

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4457546号
(P4457546)

(45) 発行日 平成22年4月28日(2010.4.28)

(24) 登録日 平成22年2月19日(2010.2.19)

(51) Int.Cl. F I
C 1 2 N 15/09 (2006.01) C 1 2 N 15/00 Z N A A

請求項の数 6 (全 17 頁)

(21) 出願番号	特願2001-530464 (P2001-530464)	(73) 特許権者	000003034 東亜合成株式会社 東京都港区西新橋1丁目14番1号
(86) (22) 出願日	平成12年10月12日(2000.10.12)	(74) 代理人	100091742 弁理士 小玉 秀男
(86) 国際出願番号	PCT/JP2000/007105	(74) 代理人	100108512 弁理士 村瀬 裕昭
(87) 国際公開番号	W02001/027259	(72) 発明者	吉田 徹彦 愛知県名古屋市港区船見町1番地の1 東 亜合成株式会社 名古屋総合研究所内
(87) 国際公開日	平成13年4月19日(2001.4.19)	審査官	小暮 道明
審査請求日	平成19年2月14日(2007.2.14)		
(31) 優先権主張番号	特願平11-291295		
(32) 優先日	平成11年10月13日(1999.10.13)		
(33) 優先権主張国	日本国(JP)		

最終頁に続く

(54) 【発明の名称】 人工プロモーター候補の選択方法

(57) 【特許請求の範囲】

【請求項1】

選択対象の人口プロモーターの塩基配列を人口プロモーターが接続される構造遺伝子の塩基配列の上流に結合し、

その結合された塩基配列から1塩基ずつずらしながら得られる3塩基から形成されるアミノ酸単位により仮想アミノ酸配列を作成し、

その作成された仮想アミノ酸配列に対して、「transfer free energy to lipophilic phase」のアミノ酸インデックス値を用いて、当該仮想アミノ酸配列における個々のアミノ酸のアミノ酸インデックス値からなるアミノ酸インデックス曲線を作成し、

その作成されたアミノ酸インデックス曲線が転写開始点近傍で正負が逆転するようなものであるとき、前記人工プロモーターを当該構造遺伝子のプロモーター候補として選択することを特徴とする人工プロモーター候補の選択方法。

【請求項2】

有効な人工プロモーター候補を選択するに際して、選択対象の人工プロモーターの塩基配列と構造遺伝子上流の塩基配列とを結合して塩基配列を得る工程、

当該塩基配列の先頭部分に存在する人工プロモーターの塩基配列から1塩基ずつずらしながら3塩基単位を抽出して仮想アミノ酸配列を作成し、「transfer free energy to lipophilic phase」のアミノ酸インデックス値を用いて、その仮想アミノ酸配列の個々のアミノ酸のインデックス値からなる第1の数列

10

20

を得る工程、

前記第1の数列をスムージング化することにより第2の数列を得る工程、
および、

前記第2の数列が転写開始点近傍で正負の反転をしている場合に、前記人工プロモーターを有効な人工プロモーター候補として選択する工程を有する人工プロモーター候補の選択方法。

【請求項3】

選択対象の人口プロモーターの塩基配列と人口プロモーターが接続される構造遺伝子の上流の塩基配列とを結合して得られた塩基配列から1塩基ずつずらしながら得られる3塩基から形成されるアミノ酸単位により作成した仮想アミノ酸配列に対して、「transfer free energy to lipophilic phase」のアミノ酸インデックス値を用いて、その仮想アミノ酸配列における個々のアミノ酸のアミノ酸インデックス値からなるアミノ酸インデックス曲線を得る手段と、

そのアミノ酸インデックス曲線取得手段で得られたアミノ酸インデックス曲線が転写開始点近傍で正負が逆転するようなものであるとき、前記人工プロモーターを当該構造遺伝子のプロモーター候補として選択するプロモーター候補選択手段とを備えた人工プロモーターの選択装置。

【請求項4】

有効な人工プロモーター候補を選択するに際して、選択対象の人工プロモーターの塩基配列と構造遺伝子の上流の塩基配列とを結合した塩基配列を得る塩基配列取得手段、

その塩基配列取得手段で得られた塩基配列の先頭部分に存在する人工プロモーターの塩基配列から1塩基ずつずらしながら3塩基単位で抽出して仮想アミノ酸配列を形成し、「transfer free energy to lipophilic phase」のアミノ酸インデックス値を用いて、その仮想アミノ酸配列の個々のアミノ酸のインデックス値からなる第1の数列を得る第1数列取得手段、

その第1数列取得手段で得られた第1の数列をスムージング化することにより第2の数列を得る第2数列取得手段、

および、

前記第2数列取得手段で得られた第2の数列が転写開始点近傍で正負の反転をしている場合に、前記人工プロモーターを有効な人工プロモーター候補として選択する選択手段とを有する人工プロモーター候補の選択装置。

【請求項5】

コンピュータに、選択対象の人口プロモーターの塩基配列と人口プロモーターが接続される構造遺伝子の上流の塩基配列とを結合した塩基配列を入力装置から得る工程、

入力装置から得た塩基配列から1塩基ずつずらしながら得られる3塩基から形成されるアミノ酸単位により仮想アミノ酸配列を作成する工程、

その作成された仮想アミノ酸配列に対して、「transfer free energy to lipophilic phase」のアミノ酸インデックス値を用いて、当該仮想アミノ酸配列における個々のアミノ酸のアミノ酸インデックス値からなるアミノ酸インデックス曲線を作成する工程、

その作成されたアミノ酸インデックス曲線が転写開始点近傍で正負が逆転するようなものであるとき、前記人工プロモーターを当該構造遺伝子のプロモーター候補として選択する工程、を実行させるためのプログラムを記憶したコンピュータ読み取り可能な記憶媒体

【請求項6】

コンピュータに、有効な人工プロモーター候補を選択するに際して、選択対象の人工プロモーターの塩基配列と構造遺伝子の上流の塩基配列と結合した塩基配列を入力装置から得る工程、

入力装置から得た塩基配列の先頭部分に存在する人工プロモーターの塩基配列から1塩基ずつずらしながら3塩基単位で抽出し、「transfer free energy

10

20

30

40

50

to lipophilic phase」のアミノ酸インデックス値を用いて、抽出した3塩基単位により形成されるアミノ酸配列の個々のアミノ酸のインデックス値からなる第1の数値を得る工程、

前記第1の数値をスムージング化することにより第2の数値を得る工程、
および

前記第2の数値が転写開始点近傍で正負の反転をしている場合に、前記人工プロモーターを有効な人工プロモーター候補として選択する工程、を実行させるためのプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

技術分野

本発明は、人工プロモーター候補の選択技術に係り、さらに詳しくは仮想的に作成したプロモーター候補の中から有効な人工プロモーター候補を選択する、人工プロモーター候補の選択技術に関する。

背景技術

生命科学の発展に伴い染色体の構造が分子レベルで解明されている。それによると染色体中に存在するDNA（デオキシリボ核酸）分子は、DNAの連続する一本の糸のようなものであることが判明している。このDNAには、アデニン（a）、グアニン（g）、シトシン（c）、チミン（t）という主として4種類の塩基があり、染色体中に存在するDNA分子はそれらの組み合わせからなり、糖、リン酸とが結合して構成されており、染色体中に存在するDNA分子は塩基配列が異なることにより種々の情報を有することになる。

