

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号  
特許第4094679号  
(P4094679)

(45) 発行日 平成20年6月4日(2008.6.4)

(24) 登録日 平成20年3月14日(2008.3.14)

(51) Int.Cl.

F I

C 1 2 N 15/09 (2006.01)

C 1 2 N 15/00 Z N A A

A O 1 K 67/027 (2006.01)

A O 1 K 67/027

A 6 1 K 39/29 (2006.01)

A 6 1 K 39/29

A 6 1 K 48/00 (2006.01)

A 6 1 K 48/00

A 6 1 P 31/12 (2006.01)

A 6 1 P 31/12

請求項の数 40 (全 125 頁) 最終頁に続く

(21) 出願番号	特願平10-538854	(73) 特許権者	ワシントン ユニヴァーシティ
(86) (22) 出願日	平成10年2月26日(1998.2.26)		アメリカ合衆国, ミズーリ州 63130
(65) 公表番号	特表2001-515353(P2001-515353A)		, セントルイス, ワン・ブルッキングス・
(43) 公表日	平成13年9月18日(2001.9.18)		ドライヴ(番地なし)
(86) 国際出願番号	PCT/US1998/004428	(74) 代理人	弁理士 川口 義雄
(87) 国際公開番号	W01998/039031	(74) 代理人	弁理士 小野 誠
(87) 国際公開日	平成10年9月11日(1998.9.11)	(74) 代理人	弁理士 金山 賢敦
審査請求日	平成16年10月1日(2004.10.1)	(74) 代理人	弁理士 大崎 勝真
(31) 優先権主張番号	08/811,566	(74) 代理人	弁理士 坪倉 道明
(32) 優先日	平成9年3月4日(1997.3.4)		
(33) 優先権主張国	米国(US)		
早期審査対象出願			
			最終頁に続く

(54) 【発明の名称】 C型肝炎ウイルス(HCV)用の機能性DNAクローン及びその使用

(57) 【特許請求の範囲】

【請求項1】

ポジティブセンス核酸上の5'から3'方向に向かって、一番端の5'末端保存配列を含む機能性5'非翻訳領域(NTR)であって当該機能性5'非翻訳領域はNCCAGCCからなり、Nは1~3の任意のヌクレオチドである5'非翻訳領域と、HCV非構造タンパク質コーディング領域コンセンサス配列をコードするオープンリーディングフレーム(ORF)であって、当該非構造タンパク質コーディング領域コンセンサス配列は、少なくとも3つのHCV分離体のクローン由来のコンセンサス残基を同定することにより決定され、又は該コンセンサス配列は配列番号1の非構造タンパク質コーディング領域ヌクレオチド配列において一個又は数個のヌクレオチドが置換、欠失、挿入、又は付加したものであり、及び、配列番号4と少なくとも90%の同一性を有する一番端の3'末端保存配列を含む3'非翻訳領域(NTR)からなる、宿主細胞において複製能を有する遺伝子操作されたC型肝炎(HCV)核酸クローンまたはRNAである該核酸クローン。

【請求項2】

配列番号1で表される配列のうち、少なくとも、機能性部分である一番端の5'末端保存配列、HCV非構造タンパクコード領域、又は3'非翻訳領域(NTR)を含む請求項1に記載のHCV核酸クローン。

【請求項3】

配列番号1で表される完全長配列からなる請求項1又は2に記載のHCV核酸クローン。

【請求項4】



、感染量の請求項 1 3 に記載の H C V R N A と接触させることを含む H C V の感染に許容である株化細胞を同定する方法。

【請求項 1 6】

培養中の株化細胞を、感染量の請求項 1 0 記載の H C V R N A と接触させ、株化細胞内の H C V R N A レベルが前進的に増加するのを検出することを含む許容株化細胞における高いレベルの H C V 複製を可能にする適応突然変異のある H C V を選択する方法。

【請求項 1 7】

請求項 1 に記載の H C V 核酸クローンでトランスフェクト、トランスフォーム又はトランスデューズされた宿主株化細胞。

【請求項 1 8】

バクテリア細胞、イースト細胞、植物細胞、昆虫細胞と哺乳類細胞からなる群から選ばれる請求項 1 7 に記載の宿主株化細胞。

【請求項 1 9】

請求項 1 0 に記載の H C V R N A の感染供与量をヒトを除く動物に投与することを含む動物を H C V に感染させる方法。

【請求項 2 0】

請求項 1 3 に記載の H C V R N A の感染供与量をヒトを除く動物に投与することを含む動物を H C V に感染させる方法。

【請求項 2 1】

請求項 1 又は 2 に記載の H C V 核酸クローンのゲノム R N A 配列からなる H C V で感染させた非ヒト動物。

【請求項 2 2】

H C V R N A の複製を可能とする条件下において、請求項 1 0 記載の H C V R N A の感染量で接触させた株化細胞を培養することを含む生体外で H C V を増殖させる方法。

【請求項 2 3】

H C V R N A の複製を可能とする条件下において請求項 1 3 記載の H C V R N A の感染量で接触させた株化細胞を培養することを含む生体外で H C V を増殖させる方法。

【請求項 2 4】

請求項 1 又は 2 に記載の H C V 核酸クローンのゲノム R N A 配列からなる H C V で感染させた生体外株化細胞。

【請求項 2 5】

肝臓株化細胞である請求項 2 4 の株化細胞。

【請求項 2 6】

H C V で非ヒト動物を感染させるのに効果的な量の請求項 6 記載の H C V 核酸クローンを非ヒト動物に投与することを含む、異種遺伝子で非ヒト動物をトランスデューズする方法。

【請求項 2 7】

H C V R N A で非ヒト動物を感染させるのに効果的な量の請求項 1 3 記載の H C V R N A を非ヒト動物に投与することを含む、異種遺伝子で非ヒト動物をトランスデューズする方法。

【請求項 2 8】

a ) H C V 複製とウイルス粒子形成を可能にする条件下において、請求項 1 7 記載の株化細胞を培養し、かつ

b ) 前記株化細胞培養物から H C V ウイルス粒子を分離することを含む H C V ウイルス粒子を生産する方法。

【請求項 2 9】

a ) H C V 複製とウイルス粒子形成を可能にする条件下において、請求項 2 4 記載の株化細胞を培養し、かつ

b ) 前記株化細胞培養物から H C V ウイルス粒子を分離することを含む H C V ウイルス粒子を生産する方法。

10

20

30

40

50

## 【請求項 3 0】

- a) H C V 粒子タンパク質の発現を可能にする条件下において、請求項 1 記載の H C V 核酸でトランスフェクトされた宿主発現株化細胞を培養し、かつ
- b) 前記細胞培養物から H C V 粒子タンパク質を分離することを含む H C V 粒子タンパク質を生産する方法。

## 【請求項 3 1】

請求項 1 又は 2 に記載の H C V 核酸クローンであって、複製能のある H C V ゲノム RNA である H C V 核酸クローンを含む H C V ウイルス粒子。

## 【請求項 3 2】

請求項 1 0 記載の H C V \_\_ RNA のゲノムテンプレート、機能性 H C V レプリカーゼ要素及びリボヌクレオチドトリホスフェート塩基を含む等張緩衝培地を含む H C V 用の生体外無細胞アッセイキット。

10

## 【請求項 3 3】

請求項 1 3 記載の H C V \_\_ RNA のゲノムテンプレート、機能性 H C V レプリカーゼ要素及びリボヌクレオチドトリホスフェート塩基を含む等張緩衝培地を含む H C V 用の生体外無細胞アッセイキット。

## 【請求項 3 4】

- a) 候補剤を請求項 2 1 記載の H C V 感染動物に投与し、及び
- b) 候補剤の投与前の動物における H C V 感染又は活性レベルと比較して H C V 感染又は活性レベルの増加又は減少をテストする、
- ことを含む H C V 複製の調節可能な剤をスクリーニングする方法であって、
- ここで候補剤の投与前の動物における H C V 感染又は活性レベルと比較して H C V 感染又は活性レベルの減少は H C V 感染又は活性を阻害させる能力がある剤であることを示すものであるスクリーニング方法。

20

## 【請求項 3 5】

- a) 動物からの組織サンプル中のウイルスタイターを測定すること、
- b) 動物からの組織サンプル中のウイルスタンパク質を測定すること、及び
- c) 肝臓酵素を測定することから成る群から選ばれる、H C V 感染レベルをテストする方法である請求項 3 4 に記載の方法。

## 【請求項 3 6】

動物を感染させるために使用する H C V ゲノムが発現調節配列と操作上関連した異種遺伝子を含み、ここで異種遺伝子及び発現調節配列はポジティブ鎖核酸分子上に配向され、H C V 活性レベルのテストは動物からの組織サンプル中のマーカータンパク質のレベルを測定することを含む請求項 3 4 に記載の方法。

30

## 【請求項 3 7】

- a) 請求項 2 4 の株化細胞を候補剤と接触させること、及び
- b) 候補剤の投与前のコントロール株化細胞又は株化細胞での H C V 感染レベル又は H C V 活性レベルと比較して H C V 感染又は H C V 活性のレベルの増加又は減少をテストすること、
- を含む H C V 複製を調節可能な剤のスクリーニング方法であって、
- ここで候補剤の投与前のコントロール株化細胞または株化細胞での H C V 感染レベル又は H C V 活性レベルと比較して H C V 感染又は H C V 活性のレベルが減少していることは、H C V 感染又は H C V 活性を阻害する能力がある剤であることを示すものであるスクリーニング方法。

40

## 【請求項 3 8】

- H C V 感染のレベルのテストが、
- a) 細胞、培地、又はその両方におけるウイルスタイターを測定すること、及び
- b) 細胞、培地、又はその両方におけるウイルスタンパク質を測定すること
- からなる群から選ばれてなる請求項 3 7 に記載の方法。

## 【請求項 3 9】

50

株化細胞を感染させるために使用するHCVゲノムが発現制御配列に操作可能に結合した異種遺伝子を含み、当該異種遺伝子及び発現制御配列はポジティブ鎖核酸分子上に配置され、かつHCV活性のレベルのテストは動物由来の組織サンプルのマーカータンパク質のレベルを測定することを含むものである、請求項37に記載の方法。

#### 【請求項40】

a) 請求項32又は33記載の生体外アッセイキットを候補剤と接触させること、及び  
b) 候補剤を投与する前のコントロール細胞系又は系におけるHCV複製のレベルに比べ、HCV複製のレベルの増減をテストすること  
を含むHCV複製を調節することができる剤をスクリーニングする方法であって、候補剤を投与する前のコントロール株化細胞又は株化細胞におけるHCV複製のレベルに比べ、HCV複製のレベルが減少することは、HCV感染又はHCV活性を阻害する能力を有する剤であることを示すものであるHCV複製を調節することができる剤をスクリーニングする方法。

10

#### 【発明の詳細な説明】

##### 政府援助

本発明に先行する研究は、少なくとも部分的に米国パブリックヘルスサービスの研究助成第CA57973号及びAI31501号からの助成金により援助された。したがって、政府は本発明にある種の権利を有する。

##### 本発明の分野

本発明は機能性HCVウイルスゲノムRNA配列の決定に関し、感染HCV DNAクローンの構造に関し、治療、ワクチン及び診断応用におけるクローン、若しくはその誘導体の使用に関する。さらに本発明はHCVベクター、例えば遺伝子治療若しくは遺伝子ワクチンに関する。

20

##### 発明の背景

##### C型肝炎ウイルスの簡潔な一般的概要

A型肝炎ウイルス及びB型肝炎ウイルス用の診断テストの開発の後、チンパンジーへ実験的に伝染させる更なる作用物質(AlterらによるLancet 1, 459-463 (1978); HollingerらによるIntervirology 10, 60-68 (1978); TaborらによるLancet 1, 463-466 (1978))が、輸血による肝炎の主な原因として認識されるようになった。C型肝炎ウイルスと呼ばれる非A型非B型(NANB)肝炎の原因に相当するcDNAクローンは1989年に報告された(ChooらによるScience 244, 359-362 (1989))。このブレイクスルーはHCVの診断法、及び疫学、病原論及び分子ウイルス学の理解に急速な進展をもたらした(概観のためにはHoughtonらによるCurr Stud Hematol Blood Transfus 61, 1-11 (1994)を参照)。HCV感染の証明は世界中で発見され、HCV特異抗体の普及率は、多くの国で0.4-2%から、エジプトでの14%までの範囲にある(HibbsらによるJ. Inf. Dis. 168, 789-790 (1993))。血液若しくは血液生成物、又はあまり頻繁ではない性的及び先天性ルートを経由する伝染に加えて、周知の危険要因と関連しない散发性のケースは、HCVの場合の40%以上を占める(AlterらによるJ. Am. Med. Assoc. 264, 2231-2235 (1990); MastとAlterによるSemin. Virol. 4, 273-283 (1993))。感染は通常慢性上の(AlterらによるN. Eng. J. Med. 327, 1899-1905 (1992))、臨床上的の不適切なキャリア状態から急性肝炎、慢性活性肝炎及び肝硬変にわたる結果であり、肝硬変は肝細胞癌の進展とかなり関連している。インターフェロン(IFN)- $\alpha$ は、慢性HCV感染(DavisらによるN. Engl. J. Med. 321, 1501-1506 (1989); DiBisceglieらによるNew Engl. J. Med. 321, 1506-1510 (1989))とチンパンジーモデル(ChooらによるProc. Natl. Acad. Sci. USA 91, 1294-1298 (1994))においてある種の兆候を示すサブユニットワクチンのある少数の患者の治療には有用であることが示されており、より有効な治療及びワクチンを開発するために更なる効果が必要とされる。さまざまなHCV分離体の間に観測されたかなりの多様性(概観のためには、BukhらによるSem. Liver Dis. 15, 41-63 (1995)を参照)、慢性的に感染した個体における遺伝的変異体の発生(EnomotoらによるJ. Hepatol. 17, 415-416 (1993); HijikataらによるBiochem. Biophys. Res Comm. 1

30

40

50

75, 220-228 (1991); KatoらによるBiochem. Biophys. Res. Comm. 189, 119-127 (1992); KatoらによるJ. Virol. 67, 3923-3930 (1993); KurosakiらによるHepatology 18, 1293-1299 (1993); LesniewskiらによるJ. Med. Virol. 40, 150-156 (1993); OgataらによるProc. Natl. Acad. Sci. USA 88, 3392-3396 (1991); WeinerらによるVirology 180, 842-848 (1991); WeinerらによるProc. Natl. Acad. Sci. USA 89, 3468-3472 (1992)、及びH C V感染後に惹起された保護免疫の欠如 (FarciらによるScience 258, 135-140 (1992); PrinceらによるJ. Infect. Dis. 165, 438-443 (1992)) は、上記目的に対する大きなチャレンジである。

#### H C Vの分子生物学

分類。ゲノム構造及びビリオン性質に基づいて、H C Vは、二つの他の属であるフラビウイルス (例えば黄熱 (Y F)) と動物ペスチウイルス (例えばウシのウイルス下痢ウイルス (B V D V)) と古典的ブタ熱ウイルス (C S F V) (FranckiらによるArch. Virol. Suppl. 2, 223 (1991)) を含むフラビウイルス系統における分離属として分類されてきた。上記系統の全てのメンバーは、単一の長いオープンリーディングフレーム (O R F) の翻訳を経由して、全ての周知のウイルス特異タンパク質をコード化する (encoding) プラス鎖 R N Aゲノムを含むビリオンを包み込む。

ビリオンの構造及び物理的性質。H C Vの構造と複製に関しては殆ど情報が入手できない。効率的なウイルス複製を支えることが可能な細胞培養の欠如と、一般的には血清に存在する感染性ウイルスの低いタイターにより研究は邪魔される。濾過実験に基づく感染性ウイルスのサイズは、30-80 nmの間である (BradleyらによるGastroenterology 88, 773-779 (1995); HeらによるJ. Infect. Dis. 156, 636-640 (1987); YuasaらによるJ. Gen. Virol. 72, 2021-2024 (1991))。スクロースにおける感染物質の浮遊密度の初期の測定により、1.1 g/ml以下の低密度のものが、ある値の範囲で主に存在していることが分かった (BradleyらによるJ. Med. Virol. 34, 206-208 (1991))。その後研究では、慢性的に感染したヒト若しくは実験的に感染させたチンパンジーからの血清に存在する潜在的感染性ウイルスの間接的測定として、H C V特異R N Aを検出するためにR T / P C Rが利用された。これらの研究から、さまざまな臨床サンプルの間にはかなりの不均一性が存在し、しかも多くの要因がH C V R N Aを含む粒子の挙動に影響を与えることが、ますます明らかになってきた (HijikataらによるJ. Virol. 67, 1953-1958 (1993); ThomsenらによるMed. Microbiol. Immunol. 81, 293-300 (1992))。かかる要因は免疫グロブリン (Hijikataらによる (1993) の上記の文献) 若しくは低密度リボタンパク質 (Thomsenらによる1992の上記文献; ThomsenらによるMed. Microbiol. Immunol. 182, 329-334 (1993)) と関連している。かなりの感染性急性期チンパンジーの血清において、H C V特異R N Aは低浮遊密度 (1.03-1.1 g/ml) の感染で通常検出される (CarrickらによるJ. Virol. Meth. 39, 279-289 (1992); Hijikataらによる (1993) の上記文献)。他のサンプルにおいて、H C V抗体の存在及び免疫複合体の形成は、高い密度と低い感染力の粒子と相関関係がある (Hijikataらによる (1993) の上記文献)。感染力を破壊するクロロホルムによる (BradleyらによるJ. Infect. Dis. 148, 245-265 (1983); FeinstoneらによるInfect. Immun. 41, 816-821 (1983))、若しくは非イオン性界面活性剤による粒子の処理は、H C Vヌクレオキャプシドを表わすと信じられている高密度 (1.17-1.25 g/ml) の粒子を含むR N Aを生じさせる (Hijikataらによる (1993) の上記文献; KantoらによるHepatology 19, 296-302 (1994); MiyamotoらによるJ. Gen. Virol. 73, 715-718 (1992))。

血清及び血漿にあるネガティブ-センスH C V特異R N Aの報告がある (FongらによるJournal of Clinical Investigation 88: 1058-60 (1991) を参照)。しかしながら、高感染力を有する血清は、ネガティブ-センスR N Aの低レベル若しくは検出不可能なレベルであるので、かかるR N Aは感染粒子の本質的成分であるとは思われない (SimizuらによるProc. Natl. Acad. USA 90: 6037-6041 (1993))。

ビリオンタンパク質成分は活発には決定されていないが、推定されるH C V構造タンパク質は、塩基性Cタンパク質と二つの細胞膜糖タンパク質E 1及びE 2を含む。

10

20

30

40

50

H C V 複製。H C V 複製の初期の出来事は殆ど理解されていない。H C V 糖タンパク質の細胞受容体は確認されていない。H C V 粒子とベータリポタンパク質及び免疫グロブリンの関連性は、これらの宿主 (host) 分子がウイルス摂取と組織向性 (tropism) を調整する可能性を高める。H C V の複製を調べる研究は、ヒト患者若しくは実験的に接種させたチンパンジーに主として限られていた。チンパンジーモデルにおいて、H C V R N A は接種後三日目に速くも血清に検出され、血清アラニンアミノトランスフェラーゼ (A L T) レベル (肝臓損傷の指標として) のピークから持続する (ShimizuらによるProc. Natl. Acad. Sci. USA 87: 6441-6444 (1990))。ウイルス血症の兆候は、肝臓のH C V 感染の間接的な特徴の発生を伴う。このことは細胞質抗原の発生 (Shimizuらによる (1990) の上記文献) と、クロロホルムに敏感な“作用物質を形成する細管 (tubule)”若しくは“T F A” (Bradleyらにより概観されているProg. Med. Virol. 37: 101-135 (1990)) と従来から呼ばれるH C V に対する微小管凝集物の形成のような肝細胞における超微構造の変化の発生を伴う。ウイルス抗原の発生 (BlightらによるAmer. J. Path. 143: 1568-1573 (1993)) ; HiramatsuらによるHepatology 16: 306-311 (1992) ; KrawczynskiらによるGastroenterology 103: 622-629 (1992) ; YamadaらによるDigest. Dis. Sci. 38: 882-887 (1993)) 及びポジティブネガティブセンスR N A の検出 (Fongらによる (1991) の上記文献 ; GunjiらによるArch. Virol. 134: 293-302 (1994) ; HarumaらによるJ. Hepatol. 18: 96-100 (1993) ; LamasらによるJ. Hepatol. 16: 219-223 (1992) ; Nouri AriaらによるJ. Clin. Inves. 91: 2226-34 (1993) ; SherkerらによるJ. Med. Virol. 39: 91-96 (1993) ; TakeharaらによるHepatology 15: 387-390 (1992) ; TanakaらによるLiver 13: 203-208 (1993)) により示されるように、特に急性感染中では、肝細胞はH C V 複製の主な部位であるらしい (NegroらによるProc. Natl. Acad. Sci. USA 89: 2247-2251 (1992))。H C V 感染の後段階では、H C V 特異抗原の発生、ウイルス血症の持続若しくは分割、及び肝臓病の発病度は、チンパンジーとヒト患者の双方では大きく変化する。ある種の肝臓損傷は、H C V 感染及び細胞変性の直接的な結果として発生するが、新たに出てきているコンセンサスは宿主免疫応答、特にウイルス特異細胞毒性Tリンパ球は細胞損傷の媒介においてより支配的な役割をする。

さらにH C V は肝臓外の貯蔵所においても複製することは推測されてきた。ある場合には、R T / P C R 若しくは切片上ハイブリッドは、H C V R N A とT細胞、B細胞及び単核細胞を含む周辺の血液単核細胞との関連を示す (BlightとGowansによりViral Hepatitis Rev 1: 143-155 (1995) に概観されている)。かかる細胞向性は慢性感染の確率に関連しており、H C V 感染と混合寒冷グロブリン血症 (FerriらによるEur. J. Clin. Invest. 23: 399-405 (1993)) 系球体腎炎、及び稀な非ホジキンBリンパ腫 (Ferriらによる (1993) の上記文献 ; KagawaらによるLancet 341: 316-317 (1993)) のようなある種の免疫異常との間の関連性において役割を担っている。しかしながら、血清中の循環性ネガティブ (マイナス) 鎖R N A の検出、本当の鎖特異R T / P C R を得る困難性 (Gunjiらによる (1994) の上記文献)、及び見かけ上の感染細胞の数が少ないことは、生体外での上記組織における複製の明らかな証拠を得ることを難しくしている。

ゲノム構造。数多くのH C V 分離体の完全長若しくは殆ど完全長に近いゲノム配列が、報告されている (LinらによるJ. Virol. 68: 5063-5073 (1994a)) ; OkamotoらによるJ. Gen. Virol. 75: 629-635 (1994) ; SakamotoらによるJ. Gen. Virol. 75: 1761-1768 (1994) 及び上記文献の引用文献を参照)。分離体の間でかなりの遺伝的多様性があると仮定すると、幾つかの主なH C V 遺伝子型が世界中に分布していることが明らかである。米国でかなり重要なタイプは、遺伝子型1、サブタイプ1a及び1bである (遺伝子型の有病率と分布は以下を参照し、Bukhらによる (1995) の上記文献を参照)。H C V ゲノムR N A は長さが約9.6キロボース (k b、図1) である。5' N T R は341-344ベース (base) の長さであり、かなり保存されている。長いO R F の長さは、3030、3031のポリタンパク質若しくは3033個のアミノ酸をコード化する分離体の間では僅かに変化している。報告された3' N T R 構造は成分と長さ (28-42の塩基) の双方においてかなりの多様性を示し、3' 末端ポリ (A) トラクト (tract) を含むある場合 (H C

10

20

30

40

50

V-1、タイプ1a) 以外は、ポリ(U)で終結するようだ(ChenらによるVirology 188: 102-113 (1992); OkamotoらによるGen. Virol. 72: 2697-2704 (1994); TokitaらによるJ. Gen. Virol. 66: 1476-83 (1994)を参照)。対照的に、我々の最近の分析により、H-菌株(タイプ1aも)のゲノムRNAは、内部ポリピリミジントラクトとそれに続く新規なRNA要素を含むことを示唆している(係属中の1995年8月29日に出願した米国特許出願08/520、678号及び1996年8月28日に出願した国際特許出願第PCT/US96/14033号)。係属中の米国特許出願第08/520、678号に示した結果からは、このタイプ1a分離体のゲノムRNAは、従来考えられていたようにホモポリマートラクトで終結せずに、むしろ約98の塩基の新規な配列で終結することが分かる。さらに、この3' NTR構造及び新規な3'末端要素は、これまでに調査された全てのHCV遺伝子型に共通の特徴である(KolykhalovらによるJ. Virol. 70: 3363-3371 (1996); TanakaらによるBiochem. Biophys. Res. Comm. 215: 744-749 (1996); TanakaらによるJ. Virol. 70: 3307-12 (1996); YamadaらによるVirology 223: 255-261 (1996))。

翻訳とタンパク質分解プロセッシング(processing)。幾つかの研究は、翻訳開始における5' NTRの役割の調べるために、無細胞翻訳及び細胞培養での一過性発現を利用した(FukushiらによるBiochem. Biophys. Res. Comm. 199:425-432 (1994); Tsukiyama-KoharaらによるJ. Virol. 66: 1476-1483 (1992); WangらによるJ. Virol. 67: 3338-3344 (1993); YooらによるVirology 1991:889-899 (1992))。このかなり保存された配列は多重の短いAUG-開始PRFを含み、ペスチウイルス属の5' NTR領域と重要な相同性を示す(BuhkらによるProc. Natl. Acad. Sci. USA 89: 4942-4946 (1992); Hanらによる(1991)の上記文献)。一連のステムループ構造は、コンピュータモデリングとさまざまなリボ核酸による消化に対する感度に基づいて提案されている(BrownらによるNucl. Acids Res. 20: 5041-5045 (1992); Tsukiyama-Koharaらによる(1992)の上記文献)。幾つかのグループからの結果により、この要素は長いORFの第一のAUGで効率的な翻訳の開始を可能にする内部リボソームエントリ部位(IRES)として機能することが分かる(Fukushiらによる(1994)の上記文献; Tsukiyama-Koharaらによる(1992)の上記文献; Wangらによる(1993)の上記文献; Yooらによる(1992)の上記文献)。HCVとペスチウイルスIRES要素の幾つかの予想される特徴は、お互いに同じである(Brownらによる(1992)の上記文献)。IRESとして機能するこの要素の能力から、HCVゲノムRNAは、5' キャップ構造が欠いていることを示唆している。

HCVポリタンパク質(図1)の組織とプロセッシングは、ペスチウイルスのそれと殆ど同じである。少なくとも10個のポリペプチドは確認され、ポリタンパク質でのこれらの切断生成物の順序は、NH<sub>2</sub>-C-E1-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B-COOHである。図1に示すように、タンパク質分解プロセッシングは、宿主単一ペプチダーゼ及び二つのHCVコード化(encoded)プロテイナーゼ、NS2-3オートプロテイナーゼ並びにNS3-4Aセリンプロテイナーゼにより媒介される(New YorkにあるRaven Pressにより発行された“Fields Virology”(B. N. Fields, D. M. KnipeとP. M. Howleyによる編集)のRiceによるpp.931-960; 概観のためにはShimotohnoらによるJ. Hepatol. 22: 87-92 (1995)を参照)。Cはウイルスコア若しくはキャプシドタンパク質であると信じられている塩基性タンパク質であり、E1及びE2は推定されるビリオンエンベロープ糖タンパク質であり、p7はE2糖タンパク質から非効率的に切断された未知の機能を有する疎水性タンパク質であり(Linらによる(1994a)の上記文献; MizushimaらによるJ. Virol. 68: 6215-6222 (1994); SelbyらによるVirology 204: 114-122 (1994))、同様にNS2-NS5BはウイルスRNA複製複合体において機能する非構造(NS)タンパク質である。特にN末端セリンプロテイナーゼドメインのほかにも、NS3はRNAヘリカーゼの特徴的なモチーフを含み、RNA刺激NTPase活性を有することを示し(SuzichらによるJ. Virol. 67, 6152-6158 (1993))、NS5Bは、プラス鎖RNAウイルスのRNA依存RNAポリメラーゼの特徴的なGDDモチーフを含む。

HCV RNA複製。フラビウイルスとの類似により、プラスセンスHCVビリオンRNA

10

20

30

40

50



Aの複製はマイナス鎖中間体を經由して起こるものと考えられる。この戦略は以下のように簡潔に説明される：(i) 進入するウイルス粒子の脱外被は、たぶん共翻訳と翻訳後に処理される単一の長いポリタンパク質を生じるように翻訳されるゲノムプラス鎖を放出し、個々の構造及び非構造タンパク質が生じる；(ii) 非構造タンパク質はマイナス鎖の合成用のテンプレートとしてビリオンRNAを利用する複製複合体をたぶん形成する；(iii) 上記のマイナス鎖はプラス鎖の合成のテンプレートとして働き、ウイルスタンパク質、マイナス鎖合成の更なる翻訳に利用される若しくは後代ビリオンにパッケージングされる。ウイルス増殖用の良好な実験システムがないので、HCV複製過程についての詳細はほとんど入手できない。本物のHCV複製とウイルス生命サイクルにおける他の段階の詳細な解析は、細胞培養でのHCV複製の効率的なシステムの開発により容易になるであらう。

多くの試みが、HCV感染個体から集められた血清のある培養細胞を感染させるために行われてきており、低レベルの複製はB細胞(BertoliniらによるRes. Virol. 144: 281-285 (1993); NakajimaらによるJ. Virol. 70: 9925-9 (1996); ValliらによるRes. Virol. 146: 285-288 (1995))、T細胞(KatoらによるBiochem. Biophys. Res. Commun. 206: 863-9 (1996); MizutaniらによるBiochem. Biophys. Res. Comm. 227: 822-826; MizutaniらによるJ. Virol. 70: 7219-7223 (1996); Nakajimaらによる(1996)の上記文献; ShimizuとYoshikuraによるJ. Virol. 68: 8406-8408 (1994); ShimizuらによるProc. Natl. Acad. Sci. USA 89:5477-5481 (1992); ShimizuらによるProc. Natl. Acad. Sci. USA 90: 6037-6041 (1993))と肝細胞(KatoらによるJpn. J. Cancer Res., 87: 787-92 (1996); Tagawa, J. Gastroenterol. and Hepatol., 10: 523-527 (1995))、細胞系統と同様に周辺の血液単核細胞(PBMCs)(CribierらによるJ. Gen. Virol., 76: 2485-2491 (1996))とヒト致死肝細胞の一次培養(CarlioniらによるArch. Virol. Suppl. 8: 31-39 (1993); Cribierらによる(1995)の上記文献; IacovacciらによるRes. Virol., 144: 275-279 (1993))若しくは大人のチンパンジーからの肝細胞(LanfordらによるVirology 202: 606-14 (1994))を含む方法により感染された多くの細胞タイプに報告されている。さらにHCV複製は、培養前に(ItoらによるJ. Gen. Virol. 77: 1043-1054 (1996))生体内でウイルスに感染したヒトHCV患者から取出された一次肝細胞において、HCV-1 cDNAクローンからの生体外で転写されたRNAでのトランスフェクションに続くヒト肝細胞系統Hu h7において(YooらによるJ. Virol., 69: 32-38 (1995))検出された。HCV-1クローンはホモポリマーの3' NTR配列下流を欠いているので、HCV-1クローンから誘導されたRNAでトランスフェクト(transfect)された細胞での複製の観察報告はわけがわからぬ(以下を参照)。HCV複製用の最もよく特徴付けられる細胞培養は、レトロウイルス(HPB-Ma若しくはMT-2)で持続的に感染されるB細胞系統(Daudi)若しくはT細胞系統を利用する(Katoらによる(1995)の上記文献; MizutaniらによるBiochem. Biophys. Res. Comm., 227: 822-826 (1996a); Mizutaniらによる(1996)の上記文献; Nakajimaらによる(1996)の上記文献; ShimizuとYoshikuraによる(1994)の上記文献; ShimizuによるProc. Natl. Acad. Sci. USA, 90: 6037-6041 (1993))。HPB Maはマウス肉腫ウイルスの両栄養性マウス白血病ウイルスで感染するのに対し、MT-2はヒトT細胞リンパ栄養性ウイルスタイプI(HTLV-I)で感染する。非クローン集団より効率的にHCV複製を助けるクローン(HPB Ma 10-2及びMT-2 C)は、二つのT細胞系統HPB Ma及びMT-2のために分離された(MizutaniらによるJ. Virol., (1996)の上記文献; Shimizuらによる(1993)の上記文献)。しかしながら、インプットRNAの分解後の上記細胞系統若しくはDaudi系統において得られたRNA複製の最大レベルは、依然として $10^6$ 細胞当たりたったの約 $5 \times 10^4$  RNA分子(Mizutaniらによる(1996)の上記文献; Mizutaniらによる(1996)の上記文献)若しくは培地のml当たり $10^4$  RNA分子(Nakajimaらによる(1996)の上記文献)である。複製のレベルは低い、一つのシステム(MizutaniらによるBiochem. Biophys. Res. Comm. 227: 822-826 (1996a))において198日までの長期間の感染、別のシステム(Nakajimaらによる(1996)の上記文献)において一年以上の感染が報告されてお

10

20

30

40

50

り、感染ウイルス生産は固有 (naive) 細胞へのウイルスの系列無細胞若しくは細胞媒介継代により実証された。

しかしながら、効率的な H C V 複製は今までのところいずれの細胞培養でも観測されておらず、複製、マイナス鎖の誤った検出、及び複製された R N A の一般的に低いタイターにより生成したプラス鎖からインプット R N A を識別する困難さを含む多くの問題に遭遇するかかるシステムを確立するように、全てのグループにより試されている。よって、以上の進展にもかかわらず、H C V 増殖用のより効率的な細胞培養システムは、高濃度のウイルスストックの生産、ビリオン成分の構造解析、及び R N A 複製を含む細胞内でのウイルス過程の改善された解析のために必要である。

ビリオン構築及び放出。この過程は直接調べられておらず、複合多糖の欠如、発現 H C V 糖タンパク質の E R 局在化 (Dubuissonらによる J. Virol. 68: 6147-6160 (1994); Rals onらによる J. Virol. 67: 6753-6761 (1993)) 及び細胞表面に上記タンパク質の不存在 (Dubuissonらによる (1994) の上記文献; Spaeteらによる Virology 188: 819-830 (1992)) により、初期のビリオンの形態形成は、細胞内のベシクルにバディング (budding) することにより起こることを示唆している。これまでのところ、効率的な粒子形成と放出は一過性発現検定において観測されておらず、本質的なウイルス若しくは宿主要因は存在しない又はブロックされていることを示唆している。ウイルスの実質的なフラクション (fraction) は、ペスチウイルスで発見されたように細胞に関連したままであるので、H C V ビリオン形成と放出は非効率的である。最近の研究により、ヒト血漿から部分的に精製された細胞外 H C V 粒子は複合 N 結合多糖を含むが、これらの炭水化物部分は E 1 又は E 2 と特別に関連していないことを示す (Satoらによる Virology 196: 354-357 (1993))。放出ビリオン上の糖タンパク質と関連する複合多糖は、トランスゴルジ体を通して移動し、宿主分泌経路を通してビリオンの移動を示唆している。このことが正しければ、H C V 糖タンパク質の細胞内セクエストレーション (sequestration) 及びビリオン形成は、免疫監視を最小化し、抗体及び補体 (complement) によりウイルス感染細胞の溶菌 (lysis) を防止することにより慢性感染の確立に重要な役割を担う。

遺伝的変異性。全てのプラス鎖 R N A ウイルスに関して、H C V (N S 5 B) の R N A 依存 R N A ポリメラーゼ (R D R P) は、取込み違いの塩基の除去のための 3' - 5' エキソヌクレアーゼ校正活性が欠如していると信じられている。したがって、複製は誤りがちであり、多くの数の変異体から成る“準種 (quasi-species)”ウイルス集合体へ導く (Martellらによる J. Virol. 66: 3225-3229 (1992); Martellらによる J. Virol. 68: 3425-3436 (1994))。この変異性はマルチレベルにおいて明らかである。第一に、慢性的に感染した個体において、ウイルス集合体における変化は時がたつにつれて起こり (Ogataらによる (1991) の上記文献; Okamotoらによる Virology 190: 894-899 (1992))、上記変化は病気に重要な結果をもたらす。特定の興味ある例としては、E 2 糖タンパク質の N 末端の 30 残基セグメントがあり、これはポリタンパク質の残りよりもかなり高い変異性を示す (例えば、Higashiらによる Virology 197, 659-668, 1993; Hijikataらによる (1992) の上記文献; Weinerらによる (1991) の上記文献)。たぶん H I V - 1 g p 1 2 0 の V 3 ドメインに類似するこの超可変性領域は、H C V 特異抗体を循環させることによる免疫選択下にあることを証明するものが蓄積されている (Katoらによる (1993) の上記文献; Taniguchiらによる Virology 195: 297-301 (1993); Weinerらによる (1992) の上記文献)。このモデルにおいて、E 2 の上記部分に対して向けられた抗体は、ウイルス中和に寄与し、よって中和から逃げることを可能にする置換により変異体の選択を引き起こす。この柔軟性は、E 2 超可変性領域における特異アミノ酸配列はビリオン吸着 (attachment)、浸透若しくは集合のようなタンパク質の他の機能に対して必須ではないことを示唆している。

さらに遺伝的変異性は、慢性感染患者の I F N - 治療後に観測されたさまざまな応答のスペクトルに寄与する。減少血清 A L T レベルと、循環する H C V R N A のレベルの減少と通常相関関係にある改善肝組織学は、約 40% の治療で確認される (Greiser-Wilkeらによる J. Gen. Virol. 72: 2105-2019 (1991))。治療後、約 70% の応答個体が再発

10

20

30

40

50

する。ある場合には、循環ウイルスRNAの一過性損失後に、治療中若しくは治療後に新たなウイルス血症が観測される。このことは、IFN耐性HCV遺伝子型若しくは変種体の存在若しくは発生を示唆するが、更なる研究が、ウイルス遺伝子型の相対的寄与及び免疫応答における宿主特異的差異を決定するために必要である。

最終的に、世界中のさまざまなHCV分離体の配列比較により、莫大な数の遺伝的多様性を示している(Bukhらによる(1995)の上記文献を参照して概観される)。交差中和テストのような生物学的に関連する血清学的検定がないので、HCVタイプ(数により示される)、サブタイプ(文字により示される)、及び分離体は、ヌクレオチド若しくはアミノ酸配列類似性に基づいて、現在分類されている。多くの多様な遺伝子型間のアミノ酸配列類似性は、比較されたタンパク質に依存して、約50%ほどである。この多様性は、特に診断、ワクチン設計及び治療によっては重要な生物学的意味を有する。

cDNAからの感染性HCV転写物(transcript)を生じさせるための他人による試み  
最近の論文(YooらによるJ. Virol. 69: 32-38 (1995))には、Hu h 7細胞のトランスフェクション後の転写HCV-1 RNAの複製が報告されている。この論文では、HCV-1 cDNAクローンの種々の誘導体からのT7転写物は、ヒト肝癌細胞系統、Hu h 7の次に続くトランスフェクションを複製させる能力をテストした。起こり得るHCV複製は、鎖特異RT/PCR(5' NTRプライマーを利用して)と<sup>3</sup>H-ウリジンのあるHCV特異RNAの代謝ラベリングにより評価された。明らかに、ポリ(A)若しくはポリ(U)で終結している完全長転写物は、上記評価によりポジティブ(プラス)であったが、5'末端144塩基に欠失のある転写物は存在しなかった。ある培養では、HCV特異RNAは培地で検出され、この推定されるウイルスはフレッシュなHu h 7細胞を再感染させるために利用された。

本発明者らは、上記結果を再現させることができなかった。上記報告は、複製とウイルス生産の本物のHCV感染よりもむしろ一過性複製を報告しているようである。例えば、上記論文で報告されているポジティブコントロール(positive control)は、高いタイターのチンパンジー血漿から抽出されたRNAのあるHu h 7細胞の生産的トランスフェクションであった。この抽出サンプルは、潜在的感染性のある完全長のHCV RNAの最大 $10^7$ 個の分子を含んでいたであろう。(顕微注射以外の)最適トランスフェクション条件下において、一般にはビリオンRNAの $10^5$ 個以上のRNA分子(少なくともポリオウイルス、シンドビスウイルス若しくはYFによっては)は、一つの感染イベントを開始させる必要がある。このことは100個よりも少なく報告されたHCV-1実験では、細胞は生産的にトランスフェクトされるであろうことを示唆している。さらに、トランスフェクションの16日後、ポジティブ及びネガティブ鎖RNAの双方は、代謝ラベリングの8時間後に繰り返して検出された。この方法によるマイナス鎖RNAの検出(トランスフェクトされたビリオンRNAと転写RNAの双方)は、HCVは効率的な複製とスプレッド(spread)の双方が可能であり、しかもHCV RNA合成のレベルは、YFのようなより強いフラビウイルス(高い多重感染のピークで)で予期されるレベルと同じである。しかしYooらは多様な抗血清を利用した上記細胞においてHCV抗原の検出は報告しておらず、さらにノーザン分析(<sup>3</sup>H-ウリジンによる代謝ラベリングよりもかなり高感度である)による完全長のポジティブ及びネガティブ鎖の検出も報告していない。最終的に、RNA若しくはHCV-1クローンから誘導されたウイルスが、チンパンジーモデルで感染されたことを示す重大な実験は報告されていない。

#### HCV研究用の感染クローン技術の重要性

ここ最近の数年における大きな進歩にもかかわらず、HCV複製、病原論及び免疫に関する数多くの疑問は未解決のままである。この分野は、我々にHCV RNAゲノム及びそのコード化されたタンパク質の機能のある面を理解させ、本物のウイルスの複製関係において構造/機能の疑問を実験的にテストさせる方法を全く有しないボトルネックに急速に近づいている。かかる分析は、保護ワクチン、有効な治療法及びHCV診断薬の設計を可能にするために、ウイルス生命サイクルにおける各段階を理解することは重要である。よって、この技術分野では感染HCV RNAの発現のための本物のHCV遺伝物質が必要

である。

さらにこの技術分野では、固有 (naive) HCV ビリオン及びHCV ビリオン構造の特徴付けを可能にするウイルス粒子タンパク質の発現のための本当の遺伝物質が必要である。さらにこの技術分野には、HCV レセプター結合、細胞感染、複製、ビリオン集合体及び放出の分析を可能にする感染性HCV に対する生体外での培養方法が必要である。

この技術分野における上記及び他の必要なことは、本発明で扱われる。

本願での参考文献の引用は、かかる参考文献が緊急出願の“先行技術”として利用可能であることの承認として、解釈されるべきではない。

#### 発明の要約

本発明は複製、機能的HCV タンパク質の発現及び抗ウイルス性治療及び診断薬の開発のための生体内と生体外での感染が可能な本物のC型肝炎ウイルス(HCV) DNA クローンを提供する。

広い態様において、本発明は、ポジティブ-センス核酸上の5'から3'と、一番端の(extreme) 5'末端保存配列から成る機能性5'非翻訳領域(NTT)と、切断生成物がHCV ウイルス粒子とRNA複製機械装置(replication machinery)の機能構成部品を形成するHCV ポリタンパク質の少なくとも一部分をコード化するオープンリーディングフレーム(ORF)と、一番端の3'末端保存配列から成る3'非翻訳領域(NTT)と、若しくは適応ウイルス、弱毒化ウイルス、複製能のある非感染ウイルス、及び欠損ウイルスから成る群から選ばれる誘導体から成る遺伝子的に特別設計された(engineered) C型肝炎ウイルス(HCV) 核酸クローンに関する。本発明者らにより、遺伝子工学技術を利用して実施される多くの操作は、本物のHCV 核酸、例えば感染HCV RNAを生じるように転写されるcDNAを発生させることが必要であることが分かった。特別設計された本物のHCV 核酸を提供することにより、本発明者らは、HCV複製機械装置とタンパク質活性の解剖、及び多様なHCV 誘導体の調製を始めて可能にした。従来では、ある一定のHCV クローンが機能することが不可能になる誤り若しくは突出然変異を含んでいるかどうか不確実であったので、HCVの更なる分析用の出発物質が、人工物(artifact)により有用であるか又は簡単であるか確かではなかった。よって、本発明の大きな利点は、本物のHCVを提供し、よって変種が先行技術で提供されたクローンでの誤りによる人工物よりもむしろ本当の変化をもたらすことを保証することにある。

本発明の更なる利点は、感染HCV ゲノム、特にポリタンパク質暗号づけ(coding)領域の特徴の認識にある。特定実施例において、HCV 核酸は、HCV 分離体若しくは遺伝子型の少なくとも三つのクローンの大部分の配列から決定されたコンセンサス核酸配列を有する。HCV 核酸は、本願で実証された本発明の特定実施例を表わすSEQ ID NO: 1に示すような配列の少なくとも機能部分を有することが好ましい。SEQ ID NO: 1はDNA配列であるが、本発明は対応するRNA配列、及びDNAとRNAの相補的配列のことも考えていることを指摘しておくべきである。更なる実施例において、HCV 分離体からの領域は、例えばSEQ ID NO: 1の配列を有するHCV 核酸の相同領域と置換される。本願で実証された更に好ましい実施例において、HCV 核酸は、複製能のあるHCV RNA複製用の発現上に遺伝暗号を指定する(code) DNAであり、又は自身は複製能のあるHCV RNAレプリコンである。以下の特定例において、本発明のHCV 核酸はSEQ ID NO: 1に示された若しくは対応する完全長配列を有する。さらに5'及び3'のさまざまな変異は、本発明により熟慮される。例えば5'末端配列は、GCCAGCC; GGCCAGCC; UGCCAGCC; AGCCAGCC; AAGCCAGCC; GAGCCAGCC; GUGCCAGCC; 及びGCGCCAGCCから成る群から選ばれるRNA配列に相同若しくは相補的であり、ここで配列GCCAGCCはSEQ ID NO: 3の5'末端である。本発明の更に別の利点は、感染HCV クローン用の完全3'-NTTの重要性のデモンストレーションである。3'-NTT、特にHCV 遺伝子型の間でかなり保存されている約98の塩基の一番端の末端配列は、本願の参考文献中に引用されている1995年8月29日に出願された米国特許出願番号08/520、678号と、これも本願の全体の参考文献中に引用されている1996年8

10

20

30

40

50

月28日に出願されたPCT国際特許出願第PCT/US96/14033号の主題である。よって、好ましい態様において、機能性3'NTRは、HCV遺伝子型の中でかなり保存されている約98の塩基の3'末端配列から成る。特定実施例において、3'NTRの一番端にある末端は、配列5'-GGTGGCTCCATCTTAGCCCTAGTCA  
CGGCTAGCTGTGAAGGTCCGTGAGCCGCATGACTGCAGAGAGTGTCTGATACTGGCCTCTCTGCTGATCATGT-3'(SEQ  
ID NO:4)を有するDNAに相同若しくは相補的である。SEQ ID NO:  
1で実証された特定実施例において、3'NTRは長いポリピリミジン領域(例えば約133の塩基)から成るが、代替のポリピリミジン領域は、短い領域(約75の塩基)、  
若しくは短い又は長い領域を含んでさらに包み込まれる。当然、ポジティブ(プラス)鎖HCV DNA核酸において、ポリピリミジン領域はポリ(T/TC)領域であり、ポジ  
ティブ鎖HCV RNA核酸において、ポリピリミジン領域はポリ(U/UC)領域であ  
る。

本発明のさまざまな態様によれば、ポリタンパク質暗号づけ領域を含むHCV核酸は突然変異し、若しくは特別設計されて、例えばサイレント突然変異、保存突然変異などのある変異体若しくは誘導体が生じる。さらにかかるクローンは、例えば動物若しくは生体外で増殖のための選択により適応される。さらに本発明は、HCVキメラの生成を可能とし、他の遺伝子型若しくは分離体のゲノム部分は、SEQ ID NO:1又は以下の付着(deposited)実施例のようなHCVクローンの相同領域と置換される。更に別の実施例において、本発明はHCV-1、HCV-1a、HCV-1b、HCV-1c、HCV-2a、  
HCV-2b、HCV-2c、HCV-3a、及びそれらの“準種”変異体から成る群から  
選ばれるHCV遺伝子型からのポリタンパク質暗号づけ配列から成るクローンの調製方法を提供する。更なる好ましい実施例において、サイレントヌクレオチドは、特定HCVク  
ローンのマーカーとして取込まれたポリタンパク質暗号づけ領域において変化する(つまり同じアミノ酸をコード化するコドンの3つの塩基の変異)。

本発明の更なる態様において、弱毒化および欠損変異体を含むHCV核酸は、発現制御配列と操作上関連した異種遺伝子から成り、ここで異種遺伝子及び発現制御配列はプラス鎖核酸分子上に配向される。特定の実施例において、異種遺伝子は、HCVポリタンパク質暗号づけ配列とのインフレーム融合と追加的シストロンの生成から成る群から選ばれる戦略により挿入される。異種遺伝子は抗生物質耐性遺伝子又はレポーター遺伝子である。あるいは、異種遺伝子は、治療遺伝子若しくはワクチン抗原をコード化する遺伝子、つまり夫々遺伝子治療若しくは遺伝子ワクチン応用のための遺伝子である。異種遺伝子が抗生物質耐性遺伝子である特定の実施例において、抗生物質耐性遺伝子は、3'NTRにおけるSfiI部位に挿入された内部リボソームエントリ部位(IRES)と操作上関連するネオマイシン耐性遺伝子である。

上記したように当然、本発明のHCV核酸は、二本鎖DNA、ポジティブ-センスcDNA、若しくはネガティブ-センスcDNA、又はポジティブ-センスRNA又はネガティブ-RNAから成る群から選ばれる。よって本発明の核酸の特定の配列が明らかにされると、DNA及び対応するRNA双方は、ポジティブ及びネガティブ鎖のDNA及びRNAを含むことが意図される。

HCV DNAは対応するHCV RNAの翻訳用のプラスミドベクターに挿入される。よって、HCV DNAはポジティブ-センスDNA上の5'-NTRのプロモーター5'から成り、ここでプロモーターからのテンプレートDNAの転写は、複製能のあるRNAを生じさせる。プロモーターは真核性プロモーター、イーストプロモーター、植物プロモーター、バクテリアプロモーター、若しくはウイルスプロモーターから成る群から選ばれる。以下の特定例において、ファージT7及びSP6プロモーターが利用される。特定実施例において、本発明はプラスミドクローン、p90/HCVFL(長いポリ(U))に関し、これは1997年2月13日に米国のメリーランド州ロックビル、パークローンドライブにあるAmerican Type Culture Collection(アメリカンタイプカルチャーコレクション)に寄託された、SEQ ID NO:5に示される配列を有する受託番号978

10

20

30

40

50

79号と帰属される感染HCV RNA転写物を生じるように転写された完全長のHCV cDNAに宿る。当然、さらに本発明は、5'末端配列がGCCAGCC, GGCCA GCC、UGCCA GCC、AGCCA GCC, AAGCCA GCC、GAGCCA G C、GUGCCA GCC、及びGCGCCA GCCから成る群から選ばれるRNA配列に相同若しくは相補的であるプラスミドの誘導体を含み、ここで配列GCCAGCCは、SEQ ID NO: 3の5'末端と3'NTRが短いポリピリミジン領域から成る誘導体である（好ましい付着実施例では長いポリピリミジンを有するからである）。更なる実施例において、付着実施例の誘導体は、他のHCV分離体又は遺伝子型からの相同領域の置換により生じた誘導体、突然変異誘発により生じた誘導体、適応された、弱毒化された、複製能のある非感染性、及び欠損変異体から成る群から選ばれた誘導体、発現制御配列と操作上関連する異種遺伝子から成る誘導体、及び上述の誘導体のいずれかの機能フラグメントから成る誘導体から成る群から選ばれる。あるいは、5'NTR、ポリタンパク質暗号づけ領域、3'NTR若しくはHCVゲノムのより一般的な暗号づけ又は非翻訳領域のような付着DNAクローンの部分は、さまざまなHCV遺伝子型からの対応する領域で置換され、新しいキメラ感染クローンが生じる、若しくは他の分離体及び遺伝子型の拡張感染クローンにより生じる。例えば、HCV-1b若しくは-2aポリタンパク質暗号づけ領域（あるいは共通ポリタンパク質暗号づけ領域）は、付着クローンのHCV-H（1a菌株）ポリタンパク質に置換される。

当然、更に本発明は、HCV DNA若しくは上記に示したプラスミドクローンに宿る完全長のHCV cDNAから転写されたRNAを提供する。

よって、特異HCVゲノム自身はHCVの変種変異体を誘導するための優れた出発物質を提供する。なぜなら変種は変化及び誤りの蓄積からの結果から生じる人工物よりむしろ、本物のウイルスへの変化から生じるからである。本発明のHCV DNAクローン若しくはRNAは多くの方法、又は以下に明らかにするように本物のHCV成分を誘導するために利用される。

例えば、本発明は上記に再度引用されたプラスミドクローンから生産されるHCV RNAの感染量で組織培養中の細胞系統と接触させることから成るHCVとの感染を許容する細胞系統を確認する方法を提供し、細胞系統の細胞中のHCVの複製を検出する。当然、本発明は、例えば上記プラスミドにより生産されるHCV RNAの感染量を動物に導入することから成るHCVで感染を許容する動物を確認する方法にも拡張され、動物中でのHCVの複製を検出する。優性選択可能なマーカー（dominant selectable marker）から成ることが好ましい本物の感染性HCVを与えることにより、本発明は、培養中の細胞系統に接触させる、若しくはHCV RNAの感染量を動物に導入し、細胞系統又は動物におけるHCV RNAの進行的に増加するレベルを検出することから成る許容細胞系統又は動物での高いレベルのHCV複製を可能にする適応突然変異でHCVを選択する方法をさらに提供する。特定実施例において、適応突然変異はHCV向性の変形を可能にする。本発明の上記態様の直接的に意味することは、HCV感染に対する新しい有効な動物モデルの生産である。

本発明の核酸を利用して確認された許容細胞系統又は動物は、とりわけHCV感染の自然の歴史を研究する、HCVの機能成分を分離するため、及び本物のHCVウイルス若しくはそれらの成分を生産することに加えて、感度よく迅速に診断を行うために、非常に有用である。上述したように、本発明の特定の利点は、動物又は細胞系統で生産的感染を開始させることが可能なHCV DNAクローンを、始めてうまく調製できたことにある。

