

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-502445

(P2004-502445A)

(43) 公表日 平成16年1月29日(2004.1.29)

(51) Int.Cl. ⁷		F I		テーマコード (参考)	
C 1 2 N 15/09		C 1 2 N 15/00		Z N A A 4 B O 2 4	
A 6 1 K 35/76		A 6 1 K 35/76		4 B O 6 5	
A 6 1 K 38/00		A 6 1 K 39/00		H 4 C O 8 4	
A 6 1 K 39/00		A 6 1 K 48/00		4 C O 8 5	
A 6 1 K 48/00		A 6 1 P 31/12		4 C O 8 7	
		審査請求 未請求 予備審査請求 有		(全 336 頁) 最終頁に続く	
(21) 出願番号	特願2002-509356 (P2002-509356)	(71) 出願人	591076811	最終頁に続く	
(86) (22) 出願日	平成13年7月5日 (2001.7.5)		カイロン コーポレーション		
(85) 翻訳文提出日	平成14年12月26日 (2002.12.26)		アメリカ合衆国, カリフォルニア 9 4 6		
(86) 国際出願番号	PCT/US2001/021241		0 8, エミリービル, ホートン ストリー		
(87) 国際公開番号	W02002/004493		ト 4 5 6 0		
(87) 国際公開日	平成14年1月17日 (2002.1.17)	(71) 出願人	503009225	最終頁に続く	
(31) 優先権主張番号	09/610, 313		ユニバーシティ オブ ステレンボスク		
(32) 優先日	平成12年7月5日 (2000.7.5)		南アフリカ 7 5 0 5 タイガーバーグ,		
(33) 優先権主張国	米国 (US)		ビー. オー. ボックス 1 9 0 6 3		
		(74) 代理人	100078282		
			弁理士 山本 秀策	最終頁に続く	
		(74) 代理人	100062409		
			弁理士 安村 高明		
		(74) 代理人	100113413		
			弁理士 森下 夏樹		

(54) 【発明の名称】 抗原性H I V C型ポリペプチドをコードするポリヌクレオチド、ポリペプチドおよびその使用

(57) 【要約】

本発明は、免疫原性H I V C型ポリペプチドをコードするポリヌクレオチドに関する。DNA免疫化、パッケージング細胞株の作成、およびH I V C型タンパク質の産生を含む用途におけるこのポリヌクレオチドの使用もまた、記載される。また、本明細書中には、新規のC型H I V配列（例えば、8 _ 5 _ T V 1 _ C . Z A、8 _ 2 _ T V 1 _ C . Z Aおよび1 2 _ 5 _ 1 _ T V 2 _ C . Z A）、これらの新規配列によってコードされるポリペプチド、およびこれらまたは他のC型H I V配列から生成される合成発現カセットが記載される。

【特許請求の範囲】

【請求項 1】

H I V P o l ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含む発現カセットであって、ここで、該 P o l ポリペプチドをコードする該ポリヌクレオチド配列が、図 8（配列番号 30）；図 9（配列番号 31）または図 10（配列番号 32）に示される配列に対して少なくとも 90% の配列同一性を有する配列を含む、発現カセット。

【請求項 2】

X 個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、（i）該 X 個連続したヌクレオチドが、配列番号 46 の Y 個連続したヌクレオチドに対して少なくとも 90% のパーセント同一性を有し、（i i）X が、Y と等しく、そして（i i i）Y が、少なくとも 97 である、発現カセット。 10

【請求項 3】

X 個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項 2 に記載の発現カセットであって、ここで、（i）該 X 個連続したヌクレオチドが、配列番号 47 の Y 個連続したヌクレオチド配列に対して、少なくとも 90% のパーセント同一性を有し、（i i）X が、Y と等しく、そして（i i i）Y が、少なくとも 144 である、発現カセット。

【請求項 4】

X 個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項 3 に記載の発現カセットであって、ここで、（i）該 X 個連続したヌクレオチドが、配列番号 49 または配列番号 97 の Y 個連続したヌクレオチドに対して少なくとも 90% のパーセント同一性を有し、（i i）X が、Y と等しく、そして（i i i）Y が、少なくとも 300 である、発現カセット。 20

【請求項 5】

X 個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項 4 に記載の発現カセットであって、ここで、（i）該 X 個連続したヌクレオチドが、配列番号 49 の Y 個連続したヌクレオチドに対して少なくとも 90% のパーセント同一性を有し、（i i）X が、Y と等しく、そして（i i i）Y が、2610 である、発現カセット。

【請求項 6】

X 個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項 4 に記載の発現カセットであって、ここで、（i）該 X 個連続したヌクレオチドが、配列番号 97 の Y 個連続したヌクレオチドに対して少なくとも 90% のパーセント同一性を有し、（i i）X が、Y と等しく、そして（i i i）Y が、2565 である、発現カセット。 30

【請求項 7】

X 個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、（i）該 X 個連続したヌクレオチドが、配列番号 51 の Y 個連続したヌクレオチドに対して少なくとも 90% のパーセント同一性を有し、（i i）X が、Y と等しく、そして（i i i）Y が、1494 である、発現カセット。

【請求項 8】

X 個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、（i）該 X 個連続したヌクレオチドが、配列番号 99 の Y 個連続したヌクレオチドに対して少なくとも 90% のパーセント同一性を有し、（i i）X が、Y と等しく、そして（i i i）Y が、1491 である、発現カセット。 40

【請求項 9】

X 個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、（i）該 X 個連続したヌクレオチドが、配列番号 55；配列番号 57；配列番号 101；配列番号 96；配列番号 134 または配列番号 135 の Y 個連続したヌクレオチドに対して少なくとも 90% のパーセント同一性を有し、（i i）X が、Y と等しく、そして（i i i）Y が、少なくとも 60 である、発現カセット。

【請求項 10】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項9に記載の発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号55；配列番号57；配列番号101；配列番号96；配列番号134または配列番号135のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、624である、発現カセット。

【請求項11】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号58のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、354である、発現カセット。

10

【請求項12】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号60のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、876である、発現カセット。

【請求項13】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号62のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、3015である、発現カセット。

20

【請求項14】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号103のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、3009である、発現カセット。

【請求項15】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号64または配列番号66のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、297である、発現カセット。

30

【請求項16】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号68のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、1965である、発現カセット。

【請求項17】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号70のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、1977である、発現カセット。

40

【請求項18】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号72または配列番号105のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、少なくとも30である、発現カセット。

【請求項19】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項18に記載の発現カセットであって、ここで、(i)該X個連続したヌクレオチドが、配列番号72または配列番号105のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii)Xが、Yと等しく、そして(iii)Yが、75である、発現カセット。

50

【請求項 20】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号74または107のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、少なくとも30である、発現カセット。

【請求項 21】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項20に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号74または配列番号107のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、246である、発現カセット 10

【請求項 22】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号76のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、1680である、発現カセット。

【請求項 23】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号78のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、1668である、発現カセット。 20

【請求項 24】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号80、配列番号81または配列番号109のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、216である、発現カセット。

【請求項 25】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号83のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、93である、発現カセット。 30

【請求項 26】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号111のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、90である、発現カセット。

【請求項 27】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号85または配列番号113のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、579である、発現カセット。 40

【請求項 28】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号87のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、288である、発現カセット。

【請求項 29】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号115のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして 50

(i i i) Yが、287である、発現カセット。

【請求項30】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号89または配列番号117のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、少なくとも30である、発現カセット。

【請求項31】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項30に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号89のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、267である、発現カセット。 10

【請求項32】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項30に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号117のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、261である、発現カセット。

【請求項33】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号91のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、少なくとも30である、発現カセット。 20

【請求項34】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項33に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号91のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、321である、発現カセット。

【請求項35】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号93または配列番号94のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、309である、発現カセット。 30

【請求項36】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号96のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、少なくとも60である、発現カセット。

【請求項37】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項36に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号96のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、624である、発現カセット。 40

【請求項38】

X個連続したヌクレオチドを含むポリヌクレオチドを含む発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号119、配列番号120；配列番号121；配列番号122；配列番号123；配列番号124；配列番号125；配列番号126；配列番号127；配列番号131；配列番号132または配列番号133のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(i i) Xが、Yと等しく、そして(i i i) Yが、少なくとも60である、発現カセット。

【請求項39】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項38に記載の発現カセ 50

ットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号119、配列番号120；配列番号121；配列番号122；配列番号123；配列番号124；配列番号125；配列番号126；配列番号127；配列番号131；配列番号132または配列番号133のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、少なくとも300である、発現カセット。

【請求項40】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項39に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号123または配列番号124のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、2433である、発現カセット。 10

【請求項41】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項39に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号122のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、2301である、発現カセット。

【請求項42】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項39に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号125のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、2517である、発現カセット。 20

【請求項43】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項39に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号126または配列番号127のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、2520である、発現カセット。

【請求項44】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項39に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号119のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、1377である、発現カセット。 30

【請求項45】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項39に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号120または配列番号121のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、1839である、発現カセット。

【請求項46】

X個連続したヌクレオチドを含むポリヌクレオチドを含む、請求項39に記載の発現カセットであって、ここで、(i) 該X個連続したヌクレオチドが、配列番号132または配列番号133のY個連続したヌクレオチドに対して少なくとも90%のパーセント同一性を有し、(ii) Xが、Yと等しく、そして(iii) Yが、1890である、発現カセット。 40

【請求項47】

配列番号33に示される配列またはそれから誘導されるフラグメントを含む、ポリヌクレオチド。

【請求項48】

請求項47に記載のポリヌクレオチドであって、ここで、前記フラグメントが、G a g、 50

P o l、V i f、V p r、T a t、R e v、V p u、E n v および N e f からなる群より選択される遺伝子産物についてのコード配列を含む、ポリヌクレオチド。

【請求項 49】

前記フラグメントが、G a g 遺伝子産物を含む、請求項 48 に記載のポリヌクレオチド。

【請求項 50】

前記フラグメントが、E n v 遺伝子産物を含む、請求項 48 に記載のポリヌクレオチド。

【請求項 51】

前記 E n v 遺伝子産物が、g p l 60、g p l 40 または g p l 20 である、請求項 50 に記載のポリヌクレオチド。

【請求項 52】

配列番号 45 に示される配列またはそれから誘導されるフラグメントを含む、ポリヌクレオチド。

【請求項 53】

請求項 52 に記載のポリヌクレオチドであって、ここで、前記フラグメントが、G a g、P o l、V i f、V p r、T a t、R e v、V p u、E n v および N e f からなる群より選択される遺伝子産物についてのコード領域を含む、ポリヌクレオチド。

【請求項 54】

前記フラグメントが、G a g 遺伝子産物を含む、請求項 53 に記載のポリヌクレオチド。

【請求項 55】

前記フラグメントが、E n v 遺伝子産物を含む、請求項 53 に記載のポリヌクレオチド。

【請求項 56】

前記 E n v 遺伝子産物が、g p l 60、g p l 40 または g p l 20 である、請求項 55 に記載のポリヌクレオチド。

【請求項 57】

配列番号 128 に示される配列またはそれから誘導されるフラグメントを含む、ポリヌクレオチド。

【請求項 58】

前記フラグメントが、E n v 遺伝子産物 g p l 60、g p l 40 または g p l 20 についてのコード配列を含む、請求項 57 に記載のポリヌクレオチド。

【請求項 59】

1 つ以上のウイルスポリペプチドまたは抗原をコードする 1 つ以上の核酸をさらに含む、請求項 1 ～ 46 のいずれか 1 項に記載の発現カセット。

【請求項 60】

前記ウイルスポリペプチドまたは抗原が、G a g、E n v、v i f、v p r、t a t、r e v、v p u、n e f およびその組み合わせからなる群より選択される、請求項 59 に記載の発現カセット。

【請求項 61】

1 つ以上のサイトカインをコードする 1 つ以上の核酸をさらに含む、請求項 1 ～ 46 のいずれかに記載の発現カセット。

【請求項 62】

請求項 1 ～ 46 のいずれかに記載の発現カセットを含む、選択された宿主細胞における使用のための組換え発現系であって、前記ポリヌクレオチド配列が、該選択された宿主細胞における発現を駆動し得る制御エレメントをさらに含む、組換え発現系。

【請求項 63】

請求項 62 に記載の組換え発現系であって、ここで、前記制御エレメントが、転写プロモーター、転写エンハンサーエレメント、転写終止シグナル、ポリアデニル化配列、翻訳の開始を最適化するための配列、および翻訳終止配列からなる群より選択される、組換え発現系。

【請求項 64】

請求項 62 に記載の組換え発現系であって、ここで、前記転写プロモーターが、C M V、

10

20

30

40

50

CMV+イントロンA、SV40、RSV、HIV-Ltr、MMLV-ltr、およびメタロチオネインからなる群より選択される、組換え発現系。

【請求項65】

請求項1～46のいずれかに記載の発現カセットを含む細胞であって、ここで、前記ポリヌクレオチド配列が、該選択された細胞における発現と適合性の制御エレメントをさらに含む、細胞。

【請求項66】

前記細胞が、哺乳動物細胞、昆虫細胞、細菌細胞、酵母細胞、植物、抗原提示細胞、一次細胞、不死化細胞、および腫瘍由来の細胞からなる群より選択される、請求項65に記載の細胞。

10

【請求項67】

前記細胞が、BHK細胞、VERO細胞、HT1080細胞、293細胞、RD細胞、COS-7細胞、およびCHO細胞からなる群より選択される、請求項66に記載の細胞。

【請求項68】

前記細胞が、CHO細胞である、請求項67に記載の細胞。

【請求項69】

前記細胞が、Trichoplusia ni (Tn5) または Sf9 昆虫細胞のいずれかである、請求項66に記載の細胞。

【請求項70】

前記抗原提示細胞が、マクロファージ、単球、樹状細胞、B細胞、T細胞、幹細胞、およびその前駆細胞からなる群より選択されるリンパ系細胞である、請求項66に記載の細胞。

20

【請求項71】

請求項1～46のいずれかに記載の発現カセットを含む、免疫学的応答を生じるための組成物。

【請求項72】

1つ以上のPolポリペプチドをさらに含む、請求項71に記載の組成物。

【請求項73】

アジュバントをさらに含む、請求項72に記載の組成物。

【請求項74】

請求項52に記載の発現カセットを含む、免疫学的応答を生じるための組成物。

30

【請求項75】

Polポリペプチドをさらに含む、請求項74に記載の組成物。

【請求項76】

請求項60に記載の核酸分子によってコードされる1つ以上のポリペプチドをさらに含む、請求項74に記載の組成物。

【請求項77】

アジュバントをさらに含む、請求項76に記載の組成物。

【請求項78】

被験体の免疫方法であって、該方法は、請求項71に記載の組成物を、該被験体における前記発現カセットの発現と適合性である条件下で該被験体に導入する工程を包含する、方法。

40

【請求項79】

前記発現カセットが、遺伝子送達ベクターを使用して導入される、請求項78に記載の方法。

【請求項80】

前記遺伝子送達ベクターが、非ウイルスベクターである、請求項79に記載の方法。

【請求項81】

前記遺伝子送達ベクターが、ウイルスベクターである、請求項79に記載の方法。

【請求項82】

50

前記遺伝子送達ベクターが、アデノウイルスベクター、ワクシニアウイルスベクター、AAVベクター、レトロウイルスベクター、レンチウイルスベクター、およびアルファウイルスベクターからなる群より選択される、請求項79に記載の方法。

【請求項83】

前記遺伝子送達ベクターが、シンドビスウイルス由来のベクターである、請求項82に記載の方法。

【請求項84】

前記遺伝子送達ベクターが、cDNAベクターである、請求項82に記載の方法。

【請求項85】

前記遺伝子送達ベクターが、真核生物の層状ウイルス開始系(ELVIS)である、請求項82に記載の方法。 10

【請求項86】

前記組成物が、粒子キャリアを使用して送達される、請求項79に記載の方法。

【請求項87】

前記組成物が、金またはタングステンの粒子上にコーティングされ、そして該コーティングされた粒子が、遺伝子銃を使用して前記被験体に送達される、請求項79に記載の方法。

【請求項88】

前記組成物が、リポソーム調製物にカプセル化されている、請求項79に記載の方法。

【請求項89】

前記被験体が、哺乳動物である、請求項79に記載の方法。 20

【請求項90】

前記哺乳動物が、ヒトである、請求項89に記載の方法。

【請求項91】

被験体における免疫応答を生じる方法であって、該方法は、以下：

請求項1～46のいずれかに記載の発現カセットを提供する工程、

適切な宿主細胞中で前記ポリペプチドを発現する工程、

該ポリペプチドを単離する工程、および

該ポリペプチドを、免疫応答を惹起するのに十分な量で該被験体に投与する工程、
を包含する、方法。 30

【請求項92】

被験体における免疫応答を生じる方法であって、該方法は、該被験体の細胞に、請求項1～46のいずれか1項に記載の発現カセットを、前記ポリヌクレオチドの発現および前記ポリペプチドの産生を可能にする条件下で導入し、これによって該ポリペプチドに対する免疫学的応答を誘発する工程を包含する、方法。

【請求項93】

前記方法が、HIVポリペプチドの同時投与を包含する、請求項92に記載の方法。

【請求項94】

前記被験体への前記ポリペプチドの同時投与が、前記発現カセットの導入の前に行われる、請求項93に記載の方法。 40

【請求項95】

前記被験体への前記ポリペプチドの同時投与が、前記発現カセットの導入と同時に進行される、請求項93に記載の方法。

【請求項96】

前記被験体への前記ポリペプチドの同時投与が、前記発現カセットの導入の後に行われる、請求項93に記載の方法。

【請求項97】

前記ウイルスポリペプチドまたは抗原が、B型肝炎由来のポリペプチド、C型肝炎由来のポリペプチド、およびその組み合わせからなる群より選択される、請求項59に記載の発現カセット。 50

【発明の詳細な説明】

【0001】

(技術分野)

抗原性C型HIVポリペプチドをコードするポリヌクレオチド（例えば、Gag、pol、vif、vpr、tat、rev、vpu、env、およびnef）が記載され、免疫原性組成物におけるこれらのポリヌクレオチドおよびポリペプチド産物の使用も記載される。C型HIVの南アフリカ改変体由来のポリヌクレオチド配列もまた、記載される。

【0002】

(発明の背景)

後天性免疫不全症候群（AIDS）は、現代医学が直面している最も巨大な健康の脅威の1つとして認識されている。この疾患に対する治療法は、未だ存在しない。1983年から1984年において、3つのグループが、AIDSについて疑われる病原性因子を独立して同定した。例えば、Barre-Sinoussiら（1983）*Science* 220:868-871; Montagnierら、Human T-Cell Leukemia Viruses (Gallo, EssexおよびGross編、1984); Vilmerら（1984）*The Lancet* 1:753; Popovicら（1984）*Science* 224:497-500; Levyら（1984）*Science* 225:840-842を参照のこと。これらの単離物は、リンパ節腫脹関連ウイルス（LAV）、ヒトT細胞リンパ指向性ウイルスIII型（HTLV-III）、またはAIDS関連レトロウイルス（ARV）と多様に呼称されている。これらの単離物の全ては、同じウイルスの株であり、そして、これらの単離物の全ては、ヒト免疫不全ウイルス（HIV）と集合的に後に命名された。関連するAIDSを引き起こすウイルスの単離によって、元来HIVと呼称された株は、現在HIV-1と称され、そしてこの関連するウイルスは、HIV-2と呼称される。例えば、Guyaderら（1987）*Nature* 326:662-669; Brun-Vezinetら（1986）*Science* 233:343-346; Clavelら（1986）*Nature* 324:691-695を参照のこと。

【0003】

HIVウイルスについての膨大な情報が集められたが、現在までに効果的なワクチンは同定されていない。ワクチン開発のためのいくつかの標的が調べられ、この標的としては、HIVによってコードされるenv遺伝子産物およびGag遺伝子産物が挙げられる。Gag遺伝子産物としては、Gag-ポリメラーゼおよびGag-プロテアーゼが挙げられるが、これらに限定されない。Env遺伝子産物としては、モノマー性gp120ポリペプチド、オリゴマー性gp140ポリペプチドおよびgp160ポリペプチドが挙げられるがこれらに限定されない。

【0004】

Haasら（*Current Biology* 6(3):315-324, 1996）は、HIV-1による選択的なコドン使用頻度が、ウイルス性タンパク質合成能がないことの実質的な部分の原因と思われることを示唆した。Andreら（*J. Virol.* 72(2):1497-1503, 1998）は、改変したコドン使用頻度と共に合成gp120配列を使用するDNAワクチン接種によって惹起される免疫応答の増大を記載した。Schneiderら（*J. Virol.* 71(7):4892-4903, 1997）は、Gagのコード配列およびGag-プロテアーゼコード配列のコード配列中に位置づけられる阻害（または、不安定化）エレメント（INS）の不活性化を議論する。

【0005】

HIV-1のGagタンパク質は、ウイルス様粒子の集合に必要である。HIV-1 Gagタンパク質は、集合（assembly）、粒子放出後のビリオン成熟、およびウイルス複製における初期の侵入後の段階を含むウイルスの生活環の多くの段階に関連する。このHIV-1 Gagタンパク質の役割は、多数であり、そして複雑である（Free, E. O., *Virology* 251:1-15, 1998）。

【0006】

Wolfr (PCT国際出願、WO 96/30523 (1996年10月3日に発行) ; 欧州特許出願、公開番号0 449 116 A1 (1991年10月2日に発行)) は、特に、免疫学的に重要なエピトープの提示のための非感染性レトロウイルス様粒子状キャリアとして作用するHIV-1の変更したpr55 Gagの用途を記載した。Wang (Virology 200: 524-534, 1994) は、ビリオン中への、HIV Gag-β-ガラクトシダーゼ融合タンパク質の集合を研究するための系を記載する。彼らは、HIV Gag-β-ガラクトシダーゼ融合タンパク質をコードする配列の構築、HIV Gagタンパク質の存在下のこのような配列の発現、およびこれらのタンパク質のウイルス粒子への集合を記載する。

10

【0007】

Shiver (PCT国際出願、WO 98/34640、1998年8月13日発行) は、HIV-1 (CAM1) Gagコード配列を改変して、HIV Gagをコードする合成DNA分子を産生すること、およびHIV Gagの改変を記載する。この合成分子のコドンは、計画された (projected) 宿主細胞によって好まれるコドンであった。

【0008】

最近、免疫原性組成物におけるHIV Envポリペプチドの使用が、記載さる (Hurtwitzらへの米国特許番号第5, 846, 546号 (1998年12月8日発行) (この特許は、各々が異なったHIV env改変体を発現する4つの異なる組み換えウイルスの混合物を含む免疫原性組成物を記載する) ; およびVahlneらへの米国特許番号第5, 840, 313号 (1998年11月24日発行) (この特許は、HIV-1 gp120タンパク質のエピトープに対応するペプチドを記載する))。さらに、Siaらへの米国特許番号第5, 876, 731号 (1999年3月2日発行) は、GPGR配列を含有するHIV-1単離物のV3ループタンパク質のB細胞エピトープのアミノ酸配列に直接的に連結されたGagのT細胞エピトープのアミノ酸配列を含むHIVに対する候補ワクチンを記載する。抗原性HIVポリペプチド (特にC型単離物) に対する必要性がなお存在する。