そして、DNAを解析することにより、DNA分子を構成する一部分が遺伝情報を担う構造単位で遺伝形質を規定する因子、即ち遺伝子であることが判明しており、特にタンパク質やtRNA、rRNAなどの一次構造を規定している遺伝子が構造遺伝子として定義されている。遺伝子によるタンパク質合成のメカニズムについて図18に基づいて説明する。

図18に示すように遺伝子は、プロモーター領域と、転写領域（タンパク質のアミノ酸配列情報を有する構造遺伝子の領域）と、ターミネーター領域の3つの領域に区分けされる。プロモーター領域は、転写の開始をコントロールする役割を持ち、転写領域は実際に転写される領域であり、ターミネーター領域は転写の終了を制御する役割を持っている。なお、転写が開始される部位を転写開始点という。

このような構造を有する遺伝子においてタンパク質の合成は以下の手順で行われる。まず、RNAポリメラーゼ（DNA配列を転写する酵素）が、プロモーター領域の少し手前のDNA配列に結合し、プロモーター領域に向かって移動する。そして、RNAポリメラーゼが、プロモーター領域を越えて転写領域側に移動すると、転写領域のDNA配列に対応したメッセンジャーRNAを合成する。そして、RNAポリメラーゼがターミネーター領域にくるとDNA配列の転写を終了する。その後、このメッセンジャーRNAは、核から細胞質へ移動し、リボゾームと結合する。リボゾームは、メッセンジャーRNAに基づいてタンパク質の合成を行う。このように、遺伝子のDNA配列が転写され、その転写されたものに従って所定のタンパク質が合成される。

ここで、上述したタンパク質の合成を行う頻度は、RNAポリメラーゼがメッセンジャーRNAを転写する頻度に係り、メッセンジャーRNAが転写される頻度は遺伝子のプロモーター領域により制御される。

したがって、プロモーター領域のDNA配列を人工的に操作することにより、タンパク質合成頻度の高いプロモーターや、タンパク質合成頻度の遅いプロモーターを得ることが理論的に可能となる。

このため、近年このプロモーター領域のDNA配列を人工的に作り替えてタンパク質合成頻度を積極的に制御しようとする試みが盛んに行なわれるようになってきている。なぜなら、タンパク質の合成頻度を制御することができれば、例えば転写活性の高いプロモーターの利用によって特定のタンパク質を人工的に大量発現することが可能になるからである。さらに、人工的に特定のタンパク質を大量発現することが可能となれば、例えば遺伝子治療

10

20

30

40

50

において、がん細胞に十分なウイルスを感染させ（遺伝子治療においては、弱毒されたウイルスが正常遺伝子を細胞内に運搬するウイルスベクターとして使用される）、導入した正常遺伝子（強力な人工プロモーターを組み込んだもの）から、がんを抑えるのに十分な量のタンパク質を特異的に発現させるという分子レベルでの治療が可能になる。

しかしながら、現在のところ、このような転写活性の高い人工プロモーター、すなわち、その様な特性を発揮するDNA配列を有する人工プロモーターを設計する手法は確立されておらず、構造遺伝子（検定用遺伝子）の前にランダムに塩基を配列させて合成したもの（プロモーター候補）を結合して、検定用遺伝子が発現したかどうかで、プロモーター候補の有効性を見ることが行われている。

このような方法では、実験を行なう人工プロモーター候補の塩基配列をランダムに決定するため、無限とも考えられる塩基配列（4の階乗）のなかから、効果の高い人工プロモーターに適した塩基配列を見つけることは困難なことである。

そこで、本発明は上述した問題を解決するためになされたものであり、その目的は、実験を行なう前に、考えられる塩基配列の中から予めプロモーターとして有効に機能する可能性の高い塩基配列を選出することにより、実際に実験を行なう塩基配列の数を少なくすることができる人工プロモーター候補の選択方法を提供することを目的とする。

発明の開示

本発明者は、今までに知られているプロモーターについてその塩基配列に関する特徴を検討した結果、プロモーターと構造遺伝子を結合させた塩基配列から所定のパターンで抽出して得た仮想アミノ酸配列における個々のアミノ酸に基づくインデックス値による曲線を作成すると、その曲線は転写開始点近傍で正負の逆転を起こす傾向が非常に高いことを見出し、本発明を完成したのである。

すなわち、本発明の人工プロモーター候補の選択方法は、構造遺伝子の転写領域における転写開始点近傍の仮想アミノ酸配列から得られるアミノ酸インデックス曲線に対し、正負が逆転したアミノ酸インデックス曲線となる仮想アミノ酸配列から得られる塩基配列を有するDNAを人工プロモーター候補として選択することを特徴とする。

本発明の人工プロモーター候補の選択方法は、プロモーターと構造遺伝子を含む所与の塩基配列の転写開始点近傍における構造遺伝子の転写領域における転写開始点近傍の仮想アミノ酸配列から得られるアミノ酸インデックス曲線とプロモーターの塩基配列から得られる仮想アミノ酸配列のアミノ酸インデックス曲線が、転写開始点近傍で正負が逆転する傾向が極めて高いという新規な事実に基づくものであり、例えば、構造遺伝子の転写開始点近傍における仮想アミノ酸配列のアミノ酸のインデックス値から構造遺伝子側の曲線を求めておき、転写開始点近傍で正負が反転するようにプロモーター領域の仮想アミノ酸配列を求めて塩基配列を決定することもでき、また、例えば、予め任意に設定された塩基配列を構造遺伝子の上流に結合し、その結合した塩基配列から求められる仮想アミノ酸配列のインデックス曲線が転写開始点近傍における構造遺伝子側の曲線と正負が反転しているかどうかで、任意に設定した仮想アミノ酸配列の有効性を判定することもできる。

また、本発明の他の人工プロモーター候補の選択方法は、選択対象の人工プロモーターの塩基配列を人工プロモーターが接続される構造遺伝子の塩基配列の上流に結合し、その結合した塩基配列から得られる仮想アミノ酸配列に基づくアミノ酸インデックス曲線が転写開始点近傍で正負が逆転するようなものであるとき、前記人工プロモーターを当該構造遺伝子のプロモーター候補として選択することを特徴とする。

すなわち、塩基配列の判明している構造遺伝子に対する有効な人工プロモーター候補を選択する際には、任意に設定された塩基配列を前記構造遺伝子の上流に結合し、その結合した塩基配列の少なくとも転写開始点近傍を含む所定の領域における塩基配列から所定のパターンで抽出して得た仮想アミノ酸配列に基づくインデックス曲線が転写開始点近傍で正負が逆転するようなものであるとき、前記任意に設定された塩基配列を人工プロモーター候補として選択するというものである。

なお、所与の塩基配列から仮想アミノ酸配列を抽出する方法及びその仮想アミノ酸配列に基づくインデックス曲線の表現方法は、転写開始点における変化を見やすくする目的で適

10

20

30

40

50

宜設定することができる。

例えば、所与の塩基配列から仮想アミノ酸配列を抽出する一つの方法としては、選択対象である人工プロモーターの塩基配列から得られる仮想アミノ酸配列が、当該塩基配列から1塩基ずつずらしながら得られる3塩基から形成されるアミノ酸単位により作成されていても良い。

