

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号

特表2023-521663

(P2023-521663A)

(43)公表日 令和5年5月25日(2023.5.25)

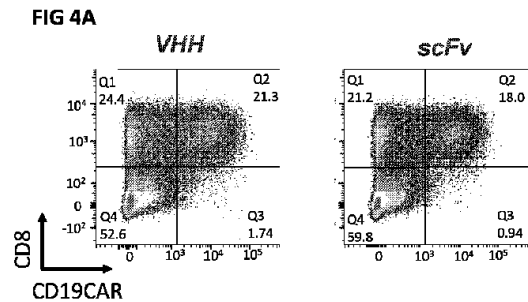
(51)国際特許分類		F I	テーマコード(参考)	
C 1 2 N	15/867 (2006.01)	C 1 2 N	15/867	Z Z N A 4 B 0 6 5
C 1 2 N	7/01 (2006.01)	C 1 2 N	7/01	4 C 0 7 6
C 1 2 N	15/45 (2006.01)	C 1 2 N	15/45	4 C 0 8 4
C 1 2 N	15/12 (2006.01)	C 1 2 N	15/12	4 C 0 8 7
C 1 2 N	15/13 (2006.01)	C 1 2 N	15/13	
審査請求 未請求 予備審査請求 未請求 (全376頁) 最終頁に続く				
(21)出願番号	特願2022-560227(P2022-560227)	(71)出願人	522043688	
(86)(22)出願日	令和3年3月30日(2021.3.30)		サナ バイオテクノロジー, インコーポ レイテッド	
(85)翻訳文提出日	令和4年11月29日(2022.11.29)		アメリカ合衆国 ワシントン州 9 8 1 0 2 シアトル イースト プレイン ストリ ート 1 8 8 スイート 4 0 0	
(86)国際出願番号	PCT/US2021/024993	(71)出願人	519397987	
(87)国際公開番号	WO2021/202604		フラッグシップ パイオニアリング イノ ベーションズ ブイ, インコーポレイテ ッド	
(87)国際公開日	令和3年10月7日(2021.10.7)		アメリカ合衆国 マサチューセッツ 0 2 1 4 2, ケンブリッジ, ケンブリッジ パークウェイ 5 5, 8 ティーエイチ フロア	
(31)優先権主張番号	63/003,168	(74)代理人	100102978	
(32)優先日	令和2年3月31日(2020.3.31)			最終頁に続く
(33)優先権主張国・地域又は機関	米国(US)			
(31)優先権主張番号	63/154,341			
(32)優先日	令和3年2月26日(2021.2.26)			
(33)優先権主張国・地域又は機関	米国(US)			
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA 最終頁に続く			

(54)【発明の名称】 標的化脂質粒子及び組成物ならびにその使用

(57)【要約】

本明細書では、内腔またはキャビティを包囲する脂質二重層と、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び結合ドメイン、例えば、シングルドメイン抗体(s d A b)可変ドメインを含有する標的化エンベロープタンパク質とを含有する脂質粒子が提供される。また、本明細書では、結合ドメイン、例えば、s d A b可変ドメインに融合または連結されたGタンパク質を含有する標的化エンベロープタンパク質、及びそのようなタンパク質をコードするポリヌクレオチドが提供される。また、そのような標的化脂質粒子を含有するプロデューサー細胞及び組成物ならびに標的化脂質粒子を作製及び使用する方法が提供される。

【選択図】図4A



## 【特許請求の範囲】

## 【請求項 1】

標的化脂質粒子であって、

( a ) 内腔を包囲する脂質二重層と、

( b ) ヘニパウウイルス F タンパク質分子またはその生物学的に活性な部分と、

( c ) ( i ) ヘニパウウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分及び ( i i ) シングルドメイン抗体 ( s d A b ) 可変ドメインを含む、標的化エンベロープタンパク質であって、前記 s d A b 可変ドメインが、前記 G タンパク質もしくはその生物学的に活性な部分の C 末端に結合されており、及び / または前記 s d A b が、ペプチドリンカーを介して前記 G タンパク質もしくはその生物学的に活

10

性な部分に結合されており、前記 s d A b が、標的細胞の細胞表面分子に結合する、前記標的化エンベロープタンパク質と

を含み、前記 F タンパク質分子またはその生物学的に活性な部分及び前記標的化エンベロープタンパク質が、脂質二重層に埋め込まれている、前記標的化脂質粒子。

## 【請求項 2】

前記細胞表面分子が、タンパク質、グリカン、脂質または低分子量分子である、請求項 1 に記載の標的化脂質粒子。

## 【請求項 3】

前記標的細胞が、腫瘍浸潤リンパ球、T 細胞、新生物または腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系 ( C N S ) 細胞、造血幹細胞 ( H S C )、肝臓細胞または完全に分化した細胞からなる群から選択される、請求項 1 または請求項 2 に記載の標的化脂質粒子。

20

## 【請求項 4】

前記標的細胞が、C D 3 + T 細胞、C D 4 + T 細胞、C D 8 + T 細胞、肝細胞、造血幹細胞、C D 3 4 + 造血幹細胞、C D 1 0 5 + 造血幹細胞、C D 1 1 7 + 造血幹細胞、C D 1 0 5 + 内皮細胞、B 細胞、C D 2 0 + B 細胞、C D 1 9 + B 細胞、がん細胞、C D 1 3 3 + がん細胞、E p C A M + がん細胞、C D 1 9 + がん細胞、H e r 2 / N e u + がん細胞、G l u A 2 + ニューロン、G l u A 4 + ニューロン、N K G 2 D + ナチュラルキラー細胞、S L C 1 A 3 + 星状細胞、S L C 7 A 1 0 + 脂肪細胞、または C D 3 0 + 肺上皮細胞からなる群から選択される、請求項 1 ~ 3 のいずれか一項に記載の標的化脂質粒子。

30

## 【請求項 5】

前記シングルドメイン抗体が、肝細胞上に存在する抗原またはその一部に結合する、請求項 1 ~ 4 のいずれか一項に記載の標的化脂質粒子。

## 【請求項 6】

前記細胞表面分子または抗原が、A S G R 1、A S G R 2 及び T M 4 S F からなる群から選択される、請求項 1 ~ 5 のいずれか一項に記載の標的化脂質粒子。

## 【請求項 7】

前記シングルドメイン抗体が、T 細胞上に存在する抗原またはその一部に結合する、請求項 1 ~ 4 のいずれか一項に記載の標的化脂質粒子。

40

## 【請求項 8】

前記細胞表面分子または抗原が、C D 8 または C D 4 である、請求項 1 ~ 4 及び 7 のいずれか一項に記載の標的化脂質粒子。

## 【請求項 9】

前記細胞表面分子または抗原が、L D L - R である、請求項 1 ~ 4 のいずれか一項に記載の標的化脂質粒子。

## 【請求項 10】

標的化脂質粒子であって、

( a ) 内腔を包囲する脂質二重層と、

( b ) ヘニパウウイルス F タンパク質分子またはその生物学的に活性な部分と、

50

(c)(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び(ii)結合ドメインを含む、標的化エンベロープタンパク質であって、前記結合ドメインが、前記Gタンパク質またはその生物学的に活性な部分のC末端に結合されており、前記結合ドメインが、ASGR1、ASGR2、TM4SF5、CD8、CD4及びLDL-Rからなる群から選択される細胞表面分子と結合する、前記標的化エンベロープタンパク質と

を含み、前記Fタンパク質分子またはその生物学的に活性な部分及び前記標的化エンベロープタンパク質が、前記脂質二重層に埋め込まれている、前記標的化脂質粒子。

【請求項11】

前記結合ドメインが、リンカーを介して前記Gタンパク質に結合されている、請求項10に記載の標的化脂質粒子。 10

【請求項12】

前記リンカーが、ペプチドリinkerである、請求項11に記載の標的化脂質粒子。

【請求項13】

レンチウイルスベクターである、請求項1~12のいずれか一項に記載の標的化脂質粒子。

【請求項14】

(a)ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、

(b)(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び(ii)結合ドメインを含む、標的化エンベロープタンパク質であって、前記結合ドメインが、前記Gタンパク質またはその生物学的に活性な部分のC末端に結合されており、前記結合ドメインが、CD4と結合する、前記標的化エンベロープタンパク質と、 20

(c)キメラ抗原受容体(CAR)をコードする核酸を含むカーゴであって、前記CARが、

(i)CD19と結合する細胞外抗原結合ドメインであって、任意に前記細胞外抗原結合ドメインがscFvである、前記細胞外抗原結合ドメイン、

(ii)膜貫通ドメイン、及び

(iii)CD3ゼータシグナル伝達ドメインを含む細胞内シグナル伝達領域

を含む、前記カーゴと

を含む、レンチウイルスベクター。 30

【請求項15】

前記CARの前記細胞内シグナル伝達領域が、4-1BB共刺激シグナル伝達ドメインをさらに含む、請求項14に記載のレンチウイルスベクター。

【請求項16】

前記レンチウイルスベクターが、前記CARをコードする前記核酸をT細胞に送達することが可能であり、任意に前記T細胞が、対象においてin vivoである、請求項14または請求項15に記載のレンチウイルスベクター。

【請求項17】

(a)ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、 40

(b)(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び(ii)結合ドメインを含む、標的化エンベロープタンパク質であって、前記結合ドメインが、前記Gタンパク質またはその生物学的に活性な部分のC末端に結合されており、前記結合ドメインが、ASGR1、ASGR2及びTM4SF5からなる群から選択される細胞表面分子と結合する、前記標的化エンベロープタンパク質と

を含む、レンチウイルスベクター。

【請求項18】

肝細胞を標的とすることが可能である、請求項17に記載のレンチウイルスベクター。

【請求項19】

肝細胞への送達のための外因性物質をさらに含み、及び/または前記レンチウイルスベクターが前記外因性物質を肝細胞に送達することが可能であり、任意に前記肝細胞が対象において *in vivo* である、請求項 17 及び請求項 18 のいずれか一項に記載のレンチウイルスベクター。

【請求項 20】

前記結合ドメインが、リンカーを介して前記 G タンパク質に結合されており、任意に前記リンカーが、ペプチドリinkerである、請求項 14 ~ 19 のいずれか一項に記載のレンチウイルスベクター。

【請求項 21】

前記結合ドメインが、シングルドメイン抗体であるか、または一本鎖可変断片 ( *scFv* ) である、請求項 10 ~ 13 のいずれか一項に記載の標的化脂質粒子または請求項 14 ~ 20 のいずれか一項に記載のレンチウイルスベクター。 10

【請求項 22】

前記ペプチドリinkerが、最大で 65 アミノ酸の長さ、任意に 2 ~ 65 アミノ酸の長さを含む、請求項 1 ~ 9、12、及び 13 のいずれか一項に記載の標的化脂質粒子または請求項 20 もしくは請求項 21 に記載のレンチウイルスベクター。

【請求項 23】

前記ペプチドリinkerが、GS、GGS、GGGS ( 配列番号 43 )、GGGG ( 配列番号 41 ) もしくはそれらの組み合わせを含むフレキシブルリンカーであるか、または前記ペプチドリinkerが、*n* が 1 ~ 10 である ( GGS ) *n*、*n* が 1 ~ 10 である ( GGGG ) *n* ( 配列番号 42 )、もしくは *n* が 1 ~ 6 である ( GGGGG ) *n* ( 配列番号 27 ) を含む、請求項 1 ~ 9、12、13 及び 22 のいずれか一項に記載の標的化脂質粒子または請求項 20 ~ 22 のいずれか一項に記載のレンチウイルスベクター。 20

【請求項 24】

前記 G タンパク質またはその生物学的に活性な部分が、野生型ニパウイルス G ( NiV - G ) タンパク質もしくはヘンドラウイルス G タンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項 1 ~ 13、22 及び 23 のいずれか一項に記載の標的化脂質粒子または請求項 14 ~ 23 のいずれか一項に記載のレンチウイルスベクター。

【請求項 25】 30

前記 G タンパク質またはその生物学的に活性な部分が、野生型 NiV - G タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項 1 ~ 13 及び 22 ~ 24 のいずれかに記載の標的化脂質粒子または請求項 14 ~ 24 のいずれか一項に記載のレンチウイルスベクター。

【請求項 26】

前記 NiV - G タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分が、配列番号 9、配列番号 28 もしくは配列番号 44 に示されるアミノ酸配列、または配列番号 9、配列番号 28 もしくは配列番号 44 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、少なくとも 84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を含む、請求項 24 または請求項 25 に記載の標的化脂質粒子またはレンチウイルスベクター。 40

【請求項 27】

前記 NiV - G タンパク質が、 50



切断型でありかつ前記野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの最大で 40 個の連続アミノ酸残基を欠く生物学的に活性な部分

である、請求項 24 ~ 26 のいずれか一項に記載の標的化脂質粒子またはレンチウイルス粒子。

【請求項 28】

前記 NiV - G タンパク質が、野生型 NiV - G の N 末端で切断された生物学的に活性な部分であり、配列番号 10 ~ 15、35 ~ 40 もしくは 45 ~ 50 のいずれかに示される配列、または配列番号 10 ~ 15、35 ~ 40 もしくは 45 ~ 50 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、少なくとも 84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、請求項 24 ~ 27 のいずれか一項に記載の標的化脂質粒子またはレンチウイルス粒子。

10

【請求項 29】

前記 NiV - G タンパク質が、前記野生型 NiV - G タンパク質の N 末端におけるまたはその近くの 5 アミノ酸切断、前記野生型 NiV - G タンパク質の N 末端におけるまたはその近くの 10 アミノ酸切断、前記 N 末端におけるまたはその近くの 15 アミノ酸切断、前記野生型 NiV - G タンパク質の N 末端におけるまたはその近くの 20 アミノ酸切断、前記野生型 NiV - G タンパク質の N 末端におけるまたはその近くの 25 アミノ酸切断、前記野生型 NiV - G タンパク質の N 末端におけるまたはその近くの 30 アミノ酸切断、または前記野生型 NiV - G タンパク質の N 末端におけるまたはその近くの 34 アミノ酸切断からなる群から選択される前記野生型 NiV - G の N 末端におけるまたはその近くの切断を有する生物学的に活性な部分であり、任意に前記野生型 NiV - G タンパク質が、配列番号 9、配列番号 28 または配列番号 44 に示される、請求項 24 ~ 28 のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクター。

20

30

【請求項 30】

前記 NiV - G タンパク質が、前記野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 34 アミノ酸切断を有する生物学的に活性な部分である、請求項 24 ~ 29 のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクター。

【請求項 31】

前記 G タンパク質またはその生物学的に活性な部分が、エフリン B2 またはエフリン B3 に対する減少した結合を示す変異体 NiV - G タンパク質である、請求項 1 ~ 13、及び 22 ~ 30 のいずれか一項に記載の標的化脂質粒子または請求項 14 ~ 30 のいずれか一項に記載のレンチウイルスベクター。

40

【請求項 32】

前記変異体 NiV - G タンパク質が、配列番号 28 に示される付番を参照して E501A、W504A、Q530A 及び E533A からなる群から選択されるアミノ酸置換に対応する 1 つ以上のアミノ酸置換を含む、請求項 31 に記載の標的化脂質粒子またはレンチウイルスベクター。

【請求項 33】

前記変異体 NiV - G タンパク質が、配列番号 28 に示される付番を参照してアミノ酸置換 E501A、W504A、Q530A 及び E533A

50

を含む、請求項 3 1 または請求項 3 2 に記載の標的化脂質粒子またはレンチウイルスベクター。

【請求項 3 4】

前記変異体 N i V - G タンパク質または前記生物学的に活性な部分が、配列番号 1 6 に示されるアミノ酸配列、または配列番号 1 6 と 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列を有する、請求項 3 1 ~ 3 3 のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクター。

10

【請求項 3 5】

前記 F タンパク質またはその生物学的に活性な部分が、野生型ニパウイルス F ( N i V - F ) タンパク質もしくはヘンドラウイルス F タンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項 1 ~ 1 3 及び 2 2 ~ 3 4 のいずれか一項に記載の標的化脂質粒子または請求項 1 4 ~ 3 4 のいずれか一項に記載のレンチウイルスベクター。

20

【請求項 3 6】

前記 F タンパク質またはその生物学的に活性な部分が、野生型 N i V - F タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項 1 ~ 1 3、及び 2 2 ~ 3 5 のいずれか一項に記載の標的化脂質粒子または請求項 1 4 ~ 3 5 のいずれか一項に記載のレンチウイルスベクター。

【請求項 3 7】

前記 N i V - F タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分が、配列番号 2 に示されるアミノ酸配列、または配列番号 2 と 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列を含む、請求項 1 ~ 1 3 及び 2 2 ~ 3 6 のいずれか一項に記載の標的化脂質粒子または請求項 1 4 ~ 3 6 のいずれか一項に記載のレンチウイルスベクター。

30

40

【請求項 3 8】

前記 N i V - F タンパク質が、

i ) 前記野生型 N i V - F タンパク質 ( 配列番号 2 ) の C 末端におけるもしくはその近くの 2 0 アミノ酸切断、及び / または

i i ) N 結合型グリコシル化部位における点変異

を含むその生物学的に活性な部分である、請求項 1 ~ 1 3、及び 2 2 ~ 3 7 のいずれか一項に記載の標的化脂質粒子または請求項 1 4 ~ 3 7 のいずれか一項に記載のレンチウイルスベクター。

【請求項 3 9】

前記 N i V - F タンパク質または前記生物学的に活性な部分が、配列番号 5 もしくは配

50

列番号 7 に示される配列、または配列番号 5 もしくは配列番号 7 と 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列を有する、請求項 3 8 に記載の標的化脂質粒子またはレンチウイルスベクター。 10

【請求項 4 0】

前記 NiV - F タンパク質が、前記野生型 NiV - F タンパク質（配列番号 2）の C 末端におけるまたはその近くの 2 2 アミノ酸切断を有するその生物学的に活性な部分である、請求項 1 ~ 1 3、及び 2 2 ~ 3 7 のいずれか一項に記載の標的化脂質粒子または請求項 1 4 ~ 3 7 のいずれか一項に記載のレンチウイルスベクター。

【請求項 4 1】

前記 NiV - F タンパク質または前記生物学的に活性な部分が、配列番号 8 もしくは配列番号 2 3 に示される配列、または配列番号 8 もしくは配列番号 2 3 と 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有する配列をコードするヌクレオチドの配列によってコードされるアミノ酸配列を有する、請求項 4 0 に記載の標的化脂質粒子またはレンチウイルスベクター。 20 30

【請求項 4 2】

前記 F タンパク質が、配列番号 2 3 に示される配列を含み、前記 G タンパク質が、配列番号 1 6 に示される配列を含む、請求項 1 ~ 1 3、及び 2 2 ~ 4 1 のいずれか一項に記載の標的化脂質粒子または請求項 1 4 ~ 4 1 のいずれか一項に記載のレンチウイルスベクター。

【請求項 4 3】

前記脂質二重層が、レトロウイルスまたはレトロウイルス様粒子を生成するために使用される宿主細胞の膜に由来する、請求項 1 ~ 1 3、及び 2 2 ~ 4 2 のいずれか一項に記載の標的化脂質粒子。

【請求項 4 4】

前記宿主細胞が、CHO 細胞、BHK 細胞、MDCK 細胞、C3H 10T1/2 細胞、FLY 細胞、Psi-2 細胞、BOSC 23 細胞、PA317 細胞、WEHI 細胞、COS 細胞、BSC 1 細胞、BSC 40 細胞、BMT 10 細胞、VERO 細胞、W138 細胞、MRC5 細胞、A549 細胞、HT1080 細胞、293 細胞、293T 細胞、B-50 細胞、3T3 細胞、NIH3T3 細胞、HepG2 細胞、Saos-2 細胞、Huh7 細胞、HeLa 細胞、W163 細胞、211 細胞、及び 211A 細胞からなる群から選択される、請求項 4 3 に記載の標的化脂質粒子。 40

【請求項 4 5】

前記脂質二重層が、ウイルスエンベロープであるか、またはそれを含む、請求項 1 ~ 1 3、及び 2 2 ~ 4 4 のいずれか一項に記載の標的化脂質粒子。 50

## 【請求項 46】

前記 F タンパク質分子及び前記 G タンパク質以外の 1 つ以上のウイルス成分を含む、請求項 1 ~ 13 及び 22 ~ 45 のいずれか一項に記載の標的化脂質粒子。

## 【請求項 47】

前記 1 つ以上のウイルス成分が、レトロウイルスからのものである、請求項 46 に記載の標的化脂質粒子。

## 【請求項 48】

前記レトロウイルスが、レンチウイルスまたはレンチウイルス様粒子である、請求項 43 ~ 47 のいずれか一項に記載の標的化脂質粒子。

## 【請求項 49】

外因性物質をさらに含む、請求項 1 ~ 13 及び 22 ~ 48 のいずれか一項に記載の標的化脂質粒子または請求項 14 ~ 42 のいずれか一項に記載のレンチウイルスベクター。

## 【請求項 50】

前記外因性物質が、前記内腔に存在する、請求項 49 に記載の標的化脂質粒子またはレンチウイルスベクター。

## 【請求項 51】

前記外因性物質がタンパク質または核酸であり、任意に前記核酸が、DNA または RNA である、請求項 49 または請求項 50 に記載の標的化脂質粒子またはレンチウイルスベクター。

## 【請求項 52】

前記外因性物質が、前記標的細胞への送達のためのカーゴをコードする核酸である、請求項 49 ~ 51 のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクター。

## 【請求項 53】

前記外因性物質が、治療剤または診断剤をコードする、請求項 49 ~ 52 のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクター。

## 【請求項 54】

前記外因性物質が膜タンパク質をコードし、任意に前記膜タンパク質が、疾患または病態によって発現するまたはそれに関連する細胞を標的とするための抗原受容体である、請求項 49 ~ 53 のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクター。

## 【請求項 55】

前記膜タンパク質がキメラ抗原受容体 (CAR) である、請求項 53 に記載の標的化脂質粒子またはレンチウイルスベクター。

## 【請求項 56】

前記標的細胞が T 細胞である、請求項 52 ~ 55 のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクター。

## 【請求項 57】

前記外因性物質が、遺伝子欠損、任意に前記標的細胞における遺伝子欠損を修正するためのペイロード遺伝子を含む核酸であり、任意に前記遺伝子欠損が、肝臓細胞または肝細胞に関連する、請求項 49 ~ 53 のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクター。

## 【請求項 58】

(i) ヘニバウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分と (ii) シングルドメイン抗体 (sdAb) 可変ドメインとをコードする核酸配列を含む、ポリヌクレオチドであって、前記 sdAb 可変ドメインが、前記 G タンパク質またはその生物学的に活性な部分の C 末端に結合されている、前記ポリヌクレオチド。

## 【請求項 59】

前記シングルドメイン抗体が、標的細胞上に存在する細胞表面分子と結合する、請求項 58 に記載のポリヌクレオチド。

## 【請求項 60】

10

20

30

40

50

前記細胞表面分子が、タンパク質、グリカン、脂質または低分子量分子である、請求項 59 に記載のポリヌクレオチド。

【請求項 61】

前記標的細胞が、腫瘍浸潤リンパ球、T細胞、新生物もしくは腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系(CNS)細胞、造血幹細胞(HSC)、肝臓細胞または完全に分化した細胞からなる群から選択される、請求項 59 または請求項 60 に記載のポリヌクレオチド。

【請求項 62】

前記標的細胞が、CD3+ T細胞、CD4+ T細胞、CD8+ T細胞、肝細胞、造血幹細胞、CD34+ 造血幹細胞、CD105+ 造血幹細胞、CD117+ 造血幹細胞、CD105+ 内皮細胞、B細胞、CD20+ B細胞、CD19+ B細胞、がん細胞、CD133+ がん細胞、EpCAM+ がん細胞、CD19+ がん細胞、Her2/Neu+ がん細胞、GluA2+ ニューロン、GluA4+ ニューロン、NKGD2+ ナチュラルキラー細胞、SLC1A3+ 星状細胞、SLC7A10+ 脂肪細胞、またはCD30+ 肺上皮細胞からなる群から選択される、請求項 59 ~ 61 のいずれか一項に記載のポリヌクレオチド。

10

【請求項 63】

前記細胞表面分子または抗原が、ASGR1、ASGR2、TM4SF5、CD8、CD4、及び低密度リポタンパク質受容体(LDL-R)からなる群から選択される、請求項 59 ~ 62 のいずれか一項に記載のポリヌクレオチド。

20

【請求項 64】

(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分と(ii)ASGR1、ASGR2、TM4SF5、CD4、CD8、及び低密度リポタンパク質受容体(LDL-R)からなる群から選択される細胞表面分子に結合する結合ドメインとをコードする核酸配列を含む、ポリヌクレオチド。

【請求項 65】

前記結合ドメインが、シングルドメイン抗体(sdAb)であるか、または一本鎖可変断片(scFv)である、請求項 64 に記載のポリヌクレオチド。

【請求項 66】

前記核酸配列が、第1の核酸配列であり、前記ポリヌクレオチドが、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする第2の核酸配列をさらに含む、請求項 58 ~ 65 のいずれか一項に記載のポリヌクレオチド。

30

【請求項 67】

前記ポリヌクレオチドが、IRESまたは前記第1及び第2の核酸配列間の連結ペプチドをコードする配列を含み、任意に、前記連結ペプチドが、自己切断性ペプチドまたはリボソームスキッピングを引き起こすペプチド、任意にT2Aペプチドである、請求項 66 に記載のポリヌクレオチド。

【請求項 68】

前記核酸の発現、任意に前記第1の核酸配列及び前記第2の核酸配列の発現をコントロールするために機能可能に連結された少なくとも1つのプロモーターをさらに含む、請求項 58 ~ 67 のいずれか一項に記載のポリヌクレオチド。

40

【請求項 69】

前記sdAb可変ドメインまたは前記結合ドメインが、コードされるペプチドリンカーを介して前記Gタンパク質に結合されている、請求項 58 ~ 68 のいずれか一項に記載のポリヌクレオチド。

【請求項 70】

前記コードされるペプチドリンカーが、最大で65アミノ酸の長さ、任意に2~65アミノ酸の長さを含む、請求項 69 に記載のポリヌクレオチド。

【請求項 71】

前記コードされるペプチドリンカーが、GS、GGS、GGGS(配列番号43)、

50

G G G G G S (配列番号 4 1) 及びそれらの組み合わせを含み、または前記コードされるペプチドリンカーが、 $n$  が 1 ~ 10 である (G G S)  $n$ 、 $n$  が 1 ~ 10 である (G G G G S)  $n$  (配列番号 4 2)、もしくは  $n$  が 1 ~ 4 である (G G G G G S)  $n$  (配列番号 2 7) を含む、請求項 6 9 または請求項 7 0 に記載のポリヌクレオチド。

【請求項 7 2】

前記コードされる G タンパク質が、野生型ニパウイルス G (N i V - G) タンパク質もしくはヘンドラウイルス G タンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分であり、任意に前記バリエーションが、ネイティブ結合パートナーについて減少した結合を示すそのバリエーションである、請求項 5 8 ~ 7 1 のいずれか一項に記載のポリヌクレオチド。

10

【請求項 7 3】

前記コードされる G タンパク質が、野生型 N i V - G タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項 5 8 ~ 7 2 のいずれか一項に記載のポリヌクレオチド。

【請求項 7 4】

前記 N i V - G タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分が、配列番号 9、配列番号 2 8 もしくは配列番号 4 4 に示されるアミノ酸配列を含み、または配列番号 9、配列番号 2 8 もしくは配列番号 4 4 と少なくとも 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列を含む、請求項 7 2 または請求項 7 3 に記載のポリヌクレオチド。

20

【請求項 7 5】

前記 N i V - G タンパク質が、

30

切断型でありかつ前記野生型 N i V - G タンパク質 (配列番号 9、配列番号 2 8 または配列番号 4 4) の N 末端におけるまたはその近くの最大で 4 0 個の連続アミノ酸残基を欠く生物学的に活性な部分である、請求項 7 2 ~ 7 4 のいずれか一項に記載のポリヌクレオチド。

【請求項 7 6】

前記 N i V - G タンパク質が、野生型 N i V - G の N 末端で切断された生物学的に活性な部分であり、配列番号 1 0 ~ 1 5、3 5 ~ 4 0 もしくは 4 5 ~ 5 0 のいずれかに示される配列、または配列番号 1 0 ~ 1 5、3 5 ~ 4 0 もしくは 4 5 ~ 5 0 と少なくとも 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列を含む、請求項 7 2 ~ 7 5 のいずれか一項に記載のポリヌクレオチド。

40

【請求項 7 7】

前記 N i V - G タンパク質が、前記野生型 N i V - G タンパク質の N 末端におけるまた

50

はその近くの5アミノ酸切断、前記野生型NiV-Gタンパク質のN末端におけるまたはその近くの10アミノ酸切断、前記N末端におけるまたはその近くの15アミノ酸切断、前記野生型NiV-Gタンパク質のN末端におけるまたはその近くの20アミノ酸切断、前記野生型NiV-Gタンパク質のN末端におけるまたはその近くの25アミノ酸切断、前記野生型NiV-Gタンパク質のN末端におけるまたはその近くの30アミノ酸切断、または前記野生型NiV-Gタンパク質のN末端におけるまたはその近くの34アミノ酸切断からなる群から選択される前記野生型NiV-GのN末端におけるまたはその近くの切断を有する生物学的に活性な部分であり、任意に前記野生型NiV-Gタンパク質が、配列番号9、配列番号28または配列番号44に示される、請求項72~76のいずれか一項に記載のポリヌクレオチド。

10

**【請求項78】**

前記NiV-Gタンパク質が、前記野生型NiV-Gタンパク質(配列番号9、配列番号28または配列番号44)のN末端におけるまたはその近くの34アミノ酸切断を有する生物学的に活性な部分である、請求項72~77のいずれか一項に記載のポリヌクレオチド。

**【請求項79】**

前記Gタンパク質が、エフリンB2またはエフリンB3に対する減少した結合を示す変異体NiV-Gタンパク質である、請求項58~78のいずれか一項に記載のポリヌクレオチド。

**【請求項80】**

前記変異体NiV-Gタンパク質が、配列番号28に示される付番を参照してE501A、W504A、Q530A及びE533Aからなる群から選択されるアミノ酸置換に対応する1つ以上のアミノ酸置換を含む、請求項79に記載のポリヌクレオチド。

20

**【請求項81】**

前記変異体NiV-Gタンパク質が、配列番号28に示される付番を参照してアミノ酸置換E501A、W504A、Q530A及びE533Aを含む、請求項79または請求項80に記載のポリヌクレオチド。

**【請求項82】**

前記変異体NiV-Gタンパク質が、配列番号16に示されるアミノ酸配列、または配列番号16と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む、請求項79~81のいずれか一項に記載のポリヌクレオチド。

30

**【請求項83】**

前記Fタンパク質またはその生物学的に活性な部分が、野生型ニパウイルスF(NiV-F)タンパク質もしくはヘンドラウイルスFタンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項66~82のいずれか一項に記載のポリヌクレオチド。

40

**【請求項84】**

前記Fタンパク質またはその生物学的に活性な部分が、野生型NiV-Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項66~83のいずれか一項に記載のポリヌクレオチド。

**【請求項85】**

前記NiV-Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活

50

性な部分が、配列番号 2 に示されるアミノ酸配列、または配列番号 2 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、請求項 83 または請求項 84 に記載のポリヌクレオチド。 10

【請求項 86】

前記 NiV - F タンパク質が、

i) 前記野生型 NiV - F タンパク質 (配列番号 2) の C 末端におけるもしくはその近くの 20 アミノ酸切断、及び / または

ii) N 結合型グリコシル化部位における点変異

を含むその生物学的に活性な部分である、請求項 85 または請求項 85 に記載のポリヌクレオチド。

【請求項 87】

前記 NiV - F タンパク質または前記生物学的に活性な部分が、配列番号 5 もしくは配列番号 7 に示される配列、または配列番号 5 もしくは配列番号 7 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、請求項 86 に記載のポリヌクレオチド。 20 30

【請求項 88】

前記 NiV - F タンパク質が、前記野生型 NiV - F タンパク質 (配列番号 2) の C 末端におけるまたはその近くの 22 アミノ酸切断を有するその生物学的に活性な部分である、請求項 85 または請求項 86 に記載のポリヌクレオチド。

【請求項 89】

前記 NiV - F タンパク質または前記生物学的に活性な部分が、配列番号 8 もしくは配列番号 23 に示される配列、または配列番号 8 もしくは配列番号 23 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有する配列をコードするヌクレオチドの配列によってコードされるアミノ酸配列を有する、請求項 88 に記載のポリヌクレオチド。 40

【請求項 90】

前記 F タンパク質が、配列番号 23 に示される配列を含み、前記 G タンパク質が、配列 50



番号 16 に示される配列を含む、請求項 66 ~ 89 のいずれか一項に記載のポリヌクレオチド。

【請求項 91】

請求項 58 ~ 90 のいずれか一項に記載のポリヌクレオチドを含む、ベクター。

【請求項 92】

前記ベクターが、哺乳動物ベクター、ウイルスベクターまたは人工染色体であり、任意に前記人工染色体が、細菌人工染色体 (BAC) である、請求項 91 に記載のベクター。

【請求項 93】

請求項 58 ~ 90 のいずれか一項に記載のポリヌクレオチドを含む、プラスミド。

【請求項 94】

レンチウイルス生成のためのタンパク質をコードする 1 つ以上の核酸をさらに含む、請求項 93 に記載のプラスミド。

【請求項 95】

請求項 58 ~ 90 のいずれか一項に記載のポリヌクレオチド、または請求項 91 もしくは請求項 92 に記載のベクター、または請求項 93 もしくは 94 に記載のプラスミドを含む、細胞。

【請求項 96】

ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分及びシングルドメイン抗体 (sdAb) 可変ドメインを含む標的化エンベロープタンパク質とを含む、標的化脂質粒子を作製する方法であって、

a) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分をコードする核酸と、標的化エンベロープタンパク質をコードする核酸とを含む、細胞を提供することであって、前記標的化エンベロープタンパク質が、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分及びシングルドメイン抗体 (sdAb) 可変ドメインを含む、前記提供すること、

b) 標的化脂質粒子の生成を可能とする条件下で前記細胞を培養すること、及び

c) 前記細胞から前記標的化脂質粒子を分離、濃縮、または精製し、それにより前記標的化脂質粒子を作製すること

を含む、前記方法。

【請求項 97】

偽型化レンチウイルスベクターを作製する方法であって、

a) レンチウイルスウイルス核酸 (複数可) と、ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分をコードする核酸と、標的化エンベロープタンパク質をコードする核酸とを含む、プロデューサー細胞を提供することであって、前記標的化エンベロープタンパク質が、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分及びシングルドメイン抗体を含む、前記提供すること、

b) 前記レンチウイルスベクターの生成を可能とする条件下で前記細胞を培養すること、及び

c) 前記細胞から前記レンチウイルスベクターを分離、濃縮、または精製し、それにより前記偽型化レンチウイルスベクターを作製すること

を含む、前記方法。

【請求項 98】

ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分及び結合ドメインを含む標的化エンベロープタンパク質とを含む、標的化脂質粒子を作製する方法であって、

a) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分をコードする核酸と、標的化エンベロープタンパク質をコードする核酸とを含む、細胞を提供することであって、前記標的化エンベロープタンパク質が、ヘニパウイルスエンベロープ結合糖タン

10

20

30

40

50

パク質 G ( G タンパク質 ) またはその生物学的に活性な部分及び結合ドメインを含み、前記結合ドメインが、

( i ) A S G R 1、A S G R 2、及び T M 4 S F 5、任意にヒト A S G R 1、ヒト A S G R 2 及びヒト A S G R 2 からなる群から選択される細胞表面分子と結合するか、

( i i ) C D 4 または C D 8、任意にヒト C D 4 またはヒト C D 8 からなる群から選択される細胞表面分子と結合するか、または

( i i i ) 低密度リポタンパク質受容体 ( L D L - R )、任意にヒト L D L - R である細胞表面分子と結合する、

前記提供すること、

b ) 標的化脂質粒子の生成を可能とする条件下で前記細胞を培養すること、及び

c ) 前記細胞から前記標的化脂質粒子を分離、濃縮、または精製し、それにより前記標的化脂質粒子を作製すること

を含む、前記方法。

【請求項 99】

偽型化レンチウイルスベクターを作製する方法であって、

a ) レンチウイルス核酸 ( 複数可 ) と、ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分をコードする核酸と、標的化エンベロープタンパク質をコードする核酸とを含む、プロデューサー細胞を提供することであって、前記標的化エンベロープタンパク質が、ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分及び結合ドメインを含み、前記結合ドメインが、

( i ) A S G R 1、A S G R 2、及び T M 4 S F 5、任意にヒト A S G R 1、ヒト A S G R 2 及びヒト A S G R 2 からなる群から選択される細胞表面分子と結合するか、

( i i ) C D 4 または C D 8、任意にヒト C D 4 またはヒト C D 8 からなる群から選択される細胞表面分子と結合するか、または

( i i i ) 低密度リポタンパク質受容体 ( L D L - R )、任意にヒト L D L - R である細胞表面分子と結合する、

前記提供すること、

b ) レンチウイルスベクターの生成を可能とする条件下で前記プロデューサー細胞を培養すること、及び

c ) 前記細胞から前記レンチウイルスベクターを分離、濃縮、または精製し、それにより前記偽型化レンチウイルスベクターを作製すること

を含む、前記方法。

【請求項 100】

前記結合ドメインが、シングルドメイン抗体であるか、または一本鎖可変断片 ( s c F v ) である、請求項 98 または請求項 99 に記載の方法。

【請求項 101】

ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、標的化エンベロープタンパク質とを含む、標的化脂質粒子を作製する方法であって、

a ) 請求項 58 ~ 90 のいずれか一項に記載のポリヌクレオチド、または請求項 91 もしくは請求項 92 に記載のベクター、または請求項 93 もしくは請求項 94 に記載のプラスミドを含む細胞を提供すること、

b ) 標的化脂質粒子の生成を可能とする条件下で前記細胞を培養すること、及び

c ) 前記細胞から前記標的化脂質粒子を分離、濃縮、または精製し、それにより前記標的化脂質粒子を作製すること

を含む、前記方法。

【請求項 102】

偽型化レンチウイルスベクターを作製する方法であって、

a ) レンチウイルス核酸 ( 複数可 ) と、請求項 58 ~ 90 のいずれか一項に記載のポリヌクレオチド、または請求項 91 もしくは請求項 92 に記載のベクター、または請求項 93 もしくは請求項 94 に記載のプラスミドとを含む、プロデューサー細胞を提供

10

20

30

40

50

すること、

b) 前記レンチウイルスベクターの生成を可能とする条件下で前記細胞を培養すること、及び

c) 前記細胞から前記レンチウイルスベクターを分離、濃縮、または精製し、それにより前記偽型化レンチウイルスベクターを作製すること

を含む、前記方法。

【請求項 103】

ステップ (b) の前に、ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分をコードするポリヌクレオチドを前記細胞に提供することをさらに含む、請求項 101 または請求項 102 に記載の方法。

10

【請求項 104】

前記細胞が、哺乳動物細胞である、請求項 96 ~ 103 のいずれか一項に記載の方法。

【請求項 105】

前記細胞が、ウイルス核酸、任意にレトロウイルス核酸またはレンチウイルス核酸を含むプロデューサー細胞であり、前記標的化脂質粒子が、ウイルス粒子またはウイルス様粒子、任意にレトロウイルス粒子またはレトロウイルス様粒子、任意にレンチウイルス粒子またはレンチウイルス様粒子である、請求項 96 ~ 104 のいずれか一項に記載の方法。

【請求項 106】

請求項 58 ~ 90 のいずれか一項に記載のポリヌクレオチド、または請求項 91 もしくは請求項 92 に記載のベクター、または請求項 93 もしくは請求項 94 に記載のプラスミドを含む、プロデューサー細胞。

20

【請求項 107】

ヘニパウイルス F タンパク質またはその生物学的に活性な部分をコードする核酸をさらに含む、請求項 106 に記載のプロデューサー細胞。

【請求項 108】

前記細胞がウイルス核酸をさらに含み、任意に前記ウイルス核酸が、レンチウイルス核酸である、請求項 106 または請求項 107 に記載のプロデューサー細胞。

【請求項 109】

(i) ウイルス核酸 (複数可) と、(ii) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分をコードする核酸と、(iii) ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分及びシングルドメイン抗体 (sdAb) 可変ドメインを含む標的化エンベロープタンパク質をコードする核酸とを含む、プロデューサー細胞であって、任意に前記ウイルス核酸 (複数可) が、レンチウイルス核酸である、前記プロデューサー細胞。

30

【請求項 110】

前記シングルドメイン抗体が、標的細胞上に存在する細胞表面分子と結合する、請求項 109 に記載のプロデューサー細胞。

【請求項 111】

前記細胞表面分子が、タンパク質、グリカン、脂質または低分子量分子である、請求項 110 に記載のプロデューサー細胞。

40

【請求項 112】

前記標的細胞が、腫瘍浸潤リンパ球、T細胞、新生物または腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系 (CNS) 細胞、造血幹細胞 (HSC)、肝臓細胞または完全に分化した細胞からなる群から選択される、請求項 110 または請求項 111 に記載のプロデューサー細胞。

【請求項 113】

前記標的細胞が、CD3+ T細胞、CD4+ T細胞、CD8+ T細胞、肝細胞、造血幹細胞、CD34+ 造血幹細胞、CD105+ 造血幹細胞、CD117+ 造血幹細胞、CD105+ 内皮細胞、B細胞、CD20+ B細胞、CD19+ B細胞、がん細胞、CD133+ がん細胞、EpCAM+ がん細胞、CD19+ がん細胞、

50

Her2/Neu+ がん細胞、GluA2+ ニューロン、GluA4+ ニューロン、NKGD2+ ナチュラルキラー細胞、SLC1A3+ 星状細胞、SLC7A10+ 脂肪細胞、またはCD30+ 肺上皮細胞からなる群から選択される、請求項110~112のいずれか一項に記載のプロデューサー細胞。

【請求項114】

(i) ウイルス核酸(複数可)と、(ii) ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸と、(iii) ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び結合ドメインを含む標的化エンベロープタンパク質をコードする核酸とを含む、プロデューサー細胞であって、前記結合ドメインが、

10

(i) ASGR1、ASGR2、及びTM4SF5、任意にヒトASGR1、ヒトASGR2及びヒトASGR2からなる群から選択される細胞表面分子と結合するか、

(ii) CD4またはCD8、任意にヒトCD4またはヒトCD8からなる群から選択される細胞表面分子と結合するか、または

(iii) 低密度リポタンパク質受容体(LDL-R)、任意にヒトLDL-Rである細胞表面分子と結合し、

任意に前記ウイルス核酸(複数可)が、レンチウイルス核酸である、前記プロデューサー細胞。

【請求項115】

前記結合ドメインが、シングルドメイン抗体であるか、または一本鎖可変断片である、請求項114に記載のプロデューサー細胞。

20

【請求項116】

前記Fタンパク質またはその生物学的に活性な部分が、野生型ニパウイルスF(NiV-F)タンパク質もしくはヘンドラウイルスFタンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項109~115のいずれか一項に記載のプロデューサー細胞。

【請求項117】

前記Fタンパク質またはその生物学的に活性な部分が、野生型NiV-Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項109~116のいずれか一項に記載のプロデューサー。

30

【請求項118】

前記ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分が、

(i) 配列番号2、配列番号5、配列番号7、配列番号8または配列番号23に示される配列、

(ii) 配列番号2、配列番号5、配列番号7、配列番号8または配列番号23と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列

40

を含む、請求項109~117のいずれか一項に記載のプロデューサー細胞。

【請求項119】

前記コードされるGタンパク質が、野生型ニパウイルスG(NiV-G)タンパク質もしくはヘンドラウイルスGタンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分であり、任意に前記バリエーションが、ネイティブ結合パートナーについて減少した結合を示すそのバリエーションである、請求項109~118のいずれか

50

一項に記載のプロデューサー細胞。

【請求項 1 2 0】

前記コードされる G タンパク質が、野生型 NiV - G タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、請求項 1 0 9 ~ 1 1 9 のいずれか一項に記載のプロデューサー細胞。

【請求項 1 2 1】

前記ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分が、

( i ) 配列番号 9、配列番号 2 8 または配列番号 4 4 に示される配列、

( i i ) 配列番号 9、配列番号 2 8 または配列番号 4 4 と 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、少なくとも 9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列

を含む、請求項 1 0 9 ~ 1 2 0 のいずれか一項に記載のプロデューサー細胞。

【請求項 1 2 2】

前記ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分が、

( i ) 配列番号 1 0 ~ 1 6、3 5 ~ 4 0 または 4 5 ~ 5 1 に示される配列、

( i i ) 配列番号 1 0 ~ 1 6、3 5 ~ 4 0 または 4 5 ~ 5 1 と少なくとも 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、少なくとも 9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列

を含む、請求項 1 0 9 ~ 1 2 1 のいずれか一項に記載のプロデューサー細胞。

【請求項 1 2 3】

膜 ( 例えば、細胞膜 ) に同じエンベロープタンパク質が組み込まれているがそれが代替的標的化部位に融合されている参照プロデューサー細胞と比較して多い、前記標的化エンベロープタンパク質の膜 ( 例えば、細胞膜 ) 発現を、前記プロデューサー細胞が有し、任意に前記代替的標的化部位が、一本鎖可変断片 ( s c F v ) である、ならびに / または

平方ミクロン当たり少なくとも 2 0 のタンパク質 ( 例えば、少なくとも 5 0、1 0 0、2 0 0、5 0 0、1 0 0 0、2 0 0 0、5 0 0 0、もしくは 1 0、0 0 0 のタンパク質 ) である、前記プロデューサー細胞の膜 ( 例えば、細胞膜 ) 上の前記標的化エンベロープタンパク質の発現を、前記プロデューサー細胞が有する、及び / もしくは前記標的化エンベロープタンパク質が、前記プロデューサー細胞の総膜 ( 例えば、細胞膜 ) タンパク質の少なくとも 0 . 1 % ( 例えば、少なくとも 0 . 2 %、0 . 5 %、1 %、2 %、3 %、4 %、5 %、6 %、7 %、8 %、9 %、もしくは 1 0 % ) ( 例えば、総タンパク質量に対して ) を構成する、

請求項 1 0 9 ~ 1 2 2 のいずれか 1 項に記載のプロデューサー細胞。

【請求項 1 2 4】

請求項 96 ~ 105 のいずれか一項に記載の方法によってまたは請求項 106 ~ 122 のいずれか一項に記載のプロデューサー細胞から生成された、標的化脂質粒子または偽型化レンチウイルスベクター。

【請求項 125】

前記標的化脂質粒子が、同様の脂質二重層に同じエンベロープタンパク質が組み込まれているがそれが代替的標的化部位に融合されている参照脂質粒子と比較して多い前記標的化エンベロープタンパク質発現を有し、任意に前記代替的標的化部位が、一本鎖可変断片 (s c F v) である、請求項 1 ~ 13、21 ~ 57 及び 124 のいずれか一項に記載の標的化脂質粒子。

【請求項 126】

形質導入後の標的細胞における力価が、 $1 \times 10^6$  形質導入単位 (T U) / m L 以上、 $2 \times 10^6$  T U / m L 以上、 $3 \times 10^6$  T U / m L 以上、 $4 \times 10^6$  T U / m L 以上、 $5 \times 10^6$  T U / m L 以上、 $6 \times 10^6$  T U / m L 以上、 $7 \times 10^6$  T U / m L 以上、 $8 \times 10^6$  T U / m L 以上、 $9 \times 10^6$  T U / m L 以上、もしくは  $1 \times 10^7$  T U / m L 以上である、及び / または

前記標的化エンベロープタンパク質が、少なくとも約 (0.001、0.002、0.005、0.01、0.02、0.05、0.1、0.2 もしくは 0.5) の標的化エンベロープタンパク質 /  $\text{nm}^2$  の密度で前記標的化脂質粒子の表面上に存在する、請求項 1 ~ 13、21 ~ 57、124 及び 125 のいずれか一項に記載の標的化脂質粒子、請求項 13 ~ 42、49 ~ 57 及び 124 のいずれか一項に記載のレンチウイルスベクター。

【請求項 127】

請求項 1 ~ 57 及び 124 ~ 126 のいずれか一項に記載の複数の標的化脂質粒子または複数のレンチウイルスベクターを含む、組成物。

【請求項 128】

薬学的に許容可能な担体をさらに含む、請求項 127 に記載の組成物。

【請求項 129】

前記組成物における脂質粒子またはレンチウイルスベクターの集団の中で、50% もしくは約 50% 超、55% もしくは約 55% 超、60% もしくは約 60% 超、65% もしくは約 65% 超、70% もしくは約 70% 超、または 75% もしくは約 75% 超が、前記標的化エンベロープタンパク質について表面陽性である、及び / または

前記標的化エンベロープタンパク質が、少なくとも約 (0.001、0.002、0.005、0.01、0.02、0.05、0.1、0.2 もしくは 0.5) の標的化エンベロープタンパク質 /  $\text{nm}^2$  の平均密度で前記標的化脂質粒子の表面上に存在する、請求項 127 または請求項 128 に記載の組成物。

【請求項 130】

細胞に形質導入する方法であって、請求項 13 ~ 42、49 ~ 57 及び 124 のいずれか一項に記載のレンチウイルスベクターまたは請求項 13 ~ 42、49 ~ 57 及び 124 のいずれか一項に記載のレンチウイルスベクターもしくは複数のレンチウイルスベクターを含む組成物で細胞に形質導入することを含む、前記方法。

【請求項 131】

前記レンチウイルスベクターもしくは標的化脂質粒子の前記標的化エンベロープタンパク質が、CD4 を標的とし、前記細胞が、CD4 + 細胞であるか、または

前記レンチウイルスベクターの前記標的化エンベロープタンパク質が、CD8 を標的とし、前記細胞が、CD8 + 細胞であるか、または

前記レンチウイルスベクターの前記標的化エンベロープタンパク質が、ASGR1、ASGR2 もしくは TM4SF5 を標的とし、前記細胞が、肝細胞である、請求項 130 に記載の方法。

【請求項 132】

外因性物質を対象 (例えば、ヒト対象) に送達する方法であって、請求項 49 ~ 57、

10

20

30

40

50

125及び126のいずれか一項に記載の標的化脂質粒子またはレンチウイルスベクターを前記対象に投与することを含み、標的化脂質粒子またはレンチウイルスベクターが、前記外因性物質を含む、前記方法。

【請求項133】

外因性物質を対象（例えば、ヒト対象）に送達する方法であって、請求項127～129のいずれか一項に記載の組成物を前記対象に投与することを含み、複数の標的化脂質粒子またはレンチウイルスベクターが、前記外因性物質を含む、前記方法。

【請求項134】

キメラ抗原受容体（CAR）を細胞に送達する方法であって、細胞を、請求項14～16及び19～42のいずれか一項に記載のレンチウイルスベクターまたは請求項55もしくは請求項56に記載の標的化脂質粒子と接触させることを含み、前記レンチウイルスベクターまたは標的化脂質粒子が、前記CARをコードする核酸を含む、前記方法。

10

【請求項135】

キメラ抗原受容体（CAR）を細胞に送達する方法であって、細胞を、請求項127～129のいずれか一項に記載の組成物と接触させることを含み、複数のレンチウイルスベクターまたは標的化脂質粒子が、前記CARをコードする核酸を含む、前記方法。

【請求項136】

外因性物質を肝細胞に送達する方法であって、細胞を、請求項17～42のいずれか一項に記載のレンチウイルスベクターまたは請求項57に記載の標的化脂質粒子もしくはレンチウイルスベクターと接触させることを含み、前記方法。

20

【請求項137】

外因性物質を肝細胞に送達する方法であって、細胞を、請求項127～129のいずれか一項に記載の組成物と接触させることを含み、複数のレンチウイルスベクターまたは標的化脂質粒子が、前記肝細胞への送達のための外因性物質を含む、前記方法。

【請求項138】

前記接触させることが、レンチウイルスベクターまたは前記標的化脂質粒子で前記細胞に形質導入することを含み、任意に前記接触させることが、対象において*in vivo*である、請求項134～137のいずれか一項に記載の方法。

【請求項139】

対象（例えば、ヒト対象）における疾患または障害を処置する方法であって、請求項1～13、21～57、及び124～126のいずれか一項に記載の標的化脂質粒子、または請求項13～42、49～57、124及び126のいずれか一項に記載のレンチウイルスベクター、または請求項127～129のいずれか一項に記載の組成物を前記対象に投与することを含み、前記方法。

30

【請求項140】

哺乳動物細胞を標的化脂質粒子に融合させる方法であって、請求項1～13、21～57、及び124～126のいずれか一項に記載の標的化脂質粒子、請求項13～42、49～57、124及び126のいずれか一項に記載のレンチウイルスベクター、または請求項127～129のいずれか一項に記載の組成物を前記対象に投与することを含み、前記方法。

40

【請求項141】

前記哺乳動物細胞を前記標的化脂質粒子に前記融合させることが、外因性物質を対象（例えば、ヒト対象）に送達する、請求項140に記載の方法。

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、2020年3月31日に出願された「Targeted Lipid Particles and Compositions and Uses Thereof」と題する米国仮出願63/003,168、及び2021年2月26日に出願された

50

「Targeted Lipid Particles and Compositions and Uses Thereof」と題する米国仮出願 63 / 154 , 341 の利益を主張し、その各々の内容は、すべての目的のためにそれらの全体が参照により組み込まれる。

【0002】

配列表の参照による組み込み

本出願は、電子形式の配列表とともに出願されている。この配列表は、2021年3月29日に作成された2,071,683バイトのサイズである186152003640 SeqList.TXTと題されるファイルとして提供される。配列表の電子形式の情報は、その全体が参照により組み込まれる。

10

【0003】

分野

本開示は、内腔またはキャビティを包囲する脂質二重層と、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び結合ドメイン、例えば、シングルドメイン抗体(sdAb)可変ドメインを含有する標的化エンベロープタンパク質とを含有する脂質粒子に関する。本開示はまた、結合ドメイン、例えば、sdAb可変ドメインに融合または連結されたGタンパク質を含有する標的化エンベロープタンパク質、及びそのようなタンパク質をコードするポリヌクレオチドを提供する。また、そのような標的化脂質粒子を含有するプロデューサー細胞及び組成物ならびに標的化脂質粒子を作製及び使用する方法が開示される。

20

【背景技術】

【0004】

背景

ウイルス様粒子及びウイルスベクターを含む脂質粒子は、外因性物質の細胞への送達のために一般的に使用される。しかしながら、脂質粒子の所定の標的細胞への送達は、課題がある場合がある。レンチウイルスベクターの場合、宿主範囲は、異種エンベロープタンパク質で偽型化することによって変更され得る。所定の再標的化エンベロープタンパク質は、脂質粒子の表面上で十分に安定ではなく、発現もしない場合がある。所望の細胞を標的とするための、ウイルス様粒子及びウイルスベクターを含む改善された脂質粒子が必要とされている。提供される開示は、この必要性に対処する。

30

【発明の概要】

【0005】

概要

本明細書では、(a)内腔を包囲する脂質二重層と、(b)ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、(c)(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び(ii)シングルドメイン抗体(sdAb)可変ドメインを含む標的化エンベロープタンパク質とを含む、標的化脂質粒子であって、sdAb可変ドメインが、Gタンパク質またはその生物学的に活性な部分のC末端に結合されており、Fタンパク質分子またはその生物学的に活性な部分及び標的化エンベロープタンパク質が、脂質二重層に埋め込まれている、標的化脂質粒子が提供される。いくつかの実施形態では、シングルドメイン抗体は、リンカーを介してGタンパク質に結合されている。いくつかの実施形態では、リンカーは、ペプチドリンカーである。

40

【0006】

本明細書では、(a)内腔を包囲する脂質二重層と、(b)ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、(c)ペプチドリンカーを介してシングルドメイン抗体(sdAb)可変ドメインに結合したヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分を含む標的化エンベロープタンパク質とを含む、標的化脂質粒子であって、シングルドメイン抗体が、標的細胞の細胞

50



表面分子に結合し、Fタンパク質分子またはその生物学的に活性な部分及び標的化エンペロープタンパク質が、脂質二重層に埋め込まれている、標的化脂質粒子が提供される。いくつかの実施形態では、Fタンパク質分子またはその生物学的に活性な部分のN末端は、脂質二重層の外側に露出されている。いくつかの実施形態では、Gタンパク質のC末端は、脂質二重層の外側に露出されている。

【0007】

いくつかの実施形態では、シングルドメイン抗体は、標的細胞上に存在する細胞表面分子と結合する。いくつかの実施形態では、細胞表面分子は、タンパク質、グリカン、脂質または低分子量分子である。任意の実施形態のうちの一部では、シングルドメイン抗体は、標的細胞上に存在する抗原またはその一部と結合する。いくつかの実施形態では、抗原は、シングルドメイン抗体によって認識されるエピトープを含有する細胞表面分子または細胞表面分子の一部である。任意の実施形態のうちの一部では、標的細胞は、腫瘍浸潤リンパ球、T細胞、新生物または腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系(CNS)細胞、造血幹細胞(HSC)、肝臓細胞または完全に分化した細胞からなる群から選択される。いくつかの実施形態では、標的細胞は、CD3+ T細胞、CD4+ T細胞、CD8+ T細胞、肝細胞、造血幹細胞、CD34+ 造血幹細胞、CD105+ 造血幹細胞、CD117+ 造血幹細胞、CD105+ 内皮細胞、B細胞、CD20+ B細胞、CD19+ B細胞、がん細胞、CD133+ がん細胞、EpCAM+ がん細胞、CD19+ がん細胞、Her2/Neu+ がん細胞、GlucA2+ ニューロン、GlucA4+ ニューロン、NKGD2+ ナチュラルキラー細胞、SLC1A3+ 星状細胞、SLC7A10+ 脂肪細胞、またはCD30+ 肺上皮細胞からなる群から選択される。任意の実施形態のうちの一部では、標的細胞は、肝細胞である。任意の実施形態のうちの一部では、細胞表面分子または抗原は、ASGR1、ASGR2及びTM4SF5からなる群から選択される。

10

20

【0008】

任意の実施形態のうちの一部では、標的細胞は、T細胞である。任意の実施形態のうちの一部では、細胞表面分子または抗原は、CD8またはCD4である。

【0009】

任意の実施形態のうちの一部では、細胞表面分子または抗原は、LDL-Rである。

【0010】

本明細書では、(a)内腔を包囲する脂質二重層と、(b)ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、(c)(i)ヘニパウイルスエンペロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び(ii)結合ドメインを含む標的化エンペロープタンパク質とを含む、標的化脂質粒子であって、結合ドメインが、Gタンパク質またはその生物学的に活性な部分のC末端に結合されており、結合ドメインが、ASGR1、ASGR2、及びTM4SF5、任意にヒトASGR1、ヒトASGR2及びヒトASGR2からなる群から選択される細胞表面分子と結合し、Fタンパク質分子またはその生物学的に活性な部分及び標的化エンペロープタンパク質が、脂質二重層に埋め込まれている、標的化脂質粒子が提供される。

30

【0011】

本明細書では、(a)内腔を包囲する脂質二重層と、(b)ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、(c)(i)ヘニパウイルスエンペロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び(ii)結合ドメインを含む標的化エンペロープタンパク質とを含む、標的化脂質粒子であって、結合ドメインが、Gタンパク質またはその生物学的に活性な部分のC末端に結合されており、結合ドメインが、CD8及びCD4、任意にヒトCD8またはヒトCD4からなる群から選択される細胞表面分子と結合し、Fタンパク質分子またはその生物学的に活性な部分及び標的化エンペロープタンパク質が、脂質二重層に埋め込まれている、標的化脂質粒子が提供される。

40

【0012】

50

本明細書では、(a) 内腔を包囲する脂質二重層と、(b) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、(c) (i) ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分及び (ii) 結合ドメインを含む標的化エンベロープタンパク質とを含む、標的化脂質粒子であって、結合ドメインが、G タンパク質またはその生物学的に活性な部分の C 末端に結合されており、結合ドメインが、低密度リポタンパク質受容体 (LDL-R)、任意にヒト LDL-R である細胞表面分子と結合し、F タンパク質分子またはその生物学的に活性な部分及び標的化エンベロープタンパク質が、脂質二重層に埋め込まれている、標的化脂質粒子が提供される。

【0013】

10

任意の実施形態のうちの一部では、脂質粒子は、レンチウイルスベクターである。任意の実施形態のうちの一部では、結合ドメインは、リンカーを介して G タンパク質に結合されている。任意の実施形態のうちの一部では、リンカーは、ペプチドリンカーである。

【0014】

本明細書では、ASGR1、ASGR2 及び TM4SF5、任意にヒト ASGR1、ヒト ASGR2 及びヒト TM4SF5 からなる群から選択される細胞表面分子を標的とする結合ドメインを含むレンチウイルスベクターであって、再標的化ウイルス融合タンパク質で偽型化されており、前記再標的化ウイルス融合タンパク質が、(a) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、(b) ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分に結合した結合ドメインを含む標的化エンベロープタンパク質とを含む、レンチウイルスベクターが、提供される。

20

【0015】

本明細書では、CD8 及び CD4、任意にヒト CD8 及びヒト CD4 からなる群から選択される細胞表面分子を標的とする結合ドメインを含むレンチウイルスベクターであって、再標的化ウイルス融合タンパク質で偽型化されており、前記再標的化ウイルス融合タンパク質が、(a) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、(b) ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分に結合した結合ドメインを含む標的化エンベロープタンパク質とを含む、レンチウイルスベクターが、提供される。

30

【0016】

本明細書では、低密度リポタンパク質受容体 (LDL-R) を標的とする結合ドメインを含むレンチウイルスベクターであって、任意に LDL-R は、ヒト LDL-R であり、レンチウイルスベクターは、(a) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、(b) ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分に結合した結合ドメインを含む標的化エンベロープタンパク質とを含む再標的化ウイルス融合タンパク質で偽型化されている、レンチウイルスベクターが、提供される。

【0017】

任意の実施形態のうちの一部では、結合ドメインは、G タンパク質またはその生物学的に活性な部分の C 末端に結合されている。

40

【0018】

本明細書では、

(a) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、

(b) (i) ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分及び (ii) 結合ドメインを含む、標的化エンベロープタンパク質であって、結合ドメインが、G タンパク質またはその生物学的に活性な部分の C 末端に結合されており、結合ドメインが、CD4 と結合する、標的化エンベロープタンパク質と、

(c) キメラ抗原受容体 (CAR) をコードする核酸を含むカーゴであって、CAR が

50

( i ) 細胞外抗原 ( 例えば、 C D 1 9 または B C M A ) と結合する細胞外抗原結合ドメイン、ならびに

( i i ) C D 3 ゼータシグナル伝達ドメイン及び、任意に 4 - 1 B B または C D 2 8 共刺激シグナル伝達ドメインを含む細胞内シグナル伝達領域

を含む、カーゴと

を含む、レンチウイルスベクターが提供される。いくつかの実施形態では、 C A R の細胞外抗原結合ドメインは、 s c F v である。

【 0 0 1 9 】

任意の実施形態のうちの一部では、レンチウイルスベクターは、 C A R をコードする核酸を T 細胞に送達することが可能である。いくつかの実施形態では、 T 細胞は、対象において *i n v i v o* である。

【 0 0 2 0 】

本明細書では、( a ) ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、( b ) ( i ) ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分及び ( i i ) 結合ドメインを含む標的化エンベロープタンパク質とを含む、レンチウイルスベクターであって、結合ドメインが、 G タンパク質またはその生物学的に活性な部分の C 末端に結合されており、結合ドメインが、 A S G R 1 と結合し、レンチウイルスベクターが、肝細胞を標的とすることが可能である、レンチウイルスベクターが提供される。任意の実施形態のうちの一部では、レンチウイルスベクターは、肝細胞への送達のための外因性物質をさらに含む。

【 0 0 2 1 】

任意の実施形態のうちの一部では、レンチウイルスベクターは外因性物質を肝細胞に送達することが可能であり、任意に肝細胞は対象において *i n v i v o* である。

【 0 0 2 2 】

任意の実施形態のうちの一部では、結合ドメインは、リンカーを介して G タンパク質に結合されている。任意の実施形態のうちの一部では、リンカーは、ペプチドリinker である。任意の実施形態のうちの一部では、結合ドメインは、シングルドメイン抗体である。任意の実施形態のうちの一部では、結合ドメインは、一本鎖可変断片 ( s c F v ) である。

【 0 0 2 3 】

任意の実施形態のうちの一部では、ペプチドリinker は、最大で 6 5 アミノ酸の長さを含む。任意の実施形態のうちの一部では、ペプチドリinker は、最大で 5 0 アミノ酸の長さを含む。任意の実施形態のうちの一部では、ペプチドリinker は、約 2 ~ 6 5 アミノ酸、 2 ~ 6 0 アミノ酸、 2 ~ 5 6 アミノ酸、 2 ~ 5 2 アミノ酸、 2 ~ 4 8 アミノ酸、 2 ~ 4 4 アミノ酸、 2 ~ 4 0 アミノ酸、 2 ~ 3 6 アミノ酸、 2 ~ 3 2 アミノ酸、 2 ~ 2 8 アミノ酸、 2 ~ 2 4 アミノ酸、 2 ~ 2 0 アミノ酸、 2 ~ 1 8 アミノ酸、 2 ~ 1 4 アミノ酸、 2 ~ 1 2 アミノ酸、 2 ~ 1 0 アミノ酸、 2 ~ 8 アミノ酸、 2 ~ 6 アミノ酸、 6 ~ 6 5 アミノ酸、 6 ~ 6 0 アミノ酸、 6 ~ 5 6 アミノ酸、 6 ~ 5 2 アミノ酸、 6 ~ 4 8 アミノ酸、 6 ~ 4 4 アミノ酸、 6 ~ 4 0 アミノ酸、 6 ~ 3 6 アミノ酸、 6 ~ 3 2 アミノ酸、 6 ~ 2 8 アミノ酸、 6 ~ 2 4 アミノ酸、 6 ~ 2 0 アミノ酸、 6 ~ 1 8 アミノ酸、 6 ~ 1 4 アミノ酸、 6 ~ 1 2 アミノ酸、 6 ~ 1 0 アミノ酸、 8 ~ 6 5 アミノ酸、 8 ~ 6 0 アミノ酸、 8 ~ 5 6 アミノ酸、 8 ~ 5 2 アミノ酸、 8 ~ 4 8 アミノ酸、 8 ~ 4 4 アミノ酸、 8 ~ 4 0 アミノ酸、 8 ~ 3 6 アミノ酸、 8 ~ 3 2 アミノ酸、 8 ~ 2 8 アミノ酸、 8 ~ 2 4 アミノ酸、 8 ~ 2 0 アミノ酸、 8 ~ 1 8 アミノ酸、 8 ~ 1 4 アミノ酸、 8 ~ 1 2 アミノ酸、 8 ~ 1 0 アミノ酸、 1 0 ~ 6 5 アミノ酸、 1 0 ~ 6 0 アミノ酸、 1 0 ~ 5 6 アミノ酸、 1 0 ~ 5 2 アミノ酸、 1 0 ~ 4 8 アミノ酸、 1 0 ~ 4 4 アミノ酸、 1 0 ~ 4 0 アミノ酸、 1 0 ~ 3 6 アミノ酸、 1 0 ~ 3 2 アミノ酸、 1 0 ~ 2 8 アミノ酸、 1 0 ~ 2 4 アミノ酸、 1 0 ~ 2 0 アミノ酸、 1 0 ~ 1 8 アミノ酸、 1 0 ~ 1 4 アミノ酸、 1 0 ~ 1 2 アミノ酸、 1 2 ~ 6 5 アミノ酸、 1 2 ~ 6 0 アミノ酸、 1 2 ~ 5 6 アミノ酸、 1 2 ~ 5 2 アミノ酸、 1 2

10

20

30

40

50

~ 4 8 アミノ酸、 1 2 ~ 4 4 アミノ酸、 1 2 ~ 4 0 アミノ酸、 1 2 ~ 3 6 アミノ酸、 1 2  
 ~ 3 2 アミノ酸、 1 2 ~ 2 8 アミノ酸、 1 2 ~ 2 4 アミノ酸、 1 2 ~ 2 0 アミノ酸、 1 2  
 ~ 1 8 アミノ酸、 1 2 ~ 1 4 アミノ酸、 1 4 ~ 6 5 アミノ酸、 1 4 ~ 6 0 アミノ酸、 1 4  
 ~ 5 6 アミノ酸、 1 4 ~ 5 2 アミノ酸、 1 4 ~ 4 8 アミノ酸、 1 4 ~ 4 4 アミノ酸、 1 4  
 ~ 4 0 アミノ酸、 1 4 ~ 3 6 アミノ酸、 1 4 ~ 3 2 アミノ酸、 1 4 ~ 2 8 アミノ酸、 1 4  
 ~ 2 4 アミノ酸、 1 4 ~ 2 0 アミノ酸、 1 4 ~ 1 8 アミノ酸、 1 8 ~ 6 5 アミノ酸、 1 8  
 ~ 6 0 アミノ酸、 1 8 ~ 5 6 アミノ酸、 1 8 ~ 5 2 アミノ酸、 1 8 ~ 4 8 アミノ酸、 1 8  
 ~ 4 4 アミノ酸、 1 8 ~ 4 0 アミノ酸、 1 8 ~ 3 6 アミノ酸、 1 8 ~ 3 2 アミノ酸、 1 8  
 ~ 2 8 アミノ酸、 1 8 ~ 2 4 アミノ酸、 1 8 ~ 2 0 アミノ酸、 2 0 ~ 6 5 アミノ酸、 2 0  
 ~ 6 0 アミノ酸、 2 0 ~ 5 6 アミノ酸、 2 0 ~ 5 2 アミノ酸、 2 0 ~ 4 8 アミノ酸、 2 0 10  
 ~ 4 4 アミノ酸、 2 0 ~ 4 0 アミノ酸、 2 0 ~ 3 6 アミノ酸、 2 0 ~ 3 2 アミノ酸、 2 0  
 ~ 2 8 アミノ酸、 2 0 ~ 2 6 アミノ酸、 2 0 ~ 2 4 アミノ酸、 2 4 ~ 6 5 アミノ酸、 2 4  
 ~ 6 0 アミノ酸、 2 4 ~ 5 6 アミノ酸、 2 4 ~ 5 2 アミノ酸、 2 4 ~ 4 8 アミノ酸、 2 4  
 ~ 4 4 アミノ酸、 2 4 ~ 4 0 アミノ酸、 2 4 ~ 3 6 アミノ酸、 2 4 ~ 3 2 アミノ酸、 2 4  
 ~ 3 0 アミノ酸、 2 4 ~ 2 8 アミノ酸、 2 8 ~ 6 5 アミノ酸、 2 8 ~ 6 0 アミノ酸、 2 8  
 ~ 5 6 アミノ酸、 2 8 ~ 5 2 アミノ酸、 2 8 ~ 4 8 アミノ酸、 2 8 ~ 4 4 アミノ酸、 2 8  
 ~ 4 0 アミノ酸、 2 8 ~ 3 6 アミノ酸、 2 8 ~ 3 4 アミノ酸、 2 8 ~ 3 2 アミノ酸、 3 2  
 ~ 6 5 アミノ酸、 3 2 ~ 6 0 アミノ酸、 3 2 ~ 5 6 アミノ酸、 3 2 ~ 5 2 アミノ酸、 3 2  
 ~ 4 8 アミノ酸、 3 2 ~ 4 4 アミノ酸、 3 2 ~ 4 0 アミノ酸、 3 2 ~ 3 8 アミノ酸、 3 2  
 ~ 3 6 アミノ酸、 3 6 ~ 6 5 アミノ酸、 3 6 ~ 6 0 アミノ酸、 3 6 ~ 5 6 アミノ酸、 3 6 20  
 ~ 5 2 アミノ酸、 3 6 ~ 4 8 アミノ酸、 3 6 ~ 4 4 アミノ酸、 3 6 ~ 4 0 アミノ酸、 4 0  
 ~ 6 5 アミノ酸、 4 0 ~ 6 0 アミノ酸、 4 0 ~ 5 6 アミノ酸、 4 0 ~ 5 2 アミノ酸、 4 0  
 ~ 4 8 アミノ酸、 4 0 ~ 4 4 アミノ酸、 4 4 ~ 6 5 アミノ酸、 4 4 ~ 6 0 アミノ酸、 4 4  
 ~ 5 6 アミノ酸、 4 4 ~ 5 2 アミノ酸、 4 4 ~ 4 8 アミノ酸、 4 8 ~ 6 5 アミノ酸、 4 8  
 ~ 6 0 アミノ酸、 4 8 ~ 5 6 アミノ酸、 4 8 ~ 5 2 アミノ酸、 5 0 ~ 6 5 アミノ酸、 5 0  
 ~ 6 0 アミノ酸、 5 0 ~ 5 6 アミノ酸、 5 0 ~ 5 2 アミノ酸、 5 4 ~ 6 5 アミノ酸、 5 4  
 ~ 6 0 アミノ酸、 5 4 ~ 5 6 アミノ酸、 5 8 ~ 6 5 アミノ酸、 5 8 ~ 6 0 アミノ酸、 また  
 は 6 0 ~ 6 5 アミノ酸を含む。任意の実施形態のうちの一部では、ペプチドリンカーは、  
 長さが 3、4、5、6、7、8、9、10、11、12、13、14、15、16、17  
 、18、19、20、21、22、23、24、25、26、27、28、29、30、 30  
 31、32、33、34、35、36、37、38、39、40、41、42、43、4  
 4、45、46、47、48、49、50、51、52、53、54、55、56、57  
 、58、59、60、61、62、63、64 または 6 5 アミノ酸であるポリペプチドを  
 含む。任意の実施形態のうちの一部では、ペプチドリンカーは、G S、G G S、G G G  
 S (配列番号 4 3)、G G G G S (配列番号 4 1) またはそれらの組み合わせを含むフ  
 レキシブルリンカーである。任意の実施形態のうちの一部では、ペプチドリンカーは、  
 (G G S) n (n は 1 ~ 1 0 である) を含む。任意の実施形態のうちの一部では、ペプチド  
 リンカーは、(G G G G S) n (配列番号 4 2) (n は 1 ~ 1 0 である) を含む。任意の  
 実施形態のうちの一部では、ペプチドリンカーは、(G G G G G S) n (配列番号 2 7)  
 (n は 1 ~ 6 である) を含む。 40

【0024】

任意の実施形態のうちの一部では、G タンパク質またはその生物学的に活性な部分は、  
 野生型ニパウイルス G (N i V - G) タンパク質またはヘンドラウイルス G タンパク質で  
 ある。任意の実施形態のうちの一部では、G タンパク質またはその生物学的に活性な部分  
 は、野生型 N i V - G タンパク質またはその機能的に活性なバリエーションもしくは生物学的  
 に活性な部分である。任意の実施形態のうちの一部では、変異体 N i V - G タンパク質ま  
 たはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、配列番号 9、配列  
 番号 2 8 または配列番号 4 4 と少なくとも 8 0 % もしくは約 8 0 %、少なくとも 8 1 % も  
 しくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3  
 %、少なくとも 8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくと 50

も 86% もしくは約 86%、少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

【0025】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、切断型でありかつ野生型 NiV-Gタンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの最大で 40 個の連続アミノ酸残基を欠く生物学的に活性な部分である。

10

【0026】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型 NiV-G の N 末端で切断された生物学的に活性な部分であり、配列番号 10 ~ 15、35 ~ 40 もしくは 45 ~ 50 のいずれかに示される配列、または配列番号 10 ~ 15、35 ~ 40 もしくは 45 ~ 50 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、少なくとも 84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する。

20

【0027】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型 NiV-Gタンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 5 アミノ酸切断を有する生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号 10 に示されるアミノ酸配列、または配列番号 10 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号 35 に示されるアミノ酸配列、または配列番号 35 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98

30

40

50

%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号45に示されるアミノ酸配列、または配列番号45と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。

10

【0028】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの10アミノ酸切断を有する生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号36に示されるアミノ酸配列、または配列番号36と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号11に示されるアミノ酸配列、または配列番号11と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号46に示されるアミノ酸配列、または配列番号46と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。

20

30

40

【0029】

任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分

50

は、野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 15 アミノ酸切断を有する。任意の実施形態のうちの一部では、NiV - G タンパク質または生物学的に活性な部分は、配列番号 12 に示されるアミノ酸配列、または配列番号 12 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV - G タンパク質または生物学的に活性な部分は、配列番号 37 に示されるアミノ酸配列、または配列番号 37 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV - G タンパク質または生物学的に活性な部分は、配列番号 47 に示されるアミノ酸配列、または配列番号 47 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する。

【0030】

任意の実施形態のうちの一部では、NiV - G タンパク質は、野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 20 アミノ酸切断を有する生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV - G タンパク質または生物学的に活性な部分は、配列番号 13 に示されるアミノ酸配列、または配列番号 13 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV - G タンパク質または生物学的に活性な部分は、配列番号 38 に示されるアミノ酸配列、または配列番

号38と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分

10

任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号48に示されるアミノ酸配列、または配列番号48と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。

20

【0031】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型NiV-Gタンパク質(配列番号9、配列番号28または配列番号44)のN末端におけるまたはその近くの25アミノ酸切断を有する生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質は、配列番号14に示されるアミノ酸配列、または配列番号14と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分

30

任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号39に示されるアミノ酸配列、または配列番号39と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号49に示されるアミノ酸配列、または配列番号49と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88

40

50



%もしくは約 88%、または少なくとも 89%もしくは約 89%、少なくとも 90%もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するアミノ酸配列を有する。

【0032】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの30アミノ酸切断を有する生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号15に示されるアミノ酸配列、または配列番号15と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号40に示されるアミノ酸配列、または配列番号40と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。

【0033】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの34アミノ酸切断を有する生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号22に示されるアミノ酸配列、または配列番号22と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、少なくとも89%もしくは約89%、90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号53に示されるアミノ酸配列、または配列番号53と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約8

4%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、少なくとも89%もしくは約89%、90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。

【0034】

任意の実施形態のうちの一部では、Gタンパク質、その生物学的に活性な部分は、エフリンB2またはエフリンB3に対する減少した結合を示す変異体NiV-Gタンパク質である機能的に活性なバリエーションである。

10

【0035】

任意の実施形態のうちの一部では、変異体NiV-Gタンパク質には、配列番号28に示される付番を参照してE501A、W504A、Q530A及びE533Aからなる群から選択されるアミノ酸置換に対応する1つ以上のアミノ酸置換が含まれる。任意の実施形態のうちの一部では、変異体NiV-Gタンパク質には、配列番号28に示される付番を参照してアミノ酸置換E501A、W504A、Q530A及びE533Aが含まれる。

【0036】

任意の実施形態のうちの一部では、変異体NiV-Gタンパク質または生物学的に活性な部分は、配列番号16に示されるアミノ酸配列、または配列番号16と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、変異体NiV-Gタンパク質または生物学的に活性な部分は、配列番号51に示されるアミノ酸配列、または配列番号51と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。

20

30

40

【0037】

任意の実施形態のうちの一部では、Fタンパク質またはその生物学的に活性な部分は、野生型ニパウイルスF(NiV-F)タンパク質もしくはヘンドラウイルスFタンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である。任意の実施形態のうちの一部では、Fタンパク質またはその生物学的に活性な部分は、野生型NiV-Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、配列番号2に示されるアミノ酸配列、または配列番号2と少なくとも80%もしくは約80%、少なくとも81%もしくは

50

は約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、少なくとも 84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

【0038】

任意の実施形態のうちの一部では、NiV-Fタンパク質は、野生型NiV-Fタンパク質（配列番号2）のC末端におけるまたはその近くの20アミノ酸切断を有するその生物学的に活性な部分である。

【0039】

任意の実施形態のうちの一部では、NiV-Fタンパク質または生物学的に活性な部分は、配列番号5に示される配列、または配列番号5と少なくとも80% もしくは約80%、少なくとも81% もしくは約81%、少なくとも82% もしくは約82%、少なくとも83% もしくは約83%、少なくとも84% もしくは約84%、少なくとも85% もしくは約85%、少なくとも86% もしくは約86%、少なくとも87% もしくは約87%、少なくとも88% もしくは約88%、少なくとも89% もしくは約89%、少なくとも90% もしくは約90%、少なくとも91% もしくは約91%、少なくとも92% もしくは約92%、少なくとも93% もしくは約93%、少なくとも94% もしくは約94%、少なくとも95% もしくは約95%、96% もしくは約96%、少なくとも97% もしくは約97%、少なくとも98% もしくは約98%、または少なくとも99% もしくは約99% の配列同一性を有するアミノ酸配列を有する。

【0040】

任意の実施形態のうちの一部では、NiV-Fタンパク質は、i) 野生型NiV-Fタンパク質（配列番号2）のC末端におけるまたはその近くの20アミノ酸切断、及びii) N結合型グリコシル化部位における点変異を含むその生物学的に活性な部分である。

【0041】

任意の実施形態のうちの一部では、NiV-Fタンパク質または生物学的に活性な部分は、配列番号7に示される配列、または配列番号7と少なくとも80% もしくは約80%、少なくとも81% もしくは約81%、少なくとも82% もしくは約82%、少なくとも83% もしくは約83%、少なくとも84% もしくは約84%、少なくとも85% もしくは約85%、少なくとも86% もしくは約86%、少なくとも87% もしくは約87%、少なくとも88% もしくは約88%、少なくとも89% もしくは約89%、少なくとも90% もしくは約90%、少なくとも91% もしくは約91%、少なくとも92% もしくは約92%、少なくとも93% もしくは約93%、少なくとも94% もしくは約94%、少なくとも95% もしくは約95%、96% もしくは約96%、少なくとも97% もしくは約97%、少なくとも98% もしくは約98%、または少なくとも99% もしくは約99% の配列同一性を有するアミノ酸配列を有する。

【0042】

任意の実施形態のうちの一部では、NiV-Fタンパク質は、野生型NiV-Fタンパク質（配列番号2）のC末端におけるまたはその近くの22アミノ酸切断を有するその生物学的に活性な部分である。

【0043】

任意の実施形態のうちの一部では、NiV-Fタンパク質または生物学的に活性な部分は、配列番号8に示される配列、または配列番号8と少なくとも80% もしくは約80%、少なくとも81% もしくは約81%、少なくとも82% もしくは約82%、少なくとも83% もしくは約83%、少なくとも84% もしくは約84%、少なくとも85% もしくは

10

20

30

40

50

は約 85%、少なくとも 86% もしくは約 86%、少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有する配列をコードするヌクレオチドの配列によってコードされるアミノ酸配列を有する。

【0044】

任意の実施形態のうちの一部では、NiV-Fタンパク質または生物学的に活性な部分は、配列番号23に示される配列、または配列番号23と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、F-タンパク質またはその生物学的に活性な部分は、F1サブユニットまたはその融合部分を含む。

10

20

【0045】

任意の実施形態のうちの一部では、Fタンパク質は、配列番号23に示される配列を含み、Gタンパク質は、配列番号16に示される配列を含む。

【0046】

任意の実施形態のうちの一部では、Fタンパク質は、配列番号23に示される配列からなり、もしくはそれから本質的になり、及び/またはGタンパク質は、配列番号16に示される配列からなり、もしくはそれから本質的になる。

【0047】

任意の実施形態のうちの一部では、F1サブユニットは、F0前駆体のタンパク質分解的に切断された部分である。任意の実施形態のうちの一部では、F1サブユニットは、配列番号4に示される配列、または配列番号4と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。

30

40

【0048】

任意の実施形態のうちの一部では、脂質二重層は、レトロウイルスまたはレトロウイルス様粒子を生成するために使用される宿主細胞の膜に由来する。任意の実施形態のうちの一部では、宿主細胞は、CHO細胞、BHK細胞、MDCK細胞、C3H 10T1/2細胞、FLY細胞、Psi-2細胞、BOSC 23細胞、PA317細胞、WEHI細胞、COS細胞、BSC 1細胞、BSC 40細胞、BMT 10細胞、VERO細胞、W138細胞、MRC5細胞、A549細胞、HT1080細胞、293細胞、293T細胞、B-50細胞、3T3細胞、NIH3T3細胞、HepG2細胞、Saos-2

50

細胞、Huh7細胞、HeLa細胞、W163細胞、211細胞、及び211A細胞からなる群から選択される。任意の実施形態のうちの一部では、宿主細胞は、293T細胞を含む。任意の実施形態のうちの一部では、脂質二重層は、ウイルスエンベロープであるか、またはそれを含む。任意の実施形態のうちの一部では、レトロウイルス様粒子は、複製欠損性である。

【0049】

任意の実施形態のうちの一部では、標的化脂質粒子は、Fタンパク質分子及びGタンパク質以外の1つ以上のウイルス成分を含む。任意の実施形態のうちの一部では、1つ以上のウイルス成分は、レトロウイルスからのものである。任意の実施形態のうちの一部では、レトロウイルスは、レンチウイルスである。任意の実施形態のうちの一部では、1つ以上のウイルス成分は、Gag、Pol、Rev及びTatのうちの一つ以上から選択されるウイルスパッケージングタンパク質を含む。任意の実施形態のうちの一部では、1つ以上のウイルス成分は、以下の核酸配列：5'LTR（例えば、U5を含み、機能性U3ドメインを欠く）、プサイパッケージング要素（プサイ）、セントラルポリプリン配列（cPPT）/セントラル終結配列（CTS）（例えば、DNAフラップ）、ポリAテール配列、転写後制御要素（例えば、WPRE）、Rev反応要素（RRE）、及び3'LTR（例えば、U5を含み、機能性U3を欠く）のうちの一つ以上（例えば、すべて）を含む。

10

【0050】

任意の実施形態のうちの一部では、標的化脂質粒子は、レンチウイルスベクターである。

20

【0051】

任意の実施形態のうちの一部では、標的化脂質粒子またはレンチウイルスベクターは、複製欠損性である。

【0052】

任意の実施形態のうちの一部では、標的化脂質粒子またはレンチウイルスベクターは、外因性物質をさらに含む。任意の実施形態のうちの一部では、標的化脂質粒子は、外因性物質をさらに含む。いくつかの実施形態では、レンチウイルスベクターは、外因性物質をさらに含む。

【0053】

任意の実施形態のうちの一部では、外因性物質は、内腔に存在する。任意の実施形態のうちの一部では、外因性物質は、タンパク質または核酸である。いくつかの実施形態では、核酸は、DNAまたはRNAである。

30

【0054】

任意の実施形態のうちの一部では、外因性物質は、標的細胞への送達のためのカーゴをコードする核酸である。任意の実施形態のうちの一部では、外因性物質は、治療剤または診断剤をコードする。

【0055】

任意の実施形態のうちの一部では、外因性物質は、膜タンパク質をコードする。いくつかの実施形態では、タンパク質は、疾患または病態によって発現するまたはそれに関連する細胞を標的とするための抗原受容体である。いくつかの実施形態では、膜タンパク質は、キメラ抗原受容体（CAR）である。いくつかの実施形態では、CARは、(i)細胞外抗原（例えば、CD19またはBCMA）と結合する細胞外抗原結合ドメインであって、任意に細胞外抗原結合ドメインは、scFvである、細胞外抗原結合ドメインと、(ii)膜貫通ドメインと、(iii)CD3ゼータシグナル伝達ドメイン及び、任意に共刺激シグナル伝達ドメイン、例えば、4-1BBまたはCD28共刺激シグナル伝達ドメインを含む細胞内シグナル伝達領域とを含む。いくつかの実施形態では、標的細胞は、T細胞である。いくつかの実施形態では、標的細胞上の細胞表面分子は、CD4またはCD8である。いくつかの実施形態では、結合ドメインは、CD4（例えば、ヒトCD4）と結合するscFvである。いくつかの実施形態では、結合ドメインは、CD4（例えば、ヒ

40

50

トCD4)と結合するシングルドメイン抗体である。いくつかの実施形態では、結合ドメインは、CD8(例えば、ヒトCD8)と結合するscFvである。いくつかの実施形態では、結合ドメインは、CD8(例えば、ヒトCD8)と結合するシングルドメイン抗体である。

【0056】

任意の実施形態のうちの一部では、外因性物質は、遺伝子欠損、任意に標的細胞における遺伝子欠損を修正するためのペイロード遺伝子を含む核酸である。いくつかの実施形態では、遺伝子欠損は、肝臓細胞または肝細胞に関連する。いくつかの実施形態では、標的細胞は、肝細胞である。いくつかの実施形態では、細胞表面分子は、ASGR1、ASGR2及びTM4SF5からなる群から選択される分子である。いくつかの実施形態では、結合ドメインは、ASGR1(例えば、ヒトASGR1)と結合するscFvである。いくつかの実施形態では、結合ドメインは、ASGR1(例えば、ヒトASGR1)と結合するシングルドメイン抗体である。いくつかの実施形態では、結合ドメインは、ASGR2(例えば、ヒトASGR2)と結合するscFvである。いくつかの実施形態では、結合ドメインは、ASGR2(例えば、ヒトASGR2)と結合するシングルドメイン抗体である。いくつかの実施形態では、結合ドメインは、TM4SF5(例えば、ヒトTM4SF5)と結合するscFvである。いくつかの実施形態では、結合ドメインは、TM4SF5(例えば、ヒトTM4SF5)と結合するシングルドメイン抗体である。

10

【0057】

任意の実施形態のうちの一部では、シングルドメイン抗体は、標的細胞上に存在する細胞表面分子と結合する。任意の実施形態のうちの一部では、細胞表面分子は、タンパク質、グリカン、脂質または低分子量分子である。任意の実施形態のうちの一部では、標的細胞は、腫瘍浸潤リンパ球、T細胞、新生物または腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系(CNS)細胞、造血幹細胞(HSC)、肝臓細胞または完全に分化した細胞からなる群から選択される。任意の実施形態のうちの一部では、標的細胞は、CD3+T細胞、CD4+T細胞、CD8+T細胞、肝細胞、造血幹細胞、CD34+造血幹細胞、CD105+造血幹細胞、CD117+造血幹細胞、CD105+内皮細胞、B細胞、CD20+B細胞、CD19+B細胞、がん細胞、CD133+がん細胞、EpCAM+がん細胞、CD19+がん細胞、Her2/Neu+がん細胞、GlucA2+ニューロン、GlucA4+ニューロン、NKGD2+ナチュラルキラー細胞、SLC1A3+星状細胞、SLC7A10+脂肪細胞、またはCD30+肺上皮細胞からなる群から選択される。

20

30

【0058】

任意の実施形態のうちの一部では、シングルドメイン抗体は、標的細胞上に存在する抗原またはその一部と結合する。任意の実施形態のうちの一部では、細胞表面分子または抗原は、ASGR1、ASGR2及びTM4SF5からなる群から選択される。いくつかの実施形態では、抗原またはその一部は、ヒトASGR1である。いくつかの実施形態では、抗原またはその一部は、ヒトASGR2である。いくつかの実施形態では、抗原またはその一部は、ヒトTM4SF5である。

【0059】

本明細書では、(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分と(ii)ASGR1、ASGR2、及びTM4SF5からなる群から選択される細胞表面分子と結合する結合ドメインとをコードする核酸配列を含む、ポリヌクレオチドが提供される。いくつかの実施形態では、細胞表面分子は、ヒトASGR1である。いくつかの実施形態では、細胞表面分子は、ヒトASGR2である。いくつかの実施形態では、細胞表面分子は、ヒトTM4SF5である。任意の実施形態のうちの一部では、細胞表面分子または抗原は、CD8またはCD4である。

40

【0060】

本明細書では、(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分と(ii)CD4及びCD8からなる群から選択され

50

る細胞表面分子と結合する結合ドメインとをコードする、核酸配列が提供される。いくつかの実施形態では、細胞表面分子は、ヒトCD4である。いくつかの実施形態では、細胞表面分子は、ヒトCD8である。いくつかの実施形態では、細胞表面分子または抗原は、低密度リポタンパク質受容体(LDL-R)である。いくつかの実施形態では、細胞表面分子または抗原は、ヒトLDL-Rである。

【0061】

本明細書では、(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分と(ii)低密度リポタンパク質受容体(LDL-R)と結合する結合ドメインとをコードする核酸配列を含む、ポリヌクレオチドが提供される。いくつかの実施形態では、結合ドメインは、ヒトLDL-Rと結合する。任意の実施形態のうちの一部では、結合ドメインは、シングルドメイン抗体(sdAb)である。任意の実施形態のうちの一部では、結合ドメインは、一本鎖可変断片(scFv)である。

10

【0062】

本明細書では、(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分と(ii)シングルドメイン抗体(sdAb)可変ドメインとをコードする核酸配列を含む、ポリヌクレオチドであって、sdAb可変ドメインが、Gタンパク質またはその生物学的に活性な部分のC末端に結合されている、ポリヌクレオチドが提供される。任意の実施形態のうちの一部では、ポリヌクレオチドは、(iii)ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸配列をさらに含む。

20

【0063】

いくつかの実施形態では、核酸配列は、第1の核酸配列であり、ポリヌクレオチドは、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする第2の核酸配列をさらに含む。いくつかの実施形態では、ポリヌクレオチドは、IRESまたは第1及び第2の核酸配列間の連結ペプチドをコードする配列を含む。いくつかの実施形態では、連結ペプチドは、自己切断性ペプチドまたはリボソームスキッピングを引き起こすペプチド、任意にT2Aペプチドである。

【0064】

任意の実施形態のうちの一部では、ポリヌクレオチドには、核酸の発現をコントロールするために機能可能に連結された少なくとも1つのプロモーターが含まれる。任意の実施形態のうちの一部では、プロモーターは、第1の核酸配列及び第2の核酸配列の発現をコントロールするように機能可能に連結される。任意の実施形態のうちの一部では、プロモーターは、構成的プロモーターである。任意の実施形態のうちの一部では、プロモーターは、誘導性プロモーターである。

30

【0065】

任意の実施形態のうちの一部では、sdAb可変ドメインは、コードされるペプチドリンカーを介してGタンパク質に結合されている。いくつかの実施形態では、結合ドメインは、コードされるペプチドリンカーを介してGタンパク質に結合されている。任意の実施形態のうちの一部では、コードされるペプチドリンカーは、最大で25アミノ酸の長さを含む。任意の実施形態のうちの一部では、コードされるペプチドリンカーは、最大で65アミノ酸の長さを含む。任意の実施形態のうちの一部では、コードされるペプチドリンカーは、約2~65アミノ酸、2~60アミノ酸、2~56アミノ酸、2~52アミノ酸、2~48アミノ酸、2~44アミノ酸、2~40アミノ酸、2~36アミノ酸、2~32アミノ酸、2~28アミノ酸、2~24アミノ酸、2~20アミノ酸、2~18アミノ酸、2~14アミノ酸、2~12アミノ酸、2~10アミノ酸、2~8アミノ酸、2~6アミノ酸、6~65アミノ酸、6~60アミノ酸、6~56アミノ酸、6~52アミノ酸、6~48アミノ酸、6~44アミノ酸、6~40アミノ酸、6~36アミノ酸、6~32アミノ酸、6~28アミノ酸、6~24アミノ酸、6~20アミノ酸、6~18アミノ酸、6~14アミノ酸、6~12アミノ酸、6~10アミノ酸、6~8アミノ酸、8~65アミノ酸、8~60アミノ酸、8~56アミノ酸、8~52アミノ酸、8~48アミノ酸

40

50

、 8 ~ 4 4 アミノ酸、 8 ~ 4 0 アミノ酸、 8 ~ 3 6 アミノ酸、 8 ~ 3 2 アミノ酸、 8 ~ 2 8 アミノ酸、 8 ~ 2 4 アミノ酸、 8 ~ 2 0 アミノ酸、 8 ~ 1 8 アミノ酸、 8 ~ 1 4 アミノ酸、 8 ~ 1 2 アミノ酸、 8 ~ 1 0 アミノ酸、 1 0 ~ 6 5 アミノ酸、 1 0 ~ 6 0 アミノ酸、 1 0 ~ 5 6 アミノ酸、 1 0 ~ 5 2 アミノ酸、 1 0 ~ 4 8 アミノ酸、 1 0 ~ 4 4 アミノ酸、 1 0 ~ 4 0 アミノ酸、 1 0 ~ 3 6 アミノ酸、 1 0 ~ 3 2 アミノ酸、 1 0 ~ 2 8 アミノ酸、 1 0 ~ 2 4 アミノ酸、 1 0 ~ 2 0 アミノ酸、 1 0 ~ 1 8 アミノ酸、 1 0 ~ 1 4 アミノ酸、 1 0 ~ 1 2 アミノ酸、 1 2 ~ 6 5 アミノ酸、 1 2 ~ 6 0 アミノ酸、 1 2 ~ 5 6 アミノ酸、 1 2 ~ 5 2 アミノ酸、 1 2 ~ 4 8 アミノ酸、 1 2 ~ 4 4 アミノ酸、 1 2 ~ 4 0 アミノ酸、 1 2 ~ 3 6 アミノ酸、 1 2 ~ 3 2 アミノ酸、 1 2 ~ 2 8 アミノ酸、 1 2 ~ 2 4 アミノ酸、 1 2 ~ 2 0 アミノ酸、 1 2 ~ 1 8 アミノ酸、 1 2 ~ 1 4 アミノ酸、 1 4 ~ 6 5 アミノ酸、 1 4 ~ 6 0 アミノ酸、 1 4 ~ 5 6 アミノ酸、 1 4 ~ 5 2 アミノ酸、 1 4 ~ 4 8 アミノ酸、 1 4 ~ 4 4 アミノ酸、 1 4 ~ 4 0 アミノ酸、 1 4 ~ 3 6 アミノ酸、 1 4 ~ 3 2 アミノ酸、 1 4 ~ 2 8 アミノ酸、 1 4 ~ 2 4 アミノ酸、 1 4 ~ 2 0 アミノ酸、 1 4 ~ 1 8 アミノ酸、 1 8 ~ 6 5 アミノ酸、 1 8 ~ 6 0 アミノ酸、 1 8 ~ 5 6 アミノ酸、 1 8 ~ 5 2 アミノ酸、 1 8 ~ 4 8 アミノ酸、 1 8 ~ 4 4 アミノ酸、 1 8 ~ 4 0 アミノ酸、 1 8 ~ 3 6 アミノ酸、 1 8 ~ 3 2 アミノ酸、 1 8 ~ 2 8 アミノ酸、 1 8 ~ 2 4 アミノ酸、 1 8 ~ 2 0 アミノ酸、 2 0 ~ 6 5 アミノ酸、 2 0 ~ 6 0 アミノ酸、 2 0 ~ 5 6 アミノ酸、 2 0 ~ 5 2 アミノ酸、 2 0 ~ 4 8 アミノ酸、 2 0 ~ 4 4 アミノ酸、 2 0 ~ 4 0 アミノ酸、 2 0 ~ 3 6 アミノ酸、 2 0 ~ 3 2 アミノ酸、 2 0 ~ 2 8 アミノ酸、 2 0 ~ 2 6 アミノ酸、 2 0 ~ 2 4 アミノ酸、 2 4 ~ 6 5 アミノ酸、 2 4 ~ 6 0 アミノ酸、 2 4 ~ 5 6 アミノ酸、 2 4 ~ 5 2 アミノ酸、 2 4 ~ 4 8 アミノ酸、 2 4 ~ 4 4 アミノ酸、 2 4 ~ 4 0 アミノ酸、 2 4 ~ 3 6 アミノ酸、 2 4 ~ 3 2 アミノ酸、 2 4 ~ 3 0 アミノ酸、 2 4 ~ 2 8 アミノ酸、 2 8 ~ 6 5 アミノ酸、 2 8 ~ 6 0 アミノ酸、 2 8 ~ 5 6 アミノ酸、 2 8 ~ 5 2 アミノ酸、 2 8 ~ 4 8 アミノ酸、 2 8 ~ 4 4 アミノ酸、 2 8 ~ 4 0 アミノ酸、 2 8 ~ 3 6 アミノ酸、 2 8 ~ 3 4 アミノ酸、 2 8 ~ 3 2 アミノ酸、 3 2 ~ 6 5 アミノ酸、 3 2 ~ 6 0 アミノ酸、 3 2 ~ 5 6 アミノ酸、 3 2 ~ 5 2 アミノ酸、 3 2 ~ 4 8 アミノ酸、 3 2 ~ 4 4 アミノ酸、 3 2 ~ 4 0 アミノ酸、 3 2 ~ 3 8 アミノ酸、 3 2 ~ 3 6 アミノ酸、 3 6 ~ 6 5 アミノ酸、 3 6 ~ 6 0 アミノ酸、 3 6 ~ 5 6 アミノ酸、 3 6 ~ 5 2 アミノ酸、 3 6 ~ 4 8 アミノ酸、 3 6 ~ 4 4 アミノ酸、 3 6 ~ 4 0 アミノ酸、 4 0 ~ 6 5 アミノ酸、 4 0 ~ 6 0 アミノ酸、 4 0 ~ 5 6 アミノ酸、 4 0 ~ 5 2 アミノ酸、 4 0 ~ 4 8 アミノ酸、 4 0 ~ 4 4 アミノ酸、 4 4 ~ 6 5 アミノ酸、 4 4 ~ 6 0 アミノ酸、 4 4 ~ 5 6 アミノ酸、 4 4 ~ 5 2 アミノ酸、 4 4 ~ 4 8 アミノ酸、 4 8 ~ 6 5 アミノ酸、 4 8 ~ 6 0 アミノ酸、 4 8 ~ 5 6 アミノ酸、 4 8 ~ 5 2 アミノ酸、 5 0 ~ 6 5 アミノ酸、 5 0 ~ 6 0 アミノ酸、 5 0 ~ 5 6 アミノ酸、 5 0 ~ 5 2 アミノ酸、 5 4 ~ 6 5 アミノ酸、 5 4 ~ 6 0 アミノ酸、 5 4 ~ 5 6 アミノ酸、 5 8 ~ 6 5 アミノ酸、 5 8 ~ 6 0 アミノ酸、 または 6 0 ~ 6 5 アミノ酸を含む。

【 0 0 6 6 】

任意の実施形態のうちの一部では、コードされるペプチドリンカーは、長さが 3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64 または 65 アミノ酸であるポリペプチドを含む。任意の実施形態のうちの一部では、コードされるペプチドリンカーは、GS、GGS、GGGS (配列番号 43)、GGGG (配列番号 41) 及びそれらの組み合わせを含む。任意の実施形態のうちの一部では、コードされるペプチドリンカーは、(GGS)n (n は 1 ~ 10 である) を含む。任意の実施形態のうちの一部では、コードされるペプチドリンカーは、(GGGS)n (配列番号 42) (n は 1 ~ 10 である) を含む。任意の実施形態のうちの一部では、コードされるペプチドリンカーは、(GGGGGS)n (配列番号 27) (n は 1 ~ 4 である) を含む。任意の実施形態のうちの一部では、Gタンパク質をコードする配列は、野生型ニパウイルス G (NiV-G) タンパク質もしくはヘンドラウ

10

20

30

40

50



イルス G タンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である。いくつかの実施形態では、バリエーションは、ネイティブ結合パートナーについて減少した結合を示すそのバリエーションである。任意の実施形態のうちの一部では、G タンパク質をコードする核酸配列は、野生型ニパウイルス G ( N i V - G ) タンパク質もしくはヘンドラウイルス G タンパク質であるか、またはネイティブ結合パートナーについて減少した結合を示すそのバリエーションである。いくつかの実施形態では、コードされる G タンパク質は、野生型 N i V - G タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である。任意の実施形態のうちの一部では、G タンパク質をコードする核酸配列は、野生型 N i V - G タンパク質である。任意の実施形態のうちの一部では、G タンパク質をコードする核酸配列は、エフリン B 2 またはエフリン B 3 に対する減少した結合を示す変異体 N i V - G タンパク質である。

10

【 0 0 6 7 】

任意の実施形態のうちの一部では、N i V - G タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、配列番号 9、配列番号 28 もしくは配列番号 44 に示されるアミノ酸配列を含み、または配列番号 9、配列番号 28 もしくは配列番号 44 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、N i V - G タンパク質は、切断型でありかつ野生型 N i V - G タンパク質 ( 配列番号 9、配列番号 28 または配列番号 44 ) の N 末端におけるまたはその近くの最大で 40 個の連続アミノ酸残基を欠く生物学的に活性な部分である。任意の実施形態のうちの一部では、N i V - G タンパク質は、野生型 N i V - G の N 末端で切断された生物学的に活性な部分であり、配列番号 10 ~ 15、35 ~ 40 もしくは 45 ~ 50 のいずれかに示される配列、または配列番号 10 ~ 15、35 ~ 40 もしくは 45 ~ 50 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

20

30

40

【 0 0 6 8 】

任意の実施形態のうちの一部では、N i V - G タンパク質は、野生型 N i V - G タンパク質 ( 配列番号 9、配列番号 28 または配列番号 44 ) の N 末端におけるまたはその近くの 5 アミノ酸切断を含む生物学的に活性な部分である。任意の実施形態のうちの一部では、N i V - G タンパク質または生物学的に活性な部分は、配列番号 10 に示されるアミノ酸配列、または配列番号 10 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90

50

%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号35に示されるアミノ酸配列、または配列番号35と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号45に示されるアミノ酸配列、または配列番号45と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

10

20

【0069】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの10アミノ酸切断を含む生物学的に活性な部分である。任意の実施形態のうちの一部では、変異体NiV-Gタンパク質または生物学的に活性な部分は、配列番号11に示されるアミノ酸配列、または配列番号11と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号36に示されるアミノ酸配列、または配列番号36と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%

30

40

50

98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号46に示されるアミノ酸配列、または配列番号46と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

10

## 【0070】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの15アミノ酸切断を含む生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号12に示されるアミノ酸配列、または配列番号12と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号37に示されるアミノ酸配列、または配列番号37と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号47に示されるアミノ酸配列、または配列番号47と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

20

30

40

## 【0071】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型NiV-Gタンバ

50

ク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 20 アミノ酸切断を含む生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号 13 に示されるアミノ酸配列、または配列番号 13 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号 38 に示されるアミノ酸配列、または配列番号 38 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号 48 に示されるアミノ酸配列、または配列番号 48 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

【0072】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型 NiV-Gタンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 25 アミノ酸切断を含む生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号 14 に示されるアミノ酸配列、または配列番号 14 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号 39 に示されるアミノ酸配列、または配列番号 3

9 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

任意の実施形態のうちの一部では、NiV - Gタンパク質または生物学的に活性な部分は、配列番号 49 に示されるアミノ酸配列、または配列番号 49 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

10

20

### 【0073】

任意の実施形態のうちの一部では、NiV - Gタンパク質は、野生型 NiV - Gタンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 30 アミノ酸切断を含む生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV - Gタンパク質または生物学的に活性な部分は、配列番号 15 に示されるアミノ酸配列、または配列番号 15 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV - Gタンパク質または生物学的に活性な部分は、配列番号 40 に示されるアミノ酸配列、または配列番号 40 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、NiV - Gタンパク質または生物学的に活性な部分は、配列番号 50 に示されるアミノ酸配列、または配列番号 50 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 8

30

40

50

7%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

【0074】

任意の実施形態のうちの一部では、NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの34アミノ酸切断を有する生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号22に示されるアミノ酸配列、または配列番号22と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Gタンパク質または生物学的に活性な部分は、配列番号53に示されるアミノ酸配列、または配列番号53と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。

【0075】

任意の実施形態のうちの一部では、Gタンパク質は、エフリンB2またはエフリンB3に対する減少した結合を示す変異体NiV-Gタンパク質である。任意の実施形態のうちの一部では、変異体NiV-Gタンパク質は、配列番号28に示される付番を参照してE501A、W504A、Q530A及びE533Aからなる群から選択されるアミノ酸置換に対応する1つ以上のアミノ酸置換を含む。任意の実施形態のうちの一部では、変異体NiV-Gタンパク質は、配列番号28に示される付番を参照してアミノ酸置換E501A、W504A、Q530A及びE533Aを含む。

【0076】

任意の実施形態のうちの一部では、変異体NiV-Gタンパク質は、i) N末端におけるまたはその近くの切断、及びii) E501A、W504A、Q530A及びE533Aからなる群から選択される点変異を含む。任意の実施形態のうちの一部では、変異体NiV-Gタンパク質は、配列番号16に示されるアミノ酸配列、または配列番号16と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少な

くとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、変異体NiV-Gタンパク質は、配列番号51に示されるアミノ酸配列、または配列番号51と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

10

## 【0077】

任意の実施形態のうちの一部では、Fタンパク質またはその生物学的に活性な部分は、野生型ニパウイルスF(NiV-F)タンパク質もしくはヘンドラウイルスFタンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である。任意の実施形態のうちの一部では、Fタンパク質またはその生物学的に活性な部分は、野生型NiV-Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、配列番号2に示されるアミノ酸配列、または配列番号2と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

20

30

## 【0078】

任意の実施形態のうちの一部では、NiV-Fタンパク質は、野生型NiV-Fタンパク質(配列番号2)のC末端におけるまたはその近くの20アミノ酸切断を有するその生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Fタンパク質または生物学的に活性な部分は、配列番号5に示される配列、または配列番号5と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、NiV-Fタンパク質は、i)野生型NiV-Fタンパク質(配列番号2)のC末端におけるまたはその近くの20アミノ酸切断、及びii)N結合型グリコシル化部位における点変異を含むその生物学的に活性な部分である。

40

## 【0079】

50

任意の実施形態のうちの一部では、NiV-Fタンパク質または生物学的に活性な部分は、配列番号7に示される配列、または配列番号7と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。 10

【0080】

任意の実施形態のうちの一部では、NiV-Fタンパク質は、野生型NiV-Fタンパク質（配列番号2）のC末端におけるまたはその近くの22アミノ酸切断を有するその生物学的に活性な部分である。任意の実施形態のうちの一部では、NiV-Fタンパク質または生物学的に活性な部分は、配列番号8に示される配列、または配列番号8と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有する配列をコードするヌクレオチドの配列によってコードされるアミノ酸配列を有する。 20

【0081】

任意の実施形態のうちの一部では、NiV-Fタンパク質は、配列番号23に示される配列、または配列番号23と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。任意の実施形態のうちの一部では、Fタンパク質は、配列番号23に示される配列を含み、Gタンパク質は、配列番号16に示される配列を含む。任意の実施形態のうちの一部では、Fタンパク質は、配列番号23に示される配列からなり、またはそれから本質的になり、Gタンパク質は、配列番号16に示される配列からなり、またはそれから本質的になる。 30 40

【0082】

本明細書では、本明細書に記載の実施形態のいずれかのポリヌクレオチドを含むベクターが提供される。任意の実施形態のうちの一部では、ベクターは、哺乳動物ベクター、ウイルスベクターまたは人工染色体であり、任意に人工染色体は、細菌人工染色体（BAC）である。

【0083】

本明細書では、本明細書に記載の実施形態のいずれかのポリヌクレオチドを含むプラスミドが提供される。任意の実施形態のうちの一部では、プラスミドは、レンチウイルス生 50



成のためのタンパク質をコードする1つ以上の核酸をさらに含む。

【0084】

本明細書では、本明細書に記載の実施形態のいずれかのポリヌクレオチドまたは本明細書に記載の実施形態のいずれかのベクター、または本明細書に記載の実施形態のいずれかのプラスミドを含む細胞が提供される。

【0085】

本明細書では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分及びシングルドメイン抗体（sdAb）可変ドメインを含む標的化エンベロープタンパク質とを含む、標的化脂質粒子を作製する方法であって、

a) ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸と、標的化エンベロープタンパク質をコードする核酸とを含む、細胞を提供することであって、標的化エンベロープタンパク質は、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分及びシングルドメイン抗体（sdAb）可変ドメインを含む、提供すること、

b) 標的化脂質粒子の生成を可能とする条件下で細胞を培養すること、及び

c) 細胞から標的化脂質粒子を分離、濃縮、または精製し、それにより標的化脂質粒子を作製することを含む方法が提供される。

【0086】

本明細書では、偽型化レンチウイルスベクターを作製する方法であって、

a) レンチウイルス核酸（複数可）と、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸と、標的化エンベロープタンパク質をコードする核酸とを含む、プロデューサー細胞を提供することであって、前記標的化エンベロープタンパク質は、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分及びシングルドメイン抗体を含む、提供すること、

b) レンチウイルスベクターの生成を可能とする条件下で細胞を培養すること、及び

c) 細胞からレンチウイルスベクターを分離、濃縮、または精製し、それにより偽型化レンチウイルスベクターを作製することを含む方法が提供される。

【0087】

任意の実施形態のうちの一部では、シングルドメイン抗体は、標的細胞上に存在する細胞表面分子と結合する。任意の実施形態のうちの一部では、細胞表面分子は、タンパク質、グリカン、脂質または低分子量分子である。任意の実施形態のうちの一部では、標的細胞は、腫瘍浸潤リンパ球、T細胞、新生物または腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系（CNS）細胞、造血幹細胞（HSC）、肝臓細胞または完全に分化した細胞からなる群から選択される。任意の実施形態のうちの一部では、標的細胞は、CD3+ T細胞、CD4+ T細胞、CD8+ T細胞、肝細胞、造血幹細胞、CD34+ 造血幹細胞、CD105+ 造血幹細胞、CD117+ 造血幹細胞、CD105+ 内皮細胞、B細胞、CD20+ B細胞、CD19+ B細胞、がん細胞、CD133+ がん細胞、EpCAM+ がん細胞、CD19+ がん細胞、Her2/Neu+ がん細胞、GLA2+ ニューロン、GLA4+ ニューロン、NKGD2+ ナチュラルキラー細胞、SLC1A3+ 星状細胞、SLC7A10+ 脂肪細胞、またはCD30+ 肺上皮細胞からなる群から選択される。任意の実施形態のうちの一部では、シングルドメイン抗体は、標的細胞上に存在する抗原またはその一部と結合する。

【0088】

本明細書では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分及び結合ドメインを含む標的化エンベロープタンパク質とを含む、標的化脂質粒子を作製する方法であって、

10

20

30

40

50

a) ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸と、標的化エンベロープタンパク質をコードする核酸とを含む、細胞を提供することであって、標的化エンベロープタンパク質は、ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分及び結合ドメインを含み、結合ドメインは、(i) ASGR1、ASGR2、及びTM4SF5、任意にヒトASGR1、ヒトASGR2及びヒトASGR2からなる群から選択される細胞表面分子と結合するか、(ii) CD4またはCD8、任意にヒトCD4またはヒトCD8からなる群から選択される細胞表面分子と結合するか、または(iii) 低密度リポタンパク質受容体(LDL-R)、任意にヒトLDL-Rである細胞表面分子と結合する、提供すること、

b) 標的化脂質粒子の生成を可能とする条件下で細胞を培養すること、及び

c) 細胞から標的化脂質粒子を分離、濃縮、または精製し、それにより標的化脂質粒子を作製すること

を含む方法が提供される。

10

#### 【0089】

本明細書では、偽型化レンチウイルスベクターを作製する方法であって、a) レンチウイルスウイルス核酸(複数可)と、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸と、標的化エンベロープタンパク質をコードする核酸とを含む、プロデューサー細胞を提供することであって、前記標的化エンベロープタンパク質は、ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分及び結合ドメインを含み、結合ドメインは、(i) ASGR1、ASGR2、及びTM4SF5、任意にヒトASGR1、ヒトASGR2及びヒトASGR2からなる群から選択される細胞表面分子と結合するか、(ii) CD4またはCD8、任意にヒトCD4またはヒトCD8からなる群から選択される細胞表面分子と結合するか、または(iii) 低密度リポタンパク質受容体(LDL-R)、任意にヒトLDL-Rである細胞表面分子と結合する、提供すること、b) レンチウイルスベクターの生成を可能とする条件下でプロデューサー細胞を培養すること、及びc) 細胞からレンチウイルスベクターを分離、濃縮、または精製し、それにより偽型化レンチウイルスベクターを作製することを含む方法が提供される。

20

#### 【0090】

任意の実施形態のうちの一部では、結合ドメインは、シングルドメイン抗体である。任意の実施形態のうちの一部では、結合ドメインは、一本鎖可変断片(scfv)である。任意の実施形態のうちの一部では、細胞表面分子は、ASGR1、ASGR2及びTM4SF5からなる群から選択される。任意の実施形態のうちの一部では、細胞表面分子は、CD8またはCD4である。任意の実施形態のうちの一部では、細胞表面分子は、LDL-Rである。

30

#### 【0091】

本明細書では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、標的化エンベロープタンパク質とを含む、標的化脂質粒子を作製する方法であって、

a) 本明細書で提供される実施形態のいずれかのポリヌクレオチド、本明細書に記載の実施形態のいずれかのベクター、または本明細書に記載の実施形態のいずれかのプラスミドを含む細胞を提供すること、

40

b) 標的化脂質粒子の生成を可能とする条件下で細胞を培養すること、及び

c) 細胞から標的化脂質粒子を分離、濃縮、または精製し、それにより標的化脂質粒子を作製すること

を含む方法が提供される。

#### 【0092】

本明細書では、偽型化レンチウイルスベクターを作製する方法であって、

a) レンチウイルスウイルス核酸(複数可)と、本明細書で列挙される実施形態のいずれかのポリヌクレオチドまたは本明細書で列挙される実施形態のいずれかのベクターとを含む、プロデューサー細胞を提供すること、

50

b) レンチウイルスベクターの生成を可能とする条件下で細胞を培養すること、及び  
 c) 細胞からレンチウイルスベクターを分離、濃縮、または精製し、それにより偽型化  
 レンチウイルスベクターを作製すること  
 を含む方法が提供される。任意の実施形態のうちの一部では、ステップ (b) の前に、方法  
 は、ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分をコードするポリ  
 リヌクレオチドを細胞に提供することをさらに含む。

## 【0093】

任意の実施形態のうちの一部では、細胞は、哺乳動物細胞である。

## 【0094】

任意の実施形態のうちの一部では、細胞は、ウイルス核酸を含むプロデューサー細胞で  
 ある。任意の実施形態のうちの一部では、ウイルス核酸は、レトロウイルス核酸またはレ  
 ンチウイルス核酸であり、標的化脂質粒子は、ウイルス粒子またはウイルス様粒子である  
 。任意の実施形態のうちの一部では、ウイルス粒子またはウイルス様粒子は、レトロウイル  
 ス粒子またはレトロウイルス様粒子である。いくつかの実施形態では、ウイルス粒子また  
 はウイルス様粒子は、レンチウイルス粒子またはレンチウイルス様粒子である。

10

## 【0095】

任意の実施形態のうちの一部では、ウイルス核酸 (複数可) は、ウイルス複製に関与す  
 る1つ以上の遺伝子を欠く。任意の実施形態のうちの一部では、ウイルス核酸は、G a g  
 、P o l、R e v 及びT a t のうちの1つ以上から選択されるウイルスパッケージングタ  
 ンパク質をコードする核酸を含む。任意の実施形態のうちの一部では、ウイルス核酸は、  
 以下の核酸配列: 5' L T R (例えば、U 5 を含み、機能性U 3 ドメインを欠く)、プサイ  
 パッケージング要素 (プサイ)、セントラルポリプリン配列 (c P P T) / セントラル  
 終結配列 (C T S) (例えば、D N A フラップ)、ポリA テール配列、転写後制御要素 (例  
 えば、W P R E)、R e v 反応要素 (R R E)、及び3' L T R (例えば、U 5 を含み  
 、機能性U 3 を欠く) のうちの1つ以上 (例えば、すべて) を含む。

20

## 【0096】

本明細書では、本明細書で列挙される実施形態のいずれかのポリヌクレオチドまたは本  
 明細書で列挙される実施形態のいずれかのベクター、または本明細書に記載の実施形態の  
 いずれかのプラスミドを含むプロデューサー細胞が提供される。

## 【0097】

任意の実施形態のうちの一部では、プロデューサー細胞は、ヘニパウイルス F タンパク  
 質またはその生物学的に活性な部分をコードする核酸をさらに含む。

30

## 【0098】

任意の実施形態のうちの一部では、細胞は、ウイルス核酸をさらに含む。任意の実施形  
 態のうちの一部では、ウイルス核酸は、レンチウイルス核酸である。本明細書では、( i )  
 ウイルス核酸 (複数可) と、( i i ) ヘニパウイルス F タンパク質分子またはその生物  
 学的に活性な部分をコードする核酸と、( i i i ) ヘニパウイルスエンベロープ結合糖タ  
 ンパク質 G ( G タンパク質) またはその生物学的に活性な部分及びシングルドメイン抗体  
 ( s d A b ) 可変ドメインを含む標的化エンベロープタンパク質をコードする核酸とを含  
 む、プロデューサー細胞であって、任意にウイルス核酸 (複数可) が、レンチウイルス核  
 酸である、プロデューサー細胞が提供される。任意の実施形態のうちの一部では、シ  
 ングルドメイン抗体は、標的細胞上に存在する細胞表面分子と結合する。任意の実施形  
 態のうちの一部では、細胞表面分子は、タンパク質、グリカン、脂質または低分子量分子である  
 。

40

## 【0099】

任意の実施形態のうちの一部では、標的細胞は、腫瘍浸潤リンパ球、T 細胞、新生物ま  
 たは腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系 ( C N S ) 細胞、造血幹細胞 ( H  
 S C )、肝臓細胞または完全に分化した細胞からなる群から選択される。任意の実施形  
 態のうちの一部では、標的細胞は、C D 3 + T 細胞、C D 4 + T 細胞、C D 8 + T 細  
 胞、肝細胞、造血幹細胞、C D 3 4 + 造血幹細胞、C D 1 0 5 + 造血幹細胞、C D 1

50

17 + 造血幹細胞、CD105 + 内皮細胞、B細胞、CD20 + B細胞、CD19 + B細胞、がん細胞、CD133 + がん細胞、EpCAM + がん細胞、CD19 + がん細胞、Her2/Neu + がん細胞、GluA2 + ニューロン、GluA4 + ニューロン、NKGD2 + ナチュラルキラー細胞、SLC1A3 + 星状細胞、SLC7A10 + 脂肪細胞、またはCD30 + 肺上皮細胞からなる群から選択される。任意の実施形態のうちの一部では、シングルドメイン抗体は、標的細胞上に存在する抗原またはその一部と結合する。

#### 【0100】

本明細書では、(i) ウイルス核酸(複数可)と、(ii) ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸と、(iii) ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び結合ドメインを含む標的化エンベロープタンパク質をコードする核酸とを含む、プロデューサー細胞であって、結合ドメインが、(i) ASGR1、ASGR2、及びTM4SF5、任意にヒトASGR1、ヒトASGR2及びヒトASGR2からなる群から選択される細胞表面分子と結合するか、(ii) CD4またはCD8、任意にヒトCD4またはヒトCD8からなる群から選択される細胞表面分子と結合するか、または(iii) 低密度リポタンパク質受容体(LDL-R)、任意にヒトLDL-Rである細胞表面分子と結合する、プロデューサー細胞が提供される。任意の実施形態のうちの一部では、ウイルス核酸(複数可)は、レンチウイルス核酸である。

10

#### 【0101】

任意の実施形態のうちの一部では、細胞表面分子または抗原は、ASGR1、ASGR2及びTM4SF5からなる群から選択される。任意の実施形態のうちの一部では、細胞表面分子または抗原は、CD8またはCD4である。任意の実施形態のうちの一部では、細胞表面分子または抗原は、LDL-Rである。

20

#### 【0102】

任意の実施形態のうちの一部では、ウイルス核酸(複数可)は、ウイルス複製に關与する1つ以上の遺伝子を欠く。任意の実施形態のうちの一部では、ウイルス核酸は、Gag、Pol、Rev及びTatのうちの一つ以上から選択されるウイルスパッケージングタンパク質をコードする核酸を含む。

#### 【0103】

任意の実施形態のうちの一部では、ウイルス核酸は、以下の核酸配列：5'LTR(例えば、U5を含み、機能性U3ドメインを欠く)、プサイパッケージング要素(プサイ)、セントラルポリプリン配列(cPPT)/セントラル終結配列(CTS)(例えば、DNAフラップ)、ポリAテール配列、転写後制御要素(例えば、WPRE)、Rev反応要素(RRE)、及び3'LTR(例えば、U5を含み、機能性U3を欠く)のうちの一つ以上(例えば、すべて)を含む。

30

#### 【0104】

任意の実施形態のうちの一部では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分は、(i) 配列番号2に示される配列、(ii) 配列番号2と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分は、(i) 配列番号5に示される配列、(ii) 配列番号5と少なくとも90%

40

50

もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、少なくとも 96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するアミノ酸配列を含む。

【0105】

任意の実施形態のうちの一部では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分は、(i)配列番号7に示される配列、(ii)配列番号7と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。任意の実施形態のうちの一部では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分は、(i)配列番号8に示される配列をコードするヌクレオチド配列によってコードする配列、(ii)配列番号8と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有する配列をコードするヌクレオチド配列によってコードされるアミノ酸配列を含む。

【0106】

任意の実施形態のうちの一部では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分は、(i)配列番号23に示される配列、(ii)配列番号23と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

【0107】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分は、(i)配列番号9、配列番号28または配列番号44に示される配列、(ii)配列番号9、配列番号28または配列番号44と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくと

10

20

30

40

50

も 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

【0108】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、(i) 配列番号 10 に示される配列、(ii) 配列番号 10 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

10

【0109】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、(i) 配列番号 35 に示される配列、(ii) 配列番号 35 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

20

30

【0110】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、(i) 配列番号 45 に示される配列、(ii) 配列番号 45 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

40

【0111】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、(i) 配列番号 11 に示される配列、(ii) 配列番号 11 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、

50

84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

【0112】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分は、(i) 配列番号36に示される配列、(ii) 配列番号36と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

10

20

【0113】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分は、(i) 配列番号46に示される配列、(ii) 配列番号46と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

30

【0114】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分は、(i) 配列番号12に示される配列、(ii) 配列番号12と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

40

【0115】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分は、(i) 配列番号37に示される配列

50

、( i i ) 配列番号 37 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を含む。

10

**【 0 1 1 6 】**

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、( i ) 配列番号 47 に示される配列、( i i ) 配列番号 47 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を含む。

20

**【 0 1 1 7 】**

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、( i ) 配列番号 13 に示される配列、( i i ) 配列番号 13 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を含む。

30

**【 0 1 1 8 】**

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、( i ) 配列番号 38 に示される配列、( i i ) 配列番号 38 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を含む。

40

**【 0 1 1 9 】**

50



任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、( i ) 配列番号 48 に示される配列、( i i ) 配列番号 48 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を含む。

10

**【 0 1 2 0 】**

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、( i ) 配列番号 14 に示される配列、( i i ) 配列番号 14 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を含む。

20

**【 0 1 2 1 】**

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、( i ) 配列番号 39 に示される配列、( i i ) 配列番号 39 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を含む。

30

**【 0 1 2 2 】**

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、( i ) 配列番号 49 に示される配列、( i i ) 配列番号 49 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一

40

50

性を有するアミノ酸配列を含む。

【0123】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分は、（i）配列番号15に示される配列、（ii）配列番号15と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

10

【0124】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分は、（i）配列番号40に示される配列、（ii）配列番号40と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

20

【0125】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分は、（i）配列番号50に示される配列、（ii）配列番号50と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む。

30

40

【0126】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分は、（i）配列番号16に示される配列、（ii）配列番号16と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくはは

50

約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

【0127】

任意の実施形態のうちの一部では、ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、(i) 配列番号 51 に示される配列、(ii) 配列番号 51 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む。

10

【0128】

提供される実施形態のいくつかの態様では、標的化脂質粒子は、同様の脂質二重層に同じエンベロープタンパク質が組み込まれているがそれが代替的標的化部位に融合されている参照脂質粒子と比較して多い標的化エンベロープタンパク質発現を有し、任意に代替的標的化部位は、一本鎖可変断片 (scFv) である。任意の実施形態のうちの一部では、発現は、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%、300%、400%、500% 以上またはそれを超えて増加する。いくつかの実施形態では、発現は、1.5 倍、2 倍、3 倍、4 倍、5 倍、6 倍、7 倍、8 倍、9 倍、10 倍、15 倍、20 倍、30 倍以上またはそれを超えて、好ましくは 10 倍もしくは約 10 倍もしくは約 10 倍超またはそれを超えて増加する。任意の実施形態のうちの一部では、形質導入後の標的細胞における力価は、 $1 \times 10^6$  形質導入単位 (TU) / mL 以上、 $2 \times 10^6$  TU / mL 以上、 $3 \times 10^6$  TU / mL 以上、 $4 \times 10^6$  TU / mL 以上、 $5 \times 10^6$  TU / mL 以上、 $6 \times 10^6$  TU / mL 以上、 $7 \times 10^6$  TU / mL 以上、 $8 \times 10^6$  TU / mL 以上、 $9 \times 10^6$  TU / mL 以上、または  $1 \times 10^7$  TU / mL 以上である。また、本明細書では、脂質粒子の集団の中で、50% もしくは約 50% 超、55% もしくは約 55% 超、60% もしくは約 60% 超、65% もしくは約 65% 超、70% もしくは約 70% 超、または 75% もしくは約 75% 超は標的化エンベロープタンパク質について表面陽性である組成物が提供される。任意の実施形態のうちの一部では、標的化エンベロープタンパク質は、少なくとも約 (0.001、0.002、0.005、0.01、0.02、0.05、0.1、0.2 または 0.5) の標的化エンベロープタンパク質 /  $\text{nm}^2$  の密度で標的化脂質粒子の表面上に存在する。

20

30

【0129】

本明細書では、本明細書で提供される実施形態のいずれかのプロデューサー細胞から生成されたウイルスベクター粒子またはウイルス様粒子が提供される。

40

【0130】

本明細書では、本明細書で提供される実施形態のいずれかの複数の標的化脂質粒子を含む組成物が提供される。いくつかの実施形態では、組成物は、薬学的に許容可能な担体をさらに含む。任意の実施形態のうちの一部では、標的化脂質粒子は、1  $\mu\text{m}$  未満の平均直径を含む。任意の実施形態のうちの一部では、組成物は、少なくとも約 (0.001、0.002、0.005、0.01、0.02、0.05、0.1、0.2 または 0.5) の標的化エンベロープタンパク質 /  $\text{nm}^2$  の平均密度で標的化脂質粒子の表面上に存在する標的化エンベロープタンパク質をさらに含む。

【0131】

50

本明細書では、膜（例えば、細胞膜）に同じエンベロープタンパク質が組み込まれているがそれが代替的標的化部位に融合されている参照プロデューサー細胞と比較して多い、標的化エンベロープタンパク質の膜（例えば、細胞膜）発現を含有するプロデューサー細胞であって、任意に代替的標的化部位は、一本鎖可変断片（s c F v）である、プロデューサー細胞が提供される。いくつかの実施形態では、発現は、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%、300%、400%、500%以上またはそれを超えて増加する。いくつかの実施形態では、発現は、1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍以上またはそれを超えて、好ましくは10倍もしくは約10倍もしくは10倍超またはそれを超えて増加する。いくつかの実施形態では、平方ミクロン当たり少なくとも20のタンパク質（例えば、少なくとも50、100、200、500、1000、2000、5000、または10,000のタンパク質）である、プロデューサー細胞の膜（例えば、細胞膜）上の標的化エンベロープタンパク質の発現を、プロデューサー細胞は有する。任意の実施形態のうちの一部では、標的化エンベロープタンパク質は、プロデューサー細胞の総膜（例えば、細胞膜）タンパク質の少なくとも0.1%（例えば、少なくとも0.2%、0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、または10%）（例えば、総タンパク質量に対して）を構成する。

10

**【0132】**

本明細書では、細胞に形質導入する方法であって、本明細書に記載のウイルスベクターのいずれかまたは本明細書に記載の組成物のいずれかで細胞に形質導入することを含む、方法が提供される。任意の実施形態のうちの一部では、レンチウイルスベクターまたは標的化脂質粒子の標的化エンベロープタンパク質は、CD4を標的とし、細胞は、CD4+細胞である。任意の実施形態のうちの一部では、レンチウイルスベクターの標的化エンベロープタンパク質は、CD8を標的とし、細胞は、CD8+細胞である。任意の実施形態のうちの一部では、レンチウイルスベクターの標的化エンベロープタンパク質は、ASGR1、ASGR2またはTM4SF5を標的とし、細胞は、肝細胞である。

20

**【0133】**

本明細書では、外因性物質を対象（例えば、ヒト対象）に送達する方法であって、本明細書で提供される実施形態のいずれかの標的化脂質粒子または本明細書で提供される実施形態のいずれかの組成物を対象に投与することを含み、標的化脂質粒子またはレンチウイルスベクターが外因性物質を含む、方法が提供される。

30

**【0134】**

本明細書では、外因性物質を対象（例えば、ヒト対象）に送達する方法であって、本明細書に記載の組成物のいずれかを対象に投与することを含み、複数の標的化脂質粒子またはレンチウイルスベクターが外因性物質を含む、方法が提供される。

**【0135】**

本明細書では、キメラ抗原受容体（CAR）を細胞に送達する方法であって、細胞を本明細書に記載のレンチウイルスベクターのいずれかまたは本明細書に記載の実施形態のいずれかの標的化脂質粒子と接触させることを含み、レンチウイルスベクターまたは標的化脂質粒子が、CARをコードする核酸を含む、方法が提供される。

40

**【0136】**

本明細書では、キメラ抗原受容体（CAR）を細胞に送達する方法であって、細胞を本明細書に記載の組成物のいずれかと接触させることを含み、複数のレンチウイルスベクターまたは標的化脂質粒子が、CARをコードする核酸を含む、方法が提供される。

**【0137】**

本明細書では、外因性物質を肝細胞に送達する方法であって、細胞を、本明細書に記載のレンチウイルスベクターのいずれか、または本明細書に記載の実施形態のいずれかの標的化脂質粒子またはレンチウイルスベクターと接触させることを含む、方法が提供される。

50

## 【0138】

本明細書では、外因性物質を肝細胞に送達する方法であって、細胞を本明細書に記載の組成物のいずれかと接触させることを含み、複数のレンチウイルスベクターまたは標的化脂質粒子が、肝細胞への送達のための外因性物質を含む、方法が提供される。任意の実施形態のうちの一部では、接触させることは、レンチウイルスベクターまたは標的化脂質粒子で細胞に形質導入する。

## 【0139】

本明細書では、対象（例えば、ヒト対象）における疾患または障害を処置する方法であって、本明細書で提供される実施形態のいずれかの標的化脂質粒子または本明細書で提供される実施形態のいずれかの組成物を対象に投与することを含む、方法が提供される。

10

## 【0140】

本明細書では、哺乳動物細胞を標的化脂質粒子に融合させる方法であって、本明細書で提供される実施形態のいずれかの標的化脂質粒子または本明細書で提供される実施形態のいずれかの組成物を対象に投与することを含む、方法が提供される。任意の実施形態のうちの一部では、哺乳動物細胞を標的化脂質粒子に融合させることは、外因性物質を対象（例えば、ヒト対象）に送達する。任意の実施形態のうちの一部では、哺乳動物細胞を標的化脂質粒子に融合させることは、対象（例えば、ヒト対象）における疾患または障害を処置する。任意の実施形態のうちの一部では、レンチウイルスベクターまたは標的化脂質粒子の標的化エンベロープタンパク質は、CD4を標的とし、細胞は、CD4 + 細胞である。任意の実施形態のうちの一部では、レンチウイルスベクターの標的化エンベロープタン

20

## 【0141】

任意の実施形態のうちの一部では、標的化脂質粒子は、同様の脂質二重層に同じエンベロープタンパク質が組み込まれているがそれが代替的標的化部位に融合されている参照脂質粒子と比較して多い標的化エンベロープタンパク質発現を有する。いくつかの実施形態では、代替的標的化部位は、一本鎖可変断片（scFv）である。任意の実施形態のうちの一部では、発現は、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%、300%、400%、500%以上またはそれを超えて増加する。任意の実施形態のうちの一部では、発現は、1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍、30倍以上またはそれを超えて、好ましくは10倍もしくは約10倍もしくは10倍超またはそれを超えて増加する。

30

## 【0142】

任意の実施形態のうちの一部では、形質導入後の標的細胞における力価は、 $1 \times 10^6$  形質導入単位（TU）/mL以上、 $2 \times 10^6$  TU/mL以上、 $3 \times 10^6$  TU/mL以上、 $4 \times 10^6$  TU/mL以上、 $5 \times 10^6$  TU/mL以上、 $6 \times 10^6$  TU/mL以上、 $7 \times 10^6$  TU/mL以上、 $8 \times 10^6$  TU/mL以上、 $9 \times 10^6$  TU/mL以上、または $1 \times 10^7$  TU/mL以上である。