20

【0009】

(発明の要旨)

本明細書中に記載されるものは、新規のC型HIV配列 (例えば、8__5__TV1__C. ZA、8__2__TV1__C. ZAおよび12-5__1__TV2__C. ZA)、これらの新規配列によってコードされるポリペプチド、およびこれらまたは他のC型HIV配列から生成される合成発現カセットである。

30

【0010】

特定の実施形態において、本発明は、HIV C型ポリペプチド (Env、Gag、Pol、Prot、Vpr、Vpu、Vif、Nef、Tat、Revおよび/またはそれらのフラグメント) をコードする合成発現カセットに関する。さらに、本発明また、HIV C型ポリペプチドの発現の改善、およびウイルス様粒子の産生に関する。HIVポリペプチド (例えば、Gag-、pol-、プロテアーゼ (prot)-、逆転写酵素、インテグラーゼ、RNAseH、Tat、Rev、Nef、Vpr、Vpu、Vifおよび/またはEnvを含有するポリペプチド) をコードする合成発現カセットが、記載され、この発現カセットの使用も記載される。

40

【0011】

従って、本発明の1つの局面は、発現カセットおよびこの中に含まれるポリヌクレオチドに関する。この発現カセットは、代表的には、発現ベクター骨格に挿入されたHIVポリペプチドコード配列を含む。1つの実施形態において、発現カセットは、1以上のPol含有ポリペプチドをコードするポリヌクレオチド配列を含み、ここで、このポリヌクレオチド配列は、本明細書中で教示される少なくとも約85%、好ましくは約90%、より好ましくは約95%、およびより好ましくは約98% (およびこれらの値の間の任意の整数

50

）の配列同一性を有する配列を含む。P o l 含有ポリペプチドをコードするポリヌクレオチド配列としては、配列番号 3 0，配列番号 3 1；配列番号 3 2；配列番号 6 2；配列番号 1 0 3；配列番号 5 8；配列番号 6 0；配列番号 6 4；配列番号 6 6；配列番号 6 8；配列番号 7 0；配列番号 7 6；および配列番号 7 8 に示される配列が挙げられるが、これらに限定されない。

【0012】

本発明の H I V ポリペプチドをコードするポリヌクレオチドはまた、さらなるポリペプチドをコードする配列を含み得る。ポリペプチドをコードするこのようなさらなるポリヌクレオチドは、例えば、他のウイルスタンパク質（例えば、B型肝炎またはC型肝炎あるいは他の H I V のタンパク質）についてのコード配列（例えば、H I V G a g ポリペプチドをコードするポリヌクレオチド配列、H I V E n v ポリペプチドをコードするポリヌクレオチド配列および／または v i f、v p r、t a t、r e v、v p u および n e f のうちの 1 以上をコードするポリヌクレオチド）；サイトカインまたは他の導入遺伝子を含み得る。1つの実施形態において、H I V P o l ポリペプチドをコードする配列は、逆転写酵素およびインテグラーゼに対応するコード領域の欠失によって改変され得る。このポリメラーゼポリペプチドにおけるこのような欠失はまた、作製され得、その結果、このポリヌクレオチド配列は、ヘルパー T 細胞および C T L のエピトープを維持する。目的の他の抗原は、同様にポリメラーゼに挿入され得る。

10

【0013】

別の実施形態において、発現カセットは、H I V G a g 含有ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含み、ここで、G a g ポリペプチドをコードするポリヌクレオチド配列は、本明細書中で教示される配列に対して、少なくとも約 8 5 %、好ましくは約 9 0 %、より好ましくは約 9 5 %、そして最も好ましくは約 9 8 % 配列同一性を有する配列を含む。G a g 含有ポリペプチドをコードするポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない：図 1 のヌクレオチド 8 4 4－9 0 3（G a g の主要な相同性領域）（配列番号 1）；図 2 のヌクレオチド 8 4 1－9 0 0（G a g の主要な相同性領域）（配列番号 2）；図 2 4（配列番号 5 3、G a g の主要な相同性領域）；図 1 として提示される配列（配列番号 3）；図 2 2 として提示される配列（配列番号 5 1）；図 7 0 として提示される配列（配列番号 9 9）；および図 2 として提示される配列（配列番号 4）。上述のように、本発明の G a g 含有ポリペプチドをコードするポリヌクレオチドはまた、さらなるポリペプチドをコードする配列を含み得る。

20

30

【0014】

別の実施形態において、発現カセットは、H I V E n v 含有ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含有し、ここで、E n v ポリペプチドをコードするポリヌクレオチド配列は、本明細書中で教示される配列に対して、少なくとも約 8 5 %、好ましくは約 9 0 %、より好ましくは約 9 5 %、そして最も好ましくは約 9 8 % の配列同一性を有する配列を含む。E n v 含有ポリペプチドをコードするポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない：図 3 のヌクレオチド 1 2 1 3～1 3 5 3（配列番号 5）（これは、E n v 共通領域をコードする）；図 1 7 として提示される配列（配列番号 4 6）（これは、9 7 ヌクレオチド長の E n v 共通領域をコードする）；配列番号 4 7（これは、1 4 4 ヌクレオチド長の E n v 共通領域をコードする）；図 3 のヌクレオチド 8 2～1 5 1 2（配列番号 6）（これは、g p 1 2 0 ポリペプチドをコードする）；図 3 のヌクレオチド 8 2～2 0 2 5（配列番号 7）（これは、g p 1 4 0 ポリペプチドをコードする）；図 3 のヌクレオチド 8 2～2 5 4 7（配列番号 8）（これは、g p 1 6 0 ポリペプチドをコードする）；配列番号 4 9（これは、g p 1 6 0 ポリペプチドをコードする）；図 3 のヌクレオチド 1～2 5 4 7（配列番号 9）（これは、シグナル配列を有する g p 1 6 0 ポリペプチドをコードする）；図 3 のヌクレオチド 1 5 1 3～2 5 4 7（配列番号 1 0）（これは、g p 4 1 ポリペプチドをコードする）；図 4 のヌクレオチド 1 2 1 0～1 3 5 3（配列番号 1 1）（これは、E n

40

50

v 共通領域をコードする) ; 図4のヌクレオチド73~1509 (配列番号12) (これは、g p l 2 0 ポリペプチドをコードする) ; 図4のヌクレオチド73~2022 (配列番号13) (これは、g p l 4 0 ポリペプチドをコードする) ; 図4のヌクレオチド73~2565 (配列番号14) (これは、g p l 6 0 ポリペプチドをコードする) ; 図4のヌクレオチド1~2565 (配列番号15) (これは、シグナル配列を有するp l 6 0 ポリペプチドをコードする) ; 図20として提示される配列 (配列番号49) (これは、g p l 6 0 ポリペプチドをコードする) ; 図68として提示される配列 (配列番号97) (これは、g p l 6 0 ポリペプチドをコードする) ; 図4のヌクレオチド1510-2565 (配列番号16) (これは、g p 4 1 ポリペプチドをコードする) ; 図90のヌクレオチド7~1464 (配列番号119) (これは、改変された野生型シグナル配列を有するg p l 2 0 ポリペプチドをコードする) ; 図91のヌクレオチド7~1977 (配列番号120) (これは、野生型8__2__TV1__C. ZAから改変されたシグナル配列 (例えば、「改変された野生型リーダー配列」) を含むg p l 4 0 ポリペプチドをコードする) ; 図92のヌクレオチド7~1977 (配列番号121) (これは、改変された野生型8__2__TV1__C. ZAシグナル配列を有するg p l 4 0 ポリペプチドをコードする) ; 図93のヌクレオチド7~2388 (配列番号122) (これは、改変された野生型シグナル配列を有するg p l 6 0 ポリペプチドをコードする) ; 図94のヌクレオチド7~2520 (配列番号123) (これは、改変された野生型8__2__TV1__C. ZAシグナル配列を有するg p l 6 0 ポリペプチドをコードする) ; 図95のヌクレオチド7~2520 (配列番号124) (これは、改変された野生型8__2__TV1__C. ZAシグナル配列を有するg p l 6 0 ポリペプチドをコードする) ; 図96のヌクレオチド13~2604 (配列番号125) (これは、TPA1シグナル配列を有するg p l 6 0 ポリペプチドをコードする) ; 図97のヌクレオチド7~2607 (配列番号126) (これは、改変された野生型8__2__TV1__C. ZAシグナル配列を有するg p l 6 0 ポリペプチドをコードする) ; 図100のヌクレオチド1~2049 (配列番号131) (これは、TPA1シグナル配列を有するg p l 4 0 ポリペプチドをコードする) ; 図98のヌクレオチド7~1607 (配列番号126) (これは、野生型8__2__TV1__C. ZAシグナル配列を有するg p l 6 0 ポリペプチドをコードする) ; 配列番号132のヌクレオチド7~2064 (これは、野生型8__2__TV1__C. ZAリーダー配列を有するg p l 4 0 ポリペプチドをコードする) ; および配列番号133のヌクレオチド7~2064 (これは、改変された野生型8__2__TV1__C. ZAリーダー配列を有するg p l 4 0 ポリペプチドをコードする)。

【0015】

特定の実施形態において、Envコード配列は、さらなる改変、例えば、g p l 2 0 ポリペプチドおよびg p 4 1 ポリペプチド (配列番号121および配列番号124) へのg p l 4 0 ポリペプチドの切断または可変領域V1および/またはV2 (配列番号119; 配列番号120; 配列番号121; 配列番号122; 配列番号123; および配列番号124) の欠失を防ぐための切断部位の変異を含む。

【0016】

別の実施形態において、発現カセットは、HIV Nef含有ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含み、ここで、Nefポリペプチドをコードするポリヌクレオチド配列は、本明細書中で教示される配列に対して、少なくとも約85%、好ましくは約90%、より好ましくは約95%、そして最も好ましくは約98%の配列同一性を有する配列を含む。Nef含有ポリペプチドをコードするポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない: 図26において提示される配列 (配列番号55) ; 図72で提示される配列 (配列番号101) ; 図28で提示される配列 (配列番号57) ; 図67で提示される配列 (配列番号96) ; 図103で提示される配列 (配列番号134) ; および図104で提示される配列 (配列番号135)。

【0017】

別の実施形態において、発現カセットは、H I V R e v 含有ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含み、ここで、R e v ポリペプチドをコードするポリヌクレオチド配列は、本明細書中で教示される配列に対して、少なくとも約 85 %、好ましくは約 90 %、より好ましくは約 95 %、そして最も好ましくは約 98 % の配列同一性を有する配列を含む。R e v 含有ポリペプチドをコードするポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない：図 43 で提示される配列（配列番号 72）；図 76 で提示される配列（配列番号 105）；図 45 で提示される配列（配列番号 74）；図 78 で提示される配列（配列番号 107）；および図 62 で提示される配列（配列番号 91）。

【0018】

別の実施形態において、発現カセットは、H I V T a t 含有ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含み、ここで、T a t ポリペプチドをコードするポリヌクレオチド配列は、本明細書中で教示される配列に対して、少なくとも約 85 %、好ましくは約 90 %、より好ましくは約 95 %、そして最も好ましくは約 98 % の配列同一性を有する配列を含む。T a t 含有ポリペプチドをコードするポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない：図 51 で提示される配列（配列番号 80）；図 80 で提示される配列（配列番号 109）；図 52 で提示される配列（配列番号 81）；図 54 で提示される配列（配列番号 83）；および図 82 で提示される配列（配列番号 111）。

【0019】

別の実施形態において、発現カセットは、H I V V i f 含有ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含み、ここで、V i f ポリペプチドをコードするポリヌクレオチド配列は、本明細書中で教示される配列に対して、少なくとも約 85 %、好ましくは約 90 %、より好ましくは約 95 %、そして最も好ましくは約 98 % の配列同一性を有する配列を含む。V i f 含有ポリペプチドをコードするポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない：図 56 で提示される配列（配列番号 85）；および図 84 で提示される配列（配列番号 113）。

【0020】

別の実施形態において、発現カセットは、H I V V p r 含有ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含み、ここで、V p r ポリペプチドをコードするポリヌクレオチド配列は、本明細書中で教示される配列に対して、少なくとも約 85 %、好ましくは約 90 %、より好ましくは約 95 %、そして最も好ましくは約 98 % の配列同一性を有する配列を含む。V p r 含有ポリペプチドをコードするポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるがこれらに限定されない：図 58 で提示される配列（配列番号 87）；および図 86 で提示される配列（配列番号 115）。

【0021】

別の実施形態において、発現カセットは、H I V V p u 含有ポリペプチドを含むポリペプチドをコードするポリヌクレオチド配列を含み、ここで、V p u ポリペプチドをコードするポリヌクレオチド配列は、本明細書中で教示される配列に対して、少なくとも約 85 %、好ましくは約 90 %、より好ましくは約 95 %、そして最も好ましくは約 98 % の配列同一性を有する配列を含む。V p u 含有ポリペプチドをコードするポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない：図 60 で提示される配列（配列番号 89）；および図 88 で提示される配列（配列番号 117）。

【0022】

本発明のさらなる実施形態としては、本明細書中で記載された配列のいずれかの精製されたポリヌクレオチドが挙げられる。G a g 含有ポリペプチドをコードする例示的なポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない：図 1 のヌクレオチド 844～903（配列番号 1）（G a g の主要な相同性領域）；図 2 のヌクレオチド 841～900（配列番号 2）（G a g の主要な相同性領域）；図 1 として提示される配列（配列番号 3）；図 2 として提示される配列（配列番号 4）；図

10

20

30

40

50

22として提示される配列（配列番号51）；図70として提示される配列（配列番号99）；および図24として提示される（配列番号53）（Gagの主要な相同性領域）。

【0023】

Env含有ポリペプチドをコードする、例示的なポリヌクレオチド配列としては、以下のポリヌクレオチドが挙げられるが、これらに限定されない：図3のヌクレオチド1213～1353（配列番号5）（Env共通領域をコードする）；図17として示される配列（配列番号46）（97ヌクレオチド長のEnv共通領域をコードする）；配列番号47（144ヌクレオチド長のEnv共通領域をコードする）；図3のヌクレオチド82～1512（配列番号6）（gp120ポリペプチドをコードする）；図3のヌクレオチド82～2025（配列番号7）（gp140ポリペプチドをコードする）；図3のヌクレオチド82～2547（配列番号8）（gp160ポリペプチドをコードする）；配列番号49（gp160ポリペプチドをコードする）；図3のヌクレオチド1～2547（配列番号9）（シグナル配列を有するgp160ポリペプチドをコードする）；図3のヌクレオチド1513～2547（配列番号10）（gp41ポリペプチドをコードする）；図4のヌクレオチド1210～1353（配列番号11）（Env共通領域をコードする）；図4のヌクレオチド73～1509（配列番号12）（gp120ポリペプチドをコードする）；図4のヌクレオチド73～2022（配列番号13）（gp140ポリペプチドをコードする）；図4のヌクレオチド73～2565（配列番号14）（gp160ポリペプチドをコードする）；図4のヌクレオチド1～2565（配列番号15）（シグナル配列を有するgp160ポリペプチドをコードする）；図20として示される配列（配列番号49）（gp160ポリペプチドをコードする）；図68として示される配列（配列番号97）（gp160ポリペプチドをコードする）；図4のヌクレオチド1510～2565（配列番号16）（gp41ポリペプチドをコードする）；図90のヌクレオチド7～1464（配列番号119）（改変した野生型シグナル配列を有するgp120ポリペプチドをコードする）；図91のヌクレオチド7～1977（配列番号120）（野生型8__2__TV1__C．ZAから改変されたシグナル配列（例えば、「改変した野生型リーダー配列」）を含む、gp140ポリペプチドをコードする）；図92のヌクレオチド7～1977（配列番号121）（改変した野生型8__2__TV1__C．ZAシグナル配列を有するgp140ポリペプチドをコードする）；図93のヌクレオチド7～2388（配列番号122）（改変した野生型シグナル配列を有するgp160ポリペプチドをコードする）；図94のヌクレオチド7～2520（配列番号123）（改変した野生型8__2__TV1__C．ZAシグナル配列を有するgp160ポリペプチドをコードする）；図95のヌクレオチド7～2520（配列番号124）（改変した野生型8__2__TV1__C．ZAシグナル配列を有するgp160ポリペプチドをコードする）；図96のヌクレオチド13～2604（配列番号125）（TPA1シグナル配列を有するgp160ポリペプチドをコードする）；図97のヌクレオチド7～2607（配列番号126）（改変した野生型8__2__TV1__C．ZAシグナル配列を有するgp160ポリペプチドをコードする）；図100のヌクレオチド1～2049（配列番号131）（TPA1シグナル配列を有するgp140ポリペプチドをコードする）；図98のヌクレオチド7～1607（配列番号126）（野生型8__2__TV1__C．ZAシグナル配列を有するgp160ポリペプチドをコードする）；配列番号132のヌクレオチド7～2064（改変した野生型8__2__TV1__C．ZAリーダー配列を有するgp140ポリペプチドをコードする）；および配列番号133のヌクレオチド7～2064（野生型8__2__TV1__C．ZAリーダー配列を有するgp140ポリペプチドをコードする）。

【0024】

さらなるHIVポリヌクレオチドをコードする、例示的な精製されたポリヌクレオチドとしては、以下が挙げられる：Polをコードするポリヌクレオチド（例えば、配列番号30、配列番号31；配列番号32；配列番号62；配列番号103；配列番号58；配列番号60；配列番号64；配列番号66；配列番号68；配列番号70；配列番号76；および配列番号78）；Nefをコードするポリヌクレオチド（例えば、配列番号55；

10

20

30

40

50

配列番号 101；配列番号 57；配列番号 96）；Rev をコードするポリヌクレオチド（例えば、配列番号 72；配列番号 105；配列番号 74；配列番号 107；配列番号 91）；Tat をコードするポリヌクレオチド（例えば、配列番号 80；配列番号 109；配列番号 81；配列番号 83；配列番号 111）；Vif をコードするポリヌクレオチド（例えば、配列番号 85；配列番号 113）；ならびに Vpr をコードするポリヌクレオチド（例えば、配列番号 87；配列番号 115）；Vpu をコードするポリヌクレオチド（例えば、配列番号 89；配列番号 117）。

【0025】

他の実施形態において、本発明は、新規の C 型株より得られるネイティブの HIV ポリペプチドをコードする配列；これらのネイティブの配列のフラグメント；これらの野生型配列を含む発現カセット；ならびに、これらの配列、フラグメントおよび発現カセットの使用に関する。例示的な全長配列は、配列番号 33 および配列番号 45 に示される。種々の HIV 遺伝子産物をコードする例示的なフラグメントとしては、以下が挙げられる：図 19 に示される配列（配列番号 48）（Env をコードする配列）；図 69 に示される配列（配列番号 98）（Env をコードする配列）；図 21 に示される配列（配列番号 50）（gp160 ポリペプチド）；図 23 に示される配列（配列番号 52）（Gag ポリペプチド）；図 71 に示される配列（配列番号 100）（Gag ポリペプチド）；図 25 に示される配列（配列番号 54）（Gag ポリペプチド）；図 27 に示される配列（配列番号 56）（Nef ポリペプチド）；図 73 に示される配列（配列番号 102）（Nef ポリペプチド）；図 30 に示される配列（配列番号 59）（p15 RNAseH ポリペプチド）；図 32 に示される配列（配列番号 61）（p31 インテグラーゼポリペプチド）；図 34 に示される配列（配列番号 63）（Pol ポリペプチド）；図 75 に示される配列（配列番号 104）（Pol ポリペプチド）；図 36 に示される配列（配列番号 65）（Prot ポリペプチド）；図 38 に示される配列（配列番号 67）（不活化 Prot ポリペプチド）；図 40 に示される配列（配列番号 69）（不活化 Prot ポリペプチドおよび RT ポリペプチド）；図 42 に示される配列（配列番号 71）（Prot ポリペプチドおよび RT ポリペプチド）；図 44 に示される配列（配列番号 73）（Rev ポリペプチド）；図 77 に示される配列（配列番号 106）（Rev ポリペプチド）；図 46 に示される配列（配列番号 75）（Rev ポリペプチド）；図 79 に示される配列（配列番号 108）（Rev ポリペプチド）；図 48 に示される配列（配列番号 77）（RT ポリペプチド）；図 50 に示される配列（配列番号 79）（変異した RT ポリペプチド）；図 53 に示される配列（配列番号 82）（Tat ポリペプチド）；図 81 に示される配列（配列番号 110）（Tat ポリペプチド）；図 55 に示される配列（配列番号 84）（Tat ポリペプチド）；図 83 に示される配列（配列番号 112）（Tat ポリペプチド）；図 57 に示される配列（配列番号 86）（Vif ポリペプチド）；図 85 に示される配列（配列番号 114）（Vif ポリペプチド）；図 59 に示される配列（配列番号 88）（Vpr ポリペプチド）；図 82 に示される配列（配列番号 116）（Vpr ポリペプチド）；図 61 に示される配列（配列番号 90）（Vpu ポリペプチド）；図 89 に示される配列（配列番号 118）（Vpu ポリペプチド）；図 63 に示される配列（配列番号 92）（Rev ポリペプチド）；ならびに図 66 に示される配列（配列番号 95）（Tat ポリペプチド）。

【0026】

本発明の HIV ポリペプチドをコードするネイティブなポリヌクレオチド配列および合成ポリヌクレオチド配列は、代表的には、本明細書中に教示される配列と、少なくとも約 85%、好ましくは、約 90%、より好ましくは、約 95%、そして最も好ましくは、約 98% の配列同一性を有する。さらに、特定の実施形態において、本発明の HIV ポリペプチドをコードするポリヌクレオチド配列は、本明細書中に教示される配列と、100% の配列同一性を示す。

【0027】

本発明のポリヌクレオチドは、組換え技術、合成技術またはそれらの組み合わせによって

生成され得る。

【0028】

本発明は、さらに、選択した宿主細胞における使用のための組換え発現系を包含し、ここで、この組換え発現系は、本発明の1つ以上のポリヌクレオチドおよび発現カセットを用いる。このような系において、ポリヌクレオチド配列は、選択した宿主細胞における発現と適合性のコントロールエレメントに作動可能に連結される。多数の発現コントロールエレメントが、当該分野で公知であり、これらとしては、以下が挙げられるがこれらに限定されない：転写プロモーター、転写エンハンサーエレメント、転写終止シグナル、ポリアダニル化配列、翻訳の開始を最適化するための配列、および翻訳終止配列。例示的な転写プロモーターとしては、CMV、CMV+イントロンA、SV40、RSV、HIV-Ltr、MMLV-Ltrおよびメタロチオネンから誘導されるものが挙げられるが、これらに限定されない。