ただし、仮想アミノ酸配列を抽出する方法としては、上記した塩基配列から1塩基ずつずらしながら3塩基単位を抽出する方法以外にも様々な方法で抽出ことができ、例えば、塩基配列の中から連続して隣合う3塩基を抽出したり、1塩基ずつ飛ばして3塩基を抽出したり、または2塩基ずつ飛ばして3塩基を抽出するようにしても良い。

また、本発明の他の人工プロモーター候補の選択方法は、有効な人工プロモーター候補を選択するに際して、選択対象の人工プロモーターの塩基配列と構造遺伝子の上流の塩基配列とを結合して塩基配列を得る工程、当該塩基配列の先頭部分に存在する人工プロモーターの塩基配列から1塩基ずつずらしながら3塩基単位を抽出し、それにより形成される仮想アミノ酸配列における個々のアミノ酸のインデックス値からなる第1の数列を得る工程、前記第1の数列をスムージング化することにより第2の数列を得る工程、および、前記第2の数列が転写開始点近傍で正負の反転をしている場合に、前記人工プロモーターを有効な人工プロモーター候補として選択する工程を有する。

この人工プロモーター候補を選択する方法では、まず仮想的に作成した人工プロモーターの塩基配列を構造遺伝子の塩基配列の上流に結合した塩基配列を作成する。具体的には、転写対象とする構造遺伝子の塩基配列の上流に人工プロモーターの塩基配列を結合することにより一定の塩基配列が得られる。

既に説明したように人工プロモーターとして考えられる塩基配列の数は非常に多くの数となるため、この作業はコンピュータを利用して行なうことが好ましい。コンピュータにより自動的に人工プロモーターの塩基配列を作成することにより効率的に作業を進めることができる。ただし、ある程度効果がありそうな人工プロモーターの塩基配列が予測されるときには、研究者等が手作業にて塩基配列を設計しても良い。例えば、既知のプロモーターの塩基配列の一部を修正して、人工プロモーターの塩基配列を設計する場合である。

次に、この塩基配列の先頭部分に存在する人工プロモーターの塩基配列から一定の方式で3塩基単位ずつ抽出して仮想アミノ酸配列を作成し、それらの個々のアミノ酸のインデックス値を求めてアミノ酸インデックス値からなる第1の数列を作成する。

ここで、アミノ酸インデックス値（例えば、transfer free energy to lipophilic phase von Heijne-Blomberg 1979等）とは、アミノ酸の物性を様々な角度から計測し、数値化したものであり、その種類は現在434種あり、インターネット上で公開されている。このようにアミノ酸インデックス値に変換するのは、塩基配列のみの解析からではわからない何らかの高次の情報が事象の背後に潜んでいると推察されるためである。

次に、この第1の数列をスムージング化することにより第2の数列を作成する。スムージング化は、例えばその数値の後ろに続く所定個数（例えば、3～6個）の数値を足し合わせその平均を採る方法、又はその数値の前に続く所定個数（例えば、3～6個）の数値を足し合わせその平均を採る方法、さらにその数値の前後に連続する所定の個数（3～6個）の数値を足し合わせその平均を採る方法等を採用することができる。また、上述した各方法においては、足し合わせた数値の和を足し合わせた数値の数で割ることにより平均を求めたが、足し合わせた数値をそのまま採用しても良い。このようにしても、後述するグラフ化の段階で、グラフの縮尺を調整することにより、そのグラフは同一のパターンを有するためである。さらに、スムージング化は、上述した方法に限られず、データのスムージング化に関する種々の数学的手法をも使用することができ、例えば移動平均法等により行うこともできる。すなわち、後述するスムージング化された数列をグラフ化した際に、そのグラフのパターン（変動の様子）が認識できる程度にスムージング化できればどのような手法であっても良い。

そして、この第2の数列が転写開始点近傍で正負の反転をしているか否かを判定し、正負

10

20

30

40

50

の反転をしている場合に、前記人工プロモーターを有効な人工プロモーター候補として選出する。

なお、第2の数列が転写開始点近傍で正負の反転をしているか否かを判定する際、数列とそれに対応する数値を図式化してインデックス曲線を作成すると、転写開始点近傍の領域でインデックス曲線の正負が逆転（反転）する傾向が明瞭になり判定が容易になる。

上述した本発明の人工プロモーター候補の選択方法によれば、数多くの仮想人工プロモーターの中からプロモーターとして機能する可能性の高い人工プロモーター候補を選出することができ、これにより、実際に効果を確認するために実験しなければならないプロモーターの数を少なくすることができる。

すなわち、従来のプロモーターの設計方法では、図1の下段に示すようにプロモーター候補をランダムな塩基配列の基に合成し（S1）、そのプロモーター候補を構造遺伝子（検定用遺伝子）と結合し（S2）、その結合したものを細胞内に導入し、その遺伝子発現を確認することによりそのプロモーター候補の有用性を決定していた。一方、本発明によれば、図1の上段に示すように、ステップS1で上述したプロモーター候補の選出をした後に、プロモーター候補の合成（S2）・構造遺伝子（検定用遺伝子）の結合（S3）・実験（S4）というステップで行うことができ、実際に合成し実験するプロモーター候補の数を少なくすることができる。

発明を実施するための最良の形態

以下、本発明の人工プロモーター候補の選択方法を具現化した実施の形態について説明する。図2に人工プロモーター候補を選択するための手順を示す。

図2に示すように本発明に係る人工プロモーター候補選択方法では、まず仮想人工プロモーターの塩基配列と構造遺伝子（検定遺伝子）の塩基配列とを結合した塩基配列を作成する（S1）。

図3に基づいてさらに具体的に説明する。図3に示すように、仮想的に作成した人工プロモーターの塩基配列を $g a t c t \cdot \cdot c a$ とし、構造遺伝子の塩基配列を $c c c g t c c a g \cdot \cdot$ とする。このステップS1の工程では、構造遺伝子の塩基配列 $c c c g t c c a g \cdot \cdot$ の前に、人工プロモーターの塩基配列 $g a t c t \cdot \cdot c a$ を結合し、塩基配列 $g a t c t \cdot \cdot c a c c c g t c c a g \cdot \cdot$ を作成する。

次に、この塩基配列の先頭の塩基から所定のパターンで3塩基単位ずつ抽出して得られる仮想アミノ酸配列から、それらを構成するアミノ酸のインデックス値からなる第1の数列を作成する（S2）。

塩基配列から3塩基単位を連続して抽出する方法を例に説明する。まずパターン1は上述した塩基配列 $g a t c t \cdot \cdot$ の先頭から1塩基ずつずらしながら3塩基単位を連続して抽出する方法で、1番目の3塩基は $g a t$ 、2番目の3塩基は $a t c$ 、3番目の3塩基は $t c t$ 、 $\cdot \cdot$ というように抽出する方法である。

ここで、図4（遺伝コード表）に示すように、 $g a t$ という3つの塩基は、アスパラギン酸（D：以下、アミノ酸をアミノ酸の一字記号で表す。）に対応し、 $a t c$ はイソロイシン（I）に対応し、同様に $t c t$ はセリン（S）に対応する。したがって、塩基配列から1塩基ずつずらしながら3塩基単位で抽出したものは、 $D, I, S, L, S \cdot \cdot$ という仮想アミノ酸配列で表すことができる。