40

## 【0143】

任意の実施形態のうちの一部では、組成物における脂質粒子またはレンチウイルスベクターの集団の中で、50%もしくは約50%超、55%もしくは約55%超、60%もしくは約60%超、65%もしくは約65%超、70%もしくは約70%超、または75%もしくは約75%超は、標的化エンベロープタンパク質について表面陽性である。任意の実施形態のうちの一部では、標的化エンベロープタンパク質は、少なくとも約（0.001、0.002、0.005、0.01、0.02、0.05、0.1、0.2または0.5）の標的化エンベロープタンパク質/nm<sup>2</sup>の密度で標的化脂質粒子の表面上に存在する。

## 【0144】

50

本明細書では、本明細書に記載の実施形態のいずれかの複数の標的化脂質粒子または本明細書に記載の実施形態のいずれかの複数のレンチウイルスベクターを含む組成物であって、標的化エンベロープタンパク質は、少なくとも約(0.001、0.002、0.005、0.01、0.02、0.05、0.1、0.2または0.5)の標的化エンベロープタンパク質/nm<sup>2</sup>の平均密度で標的化脂質粒子の表面上に存在する、組成物が提供される。

#### 【0145】

任意の実施形態のうちの一部では、膜(例えば、細胞膜)に同じエンベロープタンパク質が組み込まれているがそれが代替的標的化部位に融合されている参照プロドューサー細胞と比較して多い、標的化エンベロープタンパク質の膜(例えば、細胞膜)発現を、プロドューサー細胞は有し、任意に代替的標的化部位は、一本鎖可変断片(s c F v)である。任意の実施形態のうちの一部では、発現は、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%、300%、400%、500%以上またはそれを超えて増加する。任意の実施形態のうちの一部では、発現は、1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍、30倍以上またはそれを超えて、好ましくは10倍もしくは約10倍もしくは10倍超またはそれを超えて増加する。任意の実施形態のうちの一部では、平方ミクロン当たり少なくとも20のタンパク質(例えば、少なくとも50、100、200、500、1000、2000、5000、または10,000のタンパク質)である、プロドューサー細胞の膜(例えば、細胞膜)上の標的化エンベロープタンパク質の発現を、プロドューサー細胞は有する。任意の実施形態のうちの一部では、標的化エンベロープタンパク質は、プロドューサー細胞の総膜(例えば、細胞膜)タンパク質の少なくとも0.1%(例えば、少なくとも0.2%、0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、または10%)(例えば、総タンパク質量に対して)を構成する。

#### 【0146】

詳細な説明

本明細書では、内腔またはキャビティを包囲する脂質二重層と、(1)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び(2)結合ドメイン、例えば、シングルドメイン抗体(s d A b)可変ドメインを含有する標的化エンベロープタンパク質とを含有する、標的化脂質粒子であって、標的化エンベロープタンパク質が、脂質粒子の脂質二重層に埋め込まれている、標的化脂質粒子が提供される。特定の実施形態では、結合ドメイン、例えば、シングルドメイン抗体は、所望の標的分子に結合する、例えば、特異的に結合する能力を有する抗体である。例示的な結合ドメインは、セクションII.A.2に記載されている。いくつかの実施形態では、標的化脂質粒子はまた、脂質二重層に埋め込まれたヘニパウイルス融合(F)タンパク質分子またはその生物学的に活性な部分を含有する。特定の実施形態では、脂質粒子は、ウイルス様粒子、ウイルス、またはウイルスベクター、例えば、レンチウイルスベクターであり得る。

#### 【0147】

いくつかの実施形態では、Gタンパク質及びFタンパク質の一方または両方は、ヘンドラ(HeV)もしくはニパ(NiV)ウイルスからのものであるか、またはその生物学的に活性な部分であるか、またはそのパリアントもしくは変異体である。特定の実施形態では、Gタンパク質及びFタンパク質の両方は、ヘンドラ(HeV)またはニパ(NiV)ウイルスからのものである。いくつかの実施形態では、融合及び結合糖タンパク質は、ニパウイルスの細胞侵入を媒介する。

#### 【0148】

Fタンパク質、例えば、NiV-Fは、疎水性融合ペプチド及び2つのヘプタッドリピート領域を有するエクストドメインなどの、多くのファミリーの融合タンパク質(例えば、HIV-1 gp41またはインフルエンザウイルス血球凝集素[HA])と共通する構

造的及び機能的特徴を有するクラスI融合タンパク質である(White JM et al. 2008. Crit Rev Biochem Mol Biol 43:189-219)。Fタンパク質は、不活性な前駆体F<sub>0</sub>として合成され、2つのジスルフィド連結サブユニットF<sub>1</sub>及びF<sub>2</sub>へのタンパク質分解的切断によって活性化される(Moll M. et al. 2004. J. Virol. 78(18):9705-9712)。  
【0149】

Gタンパク質は、N末端細胞質尾部、膜貫通ドメイン、細胞外茎部、及び球状頭部を含有するII型膜貫通糖タンパク質である、ヘニパウイルス(例えば、ニパウイルスまたはヘンドラウイルス)の結合タンパク質である(Liu, Q. et al. 2015. Journal of Virology, 89(3):1838-1850)。結合タンパク質であるNiV-Gは、受容体エフリンB2及びエフリンB3を認識する。NiV-Gに対する受容体の結合は、NiV-Fの作動を最終的にもたらず一連の構造的変化を誘発し、NiV-Fの融合ペプチドを露出させ、ウイルス細胞膜融合をもたらず別の一連の構造的変化を可能とする(Stone J. A. et al. 2016. J Virol. 90(23):10762-10773)。エフリンB2は、主要なNiV受容体として(Negrete et al., 2005)、エフリンB3は、代替的受容体として(Negrete et al., 2006)過去に同定された。実際、NiV-Gは、ピコモル範囲の親和性結合定数(K<sub>d</sub>)でエフリンB2及びB3に対する高い親和性を有する(Negrete et al., 2006)(細胞表面に発現したエフリンB2及びB3についてそれぞれK<sub>d</sub>=0.06nM及び0.58nM)。  
【0150】

標的化脂質粒子の形質導入の効率は、NiV-F及びNiV-Gの一方または両方において超融合変異を操作することによって改善され得る。いくつかのそのような変異は、先行記載されている(例えば、Lee et al., 2011, Trends in Microbiologyを参照されたい)。これは、例えば、エフリンB2及び/またはB3に対するNiV-Gの特異性及びピコモル親和性を維持するために有用であり得る。また、エフリンB2及びB3結合を完全に無効にするが、このNiV-GとNiV-Fとの会合に影響を与えないNiV-Gにおける変異が同定されている。脂質粒子の標的化を改善するための方法は、結合分子とGタンパク質(例えば、エフリンB2及びエフリンB3結合を無効にするための変異を有するNiV-Gを含むNiV-G)との融合によって達成され得る。これにより、異なる細胞表面分子に対して仕向けられた結合分子の添加を介してエフリンB2+ではない他の所望の細胞タイプの標的化を可能とする改変されたGタンパク質指向性が可能となり得る。  
【0151】

Gタンパク質に融合されたそのような結合分子を組み込んでいる再標的化脂質粒子が生成されたが、いくつかの結合分子は、Gタンパク質(例えば、NiV-G)と融合された場合、他のものよりも脂質粒子の表面上で良好に発現することが本明細書で見出された。例えば、シングルドメイン抗体(sdAb)、例えば、VHHは、一本鎖可変断片(scFv)よりも10倍良好に発現し得ることが見出されている。理論に縛られることを望むものではないが、発現の増加は、脂質粒子の表面上の再標的化Gタンパク質の安定性の増加に起因し得る。このより多い発現は、同じ標的分子に対する代替的結合ドメイン、例えば、scFvを含有する同様の脂質粒子と比較して脂質粒子が標的分子(例えば、細胞表面分子)を標的とする能力を改善し得る。  
【0152】

よって、本明細書では、標的細胞上の細胞表面分子に対して仕向けられた、またはそれに結合することができるsdAb可変ドメインに結合したヘニパウイルスのGタンパク質(例えば、HendraまたはNipah、例えば、NiV-G)を含有する標的化脂質粒子が提供される。sdAb可変ドメインには、VLまたはVHのみのsdAb、ナノボディ、ラクダ科動物VHHドメイン、サメIgNARまたはその断片のものが含まれ得る。いくつかの実施形態では、sdAbは、VHHである。

10

20

30

40

50

## 【 0 1 5 3 】

提供される実施形態の態様では、標的化脂質粒子は、ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分と、( i )ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分及び ( i i ) シングルドメイン抗体 ( s d A b ) 可変ドメインを含む標的化エンベロープタンパク質とを発現するように操作され得、F タンパク質分子またはその生物学的に活性な部分及び標的化エンベロープタンパク質は、脂質二重層に埋め込まれる。いくつかの実施形態では、s d A b 可変ドメインは、G タンパク質またはその生物学的に活性な部分の C 末端に結合されている。いくつかの実施形態では、s d A b 可変ドメインは、リンカーを介して G タンパク質に結合されている。

10

## 【 0 1 5 4 】

また、対象への i n v i v o 投与後を含む、例えば、診断または治療剤の細胞への送達のための、1つ以上の外因性物質を追加的に含有する標的化脂質粒子が提供される。また、本明細書では、診断及び治療方法における、標的化脂質粒子の方法及び使用が提供される。また、ポリヌクレオチド、標的化脂質非細胞粒子を操作、調製、及び生成するための方法、粒子を含有する組成物、ならびに粒子を使用、生成及び投与することを含む及びそのためのキット及びデバイスが提供される。

## 【 0 1 5 5 】

本出願で言及される特許文献、科学記事及びデータベースを含むすべての刊行物は、各々の個々の刊行物が参照により個々に開示されているのと同じ程度にすべての目的のためにそれらの全体が参照により組み込まれる。本明細書に示される定義が参照により本明細書に組み込まれる特許、出願、公開された出願及び他の刊行物に示される定義に反している場合または別様に矛盾する場合、本明細書に示される定義は、参照により本明細書に組み込まれる定義より優先する。

20

## 【 0 1 5 6 】

本明細書で使用される節の見出しは、構成目的のみのためであり、記載される主題を限定するものと解釈されるべきではない。

## 【 図面の簡単な説明 】

## 【 0 1 5 7 】

【 図 1 A 】 図 1 A ~ 1 C は、s c F v または V H H 結合モダリティを含有するコンストラクトでトランスフェクションされた細胞の特徴づけを示している。図 1 A は、フローサイトメトリーによって分析され、H i s + 細胞の % によって定量された蛍光強度中央値 ( M F I ) として示される、s c F v または V H H 結合モダリティを含有するコンストラクトでトランスフェクションされた細胞の表面発現を示している。空のベクター及びバインダードメインを有しない発現ベクターを陰性対照として使用した。

30

【 図 1 B 】 図 1 A ~ 1 C は、s c F v または V H H 結合モダリティを含有するコンストラクトでトランスフェクションされた細胞の特徴づけを示している。図 1 B は、フローサイトメトリーによって分析され、% F c + 細胞によって定量された蛍光強度中央値 ( M F I ) として示される、V H H 結合モダリティの s c F v を含有するコンストラクトでトランスフェクションされた細胞の可溶性 h C D 4 - F c タンパク質に対する結合を示している。空のベクター及びバインダードメインを有しない発現ベクターを陰性対照として使用した。

40

【 図 1 C 】 図 1 A ~ 1 C は、s c F v または V H H 結合モダリティを含有するコンストラクトでトランスフェクションされた細胞の特徴づけを示している。図 1 C は、フローサイトメトリーによって分析され、H i s + 細胞の % によって定量される蛍光強度中央値 ( M F I ) として示される、s c F v 結合モダリティと比較した、V H H 結合モダリティを含有するコンストラクトでトランスフェクションされた細胞についての 2 9 3 細胞上の標的化結合配列の表面発現を示している。空のベクター及びバインダードメインを有しない発現ベクターを陰性対照として使用した。

【 図 2 】 T 細胞について濃縮するためにネガティブ選択し、解凍し、抗 C D 3 / 抗 C D 2

50



8で活性化した末梢血からのPanT細胞に対するscFVまたはVHH結合モダリティを含有する4つの例示的なコンストラクトの形質導入有効性を示している。細胞をフローサイトメトリーによって分析し、力価を、GFP+であったCD4陽性細胞の%によって決定した。

【図3A】図3A~Bは、フローサイトメトリーによって分析される、NOD-scid-IL2r<sup>null</sup>マウスに腹腔内注入された活性化PBMCを使用したin vivoモデルでのCD8再標的化偽型化レンチウイルスの形質導入効率を示している。CD8再標的化偽型化レンチウイルスの形質導入効率がCD8+ (図3A)またはCD8- (図3B) T細胞について示されており、力価は、GFP+であったCD8陽性または陰性細胞の%によって決定した。

10

【図3B】図3Aの説明を参照のこと。

【図4A】図4A~Bは、in vitroで白血病細胞の殺傷をもたらすためのキメラ抗原受容体(CAR)を含有するCD8再標的化偽型化レンチウイルスの能力を示している。図4Aは、形質導入後4日目でCD8+細胞上のCD19+CAR発現を検出するための能力を示している。図4Bは、フローサイトメトリーによって分析された、インキュベート後18時間の時点で評価されたNa1m6細胞の排除を示している。

【図4B】図4Aの説明を参照のこと。

【発明を実施するための形態】

【0158】

#### I. 定義

20

別段定義されない限り、本明細書で使用される当該技術分野のすべての用語、表記ならびに他の技術的及び科学的用語または専門用語は、請求される主題が関連する当該技術分野の当業者によって一般的に理解されるものと同じ意味を有することが意図されている。いくつかの場合では、一般的に理解される意味を有する用語は、明確性のため及び/またはすぐに参照できるようにするために本明細書で定義されており、本明細書におけるそのような定義の包含は、かならずしも、当該技術分野で通常理解されるものとの実質的な差を示すものと解釈されるべきではない。

【0159】

別段定義されない限り、本明細書で使用されるすべての技術的及び科学的用語、頭字語、ならびに略語は、本発明が関連する技術分野の当業者によって一般的に理解されるものと同じ意味を有する。別段示されない限り、化学及び生化学名称のための略語及び記号は、IUPAC-IUB命名法に従う。別段示されない限り、すべての数値範囲は、範囲を定義する値ならびにそれらの間のすべての整数値を含む。

30

【0160】

本明細書で使用される場合、冠詞「a」及び「an」は、冠詞の文法上の目的語の1つまたは複数(すなわち、少なくとも1つの)を指す。例として、「要素(an element)」は、1つの要素または複数の要素を意味する。

【0161】

本明細書で使用される場合、「約」は、当業者によって理解され、それが使用される文脈である程度異なる。本明細書で使用される場合、「約」は、測定可能な値、例えば、量、時間的継続時間などに言及する場合、指定された値から±20%または±10%、より好ましくは±5%、さらにより好ましくは±1%、またより好ましくは±0.1%の変動を包含するが、それは、そのような変動が、開示される方法を実施するために適切であるからである。

40

【0162】

本明細書で使用される場合、「脂質粒子」は、内腔またはキャビティを包囲する両親媒性脂質の二重層を含有する任意の生物学的または合成粒子を指す。典型的には脂質粒子は、核を含有しない。脂質粒子の例には、固体粒子、例えば、ナノ粒子、ウイルス由来粒子または細胞由来粒子が含まれる。そのような脂質粒子には、ウイルス粒子(例えば、レンチウイルス粒子)、ウイルス様粒子、ウイルスベクター(例えば、レンチウイルスベクター

50

ー) エクソソーム、除核された細胞、様々な小胞、例えば、微小胞、膜小胞、細胞外膜小胞、細胞膜小胞、巨大細胞膜小胞、アポトーシス小体、ミト粒子 (mitoparticle)、ピレノサイト、またはリソソームが含まれるがこれらに限定されない。いくつかの実施形態では、脂質粒子は、フソソームであり得る。いくつかの実施形態では、脂質粒子は、血小板ではない。

【0163】

本明細書で使用される場合、例えば、Gタンパク質またはFタンパク質などのタンパク質に関して、「生物学的に活性な部分」は、タンパク質の全長の活性または特性を示すまたは保持するタンパク質の一部を指す。例えば、Fタンパク質の生物学的に活性な部分は、Gタンパク質と併せて、各々が脂質二重層に埋め込まれている場合、融合活性を保持する。Gタンパク質の生物学的に活性な部分は、Fタンパク質と併せて、各々が脂質二重層に埋め込まれている場合に融合活性を保持する。保持される活性には、全長または野生型Fタンパク質またはGタンパク質の活性の10%~150%またはそれ以上が含まれる。F及びGタンパク質の生物学的に活性な部分の例には、細胞質ドメインの切断、例えば、最大で1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、20、25、30、35またはそれ以上の連続アミノ酸の切断が含まれ、例えば、Khetawat and Broder 2010 Virology Journal 7: 312、Witting et al. 2013 Gene Therapy 20: 997-1005、特許出願番号WO/2013/148327を参照されたい。

10

【0164】

本明細書で使用される場合、「フソソーム」は、内腔またはキャビティを包囲する両親媒性脂質の二重層及び両親媒性脂質二重層と相互作用するフソゲンを含む粒子を指す。実施形態では、フソソームは、核酸を含む。いくつかの実施形態では、フソソームは、膜包囲調製物である。いくつかの実施形態では、フソソームは、源細胞に由来する。

20

【0165】

本明細書で使用される場合、「フソソーム組成物」は、1つ以上のフソソームを含む組成物を指す。

【0166】

本明細書で使用される場合、「フソゲン」は、2つの膜包囲内腔間で相互作用を作り出す物質または分子を指す。実施形態では、フソゲンは、膜の融合を容易化する。他の実施形態では、フソゲンは、2つの内腔(例えば、レトロウイルスベクターの内腔及び標的細胞の細胞質)間の接続、例えば、孔部を生成する。いくつかの実施形態では、フソゲンは、2つ以上のタンパク質の複合体を含み、例えば、いずれのタンパク質も融合活性を単独で有しない。いくつかの実施形態では、フソゲンは、標的化ドメインを含む。

30

【0167】

本明細書で使用される場合、「再標的化フソゲン」は、フソゲンの天然に存在する形態の一部ではない配列を有する標的化部位を含むフソゲンを指す。実施形態では、フソゲンは、フソゲンの天然に存在する形態における標的化部位と比較して異なる標的化部位を含む。実施形態では、フソゲンの天然に存在する形態は、再標的化メインを欠き、再標的化フソゲンは、フソゲンの天然に存在する形態にはない標的化部位を含む。実施形態では、フソゲンは、標的化部位を含むように修飾される。実施形態では、フソゲンは、例えば、膜貫通ドメイン、融合的に活性なドメイン、または細胞質ドメインにおいてフソゲンの天然に存在する形態に対して標的化部位の外側に1つ以上の配列改変を含む。

40

【0168】

本明細書で使用される場合、「標的化エンベロープタンパク質」は、分子を所望の細胞タイプに標的化するシングルドメイン抗体(sdAb)可変ドメイン、例えば、VLまたはVHのみのsdAb、ナノボディ、ラクダ科動物VHHドメイン、サメIgNARまたはその断片に結合したヘニパウイルスGタンパク質を含むポリペプチドを指す。いくつかのそのような実施形態では、結合は、直接的、またはペプチドリンカーなどのリンカーを介して間接的であり得る。

50

## 【0169】

本明細書で使用される場合、「標的化脂質粒子」は、脂質二重層に埋め込まれた標的化エンベロープタンパク質を含有する脂質粒子を指す。

## 【0170】

本明細書で使用される場合、「レトロウイルス核酸」は、レトロウイルスまたはレトロウイルスベクターに、単独で、またはヘルパー細胞、ヘルパーウイルス、もしくはヘルパープラスミドと組み合わせてパッケージングするための少なくとも最小配列要件を含有する核酸を指す。いくつかの実施形態では、レトロウイルス核酸は、外因性物質、陽性標的細胞特異的制御要素、非標的細胞特異的制御要素、または陰性TCSREをさらに含み、またはコードする。いくつかの実施形態では、レトロウイルス核酸は、5'LTR（例えば、組み込みを促進するためのもの）、U3（例えば、ウイルスゲノムRNA転写を活性化するためのもの）、R（例えば、Tat結合領域）、U5、3'LTR（例えば、組み込みを促進するためのもの）、パッケージング部位（例えば、プサイ（ $\Psi$ ））、RRE（例えば、Revに結合し、核外搬出を促進するためのもの）のうち1つ以上（例えば、すべて）を含む。レトロウイルス核酸は、RNA（例えば、ピリオンの一部である場合）またはDNA（例えば、源細胞に導入されている場合またはレシピエント細胞における逆転写後）を含み得る。いくつかの実施形態では、レトロウイルス核酸は、gag、pol、及びenvのうち1つ以上（例えば、すべて）を含むヘルパー細胞、ヘルパーウイルス、またはヘルパープラスミドを使用してパッケージングされる。

10

## 【0171】

本明細書で使用される場合、「標的細胞」は、標的化脂質粒子が外因性物質を送達することが所望であるタイプの細胞を指す。実施形態では、標的細胞は、特定の組織タイプまたはクラスの細胞、例えば、免疫エフェクター細胞、例えば、T細胞である。いくつかの実施形態では、標的細胞は、罹患細胞、例えば、がん細胞である。いくつかの実施形態では、フソゲン、例えば、再標的化フソゲンは、非標的細胞と比較して標的細胞への外因性物質の優先的送達をもたらす。

20

## 【0172】

本明細書で使用される場合、「非標的細胞」は、標的化脂質粒子が外因性物質を送達することが所望ではない細胞のタイプを指す。いくつかの実施形態では、非標的細胞は、特定の組織タイプまたはクラスの細胞である。いくつかの実施形態では、非標的細胞は、非罹患細胞、例えば、非がん細胞である。いくつかの実施形態では、フソゲン、例えば、再標的化フソゲンは、標的細胞と比較して非標的細胞への外因性物質のより低い送達をもたらす。

30

## 【0173】

本明細書で使用される場合、「シングルドメイン抗体」または「sdAb」は、単一の単量体ドメイン抗原結合/認識ドメインを有する抗体を指す。そのような抗体には、ナノボディ、ラクダ科動物抗体（例えば、VHH）、またはサメ抗体（例えば、IgNAR）が含まれる。いくつかの実施形態では、sdAbの変域ドメインは、3つのCDR及び4つのフレームワーク領域（FR1、CDR1、FR2、CDR2、FR3、CDR3、及びFR4と指定される）を含む。いくつかの実施形態では、sdAb変域ドメインが抗原結合及び特異性を実質的に維持する限り、sdAb変域ドメインは、部分的FR1及び/またはFR4のみを含むようにN末端またはC末端で切断され得、またはそれらのフレームワーク領域の一方もしくは両方を欠く。

40

## 【0174】

「CDR」という用語は、当業者に対する少なくとも1つの同定手法によって定義される相補性決定領域を示す。所与のCDRまたはFRの正確なアミノ酸配列境界は、Kabata et al. (1991), "Sequences of Proteins of Immunological Interest," 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD ("Kabata" 付番スキーム)、Al-Lazikani 50

et al., (1997) JMB 273, 927-948 (“Chothia”付番スキーム)、MacCallum et al., J. Mol. Biol. 262:732-745 (1996), “Antibody-antigen interactions: Contact analysis and binding site topography,” J. Mol. Biol. 262, 732-745.” (“Contact”付番スキーム)、Lefranc MP et al., “IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains,” Dev Comp Immunol, 2003 Jan; 27(1):55-77 (“IMGT”付番スキーム)、Honegger A and Pluckthun A, “Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool,” J Mol Biol, 2001 Jun 8; 309(3):657-70, (“Aho”付番スキーム)、及びMartin et al., “Modeling antibody hypervariable loops: a combined algorithm,” PNAS, 1989, 86(23):9268-9272, (“AbM”付番スキーム)に記載されているものを含む、いくつかのよく知られているスキームのいずれかを使用して容易に決定され得る。

#### 【0175】

所与のCDRまたはFRの境界は、同定に使用されるスキームに応じて変わり得る。例えば、Kabatsスキームは、構造的アライメントに基づくのに対し、Chothiasスキームは、構造的情報に基づく。Kabats及びChothiasスキームの両方について付番は、挿入文字、例えば、「30a」によって提供される挿入、及びいくつかの抗体で現れる欠失を伴って、最も共通の抗体領域配列長さに基づく。これらの2つのスキームは、異なる位置で所定の挿入及び欠失(「インデル」)を配置し、異なる付番をもたらす。Contactスキームは、複雑な結晶構造の分析に基づいており、Chothia付番スキームと多くの点で類似する。AbMスキームは、Oxford MolecularのAbM抗体モデリングソフトウェアによって使用されるものに基づくKabats定義とChothia定義との間の折衷案である。

#### 【0176】

いくつかの実施形態では、CDRは、Chothia付番スキーム、Kabats付番スキーム、Kabats及びChothiaの組み合わせ、AbM定義、及び/またはContact定義のいずれかに従って定義され得る。sdAb可変ドメインは、CDR1、CDR2、及びCDR3と指定された3つのCDRを含む。以下の表1は、Kabats、Chothia、AbM、及びContactスキームによってそれぞれ同定されるCDR-H1、CDR-H2、CDR-H3の例示的な位置境界を列挙している。CDR-H1の場合、残基付番は、Kabats及びChothia付番スキームの両方を使用して列挙される。FRは、CDR間に位置し、例えば、FR-H1は、CDR-H1の前に位置し、FR-H2は、CDR-H1とCDR-H2との間に位置し、FR-H3は、CDR-H2とCDR-H3との間に位置するなどである。示されるKabats付番スキームは、H35A及びH35Bで挿入を配置するので、Chothia CDR-H1ループの端部は、示されるKabats付番規則を使用して付番された場合、ループの長さに応じて、H32とH34との間で変わることを留意されたい。

#### 【0177】

(表1) 様々な付番スキームに従うCDRの境界

20

30

40

50

CDR	Kabat	Chothia	AbM	Contact
CDR-H1 (Kabat 付番 <sup>1</sup> )	H31--H35B	H26--H32..34	H26--H35B	H30--H35B
CDR-H1 (Chothia 付番 <sup>2</sup> )	H31--H35	H26--H32	H26--H35	H30--H35
CDR-H2	H50--H65	H52--H56	H50--H58	H47--H58
CDR-H3	H95--H102	H95--H102	H95--H102	H93--H101

1 - Kabat et al. (1991), "Sequences of Proteins of Immunological Interest," 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD

2 - Al-Lazikani et al., (1997) JMB 273, 927-948

10

#### 【0178】

よって、別段特定されない限り、所与の抗体またはその可変領域などのその領域の「CDR」もしくは「相補性決定領域」、または個々の特定のCDR（例えば、CDR-H1、CDR-H2、CDR-H3）は、前述したスキームのいずれかによって定義される、ある（または特定の）相補性決定領域を包含するものと理解されるべきである。例えば、特定のCDR（例えば、CDR-H3）が、所与のsdAbアミノ酸配列における対応するCDRのアミノ酸配列を含有することが記述される場合、そのようなCDRは、前述したスキームのいずれかによって定義される、sdAb内の対応するCDR（例えば、CDR-H3）の配列を有することが理解される。任意の抗体、例えば、sdAbは、CDRを含み、それは、他の前述した付番スキームまたは当業者に知られている他の付番スキームのいずれかに従って同定され得ることが理解される。

20

#### 【0179】

本明細書で使用される場合、抗原などの標的分子に「特異的に結合する」という用語は、結合分子、例えば、シングルドメイン抗体は、代替的分子より頻繁に、より迅速に、より長い継続期間で及び/または特定の標的分子とより高い親和性で反応または会合することを意味する。結合分子、例えば、sdAb可変ドメインは、それが他の分子に結合する場合よりも高い親和性、アビディティで、より容易に、及び/またはより長い継続期間で結合する場合、標的分子に「特異的に結合する」。第1の標的に特異的に結合するsdAbなどの結合分子は、第2の標的に特異的に結合する場合があります、またはしない場合がありますことが理解される。このように、「特異的結合」は、かならずしも、排他的結合を必要とするわけではない（が、それを含み得る）。

30

#### 【0180】

本明細書で使用される場合、ペプチド、ポリペプチドまたは抗体配列に関して、「アミノ酸配列同一性パーセント（％）」及び「相同性」は、配列同一性最大パーセントを達成するために、配列を整合させ、必要に応じてギャップを導入した後の、配列同一性の一部としていずれの保存的置換も考慮しない、特定のペプチドもしくはポリペプチド配列におけるアミノ酸残基と同一である、候補配列におけるアミノ酸残基の百分率として定義される。アミノ酸配列同一性パーセントを決定する目的のためのアライメントは、当該技術分野における技術の範囲内である種々の方式で、例えばBLAST、BLAST-2、ALIGN、またはMEGALIGNTM（DNASTAR）ソフトウェア等の、公的に利用可能なコンピュータソフトウェアを使用して、達成することができる。当業者は、比較される配列の全長にわたって最大のアライメントを達成するために必要な任意のアルゴリズムを含む、アライメントを測定するための適切なパラメータを決定することができる。

40

50

## 【 0 1 8 1 】

アミノ酸置換には、ポリペプチドにおけるあるアミノ酸の別のアミノ酸での置き換えが含まれ得るがこれらに限定されない。例示的な置換が表 2 に示されている。アミノ酸置換は、対象となる抗体に導入され、産物は、所望の活性、例えば、保持 / 改善された結合について、スクリーニングされ得る。

## 【 0 1 8 2 】

(表 2)

元の残基	例示的な置換
Ala (A)	Val、Leu、Ile
Arg (R)	Lys、Gln、Asn
Asn (N)	Gln、His、Asp、Lys、Arg
Asp (D)	Glu、Asn
Cys (C)	Ser、Ala
Gln (Q)	Asn、Glu
Glu (E)	Asp、Gln
Gly (G)	Ala
His (H)	Asn、Gln、Lys、Arg
Ile (I)	Leu、Val、Met、Ala、Phe、ノルロイシン
Leu (L)	ノルロイシン、Ile、Val、Met、Ala、Phe
Lys (K)	Arg、Gln、Asn
Met (M)	Leu、Phe、Ile
Phe (F)	Trp、Leu、Val、Ile、Ala、Tyr
Pro (P)	Ala
Ser (S)	Thr
Thr (T)	Val、Ser
Trp (W)	Tyr、Phe
Tyr (Y)	Trp、Phe、Thr、Ser
Val (V)	Ile、Leu、Met、Phe、Ala、ノルロイシン

10

20

30

## 【 0 1 8 3 】

アミノ酸は、共通の側鎖特性に従って分類され得る：

- (1) 疎水性：ノルロイシン、Met、Ala、Val、Leu、Ile、
- (2) 中性親水性：Cys、Ser、Thr、Asn、Gln、
- (3) 酸性：Asp、Glu、
- (4) 塩基性：His、Lys、Arg、
- (5) 鎖配向に影響を及ぼす残基：Gly、Pro、
- (6) 芳香族：Trp、Tyr、Phe。

40

## 【 0 1 8 4 】

非保存的置換は、これらのクラスのうちの 1 つのメンバーを別のクラスと交換することを伴うことになる。

## 【 0 1 8 5 】

タンパク質の位置を参照して「に対応する」という用語、例えば、ヌクレオチドまたはアミノ酸位置が、配列表に示されるものなどの開示される配列におけるヌクレオチドまたはアミノ酸位置「に対応する」という記述は、構造的配列アライメントに基づいてまたは G A P アルゴリズムなどの標準的なアライメントアルゴリズムを使用して開示される配列

50

とのアライメントで同定されるヌクレオチドまたはアミノ酸位置を指す。例えば、同様の配列（例えば、断片または種バリエーション）の対応する残基は、構造的アライメント法によって参照配列に対するアライメントによって決定され得る。配列をアライメントすることにより、当業者は、例えば、ガイドとして保存された及び同一のアミノ酸残基を使用して、対応する残基を同定し得る。

**【0186】**

「単離された」という用語は、本明細書で使用される場合、天然で典型的に見られるまたは生成される成分の少なくとも一部から分離された分子を指す。例えば、ポリペプチドは、それが生成された細胞の成分の少なくとも一部から分離された場合、「単離された」と称される。ポリペプチドが発現後に細胞によって分泌される場合、生成された細胞からポリペプチドを含有する上清を物理的に分離することは、ポリペプチドを「単離すること」とみなされる。同様に、ポリヌクレオチドは、天然で典型的に見られるより大きなポリヌクレオチド（例えば、DNAポリヌクレオチドの場合、ゲノムDNAまたはミトコンドリアDNAなど）の一部ではない場合、または例えば、RNAポリヌクレオチドの場合、それが生成された細胞の成分の少なくとも一部から分離された場合、「単離された」と称される。よって、宿主細胞内部のベクターに含有されるDNAポリヌクレオチドは、「単離された」と称され得る。

10

**【0187】**

「有効量」という用語は、本明細書で使用される場合、処置される症状及び/または病態を有意にかつ良好に改変する（例えば、良好な臨床的反応を提供する）のに十分な薬学的組成物の量を意味する。薬学的組成物において使用するための活性成分の有効量は、処置されている特定の病態、病態の重症度、処置の継続期間、併用療法の特質、用いられる特定の活性成分（複数可）、利用される特定の薬学的に許容可能な賦形剤（複数可）及び/または担体（複数可）、及び同様の要因に応じて担当医の知識及び専門的技術を用いて変わる。

20

**【0188】**

「外因性物質」は、標的脂質粒子に関して本明細書で使用される場合、対応する野生型源細胞から作製された対応する野生型ウイルスまたはフソゲンに含まれず、コードもされない物質を指す。いくつかの実施形態では、外因性物質は、天然に存在せず、例えば、天然に存在するタンパク質に対して（例えば、挿入、付加、または置換によって）改変された配列を有するタンパク質または核酸である。いくつかの実施形態では、外因性物質は、源細胞に天然に存在しない。いくつかの実施形態では、外因性物質は、源細胞に天然に存在するが、ウイルスに対して外因性である。いくつかの実施形態では、外因性物質は、レシピエント細胞に天然に存在しない。いくつかの実施形態では、外因性物質は、レシピエント細胞に天然に存在するが、所望のレベルでまたは所望の時間で存在しない。いくつかの実施形態では、外因性物質は、RNAまたはタンパク質を含む。

30

**【0189】**

本明細書で使用される場合、「プロモーター」は、遺伝子コード配列に機能可能に連結された場合、遺伝子の転写を誘導するシス作用性DNA配列を指す。プロモーターは、転写因子結合部位を含み得る。いくつかの実施形態では、プロモーターは、遺伝子と遠位の1つ以上のエンハンサーと協働して作用する。

40

**【0190】**

本明細書で使用される場合、組成物は、細胞を含む、2つ以上の生成物、物質、または化合物の任意の混合物を指す。それは、溶液、懸濁液、液体、粉末、ペースト、水性、非水性またはそれらの任意の組み合わせであり得る。

**【0191】**

本明細書で使用される場合、「薬学的に許容可能な」という用語は、化合物の生物学的活性または特性を抑止せず、かつ比較的非毒性である担体または希釈剤などの物質を指し、すなわち、物質は、望ましくない生物学的効果を引き起こすことも、それが含有される組成物の成分のいずれとも有害な状態で相互作用することもなく、個体に投与され得る。

50

## 【0192】

本明細書で使用される場合、「薬学的組成物」という用語は、本発明の少なくとも1つの化合物と、他の化学的成分、例えば、担体、安定化剤、希釈剤、分散剤、懸濁剤、増粘剤、及び/または賦形剤との混合物を指す。薬学的組成物は、化合物の生物への投与を容易化する。静脈内、経口、エアロゾル、非経口、眼内、肺、及び局所投与を含むがこれらに限定されない化合物を投与する複数の技術が当該技術分野で存在する。

## 【0193】

「疾患」または「障害」は、本明細書で使用される場合、処置が必要とされる及び/または所望である状態を指す。

## 【0194】

本明細書で使用される場合、「処置する」、「処置すること」、または「処置」は、疾患または障害を改善すること、例えば、疾患または障害の進行を遅らせることもしくは停止させることもしくは減少させることまたはその臨床的症状の少なくとも1つを減少させることを指す。本開示の目的のため、疾患または障害を改善することは、1つ以上の症状の軽減、疾患の程度の減弱、疾患の拡大（例えば、転移、例えば、肺またはリンパ節への転移）を防止することまたは遅延させること、疾患の再発を防止することまたは遅延させること、疾患進行を遅延させることまたは遅らせること、疾患状態の改善、疾患または疾患の進行を阻害すること、疾患またはその進行の阻害することまたは遅らせること、その進展を停止させること、及び寛解（部分的または全体的にかかわらない）のうちの任意の1つ以上を含むがこれらに限定されない有益なまたは所望の臨床的結果を得ることを含み得る。

## 【0195】

「個体」及び「対象」という用語は、動物、例えば哺乳動物を指すために本明細書で互換的に使用される。患者という用語には、ヒト及び獣医学的对象が含まれる。いくつかの実施形態では、ヒト、齧歯類、サル、ネコ、イヌ、ウマ、ウシ、ブタ、ヒツジ、ヤギ、哺乳類実験動物、哺乳類牧場動物、哺乳類スポーツ動物、及び哺乳類ペットを含むがこれらに限定されない哺乳動物を処置する方法が提供される。対象は、男性または女性であり得、幼児、若年、青年、成人、及び老年対象を含む任意の好適な年齢であり得る。いくつかの例では、「個体」または「対象」は、疾患または障害のための処置を必要とする個体または対象を指す。いくつかの実施形態では、処置を受ける対象は、対象が処置に関連する傷害を有する、または障害にかかるのに十分なリスクがあるものとして特定されているという事実を指定する患者であり得る。特定の実施形態では、対象は、ヒト、例えば、ヒト患者である。

## 【0196】

II. 標的化脂質粒子（例えば、レンチウイルスベクター）

本明細書では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、(i)ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分及び(ii)結合ドメインを含む標的化エンベロープタンパク質とを含む、標的化脂質粒子であって、結合ドメインが、Gタンパク質または生物学的に活性な部分のC末端に結合されており、(i)及び(ii)の各々が、標的化脂質粒子の外側表面上に露出されている、標的化脂質粒子が提供される。いくつかの実施形態では、結合ドメインは、シングルドメイン抗体である。いくつかの実施形態では、結合ドメインは、一本鎖可変断片である。特定の実施形態では、提供される脂質粒子は、標的細胞に対する結合を容易化し、Gタンパク質またはその生物学的に活性な部分、ならびに脂質粒子及び標的細胞膜の2つの内腔の統合または融合の容易化に關与するF糖タンパク質を含有する標的化エンベロープタンパク質によって媒介される融合活性を示す。

## 【0197】

本明細書では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、(i)ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分及び(ii)シングルドメイン抗体(s d A b)可変ドメインを含む標

10

20

30

40

50



的化エンベロープタンパク質とを含む、標的化脂質粒子であって、シングルドメイン抗体が、Gタンパク質または生物学的に活性な部分のC末端に結合されており、(i)及び(ii)の各々が、標的化脂質粒子の外側表面上に露出されている、標的化脂質粒子が提供される。特定の実施形態では、提供される脂質粒子は、標的細胞に対する結合を容易化し、Gタンパク質またはその生物学的に活性な部分、ならびに脂質粒子及び標的細胞膜の2つの内腔の統合または融合の容易化に關与するF糖タンパク質を含有する標的化エンベロープタンパク質によって媒介される融合活性を示す。

【0198】

任意の実施形態のうちの一部では、標的化脂質粒子は、ウイルス粒子またはウイルス様粒子である。いくつかの態様では、そのような標的化脂質粒子は、ウイルス核酸、例えば、レトロウイルス核酸、例えばレンチウイルス核酸を含有する。特定の実施形態では、任意の提供される標的化脂質粒子、例えば、ウイルス粒子またはウイルス様粒子は、複製欠損性である。いくつかの実施形態では、標的化脂質粒子は、レンチウイルスベクターであり、レンチウイルスベクターは、ヘニパウイルスFタンパク質及び標的化エンベロープタンパク質で偽型化されている。

10

【0199】

例えば、本明細書では、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、(i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分及び(ii)結合ドメインを含む標的化エンベロープタンパク質とを含む、偽型化レンチウイルスベクターであって、結合ドメインが、Gタンパク質または生物学的に活性な部分のC末端に結合されており、(i)及び(ii)の各々が、標的化脂質粒子の外側表面上に露出されている、偽型化レンチウイルスベクターが提供される。いくつかの実施形態では、結合ドメインは、シングルドメイン抗体である。いくつかの実施形態では、結合ドメインは、一本鎖可変断片である。

20

【0200】

いくつかの実施形態では、本明細書で提供される標的化脂質粒子(例えば、標的化レンチウイルスベクター)は、同様のエンベロープタンパク質を組み込んでいるがそれが一本鎖可変断片(scFv)などのsdAb可変ドメイン以外の代替的標的化部位に融合されている参照脂質粒子(例えば、参照レンチウイルスベクター)と比較して増加したまたは多い標的化エンベロープタンパク質発現を有する。いくつかの実施形態では、そのような標的化脂質粒子は、トランスファー、エンベロープ、及びgag-polプラスミドでのパッケージング細胞のコトランスフェクションの後に脂質粒子(例えば、レンチウイルス粒子)の偽型化によって生成される。

30

【0201】

いくつかの実施形態では、発現は、参照脂質粒子(例えば、参照レンチウイルスベクター)、例えば、同様のエンベロープタンパク質を含有するがscFvに融合された参照脂質粒子と比較して、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%、300%、400%、500%以上またはそれを超えて増加する。いくつかの例では、発現は、参照脂質粒子(例えば、参照レンチウイルスベクター)、例えば、同様のエンベロープタンパク質を含有するがscFvに融合された参照脂質粒子と比較して、1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍以上またはそれを超えて増加する。いくつかの実施形態では、発現は、フローサイトメトリー、例えば、FACを使用してin vitroでアッセイされ得る。いくつかの実施形態では、発現は、標的化脂質粒子(例えば、標的化レンチウイルスベクター)の表面上の標的化エンベロープタンパク質の数または密度として示され得る。いくつかの実施形態では、発現は、標的化脂質粒子(例えば、標的化レンチウイルスベクター)の表面上の標的化エンベロープタンパク質の表面発現の平均蛍光強度(MFI)として示され得る。いくつかの実施形態では、発現は、標的化エンベロープタンパク質について表面陽性である集団における脂質粒子(例えば、レンチウイルスベクター)のパーセントとして示され得る。

40

50

## 【0202】

いくつかの実施形態では、標的化脂質粒子（例えば、標的化レンチウイルスベクター）の集団において、脂質粒子の50%または約50%超は、標的化エンベロープタンパク質について表面陽性である。例えば、提供される標的化脂質粒子（例えば、標的化レンチウイルスベクター）の集団において、集団における細胞の55%または約55%超、60%または約60%超、65%または約65%超、70%または約70%超、75%または約75%超が、標的化エンベロープタンパク質について表面陽性である。

## 【0203】

いくつかの実施形態では、形質導入（例えば、形質導入細胞）などによる標的細胞への導入後の標的化脂質粒子の力価は、同様のエンベロープタンパク質を組み込んでいるが一本鎖可変断片（s c F v）などのs d A b可変ドメイン以外の代替的標的化部位に融合されている参照脂質粒子（例えば、参照レンチウイルスベクター）の同じ標的細胞への力価と比較して増加する。典型的には、代替的標的化部位は、標的化脂質粒子の標的化エンベロープタンパク質のs d A b可変ドメインと同じ標的分子を認識し、またはそれと結合する。いくつかの実施形態では、力価は、参照脂質粒子（例えば、参照レンチウイルスベクター）、例えば、同様のエンベロープタンパク質を含有するがs c F vに融合された参照脂質粒子の力価と比較して、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%、300%、400%、500%以上またはそれを超えて増加する。いくつかの例では、力価は、参照脂質粒子（例えば、参照レンチウイルスベクター）、例えば、同様のエンベロープタンパク質を含有するがs c F vに融合された参照脂質粒子の力価と比較して、1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍、30倍以上またはそれを超えて増加する。いくつかの実施形態では、標的細胞（例えば、形質導入細胞）における標的化脂質粒子の力価は、 $1 \times 10^6$  形質導入単位（TU）/mLまたは約 $1 \times 10^6$  形質導入単位（TU）/mL超である。例えば、標的細胞（例えば、形質導入細胞）における標的化脂質粒子の力価は、 $2 \times 10^6$  TU/mLもしくは約 $2 \times 10^6$  TU/mL超、 $3 \times 10^6$  TU/mLもしくは約 $3 \times 10^6$  TU/mL超、 $4 \times 10^6$  TU/mLもしくは約 $4 \times 10^6$  TU/mL超、 $5 \times 10^6$  TU/mLもしくは約 $5 \times 10^6$  TU/mL超、 $6 \times 10^6$  TU/mLもしくは約 $6 \times 10^6$  TU/mL超、 $7 \times 10^6$  TU/mLもしくは約 $7 \times 10^6$  TU/mL超、 $8 \times 10^6$  TU/mLもしくは約 $8 \times 10^6$  TU/mL超、 $9 \times 10^6$  TU/mLもしくは約 $9 \times 10^6$  TU/mL超、または $1 \times 10^7$  TU/mLもしくは約 $1 \times 10^7$  TU/mL超である。

## 【0204】

A. 標的化エンベロープタンパク質（例えば、ヘニパウイルスと結合ドメイン）

いくつかの実施形態では、標的化脂質粒子（例えば、レンチウイルスベクター）には、標的化脂質粒子（例えば、レンチウイルスベクター）の表面上に露出した標的化エンベロープタンパク質が含まれる。

## 【0205】

いくつかの実施形態では、標的化エンベロープタンパク質は、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分及び標的細胞上の細胞表面分子に結合する結合ドメインを含有する。いくつかの実施形態では、結合ドメインは、シングルドメイン抗体（s d A b）である。いくつかの実施形態では、結合ドメインは、一本鎖可変断片（s c F v）である。結合ドメインは、Gタンパク質に直接的にまたは間接的に連結され得る。特定の実施形態では、結合ドメインは、Gタンパク質またはその生物学的に活性な部分のC末端（C末端アミノ酸）に連結される。連結は、ペプチドリンカー、例えば、フレキシブルペプチドリンカーを介するものであり得る。

## 【0206】

1. Gタンパク質

いくつかの実施形態では、標的化エンベロープタンパク質は、ヘニパウイルスエンベロープ結合糖タンパク質G（Gタンパク質）またはその生物学的に活性な部分及びシングル

10

20

30

40

50

ドメイン抗体 ( s d A b ) 可変ドメインまたはその生物学的に活性な部分を含有する。いくつかの実施形態では、 s d A b は、標的細胞上の細胞表面分子に結合する。 s d A b 可変ドメインは、 G タンパク質に直接的にまたは間接的に連結され得る。特定の実施形態では、 s d A b 可変ドメインは、 G タンパク質またはその生物学的に活性な部分の C 末端 ( C 末端アミノ酸 ) に連結される。連結は、ペプチドリンカー、例えば、フレキシブルペプチドリンカーを介するものであり得る。

【 0 2 0 7 】

いくつかの実施形態では、結合ドメイン ( 例えば、 s d A b ) は、細胞の細胞表面抗原に結合する。いくつかの実施形態では、細胞表面抗原は、 1 つのタイプの細胞に特徴的である。いくつかの実施形態では、細胞表面抗原は、複数のタイプの細胞に特徴的である。

10

【 0 2 0 8 】

いくつかの実施形態では、結合ドメイン ( 例えば、 s d A b ) 可変ドメインは、細胞表面分子または抗原に結合する。いくつかの実施形態では、細胞表面分子は、 A S G R 1、 A S G R 2、 T M 4 S F 5、 C D 8、 C D 4、または低密度リポタンパク質受容体 ( L D L - R ) である。いくつかの実施形態では、細胞表面分子は、 A S G R 1 である。いくつかの実施形態では、細胞表面分子は、 A S G R 2 である。いくつかの実施形態では、細胞表面分子は、 T M 4 S F 5 である。いくつかの実施形態では、細胞表面分子は、 C D 8 である。いくつかの実施形態では、細胞表面分子は、 C D 4 である。いくつかの実施形態では、細胞表面分子は、 L D L - R である。

【 0 2 0 9 】

いくつかの実施形態では、 G タンパク質は、ヘニパウイルス G タンパク質またはその生物学的に活性な部分である。いくつかの実施形態では、ヘニパウイルス G タンパク質は、ヘンドラ ( H e V ) ウイルス G タンパク質、ニパ ( N i V ) ウイルス G タンパク質 ( N i V - G )、セダー ( C e d P V ) ウイルス G タンパク質、 M o j i a n g ウイルス G タンパク質、コウモリパラミクソウイルス G タンパク質またはその生物学的に活性な部分である。表 3 は、 G タンパク質の非限定的な例を提供する。

20

【 0 2 1 0 】

結合 G タンパク質は、 N 末端細胞質尾部 ( 例えば、配列番号 9 のアミノ酸 1 ~ 4 9 に対応する )、膜貫通ドメイン ( 例えば、配列番号 9 のアミノ酸 5 0 ~ 7 0 に対応する )、及び細胞外茎部 ( 例えば、配列番号 9 のアミノ酸 7 1 ~ 1 8 7 に対応する ) を含有する細胞外ドメイン、及び球状頭部 ( 配列番号 9 のアミノ酸 1 8 8 ~ 6 0 2 に対応する ) を含む I I 型膜貫通糖タンパク質である。 N 末端細胞質ドメインは、脂質二重層の内側内腔内にあり、 C 末端部分は、脂質二重層の外側で露出した細胞外ドメインである。 C 末端領域における茎部の領域 ( 例えば、 N i V - G のアミノ酸 1 5 9 ~ 1 6 7 に対応する ) は、 F タンパク質との相互作用及び F タンパク質融合の誘発に関与することが示されている ( L i u e t a l . 2 0 1 5 J o f V i r o l o g y 8 9 : 1 8 3 8 )。野生型 G タンパク質において、球状頭部は、ヘニパウイルス侵入受容体エフリン B 2 及びエフリン B 3 に対する受容体結合を媒介するが、膜融合にとって不必要である ( B r a n d e l - T r e t h e w a y e t a l . J o u r n a l o f V i r o l o g y . 2 0 1 9 . 9 3 ( 1 3 ) e 0 0 5 7 7 - 1 9 )。本明細書における特定の実施形態では、 G タンパク質の指向性は、 s d A b 可変ドメインに対する G タンパク質またはその生物学的に活性な断片 ( 例えば、細胞質切断 ) の連結によって改変される。結合パートナーに対する G タンパク質の結合は、適合性 F タンパク質またはその生物学的に活性な部分によって媒介される融合を誘発し得る。本明細書に開示される G タンパク質配列は主に、翻訳の開始に必要とされる N 末端メチオニンを含む発現した配列として開示される。そのような N 末端メチオニンは、一般的に、翻訳と共にまたは翻訳後に切断されるので、本明細書に開示されるすべての G タンパク質配列についての成熟タンパク質配列はまた、 N 末端メチオニンを欠くものとして企図される。

30

40

【 0 2 1 1 】

G 糖タンパク質は、ヘニパウイルス種間で高度に保存されている。例えば、 N i V 及び

50

HeVウイルスのGタンパク質は、79%のアミノ酸同一性を共有する。研究は、異型融合活性化によって実証されるように、Gタンパク質の中で異なる種のFタンパク質との高い適合度を示した(Brandel-Tretheway et al. Journal of Virology. 2019)。以下にさらに記載されるように、再標的化脂質粒子は、異なる種からの異種のG及びFタンパク質を含有し得る。

【0212】

(表3)ヘニパウイルスタンパク質G配列クラスター

列1: Genbank IDには、クラスターの重心配列であるウイルスのゲノム配列全体のGenbank IDが含まれる。

列2: CDSのヌクレオチドは、ゲノム全体における遺伝子のCDSに対応するヌクレオチドを提供する。

列3: 完全遺伝子名であって、Genbank ID、ウイルス種、株、及びタンパク質名を含む遺伝子の完全な名称を提供する。

列4: 配列であって、遺伝子のアミノ酸配列を提供する。

列5: 配列数/クラスターであって、この重心配列とクラスター化する配列の数を提供する。

列6: 記載される配列についての配列番号を提供する。

Genbank ID	CDSのヌクレオチド	完全配列 ID	配列	#配列/クラスター	配列番号	配列番号(N末端メチオニンを有しない)
AF017149	8913-10727	gb:AF017149 生物:ヘンドラウイルス 株名:不明-AF017149 タンパク質名:糖タンパク質 遺伝子記号:G	MMADSKLVSLNNNLSGKIKDQGK VIKNYYGTMDIKKINDGLLDISKIL GAFNTVIALLGSIIVMNIMIIQNY TRTTDNQALIKESLQSVQQQIKAL TDKIGTEIGPKVSLIDTSSTITIPANI GLLGSKISQSTSSINENVNDKCKFT LPPLKIHECNISCPNPLPFREYRPIS QGVSDLVGLPNQICLQKTTSTILKP RLISYTLPI NTREGVCITDPLLAVD NGFFAYSHLEKIGSCTRGIAKQRIL GVGEVLD RGDKVPSMFMTNVWT PPNPSTIHHCSSYHEDFYITLCA VSHVGDPI LNSTSWTESLSLIRLAV RPKSDSGDYNQKYIAITKVERGKY DKVMPYGPSGIKQGD TLYFPAVG FLPRTEFQYNDSCPIIHCYKAE NCRLSMGVNSKSHYILRSGLLKY NLSLGGDIILQFIEIADNRLTIGSPS KIYNSLGQPVFYQASYSWDTMIKL GDVDTVDPLRVQWRNNSVISRPG	14	18	52

10

20

30

40

50

			QSQCPRFNVCEVCWEGTYNDAF LIDRLNWVSAGVYLNSNQTAENP VFAVFKDNEILYQVPLAEDDTNA QKTITDCFLLENIWCISLVEIYDT GDSVIRPKLFAVKIPAQCSES			
AF21 2302	8943- 10751	gb:AF212302  生物:ニパウ イルス 株名: 不明- AF212302 タ ンパク質名: 結合糖タン パク質 遺伝 子記号:G	MPAENKKVRFENTTSDKGKIPSKV IKSYYGTMDIKKINEGLLDSKILSA FNTVIALLGSIIVMNIMIIQNYTR STDNQAVIKDALQGIQQIKGLAD KIGTEIGPKVSLIDTSSTITIPANIGL LGSKISQSTASINENVNEKCKFTLP PLKIHECNISCPNPLPFREYRPQTE GVSNLVGLPNNICLQKTSNQILKP KLISYTLPPVVGQSGTCITDPLLAM DEGYFAYSHLERIGSCSRGVSKQR IIGVGEVLDRGDEVPSLFMTNVWT PPNPNTVYHCSAVYNNEFYVLC AVSTVGDPILNSTYWSGSLMMTR LAVKPKSNGGGYNQHQLALRSIE KGRYDKVMPYGPSGIKQGDLYF PAVGFLVRTEFKYNDNSNCPITKCQ YSKPENCRLSMGIRPNSHYILRSGL LKYNLSDGENPKVVFIEISDQRLSI GSPSKIYDSLQPVFYQASFSWDT MIKFGDVLTVNPLVVNWRNNTVI SRPGQSQCPRFNTCPEICWEGVYN DAFLIDRINWISAGVFLDSNQTAE NPVFTVFKDNEILYRAQLASEDTN AQKTITNCFLLNKIWCISLVEIYD TGDNVIRPKLFAVKIPEQCT	14	28	44
JQ001 776	8170- 10275	gb:JQ001776 :8170-10275  生物:セダー ウイルス 株 名:CG1a タ ンパク質名: 結合糖タン パク質 遺伝 子記号:G	MLSQLQKNYLDNSNQQGDKMNN PDKKLSVNFNPLELDKGQKDLNK SYYVKNKNYNVSNLLNESLHDIK FCIYCIFSLLIITIINIITISIVITRLKV HEENNGMESPNLQSIQDSLSSLTN MINTEITPRIGILVTATSVTLSSSIN YVGTKTNQLVNELKDYITKSCGF KVPELKLHECNISCADPKISKSAM YSTNAYAELAGPPKIFCKSVSKDP DFRLKQIDYVIPVQQDRSICMNNP	3	29	54

10

20

30

40

50

			LLDISDGFFTYIHYESGINSCKKSDS FKVLLSHGEIVDRGDYRPSLYLLS SHYHPYSMQVINCVPTCNQSSFV FCHISNNTKTLDNSDYSSDEYYIT YFNGIDRPKTKKIPINMTADNRY IHFTFSGGGGVCLGEEFIIPVTTVIN TDVFTHDYCESFNCSVQTGKSLKE ICSESLRSPNTSSRYNLNGIMIISQN NMTDFKIQLNGITYNKLSFGSPGR LSKTLGQVLYYQSSMSWDTYLKA GFVEKWKPFTPNWMNNTVISRPN QGNCPRYHKCEICYGGTYNDIAP LDLGKDMYVSVILDSQLAENPEI TVFNSTTILYKERVSKDELNTRSTT TSCFLFLDEPWCISVLETNRFNGKS IRPEIYSYKIPKYC			
NC_0 25256	9117- 11015	gb:NC_0252 56:9117- 11015 生物: コウモリパ ラミクソウ イルス Eid_hel/GH- M74a/GHA/2 009 株 名:BatPV/Ei d_hel/GH- M74a/GHA/2 009 タンパ ク質名:糖タ ンパク質 遺 伝子記号:G	MPQKTVEFINMNSPLERGVSTLSD KKTLNQSKITKQGYFGLGSHSERN WKKQKNQNDHYMTVSTMILEILV VLGIMFNLIVLTMVYYQNDNINQ RMAELTSNITVLNLNLNQLTNKIQ REIIPRITLIDTATTITIPSAITYLAT LTTRISELLPSINQKCEFPTLVL NDCRINCTPPLNPSDGVKMSSLAT NLVAHGSPPCRNFSVPTIYYRIP GLYNRTALDERCILNPRLTISSTKF AYVHSEYDKNCTRGFKYELMTF GEILEGPEKEPRMFSRSFYSPNAV NYHSCTPIVTVNEGYFLCLECTSS DPLYKANLSNSTFHLVILRHNKDE KIVSMPSFNLSTDQEYVQIIPAEGG GTAESGNLYFPCIGRLLHKRVTHP LCKKSNCSRTDDESLKSYYNQG SPQHQQVNCLIRIRNAQRDNPTW DVITVDLTNTYPGSRSRIFGSFSKP MLYQSSVSWHTLLQVAEITDLDK YQLDWLDTPYISRPGGSECPFGNY CPTVCWEGTYNDVYSLTPNNDLF	2	30	55

10

20

30

40

50

			VTVYLKSEQVAENPYFAIFSRDQI LKEFPLDAWISSARTTTISCFMFNN EIWCIAALEITRLNDDIIRPIIYYSFW LPTDCRTPYPHTGKMTRVPLRSTY NY			
NC_0 25352	8716- 11257	gb:NC_0253 52:8716- 11257 生 物:Mojiang ウイルス 株 名:Tongguan 1 タンパク 質名:結合糖 タンパク質  遺伝子記 号:G	MATNRDNTITSAEVSQEDKVKKY YGVETAEKVADSISGNKVFILMNT LLILTGAIITITLNITNLTAAKSQQN MLKIIQDDVNAKLEMFVNLDQLV KGEIKPKVSLINTAVSVSIPGQISNL QTKFLQKYVYLEESITKQCTCNPL SGIFPTSGPTYPPPTDKPDDDTTDDD KVDTTIKPIEYKPDGCNRTGDHF TMEPGANFYTVPNLGPASSNSDEC YTNPFSFSIGSSIYMFSSQEIRKTDCT AGEILSIQIVLGRIVDKGQQGPQAS PLLVWAVPNPKIINSCAVAAGDE MGWVLCVTLTAASGEPHMFDF GFWLYKLEPDTEVVSRYRITGYAY LLDKQYDSVFIGKGGGIQKGNL YFQMYGLSRNRQSFKALCEHGSC LGTGGGGYQVLCRAVMSFGSEE SLITNAYLKVNDLASGKPVIIIGQTF PPSDSYKGSNGRMYTIGDKYGLY LAPSSWNRYLRFGITPDISVRSTT WLKSQDPIMKILSTCTNTDRDMCP EICNTRGYQDIFPLSESEYYTYIGI TPNNGGTKNFVAVRDSGDGHASID ILQNYYSITSATISCFMYKDEIWC AITEGKKQKDNPQRIYAHSYKIRQ MCYNMKSATVTVGNAKNITIRRY	2	31	56

10

20

30

40

50

## 【 0 2 1 3 】

いくつかの実施形態では、Gタンパク質は、配列番号9、18、28、29、30、31、44、52、または54～56のいずれかに示される配列を有し、または配列番号9、18、28、29、30、31、44、52、もしくは54～56のいずれか1つと少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%同一である配列を有するその機能的に活性なバリエーションもしくは生物学的に活性な部分である。特定の実施形態では、Gタンパク質または機能的に活性なバリエーションもしくは生物学的に活性な部分は、ヘニパウイルスFタンパク質、例えば、セクションI.Bに示されるFタンパク質（例えば、NiV-Fまた

は H e V - F ) と併せて融合活性を保持するタンパク質である。融合活性には、2つの膜内腔、例えば、その脂質二重層に埋め込まれたヘニパウイルス F 及び G タンパク質を有する標的化脂質粒子の内腔、ならびに標的細胞、例えば、標的化エンベロープタンパク質によって認識されるまたは結合される表面受容体または分子を含有する細胞の細胞質の融合を促進または容易化するためのヘニパウイルス F タンパク質と併せた G タンパク質の活性が含まれる。いくつかの実施形態では、F タンパク質及び G タンパク質は、同じヘニパウイルス種からのものである (例えば、N i V - G 及び N i V - F )。いくつかの実施形態では、F タンパク質及び G タンパク質は、異なるヘニパウイルス種からのものである (例えば、N i V - G 及び H e V - F )。

#### 【 0 2 1 4 】

特定の実施形態では、G タンパク質は、配列番号 9、配列番号 28、配列番号 18、配列番号 30、配列番号 31、配列番号 44、配列番号 52 もしくは配列番号 54 ~ 56 に示されるアミノ酸の配列を有し、または融合活性を保持するその機能的に活性なバリエーションもしくはその生物学的に活性な部分である。いくつかの実施形態では、機能的に活性なバリエーションは、配列番号 9、配列番号 28、配列番号 18、配列番号 30、配列番号 31、配列番号 44、配列番号 52 または配列番号 54 ~ 56 と少なくとも 80% もしくは約 80%、少なくとも 85% しくは約 85%、少なくとも 90% しくは約 90%、少なくとも 91% しくは約 91%、少なくとも 92% しくは約 92%、少なくとも 93% しくは約 93%、少なくとも 94% しくは約 94%、少なくとも 95% しくは約 95%、96% しくは約 96%、少なくとも 97% しくは約 97%、少なくとも 98% しくは約 98%、または少なくとも 99% しくは約 99% の配列同一性を有するアミノ酸配列を含み、ヘニパウイルス F タンパク質 (例えば、N i V - F または H e V - F ) と併せて融合活性を保持する。いくつかの実施形態では、生物学的に活性な部分は、配列番号 9、配列番号 28、配列番号 18、配列番号 30 配列番号 31、配列番号 44、配列番号 52 または配列番号 54 ~ 56 と少なくとも 80% しくは約 80%、少なくとも 85% しくは約 85%、少なくとも 90% しくは約 90%、少なくとも 91% しくは約 91%、少なくとも 92% しくは約 92%、少なくとも 93% しくは約 93%、少なくとも 94% しくは約 94%、少なくとも 95% しくは約 95%、96% しくは約 96%、少なくとも 97% しくは約 97%、少なくとも 98% しくは約 98%、または少なくとも 99% しくは約 99% の配列同一性を有するアミノ酸配列を有し、ヘニパウイルス F タンパク質 (例えば、N i V - F または H e V - F ) と併せて融合活性を保持する。

#### 【 0 2 1 5 】

融合活性を保持することに対する言及には、配列番号 9、配列番号 28、配列番号 18、配列番号 30、配列番号 31、配列番号 44、配列番号 52 または配列番号 54 ~ 56 に示されるような対応する野生型 G タンパク質の結合のレベルまたは程度の 10% しくは約 10% から 150% しくは約 150% の間またはそれを超え、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 10%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 15%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 20%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 25%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 30%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 35%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 40%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 45%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 50%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 55%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは

10

20

30

40

50



程度の少なくともまたは少なくとも約 60%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 65%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 70%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 75%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 80%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 85%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 90%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 95%、例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 100%、または例えば、対応する野生型 G タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 120% である（ニパウイルス F タンパク質と併せた）活性が含まれる。

#### 【0216】

いくつかの実施形態では、G タンパク質は、1 つ以上のアミノ酸変異、例えば、1 つ以上のアミノ酸挿入、欠失、置換または切断を含有する機能的に活性なバリエーションまたは生物学的に活性な部分である変異体 G タンパク質である。いくつかの実施形態では、本明細書に記載の変異は、参照 G タンパク質配列と比較したアミノ酸のアミノ酸挿入、欠失、置換または切断に関する。いくつかの実施形態では、参照 G タンパク質配列は、G タンパク質またはその生物学的に活性な部分の野生型配列である。いくつかの実施形態では、その機能的に活性なバリエーションまたは生物学的に活性な部分は、野生型ヘンドラ (HeV) ウイルス G タンパク質、野生型ニパ (NiV) ウイルス G タンパク質 (NiV-G)、野生型セダー (CedPV) ウイルス G タンパク質、野生型 Mojiang ウイルス G タンパク質、野生型コウモリパラミクソウイルス G タンパク質またはその生物学的に活性な部分の変異体である。いくつかの実施形態では、野生型 G タンパク質は、配列番号 9、18、28、29、30、31、配列番号 44、配列番号 52 または配列番号 54 ~ 56 のいずれか 1 つに示される配列を有する。

#### 【0217】

いくつかの実施形態では、G タンパク質は、野生型ヘンドラ (HeV) ウイルス G タンパク質、野生型ニパ (NiV) ウイルス G タンパク質 (NiV-G)、野生型セダー (CedPV) ウイルス G タンパク質、野生型 Mojiang ウイルス G タンパク質、野生型コウモリパラミクソウイルス G タンパク質の N 末端及び / または C 末端で切断された断片である生物学的に活性な部分である変異体 G タンパク質である。特定の実施形態では、切断は、細胞質ドメインのすべてまたは一部の N 末端切断である。いくつかの実施形態では、変異体 G タンパク質は、切断型でありかつ配列番号 9、18、28、29、30、31、配列番号 44、配列番号 52 または配列番号 54 ~ 56 のいずれか 1 つに示される野生型 G タンパク質などの野生型 G タンパク質の N 末端におけるまたはその近くの最大で 49 個の連続アミノ酸残基を欠く生物学的に活性な部分である。いくつかの実施形態では、変異体 F タンパク質は、切断型であり、野生型 G タンパク質の N 末端における最大で 49 個の連続アミノ酸、例えば、最大で 49、48、47、46、45、44、43、42、41、40、30、38、37、36、35、34、33、32、31、30、29、28、27、26、25、24、23、22、21、20、19、18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2 または 1 個の連続アミノ酸を欠く。

#### 【0218】

いくつかの実施形態では、G タンパク質は、野生型ニパウイルス G (NiV-G) タンパク質もしくはヘンドラウイルス G タンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である。いくつかの実施形態では、G タンパク質は、配列番号 9、配列番号 28 または配列番号 44 に示される配列を有する NiV-G タンパク質であるか、または配列番号 9、配列番号 28 または配列番号 44 と少なくとも 80

%または約80%、少なくとも81%または約81%、少なくとも82%または約82%、少なくとも83%または約83%、少なくとも84%または約84%、少なくとも85%または約85%、少なくとも86%または約86%、少なくとも87%または約87%、少なくとも88%または約88%、少なくとも89%または約89%、少なくとも90%または約90%、少なくとも91%または約91%、少なくとも92%または約92%、少なくとも93%または約93%、少なくとも94%または約94%、少なくとも95%または約95%、96%または約96%、少なくとも97%または約97%、少なくとも98%または約98%、少なくとも99%または約99%の配列同一性を有するアミノ酸配列を有するその機能的パリアントまたは生物学的に活性な部分である。

【0219】

いくつかの実施形態では、Gタンパク質は、野生型NiV-Gの生物学的に活性な部分である変異体NiV-Gタンパク質である。いくつかの実施形態では、生物学的に活性な部分は、N末端で切断された断片である。いくつかの実施形態では、変異体NiV-Gタンパク質は、切断型であり、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で5個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で6個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で7個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で8個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で9個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で10個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で11個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で12個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で13個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で14個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で15個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で16個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で17個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で18個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で19個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で20個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で21個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で22個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で23個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で24個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で25個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端における

10

20

30

40

50

またはその近くの最大で26個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で27個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で28個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で29個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で30個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で31個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で32個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で33個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で34個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で35個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で36個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で37個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で38個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で39個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で40個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で41個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で42個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で43個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で44個の連続アミノ酸残基、または野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの最大で45個の連続アミノ酸残基（配列番号9、配列番号28または配列番号44）を欠く。

【0220】

いくつかの実施形態では、NiV-Gタンパク質は、細胞質ドメインを含有しない生物学的に活性な部分である。いくつかの実施形態では、細胞質ドメインを有しないNiV-Gタンパク質は、配列番号32によってコードされる。

【0221】

いくつかの実施形態では、変異体NiV-Gタンパク質は、配列番号10～15、35～40、45～50、22、53または配列番号32のいずれかに示される配列を含み、または配列番号10～15、35～40、45～50、22、53または配列番号32と少なくとも80%または80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくはは

約 99% の配列同一性を有するアミノ酸配列を有するその機能的バリエーションである。

【0222】

いくつかの実施形態では、変異体 NiV - Gタンパク質は、野生型 NiV - Gタンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 5 アミノ酸切断、例えば、配列番号 10 に示されるものまたは配列番号 10 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、  
 少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーション、または例えば、配列番号 35 に示されるものまたは配列番号 35 と少なくとも 80% もしくは約 80%、  
 少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、  
 少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、  
 少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーション、または例えば、配列番号 45 に示されるものまたは配列番号 45 と少なくとも 80% もしくは約 80%、  
 少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、  
 または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、  
 少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、  
 少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、  
 少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーションを有する。いくつかの実施形態では、変異体 NiV - Gタンパク質は、野生型 NiV - Gタンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 10 アミノ酸切断、例えば、配列番号 11 に示されるものまたは配列番号 11 と少なくとも 80% もしくは約 80%、  
 少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、  
 84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、  
 または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、  
 または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、  
 少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、  
 少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、  
 少なくとも 95% もしくは約 95%、96% もしくは約 96%、  
 少なくとも 97% もしくは約 97%、  
 少なくとも 98% もしくは約 98%、  
 または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーション、  
 または例えば、配列番号 36 に示されるものまたは配列番号 36 と  
 少なくとも 80% もしくは約 80%、  
 少なくとも 81% もしくは約 81%、  
 少なくとも 82% もしくは約 82%、  
 少なくとも 83% もしくは約 83%、  
 84% もしくは約 84%、  
 少なくとも 85% もしくは約 85%、  
 少なくとも 86% もしくは約 86%、  
 または少なくとも 87% もしくは約 87%、  
 少なくとも 88% もしくは約 88%、  
 または少なくとも 8

10

20

30

40

50

9%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーション、または例えば、配列番号46に示されるものまたは配列番号46と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーションを有する。

10

### 【0223】

いくつかの実施形態では、変異体NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの15アミノ酸切断、例えば、配列番号12に示されるものまたは配列番号12と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有するその機能的バリエーション、または例えば、配列番号37に示されるものまたは配列番号37と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーション、または例えば、配列番号47に示されるものまたは配列番号47と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーションを有する。

20

30

40

いくつかの実施形態では、変異体NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの

50

20 アミノ酸切断、例えば、配列番号13に示されるものまたは配列番号13と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーション、または例えば、配列番号38に示されるものまたは配列番号38と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーション、または例えば、配列番号48に示されるものまたは配列番号48と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーションを有する。

いくつかの実施形態では、変異体NiV-Gタンパク質は、野生型NiV-Gタンパク質（配列番号9、配列番号28または配列番号44）のN末端におけるまたはその近くの25アミノ酸切断、例えば、配列番号14に示されるものまたは配列番号14と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーション、または例えば、配列番号39に示されるものまたは配列番号39と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もし

くは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーション、または例えば、配列番号 49 に示されるものまたは配列番号 49 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーションを有する。

10

いくつかの実施形態では、変異体 NiV - G タンパク質は、野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 30 アミノ酸切断、例えば、配列番号 15 に示されるものまたは配列番号 15 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーション、または例えば、配列番号 40 に示されるものまたは配列番号 40 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーション、または例えば、配列番号 50 に示されるものまたは配列番号 50 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するその機能的バリエーションを有する。

20

30

40

いくつかの実施形態では、変異体 NiV - G タンパク質は、野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 34 アミノ酸切断、例えば、配列番号 22 に示されるものまたは配列番号 22 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87

50

%もしくは約 87%、少なくとも 88%もしくは約 88%、または少なくとも 89%もしくは約 89%、少なくとも 90%もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するその機能的バリエーション、または例えば、配列番号 53 に示されるものまたは配列番号 53 と少なくとも 80%もしくは約 80%、少なくとも 81%もしくは約 81%、少なくとも 82%もしくは約 82%、少なくとも 83%もしくは約 83%、84%もしくは約 84%、少なくとも 85%もしくは約 85%、少なくとも 86%もしくは約 86%、または少なくとも 87%もしくは約 87%、少なくとも 88%もしくは約 88%、または少なくとも 89%もしくは約 89%、少なくとも 90%もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するその機能的バリエーションを有する。

10

いくつかの実施形態では、変異体 HiV - G タンパク質は、配列番号 32 または配列番号 32 と少なくとも 80%もしくは約 80%、少なくとも 81%もしくは約 81%、少なくとも 82%もしくは約 82%、少なくとも 83%もしくは約 83%、84%もしくは約 84%、少なくとも 85%もしくは約 85%、少なくとも 86%もしくは約 86%、または少なくとも 87%もしくは約 87%、少なくとも 88%もしくは約 88%、または少なくとも 89%もしくは約 89%、少なくとも 90%もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するその機能的バリエーションに示されるように、野生型 HiV - G タンパク質 (配列番号 9、配列番号 28 または配列番号 44) の N 末端細胞質ドメインを欠く。

20

#### 【0224】

いくつかの実施形態では、変異体 G タンパク質は、配列番号 18 または 52 に示される配列を有する変異体 HeV - G タンパク質であるか、または配列番号 18 または 52 と少なくとも 80%もしくは約 80%、少なくとも 81%もしくは約 81%、少なくとも 82%もしくは約 82%、少なくとも 83%もしくは約 83%、少なくとも 84%もしくは約 84%、85%もしくは約 85%、少なくとも 86%もしくは約 86%、少なくとも 87%もしくは約 87%、88%もしくは約 88%、少なくとも 89%もしくは約 89%、少なくとも 90%もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するアミノ酸配列を有するその機能的バリエーションまたは生物学的に活性な部分である。

30

40

#### 【0225】

いくつかの実施形態では、G タンパク質は、野生型 HeV - G の生物学的に活性な部分である変異体 HeV - G タンパク質である。いくつかの実施形態では、生物学的に活性な部分は、N 末端で切断された断片である。いくつかの実施形態では、変異体 HeV - G タンパク質は、切断型であり、野生型 HeV - G タンパク質 (配列番号 18 または 52) の N 末端におけるまたはその近くの最大で 5 個の連続アミノ酸残基、野生型 HeV - G タンパク質 (配列番号 18 または 52) の N 末端におけるまたはその近くの最大で 6 個の連続アミノ酸残基、野生型 HeV - G タンパク質 (配列番号 18 または 52) の N 末端におけるまたはその近くの最大で 7 個の連続アミノ酸残基、野生型 HeV - G タンパク質 (配列

50





2個の連続アミノ酸残基、野生型HeV-Gタンパク質（配列番号18または52）のN末端におけるまたはその近くの最大で43個の連続アミノ酸残基、野生型HeV-Gタンパク質（配列番号18または52）のN末端におけるまたはその近くの最大で44個の連続アミノ酸残基、または野生型HeV-Gタンパク質（配列番号18または52）のN末端におけるまたはその近くの最大で45個の連続アミノ酸残基を欠く。いくつかの実施形態では、HeV-Gタンパク質は、細胞質ドメインを含有しない生物学的に活性な部分である。いくつかの実施形態では、変異体HeV-Gタンパク質は、配列番号33または配列番号33と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するその機能的バリエーションに示されるように、野生型HeV-Gタンパク質（配列番号18または52）のN末端細胞質ドメインを欠く。

10

#### 【0226】

いくつかの実施形態では、Gタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、エフリンB2またはエフリンB3に結合する。いくつかの態様では、Gタンパク質は、配列番号9、配列番号18または配列番号28、配列番号29、配列番号44、配列番号30もしくは配列番号31のいずれか1つに示されるアミノ酸の配列を有し、またはエフリンB2もしくはエフリンB3に結合することができるその機能的に活性なバリエーションもしくはその生物学的に活性な部分である。いくつかの実施形態では、機能的に活性なバリエーションまたは生物学的に活性な部分は、配列番号9、配列番号18または配列番号28、配列番号29、配列番号44、配列番号30または配列番号31と少なくとも約80%、少なくとも約85%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分を有し、エフリンB2またはB3に対する結合を保持する。エフリンB2またはB3に対する結合を保持することに対する言及には、配列番号9、配列番号18または配列番号28、配列番号29、配列番号44、配列番号30または配列番号31に示されるような対応する野生型Gタンパク質、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分の結合のレベルまたは程度の少なくともまたは少なくとも約5%、配列番号9、配列番号18または配列番号28、配列番号29、配列番号44、配列番号30または配列番号31に示されるような対応する野生型Gタンパク質、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分の結合のレベルまたは程度の10%、配列番号9、配列番号18または配列番号28、配列番号29、配列番号44、配列番号30または配列番号31に示されるような対応する野生型Gタンパク質、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分の結合のレベルまたは程度の15%、配列番号9、配列番号18または配列番号28、配列番号29、配列番号44、配列番号30または配列番号31に示されるような対応する野生型Gタンパク質、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分の結合のレベルまたは程度の20%、配列番号9、配列番号18または配列番号28、配列番号29、配列番号44、配列番号30または配列番号31に示されるような対応する野生型Gタンパク質、または機能的に活性なバリエーションもしくは生物学的に活性な部分の結合のレベルまたは程度の25%、配列番号9、配列

20

30

40

50



機能的に活性なバリエーションまたは生物学的に活性な部分は、配列番号 9、配列番号 28 または配列番号 44 と少なくとも約 80%、少なくとも約 85%、少なくとも約 90% もしくは約 90%、少なくとも約 91% もしくは約 91%、少なくとも約 92% もしくは約 92%、少なくとも約 93% もしくは約 93%、少なくとも約 94% もしくは約 94%、少なくとも約 95% もしくは約 95%、96% もしくは約 96%、少なくとも約 97% もしくは約 97%、少なくとも約 98% もしくは約 98%、または少なくとも約 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有し、エフリン B2 または B3 に対する結合を保持する。例示的な生物学的に活性な部分には、例えば、配列番号 10 ~ 15、35 ~ 40、45 ~ 50 及び 32 のいずれか 1 つに示される、細胞質ドメインのすべてまたは一部、例えば、1 つ以上、例えば、1 ~ 49 個の連続 N 末端アミノ酸残基を欠く N 末端切断バリエーションが含まれる。エフリン B2 または B3 に対する結合を保持することに対する言及には、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 5%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 10%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 15%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 20%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 25%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 30%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 35%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 40%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 45%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 50%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 55%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 60%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 65%、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の 70%、例えば、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 75%、例えば、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 80%、例えば、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 85%、例えば、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 90%、または例えば、配列番号 9、配列番号 28 または配列番号 44 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 95% である結合が含まれる。

#### 【0227】

いくつかの実施形態では、G タンパク質は、HeV - G またはその機能的に活性なバリエーションもしくは生物学的に活性な部分であり、エフリン B2 またはエフリン B3 に結合する。いくつかの態様では、HeV - G は、配列番号 18 もしくは 52 に示されるアミノ酸の配列を有し、またはエフリン B2 もしくはエフリン B3 に結合することができるその機能的に活性なバリエーションもしくはその生物学的に活性な部分である。いくつかの実施形態では、機能的に活性なバリエーションまたは生物学的に活性な部分は、配列番号 18 または 52 と少なくとも約 80%、少なくとも約 85%、少なくとも約 90% もしくは約 90%、少

なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有し、エフリン B2 または B3 に対する結合を保持する。例示的な生物学的に活性な部分には、例えば、配列番号 33 のいずれかが 1 つに示される、細胞質ドメインのすべてまたは一部、例えば、1 つ以上、例えば、1 ~ 49 個の連続 N 末端アミノ酸残基を欠く N 末端切断バリエーションが含まれる。エフリン B2 または B3 に対する結合を保持することに対する言及には、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 5%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 10%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 15%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 20%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 25%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 30%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 35%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 40%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 45%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 50%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 55%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 60%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 65%、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の 70%、例えば、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 75%、例えば、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 80%、例えば、配列番号 18 または 52 に示されるような対応する野生型 NiV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 85%、例えば、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 90%、または例えば、配列番号 18 または 52 に示されるような対応する野生型 HeV - G の結合のレベルまたは程度の少なくともまたは少なくとも約 95% である結合が含まれる。

【0228】

いくつかの実施形態では、G タンパク質またはその生物学的は、野生型 G タンパク質のネイティブ結合パートナーについて減少した結合を示す変異体 G タンパク質である。いくつかの実施形態では、変異体 G タンパク質またはその生物学的に活性な部分には、野生型 NiV - G の変異体であり、ネイティブ結合パートナーエフリン B2 またはエフリン B3 の一方または両方に対する減少した結合を示す。いくつかの実施形態では、変異体 G タンパク質または生物学的に活性な部分、例えば、変異体 NiV - G タンパク質は、ネイティブ結合パートナーに対する減少した結合を示す。いくつかの実施形態では、エフリン B2 またはエフリン B3 に対する減少した結合は、5% もしくは約 5%、10% もしくは約 10%、15% もしくは約 15%、20% もしくは約 20%、25% もしくは約 25%、30% もしくは約 30%、40% もしくは約 40%、50% もしくは約 50%、60% もしくは約 60%、70% もしくは約 70%、80% もしくは約 80%、90% もしくは約 90%、または 100% もしくは約 100% を超えて減少する。

【0229】

いくつかの実施形態では、本明細書に記載の変異は、形質導入効率を改善し得る。いくつかの実施形態では、本明細書に記載の変異は、エフリン B2 またはエフリン B3 ではな

い他の所望の細胞タイプの特異的標的化を可能とする。いくつかの実施形態では、本明細書に記載の変異は、少なくとも1つの天然受容体との結合を少なくとも部分的に不可能にし、例えば、エフリンB2またはエフリンB3のうちの少なくとも1つに対する結合を減少させた。いくつかの実施形態では、本明細書に記載の変異は、天然受容体認識と干渉する。

#### 【0230】

いくつかの実施形態では、Gタンパク質は、エフリンB2及びエフリンB3の一方または両方との相互作用に関与する残基における1つ以上のアミノ酸置換を含有する。いくつかの実施形態では、アミノ酸置換は、配列番号28に示される付番を参照して変異E501A、W504A、Q530A及びE533Aに対応する。

10

#### 【0231】

いくつかの実施形態では、Gタンパク質は、配列番号28に示される付番を参照してE501A、W504A、Q530A及びE533Aからなる群から選択される1つ以上のアミノ酸置換を含有する変異体Gタンパク質である。いくつかの実施形態では、Gタンパク質は、配列番号28を参照してE501A、W504A、Q530A及びE533Aからなる群から選択される1つ以上のアミノ酸置換を含有する変異体Gタンパク質であり、N末端切断を含有するその生物学的に活性な部分である。いくつかの実施形態では、変異体NiV-Gタンパク質またはその生物学的に活性な部分は、切断型であり、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの最大で5個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの6個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの7個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの8個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの9個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの最大で10個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの11個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの12個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの13個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの14個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの最大で15個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの16個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの17個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの18個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの19個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの最大で20個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの21個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの22個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの23個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの24個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの最大で25個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの26個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの27個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの28個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端におけるまたはその近くの29個の連続アミノ酸残基、野生型NiV-Gタンパク質(配列番号28)のN末端

20

30

40

50

におけるまたはその近くの最大で30個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの最大で31個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの32個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの33個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの34個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの35個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの最大で36個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの最大で37個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの最大で38個の連続アミノ酸残基、野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの最大で39個の連続アミノ酸残基、または野生型NiV-Gタンパク質（配列番号28）のN末端におけるまたはその近くの最大で40個の連続アミノ酸残基を欠く。

#### 【0232】

いくつかの実施形態では、変異体NiV-Gタンパク質は、配列番号16もしくは51に示されるアミノ酸配列、または配列番号16もしくは51と少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する。特定の实施形態では、Gタンパク質は、配列番号16または51に示されるアミノ酸の配列を有する。

#### 【0233】

いくつかの実施形態では、標的化エンベロープタンパク質は、Gタンパク質または機能的に活性なバリエーションもしくは生物学的に活性な部分及びsdAb可変ドメインを含有し、標的化エンベロープタンパク質は、野生型Gタンパク質のネイティブ結合パートナーとは異なる別の分子について増加した結合を示す。いくつかの実施形態では、分子は、所望の標的細胞の表面上に発現したタンパク質であり得る。いくつかの実施形態では、他の分子に対する増加した結合は、25%もしくは約25%、30%もしくは約30%、40%もしくは約40%、50%もしくは約50%、60%もしくは約60%、70%もしくは約70%、80%もしくは約80%、90%もしくは約90%、または100%もしくは約100%を超えて増加する。特定の实施形態では、結合は、野生型Gタンパク質の結合と比較して再標的化された結合を付与し、その場合、新たなまたは異なる結合活性が付与される。

#### 【0234】

##### 2. 結合ドメイン

いくつかの実施形態では、結合ドメインは、標的細胞上の細胞表面分子に結合する任意の物質であり得る。いくつかの実施形態では、結合ドメインは、抗体または抗体部分または断片であり得る。

#### 【0235】

結合ドメインは、異なる結合強度を有するように調節され得る。例えば、様々な結合強度を有するscFv及び抗体が、多いまたは少ない量の標的抗原を示す細胞に対するキメラ結合タンパク質の融合活性を改変するために使用され得る。例えば、異なる親和性を有するDARPinが、多いまたは少ない量の標的抗原を示す細胞に対する融合活性を改変するために使用され得る。結合ドメインはまた、標的リガンド上の異なる領域を標的とするように調節され得、これは、標的を提示する細胞との融合速度に影響を及ぼす。

#### 【0236】

結合ドメインは、ヒト化抗体分子、インタクトIgA、IgG、IgEまたはIgM抗体；二重または多重特異性抗体（例えば、ZyboDies（登録商標）など）；抗体断

片、例えば、Fab断片、Fab'断片、F(ab')<sub>2</sub>断片、Fd'断片、Fd断片、及び単離されたCDRまたはそのセット；一本鎖Fvs；ポリペプチド-Fc融合体；シングルドメイン抗体（例えば、サメシングルドメイン抗体、例えば、IgNARまたはその断片）；ラクダ様抗体；マスク化抗体（例えば、Probodies（登録商標））；小モジュール免疫薬剤（「SMIPSTM」）；一本鎖またはタンデムダイアボディ（TandAb（登録商標））；VHH；Anticalins（登録商標）；Nanobodies（登録商標）；ミニボディ；BiTE（登録商標）；アンキリンリピートタンパク質またはDARPIN（登録商標）；Avimer（登録商標）；DART；TCR様抗体；Adnectin（登録商標）；Affilin（登録商標）；Trans-bodies（登録商標）；Affibodies（登録商標）；TrimerX（登録商標）；MicroProteins；Fynomers（登録商標）、Centyrins（登録商標）；及びKALBITOR（登録商標）を含み得る。標的化部位にはまた、抗体またはその抗原結合断片（例えば、Fab、Fab'、F(ab')<sub>2</sub>、Fv断片、scFv抗体断片、ジスルフィド連結したFv(sdFv)、VH及びCH1ドメインからなるFd断片、線状抗体、シングルドメイン抗体、例えば、sdAb（VLまたはVHのいずれか）、ナノボディ、またはラクダ科動物VHHドメイン）、抗原結合フィブロネクチンIII型（Fn3）骨格、例えば、フィブロネクチンポリペプチドミニボディ、リガンド、サイトカイン、ケモカイン、またはT細胞受容体（TCR）が含まれ得る。

10

#### 【0237】

いくつかの実施形態では、結合ドメインは、一本鎖分子である。いくつかの実施形態では、結合ドメインは、シングルドメイン抗体である。いくつかの実施形態では、結合ドメインは、一本鎖可変断片である。特定の実施形態では、結合ドメインは、ヒトまたはヒト化された抗体可変配列（複数可）を含有する。

20

#### 【0238】

いくつかの実施形態では、結合ドメインは、シングルドメイン抗体である。いくつかの実施形態では、シングルドメイン抗体は、ヒトまたはヒト化であり得る。いくつかの実施形態では、シングルドメイン抗体またはその部分は、天然に存在する。いくつかの実施形態では、シングルドメイン抗体またはその部分は、合成である。

#### 【0239】

いくつかの実施形態では、シングルドメイン抗体は、相補性決定領域がシングルドメインポリペプチドの一部である抗体である。いくつかの実施形態では、シングルドメイン抗体は、重鎖のみの抗体可変ドメインである。いくつかの実施形態では、シングルドメイン抗体は、軽鎖を含まない。

30

#### 【0240】

いくつかの実施形態では、軽鎖を欠く重鎖抗体は、VHHと称される。いくつかの実施形態では、シングルドメイン抗体は、12～15kDaの分子量を有する。いくつかの実施形態では、シングルドメイン抗体には、ラクダ科動物抗体またはサメ抗体が含まれる。いくつかの実施形態では、シングルドメイン抗体分子は、ラクダ種、例えば、ラクダ、リヤマ、ヒトコブラクダ、アルパカ、ビクーナ及びグアナコにおいて樹立された抗体に由来する。いくつかの実施形態では、シングルドメイン抗体は、免疫グロブリン新抗原受容体（IgNAR）と称され、軟骨魚類に由来する。いくつかの実施形態では、シングルドメイン抗体は、ヒトまたはマウスIgGの二量体可変ドメインを単量体に分割し、重要な残基をラクダ化することによって生成される。

40

#### 【0241】

いくつかの実施形態では、シングルドメイン抗体は、ファージディスプレイライブラリから生成され得る。いくつかの実施形態では、ファージディスプレイライブラリは、Arbabi et al., FEBS Letters, 414, 521-526 (1997)、Lauwereys et al., EMBO J., 17, 3512-3520 (1998)、Decanniere et al., Structure, 7, 361-370 (1999)に記載されているように、様々な抗原で免疫されたラクダ科動物V

50



HHレパートリーから生成される。いくつかの実施形態では、ファージディスプレイライブラリは、非免疫ラクダ科動物の抗体断片を含めて生成される。いくつかの実施形態では、ヒトシングルドメイン抗体のシングルドメイン抗体ライブラリは、1つ以上の骨格に多様性を導入することによって合成的に生成される。

【0242】

いくつかの実施形態では、シングルドメイン抗体のC末端は、Gタンパク質またはその生物学的に活性な部分のC末端に結合されている。いくつかの実施形態では、シングルドメイン抗体のN末端は、脂質二重層の外部表面上に露出されている。いくつかの実施形態では、シングルドメイン抗体のN末端は、標的細胞の細胞表面分子に結合する。いくつかの実施形態では、シングルドメイン抗体は、標的細胞上に存在する細胞表面分子に特異的に結合する。いくつかの実施形態では、細胞表面分子は、タンパク質、グリカン、脂質または低分子量分子である。

10

【0243】

いくつかの実施形態では、標的細胞の細胞表面分子は、抗原またはその一部である。いくつかの実施形態では、シングルドメイン抗体またはその一部は、特定の抗原に選択的に結合することができる単一の単量体ドメイン抗原結合/認識ドメインを有する抗体である。いくつかの実施形態では、シングルドメイン抗体は、標的細胞上に存在する抗原と結合する。

【0244】

例示的な細胞には、多形核細胞(PMN、PML、PMNL、または顆粒球としても知られている)、幹細胞、胚性幹細胞、神経幹細胞、間葉幹細胞(MSC)、造血幹細胞(HSC)、ヒト筋原幹細胞、筋肉由来幹細胞(Mustem)、胚性幹細胞(ESまたはESC)、縁上皮幹細胞、心筋原幹細胞、心筋細胞、前駆細胞、免疫エフェクター細胞、リンパ球、マクロファージ、樹状細胞、ナチュラルキラー細胞、T細胞、細胞傷害性Tリンパ球、同種異系細胞、常在性心臓細胞、誘導多能性幹細胞(iPS)、アジポーズ由来または表現型修飾幹または前駆細胞、CD133+細胞、アルデヒドデヒドロゲナーゼ陽性細胞(ALDH+)、臍帯血液(UCB)細胞、末梢血幹細胞(PBSC)、ニューロン、神経前駆細胞、膵臓ベータ細胞、グリア細胞、または肝細胞が含まれる。

20

【0245】

いくつかの実施形態では、標的細胞は、標的組織の細胞である。標的組織には、肝臓、肺、心臓、脾臓、膵臓、消化管、腎臓、精巣、卵巣、脳、生殖器官、中枢神経系、末梢神経系、骨格筋、内皮、内耳、または眼が含まれ得る。

30

【0246】

いくつかの実施形態では、標的細胞は、筋肉細胞(例えば、骨格筋細胞)、腎臓細胞、肝臓細胞(例えば、肝細胞)、または心臓細胞(例えば、心筋細胞)である。いくつかの実施形態では、標的細胞は、心臓細胞、例えば、心筋細胞(例えば、静止心筋細胞)、肝芽細胞(例えば、胆管肝芽細胞)、上皮細胞、T細胞(例えば、ナイーブT細胞)、マクロファージ(例えば、腫瘍浸潤マクロファージ)、または線維芽細胞(例えば、心臓線維芽細胞)である。

【0247】

いくつかの実施形態では、標的細胞は、腫瘍浸潤リンパ球、T細胞、新生物または腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系(CNS)細胞、造血幹細胞(HSC)、肝臓細胞または完全に分化した細胞である。いくつかの実施形態では、標的細胞は、CD3+ T細胞、CD4+ T細胞、CD8+ T細胞、肝細胞、造血幹細胞、CD34+造血幹細胞、CD105+造血幹細胞、CD117+造血幹細胞、CD105+内皮細胞、B細胞、CD20+ B細胞、CD19+ B細胞、がん細胞、CD133+がん細胞、EpCAM+がん細胞、CD19+がん細胞、Her2/Neu+がん細胞、GluA2+ニューロン、GluA4+ニューロン、NKGD2+ナチュラルキラー細胞、SLC1A3+星状細胞、SLC7A10+脂肪細胞、またはCD30+肺上皮細胞である。

40

50

【0248】

いくつかの実施形態では、標的細胞は、抗原提示細胞、MHCクラスII+細胞、プロフェッショナル抗原提示細胞、非定型抗原提示細胞、マクロファージ、樹状細胞、骨髄樹状細胞、形質細胞様樹状細胞、CD11c+細胞、CD11b+細胞、脾細胞、B細胞、肝細胞、内皮細胞、または非がん細胞)である。

【0249】

いくつかの実施形態では、細胞表面分子は、CD8、CD4、アシアログリコプロテイン受容体2(ASGR2)、膜貫通4L6ファミリーメンバー5(TM4SF5)、低密度リポタンパク質受容体(LDLR)またはアシアログリコプロテイン1(ASGR1)のいずれか1つである。

10

【0250】

いくつかの実施形態では、Gタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、sdAb可変ドメインに直接的に連結される。いくつかの実施形態では、標的化エンベロープタンパク質は、以下の構造:(N'-シングルドメイン抗体-C')-(C'-Gタンパク質-N')を有する融合タンパク質である。

【0251】

いくつかの実施形態では、Gタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、リンカーを介してsdAb可変ドメインに間接的に連結される。いくつかの実施形態では、リンカーは、ペプチドリンカーである。いくつかの実施形態では、リンカーは、化学的リンカーである。

20

【0252】

いくつかの実施形態では、リンカーは、ペプチドリンカーであり、標的化エンベロープタンパク質は、ペプチドリンカーを介してsdAb可変ドメインに連結されたGタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分を含む融合タンパク質である。いくつかの実施形態では、標的化エンベロープタンパク質は、以下の構造:(N'-シングルドメイン抗体-C')-リンカー-(C'-Gタンパク質-N')を有する融合タンパク質である。

【0253】

いくつかの実施形態では、ペプチドリンカーは、長さが最大で65アミノ酸である。いくつかの実施形態では、ペプチドリンカーは、約2~65アミノ酸、2~60アミノ酸、2~56アミノ酸、2~52アミノ酸、2~48アミノ酸、2~44アミノ酸、2~40アミノ酸、2~36アミノ酸、2~32アミノ酸、2~28アミノ酸、2~24アミノ酸、2~20アミノ酸、2~18アミノ酸、2~14アミノ酸、2~12アミノ酸、2~10アミノ酸、2~8アミノ酸、2~6アミノ酸、6~65アミノ酸、6~60アミノ酸、6~56アミノ酸、6~52アミノ酸、6~48アミノ酸、6~44アミノ酸、6~40アミノ酸、6~36アミノ酸、6~32アミノ酸、6~28アミノ酸、6~24アミノ酸、6~20アミノ酸、6~18アミノ酸、6~14アミノ酸、6~12アミノ酸、6~10アミノ酸、6~8アミノ酸、8~65アミノ酸、8~60アミノ酸、8~56アミノ酸、8~52アミノ酸、8~48アミノ酸、8~44アミノ酸、8~40アミノ酸、8~36アミノ酸、8~32アミノ酸、8~28アミノ酸、8~24アミノ酸、8~20アミノ酸、8~18アミノ酸、8~14アミノ酸、8~12アミノ酸、8~10アミノ酸、10~65アミノ酸、10~60アミノ酸、10~56アミノ酸、10~52アミノ酸、10~48アミノ酸、10~44アミノ酸、10~40アミノ酸、10~36アミノ酸、10~32アミノ酸、10~28アミノ酸、10~24アミノ酸、10~20アミノ酸、10~18アミノ酸、10~14アミノ酸、10~12アミノ酸、12~65アミノ酸、12~60アミノ酸、12~56アミノ酸、12~52アミノ酸、12~48アミノ酸、12~44アミノ酸、12~40アミノ酸、12~36アミノ酸、12~32アミノ酸、12~28アミノ酸、12~24アミノ酸、12~20アミノ酸、12~18アミノ酸、12~14アミノ酸、14~65アミノ酸、14~60アミノ酸、14~56アミノ酸、14~52アミノ酸、14~48アミノ酸、14~44アミノ酸、14~40アミノ酸、14

30

40

50

~ 3 6 アミノ酸、 1 4 ~ 3 2 アミノ酸、 1 4 ~ 2 8 アミノ酸、 1 4 ~ 2 4 アミノ酸、 1 4  
 ~ 2 0 アミノ酸、 1 4 ~ 1 8 アミノ酸、 1 8 ~ 6 5 アミノ酸、 1 8 ~ 6 0 アミノ酸、 1 8  
 ~ 5 6 アミノ酸、 1 8 ~ 5 2 アミノ酸、 1 8 ~ 4 8 アミノ酸、 1 8 ~ 4 4 アミノ酸、 1 8  
 ~ 4 0 アミノ酸、 1 8 ~ 3 6 アミノ酸、 1 8 ~ 3 2 アミノ酸、 1 8 ~ 2 8 アミノ酸、 1 8  
 ~ 2 4 アミノ酸、 1 8 ~ 2 0 アミノ酸、 2 0 ~ 6 5 アミノ酸、 2 0 ~ 6 0 アミノ酸、 2 0  
 ~ 5 6 アミノ酸、 2 0 ~ 5 2 アミノ酸、 2 0 ~ 4 8 アミノ酸、 2 0 ~ 4 4 アミノ酸、 2 0  
 ~ 4 0 アミノ酸、 2 0 ~ 3 6 アミノ酸、 2 0 ~ 3 2 アミノ酸、 2 0 ~ 2 8 アミノ酸、 2 0  
 ~ 2 6 アミノ酸、 2 0 ~ 2 4 アミノ酸、 2 4 ~ 6 5 アミノ酸、 2 4 ~ 6 0 アミノ酸、 2 4  
 ~ 5 6 アミノ酸、 2 4 ~ 5 2 アミノ酸、 2 4 ~ 4 8 アミノ酸、 2 4 ~ 4 4 アミノ酸、 2 4  
 ~ 4 0 アミノ酸、 2 4 ~ 3 6 アミノ酸、 2 4 ~ 3 2 アミノ酸、 2 4 ~ 3 0 アミノ酸、 2 4  
 ~ 2 8 アミノ酸、 2 8 ~ 6 5 アミノ酸、 2 8 ~ 6 0 アミノ酸、 2 8 ~ 5 6 アミノ酸、 2 8  
 ~ 5 2 アミノ酸、 2 8 ~ 4 8 アミノ酸、 2 8 ~ 4 4 アミノ酸、 2 8 ~ 4 0 アミノ酸、 2 8  
 ~ 3 6 アミノ酸、 2 8 ~ 3 4 アミノ酸、 2 8 ~ 3 2 アミノ酸、 3 2 ~ 6 5 アミノ酸、 3 2  
 ~ 6 0 アミノ酸、 3 2 ~ 5 6 アミノ酸、 3 2 ~ 5 2 アミノ酸、 3 2 ~ 4 8 アミノ酸、 3 2  
 ~ 4 4 アミノ酸、 3 2 ~ 4 0 アミノ酸、 3 2 ~ 3 8 アミノ酸、 3 2 ~ 3 6 アミノ酸、 3 6  
 ~ 6 5 アミノ酸、 3 6 ~ 6 0 アミノ酸、 3 6 ~ 5 6 アミノ酸、 3 6 ~ 5 2 アミノ酸、 3 6  
 ~ 4 8 アミノ酸、 3 6 ~ 4 4 アミノ酸、 3 6 ~ 4 0 アミノ酸、 4 0 ~ 6 5 アミノ酸、 4 0  
 ~ 6 0 アミノ酸、 4 0 ~ 5 6 アミノ酸、 4 0 ~ 5 2 アミノ酸、 4 0 ~ 4 8 アミノ酸、 4 0  
 ~ 4 4 アミノ酸、 4 4 ~ 6 5 アミノ酸、 4 4 ~ 6 0 アミノ酸、 4 4 ~ 5 6 アミノ酸、 4 4  
 ~ 5 2 アミノ酸、 4 4 ~ 4 8 アミノ酸、 4 8 ~ 6 5 アミノ酸、 4 8 ~ 6 0 アミノ酸、 4 8  
 ~ 5 6 アミノ酸、 4 8 ~ 5 2 アミノ酸、 5 0 ~ 6 5 アミノ酸、 5 0 ~ 6 0 アミノ酸、 5 0  
 ~ 5 6 アミノ酸、 5 0 ~ 5 2 アミノ酸、 5 4 ~ 6 5 アミノ酸、 5 4 ~ 6 0 アミノ酸、 5 4  
 ~ 5 6 アミノ酸、 5 8 ~ 6 5 アミノ酸、 5 8 ~ 6 0 アミノ酸、 または 6 0 ~ 6 5 アミノ酸  
 を含む。いくつかの実施形態では、ペプチドリンカーは、長さが 3、4、5、6、7、8  
 、9、10、11、12、13、14、15、16、17、18、19、20、21、2  
 2、23、24、25、26、27、28、29、30、31、32、33、34、35  
 、36、37、38、39、40、41、42、43、44、45、46、47、48、  
 49、50、51、52、53、54、55、56、57、58、59、60、61、6  
 2、63、64、または65アミノ酸であるポリペプチドである。

10

20

30

【0254】

特定の実施形態では、リンカーは、フレキシブルペプチドリンカーである。いくつかの  
 そのような実施形態では、リンカーは、1~20アミノ酸、例えば、主にグリシンから構  
 成される1~20アミノ酸である。いくつかの実施形態では、リンカーは、1~20アミ  
 ノ酸、例えば、主にグリシン及びセリンから構成される1~20アミノ酸である。いくつ  
 かの実施形態では、リンカーは、GS-リンカーと称される、アミノ酸グリシン及びセリ  
 ンを含有するフレキシブルペプチドリンカーである。いくつかの実施形態では、ペプチド  
 リンカーには、配列GS、GGS、GGGS(配列番号43)、GGGG(配列番  
 号41)またはそれらの組み合わせが含まれる。いくつかの実施形態では、ポリペプチド  
 リンカーは、配列(GGS)n(nは1~10である)を有する。いくつかの実施形態で  
 は、ポリペプチドリンカーは、配列(GGGGS)n、(配列番号42)(nは1~10  
 である)を有する。いくつかの実施形態では、ポリペプチドリンカーは、配列(GGGG  
 GS)n(配列番号27)(nは1~6である)を有する。

40

【0255】

3. ポリヌクレオチド

本明細書では、標的化エンペロプタンパク質をコードする核酸配列を含むポリヌクレ  
 オチドが提供される。いくつかの実施形態では、ポリヌクレオチドは、Gタンパク質また  
 はその生物学的に活性な部分をコードする核酸配列を含む。いくつかの実施形態では、ポ  
 リヌクレオチドは、シングルドメイン抗体(sdAb)可変ドメインまたはその生物学的  
 に活性な部分をコードする核酸配列をさらに含む。ポリヌクレオチドには、上述した標的  
 化エンペロプタンパク質のいずれかをコードするヌクレオチドの配列が含まれ得る。ポ

50

リヌクレオチドは、合成核酸であり得る。提供されるポリヌクレオチドのいずれかを含有する発現ベクターも提供される。

【0256】

任意の実施形態のうちの一部では、天然または合成核酸の発現は、典型的には、対象となる遺伝子をコードする核酸をプロモーターに機能可能に連結し、コンストラクトを発現ベクターに組み込むことによって達成される。いくつかの実施形態では、ベクターは、真核生物における複製及び組み込みに好適であり得る。いくつかの実施形態では、クローニングベクターは、所望の核酸配列の発現に有用な転写及び翻訳ターミネーター、開始配列、及びプロモーターを含有する。任意の実施形態のうちの一部では、プラスミドは、細胞における発現に好適なプロモーターを含む。

10

【0257】

いくつかの実施形態では、ポリヌクレオチドは、Gタンパク質及びシングルドメイン抗体(s d A b)可変ドメインを含有する標的化エンベロープタンパク質の発現をコントロールするために機能可能に連結された少なくとも1つのプロモーターを含有する。標的化エンベロープタンパク質の発現のため、各プロモーターにおける少なくとも1つのモジュールが、RNA合成のための開始部位を配置するように機能する。この最も知られている例は、TATAボックスであるが、TATAボックスを欠くいくつかのプロモーター、例えば、哺乳動物末端デオキシヌクレオチジルトランスフェラーゼ遺伝子のためのプロモーター及びSV40遺伝子のためのプロモーターでは、開始部位自体に重なっている別の要素が、開始の箇所を固定させるのに役立つ。

20

【0258】

いくつかの実施形態では、追加のプロモーター要素、例えば、エンハンサーが、転写開始の頻度を制御する。いくつかの実施形態では、追加のプロモーター要素は、開始部位の30~110bp上流の領域に位置するが、いくつかのプロモーターは、開始部位の下流にも同様に機能的要素を含有することが最近示された。いくつかの実施形態では、プロモーター要素間の間隔はしばしばフレキシブルであるため、プロモーター機能は、要素が互いに反転または移動した場合に維持される。いくつかの実施形態では、チミジンキナーゼ(tk)プロモーター、プロモーター要素間の間隔は、活性が低下し始める前に50bp間隔まで増加され得る。いくつかの実施形態では、プロモーターに応じて、個々の要素は、協働的にまたは独立してのいずれかで機能して転写を活性化し得る。

30

【0259】

プロモーターは、コードセグメント及び/またはエクソンの上流に位置する5'非コード配列を単離することによって得られ得るように、遺伝子またはポリヌクレオチド配列に天然に関連するものであり得る。そのようなプロモーターは、「内因性」と称され得る。同様に、エンハンサーは、その配列の下流または上流のいずれかに位置するポリヌクレオチド配列に天然に関連するものであり得る。代替的には、その天然環境においてポリヌクレオチド配列に通常は関連しないプロモーターを指す、組換えまたは異種プロモーターのコントロール下でコードポリヌクレオチドセグメントを配置することによって所定の利点を得られる。組換えまたは異種エンハンサーはまた、その天然環境におけるポリヌクレオチド配列に通常関連しないエンハンサーを指す。そのようなプロモーターまたはエンハンサーには、他の遺伝子のプロモーターまたはエンハンサー、及び任意の他の原核生物、ウイルス、または真核細胞から単離されたプロモーターまたはエンハンサー、及び「天然に存在」しない、すなわち、異なる転写制御領域の異なる要素、及び/または発現を改変する変異を含有するプロモーターまたはエンハンサーが含まれ得る。プロモーター及びエンハンサーの核酸配列を合成的に生成することに加えて、配列は、本明細書に開示される組成物と組み合わせて、組換えクローニング及び/またはPCRを含む核酸増幅技術を使用して生成され得る(米国特許番号4,683,202及び5,928,906)。

40

【0260】

いくつかの実施形態では、好適なプロモーターは、最初期サイトメガロウイルス(CMV)プロモーター配列である。いくつかの実施形態では、プロモーター配列は、それに機

50

能可能に連結された任意のポリヌクレオチド配列の高レベルの発現を誘導することが可能な強力な構成的プロモーター配列である。いくつかの実施形態では、好適なプロモーターは、伸長成長因子-1 $\alpha$  (EF-1 $\alpha$ )である。いくつかの実施形態では、サルウイルス40 (SV40) 初期プロモーター、マウス乳房腫瘍ウイルス (MMTV)、ヒト免疫不全ウイルス (HIV) 長い末端反復 (LTR) プロモーター、MoMuLVプロモーター、トリ白血病ウイルスプロモーター、エプスタイン・バーウイルス最初期プロモーター、ラウス肉腫ウイルスプロモーター、ならびにヒト遺伝子プロモーター、例えば、限定されないが、アクチンプロモーター、ミオシンプロモーター、ヘモグロビンプロモーター、及びクレアチンキナーゼプロモーターを含むがこれらに限定されない他の構成的プロモーター配列も使用され得る。

10

## 【0261】

いくつかの実施形態では、プロモーターは、誘導性プロモーターである。いくつかの実施形態では、誘導性プロモーターは、そのような発現が所望である場合に機能可能に連結されたポリヌクレオチド配列の発現を作動させ、または発現が所望ではない場合に発現を停止することが可能な分子スイッチを提供する。いくつかの実施形態では、誘導性プロモーターは、メタロチオニンプロモーター、グルココルチコイドプロモーター、プロゲステロンプロモーター、及びテトラサイクリンプロモーターを含む。

## 【0262】

いくつかの実施形態では、外因的にコントロールされる誘導性プロモーターは、Gタンパク質及びシングルドメイン抗体 (sdAb) 可変ドメインの発現を制御するために使用され得る。例えば、放射線誘導性プロモーター、熱誘導性プロモーター、及び/または薬物誘導性プロモーターを、例えば、標的とされる領域における導入遺伝子発現を選択的に誘導するために使用することができる。そのような実施形態では、導入遺伝子発現の位置、継続期間、及びレベルは、誘導の外因性源の投与によって制御され得る。

20

## 【0263】

いくつかの実施形態では、Gタンパク質及びシングルドメイン抗体 (sdAb) 可変ドメインを含有する標的化エンベロープタンパク質の発現は、薬物誘導性プロモーターを使用して制御される。例えば、いくつかの場合では、プロモーター、エンハンサー、またはトランス活性化因子は、Lacオペレーター配列、テトラサイクリンオペレーター配列、ガラクトースオペレーター配列、ドキシサイクリンオペレーター配列、ラパマイシンオペレーター配列、タモキシフェンオペレーター配列、もしくはホルモン反応性オペレーター配列、またはそれらのアナログを含む。いくつかの例では、誘導性プロモーターは、テトラサイクリン反応要素 (TRE) を含む。いくつかの実施形態では、誘導性プロモーターは、タモキシフェンの存在下で遺伝子発現を活性化し得るエストロゲン反応要素 (ERE) を含む。いくつかの例では、薬物誘導性要素、例えば、TREは、薬物、例えば、ドキシサイクリンの存在下で転写を向上させるために選択されたプロモーターと組み合わせられ得る。いくつかの実施形態では、薬物誘導性プロモーターは、小分子誘導性プロモーターである。

30

## 【0264】

提供されるポリヌクレオチドのいずれも、特定の種、例えば、ヒト、イヌ、ネコ、ウマ、ヒツジ、ウシなどの種における翻訳のためにCpGモチーフを除去するために及び/またはコドン最適化のために修飾され得る。いくつかの実施形態では、ポリヌクレオチドは、ヒトコドン使用頻度について最適化される (すなわち、ヒトコドン最適化される)。いくつかの実施形態では、ポリヌクレオチドは、CpGモチーフを除去するように修飾される。他の実施形態では、提供されるポリヌクレオチドは、CpGモチーフを除去するように修飾され、コドン最適化、例えば、ヒトコドン最適化される。コドン最適化ならびにCpGモチーフ検出及び修飾の方法は、よく知られている。典型的には、ポリヌクレオチド最適化は、導入遺伝子発現を向上させ、導入遺伝子安定性を増加させ、コードされるポリペプチドのアミノ酸配列を保存する。

40

## 【0265】

50

標的化エンベロープタンパク質の発現を評価するために、細胞に導入される発現ベクターはまた、発現する粒子、例えば、ウイルス粒子の特定及び選択を容易化するために選択マーカー遺伝子またはレポーター遺伝子のいずれかまたはその両方を含有し得る。他の実施形態では、選択マーカーは、DNAの分離断片上に搭載され、コトランスフェクション手順において使用され得る。選択マーカー及びレポーター遺伝子の両方は、宿主細胞における発現を可能にするために適切な制御配列に隣接し得る。有用な選択マーカーは、当該技術分野で知られており、例えば、抗生物質耐性遺伝子、例えば、neoなどが含まれる。

#### 【0266】

レポーター遺伝子は、トランスフェクションされた可能性のある細胞を特定するために及び制御配列の機能性を評価するために使用される。容易にアッセイ可能なタンパク質をコードするレポーター遺伝子は、当該技術分野でよく知られている。通常、レポーター遺伝子は、レシピエント生物または組織では存在しないまたは発現せず、かつ発現がいくつかの容易に検出可能な特性、例えば、酵素活性によって示されるタンパク質をコードする遺伝子である。レポーター遺伝子の発現は、DNAがレシピエント細胞に導入された後の好適な時点でアッセイされる。

#### 【0267】

好適なレポーター遺伝子には、ルシフェラーゼ、ベータ-ガラクトシダーゼ、クロラムフェニコールアセチルトランスフェラーゼ、分泌アルカリホスファターゼをコードする遺伝子、または緑色蛍光タンパク質遺伝子が含まれ得る（例えば、U i - T e i e t a l . , 2 0 0 0 , F E B S L e t t . 4 7 9 : 7 9 - 8 2 を参照されたい）。好適な発現システムはよく知られており、よく知られている技術を使用して調製され、または商業的に入手され得る。内部欠失コンストラクトは、独自内部制限部位を使用してまたは非独自制限部位の部分的消化によって生成され得る。次いでコンストラクトは、高レベルの所望のポリヌクレオチド及び/またはポリペプチド発現を示す細胞にトランスフェクションされ得る。通常、レポーター遺伝子の最も高いレベルの発現を示す最小5 フランキング領域を有するコンストラクトは、プロモーターとして識別される。そのようなプロモーター領域は、レポーター遺伝子に連結され、プロモーター誘導転写を調節する能力について薬剤を評価するために使用され得る。

#### 【0268】

B . フソゲン（例えば、ヘニパウイルスFタンパク質）

いくつかの実施形態では、標的化脂質粒子は、1つ以上のフソゲンを含む。いくつかの実施形態では、標的化脂質粒子は、外因性のまたは過剰発現したフソゲンを含有する。いくつかの実施形態では、フソゲンは、脂質二重層に配置される。いくつかの実施形態では、フソゲンは、標的化脂質粒子の膜への融合を容易化する。いくつかの実施形態では、膜は、血漿細胞膜である。

#### 【0269】

いくつかの実施形態では、フソゲンは、タンパク質ベース、脂質ベース、及び化学ベースのフソゲンを含む。いくつかの実施形態では、標的化脂質粒子は、タンパク質フソゲンを含む第1のフソゲン及び脂質フソゲンまたは化学フソゲンを含む第2のフソゲンを含む。いくつかの実施形態では、フソゲンは、標的細胞表面上のフソゲン結合パートナーと結合する。

#### 【0270】

いくつかの実施形態では、フソゲンは、疎水性融合ペプチドドメインを有するタンパク質を含む。いくつかの実施形態では、フソゲンは、ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分を含む。いくつかの実施形態では、ヘニパウイルスFタンパク質は、ヘンドラ（H e v）ウイルスFタンパク質、ニパ（N i V）ウイルスF-タンパク質、セダー（C e d P V）ウイルスFタンパク質、M o j i a n g ウイルスFタンパク質またはコウモリパラミクスウイルスFタンパク質またはその生物学的に活性な部分である。

10

20

30

40

50

## 【0271】

表4は、Fタンパク質の非限定的な例を提供する。いくつかの実施形態では、Fタンパク質分子またはその生物学的に活性な部分のN末端疎水性融合ペプチドドメインは、脂質二重層の外側に露出されている。

## 【0272】

ヘニパウウイルスのFタンパク質は、シグナルペプチド（例えば、配列番号1のアミノ酸残基1～26に対応する）を含有するF<sub>0</sub>前駆体としてコードされる。シグナルペプチドの切断後、成熟F<sub>0</sub>（例えば、配列番号2）は、細胞表面に輸送され、次いで取り込まれ、カテプシンLによって（配列番号1のアミノ酸109～110の間で）成熟融合サブユニットF1（例えば、配列番号1のアミノ酸110～546に対応、配列番号4に示される）及びF2（例えば、配列番号1のアミノ酸残基27～109に対応、配列番号3に示される）に切断される。F1及びF2サブユニットは、ジスルフィド結合によって会合し、細胞表面に戻されてリサイクルされる。F1サブユニットは、融合を誘導するために細胞膜に挿入することができる、F1サブユニットのN末端に位置する融合ペプチドドメイン（例えば、配列番号1のアミノ酸110～129に対応する）を含有する。特定の場合では、Gが標的分子と係合し、Fからのその解離及び膜融合を媒介するための融合ペプチドの曝露をもたらすまで、融合活性は、Fタンパク質とGタンパク質との会合によってブロックされる。

10

## 【0273】

異なるヘニパウウイルス種間で、Fタンパク質の配列及び活性は、高度に保存されている。例えば、NiV及びHeVウイルスのFタンパク質は、89%のアミノ酸配列同一性を共有する。さらに、いくつかの場合では、ヘニパウウイルスFタンパク質は、融合を誘発するための他の種からのGタンパク質との適合性を示す（Brandel-Tretheway et al. Journal of Virology. 2019. 93(13): e00577-19）。いくつかの態様または提供される再標的化脂質粒子において、Fタンパク質は、Gタンパク質に対して異種であり、すなわち、F及びGタンパク質または生物学的に活性な部分は、異なるヘニパウウイルス種からのものである。例えば、Fタンパク質は、ヘンドラウイルスからのものであり、Gタンパク質は、ニパウウイルスからのものである。他の態様では、Fタンパク質は、ヘニパウウイルスの異なる種からのFタンパク質の領域を含有するキメラFタンパク質であり得る。いくつかの実施形態では、Fタンパク質のアミノ酸残基の領域をヘニパウウイルスのある種から別のものに交換すると、アミノ酸挿入を含む種のGタンパク質への融合をもたらし得る。（Brandel-Tretheway et al. 2019）。いくつかの場合では、キメラFタンパク質は、あるヘニパウウイルス種からの細胞外ドメイン及び異なるヘニパウウイルス種からの膜貫通及び/または細胞質ドメインを含有する。例えば、Fタンパク質は、ヘンドラウイルスの細胞外ドメイン及びニパウウイルスの膜貫通/細胞質ドメインを含有する。本明細書に開示されるFタンパク質配列は主に、N末端シグナル配列を含む発現した配列として開示される。そのようなN末端シグナル配列は、一般的に、翻訳と共にまたは翻訳後に切断されるので、本明細書に開示されるすべてのFタンパク質配列についての成熟タンパク質配列はまた、N末端シグナル配列を欠くものとして企図される。

20

30

40

## 【0274】

（表4）ヘニパウウイルスタンパク質F配列クラスター

列1：Genbank IDには、クラスターの重心配列であるウイルスのゲノム配列全体のGenbank IDが含まれる。

列2：CDSのヌクレオチドは、ゲノム全体における遺伝子のCDSに対応するヌクレオチドを提供する。

列3：完全遺伝子名であって、Genbank ID、ウイルス種、株、及びタンパク質名を含む遺伝子の完全な名称を提供する。ニパウウイルスFタンパク質は、ヘンドラウイルスのものと80%超同一であり、同じ配列クラスター内で見られる。

列4：配列であって、遺伝子のアミノ酸配列を提供する。

50

列 5 : 配列数 / クラスターであって、この重心配列とクラスター化する配列の数を提供する。

列 6 : 記載される配列についての配列番号を提供する。

Genbank ID	CD のスクレオチド	完全遺伝子名	配列	#配列 / クラスター	配列番号	配列番号 (シグナル配列を有しない)
AF017149	6618-8258	gb:AF017149 生物:ヘンドラウイルス 株名:不明-AF017149 タンパク質名:融合 遺伝子記号:F	MATQEVRLKCLLCGIIVLVLSLEGLGILH YEKLSKIGLVKGITRKYKIKSNPLTKDIVI KMIPNVSNSKCTGTVMENYKSRLTGILS PIKGAIELYNNNTHDLVGDVKLAVVMA GIAIGIATAAQITAGVALYEAMKNADNIN KLKSSSIESTNEAVVVLQETAEKTVYVLTA LQDYINTNLVPTIDQISCKQTEALDLALS KYLSDLLFVFGPNLQDPVSNMSTIQAISQ AFGGNYETLLRTLGYATEDFDDLLESDSI	29	17	59

10

20

30

40

50



			AGQIVYVDLSSYYIIVRVYFPILTEIQQAY VQELLPVSFNNDNSEWISIVPNFVLRNTL ISNIEVKYCLITKKSVCINQDYATPMTASV RECLTGSTDKCPRELVVSSHVPRFALSGG VLFANCISVTCQCQTTGRAISQSGETLL MIDNTTCTTVVLGNIIISLGKYLGSINYNS ESIAVGPPVYTDKVDISSQISSMNQSLQQS KDYIKEAQKILDVTNPSLISMLSMIILYVL SIAALCIGLITFISFVIVEKKRGNYSRLDDR QVRPVSNGDLYYIGT			
Q9 IH 63	クラスター における追 加: sp Q9IH63 F US_NIPAV 融合糖タン パク質 F0 OS=ニパウ イルス	MVVILDKRCYCNLLILMISECSVGILHY EKLSKIGLVKGVTRKYKIKSNPLTKDIVIK MIPNVSNSMQCTGSVMENYKTRLNGILT PIKGALEIYKNNTHDLVGDVRLAGVIMA GVAIGIATAAQITAGVALYEAMKNADNI NKLKSSIESTNEAVVKLQETAECTVYVLT ALQDYINTNLVPTIDKISCKQTELSLDLAL SKYLSDLLFVFGPNLQDPVNSMTIQAISQ AFGGNYETLLRTLGYATEDFDDLLESDSI TGQIIVYVDLSSYYIIVRVYFPILTEIQQAYI QELLPVSFNNDNSEWISIVPNFILVRNTLIS NIEIGFCLITKRSVCINQDYATPMTNMR ECLTGSTEKCPRELVVSSHVPRFALSNGV LFANCISVTCQCQTTGRAISQSGETLLMI DNTTCPTAVLGNVIISLGKYLGSVNYNSE GIAIGPPVFTDKVDISSQISSMNQSLQQSK DYIKEAQRLDVTNPSLISMLSMIILYVLSI ASLCIGLITFISFIIVEKKRNTYSRLEDRRV RPTSSGDLYYIGT		1	2	

10

20

30

40

50

JQ 00 17 76	6129 - 8166	gb:JQ001776: 6129-8166 生 物:セダーウ イルス 株 名:CG1a タ ンパク質名: 融合糖タン パク質 遺伝 子記号:F	MSNKRTTVLIIISYTLFYLNNAIVGFDFD KLNKIGVVQGRVLNYKIKGDPMTKDLVL KFIPNIVNITECVREPLSRYNETVRLLLLPI HNMLGLYLNNTNAKMTGLMIAGVIMGG IAIGIATAAQITAGFALYEAKKNTENIQKL TDSIMKTQDSIDKLTDSVGTSLILNKLQT YINNQLVPNLELLSCRQNKIEFDLMLTKY LVDLMTVIGPNINNPVNKDMTIQSLSLF DGNYDIMMSELGYTPQDFLDLIESKSITG QIIYVDMENLYVVIRTYLPTLIEVPDAQIY EFNKITMSSNGGEYLSTIPNFILIRGNYS NIDVATCYMTKASVICNQDYSLPMSQNL RSCYQGETEYCPVEAVIASHSPRFALTNG VIFANCINTICRCQDNGKTITQINQFVSM IDNSTCNDVMVDKFTIKVGYMGRKDIN NINIQIGPQIHDKVDLSNEINKMNQSLKDS IFYLREAKRILDSVNISLISPSVQLFIIISVL SFIILLIIIVLYCKSKHSYKYNKFIDDPDY YNDYKRERINGKASKSNNIYYVGD	3	24	57
NC _0 25 35 2	5950 - 8712	gb:NC_02535 2:5950-8712  生物:Mojiang ウイルス 株 名:Tongguan 1 タンパク 質名:融合タ ンパク質 遺 伝子記号:F	MALNKNMFSSLFLGYLLVYATTVQSSIH YDSLKVGVIKGLTYNYKIKGSPSTKLMV VKLIPNIDSVKNCTQKQYDEYKNLVRKA LEPVKMAIDTMLNNVKSGNNKYRFAGAI MAGVALGVATAATVTAGIALHRSNENA QAIANMKSIAQNTNEAVKQLQLANKQTL AVIDTIRGEINNNIIPVINQLSCDTIGLSVGI RLTQYYSEIITAFGPALQNPVNTRITQAI SVFNGNFDELLKIMGYTSGDLYEILHSELI RGNIIDVDVDAGYIALEIEFPNLTLV VQELMPISYNIDGDEWVTLVPRFVLRTRT LLSNIDTSRCTITDSSVICDNDYALPMSHE LIGCLQGDTSKCAREKVVSSYVPKFALSD GLVYANCLNTICRCMDTDTPIQSGLGATV SLLDNKRCSVYQVGDVLISVGSYLGDE YNADNVELGPPVIDKIDIGNQLAGINQTL QEAEDYIEKSEEFLKGVNPSIITLGSMVVL YIFMILIAIVSVIALVLSIKLTVKGNVVRQ QFTYTQHVPSPMENINYVSH	2	25	60

10

20

30

40

50

NC _0 25 25 6	6865 - 8853	gb:NC_02525 6:6865-8853  生物:コウモ リパラミク ソウイルス Eid_hel/GH- M74a/GHA/2 009 株 名:BatPV/Eid _hel/GH- M74a/GHA/2 009 タンパ ク質名:融合 タンパク質  遺伝子記 号:F	MKKKTDNPTISKRGHNHSRGIKSRALLRE TDNYSNGLIVENLVRNCHHPSKNNLNYT KTQKRSTIPYRVEERKGHPKIKHLIDK SYKHIKRGKRRNGHNGNIITILLILILKT QMSEGAIHYETLSKIGLIKITREYKVKGT PSSKDIVIKLIPNVTGLNKCTNISMENYKE QLDKILIPINNIELYANSTKSAPGNARFAG VIIAGVALGVAAAAQITAGIALHEARQNA ERINLLKDSISATNNAVAELQEATGGIVN VITGMQDYINTNLVPQIDKLQCSQIKTAL DISLSQYYSEILTVFGPNLQNPVTTSMSIQ AISQSFGGNIDLLNLLGYTANDLLDLLES KSITGQITYINLEHYFMVIRVYYPIMTTISN AYVQELIKISFNVDGSEWVSLVPSYILIRN SYLSNIDISECLITKNSVICRHFAMPMSY TLKECLTGDTEKCPREAVVTSYVPRFAIS GGVIYANCLSTTCQCYQTGK VIAQDGSQ TLMMDNQTCSIVRIEILISTGKYLGSQE YNTMHVSVGNPVFTDKLDITSQISNINQSI EQSKFYLDKSKAILDKINLNLIGSVPISILFI IALSLILSIITFVIVMIIVRRYNYKYTPLINSD PSSRRSTIQDVYIIPNPGEHSIRSAARSIDR DRD	2	26	58
---------------------------	-------------------	--	---	---	----	----

10

20

## 【 0 2 7 5 】

いくつかの実施形態では、Fタンパク質は、配列番号1、2、17、24、25、26  
または57～60のいずれか1つによって示される配列をコードするヌクレオチド配列に  
よってコードされ、または配列番号1、2、17、24、25、26または57～60の  
いずれか1つと少なくとも80%もしくは約80%、少なくとも85%もしくは約85%  
、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも  
92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは  
は約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、  
少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なく  
とも99%もしくは約99%同一である配列を有するその機能的に活性なバリエーション  
もしくは生物学的に活性な部分である。特定の実施形態では、Fタンパク質またはその機能的  
に活性なバリエーションもしくは生物学的に活性な部分は、ヘニパウイルスGタンパク質、例  
えば、セクションI . Aに示されるGタンパク質（例えば、NiV-GまたはHeV-G  
）と併せて融合活性を保持する。融合活性には、2つの膜内腔、例えば、その脂質二重層  
に埋め込まれたヘニパウイルスF及びGタンパク質を有する標的化脂質粒子の内腔、なら  
びに標的細胞、例えば、標的化エンベロープタンパク質によって認識されるまたはそれが  
結合する表面受容体または分子を含有する細胞の細胞質の融合を促進または容易化するた  
めのヘニパウイルスGタンパク質と併せたFタンパク質の活性が含まれる。いくつかの実  
施形態では、Fタンパク質及びGタンパク質は、同じヘニパウイルス種からのものである  
（例えば、NiV-G及びNiV-F）。いくつかの実施形態では、Fタンパク質及びG  
タンパク質は、異なるヘニパウイルス種からのものである（例えば、NiV-G及びHe  
V-F）。特定の実施形態では、Fタンパク質またはその機能的に活性なバリエーション  
もしくは生物学的に活性な部分は、カテプシンLによって切断される切断部位（例えば、配  
列番号1のアミノ酸109～110の間の切断部位に対応する）を保持する。

30

40

50

## 【 0 2 7 6 】

特定の実施形態では、Fタンパク質は、配列番号1、配列番号2、配列番号17、配列番号24、配列番号25、配列番号26、配列番号57、配列番号57、配列番号58、配列番号59、もしくは配列番号60に示されるアミノ酸の配列を有し、または融合活性を保持するその機能的に活性なバリエーションもしくはその生物学的に活性な部分である。いくつかの実施形態では、機能的に活性なバリエーションは、配列番号1、配列番号2、配列番号17、配列番号24、配列番号25、配列番号26、配列番号57、配列番号57、配列番号58、配列番号59、または配列番号60と少なくとも80%もしくは約80%、少なくとも85%もしくは約85%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含み、ヘニパウイルスGタンパク質（例えば、NiV-GまたはHeV-G）と併せて融合活性を保持する。いくつかの実施形態では、生物学的に活性な部分は、配列番号1、配列番号2、配列番号17、配列番号24、配列番号25、配列番号26、配列番号57、配列番号57、配列番号58、配列番号59、または配列番号60と少なくとも80%もしくは約80%、少なくとも85%もしくは約85%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有し、ヘニパウイルスGタンパク質（例えば、NiV-GまたはHeV-G）と併せて融合活性を保持する。

10

20

## 【 0 2 7 7 】

融合活性を保持することに対する言及には、配列番号1、配列番号2、配列番号17、配列番号24、配列番号25、配列番号26、配列番号57、配列番号57、配列番号58、配列番号59、または配列番号60に示されるような対応する野生型Fタンパク質の結合のレベルまたは程度の10%もしくは約10%から150%もしくは約150%の間またはそれを超え、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約10%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約15%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約20%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約25%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約30%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約35%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約40%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約45%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約50%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約55%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約60%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約65%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約70%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約75%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約80%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約85%、例えば、対応する野生型Fタンパク質の融合活性のレベルまたは程度の

30

40

50

少なくともまたは少なくとも約 90%、例えば、対応する野生型 F タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 95%、例えば、対応する野生型 F タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 100%、または例えば、対応する野生型 F タンパク質の融合活性のレベルまたは程度の少なくともまたは少なくとも約 120%である（ヘニパウイルス G タンパク質と併せた）活性が含まれる。

【0278】

いくつかの実施形態では、F タンパク質は、1つ以上のアミノ酸変異、例えば、1つ以上のアミノ酸挿入、欠失、置換または切断を含有する機能的に活性な断片または生物学的に活性な部分である変異体 F タンパク質である。いくつかの実施形態では、本明細書に記載の変異は、参照 F タンパク質配列と比較したアミノ酸のアミノ酸挿入、欠失、置換または切断に関する。いくつかの実施形態では、参照 F タンパク質配列は、F タンパク質またはその生物学的に活性な部分の野生型配列である。いくつかの実施形態では、変異体 F タンパク質またはその生物学的に活性な部分は、野生型ヘンドラ (Hev) ウイルス F タンパク質、ニバ (NiV) ウイルス F - タンパク質、セダー (CedPV) ウイルス F タンパク質、Mojiang ウイルス F タンパク質またはコウモリパラミクソウイルス F タンパク質の変異体である。いくつかの実施形態では、野生型 F タンパク質は、配列番号 1、2、17、24、25、26、または 57 ~ 60 のいずれか 1つをコードするヌクレオチドの配列によってコードされる。

【0279】

いくつかの実施形態では、変異体 F タンパク質は、N 末端及び/または C 末端で切断された断片である野生型 F タンパク質の生物学的に活性な部分である。いくつかの実施形態では、変異体 F タンパク質、またはその野生型 F タンパク質の生物学的に活性な部分は、1つ以上のアミノ酸置換を含む。いくつかの実施形態では、本明細書に記載の変異は、形質導入効率を改善し得る。いくつかの実施形態では、本明細書に記載の変異は、融合能力を増加させ得る。例示的な変異には、記載されている任意のものが含まれ、例えば、Khetawat and Broder 2010 Virology Journal 7:312、Witting et al. 2013 Gene Therapy 20:997-1005、公開国際、特許出願番号 WO/2013/148327 を参照されたい。

【0280】

いくつかの実施形態では、変異体 F タンパク質は、切断型でありかつ配列番号 1、17、24、25 または 26 のいずれか 1つに示される F タンパク質をコードするヌクレオチドの配列によってコードされる野生型 F タンパク質などの野生型 F タンパク質の C 末端におけるまたはその近くの最大で 20 個の連続アミノ酸残基を欠く生物学的に活性な部分である。いくつかの実施形態では、変異体 F タンパク質は、切断型であり、野生型 F タンパク質の C 末端における最大で 19 個の連続アミノ酸、例えば、最大で 18、17、16、15、14、13、12、11、10、9、8、7、6、5、4、3、2、または 1 個の連続アミノ酸を欠く。

【0281】

いくつかの実施形態では、F タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、F1 サブユニットまたはその融合部分を含む。いくつかの実施形態では、F1 サブユニットは、F0 前駆体のタンパク質分解的に切断された部分である。いくつかの実施形態では、F0 前駆体は、不活性である。いくつかの実施形態では、F0 前駆体の切断は、ジスルフィド連結 F1 + F2 ヘテロ二量体を形成する。いくつかの実施形態では、切断は、融合ペプチドを露出させ、成熟 F タンパク質を生成する。いくつかの実施形態では、切断は、単一の塩基性残基でまたはその周りで生じる。いくつかの実施形態では、切断は、NiV - F タンパク質のアルギニン 109 で生じる。いくつかの実施形態では、切断は、ヘンドラウイルス F タンパク質のリジン 109 で生じる。