本発明のHCV DNA、例えばプラスミドベクターは本物のHCV成分をコード化するので、HCV DNAによりトランスフェクト、トランスフォーム、若しくはトランスデュースされた宿主細胞系統におけるかかるベクターの発現は実行される。例えば、バキュロウイルス又は植物発現システムは利用され、HCVウイルス粒子若しくはそれらの成分を発現する。よって、宿主細胞系統は細菌細胞、酵母細胞、植物細胞、昆虫細胞及び哺乳類細胞から成る群から選ばれる。

本発明は、とりわけ感染HCV RNAを提供するので、本発明は上述したプラスミドか

ら転写されたHCV RNAのようなHCV DNAの感染供給量(dose)を動物に投与することから成るHCVで動物を感染させる方法を提供する。当然、本発明は本発明のHCVで感染された非ヒト動物を提供し、非ヒト動物は前述の方法により調製される。

本発明の更なる利点は、完全な機能性HCVゲノムを与えることにより、サブユニット発現システムにおいて不可能である方法による固有HCVタンパク質又はRNAで生産された本物のHCVウイルス粒子又はそれらの成分を調製することにある。加えて、本発明のHCVの各成分は機能的であり(よって本物のHCVが発生する)、ある特定成分は本物の成分である、つまり少なくとも一部は先行技術のクローンに影響を与える誤りが欠けている。実際に、本発明の更なる利点は、天然HCVピリオン若しくはタンパク質と構造的に同じ又はかなり関連しているHCVウイルス粒子又はウイルス粒子を発生させる能力にある。よって、更なる実施例において、本発明は、本発明のHCV RNAの感染量と、例えばHCV RNAの複製を可能とする条件下において上述したプラスミドから翻訳されたHCV RNAと接触させた細胞系統を培養させることから成る生体外でHCVを増殖させる方法を提供する。当然、本発明はHCVで感染された生体内の細胞系統にまで拡張され、ここでHCVは上述したようにゲノムRNA配列を有する。特定実施例において、細胞系統は肝細胞系統である。更に本発明は、本発明のHCV感染させた非ヒト動物からのHCVウイルス粒子を分離し、HCV複製及びウイルス粒子生成を可能にする条件下で本発明の細胞系統を培養させる、又はHCV粒子タンパク質の発現を可能にする条件下で、HCV DNAでトランスフェクトさせた宿主発現細胞系統を培養し、細胞培養からHCV粒子若しくは粒子タンパク質を分離することを含み、HCVウイルス粒子を生産させるさまざまな方法を提供する。本発明は、複製成分HCVゲノムRNA、若しくは本発明のHCV核酸に対応する複製欠損HCVゲノムRNAから成るHCVウイルス粒子に拡張される。

HCV核酸、例えばDNA若しくはRNAベクターにおける異種遺伝子を挿入させることにより、本発明は、例えばHCV RNAで動物を効果的に感染させるために、HCV RNAのある量を動物へ投与することによる遺伝子治療若しくは遺伝子予防接種のために、異種遺伝子でHCV感染しやすい動物をトランスデュースする方法を提供する。特定実施例において、かかるHCVベクターは上述したプラスミドに宿るHCVにおいて発生する。

例えば前述した本発明のプラスミドから転写されたように、本発明のHCVゲノムテンプレートRNAから成るHCV用の生体外無細胞システムは、機能性HCVレピリカーゼ成分及びリボヌクレオチドリホスフェート塩基から成る等張性緩衝培地を提供する。上記要素は複製機械装置及び原材料(NTTP)を提供する。

本物のHCVウイルス粒子及びウイルス粒子タンパク質は、HCV抗原として好ましい出発物質である。よって、更なる実施例において、本発明はHCVウイルス粒子の免疫原性を動物に投与し、動物からアンチHCV抗体を分離させることから成るHCVへの抗体を生産させる方法を提供する。かかる抗体は、例えばHCVの存在を検出するために、診断的に利用される、若しくは例えば受身免疫治療において治療的に利用される。HCVへの抗体を生産させる更なる方法は、本発明のHCVウイルス粒子との反応性のヒト抗体ライブラリーをスクリーニングすることと、HCVウイルス粒子と反応性のある抗体を発現させるライブラリーからのクローンを選択することから成る。当然、抗体を発生させることに加えて、本発明の本物のHCVウイルス粒子及びタンパク質はHCVワクチン用の好ましい出発物質である。好ましくは、本発明のワクチンは薬学的に許容なアジュバントを含む。

本願で提供された本物の物質は、生体外及び生体内でHCV複製を調節することが可能な作用物質用のスクリーニング法を提供する。かかる方法は、候補物質を本発明のHCV感染動物へ投与し、HCV感染若しくは活性レベルの増加又は減少を、候補物質投与前の動物におけるHCV感染若しくは活性レベルと比較してテストすることを含み、ここで候補物質の投与前の動物におけるHCV感染もしくは活性レベルと比較して、HCV感染若しくは活性レベルの減少は、HCV感染若しくは活性を阻害する能力があることを示してい

10

20

30

40

50

る。HCV感染レベルをテストすることは、動物からの組織サンプル中のウイルスタイターを測定し、動物からの組織サンプル中のウイルスタンパク質を測定し、若しくは肝臓酵素を測定することにより実施される。あるいは、動物を感染させるのに利用されたHCVゲノムは、発現制御配列に操作上関連する異種遺伝子を含み、ここで異種遺伝子及び発現制御配列はプラス鎖核酸分子上に配向され、HCV活性レベルのテスト法は、動物からの組織サンプル中のマーカータンパク質のレベルを測定することから成る。

あるいは、かかる分析は、例えば候補物質で請求項32の細胞系統と接触させ、候補物質を投与する前のコントロール細胞系統若しくは細胞系統におけるHCV感染又は活性のレベルと比較して、HCV感染又は活性レベルの増加又は減少をテストすることにより生体外で行われ、ここで候補物質の投与前のコントロール細胞系統若しくは細胞系統におけるHCV感染又は活性レベルと比較して、HCV感染又は活性レベルが減少していることは、HCV感染又は活性を阻害させる能力があることを示している。生体外でHCV感染レベルをテストすることは、細胞、培地、若しくはその両方でウイルスタイターを測定することと、細胞、培地、若しくはその両方でウイルスタンパク質を測定することにより実施される。あるいは、細胞系統を感染させるために利用されたHCVゲノムが発現制御配列と操作上関連する異種遺伝子を含むとき、ここで異種遺伝子及び発現制御配列がポジティブ鎖核酸分子上に配向されているが、HCV活性レベルをテストする方法は、動物からの組織サンプル中のマーカータンパク質レベルを測定することから成る。

HCV複製の調節可能な作用物質をスクリーニングする更なる方法は、上述の無細胞システムを含む。この方法は、本発明の生体外でのシステムを接触され、候補物質の投与前のコントロール細胞システム若しくはシステム中のHCV複製レベルと比較して、HCV複製レベルの増加若しくは減少をテストすることから成り、ここで候補作用物質の投与前のコントロール細胞系統若しくは細胞系統中でのHCV複製レベルと比較して、HCV複製レベルが減少していることは、HCV感染又は活性を阻害する能力があることを示している。

本発明は、プラスセンスDNA上の5'から3'を接合し、一番端の5'末端保存配列から成る機能性5'非翻訳領域(NTT)と、機能性HCVタンパク質の発現のために与えるHCVタンパク質をコード化するポリタンパク質暗号づけ領域と、一番端の3'末端保存配列からなる3'非翻訳領域(NTT)から成るHCV核酸を調製する方法を含む。更に上記方法は、5'-NTT、ポリタンパク質暗号づけ配列、及びHCV分離体又は遺伝子型の少なくとも三つのクローンの大部分の配列からの3'-NTTのコンセンサス配列を決定することから成る。特定実施例において、3'-NTTは、配列5'-GGTGGCTCCATCTTTAGCCCTAGTCACTGGCTAGCTGTGAAGGTCCGTGAGCCGCGCATGACTGCGAGAGAGTGTCTGATACTGGCCCTCTCTGCTGATCATGT-3'(SEQ ID NO:4)を有するDNAに相同である一番端の末端配列から成る。更なる特定実施例において、HCV核酸は、描かれた若しくは他のHCV分離体又は遺伝子型からの相同領域の置換から成るSEQ ID NO:1に対応するポジティブ鎖配列を有する。

さらに本発明は、重要な診断上の意味を有する。一実施例において、例えばHCVウイルス粒子とサンプル中のHCV特異抗体とを結合させることを可能にする条件下において、上述のように調製された本発明のHCVウイルス粒子で被験者からの生物学的サンプルを接触させ、HCVウイルス粒子とサンプル中の抗体との結合を検出することから成る、生物学的サンプルにおけるHCVへの抗体を生体外で検出する方法を提供し、ここでHCVウイルス粒子とサンプル中の抗体との結合を検出することは、サンプル中のHCVへの抗体の存在を示す。

被験者からの生物学的サンプル中のHCVの存在を検出する生体外での代替の方法は、生産的HCV感染を許容する細胞系統と生物学的サンプルとを接触させることから成り、ここで細胞系統はHCVにより生じたトランス-作用要因(trans-acting factor)の制御下において発現されたレポーター遺伝子生成物を発現する導入遺伝子(transgene)を含むように変種され、レポーター遺伝子生成物の発現を検出し、さらにここでは、レポータ

10

20

30

40

50



ー遺伝子生成物の発現を検出することは、被験者からの生物学的サンプル中にHCVが存在することを示している。関連した実施例において、本発明は、生物学的サンプルと生産的HCV感染を許容する細胞系統とを接触させ、レポーター遺伝子生成物の発現を検出することから成る被験者からの生物学的サンプル中のHCVの存在を生体外で検出する方法を提供し、ここで細胞系統はHCVにより生じるトランス-作用要因の制御下において高いレベルでレポーター遺伝子生成物を発現する欠陥ウイルス導入遺伝子を含むように変種され、レポーター遺伝子生成物の発現の検出は被験者からの生物学的サンプル中にHCVが存在していることを示す。よって、本発明の重要な利点は、上記の生体外での診断用の許容（感染しやすい）細胞系統を提供することにある。請求項64による方法では、欠陥ウイルス導入遺伝子は特別に設計されたアルファウイルスを生じ、トランス-作用ヘルパー

10

要因はアルファウイルスnsP4ポリメラーゼであり、ここでアルファウイルスnsP4ポリメラーゼは、HCV NS4Aとのキメラ融合タンパク質として発現され、アルファウイルスnsP4ポリメラーゼ-HCV NS4Aキメラ融合タンパク質はHCV NSプロテイナーゼにより切断され、機能性アルファウイルスnsP4ポリメラーゼを放出する。前述の方法において生物学的サンプルは血液、血清、血漿、血球、リンパ球、及び肝臓組織バイオプシーから成る群から選ばれる。

関連した態様において、さらに本発明は本物のHCVウイルス成分から成るHCV用テストキットと、本物のHCVウイルスから誘導される成分から成るHCV用診断テストキットを提供する。

よって本発明の主要目的は、感染HCVをコード化するDNAを提供しなければならない

20

。本発明の関連目的は、DNAクローンからの感染HCVゲノムRNAを提供することである。

本発明の更に別の目的は、細胞に侵入するが感染ウイルスを増殖させないワクチン開発に適する弱毒化されたHCV DNA若しくはゲノムRNAを提供することである。

本発明の別の目的は、アンチHCV（または抗ウイルス性）薬剤をテストし、薬剤耐性を評価し、弱毒化HCVウイルスワクチンをテストするためのHCV感染の生体外及び生体内モデルを提供することである。

本発明の更に別の目的は、中和検定においてHCVレポーター、アンタゴニストに結合するレポーターを確認するのに利用されたHCVビリオン若しくはウイルス粒子タンパク質の発現を提供することである。加えて、発現HCVビリオン若しくはウイルス粒子タンパク質は、固有HCVと構造的に同じ若しくはほとんど関連している抗原により、より効果的なHCVワクチンを開発するために利用される。


30

本発明の更なる目的は、特別設計されたレポーター細胞を用いて感染HCVを検出する能力に基づくHCV診断薬を提供することである。

さらに別の目的は、HCV特異抗体を分析し、HCV特異抗体を発生させるために、本物のウイルス抗原、特にウイルス粒子を提供することである。

本発明の上記及び他の目的は、図面と本発明の詳細な説明により詳述される。

#### 【図面の簡単な説明】

図1（先行技術）。HCVゲノム構造、ポリタンパク質プロセッシング、とタンパク質の特徴。上には構造及び非構造タンパク質暗号づけ領域と、3'及び5'NTRと、推定される3'の第二の構造のあるウイルスゲノムを示す。ゲノムの下の箱は、タンパク質分解プロセッシングカスケードにより生じたタンパク質を示す。推定される構造タンパク質は陰のある箱（shaded box）で示され、非構造タンパク質はオープンな箱（open box）で示される。電荷のないアミノ酸の隣接した部分の広がり黒のバーで示される。星印はN結合グリカンを示すが、位置若しくは利用した部位の数を必ずしも示さない。図示した切断部位は、宿主シグナラーゼ（）、NS2-3プロテイナーゼ（曲がった矢印）、NS3-4Aセリンプロテイナーゼ（下向き矢印）に切断させた。

40

図2。HCVベクターを利用する異種RNA及びタンパク質の発現の戦略。図2の上部は、単一の長いORF及びタンパク質分解プロセッシングの翻訳により成熟ウイルスタンパク

50

質を発現させるプラス極性RNAウイルスHCVの図である。構造タンパク質 (STRUCTURAL) 及び非構造タンパク質 (REPLICASE) をコード化するポリタンパク質の領域は、夫々薄く陰のある、オープンな箱として示される。それ以下には、多くの提案された複製能のある“レプリコン”発現構成物 (construct) を示す。第一の四つの構成物 (A-D) は構造遺伝子が欠けており、したがってパッケージングを感染ウイルスへ入れることを可能とするヘルパーシステムは必要である。構成物 E-G は複製若しくはパッケージング用のヘルパー機能は必要ない。黒く陰のある箱は異種若しくは外来遺伝子配列 (FG) を示す。翻訳開始 (aug) と終結シグナル (term) はオープン三角形と黒い菱形で夫々示される。内部リボソームエントリ部位 (IRES) は垂直なストライプのある箱として示す。構成物 A 及び H は HCV ポリタンパク質とインフレーション融合としての異種生成物の発現を示す。かかるタンパク質融合連結は、プロセシングが宿主若しくはウイルスプロテイナーゼ (矢印で示す) のいずれかにより媒介されるように特別に設計された。

10

図 3。HCV 感染を分析するための特別設計された細胞系統。パネル A は三つのサイレント導入遺伝子が発現させる細胞を示す。(i) シンドビスウイルス (Sin) 非構造タンパク質 4 (nsP4) と融合した HCV NS4A から成るポリタンパク質のタンパク質を発現する mRNA と、(ii) nsP4 暗号づけ領域を欠く欠損シンドビスウイルスレプリコンであるが、レポーター遺伝子の発現を駆動させるサブゲノミックプロモーター (矢印) と、(iii) nsP4 を欠くが、サブゲノミック RNA プロモーターの制御下においてユビキチン-nsP4 融合遺伝子を含む欠損シンドビスウイルス RNA は、核プロモータ要素により駆動される。シンドビスレプリコン及び欠損 RNA は、シンドビスウイルス特異 RNA 複製に必要な全てのシグナルを含み、転写及びパッケージングシグナル (ステムループ構造) は活性 nsP4 が存在しないとサイレントである。パネル B。HCV により感染しやすい細胞の生産的感染において、ウイルスは脱外被され、翻訳され、複製を始める (段階 1)。このことにより、活性 NS3 セリンプロテイナーゼ (段階 2) を生産し、HCV NS4A-シンドビス nsP4 連結 (段階 3) で切断させ、活性 nsP4 が生じる。他の三つのシンドビス nsP4 のある nsP4 集合は、活性シンドビス複製複合体を形成し (段階 4)、両方のシンドビス特定 RNA を複製し、シンドビスウイルスサブゲノミックプロモーターからの転写を導く (段階 5)。欠損 RNA のサブゲノミック RNA から発現された Ub-nsP4 は切断され、nsP4 ポリメラーゼのより活性な形を形成し、さらにシンドビス特異 RNA の複製および転写を増幅させる (段階 6)。これらレポーターの高いレベルの遺伝子発現をもたらす (段階 7)。

20

30

図 4。チンパンジーモデルでテストされた構成物の初期の組 (チンパンジー実験 I)。正しい HCV 5' 及び 3' 末端がクローン化され決定される前に、チンパンジーモデルでクローンのテストを行った。図は T7 若しくは SP6 プロモーター要素を示し、HCV cDNA 及びランオフ (run-off) 部位は、ポリ (A) 若しくはポリ (U) で終結する転写物の生産のために利用された。

図 5 (A 及び B)。(A) コンビナトリアルライブラリー用に増幅された HCV H77 の領域。上部には、HCV H cDNA のダイアグラムを、示されたコンビナトリアルライブラリー (KpnI と NotI: オープンな箱) をクローン化させるために利用された制限部位と共に示す。領域はレシピエントベクター、pTET/HCV BgIII/5' +3' corrへクローン化された。このレシピエントベクターは黒で示す 5' 及び 3' 末端領域の HCV H77 のコンセンサス配列を含む。おおよそのタンパク質境界も示される。以下には、HCV H77 RNA から RT-PCR により増幅されたフラグメントを、A から G として示す。各セグメントの上にある数字は、ライブラリーにおける領域の最小の複雑度 (complexity) を示す。プライマーペアと正確な位置を表 2 及び 3 に示す。(B) コンビナトリアルライブラリーの構築における中間体と最終フラグメント。表 2 及び 3 に示すように、とりわけ、アセンブリ PCR 過程での中間体と HCV cDNA におけるおおよその位置を示す。

40

図 6。アセンブリ PCR 法。アセンブリ PCR 法の一般的なスキームを示す。アセンブリ

50

に利用した特異 H C V フラグメント及びプライマーを表 3 に示す。

図 7。c D N A 希釈の P C R による複雑度決定の例。増幅領域 A、D、及び G において、第一の鎖 c D N A のさまざまな希釈溶液は、P C R によるうまい増幅でチェックされた。生成物はアガロースゲル上にて分析された。この分析から、コンビナトリアルライブラリーにおける上記領域の最小の複雑度は、夫々 c D N A の 8 0、1 0 及び 1 0 分子であった。

図 8。長いポリ ( U / U C ) トラクトを通した効率的な転写の分析。生体外での H C V R N A s の最適転写用の条件を利用して、幾つかのテンプレート D N A s からの転写生成物を示す。( A ) レーン 1、スーパーコイル p T E T / H C V F L C M R / 5 ' 3 ' corr . D N A ; レーン 2、X m m - 消化 p T E T / H C V F L / C M R / 5 ' 3 ' corr . テンプレート ( 予想されるサイズは 1 1 7 4 0 の塩基である ) ; レーン 3、H p a I - 消化 p T E T / H C V F L C M R / 5 ' 3 ' corr . テンプレート ( 予想されるサイズは 9 6 0 0 の塩基である ) ; レーン 4 及び 5、夫々 1 1、7 5 0 と 9 4 0 0 の塩基の転写された R N A サイズマーカー。転写反応は 3 m M U T P 及び 1 m M A、G、と C T P を含む。( B ) レーン 1、B s m I - 消化 p 9 2 / H C V F L 長い p U / 5 ' G G D N A ( 予想されるサイズは約 9 6 0 0 の塩基である ) ; レーン 2、X b a I - 消化 p 9 2 / H C V F L 長い p U / 5 ' G G D N A ( 予想されるサイズは約 1 3 0 0 0 塩基である )。パネル B における転写反応は、3 m M での全て四つの N T P を含む。両パネルとも、ポリ ( U / U C ) トラクトで終結する H C V R N A 転写物は長さ約 9 5 0 0 の塩基であろう。両パネルにおけるレーン M は H i n d I I I - 消化ラムダ D N A サイズマーカーである。

図 9。H C V H 7 7 コンセンサス配列の決定用の配列アラインメント。決定された H C V H 配列のアラインメントを示す。各ブロックの下部でのヌクレオチド及びアミノ酸配列は、H C V H C M R 基本形配列である。コンビナトリアルライブラリーからの配列クローンの数は左に示す ( S E Q I D N O : 1 9、2 0 )。G e n B a n k は Inchausti によって決定された H C V - H 配列である ( Proc. Natl. Acad. Sci. USA 88: 10292, 1991; Accession #M67463 )。" cons. " は H C V H 7 7 コンセンサス配列 ( S E Q I D N O : 1 ) を示す。H C V H C M R 配列に同じ位置はドットで示し、ある配列領域のギャップは点線で示す。差が発見されたところでは、下のケースレターはサイレントヌクレオチド置換を示し、上のケースレターは特定ヌクレオチド置換により遺伝暗号指定の変化を招くことを示す。

図 1 0。コンセンサスクローンの直接構築における段階。図はコンセンサスクローンの直接構築に利用された各配列クローンの領域を示す。各クローンからの一次フラグメントは斜め線が引かれた箱により示され、中間構築サブクローンはオープン箱として示され、完全長のコンセンサスクローンの構築に利用された最終のクローン及び領域は、陰のある箱で示される。表 4 はクローニング段階の詳細をまとめたものである。

図 1 1。チンパンジーでテストされた 1 0 の完全長クローンの特徴 / マーカー 実験 I I I。上部には、H C V H 7 7 c D N A コンセンサス R N A のスキームを示す。うまくチンパンジーに接種させた実験に利用された 1 0 の R N A 転写物を以下に図解する。追加の 5 ' ヌクレオチドと " 短い " 対 " 長い " ポリ ( U / U C ) トラクトを示す。全てのクローン / 転写物は、マーカーとして 2 つのサイレントヌクレオチド置換を含む。位置 8 9 9 ( T に代わり C ; 星印で示す ) ; 位置 5 9 3 6 ( A に代わり C ; 円の星印で示す )。追加の 5 ' 塩基のあるクローンは、位置 5 1 4 ( 三角形 ) の X h o I を不活性化させる突然変異を含む。" 短い " 対 " 長い " ポリ ( U / U C ) トラクトのあるクローンは、A ( 黒点 ) 対位置 8 0 5 4 での G により夫々区別される。

図 1 2。接種動物からの血清サンプルには繰越 ( carryover ) テンプレート D N A を含まない。図示するように、添加した競争者 R N A のない ( レーン 1 )、若しくは  $10^2$  ( レーン 2 から 7 ) 又は  $10^3$  ( レーン 8 から 1 4 ) 分子のある接種後の指摘した週間からの重複 R N A サンプル (  $10 \mu l$  血清 ) は、逆転写段階での ( + ) 若しくは ( - ) 酵素なしで R T - P C R により増幅された ( Kolykhalov らによる J. Virlo. 70: 3363 ( 1996 ) )。c D N A 合成が存在しないと、特定 P C R 帯は検出されず、H C V 特異核酸シグナルは R N

Aによるものであることを示す。図示した分析はチンパンジー # 1 5 3 5 に対してであり、接種 H C V R N A の高いレベルを受け取り、テンプレート D N A は D N a s e I による消化では分解されない。

図 1 3。接種動物からの循環 H C V R N A は R N A s e から保護される。レーン 1 において、 $10 \mu\text{l}$  の血清は  $3 \times 10^5$  分子の競争者 R N A と混合され、 $0.5 \mu\text{g}$  の R N a s e A で室温 1 5 分間消化させ、R N A z o l で抽出され、(Kolychalov による 1996 の上記文献) に記載されている入れ子 (nested) R T - P C R に利用された。レーン 2 に示すサンプルでは、競争者 R N A は R N A z o l ( R N A s e 処理されていない) による溶菌後に添加された。レーン 3 において、競争者 R N A なしの  $10 \mu\text{l}$  血清は、レーン 1 に示すような R N A z o l による抽出前に R N A s e A で予め消化させた。レーン 4 は R T - P C R のネガティブコントロール (negative control) である。実験から、トランスフェクトされたチンパンジーからの物質を含む H C V R N A は、過剰の競争者 R N A が完全破壊された (destroyed) 条件下において R N a s e 耐性があることを示している。分析されたサンプルは、R N A タイターが  $6 \times 10^6$  分子 / m l であった 6 週間後のチンパンジー # 1 5 3 6 からのものであった。

#### 発明の詳細な説明

以上に指摘したように、好ましくは、本発明は本物の C 型肝炎ウイルス ( H C V ) 核酸、例えば D N A 又は R N A、クローンを提供する。本発明の機能性 H C V 核酸は、感染しやすい動物及び細胞系統を提供することが好ましい。根気強い努力にもかかわらず、従来から感染性のある H C V はうまくクローン化されておらず、よってウイルスの複製機構の系統的評価、レポーター結合及び細胞侵入、生体外及び生体内での分析システムを利用した坑ウイルス治療薬の開発、及び生体外での診断分析システムの開発を妨げてきた。加えて、本発明のクローンにより、適切なプロセッシングを可能とする条件下において、H C V 粒子及び粒子タンパク質の発現、よって固有 H C V にかなり近い構造類似性を有するタンパク質の発現を可能とする。かかる粒子及びタンパク質は、アンチ H C V ワクチン開発にとって好ましい。加えて、感染に必要とされる H C V ゲノムの要素を確認することにより、本発明者らは遺伝子治療ベクターの調製のために、慢性肝臓感染を導く H C V の性質を利用する。多くのタンパク質源であり、よって被験者の中で欠失を補うための溶解性要因の発現を好ましい器官で行うベクターは肝臓を目標とするので、かかるベクターは特に有用である。

本発明は、一部機能性遺伝子型 1 a c D N A クローンの発生を基礎としており、他の H C V 遺伝子型 (例えば同様な方法を利用して組立られ調べられる) のための調製の基礎として、利用可能である。上記生成物は、( i ) より効率的な H C V 治療; ( i i ) H C V ワクチン; ( i i i ) H C V 診断薬と; ( i v ) H C V に基づく遺伝子発現ベクターの開発のためのさまざまな応用がある。上記応用例は以下に説明する。

本発明は、H C V コンセンサス配列の決定と複製能のある感染 R N A 転写物を産することが可能な完全長 H C V c D N A クローンを作り上げるための情報の利用について説明する。 $5'$  及び  $3'$  末端でのかなり保存された配列を含む末端配列の決定、H C V c D N A クローンを増幅させ構築させるための誤りの少ない方法の利用、及びコンセンサス配列を反映させたクローンの構築の全ては、本発明の成功に寄与する。

用語 “ 本物の ” とは、本願では H C V 核酸のことを言い、D N A (つまり c D N A) 若しくは R N A は、完全なゲノム複製及び機能性 H C V タンパク質の生産、若しくはそれらの成分を与える。特定実施例において、本物の H C V 核酸は、例えばチンパンジーモデル若しくは組織培養中で感染性があり、ウイルス粒子 (つまりビリオン) がチンパンジーモデルと組織培養の双方で形成される。しかしながら、本発明の本物の H C V 核酸は、幾つかの (全てではないが) 機能性 H C V タンパク質を生産するように、若しくはヘルパー細胞系統又はプラスミドなどの存在なしに、複製せずに細胞を生産的に感染させるように弱毒化される。本出願で例証された本物の H C V は、生体内での野生型ウイルスの複製に相当する H C V 複製サイクルの開始に必要である、全てのウイルスをコード化させる (virus-encoded) 情報を R N A 要素若しくはコード化されたタンパク質のいずれかを有する。本

願で説明した又は例証された A T C C 及びその変異体で付着された実施例を含む、本願で説明した特異 H C V クローンは、H C V 治療、ワクチン、診断薬及び発現ベクターを開発するための好ましい出発物質である。特に、本発明の H C V 核酸の利用は、本物の H C V 成分が含まれていることを保証する。なぜなら、先行技術のクローン化された H C V とは異なり、上記成分は共同して感染性タンパク質を与えるからである。本願で説明する特定出発物質、及び好ましくは本物の H C V c D N A に宿る付着プラスミドクロンは、本願で説明するように変種され、例えば部位特異的変異の導入 (sited-directed mutagenesis) により変種され、欠損又は弱毒化された誘導体が生じる。あるいは、他の遺伝子型又は分離体からの配列は、本願で説明された特定実施例の相同配列にとって置換される。例えば、本発明の本物の H C V 核酸は、本願で開示された、例えばレシピエントプラスミド上のコンセンサス 5' 及び 3' 配列から成り、他の分離体又は遺伝子型 (コンセンサス領域若しくは非常に高い忠実度のクローニングにより得られた領域) からのポリタンパク質暗号づけ領域は、本願で例証された H C V の相同性ポリタンパク質暗号づけ領域に置換される。加えて、一番端の 5' 若しくは 3' 配列、又は双方を含みことに限らず、特定実施例における H C V ウイルス粒子及び R N A 複製機械装置の機能性成分からの切断生成物のポリタンパク質をコード化する O R F を含むことに限らず、本願で説明した本当の H C V の一般的特性は、本物の H C V クローンを得るために提供された特定分離体又は遺伝子型のコンセンサス配列を取込む。

10

特に、本発明は、例えば正真正銘の複製をすることができず、H C V タンパク質を生産することができず、ノーザン分析により検出される H C V R N A を生産せず、動物若しくは生体外の細胞系統を感染させることができない、変種若しくは“補正 (correcting)”非機能性 H C V クローンを提供する。例えば S E Q I D N O : 1 の c D N A 配列のような本発明の本物の H C V 核酸配列と非機能性 H C V クローンの配列とを比較することにより、非機能性クローンの欠損は確認され補正される。当業者に入手可能な核酸配列を修正させる全ての方法は、部位特異的変異の導入と、例えば非機能性クローンにおける相同配列の S E Q I D N O : 1 の本当の H C V クローンからの機能性配列の置換とに限られないものを含む非機能性 H C V ゲノムでの変種をもたらす。

20

用語“コンセンサス配列”とは、5'-N T R、ポリタンパク質暗号づけ配列機能又はその部分及び H C V の菌株又は遺伝子型のクローンに無関係に、三つ又はそれ以上、好ましくは六つ又はそれ以上からのコンセンサス残基を確認することにより決定された 3'-N T R を含む H C V ゲノム配列又はその部分のことを言及するために本願では利用する。以下の例において、5'-N T R (ポリタンパク質暗号づけ領域からの幾つかのキャプシドタンパク質を含む) 及び 3'-N T R (ポリタンパク質の C 末端をコード化するゲノムの幾つかの部分を含む) コンセンサス配列は、レシピエントプラスミド (例 3) において決定され、取込まれた。K p n I 部位から N o t I 部位までのポリタンパク質暗号づけ領域の大部分のコンセンサス配列も、以下の図 8 及び例 4 にて示されるように決定されコンセンサス配列を与える。ポリタンパク質暗号づけ配列の K p n I 及び N o t I 部分の挿入は、コンセンサス 5' 及び 3' コンセンサス配列を含むレシピエントプラスミドにて行われ、本物の H C V ゲノム D N A クローンを与える。

30

本発明の本物の H C V 核酸は、H C V 核酸の機能性活性に影響を与えずに保存配列の追加の塩基上流を有する、5' 末端配列 G C C A G C C から成る 5'-N T R の一番端の保存配列を含むことが好ましい。好ましい実施例において、5'-G C C A G C C は 0 から 10 の追加の上流塩基、好ましくは 0 から 5 の上流塩基を、さらに好ましくは 0、一つ、二つの上流塩基を含む。特定実施例において、一番端の 5' 末端配列は G C C A G C C ; G G C C A G C C ; U G C C A G C C ; A G C C A G C C ; A A G C C A G C C ; G A G C C A G C C ; G U G C C A G C C ; 若しくは G C G C A G C C であり、ここで配列 G C C A G C C は S E Q I D N O : 3 の 5' 末端である。

40

本発明の本物の H C V 核酸において、3'-N T R は長いポリピリミジン領域から成る。ポジティブ鎖 H C V R N A において、その領域はポリ (U) /ポリ (U C) トラクトに相当する。当然、ポジティブ鎖 H C V D N A において、これはポリ (T) /ポリ (T C

50

）トラクトである。以下の例では、ポリピリミジントラクトは可変の長さである；長いポリ（U/U C）トラクトを含むH C V クローンがかなり感染力があると分かったが、短い（約75の塩基）と長い（133の塩基）ものは効果的である。長いトラクトは自然にH C V 分離体が発生するときに発見された。よって、本発明の本当のH C V 核酸可変長さのポリピリミジントラクトを有する。

本発明の特定実施例において、ファージプロモータの制御下において感染性H C V R N A をコード化するc D N A に宿るプラスミドp 9 0 / H C V F L （長いポリ（U））は、37 C . F . R . セクション1 . 8 0 1 以下参照の規定にしたがい、特許保護を目的とした微生物の寄託の国際認識であるブダベスト条約を順守する目的で、ワシントン大学のメディカルスクールを代表して、1997年2月13日に米国メリーランド州のロックビル、パークローンドライブ 12301にあるAmerican Type Culture Collection (ATCC) に寄託した。

上記技術の利益はものすごく、簡単には到達できない。直接的な重要性は、生物化学、細胞培養、及び遺伝子導入動物手法を利用して、個々のH C V タンパク質及びR N A 要素の機能の研究用の出発物質として、上記機能性クローンからのH C V c D N A の使用である。c D N A 合成若しくはP C R 増幅中に誤って導入されることがないので、ネガティブ若しくは誤解を招く結果を得る可能性は最小化されるであろう。さらにかかるクローンは、本物のウイルス複製に関係するH C V 生物学の多くの面にわたる将来の分子遺伝子研究用の明確な出発物質を提供する。治療及びワクチン開発に関係する利用は、( i ) ウイルス中和、吸着、浸透及びエントリの生体外及び生体内検定を開発するための明確なH C V ウイルスストックの発生；( i i ) H C V タンパク質、R N A 要素及び新しい坑ウイルスターゲットの構造/機能の研究；( i i i ) H C V R N A 複製及び粒子放出を支援するシステムや条件を確認するための細胞培養システム及び条件の系統的調査；( i v ) 細胞培養中でのより効率的な複製を可能とする適合H C V 変種体の生産；( v ) 変質組織若しくは種向性 (species tropism) のあるH C V 変種体の生産；( v i ) H C V 複製を支援するモデルを含む阻害評価の代替の動物モデルの確立；( v i i ) 無細胞H C V 複製分析の開発；( v i i i ) ワクチン用の免疫H C V 粒子の生産；( i x ) 可能なワクチン候補としての弱毒化されたH C V 誘導体のエンジニアリング；( x ) 遺伝子治療及びワクチン応用のための異種遺伝子生成物の発現用の弱毒化された若しくは欠損H C V 誘導体；( x i ) 適切な受容体による肝臓若しくは他の細胞系統への治療剤の目標へのデリバリーのためのH C V 糖タンパク質の利用を含む。

本願で利用されたさまざまな用語は以下に定義される。

フレーズ“薬学的に許容な”とは、ヒトに投与した際に、生理的に許容であり、胃の不調、めまいなどのアレルギー性若しくは同様な不適当な反応を一般には生じない分子構成要素及び成分のことを言う。本願で用いられるように、用語“薬学的に許容な”とは、連邦政府若しくは州政府、又は米国薬局に掲載された、又は動物、特にヒトに使用するための薬局で一般的に認められ、調整委員 (regulatory agency) により承認されていることを意味している。用語“キャリア”とは、投与される化合物との希釈剤、アジュバント、賦形剤、又は溶剤のことである。かかる薬学的なキャリアは、水や、石油、ピーナッツオイル、大豆オイル、ミネラルオイル、ゴマオイルなどのような動物、野菜又は合成物であるオイルのような、無菌溶液である。水又は水溶液、生理食塩水や水溶性デキストローズやグリセリン溶液は、キャリアとして、特に注射可能な溶液として用いられることが好ましい。適切な薬学的キャリアはE. W. Martinによる“Reminton's Pharmaceutical Sciences”に記載されている。

フレーズ“治療的に効果のある量”とは、活性、機能及び宿主の応答における臨床的に重要な欠損を、少なくとも約15パーセント、好ましくは少なくとも50パーセント、より好ましくは少なくとも90パーセント減少させ、最も好ましくは防止させるのに十分な量を意味する。あるいは、治療的に効果のある量は、宿主における臨床的に重要な状態の改善を生じさせるのに十分な量である。

用語“アジュバント”とは、抗原に対する免疫応答を増大させる化合物若しくは混合物の

ことである。アジュバントは、抗原をゆっくりと放出する組織デポー剤として、さらに免疫応答を特別にではなく増大させるリンパ系活性化剤として働く。(HoodらによるImmunology, Second Ed., 1984, Benjamin/Cummings: Menlo Park, California, p.384)。しばしば、アジュバントの存在しない場合での抗原とのみの主なチャレンジは、ヒト又は細胞免疫応答を引き出せない。アジュバントは、これらに限らないが、フロイント完全アジュバント、不完全なフロイントアジュバント、サポニン、水酸化アルミニウムのようなミネラルゲル、リゾレシチン、プルロニック(pluronic)ポリオール、ポリアニオン、ペプチド、オイル又は炭化水素エマルジョン、キーホールリンペット(keyhole limpet)ヘモシアン、ジニトロフェノールのような表面活性化物質とBCG(bacille Calmette-Guerin)及びコリネバクテリウムパルブムのような潜在的に有用なヒトアジュバントを含む。アジュバントは薬学的に許容であることが好ましい。

10

特定実施例において、用語“約”若しくは“おおよそ”とは、特定の値若しくは範囲の20%以内、好ましくは10%以内、より好ましくは5%以内であることを意味する。前述に開示したものを更に詳述する本願の以下のサブセクションは、発明を限定させるものではなく、都合を考慮して提供される。

#### 他のHCV分離体及び遺伝子型の機能性完全長クローン

本願で説明した手法を利用して、他のHCV遺伝子型の機能性完全長クローンが形成され、生物学的研究と坑ウイルススクリーニング及び評価に利用される。本発明の拡張において、ライブラリーは、高いRNAタイター( $10^6$ 以上)と周知の臨床歴史のある単一エクスポート患者からのRNAを利用して組立てられた。分離体のコンセンサス配列は、ライブラリーの個々のクローンの配列から生じる。プロモータと、5'及び3'末端コンセンサス配列(分離体から若しくは例えばHCV-H77のような異なる分離体のいずれかから決定された)とランオフ転写物の生産の3'制限部位を含む新しいレシピエントプラスミドが作られる。

20

コンビナトリアルライブラリーからの限定数のクローンのスクリーニングが、あまり誤りのない方法として出現し機能性クローンを与える。あるいは、本願に説明したように、多重クローンと方向性のある(directed)構築から誘導された配列は、機能性コンセンサスクローンを生産するのに利用される。

よって、本発明は他のHCVゲノム配列、若しくはコンセンサスゲノム配列の分離を考えている。本発明によれば、従来の分子生物学、微生物学、と現在の技術水準にある組換えDNA技術を利用する。かかる技術は文献に十分に説明されている。例えばSambrook, Fritsch&Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York(本願ではSambrookらによる1989の文献); DNA Cloning: A Practical Approach, Volumes IとII(D. N. Glover編集1985); Oligonucleotide Synthesis(M. J. Gait編集1984); Nucleic Acid Hybridization(B. D. Hames&S. J. Higgins、編集(1985)); Transcription And Translation(B. D. Hames & S. J. Higgins編集(1984)9; Animal Cell Culture(R. I. Freshney, 編集(1986)); Immobilized Cells And Enzymes(IRL Press, (1986); B. Perbal, A Practical Guide To Molecular Cloning(1984); F. M. Ausubelらによる編集, Current Protocols in Molecular Biology, John Wiley & Sons, Inc.(1994)。

30

40

したがって、本願で現れるならば、以下の用語は次に示す定義を有する。“3'末端配列要素”、“3'末端”“3'配列要素”のような用語HCV配列は、以下の配列を全て含むことを意味する;(i)ポジティブ-センスゲノムRNAのRNA配列;(ii)例えばHCVネガティブ-センスRNAのような上記RNA配列の補体;(iii)RNA要素のポジティブ-センス配列に対応するDNA配列;(iv)RNA要素のネガティブ-センス配列に対応するDNA配列である。したがって、実質的に等価若しくは変質を示すヌクレオチド配列も同様に考えられる。上記変種体は、例えば部位特異的変異の導入を介して得られた変種体のように計画的であり、若しくは複合体又はその名のサブユニットの生産者である宿主における突然変異を介して得られた変種体のように偶発的である。

“ベクター”は、他のDNA(RNA)セグメントが参加し、付着セグメントの複製が起

50

こるプラスミド、ファージ、又はコスミッドのようなレプリコンである。“カセット”とは、ベクターの特定制限部位に挿入されたDNA RNAのセグメントのことである。DNA又はRNAのセグメントは関心のあるポリペプチド若しくはRNAをコード化し、カセット及び制限部位は転写及び翻訳の適切なリーディングフレーム中のカセットに挿入されることを確実にするように設計される。

転写及び翻訳制御配列は、プロモーター、エンハンサー、ポリアデニレーションシグナル、ターミネーター、IRES要素などのようなDNA若しくはRNA調節配列であり、宿主細胞暗号づけ配列の発現が生じる。暗号づけ配列は、RNAポリメラーゼが暗号づけ配列をRNAへ転写するときに細胞内の転写及び翻訳制御配列の“制御下に”若しくは“操作上関連して”いる。さらにRNA配列は、翻訳、RNA安定性、RNA複製、及びRNA転写(RNAウイルスのために)を調節する能力により発現制御配列として働く。

“プロモーター配列”は、細胞中でRNAポリメラーゼと結合し、下流(3'方向)暗号づけ配列又は非暗号づけ配列の転写を開始することができるDNA又はRNA調節領域である。よって、プロモーター配列は、RNAウイルス複製と転写における類似のRNA配列又は同様な機能の構造のことを言うことにも利用される。本発明の感染性HCV DNAクローンの無細胞又はバクテリア発現のための好ましいプロモーターは、ファージプロモーターT7、T3とSP6である。あるいは、サイトメガロウイルスの直接早期の(immediate-early)プロモーターのような核プロモーターが利用される。実際には、利用したシステムに依存して、発現は真核、原核又はウイルスプロモーター要素から引き起こされる。HCV RNA発現用のプロモーターはキャップのある(capped)又はキャップのない(uncapped)転写物を与える。

本願で利用されているように、全ての文法的形及びスペルの変形における用語“相同である(homologous)”とは、スーパーファミリー(例えば免疫グロブリンスーパーファミリー)からのタンパク質と、異なる種(例えばミオシンI鎖など)からの相同なタンパク質を含む“共通の進化起源”を有するタンパク質間の関係をさすことをいう(ReekらによるCell 50: 667(1987))。かかるタンパク質(及び遺伝子をコード化するもの)は、高度な配列類似性を有している。全ての文法的形における用語“配列類似性”とは、核酸配列間若しくは共通の進化起源を所有するか否かのタンパク質のアミノ酸配列間の同一又は一致の程度のことをいう(Reekらによる上記文献を参照)。しかしながら、共通の使用及び緊急の出願において、“実質的に”若しくは“かなり”のような副詞と共に修飾されたときには、用語“相同である”とは配列類似性のことをいい、共通の進化起源を有するものではない。

特定実施例において、二つのDNA又はRNA配列は、DNA配列の定義された長さ上のヌクレオチドの少なくとも約50%(好ましくは少なくとも約75%であり、さらに好ましくは少なくとも約90又は95%である)がマッチすれば、“相同であり”若しくは“実質的に類似”している。実質的に相同である配列は、配列データベース入手可能な標準ソフトウェアを利用して、又は例えば特定のシステムで定義される緊縮条件下でのザンハイブリダイゼーション実験から、配列を比較することにより確認される。適切なハイブリダイゼーションの条件を制限することは、当業者には可能である。Maniatisらによる上記文献; DNA Cloning, Vols. I & II, 上記文献; Nucleic Acid Hybridization, 上記文献。

同様に、特定実施例において、30%以上のアミノ酸が同じである、若しくは約60%以上が類似である(機能的に同一である)ときに、二つのアミノ酸配列は“相同である”若しくは“実質的に類似”である。好ましくは、類似若しくは相同である配列は、例えばGCG((Genetic Computer Group, Program Manual for the GCG Package, Version 7, Madison, Wisconsin)パイルアップを利用したアライメントにより確認される。

核酸又はアミノ酸に関して、用語“対応する”とは、本願では類似若しくは相同な配列を言うときに利用され、正確な位置が同一であるか、分子から類似性又は相同性が異なるかどうか測定される。核酸若しくはアミノ酸配列アライメントはギャップを含む。よって、用語“対応する”とは配列類似性若しくは相同性の領域のことをいい、アミノ酸残基若

10

20

30

40

50



しくはヌクレオチド塩基のナンバリングではない。

H C V ゲノム核酸は感染性 H C V 源から、特に組織サンプル（血、血漿、血清、肝臓検体、白血球）、感染ヒト若しくはサル、又は許容動物種から分離される。ゲノム H C V クローン又はそれらの部分を得る方法は、上述したように（Sambrookらによる1989の上記文献を参照）本技術分野においては周知である。ポリタンパク質暗号づけ領域配列を含む H C V 分離体は、例えばHoughtonらによる1989年6月1日発行の国際特許出願第W O 8 9 / 0 4 6 6 9 号に；Houghtonらによる1990年10月4日発行の国際特許出願第W O 9 0 / 1 1 0 8 9 号に；Houghtonらによる1994年9月27日発行の米国特許第5、350、671号に；Miyamotoらによる1994年12月13日発行の米国特許第5、372、9238号に；Choらによる1993年1月7日発行の欧州特許第E P 0 5 2 1 3 1 8

10

A 2 号に；及び1992年10月28日発行の欧州特許第E P 0 5 1 0 9 5 2 A 1 号に記載されており、各文献は本願の引用文献全体の中に取込まれる。さらに代表的な遺伝子型は、決してこれに制限されないが、他の分離体1b、1c、2a、2b、2c、3aなどを含む（Bukhらによる（1995）の上記文献；Simmonds, Hepatology 21: 570-83（1995）；SimmondsらによるHepatology 19: 1321-1324（1994）；SimmondsらによるJ. Gen. Virol. 77: 3013-3024（1996））。多くのサブタイプと遺伝子型にとって、十分な配列データはR T / P C R と P C R アセンブリのプライマーを設計するために利用可能である。

分子クローニングゲノム H C V R N A 又は D N A において、D N A フラグメントが、例えば c D N A への逆転写と P C R により発生する。上記フラグメントは集められ完全長の配列を形成する。多くのかかるフラグメントを調製することにより、H C V クローンのコンピナトリアルライブラリーが得られる。かかるライブラリーは感染性のあるクローンを生じさせ、同様にコンセンサス配列は、かかるライブラリーからの全ての若しくは重要な数のクローンの配列を比較することにより決定すべきである。十分なクローンが評価されるべきであり、その結果多様な位置にある大部分の塩基は同じである。よって、コンセンサスは少なくとも三つのクローン、好ましくは約五つのクローン、より好ましくは六つ又はそれ以上のクローンの配列を解析することにより決定される。当然、誤りがちなクローニング方法を用いれば用いるほど、配列されるべきクローンの数は増え、本物の H C V コンセンサス配列を与える。

20

それからコンセンサス配列は感染性 H C V D N A クローンを調製するのに利用される。結果生じたクローンの忠実度は、配列決定により確立されることが好ましい。しかしながら、例えばクローンが感染性 H C V R N A をコード化するならば、クローンの性質に基づいて選択が実行される。よって、感染性 H C V D N A クローンをうまく調製することは、クローンでトランスフェクトされた又は感染された動物若しくは細胞培養の物理的、病理的又は免疫学的性質に基づく分析により検出される。例えば、c D N A クローンは、類似又は同一の物理-化学的、電気泳動移動、等電点電気泳動、又は非平衡 p H ゲル電気泳動拳動、タンパク質分解性消化マップ、又は固有 H C V 又は H C V ウイルス粒子タンパク質として知られているような抗原性質を有する H C V ビリオン又はウイルス粒子タンパク質を生産するように選択される。

30

機能性 H C V c D N A クローンの成分。本発明で説明される機能性 H C V c D N A の成分は、以下に説明するように、既知又は新しく確認された H C V 坑ウイルスターゲットのための無細胞、細胞培養、キメラウイルスと動物に基づくスクリーニング検定法を開発するために利用される。既知又は仮のターゲット及び分析法の例は、これに限定されないが以下のものを含む（Houghton, In "Field Virology"（概観のためにB. N. Fields, D. M. KnipeとP. M. Howley, 編集）, Vol. pp. 1035-1058. Raven Press, New York（1996）；Rice（1996上記文献；RiceらによるAntiviral Therapy 1, Suppl. 4, 11-17（1997）；Shimotohno, Hepatology 21; 887-8（1995）を参照）。

40

入ってくる H C V ゲノム R N A の翻訳に必須な要素を含むかなり保存された5' N T R は、一つのターゲットである。この配列又はその補体は、R N A 複製及び/又はパッケージングに重要な R N A 要素を含むであろう。潜在的な治療戦略は、アンチセンスオリゴヌクレオチド（上記）；トランス-作用リボザイム（上記）；R N A デコイ；上記要素の機能

50

を干渉する小さな分子化合物（これらはRNA要素自身と、又は同族体ウイルス若しくは活性に必要とされる細胞要因と結合することにより作用する）を含む。

別のターゲットは、かなり保存され以下の機能と関連するHCV C（キャプシド又はコア）タンパク質である：RNA結合とHCVゲノムRNAの特異包膜；細胞（ReyらによるVirus Res. 37: 209-220（1995））及び他のウイルス（ShihらによるJ. Virol. 69: 1160-1171（1995）；ShihらによるJ. Virol. 67: 5823-5832（1993））遺伝子の転写モジュレーション；細胞形質転換（RayらによるJ. Virol. 70: 4438-4443（1996a））；アポトーシスの防止（RayらによるVirology 226: 176-182（1996b））；TNF受容体スーパーファミリーのメンバーとの結合を介した宿主免疫応答のモジュレーション（MatsumotoらによるJ. Virol. 71: 1301-1309（1997））。

10

ビリオンエンベロープの成分を形成するE1、E2及びE2-p-7糖タンパク質は、潜在的な中和抗体のターゲットである。介入のキー段階は以下を含む。；シグナルペプチダーゼがポリタンパク質からのこれら前駆体の切断を媒介させる（Linらによる（1994a）の上記文献）；E1E2糖タンパク質複合体のER集合と細胞シャペロンと上記タンパク質に関連し、機械装置を折重ねる。（Dubuissonらによる（1994）の上記文献；DubuissonとRice, J. Virol. 70: 778-786（1996））；ウイルス粒子集合はヌクレオキャプシドとビリオンエンベロープとの間の相互作用を含む；ウイルス粒子の運搬と放出；免疫監視の回避若しくは結合及びLDL受容体を発現させる細胞のエントリにおいて役割を担うVLDLのような宿主成分とウイルス粒子の関連（Hijikataらによる（1993）上記文献；Thomssenらによる（1992）上記文献；ThomssenらによるMed. Microbiol. Immunol. 182: 329-334（1993））；ビリオン中の保存され可変な決定因子は、抗体による中和のため、又は抗体と結合し、同族体Fc受容体との相互作用により細胞の免疫増大感染を容易にするターゲットである；受容体結合及びエントリに重要なビリオン中の保存され可変な決定因子；ビリオン決定因子がエントリ、細胞膜との融合に参加し、入ってくるウイルスヌクレオキャプシドを脱外被させる。

20

2/3部位での切断に必要であるNS2-3オートプロテアーゼが、さらなるターゲットである。

複合体を形成し、HCVポリタンパク質（概観のためにRice,（1997）の上記文献参照）での四つの切断を媒介させるNS3セリンプロテアーゼとNS4A補助因子は、更に適当なターゲットである。ターゲットはセリンプロテアーゼ活性自身を含む；セリンプロテアーゼのC末端ドメインでの四面体Zn<sup>2+</sup>配位部位；NS3-NS4A補助因子相互作用；NS4Aの膜関連；NS4AによるNS3の安定性；NS3プロテアーゼ領域の形質変換能力（SakamuroらによるJ. Virol. 69: 3893-6（1995））である。

30

NS3RNA刺激NTase（Suzichらによる（1993）の上記文献）、RNAヘリカーゼ（JinとPeterson, Arch Biochem Biophys 323: 47-53（1995）；KimらによるBiochem. Biophys. Res. Commun. 215: 160-6（1995））、RNA結合（KanaiらによるFEBS Lett 376: 221-4（1995））活性；まだ明確でない機能のRNA複製複合体の成分としてのNS4Aタンパク質；他の推定される複製成分のNS5Aタンパク質は、セリン残基上で主にホスホリレート（phosphorylate）され（TanjiらによるJ. Virol. 69: 3980-3986（1995））、薬開発の全てのターゲットである。治療用のターゲットであるNS5Aの可能な特性は、NS5Aホスホリレーション（phosphorylation）の責任を担うキナーゼと、NS5Aとの相互作用を含む；NS5Aとの相互作用とHCV複製複合体の他の成分である。

40

HCVポジティブ及びネガティブ鎖RNAの実際の合成の責任を担う酵素であるNS5BRDRPは別のターゲットである。その活性の特異な面はポリメラーゼ活性自身を含む（BehrensらによるEMBO J. 15: 12-22（1996））；HCV RNAを含み、他のレプリカーゼ成分とNS5Bの相互作用；ネガティブ及びポジティブ鎖RNA合成の開始に関連する段階；NS5Bのホスホリレーション（HwangらによるVirology 227: 438（1997））である。

他のターゲットは宿主細胞機能のHCV RNA複製及び/又はモジュレーションに重要

50

な構造又は非構造タンパク質機能を含む。ウイルスが入ること、出ること若しくは宿主細胞遺伝子発現のモジュレーションによって重要なチャンネルを形成することが可能な可能な疎水性タンパク質成分はターゲットである。

3' NTR、特別にかなり保存された要素(ポリ(U/UC)トラクト; 98の塩基末端配列)はターゲットである。治療的手法は、ゲノムの上記部分がマイナス鎖合成の開始に重要な役割をするであろう以外は、5' NTRに説明されたものに平行である。翻訳、RNA安定性、若しくはパッケージングを含むHCV RNA複製の他の面にも関連する。機能性HCV cDNAクローンは、全てのウイルスタンパク質及びRNAパッケージングに必要とされるRNA要素をコード化する。上記要素は坑ウイルス化合物の開発にとってターゲットである。電気泳動移動シフト、UV交差結合、フィルター結合、及び三つのハイブリッド(SenGuptaらによるProc. Natl. Acad. Sci. USA 93: 8496-8501 (1996)) 検定法は、HCV RNAパッケージングに重要なタンパク質とRNA要素を定義し、上記過程の阻害剤のスクリーニングのための分析を確立させるために利用される。かかる阻害剤は生体外で生産された小さな分子又はRNAデコイを含む(Goldらによる(1995)の上記文献)。