仮想アミノ酸配列を作成したら、次にこの仮想アミノ酸配列を形成する個々のアミノ酸をそれぞれのインデックス値に変換した数列を作成する。各アミノ酸（20種類）のアミノ酸インデックス値の一例として von Heijne - Blomberg の *transfer free energy to lipophilic phase* を、図5に例示した。したがって、上記のように作成したアミノ酸配列から、 $23.22, -18.32, -1.54, -17.79, -1.54, \dots$ という数列（第1数列）が形成される。これが、請求項にいう第1の数列に相当する。

ここで、アミノ酸配列において終止コドン（X）が存在する場合には、アミノ酸に対応せずアミノ酸インデックス値が存在しないため、本実施の形態では、アミノ酸インデックス値を0として第1数列を作成した。終止コドン（X）を取り除いてしまうと、結果的に塩

10

20

30

40

50

基配列からある塩基を取り除くこととなり、アミノ酸インデックス値が変わってしまうからである。

次に、この第1数列(23.22, -18.32, -1.54, -17.79, -1.54, ...)のスムージング化を行う(S3)。ここで、第1数列のスムージング化を行うのは、第1数列のアミノ酸インデックス値は、隣接する数値の変動が激しいためアミノ酸インデックス値の変化の様子(パターン)が認識し難いためである。したがって、スムージング化は、スムージング化された数値をグラフ化した際に、そのグラフがある一定のパターンを有することが判別できる程度に行えば良い。

スムージング化の手法としては、データ処理の分野で既知の様々な手法を使用することができるが、ここでは、上述した第1数列の数値を、その数値から後ろにある所定の個数(3個)の数値の平均値に置き換えることによりスムージング化した数列を得る場合を例に説明する。

第1数列((23.22, -18.32, -1.54, -17.79, -1.54, ...))最初の3つの数値の平均は $\{23.22 + (-18.32) + (-1.54)\} / 3 = 1.12$ であり、次の3つの数値の平均は $\{-18.32 + (-1.54) + (-17.79)\} / 3 = -12.55$ であり、その次の3つの数値の平均は $\{(-1.54) + (-17.79) + (-1.54)\} / 3 = -6.96$ である。これ以降も同様の手順により平均値を求めることができる。

このように第1数列を所定個数毎に平均値化し置換することにより、第1数列をスムージング化した第2数列(1.12, -12.55, -6.96, ...)を得る。これが、請求項の第2の数列に相当する。

次いで、第2数列が転写開始点近傍で正負の反転をしているか否かを判定し、正負の反転をしている場合に、前記人工プロモーターを有効な人工プロモーター候補として選出する。好ましい判定方法は、この第2数列をグラフ化することによりインデックス曲線を得て、そのインデックス曲線のパターンにより、その人工プロモーターを評価する方法である(S4)。

即ち、上述した第2数列(1.12, -12.55, -6.96, ...)に基づいて、横軸に数列の番号、縦軸にスムージング化されたアミノ酸インデックス値をプロットすることによりインデックス曲線を得るのである。図6に転写開始点近傍におけるインデックス曲線の例を概念的に示している。

プロモーターとして機能する可能性が高い人工プロモーター候補を選出する際には、転写開始点近傍の仮想アミノ酸配列のインデックス曲線パターンで判定することができる。

すなわち、既に説明したように、既知のプロモーターのインデックス曲線は転写開始点近傍において、後述する図15に認められる様にその正負が反転している傾向が極めて高い。したがって、仮想人工プロモーターの塩基配列と構造遺伝子の塩基配列を結合してできる仮想アミノ酸配列のインデックス曲線を作成した際に、転写開始点近傍でインデックス曲線の正負が反転している場合、その塩基配列を有する仮想人工プロモーターは、当該構造遺伝子に有効な人工プロモーター候補として選出することができる。

以上説明した人工プロモーター候補の選択方法が有効であることを塩基配列既知の遺伝子(プロモーター領域部位)を用いて以下に説明する。

既知の遺伝子(プロモーター領域部位)として、J01567 Plasmid Colicin E1 (from E. coli) strong promoter region DNA、X87994 C. xyli DNA for strong promoter (569 bp)、EPD31005 (+) Ph EPSP synthase P2+ 及び EPD35038 (+) Le LAT52; range -499 to 100 を選択した。

J01567とX87994の塩基配列についてはDDBJ(DNA Data Bank of Japan)に登録されているデータを利用し、他の塩基配列についてはEPD(Eukaryotic Promoter Database)に登録されているデータを利用した(但し、X87994については481番目以降の配列データを用いた。

10

20

30

40

50

)。

それらの塩基配列は図7～図10に示す通りである。すなわち、J01567は、図7に示す塩基配列を有し、その転写開始点は84番目である。また、X87994は、図8に示す塩基配列を有し、その転写開始点は542番目である(但し、後述の図15では481番目以降が示されているので、図15(b)において転写開始点は62番目である)。また、EPD31005は、図9に示す塩基配列を有し、その転写開始点は140番目である。さらに、EPD35038は、図10に示す塩基配列を有し、その転写開始点は140番目である。

上記の塩基配列からインデックス曲線を上記記載の手順で作成した。具体的な手順を図7に示される塩基配列を有するJ01567を例に、図11～図14を用いて説明する。なお、以下に説明するインデックス曲線の作成方法は単なる例示であって、本発明を限定するものではない。

(1) 図11に示したJ01567の塩基配列の先頭から、連続する3個の塩基を単位として抽出する操作を1塩基ずつずらしながら最後の塩基まで繰り返して3個の塩基を単位とする図12に示される配列を作成する。

(2) 上記(1)で得られた配列(図12)から、3個の塩基からなる単位をアミノ酸に変換して仮想アミノ酸配列を作成する。

(3) 上記(2)で得られた配列における各アミノ酸に対して疎水性アミノ酸インデックス値を対応させたアミノ酸インデックス値からなる図13に示される数列を作成する。

(4) 上記(3)で得られた数列(図13)において各数値の前方5個の数値を合計した値を対応させた図14で示される数列を作成する。

(5) 上記(4)で得られた数列(図14)を塩基の先頭からの位置を横軸にしてその位置に対応するアミノ酸インデックス値を縦軸に表わすことによりインデックス曲線を得る。

作成したインデックス曲線を図15に示す(但し、これらの図は正負の逆転を見やすくするために棒グラフ形式で示してある)。

図15には上記4種の遺伝子の塩基配列から求めたインデックス曲線をそれぞれ示すが、いずれも転写開始点近傍で正負が反転していることを示している。

以上のことから明らかな様に、仮想人工プロモーターの塩基配列と構造遺伝子の塩基配列を結合してできる仮想アミノ酸配列のインデックス曲線を作成した際に、転写開始点近傍でインデックス曲線の正負が反転している場合、その塩基配列を有する仮想人工プロモーターを当該構造遺伝子の有効な人工プロモーター候補として選出することができる。そして、選出された人工プロモーター候補については、実際にそのDNAを合成し、構造遺伝子と結合して実験を行い最終的な評価を行うこととなる。

以上詳述したように、上述した手順により人工プロモーター候補を選出すれば、実際に合成して実験を行う人工プロモーター候補の数を格段に減らすことができ、また、有効な人工プロモーター候補延いてはその塩基配列を効率良く設計することが可能となる。

