTT TCTTACTCCT CCCTTTGTAT CCCCCAATGG GTTTCAAGAG  | 4500 |    |
| AGTCCCCCTG GGGTACTCTC TTTGCGCCTA TCCGAACCTC TAGTTACCTC CAATGGCATG | 4560 |    |
| CTTGCGCTCA AAATGGGCAA CGGCCTCTCT CTGGACGAGG CCGGCAACCT TACCTCCCAA | 4620 |    |
| AATGTAACCA CTGTGAGCCC ACCTCTCAAA AAAACCAAGT CAAACATAAA CCTGGAAATA | 4680 |    |
| TCTGCACCCC TCACAGTTAC CTCAGAAGCC CTAAGTGTGG CTGCCGCCGC ACCTCTAATG | 4740 |    |
| GTCGCGGGCA ACACACTCAC CATGCAATCA CAGGCCCCGC TAACCGTGCA CGACTCCAAA | 4800 | 40 |
| CTTAGCATTG CCACCCAAGG ACCCCTCACA GTGTCAGAAG GAAAGCTAGC CCTGCAAACA | 4860 |    |
| TCAGGCCCCC TCACCACCAC CGATAGCAGT ACCCTTACTA TCACTGCCTC ACCCCTCTA  | 4920 |    |
| ACTACTGCCA CTGGTAGCTT GGGCATTGAC TTGAAAGAGC CCATTTATAC ACAAATGGA  | 4980 |    |
| AAACTAGGAC TAAAGTACGG GGCTCCTTTG CATGTAACAG ACGACCTAAA CACTTTGACC | 5040 |    |

|   |      |    |
|---|------|----|
| GTAGCAACTG GTCCAGGTGT GACTATTAAT AATACTTCCT TGCAAACATA AGTTACTGGA | 5100 |    |
| GCCTTGGGTT TTGATTCACA AGGCAATATG CAACTTAATG TAGCAGGAGG ACTAAGGATT | 5160 |    |
| GATTCTCAAA ACAGACGCCT TATACTTGAT GTTAGTTATC CGTTTGATGC TCAAAACCAA | 5220 |    |
| CTAAATCTAA GACTAGGACA GGGCCCTCTT TTTATAAACT CAGCCCACAA CTTGGATATT | 5280 |    |
| AACTACAACA AAGGCCTTTA CTTGTTTACA GCTTCAAACA ATTCCAAAAA GCTTGAGGTT | 5340 |    |
| AACCTAAGCA CTGCCAAGGG GTTGATGTTT GACGCTACAG CCATAGCCAT TAATGCAGGA | 5400 | 10 |
| GATGGGCTTG AATTTGGTTC ACCTAATGCA CCAAACACAA ATCCCCTCAA AACAAAAATT | 5460 |    |
| GGCCATGGCC TAGAATTTGA TTCAAACAAG GCTATGGTTC CTAAACTAGG AACTGGCCTT | 5520 |    |
| AGTTTTGACA GCACAGGTGC CATTACAGTA GGAAACAAAA ATAATGATAA GCTAACTTTG | 5580 |    |
| TGGACCACAC CAGCTCCATC TCCTAACTGT AGACTAAATG CAGAGAAAGA TGCTAAACTC | 5640 |    |
| ACTTTGGTCT TAACAAAATG TGGCAGTCAA AACTTTGCTA CAGTTTCAGT TTTGGCTGTT | 5700 |    |
| AAAGGCAGTT TGGCTCCAAT ATCTGGAACA GTTCAAAGTG CTCATCTTAT TATAAGATTT | 5760 | 20 |
| GACGAAAATG GAGTGCTACT AAACAATTCC TTCCTGGACC CAGAATATTG GAACTTTAGA | 5820 |    |
| AATGGAGATC TTAAGGAAGG CACAGCCTAT ACAAACGCTG TTGGATTTAT GCCTAACCTA | 5880 |    |
| TCAGCTTATC CAAAATCTCA CGGTAAAACT GCCAAAAGTA ACATTGTCAG TCAAGTTTAC | 5940 |    |
| TTAAACGGAG ACAAACATAA ACCTGTAACA CTAACCATTA CACTAAACGG TACACAGGAA | 6000 |    |
| ACAGGAGACA CAACTCCAAG TGCATACTCT ATGTCATTTT CATGGGACTG GTCTGGCCAC | 6060 |    |
| AACTACATTA ATGAAATATT TGCCACATCC TCTTACACTT TTTCATACAT TGCCCAAGAA | 6120 |    |
| TAAAGAAGCG GCCGCTCGAG CATGCATCTA GAGGGCCCTA TTCTATAGTG TCACCTAAAT | 6180 | 30 |
| GCTAGAGCTC GCTGATCAGC CTCGACTGTG CCTTCTAGTT GCCAGCCATC TGTTGTTTGC | 6240 |    |
| CCCTCCCCCG TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA | 6300 |    |
| AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG GGGTGGGGTG | 6360 |    |
| GGGCAGGACA GCAAGGGGGA GGATTGGGAA GACAATAGCA GGCATGCTGG GGATGCGGTG | 6420 |    |
| GGCTCTATGG CTTCTGAGGC GGAAAGAACC AGCTGGGGCT CTAGGGGGTA TCCCCACGCG | 6480 |    |
| CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGTTTA CGCGCAGCGT GACCGCTACA | 6540 | 40 |
| CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTCTTCC CTTCCTTTCT CGCCACGTTT  | 6600 |    |
| GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGCATCCCTT TAGGGTTCCG ATTTAGTGCT | 6660 |    |
| TTACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTGATG GTTCACGTAG TGGGCCATCG | 6720 |    |
| CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA TAGTGGACTC | 6780 |    |