複雑なHCVライブラリーはPCRシャッフリング(sheffling)を利用して、若しくは“ペプチドディスプレイ”ライブラリーに発生したようなランダム(randomized)配列を取込むことにより調製可能である。“ファージ法”を利用して(ScottとSmith, 1990, Science 249: 386-390 (1990); CwirlaらによるProc. Natl. Acad. Sci., 87: 6378-6382 (1990); DevlinらによるScience, 249: 404-406 (1990))、かなり大規模なライブラリーは構築される( $10^6$ - $10^8$ の化学物質)。上述し以下に例証するように、かかるライブラリーからのクローンはコンセンサスゲノム配列を発生させるために利用される。ヌクレオチド暗号づけ配列の変質により、HCVポリタンパク質暗号づけ領域として実質的に同じアミノ酸配列をコード化する他のDNA配列は、本発明で実際に利用される。このことは、以下のことに限らないが、他の種からの相同性遺伝子、配列内の同じアミノ酸残基をコード化するさまざまなコドンの置換により変質されたHCVポリタンパク質遺伝子の全て又は一部から成るヌクレオチド配列を含み、サイレント変化が生じる。かかるサイレント変化は、多重感染動物モデルでの特定の感染分離体を確認するために利用されるゲノムマーカーの生成を可能とする。同様に、本発明のHCVゲノム誘導体は、これらに限らないが、一次アミノ酸配列として保存アミノ酸置換を招く配列内の残基に置換される機能的に等価なアミノ酸残基での変質配列を含むHCVポリタンパク質のアミノ酸配列の全て又は一部を含む誘導体を包含する。例えば、配列内の一つ又はそれ以上のアミノ酸残基は類似の極性のアミノ酸により置換され、機能的に等価として振る舞い、サイレント変化をもたらす。配列内でのアミノ酸の置換は、アミノ酸の属する分類の他のメンバーから選択される。例えば、非極性(疎水性)アミノ酸には、アラニン、ロイシン、イソロイシン、バリン、プロリン、フェニルアラニン、トリプトファンとメチオニンがある。芳香族環構造を有するアミノ酸には、フェニルアラニン、トリプトファンとチロシンがある。極性のある中性アミノ酸には、グリシン、セリン、トレオニン、システイン、チロシン、アスパラギンとグルタミンがある。正に帯電した(塩基性)アミノ酸には、アルギニン、リジンとヒスチジンがある。負に帯電した(酸性)アミノ酸には、アスパラギン酸とグルタミン酸がある。

特に好ましい置換は

正の電荷が維持されるように、Lysに対してArg、又はその逆と;

負の電荷が維持されるように、Gluに対してAsp、又はその逆と;

フリーな水酸基が維持されるようにSerに対してThrと;

フリーなアミノ基が維持されうようにGlnに対してAsnとの置換が挙げられる。

別の実施例において、本物のHCVクローンが変種され、タンパク質機能が減少する若しくは排除されるアミノ酸置換が導入される。さらに本物のHCVクローンが変種され、ウイルス向性を変化させるアミノ酸置換が導入される。

さらに、HCVは校正活性がないので、ウイルス自身は容易に突然変異し、本発明内でさ

10

20

30

40

50

らに考えられたH C Vの突然変異体“準種”が形成される。かかる突然変異体は、本願で説明したように被験者からの分離体の配列決定により容易に確認される。

H C V誘導体をコード化するクローンと本発明の類似体は、本技術分野で周知のさまざまな方法により生産される。そのような生産を招く操作は、遺伝子若しくはタンパク質レベルで起きる。例えば、クローン化されたH C Vゲノム配列は、本技術分野では周知の数多くの戦略により変形される(Sambrookらによる1998の上記文献)。ゲノム配列は制限エンドヌクレアーゼで適当な部位で切断され、必要ならば続いて酵素変種され、生体外で分離され、ライゲート(ligate)される。あるいは、ゲノムフラグメントは、例えばPCRで接合され、H C Vゲノムが生じる。ゲノム核酸誘導体若しくはH C Vの類似体の生産において、変種ゲノムが固有H C Vゲノムとして同じ翻訳リーディングフレーム内に残り、所望の活性がコード化される領域において翻訳停止シグナルにより中断されないことを保証するように注意しなければならない。

H C Vポリタンパク質コード化核酸配列は、生体外若しくは生体内で突然変異し、翻訳、開始、及び/又は終結配列を生じさせ及び/又は破壊し、若しくは暗号づけ領域における変種体を生じさせる及び/又は新しい制限エンドヌクレアーゼ部位を形成する、若しくは予め存在したエンドヌクレアーゼ部位を破壊し、生体外での変種を容易にする。かかる突然変異は、H C Vの機能活性の変質をもたらし、例えばウイルス活性の弱毒化、又は以下に示すような欠損ウイルスを発生させる。利用される本技術分野で周知の突然変異形成の技法は、これらに限らないが、生体外での部位特異的変異の導入(Hutchinsonやによる1978, J. Biol. Chem. 253: 6551; ZollerとSmith, 1984, DNA 3: 479-488; Oliphantらによる1986, Gene 44: 177; Hutchinsonらによる1986, Proc. Natl. Acad. Sci. USA 83: 710)、T A Bリンカー(Pharmacia)の使用がある。PCR技術は部位特異的変異の導入には好ましい(Higuchi, 1989, “Using PCR to Engineer DNA” in PCR Technology: Principles and Applications for DNA Amplification, H. Erlich編集., Stockton Press, Chapter 6, pp.61-70)。

細胞培養若しくは代替りの宿主におけるより効率的な複製のためのH C Vの適合。以前に述べたように、細胞培養中でのH C V複製は非効率である。さらにH C V複製機械装置の制御下での優性選択可能なマーカーは、H C V複製機械装置における適応突然変異を選択するために利用される。かかる適応突然変異が現れるが、これらに限定されない、：(i) H C V RNA複製の向性を変質させる；(i i) 宿主細胞への有毒効果の原因であるウイルス生成物を変質させる；(i i i) H C V RNA複製効率を増加させる又は減少させる；(i v) H C V RNAパッケージング効率及び/又はH C V粒子の構築及び放出を増加させる又は減少させる；(v) 受容体結合とエントリのレベルでの細胞向性を変質させる。H C Vオリジナルc D N Aクローンの配列は、特定細胞タイプ中で複製を確立させることと適合性がたとえなくても、突然変異は生体外転写、H C V媒介RNA合成の初期段階中に起こる、又は上記化学的若しくは生物学的な多様な方法により、テンプレートD N Aに取込まれ、特定細胞環境又は動物宿主において複製を可能にする。発現が生産的H C V RNA複製に依存する特別設計された優性選択可能なマーカーは、H C V複製機械装置若しくはトランスフェクトされた宿主細胞のいずれか、又はその双方において適応突然変異を選択するために利用される。

キメラH C Vクローン。さらに上記機能性クローンの成分は、H C V遺伝子機能とその阻害剤の分析用のキメラウイルスを作り上げるために利用される(FilocamoらによるJ. Vir ol. 71: 1417-1427 (1997)；HahnらによるVirology 226:318-326 (1996)；LuとWimmer, Proc. Natl. Acad. Sci. USA 93: 1412-7 (1996))。本発明の一つのそのような拡張において、5' I R E S、プロテアーゼ、RNAヘリカーゼ、ポリメラーゼ又は3' N T Rのような機能性H C V要素は、複製が一つ又はそれ以上の上記H C V要素に依存するB V D Vのキメラ誘導体を生じさせるために利用される。かかるB V D V/H C Vキメラは上記機能性成分に対する坑ウイルス戦略用のスクリーニング及び評価するために利用される。

加えて、優性選択可能なマーカーは、RNA複製若しくは粒子生成の高いレベルを可能に

10

20

30

40

50

するHCV複製機械装置において突然変異を選択させるために利用される。一つの例において、DHFRの突然変異形を発現する特別設計されたHCV誘導体は、メトトレキセート(MTX)に対する抵抗性を比較するために利用される。優性選択可能なマーカーとして、殆ど化学量論的量がMTX抵抗性には必要であるので、突然変異体DHFRは非効率である。培地中にてMTXの濃度をうまく増加させることにより、DHFRの増加した量が、複製するHCV RNAが宿る細胞の継続生存に必要である。この選択スキーム、又はこの概念に基づく同様なスキームは、HCV RNA複製とRNA蓄積の高いレベルを可能にするHCV RNA複製機械装置において、突然変異の選択を招く。同様な選択はHCV粒子の高収量の生産を可能にする突然変異に、若しくは変質された細胞向性での突然変異HCV粒子に応用可能である。かかる選択スキームは、培養上澄からのHCV粒子を宿し、又は細胞破壊(cell disruption)後MTX抵抗性のために選択し、固有細胞の再感染による粒子をトランスデュースさせることを含む。

10

確認され分離されたゲノムRNAはcDNAへ逆転写される。さらにcDNAは、プロモーター及びランオフ部位を含むように“長い”PCRにより作られる、又は適当なレシピエントベクターにおける挿入用の3'末端コンセンサス配列特異プライマーを用いて作られる。上記cDNAは適当なクローニングベクター、例えば適当なプロモータと3'ランオフ配列に沿ってコンセンサス5'及び3'-NTRから成るベクターに挿入される。プライマーとランオフ配列を含むクローンは、機能性HCV RNAの生産に直接利用される。本技術分野において周知の数多くのベクター-宿主システムが利用される。ベクターの例としては、以下のものに限られないが、E. coli、ラムダ誘導体のようなバクテリオファージ、pBR322誘導体又はpUCプラスミド誘導体のようなプラスミド、例えばpGEXベクター、pmal-c、pFLAG、pTETなどがある。例えばクローニングベクターへの挿入は、DNAフラグメントを相補的粘着末端を有するクローニングベクターヘライゲートさせることにより達成される。しかしながら、DNAをフラグメント化させるのに利用された相補的制限部位がクローニングベクターに存在しないのなら、DNA分子の端は酵素的に修飾される。あるいは、望ましい部位は、DNA末端上のヌクレオチド配列(リンカー)をライゲートすることにより作られる。上記ライゲートリンカーは、制限エンドヌクレアーゼ認識配列をコード化する特別に化学的に合成されたオリゴヌクレオチドから成る。組換え分子は、遺伝子配列の多くのコピーが発生するように、トランスフォーメーション、トランスフェクション、感染、エレクトロポレーションなどにより宿主へ導入される。

20

30

#### HCV RNA及びポリタンパク質の発現

HCV RNAとHCVタンパク質、特にHCV RNAレプリカーゼ又はビリオンタンパク質の遺伝暗号を指定するHCV DNAは、適当な発現ベクター、つまり挿入されたタンパク質暗号づけ配列の転写及び翻訳用の必要な要素を含むベクターへ挿入される。かかる要素は本願では“プロモーター”という。よって、本発明のHCV DNAは本発明の発現ベクター内のプロモーターと操作上(又はoperably)関連する。さらに発現ベクターは複製起点を含むことが好ましい。必要な転写及び翻訳シグナルは、組換え発現ベクター上に与えられる。機能性RNAの生体外合成の好ましい実施例において、T7、T3若しくはSP6プロモーターが利用される。

40

潜在的宿主-ベクターシステムは、これらに限らないが、ウイルス組換え体(例えばワクチンウイルス、アデノウイルス、シンドビスウイルス、セムリキ森林熱ウイルスなど)で感染された哺乳類細胞システム;組換えウイルス(例えばバキュロウイルス)で感染された昆虫;酵母ベクターを含む酵母のような微生物;植物細胞;又はバクテリオファージでトランスフォームされたバクテリア、DNA、プラスミドDNA、若しくはコスミッドDNAを含む。ベクターの発現要素は抵抗性と特異性が変化する。利用した宿主-ベクターシステムに依存して、多くの適当な転写及び翻訳要素のいずれか一つが利用される。

HCV DNAクローンが導入された細胞から成る組換えベクターは、HCV RNAの発現又は細胞によるかかるHCVタンパク質を与える条件下で、適当な細胞培養地に培養される。DNAフラグメントをクローニングベクターへ挿入させる前述したいずれの方

50

法も、適当な転写/翻訳制御シグナルとタンパク質暗号づけ配列から成る遺伝子を含む発現ベクターを作り上げるために利用される。上記方法は、生体外組換えDNAと合成技術と生体内組換え（遺伝子的組換え）を含む。

HCV RNAの発現又はタンパク質は本技術分野で知られているプロモーター/エンハンサーにより制御されるが、上記調節要素は発現のために選択された宿主で機能的でなければならない。発現を制御するために利用されるプロモーターは、これらに限定されないが、SV40早期プロモーター領域（BenoistとChambon, 1981, Nature 290: 304-310）、ラウス肉腫ウイルスの3'の長い末端リピートに含まれるプロモーター（Yamamotoらによる1980, Cell 22: 787-797）、ヘルペスチミジンキナーゼプロモーター（Wangerらによる1981, Proc. Natl. Acad. Sci. USA. 78: 1441-1445）、メタロチオネイン遺伝子の調整配列（Brinsterらによる1982, Nature 296:39-42）、 $\gamma$ -ラクタマーゼプロモーターのような原核発現ベクター（Villa-Komaroffらによる1978, Proc. Natl. Acad. Sci. USA, 75: 3727-3731）、又はtacプロモーター（DeBoerらによる1983, Proc. Natl. Acad. Sci. USA, 80: 21-25; "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242: 74-94も参照）、Gal 4プロモーターのような酵母又はカビからのプロモーター要素、ADC（アルコールデヒドロゲナーゼ）プロモーター、PGK（ホスホグリセロールキナーゼ）プロモーター、アルカリ性ホスファターゼプロモーター、組織特異性を示し遺伝子導入動物において利用される動物転写制御領域、膵臓腺房細胞において活性であるエラスターゼI遺伝子制御領域（Swiftらによる1984, Cell 38: 639-646; OrnitzらによるCold Spring Harbor Symp. Quant. Biol. 50: 399-409; MacDonald, 1987, Hepatology 7: 425-515）、膵臓ベータ細胞において活性であるインシュリン遺伝子制御領域（Hananan, 1985, Nature 351: 115-122）、リンパ系細胞において活性である免疫グロブリン遺伝子制御領域（Grosschedlらによる1984, Cell 38: 647-658; Adamesらによる1985, Nature 318: 533-538; Alexanderらによる1987, Mol. Cell. Biol. 7: 1436-1444）、精巣、胸部、リンパ及びマスト細胞において活性であるマウス乳腺癌ウイルス制御領域（Ledererらによる1986, Cell 45: 485-495）、肝臓において活性であるアルブミン遺伝子制御領域（Pinkertらによる1987, Genes and Devel. 1: 268-276）、肝臓において活性であるアルファ胎児性タンパク質遺伝子制御領域（Krumlaufらによる1985, Mol. Cell. Biol. 5: 1639-1648; Hammerらによる1987, Science 235: 53-58）、肝臓において活性であるアルファ1-トリプシン遺伝子制御領域（Kelleyらによる1987, Genes and Devel. 1: 161-171）、骨髄細胞において活性であるベータグロブリン遺伝子制御領域（Mogamらによる1985, Nature 315: 338-340; Kolliasらによる1986, Cell 46: 89-94）、脳の乏突起細胞において活性であるミエリン塩基性タンパク質遺伝子制御領域（Readheadらによる1987, Cell 48: 703-712）、骨格筋において活性であるミオシンL鎖遺伝子制御領域（Sani, 1985, Nature 314: 283-286）と視床下部において活性であるゴナドトロピン放出ホルモン遺伝子制御領域（Mansonらによる1986, Science 234: 1372-1378）を含む。

宿主/発現ベクターの幅広い組合せは、本発明のDNA配列を発現する時に利用される。例えば有用な発現ベクターは、染色体セグメント、非染色体と合成DNA配列を含む。適当なベクターはSV40n誘導体と周知のバクテリアプラスミド、例えばE. coli プラスミド、colE1、pCR1、pBR322、pMal-C2、pET、pGEX（Smithらによる1988, Gene 67: 31-40）、pMB9とその誘導体、RP4のようなプラスミド、例えばNM989のファージの数多くの誘導体のようなファージDNAと他のファージDNA、例えばM13、糸状一本鎖ファージDNA、2 $\mu$ プラスミド又はその誘導体のようなイーストプラスミド、真核細胞において有用なベクター、昆虫又は哺乳類細胞において有用なベクター、プラスミドとファージDNAの組合せから誘導されるベクター、ファージDNA又は他の発現制御配列を利用して修飾されたプラスミド、及び本技術分野で周知なものから成る。

好ましい配列分析に加えて、本発明のHCV DNAクローンを含む発現ベクターは、以下の五つの手法により確認される：（a）所望のプラスミドDNA若しくは特異mRNAのPCR増幅、（b）核酸ハイブリダイゼーション、（c）選択マーカー遺伝子機能の存

10

20

30

40

50

否、(d) 適当な制限エンドヌクレアーゼによる分析、(e) 挿入配列の発現である。第一の手法において、核酸はPCRにより増幅されて増幅生成物の検出を可能にする。第二の手法において、発現ベクターに挿入された外来遺伝子の存在を、HCV DNAと相同である配列から成るプローブを利用して、核酸ハイブリダイゼーションにより検出する。第三の手法において、組換えベクター/宿主システムはベクターにある外来遺伝子の挿入により引き起こされるある“選択マーカー”遺伝子機能( -ガラクトシダーゼ、チミジンキナーゼ活性、抗生物質に対する抵抗、形質変換表現型、バキュロウイルスにおいてオクルージョンボディ(occlusion body)形成)の存否に基づいて確認され選択される。第四の手法において、組換え発現ベクターは適当な制限酵素で消化されることにより同じである。第五の手法において、組換え発現ベクターは、組換えにより発現された遺伝子生成物、例えばHCV RNA HCVビリオン若しくはHCVウイルスタンパク質の活性、生化学的、若しくは免疫特性を分析することにより確認される。

例えば、バキュロウイルス発現システムにおいて、非融合トランスファーベクターは、これらには限定されないが、pVL941(BamHIクロニング部位; Summers)、pVL1393(BamHI、SmaI、XbaI、EcoRI、NotI、XmaIII、BglIIとPstIクロニング部位; Invitrogen)、pVL1392(BglII、PstI、NotI、XmaIII、EcoRI、XbaI、SmaI、とBamHIクロニング部位; SummersとInvitrogen)とpBlueBacIII(BamHI、BglII、PstI、NcoI、とHindIIIクロニング部位、青(blue)/白(white)組換えスクリーニングが可能; Invitrogen)が利用され、融合トランスファーベクターは、これらには限定されないがpAc700(BamHIとKpnIクロニング部位、BamHI認識部位は開始コドンで始まる; Summers)、pAc701とpAc702(pAc700と同じで異なるリーディングフレームを有する)、pAc360(ポリヘドリン開始コドンのBamHIクロニング部位36ベースペア下流; Invitrogen(195))、pBlueBacHisA、B、C(BamHI、BglII、PstI、NcoIとHindIIIクロニング部位で、三つの異なるリーディングフレーム、ProBond精製のN末端ペプチド、とブランクの青/白組換えスクリーニング; Invitrogen)が利用される。

本発明において利用されることが考えられた哺乳類発現ベクターの例は、ジヒドロ葉酸レダクターゼ(DHFR)プロモーターのような誘発性プロモーター、例えばpED(PstI、SalI、SbaI、SmaIとEcoRIクロニング部位、クローン化された遺伝子とDHFRの双方を発現させるベクター)(Kaufman, Current Protocols in Molecular Biology, 16.12(1991))のようなDHFR発現ベクター又はDHFR/メトトレキセート共増幅ベクターのある発現ベクターを含む。あるいは、pEE14(HindIII、XbaI、SmaI、SbaI、EcoRIとBclIクロニング部位、ベクターはグルタミンシンセターゼとクローン化された遺伝子を発現する; Celltech)のようなグルタミンシンセターゼ/メチオニンスルフォキシイミン共増幅ベクターも含まれる。別の実施例において、pREP4(BamHI、SfiI、XhoI、NotI、NheI、HindIII、NheI、PvuIIとKpnIクロニング部位、PSV-LTRプロモーター、ハイグロマイシン選択可能マーカーを構成する; Invitrogen)、pCEP4(BamHI、SfiI、XhoI、NotI、NheI、HindIII、NheI、PvuIIとKpnIクロニング部位、hCMV直接早期遺伝子、ハイグロマイシン選択可能マーカーを構成する; Invitrogen)、pMEP4(KpnI、PvuI、NheI、HindIII、NotI、XhoI、SfiI、BamHIクロニング部位、誘発性メタロチオネインIIa遺伝子プロモーター、ハイグロマイシン選択可能マーカーを構成する; Invitrogen)、pREP8(BamHI、XhoI、NotI、HindIII、NheIとKpnIクロニング部位、PSV-LTRプロモーター、ヒスチジノール選択可能マーカーを構成する; Invitrogen)、pREP9(KpnI、NheI、HindIII、NotI、XhoI、SfiIとBamHIクロニング部位、PSV-LTRプロモーター、G418選択可能マーカーを構成する; Invitrogen)とpBVHis

10

20

30

40

50

( P S V - L T R プロモーター、ハイグロマイシン選択可能マーカー、ProBond樹脂により精製可能で、エンテロキナーゼにより切断されたN末端ペプチド)のような、エプスタインバ-ウイルス(EBV)の制御下でエピソーム発現を導くベクターが利用される。TettとTett(GossenとBujard, Proc. Natl. Acad. Sci. USA89: 5547-51 (1992); GossenらによるScience 268: 1766-1769 (1995))のような調節可能な哺乳類発現ベクターが利用される。本発明に利用される選択可能な哺乳類発現ベクターは、pRC/CMV(HindIII, BstXI, NotI, SbaIとApaIクロニング部位、G418選択; Invitrogen)、pRC/RSV(HindIII、SeI, BstXI、NotI, XbaIクロニング部位; G418選択; Invitrogen)と他を含む。本発明による利用されるワクチンウイルス哺乳類発現ベクター(Kaufman(1991)上記文献参照)は、これらに限らないが、pSC11(SmaIクロニング部位、TK-と-gal選択)、pMJ601(SalI, SmaI, AflI, NarI, BspMII, BamHI, ApaI, NheI, SacI, KpnI, とHindIIIクロニング部位; TK-と-gal選択)とpTKgptF1S(EoRI, PstI, SalI, AccI, HindII, SbaI, BamHIとHpaIクロニング部位、TK又はXPRT選択)を含む。

本発明によるイースト発現システムの例は、非融合pYES2ベクター(XbaI, SphI, ShoI, NotI, GstXI, EcoRI, BstXI, BamHI, SacI, KpnI, とHindIIIクロニング部位; Invitrogen)又は融合pYESHisA, B, C(XbaI, SphI, ShoI, NoyI, BstXI, EcoRI, BamHI, SacI, KpnI, とHindIIIクロニング部位、ProBond樹脂で精製されエンテロキナーゼで切断されるN末端ペプチド)を含み、利用される。加えて、宿主細胞菌株は挿入された配列の発現を調節する、又は望ましい特定の方法で遺伝子生成物を変種され処理するように選択される。さまざまな宿主細胞は翻訳と翻訳後のタンパク質のプロセッシング及び変種(例えば、糖鎖形成、切断(例えば配列のシグナル))の特徴的な特異機能を有する。イーストにおける発現は糖結合生成物が生成する。真核細胞における発現は“固有”糖鎖形成及びHCVタンパク質の折畳み(folding)の可能性を増大させる。さらに、哺乳類細胞における発現は、ウイルス粒子タンパク質の固有HCVビリオンを再組立若しくは組立てる手段を提供する。

さらに、さまざまなベクター/宿主発現システムは、タンパク質分解切断のようなプロセッシング反応に、異なる程度で、影響を及ぼす。

他のRNAウイルス研究に有用である多くのトランスフェクション法は、本願では利用できる。例としては、微小注射、細胞融合、lipofectinのようなカルシウム-ホスフェイトカチオンリポソーム(RiceらによるNew Biol. 1: 285-296 (1989); “HCV-based Gene Expression Vector” 下記文献を参照)、DE-デキストラン(RiceらによるJ. Virol. 61: 3809-3819 (1987))とエレクトロポレーション(BredenbeekらによるJ. Virol. 67: 6439-6446 (1993); LiljestronらによるJ. Virol. 65: 4107-4113 (1991))がある。上記他の方法によるトランスフェクションに対する細胞タイプ治療抵抗性のためにスクレイプローディング(scrape loading)(KumarらによるBiochem. Mo. Biol. Int. 32: 1059-1066 (1994))及びバリスティック法(Ballistic method)(BurkholderらによるJ. Immunol. Meth. 165: 149-156 (1993))もさらに検討される。DNAベクタートランスポーターが検討される(Wuらによる1992, J. Biol. Chem. 267: 963-967; WuとWu, 1988, J. Biol. Chem. 263: 14621-14624; Hartmutらによる1990年3月15日に出願されたカナダ特許出願第2,012,311号)。

#### HCVによる生体外感染

HCV複製を支持する細胞系統の確認。本発明の重要な態様は、本物の標準化された生体外HCV複製システムを利用して、さまざまな治療的戦略の効力を評価する能力を比較することによる、新しくより効率的なアンチHCV治療を開発する方法を提供することである。かかる検定法はチンパンジーやHCV感染ヒト患者のような稀で価値のある実験動物を利用して試す前には、非常に貴重である。発明の背景で説明したように、せいぜい僅か



な量のHCV複製が細胞培養中で観察されており、報告された多くのシステムはドラッグスクリーニング又は評価には適応できない。報告された中で最も見込みのあるシステムは、HTLV-1感染MT-2CTリンパ球サブライン(subline)であり、約1000:1のSN比を有するHCV複製を支持することが分かった(MizutaniらによるJ. Virol., 70: 7219-23 (1996))。しかしながら、上記システムにおける複製は患者接種での感染により開始されていることを指摘しておく。かかるシステムは有用ではあるが、細胞向性及び複製の検出に影響を与える接種源間の差により限定されてしまう。

HCV感染性クローン技術は、HCV複製とパッケージングの分析用の生体外及び生体内システムを確立するために利用される。上記技術は、これらに制限されるわけではないが、(i)許容細胞タイプの確認と選択(RNA複製、ビリオン構築及び放出のため); (ii)細胞培養パラメータ(例えば培養条件、細胞活性化を変化させる)若しくは細胞培養中でのHCV複製の効率を増加させる適合突然変異の選択; (iii)感染性HCV粒子の効率生産用の条件の明確(細胞上澄へ放出する若しくは細胞破壊(cell disruption)後に得るのいずれか)を含む。上記及び本発明の他に容易に拡張したものは、HCV治療、ワクチン、及び診断開発に幅広い有用性を有する。

許容細胞タイプを確認するための一般的手法を以下に概説する。RNAトランスフェクション(上記を参照)の最適方法は、細胞タイプにより変化し、RNAレポーター構成物を利用して決定される。上記方法は、例えばトランスフェクション条件を最適化し(CAT; クロラムフェニコールアセチルトランスフェラーゼ活性)、細胞タイプがHCV IRES媒介翻訳に許容かどうかを決定するのに(LUC; ルシフェラーゼ活性)、双方で利用された5'-CAT-HCV IRES-LUC-3'構造のあるビストロニック(bicistronic)RNA(WangらによるJ. Virol. 67: 3338-44 (1993))を含む。実際のHCV

RNAトランスフェクション実験では、5'キャップルシフェラーゼレポーターRNA(Wangらによる(1993)の上記文献)のあるコトランスフェクションは、生産的トランスフェクションと翻訳の内部標準を提供する。HCV複製の潜在的に許容な細胞タイプの例には、これらに制限されないが、一次ヒト細胞(例えば肝細胞、T細胞、B細胞、包皮繊維芽細胞)だけでなく連続ヒト細胞系統(例えばHe p G 2, Hu h 7, HUT 7 8, HPB-Ma, MT-2, MT-2Cと他のHTLV-1とHTLV-II感染T細胞系統、Namalawa, Daudi, EBV-トランスフォームされたLCL)がある。加えて、他の種、特にRNAで容易にトランスフェクトされ、フラビウイルス若しくはペスチウイルス(例えばSW-13、ベロ、BHK-21、COS, PK-15, MBCKなど)に許容である細胞系統もテストされた。細胞は上記に説明した方法を用いてトランスフェクトされる。

複製検定では、RNA転写物は機能性クローンと対応する非機能性である、例えばGDD(例参照)誘導体を利用して合成され、HCV RNAの存続のネガティブコントロールと生産的複製がない場合の抗原として利用される。テンプレートDNA(コンプリケイト(complicates)は後で分析される)はDNase I処理の繰返しサイクルと酸フェノール抽出により取除かれ、その後続くゲル電気泳動若しくはゲル濾過のいずれかにより精製された(転写RNAの $10^9$ 分子につき増幅可能なDNAの1分子以下に)。DNAフリーなRNA転写物はLUCレポーターRNAと混合され、上記で決定された最適条件を利用して細胞培養をトランスフェクトさせるのに利用した。細胞回収後に、RNAase Aが培地に添加され過剰のインプットRNAが消化され、さまざまな時間インキュベートされた。早い段階で(トランスフェクト後約1日)収集され、LUC活性と(生産的トランスフェクションを確認するために)細胞と上澄液(ベースラインとして)中でのポジティブ鎖RNAレベルを分析した。サンプルは2-3週間ごとに定期的に集められ、QC-RT/PCR(Kolykhalovらによる(1996)の上記文献参照)によりプラス鎖RNAレベルを検定した。無傷の感染転写物と非機能性誘導体、例えばGDDの間の明らかで再現可能な差を示す細胞タイプ、欠失、制御は、本当の複製を確認するためにもっと徹底的な分析が行われた。かかる検定法は、QC-RT/PCR(Gunjiらによる(1994)の上記文献; LanfordらによるVirology 202: 606-14 (1994))、ノーザンブロットハイブリダイゼーション、若しくは代謝ラベリング(Yooらによる(1995)の上記文献)及びin situハイ

10

20

30

40

50

ブリダイゼーション (ISH; Gowansらによる In “Nucleic Acid Probes” (R. H. Symons, 編集). Vol. pp.139-158. CRC Press, Boca Raton. (1989))、in situ PCR (ISHに続いて HCV 特異増幅生成物のみを検出する; Haaseらによる Proc. Natl. Acad. Sci. USA 87:4941-4975 (1990))、と免疫組織化学によるネガティブ鎖 HCV RNA の蓄積の測定を含む。

ウイルス-受容体相互作用の研究のための HCV 粒子。HCV 感染と複製に許容である細胞系統の確認と組合せて、感染性クローン技術を用いて生産された明確な HCV ストックは、細胞受容体と HCV の相互作用を評価するために利用された。検定は感染しやすい細胞へのウイルスの結合若しくは生産的感染を測定するように設けられ、上記過程の阻害剤のスクリーニングに利用される。

HCV 受容体の特徴付けのための細胞系統の確認。RNA トランスフェクションにより検定された HCV RNA 複製に許容な細胞系統は、ウイルスにより感染される能力に対してスクリーニングされた。RNA 複製に許容であるが相同なウイルスにより感染しない細胞系統は、HCV 結合及びエントリに必要な一つ又はそれ以上の宿主受容体が欠けている。かかる細胞は、(i) HCV 受容体と共受容体の機能性確認と分子クローニング; (ii) ウイルス-受容体相互作用の特徴付け; (iii) 上記相互作用を阻害する化合物若しくは生物製剤 (例えば抗体、SELEX RNAs (Bartel と Szostak, In “RNA-protein interaction” (K. Nagai と I. W. Mattaj による編集), Vol. pp. 82-102. IRL Press, Oxford (1995); Goldらによる Annu. Rev. Biochem. 64: 763-797 (1995), など) のスクリーニングするための検定法の開発のための価値ある道具を提供する。

このような方法で一旦明確になると、上記 HCV 受容体は治療のターゲットとしてだけでなく、HCV に感染しやすくされた遺伝子導入された動物において発現させる働きがある (Koikeらによる Dev. Biol. Stand. 78: 101-7 (1993); Ren と Racaniello, J. Virol. 66: 296-304 (1992))。HCV 複製とスプレッドを支持するかかる遺伝子導入動物モデルは、アンチ HCV 薬を評価するのに重要な応用を有する。

許容細胞系統を確認するための代替りの手法。機能性クローンから誘導された未変種 HCV RNA 転写物を利用することに加えて、上記機能性 HCV クローンは HCV 複製用の選択可能なマーカーを与えるように特別に設計される。例えば、優性選択可能なマーカーをコード化する遺伝子は、HCV ポリタンパク質の一部として、若しくは HCV RNA ゲノムの許容領域に位置する分離シストロンとして発現される。かかる特別設計された誘導体 (概観のために、Bredenbeek と Rice, Semin. Virol. 3: 297-310 (1992) を参照) は、シンドビスウイルス (Frolovらによる, Proc. Natl. Acad. Sci. USA, 93: 11371-11377 (1996)) 又はフラビウイルス (Khromykh と Westaway, J. Virol. 71: 1497-1505 (1997)) のような他の RNA ウイルス用にうまく構築された。哺乳類細胞用の選択可能なマーカーの例には、これらに限定されないが、ジヒドロ葉酸レダクターゼ (DHFR; メトトレキセート耐性)、チミジンキナーゼ (tk: メトトレキセート耐性)、ピューロマイシンアセチルトランスフェラーゼ (pac: プューロマイシン耐性)、ネオマイシン耐性 (neo: ネオマイシン若しくは G418 に対する耐性)、ミコフェノール酸耐性 (gpt)、ハイグロマイシン耐性、及びゼオシン (zeocin) に対する耐性を含む。他の選択可能なマーカーは、イースト (ura3, his3, trp1) のようなさまざまな宿主に利用される。異種遺伝子の機能性発現の戦略は記述されている (概観のために Bredenbeek と Rice による (1992) の上記文献参照)。例 (図 2) は、(i) 細胞若しくはウイルスプロテアーゼにより媒介される選択可能なタンパク質を生じるように、切断によるウイルスポリタンパク質へのインフレーム挿入; (ii) 特別設計された転写開始と停止シグナルを利用した分離シストロンの生産を含む。例は、これらに限定されないが、細胞若しくはウイルス mRNA から誘導された内部リボソームエントリ部位 (IRES) RNA 要素の使用 (Jangらによる Enzyme 44: 292-309 (1991); Macejak と Sarnow, Nature 353: 90-94 (1991); Mollaらによる Nature 356: 255-257 (1992)) を含む。特別な明示において、EMCV IRES 要素とネオマイシン耐性遺伝子を含むカセットを HCV H773' NTR 超可変領域に挿入される。転写された RNA は、ヒト肝細胞若しくは他の細胞系統及び耐性細胞集団

10

20

30

40

50

を選択するために利用された抗生物質 G 4 1 8 をトランスフェクトさせるために利用される。この手法の一つの明示において、p H C V F L / 3 ' E M C V I R E S n e o ( 以下に示す ) からの転写物は、多種多様な細胞系統をトランスフェクトさせるのに利用される。

H C V c D N A の変種は、H C V 複製の間接的な指標として便利な検定可能なマーカーを発現する系統を発生させるように作られる。かかる自己複製 R N A は、全体の H C V ゲノム R N A 若しくは R N A レプリコンを含み、R N A 複製に必須でない領域は削除される。検定可能な遺伝子は第二の選択可能なマーカー、又は便利な分析によりタンパク質をコード化する遺伝子を含む。例は、これらに限定されないが、 $\beta$ -ガラクトシダーゼ、 $\beta$ -グルクロニダーゼ、ホタル若しくはバクテリアルシフェラーゼ、グリーン蛍光タンパク質 ( G F ) とそれらのヒトでの誘導体、細胞表面マーカーと分泌 ( secreted ) マーカーを含む。かかる生成物は直接検定される、又は追加のレポーターの発現又は活性を活性化させる。

10

#### H C V 感染及び複製用の動物モデル

チンパンジーに加えて、本発明は H C V 複製を研究し、新規な治療を評価するための代わりの動物モデルの開発を可能にする。出発物質として本発明に記載された本物の H C V c D N A クローンを利用して、H C V 複製の代わりの動物モデルを確立するために、多くの手法が検討された。一つの明示において、チンパンジーのトランスフェクションにより、又は細胞培養中での複製により生産されたよく定義された H C V ストックは、H C V 複製を支えることが可能なヒト組織に宿る免疫不全マウスを接種させるために利用される。本技術の例は、S C I D : H u マウスであり、厳しく組合された免疫不全のあるマウスは、これらに限定されないが、胎児肝臓、大人の肝臓、脾臓、又は周辺の血液単核細胞を含むさまざまなヒト ( 又はチンパンジー ) 組織と合体させる。S C I D マウスに加えて、正常な放射マウスはヒト又はチンパンジー組織の合体のレシピエントとして働く。上記キメラ動物は接種源を含む明確なウイルスで半ビボ ( ex vivo ) 又はインビボ ( in vivo ) で感染された後に、H C V 複製と置換されるであろう。

20

別の明示において、代わりの種にて H C V 複製を可能にする適応突然変異は、上記動物における複製に対して許容である変種体を生じる。例えば、連続げっ歯類細胞系統又は一次組織 ( 肝細胞のような ) のいずれかにおける複製及びスプレッドに対する適応 H C V は、小さなげっ歯類モデルにおけるウイルスの複製を可能にする、あるいは、化学若しくは生物学的 ( Stemmer, Proc. Natl. Acad. Sci. USA 91: 10747 ( 1994 ) ) 方法により発生した H C V 変種体の複雑なライブラリーは、潜在的に感染しやすい動物の接種用に作られ、利用される。かかる動物は、前述したように免疫適格若しくは免疫不全のいずれかである。複製可能な変種体は分離され、分子的にクローン化され、選択された非ヒト種の複製に対して機能的である完全長クローンへ適応突然変異が取込まれる。

30

H C V の機能活性は遺伝子導入的に ( transgenically ) 評価される。この点に関して、遺伝子導入されたマウスモデルが利用される ( Wilmut らによる Experientia 47: 905 ( 1991 ) を参照 ) 。 H C V R N A 若しくは D N A は、ウイルスベクター又はコスミッドクローン ( 又はファージクローン ) を含む遺伝子導入ベクターを調製するために利用される。コスミッドは公開された手順 ( Jaenisch, Science 240: 1468-1474 ( 1988 ) ) を利用して、遺伝子導入マウスへ導入される。遺伝子導入マウスの調製において、胚幹細胞は胚盤胞胚 ( blastocyst embryos ) ( Joyner, In Gene Targeting: A Practical Approach. The Practical Approach Series, Rickwood, D. と Hames, B. D. 編集 IRL Press: Oxford ( 1993 ) から得られ、H C V D N A 若しくは R N A によりトランスフェクトされる。トランスフェクトされた細胞は、早期胚、例えば文献 ( Hammer らによる Nature 315: 680 ( 1985 ) ) に説明されたようなマウス胚へ注射される。遺伝子導入された動物の調製用のさまざまな技術が説明されている ( 1996 年 6 月 25 日に発行された米国特許第 5, 530, 177 号 ; 1996 年 12 月 31 日に発行された米国特許第 5, 898, 604 号 ) 。特に、関心は遺伝子導入動物モデルにあり、遺伝子導入の表現及び病原効果が研究される。例えば、ラットホスホエノールビルビン酸カルボキシキナーゼ-ウシ成長ホルモン融合遺伝

40

50

子の効果はブタにおいて研究された (WieghartらによるJ. Reprod. Fert., Suppl. 41: 89-96 (1996))。アルツハイマー病と関連するヒトアミロイド前駆体タンパク質をコード化する遺伝子を発現する遺伝子導入マウスは、アルツハイマー病と他の疾患を研究するために利用される (1996年3月7日に発行された国際特許出願第WO 96/06927号; QuonらによるNature 352: 239 (1991))。さらに遺伝子導入されたマウスは肝炎デルタ剤 (PoloらによるJ. Virol. 69: 5203 (1995)) とB型肝炎ウイルス (Chisar, Curr. Top. Microbiol. Immunol. 206: 149 (1996)) を発生させ、上記特別設計された動物内で複製は起こる。

よって、本願で説明される機能性cDNAまたはその部分は、HCV複製及び病原に関連する遺伝子導入モデルを作るために利用される。一例にて、全体のHCVゲノムに宿る遺伝子導入動物が作られる。本発明の遺伝子導入マウスにおける全体のHCVゲノムの遺伝子導入発現の適当な構成物は、適当な5'末端、完全長HCV cDNA配列、シス-切断デルタリボザイム (Ball, J. Virol. 66: 2335-2345 (1992); PattnaikらによるCell 69: 1011-1020 (1992)) のある転写物を生じるように特別設計された核プロモータを含み、本物の3'末端が、続いて正しい核プロセッシングを促進し細胞質 (HCV RNA複製が起こる場所) へ運搬させるシグナルが生じる。全体のHCVゲノムに加えて、動物が特別に設計されて、HCVタンパク質とRNA要素の個々の若しくはさまざまな組合せを発現する。例えば、HCV IRESの制御下で、HCV遺伝子生成物若しくはレポーター遺伝子を発現するように特別設計された動物は、特異RNAターゲットに対して向けられた治療を評価するのに利用される。同様の動物モデルにより、多くの周知HCVターゲットを検討される。

かかる代替りの動物モデルは、(i) 全動物システムにおけるHCV複製へのさまざまな抗ウイルス剤の効果を研究するのに、(ii) 基礎をなす関連する反応機構を明示し、治療に関与するための戦略を確認しテストして、肝細胞及び他の細胞タイプへのHCV遺伝子生成物の潜在的な直接細胞障害作用を調べ、(iii) HCV病原に関連する細胞及び組織損傷の免疫媒介機構を研究し、上記過程を干渉する戦略を確認しテストすることによって有用である。

#### 薬剤耐性変種体の選択と分析

HCV複製を支える細胞系統及び動物モデルは、従来と新規な治療に抵抗性のあるHCV変種体の発生を調べるために利用される。全てのRNAウイルスと同様に、HCVレピリカーゼは校正活性が欠けていると推定され、したがってRNA複製は誤りがちであり、高いレベルの変種が生じる (Bukhらによる (1995) 上記文献)。変異性は、時間の経過にわたって感染患者において及びさまざまな分離体の間で観測されたかなりの多様性において、それ自身が明らかになる。薬剤耐性変種体の発生は、HCVモノ及び組合せ治療の設計及び評価において、重要な考察であろう。本発明のHCV複製システムは、さまざまな治療配合物下において変種体の発生を研究するために利用される。このことは、単一の治療 (monotherapy) 若しくはさまざまな治療の組合せ (例えばIFN- $\alpha$ 、リバビリン、及び新しい抗ウイルス化合物) をもたらす。耐性突然変異は耐性の分子的及び構造的基礎を明確にし、新しい治療配合物、若しくは効果的なアンチHCV薬剤 (以下に) のスクリーニング検定法における配合物を評価するのに利用される。

#### アンチHCV剤のスクリーニング

HCV許容細胞系統又は動物モデル (好ましくは、げっ歯類モデル) は、新規な阻害剤のスクリーニングに、若しくはアンチHCV治療の候補を評価するのに利用される。かかる治療は、これらに限定されないが、(i) 保存HCV DNAターゲットを目標とするアンチセンスオリゴヌクレオチド; (ii) HCV複製を阻害できる注射可能な化合物; (iii) HCV複製を阻害できる生物学的に利用可能な経口化合物を含む。かかる配合物のターゲットは、これらに制限されないが、(i) RNA複製及びRNAパッケージングによって重要な保存HCV RNA要素; (ii) HCV-コード化酵素; (iii) HCV RNA複製、ウイルス集合、ウイルス放出、ウイルス受容体結合、ウイルスエントリ、及びウイルスRNA複製の開始によって重要なタンパク質-タンパク質及びタンパク質-RNA相

相互作用；(iv)慢性感染を確立させるようにHCV能力を調節するウイルス-宿主相互作用；(v)肝臓発病度を調節し、アポトーシス及び肝毒性に影響を与える要因を含むウイルス-宿主相互作用；(vi)肝硬変及び肝細胞癌を含むより重症な臨床的結果の展開を導くウイルス-宿主相互作用と；(vii)他のあまり頻繁でなく、HCVと関連するヒト疾病を招くウイルス-宿主相互作用を含む。

アンチセンス及びリボザイム治療の評価。本発明はHCV複製を干渉する能力をテストされたアンチセンスヌクレオチド及びリボザイムの調製に拡張される。この手法は特異mRNAの翻訳を、アンチセンス核酸によるマスク若しくはリボゾームによる切断のいずれかによりブロックするためのアンチセンス核酸及びリボザイムを利用する。

アンチセンス核酸は、特異mRNAの少なくともある部分と相補的であるDNA若しくはRNAである(Marcus-Sekura, Anal. Biochem. 172: 298 (1988))。細胞内において、アンチセンス核酸はmRNAとハイブリダイゼーションして二本鎖DNA:RNA若しくはRNA:DNA分子を形成する。この二本鎖形のmRNAを翻訳しない。したがって、アンチセンス核酸はタンパク質へのmRNAの発現を干渉する。約15のヌクレオチドのオリゴマーとAUG開始コドンとハイブリダイゼーションする分子は特に効率的である。なぜなら上記ヌクレオチドと分子は合成が容易であり器官細胞に導入されたときには、大きな分子に問題を引き起こさないからである。アンチセンス法は、生体外の多くの遺伝子の発現を阻害するのに利用される(Marcus-Sekura, 1988, 上記文献; HamborらによるJ. Exp. Med. 168: 1237 (1988))。合成アンチセンスヌクレオチドは、天然ホスホエステル結合よりもむしろホスホロチオレート、又はチオエステルのようなホスホエステル類似体を含むことが好ましい。かかるホスホエステル結合類似体は分解に対して抵抗性があり、安定性が増加し、したがってアンチセンス核酸の効力が増す。

遺伝アンチセンス手法において、野生型対立遺伝子の発現は、アンチセンスRNAの発現のために抑制される。この技術は組織培養におけるTK合成を阻害し、ショウジョウバエにおけるクラッペル(Kruppel)変異と、マウスにおけるシバラー変異の表現型を発生させるように利用される(IzantらによるCell, 36: 1007-1015 (1984); GreenらによるAnn u. Rev. Biochem., 55: 569-597 (1986); KatsukiらによるScience, 241: 593-595 (1988))。この手法の重要な利点は、極僅かな遺伝子の部分が、全mRNAの同族体の発現の効率的な抑制のために発現される必要がある。アンチセンス遺伝子導入は、正しい細胞タイプにおいて発現された自身のプロモータ若しくは他のプロモータの制御下に置かれ、SV40ポリA部位の上流に配置される。

リボザイムは、DNA制限エンドヌクレアーゼとある程度類似する方法で、他の一本鎖RNA分子を特別に切断させる能力を有するRNAである。リボザイムは、あるmRNAが自身のイントロンを切除する(excise)能力を有するという観察から発見された。上記RNAsのヌクレオチド配列を変質させることにより、研究者らはRNA分子内の認識特異ヌクレオチドを配列させ、切断する分子を設計できるようになる(Cech, J. Am. Med. Assoc. 260: 3030 (1988))。配列特異であるので、特定配列のあるmRNAのみが不活性化される。

研究者らは二つのタイプのリボザイム、テトラヒメラタイプと“ハンマーヘッド”タイプを確認した。テトラヒメラタイプリボザイムは、四つの塩基配列を認識するが、“ハンマーヘッド”タイプは11から18の塩基配列を認識する。認識配列が長くなればなるほど、ターゲットmRNA種において専ら起こるのであろう。したがって、“ハンマーヘッド”タイプリボザイムは、特異mRNA種を不活性化させるにはテトラヒメラタイプよりは好ましく、18の塩基認識配列は認識配列を短くするのに好ましい。

化合物ライブラリーのアンチHCV活性に対するスクリーニング。さまざまな天然生成物若しくは合成ライブラリーは、本発明により与えられる生体外若しくは生体内モデルにおいて、アンチHCV活性に対してスクリーニングされる。コンビナトリアルライブラリーを調製する一つの手法は、主にGeysen法(GeysenらによるMolecular Immunology 23: 709-715 (1986); GeysenらによるJ. Immunologic. Method 102: 259-274 (1987))の化学方法とFodorらの方法(FodorらによるScience 251: 767-773 (1991))を利用する。例と

してはFurkaらによるもの(14th International Congress of Biocchemistry, Volume 5 Abstract FR:013 (1988); Int. J. Peptide Protein Res. 37: 487-493 (1991))、Houghtonによるもの(1986年12月に発行された米国特許第4,631,211号)、Riceらによるもの(1991年4月23日に発行された米国特許第5,010,175号)があり、アンチHCV活性がテストされるペプチド混合物を生じさせる方法が記載されている。

別の態様には、合成ライブラリー(NeedelsらによるProc. Natl. Acad. Sci. USA 90: 10700-4 (1993); OhlmeverらによるProc. Natl. Acad. Sci. USA 90: 10922-10926 (1993); Lamらによる国際特許出願第WO 92/00252号; Kocisらによる国際特許出願第WO 9428028号)などが、本発明によるアンチHCV化合物のスクリーニング用に利用された。上記文献には生物学的検定法におけるライブラリースクリーニングへの適応が説明されている。

中和検定法用の明確な/特別設計されたHCVウイルス粒子。本願で説明した機能性クローンは、感染力と中和検定法用のHCV-H粒子の明確なストックを生産するように利用される。均質なストックは、チンパンジーモデル、細胞培養システムにおいて、さまざまな非相同発現システムを利用して生産される(例えばバキュロウイルス、イースト、哺乳類細胞、上記参照)。前述したように、HCV-Hの均質ウイルス調製に加えて、他の遺伝子型又は分離体のストックが生産される。上記のストックは細胞培養若しくは生体内分析に利用され、HCV粒子生産又は感染性を中和することができる分子若しくは遺伝子治療の手法を明確にする。かかる分子の例には、これらに制限されないが、ポリクローナル抗体、モノクローナル抗体、特別設計された/最適化された特異性のある人工抗体、単鎖抗体(以下の抗体のセクション参照)、核酸若しくは特異結合及び中和用に選択された核酸誘導体、少量の生物学的に利用可能な経口化合物などがある。保存ウイルス若しくは細胞ターゲットに向けられるかかる中和剤は、遺伝子型である、若しくは分離体特異である、又は広く交差反応性があるのいずれかである。中和剤は予防的に利用される、又はウイルス負荷を減少させ、他の抗ウイルス剤(IFN- $\alpha$ 、リバビリン)と組合せてより効果的な治療の可能性を増加させる受身免疫治療が利用される。さらにHCV感染クローンの方向性のある(directed)操作は、糖タンパク質超可変領域における、又は抗体中和、CTL認識、免疫エスケープ及び免疫エンハンスメントの機構を研究するためのエピトープにおける明確な変化のあるHCVストックを生産するために利用される。上記研究は抗ウイルス治療の他のウイルス特異機能の確認をもたらす。

#### HCV複製の解剖

他のHCV複製検定法。本発明はHCV複製の方向性のある分子遺伝子解剖を始めて可能とする。かかる分析により、(i)現在追跡されている抗ウイルスターゲットを確認すること;(ii)治療介入の影響を受けやすいHCV複製の予期せぬ新しい面を発見することが期待される。突然変異誘発研究によるターゲットの直接確認には、以下の5'NTR、HCVポリタンパク質と切断生成物、及び3'NTRが含まれる。前述したように、感染性クローン技術と許容細胞培養を利用する分析は、感染性RNAで細胞培養をトランスフェクトさせた後の親と突然変異複製表現型を比較することに利用される。たとえPCRがウイルスRNA蓄積の感度のよい検出を可能にしても、条件突然変異が回復しなければ、RNA複製の効率を減少させる突然変異を解析するのが難しい。第一サイクルへの補体として、トランス-相補性検定法(trans-complementation assays)は、HCV突然変異表現型の分析と阻害剤のスクリーニングを容易にするために利用される。非相同性システム(ワクシニア、シンドビス、又は非ウイルス)は、HCV RNAレプリカゼ及び/又はパッケージング機械装置の発現を駆動させるために利用される(LemmとRice, J. Virol. 67: 1905-1915 (1993a); LemmとRice, J. Virol. 67: 1916-1926 (1993b); LemmらによるEmbo J. 13: 2925-2934 (1994); LiらによるJ. Virol. 65: 6714-6723 (1991))。仮にこれらの要素がトランスに機能するならば、適当なシス要素のあるRNAの共発現はRNA複製/パッケージングをもたらす。したがって、かかるシステムは本当のRNA複製とビリオン構築の段階を模倣するが、HCV複製からのウイルス成分のアン

カップル (uncouple) 生産は真似ない。H C V 複製がなんとか自己制御的であるならば、非相同システムはかなり高いレベルの R N A 複製若しくは粒子生産を引き起こし、突然変異表現型の解析と坑ウイルススクリーニングを容易にする。第三の手法は、H C V テンプレート依存 R N A 複製の無細胞システムを考案することである。結合 (coupled) 翻訳/複製と構築システムは、H e L a 細胞中のポリオウイルスのために説明され (Barton と Flanagan, J. Virol. 67:822-831 (1993); Molla らによる Science 254:1647-1651 (1991))、ネガティブ鎖合成開始のテンプレート依存生体外検定法はシンドビスウイルスに対して確立される。H C V 用の同様な生体外システムは、阻害剤スクリーニングと評価だけでなく、H C V 複製の多くの面を研究することにとっても価値がない。それぞれの戦略例は以下のようなものである。

10

ウイルス又は非ウイルス発現システムを利用した H C V R N A 複製及び/又はパッケージングのトランス相補性。非相同性システムは H C V 複製を駆動させるために利用される。例えば、ワクシニア/T 7 細胞質発現システムは、R N A ウイルスレプリカーゼとパッケージング機能のトランス相補性には非常に有用である (Ball, (1992) の上記文献; Lemm と Rice, (1993a) の上記文献; Lemm と Rice, (1993b) の上記文献; Lemm らによる (1994) の上記文献; Pattnaik らによる (1992) の上記文献; Pattnaik らによる Virology 206:760-4 (1995); Porter らによる J. Virol. 69:1548-1555 (1995))。簡単には、ワクシニア組換え体 (v T F 7-3) は、関心のある細胞タイプ中で T 7 R N A ポリメラーゼを発現させるために利用される。T 7 プロモーターから下流に位置するターゲット c D N A は、ワクシニア組換え体として、又はプラスミドトランスフェクションによるのいずれかで運搬される。このシステムは高いレベルの R N A とタンパク質発現を導く。ワクシニアの必要性がない (H C V R N A レプリカーゼ又はピリオン形成を干渉する) 上記手法の変形は p T 7 T 7 システムであり、T 7 プロモーターは T 7 R N A p o l の発現を引き起こさせる (Chen らによる Nucleic Acids Res. 22: 2114-2120 (1994))。p T 7 T 7 は T 7 R N A p o l (タンパク質) と混合し、関心のある T 7 駆動ターゲットプラスミドとコトランスフェクトされた。添加された T 7 R N A p o l は転写を開始し、自身の生産と高いレベルのターゲット遺伝子の発現を導く。いずれかの手法を用いて、正確な 5' 及び 3' 末端を有する R N A 転写物は、T 7 転写開始部位 (5') とシス切断 H C V リボザイム (R z) (3') を利用して生産される (Ball, (1992) の上記文献; Pattnaik らによる (1992) の上記文献)。