また、有効な人工プロモーター候補の塩基配列を設計する場合、判定すべき領域を転写開始点近傍の領域のインデックス曲線に限定できるため、その判定を容易に行うことができる。

本発明の方法により、転写活性の高い人工プロモーター候補を効率良く設計することができれば、例えば遺伝子治療の分野において有用なプロモーターを効率的に設計することができる。例えば、ガン細胞に十分なウイルス(正常な遺伝子を運ぶベクターウイルス)を感染させ、導入した正常遺伝子からガンを抑えるのに十分な量のタンパク質を特異的に発現させる治療に応用することができる。これは、正常遺伝子の特異的かつ強力に発現させるためには、人工的に強いプロモーターを設計することが必要となるからである。

上述した本発明に係る人工プロモーター候補の選出を、好適に実施することができる人工プロモーター候補選出装置の一実施の形態について説明する。図16に本実施の形態に係る人工プロモーター候補選出装置のハードウェア構成図を示す。

本実施の形態に係る人工プロモーター候補選出装置は、図16に示すように人工プロモ-

10

20

30

40

50

ター候補選出装置全体を統括的に制御するプログラムされた主制御部（以下、主制御部という）10と、主制御部10に接続され各種データを記憶する記憶装置20と、入出力制御部14を介して主制御部10に接続されるキーボードやマウス等のポインティングデバイスからなる入力装置11と、インデックス曲線等の画像を表示するディスプレイ等の表示装置12と、選出されたプロモーター候補の塩基配列をプリントアウトするプリンター等の出力装置13を備える。

主制御部10は、OS (Operating System) 等の制御プログラム、入力された塩基配列からアミノ酸インデックス値に変換した第1の数値を作成するための変換プログラム、第1の数値をスムージング化するためのプログラム、スムージング化されたデータを表示装置12に表示するための画像処理プログラムや所要データを格納するための内部メモリを有している。

10

記憶装置20は、ハードディスクやフレキシブルディスク、あるいは光ディスク等の記憶手段であり、塩基配列ファイル21と、塩基-アミノ酸ファイル22、アミノ酸インデックス値ファイル23と、選出された人工プロモーター候補の塩基配列を記憶する人工プロモーター候補ファイル24が格納されている。

塩基配列ファイル21は、仮想的に作成した人工プロモーター候補の塩基配列と構造遺伝子の塩基配列とを結合した塩基配列を記憶するファイルである。すなわち、入力装置11のキーボード等からオペレータにより入力された塩基配列のデータが、記憶装置22の中に格納される。

塩基-アミノ酸ファイル22は、3塩基を組み合わせることにより、どのようなアミノ酸が合成されるかのデータを記憶したファイルであり、具体的には図4に示すようなデータが記憶されている。

20

アミノ酸インデックス値ファイル23は、アラニン、アルギニン等の各種アミノ酸のアミノ酸インデックス値を記憶するファイルである。具体的には、図5に示すようなデータにより、アミノ酸インデックス値ファイル23が構成される。

人工プロモーター候補ファイル24は、インデックス曲線を判定した結果、人工プロモーターとして有効に機能する可能性が高いものとして選出された人工プロモーターの塩基配列を記憶するファイルである。

次に、本実施の形態の人工プロモーター候補選出装置を使用して入力された塩基配列から人工プロモーター候補を選出する手順を図17に基づいて説明する。図17は人工プロモーター候補を選出する時の手順を説明するためのフローチャートである。

30

まず、オペレータは、入力装置11から、仮想プロモーターの塩基配列と構造遺伝子の塩基配列を結合した仮想塩基配列を入力する(S1)。なお、オペレータにより入力された塩基は表示装置12上に、例えば、a, g, c, t等と表示され、オペレータは入力内容を確認しながら入力作業を行うことができる。この塩基配列は、入出力制御部14及び主制御部10を介して記憶装置20内の塩基配列ファイル21の中に格納される。ここでは入力された塩基配列が1つ(人工プロモーター候補が一つ)である場合を例に説明する。次に、オペレータは、入力装置11から、ステップS1で入力した塩基配列から3塩基単位を抽出する方式を入力する(S2)。本実施の形態のプロモーター候補選出装置では、図3に示すように、連続して3塩基を抽出する方式1と、1塩基ずつ飛ばして3塩基を抽出する方式2と、2塩基ずつ飛ばして3塩基を抽出する方式3が選択できるようになっている(図3参照)。

40

さらに、オペレータは、入力装置11から、スムージングパターンを入力する(S3)。本実施の形態の人工プロモーター候補選出装置においては、平均値を取るための数値の数(3~10個)を入力するようになっている。

ステップS1~S3により入力事項が入力されると、主制御部10は、ステップS2で入力された抽出パターンに従って、ステップS1で入力された塩基配列から3塩基を抽出し、その塩基配列から形成される仮想アミノ酸配列をアミノ酸インデックス値に変換することにより第1の数値を作成する(S4)。すなわち、塩基配列ファイル21に格納されている塩基配列から3塩基単位ずつ抽出し、その塩基配列に対応するアミノ酸を塩基-アミ

50

ノ酸ファイル 2 2 のデータに基づいて判断し、アミノ酸インデックス値ファイル 2 2 のデータを基に、第 1 の数列に変換する。作成された第 1 の数列は、主制御部 1 0 内の内部メモリに格納される。

次に、主制御部 1 0 は、内部メモリに格納している第 1 の数列に対してスムージングを行う (S 5)。そして、主制御部 1 0 は、スムージング化された数値に基づいて表示装置 1 2 上にインデックス曲線を表示する (S 6)。本実施の形態のスムージングは、 S 3 で入力された個数 (例えば、 3 個) の数値の平均値を取ることにより行っている。具体的な手順に付いては、既に説明した通りである。

そして、表示装置 1 2 上に表示されたインデックス曲線に基づいて、オペレータが、その仮想塩基配列を有するプロモーターの有効性を判定する (S 7)。具体的には、入力された塩基配列のインデックス曲線を表示装置 1 2 上に表示し、オペレータが表示されているインデックス曲線を見て、転写開始点近傍で正負が反転しているか否かを判断する。

S 7 のステップで、人工プロモーターとして機能する可能性が高いと判断された場合には、その塩基配列を有するものを人工プロモーター候補として選出される (S 8)。選出された人工プロモーター候補の塩基配列は、記憶装置 2 0 内の人工プロモーター候補ファイル 2 4 内に格納され、オペレータの要求により随時、表示装置 1 2 上に表示、又は出力装置 (プリンター) に出力されるようになっている。

以上、詳述したように、本実施形態に係る人工プロモーター候補選出装置では、オペレータにより仮想プロモーター候補の塩基配列と目的とする構造遺伝子の塩基配列とを結合した塩基配列を入力するだけで、その仮想プロモーター候補の塩基配列が人工プロモーターとして機能する可能性を有するか否かを判断できる。したがって、実際に合成して実験する人工プロモーター候補をある程度絞り込むことができるので、これにより効率的に人工プロモーターの設計を行うことができる。