|  |      |    |
|--|------|----|
| TTGTTCCAAA CTGGAACAAC ACTCAACCT ATCTCGGTCT ATTCTTTTGA TTTATAAGGG   | 6840 |    |
| ATTTTGGGGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTAAACAAAA ATTTAACGCG  | 6900 |    |
| AATTAATTCT GTGGAATGTG TGTCAGTTAG GGTGTGAAAA GTCCCCAGGC TCCCCAGGCA  | 6960 |    |
| GGCAGAAGTA TGCAAAGCAT GCATCTCAAT TAGTCAGCAA CCAGGTGTGG AAAGTCCCCA  | 7020 |    |
| GGCTCCCCAG CAGGCAGAAG TATGCAAAGC ATGCATCTCA ATTAGTCAGC AACCATAGTC  | 7080 |    |
| CCGCCCCCTAA CTCCGCCCAT CCGCCCCCTA ACTCCGCCCA GTTCCGCCCA TTCTCCGCC  | 7140 | 10 |
| CATGGCTGAC TAATTTTTTT TATTTATGCA GAGGCCGAGG CCGCTCTGC CTCTGAGCTA   | 7200 |    |
| TTCCAGAAGT AGTGAGGAGG CTTTTTTGGA GGCCTAGGCT TTTGCAAAAA GCTCCCGGGA  | 7260 |    |
| GCTTGTATAT CCATTTTCGG ATCTGATCAA GAGACAGGAT GAGGATCGTT TCGCATGATT  | 7320 |    |
| GAACAAGATG GATTGCACGC AGGTTCTCCG GCGCTTGGG TGGAGAGGCT ATTCGGCTAT   | 7380 |    |
| GACTGGGCAC AACAGACAAT CGGCTGCTCT GATGCCGCCG TGTTCCGGCT GTCAGCGCAG  | 7440 |    |
| GGGCGCCCGG TTCTTTTTGT CAAGACCGAC CTGTCCGGTG CCCTGAATGA ACTGCAGGAC  | 7500 | 20 |
| GAGGCAGCGC GGCTATCGTG GCTGGCCACG ACGGGCGTTC CTTGCGCAGC TGTGCTCGAC  | 7560 |    |
| GTTGTCACTG AAGCGGGAAG GGA CTGGCTG CTATTGGGCG AAGTGCCGGG GCAGGATCTC | 7620 |    |
| CTGTCATCTC ACCTTGCTCC TGCCGAGAAA GTATCCATCA TGGCTGATGC AATGCGGCGG  | 7680 |    |
| CTGCATACGC TTGATCCGGC TACCTGCCCA TTCGACCACC AAGCGAAACA TCGCATCGAG  | 7740 |    |
| CGAGCACGTA CTCGGATGGA AGCGGTCTT GTCGATCAGG ATGATCTGGA CGAAGAGCAT   | 7800 |    |
| CAGGGGCTCG CGCCAGCCGA ACTGTTCCGC AGGCTCAAGG CGCGCATGCC CGACGGCGAG  | 7860 |    |
| GATCTCGTCG TGACCCATGG CGATGCCTGC TTGCCGAATA TCATGGTGGA AAATGGCCGC  | 7920 | 30 |
| TTTTCTGGAT TCATCGACTG TGGCCGGCTG GGTGTGGCGG ACCGCTATCA GGACATAGCG  | 7980 |    |
| TTGGCTACCC GTGATATTGC TGAAGAGCTT GGC GGCGAAT GGGCTGACCG CTCCTCGTG  | 8040 |    |
| CTTTACGGTA TCGCGCTCC CGATTGCGAG CGCATCGCCT TCTATCGCCT TCTTGACGAG   | 8100 |    |
| TTCTTCTGAG CGGGACTCTG GGGTTCGAAA TGACCGACCA AGCGACGCC AACCTGCCAT   | 8160 |    |
| CACGAGATTT CGATTCCACC GCCGCTTCT ATGAAAGGTT GGGCTTCGGA ATCGTTTTCC   | 8220 |    |
| GGGACGCCGG CTGGATGATC CTCCAGCGCG GGGATCTCAT GCTGGAGTTC TTCGCCCACC  | 8280 | 40 |
| CCA ACTTGT TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA  | 8340 |    |
| CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAACTC ATCAATGTAT   | 8400 |    |
| CTTATCATGT CTGTATACCG TCGACCTCTA GCTAGAGCTT GCGTAATCA TGGTCATAGC   | 8460 |    |
| TGTTTCCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA CAACATACGA GCCGGAAGCA  | 8520 |    |