20

30

上記又は同様な発現システムは、H C V R N A 複製と粒子形成と上記過程を阻害する化合物の評価の検定法を確立するために利用される。H C V 機能性クローン技術の別の拡張において、T 7 駆動タンパク質発現構成物と 3' N T R に続く H C V リボザイムを取込んだ完全長 H C V クローンが利用される。検定を確認する典型的な実験プランは、p T 7 T 7 で説明されるが、本質的に同様な検定は v T F 7-3 又は T 7 R N A ポリメラーゼを発現する細胞システムを用いて検討される。H C V 許容細胞は p T 7 T 7 + T 7 R N A p o l + p 90/H C V F L 長い p U R z (又は G D D のようなネガティブコントロール) でコトランスフェクトされる。トランスフェクション後のさまざまな時間で、p T 7 T 7 システムにより駆動された H C V タンパク質と R N A の蓄積は、それぞれウエスタン及びノーザンブロッティングにより追跡される。H C V 特異レプリカーゼ機能を検定するためには、A c t . D が D N A 依存 T 7 転写をブロックするため (Lemm と Rice による (1993a) の上記文献) に添加され、A c t . D 耐性 R N A 合成は代謝ラベリングにより監視される。放射能が p 90/H C V F L 長い p U / R z の完全長 H C V R N A に取込まれるが、p 90/H C V F L G D D / R z には取込まれない。この検定システム又は精巧な誘導体は、阻害剤をスクリーニングさせるために、H C V R N A 複製への効果を研究するために利用される。

40

H C V 複製とその阻害剤を検定するための無細胞システム。H C V R N A 複製と阻害剤のスクリーニングを研究するための無細胞検定は、本発明に説明された機能性 c D N A クローンを利用してさらに確立される。ピリオン若しくは転写 R N A のいずれかが、基質 R N A として利用される。H C V によっては、生体外でトランスフェクトされた完全長 H C

50

V RNAは、ポリオウイルスで説明された生体外システム及び検定された複製を計画するように利用される(Bartonらによる(1995)の上記文献参照)。HCV RNA複製には肝細胞特異性又は他の要因が必要であり、システムは肝細胞又は他の細胞抽出物が補充され、あるいは比較可能なシステムは、HCV複製許容であると分かった細胞システムを利用して確立される。

上記手法の一つの心配事は、正しい無細胞合成とHCVポリタンパク質のプロセッシングが発生しなければならないことである。正しくプロセッシングされたレプリカーゼ成分の十分な量が生じることは難しい。この問題を克服するために、T7発現システムが適当な細胞中で高いレベルのHCVレプリカーゼ成分を発現させるために利用される(Lemmらによる(1997)の上記文献参照)。(緩衝液、 $Mg^{2+}$ 、ATP再生システムとNTRを添加した)上記細胞からのP15膜フラクションは、HCV特異テンプレートRNAを添加したときに、完全長ネガティブ鎖RNAを開始し合成されるべきである。

上記検定法のいずれか若しくは双方とも確立することにより、HCV RNA複製におけるさまざまな段階での複製と阻害剤が関係するHCV突然変異の効果、宿主要因の速く正確な分析を可能にする。さらに上記システムは、複製欠損HCVベクターを調製するヘルパーシステムの条件を確立する。

#### 予防接種と防御免疫

効果的なHCVワクチンの開発に悪影響を与える多くの未知のパラメータがいまだに存在する。幾つかの個体は感染を克服することが、ヒト及びチンパンジーの双方で明らかである。さらに10-20%のIFNで処理した個体はHCV RNAの循環が行われないことの証明として、持続的応答を示すらしい。他の研究から、相同なタイプで再感染されるだけでなく、関連するHCVタイプよりもかなり離れているタイプでも感染することで証明されるように、防御免疫の欠如を示す(Farciらによる(1992)の上記文献; Princeらによる(1992)の上記文献)。にもかかわらず、E1E2オリゴマーと上記タンパク質を発現するワクシニア組換え体から成るサブユニットワクチンで免疫されたチンパンジーは、低投与対抗量(challenge)に対して部分的に保護される(ChooらによるProc. Natl. Acad. Sci. USA 91: 1294 (1994))。本発明で説明される感染性クローン技術は、HVCに対する防御免疫応答の性質の理解を目的とする基礎研究としてだけでなく、新規なワクチン生産方法にとっても有用である。

HCVに対する活性免疫は、弱毒化された又は不活性化されたHCVビリオン、又はHCVウイルス粒子タンパク質の免疫原性量で、好ましくは免疫的に効果的なアジュバントで免疫化(予防接種)されることにより誘発される。“免疫的に効果的なアジュバント”は免疫応答を促進させる物質である。

アジュバントの選択はワクチン予防接種される被験者に依存する。好ましくは、薬学的に許容なアジュバントが利用される。例えば、ヒト用のワクチンは完全及び不完全なフロイントアジュバントを含むオイル若しくは炭化水素エマルジョンアジュバントは避けるべきである。ヒトへの使用に適するアジュバントの一例には、みょうばん(alum、アルミナゲル)がある。しかしながら、動物へのワクチンには、ヒトの使用に適さないアジュバントを含む。

抗原とアジュバントから成る伝統的ワクチンの代わりのものには、被験者組織の細胞による抗原の発現のために、抗原をコード化するDNA若しくはRNAの被験者への生体内での直接導入を含む。かかるワクチンは、本願では“DNAワクチン”、“遺伝子ワクチン”若しくは“核酸に基づくワクチン”と呼ぶ。DNAベクター若しくはベクタートランスポーターのような前述したトランスフェクションの方法は、DNAワクチンに利用される。

DNAワクチンは国際特許出願第WO 95/20660号と、国際特許出願第WO 93/19183号に説明されており、上記文献の開示内容は全体の引用文献に含まれる。防御免疫応答を顕在化させるようにウイルスタンパク質又はゲノムをコード化する、直接注射されたDNAの能力は、数多くの実験システムにおいて説明されてきた(ConryらによるCancer Res., 54: 1164-1168 (1994); CoxらによるVirol, 67: 5664-5667 (1993); Davis

10

20

30

40

50



らによるHum. Mole. Genet., 2: 1847-1851 (1993); SedegahらによるProc. Natl. Acad. Sci., 91: 9866-9870 (1994); MontgomeryらによるDNA Cell Biol., 12: 777-783 (1993); UlmerらによるScience, 259: 1745-1749 (1993); WangらによるProc. Natl. Acad. Sci., 90:4156-4160 (1993); XiangらによるVirology, 199: 132-140 (1994))。インフルエンザウイルスの中和におけるこの戦略へアクセスするための研究は、抗体の生産を誘発するようにエンベロープ及び内部ウイルスタンパク質の双方を利用するが、特にウイルス赤血球凝集素タンパク質(HA)に集中していた(FynanらによるDNA Cell Biol., 12: 785-789 (1993A); FynanらによるProc. Natl. Acad. Sci., 90: 11478-11482 (1993B); RobinsonらによるVaccine, 11: 957 (1993); WebsterらによるVaccine, 12: 1495-1498 (1994))。

防御免疫応答を顕在化させるようにタンパク質をコード化するDNA若しくはRNAを直接注射する予防接種は、細胞媒介及び体液性応答の双方を生じさせる。このことは肝臓ウイルスで得られた結果と類似する(RazらによるProc. Natl. Acad. Sci., 91: 9519-9523 (1994); Ulmer, 1993の上記文献; Wang, 1993の上記文献; Xiang, 1994の上記文献)。ケナガイタチでの研究により、インフルエンザの保存内部ウイルスタンパク質に対するDNAワクチンは、表面糖タンパク質と協力して、不活性化された又はサブピリオンワクチンよりインフルエンザの変種体に対してより効果的である(DonnellyらによるNat. Medicine, 6: 583-587 (1995))。実際には、ヌクレオタンパク質をコード化するDNAへの再現可能な免疫応答は、動物の寿命を本質的に持続させるには必要であることが、マウスにおいて報告されている(YankauchasらによるDNA Cell Biol., 12: 771-776 (1993))

。本発明のワクチンは、これらに限定されないが、筋肉内、腹腔内、静脈内、動脈内(例えば肝動脈)などを含む非経口ルートを経由して投与される。好ましくは予防接種の所望結果は、直接投与若しくはウイルスベクターのターゲティング又は選択、間接的にリンパ系細胞へ、例えばリンパ節や脾臓でのHCVの免疫応答を説明する。免疫細胞は連続して複製されるので、免疫細胞はレトロウイルスに基づく核酸ワクチンの理想的目標である。なぜならレトロウイルスは複製細胞を必要とするからである。

受動免疫は、抗血清を投与する、ポリクローナル抗体を中和する、若しくは患者へのHCVに対するモノクローナル抗体を中和することにより、HCVによる感染を患っていると疑われる動物検体へ与えられる。受動免疫は長期間の保護を与えないが、予防接種されていない被験者の急性感染の治療には価値ある手段である。受動免疫治療に投与された抗体は、自己由来の抗体であることが好ましい。例えば、被験者がヒトならば、抗体はヒト器官である、若しくは抗体に対する免疫応答の可能性を最小にするように“ヒト化”されていることが好ましい。加えて、抗体の中和をコード化する遺伝子は生体内、例えば肝細胞で発現のベクターに導入される。

受動免疫治療用の抗体。前述のように調製されたHCVピリオン若しくはピリオン粒子は、HCVを認識する抗体を発生させる免疫原として利用される。かかる抗体は、これらに限定されないが、ポリクローナル、モノクローナル、キメラ、単鎖(single chain)、Fabフラグメント、及びFab発現ライブラリーを含む。本技術分野で知られているさまざまな手順は、HCVへのポリクローナル抗体の生産に利用される。抗体に生産にとっては、さまざまな宿主動物は、例えば以下に説明されるように、これらに限定されないが、ウサギ、マウス、ラット、ヒツジ、ヤギなどを含むHCVピリオン又はポリタンパク質で注射されて免疫される。さまざまなアジュバントは、これらに限定されないが、フロイントアジュバント(完全及び不完全)、水酸化アルミニウムのようなミネラルゲル、リゾシチレンのような表面活性物質、ブルロニックポリオール、ポリアニオン、ペプチド、オイルエマルジョン、キーホールリンペット、ヘモシアニン、ジニトロフェノール、及びBCG(bacille Calmette-Guerin)やコリネバクテリウムパルブムのような潜在的にヒトに有用なアジュバントを含む宿主種に依存して、免疫応答を促進させる。

前述したようにHCVに向かうモノクローナル抗体の調製には、培地における連続細胞系統による抗体分子の生産をもたらす技術が利用される。上記のことは、これらに限定され

ないが、KohlerとMilsteinにより独自に開発されたハイブリドーマ技術 (Nature 256: 495-497 (1975)) だけでなくトリオマ (trioma) 技術、ヒト B 細胞ハイブリドーマ技術 (Kozborらによる Immunology Today 4: 72 (1983); Coteらによる Proc. Natl. Acad. Sci. USA 80: 2026-2030 (1983)) とヒトモノクローナル抗体を発生させるための E B V-ハイブリドーマ技術 (Coleらによる in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp.77-96 (1985)) を含む。本発明のさらなる実施例において、モノクローナル抗体は無菌動物においても生産される (1989年12月28に発行された国際特許出願第 WO 89/12690号)。実際に、本発明によれば、適当な生物活性のあるヒト抗体分子からの遺伝子と共に、H C V 特異マウス抗体分子からの遺伝子をスプライシングすることによる “キメラ抗体” (Morrisonらによる J. Bacteriol. 159: 870 (1984); Neubergerらによる Nature 312: 604-608 (1994); Takedaらによる Nature 314: 452-454 (1985)) の生産用に開発された技術が利用できる。かかる抗体は本発明の範囲内である。ヒト又はヒト化された抗体は特にアレルギー応答において、外因性抗体よりはあまり免疫応答を誘発しないので、かかるヒト又はヒト化されたキメラ抗体は、ヒト疾病及び疾患の治療 (以下に説明) において好ましく利用される。

本発明によれば、単鎖抗体 (Hustonへの米国特許第 5、476、786号と第 5、132、405号; 米国特許第 4、946、778号) の生産に説明された技術は、H C V 特異単鎖抗体を生産するように適応される。本発明の追加の実施例は、所望の特異性でモノクローナル F a b フラグメントの迅速で容易な確認を可能にする F a b 発現ライブラリーの構築用に説明された技術を利用する (Huseらによる Science 246: 1275-1281 (1989))。抗体分子のイディオタイプを含む抗体フラグメントは、周知の技術で発生される。例えば、かかるフラグメントは、これらに限定されないが、抗体分子のペプシン消化により生じる  $F(a b')_2$  フラグメント;  $F(a b')_2$  フラグメントのジスルフィドブリッジを還元させることにより発生する F a b' フラグメント; 抗体分子をパバインと還元剤で処理することにより発生する F a b フラグメントを含む。

サブユニット予防接種のための H C V 粒子。機能性 H C V-H c D N A クローンと同様に構築され確認された他の遺伝子型のクローンは、予防接種用の H C V 様 (like) 粒子を生産するために利用される。正しい糖鎖形成、折畳み、及び H C V 粒子の構築は、適当な抗原及び防御サブユニットワクチンを発生させるためには重要である。幾つかの方法が粒子生産に利用される。上記方法には、H C V 様粒子 (バクテリア、イースト、哺乳類細胞を用いて) の誘発性又は構成性発現のための安定な細胞システムのエンジニアリング、若しくは組換えバキュロウイルス、ワクシニアウイルス (Moss, Proc. Natl. Acad. Sci. USA. 93: 11341-11348 (1996)) 又はアルファウイルス (Frolovらによる (1996) の上記文献) のような高いレベルの真核異種発現システムの使用を含む。免疫化用の H C V 粒子は局在化に依存して、培地若しくは破壊 (disrupted) 細胞のいずれかから精製される。かかる精製 H C V 粒子又は H C V 遺伝子型のスペクトルを表わす粒子の混合物は、免疫抗原性を高めるためにさまざまなアジュバントと共に若しくはなしで注射される。

感染性非複製 H C V 粒子。別の明示において、受容体結合、エントリとゲノム R N A の翻訳が可能な H C V 粒子が生産される。かかる粒子の異種発現手法は、これらに制限されないが、E c o l i、イースト、又は哺乳動物細胞系統、感染された若しくは組換えバキュロウイルスに宿る適当な宿主細胞、組換えワクシニア、組換えアルファウイルス又は R N A レプリコン、若しくは組換えアデノウイルス、適当な H C V R N A 及びタンパク質を発現するように特別設計されたものを含む。一例において、二つの組換えバキュロウイルスが特別設計される。あるバキュロウイルスは、H C V 粒子の構築に必要とされる H C V 構造タンパク質 (例えば C-E 1-E 2-p 7) を発現する。第二の組換え体は、G D D のような欠失が H C V N S 5 B R D P R を不活性化させるように含まれている以外は、正確な 5' と 3' 末端のある全 H C V ゲノム R N A を発現する。さらに生産的 H C V 複製を破壊する他の突然変異は、代わりに若しくは組合せて利用させる。二つの組換え体との適当な宿主細胞の共感染 (S f 9, S f 21 など) は、高いレベルの H C V 構造タンパク質と H C V 様粒子へのパッケージングのためのゲノム R N A が発生する。かかる粒子

10

20

30

40

50

は高いレベルで発生し、精製され、予防接種に利用される。ワクシニアへ一旦導入されると、かかる粒子は正常な受容体結合とHCV感染しやすい細胞の感染を示す。エントリが起こり、ゲノムRNAは翻訳されて、ゲノムの更なる複製が不活性化された5Bポリメラーゼで完全にブロックされる以外は、全ての正常なHCV抗原が発生する。かかる粒子は、構造及び非構造HCVタンパク質抗原に対する効果的なCTL応答を顕在化させることが期待される。この予防接種戦略のみが、若しくは好ましくは前述したサブユニット戦略と共に、中和抗体とウイルスをクリアにさせるのを助けるCTL応答の双方の高いレベルを顕在化させるのに利用される。多種多様なHCVゲノム配列が、広い交差反応性及び防御免疫応答を確実にするために利用される。加えて、遺伝子工学、又は生体外での誘導化のいずれかによるHCV粒子の変種は、防御及び長く持続する免疫応答を顕在化させるときに、最も効果的に感染を細胞へターゲットさせるために利用される。

10

弱毒化HCV誘導体。HCVゲノムRNA配列を操作し、それによって変質病原性を有する突然変異体が発生する能力は、予防接種に適当な弱毒化されたHCV突然変異体を構築するための手段を提供する。かかるワクチン候補は防御抗原を発現するが、その能力を損なうと、病気を発症させ、慢性感染を確立させ、自己免疫応答を誘発し、細胞をトランスフォームさせる。当然、本発明の感染性HCVウイルスは弱毒化され、不活性化され、化学的又は熱処理により死滅される。

#### HCVに基づく遺伝子発現ベクター

ヒトの慢性肝臓感染を導くHCVの幾つかの同じ性質は、細胞培養システム、遺伝子予防接種、及び遺伝子治療における遺伝子発現ベクターを設計するのにかなり有用である。本願で説明した機能性クローンは、異種遺伝子生成物(RNAとタンパク質)の発現のために設計されたキメラRNAを生産するように特別設計される。戦略は上記及び他の文献(BredenbeekとRice, (1992)の上記文献; Frolovらによる(1996)の上記文献)に説明され、これらに限定されないが、(i) HCVポリタンパク質のある異種暗号づけ配列のインフレーム融合；(ii) HCVゲノムRNAにおける追加のシストロンの生成と；(iii) 一つ又はそれ以上の異種遺伝子(図2)を発現可能なマルチシストロン自己複製HCVベクターRNAを発生させるためのURES要素の封入を含む。かかるベクターに利用される機能性HCV RNAバックボーンは、これらに限定されないが、(i) 複製及びスプレッド可能な弱毒化誘導体；(ii) ウイルススプレッドに必要な一つ又はそれ以上のウイルス成分(例えば構造タンパク質)を欠いているRNA複製能のある“デッドエンド(dead end)”誘導体；(iii) 高い及び低いレベルのHCV特異RNA合成と蓄積が可能な突然変異誘導体；(iv) さまざまなヒト細胞タイプでの複製に適応された突然変異誘導体；(v) ヒト細胞中での長い非細胞変性複製が可能な特別設計された若しくは選択された突然変異誘導体を含む。RNA複製能はあるがパッケージング又はスプレッド能はないベクターは、裸のRNA、DNA若しくはパッケージRNA、DNAとしていずれかがウイルス様粒子へ導入される。かかるウイルス様粒子は前述されたように生産され、未変種若しくは肝細胞若しくは他のヒト細胞タイプのターゲット感染のために設計された変種HCVビリオン成分から成る。あるいは、HCV RNAベクターは異種ウイルスパッケージング機械装置を用いて包膜され運ばれる、又は効率的な遺伝子運搬用に変形されたりリポソームへカルセル化される。上記パッケージング戦略とその変形は、HCVベクターRNAを特異細胞タイプへ効率よく狙わせことに利用される。前述の方法を利用して、他の種において複製及び発現能のある同様なHCV誘導ベクターシステムが導かれる。

20

30

40

例えば上記に示したように、細胞とDNAワクチンのトランスフェクションと関連したさまざまな方法は、本発明のHCVベクターを導入させるために利用される。主に関心があるのは、機能性HCV RNA又はビリオンを、例えば肝臓に直接注入することである。ターゲット遺伝子運搬は、1995年10月に発行された国際特許出願第WO 95/28494号に記述されている。あるいは、ベクターはリポフェクションにより生体内に導入される。過去10年間の間に、生体外で核酸のカルセル化及びトランスフェクション用のリポソームの使用は増加している。トランスフェクションを媒介させるリポソームと遭遇する困難さと危険を制限するために設計させた合成カチオン脂質は、マーカをコード化

50

する遺伝子の生体内トランスフェクション用のリボソームを調製するために利用される (FelgnerらによるProc. Natl. Acad. Sci. USA 84:7413-7417 (1987); MackeyらによるProc. Natl. Acad. Sci. USA 85: 8027-8031 (1988) 参照; UlmerらによるScience 259: 1745-1748 (1993))。カチオン脂質の使用は負の電荷を帯びた核酸のカルセル化を促進し、さらに負に電荷を帯びた細胞膜との融合をも促進する (FelgerとRingold, Science 337: 387-388 (1989))。外因性遺伝子を生体内で特定器官へ導入させるためにリポフェクションを使用することは、実用的には利点がある。特定細胞へリボソームを分子的に狙うことは、その方法の恩恵の一面である。特定細胞タイプへトランスフェクションを向かわせることは、脾臓、肝臓、腎臓及び脳のような細胞不均一性のある組織においては、特に利点がある。脂質はターゲティングのために他の分子と化学的に結合する (Mackeyらによる上記文献参照)。標的ペプチド、例えばホルモン或いは神経伝達物質、及び抗体のようなタンパク質又は非タンパク質は、リボソームと化学的に結合する。さらにレポーター媒介DNA運搬手法も利用される (CurieらによるHum. Gene. Ther. 3: 147-154 (1992); WuとWu, J. Biol. Chem. 262: 4429-4432 (1987))。

遺伝子治療の応用例は、これらに限定されないが、(i) 生来の又は後天性代謝欠陥を正すために酵素又は他の分子の発現; (ii) 創傷治癒を促進させるための分子の発現; (iii) ヒト癌の免疫媒介後退若しくは排除を促進させるための免疫調節分子の発現; (iv) 腫瘍における細胞毒を活性化させることができる毒性分子又は酵素の標的発現; (v) 病原感染細胞における抗ウイルス若しくは抗微生物剤の標的発現を含む。さまざまな治療異種遺伝子は本発明の遺伝子治療に組込まれ、以下に限定されないが、重症複合型免疫不全 (SCID) を治療するためのアデノシンデアミナーゼ (ADA); 腫瘍浸潤 (TIL) T細胞へのマーカー遺伝子又はリンフォカイン遺伝子 (KasisらによるProc. Natl. Acad. Sci. USA 87: 473 (1990); Culverらによる同様雑誌 88:3155 (1991)); 血友病を治療するためのFactor VIIIとFactor IXのような凝固因子用の遺伝子 (DwarkiらによるProc. Natl. Acad. Sci. USA, 92: 1023-1027 (1995); Thompson, Thromb and Haemostatis, 66: 119-122 (1991)) があり、以下に限定されないが、 $\alpha$ -グロビン、ジストロフィン、インシュリン、エリスロポエチン、成長ホルモン、グルコセレブロシド、 $\alpha$ -グルクロニダーゼ、 $\alpha$ -アンチトリプシン、フェニルアラニン、ヒドロキシラーゼ、チロシンヒドロキシラーゼ、オルニチントランスカルバミラーゼ、アポリポタンパク質などのようなさまざまな他の既知治療遺伝子がある。一般にはAndersonらへの米国特許第5、399、346号を参照。

遺伝的予防接種 (HCV以外の病原からの保護) の応用例には、これらには限定されないが、バクテリア (例えばプラスモジウム、リーシュマニア、トキソプラズマ)、カビ (例えばカンジダ、ヒストプラズマ)、とウイルスヒト病原 (例えばHIV, HSV, CMV, インフルエンザ) からの防御抗原の発現を含む。HCV-誘導RNA発現ベクターを利用して発現された防御抗原の免疫原性は、所望のTh1対Th2応答の展開を容易にするサイトカイン (例えばIL-2、GM-CSF) のような免疫調節分子の共発現を含み、アジュバントを利用して増大される。かかるアジュバントは、HCVベクター自身による取込みと共発現されるか、又は他の方法を用いて上記ベクターと組合せて投与される。

#### 感染性HCVの診断方法

診断細胞系統。本願で説明する本発明は、患者サンプル内の感染性HCVの感度良好な診断のための細胞系統を引き出すためにも利用される。概念としては、機能性HCV成分は感染しやすい細胞系統をテストし、発生させるために利用され、容易に検定されたレポーターシステムはHCV感染により選択的に活性化される。これらに制限されないが、例には、(i) 導入遺伝子として取込まれ、複製がアップレギュレートされる又はHCV感染により誘発される、レピリカーゼ成分を欠いた欠損HCV RNAs; (ii) HCV感染により活性化される敏感な異種増幅可能なレポーターシステムがある。第一の明示では、HCV RNA増幅に必要なシスRNAシグナルは、ルシフェラーゼ、グリーン蛍光タンパク質 (GFP)、 $\alpha$ -ガラクトシダーゼ、又は選択可能なマーカー (上記説明) のような便利なレポーター遺伝子に接する。かかるキメラRNAの発現は、適当な核プロモータ

ーと適切な核プロセッシングに必要な要素と細胞質への運搬により駆動される。HCVで特別設計された細胞系統が感染すると、導入遺伝子の細胞質複製と増幅が誘発され、生産的HCV感染の指標として、高いレベルのレポーター発現が引き起こされる。

第二の例において、細胞系統はよりしっかりと制御されるが、HCV感染の際にかなり誘発可能なレポーター遺伝子の増幅及び発現がされるように設計される。この増幅システムは特異成分に関係して説明されるが、他の等価な成分も利用される。一つのかかるシステムにおいて、図3に図解されるように、特別設計されたアルファウイルスレプリコン導入遺伝子は、アルファウイルスnsP4ポリメラーゼがなく生産され、アルファウイルスRNA増幅に絶対必要とされる酵素は、通常非構造タンパク質からの切断により生じる。この欠損アルファウイルスレプリコンの他の特徴は、サブゲノミックRNAプロモーターを含み、ルシフェラーゼ若しくはGFPレポーター遺伝子の発現を駆動させる。このプロモーター要素は、生産的細胞質アルファウイルスレプリコンがない場合に静止状態である。細胞系統は、HCV NS4Aタンパク質とアルファウイルスnsP4 RDRPから成る遺伝子融合の発現用の第二の導入遺伝子を含む。この融合遺伝子は発現され、細胞質膜区画に向かうが、nsP4活性には正確なN末端のある孤立nsP4タンパク質は必要であるので、nsP4の形はアルファウイルス複製複合体の機能成分としては不活性である（LemmらによるEMBO J. 13: 2925 (1994)）。任意の第三の導入遺伝子は、複製用のシグナル、ユビキチン-nsP4融合をコード化するサブゲノミックRNAの転写と、アルファウイルスパッケージングシグナルのある欠損アルファウイルスRNAを発現する。HCVによるかかる細胞系統の感染により、HCV NS3プロテイナーゼが生成し、NS4A-nsP4融合タンパク質のトランス切断を媒介し、nsP4ポリメラーゼを活性化させる。この活性ポリメラーゼはトランスに機能し、少量で効果的であり、欠損アルファウイルスレプリコンだけでなくユビキチン-nsP4をコード化する欠損アルファウイルスRNAの増幅を導く機能性アルファウイルス複製複合体を形成する。サブゲノミックRNAから発現されたユビキチン-nsP4は、細胞ユビキチンカルピキシ末端ヒドロラーゼにより効率よく切断され、追加のnsP4が生じるが、この場合酵素は限られる。一旦活性化されると、このシステムは非常に高レベルのレポータータンパク質を生じる。かかるHCV感染力検定の時間スケールはちょうど時間の単位が必要である（十分なレポーター遺伝子発現のためには）。

抗体診断薬。本願で説明した細胞系統に加えて、細胞系統をトランスフェクトされ若しくは感染され、又感染動物から分離されて発生したHCVウイルス粒子（ビリオン）は、抗原として、患者血液若しくは血液生成物におけるアンチHCV抗体を検出するために利用される。HCVウイルス粒子は本物のHCVゲノムから誘導されるので、粒子は天然HCVウイルスに非常に近い類似性又は同じ構造的特性を有するようである。上記反応剤は、患者が血清変換、つまりHCV特異抗体の集団発生を検出することにより、HCVに感染することを確立するために利用される。

あるいは、本願で説明したように調製された本物のHCV生成物を与える抗体は、被験者からの生物学的サンプルにおけるHCVの存在を検出するために利用される。

発明の典型として提供される、以下の非限定例を参照して、本発明はより理解される。

#### 例

以下の例は実験の背景、最初に感染性HCV RNAをコード化するHCV DNAを調整することにうまくいかなかった試みや、最終的には機能性クローンの発生について報告する。

#### 例1. HCV-Hゲノム構造及び発現の分析

HCV-H菌株、cDNAクローニング、配列分析と殆ど完全長のcDNAクローンの構築の原理。HCV-H菌株はPruceIIとその同僚らによりチンパンジーにおいて広範囲にわたり特徴付けられ（Shimizuらによる（1990）の上記文献）、ごく最近ではShimizuとその共同研究者により生体外（Hijikataらによる（1993）の上記文献；ShimizuらによるJ. Virol. 68: 1494-1500（1994）；ShimizuらによるProc. Natl. Acad. Sci. USA 89: 5477-5481（1992）；ShimizuらによるProc. Natl. Acad. Sci. USA 90, 6037-6041（1993））で

10

20

30

40

50

研究されているので、最初の研究用に選択された。H C V - H は輸血後の N A N B 肝炎のアメリカ人からの遺伝子型 1 a ヒト分離体である (Feinstone らによる J. Infect. Dis. 144: 588-598 (1981))。

H C V - H の最初の c D N A クローニングと配列分析。オリジナルの H C V - H 77 分離体はチンパンジーに 2 回継代接種 (passage) させ、高い血清 A L T レベルと急性肝炎を発生させた。第二のチンパンジー継代接種からの肝臓組織は、c D N A 合成と入れ子 (nested) P C R 増幅に適する粗 R N A の調製に利用した。P C R 増幅 c D N A はプラスミド発現ベクターへクローン化され、幾つかの独立したクローンは分離され、配列分析、発現研究及び再構築させる長い c D N A クローンに利用された。部分配列データと制限酵素マッピングを利用して、p T E T / T 7 H C V F L C M R と呼ばれる殆ど全体の H C V - H c D N A を含むクローンが集められ配列決定された (Daemer らによる未発表結果; Grakoui らによる J. Virol. 67: 1385-1395 (1993c))。このプラスミドに含まれる H C V 配列はその後 H C V - H C M R (S E Q I D N O : 19) と呼ばれる。このクローンの配列は同一直線状であり、Inachauspe らにより (Inachauspe らによる Proc. Natl. Acad. Sci. USA 88: 10292-10296 (1991)) 発表されたチンパンジー継代接種された H C V - H 77 配列に 98.5% 相同であり、Ogata らにより (Ogata らによる (1991) の上記文献) 発表された部分 H C V - H 90 配列にかなりの類似性を示す。

プロトタイプ H C V - H クローンの特徴。H C V - H c D N A クローンと免疫試薬は、無細胞翻訳と細胞培養一過性発現検定に利用され、H C V - H 遺伝子発現のかなり詳細な絵を提供する。一般的には、上記結果は異なる H C V 遺伝子型の他の結果と同様である。この研究は、(i) H C V - H ポリタンパク質切断生成物の確認とマッピング (Grakouli らによる (1993c) の上記文献; Lin らによる (1994a) の上記文献); (ii) タンパク質分解プロセッシング部位の決定 (Grakouli らによる J. Virol. 67: 2832-2843 (1993a); Grakouli らによる Proc. Natl. Acad. Sci. USA 90: 10583-10587 (1993b); Lin らによる (1994a) の上記文献); (iii) N S 2 - 3 オートプロテイナーゼ (Grakouli らによる (1993b) の上記文献; Reed らによる J. Virol. 69: 4127-4136 (1995))、N S 3 - 4 A セリンプロテイナーゼ (Grakouli らによる (1993a) の上記文献; Lin らによる J. Virol. 68: 8147-8157 (1994b); Lin と Rice, Proc. Natl. Acad. Sci. USA 92: 7622-7626 (1995); Lin らによる J. Virol. 69: 4373-4380 (1995)) とそれらの切断条件 (Kolykhalov らによる J. Virol. 68: 7525-7533 (1994); Reed らによる (1995) の上記文献) の決定; (iv) N S 4 A セリンプロテイナーゼ共同因子と N S 3 との関連の研究 (Lin らによる (1994b) の上記文献; Lin と Rice (1995) の上記文献; Lin らによる (1995) の上記文献) と (v) 折畳みとカルネキシン、オリゴマー形成及びサブセルラーローカリゼーション (subcellular localization) との関連を含む H C V 糖タンパク質生合成の調査 (Dubuisson らによる (1994) の上記文献; Dubuisson と Rice, (1996) の上記文献) を含んでいた。他の生物学的に重要な活性の検定法は、部分的に精製された N S 3 (Suzich らによる (1993) の上記文献) と関連する R N A 刺激 N T P a s e と R N A ヘリカーゼ活性と R N A 依存 R N A ポリメラーゼ活性を含むプロトタイプ H C V - H c D N A クローンをを用いて開発された。さらにクローン化された c D N A から発現された抗原は、慢性 H C V 感染の患者からの血清 (Grakouli らによる (1993c) の上記文献参照) と細胞毒性 T リンパ球 (Battegay らによる J. Virol. 69: 2462-2470 (1995); Koziel らによる J. Clin. Invest. 96: 2311-21 (1995)) により認識される。

本発明にとっては、H C V ポリタンパク質プロセッシングの研究により、機能性 I R E S 要素、無傷 O R F、適切な膜トポロジーと全ての 10 のポリタンパク質切断生成物により証明された活性ウイルスプロテイナーゼの予めスクリーニングするための候補の完全長クローンの手段を提供した。

#### 例 2 . c D N A から機能性 H C V の回収の第一の試み

プラスミド構築。高特異感染力のある潜在的に感染性のある H C V R N A 転写物 (Ahluwist らによる Proc. Natl. Acad. Sci. USA 81: 7066-7070 (1984); Rice らによる New Biol. 1: 285-296 (1989); Rice らによる (1987) の上記文献とその引用文献を参照) の生

10

20

30

40

50

産の好ましい戦略は、ランオフRNA転写物(図4)の生産のHCV3'末端に続く制限部位のあるバクテリオファージプロモーター(SP6又はT7)からの候補の完全長HCV cDNA直後の下流のクローニングと関係していた。T7若しくはSP6転写システムは、幾つかの理由のために潜在的に感染性のあるRNAの生産に選択された。まず、T7又はSP6のいずれかが比較的高い特異感染性のキャップのある又はキャップのないRNA転写物の高収量を転写させるのに利用された他のRNAウイルスの数多くの例が存在する(BoyerとHaenni, J. Gen. Virol. 198: 415-426 (1994))。加えて、T7システムはトランスフェクション用の明確なRNAの生体外での合成だけでなく、プラスミドDNAのトランスフェクションを用いた幾つかの生体内手法を可能にするので、T7システムは特に有用である。一例にはワクシニアT7システムがあり、T7 RNAポリメラーゼを発現するワクシニア組換え体はトランスフェクトされたプラスミドテンプレートの細胞質転写を可能にする(FuerstらによるProc. Natl. Acad. Sci. USA 83: 8122-8126 (1986))。ワクシニアウイルスの必要性を回避する第二の生体内での手法には、T7 RNAポリメラーゼを発現するプラスミドのコトランスフェクションである(Chenらによる(1994)の上記文献)。明確な5'と3'末端のある転写物の生産用に設計されたHCVプラスミドDNAでのトランスフェクションは、トランスフェクション手順中に長いRNAの感染性が分解するならば、有利であろう(Ball (1992)の上記文献; Pattnaikらによる(1992)の上記文献)。しかしながら、生体内での上記方法は、転写RNAの構造上での正確な制御とHCV RNA複製が起こると信じられている細胞質への輸出を可能としない。よって、生体外転写法が通常の我々の研究に利用された。

プロセッシング研究の大部分に利用された配列決定されたプロトタイプHCV-H cDNAクローンは、これらの構築用の出発物質であった。上記実験が開始したときにはHCV-HゲノムRNAの末端配列は未知であったので、他の分離体で報告された配列がPCRによる5'及び3'末端を設計するために利用された。テストされた第一組の構成物(図4)では、更なる5'末端配列はHCV-1分離体(Hanらによる(1991)の上記文献)から誘導された。3'NTRにとっては、二つの代わりの構造のあるプラスミドが構築された。3'NTRと末端ポリ(A)トラクトを含む一組(SP6若しくはT7)は、HanらによるHCV-1に報告された(Hanらによる(1991)の上記文献)。第二の組は、全ての他の分離体のコンセンサス3'NTR配列とそれに続く3'末端ポリ(U)トラクトを用いて構築された。

HCV RNAの感染力を検定する方法。潜在的に機能性のあるクローンの初期の確認用の所望の方法は、許容細胞培養のトランスフェクション後のRNA複製のスクリーニングであろう。幾つかの研究所からさまざまな細胞培養中で感染と複製が報告されているが(上記及び下記の発明の背景を参照)、上記システムは極端に非効率的であり、殆ど特徴付けられておらず、再現させることが難しい。生体外で効率的な複製を妨げている要因は知られていないが、ウイルス生命サイクル(吸着、エントリ、RNA複製、構築又は放出)における一つ又は多重の段階を含んでいる。さらに、細胞培養で発生したHCVが“本物である”、例えばチンパンジーモデルにおいて病気を誘発可能であることをだれも示していない。上記理由のために、RNAトランスフェクション後の複製を明らかに示すことに関連した技術的に困難さのために、チンパンジーモデルはライブラリーからの機能性クローンを確認するために利用された。外科的手順及び直接的肝臓内接種が利用された。なぜなら上記技術はラビット出血性疾患ウイルスビリオンRNA(OhlingerらによるJ. Virol. 64: 3331-3336 (1990))の感染性をうまく示し、生体外転写により肝炎AウイルスRNA(EmersonらによるJ. Virol. 66: 6649-6654 (1992))をうまく発生させるからである。

#### チンパンジー実験 I

キャップのある若しくはキャップのない完全長RNA転写物は、各四つの直線化されたプラスミドテンプレートから合成され、経皮肝臓バイオプシー技術を利用して、チンパンジー肝臓の直接肝臓内接種により感染力を検定された。簡単に言えば、RNA転写後に反応はDNaseで消化され、フェノールで抽出され、RNAはエタノール沈殿により集めら

れた。各転写RNAの収量と完全性は、変性条件下でアガロース電気泳動により求められた。転写物（SP6、T7、キャップのある、キャップのない）を含む夫々のポリ（U）若しくはポリ（A）の等量はプール（pool）され、二つの動物に別々に検定された。上記動物は以前はHCV若しくはプールされた血液生成物に晒されておらず、HCV抗体であり、RNAネガティブであった。各動物にとって、二つの注入部位が利用された。一つの部位において、1mlのRNaseのないPBS中の200μgのプールされたRNAが注入された。第二の部位において、0.8mlのRNaseのないPBSと200μlのLIPOFECTIN（BRL）とを混合させた200μgのプールされたRNAが注入された。接種前後の血漿と肝臓バイオプシーサンプルが週単位で集められた。血漿サンプルはALT及びGGTP（肝臓損傷の指標）を、利用可能な血清学的検定法を用いてHCV特異抗体を、及びRT/PCRによるHCV RNAを循環させる証拠を検定された。肝臓バイオプシー組織の組織学的調査に加えて、さらにサンプルは免疫蛍光検査及び電子顕微鏡による可能な分析のために保存された。6ヶ月間動物を追跡したにもかかわらず、上記検定法を利用して生産的HCV感染の証拠は得られなかった。

10

以下に十分に説明される方法を利用して、さらに上記クローンからの転写物は、幾つかの異なる細胞タイプにおける感染力が測定された。ある場合には、HCV抗原は数日間トランスフェクトされた細胞において検出された。しかしながら、同様な結果はNS5B RDRPにおける欠失を含むコントロールHCV転写物を利用して得られ、複製に対して不活性であるべきである。よって、複製の説得力のある証拠は、第一組の実験において得られなかった。

20

### 例3. cDNAからHCVを回収させる第二の試み

試みIの失敗の可能な理由。幾つかの可能な説明は、単独で若しくは組合せて、プロタイプHCV-Hクローン（pTET/HCVFLCMR）からの感染性のあるHCV RNAを回収させるための以前のうまくいかなかった試みの原因に当てはまる。これらには、欠けている若しくは正しくない末端配列、HCV複製にとって内部の誤りのある有害若しくは致死、又は感染及び複製を検定する不十分な方法が含まれる。第一の心配事を扱うために、HCV-H 5' 及び3' 末端配列は正確に決定された。有害な誤りのない完全長クローンを回収するための可能性を増加させるために、高忠実度RT/PCRとアセンブリPCRが、新しい末端配列を含む完全長HCV-Hの新しいライブラリーを構築するために利用された。ライブラリーからの多重クローンは、チンパンジーモデルにおける感染をテストされた。

30

HCV-H末端を正確に決定する原理。上述したように、HCV-Hの5' 及び3' 末端配列は未知であり、機能性転写物を発生させる以前の試みは（例2）、他のHCV分離体に対して決定された末端配列を有するcDNAクローンからのものであった。他のRNAウイルスシステムでの研究により、特異末端配列は機能性である複製能RNAの発生には重要であることが示された（BoyerとHaenni, (1994) の上記文献を参照）。かかる配列はマイナス及びプラス鎖RNA合成の開始に関係していると信じられている。ある場合には、2、3の追加塩基、若しくはさらに長い非ウイルス配列が、5' 及び3' 末端で許容される；典型的には上記配列はなくなる若しくは本物のウイルス複製中に選択される。他のRNAウイルスにとって、余分の塩基、特に5' 末端での塩基は有害である。対照的に、本物の末端配列を欠いた転写物は、通常非機能的である。例えば、3' 末端の第二の構造の欠失又はフラビウイルスゲノムRNAの3' NTRにおける保存配列要素は、YF又はTBE RNA複製には致命的である。他のウイルスにとって上記配列が重要であるならば、我々はHCV-H末端配列をより正確に決定するように試みた。

40

HCV-H 5' NTRの構造。RNAの一番端の5' 末端を増幅させクローン化させるのに利用された方法は、第一の鎖cDNA（5' RACE）へのホモポリマーのテーリング若しくは合成オリゴヌクレオチドの連結合成（ligation）（Schaefer, Anal. Biochem. 27: 255-273 (1995)）、第一の鎖cDNAの環化に続く逆PCR（ZeinerとGehring, Biotechniques 17: 1051-1053 (1994)）又はRNAリガーゼによるゲノムRNAの環化（必要ならば5' キャップ構造を取除くように処理した後）に続くcDNA合成と5' - 3'

50



接合にわたるPCR増幅(MandlらによるBiotechniques 10: 486 (1991))を含む。上記手法の夫々は問題を有し、特に珍しいRNAには問題がある。にもかかわらず、5'末端配列は多くのHCV分離体に対して決定されており、一般的には一致している。HCV-H RNAに対しては、環化/逆PCR及び5' RACE法は、高いタイター血漿からのHCV-H RNAに対して5'末端コンセンサス配列を決定するために利用された(HCV-Hの新しいデータを強調文字で示す): 5'-GCCAGCCCCCTGATGGG G GCGACACTCCACCATGAATC...-3' (SEQ ID NO: 3)。この配列は他の分離体で決定されたものとかかなり相同であるが、二つの位置(下線部)で我々のプロトタイプの完全長cDNA配列とは異なる。低い頻度で、さらに追加の5'残基(通常1つの追加のG)が回収される。表1には5'末端分析の結果を示す。

表1 HCV-H cDNAクローンの5'末端分析の結果

クローンの数	5'末端
18	GCCAGCC...
3*	NCCAGCC...
18*	NNCCAGCC...
9	GGCCAGCC...
3	TGCCAGCC...
1	AGCCAGCC...
2	AAGCCAGCC...
1	GCGCCAGCC...

\* 配列は決定されなかった; 5'末端上のヌクレオチドの数は制限フラグメントの相対的電気泳動移動度により決定された。

18のクローンは配列5'-GCCAGCC...-3'で始まり; 9つのクローンは配列5'-GGCCAGCC...-3'で; 3つのクローンは配列5'-UGCCAGCC...-3'で; 1つのクローンは配列5'-AGCCAGCC...-3'で; 2つのクローンは配列5'-AAGCCAGCC...-3'で; 3つのクローンは配列5'-GCGCCAGCC...-3'で始まる。上記配列決定されたクローンに加えて、1つの追加の5'塩基のある18のクローンは制限分析により確認された。遺伝子型1b分離体に報告された配列は、U残基(5'-UGCCA...-3')で開始するという観察は注目すべきである。上記結果はHCV 5'末端における追加の配列若しくは異質性の存在を示すが、追加の塩基は人工産物であり、5'キャップ構造の部分コピーにより、若しくは第一の鎖cDNA合成中の逆転写酵素による非テンプレート3'塩基の添加により発生する。HCVゲノムRNAの5'末端が5'キャップ構造、又はピコルナウイルスのVPgのような共有結合的に結合した末端タンパク質を含むことを排除することはできない(VartapetianとBogdanov, Prog Nucleic Acid Res Mol Biol 34: 209-51 (1987))。上記可能性は、HCVゲノムRNAの5'末端の構造を直接決定できるようになるまで未解決のままであろう。ペスチウイルスでは、最近の結果からゲノムRNAは5'キャップ(BrockらによるJ. Virol. Meth. 38: 39-46 (1992))を含まず、しかもその構造は転写RNAの感染に必要なではない(MeyersらによるJ. Virol. 70: 8606-8613 (1996a); MeyersらによるJ. Virol. 70: 1588-95 (1996b); MoormannらによるJ. Virol. 70: 763-70 (1996); RuggliらによるJ. Virol. 70: 3478-87 (1996); VassilevらによるJ. Virol. 71: 471-478 (1997))ことを示唆している。

HCV-H 3' NTRの構造。一番端の3'末端HCV配列の決定は、本願の全体の引用文献に含まれる同時係属中の、共同出願人による1995年8月29日に出願された米国特許出願第08/520、678号と、1996年8月28日に出願されたPCT国際特

10

20

30

40

50

許出願第 P C T / U S 9 6 / 1 4 0 3 3 号に説明されている。簡単に言えば、上記結果は H C V 3 ' N T R は三つの要素（プラスセンス、5 ' から 3 ' ）：（i）遺伝子型の間でかなりの変異性のある短い配列；（ii）ホモポリマーポリ（U）トラクトに続く C 残基が散在し、主に U から成るポリピリミジンストレッチ；（iii）9 8 の塩基の新規な配列から成ることを示す。上記新規な 9 8 の塩基配列はヒトゲノム D N A には存在せず、H C V 遺伝子型の間でかなり保存されている。3 ' 末端 4 6 の塩基は安定なステム-ループ構造を形成することが予想される。定量的競争 R T / P C R 検定法を利用して、高特異感染接種物からの H C V ゲノム R N A のかなりのフラクションは、この 3 ' 末端配列要素を含むことが分かった。上記結果より、H C V ゲノム R N A は、本物の H C V 複製に、したがって c D N A から感染性の R N A の回収に必要とされるであろうかなり保存された R N A 要素で終結することを示す。上記結果は二つの他のグループにより確認された（Tanakaらによる（1995）の上記文献；Tanakaらによる（1996）の上記文献；Yamadaらによる（1996）の上記文献）。さらにかかなりの数の臨床分離体は調査され、新規な保存 3 ' 末端要素を含むことが分かった（Umlauftらによる J. Clin. Invest. 34: 2552-2558（1996））。

H C V H 7 7 5 ' と 3 ' コンセンサス配列を含むレシピエントベクター。H C V H 末端配列の我々の分析を基にして、レシピエントベクターは決定されたコンセンサス H 7 7 配列 5 ' の K p n I（5 8 0）と 3 ' から N o t I（9 2 1 9）部位（上記末端 H C V 配列は p 9 0 / H C V F 長い p U の配列と同じである、S E Q I D N O : 5 の以下を参照）を含むように構築された。このベクターは p T E T / T 7 H C V B g l I I / 5 ' 3 ' corr . と称し、以下に説明するコンビナトリアル完全長ライブラリーの構築に利用された。

H C V - H 菌株の完全長 c D N A ライブラリーの構築の更なる検討。先の試み（例 2）に関して、第二回目のトライは、バクテリオファージ D N A 依存 R N A ポリメラーゼを用いて、生体外及び生体内で転写されるプラスミドベクターでの完全長 c D N A テンプレートの構築を含む。正しい 5 ' 及び 3 ' 末端を有することに加えて、さらに R N A 転写物は機能性 H C V ポリペプチドの完全な補体をコード化しなければならない。欠損 H C V ゲノムをクローン化させる可能性を最小にするために、高特異感染性 H C V - H 血漿（H 7 7）は我々の新しいライブラリーのためのピリオン R N A 源として利用される（初期に述べたように、先のクローンは感染チンパンジーの肝臓 R N A から作られた c D N A から集められた）。しかしながら、逆転写と c D N A クローニング前の増幅の多重サイクルは、H C V c D N A テンプレートがウイルス複製に対して有害である一つ又はそれ以上の突然変異を含む可能性を高めた。上記理由のために、完全長クローンの複雑なライブラリーが、高忠実度アセンブリ P C R を利用して、それから感染性 R N A の生産用のプールにおいてスクリーニングされて構築された。

完全長 H C V - H c D N A クローンの新しいライブラリーの構築。我々は 4 1 の H C V プライマーペアをスクリーニングして、ゲノム R N A のオーバーラップ 1 - 4 k b 部分を増幅させるのに有用である 1 1 の組を発見した（図 5 と表 2 及び 3）。

表 2 HCV-H cDNAの増幅に利用されたオリゴヌクレオチド

名前	配列 (5' から 3' へ)	SEQ ID NO:	HCV-Hにおける位置と配向
SF49	GGCGACACTCCACCATAGATC	6	(+) 18-38
SF128	TGGCACTACCCTCCAAGACC	7	(+) 1800-1819
SF162	ATGACACAAGGGGGCGCTCCG CACACT	8	(-) 2027-2053
SF131	TCCTGCTTGTGGATGATG	9	(+) 2538-2555
SF152	TAGTTTGGTGATGTCA	10	(-) 2999-3014
PCL10067	ACATAGGTGCCAGTAAG	11	(-) 3171-3188
PCL10066	CTGGCAACGTGCATCA	12	(+) 3549-3564
CMR115	GGGTGAGAACAATTACCA	13	(+) 4183-4200
CMR117	ATTGATGCCCAATGCG	14	(-) 4565-4580
SF140	ACTGCCTGGGATTCCCT	15	(+) 6347-6363
SF155	CCACAGTGGCAGCGAGTG	16	(-) 6419-6436
SF156	CATGGACGTCAACACG	17	(-) 6848-6863
SF1045	AATCTTCACCGGTTGGGGAGG AGGTAGATG	18	(-) 9353-9391

10

20

表 3 オリジナル及びアセンブリPCRに利用したフラグメントとプライマー

アセンブリにおけるフラグメント	プライマーペア	結果生じたフラグメント	スタート* における位置	HCVゲノム末端*
Original PCR	SF49, SF162	A	39	2026
Original PCR	SF128, SF152	B	1820	2998
Original PCR	SF128, PLC10067	C	1820	3170
Original PCR	SF131, CMR117	D	2556	4564
Original PCR	PCL10066, SF155	E	3565	6418
Original PCR	CMR115, SF156	F	4201	6847
Original PCR	SF140, SF1045	G	6364	9352
A+B	SF49, SF152	H	39	2998
A+C	SF49, PCL10067	J	39	3170
B+D	SF128, CMR117	L	1820	4564
J+L	SF49, CMR117	K	39	4564
F+G	CMR115, SF1045	M	4201	9352
E+G	PCL10066, SF1045	N	3565	9352
L+M	SF128, SF1045	O	1820	9352
H+O	SF49, SF1045	#2	39	9352
J+O	SF49, SF1045	#3	39	9352
K+N	SF49, SF1045	#5	39	9352
K+M	SF49, SF1045	#6	39	9352

\* プライマーを排除

‡ 図5参照

熱安定酵素の混合物は過ち頻度を減少させ、完全長生成物の合成を促進させるために利用された (Barnes, Proc. Natl. Acad. Sci. USA 91: 2216-2220 (1994); Lunbergらによる Gene 108: 1-6 (1991))。かかる中間PCR生成物は結合され、アセンブリPCRの順次ラウンドを用いて完全長HCV cDNAが生じる (Mullisらによる Cold Spring Harbor Symp. 51: 263-273 (1986); Stemmer, (1994) の上記文献)。アセンブリPCRは結合される二つのオーバーラップフラグメントの一番端の末端のプライマーと、限定数の増幅サイクルを利用する (図6)。この手法は、機能性の誤りのないHCV cDNAテンプレートのあるフラクションを含むべき複雑なコンビナトリアルライブラリーを生じさせるという利点を有する。この手法の主な検討は、ライブラリーがあるクローンは誤りがないことを保証するのに十分な複雑度を含むことを確かめることである。初期の夫々の増幅反応では、第一の鎖cDNAの希釈はテストされ (図7)、多重の独立cDNA分子が増幅されたことを示す (7以上から100; 図5に示す)。図7に示すように、完全長ライブラリーは  $5.6 \times 10^5$  ( $80 \times 7 \times 10 \times 10 \times 10$ ) 以上の異なる組合せを含む。プライマー配列がPCR増幅のために選択され、構築が不適当であれば、起こり得る有害な突然変異はクローンの半分へ導入された。しかしながら、不均一性はPCRに利用されたプライマーに相当する配列には存在しないことが後で確認された。

HCV-H77ゲノムの大部分は (ヌクレオチド39-9352から) 集められ上記のように増幅され、レシピエントプラスミド (pTET/T7HCV BglII 5' 3' corr.) へのKpnI (580)-NotI (9219) フラグメントとしてクローン化され

、完全長ライブラリーができる。前述したように、p T E T / T 7 H C V B g l I I 5 ' 3 ' corr . は T 7 プロモーター、コンセンサス H C V - H 5 ' と 3 ' 末端配列、K p n I 部位への 5 ' と N o t I 部位からの 3 ' と、ランオフ R N A 転写物のテンプレート直線化及び生産用の H p a I 部位とを含む。H p a I による直線化は一つの余分な 3 ' U 残基を含むランオフ転写物を生じるように予測されることを指摘しておく。