【0029】

本発明の別の局面は、本発明の1つ以上の発現カセットを含む細胞を含み、ここで、ポリヌクレオチド配列は、選択された細胞における発現と適合性の制御エレメントに作動可能に連結される。1つの実施形態において、このような細胞は、哺乳動物細胞である。例示的な哺乳動物細胞としては、BHK細胞、VERO細胞、HT1080細胞、293細胞、RD細胞、COS-7細胞およびCHO細胞が挙げられるがこれらに限定されない。本発明の実施において有用であり得る他の細胞、細胞型、組織型などとしては、以下から得られるものが挙げられるがこれらに限定されない：昆虫（例えば、Trichoplusia ni (Tn5) および Sf9）、細菌、酵母、植物、抗原提示細胞（例えば、マクロファージ、単球、樹状細胞、B細胞、T細胞、幹細胞およびそれらの前駆細胞）、一次細胞、不死化細胞、腫瘍誘導細胞。

【0030】

さらなる局面において、本発明は、免疫学的応答を産生するための組成物を包含し、ここで、この組成物は、代表的には、本発明の少なくとも1つの発現カセットを含有し、そして例えば、発現カセット（例えば、Polポリペプチドをコードするポリヌクレオチドを有する1つ以上の発現カセット、Gagポリペプチドをコードするポリヌクレオチドを有する1つ以上の発現カセット、付随ポリペプチドをコードするポリヌクレオチド（例えば、ネイティブもしくは合成の、vpv、vpr、nef、vif、tat、rev）を有する1つ以上の発現カセット、および／またはEnvポリペプチドをコードするポリヌクレオチドを有する1つ以上の発現カセット）の組み合わせを含有し得る。このような組成物は、さらに、アジュバント（単数または複数）を含有し得る。これらの組成物はまた、1つ以上のC型HIVポリペプチドを含有し得る。このC型HIVポリペプチドは、組成物中の発現カセットによってコードされるポリペプチドに対応し得るか、または発現カセットによってコードされるものとは異なり得る。組成物中に提供されるように、同じポリペプチドをコードする発現カセットにおけるポリヌクレオチドの例は、以下の通りである：この発現カセットにおけるポリヌクレオチドは、図1のGagポリペプチド（配列番号3）をコードし、そしてこのポリペプチド（配列番号17）は、図1に示される配列によってコードされるポリペプチドである。組成物中に提供されるように、異なるポリペプチドをコードする発現カセットにおけるポリヌクレオチドの例は、以下の通りである：Gagポリメラーゼポリペプチドをコードするポリヌクレオチドを有する発現カセット。そして、組成物中に提供されるこのポリペプチドは、Gagポリペプチドおよび／またはGagプロテアーゼポリペプチドであり得る。発現カセット（または、本発明のポリヌクレオチド）およびポリペプチドの両方を含有する組成物において、本発明の種々の発現カセットが混合され得、そして／または本明細書中に記載される種々のC型HIVポリペプチドと適合され得る。

【0031】

別の局面において、本発明は、被験体の免疫化の方法を包含する。本方法において、上記の組成物のいずれかは、被験体中の発現カセットの発現に適合性である条件下で、被験体

中にある。1つの実施形態において、発現カセット（または、本発明のポリヌクレオチド）は、遺伝子送達ベクターを使用して導入され得る。この遺伝子送達ベクターは、例えば、非ウイルスベクター、またはウイルスベクターであり得る。例示的なウイルスベクターとしては、S i n d b i s ウイルス由来ベクター、レトロウイルスベクターおよびレンチウイルスベクターが挙げられるがこれらに限定されない。免疫学的応答を生じるために有用な組成物はまた、特定のキャリアを使用して送達され得る。さらに、このような組成物は、例えば、ゴールド粒子またはタングステン粒子、および例えば、遺伝子銃を使用して被験体に送達されるコーティングされた粒子上にコーティングされ得る。これらの組成物はまた、リポソームとして処方され得る。この方法の1つの実施形態において、被験体は、哺乳動物であり、そして例えば、ヒトであり得る。

10

【0032】

さらなる局面において、本発明は、被験体において免疫応答を生じる方法を包含する。本明細書中に記載される発現カセットのいずれかは、適切な細胞において発現されて、本発明のポリヌクレオチドによってコードされるC型H I Vポリペプチドの発現を提供し得る。次いで、ポリペプチドは単離（例えば、続いて精製）され、免疫応答を誘発するのに十分な量で被験体に投与される。特定の実施形態において、本方法は、本明細書中に記載される遺伝子送達技術のいずれかを使用して、本発明の1つ以上の発現カセットまたはポリヌクレオチドを投与する工程を包含する。他の実施形態において、本方法は、本発明の1つ以上の発現カセットまたはポリヌクレオチドおよび1つ以上のポリペプチドを同時投与する工程を包含し、ここで、これらのポリペプチドは、これらのポリヌクレオチドから発

20

【0033】

本発明はさらに、被験体において免疫応答を生じる方法を包含し、ここで、被験体の細胞は、選択されたポリヌクレオチドの発現および目的のポリペプチド（例えば、本発明の任意の発現カセットによってコードされる）の産生が可能な条件下で、上記の発現カセットまたは本発明のポリヌクレオチドのいずれかを用いてトランスフェクトされる。この方法によって、ポリペプチドに対する免疫応答が、被験体で誘発される。細胞のトランスフェクションは、エキソピボで行われ得、そしてトランスフェクトされた細胞は、被験体に再導入される。あるいは、またはさらに、細胞は被験体においてインピボでトランスフェクトされ得る。免疫応答は、体液性および／または細胞媒介性（細胞性）であり得る。さらなる実施形態において、この方法はまた、被験体への発現カセットの導入前、導入時および／または導入後に、C型H I Vポリペプチドを投与する工程を包含し得る。

30

【0034】

本発明のこれらおよび他の実施形態は、本明細書中の開示の点から、当業者に容易に思い浮かぶことである。

【0035】

（発明の詳細な説明）

本発明の実施は、他で示されない限り、当業者の範囲内の、化学、生化学、分子生物学、免疫学および薬理学の従来の方法を用いる。このような技術は、文献において十分に説明される。例えば、Remington's Pharmaceutical Sciences, 第18版 (Easton, Pennsylvania: Mack Publishing Company, 1990); Methods In Enzymology (S. ColowickおよびN. Kaplan編、Academic Press, Inc.) ; および Handbook of Experimental Immunology, 第I~IV巻 (D. M. WeirおよびC. C. Blackwell編、1986, Blackwell Scientific Publications) ; Samb

40

50

rookら、Molecular Cloning: A Laboratory Manual (第2版、1989); Short Protocols in Molecular Biology、第4版 (Ausubelら編、1999、John Wiley & Sons); Molecular Biology Techniques: An Intensive Laboratory Course, (Reamら編、1998, Academic Press); PCR (Introduction to Biotechniques Series)、第二版 (Newton & Graham編、1997, Springer Verlag) を参照のこと。

【0036】

本明細書中および添付の特許請求の範囲で使用される場合、単数形、「1つの (a、an)」および「この、その (the)」は、内容が別に明確に指示されない限り、複数の対象を含む。従って、例えば、参考として「1つの抗原 (an antigen)」は、このような因子の2以上の混合物を含む。

【0037】

(1. 定義)

本発明の記載において、以下の用語が用いられ、そして以下に示すように規定されることが意図される。

【0038】

本明細書において用いる場合、「合成」配列とは、C型HIVポリペプチドをコードするポリヌクレオチドであって、その発現が本明細書において記載されるように、例えば、コード置換および阻害性配列の不活性化によって、改変されているポリヌクレオチドをいう。「野生型」または「ネイティブ」の配列とは、本明細書において用いる場合、ポリペプチドコード配列であって、それが天然に見出されるかのように本質的であるポリペプチドコード配列をいう。例えば、C型単離体において見出されるような、Gag、Pol、Vif、Vpr、Tat、Rev、Vpu、Env、および／またはNefのコード配列 (例えば、AF110965、AF110967、AF110968、AF110975、8_5_TV1_C.ZA、8_2_TV1_C.ZAまたは12-5_1_TV2_C.ZA。HIVゲノムの種々の領域を、8_5_TV1_C.ZA (配列番号33) に対して番号付けをして、表Aに示す。従って、用語「Pol」は、以下のポリペプチドの1つ以上をいう：ポリメラーゼ (p6Pol); プロテアーゼ (prot); 逆転写酵素 (p66RTまたはRT); RNaseH (p15RNaseH); および／またはインテグラーゼ (p31IntまたはInt)。

【0039】

本明細書において用いる場合、用語「ウイルス様粒子」または「VLP」とは、以下にさらに考察する任意のいくつかのウイルス由来の非複製性、ウイルスシェルをいう。VLPは、一般に1つ以上のウイルスタンパク質 (例えば、キャプシドタンパク質、コートタンパク質、シェルタンパク質、表面タンパク質、および／またはエンベロープタンパク質と呼ばれるタンパク質が挙げられるがこれらに限定されない)、またはこれらのタンパク質由来の粒子形成ポリペプチドから構成される。VLPは、適切な発現系におけるタンパク質の組み換え発現の際に自然に形成され得る。特定のVLPを生成するための方法は当該分野で公知であり、そして以下にさらに詳細に考察される。ウイルスタンパク質の組み換え発現後のVLPの存在は、当該分野で公知の従来の技術を用いて (例えば、顕微鏡検査 (法)、X線結晶学、などによって) 検出され得る。例えば、Bakerら、Biophys. J. (1991) 60: 1445~1456; Hagenseeら、J. Virol. (1994) 68: 4503~4505を参照のこと。例えば、VLPを、密度勾配遠心分離によって単離するか、および／または特徴的密度バンド形成によって同定し得る。あるいは、低温電子顕微鏡検査 (法) を、問題のVLP調製物のガラス化されたサンプル上で実施し、そして適切な曝露条件下で画像を記録し得る。

【0040】

特定のウイルスタンパク質由来の「粒子形成ポリペプチド」とは、VLP形成を支持する

条件下でVLPを形成する能力を有する、全長またはほぼ全長のウイルスタンパク質、ならびにそのフラグメント、または内部欠失を有するウイルスタンパク質を意味する。従って、このポリペプチドは、参照分子の全長配列、フラグメント、短縮配列および部分配列、ならびにアナログ形態および前駆体形態を含み得る。従って、この用語は、このポリペプチドがVLPを形成する能力を保持する限り、この配列に対して欠失、付加、および置換を意図する。従ってこの用語は、特定化されたポリペプチドの天然のバリエーションを含む。なぜならコートタンパク質におけるバリエーションがしばしば、ウイルス単離体の間で生じるからである。この用語は、このタンパク質がVLPを形成する能力を保持する限り、参照タンパク質において天然に存在しない欠失、付加、および置換を含む。好ましい置換は、事実上保存的である置換（すなわち、アミノ酸の側鎖に関連してアミノ酸のファミリー内に生じる置換）である。詳細には、アミノ酸は、一般に、以下の4つのファミリーに分類される：（1）酸性－アスパラギン酸、およびグルタミン酸；（2）塩基性－リジン、アルギニン、ヒスチジン；（3）非極性－アラニン、バリン、ロイシン、イソロイシン、プロリン、フェニルアラニン、メチオニン、トリプトファン；および（4）非荷電極性－グリシン、アスパラギン、グルタミン、システイン、セリン、トレオニン、チロシン。フェニルアラニン、トリプトファン、およびチロシンは、時々、芳香族アミノ酸として分類される。

【0041】

「抗原」とは、宿主の免疫系を刺激して、体液性および／または細胞性抗原特異的応答を起こす1以上のエпитープ（線状、高次構造的、またはその両方のいずれか）を含む分子をいう。この用語は、用語「免疫原」と交換可能に用いられる。通常、B細胞エпитープは、少なくとも約5アミノ酸を含むが、3～4アミノ酸まで小さくても可能である。T細胞エпитープ（例えば、CTLエпитープ）は、少なくとも約7～9アミノ酸を含み、そしてヘルパーT細胞エпитープは、少なくとも約12～20アミノ酸を含む。通常、エпитープは、約7と15の間のアミノ酸（例えば、9、10、12、または15アミノ酸）を含む。用語「抗原」とは、サブユニット抗原（すなわち、その抗原が天然に関連している生物全体から分離し、区別される抗原）ならびに、殺傷されるか、弱毒化されるか、または不活性化された細菌、ウイルス、真菌、寄生生物、もしくは他の微生物の両方を示す。抗体（例えば、抗イディオタイプ抗体またはそのフラグメント、および抗原または抗原決定基を模倣し得る合成ペプチドミモトープ）もまた、本明細書中で使用され得る抗原の定義に入る。同様に、例えば、遺伝子治療およびDNA免疫適用においてインビボで抗原または抗原性決定基を発現するオリゴヌクレオチドまたはポリペプチドもまた、本明細書中の抗原の定義に含まれる。

【0042】

本発明の目的のために、抗原は、以下でより十分に記載されるように、いくつかの公知のウイルス、細菌、寄生生物、および真菌に由来し得る。この用語はまた、種々の腫瘍抗原のいずれかを意図し得る。さらに、本発明の目的のために、「抗原」とは、本明細書中に定義しているように、タンパク質が免疫学的応答を誘発する能力を維持する限り、ネイティブ配列に対する改変（例えば、欠失、付加、および置換（一般的には、事実上保存性））を含むタンパク質をいう。これらの改変は、部位特異的変異誘発を通してのような意図的なものであり得るか、または抗原を産生する宿主の変異を通してのような偶然によるものであり得る。

【0043】

抗原または組成物に対する「免疫応答」は、目的の組成物中に存在する抗原に対する体液性および／または細胞性免疫応答の被験体における発生である。本発明の目的のために、「体液性免疫応答」とは、抗体分子によって媒介される免疫応答をいうが、一方、「細胞性免疫応答」は、Tリンパ球および／または他の白血球によって媒介される免疫応答である。細胞性免疫の1つの重要な局面は、細胞溶解性T細胞（「CTL」）による抗原特異的応答を含む。CTLは、主要組織適合遺伝子複合体（MHC）によってコードされるタンパク質に付随して存在するペプチド抗原、および細胞表面上で発現するペプチド抗原に

対する特異性を有する。CTLは、細胞内微生物の破壊の誘導および促進を補助するか、またはこのような微生物に感染した細胞の溶解を補助する。細胞性免疫の別の局面は、ヘルパーT細胞による抗原特異的応答を含む。ヘルパーT細胞は、細胞表面上でMHC分子に結合するペプチド抗原を提示する細胞に対して、その機能を刺激すること、および非特異的エフェクター細胞の活性を集中することを補助するように作用する。「細胞性免疫応答」はまた、サイトカイン、ケモカイン、および活性化されたT細胞および／または他の白血球（CD4＋およびCD8＋ T細胞から誘導される白血球を含む）によって産生される他のそのような分子の産生をいう。

【0044】

細胞性免疫応答を誘発する組成物またはワクチンは、細胞表面でのMHC分子と結合する抗原の提示によって、脊椎動物被験体感を作するように働き得る。細胞媒介性免疫応答は、それらの表面で抗原を提示する細胞に、またはその細胞の近傍に方向付けられる。さらに、抗原特異的Tリンパ球は、免疫された宿主の将来の防御を可能にするために産生され得る。

【0045】

特定の抗原が細胞媒介性の免疫学的応答を刺激する能力は、多くのアッセイ（例えば、リンパ球増殖（リンパ球活性化）アッセイ、CTL細胞毒性細胞アッセイ、または感作された被験体中で抗原について特異的なTリンパ球についてアッセイすること）によって決定され得る。このようなアッセイは、当該分野で周知である。例えば、Ericssonら、J. Immunol. (1993) 151: 4189-4199; Doeら、Eur. J. Immunol. (1994) 24: 2369-2376を参照のこと。細胞媒介免疫応答を測定する最近の方法には、細胞内サイトカインまたはT細胞集団によるサイトカイン分泌の測定、またはエプトープ特異的T細胞の測定（例えば、テトラマー技術による）(McMichael, A. J. およびO'Callaghan, C. A., J. Exp. Med. 187 (9): 1367-1371, 1998; Mcheyzer-Williams, M. G. ら、Immunol. Rev. 150: 5-21, 1996; Lallvani, A. ら、J. Exp. Med. 186: 859-865, 1997によって概説される)が含まれる。

【0046】

従って、本明細書中で使用されるような免疫学的応答は、CTLの産生および／またはヘルパーT細胞の産生または活性化を刺激するものであり得る。目的の抗原はまた、抗体媒介免疫応答を誘発し得る。従って、免疫学的応答は、以下の1つ以上の効果を含み得る：B細胞による抗体の産生；および／またはサプレッサーT細胞の活性化および／または目的の組成物もしくはワクチン中に存在する抗原に特異的に方向付けられた $\gamma\delta$ T細胞。これらの応答は、感染性を中和するように、および／または抗体-補体もしくは抗体依存性細胞毒性(ADCC)を媒介するように働き、免疫された宿主に対する防御を提供し得る。このような応答は、当該分野で周知の標準的なイムノアッセイおよび中和アッセイを用いて決定され得る。

【0047】

「免疫原性組成物」とは、被験体への組成物の投与が、目的の抗原性分子に対して体液性免疫応答および／または細胞性免疫応答のこの被験体における発生を生じる抗原性分子を含む組成物である。この免疫原性組成物は、注射投与、吸入投与、経口投与、経鼻投与、および粘膜投与（例えば、経直腸的または経膈的に）などによって、レシピエント被験体に直接導入され得る。

【0048】

「サブユニットワクチン」とは、目的の病原体由来（例えば、ウイルス、細菌、寄生生物、または真菌由来）の抗原に由来するか、またはそれに対して相同な、1つ以上の選択された抗原（ただし全ての抗原ではない）を含むワクチン組成物を意味する。このような組成物は、インタクトな病原性細胞もしくは病原性粒子、またはこのような細胞もしくは粒子の溶解物を実質的に含まない。従って、「サブユニットワクチン」は、この病原体由来

の、少なくとも部分的に精製された（好ましくは実質的に精製された）免疫原性ポリペプチド、またはそのアナログから調製され得る。従って、このサブユニットワクチンに含まれる抗原を得る方法は、標準的な精製技術、組み換え生成、または合成生成を含み得る。

【0049】

「実質的に精製された」とは、一般に、物質（化合物、ポリヌクレオチド、タンパク質、ポリペプチド、ポリペプチド組成物）が、サンプル（この物質が存在する）の大部分のパーセントを含むようなこの物質の単離をいう。代表的に、実質的に精製された成分は、サンプル中で、このサンプルの50%、好ましくは80～85%、より好ましくは90～95%を含む。目的のポリヌクレオチドおよびポリペプチドの精製のための技術は、当該分野で周知であり、そして例えば、イオン交換クロマトグラフィー、アフィニティークロマトグラフィー、および密度による沈降を含む。 10

【0050】

「コード配列」または選択されたポリペプチドを「コードする」配列とは、適切な調節配列（すなわち、「制御エレメント」）の制御下に配置された場合、インビボで転写され（DNAの場合）、そしてポリペプチドに翻訳される（mRNAの場合）、核酸分子である。コード配列の境界は、5'（アミノ）末端の開始コドン、および3'（カルボキシ）末端の翻訳終止コドンによって決定される。コード配列としては、ウイルス、原核生物のmRNAまたは真核生物のmRNA由来のcDNA、ウイルスDNAまたは原核生物DNA由来のゲノムDNA配列、および合成DNA配列さえ挙げられるがこれらに限定されない。終止コドンのような転写終止配列は、コード配列の3'に配置され得る。 20

【0051】

代表的な「制御エレメント」としては、転写プロモーター、転写エンハンサーエレメント、転写終止シグナル、ポリアデニル化配列（翻訳終止コドンの3'側に位置する）、翻訳の開始の至適化のための配列（コード配列の5'側に位置する）、および翻訳終止配列が挙げられるがこれらに限定されない。

【0052】

「ポリヌクレオチドコード配列」、または選択されたポリペプチドを「コードする」配列とは、適切な調節配列（すなわち、「制御エレメント」）の制御下に配置された場合、インビボで転写され（DNAの場合）、そしてポリペプチドに翻訳される（mRNAの場合）、核酸分子である。コード配列の境界は、5'（アミノ）末端の開始コドン、および3'（カルボキシ）末端の翻訳終止コドンによって決定される。例示的なコード配列は、本発明の改変されたウイルスポリペプチドコード配列である。転写終止配列は、コード配列の3'に配置され得る。代表的な「制御エレメント」としては、転写調節因子（例えば、プロモーター、転写エンハンサーエレメント、転写終止シグナル、およびポリアデニル化配列）；および翻訳の開始の至適化のための配列のような翻訳調節因子（例えば、シャインダーガーノ（リボソーム結合部位）配列、Kozak配列（すなわち、例えば、コード配列の5'側に位置する、翻訳の至適化のための配列）、リーダー配列、翻訳開始コドン（例えば、ATG）、ならびに翻訳終止配列が挙げられるがこれらに限定されない。特定の実施形態において、1つ以上の翻訳調節配列または開始配列（例えば、リーダー配列）が、野生型の翻訳開始配列（すなわち、ネイティブ状態でコード領域の翻訳を調節する配列）に由来する。本明細書において記載される方法を用いて改変された野生型のリーダー配列はまた、本発明における用途を見出す。プロモーターとしては、誘導性プロモーター（このプロモーターに作動可能に連結されたポリヌクレオチド配列の発現は、分析物、補因子、調節性タンパク質などによって誘導される）、抑制性プロモーター（このプロモーターに作動可能に連結されたポリヌクレオチド配列の発現は、分析物、補因子、調節性タンパク質などによって抑制される）、および構成的プロモーターが挙げられる。 30 40

【0053】

「核酸」分子としては、原核生物配列、真核生物mRNA、真核生物mRNA由来のcDNA、真核生物（例えば、哺乳動物）DNA由来のゲノムDNA配列、および合成DNA配列でさえ挙げられるがこれらに限定されない。この用語はまた、DNAおよびRNAの 50

任意の既知の塩基性アナログを含む配列を包含する。

【0054】

「作動可能に連結した（された）」とは、このように記載される成分が、それらの通常の機能を実行するように構成されるエレメントの配置をいう。従って、コード配列に作動可能に連結される所定のプロモーターは、適切な酵素が存在する場合、コード配列の発現を行い得る。プロモーターは、その発現を指向するように機能する限り、コード配列と連続である必要はない。従って、例えば、翻訳されないが転写される、介在する配列は、プロモーター配列とコード配列との間に存在し得、そしてプロモーター配列は、なお、コード配列に「作動可能に連結される」と考えられ得る。

【0055】

「組換え体」は、核酸分子を記載するために本明細書中で使用される場合、その起源または操作によって：（１）天然には関連していないポリヌクレオチドの全部または一部と関連していないか、および／または、（２）現実には連結されているポリヌクレオチド以外のポリヌクレオチドに連結している、ゲノム、cDNA、半合成、または合成起源のポリヌクレオチドを意味する。用語、「組換え体」は、タンパク質またはポリペプチドに関して使用される場合、組換えポリヌクレオチドの発現によって産生されるポリペプチドを意味する。「組み換え宿主細胞」、「宿主細胞」、「細胞」、「細胞株」、「細胞培養物」、および単細胞物として培養される原核生物微生物または真核生物細胞株を意味する他のこのような用語は、互換可能に用いられ、そして組み換えベクターまたは他の転移RNAのためのレシピエントとして用いられ得るか、またはレシピエントとして用いられた細胞をいい、そしてトランスフェクトされたもとの細胞の子孫を含む。単一の親細胞の子孫は、偶発的な変異または意図的な変異に起因して、形態もしくはゲノムにおいて、またはもとの親に相補的な総DNAにおいて、完全に同一である必要はないことが理解される。関連の特性（例えば、所望のペプチドをコードするヌクレオチド配列の存在）によって特徴付けられるように親に対して十分類似している親細胞の子孫は、この定義によって意図される子孫に含まれ、そして上記の用語によって包含される。