|            |            |            |            |            |            |       |
|------------|------------|------------|------------|------------|------------|-------|
| TAAAGTGTA  | AGCCTGGGGT | GCCTAATGAG | TGAGCTAACT | CACATTAATT | GCGTTGCGCT | 8580  |
| CACTGCCCCG | TTTCCAGTCG | GGAAACCTGT | CGTGCCAGCT | GCATTAATGA | ATCGGCCAAC | 8640  |
| GCGCGGGGAG | AGGCGGTTTG | CGTATTGGGC | GCTCTTCCGC | TTCCTCGCTC | ACTGACTCGC | 8700  |
| TGCGCTCGGT | CGTTCGGCTG | CGGCGAGCGG | TATCAGCTCA | CTCAAAGGCG | GTAATACGGT | 8760  |
| TATCCACAGA | ATCAGGGGAT | AACGCAGGAA | AGAACATGTG | AGCAAAAGGC | CAGCAAAAGG | 8820  |
| CCAGGAACCG | TAAAAAGGCC | GCGTTGCTGG | CGTTTTTCCA | TAGGCTCCGC | CCCCCTGACG | 8880  |
| AGCATCACAA | AAATCGACGC | TCAAGTCAGA | GGTGCGGAAA | CCCGACAGGA | CTATAAAGAT | 8940  |
| ACCAGGCGTT | TCCCCCTGGA | AGCTCCCTCG | TGCGCTCTCC | TGTTCCGACC | CTGCCGCTTA | 9000  |
| CCGGATACCT | GTCCGCCTTT | CTCCCTTCGG | GAAGCGTGGC | GCTTTCTCAA | TGCTCACGCT | 9060  |
| GTAGGTATCT | CAGTTCGGTG | TAGGTCGTTT | GCTCCAAGCT | GGGCTGTGTG | CACGAACCCC | 9120  |
| CCGTTACAGC | CGACCGCTGC | GCCTTATCCG | GTAACATATC | TCTTGAGTCC | AACCCGGTAA | 9180  |
| GACACGACTT | ATCGCCACTG | GCAGCAGCCA | CTGGTAACAG | GATTAGCAGA | GCGAGGTATG | 9240  |
| TAGGCGGTGC | TACAGAGTTC | TTGAAGTGGT | GGCCTAACTA | CGGCTACACT | AGAAGGACAG | 9300  |
| TATTTGGTAT | CTGCGCTCTG | CTGAAGCCAG | TTACCTTCGG | AAAAAGAGTT | GGTAGCTCTT | 9360  |
| GATCCGGCAA | ACAAACCACC | GCTGGTAGCG | GTGGTTTTTT | TGTTTGCAAG | CAGCAGATTA | 9420  |
| CGCGCAGAAA | AAAAGGATCT | CAAGAAGATC | CTTTGATCTT | TTCTACGGGG | TCTGACGCTC | 9480  |
| AGTGGAACGA | AAACTCACGT | TAAGGGATTT | TGGTCATGAG | ATTATCAAAA | AGGATCTTCA | 9540  |
| CCTAGATCCT | TTTAAATTAA | AAATGAAGTT | TTAAATCAAT | CTAAAGTATA | TATGAGTAAA | 9600  |
| CTTGGTCTGA | CAGTTACCAA | TGCTTAATCA | GTGAGGCACC | TATCTCAGCG | ATCTGTCTAT | 9660  |
| TTCGTTCATC | CATAGTTGCC | TGACTCCCCG | TCGTGTAGAT | AACTACGATA | CGGGAGGGCT | 9720  |
| TACCATCTGG | CCCCAGTGCT | GCAATGATAC | CGCGAGACCC | ACGCTCACC  | GCTCCAGATT | 9780  |
| TATCAGCAAT | AAACCAGCCA | GCCGGAAGGG | CCGAGCGCAG | AAGTGGTCCT | GCAACTTTAT | 9840  |
| CCGCCTCCAT | CCAGTCTATT | AATTGTTGCC | GGGAAGCTAG | AGTAAGTAGT | TCGCCAGTTA | 9900  |
| ATAGTTTGCG | CAACGTTGTT | GCCATTGCTA | CAGGCATCGT | GGTGTACGCG | TCGTGTTTTG | 9960  |
| GTATGGCTTC | ATTCAGCTCC | GGTTCCCAAC | GATCAAGGCG | AGTTACATGA | TCCCCCATGT | 10020 |
| TGTGCAAAAA | AGCGGTTAGC | TCCTTCGGTC | CTCCGATCGT | TGTCAGAAGT | AAGTTGGCCG | 10080 |
| CAGTGTTATC | ACTCATGGTT | ATGGCAGCAC | TGCATAATTC | TCTTACTGTC | ATGCCATCCG | 10140 |
| TAAGATGCTT | TTCTGTGACT | GGTGAGTACT | CAACCAAGTC | ATTCTGAGAA | TAGTGTATGC | 10200 |
| GGCGACCGAG | TTGCTCTTGC | CCGGCGTCAA | TACGGGATAA | TACCGCGCCA | CATAGCAGAA | 10260 |