ライブラリーからのクローンは二つの基準に基づく感染検定法のために選択された。まず一連の制限消化が実行され、明らかに欠失若しくは H C V c D N A への挿入を有するクローンを排除した。233のクローンを分析し、このスクリーンを通過したクローンは、それから予想される H C V ポリタンパク質切断生成物の生産のためにワクシニア T 7 一過性発現システム (Grakouliらによる (1993a) の上記文献; Grakouliらによる (1993c) の上記文献を参照) を用いて分析された。B H K 細胞における予備研究により、H C V I R E S は H C V ポリペプチドの発現の E M C V I R E S と殆ど同じように効率的に機能することが分かったので、完全長クローンは上記技術を用いて直接分析された。129のクローンは慢性 H C V (JHF: Grakouliらによる 1993c の上記文献) の患者からのポリクローナル抗血清を利用してスクリーニングされた。49のクローンは、N S 5 B、H C V - H O R F における C 末端タンパク質の生産のために分析された (Grakouliらによる 1993a の上記文献; Grakouliらによる 1993c の上記文献)。上記テストを通過した 34 のクローン (予想制限パターン; 無傷の O R F 及び正しいプロセッシング; N S 5 B 生産) は、潜在的な感染力のある R N A の生体外での転写と感染分析のために選択された。

内部ポリ (U / U C) トラクトと 98 の塩基要素を含む完全長 H C V R N A の転写の特別な条件。T 7 駆動転写では、生体外転写条件は最適化され、結果生じた R N A は一番端の 3 ' 末端配列を含むことを示した。T 7 R N A ポリメラーゼ終結シグナル (第二の構造とそれに続くポリ U) は 3 ' の新規な要素に先行する H C V 配列と似ており、我々はこの部位で終結を観測したので、上記のことは特に関心があった。加えて、酵素はポリ (U / U C) トラクト内部の終結を早期に始める傾向があるようだ。図 8 A に示すように、転写反応中の U T P 濃度を 3 m M まで増大させることにより、高収量の完全長 H C V R N A 転写物を得た。この点に関して、T 7 ポリメラーゼは S P 6 ポリメラーゼよりも明らかに良く、U T P の比較的高濃度でさえも、ポリ (U) トラクトでのかなり早い終結を示した。

#### チンパンジー実験 I I

本質的には前述したように (例 2)、外科的手順と直接肝臓内接種が転写 R N A の感染力を検定するのに利用された。先の H C V 研究には利用されなく、H C V 血清学及び R N A ではネガティブであった三つの動物が接種された。夫々二つの動物は 17 の独立のクローンからの R N A 転写物で、肝臓の 34 の個々の部位に注入され接種された。各転写物調製に利用された二つの個々の接種は、ある部位に注入された P B S 中の 50 - 100  $\mu$  g の R N A と、第二の部位への 10  $\mu$  g の lipofectin (R N A トランスフェクションを促進させるカチオン性リポソーム (Riceらによる (1989) の上記文献参照)) と混合させた 1  $\mu$  g の R N A であった。この手順は各クローン / R N A 調製用の生産的トランスフェクションの可能性を最大限にする目的であった。ネガティブコントロールとして、第三の動物 (チンパンジー 1557) は、H C V R N A 依存 R N A ポリメラーゼ (G D D と呼ばれる) の活性部位を取り囲む N S 5 B における 21 の残基のインフレーム欠失を含む転写物 (約 1500  $\mu$  g) で 34 の部位を同様に接種させた。接種に続き、血清サンプルを集めて (一週間間隔で) H C V R N A、肝臓トランスアミラーゼの上昇及び H C V 特異抗体を分析した。実験動物又はネガティブコントロール動物 (G D D) のいずれも生産的感染の兆候を示さなかった (循環 H C V R N A、上昇した肝臓酵素、組織病理学)。更なる実験では、接種後の一週間ぐらいの早い段階でさえも検出可能な循環 H C V R N A を全く存在しなかったことは注目すべきことであった。

#### 例 4 : c D N A から感染性のある H C V の成功した回収

H C V - H コンセンサス配列の決定。限定されたプールのスクリーニング手法はうまくいかなかったので、我々は H C V - H 菌株の完全なコンセンサス配列を決定した。上記配列

決定されたクローンのセグメントは、コンセンサス配列を有する完全長HCV-Hクローンの方向性のある集合に対して利用された。この手順は、cDNA合成若しくはPCR増幅中に起こるであろう、又はオリジナルHCV集団に存在する致死突然変異を排除するように期待される。したがって、コンセンサス方法は機能性HCVを発生させる大きな可能性を有していた。

表 4 HCV-Hコンセンサス配列を決定するために利用した配列情報

名称	説明	
HCV-H CMR	CMRプロトタイプHCV-H cDNA; 感染されたチンパンジー肝臓RNA (SEQ ID NO: 19)	10
HCV-H GenBank	HCV-H配列	
AAK#83	コンビナトリアルライブラリークローン#83; H77血清	
AAK#84	コンビナトリアルライブラリークローン#84; H77血清	
AAK#86	コンビナトリアルライブラリークローン#86; H77血清	20
AAK#87	コンビナトリアルライブラリークローン#87; H77血清	
AAK#89	コンビナトリアルライブラリークローン#89; H77血清	
AAK#90	コンビナトリアルライブラリークローン#90; H77血清	
AAK#92	コンビナトリアルライブラリークローン#92; H77血清	
AAK#93	コンビナトリアルライブラリークローン#93; H77血清	30
AAK#96	コンビナトリアルライブラリークローン#96; H77血清	
AAK#99	コンビナトリアルライブラリークローン#99; H77血清	
AAK#101	コンビナトリアルライブラリークローン#101; H77血清	
AAK#248	コンビナトリアルライブラリークローン#248; H77血清	
AAK#227	コンビナトリアルライブラリークローン#227; H77血清	40
AAK#213	コンビナトリアルライブラリークローン#213; H77血清	
AAK#211	コンビナトリアルライブラリークローン#211; H77血清	
AAK#209	コンビナトリアルライブラリークローン#209; H77血清	
AAK#12	コンビナトリアルライブラリークローン#12; H77血清	
HCV cDNAにおけるKpnI(580)とNotI(9219)部位の間の完全な配列は、クローンAAK#248、AAK#227、AAK#213、AAK#211、AAK#209、とAAK#12に対して決定された。プロトタイプHCV-H CMR		
		50

の配列 (Daemerらによる上記文献; Grakouliらによる (1993c) の上記文献) と H C V - H G e n B a n k (Inchauspeらによる (1991) の上記文献) は以前に決定された。上記配列は図9に並べる。ドットは、下 (S E Q I D N O S : 19と20) に示す H C V - H C N R 配列に同じ位置を示し、点線はギャップを示し、配列 “ P C R s e q ” は P C R 増幅 H C V - H 7 7 c D N A の直接配列決定により求めた。我々のコンピナトリアルライブラリー (A A K # 83、# 84、# 86、# 87、# 89、# 90、# 92、# 93、# 95、# 96、# 99、# 101) からの追加のクローンの配列は、E 2 (多くはヌクレオチド 1464-1823の間で配列決定された。) における H V R 1 超可変領域に対して決定された。アライメントの検査により、多くの位置で H C V H 7 7 共通配列 (S E Q I D N O : 1) を示す。しかしながら、ある位置では明らかなコンセンサス配列は出現しなかった。上記可変な位置は、2170 (G a c 対 A a c ; 可変塩基が上のケースタイプで示される)、3940 (g A g 対 g G g) と 5560 (c a A 対 c a T) であった。上記の場合には、コンセンサスクローンに利用された配列は、配列決定された H C V 分離体の大部分の位置にて発見されたアミノ酸を生じるヌクレオチドに相当した。コンセンサス配列の決定に関して、H C V ゲノムの追加の領域はさらなるコメントをするのに値する。第一に、E 2 の N 末端部分はかなり可変であり、免疫選択のターゲットであると信じられている (Houghton, (1996) の上記文献)。H 7 7 サンプルにおいて、かなりの変異性が H V R 1 に存在する (Nakajimaらによる J. Virol. 70: 3325-9 (1996) ; Ogataらによる (1991) の上記文献)。この領域からの多重の独立クローンは配列決定され、各位置における主要な H V R 1 配列はコンセンサスクローンにおいて利用された。利用された主要な配列は他人により決定されたもの (Inchauspeらによる (1991) の上記文献; Nakajimaらによる (1996) の上記文献; Ogataらによる (1991) の上記文献) とは位置が異なる。しかしながら、上記配列は、H 7 7 接種チンパンジーから分離された肝臓 R N A から誘導されたプロトタイプ H C V - H クローンのそれとかなり類似している。よって、上記配列はチンパンジーの H C V 複製には許容であろう。以下に示すように、上記配列は機能的であるが、さらに多くの他の H V R 配列変異も許容であろう。

H C V - H 配列の第二の領域である 3' N T R ポリ (U/U C) トラクトの長さや成分ははっきりとは決定されていなかった。二本鎖 c D N A の十分な量が、P C R 増幅にたよらずに上記領域の直接クローニングに対しては得られなかった。P C R 増幅は上記ホモポリマートラクトの長さを収縮及び拡張させることができる。よって、この手順から生じるクローンは固有 H C V ゲノム R N A 構造を反映しない。P C R 増幅により誘導された多重独立クローンにおいて、上記トラクトの長さは 41 から 133 のヌクレオチドへ変化した (Kolykhalov による (1996) と特許出願番号第 08/520、678 号を参照)。よって、ポリ (U/U C) トラクトの二つの異なる長さがテストされた: “ 短い (75 の塩基) ” 又は “ 長い (113 の塩基) ”。 “ 短い ” トラクトの長さは、実際には我々 (Kolykhalov による (1996) の上記文献) 又は他の研究者 (Tanakaらによる (1995) の上記文献; Tanakaらによる (1996) の上記文献; Yamadaらによる (1996) の上記文献) により報告された全ての配列 (異なる遺伝子型) のほぼ中間の長さである。 “ 長い ” トラクトは一つの H C V - H クローン (p G E M 3 Z f (-) H C V - H 3' N T R # 10) のみにおいて回収された; 同様な長さのトラクトは、遺伝子型 4 分離体 W D (Kolykhalov による (1996) の上記文献) の一つのクローンにおいて回収された。かかる長いポリ (U/U C) トラクトは他の研究者らにより報告されていない (Tanakaらによる (1995) の上記文献; Tanakaらによる (1996) の上記文献; Yamadaらによる (1996) の上記文献)。

5' 末端配列、サイレントマーカー、3' N T R ポリ (U/U C) トラクトの長さや 3' ランオフ部位における変異。追加の塩基が幾つかの H C V c D N A クローンの 5' 末端において発見され、ポリ (U/U C) トラクトの長さについて不確実であるならば、幾つかの代替のクローンが作られた。誘導体が後の分析において機能的であることを確認するためのマーカーとして働き、複製ウイルスが構築 c D N A クローンから実際に回収されたことを示すために、サイレントヌクレオチド置換が O R F に組込まれた。先に使用した H p a I 部位を置き換えることは、B s m I 部位が H C V c D N A の 3' 末端に続いて

10

20

30

40

50

生じ、HCVゲノムRNAの正確な3'末端に対応するランオフ転写物の生産を可能にする。上記構築の詳細な説明は以下のものである。

5'末端での追加の塩基。pTET/T7HCV BglII+3'corr.と呼ばれる最も頻繁に出現する5'末端配列(5'-GCCA...-3')を含むレシピエントクローンは、チンパンジーにおいてテストされたプロトタイプHCV-H cDNAクローンの一つであるpTET/HCV5'T7G3'AFIのBssHII(479)からKpnI(580)フラグメントをサブクローニングすることにより変種され、p67/HCV BglII/5'+3'/XhoI-が生じた。上記クローンは514位置におけるXhoI部位の存在(pTET/T7HCV BglII/5'+3'corr.)又は(p67/HCV BglII/5'+3'/XhoI-)存在しないにことにより異なる。それからp67/HCV BglII/5'+3'/XhoI-は、異なる5'末端配列のある四つの誘導体を構築させるためのベクターとして利用された。

プラスミド	T7転写物の5'配列	マーカー (位置)
p70/HCVΔBglII/5'+3'/XhoI-/GG	5'-GGCCA...-3'	XhoI- (514)
p71/HCVΔBglII/5'+3'/XhoI-/GAG	5'-GAGCCA...-3'	XhoI- (514)
p72/HCVΔBglII/5'+3'/XhoI-/GUG	5'-GUGCCA...-3'	XhoI- (514)
p73/HCVΔBglII/5'+3'/XhoI-/GCG	5'-GCGCCA...-3'	XhoI- (514)

上記誘導体は適当な合成オリゴヌクレオチドとPCR増幅と配列分析により確認された構造を利用して構築された。

KpnI(580)とNotI(9219)の間のコンセンサス配列を含むクローンの構築。構築段階のスキームを図10に示す。7つの配列決定されたHCV-Hクローンを、プロトタイプコンセンサスクローンを構築するのに利用された。プラスミド源、HCV cDNAにおける位置、構築に利用された制限部位を表5にまとめる。



表 5 コンセンサスクローン構築に利用されたクローン、フラグメントと制限部位

クローンの フラグメント数のソース	HCVゲノムにおける位置	利用した制限部位
313	580-1046	<i>KpnI-Xho I</i>
248	1046-1174	<i>Xho I-PpuM I</i>
12	1174-1357	<i>PpuM I-BamH I</i>
209	1357-1482	<i>BamH I-Sal I</i>
227	1482-1748	<i>Sal I-PpuM I</i>
209	1748-1908	<i>PpuM I-Asc I</i>
227	1908-2108	<i>Asc I-BspE I</i>
312	2108-2322	<i>BspE I-Sst I</i>
CMR	2322-2440	<i>Sst I-Sca I</i>
213	2440-2526	<i>Sca I-BssH II</i>
CMR	2526-2828	<i>BssH II-Hinf I</i>
211	2828-2978	<i>Hinf I-BsrG I</i>
209	2978-3236	<i>BsrG I-Bgl II</i>
227	3236-3478	<i>Bgl II-Bgl I</i>
209	3478-3733	<i>Bgl I-SexA I</i>
12	3733-3942	<i>SexA I-Bfa I</i>
211	3942-4069	<i>Bfa I-Spl I</i>
227	4069-4545	<i>Spl I-Sst I</i>
248	4545-4646	<i>Sst I-Sal I</i>
211	4646-4976	<i>Sal I-Sma I</i>
227	4976-5610	<i>Sma I-Xho I</i>
209	5610-5750	<i>Xho I-Eae I</i>
CMR	5750-6209	<i>Eae I-Bsu36 I</i>
213	6209-6302	<i>Bsu36 I-Blp I</i>
227	6302-7529	<i>Blp I-Blp I-BamH I</i>
213	7529-9219	<i>BamH I-Not I</i>
209	7861-8205	<i>Hind III-EcoR I</i>

構築の最終段階ではKpnI-NotIコンセンサ領域をレシピエントベクターpTE  
T/T7HCV BglII/5' +3' corrへサブクロニングさせることを含み、  
p61/HCVFLconsが生成した。

HCV cDNAとBsmIランオフ部位におけるBsmI-置換の導入。先に使用した

10

20

30

40

50

HpaIランオフ部位は有害である追加の3'末端U残基のある転写物が生じたので、クローンは正確なHCV3'ヌクレオチドで終結する転写物が合成されるように再度特別に設計された。このことはBsmI部位をHCV3'末端からの適当な下流位置に位置決めすることにより達成された。BsmIによる切断により、HCV3'末端に相当する位置で終結するテンプレート鎖が生じる。H77コンセンサス配列は位置5934にBsmI部位を含むので、この部位は部位特異的変異の導入により特別設計された翻訳的サイレント置換により不活性化された。

構築の本シリーズの第一の段階は、HCV H77 cDNAにおけるBsmI部位を不活性化させることであった。p62/HCVFLcons/Bsm(-)と呼ばれるこのクローンは、以下に含まれる四つのフラグメント連結反応において発生した：(1) BsmI部位(位置5934でのAに代わりにC)を不活性化するサイレント置換を含むSacI(5923)とSau3AI(5942)の間のアニールされた合成オリゴ；(2) p61/HCVFLconsのNsiI(5282)からSacI(5923)のフラグメント；(3) p61/HCVFLconsのSau3AI(5942)からBsu36I(6209)；(4) p61/HCVFLconsを消化させたBsu36I(6209)とNsiI(5282)。p62/HCVFLcons/Bsm(-)は、構築されたコンセンサスクローンの構造と、位置899(Tに代わりにC)でのサイレントマーカース突然変異、切断された(ablated)BsmI部位、及び位置8054(以下参照)でのサイレントマーカース突然変異の存在とを確認して、完全に配列決定された。

3' BsmIランオフ部位を含む中間体プラスミドp65/3'HCVBsm(+)/Not-Mluは、以下の三つのフラグメント連結反応により発生した：(1) BsmI部位(5'-tgTcgcatctc-3'(SEQ ID NO:21)ここで強調文字で示したヌクレオチドはBsmI部位を指し、上の場合ヌクレオチドはHCVゲノムの3'末端塩基に相当する)を含むSau3AI(9639)とMluI(9656)の間のアニールされた合成オリゴ；(2) p62/HCVFLcons/Bsm(-)のNotI(9219)からSau3AI(9639)のフラグメント；(3) p61/HCVFLconsのMluI(9656)からNotI(9219)。上記クローンは内部BsmI部位(5934)と特別設計されたBsmIランオフ部位の双方を含むことを指摘しておく。オリジナルコンセンサス完全長クローンであるp61/HCVFLconsは、NS5B暗号づけ領域においてサイレント置換を含む(位置8054でのGに代わりA)。この置換は“短い”ポリ(U/UC)トラクト(三つのクローンは位置8054にAを含む)又は“長い”ポリ(U/UC)トラクト(位置8054にはGを有する)を含むクローン間を区別させるためのマーカーとして利用された。長いポリ(U/UC)と位置8054でのGを含むp90/HCVFL長いpU(SEQ ID NO:5)は、四つのフラグメントの連結反応により構築された：(1) p62/HCVFLcons/Bsm(-)のXbaI(-20)からHindIII(7861)；(2) 位置8054にG残基を含むライブラリークローンAAK#209(図9)のHindIII(7861)からEcoRI(8025)；(3) p62/HCVFLcons/Bsm(-)のEcoRI(8205)からNotI(9219)；(4) p65/3'HCVBsm(+)/Not-MluのNotI(9219)からXbaI(-20)。

“短い”ポリ(U/UC)トラクトと位置8054でサイレントマーカーAを含む誘導体であるp91/HCVFL短いpUは、以下のフラグメントの連結反応により発生した：(1) pGEM3Zf(-)HCV-H3'NTR#8のBglI(9398)からNheI(9520)；(2) p65/3'HCVBsm(+)/Not-MluのNheI(9520)からMluI(9597)；(3) p62/HCVFLcons/Bsm(-)のMluI(9597)からNotI(9219)。上記構築のナンバリングは最終のp91/HCVFL短いpU配列について言う。

長いポリ(U/UC)と5'末端での追加のヌクレオチドのある完全長構成物の最終セットを発生させるために、p90/HCVFL長いpUのKpnI(580)からMluI(9656)フラグメントが、p70/HCV BglII/5'+3'/XhoI-/GG、

10

20

30

40

50

p 7 1 / H C V B g l I I / 5 ' + 3 ' / X h o I - / G A G、p 7 2 / H C V B g l I I / 5 ' + 3 ' / X h o I - / G U G と、p 7 3 / H C V B g l I I / 5 ' + 3 ' / X h o I - / G C G ヘクロン化され、それぞれ p 9 2 / H C V F L 長い p U / 5 ' G G、p 9 3 / H C V F L 長い p U / 5 ' G A G、p 9 4 / H C V F L 長い p U / 5 ' G U G、p 9 5 / H C V F L 長い p U / 5 ' G C G G C G が生じた。

短いポリ ( U / U C ) のある完全長構成物の類似のセットを発生させるために、p 9 1 / H C V F L 短い p U の K p n I ( 5 8 0 ) から M l u I ( 9 5 9 7 ) フラグメントが、p 7 0 / H C V B g l I I / 5 ' + 3 ' / X h o I - / G G、p 7 1 / H C V B g l I I / 5 ' + 3 ' / X h o I - / G A G、p 7 2 / H C V B g l I I / 5 ' + 3 ' / X h o I - / G U G と、p 7 3 / H C V B g l I I / 5 ' + 3 ' / X h o I - / G C G ヘクロン化され、それぞれ p 9 6 / H C V F L 短い p U / 5 ' G G、p 9 7 / H C V F L 短い p U / 5 ' G A G、p 9 8 / H C V F L 短い p U / 5 ' G U G、p 9 9 / H C V F L 短い p U / 5 ' G C G が生じた。

上記 1 0 のクロンの目立った特徴を ( 5 ' 塩基、サイレントマーカー、ポリ ( U / U C ) の長さ)、図 1 1 にまとめる。プラスミドは E . c o l i ( t e t <sup>s</sup> S U R E 菌株 ) 中で増殖され、精製プラスミド DNA は、C s C l 勾配での 2 回バンディング ( banding ) を含む標準的な方法により調製された ( Ausubel らによる Current protocols in molecular biology. eds. Greene Publishing Associates, New York ( 1993 ) ; Sambrook らによる Cold Spring Harbor laboratory, Cold Spring Harbor, NY ( 1989 ) ) 。

完全長 RNA の転写。上述したように、T 7 転写反応における U T P 濃度が 3 m M まで増加することは、ポリ ( U / U C ) トラクトの読み過しを容易にすることにより、完全長 H C V RNA の収量が増加する。U T P ( 3 m M ) に対する他の r N T R ( 1 m M ) 斜傾 ( skewed ) 比は U 残基の誤った取込みを導き、特に他の N T R が実質的に除去されたときに転写反応を遅くする。この心配事は、他の三つの N T R の濃度を 3 m M まで増加させることにより回避された。精製プラスミド DNA は B s m I により消化され完成し、フェノール-クロロホルムにより一度抽出され、エタノールで析出させた ( Ausubel らによる ( 1993 ) の上記文献 ; Sambrook らによる ( 1989 ) の上記文献 ) 。DNA ペレットを E t O H で洗浄し、塩を取除き、R N a s e のない H <sub>2</sub> O 中で再懸濁させた。転写反応 ( 1 0 0 μ l ) は以下の成分を含む : 1 0 μ g の B s m I 直線化されたテンプレート DNA , 4 0 m M の Tris-Cl、p H 7 . 8、1 6 m M の M g C l <sub>2</sub>、5 m M の D T T、1 0 m M の N a C l、各 3 m M の r N T R、1 0 0 ユニット T 7 RNA ポリメラーゼと 0 . 0 2 U 無機ピロホスファターゼ。3 7 °C での一時間のインキュベーション後に、典型的収量は、ゲル電気泳動により見積られた 8 0 % 以上の完全長 RNA の約 3 0 0 μ g であった。

#### チンパンジー実験 I I I

1 0 のコンセンサスクローンからの転写物は、前述した本質的には同じ外科的手順を利用して、二つの異なる動物へ接種させるのに利用した。実験手順は F D A 及び N I H 動物研究委員会により評価され認可された。動物は全ての肝炎ウイルスに対して血清陰性であり、入れ子 R T - P C R により H C V RNA に対して陰性であり、肝臓酵素の正常なベースラインを有していた。二つの異なる接種/トランスフェクション実験手順を利用した。

チンパンジー # 1 5 3 5 では、1 0 0 μ l の転写反応は 4 0 0 μ l P B S で希釈され、接種に利用されるまで - 8 0 度で冷凍保存された。これらの保存条件がテストされ、H C V

RNA 転写物の完全性に対する観察可能な効果は何も示さなかった。接種前に、サンプルは解凍され、各サンプルは肝臓内の 2 箇所に入力された ( 1 箇所につき約 0 . 2 5 m l ) 。1 0 のクロンの注入箇所は肝臓の三つのローブ ( lobe ) に分布された。この手順のポジティブコントロールとして、チンパンジー # 1 5 5 7 は二つの異なる A 型肝炎ウイルスクローンからの RNA 転写物により、同様に接種させた。この場合には、クローン当たり 8 0 - 1 0 0 μ g の転写 RNA が 2 箇所に接種された。第三の動物であるチンパンジー # 1 5 3 6 は、lipofectin と混合させた RNA の少量で接種させた。この場合には、1 0 の完全長 H C V - H 7 7 クローンからの同じ転写 RNA は、D N a s e I で処理され、テンプレート DNA を取除き、0 . 1 5 μ g、0 . 5 μ g と 1 . 5 μ g 部分が P B S で 5 0 μ l へ希釈され、接種に利用されるまで - 8 0 度で保存された。解凍後、9 μ g の lipofec

tin (Besthesda Research laboratory) を含む  $100\mu\text{l}$  の P B S が各サンプルに添加され、混合し単一の部位へ注入された。よって、異なる R N A / lipofectin 比を有する各クローン/転写調製物が三つのさまざまな部位へ注入された。

血清サンプル及び肝臓バイオプシーは接種前に取られ、その後1週間間隔で取られた。接種後殆ど2ヶ月間、サンプルは肝臓酵素 ( A L T , I C D , G G T P ) 肝炎ウイルス血清学と、定量的競争 R T - P C R ( Kolykhalov らによる ( 1996 ) の上記文献 ) によるウイルス血症を検定された。

感染と複製がうまく開始されたことの証明。これまでの我々の結果を表6にまとめる。

表 6 チンパンジー実験IIIの結果

チンパンジー#1535 (PBS中のRNA-DNA)						
週	ALT	ICD	GGTP	アンチー HCV a b	HCV RNA bDNA (Meg/ml)	定量的競争 P C R
-5	43	453	28	0.2	-	-
-2-3	32	325	27	0.1	-	-
-1	36	600	27	0.2	-	-
0	40	430	28	0.1	<0.2	<10 <sup>2</sup> /ml
1	42	490	24	0	0.445	1x10 <sup>5</sup> /ml
2	96C	1000	53	0	0.283	3x10 <sup>5</sup> /ml
3	81C	780	55	0	0.593	6x10 <sup>5</sup> /ml
4	78	640	52	0.2	2.026	1x10 <sup>6</sup> /ml
5	60	510	57	0.1	2.609	2x10 <sup>6</sup> /ml
6	49	670	50	0.1	3.286	T.B.D.
7	49	525	44	0	5.708	T.B.D.
8	56	485	50	.01	T.B.D.	T.B.D.
9	67	500	67	0.1	T.B.D.	T.B.D.
10	98	725	79	0.2	T.B.D.	T.B.D.
11	86	525	85	0.2	T.B.D.	T.B.D.

10

20

30

チンパンジー # 1 5 3 6 (RNA + lipofectin):						
週	ALT	ICD	GGTP	アンチー HCV a b	HCB RNA bDNA (Meg/ml)	定量的競争 P C R
-9	27	368	33	0.1	-	-
-5	45/4 5	524/49 6	82/77R	0.2	-	-
-2-3	28	375	52	0.1	-	-
-1	34	475	41	0.1	-	-
0	36	680	44	0.1	<0.2	<10 <sup>2</sup> /ml
1	45	660	42	0	<0.2	1x10 <sup>4</sup> /ml
2	44	875	51	0	0.252	3x10 <sup>5</sup> /ml
3	49	760	55	0	0.469	1x10 <sup>6</sup> /ml
4	41	465	52	0.2	0.862	2x10 <sup>6</sup> /ml
5	42	500	49	0.1	0.904	3x10 <sup>6</sup> /ml
6	50	730	60	0.00	1.489	6x10 <sup>6</sup> /ml
7	43	490	55	0.1	3.413	T.B.D.
8	53	700	64	0.1	13.00	T.B.D.
9	38	505	65	0.1	3.271	T.B.D.
10	133	1270	120	0.4	T.B.D.	T.B.D.
11	324	1485	258	1.3	T.B.D.	T.B.D.

チンパンジー # 1 5 5 7 (PBS中のHAV RNA+DNA) 、 ポジティブコントロール				
週	ALT	ICD	GGTP	アンチー-HAV
0	33	405	19	(-)
1	42	360	14	(-)
2	33	345	16	0.6
3	26	520	14	0.7
4	62	1330	24	3.5
5	43	700	28	21.4
6	23	650	27	27.9
7	22	540	25	14.6
8	20	490	22	T.B.D.

R=繰り返し

C=確認

T. B. D. =未定

チンパンジー # 1 5 3 5 は接種後 2 週間で肝臓酵素はピークを示し、徐々に接種前のベ

10

20

30

40

50

スラインに減少した。10週間後に、肝臓酵素の第二のピークが観察された。HCV RNA タイターは接種前は我々の検出限界以下であったが( $< 10^2$ )、1週間で $10^5/\text{ml}$ まで増加し、5週間で $2 \times 10^6/\text{ml}$ に達するまでに着実に増加した。これは1週間後と比較して20倍の増加を示す。

チンパンジー # 1536 は、2週間後でのICDレベルに僅かなピークを示して初期に肝臓損傷のはっきりとした証拠は示さず、その後値は変動した。しかしながら、酵素のかなり増大したレベルが10及び11週間後に観察された。さらにこの動物は、10及び11週間後にHCV血清陽性になった。1週間では、HCV RNA タイターは $10^4/\text{ml}$ であったが、6週間後には $6 \times 10^6/\text{ml}$ まで増加した。このことは1週間と比較して600倍増加したことを示す。

HAV転写物で接種させたポジティブコントロール(チンパンジー # 1557)は、4週間で肝臓酵素に鋭いピークを示し、この時期に明らかに血清学的に変換した(seroconverted)ことを示した。HAV特異免疫反応は5週間で鋭く増加し、その後高レベルを持続させた。上記の結果はHAV感染の明らかな証拠を示し、チンパンジー # 1535 に利用した接種方法は有効である。

さらにHCV RNAに対して分析された全てのサンプルは、逆転写段階での酵素を無視することにより残留テンプレートDNAの存在を測定された。生成物は全く得られず、定量的競争PCR検定法において検出されたシグナルはRNAに由来するものであることを示す(図12)。加えて、上記サンプル物質を含むHCV RNAは、完全に分解された裸の競争者RNAが分析される血清と混合している同じ条件下でRNase消化に抵抗性があつた(図13)。残留接種RNAとは反対に、RNAがエンベロープRNase耐性ウイルス粒子にパッケージされたなら、このことは予想された結果である。さらに、接種に利用された転写RNAの全量はチンパンジー # 1535 で $3000 \mu\text{g}$ 、チンパンジー # 1536 でたったの約 $22 \mu\text{g}$ であった。約150倍以下のRNA接種にもかかわらず、チンパンジー # 1536 はチンパンジー # 1535 よりも高いレベルのウイルス血症を示した。よって、ウイルス血症のレベルはインプットRNAとは相関関係にはなく、ウイルス増幅とスプレッドの暗示する。最終的には、非コンセンサスコンビナトリアルライブラリークローンとGDDネガティブコントロールを利用した先のネガティブな実験において(例3)、 $1000-2000 \mu\text{g}$ のHCV特異RNAが、同様の手順を利用して動物当たりにつき接種された。1週間後又はその後でもHCV RNAは検出されておらず、この実験で観測されたシグナルは本物のウイルス複製と、血清に放出によるものであることを示唆している。

接種転写RNAに由来する上記動物で観測された感染の証拠は、制限酵素と特別設計されたマーカーの存在に対する回収ウイルスの配列分析により得られた。二つのサイレント突然変異は全てのトランスフェクトされたRNAをマークさせた。このことは位置899での置換(Tに代わりC)であり、内部BsmI部位(5934)を切断する位置5936での置換(Aに代わりC)であった。ヌクレオチド899マーカーにとって、466から950の間の領域は入れ子RT-PCRにより増幅され、直接配列決定され、位置899でのサイレントC(Tに代わり)マーカーを含む予想されたH77配列を有することを示した。さらに5801から6257の領域は入れ子RT-PCRにより増幅され、BsmIにより消化に対して耐性があることが分かった。しかしながら、H77 cDNA配列の上記領域(SstI(5923); BspHI(5944); Bsu36I(6209); RsaI(6244))において切断する他の四つの酵素に対して、予想される生成物が得られた。この分析はチンパンジー # 1535 (5週間)とチンパンジー # 1536 (6週間)の双方で行われた。

RNA接種動物の病原プロファイルは、H77物質若しくは他のHCVを含むサンプルで接種させたチンパンジーでの先の実験で得られたものを暗示させる。ヒトと同様に、チンパンジーでの病気の進行は、肝臓の損傷の程度、慢性への進行、ウイルス血症のレベル、血清変換(seroconversion)のタイミングに関して、かなり可変である。

感染動物から回収されたウイルスに存在するサイレントマーカーを評価することにより機

10

20

30

40

50

能性“感染性のある”クローンの確認。前述のように、追加のサイレントマーカ―は、5'末端配列と感染を開始させるに必要な若しくは好ましいポリ(U/UC)トラクトの長さを確認するのを助けるために組込まれた。

単一のG(5'-GCCA...-3')を含む転写物は、Cタンパク質暗号づけ領域でのXhoI(514)サイレントマーカ―の存在により、追加の5'残基と区別された。上記マーカ―を含む領域は、代表的数の独立cDNAが分析される(この場合には50以上)ことを保証する条件下でRT-PCRにより増幅された。結果生じた生成物は、XhoI若しくはコントロールとしてのAccI、全ての入力クローンのフラグメントを消化すべき酵素のいずれかによる消化を分析された。チンパンジー#1535(3週間後のサンプル)では、XhoIで消化された生成物のフラクションは入力接種物に平行であった: 約20%はXhoI(4Uと30Uの双方)で消化され、80%は消化に対して抵抗があった(値はIC1000イメージングシステムにより臭化エチジウム染色消化を走査させることにより決定された)。完全な消化はAccIで観測された。チンパンジー#1536で分析された4週間後のサンプルにおいて、55%はXhoIで消化され、45%は消化に対して抵抗性があった。また、完全な消化はAccIで観測された。よって、第二の動物において、単一G(5'-GCCA...-3')のみの転写物で利点が観測された。特異感染性での在り得る差異に関する上記データから確固たる定量結果を引き出すことは可能ではないが、明らかに結果は、追加のヌクレオチドのない転写物は感染力があることを示している(クローンp90/HCVFL長いpUとp91/HCVFL短いpU)。これまでの我々の分析はさまざまなクローンを区別することはできないが、さらに追加のヌクレオチドのある転写物も感染を生じさせる。

“短い”若しくは“長い”ポリ(U/UC)トラクトを含む転写物は、NS5B暗号づけ領域の位置8054にあるサイレントマーカ―により区別される。7955と8088の間の領域は、100以上の分子的にクローン化された独立cDNA分子の増幅を保証する十分なcDNAを利用して、RT-PCRにより増幅された。10若しくは9の独立クローンの配列は、それぞれチンパンジー#1535(3週間後)とチンパンジー#1536(4週間後)で決定された。チンパンジー#1535の10のクローンのうち9つ(90%)は、位置8054でGを含み、“長い”ポリ(U/UC)トラクトを示唆している。チンパンジー#1536の9つのクローンのうち6つ(66%)は、位置8054でGを含み、“長い”ポリ(U/UC)トラクトを示唆している。結果は“短い”若しくは“長い”ポリ(U/UC)トラクトのいずれかを含む転写物は感染力があるが、“長い”ポリ(U/UC)トラクトが好ましいようだ。しかしながら、我々は上記効果が8054でのマーカ―突然変異の有害な効果によるものである可能性を排除することはできない。上記の追加の分析は、これらの動物で観測されたウイルス血症は我々の完全長クローンから誘導された転写物により開始されたことをさらに確認することを与える。

本例で説明された機能性遺伝子型1a cDNAクローン、又は他のHCV遺伝子型の機能性クローン(同様な方法で構築され確認された)には、(i)より効果的なHCV治療; (ii)HCVワクチン; (iii)HCV診断薬と; (iv)HCVに基づく遺伝子発現ベクターの開発のさまざまな応用がある。

#### 例5: 肝細胞系統の生産的HCV感染

pCENからのEcoRI-BstBIフラグメントは、p90/HCVFL長いpUのユニークなSfiI部位へクローン化された。連結反応前に、突出した末端はdNTRの存在下でT4 DNAポリメラーゼを用いて、鈍く終結した(blunt end)。pCENからのEcoRI-BstBIフラグメントは、EMCV IRES要素とそれに続くネオマイシン耐性(NEO)暗号づけ領域を含む。このIRES NEOカセットはGhattasらにより説明されたものと本質的には同じである(Mol. Cell Biol. 11: 5848 (1991))。正しい配向(HCVゲノムRNAに関してプラスセンス(positive-sense))の上記カセットを含むクローンは、適当な制限酵素での消化により確認された。

EMCV IRES NEOカセットはp90/HCVFL長いpUの3'NTRにおけるSfiI部位へ挿入された。この転写RNAはヒト肝細胞系統をトランスフェクトさせ

10

20

30

40

50

るのに利用され、それから G 4 1 8 を利用してネオマイシン耐性に対して選択された。多くの細胞は死滅したが、G 4 1 8 集団は数ヶ月以上の間成長した。驚くことに、H C V R N A は細胞当たり約 1 0 0 0 R N A 分子のコピー数で、上記細胞にいまだに存在しているらしい。ネオマイシン耐性は H C V R N A により媒介されると信じられている。なぜなら上記細胞のゲノムには混入テンプレート D N A の組込み ( integration ) の証拠がないからである。

本発明は本願で説明した特定の実施例により範囲を限定されるものではない。実際に、本願で説明した例に加えて、本発明のさまざまな変形例は、前述の説明と添付図面から当業者には明らかである。かかる変形例は、添付した特許請求の範囲内にある。

核酸若しくはポリペプチドに対して与えられた全ての塩基サイズ若しくはアミノ酸サイズ、及び全ての分子量又は分子質量値はおおよそであり、説明のために提供された。

さまざまな刊行物が本願では引用され、それらの開示は全体の引用文献に組込まれる。

#### 配列リスト

##### ( 1 ) 一般情報 :

( i ) 出願人 : Rice, Charles ら

( ii ) 発明の名称 : C 型肝炎ウイルス ( H C V ) 用の機能性 D N A クローン及びその使用

( iii ) 配列数 : 21

##### ( iv ) 通信住所 :

( A ) 受信人 : David A. Jackson, Esq.

( B ) ストリート : 411 Hackensack Ave, Continental Plaza, 4th Floor

( C ) 市 : Hackensack

( D ) 州 : New Jersey

( E ) 国 : USA

( F ) 郵便番号 : 07601

##### ( V ) コンピュータ読取可能形態 :

( A ) 媒体タイプ : フロッピーディスク

( B ) コンピュータ : IBM PC 互換機

( C ) オペレーティングシステム : PC-DOS/MS-DOS

( D ) ソフトウェア : パテントインリリース 1.0, バージョン #1.30

##### ( vi ) 現在の出願データ :

( A ) 出願番号 : US

( B ) 出願日 : 1997年3月3日

##### ( C ) 分類 :

##### ( viii ) 弁護士 / エージェント情報 :

( A ) 名前 : Jackson Esq., David A.

( B ) 登録番号 : 26,742

( C ) 参照 / 整理番号 : 1113-1006

##### ( ix ) 通信情報 :

( A ) 電話 : 201-487-5800

( B ) ファックス : 201-343-1684

##### ( 2 ) SEQ ID NO: 1 の情報 :

##### ( i ) 配列特徴 :

( A ) 長さ : 9646 ベースペア

( B ) タイプ : 核酸

( C ) スtrand ネス : 二重鎖

( D ) トポロジー : 直線

( ii ) 分子タイプ : cDNA

( iii ) 仮説 : なし

( iv ) アンチセンス : なし

( xi ) 配列記述 : SEQ ID NO: 1 :

10

20

30

40

50



GCCAGCCCCC TGATGGGGGC GACACTCCAC CATGAATCAC TCCCCTGTGA GGAACACTG	60
TCTTCACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG TGTCGTGCAG CCTCCAGGAC	120
CCCCCTCCC GGGAGAGCCA TAGTGGTCTG CGGAACCGGT GAGTACACCG GAATTGCCAG	180
GACGACCGGG TCCTTTCTTG GATAAACCCG CTCAATGCCT GGAGATTTGG GCGTGCCCCC	240
GCAAGACTGC TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG	300
GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGCAC CATGAGCACG AATCCTAAAC	360
CTCAAAGAAA AACCAAACGT AACACCAACC GTCGCCCACA GGACGTCAAG TTCCCGGGTG	420
GCGGTCAGAT CGTTGGTGGA GTTTACTTGT TGCCGCGCAG GGGCCCTAGA TTGGGTGTGC	480
GCGCGACGAG GAAGACTTCC GAGCGGTCGC AACCTCGAGG TAGACGTCAG CCTATCCCCA	540
AGGCACGTCG GCCCGAGGGC AGGACCTGGG CTCAGCCCGG GTACCCTTGG CCCCTCTATG	600
GCAATGAGGG TTGCGGGTGG GCGGGATGGC TCCTGTCTCC CCGTGGCTCT CGGCCTAGCT	660
GGGGCCCCAC AGACCCCCGG CGTAGGTCGC GCAATTTGGG TAAGGTCATC GATACCCCTTA	720
CGTGCGGCTT CGCCGACCTC ATGGGGTACA TACCGCTCGT CGGCGCCCCT CTTGGAGGCG	780
CTGCCAGGGC CCTGGCGCAT GGCGTCCGGG TTCTGGAAGA CGGCGTGAAC TATGCAACAG	840
GGAACCTTCC TGGTTGCTCT TTCTCTATCT TCCTTCTGGC CCTGCTCTCT TGCCTGACTG	900
TGCCCCGCTT AGCCTACCAA GTGCGCAATT CCTCGGGGCT TTACCATGTC ACCAATGATT	960
GCCCTAACTC GAGTATTGTG TACGAGGCGG CCGATGCCAT CCTGCACACT CCGGGGTGTG	1020
TCCCTTGCGT TCGCGAGGGT AACGCCTCGA GGTGTTGGGT GGCGGTGACC CCCACGGTGG	1080
CCACCAGGGA CGGCAAACTC CCCACAACGC AGCTTCGACG TCATATCGAT CTGCTTGTCG	1140
GGAGCGCCAC CCTCTGCTCG GCCCTCTACG TGGGGGACCT GTGCGGGTCT GTCTTTCTTG	1200

10

20

30

TTGGTCAACT	GTTTACCTTC	TCTCCCAGGC	GCCACTGGAC	GACGCAAGAC	TGCAATTGTT	1260
CTATCTATCC	CGGCCATATA	ACGGGTCATC	GCATGGCATG	GGATATGATG	ATGAACTGGT	1320
CCCCACGGC	AGCGTTGGTG	GTAGCTCAGC	TGCTCCGGAT	CCCACAAGCC	ATCATGGACA	1380
TGATCGCTGG	TGCTCACTGG	GGAGTCCTGG	CGGGCATAGC	GTATTTCTTC	ATGGTGGGGA	1440
ACTGGGCGAA	GGTCCTGGTA	GTGCTGCTGC	TATTTGCCGG	CGTCGACGCG	GAAACCCACG	1500
TCACCGGGGG	AAGTGCCGGC	CGCACCACGG	CTGGGCTTGT	TGGTCTCCTT	ACACCAGGCG	1560
CCAAGCAGAA	CATCCAACCTG	ATCAACACCA	ACGGCAGTTG	GCACATCAAT	AGCACGGCCT	1620
TGAACTGCAA	TGAAAGCCTT	AACACCGGCT	GGTTAGCAGG	GCTCTTCTAT	CAGCACAAAT	1680
TCAACTCTTC	AGGCTGTCCT	GAGAGGTTGG	CCAGCTGCCG	ACGCCTTACC	GATTTTGCCC	1740
AGGGCTGGGG	TCCTATCAGT	TATGCCAACG	GAAGCGGCCT	CGACGAACGC	CCCTACTGCT	1800
GGCACTACCC	TCCAAGACCT	TGTGGCATTG	TGCCCCGAAA	GAGCGTGTGT	GGCCCCGTAT	1860
ATTGCTTCAC	TCCCAGCCCC	GTGGTGGTGG	GAACGACCGA	CAGGTCGGGC	GCGCCTACCT	1920
ACAGCTGGGG	TGCAAATGAT	ACGGATGTCT	TCGTCCTTAA	CAACACCAGG	CCACCGCTGG	1980
GCAATTGGTT	CGGTTGTACC	TGGATGAACT	CAACTGGATT	CACCAAAGTG	TGCGGAGCGC	2040
CCCCTTGTGT	CATCGGAGGG	GTGGGCAACA	ACACCTTGCT	CTGCCCCACT	GATTGTTTCC	2100
GCAAGCATCC	GGAAGCCACA	TACTCTCGGT	GCGGCTCCGG	TCCCTGGATT	ACACCCAGGT	2160
GCATGGTCGA	CTACCCGTAT	AGGCTTTGGC	ACTATCCTTG	TACCATCAAT	TACACCATAT	2220
TCAAAGTCAG	GATGTACGTG	GGAGGGGTCTG	AGCACAGGCT	GGAAGCGGCC	TGCAACTGGA	2280
CGCGGGGCGA	ACGCTGTGAT	CTGGAAGACA	GGGACAGGTC	CGAGCTCAGC	CCATTGCTGC	2340
TGTCCACCAC	ACAGTGGCAG	GTCCTTCCGT	GTTCTTTTAC	GACCCTGCCA	GCCTTGTTCA	2400
CCGGCCTCAT	CCACCTCCAC	CAGAACATTG	TGGACGTGCA	GTACTTGTTAC	GGGGTAGGGT	2460
CAAGCATCGC	GTCCTGGGCC	ATTAAGTGGG	AGTACGTCGT	TCTCCTGTTT	CTCCTGCTTG	2520
CAGACGCGCG	CGTCTGCTCC	TGCTTGTGGA	TGATGTTACT	CATATCCCAA	GCGGAGGCGG	2580
CTTTGGAGAA	CCTCGTAATA	CTCAATGCAG	CATCCCTGGC	CGGGACGCAC	GGTCTTGTGT	2640
CCTTCCTCGT	GTTCTTCTGC	TTTGCGTGGT	ATCTGAAGGG	TAGGTGGGTG	CCCGGAGCGG	2700
TCTACGCCTT	CTACGGGATG	TGGCCTCTCC	TCCTGCTCCT	GCTGGCGTTG	CCTCAGCGGG	2760

10

20

30

40

CATACGCACT	GGACACGGAG	GTGGCCGCGT	CGTGTGGCGG	CGTTGTTCTT	GTCGGGTAA	2820
TGGCGCTGAC	TCTGTCGCCA	TATTACAAGC	GCTACATCAG	CTGGTGCATG	TGGTGGCTTC	2880
AGTATTTTCT	GACCAGAGTA	GAAGCGCAAC	TGCACGTGTG	GGTTCCCCC	CTCAACGTCC	2940
GGGGGGGGCG	CGATGCCGTC	ATCTTACTCA	TGTGTGTTGT	ACACCCGACT	CTGGTATTTG	3000
ACATCACCAA	ACTACTCCTG	GCCATCTTCG	GACCCCTTTG	GATTCTTCAA	GCCAGTTTGC	3060
TTAAAGTCCC	CTACTTCGTG	CGCGTTCAAG	GCCTTCTCCG	GATCTGCGCG	CTAGCGCGGA	3120
AGATAGCCGG	AGGTCATTAC	GTGCAAATGG	CCATCATCAA	GTTAGGGGCG	CTTACTGGCA	3180
CCTATGTGTA	TAACCATCTC	ACCCCTCTTC	GAGACTGGGC	GCACAACGGC	CTGCGAGATC	3240
TGGCCGTGGC	TGTGGAACCA	GTCGTCTTCT	CCCGAATGGA	GACCAAGCTC	ATCACGTGGG	3300
GGGCAGATAC	CGCCGCGTGC	GGTGACATCA	TCAACGGCTT	GCCCCGTCTCT	GCCCGTAGGG	3360
GCCAGGAGAT	ACTGCTTGGG	CCAGCCGACG	GAATGGTCTC	CAAGGGGTGG	AGGTTGCTGG	3420
CGCCCATCAC	GGCGTACGCC	CAGCAGACGA	GAGGCCTCCT	AGGGTGTATA	ATCACCAGCC	3480
TGACTGGCCG	GGACAAAAAC	CAAGTGGAGG	GTGAGGTCCA	GATCGTGTCA	ACTGCTACCC	3540
AAACCTTCCT	GGCAACGTGC	ATCAATGGGG	TATGCTGGAC	TGTCTACCAC	GGGGCCGGAA	3600
CGAGGACCAT	CGCATCACCC	AAGGGTCCTG	TCATCCAGAT	GTATACCAAT	GTGGACCAAG	3660
ACCTTGTGGG	CTGGCCCGCT	CCTCAAGGTT	CCCGCTCATT	GACACCCTGC	ACCTGCGGCT	3720
CCTCGGACCT	TTACCTGGTC	ACGAGGCACG	CCGATGTCAT	TCCCGTGCGC	CGGCGAGGTG	3780
ATAGCAGGGG	TAGCCTGCTT	TCGCCCCGGC	CCATTTCCTA	CTTGAAAGGC	TCCTCGGGGG	3840
GTCCGCTGTT	GTGCCCCGCG	GGACACGCCG	TGGGCCTATT	CAGGGCCGCG	GTGTGCACCC	3900
GTGGAGTGGC	TAAGGCGGTG	GACTTTATCC	CTGTGGAGAA	CCTAGAGACA	ACCATGAGAT	3960
CCCCGGTGTT	CACGGACAAC	TCCTCTCCAC	CAGCAGTGCC	CCAGAGCTTC	CAGGTGGCCC	4020
ACCTGCATGC	TCCCACCGGC	AGCGGTAAGA	GCACCAAGGT	CCCGGCTGCG	TACGCAGCCC	4080
AGGGCTACAA	GGTGTGGGTG	CTCAACCCCT	CTGTTGCTGC	AACGCTGGGC	TTTGGTGCTT	4140
ACATGTCCAA	GGCCCATGGG	GTGATCCTA	ATATCAGGAC	CGGGGTGAGA	ACAATTACCA	4200
CTGGCAGCCC	CATCACGTAC	TCCACCTACG	GCAAGTTCCT	TGCCGACGGC	GGGTGCTCAG	4260
GAGGTGCTTA	TGACATAATA	ATTTGTGACG	AGTGCCACTC	CACGGATGCC	ACATCCATCT	4320

10

20

30

40

TGGGCATCGG	CACTGTCTTT	GACCAAGCAG	AGACTGCGGG	GGCGAGACTG	GTTGTGCTCG	4380
CCACTGCTAC	CCCTCCGGGC	TCCGTCACCTG	TGTCCCATCC	TAACATCGAG	GAGGTTGCTC	4440
TGTCCACCAC	CGGAGAGATC	CCTTTTTTACG	GCAAGGCTAT	CCCCCTCGAG	GTGATCAAGG	4500
GGGGAAGACA	TCTCATCTTC	TGCCACTCAA	AGAAGAAGTG	CGACGAGCTC	GCCGCGAAGC	4560
TGGTCGCATT	GGGCATCAAT	GCCGTGGCCT	ACTACCGCGG	TCTTGACGTG	TCTGTCATCC	4620
CGACCAGCGG	CGATGTTGTC	GTCGTGTCGA	CCGATGCTCT	CATGACTGGC	TTTACCGGCG	4680
ACTTCGACTC	TGTGATAGAC	TGCAACACGT	GTGTCACTCA	GACAGTCGAT	TTTACCGGCG	4740
ACCTTACCTT	TACCATTGAG	ACAACCACGC	TCCCCCAGGA	TGCTGTCTCC	AGGACTCAAC	4800
GCCGGGGCAG	GACTGGCAGG	GGGAAGCCAG	GCATCTACAG	ATTTGTGGCA	CCGGGGGAGC	4860
GCCCCCTCCG	CATGTTTCGAC	TCGTCCGTCC	TCTGTGAGTG	CTATGACGCG	GGCTGTGCTT	4920
GGTATGAGCT	CACGCCCCGCC	GAGACTACAG	TTAGGCTACG	AGCGTACATG	AACACCCCGG	4980
GGCTTCCCGT	GTGCCAGGAC	CATCTTGAAT	TTTGGGAGGG	CGTCTTTACG	GGCCTCACTC	5040
ATATAGATGC	CCACTTTCTA	TCCCAGACAA	AGCAGAGTGG	GGAGAACTTT	CCTTACCTGG	5100
TAGCGTACCA	AGCCACCGTG	TGCGCTAGGG	CTCAAGCCCC	TCCCCCATCG	TGGGACCAGA	5160
TGTGGAAGTG	TTTGATCCGC	CTTAAACCCA	CCCTCCATGG	GCCAACACCC	CTGCTATACA	5220
GACTGGGCGC	TGTTCAGAAT	GAAGTCACCC	TGACGCACCC	AATCACCAAA	TACATCATGA	5280
CATGCATGTC	GGCCGACCTG	GAGGTCGTCA	CGAGCACCTG	GGTGCTCGTT	GGCGGCGTCC	5340
TGGCTGCTCT	GGCCGCGTAT	TGCCTGTCAA	CAGGCTGCGT	GGTCATAGTG	GGCAGGATTG	5400
TCTTGTCCGG	GAAGCCGGCA	ATTATACCTG	ACAGGGAGGT	TCTCTACCAG	GAGTTCGATG	5460
AGATGGAAGA	GTGCTCTCAG	CACTTACCGT	ACATCGAGCA	AGGGATGATG	CTCGCTGAGC	5520
AGTTCAAGCA	GAAGGCCCTC	GGCCTCCTGC	AGACCGCGTC	CCGCCAAGCA	GAGGTTATCA	5580
CCCCTGCTGT	CCAGACCAAC	TGGCAGAAAC	TCGAGGTCTT	CTGGGCGAAG	CACATGTGGA	5640
ATTTTCATCAG	TGGGATACAA	TACTTGGCGG	GCCTGTCAAC	GCTGCCTGGT	AACCCCGCCA	5700
TTGCTTCATT	GATGGCTTTT	ACAGCTGCCG	TCACCAGCCC	ACTAACCCT	GGCCAAACCC	5760
TCCTCTTCAA	CATATTGGGG	GGGTGGGTGG	CTGCCAGCT	CGCCGCCCCC	GGTGCCGCTA	5820
CCGCCTTTGT	GGGCGCTGGC	TTAGCTGGCG	CCGCCATCGG	CAGCGTTGGA	CTGGGGAAGG	5880