【0282】

10

20

30

40

50

いくつかの実施形態では、Fタンパク質は、野生型ニパウイルスF(NiV-F)タンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である。いくつかの実施形態では、F<sub>0</sub>前駆体は、配列番号1に示される配列をコードするヌクレオチドの配列によってコードされる。コードする核酸は、配列M V V I L D K R C Y C N L L I L I L M I S E C S V G(配列番号34)を有するシグナルペプチド配列をコードし得る。いくつかの実施形態では、Fタンパク質は、配列番号2に示される配列を有する。いくつかの例では、Fタンパク質は、配列番号4に示される配列を含むF1サブユニット及び配列番号3に示される配列を含むF2サブユニットに切断される。

#### 【0283】

いくつかの実施形態では、Fタンパク質は、配列番号1に示される配列をコードするヌクレオチドの配列によってコードされるNiV-Fタンパク質であるか、または配列番号1と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有するその機能的に活性なバリエーションもしくは生物学的に活性な部分である。いくつかの実施形態では、NiV-Fタンパク質は、配列番号2に示される配列を有し、または配列番号2と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有するその機能的に活性なバリエーションもしくは生物学的に活性な部分である。特定の実施形態では、Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、カテプシンLによって切断される切断部位(例えば、配列番号1のアミノ酸109~110の間の切断部位に対応する)を保持する。

#### 【0284】

いくつかの実施形態では、Fタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分には、配列番号4に示される配列、または配列番号4と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有するF1サブユニットが含まれる。

#### 【0285】

いくつかの実施形態では、Fタンパク質またはその機能的に活性なバリエーションもしくは

10

20

30

40

50

生物学的に活性な部分には、配列番号 3 に示される配列、または配列番号 3 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、少なくとも 84% もしくは約 84%、少なくとも 85% もしくは約 85%、86% もしくは約 86%、少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する F2 サブユニットが含まれる。

10

**【0286】**

いくつかの実施形態では、Fタンパク質は、切断型でありかつ野生型 NiV-Fタンパク質（例えば、配列番号 2）の C 末端におけるまたはその近くの最大で 20 個の連続アミノ酸残基を欠くその生物学的に活性な部分である変異体 NiV-Fタンパク質である。いくつかの実施形態では、変異体 NiV-Fタンパク質は、配列番号 5 に示されるアミノ酸配列を含む。いくつかの実施形態では、変異体 NiV-Fタンパク質は、配列番号 5 と少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有する配列を有する。いくつかの実施形態では、変異体 Fタンパク質は、配列番号 6 に示される配列を有する F1タンパク質を含有する。いくつかの実施形態では、変異体 Fタンパク質は、配列番号 6 と少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有する配列を有する。

20

**【0287】**

いくつかの実施形態では、Fタンパク質は、野生型 NiV-Fタンパク質（配列番号 2）の C 末端におけるまたはその近くの 20 アミノ酸切断；及び N 結合型グリコシル化部位における点変異を有するその生物学的に活性な部分である変異体 NiV-Fタンパク質である。いくつかの実施形態では、変異体 NiV-Fタンパク質は、配列番号 7 に示されるアミノ酸配列を含む。いくつかの実施形態では、変異体 NiV-Fタンパク質は、配列番号 7 と少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有する配列を有する。

30

40

**【0288】**

いくつかの実施形態では、Fタンパク質は、野生型 NiV-Fタンパク質（配列番号 2）の C 末端におけるまたはその近くの 22 アミノ酸切断を有するその生物学的に活性な部分である変異体 NiV-Fタンパク質である。いくつかの実施形態では、NiV-Fタンパク質は、配列番号 8 に示される配列をコードするヌクレオチド配列によってコードされる。いくつかの実施形態では、NiV-Fタンパク質は、配列番号 8 と少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の

50

配列同一性を有する配列をコードするヌクレオチド配列によってコードされる。特定の実施形態では、バリエーションタンパク質は、配列番号 23 に示されるアミノ酸の配列を有する変異体 Ni v - F タンパク質である。いくつかの実施形態では、Ni V - F タンパク質は、配列番号 23 と少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有する配列によってコードされる。

#### 【0289】

##### C. 脂質二重層

いくつかの実施形態では、標的化脂質粒子には、内腔またはキャビティを包囲する両親媒性脂質の天然に由来する二重層が含まれる。いくつかの実施形態では、標的化脂質粒子は、最外表面として脂質二重層を含む。いくつかの実施形態では、脂質二重層は、内腔を包囲する。いくつかの実施形態では、内腔は、水性である。いくつかの実施形態では、内腔は、脂質二重層の内部の親水性頭部基と接触している。いくつかの実施形態では、内腔は、サイトゾルである。いくつかの実施形態では、サイトゾルは、源細胞に存在する細胞成分を含有する。いくつかの実施形態では、サイトゾルは、源細胞に存在する成分を含有しない。いくつかの実施形態では、内腔は、キャビティである。いくつかの実施形態では、キャビティは、水性環境を含有する。いくつかの実施形態では、キャビティは、水性環境を含有しない。

#### 【0290】

いくつかの態様では、脂質二重層は、脂質含有粒子を生成するためのプロセス中に源細胞から誘導される。脂質含有粒子を生成するための例示的な方法は、セクション I . E で提供される。いくつかの実施形態では、脂質二重層には、脂質二重層が生成される細胞の膜成分、例えば、リン脂質、膜タンパク質などが含まれる。いくつかの実施形態では、脂質二重層には、微小胞が生成される細胞に見られる成分を含むサイトゾル、例えば、溶質、タンパク質、核酸などが含まれるが、細胞の成分のすべてではなく、例えば、それらは核を欠く。いくつかの実施形態では、脂質二重層は、エクソソーム様とみなされる。脂質二重層は、サイズが様々であり得、いくつかの例では、40 ~ 100 nm を含む、30 から 150 nm などの 30 から 300 nm の範囲の直径を有する。

#### 【0291】

いくつかの実施形態では、脂質二重層は、ウイルスエンベロープである。いくつかの実施形態では、ウイルスエンベロープは、源細胞から得られる。いくつかの実施形態では、ウイルスエンベロープは、源細胞の細胞膜からのウイルスカプシドによって得られる。いくつかの実施形態では、脂質二重層は、宿主細胞の細胞膜以外の膜から得られる。いくつかの実施形態では、ウイルスエンベロープ脂質二重層は、ウイルス糖タンパク質を含むウイルスタンパク質と共に埋め込まれている。

#### 【0292】

他の態様では、脂質二重層には、合成脂質複合体が含まれる。いくつかの実施形態では、合成脂質複合体は、リポソームである。いくつかの実施形態では、脂質二重層は、リン脂質二重層膜及び内側水性媒体によって特性化される小胞構造である。いくつかの実施形態では、脂質二重層は、水性媒体によって分離された複数の脂質層を有する。いくつかの実施形態では、脂質二重層は、リン脂質が過剰の水溶液に懸濁された場合に自発的に形成する。いくつかの例では、脂質成分は、近接した構造の形成前に自己再構成を受け、脂質二重層間の水及び溶解した溶質を捕捉する。

#### 【0293】

いくつかの実施形態では、源細胞に対して外因性のまたは過剰発現したものを含む、上述したものなどの標的化エンベロープタンパク質及びフソゲンは、脂質二重層に配置される。

#### 【0294】

10

20

30

40

50



いくつかの実施形態では、標的化脂質粒子は、いくつかの異なるタイプの脂質を含む。いくつかの実施形態では、脂質は、両親媒性脂質である。いくつかの実施形態では、両親媒性脂質は、リン脂質である。いくつかの実施形態では、リン脂質は、ホスファチジルコリン、ホスファチジルエタノールアミン、ホスファチジリンイノシトール、及びホスファチジルセリンを含む。いくつかの実施形態では、脂質は、ホスホコリン及びホスホイノシトールなどのリン脂質を含む。いくつかの実施形態では、脂質は、DMPC、DOPC、及びDSPCを含む。

【0295】

いくつかの実施形態では、二重層は、同じまたは異なるタイプの1つ以上の脂質で構成され。いくつかの実施形態では、源細胞は、CHO細胞、BHK細胞、MDCK細胞、C3H 10T1/2細胞、FLY細胞、Psi-2細胞、BOSC 23細胞、PA317細胞、WEHI細胞、COS細胞、BSC 1細胞、BSC 40細胞、BMT 10細胞、VERO細胞、W138細胞、MRC5細胞、A549細胞、HT1080細胞、293細胞、293T細胞、B-50細胞、3T3細胞、NIH3T3細胞、HepG2細胞、Saos-2細胞、Huh7細胞、HeLa細胞、W163細胞、211細胞、及び211A細胞から選択される細胞を含む。

10

【0296】

D. 外因性物質

実施形態では、レンチウイルスベクターなどの標的化脂質粒子は、源細胞に対して外因性である物質（以下「カーゴ」または「ペイロード」とも呼ばれる）をさらに含む。いくつかの実施形態では、外因性物質は、タンパク質または核酸（例えば、DNA、染色体（例えば、ヒト人工染色体）、RNA、例えば、mRNAまたはmiRNA）である。いくつかの実施形態では、外因性物質は、タンパク質をコードする核酸である。タンパク質は、標的細胞に対する標的化送達のために所望であるような任意のタンパク質であり得る。いくつかの実施形態では、タンパク質は、治療剤または診断剤である。いくつかの実施形態では、タンパク質は、疾患または病態によって発現するまたはそれに関連する細胞を標的化するための抗原受容体、例えばキメラ抗原受容体（CAR）またはT細胞受容体（TCR）である。タンパク質をコードする核酸のコード配列に対する言及はまた、本明細書でペイロード遺伝子と称される。いくつかの実施形態では、外因性物質または外因性物質をコードする核酸は、非細胞粒子の内腔に存在する。

20

30

【0297】

いくつかの実施形態では、外因性物質またはカーゴは、サイトゾルタンパク質を含みまたはコードする。いくつかの実施形態では、外因性物質またはカーゴは、膜タンパク質を含みまたはコードする。いくつかの実施形態では、外因性物質またはカーゴは、治療剤を含みまたはコードする。いくつかの実施形態では、治療剤は、タンパク質、例えば、酵素、膜貫通タンパク質、受容体、抗体；核酸、例えば、DNA、染色体（例えば、ヒト人工染色体）、RNA、mRNA、siRNA、miRNA、または小分子のうちの1つ以上から選択される。

【0298】

実施形態では、外因性物質は、少なくとも10、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、50,000、100,000、200,000、500,000、1,000,000、5,000,000、10,000,000、20,000,000、50,000,000、100,000,000、200,000,000、500,000,000、1,000,000,000、または1,000,000,000コピー、または10、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、50,000、100,000、200,000、500,000、1,000,000、5,000,000、10,000,000、20,000,000、50,000,000、100,000,000、200,000,000、500,000,000、1,000,000,000、または1,000,000,000コピー以下で存在する。実施形態では、標的化脂質粒子は、例えば、siRNAまたは遺伝子編集酵素での源細胞、例えば、哺乳動物源細胞の処置により、1つ以上の内因性分子、

40

50

例えば、タンパク質または核酸（例えば、いくつかの実施形態では、源細胞に対して内因性、いくつかの実施形態では、標的細胞に対して内因性）の改変、例えば、増加または低下したレベルを有する。実施形態では、内因性分子は、少なくとも10、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、50,000、100,000、200,000、500,000、1,000,000、2,000,000、5,000,000、10,000,000、20,000,000、50,000,000、100,000,000、または1,000,000,000コピー、または10、20、50、100、200、500、1,000、2,000、5,000、10,000、20,000、50,000、100,000、200,000、500,000、1,000,000、2,000,000、5,000,000、10,000,000、20,000,000、50,000,000、100,000,000、または1,000,000,000コピー以下で存在する。実施形態では、内因性分子（例えば、RNAまたはタンパク質）は、源細胞におけるその濃度よりも少なくとも1、2、3、4、5、10、20、50、100、500、 $10^3$ 、 $5.0 \times 10^3$ 、 $10^4$ 、 $5.0 \times 10^4$ 、 $10^5$ 、 $5.0 \times 10^5$ 、 $10^6$ 、 $5.0 \times 10^6$ 、 $1.0 \times 10^7$ 、 $5.0 \times 10^7$ 、または $1.0 \times 10^8$ 高い濃度で存在する。実施形態では、内因性分子（例えば、RNAまたはタンパク質）は、源細胞におけるその濃度よりも少なくとも1、2、3、4、5、10、20、50、100、500、 $10^3$ 、 $5.0 \times 10^3$ 、 $10^4$ 、 $5.0 \times 10^4$ 、 $10^5$ 、 $5.0 \times 10^5$ 、 $10^6$ 、 $5.0 \times 10^6$ 、 $1.0 \times 10^7$ 、 $5.0 \times 10^7$ 、または $1.0 \times 10^8$ 低い濃度で存在する。

#### 【0299】

いくつかの実施形態では、標的化脂質粒子は、フソソームに含まれるカーゴ（例えば、治療剤、例えば、外因性治療剤）の少なくとも10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、96%、97%、98%、または99%を標的細胞に送達する。いくつかの実施形態では、標的細胞（複数可）と融合する標的化脂質粒子は、標的細胞（複数可）と融合する脂質粒子に含まれるカーゴ（例えば、治療剤、例えば、外因性治療剤）の平均で少なくとも10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、96%、97%、98%、または99%を標的細胞に送達する。いくつかの実施形態では、標的化脂質粒子組成物は、標的化脂質粒子組成物に含まれるカーゴ（例えば、治療剤、例えば、外因性治療剤）の少なくとも10%、20%、30%、40%、50%、60%、70%、80%、90%、95%、96%、97%、98%、または99%を標的組織に送達する。

#### 【0300】

いくつかの実施形態では、外因性物質またはカーゴは、標的化脂質粒子の由来となった細胞において自然に発現しない。いくつかの実施形態では、外因性物質またはカーゴは、標的化脂質粒子の由来となった細胞において自然に発現する。いくつかの実施形態では、外因性物質またはカーゴは、脂質粒子の由来となった細胞における発現（例えば、トランスフェクション、形質導入、またはエレクトロポレーションを介して導入されたDNAまたはmRNAからの発現）を介して標的化脂質粒子に搭載される。いくつかの実施形態では、外因性物質またはカーゴは、ゲノムに組み込まれたまたはエピソーム的に維持されたDNAから発現する。いくつかの実施形態では、外因性物質またはカーゴの発現は、構成的である。いくつかの実施形態では、外因性物質またはカーゴの発現は、誘導される。いくつかの実施形態では、外因性物質またはカーゴの発現は、標的化脂質粒子を生成する直前に誘導される。いくつかの実施形態では、外因性物質またはカーゴの発現は、フソゲンの発現と同時に誘導される。

#### 【0301】

いくつかの実施形態では、外因性物質またはカーゴは、脂質粒子自体またはフソソームの由来となった細胞へのエレクトロポレーションを介して脂質粒子に搭載される。いくつかの実施形態では、外因性物質またはカーゴは、脂質粒子自体または脂質粒子の由来となった細胞への（例えば、カーゴをコードするDNAまたはmRNAの）トランスフェクシ

ヨンを介して脂質粒子に搭載される。

【0302】

いくつかの実施形態では、外因性物質またはカーゴには、1つ以上の核酸配列、1つ以上のポリペプチド、核酸配列及び/またはポリペプチドの組み合わせ、1つ以上の細胞小器官、及びそれらの任意の組み合わせが含まれ得る。いくつかの実施形態では、外因性物質またはカーゴには、1つ以上の細胞成分が含まれ得る。いくつかの実施形態では、外因性物質またはカーゴには、1つ以上のサイトゾル及び/または核成分が含まれる。

【0303】

いくつかの実施形態では、外因性物質またはカーゴには、核酸、例えば、DNA、nDNA (核DNA)、mtDNA (ミトコンドリアDNA)、タンパク質コードDNA、遺伝子、オペロン、染色体、ゲノム、トランスポゾン、レトロトランスポゾン、ウイルスゲノム、イントロン、エクソン、修飾DNA、mRNA (メッセンジャーRNA)、tRNA (トランスファーRNA)、修飾RNA、マイクロRNA、siRNA (低分子干渉)、tmRNA (トランスフェームッセンジャーRNA)、rRNA (リボソームRNA)、mtRNA (ミトコンドリアRNA)、snRNA (小核RNA)、小核小体RNA (snORNA)、SmY RNA (mRNAトランス-スプライシングRNA)、gRNA (ガイドRNA)、TERC (テロメラーゼRNA成分)、aRNA (アンチセンスRNA)、cis-NAT (シス-天然アンチセンス転写産物)、CRISPR RNA (crRNA)、IncRNA (長鎖ノンコーディングRNA)、piRNA (piwi相互作用RNA)、shRNA (短ヘアピンRNA)、tasirRNA (トランス作用性siRNA)、eRNA (エンハンサーRNA)、サテライトRNA、pcRNA (タンパク質コードRNA)、dsRNA (二本鎖RNA)、RNAi (干渉RNA)、circRNA (環状RNA)、リプログラミングRNA、アプタマー、及びそれらの任意の組み合わせが含まれる。いくつかの実施形態では、核酸は、野生型核酸である。いくつかの実施形態では、タンパク質は、変異体核酸である。いくつかの実施形態では、核酸は、複数の核酸配列の融合体またはキメラである。

【0304】

いくつかの実施形態では、外因性物質またはカーゴには、核酸が含まれ得る。例えば、外因性物質またはカーゴは、内因性タンパク質の発現を向上させるためのRNA、または内因性タンパク質のタンパク質発現を阻害するsiRNAまたはmiRNAを含み得る。例えば、内因性タンパク質は、標的細胞において構造または機能を調節し得る。いくつかの実施形態では、カーゴには、標的細胞において構造または機能を調節する操作されたタンパク質をコードする核酸が含まれ得る。いくつかの実施形態では、外因性物質またはカーゴは、標的細胞において構造または機能を調節する転写活性化因子を標的とする核酸である。

【0305】

いくつかの実施形態では、外因性物質またはカーゴは、ポリペプチド、例えば、酵素、構造ポリペプチド、シグナル伝達ポリペプチド、制御性ポリペプチド、輸送ポリペプチド、感覚ポリペプチド、運動ポリペプチド、防御ポリペプチド、保存ポリペプチド、転写因子、抗体、サイトカイン、ホルモン、異化ポリペプチド、同化ポリペプチド、タンパク質分解ポリペプチド、代謝ポリペプチド、キナーゼ、トランスフェラーゼ、ヒドロラーゼ、リアーゼ、イソメラーゼ、リガーゼ、酵素調節因子ポリペプチド、タンパク質結合ポリペプチド、脂質結合ポリペプチド、膜融合ポリペプチド、細胞分化ポリペプチド、エピジェネティックポリペプチド、細胞死ポリペプチド、核輸送ポリペプチド、核酸融合ポリペプチド、再プログラミングポリペプチド、DNA編集ポリペプチド、DNA修復ポリペプチド、DNA組換えポリペプチド、トランスポサーゼポリペプチド、DNA組み込みポリペプチド、標的化エンドヌクレアーゼ (例えば、ジンクフィンガーヌクレアーゼ、転写活性化因子様ヌクレアーゼ (TALEN)、cas9及びそのホモログ)、リコンビナーゼ、及びそれらの任意の組み合わせであるか、またはそれをコードする。いくつかの実施形態では、タンパク質は、分解のための細胞におけるタンパク質を標的とする。いくつかの実施

10

20

30

40

50

形態では、タンパク質は、タンパク質をプロテアソームに局在化することによって分解のための細胞におけるタンパク質を標的とする。いくつかの実施形態では、タンパク質は、野生型タンパク質である。いくつかの実施形態では、タンパク質は、変異体タンパク質である。いくつかの実施形態ではタンパク質は、融合またはキメラタンパク質である。

【0306】

いくつかの実施形態では、外因性物質またはカーゴは、小分子、例えば、イオン（例えば、 $Ca^{2+}$ 、 $Cl^-$ 、 $Fe^{2+}$ ）、炭水化物、脂質、反応性酸素種、反応性窒素種、イソプレノイド、シグナル伝達分子、ヘム、ポリペプチド補因子、電子求引性化合物、電子供与性化合物、代謝産物、リガンド、及びそれらの任意の組み合わせである。いくつかの実施形態では、小分子は、細胞において標的と相互作用する薬剤である。いくつかの実施形態では、小分子は、分解のための細胞におけるタンパク質を標的とする。いくつかの実施形態では、小分子は、タンパク質をプロテアソームに局在化することによって分解のための細胞におけるタンパク質を標的とする。いくつかの実施形態では、その小分子は、タンパク質分解標的化キメラ分子（PROTAC）である。

10

【0307】

いくつかの実施形態では、外因性物質またはカーゴには、タンパク質、核酸、または代謝産物の混合物、例えば、複数のポリペプチド、複数の核酸、複数の小分子；核酸、ポリペプチド、及び小分子の組み合わせ；リボヌクレオタンパク質複合体（例えば、Cas9-gRNA複合体）；複数の転写因子、複数のエピジェネティック因子、再プログラム化因子（例えば、Oct4、Sox2、cMyc、及びKlf4）；複数の制御性RNA；及びそれらの任意の組み合わせが含まれる。

20

【0308】

いくつかの実施形態では、外因性物質またはカーゴには、1つ以上の細胞小器官、例えば、コンドリソーム、ミトコンドリア、リソソーム、核、細胞膜、細胞質、小胞体、リボソーム、液胞、エンドソーム、スプライセオソーム、ポリメラーゼ、カプシド、アクロソーム、オートファゴソーム、中心小体、グリコソーム、グリオキシソーム、ハイドロジェノソーム、メラノソーム、マイトソーム、マイオファイブリル、刺胞（cnidocyst）、ペルオキシソーム、プロテアソーム、小胞、ストレス顆粒、細胞小器官のネットワーク、及びそれらの任意の組み合わせが含まれる。

【0309】

いくつかの実施形態では、外因性物質は、サイトゾルタンパク質、例えば、レシピエント細胞において生成され、レシピエント細胞質に局在化するタンパク質であるか、またはそれをコードする。いくつかの実施形態では、外因性物質は、分泌タンパク質、例えば、レシピエント細胞によって生成され、分泌されるタンパク質であるか、またはそれをコードする。いくつかの実施形態では、外因性物質は、核タンパク質、例えば、レシピエント細胞において生成され、レシピエント細胞の核に搬入されるタンパク質であるか、またはそれをコードする。いくつかの実施形態では、外因性物質は、細胞小器官タンパク質（例えば、ミトコンドリアタンパク質）、例えば、レシピエント細胞において生成され、レシピエント細胞の細胞小器官（例えば、ミトコンドリア）に搬入されるタンパク質であるか、またはそれをコードする。いくつかの実施形態では、タンパク質は、野生型タンパク質または変異体タンパク質である。いくつかの実施形態ではタンパク質は、融合またはキメラタンパク質である。

30

40

【0310】

いくつかの実施形態では、外因性物質は、肝細胞または肝臓細胞に送達されることが可能である。いくつかの実施形態では、外因性物質またはカーゴは、肝細胞または肝臓細胞における疾患または障害を処置するために送達され得る。

【0311】

いくつかの実施形態では、外因性物質は、OTC、CPS1、NAGS、BCKDHA、BCKDHB、DBT、DLLD、MUT、MMAA、MMAB、MMACHC、MMAADHC、MCEE、PCCA、PCCB、UGT1A1、ASS1、PAH、PAL、A

50

TP8B1、 ABCB11、 ABCB4、 TJP2、 IVD、 GCDH、 ETF A、 ET  
 FB、 ETFDH、 ASL、 D2HGDH、 HMGCL、 MCCC1、 MCCC2、 AB  
 CD4、 HCFC1、 LNBRD1、 ARG1、 SLC25A15、 SLC25A13、  
 ALAD、 CPOX、 HMBS、 PPOX、 BT D、 HLC S、 PC、 SLC7A7、 C  
 PT2、 ACADM、 ACADS、 ACADV L、 AGL、 G6PC、 GBE1、 PHK  
 A1、 PHKA2、 PHKB、 PHKG2、 SLC37A4、 PMM2、 CBS、 FAH  
 、 TAT、 GALT、 GALK1、 GALE、 G6PD、 SLC3A1、 SLC7A9、  
 MTHFR、 MTR、 MTRR、 ATP7B、 HPRT1、 HJV、 HAMP、 JAG1  
 、 TTR、 AGXT、 LIPA、 SERPING1、 HSD17B4、 UROD、 HFE  
 、 LPL、 GRHPR、 HOGA1、 LDLR、 ACAD8、 ACADSB、 ACAT1 10  
 、 ACSF3、 ASPA、 AUH、 DNAJC19、 ETHE1、 FBP1、 FTCD、  
 GSS、 HIBCH、 IDH2、 L2HGDH、 MLYCD、 OPA3、 OPLAH、 O  
 XCT1、 POLG、 PPM1K、 SERAC1、 SLC25A1、 SUCLA2、 SU  
 CLG1、 TAZ、 AGK、 CLPB、 TME M70、 ALDH18A1、 OAT、 CA  
 5A、 GLUD1、 GLUL、 UMPS、 SLC22A5、 CPT1A、 HADHA、 H  
 ADH、 SLC52A1、 SLC52A2、 SLC52A3、 HADHB、 GYS2、 P  
 YGL、 SLC2A2、 ALG1、 ALG2、 ALG3、 ALG6、 ALG8、 ALG9  
 、 ALG11、 ALG12、 ALG13、 ATP6V0A2、 B3GLCT、 CHST1  
 4、 COG1、 COG2、 COG4、 COG5、 COG6、 COG7、 COG8、 DOL  
 K、 DHDDS、 DPAGT1、 DPM1、 DPM2、 DPM3、 G6PC3、 GFPT 20  
 1、 GMPPA、 GMPPB、 MAGT1、 MAN1B1、 MGAT2、 MOGS、 MP  
 DU1、 MPI、 NGLY1、 PGM1、 PGM3、 RFT1、 SEC23B、 SLC3  
 5A1、 SLC35A2、 SLC35C1、 SSR4、 SRD5A3、 TME M165、  
 TRIP11、 TUSC3、 ALG14、 B4GALT1、 DDOST、 NUS1、 RP  
 N2、 SEC23A、 SLC35A3、 ST3GAL3、 STT3A、 STT3B、 AG  
 A、 ARSA、 ARSB、 ASAH1、 ATP13A2、 CLN3、 CLN5、 CLN6  
 、 CLN8、 CTNS、 CTSA、 CTSD、 CTSF、 CTSK、 DNAJC5、 FU  
 CA1、 GAA、 GALC、 GALNS、 GLA、 GLB1、 GM2A、 GNPTAB、  
 GNPTG、 GNS、 GRN、 GUSB、 HEXA、 HEXB、 HGSNAT、 HYAL 30  
 1、 IDS、 IDUA、 KCTD7、 LAMP2、 MAN2B1、 MANBA、 MCOL  
 N1、 MFSD8、 NAGA、 NAGLU、 NEU1、 NPC1、 NPC2、 SGSH、  
 PPT1、 PSAP、 SLC17A5、 SMPD1、 SUMF1、 TPP1、 AHCY、  
 GNMT、 MAT1A、 GCH1、 PCBD1、 PTS、 QDPR、 SPR、 DNAJC  
 12、 ALDH4A1、 PRODH、 HPD、 GBA、 HGD、 AMN、 CD320、 C  
 UBN、 GIF、 TCN1、 TCN2、 PREPL、 PHGDH、 PSAT1、 PSPH  
 、 AMT、 GCSH、 GLDC、 LIAS、 NFU1、 SLC6A9、 SLC2A1、 A  
 TP7A、 AP1S1、 CP、 SLC33A1、 PEX7、  
 PHYH、 AGPS、 GNPAT、 ABCD1、 ACOX1、 PEX1、 PEX2、 PE  
 X3、 PEX5、 PEX6、 PEX10、 PEX12、 PEX13、 PEX14、 PEX  
 16、 PEX19、 PEX26、 AMACR、 ADA、 ADSL、 AMPD1、 GPHN 40  
 、 MOCOS、 MOCS1、 PNP、 XDH、 SUOX、 OGDH、 SLC25A19、  
 DHTKD1、 SLC13A5、 FH、 DLAT、 MPC1、 PDHA1、 PDHB、 P  
 DHX、 PDP1、 ABCC2、 SLCO1B1、 SLCO1B3、 HFE2、 ADAM  
 TS13、 PYGM、 COL1A2、 TNFRSF11B、 TSC1、 TSC2、 DHC  
 R7、 PGK1、 VLDLR、 KYNU、 F5、 C3、 COL4A1、 CFH、 SLC1  
 2A2、 GK、 SFTPC、 CRTAP、 P3H1、 COL7A1、 PKLR、 TALD  
 O1、 TF、 EPCAM、 VHL、 GC、 SERPINA1、 ABCC6、 F8、 F9、  
 ApoB、 PCSK9、 LDLRAP1、 ABCG5、 ABCG8、 LCAT、 SPIN  
 K5、 またはGNEからの遺伝子によってコードされる。

【0312】

いくつかの実施形態では、外因性物質は、OTC、CPS1、NAGS、BCKDHA、BCKDHB、DBT、DL D、MUT、MMAA、MMAB、MMACHC、MMA DHC、MCEE、PCCA、PCCB、UGT1A1、ASS1、PAL、PAH、ATP8B1、ABCB11、ABCB4、TJP2、IVD、GCDH、ETFA、ETFB、ETFDH、ASL、D2HGDH、HMGCL、MCCC1、MCCC2、ABCD4、HCFC1、LMBRD1、ARG1、SLC25A15、SLC25A13、ALAD、CPOX、HMBS、PPOX、BTD、HLC S、PC、SLC7A7、CPT2、ACADM、ACADS、ACADVL、AGL、G6PC、GBE1、PHKA1、PHKA2、PHKB、PHKG2、SLC37A4、PMM2、CBS、FAH、TAT、GALT、GALK1、GALE、G6PD、SLC3A1、SLC7A9、MTHFR、MTR、MTRR、ATP7B、HPRT1、HJV、HAMP、JAG1、TTR、AGXT、LIPA、SERPING1、HSD17B4、UROD、HFE、LPL、GRHPR、HOGA1、またはLDLRからの遺伝子によってコードされる。いくつかの実施形態では、外因性物質は、酵素フェニルアラニンアンモニアリアーゼ(PAL)である。

10

## 【0313】

いくつかの実施形態では、外因性物質またはカーゴは、表5に列挙される疾患または適応症を処置するために送達され得る。いくつかの実施形態では、適応症は、肝臓細胞または肝細胞に特定のである。

## 【0314】

20

いくつかの実施形態では、外因性物質は、以下の表5のタンパク質を含む。いくつかの実施形態では、外因性物質は、表5のタンパク質のいずれかの野生型ヒト配列、その機能的断片(例えば、その酵素的に活性な断片)、またはその機能的バリエーションを含む。いくつかの実施形態では、外因性物質は、表5のアミノ酸配列、例えば、表5の列4のUniProtタンパク質アクセッション番号配列または表5の列5のアミノ酸配列と少なくとも70%、75%、80%、85%、90%、95%、96%、97%、98%、または99%の同一性を有するアミノ酸配列を含む。いくつかの実施形態では、外因性物質をコードするペイロード遺伝子は、表5のアミノ酸配列と少なくとも70%、75%、80%、85%、90%、95%、96%、97%、98%、または99%の同一性を有するアミノ酸配列をコードする。いくつかの実施形態では、外因性物質をコードするペイロード

30

## 【0315】

(表5)

第1の列は、本明細書における方法及び使用に従って、第6の列における適応症を処置するために送達され得る外因性物質を列挙している。表5の各UniProtアクセッション番号は、その全体が参照により本明細書に組み込まれる。

40

50

遺伝子	Entrez ア クセッ ション 番号	Ensembl 遺 伝子(複数 可)アクセ ッション番 号 (ENSG0000 +示される 番号)	Uniprot タ ンパク質 (複数可)ア クセッシ ョン番号	アミノ酸 配列(第1 の Uniprot アクセッ ション番 号) 配列番 号	疾患/障害	カテゴリー
OTC	5009	0036473	P00480	61	オルニチン トランスカ ルバミラー ゼ(OTC)欠損 症	尿素回路障害
CPS1	1373	0021826	P31327, Q6PEK7, B7ZAW0, A0A024R4 54	62	カルバモイ ルリン酸シ ンテターゼ I(CPSI)欠損 症	尿素回路障害
NAG S	162417	0161653	Q8N159, Q2NKP2	63	N-アセチル グルタミン 酸シタ ーゼ(NAGS) 欠損症	尿素回路障害
BCK DHA	593	0248098	A0A024R0 K3, P12694, Q59EI3	64	メープルシ ロップ尿症 (MSUD); 古典的メー プルシロッ プ尿症 (CMSUD)	有機酸血症

10

20

30

40

50

BCK DHB	594	0083123	A0A140V KB3, P21953, B4E2N3, B7ZB80	65	メープルシ ロップ尿症 (MSUD); 古典的メー プルシロッ プ尿症 (CMSUD)	有機酸血症
DBT	1629	0137992	P11182	66	メープルシ ロップ尿症 (MSUD); 古典的メー プルシロッ プ尿症 (CMSUD)	有機酸血症
DLD	1738	0091140	A0A024R7 13, P09622, E9PEX6	67	メープルシ ロップ尿症 (MSUD) ジヒドロリ ポアミドデ ヒドロゲナ ーゼ欠損症	尿素回路障害
MUT	4594	0146085	A0A024R D82, B2R6K1, P22033	68	メチルマロ ニル-CoA ム ターゼ欠損 症によるメ チルマロン 酸血症	有機酸血症
MMA A	166785	0151611	Q8IVH4	69	コバラミン A 欠損症(メ チルマロン 酸血症)	有機酸血症
MMA B	326625	0139428	Q96EY8	70	コバラミン B 欠損症(メ チルマロン 酸血症)	有機酸血症

10

20

30

40

50



MMA CHC	25974	0132763	A0A0C4D GU2, Q9Y4U1	71	コバラミン C欠損症(メ チルマロン 酸血症); ホモシスチ ン尿症を伴 うメチルマ ロン酸血症	有機酸血症	10
MMA DHC	27249	0168288	Q9H3L0	72	コバラミン D欠損症(メ チルマロン 酸血症); ホモシスチ ン尿症を伴 うメチルマ ロン酸血症; ホモシスチ ン尿症; コバラミン C欠損症	有機酸血症	20
MCE E	84693	0124370	Q96PE7	73	メチルマロ ン酸血症; コバラミン D欠損症	有機酸血症	30
PCCA	5095	0175198	P05165	74	プロピオン 酸血症	有機酸血症	30
PCCB	5096	0114054	P05166	75	プロピオン 酸血症	有機酸血症	40

UGT1 A1	54658	0241635	P22309, Q5DT03	76	クリグラ ー・ナジャ ール症候群 1 型 クリグラ ー・ナジャ ール症候群 2 型, ジルペー ル症候群	
ASS1	445	0130707	P00966, Q5T6L4	77	シトルリン 血症 I 型	尿素回路障害
PAH	5053	0171759	A0A024R BG4, P00439	78	フェニルア ラニンヒド ロキシラー ゼ欠損症	アミノ酸代謝異 常
PAL				79	フェニルア ラニンヒド ロキシラー ゼ欠損症	アミノ酸代謝異 常
ATP8 B1	5205	0081923	O43520	80	進行性家族 性肝内胆汁 鬱滞 1 型	
ABC B11	8647	0073734,02 76582	O95342	81	進行性家族 性肝内胆汁 鬱滞 2 型 ; 進行性家族 性肝内胆汁 鬱滞 3 型	
ABC B4	5244	0005471	P21439	82	進行性家族 性肝内胆汁 鬱滞 3 型 ; 進行性家族 性肝内胆汁 鬱滞 2 型	

10

20

30

40

50

TJP2	9414	0119139	B7Z2R3, Q9UDY2, B7Z954	83	進行性家族 性肝内胆汁 鬱滞 4 型	
IVD	3712	0128928	P26440, A0A0A0M T83	84	イソ吉草酸 血症(IVD)	有機酸血症
GCD H	2639	0105607	A0A024R7 F9, Q92947	85	グルタル酸 血症 I 型	有機酸血症
ETFA	2108	0140374	A0A0S2Z3 L0, P13804	86	多重アシル- CoA デヒド ロゲナーゼ 欠損症 (グル タル酸尿症 II 型としても 知られてい る)	有機酸血症
ETFB	2109	0105379	P38117	87	多重アシル- CoA デヒド ロゲナーゼ 欠損症 (グル タル酸尿症 II 型としても 知られてい る)	有機酸血症
ETFD H	2110	0171503	B4DEQ0, Q16134	88	多重アシル- CoA デヒド ロゲナーゼ 欠損症 (グル タル酸尿症 II 型としても 知られてい る)	有機酸血症

10

20

30

40

50

ASL	435	0126522	A0A024R DL8, P04424, A0A0S2Z3 16	89	アルギニノ コハク酸リ アーゼ(ASL) 欠損症	尿素回路障害
D2H GDH	728294	0180902	B3KSR6, B4E3K7, B5MCV2, Q8N465	90	D-2-ヒドロキ シグルタル 酸尿症 I 型	有機酸血症
HMG CL	3155	0117305	P35914	91	3-ヒドロキシ -3-メチルグ ルタリル- CoA リアー ゼ(3HMG)欠 損症	有機酸血症 尿素回路障害
MCC C1	56922	0078070	Q68D27, Q96RQ3, A0A0S2Z6 93, E9PHF7	92	3-メチルクロ トニル-CoA カルボキシ ラーゼ (3MCC)欠損 症	有機酸血症
MCC C2	64087	0131844,02 81742,0275 300	A0A140V K29, Q9HCC0	93	3-メチルクロ トニル-CoA カルボキシ ラーゼ (3MCC)欠損 症	有機酸血症
ABC D4	5826	0119688	A0A024R6 B9, O14678, A0A024R6 C8	94	ホモシスチ ン尿症を伴 うメチルマ ロン酸血症	有機酸血症

10

20

30

40

50

HCFC 1	3054	0172534	P51610, A6NEM2	95	ホモシスチ ン尿症を伴 うメチルマ ロン酸血症	有機酸血症
LMB RD1	55788	0168216	Q9NUN5	96	ホモシスチ ン尿症を伴 うメチルマ ロン酸血症	有機酸血症
ARG1	383	0118520	P05089	97	アルギナー ゼ(ARG1)欠 損症	尿素回路障害
SLC2 5A15	10166	0102743	Q9Y619	98	高アンモニ ア血症・高 オルニチン 血症・ホモ シトルリン 尿症(HHH)症 候群	尿素回路障害
SLC2 5A13	10165	0004864	Q9UJS0	99	シトルリン欠 損症 シトルリン 血症 II 型	尿素回路障害
ALA D	210	0148218	P13716	100	急性肝臓ポ ルフィリン 症	ポルフィリン症
CPO X	1371	0080819	P36551	101	急性肝臓ポ ルフィリン 症	ポルフィリン症

10

20

30

40

50

HMB S	3145	0256269,02 81702	P08397	102	急性肝臓ポ ルフィリン 症; 急性間欠性 ポルフィリ ン症	ポルフィリン症	
PPOX	5498	0143224	P50336, B4DY76	103	急性肝臓ポ ルフィリン 症	ポルフィリン症	10
BTD	686	0169814	P43251	104	ビオチニダ ーゼ欠損症	有機酸血症	
HLCS	3141	0159267	P50747	105	ホロカルボ キシラーゼ シンテター ゼ欠損症	有機酸血症	20
PC	5091	0173599	P11498 A0A024R5 C5	106	ピルビン酸 カルボキシ ラーゼ欠損 症	尿素回路障害	
SLC7 A7	9056	0155465	Q9UM01 A0A0S2Z5 02	107	リジン尿性 タンパク不 耐症	尿素回路障害	
CPT2	1376	0157184	P23786 A0A140V K13 A0A1B0G TB8	108	カルニチン パルミトイ ルトランス フェラーゼ II型(CPT II) 欠損症	脂肪酸酸化	30

ACA DM	34	0117054	P11310 A0A0S2Z3 66, B7Z911, Q5HYG7, Q5T4U5, B4DJE7	109	中鎖アシル- CoA デヒド ロゲナーゼ (MCAD)欠損 症	脂肪酸酸化	
ACA DS	35	0122971	P16219 E5KSD5, B4DUH1, E9PE82	110	短鎖アシル- CoA(SCAD) デヒドロゲ ナーゼ欠損 症	脂肪酸酸化	10
ACA DVL	37	0072778	P49748 B3KPA6	111	極長鎖アシ ル-CoA デヒ ドロゲナー ゼ(VLCAD) 欠損症	脂肪酸酸化	20
AGL	178	0162688	P35573 A0A0S2A 4E4	112	GSD III (コリ /フォルベ病 またはデブ ランチャー)	肝臓グリコーゲ ン蓄積障害	
G6PC	2538	0131482	P35575	113	GSDIa (フォ ンギールケ 病)	肝臓グリコーゲ ン蓄積障害	
GBE1	2632	0114480	Q04446 Q59ET0	114	GSD IV (アン ダーソン病, 分枝酵素)	肝臓グリコーゲ ン蓄積障害	30
PHK A1	5255	0067177	P46020	115	GSD IXa		
PHK A2	0044446 5256	5256 0044446	P46019	116	GSD IXa	肝臓グリコーゲ ン蓄積障害	
PHK B	5257	0102893	Q93100	117	GSD IXb	肝臓グリコーゲ ン蓄積障害	40

PHK G2	5261	0156873	P15735	118	GSD IXc	肝臓グリコーゲン蓄積障害
SLC3 7A4	2542	0281500 0137700	O43826 A0A024R3 H9, A8K0S7, A0A024R3 L1, B4DUH2	119	GSDIb. c, d	肝臓グリコーゲン蓄積障害
PMM 2	5373	0140650	O15305, A0A0S2Z4 J6, Q59F02	120	PMM2-CDG	グリコシル化障害
CBS	1027245 60, 875	0160200	P35520, P0DN79, Q9NTF0, B7Z2D6	121	シスタチオニンペーターシンターゼ欠損症 (古典的ホモシスチン尿症); ホモシスチン尿症	アミノ酸代謝異常
FAH	2184	0103876	P16930	122	チロシン血症 I 型	アミノ酸代謝異常
TAT	6898	0198650	P17735, A0A140V KB7	123	チロシン血症 II 型 チロシン血症 III 型	アミノ酸代謝異常

10

20

30

40

50



GAL T	2592	0213930	P07902, A0A0S2Z3 Y7, B2RAT6	124	ガラクトー ス-1-ホスフ ェートウリ ジリルトラ ンスフェラ ーゼ(GALT) 欠損症に起 因するガラ クトース血 症	炭水化物障害	10
GAL K1	2584	0108479	P51570	125	ガラクトー ス血症	炭水化物障害	
GAL E	2582	0117308	Q14376	126	ガラクトー ス血症	炭水化物障害	
G6PD	2539	0160211	P11413	127	グルコース- 6-ホスフェー トデヒドロ ゲナーゼ (G6PD)欠損 症	炭水化物障害	20
SLC3 A1	6519	0138079	Q07837, A0A0S2Z4 E1, B8ZZK1	128	シスチン尿 症	アミノ酸代謝異 常	
SLC7 A9	11136	0021488	P82251	129	シスチン尿 症	アミノ酸代謝異 常	30
MTH FR	4524	0177000	P42898, Q59GJ6, Q81U67	130	ホモシスチ ン尿症	アミノ酸代謝異 常	
MTR	4548	0116984	Q99707	131	ホモシスチ ン尿症	アミノ酸代謝異 常	
MTR R	4552	0124275	Q9UBK8	132	ホモシスチ ン尿症	アミノ酸代謝異 常	40

ATP7B	540	0123191	P35670, A0A024RDX3, B7ZLR4, B7ZLR3, E7ET55	133	ウィルソン病 銅代謝障害	金属輸送障害
HPRT1	3251	0165704	P00492, A0A140VJL3	134	レッシュ・ナイハン症候群 プリン代謝障害	プリン代謝障害
HJV	148738	0168509	Q6ZVN8	135	ヘモクロマトーシス, 2A型	
HAMP	57817	0105697	P81172	136	ヘモクロマトーシス 2B型:原発性ヘモクロマトーシス	
JAG1	182	0101384	P78504, Q99740	137	アラジール症候群 1	
TTR	7276	0118271	P02766, E9KL36	138	家族性 TTR アミロイドーシス;家族性アミロイド多発ニューロパシー	
AGXT	189	0172482	P21549	139	原発性シュウ酸尿症 I 型	
LIPA	3988	0107798	P38571 A0A0A0MT32	140	リソソーム酸リパーゼ欠損症	リソソーム蓄積障害

10

20

30

40

50

SERP ING1	710	0149131	P05155, A0A0S2Z4 J1, B2R659, E7EWE5, B3KSP2, G5E9S2	141	遺伝性血管 浮腫	
HSD1 7B4	3295	0133835	P51659	142	D-二頭酵素 欠損症 X連鎖副腎 白質ジスト ロフィー	10 ペルオキシソー ム障害
URO D	7389	0126088	P06132	143	晩発性皮膚 ポルフィリ ン症	
HFE	3077	0010704	Q30201	144	晩発性皮膚 ポルフィリ ン症	20
LPL	4023	0175445	P06858, A0A1B1R VA9	145	リポタンパ ク質リパー ゼ欠損症 (「高リポタ ンパク血症 Ia型;バー ジャー・グル ーツ症候群, または家族 性高カイロ ミクロン血 症)	30
GRH PR	9380	0137106	Q9UBQ7	146	原発性シュ ウ酸尿症 II 型	40

10

20

30

40

50

HOG A1	112817	0241935	Q86XE5	147	原発性シュ ウ酸尿症 III 型	
LDLR	3949	0130164	P01130, A0A024R7 D5	148	ホモ接合性 家族性高コ レステロー ル血症	
ACA D8	27034	0151498	Q9UKU7	149	イソブチリ ル-CoA デヒ ドロゲナー ゼ(IBD)欠損 症	有機酸血症
ACA DSB	36	0196177	P45954, A0A0S2Z3 P9	150	短分岐鎖ア シル-CoA デ ヒドロゲナ ーゼ (SBCAD)欠 損症	有機酸血症
ACA T1	38	0075239	A0A140VJ X1, P24752	151	ベータ-ケト チオラーゼ 欠損症	有機酸血症
ACSF 3	197322	0176715	Q4G176, F5H5A1	152	混合型マロ ン酸及びメ チルマロン 酸尿症	有機酸血症
ASPA	443	0108381	P45381, Q6FH48	153	カナバン病	有機酸血症
AUH	549	0148090	Q13825, B4DY16	154	3-メチルグル タコン酸血 症 I 型	有機酸血症

10

20

30

40

50

DNAJ C19	131118	0205981	Q96DA6, A0A0S2Z5 X1	155	運動失調症 候群を伴う 拡張型心筋 症 (3-メチル グルタコン 酸尿症を引 き起こす)	有機酸血症
ETHE 1	23474	0105755	A0A0S2Z5 80, O95571, A0A0S2Z5 N8, A0A0S2Z5 B3, B2RCZ7	156	エチルマロ ン酸脳症	有機酸血症
FBP1	2203	0165140	P09467, Q2TU34	157	フルクトー ス 1,6-ビスホ スファター ゼ欠損症	有機酸血症
FTCD	10841	0160282,02 81775	O95954	158	グルタミン 酸ホルムイ ミノトラン スフェラー ゼ欠損症 (FIGLU	有機酸血症
GSS	2937	0100983	P48637, V9HWJ1	159	グルタチオ ンシンテタ ーゼ欠損症	有機酸血症
HIBC H	26275	0198130	A0A140VJ L0, Q6NVY1	160	3-ヒドロキシ イソブチリ ル-CoA ヒド ロラーゼ欠 損症	有機酸血症

10

20

30

40

50

IDH2	3418	0182054	P48735, B4DSZ6	161	D-2-ヒドロキ シグルタル 酸尿症 II 型	有機酸血症
L2HG DH	79944	0087299	Q9H9P8	162	L-2-ヒドロキ シグルタル 酸尿症	有機酸血症
MLY CD	23417	0103150	O95822	163	マロン酸血 症	有機酸血症
OPA3	80207	0125741	Q9H6K4, B4DK77	164	コステフ症 候群/3-メチ ルグルタコ ン酸尿症 III 型	有機酸血症
OPLA H	26873	0178814	O14841	165	5-オキソプロ リナーゼ欠 損症	有機酸血症
OXC T1	5019	0083720	A0A024R0 40, P55809	166	SCOT 欠損症	有機酸血症
POLG	5428	0140521	E5KNU5, P54098	167	3-メチルグル タコン酸尿 症	有機酸血症
PPM1 K	152926	0163644	Q8N3J5	168	メープルシ ロップ尿症 (MSUD), バ リエント型	有機酸血症
SERA C1	84947	0122335	Q96JX3	169	メグデル症 候群	有機酸血症
SLC2 5A1	6576	0100075	D9HTE9, B4DP62, P53007	170	D,L-2-ヒドロ キシグルタ ル酸尿症	有機酸血症

10

20

30

40

50

SUCL A2	8803	0136143	E5KS60, Q9P2R7, Q9Y4T0	171	コハク酸塩- CoA リガー ゼ欠損症, メ チルマロン 酸尿症	有機酸血症
SUCL G1	8802	0163541	P53597	172	コハク酸塩- CoA リガー ゼ欠損症, メ チルマロン 酸尿症	有機酸血症
TAZ	6901	0102125	A0A0S2Z4 K0, Q16635, A6XNE1, A0A0S2Z4 E6, A0A0S2Z4 K9, A0A0S2Z4 F4	173	バース症候 群	有機酸血症
AGK	55750	0006530,02 62327	A4D1U5, Q53H12	174	3-メチルグル タコン酸尿 症	有機酸血症
CLPB	81570	0162129	Q9H078, A0A140V K11	175	3-メチルグル タコン酸尿 症	有機酸血症
TME M70	54968	0175606	Q9BUB7	176	3-メチルグル タコン酸尿 症	有機酸血症
ALD H18A 1	5832	0059573	P54886	177	ALDH18A1 関連皮膚弛 緩症	尿素回路障害

10

20

30

40

50

OAT	4942	0065154	A0A140VJ Q4, P04181	178	脳回転状萎縮 (OAT)	尿素回路障害
CA5A	763	0174990	P35218	179	炭酸脱水酵素欠損症	尿素回路障害
GLU D1	2746	0148672	P00367, E9KL48	180	グルタミン酸デヒドロゲナーゼ欠損症	尿素回路障害
GLU L	2752	0135821	A8YXX4, P15104	181	グルタミンシンテターゼ欠損症	尿素回路障害
UMP S	7372	0114491	A8K5J1, P11172	182	オロト酸尿症	尿素回路障害
SLC2 2A5	6584	0197375	O76082	183	カルニチン-アシルカルニチントランスロカーゼ(CACT)欠損症	脂肪酸酸化
CPT1 A	1374	0110090	P50416, A0A024R5 F4, B2RAQ8, Q8WZ48	184	カルニチンパルミトイルトランスフェラーゼ I 型(CPT I)欠損症	脂肪酸酸化
HAD HA	3030	0084754	E9KL44, P40939	185	長鎖 3-ヒドロキシアシル-CoA デヒドロゲナーゼ(LCHAD)欠損症	脂肪酸酸化

10

20

30

40

50



HAD H	3033	0138796	Q16836, B3KTT6	186	中/短鎖アシル-CoA デヒドロゲナーゼ (M/SCHAD) 欠損症	脂肪酸酸化	
SLC5 2A1	55065	0132517	Q9NWF4	187	リボフラビントランスポーター欠損症	脂肪酸酸化	10
SLC5 2A2	79581	0185803	Q9HAB3	188	リボフラビントランスポーター欠損症	脂肪酸酸化	
SLC5 2A3	113278	0101276	K0A6P4, Q9NQ40	189	リボフラビントランスポーター欠損症	脂肪酸酸化	20
HAD HB	3032	0138029	P55084, F5GZQ3	190	三頭酵素欠損症	脂肪酸酸化	
GYS2	2998	0111713	P54840	191	GSD 0 (グリコーゲンシンターゼ, 肝臓アイソフォーム)	肝臓グリコーゲン蓄積障害	30
PYGL	5836	0100504	P06737	192	GSD VI (ハーズ病)	肝臓グリコーゲン蓄積障害	
SLC2 A2	6514	0163581	P11168, Q6PAU8	193	ファンコニ・ビッケル症候群	肝臓グリコーゲン蓄積障害	
ALG1	56052	0033011	Q9BT22	194	ALG1-CDG	グリコシル化障害	40

ALG2	85365	0119523	A0A024R1 84, Q9H553	195	ALG2 関連筋 無力症候群	グリコシル化障 害
ALG3	10195	0214160	Q92685, C9J7S5	196	ALG3-CDG	グリコシル化障 害
ALG6	29929	0088035	Q9Y672	197	ALG6-CDG	グリコシル化障 害
ALG8	79053	0159063	Q9BVK2, A0A024R5 K5	198	ALG8-CDG	グリコシル化障 害
ALG9	79796	0086848	Q9H6U8	199	ALG9-CDG	グリコシル化障 害
ALG1 1	440138	0253710	Q2TAA5	200	ALG11-CDG	グリコシル化障 害
ALG1 2	79087	0182858	A0A024R4 V6, Q9BV10	201	ALG12-CDG	グリコシル化障 害
ALG1 3	79868	0101901	Q9NP73, A0A087W X43, A0A087W T15	202	ALG13-CDG	グリコシル化障 害
ATP6 V0A2	23545	0185344	Q9Y487	203	ATP6V0A2 関連皮膚弛 緩症	グリコシル化障 害
B3GL CT	145173	0187676	Q6Y288	204	B3GLCT- CDG	グリコシル化障 害
CHST 14	113189	0169105	Q8NCH0	205	CHST14- CDG	グリコシル化障 害
COG1	9382	0166685	Q8WTW3	206	COG1-CDG	グリコシル化障 害
COG2	22796	0135775	Q14746, B1ALW7	207	COG2-CDG	グリコシル化障 害

10

20

30

40

50

COG4	25839	0103051	A0A0A0M S45, Q8N8L9, Q9H9E3, J3KNI1	208	COG4-CDG	グリコシル化障 害
COG5	10466	0164597,02 84369	Q9UP83	209	COG5-CDG	グリコシル化障 害
COG6	57511	0133103	A0A140VJ G7, Q9Y2V7, A0A024R DW5	210	COG6-CDG	グリコシル化障 害
COG7	91949	0168434	A0A0S2Z6 52, P83436	211	COG7-CDG	グリコシル化障 害
COG8	84342	0272617	A0A024R6 Z6, Q96MW5	212	COG8-CDG	グリコシル化障 害
DOL K	22845	0175283	A0A0S2Z5 97, Q9UPQ8	213	DOLK-CDG	グリコシル化障 害
DHD DS	79947	0117682	Q86SQ9	214	DHDDS-CDG	グリコシル化障 害
DPA GT1	1798	0172269	A0A024R3 H8, Q9H3H5	215	DPAGT1- CDG	グリコシル化障 害
DPM 1	8813	0000419	O60762, Q5QPK2, A0A0S2Z4 Y5	216	DPM1-CDG	グリコシル化障 害
DPM 2	8818	0136908	O94777	217	DPM2-CDG	グリコシル化障 害
DPM 3	54344	0179085	A0A140VJ I4, Q9P2X0, Q86TM7	218	DPM3-CDG	グリコシル化障 害

10

20

30

40

50

G6PC3	92579	0141349	Q9BUM1	219	先天性好中球減少症	グリコシル化障害
GFPT1	2673	0198380	Q06210	220	先天性筋無力症候群	グリコシル化障害
GMPPA	29926	0144591	A0A024R482, Q96IJ6	221	GMPPA-CDG	グリコシル化障害
GMPPB	29925	0173540	Q9Y5P6	222	先天性筋ジストロフィー, 先天性筋無力症候群, 及びジストログリカノパチー	グリコシル化障害
MAGT1	84061	0102158	A0A087WU53, Q9H0U3	223	MAGT1-CDG; マグネシウム欠乏, エプスタイン・バーウイルス感染症及び腫瘍(XMEN)症候群を伴う X連鎖免疫不全	グリコシル化障害
MAN1B1	11253	0177239	Q9UKM7	224	MAN1B1-CDG	グリコシル化障害
MGAT2	4247	0168282	Q10469	225	MGAT2-CDG	グリコシル化障害
MOGS	7841	0115275	Q13724, Q58F09	226	MOGS-CDG	グリコシル化障害

10

20

30

40

50

MPDU1	9526	0129255	J3QW43, O75352, A0A0S2Z4 W8, B4DLH7	227	MPDU1-CDG	グリコシル化障 害
MPI	4351	0178802	H3BPP3, Q8NHZ6, B4DW50, F5GX71, P34949, H3BPB8	228	MPI-CDG	グリコシル化障 害
NGLY1	55768	0151092	Q96IV0	229	NGLY1-CDG	グリコシル化障 害
PGM1	5236	0079739	B7Z6C2, P36871, B4DDQ8	230	PGM1-CDG	グリコシル化障 害
PGM3	5238	0013375	O95394, A0A087W T27	231	PGM3-CDG	グリコシル化障 害
RFT1	91869	0163933	Q96AA3	232	RFT1-CDG	グリコシル化障 害
SEC23B	10483	0101310	Q15437, B4DJW8	233	SEC23B- CDG	グリコシル化障 害
SLC35A1	10559	0164414	P78382	234	SLC35A1- CDG	グリコシル化障 害
SLC35A2	7355	0102100	P78381, A6NFI1, A6NKM8, B4DE15	235	SLC35A2- CDG	グリコシル化障 害
SLC35C1	55343	0181830	Q96A29, B3KQH0	236	SLC35C1- CDG	グリコシル化障 害
SSR4	6748	0180879	P51571	237	SSR4-CDG	グリコシル化障 害

10

20

30

40

50

SRD5 A3	79644	0128039	Q9H8P0	238	SRD5A3- CDG	グリコシル化障 害
TME M165	55858	0134851	Q9HC07	239	TMEM165- CDG	グリコシル化障 害
TRIP 11	9321	0100815	Q15643	240	TRIP11-CDG	グリコシル化障 害
TUSC 3	7991	0104723	Q13454	241	TUSC3-CDG	グリコシル化障 害
ALG1 4	199857	0172339	Q96F25	242	ALG14-CDG	グリコシル化障 害
B4GA LT1	2683	0086062	P15291, W6MEN3	243	B4GALT1- CDG	グリコシル化障 害
DDO ST	1650	0244038	A0A024R AD5, P39656	244	DDOST-CDG	グリコシル化障 害
NUS1	116150	0153989	Q96E22	245	NUS1-CDG	グリコシル化障 害
RPN2	6185	0118705	P04844	246	RPN2-CDG	グリコシル化障 害
SEC2 3A	10484	0100934	Q15436	247	SEC23A- CDG	グリコシル化障 害
SLC3 5A3	23443	0117620	Q9Y2D2, A0A1W2P RT7, A0A1W2P SD1, A0A1W2P QL8	248	SLC35A3- CDG	グリコシル化障 害
ST3G AL3	6487	0126091	Q11203	249	ST3GAL3- CDG	グリコシル化障 害

10

20

30

40

50

STT3 A	3703	0134910	P46977	250	STT3A-CDG	グリコシル化障 害
STT3 B	201595	0163527	Q8TCJ2	251	STT3B-CDG	グリコシル化障 害
AGA	175	0038002	P20933	252	アスパルチ ルグルコサ ミン尿症	リソソーム蓄積 障害
ARS A	410	0100299	A0A0C4D FZ2, B4DVI5, P15289	253	異染性白質 ジストロフ イー	リソソーム蓄積 障害
ARSB	411	0113273	A0A024R AJ9, P15848, A8K4A0	254	ムコ多糖症 VI型	リソソーム蓄積 障害
ASA H1	427	0104763	A8K0B6, Q13510, Q53H01	255	ファーバー 病	リソソーム蓄積 障害
ATP1 3A2	23400	0159363	Q8N4D4, Q9NQ11, Q8NBS1	256	ニューロン セロイドリ ポフスチン 症 12(CLN12), クフォー・ ラケブ症候 群(KRS)	リソソーム蓄積 障害
CLN3	1201	0188603,02 61832	A0A024Q ZB8, Q13286, B4DMY6, Q2TA70, B4DFF3	257	ニューロン セロイドリ ポフスチン 症 3(CLN3)	リソソーム蓄積 障害

10

20

30

40

50

CLN5	1203	0102805	A0A024R6 44, O75503	258	ニューロン セロイドリ ポフスチン 症 5(CLN5)	リソソーム蓄積 障害
CLN6	54982	0128973	A0A024R6 01, Q9NWW5	259	ニューロン セロイドリ ポフスチン 症 6 (CLN6)	リソソーム蓄積 障害
CLN8	2055	0182372,02 78220	A0A024Q Z57, Q9UBY8	260	ニューロン セロイドリ ポフスチン 症 8 (CLN8)	リソソーム蓄積 障害
CTNS	1497	0040531	A0A0S2Z3 I9, O60931, A0A0S2Z3 K3	261	シスチン症	リソソーム蓄積 障害
CTSA	5476	0064601	P10619, X6R8A1, B4E324, X6R5C5	262	ガラクトシ アリドーシ ス	リソソーム蓄積 障害
CTSD	1509	0117984	P07339, V9HWI3	263	ニューロン セロイドリ ポフスチン 症 10(CLN10)	リソソーム蓄積 障害
CTSF	8722	0174080	Q9UBX1	264	ニューロン セロイドリ ポフスチン 症 13(CLN13)	リソソーム蓄積 障害
CTSK	1513	0143387	P43235	265	濃化異骨症	リソソーム蓄積 障害

10

20

30

40

50



DNAJ C5	80331	0101152	Q6AHX3, Q9H3Z4	266	ニューロン セロイドリ ポフスチン 症 4(CLN4)	リソソーム蓄積 障害
FUC A1	2517	0179163	P04066, B5MDC5	267	フコシドー シス	リソソーム蓄積 障害
GAA	2548	0171298	P10253	268	ポンペ病	リソソーム蓄積 障害
GAL C	2581	0054983	A0A0A0M QV0, P54803	269	クラッペ病	リソソーム蓄積 障害
GAL NS	2588	0141012	P34059, Q96I49, Q6YL38	270	ムコ多糖症 IVa 型	リソソーム蓄積 障害
GLA	2717	0102393	P06280, Q53Y83	271	ファブリー 病	リソソーム蓄積 障害
GLB1	2720	0170266	P16278, B7Z6Q5	272	GM1 ガング リオシドー シス, ムコ多 糖症 IVb	リソソーム蓄積 障害
GM2 A	2760	0196743	P17900	273	GM2-ガング リオシドー シス, AB バ リエーション	リソソーム蓄積 障害
GNPT AB	79158	0111670	Q3T906	274	ムコリピド ーシス II 型 アルファ/ベ ータ, ムコリ ピドーシス III アルファ/ ベータ	リソソーム蓄積 障害

10

20

30

40

50

GNPT G	84572	0090581	Q9UJJ9	275	ムコリピド ーシス III ガ ンマ	リソソーム蓄積 障害
GNS	2799	0135677	A0A024R BC5, P15586, Q7Z3X3	276	ムコ多糖症 IIID 型	リソソーム蓄積 障害
GRN	2896	0030582	P28799	277	ニューロン セロイドリ ポフスチン 症 11(CLN11), 前頭側頭型 認知症	リソソーム蓄積 障害
GUS B	2990	0169919	P08236	278	ムコ多糖症 VII 型	リソソーム蓄積 障害
HEX A	3073	0213614	A0A0S2Z3 W3, P06865, B4DVA7, H3BP20	279	テイ・サッ クス病	リソソーム蓄積 障害
HEX B	3074	0049860	A0A024R AJ6, P07686, Q5URX0	280	サンドホフ 病	リソソーム蓄積 障害
HGS NAT	138050	0165102	Q68CP4, Q8IVU6	281	ムコ多糖症 IIIC 型	リソソーム蓄積 障害
HYA L1	3373	0114378	A0A024R2 X3, Q12794, B3KUI5, A0A0S2Z3 Q0	282	ムコ多糖症 IX 型	リソソーム蓄積 障害

10

20

30

40

50

IDS	3423	0010404	P22304, B4DGD7	283	ムコ多糖症 II型	リソソーム蓄積 障害
IDUA	3425	0127415	P35475	284	ムコ多糖症 I 型	リソソーム蓄積 障害
KCT D7	154881	0243335	Q96MP8, A0A024R DN7	285	ニューロン セロイドリ ポフスチン 症 14(CLN14)	リソソーム蓄積 障害
LAM P2	3920	0005893	P13473	286	ダノン病	リソソーム蓄積 障害
MAN 2B1	4125	0104774	O00754, A8K6A7	287	アルファ-マ ンノシドー シス	リソソーム蓄積 障害
MAN BA	4126	0109323	O00462	288	ベータ-マン ノシドーシ ス	リソソーム蓄積 障害
MCO LN1	57192	0090674	Q9GZU1	289	ムコリピド ーシス IV型	リソソーム蓄積 障害
MFS D8	256471	0164073	Q8NHS3	290	ニューロン セロイドリ ポフスチン 症 7(CLN7)	リソソーム蓄積 障害
NAG A	4668	0198951	A0A024R1 Q5, P17050	291	シンドラー 病	リソソーム蓄積 障害
NAG LU	4669	0108784	A0A140VJ E4, P54802	292	ムコ多糖症 IIIB	リソソーム蓄積 障害

10

20

30

40

50

NEU1	4758	0204386,02 27315,0227 129,022395 7,0234846,0 184494,022 8691,02343 43	Q5JQI0, Q99519	293	ムコリピド ーシス I 型, シアリド ーシス I	リソソーム蓄積 障害	
NPC1	4864	0141458	O15118	294	ニーマン・ ピック C 型	リソソーム蓄積 障害	10
NPC2	10577	0119655	A0A024R6 C0, P61916, G3V3E8	295	ニーマン・ ピック C 型	リソソーム蓄積 障害	
SGSH	6448	0181523	P51688	296	ムコ多糖症 IIIA	リソソーム蓄積 障害	
PPT1	5538	0131238	P50897	297	ニューロン セロイドリ ポフスチン 症 1(CLN1)	リソソーム蓄積 障害	20
PSAP	5660	0197746	P07602, A0A024Q ZQ2	298	プロサポシ ン欠損症, SapA 欠損症 (クラッペ バリエント), SapB 欠損症 (MLD バリア ント), SapC 欠損症(ゴー シェバリア ント)	リソソーム蓄積 障害	30

40

50

SLC17A5	26503	0119899	Q9NRA2	299	幼児期シアル酸蓄積症, サラ病	リソソーム蓄積障害
SMPD1	6609	0166311	P17405, Q59EN6, E9LUE8, Q8IUN0, E9LUE9	300	ニーマンピック A 及び B 型	リソソーム蓄積障害
SUMF1	285362	0144455	Q8NBK3	301	多重スルファターゼ欠損症	リソソーム蓄積障害
TPP1	1200	0166340	O14773	302	ニューロンセロイドリポフスチン症 2(CLN2)	リソソーム蓄積障害
AHCY	191	0101444	P23526, Q1RMG2	303	高メチオニン血症	アミノ酸症
GNMT	27232	0124713	A0A0S2Z5F2, Q14749, V9HW60	304	高メチオニン血症	アミノ酸症
MAT1A	4143	0151224	Q00266	305	高メチオニン血症	アミノ酸症
GCH1	2643	0131979	A0A024R642, P30793, Q8IZH9	306	BH4 補因子欠損症	アミノ酸症
PCBD1	5092	0166228	P61457	307	BH4 補因子欠損症	アミノ酸症
PTS	5805	0150787	Q03393	308	BH4 補因子欠損症	アミノ酸症

10

20

30

40

50

QDP R	5860	0151552	A0A140V KA9, P09417	309	BH4 補因子 欠損症	アミノ酸症
SPR	6697	0116096	P35270	310	BH4 補因子 欠損症	アミノ酸症
DNAJ C12	56521	0108176	Q6IAH1, Q9UKB3	311	フェニルア ラニン, チロ シン, 及びト リプトファ ンヒドロキ シラーゼヒ ートショッ クコシャペ ロン欠損症	アミノ酸症
ALD H4A1	8659	0159423	P30038, A0A024R AD8	312	高プロリン 血症	アミノ酸症
PRO DH	5625	0100033	O43272	313	高プロリン 血症	アミノ酸症
HPD	3242	0158104	P32754	314	チロシン血 症 II 型	アミノ酸症
GBA	2629	0177628,02 62446	A0A068F6 58, P04062, B7Z6S9	315	ゴーシェ病	
HGD	3081	0113924	Q93099, B3KW64	316	アルカプト ン尿症	
AMN	81693	0166126	Q9BXJ7, B3KP64	317	混合型メチ ルマロン酸 血症及びホ モシスチン 尿症	有機酸血症

10

20

30

40

50

CD320	51293	0167775	Q9NPF0	318	混合型メチルマロン酸血症及びホモシスチン尿症	有機酸血症
CUBN	8029	0107611	O60494	319	混合型メチルマロン酸血症及びホモシスチン尿症	有機酸血症
GIF	2694	0134812	P27352	320	混合型メチルマロン酸血症及びホモシスチン尿症	有機酸血症
TCN1	6947	0134827	P20061	321	混合型メチルマロン酸血症及びホモシスチン尿症	有機酸血症
TCN2	6948	0185339	P20062	322	混合型メチルマロン酸血症及びホモシスチン尿症	有機酸血症
PREPL	9581	0138078	Q4J6C6	323	シスチン尿症	アミノ酸症
PHGDH	26227	0092621	O43175	324	セリン生合成の障害	アミノ酸症

10

20

30

40

50

PSAT 1	29968	0135069	A0A024R2 80, Q9Y617, A0A024R2 22	325	セリン生合 成の障害	アミノ酸症
PSPH	5723	0146733	A0A024R DL3, P78330	326	セリン生合 成の障害	アミノ酸症
AMT	275	0145020	A0A024R2 U7, P48728	327	グリシン脳 症	アミノ酸症
GCS H	2653	0140905	P23434	328	グリシン脳 症	アミノ酸症
GLD C	2731	0178445	P23378	329	グリシン脳 症	アミノ酸症
LIAS	11019	0121897	O43766, Q6P5Q6, B4E0L7, A0A024R9 W0, A0A1W2P QE9, A0A1X7S BR7	330	グリシン脳 症	アミノ酸症
NFU1	27247	0169599	Q9UMS0	331	グリシン脳 症	アミノ酸症
SLC6 A9	6536	0196517	P48067, B7Z3W8, B7Z589	332	グリシン脳 症	アミノ酸症
SLC2 A1	6513	0117394	P11166, Q59GX2	333	グルコース トランスポ ーター1型欠 損症	炭水化物障害

10

20

30

40

50



ATP7A	538	0165240	B4DRW0, Q04656, Q762B6	334	ATP7A 関連障害 銅代謝障害	金属輸送障害	
AP1S1	1174	0106367	A0A024QYT6, P61966	335	銅代謝障害	金属輸送障害	
CP	1356	0047457	A5PL27, P00450	336	銅代謝障害	金属輸送障害	10
SLC33A1	9197	0169359	O00400	337	銅代謝障害	金属輸送障害	
PEX7	5191	0112357	O00628, Q6FGN1	338	成人レフサム病 根性点状軟骨異形成症 スペクトル	ペルオキシソーム障害	
PHYH	5264	0107537	O14832	339	成人レフサム病	ペルオキシソーム障害	20
AGPS	8540	0018510	O00116, B7Z3Q4	340	根性点状軟骨異形成症 スペクトル	ペルオキシソーム障害	
GNPAT	8443	0116906	O15228	341	根性点状軟骨異形成症 スペクトル	ペルオキシソーム障害	
ABCD1	215	0101986	P33897	342	X連鎖副腎白質ジストロフィー	ペルオキシソーム障害	30
ACOX1	51	0161533	Q15067	343	X連鎖副腎白質ジストロフィー	ペルオキシソーム障害	

PEX1	5189	0127980	O43933, A0A0C4D G33, B4DER6	344	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX2	5828	0164751	P28328	345	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX3	8504	0034693	P56589	346	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX5	5830	0139197	A0A0S2Z4 80, P50542, B4DR50, A0A0S2Z4 F3, A0A0S2Z4 H1, B4E0T2	347	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX6	5190	0124587	A0A024R D09, Q13608	348	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX1 0	5192	0157911	A0A024R0 68, O60683, A0A024R0 A4	349	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX1 2	5193	0108733	O00623	350	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX1 3	5194	0162928	Q92968	351	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害

10

20

30

40

50

PEX1 4	5195	0142655	O75381	352	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX1 6	9409	0121680	Q9Y5Y5	353	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX1 9	5824	0162735	P40855, A0A0S2Z4 97	354	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
PEX2 6	55670	0215193	A0A024R1 00, Q7Z412, A0A0S2Z5 M7, Q7Z2D7	355	X連鎖副腎 白質ジスト ロフィー	ペルオキシソー ム障害
AMA CR	23600	0242110	Q9UHK6	356	ツェルウェ ガースペク トル障害	ペルオキシソー ム障害
ADA	100	0196839	A0A0S2Z3 81, P00813, F5GW14	357	プリン代謝 障害	プリン代謝障害
ADSL	158	0239900	P30566, X5D8S6, X5D7W4, A0A1B0G WJ0	358	プリン代謝 障害	プリン代謝障害
AMP D1	270	0116748	P23109	359	プリン代謝 障害	プリン代謝障害
GPH N	10243	0171723	Q9NQX3	360	プリン代謝 障害	プリン代謝障害
MOC OS	55034	0075643	Q96EN8	361	プリン代謝 障害	プリン代謝障害

10

20

30

40

50

MOC S1	4337	0124615	A0A024R D17, Q9NZB8	362	プリン代謝障害	プリン代謝障害
PNP	4860	0198805	P00491, V9HWH6	363	プリン代謝障害	プリン代謝障害
XDH	7498	0158125	P47989	364	プリン代謝障害	プリン代謝障害
SUOX	6821	0139531	A0A024R B79, P51687	365	プリン代謝障害	プリン代謝障害
OGDH	4967	0105953	A0A140VJ Q5, Q02218, B4E3E9, E9PCR7, E9PDF2	366	2-ケトグルタル酸デヒドロゲナーゼ欠損症	ピルビン酸代謝及びトリカルボン酸サイクル欠陥
SLC25A19	60386	0125454	Q5JPC1, Q9HC21	367	2-ケトグルタル酸デヒドロゲナーゼ欠損症	ピルビン酸代謝及びトリカルボン酸サイクル欠陥
DHTKD1	55526	0181192	Q96HY7	368	2-ケトグルタル酸デヒドロゲナーゼ欠損症	ピルビン酸代謝及びトリカルボン酸サイクル欠陥
SLC13A5	284111	0141485	Q68D44, Q86YT5	369	クエン酸塩トランスポーター欠損症	ピルビン酸代謝及びトリカルボン酸サイクル欠陥
FH	2271	0091483	A0A0S2Z4 C3, P07954	370	フマラーゼ欠損症	ピルビン酸代謝及びトリカルボン酸サイクル欠陥

10

20

30

40

50

DLA T	1737	0150768	P10515, Q86YI5	371	ピルビン酸 デヒドロゲ ナーゼ欠損 症	ピルビン酸代謝 及びトリカルボ ン酸サイクル欠 陥
MPC1	51660	0060762	Q5TI65, Q9Y5U8	372	ピルビン酸 デヒドロゲ ナーゼ欠損 症	ピルビン酸代謝 及びトリカルボ ン酸サイクル欠 陥
PDH A1	5160	0131828	A0A024R BX9, P08559	373	ピルビン酸 デヒドロゲ ナーゼ欠損 症	ピルビン酸代謝 及びトリカルボ ン酸サイクル欠 陥
PDH B	5162	0168291	P11177	374	ピルビン酸 デヒドロゲ ナーゼ欠損 症	ピルビン酸代謝 及びトリカルボ ン酸サイクル欠 陥
PDH X	8050	0110435	O00330	375	ピルビン酸 デヒドロゲ ナーゼ欠損 症	ピルビン酸代謝 及びトリカルボ ン酸サイクル欠 陥
PDP1	54704	0164951	Q9P0J1, Q6P1N1, A0A024R9 C0	376	ピルビン酸 デヒドロゲ ナーゼ欠損 症	ピルビン酸代謝 及びトリカルボ ン酸サイクル欠 陥
ABC C2	1244	0023839	Q92887	377	デュビン・ ジョンソン 症候群	
SLCO 1B1	10599	0134538	A0A024R AU7, Q05CV5, Q9Y6L6	378	ローター症 候群	

10

20

30

40

50

SLCO1B3	28234	0111700	B3KP78, Q9NPD5	379	ローター症候群	
HFE2	148738	0168509	Q6ZVN8, A8K466, A0A024R4 F5	380	ヘモクロマトーシス, 2A型	
ADAMTS13	11093	0160323,02 81244	Q76LX8	381	ADAMTS-13欠損症による先天性血栓性血小板減少性紫斑病	10
PYGM	5837	0068976	P11217	382	マッカードール病	
COL1A2	1278	0164692	A0A0S2Z3 H5, P08123	383	エーラス・ダンロス症候群, 心臓弁型	20
TNFRSF11B	4982	0164761	O00300	384	若年性パジエット病	
TSC1	7248	0165699	Q86WV8, Q92574, X5D9D2, Q32NF0	385	結節性硬化症	
TSC2	7249	0103197	P49815, X5D7Q2, B3KWH7, Q5HYF7, H3BMQ0, X5D2U8	386	結節性硬化症	30
DHCR7	1717	0172893	A0A024R5 F7, Q9UBM7	387	スミス・レムリ・オピッツ症候群	

PGK1	5230	0102144	P00558, V9HWF4	388	D-グリセリン酸血症	
VLDLR	7436	0147852	P98155, Q5VVF5	389	平衡障害症候群	
KYNU	8942	0115919	Q16719	390	ヒドロキシキヌレニン尿症による脳症	
F5	2153	0198734	P12259	391	第5因子欠損症	
C3	718	0125730	B4DR57, P01024, V9HWA9	392	C3異常(anomaly)を伴う非典型溶血性尿毒症症候群	
COL4A1	1282	0187498	A5PKV2, F5H5K0, P02462	393	常染色体優性家族性血尿症-網膜細動脈屈曲-拘縮	
CFH	3075	0000971	A0A024R962, P08603, A0A0D9SG88	394	非典型溶血性尿毒症症候群	
SLC12A2	6558	0064651	P55011, Q53ZR1, B7ZM24	395	バーター症候群I型(新生児)	

10

20

30

40

50

GK	2710	0198814	B4DH54, P32189	396	グリセロール キナーゼ 欠損症	
SFTP C	6440	0168484	A0A0A0M TC9, P11686, A0A0S2Z4 Q0, E5RI64	397	サーファク タント代謝 欠損症を伴 う慢性呼吸 窮迫症	
CRT AP	10491	0170275	O75718	398	骨形成不全 症 VII	
P3H1	64175	0117385	Q32P28	399	骨形成不全 症 VIII	
COL7 A1	1294	0114270	Q02388, Q59F16	400	常染色体劣 性栄養障害 性表皮水疱 症	
PKLR	5313	0143627	P30613	401	ピルビン酸 キナーゼ欠 損症	
TAL DO1	6888	0177156	A0A140V K56, P37837	402	トランスア ルドラーゼ 欠損症	
TF	7018	0091513	A0PJA6, P02787, Q06AH7	403	無トランス フェリン血 症(家族性低 トランスフ ェリン血症)	
EPCA M	4072	0119888	P16422	404	腸上皮異形 成	

10

20

30

40

50



VHL	7428	0134086	A0A024R2 F2, P40337, A0A0S2Z4 K1	405	家族性赤血 球増加症 2 型;フォンヒ ッペルリン ドウ病	
GC	2638	0145321	P02774	406	ビタミン D 欠損症	
SERP INA1	5265	0197249,02 77377	E9KL23, P01009	407	アルファ-1 アンチトリ プシン欠損 症	
ABC C6	368	0091262,02 75331	O95255	408	弾性線維性 仮性黄色腫	
F8	2157	0185010	P00451	409	血友病 A	
F9	2158	0101981	P00740	410	血友病 B	
ApoB	338	0084674	P04114	411	家族性高コ レステロー ル血症	
PCSK 9	255738	0169174	Q8NBP7	412	家族性高コ レステロー ル血症	
LDLR AP1	26119	0157978	B3KR97, Q5SW96	413	家族性高コ レステロー ル血症	
ABC G5	64240	0138075	Q9H222	414	シトステロ ール血症	
ABC G8	64241	0143921	Q9H221	415	シトステロ ール血症	

10

20

30

40

50

LCAT	3931	0213398	A0A140V K24, P04180	416	レシチンコ レステロー ルアシルト ランスフェ ラーゼ欠損 症	
SPIN K5	11005	0133710	Q9NQ38	417	ネザートン 症候群	
GNE	10020	0159921	Q9Y223	418	封入体ミオ パシー2	

10

## 【 0 3 1 6 】

いくつかの実施形態では、標的化脂質粒子またはレンチウイルスベクターは、T細胞を標的とすることが可能な外因性物質を含有する。いくつかの実施形態では、T細胞を標的とすることが可能な外因性物質は、キメラ抗原受容体（CAR）、T細胞受容体、インテグリン、イオンチャネル、孔形成タンパク質、トル様受容体、インターロイキン受容体、細胞接着タンパク質、または輸送タンパク質である。

20

## 【 0 3 1 7 】

いくつかの実施形態では、CARは、抗原結合ドメイン、膜貫通ドメイン、及びシグナル伝達ドメイン（例えば、1、2または3つのシグナル伝達ドメイン）を含む第一世代CARであるか、またはそれを含む。いくつかの実施形態では、CARは、抗原結合ドメイン、膜貫通ドメイン、及び少なくとも3つのシグナル伝達ドメインを含む第三世代CARを含む。いくつかの実施形態では、第四世代CARは、抗原結合ドメイン、膜貫通ドメイン、3または4つのシグナル伝達ドメイン、及びCARの成功裏のシグナル伝達時にサイトカイン遺伝子の発現を誘導するドメインを含む。いくつかの実施形態では、抗原結合ドメインは、scFvまたはFabであるか、またはそれを含む。

30

## 【 0 3 1 8 】

いくつかの実施形態では、CAR抗原結合ドメインは、抗体またはその抗原結合部分であるか、またはそれを含む。いくつかの実施形態では、CAR抗原結合ドメインは、scFvまたはFabであるか、またはそれを含む。いくつかの実施形態では、CAR抗原結合ドメインは、T細胞アルファ鎖抗体；T細胞鎖抗体；T細胞鎖抗体；T細胞鎖抗体；CCR7抗体；CD3抗体；CD4抗体；CD5抗体；CD7抗体；CD8抗体；CD11b抗体；CD11c抗体；CD16抗体；CD19抗体；CD20抗体；CD21抗体；CD22抗体；CD25抗体；CD28抗体；CD34抗体；CD35抗体；CD40抗体；CD45RA抗体；CD45RO抗体；CD52抗体；CD56抗体；CD62L抗体；CD68抗体；CD80抗体；CD95抗体；CD117抗体；CD127抗体；CD133抗体；CD137（4-1BB）抗体；CD163抗体；F4/80抗体；IL-4Ra抗体；Sca-1抗体；CTLA-4抗体；GITR抗体；GARP抗体；LAP抗体；グランザイムB抗体；LFA-1抗体；MR1抗体；uPAR抗体；またはトランスフェリン受容体抗体のscFvまたはFab断片を含む。

40

## 【 0 3 1 9 】

いくつかの実施形態では、CAR結合ドメインは、細胞の細胞表面抗原に結合する。いくつかの実施形態では、細胞表面抗原は、1つのタイプの細胞に特徴的である。いくつかの実施形態では、細胞表面抗原は、複数のタイプの細胞に特徴的である。

## 【 0 3 2 0 】

50

いくつかの実施形態では、CARの抗原結合ドメインは、T細胞に特徴的な抗原を標的とする。いくつかの実施形態では、T細胞に特徴的な抗原は、T細胞に特徴的な細胞表面受容体、膜輸送タンパク質（例えば、能動的または受動的輸送タンパク質、例えば、イオンチャネルタンパク質、孔形成タンパク質など）、膜貫通受容体、膜酵素、及び/または細胞接着タンパク質から選択される。いくつかの実施形態では、T細胞に特徴的な抗原は、Gタンパク質共役型受容体、受容体チロシンキナーゼ、チロシンキナーゼ関連受容体、受容体様チロシンホスファターゼ、受容体セリン/トレオニンキナーゼ、受容体グアニル酸シクラーゼ、ヒスチジンキナーゼ関連受容体、AKT1；AKT2；AKT3；ATF2；BCL10；CALM1；CD3D（CD3）；CD3E（CD3）；CD3G（CD3）；CD4；CD8；CD28；CD45；CD80（B7-1）；CD86（B7-2）；CD247（CD3）；CTLA4（CD152）；ELK1；ERK1（MAPK3）；ERK2；FOS；FYN；GRAP2（GADS）；GRB2；HLA-DRA；HLA-DRB1；HLA-DRB3；HLA-DRB4；HLA-DRB5；HRAS；IKBKA（CHUK）；IKBKB；IKBKE；IKBKG（NEMO）；IL2；ITPR1；ITK；JUN；KRAS2；LAT；LCK；MAP2K1（MEK1）；MAP2K2（MEK2）；MAP2K3（MKK3）；MAP2K4（MKK4）；MAP2K6（MKK6）；MAP2K7（MKK7）；MAP3K1（MEKK1）；MAP3K3；MAP3K4；MAP3K5；MAP3K8；MAP3K14（NIK）；MAPK8（JNK1）；MAPK9（JNK2）；MAPK10（JNK3）；MAPK11（p38）；MAPK12（p38）；MAPK13（p38）；MAPK14（p38）；NCK；NFAT1；NFAT2；NFKB1；NFKB2；NFKBIA；NRAS；PAK1；PAK2；PAK3；PAK4；PIK3C2B；PIK3C3（VPS34）；PIK3CA；PIK3CB；PIK3CD；PIK3R1；PKCA；PKCB；PKCM；PKCQ；PLCY1；PRF1（パーフォリン）；PTEN；RAC1；RAF1；RELA；SDF1；SHP2；SLP76；SOS；SRC；TBK1；TCRA；TEC；TRAF6；VAV1；VAV2；またはZAP70であり得る。

#### 【0321】

いくつかの実施形態では、CARの抗原結合ドメインは、障害に特徴的な抗原を標的とする。いくつかの実施形態では、疾患または障害は、CD4+ T細胞に関連する。いくつかの実施形態では、疾患または障害は、CD8+ T細胞に関連する。

#### 【0322】

いくつかの実施形態では、CAR膜貫通ドメインは、T細胞受容体、CD28、CD3イプシロン、CD45、CD4、CD5、CD8、CD9、CD16、CD22、CD33、CD37、CD64、CD80、CD86、CD134、CD137、CD154、またはその機能的バリエーションのアルファ、ベータまたはゼータ鎖の少なくとも膜貫通領域を含む。いくつかの実施形態では、膜貫通ドメインは、CD8、CD8、4-1BB/CD137、CD28、CD34、CD4、FcRI、CD16、OX40/CD134、CD3、CD3、CD3、CD3、TCR、TCR、TCR、CD32、CD64、CD64、CD45、CD5、CD9、CD22、CD37、CD80、CD86、CD40、CD40L/CD154、VEGFR2、FAS、及びFGFR2B、またはその機能的バリエーションの少なくとも膜貫通領域（複数可）を含む。

#### 【0323】

いくつかの実施形態では、CARは、B7-1/CD80；B7-2/CD86；B7-H1/PD-L1；B7-H2；B7-H3；B7-H4；B7-H6；B7-H7；BTLA/CD272；CD28；CTLA-4；Gi24/VISTA/B7-H5；ICOS/CD278；PD-1；PD-L2/B7-DC；PDCD6）；4-1BB/TNFSF9/CD137；4-1BBリガンド/TNFSF9；BAFF/BlyS/TNFSF13B；BAFFR/TNFRSF13C；CD27/TNFRSF7；CD27リガンド/TNFSF7；CD30/TNFRSF8；CD30リガンド/TN

F S F 8 ; C D 4 0 / T N F R S F 5 ; C D 4 0 / T N F S F 5 ; C D 4 0 リガンド / T N F S F 5 ; D R 3 / T N F R S F 2 5 ; G I T R / T N F R S F 1 8 ; G I T R リガンド / T N F S F 1 8 ; H V E M / T N F R S F 1 4 ; L I G H T / T N F S F 1 4 ; リンホトキシン - アルファ / T N F - ベータ ; O X 4 0 / T N F R S F 4 ; O X 4 0 リガンド / T N F S F 4 ; R E L T / T N F R S F 1 9 L ; T A C I / T N F R S F 1 3 B ; T L 1 A / T N F S F 1 5 ; T N F - アルファ ; T N F R I I / T N F R S F 1 B ) ; 2 B 4 / C D 2 4 4 / S L A M F 4 ; B L A M E / S L A M F 8 ; C D 2 ; C D 2 F - 1 0 / S L A M F 9 ; C D 4 8 / S L A M F 2 ; C D 5 8 / L F A - 3 ; C D 8 4 / S L A M F 5 ; C D 2 2 9 / S L A M F 3 ; C R A C C / S L A M F 7 ; N T B - A / S L A M F 6 ; S L A M / C D 1 5 0 ) ; C D 2 ; C D 7 ; C D 5 3 ; C D 8 2 / K a i - 1 ; C D 9 0 / T h y 1 ; C D 9 6 ; C D 1 6 0 ; C D 2 0 0 ; C D 3 0 0 a / L M I R 1 ; H L A C l a s s I ; H L A - D R ; I k a r o s ; インテグリンアルファ 4 / C D 4 9 d ; インテグリンアルファ 4 ベータ 1 ; インテグリンアルファ 4 ベータ 7 / L P A M - 1 ; L A G - 3 ; T C L 1 A ; T C L 1 B ; C R T A M ; D A P 1 2 ; デクチン - 1 / C L E C 7 A ; D P P I V / C D 2 6 ; E p h B 6 ; T I M - 1 / K I M - 1 / H A V C R ; T I M - 4 ; T S L P ; T S L P R ; リンパ球機能関連抗原 - 1 ( L F A - 1 ) ; N K G 2 C 、 C D 3 ゼータドメイン、免疫受容体チロシンベースの活性化モチーフ ( I T A M ) 、 C D 2 7 、 C D 2 8 、 4 - 1 B B 、 C D 1 3 4 / O X 4 0 、 C D 3 0 、 C D 4 0 、 P D - 1 、 I C O S 、 リンパ球機能関連抗原 - 1 ( L F A - 1 ) 、 C D 2 、 C D 7 、 L I G H T 、 N K G 2 C 、 B 7 - H 3 、 C D 8 3 と特異的に結合するリガンド、またはそれらの機能的断片のうちの一つ以上から選択される少なくとも一つのシグナル伝達ドメインを含む。

10

20

#### 【 0 3 2 4 】

いくつかの実施形態では、C A R は、C D 3 ゼータドメインまたは免疫受容体チロシンベースの活性化モチーフ ( I T A M ) 、またはその機能的バリエーションを含む。いくつかの実施形態では、C A R は、( i ) C D 3 ゼータドメイン、または免疫受容体チロシンベースの活性化モチーフ ( I T A M ) 、またはその機能的バリエーション、及び ( i i ) C D 2 8 ドメイン、または 4 - 1 B B ドメイン、またはその機能的バリエーションを含む。いくつかの実施形態では、C A R は、( i ) C D 3 ゼータドメイン、または免疫受容体チロシンベースの活性化モチーフ ( I T A M ) 、またはその機能的バリエーション、( i i ) C D 2 8 ドメインまたはその機能的バリエーション、及び ( i i i ) 4 - 1 B B ドメイン、または C D 1 3 4 ドメイン、またはその機能的バリエーションを含む。いくつかの実施形態では、C A R は、( i ) C D 3 ゼータドメイン、もしくは免疫受容体チロシンベースの活性化モチーフ ( I T A M ) 、もしくはその機能的バリエーション、( i i ) C D 2 8 ドメイン、もしくは 4 - 1 B B ドメイン、もしくはその機能的バリエーション、及び / または ( i i i ) 4 - 1 B B ドメイン、もしくは C D 1 3 4 ドメイン、もしくはその機能的バリエーションを含む。いくつかの実施形態では、C A R は、( i ) C D 3 ゼータドメイン、または免疫受容体チロシンベースの活性化モチーフ ( I T A M ) 、またはその機能的バリエーション、( i i ) C D 2 8 ドメインまたはその機能的バリエーション、( i i i ) 4 - 1 B B ドメイン、または C D 1 3 4 ドメイン、またはその機能的バリエーション、及び ( i v ) サイトカインまたは共刺激リガンド導入遺伝子を含む。

30

40

#### 【 0 3 2 5 】

所定の実施形態では、細胞内シグナル伝達ドメインは、C D 3 ( 例えば、C D 3 - ゼータ ) 細胞内ドメインに連結された C D 2 8 膜貫通及びシグナル伝達ドメインを含む。いくつかの実施形態では、細胞内シグナル伝達ドメインは、C D 3 ゼータ細胞内ドメインに連結された、キメラ C D 2 8 及び C D 1 3 7 ( 4 - 1 B B 、 T N F R S F 9 ) 共刺激ドメインを含む。

#### 【 0 3 2 6 】

いくつかの実施形態では、C A R は、細胞質部分において、一つ以上の、例えば、二つ以上の、共刺激ドメイン及び活性化ドメイン、例えば、一次活性化ドメインを包含する。例示的な C A R には、C D 3 - ゼータ、C D 2 8 、及び 4 - 1 B B の細胞内成分が含まれ

50

る。

【 0 3 2 7 】

いくつかの実施形態では、細胞内シグナル伝達ドメインには、4 - 1 B Bシグナル伝達ドメイン及びCD3 - ゼータシグナル伝達ドメインの細胞内成分が含まれる。いくつかの実施形態では、細胞内シグナル伝達ドメインには、CD28シグナル伝達ドメイン及びCD3ゼータシグナル伝達ドメインの細胞内成分が含まれる。

【 0 3 2 8 】

いくつかの実施形態では、CARは、抗原（例えば、腫瘍抗原）に結合する細胞外抗原結合ドメイン（例えば、抗体または抗体断片、例えば、scFv）、スペーサー（例えば、本明細書に記載される任意のものなどのヒンジドメインを含有する）、膜貫通ドメイン（例えば、本明細書に記載される任意のもの）、及び細胞内シグナル伝達ドメイン（例えば、任意の細胞内シグナル伝達ドメイン、例えば、本明細書に記載される一次シグナル伝達ドメインまたは共刺激シグナル伝達ドメイン）を含む。いくつかの実施形態では、細胞内シグナル伝達ドメインは、一次細胞質シグナル伝達ドメインであるか、またはそれを含む。いくつかの実施形態では、細胞内シグナル伝達ドメインには、共刺激分子の細胞内シグナル伝達ドメイン（例えば、共刺激ドメイン）が追加的に含まれる。CARの例示的な成分の例は、表6に記載されている。提供される態様では、CARにおける各成分の配列には、表6に列挙される任意の組み合わせが含まれ得る。

10

【 0 3 2 9 】

（表6）CAR成分及び例示的な配列

20

30

40

50

成分	配列	配列番号
細胞外結合ドメイン		
抗 CD19 scFv (FMC63)	DIQMTQTTSSLSASLGDRVTISCRASQDISKYL NWXQQKPDGTVKLLIYHTSRLHSGVPSRFSGS GSGTDYSLTISNLEQEDIATYFCQQGNTLPYTF GGGKLEITGSTSGSGKPGSGEGSTKGEVKLQ ESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWI RQPPRKLEWLGVIWGSETTYNSALKSRLTI IKDNSKSQVFLKMNSLQTDDTAIYYCAKHYY YGGSYAMDYWGQGTSVTVSS	419
抗 CD19 scFv (FMC63)	DIQMTQTTSSLSASLGDRVTISCRASQDISKYL NWXQQKPDGTVKLLIYHTSRLHSGVPSRFSGS GSGTDYSLTISNLEQEDIATYFCQQGNTLPYTF GGGKLEITGGGGSGGGGSGGGGSEVKLQES GPGLVAPSQSLSVTCTVSGVSLPDYGVSIRQ PPRKLEWLGVIWGSETTYNSALKSRLTIHK DNSKSQVFLKMNSLQTDDTAIYYCAKHYYYG GSYAMDYWGQGTSVTVSS	420
スペーサー (例えば、ヒンジ)		
IgG4 ヒンジ	ESKYGPPCPPCP	421
CD8 ヒンジ	TTTTAPRPPTPAPTIASQPLSLRPE	422
CD28	IEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPLF PGPSKP	423
膜貫通		
CD8	ACRPAAGGAVHTRGLDFACDIYWAPLAGTC GVLLLSLVITLYC	424
CD28	FWVLVVVGGVLACYSLLVTVAFIIFWV	425
CD28	FWVLVVVGGVLACYSLLVTVAFIIFWV	426
共刺激ドメイン		
CD28	RSKRSRLLHSDYMNMTPRRPGPTRKHYPYA PPRDFAAAYS	427
4-1BB	KRGRKLLYIFKQPFMRPVQTTQEEDGCSCRF PEEEGGCEL	428

10

20

30

40

50

一次シグナル伝達ドメイン		
CD3 ゼータ	RVKFSRSADAPAYQQGQNQLYNELNLGRREE YDVLDKRRGRDPENGGKPRRKNPQEGLYNE LQKDKMAEAYSEIGMKGERRRGKGGHDGLYQ GLSTATKDTYDALHMQUALPPR	429
CD3 ゼータ	RVKFSRSADAPAYKQGQNQLYNELNLGRREE YDVLDKRRGRDPENGGKPRRKNPQEGLYNE LQKDKMAEAYSEIGMKGERRRGKGGHDGLYQ GLSTATKDTYDALHMQUALPPR	430

10

## 【0330】

いくつかの実施形態では、CARは、1つ以上のスペーサーをさらに含み、例えば、スペーサーは、抗原結合ドメインと膜貫通ドメインとの間の第1のスペーサーである。いくつかの実施形態では、第1のスペーサーには、免疫グロブリン定常領域もしくはバリエーションまたはその修飾されたバージョンの少なくとも一部が含まれる。いくつかの実施形態では、スペーサーは、膜貫通ドメインとシグナル伝達ドメインとの間の第2のスペーサーである。いくつかの実施形態では、第2のスペーサーは、オリゴペプチドであり、例えば、オリゴペプチドは、グリシン-セリンダブレットを含む。

20

## 【0331】

本明細書に記載のCARに加えて、様々なキメラ抗原受容体及びそれをコードするヌクレオチド配列が知られており、本明細書に記載される *in vivo* 及び *in vitro* での標的細胞のフソソーム送達及び再プログラム化に好適であろう。例えば、WO2013040557、WO2012079000、WO2016030414、Smith T, et al., Nature Nanotechnology, 2017. (DOI: 10.1038/NNANO.2017.57) (その開示は、それらの全体が参照により本明細書に組み込まれる)を参照されたい。

## 【0332】

いくつかの実施形態では、CARまたはCARをコードする核酸(例えば、DNA、gDNA、cDNA、RNA、pre-mRNA、mRNA、miRNA、siRNAなど)を含む標的化脂質粒子が標的細胞に送達される。いくつかの実施形態では、標的細胞は、1つ以上のFc受容体を発現し、かつ1つ以上のエフェクター機能を媒介するエフェクター細胞、例えば、免疫系の細胞である。いくつかの実施形態では、標的細胞には、単球、マクロファージ、好中球、樹状細胞、好酸球、肥満細胞、血小板、大型顆粒リンパ球、ランゲルハンス細胞、ナチュラルキラー(NK)細胞、Tリンパ球(例えば、T細胞)、ガンマデルタT細胞、Bリンパ球(例えば、B細胞)のうちの一つ以上が含まれ得るがこれらに限定されない場合があり、ヒト、マウス、ラット、ウサギ、及びサルを含むがこれらに限定されない任意の生物からのものであり得る。

30

40

## 【0333】

## E. 標的化脂質粒子を生成する方法

本明細書では、脂質二重層と、脂質二重層によって囲まれた内腔と、標的化エンベロプタンパク質と、フソゲンとを含む、標的化脂質粒子であって、標的化エンベロプタンパク質及びフソゲンが、脂質二重層に埋め込まれている、標的化脂質粒子が提供される。いくつかの実施形態では、標的化脂質粒子は、ウイルス粒子、ウイルス様粒子、ナノ粒子、小胞、エクソソーム、デンドリマー、レンチウイルス、ウイルスベクター、除核細胞、微小胞、膜小胞、細胞外膜小胞、細胞膜小胞、巨大細胞膜小胞、アポトーシス小体、ミト粒子、ピレノサイト、リソソーム、別の膜包囲小胞、またはレンチウイルスベクター、ウイルスベースの粒子、ウイルス様粒子(VLP)または細胞由来粒子であり得る。

50

## 【0334】

## 1. ウイルス様粒子

本明細書では、ウイルス粒子またはウイルス様粒子などのウイルスに由来する標的化脂質粒子（レトロウイルスまたはレンチウイルスに由来するものを含む）が提供される。いくつかの実施形態では、両親媒性脂質の標的化脂質粒子の二重層は、ウイルスエンベロープであるか、またはそれを含む。いくつかの実施形態では、両親媒性脂質の標的化脂質粒子の二重層は、プロデューサー細胞に由来する脂質であるか、またはそれを含む。いくつかの実施形態では、ウイルスエンベロープは、フソゲン、例えば、ウイルスに対して内因性であるフソゲンまたは偽型化フソゲンを含み得る。いくつかの実施形態では、標的化脂質粒子の内腔またはキャビティは、ウイルス核酸、例えば、レトロウイルス核酸、例えば、レンチウイルス核酸を含む。いくつかの実施形態では、ウイルス核酸は、ウイルスゲノムであり得る。いくつかの実施形態では、標的化脂質粒子は、1つ以上のウイルス非構造タンパク質を、例えば、そのキャビティまたは内腔にさらに含む。いくつかの実施形態では、標的化脂質粒子は、ウイルス様粒子（VLP）であるか、またはそれを含む。いくつかの実施形態では、VLPは、エンベロープを含まない。いくつかの実施形態では、VLPは、エンベロープを含む。

10

## 【0335】

いくつかの実施形態では、ウイルス粒子またはウイルス様粒子、例えば、レトロウイルスまたはレトロウイルス様粒子は、gagポリタンパク質、ポリメラーゼ（例えば、pol）、インテグラーゼ（例えば、非機能性または非機能性バリエーション）、プロテアーゼ、及びフソゲンのうちの1つ以上を含む。いくつかの実施形態では、標的化脂質粒子は、revをさらに含む。いくつかの実施形態では、前述のタンパク質の1つ以上は、レトロウイルスゲノムにコードされ、いくつかの実施形態では、前述のタンパク質の1つ以上は、例えば、ヘルパー細胞、ヘルパーウイルス、またはヘルパープラスミドによってトランスで提供される。いくつかの実施形態では、標的化脂質粒子核酸（例えば、レトロウイルス核酸）は、以下の核酸配列：5'LTR（例えば、U5を含み、機能性U3ドメインを欠く）、プサイパッケージング要素（プサイ）、ペイロード遺伝子に機能可能に連結されたセントラルポリプリン配列（cPPT）プロモーター、ペイロード遺伝子（任意にオープンリーディングフレームの前のイントロンを含む）、ポリAテール配列、WPRE、及び3'LTR（例えば、U5を含み、機能性U3を欠く）のうちの1つ以上を含む。いくつかの実施形態では、標的化脂質粒子核酸は、1つ以上のインシュレーター要素をさらに含む。いくつかの実施形態では、認識部位は、ポリAテール配列とWPREとの間に位置する。

20

30

## 【0336】

いくつかの実施形態では、標的化脂質粒子は、カプシドへ自己構築するウイルスタンパク質によって形成される超分子複合体を含む。いくつかの実施形態では、標的化脂質粒子は、ウイルスカプシドに由来するウイルス粒子またはウイルス様粒子である。いくつかの実施形態では、標的化脂質粒子は、ウイルスヌクレオカプシドに由来するウイルス粒子またはウイルス様粒子である。いくつかの実施形態では、標的化脂質粒子は、核酸をパッケージングする特性を保持するヌクレオカプシド由来を含む。いくつかの実施形態では、ウイルス粒子またはウイルス様粒子は、ウイルス構造糖タンパク質のみを含む。いくつかの実施形態では、標的化脂質粒子は、ウイルスゲノムを含有しない。

40

## 【0337】

いくつかの実施形態では、標的化脂質粒子は、発現プロセス中に宿主細胞からの核酸をパッケージングする。いくつかの実施形態では、核酸は、ウイルス複製に関するいかなる遺伝子もコードしない。特定の実施形態では、標的化脂質粒子は、例えば、複製欠損性であるウイルス様粒子、例えば、レンチウイルス様粒子などのレトロウイルス様粒子である。

## 【0338】

いくつかの場合では、標的化脂質粒子は、野生型感染性ウイルスから形態的に区別不可

50



能なウイルス粒子である。いくつかの実施形態では、ウイルス粒子は、ウイルスプロテオーム全体を抗原として提示する。いくつかの実施形態では、ウイルス粒子は、プロテオームの一部のみを抗原として提示する。

【0339】

いくつかの実施形態では、ウイルス粒子またはウイルス様粒子は、パラミクソウイルス科内のウイルスからのタンパク質（例えば、エンペローブタンパク質）を利用して生成される。いくつかの実施形態では、パラミクソウイルス科は、ヘニパウイルス属内のメンバーを含む。いくつかの実施形態では、ヘニパウイルスは、ヘンドラ（HeV）またはニパ（NiV）ウイルスであるか、またはそれを含む。特定の実施形態では、ウイルス粒子またはウイルス様粒子は、セクションI.A.及び1.B.に記載される標的化エンペローブタンパク質及びフソゲンを組み込む。

10

【0340】

いくつかの実施形態では、ウイルス粒子またはウイルス様粒子は、細菌、哺乳動物細胞株、昆虫細胞株、酵母及び植物細胞を含む複数の細胞培養システムにおいて生成され得る。

【0341】

いくつかの実施形態では、ウイルス粒子またはウイルス様粒子の構築は、ウイルスゲノム内の独自のカプシド形成配列（例えば、ステムループ構造を有するUTR）へのコアタンパク質の結合によって開始する。いくつかの実施形態では、コアとカプシド形成配列との相互作用は、オリゴマー化を容易化する。

20

【0342】

いくつかの実施形態では、標的化脂質粒子は、配列からウイルスRNAを除去または排除する結果であり得るウイルスRNAがないまたはそれを欠く配列を含むウイルス様粒子である。いくつかの実施形態では、これは、gag上の内因性パッケージングシグナル結合部位を使用することによって達成され得る。いくつかの実施形態では、内因性パッケージングシグナル結合部位は、pol上にある。いくつかの実施形態では、送達されるRNAは、同種パッケージングシグナルを含有する。いくつかの実施形態では、送達されるRNA上に位置する異種結合ドメイン（gagに対して異種である）、及びgagまたはpol上に位置する同種結合部位は、送達されるRNAのパッケージングを確保するために使用され得る。いくつかの実施形態では、異種配列は、非ウイルス性であり得、またはそれは、ウイルス性であり得、その場合、それは異なるウイルスに由来し得る。いくつかの実施形態では、ベクター粒子は、治療RNAを送達するため使用され得、その場合、機能性インテグラーゼ及び/または逆転写酵素は必要とされない。いくつかの実施形態では、ベクター粒子はまた、対象となる治療遺伝子を送達するため使用され得、その場合、polが典型的に含まれる。

30

【0343】

a. トランスファーベクター

いくつかの実施形態では、レトロウイルス核酸は、5'プロモーター（例えば、パッケージングされたRNA全体の発現をコントロールするためのもの）、5'LTR（例えば、R（ポリアデニル化テールシグナルを含む）及び/またはプライマー活性化シグナルを含むU5）、プライマー結合部位、プサイパッケージングシグナル、核外搬出のためのRRE要素、導入遺伝子発現をコントロールするための導入遺伝子の直接的上流のプロモーター、導入遺伝子（または他の外因性物質要素）、ポリプリントラクト、及び3'LTR（例えば、変異したU3、R、及びU5を含む）のうちの一つ以上（例えば、すべて）を含む。いくつかの実施形態では、レトロウイルス核酸は、cPPT、WPRE、及び/またはインシュレーター要素のうちの一つ以上をさらに含む。

40

【0344】

レトロウイルスは、典型的には、そのゲノムRNAの線状二本鎖DNAコピーへの逆転写によって複製し、その後、そのゲノムDNAを宿主ゲノムに共有結合的に組み込む。特定の実施形態において使用するために好適な例示的なレトロウイルスには、モロニーマウ

50

ス白血病ウイルス (M - M u L V)、モロニー Maus 肉腫ウイルス (M o M S V)、ハーベイ Maus 肉腫ウイルス (H a M u S V)、マウス乳腫瘍ウイルス (M u M T V)、テナガザル白血病ウイルス (G a L V)、ネコ白血病ウイルス (F L V)、スピューマウイルス、フレンドマウス白血病ウイルス、マウス幹細胞ウイルス (M S C V) 及びラウス肉腫ウイルス (R S V)、及びレンチウイルスが含まれるがこれらに限定されない。

【0345】

いくつかの実施形態ではレトロウイルスは、ガンマレトロウイルスである。いくつかの実施形態ではレトロウイルスは、イプシロンレトロウイルスである。いくつかの実施形態ではレトロウイルスは、アルファレトロウイルスである。いくつかの実施形態ではレトロウイルスは、ベータレトロウイルスである。いくつかの実施形態ではレトロウイルスは、デルタレトロウイルスである。いくつかの実施形態ではレトロウイルスは、レンチウイルスである。いくつかの実施形態では、レトロウイルスは、スプマレトロウイルスである。いくつかの実施形態ではレトロウイルスは、内因性レトロウイルスである。

10

【0346】

例示的なレンチウイルスには、H I V (ヒト免疫不全ウイルス; H I V 1 型、及び H I V 2 型を含む)、ピスナ・マエディウイルス (V M V) ウイルス、ヤギ関節炎 - 脳炎ウイルス (C A E V)、ウマ伝染性貧血ウイルス (E I A V)、ネコ免疫不全ウイルス (F I V)、ウシ免疫不全ウイルス (B I V)、及びサル免疫不全ウイルス (S I V) が含まれるがこれらに限定されない。いくつかの実施形態では、H I V ベースのベクター骨格 (すなわち、H I V シス作用性配列要素) が使用される。

20

【0347】

いくつかの実施形態では、本明細書におけるベクターは、別の核酸分子を移行または輸送することが可能な核酸分子である。移行される核酸は通常、ベクター核酸分子に連結され、例えば、挿入される。ベクターには、細胞において自己複製を誘導する配列が含まれ得、または宿主細胞 D N A への組み込みを可能とするのに十分な配列が含まれ得る。有用なベクターには、例えば、プラスミド (例えば、D N A プラスミドまたは R N A プラスミド)、トランスポゾン、コスミド、細菌人工染色体、及びウイルスベクターが含まれる。有用なウイルスベクターには、例えば、複製欠損性レトロウイルス及びレンチウイルスが含まれる。

【0348】

いくつかの実施形態では、ウイルスベクターは、典型的には核酸分子の移行または細胞のゲノムもしくは核酸移行を媒介するウイルス粒子への組み込みを容易化するウイルス由来核酸要素を含む核酸分子 (例えば、トランスファープラスミド) を含む。ウイルス粒子は、典型的には、核酸 (複数可) に加えて様々なウイルス成分及び時に宿主細胞成分も含み得る。いくつかの実施形態では、ウイルスベクターは、例えば、核酸を細胞に、または移行した核酸 (例えば、むき出しの D N A) に移行させることが可能なウイルスまたはウイルス粒子を含む。いくつかの実施形態では、ウイルスベクター及びトランスファープラスミドは、主にウイルスに由来する構造的及び/または機能的遺伝子要素を含む。レトロウイルスベクターは、主にレトロウイルスに由来する構造的及び/または機能的遺伝子要素、またはその一部を含有するウイルスベクターまたはプラスミドを含み得る。レンチウイルスベクターは、主にレンチウイルスに由来する、L T R を含む、構造的及び/または機能的遺伝子要素、またはその一部を含有するウイルスベクターまたはプラスミドを含み得る。

30

40

【0349】

実施形態では、レンチウイルスベクター (例えば、レンチウイルス発現ベクター) は、レンチウイルストランスファープラスミド (例えば、むき出しの D N A として) または感染性レンチウイルス粒子を含み得る。クローニング部位、プロモーター、制御要素、異種核酸などの要素に関して、これらの要素の配列は、R N A 形態でレンチウイルス粒子に存在し得、D N A 形態で D N A プラスミドに存在し得ることが理解されるべきである。

【0350】

50

いくつかの実施形態では、本明細書に記載のベクターにおいて、複製に寄与する、または必須である1つ以上のタンパク質コード領域の少なくとも一部は、対応する野生型ウイルスと比較して存在しない場合がある。いくつかの実施形態では、ウイルスベクターは、複製欠損性である。いくつかの実施形態では、ベクターは、標的非分裂宿主細胞の形質導入、及び/またはそのゲノムを宿主ゲノムに組み込むことが可能である。

#### 【0351】

いくつかの実施形態では、野生型レトロウイルスゲノムの構造は、しばしば、5'長い末端反復(LTR)及び3'LTRを含み、その間またはその中に、ゲノムがパッケージングされることを可能にするためのパッケージングシグナル、プライマー結合部位、宿主細胞ゲノムへの組み込みを可能とするための組み込み部位ならびにウイルス粒子の構築を促進するパッケージング成分をコードするgag、pol及びenv遺伝子が配置されている。より複雑なレトロウイルスは、追加の特徴、例えば、HIVにおけるrev及びRRE配列を有し、これは、感染した標的細胞の核から細胞質への組み込まれたプロウイルスのRNA転写産物の効率的な搬出を可能とする。プロウイルスでは、ウイルス遺伝子は、長い末端反復(LTR)と呼ばれる領域によって両端で挟まれている。いくつかの実施形態では、LTRは、プロウイルスの組み込み及び転写に關与する。いくつかの実施形態では、LTRは、エンハンサー-プロモーター配列として機能し、ウイルス遺伝子の発現をコントロールし得る。いくつかの実施形態では、レトロウイルスRNAのカプシド形成は、ウイルスゲノムの5'末端に位置するプサイ配列により生じる。

#### 【0352】

いくつかの実施形態では、LTRは、U3、R及びU5と呼ばれる3つの要素に分けられ得る同様の配列である。U3は、RNAの3'末端に特有の配列に由来する。Rは、RNAの両端で繰り返される配列に由来し、U5は、RNAの5'末端に特有の配列に由来する。3つの要素のサイズは、異なるレトロウイルス間で大幅に変化し得る。

#### 【0353】

いくつかの実施形態では、ウイルスゲノムについては、転写開始の部位は、典型的には、一方のLTRにおけるU3とRとの間の境界であり、ポリ(A)付加(終結)の部位は、他方のLTRにおけるRとU5との間の境界である。U3は、細胞性、及びいくつかの場合では、ウイルス転写活性化因子タンパク質に対して反応性のプロモーター及び複数のエンハンサー配列を含む、プロウイルスの転写コントロール要素のほとんどを含有する。いくつかの実施形態では、レトロウイルスは、遺伝子発現の制御に關与するタンパク質をコードする次の遺伝子:tat、rev、tax及びrexのうちの任意の1つ以上を含む。

#### 【0354】

いくつかの実施形態では、構造遺伝子gag、pol及びenv、gagは、ウイルスの内部構造タンパク質をコードする。いくつかの実施形態では、Gagタンパク質は、成熟タンパク質MA(マトリックス)、CA(カプシド)及びNC(ヌクレオカプシド)にタンパク質分解的に処理される。いくつかの実施形態では、pol遺伝子は、逆転写酵素(RT)をコードし、これは、ゲノムの複製を媒介する、DNAポリメラーゼ、関連するRNase H及びインテグラーゼ(IN)を含有する。いくつかの実施形態では、env遺伝子は、細胞受容体タンパク質と特異的に相互作用する複合体を形成する、ビリオンの表面(SU)糖タンパク質及び膜貫通(TM)タンパク質をコードする。いくつかの実施形態では、相互作用は、ウイルス膜と細胞膜との融合によって感染症を促進する。

#### 【0355】

いくつかの実施形態では、複製欠損性レトロウイルスベクターゲノムgag、pol及びenvは、存在しない場合があり、または機能性ではない場合がある。いくつかの実施形態では、RNAの両末端におけるR領域は、典型的には繰り返し配列である。いくつかの実施形態では、U5及びU3は、RNAゲノムの5'及び3'末端における独自の配列をそれぞれ表す。

#### 【0356】

10

20

30

40

50

いくつかの実施形態では、レトロウイルスはまた、gag、pol及びenv以外のタンパク質をコードする追加の遺伝子を含み得る。追加の遺伝子の例には、(HIVにおける)、vif、vpr、vpx、vpu、tat、rev及びnefのうちの一つ以上が含まれる。EIAVは、(とりわけ)追加の遺伝子S2を有する。いくつかの実施形態では、追加の遺伝子によってコードされるタンパク質は、様々な機能を果たし、その一部は、細胞性タンパク質によって提供される機能の重複であり得る。EIAVにおいて、例えば、tatは、ウイルスLTRの転写活性化因子として作用する(Derse and Newbold 1993 Virology 194:530-6; Maury et al. 1994 Virology 200:632-42)。それは、TARと称される安定なステムループRNA二次構造に結合する。Revは、rev-反応要素(RRE)を介してウイルス遺伝子の発現を制御し、調整する(Martarano et al. 1994 J. Virol. 68:3102-11)。

10

## 【0357】

いくつかの実施形態では、プロテアーゼ、逆転写酵素及びインテグラーゼに加えて、非霊長類レンチウイルスは、dUTPaseをコードする第4のpol遺伝子産物を含み得る。いくつかの実施形態では、これは、これらのレンチウイルスが、所定の非分裂細胞または徐々に分裂する細胞タイプに感染する能力において役割を果たす。

## 【0358】

実施形態では、組換えレンチウイルスベクター(RLV)は、RNAゲノムを、パッケージング成分の存在下で、標的細胞に感染することが可能なウイルス粒子にパッケージングすることが可能となるように十分なレトロウイルス遺伝子情報を有するベクターである。いくつかの実施形態では、標的細胞の感染は、逆転写及び標的細胞ゲノムへの組み込みを含み得る。いくつかの実施形態では、RLVは、典型的には、ベクターによって標的細胞に送達される非ウイルスコーディング配列を保有する。いくつかの実施形態では、RLVは、標的細胞内で感染性レトロウイルス粒子を生成するための独立した複製が不可能である。いくつかの実施形態では、RLVは、機能性gag-pol及び/またはenv遺伝子及び/または複製に関与する他の遺伝子を欠く。いくつかの実施形態では、ベクターは、例えば、PCT特許出願WO99/15683(その全体が参照により本明細書に組み込まれる)に記載されているように、スプリット-イントロンベクターとして構成され得る。

20

30

## 【0359】

いくつかの実施形態では、レンチウイルスベクターは、最小ウイルスゲノムを含み、例えば、ウイルスベクターは、例えば、WO98/17815(その全体が参照により本明細書に組み込まれる)に記載されているように、対象となるヌクレオチド配列を標的宿主細胞に感染させ、形質導入し、送達するための必要とされる機能性を提供するために、非必須要素を除去し、必須要素を保持するように操作されている。

## 【0360】

いくつかの実施形態では、最小レンチウイルスゲノムは、例えば、(5')R-U5-1つ以上の第1のヌクレオチド配列-U3-R(3')を含み得る。いくつかの実施形態では、源細胞内でレンチウイルスゲノムを生成するために使用されるプラスミドベクターにはまた、源細胞におけるゲノムの転写を誘導するためのレンチウイルスゲノムに機能可能に連結された転写制御コントロール配列が含まれ得る。いくつかの実施形態では、制御配列は、転写されたレトロウイルス配列、例えば、5'U3領域に関連する天然配列を含み得、またはそれらは、異種プロモーター、例えば、別のウイルスプロモーター、例えばCMVプロモーターを含み得る。いくつかの実施形態では、レンチウイルスゲノムは、効率的なウイルス生成を促進するための追加の配列を含む。いくつかの実施形態では、HIVの場合、rev及びRRE配列が含まれ得る。いくつかの実施形態では、代替的にまたは組み合わせて、コドン最適化が使用され得、例えば、外因性物質をコードする遺伝子は、例えば、WO01/79518(その全体が参照により本明細書に組み込まれる)に記載されているように、コドン最適化され得る。いくつかの実施形態では、rev/RRE

40

50

システムと類似するまたは同じ機能を発揮する代替的配列も使用され得る。いくつかの実施形態では、rev/RREシステムの機能性アナログは、Mason Pfizerサルウイルスに見られる。いくつかの実施形態では、これは、CTEとして知られており、感染細胞における因子と相互作用すると考えられるゲノムにおけるRRE型配列を含む。細胞因子は、revアナログと考えられ得る。いくつかの実施形態では、CTEは、rev/RREシステムの代替手段として使用され得る。いくつかの実施形態では、HTLV-IのRexタンパク質は、HIV-IのRevタンパク質に機能的に取って代わり得る。Rev及びRexは、IRE-BPに対して類似する効果を有する。

#### 【0361】

いくつかの実施形態では、レトロウイルス核酸（例えば、レンチウイルス核酸、例えば、霊長類または非霊長類レンチウイルス核酸）は、（1）欠失したgag遺伝子であって、gag遺伝子における欠失が、gagコード配列の約ヌクレオチド350または354の下流の1つ以上のヌクレオチドを除去する、欠失したgag遺伝子を含む、（2）レトロウイルス核酸に存在しない1つ以上のアクセサリ遺伝子を有する、（3）tat遺伝子を欠くが、5'LTRの末端とgagのATGとの間にリーダー配列を含む、及び（4）（1）、（2）及び（3）の組み合わせである。実施形態では、レンチウイルスベクターは、特徴（1）及び（2）及び（3）のすべてを含む。この戦略は、W099/32646（その全体が参照により本明細書に組み込まれる）においてより詳細に記載されている。

#### 【0362】

いくつかの実施形態では、霊長類レンチウイルス最小システムは、ベクター生成のためにまたは分裂細胞及び非分裂細胞の形質導入のためにHIV/SIVの追加の遺伝子vif、vpr、vpx、vpu、tat、rev及びnefのいずれも必要としない。いくつかの実施形態では、EIAV最小ベクターシステムは、ベクター生成のためにまたは分裂細胞及び非分裂細胞の形質導入のためにS2を必要としない。

#### 【0363】

いくつかの実施形態では、追加の遺伝子の欠失により、レンチウイルス（例えば、HIV）感染症における疾患に関連する遺伝子を伴わずにベクターを生成することが可能となり得る。いくつかの実施形態では、tatは、疾患に関連する。いくつかの実施形態では、追加の遺伝子の欠失により、ベクターは、より多くの異種DNAをパッケージングすることが可能となる。いくつかの実施形態では、S2などの機能が不明である遺伝子は、省略され、よって、不要な効果を引き起こすリスクを減少させ得る。最小レンチウイルスベクターの例は、W099/32646及びW098/17815に開示されている。

#### 【0364】

いくつかの実施形態では、レトロウイルス核酸は、少なくともtat及びS2（それがEIAVベクターシステムである場合）を欠き、場合によってはvif、vpr、vpx、vpu及びnefも欠く。いくつかの実施形態では、レトロウイルス核酸はまた、rev、RRE、またはその両方を欠く。

#### 【0365】

いくつかの実施形態ではレトロウイルス核酸は、vpxを含む。Vpxポリペプチドは、細胞質における遊離dNTPを分解するSAMHD1制限因子に結合し、その分解を誘導する。いくつかの実施形態では、細胞質における遊離dNTPの濃度は、VpxがSAMHD1を分解し、逆転写活性が増加するとともに増加し、よって、レトロウイルスゲノムの逆転写及び標的細胞ゲノムへの組み込みを容易化する。

#### 【0366】

いくつかの実施形態では、異なる細胞は、特定のコドンのそれらの使用頻度が異なる。いくつかの実施形態では、このコドンバイアスは、細胞タイプにおける特定のtRNAの相対的存在量におけるバイアスに対応する。いくつかの実施形態では、対応するtRNAの相対的存在量に一致するように調整されるように配列におけるコドンを変更することにより、発現を増加させることが可能である。いくつかの実施形態では、対応するtRNA

10

20

30

40

50

が特定の細胞タイプにおいて稀であることが知られているコドンを選択することによって発現を減少させることが可能である。いくつかの実施形態では、追加の程度の翻訳コントロールが可能である。コドン最適化の追加的説明は、例えば、W O 9 9 / 4 1 3 9 7 (その全体が参照により本明細書に組み込まれる)に記載されている。

【0367】

いくつかの実施形態では、H I V 及び他のレンチウイルスを含むウイルスは、多数の稀なコドンを使用し、これらを一般的に使用される哺乳動物コドンに対応するように変化させることによって、哺乳動物プロデューサー細胞におけるパッケージング成分の増加した発現が達成され得る。

【0368】

いくつかの実施形態では、コドン最適化は、多数の他の利点を有する。いくつかの実施形態では、それらの配列における変化により、パッケージング成分をコードするヌクレオチド配列は、RNA不安定性配列 (INS) が減少、またはそれから排除され得る。同時に、パッケージング成分のためのアミノ酸配列コード配列は、その配列によってコードされるウイルス成分が、パッケージング成分の機能が損なわれないように同じ、または十分に類似したままとなるように保持される。いくつかの実施形態では、コドン最適化はまた、搬出のための Rev / RRE 要件を克服し、最適化された配列を Rev 非依存とする。いくつかの実施形態では、コドン最適化はまた、ベクターシステム内の異なるコンストラクト間 (例えば、gag-pol 及び env オープンリーディングフレームにおける重複の領域間) の相同組換えを減少させる。いくつかの実施形態では、コドン最適化は、ウイルス力価の増加及び / または改善した安全性をもたらす。

【0369】

いくつかの実施形態では、INS に関するコドンのみが、コドン最適化される。他の実施形態では、配列は、gag-pol のフレームシフト部位を包含する配列を除き、それらの全体がコドン最適化される。

【0370】

gag-pol 遺伝子は、gag-pol タンパク質をコードする2つの重複するリーディングフレームを含む。両方のタンパク質の発現は、翻訳中のフレームシフトに依存する。このフレームシフトは、翻訳中のリボソーム「滑り (slippage)」の結果として生じる。この滑りは、少なくとも部分的にリボソーム停止 RNA 二次構造によって引き起こされると考えられる。そのような二次構造は、gag-pol 遺伝子におけるフレームシフト部位の下流に存在する。H I V の場合、重複の領域は、gag の開始の 1 2 2 2 下流のヌクレオチド (ヌクレオチド 1 は、gag ATG の A である) から gag の最後 (nt 1 5 0 3) まで伸長する。結果的に、フレームシフト部位及び2つのリーディングフレームの重複領域に及ぶ 2 8 1 b p 断片は、好ましくはコドン最適化されない。いくつかの実施形態では、この断片の保持は、gag-pol タンパク質の発現をより効率的なものとする。E I A V の場合、重複の開始は、nt 1 2 6 2 (ヌクレオチド 1 は gag ATG の A である) である。重複の最後は、nt 1 4 6 1 である。フレームシフト部位及び gag-pol 重複が保存されることを確保するために、野生型配列は、nt 1 1 5 6 から 1 4 6 5 まで保持され得る。

【0371】

いくつかの実施形態では、例えば、好都合な制限部位を提供するために、最適なコドン使用頻度からの逸脱がなされ得、保存的アミノ酸変化が gag-pol タンパク質に導入され得る。

【0372】

いくつかの実施形態では、コドン最適化は、哺乳動物システムにおける不十分なコドン使用頻度を伴うコドンに基づく。3番目ならびに時に2番目及び3番目の塩基が変化され得る。

【0373】

いくつかの実施形態では、遺伝子コードの縮重特質により、多数の gag-pol 配列

10

20

30

40

50

が当業者によって達成され得ることが理解される。また、コドン最適化 gag - pol 配列を生成するための足掛かりとして使用され得る記載される多くのレトロウイルスバリエーションが存在する。レンチウイルスゲノムは、非常に可変性である。例えば、依然として機能性である HIV - I の多くの準種が存在する。これは、EIAV についても当てはまる。これらのバリエーションは、形質導入プロセスの特定の部分を向上させるために使用され得る。HIV - I バリエーションの例は、Los Alamos National Laboratory によって維持されている HIV データベースで見ることができる。EIAV クローンの詳細は、National Institutes of Health によって維持されている NCBI データベースで見ることができる。

【0374】

いくつかの実施形態では、コドン最適化 gag - pol 配列のためのストラテジーが、任意のレトロウイルス、例えば、EIAV、FIV、BIV、CAEV、VMR、SIV、HIV - I 及び HIV - 2 に関して使用され得る。また、この方法は、HTLV - I、HTLV - 2、HFV、HSRV 及びヒト内因性レトロウイルス (HERV)、MLV 及び他のレトロウイルスからの遺伝子の発現を増加させるために使用され得る。

【0375】

実施形態では、レトロウイルスベクターは、env 配列を依然として保持するベクターにおける gag の 255 ~ 360 ヌクレオチド、またはスプライドナー変異、gag 及び env 欠失の特定の組み合わせにおける gag の約 40 ヌクレオチド含むパッケージングシグナルを含む。いくつかの実施形態では、レトロウイルスベクターは、1 つ以上の欠失を含む gag 配列を含み、例えば、gag 配列は、N 末端から誘導可能な約 360 ヌクレオチドを含む。

【0376】

いくつかの実施形態では、レトロウイルスベクター、ヘルパー細胞、ヘルパーウイルス、またはヘルパープラスミドは、レトロウイルス構造及びアクセサリタンパク質、例えば gag、pol、env、tat、rev、vif、vpr、vpu、vpx、または nef タンパク質または他のレトロウイルスタンパク質を含み得る。いくつかの実施形態ではレトロウイルスタンパク質は、同じレトロウイルスに由来する。いくつかの実施形態では、レトロウイルスタンパク質は、複数のレトロウイルス、例えば、2、3、4、またはそれ以上のレトロウイルスに由来する。

【0377】

いくつかの実施形態では、gag 及び pol をコードする配列は通常、ネイティブレンチウイルスにおいて Gag - Pol 前駆体として組織化される。gag 配列は、p55 と呼ばれる、55 kD の Gag 前駆体タンパク質をコードする。p55 は、MA (マトリックス [p17])、CA (カプシド [p24])、NC (ヌクレオカプシド [p9])、及び p6 と指定された 4 つのより小さなタンパク質への成熟のプロセス中にウイルスにコードされるプロテアーゼ (pol 遺伝子の生成物) によって切断される。pol 前駆体タンパク質は、ウイルスにコードされるプロテアーゼによって Gag から切断され、さらに消化されてプロテアーゼ (p10)、RT (p50)、RNase H (p15)、及びインテグラーゼ (p31) 活性を分離する。

【0378】

いくつかの実施形態では、レンチウイルスベクターは、組み込み欠損性である。いくつかの実施形態では、pol は、インテグラーゼ遺伝子における変異によりコードすることなどによってインテグラーゼ欠損性となる。例えば、pol コード配列は、例えば、触媒活性に關与するアミノ酸の 1 つ以上の変異、すなわち、アスパラギン 64、アスパラギン酸 116 及び / またはグルタミン酸 152 のうちの 1 つ以上の変異によって、インテグラーゼにおける不活性化変異を含有し得る。いくつかの実施形態では、インテグラーゼ変異は、D64V 変異である。いくつかの実施形態では、インテグラーゼにおける変異は、ウイルス RNA のレンチウイルスへのパッケージングを可能とする。いくつかの実施形態では、インテグラーゼにおける変異は、ウイルスタンパク質のレンチウイルスへのパッケー

10

20

30

40

50

ジングを可能とする。いくつかの実施形態では、インテグラーゼにおける変異は、挿入変異生成の可能性を減少させる。いくつかの実施形態では、インテグラーゼにおける変異は、複製能力のあるリコンビナント (RCR) を生成する可能性を低下させる (Wanisch et al., 2009, Mol Ther., 1798): 1316 - 1332)。いくつかの実施形態では、ネイティブ Gag - Pol 配列がヘルパーベクター (例えば、ヘルパープラスミドまたはヘルパーウイルス) において利用され得、または修飾がなされ得る。これらの修飾には、キメラ Gag - Pol であって、Gag 及び Pol 配列が、異なるウイルス (例えば、異なる種、亜種、株、分岐群など) から得られ得、及び/または配列が、転写及び/または翻訳を改善するように、及び/または組換えを減少させるように修飾されているものが含まれる。

10

## 【0379】

いくつかの実施形態では、レトロウイルス核酸には、(i) 野生型 INS 1 に対して RNA の核外搬出の制限を減少させる変異した INS 1 阻害配列を含む、(ii) フレームシフト及び未完熟終結をもたらす2つのヌクレオチド挿入を含有する、及び/または (iii) gag の INS 2、INS 3、及び INS 4 阻害配列を含まない gag タンパク質の 150 ~ 250 (例えば、168) ヌクレオチド部分をコードするポリヌクレオチドが含まれる。

## 【0380】

いくつかの実施形態では、本明細書に記載のベクターは、レトロウイルス (例えば、レンチウイルス) 配列及び非レンチウイルス性ウイルス配列の両方を含むハイブリッドベクターである。いくつかの実施形態では、ハイブリッドベクターは、逆転写、複製、組み込み及び/またはパッケージングのためのレトロウイルス、例えば、レンチウイルスの配列を含む。

20

## 【0381】

いくつかの実施形態では、ウイルスベクター骨格配列のほとんどまたはすべては、レンチウイルス、例えば、HIV - 1 に由来する。しかしながら、レトロウイルス及び/またはレンチウイルス配列の多くの異なる源が使用または組み合わせられ得、所定のレンチウイルス配列における多数の置換及び改変が、トランスファーベクターが本明細書に記載の機能を発揮する能力を損なうことなく提供され得ることが理解されるべきである。多様なレンチウイルスベクターが Naldini et al., (1996a, 1996b, and 1998)、Zufferey et al., (1997)、Dull et al., 1998, 米国特許番号 6,013,516、及び 5,994,136 に記載されており、その多くは、レトロウイルス核酸を生成するために適応され得る。

30

## 【0382】

いくつかの実施形態では、プロウイルスの各末端には、長い末端反復 (LTR) が典型的に見られる。LTR は、典型的には、それらの天然配列状況において、ダイレトリートであり、かつ U3、R 及び U5 領域を含有するレトロウイルス核酸の末端に位置するドメインを含む。LTR は通常、レトロウイルス遺伝子の発現 (例えば、遺伝子転写産物の促進、開始及びポリアデニル化) 及びウイルス複製を促進する。LTR は、転写コントロール要素、ポリアデニル化シグナルならびにウイルスゲノムの複製及び組み込みのための配列を含む、多数の制御シグナルを含み得る。ウイルス LTR は、典型的には、U3、R 及び U5 と呼ばれる3つの領域に分けられる U3 領域は、典型的には、エンハンサー及びプロモーター要素を含有する。U5 領域は、典型的には、プライマー結合部位と R 領域との間の配列であり、ポリアデニル化配列を含有し得る。R (リピート) 領域は、U3 及び U5 領域によって挟まれ得る。LTR は、典型的には、U3、R 及び U5 領域から構成され、ウイルスゲノムの 5' 及び 3' 末端の両方に現れ得る。いくつかの実施形態では、5' LTR には、ゲノムの逆転写のための配列 (tRNA プライマー結合部位) 及びウイルス RNA の粒子への効率的なパッケージングのための配列 (プサイ部位) が隣接する。

40

## 【0383】

いくつかの実施形態では、パッケージングシグナルは、ウイルス RNA のウイルスカプ

50



シドまたは粒子への挿入を媒介するレトロウイルスゲノム内に配置された配列を含み得、例えば、Clever et al., 1995. J. of Virology, Vol. 69, No. 4; pp. 2101 - 2109を参照されたい。いくつかのレトロウイルスベクターは、ウイルスゲノムのカプシド形成のための最小パッケージングシグナル（プサイ [ ] 配列）を使用する。