10

20

30

40

CTTTAAAAGT GCTCATCATT GGAAAACGTT CTCGGGGCG AAAACTCTCA AGGATCTTAC 10320  
 CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT 10380  
 TTACTTTCAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC GCAAAAAAGG 10440  
 GAATAAGGGC GACACGGAAG TGTGAATAC TCATACTCTT CCTTTTCAA TATTATTGAA 10500  
 GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA 10560  
 AACAAATAGG GGTTCGCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC 10610

10

- ( 2 ) 配列番号 : 1 7 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 2 4 塩基対  
 ( B ) 配列の型 : 核酸  
 ( C ) 鎖の数 : 一本鎖  
 ( D ) トポロジー : 直鎖状  
 ( ii ) 配列の種類 : D N A ( genomic )  
 ( iii ) ハイボセティカル : NO  
 ( iii ) アンチセンス : NO  
 ( xi ) 配列の記載 : 配列番号 : 1 7 :

20

TGTACACCGG ATCCGGCGCA CACC

24

- ( 2 ) 配列番号 : 1 8 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 3 5 塩基対  
 ( B ) 配列の型 : 核酸  
 ( C ) 鎖の数 : 一本鎖  
 ( D ) トポロジー : 直鎖状  
 ( ii ) 配列の種類 : D N A ( genomic )  
 ( iii ) ハイボセティカル : NO  
 ( iii ) アンチセンス : NO  
 ( xi ) 配列の記載 : 配列番号 : 1 8 :