10

20

30

40

TCCTCGTGGA	CATTCTTGCA	GGGTATGGCG	CGGGCGTGGC	GGGAGCTCTT	GTAGCATTCA	5940
AGATCATGAG	CGGTGAGGTC	CCCTCCACGG	AGGACCTGGT	CAATCTGCTG	CCCGCCATCC	6000
TCTCGCCTGG	AGCCCTTGTA	GTCGGTGTGG	TCTGCGCAGC	AATACTGCGC	CGGCACGTTG	6060
GCCCCGGCGA	GGGGGCAGTG	CAATGGATGA	ACCGGCTAAT	AGCCTTCGCC	TCCCGGGGGA	6120
ACCATGTTTC	CCCCACGCAC	TACGTGCCGG	AGAGCGATGC	AGCCGCCCCG	GTCAGTGCCA	6180
TACTCAGCAG	CCTCACTGTA	ACCCAGCTCC	TGAGGCGACT	GCATCAGTGG	ATAAGCTCGG	6240
AGTGTAACCAC	TCCATGCTCC	GGTTCCTGGC	TAAGGGACAT	CTGGGACTGG	ATATGCGAGG	6300
TGCTGAGCGA	CTTTAAGACC	TGGCTGAAAG	CCAAGCTCAT	GCCACAACCTG	CCTGGGATTC	6360
CCTTTGTGTC	CTGCCAGCGC	GGGTATAGGG	GGGTCTGGCG	AGGAGACGGC	ATTATGCACA	6420
CTCGCTGCCA	CTGTGGAGCT	GAGATCACTG	GACATGTCAA	AAACGGGACG	ATGAGGATCG	6480
TCGGTCCTAG	GACCTGCAGG	AACATGTGGA	GTGGGACGTT	CCCCATTAAAC	GCCTACACCA	6540
CGGGCCCCCTG	TACTCCCCCTT	CCTGCGCCGA	ACTATAAGTT	CGCGCTGTGG	AGGGTGTCTG	6600
CAGAGGAATA	CGTGGAGATA	AGGCGGGTGG	GGGACTTCCA	CTACGTATCG	GGTATGACTA	6660
CTGACAATCT	TAAATGCCCCG	TGCCAGATCC	CATCGCCCCG	ATTTTTCACA	GAATTGGACG	6720
GGGTGCGCCT	ACATAGGTTT	GCGCCCCCTT	GCAAGCCCTT	GCTGCGGGAG	GAGGTATCAT	6780
TCAGAGTAGG	ACTCCACGAG	TACCCGGTGG	GGTCGCAATT	ACCTTGCGAG	CCCGAACCGG	6840
ACGTAGCCGT	GTTGACGTCC	ATGCTCACTG	ATCCCTCCCA	TATAACAGCA	GAGGCGGCCG	6900
GGAGAAGGTT	GGCGAGAGGG	TCACCCCTT	CTATGGCCAG	CTCCTCGGCC	AGCCAGCTGT	6960
CCGCTCCATC	TCTCAAGGCA	ACTTGCACCG	CCAACCATGA	CTCCCCTGAC	GCCGAGCTCA	7020
TAGAGGCTAA	CCTCCTGTGG	AGGCAGGAGA	TGGGCGGCAA	CATCACCAGG	GTTGAGTCAG	7080
AGAACAAAGT	GGTGATTCTG	GACTCCTTCG	ATCCGCTTGT	GGCAGAGGAG	GATGAGCGGG	7140
AGGTCTCCGT	ACCCGCAGAA	ATTCTGCGGA	AGTCTCGGAG	ATTGCCCCGG	GCCCTGCCCCG	7200
TTTGGGCGCG	GCCGGACTAC	AACCCCCCGC	TAGTAGAGAC	GTGGAAAAG	CCTGACTACG	7260
AACCACCTGT	GGTCCATGGC	TGCCCCTAC	CACCTCCACG	GTCCCCTCCT	GTGCCTCCGC	7320
CTCGGAAAAA	GCGTACGGTG	GTCCTCACCG	AATCAACCCT	ATCTACTGCC	TTGGCCGAGC	7380
TTGCCACCAA	AAGTTTTGGC	AGCTCCTCAA	CTTCCGGCAT	TACGGGCGAC	AATACGACAA	7440

10

20

30

40

CATCCTCTGA	CCCCGCCCCCT	TCTGGCTGCC	CCCCCGACTC	CGACGTTGAG	TCCTATTCTT	7500
CCATGCCCCC	CCTGGAGGGG	GAGCCTGGGG	ATCCGATCT	CAGCGACGGG	TCATGGTCGA	7560
CGGTCAGTAG	TGGGGCCGAC	ACGGAAGATG	TCGTGTGCTG	CTCAATGTCT	TATTCCTGGA	7620
CAGGCGCACT	CGTCACCCCG	TGCGCTGCGG	AAGAACAAAA	ACTGCCCAIC	AACGCACTGA	7680
GCAACTCGTT	GCTACGCCAT	CACAATCTGG	TGTATTCCAC	CACTTCACGC	AGTGCTTGCC	7740
AAAGGCAGAA	GAAAGTCACA	TTTGACAGAC	TGCAAGTTCT	GGACAGCCAT	TACCAGGACG	7800
TGCTCAAGGA	GGTCAAAGCA	GCGGCGTCAA	AAGTGAAGGC	TAAGTTGCTA	TCCGTAGAGG	7860
AAGCTTGCA	CCTGACGCCC	CCACATTCAG	CCAAATCCAA	GTTTGGCTAT	GGGGCAAAAG	7920
ACGTCCGTTG	CCATGCCAGA	AAGGCCGTAG	CCCACATCAA	CTCCGTGTGG	AAAGACCTTC	7980
TGGAAGACAG	TGTAACACCA	ATAGACACTA	CCATCATGGC	CAAGAACGAG	GTTTTCTGCG	8040
TTCAGCCTGA	GAAGGGGGGT	CGTAAGCCAG	CTCGTCTCAT	CGTGTTCCCC	GACCTGGGCG	8100
TGCGCGTGTG	CGAGAAGATG	GCCCTGTACG	ACGTGGTTAG	CAAGCTCCCC	CTGGCCGTGA	8160
TGGGAAGCTC	CTACGGATTG	CAATACTCAC	CAGGACAGCG	GGTTGAATTC	CTCGTGCAAG	8220
CGTGGAAGTC	CAAGAAGACC	CCGATGGGGT	TCTCGTATGA	TACCCGCTGT	TTTGACTCCA	8280
CAGTCACTGA	GAGCGACATC	CGTACGGAGG	AGGCAATTTA	CCAATGTTGT	GACCTGGACC	8340
CCCAAGCCCC	CGTGCCCATC	AAGTCCCTCA	CTGAGAGGCT	TTATGTTGGG	GGCCCTCTTA	8400
CCAATTCAAG	GGGGGAAAAC	TGCGGCTACC	GCAGGTGCCG	CGCGAGCGGC	GTAAGTACAA	8460
CTAGCTGTGG	TAACACCCCTC	ACTTGCTACA	TCAAGGCCCC	GGCAGCCTGT	CGAGCCGACG	8520
GGCTCCAGGA	CTGCACCATG	CTCGTGTGTG	GCGACGACTT	AGTCGTTATC	TGTGAAAGTG	8580
CGGGGGTCCA	GGAGGACGCG	GCGAGCCTGA	GAGCCTTCAC	GGAGGCTATG	ACCAGGTACT	8640
CCGCCCCCCC	CGGGGACCCC	CCACAACCAG	AATACGACTT	GGAGCTTATA	ACATCATGCT	8700
CCTCCAACGT	GTCAGTCGCC	CACGACGGCG	CTGGAAGAG	GGTCTACTAC	CTTACCCGTG	8760
ACCCTACAAC	CCCCCTCGCG	AGAGCCGCGT	GGGAGACAGC	AAGACACACT	CCAGTCAATT	8820
CCTGGCTAGG	CAACATAATC	ATGTTTGCCC	CCACACTGTG	GGCGAGGATG	ATACTGATGA	8880
CCCATTCTTT	TAGCGTCCTC	ATAGCCAGGG	ATCAGCTTGA	ACAGGCTCTT	AACTGTGAGA	8940
TCTACGGAGC	CTGCTACTCC	ATAGAACCAC	TGGATCTACC	TCCAATCATT	CAAAGACTCC	9000

10

20

30

40

ATGGCCTCAG CGCATTTTCA CTCCACAGTT ACTCTCCAGG TGAAATCAAT AGGGTGGCCG 9060  
 CATGCCTCAG AAAAATTGGG GTCCCGCCCT TGCAGCTTG GAGACACCGG GCCCGGAGCG 9120  
 TCCGCGCTAG GCTTCTGTCC AGAGGAGGCA GGGCTGCCAT ATGTGGCAAG TACCTCTTCA 9180  
 ACTGGGCAGT AAGAACAAAG CTCAACTCA CTCCAATAGC GGCCGCTGGC CGGCTGGACT 9240  
 TGTCCGTTG GTTCACGGCT GGCTACAGCG GGGGAGACAT TTATCACAGC GTGTCTCATG 9300  
 CCCGGCCCCG CTGGTTCTGG TTTTGCCTAC TCCTGCTCGC TGCAGGGGTA GGCATCTACC 9360  
 TCCTCCCCAA CCGATGAAGG TTGGGGTAAA CACTCCGGCC TCTTAGGCCA TTTCCTGTTT 9420  
 TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTCT TTTTTTTTTT 9480  
 TTTTTTCCTT TTTTTTTTTT TTTTTTTTCT TTCCTTCTTT TTTCTTTTCT TTTCTTCTT 9540  
 TCTTTAATGG TGGCTCCATC TTAGCCCTAG TCACGGCTAG CTGTGAAAGG TCCGTGAGCC 9600  
 GCATGACTGC AGAGAGTGCT GATACTGGCC TCTCTGCAGA TCATGT 9646

10

( 2 ) SEQ ID NO: 2 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 3012 アミノ酸

20

( B ) タイプ : アミノ酸

( C ) スtrandネス : 一本鎖

( D ) トポロジー : 直線

( ii ) 分子タイプ : タンパク質

( iii ) 仮説 : なし

( v ) フラグメントタイプ : N末端

( xi ) 配列記述 : SEQ ID NO: 2 :

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
 1 5 10 15

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly  
 20 25 30

30

Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
 35 40 45

Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro

50	55	60	
Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly			
65	70	75	80
Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp			
	85	90	95
Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro			
	100	105	110
Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys			
	115	120	125
Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu			
	130	135	140
Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp			
	145	150	155
Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile			
	165	170	175
Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr			
	180	185	190
Gln Val Arg Asn Ser Ser Gly Leu Tyr His Val Thr Asn Asp Cys Pro			
	195	200	205
Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Ala Ile Leu His Thr Pro			
	210	215	220
Gly Cys Val Pro Cys Val Arg Glu Gly Asn Ala Ser Arg Cys Trp Val			
	225	230	235
Ala Val Thr Pro Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Thr Thr			
	245	250	255
Gln Leu Arg Arg His Ile Asp Leu Leu Val Gly Ser Ala Thr Leu Cys			
	260	265	270
Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val Gly			
	275	280	285
Gln Leu Phe Thr Phe Ser Pro Arg Arg His Trp Thr Thr Gln Asp Cys			
	290	295	300
Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala Trp			
	305	310	315
Asp Met Met Met Asn Trp Ser Pro Thr Ala Ala Leu Val Val Ala Gln			

10

20

30

40



[illegible]

10

20

30

40

595					600					605					
Val	Asp	Tyr	Pro	Tyr	Arg	Leu	Trp	His	Tyr	Pro	Cys	Thr	Ile	Asn	Tyr
610						615					620				
Thr	Ile	Phe	Lys	Val	Arg	Met	Tyr	Val	Gly	Gly	Val	Glu	His	Arg	Leu
625					630				635						640
Glu	Ala	Ala	Cys	Asn	Trp	Thr	Arg	Gly	Glu	Arg	Cys	Asp	Leu	Glu	Asp
				645					650					655	
Arg	Asp	Arg	Ser	Glu	Leu	Ser	Pro	Leu	Leu	Leu	Ser	Thr	Thr	Gln	Trp
			660					665					670		
Gln	Val	Leu	Pro	Cys	Ser	Phe	Thr	Thr	Leu	Pro	Ala	Leu	Ser	Thr	Gly
	675						680					685			
Leu	Ile	His	Leu	His	Gln	Asn	Ile	Val	Asp	Val	Gln	Tyr	Leu	Tyr	Gly
	690					695					700				
Val	Gly	Ser	Ser	Ile	Ala	Ser	Trp	Ala	Ile	Lys	Trp	Glu	Tyr	Val	Val
705					710					715					720
Leu	Leu	Phe	Leu	Leu	Leu	Ala	Asp	Ala	Arg	Val	Cys	Ser	Cys	Leu	Trp
					725				730					735	
Met	Met	Leu	Leu	Ile	Ser	Gln	Ala	Glu	Ala	Ala	Leu	Glu	Asn	Leu	Val
					740				745					750	
Ile	Leu	Asn	Ala	Ala	Ser	Leu	Ala	Gly	Thr	His	Gly	Leu	Val	Ser	Phe
		755					760					765			
Leu	Val	Phe	Phe	Cys	Phe	Ala	Trp	Tyr	Leu	Lys	Gly	Arg	Trp	Val	Pro
	770					775					780				
Gly	Ala	Val	Tyr	Ala	Phe	Tyr	Gly	Met	Trp	Pro	Leu	Leu	Leu	Leu	Leu
785					790					795					800
Leu	Ala	Leu	Pro	Gln	Arg	Ala	Tyr	Ala	Leu	Asp	Thr	Glu	Val	Ala	Ala
				805					810					815	
Ser	Cys	Gly	Gly	Val	Val	Leu	Val	Gly	Leu	Met	Ala	Leu	Thr	Leu	Ser
			820					825					830		
Pro	Tyr	Tyr	Lys	Arg	Tyr	Ile	Ser	Trp	Cys	Met	Trp	Trp	Leu	Gln	Tyr
	835						840						845		
Phe	Leu	Thr	Arg	Val	Glu	Ala	Gln	Leu	His	Val	Trp	Val	Pro	Pro	Leu
	850					855					860				
Asn	Val	Arg	Gly	Gly	Arg	Asp	Ala	Val	Ile	Leu	Leu	Met	Cys	Val	Val

10

20

30

40

865	870	875	880	
His Pro Thr Leu Val Phe Asp Ile Thr Lys Leu Leu Leu Ala Ile Phe	885	890	895	
Gly Pro Leu Trp Ile Leu Gln Ala Ser Leu Leu Lys Val Pro Tyr Phe	900	905	910	
Val Arg Val Gln Gly Leu Leu Arg Ile Cys Ala Leu Ala Arg Lys Ile	915	920	925	
Ala Gly Gly His Tyr Val Gln Met Ala Ile Ile Lys Leu Gly Ala Leu	930	935	940	10
Thr Gly Thr Tyr Val Tyr Asn His Leu Thr Pro Leu Arg Asp Trp Ala	945	950	955	960
His Asn Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe	965	970	975	
Ser Arg Met Glu Thr Lys Leu Ile Thr Trp Gly Ala Asp Thr Ala Ala	980	985	990	
Cys Gly Asp Ile Ile Asn Gly Leu Pro Val Ser Ala Arg Arg Gly Gln	995	1000	1005	20
Glu Ile Leu Leu Gly Pro Ala Asp Gly Met Val Ser Lys Gly Trp Arg	1010	1015	1020	
Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu	1025	1030	1035	1040
Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu	1045	1050	1055	
Gly Glu Val Gln Ile Val Ser Thr Ala Thr Gln Thr Phe Leu Ala Thr	1060	1065	1070	
Cys Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg	1075	1080	1085	30
Thr Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val	1090	1095	1100	
Asp Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg Ser Leu	1105	1110	1115	1120
Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His	1125	1130	1135	
Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu				40

1140	1145	1150
Leu Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro 1155	1160	1165
Leu Leu Cys Pro Ala Gly His Ala Val Gly Leu Phe Arg Ala Ala Val 1170	1175	1180
Cys Thr Arg Gly Val Ala Lys Ala Val Asp Phe Ile Pro Val Glu Asn 1185	1190	1195 1200
Leu Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro 1205	1210	1215
Pro Ala Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr 1220	1225	1230
Gly Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly 1235	1240	1245
Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe 1250	1255	1260
Gly Ala Tyr Met Ser Lys Ala His Gly Val Asp Pro Asn Ile Arg Thr 1265	1270	1275 1280
Gly Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr 1285	1290	1295
Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile 1300	1305	1310
Ile Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly 1315	1320	1325
Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val 1330	1335	1340
Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Ser His Pro 1345	1350	1355 1360
Asn Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr 1365	1370	1375
Gly Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile 1380	1385	1390
Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val 1395	1400	1405
Ala Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser		

10

20

30

40

1410	1415	1420	
Val Ile Pro Thr Ser Gly Asp Val Val Val Val Ser Thr Asp Ala Leu			
1425	1430	1435	1440
Met Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr			
	1445	1450	1455
Cys Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile			
	1460	1465	1470
Glu Thr Thr Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg			
	1475	1480	1485
Gly Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro			
	1490	1495	1500
Gly Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys			
	1505	1510	1515
Tyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu Thr Pro Ala Glu Thr Thr			
	1525	1530	1535
Val Arg Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln			
	1540	1545	1550
Asp His Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile			
	1555	1560	1565
Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Phe Pro			
	1570	1575	1580
Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro			
	1585	1590	1595
Pro Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro			
	1605	1610	1615
Thr Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln			
	1620	1625	1630
Asn Glu Val Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Thr Cys			
	1635	1640	1645
Met Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly			
	1650	1655	1660
Gly Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val			
	1665	1670	1675
Val Ile Val Gly Arg Ile Val Leu Ser Gly Lys Pro Ala Ile Ile Pro			

10

20

30

40

1685	1690	1695	
Asp Arg Glu Val Leu Tyr Gln Glu Phe Asp Glu Met Glu Glu Cys Ser			
1700	1705	1710	
Gln His Leu Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu Gln Phe			
1715	1720	1725	
Lys Gln Lys Ala Leu Gly Leu Leu Gln Thr Ala Ser Arg Gln Ala Glu			
1730	1735	1740	
Val Ile Thr Pro Ala Val Gln Thr Asn Trp Gln Lys Leu Glu Val Phe			
1745	1750	1755	10
		1760	
Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala			
1765	1770	1775	
Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala			
1780	1785	1790	
Phe Thr Ala Ala Val Thr Ser Pro Leu Thr Thr Gly Gln Thr Leu Leu			
1795	1800	1805	
Phe Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Ala Pro Gly			
1810	1815	1820	20
Ala Ala Thr Ala Phe Val Gly Ala Gly Leu Ala Gly Ala Ala Ile Gly			
1825	1830	1835	
		1840	
Ser Val Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly			
1845	1850	1855	
Ala Gly Val Ala Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu			
1860	1865	1870	
Val Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser			
1875	1880	1885	
Pro Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg			
1890	1895	1900	30
His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile			
1905	1910	1915	
		1920	
Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro			
1925	1930	1935	
Glu Ser Asp Ala Ala Ala Arg Val Thr Ala Ile Leu Ser Ser Leu Thr			
1940	1945	1950	
Val Thr Gln Leu Leu Arg Arg Leu His Gln Trp Ile Ser Ser Glu Cys			
			40

1955					1960					1965						
Thr	Thr	Pro	Cys	Ser	Gly	Ser	Trp	Leu	Arg	Asp	Ile	Trp	Asp	Trp	Ile	
1970					1975					1980						
Cys	Glu	Val	Leu	Ser	Asp	Phe	Lys	Thr	Trp	Leu	Lys	Ala	Lys	Leu	Met	
1985					1990					1995					2000	
Pro	Gln	Leu	Pro	Gly	Ile	Pro	Phe	Val	Ser	Cys	Gln	Arg	Gly	Tyr	Arg	
2005					2010					2015						
Gly	Val	Trp	Arg	Gly	Asp	Gly	Ile	Met	His	Thr	Arg	Cys	His	Cys	Gly	
2020					2025					2030						
Ala	Glu	Ile	Thr	Gly	His	Val	Lys	Asn	Gly	Thr	Met	Arg	Ile	Val	Gly	
2035					2040					2045						
Pro	Arg	Thr	Cys	Arg	Asn	Met	Trp	Ser	Gly	Thr	Phe	Pro	Ile	Asn	Ala	
2050					2055					2060						
Tyr	Thr	Thr	Gly	Pro	Cys	Thr	Pro	Leu	Pro	Ala	Pro	Asn	Tyr	Lys	Phe	
2065					2070					2075					2080	
Ala	Leu	Trp	Arg	Val	Ser	Ala	Glu	Glu	Tyr	Val	Glu	Ile	Arg	Arg	Val	
2085					2090					2095						
Gly	Asp	Phe	His	Tyr	Val	Ser	Gly	Met	Thr	Thr	Asp	Asn	Leu	Lys	Cys	
2100					2105					2110						
Pro	Cys	Gln	Ile	Pro	Ser	Pro	Glu	Phe	Phe	Thr	Glu	Leu	Asp	Gly	Val	
2115					2120					2125						
Arg	Leu	His	Arg	Phe	Ala	Pro	Pro	Cys	Lys	Pro	Leu	Leu	Arg	Glu	Glu	
2130					2135					2140						
Val	Ser	Phe	Arg	Val	Gly	Leu	His	Glu	Tyr	Pro	Val	Gly	Ser	Gln	Leu	
2145					2150					2155					2160	
Pro	Cys	Glu	Pro	Glu	Pro	Asp	Val	Ala	Val	Leu	Thr	Ser	Met	Leu	Thr	
2165					2170					2175						
Asp	Pro	Ser	His	Ile	Thr	Ala	Glu	Ala	Ala	Gly	Arg	Arg	Leu	Ala	Arg	
2180					2185					2190						
Gly	Ser	Pro	Pro	Ser	Met	Ala	Ser	Ser	Ser	Ala	Ser	Gln	Leu	Ser	Ala	
2195					2200					2205						
Pro	Ser	Leu	Lys	Ala	Thr	Cys	Thr	Ala	Asn	His	Asp	Ser	Pro	Asp	Ala	
2210					2215					2220						
Glu	Leu	Ile	Glu	Ala	Asn	Leu	Leu	Trp	Arg	Gln	Glu	Met	Gly	Gly	Asn	

10

20

30

40

2225	2230	2235	2240
Ile Thr Arg Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe			
2245	2250	2255	
Asp Pro Leu Val Ala Glu Glu Asp Glu Arg Glu Val Ser Val Pro Ala			
2260	2265	2270	
Glu Ile Leu Arg Lys Ser Arg Arg Phe Ala Arg Ala Leu Pro Val Trp			
2275	2280	2285	
Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro			
2290	2295	2300	
Asp Tyr Glu Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Arg			
2305	2310	2315	2320
Ser Pro Pro Val Pro Pro Pro Arg Lys Lys Arg Thr Val Val Leu Thr			
2325	2330	2335	
Glu Ser Thr Leu Ser Thr Ala Leu Ala Glu Leu Ala Thr Lys Ser Phe			
2340	2345	2350	
Gly Ser Ser Ser Thr Ser Gly Ile Thr Gly Asp Asn Thr Thr Thr Ser			
2355	2360	2365	
Ser Glu Pro Ala Pro Ser Gly Cys Pro Pro Asp Ser Asp Val Glu Ser			
2370	2375	2380	
Tyr Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu			
2385	2390	2395	2400
Ser Asp Gly Ser Trp Ser Thr Val Ser Ser Gly Ala Asp Thr Glu Asp			
2405	2410	2415	
Val Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Val Thr			
2420	2425	2430	
Pro Cys Ala Ala Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn			
2435	2440	2445	
Ser Leu Leu Arg His His Asn Leu Val Tyr Ser Thr Thr Ser Arg Ser			
2450	2455	2460	
Ala Cys Gln Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu			
2465	2470	2475	2480
Asp Ser His Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ala Ser			
2485	2490	2495	
Lys Val Lys Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr			

10

20

30

40



2500	2505	2510	
Pro Pro His Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val 2515 2520 2525			
Arg Cys His Ala Arg Lys Ala Val Ala His Ile Asn Ser Val Trp Lys 2530 2535 2540			
Asp Leu Leu Glu Asp Ser Val Thr Pro Ile Asp Thr Thr Ile Met Ala 2545 2550 2555 2560			
Lys Asn Glu Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro 2565 2570 2575			10
Ala Arg Leu Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys 2580 2585 2590			
Met Ala Leu Tyr Asp Val Val Ser Lys Leu Pro Leu Ala Val Met Gly 2595 2600 2605			
Ser Ser Tyr Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu 2610 2615 2620			
Val Gln Ala Trp Lys Ser Lys Lys Thr Pro Met Gly Phe Ser Tyr Asp 2625 2630 2635 2640			20
Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu 2645 2650 2655			
Glu Ala Ile Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala 2660 2665 2670			
Ile Lys Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn 2675 2680 2685			
Ser Arg Gly Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val 2690 2695 2700			
Leu Thr Thr Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg 2705 2710 2715 2720			30
Ala Ala Cys Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys 2725 2730 2735			
Gly Asp Asp Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp 2740 2745 2750			
Ala Ala Ser Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala 2755 2760 2765			
Pro Pro Gly Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr			40

2770                      2775                      2780  
 Ser Cys Ser Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg  
 2785                      2790                      2795                      2800  
 Val Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala  
                                  2805                      2810                      2815  
 Trp Glu Thr Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile  
                                  2820                      2825                      2830  
 Ile Met Phe Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His  
                                  2835                      2840                      2845  
 Phe Phe Ser Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asn  
                                  2850                      2855                      2860  
 Cys Glu Ile Tyr Gly Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro  
 2865                      2870                      2875                      2880  
 Pro Ile Ile Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His Ser  
                                  2885                      2890                      2895  
 Tyr Ser Pro Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg Lys Leu  
                                  2900                      2905                      2910  
 Gly Val Pro Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg  
                                  2915                      2920                      2925  
 Ala Arg Leu Leu Ser Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr  
                                  2930                      2935                      2940  
 Leu Phe Asn Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala  
 2945                      2950                      2955                      2960  
 Ala Ala Gly Arg Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser  
                                  2965                      2970                      2975  
 Gly Gly Asp Ile Tyr His Ser Val Ser His Ala Arg Pro Arg Trp Phe  
                                  2980                      2985                      2990  
 Trp Phe Cys Leu Leu Leu Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu  
                                  2995                      3000                      3005  
 Pro Asn Arg Glx  
 3010

10

20

30

( 2 ) SEQ ID NO: 3 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 38ベースペア

( B ) タイプ : 核酸

( C ) ストランドネス : 二重鎖

( D ) トポロジー : 直線

( ii ) 分子タイプ : DNA ( ゲノム )

( iii ) 仮説 : なし

( xi ) 配列記述 : SEQ ID NO: 3 :

GCCAGCCCCC TGATGGGGGC GACACTCCAC CATGAATC

38

( 2 ) SEQ ID NO: 4 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 101ベースペア

40

50

- (B) タイプ： 核酸
- (C) スtrandネス： 二重鎖
- (D) トポロジー： 直線
- (ii) 分子タイプ： DNA (ゲノム)
- (iii) 仮説： なし

(xi) 配列記述： SEQ ID NO: 4 :

```
GGTGGCT CCATCTTAGC CCTAGTCACG GCTAGCTGTG AAAGGTCCGT GAGCCGCATG 60
ACTGCAGAGA GTGCTGATAC TGGCCTCTCT GCTGATCATG T 101
```

( 2 ) SEQ ID NO: 5 の情報：

10

(i) 配列特徴：

- (A) 長さ： 12980 ペースペア
- (B) タイプ： 核酸
- (C) スtrandネス： 二重鎖
- (D) トポロジー： 直線
- (ii) 分子タイプ： cDNA
- (iii) 仮説： なし
- (iv) アンチセンス： なし
- (xi) 配列記述： SEQ ID NO: 5 :

GCCAGCCCCC	TGATGGGGGC	GACACTCCAC	CATGAATCAC	TCCCCTGTGA	GGAAC TACTG	60
TCTTCACGCA	GAAAGCGTCT	AGCCATGGCG	TTAGTATGAG	TGTCGTGCAG	CCTCCAGGAC	120
CCCCCTCCC	GGGAGAGCCA	TAGTGGTCTG	CGGAACCGGT	GAGTACACCG	GAATTGCCAG	180
GACGACCGGG	TCCTTTCTTG	GATAAACCCG	CTCAATGCCT	GGAGATTTGG	GCGTGCCCCC	240
GCAAGACTGC	TAGCCGAGTA	GTGTTGGGTC	GCGAAAGGCC	TTGTGGTACT	GCCTGATAGG	300
GTGCTTGCGA	GTGCCCCGGG	AGGTCTCGTA	GACCGTGCAC	CATGAGCACG	AATCCTAAAC	360
CTCAAAGAAA	AACCAAACGT	AACACCAACC	GTGCCCCACA	GGACGTCAAG	TTCCCGGGTG	420
GCGGTCAGAT	CGTTGGTGGA	GTTTACTTGT	TGCCGCGCAG	GGGCCCTAGA	TTGGGTGTGC	480
GCGCGACGAG	GAAGACTTCC	GAGCGGTCGC	AACCTCGAGG	TAGACGTCAG	CCTATCCCCA	540
AGGCACGTCG	GCCCCGAGGG	AGGACCTGGG	CTCAGCCCCG	GTACCCTTGG	CCCCCTATG	600
GCAATGAGGG	TTGCGGGTGG	GCGGGATGGC	TCCTGTCTCC	CCGTGGCTCT	CGGCCTAGCT	660
GGGGCCCCAC	AGACCCCCGG	CGTAGGTCGC	GCAATTGTTG	TAAGGTCATC	GATACCCTTA	720
CGTGCGGCTT	CGCCGACCTC	ATGGGGTACA	TACCGCTCGT	CGGCGCCCCT	CTTGAGGGCG	780
CTGCCAGGGC	CCTGGCGCAT	GGCGTCCGGG	TTCTGGAAGA	CGGCGTGAAC	TATGCAACAG	840
GGAACCTTCC	TGGTTGCTCT	TTCTCTATCT	TCCTTCTGGC	CCTGCTCTCT	TGCCTGACCG	900
TGCCCCGCTC	AGCCTACCAA	GTGCGCAATT	CCTCGGGGCT	TTACCATGTC	ACCAATGATT	960
GCCCTAACTC	GAGTATTGTG	TACGAGGCGG	CCGATGCCAT	CCTGCACACT	CCGGGGTGTG	1020
TCCCTTGCGT	TCGCGAGGGT	AACGCCTCGA	GGTGTGGGGT	GGCGGTGACC	CCCACGGTGG	1080
CCACCAGGGA	CGGCAAACCT	CCCACAACGC	AGCTTCGACG	TCATATCGAT	CTGCTTGTCG	1140
GGAGCGCCAC	CCTCTGCTCG	GCCCTCTACG	TGGGGGACCT	GTGCGGGTCT	GTCTTTCTTG	1200
TTGGTCAACT	GTTTACCTTC	TCTCCCAGGC	GCCACTGGAC	GACGCAAGAC	TGCAATTGTT	1260
CTATCTATCC	CGGCCATATA	ACGGGTCATC	GCATGGCATG	GGATATGATG	ATGAACTGGT	1320
CCCCTACGGC	AGCGTTGGTG	GTAGCTCAGC	TGCTCCGGAT	CCCACAAGCC	ATCATGGACA	1380
TGATCGCTGG	TGCTCACTGG	GGAGTCCTGG	CGGGCATAGC	GTATTTCTCC	ATGGTGGGGA	1440

10

20

30

ACTGGGCGAA GGTCTGGTA GTGCTGCTGC TATTTGCCGG CGTCGACGCG GAAACCCACG	1500
TCACCGGGGG AAGTGCCGGC CGCACCACGG CTGGGCTTGT TGGTCTCCTT ACACCAGGCG	1560
CCAAGCAGAA CATCCAACG ATCAACACCA ACGGCAGTTG GCACATCAAT AGCACGGCCT	1620
TGAACTGCAA TGAAAGCCTT AACACCGGCT GGTTAGCAGG GCTCTTCTAT CAGCACAAAT	1680
TCAACTCTTC AGGCTGTCCT GAGAGGTTGG CCAGCTGCCG ACGCCTTACC GATTTTGCCC	1740
AGGGCTGGGG TCCTATCAGT TATGCCAACG GAAGCGGCCT CGACGAACGC CCCTACTGCT	1800
GGCACTACCC TCCAAGACCT TGTGGCATTG TGCCCGCAAA GAGCGTGTGT GGCCCGGTAT	1860
ATTGCTTCAC TCCCAGCCCC GTGGTGGTGG GAACGACCGA CAGGTCGGGC GCGCCTACCT	1920
ACAGCTGGGG TGCAAATGAT ACGGATGTCT TCGTCCTTAA CAACACCAGG CCACCGCTGG	1980
GCAATTGGTT CGGTTGTACC TGGATGAACT CAACTGGATT CACCAAAGTG TGCGGAGCGC	2040
CCCCTTGTGT CATCGGAGGG GTGGGCAACA ACACCTTGCT CTGCCCCACT GATTGTTTCC	2100
GCAAGCATCC GGAAGCCACA TACTCTCGGT GCGGCTCCGG TCCCTGGATT ACACCCAGGT	2160
GCATGGTCGA CTACCCGTAT AGGCTTTGGC ACTATCCTTG TACCATCAAT TACACCATAT	2220
TCAAAGTCAG GATGTACGTG GGAGGGGTCG AGCACAGGCT GGAAGCGGCC TGCAACTGGA	2280
CGCGGGGCGA ACGCTGTGAT CTGGAAGACA GGGACAGGTC CGAGCTCAGC CCATTGCTGC	2340
TGTCCACCAC ACAGTGGCAG GTCCTTCCGT GTTCTTTTAC GACCCTGCCA GCCTTGTTCA	2400
CCGGCCTCAT CCACCTCCAC CAGAACATTG TGGACGTGCA GTACTTGTA GGGGTAGGGT	2460
CAAGCATCGC GTCCTGGGCC ATTAAGTGGG AGTACGTCGT TCTCCTGTTC CTCCTGCTTG	2520
CAGACGCGCG CGTCTGCTCC TGCTTGTGGA TGATGTTACT CATATCCCAA GCGGAGGCGG	2580
CTTTGGAGAA CCTCGTAATA CTCAATGCAG CATCCCTGGC CGGGACGCAC GGTCTTGTGT	2640
CCTTCCTCGT GTTCTTCTGC TTTGCGTGGT ATCTGAAGGG TAGGTGGGTG CCCGGAGCGG	2700
TCTACGCCTT CTACGGGATG TGGCCTCTCC TCCTGCTCCT GCTGGCGTTG CCTCAGCGGG	2760
CATACGCACT GGACACGGAG GTGGCCGCGT CGTGTGGCGG CGTTGTTCTT GTCGGGTTAA	2820
TGGCGCTGAC TCTGTCGCCA TATTACAAGC GCTACATCAG CTGGTGCATG TGGTGGCTTC	2880
AGTATTTTCT GACCAGAGTA GAAGCGCAAC TGCACGTGTG GGTTCCTCCC CTCAACGTCC	2940
GGGGGGGGCG CGATGCCGTC ATCTTACTCA TGTGTGTTGT ACACCCGACT CTGGTATTTG	3000

10

20

30

40

ACATCACCAA	ACTACTCCTG	GCCATCTTCG	GACCCCTTTG	GATTCTTCAA	GCCAGTTTGC	3060
TTAAAGTCCC	CTACTTCGTG	CGCGTTCAAG	GCCTTCTCCG	GATCTGCGCG	CTAGCGCGGA	3120
AGATAGCCCG	AGGTCATTAC	GTGCAAATGG	CCATCATCAA	GTTAGGGGCG	CTTACTGGCA	3180
CCTATGTGTA	TAACCATCTC	ACCCCTCTTC	GAGACTGGGC	GCACAAAGGC	CTGCGAGATC	3240
TGGCCGTGGC	TGTGGAACCA	GTCGTCTTCT	CCCGAATGGA	GACCAAGCTC	ATCACGTGGG	3300
GGGCAGATAC	CGCCGCGTGC	GGTGACATCA	TCAACGGCTT	GCCCGTCTCT	GCCCGTAGGG	3360
GCCAGGAGAT	ACTGCTTGGG	CCAGCCGACG	GAATGGTCTC	CAAGGGGTGG	AGGTTGCTGG	3420
CGCCCATCAC	GGCGTACGCC	CAGCAGACGA	GAGGCCTCCT	AGGGTGTATA	ATCACCAGCC	3480
TGACTGGCCG	GGACAAAAAC	CAAGTGGAGG	GTGAGGTCCA	GATCGTGTCA	ACTGCTACCC	3540
AAACCTTCCT	GGCAACGTGC	ATCAATGGGG	TATGCTGGAC	TGTCTACCAC	GGGGCCGGAA	3600
CGAGGACCAT	CGCATCACCC	AAGGGTCCTG	TCATCCAGAT	GTATACCAAT	GTGGACCAAG	3660
ACCTTGTGGG	CTGGCCCGCT	CCTCAAGGTT	CCCGCTCATT	GACACCCTGC	ACCTGCGGCT	3720
CCTCGGACCT	TTACCTGGTC	ACGAGGCACG	CCGATGTCAT	TCCCGTGCGC	CGGCGAGGTG	3780
ATAGCAGGGG	TAGCCTGCTT	TCGCCCCGGC	CCATTTCCCTA	CTTGAAAGGC	TCCTCGGGGG	3840
GTCCGCTGTT	GTGCCCCGCG	GGACACGCCG	TGGGCCTATT	CAGGGCCGCG	GTGTGCACCC	3900
GTGGAGTGGC	TAAGGCGGTG	GACTTTATCC	CTGTGGAGAA	CCTAGAGACA	ACCATGAGAT	3960
CCCCGGTGTT	CACGGACAAC	TCCTCTCCAC	CAGCAGTGCC	CCAGAGCTTC	CAGGTGGCCC	4020
ACCTGCATGC	TCCCACCGGC	AGCGGTAAGA	GCACCAAGGT	CCCGGCTGCG	TACGCAGCCC	4080
AGGGCTACAA	GGTGTGGTG	CTCAACCCCT	CTGTTGCTGC	AACGCTGGGC	TTTGGTGCTT	4140
ACATGTCCAA	GGCCCATGGG	GTTGATCCTA	ATATCAGGAC	CGGGGTGAGA	ACAATTACCA	4200
CTGGCAGCCC	CATCACGTAC	TCCACCTACG	GCAAGTTCCT	TGCCGACGGC	GGGTGCTCAG	4260
GAGGTGCTTA	TGACATAATA	ATTTGTGACG	AGTGCCACTC	CACGGATGCC	ACATCCATCT	4320
TGGGCATCGG	CACTGTCTTT	GACCAAGCAG	AGACTGCGGG	GGCGAGACTG	GTTGTGCTCG	4380
CCACTGCTAC	CCCTCCGGGC	TCCGTCACCTG	TGTCCCATCC	TAACATCGAG	GAGGTTGCTC	4440
TGTCCACCAC	CGGAGAGATC	CCCTTTTACG	GCAAGGCTAT	CCCCCTCGAG	GTGATCAAGG	4500
GGGGAAGACA	TCTCATCTTC	TGCCACTCAA	AGAAGAAGTG	CGACGAGCTC	GCCGCGAAGC	4560

10

20

30

40

TGGTCGCATT	GGGCATCAAT	GCCGTGGCCT	ACTACCGCGG	TCTTGACGTG	TCTGTCATCC	4620
CGACCAGCGG	CGATGTTGTC	GTCGTGTCGA	CCGATGCTCT	CATGACTGGC	TTTACCGGCG	4680
ACTTCGACTC	TGTGATAGAC	TGCAACACGT	GTGTCACTCA	GACAGTCGAT	TTCAGCCTTG	4740
ACCCCTACCTT	TACCATTGAG	ACAACCACGC	TCCCCCAGGA	TGCTGTCTCC	AGGACTCAAC	4800
GCCGGGGCAG	GACTGGCAGG	GGGAAGCCAG	GCATCTACAG	ATTTGTGGCA	CCGGGGGAGC	4860
GCCCCCTCCGG	CATGTTGAC	TCGTCCGTCC	TCTGTGAGTG	CTATGACGCG	GGCTGTGCTT	4920
GGTATGAGCT	CACGCCCCGCC	GAGACTACAG	TTAGGCTACG	AGCGTACATG	AACACCCCGG	4980
GGCTTCCCGT	GTGCCAGGAC	CATCTTGAAT	TTTGGGAGGG	CGTCTTTACG	GGCCTCACTC	5040
ATATAGATGC	CCACTTTCTA	TCCCAGACAA	AGCAGAGTGG	GGAGAACTTT	CCTTACCTGG	5100
TAGCGTACCA	AGCCACCGTG	TGCGCTAGGG	CTCAAGCCCC	TCCCCCATCG	TGGGACCAGA	5160
TGTGGAAGTG	TTTGATCCGC	CTTAAACCCA	CCCTCCATGG	GCCAACACCC	CTGCTATACA	5220
GACTGGGCGC	TGTTCAGAAT	GAAGTCACCC	TGACGCACCC	AATCACCAAA	TACATCATGA	5280
CATGCATGTC	GGCCGACCTG	GAGGTCGTCA	CGAGCACCTG	GGTGCTCGTT	GGCGGCGTCC	5340
TGGCTGCTCT	GGCCGCGTAT	TGCCTGTCAA	CAGGCTGCGT	GGTCATAGTG	GGCAGGATTG	5400
TCTTGTCGGG	GAAGCCGGCA	ATTATACCTG	ACAGGGAGGT	TCTCTACCAG	GAGTTCGATG	5460
AGATGGAAGA	GTGCTCTCAG	CACTTACCGT	ACATCGAGCA	AGGGATGATG	CTCGCTGAGC	5520
AGTTCAAGCA	GAAGGCCCTC	GGCCTCCTGC	AGACCGCGTC	CCGCCAAGCA	GAGGTTATCA	5580
CCCCTGCTGT	CCAGACCAAC	TGGCAGAAAC	TCGAGGTCTT	CTGGGCGAAG	CACATGTGGA	5640
ATTTTCATCAG	TGGGATACAA	TACTTGCGCG	GCCTGTCAAC	GCTGCCTGGT	AACCCCGCCA	5700
TTGCTTCATT	GATGGCTTTT	ACAGCTGCCG	TCACCAGCCC	ACTAACCCT	GGCCAAACCC	5760
TCCTCTTCAA	CATATTGGGG	GGGTGGGTGG	CTGCCCAGCT	CGCCGCCCCC	GGTGCCGCTA	5820
CCGCCTTTGT	GGGCGCTGGC	TTAGCTGGCG	CCGCCATCGG	CAGCGTTGGA	CTGGGGAAGG	5880
TCCTCGTGGA	CATTCTTGCA	GGGTATGGCG	CGGGCGTGGC	GGGAGCTCTT	GTAGCCTTCA	5940
AGATCATGAG	CGGTGAGGTC	CCCTCCACGG	AGGACCTGGT	CAATCTGCTG	CCCGCCATCC	6000
TCTCGCCTGG	AGCCCTTGTA	GTCGGTGTGG	TCTGCGCAGC	AATACTGCGC	CGGCACGTTG	6060
GCCCGGGCGA	GGGGGCAGTG	CAATGGATGA	ACCGGCTAAT	AGCCTTCGCC	TCCCGGGGGA	6120

10

20

30

40

ACCATGTTTC	CCCCACGCAC	TACGTGCCGG	AGAGCGATGC	AGCCGCCCCG	GTCAGTGCCA	6180
TACTCAGCAG	CCTCACTGTA	ACCCAGCTCC	TGAGGCGACT	GCATCAGTGG	ATAAGCTCGG	6240
AGTGTACCAC	TCCATGCTCC	GGTTCCTGGC	TAAGGGACAT	CTGGGACTGG	ATATGCGAGG	6300
TGCTGAGCGA	CTTTAAGACC	TGGCTGAAAG	CCAAGCTCAT	GCCACAACCTG	CCTGGGATTC	6360
CCTTTGTGTC	CTGCCAGCGC	GGGTATAGGG	GGGTCTGGCG	AGGAGACGGC	ATTATGCACA	6420
CTCGCTGCCA	CTGTGGAGCT	GAGATCACTG	GACATGTCAA	AAACGGGACG	ATGAGGATCG	6480
TCGGTCCTAG	GACCTGCAGG	AACATGTGGA	GTGGGACGTT	CCCCATTAAC	GCCTACACCA	6540
CGGGCCCCCTG	TACTCCCCCTT	CCTGCGCCGA	ACTATAAGTT	CGCGCTGTGG	AGGGTGTCTG	6600
CAGAGGAATA	CGTGAGAGATA	AGGCGGGTGG	GGGACTTCCA	CTACGTATCG	GGTATGACTA	6660
CTGACAATCT	TAAATGCCCCG	TGCCAGATCC	CATCGCCCGA	ATTTTTCACA	GAATTGGACG	6720
GGGTGCGCCT	ACATAGGTTT	GCGCCCCCTT	GCAAGCCCTT	GCTGCGGGAG	GAGGTATCAT	6780
TCAGAGTAGG	ACTCCACGAG	TACCCGGTGG	GGTCGCAATT	ACCTTGCGAG	CCCGAACCGG	6840
ACGTAGCCGT	GTGACGTCC	ATGCTCACTG	ATCCCTCCCA	TATAACAGCA	GAGGCGGCCG	6900
GGAGAAGGTT	GGCGAGAGGG	TCACCCCTT	CTATGGCCAG	CTCCTCGGCC	AGCCAGCTGT	6960
CCGCTCCATC	TCTCAAGGCA	ACTTGACCG	CCAACCATGA	CTCCCTGAC	GCCGAGCTCA	7020
TAGAGGCTAA	CCTCCTGTGG	AGGCAGGAGA	TGGGCGGCAA	CATCACCAGG	GTGAGTCAG	7080
AGAACAAAGT	GGTGATTCTG	GACTCCTTCG	ATCCGCTTGT	GGCAGAGGAG	GATGAGCGGG	7140
AGGTCTCCGT	ACCCGCAGAA	ATTCTGCGGA	AGTCTCGGAG	ATTGCCCCGG	GCCCTGCCCCG	7200
TTTGGGCGCG	GCCGGAACCT	AACCCCCCGC	TAGTAGAGAC	GTGGAAAAAG	CCTGACTACG	7260
AACCACCTGT	GGTCCATGGC	TGCCCCGTAC	CACCTCCACG	GTCCCTCCT	GTGCCTCCGC	7320
CTCGGAAAAA	GCGTACGGTG	GTCCTCACCG	AATCAACCCT	ATCTACTGCC	TTGGCCGAGC	7380
TTGCCACCAA	AAGTTTTTGGC	AGCTCCTCAA	CTTCCGGCAT	TACGGGCGAC	AATACGACAA	7440
CATCCTCTGA	GCCCCGCCCT	TCTGGCTGCC	CCCCCGACTC	CGACGTTGAG	TCCTATTCTT	7500
CCATGCCCCC	CCTGGAGGGG	GAGCCTGGGG	ATCCGGATCT	CAGCGACGGG	TCATGGTCGA	7560
CGGTCAGTAG	TGGGGCCGAC	ACGGAAGATG	TCGTGTGCTG	CTCAATGTCT	TATTCCTGGA	7620
CAGGCGCACT	CGTCACCCCG	TGCGCTGCGG	AAGAACAAAA	ACTGCCCATC	AACGCACTGA	7680

10

20

30

40



GCAACTCGTT	GCTACGCCAT	CACAATCTGG	TGTATTCCAC	CACTTCACGC	AGTGCTTGCC	7740
AAAGGCAGAA	GAAAGTCACA	TTTGACAGAC	TGCAAGTTCT	GGACAGCCAT	TACCAGGACG	7800
TGCTCAAGGA	GGTCAAAGCA	GCGGCGTCAA	AAGTGAAGGC	TAAGTTGCTA	TCCGTAGAGG	7860
AAGCTTGCCAG	CCTGACGCCC	CCACATTCCAG	CCAAATCCAA	GTTTGGCTAT	GGGGCAAAAG	7920
ACGTCCGTTG	CCATGCCAGA	AAGGCCGTAG	CCCACATCAA	CTCCGTGTGG	AAAGACCTTC	7980
TGGAAGACAG	TGTAACACCA	ATAGACACTA	CCATCATGGC	CAAGAACGAG	GTTTTCTGCG	8040
TTCAGCCTGA	GAAGGGGGGT	CGTAAGCCAG	CTCGTCTCAT	CGTGTTCCCC	GACCTGGGCG	8100
TGCGCGTGTG	CGAGAAGATG	GCCCTGTACG	ACGTGGTTAG	CAAGCTCCCC	CTGGCCGTGA	8160
TGGGAAGCTC	CTACGGATTC	CAATACTCAC	CAGGACAGCG	GGTTGAATTC	CTCGTGCAAG	8220
CGTGGAAGTC	CAAGAAGACC	CCGATGGGGT	TCTCGTATGA	TACCCGCTGT	TTTGACTCCA	8280
CAGTCACTGA	GAGCGACATC	CGTACGGAGG	AGGCAATTTA	CCAATGTTGT	GACCTGGACC	8340
CCCAAGCCCG	CGTGGCCATC	AAGTCCCTCA	CTGAGAGGCT	TTATGTTGGG	GGCCCTCTTA	8400
CCAATTCAAG	GGGGGAAAAC	TGCGGCTACC	GCAGGTGCCG	CGCGAGCGGC	GTACTGACAA	8460
CTAGCTGTGG	TAACACCCCTC	ACTTGCTACA	TCAAGGCCCCG	GGCAGCCTGT	CGAGCCGCAG	8520
GGCTCCAGGA	CTGCACCATG	CTCGTGTGTG	GCGACGACTT	AGTCGTTATC	TGTGAAAGTG	8580
CGGGGGTCCA	GGAGGACGCG	GCGAGCCTGA	GAGCCTTCAC	GGAGGCTATG	ACCAGGTACT	8640
CCGCCCCCCC	CGGGGACCCC	CCACAACCAG	AATACGACTT	GGAGCTTATA	ACATCATGCT	8700
CCTCCAACGT	GTCAGTCGCC	CACGACGGCG	CTGGAAAGAG	GGTCTACTAC	CTTACCCGTG	8760
ACCTTACAAC	CCCCCTCGCG	AGAGCCGCGT	GGGAGACAGC	AAGACACACT	CCAGTCAATT	8820
CCTGGCTAGG	CAACATAATC	ATGTTTGCCC	CCCACTGTG	GGCGAGGATG	ATACTGATGA	8880
CCCATTTCTT	TAGCGTCCTC	ATAGCCAGGG	ATCAGCTTGA	ACAGGCTCTT	AACTGTGAGA	8940
TCTACGGAGC	CTGCTACTCC	ATAGAACCAC	TGGATCTACC	TCCAATCATT	CAAAGACTCC	9000
ATGGCCTCAG	CGCATTTTCA	CTCCACAGTT	ACTCTCCAGG	TGAAATCAAT	AGGGTGCCCG	9060
CATGCCTCAG	AAAACCTGGG	GTCCCGCCCT	TGCGAGCTTG	GAGACACCGG	GCCCGGAGCG	9120
TCCGCGCTAG	GCTTCTGTCC	AGAGGAGGCA	GGGCTGCCAT	ATGTGGCAAG	TACCTCTTCA	9180
ACTGGGCAGT	AAGAACAAAAG	CTCAAACCTCA	CTCCAATAGC	GGCCGCTGGC	CGGCTGGACT	9240

10

20

30

40

TGTCCGGTTG GTTCACGGCT GGCTACAGCG GGGGAGACAT TTATCACAGC GTGTCTCATG	9300
CCCGGCCCGG CTGGTTCTGG TTTTGCCCTAC TCCTGCTCGC TGCAGGGGTA GGCATCTACC	9360
TCCTCCCCAA CCGATGAAGG TTGGGGTAAA CACTCCGGCC TCTTAGGCCA TTTCTGTTTT	9420
TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT CTTTTTTTTT	9480
TTTTTTTTTC TTTTTTTTTT TTTTTTTTTT CTTTCCTTCT TTTTTCCTTT CTTTTCCTTC	9540
CTTCTTTAAT GGTGGCTCCA TCTTAGCCCT AGTCACGGCT AGCTGTGAAA GGTCCGTGAG	9600
CCGCATGACT GCAGAGAGTG CTGATACTGG CCTCTCTGCA GATCATGTCTG CATTCACGCG	9660
TTCGAATTAA TTAAGTAGTG GGAATACGCG GGGTATGCCG CGTTTTAGCA TATTGACGAC	9720
CCAATTCCTCA TGTTTGACAG CTTATCATCG ATAAGCTTTA ATGCGGTAGT TTATCACAGT	9780
TAAATTGCTA ACGCAGTCAG GCACCGTGTA TGAAATCTAA CAATGCGCTC ATCGTCATCC	9840
TCGGCACCGT CACCCTGGAT GCTGTAGGCA TAGGCTTGGT TATGCCGGTA CTGCCGGGCC	9900
TCTTGCGGGA TATCGTCCAT TCCGACAGCA TCGCCAGTCA CTATGGCGTG CTGCTAGCGC	9960
TATATGCGTT GATGCAATTT CTATGCGCAC CCGTTCTCGG AGCACTGTCC GACCGCTTTG	10020
GCCGCCGCCC AGTCCTGCTC GCTTCGCTAC TTGGAGCCAC TATCGACTAC GCGATCATGG	10080
CGACCACACC CGTCCTGTGG ATCCTCTACG CCGGACGCAT CGTGGCCGGC ATCACC GGCG	10140
CCACAGGTGC GGTGCTGGC GCCTATATCG CCGACATCAC CGATGGGGAA GATCGGGCTC	10200
GCCACTTCGG GCTCATGAGC GCTTGTTTCG GCGTGGGTAT GGTGGCAGGC CCCGTGGCCG	10260
GGGGACTGTT GGGCGCCATC TCCTTGCAATG CACCATTCCT TGCGGCGGCG GTGCTCAACG	10320
GCCTCAACCT ACTACTGGGC TGCTTCCTAA TGCAGGAGTC GCATAAGGGA GAGCGTCGAC	10380
CGATGCCCTT GAGAGCCTTC AACCAGTCA GCTCCTTCCG GTGGGCGCGG GGCATGACTA	10440
TCGTCGCCGC ACTTATGACT GTCTTCTTTA TCATGCAACT CGTAGGACAG GTGCCGGCAG	10500
CGCTCTGGGT CATTTTCGGC GAGGACCGCT TTCGCTGGAG CGCGACGATG ATCGGCCTGT	10560
CGCTTGCGGT ATTCGGAATC TTGCACGCCC TCGCTCAAGC CTTTCGTCAT GTTCCCGCCA	10620
CCAAACGTTT CGGCGAGAAG CAGGCCATTA TCGCCGCGAT GCGGCGCGAC GCGCTGGGCT	10680
ACGTCTTGCT GCGGTTTCGG ACGCGAGGCT GGATGGCCTT CCCATTATG ATTCTTCTCG	10740
CTTCCGGCGG CATCGGGATG CCCGCGTTGC AGGCCATGCT GTCCAGGCAG GTAGATGACG	10800