#### 【0384】

様々な実施形態では、レトロウイルス核酸は、修飾された5'LTR及び/または3'LTRを含む。LTRのいずれかまたは両方は、1つ以上の欠失、挿入、または置換を含むがこれらに限定されない1つ以上の修飾を含み得る。3'LTRの修飾は、しばしば、ウイルスを複製欠損性とすることによってレンチウイルスまたはレトロウイルスシステムの安全性を改善するために行われる（例えば、感染性ビリオンが生成されないように完全で有効な複製が不可能なウイルス（例えば、複製欠損性レンチウイルス子孫））。

10

#### 【0385】

いくつかの実施形態では、ベクターは、自己不活性化（SIN）ベクター、例えば、複製欠損性ベクター、例えば、レトロウイルスまたはレンチウイルスベクターであって、右の（3'）LTRエンハンサー-プロモーター領域（U3領域として知られている）が、ウイルス複製の第1ラウンドを超えてウイルス転写を防止するように（例えば、欠失または置換によって）修飾されているものである。これは、右の（3'）LTR U3領域がウイルス複製中に左の（5'）LTR U3領域のためのテンプレートとして使用され得、よって、U3エンハンサー-プロモーターの非存在がウイルス複製を阻害するからである。実施形態では、3'LTRは、U5領域が除去、改変、または例えば、外因性ポリ（A）配列で置き換えられるように修飾される。3'LTR、5'LTR、または3'及び5'LTRの両方は、修飾されたLTRであり得る。

20

#### 【0386】

いくつかの実施形態では、5'LTRのU3領域は、ウイルス粒子の生成中にウイルスゲノムの転写を誘導するために異種プロモーターに置き換えられる。使用され得る異種プロモーターの例には、例えば、ウイルス性サルウイルス40（SV40）（例えば、初期または後期）、サイトメガロウイルス（CMV）（例えば、前初期）、モロニマウス白血病ウイルス（MoMLV）、ラウス肉腫ウイルス（RSV）、及び単純ヘルペスウイルス（HSV）（チミジンキナーゼ）プロモーターが含まれる。いくつかの実施形態では、プロモーターは、Tat非依存的様式で高レベルの転写を誘導することができる。所定の実施形態では、異種プロモーターは、ウイルスゲノムが転写される様式をコントロールする際に追加の利点を有する。例えば、異種プロモーターは、誘導性であり得、これによりウイルスゲノムのすべてまたは一部の転写が、誘導因子が存在する場合にのみ生じる。誘導因子には、1つ以上の化学的化合物または宿主細胞が培養される温度またはpHなどの生理学的条件が含まれるがこれらに限定されない。

30

#### 【0387】

いくつかの実施形態では、ウイルスベクターは、例えば、レンチウイルス（例えば、HIV）LTRのR領域に位置する、TAR（トランス活性化反応）要素を含む。この要素は、レンチウイルストランス活性化因子（tat）遺伝子要素と相互作用してウイルス複製を向上させる。しかしながら、この要素は、例えば、5'LTRのU3領域が異種プロモーターによって置き換えられた実施形態では必要とされない。

40

#### 【0388】

いくつかの実施形態では、R領域、例えば、キャッピング基の開始（すなわち、転写の開始）で始まり、ポリAトラクトの開始の直前に終了するレトロウイルスLTR内の領域は、U3及びU5領域によって挟まれ得る。R領域は、ゲノムの一端から他端への発生前DNAの移行において逆転写中に役割を果たす。

#### 【0389】

いくつかの実施形態では、レトロウイルス核酸はまた、FLAP要素、例えば、配列がレトロウイルス、例えば、HIV-1またはHIV-2のセントラルポリプリン配列及び

50

セントラル終結配列 (cPPT及びCTS) を含む核酸を含み得る。好適なFLAP要素は、米国特許番号6,682,907及びZennou, et al., 2000, Cell, 101:173 (それらの全体が参照により本明細書に組み込まれる) に記載されている。HIV-1逆転写中に、セントラルポリプリン配列 (cPPT) におけるプラス鎖DNAの中央開始及びセントラル終結配列 (CTS) における中央終結は、3本鎖DNA構造: HIV-1セントラルDNAフラップの形成をもたらす。いくつかの実施形態では、レトロウイルスまたはレンチウイルスベクター骨格は、外因性物質をコードする遺伝子の上流または下流に1つ以上のFLAP要素を含む。例えば、いくつかの実施形態では、トランスファープラスミドには、FLAP要素、例えば、HIV-1に由来する、またはそれから単離されたFLAP要素が含まれる。

10

## 【0390】

実施形態では、レトロウイルスまたはレンチウイルス核酸は、1つ以上の搬出要素、例えば、細胞の核から細胞質へのRNA転写産物の輸送を制御するシス作用性転写後制御要素を含む。RNA搬出要素の例には、ヒト免疫不全ウイルス (HIV) rev反応要素 (RRE) (例えば、Cullen et al., 1991, J. Virol. 65:1053; 及びCullen et al., 1991, Cell 58:423を参照されたい)、及びB型肝炎ウイルス転写後制御要素 (HPRE) が含まれるがこれらに限定されず、それらの全体が参照により本明細書に組み込まれる。通常、RNA搬出要素は、遺伝子の3' UTR内に配置され、1つまたは複数のコピーとして挿入され得る。

20

## 【0391】

いくつかの実施形態では、ウイルスベクターにおける異種配列の発現は、転写後制御要素、ポリアデニル化部位、及び転写終結シグナルのうちの一つ以上、例えば、すべてをベクターに組み込むことによって増加する。多様な転写後制御要素、例えば、ウッドチャック肝炎ウイルス転写後制御要素 (WPRE; Zufferey et al., 1999, J. Virol., 73:2886)、B型肝炎ウイルスに存在する転写後制御要素 (HPRE) (Huang et al., Mol. Cell. Biol., 5:3864)、及び類似物 (Liu et al., 1995, Genes Dev., 9:1766) (その各々は、その全体が参照により本明細書に組み込まれる) がタンパク質での異種核酸の発現を増加させ得る。いくつかの実施形態では、本明細書に記載のレトロウイルス核酸は、転写後制御要素、例えば、WPREまたはHPREを含む。

30

## 【0392】

いくつかの実施形態では、本明細書に記載のレトロウイルス核酸は、転写後制御要素、例えば、WPREまたはHPREを欠くまたは含まない。

## 【0393】

いくつかの実施形態では、異種核酸転写産物の終結及びポリアデニル化を指示する要素が、例えば、外因性物質の発現を増加させるために、含まれ得る。転写終結シグナルは、ポリアデニル化シグナルの下流で見ることができる。いくつかの実施形態では、ベクターは、外因性物質をコードするポリヌクレオチドの3' にポリアデニル化配列を含む。ポリA部位は、RNAポリメラーゼIIによって発生期RNA転写産物の終結及びポリアデニル化の両方を指示するDNA配列を含み得る。ポリアデニル化配列は、コード配列の3' 40  
末端へのポリAテールの付加によってmRNA安定性を促進し、よって、増加した翻訳効率に寄与し得る。レトロウイルス核酸において使用され得るポリAシグナルの例示的な例には、AATAAA、ATTAAA、AGTAAA、ウシ成長ホルモンポリA配列 (BGHpA)、ウサギ - グロビンポリA配列 (rgpA)、または別の好適な異種もしくは内因性ポリA配列が含まれる。

40

## 【0394】

いくつかの実施形態では、レトロウイルスまたはレンチウイルスベクターは、1つ以上のインシュレーター要素、例えば、本明細書に記載のインシュレーター要素をさらに含む。

## 【0395】

50

様々な実施形態では、ベクターは、外因性物質をコードするポリヌクレオチドに機能可能に連結されたプロモーターを含む。ベクターは、1つ以上のLTRを有し得、いずれかのLTRは、1つ以上の修飾、例えば、1つ以上のヌクレオチド置換、付加、または欠失を含む。ベクターは、形質導入効率（例えば、cPPT/FLAP）を増加させるためのより多くのアクセサリ要素、ウイルスパッケージング（例えば、プサイ（ ）パッケージングシグナル、RRE）、及び/または外因性遺伝子発現を増加させる他の要素（例えば、ポリ（A）配列）のうちの1つをさらに含み得、任意にWPREまたはHPREを含み得る。

#### 【0396】

いくつかの実施形態では、レンチウイルス核酸は、例えば、5'から3'まで、プロモーター（例えば、CMV）、R配列（例えば、TARを含む）、U5配列（例えば、組み込みのため）、PBS配列（例えば、逆転写のため）、DIS配列（例えば、ゲノム二量化のため）、プサイパッケージングシグナル、部分的gag配列、RRE配列（例えば、核外搬出のため）、cPPT配列（例えば、核搬入のため）、外因性物質の発現を誘導するためのプロモーター、外因性物質をコードする遺伝子、WPRE配列（例えば、効率的な導入遺伝子発現のため）、PPT配列（例えば、逆転写のため）、R配列（例えば、ポリアデニル化及び終結のため）、及びU5シグナル（例えば、組み込みのため）のうちの1つ以上、例えば、すべてを含む。

#### 【0397】

b. パッケージングベクター及びプロデューサー細胞

大規模なウイルス粒子生成は、しばしば、所望のウイルス力価を達成するために有用である。ウイルス粒子は、ウイルス構造及び/またはアクセサリ遺伝子、例えば、gag、pol、env、tat、rev、vif、vpr、vpu、vpx、もしくはnef遺伝子または他のレトロウイルス遺伝子を含むパッケージング細胞株にトランスファーベクターをトランスフェクションすることによって生成され得る。

#### 【0398】

いくつかの実施形態では、パッケージングベクターは、パッケージングシグナルを欠き、かつ1、2、3、4またはそれ以上のウイルス構造及び/またはアクセサリ遺伝子をコードするポリヌクレオチドを含む発現ベクターまたはウイルスベクターである。典型的には、パッケージングベクターは、プロデューサー細胞に含まれ、トランスフェクション、形質導入または感染を介して細胞に導入される。レトロウイルス、例えば、レンチウイルストランスファーベクターは、源細胞または細胞株を生成するためにトランスフェクション、形質導入または感染を介してプロデューサー細胞株に導入され得る。パッケージングベクターは、例えば、リン酸カルシウムトランスフェクション、リポフェクションまたはエレクトロポレーションを含む標準的な方法によってヒト細胞または細胞株に導入され得る。いくつかの実施形態では、パッケージングベクターは、顕性選択マーカー、例えば、ネオマイシン、ハイグロマイシン、ピューロマイシン、ブラストサイジン、ゼオシン、チミジンキナーゼ、DHFR、GlnシンテターゼまたはADAと共に細胞に導入され、続いて適切な薬物の存在下で選択され、クローンが単離される。選択マーカー遺伝子は、パッケージングベクターによって、例えば、IRESまたは自己切断性ウイルスペプチドによってコードする遺伝子に物理的に連結され得る。

#### 【0399】

いくつかの実施形態では、プロデューサー細胞株には、パッケージングシグナルを含有しないが、ウイルス粒子をパッケージングし得るウイルス構造タンパク質及び複製酵素（例えば、gag、pol及びenv）を安定的にまたは一過性で発現する細胞株が含まれる。任意の好適な細胞株、例えば、哺乳動物細胞、例えば、ヒト細胞が採用され得る。使用され得る好適な細胞株には、例えば、CHO細胞、BHK細胞、MDCK細胞、C3H10T1/2細胞、FLY細胞、Psi-2細胞、BOSC 23細胞、PA317細胞、WEHI細胞、COS細胞、BSC 1細胞、BSC 40細胞、BMT 10細胞、VERO細胞、W138細胞、MRC5細胞、A549細胞、HT1080細胞、293

10

20

30

40

50

細胞、293T細胞、B-50細胞、3T3細胞、NIH3T3細胞、HepG2細胞、Saos-2細胞、Huh7細胞、HeLa細胞、W163細胞、211細胞、及び211A細胞が含まれる。実施形態では、パッケージング細胞は、293細胞、293T細胞、またはA549細胞である。

#### 【0400】

いくつかの実施形態では、源細胞株には、プロデューサー細胞株及びパッケージングシグナルを含むトランスファーベクターコンストを含む、組換えレトロウイルス粒子を生成することが可能な細胞株が含まれる。ウイルスストック溶液を調製する方法は、例えば、Y. Soneoka et al. (1995) Nucl. Acids Res. 23: 628-633、及びN. R. Landau et al. (1992) J. Virol . 66: 5110-5113 (参照により本明細書に組み込まれる)によって説明されている。感染性ウイルス粒子は、例えば、細胞溶解または、細胞培養物の上清の収集によって、プロデューサー細胞から収集され得る。任意に、収集されたウイルス粒子は、濃縮または精製され得る。

10

#### 【0401】

いくつかの実施形態では、源細胞は、ウイルス粒子をパッケージングし得るウイルス構造タンパク質及び複製酵素(例えば、gag、pol及びenv)をコードする1つ以上のプラスミドを含む。いくつかの実施形態では、gag、pol、及びenv前駆体のうちの少なくとも2つをコードする配列は、同じプラスミド上にある。いくつかの実施形態では、gag、pol、及びenv前駆体をコードする配列は、異なるプラスミド上にある。いくつかの実施形態では、gag、pol、及びenv前駆体をコードする配列は、同じ発現シグナル、例えば、プロモーターを有する。いくつかの実施形態では、gag、pol、及びenv前駆体をコードする配列は、異なる発現シグナル、例えば、異なるプロモーターを有する。いくつかの実施形態では、gag、pol、及びenv前駆体の発現は、誘導性である。いくつかの実施形態では、ウイルス構造タンパク質及び複製酵素をコードするプラスミドは、同時にまたは異なる時間でトランスフェクションされる。いくつかの実施形態では、ウイルス構造タンパク質及び複製酵素をコードするプラスミドは、パッケージングベクターと同時に、または異なる時間にトランスフェクションされる。

20

#### 【0402】

いくつかの実施形態では、源細胞株は、1つ以上の安定に組み込まれたウイルス構造遺伝子を含む。いくつかの実施形態では、安定に組み込まれたウイルス構造遺伝子の発現は、誘導性である。

30

#### 【0403】

いくつかの実施形態では、ウイルス構造遺伝子の発現は、転写レベルで制御される。いくつかの実施形態では、ウイルス構造遺伝子の発現は、翻訳レベルで制御される。いくつかの実施形態では、ウイルス構造遺伝子の発現は、翻訳後レベルで制御される。

#### 【0404】

いくつかの実施形態では、ウイルス構造遺伝子の発現は、テトラサイクリン(Tet)依存性システムによって制御され、Tet制御転写リプレッサー(Tet-R)が、プロモーターに含まれるDNA配列に結合し、立体障害によって転写を抑制する(Yao et al, 1998、Jones et al, 2005)。ドキシサイクリン(dox)を添加すると、Tet-Rが放出され、転写が可能となる。複数の他の好適な転写制御プロモーター、転写因子、及び小分子誘導因子は、ウイルス構造遺伝子の転写を制御するために好適である。

40

#### 【0405】

いくつかの実施形態では、Tet制御プロモーターのコントロール下にありかつ抗生物質耐性カセットと結合した第三世代レンチウイルス成分であるヒト免疫不全ウイルス1型(HIV)Rev、Gag/Pol、及びエンベロープが源細胞ゲノムに別々に組み込まれる。いくつかの実施形態では、源細胞は、ゲノムに組み込まれたRev、Gag/Pol、及びエンベロープタンパク質の各々の1コピーのみを有する。

50

## 【0406】

いくつかの実施形態では、外因性物質をコードする核酸（例えば、外因性物質をコードするレトロウイルス核酸）もまた、源細胞ゲノムに組み込まれる。

## 【0407】

いくつかの実施形態では、本明細書に記載のレトロウイルス核酸は、逆転写を受けることができない。そのような核酸は、実施形態では、外因性物質を一過性に発現することができる。レトロウイルスまたはVLPは、無効化された逆転写酵素タンパク質を含む場合があり、または逆転写酵素タンパク質を含まない場合がある。実施形態では、レトロウイルス核酸は、無効化されたプライマー結合部位（PBS）及び/またはatt部位を含む。実施形態では、rev、tat、vif、nef、vpr、vpu、vpx及びS2またはその機能的同等物を含む1つ以上のウイルスアクセサリ遺伝子は、機能不全とされており、またはレトロウイルス核酸に存在しない。実施形態では、S2、rev及びtatから選択される1つ以上のアクセサリ遺伝子は、機能不全とされており、またはレトロウイルス核酸には存在しない。

10

## 【0408】

## 2. 細胞由来粒子

本明細書では、天然由来の膜を含む標的化脂質粒子が提供される。いくつかの実施形態では、天然由来の膜は、細胞または組織から調製された膜小胞を含む。いくつかの実施形態では、標的化脂質粒子は、細胞から得ることができる小胞を含む。いくつかの実施形態では、標的化脂質粒子は、微小胞、エクソソーム、膜包囲体、アポトーシス小体（アポトーシス細胞からのもの）、粒子（例えば、血小板に由来し得る）、エクソソーム（例えば、血清における好中球及び単球に由来し得る）、プロスタトソーム（前立腺癌細胞から得ることができる）、またはカーディオソーム（心臓細胞に由来し得る）を含む。

20

## 【0409】

いくつかの実施形態では、源細胞は、内皮細胞、線維芽細胞、血液細胞（例えば、マクロファージ、好中球、顆粒球、白血球）、幹細胞（例えば、間葉幹細胞、臍帯幹細胞、骨髄幹細胞、造血幹細胞、誘導多能性幹細胞、例えば、対象の細胞に由来する誘導多能性幹細胞）、胚性幹細胞（例えば、胚卵黄嚢、胎盤、臍帯、胎児皮膚、青年皮膚、血液、骨髄、脂肪組織、赤血球生成組織、造血組織からの幹細胞）、筋芽細胞、実質細胞（例えば、肝細胞）、肺胞細胞、ニューロン（例えば、網膜神経細胞）前駆体細胞（例えば、網膜前駆体細胞、骨髄芽球、骨髄前駆体細胞、胸腺細胞、減数分裂細胞、巨核芽球、前巨核芽球、メラノblast、リンパ芽球、骨髄前駆体細胞、正赤芽球、または血管芽細胞）、前駆細胞（例えば、心臓前駆細胞、衛星細胞、放射状グリア細胞、骨髄間質細胞、膵臓前駆細胞、内皮前駆細胞、芽細胞）、または不死化細胞（例えば、HeEa、HEK293、HFF-1、MRC-5、WI-38、IMR90、IMR91、PER.C6、HT-1080、またはBJ細胞）である。いくつかの実施形態では、源細胞は、293細胞、HEK細胞、ヒト内皮細胞、またはヒト上皮細胞、単球、マクロファージ、樹状細胞、または幹細胞以外である。

30

## 【0410】

いくつかの実施形態では、標的化脂質粒子は、1未満、1~1.1、1.05~1.15、1.1~1.2、1.15~1.25、1.2~1.3、1.25~1.35、または1.35超g/mlの密度を有する。いくつかの実施形態では、標的化脂質粒子組成物は、タンパク質量に基づいて0.01%、0.05%、0.1%、0.5%、1%、1.5%、2%、2.5%、3%、4%、5%、または10%未満の源細胞または0.01%、0.05%、0.1%、0.5%、1%、1.5%、2%、2.5%、3%、4%、5%、または10%未満の機能性核を有する細胞を含む。

40

## 【0411】

実施形態では、標的化脂質粒子は、サイズを有し、または標的化脂質粒子の集団は、源細胞のものの約0.01%、0.05%、0.1%、0.5%、1%、2%、3%、4%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%未

50

満である平均サイズを有する。

【0412】

いくつかの実施形態では、標的化脂質粒子は、細胞外小胞、例えば、内部空間を包囲し、かつその由来となった細胞よりも小さな直径を有する膜を含む細胞由来小胞を含む。実施形態では、細胞外小胞は、20 nm ~ 1000 nmの直径を有する。実施形態では、標的化脂質粒子は、アポトーシス小体、細胞の断片、直接的または間接的操作によって細胞から誘導された小胞、小胞化細胞小器官、及び生存細胞によって（例えば、直接的細胞膜出芽または後期エンドソームと細胞膜との融合によって）生成される小胞を含む。実施形態では、細胞外小胞は、生存もしくは死生物、移植された組織もしくは器官、または培養された細胞に由来する。

10

【0413】

実施形態では、標的化脂質粒子は、ナノ小胞、例えば、内部空間を包囲し、かつ直接的または間接的操作によって前記細胞から生成された膜を含む細胞由来の小さな（例えば、直径が20 ~ 250 nm、または直径が30 ~ 150 nm）小胞を含む。ナノ小胞の生成は、いくつかの例では、源細胞の破壊をもたらし得る。ナノ小胞は、脂質または脂肪酸及びポリペプチドを含み得る。

【0414】

実施形態では、標的化脂質粒子は、エクソソームを含む。実施形態では、エクソソームは、内部空間を包囲し、かつ直接的細胞膜出芽によってまたは後期エンドソームと細胞膜との融合によって前記細胞から生成された膜を含む細胞由来の小さな（例えば、直径が20 ~ 300 nm、または直径が40 ~ 200 nm）小胞である。実施形態では、エクソソームの生成は、源細胞の破壊をもたらさない。実施形態では、エクソソームは、脂質または脂肪酸及びポリペプチドを含む。例示的なエクソソーム及び他の膜包囲体はまた、WO / 2017 / 161010、WO / 2016 / 077639、US 20160168572、US 20150290343、及びUS 20070298118（その各々は、その全体が参照により本明細書に組み込まれる）に記載されている。

20

【0415】

いくつかの実施形態では、標的化脂質粒子は、免疫調節因子の増加した発現をもたらす遺伝子修飾を有する源細胞に由来する。いくつかの実施形態では、免疫抑制因子は、細胞の外部表面にある。いくつかの実施形態では、免疫抑制因子は、標的化脂質粒子の外部表面に組み込まれる。いくつかの実施形態では、標的化脂質粒子は、共有結合または非共有結合によって固体粒子の表面に結合した免疫調節因子を含む。

30

【0416】

c . a . 細胞由来粒子の生成

いくつかの実施形態では、標的化脂質粒子は、エクソソーム、微小胞、膜小胞、細胞外膜小胞、細胞膜小胞、巨大細胞膜小胞、アポトーシス小体、ミト粒子、ピレノサイト、リソソーム、または他の膜包囲小胞の出芽を誘導することによって生成される。

【0417】

いくつかの実施形態では、標的化脂質粒子は、細胞除核を誘導することによって生成される。除核は、遺伝子的、化学的（例えば、アクチノマイシンDを使用する、Bayona - Bafaluyet al . , " A chemical enucleation method for the transfer of mitochondrial DNA to cells " Nucleic Acids Res . 2003 Aug 15 ; 31 ( 16 ) : e98を参照されたい）、機械的方法（例えば、絞りまたは吸入、Lee et al . , " A comparative study on the efficiency of two enucleation methods in pig somatic cell nuclear transfer : effects of the squeezing and the aspiration methods . " Anim Biotechnol . 2008 ; 19 ( 2 ) : 71 - 9を参照されたい）、またはそれらの組み合わせなどのアッセイを使用して実施され得る。

40

50

## 【0418】

いくつかの実施形態では、標的化脂質粒子は、細胞断片化を誘導することによって生成される。いくつかの実施形態では、細胞断片化は、以下の方法（化学的方法、機械的方法（例えば、遠心分離（例えば、超遠心分離、または密度遠心分離）、凍結解凍、または超音波処理）、またはそれらの組み合わせを含むがこれらに限定されない）を使用して実施され得る。

## 【0419】

いくつかの実施形態では、標的化脂質粒子は、微小胞である。いくつかの実施形態では微小胞は、約100nm～約2000nmの直径を有する。いくつかの実施形態では、標的化脂質粒子は、細胞ゴーストを含む。いくつかの実施形態では、小胞は、細胞膜小胞、例えば、巨大細胞膜小胞である。

10

## 【0420】

いくつかの実施形態では、標的化脂質粒子を作製するために使用された源細胞は、標的化脂質粒子が作製された後の試験のために利用可能ではない。

## 【0421】

いくつかの実施形態では、標的化脂質粒子の特徴は、参照細胞に対する比較によって説明される。実施形態では、参照細胞は、源細胞である。実施形態では、参照細胞は、HeLa、HEK293、HFF-1、MRC-5、WI-38、IMR90、IMR91、PER.C6、HT-1080、またはBJ細胞である。いくつかの実施形態では、標的化脂質粒子の集団の特徴は、参照細胞の集団、例えば、源細胞の集団、またはHeLa、HEK293、HFF-1、MRC-5、WI-38、IMR90、IMR91、PER.C6、HT-1080、またはBJ細胞の集団に対する比較によって記載される。

20

## 【0422】

## III. 薬学的組成物

本開示はまた、いくつかの態様では、本明細書に記載の標的化脂質粒子組成物及び薬学的に許容可能な担体を含む薬学的組成物を提供する。薬学的組成物は、記載される標的化脂質粒子のいずれかを含み得る。

## 【0423】

いくつかの実施形態では、標的化脂質粒子は、薬学的または適正製造基準（GMP）標準を満たす。いくつかの実施形態では、標的化脂質粒子は、適正製造基準（GMP）に従って作製された。いくつかの実施形態では、標的化脂質粒子は、既定の参照値未満の病原体レベルを有し、例えば、病原体を実質的に含まない。いくつかの実施形態では、標的化脂質粒子は、既定の参照値未満の混入物質レベルを有し、例えば、混入物質を実質的に含まない。いくつかの実施形態では、標的化脂質粒子は、低い免疫原性を有する。

30

## 【0424】

いくつかの実施形態では、本明細書では、本発明の方法を実施するための本発明の薬学的組成物またはその塩の使用が提供される。そのような薬学的組成物は、対象への投与に好適な形態の本発明の少なくとも1つの化合物もしくはコンジュゲートまたはその塩からなり得、または薬学的組成物は、本発明の少なくとも1つの化合物もしくはコンジュゲートまたはその塩、及び1つ以上の薬学的に許容可能な担体、1つ以上の追加の成分、またはこれらのいくつかの組み合わせを含み得る。いくつかの実施形態では、本発明の化合物またはコンジュゲートは、例えば、当該技術分野でよく知られているように、生理的に許容可能なカチオンまたはアニオンと組み合わせて、生理的に許容可能な塩の形態で薬学的組成物に存在し得る。

40

## 【0425】

いくつかの実施形態では、本発明の方法を実施するのに有用な薬学的組成物は、1ng/kg/日～100mg/kg/日の間の用量で送達するために投与され得る。別の実施形態では、本発明を実施するのに有用な薬学的組成物は、1ng/kg/日～500mg/kg/日の間の用量で送達するために投与され得る。

## 【0426】

50

いくつかの実施形態では、本発明の薬学的組成物における活性成分、薬学的に許容可能な担体、及び任意の追加の成分の相対量は、処置される対象が何であるか、そのサイズ、及び状態に応じて、さらに組成物が投与される経路に応じて変化する。いくつかの実施形態では、組成物は、0.1%~100%(w/w)の活性成分を含み得る。

【0427】

いくつかの実施形態では、本発明の方法において有用である薬学的組成物は、経口、直腸、腔、非経口、局所、肺、鼻腔内、口腔内、眼、または別の投与経路のために好適に開発され得る。いくつかの実施形態では、本発明の方法において有用な組成物は、哺乳動物の皮膚、腔または任意の他の組織に直接的に投与され得る。いくつかの実施形態では、製剤には、リポソーム調製物、活性成分を含有する再封止された赤血球、及び免疫学ベースの製剤が含まれる。いくつかの実施形態では、投与の経路(複数可)は、当業者に容易に明らかとなり、処置されている疾患のタイプ及び重症度、処置されている獣医学的またはヒト対象のタイプ及び年齢などを含む多数の要素に依存する。

10

【0428】

いくつかの実施形態では、本明細書に記載の薬学的組成物の製剤は、薬理学の技術分野で知られている、または後に開発され得る任意の方法によって調製され得る。いくつかの実施形態では、調製方法には、活性成分を担体または1つ以上の他の付随成分と会合させ、次いで、必要に応じてまたは所望により、生成物を所望の単一または複数の用量単位に成形または包装するステップが含まれる。

【0429】

いくつかの実施形態では、「単位用量」は、既定の量の活性成分を含む薬学的組成物の個別量である。いくつかの実施形態では、活性成分の量は通常、対象に投与されるであろう活性成分の投薬量または例えば、そのような投薬量の2分の1または3分の1などのそのような投薬量の好都合な一部と同等である。いくつかの実施形態では、単位投薬形態は、単回1日用量または複数の1日用量の1つ(例えば、1日当たり約1~4回以上)のためのものであり得る。いくつかの実施形態では、複数の1日用量が使用される場合、単位投薬形態は、各用量について同じまたは異なり得る。

20

【0430】

いくつかの実施形態では、本明細書で提供される薬学的組成物の説明は主に、ヒトへの医療用投与に好適な薬学的組成物を対象としているが、そのような組成物は通常、すべての種類の動物への投与に好適であることが当業者によって理解される。いくつかの実施形態では、組成物を様々な動物に対する投与に好適なものとするためのヒトへの投与に好適な薬学的組成物の変更はよく理解されており、通常、熟練された獣医学的薬理学者は、存在する場合でも単なる通常の実験を伴ってそのような修飾を設計及び実施し得る。いくつかの実施形態では、本発明の薬学的組成物の投与が企図される対象には、ヒト及び他の霊長類、商業的に関連する哺乳動物、例えば、ウシ、ブタ、ウマ、ヒツジ、ネコ、及びイヌを含む哺乳動物が含まれる。

30

【0431】

任意の実施形態のうちの一部では、本発明の組成物は、1つ以上の薬学的に許容可能な賦形剤または担体を使用して製剤化される。一実施形態では、本発明の薬学的組成物は、治療的有効量の本発明の化合物またはコンジュゲート及び薬学的に許容可能な担体を含む。いくつかの実施形態では、有用な薬学的に許容可能な担体には、グリセロール、水、生理食塩水、エタノール及び他の薬学的に許容可能な塩溶液、例えば、リン酸塩及び有機酸の塩が含まれるがこれらに限定されない。これらの及び他の薬学的に許容可能な担体の例は、Remington's Pharmaceutical Sciences(1991, Mack Publication Co., New Jersey)に記載されている。

40

【0432】

いくつかの実施形態では、担体は、例えば、水、エタノール、ポリオール(例えば、グリセロール、プロピレングリコール、及び液体ポリエチレングリコールなど)、好適なそ

50



これらの混合物、及び植物油を含有する溶媒または分散媒体であり得る。いくつかの実施形態では、適切な流動性は、例えば、レシチンなどのコーティングの使用によって、分散の場合には必要とされる粒子サイズの維持によって及び界面活性剤の使用によって維持され得る。いくつかの実施形態では、微生物の作用の防止は、様々な抗細菌及び抗真菌剤、例えば、パラベン、クロロブタノール、フェノール、アスコルビン酸、チメロサルなどによって達成され得る。いくつかの実施形態では、組成物において等張剤、例えば、糖類、塩化ナトリウム、またはポリアルコール、例えば、マンニトール及びソルビトールを含むことが好ましい。いくつかの実施形態では、注射用組成物の長期吸収は、吸収を遅らせる薬剤、例えば、モノステアリン酸アルミニウムまたはゼラチンを組成物に含めることによってもたらされ得る。一実施形態では、薬学的に許容可能な担体は、DMSO単独ではない。

10

#### 【0433】

いくつかの実施形態では、従来賦形剤、すなわち、当該技術分野で知られている、経口、腔、非経口、鼻、静脈内、皮下、腸内、または任意の他の好適な投与様式に好適な薬学的に許容可能な有機または無機担体物質と併せて用いられ得る。いくつかの実施形態では、薬学的調製物は、滅菌され得、所望の場合、補助剤、例えば、滑沢剤、防腐剤、安定化剤、湿潤剤、乳化剤、浸透圧緩衝液に影響を及ぼすための塩、着色、香味及び/または芳香物質などと混合され得る。いくつかの実施形態では、薬学的調製物はまた、所望の場合、他の活性剤、例えば、他の鎮痛剤と組み合わせられ得る。

#### 【0434】

いくつかの実施形態では、「追加の成分」には、以下のうちの1つ以上が含まれるがこれらに限定されない：賦形剤；表面活性剤；分散剤；不活性希釈剤；顆粒化及び崩壊剤；結合剤；滑沢剤；甘味剤；香味剤；着色剤；防腐剤；生理的分解性組成物、例えば、ゼラチン；水性ビヒクル及び溶媒；油性ビヒクル及び溶媒；懸濁剤；分散または湿潤剤；乳化剤、粘滑剤；緩衝液；塩；増粘剤；充填剤；乳化剤；抗酸化剤；抗生物質；抗真菌剤；安定化剤；及び薬学的に許容可能なポリマー性または疎水性物質。いくつかの実施形態では、本発明の薬学的組成物に含まれ得る「追加の成分」は、当該技術分野で知られており、例えば、Genaro, ed. (1985, Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa.) (参照により本明細書に組み込まれる)に記載されている。

20

#### 【0435】

いくつかの実施形態では、本発明の組成物は、組成物の総重量に基づいて約0.005%~2.0%の防腐剤を含み得る。いくつかの実施形態では、防腐剤は、環境における混入物質に曝露した場合に損傷を防止するために使用される。いくつかの実施形態では、本発明に従って有用な防腐剤の例には、ベンジルアルコール、ソルビン酸、パラベン、イミドウレア及びそれらの組み合わせからなる群から選択されるものが含まれるがこれらに限定されない。いくつかの実施形態では、特に好ましい防腐剤は、約0.5%~2.0%のベンジルアルコール及び0.05%~0.5%のソルビン酸の組み合わせである。

30

#### 【0436】

いくつかの実施形態では、組成物には、好ましくは、抗酸化剤及び化合物の分解を阻害するキレート剤が含まれる。いくつかの実施形態では、いくつかの化合物のための抗酸化剤は、組成物の総重量に基づいて重量で約0.01%~0.3%の好ましい範囲のBHT、BHA、アルファ-トコフェロール及びアスコルビン酸、より好ましくは0.03%~0.1%の範囲のBHTである。いくつかの実施形態では、キレート剤は、組成物の総重量に基づいて重量で0.01%~0.5%の量で存在する。特に好ましいキレート剤には、組成物の総重量に基づいて重量で約0.01%~0.20%の重量範囲及びより好ましくは0.02%~0.10%の範囲のエドト酸塩(例えば、エドト酸二ナトリウム)及びクエン酸が含まれる。いくつかの実施形態では、キレート剤は、製剤の保存可能期間に有害であり得る組成物中の金属イオンをキレートするのに有用である。いくつかの実施形態では、他の好適な及び同等の抗酸化剤及びキレート剤は、そのため、当業者に知られるように代用され得る。

40

50

## 【 0 4 3 7 】

いくつかの実施形態では、水性または油性ビヒクルにおける活性成分の懸濁液を達成するために従来の方法を使用して液体懸濁液が調製され得る。いくつかの実施形態では、水性ビヒクルには、例えば、水、及び等張性生理食塩水が含まれる。いくつかの実施形態では、油性ビヒクルには、例えば、アーモンド油、油性エステル、エチルアルコール、植物油、例えば、ラッカセイ、オリーブ、ゴマ、またはココナッツ油、分画植物油、及び鉱油、例えば、液体パラフィンが含まれる。いくつかの実施形態では、液体懸濁液は、懸濁剤、分散または湿潤剤、乳化剤、粘滑剤、防腐剤、緩衝液、塩、香味剤、着色剤、及び甘味剤を含むがこれらに限定されない1つ以上の追加の成分をさらに含み得る。いくつかの実施形態では、油性懸濁液は、増粘剤をさらに含み得る。いくつかの実施形態では、懸濁剤には、ソルビトールシロップ、水素化食用脂肪、アルギン酸ナトリウム、ポリビニルピロリドン、トラガカントガム、アカシアガム、及びセルロース誘導体、例えば、カルボキシメチルセルロースナトリウム、メチルセルロース、ヒドロキシプロピルメチルセルロースが含まれるがこれらに限定されない。いくつかの実施形態では、分散または湿潤剤には、天然に存在するホスファチド、例えば、レシチン、アルキレンオキシドと脂肪酸との、長鎖脂肪族アルコールとの、脂肪酸及びヘキシトールに由来する部分エステルとの、または脂肪酸及び無水ヘキシトール（例えば、それぞれ、ステアリン酸ポリオキシエチレン、ヘプタデカエチレンオキシセタノール、モノオレイン酸ポリオキシエチレンソルビトール、及びモノオレイン酸ポリオキシエチレンソルビタン）に由来する部分エステルとの縮合生成物が含まれるがこれらに限定されない。既知の乳化剤には、レシチン、及びアカシアが含まれるがこれらに限定されない。既知の防腐剤には、メチル、エチル、またはn-プロピル-バラ-ヒドロキシベンゾエート、アスコルビン酸、及びソルビン酸が含まれるがこれらに限定されない。既知の甘味剤には、例えば、グリセロール、プロピレングリコール、ソルビトール、スクロース、及びサッカリンが含まれる。油性懸濁液のための既知の増粘剤には、例えば、蜜ろう、固形パラフィン、及びセチルアルコールが含まれる。

10

20

## 【 0 4 3 8 】

いくつかの実施形態では、液体懸濁液と実質的に同じ手法（主な違いは、活性成分が溶媒に懸濁されるのではなく溶解されることである）で水性または油性溶媒中の活性成分の液体溶液が調製され得る。本明細書で使用される場合、「油性」液体は、炭素含有液体分子を含み、かつ水よりも低い極性特性を示すものである。いくつかの実施形態では、本発明の薬学的組成物の液体溶液は、液体懸濁液に関して記載される成分の各々を含み得、懸濁剤はかならずしも、溶媒中の活性成分の溶解を補助するわけではないことが理解される。いくつかの実施形態では、水性溶媒には、例えば、水、及び等張性生理食塩水が含まれる。いくつかの実施形態では、油性溶媒には、例えば、アーモンド油、油性エステル、エチルアルコール、植物油、例えば、ラッカセイ、オリーブ、ゴマ、またはココナッツ油、分画植物油、及び鉱油、例えば、液体パラフィンが含まれる。

30

## 【 0 4 3 9 】

いくつかの実施形態では、本発明の薬学的調製物の粉末及び顆粒製剤が既知の方法を使用して調製され得る。いくつかの実施形態では、製剤は、対象に直接的に投与され、例えば、錠剤を形成するため、カプセルを充填するため、または水性もしくは油性ビヒクルをそれに添加することによって水性もしくは油性懸濁液もしくは溶液を調製するために使用される。任意の実施形態のうちの一部では、製剤は、分散または湿潤剤、懸濁剤、及び防腐剤のうち1つ以上をさらに含み得る。追加の賦形剤、例えば、充填剤及び甘味剤、香味剤、または着色剤もまた、これらの製剤に含まれ得る。

40

## 【 0 4 4 0 】

いくつかの実施形態では、本発明の薬学的組成物はまた、水中油エマルションまたは油中水エマルションの形態で調製、包装、または販売され得る。いくつかの実施形態では、油性相は、植物油、例えば、オリーブまたはラッカセイ油、鉱油、例えば、液体パラフィン、またはこれらの組み合わせであり得る。いくつかの実施形態では、組成物は、1つ以上の乳化剤、例えば、天然に存在するガム、例えば、アカシアガムまたはトラガカントガ

50

ム、天然に存在するホスファチド、例えば、大豆またはレシチンホスファチド、脂肪酸及び無水ヘキシトールの組み合わせに由来するエステルまたは部分エステル、例えば、モノオレイン酸ソルピタン、ならびにそのような部分エステルとエチレンオキシドとの縮合生成物、例えば、モノオレイン酸ポリオキシエチレンソルピタンをさらに含む。いくつかの実施形態では、エマルションはまた、例えば、甘味剤または香味剤を含む追加の成分を含有し得る。

#### 【0441】

##### IV. 処置の方法

いくつかの実施形態では、本明細書で提供される標的化脂質粒子、または本明細書に記載されるその薬学的組成物は、対象、例えば、哺乳動物、例えば、ヒトに投与され得る。そのような実施形態では、対象は、特定の疾患または病態のリスクがある、その症状を有する、またはその診断がされている、またはそれを有するものとして特定されている場合がある。一実施形態では、対象は、がんを有する。一実施形態では、対象は、感染性疾患を有する。いくつかの実施形態では、標的化脂質粒子は、対象における疾患または病態を処置するための外因性物質をコードする核酸配列を含有する。例えば、外因性物質は、新生物細胞のタンパク質を標的とし、またはそれに特異的なものであり、標的化脂質粒子は、対象における腫瘍またはがんを処置するために対象に投与される。別の例では、外因性物質は、炎症性メディエーターまたは免疫分子、例えば、サイトカインであり、標的化脂質粒子は、免疫反応を調節する（例えば、増加させる）ことが所望である任意の病態、例えば、がんまたは感染性疾患を処置するために対象に投与される。いくつかの実施形態では、標的化脂質粒子は、疾患、病態または障害の処置を達成するための有効量または用量で投与される。本明細書では、そのような方法及び処置における、ならびにそのような治療方法を行うための薬品の調製における提供される標的化脂質粒子のいずれかの使用が提供される。いくつかの実施形態では、方法は、標的化脂質粒子またはそれを含む組成物を、疾患または病態または障害を有する、有していた、またはそれが疑われる対象に投与することによって行われる。いくつかの実施形態では、方法は、それにより対象における疾患または病態または障害を処置する。また、本明細書では、外因性物質によって標的とされるまたは提供される特定の遺伝子またはタンパク質に関連する疾患、病態または障害の処置のための、本明細書で提供される薬学的組成物などの組成物のいずれかの使用が提供される。

#### 【0442】

いくつかの実施形態では、提供される方法または使用は、経口、吸入、経皮または非経口（静脈内、腫瘍内、腹腔内、筋肉内、腔内、及び皮下を含む）投与を含む薬学的組成物の投与を伴う。いくつかの実施形態では、標的化脂質粒子は、単独で投与され得るか、または薬学的組成物として製剤化され得る。いくつかの実施形態では、本明細書に記載の標的化脂質粒子または組成物は、対象、例えば、哺乳動物、例えば、ヒトに投与され得る。任意の実施形態のうちの一部では、対象は、特定の疾患または病態（例えば、本明細書に記載の疾患または病態）のリスクがある、その症状を有する、またはその診断がされている、またはそれを有するものとして特定されている場合がある。いくつかの実施形態では、疾患は、疾患または障害である。

#### 【0443】

いくつかの実施形態では、標的化脂質粒子は、単位用量組成物の形態、例えば、単位用量の経口、非経口、経皮または吸入組成物で投与され得る。いくつかの実施形態では、組成物は、混和によって調製され、経口、吸入、経皮または非経口投与のために適応され、したがって、錠剤、カプセル、経口液体調製物、粉末、顆粒、ロゼンジ、再構成可能な粉末、注射可能な及び注入可能な溶液または懸濁液または座剤またはエアロゾルの形態であり得る。

#### 【0444】

いくつかの実施形態では、投与のレジメンは、有効量を構成するものに影響を及ぼし得る。いくつかの実施形態では、治療製剤は、疾患の診断の前または後のいずれかで対象に

10

20

30

40

50

投与され得る。いくつかの実施形態では、いくつかの分割された投薬量、ならびに調整された投薬量が、毎日または順次に投与され得、または用量は、連続的に注入され得、またはボラス注射であり得る。いくつかの実施形態では、治療製剤の投薬量は、治療的または予防的状況の要件によって示されるように比例的に増加または減少され得る。

【0445】

いくつかの実施形態では、対象、好ましくは哺乳動物、より好ましくはヒトへの本発明の組成物の投与は、疾患を予防または処置するのに有効な投薬量及び期間で、既知の手順を使用して行われ得る。いくつかの実施形態では、治療効果を達成するために必要な治療化合物の有効量は、用いられる特定の化合物の活性；投与の時間；化合物の排泄の速度；処置の継続期間；化合物と組み合わせて使用される他の薬物、化合物または物質；処置されている対象の疾患または障害の状態、年齢、性別、体重、状態、全般的健康状態及び既往歴、ならびに医学の分野でよく知られている同様の要素などの要素に応じて変わり得る。いくつかの実施形態では、投薬レジメンは、最適な治療反応を提供するために調整され得る。いくつかの実施形態では、いくつかの分割用量が毎日投与され得、または用量は、治療的状況の要件によって示されるように比例的に減少され得る。いくつかの実施形態では、本発明の治療化合物のための有効な用量範囲は、1日当たり約1から5,000 mg/kg体重である。当業者は、関連性のある要因を研究し、過度の実験を用いることなく治療化合物の有効量に関して決定を下すことができるであろう。

10

【0446】

いくつかの実施形態では、化合物は、1日数回もの頻度で対象に投与され得、またはそれは、より頻度が少なく、例えば、1日1回、週1回、2週に1回、月1回、またはさらに頻度が少なく、例えば、数ヶ月毎に1回もしくはさらに1年に1回またはそれ以下で投与され得る。いくつかの実施形態では、1日当たり投薬される化合物の量は、非限定的な例では、1日毎、1日おき、2日毎、3日毎、4日毎、または5日毎に投与され得る。いくつかの実施形態では、1日おきの投与を用いて、1日当たり5 mgの用量が月曜日に開始され得、第1の後続の1日当たり5 mgの用量が水曜日に投与され、第2の後続の1日当たり5 mgの用量が金曜日に投与されるなどである。用量の頻度は、当業者に容易に明らかとなり、例えば、限定されないが、処置されている疾患のタイプ及び重症度、動物のタイプ及び年齢などの多数の要因に依存する。

20

【0447】

いくつかの実施形態では、本発明の薬学的組成物における活性成分の投薬量レベルは、対象への毒性を伴うことなく、特定の対象、組成物、及び投与様式のための所望の治療的反応を達成するのに有効な活性成分の量を得るように変更され得る。

30

【0448】

当該技術分野における通常の技術を有する医師、例えば、内科医または獣医師は、必要とされる薬学的組成物の有効量を容易に決定し、処方し得る。いくつかの実施形態では、医師または獣医師は、所望の治療効果を達成するために必要とされるものよりも少ないレベルにて薬学的組成物で用いられる本発明の化合物の用量を開始し、所望の効果が達成されるまで投薬量を徐々に増加させ得る。

【0449】

いくつかの実施形態では、投与の容易性及び投薬量の均一性のために投薬単位形態で化合物を製剤化することが特に有利である。いくつかの実施形態では、投薬単位形態は、本明細書で使用される場合、処置される対象のための単一の投薬量として適した物理的に別個の単位を指す。各単位は、必要とされる薬学的ビヒクルと関連した所望の治療効果を生成するために計算された既定量の治療化合物を含有する。いくつかの実施形態では、本発明の投薬単位形態は、(a)治療化合物及び達成される特定の治療効果の独自の特性、及び(b)対象における疾患の処置のためのそのような治療化合物を調合/製剤化する当該技術分野に固有の制限によって影響を受け、及び/またはそれに直接的に依存する。

40

【0450】

いくつかの実施形態では、「容器」という用語には、薬学的組成物を保持するための任

50

意の入れ物が含まれる。いくつかの実施形態では、容器は、薬学的組成物を含有する包装である。他の実施形態では、容器は、薬学的組成物を含有する包装ではなく、すなわち、容器は、包装された薬学的組成物または未包装の薬学的組成物及び薬学的組成物の使用のための説明書を含有するボックスまたはバイアルなどの入れ物である。薬学的組成物の使用のための説明書は、薬学的組成物を含有する包装上に含有され得、このように説明書は、包装された製品に対する増加した機能的関係性を形成することが理解されるべきである。いくつかの実施形態では、説明書は、化合物がその意図された機能、例えば、対象における疾患を処置もしくは予防する、または画像化もしくは診断剤を対象に送達する機能を発揮する能力に関する情報を含有し得る。

#### 【0451】

いくつかの実施形態では、本明細書に開示される組成物のいずれの投与経路にも、経口、鼻、直腸、非経口、舌下、経皮、経粘膜（例えば、舌下、舌、（経）口腔内、（経）尿道、膣（例えば、経膣及び膣周囲）、鼻（内）、及び（経）直腸）、膀胱内、肺内、十二指腸内、胃内、髄腔内、皮下、筋肉内、皮内、動脈内、静脈内、気管支内、吸入、及び局所投与が含まれる。

#### 【0452】

任意の実施形態のうちの一部では、好適な組成物及び投薬形態には、例えば、錠剤、カプセル、カプレット、丸剤、ジェルキャップ、トローチ、分散体、懸濁液、溶液、シロップ、顆粒、ビーズ、経皮膚パッチ、ゲル、粉末、ペレット、マグマ、ロゼンジ、クリーム、ペースト、プラスター、ローション、ディスク、座剤、鼻または経口投与のための液体スプレー、吸入のための乾燥粉末またはエアロゾル化製剤、膀胱内投与のための組成物及び製剤などが含まれる。

#### 【0453】

いくつかの実施形態では、外因性物質またはカーゴを含む標的化脂質粒子組成物は、そのような外因性物質またはカーゴを細胞組織または対象に送達するために使用され得る。いくつかの実施形態では、本明細書に記載の標的化脂質粒子組成物の投与によるカーゴの送達は、細胞タンパク質発現レベルを改変し得る。所定の実施形態では、投与される組成物は、ポリペプチドが送達される細胞において実質的に非存在のまたは減少した機能的活性を提供する1つ以上のカーゴ（例えば、ポリペプチドまたはmRNA）の上方制御（細胞における発現、細胞における送達、または細胞内の誘導を介する）を誘導する。いくつかの実施形態では、欠落した機能的活性は、酵素的、構造的、または制御的の性質のものであり得る。いくつかの実施形態では、投与される組成物は、ポリペプチドが上方制御された細胞において存在するが、実質的に欠損している機能的活性を増加させる（例えば、相乗的に）1つ以上のポリペプチドの上方制御を誘導する。任意の実施形態のうちの一部では、投与される組成物は、ポリペプチド、siRNA、またはmiRNAが送達される細胞において存在するまたは上方制御された機能的活性を抑制する1つ以上のカーゴ（例えば、ポリペプチド、siRNA、またはmiRNA）の下方制御（細胞における発現、細胞における送達、または細胞内の誘導を介する）を誘導する。任意の実施形態のうちの一部では、上方制御された機能的活性は、事実上、酵素的、構造的、または制御的であり得る。いくつかの実施形態では、投与される組成物は、ポリペプチドが下方制御された細胞において存在するまたは上方制御された機能的活性を低下させる（例えば、相乗的に）1つ以上のポリペプチドの下方制御を誘導する。いくつかの実施形態では、投与される組成物は、所定の機能的活性の上方制御及び他の機能的活性の下方制御を誘導する。

#### 【0454】

任意の実施形態のうちの一部では、標的化脂質粒子組成物（例えば、ミトコンドリアまたはDNAを含むもの）は、標的細胞に対する効果を媒介し、効果は、少なくとも1、2、3、4、5、6、もしくは7日間、2、3、もしくは4週間、または1、2、3、6、もしくは12ヶ月間持続する。いくつかの実施形態（例えば、標的化脂質粒子組成物が外因性タンパク質を含む）では、効果は、1、2、3、4、5、6、もしくは7日、2、3、もしくは4週、または1、2、3、6、もしくは12ヶ月未満の間持続する。

10

20

30

40

50

## 【0455】

任意の実施形態のうちの一部では、本明細書に記載の標的化脂質粒子組成物は、細胞または組織、例えば、ヒト細胞または組織に *ex vivo* で送達される。実施形態では、組成物は、細胞または組織の機能を *ex vivo* で改善し、例えば、細胞生存性、呼吸、または他の機能（例えば、本明細書に記載の別の機能を改善する）。

## 【0456】

いくつかの実施形態では、組成物は、傷害状態（例えば、外傷、疾患、低酸素症、虚血または他の損傷からのもの）である *ex vivo* 組織に送達される。

## 【0457】

いくつかの実施形態では、組成物は、*ex vivo* 移植（例えば、組織外植片または移植のための組織、例えば、ヒト静脈、筋骨格移植片、例えば、骨もしくは腱、角膜、皮膚、心臓弁、神経；または単離もしくは培養された器官、例えば、ヒトに移植される器官、例えば、ヒト心臓、肝臓、肺、腎臓、膵臓、腸、胸腺、眼）に送達される。いくつかの実施形態では、組成物は、移植の前、間及び/または後に組織に送達される。

## 【0458】

いくつかの実施形態では、組成物は、細胞、例えば、細胞調製物と共に送達、投与または接触される。いくつかの実施形態では、細胞調製物は、細胞療法調製物（ヒト対象への投与のために意図された細胞調製物）であり得る。実施形態では、細胞調製物は、キメラ抗原受容体（CAR）を発現する、例えば、組換えCARを発現する細胞を含む。CARを発現する細胞は、例えば、T細胞、ナチュラルキラー（NK）細胞、細胞傷害性Tリンパ球（CTL）、制御性T細胞であり得る。実施形態では、細胞調製物は、神経幹細胞調製物である。実施形態では、細胞調製物は、間葉幹細胞（MSC）調製物である。実施形態では、細胞調製物は、造血幹細胞（HSC）調製物である。実施形態では、細胞調製物は、膵島細胞調製物である。

## 【0459】

いくつかの実施形態では、本明細書に記載の標的化脂質粒子組成物は、対象、例えば、哺乳動物、例えば、ヒトに投与され得る。そのような実施形態では、対象は、特定の疾患または病態（例えば、本明細書に記載の疾患または病態）のリスクがある、その症状を有する、またはその診断がされている、またはそれを有するものとして特定されている場合がある。

## 【0460】

いくつかの実施形態では、標的化脂質粒子の源は、標的化脂質粒子組成物が投与されるものと同じ対象からのものである。他の実施形態では、それらは異なる。いくつかの実施形態では、標的化脂質粒子及びレシピエント組織の源は、自己移植（同じ対象からのもの）または異種（異なる対象からのもの）であり得る。いくつかの実施形態では、本明細書に記載の標的化脂質粒子組成物のためのドナー組織は、レシピエント組織とは異なる組織タイプであり得る。いくつかの実施形態では、ドナー組織は、筋肉組織であり得、レシピエント組織は、結合組織（例えば、脂肪組織）であり得る。他の実施形態では、ドナー組織及びレシピエント組織は、同じまたは異なるタイプのものであり得るが、異なる器官系からのものであり得る。

## 【0461】

いくつかの実施形態では、本明細書に記載の標的化脂質粒子組成物は、がん、自己免疫疾患、感染性疾患、代謝疾患、神経変性疾患、または遺伝子疾患（例えば、酵素欠損症）を有する対象に投与され得る。いくつかの実施形態では、対象は、再生を必要とする。

## 【0462】

いくつかの実施形態では、標的化脂質粒子は、膜融合を阻害するタンパク質の阻害剤と共に投与される。例えば、サプレッシン（Suppressyn）は、細胞-細胞融合を阻害するヒトタンパク質である（Sugimoto et al., "A novel human endogenous retroviral protein inhibits cell-cell fusion" Scientific Reports 3: 40

1462 (DOI: 10.1038/srep01462))。いくつかの実施形態では、標的化脂質粒子は、シプレシン (sypresyn) の阻害剤、例えば、siRNA または阻害性抗体と共投与される。

【0463】

V. 例示的な実施形態

提供される実施形態には以下のものがある：

【0464】

1. 標的化脂質粒子であって、

(a) 内腔を包囲する脂質二重層と、

(b) ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、

(c) (i) ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分及び (ii) シングルドメイン抗体 (sdAb) 可変ドメインを含む、標的化エンベロープタンパク質であって、前記sdAb可変ドメインは、前記Gタンパク質またはその生物学的に活性な部分のC末端に結合されている、前記標的化エンベロープタンパク質と

を含み、前記Fタンパク質分子またはその生物学的に活性な部分及び前記標的化エンベロープタンパク質は、前記脂質二重層に埋め込まれている、前記標的化脂質粒子。

【0465】

2. 前記シングルドメイン抗体は、リンカーを介して前記Gタンパク質に結合されている、実施形態1に記載の標的化脂質粒子。

【0466】

3. 前記リンカーは、ペプチドリinkerである、実施形態2に記載の標的化脂質粒子。

【0467】

4. 標的化脂質粒子であって、

(a) 内腔を包囲する脂質二重層と、

(b) ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、

(c) ペプチドリinkerを介してシングルドメイン抗体 (sdAb) 可変ドメインに結合したヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分を含む、標的化エンベロープタンパク質であって、前記シングルドメイン抗体は、標的細胞の細胞表面分子に結合する、前記標的化エンベロープタンパク質と

を含み、前記Fタンパク質分子またはその生物学的に活性な部分及び前記標的化エンベロープタンパク質は、前記脂質二重層に埋め込まれている、前記標的化脂質粒子。

【0468】

5. 前記Fタンパク質分子またはその生物学的に活性な部分のN末端は、脂質二重層の外側に露出されている、実施形態1~4のいずれかに記載の標的化脂質粒子。

【0469】

6. 前記Gタンパク質のC末端は、前記脂質二重層の外側に露出されている、実施形態1~5のいずれかに記載の標的化脂質粒子。

【0470】

7. 前記シングルドメイン抗体は、標的細胞上に存在する細胞表面分子と結合する、実施形態1~6のいずれかに記載の標的化脂質粒子。

【0471】

8. 前記細胞表面分子は、タンパク質、グリカン、脂質または低分子量分子である、実施形態7に記載の標的化脂質粒子。

【0472】

9. 前記標的細胞は、腫瘍浸潤リンパ球、T細胞、新生物もしくは腫瘍細胞、ウイルス感染細胞、幹細胞、中枢神経系 (CNS) 細胞、造血幹細胞 (HSC)、肝臓細胞または完全に分化した細胞からなる群から選択される、実施形態7に記載の標的化脂質粒子。

【0473】

10. 前記標的細胞は、CD3+ T細胞、CD4+ T細胞、CD8+ T細胞、肝

10

20

30

40

50

細胞、造血幹細胞、CD34+ 造血幹細胞、CD105+ 造血幹細胞、CD117+ 造血幹細胞、CD105+ 内皮細胞、B細胞、CD20+ B細胞、CD19+ B細胞、がん細胞、CD133+ がん細胞、EpCAM+ がん細胞、CD19+ がん細胞、Her2/Neu+ がん細胞、GluA2+ ニューロン、GluA4+ ニューロン、NKGD2+ ナチュラルキラー細胞、SLC1A3+ 星状細胞、SLC7A10+ 脂肪細胞、またはCD30+ 肺上皮細胞からなる群から選択される、実施形態9に記載の標的化脂質粒子。

【0474】

11. 前記シングルドメイン抗体は、標的細胞上に存在する抗原またはその一部と結合する、先行実施形態のいずれかに記載の標的化脂質粒子。

10

【0475】

12. 前記ペプチドリンカーは、最大で65アミノ酸の長さを含む、実施形態3~11のいずれかに記載の標的化脂質粒子。

【0476】

13. 前記ペプチドリンカーは、約2~65アミノ酸、2~60アミノ酸、2~56アミノ酸、2~52アミノ酸、2~48アミノ酸、2~44アミノ酸、2~40アミノ酸、2~36アミノ酸、2~32アミノ酸、2~28アミノ酸、2~24アミノ酸、2~20アミノ酸、2~18アミノ酸、2~14アミノ酸、2~12アミノ酸、2~10アミノ酸、2~8アミノ酸、2~6アミノ酸、6~65アミノ酸、6~60アミノ酸、6~56アミノ酸、6~52アミノ酸、6~48アミノ酸、6~44アミノ酸、6~40アミノ酸、6~36アミノ酸、6~32アミノ酸、6~28アミノ酸、6~24アミノ酸、6~20アミノ酸、6~18アミノ酸、6~14アミノ酸、6~12アミノ酸、6~10アミノ酸、6~8アミノ酸、8~65アミノ酸、8~60アミノ酸、8~56アミノ酸、8~52アミノ酸、8~48アミノ酸、8~44アミノ酸、8~40アミノ酸、8~36アミノ酸、8~32アミノ酸、8~28アミノ酸、8~24アミノ酸、8~20アミノ酸、8~18アミノ酸、8~14アミノ酸、8~12アミノ酸、8~10アミノ酸、10~65アミノ酸、10~60アミノ酸、10~56アミノ酸、10~52アミノ酸、10~48アミノ酸、10~44アミノ酸、10~40アミノ酸、10~36アミノ酸、10~32アミノ酸、10~28アミノ酸、10~24アミノ酸、10~20アミノ酸、10~18アミノ酸、10~14アミノ酸、10~12アミノ酸、12~65アミノ酸、12~60アミノ酸、12~56アミノ酸、12~52アミノ酸、12~48アミノ酸、12~44アミノ酸、12~40アミノ酸、12~36アミノ酸、12~32アミノ酸、12~28アミノ酸、12~24アミノ酸、12~20アミノ酸、12~18アミノ酸、12~14アミノ酸、14~65アミノ酸、14~60アミノ酸、14~56アミノ酸、14~52アミノ酸、14~48アミノ酸、14~44アミノ酸、14~40アミノ酸、14~36アミノ酸、14~32アミノ酸、14~28アミノ酸、14~24アミノ酸、14~20アミノ酸、14~18アミノ酸、18~65アミノ酸、18~60アミノ酸、18~56アミノ酸、18~52アミノ酸、18~48アミノ酸、18~44アミノ酸、18~40アミノ酸、18~36アミノ酸、18~32アミノ酸、18~28アミノ酸、18~24アミノ酸、18~20アミノ酸、20~65アミノ酸、20~60アミノ酸、20~56アミノ酸、20~52アミノ酸、20~48アミノ酸、20~44アミノ酸、20~40アミノ酸、20~36アミノ酸、20~32アミノ酸、20~28アミノ酸、20~26アミノ酸、20~24アミノ酸、24~65アミノ酸、24~60アミノ酸、24~56アミノ酸、24~52アミノ酸、24~48アミノ酸、24~44アミノ酸、24~40アミノ酸、24~36アミノ酸、24~32アミノ酸、24~30アミノ酸、24~28アミノ酸、28~65アミノ酸、28~60アミノ酸、28~56アミノ酸、28~52アミノ酸、28~48アミノ酸、28~44アミノ酸、28~40アミノ酸、28~36アミノ酸、28~34アミノ酸、28~32アミノ酸、32~65アミノ酸、32~60アミノ酸、32~56アミノ酸、32~52アミノ酸、32~48アミノ酸、32~44アミノ酸、32~40アミノ酸、32~38アミノ酸、32~36アミノ酸、36~65アミノ酸

20

30

40

50



ノ酸、36～60アミノ酸、36～56アミノ酸、36～52アミノ酸、36～48アミノ酸、36～44アミノ酸、36～40アミノ酸、40～65アミノ酸、40～60アミノ酸、40～56アミノ酸、40～52アミノ酸、40～48アミノ酸、40～44アミノ酸、44～65アミノ酸、44～60アミノ酸、44～56アミノ酸、44～52アミノ酸、44～48アミノ酸、48～65アミノ酸、48～60アミノ酸、48～56アミノ酸、48～52アミノ酸、50～65アミノ酸、50～60アミノ酸、50～56アミノ酸、50～52アミノ酸、54～65アミノ酸、54～60アミノ酸、54～56アミノ酸、58～65アミノ酸、58～60アミノ酸、または60～65アミノ酸を含む、実施形態3～11のいずれかに記載の標的化脂質粒子。

【0477】

14. 前記ペプチドリンカーは、長さが3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50、51、52、53、54、55、56、57、58、59、60、61、62、63、64または65アミノ酸であるポリペプチドを含む、実施形態3～11のいずれかに記載の標的化脂質粒子。

【0478】

15. 前記ペプチドリンカーは、GS、GGS、GGGS(配列番号43)、GGGG(配列番号41)またはそれらの組み合わせを含むフレキシブルリンカーである、

【0479】

16. 前記ペプチドリンカーは、nが1～10である(GGS)<sub>n</sub>を含む、実施形態3～15のいずれかに記載の標的化脂質粒子。

【0480】

17. 前記ペプチドリンカーは、nが1～10である(GGGGS)<sub>n</sub>(配列番号42)を含む、実施形態3～15のいずれかに記載の標的化脂質粒子。

【0481】

18. 前記ペプチドリンカーは、nが1～6である(GGGGG)<sub>n</sub>(配列番号27)を含む、実施形態3～15のいずれかに記載の標的化脂質粒子。

【0482】

19. 前記Gタンパク質またはその生物学的に活性な部分は、野生型ニパウイルスG(NiV-G)タンパク質またはヘンドラウイルスGタンパク質である、実施形態1～18のいずれかに記載の標的化脂質粒子。

【0483】

20. 前記Gタンパク質またはその生物学的に活性な部分は、野生型NiV-Gタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、実施形態1～19のいずれかに記載の標的化脂質粒子。

【0484】

21. 前記変異体NiV-Gタンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、配列番号9、配列番号28または配列番号44と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、少なくとも84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む、実施形態20に記載

10

20

30

40

50

の標的化脂質粒子。

【0485】

22. 前記 NiV - G タンパク質は、

切断型でありかつ前記野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの最大で 40 個の連続アミノ酸残基を欠く生物学的に活性な部分

である、実施形態 21 に記載の標的化脂質粒子。

【0486】

23. 前記 NiV - G タンパク質は、野生型 NiV - G の N 末端で切断された生物学的に活性な部分であり、配列番号 10 ~ 15、35 ~ 40 もしくは 45 ~ 50 のいずれかに示される配列、または配列番号 10 ~ 15、35 ~ 40 もしくは 45 ~ 50 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、少なくとも 84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 1 ~ 18 のいずれかに記載の標的化脂質粒子。

【0487】

24. 前記 NiV - G タンパク質は、前記野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 5 アミノ酸切断を有する、実施形態 21 ~ 23 のいずれかに記載の標的化脂質粒子。

【0488】

25. 前記 NiV - G タンパク質は、配列番号 10 に示されるアミノ酸配列、または配列番号 10 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 24 に記載の標的化脂質粒子。

【0489】

26. 前記 NiV - G タンパク質は、配列番号 35 に示されるアミノ酸配列、または配列番号 35 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 24 に記載の標的化脂質粒子。

【0490】

10

20

30

40

50

27. 前記 NiV - G タンパク質は、配列番号 45 に示されるアミノ酸配列、または配列番号 45 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 24 に記載の標的化脂質粒子。 10

【0491】

28. 前記 NiV - G タンパク質は、前記野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 10 アミノ酸切断を有する、実施形態 21 ~ 23 のいずれかに記載の標的化脂質粒子。

【0492】

29. 前記 NiV - G タンパク質は、配列番号 11 に示されるアミノ酸配列、または配列番号 11 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 28 に記載の標的化脂質粒子。 20

【0493】

30. 前記 NiV - G タンパク質は、配列番号 36 に示されるアミノ酸配列、または配列番号 36 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 28 に記載の標的化脂質粒子。 30

【0494】

31. 前記 NiV - G タンパク質は、配列番号 46 に示されるアミノ酸配列、または配列番号 46 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列 40 50

を有する、実施形態 28 に記載の標的化脂質粒子。

【0495】

32. 前記 NiV - G タンパク質は、前記野生型 NiV - G タンパク質 (配列番号 9、配列番号 28 または配列番号 44) の N 末端におけるまたはその近くの 15 アミノ酸切断を有する、実施形態 21 ~ 23 のいずれかに記載の標的化脂質粒子。

【0496】

33. 前記 NiV - G タンパク質は、配列番号 12 に示されるアミノ酸配列、または配列番号 12 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 32 に記載の標的化脂質粒子。

10

【0497】

34. 前記 NiV - G タンパク質は、配列番号 37 に示されるアミノ酸配列、または配列番号 37 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 32 に記載の標的化脂質粒子。

20

【0498】

35. 前記 NiV - G タンパク質は、配列番号 47 に示されるアミノ酸配列、または配列番号 47 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 32 に記載の標的化脂質粒子。

30

40

【0499】

36. 前記 NiV - G タンパク質は、前記野生型 NiV - G タンパク質 (配列番号 9、配列番号 28 または配列番号 44) の N 末端におけるまたはその近くの 20 アミノ酸切断を有する、実施形態 21 ~ 23 のいずれかに記載の標的化脂質粒子。

【0500】

37. 前記 NiV - G タンパク質は、配列番号 13 に示されるアミノ酸配列、または配列番号 13 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、

50

または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 36 に記載の標的化脂質粒子。

【0501】

38. 前記 NiV - G タンパク質は、配列番号 38 に示されるアミノ酸配列、または配列番号 38 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 36 に記載の標的化脂質粒子。

10

【0502】

39. 前記 NiV - G タンパク質は、配列番号 48 に示されるアミノ酸配列、または配列番号 48 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 36 に記載の標的化脂質粒子。

20

30

【0503】

40. 前記 NiV - G タンパク質は、前記野生型 NiV - G タンパク質 (配列番号 9、配列番号 28 または配列番号 44) の N 末端におけるまたはその近くの 25 アミノ酸切断を有する、実施形態 21 ~ 23 のいずれかに記載の標的化脂質粒子。

【0504】

41. 前記 NiV - G タンパク質は、配列番号 14 に示されるアミノ酸配列、または配列番号 14 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 40 に記載の標的化脂質粒子。

40

【0505】

42. 前記 NiV - G タンパク質は、配列番号 39 に示されるアミノ酸配列、または配列番号 39 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、

50

少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する、実施形態40に記載の標的化脂質粒子。

【0506】

43. 前記NiV-Gタンパク質は、配列番号49に示されるアミノ酸配列、または配列番号49と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する、実施形態40に記載の標的化脂質粒子。

【0507】

44. 前記NiV-Gタンパク質は、前記野生型NiV-Gタンパク質(配列番号9、配列番号28または配列番号44)のN末端におけるまたはその近くの30アミノ酸切断を有する、実施形態21~23のいずれかに記載の標的化脂質粒子。

【0508】

45. 前記NiV-Gタンパク質は、配列番号15に示されるアミノ酸配列、または配列番号15と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する、実施形態44に記載の標的化脂質粒子。

【0509】

46. 前記NiV-Gタンパク質は、配列番号40に示されるアミノ酸配列、または配列番号40と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する、実施形態44に記載の標的化脂質粒子。

【0510】

10

20

30

40

50

47. 前記 NiV - G タンパク質は、配列番号 50 に示されるアミノ酸配列、または配列番号 50 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を有する、実施形態 44 に記載の標的化脂質粒子。 10

【0511】

48. 前記 NiV - G タンパク質は、前記野生型 NiV - G タンパク質（配列番号 9、配列番号 28 または配列番号 44）の N 末端におけるまたはその近くの 34 アミノ酸切断を有する、実施形態 21 ~ 23 のいずれかに記載の標的化脂質粒子。

【0512】

49. 前記 NiV - G タンパク質は、配列番号 22 に示されるアミノ酸配列、または配列番号 22 と 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を有する、実施形態 48 に記載の標的化脂質粒子。 20

【0513】

50. 前記 NiV - G タンパク質は、配列番号 53 に示されるアミノ酸配列、または配列番号 53 と 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列を有する、実施形態 48 に記載の標的化脂質粒子。 30

【0514】

51. 前記 G タンパク質またはその生物学的に活性な部分は、エフリン B2 またはエフリン B3 に対する減少した結合を示す変異体 NiV - G タンパク質である、実施形態 1 ~ 48 のいずれかに記載の標的化脂質粒子。 40

【0515】

52. 前記変異体 NiV - G タンパク質は、配列番号 28 に示される付番を参照して E501A、W504A、Q530A 及び E533A からなる群から選択されるアミノ酸置換に対応する 1 つ以上のアミノ酸置換を含む、実施形態 51 に記載の標的化脂質粒子。

【0516】

53. 前記変異体 NiV - G タンパク質は、配列番号 16 に示されるアミノ酸配列、ま 50

たは配列番号 16 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 51 または実施形態 52 に記載の標的化脂質粒子。

10

## 【0517】

54. 前記変異体 NiV - G タンパク質は、配列番号 51 に示されるアミノ酸配列、または配列番号 51 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を有する、実施形態 51 または実施形態 52 に記載の標的化脂質粒子。

20

## 【0518】

55. 前記 F タンパク質またはその生物学的に活性な部分は、野生型ニパウイルス F (NiV - F) タンパク質もしくはヘンドラウイルス F タンパク質であるか、またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、実施形態 1 ~ 54 のいずれかに記載の標的化脂質粒子。

## 【0519】

56. 前記 F タンパク質またはその生物学的に活性な部分は、野生型 NiV - F タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分である、実施形態 1 ~ 55 のいずれかに記載の標的化脂質粒子。

30

## 【0520】

57. 前記 NiV - F タンパク質またはその機能的に活性なバリエーションもしくは生物学的に活性な部分は、配列番号 2 に示されるアミノ酸配列、または配列番号 2 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 1 ~ 56 のいずれかに記載の標的化脂質粒子。

40

## 【0521】

58. 前記 NiV - F タンパク質は、前記野生型 NiV - F タンパク質 (配列番号 2) の C 末端におけるまたはその近くの 20 アミノ酸切断を有するその生物学的に活性な部分である、実施形態 1 ~ 57 のいずれかに記載の標的化脂質粒子。

## 【0522】

59. 前記 NiV - F タンパク質は、配列番号 5 と 80% もしくは約 80%、少なくとも

50



も 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列を有する、実施形態 5 8 に記載の標的化脂質粒子。

## 【 0 5 2 3 】

10

6 0 . 前記 N i V - F タンパク質は、

i ) 前記野生型 N i V - F タンパク質 ( 配列番号 2 ) の C 末端におけるまたはその近くの 2 0 アミノ酸切断、及び

i i ) N 結合型グリコシル化部位における点変異

を含むその生物学的に活性な部分である、実施形態 1 ~ 5 7 のいずれかに記載の標的化脂質粒子。

## 【 0 5 2 4 】

6 1 . 前記 N i V - F タンパク質は、配列番号 7 と 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有するアミノ酸配列を有する、実施形態 6 0 に記載の標的化脂質粒子。

20

## 【 0 5 2 5 】

6 2 . 前記 N i V - F タンパク質は、前記野生型 N i V - F タンパク質 ( 配列番号 2 ) の C 末端におけるまたはその近くの 2 2 アミノ酸切断を有するその生物学的に活性な部分である、実施形態 1 ~ 5 7 のいずれかに記載の標的化脂質粒子。

30

## 【 0 5 2 6 】

6 3 . 前記 N i V - F タンパク質は、配列番号 8 と 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2 %、少なくとも 9 3 % もしくは約 9 3 %、少なくとも 9 4 % もしくは約 9 4 %、少なくとも 9 5 % もしくは約 9 5 %、9 6 % もしくは約 9 6 %、少なくとも 9 7 % もしくは約 9 7 %、少なくとも 9 8 % もしくは約 9 8 %、または少なくとも 9 9 % もしくは約 9 9 % の配列同一性を有する配列をコードするヌクレオチドの配列によってコードされるアミノ酸配列を有する、実施形態 6 2 に記載の標的化脂質粒子

40

## 【 0 5 2 7 】

6 4 . 前記 N i V - F タンパク質は、配列番号 2 3 と 8 0 % もしくは約 8 0 %、少なくとも 8 1 % もしくは約 8 1 %、少なくとも 8 2 % もしくは約 8 2 %、少なくとも 8 3 % もしくは約 8 3 %、8 4 % もしくは約 8 4 %、少なくとも 8 5 % もしくは約 8 5 %、少なくとも 8 6 % もしくは約 8 6 %、または少なくとも 8 7 % もしくは約 8 7 %、少なくとも 8 8 % もしくは約 8 8 %、または少なくとも 8 9 % もしくは約 8 9 %、少なくとも 9 0 % もしくは約 9 0 %、少なくとも 9 1 % もしくは約 9 1 %、少なくとも 9 2 % もしくは約 9 2

50

%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を有する、実施形態63に記載の標的化脂質粒子。

【0528】

65. 前記F-タンパク質またはその生物学的に活性な部分は、F1サブユニットまたはその融合部分を含む、実施形態1~57のいずれかに記載の標的化脂質粒子。

【0529】

66. 前記F1サブユニットは、F0前駆体のタンパク質分解的に切断された部分である、実施形態65に記載の標的化脂質粒子。

10

【0530】

67. 前記F1サブユニットは、配列番号4に示される配列、または配列番号4と80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む、実施形態66に記載の標的化脂質粒子。

20

【0531】

68. 前記脂質二重層は、レトロウイルスまたはレトロウイルス様粒子を生成するために使用される宿主細胞の膜に由来する、実施形態1~67のいずれかに記載の標的化脂質粒子。

【0532】

69. 前記脂質二重層は、ウイルスエンベロープであるか、またはそれを含む、実施形態1~60のいずれかに記載の標的化脂質粒子。

【0533】

70. 前記レトロウイルス様粒子は、複製欠損性である、実施形態68に記載の標的化脂質粒子。

30

【0534】

71. 前記Fタンパク質分子及び前記Gタンパク質以外の1つ以上のウイルス成分を含む、実施形態1~70のいずれかに記載の標的化脂質粒子。

【0535】

72. 前記1つ以上のウイルス成分は、レトロウイルスからのものである、実施形態71に記載の標的化脂質粒子。

【0536】

73. 前記レトロウイルスは、レンチウイルスである、実施形態72に記載の標的化脂質粒子。

40

【0537】

74. 前記1つ以上のウイルス成分は、Gag、Pol、Rev及びTatのうち1つ以上から選択されるウイルスパッケージングタンパク質を含む、実施形態71~73のいずれかに記載の標的化脂質粒子。

【0538】

75. 前記1つ以上のウイルス成分は、以下の核酸配列：5'LTR（例えば、U5を含み、機能性U3ドメインを欠く）、プサイパッケージング要素（プサイ）、セントラルポリプリン配列（cPPT）/セントラル終結配列（CTS）（例えば、DNAフラップ）、ポリAテール配列、転写後制御要素（例えば、WPRE）、Rev反応要素（RRE

50

)、及び3'LTR(例えば、U5を含み、機能性U3を欠く)のうちの1つ以上(例えば、すべて)を含む、実施形態71~74のいずれかに記載の標的化脂質粒子。

【0539】

76. 外因性物質をさらに含む、実施形態1~75のいずれかに記載の標的化脂質粒子。

【0540】

77. 前記外因性物質は、内腔に存在する、実施形態76に記載の標的化脂質粒子。

【0541】

78. 前記外因性物質は、タンパク質または核酸であり、任意に前記核酸は、DNAまたはRNAである、実施形態77に記載の標的化脂質粒子。

10

【0542】

79. 前記外因性物質は、治療剤または診断剤をコードする、実施形態76~78のいずれかに記載の標的化脂質粒子。

【0543】

80. 前記宿主細胞は、CHO細胞、BHK細胞、MDCK細胞、C3H10T1/2細胞、FLY細胞、Psi-2細胞、BOSC23細胞、PA317細胞、WEHI細胞、COS細胞、BSC1細胞、BSC40細胞、BMT10細胞、VERO細胞、W138細胞、MRC5細胞、A549細胞、HT1080細胞、293細胞、293T細胞、B-50細胞、3T3細胞、NIH3T3細胞、HepG2細胞、Saos-2細胞、Hu h7細胞、HeLa細胞、W163細胞、211細胞、及び211A細胞からなる群から選択される、実施形態68~79のいずれかに記載の標的化脂質粒子。

20

【0544】

81. 前記宿主細胞は、293T細胞を含む、実施形態68~80のいずれかに記載の標的化脂質粒子。

【0545】

82. (i)ヘニパウイルスエンベロープ結合糖タンパク質G(Gタンパク質)またはその生物学的に活性な部分と(ii)シングルドメイン抗体(sdAb)可変ドメインとをコードする核酸配列を含む、ポリヌクレオチドであって、前記sdAb可変ドメインは、前記Gタンパク質またはその生物学的に活性な部分のC末端に結合されている、前記ポリヌクレオチド。

30

【0546】

83. (iii)ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸配列をさらに含む、実施形態82に記載のポリヌクレオチド。

【0547】

84. 前記核酸の発現をコントロールするために機能可能に連結された少なくとも1つのプロモーターをさらに含む、実施形態82または実施形態83に記載のポリヌクレオチド。

【0548】

85. 前記プロモーターは、構成的プロモーターである、実施形態83~84のいずれかに記載のポリヌクレオチド。

40

【0549】

86. 前記プロモーターは、誘導性プロモーターである、実施形態83~85のいずれかに記載のポリヌクレオチド。

【0550】

87. 前記sdAb可変ドメインは、コードされるペプチドリンカーを介して前記Gタンパク質に結合されている、実施形態82~86のいずれかに記載のポリヌクレオチド。

【0551】

88. 前記コードされるペプチドリンカーは、最大で65アミノ酸の長さを含む、実施形態86~87のいずれかに記載のポリヌクレオチド。

【0552】

50

89. 前記コードされるペプチドリンカーは、約2～65アミノ酸、2～60アミノ酸、2～56アミノ酸、2～52アミノ酸、2～48アミノ酸、2～44アミノ酸、2～40アミノ酸、2～36アミノ酸、2～32アミノ酸、2～28アミノ酸、2～24アミノ酸、2～20アミノ酸、2～18アミノ酸、2～14アミノ酸、2～12アミノ酸、2～10アミノ酸、2～8アミノ酸、2～6アミノ酸、6～65アミノ酸、6～60アミノ酸、6～56アミノ酸、6～52アミノ酸、6～48アミノ酸、6～44アミノ酸、6～40アミノ酸、6～36アミノ酸、6～32アミノ酸、6～28アミノ酸、6～24アミノ酸、6～20アミノ酸、6～18アミノ酸、6～14アミノ酸、6～12アミノ酸、6～10アミノ酸、6～8アミノ酸、8～65アミノ酸、8～60アミノ酸、8～56アミノ酸、8～52アミノ酸、8～48アミノ酸、8～44アミノ酸、8～40アミノ酸、8～36アミノ酸、8～32アミノ酸、8～28アミノ酸、8～24アミノ酸、8～20アミノ酸、8～18アミノ酸、8～14アミノ酸、8～12アミノ酸、8～10アミノ酸、10～65アミノ酸、10～60アミノ酸、10～56アミノ酸、10～52アミノ酸、10～48アミノ酸、10～44アミノ酸、10～40アミノ酸、10～36アミノ酸、10～32アミノ酸、10～28アミノ酸、10～24アミノ酸、10～20アミノ酸、10～18アミノ酸、10～14アミノ酸、10～12アミノ酸、12～65アミノ酸、12～60アミノ酸、12～56アミノ酸、12～52アミノ酸、12～48アミノ酸、12～44アミノ酸、12～40アミノ酸、12～36アミノ酸、12～32アミノ酸、12～28アミノ酸、12～24アミノ酸、12～20アミノ酸、12～18アミノ酸、12～14アミノ酸、14～65アミノ酸、14～60アミノ酸、14～56アミノ酸、14～52アミノ酸、14～48アミノ酸、14～44アミノ酸、14～40アミノ酸、14～36アミノ酸、14～32アミノ酸、14～28アミノ酸、14～24アミノ酸、14～20アミノ酸、14～18アミノ酸、18～65アミノ酸、18～60アミノ酸、18～56アミノ酸、18～52アミノ酸、18～48アミノ酸、18～44アミノ酸、18～40アミノ酸、18～36アミノ酸、18～32アミノ酸、18～28アミノ酸、18～24アミノ酸、18～20アミノ酸、20～65アミノ酸、20～60アミノ酸、20～56アミノ酸、20～52アミノ酸、20～48アミノ酸、20～44アミノ酸、20～40アミノ酸、20～36アミノ酸、20～32アミノ酸、20～28アミノ酸、20～24アミノ酸、24～65アミノ酸、24～60アミノ酸、24～56アミノ酸、24～52アミノ酸、24～48アミノ酸、24～44アミノ酸、24～40アミノ酸、24～36アミノ酸、24～32アミノ酸、24～30アミノ酸、24～28アミノ酸、28～65アミノ酸、28～60アミノ酸、28～56アミノ酸、28～52アミノ酸、28～48アミノ酸、28～44アミノ酸、28～40アミノ酸、28～36アミノ酸、28～34アミノ酸、28～32アミノ酸、32～65アミノ酸、32～60アミノ酸、32～56アミノ酸、32～52アミノ酸、32～48アミノ酸、32～44アミノ酸、32～40アミノ酸、32～38アミノ酸、32～36アミノ酸、36～65アミノ酸、36～60アミノ酸、36～56アミノ酸、36～52アミノ酸、36～48アミノ酸、36～44アミノ酸、36～40アミノ酸、40～65アミノ酸、40～60アミノ酸、40～56アミノ酸、40～52アミノ酸、40～48アミノ酸、40～44アミノ酸、44～65アミノ酸、44～60アミノ酸、44～56アミノ酸、44～52アミノ酸、44～48アミノ酸、44～44アミノ酸、48～65アミノ酸、48～60アミノ酸、48～56アミノ酸、48～52アミノ酸、50～65アミノ酸、50～60アミノ酸、50～56アミノ酸、50～52アミノ酸、54～65アミノ酸、54～60アミノ酸、54～56アミノ酸、58～65アミノ酸、58～60アミノ酸、または60～65アミノ酸を含む、実施形態86～87のいずれかに記載のポリヌクレオチド。

10

20

30

40

50

【0553】

90. 前記コードされるペプチドリンカーは、長さが3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、31、32、33、34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50

0、51、52、53、54、55、56、57、58、59、60、61、62、63、64または65アミノ酸であるポリペプチドを含む、実施形態86~87のいずれかに記載のポリヌクレオチド。

【0554】

91. 前記コードされるペプチドリンカーは、GS、GGS、GGGS (配列番号43)、GGGG (配列番号41) 及びそれらの組み合わせを含む、実施形態86~87のいずれかに記載のポリヌクレオチド。

【0555】

92. 前記コードされるペプチドリンカーは、nが1~10である(GGS)<sub>n</sub>を含む、実施形態86~87のいずれかに記載のポリヌクレオチド。

10

【0556】

93. 前記コードされるペプチドリンカーは、nが1~10である(GGGGS)<sub>n</sub> (配列番号42) を含む、実施形態86~87のいずれかに記載のポリヌクレオチド。

【0557】

94. 前記コードされるペプチドリンカーは、nが1~4である(GGGGG)<sub>n</sub> (配列番号27) を含む、実施形態86~87のいずれかに記載のポリヌクレオチド。

【0558】

95. 前記Gタンパク質をコードする前記核酸配列は、野生型ニパウイルスG(NiV-G)タンパク質もしくはヘンドラウイルスGタンパク質であるか、またはネイティブ結合パートナーについて減少した結合を示すそのバリエーションである、実施形態86~87のいずれかに記載のポリヌクレオチド。

20

【0559】

96. 前記Gタンパク質をコードする前記核酸配列は、野生型NiV-Gタンパク質である、実施形態82~95のいずれかに記載のポリヌクレオチド。

【0560】

97. 前記Gタンパク質をコードする前記核酸配列は、エフリンB2またはエフリンB3に対する減少した結合を示す変異体NiV-Gタンパク質である、実施形態82~95のいずれかに記載のポリヌクレオチド。

【0561】

98. 前記変異体NiV-Gタンパク質をコードする前記核酸配列は、配列番号9、配列番号28または配列番号44と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む、実施形態97に記載のポリヌクレオチド。

30

40

【0562】

99. 前記変異体NiV-Gタンパク質をコードする前記核酸配列は、配列番号10~15、35~40もしくは45~50のいずれかに示される配列、または配列番号10~15、35~40もしくは45~50と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも9

50

5%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む、実施形態82~95及び97のいずれかに記載のポリヌクレオチド。

【0563】

100. 前記変異体NiV-Gタンパク質をコードする前記核酸配列は、前記野生型NiV-Gタンパク質(配列番号9、配列番号28または配列番号44)のN末端におけるまたはその近くの5アミノ酸切断を含む、実施形態97~99のいずれかに記載のポリヌクレオチド。

【0564】

101. 前記変異体NiV-Gタンパク質をコードする前記核酸配列は、配列番号10に示されるアミノ酸配列、または配列番号10と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む、実施形態100に記載のポリヌクレオチド。

【0565】

102. 前記変異体NiV-Gタンパク質をコードする前記核酸配列は、配列番号35に示されるアミノ酸配列、または配列番号35と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む、実施形態100に記載のポリヌクレオチド。

【0566】

103. 前記変異体NiV-Gタンパク質をコードする前記核酸配列は、配列番号45に示されるアミノ酸配列、または配列番号45と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列を含む、実施形態100に記載のポリヌクレオチド。

【0567】

104. 前記変異体NiV-Gタンパク質をコードする前記核酸配列は、前記野生型NiV-Gタンパク質(配列番号9、配列番号28または配列番号44)のN末端におけるまたはその近くの10アミノ酸切断を含む、実施形態97~99のいずれかに記載のポリヌクレオチド。

10

20

30

40

50

## 【0568】

105. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 11 に示されるアミノ酸配列、または配列番号 11 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 104 に記載のポリヌクレオチド。

## 【0569】

106. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 36 に示されるアミノ酸配列、または配列番号 36 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 104 に記載のポリヌクレオチド。

## 【0570】

107. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 46 に示されるアミノ酸配列、または配列番号 46 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 104 に記載のポリヌクレオチド。

## 【0571】

108. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、前記野生型 NiV - G タンパク質 (配列番号 9、配列番号 28 または配列番号 44) の N 末端におけるまたはその近くの 15 アミノ酸切断を含む、実施形態 97 ~ 99 のいずれかに記載のポリヌクレオチド。

## 【0572】

109. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 12 に示されるアミノ酸配列、または配列番号 12 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少な

くとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 108 に記載のポリヌクレオチド。

【0573】

110. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 37 に示されるアミノ酸配列、または配列番号 37 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 108 に記載のポリヌクレオチド。

10

【0574】

111. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 47 に示されるアミノ酸配列、または配列番号 47 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 108 に記載のポリヌクレオチド。

20

【0575】

112. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、前記野生型 NiV - G タンパク質 (配列番号 9、配列番号 28 または配列番号 44) の N 末端におけるまたはその近くの 20 アミノ酸切断を含む、実施形態 97 ~ 99 のいずれかに記載のポリヌクレオチド。

30

【0576】

113. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 13 に示されるアミノ酸配列、または配列番号 13 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 112 に記載のポリヌクレオチド。

40

【0577】

114. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 38 に示されるアミノ酸配列、または配列番号 38 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくと

50



も 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 112 に記載のポリヌクレオチド。

【0578】

115. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 48 に示されるアミノ酸配列、または配列番号 48 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 112 に記載のポリヌクレオチド。

10

【0579】

116. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、前記野生型 NiV - G タンパク質 (配列番号 9、配列番号 28 または配列番号 44) の N 末端におけるまたはその近くの 25 アミノ酸切断を含む、実施形態 97 ~ 99 のいずれかに記載のポリヌクレオチド。

20

【0580】

117. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 14 に示されるアミノ酸配列、または配列番号 14 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 116 に記載のポリヌクレオチド。

30

【0581】

118. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 39 に示されるアミノ酸配列、または配列番号 39 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 116 に記載のポリヌクレオチド。

40

【0582】

119. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 49 に示されるアミノ酸配列、または配列番号 49 と少なくとも 80% もしくは約 80%、少

50

なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 116 に記載のポリヌクレオチド。

【0583】

120. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、前記野生型 NiV - G タンパク質 (配列番号 9、配列番号 28 または配列番号 44) の N 末端におけるまたはその近くの 30 アミノ酸切断を含む、実施形態 97 ~ 99 のいずれかに記載のポリヌクレオチド。

【0584】

121. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 15 に示されるアミノ酸配列、または配列番号 15 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 120 に記載のポリヌクレオチド。

【0585】

122. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 40 に示されるアミノ酸配列、または配列番号 40 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 120 に記載のポリヌクレオチド。

【0586】

123. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 50 に示されるアミノ酸配列、または配列番号 50 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列を含む、実施形態 120 に記載のポリヌクレオチド。

10

20

30

40

50

## 【0587】

124. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、  
 i) N 末端におけるまたはその近くの切断、及び  
 ii) E501A、W504A、Q530A 及び E533A からなる群から選択される  
 点変異  
 を含む、実施形態 97 ~ 99 のいずれかに記載のポリヌクレオチド。

## 【0588】

125. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 16  
 に示されるアミノ酸配列、または配列番号 16 と少なくとも 80% もしくは約 80%、少  
 なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83  
 % もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少  
 なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくと  
 も 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90  
 % もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約  
 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少な  
 くとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約  
 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99%  
 の配列同一性を有するアミノ酸配列を含む、実施形態 124 に記載のポリヌクレオチド。

10

## 【0589】

126. 前記変異体 NiV - G タンパク質をコードする前記核酸配列は、配列番号 51  
 に示されるアミノ酸配列、または配列番号 51 と少なくとも 80% もしくは約 80%、少  
 なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83  
 % もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少  
 なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくと  
 も 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90  
 % もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約  
 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少な  
 くとも 95% もしくは約 95%、96% もしくは約 96%、少なくとも 97% もしくは約  
 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99%  
 の配列同一性を有するアミノ酸配列を含む、実施形態 124 に記載のポリヌクレオチド。

20

30

## 【0590】

127. 実施形態 82 ~ 126 のいずれかに記載のポリヌクレオチドを含む、ベクター  
 。

## 【0591】

128. 前記ベクターは、哺乳動物ベクター、ウイルスベクターまたは人工染色体であ  
 り、任意に前記人工染色体は、細菌人工染色体 (BAC) である、実施形態 127 に記載  
 のベクター。

## 【0592】

129. 実施形態 82 ~ 126 のいずれかに記載のポリヌクレオチドまたは実施形態 1  
 27 もしくは実施形態 128 に記載のベクターを含む、細胞。

40

## 【0593】

130. ヘニバウイルス F タンパク質分子またはその生物学的に活性な部分と、ヘニバ  
 ウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な  
 部分及びシングルドメイン抗体 (sdAb) 可変ドメインを含む標的化エンベロープタン  
 パク質とを含む、標的化脂質粒子を作製する方法であって、

a) ヘニバウイルス F タンパク質分子またはその生物学的に活性な部分をコードする核  
 酸と、ヘニバウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物  
 学的に活性な部分及びシングルドメイン抗体 (sdAb) 可変ドメインを含む標的化エン  
 ベロープタンパク質をコードする核酸とを含む、細胞を提供すること、

b) 標的化脂質粒子の生成を可能とする条件下で前記細胞を培養すること、及び

50

c) 前記細胞から前記標的化脂質粒子を分離、濃縮、または精製し、それにより前記標的化脂質粒子を作製することを含む、前記方法。

## 【0594】

131. ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分と、ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分及びシングルドメイン抗体 (sdAb) 可変ドメインを含む標的化エンベロープタンパク質とを含む、標的化脂質粒子を作製する方法であって、

a) 実施形態82~126のいずれかに記載のポリヌクレオチドまたは実施形態127もしくは実施形態128に記載のベクターを含む細胞を提供すること、

b) 前記細胞にヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードするポリヌクレオチドを提供すること、

c) 標的化脂質粒子の生成を可能とする条件下で前記細胞を培養すること、及び

d) 前記細胞から前記標的化脂質粒子を分離、濃縮、または精製し、それにより前記標的化脂質粒子を作製することを含む、前記方法。

## 【0595】

132. 前記細胞は、哺乳動物細胞である、実施形態130または実施形態131に記載の方法。

## 【0596】

133. 前記細胞は、プロデューサー細胞であり、前記標的化脂質粒子は、ウイルス粒子またはウイルス様粒子、任意にレトロウイルス粒子またはレトロウイルス様粒子、任意にレンチウイルス粒子またはレンチウイルス様粒子である、実施形態130~131のいずれかに記載の方法。

## 【0597】

134. (i) ウイルス核酸 (複数可) と、(ii) ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分をコードする核酸と、(iii) ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分及びシングルドメイン抗体 (sdAb) 可変ドメインを含む標的化エンベロープタンパク質をコードする核酸とを含む、プロデューサー細胞であって、任意に前記ウイルス核酸 (複数可) は、レンチウイルス核酸である、前記プロデューサー細胞。

## 【0598】

135. 前記ウイルス核酸 (複数可) は、ウイルス複製に関与する1つ以上の遺伝子を欠く、実施形態134に記載のプロデューサー細胞。

## 【0599】

136. 前記ウイルス核酸は、Gag、Pol、Rev及びTatのうちの1つ以上から選択されるウイルスパッケージングタンパク質をコードする核酸を含む、実施形態134または実施形態135に記載のプロデューサー細胞。

## 【0600】

137. 前記ウイルス核酸は、

以下の核酸配列: 5'LTR (例えば、U5を含み、機能性U3ドメインを欠く)、プサイパッケージング要素 (プサイ)、セントラルポリプリン配列 (cPPT) / セントラル終結配列 (CTS) (例えば、DNAフラップ)、ポリAテール配列、転写後制御要素 (例えば、WPRE)、Rev反応要素 (RRE)、及び3'LTR (例えば、U5を含み、機能性U3を欠く) のうちの1つ以上 (例えば、すべて) を含む、実施形態134~136のいずれかに記載のプロデューサー細胞。

## 【0601】

138. 前記ヘニパウイルスFタンパク質分子またはその生物学的に活性な部分は、

(i) 配列番号2に示される配列、

(ii) 配列番号2と80%もしくは約80%、少なくとも81%もしくは約81%、

10

20

30

40

50

少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

を含む、実施形態 134 ~ 137 のいずれかに記載のプロデューサー細胞。

10

【0602】

139. 前記ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分は、

(i) 配列番号 5 に示される配列、

(ii) 配列番号 5 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

20

を含む、実施形態 134 ~ 137 のいずれかに記載のプロデューサー細胞。

【0603】

140. 前記ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分は、

(i) 配列番号 7 に示される配列、

(ii) 配列番号 7 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

30

を含む、実施形態 134 ~ 137 のいずれかに記載のプロデューサー細胞。

【0604】

141. 前記ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分は、

(i) 配列番号 8 に示される配列をコードするヌクレオチド配列によってコードする配列、

40

(ii) 配列番号 8 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98

50

%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有する配列をコードするヌクレオチド配列によってコードされるアミノ酸配列を含む、実施形態 134 ~ 137 のいずれかに記載のプロデューサー細胞。

【0605】

142. 前記ヘニパウイルス F タンパク質分子またはその生物学的に活性な部分は、

(i) 配列番号 23 に示される配列、

(ii) 配列番号 23 と 80%もしくは約 80%、少なくとも 81%もしくは約 81%、少なくとも 82%もしくは約 82%、少なくとも 83%もしくは約 83%、84%もしくは約 84%、少なくとも 85%もしくは約 85%、少なくとも 86%もしくは約 86%、または少なくとも 87%もしくは約 87%、少なくとも 88%もしくは約 88%、または少なくとも 89%もしくは約 89%、少なくとも 90%もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、少なくとも 96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するアミノ酸配列

を含む、実施形態 134 ~ 137 のいずれかに記載のプロデューサー細胞。

【0606】

143. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、

(i) 配列番号 9、配列番号 28 または配列番号 44 に示される配列、

(ii) 配列番号 9、配列番号 28 または配列番号 44 と 80%もしくは約 80%、少なくとも 81%もしくは約 81%、少なくとも 82%もしくは約 82%、少なくとも 83%もしくは約 83%、84%もしくは約 84%、少なくとも 85%もしくは約 85%、少なくとも 86%もしくは約 86%、または少なくとも 87%もしくは約 87%、少なくとも 88%もしくは約 88%、または少なくとも 89%もしくは約 89%、少なくとも 90%もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、少なくとも 96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するアミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【0607】

144. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、

(i) 配列番号 10 に示される配列、

(ii) 配列番号 10 と 少なくとも 80%もしくは約 80%、少なくとも 81%もしくは約 81%、少なくとも 82%もしくは約 82%、少なくとも 83%もしくは約 83%、84%もしくは約 84%、少なくとも 85%もしくは約 85%、少なくとも 86%もしくは約 86%、または少なくとも 87%もしくは約 87%、少なくとも 88%もしくは約 88%、または少なくとも 89%もしくは約 89%、90%もしくは約 90%、少なくとも 91%もしくは約 91%、少なくとも 92%もしくは約 92%、少なくとも 93%もしくは約 93%、少なくとも 94%もしくは約 94%、少なくとも 95%もしくは約 95%、少なくとも 96%もしくは約 96%、少なくとも 97%もしくは約 97%、少なくとも 98%もしくは約 98%、または少なくとも 99%もしくは約 99%の配列同一性を有するアミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【0608】

145. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (G タンパク質) またはその生物学的に活性な部分は、

10

20

30

40

50

( i ) 配列番号 35 に示される配列、

( i i ) 配列番号 35 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有する  
アミノ酸配列  
を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【 0609 】

146 . 前記ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、

( i ) 配列番号 45 に示される配列、

( i i ) 配列番号 45 と少なくとも 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有する  
アミノ酸配列  
を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【 0610 】

147 . 前記ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、

( i ) 配列番号 11 に示される配列、

( i i ) 配列番号 11 と 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有する  
アミノ酸配列  
を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【 0611 】

148 . 前記ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、

( i ) 配列番号 36 に示される配列、

( i i ) 配列番号 36 と 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、また

は少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【0612】

149. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

10

(i) 配列番号 46 に示される配列、

(ii) 配列番号 46 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

20

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【0613】

150. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

(i) 配列番号 12 に示される配列、

(ii) 配列番号 12 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

30

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【0614】

151. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

40

(i) 配列番号 37 に示される配列、

(ii) 配列番号 37 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有する

50



## アミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

## 【0615】

152. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

(i) 配列番号 47 に示される配列、

(ii) 配列番号 47 と 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

## 【0616】

153. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

(i) 配列番号 13 に示される配列、

(ii) 配列番号 13 と 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

## 【0617】

154. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

(i) 配列番号 38 に示される配列、

(ii) 配列番号 38 と 80 % もしくは約 80 %、少なくとも 81 % もしくは約 81 %、少なくとも 82 % もしくは約 82 %、少なくとも 83 % もしくは約 83 %、84 % もしくは約 84 %、少なくとも 85 % もしくは約 85 %、少なくとも 86 % もしくは約 86 %、または少なくとも 87 % もしくは約 87 %、少なくとも 88 % もしくは約 88 %、または少なくとも 89 % もしくは約 89 %、少なくとも 90 % もしくは約 90 %、少なくとも 91 % もしくは約 91 %、少なくとも 92 % もしくは約 92 %、少なくとも 93 % もしくは約 93 %、少なくとも 94 % もしくは約 94 %、少なくとも 95 % もしくは約 95 %、少なくとも 96 % もしくは約 96 %、少なくとも 97 % もしくは約 97 %、少なくとも 98 % もしくは約 98 %、または少なくとも 99 % もしくは約 99 % の配列同一性を有するアミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

## 【0618】

155. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

10

20

30

40

50

( i ) 配列番号 48 に示される配列、

( i i ) 配列番号 48 と 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有する

10

アミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【 0619 】

156 . 前記ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、

( i ) 配列番号 14 に示される配列、

( i i ) 配列番号 14 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

20

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【 0620 】

157 . 前記ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、

( i ) 配列番号 39 に示される配列、

( i i ) 配列番号 39 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、少なくとも 90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

30

40

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【 0621 】

158 . 前記ヘニパウイルスエンベロープ結合糖タンパク質 G ( G タンパク質 ) またはその生物学的に活性な部分は、

( i ) 配列番号 49 に示される配列、

( i i ) 配列番号 49 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%

50

8%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列

を含む、実施形態134～142のいずれかに記載のプロデューサー細胞。

【0622】

159. 前記ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分は、

10

(i) 配列番号15に示される配列、

(ii) 配列番号15と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列

20

を含む、実施形態134～142のいずれかに記載のプロデューサー細胞。

【0623】

160. 前記ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分は、

(i) 配列番号40に示される配列、

(ii) 配列番号40と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一性を有するアミノ酸配列

30

を含む、実施形態134～142のいずれかに記載のプロデューサー細胞。

【0624】

161. 前記ヘニパウイルスエンベロープ結合糖タンパク質G (Gタンパク質) またはその生物学的に活性な部分は、

40

(i) 配列番号50に示される配列、

(ii) 配列番号50と少なくとも80%もしくは約80%、少なくとも81%もしくは約81%、少なくとも82%もしくは約82%、少なくとも83%もしくは約83%、84%もしくは約84%、少なくとも85%もしくは約85%、少なくとも86%もしくは約86%、または少なくとも87%もしくは約87%、少なくとも88%もしくは約88%、または少なくとも89%もしくは約89%、少なくとも90%もしくは約90%、少なくとも91%もしくは約91%、少なくとも92%もしくは約92%、少なくとも93%もしくは約93%、少なくとも94%もしくは約94%、少なくとも95%もしくは約95%、少なくとも96%もしくは約96%、少なくとも97%もしくは約97%、少なくとも98%もしくは約98%、または少なくとも99%もしくは約99%の配列同一

50

性を有するアミノ酸配列

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【0625】

162. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

(i) 配列番号 16 に示される配列、

(ii) 配列番号 16 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

10

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【0626】

163. 前記ヘニパウイルスエンベロープ結合糖タンパク質 G (Gタンパク質) またはその生物学的に活性な部分は、

(i) 配列番号 51 に示される配列、

(ii) 配列番号 51 と少なくとも 80% もしくは約 80%、少なくとも 81% もしくは約 81%、少なくとも 82% もしくは約 82%、少なくとも 83% もしくは約 83%、84% もしくは約 84%、少なくとも 85% もしくは約 85%、少なくとも 86% もしくは約 86%、または少なくとも 87% もしくは約 87%、少なくとも 88% もしくは約 88%、または少なくとも 89% もしくは約 89%、90% もしくは約 90%、少なくとも 91% もしくは約 91%、少なくとも 92% もしくは約 92%、少なくとも 93% もしくは約 93%、少なくとも 94% もしくは約 94%、少なくとも 95% もしくは約 95%、少なくとも 96% もしくは約 96%、少なくとも 97% もしくは約 97%、少なくとも 98% もしくは約 98%、または少なくとも 99% もしくは約 99% の配列同一性を有するアミノ酸配列

20

を含む、実施形態 134 ~ 142 のいずれかに記載のプロデューサー細胞。

【0627】

164. 実施形態 134 ~ 163 のいずれかに記載のプロデューサー細胞から生成された、ウイルスベクター粒子またはウイルス様粒子。

【0628】

165. 実施形態 1 ~ 81 及び 173 ~ 176 のいずれかに記載の複数の標的化脂質粒子を含む、組成物。

【0629】

166. 薬学的に許容可能な担体をさらに含む、実施形態 165 に記載の組成物。

40

【0630】

167. 前記標的化脂質粒子は、1 μm 未満の平均直径を含む、実施形態 165 または実施形態 166 に記載の薬学的組成物。

【0631】

168. 外因性物質を対象 (例えば、ヒト対象) に送達する方法であって、実施形態 1 ~ 81 及び 173 ~ 176 のいずれかに記載の標的化脂質粒子または実施形態 165 ~ 167 及び 177 のいずれかに記載の組成物を前記対象に投与することを含む、前記方法。

【0632】

169. 対象 (例えば、ヒト対象) における疾患または障害を処置する方法であって、実施形態 1 ~ 81 及び 173 ~ 176 のいずれかに記載の標的化脂質粒子または実施形態

50

165～167及び177のいずれかに記載の組成物を前記対象に投与することを含む、前記方法。

【0633】

170．哺乳動物細胞を標的化脂質粒子に融合させる方法であって、実施形態1～81及び173～176のいずれかに記載の標的化脂質粒子または実施形態165～167及び177のいずれかに記載の組成物を前記対象に投与することを含む、前記方法。

【0634】

171．前記哺乳動物細胞を前記標的化脂質粒子に前記融合させることは、外因性物質を対象（例えば、ヒト対象）に送達する、実施形態170に記載の方法。

【0635】

172．前記哺乳動物細胞を前記標的化脂質粒子に前記融合させることは、対象（例えば、ヒト対象）における疾患または障害を処置する、実施形態170または実施形態171に記載の方法。

【0636】

173．前記標的化脂質粒子は、同様の脂質二重層に同じエンベロープタンパク質が組み込まれているがそれが代替的標的化部位に融合されている参照脂質粒子と比較して多い前記標的化エンベロープタンパク質発現を有し、任意に前記代替的標的化部位は、一本鎖可変断片（scFv）である、実施形態1～81のいずれかに記載の標的化脂質粒子。

【0637】

174．前記発現は、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%、300%、400%、500%以上またはそれを超えて増加する、実施形態173に記載の標的化脂質粒子。

【0638】

175．前記発現は、1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍、30倍以上またはそれを超えて、好ましくは10倍もしくは約10倍もしくは10倍超またはそれを超えて増加する、実施形態173に記載の標的化脂質粒子。

【0639】

176．形質導入後の標的細胞における力価は、 $1 \times 10^6$ 形質導入単位（TU）/mL以上、 $2 \times 10^6$  TU/mL以上、 $3 \times 10^6$  TU/mL以上、 $4 \times 10^6$  TU/mL以上、 $5 \times 10^6$  TU/mL以上、 $6 \times 10^6$  TU/mL以上、 $7 \times 10^6$  TU/mL以上、 $8 \times 10^6$  TU/mL以上、 $9 \times 10^6$  TU/mL以上、または $1 \times 10^7$  TU/mL以上である、実施形態1～81及び173～175のいずれかに記載の標的化脂質粒子または実施形態164に記載のウイルスベクター粒子もしくはウイルス様粒子。

【0640】

177．前記組成物における脂質粒子の集団の中で、50%もしくは約50%超、55%もしくは約55%超、60%もしくは約60%超、65%もしくは約65%超、70%もしくは約70%超、または75%もしくは約75%超が、前記標的化エンベロープタンパク質について表面陽性である、実施形態165～167のいずれかに記載の組成物。

【0641】

178．前記標的化エンベロープタンパク質が、少なくとも約（0.001、0.002、0.005、0.01、0.02、0.05、0.1、0.2または0.5）の標的化エンベロープタンパク質/nm<sup>2</sup>の密度で前記標的化脂質粒子の表面上に存在する、実施形態1～81及び173～176のいずれかに記載の標的化脂質粒子。

【0642】

179．実施形態1～81及び173～176及び178のいずれかに記載の複数の標的化脂質粒子を含む組成物であって、前記標的化エンベロープタンパク質が、少なくとも約（0.001、0.002、0.005、0.01、0.02、0.05、0.1、0.2または0.5）の標的化エンベロープタンパク質/nm<sup>2</sup>の平均密度で前記標的化脂質粒子の表面上に存在する、前記組成物。

10

20

30

40

50

## 【0643】

180. 膜（例えば、細胞膜）に同じエンベロープタンパク質が組み込まれているがそれが代替的標的化部位に融合されている参照プロデューサー細胞と比較して多い、前記標的化エンベロープタンパク質の膜（例えば、細胞膜）発現を、前記プロデューサー細胞が有し、任意に前記代替的標的化部位は、一本鎖可変断片（s c F v）である、実施形態134～163のいずれか1つに記載のプロデューサー細胞。

## 【0644】

181. 前記発現は、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、125%、150%、200%、300%、400%、500%以上またはそれを超えて増加する、実施形態180に記載のプロデューサー細胞

10

## 【0645】

182. 前記発現は、1.5倍、2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍、15倍、20倍以上またはそれを超えて、好ましくは10倍もしくは約10倍もしくは10倍超またはそれを超えて増加する、実施形態180に記載のプロデューサー細胞。

## 【0646】

183. 平方ミクロン当たり少なくとも20のタンパク質（例えば、少なくとも50、100、200、500、1000、2000、5000、または10,000のタンパク質）である、前記プロデューサー細胞の膜（例えば、細胞膜）上の前記標的化エンベロープタンパク質の発現を、前記プロデューサー細胞が有する、実施形態134～163及び180～182のいずれか1つに記載のプロデューサー細胞。

20

## 【0647】

184. 前記標的化エンベロープタンパク質は、前記プロデューサー細胞の総膜（例えば、細胞膜）タンパク質の少なくとも0.1%（例えば、少なくとも0.2%、0.5%、1%、2%、3%、4%、5%、6%、7%、8%、9%、または10%）（例えば、総タンパク質量に対して）を構成する、実施形態134～163及び180～183のいずれか1つに記載のプロデューサー細胞。

## 【実施例】

## 【0648】

以下の例は、例示の目的のためにのみ含まれ、本発明の範囲を限定することは意図されていない。

30

## 【0649】

実施例1：標的化バインダーを含有するプロデューサー細胞の生成及び特性づけ

この実施例は、NiVGがs c F vまたはV H H結合モダリティに連結されたNiVG標的化結合配列の生成及び評価を記載する。

## 【0650】

A. CD4に仕向けられた結合モダリティ

ヒト細胞受容体CD4に対するs c F vまたはV H H結合モダリティを含有する例示的な再標的化NiVGフソゲンコンストラクトを生成した。各結合モダリティについて、独自のCDR3を含有する4つの異なる配列を評価した。各々の例示的なバインダー配列をコード最適化し、NiVG (Gc 34; Bender et al. 2016 PLoS Pathol 12(6): e1005641) をコードする配列との融合体として発現ベクターにクローニングした。得られたベクターは、NiVG (配列番号16) を含有するNiVG標的化ドメイン、フレキシブルリンカー及び結合ドメインと、それに続く検出用の6xHisタグ (NiVG-リンカー-s c F v-6xHis) をコードしていた。

40

## 【0651】

サブクローニング後、トランスフェクション試薬を使用して5 µgの各々の例示的なコンストラクトをHEK293細胞にトランスフェクションした。p c D N A 3.1プラス

50

ミド（空のベクター）及びバインダードメインを有しない発現ベクター（NiVG-リンカー-バインダーなし）を陰性対照として使用した。

【0652】

トランスフェクションから48時間後、細胞を採取し、100,000細胞を、50 nMまたは300 nMのいずれかのヒトFcタグを有する可溶性ヒトCD4タンパク質（hCD4-Fc）と4で1時間インキュベートした。インキュベート後、細胞を洗浄し、NiVG-バインダーの表面発現を検出するためのAlexa-647にコンジュゲートされた抗His抗体、及び可溶性hCD4-Fcタンパク質に対する結合を検出するためのAlexa-488にコンジュゲートされた抗ヒトFc抗体で共染色した。

【0653】

細胞をフローサイトメトリーによって分析し、His（表面発現）及びFc（CD4タンパク質結合）についてのゲートを陰性対照の空のベクター（pcDNA3.1）に基づいて設定した。VHH結合モダリティを含有するコンストラクトでトランスフェクションされた細胞の蛍光強度中央値（MFI）の評価は、scFv結合モダリティを含有するコンストラクトでトランスフェクションされた細胞よりも、His+細胞の%によって定量されるより高い表面発現（図1A）及びFc+細胞の%によって定量される可溶性hCD4-Fcタンパク質に対するより高い結合（図1B）を実証した。

【0654】

B. 複数の細胞受容体に仕向けられた結合モダリティ

全般的に上述したscFv及びVHH結合モダリティを含有するが、他の細胞受容体hCD8、CD4、ASGR2、TM4SF5、LDLRまたはASGR1に対して仕向けられた独自の配列を含有する例示的なコンストラクトを生成した。複数の配列（各々が独自のCDR3を含有する）を、異なる細胞受容体を含有する各結合モダリティについて評価した。上述したようにNiVG-リンカー-6xHis発現ベクターにサブクロニングした後、5 µgの各々の例示的なコンストラクトを約HEK293細胞にトランスフェクションした。pcDNA3.1プラスミド（空のベクター）及び結合ドメインを有しない発現ベクター（NiVG-リンカー-バインダーなし）を陰性対照として使用した。

【0655】

トランスフェクションから48時間後、細胞を採取し、100,000細胞を洗浄し、NiVG-バインダーの表面発現を検出するためのAlexa-647にコンジュゲートされた抗His抗体で染色した。細胞をフローサイトメトリーによって分析し、His（表面発現）についてのゲートを陰性対照の空のベクター（pcDNA3.1）に基づいて設定した。蛍光強度中央値（MFI）を、100に設定されたNiVG-バインダーなし対照のものに対して正規化した。scFv結合モダリティと比較して、VHH結合モダリティを含有するコンストラクトでトランスフェクションされた細胞は、His+細胞の%によって定量される293細胞に対する標的化結合配列のより高い表面発現を実証した（図1C）。

【0656】

実施例2：標的化バインダーで偽型化されたレンチウイルスの生成及び特性づけ

この実施例は、NiVG再標的化フソゲンで偽型化されたレンチウイルスの生成及び初代ヒトT細胞の形質導入の評価を記載する。

【0657】

A. NiVG偽型化レンチウイルスの生成

293細胞を10 cmの皿に $5.4 \times 10^6$ で播種し、24時間放置した。播種の24時間後、細胞を以下のプラスミドを用いてポリエチレンイミン（PEI）を使用してトランスフェクションした：scFvまたはVHH結合モダリティに連結したhCD4標的化結合配列を含有するNiVG偽型化ベクター（NiVG-リンカー-hCD4-結合モダリティ）、NiVF配列NiVFdel22をコードするヌクレオチド配列を含有するベクター（配列番号8；またはシグナル配列を有しない配列番号23；Bender et al. 2016 PLoS）、空の骨格、HIV-1 pol、HIV-1 gag、H

10

20

30

40

50

I V - 1 R e v、H I V - 1 T a t、A m p R プロモーター及びS V 4 0 プロモーターを含有するパッケージングプラスミド、ならびにS F F V プロモーターのコントロール下で高感度緑色蛍光タンパク質 ( e G F P ) をコードするレンチウイルスレポータープラスミド p L e n t i - S F F V - e G F P。陽性対照細胞を、4  $\mu$  g の V S V - G と共に上述したプラスミドを使用して生成した。

【 0 6 5 8 】

B . 初代ヒトT細胞のN i v G 偽型化レンチウイルス形質導入効率

T細胞を濃縮するためにネガティブ選択された末梢血からのP a n T細胞 ( S t e m C e l l T e c h , V a n c o u v e r , C a n a d a ) を解凍し、抗C D 3 / 抗C D 2 8 で2日間活性化した。概して上述したように生成された濃縮したレンチウイルスを0 . 0 5 希釈で開始して6倍連続希釈し、合計で4点の希釈系列とした。レンチウイルスを1 0 0 , 0 0 0 個のP a n T細胞に添加し、1 0 0 0 g にて2 5 C で9 0 分間のスピニングによって形質導入した。形質導入したP a n T細胞を形質導入後2日目及び5日目に分割し、形質導入後7日目に、細胞を採取し、A l e x a - 6 4 7 がコンジュゲートされた抗ヒトC D 4 抗体で染色した。細胞をフローサイトメトリーによって分析し、力価を、G F P + であったC D 4 陽性細胞の%によって決定した。V H H 結合モダリティを含有するコンストラクトでトランスフェクションされた細胞は、初代ヒトT細胞に対するs c F v 結合モダリティを含有するコンストラクトよりも1 0 倍増加した力価を実証した ( 図 2 ) 。

10

20

【 0 6 5 9 】

実施例 3 . C D 8 標的化バインダーで偽型化されたレンチウイルスの i n v i v o 送達

この実施例は、C D 8 N i v G 再標的化フソゲンで偽型化されたレンチウイルスの生成及び初代ヒトT細胞の形質導入の i n v i v o 評価を記載する。

【 0 6 6 0 】

C D 8 再標的化N i v G フソゲンを本質的に実施例 2 に記載されているように生成した。再標的化N i v G 偽型化フソゲンは、N i v G ( 配列番号 1 6 ) を含有するN i v G 標的化ドメイン、フレキシブルリンカー及び例示的なC D 8 結合ドメインであるV H H またはs c F v のいずれかの結合モダリティを含有していた。

【 0 6 6 1 】

ヒト末梢血単核細胞 ( P B M C ) からT細胞を抗C D 3 / 抗C D 2 8 で3日間活性化した。3日のインキュベート後、 $1 \times 10^7$ 細胞をN O D - s c i d - I L 2 r<sup>n u l 1</sup>マウスに腹腔内に注射した。注射から1日後、マウスに上述したように生成した $1 \times 10^7$ 形質導入単位 ( T U ) のC D 8 N i v G 偽型化レンチウイルス、またはレンチウイルスベクター ( L V V ) なし対照を、腹腔内注射を介して摂取させた。C D 8 N i v G 偽型化レンチウイルス注射から7日後に、腹腔細胞を採取し、フローサイトメトリーによって分析し、G F P + であったC D 8 陽性または陰性細胞の%によって力価を決定した。C D 8 再標的化偽型化レンチウイルスは、C D 8 + T細胞の有意な i n v i v o 形質導入 ( 図 3 A ) 及びC D 8 - T細胞の最小の形質導入 ( 図 3 B ) を実証した。これらの結果は、C D 8 標的化偽型化レンチウイルス媒介送達が、意図される細胞タイプ ( 例えば、C D 8 + T細胞 ) への導入遺伝子の特異的送達を可能にすることを示している。

30

40

【 0 6 6 2 】

実施例 4 . C D 8 標的化バインダーを有する偽型化レンチウイルスを含有するキメラ抗原受容体 ( C A R ) の i n v i t r o 評価

この実施例は、C D 8 再標的化フソゲンで偽型化されたレンチウイルスの i n v i t r o 腫瘍殺傷活性及びC D 1 9 に仕向けられたキメラ抗原受容体 ( C D 1 9 C A R ) の発現を記載する。e G F P またはC D 1 9 C A R のいずれかをコードするプラスミドを2 9 3 プロデューサー細胞にトランスフェクションしたこと以外は、本質的に実施例 3 に記載されているようにレンチウイルスを生成した。C D 1 9 C A R は、C D 1 9 に対して仕向けられた抗s c F v と、4 - 1 B B 及びC D 3 - ゼータである細胞内成分を含有する細胞内シグナル伝達ドメインとを含有していた。

50



## 【0663】

ヒト末梢血単核細胞（PBMＣ）を抗CD3 / 抗CD28試薬で活性化し、CD19 + CARまたはGFPを発現するCD8再標的化NiVGLレンチウイルスで様々な濃度範囲（10<sup>-10</sup> ~ 10<sup>-1000</sup>形質導入単位 / ウェル）で形質導入した。RFP + Na1m6白血球細胞を3日目に培養物に添加し、Na1m6細胞の排除をフローサイトメトリーによって18時間の時点で評価した。

## 【0664】

図4Aに示されているように、形質導入から4日後に両方のCD8再標的化フソゲンを有するCD8 + 細胞においてCD19 + CAR発現を特異的に検出した。CD19CARを発現する形質導入されたCD8 + T細胞はまた、CD19 + Na1m6白血球細胞の殺傷において、強力でレンチウイルス用量依存的な増加を媒介し、対照的に、GFPを発現するように形質導入された細胞は、標的細胞殺傷を示さなかった（図4B）。

10

## 【0665】

これらの結果は、CD19CARをコードする導入遺伝子を有するCD8再標的化偽型化レンチウイルスがCD19CARをヒトCD8 + T細胞に送達して、PBMＣの複雑な混合物におけるCD8 + T細胞の特異的形質導入を媒介し、*in vitro*での白血球細胞の殺傷によって用量依存的抗腫瘍反応を示したことを実証している。

## 【0666】

本発明は、例えば、本発明の様々な態様を例示するために提供される特定の開示される実施形態に範囲が限定されることは意図されていない。記載される組成物及び方法に対する様々な修飾が、本明細書における説明及び教示から明らかになるであろう。そのような変形は、本開示の真の範囲及び趣旨から逸脱することなく実施され得、本開示の範囲に入ることが意図されている。

20

## 【0667】

配列

30

40

50

#	配列	アノテーション
1	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQEELLVPS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALS N GVLFANCISV TCQCQTTGRA ISQSGEQTLL MIDNTTCPTA VLG NVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNP SLIS MLSMI ILYVL SIASLCIGLI TFISFIIIVEK KRNTYSRLED RRV RPTSSGD LYYIGT	シグナル配列を有するニパウイルス NiV-F (アミノ酸 1-546) Uniprot Q9IH63
2	ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNMSQ CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQEELLVPS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALS N GVLFANCISV TCQCQTTGRA ISQSGEQTLL MIDNTTCPTA VLGNVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLQQS KDYIKEAQRL LDTVNP SLIS MLSMI ILYVL SIASLCIGLI TFISFIIIVEK KRNTYSRLED RRV RPTSSGD LYYIGT	ニパウイルス NiV-F F0 (アミノ酸 27-546)
3	ILHYEKLSKIGLVKGVTRKYKIKSNPLTKDIVIKMIPNVSNMSQCTGSVMENYKT RLNGILTPIKGALEIYKNNTHDLVGDVR	ニパウイルス NiV-F F2 (アミノ酸 27-109)
4	LAGVIMAGVAIGIATAAQITAGVALYEAMKNADNINKLKSS IESTNEAVVKLQET AEKTVYVLTALQDYINTNLVPTIDKISCKQTELSLDLALSKYLSDLLFVFGPNLQ DPVSNSMTIQAISQAFGGNYETLLRTLGYATEDFDDLLESDSITGQIIYVDLSSY YIIVRVYFPILTEIQQAYIQEELLVPSFNNDNSEWISIVPNFILVRNTLISNIEIG FCLITKRSVICNQDYATPMTNNMRECLTGSTEKCPRELVVSSHVPRFALSNGVLF ANCISVTCQCQTTGRAISQSGEQTLLMIDNTTCPTAVLG NVIISLGKYLGSVNYN SEGIAIGPPVF TDKVDISSQISSMNQSLQQSKDYIKEAQRL LDTVNP SLIS MLSM I ILYVLSIASLCIGLITFISFIIIVEKKRNTYSRLEDRRV RPTSSGDLYYIGT	ニパウイルス NiV F F1 (アミノ酸 110-546)

10

20

30

40

50

5	<p>ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNSMSQ  CTGSMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI  GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY  VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFFVFGPNLQD  PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD  LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN  TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS  SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSSEQTLL MIDNTTCPTA  VLGNVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQLQQS  KDYIKEAQRLLDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIIVEK  KRNTGT</p>	<p>ニパウイ  ルス NiV-  F F0 T234  切断 (ア  ミノ酸  525-544)</p>	10
6	<p>LAGVIMAGVAIGIATAAQITAGVALYEAMKNADNINKLKSSIESTNEAVVKLQET  AEKTVYVLTALQDYINTNLVPTIDKISCKQTELSLDLALSKYLSDDLFFVFGPNLQ  DPVNSMTIQAISQAFGGNYETLLRTLGYATEDFDDLLESDSITGQIIYVDLSSY  YIIVRVYFPILTEIQQAYIQELLPVSFNNDNSEWISIVPNFILVRNTLISNIEIG  FCLITKRSVICNQDYATPMTNMMRECLTGSTEKCPRELVVSSHVPRFALSNGVLF  ANCISVTCQCQTTGRAISQSSEQTLLMIDNTTCPTAVLGNVIISLGKYLGSVNYN  SEGIAIGPPVFTDKVDISSQISSMNQLQQSKDYIKEAQRLLDTVNPSLISMLSM  IILYVLSIASLCIGLITFISFIIIVEKKRNTGT</p>	<p>ニパウイ  ルス NiV  F F1 (アミ  ノ酸 110-  546) 切断  (アミノ酸  525-544)</p>	20
7	<p>ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNSMSQ  CTGSMENYK TRLNGILTPI KGALEIYKNQ THDLVGDVRL AGVIMAGVAI  GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY  VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFFVFGPNLQD  PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD  LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN  TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS  SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSSEQTLL MIDNTTCPTA  VLGNVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQLQQS  KDYIKEAQRLLDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIIVEK  KRNTGT</p>	<p>ニパウイ  ルス NiV-  F F0 T234  切断 (ア  ミノ酸  525-544)  及び N 結  合型グリ  コシル化  部位にお  ける変異</p>	30
8	<p>MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS  NPLTKDIVIK MIPNVSNSMSQ CTGSMENYK TRLNGILTPI KGALEIYKNN  THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS  IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD  LALSKYLSDL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT  EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS  FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN  NMRECLTGST EKCPRELVVS SHVPRFALSN GVLFANCISV TCQCQTTGRA  ISQSSEQTLL MIDNTTCPTA VLGNVIISLG KYLGSVNYNS EGIAIGPPVF  TDKVDISSQI SSMNQLQQS KDYIKEAQRLLDTVNPSLIS MLSMIILYVL  SIASLCIGLI TFISFIIIVEK KRNT</p>	<p>細胞質尾  部で切断  された  NiV 融合  糖タンパ  ク質  (FcDelta22  )(シグナ  ル配列を  有する)</p>	40

9	MGPAENKKVR FENTTSDK GK IPSKVIKSY G TMDIKKINE GLLDSKILSA FNTVIAL LGS IVIIVMNIMI IQNYTRSTD N QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTL PV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPGSIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM GIRPN SHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タ ンパク質 結合糖タ ンパク質 (602 アミ ノ酸)	10
1 0	MGKVR FENTTSDK GK IPSKVIKSY G TMDIKKINE GLLDSKILSA FNTVIAL LGS IVIIVMNIMI IQNYTRSTD N QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTL PV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPGSIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM GIRPN SHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ5	20
1 1	MGN TSDK GK IPSKVIKSY G TMDIKKINE GLLDSKILSA FNTVIAL LGS IVIIVMNIMI IQNYTRSTD N QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTL PV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPGSIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM GIRPN SHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ10	30

10

20

30

40

50

12	MGK GK IPSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIAL LGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQ LALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ15	10
13	MGSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIAL LGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQ LALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ20	20
14	MGSYY GTMDIKKINE GLLDSKILSA FNTVIAL LGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQ LALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ25	30

40

50

15	MGTMDIKKINE GLLDSKILSA FNTVIALLGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ30	10	
16	MKKINEGLLDSKILSA FNTVIALLGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PAICAEGVYN DAFLIDRINW ISAGVFLDSN ATAANPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG タ ンパク質 結合糖タ ンパク質 切断及び 変異型 (E501A, W504A, Q530A, E533A) NiV G タ ンパク質 (Gc Δ 34)	20	30

17	MATQEVRLKC NPLTKDIVIK THDLVGDVKL IESTNEAVVK LALSKYLSL EDFDDLLESD FNNDNSEWIS SVRECLTGST ISQSGEQTLL TDKVDISSQI SIAALCIGLI	LLCGIIVLVL MIPNVSNSK AGVVMAGIAI LQETAECTVY LFVFGPNLQD SIAGQIVYVD IVPNFVLIRN DKCPRELVVS MIDNTTCTTV SSMNQSLQQS TFISFVIVEK	SLEGLGILHY CTGTVMENYK GIATAAQITA VLTALQDYIN PVSNSMTIQA LSSYYIIVRV TLISNIEVKY SHVPRFALS VLGNIIISLG KDYIKEAQKI KRGNYRLLD	EKLSKIGLVK SRLTGILSPI GVALYEAMKN TNLVPTIDQI ISQAFGGNYE YFPILTEIQQ CLITKKSVC GVLFANCISV KYLGSINYNS LDTVNPSLIS RQVRPVSNGD	GITRKYKIKS KGAIELYNNN ADNINKLKSS SCKQTEALD TLLRTLGYAT AYVQELLPVS NQDYATPMTA TCQCQTTGRA ESIAGPPVY MLSMIILYVL LYYIGT	ヘンドラ ウイルス Fタンパ ク質 Uniprot O89342 (シグナル 配列を有 する)	10
18	MMADSKLVSL NTVIALLSI IGTEIGPKVS PPLKIHECNI RLISYTLPIN GVGEVLDRGD GDPILNSTSW YGPSGIKQGD VNSKSHYILR PVFYQASYSW EVCWEGTYND LAEDDTNAQK CSES	NNNLSGKIKD IIIVMNIMII LIDTSSTITI SCPNPLPRE TREGVCITDP KVPSMFMTNV TESLSLIRLA TLYFPAVGFL SGLLKYNLSL DTMIKLGVD AFLIDRLNWW TITDCFLEN	QGVKIKNYG QNYTRTTDQ PANIGLLGSK YRPISQGVSD LLAVDNGFFA WTPPNPSTIH VRPKSDSGDY PRTEFQYND GGDIILQFIE TVDPLRVQR SAGVYLNNSQ VIWCISLVEI	TMDIKKINDG ALIKESLQSV ISQSTSSINE LVGLPNQICL YSHLEKIGSC HCSSTYHEDF NQKYIAITKV NCPIIHCKYS IADNRLTIGS NNSVISRPGQ TAENPVFAVF YDTGDSVIRP	LLDSKILGAF QQQIKALTDK NVNDKCKFTL QKTTSTILKP TRGIAKQRII YITLCAVSHV ERKGYDKVMP KAENCRLSMG PSKIYNSLGQ SQCPRENVCP KDNEILYQVP KLFVAVKIPAQ	ヘンドラ ウイルス Gタンパ ク質 Uniprot O89343	20
19	MVVILDKRCY NPLTKDIVIK THDLVGDVRL IESTNEAVVK LALSKYLSL EDFDDLLESD FNNDNSEWIS NMRECLTGST ISQSGEQTLL TDKVDISSQI SIASLCIGLI	CNLLILILMI MIPNVSNSQ AGVIMAGVAI LQETAECTVY LFVFGPNLQD SITGQIIYVD IVPNFILVRN EKCPRELVVS MIDNTTCPTA SSMNQSLQQS TFISFIIVEK	SECSVGILHY CTGSVMENYK GIATAAQITA VLTALQDYIN PVSNSMTIQA LSSYYIIVRV TLISNIEIGF SHVPRFALS VLGNVIISLG KDYIKEAQR KRNTGT	EKLSKIGLVK TRLNGILTPI GVALYEAMKN TNLVPTIDKI ISQAFGGNYE YFPILTEIQQ CLITKRSVIC GVLFANCISV KYLGSVNYNS LDTVNPSLIS	GVTRKYKIKS KGALEIYKNN ADNINKLKSS SCKQTELSLD TLLRTLGYAT AYIQELLPVS NQDYATPMTN TCQCQTTGRA EGIAIGPPVF MLSMIILYVL	ニパウイ ルス NiV- F F0 T234 切断(ア ミノ酸 525-544) (シグナル 配列を有 する)	30

10

20

30

40

50

2 0	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNSMQ CTGSMENYK TRLNGILTPI KGALEYKKNQ THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALSN GVLFANCISV TCQCQTGRA ISQSGEQTLT MIDNTTCPTA VLG NVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLOQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIIVEK KRNTGT	ニパウイ ルス NiV- F F0 T234 切断 (ア ミノ酸 525-544) 及び N 結 合型グリ コシル化 部位にお ける変異 (シグナル 配列を有 する)	10
2 1	MVVILDKRCY CNLLILILMI SECSVGILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNSMQ CTGSMENYK TRLNGILTPI KGALEYKKN THDLVGDVRL AGVIMAGVAI GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRRTLGYAT EDFDDLLESD SITGQIIYVD LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS SHVPRFALSN GVLFANCISV TCQCQTGRA ISQSGEQTLT MIDNTTCPTA VLG NVIISLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQSLOQS KDYIKEAQRL LDTVNPSLIS MLSMIILYVL SIASLCIGLI TFISFIIIVEK KRNT	細胞質尾 部で切断 された NiV 融 合 糖タンパ ク質 (FcDelta22 ) (シグナル 配列を 有する)	20
2 2	MKKINEGLLDKILSA FNTVIALLGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRL DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIINSTY WSGSLMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLAS EDTNAQ KTITNCFLLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG タ ンパク質 結合糖タ ンパク質 切断型 (Gc Δ 34)	30