なお、上述した実施の形態では、塩基配列を一つ (人工プロモーター候補が一つ) だけ入力する例について説明したが、複数の塩基配列 (人工プロモーター候補が複数) を一度に入力して記憶装置 2 0 の塩基配列ファイル 2 1 に格納するようにしても良い。この場合には、塩基配列ファイル 2 1 内に格納されている各塩基配列について、図 1 7 のステップ S 2 ~ S 8 を繰り返すこととなる。

また、上述した実施の形態のように塩基配列をオペレータにより入力するのではなく、人工プロモーター候補選出装置への入力事項は、仮想プロモーターを構成する塩基の数と構造遺伝子の塩基配列のみとし、主制御部 1 0 により考え得る塩基配列の全てを作成するようにしても良い。

すなわち、DNA は a , g , c , t の四種類の塩基により構成されるため、プロモーター領域の塩基の数 (n) が決まれば、考え得る塩基配列 (4^n 個) を全て作成することができ、自動的に作成されたものと構造遺伝子の塩基配列を結合すれば良い。このようないわゆる塩基配列自動作成手段を人工プロモーター候補選出装置内に設ければ、効率的に人工プロモーター候補を作成し、その有効性を判断することができるため、従来の人工プロモーターをランダムに作成したときと比較して漏れのない塩基配列の評価を行うことができる。

また、上述した実施の形態では、インデックス曲線の判定をオペレータ (研究者) が、表示装置 1 2 上に表示されたインデックス曲線を見て行うようにしたが、このような形態をとることなく、コンピュータによるパターン認識技術を利用して、主制御部 1 0 により自動的に行うようにしても良い。また、塩基配列を主制御部 1 0 により自動的に作成する塩基配列自動作成手段と組み合わせ、主制御部 1 0 により自動的にインデックス曲線の判定を行うことにより、塩基配列の数が増えても、全ての塩基配列について自動的に評価することができる。

以上、本発明の実施の形態について詳細に説明したが、これらは例示に過ぎず、本発明は当業者の知識に基づいて種々の変更、改良を施した形態で実施することができる。

【配列表】

10

20

30

40

SEQUENCE LISTING

<110> TOAGOSEI CO.,LTD.
Yoshida, Tetsuhiko

<120> A Method for Selection of a Prospective Artificial
Promoter

<130> K00-208

<140>

<141>

<150> JP P1999-291295

<151> 1999-10-13

<160> 4

<170> PatentIn Ver. 2.1

<210> 1

<211> 152

<212> DNA

<213> Plasmid Colisin E1(from E.coli)

<400> 1

```
ccggattagc agagcgatga tggcacaac ggtgotacag agttottgaa gtagtggecc 60
gactacgget aactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 120
ttcggaaaaa gagttggtag ctcttgatcc gg 152
```

<210> 2

<211> 569

<212> DNA

<213> Clavibacter xyli

<400> 2

```
gatctcctcg gcgaacacac ccgtccagct cggcttggac gggctgcega tctcaccgga 60
gagctcggtc gtggcgctga ggtgggcrag tgtgcccggc cggggtcggg gtagtcgtcc 120
gcggggacgt agatcgcgtc ggtttccgtg tccgcacggg agaatccatt gtctgcatta 180
ccgctgtgta acgttttkcg agcctcgttt tgcattcatt gcgcacggta cctgagggtc 240
agtcttatcc cgcgcgggac caggcaactct cgcacgcccg tgcsgcgcgc aacacttcgc 300
acaggaggaa cattgcacat caagggaga aaggctcggc tcgcggtcag cgcgctgcgc 360
ggtgtcgsc tctcgctctg tcagcctgca ccacgtccgg caccggtggt tcgtccgctg 420
ctaaaggggg catggtgaca gtggcggctg tcaacgacct cacctcactg aactcgcaga 480
ccccgcaggg caacctggac accaacggcc aggtcggcta cctgaagttc ctacggaacc 540
ggtttccagt acatcgacaa caactacaa 569
```

10

20

30

40

<210> 3
 <211> 175
 <212> DNA
 <213> Petunia Hybrida

<400> 3
 gagaacacag ctggaatddd ttacaaaggt agttggtgaa gctagtcagc gaateccatt 60
 accttccact ctacctaaccc cccttcacca acaacaaatt tctgtaattd aaaaactagc 120
 caaaaaagaa ctctcttdtta caaagagoca aagactcaat ctttacttdc aagaa 175

10

<210> 4
 <211> 239
 <212> DNA
 <213> Lycopersicon esculentum(tomato)

<400> 4
 tcaattgcta caatcacttc attattaatt ttaattaata tatgtggtta tatatgaaac 60
 tgtttagagaa ataatagctc caccatattd ttdtctcaat ttattdttdcac tataaaaagg 120
 ctattdtcatt ataatcaaaa caagacacac acaaaagagaa ggagcaataa aataaaaagta 180
 aacaacaatt tgtgtgttdta aaaaaaaaaa aaaagtacac acaccaaaaa aaaaattdc 239

20

【図面の簡単な説明】

- 図 1 は、本発明の効果の説明するための図面である。
 図 2 は、本発明の人工プロモーター候補を選出するための手順を説明するためのフローチャートである。
 図 3 は、塩基配列からアミノ酸インデックス配列に変換する手順を説明するための図面である。
 図 4 は、遺伝コード表を示す図面である。
 図 5 は、アミノ酸インデックス値の一例を示す図面である。
 図 6 は、インデックス曲線を概念的に説明するための図面である。
 図 7 は、遺伝子 J 0 1 5 6 7 のプロモータ領域部位における塩基配列を示す図面である。
 図 8 は、遺伝子 X 8 7 9 9 4 のプロモータ領域部位における塩基配列を示す図面である。
 図 9 は、遺伝子 E P D 3 1 0 0 5 のプロモータ領域部位における塩基配列を示す図面である。
 図 1 0 は、遺伝子 E P D 3 5 0 3 8 のプロモータ領域における塩基配列を示す図面である。
 図 1 1 は、遺伝子 J 0 1 5 6 7 のプロモータ領域部位における塩基配列を上流から順に 1 塩基ずつ示した図面である。
 図 1 2 は、図 1 1 に示す塩基配列から 3 塩基を 1 単位として抽出することで作成した配列を示す図面である。
 図 1 3 は、図 1 2 に示す配列の各 3 塩基をアミノ酸インデックス値に変換した数列を示す図面である。
 図 1 4 は、図 1 3 に示す数列をスムージング化した数列を示す図面である。
 図 1 5 は、既知のプロモーターのインデックス曲線の一例を示す図面である。
 図 1 6 は、本発明の一実施の形態に係る人工プロモーター候補選出装置のハードウェア構成図を示す図面である。
 図 1 7 は、図 1 6 に示す人工プロモーター候補選出装置において、人工プロモーター候補を選出する手順を示したフローチャートである。
 図 1 8 は、DNA の転写からタンパク質合成までを説明するための図面である。

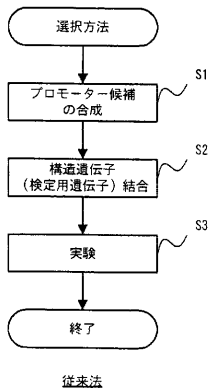
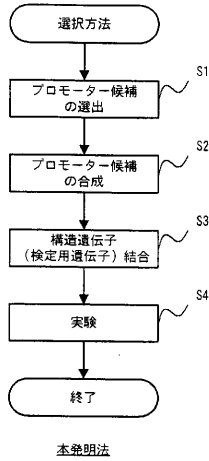
30