【0056】

アミノ酸配列「類似性」を決定する技術は、当該分野で周知である。一般に、「類似性」とは、2つ以上のポリペプチドの、適切な部位で比較される、アミノ酸とアミノ酸の的確さを意味し、ここで、アミノ酸は、同一であるか、または類似した化学的特性および／もしくは物理的特性（例えば、電荷または疎水性）を有する。次いで、いわゆる、用語「類似パーセント」とは、比較するポリペプチド配列間で決定され得る。核酸配列およびアミノ酸配列の同一性を決定する技術もまた、当該分野で周知であって、遺伝子に対するmRNAのヌクレオチド配列を決定する工程（通常、cDNA中間物を介する）およびこれらにコードされるアミノ酸配列を決定する工程、ならびにこれを第2のアミノ酸配列と比較する工程を含む。一般に、「同一性」とは、それぞれ2つのポリヌクレオチド配列またはポリペプチド配列に対応するヌクレオチドとヌクレオチドもしくはアミノ酸とアミノ酸の的確さをいう。

【0057】

2つ以上のポリヌクレオチド配列を、これらの「同一性パーセント」を決定することによって比較し得る。同様に、2つ以上のアミノ酸配列を、「同一性パーセント」を決定することによって比較し得る。核酸配列またはペプチド配列のいずれかに関わらず2つの配列の同一性パーセントは一般に、2つの整列した配列間での的確に一致する数を、短い方の配列の長さで除算して、100を掛けた数字によって、記載される。核酸配列に対する適切な整列は、SmithおよびWaterman、Advances in Applied Mathematics 2:482-489 (1981)の局所的相同性アルゴリズムにより、規定される。このアルゴリズムを拡張して、Dayhoff、Atlas of Protein Sequences and Structure (M. O. Dayhoff編、第5補遺、3:353-358、National Biomedical Research Foundation、Washington, D. C.、U

10

20

30

40

50

SA)により開発されたスコア付けマトリックスを用いて、ペプチド配列に使用し得、Grubbskov、Nucleic Acids Res. 14 (6): 6745-6763 (1986)により基準化され得る。このアルゴリズムを核酸配列およびペプチド配列に対して実行することは、Genetics Computer Group (Madison, WI)により、これらのBestFit utilityの適用において、規定される。この方法についての初期パラメータは、Wisconsin Sequence Analysis Package Program Manual, Version 8 (1995) (Genetics Computer Group, Madison, WIから入手可能)に記載される。他の等しく適切な、配列間の同一性パーセントおよび類似パーセントを計算するプログラムは、一般に当該分野で公知である。

10

【0058】

例えば、特定のヌクレオチド配列の参照配列に対する同一性パーセントは、SmithおよびWatermanの相同性アルゴリズムを、6つのヌクレオチド位置の初期スコア付け表およびギャップペナルティーと共に用いて決定し得る。本発明の脈絡における、同一性パーセントを確立する別の方法は、MPSEARCHパッケージのプログラム (University of Edinburghに著作権があり、John F. CollinsおよびShane S. Sturrockにより開発され、そしてIntelliGenetics, Inc. (Mountain View, CA)により配布される)を用いることである。このパッケージソフトから、Smith-Watermanアルゴリズムを利用し得、ここで初期パラメータをスコア付け表に使用する (例えば、12のギャップオープンペナルティー、1のギャップエクステンションペナルティーおよび6のギャップ)。生じたデータから、「一致」した値は、「配列同一性」を示す。配列間の同一性パーセントまたは類似性パーセントを計算する他の適切なプログラムは、一般に当該分野で公知であり、例えば、整列プログラムBLASTもまた初期パラメータで使用し得る。例えば、BLASTNおよびBLASTPは、以下の初期パラメータで使用され得る：genetic code=standard; filter=none; strand=both; cutoff=60; expect=10; Matrix=BLOSUM62; Descriptions=50配列; sort by=HIGH SCORE; Databases=non-redundant, GenBank+EMBL+DDBJ+PDB+GenBank CDS translations+Swiss protein+Spupdate+PIR。これらのプログラムの詳細は、以下のインターネットアドレスで見出され得る：<http://www.ncbi.nlm.gov/cgi-bin/BLAST>。

20

30

【0059】

当業者は、適切な検索パラメータ (典型的には、好ましいSmith Watermanに基づいたパラメータが上で規定される)を容易に決定し得て、所定の配列に対して使用し得る。例えば、検索パラメータは、該当する配列の大きさに基づいて変化し得る。従って、本発明のポリヌクレオチド配列に対して、本明細書で開示されるポリヌクレオチド配列の長さは、選択されるデータベースに対して検索され、そして基本的に同じ長さの配列に対して比較して、同一性パーセントを決定する。例えば、本発明の代表的な実施形態としては、X個連続するヌクレオチドを有する単離されたポリヌクレオチドが挙げられ、ここで(i) X個連続するヌクレオチドは、少なくとも、およそその選択されたレベルの同一性パーセントを、本明細書に記載される、Y個連続するヌクレオチドの配列に関して有し、そして(ii) 検索目的のために、X=Yであり、ここでYは、選択された参照ポリヌクレオチドの規定された長さである。

40

【0060】

本発明の配列としては、配列のフラグメント (例えば、約15ヌクレオチドから本明細書に記載の全長配列に存在するヌクレオチド数まで (上記の範囲の全ての整数値を含む) (例えば、配列表、図および特許請求の範囲を参照のこと))が挙げられ得る。例えば、本発明のポリヌクレオチド配列のフラグメントは、30~60ヌクレオチド、60~120

50

ヌクレオチド、120～240ヌクレオチド、240～480ヌクレオチド、480～1000ヌクレオチドおよびこれらの間の全ての整数値のヌクレオチドであり得る。

【0061】

本発明の配列が、例えば配列データベースに対して問い合わせ配列として使用される場合、本発明の合成発現カセット（および精製されたポリヌクレオチド）としては、約80%～100%の配列同一性、80～85%を超える配列同一性、好ましくは90～92%を超える配列同一性、より好ましくは95%を超える配列同一性、そして最も好ましくは98%を超え100%まで（これらの記載される範囲にある全ての整数値を含む）の配列同一性を、本明細書で開示される、合成発現カセット（および精製されたポリヌクレオチド）配列（例えば、特許請求の範囲の配列または本発明の他の配列）に対して有する、関連ポリヌクレオチド配列が挙げられる。 10

【0062】

2つの核酸フラグメントは、本明細書で記載される場合、「選択的にハイブリダイズ」するとみなされる。2つの核酸配列分子間の配列同一性の程度は、これらの分子間のハイブリダイゼーション事象の効率および強度に影響する。部分的に同一の核酸配列は、完全に同一の配列が標的分子にハイブリダイズするのを、少なくとも部分的に阻害する。完全に同一の配列のハイブリダイゼーションを阻害することは、当該分野で周知のハイブリダイゼーションアッセイ（例えば、サザンブロット、ノザンブロット、溶液ハイブリダイゼーションなど（Sambrookら、前出またはAusubelら、前出を参照のこと））を用いて評価し得る。このようなアッセイは、選択性の程度を変えること（例えば、低ストリンジェンシーから高ストリンジェンシーに変化する条件の利用）を利用して実施され得る。低ストリンジェンシーの条件を用いる場合、非特異的結合が存在しないことを、配列同一性を不完全な程度でさえも持っていない2次プローブ（例えば、標的分子に約30%未満の配列同一性を有するプローブ）を用いて評価し得、その結果、非特異的な結合事象が存在しない場合、この2次プローブは、標的にハイブリダイズしない。 20

【0063】

ハイブリダイゼーションに基づいた検出システムを利用する場合、標的核酸配列に相補的である核酸プローブが選択され、次いで、適切な条件を選択することによって、プローブと標的配列が、各々に「選択的にハイブリダイズ」、すなわち結合して、ハイブリッド分子を形成する。標的配列に、「中程度のストリンジェント」の下で選択的にハイブリダイズし得る核酸分子は、典型的に、少なくとも約70%の配列同一性を、選択された核酸プローブの配列に有する、少なくとも約10～14個のヌクレオチド長の標的核酸配列の検出が可能な条件下で、ハイブリダイズする。ストリンジェントなハイブリダイゼーション条件は、典型的に、約90～95%を超える配列同一性を、選択される核酸プローブの配列に有する、少なくとも約10～14個のヌクレオチド長の標的核酸配列の検出を可能にする。プローブ／標的のハイブリダイゼーションに有用なハイブリダイゼーション条件（プローブおよび標的は、特定の程度の配列同一性を有する）は、当該分野で公知のように決定され得る（例えば、Nucleic Acid Hybridization: A Practical Approach、B. D. HamesおよびS. J. Higgins編（1985）Oxford; Washington, DC; IRL Pressを参照のこと）。 30 40

【0064】

ハイブリダイゼーションについてのストリンジェンシー条件に関して、多くの等価な条件を用い得て、特定のストリンジェンシーを、例えば、以下の因子を変えることで確立し得ることは、当該分野で周知である：プローブおよび標的配列の長さおよび性質、種々の配列の塩基組成、塩および他のハイブリダイゼーション溶液組成物の濃度、ハイブリダイゼーション溶液中にブロック因子（例えば、ホルムアミド、硫酸デキストランおよびポリエチレングリコール）が存在するかしないか、ハイブリダイゼーション反応温度および時間のパラメータ、ならびに洗浄条件の変更。特定のセットのハイブリダイゼーション条件の選択は、当該分野の標準的方法に従って選択される（例えば、Sambrookら、 50

前出、または Ausubel ら、前出を参照のこと）。

【0065】

第1のポリヌクレオチドが、第2のポリヌクレオチド、そのcDNA、これらの相補体領域と同じかもしくは実質的に同じ塩基対配列を、有する場合または第1のポリヌクレオチドが、上記のような配列同一性を示す場合に、第1のポリヌクレオチドは、第2のポリヌクレオチドに「由来」する。

【0066】

第1のポリペプチドが、(i)第2のポリヌクレオチドに由来する第1のポリヌクレオチドによってコードされる場合、または(ii)第1のポリペプチドが、上記のような配列同一性を第2のポリペプチドに対して示す場合に、第1のポリペプチドは、第2のポリペ

10

【0067】

ウイルスポリペプチドが、(i)ウイルスのポリヌクレオチド(ウイルス性ポリヌクレオチド)のオープンリーディングフレームによってコードされる場合、または(ii)ウイルスポリペプチドが、上記のような配列同一性をウイルスのポリペプチドに対し示す場合に、一般に、ウイルスポリペプチドは、ウイルスの特定のポリペプチド(ウイルス性ポリペプチド)に「由来」する。

【0068】

「コードされる」とは、ポリペプチド配列をコードする核酸配列に関連し、ここでこのポリペプチド配列またはその一部は、この核酸配列によりコードされるポリペプチドに由来する、少なくとも3〜5個のアミノ酸、より好ましくは少なくとも8〜10個のアミノ酸、そしてさらにより好ましくは少なくとも15〜20個のアミノ酸からなる配列を含む。免疫学的にこの配列によってコードされるポリペプチドと同一視され得るポリペプチド配列もまた含まれる。さらに、ポリタンパク質は、ポリペプチドまたはペプチド産物をコードする、2つ以上のポリヌクレオチド配列をインフレーム融合することによって構築され得る。さらに、ポリシストロン(polycistronic)配列コード配列は、互いに隣接するポリペプチド産物をコードする2つ以上のポリヌクレオチド配列を、典型的には1つのプロモーターの制御下に、配置することで産生され得、ここで各ポリペプチドコード配列は、内部リボソーム結合部位に対する配列を含むよう改変され得る。

20

【0069】

「精製されたポリヌクレオチド」とは、基本的に、目的のポリヌクレオチドと自然の状態に関連するタンパク質が除かれた(例えば、そのタンパク質の約50%未満、好ましくは約70%未満、そしてより好ましくは約90%未満を含む)、目的のポリヌクレオチドまたはそのフラグメントに関する。目的のポリヌクレオチドを精製する技術は、当該分野で周知であり、例えば、ポリヌクレオチドを含む細胞を、カオトロピック剤で破壊し、そしてポリヌクレオチドおよびタンパク質を、イオン交換クロマトグラフィー、アフィニティークロマトグラフィーおよび密度に依存した沈殿法により分離する工程を含む。

30

【0070】

「核酸免疫化」とは、1つ以上の選択された抗原をコードする核酸分子を、宿主細胞に、抗原またはエピトープのインビボ発現のため導入することを意味する。核酸分子は、レシ

40

【0071】

「遺伝子移入」または「遺伝子送達」とは、目的のDNAを宿主細胞に正確に挿入する方法またはシステムをいう。このような方法は、組みこまれていない移入DNAの一過性の発現、染色体外複製および移入レプリコン(例えば、エピソーム)の発現、または移入遺伝子物質の宿主細胞ゲノムDNA中への組みこみを生じ得る。遺伝子送達発現ベクターとしては、アルファウイルス、ボックスウイルスおよびワクシニアウイルス由来のベクター

50

が挙げられるが、これらに限定されない。免疫化に用いる場合、このような遺伝子送達発現ベクターは、ワクチンまたはワクチンベクターと呼ばれ得る。

【0072】

「Tリンパ球」または「T細胞」は、抗体を産生しないリンパ球であり、免疫系の細胞性の備えの一部を構成する。T細胞は、骨髄から胸腺に遊走した未成熟なリンパ球から生じ、これらは、胸腺ホルモンの指令のもとで成熟過程を経ていく。ここで成熟リンパ球は、急速に分裂して、非常に多数に増加する。成熟T細胞は、特定の抗原を認識および結合する能力に基づいて、免疫応答性になる。抗原がリンパ球表面レセプター結合した場合、免疫応答性T細胞の活性化は、誘発される。

【0073】

用語「トランスフェクション」は、外来DNAの、細胞による取り込みを示すために使用される。外因性DNAが細胞膜内に導入された場合、細胞は、「トランスフェクト」されている。多くのトランスフェクション技術は、当該分野で一般に公知である。例えば、Grahamら、(1973) *Virology*, 52:456, Sambrookら、(1989) *Molecular Cloning, a laboratory manual*, Cold Spring Harbor Laboratories, New York, Davisら、(1986) *Basic Methods in Molecular Biology*, Elsevier、およびChuら、(1981) *Gene* 13:197を参照のこと。このような技術を用いて、1つ以上の外来DNA部分を、適切な宿主細胞中に導入し得る。この用語は、遺伝子物質の安定な取り込みおよび一過性の取り込みの両方を意味し、そしてペプチド結合DNAまたは抗体結合DNAの取り込みを含む。

【0074】

「ベクター」は、遺伝子配列を、標的細胞へ移入し得る（例えば、ウイルスベクター、非ウイルスベクター、粒子性キャリアおよびリポソーム）。典型的には、「ベクター構築物」、「発現ベクター」および「遺伝子移入ベクター」は、目的の遺伝子発現を指向し得る任意の核酸構築物をいい、これらは、遺伝子配列を、標的細胞に移入し得る。従って、この用語は、クローニングおよび発現ビヒクル、ならびにウイルスベクターを含む。

【0075】

「自殺遺伝子」（例えば、薬物感受性遺伝子）の標的細胞への移入は、正常細胞に対して比較的無毒である化合物または組成物に対して、細胞を感受性にする。Moolten, F. L. (1994) *Cancer Gene Ther.* 1:279-287. 自殺遺伝子の例は、単純疱疹ウイルスのチミジンキナーゼ (HSV-tk)、シトクロムP450 (Manomeら、(1996) *Gene Therapy* 3:513-520)、ヒトデオキシシチジンキナーゼ (Manomeら、(1996) *Nature Medicine* 2(5):567-573) および細菌性酵素シトシンデアミナーゼ (Dongraら、(1996) *Human Gene Therapy* 7:713-720) である。これらの遺伝子を発現する細胞は、比較的無毒のプロドラッグであるガンシクロビル (HSV-tk)、シクロホスファミド (シトクロムP450 2B1)、シトシンアラビノシド (ヒトデオキシシチジンキナーゼ) または5-フルオロシトシン (細菌性シトシンデアミナーゼ) の影響に対し、感受性にされる。Culverら、(1992) *Science* 256:1550-1552, Huberら、(1994) *Proc. Natl. Acad. Sci. USA* 91:8302-8306。

【0076】

「選択マーカー」または「レポーターマーカー」とは、遺伝子移入ベクターに含まれるヌクレオチド配列をいい、これらは治療的活性を持たないが、むしろ遺伝子移入ベクターのより単純な調製、製造、特徴づけまたは試験を可能にするために含まれる。

【0077】

「特異的結合因子」とは、分子の特定の結合対のメンバーをいい、ここでこれらの分子の1つは、第2の分子に対し、化学的および／または物理的手段を介して、特異的に結合す

10

20

30

40

50

る。特異的結合因子の1つの例は、選択された抗原に対して指向する抗体である。

【0078】

「被験体」とは、脊椎動物の亜門のいずれかのメンバー（ヒトおよび他の霊長類（例えば、チンパンジーおよび他の類人猿、ならびに猿種などのヒト以外の霊長類を含む）；家畜（例えば、ウシ、ヒツジ、ブタ、ヤギおよびウマ）；家庭向けの哺乳動物（例えば、イヌおよびネコ）；実験動物（例えば、マウス、ラットおよびモルモットなどのげっ歯類を含む）；鳥（家庭向けの鳥、野生の鳥および狩猟鳥を含む（例えば、ニワトリ、シチメンチョウおよび他の家禽の鳥、アヒル、ダチョウなど））を含むが、これらに限定されない）を意味する。この用語は、特定の年齢を示さない。従って、成体および新生仔の個体が意図され、含まれる。これらの脊椎動物の全ての免疫系は、同様に機能するので、上記のシステムは、上記脊椎動物種のいずれかにおける使用が意図される。

10

【0079】

「薬学的に受容可能」または「薬理学的に受容可能」は、生物学的でない物質またはさもなくば所望されない物質に関する。すなわち、この物質は、処方物または組成物で個体に投与され得て、任意の所望されない生物学的影響を引き起こすことも、有害な様式で、含まれる組成物成分のいずれかと相互作用することない。

【0080】

「生理学的pH」または「生理学的範囲にあるpH」とは、約7.2～8.0に含まれる範囲にあるpHを意味し、より典型的には、約7.2～7.6に含まれる範囲にある。

【0081】

本明細書で使用される場合、「処置」とは、以下のいずれかをいう：（I）従来のワクチンにおけるような、感染または再感染の予防、（ii）症状の軽減または除去、および（iii）該当する病原体の実質的除去または完全な除去。処置は、予防的に（感染前に）または治療的に（感染後に）実施され得る。

20

【0082】

「同時投与」とは、2つ以上の組成物または分子を投与することをいう。従って、同時投与は、同じ投与経路または異なる投与経路を介して、同時に投与する工程または連続して投与する（任意の順番で）工程を含む。同時投与レジメンの限定しない例としては、核酸およびペプチドの同時投与；異なる核酸の同時投与（例えば、本明細書に記載されるような異なる発現カセットおよび／または異なる遺伝子送達ベクター）；および異なるポリペプチドの同時投与（例えば、異なるHIVポリペプチドおよび／または異なるアジュバント）が挙げられる。この用語はまた、同時投与される分子または組成物のうちいずれか1つを複数回投与する工程を含む（例えば、本明細書に記載される1つ以上の発現カセットの複数回投与後の、ポリペプチド含有組成物の1回以上の投与）。分子または組成物が連続して送達される場合に、各投与間の時間は、本明細書の教示を考慮することで、当業者により容易に決定され得る。

30

【0083】

「レンチウイルスベクター」および「組換えレンチウイルスベクター」とは、目的の核酸分子を運ぶ核酸構築物をいい、そして特定の実施形態において、目的の核酸分子の発現を指向し得る。レンチウイルスベクターは、少なくとも1つの転写プロモーター／エンハンサーもしくは遺伝子座規定エレメント、または遺伝子発現を他の手段（例えば、選択的スプライシング、核RNA輸出、メッセンジャーの翻訳後修飾またはタンパク質の転写後修飾）により制御する他のエレメントを含む。このようなベクター構築物はまた、パッケージングシグナル、長い末端反復配列（LTRS）またはこれらの一部、ならびに使用するレトロウイルスに適した正および負のプライマー鎖結合部位を含まなければならない（これらがまだ、レトロウイルスベクターに存在していない場合）。必要に応じて、組換えレンチウイルスベクターはまた、ポリアデニル化を指向するシグナル、選択マーカー（例えば、Neo、TK、ハイグロマイシン、プレオマイシン、ヒスチジノールまたはDHFR）、ならびに1つ以上の制限酵素認識部位および翻訳終結配列を含み得る。例として、このようなベクターは、典型的に、5'LTR、tRNA結合部位、パッケージングシグナ

40

50

ル、第2鎖DNA合成の起点および3'LTRまたはこれらの一部を含む。

【0084】

「レンチウイルスベクター粒子」とは、本発明の範囲内で使用される場合、少なくとも1つの目的の遺伝子を保持するレンチウイルスをいう。このレトロウイルスはまた、選択マーカーを含み得る。この組換えレンチウイルスは、その遺伝物質(RNA)をDNAに逆転写し得、そして感染の際にこの遺伝物質を宿主細胞のDNAに組み込み得る。レンチウイルスベクター粒子は、レンチウイルスエンベロープ、非レンチウイルスエンベロープ(例えば、両性(ampho)エンベロープまたはVSV-Gエンベロープ)またはキメラエンベロープを有し得る。

【0085】

「核酸発現ベクター」または「発現カセット」とは、目的の配列または遺伝子の発現を指向し得る構築物をいう。核酸発現ベクターは、目的の配列または遺伝子に作動可能に連結されたプロモーターを含む。他の制御エレメントも同様に存在し得る。本明細書中に記載される発現カセットは、プラスミド構築物内に含まれ得る。この発現カセットの成分に加えて、このプラスミド構築物はまた、細菌複製起点、1以上の選択マーカー、プラスミド構築物が一本鎖DNAとして存在することを可能にするシグナル(例えば、M13複製起点)、マルチクロニングサイトおよび「哺乳動物」複製起点(例えば、SV40またはアデノウイルスの複製起点)を含み得る。

【0086】

「パッケージング細胞」とは、組換えレトロウイルスベクターを欠く感染性組換えレトロウイルスの産生のために必要なエレメントを含む細胞をいう。代表的には、このようなパッケージング細胞は、Gag、polおよびenvタンパク質をコードする、タンパク質を発現し得る1以上の発現カセットを含む。