30

CACAACGAGC TCAATTAATT AATTGCCACA TCCTC

35

G

- ( 2 ) 配列番号 : 1 9 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 4 アミノ酸  
 ( B ) 配列の型 : アミノ酸  
 ( D ) トポロジー : 直鎖状  
 ( ii ) 配列の種類 : ペプチド  
 ( v ) フラグメント型 : 中間部  
 ( xi ) 配列の記載 : 配列番号 : 1 9 :

40

Thr Leu Trp Thr

1

- ( 2 ) 配列番号 : 2 0 の情報 :  
 ( i ) 配列の特性 :  
 ( A ) 配列の長さ : 1 2 アミノ酸  
 ( B ) 配列の型 : アミノ酸  
 ( D ) トポロジー : 直鎖状

50

(ii) 配列の種類：ペプチド

(v) フラグメント型：中間部

(xi) 配列の記載：配列番号：20：

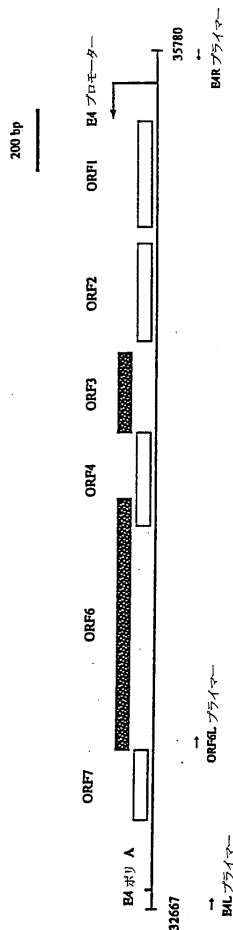
Pro Ser Ala Ser Ala Ser Ala Ser Ala Pro Gly Ser

1

5

10

【図1】



【図2】

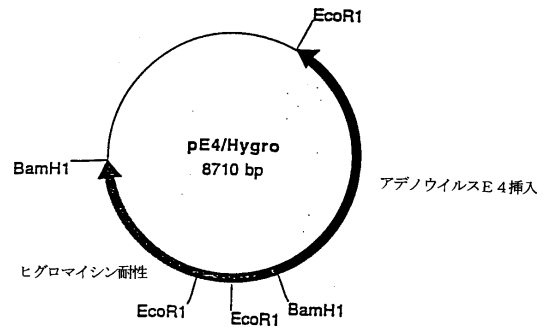
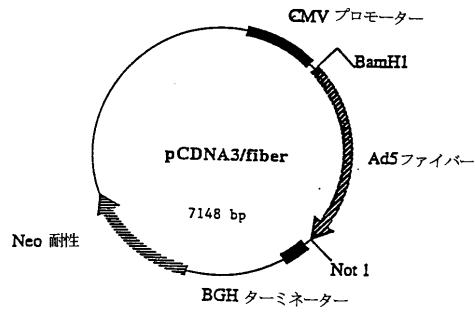