40

50

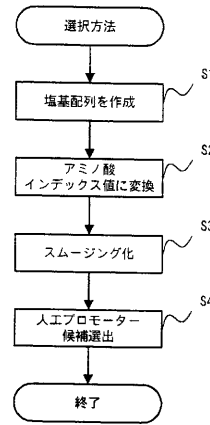
【図1】

FIG.1



【図2】

FIG.2



【図3】

FIG.3

塩基配列(例)

gatctcctc...ca/cccggtccag...
 プロモーター 構造遺伝子

(1) パターン1

・塩基配列

gatctcctcggcggaaca...
 gat → D → 23.22
 atc → I → -18.32
 tct → S → -1.54
 ctc → L → -17.79
 tcc → S → -1.54

(2) パターン2

・塩基配列

gatctcctcggcggaaca...
 g*t*t → V → -16.22
 a*c*c → T → -4.15
 t*t*c → F → -21.98

(3) パターン3

・塩基配列

gatctcctcggcggaaca...
 g**c**c → A → -12.04
 a**t**t → I → -18.32
 t**c**c → S → -1.54

【図4】

FIG.4

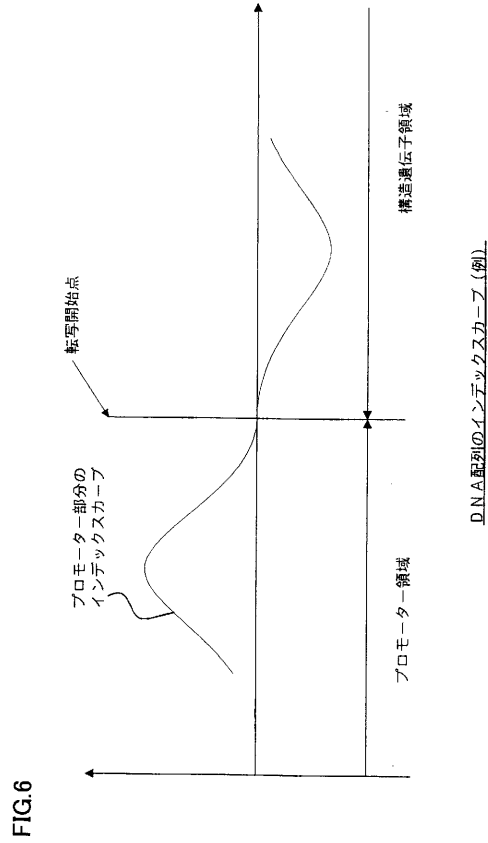
		2番目の塩基								3番目の塩基
		t	c	a	g	t	c	a	g	
1番目の塩基	t	ttt ttc tta ttg	F L	tct tcc tca tcg	S	tat tac taa tag	Y 終止	tgt tgc tga tgg	C 終止	t a g
	c	ctt ctc cta ctg	L	cct ccc cca ccg	P	cat cac caa cag	H Q	cgt cgc cga cgg	R	t c a g
	a	att atc ata atg	I M	act acc aca acg	T	aat aac aaa aag	N K	agt agc aga agg	S R	t c a g
	g	gtt gtc gta gtg	V	gct gcc gca gcg	A	gat gac gaa gag	D E	ggt ggc gga ggg	G	t c a g

【 図 5 】
FIG.5

アミノ酸文字列	アミノ酸インデックス値
A (アラニン)	-12.04
R (アルギニン)	39.23
N (アスパラギン)	4.25
D (アスパラギン酸)	23.22
C (システイン)	3.95
Q (グルタミン)	2.16
E (グルタミン酸)	16.81
G (グリシン)	-7.85
H (ヒスチジン)	6.28
I (イソロイシン)	-18.32
L (ロイシン)	-17.79
K (リジン)	9.71
M (メチオニン)	-8.86
F (フェニルアラニン)	-21.98
P (プロリン)	5.82
S (セリン)	-1.54
T (トレオニン)	-4.15
W (トリプトファン)	-16.19
Y (チロシン)	-1.51
V (バリン)	-16.22
X (終止) *1	0.00

* transfer free energy to lipophilic phase, von Heijne-Blomberg 1979
 *1: タンパク質合成の終わり (翻訳終了) を決めているコドンを表す。

【 図 6 】



【 図 7 】
FIG.7

```

1 cggattagcagagcagatgatggcacaacgggtctacagattcttgaagtagtgccc
61 gactacggctacactagaaggacagatttggatctgctgctgaagccattacc
121 ttggaaaaagagttggtagctcttgatccgg

```

【 図 10 】
FIG.10

```

1 tcaattgctacaatcacttcattattaatttaataatataatgtggtatataatgaac
61 tgttagagaaataatagctccaccatattttttctcaattttttcactataaaaagg
121 ctatttcattataatcaaaaacacacacaaaagagaaggagcaataaaaataaagta
181 aacaacaatttgggttttaaaaaaaaaaaaaagtacacacacaaaaaaaattc

```

【 図 8 】
FIG.8

```

1 gatctcctcggcgaacacaccctccagctcggcttggacgggctgcccgtcaccggga
61 gagctcggctcgtgctcagctcaggtggcragtgcggcggcggcggcggcggcggcggc
121 gcggggaactagatagctcggcttccgtgctcggcggcggcggcggcggcggcggcggc
181 ccgctgtgtaacgttttccgagcctctttgcatcattcggcggcggcggcggcggcggc
241 agcttatacccgccggaccaggcactctcgcacggcggcggcggcggcggcggcggcggc
301 acaggaggaaacattgcacatcaagggaagaaaggcggcggcggcggcggcggcggcggc
361 ggtgctcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggc
421 ctaaaggggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggcggc
481 ccccgaggcgaacctggacacacacacacacacacacacacacacacacacacacacac
541 ggtttccagtagcagcaacacacacacacacacacacacacacacacacacacacacac

```

【 図 11 】
FIG.11

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
c	c	g	g	a	t	t	a	g	c	a	g	a	g	c
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
g	a	t	g	a	t	g	g	c	a	c	a	a	a	c
31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
g	g	t	g	c	t	a	c	a	g	a	g	t	t	c
46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
t	t	g	a	a	g	t	a	g	t	g	g	c	c	c
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
g	a	c	t	a	c	g	g	c	t	a	c	a	c	t
76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
a	g	a	a	g	g	a	c	a	g	t	a	t	t	t
91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
g	g	t	a	t	c	t	g	c	g	c	t	c	t	g
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
c	t	g	a	a	g	c	c	a	g	t	t	a	c	c
121	122	123	124	125	126	127	128	129	130	131	132	133	134	135
t	t	c	g	g	a	a	a	a	a	g	a	g	t	c
136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
g	g	t	a	g	c	t	c	t	t	g	a	t	c	c
151	152													
g	g													

【 図 9 】
FIG.9

```

1 gagaacacagctggaatttttacaagtagttggtgaagctagttagcgaatccatt
61 accttccactctacctaaccccttaccacaacacaaatctgttaatttaaaactagc
121 caaaaaagaactctcttttacaagagccaaagactcaactttactttcaagaa