【0087】

「プロデューサー細胞」または「ベクター産生細胞」とは、組換えレトロウイルスベクター粒子の産生のために必要な全てのエレメントを含む細胞をいう。

【0088】

(2. 本発明を実行する様式)

本発明を詳細に記載する前に、本発明は、特定の処方またはプロセスパラメーターに限定されず、このようなものは、当然異なり得ることが理解されるべきである。本明細書中で使用される専門用語は、本発明の特定の実施形態のみを記載する目的のためであり、限定することを意図しないこともまた、理解されるべきである。

【0089】

本明細書中に記載の方法および材料に類似するかまたは等価である多くの方法および材料が、本発明の実行において使用され得るが、好ましい材料および方法が、本明細書中に記載される。

【0090】

(2. 1. HIVゲノム)

HIVゲノムおよび種々のポリペプチドコード領域を、表Aに示す。ヌクレオチド位置は、8__5__TV1__C.ZA(配列番号33、図11)に対して与えられる。しかし、本発明の開示の技術の観点から、例えば、配列比較プログラム(例えば、本明細書中に記載されるBLASTなど)または構造的特徴の同定およびアラインメント(例えば、種々の領域を同定し得る、本明細書中に記載される「ALB」プログラムのようなプログラム)を使用して、他のHIV株または改変体(例えば、単離株HIV_{IIIB}、HIV_{SF2}、HIV-1_{SF162}、HIV-1_{SF170}、HIV_{LAV}、HIV_{LAI}、HIV_{MN}、HIV-1_{CM235}、HIV-1_{US4})、種々のサブタイプ由来の他のHIV-1株(例えば、サブタイプA~G、およびO)、HIV-2株および種々のサブタイプ(例えば、HIV-2_{UC1}およびHIV-2_{UC2})、ならびにサル免疫不全ウイルス(SIV)(これらおよび他の関連のウイルスの記載について、例えば、Virology、第三版(W. K. Joklik編、1988); Fundamental Vir

10

20

30

40

50

ology、第二版 (B. N. Fields および D. M. Knipe 編、1991) ;
Virology、第三版 (Fields, BN, DM Knipe, PM Howley、
編者、1996、Lippincott-Raven、Philadelphia, PA) を参照のこと) 中の対応する領域を決定する方法は、当業者に容易に明らかである。

【0091】

【表1】

表A: 8_5_TV1_C.ZA に対する HIV ゲノムの領域

領域	ヌクレオチド配列中の位置
5'LTR	1-636
U3	1-457
R	458-553
U5	554-636
NFκB II	340-348
NFκB I	354-362
Sp1 III	379-388
Sp1 II	390-398
Sp1 I	400-410
TATA ボックス	429-433
TAR	474-499
ポリ A シグナル	529-534
PBS	638-655
p7 結合領域、パッケージングシグナル	685-791
Gag:	792-2285
p17	792-1178
p24	1179-1871
サイクロフィリンA 結合	1395-1505
MHR	1632-1694
p2	1872-1907
p7	1908-2072
フレームシフトスリップ	2072-2078
p1	2073-2120
p6Gag	2121-2285
Zn-モチーフ I	1950-1991
Zn-モチーフ II	2013-2054

10

20

30

40

(表A 続き)

Pol:	2072-5086	
p6Pol	2072-2245	
Prot	2246-2542	
p66RT	2543-4210	
p15RNaseH	3857-4210	
p31Int	4211-5086	
Vif:	5034-5612	
親水性領域	5292-5315	10
Vpr:	5552-5839	
オリゴマー化	5552-5677	
両親媒性 α -ヘリックス	5597-5653	
Tat:	5823-6038 および 8417-8509	
Tat-1 エキソン	5823-6038	
Tat-2 エキソン	8417-8509	
N-末端ドメイン	5823-5885	20
トランス活性化ドメイン	5886-5933	
形質導入ドメイン	5961-5993	
Rev:	5962-6037 および 8416-8663	
Rev-1 エキソン	5962-6037	
Rev-2 エキソン	8416-8663	
高親和性結合部位	8439-8486	
Leu-リッチエフェクタードメイン	8562-8588	
Vpu:	6060-6326	30
膜貫通ドメイン	6060-6161	
細胞質ドメイン	6162-6326	

(表A 続き)

Env (gp160):	6244-8853
シグナルペプチド	6244-6324
gp120	6325-7794
V1	6628-6729
V2	6727-6852
V3	7150-7254
V4	7411-7506
V5	7663-7674
C1	6325-6627
C2	6853-7149
C3	7255-7410
C4	7507-7662
C5	7675-7794
CD4 結合	7540-7566
gp41	7795-8853
融合ペプチド	7789-7842
オリゴマー化ドメイン	7924-7959
N-末端7反復	7921-8028
C-末端7反復	8173-8280
免疫優性領域	8023-8076
Nef:	8855-9478
ミリストイル化	8858-8875
SH3 結合	9062-9091
ポリプリン区域	9128-9154
SH3 結合	9296-9307

10

20

当業者が、表Aに示される任意の配列を整列させて、任意の特定のH I V遺伝子の相対的な位置を容易に決定し得ることは、容易に明らかである。例えば、本明細書中に記載されるアラインメントプログラムの1つ（例えば、B L A S T）を使用して、他のC型H I Vの配列を、8 _ 5 _ T V 1 _ C . Z Aと共に整列させて（表A）、遺伝子の位置を決定し得る。

30

【0092】

ポリペプチド配列も同様に整列し得る。例えば、図103は、種々の株由来のE n vポリペプチド配列の、S F - 1 6 2に対するアラインメントを示す。共有に係るW O / 3 9 3 0 3に詳細に記載されるように、E n vポリペプチド（例えば、g p 1 2 0、g p 1 4 0およびg p 1 6 0）は、βシートを形成する4つのアンチパラレルβ鎖（β - 2、β - 3、β - 2 0およびβ 2 1）からなる「架橋（b r i d g i n g）シート」を含む。第一の対のβ鎖（β - 2およびβ - 3）からの突出は、2つのループ（V 1およびV 2）である。S F - 1 6 2に対して、このβ - 2シートは、ほぼアミノ酸残基1 1 3（C y s）～アミノ酸残基1 1 7（T h r）を占めるが、β - 3は、ほぼアミノ酸残基1 9 2（S e r）～アミノ酸残基1 9 4（I l e）を占める（図103を参照のこと）。この「V 1 / V 2領域」は、S F - 1 6 2に対して、ほぼアミノ酸残基1 2 0（C y s）～アミノ酸残基1 8 9（C y s）を占める。β鎖の第二の対（β - 2 0およびβ - 2 1）からの突出は、「小ループ」構造であり、本明細書中で「架橋シート小ループ」とも称される。小ループおよび架橋シート小ループの両方の位置は、本明細書中およびW O / 3 9 3 0 3の技術に従って、H X B - 2に対して決定され得る。図103 A ~ Cにおいて、矢印によって、βシート領域からの欠失配列のおおよその部位もまた示される。「*」は、本明細書の技術に

40

50

従って変異され得るNグリコシル化部位を示す。

【0093】

(2.2 合成発現カセット)

(2.2.1 C型HIV-1の、Pol-、Prot-、Rt-、Int-、Gag、Env、Tat、Rev、Nef、RnaseH、Vif、VprおよびVpuの核酸コード配列の改変)

本発明の1つの局面は、対応する野生型配列に対して改善された発現を有する、C型HIV-1コード配列および関連配列の作製である。

【0094】

(2.2.1.1. Gagの核酸コード配列の改変)

HIV-1サブタイプCのAF110965株およびAF110967株から得られたGagタンパク質の野生型配列を改変することによる、本発明の例示的な実施形態が、本明細書中に示される(例えば、Korberら(1998) Human Retroviruses and Aids、Los Alamos、New Mexico: Los Alamos National Laboratory; Novitskyら、(1999) J. Virol. 73 (5): 4427-4432 (ボツワナ由来の種々のサブタイプCクローンの分子クローニングについて)を参照のこと)。新規単離株8_5_TV1_C. ZA (TV001またはTV1とも呼ばれる) および12_5_1_TV2_C. ZA (TV002またはTV2とも呼ばれる) に由来する野生型配列の改変もまた、本明細書中に示される。配列番号52は、8_5_TV1_C. ZA由来のGagの野生型配列を示し、そして配列番号54は、同じ株由来のGagの主要な相同領域(表Aのヌクレオチド1632~1694)の野生型配列を示す。配列番号100は、12_5_1_TV2_C. ZAのGagの野生型配列を示す。

10

20

【0095】

他のC型HIV-1改変体から得られたGag配列は、本明細書の技術に従って類似の様式で操作され得る。このような他の改変体としては、C型HIV-1の単離株から得られたGagタンパク質コード配列(例えば、Novitskyら(1999)、前出; Myersら、後出; Virology、第三版(W. K. Joklik編、1988); Fundamental Virology、第二版(B. N. FieldsおよびD. M. Knipe編、1991); Virology、第三版(Fields, BN、DM Knipe、PM Howley、編者、1996、Lippincott-Raven、Philadelphia, PA)ならびにワールドワイドウェブ(インターネット)(例えば、<http://hiv-web.lanl.gov/cgi-bin/hivDB3/public/wdb/ssampublic>および<http://hiv-web.lanl.gov>。)に記載される)が挙げられるが、これらに限定されない。

30

【0096】

最初に、HIV-1のコドン用法パターンを、得られた核酸コード配列が、高度に発現されるヒト遺伝子において見出されるコドン用法と匹敵するように、改変した(実施例1)。HIVコドン用法は、コドントリプレットのヌクレオチドAおよびTの高含量を反映する。HIV-1コドン用法の影響は、mRNAの減少した翻訳能力および不安定性を生じる、DNA配列中の高いAT含量である。対照的に、高度に発現されるヒトコドンは、ヌクレオチドGまたはCを好む。Gagコード配列を、高度に発現されるヒト遺伝子において見出されるコドン用法に匹敵するように改変した。

40

【0097】

第二に、Gagコード配列のコード配列内に位置する阻害(または不安定)エレメント(INS)が存在する。RREは、HIVがコードするRevタンパク質と相互作用して、INSの発現ダウンレギュレーション効果を克服する、二次RNA構造である。RREおよびRevの転写後活性化機構を克服するために、不安定エレメントは、コードされたタンパク質のリーディングフレームを変更しない複数の点変異を導入することによって、不活化され得る。不活化されたRRE部位を有する、サブタイプCのGagコード配列は、

50

例えば、図1（配列番号3）、図2（配列番号4）、図5（配列番号20）および図6（配列番号26）に示される。同様に、他のサブタイプC株に由来する他の合成ポリヌクレオチドは、RRE部位を不活化するために改変され得る。

【0098】

Gagポリペプチドコード配列の改変は、多くの哺乳動物細胞株（ならびに昆虫細胞を含むがこれに限定されない他の型の細胞株）において、野生型コード配列に対して改善された発現を生じる。さらに、これらの配列の発現は、これらの細胞株によるウイルス様粒子（VLP）の産生を生じる（以下を参照のこと）。

【0099】

（2.2.1.2 Env核酸コード配列の改変）

同様に、本発明はまた、合成Envコードポリヌクレオチドおよび改変されたEnvタンパク質を含む。野生型のEnv配列は、C型HIV-1のAF110968株およびAF110975株ならびに新規株8_5_TV1_C.ZA（配列番号33）および12_5_1_TV2_C.ZA（配列番号45）から得られる（例えば、ボツワナ由来の種々のサブタイプCクローンの分子コロニーングについて、Novitskyら（1999）J. Virol. 73（5）：4427-4432を参照のこと）。8_5_TV1_C.ZAの野生型Env配列は、例えば、配列番号48（野生型Env共通領域、表Aに示されるヌクレオチド7486～7629）；および配列番号50（野生型gp160、表Aに示されるヌクレオチド6244～8853）に示される。12_5_1_TV2_C.ZAの野生型Env gp160は、配列番号98に示される。本明細書中の開示から、Env gp160（例えば、gp120、gp41、gp140）のフラグメントをコードするポリヌクレオチドが、本明細書中に開示されるより大きい全長配列から容易に得られ得ることが、容易に明らかである。他の改変（例えば、V1および／またはV2領域のような領域の欠失；切断部位の変異など）がなされ得ることもまた、容易に明らかである（実施例1を参照のこと）。このような改変の例示的配列は、配列番号119～127に示される。

【0100】

さらに、C型HIV-1改変体から得られるEnv配列は、本明細書の技術に従って類似の様式で操作され得る。このような他の改変体としては、上記のC型HIV-1単離株から得られるEnvタンパク質コード配列が挙げられるが、これに限定されない。

【0101】

Envについてのコドン用法パターンを、Gagについて上記のように改変し、その結果、得られた核酸コード配列は、高度に発現されるヒト遺伝子において見出されるコドン用法に匹敵した。本発明を支持して、実行された実験は、合成Env配列がネイティブなEnv配列と比較して、高レベルのタンパク質を産生し得たことを示す。

【0102】

Envポリペプチドコード配列の改変は、多くの哺乳動物細胞株（ならびに昆虫細胞を含むがこれに限定されない他の型の細胞株）において、野生型コード配列に対して改善された発現を生じる。類似のEnvポリペプチドコード配列は、種々の単離株（Gagについて上記された株を含む）から得られ得、改変され得、そして改善された発現について試験され得る。

【0103】

Envのさらなる改変は、その中に変異および／または欠失を有するEnvポリペプチドをコードするポリヌクレオチドを作製することを含むが、これに限定されない。例えば、超可変領域（V1および／またはV2）は、本明細書中に記載のように欠失され得る。さらに、例えば、Env内の架橋シート領域および／またはN-グリコシル化部位に対する他の改変もまた、本明細書中の技術に従って実行され得る（図130A～CおよびWO/39303を参照のこと）。これらの改変の種々の組み合わせは、本明細書中で記載されるような合成発現カセットを作製するために使用され得る。

【0104】

10

20

30

40

50

(2. 2. 1. 3 HIV-1 Pol 核酸コード配列を含む配列の改変)

本発明はまた、合成 Pol 配列を含む発現カセットを包含する。上記のように、「Pol」は、図 7 に示されるようなタンパク質コード領域（例えば、ポリメラーゼ、プロテアーゼ、逆転写酵素および／またはインテグラーゼ含有配列）を含むが、これに限定されない。図 7 に示される領域は、例えば、Wanら（1996）*Biochem. J.* 316: 569-573; Kohlら（1988）*PNAS USA* 85: 4686-4690; Krausslichら（1988）*J. Virol.* 62: 4393-4397; Coffin「*Retroviridae and their Replication*」、*Virology*, 1437-1500 頁（Raven, New York, 1990）; Patelら（1995）*Biochemistry* 34: 5351-5363 に記載される。従って、本明細書中で例示される合成発現カセットは、1 以上のこれらの領域および得られるアミノ酸配列に対する 1 以上の変化を含む。

10

【0105】

野生型の Pol 配列を、C 型 HIV-1 の AF110975 株、8__5__TV1__C. ZA 株および 12-5__1__TV2__C. ZA 株から得た（例えば、ボツワナ由来の種々のサブタイプ C クローンの分子クローニングについて、Novitskyら（1999）*J. Virol.* 73 (5): 4427-4432 を参照のこと）。配列番号 34 は、Pol の p2～p7 領域由来の AF110975 の野生型配列を示す（図 7 および表 A を参照のこと）。配列番号 35 は、p1～インテグラーゼの最初の 6 アミノ酸由来の AF110975 の野生型配列を示す（図 7 および表 A を参照のこと）。配列番号 63 および配列番号 104 は、それぞれ、8__5__TV1__C. ZA および 12-5__1__TV2__C. ZA 由来の Pol の野生型配列を示す（また表 A を参照のこと）。

20

【0106】

他の C 型 HIV-1 改変体から得られた配列は、本明細書の技術に従って類似の様式で操作され得る。このような他の改変体としては、本明細書中に記載される C 型 HIV-1 単離株から得られた Pol タンパク質コード配列が挙げられるが、これに限定されない。

【0107】

Pol についてのコドン用法パターンを、Gag および Env について上記されたように改変し、その結果、得られた核酸コード配列は、高度に発現されるヒト遺伝子において見出されるコドン用法と匹敵した。

30

【0108】

表 B は、本明細書中で例示される Pol 構築物（例えば、配列番号 30～32）において見出される種々の領域のヌクレオチド位置を示す。

【0109】

【表 2】

表 B

領域	構築物中のヌクレオチド配列における位置		
	PR975(+)	PR975YM	PR975(+) YMWM
	配列番号 30	配列番号 31	配列番号 32
Sal I 制限部位	1-6	1-6	1-6
Kozak 開始コドン	7-16	7-16	7-16
p2	16-54	16-54	16-54
p7	55-219	55-219	55-219
p1/p6 pol	220-375	220-375	220-375
インフレームについての挿入変異	225	225	225
p10プロテアーゼ	376-672	376-672	376-672
p66RT	673-2352	673-2346	673-2340
p51RT	673-1992	673-1986	673-1980
p15RNaseH	1993-2352	1993-2346	1993-2340
触媒中心領域 (YMDD)	1219-1230	1219-1224	1219-1224
プライマーグリップ領域 (WMGY)	1357-1368	1351-1362	1351-1356
6aa インテグラーゼ	2353-2370	2347-2364	2341-2358
YMDD エピトープカセット (5'+3'Glyを含む)	2371-2424	2365-2418	2359-2412
MCS (マルチクローニングサイト)	2425-2463	2419-2457	2413-2451
EcoR I 制限部位	2464-2469	2458-2463	2452-2457

表 B に示すように、例示的な構築物を、種々の様式で改変した。例えば、本明細書中で例示される発現構築物は、インテグラーゼペプチドの最初の 6 つのアミノ酸をコードする配列を含む。この 6 アミノ酸の領域は、H I V プロテアーゼによって認識される、切断認識部位を提供すると考えられる（例えば、McCornack ら（1997）FEBS Lett 414:84-88 を参照のこと）。上記のように、本明細書中に例示される特定の構築物は、1 つ以上の導入遺伝子の挿入のための、マルチクローニングサイト（MCS）を、代表的にはこの構築物の 3' 末端に含む。さらに、RT において触媒中心から誘導された触媒中心エピトープをコードするカセットは、代表的に、インテグラーゼの 6

【0110】

特定の実施形態において、RT の触媒中心および／またはプライマーグリップ領域が、改変される。RT の触媒中心およびプライマーグリップ領域は、例えば、Patel ら（1995）Biochem. 34:5351 および Palaniappan ら（1997）J. Biol. Chem. 272(17):11157 に記載されている。例えば、PR975YM（配列番号 31）と称される構築物において、p66RT の 183 位～185 位（AF110975 に対して番号付けられた）においてアミノ酸 YMDD をコードする野生型配列は、アミノ酸「AP」をコードする配列と置換される。PR975YMWM（配列番号 32）と称される構築物において、YMDD における同じ変位がなされ、そしてさらに、プライマーグリップ領域（アミノ酸 WMGY、p66RT の残基 229～232（AF110975 に対して番号付けられた））が、アミノ酸「PI」をコードする配列で置換される。

【0111】

Pol 配列に対して、コドン利用の変化は、代表的に、-1 までのフレームシフトの領域

に制限されており、そして再度、G a gリーディングフレームの末端において開始する；しかし、フレームシフト翻訳領域内の領域も同様に改変され得る。最後に、プロテアーゼポリペプチドコード配列のコード配列内に位置する阻害（すなわち、不安定）エレメント（I N S）も同様に、変更され得る。

【0112】

本発明の支持によって実験を実施して、合成P o l配列がネイティブP o l配列と比較してより高レベルのタンパク質を産生し得たことを示し得る。P o lポリペプチドコード配列の改変は、多くの哺乳動物細胞株（および他の型の細胞（昆虫細胞が挙げられるがこれらに限定されない））において、野生型コード配列と比較して、改善された発現を生じる。種々の単離物（G a gおよびE n vに関して上記したものを含む）からの改善された発現のために改変および試験された、類似のP o lポリペプチドコード配列が得られ得る。 10

【0113】

（2. 2. 1. 4 他のH I V配列の改変）

本発明はまた、G a g、E n vおよびP o l以外の、合成C型H I V配列由来のH I V遺伝子（G a g、E n v、P o l、ならびに例えば、8 _ 5 _ T V 1 _ C . Z A（配列番号33）または12 _ 5 _ 1 _ T V 2 _ C . Z A（配列番号45）由来のv i f、v p r、t a t、r e v、v p u、およびn e f内の領域が挙げられるが、これらに限定されない）を含む、発現カセットを包含する。他の株から得られた配列は、本明細書の教示に従って類似の様式で操作され得る。

【0114】

上記のように、コドン利用パターンは、G a g、E n vおよびP o lに関して上記したように改変され、その結果、得られる核酸コード配列は、高度に発現されるヒト遺伝子において見出されるコドン利用に匹敵する。本発明の支持によって実験を実施して、これらの合成配列が、ネイティブの配列と比較してより高レベルのタンパク質を産生し得たこと、および野生型ポリペプチドコード配列の改変が、多くの哺乳動物細胞株（および他の型の細胞（昆虫細胞が挙げられるがこれらに限定されない））において、野生型コード配列と比較して改善された発現を生じることを示し得る。さらに、核酸配列もまた改変されて、遺伝子の1つ以上の領域に変異を導入し、例えば、その遺伝子産物を非機能的にし得、そして／またはN e fにおけるミリストイル化部位を排除し得る。 20

【0115】

本明細書中で例示される合成発現カセットは、以下を含む：配列番号49および配列番号97（それぞれ8 _ 5 _ T V 1 _ C . Z A野生型および12 _ 5 _ 1 _ T V 2 _ C . Z A野生型に基づいて改変されたE n v g p 160コード配列）；配列番号51および配列番号99（それぞれ8 _ 5 _ T V 1 _ C . Z A野生型および12 _ 5 _ 1 _ T V 2 _ C . Z A野生型に基づいて改変されたG a gコード配列）；配列番号53（8 _ 5 _ T V 1 _ C . Z A野生型に基づいて改変されたG a gの主要な相同性領域）；配列番号55および配列番号101（それぞれ8 _ 5 _ T V 1 _ C . Z A野生型および12 _ 5 _ 1 _ T V 2 _ C . Z A野生型に基づいて改変されたN e fコード配列）；配列番号57および配列番号134（それぞれ8 _ 5 _ T V 1 _ C . Z A野生型および12 _ 5 _ 1 _ T V 2 _ C . Z Aに基づいて改変された、125位における変異を有して非機能的な遺伝子産物を生じるN e fコード配列）；配列番号58（8 _ 5 _ T V 1 _ C . Z A野生型に基づいて改変されたR N A s e Hコード配列）；配列番号60（8 _ 5 _ T V 1 _ C . Z A野生型に基づいて改変されたインテグラーゼコード配列）；配列番号62および配列番号103（それぞれ8 _ 5 _ T V 1 _ C . Z A野生型および12 _ 5 _ 1 _ T V 2 _ C . Z A野生型に基づいて改変されたP o lコード配列）；配列番号64（8 _ 5 _ T V 1 _ C . Z A野生型に基づいて改変されたプロテアーゼコード配列）；配列番号66（8 _ 5 _ T V 1 _ C . Z A野生型に基づいて改変された不活化プロテアーゼコード配列）；配列番号68（8 _ 5 _ T V 1 _ C . Z A野生型に基づいて改変された不活化プロテアーゼおよびR T変異配列）；配列番号70（8 _ 5 _ T V 1 _ C . Z A野生型に基づいて改変されたプロテアーゼおよび逆転写酵素のコード配列）；配列番号72および配列番号105（それぞれ8 40 50