10

20

30

40

50



23	<p>ILHY EKLSKIGLVK GVTRKYKIKS NPLTKDIVIK MIPNVSNSMSQ  CTGSVMENYK TRLNGILTPI KGALEIYKNN THDLVGDVRL AGVIMAGVAI  GIATAAQITA GVALYEAMKN ADNINKLKSS IESTNEAVVK LQETAECTVY  VLTALQDYIN TNLVPTIDKI SCKQTELSLD LALSKYLSDL LFFVFGPNLQD  PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT EDFDDLLESD SITGQIIYVD  LSSYYIIVRV YFPILTEIQQ AYIQELLPVS FNNDNSEWIS IVPNFILVRN  TLISNIEIGF CLITKRSVIC NQDYATPMTN NMRECLTGST EKCPRELVVS  SHVPRFALSN GVLFANCISV TCQCQTTGRA ISQSGETLL MIDNTTCPTA  VLGNVILSLG KYLGSVNYNS EGIAIGPPVF TDKVDISSQI SSMNQLQQS  KDYIKEAQLR LDTVNPPLIS MLSMIILYVL SIASLCIGLI TFISFIIVEK  KRNT</p>	<p>細胞質尾 部で切断 された成 熟 NiV 融 合糖タン パク質 (FcDelta22 )</p>
24	<p>MSNKRTTVLIIISYTLFYLNNAIIVGFDFDKLNKIGVVQGRVLNYKIKGDPMTKD  LVLKFI PNIVNITECVREPLSRNETVRRLLP IHNMLGLYLNN TNAKMTGLMIA  GVIMGGIAIGIATAAQITAGFALYEAKNTENIQKLTDSIMKTQDSIDKLTDSVG  TSILILNKLQTYINNQLVPLELLSQRNKIEFDLMLTKYLVDLMTVIGPNINNP  VNKDMTIQSLSLFDGNYDIMMSELGYTPQDFDLIESKSITGQIIYVDMENLYV  VIRTYLPTLIEVPDAQIYEFNKITMSSNGGEYLSTIPNFILIRGNYSNIDVATC  YMTKASVICNQDYSLPMSONLRSCYQGETEYCPVEAVIASHSPRFALTNVIFAN  CINTICRCQDNGKTITQINQFVSMIDNSTCNDVMVDKFTIKVKGKYMGRKDINNI  NIQIGPQIIIDKVDLSNEINKMNQSLKDSIFYLREAKRILDSVNISLISPSVQLF  LIIISVLSFIIILLIIVVLYLCKSKHSYKYNKFIDDPDYNDYKRERINGKASKSN  NIYYVGD</p>	<p>gb:JQ0017 76:6129- 8166 生物: セダーウ イルス 株 名:CG1a  タンパク 質名:融合 糖タンパ ク質 遺伝 子記号:F (シグナル 配列を有 する)</p>
25	<p>MALNKNMFSSFLGLYLLVYATTVQSSIHYSLSKVGVIKGLTYNYKIKGSPSTKL  MVVKLIPNIDSVKNCTQKQYDEYKNLVRKALEPVKMAIDTMLNNVKSNNKYRFA  GAIMAGVALGVATAATVTAGIALHRSENENAQAIANMKSAIQNTNEAVKQLQLANK  QTLAVIDTIRGEINNNIIPVINQLSCDTIGLSVGIRLTQYYSEIITAFGPALQNP  VNTRITIQAISSVFNNGFDELLKIMGYTSGDLYEILHSELIRGNIIDVDVDAGYI  ALEIEFPNLTLPNAVAVQELMPISYNIDGDEWVTLVPRFVLTRTTLLSNIDTSRC  TITDSSVICDNDYALPMSHELIGCLQGDTSKAREKVVSSYVPKFALSDGLVYAN  CLNTICRCMDTDPISQSLGATVSLDNKRCVYQVGDVLSVGSYLGDEYNAD  NVELGPPVIDKIDIGNQLAGINQTLQEAEDYIEKSEEFKGVNPSIITLGSMVV  LYIFMILIAIVSVIALVLSIKLTVKGNVVRQFTYTDQHVPSMENINYVSH</p>	<p>gb:NC_02 5352:5950 -8712 生 物:Mojian g ウイル ス 株 名:Tonggu an タン パク質名: 融合タン パク質 遺 伝子記 号:F(シグ ナル配列 を有する)</p>

10

20

30

40

50

2 6	MKKKTDNPTISKRGHNHSRGIKSRALLRETDNYSNGLIVENLVRNCHHPS KNNLNYTKTQKRDSTIPYRVEERKGHYPKIKHLIDKSYKHIKRGKRRNGH NGNIITIIILLILLILILKTQMSEGAIHYETLSKIGLIKGITREYKVKGTPSS KDVIKLIIPNVTGLNKCTNISMENYKEQLDKILIPINNIIELYANSTKSA PGNARFAGVIIAGVALGVAAAAQITAGIALHEARQNAERINLLKDSISAT NNAVAELQEATGGIVNVITGMQDYINTNLVPQIDKLQCSQIKTALDISLS QYYSEILTTFVFPNLQNPVTTSMSIQAISQSFGGNIDLLLNLGLYTANDLL DLLESKSITGQITYINLEHYFMVIRVYYPIMTTISNAYVQELIKISFNVD GSEWVSLVPSYILIRNSYLSNIDISECLITKNSVICRHFAMPMSYTLKE CLTGDTEKCPREAVVTSYVPRFAISGGVIYANCLSTTCQCYQTGKVIAQD GSQTLMMIDNQTCISVRIEEILISTGKYLGSQEYNTMHVSVGNPVFTDKL DITSQISNINQSIEQSKFYLDKSKAILDKINLNLIGSVPISILFIIAILS LILSIITFVIVMIIVRRYNKYTPLINSDPSSRRSTIQDVYIIPNPGEHSI RSAARSIDRDRD	gb:NC_025256: 6865-8853 生 物:コウモリパ ラミクソウイ ルス Eid_hel/GH- M74a/GHA/20 09 株 名:BatPV/Eid_ hel/GH- M74a/GHA/20 09 タンパク質 名:融合タンパ ク質 遺伝子記 号:F (シグナル 配列を有する)
2 7	(GGGGGS)n(nは1~6である)	ペプチドリン カー
2 8	MPAENKKVRFENTTSDKGKIPSKVIKSYYGTMDIKKINEGLLDSKILSAF NTVIALLGSIIVIMNIMI IQNYTRSTDNQAVIKDALQGIQQQIKGLADK IGTEIGPKVSLIDTSSTITIPANIGLLGSKISQSTASINENVNEKCKFTL PPLKIHECNISCPNPLPFREYRPQTEGVSNLVGLPNNICLQKTSNQILKP KLISYTLPVVVGQSGTCITDPLLAMDEGYFAYSHLERIGSCSRGVSKQRII GVGEVLDRGDEVPSLFMTNVWTPPNPNTVYHCSAVYNNEFYVLCVAVSTV GDPILNSTYWSGSLMMTRLAVKPKSNGGGYNQHQLALRSIEKGRYDKVMP YGPSTYKQGDITLYFPAVGFLVRTEFKYNDNSNCPITKCQYSKPENCRLSMG IRPNSHYILRSGLLKYNLSDGENPKVVFIEISDQRLSIGSPSKIYDSLQ PVFYQASFSWDTMIKFGDVLTVNPLVVNWRNNTVISRPGQSQCPRFNTCP EICWEGVYNDAFLIDRINWISAGVFLDSNQTAENPVFTVFKDNEILYRAQ LASEDTNAQKTTITNCFLLKNKIWCISLVEIYDTGDNVIRPKLFAVKIPEQ CT	gb:AF212302  生物:ニパウイ ルス 株名:不明 -AF212302 タ ンパク質名:結 合糖タンパク 質 遺伝子記 号:G (Uniprot Q9IH62)

10

20

30

40

50

29	<p>MLSQLQKNYLDNSNQOGDKMNNPDKKLSVNFNPLELDKGQKDLNKSY  YVKNKNYNVSNLLNESLHDIKFCIYCFSLIIITINIITISIVIT  RLKVHEENGMESPQLQSIQDSLSSLTNMINTEITPRIGILVTATSV  TLSSSINYVGTKTNQLVNELKDYITKSCGFKVPELKLHECNISCADP  KISKSAMYSTNAYAELAGPPKIFCKSVSKDPDFRLKQIDYVIPVQQD  RSICMNNPLLDISDGFFTYIHYEGINSCKKSDSFKVLLSHGEIVDRG  DYRPSLYLLSSHYPYSMQVINCVPTCNQSSFFVCHISNNTKTLDN  SDYSSDEYYITYFNGIDRPKTKKIPINNMTADNRYIHFTFSGGGVC  LGEFIIIPVTTVINTDVFTHDYCESFNCSVQTGKSLKEICSESLRSP  TNSRYNLNGIMIISQNNMTDFKIQLNGITYNKLSFGSPGRLSKTLG  QVLYYQSSMSWDTYKAGFVEKWKPF'PNWMNNTVISRPNQGNCPRY  HKCEICYGGTYNDIAPLDLGKDMYVSVILDSQLAENPEITVFNST  TILYKERVSKDELNTRSTTTSCFLFLDEPWCISVLETNRFNGKSIRP  EIYSYKIPKYC</p>	<p>gb:JQ001776:817  0-10275 生物:セ  ダーウイルス 株  名:CG1a タンパ  ク質名:結合糖  タンパク質 遺伝  子記号:G</p>
30	<p>MPQKTVEFINMNSPLERGVSTLSDKKTLNQSKITKQGYFGLGSHSER  NWKKQKNQNDHYMTVSTMIILEILVVLGIMFNLIIVLTMVYYQNDNINQ  RMAELTSNITVNLNLNQLTNKIQREIIPRITLIDTATTITIPSAIT  YILATLTTRISELLPSINQKCEFKTPTLVLNDCRINCTPPLNPSDGV  KMSSLATNLVAHGSPPCRNFSVPTIYYYRIPGLYNRTALDERCILN  PRLTISSTKFAYVHSEYDKNCTRGFKYELMTFGEILEGPEKEPRMF  SRSFYSPTNAVNYHSCTPIVTVNEGYFLCLECTSSDPLYKANLSNST  FHLVILRHNKDEKIVSMPSFNLSTDQEYVQIIPAEGGTAESGNLYF  PCIGRLLHHRVTHPLCKKSNCSRTDESCLKSYNQSPQHVVNCL  IRIRNAQRDNPTWDVITVDLTNTYPGSRRIFGSFSKPMLYQSSVSW  HTLLQVAEITDLDKYQLDWLDTPIYSRPGGSECPFGNYCPTVCWEGT  YNDVYSLTPNNDLFVTVYLKSEQVAENPYFAIFSRDQILKEFPLDAW  ISSARTTTISCFMFNNEIWCIAALEITRLNDDIIRPIYYSFWLPTDC  RTPYPHTGKMTRVPLRSTYNY</p>	<p>gb:NC_025256:9  117-11015 生物:  コウモリパラミ  クソウイルス  Eid_hel/GH-  M74a/GHA/2009   株  名:BatPV/Eid_he  l/GH-  M74a/GHA/2009   タンパク質名:  糖タンパク質 遺  伝子記号:G</p>

10

20

30

40

50

31	<p>MATNRDNTITSAEVSQEDKVKKYYGVETAEKVADSISGNKVFILMNT  LLILTGAIITITLNLITNLTAAKSQQNMLKIIQDDVNAKLEMFVNLDO  LVKGEIKPKVSLINTAVSVSIPGQISNLQTKFLQKYVYLEESITKQC  TCNPLSGIFPTSGPTYPPDTDKPDDDTDDDKVDTTIKPIEYPKPDGC  NRTGDHFTMEPGANFYTVPNLGPASSNSDECYTNPSFSIGSSIYMF5  QEIRKTDCTAGEILSIQIVLGRIVDKGQOGPQASPLLVWAVPNPKII  NSCAVAAGDEMGLVCSVTLTAASGEPIPHMFDGFWLYKLEPDTEVV  SYRITGYAYLLDKQYDSVFIGKGGGIQKGNLYFQMYGLSRNRQSF6  ALCEHGSCLGTGGGGYQVLCDRAWMSFGSEESLITNAYLKVNDLASG  KPVIIIGQTFPPSDSYKGSNGRMYTIGDKYGLYLAPSSWNRYLRFGIT  PDISVRSTTWLKSQDPIMKILSTCTNTDRDMCPEICNTRGYQDIFPL  SEDSEYYTYIGITPNNGGTKNFVAVRDSDGHASIDILQNYYSITSA  TISCFMYKDEIWCIAITEGKKQKDNPQRIYAH5YKIRQMCYNMKSAT  VTVGNAKNITIRRY</p>	<p>gb:NC_025352:8  716-11257 生  物:Mojiang ウイ  ルス 株  名:Tongguan1 タ  ンパク質名:結  合糖タンパク質   遺伝子記号:G</p>	10
32	<p>FNTVIALLGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG  IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS  KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR  EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV  VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI  IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE  FYYVLCAVST VGDPILNSTY WSGSLMTRL AVKPKSNGGG  YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF  LVRTEFKYND SNCPITKCQY SKPENCRISM GIRPNSHYIL  RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSL6  QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG  QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN  QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK  NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC</p>	<p>NivG タンパク  質結合糖タンパ  ク質   細胞質尾部を有  しない  Uniprot Q9IH62</p>	20

30

40

50

33	<p>FNTVIALLGSI IIIVMNIMI QNYTRTTDNQ ALIKESLQSV QQQIKALTDK  IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTSSINE NVNDKCKFTL  PPLKIHECNI SCPNPLPFRE YRPISQGVSD LVGLPNQICL QKTTSTILKP  RLISYTLPIN TREGVCITDP LLAVDNGFFA YSHLEKIGSC TRGIAKQRII  GVGEVLDRGD KVPSMFMTNV WTPPNPSTIH HCSSTYHEDF YYTLCAVSHV  GDPILNSTSW TESLSLIRLA VRPKSDSGDY NQKYIAITKV ERGKYDKVMP  YGPSGIKQGD TLYFPAVGFL PRTEFQYNDS NCPIIHCKYS KAENCRLSMG  VNSKSHYILR SGLLKYNLSL GGDIILQFIE IADNRLTIGS PSKIYNSLGQ  PVFYQASYSW DTMIKLGDVD TVDPLRVQWR NNSVISRPGQ SQCPRFNVCP  EVCWEGTYND AFLIDRLNWW SAGVYLSNQ TAENPVFAVF KDNEILYQVP  LAEDDTNAQK TITDCFLLEN VIWCISLVEI YDTGDSVIRP KLFVAVKIPAQ  CSES</p>	<p>ヘンドラ  ウイルス  Gタンパ  ク質  Uniprot  O89343  細胞質尾  部を有し  ない</p>	10
34	<p>MVVILDKRCY CNLLILILMI SECSVG</p>	<p>シグナル  配列</p>	
35	<p>MKVR FENTTSDK GK IPSKVIKSY Y GTMDIKKINE GLLDSKILSA  FNTVIALLGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD  KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT  LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK  PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI  IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST  VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM  PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM  GIRPNSHYIL RSGLLKYNLS DGENPKV VFI EISDQRLSIG SPSKIYDSL G  QPVFYQASFS WDTMIKFGDV LTVNPLV VNW RNNTVISRPG QSQCPRFNTC  PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA  QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE  QCT</p>	<p>NiVG タ  ンパク質  結合糖タ  ンパク質  切断型 Δ5</p>	20
36	<p>MNTTSDK GK IPSKVIKSY Y GTMDIKKINE GLLDSKILSA FNTVIALLGS  IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV  SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN  ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV  VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG  DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPILNSTY  WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPSGIKQG  DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL  RSGLLKYNLS DGENPKV VFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS  WDTMIKFGDV LTVNPLV VNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN  DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ  KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT</p>	<p>NiVG タ  ンパク質  結合糖タ  ンパク質  切断型  Δ10</p>	30

37	<p>MK GK IPSKVIKSY Y GTMDIKKINE GLLDSKILSA FNTVIAL LGS  IV IIVMNIMI IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV  SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN  ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV  VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG  DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY  WSGSLMMTRL AVKPKSNGGG YNQHQ LALRS IEKGRYDKVM PYGPSGIKQG  DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL  RSGLLKYNLS DG ENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS  WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN  DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ  KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT</p>	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ15	10
38	<p>MSKVIKSY Y GTMDIKKINE GLLDSKILSA FNTVIAL LGS IV IIVMNIMI  IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT  IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR  EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD  PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN  VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY WSGSLMMTRL  AVKPKSNGGG YNQHQ LALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF  LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS  DG ENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV  LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW  ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK  NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT</p>	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ20	20
39	<p>MSYY GTMDIKKINE GLLDSKILSA FNTVIAL LGS IV IIVMNIMI  IQNYTRSTDN QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT  IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR  EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD  PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLD RG DEVPSLFMTN  VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY WSGSLMMTRL  AVKPKSNGGG YNQHQ LALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF  LVRTEFKYND SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS  DG ENPKVVFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS WDTMIKFGDV  LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW  ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK  NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT</p>	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ25	30

40

50

40	MTMDIKKINE GLLDSKILSA FNTVIALLGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRSLM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QCT	NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ30	10
41	GGGGGS	ペプチド リンカー	
42	(GGGGGS) <sub>n</sub> ( <sub>n</sub> は1~10である)	ペプチド リンカー	20
43	GGGGGS	ペプチド リンカー	
44	PAENKKVR FENTTSKDGK IPSKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALLGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCRSLM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVWNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タ ンパク質 結合糖タ ンパク質 (602 アミ ノ酸) N 末端メ チオニン を有しな い	30

45	<p>KVR FENTTSDK GK IPSKVIKSY Y GTMDIKKINE GLLDSKILSA  FNTVIAL LGS IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD  KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT  LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK  PKLISYTL PV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI  IGVGEVLDRG DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST  VGDPILNSTY WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM  PYGPSGIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM  GIRPNSHYIL RSGLLKYNLS DGENPKV VFI EISDQRLSIG SPSKIYDSL G  QPVFYQASFS WDTMIKFGDV LTVNPLV VNW RNNTVISRPG QSQCPRFNTC  PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA  QLASEDTNAQ KTITNCFL LK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE  QC</p>	<p>NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ5 N 末端メ チオニン を有しな い</p>	10
46	<p>NTTSDK GK IPSKVIKSY Y GTMDIKKINE GLLDSKILSA FNTVIAL LGS  IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV  SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN  ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTL PV  VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG  DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY  WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG  DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM GIRPNSHYIL  RSGLLKYNLS DGENPKV VFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS  WDTMIKFGDV LTVNPLV VNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN  DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ  KTITNCFL LK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC</p>	<p>NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ10 N 末端メ チオニン を有しな い</p>	20
47	<p>KGK IPSKVIKSY Y GTMDIKKINE GLLDSKILSA FNTVIAL LGS  IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV  SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN  ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTL PV  VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRG  DEVPSLFMTN VWT PPNPNTV YHCSAVYNNE FYYVLC AVST VGDPILNSTY  WSGSLMMTRL AVKPKSNGGG YNQHLALRS IEKGRYDKVM PYGPSGIKQG  DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCR LSM GIRPNSHYIL  RSGLLKYNLS DGENPKV VFI EISDQRLSIG SPSKIYDSL G QPVFYQASFS  WDTMIKFGDV LTVNPLV VNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN  DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ  KTITNCFL LK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC</p>	<p>NiVG タ ンパク質 結合糖タ ンパク質 切断型 Δ15 N 末端メ チオニン を有しな い</p>	30

40

50



48	SKVIKSYY GTMDIKKINE GLLDSKILSA FNTVIALLSG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRLG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPGSIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タンパク質 結合糖タンパク質 切断型 Δ20 N 末端メチオニンを有しない	10
49	SYG GTMDIKKINE GLLDSKILSA FNTVIALLSG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRLG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPGSIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タンパク質 結合糖タンパク質 切断型 Δ25 N 末端メチオニンを有しない	20
50	TMDIKKINE GLLDSKILSA FNTVIALLSG IVIIVMNIMI IQNYTRSTDN QAVIKDALQG IQQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF AYSHLERIGS CSRGVSKQRI IGVGEVLDRLG DEVPSLFMTN VWTTPNPNTV YHCSAVYNNE FYYVLCVAVST VGDPIILNSTY WSGSLMMTRL AVKPKSNGGG YNQHQALALRS IEKGRYDKVM PYGPGSIKQG DTLYFPAVGF LVRTEFKYND SNCPITKCQY SKPENCLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLDSN QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE IYDTGDNVIR PKLFAVKIPE QC	NiVG タンパク質 結合糖タンパク質 切断型 Δ30 N 末端メチオニンを有しない	30

40

51	<p>KKINEGLLLDSKILSA FNTVIALLGS IVIIVMNIMI IQNYTRSTDN  QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS  KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS  NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF  AYSHLERIGS CSRGVSKQRI IGVGEVLDRG DEVPSLFMTN VWTTPNPNTV  YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG  YNQHQLALRS IEKGRYDKVM PYGPSGIKQG DTLYFPAVGF LVRTEFKYND  SNCPITKCQY SKPENCRLSM GIRPNSHYIL RSGLLKYNLS DGENPKVVFI  EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVVNW  RNNTVISRPG QSQCPRFNTC PAICAEGVYN DAFLIDRINW ISAGVFLDSN  ATAANPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLLK NKIWCISLVE  IYDTGDNVIR PKLFAVKIPE QCT</p>	<p>NiVG タ  ンパク質  結合糖タ  ンパク質  切断及び  変異型  (E501A、  W504A、  Q530A、  E533A)  NiV G タ  ンパク質  (Gc Δ 34)  N 末端メ  チオニン  を有しな  い</p>
----	---	---

10

20

30

40

50

52	<p>MADSKLVSL NNNLSGKIKD QGKVIKNIYYG TMDIKKINDG LLDSKILGAF          NTVIALLGSI IIIVMNIMII QNYTRTTDNQ ALIKESLQSV QQQIKALTDK          IGTEIGPKVS LIDTSSTITI PANIGLLGSK ISQSTSSINE NVNDKCKFTL          PPLKIHECNI SCPNPLPFRE YRPISQGVSD LVGLPNQICL QKTTSTILKP          RLISYTLPIN TREGVCITDP LLAVDNGFFA YSHLEKIGSC TRGIAKQRII          GVGEVLDRGD KVPSPMFTNV WTPPNPSTIH HCSSTYHEDE YYTLCAVSHV          GDPILNSTSW TESLSLIRLA VRPKSDSGDY NQKYIAITKV ERGKYDKVMP          YGPGSIKQGD TLYFPAVGFL PRTEFQYNSD NCPIIHCKYS KAENCRLSMG          VNSKSHYILR SGLLKYNLSL GGDIIHQFIE IADNRLTIGS PSKIYNSLQ          PVFYQASYSW DTMIKLGDVD TVDPLRVQWR NNSVISRPGQ SQCPRENVCP          EVCWEGTYND AFLIDRLNWW SAGVYLNNSQ TAENPVFAVF KDNEILYQVP          LAEDDTNAQK TITDCFLEN VIWCISLVEI YDTGDSVIRP KLFAVKIPAQ          CSES</p>	<p>ヘンドラ ウイルス Gタンパ ク質 Uniprot O89343  N末端メ チオニン を有しな い</p>
53	<p>KKINEGLLDKILSA FNTVIALLGS IVIIVMNIMI IQNYTRSTDN          QAVIKDALQG IQQIKGLAD KIGTEIGPKV SLIDTSSTIT IPANIGLLGS          KISQSTASIN ENVNEKCKFT LPPLKIHECN ISCPNPLPFR EYRPQTEGVS          NLVGLPNNIC LQKTSNQILK PKLISYTLPV VGQSGTCITD PLLAMDEGYF          AYSHLERIGS CSRGVSKQRI IGVEVLDRG DEVPSLFMTN VWTTPNPNTV          YHCSAVYNNE FYYVLCVAVST VGDPILNSTY WSGSLMMTRL AVKPKSNGGG          YNQHQALALRS IEKGRYDKVM PYGPGSIKQG DTLYFPAVGFL LVRTEFKYND          SNCPITKCQY SKPENCRISM GIRPNSHYIL RSGLLKYNLS DGENPKVVF          EISDQRLSIG SPSKIYDSLQ QPVFYQASFS WDTMIKFGDV LTVNPLVNVW          RNNTVISRPG QSQCPRFNTC PEICWEGVYN DAFLIDRINW ISAGVFLEDSN          QTAENPVFTV FKDNEILYRA QLASEDTNAQ KTITNCFLK NKIWCISLVE          IYDTGDNVIR PKLFAVKIPE QCT</p>	<p>NiVGタ ンパク質 結合糖タ ンパク質 切断型 (Gc Δ 34) N末端メ チオニン を有しな い</p>
54	<p>LSQLQKNYLDNSNQQGDKMNPDKLSVNFNPLELDKGQKDLNKSYYVKNKNYNV          SNLLNESLHDIKFCIYCFSLIIITIIINIITISIVITRLKVHEENNGMESPNLQ          SIQDSLSSLTNMINTETPRIGILVTATSVTLSSSINYVGTKTNLVNELKDYIT          KSCGFKVPELKLHECNISCADPKISKSAMYSTNAYAEAGPPKIFCKSVSKDPDF          RLKQIDYVIPVQQDRSICMNNPLLDISDGFFTYIHYEGINSCKKSDSFKVLLSHG          EIVDRGDYRPSLYLLSSHYHPYSMQVINCVPTCNQSSFFVCHISNNTKTLDNSD          YSSDEYYITYFNGIDRPKTKKIPINMNTADNRYIHFTFSGGGGVCLGEEFIIPVT          TVINTDVFTHDYCESFNCSVQTGKSLKEICSESLRSPTNSSRYNLNGIMIISQNN          MTDFKIQNLGITYNKLSFGSPGRLSKTLGQVLYYQSSMSWDTYLKAGFVEKWKPF          TPNWMNNTVISRPNQGNCPRYHKCPEICYGGTYNDIAPLDLGKDMYVSVILDSQ          LAENPEITVFNSTTILYKERVSKDELNTRSTTTSCFLFLDEPWCISVLETNRFNG          KSIRPEIYSYKIPKYC</p>	<p>gb:JQ0017 76:8170- 10275 生 物:セダー ウイルス  株 名:CG1a  タンパク 質名:結合 糖タンパ ク質 遺伝 子記号:G N末端メ チオニン を有しな い</p>

10

20

30

40

55	<p>PQKTVEFINMNSPLERGVSTLSDKKTLNQSKITKQGYFGLGSHSERNWKKQKN  QNDHYMTVSTMILEILVVLGIMFNLIIVLTMVYYQNDNINQRMAELTSNITVLN  LNLNQLTNKIQREIIPRITLIDTATTITIPSAITYILATLTTRISELLPSINQ  KCEFKTPTLVLNDCRINCTPPLNPSDGVKMSLATNLVAHGSPPCRNFSSVPT  IYYYRIPGLYNRTALDERCILNPRLTISSTKFAYVHSEYDKNCTRGFKYELM  TFGEILEGPEKEPRMFSRSFYSPNAVNYHSCTPIVTVNEGYFLCLECTSSDP  LYKANLSNSTFHLVILRHNKDEKIVSMPSFNLSTDQEYVQIIPAEGGGTAESG  NLYFPCIGRLLHKRVTHPLCKKSNCSRTDDESLKSYYNQGSPOHQVNVNCLIR  IRNAQRDNPTWDVITVDLTNTYPGSRSRIFGSFSKPMLYQSSVSWHTLLQVAE  ITDLDKYQLDWLDTPIYISRPGGSECPFGNYCPTVCWEGTYNDVYSLTPNNDLF  VTVYLKSEQVAENPYFAIFSRDQILKEFPLDAWISSARTTTISCFMFNNEIWC  IAALEITRLNDDIIRPIYYSFWLPTDCRTPYPHTGKMTRVPLRSTYNY</p>	<p>gb:NC_02  5256:9117  -11015 生  物:コウモ  リパラミ  クソウイ  ルス  Eid_hel/G  H-  M74a/GH  A/2009 株  名:BatPV/  Eid_hel/G  H-  M74a/GH  A/2009 タ  ンパク質  名:糖タン  パク質 遺  伝子記  号:G  N末端メ  チオニン  を有しな  い</p>
----	---	--

10

20

30

40

50

56	<p>ATNRDNTITSAEVSQEDKVKKYYGVETAEKVADSI SGNKVFILMNTLL  IILTGAIITITLNIITNLTAAKSQQNMLKIIQDDVNAKLEMFVNLDQLVK  GEIKPKVSLINTAVSVSIPGQISNLQTKFLQKYVYLEESITKQCTCNP  LSGIFPTSGPTYPPPTDKPDDDTDDDDKVDTTIKPIEYPKPDGCNRTGD  HFTMEPGANFYTVPNLGPASSNSDECYTNPSFSIGSSIYMFSQEIRKT  DCTAGEILSIQIVLGRIVDKGQQGPQASPLLVWAVPNPKIINSCAVAA  GDEMGWVLCVTLTAASGEPIPHMFDGFWLYKLEPDTEVVSYRITGYA  YLLDKQYDSVFIGKGGGIQKGNLDYFQMYGLSRNRQSFKALCEHGSCL  GTGGGGYQVLCRAVMSFGSEESLITNAYLKVNDLASGKPVIIIGQTFP  PSDSYKGSNGRMYTIGDKYGLYLAPSSWNRYLRFGITPDISVRSTTWL  KSQDPIMKILSTCTNTDRDMCPEICNTRGYQDIFPLSEDESEYYTYIGI  TPNNGGTKNFVAVRSDGHIASIDILQNYYSITSATISCFMYKDEIWC  IAITEGKKQKDNPQRIYAHSYKIRQMCYNMKSATVTVGNAKNITIRRY</p>	<p>gb:NC_025352:  8716-11257 生  物:Mojiang ウ  イルス 株  名:Tongguan1   タンパク質名:  結合糖タンパ  ク質 遺伝子記  号:G  N末端メチオ  ニンを有しな  い</p>
57	<p>DFDKLNKIGVVQGRVLNYKIKGDPMTKDLVLFIPNIVNITECVREPL  SRYNETVRRLLLPIHNMLGLYLNNNTNAKMTGLMIAGVIMGGIAIGIAT  AAQITAGFALYEAKKNTENIQKLTDSIMKTQDSIDKLTDSVGTSLILIL  NKLQTYINNQLVLPNLELLSCRQNKIEFDLMLTKYLVDLMTVIGPNINN  PVNKDMTIQSLSLFDGNYDIMMSELGYTPQDFLDLIESKSITGQIIY  VDMENLYVVIRTYLPTLIEVPDAQIYEFNKITMSSNGGEYLSTIPNFI  LIRGNYSNIDVATCYMTKASVICNQDYSLPMSQNLRSYQGETEYCP  VEAVIASHSRPFALTNGVIFANCINTICRCQDNGKTITQININQFVSMI  DNSTCNDVMVDKFTIKVGKYMGRKDINNINIQIGPQIIIDKVDLSNEI  NKMNQLKDSIFYLREAKRILDSVNISLISPSVQLFLIIISVLSFIIL  LIIIVYLYCKSKHSYKYNKFIDDPDYNDYKRERINGKASKSNNIYYV  GD</p>	<p>gb:JQ001776:6  129-8166 生物:  セダーウイル  ス 株名:CG1a   タンパク質名:  融合糖タンパ  ク質 遺伝子記  号:F(シグナル  配列を有しな  い)</p>

10

20

30

40

50

58	<p>SRALLRETDNYSNGLIVENLVRNCHHPSKNNLNYTKTQKRDSTIPYRVEERK  GHYPKIKHLIDKSYKHIKRGKRRNGHNGNIITIILLILLILILKTQMSEGAIHY  ETLSKIGLIKGITREYKVKGTPSSKDIVIKLIPNVTGLNKCTNISMENYKEQ  LDKILIPINNIIELYANSTKSAPGNARFAGVIIAGVALGVAAAAQITAGIAL  HEARQNAERINLLKDSISATNNAVAELQEATGGIVNVITGMQDYINTNLVPQ  IDKLQCSQIKTALDISLSQYYSEILTVFGPNLQNPVTTSMSIQAISQSFGGN  IDLNLLNLLGYTANDLLDLLESKSITGQITYINLEHYFMVIRVYYPIMTTISN  AYVQELIKISFNVDGSEWVSLVPSYILIRNSYLSNIDISECLITKNSVICRH  DFAMPMSYTLKECLTGDTEKCPREAVVTSYVPRFAISGGVIYANCLSTTCQC  YQTGKVIAQDGSQTLMMIDNQTC SIVRIEEILISTGKYLGSQEYNTMHVSVG  NPVFTDKLDITSQISNINQSIEQSKFYLDKSKAILDKINLNLIGSVPI SIF  IIAILSLILSIITFVIVMIIVRRYNKYTPLINSDPSSRRSTIQDVYIIPNPG  EHSIRSAARSIDRDRD</p>	<p>gb:NC_0252  56:6865-  8853 生物:  コウモリパ  ラミクソウ  イルス  Eid_hel/GH-  M74a/GHA/  2009 株  名:BatPV/Ei  d_hel/GH-  M74a/GHA/  2009 タンパ  ク質名:融  合タンパク  質 遺伝子記  号:F(シグ  ナル配列を  有しない)</p>
----	---	--

10

20

30

40

50

59	<p>ILHY EKLSKIGLVK GITRKYKIKS  NPLTKDIVIK MIPNVSNSVSK CTGTVMENYK SRLTGILSPI KGAIELYNNN  THDLVGDVKL AGVVMAGIAI GIATAAQITA GVALYEAMKN ADNINKLKSS  IESTNEAVVK LQETAECTVY VLTALQDYIN TNLVPTIDQI SCKQTEALD  LALSKYLSDL LFVFGPNLQD PVSNSMTIQA ISQAFGGNYE TLLRTLGYAT  EDFDDLLESD SIAGQIVYVD LSSYYIIVRV YFPILTEIQQ AYVQELLPVS  FNNDNSEWIS IVPNFVLIRN TLISNIEVKY CLITKKSVC NQDYATPMTA  SVRECLTGST DKCPRELVVS SHVPRFALSG GVLFANCISV TCQCQTTGRA  ISQSGEQTLL MIDNTTCTTV VLGNIISLQ KYLGSINYNS ESIAVGPPVY  TDKVDISSQI SSMNQLQOS KDYIKEAQKI LDTVNPSSLIS MLSMIILYVL  SIAALCIGLI TFISFVIVEK KRGNYRSLDD RQVRPVSNGD LYYIGT</p>	<p>ヘンドラ  ウイルス  Fタンパ  ク質  Uniprot  O89342  (シグナル  配列を有  しない)</p>	10	
60	<p>IHYDSLKVGVIKGLTYNYKIKGSPSTKLMVVKLIPNIDSVKNCTQKQYDEYKNL  VRKALEPVKMAIDTMLNNVKSGNKNYRFAGAIMAGVALGVATAATVTAGIALHRS  NENAQAIANMKSAIQNTNEAVKQLQLANKQTLAVIDTIRGEINNNIIPVINQLSC  DTIGLSVGIRLTQYYSEIITAFGPALQNPVNTRITIQAISSVFNGNFDELLKIMG  YTSGDLYEILHSELIRGNIIDVDVDAGYIALEIEFPNLTLVNAVQELMPISYN  IDGDEWVTLVPRFVLTRTTLLSNIDTSRCTITDSSVICDNDYALPMSHELIGCLQ  GDTSKCAREKVVSSYVPKFALSDGLVYANCLNTICRCMDTDPISQSLGATVSL  DNKRCSVYQVGDVLISVGSYLGDEYNADNVELGPPIVIDKIDIGNQLAGINQTL  QEAEDYIEKSEEFKGVNPSIITLGSVVLYIFMILIAIVSVIALVLSIKLTVKG  NVVRQQFTYTTQHVPSMENINYVSH</p>	<p>gb:NC_02  5352:5950  -8712 生  物:Mojian  g ウイル  ス株  名:Tonggu  an1 タン  パク質名:  融合タン  パク質 遺  伝子記  号:F(シグ  ナル配列  を有しな  い)</p>	20	30

61	MLFNLRLILLNNAAFRNGHNFMRNFRFCGQPLQNKVQLKGRDLLTLKNFT GEEIKYMLWLSADLKFRKQKGEYPLLLQGKSLGMIFEKRSTRTRLSTETG FALLGGHPCFLTTQDIHLGVNESLTDARVLSMADAVLARVYKQSDLDL LAKEASIPINGLSLDLYHPIQILADYTLQEHYSSLKGLTSLWIGDGNLHS IMMSAAKFGMHLQAATPKGYEPDASVTKLAEQYAKENGTKLLLNDP AAHGGNVLITDTWISMGQEEKKRLQAFQGYQVTMKTAKVAASDWF LHCLPRKPEEVDDEVFYSRSLVFPEAENRKWTIMAVMVSLLDYSPQLQ KPKF	OTC
62	MTRILTAFKVVRTLKTGFGFTNVTAHQKWKFSRPGIRLLSVKAQTAHIVL EDGTMKMGYSFGHPSSVAGEVVFNTGLGGYPEAITDPAYKGQILTMANPI IGNGGAPDTTALDELGLSKYLESNGIKVSGLLVLDYSKDYNHWLTKSLG QWLQEEKVPAIYGVDTRMLTKIIRDKGTMLGKIEFEGQPVDVDPNKQNL IAEVSTKDVKVYGGKGNPTKVVAVDCGIKNNVIRLLVKRGAEVHLPWNH DFTKMEYDGLIAGGPGNPALAEPLIQNVRKILES DRKEPLFGISTGNLITG LAAGAKTYKMSMANRGQNQPVLNITNKQAFITAQNHGYALDNTLPAGW KPLFVNVDQTNNEGIMHESKPFFAVQFHPEVTPGPIDTEYLFDSFFSLIKKG KATTITSVLPKALVASRVEVSKVLILGSGLSIGQAGEFDYSGSQAVKAM KEENVKTVLMNPNIASVQTNEVGLKQADTVYFLPITPQFVTEVIKAEQPD GLILGMGGQTALNCGVELFKRGVLKEYGVKVLGTSVESIMATEDRQLFS DKLNEINEKIAPSF AVESIEDALKAADTIGYPVMIRSAYALGGLGSGICPNR ETLMDLSTKAFAMTNQILVEKSVTGWKEIEYEVVRDADDNCVTVCNME NVDAMGVHTGDSVVVAPAQTL SNAEFQMLRRTSINVVRHLGIVGECNIQ FALHPTSMEYCHIEVNARLSRSSALASKATGYPLAFIAAKIALGIPLPEIKNV VSGKTSACFEPSLDYMTKIPRWDLDRFHGTSSRIGSSMKSVGEVMAIGR TFEESFQKALRMCHPSIEGFTPRLPMNKEWPSNLDLRKELSEPSSTRIY AIA KAIDDNMSLDEIEKLYIDKWFLYKMRDILNMEKTLKGLNSESMTEETLK RAKEIGFSDKQISKCLGLTEAQTRELRLKKNHPWVKQIDTLAAEYPSVTN YLYVTYNGQEHDVNFDDHGMMLVLCGPYHIGSSVEFDWCAVSSIRTLRQ LGKKTVVVNCNPETVSTDFDECDKLYFEELSLERILDIYHQEACGGCISV GGQIPNNLAVPLYKNGVKIMGTSPLQIDRAEDRSIFSAVLDELKVAQAPW KAVNTLNEALEFAKSVDYPCLLRPSYVLSGSAMNVVFSEDEMKKFLEEA TRVSQEHVVLTKFVEGAREVEMDAVGKDG R VISHAISEHVEDAGVHSG DATLMLPTQTISQGAIEKVKDATR K IAKAFAISGPFNVQFLVKGNDVLVIE CNLRASRSPFVSKTLGVDFIDVATKVMIGENVDEKHLPTLDHPPIPADYV AIKAPMFSWPRLRDADPILRCEMASTGEVACFGEGIH TAFLKAMLSTGFKI PQKGILIGIQSFRPRFLGVAEQLHNEGFKLFATEATSDWLNANNVPATPV AWPSQEGQNPSLSSIRKLIRDGSIDLVINLPNNNTK FVHDNYVIRRTAVDS GIPLLTNFQVTKLFAEAVQKSRKVDSKSLFH YRQYSAGKAA	CPS1

10

20

30

40

50



63	MATALMAVVLRAAAVAPRLRGRGGTGGARRLSCGARRRAARGTSPGRR LSTAWSQPQPPEEYAGADDVSQSPVAEEPSWVPSRPPVPHESPEPPSGR SLVQRDIQAFLNQCASPGEARHWLTQFQTCHHSADKPFVAVIEVDEEVLK CQQGVSSALAFALAFQRMMDKPLVVLGLPAPTAPSGCLSFWEAKAQLAK SCKVLVDALRHNAAAAAPFFGGGSVLRAAEPAPHASYGGIVSVETDLLQ WCLESGSIPILCPIGETAARRSVLLDSLEVTASLAKALRPTKIIFLNNTGGLR DSSHKVLSNVNLPADLDLVCNAEWVSTKERQQMRLIVDVL SRLPHHSSA VITAASTLLTELF SNKSGT LFKNAERMLRVRS LDKLDQGRLVDLVNASF GKCLRDDYLASLRPRLHSIYVSEGYNAAILTMEPVLGGTPYLDKFFVSS SRQGGQSGQMLWECLRRDLQTLFWRSRVTNPINPWYFKHSDGSFSNKQ WIFFWFG LADIRDSYELVN HAKGLPDSFHKPASDPGS	NAGS
64	MAVAIAAARVWRLNRGLSQAALLLRQPGARGLARSHPPRQQQFSSLD DKPQFPGASAEFIDKLEFIQPNVISGIPIYRVM DRQGQIINPSEDPHLPKEKV LKLKYSMTLLNTMDRILYESQRQGRISFYMTNYGEEGTHVGSAAALDNT DLVFGQYREAGVLMYRDYPLELFMAQCYGNISDLGKGRQMPVHYGCKE RHFVTISSPLATQIPQAVGAAYA AKRANANRVVICYFGEGAASEGDAHAG FNFAATLECPHIFCRNNGYAISTPTSEQYRGDGAARGPGYGIMSIRVDGN DVFAVYNATKEARRRAVAENQPFLIEAMTYRIGHHSTSDSSAYRSVDEV NYWDKQDHPISRLRHYL LSQGWDEEQEKA WRKQSRRKVM EAFEQAE RKPKNPNLLFSVYQEMPAQLRKQ QESLARHLQTYGEHYPLDHF DK	BCKDHA
65	MAVVAAAAGWLLRLRAAGAEGHWRR LPGA LARGFLHPAATVEDAAQ RRQVAHFTFQPDPEPREYQGTQKMNL FQSVTSALDNSLAKDPTAVIFGED VAFGGVFRCTVGLRDKY GKDRVFNTP LCEQGIVGFGIGIAVTGATAIAEIQ FADYIFPAFDQIVNEAAKYRYS GDLFNCGLTIRSPWGC VGHGALYHSQ SPEAFFAHCPGKIVIPRSPFQAK GLLSCIEDKNPCIFFEPKILYRAAAEEV PIEPYNIPLSQA EVIQEGSDVTLVAWGTQVHVIREVASM AKEKLGVSCEVI DLRTHIPWDVDTICKSVIKTG RLLISHEAPLTGGFASEISSTVQEECFLNLEA PISRVCGYDTPFP HIFEPFYIPDKWKCYDALRKM INY	BCKDHB
66	MAAVRMLRTWSRNAGK LICVRYFQTCGNVHVLKPNYVCF FGYPSFKYS HPHHFLKTTAALRGQVVQFKLSDIGEGIREVTVKEWYVKEGDTVSQFDSI CEVQSDKASVTITSR YDGVIKKLYNLDDIAYVGKPLVDIETEAL KDSEED VVETPAVSHDEH THQEIKGRKTLATPAVRRLAMENNIK LSEVVGSGKDG RILKEDILNYLEKQTGAILP PSPKVEIMPPP KPKDMTVPILVSKPPVFTGK DKTEPIKGFQKAMVKTMSAALKIPHFGYCDEIDLTEL VKLREELKPIAFAR GIKLSFMPF LKAASLGLLQFPILNASVDENCQNITYKASHNIGIAMDTEQ GLIVPNVKNVQICSIFDIATELNRLQKLG SVGQLSTTDLTGGTFTLSNIGSIG GTFAKPVMPPPEVAIGALGSIKAI PRFNQKGEVYKAQIMNVSWSADHRVID GATMSRFSNLWKSYLENPAFMLLDLK	DBT

10

20

30

40

50

67	<p>MQSWSRVYCSLAKRGHFNRIHGLQGLSAVPLRTYADQPIDADVTVIGSG          PGGYVAAIKAAQLGFKTVICIEKNETLGGTCLNVGCIPSKALLNNSHYHYM          AHGKDFASRGIEMSEVRLNLDKMMEQKSTAVKALTGGIAHLFKQNKVV          HVNGYGKITGKNQVTATKADGGTQVIDTKNLIATGSEVTPFPGITIDEDTI          VSSTGALSLLKKVPEKMMVIGAGVIGVELGSVWQRLGADVTAVEFLGHVG          GVGIDMEISKNFQRILQKQGFKFLNNTKVTGATKKSDBGKIDVSIEAASGG          KAEVITCDVLLVCIGRRPFTKNLGLEELGIELDPRGRIPVNTFRFQTKIPNIYA          IGDVVAGPMLAHKAEDEGIICVEGMAGGAVHIDYNCVPSVIYTHPEVAW          VGKSEEQLKEEGIEYKVGKFPFAANSRAKTNADTDGMVKILGQKSTDRV          LGAHILGPGAGEMVNEAALALEYGASCEDIARVCHAHPTLSEAFREANLA          ASFGKSINF</p>	DLD
68	<p>MLRAKNQLFLLSPHYLRQVKESSGSRLIQRLLHQQPLHPEWAALAKK          QLKGGKNPEDLIWHTPEGISIKPLYSKRDTMDLPEELPGVKPFTRGPYPTMY          TFRPWTIRQYAGFSTVEESNKFYKDNKAGQQGLSVAFDLATHRGYSDSN          PRVRGDVGMAGVAIDTVEDTKILFDGIPLEKMSVSMTMNGAVIPVLANFI          VTGEEQGVPEKLTGTIQNDILKEFMVRNTYIFPEPSMKIADIFEYTAKH          MPKFNSISISGYHMQEAGADAILELAYTLADGLEYSRTGLQAGLTIDEFAP          RLSFFWVIGMNFYMEIAKMRAGRRLWAHLIEKMFQPKNSKSLLLRAHCQ          TSGWSLTEQDPYNNIVRTAIEAMAAVFGGTQSLHTNSFDEALGLPTVKS          RIARNTQIIQEESEGIPKVADPWGGSYMMECLTNDVYDAALKLINEIEEMG          GMAKAVAEGIPKLRIECAARRQARIDSGSEVIVGVNKYQLEKEDAVEVL          AIDNTSVRNRQIEKLLKIKSSRDQALAECLAALTECAASGDGNILALAV          DASRARTVGEITDALKKVFGEHKANDRMVSGAYRQEFGESKEITSAIKR          VHKFMEREGRRPRLLVAKMGQDGHDRGAKVIATGFADLGFVDIGPLFQ          TPREVAQQAVDADVHAVGISTLAAGHKTLPVPELIKELNSLGRPDILVMCG          GVIPPQDYEFLEFVGVSNVFGPGTRIPKAAVQVLDDIEKCLEKKQQSV</p>	MUT
69	<p>MPMLLPHPHQHFLKGLLRAPFRCYHFIFHSSTHLGSGIPCAQPFNSLGLHC          TKWMLLSDGLKRKLCVQTTLKDHTEGLSKDEQRFVDKLYTGLIQGQRAC          LAEAITLVESTHSRKKELAQVLLQKVLLYHREQEQSNGKPLAFRVGLSG          PPGAGKSTFIEYFGKMLTERGHKLSVLAVDPSSTSGSLLGDKTRMTEL          SRDMNAYIRSPTRGTLGGVTRTTNEAILLCEGAGYDIILTIETVGVGQSEFA          VADMVDMFVLLLPPAGGDELQGIKRGHEMADLVAVTKSDGDLIVPARRI          QAEYVSALKLLRKRQVWPKVIRISARSGEGISEMWDKMKDFQDLMLA          SGELTAKRRKQKQVWMWNLIQESVLEHFRTHPTVREQIPLEQKVLIGAL          SPGLAADFLLKAFKSRD</p>	MMAA
70	<p>MAVCGLGSRLGLGSRLGLRGCFGAARLLYPRFQSRGPQGVEDGDRPQPS          SKTPRIPKIYTKTGDKGFSSTFTGERRPKDDQVFEAVGTTDELSSAIGFALE          LVTEKGHTFAEELQKIQCTLQDVGSAALATPCSSAREHLKYTTFKAGPILE          LEQWIDKYTSQLPPLTAFILPSGGKISSALHFCRAVCRAERRVPLVQMG          ETDANVAKFLNRLSDYLFTLARYAAMKEGNQEKIYMKNDPSAESEGL</p>	MMAB

10

20

30

40

50

71	MFDRAKLPFLQSCHLRMLTDPVDQCVAYHLGRVRESLPELQIEIIADYEV HPNRRPKILAQTAHVAGAAAYYQRQDVEADPWGNQRISGVCIHPRFGG WFAIRGVVLLPGIEVDPDLPKRPHDCVPTRADRIALLEGFNHWRDWTYR DAVTPQERYSEEQKAYFSTPPAQLALLGLAQPSEKPSPPDLPTTPAPK KPGNPSRARSWLSRVSPASP	MMACH C
72	MANVLCNRARLVSYLPGFCSLVKRVVNPKAFSTAGSSGSDESHVAAAPP DICSRTVWPDETMGPFQDQRFQPLGNIGFDCHLNGTASQKSLVHKTL PDVLAEPLSSERHEFVMAQYVNEFQGNDAPEVEQINSAEYFESARVECA IQTPELLRKDFESLFPEVANGKLMILVTQKTKNDMTVWSEEVEIEREVL LEKFINGAKEICYALRAEGYWADFIDPSSGLAFFGPYTNNTLFETDERYRH LGFSVDDLGCCKVIRHSLWGTHVVVGSIFTNATPDSHIMKKLSGN	MMADH C
73	MARVLKAAAANAVGLFSRLQAPIPTVRASSTSQPLDQVTGSVWNLGRLN HVAIAVPDLEKAAAFYKNILGAQVSEAVPLPEHGVSVVFNLGNTKMEL LHPLGRDSPIAGFLQKNKAGGMHHICIEVDNINAAVMDLKKKKIRSLSEE VKIGAHGKPVIFLHPKDCGGVLVELEQA	MCEE
74	MAGFWVGTAPLVAAGRGRWPPQQLMLSAALRTLKHVLYYSRQCLMV SRNLGSGYDPNEKTFDKILVANRGEIACRVIRTCKKMGIKTVAIHSDVD ASSVHKMADEAVCVGPAPTSKSYLNMDAIMEAIKKTRAQAVHPGYGFL SENKEFARCLAAEDVVFIPDTHAIQAMGDKIESKLLAKKAENVNTIPGFD GVVKDAEEAVRIAREIGYPVMIKASAGGGGKGMRIA WDDEETRDGFRLS SQEAASSFGDDRLIEKFIDNPRHIEIQVLGDKHGNALWLNRECSIQRN QKVVEEAPSIFLDAETRRAMGEQAVALARAVKYSSAGTVEFLVDSKKNF YFLEMNTRLQVEHPVTECITGLDLVQEMIRVAKGYPLRHKQADIRINGWA VECRVYAEDPYKSFGLPSIGRLSQYQEPLHLPGVRVDSGIQPGSDISIYDP MISKLITYGSDRTEALKRMADALDNYVIRGVTHNIALREVIINSRFVKGD ISTKFLSDVYPDGFKGHMLTKSEKNQLLAIASSLFAFQLRAQHFQENSR MPVIKPDIANWELSVKLDKVTHTVVASNNGSVFSVEVDGSKLNVTSTWN LASPLLSVSDGTQRTVQCLSREAGGNMSIQFLGTVYKVNILTRLAAELN KFMLEKVTEDTSSVLRSPMPGVVAVSVKPGDAVAEGQEICVIEAMKMQ NSMTAGKTGTVKSVCQAGDTVGEGLLVELE	PCCA
75	MAAALRVAAVGARLSVLAASGLRAAVRSLCSQATSVNERIENKRRTALLG GGQRRIDAQHKRGLTARERISLLLDPGSFVESDMFVEHRCADFGMAAD KNKFPGDSVVTGRGRINGRLVYVFSQDFTVFGGSLSGAHAQKICKIMDQA ITVGAPVIGLNDSGGARIQEGVESLAGYADIFLRNVTASGVIPQISLIMGPC AGGAVYSPALTDFTFMVKDTSYLFITGPDVVKSVTNEDVTQEELGGAKT HTMSGVAHRAFENDVDALCNLRDFNYLPLSSQDPAPVRECHDPSDRL VPELDTIVPLESTKAYNMVDIIHSVVDEREFFEIMPNYAKNIIVGFARMNG RTVGIVGNQPKVASGCLDINSSVKGARFVRFCDAFNIPLITFVDVPGFLPG TAQEYGGIIRHGAKLLYAFEAATVPKVTVITRKAYGGAYDVMSSKHLCG DTNYAWPTAEIAVMGAKGAVEIIFKGHENVEAAQAEYIEKFANPFPAVR GFVDDIIQPSSTRARICCDLDVLASKKVQRPWRKHANIPL	PCCB

10

20

30

40

50

76	<p>MAVESQGRPLVLGLLCVLGPVVS HAGKILLIPVDGSHWLSMLGAIQQL          QQRGHEIVVLAPDASLYIRDGAFYTLKTYPPVFQREDEVKESFVSLGHNVF          ENDSFLQRVIKTYKKIKKDSAMLLSGCSHLLHNKELMASLAESSFDVMLT          DPFLPCSPIVAQYLSLPTVFFLHALPCSLEFEATQCPNPFSYVPRPLSSHSDH          MTFLQRVKNMLIAFSQNFCLDVVYSPYATLASEFLQREVTVQDLLSSASV          WLFRSDFVKDYPRPIMPNMVFGGINCLHQNPLSQEFAYINASGEHGIV          VFSLGSMVSEIPEKKAMAIADALGKIPQTVLWRYTGTRPSNLANNILVK          WLPQNDLLGHPMTRAFITHAGSHGVYESICNGVPMVMPLFGDQMDNA          KRMETKGAGVTLNVLEMTSEDLENALKAVINDKSYKENIMRSSLHKDR          PVEPLDLAVFWVEFVMRHKGAPHLRPAAHDLTWYQYHSLDVIGFLLAV          VLTVAFITFKCCAYGYRKCLGKKGRVKKAHKSKTH</p>	UGT1A1
77	<p>MSSKGSVVLA YSGGLDTSCILVWLKEQGYDVIAYLANIGQKEDFEEARK          KALKLGAKKVFIEDVSREFVEEFIWPAIQSSALYEDRYLLGTSLARPCAR          KQVEIAQREGAKYVSHGATGKGNQVRFELSCYSLAPQIKVIAPWRMPEF          YNRFKGRNDLMEYAKQHGIPIVTPKNPWSMDENLMHISYEAGILENPKN          QAPPGLYTKTQDPAKAPNTPDILEIEFKKGVVKTNVKDGTTHTSLELF          MYLNEVAGKHGVRIDIVENRFIGMKSRIYETPAGTILYHAHLDIEAFT          MDREVRKIKQGLGLKFAELVYTGFWHSPECEFVRHCIAKSQERVEGKVQ          VSVLKGQVYILGRESPLSLYNEELVSMNVQGDYEPTDATGFININSLRLKE          YHRLQSKVTAK</p>	ASS1
78	<p>MSTAVLENPGLGRKLSDFGQETS YIEDNCNQNGAISLIFSLKEEVGALAKV          LRLFEENDVNLTHIESRPSRLKKDEYEFFTHLDRSLPALTNIKILRHDIGA          TVHELSDKDDTVPWFPRTIQELDRFANQILSYGAELDADHPGFKDPVY          RARRKQFADIAYNYRHGQPIPRVEYMEEKKTWGTVFKTLKSLYKTHAC          YEYNHIFPLLEKYCGFHEDNIPQLEDVVSQFLQCTGFRLRPVAGLLSRDF          LGGLAFRVFHTQYIRHGSKPMYTPEDICHELLGHVPLFSDRSFAQFSQEI          GLASLGAPDEYIEKLATYWFVTEFGLCKQGDSIKAYGAGLLSSFELQYC          LSEKPKLLPLELEKTAIQNYTVTEFQPLYVAESFNDAKEKVRNFAATIPR          PFSVRYDPYTQRIEVLNNTQQLKILADSINSEIGILCSALQKIK</p>	PAH
79	<p>MAKTLQAQSKTSSQQFSFTGNSSANVIIGNQKLTINDVARVARNGTLVS          LTNNTDILQGIQASCDYINNAVESGEPIYGVTSFGGMANVAISREQASEL          QTNLVWFLKTGAGNKLPLADVRAAMLLRANSHMRGASGIRLELIKRMEL          FLNAGVTPYVYEFSGSIGASGDLVPLSYITGSLIGLDPSFKVDFNGKEMDAP          TALRQLNLSPLTLLPKEGLAMMNGTSMVTGIAANCYDQTILTAIAMGV          HALDIQALNGTNQSFHPFIHNSKPHPGQLWAADQMISLLANS          QLVRELDGKHDIRDHELIQDRYSLRCLPQYLGPIVDGISQIAKQIEIEINS          VTDNPLIDVDNQASYHGGNFLGQYVGMGMMDHLRYIIGLLAKHLDVQIA          LLASPEFSNGLPPSLGNRERKVNMGKGLQICGNSIMPLTFYGNISADR          FPTHAEQFNQINSQGYTSATLARRSVDIFQNYVAIALMFGVQAVDLRTY          KKTGHYDARASLSPATERLYSAVRHVVGQKPTSDRPYIWNNEQGLDEH          IARISADIAAGGVIVQAVQDILPSLH</p>	PAL

10

20

30

40

80	MSTERDSETTFDEDSQPNDDEVVPYSDDETEDELDDQGS AVEPEQNRVNRE AEENREPFRKECTWQVKANDRKYHEQPHFMNTKFLCIKESKYANNAIKT YKYNAFTFIPMNLFEQFKRAANLYFLALLILQAVPQISTLAWYTTLVPLL VLGVT AIKDLVDDVARHKMDKEINNRTCEVIKDGFRFKVAKWKEIQVGD VIRLKKND FVPADILLSSSEPNSLCYVETAELDGETNLKFKMSLEITDQY LQREDTLATFDGFIECEEPPNRLDKFTGTLFWRNTSFPLDADKILLRGCVI RNTDFCHGLVIFAGADTKIMKNSGKTRFKRTKIDYLMNYMVYTIFVVLIL LSAGLAIGHAYWEAQVGNSSWYLYDGEDDTPSYRGFLIFWGYIIVLNTM VPISLYVSVEVIRLQGS HFINWDLQMYA EKDTPAKARTTTLNEQLGQIH YIFSDKTGTLTQNI MTFKKCCINGQIYGDHRDASQHNHNKIEQVDFSWNT YADGKLAFYDHYLIEQIQSGKEPEVRQFFFLAVCHTVMVDRTDGGQ LNY QAASPDEGALVNAARNFGFAFLARTQNTITISELGTERTYNVLA ILDFNSD RKRMSIIVRTPEGNIKLYCKGADTVIYERLHRMNPTKQETQDALDIFANET LRTLCLCYKEIEEKEFTEWNKKFMAASVASTNRDEALDKVYEEIEKDLIL LGATAIEDKLQDGV PETISKLAKADIKIWVLTGDKKETAENIGFACELLTE DTTICYGEDINSL LHARMENQRNRGGVYAKFAPPVQESFFPPGGNRALIT GSWLNEILLEKKT KR NKILKLFPRTEEEERRMRTQSKRRLEAKKEQRQKN FVDLACECSAVICCRVTPKQKAMVVDLVKRYKAITLAIGDGANDVNMI KTAHIGVGISGQEGMQAVMSSDYSFAQFRYLQRLLLVHGRWSYIRMCKF LRYFFYKNFAFTLVHFWYSFFNGYSAQTAYEDWFITLYNVLYTSLPVLLM GLLDQDVSDKLSLRFPGLYIVGQRDLLFNYKRFFV SLLHGVL TSMILFFIPL GAYLQTVGQDGEAPSDYQSF AVTIASALVITVNFQIGLDTSYWTFVNAFSI FGSIALYFGIMFDFHSAGIHLFPSAFQFTGTASNALRQPYIWLTIILAVAV CLLPVVAIRFLSMTIWPSESDKIQKHRKRLKAE EQWQRRQQVFRRGVSTR RSAYAFSHQRGYADLISSGRSIRK KRSPLDAIVADGTA EYRRTGDS	ATP8B1
----	---	--------

10

20

30

40

50



82	MDLEAAKNGTAWRPTSAEGDFELGISSKQKRKKTCTVKMIGVLTFLFRYS DWQDKLFMSLGTIMAIAHGSGLPLMMIVFGEMTDKFDVDTAGNFSFPVNF SLSLLNPGKILEEMTRYAYYYSSGLGAGVLVAAYIQVSFWTLAAGRQIRK IRQKFFHAILRQEIGWFDINDTTELNTRLTDDISKISEGIGDKVGMFFQAVA TFFAGFIVGFIRGWKLTLVIMAISPILGLSAAVWAKILSAFSDKELAAAYAK AGAVAEALGAIRTVIAFGGQNKELERYQKHLENAKEIGIKKAISANISMG IAFLLIYASYALAFWYGSTLVISKEYTIGNAMTVFFSILIGAFSVGQAAPCI DAFANARGAAYVIFDIIDNNPKIDSFSEGRGHKPDSEIKGNLEFNDVHFSYPSR ANVKILKGLNLKVQSGQTVLVGSSGCGKSTTVQLIQRLYDPDEGTINID GQDIRNFNVNYLREIIGVVSQEPVLFSTTIAENICYGRGNVTMDEIKKAVK EANAYEFIMKLPQKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLD EATSALDTESEAEVQAALDKAREGRTTIVIAHRLSTVRNADVIAGFEDGVI VEQGSHELMKKEGVYFKLVNMQTSQSIIQSEEFELNDEKAATRMAPNG WKSRLFRHSTQKNLKSQMCQKSLDVETDGLEANVPPVSFLKVLKLNKT EWPYFVVGTVCAIANGGLQPAFVIFSEIIAIFGPGDDAVKQKCNIFSLIFL FLGIISFFTFQFTFGKAGEILTRRLRSMAFKAMLRQDMSWFDDHKNST GALSTRLATDAAQVQGATGTRLALIAQNIANLGTGIIISFIYGWQLTLLL AVVPIIAVSGIVEMKLLAGNAKRDKKELEAAGKIATEAIENIRTVVSLTQE RKFESMYVEKLYGPNRNSVQKAHIYGITFSISQAFMYFSYAGCFRFGAYLI VNGHMRFRDVILVFSVIFGAVLGHASSFAPDYAKAKLSAAHLFMLFER QPLIDSYSEGLKPKDFEGNITFNEVFNYPTRANVPVLQGLSLEVKKGQT LALVSSGCGKSTVVQLLERFYDPLAGTVFVDFGFQLLDGQEAKKLVQ WLRALGIVSQEPILFDCSIAENIAYGDNRSRVVSQDEIVSAAKAANIHPFIE TLPHKYETRVGDKGTQLSGGQKQRIAIARALIRQPQILLLDEATSALDTE EKVVQEALDKAREGRTCVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQL LAQKGIYFSMVSQAGTQNL	ABCB4
----	---	-------

10

20

30

40

50

83	MPVRGDRGFPPRRELSGWLRA PGMEELIWEQYTVTLQKDSKRGFGIAVS GGRDNPHFENGETSIVISDVLPGGPADGLLQENDRVVMVNGTPMEDVLH SFAVQQLRKSGKVA AIVVKRPRKVQVAALQASPLDQDDRAFEVMDEFD GRSFRSGYSERSRLNSHGGRSRSWEDSPERGRPHERARSRERDLSRDRSR GRSLERGLDQDHARTRDRSRGRSLERGLDHDGFSRDRDRDRSRGRSIDQ DYERAYHRAYPDYERAYSPEYRRGARHDARSRGPRSRSRREHPSRSPSP EPRGRPGPIGVLLMKSRANEEYGLRLGSQIFVKEMTRTGLATKDGNLHEG DIILKINGTVTENMSLTDARKLIEKSRGKLQLVVL RDSQQTLINIPSLNDS SEIEDISEIESNRSFSPEERRHQYSYDYHSSSEKLERPSSREDTPSRLSRM GATPTPFKSTGDIAGTVVPETNKEPRYQEDPPAPQPKAAPRTFLRPSPEDE AIYGPNTKMVRFKKGDSVGLRLAGGNDVGIFVAGIQEGTSAEQEGLQEG DQILKVNTQDFRGLVREDAVLYLLEIPKGEMVTILAQSRADVYRDILACG RGDSFFIRSHFECEKETPQSLAFTRGEVFRVVD TLYDGKLG NWLAVRIGN ELEKGLIPNKSRAEQMASVQNAQRDNAGDRADFWRMRGQRSGVKKNLR KSREDLTAVVSVSTKFPAYERVLLREAGFKRPVVLFGPIADIAMEKLANE LPDWFQTAKTEPKDAGSEKSTGVVRLN TVRQIIEQDKHALLDVT PKAVDL LNYTQWFPIVIFFPDSRQGVKTMRQRLNPTS NKSSRKLFDQANKLKKTC AHLFTATINLNSANDSWFGSLKDTIQHQQGEAVVWSEGKMEGMDDDDPE DRMSYLTAMGADYLSCD SRLISDFEDTDGEGGAYTDNELDEPAEEPLVSS ITRSSEPQHEESIRKPSPEPRAQMRR AASSDQLRDNSPPPAFKPEPPKAKT QNKEESYDFSKSYEYKSNPSAVAGNETPGASTKGYPPPVA AKPTFGRSILK PSTPIPPQEGEEVGESSEEQDNAPKSVLGKVKIFEKMDHKARLQRMQELQ EAQNARIEIAQKHPDIYAVPIKTHKPDGTPQHTSSRPPEPQKAPSRPYQDT RGSYGSDAE EEEYRQQLSEHSKRGYYGQSARYRDEL	TJP2
84	MATATRLLGWRVASWRLRPLAGFVSQRAHSLLPVDDAINGLSEEQRQL RQTMAKFLQEHLAPKAQEIDRSNEFKNLREFWKQLGNLGVLGITAPVQY GSGGLGYLEHVLVMEEISRASGAVGLSYGAHNL CINQLVRNGNEAQKE KYLPLKISGEYIGALAMSEPNAGSDV VSMKLKAEKKNHYILNGNKFWIT NGPDADVLIVYAKTDLAAVPASRGITAFIVEKGM PGFSTSKKLDKLGMRG SNTCELIFEDCKIPAANILGHENKGVYVLM SGLDLERLVLAGGPLGLMQA VLDHTIPYLHVREAFGQKIGHFQLMQGKMADMYTRLMACRQYVYNVA KACDEGHCTAKDCAGVILYSAECATQVALDGIQCFGGNGYINDFPMGRF LRDAKLYEIGAGTSEVRRLVIGRAFNA DFH	IVD
85	MALRGVSVRLLSRGPGLHVLRTWVSSAAQTEKGGRTQSQLAKSSRPEFD WQDPLVLEEQLTTDEILIRD TFRTYCQERLMPRILLANRNEVFHREIISEMG ELGVLGPTIKGYGCAGVSSVAYGLLARELERVDSGYRSAMSVQSSLMH PIYAYGSEEQRQKYLPLQAKGELLGCFGLTEPN SGPSSMETRAHYNSS NKS YTLNGTKTWITNSPMADLFV VWARCEDGCIRGFLLEKGMRGLSAPRI QGKFSLRASATGMIIMDGVEVPEENVLP GASSLGGPFGCLNNARYGIAWG VLGASEFCLHTARQYALDRMQFGVPLARNQLIQK KLADMLTEITLGLHA CLQLGRLKDQDKAAPMVSLKRN NCGKALDIARQARDMLGGNGISDE YHVIRHAMNLEAVNTYEGTHDIHALILGRAITGIQAF TASK	GCDH

10

20

30

40

50



86	MFRAAAPGQLRRAASLLRFQSTLVIAEHANDSLAPITLNTITAAATRLGGEV SCLVAGTKCDKVAQDLCKVAGIAKVLVAQHVDYKGLLPEELTPLILATQ KQFNETHICAGASAFGKNLLPRVAAKLEVAPISDIIAIKSPDTFVRTIYAGN ALCTVKCDEKVKVFSVRGTSFDAAATSGGSASSEKASSTSPVEISEWLDQ KLTKSDRPELTGAKVVVSGGRGLKSGENFKLLYDLADQLHAAVGASRAA VDAGFVPNDMQVGQGTGKIVAPELYIAVGISGAIQHLAGMKDSKTIVAINK DPEAPIFQVADYGIVADLFKVVPEMTEILKKK	ETFA
87	MAELRVLVAVKRVIDYAVKIRVKPDRGTGVVTDGVKHSMPFCEIAVEEA VRLKEKLVKEVIAVSCGPAQCQETIRTALAMGADRGIHVEVPPAEAERL GPLQVARVLAKLAEKEKVDLVLGKQAIDDDCNQTGQMTAGFLDWPQG TFASQVTLEGDKLKVEREIDGGLETLRRLKLPVVTADLRLNEPRYATLPNI MKAKKKKIEVIKPGDLGVDLTSKLSVISVEDPPQRTAGVKVETTEDLVAK LKEIGRI	ETFB
88	MLVPLAKLSCLAYQCFHALKIKKNYLPLCATRWSSTSTVPRITHTYIYPR DKDKRWEGVNMERFAEEADVIVGAGPAGLSAAVRLKQLAVAHEKDIR VCLVEKAAQIGAHTLSGACLDPGAFKELFPDWKEKGAPLNTPVTEDFGI LTEKYRIPVPILPGLPMNNHGNIVRLGHLVSWMGEQAEALGVEVYPGY AAAEVLFHDDGSVKGIATNDVGIQKDGAPKATFERGLELHAKVTIFAEGC HGHLAKQLYKFFDLRANCEPQTYGIGLWVWIDEKNWKPGRVDHTVG WPLDRHTYGGFLYHLNEGEPLVALGLVVGLDYQNPYLSPFREFQRWKH HPSIRPTLEGGKRIAYGARALNEGGFQSIPKLTFFGGLLIGCSPGFMNVPKI KGTHTAMKSGILAAESIFNQLTSENLSKQSKTIGLHVTEYEDNLKNSWVWKE LYSVRNIRPSCHGVLGVYGGMIYTGIFYWILRGMEPWTLKHKGSDFERLK PAKDCTPIEYKPDGQISFDLLSSVALSGTNHEHDQPAHLTLRDDSIPVNR NLSIYDGPQRFPCPAGVYEFVPEVQGDGFRQLQINAQNCVHCKTCDIKDP QNINWVPEGGGGPAYNGM	ETFDH
89	MASEGKLWGGRFVGAVDPIMEKFNASIAADRHLWEVDVQGSKAYSRG LEKAGLLTKAEMDQILHGLDKVAEEWAQGTFLKNSNDEDIHTANERRLK ELIGATAGKLHTGRSRNDQVVDLRLWMRQTCSTLSGLLWELIRTMVDR AEAERDVLFPGYTHLQRAQPIRWSHWILSHAVALTRDSERLLEVRKRINV LPLGSGAIAGNPLGVDRELLRAELNFGAITLNSMDATSERDFVAEFLFWA SLCMTHLSRMAEDLILYCTKEFSFVQLSDAYSTGSSLMPQKKNPDSLELIR SKAGR VFGR CAGLLMTL KGLPSTYNKDLQEDKEAVFEVSDTMSAVLQV ATGVISTLQIHQENMGQALSPDMLATDLAYYLVRKGM PFRQAHEASGKA VFMAETKGV ALNQLSLQELQTISPLFSGDVICVWDYGH SVEQY GALGGT ARSSVDWQIRQVRALLQAQQA	ASL

10

20

30

40

50

90	<p>MVGGSPVPFDEILSTARARMNRVLSFHVS SGILVCQAGCVLEELSRYVEERD  FIMPLDLGAKGSGCHIGGNVATNAGGLRFLRYGSLHGTVLGLEVVADGT  VLDCLTSLRKDNTGYDLKQLFIGSEGTLGIITTVSILCPPKPRAVNVAFLGC  PGFAEVLQTFSTCKGMLGEILSAFEFMDAVCMQLVGRHLHLASPVQESPF  YVLIETSGSNAGHDAEKLGHFLEHALGSGLVTDGTMATDQRKVKMLWA  LRERITEALS RDGYVYKYDLSLPVERLYDIVTDLRRLGPHAKHVVG YGH  LGDGNLHLNVTAEAFSPSLLAALEPHVYEW TAGQQGSVSAEHGVGFRKR  DVLGYSKPPGALQLMQQLKALLDPKGILNPYKTLPSQA</p>	D2HGDH
91	<p>MAAMRKALPRRLVGLASLRAVSTSSMGTLPKRVKIVEVGPRDGLQNEKN  IVSTPVKIKLIDMLSEAGLSVIETTSFVSPKWVPQMGDHTV LKGIQKFPPI  NYPVLT PNLKGFEAAVAAGAKEVVIFGAASELFTKKNINCSIEESFQRFDA  ILKAAQSANISVRGYVSCALGCPYEGKISPAKVAEVTKKFYSMGCYEISLG  DTIGVGTGPGIMKMLSAVMQEVPLAALAVHCHD TYGQALANTLMALQM  GVSVDSSVAGLGGCPYAQGASGNLATEDLVYMLEGLGIHTGVNLQKLL  EAGNFICQALNRKTSSKVAQATCKL</p>	HMGCL
92	<p>MAAASAVSVLLVAAERNRWHRLPSLLLPPRTWVWRQRTMKYTTATGRN  ITKVLIANRGEIACRVMRTAKKLG VQTVAVYSEADRNSMHVDMADEAY  SIGPAPSQQSYLSMEKIIQVAKTSA AQAIHPGCGFLSENMEFAELCKQEGII  FIGPPPSAIRDMGIKSTSKSIMAAAGVPVVEGYHGEDQSDQCLKEHARRIG  YPMVIKAVRGGGKGMRIVRSEQEFQEQLESARREAKKSFNDDAMLIEK  FVDTPRHVEVQVFGDHHGNAVYLFERDCSVQRRHQKIIEEAPAPGIKSEV  RKKLGEAAVRAAKAVNYVGAGTVEFIMDSKHNFCMEMNTRLQVEHPV  TEMITGTDLVEWQLRIAAGEKIPLSQEEITLQGHAFEARIYAEDPSNNFMP  VAGPLVHLSTPRADPSTRIETGVRQGDDEVSVHYDPMIAKL VVWAADRQA  ALTKLRYSLRQYNIVGLHTNIDFLLNLSGHPEFEAGNVHTDFIPQHHKQLL  LSRKA AAKE SL CQAALGLILKEKAMTDFTLQAHDQFSPFSSSSGRRLNIS  YTRNMTLKDGNNAIAVVTYNHDGYSM QIEDKTFQVLGNLYSEGDC T  YLKCSVNGVASKAKLIILENTIYLF SKEGSIEIDIPVPKYLSSVSSQETQGGP  LAPMTGTIEKV FVKAGDKVKAGDSL MVMIAMKMEHTIKSPKDGTVKKV  FYREGAQANRHTPLVEFE EEEESDKRESE</p>	MCCC1
93	<p>MWAVLRLALRPCARASPAGPRAYHGDSVASLGTQPD LGSALYQENYKQ  MKALVNQLHERVEHIKLGGEKARALHISRGKLLPRERIDNLIDPGSPFLE  LSQFAGYQLYDNEEVPGGGIITGIGRVSGVECMIIANDATVKGGAYYPVT  VKKQLRAQEIAMQNRLPCIYLVDSGGAYLPRQADVFPDRDHFGRTFYNQ  AIMSSKNIAQIAVVMG SCTAGGAYVPAMADENIIVRKQGTIFLAGPPLVK  AATGEEVSAEDLGGADLHCRKSGVSDHWALDDHHALHLTRKVVRNLNY  QKKLDVTIEPSEEP LPADELYGIVGANLKR SFDVREVIARIVDGSRFTEFK  AFYGD TLVTGFARIFGYPVGIVGNNGVLFSES AKKGTHFVQLCCQRNIPLL  FLQNITGFMVGREYEAEGIAKDGAKMVA AVACAQVPKITLIIGGSY GAGN  YGMCGRAYSPRFLYIWP NARISVMGGEQAANVLATITKDQRAREGKQFS  SADEAALKEPIIKKFEEEGNPYYSSARVWDDGIIDPADTRLVLGLSFS AAL  NAPIEK TDFGIFRM</p>	MCCC2

10

20

30

40

50

94	MAVAGPAPGAGARPRDLQFLQRFLQILKVLFPSSQNALMFLTLLCLT LLEQFVIYQVGLIPSQYYGVLGNKDLEGFKTLTFLAVMLIVLNSTLKSFDQ FTCNLLYVSWRKDLTEHLHRLYFRGRAYYTLNVLRDDIDNPDQRISQDVE RFCRQLSSMASKLIISPFTLVYYTYQCFQSTGWLGPVSIFGYFILGTVVNKT LMGPVIMKLVHQEKLEGDFRFKHMQIRVNAEPAAFYRAGHVEHMRTDR RLQRLQQTQRELMSKELWLYIGINTFDYLGSIYSYVVIPIFSGVYGDLS AELSTLVSKNAFVCIYLISCFTQLIDLSTLSDVAGYTHRIGQLRETLDDMS LKSQDCEILGESEWGLDTPPGWPAEPAADTAFLLERVSISAPSSDKPLIKDL SLKISEGQSLITGNTGTGKTSLLRVLGGLWTSTRGSVQMLTDFGPHGVLF LPQKPFFTDGTRELVYIPLKEVYPDSGSADDERILRFLELAGLSNLVART EGLDQQVDWNWYDVLSPGEMQRLSFARLFYLQPKYAVLDEATSALTEE VESELYRIGQQLGMTFISVGHRSLEKFSLSVLKLCGGGRWELMRIKVE	ABCD4
----	--	-------

10

20

30

40

50

95	<p>MASAVSPANLPAVLLQPRWKRVVGWSGPVPRPRHGHRAVAIKELIVVFG  GGNEGIVDELHVYNTATNQWFIPAVRGDIPPGCAAYGFVCDGTRLLVFG  GMVEYGGKYSNDLYELQASRWEWKRLKAKTPKNGPPPCPRLGHSFSLVG  NKCYLEFGLANDSEDPKNNIPRYLNDLYILELRPGSGVVAWDIPITYGVLP  PPRESHTAVVYTEKDNKSKLVYGGMSGCRLGDLWTLDDIDTLTWNKPS  LSGVAPLPRSLHSATTIGNKMYVFGGWVPLVMDDVKVATHEKEWKCTN  TLACLNLDTMAWETILMDTLEDNIPRARAGHCAVAINTRLYIWSGRDGY  RKAANNQVCKDLWYLETEKPPPPARVQLVRANTNSLEVSAGAVATAD  SYLLQLQKYDIPATAATATSPTPNPVSVPANPPKSPAPAAAAPAVQPLTQ  VGITLLPQAAPAPPTTTTIQVLPVPGSSISVPTAARTQGVPAVLKVTGPQA  TTGTPLVTMRPASQAGKAPVTVTSPLAGVRMVVPTQSAQGTVIGSSPQMS  GMAALAAAAAATQKIPPSSAPTIVLSPAGTTIVKTMAVTPGTTTLPATVK  VASSPVMVSNPATRMLKTAQAQVGTSSVSSATNTSTRPIITVHKSQVTVA  QQAQVVTTVVGGVTKTITLVKSPISVPGGSALISNLGKVMSSVQTKPVQT  SAVTGQASTGVPVQIIQTKGPLPAGTILKLVTSADGKPTTIITTTQASGAGT  KPTILGISSVSPSTTKPGTTTTIHKTIPTMSAHTQAGATGVTSSPGIKSPITITTK  VMTSGTGAPAKIITAVPKIATGHGQOGVTQVVLKQAGQPGTILRTVPMG  GVRLVTPVTVSAVKPAVTTLVVKGTTGVTTLGTVTGTVSTSLAGAGGHS  TSASLATPITTLGTIATLSSQVINPTAITVSAAQTTLTAAGGLTPTITMQPV  SQPTQVTLITAPSGVEAQPVHDLVPSILASPTTEQPTATVTIADSGQGDVQP  GTVTLVCSNPPCETHETGTTNTATTTVVANLGGHPQPTQVQFVCDRQEA  AASLVSTVGGQNGSVVRVCSNPPCETHETGTTNTATTATSNMAGQHGC  SNPPCETHETGTTNTATTAMSSVGANHQDARRACAAGTPAVIRISVATG  ALEAAQGSKSQCQTRQTSATSTMTVMATGAPCSAGPLLGPSMAREPGG  RSPAFVQLAPLSSKVRLLSSPSIKDLPAGRHSHAVSTAAMTRSSVGAGEPR  MAPVCESLQGGSPSTTVTVTALEALLCPSATVTQVCSNPPCETHETGTTNT  ATTSNAGSAQRVCSNPPCETHETGTTHTATTATSNNGGTGQPEGGQQPPAG  RPCETHQTTSTGTTMSVSVGALLPDATSSHRTVESGLEVAAPSVPQAG  TALLAPFPTQRVCSNPPCETHETGTTHTATTVTSNMSSNQDPPPAASDQGE  VESTQGDVSNITSSAITTTVSSTLTRAVTTVTQSTPVPGPSVPPPEELQVSP  GPRQQLPPRQLLQSASTALMGESAEVLSASQTPELPAAVDLSSTGEPSSGQ  ESAGSAAVAVVVQPPPTQSEVDQLSLPQELMAEAQAGTTTLMVTGLTP  EELAVTAAAEAAAQAAATEEAQALAIQAVLQAAQQAVMGTGEPMDTSE  AAATVTQAEGLHLSAEGQEGQATTIPIVLTQQELAALVQQQLQEAQAQ  QQHHHLPTALAPADSLNDPAIESNCLNELAGTVPSTVALLPSTATESLAP  SNTFVAPQPVVVASPAKLQAAATLTVANGIESLGVKPDLPKPPSKAPMK  KENQWFDVGVKGTNVMTVHYFLPPDDAVPSDDDLGTVPDYNQLKKQE  LQPGTAYKFRVAGINACGRGPFSEISAFKTCLPGFPGAPCAIKISKSPDGAH  LTWEPPSVTSKHEYSVYLAIQSSQAGGELKSSTPAQLAFMRVYCGPSPSC  LVQSSSLNAHIDYTTKPAIIFRIARNEKGYGPATQVRWLQETSKDSSGT  KPANKRPMSSPEMKSAPKKSADGQ</p>	HCFC1
----	--	-------

10

20

30

40

50

96	<p>MATSGAASAELVIGWCIFGLLLLAILAFCWIYVRKYQSRRESEVVSTITAIF  SLAIALITSALLPVDIFLVSYMKNQNGTFKDWANANVSRQIEDTVLYGYY  TLYSVILFCVFFWIPFVYFYEEKDDDDTSKCTQIKTALKYTLGFVVICAL  LLLVGAFVPLNVPNNKNSTEWKVKSLFEELGSSHGLAALSFSISLTLIG  MLAAITYTAYGMSALPLNLIKGTRSAAYERLENTEDIIEVEQHIQTIKSKS  KDGRPLPARDKRALKQFEERLRTLKKRERHLEFIENSWWTKFCGALRPLK  IVWGIFFILVALLFVISLFLSNLDKALHSAGIDSGFIIFGANLSNPLNMLLPL  LQTVFPLDYILITHIIMYFIFTSMAGIRNIGIWWFFWIRLYKIRRGTRTPQALLF  LCMILLLIVLHTSYMIYSLAPQYVMYGSQNYLIETNITSDNHKGNSTLSVP  KRCADAPEDQCTVTRTYLFLHKFWFFSAAYYFGNWAFLGVFLIGLIVSC  CKGKKSVEGVEDSDISDDEPSVYSA</p>	LMBRD1
97	<p>MSAKSRTIGIIGAPFSKGQPRGGVEEGPTVLRKAGLLEKLKEQECDVKDY  GDLPFADIPNDSPFQIVKNPRSVGKASEQLAGKVAEVKKNGRISLVLGGD  HSLAIGSISGHARVHPDLGVIWVDAHTDINTPLTTTTSGNLHGQPVSFLLKE  LKGKIPDVPGFSWVTPCISAKDIVYIGLRDVPGEHYILKTLGIKYFSMTEV  DRLGIGKVMEEETLSYLLGRKKRPIHLSFDVDGLDPSFTPATGTPVVGGLTY  REGLYTEEYKTGLLSGLDIMEVNPSSLGKTPEEVTRTVNTAVAITLACFGL  AREGNHKPIDYLNPPK</p>	ARG1
98	<p>MKSNPAIQAAIDLTAGAAGGTACVLTGQPFDTMKVKMQTFPDLYRGLTD  CCLKTYSQVGRGFYKGTSPALIANIAENSVLFMCYGFCCQVVRKVAGL  DKQAKLSDLQNAAGSFASAFALVLCPTLVKCRQLQTMEMETSGKIA  KSQNTVWSVIKSILRKDGPLGFYHGLSSTLLREVPGYFFFFGGYELSRFF  ASGRSKDELGPVPLMLSGGVGGICLWLAVYPVDCIKSRIQVLSMSGKQAG  FIRTFINNVKNEGITALYSGLKPTMIRAFPANGALFLAYEYSRKLMMNQLE  AY</p>	SLC25A1 5
99	<p>MAAAKVALTKRADPAELRTIFLKYASIEKNGEFFMSPNDFVTRYLNIFGES  QPNPKTVELLSGVVDQTKDGLISFQEFVAFESVLCAPDALFMVAFQLFDK  AGKGEVTFEDVKQVFGQTTIHQHIFPNWDSEFVQLHFGKERKRHLTYAEF  TQFLLEIQLEHAKQAFVQRDNARTGRVTAIDFRDIMVTIRPHVLTVPFVEEC  LVAAAGGTTSHQVSFSYFNGFNLLNNMELIRKIYSTLAGTRKDVEVTKE  EFVLAQKFGQVTPMEVDILFQLADLYEPRGRMTLADIERIAPLEEGTLPF  NLAEAQRQKASGDSARPVLLQVAESAYRFGLGSVAGAVGATAVYPIDL  KTRMQNQRSTGSFVGLMYKNSFDCFKKVLRYEGFFGLYRGLLPQLLGV  APEKAIKLTVNDVFRDKFMHKDGSVPLAAEILAGGCAGGSQVIFTNPLEIV  KIRLQVAGEITTGPRVSALSVRDLGFFGIYKGAACFLRDIPFSAIYFPCY  AHVKASFANEDGQVSPGSLLAGAIAAGMPAASLVTPADVIKTRLQVAAR  AGQTTYSGVIDCFRKILREEGPKALWKGAGARVFRSSPQFGVTLLTYELL  QRWFYIDFGGVKPMGSEVPKSRINLPAPNPDHVGGYKLAVATFAGIENK  FGLYLPLFKPSVSTSKAIGGGP</p>	SLC25A1 3

10

20

30

40

50

100	MQPQSVLHSGYFHPLLRWQTATTTLNASNLIYPIFVTDVDPDDIQPITSLPG VARYGVKRLEEMLRPLVEEGLRCVLIFGVPSRVPKDERGSAADSEESPAIE AIHLLRKTFFPNLLVACDVCLCPYTSHGHCGLLSENGAFRAEESRQRLAEV ALAYAKAGCQVVAPSDMMDGRVEAIKEALMAHGLGNRVSVMSYSAKF ASCFYGPFRDAAKSSPAFGDRRCYQLPPGARGLALRAVDRDVREGADML MVKPGMPYLDIVREVKDKHPDLPLAVYHVSGEFAMLWHGAQAGAFDLK AAVLEAMTAFRRAGADIIITYYTPQLLQWLKEE	ALAD
101	MALQLGRLSSGPCWLVARGGCGGPRAWWSQCGGGGLRAWWSQSAAGR CRPPGPAGTEQSRGLGHGSTRGGPWVGTGLAAALAGLVGLATAAFGHV QRAEMLPKTSGTRATSLGRPEEEDELAHRCSSFMAPPVTDLGELRRRPG DMKTKMELLILETQAQVCQALAQVDGGANFSVDRWERKEGGGGISCVL QDGCVFEEKAGVSISVVHGNLSEEAAKQMRSGKVLKTKDGKLPFCAMG VSSVIHPKNPHAPTIHFNYRYFEVEEADGNKQWWFGGGCDLTPTYLNQE DAVHFHRTLKEACDQHGPDLYPKFKKWCDYFFIAHRGERRGIGGIFDD LDSPSKEEVFRVQSCARAVVPSYIPLVKKHCDDSFPTQEKLWQLRRGR YVEFNLLYDRGTFGLFTPGSRIESILMSLPLTARWEYMHSPSENSKEAAIL EVLRHPRDWVR	CPOX
102	MSGNGNAAATAEENSPKMRVIRVGTTRKSQLARIQTDSVVATLKASYPGL QFEIIMSTTGDKILDALSKIGESLFTKELEHALEKNEVDLVVHSLKDLPL TVLPPGFTIGAICKRENPHDAVVFHFKFVGKTLETLPKSVVGTSSLRAA QLQRKFPHLEFRSIRGNLNTRLRKLDEQQEFSAILATAGLQRMGWHRV GQILHPEECMYAVGQGALGVEVRAKDQDILDVLGVLDHPETLLRCIAER AFLRHLEGGCSVPVAVHTAMKDGQLYLTGGVWSLDGSDSIQETMQATIH VPAQHEDGPEDDPQLVGITARNIPRGPQLAAQNLGISLANLLLSKGAKNIL DVARQLNDAH	HMBS
103	MGRTVVVLGGGISGLAASYHLSRAPCPPKVVLVESSERLGGWIRSVRGP GAIFELGPRGIRPAGALGARTLLLVSELGLDSEVLPVRGDHPAAQNRFLYV GGALHALPTGLRGLLRPSPFFSKPLFWAGLRELTKPRGKEPDETIVHSFAQ RRLGPEVASLAMDSLRCGVFAGNSRELSIRSCFPSLFQAEQTHRSILLGLL GAGRTPQPSALIRQALAERWSQWSLRGGLEMLPQALETHLTSRGVSVL RGQPVCGLSLQAEGRWKVSLRDSLEADHVISAI PASVLSSELLPAEAAPLA RALSAITAVSVAVVNLQYQGAHLPVQGFHLPVPSSEDPGVLGIVYDSVAF PEQDGSPPGLRVTVMLGGSWLQTLEASGCVLSQELFQQRQAQEAATQLG LKEMPSHCLVHLHKNCIPQYTLGHWQKLESARQFLTAHRLPLTLGASY EGVAVNDCIESGRQAAVSVLGTEPNS	PPOX

10

20

30

40

50

104	MAHAHIQGGRRAKSRFVVCIMSGARSKLALFLCGCYVVALGAHTGEESV ADHHEAEYYVAAVYEHPSILSLNPLALISRQEALELMNQNLDIYEQQVMT AAQKDVQIIVFPEDGIHGFNFTRTSIYPFLDFMPSPQVVRWNPCLPHRFN DTEVLQRLSCMAIRGDMFLVANLGTKEPCHSSDPRCPKDGRYQFNTNVV FSNNGTLVDRYRKHNLYFEAAFDVPLKVDLITFDTPFAGRFGIFTCFDILFF DPAIRVLRDYKVKHVYPTAWMNQLPLLAIEIQKAFVAFAFINVLAAN VHHPVLGMTGSGIHTPLESFWYHDMENPKSHLIIAQVAKNPVGLIGAENA TGETDPSHSKFLKILSGDPYCEKDAQEVHCDEATKWNVNAPPTFHSEMM YDNFTLVVWVGKEGYLHVCSNGLCCYLLYERPTLSKELYALGVFDGLHT VHGTYIQCVCALVRCGGLGFDTCGQEITEATGIFEFHLWGNFSTSYIFPLF LTSGMTLEVDPDQLGWENDHYFLRKSRLSSGLVTAALYGRLYERD	BTD
105	MEDRLHMDNGLVPQKIVSVHLQDSTLKEVKDQVSNKQAQILEPKPEPSLE IKPEQDGMHEVGRDDPKALGEEPQRRGSASGSEPAGDSDRGGGPVEHY HLHLSSCHECLELENSTIESVKFASAENIPDLPYDYSSSLESVADETSPERE GRRVNLTKAPNILLYVGSQSQUALGRFHEVRSVLADCVDIDSYILYHLL EDSALRDPWTDNCLLLVIATRESIPEDLYQKF MAYLSQGGKVLGLSSSFTF GGFQVTSK GALHKT VQNLVFSKADQSEVKLSVLSSGCRYQEGPVRLSPG RLQGHLENEDKDRMIVHVPFGTRGGEAVLCQVHLELPPSSNIVQTPEDFN LLKSSNFRRYEV LREIL TTLGLSCDMKQVPAL TPLYLLSAAEEIRDPLMQW LGKHVDSEGEIKSGQLSLRFVSSYVSEVEITPSCIPVVTNMEAFSSEHFNLEI YRQNLQTKQLGK VILFAEVTPTTMRLLDGLMFQTPQEMGLIVIAARQTEG KGRGGNVWLSPVGCALSTLLISIPLRSQLGQRIPFVQHLM SVAVVEAVRSI PEYQDINLRVKWPNDIYYSDLMKIGGVLVNSTLMGETFYILIGCGFNVTN SNPTICINDLITEYNKQHK AELKPLRADYLIARVVTVLEKLIKEFQDKGPNS VLPLYRYRWVHSGQQVHLGSAEGPKVSIVGLDDSGFLQVHQEGGEVTV HPDGNSFDMLRN LILPKRR	HLCS

10

20

30

40

50

106	MLKFRTVHGGRLLLGIRRTSTAPAASPNVRRLEYKPIKKVMVANRGEIAI RVFRACTELGIRTVAIYSEQDTGQMHRQKADEAYLIGRGLAPVQAYLHIP DIIKVAKENNVDAVHPGYGFLSERADFAQACQDAGVRFIGPSPEVVRKM GDKVEARAI AIAAGVPVPGTDAPITSLHEAHEFSNTYGFPIIFKAAAYGGG GRGMRVVHSHYEELEENYTRAYSEALAAFNGALFVEKFIEKPRHIEVQIL GDQYGNILHLYERDCSIQRRHQKVVEIAPAAHLDPQLRTRLTSDSVKLAK QVGYENAGTVEFLVDRHGKHYFIEVNSRLQVEHTVTEEITDVDLVHAQIH VAEGRSLPDLGLRQENIRINGCAIQCRVTTEDPARSFQPD TGRIEVFRSGEG MGIRLDNASAFQGA VISPHYDSSLVKVIAHGKDHPTAATKMSRALAEFRV RGVKTNIAFLQNVLNNQQFLAGTVDTQFIDENPELFLRPAQNRAQKLLH YLGHV MVNGPTTPIPVKASPSPTDPVPAVPIGPPPAGFRDILLREGPEGFA RAVRNHPGLLLMDTTFRDAHQSLLATRVRTHDLKKIAPYVAHNFSKLF S MENWGGATFDVAMRFLYECPWRR LQELRELIPNIPFQMLLRGANAVGYT NYPDNVVFKEVAKENGMDVFRVFD SLNYLPNMLLGMEAAGSAGGVV EAAISYTG DVADPSRTKYS LQYYMGLAEELVRAGTHILCIKDMAGLLKPT ACTMLVSSLRDRFPDLPLHIHTHDTSGAGVAAMLACAQAGADVVDVAA DSMSGMTSQPSMGALVACTRGTPLDTEVPMERVFDYSEY WEGARGLYA AFDCTATMKSGNSDVYENEIPGGQYTNLHFQAHS MGLGSKFKEVKKAY VEANQMLGDLIKVTPSSKIVGDLAQFMVQNGLSRAEAEAQA EELSFPRSV VEFLQGYIGVPHGGFPEPFRSKVLKDLPRVEGRPGASLPPLDLQALEKELV DRHGEEVTPEDVLSAAMY PDVFAHFKDFTATFGPLDSL NTRLFLQGPKIA EEFEVELERGKTLHIKALAVSDLNRAGQRQVFFELNGQLRSILVKDTQAM KEMHFHPKALKDVKGQIGAPMPGKVIDIKVVAGAKVAKGQPLCVLSAM KMETVVTSPMEGTVRKVHVTKDMTLEGDDLILEIE	PC
107	MVDSTEYEVASQPEVETSPLGDGASPGPEQVKL KKEISLLNGVCLIVGNMI GSGIFVSPKGVLIYSASFGLSLVIWAVGGLFSVFGALCYAELGTTIKKSGAS YAYILEAFGGFLAFIRLWTSLLIIEPTSQAIIAITFANYMVQPLFPSCFAPYA ASRLAAACICLLTFINCA YVKWGTLVQDIFTYAKVLALIAVIVAGIVRLG QGASTHFENSFEGSSFAVGDI ALALYSALFSYSGWDTLNYVTEEIKNPERN LPLSIGISMPIVTHIYILTNAVAYTVLDMRDILASDAVA VTFADQIFGIFNWII PLSVALSFCFGLNASIVAASRLFFVGSREGHLPDAICMIHVERFTPVP SLLF NGIMALIYLCVEDIFQLINYY SFSYWFFVGLSIVGQLYLRWKEPDRPRPLK LSVFFPIVFC LCTIFLVAVPLYSDTINSLIGIAIALSGLPFYFLIIRVPEHKRPL YLRRIVGSATRYLQVLCMSVAAEMDLEDGGEMPKQRDPKSN	SLC7A7

10

20

30

40

50



108	MVPRLLLRAWPRGPAVGPAPSRLSAGSGPGQYLQRSIVPTMHYQDSLPLRPIPKLEDTIRRYLSAQKPLLNDGQFRKTEQFCKSFENGIGKELHEQLVALDKQNKHTSYISGPWFDMYLSARDSVVLNFPFMAFNPDPKSEYNDQLTRATNMTVSAIRFLKTLRAGLLEPEVFHLNPAKSDTITFKRLIRFVPSLSWYGAYLVNAYPLDMSQYFRLFNSTRLPKPSRDELFTDDKARHLLVLRKGNFYIFDVLDQDGNIVSPSEIQAHLKYILSDSSPAPEFPLAYLTSENARDIWAELRQKLMSSGNEESLRKVDSAVFCLCLDDFPIKDLVHLSHNMLHGDGTNRWFDKSFNLIKADGSTAVHFEHSWGDGVAVLRFFNEVFKDSTQTPAVTPQSQPATTDSTVTVQKLN FELTDALKTGITAAKEKFDATMKTLTIDCVQFQRGGKEFLKKQKLSPDVAQALAFQMAFLRQYGGTVATYESCSTA AFKHGR TETIRPASVYTKRCSEAFVREPSRHSAGELQMMVECSKYHGQLTKEAAMGQGFDRHLFALRH LAAAKG IILPELYLDPAYGQINHNVLSTSTLSSPAVNLGGFAPVVS DGFVGYAVHDNWIGCNVSSYPGRNAREFLQCVEKALEDMFD ALEGKSIKS	CPT2
109	MAAGFGRCCRVLRISRFHWRSQHTKANRQREPGLGFSFEFTEQQKEFQATARKFAREEIIPVAAEYDKTGEYVPLIRRAWELGLMNTHIPENCGGLGLGTFDA CLISEELAYGCTGVQTAIEGNSLGQMPHIIAGNDQQKKKYLGRMTEEPLMCAYCVTEPGAGSDVAGIKTKAEKKGDEYIINGQKMWITNGGKANWYFLLARSDPDPKAPANKAFTGFIVEADTPGIQIGRKELNMGQRCS DTRGIVFEDVKVPKENVLIGDGAGFKVAMGAFDKTRPVVAAGAVGLAQRALDEATKYALERKTFGKLLVEHQ AISFMLAEMAMKVELARMSYQRAAWEVDSGRRNTYYASIAKAFAGDIANQLATDAVQILGGNGFNTEYPVEKLMRDAKIYQIYEGTSQIQRLIVAREHIDKYKN	ACADM
110	MAAALLARASGPARRALCPRAWRQLHTIYQSVELPETHQMLLQTCRDFAEKELFP IAAQVDKEHLFPAAQVKKMGGLG LLAMDVPEELGGAGLDYLA YAIAMEEISRGCASTGVIMSVNNSLYLGPILKFGSKEQKQAWVTPFTSGDKIGCFALSEPGNGSDAGAASTTARAEGDSWVLNGTKAWITNAWEASAAVVFAS TDRALQNKGISAF LVPMPPTGLTLGKKEDKLGIRGSSTANLIFEDCRIPKDSILGEPGMGFKIAMQTLDMGRIGIASQALGIAQTALDCAVNYAENRM AFGAPLTKLQVIQFKLADMALALESARLLTWRAAMLKDNKKPFIKEAAMAKLAASEAATAISHQAIQILGGMGYVTEMPAERHYRDARITEIYEGTSEIQRLVIAGHLLRSYRS	ACADS

10

20

30

40

50

111	MQAARMAASLGRQLRLRGGSSRLTALLGQPRPGPARRPYAGGAAQLA LDKSDSHPSDALTRKKPAKAESKSFAVGMFKGQLTTDQVFPYPSVLNEEQ TQFLKELVEPVSRRFEEVNDPAKNDALEMVEETTQQGLKELGAFGLQVP SELGGVGLCNTQYARLVEIVGMHDLGVGITLGAHQSIGFKGILLFGTKAQ KEKYLPKLASGETVAAFCLTEPSSGSDAASIRTSAVPSPCGKYYTLNGSKL WISNGGLADIFTVFAKTPVTDPATGAVKEKITAFVVERGFGGITHGPPEKK MGIKASNTAEVFFDGVRVPSENVLGEVGSFKVAMHILNNGRFGMAAAL AGTMRGIIAKAVDHATNRTQFGEKIHNFGLIQEKLARMVMLQYVTESMA YMVSANMDQGATDFQIEAAISKIFGSEAAWKVTDECIQIMGGMGFMKEP GVERVLRDLRIFRIFEGTNDILRLFVALQGCMDKGKELSGLSALKNPFGN AGLLLGEAGKQLRRRAGLGSGLSLSGLVHPELSRSGELAVRALEQFATVV EAKLIKHKKGIVNEQFLLQRLADGAIDLYAMVVVLSRASRSLSEGHPTAQ HEKMLCDTWCIEAAARIREGMAALQSDPWQQELYRNFKSISKALVERGG VVTSNPLGF	ACADVL
-----	--	--------

10

20

30

40

50

112	<p>MGHSKQIRILLNEMEKLEKTLFRLEQGYELQFRLGPTLQGKAVTVYTN  PFPGETFNREKFRSLDWENPTEREDDSKYCKLNLQQSGSFQYYFLQNE  KSGGGYIVVDPILRVGADNHVLPDCVTLQTFLAKCLGPFDEWESRLRVA  KESGYNMIHFTPLQTLGLSRSCYSLANQLELNPDFSRLPNRKYTWNDVGQL  VEKLLKEWNVICITDVVYNHTAANSKWIQEHPECAYNLVNSPHLKPAWV  LDRALWRFSCDVAEGKYKEKGIPALIENDHHMNSIRKIIWEDIFPKLKLWE  FFQVDVNKAVEQFRRLLTQENRRVTKSDPNQHLTHIQDPEYRRFGCTVDM  NIALTTFIPHDKGPAIEECCNWFHKRMEELNSEKHRLINYHQEQAVNCL  LGNVFIYERLAGHGPKLGPVTRKHPLVTRYFTFPFEEIDFSMEESMIHLPNK  ACFLMAHNGWVMGDDPLRNFAEPGSEVYLRRELICWGDSVKLRYGNKP  EDCPYLWAHMKKYTEITATYFQGVRLDNCHSTPLHVAEYMLDAARNLQ  PNLYVVAELFTGSELDLNVFVTRLGISSLIREAMSAYNSHEEGRLVYRYG  GEPVGSFVQPCLRPLMPAIAHALFMDITHDNECPIVHRSAYDALPSTTIVS  MACCASGSTRGYDELVPHQISVVSEERFYTKWNPEALPSNTGEVNFQSGII  AARCAISKLHQELGAKGFIQVYVDQVDEDIVAVTRHSPSIHQSVVAVSRT  AFRNPKTSFYSKEVPQMCIPGKIEEVVLEARTIERNTKPYRKDENSINGTPD  ITVEIREHIQLNESKIVKQAGVATKGPNEYIQEIEFENLSPGSVIIFRVSLDPH  AQVAVGILRNHLTQFSPHFKSGSLAVDNADPILKIPFASLASRLTLAELNQI  LYRCESEEKEDGGGCYDIPNWSALKYAGLQGLMSVLAIRPKNDLGHFP  CNNLRSGDWMIDYVSNRLISRGTIAEVGKWLQAMFFYLKQIPRYLIPCY  FDAILIGAYTTLLDTAWKQMSSFVQNGSTFVKHLSLGSVQLCGVGKFP  PILSPALMDVPYRLNEITKEKEQCCVSLAAGLPHFSSGIFRCWGRDTFI  GILLITGRYVEARNILAFAGTLRHGLIPNLLGEGIYARYNCRDAVWWWL  QCIQDYCKMVPNGLDILKCPVSRMYPTDDSAPLPAGTLDQPLFEVIQ  EAMQKHMQGIQFRERNAGPQIDRNMKDEGFNITAGVDEETGFVYGGN  RNFNCGTWMDKMGESDRARNRGIPATPRDGSAVEIVGLSKSAVRWLL  ELSKKNIFYHEVTVKRHGKAIKVSYDEWNRKIQDNFEKLFHVSEDP  SDLNEKHPNLVHKGRIYKDSYGASSPWCDYQLRPNFTIAMVVAPEL  FTTEKAWKALEIAEKLLGPLGMKTLDPDDMVYCGIYDNALDNDNYN  LAKGFNYHQPEWLWPIGYFLRAKLYFSRLMGPETTAKTIVLVKNV  LSRHYVHLERSPWKGLPELTNENAQYCPFSCETQAWSIATILETLYDL</p>	AGL
113	<p>MEEGMNVLHDFGIQSTHYLQVNYQDSQDWFILVSVIADLRN  AFYVLFPIWFHLQEAVGIKLLWVAVIGDWLNLVFKWILFGQRPY  WWVLDTDYYSNTSVPLIKQFPVTCETGPGSPSGHAMGTAGVYYV  MVTSTLSIFQGKIKPTYRF RCLNVILWLGFAVQLNVCLSRIY  LAAHFPHQVVAGVLSGIAVAETFSHIHSIYNASLKKYFLITFFL  FSFAIGFYLLLKGLGVDLLWTLEKAQRWCEQPEWVHIDTTPFAS  LLKNLGTFLFGLALNSSMYRESCKGKLSKWLPFRLSSIVASLV  LLHVFDLSKPPSQVELVFYVLSFCKSAVVPLASVSVIPYCLAQV  LGQPHKKS</p>	G6PC

10

20

30

40

50

114	MAAPMTPAARPEDYEAAALNAALADVPELARLLEIDPYLKPYAVDFQRRY KQFSQILKNIGENEGGIDKFSRGYESFGVHRCADGGLYCKEWAPGAEGVF LTGDFNGWNPFSYPYKKLDYDGKWELYIPPKQNKSVLVPHGSKLKVVITS KSGEILYRISPWAKYVVREGDNVNYDWHWDPEHSYEFKHSRPPKPRSLR IYESHVGISSHEGKVASYKHFTCNVLPRIKGLGYNCIQLMAIMEHAYYASF GYQITSFFAASSRYGTPEELQELVDTAHSMGIIVLLDVVHSHASKNSADGL NMFDTGTDSCYFHSGPRGTHDLWDSRLFAYSSWEILRFLSNIRWWLEEYR FDGFRFDGVTSMLYHHHGVGQGFSGDYSEYFGLQVDEDALTYLMLANH LVHTLCPDSITIAEDVSGMPALCSPISQGGGGFDYRLAMAIPDKWIQLLKE FKDEDWNMGDIVYTLNRRYLEKCIAYAESHQALVGDKSLAFWLMDA EMYTNSVLTPTFPVIDRGIQLHKMIRLITHGLGGEGYLNFMGNEFGHPE WLDFFPRKGNNEHYARRQFHLTDDDLLRYKFLNFDLDRDMNRLEERYG WLAAPQAYVSEKHEGNKIIAFERAGLLFIFNFHPSKSYTDYRVGTALPGKF KIVLDSDAAEYGGHQRLDHDSTDFSEAFEHNGRPYSLLVYIPSRVALILQN VDLPN	GBE1
115	MRSRSNSGVRLDGYARLVQQTILCHQNPVTGLLPASYDQKDAWVRDNDV YSILAVWGLGLAYRKNADRDEDKAKAYELEQSVVKLMRGLLHCMIRQV DKVESFKYSQSTKDSLHAKYNTKTCAVVGDDQWGHLLQDQATSVYLLF LAQMTASGLHIIHSLDEVNFIQNLVFYIEAAYKTADFGIWERGDKTNQGIS ELNASSVGMKAALAEALDELDFGVKGGPQSVIHVLADEVQHCQSILNSL LPRASTSKEVDASLLSVVSFPFAVEDSQLVELTKQEITKLQGRYGCCRF LRDGYKTPKEDPNRLYYEPAELKLFENIECEWPLFWTYFILDGVFSGNAE QVQEYKEALEAVLIKGNKGVPLPELYSVPPDRVDEEYQNPHTVDRVPM GKLPHMWGQSLYILGSLMAEGFLAPGEIDPLNRRFSTVPKPDVVVQVSIL AETEEIKTILKDKGIYVETIAEVYPIRVQPARILSHIYSSLGCNRMKLSGRP YRHMVGLGTSKLYDIRKTIFTFTFPQFIDQQQFYALDNKMIVEMLRDLS YLCSRWRMTGQPTITFPISHSMLDEDGTSLNSSILAALRKMQDGYFGGAR VQTGKLEFLTTSCTHLSFMDFGPEGKLYSEDDYDNDYDLESIGNWMND YDSTSHARCDEVARYLHLLAHTAPHPKLAPTSQKGGDLDRFQAAVQTT CDLMSLVTKAKELHVQNVHMYLPTKLFQASRPSFNLLDSPHPRQENQVP SVRVEIHLPRDQSGEVDFKALVLQKETSLSLQEQADILYMLYTMKGPDW NTELYNERSATVRELLTELYGKVGEIRHWGLIRYISGILRKKVEALDEACT DLLSHQKHLTVGLPPEPREKTISAPLPEALTQLIDEASEGDMSISILTQEIM VYLAMYMRTQPGLFAEMFRLRIGLIQVMATELAHSLRCSAEEATEGLMN LSPSAMKNLLHHILSGKEFGVERSVPRTDSNVSPAISIHEIGAVGATKERT GIMQLKSEIKQVEFRRLSISAESQSPGTSMTSPSSGFSFSAIDQSSKDSRQG QWQRRRLD GALNRVPVGFYQKVWVVLQKCHGLSVEGFVLPSSTTREM TPGEIKFSVHVESVLNRVPQPEYRQLLVEAILVLTMLADIEIHSIGSIIAVEK IVHIANDLFLQEQLTGADDTMLAKDPASGICTLLYDSAPSGRFGTMTYL SKAAATYVQEFLPHSICAMQ	PHKA1

10

20

30

40

50

116	MRSRSNSGVRLDGYARLVQQTILCYQNPVTGLLSASHEQKDAWVRDNIY SILAVWGLGMAYRKNADRDEDKAKAYELEQNVVKLMRGLLQCMMRQV AKVEKFKHTQSTKDSLHAKYNTATCGTVVGDDQWGHLLQVDATSLFLLF LAQMTASGLRIIFTLDEVAFIQNLVFIYIAAYKVADYGMWERGDKTNQGI PELNASSVGMAKAALEAIDELDLFGAHGGRKSVIHVLPDEVEHCQSILFS MLPRASTSKEIDAGLLSIISFPFAVEDVNLVNVTKNEIISKLQGRYGCCRF LRDGYKTPREDPNRLHYDPAELKLFENIECEWPVFWTYFIIDGVFSGDAV QVQEYREALEGILIRGKNGIRLVPELYAVPPNKVDEEYKNPHTVDRVPMG KVPHLWGQSLYILSSLLAEGFLAAGEIDPLNRRFSTSVKPDVVVQVTVLA ENNHKDLLRKHGVNVQSIADIHPQVQGRILSHIYAKLGRNKNMNLSGR PYRHIGVLGTSKLYVIRNQIFTFPQFTDQHFFYLALDNEMIVEMRLIELA YLCTCWRMTGRPTLTFPISRTMLTNDGSDIHS AVLSTIRKLEDGYFGGAR VKLGNLSEFLTTSFYTYLTFLDPCDEKLFDNASEGTFSPDSDSL VGYLE DTCNQESQDELDHYINHLLQSTSLRSYLPPLCKNTEDRHVFS AIHSTRDILS VMAKAKGLEVPFVPMTLPTKVL SAHRKSLNLVDSQPPLLEKVPESDFQW PRDDHGDVDCEKLVEQLKDCSNLQDQADILYILYVIKGPSWDTNLSGQH GVTVQNLLGELYGKAGLNQEWGLIRYISGLLRKKVEVLAEACTDLLSHQ KQLTVGLPPEPREKIISAPLPPEELTKLIYEASGQDISIAVLTQEIVVYLAMY VRAQPSLFVEMLRRLRIGLIIQVMATELARSLNCSGEEASESLMNLSPFDMK NLLHHILSGKEFGVERS VRPIHSSTSSPTISIHEVGHTGVTKTERSGINRLRS EMKQMTRRFS ADEQFFSVGQAASSAHSSKSARSSTPSSPTGTSSSDSGGH HIGWGERQQWLRRRRLDGAINRVPVGFYQRVWKILQKCHGLSIDGYVL PSSTREMPHEIKFAVHVESVLNRVPQPEYRQLLVEAIMVLTLLSDTEMT SIGGIIHVDQIVQMASQLFLQDQVSIGAMDTLEKDQATGICHFFYDSAPSG AYGTMTYLTRAVASYLQELLPNNGCQM	PHKA2
-----	---	-------

10

20

30

40

50

117	MAGAAGLTAEVSWKVLERRARTKRSGSVYEPLK SINLPRPDNETLWDKL DHYYRIVKSTLLLYQSPTTGLFPTKTCGGDQKAKIQDSLYCAAGAWALA LAYRRIDDDKGRTHELEHSAIKCMRGILYCYMRQADKVQQFKQDPRPTT CLHSVFNVHTGDELLSYEEYGHLQINAVSLYLLYL VEMISSGLQIHYNTDE VSFIQNLVFCVERVYRVPDFGVWERGSKYNNGSTELHSSSVGLAKAALE AINGFNLFNGQCSWSVIFVDLDAHNRNRQTLCSLLPRESRSHNTDAALL PCISYPAFALDDEVLFQSQTLDKVVVRKLGKYGFKRFLRDGYRTSLEDPNR CYYKPAEIKLFDGIECEFPFFLYMMIDGVFRGNPKQVQEYQDLLTPVLHH TTEGYPPVVKYYPADVFVEYKNNPGSQKRFPSNCGRDGKFLWGQAL YIIAKLLADELISPKDIDPVQRYVPLKDQRNVSMRFSNQGPLENDLVVHV ALIAESQRLQVFLNTYGIQTQTPQQVEPIQIWPQQELVKAYLQLGINEKLG LSGRPDRPIGCLGTSKIYRILGKTVVCYPIIFDLSDFYMSQDVFLIDDIKNA LQFIKQYWKMHGRPLFLVLIREDNIRGSRFNPILDMLAALKKGIIGGVKVVH VDRLQTLISGAVVEQLDFLRISDTEELPEFKSFEELPPKHSKVKRQSSTPS APELGQQPDVNISEWKDKPTHEILQKLNDCSCLASQAILLGILLKREGPNFI TKEGTVSDHIERVYRRAGSQKLWLAVRYGAAFTQKFSSSIAPHITTFVLVH GKQVTLGAFGHEEEVISNPLSPRVIQNIYYKCNTHDEREAVIQQELVIHIG WIISNNPELFSGMLKIRIGWIIHAMEYELQIRGGDKPALDLYQLSPSEVKQL LLDILQPQNGRCWLNRRQIDGSLNRTPTGFYDRVWQILERTPNGHIVAGK HLPQQPTLSDMTMYEMNFSLLVEDTLGNIDQPQYRQIVVELLMVVSIVLE RNPELEFQDKVDLDRLVKEAFNEFQKQDQSRLKEIEKQDDMTSFYNTPLG KRGTC SYLTKAVMNLLLEGEVKPNNDPCLIS	PHKB
118	MTLDVGPEDELPDWAAAKEFYQKYDPKDVIGRGVSSVVRRCVHRATGH EFAVKIMEVTAERLSPEQLEEVREATRRETHILRQVAGHPHIITLIDSYESSS FMFLVFDLMRKGELFDYLTEKVALSEKETRSIMRSLEAVSFLHANNIVH RDLKPENILLDDNMQIRLSDFGFSCHLEPGEKLRCLCGTPGYLAPEILKCS MDETHPGYGKEVDLWACGVILFTLLAGSPPFWHRRQILMLRMIMEGQYQ FSSPEWDDRSSTVKDLISRLQVDPEARLTAEQALQHPFFERCEGSQPWNL TPRQRFRVAVWTVLAAGRVALSTHRVRPLTKNALLRDPYALRSVRHLID NCAFRLYGHVKKGEQQNRAALFQHRPPGPFPIMGPEEEGDSAAITEDA VLVLG	PHKG2
119	MAAQGYGYRTVIFSAMFGGYSLYYFNKRTFSFVMPSLVEEIPLDKDDL FITSSQSAAYAISKFVSGVLSQMSARWLFSSGLLLVGLVNIFFAWSSTVP VFAALWFLNGLAQGLGWPPCGKVLKWFEPSSQFGTWWAILSTSMNLAG GLGPILATILAQSYSWRSTLALSGALCVVVSFLCLLIHNEPADVGLRNL PMPSEGKKGSLKEESTLQELLLSPYLWVLSTGYLVVFGVKTCCTDWGQFF LIQEKQSALVGSSYMSALEVGGVLSIAAGYLSDRAMAKAGLSNYGNP RHGLLLMMAGMTVSMYLFRTVTSVSPKLWILVLGAVFGFSSYGPIALF GVIANESAPPNLCGTSHAIVGLMANVGGFLAGLPFSTIAKHYSWSTAFWV AEVICAASTAAFFLLRNIRTKMGRVSKKAE	SLC37A4

10

20

30

40

50

120	MAAPGPALCLFDVDGTLTAPRQKITKEMDDFLQKLRQKIKIGVVGGSDFE KVQEQLGNDVVEKYDYVFPENGLVAYKDGKLLCRQNIQSHLGEALIQDL INYCLSYIAKIKLPKKRGTIEFRNGMLNVSPIGRSCSQEERIEFYELDKKE NIRQKFVADLRKEFAGKGLTFSIGGQISFDVFPDGDWVKRYCLRHVENDGY KTIYFFGDKTMPGGNDHEIFTDPRTMGYSVTAPEDTRRICELLFS	PMM2
121	MPSETPQAEVGP TGCPHRSGPHSAKGSLEKGPEDKEAKEPLWIRPDAPSR CTWQLGRPASESPHHHTAPAKSPKILPDILKKIGDTPMVRINKIGKKFGLK CELLAKCEFFNAGGSVKDRISLRMIEDAERDGLKPGDTIIEPTSGNTGIGL ALAAAVRGYRCIIVMPEKMSSEKVDVLRALGAEIVRTPTNARFDSPEHSV GVAWRLKNEIPNSHILDQYRNASNPLAHYDTTADEILQQCDGKLDMLVA SVGTGGTITGIARKLKEKCPGCRHIGVDPEGSILAEPEELNQTEQTTYEVEGI GYDFIPTVLDRTVVDKWFKSNDEEAFTFARMLIAQEGLLCGGSAGSTVAV AVKAAQELQEGQRCVVILPDSVRNYMTKFLSDRWMLQKGFLEEDLTE KKPWWWHLRVQELGLSAPLTVLPTITCGHTIEILREKGFQAPVVDEAGV ILGMVTLGNMLSSLLAGKVQPSDQVGVVIYKQFKQIRLDTLGRLSHILE MDHFALVVHEQIQYHSTGKSSQRQMVFGVVTAIDLLNFVAAQERDQK	CBS
122	MSFIPVAEDSDFPIHNLPGYVVFSTRGDPRPRIGVAIGDQILDLSIIKHLFTGP VLSKHQDVFNQPTLNSFMGLGQAAWKEARVFLQNLLSVSQAARLRDDTEL RKCAFISQASATMHLPATIGDYTDFYSSRQHATNVGIMFRDKENALMPN WLHLPVGYHGRASSVVVSGTPIRRPMGQMKPDDSKPPVYGACKLLDME LEMAFFVGPNGRLGEPPISKAHEHIFGMVLMNDWSARDIQKWEYVPLGP FLGKSFGTTVSPVVPMDALMPFAVNPQKQDPRPLPYLCHDEPYTFDINL SVNLKGEGMSQAATICKSNFKYMYWTMLQQLTHHSVNGCNLRPGDLLA SGTISGPEPENFGSMLELSWKGTKPIDLNGQTRKFLLDGDEVIITGYCQG DGYRIGFGQCAGKVLPAALLPS	FAH
123	MDPYMIQMSSKGNLPSILDVHVNVGGRSSVPGMKMKGRKARWSVRPSDM AKKTFNPIRAIVDNMKVKPNPNTMISLSIGDPTVFGNLPTDPEVTQAMK DALDSGKYNGYAPSIGFLSSREEIASYYHCPEAPLEAKDVILTSGCSQAIDL CLAVLANPGQNILVPRPGFSLYKTLAESMGIEVKLYNLLPEKSWEIDLKQL EYLIDEKTA CLIVNPNPCGSVFSKRHLQKILAVARQCVPILADEIYGD MVFSDCKYEPLATLSTDVPILSCGGLAKRWLVPGWRLGWILIHDRRDFG NEIRDGLVKLSQRILGPCTIVQGALKSILCRTPGEFYHNTLSFLKSNADLCY GALAAIPGLRPVRPSGAMYL MVGIEMEHFPEFENDVEFTERLVAEQSVHC LPATCFEYPNFIRVVITVPEVMMLEACSRIQEFCEQHYHCAEGSQEECDK	TAT
124	MSRSGTDPQQRQASEADAAAATFRANDHQHIRYNPLQDEWVLSAHR MKRPWQGQVEPQLLKTVP RHDPLNPLCPG AIRANGEVNPQYDSTFLFDN DFPALQPDAPSPGSDHPLFQAKSARGVCKVMCFHPWSDVTLPLMSVPEI RAVVDAWASVTEELGAQYPWVQIFENKGAMMGCSNPHPHCQVWASSFL PDIAQREERSQQAYKSQHGEPLLMEYSRQELLRKERLVTSEHWLVLPF WATWPYQTL LPRRHVRRLPELTPAERDDLASIMKLLTKYDNLFETSFP YSMGWHGAPTGSEAGANWNHWQLHAHYYP LLSATVRKFMVGYEML AQAQRDLTPEQAERLRALPEVHYHLGQKDRETATIA	GALT

10

20

30

40

50

125	MAALRQPQVAELLAEARRAFREFGAEPPELAVSAPGRVNLIGEHTDYNQ GLVLPMALELMTVLVGSPRKDGLVSLTTSEGADEPQRLQFPLPTAQRSL EPGTPRWANYVKGVIIQYYPAAPLPGFSAVVVSSVPLGGGLSSSASLEVAT YTFLQQLCPDSGTIAARAQVCQQAHSFAGMPCGIMDQFISLMGQKQKHA LLIDCRSLETSLVPLSDPKLAVLITNSNVRHSLASSEYPVRRRQCEEVARAL GKESLREVQLEEEAARDLVSKEGFRRARHVVEIRRTAQAAAALRRGD YRAFGRLMVESHRLRDDYEVSCPELDQLVEAALAVPGVYGSRMTGGGF GGCTVTLLLEASAAPHAMRHIQEHYGGTATFYLSQAADGAKVLCL	GALK1
126	MAEKVLVTGGAGYIGSHTVLELLEAGYLPVIDNFHNAFRGGGSLPESLR RVQELTGRSVEFEEMDILDQGALQRLFKKYSFMAVIHFAGLKAVGESVQ KPLDYRVNLTGTIQLLEIMKAHGKLNLFSSSATVYGNPQYPLPLDEAHP TGGCTNPYGKSKFFIEEMIRDLQADKTWNAVLLRYFNPTGAHASGCIGE DPQGIPNNLMPYVSQVAIGRREALNVFGNDYDTEGTVRDIYHVVDLA KGHIAALRKLKEQCGCRIYNLGTGTGYSVLQMVQAMEKASGKKIPYKVV ARREGDVAACYANPSLAQEELGWTAAALGLDRMCEDLWRWQKQNPSTGF GTQA	GALE
127	MAEQVALSRTQVCGILREELFQGDFAHQSDTHIFIMGASGDLAKKKIYPT IWWLFRDGLLPENTFIVGYARSRLTVADIRKQSEPFKATPEEKLKLEDFE ARNSYVAGQYDDAASYQRLNSHMNALHLGSQANRLFYLALPPTVYEAV TKNIHESCMSQIGWNRIIVEKPFGRDLQSSDRLSNHISLRFREDQIYRIDHY LGKEMVQNLMLVLRFANRIFGPIWNRDNIACVILTFKEPFGTEGRGGYFDE FGIIRDVMQNHLQLMLCLVAMEKPASTNSDDVRDEKVKVLKCISEVQAN NVVLGQYVGNPDGEGEATKGYLDDPTVPRGSTTATFAAVVLYVENERW DGVPFILRCGKALNERKAEVRLQFHDVAGDIFHQCKRNELVIRVQPNEA VYTKMMTKKPGMFFNPEESELDTYGNRYKNVCLPDAYERLILDVFCGS QMHFVRSDELREAWRIFTPLLHQIELEKPKPIPIYIYGSRGPTAEDELMKRV GFQYEGTYKVVNPHKL	G6PD
128	MAEDKSKRDSIEMSMKGCQTNNGFVHNEDILEQTPDPSSTDNLKHSTR GILGSQEPDFKGVQPYAGMPKEVLFQFSGQARYRIPREILFWLTVASVLVL IAATIAIALSPKCLDWWQEGPMYQIYPRSFKDSNKDGNGLKGIQDKLD YITALNIKTWITSFYKSSLKDFRYGVEDFREVDPIFGTMEDEFENLVAAIH DKGLKLIIDFIPNHTSDKHIWFQLSRTRTGKYTDYIWHDCTHENGKTIPP NNWLSVYGNSSWHFDEVRNQCYPFHQFMKEQPDNFRNPDVQEEIKEILR FWLTKGVDGFSLDAVKFLLEAKHLRDEIQVNKTQIPDVTQYSELVYHDFE TTQVGMHDIRSFRQTMQYSTEPGRYRFMGTEAYAESIDRTVMYYGLP FIQEADFPFNYSMLDVTSGNSVYEVITSWMENMPEGKWPNNMIGGPD SSRLTSRLGNQYVNVMMNMLLFTLPGTPITYYGEEIGMGNIVAANLNEYSYD INTLRKSPMQWDNSSNAGFSEASNTWLPTNSDYHTVNVVDVQKTQPRSA LKLQDLKLLHANELLNLRGWFCHLRNDSHYVVYTRELDGIDRIFIVVLN FGESTLLNLHNMISGLPAKMRIRLSTNSADKGSKVDTSIGIFLDKGEGLIFE HNTKNLLHRQTAFDRRCFVSNRACYSSVLNLYTSC	SLC3A1

10

20

30

40

50



129	<p>MGDTGLRKRREDEKSIQSQEPKTTSLQKELGLISGSIHVGTHIGSGIFVSPKS  VLSNTEAVGPCLIIWAACGVLATLGALCFAELGTMITKSGGEYPYLMEAY  GPIPAYLFSWASLIVIKPTSFAIICLSFSEYVCAPFYVGCKPPQIVVKCLAAA  AILFISTVNSLSVRLGYSYVQNIFTAAKLIVAVIIIISGLVLLAQGNTKNFDNS  FEGAQLSVGAISLAFYNGLWAYDGNWQLNYITEELRNPYRNLP LAIIIIGIP  LVTACYILMNVSYFTVMTATELLQSQAVAVTFGDRVLYPASWIVPLFVAF  STIGAANGTCFTAGRLIYVAGREGHMLKVL SYISVRRLTPAPAIIFYGIIATI  YIIPGDINSLVNYFSFAAWLFYGLTILGLIVMRFRKELERPIKVPVVIPVL  MTLISVFLVLAPIISKPTWEYLYCVLFILSGLLFYFLFVHYKFGWAQKISKPI  TMHLQMLMEVVPPEEDPE</p>	SLC7A9
130	<p>MVNEARGNSSLNPCLEGSASSGSESSKSSRCSTPGLDPERHERLREKMR  RRLESGDKWFSLEFFPPRTAEGAVNLISRFDMAAGGPLYIDVTWHPAGD  PGSDKETSSMMIASTAVNYCGLETILHMTCCRQRLEEITGHLHKAKQLGL  KNIMALRGDPIGDQWEEEEGGFNAYVDLVKHIRSEFGDYFDICVAGYPKG  HPEAGSFEADLKHLKEKVSAGADFIITQLFFEADTFFRFV KACTDMGITCPI  VPGIFPIQGYHSLRQLVKLSKLEVPQEIKDVIEPIKDNDAAIRNYGIELAVSL  CQELLASGLVPGLHFYTLNREMATTEVLKRLGMWTE DPRRPLPWALSAH  PKRREEDVRPIFWASRPKSYIYRTQEWDEFPNR WGNSSSPAFGELKDYY  LFYLKSKSPKEELLKMWGEELTSEESVFEVFLYLSGEPNRNGHKVTCLP  WNDEPLAAETSLLKEELLRVNRQGILTINSQPNINGK PSSDPIVGWGPSGG  YVFQKAYLEFFTSRETAEALLQVLK KYELRVNYHLVNVKGENITNAPELQ  PNAVTWGIFPGREIIQPTVVDPV SFMFWKDEAFALWIERWGKLYEEESPS  RTIIQYIHDNYFLVNLVDNDFPLDNCLWQVVEDTLELLNRPTQNARETEA  P</p>	MTHFR

10

20

30

40

50

131	MSPALQDLSQPEGLKKTLRDEINAILQKRIMVLDGGMGMTMIQREKLNEEH FRGQEFKDHARPLKGNNDILSITQPDVIYQIHKEYLLAGADIETNTFSSTSI AQADYGLEHLAYRMNMCSAGVARKAAEEVTLQTGIKRFVAGALGPTNK TLVSPSPVERPDYRNITFDELVEAYQEQAQKGLLDGGVDILLIETIFDTANA KAALFALQNLFEKEYAPRPIFISGTIVDKSGRTLSGQTGEGFVISVSHGEPL CIGLNCALGAAEMRPFIEIIGKCTTAYVLCYPNAGLPNTFGDYDETPSMM AKHLKDFAMDGLVNIVGGCCGSTPDHIREIAEAVKNCKPRVPPATAFEGH MLLSGLEPFRIGPYTNFVNIGERCNVAGSRKFAKLIMAGNYEEALCVAKV QVEMGAQVLDVNMDDGMLDGPSAMTRFCNLIASEPDIAKVPLCIDSSNF AVIEAGLKCCQGKCIVNSISLKEGEDDFLEKARKIKKYGAAMVVMFAFDEE GQATETDTKIRVCTRAYHLLVKKLGFNPNDIIFDPNILTIGTGMEEHNLYAI NFIHATKVIKETLPGARISGGLSNLSFSFRGMEAIREAMHGVFLYHAIKSG MDMGIVNAGNLPVYDDIHKELLQLCEDLIWNKDPEATEKLLRYAQTQGT GGKKVIQTDEWRNGPVEERLEYALVKGIEKHIEDTEEARLNQKKYPRPL NIEGPLMNGMKIVGDLFGAGKMFLPQVIKSARVMKKA VGHLPFMEKER EETRVLNGTVEEEDPYQGTIVLATVKGDVHDIGKNIVGVVLGCNNFRVID LGMVTPCDKILKAALDHDKADIIGLSGLITPSLDEMIFVAKEMERLAIRPLLI GGATTSKTHTAVKIAPRYSAPVIHVLDASKSVVCSQLLDENLKDEYFEEI MEEYEDIRQDHYESLKERRYPLSQARKSGFQMDWLSEHPVKPTFIGTQ VFEDYDLQKLVYIDWKPFFDVWQLRGKYPNRGFPKIFNDKTVGGARK VYDDAHNMLNTLISQKCLRARGVVGFWPAQSIQDDIHLIAEAAVPQAAE PIATFYGLRQQAEKDSASTEPYYCLSDFIAPLHSGIRDYLGLFAVACFGVE ELSKAYEDDGDYSSIMVKALGDRLAEFAEELHERVRRELWAYCGSEQ LDVADLRRRLRYK GIRPAPGYPSQPDHTEKLTMWRLADIEQSTGIRLTESLA MAPASAVSGLYFSNLKSKYFAVGKISKDQVEDYALRKNISVAEVEKWL G PILGYDTD	MTR
132	MGAASVRAGARLVEVALCSFTVTCLEVMRFRLLLYATQQGQAKAIAEEI CEQAVVHGFSADLHCISESDKYDLKTETAPLVVVVSTTGTGDPDPDTARKF VKEIQNQTLPVDFFAHLRYGLLGLGDSEYTYFCNGGKIIDKRLQELGARH FYDTGHADDCVGLLELVVEPWIAGLWPALRKHFRSSRGQEEISGALPVASP ASSRTDLVKSELLHIESQVELLRFDSDGRKDSEVLKQNAVNSNQSNVIE DFESSLTRSVPLSQASLNIPGLPPEYLQVHLQESLQEEESQVSVTSADPVF QVPISKAVQLTTNDAIKTTLLVELDISNTDFSYQPGDAFSVICPNSDSEVQS LLQRLQLEDKREHCVLLKIKADTKKKGATLPQHIPAGCSLQFIFTWCLEIR AIPKKAFLRALVDYTSDSAERRLQELCSKQGAADYSRFRVACACLLDL LLAFSPCQPPLSLLLEHLPLKQPRPYSCASSLPHPGKLFHVFNIVEFLSTAT TEVLRKGVCTGWLALLVASVLQPNIHASHEDSGKALAPKISISPRTTNSFH LPDDPSIPIIMVGPGTGIAPFIGFLQHREKLQEHPDGNFGAMWLFFGCRH KDRDYLFKELRHFLKHGILTHLKVFSRDPVGEESAPAKYVQDNIQLH GQVARILLQENGHYVCGDAKNMAKDVHDALVQIISKEVGVKLEAMK TLATLKEEKRYLQDIWS	MTRR

10

20

30

40

50

133	MPEQERQITAREGASRKILSKLSLPTRAWEPAAMKKSFAFDNVGYEGGLDG LGPSSQVATSTVRILGMTQCSCVKSIEDRISNLKGIISMKVSLEQGSATVKY VPSVVCLQQVCHQIGDMGFEASIAEGKAASWPSRSLPAQEAVVKLRVEG MTCQSCVSSIEGKVRKLGQVVRVKVSLSNQEAVITYQPYLIQPEDLRDHV NDMGFEAAIKSKVAPLSLGPIDIERLQSTNPKRPLSSANQNFNNSETLGHQ GSHVVTLQLRIDGMHCKSCVLNIEENIGQLLGVQSIQVSLENKTAQVKYD PSCTSPVALQRAIEALPPGNFKVSLPDGAEGSGTDHRSSSSHSPGSPPRNQV QGTCSTTLIAIAGMTCASCVHSIEGMISQLEGVQQISVSLAEGTATVLYNPS VISPEELRAAIEDMGFEASVVSESCSTNPLGNHSAGNSMVQTTDGTPTSQV EVAPHTGRLPANHAPDILAKSPQSTRAPVAPQKCFQIKGMTCASCVSNIER NLQKEAGVLSVLVALMAGKAEIKYDPEVIQPLEIAQFIQDLGFEEAVMED YAGSDGNIELTITGMTASCVHNIESKLTRTNGITYASVALATSKALVKFD PEIIGPRDIIKIEEIGFHASLAQRNPNNAHHLDHKMEIKQWKKSFLCSLVFGI PVMALMIYMLIPSNEPHQSMVLDHNIIPGLSILNLIFFILCTFVQLLGGWYF YVQAYKSLRHRANMDVLIVLATSIAVYVYSLVILVVAVAEKAERSPVTFP DTPPMLFVFIALGRWLEHLAKSKTSEALAKLMSLQATEATVVTLGEDNLII REEQVPMELVQRGDIVKVVPGGKFPVDGKVLEGNTMADESLITGEAMPV TKKPGSTVIAGSINAHGSVLIKATHVGNDDTLAQIVKLVEEAQMSKAPIQQ LADRFSGYFVFFIIIMSTLTLVWVWVIGFIDFGVQRYFPNPNKHISQTEVIIR FAFQTSITVLCIACPCSLGLATPTAVMVGTVAAQNGILIKGGKPLEMAH KIKTVMFDKTGTITHGVPRVMRVLVLLGDVATLPLRKLAVVGTAEASSE HPLGVAVTKYCKEELGTETLGYCTDFQAVPGCGIGCKVSNVEGILAHSER PLSAPASHLNEAGSLPAEKDAVPQTFVSVLIGNREWLRNGLTISSDVSDA MTDHEMKGQTAILVAIDGVLGMIADAVKQEAALAVHTLQSMGVDV VLITGDNRKTARAIATQVGINKVFAEVLPSHKVAKVQELQNKGKKVAMV GDGVNDSPALAQADMGVAIGTGTDAIEAADVVLIRNDLLDVVASIHLS KRTVRRIRINLVLALIYNLVGIPIAAGVFMPIGIVLQPWMGSAAMAASSVS VVLSSLQLKCYKKPDLERYEAQAHGHMKPLTASQVSVHIGMDDRWRDS PRATPWDQVSYVSQVSLSSLTSDKPSRHSAAADDDGDKWLLLLNGRDEE QYI	ATP7B
134	MATRSPGVVISDDEPGYDLDFCIPNHYAEDLERVFIPHGLIMDRTERLAR DVMKEMGGHHIVALCVLKGGYKFFADLLDYIKALNRNSDRSIPMTVDFI RLKSYCNDQSTGDIKVI GGDDLSTLTGKNVLIVEDIIDTGKTMQTLLSLVR QYNPKMVKVASLLVKRTPRSVGYKPDFVGFEPDKFVVGALDYNEYFR DLNHVCVISETGKAKYKA	HPRT1