10

20

30

40

ACCATCAGGG	ACAGCTTCAA	GGATCGCTCG	CGGCTCTTAC	CAGCCTAACT	TCGATCACTG	10860
GACCGCTGAT	CGTCACGGCG	ATTTATGCCG	CCTCGGCGAG	CACATGGAAC	GGGTTGGCAT	10920
GGATTGTAGG	CGCCGCCCTA	TACCTTGTCT	GCCTCCCCGC	GTTGCGTCGC	GGTGCA TGGA	10980
GCCGGGCCAC	CTCGACCTGA	ATGGAAGCCG	GCGGCACCTC	GCTAACGGAT	TCACCACTCC	11040
AAGAATTGGA	GCCAATCAAT	TCTTGCGGAG	AACTGTGAAT	GCGCAAACCA	ACCCTTGGA	11100
GAACATATCC	ATCGCGTCCG	CCATCTCCAG	CAGCCGCACG	CGGCGCATCT	CGGGCAGCGT	11160
TGGGTCCTGG	CCACGGGTGC	GCATGATCGT	GCTCCTGTCT	TTGAGGACCC	GGCTAGGCTG	11220
GCGGGGTTGC	CTTACTGGTT	AGCAGAATGA	ATCACCGATA	CGCGAGCGAA	CGTGAAGCGA	11280
CTGCTGCTGC	AAAACGTCTG	CGACCTGAGC	AACAACATGA	ATGGTCTTCG	GTTTCCGTGT	11340
TTCGTAAAGT	CTGGAAACGC	GGAAGTCAGC	GCCCTGCACC	ATTATGTTCC	GGATCTGCAT	11400
CGCAGGATGC	TGCTGGCTAC	CCTGTGGAAC	ACCTACATCT	GTATTAACGA	AGCGCTGGCA	11460
TTGACCCTGA	GTGATTTTTT	TCTGGTCCCG	CCGCATCCAT	ACCGCCAGTT	GTTTACCCTC	11520
ACAACGTTCC	AGTAACCGGG	CATGTTTCATC	ATCAGTAACC	CGTATCGTGA	GCATCCTCTC	11580
TCGTTTCATC	GGTATCATT	CCCCATGAA	CAGAAATTCC	CCCTTACACG	GAGGCATCAA	11640
GTGACCAAAC	AGGAAAAAAC	CGCCCTTAAC	ATGGCCCCGT	TTATCAGAAG	CCAGACATTA	11700
ACGCTTCTGG	AGAAACTCAA	CGAGCTGGAC	GCGGATGAAC	AGGCAGACAT	CTGTGAATCG	11760
CTTCACGACC	ACGCTGATGA	GCTTTACCGC	AGCTGCCTCG	CGCGTTTCGG	TGATGACGGT	11820
GAAAACCTCT	GACACATGCA	GCTCCCGGAG	ACGGTCACAG	CTTGTCTGTA	AGCGGATGCC	11880
GGGAGCAGAC	AAGCCCGTCA	GGGCGCGTCA	GCGGGTGTTG	GCGGGTGTCG	GGGCGCAGCC	11940
ATGACCCAGT	CACGTAGCGA	TAGCGGAGTG	TATACTGGCT	TAACATGCG	GCATCAGAGC	12000
AGATTGTACT	GAGAGTGCAC	CATATGCGGT	GTGAAATACC	GCACAGATGC	GTAAGGAGAA	12060
AATACCGCAT	CAGGCGCTCT	TCCGCTTCCT	CGCTCACTGA	CTCGCTGCGC	TCGGTCGTTC	12120
GGCTGCGGCG	AGCGGTATCA	GCTCACTCAA	AGGCGGTAAT	ACGGTTATCC	ACAGAATCAG	12180
GGGATAACGC	AGGAAAGAAC	ATGTGAGCAA	AAGGCCAGCA	AAAGGCCAGG	AACCGTAAAA	12240
AGGCCGCGTT	GCTGGCGTTT	TTCCATAGGC	TCCGCCCCCC	TGACGAGCAT	CACAAAAATC	12300
GACGCTCAAG	TCAGAGGTGG	CGAAACCCGA	CAGGACTATA	AAGATACCAG	GCGTTTCCCC	12360

10

20

30

40

CTGGAAGCTC CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG 12420  
 CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG TATCTCAGTT 12480  
 CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA ACCCCCCGTT CAGCCCGACC 12540  
 GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC GGTAAGACAC GACTTATCGC 12600  
 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG 12660  
 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG 12720  
 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA 12780  
 CCACCGCTGG TAGCGGTGGT TTTTTTGTTC GCAAGCAGCA GATTACGCGC AGAAAAAAG 12840  
 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT 12900  
 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTCT 12960  
 AGATAATACG ACTCACTATA 12980

10

( 2 ) SEQ ID NO: 6 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 21ベースペア

20

( B ) タイプ : 核酸

( C ) ストランドネス : 二重鎖

( D ) トポロジー : 直線

( ii ) 分子タイプ : DNA ( ゲノム )

( iii ) 仮説 : なし

( xi ) 配列記述 : SEQ ID NO: 6 :

GGCGACACTC CACCATAGAT C

21

( 2 ) SEQ ID NO: 7 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 20ベースペア

30

( B ) タイプ : 核酸

( C ) ストランドネス : 二重鎖

( D ) トポロジー : 直線

( ii ) 分子タイプ : DNA ( ゲノム )

( iii ) 仮説 : なし

( xi ) 配列記述 : SEQ ID NO: 7 :

TGGCACTACC CTCCAAGACC

20

( 2 ) SEQ ID NO: 8 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 27ベースペア

40

( B ) タイプ : 核酸

( C ) ストランドネス : 二重鎖

( D ) トポロジー : 直線

( ii ) 分子タイプ : DNA ( ゲノム )

( iii ) 仮説 : なし

( xi ) 配列記述 : SEQ ID NO: 8 :

ATGACACAAG GGGGCGCTCC GCACACT

27

( 2 ) SEQ ID NO: 9 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 18ベースペア

50

- (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 9 :  
**TCCTGCTTGT GGATGATG** 18
- (2) SEQ ID NO: 10 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 16 ベースペア 10  
 (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 10 :  
**TAGTTTGGTG ATGTCA** 16
- (2) SEQ ID NO: 11 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 17 ベースペア 20  
 (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 11 :  
**ACATAGGTGC CAGTAAG** 17
- (2) SEQ ID NO: 12 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 16 ベースペア 30  
 (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 12 :  
**CTGGCAACGT GCATCA** 16
- (2) SEQ ID NO: 13 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 18 ベースペア 40  
 (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 13 :  
**GGGTGAGAAC AATTACCA** 18
- (2) SEQ ID NO: 14 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 16 ベースペア 50

(B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 1 4 :  
**ATTGATGCCC AATGCG**

16

( 2 ) SEQ ID NO: 1 5 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 17ベースペア  
 (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 1 5 :  
**ACTGCCTGGG ATTCCCT**

17

( 2 ) SEQ ID NO: 1 6 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 18ベースペア  
 (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 1 6 :  
**CCACAGTGGC AGCGAGTG**

18

( 2 ) SEQ ID NO: 1 7 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 16ベースペア  
 (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 1 7 :  
**CATGGACGTC AACACG**

16

( 2 ) SEQ ID NO: 1 8 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 30ベースペア  
 (B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： DNA (ゲノム)  
 (iii) 仮説： なし  
 (xi) 配列記述： SEQ ID NO: 1 8 :  
**AATCTTCACC GGTGGGGAG GAGGTAGATG**

30

( 2 ) SEQ ID NO: 1 9 の情報：  
 (i) 配列特徴：  
 (A) 長さ： 9416ベースペア

10

20

30

40

50

(B) タイプ： 核酸  
 (C) スtrandネス： 二重鎖  
 (D) トポロジー： 直線  
 (ii) 分子タイプ： cDNA  
 (iii) 仮説： なし

(xi) 配列記述： SEQ ID NO: 19 :

GCCAGCCCECC TGATGGGGGC GAACTCCAC CATAGATCAC TCCCCTGTGA GGAACACTG	60	
TCTTCACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG TGTCGTGCAG CCTCCAGGAC	120	
CCCCCTCCC GGGAGAGCCA TAGTGGTCTG CGGAACCGGT GAGTACACCG GAATTGCCAG	180	10
GACGACCGGG TCCTTTCTTG GATAAACCCG CTCATGCCT GGAGATTG GCGTGCCCCC	240	
GCAAGACTGC TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG	300	
GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGCAC CATGAGCACG AATCCTAAAC	360	
CTCAAAGAAA AACCAAACGT AACACCAACC GTCGCCCACA GGACGTCGAG TTCCCGGGTG	420	
GCGGTCAGAT CGTTGGTGGA GTTTACTTGT TGCCGCGCAG GGGCCCTAGA TTGGGTGTGC	480	
GCGCGACGAG GAAGACTTCC GAGCGGTCGC AACCTCGTGG TAGACGTCAG CCTATCCCCA	540	
AGGCACGTCG GCCCCAGGGC AGGACCTGGG CTCAGCCCGG GTACCCTTGG CCCCTCTATG	600	20
GCAATGAGGG TTGCGGGTGG GCGGGATGGC TCCTGTCTCC CCGTGGCTCT CGGCCTAGCT	660	

GGGGCCCCAC AGACCCCCGG CGTAGGTCGC GCAATTGTTGGG TAAGGTCATC GATACCCCTTA	720
CGTGCGGCTT CGCCGACCTC ATGGGGTACA TACCGCTCGT CGGCGCCCCCT CTTGGAGGCG	780
CTGCCAGGGC CCTGGCGCAT GGCCTCCGGG TTCTGGAAGA CGGCGTGAAC TATGCAACAG	840
GGAACCTTCC TGGTTGCTCT TTCTCTATCT TCCTTCTGGC CCTGCTCTCT TGCCTGACTG	900
TGCCCCGCTTC AGCCTACCAA GTGCGCAATT CCTCGGGGCT TTACCATGTC ACCAATGATT	960
GCCCTAATTC GAGTATTGTG TACGAGGCGG CCGATGCCAT CCTGCACACT CCGGGGTGTG	1020
TCCCTTGCGT TCGCGAGGGT AACGCCTCGA GGTGTTGGGT GGCGGTGACC CCCACGGTGG	1080
CCACCAGGGA CGGCAAACTC CCCACAACGC AGCTTCGACG TCATATCGAT CTGCTTGTGG	1140
GGAGCGCCAC CCTCTGCTCA GCCCTCTACG TGGGGGACCT GTGCGGGTCT GTTTTTCTTG	1200
TTGGTCAACT GTTTACCTTC TCTCCCAGGC GCCACTGGAC GACGCAAAGC TGCAATTGTT	1260
CTATCTATCC CGGCCATATA ACGGGTCATC GCATGGCATG GGATATGATG ATGAACTGGT	1320
CCCCTACGGC AGCGTTGGTG GTAGCTCAGC TGCTCCGGAT CCCACAAGCC ATCATGGACA	1380
TGATCGCTGG TGCTCACTGG GGAGTCCTGG CGGGCATAGC GTATTTCTCC ATGGTGGGGA	1440
ACTGGGCGAA GGTCTGGTA GTGCTGCTGC TATTGCGCG CGTCGACGCG GAAACCCACG	1500
TCACCGGGGG AAGTGCCGGC CACACCACGG CTGGGCTTGT TGGTCTCCTT ACACCAGGCG	1560
CCAAGCAGAA CATCCAATG ATCAACACCA ACGGCAGTTG GCACATCAAT AGCAGGCCT	1620
TGAACTGCAA CGATAGCCTT ACCACCGGCT GGTTAGCAGG GCTCTTCTAT CGCCACAAAT	1680
TCAACTCTTC AGGCTGTCCT GAGAGGTTGG CCAGCTGCCG ACGCCTTACC GATTTTGCCC	1740
AGGGCTGGGG TCCCATCAGT TATGCCAACG GAAGCGGCCT TGACGAACGC CCCTACTGTT	1800
GGCACTACCC TCCAAGACCT TGTGGCATTG TGCCCGCAAA GAGCGTGTGT GGCCCGGTAT	1860
ATTGCTTCAC TCCCAGCCCC GTGGTGGTGG GAACGACCGA CAGGTCGGGC GCGCCTACCT	1920
ACAGCTGGGG TGCAAATGAT ACGGATGTCT TCGTCCTTAA CAACACCAGG CCACCGCTGG	1980
GCAATTGGTT CGGTTGTACC TGGATGAACT CAACTGGATT CACCAAAGTG TGCGGAGCGC	2040
CCCCTTGTTG CATCGGAGGG GTGGGCAACA ACACCTTGCT CTGCCCCACT GATTGCTTCC	2100
GCAAACATCC GGAAGCCACA TACTCTCGGT GCGGCTCCGG TCCCTGGATT ACACCCAGGT	2160
GCATGGTCGA CTACCCGTAT AGGCTTTGGC ACTATCCTTG TACTATCAAT TACACCATAT	2220

10

20

30

40



TCAAAGTCAG	GATGTACGTG	GGAGGGGTCG	AGCACAGGCT	GGAAGCGGCC	TGCAACTGGA	2280
CGCGGGGCGA	ACGCTGTGAT	CTGGAAGACA	GGGACAGGTC	CGAGCTCAGC	CCATTGCTGC	2340
TGTCCACCAC	ACAGTGGCAG	GTCCTTCCGT	GTTCTTTCAC	GACCCTGCCA	GCCTTGCTCA	2400
CCGGCCTCAT	CCACCTCCAC	CAGAACATTG	TGGACGTGCA	GTACTTGTAC	GGGGTGGGGT	2460
CAAGCATCGC	GTCCTGGGCC	ATTAAGTGGG	AGTACGTCGT	TCTCCTGTTC	CTTCTGCTTG	2520
CAGACGCGCG	CGTCTGCTCC	TGCTTGTTGA	TGATGTTACT	CATATCCCAA	GCGGAGGCGG	2580
CTTTGGAGAA	CCTCGTAATA	CTCAATGCAG	CATCCCTGGC	CGGGACGCAC	GGTCTTGTTG	2640
CCTTCCTCGT	GTTCTTCTGC	TTTGCGTGGT	ATCTGAAGGG	TAGGTGGGTG	CCCGGAGCGG	2700
TCTACGCCCT	CTACGGGATG	TGGCCTCTCC	TCCTGCTCCT	GCTGGCGTTG	CCTCAGCGGG	2760
CATACGCACT	GGACACGGAG	GTGGCCGCGT	CGTGTGGCGG	CGTTGTTCTT	GTCGGGTTAA	2820
TGGCGCTGAC	TCTGTCACCA	TATTACAAGC	GCTATATCAG	CTGGTGCATG	TGGTGGCTTC	2880
AGTATTTTCT	GACCAGAGTA	GAAGCGCAAC	TGCACGTGTG	GGTTCCCCCC	CTCAACGTCC	2940
GGGGGGGGCG	CGATGCCGTC	ATCTTACTCA	TGTGTGTTGT	ACACCCGACT	CTGGTATTTG	3000
ACATCACCAA	ACTACTCCTG	GCCATCTTCG	GACCCCTTTG	GATTCTTCAA	GCCAGTTTGC	3060
TTAAAGTCCC	CTACTTCGTG	CGCGTTCAAG	GCCTTCTCCG	GATCTGCGCG	CTAGCGCGGA	3120
AGATAGCCGG	AGGTCATTAC	GTGCAAATGG	CCATCATCAA	GTTGGGGGCG	CTTACTGGCA	3180
CCTATGTGTA	TAACCATCTC	ACCCCTCTTC	GAGACTGGGC	GCACAACGGC	CTGCGAGATC	3240
TGGCCGTGGC	TGTGGAACCA	GTCGTCTTCT	CCCGAATGGA	GACCAAGCTC	ATCACGTGGG	3300
GGGCAGATAC	CGCCGCGTGC	GGTGACATCA	TCAACGGCTT	GCCCCTCTCT	GCCCCTAGGG	3360
GCCAGGAGAT	ACTGCTTGGA	CCAGCCGACG	GAATGGTCTC	CAAGGGGTGG	AGGTTGCTGG	3420
CGCCCATCAC	GGCGTACGCC	CAGCAGACGA	GAGGCCTCCT	AGGGTGTATA	ATCACCAGCC	3480
TGACTGGCCG	GGACAAAAAC	CAAGTGGAGG	GTGAGGTCCA	GATCGTGTCA	ACTGCTACCC	3540
AAACCTTCCT	GGCAACGTGC	ATCAATGGGG	TATGCTGGAC	TGTCTACCAC	GGGGCCGGAA	3600
CGAGGACCAT	CGCATCACCC	AAGGGTCCTG	TCATCCAGAT	GTATACCAAT	GTGGACCAAG	3660
ACCTTGTTGG	CTGGCCCCGT	CCTCAAGGTT	CCCGCTCATT	GACACCCTGC	ACCTGCGGCT	3720
CCTCGGACCT	TTACCTGGTT	ACGAGGCACG	CCGACGTCAT	TCCCGTGCGC	CGGCGAGGTG	3780

10

20

30

40

ATAGCAGGGG TAGCCTGCTT TCGCCCCGGC CCATTTCTTA CCTAAAAGGC TCCTCGGGGG	3840
GTCCGCTGTT GTGCCCCGCG GGACACGCCG TGGGCCTATT CAGGGCCGCG GTGTGCACCC	3900
GTGGAGTGAC CAAGGCGGTG GACTTTATCC CTGTGGAGAA CCTAGAGACA ACCATGAGAT	3960
CCCCGGTGTG CACGGACAAC TCCTCTCCAC CAGCAGTGCC CCAGAGCTTC CAGGTGGCCC	4020
ACCTGCATGC TCCCACCGGC AGTGGTAAGA GCACCAAGGT CCCGGCTGCG TACGCAGCCC	4080
AGGGCTACAA GGTGTGGTG CTCAACCCCT CTGTTGCTGC AACGCTGGGC TTTGGTGCTT	4140
ACATGTCCAA GGCCCATGGG GTCGATCCTA ATATCAGGAC CGGGGTGAGA ACAATTACCA	4200
CTGGCAGCCC CATCACGTAC TCCACCTACG GCAAGTTCTT TGCCGACGGC GGGTGCTCAG	4260
GAGGCGCTTA TGACATAATA ATTTGTGACG AGTGCCACTC CACGGATGCC ACATCCATCT	4320
TGGGCATCGG CACTGTCTTT GACCAAGCAG AGACTGCGGG GCGGAGATTG GTTGTGCTCG	4380
CCACTGCTAC CCTCCGGGC TCCGTCACTG TGTCCCATCC TAACATCGAG GAGGTTGCTC	4440
TGTCCACCAC CGGAGAGATC CCTTTCTACG GCAAGGCTAT CCCCTCGAG GTGATCAAGG	4500
GGGGAAGACA TCTCATCTTC TGTCACCTAA AGAAGAAGTG CGACGAGCTC GCCGCGAAGC	4560
TGGTCGCATT GGGCATCAAT GCCGTGGCCT ACTACCGCGG ACTTGACGTG TCTGTCATCC	4620
CGACCAACGG CGATGTTGTC GTCGTGTCGA CCGATGCTCT CATGACTGGC TTTACCGGCG	4680
ACTTCGACTC TGTGATAGAC TGCAACACGT GTGTCACTCA GACAGTCGAT TTCAGCCTTG	4740
ACCCTACCTT TACCATTGAG ACAACCACGC TCCCCAGGA TGCTGTCTCC AGGACTCAGC	4800
GCCGGGGCAG GACTGGCAGG GGGGAAGCCAG GCATCTACAG ATTTGTGGCA CCGGGGGAGC	4860
GCCCCTCCGG CATGTTTCGAC TCGTCCGTCC TCTGTGAGTG CTATGACGCG GGCTGTGCTT	4920
GGTATGAGCT CATGCCCGCC GAGACTACAG TTAGGCTACG AGCGTACATG AACACCCCGG	4980
GGCTTCCCGT GTGCCAGGAC CATCTTGAAT TTTGGGAGGG CGTCTTTACG GGCCTCACCC	5040
ATATAGATGC CCACTTTCTA TCCCAGACAA AGCAGAGTGG GGAGAACTTT CTTTACCTGG	5100
TAGCGTACCA AGCCACCGTG TGCGCTAGGG CTCAAGCCCC TCCCCATCG TGGGACCAGA	5160
TGTGGAAGTG TTTGATCCGC CTTAAACCCA CCCTCCATGG GCCAACACCC CTGCTATACA	5220
GACTGGGCGC TGTTCAGAAT GAAGTCACCC TGACGCACCC AATCACCAA TACATCATGA	5280
CATGCATGTC GGCCGACCTG GAGGTCGTCA CGAGCACCTG GGTGCTCGTT GCGGCGTCC	5340

10

20

30

40

TGGCTGCTCT	GGCCGCGTAT	TGCCTGTCAA	CAGGCTGCGT	GGTCATAGTG	GGCAGGATTG	5400
TCTTGTCGGG	GAAGCCGGCA	ATTATACCTG	ACAGGGAGGT	TCTCTACCAG	GAGTTCGATG	5460
AGATGGAAGA	GTGCTCTCAG	CACTTACCGT	ACATCGAGCA	AGGGATGATG	CTCGCTGAGC	5520
AGTTCAAGCA	GAAGGCCCTC	GGCCTCCTGC	AGACCGCGTC	CCGCCATGCA	GAGGTTATCA	5580
CCCCTGCTGT	CCAGACCAAC	TGGCAGAAAC	TCGAGGTCTT	CTGGGCGAAG	CACATGTGGA	5640
ATTTTCATCAG	TGGGATACAA	TATTTGGCGG	GCCTGTCAAC	GCTGCCTGGT	AACCCCGCCA	5700
TTGCTTCATT	GATGGCTTTT	ACAGCTGCCG	TCACCAGCCC	ACTAACCCT	GGCCAAACCC	5760
TCCTCTTCAA	CATATTGGGG	GGGTGGGTGG	CTGCCAGCT	CGCCGCCCCC	GGTGCCGCTA	5820
CCGCCTTTGT	GGGCGCTGGC	TTAGCTGGCG	CCGCCATCGG	CAGCGTTGGA	CTGGGGAAGG	5880
TCCTCGTGGA	CATTCTTGCA	GGGTATGGCG	CGGGCGTGGC	GGGAGCTCTT	GTAGCATTCA	5940
AGATCATGAG	CGGTGAGGTC	CCCTCCACGG	AGGACCTGGT	CAATCTGCTG	CCCGCCATCC	6000
TCTCGCCTGG	AGCCCTTGTA	GTCGGTGTGG	TCTGCGCAGC	AATACTGCGC	CGGCACGTTG	6060
GCCCGGGCGA	GGGGGCAGTG	CAATGGATGA	ACCGGCTAAT	AGCCTTCGCC	TCCCGGGGGA	6120
ACCATGTTTC	CCCCACGCAC	TACGTGCCGG	AGAGCGATGC	AGCCGCCCCG	GTCAGTCCCA	6180
TACTCAGCAG	CCTCACTGTA	ACCCAGCTCC	TGAGGCGACT	ACATCAGTGG	ATAAGCTCGG	6240
AGTGTACCAC	TCCATGCTCC	GGCTCCTGGC	TAAGGGACAT	CTGGGACTGG	ATATGCGAGG	6300
TGCTGAGCGA	CTTTAAGACC	TGGCTGAAAG	CCAAGCTCAT	GCCACAACCTG	CCTGGGATTC	6360
CCTTTGTGTC	CTGCCAGCGC	GGGTATAGGG	GGGTCTGGCG	AGGAGACGGC	ATTATGCACA	6420
CTCGCTGCCA	CTGTGGAGCT	GAGATCACTG	GACATGTCAA	AAACGGGACG	ATGAGGATCG	6480
TCGGTCCTAG	GACCTGCAGG	AACATGTGGA	GTGGGACGTT	CCCCATTAAC	GCCTACACCA	6540
CGGGCCCCTG	TACTCCCCTT	CCTGCGCCGA	ACTATAAGTT	CGCGCTGTGG	AGGGTGTCTG	6600
CAGAGGAATA	CGTGGAGATA	AGGCGGGTGG	GGGACTTCCA	CTACGTATCG	GGTATGACTA	6660
CTGACAATCT	TAAATGCCCC	TGCCAGATCC	CATCGCCCCG	ATTTTTCACA	GAATTGGACG	6720
GGGTGCGCCT	ACATAGGTTT	GCGCCCCCTT	GCAAGCCCTT	GCTGCGGGAG	GAGGTATCAT	6780
TCAGAGTAGG	ACTCCACGAG	TACCCGGTGG	GGTCGCAATT	ACCTTGCGAG	CCCGAACCGG	6840
ACGTAGCCGT	GTTGACGTCC	ATGCTCACTG	ATCCCTCCCA	TATAACAGCA	GAGGCGGCCG	6900

10

20

30

40

GGAGAAGGTT	GGCGAGAGGG	TCACCCCTT	CTATGGCCAG	CTCCTCGGCC	AGCCAGCTGT	6960
CCGCTCCATC	TCTCAAGGCA	ACTTGCACCG	CCAACCATGA	CTCCCCTGAC	GCCGAGCTCA	7020
TAGAGGCTAA	CCTCCTGTGG	AGGCAGGAGA	TGGGCGGCAA	CATCACCAGG	GTTGAGTCAG	7080
AGAACAAAGT	GGTGATTCTG	GACTCCTTCG	ATCCGCTTGT	GGCAGAGGAG	GATGAGCGGG	7140
AGGTCTCCGT	ACCCGCAGAA	ATTCTGCGGA	AGTCTCGGAG	ATTGCCCCGG	GCCCTGCCCCG	7200
TTTGGGCGCG	GCCGGACTAC	AACCCCCCGC	TAGTAGAGAC	GTGGAAAAAG	CCTGACTACG	7260
AACCACCTGT	GGTCCATGGC	TGCCCCGTAC	CACCTCCACG	GTCCCCTCCT	GTGCCTCCGC	7320
CTCGGAAAAA	GCGTACGGTG	GTCCTCACCG	AATCAACCTT	ACCTACTGCC	TTGGCCGAGC	7380
TTGCCACCAA	AAGTTTGGC	AGCTCCTCAA	CTTCCGGCAT	TACGGGCGAC	AATATGACAA	7440
CATCCTCTGA	GCCCCCCCCCT	TCTGGCTGCC	CCCCCGACTC	CGACGTTGAG	TCCTATTCTT	7500
CCATGCCCCC	CCTGGAGGGG	GAGCCTGGGG	ATCCGGATTT	CAGCGACGGG	TCATGGTCGA	7560
CGGTCAGTAG	TGGGGCCGAC	ACGGAAGATG	TCGTGTGCTG	CTCAATGTCT	TATACCTGGA	7620
CAGGCGCACT	CGTCACCCCG	TGCGCTGCGG	AAGAACAAAA	ACTGCCCATC	AACGCACTGA	7680
GCAACTCGTT	GCTACGCCAT	CACAATCTGG	TATATTCCAC	CACTTCACGC	AGTGCTTGCC	7740
AAAGGCAGAA	GAAAGTCACA	TTTGACAGAC	TGCAAGTTCT	GGACAGCCAT	TACCAGGACG	7800
TGCTCAAGGA	GGTCAAAGCA	GCGGCGTCAA	AAGTGAAGGC	TAAGTTGCTA	TCCGTAGAGG	7860
AAGCTTGACG	CCTGACGCCC	CCACATTGAG	CCAAATCCAA	GTTTGGCTAT	GGGGCAAAAG	7920
ACGTCCGTTG	CCATGCCAGA	AAGGCCGTAG	CCCACATCAA	CTCCGTGTGG	AAAGACCTTC	7980
TGGAAGACAG	TGTAACACCA	ATAGACACTA	TCATCATGGC	CAAGAACGAG	GTCTTCTGCG	8040
TTCAGCCTGA	GAAGGGGGGT	CGTAAGCCAG	CTCGTCTCAT	CGTGTTCCCC	GACCTGGGCG	8100
TGCGCGTGTG	CGAGAAGATG	GCCCTGTACG	ACGTGGTTAG	CAAACTCCCC	CTGGCCGTGA	8160
TGGGAAGCTC	CTACGGATTC	CAATACTCAC	CAGGACAGCG	GGTTGAATTC	CTCGTGCAAG	8220
CGTGGAAGTC	CAAGAAGACC	CCGATGGGGT	TCCCGTATGA	TACCCGCTGT	TTTGACTCCA	8280
CAGTCACTGA	GAGCGACATC	CGTACGGAGG	AGGCAATTTA	CCAATGTTGT	GACCTGGACC	8340
CCCAAGCCCG	CGTGGCCATC	AAGTCCCTCA	CTGAGAGGCT	TTATGTTGGG	GGCCCTCTTA	8400
CCAATTCAAG	GGGGGAAAAC	TGCGGCTATC	GCAGGTGCCG	CGCGAGCGGC	GTAAGTACAA	8460

10

20

30

40

CTAGCTGTGG TAACACCCCTC ACTTGCTACA TCAAGGCCCG GGCAGCCCGT CGAGCCGCAG	8520
GGCTCCAGGA CTGCACCATG CTCGTGTGTG GCGACGACTT AGTCGTTATC TGTGAAAGTG	8580
CGGGGGTCCA GGAGGACGCG GCGAGCCTGA GAGCCTTTAC GGAGGCTATG ACCAGGTACT	8640
CCGCCCCCCC CGGGGACCCC CCACAACCAG AATACGACTT GGAGCTTATA ACATCATGCT	8700
CCTCCAACGT GTCAGTCGCC CACGACGGCG CTGGAAAAAG GGTCTACTAC CTTACCCGTG	8760
ACCCTACAAC CCCCCTCGCG AGAGCCGCGT GGGAGACAGC AAGACACACT CCAGTCAATT	8820
CCTGGCTAGG CAACATAATC ATGTTTGCCC CCACACTGTG GGCGAGGATG ATACTGATGA	8880
CCCATTTCTT TAGCGTCTC ATAGCCAGGG ATCAGCTTGA ACAGGCTCTT AACTGTGAGA	8940
TCTACGCAGC CTGCTACTCC ATAGAACCAC TGGATCTACC TCCAATCATT CAAAGACTCC	9000
ATGGCCTCAG CGCATTTTTA CTCCACAGTT ACTCTCCAGG TGAAGTCAAT AGGGTGGCCG	9060
CATGCCTCAG AAAACTTGGG GTCCCGCCCT TGCAGCTTG GAGACACCGG GCCCGGAGCG	9120
TCCGCGCTAG GCTTCTGTCC AGGGGAGGCA GGGCTGCCAT ATGTGGCAAG TACCTCTTCA	9180
ACTGGGCAGT AAGAACAAAG CTCAACTCA CTCCAATAGC GGCCGCTGGC CGGCTGGACT	9240
TGTCCGTTG GTTCACGGCT GGCTACAGCG GGGGAGACAT TTATCACAGC GTGTCTCATG	9300
CCCGGCCCCG CTGGTTCTGG TTTTGCTTAC TCCTGCTCGC TGCAGGGGTA GGCATCTACC	9360
TCCTCCCCAA CCGGTGAAGA TTGGGCTAAC CACTCCAGGC CAATAGGCCA TCCCCT	9416

( 2 ) SEQ ID NO: 2 0 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 3011 アミノ酸

( B ) タイプ : アミノ酸

( C ) スtrandネス : 一本鎖

( D ) トポロジー : 直線

( ii ) 分子タイプ : タンパク質

( iii ) 仮説 : なし

( v ) フラグメントタイプ : N 末端

( xi ) 配列記述 : SEQ ID NO: 2 0 :

10

20

30

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn  
 1 5 10 15  
 Arg Arg Pro Gln Asp Val Glu Phe Pro Gly Gly Gly Gln Ile Val Gly  
 20 25 30  
 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala  
 35 40 45  
 Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro  
 50 55 60  
 Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Thr Trp Ala Gln Pro Gly  
 65 70 75 80  
 Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Cys Gly Trp Ala Gly Trp  
 85 90 95  
 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Thr Asp Pro  
 100 105 110  
 Arg Arg Arg Ser Arg Asn Leu Gly Lys Val Ile Asp Thr Leu Thr Cys  
 115 120 125  
 Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro Leu Val Gly Ala Pro Leu  
 130 135 140  
 Gly Gly Ala Ala Arg Ala Leu Ala His Gly Val Arg Val Leu Glu Asp  
 145 150 155 160  
 Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile  
 165 170 175  
 Phe Leu Leu Ala Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Tyr  
 180 185 190  
 Gln Val Arg Asn Ser Ser Gly Leu Tyr His Val Thr Asn Asp Cys Pro  
 195 200 205  
 Asn Ser Ser Ile Val Tyr Glu Ala Ala Asp Ala Ile Leu His Thr Pro  
 210 215 220  
 Gly Cys Val Pro Cys Val Arg Glu Gly Asn Ala Ser Arg Cys Trp Val  
 225 230 235 240  
 Ala Val Thr Pro Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Thr Thr  
 245 250 255  
 Gln Leu Arg Arg His Ile Asp Leu Leu Val Gly Ser Ala Thr Leu Cys  
 260 265 270

10

20

30

Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val Gly  
 275 280 285  
 Gln Leu Phe Thr Phe Ser Pro Arg Arg His Trp Thr Thr Gln Ser Cys  
 290 295 300  
 Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala Trp  
 305 310 315 320  
 Asp Met Met Met Asn Trp Ser Pro Thr Ala Ala Leu Val Val Ala Gln  
 325 330 335  
 Leu Leu Arg Ile Pro Gln Ala Ile Met Asp Met Ile Ala Gly Ala His  
 340 345 350  
 Trp Gly Val Leu Ala Gly Ile Ala Tyr Phe Ser Met Val Gly Asn Trp  
 355 360 365  
 Ala Lys Val Leu Val Val Leu Leu Leu Phe Ala Gly Val Asp Ala Glu  
 370 375 380  
 Thr His Val Thr Gly Gly Ser Ala Gly His Thr Thr Ala Gly Leu Val  
 385 390 395 400  
 Gly Leu Leu Thr Pro Gly Ala Lys Gln Asn Ile Gln Leu Ile Asn Thr  
 405 410 415  
 Asn Gly Ser Trp His Ile Asn Ser Thr Ala Leu Asn Cys Asn Asp Ser  
 420 425 430  
 Leu Thr Thr Gly Trp Leu Ala Gly Leu Phe Tyr Arg His Lys Phe Asn  
 435 440 445  
 Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Arg Leu Thr Asp  
 450 455 460  
 Phe Ala Gln Gly Trp Gly Pro Ile Ser Tyr Ala Asn Gly Ser Gly Leu  
 465 470 475 480  
 Asp Glu Arg Pro Tyr Cys Trp His Tyr Pro Pro Arg Pro Cys Gly Ile  
 485 490 495  
 Val Pro Ala Lys Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro Ser  
 500 505 510  
 Pro Val Val Val Gly Thr Thr Asp Arg Ser Gly Ala Pro Thr Tyr Ser  
 515 520 525  
 Trp Gly Ala Asn Asp Thr Asp Val Phe Val Leu Asn Asn Thr Arg Pro  
 530 535 540

10

20

30

Pro Leu Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly Phe  
 545 550 555 560  
 Thr Lys Val Cys Gly Ala Pro Pro Cys Val Ile Gly Gly Val Gly Asn  
 565 570 575  
 Asn Thr Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys His Pro Glu Ala  
 580 585 590  
 Thr Tyr Ser Arg Cys Gly Ser Gly Pro Trp Ile Thr Pro Arg Cys Met  
 595 600 605  
 Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Ile Asn Tyr  
 610 615 620  
 Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg Leu  
 625 630 635 640  
 Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu Asp  
 645 650 655  
 Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Gln Trp  
 660 665 670  
 Gln Val Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly  
 675 680 685  
 Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly  
 690 695 700  
 Val Gly Ser Ser Ile Ala Ser Trp Ala Ile Lys Trp Glu Tyr Val Val  
 705 710 715 720  
 Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ser Cys Leu Trp  
 725 730 735  
 Met Met Leu Leu Ile Ser Gln Ala Glu Ala Ala Leu Glu Asn Leu Val  
 740 745 750  
 Ile Leu Asn Ala Ala Ser Leu Ala Gly Thr His Gly Leu Val Ser Phe  
 755 760 765  
 Leu Val Phe Phe Cys Phe Ala Trp Tyr Leu Lys Gly Arg Trp Val Pro  
 770 775 780  
 Gly Ala Val Tyr Ala Phe Tyr Gly Met Trp Pro Leu Leu Leu Leu Leu  
 785 790 795 800  
 Leu Ala Leu Pro Gln Arg Ala Tyr Ala Leu Asp Thr Glu Val Ala Ala  
 805 810 815

10

20

30



Ser Cys Gly Gly Val Val Leu Val Gly Leu Met Ala Leu Thr Leu Ser  
 820 825 830  
 Pro Tyr Tyr Lys Arg Tyr Ile Ser Trp Cys Met Trp Trp Leu Gln Tyr  
 835 840 845  
 Phe Leu Thr Arg Val Glu Ala Gln Leu His Val Trp Val Pro Pro Leu  
 850 855 860  
 Asn Val Arg Gly Gly Arg Asp Ala Val Ile Leu Leu Met Cys Val Val  
 865 870 875 880  
 His Pro Thr Leu Val Phe Asp Ile Thr Lys Leu Leu Leu Ala Ile Phe  
 885 890 895  
 Gly Pro Leu Trp Ile Leu Gln Ala Ser Leu Leu Lys Val Pro Tyr Phe  
 900 905 910  
 Val Arg Val Gln Gly Leu Leu Arg Ile Cys Ala Leu Ala Arg Lys Ile  
 915 920 925  
 Ala Gly Gly His Tyr Val Gln Met Ala Ile Ile Lys Leu Gly Ala Leu  
 930 935 940  
 Thr Gly Thr Tyr Val Tyr Asn His Leu Thr Pro Leu Arg Asp Trp Ala  
 945 950 955 960  
 His Asn Gly Leu Arg Asp Leu Ala Val Ala Val Glu Pro Val Val Phe  
 965 970 975  
 Ser Arg Met Glu Thr Lys Leu Ile Thr Trp Gly Ala Asp Thr Ala Ala  
 980 985 990  
 Cys Gly Asp Ile Ile Asn Gly Leu Pro Val Ser Ala Arg Arg Gly Gln  
 995 1000 1005  
 Glu Ile Leu Leu Gly Pro Ala Asp Gly Met Val Ser Lys Gly Trp Arg  
 1010 1015 1020  
 Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr Arg Gly Leu Leu  
 1025 1030 1035 1040  
 Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg Asp Lys Asn Gln Val Glu  
 1045 1050 1055  
 Gly Glu Val Gln Ile Val Ser Thr Ala Thr Gln Thr Phe Leu Ala Thr  
 1060 1065 1070  
 Cys Ile Asn Gly Val Cys Trp Thr Val Tyr His Gly Ala Gly Thr Arg  
 1075 1080 1085

10

20

30

Thr Ile Ala Ser Pro Lys Gly Pro Val Ile Gln Met Tyr Thr Asn Val  
 1090 1095 1100  
 Asp Gln Asp Leu Val Gly Trp Pro Ala Pro Gln Gly Ser Arg Ser Leu  
 1105 1110 1115 1120  
 Thr Pro Cys Thr Cys Gly Ser Ser Asp Leu Tyr Leu Val Thr Arg His  
 1125 1130 1135  
 Ala Asp Val Ile Pro Val Arg Arg Arg Gly Asp Ser Arg Gly Ser Leu  
 1140 1145 1150  
 Leu Ser Pro Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro  
 1155 1160 1165  
 Leu Leu Cys Pro Ala Gly His Ala Val Gly Leu Phe Arg Ala Ala Val  
 1170 1175 1180  
 Cys Thr Arg Gly Val Thr Lys Ala Val Asp Phe Ile Pro Val Glu Asn  
 1185 1190 1195 1200  
 Leu Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser Pro  
 1205 1210 1215  
 Pro Ala Val Pro Gln Ser Phe Gln Val Ala His Leu His Ala Pro Thr  
 1220 1225 1230  
 Gly Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala Ala Gln Gly  
 1235 1240 1245  
 Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe  
 1250 1255 1260  
 Gly Ala Tyr Met Ser Lys Ala His Gly Val Asp Pro Asn Ile Arg Thr  
 1265 1270 1275 1280  
 Gly Val Arg Thr Ile Thr Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr  
 1285 1290 1295  
 Gly Lys Phe Leu Ala Asp Gly Gly Cys Ser Gly Gly Ala Tyr Asp Ile  
 1300 1305 1310  
 Ile Ile Cys Asp Glu Cys His Ser Thr Asp Ala Thr Ser Ile Leu Gly  
 1315 1320 1325  
 Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Ala Arg Leu Val  
 1330 1335 1340  
 Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Val Ser His Pro  
 1345 1350 1355 1360

10

20

30

Asn Ile Glu Glu Val Ala Leu Ser Thr Thr Gly Glu Ile Pro Phe Tyr  
 1365 1370 1375  
 Gly Lys Ala Ile Pro Leu Glu Val Ile Lys Gly Gly Arg His Leu Ile  
 1380 1385 1390  
 Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val  
 1395 1400 1405  
 Ala Leu Gly Ile Asn Ala Val Ala Tyr Tyr Arg Gly Leu Asp Val Ser  
 1410 1415 1420  
 Val Ile Pro Thr Asn Gly Asp Val Val Val Val Ser Thr Asp Ala Leu  
 1425 1430 1435 1440  
 Met Thr Gly Phe Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn Thr  
 1445 1450 1455  
 Cys Val Thr Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe Thr Ile  
 1460 1465 1470  
 Glu Thr Thr Thr Leu Pro Gln Asp Ala Val Ser Arg Thr Gln Arg Arg  
 1475 1480 1485  
 Gly Arg Thr Gly Arg Gly Lys Pro Gly Ile Tyr Arg Phe Val Ala Pro  
 1490 1495 1500  
 Gly Glu Arg Pro Ser Gly Met Phe Asp Ser Ser Val Leu Cys Glu Cys  
 1505 1510 1515 1520  
 Tyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu Met Pro Ala Glu Thr Thr  
 1525 1530 1535  
 Val Arg Leu Arg Ala Tyr Met Asn Thr Pro Gly Leu Pro Val Cys Gln  
 1540 1545 1550  
 Asp His Leu Glu Phe Trp Glu Gly Val Phe Thr Gly Leu Thr His Ile  
 1555 1560 1565  
 Asp Ala His Phe Leu Ser Gln Thr Lys Gln Ser Gly Glu Asn Phe Pro  
 1570 1575 1580  
 Tyr Leu Val Ala Tyr Gln Ala Thr Val Cys Ala Arg Ala Gln Ala Pro  
 1585 1590 1595 1600  
 Pro Pro Ser Trp Asp Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro  
 1605 1610 1615  
 Thr Leu His Gly Pro Thr Pro Leu Leu Tyr Arg Leu Gly Ala Val Gln  
 1620 1625 1630

10

20

30

Asn Glu Val Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Thr Cys  
 1635 1640 1645  
 Met Ser Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly  
 1650 1655 1660  
 Gly Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Cys Val  
 1665 1670 1675 1680  
 Val Ile Val Gly Arg Ile Val Leu Ser Gly Lys Pro Ala Ile Ile Pro  
 1685 1690 1695  
 Asp Arg Glu Val Leu Tyr Gln Glu Phe Asp Glu Met Glu Glu Cys Ser  
 1700 1705 1710  
 Gln His Leu Pro Tyr Ile Glu Gln Gly Met Met Leu Ala Glu Gln Phe  
 1715 1720 1725  
 Lys Gln Lys Ala Leu Gly Leu Leu Gln Thr Ala Ser Arg His Ala Glu  
 1730 1735 1740  
 Val Ile Thr Pro Ala Val Gln Thr Asn Trp Gln Lys Leu Glu Val Phe  
 1745 1750 1755 1760  
 Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile Gln Tyr Leu Ala  
 1765 1770 1775  
 Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Ile Ala Ser Leu Met Ala  
 1780 1785 1790  
 Phe Thr Ala Ala Val Thr Ser Pro Leu Thr Thr Gly Gln Thr Leu Leu  
 1795 1800 1805  
 Phe Asn Ile Leu Gly Gly Trp Val Ala Ala Gln Leu Ala Ala Pro Gly  
 1810 1815 1820  
 Ala Ala Thr Ala Phe Val Gly Ala Gly Leu Ala Gly Ala Ala Ile Gly  
 1825 1830 1835 1840  
 Ser Val Gly Leu Gly Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly  
 1845 1850 1855  
 Ala Gly Val Ala Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu  
 1860 1865 1870  
 Val Pro Ser Thr Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser  
 1875 1880 1885  
 Pro Gly Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg  
 1890 1895 1900

10

20

30

His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile  
 1905 1910 1915 1920  
 Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val Pro  
 1925 1930 1935  
 Glu Ser Asp Ala Ala Ala Arg Val Thr Ala Ile Leu Ser Ser Leu Thr  
 1940 1945 1950  
 Val Thr Gln Leu Leu Arg Arg Leu His Gln Trp Ile Ser Ser Glu Cys  
 1955 1960 1965  
 Thr Thr Pro Cys Ser Gly Ser Trp Leu Arg Asp Ile Trp Asp Trp Ile  
 1970 1975 1980  
 Cys Glu Val Leu Ser Asp Phe Lys Thr Trp Leu Lys Ala Lys Leu Met  
 1985 1990 1995 2000  
 Pro Gln Leu Pro Gly Ile Pro Phe Val Ser Cys Gln Arg Gly Tyr Arg  
 2005 2010 2015  
 Gly Val Trp Arg Gly Asp Gly Ile Met His Thr Arg Cys His Cys Gly  
 2020 2025 2030  
 Ala Glu Ile Thr Gly His Val Lys Asn Gly Thr Met Arg Ile Val Gly  
 2035 2040 2045  
 Pro Arg Thr Cys Arg Asn Met Trp Ser Gly Thr Phe Pro Ile Asn Ala  
 2050 2055 2060  
 Tyr Thr Thr Gly Pro Cys Thr Pro Leu Pro Ala Pro Asn Tyr Lys Phe  
 2065 2070 2075 2080  
 Ala Leu Trp Arg Val Ser Ala Glu Glu Tyr Val Glu Ile Arg Arg Val  
 2085 2090 2095  
 Gly Asp Phe His Tyr Val Ser Gly Met Thr Thr Asp Asn Leu Lys Cys  
 2100 2105 2110  
 Pro Cys Gln Ile Pro Ser Pro Glu Phe Phe Thr Glu Leu Asp Gly Val  
 2115 2120 2125  
 Arg Leu His Arg Phe Ala Pro Pro Cys Lys Pro Leu Leu Arg Glu Glu  
 2130 2135 2140  
 Val Ser Phe Arg Val Gly Leu His Glu Tyr Pro Val Gly Ser Gln Leu  
 2145 2150 2155 2160  
 Pro Cys Glu Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu Thr  
 2165 2170 2175

10

20

30

Asp Pro Ser His Ile Thr Ala Glu Ala Ala Gly Arg Arg Leu Ala Arg  
 2180 2185 2190  
 Gly Ser Pro Pro Ser Met Ala Ser Ser Ser Ala Ser Gln Leu Ser Ala  
 2195 2200 2205  
 Pro Ser Leu Lys Ala Thr Cys Thr Ala Asn His Asp Ser Pro Asp Ala  
 2210 2215 2220  
 Glu Leu Ile Glu Ala Asn Leu Leu Trp Arg Gln Glu Met Gly Gly Asn  
 2225 2230 2235 2240  
 Ile Thr Arg Val Glu Ser Glu Asn Lys Val Val Ile Leu Asp Ser Phe  
 2245 2250 2255  
 Asp Pro Leu Val Ala Glu Glu Asp Glu Arg Glu Val Ser Val Pro Ala  
 2260 2265 2270  
 Glu Ile Leu Arg Lys Ser Arg Arg Phe Ala Arg Ala Leu Pro Val Trp  
 2275 2280 2285  
 Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Thr Trp Lys Lys Pro  
 2290 2295 2300  
 Asp Tyr Glu Pro Pro Val Val His Gly Cys Pro Leu Pro Pro Pro Arg  
 2305 2310 2315 2320  
 Ser Pro Pro Val Pro Pro Pro Arg Lys Lys Arg Thr Val Val Leu Thr  
 2325 2330 2335  
 Glu Ser Thr Leu Pro Thr Ala Leu Ala Glu Leu Ala Thr Lys Ser Phe  
 2340 2345 2350  
 Gly Ser Ser Ser Thr Ser Gly Ile Thr Gly Asp Asn Met Thr Thr Ser  
 2355 2360 2365  
 Ser Glu Pro Ala Pro Ser Gly Cys Pro Pro Asp Ser Asp Val Glu Ser  
 2370 2375 2380  
 Tyr Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Phe  
 2385 2390 2395 2400  
 Ser Asp Gly Ser Trp Ser Thr Val Ser Ser Gly Ala Asp Thr Glu Asp  
 2405 2410 2415  
 Val Val Cys Cys Ser Met Ser Tyr Thr Trp Thr Gly Ala Leu Val Thr  
 2420 2425 2430  
 Pro Cys Ala Ala Glu Glu Gln Lys Leu Pro Ile Asn Ala Leu Ser Asn  
 2435 2440 2445

10

20

30

Ser Leu Leu Arg His His Asn Leu Val Tyr Ser Thr Thr Ser Arg Ser  
 2450 2455 2460  
 Ala Cys Gln Arg Gln Lys Lys Val Thr Phe Asp Arg Leu Gln Val Leu  
 2465 2470 2475 2480  
 Asp Ser His Tyr Gln Asp Val Leu Lys Glu Val Lys Ala Ala Ala Ser  
 2485 2490 2495  
 Lys Val Lys Ala Asn Leu Leu Ser Val Glu Glu Ala Cys Ser Leu Thr  
 2500 2505 2510  
 Pro Pro His Ser Ala Lys Ser Lys Phe Gly Tyr Gly Ala Lys Asp Val  
 2515 2520 2525  
 Arg Cys His Ala Arg Lys Ala Val Ala His Ile Asn Ser Val Trp Lys  
 2530 2535 2540  
 Asp Leu Leu Glu Asp Ser Val Thr Pro Ile Asp Thr Ile Ile Met Ala  
 2545 2550 2555 2560  
 Lys Asn Glu Val Phe Cys Val Gln Pro Glu Lys Gly Gly Arg Lys Pro  
 2565 2570 2575  
 Ala Arg Leu Ile Val Phe Pro Asp Leu Gly Val Arg Val Cys Glu Lys  
 2580 2585 2590  
 Met Ala Leu Tyr Asp Val Val Ser Lys Leu Pro Leu Ala Val Met Gly  
 2595 2600 2605  
 Ser Ser Tyr Gly Phe Gln Tyr Ser Pro Gly Gln Arg Val Glu Phe Leu  
 2610 2615 2620  
 Val Gln Ala Trp Lys Ser Lys Lys Thr Pro Met Gly Phe Pro Tyr Asp  
 2625 2630 2635 2640  
 Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Ser Asp Ile Arg Thr Glu  
 2645 2650 2655  
 Glu Ala Ile Tyr Gln Cys Cys Asp Leu Asp Pro Gln Ala Arg Val Ala  
 2660 2665 2670  
 Ile Lys Ser Leu Thr Glu Arg Leu Tyr Val Gly Gly Pro Leu Thr Asn  
 2675 2680 2685  
 Ser Arg Gly Glu Asn Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val  
 2690 2695 2700  
 Leu Thr Thr Ser Cys Gly Asn Thr Leu Thr Cys Tyr Ile Lys Ala Arg  
 2705 2710 2715 2720

10

20

30

Ala Ala Arg Arg Ala Ala Gly Leu Gln Asp Cys Thr Met Leu Val Cys  
 2725 2730 2735

Gly Asp Asp Leu Val Val Ile Cys Glu Ser Ala Gly Val Gln Glu Asp  
 2740 2745 2750

Ala Ala Ser Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser Ala  
 2755 2760 2765

Pro Pro Gly Asp Pro Pro Gln Pro Glu Tyr Asp Leu Glu Leu Ile Thr  
 2770 2775 2780

Ser Cys Ser Ser Asn Val Ser Val Ala His Asp Gly Ala Gly Lys Arg  
 2785 2790 2795 2800

Val Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala Ala  
 2805 2810 2815

Trp Glu Thr Ala Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile  
 2820 2825 2830

Ile Met Phe Ala Pro Thr Leu Trp Ala Arg Met Ile Leu Met Thr His  
 2835 2840 2845

Phe Phe Ser Val Leu Ile Ala Arg Asp Gln Leu Glu Gln Ala Leu Asn  
 2850 2855 2860

Cys Glu Ile Tyr Ala Ala Cys Tyr Ser Ile Glu Pro Leu Asp Leu Pro  
 2865 2870 2875 2880

Pro Ile Ile Gln Arg Leu His Gly Leu Ser Ala Phe Leu Leu His Ser  
 2885 2890 2895

Tyr Ser Pro Gly Glu Val Asn Arg Val Ala Ala Cys Leu Arg Lys Leu  
 2900 2905 2910

Gly Val Pro Pro Leu Arg Ala Trp Arg His Arg Ala Arg Ser Val Arg  
 2915 2920 2925

Ala Arg Leu Leu Ser Arg Gly Gly Arg Ala Ala Ile Cys Gly Lys Tyr  
 2930 2935 2940

Leu Phe Asn Trp Ala Val Arg Thr Lys Leu Lys Leu Thr Pro Ile Ala  
 2945 2950 2955 2960

Ala Ala Gly Arg Leu Asp Leu Ser Gly Trp Phe Thr Ala Gly Tyr Ser  
 2965 2970 2975

Gly Gly Asp Ile Tyr His Ser Val Ser His Ala Arg Pro Arg Trp Phe  
 2980 2985 2990

Trp Phe Cys Leu Leu Leu Leu Ala Ala Gly Val Gly Ile Tyr Leu Leu  
 2995 3000 3005

Pro Asn Arg  
 3010

10

20

30

40

( 2 ) SEQ ID NO: 2 1 の情報 :

( i ) 配列特徴 :

( A ) 長さ : 10ベースペア

( B ) タイプ : 核酸

( C ) ストランドネス : 二重鎖

( D ) トポロジー : 直線

50



( ii ) 分子タイプ： DNA ( ゲノム )  
( iii ) 仮説： なし  
( xi ) 配列記述： SEQ ID NO: 2 1 :