5 TV 1 C. Z A 野生型および 1 2 - 5 1 TV 2 C. Z A 野生型に基づいて
 改変された、Rev のエキソン 1) ; 配列番号 7 4 および配列番号 1 0 7 (それぞれ 8
 5 TV 1 C. Z A 野生型および 1 2 - 5 1 TV 2 C. Z A 野生型に基づいて改
 変された、Rev のエキソン 2) ; 配列番号 7 6 (8 5 TV 1 C. Z A 野生型に基
 づいて改変された逆転写酵素コード配列) ; 配列番号 7 8 (8 5 TV 1 C. Z A 野
 生型に基づいて改変された変異逆転写酵素) ; 配列番号 8 0 (8 5 TV 1 C. Z A
 野生型に基づいて改変された、非機能的 Tat を生じる変異を含む Tat のエキソン 1)
 ; 配列番号 8 1 および配列番号 1 0 9 (それぞれ 8 5 TV 1 C. Z A 野生型および
 1 2 - 5 1 TV 2 C. Z A 野生型に基づいて改変された、Tat のエキソン 1) ;
 配列番号 8 3 および配列番号 1 1 1 (それぞれ 8 5 TV 1 C. Z A 野生型および 1 2 - 5 1 TV 2 C. Z A 野生型に基づいて改変された、Tat のエキソン 2) ; 配
 列番号 8 5 および配列番号 1 1 3 (それぞれ 8 5 TV 1 C. Z A 野生型および 1 2 - 5 1 TV 2 C. Z A 野生型に基づいて改変された Vif コード配列) ; 配列番
 号 8 7 および配列番号 1 1 5 (それぞれ 8 5 TV 1 C. Z A 野生型および 1 2 - 5 1 TV 2 C. Z A 野生型に基づいて改変された Vpr コード配列) ; 配列番号 8 9
 および配列番号 1 1 7 (それぞれ 8 5 TV 1 C. Z A 野生型および 1 2 - 5 1 TV 2 C. Z A 野生型に基づいて改変された Vpu コード配列) ; 配列番号 9 1 (8 5 TV 1 C. Z A 野生型に基づいて改変された、Rev のエキソン 1 およびエキソン
 2 の配列) ; 配列番号 9 3 (8 5 TV 1 C. Z A 野生型に基づいて改変された、T
 at の変異エキソン 1 および Tat のエキソン 2 の配列 (ここでエキソン 1 の変異は、非
 機能的 Tat を生じる)) ; 配列番号 9 4 (8 5 TV 1 C. Z A 野生型に基づいて
 改変された、Tat のエキソン 1 およびエキソン 2 の配列) ; 配列番号 9 6 および配列番
 号 1 3 5 (それぞれ 8 5 TV 1 C. Z A 野生型および 1 2 - 5 1 TV 2 C. Z A
 に基づいて改変された、ミリストイル化部位を排除する変異を含む Nef コード配
 列) 。

【0116】

(2. 2. 1. 5 HIV-1 核酸コード配列を含む配列のさらなる改変)

本明細書中に記載される、C 型 HIV ポリペプチドコード発現カセットはまた、例えば、
 1 つ以上の導入遺伝子をコードする、1 つ以上のさらなる配列を含み得る。本発明の実施
 において有用なさらなる配列 (例えば、導入遺伝子) としては、以下が挙げられるが、こ
 れらに限定されない : さらなるウイルスエピトープ / 抗原 {HCV 抗原 (例えば、E1、
 E2 ; Houghton, M. . ら、米国特許第 5, 714, 596 号 (1998 年 2 月
 3 日発行) ; Houghton, M. . ら、米国特許第 5, 712, 088 号 (1998
 年 1 月 27 日発行) ; Houghton, M. . ら、米国特許第 5, 683, 864 号 (1997 年 11 月 4 日発行) ; Weiner, A. J. ら、米国特許第 5, 728, 520 号 (1998 年 3 月 17 日発行) ; Weiner, A. J. ら、米国特許第 5, 766, 845 号 (1998 年 6 月 16 日発行) ; Weiner, A. J. ら、米国特許第 5, 670, 152 号 (1997 年 9 月 23 日発行))、HIV 抗原 (例えば、tat、rev、nef および / または env 由来のもの) を含むがこれらに限定されない} をコード
 するさらなる配列 ; および腫瘍抗原 / エピトープをコードする配列。さらなる配列はまた
 、非ウイルス供給源に由来し得る (例えば、インターロイキン-2 (IL-2)、幹細胞
 因子 (SCF)、インターロイキン 3 (IL-3)、インターロイキン 6 (IL-6)、
 インターロイキン 12 (IL-12)、G-CSF、顆粒球マクロファージコロニー刺激
 因子 (GM-CSF)、インターロイキン-1α (IL-1I)、インターロイキン-1
 1 (IL-11)、MIP-1I、腫瘍壊死因子 (TNF)、白血病抑制因子 (LIF)
 、c-kit リガンド、トロンプオエチン (TPO) および flt3 リガンドのようなサ
 イトカイン (例えば、Genzyme (Framingham, MA)、Genentech (South San Francisco, CA)、Amgen (Thousand Oaks, CA)、R&D Systems and Immunex (Seattle, WA) のようないくつかの販売元から市販されている) をコードする配列) 。さら

なる配列は、例えば 2. 3 節で、以下に記載されている。また、Gag および他のコード配列の、互いに対する配向のバリエーションは、以下に記載されている。

【0117】

HIV ポリペプチドコード配列は、他の C 型 HIV 単離物から得られ得る。例えば、Myers ら、Los Alamos Database, Los Alamos National Laboratory, Los Alamos, New Mexico (1992); Myers ら、Human Retroviruses and Aids, 1997, Los Alamos, New Mexico: Los Alamos National Laboratory を参照のこと。合成発現カセットは、このようなコード配列を出発物質として使用して、本明細書の教示に従って、生成され得る（例えば、実施例 1 を参照のこと）。 10

【0118】

さらに、本発明の合成発現カセットは、本明細書中に開示される合成発現カセット配列（例えば、配列番号 30～32；配列番号 3、4、20、および 21、ならびに配列番号 5～17）に対して、85% より大きい、好ましくは 90% より大きい、より好ましくは 95% より大きい、そして最も好ましくは 98% より大きい配列同一性を有する、関連するポリペプチド配列を含む。種々のコード領域が、図 3 および 4 に示されており、例えば図 3 (AF110968) においては、ヌクレオチド 1～81 (配列番号 18)；ヌクレオチド 82～1512 (配列番号 6) は、gp120 ポリペプチドをコードし、ヌクレオチド 1513～2547 (配列番号 10) は、gp41 ポリペプチドをコードし、ヌクレオチド 82～2025 (配列番号 7) は、gp140 ポリペプチドをコードし、そしてヌクレオチド 82～2547 (配列番号 8) は、gp160 ポリペプチドをコードする。同様に、図 98 (配列番号 127、株 8__2__TV1__C. ZA) においては、ヌクレオチド 1～6 は、EcoRI 制限部位であり；ヌクレオチド 7～87 は、野生型 (8__2__TV1__C. ZA 由来) リーダーシグナルペプチドをコードし；ヌクレオチド 88～1563 は、gp120 ポリペプチドをコードし；ヌクレオチド 88～2064 は、gp140 ポリペプチドをコードし；ヌクレオチド 88～2607 は、gp160 ポリペプチドをコードする。 20

【0119】

(2. 3. 3 HIV-1 サブタイプ C および関連するポリペプチドをコードする、合成配列の発現) 30

本発明の合成 HIV コード配列 (発現カセット) は、発現のレベルを評価するため、および Gag の場合には、VLP の産生を評価するために、多数の異なる発現ベクターにクローニングされ得る。HIV ポリペプチドに対する合成 DNA フラグメントは、真核生物発現ベクター (一過性発現ベクター、CMV プロモーターに基づく哺乳動物ベクター、およびバキュロウイルス発現系において使用するためのシャトルベクターが挙げられる) にクローニングされ得る。対応する野生型配列もまた、同じベクターにクローニングされ得る。

【0120】

次いで、これらのベクターは、いくつかの異なる細胞型 (種々の哺乳動物細胞株 (293、RD、COS-7、および CHO、例えば、A. T. C. C. から入手可能な細胞株) が挙げられる) にトランスフェクトされ得る。次いで、細胞株は、適切な条件下で培養され、そして任意の適切なポリペプチド産物のレベルが、上清中で評価され得る (表 A および実施例 2 を参照のこと)。例えば、p24 は、Gag 発現を評価するために使用され得る；gp160、gp140 または gp120 は、Env 発現を評価するために使用され得る；p6pol は、Pol 発現を評価するために使用され得る；prot は、プロテアーゼを評価するために使用され得る；p15 は、RNase H を評価するために使用され得る；p31 は、インテグラーゼを評価するために使用され得る；そして他の適切なポリペプチドは、Vif、Vpr、Tat、Rev、Vpu および Nef を評価するために使用され得る。さらに、改変されたポリペプチドがまた使用され得、例えば、他の Env ポ 40 50

リペプチドとしては、例えば、ネイティブ gp160、オリゴマー gp140、モノマー gp120、ならびにこれらのポリペプチドの改変配列および／または合成配列が挙げられるが、これらに限定されない。これらのアッセイの結果は、合成 HIV ポリペプチドコード配列の発現が、対応する野生型配列より有意に高いことを実証する。

【0121】

さらに、ウエスタンブロット分析を使用して、合成発現カセットを含む細胞が、ネイティブ発現カセットを含む細胞より高い細胞あたりの濃度で、予測されたタンパク質を発現することを示し得る。HIV タンパク質は、細胞溶解物と上清との両方において見られ得る。本発明の合成発現カセットでトランスフェクトされた細胞については、産生のレベルは、細胞上清において有意により高い。

10

【0122】

合成発現カセットでトランスフェクトされた哺乳動物細胞からの上清の分画を使用して、これらのカセットが、野生型配列と比較して、HIV タンパク質の優れた産生（そして Gag の場合には、VLP の産生）を提供することを示し得る。

【0123】

これらの HIV 含有ポリペプチドの、哺乳動物細胞株における効率的な発現は、以下の利点を提供する：ポリペプチドは、バキュロウイルス夾雑物を含まない；確立された方法による産生が FDA によって認可されている；増加した純度；より高い収率（ネイティブコード配列と比較して）；および CHO 細胞におけるサブタイプ C HIV 含有ポリペプチドの産生（これは、本発明の構築物を使用して得られる増加した発現の非存在下では実行不可能である）の新規方法。例示的な哺乳動物細胞株としては、BHK、VERO、HT1080、293、293T、RD、COS-7、CHO、Jurkat、HUT、SUP T、C8166、MOLT4／クローン8、MT-2、MT-4、H9、PM1、CEM、および CEMX174 が挙げられるが、これらに限定されず、このような細胞株は、例えば、A. T. C. C. から入手可能である。

20

【0124】

本発明の合成 Gag 発現カセットはまた、昆虫細胞にトランスフェクトされる場合に、高レベルの発現および VLP 産生を示す。本明細書中に記載される合成発現カセットはまた、昆虫細胞において高レベルの発現を示す。さらに、より高い全タンパク質収率に加えて、合成ポリペプチドからの最終生成物は、一貫して、ネイティブの C 型配列からの最終生成物より低い量の夾雑バキュロウイルスタンパク質を含む。

30

【0125】

さらに、本発明の合成発現カセットはまた、酵母ベクターに導入され得、このベクターは次に、酵母細胞 (*Saccharomyces cerevisiae*; Rosenberg, S. および Tekamp-Olson, P., 米国再発行特許発明第 35, 749 号 (1998 年 3 月 17 日発行) に記載されるようなベクターを使用する) に形質転換され得、そして酵母細胞によって効率的に発現され得る。

【0126】

哺乳動物ベクターおよび昆虫ベクターに加えて、本発明の合成発現カセットは、選択された発現制御エレメントを使用して、種々の発現ベクターに組み込まれ得る。任意の所定の細胞型に適切なベクターおよび制御エレメントは、本明細書の教示、および発現ベクターについて当該分野において公知の情報を考慮して、当業者によって選択され得る。

40

【0127】

例えば、合成発現カセットは、所望のコード配列に作動可能に連結された制御エレメントを含むベクターに挿入され得、これによって、選択された細胞型における遺伝子の発現が可能となる。例えば、哺乳動物細胞発現のための代表的なプロモーターとしては、とりわけ、SV40 初期プロモーター、CMV プロモーター（例えば、CMV 極初期プロモーター (CMV プロモーターはイントロン A を含み得る)）、RSV、HIV-Ltr、マウス乳腺癌ウイルス LTR プロモーター (MMLV-Ltr)、アデノウイルスの主要後期プロモーター (Ad MLP)、および単純疱疹ウイルスプロモーターが挙げられる。他

50

の非ウイルスプロモーター（例えば、マウスメタロチオネイン遺伝子由来のプロモーター）もまた、哺乳動物発現のために用途を見出す。代表的に、転写終結配列およびポリアデニル化配列もまた存在し、翻訳終止コドンに対して3'に位置する。好ましくは、翻訳の開始の最適化のための配列もまた存在し、コード配列に対して5'に位置する。転写終結シグナル／ポリアデニル化シグナルの例としては、Sambrookら、前出に記載されるようなSV40由来のシグナル、およびウシ成長ホルモントーミネーター配列が挙げられる。スプライスドナー部位およびアクセプター部位を含むイントロンもまた、本発明と共に使用するための構築物中に設計され得る（Chapmanら、Nuc. Acids Res. (1991) 19:3979-3986）。

【0128】

エンハンサーエレメントもまた、本明細書中において使用されて、哺乳動物構築物の発現レベルを増加させ得る。例としては、Dijkemaら、EMBO J. (1985) 4:761に記載されるようなSV40初期遺伝子エンハンサー、Gormanら、Proc. Natl. Acad. Sci. USA (1982b) 79:6777に記載されるようなラウス肉腫ウイルスの長末端反復（LTR）由来のエンハンサー／プロモーター、およびBoshartら、Cell (1985) 41:521に記載されるようなヒトCMV由来のエレメント（例えば、CMVイントロンA配列に含まれるエレメント（Chapmanら、Nuc. Acids Res. (1991) 19:3979-3986））が挙げられる。

【0129】

所望の合成ポリペプチドをコードする配列は、多数の市販のベクターにクローニングされて、このポリペプチドの、適切な宿主系における発現を生じさせ得る。これらの系としては、以下が挙げられるが、これらに限定されない：バキュロウイルス発現 {Reilly, P. R. ら、BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992); Beamesら、Biotechniques 11:378 (1991); Pharmingen; Clontech, Palo Alto, CA}、ワクシニア発現 {Earl, P. L. ら、「Expression of proteins in mammalian cells using vaccinia」Current Protocols in Molecular Biology. (F. M. Ausubelら編), Greene Publishing Associates & Wiley Interscience, New York (1991); Moss, B. ら、米国特許第5,135,855号 (1992年8月4日発行)}、細菌における発現 {Ausubel, F. M. ら、CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley and Sons, Inc., Media PA; Clontech}、酵母における発現 {Rosenberg, S. およびTekamp-Olson, P., 米国再発行特許発明第35,749号 (1998年3月17日発行); Shuster, J. R., 米国特許第5,629,203号 (1997年5月13日発行); Gellissen, G., ら、Antonie Van Leeuwenhoek, 62(1-2):79-93 (1992); Romanos, M. A., ら、Yeast 8(6):423-488 (1992); Goeddel, D. V., Methods in Enzymology 185 (1990); Guthrie, C., およびG. R. Fink, Methods in Enzymology 194 (1991)}、哺乳動物細胞における発現 {Clontech; Gibco-BRL, Ground Island, NY; 例えば、チャイニーズハムスター卵巣（CHO）細胞株 (Haynes, J. ら、Nuc. Acid. Res. 11:687-706 (1983); 1983, Lau, Y. F. ら、Mol. Cell. Biol. 4:1469-1475 (1984); Kaufman, R. J., 「Selection and coamplification of heterologous genes in mammalian cells」、Methods in Enzymology, 第185巻, 537-566頁、Aca

10

20

30

40

50

demic Press, Inc., San Diego CA (1991)}、および植物細胞における発現 {植物クローニングベクター、Clontech Laboratories, Inc., Palo Alto, CA、およびPharmacia LKB Biotechnology, Inc., Piscataway, NJ; Hood, E. ら, J Bacteriol. 168:1291-1301 (1986); Nagel, R. ら, FEMS Microbiol. Lett. 67:325 (1990); An ら, 「Binary Vectors」、およびPlant Molecular Biology Manual A3:1-19 (1988); Miki, B. L. A. ら, 249-265頁における他のもの、ならびにPlant DNA Infectious Agents (Hohn, T. ら編) Springer-Verlag, Wien, Austria, (1987); Plant Molecular Biology: Essential Techniques, P. G. Jones および J. M. Sutton, New York, J. Wiley, 1997; Miglani, Gurbachan Dictionary of Plant Genetics and Molecular Biology, New York, Food Products Press, 1998; Henry, R. J., Practical Applications of Plant Molecular Biology, New York, Chapman & Hall, 1997における他のもの}。 10

【0130】

本発明にはまた、コード配列、および適切な宿主におけるコード領域の発現を可能にする発現制御エレメントを含む発現ベクターが含まれる。これらの制御エレメントは、一般に、プロモーター、翻訳開始コドン、および翻訳終止配列および転写終止配列、ならびにベクターへ挿入を導入するための挿入部位を含む。転写制御エレメントは、M. Kozak によって概説されている (例えば、Kozak, M., Mamm. Genome 7 (8): 563-574, 1996; Kozak, M., Biochimie 76 (9): 815-821, 1994; Kozak, M., J Cell Biol 108 (2): 229-241, 1989; Kozak, M., および Shatkin, A. J., Methods Enzymol 60:360-375, 1979)。 20

【0131】

酵母系における発現は、工業的生産の利点を有する。ワクシニアおよびCHO細胞株による組換えタンパク質の産生は、哺乳動物発現系であるという利点を有する。さらに、ワクシニアウイルスの発現は、以下を含むいくつかの利点を有する: (i) その広範な宿主範囲; (ii) 組換えタンパク質の正確な転写後修飾、プロセッシング、折り畳み、輸送、分泌および構築; (iii) 比較的可溶性の組換えタンパク質の高レベルな発現; ならびに (iv) 外来DNAに適合する大きな能力。 30

【0132】

合成HIVポリペプチドコード発現カセットから組換え的に発現されたポリペプチドは、典型的に、溶解された細胞および培養培地から単離される。精製は、当該分野で公知の方法 (塩分画、イオン交換クロマトグラフィー、ゲル濾過、サイズ排除クロマトグラフィー、サイズ分画およびアフィニティークロマトグラフィーを含む) によって、実施され得る。免疫親和性クロマトグラフィーは、例えば、HIV抗原に基づいて作製された抗体を使用して、用いられ得る。 40

【0133】

哺乳動物細胞を使用して本発明のタンパク質を発現させる利点としては、以下が挙げられるが、これらに限定されない: 産生のスケールアップのための十分に確立されたプロトコル; VLPを生成する能力; 細胞株が優良製造プロセス (GMP) 基準を満たすように適合されること; 哺乳動物細胞の培養条件が当該分野で公知であること。

【0134】

本明細書中に記載される本発明の異なる実施形態の様々な形態は、組み合わされてもよい。

【0135】

(2.3 ウイルス様粒子の生成およびパッケージング細胞株を作製するための本発明の構築物の使用)

ヒト免疫不全ウイルス1型(HIV-1)の群特異的抗原(Gag)は、発芽によって種々の真核生物細胞から放出される非感染性ウイルス様粒子(VLP)に自己構築する(Freed, E. O., Virology 251:1-15, 1998による総説)。本発明の合成発現カセットは、種々の異なる細胞型(哺乳動物細胞が挙げられるが、これに限定されない)を使用するHIV-Gagウイルス様粒子(VLP)の生成に有効な手段を提供する。

【0136】

ウイルス粒子は、宿主の免疫系に対する、この粒子に捕捉または結合された抗原の適切な提示のためのマトリクスとして、使用され得る。

【0137】

(2.3.1 本発明の合成発現カセットを使用するVLP生成)

本発明を支持して実験を行って、本発明の合成発現カセットが、ネイティブのGagコード配列と比較して、Gagタンパク質およびVLPの両方の優れた生成を提供することを証明し得る。さらに、VLP生成の電子顕微鏡評価により、推定サイズの遊離の発芽している未成熟ウイルス粒子は、この合成発現カセットを含む細胞によって生成されることが示され得る。

【0138】

ウイルス様粒子の生成のために、ネイティブのGagコード配列ではなく、本発明の合成発現カセットを使用することは、いくつかの利点を提供する。第1に、VLPは、増大した量で生成され得、これによりこのVLPの単離および精製がより容易になり得る。第2に、VLPは、この合成発現カセットを使用して種々の細胞型で生成され得、特に、哺乳動物細胞株(例えば、CHO細胞)は、VLP生成のために使用され得る。CHO細胞を使用する生成は、以下を提供する:(i) VLP形成;(ii) 正確なミリストイル化および発芽;(iii) 非哺乳動物細胞混入物(例えば、昆虫ウイルスおよび/または細胞)の非存在;および(iv) 精製の容易さ。本発明の合成発現カセットはまた、哺乳動物細胞株以外の細胞型における発現を増大するために有用である。例えば、この合成発現カセットをコードするバキュロウイルスベクターで昆虫細胞を感染することにより、高レベルのGagタンパク質の全収率およびより高いレベルのVLP生成が生じる(野生型コード配列と比較して)。さらに、バキュロウイルスGag合成発現カセットに感染した昆虫由来の最終産物は、野生型コード配列を使用した場合の最終産物より少ない量の混入昆虫タンパク質を一貫して含む。

【0139】

VLPは、目的の粒子形成ポリペプチドが適切な宿主細胞において組換え発現される場合に、自然に形成され得る。従って、本発明の合成発現カセットを使用して生成されるVLPは、組換え技術を使用して簡便に調製される。下記のように、本発明のGagポリペプチドコード合成発現カセットは、目的の他のポリペプチドコード配列(例えば、HIVプロテアーゼ、HIVポリメラーゼ、HCVコア、Env、合成Env;実施例1を参照のこと)を含み得る。このような合成発現カセットの発現は、Gagポリペプチドを含むVLPならびに目的のポリペプチドを生じる。