10

20

30

40

50

135	MGEPEGQSPSPRSSHGSPPTLSTLTLLLLCGHAHSQCKILRCNAEYVSSTLS LRGGGSSGALRGGGGGGRRGGVGSGLCRALRSYALCTRRTARTCRGD LAFHSAVHGIEDLMIQHNC SRQGPTAPPPRGPALPGAGSGLPAPDPCDYE GRFSRLHGRPPGFLHCASFDPHVRSFHHHFHTCRVQGAWPLLDNDFLV QATSSPMALGANATATRKLTIIFKNMQECIDQKVYQAEVDNLPVAFEDGS INGDRPGGSSLSIQTANPGNHVEIQAA YIGTTIIRQTAGQLSFSIKVAEDV AMAFSAEQDLQLCVGGCPPSQRLSRSENRRRGAITIDTARRLCKEGLPVE DAYFHSCVFDVLISGDPNFTVAAQA ALEDARAFLPDLEKLHLFPSDAGVP LSSATLLAPLLSGLFVLWLCIQ	HJV
136	MALSSQIWAACL LLLLLL ASLTSGSVFPQQTGQLAELQPQDRAGARASW MPMFQRRRRRDTHFPICIFCCGCCHRSKCGMCKT	HAMP
137	MRSPRTRGRSGRPLSLLLALLCALRAKVC GASGQFELEILSMQNVNGELQ NGNCCGGARNPGDRKCTRDECDTYFKVCLKEYQSRVTAGGPCSFGSGST PVIGGNTFNLKASRGNDRNRIVLPFSFAWPRS YTLLVEAWDSSNDTVQPD SII EKASHSGMINPSRQWQTLKQNTGVAHF EYQIRVTCDDYYYGFGCNKF CRPRDDFFGHYACDQNGNKTCEGWMGPECNRAICRQGCS PKHGSCKL PGDCRCQYGWQGLYCDKCIHPGCVH GICNEPWQCLCETNWGGQLCDK DLNYCGTHQPCLNGGTC SNTGPDKYQCSCPEGYS GPNCEIAEHA CLSDPC HNRGSCKETSLGFECECSPGW TGPTCSTNIDDCSPNCSHGGTCQDLVNG FKCVCPPQWTGKTCQLDANECEAKPCVNAK SCKNLIASYYCDCLPGWM GQNC DININDCLGQCQNDASCRDLVNGYRCICPPGYAGDHCERDIDECAS NPCLNGGHCQNEINRFQCLCPTGFSGNLCQLDIDYCEPNPCQNGAQCYNR ASDYFCKCPEDYEGKNCSHLKDHCRTTPCEVIDSCTVAMASNDTPEGVR YISSNVC GPHGKCKSQSGGKFTCDCNKGFTGT YCHENINDCESNPCRNGG TCIDGVNSYKICSDGWEGAYCETNINDCSQNPCHNGGTCRDLVNDFYC DCKNGWKGKTCHSRDSQCDEATC NNGGTCYDEGDAFKCMCPGGWEGT TCNIARNSSCLPNPCHNGGTCV VNGESFTCVCKEGWEGPICAQNTNDCSP HPCYNSGTCVDGDNWYRCECAPGFAGPDCRININECQSSPCAFGATCVDE INGYRCVCPPGHSGAKCQEVSGRPCITMG SVIPDGA KWDDDCNTCQCLN GRIACSKVWCGRPRCLLHKGHSECP SGQSCIPILDDQCFVHPCTGVGECRS SSLQPVKTKCTSDSYQDNCANITFTFNKEMMSPGLTTEHICSELRNLNIL KNVSAEYSIYIACEPSANNEIHVAISAEDIRDDGNPIKEITDKIIDLVSKR DGNSSLIAAVA EVRVQRRPLKNRTDFLVPL LSSVLTVAWICCLVTA FYWC LRKRRKPGSHTHSASEDNTTNNVREQLNQIKNPIEKHGANTVPIKDYENK NSKMSKIRTHNSEVEEDMDKHQQKARFAKQPAYTLVDREKPPNGTPT KHPNWTNKQDNRDLESAQSLNRMEYIV	JAG1
138	MASHRLLLLCLAGLVFVSEAGPTGTGESK CPLMVKVLD AVRGSPAINVA VHVFRKAADDTWEPFASGKTSESGELHGLTTEEEFVEGIYKVEIDTKSYW KALGISPFHEHA EVVFTANDSGPRRYTIAALLSPYSYSTTAVVTNPKE	TTR

10

20

30

40

50

139	MASHKLLVTPPKALLKPLSIPNQLLLGPSPNLPPRIMAAGGLQMIGSMSK DMYQIMDEIKEGIQYVFQTRNPLTLVISGSGHCALEAALVNVLEPGDSFLV GANGIWQRAVDIGERIGARVHPMTKDPGGHYTLQEVEEGLAQHKPVLL FLTHGESSTGVLQPLDGFGECHRYKCLLLVDSVASLGGTPLYMDRQID ILYSGSQKALNAPPGTSLISFSDKAKKKMYSRKTTPFSFYLDIKWLANFW GCDDQPRMYHHTIPVISLYSLRESLALIAEQGLENSWRQHREAAAYLHGR LQALGLQLFVKDPALRLPTVTTVAVPAGYDWRDIVSYVIDHFDIEIMGGL GPSTGKVLRIKLLGCNATRENVDRVTEALRAALQHCPKKKL	AGXT
140	MKMRFLGLVVCLVLTWLHSESGGKLTAVDPETNMNVSEIISYWGFPSE EYLVEDEDGYILCLNRIPHGRKNHSDKGPKPVVFLQHGLLADSSNWVTNL ANSSLGFIADAGFDVWMGNSRGNTWSRKHKTLSVVSQDEFWAFSYDEM AKYDLPASINFILNKTGQEQVYVVGHSQGTIGFIAFSQIPELAKRIKMFFA LGPVASVAFACTSPMAKLGRLPDHLIKDLFGDKEFLPQSAFLKWLGHVCT HVILKELCGNLCFLLCGFNERNLMSRVVYTTHTSPAGTSVQNMLHWSQ AVKFQKFQAFDWGSSAKNYFHYNQSYPTYNVKDMLVPTAVWSSGGHD WLADVVDVNILLTQITNLVFHESIPEWEHLDFIWGLDAPWRLYNKIINLM RKYQ	LIPA
141	MASRLTLLTLLLLLAGDRASSNPATSSSSQDPESLQDRGEGKVATTVIS KMLFVEPILEVSSLPTTNSSTNSATKITANTTDEPTTQPTTEPTTQPTIQPTQ PTTQLPTDSPTQPTTGSFCPGPVTLCSDLESHSTEAVLGDALVDFSLKLYH AFSAMKKVETNMAFSPFSIASLLTQVLLGAGENTKTNLESILSYPKDFTCV HQALKGFTTKGVTSVSQIFHSPDLAIRDTFVNASRTLYSSSPRVLSNNSDA NLELINTWVAKNTNNKISRLDLSLPSDTRLVLLNAIYLSAKWKTTFDPKK TRMEPFHFKNVSVIKVPMMNSSKKYPVAHFIDQTLKAKVGQLQLSHNLSLVI LVPQNLKHRLEDMEQALSPSVFKAIMEKLEMSKFQPTLLTLPRIKVTTSQD MLSIMEKLEFFDFSYDLNLCGLTEDPDLQVSAMQHQTIVLELTETGVEAAA ASAISVARTLLVFEVQQPFLFVLWDQQHKFPVFMGRVYDRA	SERPING 1
142	MGSPLRFDRVVLVTGAGAGLGRAYALAFERAGLVVVNDLGGDFKGV GKGSAAADKVVVEIRRRGGKAVANYDSVEEGEKVVKTALDAFGRIDVVV NNAGILRDRSFARISDEDWDIIHRVHLRGSFQVTRAAWEHMKKQKYGRII MTSSASGIYGNFGQANYSAAKLGLLGLANSLAIEGRKSNIHCNTIAPNAGS RMTQTVMPEDLVEALKPEYVAPLVLWLCHESEENGGLFEVGAGWIGKL RWERTLGAIVRQKNHPMTPEAVKANWKKICDFENASKPQSIQESTGSHIEV LSKIDSEGGVSANHTSRATSTATSGFAGAIGQKLPPFSYAYTELEAIMYAL GVGASIKDPKDLKFIYEGSSDFCLPTFGVIIGQKSMGGGLAEIPGLSINF AKVLHGEQYLELYKPLPRAGKLCVAVVADVLDKGGVVIIMDVYSYSE KELICHNQFSLFLVGGSGFGGKRTSDKVKVAVAIPNRPPDAVLTDTTSLN QAALYRLSGDWNPLHIDPNFASLAGFDKPIHLGLCTFGFSARRVLQQFAD NDVSRFKAIKARFAKPVYPGQTLQTEMWKEGNRIHFQTKVQETGDIVISN AYVDLAPTSAGTSAKTPSEGGKLQSTFVFEEIGRRLKDIGPEVVKKVNAVFE WHITKGGNIGAKWTIDLKSGSGKVYQGPAGKAADTTIILSDEDFMEVVLG KLDPPQKAFFSGRLKARGNIMLSQKLQMLKDYAKL	HSD17B4

10

20

30

40

143	MEANGLGPQGFPELKNDFLRAAWGEETDYPVWCMRQAGRYLPEFRE TRAAQDFSTCRSPEACCELTLQPLRRFPLDAAIIFSDILVVPQALGMEVT MVPGKGPSFPEPLREEQDLERLRDPEVVASELGYVFQAITLTRQRLAGRV PLIGFAGAPWTLMTYMVEGGSSMAQAKRWLYQRPQASHQLLRILTDA LVPYLVGQVVAGAQUALQFESHAGHLGPQLFNKFALPYIRDVAKQVKAR LREAGLAPVPMIIFAKDGHFALEELAQAGYEVVGLDWTVPAPKKARECVG KTVTLQGNLDPICALYASEEEIGQLVKQMLDDFGPHRYIANLGHGLYPDM DPEHVGA FVDAVHKHSRLLRQN	UROD
144	MGPRARPALLLMLLQTA VLQGRLLRSHSLHYLFMGASEQDLGLSLFEA LGYVDDQLFVFDHESRRVEPRTPWVSSRISSQMWLQLSLSLKGWDHMF TVDFWTIMENHNHSHKESHTLQVILGCEMQEDNSTEGYWKYGYDGDHDL EFCPDTLDWRAAEPRAWPTKLEWERHKIRARQNRAYLERDCPAQLQQLL ELGRGVLDDQVPLVKVTHHVTSSVTTLRCRALNYYPQNITMKWLKDKQ PMDAKEFEPKDVLPNGDGT YQG WITLAVPPGEEQRYTCQVEHPGLDQPLI VIWEPSPSGTLVIGVISGIAVFVILFIGILFIILRKRQGSRGAMGHYVLAER E	HFE
145	MESKALLVLTAVWLQSLTASRGGVAAADQRRDFIDIESKFALRTPEDA EDTCHLIPGVAESVATCHFNHSSKTFMVIHGWTVTGMYESWVPKLVAAL YKREPDSNVIVVDWLSRAQEHYPVSAGYTKLVGQDVARFINWMEEEFN PLDNVHLLGYSLGAHAAGIAGSLTNKKNRITGLDPAGPNFEYAEAPSRL SPDDADFVDVLHTFTRGSPGRSIGIQKPVGHVDIYPNGGTFQPGCNIGEAIR VIAERGLGDVDQLVKCSHERSIHLFIDSLNNEENPSKAYRCSSKEAFEKGL CLSCRKNRCNNLGYEINKVRAKRSSKMYLKTRSQMPYKVFHYQVKIHFS GTESEHTNQAFEISLYGTVAESENIPFTLPEVSTNKTY SFLIYTEVDIGELL MLKWKWSDSYFSDWSSPGFAIQKIRVKAGETQKKVIFCSREK VSH LQKGKAPAVFVKCHDKSLNKKSG	LPL
146	MRPVRLMKVVFVTRRIPAEGRVALARAADCEVEQWDSDEPIPAKELERGV AGAHGLLCLLSDHVDKRILDAAGANLKVISTMSVGIDHLALDEIKKRGIR VGYTPDVLTDTTAELAVSLLTTTCRRLPEAIEEVKNGGWTSWKPLWLCG YGLTQSTVGIIGLGRIGQAIARRLKPFGVQRFLYTGRQRPPEEAAEFQAEF VSTPELAAQSDFIVVACSLTPATEGLCNKDFQKMKETA VFINISRGDVVN QDDLYQALASGKIAAAGLDVTSPEPLPTNHPLLTLKNCVILPHIGSATHRT RNTMSLLAANNLLAGLRGEPMPSELKL	GRHPR
147	MLGPQVWSSVRQGLSRSLSRNVGVWASGEGKKVDIAGIYPPVTTPTAT AEVDYGKLEENLHKLGTFFRFGFVVQGSNGEFPFLTSSERLEVVSRRVQA MPKNRLLLAGSGCESTQATVEMTVSMAQVGADAAMVVTPCYRGRMS SAALIHHTYTKVADLSPIPVVLYSVPANTGLDLPVDAVVTL SQHPNIVGMK DSGGDVTRIGLIVHKTRKQDFQVLASAGFLMASYALGAVGGVCALANV LGAQVCQLERLCCTGQWEDAQKLQHRLIEPNAAVTRRFGIPGLKKIMDW FGYYGGPCRAPLQELSPAEEEEALRMDFTSNGWL	HOGA1

10

20

30

40

148	MGPWGWKLRWTVALLLAAAGTAVGDR CERNEFQCQDGK CISYKWVCD GSAECQDGSEDESQETCLSVTCKSGDFSCGGRVNR CIPQFWRC DGQVDCD NGSDEQGCPPKTC SQDEF RCHDGK CISRQFVCDSDRDCLDGSDEASCPVL TCGPASFQCNSSTCIPQLWACDNDPDCEDGSDEWPQR CRGLYVFQGDSSP CSAFEFHCLSGECIHSSWRCDGGPDCKDKSDEENCAVATCRPDEFQCS DG NCIHGSRQC DREYDCKDMSDEVGC VNVTLCEGPNKFKCHSGECITLDKV CNMARDCRDWSDEPIKECGTNECLDNNGGC SHVCNDLKIGYECLCPDGF QLVAQR CEDIDECQDPDTCSQLCVNLEGGYKQCCEEGFQLDPHTKACK AVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYW SDLSQRMICSTQLDRAHGVS SYDTVISRDIQAPDGLAVDWIHSNIYWTDS VLGTVSVADTKGVKRKTLFRENGSKPRAIVDPVHGFMYWTDWGTPAKI KKGGLNGVDIYSLVTENIQWPNGITLDLLSGRLYWVDSKLHSSIDVNGG NRKTILEDEKRLAHPFSLAVFEDKVFWDIINEAIFSANRLTGSDVNLLAE NLLSPEDMVL FHNLTQPRGVNWCERTTLSNGGCQYLCLPAPQINPHSPKF TCACPDGMLLARDMRSCLTEAEAAVATQETSTVRLKVSSTAVRTQH TTT RPVPDTSRLPGATPGLT TVEIVTMSHQALGDVAGRGNEKKPSSVRALSIVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDPVYQKTTEDEVHICHNQDG YSYPSRQMVSLEDDVA	LDLR
149	MLWSGCRRFGARLGCLPGGLRVLVQTGHR S LTSCIDPSMGLNEEQKEFQKVAFDFAAREM APNMAEWDQKELFPVDVMRKA AQLGFGGVY IQTDVGGSGLSRLDTSVIFEALATGCTSTT AYISIHNMCAWMID SFGNEE QRHKFCPPLCTMEKFASYCLTEPGSGSDAA SLLTSAKKQGDHYILNGSKAFISGAGESDI YVVMCRTGGPGPKGIS CIVVEKGTPLSFG KKEKKV GWSQPTRAVIFEDCAVPVANRIG SEGQGFLIAVRGLNGGRINIASCSLGA AHA SVILTRDHLNVRKQFGEPLASNQYLQFTLADMATRLVAARLMVRNAAVA LQEERKDAVALCSMAKLFATDECFAICNQALQM HGGYGYLKDYAVQQY VRDSRVHQILEGSNEVMRILISRLLQE	ACAD8
150	MEGLAVRLLRGSRLRRNFLTCLSSWKIPPHVSKSSQSEALLNITNNGIHF APLQTTFTDEEMMIKSSVKKFAQE QIAPLVSTMDENSKMEKSVIQGLFQQG LMGIEVDPEYGGTGASFLSTVLVIEELAKVDASVAVFCEIQNTLINTLIRKH GTEEQKATYLPQLTTEKVGSFCLSEAGAGSDSFALKTRADKEGDYYVLN GSKMWISSAEHAGLFLV MANVDPTIGYKGITSFLVDRDTPGLHIGKPENK LGLRASSTCPLTFENVKVPEANILGQIGHGYKYAIGSLNEGRIGIAAQMLG LAQGCFDYTIPYIKERIQFGKRLFD FQGLQHQAHVATQLEAARLLTYNA ARLLEAGKPFIKEASMAKY YASEIAGQTTSKCIEWMGGVGYTKDYPVEK YFRDAKIGTIYEGASNIQLNTIAKHIDA EY	ACADSB

10

20

30

40

50

151	MAVLAALLRSGARSRSPLLRRLVQEIRYVERSIVSKPTLKEVVIVSATRTP IGSFLGSLSELLPATKLGSAIQGAIEKAGIPKEEVKEAYMGNVLQGGEGQA PTRQAVLQAGLPSTPCTTINKVCASGMKAIMMASQSLMCGHQDVMVAG GMESMSNVPPYVMNRGSTPYGGVKLEDLIVKDGLTDVYNKIHMGSCAEN TAKKLNARNEQDAYAINSYTRSKAAWEAGKFGNEVIPVTVTVKGPQDV VVKEDDEEYKRVDVFSKVPKLTQVFKENGTVTAANASTLNDGAAALVLM TADAAKRLNVTPLARIVAFADAAVEPIDFPIAPVYAASMLVKDVGLKKED IAMWEVNEAFSLVVLANIKMLEIDPQKVNINGGAVSLGHPIGMSGARIVG HLTHALKQGEYGLASICNNGGGGASAMLIQKL	ACAT1
152	MLPHVVLTFRRLGCALASCRLAPARHRGSGLLHTAPVARSDRSAPVFTRA LAFGDRIALDQHGRHTYRELYSRSLRSLQIEICRLCGCVGGDLREERVSLFC ANDASYVVAQWASWMSGGVAVPLYRKHPAAQLEYVICDSQSSVVLASQ EYLELLSPVVRKLGVPPLPLTPAIYTGAVEEPAEVPVPEQGWRNKGAMIIY TSGTTGRPKGVLSTHQNIRAVVTGLVHKWAWTKDDVILHVLPLHHVHG VVNALLCPLWVGATCVMMPFEFSPQQVWEKFLSSETPRINVFMAVPTIYT KLMEYYDRHFTQPHAQDFLRVCEEKIRLMVSGSAAALPLPVLEKWKNT GHTLLERYGMTEIGMALSGPLTTAVRLPGSVGTPLPGVQVRIVSENQRE ACSYTIHAEGDERGKVTQPGFEEKEGELLVRGPSVFREYWNKPEETKSAF TLDGWFKTGDTVVFQKQYWIRGRTSVDIIKTGGYKVSALVEVHLLAH PSITDVAVIGVPDMTWGQRVTAVVTLREGHSLSHRELKEWARNVLAPYA VPSELVVEEIPRNQMGKIDKKALIRHFHPS	ACSF3
153	MTSCHIAEEHIQKVAIFGGTHGNELTGVLVVKHWLENGAEIQRGTGLEVKP FITNPRAVKKCTRYIDCDLNRIFDLENLGGKMSEDLPYEVRRAQEINHLFG PKDSEDSYDIIFDLHNTTSNMGCTLILEDSRNNFLIQMFHYIKTSLAPLPCY VYLIEHPSLKYATTRSIAKYVPGIEVGPQPGVLRADILDQMRKMIKHAL DFIHNFNEGKEFPFCAIEVYKIIKVDYPRDENGEIA AIIHPNLQDQDWKPLHPGDPMFLTLDGKTIPLGGDCTVYPVVFVNEAAYYE KKEAFAKTTKLTNAKSIRCCLH	ASPA
154	MAAAVAAAPGALGSLHAGGARLVAACSAWLCPLRPLGSLAGRRAGPA IWAQGWVPAAGGPAPKRGYSSEMKTEDELVRHLEENRGIVVLGINRA YGKNSLSKNLIKMLSKAVDALKSDKKVTRTHIRSEVPGIFCAGADLKERAK MSSSEVGPVFSKIRAVINDIANLPVPTIAAIDGLALGGGLELALACDIRVAA SSAKMGLVETKLAIPGGGGTQRLPRAIGMSLAKELIFSARVLDGKEAKA VGLISHVLEQNQEGDAAYRKALDLAREFLPQGPVAMRVAKLAINQGMEV DLVTGLAIEEACYAQTIPTKDRLEGLLAFKEKRPPRYKGE	AUH
155	MASTVVAVGLTIAAAGFAGRYVLQAMKHMEPQVKQVFQSLPKSAFSGG YYRGGFEPKMTKREAALILGVSPTANKGKIRDAHRRIMLLNHPDKGGSPY IAAKINEAKDLLEGQAKK	DNAJC19

10

20

30

40

50



156	MAEAVLRVARRQLSQRGGSGAPILLRQMFEPVSCTFTYLLGDRESREAVL IDPVLETAPRDAQLIKELGLRLLYAVNTHCHADHITGSGLLRSL LPGCQSV ISRLSGAQADLHIEDGDSIRFGRFALETRASPGHTPGCVTFVLNDHSMAFT GDALLIRGCGRTDFQQCAKTLYHSVHEKIFTLPGDCLIYPAHDYHGFTV STVEEERTLNPRLTLSCEEFVKIMGNLNLPKPQQIDFAVPANMRCGVQTP A	ETHE1
157	MADQAPFDTDVNTLTRFVMEEGRKARGTGELTQLLNSLCTAVKAISSAV RKAGIAHLYGIAGSTNVTGDQVKKLDVLSNDLVMNMLKSSFATCVLVSE EDKHAIIVEPEKRGKYVVCDFPLDGSSNIDCLVSVGTIFGIYRKKSTDEPSE KDALQPGRNLVAAGYALYGSATMLVLAMDCGVNCFMLDPAIGEFILVD KDVKIKKKGKIYSLNEGYARDFDPAVTEYIQRKKFPPDNSAPYGARYVGS MVADVHRTL VYGGIFLYPANKKSPNGKLRLLYECNPMAYVMEKAGGM ATTGKEAVLDVIPTDIHQRAPVILGSPDDVLEFLKVYEKHSAQ	FBP1
158	MSQLVECVPNFSEGNQEVIDAISGAIQTGCVLLD D V D A G P S T N R T V Y T F VGPPECVVEGALNAARVASRLIDMSRHQGEHPRMGALDVC PFIPVRGVS VDECVLCAQAFGQRLAEELDVPVYLYGEAARMDSRRTLPAIRAGEYEAL PKKLQQADWAPDFGPSSFVPSWGATATGARKFLIAFNINLLGTKEQAHRI ALNREQGRGKDQPGRLKKVQGIGWYLDEKNLAQVSTNLLDFEVTALHT VYEETCREAQELSLPVVGSQLVGLVPLKALLDAAAFYCEKENLFILEEEQ RI RLVVSRLGLDSLCPFSPKERIIEYLVPERGPERGLGSKSLRAFVGEVGARSA APGGGSVAAAAAAMGAALGSMVGLMTYGRRQFQSLD T T M R R L I P P F R E ASAKLTTLVDADAEFTAYLEAMRLPKNTPEEKDRRTAALQEGLRRAVS VPLTLAETVASLWPALQELARCGNLACRSDLQVAAKALEMGVFGAYFN VLINLRDITDEAFKDQIHHRVSSLLQEAKTQAALVLDCLETRQE	FTCD
159	MATNWGSL L Q D K Q Q L E E L A R Q A V D R A L A E G V L L R T S Q E P T S S E V V S Y A P FTLFPSLVPSALLEQAYAVQMDFNLLVDAVSQNAAFLEQTLSSTIKQDDF TARLFDIHKQVLKEGIAQTVFLGLNRS D Y M F Q R S A D G S P A L K Q I E I N T I S A S FGGLASRTPAVHRHVL SVL SK TKEAGKILSNNPSKGLALGIAKAWELYGS PNALVLLIAQEKERNIFDQRAIENELLARNIHVIRRTFEDISEKGLDQDRR LFVDGQEI AVVYFRDGYM PRQYSLQNWEARLLERSHAAKCPDIATQLA GTKKVQQELSRPGMLEMLLPGQPEAVARLRATFAGLYSLDVGEEGDQAI AEALAAPSRFV L K P Q R E G G G N N L Y G E E M V Q A L K Q L K D S E E R A S Y I L M E K IEPEPFENCLLRPGSPARVVQCISELGIFGVYVRQEKTLMNKHVGHLLRT KAIEHADGGVAAGVAVLDPYPV	GSS

10

20

30

40

50

160	<p>MGQREMWRLMSRFNAFKRTNTILHHLRMSKHTDAAEEVLEKKGCTGV  ITLNRPKFLNALTNLNMIHQIYPQLKKWEQDPETFLIIIKAGGKAFCAGGDI  RVISEAEKAKQKIAPVFFREEYMLNNAVGCQKPYVALIHGITMGGGVGL  SVHGQFRVATEKCLFAMPETAIGLFPDVGGYFLPRLQGKLGFLALTGF  RLKGRDVYRAGIATHFVDSEKLAMLEEDLLALKSPSKENIASVLENYHTE  SKIDRDKSFILEEHMDKINSCFSANTVEEIIENLQQDGSSFALEQLKVINKM  SPTSLKITLRQLMEGSSKTLQEVLTMEYRLSQACMRGHDFHEGVRAVLID  KDQSPKWKPADLKEVTEEDLNNHFKSLGSSDLKF</p>	HIBCH
161	<p>MAGYLRVVRSLCRASGSRPAWAPAALTAPTSQEQPRRHYADKRIKVAKP  VVEMDGDDEMTRIIWQFIKEKLILPHVDIQLKYFDLGLPNRDQTDQVITIDS  ALATQKYSVAVKCATITPDEARVEEFKLLKMWKSPNGTIRNILGGTVFRE  PIICKNIPRLVPGWTKPITIGRHAHGDQYKATDFVADRAGTFKMVFTPDK  GSGVKEWEVYNFPAGGVGMGMYNNTDESISGFAHSCFYAIQKKWPLYM  STKNITLKAYDGRFKDIFQEIFDKHYKTDFDKNKIWIYEHRLIDDMVAQVL  KSSGGVFWACKNYDGDVQSDILAQGFGLGLMTSVLVCPDGKTIEAEEA  HGTVTRHYREHQGRPTSTNPIASIFAWTRGLEHRGKLDGNQDLIRFAQM  LEKVCVETVESGAMTKDLAGCIHGLSNVKNLNEHFLNTDFLDTIKSNLDR  ALGRQ</p>	IDH2
162	<p>MVPALRYLVGACGRARGLFAGGSPGACGFASGRPRPLCGGSRASSTSSFD  IVIVGGGIVGLASARALILRHPSLSIGVLEKEKDLAVHQTGHNSGVIHSGIY  YKPESLKAKLCVQGAALLYEYCQKQKISYKQCGKLIVAVEQEEIPRLQAL  YEKGLQNGVPLRLIQQEDIKKKEPYCRGLMAIDCPHTGIVDYRQVALSF  AQDFQEAGGSVLTNFEVKGIEMAKESPSRSIDGMQYPIVIKNTKGEEIRCQ  YVVTCAGLYSDRISELSGCTPDPRIVPFRGDYLLLKPEKCYLVKGNIYPVP  DSRFPFLGVHFTPRMDGSIWLGPNVAVLAFKREGYRPFDFSATDVMDDIINS  GLIKLASQNFYSYGVTEMYKACFLGATVKYLQKFIPEITISDILRGPAGVRA  QALDRDGNLVEDFVFDAGVGDIGNRILHVRNAPSPAATSSIAISGMIADDEV  QQRFEL</p>	L2HGDH
163	<p>MRGFGPGLTARRLLPLRLPPRPPGPRLASGQAAGALERAMDELLRRAVPP  TPAYELREKTPAPAEGQCADFVSFYGGLAETAQRAELLGRLARGFGVDH  GQVAEQSAGVLHLRQQQREAAVLLQAEDRLRYALVPRYRGLFHHISKLD  GGVRFLVQLRADLLEAQALKLVEGPDVREMNGVLKGMLSEWFSSGFLN  LERVTWHSPEVLQKISEAEAVHPVKNWMDMKRRVGPYRRCYFFSHCST  PGEPLVVLHVALTGDISSNIQAIVKEHPPSETTEKNKITAIFYSISLTQQGL  QG  VELGTFLIKRVVKELQREFPHLGVFSSLSPIPGFTKWLLGLLNSQTKEHGR  NELFTDSECKEISEITGGPINETLKLSSSEWVQSEKLVRALQTPLMRLCA  WYLYGEKHRGYALNPVANFHLQNGAVLWRINWMADVSLRGITGSCGL  MANYRYFLEETGPNSTSYLGSKIHKASEQVLSLVAQFQKNSKL</p>	MLYCD

10

20

30

40

50

164	<p>MVVGAFPMAKLLYLGRVSKPLANRIKEAARRSEFFKTYICLPPAQLYH  WVEMRTKMRIMGFRGTVIKPLNEEAAAELGAELLGEATIFIVGGGCLVLE  YWRHQAQQRHKEEEQRAAWNLRDEVGHLALALEALQAQVQAAPPQG  ALEELRTELQEVRAQLCNPGRSASHAVPASKK</p>	OPA3
165	<p>MGSPEGRFHFHAIDRGGTFTDVFACPGGHVVRVLKLLSEDPANYADAPTEG  IRRILEQEAGMLLPRDQPLDSSHASIRMGTTVATNALLERKGERVALLVT  RGRDLLHIGTQARGDLFDLAVPMPEVLYEEVLEVDERVVLHRGEAGTG  TPVKGRTGDLLEVQQPVLDLALRGKLEGLLSRSLAVVLMHSYTWAG  HEQQVGVLAARELGFTHVLSSEAMPVVRIVPRGHTACADAYLTPAIQRY  VQGFRCRQFQGLKDVQVLFMRSDGGLAPMDTFSGSSAVLSGPAGGVV  YSATTYQQEGGQPVIGFDMGGTSTDVSRVAGEFEHVFEASTAGVTLQAP  QLDINTVAAGGSRLLFFRSGLFVVGPEASAGHPGPACYRKGGPVTVTDA  NLVLRLLPASFPICFGPGENQPLSPEASRKALEAVATEVNSFLTNGPCPAS  PLSLEEVAMGFVRVANEAMCRPIRALTQARGHDPSAHVLACFGGAGGQH  ACAIARALGMDTVHHRHSGLLSALGLALADVVEAQPESLLYAPETV  QLDQRLSRLEEQCVDALQAQGFPRSQISTESFLHLRYQGTDCALMVSAHQ  HPATA  RSPRAGDFGAAFVERYMREFGFVIPERPVVDDVVRVGTGRSGLRLEDAP  KAQTGPPRVDKMTQCYFEGGYQETPVYLLAELGYGHKLHGPCIIDSNST  ILVEPGCAEVTKTGDICISVGAEPGTVPQDPIQLSIFSHRFMSIAEQM  GRILQRTAISTNIKERLDFSCALFGPDGGLVSNAPHIPVHLGAMQETVQFQI  QHLGADLHPGDVLLSNHPSAGGSHLPDLTVITPVFWPGQTRPVFYVASRG  HHADIGGITPGSMPPHSTMLQQEGAVFLSFKLVQGGVFQEEAVTEALRAP  GKVPNCSGTRNLHDNLSDLRAQVAANQKGIQLVGELIGQYGLDVVQAY  MGHIQANAELAVRDMLEAFGTSRQARGLPLEVSSDHMDGSPIRLRVQI  SLSQGSVDFDGTGPEVFGNLNAPRAVTLALYCLRCLVGRDIPLNQGC  LAPVRVVIIPRGSILDPSPEAAVGGNVLTSQRVVDVILGAFGACAASQGC  MNNVTLGNAHMGYYETVAGGAGAGPSWHGRSGVHSHMTNTRITDPEIL  ESRYPVILRRFELRRGSGRGRFRGGDGTRELLFREEALLSVLTERRAFR  PYGLHGGEPGARGLNLLIRKNGRTVNLGGKTSVTVYPGDVFLHTPGGG  GYGDPEDPAPPPGSPQALAFPEHGSVYERYRAQEAV</p>	OPLAH
166	<p>MAALKLLSSGLRLCASARGSGATWYKGCVCSFSTSAHRHTKFYTDVPEA  VKDIPDGATVVLVGGFGLCGIPENLIDALLKTGVKGLTAVSNNAGVDNFG  GLLRSKQIKRMVSSYVGENAEFERQYLSGELEVELTPQGTLAERIRAGG  AGVPAFYTPTGYGTLVQEGGSPIKYNKDGSVAIASKPREVREFNGQHFILE  EAITGDFALVKAWKADRAGNVIFRKSARNFNLMCKAAETTVVEVEEIV  DIGAFAPEDIHIPQIYVHRLIKGEKYEKRIERLSIRKEGDGEAKSAKPGDDV  RERIKRAALEFEDGMYANLIGIPLASNFIENITVHLQSENGVLGLGPY  PRQHEADADLINAGKETVTLPGASFFSSDESAMIRGGHVDLTMLGAMQ  VSKYGLANWMIKGMVKGMMGAMDVSSAKTKVVVTMEHSAKGNA  HKIMEKCTLPLTGKQCVNRIITEKAVFDVDKGLTLIELWEGLTVDDVQ  KSTGCDFAVSPKLMMPMQIAN</p>	OXCT1

10

20

30

40

50

167	MSRLLWRKVAGATVGPVPAPGRWVSSVSPASDPSDGQRRRQQQQQQ QQQQQQQPQPQVLSSEGGQLRHNPLDIQMLSRGLHEQIFGQGGEMPGE AAVRRSVEHLQKHGLWGQPAVPLPDVELRLPPLYGDNLDQHFRLLAQKQ SLPYLEAANLLQAQLPPKPPAWAWAEGWTRYGPEGEAVPVAIPEERAL VFDVEVCLAEGTCPTLAVAISSAWYSWCSQRLVEERYSWTSQLSPADLI PLEVPTGASSPTQRDWQEQLVVGHNVSFDRAHIREQYLIQGSRMRFLLDTM SMHMAISGLSSFQSLWIAAKQKGKHKVQPPTKQGQKSQRKARRGPAISS WDWLDISSVNSLAEVHRLYVGGPPLEKEPREL FVKGTMKDIRENFQDLM QYCAQDVWATHEVFQQQLPLFLERCPHPVTLAGMLEMGVSYLPVNQNW ERYLAEAQGTYEELQREMKKSMDLANDACQLLSGERYKEDPWLWdle WDLQEFKQKKAKKVKKEPATASKLPIEGAGAPGDPMDQEDLGPCSEEEE FQQDVMARACLQKLKGTTELLPKRPQHLPGHGWYRKLCPRLDDPAWT PGPSLLSLQMRVTPKLMALTWDGFPLHYSERHGWGYLVPGRRDNLAKLP TGTTLESAGVVCPYRAIESLYRKHCLQGGKQQLMPQEAGLAEFLLTDNS AIWQTVVEELDYLEVEAEAKMENLRAAVPGQPLALTARGGPKDTQPSYHH GNGPYNDVDIPGCWFFKLPHKDGNSCNVGSFPAKDFLPKMEDGTLQAGP GGASGPRALEINKMISFWRNAHKRISSQMVVWLPRALPRAVIRHPDYDE EGLYGAILPQVVTAGTITRAVEPTWLTASNARPDRVGSELKAMVQAPPG YTLVGADVDSQELWIAAVLGDAAHFAGMHGCTAFGWMTLQGRKSRGTD LHSKTATTVGISREHAKIFNYGRIYGAGQPFAERLLMQFNHRLTQQEAAE KAQQMYAATKGLRWYRLSDEGEWL VRELNLPVDRTEGGWISLQDLRKY QRETARKSQWKKWEVVAERAWKGGTESEM FNKLESIATS DIPRTPVLGC CISRALEPSAVQEEFMTSRVNWVQSSAVDYLHMLLVAMKWLFEFFAID GRFCISIHDEVRYLVREEDRYRAALALQITNLLTRCMFAYKLGNDLPQS VAFFSAVDIDRCLRKEVTMDCKTPSNPTGMERRYGIPQGEALDIYQIIELT KGSLEKRSQPGP	POLG
168	MSTAALITLVRSGGNQVRRRVLLSSRLLQDDRRVTPTCHSSTSEPRCSRFD PDGSGSPATWDFNGIWDNRIDEPIPPSIKYGKPIPKISLENVGCASQIGKR KENEDRFDAQLTDEVLYFAVYDGHGGPAAADFCPTHMEKCIMDLLPKE KNLETLLTLAFLEIDKAFSSHARLSADATLLTSGTTATVALLRDGIELVVA SVGDSRAILCRKKGKPMKLTIDHTPERKDEKERIKKCGGFVAWNSLGQPHV NGRLAMTRSIGDLDLKTSGVIAEPETKRIKLHHADDSFLVLT TDGINFMVN SQEICDFVNQCHDPNEAAHAVTEQAIQYGTEDNSTAVVVPFGAWGKYKN SEINFSFSRSFASSGRWA	PPMIK

10

20

30

40

50

169	MSLAAYCVICRRIGTSTSPKSGTHWRDIRNIIKFTGSLILGGSLFLTYEVL ALKKAVTLDTQVVEREKMKSYYVHTVSLDKGENHGIAWQARKELHKA VRKVLATSakilRNPFADPFSTVDIEDHECAVWLLLRKSKSDDKTTRLEA VREMSETHHWHDYQYRIIAQACDPKTLIGLARSEESDLRFFLLPPPLPSLK EDSSTEEELRQLLASLPQTELDECIQYFTSLALSESSQ SLAAQKGGWLWCFGGNGLPYAESFGEVPSATVEMFCLEAIVKHSEISTHCD KIEANGGLQLLQRLYRLHKDCPKVQRNIMRVIGNMALNEHLHSSIVRS WVSIMAEAMKSPHIMESSHAARILANLDRETVQEKYQDGVVYLHPQYRT SQPIKADVLFHIGLMGAAFKTWRQQDSEQAVIEKPMEDEDRYTTCWPKT WLAKDCPALRIISVEYDTSLSDWRARCPMERKSIAFRSNELLRKLRAAGV GDRPVVWISHSMGGLLVKKMLLEASTKPEMSTVINNTRGIIFYVPHHGS RLAEYSVNIRYLLFPSLEVKELSKDSPALKTLQDDFLEFAKDNFQVLNLFV ETLPTYIGSMIKLHVVPVESADLGIGDLIPVDVNHNLNICKPKKKDAFLYQR TLQFIREALAKDLEN	SERAC1
170	MPAPRAPRALAAAAPASGKAKLTHPGKAILAGGLAGGIEICITFPTEYVKT QLQLDERSHPPRYRGIGDCVRQTVRSHGVLGLYRGLSSLLYGSIPKAAVR FGMFEFLSNHMRDAQGRLDSTRGLLCLGAGVAEAVVVVCPMETIKVKF IHDQTSNPKNKYRGGFFHGVREIVREQGLKGTQGLTATVLKQGSNQAIRFF VMTSLRNWYRGDNPNKPMNPLITGVFGAIAAGASVFGNTPLDVIKTRMQ GLEAHKYRNTWDCGLQILKKEGLKAFYKGTVPRLGRVCLDVAIVFVIYD EV VKLLNKVWKT	SLC25A1
171	MAASMFYGRVLAVATLRNHRPRTAQRAAAQVLGSSGLFNNHGLQVQQQ QQRNLSLHEYMSMELLQEAGVSVPKGYVAKSPDEAYAIKKLGSKDVVI KAQVLAGGRGKGTFFESGLKGGVKIVFSPEEAKAVSSQMIGKKLFTKQTGE KGRICNQVLVCERKYPRREYYFAITMERSFQGPVLIGSSHGGVNIEDVAAE SPEAIKEPIDIEEGIKKEQALQLAQKMGFPNIVESAAENMVKLYSLFLKY DATMIEINPMVEDSDGAVLCMDAKINFDSNSAYRQKKIFDLQDWTQEDE RDKDAAKANLNYIGLDGNIGCLVNGAGLAMATMDIIKLHGGTPANFLDV GGGATVHQVTEAFKLITSDKKVLAILVNIFGGIMRCDVIAQGIVMAVKDL EIKIPVVVRLQGTRVDDAKALIADSGLKILACDDLDEAARMVVKLSEIVTL AKQAHVDVKFQLPI	SUCLA2
172	MTATLAAAADIATMVSGSSGLAAARLLRSFLLPQNGIRHCSYTASRQHL YVDKNTKIICQGFTGKQGTFFHSQQALEYGTCLVGGTTPGKGGQTHLGLP VFNTVKEAKEQTGATASVIYVPPPFAAAINEAIEAEIPLVVCITEGIPQQD MVRVKHKLLRQEKTRLIGPNCPGVINPGECKIGIMPGHIHKKGRIGIVSRS GTLTYEAVHQTTQVGLGQSLCVGIGGDPFNGTDFIDCLEIFLNDSETEGIIIL IGEIGGNAEENAAEFLKQHNSGPNSKPVVSFIAGLTAPPGRRMGHAGAIIA GGKGGAKEKISALQSAGVVVSMSPAQLGTTIYKEFEKRKML	SUCLG1

10

20

30

40

50

173	MPLHVKWPFPAVPPLTWTLASSVVMGLVGTYSFCFWTKYMNHLTVHNRE VLYELIEKRGPATPLITVSNHQSCMDDPHLWGILKLRHIWNLKLMRWTPA AADICFTKELHSHFFSLGKCVPCRGAEFFQAENEGKGVLDTRHMPGA GKRREKGDGVYQKGMDFILEKLNHGDWVHIFPEGKVNMSSEFLRFKWI GRLIAECHLNPIILPLWHVGMNDVLPNSPPYFPRFGQKITVLIGKPFSALPV LERLRAENKSAVEMRKALTDIFIQEEFQHLKTQAEQLHNHLQPGR	TAZ
174	MTVFFKTLRNHWKKTATAGLCLLTWGGHWLYGKHCDNLLRRAACQEAQ VFGNQLIPPNAQVKKATVFLNPAACKGKARTLFEKNAAPILHLSGMDVTI VKTDYEGQAKKLELEMENTDVIIVAGGDGTLQEVVTGVLRRRTDEATFSKI PIGFIPLGETSSLSHTLFAESGNKVQHITDATLAIVKGETVPLDVLQIKGEK EQPVFAMTGLRWGSRDAGVKVSKYWYLGPLKIKAAHFFSTLKEWPQT HQASISYTGPTERPPNEPEETPVQRPSLYRRILRRLASYWAQPQDALSQEV SPEVWKDVQLSTIELSITTRNNQLDPTSKEDFLNICIEPDTISKGDFITIGSRK VRNPKLHVEGTECLQASQCTLLIPEGAGGSFSIDSEEYEAMPVEVKLLPRK LQFFCDPRKREQMLTSPTQ	AGK
175	MLGSLVLRKALAPRLLRLLRSPTLRGHGGASGRNVTTGSLGEPQWLR VATGGRPGTSPALFSGRGAATGGRQGGFRDTKCLAAATWGRLPGPEETL PGQDSWNGVPSRAGLGMCALAAALVVHCYKSPSNKDAALLEAARANN MQEVSRLLESEGADVNAKHRLGWTALMVAAINRNNSSVVQVLLAAGADPN LGDDFSSVYKTAKEQGIHSLEDGGQDGASRHITNQWTSALEFRRWLGLP AGVLITREDDFNRLNRRASFKGCTALHYAVLADDYRTVKELLDGGANP LQRNEMGHTPLDYAREGEVMKLLRTSEAKYQEKQRKREAEERRRFPLEQ RLKEHIIGQESAIATVGAA IRRKENGWYDEEHPLVFLFLGSSGIGKTELAKQTAKYMHKDAKKGFIRD MSEFQERHEVAKFIGSPPGYVGHEEGQLTKKLKQCPNAVVLDFEVDKA HPDVLTIMLQLFDEGRLTDGKGTIDCKDAIFIMTSNVASDEIAQHALQLR QEALEMSRNRIAENLGDVQISDKITISKNFKENVIRPILKAHFRRDEFLGRI NEIVYFLPFCHSELIQLVNKELNFWAKRAKQRHNITLLWDREVADVLVDG YNVHYGARSIKHEVERRVVNQLAAAYEQDLLPGGCTLRITVEDSDKQLL KSPPELSPQAEKRLPKLRLEIIDKDSKTRRLDIRAPLHPEKVCNTI	CLPB
176	MLFLALGSPWAVELPLCGRRTALCAAALRGPRASVSRASSSSGSPGPVA GWSTGSPGAARLLRRPGRAQIPVYWEGYVRFLNTPSDKSEDGRLIYTG MARAVFGVKCFSYSTSLIGLTFLPYIFTQNNAISESVPPIQIIFYGIMGSFTV ITPVLLHFITKGYVIRLYHEATTDYKAITYNAMLAETSTVFHQNDVKIPD AKHVFTTFYAKTKSLLVNPVLPNREDYIHLMGYDKEEFILYMEETSEEK RHKDDK	TMEM70

10

20

30

40

50

177	MLSQVYRCGFQPFNQHLLPWVKCTTVFRSHCIQPSVIRHVRSWSNIPFITV PLSRTHGKSFAHRSELKHAKRIVVKLGSAVVTRGDECGLALGRLASIVEQ VSVLQNGQREMMLVTSGAVAFGKQRLRHEILLSQSVRQALHSGQNQLKE MAIPVLEARACAAAGQSGMLALYEAMFTQYSICAAQILVTNLDHFHDEQK RRNLNGTLHELLRMNIVPIVNTNDAAVPPAEPNSDLQGVNVISVKDNDL AARLAVEMKTDLLIVLSDVEGLFDSPPGSDDAKLIDIFYPGDQQSVTFGTK SRVGMGGMEAKVKAALWALQGGTSVVIANGTHPKVSGHVITDIVEGKK VGTFFSEVKPAGPTVEQQGEMARSGGRMLATLEPEQRAEIIHHLADLLTD QRDEILLANKKDLEEAEGRLAAPLLKRLSLSTSKLNSLAIGLRQIAASSQD SVGRVLRRTRIAKNLELEQVTVPIGVLLVIFESRPDCLPQVAALAIASNGG LLLKGGKEAAHSNRILHLLTQEALSIHGVKEAVQLVNTREEVEDLCRLDK MIDLIIPRGSSQLVRDIQKAAKGIPVMGHSEGICHMYVDSEASVDKVRTLV RDSKCEYPAACNALETLLIHRDLLRTPFDQIIDMLRVEQVKIHAGPKFAS YLTFSPSEVKSLRTEYGDLELCIEVVDNVQDAIDHIIHKYGSSTDVIVTED ENTAEFFLQHVDSACVFWNASTRFSQGYRFGGLGAEVGISTSRIHARGPVG LEGLLTTKWLLRGKDHVVSDFSEHGSLKYLHENLPIQRNTN	ALDH18A 1
178	MFSKLAHLQRFAVLRSRVHSSVASATSVATKKTQGPPTSDDIFEREYKY GAHNYHPLPVALERGGKIYLWVVEGRKYFDLSSYSAVNQGHCHPKIVN ALKSQVDKLTLSRAFYNVNLGEYEEYITKLFNYHKVLPMTGVEAGET ACKLARKWGYTVKGIQKYKAKIVFAAGNFWGRTLSAISSTDPSTSYDGF GPFMPGFDIIPYNDLPALERALQDPNVAAFMVEPIQGEAGVVVPDPGYLM GVRELCTRHQVLFIADEIQTGLARTGRWLAVDYENVRPDIPLLKALSGG LYPVSAVLCDDDIMLTIKPGEHGSTYGGNPLGCRVAIAALEVLEENLAE NADKLGILRNELMKLPSDVVTAVRGKGLLNAIVIKETKDWDWVCLR LRDNGLLAKPTHGDIIRFAPPLVIKEDDELRESIEIINKTILSF	OAT
179	MLGRNTWKTSAFSFLVEQMWAPLWSRSMRPGRWCSQRSCAWQTSNNT LHPLWTVPVSVPGGTRQSPINIQWRDSVYDPQLKPLRVSYEAASCLYIWN TGYLQVFEFDDATEASGISGGPLENHRYLQKQFHFHWGAVNEGGSEHTVD GHAYPAELHLVHWNSVKYQNYKEAVVGENGLAVIGVFLKLGAAHQTLQ RLVDILPEIKHKDARAAMRPFDPSTLLPTCWDYWTYAGSLTTPPLTESVT WIIQKEPVEVAPSQLSAFRLLFSALGEEKMMVNRYRPLQPLMNRKVV ASFQATNEGTRS	CA5A

10

20

30

40

50

180	MYRYLGEALLLSRAGPAALGSASADSAALLGWARGQPAAAPQPGLALA ARRHYSEA VADREDDPNFFKMVEGFFDRGASIVEDKLVEDLRTRESEEQK RNRVRGILRIIKPCNHVLSLSPIRRDDGSWEVIEGYRAQHSQHRTPCGGI RYSTDVSVDEVKALASLMTYKCAVVDVPPFGGAKAGVKINPKNYTDNEL EKITRRFTMELAKKGFIPGIDVPAPDMSTGEREMSWIADTYASTIGHYDI NAHACVTGKPISQGGIHRISATGRGVFHGIENFINEASYMSILGMTPGFG DKTFVVQGFQGNVGLHSMRYLHRFGAKCIAVGESDGSIWNPDGIDPKELE DFKLQHGSIKPKAKPYEGSILEADCILIPAASEKQLTKSNAPRVKAKII AEGANGPTTPEADKIFLERNIMVIPDLYLNAGGVTVSYFEWLKLNHNHVS Y GRLTFKYERDSNYHLLMSVQESLERKFGKHGGTIPIVPTAEFQDRISGASE KDIVHSGLAYTMERSARQIMRTAMKYNLGLDLRTAAYVNAIEKVFVYN EAGVTFT	GLUD1
181	MTTSASSHLNKGKQVYMSLPQGEKVQAMYIWIWIDGTGEGLRCKTRTLDS EPKCVEELPEWNFDGSSTLQSEGSNSDMYLVAAMFRDPFRKDPNKLVL CEVFKNRRPAETNLRHTCKRIMDMVSNQHPWFGMEQEYTLMGTDGHP FGWPSNGFPQPQPYCGVGADRAYGRDIVEAHYRACLYAGVKIAGTN AEVMPAQWEFQIGPEGISMGDHLWVARFILHRVCEDFGVIAFDPKPIPG NWNAGAGCHTNFSTKAMREENGLKYIEEAIEKLSKRHQYHIRAYDPKGGL DNARRLTGFHETSININDFSAGVANRSASIRIPRTVGQEKKGYFEDRRPSAN CDPFSVTEALIRTCLLNETGDEPFQYKN	GLUL
182	MAVARAALGPLVTGLYDVQAFKFGDFVLKSGLSPIYIDLRGIVSRPRLS QVADILFQTAQNAGISFDTVCGVPYALPLATVICSTNQIPMLIRRKETKD YGTKRLVEGTINPGETCLIEDVVTSGSSVLETVEVLQKEGLKVTDIVLL DREQGKDKLQAHGIRLHSVCTLSKMLEILEQQKKVDAETVGRVRFIQE NVFVAANHNGSPLSIKEAPKELSFGARAEHPRIHPVA SKLLRLMQKKTENLCLADVSLARELLQLADALGPSICMLKTHVDILNDF TLDVMKELITLAKCHEFLIFEDRKFADIGNTVKKQYEGGIFKIASWADLVN AHVVPGSGVVKGLQEVGLPLHRGCLLIAEMSSTGSLATGDYTRAAVRMA EEHSEFVVGFIGSRVSMKPEFLHLTPGVQLEAGGDNLGQQYNSPQEVIG KRGSDIIVGRGIISAADRLEAAEMYRKAWEAYLSRLGV	UMPS
183	MRDYDEVTAFLGEWGPQRLIFFLLSASHIPNGFTGLSSVFLIATPEHRCRV PDAANLSSAWRNHTVPLRLRDGREVPHSCRRYRLATIANFSALEPGRD VDLGGLEQESCLDGFWEFSQDVYLSTIVTEWNLVCEDDWKAPLTISLFFVG VLLGSFISGQLSDRFRGNVLFVTMGMQTGFSFLQIFSKNFEMFVVLVFLV GMGQISNYVAAFVLGTEILGKSVRIIFSTLGVCFYAFGYMVLPLFAYFIRD WRMLLVALTMPGVLCVALWWFIPESPRWLISQGRFEEAEVIIRKAAKAN GIVVPSTIFDPSELQDLSSKKQQSHNILDLLRTWNIRMVTIMSIMLWMTISV GYFGLSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLQYLPRRYSMAT ALFLGGSVLLFMQLVPPDLYLATVLMVVGKFGVTAAFSMVYVYTAEL YPTVVRNMGVGVSSSTASRLGSILSPYFVYLGAYDRFLPYILMGSLTILTAIL TLFLPESFGTLPDITIDQMLRVKGMKHRKTPSHTR MLKDGQERPTILKSTAF	SLC22A5

10

20

30

40

50



184	MAEAHQAVAFQFTVTPDGIDLRLSHEALRQIYLSGLHSWKKKFIRFKNGII TGVYPASPSSWLIVVVGVMTTMYAKIDPSLGIIAKINRTLETANCMSSQTK NVVSGVLFGTGLWVALIVTMRYSLKVLSSYHGWMFTEHGKMSRATKIW MGMVKIFSGRKPMLYSFQTSPLRPLVPAVKDTVNRYLQSVRPLMKEEDF KRMTALAQDFAVGLGPRLQWYLKLSWWATNYVSDWWEYIYLRGRG PLMVNSNYAMDLLYILPHTIQAARAGNAIHAILLYRRKLDREEIKPIRLL GSTIPLCSAQWERMFNTRSIPGEETDTIQHMRDSKHIVVYHRGRYFKVWL YHDGRLLKPREMEQQMQRILDNTSEPQPGEARLAALTAGDRVPWARCR QAYFGRGKNKQSLDAVEKAAFFVTLDETEEGYRSEDPDTSMDSYAKSLL HGRCYDRWFDKSFTFVVFKNKMGMLNAEHSWADAPIVAHLWEYVMSID SLQLGYAEDGHCKGDINPNIPYPTLQWDIPGECQEVIENTSLNTANLLAND VDFHSFPFVAFGKGIKKCRTSPDAFVQLALQLAHYKDMGKFCLTYEASM TRLFREGRTETVRSCTTESCDFVRAMVDPAQTVQRLKLFKLASEKHQH MYRLAMTGSIDRHLFCLYVVSXYLAVESPFLKEVLSEPWRLSTSQTPOQ QVELFDLENNPEYVSSGGGFGPVADDGYGVSYILVGENLINFHISSKFSCP ETDSHRFGRHLKEAMTDIITLFLSSNSKK	CPT1A
185	MVACRAIGILSRFSAFRILRSRGIYICRNFTGSSALLTRTHINYGVKGDVAV VRINSPNSKVNLSKELHSEFSEVMNEIWASDQIRSAVLISSKPGCFIAGAD INMLAACKTLQEVTLQSQEAQRIVEKLEKSTKPIVAAINGSCLGGGLEVAI SCQYRIATKDRKTVLGTPEVLLGALPGAGGTQRLPKMVGVPAAALDMMLT GRSIRADRAKMGMLVDQLVEPLGPGLKPEERTIEYLEEVAITFAKGLAD KKISPKRDKGLVEKLTAYAMTIPFVRQQVYKKEEKVRKQTKGLYPAPL KIIDVVKTGIEQGS DAGYLCESQKFGELVMTKESKALMGLYHGQVLCCK NKFQAPQKDVKHLAILGAGLMGAGIAQVSVDKGLKTKILKDATLALDRG QQQVFKGLNDKVKKKALTSFERDSIFSNTLGTLDYQGFEEKADMVIEAVFE DLSLKHRLKEVEAVIPDHCIFASNTSALPISEIAAVSKRPEKVIKMHYFSP VDKMQLLEIITTEKTSKDTASAVAVGLKQGVIIIVK DGPFGYTTTRCLAPMMSEVIRILQEGVDPKKLDSLTTSTFGFPVGAATLVDE VGVDVAKHVAEDLGKVFGERFGGGNPELLTQMVSKGFLGRKSGKGFYIY QEGVKKDLNSDMDSILASLKLPPKSEVSSDEDIQFRLVTRFVNEAVMCL QEGILATPAEGDIGAVFGLGFPCLGGPFRFVDLYGAQKIVDRLKKEYEAA YGKQFTPCQLLADHANSNPKKFYQ	HADHA
186	MAFVTRQFMRSVSSSSTASASAKKIIVKHVTVIGGGLMGAGIAQVAAATG HTVVLVDQTEDILAKSKKGIIESLRKVAKKKFAENLKAGDEFVEKTLSTI ATSTDAASVVHSTDLVVEAIVENLKVKNELFKRLDKFAAEHTIFASNTSSL QITSIANATTRQDRFAGLHFFNPVPMKLEVIKTPMSTQKTFESLVDFSK ALGKHPVCKDTPGFIVNRLLPYLMEAIRLYERGDASKEDIDTAMKLGGA GYPMGPFELLDYVGLDITTKFIVDGDWHEMDAENPLHQSPSLNKLVAENK FGKKTGEGFYKYK	HADH

10

20

30

40

50

187	MAAPTLGRLVLTHLLVALFGMGSWAAVNGIWVELPVVVKDLPEGWSLP SYLSVVVALGNLGLLVVTLWRQLAPGKGEQVPIQVVQVLSVVG TALLAP LWHHVAPVAGQLHSVAFLTLALVLAMACCTSNVTFLPFLSHLPPFLRSF FLGQGLSALLPCVLALVQGVGRLECPPAPTNGTSGPPLDFPERFPASTFFW ALTALLVTSAAAFRGLLLLLPSLPSVTTGGSGPELQLGSPGAE EEEKEEEE ALPLQEPPSQAAGTIPGPDPEAHQLFSAHGAFLLGLMAFTSAV TNGVLP QSFSCLPYGRLAYHLAVVLGSAANPLACFLAMGVLCRSLAGLVGLSLLG MLFGAYLMALAILSPCPPLVGTTAGVVLVVLSWVLCVFSYVKVAASS LLHGGGRPALLAAGVAIQVGSLLGAGAMFPPTSIIYHVFSRKDCVDPCGP	SLC52A1	10
188	MAAPTPARPVLTHLLVALFGMGSWAAVNGIWVELPVVVKELPEGWSLPS YVSVLVALGNLGLLVVTLWRRLAPGKDEQVPIRVVQVLGMVGTALLASL WHHVAPVAGQLHSVAFLALAFVLALACCASNVTFLPFLSHLPPRFLRSFF LGQGLSALLPCVLALVQGVGRLECPPAPINGTPGPPLDFLERFPASTFFWA LTALLVASAAAFQGLLLLLPPPPSVPTGELGSGLQVGAPGAE EEEVEESSPL QEPPSQAAGTTPGDPKAYQLLSARSACLLGLLAATNAL TNGVLP SCLPYGRLAYHLAVVLGSAANPLACFLAMGVLCRSLAGLGGLSLLGVFC GGYLMALAVLSPCPPLVGTSAGVVLVVLSWVLCVFSYVKVAASSLLH GGGRPALLAAGVAIQVGSLLGAVAMFPPTSIIYHVHFSRKDCADPCDS	SLC52A2	20
189	MAFLMHLLVCVFGMGSWVTINGLWVELPLLVMELPEGWYLP SYLTVVI QLANIGPLLVTLLHFRPSCLSEVPIIFTLLGVGTVTCIIF AFLWNMTSWVL DGHHSIAFLVLTFFLALVDCTSSVTFLPFMSRLPTY YLTFFVGEGLSGLLP ALVALAQGSGLTTCVNVTEISDSVSPVPTRET DIAQGVPRALVSALPGME APLSHLESRYLPAHFSPLVFFLLS SIMMACCLVAFFV LQRQPCWEASVEDLLNDQVTLHSIRPREENDLGPAGTV DSSQGQGYLEE KAAPCCPAHLAFIYTLVAFVNAL TNGMLPSVQTY SCLSYGPVAYHLAATL SIVANPLASLVSMFLPNRSLFLGVLSVLGTCFGGYNMAMAVMSPCPLLQ GHWGGEVLIVASWVLFSGCLSYVKVMLGVVLRDL SRSALLWCGAAVQL GSLLGALLMFPLVNVLRLFSSADFCNLHCPA	SLC52A3	30
190	MTILTYPFKNLPTASKWALRFSIRPLSCSSQLRAAPAVQTKTKKTLAKPNI RNVVVVDGVRTPFLLSGTSYKDLMPHDLARAALTGL LHRTSVPKEVVDY IIFGTVIQEVKTSNVAREEALGAGFSDKTPAHTVTMACISANQAMTTGVG LIASGQCDVIVAGGVELMSDVPIRHSRKMRLMLDLNKAKSMGQRLSLIS KFRFNFLAPELPAVSEFSTSETMGHSADRLAAAF AVSRLEQDEYALRSHSL AKKAQDEGLLSDVVPFKVPGKDTVTKDNGIRPSSLEQMAKLKPAFIKPY GTVTAANSSFLTDGASAMLIMAEKALAMGYKPKAYLRDFMYVSQDPK DQLLLGPTYATPKVLEKAGLTMNDIDAFEFHEAFSGQILANFKAMDSDW FAENYMRKTKVGLPPELKFNNWGGSLSLGHPFGATGCRLVMAAANRL RKEGGQYGLVAACAAGGQGHAMIVEAYPK	HADHB	40

191	<p>MLRGRSLSVTSLGGLPQWEVEELPVEELLFEVAWEVTNKVGGIYTVIQT  KAKTTADEWGENYFLIGPYFEHNMKTQVEQCEPVNDAVRRRAVDAMNKH  GCQVHFGRWLIEGSPYVVLFDIGYSAWNLDRWKGDLEWACSVGIPYHDR  EANDMLIFGSLTAWFLKEVTDHADGKYVVAQFHEWQAGIGLILSRARKL  PIATIFTTHATLLGRYLCAANIDFYNHLDKFNIDKEAGERQIYHRYCMERA  SVHCAHVFTTVSEITAIEAEHMLKRKPDVVTNGLNVKKFSAVHEFQNLH  AMYKARIQDFVRGHFYGHLDLDFLEKTLFLFIAGRYEFSNKGADIFLESLSR  LNFLLRMHKSDITVMVFFIMPAKTNNFNVTETLKGQAVRKQLWDVAHSV  KEKFGKLYDALLRGEIPDLNDILDRDDLTIMKRAIFSTQRQSLPPVTTHN  MIDDSTDPILSTIRRIGLFNNRTRDRVKVILHPEFLSSTSPLLPMDYEEFVRGC  HLGVFSPSYEPWGYTPAECTVMGIPSVTTNLSGFGCFMQEHVADPTAYGI  YIVDRRFRSPDDSCNQLTKFLYGFCQSRQRRIQRNRTERLSDLLDWRYL  GRYYQHARHLTISRAFPDKFHVELTSPPTTEGFKYPRPSSVPPSPSGSQASS  PQSSDVEDEVEDERYDEEEEEAERDRLNIKSPFSLSHVPHGKKKLHGEYKN</p>	GYS2
192	<p>MAKPLTDQEKRRQISIRGIVGVENVAELKKSFNRLHFTLVKDRNVATTR  DYYFALAHTVRDHLVGRWIRTQQHYDYKCPKRVYYLSLEFYMGRTLQN  TMINLGLQACDEAIYQLGLDIEELEIEEDAGLGNGLGRLAACFLDSM  ATLGLAAYGYGIRYEGIFNQKIRDGWQVEEADDWLRYPNPWEKSRPEF  MLPVHFYKVEHTNTGKWIDTQVVLALPYDTPVPGYMNNTVNTMRLW  SARAPNDFNLRDFNVGDYIQAVLDRNLAENISRVLYPNDNFFEGKELRLK  QEYFVVAATLQDIIRRFKASKFGSTRGAGTVFDFAPDQVAIQLNDTHPAL  AIPELMRIFVDIEKL  PWSKAWELTQKTFAYTNHTVLPEALERWPVDLVEKLLPRHLEIIEYNQK  HLDRIVALFPKDVDRRLRRMSLIEEESKRINMAHLICVGSNAVNGVAKIHS  DIVKTKVFKDFSELEPKFQNKTNGITPRRWLLCNPGLAELIAEKIGEDY  VKDLSQLTKLHSFLGDDVFLRELAKVKQENKLFKFSQFLETEYKVKINPSS  MFDVQVKRIHEYKRQLLNCLHVITMYNRIKKDPKLFVPRTVIIGGKAAP  GYHMAKMIKLITSVADVNNNDPMVGSKLKVFLENYRVSLAEKVIPATD  LSEQISTAGTEASGTGNMKFMLNGALTIGTMDGANVEMAEAGEENLFIF  GMRIDDVAALDKKGYEAKEYEALPELKLVIDQIDNGFFSPKQPDLFKDII  NMLFYHDRFKVFADYEAAYVKCQDKVSQLYMNPKAWNTMVLKNIAASG  KFSSDRITIKYAQNIWNVESDLKISLSNESNKVNGN</p>	PYGL
193	<p>MTEDKVTGTLVFTVITAVLGSFQFGYDIGVINAPQQVIISHYRHVLGVPLD  DRKAINNYVINSTDELPTISYSMNPKPTWAEETVAAAQLITMLWSLSVS  SFAVGGMTASFFGGWLGDTLGRIKAMLVANILSLVGALLMGFSKLGPSHI  LIIAGRSISGLYCGLISGLVPMYIGEIAPTALRGALGTFHQLAIVTGILISQIIG  LEFILGNYDLWHILLGLSGVRAILQSLLLFFCPESPRYLYIKLDEEVKAKQS  LKRLRGYDDVTKDINEMRKEREEASSEQKVSIIQLFTNSSYRQPILVALML  HVAQQFSGINGIFYYSTSIFQTAGISKPVYATIGVGAVNMVFTA VSVFLVE  KAGRRSLFLIGMSGMFVCAIFMSVGLVLLNKFSWMSYVSMIAIFLVSFFE  IGPGPIPWFMAEFFSQGPRPAALAIAAFSNWT CNFIVALCFQYIADF CGP  YVFFLFAGVLLAFTLFTFFKVPETKGSFEEIAAEFQKKS GSAHRPKAAVE  MKFLGATETV</p>	SLC2A2

10

20

30

40

50

194	MAASCLVLLALCLLLPLLLLGGWKRWRGRRAARHVAVVVLGDVGRSPR MQYHALSLAMHGFSVTLLGFCNSKPHDELLQNNRIQIVGLTELQSLAVGP RVFQYGVKVVVLQAMYLLWKLWREPGAYIFLQNPGLPSIAVCWFVGC LCGSKLVIDWHNYGYSIMGLVHGPNHPLVLLAKWYEKFFGRLSHLNLCV TNAMREDLADNWHIRAVTVYDKPASFFKETPLDLQHRLFMKLGSMHSPF RARSEPDPVTERSATERDAGSGLVTRLRERPALLVSSTSWTEDEDFE SILLAALEKFEQLTLDGHNLPVLCVITGKGPLREYYSRLIHQKHQHIQVCTP WLEAEDYPLLLGSADLGVCLHTSSSGLDLPKVVDMFVGCCLPVCVAVNFK CLHELVKHEENGLVFEDSEELAAQLQMLFSNFPDPAGKLNQFRKNLRESQ QLRWDESWWQTVLPLVMDT	ALG1
195	MAEEQGRERDSVPKPSVFLHPLDLGVGGAERLVLDAALALQARGCSVKI WTAHYDPGHCFAESRELPVRCAGDWLPRGLGWGGRGAAVCAYVRMVF LALYVFLADEEFDVVCDQVSACIPVFRRLARRRKKILFYCHFPDLLLTKR DSFLKRLYRAPIDWIEEYTTGMADCILVNSQFTA AVFKETFKSLSHIDPDV LYPSLNVTSFDSVPEKLLDLPKGGKFLLLSINRYERKKNLTLALEALV QLRGRLTSQDWERVHLIVAGGYDERVLENVEHYQELKKMVQQSDLGQY VTFLRSFSDKQKISLLHSCCTVLYTPSNEHFGIVPLEAMYMQCPVIAVNSG GPLESIDHSVTGFLCEPDPVHFSEAIEKFIREFSLKATMGLAGRARVKEKFS PEAFTEQLYRYVTKLLV	ALG2
196	MAAGLRKRGRSGSAAQAEGLCQWLQRAWQERRLLLREPRYTLLVAAC LCLAEVGITFWVIHRVAYTEIDWKAYMAEVEGVINGTYDYTQLQGDGTP LVYPAGFVYIFMGLYYATSRGTDIRMAQNIFAVLYLATLLLFLIYHQTC KVPPFVFFFMCCASYRVHSIFVLRFLNDPVAMVLLFLSINLLAQRWGWG CCFFSLAVSVKMNVLFPAGLLFLLLTQFGFRGALPKLGCAGLQVVLGLP FLLNPSGYLSRFDLGRQFLFHWTVNWRFLPEALFLHRAFHLALLTAHL TL LLLFALCRWHRTGESILSLLRDPKRRKVPQPLTPNQIVSTLFTSNFIGICFS RSLHYQFYVWYFHTLPYLLWAMPARWLTHLLRLLVLGLIELSWNTYPST SCSSAALHICHA VILLQLWLGPQFPKSTQHSKKAH	ALG3
197	MEKWYLMTVVVLIGLTVRWTVSLNSYSGAGKPPMFGDYEAQRHWQEIT FNLPVKQWYFNSSDNNLQYWGLDYPPLTAYHSLLCAYVAKFINPDWIAL HTSRGYESQAHKLFMRRTVLIADLLIYIPAVVLYCCCLKEISTKKKIANAL CILLYPGLILIDYGHFQYNSVSLGFALWGVLGISDCDLLGSLAFCLAINY KQMELYHALPFFCFLGKCFKGLKGGKGFVLLVKLACIVVASFVLCWLPF FTEREQTLQVLRRLFPVDRGLFEDKVANIWC SFNVFLKIKDILPRHIQLIMS FCSTFLSLLPACIKLILQSSKGFKFTLVSCALSFVLFVQVHEKSILLVSLPV CLVLSEIPFMSTWFLLVSTFSMLPLLLKDELMPVSVTTMAFFIACVTSFSI FEKTSEEELQLKSFSISVRKYLPCFTFLSRIIQYFLISVITMVLLTLMTVTL DPPQKLPDLFSVLVCFVSCLNFLFVYFNIIIMWDSKSGRNQKKIS	ALG6

10

20

30

40

50

198	MAALTIATGTGNWFSALALGVTLKCLLIPTYHSTDFEVHRNWLAITHSL PISQWYYEATSEWTLDYPPFFAWFEYILSHVAKYFDQEMLNVHNLNYSSS RTLLFQRFVSIVFMDVLFVYAVRECKCIDGKKVVGKELTEKPKFILSVLLLW NFGLLIVDHIHFQYNGFLFGLMLLSIARLFQKRHMEGAFLFAVLLHFKHIY LYVAPAYGVYLLRSYCFTANKPDGSIRWKSFSFVRVISLGLVVFLVSALS LPFLALNQLPQVFSRFLPFKRGGLCHAYWAPNFWALYNALDKVLSVIGLK LKFLDPNNIPKASMTSGLVQQFQHTVLPSTPLATLICTLIAILPSIFCLWFK PQGPRGFLRCLTLCALSSFMFGWHVHEKAILLAILPMSLLSVGKAGDASIF LILTTTGHYSLFPLLFTAPELPIKILLMLLFTIYSISSLKTLFRKEKPLFNWME TFYLLGLGPLEVCCEFVFPFTSWKVKYPPFIPLLLTSVYCAVGITYAWFKLY VSVLIDSAIGKTKKQ	ALG8
199	MASRGARQRLKSGGASSGDTAPAADKLRELLGSREAGGAEHRTLSGNK AGQVWAPEGSTAFKCLLSARLCAALLSNISDCDETFNYWEPHYLIYGE FQTWEYSPAYAIRSYAYLLLHAWPAAFHARILQTNKILVFYFLRCLLAFVS CICELYFYKAVCKKFGHLHVSRRMMLAFLVLSTGMFCSSSAFLPSSFCMYTT LIAMTGWYMDKTSIAVLGVAAGAILGWPFSAALGLPIAFDLLVMKHRWK SFFHWSLMALILFLVPVVVIDSYYYGKLVIAPLNIVLYNVFTPHGPDLYGT EPWYFYLINGFLNFVAFALALLVPLTSLMEYLLQRFHVQNLGHPYWLT LAPMYIWFIIFFIQPHKEERFLFPVYPLICLCGAVALSALQKCYHFVFQRYR LEHYTVTSNWLALGTVFLFGLLSFSRSVALFRGYHGPLDLYPEFYRIATDP TIHTVPEGRPVNVCVGKEWYRFPSSFLLPDNWQLQFIPSEFRGQLPKPFAE GPLATRIVPTDMNDQNLEEPSRYIDISKCHYLVDLDTMRETREPKEYSSNK EEWISLAYRPFLDASRSSKLLRAFYPFLSDQYTVYVNYTILKPRKAKQIR KKSGG	ALG9
200	MAAGERSWCLCKLLRFFYSLFFPGLIVCGTLCVCLVIVLWGIKLLLQRKK KLVSTSKNGKNQMVI AFFHPYCNAGGGGERVLWCALRALQKKYPEAVY VVYTGDNVNVNGQQILEGAFRRFNIRLIHPVQFVFLRKRYLVEDSLYPHFTL LGQSLGSIFLGWEALMQCVPDVYIDSMGYAFTLPLFKYIGGCQVGSYVHY PTISTDMLS VVKNQNI GFNNAAFITRNPF LSKVKLIYYLFAFIYGLV GSCS DVVMVNSSWTLNHILSLWKVG NCTNIVYPPCDVQTFLDIPLHEKKMTPG HLLVSVGQFRPEKNHPLQIRAFKLLNKKMVESPPSLKLV LIGGCRNKDD ELRVNQLRRLSED LGVQEYVEFKINIPFDELKNYLSEATIGLHTMWNEHF GIGVVECMAAGTIILAHNSGGPKLDIVVPHEGDITGFLAESEEDY AETIAHI LSMSAEKRLQIRKSARASVSRFSDQEFVTF LSSVEKLFK	ALG11

10

20

30

40

50

201	<p>MAGKGSSGRRPLLLGLLVAVATVHLVICPYTKVEESFNLQATHDLLYHW          QDLEQYDHLEFPGVVPRFTLGPVVIAVFSSPAVYVLSLLEMSKFYSQLIVR          GVLGLGVIFGLWTLQKEVRRHFGAMVATMFCWVTAMQFHLMFYCTRTL          PNVLALPVVLLALAAWLRHEWARFIWLSAFIIVFRVELCLFLGLLLLLAL          GNRKVSVVRLRHAVPAGILCLGLTVAVDSYFWRQLTWPEGKVLWYNT          VLNKSSNWGTSPLLWYFYFYSALPRGLGCSLLFIPLGLVDRRTHAPTVALG          FMALYSLLPHKELRFIYAFPMLNITAARGCSYLLNNYKKSPLYKAGSLL          VIGHLVVNAAYSATALYVSHFNYPGGVAMQRLHQLVPPQTDVLLHIDVA          AAQTGVSRLFQVNSAWRYDKREDVQPGTGMLAYTHILMEAAPGLLALY          RDTHRVLASVVGTTGVSLLNTQLPPFNHLQTKLVLLERLPRPS</p>	ALG12
202	<p>MKCVFVTVGTTSFDDLIACVSAPDSLQKIESLGYNRLILQIGRGTVVPEPFS          TESFTLDVYRYKDSLKEDIQKADLVISHAGAGSLETLEK GKPLVVINEK          LMNNHQLELAKQLHKEGHLFYCTCRVLTCPGQAKSIASAPGKCQDSAAL          TSTAFSGLDFGLLSGYLHKQALVTATHPTCTLLFPSCHAFFPLPTPTLYK          MHKGWKNYCSQKSLNEASMDEYLGSLGLFRKLTAKDASCLFRAISEQLF          CSQVHHLEIRKACVSYMRENQQTFSYVEGSFEKYLRLGDPKESAGQLE          IRALSILIYNRDFILYRFPKPTTYVTDNGYEDKILLCYSSSGHYDSVYSKQF          QSSAAVCQAVLYEILYKDFVVDDEELKTAIKLFRSGSKNRRNAVTVGSE          DAHTDYKSSNQNRMEEWGACYNENIPEGYNKGTEETKSPENPSKMPFP          YKVLKALDPEIYRNVEFDVWLDNRKELQKSDYMEYAGRQYYLGDKCQV          CLESEGRYYNAHIQEVGNENNSVTVFIEELAEKHVVPLANLKPVTQVMSV          PAWNAMP SRKGRGYQKMPGGYVPEIVISEMDIKQKMKMFKKIRGKEVY          M          TMAYGKGDPLLPPRLQHSMHYGHDPMPMHYSQTAGNVMSNEHFHPQHPS          PRQGRGYGMPRNSRFINRHNMPGPKVDFYPGPGKRCCQSYDNFSYRSRS          FRRSHRQMSCVNKESQYGFTPGNGQMPRGLEETITFYEEVEEGDETAIPTL          PNHGGPSTMPATSGYCVGRRGHSSGKQTLNLEEGNGQSENGRYHEEYL          YRAEPDYETSGVYSTTASTANLSLQDRKSCSMSPQDTVTSYNYPOKMMG          NIAAVAASCANNVPAPVLSNGAAANQAISTTSVSSQNAIQPLFVSPPTHGR          PVIASPSYPCHSAIPHAGASLPPPPPPPPPPPPPPPPPPPPPPPPPPALDVG          ET          SNLQPPPPPLPPPPYSCDPSGSDLPQDTKVLQYYFNLGLQCYHSHYWHSMV          YVPQMQQQLHVENYPVYTEPPLVDQTVPCYSEVRREDGIQAEASANDT          FPNADSSSVPHGAVYYPVMSDPYGPPLPGFDSCLPVVPDYSCVPPWHPV          GTAYGGSSQIHGAINPGPIGCIAPSPASHYVPQGM</p>	ALG13

10

20

30

40

50

203	<p>MGSLFRSETMCLAQLFLQSGTAYECLSALGEKGLVQFRDLNQNVSFQR  KFGVEVKRCEELERILVYLVQEINRADIPLPEGEASPPAPPLKQVLEMQEQ  LQKLEVELREVTKNKEKLRKNLLELIEYTHMLRVTKTFVKRNVEFEPTYE  EFPSLESDSLDDYSCMQRLGAKLGFVSGLINQKGKVEAFEKMLWRVCKGY  TIVSYAELDESLEDPETGEVIKWYVFLISFWGEQIGHKVKKICDCYHCHVY  PYPNTAEERREIQEGLNTRIQLDYTVLHKTEDYLRQVLCCKAAESVYSRVIQ  VKMKKAIYHMLNMCSFDVTNKCLIAEVCPEADLQDLRRALEEGSRESG  ATIPSEFMNIPTKETPPTRIRTNKFTEGFQNIVDAYGVGSYREVNPAFTIITF  PFLFAVMFGDFGHGFVMFLFALLLVLNENHPRLNQSQEIMRMFFNGRYIL  LLMGLFSVYTGLIYNDCFSKSVNLFGSGWNVSAMYSSSHPPAEHKKMVL  WNDSVVRHNSILQLDPSIPGVFRGPYPLGIDPIWNLATNRLTFLNSFKMKM  SVILGIIHMTFGVILGIFNHLHFRKKFNIYLSIPELLFMLCIFGYLIFMIFYK  WLVFSAETSRVAPSILIEFINMFLFPASKTSGLYTGQEYVQRVLLVVTALS  VPVFLGKPLFLLWLHNGRSCFGVNRSGYTLIRKDSEEEVSLGSDIEEG  NHQVEDGCREMACEEFNFGIEMTQVIHSIEYCLGCISNTASYLRLWALS  AHAQLSDVLWAMLMRVGLRVDTTYGVLLLLPVIALFAVLTIFILLIMEGL  SAFLHAIRLHWVEFQNKFYVGAGTKFVPP  SFSLLSSKFNNDDSDVA</p>	ATP6V0A 2
204	<p>MRPPACWLLAPPALLALLTCSLAFGLASEDTKKEVKQSQDLEKSGISRK  NDIDLKGIQVFIQSQNSFHAKRAEQLKKSILKQAADLTQELPSVLLLHQL  AKQEGAWTILPLLPHFSVTYSRNSSWIFFCEEETRIQIPKLETLRRYDPSK  EWFLGKALHDEEATIIHHYAFSENPTVFKYPDFAAGWALSIPLVNKLTKR  LKSESLKSDFTIDLKHEIALYIWDKGGGPPLTPVPEF  CTNDVDFYCATTFHSFLPLCRKPKKKDIFVAVKTCKKFHGDRIPIVKQT  WESQASLIEYSDYTENSIPTVDLGIPNTDRGHCCKTFAILERFLNRSQDK  TAWLVIVDDDTLISISRLQHLLSCYDSGEPVFLGERYGYGLGTGGYSYITG  GGGMVFSREAVRLLASKCRCYSNDAPDDMVLGMCFSGLGIPVTHSPLF  HQARPVDYPKDYLSHQVPISFHKHWNIQPKVYFTWLAPSDDEKARQET  QKGFREEL</p>	B3GLCT
205	<p>MFPRPLTPLAAPNGAEPLGRALRRAPLGRARAGLGGPPLLLPSMLMFAVI  VASSGLLLMIERGILAEKPLPLHPPGREGTAWRGKAPKPGGLSLRAGDA  DLQVRQDVRNRTLRAVCGQPMPDPDLPVGGQRRTLLRHILVSDRYRF  LYCYVPKVACSNWKRVMKVLAVGLDSVDVRLKMDHRSDLVFLADLRPE  EIRYRLQHYFKFLVREPLERLLSAYRNKFGIERYQQRYGAEIVRRYRAG  AGPSPAGDDVTFPEFLRYLVEDDPERMNEHWMPVYHLCQPCAVHYDFV  GSYERLEADANQVLEWVRAPPHVRFPARQAWYRPASPESLHYHLCAPR  ALLQDVLPKYILDFSLFAYPLPNVTKEACQQ</p>	CHST14

10

20

30

40

50

206	<p>MATAATSPALKRDLRDPAAALFETHGAEIEIRGLERQVRAEIEHKKEELRQ  MVGERYRDLIEAADTIGQMRRCAVGLVDAVKATDQYCARLRQAGSAAP  RPPRAQQPQQPSQEKFY SMAAQIKLLEIPEKIWSSMEASQCLHATQLYLL  CCHLHSLQLDSSSSRYSPVLSRFPILIRQVAAASHFRSTILHESKMLLKCC  GVSDQAVAEALCSIMLLEESSRQALTDFFLLARKATIQLLNQPHHGAGI  KAQICSLVELLATTLKQAHALFYTLPEGLLPDPALPCGLLFSTLETITGQHP  AGKGTGVLQEEMKLC SWFKHLPASIVEFQPTLR TL AHPISQEY LKDTLQK  WIHMCNEDIKNGITNLLMYVKSMKGLAGIRDAMWELLTNESTNHSWDV  LCRRLLEKPLLFWEDMMQQLFLDRLQTLTKEGFDSISSSSKELLVSALQEL  ESSTSNPSNKHIFEYNMSLFLWSESPNDLPSDAAWVSVANRGQFASSG  LSMKAQAISPCVQNFCSALDSKLVKLDLDDLLAYLPSDD  SSLPKDVSPQTAKSSAFDRYADAGTVQEMLRTQSVACIKHIVDCIRAELOQ  IEEGVQGGQDALNSAKLHSVLFMARLCQSLGELCPHLKQCILGKSESSEK  PAREFRALRKQGGKVKTQEIIPTQAKWQEVKEVLLQSSVMGYQVWSSAV  VKVLIHGFTQSLLLDDAGSVLATATSWDELEIQEEAESGSSVTSKIRLPAQ  PSWYVQSFLFSLCQEINRVGGHALPKVTLQEMLKSCMVQVVAAYEKLSE  EKQIKKEGAFPVTQNRALQLLYDLRYLNIVLTAKGDEVKSGRSKPDSRIE  K  VTDHLEALIDPFDLDVFTPHLNSNLHRLVQRTSVLFGVLTGTENQLAPRSS  TFNSQEPHNILPLASSQIRFGLLPLSMTSTRKAKSTRNIETKAQVPPARST  AGDPTVPGSLFRQLVSEEDNTSAPSLFKLGWLSSMTK</p>	COG1
207	<p>MEKSRMNLPGPDTLCFDKDEFMKEDFDVDHFVSDCRKRVQLEELRDDL  ELYYKLLKTAMVELINKDYADFNVLSTNLVGMKALNQLSVPLGQLREE  VLSLRSSVSEGIRAVDERMSKQEDIRKKKMCVLR LIQVIRSV EKIEKILNSQ  SSKETSALASSPLL TGQILER IATEFNQLQFHAVQSKGMPLLDKVRPRIAG  ITAMLQQSLEGLLLEGLQTSVDIIRHCLRTYATIDKTRDAEALVGQVLVK  PYIDEVIEQFVESHVNGLQVMYNKLLFVPHHCRLREVTGGAISSSEKGN  TVPGYDFLVNSVWPQIVQGLEEKLP SLFNPGNPDAFHEKYTISMDFVRRLL  ERQCGSQASVKRLRAHPAYHSFNKKWNLVYFQIRFREIAGSLEAALTDV  LEDAPAESPYCLLASHRTWSSLRRCWSDEMFLPLL VHRLWRLTLQILARY  SVFVNELSLRPISNESPKKPLVTGSKEPSITQGNTE DQGGPSETKPVVS  ISRTQLVYVVADLDKLQEQLPELLEIHKPKLEMIGFKNFSSISAALEDSQSSF  SACVPSLSSKIIQDLSDFGFLKSALEVPRLYRRTNKEVPTTASSYVDSAL  KPLFQLQSGHKDKLQAIHQWLEGLTSESTHKYYETVSDVLNSVKKMEE  SLKRLKQARKTTPANVPGPSGGMSDDDKIRLQLALDVEYLGEIQIKLGLQ  ASDIKSFSALAEVAAAKDQATAEQP</p>	COG2

10

20

30

40

50



208	MADLDSPPKLSGVQQPSEGVGGGRCSEISAELIRSLTELQELEAVYERLCG EEKVVERELDALLEQQNTIESKMVTLHRMGPNLQIEGDAKQLAGMITFT CNLAENVSSKVRQLDLAKNRLYQAIQRADDILDLKFCMDGVQTALRSED YEQAAAHTHRYLCLDKSVIELSRQGKEGSMIDANLKLQEAQRLKAIVA EKFAIATKEGDLQVERFFKIFPLLGLHEEGLRKFSEYLCKQVASKAEENL LMVLGTDMSDRRAAVIFADTLTLFEGIARIVETHQPIVETYYPGRLYTL IKYLQVECDRQVEKVVDKFIKQRDYHQFRHVQNNLMRNSTTEKIEPREL DPILTEVTLMNARSELYLRFLKKRISSDFEVGDSMASEEVKQEHQKCLDK LLNNCLLSCTMQELIGLYVTMEEYFMRETVNKAVAALDITYEKGQLTSSMV DDVFYIVKKCIGRALSSSIDCLCAMINLATTELESDFRDVLCNKLRMGFP ATTFQDIQRGVTSAVNMHSSLQQKFDTKGIESTDEAKMSFLVTLNNVE VCSENISTLKKTTLESDCTKLFSQIGGEQAQAKFDSCLSDLA AVSNKFRDL LQEGLTELNSTAIKPQVQPWINSFFSVSHNIEEEEFNDYEANDPWVQQFIL NLEQQMAEFKASLSPVIYDSL TGLMTSLVAVELEKVVLKSTFNRLGGLQF DKELRSLIAYLTTVTWTWIRDKFARLSQMATILNLERVTEILDYWGPNSSGP LTWRLTPAEVRQVLALRIDFRSEDIKRLRL	COG4
209	MGWVGRRRDSASPPGRSRSAAADDINPAPANMEGGGGSVAVAGLGARG SGAAAATVRELLQDGCYSDFLNEDFDVKTYTSQSIHQAVIAEQLAKLAQG ISQLDRELHLQVVARHEDLLAQTGIESLEGVLQMMQTRIGALQGA VDRI KAKIVEPYNKIVARTACLARLQVACDLLRRIIRILNLSKRLQGQLQGG SRE ITKAAQSLNELDYL SQIDLSGIEVIENDLLFIARARLEVENQAKRLL EQGL ETQNPTQVGTALQVFYNLGLKDTITSVVDGYCATLEENINSALDIKVLT QPSQSAVRGGPGRSTMPGNTAALRASFWTNMEKLMMDHIYAVCGQVQ HLQKVLAKKRDPVSHICFIEEIVKDGQPEIFYTFWNSVTQALSSQFHMATN SSMFLKQAFEGEYPKLLRLYN DLWKRLQQYSQHIQGNFNASGTTDLYVD LQHMEDDAQDIFIPKPDYDPEKALKDSLQPYEAYLSKSLSR LFDPINLV FPPGGRNPPSSDEL DGIKTIASELNVA AVDTNLT LAVSKNVAKTIQLYSV KSEQLLSTQGDASQVIGPLTEGQRRNVAVVNSLYKLHQS VTKAIHALME NAVQPLLT SVGDAIEAIIITMHQEDFSGSLSSSGKPDVPCSLYMKELQG FIA RVMSDYFKHFECLDFVFDNTEAIAQRAVELFIRHASLIRPLGEGGKMRLA ADFAQMELAVGPF CRRVSDLGKSYRMLRSFRPLL FQASEHVASSPALGD VIPFSIIIQFLFTRAPAELKSPFQRAEWSHTRFSQWLDDHPSEKDRLLLIRGA LEAYVQSVRSREGKEFAPVYPIMVQLLQKAMSALQ	COG5

10

20

30

40

50

210	MAEGSGEVVAVSATGAANGLNNGAGGTSATTNPLSRKHLKILETRLDN DKEMLEALKALSTFFVENSLRTRRNLRGDIERKSLAINEEFVSIFKEVKEEL ESISEDVQAMSNCCQDMTSRLQAAKEQTQDLIVKTTKLQSESQKLEIRAQ VADAFLSKFQLTSDEMSLLRGTREGPITEDFFKALGRVKQHNDVKVLLR TNQQTAGLEIMEQMALLQETAYERLYRWAQSECRILTQESCDVSPVLTQ AMEALQDRPVLYKYTLDEFGTARRSTVVRGFIDALTRGGPGGTPRPIEMH SHDPLRYVGDMLAWLHQATASEKEHLEALLKHVTTQGVENIQEVVGGHI TEGVCRPLKVRIEQVIVAEPGAVLLYKISNLLKFYHHTISGIVGNSATALLT TIEEMHLLSKKIFFNSLSLHASKLMDKVELPPPDLGPSSALNQTLMMLREV LASHDSSVPLDARQADFVQVLSVLDPLLQMCTVSASNLGTADMATFM VNSLYMMKTTLALFEFTDRRLEMLQFQIEAHLDTLINEQASYVLTRVGLS YIYNTVQQHKPEQGLANMPNLDVTLKAAMVQFDRYLSAPDNLLIPQL NFLLSATVKEQIVKQSTELVCRAYGEVYAAMVNPINEYKDPENILHRSPQ QVQTLLS	COG6
211	MDFSFLADDVKEWINAAFRAGSKEAASGKADGHAATLVMKLQLFIQ EVNHAVEETSHQALQNPVKVLRDVEALKQEASFLKEQMILVKEDIKFFE QDTSQSMQVLVEIDQVKSRLQAAESLQEADKWTLSADIEETFKTQDIA VISAKLTGMQNSLMMLVDTPDYSEKCVHLEALKNRLEALASPQIVAAFTS QAVDQSKVFKVFTVIDRMPQLLAYYYKCHKVQLLAAWQELCQSDLSL DRQLTGLYDALLGAWHTQIQWATQVFQKPHEVVMVLLIQTLGALMPSLP SCLSNQVERAGPEQELTRLLEFYDATAHFAKGLEMALLPHLHEHNLVKV TELVDAVYDPYKPYQLKYGDMEESNLLIQMSAVPLEHGEVIDCVQELSHS VNKFLGLASAAVDRCVRFNTGLGTCGLLSALKSLFAKYVSDFTSTLQ SIR KKCKLDHIPNSLFQEDWTAFAQNSIRIIATCGELLRHCGDFEQQLANRILST AGKYLSDSCSPRSLAGFQESILTDDKNSAKNPWQEYNYLQKDNPAEYASL MEILYTLKEKGSSNHNLLAAPRAALTRLNQQAHQLAFDSVFLRIKQQLLI SKMDSWNTAGIGETLTDELPAFSLTPLEYISNIGQYIMSLPLNLEPFVTQED SALELALHAGKLPFPPEQGDELPELDNMADNWLGSIARATMQTYCDAIL QIPELSPHSAKQLATDIDYLVMDALGLQPSRTLQHIVTLLKTRPEDYRQ VSKGLPRRLATTVATMRSVNY	COG7
212	MATAATIPSVATATAAALGEVEDEGLLASLFRDRFPEAQWRERPDVGRY LRELSGGLERLRREPERLAEERAQLLQQRDLAFANYKTFIRGAECTERI HRLFGDVEASLGRLLDRLPSFQQSCRNFVKEAEEISSNRRMNSLTLNRHTE ILEILEIPQLMDTCVRNSYEEALELAAYVRRLERKYSSIPVIQGIVNEVRQ SMQLMLSQLIQQLRTNIQLPACLR VIGYLRRMDVFTAEELRVKFLQARDA WLR SILTAIPNDDPYFHITKTIEASRVHLFDIITQYRAIFSDDEDPLPPAMGE HTVNESAIFHGWVLQKVSQFLQVLETDLYRGIGGHLDSLLGQCMYFGLSF SRVGADFRGQLAPVFQRVAISTFQKAIQETVEKFQEEMNSYMLISAPAILG TSNMPAAVPATQPGTLQPPMVLLDFPPLACFLNNILVAFNDLRLCCPVAL AQDVTGALEDALAKVTKIILAFHRAEEAAFSSGEQELFVQFCTVFLEDLVP YLNRLQVLFPPAQIAQTLGIPPTQLSKYGNLGHVNIGAIQEPLAFILPKRE TLFTLDDQALGPELTAPAPEPPAEPRLEPAGPACPEGGRAETQAEPSPVG P	COG8

10

20

30

40

50

213	DRLLQQGSAVFQFRMSANSGLLPASVMPLLGVMKERCQTAGNPFFER FGIVVAATGMAVALFSSVLALGITRPVPTNTCVILGLAGGVIIYMKHSLSV GEVIEVLEVLLIFVYLNMILLYLLPRCFTPGEALLVLGGISFVLNQLIKRSLT LVESQGDVPVDFLLVVVVGVMVLMGIFSTLFFVFMDSGTWASSIFFHLMTC VLSLGVVLPWLHRLIRRNP LLWLLQFLFQTDTRIYLLAYWSSLATLACL VLYQNAKRSSSESKKHQAPTIARKYFHLIVVATYIPGIIFDRPLLYVAATVC LAVFIFLEYVRYFRIKPLGHTLRSFSLFLDERDSGPLILTHIYLLLGMSLPI WLIPRPCTQKGSLLGARALVPYAGVLA VGVGD TVASIFGSTMGEIRWPGT KKT FEGTMTSIFAQHSVALILIFDSGVLDLNYSYAWILGSISTVSLLEAYTTQ IDNLLLPLYLLILLMA	DOLK
214	MSWIKEGELSLWERFCANIIKAGPMPKHIAFIMDGNRRYAKKCQVERQEG HSQGFNKLAE TLRWCLNLGILEVTVYAFSIENFKRSKSEVDGLMDLARQK FSRLMEEKEKLQKHGVCIRVLGDLHLLPLDLQELIAQAVQATKNYNKCF NVCFA YTSRHEISNAVREMAWGVEQGLLDPSESLDKCLYTNRSPHP DILIRTSGEVRLSDFLLWQTS HSLVFPVLPWEYTFWNLFEAILQFQMNH SVLQKARDMYAEERKRQQLERDQATVTEQLLREGLQASGDAQLRRL HKL SARREERVQGF LQALELKRADWLARLGTASA	DHDDS
215	MWAFSELPMPLLINLIVSLLGFVATVTLIPAFRGHFIAARLCGQDLNKT QQIPESQGVISGAVFLIILFCFIPFPFLNCFVKEQCKAFPHHEFVALIGALLAI CCMIFLGFADDVNLNRWRHKLLLP TAASPLLMVYFTNFGNTTIVVPKPF RPILGLHLDLILYVYMGLLAVFCTNAINILAGINGLEAGQSLVISASIIVF NLVELEGDCRDDHVFSLYFMIPFFFTLGLLYHNWYPSRVFVGD TFCYFA GMTFAVVGILGHFSKTMLLFFMPQVFNFYSLPQLLHIIPCRRHRIPRLNIK TGKLEMSYSKFKTKSLSFLGT FILKVAESLQLVTVHQSETEDGEFTECNN MTLINLLLKVLGPIHERNL TLLLLLQILGSAITFSIRYQLVRLFYDV	DPAGT1
216	MASLEVSRSRPRRSRRELEVRSPRQNKYSVLLPTYNERENLPLIVWLLVKS SEGINYEIIIIIDDGSPDGRDVAEQLEKIYGS DRILLRPREKKLGLGTAYIH GMKHATGNYIIIMDADLSHHPKFIPEFIRKQKEGNFDIVSGTRYKGNGGVY GWDLKRKII SRGANFLTQILLRPGASDLTGSFRLYRKEVLEKLIKCVSKG YVFQMEMIVRARQLNYTIGEVPI SFVDRVYGESK LGGNEIVSFLKGLLTLFATT	DPM1
217	MATGTDQVVGLGLVAVSLIIFTYYTAWVILLPFIDSQHVHVKYFLPRAYAV AIPLAAGLLLLLVGLFISYVMLKTKRVTKKAQ	DPM2
218	MTKLAQWLWGLAILGSTWVALTTGALGLELPLSCQEVLPWPAYLLVSA GCYALGTVG YRVATFHDCEDAARELQSQIQEARADLARRGLRF	DPM3
219	MESTLGAGIVIAEALQNQLAWLENVWLWITFLGDPKILFLFYFPAAYYAS RRVGI AVLWISLITEWLNLIKWFLFGDRPFWWHESGYYSQAPAVHQF PSSCETGPGSPSGHCMITGAALWPIMTALSSQVATRARSRWRVVMPSLAY CTFLAVGLSRIFILAHFPHQVLAGLITGAVLGWLMTPRVPMERELSFYGL TALALMLGTS LIYWTLFTLGLDLSWSISLAFKW CERPEWIHVDSRPFASLS RDSGAALGLGIALHSPCYAQVRRRAQLGNGQKIACLVLAMGLLGPLDWLG HPPQISLFYIFNFKYTLWPCLVLALVPWAVHMFSAQEAPPIHSS	G6PC3

10

20

30

40

50

220	<p>MCGIFAYLNYHVPRTREILETLIKGLQRLEYRGYDSAGVGFDDGGNDKD  WEANACKIQLIKKKGKVKALDEEVHKQQDMDLDIEFDVHLGIAHTRWA  THGEPSPVNSHPQRSKNNFIVIHNGIITNYKDLKKFLESKGYDFESETDT  ETIAKLVKMYMDNRESQDTSFTTLVERVIQQLEGAFALVFKSVHFPQAV  GTRRGSPLLIGVRSEHKLSTDHIPILYRTARTQIGSKFTRWGSQGERGKDK  KGSCNLSRVDDSTCLFPVEEKAVEYYFASDASAVIEHTNRVIFLEDDDDVA  AVVDGRLSIHRIKRTAGDHPGRAVQTLQMELOQIMKGNFSSFMQKEIFEQ  PESVVNTMRGRVNFDDYTVNLGGLKDHKEIQRRLILACGTSYHAGV  ATRQVLEELTELPVMVELASDFLDRNTPVFRDDVCFFLSQSGETADTLMG  LRYCKERGALTVGITNTVGSISRETDCGVHINAGPEIGVASTKAYTSQFV  SLVMFALMMCDRISMQERRKEIMLGLKRLPDLIKEVLSMDDEIQKLATE  LYHQKSVLIMGRGYHYATCLEGALKIKEITYMHSEGILAGELKHGPLALV  DKLMPVIMIIMRDHTYAKCQNALQQVVARQGRPVVICDKEDTETIKNTK  RTIKVPHSVDCLQGILSVIPLQLLAFHLAVLRGYDVFPRNLAKSVTVE</p>	GFPT1
221	<p>MLKAVILIGGPQKGTFRPLSFEVPKPLFPVAGVPMIQHHIEACAQVPGMQ  EILLIGFYQPDEPLTQFLEAAQQEFNLPVRYLQEFAPLGTGGGLYHFRDQIL  AGSPEAFFVLNADVCSDPFSAMLEAHRQRHPFLLGTTANRTQSLNYG  CIVENPQTHEVLHYVEKPSTFISDIINCGIYLFSPALKPLRDVQQRNQQDG  QLEDSPGLWPGAGTIRLEQDVFSALAGQGQIYVHL  TDGIWSQIKSAGSALYASRLYLSRYQDTHPERLAKHTPGGPWIRGNVYIH  PTAKVAPSAVLGPNVSIKGVTVGEGVRLRESIVLHGATLQEHTCVLHSIV  GWGSTVGRWARVEGTPSDPNPNDRARMSESLFKDGKLLPAITILGCRV  RIPAEVLILNSIVLPHKELSRSTNQIIL</p>	GMPPA
222	<p>MKALILVGGYGTRLRPLTLSTPKPLVDFCNKPILLHQVEALAAAGVDHVI  LAVSYMSQVLEKEMKAQEQRLGIRISMSHEEEPLGTAGPLALARDLLSET  ADPFFVLNSDVICDFPFQAMVQFHRHHGQEGSILVTKVEEPSKYGVVCE  ADTGRIHRFVEKPVFVSNKINAGMYILSPAQLRIQLQPTSIEKEVFPIMA  KEGQLYAMELQGFWMDIGQPKDFLTGMCLFLQSLRQKQPERLCSGPGIV  GNVLVDPSARIGQNCISGPNVSLGPGVVVEDGVCIRRCTVLRDARIRSHS  WLESCIVGWRCRVGQWVRMENVTVLGEDVIVNDELNLNGASVLPKHSIG  ESVPEPRIIM</p>	GMPPB
223	<p>MAARWRFWCVSVTMVVALLIVCDVPSASAQRKKEMVLSEKVSQMLMEW  TNKRPVIRMNGDKFRRLVKAPPRNYSVIVMFTALQLHRQCVVCKQADEE  FQILANSWRYSSAFTNRIFAMVDFDEGSDVFQMLNMNSAPTFINFPAKG  KPKRGDTYELQVRGFSAEQIARWIADRTDVNIRVIRPPNYAGPLMLGLLL  AVIGGLVYLRRSNMEFLFNKTGWAFALCFVLAMTSGQMWNHIRGPPY  AHKNPHTGHVNYIHGSSQAQFVAETHIVLLFNGGVTLGMVLLCEAATSD  MDIGKRKIMCVAGIGLVVLLFFSWMLSIFRSKYHGYPYSFLMS</p>	MAGT1

10

20

30

40

50

224	MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVVMYPPPPPPHRDFI SVTLSFGENYDNSKSWRRRSCWRKWKQLSRLQRNMILFLLAFLFCGLLF YINLADHWKALAFRLEEEQKMRPEIAGLKPANPPVLPAPQKADTDPENLP EISSQKTQRHIQRGPPHLQIRPPSQDLKDGTEEAATKRQEAPVDPRPEGDP QRTVISWRGAVIEPEQGTLPSSRAEVPTKPLPPARTQGTPVHLNRYRQKG VIDVFLHAWKGYRKFAWGHDELKPVSRSEWFGGLTLIDALDTMWIL GLRKEFEEARKWVSKKLHFEKDVDVNLFESTIRILGGLLSAYHLSGDSLFL RKAEDFGNRLMPAFRTPSKIPYSDVNIGTGVAHPPRWTSSTVAEVTSIQL EFRELSRLTGDKKFQEAWEKVTQHIHGLSGKKDGLVPMFINTHSGLFTHL GVFTLGARADSYEYLLKQWIQGGKQETQLLEDYVEAIEGVRTHLLRHS EPSKLTfVGEIAGHGRFSAKMDHLVCFLPGTLALGVYHGLPASHMELAQE LMETCYQMNRMETGLSPEIVHFNLYPQGRRDVEVKPADRHNLRLPET VESLFYLYRVTGDRKYQDWGWEILQSFSRFTRVPSGGYSSINNVQDPQKP EPRDKMESFFLGETLKYFLLLFSDDPNLLSLDAYVFNTEAHPLPIWTPA	MAN1B1
225	MRFRIYKRKVLILTLVVAACGFVLWSSNGRQRKNEALAPLLDAEPARG AGGRGGDHPSVAVGIRRVSNVSAASLPAVPQPEADNLTLRYRSLVYQL NFDQTLRNVDKAGTWAPRELVLVVQVHNRPEYLRLLLDLSLKAQGIDNV LVIFSHDFWSTEINQLIAGVNFPCVLQVFFPFSIQLYPNEFFGSDPRDCPRDL PKNAALKLGCINAEYPDSFGHYREAKFSQTKHHWWKLFHFVWERVKIL RDYAGLILFLEEDHYLAPDFYHVFKKMWKLKQCEPCDVLSLGTYSAS RSF YGMADKVDVKTWKSTEHNMGLALTRNAYQKLIECTDTFCTYDDYNWD WTLQYLTVSCLPKFWKVLVPQIPRIFHAGDCGMHHKKTCPSTQSAQIES LLNNNKQYMFPELTISEKFTVVAISPPRKNNGGWGDIRDHELCKSYRRLQ	MGAT2
226	MARGERRRAVPAEGVRTAERAARGGPRRDGRGGGPRSTAGGVALAV VVLSLALGMSGRWVLAWYRARRAVTLHSAPPVLPADSSSPAVAPDLFW GTYRPHVYFGMKTRSPKPLLTGLMWAQQGTTTPGTPKLRHTCEQGDGVG PYGWEFHDGLSFGRQHIQDGALRLTTEFVKRPGGQHGGDWSWRVTVEP QDSGTSALPLVSLFFYVVDGKEVLLPEVGAKGQLKFISGHTSELGDFRFT LLPPTSPGDTAPKYGSYNVFWTSNPGPLLLTEMVKSRLNSWFQHRPPGAP PERYLGLPGSLKWEDRGPSGQGQQLIQVTLKIPISIEFVFESGSAQAGG NQALPRLAGSLLTQALESHEGFRERFEKTFQLKEKGLSSGEQVLGQAAL SGLLGGIGYFYGQGLVLPDIGVEGSEQKVDPALFPPVPLFTA VPSRSFFPRG FLWDEGFHQLVVQRWDPSLTREALGHWLGLLNADGWIGREQILGDEAR ARVPPEFLVQRAVHANPPTLLLPAVHMLEVGD PDDLAFLRKALPRLHAW FSWLHQSQAGPLPLSYRWRGRDPALPTLLNPKTLPSGLDDYPRASHPSVT ERHLDLRCWVALGARVLTRELAEHLGAEVAAELGPLAASLEAAESLDEL HWAPELGVFADFGNHTKAVQLKPRPPQGLVRVVGRPQPQLQYVDALGY VSLFPLLLRLLDPTSSRLGPLLDILADSRHLWSPFGLRSLAASSSFYQQRNS EHDPPYWRGAVWLVNPNYALGALHGHYGHLEGPHQARA AKLHGELRAN VVGNVWRQYQATGFLWEQYSDRDGRGMGCRPFHGWTSLVLLAMAEDY	MOGS

10

20

30

40

50

227	MAAEADGPKRLLVPILLPEKCYDQLFVQWDLHVPCLKILLSKGLGLGI VAGSLLVKLPQVFKILGAKSAEGLSLQSVMLELVALTGTMVYSITNFPFS SWGREALFLMLQTITICFLVMHYRGQTVKGVAFACYGLVLLVLLSPLTPL TVVTLLQASNPVAVVGRLLQAATNYHNGHTGQLSAITVFLFVGGSLARI FTSIQETGDPLMAGTFVSSLCNGLIAAQLLFYWNAKPPHKQKKAQ	MPDU1
228	MAAPRVFPLSCAVQQYAWGKMGSNSEVARLLASSDPLAQIAEDKPYAEL WMGTHPRGDAKILDNRISQKTLQSWIAENQDSLGSVKKDTFNGNLPFLFK VLSVETPLSIQAHPNKELAEKHLHLQAPQHYPDANHKPEMAIALTPFQGLC GFRPVEEIVTFLKVKPEFQFLIGDEAATHLKQTMSHDSQAVASSLQSCFSH LMKSEKVVVEQLNLLVKRISQQAAGNNMEDIFGELLQLHQYYPGDI GCFAYFLNLLTLKPGEAMFLEANVPHAYLKGDCVECMACSDNTVRAGL TP KFIDVPTLCEMSYTPSSSKDRLFLPTRSQEDPYLSIYDPPVPDFTIMKTEVP GSVTEYKVLALDSASILLMVQGTVIASPTTQTPIPLQRGGVLFIGNESVS LKLTEPKDLLIFRACCLL	MPI
229	MAAAALGSSSGSASPAVAELCQNTPETFLEASKLLLTYADNLRNPNDEK YRSIRIGNTAFSTRLLPVRGAVECLFEMGFEEGETHLIFPKASVEQLQKIR DLIAIERSRDLGDSNKSHKVKSSQPAASTQLPTTPSSNPSGLNQHTRNRQ GQSSDPPSASTVAADSAILEVLQSNIQHVLVYENPALQEKALACIPVQELK RKSQEKLSRARKLDKGINISDEDFLLLELLHWFKEE FFHWVNNVLCSKCGGQTRSRDRSLLPSDDELKWGAKEVEDHYCDACQF SNRFPRYNNEPKLLETRCGRGGEWANCFTLCCRAVGFEARYVWDYTDH VWTEVYSPSQRWLHCDACEDVCDKPLL YEIGWGKLSYVIAFSKDEVV DVTWRYSCHEEVIARRTKVKEALLRDTINGLNKQRQLFSENRRKELLQ RIIVELVEFISPKTPKPGELGGRISGSVAWRVARGEMGLQRKETFIPCENE KISKQLHLCYNIVKDRYVRVSNNTISGWENGWKMESIFRKVETDWH MVYLARKEGSSFAYISWKFECGSVGLKVDISIRTSSQTFQTGTVEWKLRS DTAQVELTGDNSLHSYADFSGATEVILEAELSRGDGDVAWQHTQLFRQS LNDHEENCLEIIKFSDL	NGLY1
230	MVKIVTVKTQAYQDQKPGTSGLRKRVKVFQSSANYAENFIQSIISTVEPA QRQEATLVVGGDGRFYMKEAIQLIARIAAANGIGRLVIGQNGILSTPAVSC IIRKIKAIIGGIILTASHNPGGPNDFGIKFNISNGGPAPEAITDKIFQISKTIEE YAVCPDLKVDLGLVGLKQFDLENKFKPFTVEIVDSVEAYATMLRSIFDFS ALKELLSGPNRLKIRIDAMHGTVGYPYVKKILCEELGAPANSVAVNCVPLED FGGHHDPNLTYAADLVETMKSGEHDFGAADFDDGDRNMILGKHGFFV NPSDSVAVIAANIFSIPYFQQTGVRGFARSMPTSGALDRVASATKIALYET PTGWKFFGNLMDASKLSLCGEESFGTSDHIREKDGLWAVLAWLSILAT RKQSVEDILKDHVQYGRNFFTRYDYEEVEAEGANKMMKDLEALMFDR SFVGKQFSANDKVYTVKADNFEYSDPVDGSISRNOGLRLIFTDGSRIVFR LSGTGSAGATIRLYIDSYEKDVAKINQDPQVMLAPLISIALKVSQQLQERTG RTAPTUIT	PGM1

10

20

30

40

50

231	<p>MDLGAITKYSALHAKPNGLILQYGTAGFRTKAEHLDHVMFRMGLLAVLR          SKQTKSTIGVMVTASHNPEEDNGVKLVDPLGEMLAPSWEEHATCLANAE          EQDMQRVLIDISEKEAVNLQQDAFVVIGRDTRPSSEKLSQSVIDGVTVLGG          QFHGYGLLTPQLHYMVYCRNTGGRYGKATIEGYYQKLSKAFVELTKQA          SCSGDEYRSLKVDCANGIGALKLREMEHYFSQGLSVQLFNDGSKGKLNH          LCGADFKSHQKPPQGMEIKSNERCCSFDGDADRIVYYYHDADGHFHLI          DGDKIATLISSFLKELLVEIGESLNIGVVQTAYANGSSTRYLEEVMKVPVY          CTKTGVKHLHHAQEFDIGVYFEANGHGTALEFSTAVEMKIKQSAEQLED          KKRKAAMLENIIDLNFQAAGDAISDMLVIEAILALKGLTVQQWDALYT          DLPNRQLKVQVADRRVISTTDAERQAVTPPGLQEAINDLVKKYKLSRAFV          RPSGTEDVVRVYAEADSQESADHLAHEVSLAVFQLAGGIGERPQPGF</p>	PGM3
232	<p>MGSQEVLGHAARLASSGLLLQVLFRLITFVLNAFILRFLSKEIVGVVNVRL          TLLYSTTLFLAREAFRACLSSGGTQRDWSQTLNLLWLTVPVPLGVFWSLFLG          WIWLQLEVPDPNVPHYATGVVLFGLSAVVELLGEPFVWLAQAHMV          KLK VIAESLSVILKSVLTAFLVLWLPHWGLYIFSLAQLFYTTVLVLCYVIY          FTKLLGSPESTKLQTLVSRITDLLPNITRNGAFINWKEAKLTWSFFKQSFL          KQILTEGERYVMTFLNVLNFGDQGVYDIVNNLGSVARLIFQPIEESFYIFF          AKVLERGKDATLQKQEDVAVA AVLESLLKLALLAGLTITVFGFAYSQL          ALDIYGGTMLSSGSPVLLRSYCLYVLLLAINGVTECFFAAMSKEEVDR          YNFVMLALSSSFLVLSYLLTRWCGSVGFILANCFNMGIRITQSLCFIHRYY          RRSPhRPLAGLHLSPVLLGTFALSSGGVTAVSEVFLCCEQGWPARLAHIAV          GAFCLGATLGTAFLETETKLIHFLRTQLGVPRRTDKMT</p>	RFT1
233	<p>MATYLEFIQQNEERDGVRFVSWNVWSSRLEATRMVVPLACLLTPLKERP          DLPPVQYEPVLCRPTCKAVLNPLCQVDYRAKLWACNFCFQRNQFPAY          GGISEVNQPAELMPQFSTIEYVIQRGAQSPLIFLYVVDTCLEEDDLQALKES          LQMSLSLLPPDALVGLITFGRMVQVHELSCGSKSYVFRGTDLTAKQIQ          DMLGLTKPAMPMQARPAQPQEHFPAASSRFLQPVHKIDMNLTDLLGELQ          RDPWPVTQGKRPLRSTGVALSIAVGLLEGTFPNTGARIMLFTGGPPTQGP          GMVVGDELKIPRSWHDIEKDNARFMKKATKHYEMLANRTAANGHCIDI          YACALDQTGLLEMKCCANLTGGYVMGDSFNLSLQTFQRFQIFTKDFNG          DFRMAFGATLDVKTSRELKIAGAIGPCVSLNVKGPCVSENELGVGGTSQ          WKICGLDPTSTLGIYFEVVNQHNTPIPQGGRGAIQFVTHYQHSSTQRRIRV          TTIARNWADVQSLRHIEAAFDQEA AVLMARLGVFRAESEEGPDVLRW          LDRQLIRLCQKFGQYNKEDPTSFRLSDFSLSYPQFMFHLRRSPFLQVFNNS          PDESSYRHHFARQDLTQSLIMIQPILYSYSFHGPPEPVLLDSSSILADRILL          MDTFFQIVIYLGETIAQWRKAGYQDMPEYENFKHLLQAPLDDAQEILQAR          FPMPRYINTEHGGSQARFLLSKVNPSQTHNNLYAWGQETGAPILTDDVSL          QVFMHDLKKLAVSSAC</p>	SEC23B

10

20

30

40

50

234	MAAPRDNVTLLFKLYCLAVMTLMAAVYTIALRYTRTSDKELYFSTTAVC ITEVIKLLLSVIGILAKETGSLGRFKASLRENVLGSPKELLKLSVPSLVYAVQ NNMAFLALSNDAAVYQVTYQLKIPCTALCTVLMNRTLSKLQWVSF MLCAGVTLVQWKAQATKVVVEQNPLLGFGAIAI AVLCSGFAGVYFEKV LKSSDTSLVWRNIQMYLSGIIVTLAGVYLSDGAEIKEKGFYGYTYVWF VIFLASVGLYTSVVVKYTDNIMKGFSA AAAI VLSSTIASVMLFGLQITLTF ALGTLVCVSIYLYGLPRQDTTSIQOGETASKERVIGV	SLC35A1
235	MAAVGAGGSTAAPGPGAVSAGALEPGTASAAHRRLKYISLAVLVVQNAS LILSIRYARTLPGDRFFATTAVVMAEVLKGLTCLLLLFAQKRGVVKHLVL FLHEAVLVQYVDTLKLA VPSLIYTLQNNLQYVAISNLPAATFQVTYQLKIL TTALFSVLMNRSLSRLQWASLLLLFTGVAIVQAQQAGGGGPRPLDQNP AGLA AVVASCLSSGFAGVYFEKILKSSGSVWLRNLQGLFGTALGLVGL WWAEGTAVATRGGFFGYTPAVWGVVNLNQA FGGLLVAVVVKYADNILK GFATSLSIVLSTVASIRLFGFHVDP L FALGAGLVIGAVYLYSLPRGAAKAIA SASASASGPCVHQPPGQPPPPQLSSHRGDLITEPFLPKLLTKVKGS	SLC35A2
236	MNRAPLKRSRILHMAL TGASDPSAEAE ANGEKPFLLRALQIALVVS LYWV TSISMVFLNKYLLDSPSLRLDTPIFVTFYQCLVTTLCKGLSALAACCPGA VDFPSRLDLRVARSVLPLSVVFIGMITFNCLKYVGVAFYNVGRSLTT VFNVL SYLLKQTT SFYALLTCGIIHGGFWLGV DQEGAEGTLSWLGTVFG VLASLCVSLNAIYTTKVLPAVDGSIWRLTFYNNVNACILFLPLLLL GELQ ALRDF AQLGSAHF WGM T L GGLFGFAIGYVTGLQIKFTSPLTHNVSG TAKACAQTVLAVLYYEETKSFLWWT SNMMVLGGSSAYTWVRGWEMK KTPEEPSPKDSEKSAMGV	SLC35C1
237	MAAMASLGALALLLSSLSRCSAEACLEPQITPSYYTTS DAVISTETVFIVE ISLTCKNRVQNMALYADVGGKQFPVTRGQDVGRYQVSWSLDHKSAHAG TYEVRFFDEESYLLRKAQRN NEDISIIPPLFTVSDHRGTWNGPWVSTEV LAAAIGLVIIYLAFSAKSHIQA	SSR4
238	MAPWAEAEHSALNPLRAVWLT LTA AFLTLLQLLPGLLP GCAIFQDLI RYGKTKCGEPSRPAACRAFDVPKRYFSHFYIISVLWNGFLLWCLTQSLFL GAPFPSWLHGLLRILGAAQFQGGELALS AFLVLVFLWLHSLRRLFECLYV SVFSNVMIHVVQYCFGLVYYVLVGLTVLSQVPM DGRNAYITGKNLLMQ ARWFHILGMMFIWSSAHQYKCHVILGNLRKNKAGVVIHCNHRIPFGDW FEYVSSPNYLAELMIYVSM AVTFGFHNL TWLVLVTNVFFNQALSAFLSH QFYKSKFVSYPKHRKAFLPFLF	SRD5A3
239	MAAAAPGNRASAPRLLLLFLVPLLWAPA AVRAGPDEDLSHRNKEPPAP AQLQPPVAVQGPEPARVEKIFTPAAPVHTNKEDPATQTNLGFIFHAFVA AISVIIVSELGDKTFFIAAIMAMRYNRLTVLAGAMLALGLMTCLSVLFGY ATTVIPRVYTYVSTVLF AIFGIRMLREGLKMSPDEGQEELEEVQAELKK KDEEFQRTKLLNGPGDVETGTSITVPQKKWLHFISPIFVQALTLTFLAEWG DRSQLTTIVLAAREDPYGVAVGGTVGHCLCTGLAVIGGRMIAQKISVRTV THGGIVFLAFAFSALFISPDSGF	TMEM165

10

20

30

40

50



240	MSSWLGGLGSLGQSLGQVGGSLASLTGQISNFTKDMLMEGTEEVEAEL PDSRTKEIEAIHAILRSENERLKKLCTDLEEKHEASEIQIKQQSTSYRNQLQ QKEVEISHLKARQIALQDQLLKLQSAQAQSVPSGAGVPATTASSSFAYGISH HPSAFHDDDDMDFGDISSQEQEINRLSNEVSRLESEVGHWRHIAQTSKAQGT DNSDQSEICKLQNIKELKQNRSQEIDDHQHEMSVLQNAHQKLTETISRRH REELSDYEERIEELENLLQGGSGVIETDLSKIYEMQKTIQVLQIEKVESTK KMEQLEDKIKDINKKLSSAENDRDILRREQEQLNVEKRQIMEECENLKLE CSKLQPSAVKQSDTMTEKERILAQASVVEEVFRLQQALSDAENEIMRSSL NQDNSLAEDNLKLMRIEVLEKEKSLLSQEKEELQMSLLKLNNEYEVIKS TATRDISLSELHDLRLNLEAKEQELNQSISEKETLIAEIEELDRQNQEATK HMILIKDQLSKQQNEGDSIISKLKQDLNDEKKRVHQLEDDKMDITKELDV QKEKLIQSEVALNDLHLTKQKLEDKVENLVDQLNKSQESNVSQIKENLEL KEHIRQNEEELSRIRNELMQSLNQDSNSNFKDTLLKEREAEVRNLKQNL ELEQLNENLKKVAFDVKMENEKLVLACEDVRHQLEECLAGNNQLSLEK NTIVETLMEKGEIEAELCWAKKRLLEEANKYEKTIEELSNARNLNTSAL QLEHEHLIKLNQKKDMEIAELKKNIEQMDTDHKETKDVLSSSLEEQKQLT QLINKKEIFIEKLERSSKLQEELDKYSQALRKNEILRQTIEEKDRSLGSMK EENHLQEELERLREEQSRTAPVADPKTLDVTELASEVSQLNTIKEHLEE EIKHHQKIHEDQNQSKMQLLQSLQEKKEMDEFQYQHEQMNAHTHTQLFL EKDEEIKSLQKTIEQIKTQLHEERQDIQTDNSDIFQETKVQSLNIENGSEKH DLSKAETERLVKGIKERELEIKLLNEKNISLTKQIDQLSKDEVGKLTQIIQQ KDLEIQALHARISSTSHTQDVVYLQQQLQAYAMEREKVFVAVLNEKTREN SHLKTEYHKMMDIVAAKEAALIKLQDENKKLSTRFESSGQDMFRETIONL SRIREKDIEIDALSQKCQTLAVLQTSSTGNEAGGVNSNQFEELLQERDKL KQQVKMEEWKQQVMTTVQNMQHESAQLQEELHQLQAQVLVDSDNNS KLQVDYTGLIQSYEQNETKLKNFGQELAQVQHSIGQLCNTKDLLLGKLDI ISPQLSSASLLTPQSAECLRASKSEVLSESSELLQQELEELRKSLEKDATIR TLQENNHRLSDSIAATSELERKEHEQTDSEIKQLKEKQDVLQKLLKEKDL LIKAKSDQLLSSNENFTNKVNENELLRQAVTNLKERILILEMDIGKLGGEN EKIVETYRGKETEQALQETNMKFSMMLREKEFECHSMKEKALAFEQLL KEKEQGKTGELNQLLNAVKSMQEKTVVFQQRDQVMLALKQKQMENT ALQNEVQRLRDKEFRSNQELERLRNHLLESEDSYTREALAAEDREAKLRK KVTVLEEKLVSNNAMENASHQASVQVESLQEQLNVVSKQRDETALQLS VSQEQVKQYALSLANLQMVLEHFQQEEKAMYSAELEKQKQLIAEWKKN AENLEGKVISLQECLDEANAALDSASRLTEQLDVKEEQIEELKRQNELRQ EMLDDVQKKLMSLANSSSEKVDKVLNRNLFHGFHTPKNQRHEVLRLM GSILGVRREEMEQLFHDDQGGVTRWMTGWLGGGSKSVPNTPLRPNQQS VVNSSFSELVFKFLETESHPSIPPKLSVHDMKPLDSPGRRKRDTNAPESFK DTAESRSGRRTDVNPFLAPRSAAVPLINPAGLPGGGPHLLLKPISDVLPF TPLPALPDNSAGVVLKDLLKQ	TRIP11
-----	--	--------

10

20

30

40

50

241	<p>MGARGAPSRRRQAGRRLRYLPTGSFPFLLLLLLLLCIQLGGGQKKKENLLA  EKVEQLMEWSSRRSIFRMNGDKFRKFIKAPPRNYSMIVMFTALQPQRQCS  VCRQANEYQILANSWRYSSAFCNKLFFSMVDYDEGTDVVFQQLNMNSAP  TFMHFPPKGRPKRADTFDLQRIGFAAEQLAKWIADRTDVHIRVFRPPNYS  GTIALALLVSLVGGLLYLRRNNLEFIYNKTGWAMVSLCIVFAMTSGQMW  NHIRGPPYAHKNPHNGQVSYIHGSSQAQFVAESHIILVLNAAITMGMVLL  NE  AATSKGDVVGKRRICLVGLGLVVFSSFLLSIFRSKYHGYPSDLDFE</p>	TUSC3
242	<p>MVCVLVLAAGAVAVFLILRIWVLRSMVTPRESLSILVVAGSGGHT  TEILRLGSLSNAYSPRHYVIADTDEMSANKINSFELDRADRDPSNMYTK  YYIHRIPRSREVQQSWPSTVFTTLHSMWLSFPLIHRVKPDLVLCNGPGTCV  PICVSALLLGILGIKKVIIVYVESICRVETLSMSGKILFHLSDYFIVQWPALK  EKYPKSVYLGRIV</p>	ALG14
243	<p>MRLREPLLSGSAAMPGASLQRACRLLVAVCALHLGVTLVYYLAGRDLSR  LPQLVGVSTPLQGGSNSAAAIGQSSGELRTGGARPPPPLGASSQPRPGGDS  SPVVDSPGPASNLTSPVPHTTALSPLACPEESPLLVGPMLEFNMPVDL  ELVAKQNPVVKMGGRYAPRDCVSPHKVAIIIPFRNRQEHLKYWLYYLP  VLQRQQLDYGIYVINQAGDTIFNRAKLLNVGFQEALKDYDYTCFVFS  DLIPMNDHNAYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSQQFLT  INGFPNNYWGWWGEDDDIFNRLVFRGMSISRPNVVGRCRMIRHSRDKK  NEPNPQRFDRIAHTKETMLSDGLNSLTYQVLDVQRYPLYTQITVDIGTPS</p>	B4GALT1
244	<p>MGYFRCARAGSFRRRKMEPSTAARAWALFWLLPLLGAVCASGPRTL  VLLDNLNVRETHSLFFRSLKDRGFELTFKTADDPSSLIKYGEFLYDNLHIF  SPSVEDFGGNINVETISAFIDGGGSVLVAASSDIGDPLRELGSECGIEFDEEK  TAVIDHHNYDISDLGQHTLIVADTENLLKAPTIVGKSSLNPILFRGVGMVA  DPDNPLVLDILTGSSTSYSFFPKPITQYPHAVGKNTLLIAGLQARNARVI  FSGSLDFFSDSFFNSAVQKAAPGSQRYSQTGNYELAVALSRRWVFKEEGVL  RVGPVSHHRVGETAPPNAYTVTDLVEYSIVIQQLSNGKWVPFDGDDIQLE  FVRIDPFVRTFLKKKGGKYSVQFKLPDVYGVFQFKVDYNRLGYTHLYSST  QVSVRPLQHTQYERFIPSAYPYASAFSMMLGLFIFSIVFLHMKEKEKSD</p>	DDOST
245	<p>MTGLYELVWRVLHALLCLHRTLTSWLRVRFGTWNWIWRRCCRAASAA  VLAPLGFTLRKPPAVGRNRRHHRHPRGGSCLA AAHHRMRWRADGRSLE  KLPVHMGLVITEVEQEPSFSDIASLVVWCMAVGISYISVYDHQGIFKRNS  RLMDEILKQQQELLGLDCSKYSPEFANSNDKDDQVLNCHLAVKVLSPED  GKADIVRAAQDFCQLVAQKQKRPTDLVDVDTLASLLSSNGCPDPDLVLKF  GPVDSTLGLPWHIRLTEIVSLPSHLNISYEDFFSALRQYAACEQRLGK</p>	NUS1

10

20

30

40

50

246	MAPPGSSTVFLALTIHASTWALTPHYLTKHDVERLKASLDRPFTNLESA FYSIVGLSSLGAQVPDAKKACTYIRSNLDPSNVDSLFYAAQASQALSGCEI SISNETKDLLLAUVSEDSSTQIYHAVAALSFGFLPLASQEALSALTARLS KEETVLATVQALQTASHLSQQADLRISIVEEIEDLVARLDELGGVYVLQFEE GLETTALFVAATYKLMDHVGTPEPSIKEDQVIQLMNAIFSKKNFESLSEAFS VASAAAVLSHNRYPVVPVVVPEGSASDTHEQAILRLQVTNVLSQPLTQA TVKLEHAKSVASRATVLQKTSFTPVGDFELNFMNVKFSSGYYDFLVEV EGDNRYIANTVELRVKISTEVGITNVDLSTVDKDQSIAPKTTRVTYPAKAK GTFIADSHQNFALFFQLVDVNTGAELTPHQTFVRLHNQKTGQEVVFAEP DNKNVYKFELDTSERKIEFDSASGTYTLYLIIGDATLKNPILWNVADVVIK FPEEEAPSTVLSQNLFTPQKQEIQHFLFREPEKRPPTV VSNTFTALILSPLLLLFALWIRIGANVSNFTFAPSTIIFHLGHAAMLGLMYV YWTQLNMFQTLKYLAAILGSVTFLAGNRMLAQQAVKRTAH	RPN2
247	MTTYLEFIQQNEERDGVRFSSWNVWPSSRLEATRMVVPVAALFTPLKERP DLPPIQYEPVLCSTRITCRAVLNPLCQVDYRAKLWACNFCYQRNQFPPSYA GISELNQPAELLQFSSIEYVVLGRQPMLIFLYVVDTCMEDEDLQALKES MQMSLSLLPPTALVGLITFGRMVQVHELGCESKSYVFRGTKDLSAKQL QEMGLSKVPLTQATRGPVQVQPPPSNRFLQPVQKIDMNLTDLLGELQRD PWPVPQGRPLRSSGVALSIAVGLLECTFPNTGARIMMFIGGPATQGGPM VVGDELKTPIRSWHDIDKDNAYVKKGKTKHFEALANRAATTGHVIDIYA CALDQTGLLEMKCCPNLTGGYVMVGDSEFNTSLFKQTFQRVFTKDMHGQ FKMGFGGTLEIKTSREIKISGAIGPCVSLNSKGPCVSENEIGTGGTCQWKIC GLSPTTTLAIYFEVNVQHNAPIPQGGRGAIQFVTQYQHSSGQRRIRVTTIAR NWADAQTQIQNIAASFDQEAAILMARLAIYRAETEEGPDVLRWLDRQLI RLCQKFGYHKKDDPSSFRFSETFSLYPQFMFHLRRSSFLQVFNNSPDESSY YRHHFMRQDLTQSLIMIPIYAYSFSGPPEPVLLDSSSILADRILLMDTFF QILYHGETIAQWRKSGYQDMPEYENFRHLLQAPVDDAQEILHSRFPMPR YIDTEHGGSQARFLLSKVNPSQTHNNMYAWGQESGAPILTDDVSLQVFM DHLKKLAVSSAA	SEC23A
248	MFANLKYVSLGILVFQTTSLVLTMRYSRTLKEEGPRYLSSTAVVVAELKI MACILLVYKDSKCSLRALNRVLHDEILNKP METLKLAIPSGIYTLQNNLLY VALSNLDAATYQVTYQLKILTTALFSVSMLS KKLGVYQWLSLVILMTGV AFVQWPSDSQLDSKELSAGSQFVGLMAVLTACFSSGFAGVYFEKILKETK QSVWIRNIQLGFFGSIFGLMGVYIYD GELVSKNGFFQGYNRLTWIVVVLQ ALGGLVIAAVIKYADNILKGFATSLSIILSTLISYFWLQDFVPTS VFFLGAIL VITATFLYGYDPKPAGNPTKA	SLC35A3