FIGURE 1

FIGURE 2

【図 3】



【図 4】

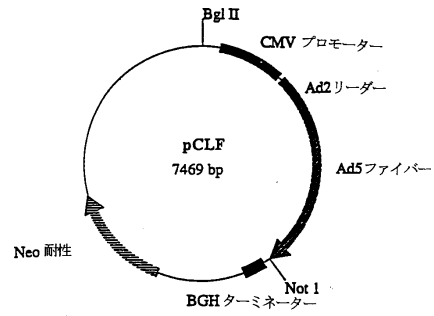


FIGURE 3

【図 5】

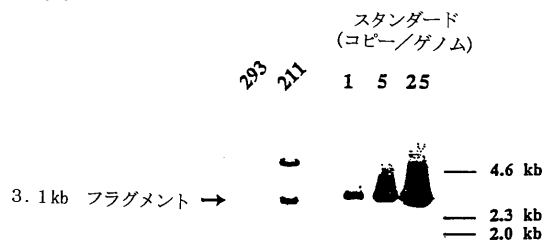


FIGURE 5

FIGURE 4

【図 6】

B

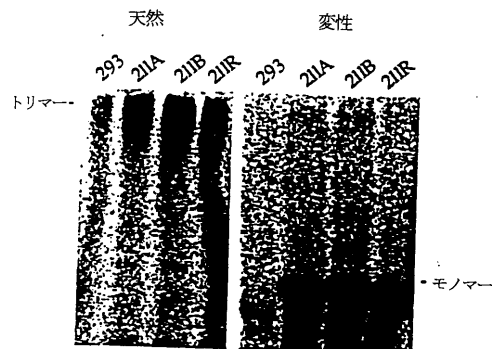
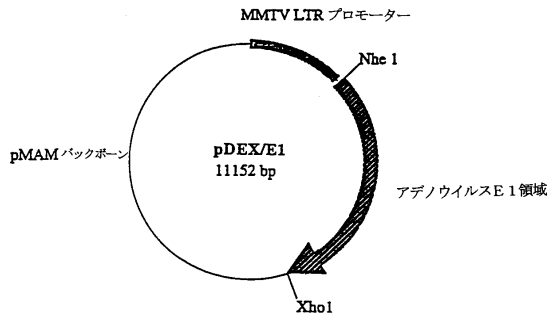