【0140】

一旦、所望の粒子形成ポリペプチドのコード配列が単離または合成されると、これらのコード配列は、発現に適切な任意のベクターまたはレプリコンにクローニングされ得る。多数のクローニングベクターが、当業者に公知であり、適切なクローニングベクターの選択が可能である。一般的には、Sambrookら(前出)を参照のこと。次いで、このベクターは、適切な宿主細胞を形質転換するために使用される。適切な組換え発現系としては、細菌発現系、哺乳動物発現系、バキュロウイルス/昆虫発現系、ワクシニア発現系、セムリキ森林熱ウイルス(SFV)発現系、アルファウイルス(例えば、シンドビスウイ

10

20

30

40

50

ルス、ベネズエラウマ脳脊髄炎（V E E）ウイルス）発現系、哺乳動物発現系、酵母発現系およびゼノブス発現系が挙げられるが、これらに限定されず、これらは当該分野で周知である。特に好まし発現系は、哺乳動物細胞株、ワクシニア、シンドビス、昆虫および酵母発現系である。

【0141】

例えば、多数の哺乳動物細胞株が当該分野で公知であり、そしてAmerican Type Culture Collection (A. T. C. C.) から入手可能な不死化細胞株を含み、この不死化細胞株としては、チャイニーズハムスター卵巣 (CHO) 細胞、HeLa細胞、新生ハムスター腎臓 (BHK) 細胞、サル腎臓細胞 (COS) などが挙げられるが、これらに限定されない。同様に、細菌宿主 (例えば、E. coli、Bacillus subtilis および Streptococcus spp.) は、本発明の発現構築物を用いる用途を見出す。本発明において有用な酵母宿主としては、特に、Saccharomyces cerevisiae、Candida albicans、Candida maltosa、Hansenula polymorpha、Kluyveromyces fragilis、Kluyveromyces lactis、Pichia guillierimondii、Pichia pastoris、Schizosaccharomyces pombe および Yarrowia lipolytica が挙げられる。バキュロウイルス発現ベクターと共に使用するための昆虫細胞としては、特に、Aedes aegypti、Autographa californica、Bombyx mori、Drosophila melanogaster、Spodoptera frugiperda および Trichoplusia ni. が挙げられる。例えば、Summers および Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987) を参照のこと。

【0142】

ウイルスベクター (例えば、ポックスファミリーのウイルス (ワクシニアウイルスおよび鳥類ポックスウイルスを含む) 由来のベクター) は、真核生物細胞における粒子の生成のために使用され得る。さらに、ワクシニアベースの感染／トランスフェクション系はまた、Tomei ら、J. Virol. (1993) 67: 4017-4026 および Selby ら、J. Gen. Virol. (1993) 74: 1103-1113 に記載されるように、本発明を伴う用途を見出す。この系において、細胞を、始めに、バクテリオファージ T7 RNA ポリメラーゼをコードするワクシニアウイルス組換え体で、インビトロで感染させる。このポリメラーゼは、T7 プロモーターを保有するテンプレートを転写するのみであるという点で、優れた特異性を示す。感染後、細胞を、T7 プロモーターにより駆動される目的の DNA でトランスフェクトする。ワクシニアウイルス組換え体由来の細胞質において発現されるポリメラーゼは、このトランスフェクトされた DNA を RNA に転写し、この RNA は次いで、宿主翻訳機構によりタンパク質に翻訳される。あるいは、T7 は、精製タンパク質として添加され得るか、または「前駆体」系の場合には酵素として添加される (Studier および Moffatt, J. Mol. Biol. (1986) 189: 113-130)。この方法は、多量の RNA およびその翻訳産物の高レベルの一過性の細胞質生成を提供する。

【0143】

選択された発現系および宿主に依存して、VLP は、粒子形成ポリペプチドが発現されて VLP が形成され得る条件下で、発現ベクターにより形質転換された宿主細胞を増殖することにより生成される。適切な増殖条件の選択は、当業者の範囲内である。VLP が細胞内で形成される場合、この細胞は、細胞を溶解するが VLP を実質的にインタクトなまま保つ、化学的手段、物理的手段または機械的手段を使用して破壊される。このような方法は、当業者に公知であり、そして例えば、Protein Purification Applications: A Practical Approach (E. L. V. Harris および S. Angal 編、1990) に記載される。

10

20

30

40

50

【0144】

次いで、この粒子は、それらの完全性を保存する方法（例えば、勾配遠心分離（例えば、塩化セシウム（CsCl）スクロース勾配、ペレット化など））（例えば、Kirnbaueら、J. Virol.（1993）67：6939-6936を参照のこと）、ならびに標準的な精製技術（例えば、イオン交換クロマトグラフィーおよびゲル濾過クロマトグラフィー）によって、単離される（または実質的に精製される）。

【0145】

本発明の合成発現カセットを含む細胞により生成されるVLPは、被験体に投与された場合に免疫応答を引き起こすために、使用され得る。本発明の1つの利点は、VLPが、以前には不可能であったレベルで、合成発現カセットを有する哺乳動物細胞により生成され得るということである。上記のように、このVLPは、Gagポリペプチド（例えば、Gagプロテアーゼ、Gagポリメラーゼ、Env、合成Envなど）に加えて、種々の抗原を含み得る。本発明の合成発現カセットを使用して生成される、精製VLPは、通常はワクチン組成物の形態で、脊椎動物被験体に投与され得る。組み合わせワクチンもまた使用され得、ここで、このようなワクチンは、例えば、アジュバントサブユニットタンパク質（例えば、Env）を含む。投与は、単独で処方されたVLPまたは他の抗原と共に処方されたVLPを使用して、行われ得る。さらに、このVLPは、DNA免疫（以下を参照のこと）および／または他のワクチンの送達のための合成発現カセットの送達の前か、これと同時に、またはこれに続いて投与され得る。また、VLP投与の部位は、投与されている他のワクチン組成物と同じであっても異なってもよい。遺伝子送達は、多数の方法によって達成され得、この方法としては、DNA、アルファウイルスベクター、ボックスウイルスベクターおよびワクシニアウイルスベクターでの免疫が挙げられるが、これらに限定されない。

10

20

【0146】

VLP免疫刺激（または、ワクチン）組成物は、種々の賦形剤、アジュバント、キャリア、補助物質、調節剤などを含み得る。免疫刺激組成物は、免疫学的応答を開始するのに十分な量のVLP／抗原を含む。適切な有効量は、当業者によって決定され得る。このような量は、慣用的な試験により決定され得る比較的広い範囲内に含まれ、一般に、約0.1 μg～約1000 μg、より好ましくは約1 μg～約300 μgの程度のVLP／抗原の量である。

30

【0147】

キャリアが必要に応じて存在し、ここでこのキャリアは、組成物を受ける個体に対して有害な抗体の生成をそれ自体は引き起こさない分子である。適切なキャリアは、典型的には、ゆっくりと代謝される大きな高分子（例えば、タンパク質、多糖類、ポリ乳酸、ポリグリコール酸、高分子アミノ酸、アミノ酸コポリマー、脂質凝集体（例えば、油滴またはリポソーム）、および不活性なウイルス粒子）である。粒子性キャリアの例としては、ポリメチルメタクリレートポリマーから誘導されるキャリア、ならびにポリ（ラクチド）およびポリ（ラクチド-co-グリコリド）（PLGとして公知）から誘導される微粒子が挙げられる。例えば、Jefferyら、Pharm. Res.（1993）10：362-368；McGee JPら、J. Microencapsul. 14（2）：197-210，1997；O'Hagan DTら、Vaccine 11（2）：149-54，1993を参照のこと。このようなキャリアは、当業者に周知である。さらに、これらのキャリアは、免疫刺激因子（「アジュバント」）として機能し得る。さらに、抗原は、細菌毒素（例えば、ジフテリア、テタヌス、コレラなど由来の毒素）、およびE. coli由来の毒素に結合され得る。

40

【0148】

アジュバントはまた、組成物の効力を増大するために使用され得る。このようなアジュバントとしては、以下が挙げられるが、これらに限定されない：（1）アルミニウム塩（ミョウバン）（例えば、水酸化アルミニウム、リン酸アルミニウム、硫酸アルミニウムなど）；（2）水中油型エマルジョン処方物（ムラミルペプチド（以下を参照のこと）または

50

細菌細胞壁成分のような他の特定の免疫刺激因子を含むかまたは含まない) (例えば、(a) MF59 (国際公開番号WO90/14837) (これは5% スクアレン、0.5% Tween 80および0.5% Span 85を含み(必要に応じて種々の量のMTP-PEを含む(以下を参照のこと)が、必要ではない)、Model 110Yマイクロフルイダイザー(Microfluidics, Newton, MA)のようなマイクロフルイダイザーを使用してサブミクロンの粒子に処方される)、(b) SAF (これは、10% スクアレン、0.4% Tween 80、5% プルロニックブロックポリマー(pluronic-blocked Polymer) L121、およびthr-MDP (以下を参照のこと)を含み、サブミクロンエマルジョンに微小流体化されるか、または大きな粒子サイズのエマルジョンを作製するためにボルテックスされるかのいずれかである)、および(c) RibiTM アジュバント系(RAS) (Ribi Immunochem, Hamilton, MT) (これは2% スクアレン、0.2% Tween 80、およびモノホスホリルリピド A (MPL)、トレハロースジミコレート(TDM)および細胞壁骨格(CWS)からなる群から選択される1つ以上の細菌細胞壁成分(好ましくは、MPL+CWS (DetoxTM)を含む)) ; (3) StimulonTM (Cambridge Bioscience, Worcester, MA)のようなサポニンアジュバントが使用され得るか、またはこれから作製された粒子(例えば、ISCOM (免疫刺激複合体)が使用され得る ; (4) 完全フロイントアジュバント(CFA)および不完全フロイントアジュバント(IFA) ; (5) サイトカイン(例えば、インターロイキン(IL-1、IL-2など)、マクロファージコロニー刺激因子(M-CSF)、腫瘍壊死因子(TNF)など ; (6) 免疫刺激性CpGモチーフをコードするオリゴヌクレオチドまたはポリマー分子(Davis, H. L. ら、J. Immunology 160:870-876, 1998; Sato, Y. ら、Science 273:352-354, 1996)、または抗原/オリゴヌクレオチドの複合体{ポリマー分子は、二本鎖および一本鎖のRNAおよびDNA、それらの骨格改変体(例えば、メチルホスホネート結合)を含む ; あるいは(7) 細菌ADP-リボシル化毒素(LT) (例えば、コレラ毒素(CT)、百日咳毒素(PT)もしくはE. coli 熱不安定性毒素(LT)、特に、LT-K63 (ここでリジンが、63位の野生型アミノ酸と置換される)、LT-R72 (ここでアルギニンが、72位の野生型アミノ酸と置換される)、CT-S109 (ここでセリンが、109位の野生型アミノ酸と置換される)、およびPT-K9/G129 (ここでリジンが、9位の野生型アミノ酸と置換され、グリシンが129位で置換される))の無毒化変異体(例えば、国際公開番号WO93/13202およびWO92/19265を参照のこと) ; および(8) 免疫刺激因子として働き、組成物の効力を増大する他の物質。さらに、このようなポリマー分子は、代替のポリマー骨格構造(例えば、ポリビニル骨格(Pitha, Biochem Biophys Acta, 204:39, 1970a; Pitha, Biopolymers, 9:965, 1970b)、およびモルホリン骨格(Summerton, J. ら、米国特許第5, 142, 047号、08/25/92発行; Summerton, J. ら、米国特許第5, 185, 444号、02/09/93発行)(これらに限定されない))を含む。種々の他の電荷および無電荷のポリヌクレオチドアナログが報告されている。多数の骨格修飾が当該分野で公知であり、この修飾には、無電荷結合(例えば、メチルホスホネート、ホスホトリエステル、ホスホアミデートおよびカルバメート)および電荷結合(例えば、ホスホロチオエートおよびホスホロジチオエート)が挙げられるが、これらに限定されない。} ; ならびに(7) 免疫刺激因子として働いてVLP免疫刺激(またはワクチン)組成物の有効性を増大する他の物質。ミョウバン、CpGオリゴヌクレオチドおよびMF59が好ましい。

【0149】

ムラミルペプチドとしては、N-アセチル-ムラミル-L-トレオニル-D-イソグルタミン(thr-MDP)、N-アセチル-ノルムラミル-L-アラニル-D-イソグルタミン(nor-MDP)、N-アセチル-ムラミル-L-アラニル-D-イソグルタミル-

10

20

30

40

50

L-アラニン-2-(1'-2'-ジパルミトイル-sn-グリセロ-3-ヒドロキシホスホリルオキシ)-エチルアミン(MTP-PE)などが挙げられるが、これらに限定されない。

【0150】

VLP組成物を用いる投薬処置は、単回用量スケジュールまたは複数用量スケジュールであり得る。複数用量スケジュールは、ワクチン接種の一次過程に、1～10の別個の用量が用いられ、続いて他の用量が後の時間間隔（これは、免疫応答を維持および／または強化するように選択される、例えば、第2の用量について1～4ヶ月）で与えられ、そして必要に応じて、数ヶ月後に引き続く用量（単数または複数）が与えられる、スケジュールである。投薬レジメンはまた、少なくとも部分的に、被験体の要求によって決定され、そしてこれは、医師の判断に依存する。 10

【0151】

疾患の予防が所望される場合、抗原保有VLPは一般に、目的の病原体の一次感染の前に投与される。処置（例えば、症状および再発の減少）が所望される場合、VLP組成物は一般に、一次感染の後に投与される。

【0152】

(2.3.2 パッケージング細胞株を作製するための本発明の合成発現カセットの使用)

多数のウイルスベースの系は、哺乳動物宿主細胞に対する遺伝子移入ベクターとして使用するために開発されてきた。例えば、レトロウイルス（特に、レンチウイルスベクター）は、遺伝子送達系に好都合なプラットフォームを提供する。目的のコード配列（例えば、遺伝子治療の用途に有用な配列）は、遺伝子送達ベクターに挿入され得、当該分野で公知の技術を使用してレトロウイルス粒子中にパッケージングされる。次いで、組換えウイルスが単離され、そしてインビボまたはエキソビボのいずれかで、被験体の細胞に送達され得る。多数のレトロウイルス系は、例えば、以下に記載される：米国特許第5,219,740号；Millerら(1989) *BioTechniques* 7:980；Miller, A. D. (1990) *Human Gene Therapy* 1:5；Scarpaら(1991) *Virology* 180:849；Burnsら(1993) *Proc. Natl. Acad. Sci. USA* 90:8033；Boris-Lawrieら、(1993) *Cure. Opin. Genet. Develop.* 3:102；GB 2200651；EP 0415731；EP 0345242；WO 89/02468；WO 89/05349；WO 89/09271；WO 90/02806；WO 90/07936；WO 90/07936；WO 94/03622；WO 93/25698；WO 93/25234；WO 93/11230；WO 93/10218；WO 91/02805；U. S. 5,219,740；；U. S. 4,405,712；U. S. 4,861,719；U. S. 4,980,289およびU. S. 4,777,127；米国出願番号07/800,921；およびVile(1993) *Cancer Res* 53:3860-3864；Vile(1993) *Cancer Res* 53:962-967；Ram(1993) *Cancer Res* 53:83-88；Takamiya(1992) *J Neurosci Res* 33:493-503；Baba(1993) *J Neurosurg* 79:729-735；Mann(1983) *Cell* 33:153；Cane(1984) *Proc Natl Acad Sci USA* 81:6349；ならびにMiller(1990) *Human Gene Therapy* 1。 20 30 40

【0153】

他の実施形態において、遺伝子移入ベクターは、サイトカインまたは他の免疫調節性分子をコードするように構築され得る。例えば、ネイティブのIL-2およびγ-インターフェロンをコードする核酸配列は、それぞれ、米国特許第4,738,927号および同第5,326,859号に記載される様にして、得られ得、一方、これらのタンパク質の有用なムテインは、米国特許第4,853,332号に記載されるようにして得られ得る。 50

mC S Fの短い形態および長い形態をコードする核酸配列は、それぞれ、米国特許第4, 847, 201号および同第4, 879, 227号に記載されるようにして得られ得る。本発明の特定の局面において、サイトカインを発現するレトロウイルスベクターまたは免疫調節性遺伝子は、本明細書中（例えば、本発明のパッケージング細胞株を使用する）および国際出願番号PCT US 94/02951（表題「Compositions and Methods for Cancer Immunotherapy」）に記載されるようにして、生成され得る。

【0154】

本明細書中で使用するために適切な免疫調節性分子の例としては、以下が挙げられる：IL-1およびIL-2（Karupiahら、(1990) J. Immunology 144:290-298, Weberら、(1987) J. Exp. Med. 166:1716-1733, Gansbacherら、(1990) J. Exp. Med. 172:1217-1224および米国特許第4, 738, 927号）；IL-3およびIL-4（Tepperら、(1989) Cell 503-512, Golumbekら、(1991) Science 254:713-716および米国特許第5, 017, 691）；IL-5およびIL-6（Brakenhofら(1987) J. Immunol. 139:4116-4121および国際公開番号WO90/06370）；IL-7（米国特許第4, 965, 195号）；IL-8、IL-9、IL-10、IL-11、IL-12およびIL-13（Cytokine Bulletin, Summer 1994）；IL-14およびIL-15； α -インターフェロン（Finterら、(1991) Drugs 42:749-765, 米国特許第4, 892, 743号および同第4, 966, 843号、国際公開番号WO85/02862, Nagataら(1980) Nature 284:316-320, Familettaら(1981) Methods in Enz. 78:387-394, Twuら(1989) Proc. Natl. Acad. Sci. USA 86:2046-2050, およびFaktorら、(1990) Oncogene 5:867-872）； β -インターフェロン（Seifら、(1991) J. Virol. 65:664-671）； γ -インターフェロン（Radfordら(1991) The American Society of Hepatology 20082015, Watanabeら(1989) Proc. Natl. Acad. Sci. USA 86:9456-9460, Gansbacherら(1990) Cancer Research 50:7820-7825, Maiora、(1989) Can. Immunol. Immunother. 30:34-42、ならびに米国特許第4, 762, 791号および同第4, 727, 138号）；G-C S F（米国特許第4, 999, 291号および同第4, 810, 643号）；GM-C S F（国際公開番号WO85/04188）。

【0155】

免疫調節性因子はまた、これらの分子に対するアゴニスト、アンタゴニストまたはリガンドであり得る。例えば、可溶性形態のレセプターは、しばしば、変異形態の因子自体と同様に、これらの型の因子に対するアンタゴニストとして挙動し得る。

【0156】

上記の基質をコードする核酸分子、本発明での使用に有利である他の核酸分子は、種々の供給源（例えば、American Type Collectionのような寄託所を含む）から、またはBritish Bio-Technology Limited（Cowley, Oxford England）のような商業的供給源から容易に得られ得る。代表的な例としては、BBG12（127個のアミノ酸の成熟タンパク質をコードするGM-C S F遺伝子を含む）、BBG6（これは、 γ インターフェロンをコードする配列を含む）、A. T. C. C. 寄託番号39656（これは、TNFをコードする配列を含む）、A. T. C. C. 寄託番号20663（これは、 α インターフェロンをコードする配列を含む）、A. T. C. C. 寄託番号31902、31902および39517（これは、 β インターフェロンをコードする配列を含む）、A. T. C. C. 寄託番号6

10

20

30

40

50

7024（これは、インターロイキン1bをコードする配列を含む）、A. T. C. C. 寄託番号39405、39452、39516、39626および39673（これは、インターロイキン2をコードする配列を含む）、59399、59398および67326（これは、インターロイキン3をコードする配列を含む）、A. T. C. C. 寄託番号57592（これは、インターロイキン4をコードする配列を含む）、A. T. C. C. 寄託番号59394および59395（これは、インターロイキン5をコードする配列を含む）、ならびにA. T. C. C. 寄託番号67153（これは、インターロイキン6をコードする配列を含む）が挙げられる。

【0157】

サイトカイン遺伝子または免疫調節遺伝子を含むプラスミド（国際公開番号WO94/02951およびWO96/21015）は、適切な制限酵素で消化され得、そして目的の特定の遺伝子を含むDNAフラグメントは、標準的な分子生物学的技術（例えば、Sambrookら、前出、またはAusbelら（編）Current Protocols in Molecular Biology、Greene Publishing and Wiley-Interscience）を使用して、遺伝子伝達ベクターに挿入され得る。

【0158】

上記の分子をコードするポリヌクレオチド配列は、例えば、遺伝子を発現するcDNAおよびゲノムライブラリーをスクリーニングすることによって、またはこれを含むことで公知のベクターから遺伝子を誘導することによって得られ得る。例えば、変更された細胞産物をコードする配列を含むプラスミドは、A. T. C. C. のような寄託所からまたは商業的供給源から得られ得る。目的のヌクレオチド配列を含むプラスミドは、適切な制限酵素で消化され得、そしてこのヌクレオチドを含むDNAフラグメントは、標準的な分子生物学的な技術を使用して遺伝子伝達ベクターに挿入され得る。

【0159】

あるいは、本発明での使用のためのcDNA配列は、標準的な技術（例えば、フェノール抽出およびcDNAまたはゲノムDNAのPCR）を使用して、配列を発現するか、または含む細胞から得られ得る。例えば、Sambrookら、前出（DNAを得るためおよびDNAを単離するために使用される技術の記載）を参照のこと。簡潔には、目的の遺伝子を発現する細胞由来のmRNAは、オリゴdTまたはランダムポリマーを使用して、逆転写酵素で逆転写され得る。次いで、この単鎖cDNAは、所望の配列のいずれかの側の配列に相補的なオリゴヌクレオチドプライマーを使用して、PCRによって増幅され得る（米国特許第4,683,202号、銅第4,683,195号、および同第4,800,159号を参照のこと、またPCR Technology: Principles and Applications for DNA Amplification, Erlich（編）、Stockton Press, 1989を参照のこと）。

【0160】

目的のヌクレオチド配列はまた、DNA合成器（例えば、Applied Biosystems Model 392 DNA Synthesizer（ABI、Foster City, Californiaから市販されている））を使用して、クローン化ではなく合成により生成され得る。このヌクレオチド配列は、所望される発現産物の適切なコドンを用いて設計され得る。完全な配列が、標準的な方法によって調製されたオーバーラップオリゴヌクレオチドから構築され、そして完全なコード配列に構築される。例えば、Edge（1981）Nature 292:756; Nambairら（1984）Science 223:1299; Jayら（1984）J. Biol. Chem. 259:6311を参照のこと。

【0161】

本発明の合成発現カセットは、レンチウイルスベクターを使用するために、パッケージング細胞株を構築において使用され得る。

【0162】

10

20

30

40

50

1つの型のレトロウイルス、マウスの白血病ウイルス、または「MLV」は、遺伝子治療適用のために広範に利用されている（一般に、Mannら（Cell 33:153, 1993）、CaneおよびMulligan（Proc, Nat'l Acad. Sci. USA 81:6349, 1984）、ならびにMillerら、Human Gene Therapy 1:5-14, 1990を参照のこと）。