10

20

30

40

50

249	MGLLVFVRNLLLALCLFLVLGFLYSSAWKHLHLLQWEEDSNSVVLVSFDSA GQTLGSEYDRLGFLNLDLPAELATKYANFSEGACKPGYASALMTAIF PRFSKPAPMFLDDSRKWARIREFVPPFGIKGQDNLIKAILSVTKEYRLTPA LDSLRCRRCHIVGNNGVLANKSLGSRIDYDIVVRLNSAPVKGFEKDVGS KTTLRITYPEGAMQRPEQYERDSLFLAGFKWQDFKWLKYIVYKERVSA SDGFWKSVA TRVPKEPPEIRILNPYFIQEAFTLIGLFPNGLMGRGNIPTL GSVAVTMALHGCDEVAVAGFGYDMSTPNAPLHYYETVRMAAIKESWTH NIQREKEFLRKLVKARVITDLSSGI	ST3GAL3
250	MTKFGFLRLSYEKQDTLLKLLLSMAAVLSFSTRLFAVLRVESVIHEFDPYF NYRTTRFLAEEGFYKFHNWFDDRAWYPLGRIIGGTIYPGLMITSAAIYHVL HFFHITIDIRNVCVFLAPLFSSTTIVTYHLTKELKDAGAGLLAAAMIAVVP GYISRSVAGSYDNEGIAIFCMLTYMVIKAVKTGSICWAAKCALAYFY MVSSWGGYVFLINLIPLHVLVLMGTGRFSHRIYVAYCTVYCLGTILSMQIS FVGFQPVLSSEHMAAFGVFGLCQIHAFVDYLRSKLNPQQFEVLFRSVISLV GFVLLTVGALLMLTGKISPWTRGFYSLDPSYAKNNIPIIASVSEHQPTTW SSYYFDLQLLVFMFPVGLYYCFNSLSDARIFIIMYGVTSMYFSAVMVRML LVLAPVMCILSGIGVSQVLSTYMKNLDIRPDKSKKQDSTYPIKNEVAS GMILVMAFFLITYTFHSTWVTSEAYSSPSIVLSARGGDSRIIFDDFREAYY WLRHNTPEDAKVMSWWDYGYQITAMANRTILVDNNTWNNTHISRVGQ AMASTEELKAYEIMRELDVSYVLVIFGGTGYSSDDINKFLWMVRIGGSTD TGKHIKENDYYTPTGEFRVDREGSPVLLNCLMYKMCYYRFGQVYTEAKR PPGFDRVRNAEIGNKDFELDVLEEAYTTEHWLVRIYKVKDLNDRGLSRT	STT3A
251	MAEPSAPESKHKSSLNPPWGLMALGNSRHGHGPGAQCAHKAAGGA APPKPAPAGLSGGLSQPAGWQSLLSFTILFLAWLAGFSSRLFVIRFESIHE FDPWFNYRSTHHLASHGFYEFLNWFDERAWYPLGRIVGGTVYPGLMITA GLIHWILNTLNTVHIRDVCVFLAPTFSGLTSISTFLLTRELWNQGAGLLAA CFIAIVPGYISRSVAGSFDNEGIAIFALQFTYYLWVKS VKTGSVFWTMCCL LSYFYMVSAWGGYVFIINLIPLHVFVLLMQRYSKRVYIAYSTFYIVGLILS MQIPFVGFQPIRTSEHMAAAGVFALLQAY AFLQYLRDRLTKQEFQTLFFL GVSLAAGAVFLSVIYLYTYGYIAPWSGRFYSLWDTGYAKIHIPIIASVSEH QPTTWVSFFDLHILVCTFPAGLWFCIKNINDERVFVALY AISA VYFAGVM VRLMLTLTPVVCMLSAIAFSNVFEHYLGDDMKRENPPVEDSSDEDDKRN QGNLYDKAGKVRKHATEQEKT EELGPNIKSIVTMLMLMLLMMFAVHC TWVTSNAYSSPSVVLASYNHDGTRNLDLDFREAYFWLRQNTDEHARVMS WWDYGYQIAGMANRTTLVDNNTWNNSHIALVGKAMSSNETAAYKIMR TLDVDYVLVIFGGVIGYSGDDINKFLWMVRIAEGEHPKDIRESDYFTPQGE FRVDKAGSPTLLNCLMYKMSYYRFGEMQLDFRTPPGFDRTRNAEIGNKDI KFKHLEEAFTEHWLVRIYKVKAPDNRETLDHKPRVTNIFPKQKYL SKKT TKRKRGYIKNKL VFKKGKKISKKT V	STT3B

10

20

30

40

50

252	MARKSNLPVLLVPFLLCQALVRCSSPLPLVVNTWPFKNATEAAWRALAS GGSALDAVESGCAMCEREQCDGSGVFGGSPDELGETTLDAMIMDGTTM DVGAVGDLRRIKNAIGVARKVLEHTTHTLLVGESATTFASMGFINEDLS TTASQALHSDWLARNCPNYWRNVIPDPSKYCGPYKPPGILKQDIPIHKE TEDDRGHDITIGMVVIHKTGHIAAGTSTNGIKFKIHGRVGDSPGAGAYA DDTAGAAAATGNGDILMRFLPSYQAVEYMRRGEDPTIACQKVISRIQKHF PEF FGAVICANVTGSYGAACNKLSTFTQFSFMVYNSEKNQPTEEKVDCI	AGA
253	MGAPRSLLLALAAGLAVARPPNIVLIFADDLGYGDLGCYGHPSSTTPNLD QLAAGGLRFTDFYVPVSLCTPSRAALLTGRLPVRMGMYPGVLPSSRGG LPLEEVTVAEVLAAARGYLTMAGKWHLGVGPEGAFPPHQGFHRLGIP YSHDQGPCQNLTCFPPATPCDGGCDQGLVPIPLLANSVEAQPPWLPGLE ARYMAFAHDLMADAQRQDRPFLYYASHHHTYPQFSGQSFASRGRGPF GDSLMELDAAVGTLMTAIGDLGLEETLVIFTADNGPETMRMSRGGCSG LLRC GKGTTYEGGVREPALAFWPGHIAPGVTHELASSDLLPTLAALAGAPLPN VTLDGFDLSPLLLGTGKSPRQSLFFYPSYPDEVRGVFAVRTGKYKAHFFT QGSASHDTTADPACHASSSLTAHEPPLLVDLSKDPGENYNLLGGVAGATP EVLQALKQLQLLKAQLDAAVTFGPSQVARGEDPALQICCHPGCTPRPACC HCPDPA	ARSA
254	MGPRGAASLPRGPGRRLLLPPVPLLLLLLLLAPPGSGAGASRPPHLVFL ADDLGWNDVGFHGSRI RTPHLDALAAGGVLLDNYYTQPLCTPSRSQLLT GRYQIRTGLQHIIWPCQPSCVPLDEKLLPQLLKEAGYTTHMVGKWHLG MYRKECLPTRRGFDTYFGYLLGSEDIYSHERCTLIDALNVTRCALDFRDG EEVATGYKNMYSTNIFTKRAIALITNHPPEKPLFLYLALQSVHEPLQVPEE YLKPYDFIQDKNRHHYAGMVSLMDEAVGNVTAALKSSGLWNNTVFIFST DNGGQTLAGGNNWPLRGRKWSLWEGGVRGVGFVASPLLKQKGVKNRE LIHISDWLPTLVKLARGHTNGTKPLDGFVWKTISEGSPSPRIELHNDPN FVDSSPCPRNSMAPAKDDSSLPEYSAFNTSVHAAIRHGNWKLTTGYPGCG YWFPPSQYNVSEIPSSDPPTKTLWLFDIDRDPEERHDLSREYPHIVTKLLS RLQFYHKHSVPVYFPAQDPRCDPKATGVWGPWM	ARSB
255	MPGRSCVALVLLAAVSCAVAQHAPPWTEDCRKSTYPPSGPTYRGAVP WYTINLDLPPYKRWHELMMLDKAPVLKIVNSLKNMINTFVPSGKIMQVV DEKLPGLLGNFPGPFEEEMKGIAAVTDIPLGEIISFNIFYELFTICTSIVAEDK KGHLIHGRNMDFGVFLGWNINNDTWVITEQLKPLTVNLDQFQNNKTVFK ASSFAGYVGMGTGFKPGLFSLTLNERFSINGGYLGILEWILGKKDVMWIG FLTRTVLENSTSYEEAKNLLTKTKILAPAYFILGGNQSSEGCVITRDRKESL DVYELDAKQGRWYVVQTNYDRWKHPFLDDRTPAKMCLNRTSQENIS FETMYDVLSTKPVLNKLTVYTTLIDVTKGQFETYLRDCPDPCIGW	ASAHI

10

20

30

40

50

256	<p>MSADSSPLVGSTPTGYGTLTIGTSIDPLSSSVSSVRLSGYCGSPWRVIGYHV  VVWMMAGIPLLLFRWKPLWGVRLRLRPCNLAHAETLVIEIRDKEDSSWQ  LFTVQVQTEAIGEGSLESPQSQAEEDGRSQAAVGAVPEGAWKDTAQLHK  SEEAVSVGQKRVLRYLRFQGRYIWIETQQAFYQVSLLDHGRSCDDVHR  SRHGLSLQDMVRKAIYGNVISIPVKSYPQLLVDEALNPYYGFQAFSIAL  WLADHYWYALCIFLISSISICLSLYKTRKQSQTLRDMVKLSMRVCVCRP  GGEEEVWDSSELVPGDCLVLPQEGGLMPCDAALVAGECMVNESSLTGES  IPVLKTALPEGLGPYCAETHRRHTLFCGTLILQARAYVGPVHLAVVTRTG  FCTAKGGLVSSILHPRPINFKFYKHSMKFVAALSVLALLGTIYSIFILYRNR  VPLNEIVIRALDLVTVVVPPALPAAMTVCTLYAQSRLRRQGIFCIHPLRINL  GGKLQLVCFDKTGTLTEDGLDVMGVVPLKGGQAFPLV  PEPRRLPVGPLLRALATCHALSRLQDTPVGDPMDLKMVESTGWVLEEEP  AADSAGFTQVLAVMRPPLWEPQLQAMEEPPVPVSVLHRFPFSSALQRMS  VVVAWPGATQPEAYVKGSPELVAGLCNPETVPTDFAQMLQSYTAAGYR  VVALASKPLPTVPSLEAAQQLTRDTVEGDLSLLGLLVMRNLKPKQTPVI  QALRRTRIRAVMVTGDNLQTAVTVARCGGMVAPQEHLIIVHATHPERGQ  PASLEFLPMESPTAVNGVKDPDQAASYTVEPDPRSRHLALSGPTFGIIVKH  FPKL  LPKVLVQGTVFARMAPEQKTELVCCELQKLQYCVGMCGDGANDCGALKA  ADVGISLSQAEASVSPFTSSMASIECVPMVIREGRCSLDTSFVFKYMAL  YSLTQFISVLILYTINTNLGDLQFLAIDLVITTTVAVLMSRTGPALVLGRVR  PPGALLSVPVLSLLLQMVLTGVVQLGGYFLTLAQPFVPLNRTVAAPD  NLPNYENTVVFSLSSFQYLILAAVSKGAPFRRPLYTNVPFLVALALLSSV  LVGLVLPGLLQGPLALRNITDTGFKLLLLGLVTLNFVGFAMLESVLDQC  LPACLRRLRPKRASKKRFKQLERELAEQPWPPLPAGPLR</p>	ATP13A2
257	<p>MGGCAGSRRRFSDEGEETVPEPRLPLLDHQGAHWKNAVGFWLLGLCN  NFSYVVMLSAAHDILSHKRTSGNQSHVDPGPTPIPHNSSSRFDCNSVSTAA  VLLADILPTLVIKLLAPLGLHLLPYSRVLVSGICAAGSFVLVAFSHSVGTS  LCGVVFASISSGLGEVTFLSLTAFYPRAVISWWSSGTGGAGLLGALSYLGL  TQAGLSPQQTLLSMLGIPALLASYFLLTSPEAQDPGEEEAESAARQPLI  RTEAPESKPGSSSSLSLRERWTVFKGLLWYIVPLVVVYFAEYFINQGLFEL  LFFWNTSLSHAQQYRWYQMLYQAGVFASRSSLRCCRIRFTWALALLQCL  NLVFLADVWFGFLPSIYLVFLIILYEGLLGGAAVYNTFHNIALETSDEHR  EFAMAATCISDTLGISLSGLLALPLHDFLCQLS</p>	CLN3
258	<p>MAQEVDTAQAEMRRGAGAARGRASWCWALALLWLA VVPGWSRVSGI  PSRRHWPVPYKRDFRPKDPYQCAKYTFCTGSPVMEGDDDDIEVFRLQ  APVWEFKYGDLLGHLKIMHDAIGFRSTLTGKNYTMWYELFQLGNCTFP  HLRPEMDAPFWCNQGAACFFEGIDDVHWKENGTLVQVATISGNMFMNQ  AKWVKQDNETGIYYETWNVKASPEKGAETWFDSYDCSKFVLRFTFNKLA  EFGAEFKNIETNYTRIFLYSGEPTYLGNETS VFGPTGNKTLGLAIKRFYYPF  KPHLPTKEFLLSLLQIFDAVIVHKQFYLFYNFEYWFLPMKFPFIKITYEEIPL  PIRNKTL SGL</p>	CLN5

10

20

30

40

50

259	MEATRRRQHLGATGGPGAQLGASFLQARHGVSADAARTAPFHLDLW FYFTLQNWVLDVDFGRPIAMLVFPLEWFLNPKPSVGDYFHMAYNVITPFLLL KLIERSPRTLPRSTITYVSIHIFIMGASIHVLGDSVNHRLLFSGYQHHLVSVREN PIKLNKPETLIDSFELLYYYDEYLGHCMWYIPFFLILFMYFSGCFTASKAE SLIPGPALLLVAPSGLYYWYLVTEGQIFILFIFTFAMLALVLHQKRKRLFL DSNGLFLFSSFALTLLLVALWVAWLWNDPVLRRKKYPGVYVPEPWAFYT LHVSSRH	CLN6
260	MNPASDGGTSESIFDLDYASWGIKSTLMVAGFVYLGVFVCHQLSSSLN ATYRSLVAREKVFWDLAATRAVFGVQSTAAGLWALLGDPVLHADKARG QQNWCWFHITATGFFCFENVAVHLSNLIFRTFDLFLVIHHLFAFLGFLGC LVNLQAGHYLAMTTLLLEMSTPFTCVSWMLLKAGWSESLFWKLNQWL MIHMFHCRMVLYHMMWVCFWHWDGLVSSLYLPHLTLFLVGLALLTLI INPYWTHKKTQQLNPDWNFAQPEAKSRPEGNQQLRKKRP	CLN8
261	MIRNWLTIIFLPLKLVKCESSVSLVPPVVKLENGSSTNVSLTLRPLNA TLVITFEITFRSKNITILELPDEVVPPGVTNSSFQVTSQNVGQLTVYLHGN HSNQTGPRIRFLVIRSSAISINQVIGWIYFVAWSISFYQPQVIMNWRKSVIG LSDFVALNLTGFVAYSVFNIGLLWVPYIKEQFLKYPNGVNPVNSNDVF FSLHAVVLTLLIIVQCCLYERGGQVSWPAIGFLVLAWLFAFVTMIVA AVG VTTWLQFLFCFSYIKLAVTLVKYFPQAYMNFYKSTEGWSIGNVLLDFTG GSFSLQMFQSYNNDQWTLIFGDPTKFGVFSIVFDVFFIQHFCLYRK RPGYDQLN	CTNS
262	MIRAAPPPLFLLLLLLLLLLVSASRGEAAPDQDEIQRPLGLAKQPSFRQYS GYLKGSGSKHLHYWVVESQKDPENSPVVLWLNCGPGCSSLDGLLTHEGP FLVQPDGVTLEYNPYSWNLIANVLYLESPAGVGFYSDDKFYATNDTEV AQSNFEALQDFRFLPEYKNNKLFLLTGESYAGIYIPTLAVLVMQDPSMNL QGLAVGNLSSYEQNDNSLVYFAYYHGLLGNRLWSSLQTHCCSQNKCN FYDNKDLECVTNLQEVARIVGNISGLNIYNYAPCAGGVPSHFVRYEKDTV VVQD LGNIFTRLPLKRMWHQALLRSGDKVRMDPPCTNTTAASTYLNNPYVRKA LNIPQLPQWDMCNFLVNLQYRRLYRSMNSQYLKLLSSQYQILLYNGD VDMACNFMGDEWFVDSLQKMEVQRRPWLVKYGDSGEQIAGFVKEFS HIAFLTIKAGHMPVTDKPLAAFTMFSRFLNKQPY	CTSA
263	MQPSSLLPLALCCLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGP VSKYSQAVPAVTEGPIPEVLKNYMDAQYYGEIGITPPQCFTVVFDTGSS NLWVPSIHCKLLDIACWIHHKYNSDKSSTYVKNGTSTFDIHYGSGLSGYLS QDTSVPCQSASSALGGVKVERQVFG EATKQPGITFIAAKFDGILGMAYPRISVNNVLPVFDNLMQQKLVQDNIFSF YLSRDPDAQPGGELMLGGTDSKYKGSLSYLVNTRKAYWQVHLDQVEV ASGLTLCKEGCEAIVDTGTSLMVGPVDEVRELQKAIGAVPLIQGEYMIPCE KVSTLPAILKLGKGYKLSPEDYTLKVSQAGKTLCLSGFMGMDIPPSGP LWILGDVFIGRYTYVFDNRDNRVGFEEAARL	CTSD

10

20

30

40

50

264	MAPWLQLLSLLGLLPGAVAAPAQPRAASFQAWGPPSPELLAPTRFALEMF NRGRAAGTRAVLGLVRGRVRRAGQGSLSLEATLEPPCNDPMVCRLPV SKKTLLCSFQVLDLGRHVLLRKDCGPVDTKVPGAGEPKSAFTQGSAMIS SLSQNHDPNRNETFSSVISLLNEDPLSQDLPVKMASIFKNFVITYNRTYESK EEARWRLSVFVNNMVRAQKIQALDRGTAQYGVTKFSDLTEEFRTIYLN TLLRKEPGNKMKQAKSVGDLAPPEWDWRSGAVTKVKDQGMCGSCWA FSVTGNVEGQWFLNQGTLTLLSLSEQELLDCKMDKACMGGLPSNAYSIAK NLGGLLETEDDYSYQGHMQSCNFSAEKAKVYINDSVELSQNEQKLAAWL AKRGPISVAINAFGMQFYRHGISRPLRPLCSPWLIDHAVLLVGYGNRSDVP FWAIKNSWGTDWGEKGYYYLHRGSGACGVNTMASSAVVD	CTSF
265	MWGLKVLLLPVVSFALYPEEILDTHWELWKKTHRKYNNKVDEISRRLI WEKNLKYISIHNEASLGVHTYELAMNHLGDMTSEEVVQKMTGLKVPLS HSRNDTLYIPEWEGRAPDSVDYRKKGYVTPVKNQGQCGSCWAFSSVGA LEGQLKKKTGKLLNLSPQNLVDCVSENDGCGGGYMTNAFQYVQKNRGI DSEDAYPYVQGEESMYNPTGKAAKCRGYREIPEGNEKALKRAVARVGP VSVAIIDASLTSFQFYSGVYVYDESCNSDNLNHAVLAVGYGIQKGNKHVII KNSWGENWGNKGYILMARNKNNACGIANLASFPKM	CTSK
266	MADQRQRSLSTSGESLYHVLGLDKNATSDDIKKSYRKLALKYHPDKNPD NPEAADKFKEINNAHAILTDATKRNIYDKYGSGLGYVAEQFGEENVNTYF VLSSWWAKALFVFCGLLTCCYCCCCLCCCFNCCCCKKCPKAPEGEETEF YVSPEDLEAQLQSDEREATDTPIVIQPASATETTQLTADSHPSYHTDGFN	DNAJC5
267	MRAPGMRSRPAGPALLLLLLFLGAAESVRRAPPRRYTPDWPSLDSRPLP AWFDEAKFGVFIHWGVFVPAWGSEWFWWHWQGEGRPQYQRFMRDN YPPGFSYADFGPQFTARFFHPEEWADLFQAAGAKYVVLTTKHHEGFTNW PSPVSWNWNKDVGPHRDLVGELGTALRKRNIYGLYHSLLEWFHPLYL LDKKNGFKTQHFVSAKTMPELYDLVNSYKPDLIWSDGEWPCPTYWNS TNFLSWLYNDSPVKDEVVNDRWGQNCSCHHGGYNCEDKFKPQSLPD HKWEMCTSIDKFSWGYRRDMALSDVTEESEIISELVQTVSLGGNYLLNIG PTKDGLIVPIFQERLLAVGK WLSINGEAIYASKPWRVQWEKNTTSVWYTSKGSVYAIFLHWPENGVLN LESPITTSTTKITMLGIQGLKWSTRDPDKGLFISLPQLPPSAVPAEFAWTIKL TGVK	FUCA1

10

20

30

40

50



268	<p>MGVRHPPCSHRLLAVCALVSLATAALLGHILLHDFLLVPRELSGSSPVLEE          THPAHQQGASRPGRDAQAHPGRPRAVPTQCDVPPNSRFDCAPDKAITQE          QCEARGCCYIPAKQGLQGAQMGQPWCFFPPSYPSYKLENLSSEMGYTA          TLTRTPTFFPKDILTLRLDVMETENRLHFTIKDPANRRYEVPLETPHVVH          SRAPSPLYSVEFSEEPFGVIVRRQLDGRVLLNTTVAPLFFADQFLQLSTSLP          SQYITGLAEHLSPLMLSTSWTRITLWNRDLAPTGANLYGSHPPFYLALED          GGSAGHVLLNSNAMDVVLQSPALSWRSTGGILDVYIFLGPEPKSVVQQ          YLDVVGYPFMPPYWGLGFHLCRWGYSSAITRQVVENMTRAHFPLDVQ          WNDLDYMDSRRDFTFNKDGFRDFPAMVQELHQGGRRYMMIVDPAISS          GPAGSYRPYDEGLRRGVFITNETGQPLIGKVWPGSTAFPDTNPTALAWW          EDMVAEFHDQVPFDGMWIDMNEPSNFIRGSEDGCPNNELENPPYVPGVV          GGTLQAATICASSHQFLSTHYNLHNLVYGLTEAIASHRALVKARGTRPFVIS          RSTFAGHGRYAGHWTGDVWSSWEQLASSVPEILQFNLLGVPLVGADVCG          FLGNTSEELCVRWTQLGAFYPMRNHNSLLSLPQEPYSFSEPAQQAMRKA          LTLRYALLPHLYTLFHQAHVAGETVARPLFLEFPKDSSTWTVDHQLLWG          EALLITPVLQAGKAEVTGYFPLGTWYDLQTPVVEALGSLPPPPAAPREPAI          HSEGQWVTLPAPLDTINVHLRAGYIIPLQGPGLTTESRQQPMALAVALT          KGGEARGELFWDDGESLEVLERGAYTQVIFLARNNTIVNELVRVTSEGAG          LQLQKVTVLGVATAPQQVLSNGVPVSNFTYSPDTKVLDICVSLLMGEQFL          VSWC</p>	GAA
269	<p>MAEWLLSASWQRRAKAMTAAAGSAGRAAVPLLLCALLAPGGAYVLDD          SDGLGREFDGIGAVSGGGATSRLLVNYPEPYRSQILDYLFKPNFGASLHIL          KVEIGDGGQTTDGTESPHMHYALDENYFRGYEWWMKEAKKRNPNTLI          GLPWSFPGWLKGKFDWPYVNLQLTAYYVVTWIVGAKRYHDLDDIDYIGI          WNERSYNANYIKILRKMLNYQGLQRVKIIASDNLWESISASMLLDAELFK          VVDVIGAHYPGTHSAKDAKLTGKKLWSSSEDFSTLNSDMGAGCWGRILN          QNYINGYMTSTIAWNLVASYEQLPYGRCGLMTAQEPWSGHYVVEPV          WWSAHTTQFTQPGWYYLKTVGHLEKGGSYVALTDGLGNLTHIETMSHK          HSKCIRPFLPYFNVSQQFATFVLKGSFSEIPELQVWYTKLGKTSERFLFKQL          DSLWLLSDSGSFTLSLHEDELFTLTTLTTGRKGSYPLPPKSQPPSTYKDD          FNVDYPPFFSEAPNFADQTGVFEYFTNIEDPGEHHFTLRQVLNQRPIWAA          DASNTISIIGDYNWNTNLTIKCDVYIETPDTGGVFIAGRNVKGGILIRSARGIF          FWIFANGSYRVTGDLAGWIIYALGRVEVTAKKWYTLTLTIKGHFTSGML          NDKSLWTDIPVNFPKNGWAAIGTHSFEFAQFDNFLVEATR</p>	GALC

10

20

30

40

50

270	MAAVVAATRWWQLLLVLVLSAAGMGASGAPQPPNILLLLMDDMGWGDG VYGEPSRETPNLDRAAEGLLFPNFYSANPLCSPSRAALLTGRLPIRNGFY TTNAHARNAYTPQEIVGGIPDSEQLPELLKKAGYVSKIVGKWHLGHRPQ FHPLKHGFDEWFGSPNCHFGPYDNKARNIPVYRDWEMVGRYYEEFPIN LKTGEANLTQIYLQEALDFIKRQARHHPFFLYWAVDATHAPVYASKPFLG TSQRGRYGDVREIDDSIGKILELLQDLHVADNTFVFFTSNDGAALISAPE QGGSNPFLCGKQTFEGGMREPALAWWPGHVTAQVSHQLGSIMDLF TTSLALAGLTPPSDRAIDGLNLLPTLLQGRLMDRPIFYRGTLMMAATLG QHKAHFWTWTNSWENFRQIDFCPGQNVSGVTTHNLEDHTKLPLIFHLG RDPGERFPLSFASAEYQEALSRTSVVQQHQEALVPAQPQLNVCNWAVM NWAPPGEKLGKCLTPPESIPKKCLWSH	GALNS
271	MQLRNPELHLGCALALRFLALVSWDIPGARALDNGLARTPTMGWLHWE RFMCNLDCQEEPDISCEKLFMEMAELMVSEGWKDAGYEYLCIDDCWM APQRDSEGRLQADPQRFPHGIRQLANYVHSGKGLKGIYADVGNKTCAGF PGSFGYYDIDAQTFADWGVDLLKFDGCYCDLENLADGYKHMSLALNR TGRSIVYSCEWPLYMWPFPKPNYTEIRQYCNHWRNFADIDDSWKSISIL DWTSFNQERIVDVAGPGGWNDPMLVIGNFGLSWNQQVTQMALWAIM AAPLFMSNDRHISPQAKALLQDKDVIQDPLGKQGYQLRQGDNFEV WERPLSGLAWAVAMINRQEIG GPRSYTIAVASLGKGVACNPACFITQLLPVKRKLGFYEWTSRLRSHINPTG TVLLQLENTMQMSLKDLL	GLA
272	MPGFLVRILPLLLVLLLLGPTRGLRNATQRMFEIDYSRDSFLKDGQPFRYIS GSIHYSRVPRFYWKDRLLKMKMAGLNAIQTYVPWNFHEPWPGYQFSE DHDVEYFLRLAHELGLLVILRPGPYICAEWEMGGLPAWLEKESILLRSSD PDYLAADVDKWLGVLLPKMKPLLYQNGGPVITVQVENEYGSYFACDFDY LRFLQKRFRHHLGDDVVLFTTDGAHKTFKCGALQGLYTTVDFGTGSNIT DAFLSQRKCEPKGPLINSEFYTGWLDHWGQPHSTIKTEAVASSLYDILAR G ASVNLVYMFIGGTNFAYWNGANSPYAAQPTSVDYDAPLSEAGDLTEKYFA LRNIIQKFEKVPEGPIPPSTPKFAYGKVTLEKLTVGAALDILCPSGPIKSLY PLTFIQVKQHYGFVLYRTTLPQDCSNPAPLSSPLNGVHDRAVAVDGIQ GVLERNNVITLNTGKAGATDLLVENMGRVNYGAYINDFKGLVSNLTL SNILTDWTIFPLDTEDAVRSHLGGWHRDSGHHDEAWAHNSSNYTLPAF YMGNFSIPSGIPDLPQDTFIQFPGWTKGQVWINGFNLGRYWPARGPQLTL FVPQHILMTSAPNTITVLELEWAPCSSDDPELCAVTFVDRPVIGSSVTYDH PSKPVEKRLMPPPPQKNKDSWLDHV	GLB1
273	MQSLMQAPLLIALGLLLAAPAQAHLKKPSQLSSFSWDNCDEGKDPVIRS LTLEPDPIVPGNVTLVSMGSTSVPLSSPLKVDLVLEKEVAGLWIKIPCTDY IGSCTFEHFCVDLMLIPTGEPCEPLRXYGLPCHCPFKEGTYSLPKSEFVV PDLELPSWLTGNYRIESVLSSSGKRLGCIKIAASLKI	GM2A

10

20

30

40

50

274	<p>MLFKLLQRQTYTCLSHRYGLYVCFLGVVVTVSAFQFGEVVLEWSRDQY  HVLFDSDYRDNIAGKSFQNRCLCLPMPIDVVYTWVNGTDLELLKELQQVRE  QMEEEEQKAMREILGKNTTEPTKKSEKQLECLLTHCIKVPMLVLDPALPAN  ITLKDLPSPSYPSFHSASDIFNVAKPKNPSTNVSVVVFDDSTKDVEDAHSGLL  KGNSRQTVWRGYLTTDKEVPGLVLMQDLAFLSGFPPTFKETNQLKTKLP  ENLSSKVLLQLYSEASVALLKLNNPKDFQELNKQTKKNMTIDGKELTIS  PA  YLLWDLAISQSKQDEDISASRFEDNEELRYSLSIERHAPWVRNIFIVTNG  QIPSWLNLNPNRVTVTHQDVFRNLSHLPTFSSPAIESHIHRIEGLSQKFIYL  NDDVMFGKDVWPDDFYSHSKGQKVYLTWPVPNCAEGCPGSWIKDGYC  DKACNNSACDWDGGDCSGNSGGSRYIAGGGGTGSIGVGPWPQFGGGINS  VSYCNQGCANSWLADKFCQACNVLSCGFDAGDCGQDHFHELKYLKIVILL  PNQTHYIIPKGECLPYFSFAEVAKRGVEGAYSNDPIIRHASIANKWKTIIHLI  MHSGMNATTIHFNLTFQNTNDEEFKMQITVEVDTREGPKLNSTAQKGYE  NLVSPITLLPEAEILFEDIPKEKRFKPKFRHDVNSTRAQEEVKIPLVNISSL  PKDAQLSLNTLDLQLEHGDITLKGYNLSKSALLRSFLMNSQHAKIKNQAIH  TDETNSDLVAPQEKQVHKSILPNSLGVSERLQRLTFPAVSVKVNNGHDQGG  NPPLDLETTARFRVETHTQKTIGGNVTKEKPPSLIVPLESQMTKEKKITGK  EKENSRMEENAENHIGVTEVLLGRKLQHYTDSYLGFPLWEKKKYFQDLL  DEEESLKTQLAYFTDSKNTGRQLKDTFADSLRYVNKILNSKFGFTSRKVP  AHMPHMIDRIVMQELQDMFPEEFDKTSFHKVRHSEDMQFAFSYFYLLMS  AVQPLNISQVFDEVDTDQSGVLSDREIRTLATRIHELPLSLQDLTGLEHMLI  NCSKMLPADITQLNNIPPTQESYYDPNLPPVTKSLVTNCKPVTDKIHKAYK  DKNKYRFEIMGEEIEAFKMIRTNVSHVVGQLDDIRKNPRKFVCLNDNIDH  NHKDAQTVKAVLRDFYESMFPIPSQFELPREYRNRFLHMHELQEWRAIR  DKLKFWTHCVLATLIMFTIFSFFAEQLIALKRKIFPRRRIHKEASPNRIRV</p>	GNPTAB
275	<p>MAAGLARLLLLLGLSAGGPAPAGAAKMKVVEEPNAFGVNNPFLPQASRL  QAKRDPSPVSGPVHLFRLSGKCFSLVESTYKYEFCPFHNVTQHEQTFRWN  AYSGILGIWHEWEIANNTFTGMWMRDGDACRSRSRQSKVELACGKSNRL  AHVSEPSTCVYALTFETPLVCHPHALLVYPTLPEALQRQWDQVEQDLAD  ELITPQGHEKLLRTLFEADAGYLKTPENEPTQLEGGPDSLGFETLENCRKA  HKELSKEIKRLKGLLTQHGIPTTRPTETSNEHLGHETPRAKSPEQLRGDP  G  LRGSL</p>	GNPTG

10

20

30

40

50

276	<p>MRLPLAPGRLRRGSPRHLPS CSPALLLVGGCLGVFGVAAGTRRPNVV  LLLTDQDEVLGGMTPLKKTALIGEMGMTFSSAYVPSALCCPSRASILT  GKYPHNHVVNNTLEGNCSSKSWQKIQEPNTFPAILRSMCGYQTFAGK  YLNEYGAPDAGGLEHVPLGWSYWYALEKNSKYNYTLSINGKARKHGE  NYSVDYLTVDLANVSLDFLDYKSNFEPFFMMIATPAPHSPWTAAPQYQK  AFQNVFAPRNKNFNHGTNKHVLRQAKTPMTNSSIQFLDNAFRKRWQT  LLSVD  DLVEKLVKRLEFTGELNNTYIFYTSDNGYHTGQFSLPIDKRQLYEFDIKVP  LLVRGPGIKPNQTSKMLVANIDLGPTILDIAGYDLNKTQMDGMSLLPILRG  ASNLTWRSVDLVEYQGEGRNVTDPCTPSLSPGVSQCFPDCVEDAYNNT  YACVRTMSALWNLQYCEFDQEVFVEVYNLTADPDQITNIAKTIDPELLG  KMNYRLMMLQSCSGPTCRTPGVDFPGYRFDPRMLMFSNRGSVRTRRFKSH  LL</p>	GNS
277	<p>MWTLVSWVALTAGLVAGTRCPDGQFCPVACCLDPGGASYSCCRPLLDK  WPTTLRHLGGPCQVDAHCSAGHSCIFTVSGTSSCCPFPEAVACGDGHHHC  CPRGFHCSADGRSCFQSRGNNSVGAIQCPDSQFECPDFSTCCVMVDGSWG  CCPMPQASCCEDRVHCCPHGAFCDLVHTRCITPTGTHPLAKKLPARTNR  AVALSSVMCPDARSRCPDGSTCCELPKGKYGCCPMPNATCCSDHLHCCP  QDTVCDLIQSKCLSKENATDILLTKLPAHTVGDVKCDMEVSCPDPGYTCC  RLQSGAWGCCPFTQAVCCEDHIHCCPAGFTCDTQKGTCEQGPVQVPM  EKAPAHLSLPDPQALKRDVPCDNVSSCPSSDTCCQLTSGEWGCCPIPEAV  CCSDHQHCCPQGYTCVAEGQCQRGSEIVAGLEKMPARRASLSHPRDIGC  DQHTSCPVGQTCPSLGGSWACCQLPHAVCCEDRQHCCPAGYTCNVKA  RSCEKEVSAQPATFLARSPHVGVKDVCEGEGHFCHDNQTCRDNRQG  WACCPYRQGVCCADRRHCCPAGFRCAARGTKCLRREAPRWDAPLRDPA  LRQLL</p>	GRN
278	<p>MARSAVAWAALGPLLWGCALGLQGGMLYPQESPSRECKELDGLWSFR  ADFSNRRRGFEEQWYRRPLWESGPTVDMVPSSFNDISQDWRLRHFVG  WVWYEREVILPERWTQDLRTRVVLRIKSAHSYAIVWVNGVDTLEHEGGY  LPFEADISNLVQVGPLPSRLRITIAINNTLTPTTLPPGTIQYLTDTSKYPKGY  FVQNTYFDFNYAGLQRSVLLYTTPTTYIDDITVTTTSVEQDGLVNYQISV  KGSNLFKLEVRLDAENKVVANGTGTQGLKVPVGVSLWVWYLMHERPA  YL  YSLEVQLTAQTSGLPVSDFYTLPVGIRTVAVTKSQFLINGKPFYFHGVNKH  EDADIRGKGFDPVLLVKDFNLLRWLGANAFRTSHYPYAEVVMQCDRY  GIVVIDECPVGLALPQFFNNSLHHHMVMEEVRRDKNHPAVVMWS  VANEPASHLESAGYKLVIAHTKSLDPSRPVTFVSNSNYAADKGAPYV  DVICLNSYYSWYHDYGHLELIQLATQFENWYKQKPIQSEYGAETI  AGFHQDPPLMFTEEYQKSLLEQYHLGLDQKRRKYVVGELIWNFADFMT  QSPTRVLGNKKGIFTRQRPKSAFLLRERYWKIANETRYPHSVAKSQCL  ENSLFT</p>	GUSB

10

20

30

40

50

279	<p>MTSSRLWFSLLAAAFAGRATALWPWPQNFQTSQRYVLYPNNFQFQY  DVSSAAQPGCSVLDEAFQRYRDLDFGSGSWPRPYLTGKRHTLEKNVLVV  SVVTPGCNQLPTLESVENYTLTINDDQCLLSETVWGALRGLETFSQLVW  KSAEGTFFINKTEIEDFPRFPHRGLLLDTSRHYLPLSSILDTLDVMAYNKLN  VFHWHLVDDPSFPYESFTPELMRKGSYNPVTHIYTAQDVKEVIEYARLR  GIRVLAEFDTPGHTLSWGPPIPGLLTPCYSGSEPSGTFGPVNPSLNNTYEF  MSTFFLEVSSVFPDFYLHLGGDEVDFTCWKSNIQDFMRKKGFGEDFKQ  LESFYIQTLTLDIVSSYGKGYVWVQEVFDNKVKIQPDTIIQVWREDIPVNYM  KELELVTKAGFRALLSAPWYLNRISSYGPDWKDFYIVEPLAFEGTPEQKAL  VIGGEACMWGEYVDNTNLVPRLWPRAGAVAERLWSNKLTSDLTFAYER  LSHFRCELLRRGVQAQPLNVGFCEQEFEQT</p>	HEXA
280	<p>MELCGLGLPRPPMLLALLLATLLAAMLALLTQVALVVQVAEAAARAPSVS  AKPGPALWPLPLSVKMTPNLLHLAPENFYISHSPNSTAGPSCTLLEEAFRR  YHGYIFGFYKWHHEPAEFQAKTQVQQLLVSITLQSECDAFPNISSDESYTL  LVKEPVAVLKANRVWGALRGLETFSQLVYQDSYGTFTINESTIIDSPRFSH  RGILIDTSRHYLPVKIILKTLDAMAFNKFVNLHWHIVDDQSFYQSIITPEL  SNKGSYSLSHVYTPNDVRMVIEYARLRGIRVLPFDTPGHTLSWKGKQKD  LLTPCYSRQNKLDSFGPINPTLNTTYSFLTTFKKEISEVFPDQFIHLGGDEVE  FKCWESNPKIQDFMRQKGFGTDFKKLESFYIQKVLDIATINKGSIVWQEV  FDDKAKLAPGTIVEVWKDSAYPEELSRVTASGFPVILSAPWYLDLISYGG  DWRKYYKVEPLDFGGTQKQKQLFIGGEACLWGEYVDA TNLTPRLWPR  SAVGERLWSSKDVRDMDDAYDRLTRHRCRMVERG  IAAQPLYAGYCNHENM</p>	HEXB
281	<p>MTGARASAAEQRRAGRSGQARAERAAGMSGAGRALAALLLAASVLSA  ALLAPGGSSGRDAQAAPRDLDKRRHAELKMDQALLIHNELLWTNLT  YWKSECCYHCLFQVLVNVPQSPKAGKPSAAAASVSTQHGSILQLNDTLEE  KEVCRLEYRFGFEGNYSLLVKNIHNGVSEIACDLAVNEDPVDSNLPVSI  LIGLAVIIVISFLRLLLSLDDFNNWISKAISSRETDRLINSELGSPSRDPLDG  DVQPATWRLSALPPRLRSVDTFRGIALILMVFNNGGGKYWYFKHASWN  GLTVADLVFPWFVIMGSSIFLSMTSILQRCGSKFRLLGKIAWRSFLICIGI  IIVNPNYCLGPLSWDKVRIPGVLQRLGVTYFVAVLELLFAKPVPEHCASE  RSCLSLRDITSSWPQWLLILVLEGLWLGLTFLLPVPGCPTGYLGGGIGDF  GKYPNCTGGAAGYIDRLLLGDHLYQHPSSAVLYHTEVAYDPEGILGTIN  SIVMAFLGVQAGKILLYKARTKDILIRFTA WCC  ILGLISVALTKVSENEGFI PVNKNLWLSYVTTLSSFAFFILLVLYPVVDVK  GLWTGTPFFYPGMNSILVYVGHEVFENYFPFQWKLKDNQSHKEHLTQNI  VATALWVLIAYILYRKKIFWKI</p>	HGSNAT

10

20

30

40

50

282	MAAHLLPICALFLTLLDMAQQFRGPLLPNRPFTTVWNANTQWCLERHGV DVDVSVFDVVANPGQTFRGPDMTIFYSSQLGTYPYTPTGEPVFGGLPQN ASLIAHLARTFQDILAAIPAPDFSGLAVIDWEAWRPRWAFNWDTKDIYRQ RSRALVQAQHPDWPAPQVEAVAQDQFQGAARAWMAGTLQLGRALRPR GLWGFYGFPCYNYDFLSPNYTGQCPSGIRAQNDQLGWLWGQSRALYPS IYMPAVLEGTGKSQMYVQHRVAEAFRVAVAAGDPNLPVLPYVQIFYDTT NHFLPLDELEHSLGESAAQGAAGVVLWVSWENTRTKESCQAIKEYMDTT LGPFILNVTSGALLCSQ ALCSGHGRCVVRTSHPKALLLNPAFSIQLTPGGGPLSLRGALSLEDQAQ MAVEFKCRCYPGWQAPWCERKSMW	HYAL1
283	MPPPRTGRGLLWLGLVLSSVCVALGSETQANSTTDALNVLLIIVDDLRLPSL GCYGDKLVRSPNIDQLASHSLFQNAFAQAVCAPSRVSFLTGRRPDTR LYDFNSYWRVHAGNFSTIPQYFKENGYVTMSVGKVFHPGISSNHTDDSPY SWSFPPYHPSSEKYENTKTCRGPDGELHANLLCPVDVLDVPEGTLPKQS TEQAIQLLEKMKTSASPFFLAVGYHKPHIPFRYPKEFQKLYPLENITLAPDP EVPDGLPPVAYNPWMDIRQREDVQALNISVPYGPVDFQRKIRQSYFASV SYLDTQVGRLLSALDDLQLANSTIIAFTSDHGWALGEHGEWAKYSNFDV ATHVPLIFYVPGRTASLPEAGEKLPYLDPFDSASQLMEPGRQSMDELVEL VSLFPTLAGLAGLQVPPRCVPSFHVELCREGKNLLKHFRFRDLEEDPYLP GNPRELIAYSQYPRPSDIPQWNSDKPSLKDIIKIMGYSIRTIDYRYTVWVGF NPDEFLANFSDIHAGELYFVDSPLQDHNMYNDSQGGDLFQLLMP	IDS
284	MRPLRPRAALLALLASLLAAPPVAPAEAPHLVHVDAARALWPLRRFWRS TGFCPPLPHSQADQYVLSWDQQLNLAYVGAVPHRGIKQVRTHWLELVT TRGSTGRGLSYNFTHLDGYLDLLRENQLLPGFELMGSASGHFTDFEDKQQ VFEWKDLVSSLARRYIGRYGLAHVSKWNFETWNEPDHHDNDVNSMTMQ GFLNYYDACSEGLRAASPALRLGGPGDSFHTPPRSPLSWGLLRHCHDGTN FFTGEAGVRLDYISLHRKGARSSISILEQEKVVAQQIRQLFPKFADTPIYND EADPLVGWSLPQPWRADVYAAMVVKVIAQHQNLLANTTSAPYALLS NDNAFLSYHPPFAQRTLARFQVNNTRPPHVQLLRKPVLAMGLLALL DEEQLWAEVSQAGTVLDSNHTVGLASAHQPQGPADAWRAAVLIYASD DTRAHPNRSVAVTLRLRGVPPGGLVYVTRYLDNGLCSPDGEWRRLGRP VFPTAEQFRMRRAAEDPVAAPRPLPAGGRLTLRPAALRLPSLLLHVHVCAR PEKPPGQVTRLRALPLTQGQLVLVWSDEHVGSKCLWYIEIQFSQDGKAY TPVSRKPSTFNLVFPDGTGAVSGSYRVRALDYWARPGPFSDPVPYLEVP VPRGPPSPGNP	IDUA
285	MVVVTGREPDSRRQDGAMSSSDAEDDFLEPATPTATQAGHALPLLQEFEP EVVPLNIGGAHFTTRLSTLRCEYEDTMLAAMFSGRHYIPTDSEGRYFIDRDG THFGDVLNFLRSGDLPPRERVRAVYKEAQYYAIGPLLEQLENMQPLKGE KVRQAFLGLMPYYKDHLEIRIVEIARLRAVQRKARFAKLVKCVFKEEMPIT PYECPLLSLRFERSESDGQLFEHHCEVDVSFGPWEAVADVYDLLHCLVT DLSAQGLTVDHQCIGVCDKHLVNHYECKRPIYEFKITWW	KCTD7

10

20

30

40

286	MVCFRLFPVPGSGLVLVCLVLGAVRSYALELNLTDSENATCLYAKWQM NFTVRYETTNTKTYKTVTISDHGTVTYNGSICGDDQNGPKIAVQFGPGFSW IANFTKAASTYSIDSVSFSYNTGDNTTFPDAEDKGILTVDELLAIRIPLNDF RCNSLSTLEKNDVVQHYWDVLVQAFVQNGTVSTNEFLCDKDKTSTVAPT IHTTVPSPTTTPKPEKPEAGTYSVNNNGNDTCLLATMGLQLNITQDKVAS VININPNTTHSTGSCRSHALLRLNSSTIKYLDFVFAVKNNRFYLKEVNIS MYLVNGSVFSIANNLSYWDAPLGSYMCNKEQTVSVSGAFQINTFDLR VQPFNVTQGKYSTAQDCSADDDNFLVPIAVGAALAGVLILVLLAYFIGLK HHHAGYEQF	LAMP2
287	MGAYARASGVCARGCLDSAGPWTMSRALRPPLPPLCFFLLLLAAAGARA GGYETCPTVQPMLNVHLLPHTHDDVGVWLTVDQYFYGIKNDIQHAGV QYILDSVISALLADPTRRFIYVEIAFFSRWWHQQTNATQEVVRDLVRQGR LEFANGGWVMNDEAATHYGAIVDQMTLGLRFLDFTFGNDGRPRVAWHI DPFGHSREQASLFAQMFGDFGFFGRLDYQDKWVRMQKLEMEQVWRAST SLKPPTADLFTGVLPNGYNPPRNLCDVLCVDQPLVEDPRSPEYNAKELV DYFLNVATAQGRYYRTNHTVMTMGSDFYENANMWFKNLDKLRILVN AQQAKGSSVHVLYSTPACYLWELNKANLTSVVKHDDFFPYADGPHQFW TGYFSSRPALKRYERLSYNFLQVCNQLEALVGLAANVGPYSGSDSAPLNE AMAVLQHHDVAVSGTSRQHVANDYARQLAAGWGPCEVLLSNALARLRG FKDHFTFCQQLNISICPLSQTAAARFQVIVYNPLGRKVNWMVRLPVSEGVF VVKDPNGRTVPSDVVIFPSSDSQAHPPPELLFSASLPALGFSTYSVAQVPRW KPQARAPQPIPRRSWSPALTIENEHIRATFDPDTGLLMEIMNMNQQLLPV RQTFFWYNASIGDNESDQASGAYIFRPNQKPLPVSRAWAQLVKTPLVQ EVHQNFSAWCSQVVRLYPGQRHLELEWSVGPIPVGDTWGKEVISRFDTPL ETKGRFYTDSNGREILERRRDYRPTWKLNQTEPVAGNYYPVNTRIYITDG NMQLTVLTDRSQGGSSLRDGSLELMVHRRLKDDGRGVSEPLMENGSGA WVRGRHLVLLDTAQAAAAGHRLLAEQEVLAPQVVLAPGGGAAYNLGAP PRTQFSGLRDLPPSVHLLTLASWGPPEMVLLRLEHQFAVGEDSGRNLSAP VTLNLRDLFSTFTITRLQETTLVANQLREAASRLKWTNTGTTPHQTPYQL DPANITLPEMEIRTFILASVQWKEVDG	MAN2B1

10

20

30

40

50

288	MRLHLLLLLALCGAGTTAAELSYSLRGNWSICNGSLELPGAVPGCVHS ALFQQGLIQDSYRFNDLNWRVSLDNWTYSKEFKIPFEISKWQVNLIL EGVDTVSKILFNEVTIGETDNMNFNRYSDITNVVRDVNSIELRFQSAVLYA AQQSKAHTRYQVPPDCPPLVQKGECHVNFVRKEQCSFSWDWGPSFPTQG IWKDVRIEAYNICHNLNYFTFSPIYDKSAQEWNLIEIESTFDVVSSKPVGGQVI VAIPKLQTQQTYSIELQPGKRIVELFVNISKNITVETWWPHGHGNQTGYN MTVLFELDGGLNIEKSAKVYFRTVELIEEPIKGGSPGLSFYFKINGFPIFLKGS NWIPADSFQDRVTSELLRLLLQSVVDANMNTLRVWGGGIYEQDEFYELC DELGIMVWQDFMFACALYPTDQGFLLDSVTAEVAYQIKRLKSHPSIIIWSG NNENEEALMMNWHYHISFTDRPIYIKDYVTLVYVKNIRELVLGDKSRPFITS SPTNGAETVAEAWVSQNPNSNYFGDVHFYDYISDC WNWKFVFPKARFASEYGYQSWPSFSTLEKVSSTEDWSFNKFSLHRQHHE GGNKQMLYQAGLHFKLPQSTDPLRFTKDIYLTQVMQAQCVKTEFETFYR RSRSEIVDQQGHTMGALYWQLNDIWQAPSWASLEYGGKWKMLHYFAQ NFFAPLLPVGFENENTFYIYGVSDDLHSDYSMTLSVRVHTWSSLEPVCSRVT ERFVMKGGAEVCLYEVPSELLRRCGNCTRESCVVSFYLSADHELLSPTN YHFLSSPKEAVGLCKAQITAIISQQGDIFVFDLETSAPFVWLDVGSIPGR FSDNGFLMTEKTRTILFYPEPTSKNELEQSFHVTSITDIY	MANBA
289	MTAPAGPRGSETERLLTPNPGYGTQAGPSPAPPTPPEEEDLRRRLKYFFMS PCDKFRAKGRKPKLMLQVVKILVTVQLILFGLSNQLAVTFREENTIAFR HLFLLGYSDGADDTFAAYTREQLYQAIFHAVDQYLALPDVSLGRYAYVR GGGDPWTNGSGLALCQRYHRGHVDPANDTFDIDPMVVTDCIQVDPPER PPPPSDDLTLLESSSYKNLTLKFHKLNVNVIHFRLKTINLQSLINNEIPDC YTFSVLITFDNKAHSGRIPISLETQAHIQECHKPSVFQHGDNFRLLFDVVV ILTCSLFLLCARSLLRGFLQNEFVGMWRQRGRVISLWERLEFVNGWYI LLVTSVLTISGTIMKIGIEAKNLASYDVCSILLGTSTLLVWVGVIYRITFF HNYNIIATLRVALPSVMRFFFFAVIYLYGFCGWIWVLPYHVKFRSLS MVSECLFSLINGDDMFVTFAMQAQQGRSSLVWLFSQLYLYSFISLFIYM VLSLFIALITGAYDTIKHPGGAGAEESLQAYIAQCQDSPTSQKFRRGSGS ACSLCCCGRDPSEEHSLLVN	MCOLN1
290	MAGLRNESEQEPLLGDTPGSREWDILETEEHYKSRWRSIRILYLTMFLSSV GFSVVMMSIWPYLQKIDPTADTSFLGWVIASYSLGQMVASPIFGLWSNYR PRKEPLIVSILISVAANCLYAYLHIPASHNKYYMLVARGLLGIGAGNVAVV RSYTAGATSLQERTSSMANISMCQALGFILGPVFQTCFTFLGKGVTDVVI KLQINMYTTPVLLSAFLGILNIIILAILREHRVDDS GRQCKSINFEEASTDEAQVPQGNIDQVAVVAINVLFVTLFIFALFETIITPL TMDMYAWTQEQAFLYNGIILAALGVEAVVIFLGVKLLSKKIGERAILLGG LIVVWVGFILLPWGNQFPKIQWEDLHNSIPNTTFGEIIIWLVKSPMEDD NERPTGCSIEQAWCLYTPVIHLAQFLTSVAVLIGLGYPCNLMSTLYSKIL GPKPQGVYMGWLTASGSGARILGPMFISQVYAHWGPRWAFSLVCGIIVL TITLLGVVYKRLIALSVRYGRIQE	MFSD8

10

20

30

40

50



291	<p>MLLKTVLLLGHVAQVLMMLDNGLLQTPPMGWLAWERFRCNINCDEDPKN  CISEQLFMEMADRMAQDGRDMGYTYLNIDDCWIGGRDASGRMLPDPK  RFPHGIPFLADYVHSLGLKLGiyADMGNFTCMGYPGTTLDKVVQDAQTF  AEWKVDMLKLDGCFSTPEERAQGYPKMAAALNATGRPIAFSCSWPAYE  GGLPPRVNYSLLADICNLWRNYDDIQDSWWSVLSILNWFVEHQDILQPV  AGPGHWNDPDMMLLIGNFGLSLEQSRQMALWTVLAAPLLMSTDLRTISA  QNMDILQNPLMIKINQDPLGIQGRRIHKEKSLIEVYMRPLSNKASALVFFS  CRTDMPYRYHSSLGQLNFTGSVIYEAQDVYSGDIISGLRDETNFTVIINPSG  VVMWYLYPIKNLEMSQQ</p>	NAGA
292	<p>MEAVAVAAAVGVLLLAGAGGAAGDEAREAAAVRALVARLLGPGPAAD  FSVVERALAAKPLDTSYSLGGGGAARVRVRGSTGVAAAAGLHRYLRDF  CGCHVAWSGSQRLRPLPAVPGELTEATPNRYRYQNVCTQSYSFVWW  DWARWEREIDWMALNGINLALAWSGQEAIWQRVYLALGLTQAEINEFFT  GPAFLAWGRMGNLHTWDGPLPSSWHIKQLYLQHRVLDQMRSFGMTPVL  PAFAGHVPEAVTRVFPQVNVTKMGSWGHFNCSYSCSFLAPEDPIFPIIGS  LFLRELIKEFGTDHIYGADTFNEMQPPSSEPSYLAATTAVYEAMTAVDT  EAVWLLQGWLQFHQPQF  WGPAQIRAVLGA VPRGRLLVLDLFAESQPVYTRTASFQGPFIWCMLHNF  GGNHGLFGALEAVNGGPEAARLFPNSTMVGTMapeGISQNEVVYSLMA  ELGWRKDPVPLA AAVVTSFAARRYGVSHPDAGAAWRLLLRSVYNCSGE  ACRGHNRSPLVRRPSLQMNTSIWYNRSVFEAWRLLLTSAPSLATSPAFR  YDLLDLTRQAVQELVSLYEEARSAYLSKELASLLRAGGVLAYELLPALD  EVLASDSRFLGSWLEQARAAA VSEAEADFYEQNSRYQLTLWGPEGNIL  DYANKQLAGLVANYTTPRWRLFLEALVDSVAQGIPFQQHQFDKNVFQLE  QAFVLSKQRYPSQPRGDTVDLAKKIFLKYPRWVAGSW</p>	NAGLU
293	<p>MTGERPSTALPDRRWGPRILGFWGGCRVWVFAAIFLLLSLAASWSKAEN  DFGLVQPLVTMEQLLWVSGRQIGSVDTFRIPLITATPRGTLFAAEARKMS  SSDEGAKFIALRRSMDQGSTWSPTAFIVNDGDVPDGLNLGAVVSDVETG  VVFLFYSLCAHKAGCQVASTMLVWSKDDGVSWSWSTPRNLSLDIGTEVFAP  GPGSGIQKQREPRKGRLIVCGHGTlerDGVFCLLSDDHGASWRYGSGVS  GIPYGQPKQENDFNpDECQPYELPDGSVVINARNQNNYHCHCRIVLRSYD  ACDTLRPRDVTDFPELVDPVVAAGAVVTSSGIVFFSNPAHPEFRVNLTLR  WSFSNGTSWRKET  VQLWPGPSGYSSLATLEGSMDGEEQAPQLYVLYEKGRNHYTESISVAKIS  VYGTL</p>	NEU1

10

20

30

40

50

294	<p>MTARGLALGLLLLLCPAQVFSQSCVWYGECGAIYGDKRYNCEYSGPPK  PLPKDGYDLVQELCPGFFFGNVSLCCDVRQLQTLKDNLQLPLQFLSRCPS  CFYNLLNLFCELTCSQRSQFLNVTATEDYVDPVTNQTKTNVKELQYYV  GQSFANAMYNACRDVEAPSSNDKALGLLCGKDADACNATNWIEYMFNK  DNGQAPFTITPVFSDFPVHGMEPMNNATKGCDESVDEVTAPCSCQDCSIV  CGPKPQPPPPAPWTILGLDAMYVIMWITYMAFLLVFFGAFFAVWCYRK  RYFVSEYTPIDSNIAFSVNASDKGEASCCDPVSAAFEGCLRRLFTRWGSFC  VRNPGCVIFFSLVFITACSSGLVFRVTTNPVDLWSAPSSQARLEKEYFDQ  HFGPFRTEQLIIRAPLTDKHIYQPYPSGADVFPFGPPLDIQILHQVLDLQIAI  ENITASYDNETVTLQDICLAPLSPYNTNCTILSVLNYFQNSHSLDHHKGD  DFFVYADYHHTFLYCVRAPASLNDTSLLDHDPCLGTGGPVPFWLVLGGY  DDQNYNNATALVITFPVNNYYNDTEKLQRAQAWKEFINFVKNYKNPNL  TISFTAERSIEDELNRESDDVFTVVISYAIMFLYISLALGHMKSCRLLVD  SKVSLGIAGILIVLSSVACSLGVFSYIGLPLTLIVIEVIPFLVLAVGVDNIFILV  QAYQRDERLQGETLDQQLGRVLGEVAPSMFLSSFSETVAFFLGALSVMVA  VHTFSLFAGLAVFIDFLLQITCFV  SLLGLDIKRQEKNRLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPLLLK  DWMRPIVIAIFVGLSFSIAVLNKVDIGLDQSLSMPPDDSYMVDYFKSISQY  LHAGPPVYFVLEEGHDYTSSKGQNMVCGGMGCNNDLQVQIFNAAQLD  NYTRIGFAPSSWIDDYFDWVKPQSSCCRVNDITDQFCNASVVDPAVRCR  PLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGGHAAAYSSAVNILLGH  GTRVGATYFMTYHTVLQTSADFIDALKKARLIASNVTETMGINGSAYRVF  PYSVFYVFYEQYLTIIDDTIFNLGVSGLAIFLVTMVLGCELWSAVIMCATI  AMVLVNMFGVMWLWGISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMK  GSRVERAEEALAHMGSSVFSGITLTKFGGIVVLAFAKSQIFQIFYFRMYLA  MVLLGATHGLIFLPVLLSYIGPSVNKAKSCATEERYKGTERRLLNF</p>	NPC1
295	<p>MRFLAATFLLLALSTAAQAEPVQFKDCGSVDGVIKEVNVSPCPTQPCQLS  KGQSYSVNVTFTSNIQSKSSKAVVHGILMGVVPVFPPIPEPDGCKSGINCP  KDKTYSYLNKLPVKSEYPSIKLVVEWQLQDDKNQSLFCWEIPVQIVSHL</p>	NPC2
296	<p>MSCPVPACCALLLVLGLCRARPRNALLLADDGGFESGAYNNSAIATPHL  DALARRSLLFRNAFTSVSSCSPSRASLLTGLPQHONGMYGLHQDVHHFNS  FDKVRSLPLLSQAGVRTGIIGKKHVGPETVYPDFAYTEENGSVLQVGR  NITRIKLLVRKFLQTQDDRPFFLYVAFHDPHRCGHSQPQYGTFCFKFGNG  ESGMGRIPDWTPQAYDPLDVLVPYFVPNTPAARADLAAQYTTVGRMDQ  GVGLVLQELRDAGVLNDTLVIFTSDNGIPFSGRTNLYWPGTAEPLLVSSP  E  HPKRWGQVSEAYVSLDLTPTILDWFSIPYPSYAIFGSKTIHLTGRSLLPAL  EAEPLWATVFGSQSHHEVTMSYPMRSVQHRHFRLVHNLNFKMPFPIDQD  FYVSPTFQDLLNRRTAGQPTGWYKDLRHYYRARWELYDRSRDPHETQ  NLATDPRFAQLLEMLRDQLAKWQWETHDPWVCAPDGVLEEKLSPQCQP  LHNEL</p>	SGSH

10

20

30

40

50

297	MASPGCLWLLAVALLPWTCASRALQHLDPPAPLPLVIWHGMGDSCCNPL SMGAIKKMVEKKIPGIYVLSLEIGKTLMEDVENSFFLNVNSQVTTVCQAL AKDPKLQQGYNAMGFSQGGQFLRAVAQRCPSPPMINLISVGGQHOGVFG LPRCPGESSHICDFIRKTLNAGAYSKVVQERLVQA EYWH D PIKEDVYRNH SIFLADINQERGINESYKKNLMALKKFVMVKFLNDSIVDPVDSEWFGFYR SGQAKETIPLQETS LYTQDRLGLKEMDNAGQLVFLATEGDHLQLSEEW YAHIIIFLG	PPT1
298	MYALFLLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHC LQTVWNKPTVKSLPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWL PKPNMSASCKEIVDSYLPVILDIK GEMSRPGEVCSALNLCE SLQKHLAEL NHQKQLESNKIPELDMTEVVAPFMANIPLLLYPQDGPRSKPQPKDNGDVC QDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISQY SEIAIQMMM HMQPKEICALVGFCEVKEMPMQTLVPAKVASKNVIPALE LVEPIKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMCS KLPKSLSEECQEVVD TYGSSILSILLEEV SPELVCSMLHLCSGTRLPALTVH VTQPKDGGFCEVCKKLVGYLDRNLEKNSTKQEILAALEK GCSFLPDPYQK QCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAHKPLLGT E KCIWGPS YWCQNTETAAQCN AVEHCKRHVWN	PSAP
299	MRSPVRDLARNDGEESTDRTPLLPGAPRAEAAPVCCSARYNLAILAFFGF FIVYALRVNLSVALVDMVDSNTTLEDNRTSKACPEHSAPIKVHHNQTKG KYQWDAETQG WILGSFFYGYIITQIPGGYVASKIGGKMLLGF GILGTAVLT LFTPIAADLGVGPLIVLRALEGLGEGVTFPAMHAMWSSWAPPLERSKLLSI SYAGAQLGTVISLPLSGIICYMNW TYVFYFFGTIGIFWFLW IWLVS DTP QHKKRISHYEKEYILSSLRNQLSSQKSVWPVILKSLPLWAI VVAHFSYNW TFYTLLTLLPTYMKEILRFNVQENGLSSLPYLGSWLCMILSGQAADNLR AKWNFSTLCVRRIFSLIGMIGPAVFLVAAGFIGCDYSLAVAFLTISTTLGGF CSSGFSINHLDIAPSYAGILLGITNTFATIPGMVGPVIAKSLTPDNTVGEWQ TVFYIAAAINVFGAIFFTLFAKGEVQNWALNDHHGHRH	SLC17A5
300	MPRYGASLRQSCPRSGREQGDGTAGAPGLLWMGLVLALALALALALA LSDSRVLWAPAEAHPLSPQGH PARLHRIVPRLRDVFGWGNLTC PICKGLF TAINLGLKKEPNVARVGSVAIKLCNLLKIAPPAVCQSIVHLFEDDMVEVW RRSVLSPSEACGLLLGSTCGHWDIFSSWNISLPTVPKPPP KPPSPAPGAPV SRILFLDHLWDHDYLEGTDPCADPLCCRRGSGLPPASRPGAGYWGEYS KCDLPLRTLESLLSGLGPA GPFDMVYWTGDIPAHDVWHQTRQDQLRAL T TVTALVRKFLGPVPVYPAVGNH ESTPVNSFPPPFIEGNHSSRWLYEAMAK AWEPWLPAEALRTL RIGGFYALSPYPGLRLISLNMNFC SRENFWLLINSTD PAGQLQWL VGELQAAEDRGDKVHIIGHIPP GHCLKSWSWNYRIVARYE NTLAAQFFGH THVDEFEVFYDEETLSRPLAVAF LAPSATTYIGLNPGYRV YQIDGNYS GSSHVLDHETYILNLTQANIPGAIPH WQLLYRARETYGLPN TLPTAWHNLVYRMRGDMQLFQTFWFLYHKGHPPSEPCGTPCRLATLCAQ LSARADSPALCRHLM PDGSLPEAQSLWPRPLFC	SMPD1

10

20

30

40

50

301	MAAPALGLVCGRCPELGLVLLLLLLSLLCGAAGSQEAGTGAGAGSLAGS CGCGTPQRPGAHGSSAAAHRYREANAPGPVPGERQLAHSKMVPIAGV FTMGTDDPQIKQDGEAPARRVTIDAFYMDAYEVSNTFEFEKFNSTGYLTE AEKFGDSFVFEGMLSEQVKTNIQQA VAAAPWWLPVKGANWRHPEGPDS TILHRPDHPVLHVSWNDAVAYCTWAGKRLPTEAEWEYSCRGLHNRLFP WGNKLQPKGQHYANIWQGEFPVTNTGEDGFQGTAPVDAFPNGYGLYNI VGNAWEWTSWWTVHHSVEETLNPKGPPSGKDRVKKGGSYMCHRSYC YRRCARSQNTPDSSASN LGFRCAADRLPTMD	SUMF1
302	MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTF ALRQQNVERLSELVQAVSDPSSPQYGKYLLENVADLVRPSPLTLHTVQK WLLAAGA QKCHSVITQDFLTCWLSIRQAELLPGAEFHYYVGGPTETHV VRSPHPYQLPQALAPHVDFVGG LHRFPPTSSLRQRPEPQVTGTVGLHLGV TPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHSDLAQFMRLFGGN FAHQASVARVVGQQRGRAGIEASLDVQYLM SAGANISTWVYSSPGRHE G QEPFLQWLMLLSNESALPHVHTVSYGDD EDSLSSAYIQRVNTELMKAAA RGLTLLFASGDSGAGCWSVSGRHQFRPTFPASSPYVTTVGGTSFQEPFLIT NEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSPHLPSSYFNASGRA YPDV AALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSGRPPLGFLN PRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPNF PALLKTLLNP	TPP1
303	MSDKLPYKVADIGLAAWGRKALDIAENEMPGLMRMRERY SASKPLKGA RIAGCLHMTVETAVLIETLVTLGAEVQWSSCNIFSTQDHAAAIAKAGIPV YAWKGETDEEYLWCIEQTL YFKDGPLNMILDDGGDLTNLIHTKYPQLLP GIRGISEETTTGVHNL YKMMANGILKVPAINVND SVTKSKFDNLYGCRES LIDGIKRATDVMIAGKVAVVAGYGDVKGKCAQALRGFGARVIITEIDPIN ALQAAMEGYEVTTMDEACQEGNIFVTTTGCIDIILGRHFEQMKDDAIVCN IG HFDVEIDVKWLNENAVEKVNIKPQVD RYRLKNGRRIILLAEGRVLNLGCA MGHPSFVMSNSFTNQVMAQIELWTHPDKYPVGVHFLPKKLDEAVAEAH LGKLVNKLTKLTEKQAQYLGMSCDGPFKPDHYRY	AHCY
304	MVDSVYRTRSLGVAAEGLPDQYADGEAARVWQLYIGDTRSR TA EYKAW LLGLLRQHGCQRVLDVACGTGVDSIMLVEEGFSVTSVDASDKMLKYALK ERWNRREHAPFDKWWIEEANWMTLDKDVPSAEGGFDAVICLGNSFAH LPDCKGDQSEHRLALKNIASMRVAGGLVIDHRNYDHILSTGCAPP GKNI YYKSDLTKDVTTSVLIVNNKAHMTLDYTVQVPGAGQDGSPGLSKFRLS YYPHCLASFTELLQAAFGGKQCQHSVLGDFKPYKPGQTYIPCYFIHV LKRT D	GNMT

10

20

30

40

50

305	MNGPVDGLCDHSLSEGVMFTSESVGEGHPDKICDQISDAVLDAHLKQDP NAKVACETVCKTGMVLLCGEITSMAMVDYQRVVRDTIKHIGYDDSAKG FDFKTCNVLVALEQQSPDIAQCVHLDRNEEDVGAGDQGLMFGYATDETE ECMPLTII LAHKL NARMADLRRSGLLPWLRPDSKTQVTVQYMQDNGAVI PVRIHTIVISVQHNEEDITLEEMRRALKEQVIRAVVPAKYLDEDTVYHLQPS GRFVIGGPQGDAGVTGRKIIVDTYGGWGAHGGGAFSGKDYTKVDRSAA YAARWVAKSLVKAGLCRRVLVQVSYAIGVAEPLSISIFTYGT SQKTEREL LDVVHKNFDLRPGVIVRDLDLKKPIYQKTACYGHFGRSEFPWEVPRKLVF	MAT1A
306	MEKGPVRAPA EKPRGARCSNGFPERDPPRPGPSRPAEKPPRPEAKSAQPA DGWKG ERPRSEEDNELNLPNLAAA YSSILSSLGENPQRQGLL KTPWRAAS AMQFFTKGYQETISDVLNDAIFDEHDDEMVIK DIDMFSMCEHHLVPFVG KVHIGYLPNKQVLGLSKLARIVEIYSRRLQVQERLTKQIAVAITEALRPAG VGVVVEATHMCMVMRGVQKMNSKTVTSTMLGVFREDPKTREEFLTLIR S	GCH1
307	MAGKAHRLSAEERDQLLPNLRAVGWNELEGRDAIFKQFHFKDFNRAFGE MTRVALQAEKLDHHPWFVFNKVVHITLSTHECAGLSERDINLASFIEQV AVSMT	PCBD1
308	MSTEGGRRRCQAQVSRISFSASHRLYSKFLSDEENLKLFGKCNNPNHG HNYKVVVTVHGEIDPATGMVMNLADLKKYMEEAIMQPLDHNLDMDV PYFADVSTTENVAVYIWDNLQKVLVPGVLYKVKVYETDNNIVVYKGE	PTS
309	MAAAAAAGEARRVLVYGGRGALGSRVCQAFRARNWWWASVDVVENEE ASASIIVKMTDSFTEQADQVTAEVGKLLGEEKVDAILCVAGGWAGGNAK SKSLFKNCDLMWKQSIWTSTISSHLATKHLKEGGLLTLAGAKAALDGT MIGYGMAGKAVHQLCQSLAGKNSGMPPGAAAIAVLPVTLDTMPNRS PEADFSWTPLEFLVETFDHWITGKNRPSSGSLIQVVTTEGRTELTPAYF	QDPR
310	MEGGLGRAVCLLTGASRGFGR TLAPLLASLLSPGSVLVLSARNDEALRQL EAE LGAERSGLRVVRVPADLGA EAGLQQLL GALRELPRPKGLQRLLINN AGSLGDVSKGFVDLSDSTQVNNY WALNLT SMLCLTSSVLKAFPDSPLN RTVVNISSCALQPFKGWALYCAGKAARDMLFQVLALEEPNVRVNLNYAP GPLD TDMQQLARETSVDPDMRKGLQELKAKGKLV DCKVSAQKLLSLE KDEFKSGAHVDFYDK	SPR
311	MDAILNYRSEDTE DY TLLGCDELSSVEQILAEFKVRALECHPKHPENP KAVETFQKLQKAKEILTNEESRARYDHWRRSQMSMPFQQWEALNDSVK TSMHWVVRGK KDLMLEESDKTHTTKMENE ECNEQRERKKEELASTAEK TEQKEPKPLEKSVSPQNSDSSGFADVNGWHLRFRWSKDAPSELLRKF RNYEI	DNAJC12

10

20

30

40

50

312	<p>MLLPAPALRRALLSRPWTGAGLRWKHTSSLKVANEPVLAFTQGS PERDA  LQKALKDLKGRMEAI PCVVGDEEVWTS DVQYQVSPFNHGHKVAKFCYA  DKSLLNKAIEAALAAARKEWDLKPIADRAQIFLKAADMLSGPRRAEILAKT  MVGQGGKTVIQAEIDAAAELIDFFRFNAKYAVELEGQQPISVPPSTNSTVYR  GLEGFVAAISPFNFTAIGGNLAGAPALMGNVVLWKPSDTAMLASYAVYR  ILREAGLPPNIIQFVPADGPLFGDTVTSSEHL CGINFTG SVPTFKHLWKQVA  Q  NLDRFHTFPRLAGECGGKNFHFVHRSADVESV VSGTLRS AF EYGGQKCS  ACSRLYVPHSLWPQIKGRLL EHSRIKVGDP AEDFGTFFSAVIDAKSFARI  KKWLEHARSSPSLTILAGGKCDDSVGYFVEPCIVESKDPQEPIMKEEIFGP  VLSVYVYPDDKYKETLQLVDSTTSYGLTGAVFSQDKDVVQEATKVL RNA  AGNFYINDKSTGSIVGQQPFGGARASGTNDKPGGPHYILRWTSPQVIKET  HKPLGDWSYAYMQ</p>	ALDH4A1
313	<p>MALRRALPALRPCIPRFVQLSTAPASREQPAAGPAAVPGGGSATAVRPPV  PAVDFGNAQEAYRSRRTWELARSLVLRLCAWPALLARHEQLLYVSRKL  LGQRLFNKLMKMTFYGHFVAGEDQESIQLLRHYRAFGVSAILDYGV EEE  DLSPEEAH KEMESCTSA AERDGS GTNKRDKQYQAHRAFGDRRNGVISA  RTYFYANEAKCDSHMETFLRCIEASGRVSDDGFIAIKLTALGRPQFLQFS  EVLAKWRCFFHQMAVEQQAGLAAMDTKLEVAVLQESVAKLGIASRAE  IEDW  FTAETLGVSGTMDLLDWSS LIDSR TKLSKHLVVPNAQTGQLEPLLSRFTEE  EELQMTRMLQRMDVLAKKATEMGMVRLMVD AEQTYFQPAISR LTLEMQR  KFNVEKPLIFNTYQCYLKDAYDNVTL DVELARREGWCFGAKLVRGAYL  AQERARAAEIGYEDPINPTYEATNAMYHRCLDYVLEELKHNAKAKVMV  ASHNEDTVRFALRRMEELGLHPADHQVYFGQLLGMCDQISFPLGQAGYP  VYKYVPYGPVMEVLPYLSRRALENSSLMKGTHRERQLLWLELLRRLRTG  NLFHRPA</p>	PRODH
314	<p>MTTYS DKGAKPERGRFLHFHSVTFWVGNAKQAASFYCSKMGFEPLAYR  GLETGSREVVSHVIKQGKIVFVLSSALNPWNKEMGDHLVKHGDGVKDIA  FEVEDCDYIVQKARER GAKIMREPWVEQDKFGKVKFAVLQTYGDTHTL  VEKMNYIGQFLPGYEAPAFMDPLL PKL PKCSLEMIDHIVGNQPDQEMVSA  SEWYLKNLQFHRFWSVDDTQVHTEYSSLSIVVANYEESIKMPINEPAPG  KKKSQIQEYVDYNGGAGVQHIALKTEDIITAIRHLRERGLEFLSVPSTYYK  QLREKLKTAKIKVKENIDALEELKILVDYDEKGYLLQIFTKPVQDRPTLFL  EVIQRH NHQGF GAGNFNSL FKA FEEEEQNLRGNLTN METNGVVPGM</p>	HPD

10

20

30

40

50

315	MEFSSPSREECPKPLSRVSIMAGSLTGLLLLQAVSWASGARPCIPKSFQYSS VVCVCNATYCDSDPPTFPALGTFSTRYESTRSGRRMELSMGPIQANHTGT GLLLTLQPEQKFQKVKFGGAMTDAAALNILALSPPAQNLLLKSYFSEEG IGYNIIRVPMASCDIFSIRTYTYADTPDDFQLHNFSLPEEDTKLKIPLIHRALQ LAQRPVSLASPWTSPTWLKTNGAVNGKSLKGQP GDIYHQTWARYFVKFLDAYAEHKLQFWAVTAENEPSAGLLSGYPFQCLG FTPEHQRFIARDLGPTLANSTHHNVRLMLDDQRLLLPHWAKVVLTD EAAKYVHGIAVHWYLDLAPAKATLGETHRLFPNTMLFASEACVGSKFW EQSVRLGSWDRGMQYSHSIITNLLYHVVGWTDWNLALNPEGGPNWVRN FVDSPIIVDITKDTFYKQPMFYHLGHFSKFIPEGSQRVGLVASQKNDLDAV ALMHPDGSVVVVLNRSSKDVPLTIKDPVGFLETISPGYSIHTYLWRRQ	GBA
316	MAELKYISGFGNECSSEDPKCPGSLPEGQNNPQVCPYNLYAEQLSGSAFT CPRSTNKRSWL YRILPSVSHKPFESIDEGQVTHNWDEVDPDPNQLRWKPF EIPKASQKKVDFVSLHTLCGAGDIKSNGLAIHIFLCNTSMENRCFYNSD GDFLIVPQKGNLLIYTEFGKMLVQPNEICVIQRGMRFSDVFEETRGYILEV YGVHVFELPDLGPIGANLANPRDFLIPIAWYEDRQVPGGYTVINKYQGKL FAAKQDVSPFNVAWHGNYTPYKYNLKNFMVINSVAFDHADPSIFTVLT AKSVRPGVAIADVFIFPPRWGVADKTFRPPYYHRNCMSEFMGLIRGHYEA KQGGFLPGGSLHSTMTPHGPDADCFEKASKVKLAPERIADGTMAFMFE SSLSLAVTKWGLKASRCLDENYHKCWEPLKSHFTPNSRNPAEPN	HGD
317	MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGA VEFPADKMVSVLVQEGHAVSDMLPLDGLVLASGAGFGVSDVGSGLD CGAGEPAVFRDSDRFSWHDPHLWRSDEAPGLFFVDAERVPCRHDDVFF PPSASFRVGLGPGASPVVRVSISALGRTRFTRDEDLAVFLASRAGRFRHGP GALSVGPEDCADPSGCVCGNAEAQPWICAALLQPLGGRCQAACHSALR PQGQCCDLGAVVLLTHGPAFDLERYRARILDFTLGLPQYHGLQVAVSK VPRSSRLREADTEIQVVLVENGPEPTEGGAGRLARALLADVAENGEALGVLE ATMRESGAHVWGSSAAGLAGGVAAVLLALLVLLVAPPLRRRAGRRLRW RRHEAAAPAGAPLGFNPVFDVTASEELPLPRRLSLVPKAAADSTSHSYF VNPLFAGAEAEA	AMN
318	MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAG PSSGSCPPTKFCRTSGLCVPLTWRCRDLDCSDGSDEEEECRIEPTQKGGQ CPPPPGLPCPCTGVSDCSGGTDKCLRNCRLACLAGELRCTLSDDCIPLTW RCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVTLESVTSLRNATTMG PPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVATLLLLS WLRQERLRPLGLLVAMKESLLLSEQKTSLP	CD320

10

20

30

40

50

319	<p>MMNMSLPFLWSLLTLLIFAEVNGEAGELELQRQKRSINLQQPRMATERG  NLVFLTGSAQNIIEFRFTGSLGKIKLNDEDLSECLHQIQKNKEDIHELKGS  AIGLPQNISSQIYQLNSKLVDLERKFQGLQQTVDKVKCSSNPCQNGGTCLNLH  DSFFCICPPQWKGPLCSADVNECEIYSGTPLSCQNGGTVCVNTMGSYSCHC  PPETYGPQASKYDDCEGGSVARCVHGICEDLMREQAGEPKYSCVCDAG  WMFSPNSPACTLDRDECSFQPGPCSTLVQCFNTQGSFYCGACPTGWQGN  GYICEDINECEINNGGCSVAPPVECVNTPGSSHCQACPPGYQGDGRVCTLT  DICSVSNGGCHPDASCSSTLGSLPLCTCLPGYTGNGYGPNGCVQLSNICLS  HPCLNGQCIDTVSGYFCKCDSGWTGVNCTENINECLSNPCLNGGTCDVG  VDSFSCECTRLWTGALCQVPQVCGESLSGINGSFSYRSPDVGYVHDVNC  FWVIKTEMGKVLRTFTFRLESMDNCPHEFLQVYDGDSSAFQLGRFCG  SSLPHELLSSDNALYFHLYSEHLRNGRGFTVRWETQQPECGGILTGPYGSI  KSPGYPGNYPPGRDCVWIVVTSPDLLVTFTFGTLSLEHHDDCNKDYLEIR  DGPLYQDPLLKGFCTTFSVPPLQTTGPFARIHFHSDSQISDQGFHITYLTSPS  DLRCGGNYTDPEGELFLPELSPGFTHTRQCVYMMKQPQGEQIQINFTHVE  LQCQSDSSQNYIEVRDGETLLGKVCNGTISHIKSITNSVWIRFKIDASVEK  ASFRAVYQVACGDELTEGEGVIRSPFPNVYPGERTCRWTHIQPSQVILLN  FTVFEIGSSAHCETDYVEIGSSSILGSPENKKYCGTDIPSFITSVYNFLYVTF  VKSSSTENHGFMAKFAEDLACGEILTESTGTIQSPGHPNVYPHGINCTWH  ILVQPNHLIHLMFETFHLEFHYNCTNDYLEVYD TDSETSLGRYCGKSIPPS  LTSSGNSL</p> <p>MLVFVTDSDLAYEGFLINYEAISATAACLQDYTDDLGTFTSPNFPNNYPN  NWECIYRITVRTGQLIAVHFTNFSLEEAIGNYYTDFLEIRDGGYEKSPLLGI  FYGSNLPPTIISHSNKLWLKFKSDQIDTRSGFSAYWDGSSTGCGGNLTSS  GTFISPNYPMPYYHSSECYWWLKSSHGSAFELEFKDFHLEHHPNCTLDYL  AVYDGPSSNSHLLTQLCGDEKPLIRSSGDSMFIKLR</p> <p>TDEGQQGRGFKAEYRQTCENVVIVNQTYGILESIGYPNPYSENQHCNWTI  RATTGNTVNYTFLAFDLEHHINCSTDYLELYDGPQRMGRYCGVDLPPPGS  TTSSKLQVLLLTDGVGRREKGFQMWFVYGGGELSGATGSFSSPGFPN  RYPPNKECIWYIRTDPGSSIQLTIHDFDVEYHSRCNFDVLEIYGGPDFHSPR  IAQLCTQRSPENPMQVSSTGNELAIRFKTDLSINGRGNASWQAVTGCGG  GIFQAPSGEIHSPNYSPYRSNTDCSWVIRVDRNHRVLLNFTDFDLEPQDS  CIMAYDGLSSTMSRLARTCGREQLANPIVSSGNSLFLRFQSGPSRQNRGFR  AQFRQACGGHILTSSFDTVSSPRFPANYPNNQNCSWIIQAQPPLNHITLSFT  HFELERSTTCARDFVEILDGGHEDAPLRGRYCGTDMPHPITSFSSALTLRF  VSDSSISAGGFHTTASVSACGGTFYMAEGIFNSPGYPDIYPPNVECVWN  IVSSPGNRLQLSFISFQLEDSQDCSRDFVEIREGNATGHLVGRYCGNSFPLN  YSSIVGHTLWVRFISDGSSTGTFQATFMKIFGNDNIVGTHGKVASPFWPE  NYPHNSNYQWTVNVNASHVVHGRILEMDIEEIQNCYYDKLRIYDGPSIHA  RLIGAYCGTQTESFSSTGNSLTFHFYSDSSISGKGFLEWFAVDAPDGVLP  TIAPGACGGFLRTGDAPVFLFSPGWPDYSNRVDCTWLIQAPDSTVELNIL  SLDIESHRTCAYDSLVRDGDNNLAQQLAVLCGREIPGPIRSTGEYMFIRFT</p>	CUBN
-----	---	------

10

20

30

40

50



	<p>SDSSVTRAGFNASFHKSCGGYLHADRGIITSPKYPETYPSNLNCSWHVLV          QSGLTIAVHFEQPFQIPNGDSSCNQGDYLVLRNGPDICSPPLGPPGGNGHF          CGSHASSTLFTSDNQMFVQFISDHSNEGQGFKIKYEAKSLACGGNVYIHD          ADSAGYVTSFNHPHNYPPHADCIWILAAPPETRIQLQFEDRFDIEVTPNCTS          NYLELRDGVDSADAPILSKFCGTSLPSSQWSSGEVMYLRFRSDNSPTHVGF          KAKYSIAQCGRVPGQSGVVESIGHPTLPYRDNLFCWEHLQGLSGHYLTI          SFEDFNLQNSSGCEKDFVEIWDNHTSGNILGRYCGNTIPDSIDTSSNTAVV          RFVTDGSVTASGFRLRFESSMEECGGDLQGSIGTFTSPNYPNPNPHGRICE          WRITAPEGRITLMFNRLRLATHPSCNNEHVIVFNGIRSNPQLEKLCSSV          NVSNEIKSSGNTMKVIFFTDGRPYGGFTASYTSEDVAVCGGSLPNTPEGN          FTSPGYDGVNRYSRNLNCEWTLSPNPQGNSSISIHFEFDYLESHQDCQFDV          LEFRVGDADGPLMWRLCGPSKPTLPLVIPYSQVWIHFVTNERVEHIGFHA          KYSFTDCGGIQIGDSGVITSPNYPNAYDSLTHCSSLLEAPQGHTITLTFSDF          DIEPHTTCAWDSVTVRNGGSPESPIIGQYCGNSNPRTIQSGSNQLVVTFNS          DHSLQGGGFYATWNTQTLGCGGIFHSDNGTIRSPHWPQNFPENSRCSWT          AITHKSKHLEISFDNNFLIPSGDQCQNSFVKVWAGTEEVDKALLATGCG          NVAPGPVITPSNTFTA VFSQSEAPAQGFSASFVSRCGSNFTGPSGYIISPNY          PKQYDNNMNTYVIEANPLSVLLTFVFSFHLEARS AVTGSCVNDGVHIIR          GYSVMSTPFATVCG          DEMPAPLTIAGPVLLNFYSNEQITDFGFKFSYRIISCGGVFNFSGGIITSPAYS          YADYPNDMHCLYTITVSDDKVIELKFSDFDVPSTSCSHDYLAIDGANT          SDPLLKFCGSKRPPNVKSSNNSMLLVFKTDSFQTAKGWKMSFRQTLGP          QQCGGGYLTGSNNTFASPDSDSNGMYDKNLNCVWIIIPVNVKVIHLTFNT          FALEAASTRQRCLYDYVKLYDGDSEANLAGTFCGSTVPAPFISSGNFLT          VQFISDLTLEREGFNATYTIMDMPCGGTYNATWTPQNISSPNSSDPDVPFSI          CTWVIDSPPHQQVKITVWALQLTSQDCTQNYLQLQDSPQGHGNSRFQFC          GRNASA VPVFYSSMSTAMVIFKSGVVNRNSRMSFTYQIADCNRDYHKAF          GNLRSPGWPDNYDNDKDCTVTLTAPQNHTISLFFHSLGIENSVECRNDFL          EVRNGSNSNSPLLKGYCGTLLPNPVFSQNNELYLRFKSDSVTSDRGYEHW          TSSPSGCGGTLYGDRGSFTSPGYPGTYPNNTYCEWVLVAPAGRLVTINFY          FISIDDPGDCVQNYLTLYDGNASSPSSGPYCGGDTSIAPFVASSNQVFIKF          HADYARRPSAFRLTWDS</p>	
320	<p>MAWFALYLLSLLWATAGTSTQTQSSCSVPSAQEPLVNGIQVLMENSVTSS          AYPNPSILIAMNLAGAYNLKAQKLLTYQLMSSDNNDLTIGQLGLTIMALT          SSCRDPGDKVSILQRQMENWAPSSPNAEASAFYGPSLAILALCQKNSEAT          LPIAVRFAKTLLANSSPFNVDTGAMATLALTCMYNKIPVGSEEGYRSLFG          QVLKDIVEKISMKIKDNGIIGDIYSTGLAMQALSVTPEPSKKEWNCKKTTD          MILNEIKQGKFHNPMIAQILPSLKGKTYLDVPQVTCSPDHEVQPTLPSNP          GPGPTSASNITVIYTINNQLRGVELLFNETINVSVKSGSVLLVVLEEAQRKN          PMFKFETMTSWGLVVSSINNIAENVNHKTYWQFLSGVTPLNEGVADYIP          FNHEHITANFTQY</p>	GIF

10

20

30

40

50

321	MRQSHQLPLVGLLLFSFIPSQLCEICEVSEENYIRLKPLLNTMIQSNYNRGT SAVNVVLSLKLVGIIQITLMQKMIQQIKYNVKSRLSDVSSGELALHILALG VCRNAEENLIYDYHLIDKLENKFQAEIENMEAHNGTPLTNYQQSLDVLAL LCLFNGNYSTAEVVNHFTPENKNYYFGSQFSVDTGAMAVLALTCVKKSL INGQIKADEGSLKNISIIYTKSLVEKILSEKKENGLIGN TFSTGEAMQALFVSSDYNNENDWNCQQLNTVLTEISQGAFSNPNAAAQ VLPALMGKTFDINKDSSCVSASGNFNISADEPITVTPPDSQSYISVNYSVR INETYFTNVTVLNGSVFLSVMEKAQKMNDTIFGFTMEERSWGPYITCIQG LCANNNDRTYWELLSGGEPLSQGAGSYVVRNGENLEVRWSKY	TCN1
322	MRHLGAFLFLLGVLGALTEMCEIPEMDSLHVEKLGQHLLPWMDRLSLEH LNPSIYVGLRSSLQAGTKEDLYLHSLKLGYYQQCLLGSFAFSEDDGDCQGK PSMGQLALYLLALRANCEFVRGHKGDRLVSQLKWFLEDEKRAIGHDHK GHPHTSYYYQYGLGILALCLHQKRVHDSVVDKLLYAVEPFHQGHHSVDTA AMAGLAFTCLKRSNFPGRQRITMAIRTVREEILKAQTPEGHFGNVYST PLALQFLMTSPMRGAELGTACLKARVALLASLQDGAFAQNALMISQLLPV LNHKTYIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSS VLAGSTVEDVLKKAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQL LRDPNTPLLQGIADYRPKDGETIELRLVSW	TCN2
323	MQQKTKLFLQALKYSIPHLGKCMQKQHLNHYNFADHCYNRIKLKKYHL TKCLQNKPKISELARNIPSRFSCKDLQPVKQENKPLPENMDAFEKVRTK LETQPQEEYEIINVEVKHGGFVYYQEGCCLVRSKDEEADNDNYEVLNLE ELKLDQPFIDCIRVAPDEKYVAAKIRTEDSEASTCVIHKLSQPVMEASFPN VSSFEWVKDEEDEDVLFYTFQRNLRCHDVYRATFGDNKRNERFYTEKDP SYFVFLYLTKDSRFLTINIMNKTSEVWLIDGLSPWPPVLIQKRIHGVLY YVEHRDDELYILTNVGEPTFKLMRTAADTPAIMNWDLFFTMKRNTKVI DLDMFKDHCVFLFKHSNLLYVNVIGLADDSVRSKLPWACGFIMDTNS DPKNCPFQLCSPIRPPKYTYKFAEGKLFEEFGHEDPITKTSRVLRLKAS KDGLKVPMTVFHKTSEDLQKKPLLHVYGYGMDLKMNFRPERRVLV DDGWILAYCHVRGGGELGLQWHADGRLTKKLNLGLADLEACIKTLHGQG FSQPSLTTLTAFSAGGVLAGALCNSNPENVRAVTLEAPFLDVLNTMMDTT LPLT LEELEEWGNPSSDEKHKNYIKRYCPYQNIKQHYPSIHITAYENDERVPLK GIVSYTEKLKEAIAEHAHAKDTGEGYQTPNIILDIQPGGNHVIEDSHKKITAQI KFLYEELGLDSTSVFEDLKKYLF	PREPL

10

20

30

40

50

324	MAFANLRKVLISDSLDPCCRKILQDGGGLQVVEKQNLKEELIAELQDCEG LIVRSATKVTADVINA AEKLQVVGRAGTGVDNVDLEAATRKGILVMNTP NGNSLSAAELTCGMIMCLARQIPQATASMKGKWERKKFMGTENGKT LGILGLGRIGREVATRMQSFQGMKTIGYDPIISPEVSASFQVQQLPLEEIWPL CDFITVHTPLLPSTTGLLNDNTFAQCKKGVRVVCARGGIVDEGALLRAL QSGQCAGAALDVFTEEPPRDRALVDHENVISCPHLGASTKEAQSRCGEEI A VQFVDMVKGKSLTGVVNAQALTSAFSPHTKPWIGLAEALGTLMRAWAG SPKGTIQVITQGTSLKNAGNCLSPAVIVGLLKEASKQADVNLVNAKLLVK EAGLNVTTSHSPAAPGEQGFGECLLAVALAGAPYQAVGLVQGTTPVLQ LNGAVFRPEVPLRRDLPLLLFRTQTS DPAMLPTMIGLLAEAGVRLLSYQTS LVSDGETWHVMGISSLLPSLEAWKQHVTEAFQFHF	PHGDH
325	MDAPRQVVNFGPGPAKLPHSVLLEIQKELLDYKGVGISVLEMSHRSSDFA KIINNTENLVRELLAVPDNYKVI FLQGGGCGQFSAVPLNLIGLKAGRCAD YVVTGAWSAKAAEEAKKFGTINIVHPKLGSYTKIPDPSTWNLNPDASYVY YCANETVHGVVEFDIPDVKGAVLVCDMSSNFLSKPVDVSKFGVIFAGAOK NVGSAGVTVVIVRDDLLGFALRECPSVLEYKVQAGNSSLYNTPPCFISIYV MGLVLEWIKNNGGAAAMEKLSIKSQTIEIIDNSQGFYVCPVEPQNRSK MNIPFRIGNAKGDDALEKRFLDKALELNMLSLKGHRVSGGIRASLYNAV IEDVQKLA AFMKKFLEMHQL	PSATI
326	MVSHSELRKL FYSADAVCFDVDSTVIREEGIDELAKICGVEDAVSEMTRR AMGGAVPFKAALTERLALI QPSREQVQRLIAEQPPHLP GIRELVSRLQER NVQVFLISGGFRSIVEHVASKLNIPATNVFANRLKFYFNGEYAGFDETQPT AESGGKGVKIKLLKEKFHFKIIMIGDGATDMEACPPADAFIGFGGNVIRQ QVKDNAK WYITDFVELLGELEE	PSPH
327	MQRASVVARLGFRLQAFPPALCRPLSCAQEVLRRTPLYDFHLAHGGKM VAFAGWVSLPVQYRDSHTD SHLHTRQHCSLFDVSHMLQTKILGSDRVKLM ESLVVGDIAELRPNQGTLSLFTNEAGGILDDLIVTNTSEGHLV VVSNAGC WEKDLALMQDKVRELQNGRQDVGLEVLNALLALQGP TAAQVLQAGV ADDLRKLPMFMTSAVMEVFGVSGCRVTRCGYTGEDGVEISVPVAGAVHLA TAILKNPEVKLAGLAARDSLRLEAGLCLYGN DIDEHTTPVEGSLSWTLGK RRRAAMDFPGAKVIVPQLKGRVQRRRVGLMCEGAPMRAHSPILNMEGT KIGTVTSGCPSPLKKNVAMGYVPCEYSRPGTM LLEVRRKQQMAVVSK MPFVPTNYITLK	AMT
328	MALRVVRSVRALLCTLRAVPSPAAPCPPRPWQLG VGAVRTLRTGPALLS VRKFTEKHEWVTTENGIGTVGISNFAQEALGDVVYCSLPEVGTCLNKQD EFGALESVKAASELYSPLSGEVTEINEALAENPGLV NKSCYEDGWLIKMT LSNPSELDELMSEEAYEKYIKSIEE	GCSH

10

20

30

40

50

329	<p>MQSCARAWGLRLGRGVGGRRLAGGSGPCWAPRSRDSSSGGDSAAAG  ASRLLERLLPRHDDFARRHIGPGDKDQREMLQTLGLASIDELIEKTVPANI  RLKRPLKMEDPVCENEILATLHAISSKNQIWRSYIGMGYYNCSVPQTILRN  LLENSGWITQYTPYQPEVSQGRLESLLNYQTMVCDITGLDMANASLLDEG  TAAAEALQLCYRHNKRRKFLVDPRCHPQTIAVVQTRAKYTGVLTELKLP  CEMDFSGKDVSGVLFQYPDTEGKVEDFTELVERAHQSGSLACCATDLLA  LC  ILRPPGEFGVDIALGSSQRFVPLGYGGPHAAFFAVRESLVRMMPGRMVG  VTRDATGKEVYRLALQTREQHRRDKATSNICTAQALLANMAAMFAIYH  GSHGLEHIARRVHNATLILSEGLKRAGHQLQHDLFFDTLKIQCSCSVKEV  LGRAAQRQINFRLFEDGTLGISLDETVNEKDLDDLWIFGCESSAELVAES  MGEECRGIPGSVFKRTSPFLTHQVFNSYHSETNIVRYMKKLENKDISLVHS  MIPLGCTMKNLSSSELAPITWKEFANIHPFVPLDQAQGYQQLFRELEKDL  CELTGYDQVCFQPNNSGAQGEYAGLATIRAYLNQKGEHRTVCLIPKSAH  GTNPASAHMAGMKIQPVEVDKYGNIDAVHLKAMVDKHKENLAAIMITY  PSTNGVFEENISDVCDLIHQHGGQVYLDGANMNAQVVICRPGDFGSDVS  HLNLHKTFCIPHGGGGPGMGPVIGVKKHLAPFLPNHPVISLKRNEACPVG  TVSAAWPGSSSILPISWAYIKMMGGKGLKQATETAILNANYMAKRLETH  YRILFRGARGYVGFHEFILDTRPFKKSANIEAVDVAKRLQDYGFHAPTMSW  PVAGTLMVEPTESDKAELDRFCDAMISIRQEIADIEGRIDPRVNPLKMSP  HSLTCVTSSHWRPYSREVAAPLFPVKPENKFWPTIARIDDIYGDQHLVC  TCPPMEVYESPFSEQKRASS</p>	GLDC
330	<p>MSLRCGDAARTLGPRVFGRYFCSPVRPLSSLPKKKELLQNGPDLQDFVS  GDLADRSTWDEYKGNLKRQKGERLRLPPWLKTEIPMGKNYNKLNKTLR  NLNLHTVCEEARCPNIGECWGGGEYATATATIMLMGDTCTRGCRFCSVK  TARNPPPLDASEPYNTAKAIAEWGLDYVVLTSVDRDDMPDGGAEHIAKT  VSYLKERNPKILVECLTPDFRGDLKAIEKVALSGLDVYAHNVETVPELQS  KVRDPRANFDQSLRVLKHAKKQPDVSKTSIMLGLGENDEQVYATMKA  LREADVDCLTLGQYMQPTRRHLKVEEYITPEKFKYWEKVGNELGFHYTA  SGPLVRSSYKAGEFFL  KNLVAKRRTKDL</p>	LIAS
331	<p>MAATARRGWGAAAVAAGLRRRFCHMLKNPYTIKKQPLHQFVQRPLFPL  PAAFYHPVRYMFIQTQDTPNPNSLKFIPGKPVLETRTMDFPTAAAFRSPL  ARQLFRIEGVKSVEFGPDFITVTKENEELDWNLLKPDIIYATIMDFFASGLPL  VTEETPSGEAGSEEDDEVVAMIKELLDTRIRPTVQEDGGDVIIYKGFEDGIV  QLKLQGSCTSCPSSIITLKNGIQNMLQFYIPEVEGVEQVMDESDEKEANS  P</p>	NFU1

10

20

30

40

50

332	MSGGDTRAAIARPRMAAAHGPVAPSSPEQVTLLPVQRSFFLPPFSGATPST SLAESVLKVWHGAYNSGLLPQLMAQHSLAMAQNGAVPSEATKRDQNLK RGNWGNQIEFVLTSTVGYAVGLGNVWRFYLYCYRNGGGAFMFPYFIMLIF CGIPLFFMELSFQQFASQGCLGVWRISPMFKGVGYGMMVVSTYIGIYYNV VICIAFYFFSSMTHVLPWAYCNPWNTHDCAGVLDASNLNNGSRPAAL PSNLSHLLNHSLQRTSPSEEWRLYVLKLSDDIGNFGEVRLPLLGLGVS WLVVFLCLRIGVKSSGKVVYFTATFPYVVLTLFVRGVTLLEGAFDGIMYY LTPQWDKILEAKVWGDAAASQIFYSLGCWGLITMASYNKFHNNCYRDS VIISITNCATSUYAGFVIFSILGFMANHLGVDVSRVADHGPGLAFVAYPEA LTLPLISPLWSLLFFFMLLILGLGTQFCLLETLVTAIVDEVGNEWILQKITY VTLGVAVAGFLLGIPLTSQAGIYWLLMDNYAASFSLVVISICIMCVAIMYI YGHRNYFQDIQMMMLGFPPPLFFQICWRFVSPAIIFFILVFTVIQYQPITYNH YQYPGWAVAIGFLMALSSVLCIPLYAMFRLCRTDGD TLLQRLKNATKPSR DWGPALLEHRTGRYAPTIAPSPEDGFVQPLHPDKAQIPIVGSNGSSRLQD SRI	SLC6A9
333	MEPSSKKLTGRLMLAVGGAVLGSQFGYNTGVINAPQKVIEEFYNQTWV HRYGESILPTTLTTLWLSVAIFSVMIGSFSVGLFVNRFGRRNSMLMM NLLAFVSAVLMGFSKLGKSFEMLILGRFIIGVYCGLTTFVPMYVGEVSPT ALRGALGTLHQLGIVVGILIAQVFGLDSIMGNKDLWPLLLSIIFIPALLQCIV LPFCPEPRFLLINRNEENRAKSVLKKLRGTADVTHDLQEMKEESRQMMR EKKVTILELFRSPAYRQPILIAVVLQLSQQLSGINAVFYYSTSIFEKAGVQQ PVYATIGSGIVNTAFTVVSLFVVERAGRRTLHLIGLAGMAGCAILMTIALA LLEQLPWMSYLSIVAIFGFVAFFEVGPGPIWFIVAELFSQGPRPAAIAVAG FSNWTSNFIVGMCQYVEQLCGPYVFIIFTVLLVLFIFTYFKVPETKGRTE DEIASGFRQGGASQSDKTPE ELFHPLGADSQV	SLC2A1

10

20

30

40

50

334	<p>MDPSMGVNSVTISVEGMTCNCSVWTIEQQIGKVNGVHHIKVSLEEKNATI          IYDPKLTQPKTLQEAIIDMGFDAVIHNPDPVLTDTLFLVTASLTLPWD          HIQSTLLKTKGVTDIKIYPQKRTVAVTIIPSIVNANQIKELVPELSLDTGTLE          KKSGACEDHSMAQAGEVVLKMKVEGMTCHSCTSTIEGKIGKLGQVQRIK          VSLDNQEATIVYQPHLISVEEMKKQIEAMGFPAFVKKQPKYLKLGAI DVE          RLKNTPVKSSEGSQQRSPSYTNDSTATFIIDGMHCKSCVSNIESTLSALQY          VSSIVVSLENRSAIVKYNASSVTPESLRKAIEAVSPGLYRVSITSEVESTSNS          PSSSSLQKIPLNVVSQPLTQETVINIDGMTCNCSVQSIEGVISKKPGVK SIRV          SLANSNGTVEYDPLLTSPELTRGAIEDMGFDATLSDTNEPLVVIAQPSSEM          PLLTSTNEFYTKGMTPVQD          KEEGKNSSKCYIQVTGMTASCVANIERNLREEGIYSILVALMAGKAEV          RYNPAVIQPPMIAEFIRELGFATVIENADEGDGVLELVVRGMTASCVH          KIESSLTKHRGILYCSVALATNKAHIKYDPEIIGPRDIIHTIESLGFEASLVK          KDRSASHLDHKREIRQWRRSFLVSLFFCIPVMGLMIYMMVMDHHEFATLH          HNQNMSKEEMINLHSSMFLERQILPGLSVMNLLSFLLC          VPVQFFGGWYFYIQAYKALKHKTANMDVLIVLATTIAFAYSLIILLVAMY          ERAKVNPIFFDTPPMLFVFIALGRWLEHIAKGTSEALAKLISLQATEATI          VTLDSDNILLSEEQVDVELVQRGDIKVVPGGKFPVDGRVIEGHSMDVDESL          ITGEAMPVAKKPGSTVIAGSINQNGSLLICATHVGADTTLSQIVKLVEEAQ          TSKAPIQQFADKLSGYFVPFIVFVSIATLLVWIVIG          FLNFEIVETYFPGYNRSISRTETIIRFAFQASITVLCIACPCSLGLATPTAVM          VGTGVGAQNGILIKGGEPLEMAHKVKVVVFDKGTGTITHGTPVNVQVKVL          TESNRISHHKILAIVGTAESNSEHPLGTAITKYCKQELDTETLGT CIDFQVV          PGC GISCKVTNIEGLLHKNNWNIEDNNIKNASLVQIDASNEQSSTSSSMIID          AQISNALNAQYKVLIGNREWMIRNGLVINNDVN          DFMTEHERKGR TAVLVAVDDEL CGLIAIADTVKPEAELAIHILKSMGLEV          VLMTGDNSKTARSIASQVGITKVF AEVLP SHKVAKVQLQEEGKRVAMV          GDGINDSPALAMANVGIAIGTGDVAIEAADVVLIRNDLLDVVASIDL SRK          TVKRIRINFVFALIYNLVGIPIAAGVFMPIGLVLQPWMGSAAMAASSVSVV          LSSLFLKLYRKPTYESYELPARSQIGQKSPSEISVHVGIDDTSRNSPKLGLL          DRIVNYSRASINLLSDKRSLNSVVTSEPKHSLLVGDFREDDDTAL</p>	ATP7A
335	<p>MMRFMLLFSRQGKRLRLQKWYLATSDKERKKMVRELMQVVLARKPKMC          SFLEWRDLKV VYKRYASLYFCCAIEGQDNELITLELIHRYVELLDKYFGS          VCELDIIFNEKAYFILDEFMGGDVQDTSKKSVLKAIEQADLLQEEDESP          RSVLEEMGLA</p>	AP1S1

10

20

30

40

50

336	<p>MKILILGIFLFLCSTPAWAKEKHYYIGIETTWDYASDHGEKKLISVDTEHS  NIYLQNGPDRIGRLYKALYLQYTDDETFRTTIEKPVWLGLGPIKAETGD  KVYVHLKLNLASRPYTFHSHGITYYKEHEGAIYPDNTTDFQRADDKVYPG  EQYTYMLLATEEQSPGEGDGNVCVTRIVYHSHIDAPKDIASGLIGPLIICKKDS  LDKEKEKHIDREFVVMFVVDENFSWYLEDNIKTYC  SEPEKVDKDNEDFQESNRMYSVNGYTFGSLPGLSMCAEDRVKWYLFGM  GNEVDVHAFFHFGQALTNKNYRIDTINLFPATLFDAYMVAQNPGEWMLS  CQNLNHLKAGLQAFFQVQECNKSSSKDNIRGKHVRHYIAAEEIWNYP  SGIDIFTKENLTAPGSDSAVFEQGTTRIGGSYKLVYREYTDASFTNRKE  RGPEEEHLGILGPVIWAIEVGDITRVTFHNKGA YPLSIEPIGVRFNKNNEG  YYSNPNPQRSVPPSASHVAPTETFTYEWTVPKVGPVNADPVCLAKM  YY  SAVDPTKDIFTGLIGPMKICKKGLSHANGRQKDVDKEFYLFPTVFDENESL  LLEDNIRMFTTAPDQVDKEDEDFQESNKMMSMNGFMYGNQPGLTMCKG  DSVVWYLFSAAGNEADVHGIVYFSGNTYLWRGERRDANLFPQTSLLHM  WPDTEGTFNVECLTDDHYTGGMKQKYTVNQCRRQSEDSTFYLGERTYYI  AAVEVEWDYSPQREWEKELHHLQEONVSNAFLDKGEFYIGSKYKVVY  RQYTDSTFRVVERKAEHLGILGPQLHADVDGDKVKIIFKNMATRPYSIH  AHGVQTESSTVPTLPGETLTYVWKIPERSGAGTEDSACIPWAYYSTVDQ  VKDLYSGLIGPLIVCRRPYLKVFNPRRKLIFALLFLVFDENESWYLDLNIK  TYS DHPEKVNKDDEEFIESNKMHAINGRMFGNLQGLTMHVGDEVNWYL  MGMGNEIDLHTVHFHGHFSFYKHRGVYSSDVFDIFPGTYQTLEMFPRTPG  IWLHCHVTDHIHAGMETTYTVLQNEDETKSG</p>	CP
337	<p>MSPTISHKDSRQRRPGNFHSLDMKSGPLPPGGWDDSHLDSAGREGDRE  ALLGDTGTGDFLKAPQSFRAELSSILLFLYVLQGIPLGLAGSIPLILQSKN  VSYTDQAFFSFVFWPFLKLLWAPLVDVYVKNFGRKRSWLVPYQYILG  LFMIYLSLSTQVDRLLGNTDDRTPDVIALTVAFFLFEFLAATQDIAVDGVAL  TMLSRENVGYASTCNSVGQTAGYFLGNVFLALESADFCNKYLRFPQP  RGIVTSLDFLFFWGTVFLITTTLVALLKKENEVSVVKEETQGITDTYKL  LFAIKMPAVLTFCLLILTAKIGFSAADAVTGLKLVVEGVPKEHLALLAVP  MVPLQIILPLIISKYTAGPQPLNTFYKAMPYRLLGLLEYALLVWVTPKVEH  QGGFPIYYYIVVLLSYALHQVTVYSMYVSIMAFNAKVSDPLIGGTMYMTLL  NTVSNLGGNWPSTVALWLVDPLTVKECVGASNQCRTPDAVELCKKLG  GSCVTALDGYVESIICVFIGFGWFFLGPKFKLQDEGSSSWKCKRNN</p>	SLC33A1
338	<p>MSAVCGGAARMLRTPGRHGYAAEFSPYLPRLACATAQHYGIAGCGTLL  ILDPEAGLRLFRSFDWNGDFDVTWSENNEHVLITCSGDGSLQLWDTA  KAAGPLQVYKEHAQEVYSVDWSQTRGEQLVVSWSWDQTVKLWDPTVG  KSLCTFRGHESIYSTIWSHIPGCFASASGDQTLRIWDVKAAGVRIVIPAH  QAEILSCDWCKYNENLLVTGAVDCSLRGWDLRNVRQPVFELLGHTYAIR  RVKFSFPHASVLASCSYDFTVRFWNFSKPDLSLLETVEHHTEFTCGLDFSLQ  SPTQVADCSWDETIKIYDPACLTIPA</p>	PEX7

10

20

30

40

50

339	MEQLRAAARLQIVLGHLLGRPSAGAVVAHPTSGTISSASFHPQQFQYTLDN NVLTLERKQFYEEENGFLVIKNLVPDADIQRFRNEFEKICRKEVKPLGLTVM RDVTISKSEYAPSEKMITKVQDFQEDKELFRYCTLPEILKYVECTGPNIM AMHTMLINKPPDSGKTSRHLPLHQDLHYFPRPSDLIVCAWTAMEHISRN NGCLVVLPGTHKGLKPHDYPKWEGGVNKMFGHIQDYEENKARVHLM EKGDTVFFHPLLIHGSGQNKTOGFRKAISCHFASADCHYIDVKGTSQENIE KEVVGIAHKFFGAENSVNLKDIWMFRARLVKGERTNL	PHYH
340	MAEAAAAAGGTGLGAGASYGSAADRDRDPDRAGRRLRVLSGHELLGR PREALSTNECKARRAASAATAAPTATPAAQESGTIPKKRQEVMMKWNW GYNDSKFIFNKKGQIELTGKRYPLSGMGLPTFKEWIQNTLGVNVEHKTT KASLNPSDTPPSVVNEDFLHDLKETNISYSQEADDRVFRAHGHCLHEIFLL REGMFERIPDIVLWPTCHDDVVKIVNLACKYNLCIPIGGGTSVSYGLMCP ADETRTHSLDTSQMNRLWVDENNLTAHVEAGITGQELERQLKESGYCT GH EPDSLEFSTVGGWVSTRASGMKKNYGNIEDLVVHIKMVTPRGIIEKSCQG PRMSTGPDIIHFIMGSEGTLGVITEATIKIRPVPEYQKYGSVAFPNFEQVA CLREIAKQRCAPASIRLMDNKQFQFGHALKPQVSSIFTSFLDGLKKFYITKF KGFDPNQLSVATLLFEGDREKVLQHEKQVYDIAAKFGGLAAGEDNGQRG YLLTYVIAIYRDLALEYYVLGESFETSAPWDRVVDLCRNVKERITRECKE KGVQFAPFSTCRVTQTYDAGACIYFYAFNYRGISDPLTVFEQTEAAAREE ILANGGSLSHHHGVGKLRKQWLKESISDVGFGMLKSVKEYVDPNNIFGN RNLL	AGPS
341	MESSSSNSYFSVGPTSPSAVVLLYSKELKKWDEFEDILEERRHVSDLKFA MKCYTPLVYKGITPCKPIDIKCSVLNSEEIHVYIKQLSKESLQSVDLREEV SEILDEMShKLRLGAIRFCAFTLSKVFQIFSKVCVNEEGIQKLQRAIQEHP VVLLPSHRSYIDFLMLSFLLYNYDLVPPVIAAGMDFLGMKMGVCELLRMS GAFFMRRTFGGNKLYWAVFSEYVKTMLRNGYAPVEFFLEGTRRSRSTL TPKFGLLNIVMEPFFKREVFDTYLVPISSYDKILEETLYVYELLGVPKPKE STTGLLKARKILSENFSGSIHVYFGDPVSLRSLAAGRMSRSSYNLVPRYIPQ KQSEDMHAFVTEVAYKMELLQIENMVLSPWTLIVAVLLQNRPSMDFDAL VEKTLWLKGLTQAFGGFLIWPDNKPAEEVVPASILLHSNIASLVKDQVILK VDSGDSEVVDGLMLQHITLLMCSAYRNQLLNIFVRPSLVAVALQMTPGF RKEDVYSCFRFLRDVFADEFIFLPGNTLKDFEEGCYLLCKSEAIQVTTKDIL VTEKGNTVLEFLVGLFKPFVESYQIICKYLLSEEEDHFSEEQYLAAVRKFT SQLLDQGTSQCYDVLSSDVQKNALAACVRLGVVEKKKINNNCIFVNEP ATTKLEEMLGCKTPIGKPATAKL	GNPAT

10

20

30

40

50



342	<p>MPVLSRPRPWRGNTLKRRTAVLLALAAYGAHKVYPLVRQCLAPARGLQA  PAGEPTQEASGVAAAKAGMNRVFLQRLLWLLRLLFPRVLCRETGLLALH  SAALVSRFTLSVYVARLDGRLARCIVRKDPRAFGWQLLQWLLIALPATFV  NSAIRYLEGQLALSFRSRLVAHAYRLYFSQTTYRVSNMDGRLRNPDQSL  TEDVVFAAASVAHLYSNLTKPLLDVAVTSTYLLRAARSRGAGTAWPSAI  AGLVVFLTANVLRASFSPKFGELVAEEARRKGELRYMHSRVVANSEEIAFY  GGHEVELALLQRSYQDLASQINLILLERLWYVMLEQFLMKYVWSASGLL  MVAVPIITATGYSESDAEAVKKAALKKEEELVSETEAFTIARNLLTAAA  DAIERIMSSYKEVTELAGYTARVHEMFQVFEDVQRCHFKRPRELEDAQA  GSGTIGRSGVRVEGPKIRGQVVDVEQGIICENIPIVTPSGEVVVASLNIRV  EEGMHLLITGPNCGKSSLFRILGGLWPTYGGVLYKPPPQRMFYIPQRPY  MSVGLSRDQVIYPSVEDMQRKGyseQDLEAILDVVHLHHILQREGGWE  AMCD  WKDVLSGGEKQRIGMARMFYHRPKYALLDECTSAVSIDVEGKIFQAAKD  AGIALLSITHRPSLWKYHThLLQFDGEGGWKFEKLDSAARLSLTEEKQRL  EQQLAGIPKMQRRLQELCQILGEAVAPAHVPAPSPQGGPGLQGAST</p>	ABCD1
343	<p>MNPDLRRERDSASFNPPELLTHILDGSPEKTRRRREIENMILNDPDFQHEDL  NFLTRSQRYEVAVRKSAIMVKKMREFGIADPDEIMWFKKLHLVNFVEPV  GLNYSMFIPTLLNQGTTAQKEKWLLSSKGLQIIGTYAQTEMGHGTHLRGL  ETTATYDPETQEFILNSPTVTSIKWWPGGLGKTSNHAIVLAQLITK GKCYG  LHAFIVPIREIGTHKPLPGITVGDIGPKFGYDEIDNGYLKMDNHRIPRENML  MKYAQVKPDGTYVKPLSNKLTYGTMVVFVRSFLVGEAARALSKACTIAIR  YSAVRHQSEIKPGEPEPQILDFQTQQYKLFLLATAYAFQFVGAYMKETY  HRINEGIGQGDLSELPELHALTAGLKAFTSWTANTGIEACRMACGGHGYS  HCSGLPNYVNFTPSCTFEGENTVMMLQTARFLMKSYDQVHSGKLVCGM  VSYLNDLPSQRIQPQVAVWPTMVDINSPELSTEAYKLRAARLVEIAAKN  LQKEVIHRKSKEVAWNLTsvDLVRASEAHCHYVVVKLFSEKLLKIQDKAI  QAVLRSCLLysLYGISQnAGDFLQGSIMTEPQITQVNQRVKELLTLIRSD  AVALVDAFDQDVTLGsvLGRYDGNVYENLFEWAKNSPLNKAevHESY  KHLKSLQSKL</p>	ACOX1

10

20

30

40

50

344	<p>MWGS DR LAGAGGGAAVTVAF T NARD CFLHLPRRLVAQLHLLQNQAIE  VVWSHQPAFLSWVEGRHFS DQGENVAEINRQVGQKLGLSNGGQVFLKPC  SHVWSCQQVEVEPLSADDWEILELH AVSLEQHLLDQIRIVFPKAIFPVWVD  QQTYIFIQIVALIPAASYGRLETDTKLLIQPKTRAKENTFSKADA EYKKLH  SYGRDQK GMMKELQTKQLQSN TVGITESNENESEIPVDSSSVASLWTMIG  SIFSFQSEKKQETS WGLTEINAFKNMQSKV VPLDNIFRVCKSQPPSIYNASA  TSVFH KHCAIHVFPWDQEYFDVEPSFTVTY GKLVKLLSPKQQQSKTKQN  VLSPEKEKQ MSEPLDQKKIRSDHNEEDEKACVLQVVWNGLEELNNAIKY  TKNVEVLHLGKVWIPDDL RKR LNIEMH AVVRITPVEVTPKIPRSLK LQPRE  NLPKDISEEDIKTVFYSWLQSTTTMLPLVISEEEFIKLETKDGLKEFSLSIV  HSWEKEKDKNIFLLSPNLLQKTTIQVLLDPMVKEEN  SEEIDFILPFLKLSLGGVNSLGVSSLEHITHSLLGRPLSRQLMSLVAGLRN  GALLTG GKGSGKSTLAKAICKEAFDKLDAHVERVDCKALRGKRLENIQ  KTLEVA FSEAVWMQPSVLLDDLDLIAGLPAVPEHEHSPDAVQSQR LAH  ALNDMIKEFISMGSLVALIATSQSQSLHPLL VSAQGVHIFQCVQHIQPPN  QEQRCEILCNVIK NKLDCDINKFTDL DLQHVAKETGGFVARDFTVLV DRA  IHSRLSRQSISTREKLVLTTLDFQKALRGFLPASLRSVNLHKPRDLGW DKI  GGLHEVRQILMDTIQLPAKYPEL FANLPIRQRTGILLYGPPGTGKTLLAGVI  ARESRMNFISVKGPELLSKYIGASEQAVRDIFIRAQA AKPCILFFDEFESIAP  RRGHDNTGV TDRVVNQLLTQLDGV EQLQGVYVLAATSRPDLIDPALLRP  GRLDKCVYCPPPDQVSRLEILNVLSDSLPLADDVDLQHVASVTDSFTGAD  LKALLYNAQLEALHGMLLSSGLQDGSSSSDSDLSSSMVFLNHSSGSDDS  AGDGECGLDQSLV SLEMSEILPDESKFNMYRLYFGSSYESELGNGTSSDLS  SQCLSAPSMTQDLPGVPGKDQLFSQPPVLR T ASQEGCQELTQEQRDQLR  ADISI IKGRYRSQSGEDES MNQPGPIKTRLAISQSHLMTALGHTRPSISED D  WKNFAELYESFQNPKR RKNQSGTMFRPGQKVTLA</p>	PEX1
345	<p>MASRKENAKSANRVL RISQLDALELNKALEQLVWSQFTQCFHGFKPGLL  ARFEPEVKACLWVFLWRFTIYSKNATV GQSVLN IYKND FSPNLR YQPPS  KNQKIWYAVCTIGGRWLEERCYDLFRNHHLASFGKVKQCVNFVIGLLKL  GGLINFLIFLQRGKFATLTERLLGIHSVFCKPQNICEVGF EYMNRELLWHG  FAEFLIFLLPLINVQKLKAKLSSWCIPLTGAPNSDNTLATSGKECALCGEW  PTMPHTIGCEHIFCYFCAKSSFLFDVYFTCPKCGTEVHSLQPLKSGIEMSEV  NAL</p>	PEX2
346	<p>MLRSVWNFLKRHKK KCIFLGTVLGGVYILGKYGQKKIREIQEREA AEYIA  QARRQYHFESNQR TCNMTVLSMLPTLREALMQQLNSES LTALLKNRPSN  KLEIWEDLKIISFTRSTVA VYSTCMLVLLRVQLNIIGGYIYLDNAAVGKN  GTILAPPDVQQQYLSSIQHLLGDGLTELITVIKQAVQKVLG SVSLKHSLSL  LDLEQKLKEIRNLVEQH KSSSWINKDGSKPLLCHYMM PDEETPLAVQAC  GLSPRDITTIKLLNETRDMLESPDFSTVLNTCLNRGFSRLLDNMAE FFRPTE  QDLQHGNMNSLSSVSLPLAKIPIVNGQIHSVCSETPSHFVQDLLTMEQV  KDFAANVYEA FSTPQLEK</p>	PEX3

10

20

30

40

50

347	MAMRELVEAECGGANPLMKLAGHFTQDKALRQEGLRPGPWPPGAPASE AASKPLGVASEDELVAEFLQDQNAPLVSRAPQTFKMDDLAEQQIEQS NFRQAPQRAPGVADLALSENWAQEFLLAAGDAVDVTQDYNEDWSQEFIS EVTDPLSVSPARWAEYLEQSEEKLVLGEPEGTATDRWYDEYHPEEDLQ HTASDFVAKVDDPKLANSEFLKFVVRQIGEGQVSLESGAGSGRAQAEQWA AEFIQQGTSDAWVDQFTRPVNTSALDMEFERAKSAIESDVDFWDKLQA ELEEMAKRDAEAHPWLSYDDLTSAFYDKGYQFEEENPLRDHPQPFEEG LRRLEQEGDLPNAVLLFEAAVQQDPKHMEAWQYLGTTQAENEQELLAISA LRRCLELKPNDQNTALMALAVSFTNESLQRQACETLRDWLRYTPAYAHLV TPAEEGAGGAGLGPSKRILGSLSDSLFLEVKELFLAAVRLDPTSIDPDVQ CGLGVLFNLSGEYDKAVDCFTAALSVRPNLYLLWNKLGATLANGNQSEE AVAAARRALELQPGYIRSRYNLGISCINLGAHREAVEHFLEALNMQRKSR GPRGEGGAMSENIWSTLRLALSMLGQSDAYGAADARDLSTLLTMFGLPQ	PEX5
348	MALAVLRVLEFPPTETPPLAVLLPPGGPWPAELGLVLALRPAGESPAGP ALLVAALEGPDAGTEEQGPGPPQLLVSALLRLLALGSGAWVRARAVRR PPALGWALLGTSLGPGLPVRRVRRGETLPVPGPRVLETRPALQGLL GPGTRLAVTELGRARLCPESGDSRPPPPVSSFAVSGTVRRLQGVLLGG TGDSLGVSRSLRGLGLFQGEVWVWAQARESSNTSQPHLARVQVLEPRW DLSDRLGPGSGPLGEPLADGLALVPATLAFNLGCDPLEMGELRIQRYLEG SIAPEDKGCSSLPGPPFARELHIEIVSSPHYSTNGNYDGVLYRHFQIPRVV QEGDVLCPVTIGQVEILEGSPEKLPRWREMFVKVKKTVGEAPDGPASAYL ADTHTSLYMGVSTLSPVPWLPSEESTLWSSLSPGLEALVSELCAVLKPR LQPGGALLTGTSVLLRGPFGCKTTVVAACSHLGLHLLKVPSSLCAE SSGAVETKLQAFSRARRCRPAVLLLTAVDLLGRDRDGLGEDARVMAVL RHLLLNEDPLNSCPPLMVVATTSRAQDLPADVQTAFPHELEVPALSEGQR LSILRALTAHLPLGQEVNLAQLARRCAGFVVGDLYALLTHSSRAACTRIK NSGLAGGLTEEDEGELCAAGFLLAEDFGQALEQLQTAHSQAVGAPKIPS VSWHDVGGVLEVKKEILETIQLPLEHPELLSLGLRRSGLLLHGGPGTGKTL LAKAVATECSLTFVSVKPELINMYVGGQSEENVREVFARARAAAPCIFFD ELDSLAPSRGRSGDSGGVMDRVVSQLLAELDGLHSTQDVFVIGATNRPDL LDPALLRPGFRDKLVFVGANEDRASQLRVLSAITRKFLEPSVSLVNVLD CCPPQLTGADLYSLCSDAMTAALKRRVHDLEEGLEPGSSALMLTMDLL QAAARLQPSVSEQELLYKRIQRKFAAC	PEX6
349	MAPAAASPPEVIRAAQKDEYYRGGLRSAAGGALHSLAGARKWLEWRKE VELLSDVAYFGLTTLAGYQTLGEEYVSIHQVDPSRIHVPSLRRGVLVTLH AVLPYLLDKALLPLEQELQADPDSGRPLQGSGLPGGRGCSGARRWMRHH TATLTEQQRALLRAVFLRQGLACLQRLHVAWFYIHGVFYHLAKRLTG ITYLRVRSPLGEDLRARVSYRLLGVISLLHLVLSMGLQLYGFRRQRARK EWRLHRGLSHRRASLEERAVSRNPLCTLCEERRHPTATPCGHLFCWECI TAWCSSKAECPLCREKFPQKLIYLRHYR	PEX10

10

20

30

40

50

350	MAEHGAHFTAASVADDQPSIFEVVAQDSLMTAVRPALQHVVKVLAESNP THYGFLWRWFDEIFTLLDLLQQHYLSRTSASFSENFYGLKRIVMGDTHK SQRLASAGLPKQQLWKSIMFLVLLPYLKVKLEKLVSSLREEDEYSIHPPSS RWKRFYRAFLAAYPFVNMAWEGWFLVQQLRYILGKAQHHSPLLRLAGV QLGRLTVQDIQALEHKPAKASMMQPPARSVSEKINSALKKAVGGVALSL STGLSVGVFFLQFLDWWYSSSENQETIKSLTALPTPPPPVHLDYNSDSPLLP KMKTVCPLCRKTRVNDTVLATSQYVFCYRCVFHYVRSHQACPITGYPTE VQHLIKLYSPEN	PEX12
351	MASQPPPPKPKWETRRIPGAGPGPGPPTFQSADLGPLMTRPGQPALTRV PPPILPRPSQQTGSSSVNTRPAYSSFSYGYGAYGNSFYGGYSPYSYGYNG LGYNRLRVDDLPPSRFVQQAEESSRGAFQSIESIVHAFASVMMMDATFS AVYNSFRAVLVDVANHFSLKIHFTKVFSAFALVRTIRYLYRRLQRMGLR RGSENEHLWAESEGTVAACGAEDRAATSAKSWPIFFAVILGGPYLIWK LLSTHSDEVTDSINWASGEDDHVVARAEYDFAAVSEEEISFRAGDMLNLA LKEQQPKVRGWLLASLDGQTTGLIPANYVKILGKRKGRKTVESKVKSKQ QQSFTNPTLTKGATVADSLDEQEAFAFESVFVETNKVPVAPDSIGKDGEKQ DL	PEX13
352	MASSEQAEQPSQPSSTPGSENVLPREPLIATAVKFLQNSRVRQSPLATRR FLKKKGLTDEEIDMAFQQSGTAADEPSSLGPATQVVPVQPPHLISQPYSPA GSRWRDYGALAIIMAGIAFGFHQLYKYYLLPLILGGREDRQKLERMEAGL SELSGSVAQTVTQLQTTLASVQELLIQQQKIQELAHELAAAKATTSTNW ILESQNINELKSEINSLKGLLLNRRQFPSPSPAPKIPSWQIPVKSPSPSSPAV NHHSSDISPVSNESTSSSPGKEGHSPEGSTVTYHLLGPQEEGEGVVDVKG QVRMEVQGEEEKREDKEDEEEDDDVSHVDEEDCLGVQREDRRGGDG QINEQVEKLRPEGASNESERD	PEX14
353	MEKLRLGLRYQEYVTRHPAATAQLETAVRGFSYLLAGRFADSHSELSEL VYSASNLLVLLNDGILRKELRKKLPVLSLQKLLTWLSVLECVFVFMEM GAAKVWGEVGRWLVIALLVQLAKAVLRMLLLLWFKAGLQTSPPIVPLDRE TQAQPPDGDHSPGNHEQSYVGKRSNRVVRTLQNTPSLHSRHWGAPQRE GRQQQHHEELSATPTPLGLQETIAEFLYIARPLLHLLSLGLWGQRSWKPW LLAGVVDVTSLSLLSDRKGLTRRERRELRRRTILLLYLLRSPFYDRFSEA RILFLLQLLADHVPVGLVTRPLMDYLPTWQKIYFYSWG	PEX16
354	MAAAEEGCSVGAADRELEELLESALDDFDKAKPSPAPPSTTTAPDASGP QKRSPGDTAKDALFASQEKFFQELFDSELASQATAEFEKAMKELAESEPH LVEQFQKLSEAAGRVGSDMTSQQEFTSCLKETLSGLAKNATDLQNSSMSE EELTKAMEGLGMDEGDGEGNLPIMQSIMQNLLSKDVLYPSELKEITEKYP EWLQSHRESLPPEQFEKYQEHSVMCKICEQFEAETPTDSETTQKARFEM VLDLMQQLQDLGHPKELAGEMPPGLNFDLALNLSGPPGASGEQCLIM	PEX19

10

20

30

40

50

355	<p>MKSDSSTSAAPLRGLGGPLRSSEPVRVAVPARAPAVDLLEEAADLLVVHLD  FRAALETCEAWQSLANHAVAEEPAGTSLEVKCSLVCVVGIQALAEEMDRW  QEVLSWVLQYYQVPEKLPKVLLELCILLYSKMQEPGAVLDVVGAWLQDP  ANQNLPEYGALAEFHVQRVLLPLGCLSEAEELVVGSAAFGEERRLDVLQ  AIHTARQQKQEHSGSEEAQKPNLEGSVSHKFLSLPMLVRQLWDSAVSH  FFSLPFKKSLLAALILCLLVVRFDPASPSSLHFLYKLAQLFRWIRKAAFSRL  YQLRIRD</p>	PEX26
356	<p>MALQGISVVELSGLAPGPFCAMVLADFGARVVRVDRPGSRYDVSRLGRG  KRSVLVDLQKPRGAAVLRRLCKRSVDLLEPFRRGVMKQLQGPILQREN  PRLIYARLSGFGQSGSFCRLAGHDINYLAISGVLKIGRSGENPYAPLNL  ADFAGGGLMCALGIIMALFDRTRTGKGQVIDANMVEGTAYLSSFLWKTQ  KLSLWEAPRGQNMLDGGAPFYTTYRTADGEFMAVGAIEPQFYELLIKGL  GLKSDLPNQMSMDDWPEMKKKFADVFAEKTAEWCQIFDGTDACVTP  VLTFFEEVVHHDHNERGSFITSEEQDVSPRPAPLLNTPAIPSFKRDPFIGE  HTEEILEEFGFSREEIYQLNSDKIIESNKVKASL</p>	AMACR
357	<p>MAQTPAFDKPKVELHVHLDGSIKPETILYYGRRRGIALPANTAEGLLNVIG  MDKPLTLPDFLAKFDYYMPAIAAGCREAIKRIAYEFVEMKAKEGVVYVEV  RYSPELLANSKVEPIPWNQAEGDLTPDEVVALVGQGLQEGERDFGVKAR  SILCCMRHQPNWSPKVVELCKKYQQQTVVAIDLADGETIPGSSLLPGHVQ  AYQEAVKSGIHRTVHAGEVGSAEVVKEAVDILKTERLGHGYHTLEDQAL  YNRLRQENMHFEICPWSSYLTGAWKPDTEHAVIRLKNQANYSLNDDP  LIF  KSTLTDYQMTKRDMGFTEEEFKRLNINAAKSSFLPEDEKRELLDLYKA  YGMPPSASAGQNL</p>	ADA
358	<p>MAAGGDHGSPPSYRSPLASRYASPEMCFVFSRDYKFRTRWRQLWLWLAE  AEQTLGLPITDEQIQEMKSNLENIDFKMAAEEKRLRHVMAHVHTFGH  CCPKAAGIIHLGATSCYVGDNTDLIILRNALDLLLLPKLARVISRLADFAKE  RASLPTLGFTHFQPAQLTTVGKRCCLWIQDLCMDLQNLKRVRRDDLFRG  VKGTTGTQASFLQLFEGDDHKVEQLDKMVTEKAGFKRAFIITGQTYTRK  VDIEVLSVLASLGASVHKICTDIRLLANLKEMEPEFEKQIGSSAMPYKRN  PMRSERCCSLARHMLTLVMDPLQTASVQWFERTLDDSANRRICLAEAF  TADTILNLTQNISEGLVVYPKVIERRIRQELPFMATENIIMAMVKAGGSRQ  DCHEKIRVLSQQAASVVKQEGGDNDLIERIQVDAYFSPIHSQLDHLLDPSS  FTGRASQQVQRFLEEEVYPLLKPYESVMKVKAELCL</p>	ADSL

10

20

30

40

50

359	<p>MNVRIFYSVSQSPHSLLSLLFYCAILESRI SATMPLFKLP AEEKQIDDAMRN  FAEKVFASEVKDEGGRQEISPFDVDEICPISHHEM QAHI FHLET LSTSTEAR  RKRKFQGRKTVNLSIPLSETSSTKLSHIDEYISSSPTYQTVPDFQRVQITGD  YASGVTVEDFEIVCKGLYRALCIREKYMQKSFQRFPKTPSKYLRNIDG EA  WVANESFYPVFTPPVKKGEDPFRTDNL PENLGYHLKMKDGVVYVYPNE  AAVSKDEPKPLYPNLD TFLDDMNFL LALIAQGPVKTYTHRRLKFLSSKF  QVHQMLNEMDELKELKNNPHRDFYNCRKVDTHIHAAACMNQKHLLRFI  KKS YQIDADR VVYSTKEKNL TLKELFAKLKMHPYDLTVDSL DVHAGRQT  FQRFDKFNDKYNPVGASELRDLYLKT DNYINGEYFATI IKEVGADLVEAK  YQHAEPRLSIYGRSPDEWSKLSSWFVCNRIHCPNMTWMIQVPRIYDVFRS  KNFLPHFGKML ENIFMPVFEATINPQADPELSVFLKHIT  GFDSVDDESKHSGHMFSSKSPKPQEW TLEKNPSYTY YAYMYANIMVLN  SLRKERGMNTFLFRPHCGEAGAL THLMTAFMIADDISHGLNLK KSPVLQY  LFFLAQIPIAMSP LSNNSLFLEYAKNPFDFLQKGLMISLSTDDPMQFHFTK  EPLMEEY AIAAQVFKLSTCDMCEVARN SVLQCGISHEEKVKFLGDNYLEE  GPAGNDIRRTNVAQIRMAYRYETWCYELN LIAEGLKSTE</p>	AMPD1
360	<p>MATEGMILTNHDHQIRVGVLTVSDSCFRNLAEDRSGINLKDLVQDPSLLG  GTISAYKIVPDEIEIEIKETLIDWCDEKELNLILTTGGTGFAPRDVTPEATKE  VIEREAPGMALAMLMGSLNVTPLGMLSRPVCGIRGKTLIINLPGSKKGSQ  ECFQFILPALPHAIDL RDAIVKVKEVHDELEDLSPPPP LSPPTTSPHKQT  EDKGVQCEEEEEEEKKDSGVASTEDSSSSHITAA AIAAKIPDSIISRGVQVLP  RDTASLSTTPSESPRAQATSRLSTASCPTPKVQSRCSSKENILRASHSAVDI  TKVARRHRMSPFPLTSM DKA FITVLEMPVLGTEIIN YRDGMGRVLAQDV  YAKDNLPPFPASVKDGYAVRAADGPGDRFIIGESQAGEQPTQTVMPGQV  MRVTTGAPIPCGADAVVQVEDTELIRESDDGTEELEV RILVQARPGDIRP  IGHDIKRGECVLAKGTHMGPSEIGLLATVGVTEVEV NKFPVVAVMSTGNE  LLNPEDDLLPGKIRDSNRSTLLATI QEHGYPTINLGIVGDNPDLLNALNE  GISRADVIITSGGVSMGEKDYLKQVLDIDLHAQIHFGRVFMKPLPTTFAT  LDIDGVRKIIFALPGNPVSAVVTCNL FVVPALRKMQGILDPRPTI IKARLSC  DVKLDPREYHRCILTW HHQEPLWAQSTGNQMSSRLMSMRSANGLLM  LPPKTEQYVELHKGEVVDVMVIGRL</p>	GPHN