TGTCGCATTC

10

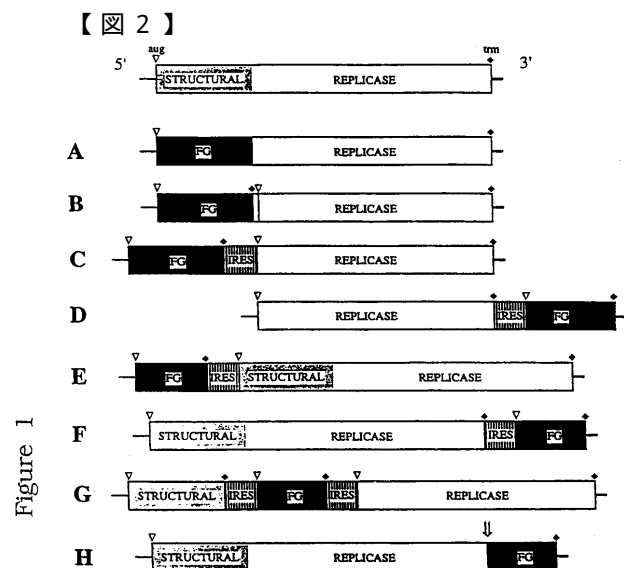
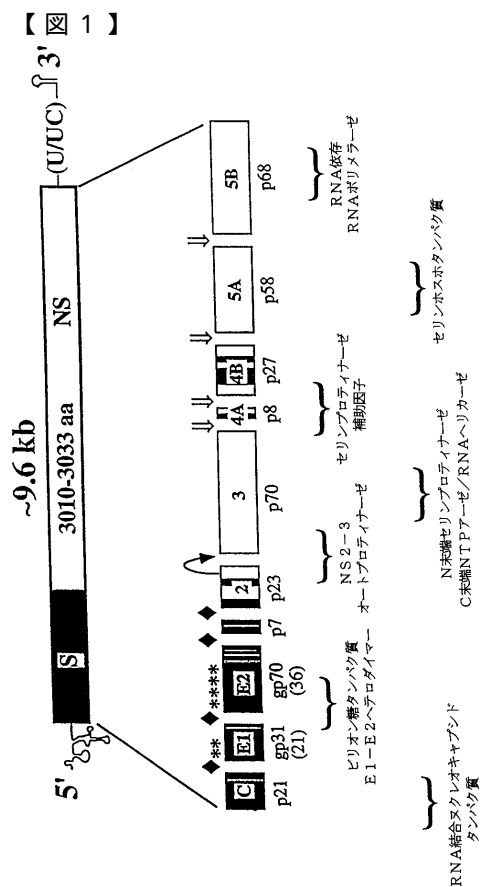


Figure 2

【図 3 a】

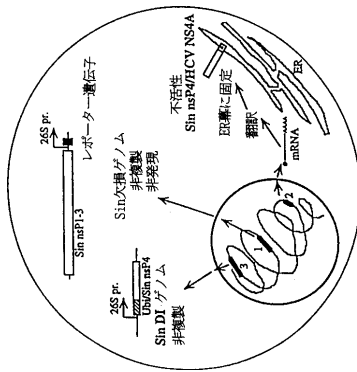


Figure 3a

【図 3 b】

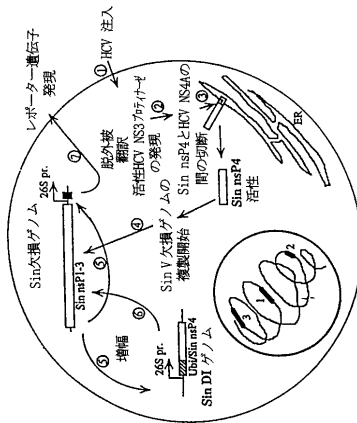


Figure 3b

【図 4】

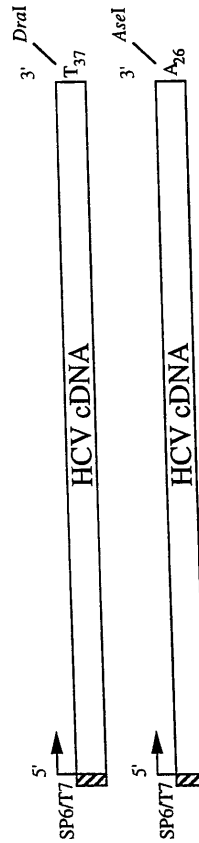


Figure 4

【図 5】

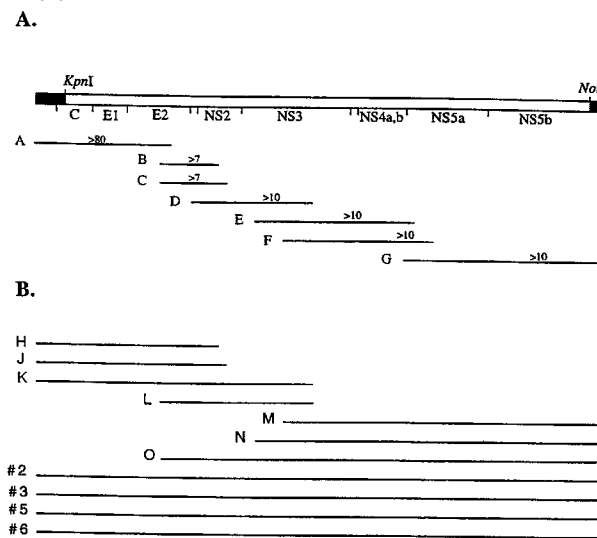


Figure 5

【図 6】

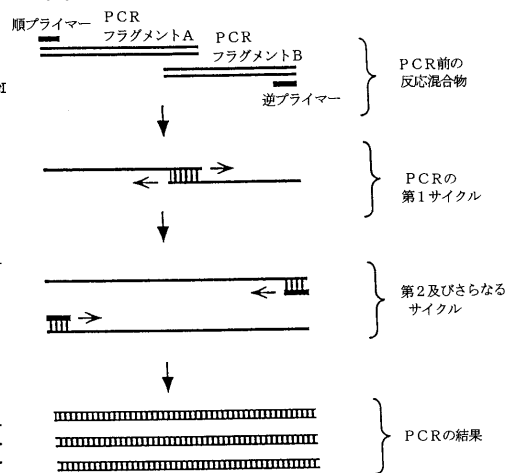


Figure 6

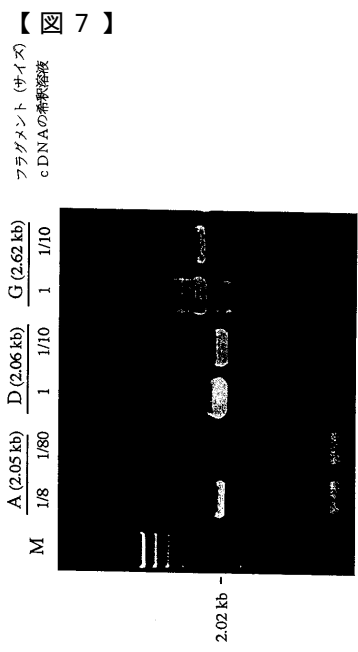


Figure 7

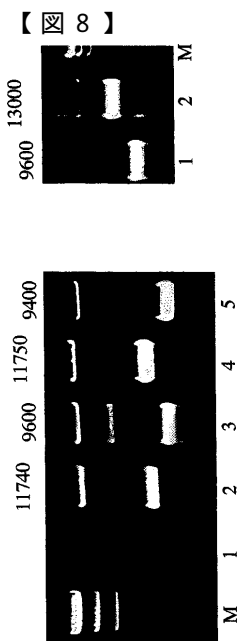


Figure 8a

Figure 8a

【図 9】  
AA

AB

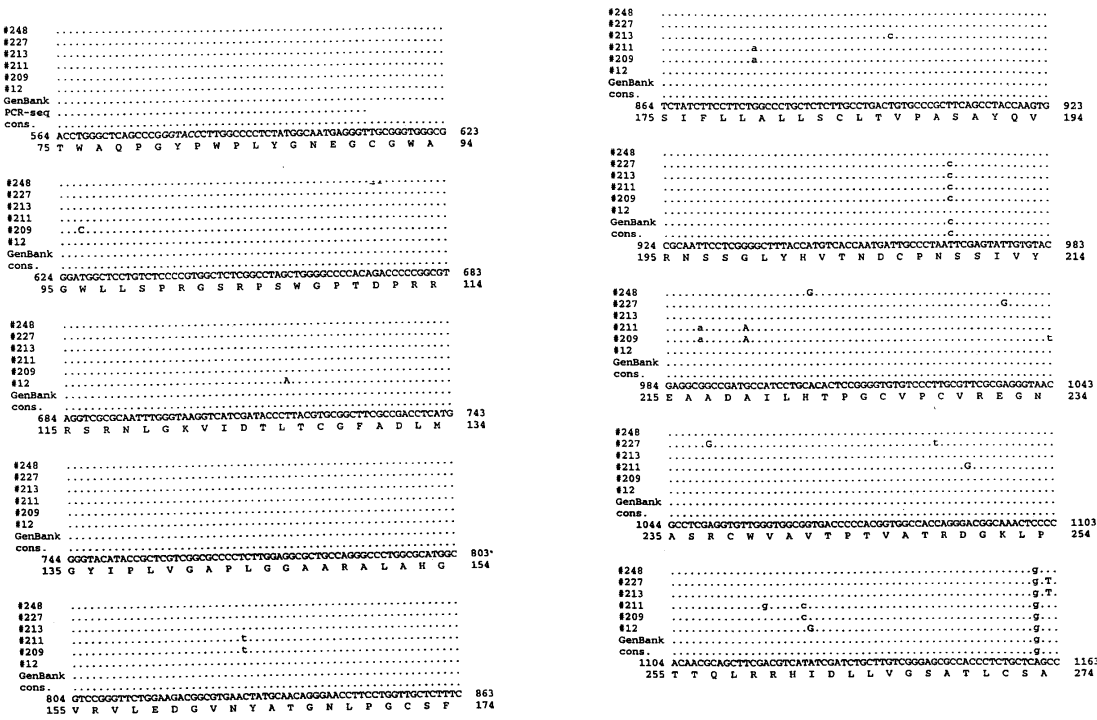


Figure 9

Figure 9

AC

AD

```
#248 .....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....
cons. ....C.....
1164 CTCCTACGTGGGGACCTGGGGGCTCTTTTACTGTGGTCAACTGTTAACTCTCT 1223
275 L Y V G D L C G S V F L V G Q L F T F S 294

#248 .....GA.....
#227 .....G.....
#213 .....GA.....
#211 .....GA.....
#209 .....G.....
#12 .....GA.....
GenBank .....GA.....
cons. ....GA.....
1224 CCCAGGGCCACTGGAGGACCAAGCTGCAATGTTCTATCTATCCCGGCATATACG 1283
295 P R R H W T T Q S C N C S I Y P G H I T 314

#248 .....
#227 .....G.....
#213 .....
#211 .....
#209 .....C.....
#12 .....
GenBank .....
cons. ....
1284 GTCATCGCATGGATATGATATGAACTGGTCCCTATGGCAGGTTTGGGATA 1343
315 G H R H A W D H M M N W S P T A A L V V 334

#248 .....C.....
#227 .....
#213 .....
#211 .....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
PCR-seq .....
cons. ....
1344 GCTCAGCTGCTCCGGATCCCAAGCATCATGGACATGATCGCTGGTCTACTGGGA 1403
335 A Q L L R I P Q A I N D H I A G A H W G 354

#248 .....
#227 .....T.....
#213 .....T.....
#211 .....
#209 .....G.....
#12 .....
GenBank .....AA.....
PCR-seq .....
cons. ....
1404 CTCCTGGCGGCAATAGCTATTTCTCCATGGTGGGGAAGTGGGCAAGGCTCTGTAGTG 1463
355 V L A G I A Y F S H V G N W A K V L V V 374
```

Figure 9

AE

AF

```
#83 .....C.....E.....
#84 .....E.A.....A.....
#86 .....E.A.....A.....
#87 .....G.....T.....
#89 .....C.....E.....A.....
#90 .....C.....E.....A.....
#92 .....C.....E.....A.....
#93 .....a.....E.A.....A.....
#95 .....C.....Y.....A.....
#96 .....C.....E.A.....A.....
#99 .....G.....E.A.....A.....
#101 .....C.....E.A.....A.....
#248 .....C.....E.A.....A.....
#227 .....E.A.....A.....
#213 .....E.A.....A.....
#211 .....C.....E.A.....G.....
#209 .....C.....E.....A.....
#12 .....
GenBank .....E.A.....A.....
PCR-seq .....E.A.....A.....
cons. ....E.A.....A.....
1584 AACACCAACGGCAGTTGGCATCATAGCAAGGCTTGAATGCAACGATAGCTTACC 1643
415 N T N G S W H I N S T A L N C N D S L T 434

#83 .....Aa.....
#84 .....Aa.....
#86 .....Aa.....
#87 .....G.....TA.....
#89 .....Aa.....
#90 .....Aa.....
#92 .....Aa.....G.....T.....G.....
#93 .....A.....A.....T.....G.....
#95 .....Aa.....
#96 .....Aa.....
#99 .....Ag.....
#101 .....Aa.....
#248 .....A.....
#227 .....A.....
#213 .....Ag.....
#211 .....Aa.....G.....
#209 .....
#12 .....A.....
GenBank .....A.....
PCR-seq .....Ag.....
cons. ....Ag.....
1644 ACCGCTGTTAGAGGGCTCTCTTCGCCAGAAATTCATCTTCAGGCTCTCTGAG 1703
435 T G W L A G L F Y R H K F N S G C P E 454
```

Figure 9

```
#83 .....G.....
#84 .....E.....
#86 .....A.....
#87 .....A.....G.....
#89 .....G.....
#90 .....G.....
#92 .....A.....
#93 .....A.....G.....
#95 .....A.....G.....
#96 .....A.....G.....
#99 .....G.....
#101 .....A.....G.....
#248 .....A.....G.....
#227 .....E.....
#213 .....G.....
#211 .....G.....
#209 .....G.....
#12 .....E.....
GenBank .....R.....G.....
PCR-seq .....
cons. ....
1464 CTCCTGCTATTTGGCGGCTGGAGGGAACCCACCTTACCCGGGAGTCCCGGCAC 1523
375 L L L F A G V D A E T H V T G G S A G H 394

#83 .....TA.....CT..AC.....
#84 .....C.....T.....
#86 .....G.....A.....C.....T.....T.....
#87 .....TA.....CT..AC.....
#89 .....TA.....CT..AC.....
#90 .....G.....C.....T.....T.....
#92 .....G.....C.....T.....T.....
#93 .....C.....
#95 .....TA.....CT..AC.....
#96 .....TA.....CT..AC.....
#99 .....A.....G.....
#101 .....C.....T.....
#248 .....C.....T.....
#227 .....C.....T.....
#213 .....A.....
#211 .....TA.....CT..AC.....
#209 .....C.....T.....
#12 .....C.....T.....
GenBank .....
PCR-seq .....
cons. ....
1524 ACCACGGCTGGGCTTTTGGTCTCTTACACAGGGCCAGGAGAACATCCAACTATC 1583
395 T T A G L V G L L T P G A K Q N I Q L I 414
```

Figure 9

```
#83 .....C.....C.....E.....
#84 .....C.....E.....
#86 .....C.....E.....
#87 .....C.....E.....C.....
#89 .....C.....E.....C.....
#90 .....C.....E.....C.....
#92 .....C.....E.....C.....
#93 .....T.....E.....E.....
#95 .....C.....Y.....E.....
#96 .....C.....Y.....E.....
#99 .....C.....E.....
#101 .....C.....E.....
#248 .....G.....E.....
#227 .....E.....
#213 .....C.....E.....
#211 .....C.....E.....
#209 .....C.....E.....G.....
#12 .....
GenBank .....
PCR-seq .....
cons. ....
1704 AGGTTGGCCAGCTGCGGCGGCTTACCGATTTCGCCAGGGCTGGGCTCCCATCAGTTAT 1763
455 R L A S C R R L T D F A Q G W G P I S Y 474

#83 .....C.....C.....T.....A.....
#84 .....C.....C.....T.....A.....
#86 .....C.....C.....T.....A.....
#87 .....C.....C.....T.....A.....
#89 .....C.....C.....T.....A.....
#90 .....C.....C.....T.....A.....
#92 .....C.....C.....T.....A.....
#93 .....C.....C.....T.....A.....
#95 .....C.....C.....T.....A.....
#96 .....C.....C.....T.....E.....
#99 .....C.....C.....T.....E.....
#101 .....C.....C.....T.....E.....
#248 .....C.....C.....T.....E.....
#227 .....C.....C.....T.....E.....
#213 .....C.....C.....T.....E.....
#211 .....C.....C.....T.....E.....
#209 .....C.....C.....T.....E.....
#12 .....C.....C.....T.....E.....
GenBank .....C.....C.....T.....E.....
PCR-seq .....C.....C.....T.....E.....
cons. ....C.....C.....T.....E.....
1764 CCACACGAGAGGGGCTGCGGAGGCTCTCTCTGAGCTCTCTCTCTCTCTCTCTCTCT 1823
475 A N G S G L D E R F Y C W H Y P P R P C 494
```

Figure 9

AH

#248	.....t.....g.....	
#227	.....g.....	
#213	.....g.....	
#211	.....g.....	
#209	.....t.....g.....	
#12	.....c.....t.....g.....	
GenBank	.....g.....	
PCR-seq	.....t.....g.....	
cons.	.....g.....	
2054	575 G N N T L L C P T D C F R K H P E A T Y	2127 594
#248	.....A.....	
#227	.....A.....	
#213	.....A.....	
#211	.....A.....	
#209	.....A.....	
#12	.....A.....	
GenBank	.....A.....	
PCR-seq	.....A.....	
cons.	.....A.....	
2124	T C T G G T G C G G C T C C G G T C C T G G A T T A C A C C C A G G T G C A T G G T C G A T A C C G T A T A G G	2183 614
595	S R C G S G S G F M I T F R C N V D Y P Y R	
#248	.....C.....	
#227	.....C.....	
#213	.....C.....	
#211	.....C.....	
#209	.....C.....	
#12	.....C.....y.....	
GenBank	.....C.....	
PCR-seq	.....C.....	
cons.	.....C.....	
2184	C T T T G G C A T A T C T T G T A C T A T C A A T T A C C A C A T A T C A A G T C A G A T G T A C T G G G A	2243 634
615	L W H Y P C T I N Y T I F K V R M Y V G	
#248	.....	
#227	.....	
#213	.....	
#211	.....	
#209	.....	
#12	.....	
GenBank	.....	
PCR-seq	.....	
cons.	.....	
2244	G G G G T C G A C C A G C T G G A G G G C C T C G C A C T G G A G C G G G G G C A C C T G T A T C T G	2303 655
635	G V E H R L E A A C N W T R G E R C D L	

Figure 9

AJ

```

#248 .....G.....
#227 .....A.....
#213 .....G.....
#211 .....G.....
#209 .....G.....
#12 .....G.....
GenBank .....
PCR-seq .....
cons. ....
2544 TTGTGGATGATGTTACTCATATCCCAAGCGAGCGCCTTTGAGACACTCTGATATAC 2603
735 LWNHMLLLISQAEAALENLVIL 754

#248 .....T.....
#227 .....C.....E.....
#213 .....E.....
#211 .....E.....
#209 .....E.....
#12 .....E.....
GenBank .....
PCR-seq .....
cons. ....
2604 AATGACGATCCTCGGCGGAGCCACGAGTCTTGTCCTTCCTCCTGTTCTTCTGCTTT 2663
755 NAAASLAGATHGLVLSFLVFFCF 774

#248 .....T.....
#227 .....T.....
#213 .....T.....
#211 .....T.....
#209 .....T.....
#12 .....T.....
GenBank .....
cons. ....
2664 CCGTGTGCTCTCTCCGCTGCGCTGATGGTGTCCCGAGACGCTTACGCTTCTACGGAGATGG 2723
775 AWWY LKGRWVPGAVYAFYGNW 794

#248 .....G.....
#227 .....G.....
#213 .....G.....
#211 .....G.....
#209 .....G.....
#12 .....G.....
GenBank .....
cons. ....
2724 CCGCTCTCTCTCTCCGCTGCGCTGATGGTGTCCCGAGACGCTTACGCTTCTACGGAGATGG 2783
795 PLLLLLLLLALPQRAYALDTEV 814

```

Figure 9

Figure 9

## AL

#248	.....	g	
#227	.....	G	g
#213	.....	g	
#209	.....	g	
#12	.....	g	
GenBank	.....	g	
PCR-seq	.....	g	
cons.	.....	g	
2784	CGCCGCTCGTGTGGGGCGCTTGTCGTCGGTTAAAGGGCTGACCTGTGCACATAT		2843
815	A A S C G G V V L V G L N A L T L S F		834
#248	.....	g	
#227	.....	g	
#213	.....	G	
#209	.....	G	
#12	.....	a	
GenBank	.....		
PCR-seq	.....		
cons.	.....		
2844	TACAGAGCGGTATACAGCTGTGCGTCTCAGTATTTCTGACGAGATAGAA		2903
835	Y K R Y I S W C H W V L Q Y F L T R V E		854
#248	.....		
#227	.....		
#213	.....		
#209	.....		
#12	.....		
GenBank	.....		
PCR-seq	.....		
cons.	.....		
2904	GCGCAATGCACAGTGTGGGTTCGCCCCCTCAAGTTCGGGGGGGCGCGATGCGCTATC		2963
855	A Q L H V W V P F L N V R G G R D A V I		874
#248	.....	a	C
#227	.....	C	
#213	.....	C	
#209	.....		T
#12	.....	a	
GenBank	.....	C	g
PCR-seq	.....	x	y
cons.	.....		
864	TTATCTATATGTTGTGTGACACCGCTCTGGTATGTGACATGACCAAACTCTCTGGG		892
875	L L L C V V H F T L V F D I T K L L L A		904

Figure 9

#248	.....	
#227	.....	
#213	.....	
#211	.....	
#209	.....	
#12	.....	
GenBank	.....	
PCR-seq	.....	
cons.	.....	
3024	ATCTTCGAGACCCCTTTGGATTCTTCAAGCCAGTTTGGCTTAAAGTCCCTACTTCGTGGGC	3083
895	I T F G G P L W I L Q A S L L K V P Y F V R	914
#248	.....G.....	
#227	.....	
#213	.....	
#211	.....	
#209	.....	
#12	.....G.....	
GenBank	.....	
PCR-seq	.....	
cons.	.....	
3084	GTTCAAGCCTCTTCCGGATCTGGCGCTAGCGCGAAGATAGCCGGAGTCATTACGTG	3143
915	V Q G L L R I C A L A R K I A G G H Y V	934
#248	.....a.....	
#227	.....a.....	
#213	.....a.....	
#211	.....a.....	
#209	.....a.....	
#12	.....G.....a.....	
GenBank	.....a.....G.....G.t	
PCR-seq	.....a.....	
cons.	.....	
3144	CTCTTCGAGACTGGCGCCAGCGGCTCGGAGATCTGGCGTGGCTGTGGAACGTC	3203
935	Q M A I I K L G A L T G T Y V Y N H L T	954
#248	.....	
#227	.....	
#213	.....C.....	
#211	.....g.....	
#209	.....	
#12	.....A.....	
GenBank	.....	
PCR-seq	.....	
cons.	.....	
3204	CCTCTTCGAGACTGGCGCCAGCGGCTCGGAGATCTGGCGTGGCTGTGGAACGTC	3263
955	P L R D W A H N G L R D L A V A P V P	974

Figure 9

## AN

[illegible]

Figure 9

#248	.....	
#227	.....a.....	
#213	.....	
#211	.....	
#209	.....	
#12	.....	
GenBank	.....g.....	
cons.	.....	
3504	CPGAGGGTGCAGCTCCGACGCTCTGACACGCTGACCAAGCTCTCCGGGACATCCGAC	3563
1055	V E G E V Q I V S T A T Q T F L A T C I	1074
#248	.....	
#227	.....	
#213	.....	
#211	.....g.....	
#209	.....	
#12	.....	
GenBank	.....	
PCR-seq	.....	
cons.	.....	
3564	AATGGGGTATCTCTGATCTGCTACGACGGGGCTCGACGACGAGCGATGATCCGAC	3623
1075	N G V C W T V Y H G A G T R T I A S P K	1094
#248	.....	
#227	.....	
#213	.....	
#211	.....G.....	
#209	.....	
#12	.....	
GenBank	.....	
PCR-seq	.....	
cons.	.....	
3624	AGTGGCTTCTCTCTGATCTGCTACGACGGGGCTCGACGACGAGCGATGATCCGAC	3683
1095	G P V I Q H Y T N V D Q D L V G W P A P	1114
#248	.....C.....C.....	
#227	.....C.....C.....	
#213	.....C.....C.....	
#211	.....C.....C.....	
#209	.....C.....C.....	
#12	.....C.....C.....	
GenBank	.....C.....C.....	
PCR-seq	.....C.....C.....	
cons.	.....C.....C.....	
3684	CAATGGCTTCTCTCTGATCTGCTACGACGGGGCTCGACGACGAGCGATGATCCGAC	3743
1115	Q G S R S L T P C T A C G S S D L Y L G T C	1134
#248	.....t.....	
#227	.....t.....	
#213	.....t.....	
#211	G.....t.....	
#209	.....t.....	
#12	.....t.....	
GenBank	.....t.....	
PCR-seq	.....t.....	
cons.	.....t.....	
3744	AGGACGCGCGGCTGATCTGCTACGACGGGGCTCGACGAGCGAGCGATGATCCGAC	3803
1153	R H A D V I P V R R R G D S R G S L L S	1194

Figure 9

AO

AP

```
#248 .....t.g.....
#227 .....t.g.....
#213 .....t.g.....
#211 .....t.g.....
#209 .....t.g.....
#12 .....t.g.....
GenBank .....A.....
PCR-seq .....t.g.....
cons. ....
3804 CCCCAGCCCAATTTCCTACCTAAAGGCTCTCGGGGGTCCGCTGTGTGCCCCCGGGA 3863
1155 P R P I S Y L K G S S G G P L L C P A G 1174

#248 .....G.t.....
#227 .....G.t.....
#213 .....G.t.....
#211 .....G.....
#209 .....G.....
#12 .....G.t.....
GenBank .....G.t.....
PCR-seq .....G.t.....
cons. ....
3864 CACGCCGTGGGCTATTTCAGGGCCGGGTGTGCACCCGTGAGTGACCAAGGCGGTGGAC 3923
1175 H A V G L F R A A V C T R G V T K A V D 1194

#248 C.....G.....
#227 .....G.....
#213 .....G.....
#211 .....G.....
#209 .....G.....
#12 .....G.....
GenBank .....C.....
cons. ....
3924 TTATCCCTGTGGAGAACCTAGAGCAACCATGAGATCCCGGTGTTCACGGACAATCC 3983
1195 F I P V E N L E T T M R S P V F T D N S 1214

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
cons. ....
3984 TCTCCACAGCAGTCCCAAGCTTCAGGTGGCCACCTGCATGCTCCACCGGAGT 4043
1215 S P P A V P Q S F Q V A H L H A P T G S 1234

#248 .....A.....
#227 .....A.....
#213 .....A.....
#211 .....A.....
#209 .....A.....
#12 .....A.....
GenBank .....A.....
cons. ....
4044 GGTAAAGACACCAAGCTCCCGCTGCGTACGACGCCAGGCTACAGGTGTGTGCTC 4103
1235 G K S T K V P A A Y A A Q G Y K V L V L 1254
```

Figure 9

```
#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
cons. ....
4104 AACCCCTCTGTGCTGCAACGCTGGGCTTGTGTCTTACATGTCCAGGCCATAGGGTC 4163
1255 N P S V A A T L G F G A Y M S K A H G V 1274

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
cons. ....
4164 GATCCTAATATCAGGACCGGGGTGAGAACATTACCACCTGGCAGCCCATACGTACTCC 4223
1275 D P N I R T G V R T I T T G S P I T Y S 1294

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
cons. ....
4224 ACCTACGGCAAGTTCCTTCCCGACGGCGGTCTCAGAGGGCGCTTATGACATATAAT 4283
1295 T Y G K F L A D G G C S G G A Y D I I I 1314

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
cons. ....
4284 TGTGACGAGTGGCACTCCACGGATGCCACATCTCTGGCATCGGCATGTCTTGAC 4343
1315 C D E C H S T D A T S I L G I G T V L D 1334

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
cons. ....
4344 CAAGCAGAGACTCGGGGGGAGATTTGTTGTCTGCCACGTGTACCCCTCCGGGCTCC 4403
1335 Q A E T A G A R L V V L A T A T F F G S 1354
```

Figure 9

AQ

AR

```
#248 .....c.....
#227 .....y.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....c.....
cons. ....
4404 GTCACGTGTGCCATCTCACTACGAGAGGTTCCTCTGTCCACCGGAGAGATCCCT 4463
1355 V T V S H P N I E E V A L S T T G E I P 1374

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
PCR-seq .....t.....
cons. ....
4464 TTCTAAGGACAGCTATCCCGCTGAGGTGATCAAGGGGGAAGACATCTCATCTCTCT 4523
1375 Y Y G K A I P L E V I K G G R H L I F C 1394

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
PCR-seq .....t.....
cons. ....
4524 CACTCAAGAGAGAGTCCAGGAGCTCCCGGAGCTGTGCTCTTGGGCAATGATGCC 4583
1395 H S K K K C D E L A A K L V A L G I N A 1414

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
PCR-seq .....t.....
cons. ....
4584 GTGCGCAACGATCTCTCAATGATGCTTACCGGAGCTCCGCTCTGTGATGAGTCTC 4643
1415 V A Y Y R G L D V S V I P T N G D V V V 1434

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
PCR-seq .....t.....
cons. ....
4644 GTGCGCAACGATCTCTCAATGATGCTTACCGGAGCTCCGCTCTGTGATGAGTCTC 4703
1435 V S T D A L H T G F T G D F D S V I D C 1454
```

Figure 9

```
#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
PCR-seq .....t.....
cons. ....
4704 AACAGTGTGTCACTCAGACGTGATTTCAGCTTGAACCTACCTTACCATTTAGACA 4763
1455 N T C V T Q T V D F S L D P T F T I E T 1474

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
PCR-seq .....t.....
cons. ....
4764 ACCAGCTCCCGAGGATGCTGTCTCCAGGCTACACCCCGGGGAGAGTCCGAGGG 4823
1475 T T L P Q D A V S R T Q R R G R T G R G 1494

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
cons. ....
4824 AAGCCAGGATCTACAGATTTGTGGACCCGGGAGGCGCCCTCCGCAATGTCTGATCG 4883
1495 K P G I Y R F V A P G E R P S G M F D S 1514

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
cons. ....
4884 TCCGCTCTCTGAGTCTATGACACCGGGCTGTGCTTGGTATGAGCTCATCCCGCGAG 4943
1515 S V L C E C Y D A G C A W Y E L H P A E 1534

#248 .....t.....
#227 .....t.....
#213 .....t.....
#211 .....t.....
#209 .....t.....
#12 .....t.....
GenBank .....t.....
cons. ....
4944 ACTACAGTTAGGCTACGAGCGTACATGAACACCGGGGCTCCCGTGTCCAGAGACAT 5003
1535 T T V R L R A Y H N T P G L P V C Q D H 1554
```

Figure 9

AS

AT

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....G.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....G.....  
5004 CTTGAAATTCTGGAGGGCTCTTTACGGGCTTACCCATATAGATGGCCACTTCTCATCC 5061  
1555 L E F W E G V F T G L T H I D A H F L S 1574

#248 .....T.....  
#227 .....T.....  
#213 .....C.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....G.....  
5064 CAGACAAAGCAGATGGGGAGACTTCTCTTAACTGTAGGTACCAAGCCACCGGTGTGC 5123  
1575 Q T K Q S G E N F P Y L V A Y Q A T V C 1594

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5124 GCTAGGGCTCAAGCCCTCCGCCATCTGTGGACAGATGTGGAAAGTCTTGATCCGGCTT 5183  
1595 A R A Q A P P P S W D Q M W K C L I R L 1614

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5184 AAACCCACCCCTCCATGGGCAACCCCTCTCTATACAGACTGGGGCTCTTCAGAAATGAA 5243  
1615 K P T L H G P T P L L Y R L G A V Q N E 1634

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5244 GTCAACCTGACGACCCCAATCAACCAATATATATGACATGCAATGTGGCCGACCTGTGAG 5303  
1635 V T L T H P I T K Y I N T C M S A D L E 1654

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....G.....  
5304 GTCTCACAGACCTGGGTCTCTGTGGGGCTCTCTGCTGCTCTGGCCGCTATATGC 5363  
1655 V V T S T W V L V G G V L A A L A A Y C 1674

#248 .....T.....  
#227 .....C.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5364 CTGTCAACAGCTGGCTGTCTATGTGGGAGGATTCTCTTCTCCGGGAAGCCGCAATT 5423  
1675 L S T G C V V I V G R I V L S G K P A I 1694

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5424 ATACCTGACAGGAGGTCTCTTACAGGAGTTCGATGAGATGGAGAGTCTCTCAGCAC 5483  
1695 I P D R E V L Y Q E F D E H E E C S Q H 1714

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5484 TTACCTTACATCGACAAAGGATGATGCTGTGCTGAGGATTCAGAGAGAGGCTCTGGC 5543  
1715 L P Y I E Q G M M L A E Q F K Q K A L G 1734

#248 .....T.....  
#227 .....A.....  
#213 .....T.....  
#211 .....A.....  
#209 .....T.....  
#12 .....C.....  
GenBank  
cons. ....A.....  
5544 CTCTGACAGCCGCTCCGGCATGACAGGATTCACCCCTGCTGTCTCAGACCACTGG 5603  
1735 L L Q T A S R H A E V I T P A V Q T N W 1754

Figure 9

Figure 9

AU

AV

#248 .....T.....  
#227 .....C.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5604 CAGAACTCAGGCTCTCTGGGCAAGCAGATGTGGAATTCATCAGTGGATACAAATAT 5663  
1755 Q K L E V F W A K H M W N F I S G I Q Y 1774

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5664 TTGGCGGGCTGTCAAGCTGCTGCTAACCCTGCTTCTTATGATGGCTTTTACA 5723  
1775 L A G L S T L P G N P A I A S L H A F T 1794

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5724 GCTGCGCTCACCAGCCACTTAACCTGGGCAAGCTCTCTCTTCAACATATGGGGGG 5783  
1795 A A V T S P L T T G Q T L L F N I L G G 1814

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5784 TGGGTGGCTGCCAGCTGGCCCGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5843  
1815 W V A A Q L A A P G A A T A F V G A G L 1834

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5844 GCTGCGCGCCATCTGCAAGCTTGGATGGGAAAGTCTCTGTTGACATCTTTCAGGG 5903  
1835 A G A A I G S V G L G K V L V D I L A G 1854

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....G.....  
5904 TATGGCGGGCTGGCGGAGCTCTTGTAGATTCATGATGATGATGATGATGATGATGATG 5963  
1855 Y G A G V A G A L V A F K I M S G E V P 1874

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....C.....  
5964 TCCAGGAGAGCTGGTCAATCTCTGCGGCAATCTCTGCGCTGGAGCCCTGTGTATC 6023  
1875 S T E D L V N L L P A I L S P G A L V V 1894

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
cons. ....T.....  
6024 GTGTGCTGCTGCGGCAATCTCTGCGGCAATCTCTGCGCTGGAGCCCTGTGTATC 6083  
1895 G V V C A A I L R R H V G P G E G A V Q 1914

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
PCR-seq  
cons. ....T.....  
6084 TGGATGAACCGCTAATAGCTTGGCTCTCCGGGGGACCATTTTCCCGCCGCTATC 6143  
1915 W M N R L I A F A S R G H V S P T H Y 1934

#248 .....T.....  
#227 .....T.....  
#213 .....T.....  
#211 .....T.....  
#209 .....T.....  
#12 .....T.....  
GenBank  
PCR-seq  
cons. ....T.....  
6144 GTGTGCTGCTGCGGCAATCTCTGCGGCAATCTCTGCGCTGGAGCCCTGTGTATC 6203  
1935 V P E S D A A A R V T A I L S S L T V T 1954

Figure 9

Figure 9



AW

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6204 CAGCTCCTGAGCGGACTACATCAGTGGATAAGCTCGAGTGTACCACTCCATGCTCCGGC 6263  
1955 Q L L R R L L H Q W I S S E C T T P C S G 1974

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6264 TCCTGGCTAAGGACATCTGGGACTGGATATGCGAGGTGCTGAGCGACTTTAAGACCTGG 6323  
1975 S W L R D I W D W I C E V L S D F K T W 1994

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6324 CTGAAGGCCAAGCTCATGCCACACTGCTGGGATTCCTTTGTGTCTCCAGCGCGGG 6383  
1995 L K A K L M P Q L P G I P F V S C Q R G 2014

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6384 TATAGGGGGTCTGGGAGGAGAGGCATATGACACATCGCTGCCACTGTGAGCTGAG 6443  
2015 Y R G V W R G D G I M H T R C H C G A E 2034

AX

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6444 ATCACTGGACATGTCAAAACGGGACGATAGGATGCTCGGTCTAGGAGCTCGAGGAAC 6503  
2035 I T G H V K N G T M R I V G P R T C R N 2054

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6504 ATGTGAGTGGGACGTTCCCATTAACGCTACACACGGGCCCCCTACTCCCTTCCT 6563  
2055 H W S G T F P I N A Y T T G P C T P L P 2074

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6564 GCGCCGAACTATAAGTTGCGCTGTGGAGGTGTCTGAGAGGATACGTGAGATAGG 6623  
2075 A P N Y K F A L W R V S A E E Y V E I R 2094

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6624 CCGGTGGGGACTTCACTACGTATCGGTATGACTCTGACAACTTAAATGCCCTGC 6683  
2095 R V G D F H Y V S G M T T D N L K C P C 2114

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6684 CAGATCCCATCGCCGAAATTTTCACAGANTGGACGGGTGCGCTACATAGTTTGGG 6743  
2115 Q I P S P E F F T E L D G V R L H R F A 2134

Figure 9

Figure 9

AY

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6744 CCGCTTCAAGCCCTCTCTGGGAGGAGGTATATTCAGAGTAGGACTCCAGAGTAC 6803  
2135 P F C K P L L R E E V S F R V G L H E Y 2154

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6804 CCGGTGGGTCCCATATTCCTGGGAGCGGACACCGGAGTAGCCGTGTGAAGTCCATG 6863  
2155 P V G S Q L P C E P E P D V A V L T S H 2174

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6864 CTCACTGATCCCTCCCATATACAGCAGAGCGCGCGGAGAGGTGGCGAGAGGTCA 6923  
2175 L T D P S H I T A E A A G R R L A R G S 2194

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6924 CCCCCCTATAGGCGAGCTCTCGGCCAGCGAGCTCTCCGCTCCATCTCTCAAGGCACT 6983  
2195 P P S H A S S S A S Q L S A P S L K A T 2214

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
6984 TGCAACCCCAACCATGACTCCCTGAGCGCGAGCTCATAGAGCTAACCTCTGTGGAG 7043  
2215 C T A N H D S F D A E L I E A N L L W R 2234

AZ

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
7044 CAGGAGATGGCGGCAACATCACAGAGGTGATGACAGACAAAGTGGTGAATCTGAC 7103  
2235 Q E M G G N I T R V E S E N K V V I L D 2254

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
7104 TCCTTCGATCCGCTTGTGGCAGAGGAGATGAGCGGAGGTCTCCGTACCGCGAATTT 7163  
2255 S F D P L V A E E D E R E V S V P A E I 2274

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
7164 CTGCGAAGCTCGGAGTGTGCGCGGCTTCCGTTTGGCGCGCGGCTACAC 7223  
2275 L R K S R R F A R A L P V W A R P D Y N 2294

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
7224 CCCCCCTATAGGCGAGCTCTCGGCCAGCGAGCTCTCCGCTCCATCTCTCAAGGCACT 7283  
2295 P P L V E T W K K P D Y E P P V V H G C 2314

#248 .....  
#227 .....  
#213 .....  
#211 .....  
#209 .....  
#12 .....  
GenBank .....  
PCR-seq .....  
cons. ....  
7284 CCGCTGACACTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7343  
2315 P L P P P R S P P V P P P R K K R T V V 2334

Figure 9

Figure 9

BA

```
#248 .....T.....
#227 .....T.....
#213 .....T.....
#211 .....T.....
#209 .....T.....
#12 .....C.....
GenBank .....T.....
cons. ....T.....
7344 CTCACCGAATCAACCTACCTATCTGCTTGGCGAGCTTCCACCAAAAGTTGGAGC 7403
2335 L T E S T L F T A L A E L A T K S F G S 2354

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
cons. ....C.....
7404 TCCTCAAGTTCCGGCATTAACGGGAGCAATATGACACATCTCTCGAGCCGCCCTCTCT 7463
2355 S S T S G I T G D N H T T S S E P A F S 2374

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons. ....
7464 GGCTGCCGCCCGACTCCGACTGAGTCTTATCTCTTCCATGCCGCCCTCTGAGGGGAG 7523
2375 G C P F D S D V E S Y S S H P P L E G E 2394

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons. ....
7524 CCTGGGAGTCCGATTTAGGACCGCTCATCTGTCGACCTCTAGTAGTGGGACGACAG 7583
2395 P G D P D F S D G S W S T V S S G A D T 2414

#248 .....T.....
#227 .....T.....
#213 .....T.....
#211 .....T.....
#209 .....T.....
#12 .....g.....
GenBank .....T.....
cons. ....T.....
7584 GAAGATGCTGCTGCTCTCAATGCTTATACCTGGACAGGCGCACTCTCTACCCGCTGC 7643
2415 E D V V C C S M S Y T W T G A L V T P C 2434
```

Figure 9

BB

```
#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons. ....
7644 GCTCCGAGAGAACAAACTGCCATCAAGCCACTGAGCAACTCTCTGCTACCCATC 7703
2435 A A E E Q K L P I N A L S N S L L R H H 2454

#248 .....
#227 .....
#213 .....g.....
#211 .....g.....
#209 .....a.....
#12 .....a.g.....
GenBank .....A.....
cons. ....g.....
7704 AATCTGGTATATTCCACCACTTCAAGCAGTCTTGCCAAAGCGAGAGAAAGTACATTT 7763
2455 N L V Y S T T T S R S A C Q R Q K K V T F 2474

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons. ....
7764 GACAGACTGCAAGTCTCTACACCCATCAAGCACTCTCTCAAGGATCAAGGACCC 7823
2475 D R L Q V L D S H Y Q D V L K E V K A A 2494

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons. ....
7824 GGCTCAAAAGTGAAGGCTAATCTCTATCTGATGAGGAGCTTGCAGCTGAGCGCCCA 7883
2495 A S K V K A N L L S V E E A C S K L T P P 2514
```

Figure 9

BC

```
#248 .....E.....
#227 .....E.....
#213 .....E.....
#211 .....E.....
#209 .....E.....
#12 .....
GenBank .....
cons. ....
7884 CATTCAGCCAAATCAAGTTTGGCTATGGGCAAAAGACCTCCGTTGCCATGCCAAG 7943
2515 H S A K S K F G Y G A K D V R C H A R K 2534

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
cons. ....C.....
7944 GCCGTAGCCCATCACTACCTCGTGTGGAAAGACCTTCTGGAAGACAGTATACACCAATA 8003
2535 A V A H I N S V W K D L L E D S V T P I 2554

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
cons. ....C.....
8004 GACACTATCATATCTGCAAGAGAGCTCTCTCTGCTTCAAGCTGAGAGGGGGTCTGT 8063
2555 D T I I N A K N E V F C V Q P E K G G R 2574

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
cons. ....C.....
8064 AAGCCAGCTGCTCTCTCGCTGCTCCGAGCTGGGCGTGGGCTGTGGAGAGATGGCC 8123
2575 K P A R L I V F P D L G V R V C E K M A 2594

#248 .....g.....
#227 .....g.....
#213 .....g.....
#211 .....g.....
#209 .....g.....
#12 .....g.....
GenBank .....g.....
cons. ....g.....
8124 CTGTACGACGTGTTAGCAAACTCCCGCTGGCCGTATGGGAAGCTCTCTACGATTCCAA 8183
2595 L Y D V V S K L P L A V H G S S Y G F Q 2614
```

Figure 9

BD

```
#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons. ....
8184 TACTCACCGAGACAGCGGTTGAATCTCTCTGCAAGCTGAGAGAGAGAGAGAGAGAG 8243
2615 Y S P G Q R V E F L V Q A W K S K K T P 2634

#248 .....T.....
#227 .....T.....
#213 .....T.....
#211 .....T.....
#209 .....T.....
#12 .....
GenBank .....T.....
cons. ....T.....
8244 ATGGGGTTCCCGTATGATACCCGCTGTTTGTACCTCCAGCTCACTGAGAGAGAGAGAG 8303
2635 H G F P Y D T R C F D S T V T E S D I R 2654

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons. ....
8304 ACAGAGGAGCAATTTACCAATGTTTGTACCTCCAGCTCACTGAGAGAGAGAGAGAGAG 8363
2655 T E E A I Y Q C C D L D P Q A R V A I K 2674

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons. ....
8364 TCCCTCACTGAGAGCTTATATGTTGGGGGCTCTCAATCAAGGGGAGAACTCTC 8423
2675 S L T E R L Y V G G P L T N S R G E N C 2694

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank .....C.....
cons. ....C.....
8424 GGCTATGCCAGGTGGCCGGGAGGGGCTACTGACACTAGCTCTGTACAGCCCTCACT 8483
2695 G Y R R C R A S G V L T T S C G N T L T 2714
```

Figure 9

BE

BF

```
#248 .....T.C.....
#227 .....T.....
#213 .....T.....
#211 .....T.....
#209 .....T.....
#12 .....T.....
GenBank C.....T.....
cons.....T.....
8484 TCCATGATCAGAGGCTTGGGACAGCCCTGAGAGCCGCTGACAGATCCACCATCTC 8543
2715 C Y I K A R A A R R A A G L Q D C T H L 2734

#248 .....A.....
#227 .....A.....
#213 .....A.....
#211 .....A.....
#209 .....A.....
#12 .....A.....
GenBank C.....A.....
cons.....A.....
8544 GGTGTGCGACGACTTATGCTTATCTCTGTAAGTGGGGGTCCAGAGGACGCGGCG 8603
2735 V C G D D L L V I C E S A G V Q E D A A 2754

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank C.....C.....
cons.....C.....
8604 AGCTGTGAGGCTTTTACGAGGCTTACGAGGCTTCTGCGCCGCGGAGCCGCCA 8663
2755 S L R A F T E A M T R Y S A P P G D P P 2774

#248 .....C.....
#227 .....C.....
#213 .....C.....
#211 .....C.....
#209 .....C.....
#12 .....C.....
GenBank C.....C.....
cons.....C.....
8664 CAAACGATGAGGCTTACGAGGCTTACGAGGCTTCTGCGCCGCGGAGCCGCCA 8723
2775 Q P E Y D L E L I T S C S S N V S V A H 2794

#248 .....g.....
#227 .....g.....
#213 .....g.....
#211 .....g.....
#209 .....g.....
#12 .....g.....
GenBank C.....g.....
cons.....g.....
8724 GAGCGCGCGGAAAGAGGCTTACGAGGCTTACGAGGCTTCTGCGCCGCGGAGCCGCCA 8783
2795 D G A G R V Y Y L T R D P T T P L A R 2814
```

```
#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons.....
8784 GCGCGTGGGAGACAGCAAGACACTCCAGTCATTTCCGCTAGGCAACATATCAG 8843
2815 A A W E T A R H T P V N S W L G N I I M 2834

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons.....
8844 TTTGCCCCACACTGTGGGAGAGTATATCTGAGAGCCATTCTTTAGCGTCTCTATA 8903
2835 F A P T L W A R M I L M T H F F S V L I 2854

#248 .....G.....
#227 .....G.....
#213 .....G.....
#211 .....G.....
#209 .....G.....
#12 .....G.....
GenBank .....G.....
cons.....G.....
8904 GCCAGGGATCAGCTTGAAGCACTCTTAAGTCTGAGATCAAGCGCTGCTACTCTATA 8963
2855 A R D Q L E Q A L N C E I Y A A C Y S I 2874

#248 .....G.....
#227 .....G.....
#213 .....G.....
#211 .....G.....
#209 .....G.....
#12 .....G.....
GenBank .....G.....
cons.....G.....
8964 GAACCACTGGATCTACTCCATCATTTCAAGACTCCATGGCTCAGCGCATTTTACTC 9023
2875 E P L D L P P I I Q R L R H G L S A F L L 2894

#248 .....A.....
#227 .....A.....
#213 .....A.....
#211 .....A.....
#209 .....A.....
#12 .....A.....
GenBank .....A.....
cons.....A.....
9024 CACAGTACTCTCCAGTGAAGTCAATAGGTTGGCGCATGCCCTGAGAAACTTGGGTC 9083
2895 H S Y S P G E V N R V A A C L R K L G V 2914
```

Figure 9

Figure 9

BG

```
#248 .....t.....
#227 .....a.....
#213 .....a.....
#211 .....a.....
#209 .....a.....
#12 .....a.....
GenBank .....a.....
PCR-seq .....a.....
cons.....a.....
9084 CCGCCCTTGGAGCTTGGAGACACCGGCGGAGGCTTCTGCGCCGCGGAGCCGCCA 9143
2915 P F L R A W R H R A R S V R A R L L S R 2934

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
cons.....
9144 GGAGGCAGGGCTCCATATAGGCGGCTGCGCGCTGGACTTCTCGGTGGTTCAAGGCTC 9203
2935 G G R A A I C G K Y L F N W A V R T K L 2954

#248 .....
#227 .....
#213 .....
#211 .....
#209 .....
#12 .....
GenBank .....
PCR-seq .....
cons.....
9204 AAACCTACTTCATAGGCGGCTGCGCGCTGGACTTCTCGGTGGTTCAAGGCTC 9263
2955 K L T P I A A A G R L D L S G W F T A G 2974
```

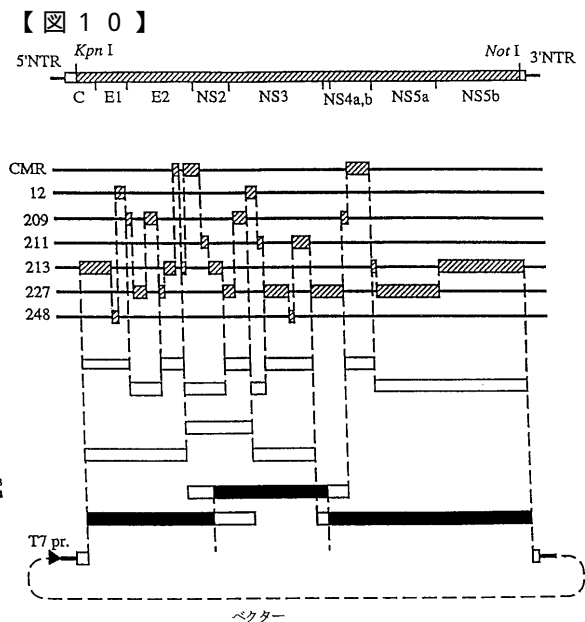


Figure 10

Figure 9

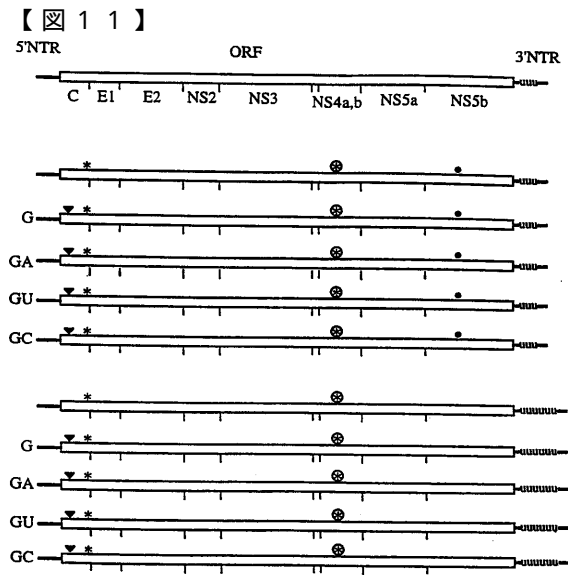


Figure 11

## フロントページの続き

(51)Int.Cl.			F I		
C 1 2 N	1/15	(2006.01)	C 1 2 N	1/15	
C 1 2 N	1/19	(2006.01)	C 1 2 N	1/19	
C 1 2 N	1/21	(2006.01)	C 1 2 N	1/21	
C 1 2 N	5/10	(2006.01)	C 1 2 N	5/00	A
C 1 2 N	7/00	(2006.01)	C 1 2 N	7/00	
C 1 2 P	21/02	(2006.01)	C 1 2 P	21/02	C
C 1 2 P	21/08	(2006.01)	C 1 2 P	21/08	
C 1 2 Q	1/68	(2006.01)	C 1 2 Q	1/68	A
C 1 2 Q	1/70	(2006.01)	C 1 2 Q	1/70	
G 0 1 N	33/15	(2006.01)	G 0 1 N	33/15	Z
G 0 1 N	33/50	(2006.01)	G 0 1 N	33/50	Z
G 0 1 N	33/576	(2006.01)	G 0 1 N	33/576	Z

- (72)発明者 ライス, チャールズ エム サード  
 アメリカ合衆国, ミズーリ州 6 3 1 3 0, ユニヴァーシティ・シティ, コルゲイト・アヴェニュー  
 7 3 1 6
- (72)発明者 コリハロフ, アレキサンダー エイ  
 アメリカ合衆国, ミズーリ州 6 3 1 3 9, セントルイス, サルファ・アヴェニュー 3 5 5 7

審査官 森井 隆信

- (56)参考文献 国際公開第 9 6 / 0 3 4 9 7 6 (WO, A 1)  
 Journal of Virology, 米国, 1 9 9 5 年, vol.69, 32-38  
 Journal of Virology, 米国, 1 9 9 6 年, vol.70, 3363-3371

## (58)調査した分野(Int.Cl., D B 名)

C12N 15/00  
 A01K 67/027  
 A61K 39/29  
 A61K 48/00  
 A61P 31/12  
 C12N 1/00  
 C12N 7/00  
 C12Q 1/68  
 G01N 33/00  
 BIOSIS/MEDLINE/WPIDS(STN)  
 GenBank/EMBL/DDBJ/GeneSeq  
 UniProt/Geneseq  
 PubMed