FIGURE 6

【図 7】



【図 8】

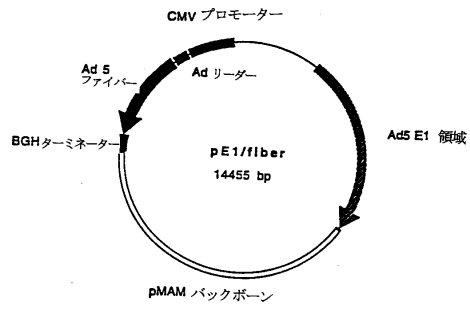


FIGURE 7

【図 9】

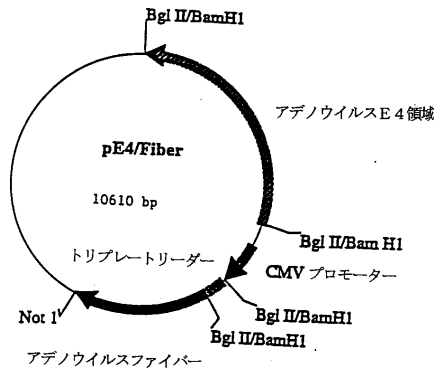


FIGURE 8

【図 10】

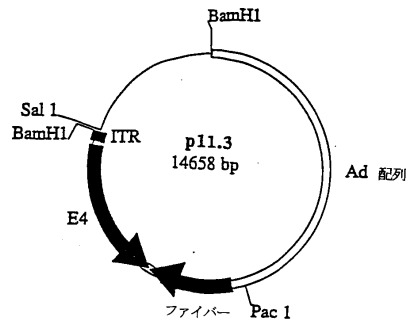
pΔE1Bβgal



FIGURE 9

FIGURE 10

【図 1 1】



【図 1 2】

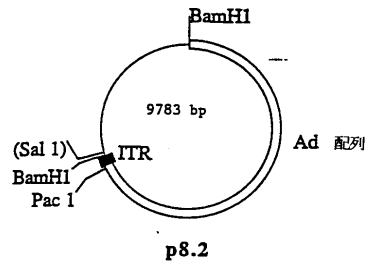


FIGURE 11

【図 1 3】

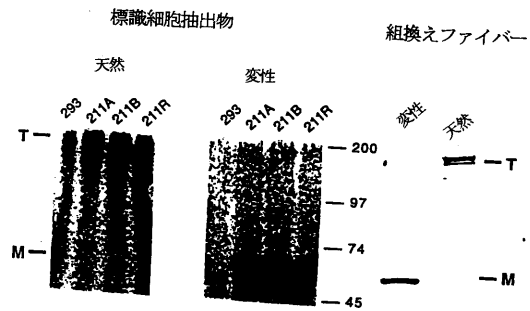


FIGURE 12

【図 1 4】

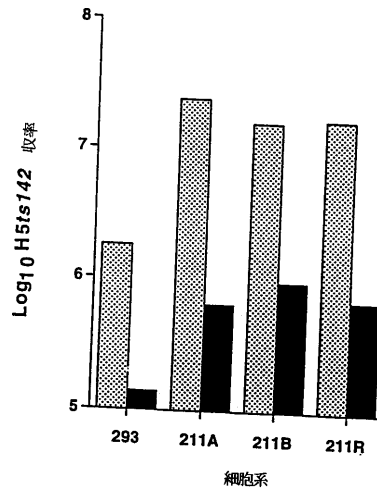


Figure 14

Figure 13

【 15 】

A

|          |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|----------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ad2/5    | Q | K | R | R | S | E | D | T | R | P | V | P | V | D | E | T | G | P | R | I | V | P | F | L | T | P | P | F | V | S | P | N | G | F | D | E | S | P |
| Ad3      | Q | K | R | R | S | E | D | T | R | P | V | P | V | D | E | T | G | P | R | I | V | P | F | L | T | P | P | F | V | S | P | N | G | F | D | E | S | P |
| Ad1827   | Q | K | R | R | S | E | D | T | R | P | V | P | V | D | E | T | G | P | R | I | V | P | F | L | T | P | P | F | V | S | P | N | G | F | D | E | S | P |
| Ad40 (1) | Q | K | R | R | S | E | D | T | R | P | V | P | V | D | E | T | G | P | R | I | V | P | F | L | T | P | P | F | V | S | P | N | G | F | D | E | S | P |
| Ad40 (2) | Q | K | R | R | S | E | D | T | R | P | V | P | V | D | E | T | G | P | R | I | V | P | F | L | T | P | P | F | V | S | P | N | G | F | D | E | S | P |

B

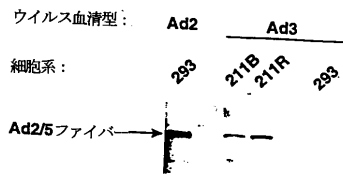


Figure 15

【 16 】

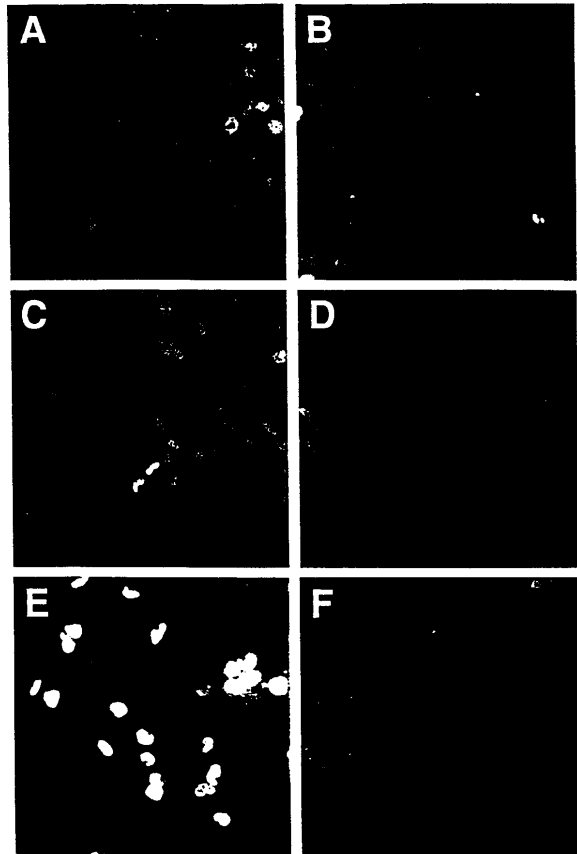


Figure 16



## フロントページの続き

(51)Int.Cl. F I  
A 6 1 K 48/00 (2006.01) A 6 1 K 48/00

(74)代理人 100067035

弁理士 岩崎 光隆

(72)発明者 ネメロー, グレン・アール

アメリカ合衆国 9 2 0 2 4 カリフォルニア州 エンシニタス、セロ・ストリート 4 6 2 番

(72)発明者 ボン・セガー、ダニエル・ジェイ

アメリカ合衆国 9 2 1 2 2 カリフォルニア州 サンディエゴ、ルイーダ・テラス 5 1 7 5 番、アパートメント 3 0

審査官 濱田 光浩

(56)参考文献 国際公開第 9 5 / 0 2 7 0 7 1 (WO, A 1)

Proc. Natl. Acad. Sci. USA, (1996), Vol. 93, p. 5731-5736

J. Virology, (1996), Vol. 70, No. 4, p. 2116-2123

(58)調査した分野(Int.Cl., D B 名)

C12N 5/10

C07K 14/075

C12N 15/09

CA/BIOSIS/MEDLINE/WPIDS(STN)

PubMed