10

20

30

40

50

361	MAGAAAESGRELWTFAGSRDPSAPRLAYGYGPGSLRELREFFSRLAGT VYLDHAGATLFSQSQLESFTSDLMENTYGNPHSQNISSKLTHDTVEQVRY RILAHFHTTAEDYTVIFTAGSTAALKLVAEAFPWVSQGPESGSRFCYLT SHTSVVGMNRVNTMAINVISTPVRPEDLWSAEERSASASNPDCQLPHLFCY PAQSNFSGVRYPLSWIEEVKSGRLHPVSTPGKWFVLLDAASYVSTSPLDLS AHQADFVPISFYKIFGFPTGLGALLVHNRAAPLLRKYFGGGTASAYLAG EDFYIPRQSVARFEDGTISFLDVIALKHGFDTLERLTGGMENIKQHTFTL AQYTYVALSSLQYPNGAPVVRIYSDSEFSSPEVQGPINFVLDKGNIIIGY SQVDKMASLYNIHLRTGCFCNTGACQRHLGISNEMVRKHFQAGHVCGD NMDLIDGQPTGSRISFGYMSTLDDVQAFLRFIIDLRLHSSGDWPVPOAH ADTGETGAPSADSQADVIPA VMGRRSLSPQEDALTGSRVWNNSSSTVNAV PVAPPVCDVARTQPTPSEKAAGVLEGALGPHVVTNLYLYPIKSCAAFEVT RWPVGNQGLLYDRSWMVVNHNGVCLSQKQEPRLCLIQPFIDLRQRIMVI KAKGMEPIEVPLEENSERTQIRQSRVCADR VSTYDCGEKISSWLSTFFGRP CHLIKQSSNSQRNAKKKHGKDQLPGTMATLSLVNEAQYLLINTSSILELH RQLNTSDENGKEELFSLKDL SLRFRANIINGKRAFEEKWEISIGSLRFQ VLGPCHRCQMIDQQTGQRNQHV FQKLSERETKVNFGMYLMHASLDL SSPCFLSVGSQVLPVLKENVEGHDLPASEKHQDVTS	MOCOS
362	MAARPLSRMLRLLRSSARSCSSGAPVTQPCPGESARAASEEVSRRRQFL REHAAPFSAFLTDSFGRQHSYLRLISL TEKCNLRCQYCMPEEGVPLTPKANL LTTEEILTLARLFVKEGIDKIRLTGGEPLIRPDVVDIVAQLQRLEGLRTIGVT TNGINLARLLPQLQKAGLSAINISLDTLVPKFEFIVRRKGFHKVMEGIHK AIELGYNPVKVNCVVMRGLNEDEL DFAALTEGLPLDVRFIEYMPFDGN KWNFKKMVSYSKEMLDTVRQQWPELEKVP EEESSTAKAFKIPGFQGGQISFI TSMSEHFCGTCNRLRITADGNLKVCLFGNSEVSLRDHLRAGASEQELLRII GAAVGRKKRQHAGMFSISQMKNRPMILIELFLMFPNSPPANPSIFSWDPLH VQGLRPRMSFSSQVATLWKGCRVPQTPPLAQQLGSGSFQRHYTSRADS DANSKCLSPGSWASAAPSGPQLTSEQLTHVDSEGRAAMVDVGRKPDTER VAVASAVVLLGPVAFKL VQQNQLKKGDALVVAQLAGVQAAKVTSQLIP LCHHVALSHIQVQLELDSTRHAVKIQASCRARGPTGVEMEAL TSAAVAA LTLYDMCKAVSRDIVLEEIKLISKTTGGQRGDFHRA	MOCS1
363	MENGYTYEDYKNTAEWLLSHTKHRPQVAIICGSGLGGLTDKLTQAQIFD YGEIPNFPRSTVPGHAGRLVFGFLNGRACVMMQGRFHMYEGYPLWKVTF PVRV FHLLGVDTLVVTNAAGGLNPKFEVGDIMLIRDHINLPFGSQNPLR GPNDERFGDRFPAMSDAYDR TMRQALSTWKQMGEQRELQEGTYVMV AGPSFETVAECRVLQKLGADAVGMSTVPEVIVARHCGLRVFGFSLITNKV IMDYESLEKANHEEVLAAGKQAAQKLEQFVSILMASIPLPKAS	PNP

10

20

30

40

50

364	<p>MTADKLVFFVNGRQVVEKNADPETLLAYLRRKLGKLSGKLGCGEGGC  GACTVMLSKYDRLQNKIVHFSANACLAPICSLHHVAVTTVEGIGSTKTRL  HPVQERIAKSHGSQCGFCTPGIVMSMYTLLRNQPEPTMEEIENAFQGNLC  RCTGYRPILQGFRTFARDGGCCGGDGNPNCCMNQKDHVSLSPLFKP  EEFTPLDPTQEPFPELLRLKDTPRKQLRFEGERV TWIQASTLKELDLKA  QHPDAKLVVGNTEIGIEMKFKNMLFPMIVCPAWIPELNSVEHGPDGISFGA  ACPLSIVEKTLVDAVAKLPAQKTEVFRGVLEQLRWFAGKQVKSVASVGG  NIITASPIDLNPVFMASGAKLTLVSRGTRRTVQMDHTFFPGYRKTLLSPE  EILLSIEIPYSREGEYFSAFKQASRREDDIAKVTSGMRVLFKPGTTEVQELA  LCYGGMANRTISALKTTQRQLSKLWKEELLQDVCAGLAEELHLPDPAPG  GMVDFRCTLTLSTFFFKFYLTVLQKLGQENLEDKCGKLDPTFASATLLFQK  DPPADVQLFQEV PKGQSEEDMVG RPLPHLAADMQASGEAVYCDDIPRYE  NELSLRLVTSTRAHAKIKSIDTSEAKKVPGFVCFISADDVPGSNITGICNDE  TVFAKDKVTCVGHII GAVVADTPEHTQRAAQGVKITYEELPAITIEDAIKN  NSFYGP ELKIEKGD LKKGFSEADNVVSGEIIYIGGQEHFYLETHCTIAV PKG  EAGEMELFVSTQNTMKTQSFVAKMLGVPANRIVVRVKRMGGGFGGKET  RSTVVSTAVALAAYKTGRPVR CMLDRDEDMLITGGR  HPFLARYKVGF MKTGTVV ALEVDFHFSNVGNTQDLSQSIMERALFHMDNC  YKIPNIRGTGRLCKTNLPSNTAFRGGFPQGMLIAECWMSEVAVTCGMPA  EEVRRKNLYKEGDLTHFNQKLEGFTLPRCWEECLASSQYHARKSEVDKF  NKENCWKKRGLCHIPTKFGISFTVPFLNQAGALLHVYTDGSVLLTHGGTE  MGQGLHTKMVQVASRALKIPTSKIYISETSTNTVPNTSPTAASVSADLNG  QAVYAACQTILKRLEPYKKNPSGSWEDWVTAAYMDTVSLSATGFYRTP  NLGYSFETNSGNPFHYFSYGVACSEVEIDCLTGDHKNLRTDIVMDVGSSL  NPAIDIGQVEGAFVQGLGLFTLEELHYSPEGLHTRGPSTYKIPAFGSIPIEF  RVSLLRDCPNKKA IYASKAVGEPPLFLAASIFFAIKDAIRAARAQHTGNNV  KELFRLDSPATPEKIRNACVDKFTTLCVTGVPENCKPWSVRV</p>	XDH
365	<p>MLLLHRAVVLRLQACRLKSIPSRICIQACSTNDSFQPQRPSLTFSGDNSST  QGWRVMGTLGLGAVLAYQDHCRAAQESTHIYTKEEVSSHTSPETGIW  VTLGSEVFDVTEFVDLHPGGPSKMLLAAGGPLEPFWALYAVHNQSHVRE  LLAQYKIGELNPEDKVAPT VETS DPYADDPVRHPALKVNSQRPFNAEPPP  ELLTENYITPNPIFFTRNHL PVPNLDPDTYRLHVVGAPGGQSLSLSLDDLH  NFPRYEITVTLQCAGNRRSEMTQVKEVKGLEWRTGAISTARWAGARLCD  VLAQAGHQLCETEAHVCFEGLSDPTGTAYGASIP LARAMDPEAEVLLA  YEMNGQPLPRDHGFVVRVVVPGVVGARHV KWLGRVSVQPEESYSHWQR  RDYKGFSPSVDWETVDFDSAPSIQELPVQSAITEPRDGETVESGEVTIKGY  AWSGGGRAVIRVDVSLDGGLTWQVAKLDGEEQRPRKAWAWRLWQLKA  PVPAGQKELNIVCKAVDDGYNVQPDTVAPIWNLRGVLSNAWHRVHVYV  SP</p>	SUOX

10

20

30

40

50



366	<p>MFHLRTCAAKLRPLTASQTVKTFSSQNRPAARTFQQIRCYSAPVAAEPFL  SGTSSNYVEEMYCAWLENPKSVHKSWDIFFRNTNAGAPPGTAYQSPLPLS  RGS�AAVAHAQSLVEAQPNVDKLVEDHLAVQSLIRAYQIRGHHVAQLDP  LGILDADLDSSVPADIISSTDKLGFYGLDESDLDKVFHLPTTTFFIGGQESAL  PLREIIRLEMAYCQHIGVEFMFINDLEQCQWIRQKFETPGIMQFTNEEKR  TLLARLVRSTRFEEFLQRKWSSEKRFGLGCEVLIPALKTIIDKSSENGVDY  VIMGMPHRGRLNVLANVIRKELEQIFCQFDSKLEAADEGSGDVKYHLGM  YHRRINRVTDNRNITLSLVANPSHLEAADPVVMGKTKAEQFYCGDTEGKK  VMSILLHGDAAFAGQGIVYETFHLSLDLPSYTTHTGTVHVVVNNQIGFTTDP  RMARSSPYPTDVARVVNAPIFHVNSDDPEAVMYVCKVAAEWRSTFHKD  VVVDLVCYRRNGHNEMDEPMFTQPLMYKQIRKQKPVLPQKYAELLVSQG  VVNQPEYEEEISKYDKICEEAFARSKDEKILHIKHVLDSPWPGFFTLDGQP  RSMSCPSTGLTEDILTHIGNVASSVPVENFTIHGGLSRILKTRGEMVKNRT  VDWALAEYMAFGSLLKEGIHIRLSGQDVERGTFSHRHHVLHDQNVDKRT  CIPMNLWPNQAPYTVCNSSLSEYGVLFELGFAMASPVALVWEAQFG  DFHNTAQCIIDQFICPGQAKWVRQNGIVLLLPHGMEGMGPEHSSARPERF  LQMCNDDPDVLPDLKEANFDINQLYDCNWVVVNCSTPGNFFHVLRRQIL  LPFRKPLIIFTPKSLLRHPEARSSFDEMLPGTHFQRVIPEDGPAQNPENVK  RLLFCTGKVYYDLTRERKARDMVGQVAITRIEQLSPFPDLLLKEVQKYP  NAELAWCQEEHKNQGYDYVKPRLRTTISRKPVWYAGRDPAAAPATG  NKKTHLTELQRLDFTAFLDLDVFKNFS</p>	OGDH
367	<p>MVGYDPKPDGRNNTKFQVAVAGSVSGLVTRALISPFDVIRFQLQHERL  SRSDPSAKYHGILQASRQILQEEGPTAFWKGHVPAQILSIGYGAVQFLSFE  MLTELVHRGSVYDAREFSVHFVCGGLAACMATLTVHPVDVLRTRFAAQ  GEPKVYNTLRHAVGTMYRSEGPQVFYKGLAPTLIAIFPYAGLQFSCYSSL  KHLYKWAIPAEGKKENLQNLCSGAGVISKTLTYPLDLFKKRLQVGG  FEHARAAFGQVRRYKGLMDCAKQVLQKEGALGFFKGLSPSLLKAALSTG  FMFFSYEFFCNVFCMNRTASQR</p>	SLC25A1 9

10

20

30

40

50

368	<p>MASATAAAARRGLGRALPLFWRGYQTERGVYGYRPRKPESREPQGALER  PPVDHGLARLVTVYCEHGHKAAKINPLFTGQALLENVPEIQALVQTLQGP  FHTAGLLNMGKEEASLEEVLVYLNQIYCGQISIETSQSQDEKDWFAKR  FEELQKETFTTEERKHLKMLLESQEFDFHFLATKVFSTVKRYGGEGAESMM  GFFHELLKMSAYSGITDVIIGMPHRGRNLNLTGLLQFPPELMFRKMRGLSE  FPENFSATGDVLSHLTSSVDLYFGAHHPLHVTMLPNPSHLEAVNPVAVGK  TRGRQQSRQDGDYSPDNSAQPGDRVICLQVHGDAFCGQGIVPETFTLSN  LPHFRIGGSVHLIVNQLGYTTPAERGRSSLYCSDIGKLVGCAIIHVNGDSP  EEVVRATRLAFEYQRQFRKDVIIDLLCYRQWGHNELDEPFYTNPIMYKIIR  ARKSIPDTYAEHLIAGGLMTQEEVSEIKSSYYAKLNDHLNNMAHYRPPAL  NLQAHWQGLAQPEAQITTWSTGVPLDLLRFVGMKSVEVPRELQMHSHLL  KTHVQSRMEKMMDGIKLDWATAEALALGSLLAQGFNVRLSGQDVGRGT  FSQRHAIVVCQETDDTYIPLNHMDPNQKGFLEVSNSPLSEEAVLGFHEYGM  SIESPKLLPLWEAQGFDFNGAQIIFDTFISGGAEKWLQSGIVILLPHGYD  GAGPDHSSCRIERFLQMCDSAEEDVDGDTVMNMFVVHPTTPAQYFHLRR  QMVRNFRKPLIVASPKMLLRPAAVSTLQEMAPGTTFNPIVIGDSSVDPKK  VKTLVFCSGKHFYSLVKQRESLGAKKHDFAIIRVEELCPFPLDSLQQEMSK  YKHVKDHIWSQEEPQNMGPWSFVSPRFEKQLACKLRLVGRPPLPVPVAV  GIGTVHLHQHEDILAKTFA</p>	DHTKD1
369	<p>MASALSYVSKFKSFVILFVTPLLLLPLVILMPAKFVRCAYVIILMAIYWCTE  VIPLAVTSLMPVLLFPLFQILDSRQVCVQYMKDTNMLFLGGLIVAVAVER  WNLHKRIARLTLWVGAKPARLMLGFMGVTALLSMWISNTATTAMMVP  IVEAILQQMEATSAAATEAGLELVDKGKAKELPGSQVIFEGPTLGQQEDQE  RKRLCKAMTLCICYAASIGGTATLTGTGPNVLLGQMNELFPDSKDLVNF  ASWFAFAFPNMLVMLLFAWLWLQFVYMRFNFKKSWGCGLESKKNEKA  ALKVLQEEYRKLGPLSFAEINVLCIFLLVILWFSRDPGFMPGWLTVAWV  EGETKYVSDATVAIFVATLLFIVPSQKPKFNFRSQTEEERKTPFYPPPLLDW  KVTQEKVPWGVILLGGGFALAKGSEASGLSVWMGKQMEPLHAVPPAAI  TLILSLLVAVFTECTSNVATTTLFLPIFASMSRSIGLNPLYIMLPCTLSASFA  FMLPVATPPNAIVFTYGHCLKVADMVKTGVIMNIIGVFCVFLAVNTWGRAI  FDLDHFPDWANVTHIET</p>	SLC13A5
370	<p>MYRALRLLARSRPLVRAPAAALASAPGLGGAAVPSFWPPNAARMASQNS  FRIEYDTFGELKVPNDKYYGAQTVRSTMNFKIGGVTERMPTPVIKAFGILK  RAAAEVNQDYGLDPKIANAIMKAADEVAEGKLNDFHPLVWVWQTGSGTQ  TNMNVNEVISNRAIEMLGELGSKIPVHPNDHVNKSQSSNDTFPTAMHIA  AAIEVHEVLLPGLQKLHDALDAKSKEFAQIHKIGRTHQTDAVPLTLGQEF  GYVQVQKYAMTRIKAAAMPRIYELAAGGTAVGTGLNTRIGFAEKVAAKV  AALTGLPFVTAPNKFEALAAHDALVELSGAMNTTACSLMKIANDIRFLGS  GPRSLGELILPENEPGSSIMPGKVNPTQCEAMTMVAAQVMGNHVAVTV  GGSNGHFELNVFKPMMIKNVLHSARLLGDASVSFTENCVVGIQANTERIN  KLMNESLMLVTALNPHIGYDKAAKIAKTAHKNGSTLKETAIELGYLTAEQ  FDEWVKPKDMLGPK</p>	FH

10

20

30

40

50

371	MWRVCARRAQNVPWAGLEARWTALQEVPGTPTVTSRSGPAPARRNSV TTGYGGVRLALCGWTPSSGATPRNRLLLQLLGSPPGRRYSLPPHQKVLPS LSPTMQAGTIARWEKKEGDKINEGDLIAEVETDKATVGFESLEECYMAKI LVAEGTRDVPVIGAIICITVVGKPEDIEAFKNYTLDSAAPTPQAAPAPTPAAT ASPPTPSAQAPGSSYPHMQVLLPALSPMTMTMGTVQRWEKKVGEKLESEG DLLAEIETDKATIGFEVQEEGYLAKILVPEGTRDVPLGTPLCIIVEKEADISA FADYRPTTEVTDLKPQVPPPTPPVAAVPPTPQPLAPTPSAPCPATPAGPKG RVFVVSPLAKKLAVEKGIDLTQVKGTGPDGRITKDDIDSFVPSKVAPAPAA VVPPTGPGMAPVPTGVFTDIPISNIRRVIAQRLMQSKQTIPHYYLSIDVNM GEVLLVRKELNKILEGRSKISVNDFIKASALACLKVPANSSWMDTVIRQ NHVVDVSVAVSTPAGLITPIVFNHAIKGVETIANDVVSLATKAREGKLP HEFQGGTFTISNLGMFGIKNFSAINPPQACILAIASEDKLVPADNEKGF VASMMSVTLSCDHRVVDGAVGAQWLAEFRKYLEKPITMLL	DLAT
372	MAGALVRKAADYVRSKDFRDYLMSTHFWGPVANWGLPIAAINDMKKSP EIIISGRMTFALCCYSLTFMRFAVKVQPRNWLLFACHATNEVAQLIQGGRL IKHEMKTASA	MPC1
373	MRKMLAAVSRVLSGASQKPARSVLVAARNFANDATFEIKKCDLHRLEEG PPVTTVLTREDGLKYRMMQTVRRMELKADQLYKQKIIRGFCHLCDGQE ACCVGLEAGINPTDHLITAYRAHGFTFTRGLSVREILAEALTGRKGGCAKG KGGSMHMYAKNFYGGNGIVGAQVPLGAGIALACKYNGKDEVCLTYGD GAANQQQIFEAYNMAALWKLPCIFICENNRYSMGTSVERAAASTDYKR GDFIPGLRVDGMDILCVREATRFAAAACRSKGKPILMELQTYRYHGHSMS DPGVSYRTREEIQEVRSKSDPIMLLKDRMVNSNLASVEELKEIDVEVRKEI EDAAQFATADPEPPLEELGYHIYSSDPPFEVRGANQWIKFKSVS	PDHA1
374	MAAVSGLVRRPLREVSGLLKRRFHWTPAALQVTVRDAINQGMDEELER DEKVFLGEEVAQYDGAAYKVSRLWKKYGDKRIIDTPISEMFGFAGIAGV AAMAGLRPICFMTFNFSMQAIDQVINSAAKTYYSMSGGLQPVPVIFRGP GASAGVAAQHSQCFAAWYGHCPGLKVVSPWSEDAKGLIKSAIRDNNP VVVLENELMYGVPFEPPEAQSKDFLIPIGKAKIERQGTHITVVSHSRPVG HCLEAAAVLSKEGVECEVINMRTIRPMDMETIEASVMKTNHLVTVEGGW PQFGVGAEICARIMEGPAFNFLDAPAVRVTGADVMPYAKILEDNSIPQV KDIIFAIKKTLNI	PDHB
375	MAASWRLGCDPRLRLRYLVGFPGRRSVGLVKGALGWSVSRGANNRWFH STQWLRGDPKILMPSLSPTMEEGNIVKWLKKEGEAVSAGDALCEIETDK AVVTLASDDGILAKIVVEEGSKNIRLGSLIGLIVEEGEDWKHVEIPKDV PPPPVSKPSEPRPSEPQISIPVKKKEHIPGTLRFRLSPAARNILEKHS LDASQG TATGPRGIFTKEDALKLVQLKQTGKITESRPTAPTAPTAPSPLQ ATAGPS YPRPVIPPVSTPGQPNVGTFTTEIPASNIRRVIAKRLTESKSTV PHAYATAD CDLGAVLKVRQDLVKDDIKVSVNDFIKA AAVTLKQMPDV NVSWDGEGP KQLPFIDISVAVATDKGLLTPIIKDAAAKGIQEIADSVK ALSCKKARDGKLLP EEYQGGSFISNLGMFGIDEFTA VINPPQACILAV GRFRPV LKLTEDDEEGNA KLQQRQLITVTMSSDSRVVDEDELAT RFLKSFKANLENPIRLA	PDHX

10

20

30

40

50

376	MPAPTQLFFPLIRNCELSRIYGTACYCHHKHLCCSSSYIPQSRLRYTPHPAY ATFCRPKENWWQYTQGRRYASTPQKFYLTTPQVNSILKANEYSFKVPEFD GKNVSSILGFDSNQLPANAPIEDRRSAATCLQTRGMLLGVFDFGHAGCAC QAVSERLFYYIAVSLLPHETLLEIENAVESGRALLPILQWHKHPNDYFSKE ASKLYFNLSRITYWQELIDLNTGESTDIDVKEALINAFKRLDNDISLEAQVG DPNSFLNYLVLRFVAFSGATACVAHVVDGVDLHVANTGDSRAMLGVQEED GSWSAVTLSNDHNAQNERELERLKLHHPKSEAKSVVKQDRLLGLLMPFR AFGDVKFKWSIDLQKRVIESGPDQLNDNEYTKFIPPNYHTPPYLTAPEV YHRLRPQDKFLVLATDGLWETMHRQDVVRIVGEYLTGMHHQQPIAVGG YKVTLGQMHGLLTERRTKMSSVFEDQNAATHLIRHAVGNNEFGTVDHER LSKMLSLPEELARMYRDDITHIVVQFNHSHVVGAYQONQE	PDP1
377	MLEKFCNSTFWNSSFLDSPEADLPLCFEQTVLVWIPLGYLWLLAPWQLLH VYKSRTKRSSTTKLYLAKQVFGVGLLILAAIELALVLTEDSGQATVPAVR YTNPSTYLGTWLLVLLIQYSRQWCQKNSWFLSLFWILSILCGTFQFQTLI RTLLQGDNSNLAYSCLFFISYGFQILILIFSAFSENNESSNNPSSIASFLSSITY SWYDSIILKGYKRPLTEDVWEVDEEMKTKTLVS KFETHMKRELQKARRALQRRQEKSSQNSGARLPGLNKNQSQSDALVL EDVEKKKKSGTKKDVPKSWLMKALFKTFYMVLLKSFLKLVNDIFTFV SPQLLKLISFASDRDYLWIGYLCAILLFTAALIQSFCLQCYFQLCFKLG KVRTAIMASVYKKALTLNLSARKEYTVGETVNLMSVDAQKLMVDVTFM HMLWSSVLQIVLSIFFLWRELGPSVLAGVGMVVLVIPINAILSTKSKIQV KNMKNKDKRLKIMNEILSGIKILKYFAWEPSFRDQVQNLKRELKLLAF S QLQCVVIFVFQTPVLVSVVTFVSVYVLDVSNLILDAQKAFTSITLNFILRFP LSMLPMMISSMLQASVSTERLEKYLGGDDLDTSAIRHDCNFDKAMQFSE ASFTWEHDSEATVRDVLNDIMAGQLVAVIGPVGSGKSSLISAMLGEMEN VHGHTIKGTTAYVPQQSWIQNGTIKDNILFGTEFNEKRYQQVLEACALLP DLEMLPGGDLAIEGKGINLSGGQKQRISLARATYQNLDIYLLDDPLSAV DAHVGKHIFNKVLGPNGLLKGKTRLLVTHSMHFLPQVDEIVVLGNGTIV EKGSYSALLAKKGEFAKNLKTFLRHTGPEEEATVHDGSEEDDDYGLISS VEEIPEDAASITMRRENSFRRTLRSRSRNGRHLKSLRNSLKTRNVNSLKE DEELVKGQKLIKKEFIETGKVKFSIYLEYLQAIGLFSIFFIILAFVMNSVAFI GSNLWLSAWTSDSKIFNSTDYPASQRDMRVGVYVYALGLAQQGIFVFAHF WSAFGFVHASNILHKQLLNNILRAPMRFFDTTPTGRIVNRFAGDISTVDDT LPQSLRSWITCFLGHSTLVMICMATPVFTIIVIPLGHYVSVQMFYVSTSRQL RRLDSVTRSPIYSHFSETVSGLPVIRAFEHQQRFLKHNEVRIDTNQKCVFS WITSNRWLAIARLELVGNLTVFVSALMMVIYRDTLSGDTVGFVLSNALNIT QTLNWLVRMTSEIETNIVAVERITEYTKVENEAPWVTDKRPDPWPSKGGK IQFNQYQVRYRPELDLVLRGITCDIGSMEKIGVVGRTGAGKSSLTNCLFRI LEAAGGQIHDGVDIASIGLHDLREKLTIPQDPILFSGSLRMNLDPFNNYS EEIWKALELAHLKSFVSLQLGLSHEVTEAGGNLSIGQRQLCLGRALLR KSKILVLDEATAAVDLETNLIQTIIQNEFAHCTVITIAHRLHTIMDSK MVLNNGKIIECGSPEELLQIPGPFYFMAKEAGIENVNSTKF	ABCC2

10

20

30

40

50

378	MDQNQHNLNKTAEAQPSSENKKTRYCNGLKMFLAALSLSFIAKTLGAIIMKS SIIHIERRFEISSSLVGFIDGSFEIGNLLVIVFVSYFGSKLHRPKLIGIGCFIMGI GGVLTALPHFFMGYYRYSKETNINSENSTSTLSTCLINQILSLNRASPEIV GKGCKESGSYMWIYVFMGNMLRGIGETPIVPLGLSYIDDFAKEGHSSLY LGILNAIAMIGPIIGFTLGLSFLSKMYVDIGYVDLSTIRITPTDSRWVGAWWL NFLVSGLFSIISSIPFFFLPQTPNKPQKERKASLSLHVLETNDEKDQTANLTN QGKNITKNVTGFFQSFKSILTNPPLYVMFVLLTLLQVSSYIGAFTYVFKYVE QQYGQPSSKANILLGVITIPFASGMFLGGYIHKFKLNTVGIKAFSCFTAV MSLSFYLLYFFILCENKSVAGLTMTYDGNPNVTSHRDVPLSYCNSDCNCD ESQWEPVCGNNGITYISPCLAGCKSSSGNKKPIVFYNCSCLEVTGLQNRNY SAHLGECPRDDACTRKFYFFVAIQVLNLFSSALGGTSHVMLIVKIVQPELK SLALGFHSMVIRALGGILAPIYFGALIDTTCIKWSTNCGTRGSCRTYNSTS FSRVYLGLSSMLRVSSLVLYIILYAMKKKYQEKDINASENGSVMDEANL ESLNKNKHVPSAGADSETHC	SLCO1B1
379	MDQHQLNKTAEASSEKKTTRRCNGFKMFLAALSFSYIAKALGGIIMKI SITQIERRFDISSLAGLIDGSFEIGNLLVIVFVSYFGSKLHRPKLIGIGCLLM GTGSILTSLPHFFMGYYRYSKETHINPENSTSSLSTCLINQTLFNGTSPEI VEKDCVKESGSHMWIYVFMGNMLRGIGETPIVPLGISYIDDFAKEGHSSLY YLGLSNAIGMIGPVIGFALGSLFAKMYVDIGYVDLSTIRITPKDSRWVGA WWLGFLVSGLFSIISSIPFFFLPKPNPKQKERKISLSLHVLTNDDRNTA NLTNQGKNVTKNVTGFFQSLKSILTNPPLYVIFLLLTLQVSSFIGSFTYVFK YMEQQYQGSASHANFLLGHIITPTVATGMFLGGFIIKKFKLSLVGIKAFSFL TSMISFLFQLLYFPLICESKSVAGLTLTYDGNNSVASHVDVPLSYCNSECN CDESQWEPVCGNNGITYLSPCLAGCKSSSGIKKHTVFYNCSCVEVTGLQN RNYSAHLGECPRDNTCTRKFFIYVAIQVINSLSFATGGTTFILLTVKIVQPE LKALAMGFQSMVIRTLGGILAPIYFGALIDKTCMKWSTNSCGAQGACRIY NSVFFGRVYLGLSIALRFPALVLYIVFIFAMKKKFQKDKTKASDNERKVM DEANLEFLNNGEHFVPSAGTDSKTCNLDMQDNAAAN	SLCO1B3
380	MGEPGQSPSPRSSHGSPTLSTLTLTLLLLCGHAHSQCKILRCNAEYVSSTLS LRGGGSSGALRGGGGGGRRGGVGSGLCRALRSYALCTRRTARTCRGD LAFHSAVHGIEDLMIQHNC SRQGPTAPPPRGPALPGAGSGLPAPDPCDYE GRFSRLHGRPPGFLHCASFGDPHVRSFHHHFHTCRVQGAWPLLDNDFLV QATSSPMALGANATATRKLTIIFKNMQECIDQKVYQAEVDNLPVAFEDGS INGGDRPGGSSLSIQTANPGNHVEIQAAAYIGTTIIRQTAGQLSFSIKVAEDV AMAFSAEQDLQLCVGGCPPSQRLSRSENRNRGAIITIDTARRLCKEGLPVE DAYFHSCVFDVLISGDPNFTVAAQALEDARAFLPDLEKLHLFPSDAGVP LSSATLLAPLLSGLFVLWLCIQ	HFE2

10

20

30

40

50

381	<p>MHQRHPRARCPLCVAGILACGFLGCVGSPSHFQQSCLQALEPQAVSSYL          SPGAPLKGRPPSPGFQRQRQRQRAAGGILHLELLVAVGPDVFAQHQEDT          ERYVLTNLNIGAELLRDPSLGAQFRVHLVKMVILTEPEGAPNITANLTSSL          LSVCGWSTINPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTQLGGA          CSPTWWSCLITEDTGFDLGVITIAHEIGHSFGLHEDGAPGSGCGPSGHVMASD          GAAPRAGLAWSPCSRRLSLLSAGRARCVDPPRPQPGSAGHPPDAQP          GLYYANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSCSRL          LVPLLDGTECGVEKWCSKGRCSRSLVELTPIAAVHGRWSSWGPRSPCSRSC          GGGVVTRRRQCNNPRPAFGGRACVGADLQAEMCNTQACEKTQLEFMSQ          QCARTDQPLRSSPGGASFYHWGAAVPHSQGDALCRHMCRAIGESFIMK          RGDSFLDGTRCMPSGPREDDGTLSCVSGSCRTFGCDGRMDSQQVWDRQC          VCGGDNSTCSPRKGSFTAGRAREYVTFITVTPNLTSVYIANHRPLFTHLA          VRIGGRYVAVAGKMSISPNTTYPSSLEDDGRVEYRVALTEDRLPRLEEIRIWG          PLQEDADIQVYRRYGEEYGNLTRPDITFTYFQPKPRQAWVWAAVRGPCS          VSCGAGLRWVNYSCLDQARKELVETVQCQGSQPPAWPEACVLEPCPPY          WAVGDFGPCSASCGGLRERPVRCVEAQGSLLKTLPPARCRAQAQPAV          ALETCPNPQPCPARWEVSEPSSTASAGGAGLALENETCVPGADGLEAPVTE          GPGSVDEKLPAPPCVGMSCPPGWGHLDATSAGEKAPSPWGSIRTGAQA          AHVWTPAAGSCSVSCGRGLMELRFLCMDALRVPVQEELCGLASKPGSR          REVCQAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGEEILL          DTQCQGLRPEPEQACSLEPCPPRWKVMISLGPCSASCGLGTARRSVACVQ          LDQGDVEVDEAACAALVRPEASVPCLADCTYRWHVGTWMECSVSCG          DGIQRRRDTCLGPQAQAPVPADFCQHLPKPVTVRGCWAGPCVGGQTPSL          VPHEEAAAPGRTTATPAGASLEWSQARGLLFSPAPQPRLLPGPQENSQVQ          SSACGRQHLEPTGTIDMRGPGQADCAVAIGRPLGEVTLRVLESSLNCSA          GDMLLLWGRLTWRKMCRLKLLDMTFSSKTNLTVRQRCGRPGGGVLLRY          GSQ LAPETFYRECDMQLFGPWGEIVSPSLSPATSNAGGCRLFINVAPHARI          AIHALATNMGAGTEGANASYILIRDTHSLRTTAFHGQQVLYWESESSQAE          MEFSEGFLKAQASLRGQYWTLSWVPEMQDPQSWKKEGT</p>	ADAMTS 13
-----	--	--------------

10

20

30

40

50

382	MSRPLSDQEKRKQISVRGLAGVENVTELKKNFNRHLHFTLVKDRNVATP RDYYFALAHTVRDHLVGRWIRTQQHYEYKDPKRIYYLSLEFYMGRTLQN TMVNLALENACDEATYQLGLDMEELIEEEDAGLGNGGLGRLAACFLDS MATLGLAAYGYGIRYEFGLFNQKISGGWQMEEADDWLRYPGPNWEKARP EFTLPVHFYGHVEHTSQGAKWVDTQVVLAMPYDTPVPGYRNNVVNTMR LWSAKAPNDFNLKDFNVGGYIQAVLDRNLAENISRVLYPNDNFFEGKEL RLKQEYFVVAATLQDIIRRFKSSKFGCRDPVRTNFDAFPDKVAIQLNTHP SLAIPELMRILVDLERMDWDKAWDVTVRTCAYTNHTVLPEALERWPVHL LETLLPRHLQIYEINQRFLNRVAAAFPGDVDRLRRMSLVEEGAVKRINM AHLCIAGSHAVNGVARIHSEILKKTIFKDFYELEPHKFQNKTNGITPRRWL VLCNPGLAEVIAERIGEDFISDLQDLRKLKLSFVDDEAFIRDVAKVKQENKL KFAAYLEREYKVHINPNSLFDIQVKRIHEYKRQLLNCLHVITLYNRIKREP NKFFVPRVTVMIGGKAAPGYHMAKMIIRLVTAIGDVVNHDPAVGDRLRVIF LENYRVSLAEKVIPAADLSEQISTAGTEASGTGNMKFMLNGALTIGTMDG ANVEMAEAGEENFFIFGMRVEDVDKLDQRGYNAQEYYDRIPELRQVIE QLSSGFFSPKQPDLFKDIVNMLMHHDRFKVFADYEDYIKCQEKVSALYK NPREWTRMVIRNIATSGKFSSDRTIAQYAREIWGVESRQRLPAPDEAI	PYGM
-----	--	------

10

20

30

40

50

383	<p>MLSFVDTRTL LLLAVTLCLATCQSLQEETVRKGPAGDRGPRGERGPPGPP  GRDGEDGPTGPPGPPGPPGPPGLGGNFAAQYDQKGVGLGPGPMGLMGPR  GPPGAAGAPGPQGFQGPAGEPEPGQTGPAGARGPAGPPGKAGEDGHPG  KPRPGERGVVGPQGARGFPGTPLPGFKGIRGHNGLDGLKGQPGAPGV  KGEPGAPGENGTGQTGARGLPGERGRVVGAPGPAGARGSDGSGVGPVGA  GPIGSAGPPGFPAPGPKGEIGAVGNAGPAGPAGPRGEVGLPGLSGPVGPP  GNPGANGLTGAKGAAGLPGVAGAPGLPGPRGIPGPVGAAGATGARGLV  GEPGPAGSKGESGNKGEPGSAGPQGPSPSGEEGKRGPNGEAGSAGPPGP  PGLRGSPPSRGLPGADGRAGVMGPPSGRGASGPAGVRGPNGDAGRPEP  GLMGPRGLPGSPGNIPAGKEGPVGLPGIDGRPGPIGPAGARGEPGNIGFP  GPKGPTGDPGKNGDKGHAGLAGARGAPDPGNNGAQGPPGPQGVQGG  KGEQGGPPGPPGFQGLPGSPGAGEVVKPGERGLHGEFGLPGPAGPRGERG  PPGESGAAGPTGPIGSRGSPGPPGPDGNKGEPGVVGA VGTAGPSGSPGLPG  ERGAAGIPGGKGEKGEPLRGEIGNPGRDGARGAPGAVGAPGPAGATGD  RGEAGAAGPAGPAGPRGSPGERGEVGPAGPNGFAGPAGAAGQPGAKGE  RGAKGPKGENGVVGTGPVGAAGPAGPNGPPGPAGSRGDGGPPGMTGFP  GAAGRTGPPGSPGISGPPGPPGAGKEGLRGPRGDQGPVGRTEVGA VGP  PGFAGEKGPSGEAGTAGPPGTGPQQLL GAPGILGLPGSRGERGLPGVAG  AVGEPGPLGIAGPPGARGPPGAVGSPGVNGAPGEAGRDGNPNDGPPGR  DGQPGHKGERGYPGNIPVGAAGAPGHPGVGPAGKHG NRGETGPSGPV  GPAGAVGPRGPSGPQGIRGDKGEPGEKGRGLPGLKGHNGLQGLPGIAGH  HGDQGAPG SVGPAGPRGPAGPSGPAGKDGRTGHPGTVPAGIRGPQGHQ  GPAGPPGPPGPPGPPGVSGGGYDFGYDGD FYRADQPRSAPSLRPKDYEVD  ATLKS LNNQIETLLTPEGSRKNPARTCRDLRLSHPEWSSGYYWIDPNQGC  TMDAIKVYCDFSTGETCIRAQPENIPAKNWYRSSKDKKHVWLGETINAGS  QFEYNVEGVTSKEMATQLAFMRLLANYASQNITYHCKNSIAYMDEETGN  LKKA VILQGSNDVELVAEGNSRFTYTVLVDGCSKKTNEWGKTHIEYKTNK  PSRLPFLDIAPLDIGGADQEFFVDIGPVC FK</p>	COL1A2
384	<p>MNLLCCALVFLDISIKWTTQETFPKY LH YDEETSHQLL CDKCPPGTYL  KQHCTAKWKTVCAPCPDHYYTDSWHTSDECLY CSPVCKELQYVKQECN  RTHNRVCECKEGRYLEIEFCLKHRSCPPGFVVQAGTPERN TVCKRCPDG  FFSNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNICSGNSESTQKCGIDV  TLCEEAFFRFAVPTKFTPNWLSVLVDNLPGTKVNAESVERIKRQHSSQEQ  TFQLLKLWKHQNKDQDIVKKI IQDIDL CENSVQRHIGHANLTFEQLRSLM  ESLPGKKVGAEDIEKTIKACKPSDQILKLLSLWRIKNGDQDTLKGLMHAL  KHSKTYHF PKTVTQSLKKTIRFLHSFTMYKLYQKLFLEMIGNQVQSVKISC  L</p>	TNFRSF1 1B

10

20

30

40

50



385	MAQQANVGELLAMLDSPLGVRDDVTAVFKENLNSDRGPMLVNTLVD YYLETSSQPALHILTTLQEPHDKHLLDRINEYVGKAATRLSILLSLLGHVIRL QPSWKHKLSQAPLLPSLLKCLKMDTDVVVLTTGVLVLITMLPMIPQSGKQ HLLDFFDIFGRLSSWCLKKPGHVAEVYLVHLHASVYALFHRLYGMYPNCN FVSFLRSHYSMKENLETFEVVVKPMMEHVRIHPELVTGSKDHELDPRRW KRLETHDVVIECAKISLDPTEASYEDGYSVSHQISARFPHRSADVTTSPYA DTQNSYGCATSTPYSTSRLMLLNMPGQLPQTLSSPSTRRLITEPPQATLWSP SMVCGMTTPTPTSPGNVPPDLSPYKVFVGTAGGKGTPLGTPATSPPPAPL CHSDDYVHISLPQATVTPRKEERMDSARPCLHRQHLLNDRGSEPPGS KGSVTLSDLPGFLGDLASEEDSIEKDKEEAISRELSEITTAEAEPVPRGG FDSPFYRDSLPGSQRKTHSAASSSQGASVNPEPLHSSL DKLGPDTPKQAFTPIDLPCGSADESPAGDRECQTSLETSIFTPSPCKIPPTR VGFSGSQPPPYDHLFEVALPKTAHHFVIRKTEELLKAKGNTEEDGVPST SPMEVLDRLIQQGADAHSKELNKLPLPSKSVDWTHFGGSPPSDEIRTLRD QLLLHNQLLYERFKRQQHALRNRRLLRKVIKAAALEEHNAAMKDQLKL QEKDIQMWKVSLSQKEQARYNQLQEQRDTMVTKLHSQIRQLQHDREEFY NQSQELQTKLEDCRNMIANELRIELKKANNKVCHTELLLSQVSQKLSNSES VQQQMEFLNRQLLVLGEVNELYEQLNKHSDTTKEVEMMKAA YRKEL EKNRSHVLQQTQRLDTSQKRILELESHLAKKDHLLLEQKKYLEDVKLQA RGQLQAAESRYEAQKRITQVFELEILDLYGRLEKDGLLKKLEEEKAEAAE AAEERLDCCNDGCSDSMVGHNEEASGHNGETKTPRPSSARGSSGSRGGG GSSSSSELSTPEKPPHQ RAGPFSSRWETTMGEASASIPTTVGSLPSSKSFLG MKARELFRNKSESQCDEDGMTSSLSLKTTELKDLGVEAKIPLNLDGPH PSPPTPDSVGQLHIMDYNETHHEHS	TSC1
-----	--	------

10

20

30

40

50

386	<p>MAKPTSKDSGLKEKFKILLGLGTPRPNRSAEGKQTEFIITAEILRELSMEC  GLNNRIRMIGQICEVAKTKKFEHVEALWKAVADLLQPERPLEARHAV  LALLKAIVQGQGERLGVLRALFFKVIKDYP SNEDLHERLEVFKALTDNGR  HITYLEEELADFLVQWMDVGLSSEFLLVLVNLVKFN SCYLDEYIARMVQ  MICLLCVRTASSVDIEVSLQVLDAVVCYNCLPAESLPLFIVTLCRTINVKEL  CEPCWKLMRNLLGTHLGHSAIYNMCHLMEDRAYMEDAPLLRGAVFFVG  MALWGAHRLYSLRNSPTSVLPSFYQAMACPNEVVSYEIVLSITRLIKKYR  KELQVVAWDILLNIIERLLQQLQTLDSPELRTIVHDLTTVEELCDQNEFH  GSQERYFELVERCADQRPESSLLNLISYRAQSIHPAKDGIQNLQALMERF  FRSESRGAVRIKVLVLSFVLLINRQFYEEELINSVVISQLSHIPEDKDHQV  RKLATQLLVDLAEGCHTHHFNSLLDIIKVMARSLSPPELEERDVAAYSA  SLEDVKTAVLGLLVILQTKLYTLPASHATRVEMLVSHIQLHYKHSYTLPI  ASSIRLQAFDFLLLLRADSLHRLGLPNKDGVVRFSPYCVCDYMEPERGSE  KKTSGPLSPPTGPPGPAPAGPAVRLGSPYSLFRVLLQCLKQESDWKVL  KLVGLRPELRYKVLIFTSPCSVDQLCSALCSMLSGPKTLERLRGAPEGF  SRTDLHLAVVPVLTALISYHNYLDKTKQREMVYCLEQGLIHRCASQCVV  ALSICSVEMPDIIKALPVLVVKLTHISATASMAVPLLEFLSTLARLPHLYR  NFAAEQYASVFAISLPYTNPSKFNQYIVCLAHHVIAMWFIRCLPFRKDFV  PFITKGLRSNVLLSFDDTPEKDSFRARSTSLNERPKSLRIARPPKQGLNNSP  PVKEFKESSAAEAFRCRSISVSEHVRSRIQTSLSASLGSADENSVAQAD  DSLKNLHLELTETCLDMMARYVFSNFTAVPKRSPVGEFLLAGGRKTWL  VGNKLVTVTTSVGTGTRSLLGLDSGELQSGPESSSSPGVHVRQTKEAPAK  LESQAGQQVSRGARDRVRSMGGHGLRVGALDVPASQFLGSATSPGPRT  APAAKPEKASAGTRVPVQEKTNLAAAYVPLL TQGWA EILVRRPTGNTSWL  MSLENPLSPFSSDINN MPLQELSNALMAAERFKEHRDTALYKSLVPAAS  TAKPPPLPRSN TVASFSSLYQSSCQQLHRSVSWADSAVVMEEGSPGEVP  VLVEPPGLEDEVAALGMDRRTDAYSRSVVSSQEEKSLHAEELVGRGPIE  RVVSSEGG RPSVDLSFQPSQPLSKSSSPELQTLQDILGDPGDKADVGR LSP  EVKARSQSGTLDGESAAWSASGEDSRGQPEGPLSSSPRSPGLRPRGYTI  SDSAPSRGKRVERDALKSRATASNAEKVPGINPSFVFLQLYHSPFFGDES  NKPILLPNESQSFERSVQLLDQIPSYDTHKIAVLYVGEGQSNSELA ILSNEH  GSYRYTEFLTGLGRLIELKDCQPKVYLGGLDVCGEDGQFTYCW HDDIM  QAVFHIA TLMPTKDVDKHRCDDK RHLGNDFVSIVYNDSGEDFKLGTIKG  QFNFVHVIVTPLDYECNLVSLQCRKDM EGLVDTSVAKIVSDRNLPFVARQ  MALHANMASQVHHSRSNPTDIYPSKWIARLRHIKRLRQRICEEAAYS NPS  LPLVHPPSHSKAPAQTPAEPTPGYEVGQRKRLISSVEDFTEFV</p>	TSC2
-----	---	------

10

20

30

40

50

387	MAAKSQPNIPKAKSLDGVTDNDRTASQGGQWGRAWEVDWFSLASVIFLLLF APFIVYYFIMACDQYSCALTGPVVDIVTGHARLSDIWAKTPPITRCAAQLY TLWVTFQVLLYTSPLDFCHKFLPGYVGGIQEGAVTPAGVVNKYQINGLQ AWLLTHLLWFANAHLSSWFSPTIIFDNWIPLLCANILGYAVSTFAMVKG YFFPTSARDCKFTGNFFYNYMMGIEFNPRIGKWFDFKLFNNGRPGIVAWT LINLSFAAKQRELHSHVTNAMVLVNVLQAIYVIDFFWNETWYLKTIDICH DHFGWYLGWDCVWLPYLYTLQGLYLVYHPVQLSTPHAVGVLLGLVG YYIFRVANHQDLFRRTDGRCLIWGRKPKVIECSYTSADGQRHHSKLLVS GFWGVARHFNYVGDLMGSLAYCLACGGGHLLPYFYIHYMAILLTHRCLR DEHRCASKYGRDWERYTAAVPYRLLPGIF	DHCR7
388	MSLSNKLTLDKLDVKGKRVMRVDNFVPMKNNQITNNQRIKAAVPSIKF CLDNGAKSVVLMShLGRPDGVPMPDKYSLEPVAVELKSLLGKDVFLKD CVGPEVEKACANPAAGSVILLENLRFHVEEEGKGDASGNKVKAEPAKIE AFRASLSKLGDVYVNDAFGTAHRAHSSMVGNLQKAGGFLMKKELNY FAKALESPPERPFLAILGGAKVADKIQLINMLDKVNEMIIGGGMAFTFLK VLNNMEIGTSLFDEEGAKIVKDLMSKAENGVKITLPVDFVTADKFDENA KTGQATVASGIPAGWMGLDCGPESKKYAEA VTRAKQIVWNGPVGVE WEAFARGTKALMDEVVKATSRGCITIIGGGDTATCCAkwntEDKVSHVS TGGGASLELLEGGVLPVDALSNI	PGK1
389	MGTSALWALWLLLALCWAPRESGATGTGRKAKCEPSQFQCTNGRCITLL WKCDGDEDCVDGSDEKNCVKKTCAESDFVCNNGQCVPsrWkCDGDPD CEDGSDESPEQCHMRTCRIHEISCgAHSTQCIPVSWRCDGENDCDSGEDE ENCGNITCSPDEFTCSSGRcISRNFVCNGQDDCSDGSDELDCAPPTCGAHE FQCSTSSCIPISWVCDDDADCSdQsDESLEQCGRQpVIHTKCPASEIQCGSG ECIHKKWRCDGDPDCKDGSDEVNcPSRTRCPDQFECEDGSCIHGSRQCNG I RDCVDGSDEVNCKNVNQCLGPGKFKCRSGECIDISKVCNQEQDCRDWSD EPLKECHINECLVNNGGCShICKDLVIGYECDAAGFELIDRKTCDGIDEC QNPgICSQICINLKGgYKCECSRGYQMDLATGVCKAVGKEPSLIFTNRRDI RKIGLERKEYIQLVEQLRNTVALDADIAAQKLFWADLSQKAIFsASIDDKV GRHVkMIDNVYNpAAIAVDWVYKTIYWTDAASKTISVATLDGTRKkFLF NSDLREPASIAVDPLSGFVYWSdWGEPAKIEKAGMNGFDRRPLVTADIQ WPNGITLDLIKSRLYWLdSKLHMLSSVDLNGQDRRIVLKSLEFLAHPLAL TIFEDRVYWIDGENEA VYGANKFTGSELATLVNNLNDaQDIIVYHELvQP SGKNWCEEDMENGgCEYLCLPAPQINDHSPKYTCSCPSGYNVEENGRDC QSTATTvTYSETKDTNTTEISATsGLVPGGINVTTAVSEVSVPKGTsAAW AILPLLLLVMaAVGGYLMWRNWQHKNMKSMNFDNPVYLKTTEEDLSID IGRHsASVGHtYPaISVVSTDDDLA	VLDLR

10

20

30

40

50

390	MEPSSLELPADTVQRIAAELKCHPTDERVALHLDEEDKLRHFRECFYIPKI QDLPPVDLSLVNKDENAIYFLGNSLGLQPKMVKTYLEEELDKWAKIAAY GHEVGRKPWITGDESIVGLMKDIVGANEKEIALMNALTVNLHLLMLSFFK PTPKRYKILLEAKAFPSDHYAIESQLQLHGLNIEESMRMIKPREGEETLRIE DILEVIEKEGDSIAVILFSGVHFYTGQHFNIPAITKAGQAKGCYVGFDLAH AVGNVELYLHDWGVDFACWCSYKYLNAGAGGIAGAFIHEKHAHTIKPA LVGWFGHELSTRFKMDNKLQLIPGVCGRISNPPILLVCSLHASLEIFKQAT MKALRKKSVLLTGYLEYLIKHNKGDKAATKKPVVNIITPSHVEERGQCL TITFSVPNKDVFQELEKRGVVCCKRNPNGIRVAPVPLYNSFHDVYKFTNL LTSILDSAETKN	KYNU
391	MFPGCPRLWVLVVLGTSWVGWGSQGTEAAQLRQFYVAAQGISWSYRPE PTNSSLNLSVTSFKKIVYREYEPYFKKEKQSTISGLLGPTLYAEVGDIIKV HFKNKADKPLSIHPQGIRYSKLSGASYLDHTFPAEKMDDAVAPGREYTY EWSISEDGPTHDDPPCLTHIYSHENLIEDFNGLIGPLLICKKGTLEGGT QKTFDKQIVLLFAVFDESKSWSQSSSLMYTVNGYVNGTMPDITVCAHDHI SWHLLGMSSGPELFSIHFNGQVLEQNHKVSATLVSATSTANMTVGPE GKWIISSLTPKHLQAGMQAYIDIKNCPKKTRNLKKITREQRHMKRWEYF IAAEEVIWDYAPVIPANMDKKYRSQHLDNFSNQIGKHYKKVMYTQYEDE SFTKHTVNPMMKEDGILGPIRAQVRDTLKIVFKNMASRPYSIYPHGVTFSP YEDEVNSSFTSGRNNTMIRAVQPGETYTYKWNILEFDEPTENDAQCLTRP YYSVDIMRDIASGLIGLLICKSRSLDRRGIQRAADIEQQAVFAVDENK SWYLEDNINKFCENPDEVKRDDPKFYESNIMSTINGYVPESITTLGFCFDD TVQWHFCSVGTQNEILTIHFTGHSFIYGKRHEDTLTLFPMRGESVTVTMD NVGTWMLTSMNSSPRSKKLRKFRDVKCIPDDDEDSYEIFEPPESTMAT RKMHDRLEPEDEESDADYDYQNRLAAALGIRSNSSLNQEENLNTAL ALENGTEFVSSNTDIIVGSNYSSPSNISKFTVNNLAEPQKAPSHQQATTAGS PLRHLIGKNSVLNNSSTAESHSPYSEDPIEDPLQPDVTGIRLLSLGAGEFKSQ EHAKHKGPKVERDQAAKHRSWMKLLAHKVGRHLSQDTGSPSGMRPW EDLPSQDTGSPSRMRPWKPPSDLLLLKQSNSSKILVGRWHLASEKGSYEI IQDTDEDTAVNNWLISPQNASRAWGESTPLANKPGKQSGHPKFPVRHKS LQVRQDGGKSRLKKSQFLIKTRKKKKEKHTHAPLSPRTFHPLRSEAYNT FSERRLKHSLVLHKSNETSLPTDLNQTLPMSDFGWIASLPDHNQNSSNDT GQASCPPGLYQTVPEEHYQTFPIQDPDQMHSTSDPSHRSSPELSEMLEY DRSHKSFPTDISQMSPESEHEVWQTVISPDLSQVTLSPELSQTNLSPDLSHT TLPSELQIRNLSPALGQMPISPDLSHTTLPDLSHTTSLDLSQTNLSPELSQ TNLSPALGQMPLSPDLSHTTSLDFSQTNLSPELSHMTLSPELSQTNLSPAL GQMPISPDLSHTTSLDFSQTNLSPELSQTNLSPALGQMPLSPDPSHTTSL DLSQTNLSPELSQTNLSPDLSEMPFADLSQIPLTPDLQMTLSPDLGETDL SPNFGQMSLSPDLSQVTLSPDISDTLLPDLISPPDLQIFYPSESSQSL LQEFNESFPYDGLGQMPSPSSPTLNDTFLSKEFNPLVIVGLSKDGTDYIEIIP KEEVQSSSEDDYAEIDYVPYDDPYKTDVRTNINSSRDPDNIAAWYLRNNG NRRNYIAAEEISWDYSEFVQRETDIEDSDDIPEDTTYKKVVFVKYLDSTF	F5

10

20

30

40

50

	<p>TKRDPRGEYEEHLGILGPIIRAEVDDVIQVRFKNLASRPYSLHAHGLSYEK  SSEGKTYEDDSPEWFKEDNAVQPNSSYTYVWHATERSGPESPGSACRAW  AYYSAVNPEKDIHSGLIGPLLICQKGILHKDSNMPMDMREFVLLFMTFDE  KKSWEYEEKSRSSWRLTSSEMKKSHEFHAINGMIYSLPGLKMYEQEWVR  LHLLNIGGSQDIHVVFHFGQTLENGNKQHQLGVWPLLPGSFKTLEMKA  SKPGWLLNTEVGENQRAGMQTPFLIMDRDCRMPMGLSTGIISDSQIKAS  EFLGYWEPRLARLNNGGSYNAWSVEKLAAEFASKPWIQVDMQKEVIITGI  QTQGAKHYLKSCYTTEFYVAYSSNQINWQIFKGNSTRNVMYFNGNSDAS  TIKENQFDPPIVARYIRISPTRAYNRPTLRLELQCEVNGCSTPLGMENGI  ENKQITASSFKKSWWGDYWEPFRARLNAQGRVNAWQAKANNKQWLE  IDLLKIKKITAHTQGCKSLSEMYVKSytiHYSEQVWPKPYRLKSSMVD  KIFEGTNTKGHVKNFFNPPIISRFIRVIPKTWNQSIARLELFGCDIY</p>	
392	<p>MGPTSGPSLLLLLTHLPLALGSPMYSIITPNILRLESEETMVLEAHDAQGD  VPVTVTVHDFPGKLVLSSEKTVLTPATNHMGNTFTIPANREFKSEKGR  NKFVTVQATFGTQVVEKVVLVSLQSGYLFIQTDKTIYTPGSTVLYRIFTVN  HKLLPVGRTVMVNIENPEGIPVKQDSLSSQNQLGVLPLSWDPELVNMGQ  WKIRAYYENSPQQVFSTEFVKEYVLPSEFVIVEPTEKFYYIYNEKGLEVTI  TARFLYGKKVEGTAFVIFGIQDGEQRISLPESLKRIPIEDGSGEVVLRSKVL  LDGVQNPRaedLVGKSLYVSATVILHSGSDMVQAERSGIPIVTSPIYQIHFT  KTPKYFKPGMPFDLMVFTNPDGSPAYRVPVAVQGEDTVQSLTQGDGV  AKLSINTHPSQKPLSITVTRTKKQELSEAEQATRMTQALPYSTVGNSNNYL  HLSVLRTELRPGETLNVNFLLRMDRAHEAKIRYYTYLIMNKGRLLKAGR  QVREPGQDLVVLPLSITDFIPSFRLVAYYTLIGASGQREVVADSVWVDV  KDSCVGLSVVKSQSEDRQPVPQGQMTLKIEGDHGARVVLVAVDKGVF  VLNKKNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFS DAGL TFTSSSGQ  QTAQRAELQCPQPAARRRRSVQLTEKRMdKVGKYPKELRKCCEDGMRE  NPMRFSCQRRTRFISLGEACKKVFLDCCNYITELRRQHARASHLGLARSN  LDEDIIAEENIVSRSEFPESWLWNVEDLKEPPKNGISTKLMNIFLKDSITTW  EILAVSMSDKKGCIVADPFEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLY  NYRQNQELKVRVELLHNP AFCSLATTKR RHQQT VTI PPKSSLSVPYVIVPL  KTGLQEVEVKA AVYHHFISDGVRKSLKVPEGIRMNKTVAVRTLDPERL  GREGVQKEDIPPADLSDQVPDTESETRILLQGTPVAQMTEDAVDAERLKH  LIVTPSGCGEQNMIGMTPTVIAVHYLDETEQWEKFGLEKRQGALELIKKG  YTQQLAFRQPSSAFAAFVKRAPSTWLTAYVVKVFS LAVNLI AID SQVLCG  AVKWLILEKQKPDGVFQEDAPVIHQEMIGGLRNNNEKDMALTAFLVLSLQ  EAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVAIAGYALAQMGR  LKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSYALLALLQLKDFDFVP  PVVRWLNEQRYYG GYGSTQATFMVFQALAQYQKDAPDHQELNLDVSL  QLPSRSSKITHRIHWESASLLRSEETKENEGFTVTAEGKGQGTLSVVTMYH  AKAKDQLTCNKFDLKVTIKAPETEKRPQDAKNTMILEICTRYRGDQDAT  MSILDISMMTGFAPD TDDLKQLANGVD RYISKYELDKAFSDRNTLIYLDK</p>	C3

10

20

30

40

50

	<p>VSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCTRFYHPEKED  GKLNKLCRDELCRCAEENCIFIQSDDKVTLEERLDKACEPGVDYVYKTR  LVKVQLSNDFDEYIMAIEQTIKSGSDEVQVGGQRTFISPIKCREALKLEEK  KHLYLMWGLSSDFWGEKPNLSYIIGKDTWVEHWPEEDECQDEENQKQCQ  DLGAFTESMVVFGCPN</p>	
393	<p>MGPRLSVWLLLLPAALLLHEEHSRAAAKGGCAGSGCGKCDCHGVKGQK  GERGLPGLQGVIQFPGMVGPEPQGGPPGQKGDGTGEPGLPGTKGTRGPPGA  SGYPGNPGLPGIPGQDGGPPGPPGIPGCNGTKGERGPLGPPGLPGFAGNPGP  PGLPGMKGDPGEILGHVPGMLLKGERGFPGIPGTPGPPGLPGLQGPVGGPP  FTGPPGPPGPPGPPGEGKGMGLSFQGPKGDKGDQVSGPPGVPGQAQVQ  EKGDFATKGEKGQKGEFQGMGPGVGEKGEFPGKPGPRGKPGKDGDGKE  KGSPPGFPEPGYPGLIGRQGPQGEKGEAGPPGPPGIVIGTGPLGEKGERGY  PGTPGPRGEPGPKGFPLPGQPGPPGLVPGQAGAPGFPERGERGEKDRGFP  GTSPLGSPGRDGLPGPPGSPGPPGQPGYTNGIVECQPGPPGDQGGPPGIPGQP  GFIGEIGEKGQKGESCLICDIDGYRGGPPGQGGPPGEIGFPGQPGAKGDRGLP  GRDGVAGVPGPQGTPLIGQPGAKGEPGEFYFDLRLKGDKGDPGFPGQP  GMPGRAGSPGRDGHPLPGPKGSPGSLKGERGPPGGVGFPGSRGDTG  PPGPPGYGPAGPIGDKGQAGFPGGPGSPGLPGPKGEPGKIVPLPGPPGAEG  LPGSPGFPGPQDRGFPGTPGRPLGPKGAVGQPGIGFPGPPGPKGVDGL  PGDMGPPGTPGRPFNGLPGNPGVQGGQKGEFVGLPGLKGLPGLPGIPGT  PGEKGSIGVPGVPEHGAIGPPGLQGIRGEPGPPGLPGSVGSPGVPGIGPPG  ARGPPGGQGGPGLSGPPGKGEKGFPGFPGLDMPGPKGDKGAQGLPGITG  QSLPGLPGQQGAPGIPGFPGSKGEMGVMGTPGQPGSPGVPVAPGLPGEK  GDHGFPGSSGPRGDPGLKGDGKDVGLPGKPGSMDKVDMGSMKGQKGD  QGEKQIGPIGEKGSRGDPGTPGVPGKDGQAGQPGQPGKGDGPGISGTPG  APGLPGPKGSVGGMGLPGTPGKGVPGIPGPPGQSPGLPGDKGAKGEKQ  AGPPGIGIPGLRGEKGDQGIAGFPGSPGEGKGEKGSIGIPGMPGSPGLKGS  SVGYPGSPGLPGEKGDGKGLPGLDGIPGVKGEAGLPGTPGTPGAGQKGE  GSDGIPGSAGEKGEPLPGRGFPGFPGAKGDKGSKGEVGFPLAGSPGIPG  SKGEQGMGPPGPPGQGPGLPGSPGHATEGPKGDRGPQGGPGLPGLPGPM  GPPGLPGIDGVKGDGKNPGWPGAPGVPGPKGDPGFQGMPIGGSPGITGS  KGDMPGPPGVPGFQGPGLPGLQGIKGDQGDQGVPGAKGLPGPPGPPGY  DIIKGEPLPGPEGPPGLKGLQGLPGPKGQQGVTGLVGIPGPPGIPGFDGAP  GQKGEAGPAGPTGPRGFPGPPGPDGLPGSMGPPGTPSVDHGFLVTRHSQT  IDDPQCPSGTKILYHGYSLLYVQGNRAHGQDLGTAGSCLRKFSTMPFLF  CNINNVCFASRNDYSYWLSTPEPMPMSMAPITGENIRPFISRCVCEAPA  MVMAVHSQTIQIPPCPSGWSLWIGYSFVMHTSAGAEGSQALASPGSCL  EEFRSAPFIECHGRGTCNYYANAYSFWLATIERSEMFKKPTPSTLKAGELR  THVSRQCVMRRT</p>	COL4A1

10

20

30

40

50

394	MRLAKIICLMLWAICVAEDCNELPPRRNTEILTGSWSDQTYPEGTQAIYK CRPGYRSLGNVIMVCRKGEWVALNPLRKCQKRPCGHPGDTDFGTFTLTG GNVFEYGVKAVYTCNEGYQLLGEINYRECDTDGWTNDIPICEVVKCLPV TAPENGGKIVSSAMEPDREYHFGQAVRFVCNSGYKIEGDEEMHCSDDGFW SKEKPKCVEISCKSPDVINGSPISQKIIYKENERFQYKCNMGYEYSERGDA VCTESGWRPLPSCEEKSCDNPYIPNGDYSPLRIKHRTGDEITYQCRNGFYF ATRGNTAKCTSTGWIPAPRCTLKPCDYPDIKHGGLYHENMRRPYFPVAV GKYYSYCYCDEHFETPSGSYWDHIHCTQDGWSPA VPCLRKCYFPYLENGY NQNYGRKFVQGKSIDVACHPGYALPKAQT TVTCMENGWSPTPRCIRVKT CSKSSIDIENGFISESQYTYALKEKAKYQCKLGYVTADGETSGSITCGKDG WSAQPTCIKSCDIPVFMNARTKNDFTWFKLNDTLDYECHDGYESNTGSTT GSIVCGYNGWSDLPICYERECELPKIDVHLPDRKKDQYKVGEVLKFSCK PGFTIVGPNSVQCYHFGLSPDLPICKEQVQSCGPPPELLNGNVKEKTKEYY GHSEVVEYYCNPRFLMKGPNKIQCVDGEWTTLPVCIVEESTCGDIPELEH GWAQLSSPPYYYGDSVEFNCSSESFTMIGHRSITCIHGVWTQLPQCVAIDKL KKCKSSNLIILEEHLKNKKEFDHNSNIRYRCRGKEGWIHTVCINGRWDPE VNCSMAQIQLCPPPQIPNSHNMTTTLNRYRDGEKVSVLCQENYLIQEGEEI TCKDGRWQSIPLCVEKIPCSQPPQIEHGTINSSRSSQESYAHGTKLSYTCG GFRISEENETTCYMGKWSSPPQCEGLPCKSPPEISHGVVAHMSDSYQYGE EVTYKCFEGFGIDGPAIAKCLGEKWSHPPSCIKTDCLSLPSFENAIPMGEK KDVKYKAGEQVITYCATYYKMDGASNVT CINSRWTGRPTCRDTSCVNPP VQNAYIVSRQMSKYPSGERVRYQCRSPYEMFGDEEVMCLNGNWTEPPQ CKDSTGKCGPPPIDNGDITSFPLSVYAPASSVEYQCQONLYQLEGNKRITC RNGQWSEPPKCLHPCVISREIMENYNIALRWTAQKLYSRTGESVEFVCK RGYRLSSRSHTLRTTCWDGKLEYPTCAKR	CFH
-----	---	-----

10

20

30

40

50

395	<p>MEPRPTAPSSGAPGLAGVGETPSAAAALAAARVELPGTAVPSVPEDAAPAS  RDGGGVRDEGPAAAGDGLGRPLGPTPSQSRFQVDLVSEENAGRAAAAAA  AAAAAAAAAAGAGAGAKQTPADGEASGESEPAKGSEEAKGRFRVNFVDP  AASSAEDSLSDAAGVVDGPNVVSFQNGGDTVLESGSSLHSGGGGGSGH  HQHYYYDTHNTYTLRFTGHNTMDAVPRIDHYRHTAAQLGEKLLRPSLA  ELHDELEKEPFEDGFANGEESTPTRDAVVITYAESKGVVKGWIKGVLVR  CMLNIWGVMLFIRLSWIVGQAGIGLSVLVIMMATVVTTITGLSTSAIATNG  FVRGGGAYYLISRSLGPEFGGAIGLIFAFANAVAVAMYVVGFAETVVVELL  KEHSILMIDEINDIRIIGAITVVILLGISVAGMEWEAKAQIVLLVILLAIGDF  VIGTFIPLESKKPKGFFGYKSEIFNENFGPDFREEETFFSVFAIFFPAATGILA  GANISGDLADPQSAIPKGTLLAILITLVYVGVIAVSVGSCVVRDATGNVND  TIVTEL TNCTSAACKLNFD FSSCESSPCSYGLMNNFQVMSMVSGFTPLISA  GIFSATLSSALASLVSA PKIFQALCKDNIYPAFQMF AKGYGKNNEPLRGIYI  LTF LIALGFILIAELNVI APIISNFFLASYALINFSVFHASLAKSPGWRPAFKY  YNMWISLLGAILCCIVMFVINWWAALLTYVIVLGLYIYVTYKKPDVNWG  SSTQALTYLNLQHSIRLSGVEDHVKNFRPQCLVMTGAPNSRPALLHLVH  DFTKNVGLMICGHVHMGP RRQAMKEMSIDQAKYQRWLIK NKMKA FYA  PVHADDLREGAQYLMQAAGLGRMKPNTLV LGFKDWLQADMRD VDM  YINLFHDAFDIQYGVVIRLKEGLDISHLQGEELLSSQEKSPGTKD VVVS  VEYSKKS DLDTSKPLSEK PITHKVEEEDGKTATQPLLK KESKGPVPLNVA  DQKLEASTQFQKKQ GKNTIDVWWLFDDGGLTLLIPYLLTTKKKWKDCK  IRVFIGGKINRIDHRRAMATLLSKFRIDFS DIMVLGDINTKPKKENIIAFEE  IIEPYRLHEDDKEQDIADKMKED EPWRITDNELELYKTKTYRQIRLNELLK  EHSSTANIIVMSLPVARKGAVSSALYMAWLEALSKDLPPILLVRGNHQSV  LTFYS</p>	SLC12A2
396	<p>MAASKKAVLGPLVGAVDQGTSSSTRFLVFNSKTAELLSHHQVEIKQEFPRE  GWVEQDPKEILHSVYECIEKTCEKLGQLNIDISNIKAIGVSNQRETTVVWD  KITGEPLYNAV VWL DLRTQSTVESLSKRIPGNNNFVKSKTGLPLSTYFSAV  KLRWLLDNVRKVQKAVEEKRALFGTIDSWLIWSLTGGVNGGVHCTDVT  NASRTMLFNIHSLEWDKQLCEFFGIPMEILPNVRSSEIYGLMKISHSVKA  GALEGVPISGCLGDQSAALVGQMC FQIQAKNTYGTGCFLLCNTGHKCV  FSDHGLLTTVAYKLGRDKPVYYALEGSVAIAGAVIRWLRDNLGIKTSEEI  EKLAKEVGTSGCYFVPAFSGLYAPYWEPSARGIICGLTQFTNKCHIAFAA  LEAVCFQTREILDAMNRDCGIPLSHLQVDGGMTSNKILMQLQADILYIPV  VKPSPPETTALGAAMAAGAAEGVGVWSLEPEDLSAVTMERFEPQINAE  SEIRYSTWKKAVMKSMGWTTQSPESGDPSIFCSLPLGFFIVSSMVMLIGA  RYISGIP</p>	GK
397	<p>MDVGSKEVLMESPPDYSAAPRGRFGIPCCPVHLKRLIVVVVVVLIVVVI  VGALLMGLHMSQKHEMVLMSIGAPEAQRLALSEHLVTTATFSIGSTG  LVVYDYQQLLIAYKPAPGTCCYIMKIAPESIPSLEALNRKVHNFQMECSLQ  AKPAVPTSKLGQAEGRDAGSAPSGGDP AFLGMAVNTLCGEVPLYI</p>	SFTPC

10

20

30

40

50



398	MEPGRRGAAALLALLCVACALRAGRAQYERYSFRRSFRDELMPLESAYR HALDKYSGEHWAESVGYLEISLRLHRLLRDSEAFCHRNCSPAAPQPEPAAG LASYPELRLFGLLRRAHCLKRCKQGLPAFRQSQPSREVLADFQRREPYK FLQFAYFKANNLPKAIAAAHTFLLKHPDDEMMKRNMAYYKSLPGAEDYI KDLETKSYESLFIKAVRAYNGENWRTSITDMELALPDFFKAFYECLAACE GSREIKDFKDFYLSIADHYVEVLECKIQCEENLTPVIGGYPVEKVFVATMYH YLQFAYYKLNLDLKNAAPCAVSYLLFDQNDKVMQQNLVYYQYHRDTWG LSDEHFQPRPEAVQFFNVTTLQKELYDFAKENIMDDDEGEVVEYVDDLLE LEETS	CRTAP
399	MAVRALKLLTTLAVVAAASQA EVESEAGWGMVTPDLLFAEGTAAYAR GDWPGVVLSEMERALSRAALRALRLRCRTQCAADFPWELDPDWSPSPA QASGAAALRDLSSFFGLLRRAACLRRCLGPPAAHSLSEEMELEFRKRSPY NYLQVAYFKINKLEKAVAAAHTFFVGNPEHMEMQQNLDYYQTMSGVKE ADFKDLETQPHMQEFRLGVRLYSEEQPEAVPHLEAALQEYFVAYEECR ALCEGPYDYGNYLEYNADLFQAITDHYIQVLNCKQNCVTELASHP SR EKPFEDFLPSHYNYLQFAYYNIGNYTQAVECAKTYLLFFPNDEVMNQNL AYYAAMLGEEHTRSIGPRESAKEYRQRSLLLEKELLFFAYDVFVIPFVDPDS WTPEEVIPKRLQEKQKSERETA VRISQEIGNLMKEIETLVEEKTESLDVS RLTREGGPLLYEGISLTMNSKLLNGSQRVVMGVDGVISDHECQELQRLTNVA ATSGDGYRGQTSPHTPNEKFGVTVFKALKLGQEGKVPLQSAHLYYNVT EKVRRIMESYFRLDTPLYFSYSHLVCRTAIEEVQAERKDDSHPVHVDNCIL NAETLVCVKEPPAYTFRDYSAILYLNDFDGGNFYFTELDKTVTAEVQP QCGRAVGFSSGTENPHGVKAVTRGQRCAIALWFTLDPRHSERDRVQADD LVKMLFSPEEMDLSQEQLDAQQGPPEPAQESLSGSESKPKDEL	P3H1

10

20

30

40

50

400	MTLRLLVAALCAGILAEAPRVRAQHRERVTCTRLYAADIVFLLDGSSSIG RSNFREVRSFLEGLVLPFSGAASAQGVRFATVQYSDDPRTEFGLDALGSG GDVIRAIRELSYKGGNTRTGAAILHVADHVFLPQLARPGVPKVCILITDGK SQDLVDTAAQRLKGGQVKLFAVGIKNADPEELKRVASQPTSDFFFFVNDP SILRLLPLVSRVCTTAGGVPVTRPPDDSTSAPRDLVLEPSSQSLRVQW TAASGPVTGYKVQYTPLTGLGQPLSERQEVNVPAGETSVRLRGLRPLTE YQVTVIALYANSIGEAVSGTARTTALEGEPELTIQNTTAHSLLVAVRSVPG ATGYRVTWRVLSGGPTQQQELGPGQGSVLLRDLEPGTDYEVTVSTLFG SVGPATSLMARTDASVEQTLRPVILGPTSILLSWNLVPEARGYRLEWRRE TGLEPPQKVLPSPDVTRYQLDGLQPGTEYRLTYTLLEGHEVATPATVVP TGPPELVPSPVTDLQATELPGQVRVRSWSPVPGATQYRIIVRSTQGVERTLV LPGSQTAFLDDVQAGLSYTVRVSARVGPREGSASVLTVRREPETPLAVP GLRVVVS DATRVRVAWGVPVPGASGFRISWSTGSGPESSQTLPPDSTATDIT GLQPGTTYQVAVSVLRGREGPAAVIVARTDPLGPVRTVHVTQASSSVT ITWTRVPGATGYRVSWSHAHGPEKSQLVSGEATVAELDGLEPDTEYTVH VRAHVAGVDGPPASVVVRTAPEPVGRVSRQLILNASSDVLRTWVGVTG ATAYRLAWGRSEGGPMRHQILPGNTDSAEIRGLEGGVSYSVRVTALVGD REGTPVSIVVTTPEAPPALGTLHVVRGEHSLRLRWEPVPRAQGFLLHW QPEGGQEQRVLGPELSSYHLDGLEPATQYRVRLSVLGPAGEGPSAEVTA RTESPRVPSIELRVVDTSIDSVTLAWTPVSRASSYILSWRPLRGPGQEVPGS PQTLPGISSQRVTGLEPGVSYIFSLTPVLDGVRGPEASVTQTPVCPRLAD VVFLPHATQDNAHRAEATRRVLERLVLALGPLGPQAVQVGLLSYSHRPS LFPLNGSHDLGILQRIRDMPYMDPSGNNLGTAVVTAHRYMLAPDAPGR QHVPGMVLLVDEPLRGDIFSPIREAQASGLNVMLGMAGADPEQLRRL APGMDSVQTFFAVDDGPSLDQAVSGLATALCQASFTTQPRPEPCPVYCPK GQKGEPEGMGLRGQVGGPDPGLPGRTGAPGPQGGPSATAKGERGFPG ADGRPGSPGRAGNPGTPGAPGLKGGPGLPGRGDPGERGPRGPKGEPGAP GQVIGGEGPGLPGRKGDGPGSGPPGPRGPLGDPGPRGPPGLPGTAMKGD GDRGERGPPGPEGGIAPGEPGLPGLPGSPGPQGPVGGPKKGEKGDSE GAPGLPGQPGSPGEQGRGPPGAIGPKGDRGFPGPLGEAGEKGERGPPG AGSRGLPGVAGRPGAKGPEGPPGPTGRQGEKGEPRGPDPAVVGPAVAG PKGEKGDVGPAGPRGATGVQGERGPPGLVLPGDGPKGDPGDRGPIGLT GRAGPPGDSGPPGEKGDGPRGPPGPVGPGRGRDGEVGEKGDGEGPPGDP PGKAGERGLRGAPVVRGPVGEKGDQGDGDRNGSPGSSGPKGDRGEP GPPGPPGRLVDTGPGAREKGEPRGDRGQEGPRGPKGDPGLPGAPGERGIE FRGPPGPQGDGVRGPAKEKGDGRGPPGLDGRSGLDGKPGAAGPSGPN AGKAGDPGRDGLPGLRGEQGLPGPSGPPGLPGKPGEDGKPLNGKNGEP GDPGEDGRKGEKGDGSGREGRDGPKGERGAPGILGPQGGPGLPGVGP PGQGFPGVGGTGPKGDRGETGSKGEQGLPGERGLRGEVGSVPNVDRLL TAGIKASALREIVETWDESSGFLVPERRRGPKGDSGEQGGPKGEPGIFP GERGLKGDGRDGPQGGPGLALGERGPPGSLAGEPGKPGIPGLPGRAG GYGEAGRPERGERGEKGERGEQGRDGPGLPGTPGPPGPPGPKVSVDEP	COL7A1
-----	--	--------

10

20

30

40

50

	<p>GPGLSGEQPPGLKGAKEPGSNGDQGPKGDRGVPGIKGDRGEPGPRGQ  DGNPGLPGERGMAGPEGKPGLOQPRGPPGPVGGHGDPPGAPGLAGPA  GPQGPSGLKGEGETGPPGRGLTGPTGAVGLPGGPSGLVGPQGSPLPG  QVGETGKPGAPGRDASGKDGDRGSPGVPGSPGLPGVPKGEPTGA  PGQAVVGLPGAKGEKGAPGGLAGDLVGEPEGAKGDRGLPGPRGEKGEAG  RAGEPGDPGEDGQKGAPGPKGFKGDPGVGVPGSPGPPGPPGVKGLGLP  GLPGAPGVVGFPGQTGPRGEMQPGPSGERGLAGPPGREGIPGLGPPGPP  GSVGGASGLKGDKGDPGVGLPGPRGERGEPGIRGEDGRPGQEGPRGLT  GPPGSRGERGEKGDVGSAGLKGDKGDSA VILGPPGPRGAKGDMGERGPR  GLDGDKGPRGDNDPGDKGSKGEPGDKGSAGLPGLRGLLGPQGQPGAA  GIPGDPGSPGKDGVPGIRGEKGDVGFMPRGLKGERGVKGACGLDGEKG  DKGEAGPPGRPGLAGHKGEMGEPGVPGQSGAPGKEGLIGPKGDRGFDGQ  PGPKGDQGEKGERGTPGIGGFPGPSGNDGSAGPPGPPGSPGPRGPEGLQ  QKGERGPPGERVVGAPGVGAPGERGEQGRPGPAGPRGEKGEAALTEDD  IRGFVRQEMSQHCACQGFASGSRPLPSYAADTAGSQLHAVPVLVSHA  EEEERVPEDDEYSEYSEYSVEEYQDPEAPWDSDDPCSLPLDEGSCTAYTL  RWYHRAVTGSTEAHPFVYGGCGGNANRFGTREACERRCPPRVVQSQG  TGTAQD</p>	
401	<p>MSIQENISSLQLRSWVSKSQRD LAKSILIGAPGGPAGYLRRASVAQLTQEL  GTAFFQQQLPAAMADTFLEHLCLLDIDSEPVAARSTSIATIGPASRSVER  LKEMIKAGMNIARLNFSHGSHEYHAESIANVREAVESFAGSPLSYRPVAIA  LDTKGPEIRTGILQGGPESEVELVKGSQVLTVDPAFRTRGNANTVWVDY  PNIVRVVPVGGRIYIDDGLISLVVQKIGPEGLVTQVENGGVLGSRKGVNLP  GAQVDLPLGSEQDVRDLRFGVEHGVDIVFASFVRKASDVAAVRAALGPE  GHGIKIISKIENHEGVKRFDEILEVSDGIMVARGDLGIEIPAENVFLAQKMM  IGRCNLAKPVCATQMLESMITKPRPTRAETSDVANAVLDGADCIMLSG  ETAKGNFPVEAVKMQHAIAREAEA AVYHRQLFEELRRAAPLSRDPTEVT  AIGAVEAAFKCCAAAIVLTTTGRSAQLLSRYRPRAAVIAVTRSAQAARQ  VHLCRGVFPLL YREPPEAIWADDVDRRVQFGIESGKLRGFLRVGDLVIVV  TGWRPGSGYTNIMRVL SIS</p>	PKLR
402	<p>MSSSPVKRQRMESALDQLKQFTTVVADTGDFHAIDEYKPQDATTNPSLIL  AAAQMPAYQELVEEAIAYGRKLGGSQEDQIKNAIDKLFVLFGAEILKKIP  GRVSTEVDARLSFDK DAMVARARRLIELYKEAGISKDRILIKLSSTWEGIQ  AGKELEEQHGHCNMTLLFSFAQAVACAEAGVTLISPFVGRILDWHVANT  DKKSYEPLDPGVKSVTKIYNYKFKFSYKTIVMGASFRNTGEIKALAGCD  FLTISP KLLGELLQDNAKLVVLSAKAAQASDLEKIHLEKSFRLHNED  QMAVEKLSDGIRKFAADAVKLERMLTERMFNAENGK</p>	TALDO1

10

20

30

40

50

403	MRLAVGALLVCAVLGLCLAVPDKTVRWCAVSEHEATKCQSFRDHMKSV IPSDGPSVACVKKASYLDCIRAIAANEADAVTL DAGLVYDAYLAPNNLKP VVAEFYGSKEDPQTFYYAVAVVKKDSGFQMNQLRGKKSCHTGLGRSAG WNIPIGLLYCDLPEPRKPLEKAVANFFSGSCAPCADGTDFPQLCQLCPGCG CSTLNQYFGYSGAFKCLKDGAGDVAFVKHSTIFENLANKADRQYELLC LDNTRKPVDEYKDCHLAQVPSHTVVARSMGGKEDLIWELLNQAQEHFG KDKSKEFQLFSSPHGKDLLFKDSA HGF LK VPPRMDAKMYLGYEYVTAIR NLREGTCPEAPTDECKP VKWCALSHHERLKCEDEWSVNSVVGKIECVSAETTEDCIAKIMNGEADAMS LDGGFVYIAGKCGLPVLAENYNKSDNCEDTPEAGYFAIAVVKKSASDL TWDNLKGGKSCHTAVGRTAGWNIPMGLLYNKINHCRFDEFFSEGCAPGS KKDSSLCKLCMGSGLNLCEPNNKEGYGYTGAFRCLVEKGDVAFVKHQ TVPQNTGGKNPDPWAKNLNEKDYELLCLDGTRKPVVEYANCHLARAPN HAVVTRKDKEACVHKILRQQHLFGSNVTDCSGNFCLFRSETKDLLFRD DTVCLAKLHDRNTYEKYLGEYVKA VGNLRKCSTSSLLEACTFRRP	TF
404	MAPPQVLAFLGLLAAATATFAAAQEEVCENYKLAVNCFVNNNRQCQC TSVGAQNTVICSKLAAKCLVMKAEMNGSKLGRRAKPEGALQNNNDGLYD PDCDESLFKAKQCNGTSMCWCVNTAGVRRTDKDEITCSERVRTYWIII ELKHKAREKPYDSKSLRTALQKEITTRYQLDPKFITSILYENNVITIDL VQN SSQKTQNDVDIADVAYYFEKDVKGESLFHSSKMDLTVNGEQLDLDPGQT LIYYVDEKAPEFSMQGLKAGVIAVIVVVVIAVVAGIVVLVISRKKRMAKY EKAEIKEMGEMHRELNA	EPCAM
405	MPRRAENWDEAEVGAEEAGVEEYGPPEEDGGEESGAEESGPEESGPEELG AEEEMEAGRPRPVLRSVNSREPSQVIFCNRSPRVVLPVWLNFDGEPQPYPT LPPGTGRRHSYRGHLWFRDAGTHDGLLVNQTELFVPSLNV DGQPIFANI TLPVYTLKERCLQVVRSLVKPENYRRLDIVRSLYEDLEDHPNVQKDLERL TQERIAHQRMGD	VHL
406	MKRVLVLLLA VAFGHALERGRDYEKNKVCKEFSLGKEDFTSLSLVLYS RKFPSGTFEQVSQLVKEVSLTEACCAEGADPCYDTRTSALSAKSCESN SPFPVHPGTAECCTKEGLERKLCMAALKHQ PQEFPTYVEPTNDEICEAFR KDPKEYANQFMWEYSTNYGQAPLSLLVSYTKSYLSMVGSCCTSASPTVC FLKERLQLKHL SLLTTLSNRVCSQY AAYGEKKSRLSNLIKLAQKVPTADL EDVLP LAEDITNILSKCESASEDCMAKELPEHTVKLCDNLSTKNSKFEDC CQEKTAMDVFVCTYFMPAAQLPELPDVELPTNKDVCDPGNTKVM DKYT FELSRRTHLPEVFLSKVLEPTL KSLGECCDVEDSTTCFNAKGPLLKKESSF IDKGQELCADYSENTFTEYKKKLAERLKAKLPDATPTELAKLVNKHSDFA SNCCSINSPLYCDSEIDAELKNIL	GC

10

20

30

40

50

407	MPSSVSWGILLLAGLCCLPVSLAEDPQGDAAQKTDTSHHHQDHPTFNKI TPNLAEFAFSLYRQLAHQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILE GLNFNLTEIPEAQIHEGFQELLRTLNPDSQLQLTTGNGLFLSEGLKLVDK FLEDVKKLYHSEAFVNFVGDTEEAQKQINDYVEKGTQGGKIVDLVKELDR DTVFALVNYIFFKQKWERPFVVDTEEEEDFHVDQVTTVKVPMKRLGM FNIQHCKKLSSWVLLMKYLGNAIAIFFLPDEGKLQHLENELTHDIITKFL NEDRRSASLHLPKLSITGTDLKSVLGLGITKVFVSNADLSGVTEEAPLK LSKAVHKAVLTIKDEKTEAAGAMFLEAIPMSIPPEVKFNKPFVFLMIEQNT KSPLFMGKVVNPTQK	SERPINA 1
408	MAAPAEPACAGQGVWNQTEPEPAATSLLSLCLFRTAGVWVPPMYLWVLG PIYLLFIHHHGRGYLRMSPLFKAKMVLGFALIVLCTSSVAVALWKIQGT PEAPEFLIHPTVWLTMSFAVFLIHTERKKGVQSSGVLFYWLLCFVLPAT NAAQQASGAGFQSDPVRHLSTYLCLSLVVAQFVLSCLADQPPFFPEDPQQ SNPCPETGAAPSKATFWVWVSGLVWRGYRRPLRPKDLWSLGRNSSEEL VSRLEKEWMRNRSAARRHNKAIKFRKGGSGMKAPETEPFLRQEGSQW RPLLKAIWQVFHSTFLLGTLSLIISDVFRFTVPKLLSLFLEFIGDPKPPAWKG YLLAVLMFLSACLQTLFEQQNMYRLKVLQMLRSITGLVYRKVLALSS GSRKASAVGDVNLVSVQVRLTESVLYLNLWLPVWVIVCFVYLWQ LLGPSALTAIAVFLSLLPLNFFISKRNHHQEEQMRQKDSRARTSSILRNS KTIKFHGWEGAFLDRVLGIRGQELGALRTSGLLFSVSLVSFQVSTFLVALV VFAVHTLVAENAMNAEKAFVTLTVLNILNKAQAFLPFSIHSVLVQARVSFD RLVTFLCLEEVDPGVVDSSSSGSAAGKDCITIHSATFAWSQESPPCLHRINL TVPQGCLLAVVGPVAGKSSLLSALLGELSKVEGFVSIEGAVAYVPQEA WVQNTSVVENVCFGQELDPPWLERVLEACALQPDVDSFPEGIHTSIGEQG MNLSSGGQKQRLSLARAVYRKAAYLLDDPLAALDAHVGQHVFNQVIGP GGLLQGTTRILVTHALHILPQADWIIVLANGAIAEMGSYQELLQRKGALM CLLDQARQPGDRGEGETEPTSTKDPRTSAGRRPELRRERSIKSVPEKDR TTSEAQTEVPLDDPDRAGWPAGKDSIQYGRVKATVHLAYLRAVGTPLCL YALFLFLCQQVASFCRGYWLSLWADDPVAVGGQQTQAALRGGIFGLLGCL QAIGLFAVMAAVLLGGARASRLLFQRLLDVVRSPISFFERTPIGHLLNRF SKETDVTVDIPDKLRSLMYAFGLLEVSLVAVATPLATVAILPLFLLYA GFQSLYVVSSQLRRLLESASYSSVCSHMAETFQGSTVVRAFRTQAPFVAQ NNARVDESQRISFPRLVADRWLAANVELLGNLVAFAATCAVLSKAHLS AGLVGFSVSAALQVTQTLQWVVRNWTDLSENSIVSVERMQDYAWTPKEA PWRLPTCAAQPPWPQGGQIEFRDFGLRYRPELPLAVQGVSFKIHAGEKVG IVGRTGAGKSSLASGLLRLQEAEEGGIWDGVPPIAHVGLHTLRSRISIIPQD PILFPGSLRMNLDLLQEHSDIAWALETVQLKALVASLPGQLQYKCADR GEDLSVGQKQLLCLARALLRKTQILILDEATAAVDPGTELQMQAMLGSW FAQCTVLLIAHRLRSVMDCARVLVMDKGQVAESGSPAQLLAQKGLFYRL AQESGLV	ABCC6

10

20

30

40

50

409	<p>MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFP  PRVPKSFPFNTSVVYKKTFLVEFTDHLFNIAPRPPWMGLLGPTIQAEVYD  TVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEDDKVFPG  GSHTYVWQVLKENGPMASDPLCLTYSYLSHVLDLVKDLNSGLIGALLVCR  EGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWP  KMHTVNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVR  NHRQASLEISPITFLTAQTLLMDLGQFLLFCHISSHQHDGMEAYVKVDSCP  EEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDNNSPSFIQIRSVAKKHPK  TWWHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNQPQRIGRKYKKVRF  MAYTDETFKTREAIQHESGILGPLLYGEVGDITLLIIFKNQASRPYNIYPHGI  TDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTR  YYSSFVNMERDLASGLIGPLLYCYKESVDQRGNQIMSDKRNVLFSVFDEN  RSWYL TENIQRFLPNPAGVQLEDEPEFQASNIMHSINGYVFDLQSLVCLHE  VAYWYILSIGAQTDFLSVFFSGYTFKHKMYEDTLTLFPFSGETVFMSE  NPGLWILGCHNSDFRNRGMTALLKVSSCDKNTGDYEDSYEDISAYLLS  KNNAIEPSFSQNSRHPSTRQKQFNATTIPENDIEKTDPWFAHRTPMPKIQ  NVSSDILLMLLRQSPTPHGLSLSDLQEAKEYETFSDDPSPGAIDSNNLSSEM  THFRPQLHHSMDMVFTPESEGLQLRLNEKLGTTAATELKKLDFKVSSTSN  LISTIPSDNLAAGTDNTSSLGPPSMPVHYDSQLDITLFGKKSPLTESGGL  SLSEENNSKLLSEGLMNSQESSWGKNVSTESGRLFKGKRAHGPAALLTK  DNALFKVISILLKTNKTSNNSATNRKTHIDGPSLLIENSPSVWQNILESDE  FKKVTPLIHDRMLMDKNATALRLNHMSNKTTSSKNMEMVQQKKEGPIPP  DAQNPDMSEFFKMLFLPESARWIQRTHGKNSLNSGQGPSKQLVSLGPEKS  VEGQNFLSEKNKVVVGKGEFTKDVGLKEMVFPSSRNLFITNLDNLHENN  THNQEKKIQEEIEKKTLIQENVVLPQIHTVTGKTNFMKNLFLSTRQNV  GSYDGAAPVLQDFRSLNDSTNRKHTAHFSKKGEEENLEGLGNQTKQ  IVEKYACTTRISPNTSQNFVTQRSKRALKQFRLPLEETELEKRIIVDDTST  QWSKNMKHLTPSTLTQIDYNEKEKGAITQSPLSDCLTRSHSIPQANRSPPI  AKVSSFPSIRPIYLTRVLFQDNSSHLPAASYRKKDSGVQESSHFLQGAKKN  NLSLAILTLEMTGDQREVGLGTSATNSVTYKKVENTVLPKPDLPKTS GK  VELLPKVHIYQKDLFPETSTNGSPGHLDLVEGSLQGTGAIKWNEANRP  GKVPFLRVATESSAKTPSKLLDPLAWDNHYGTQIPKEEWKSQEKSPKTA  FKKQDTILSLNACESNHAIAAINEGQNKPEIEVTWAKQGRTERLCSQNPV  LKRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQK  KTRHYFIAAVERLWDYGMSSSPHVLNRNAQSGSVPQFKKVVFQFTDGS  FTQPLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEE  DQRQGAEPKRNFKVNETKTYFWKVQHMMAPTKDEFDCKAWAYFSDV  DLEKDVHSLIGPLLVCHTNTLNPAGRQVTVQEFALFFITFDETKSWYFT  ENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPLGLVMAQDQRIR  WYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMLPS  KAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDFQITASG  QYGQWAPKLARLHYSGSINAWSTKEPFSWIKVDLLAPMIHGIKTQGARQ</p>	F8
-----	--	----

10

20

30

40

50

	KFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNP PIIARYIRLHPHTHYSIRSTLRMELMGCDLNSCSMPLGMESKAISDAQITASS YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVT GVTTQGVKSLTSMYVKEFLISSQDGHQWTLFFQNGKVKVFQGNQDSF TPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY	
410	MQRVNMIMAESPGLITICLLGYLLSAECTVFLDHENANKILNRPKRYNSG KLEEFVQGNLERECMEEKCSFEEAREVFENTERTTEFWKQYVDGDQCES NPCLNGGCKDDINSYECWCPFGFEGKNCELDVTCTNIKNGRCEQFCKNSA DNKVVCSCTEGYRLAENQKSCEPAVFPFCGRVSVSQTSLKTRAETVFPDV DYVNSTEAEITLDNITQSTQSFNDFTRVVGEDAKPGQFPWQVVLNGKV DAFCGGSIVNEKWIVTAAHCVETGVKITVVAGEHNIEETEHTEQKRNVIRI I PHHNYNAAINKYNHDIALLELDEPLVLNSYVTPICIADKEYTNIFLKFGSG YVSGWGRVFHKGRSALVLQYLRVPLVDRA TCLRSTKFTIYNNMFCAGFH EGGRDSCQGDSGGPHVTEVEGTSFLTGIISWGEECAMKGKYGITYTKVSRY VNWIKEKTKLT	F9
411	MDPPRPALLALLPALLLLLLLAGARAEEMLENSLVCPKDATRKFHLR KYTYNYEAESSSGVPGTADSRSATRINCKVELEVPQLCSFILKTSQCTLKE VYGFNPEGKALLKTKNSEEFAAAMSRVELKLAIPGKQVFLYPEKDEPT YILNIKRGIIISALLVPETEEAKQVFLDFTVYGNCSHFVTKTRKGNVATEI STERDLGQCDFRFPKPIRTGISPLALIKGMTRPLSTLISSSQSCQYTLDAKRKH VAEAIKCEQHFLFLPFSYKNKYGMVAQVTQTLKLEDTPKINSRFFGEGTKK MGLAFESTKSTSPPKQAEAVLKTQLQELKKTISEQNIQRANLNFNKLVTCLR GLSDEAVTSLLPQLIEVSSPITLQALVQCGQPQCSTHILQWLKR VHANPLLI DVVTYLVALIPEPSAQLREIFNMARDQRSRATLYALSHAVNNYHKTNPT GTQELLDIANYLMEQIQDDCTGDEDYTYLILRVIGNMGQTMEQLTPELKS SILKCVQSTKPSLMIQKAAIQALRKMEPKDKDQEVLLQTFLLDDASPGDKR LAAYLMLMRSPSQAINKIVQILPWEQNEQVKNFVASHIANILNSEELDIQD LKKLVKEALKESQLPTVMDFRKFSRNYQLYKSVSLPSLDPASAKIEGNLIF DPNNYLPKESMLKTTLTAFGFASADLIEIGLEGKGFEPTEALFGKQGFPP DSVNKALYWVNGQVPDGVSKVLVDHFGYTKDDKHEQDMVNGIMLSVE KLIKDLKSKEVPEARAYLRILGEELGFASLHDLQLLGKLLLMGARTLQGIP QMIGEVIKKGSKNDFFLHYIFMENAFELPTGAGLQLQISSSGVIAPGAKAG VKLEVANMQAELVAKPSVSVEFVTNMGHIIPDFARSGVQMNTNFFHESGL EAHVALKAGKLFIIIPSPKRPVKLLSGGNTLHLVSTTKTEVIPPLIENRQSW SVCKQVFPGLNYCTSGAYSNASSTDSASYYPLTGDTRLELELRPTGEIEQY SVSATYELQREDRALVDTLKFVTQAEGAKQTEATMTFKYNRQSM TLSSE VQIPDFDVLGTLRVNDESTEGKTSYRLTLDIQNKKITEVALMGHLSCDT KEERKIKGVISIPRLQAEARSEILAHWSPAKLLLQMDSSATAYGSTVSKRV AWHYDEEKIEFEWNTGTNVDTKKMTSNFPVDLSYDYPKSLHMYANRLLD HRVPQTDMTFRHVGSKLIVAMSSWLQKASGSLPYTQTLQDHLNSLKEFN LQNMGLPDFHIPENLFLKSDGRVKYTLNKNLSLKIEIPLPFGGKSSRDLM	ApoB

10

20

30

40

50

ETVRTPALHFKSVGFHLP SREFQVPTFTIPKLYQLQVPLLGVLDLSTNVYS  
 NLYNWSASYSGGNTSTDHFSLRARYHMKADSVVDLLSYNVQSGGETTY  
 DHKNTFTLSYDGSLRHKFLDSNIKFSHVEKLGNNPVSKGLLIFDASSSWG  
 QMSASVHLDSKKKQHLFVKEVKIDGQFRVSSFYAKGTYGLSCQRDPNTG  
 RLVGESNLRFNSSYLQGTNQTGRYEDGTLSTSTSDLQSGIHKNTASLKYE  
 NYELTLKSDTNGKYKNFATS NKMDMTFSKQNALRSEYQADYESLRFFS  
 LLSGSLNSHGLELNADILGTDKINSGAHKATLRIGQDGISTSATTNLKCSLL  
 VLENELNAELGLSGASMKLTTNGRFREHNAKFSLDGKAALTELSLGSAY  
 QAMILGVDSKNIFNFKVSQEGLKLSNDMMGSYAEMKFDHTNSLNIAGLS  
 LDFSSKLDNIYSSDKFYKQTVNLQLQPYSLVTTLNSDLKYNALDLTNNGK  
 LRLEPLKLVAGNLKGAYQNNEIKHIYAIS SAALSASYKADTVAKVQGV  
 FSHRLNTDIAGLASAIDMSTNYNSDSLHFSNVFRSVMAPFTMTIDAHTNG  
 NGKLALWGEHTGQLYSKFLKAEPLAFTFSDYKGGSTSHHLVSRKSISAA  
 LEHKVSALLTPAEQTGTWKLKTQFNNEYSQDLDAYNTKDKIGVELTGR  
 TLADLTLLDSPIKVPLLLSEPINIIDALEMRDAVEKPQEFTIVAFVKYDKNQ  
 DVHSINLPPFETLQEYFERNRQTIIIVLENVQRNLKHINIDQFVRKYRAAL  
 GKLPQQANDYLN SFNWERQVSHAKEKLTALTKKYRITENDIQIALDDAKI  
 NFNEKLSQLQTYMIQFDQYIKDSYDLHDLKIAIANIIDEIIEKLSLDEHYHI  
 RVNLVKTIHDLHLFIENIDFNKSGSSTASWIQNVDTKYQIRIQIQEKLQQLK  
 RHIQNIDIQHLAGKLLKQHIEAIDVRVLLDQLGTTISFERINDILEHVKHFVIN  
 LIGDFEVAEKINAFRAKVHELIERYEVDQIQVLMDKLVELAHQYKVKET  
 IQKLSNVLQQVKIKDYFEKLVGFIDDAVKKLNELSFKTFIEDVNFKFLDMLI  
 KKLKSFQDYHQFVDETNDKIREVTQRLNGEIQAELPQKAEALKLFLEETK  
 ATVAVYLESLQDTKITLIINWLQEALSSASLAHMKAKFRETLEDTRDRMY  
 QMDIQQELQRYLSLVGVVYSTLVTYISDWWTLAAKNLTDFAEQYSIQDW  
 AKRMKALVEQGFTVPEIKTILGTMPAFEVSLQALQKATFQTPDFIVPLTDL  
 RIPSQVQINFKDLKNKIPSRFSTPEFTILNTFHIPSFTIDFVEMKVKIIRTIDQM  
 LNSELQWPVPDIYLRDLKVEDIPLARITLPDFRLPEIAIPEFIIPTLNLNDFQV  
 PDLHIPEFQLPHISHTIEVPTFGKLYSILKIQSPLFTLDANADIGNGTTSANE  
 AGIAASITAKGESKLEVLNDFQANAQLSNPKINPLALKESVKFSSKYLRT  
 EHGSEMLFFGNAIEGKSNTVASLHTEKNTLELSNGVIVKINNQLTLDSENTK  
 YFHKLNIPKLD FSSQADLRNEIKTLLKAGHIAWTSSGKGSWKWACPRFSD  
 EGTHESQISFTIEGPLTSFGLSNKINSKHLRVNQNLVYESGSLNFSKLEIQSQ  
 VDSQHVGHSVLTAKGMALFGEGKAEFTGRHDAHLNGKVIGTLKNSLFFS  
 AQPFEITASTNNEG NLKVRFPRLRTGKIDFLN NYALFLSPSAQQASWQVSA  
 RFNQYKYNQNF SAGNNENIMEAHVGINGEANLDFLNIPLTIPEMRLPYTII  
 TTPPLKDFSLWEKTGLKEFLKTTKQSFDSL SVKAQYKKNKHRHSITNPLAV  
 LCEFISQSIKSFDRHFEKRN RNALDFVTKSYNETKIKFDKYKAEKSHDEL  
 PRTFQIPGYTVPVVNVEVSPFTIEMSAFGYVFPKAVSMPSFSILGSDVRVPSY  
 TLILPSLELPVLHVPRNLKLSLPDFKELCTISHIFIPAMGNITYDFSFKSSVIT  
 LNTNAELFNQSDIVAHLLSSSSVIDALQYKLEGTTRLTRKRGLKLATALS  
 LSNKFVEGSHNSTVSLTTKNMEVSVATTTKAQIPILRMNFKQELNGNTKS

10

20

30

40

50



	<p>KPTVSSSMFEFKYDFNSSMLYSTAKGAVDHLKLSLESLSYFYSIESSTKGDVK  GSVLSREYSGTIASEANTYLNSKSTRSSVKLQGTSKIDDIWNLEVKENFAG  EATLQRIYSLWEHSTKNHLQLEGLFFTNGEHTSKATLELSPWQMSALVQV  HASQPSSFHDFPDLGQEVALNANTKNQKIRWKNEVRIHSGSFQSQVELSN  DQEKAHLDIAGSLEGHLRFLKNIILPVYDKSLWDFLKLDDVTTSIGRRQHLLR  VSTAFVYTKNPNNGYSFSPVKVLADKFIIPGLKLNLDLNSVLVMPFTHVPFT  DLQVPSCKLDREIYQYKLRITSSFALNPLTLPEVKFPEVDVLTKYSPEDS  LIPFFEITVPESQLTVSQFTLPSKSVSDGIAALDLNAVANKIADFELPTIIVPEQ  TIEIPSIKFSVPAGIVIPSFQALTARFEVDSVPVYNATWSASLKNKADYVETV  LDSTCSSTVQFLEYELNVLGTHKIEDGTLASKTKGTFahrDFSAEYEEEDG  KYEGLQEWEGKAHLNIKSPAFTDLHLRYQKDKKGISTSAA SPAVGTVGM  DMDEDDDFSKWNFYSPQSSPKKLTIFKTEL RVRESDEETQIKVNWEEE  AASGLTSLKDNVPAKATGVLYDYVNKYHWEHTGLTLREVSSKLRRNLQ  NNAEWVYQGAIRQIDDIVRFQKAASGTTGTQYQEWKDKAQNLQYELLT  QEGQASQGLKDNVFDGLVRVTQEFHMKVKHLIDSLIDFLNFRPFQFPGK  PGIYTREELCTMFIREVGTVLSQVYSKVHNGSEILFSYFQDLVITLPELKR  HKLIDVISMYRELLKDLKAEQEVFKAIQSLKTTEVLRNLQDLLQFIFQLIE  DNIKQLKEMKFTYLINYIQDEINTIFSDYIPYVFKLLKENLCLNLHKFNEFI  QNELQEASQELQIQHYIMALREEYFDPSIVGWTVKYYELEEKIVSLIKNL  LVALKDFHSEYIVSASNFTSQLSSQVEQFLHRNIQEYLSILTDPDGKGKEKI  AELSATAQEIIKSQAIAATKKIISDYHQFRYKLQDFSDQLSDYYEKFIAESK  RLIDLSIQNYHTFLIYITELLKKLQSTTVMNPNYMKLAPGELTIIL</p>	
412	<p>MGTVSSRRSwwPLPLLLLLLLLLLPAGARAQEDDGDYEELVLA LRSEE  DGLAEAPEHGTTATFHRCADPWRLPGTYVVVLKEETHLSQSERTARRL  QAQAARRGYLTKILHVFHGLLPGLVKMSGDLLELALKLPHVDYIEEDSS  VFAQSPWNLERITPPRYRADEYQPPDGGSLVEVYLLDTSIQSDHREIEGR  VMVTFDENVPEEDGTRFHRQASKCDSHGTHLAGVVSGRDAGVAKGASM  RSLRVLNCQGGKTVSGTLIGLEFIRKSQLVQVGPVLVLLPLAGGYSRVLN  AACQRLARAGVVLVTAAGNFRDDACLSPASAPEVITVGATNAQDQPVT  LGTGTFNGRCDLDFAPGEDIIGASSDCSTCFVSQSGTSQAAAHVAGIAA  MMLSAEPELTLAELRQLIHFSKDVINEAWFPEDQRVLT PNLVAALPPST  HGAGWQLFCRTVWSAHSGPTRMATAVARCAPDEELLSCSSFSRSGKRRG  ERMEAQGGKLVCRAHNAFGGEGVYAIARCCLLPQANCSVHTAPPAEAS  MGTRVHCHQQGHVLTGCSHWEDLGTHKPPVLRPRGQPNQCVGHRE  ASIHASCCHAPGLECKVKEHGIPAPQEQT VACEEGWTLTGCSALPGTSH  VLGAYAVDNTCVVRSRDVSTTGSTSEGAVTAVAICCRSRHLAQASQELQ</p>	PCSK9
413	<p>MDALKSAGRALIRSPSLAKQSWG GGRHRKLPENWTDRETLLLEGMLFS  LKYLGMTLVEQPKGEELSAAAIKRIVATAKASGKKLQKVT LKVS PRGIILT  DNL TNQLIENVSIYRISYCTADKMHDKVFAYIAQSQHNSLECHAFLCTK  RKMAQAVTLTVAQAFKVAFEFWQVSKEEKEKRDKASQEGGDVLGARQ  DCTPSLKSLVATGNLLDLEETAKAPLSTVSANTTNMDEVPRPQALSGSSV  VWELDDGLDEAFSRLAQSR TNPQVLD TGLTAQDMHYAQCLSPVDWDPK  DSSGTEQDDLFSF</p>	LDLRAP1

10

20

30

40

50

414	<p>MGDLSSLTPGGSMGLQVNRGSQSSLEGAPATAPEPHSLGILHASYSVSHR  VRPWWDITSCRQQWTRQILKDVSLYVESGQIMCILGSSGSGKTLLDAMS  GRLGRAGTFLGEVYVNGRALRREQFQDCFSYVLQSDTLLSSLTVRETLHY  TALLAIRRGNPGSFQKKVEAVMAELSLSHVADRLIGNYSLGGISTGERRR  VSIAAQLLQDPKVMLFDEPTTGLDCMTANQIVVLLVELARRNRIVVLTIH  QPRSELFQLFDKIAILSFGELIFCGTPAEMLDFFNDCGYPCPEHSNPFDFYM  DLTSVDTQSKEREIETSKRVQMIESAYKKS AICHKTLKNIERMKHLKTLPM  VPFKTKDSPGVFSKLGVLRRVTRNLVRNKLA VITRLLQNLIMGLFLLFFV  LRVRSNVLKGAIQDRVGLLYQFVGATPYTGMLNAVNLFPVLRVAVSDQES  QDGLYQKWQMLLAYALHVL PFSVVATMIFSSVCYWTGLHPEVARFGY  FSAALLAPHLIGEFLTLVLLGIVQNPNI VNSV VALLSIAGVLV GSGFLRNIQ  EMPIPKIISYFTFQKYCSEILVVNEFYGLNFTCGSSNVSVTTNPMCAFTQG  IQFIEKTCPGATS RFTMNFLLYSFIPALVILGIVVFKIRDHLISR</p>	ABCG5
415	<p>MAGKAAEERGLPKGATPQDTSGLQDR LFSSES DNSLYFTYSGQPNTLEVR  DLNYQVDLASQVPWFQLAQFKMPWTSPSCQNSCELGIQNLSFKVRSGQ  MLAIIGSSGCGRASLLDVITGRGHGGKIKSGQI WINGQPSSPQLVRKCVAH  VRQHNQLLPNLTVRETLAFIAQMRLPRTFSQAQRDKRVEDVIAELRLRQC  ADTRVGNMYVRGLSGGERRRVSIGVQLLWNP GILILDEPTSGLDSFTAHN  LVKTL SRLAKGNRLVLISLHQPRSDIFRLFDL VLLMTSGTPIYLGA AQHMV  QYFTAIGYPCPRYSNPADFYVDLTSIDRRSREQELATREKAQSLAALFLEK  VRDLDDFLWKAETKDLDEDTCVESSVTP LDTNCLPSPTKMPGAVQQFTT  LIRRQISNDFRDLPTLLIHGAEA CLMSMTIGFLYFGHGSIQLSFMDTAALLF  MIGALIPFNVILDVISKCYSERAMLYYELEDGLYTTG PYFFAKILGELPEHC  AYIIIYGMPTYWLANLRPGLQPFL LHFLLVWL VVFCRIMALAAAALLPT  FHMASFFSNALYNSFYLAGGFMINLSSL WTVPAWISKVSFLRWCFEGLMK  IQFSRRTYKMPLGNLTI AVSGDKILSVMELDSYPLYAIYLIVIGLSGGFMVL  YYVSLRFIKQKPSQDW</p>	ABCG8
416	<p>MGPPGSPWQWV TLLLGLLLPPAAPFWLL NVLFPHTTPKAELSNHTRPVI  LVPGLGNQLEAKLDKPDVVNWMCYRKTE DFFTIWLDLNMFLPLGVDC  WIDNTRVVYNRSSGLVSNAPGVQIRVPGFGK TY SVEYLDSSKLAGYLHTL  VQNLVNNGYVRDET VRAAPYDWRLEPGQ QEEYRKL AGLVEEMHAAY  GKPVFLIGHSLGCLHLLYFLLRQPQAWKDR FIDGFISLGAPWGGSIKPMLV  LASGDNQGIPI MSSIKLKEEQRITTTSPWMFPSRMAWPEDHVFISTPSFN Y T  GRDFQRFFADLHFEEGWYMWLQSRDLLAGLPAPGVEVYCLYGVGLPTPR  TYIYDHGFYTD P VGVLYEDGDDTVATRSTELCGLWQGRQPQPVHLLPL  HGIQHLMVFSNLTLEHINA ILLGAYRQGPPASPTASPEPPPE</p>	LCAT

10

20

30

40

50

417	<p>MKIATVSVLLPLALCLIQDAASKNEDQEMCHEFQAFMKNGLFCPQDKK  FFQSLDGIMFINKCATCKMILEKEAKSQKRARHLARAPKATAPTELNCDD  FKKGERDGFICPDYIEAVCGTDGKTYDNRCALCAENAKTGSQIGVKSE  GECKSSNPEQDVCSAFRPFVVDGRLGCTRENDPVLGPDGKTHGNKCAMC  AELFLKEAENAKREGETRIRRNAEKDFCKEYKQVRNGRLFCTRESDPVR  GPDGRMHGNKCALCAEIFKQRFSEENSKTDQNLGKAEKTKVKREIVKL  CSQYQNQAKNGILFCTRENDPIRGPDKMHGNLCSMCQAYFQAENEKK  KAEARARNKRESGKATSYAELCSEYRKLVRNGKLACTRENDPIQGPDK  VHGNTCSMCEVFFQAEKKKKKEGKSRNKRQSKSTASFEELCSEYRCSR  KNGRLFCTRENDPIQGPDKMHGNTCSMCEAFFQEEERARAKAKREAAK  EICSEFRDQVRNGTLICTREHNPVRGPDGKMHGNKCAMCASVFKLEEEK  KNDKEEKVKVEAEKVKREAVQELCSEYRHYVRNGRLPCTRENDPIEGLD  GKIHGNTCSMCEAFFQQAEEKERAEPRAKVKREAEKETCDEFRRLLQN  GKLFCTRENDPVRGPDGKTHGNKCAMCKAVFQKENEERKRKEEEDQRN  AAGHGSSGGGGNTQDECAEYREQMKNGLSCTRESDPVRDADGKSYN  NQCTMCKAKLEREAERKNEYSRSRNGTGESGKDTCDEFRSQMKNGL  ICTRESDPVRGPDGKTHGNKCTMCKEKLERAEEKKKKEDEDRSNTGER  SNTGERSNDKEDLCREFRSMQRNGKLICTRENNPVRGPYGMHINKCAM  CQSIFDREANERKKKDEEKSSSKPSNNAKDECSEFRNYIRNNELICPREND  PVHGADGKFYTNKCYMCRAVFLTEALERAKLQEKPSHVRASQEEDSPDS  FSSLDSEMCKDYRVLPRIGYLCPKDLKPVCGDDGQTYNNPCMLCHENLIR  QTNTHIRSTGKCEESSTPGTTAASMPPSDE</p>	SPINK5
418	<p>MEKNGNNRKLRCVATCNRADYSKLAPIMFGIKTEPEFFELDVVVLGSHL  IDDYGNTYRMIEQDDFDINTRLHTIVRGEDEAAMVESVGLALVKLPDVLN  RLKPDIMIVHGDRFDALALATSAALMNIRILHIEGGEVSGTIDDSIRHAITK  LAHYHVCCTRSAEQHLISMCEDHDRILLAGCPSYDKLLSAKNKDYMHIR  MWLGDDVSKDYIVALQHPVTTDIKHSIKMFELTDALISFNKRTLVLFP  NIDAGSKEMVRVMRKKGIEHHPNFRAVKHVPFDQFIQLVAHAGCMIGNS  SCGVREVGAFGTPVINLGTRQIGRETGENVLHVRDADTQDKILQALHLQF  GKQYPCSKIYGDGNAVPRILKFLKSIDLQEPLQKKFCFPPVKENISQDIDHI  LETLSALAVDLGGTNLRVAIVSMKGEIVKKYTQFNPKTYEERINLILQMC  VEAAAEAVKLNCRILGVGISTGGRVNPREGIVLHSTKLIQEWNSVDLRTPL  SDTLHLPVWVDNDGNCAALAERKFGQGKLENFVTLITGTGIGGGIIHQH  ELIHSSFCAAELGHLVVS LDGPDSCGSHGCIEAYASGMALQREAKKLH  DEDLLLVEGMSVPKDEAVGALHLIQA AKLGNAKAQSILRTAGTALGLGV  VNILHTMNPSLVILSGVLASHYIHIVKDVIRQQALSSVQDQDQVVDVSDLVDP  ALLGAASMVLDTTRRIY</p>	GNE
419	<p>DIQMTQTTSSLSASLGDRVTISCRASQDISKYLWYQQKPDGTVKLLIYHT  SRLHSGVPSRFSGSGSGTDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGT  KLEITGSTSGSGKPGSGEGSTKGEVKLQESGPLVAPSQSLSVTCTVSGVS  LPDYGVSWIRQPPRKGLEWLGVWGSETTYNSALKSRLTIKDNSKSQVF  LKMNSLQTD DTAIYYCAKHYYYGGSYAMDYWGQTSVTVSS</p>	抗 CD19 scFv (FMC63)

10

20

30

40

50

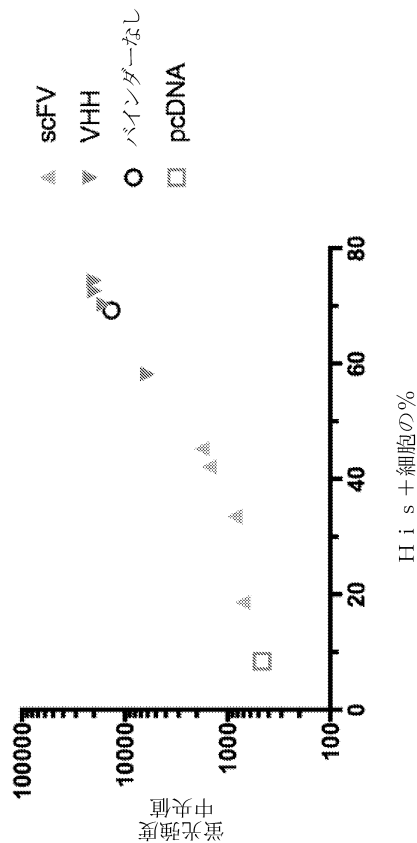
420	DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPDGTVKLLIYHT SRLHSGVPSRFSGSGGTDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGT KLEITGGGGSGGGGSGGGGSEVKLQESGPELVAPSQSLSVTCTVSGVSLP DYGVSWIRQPPRKGLEWLGVWGETTYNSALKSRLTIHKDNSKSQVFL KMNSLQTDDETAIYYCAKHYYYGGSYAMDYWGQGTSTVTVSS	抗 CD19 scFv (FMC63)
421	ESKYGPCPCP	IgG4 ヒンジ
422	TTTPAPRPPTPAPTIASQPLSLRPE	CD8 ヒンジ
423	IEVMYPPPYLDNEKSNGTIIHVKGKHLCPSPFPGPSKP	CD28
424	ACRPAAGGAVHTRGLDFACDIYWAPLAGTCGVLLLSLVITLYC	CD8
425	FWVLVVVGGVLACYSLLVTVAFIIFWV	CD28
426	FWVLVVVGGVLACYSLLVTVAFIIFWV	CD28
427	RSKRSRLHSDYMNMTPRRPGPTRKHYPYAPPRDFAAYRS	CD28
428	KRGRKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEGGCEL	4-1BB
429	RVKFSRSADAPAYQQQNQLYNELNLGRREEYDVLDKRRGRDPEMGGK PRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGKGGHDGLYQGLSTA TKDITYDALHMQUALPPR	CD3 ゼー タ
430	RVKFSRSADAPAYKQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGK PRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGKGGHDGLYQGLSTA TKDITYDALHMQUALPPR	CD3 ゼー タ

10

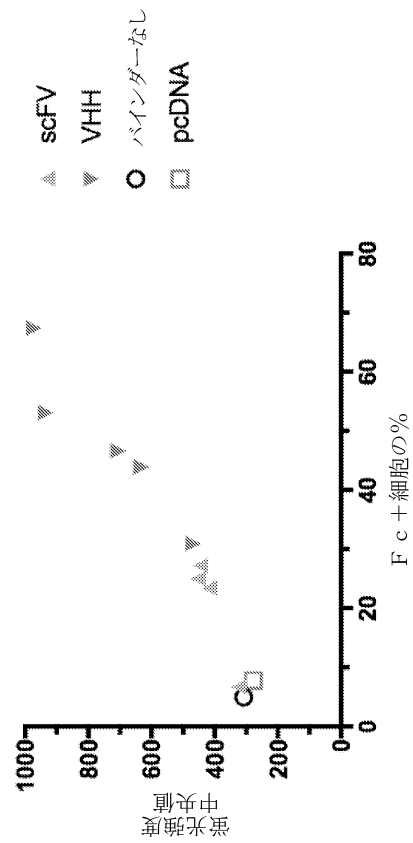
20

【 図面 】

【 図 1 A 】



【 図 1 B 】

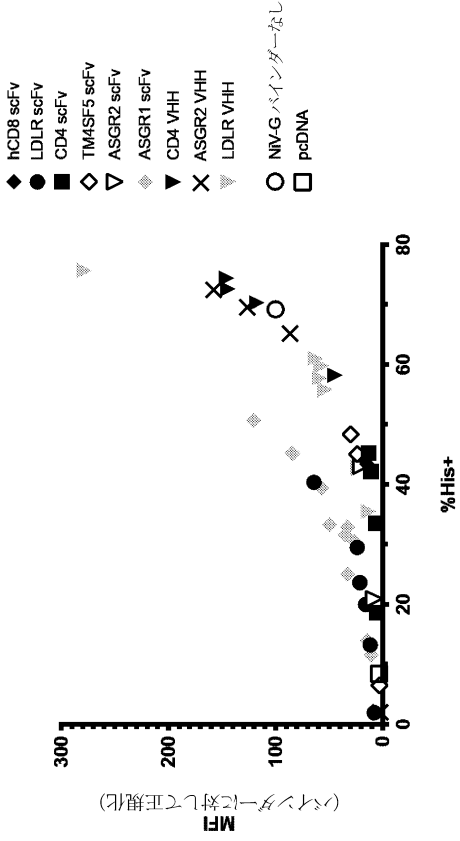


30

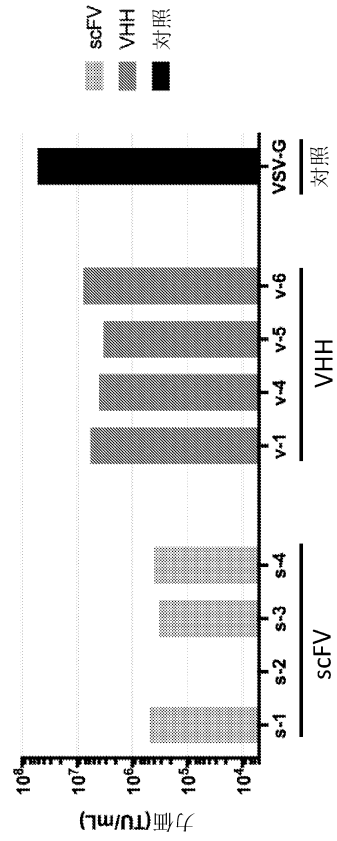
40

50

【 図 1 C 】



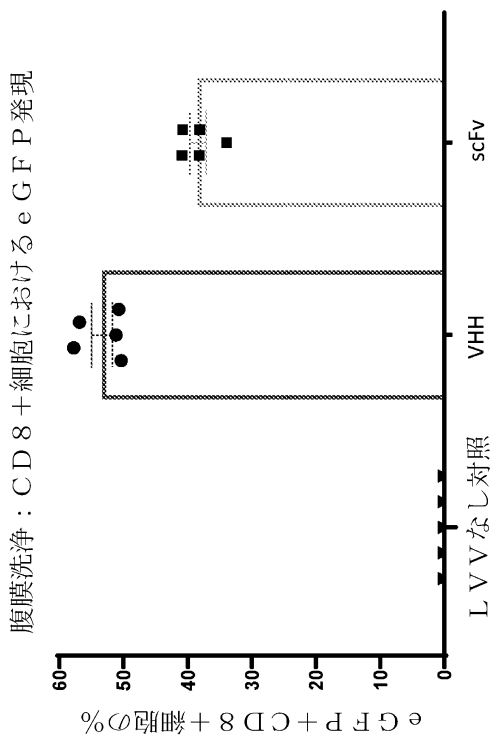
【 図 2 】



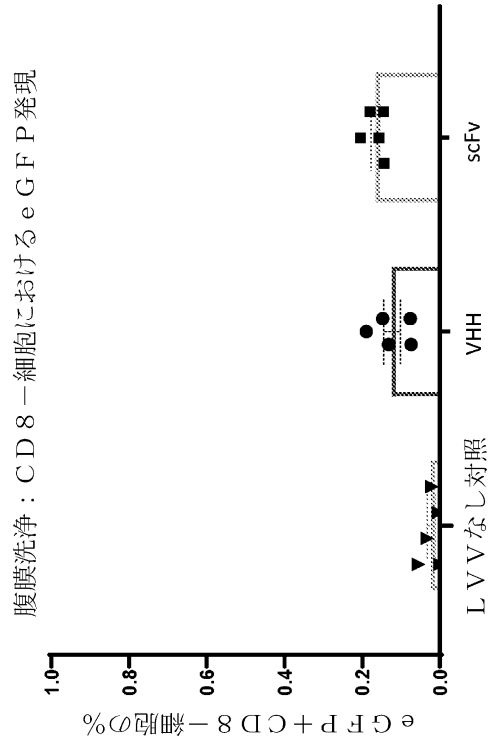
10

20

【 図 3 A 】



【 図 3 B 】

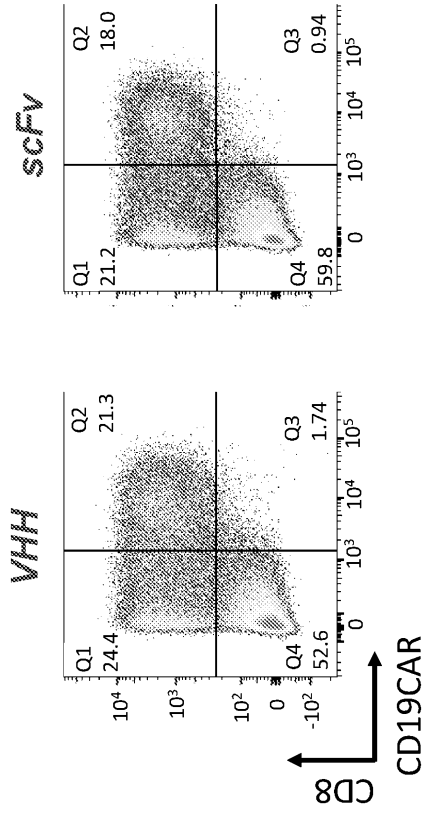


30

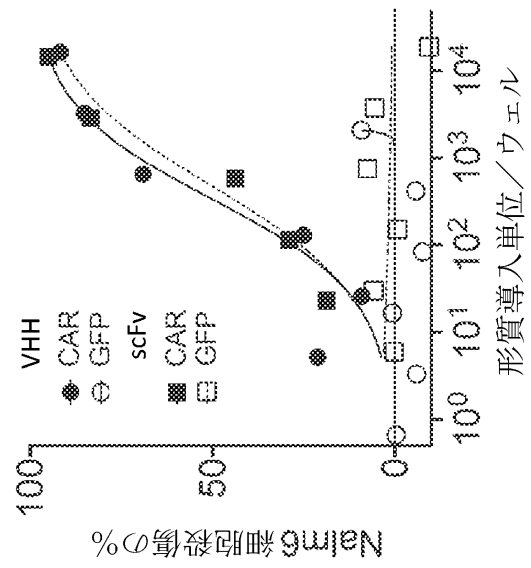
40

50

【 図 4 A 】



【 図 4 B 】



10

20

30

40

50

【 配列表 】

2023521663000001.app

## 【 国際調査報告 】

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2021/024993

A. CLASSIFICATION OF SUBJECT MATTER INV. C12N7/00 C12N15/86 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) C12N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, Sequence Search, EMBASE, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2019/086351 A1 (MILTENYI BIOTEC GMBH [DE]) 9 May 2019 (2019-05-09) page 12, line 16 - page 13, line 9; claim 1 page 16, lines 1-4,8-10,24-26 page 19, line 27 - page 20, line 4 page 37, line 22	1-141
A	----- WO 2020/014209 A1 (FLAGSHIP PIONEERING INNOVATIONS V INC [US]) 16 January 2020 (2020-01-16) claim 24 ----- -/--	1-141
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.		<input checked="" type="checkbox"/> See patent family annex.
* Special categories of cited documents :		
*A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier application or patent but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed		*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art *&* document member of the same patent family
Date of the actual completion of the international search	Date of mailing of the international search report	
11 August 2021	06/09/2021	
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer  Sprinks, Matthew	

Form PCT/ISA/210 (second sheet) (April 2005)

10

20

30

40

50

INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2021/024993

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>CHEN CHENG ET AL: "Development of T cells carrying two complementary chimeric antigen receptors against glypican-3 and asialoglycoprotein receptor 1 for the treatment of hepatocellular carcinoma", CANCER IMMUNOLOGY, IMMUNOTHERAPY, SPRINGER, BERLIN/HEIDELBERG, vol. 66, no. 4, 29 December 2016 (2016-12-29), pages 475-489, XPO36192310, ISSN: 0340-7004, DOI: 10.1007/S00262-016-1949-8 [retrieved on 2016-12-29] the whole document</p> <p style="text-align: center;">-----</p>	1-141

10

20

30

40

1

50



**INTERNATIONAL SEARCH REPORT**

International application No.  
PCT/US2021/024993

**Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

10

a.  forming part of the international application as filed:

in the form of an Annex C/ST.25 text file.  
 on paper or in the form of an image file.

b.  furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

c.  furnished subsequent to the international filing date for the purposes of international search only:

in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).  
 on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).

2.  In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

20

3. Additional comments:

30

40

50

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No  
PCT/US2021/024993

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2019086351 A1	09-05-2019	CA 3080424 A1	09-05-2019
		CN 111315761 A	19-06-2020
		EP 3704136 A1	09-09-2020
		JP 2021500909 A	14-01-2021
		KR 20200074204 A	24-06-2020
		US 2021180083 A1	17-06-2021
		WO 2019086351 A1	09-05-2019
-----			
WO 2020014209 A1	16-01-2020	AU 2019301053 A1	28-01-2021
		BR 112021000145 A2	06-04-2021
		CA 3105953 A1	16-01-2020
		CN 112955174 A	11-06-2021
		EP 3820509 A1	19-05-2021
		KR 20210043574 A	21-04-2021
		SG 11202100071Y A	25-02-2021
		WO 2020014209 A1	16-01-2020
		-----	

10

20

30

40

50

## フロントページの続き

## (51)国際特許分類

C 1 2 N 15/62 (2006.01)  
 C 1 2 N 15/85 (2006.01)  
 C 1 2 N 15/86 (2006.01)  
 C 1 2 N 5/10 (2006.01)  
 C 1 2 N 5/0783(2010.01)  
 C 1 2 N 5/0793(2010.01)  
 C 1 2 N 5/09 (2010.01)  
 A 6 1 K 9/14 (2006.01)  
 A 6 1 K 9/127(2006.01)  
 A 6 1 K 35/76 (2015.01)  
 A 6 1 K 48/00 (2006.01)  
 A 6 1 K 47/69 (2017.01)  
 A 6 1 P 43/00 (2006.01)

## F I

C 1 2 N 15/62 Z  
 C 1 2 N 15/85 Z  
 C 1 2 N 15/86 Z  
 C 1 2 N 5/10  
 C 1 2 N 5/0783  
 C 1 2 N 5/0793  
 C 1 2 N 5/09  
 A 6 1 K 9/14  
 A 6 1 K 9/127  
 A 6 1 K 35/76  
 A 6 1 K 48/00  
 A 6 1 K 47/69  
 A 6 1 P 43/00 1 1 1

## テーマコード (参考)

,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,D  
 K,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),O  
 A(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,B  
 B,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD  
 ,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,IT,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,  
 LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,  
 RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,Z  
 W

弁理士 清水 初志

(74)代理人 100102118

弁理士 春名 雅夫

(74)代理人 100160923

弁理士 山口 裕孝

(74)代理人 100119507

弁理士 刑部 俊

(74)代理人 100142929

弁理士 井上 隆一

(74)代理人 100148699

弁理士 佐藤 利光

(74)代理人 100188433

弁理士 梅村 幸輔

(74)代理人 100128048

弁理士 新見 浩一

(74)代理人 100129506

弁理士 小林 智彦

(74)代理人 100205707

弁理士 小寺 秀紀

(74)代理人 100114340

弁理士 大関 雅人

(74)代理人 100214396

弁理士 塩田 真紀

(74)代理人 100121072

弁理士 川本 和弥

(72)発明者 トルドー カイル マービン

アメリカ合衆国 9 8 1 0 2 ワシントン州 シアトル イースト ブレイン ストリート 1 8 8 ス  
 イート 4 0 0

(72)発明者 バンドロ クリストファー

- 
- アメリカ合衆国 9 8 1 0 2 ワシントン州 シアトル イースト ブレイン ストリート 1 8 8 ス  
イト 4 0 0
- (72)発明者 マッケンジー ローレン ペッパー  
アメリカ合衆国 9 8 1 0 2 ワシントン州 シアトル イースト ブレイン ストリート 1 8 8 ス  
イト 4 0 0
- (72)発明者 シャー ジャゲシュ ヴィジャヤクマール  
アメリカ合衆国 9 8 1 0 2 ワシントン州 シアトル イースト ブレイン ストリート 1 8 8 ス  
イト 4 0 0
- (72)発明者 フォン マルツァーン ジェフリー エー .  
アメリカ合衆国 0 2 1 4 2 マサチューセッツ州 ケンブリッジ ケンブリッジ パークウェイ 5 5  
エイス フロア
- (72)発明者 ルーベンス ジェイコブ ローゼンブラム  
アメリカ合衆国 0 2 1 4 2 マサチューセッツ州 ケンブリッジ ケンブリッジ パークウェイ 5 5  
エイス フロア
- (72)発明者 ミー マイケル トラヴィス  
アメリカ合衆国 0 2 1 4 2 マサチューセッツ州 ケンブリッジ ケンブリッジ パークウェイ 5 5  
エイス フロア
- F ターム ( 参考 ) 4B065 AA90X AA95X  
4C076 AA19 AA31 AA95 CC41 EE59  
4C084 AA13 MA52 MA55 NA13 ZC021 ZC022  
4C087 AA01 AA02 CA12 MA52 MA55 NA13 ZC02