```

【図12】
FIG.12

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
121	122	123	124	125	126	127	128	129	130	131	132	133	134	135
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
092	092	092	092	092	092	092	092	092	092	092	092	092	092	092
151	152	153												
092	092	092												

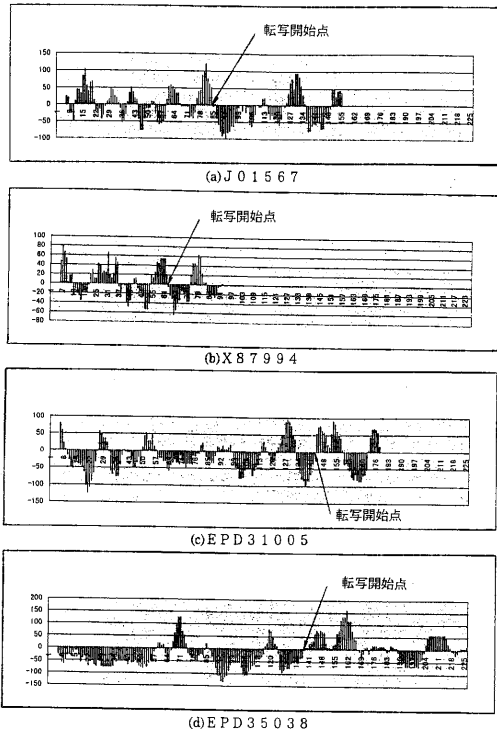
【図13】
FIG.13

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
58	392	-79	232	-183	-178	00	-15	-120	22	392	168	-15	-120	392
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
232	-89	00	232	-89	-162	-79	-120	63	-42	22	97	43	-42	392
31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
-79	-162	40	-120	-178	-15	-42	22	392	168	-15	-162	-220	-15	-178
46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
-178	00	168	97	-15	-162	00	-15	-162	-162	-79	-120	58	58	392
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
232	-42	-178	-15	-42	392	-79	-120	-178	-15	-42	63	-42	-178	00
76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
392	168	97	392	-79	232	-42	22	-15	-162	-15	-183	-220	-178	-162
91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
-79	-162	-15	-183	-15	-178	40	-120	392	-120	-178	-15	-178	40	-120
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
-178	00	168	97	-15	-120	58	22	-15	-162	-178	-15	-42	58	-178
121	122	123	124	125	126	127	128	129	130	131	132	133	134	135
-220	-15	392	-79	168	97	97	97	97	392	168	-15	-162	-178	-162
136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
-79	-162	00	-15	-120	-178	-15	-178	-178	00	232	-183	-15	58	392
151	152	153	154	155										
00	00	00	00	00										

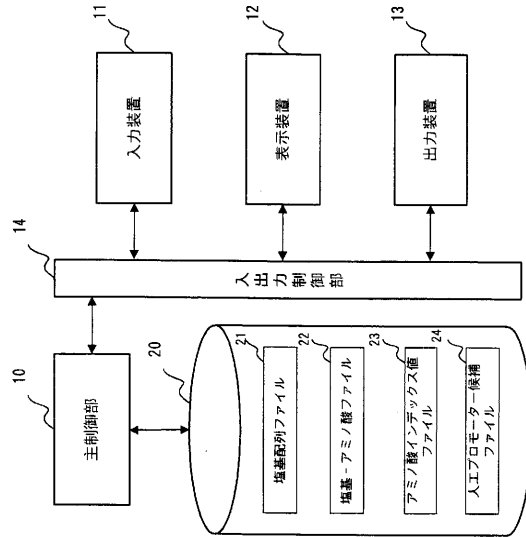
【図14】
FIG.14

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
					243	185	-223	-265	-475	100	446	431	326	839
16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
3049	568	400	648	680	125	-185	-217	-154	-428	-318	-59	62	141	471
31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
434	250	192	29	-107	-515	-478	-294	59	348	510	363	185	148	-423
46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
-769	-753	-423	-106	-106	-90	88	72	-258	-517	-80	-838	-480	-407	148
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
542	579	522	448	349	349	38	-41	-41	-41	-41	-371	-334	-391	-213
76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
194	404	438	872	971	1304	770	623	511	-44	20	-396	-574	-774	920
91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
836	-984	-815	-779	-616	-632	-514	-473	-65	-62	-165	-62	-220	-60	-573
106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
-630	-452	-269	07	-48	-48	188	209	26	-234	-396	-291	-391	-354	-516
121	122	123	124	125	126	127	128	129	130	131	132	133	134	135
-574	-412	-04	-41	69	344	661	773	478	949	949	836	577	302	43
136	137	138	139	140	141	142	143	144	145	146	147	148	149	150
-428	-758	-743	-896	-538	-554	-491	-507	-485	-669	-317	-322	-322	-86	484
151	152	153	154	155										
484	252	435	451	393										

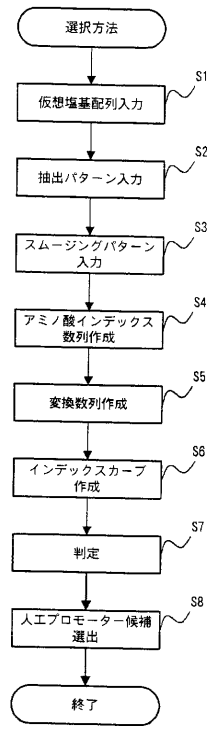
【図15】
FIG.15



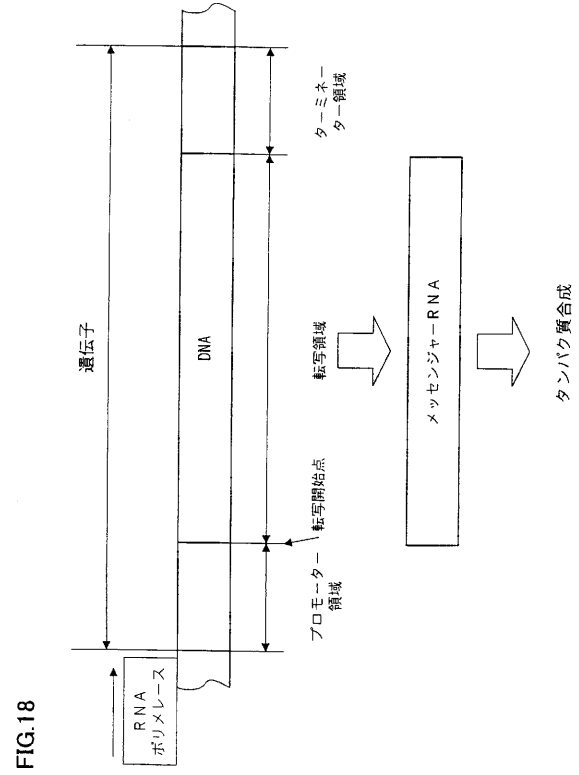
【図16】
FIG.16



【 図 17 】
FIG.17



【 図 18 】



フロントページの続き

(56)参考文献 Comput.Appl.Biosci., 13[1](1997) p.29-35
Gene, 187[2](1997) p.281-287

(58)調査した分野(Int.Cl., D B 名)

C12N15/

CA/BIOSIS/MEDLINE/WPIDS(STN)

JSTPlus/JMEDPlus/JST7580(JDreamII)