【0163】

レンチウイルスベクターは、代表的に、5'レンチウイルスLTR、tRNA結合部位、パッケージングシグナル、1個以上の目的の遺伝子に作動可能に連結したプロモーター、第2のDNA鎖合成の起点、3'レンチウイルスLTRを含み、ここで、このレンチウイルスベクターは、核輸送エレメントを含む。この核輸送エレメントは、目的のコード配列（例えば、目的の合成Gagまたは合成Envの発現カセット）の上流（5'）または下流（3'）のいずれかに配置され得る。特定の実施形態において、核輸送エレメントは、RREではない。一実施形態において、パッケージングシグナルは、伸長されたパッケージングシグナルである。他の実施形態において、このプロモーターは、組織特異的なプロモーターであるか、あるいはCMVのようなプロモーターである。他の実施形態において、このレンチウイルスベクターは、内部リボソーム導入部位（entry site）をさらに含む。

【0164】

広範な種々のレンチウイルスは、本発明の範囲内で利用され得、これには、例えば、HIV、HIV-1、HIV-2、FIVおよびSIVからなる群から選択されるレンチウイルスが挙げられる。

【0165】

本発明の一実施形態において、合成Gagポリメラーゼ発現カセット（プロモーター、ならびに合成Gagポリメラーゼおよびvpr、vpu、nefまたはvifのうち少なくとも1個をコードする配列を含む）が提供され、ここで、このプロモーターは、Gagポリメラーゼおよびvpr、vpu、nefまたはvifに作動可能に連結される。

【0166】

本発明のなお別の局面において、宿主細胞（例えば、パッケージング細胞株）（本明細書中に記載される発現カセットのいずれかを含む）が、提供される。例えば、一局面内において、パッケージング細胞株（合成Gagポリメラーゼをコードする配列、および核輸送エレメントを含む発現カセットを含む）が、提供され、ここで、このプロモーターは、Gagポリメラーゼをコードする配列に作動可能に連結される。パッケージング細胞株は、tat、rev、またはエンベロープを含むプロモーターおよび配列をさらに含み得、ここで、このプロモーターは、tat、rev、Envをコードする配列、またはこれらのタンパク質の改変バージョンをコードする配列に作動可能に連結される。このパッケージング細胞株は、nef、vif、vpuまたはvpr（野生型または合成）のうちいずれか1つ以上をコードする配列をさらに含み得る。

【0167】

一実施形態において、発現カセット（例えば、合成Gagポリメラーゼを保有する）は、安定的に組み込まれる。このパッケージング細胞株は、レンチウイルスベクターの導入により、代表的に粒子を生成する。この合成発現カセットのプロモーター調節発現は、誘導性であり得る。代表的に、このパッケージング細胞株は、レンチウイルスベクターの導入により、複製コンピテントウイルスを本質的に含まない粒子を生成する。

【0168】

パッケージング細胞株（合成Gagポリメラーゼ遺伝子の発現を指向する発現カセットを含むか、または本明細書中に記載される合成Env遺伝子の発現を指向する発現カセットを含む）が、提供される。（他の改変Env配列の記載に関して、Andre, S. ら、Journal of Virology 72(2):1497-1503, 1998; Haas, J. ら、Current Biology 6(3):315-324, 1996も参照のこと）。レンチウイルスベクターは、ベクター産生細胞株を産生するため

にパッケージング細胞株に導入される。

【0169】

上記のように、レンチウイルスベクターは、目的の選択された遺伝子または配列を保有または発現するように設計され得る。レンチウイルスベクターは、広範の種々のレンチウイルスから容易に構築され得る（RNA Tumor Viruses, 第2版, Cold Spring Harbor Laboratory, 1985を参照のこと）。レンチウイルスの代表的な例としては、HIV、HIV-1、HIV-2、FIVおよびSIVが挙げられる。このようなレンチウイルスは、患者の単離物から、またはより好ましくは、the American Type Culture Collectionのような寄託機関または収集機関から得られるか、あるいは、利用可能な技術を使用して公知の供給源から単離されるかのいずれかである。 10

【0170】

レンチウイルス遺伝子送達ベクター（またはビヒクル）の部分は、異なるウイルス由来であり得る。例えば、所定の組換えレンチウイルスベクターにおいて、LTRは、HIV、SIV由来のパッケージングシグナル、およびHrV-2由来の第2の鎖合成の起点由来であり得る。レンチウイルスベクター構築物は、5'レンチウイルスLTR、tRNA結合部位、パッケージングシグナル、1以上の異種配列、第2のDNA鎖合成の起点、および3'LTRを含み得、ここで、このレンチウイルスベクターは、RREではない核輸送エレメントを含む。

【0171】

簡潔には、末端繰返し配列（「LTR」）は、U5、RおよびU3と命名された3つのエレメントに分けられる。これらのエレメントは、レトロウイルスの生物学的な活性に寄与する種々のシグナルを含み、これには、例えば、U3内に配置されるプロモーターエレメントおよびエンハンサーエレメントが挙げられる。LTRは、このゲノムのいずれかの末端でのそれらの正確な複製に起因して、プロウイルス（組み込まれたDNA形態）において容易に同定され得る。本明細書中で使用される場合、5'LTRは、5'プロモーターエレメントおよび十分なLTR配列を含み、逆転写およびこのベクターのDNA形態の組み込みを可能にすることが理解されるべきである。この3'LTRは、ポリアデニル化シグナル、および十分なLTR配列を含み、逆転写およびこのベクターのDNA形態の組み込みを可能にすることが理解されるべきである。 20 30

【0172】

tRNA結合部位および第2のDNA鎖合成の起点はまた、レトロウイルスが生物学的に活性となるのに重要であり、そして当業者によって容易に同定され得る。例えば、レトロウイルスは、Watson-Crick塩基対によってtRNA結合部位に結合し、そしてウイルス粒子中にこのレトロウイルスゲノムを運搬する。次いでこのtRNAは、逆転写酵素によってDNA合成のためのプライマーとして利用される。このtRNA結合部位は、5'LTRからちょうど下流の位置に基づいて容易に同定され得る。同様に、第2のDNA鎖合成の起点は、その名前から暗示されるように、レトロウイルスの第2のDNA鎖合成のために重要である。この領域（この領域は、ポリプリントラクト（tract）と呼ばれる）は、3'LTRのちょうど上流に位置する。 40

【0173】

5'LTRおよび3'LTR、tRNA結合部位、および第2のDNA鎖合成の起点に加えて、組換えレトロウイルスベクター構築物はまた、パッケージングシグナル、および目的の1つ以上の遺伝子またはコード配列を含み得る。さらに、このレンチウイルスは、核輸送エレメント（好ましい実施形態において、RREではない）を有する。適切な核輸送エレメントの代表的な例としては、ラウス肉腫ウイルスのエレメント（Ogerら、J Virol 70、3834-3843、1996）、ラウス肉腫ウイルスのエレメント（Liu & Mertz、Genes & Dev.、9、1766-1789、1995）およびシミアンレトロウイルスI型のゲノムのエレメント（Zolotukhinら、J Virol. 68、7944-7952、1994）が挙げられる。他の潜在 50

的なエレメントとしては、ヒストン遺伝子のエレメント (K e d e s、A n n u. R e v. B i o c h e m. 48、837-870、1970)、 α -インターフェロン遺伝子 (N a g a t a ら、N a t u r e 287、401-408、1980)、 β -アドレナリン作用性レセプター遺伝子 (K o i l k a ら、N a t u r e 329、75-79、1987)、ならびに c-J u n 遺伝子 (H a t t o r i e ら、P r o c. N a t l. A c a d. S c i. U S A 85、9148-9152、1988) が挙げられる。

【0174】

組換えレンチウイルスベクターは、代表的に、G a g ポリメラーゼコード配列および E n v コード配列の両方を欠失している。組換えレンチウイルスベクターは、代表的に、20 個未満、好ましくは、15 個、より好ましくは、10 個、そして最も好ましくは、8 個の連続したヌクレオチド (G a g ポリメラーゼおよび E n v 遺伝子中に見出される) を含む。本発明の1つの利点は、合成 G a g ポリメラーゼ発現カセット (これは、組換えレトロウイルスベクター構築のためにパッケージング細胞株を構築するために使用され得る) が、野生型 G a g ポリメラーゼ配列に対してほとんど相同性を有さず、従って、合成配列と野生型配列との間の相同組換えの可能性をかなり減少または排除するということである。 10

【0175】

レンチウイルスベクターはまた、目的の1個以上の遺伝子または配列の発現を駆動する組織特異的プロモーターを含み得る。

【0176】

レンチウイルスベクター構築物は、目的の1個以上の遺伝子発現されるように生成され得る。これは、ジシストロンカセットまたはオリゴシストロンカセット (例えば、このコード領域は、80 個以下のヌクレオチドによって分離される (一般的に、L e v i n ら、G e n e 108:167-174、1991を参照のこと)) の使用を介して、または内部リボソーム導入部位 (I R E S) の使用を介して達成され得る。 20

【0177】

上記の組換えレトロウイルスベクター構築物と共に使用するのに適したパッケージング細胞株は、本明細書中で提供される開示から容易に調製され得る。簡潔に、このパッケージング細胞株が由来する親の細胞株は、種々の哺乳動物の細胞株から選択され得、これには、例えば、293、RD、COS-7、CHO、BHK、VERO、HT1080、および骨髄腫細胞が挙げられる。 30

【0178】

パッケージング細胞株の生成のために適切な宿主細胞の選択後、1個以上の発現カセットが、欠失されているベクター成分をトランスで捕捉または供給するために、細胞株中に導入される。

【0179】

適切な発現カセットの代表的な例は、本明細書中に記載されており、そして合成 E n v、合成 G a g、合成 G a g プロテアーゼ、および合成 G a g ポリメラーゼの発現カセットを含み、これらは、プロモーター、ならびに、例えば、G a g ポリメラーゼおよび v p r、v p u、n e f または v i f のうち少なくとも1個をコードする配列を含み、ここで、このプロモーターは、G a g ポリメラーゼおよび v p r、v p u、n e f または v i f に作動可能に連結される。上記のように、ネイティブなコード配列および／または合成コード配列はまた、これらの発現カセットにおいて利用され得る。 40

【0180】

上記の発現カセットを利用することにより、広範な種々のパッケージング細胞株が産生され得る。例えば、一局面において、パッケージング細胞株 (合成 G a g ポリメラーゼをコードする配列、および核輸送エレメントを含む発現カセットを含む) が、提供され、ここで、このプロモーターは、G a g ポリメラーゼをコードする配列に作動可能に連結される。他の局面において、パッケージング細胞株 (プロモーター、および t a t、r e v、E n v、もしくは他の H I V 抗原、またはそれら由来のエピトープをコードする配列を含む) が、提供され、ここで、このプロモーターは、t a t、r e v、E n v、もしくは他の 50

H I V 抗原またはエピトープをコードする配列に作動可能に連結される。さらなる実施形態において、このパッケージング細胞株は、*nef*、*vif*、*vpu*または*vpr*のうちいずれか1個以上をコードする配列を含み得る。例えば、このパッケージング細胞株は、*nef*、*vif*、*vpu*または*vpr*単独、*nef*と*vif*、*nef*と*vpu*、*nef*と*vpr*、*vif*と*vpu*、*vif*と*vpr*、*vpu*と*vpr*、*nef*、*vif*と*vpu*、*nef*、*vif*と*vpr*、*nef*、*vpu*と*vpr*、*vvir*、*vpu*と*vpr*、または*nef*、*vif*、*vpu*と*vpr*の4つ全てのみを含み得る。

【0181】

1つの実施形態において、この発現カセットは、安定的に組み込まれる。別の実施形態において、このパッケージング細胞株は、レンチウイルスベクターの組み込みにより、粒子を生成する。さらなる実施形態において、このプロモーターは、誘導性である。本発明の特定の好ましい実施形態において、このパッケージング細胞株は、レンチウイルスベクターの導入により、複製コンピテントウイルスを含まない粒子を生成する。

【0182】

改変されたコード配列を含む合成カセットは、選択された細胞株にトランスフェクトされる。(i) 代表的に、H I Vコード配列の組み込まれた安定的なコピーを保有し、そして(ii) これらのポリペプチドの受容可能なレベルを発現する、トランスフェクトされた細胞が選択される(発現は、当該分野で公知の方法により評価され得る(例えば、実施例1~4を参照のこと))。この細胞株がV L Pを産生する能力はまた、変化し得る。

【0183】

目的の配列は、上記のように適切なウイルスベクター中に構築される。次いで、この不完全なウイルスは、パッケージング細胞株にトランスフェクトされる。このパッケージング細胞株は、ウイルス用粒子を産生するのに必要なウイルス性機能を提供し、この粒子に目的の配列を含む不完全なウイルスゲノムをパッケージングする。次いで、これらのV L Pが単離され、そして例えば、遺伝子送達または遺伝子治療に使用され得る。

【0184】

さらに、このようなパッケージング細胞株はまた、V L Pのみを産生するために使用され得、このV L Pは、例えば、他の抗原と共に投与するためのアジュバンドとして、またはワクチン組成物に使用され得る。また、パッケージング細胞株中の、ポリペプチド(例えば、抗原)をコードする選択された目的の配列の同時発現はまた、結果として、この選択されたポリペプチドのV L P中への包埋および/またはこの選択されたポリペプチドのV L Pとの結合を生じ得る。

【0185】

本発明の異なる実施形態の種々の形態(例えば、構築物)が、組み合わせられ得る。

【0186】

(2. 4 DNA免疫および遺伝子送達)

種々のH I Vポリペプチド抗原、特にC型H I V抗原は、本発明の粒子中で使用され得る。H I V抗原は、DNA免疫構築物中に含まれ得、この構築物は、例えば、インフレーション中で、ポリペプチド抗原(合成または野生型)のコード配列に融合される合成G a g発現カセットを含み、この構築物の発現により、目的の抗原を提示するV L Pが得られる。

【0187】

本発明の実施に使用される特定の目的のH I V抗原としては、*tat*、*rev*、*nef*、*vif*、*vpu*、*vpr*および他のH I V抗原またはそれら由来のエピトープが挙げられる。これらの抗原は、合成(本明細書中で記載される)または野生型であり得る。さらに、パッケージング細胞株は、以下を含む:*nef*のみ、およびH I V-1(H T L V-1 I I、L A V、A R Vなどとしても知られる)(例えば、*gp120*、*gp41*、*gp160*(ネイティブおよび改変体の両方)が挙げられるが、これらに限定されない); G a g; ならびに種々の単離体由来の*p o l*(多様なサブタイプ由来のH I V_{I I I b}、H I V_{S F 2}、H I V-1_{S F 1 6 2}、H I V-1_{S F 1 7 0}、H I V_{L A V}、H I V_{L A I}、H I V_{M N}、H I V-1_{C M 2 3 5}、H I V-1_{U S 4}、種々のサブタイプ(例えば、

10

20

30

40

50

サブタイプ A～G、および O) 由来の他の HIV-1 株、HIV-2 株および種々のサブタイプ (例えば、HIV-2_{UC1} および HIV-2_{UC2}) が挙げられるが、これらに限定されない)。例えば、Myers ら、Los Alamos Database、Los Alamos National Laboratory、Los Alamos、New Mexico; Myers ら、Human Retroviruses and Aids、1990、Los Alamos、New Mexico; Los Alamos National Laboratory を参照のこと。

【0188】

効力を評価するために、本発明の合成発現カセットを使用して有効な DNA 免疫が実施され得、例えば、実施例 4 に記載されるように実施され得る。マウスは、Gag (および/または Env) 合成発現カセットならびに Gag (および/または Env) 野生型発現カセットの両方で免疫される。プラスミド DNA を用いたマウスの免疫化は、この合成発現カセットがネイティブな発現カセットに対して免疫原性の明瞭な改良を提供することを示す。また、第 2 のブースト免疫は、例えば、約 2 週間後に、第 2 の免疫応答を誘導する。さらに、CTL アッセイの結果は、DNA 免疫化による細胞毒性 T リンパ球 (CTL) 応答の誘導について、合成 Gag (および/または Env) 発現カセットの増強した効力を示す。 10

【0189】

本発明は、広範の種々の抗原に対する免疫応答を固定し、そしてこれにより、HIV 感染、特に C 型 HIV 感染を処置または予防するために使用され得ることが容易に明らかである。 20

【0190】

(2. 4. 1 本発明の合成発現カセットの送達)

上記の分子をコードするポリヌクレオチド配列は、組換え方法を使用して、例えば、この遺伝子を発現する細胞からの cDNA およびゲノムライブラリーをスクリーニングすることによって、または遺伝子を含むことが知られているベクターから遺伝子を誘導することによって、得られ得る。さらに、所望の遺伝子は、標準的な技術 (例えば、cDNA またはゲノム DNA のフェノール抽出および PCR) を使用して、この遺伝子を含む細胞および組織から直接的に単離され得ることが容易に理解される。DNA を得るためおよび単離するために使用される技術の記載について、例えば、Sambrook ら、前出を参照のこと。 30
目的の遺伝子はまた、クローン化ではなく、合成により産生され得る。このヌクレオチド配列は、所望される特定のアミノ酸配列について、適切なコドンで設計され得る。一般に、配列が発現される、意図される宿主に対して、好ましいコドンを選択する。完全な配列は、標準的な方法によって調製されるオーバーラップオリゴヌクレオチドから構築され、そして完全なコード配列へと構築される。例えば、Edge、Nature (1981) 292:756; Nambair ら、Science (1984) 223:1299; Jay ら、J. Biol. Chem. (1984) 259:6311; Stemmer、W. P. C.、(1995) Gene 164:49~53 を参照のこと。

【0191】

次に、所望の抗原をコードする遺伝子配列は、本発明の合成発現カセットを含むベクターに挿入され得る。特定の実施形態において、抗原は、Gag コード配列に挿入され、その結果、複合配列が発現される場合に、Gag ポリペプチドおよび目的の抗原 (例えば、Env (ネイティブまたは改変体) または HIV 由来の他の抗原 (ネイティブまたは改変体)) を含む VLP の産物を生じる。挿入は、コード配列内またはコード配列のいずれかの末端 (発現された Gag ポリペプチドの 5' アミノ末端; または発現された Gag ポリペプチドの 3' カルボキシル末端) で作製され得る (Wagner, R ら、Arch Virol. 127:117-137、1992; Wagner, R. ら、Virology 200:162-175、1994; Wu, X. ら、J. Virol. 69 (6): 3389-3398、1995; Wang, C-T. ら、Virology 200:524-534、1994; Chazal, N. ら、Virology 68 (1): 111 40 50

－122、1994；Griffiths, J. C. ら、J. Virol. 67 (6) : 3191－3198、1993；Reicin, A. S. ら、J. Virol. 69 (2) : 642－650、1995)。

【0192】

p55 Gagのコード配列の50%までは、ウイルス様粒子への構築および発現効率に影響を与えることなく欠失され得る (Borsetti, A. ら、J. Virol. 72 (11) : 9313－9317、1998；Gamier, L. ら、J. Virol. 72 (6) : 4667－4677、1998；Zhang, Y. ら、J. Virol. 72 (3) : 1782－1789、1998；Wang, C. ら、J. Virol. 72 (10) : 7950－7959、1998)。本発明の一実施形態において、合成Gag発現カセットを発現する高いレベルの免疫原性は、異なる構造または非構造のHIV抗原、マルチエピトープカセット、またはサイトカイン配列をGag配列の欠失領域に挿入することによって増加され得る。このような欠失は、本発明の技術および当業者に利用可能な情報に従って作製され得る。異種ポリペプチドに融合された全長配列を使用するのに対する、本アプローチの1つの可能な利点は、発現産物のより高い発現効率／分泌効率であり得る。

10

【0193】

配列が、Gagのアミノ末端に付加される場合、ポリヌクレオチドは、ミリスチン部分のGag含有ポリペプチドへの付加のためのシグナルをコードする、5'末端のコード配列 (例えば、Met-Glyをコードする配列) を含み得る。

20

【0194】

Gag含有ポリペプチド構築物がVLPを形成する能力は、本明細書の教示に従って実験的に決定され得る。

【0195】

合成発現カセットはまた、コード配列に作動可能に連結された制御エレメントを含み得、これは、目的の種においてインビボで遺伝子の発現を可能にする。例えば、哺乳動物細胞発現の代表的なプロモーターとしては、特に、SV40初期プロモーター、CMVプロモーター (例えば、CMV極初期プロモーター)、マウス乳腺ガンウイルスLTRプロモーター、アデノウイルス主要後期プロモーター (Ad MLP)、および単純疱疹ウイルスプロモーターが挙げられる。他の非ウイルスプロモーター (例えば、マウスメタロチオネイン (metallothionein) 遺伝子) に由来するプロモーター) はまた、哺乳動物発現についての使用を見出す。代表的には、転写終止およびポリアデニル化配列もまた存在し、翻訳終止コドンの3'側に位置する。好ましくは、翻訳の開始の最適化のための配列 (コード配列の5'側に配置する) もまた存在する。転写ターミネーター／ポリアデニル化シグナルの例としては、上記Sambrookらに記載される、SV40由来のもの、ならびにウシ成長ホルモントーミネーター配列が挙げられる。

30

【0196】

エンハンサーエレメントはまた、哺乳動物構築物の発現レベルを増加するために本明細書中で使用され得る。例としては、SV40初期遺伝子エンハンサー (Dijkemaら、EMBO J. (1985) 4 : 761に記載される)、Rous Sarcomaウイルスの長い末端反復 (LTR) から誘導されるエンハンサー／プロモーター (Gormanら、Proc. Natl. Acad. Sci. USA (1982b) 79 : 6777に記載される) およびヒトCMVに由来するエレメント (Boshartら、Cell (1985) 41 : 521に記載される) (例えば、CMVイントロンA配列に含まれるエレメント) が挙げられる。

40

【0197】

さらに、キメラ抗原－コード遺伝子配列を含み、例えば、1つより多くのウイルス単離体に由来する、例えば、目的の複数抗原／エピトープをコードする、プラスミドが構築され得る。

【0198】

50

代表的には、抗原コード配列は、合成コード配列より先にあるか、合成コード配列に続き、そしてキメラ転写ユニットは、目的の抗原および合成コード配列の両方をコードする単一のオープンリーディングフレームを有する。あるいは、複数シストロンカセット（例えば、二シストロンカセット）が構築され得、E M C V I R E Sなどを使用して、単一のmRNAから複数の抗原の発現を可能にする。

【0199】

一旦完了すると、構築物は、標準的な遺伝子送達プロトコルを使用して、核酸免疫化のために使用される。遺伝子送達のための方法は、当該分野において公知である。例えば、米国特許第5,399,346号、同第5,580,859号、同第5,589,466号を参照のこと。遺伝子は、直接、脊椎動物被験体に送達されるか、あるいは、被験体由来の細胞に、エキソビボに送達され、そして細胞が被験体に再移植され得るかのいずれかであり得る。

10

【0200】

多数のウイルスに基づく系が、哺乳動物細胞への遺伝子移入のために開発された。例えば、レトロウイルスは、遺伝子送達系のための従来のプラットフォームを提供する。選択された配列は、ベクターに挿入され得、そして当該分野で公知の技術を使用してレトロウイルス粒子にパッケージングされ得る。次いで、組換えウイルスは、単離され、そしてインビボまたはエキソビボのいずれかで、被検体の細胞に送達される。多数のレトロウイルス系が記載されている（米国特許第5,219,740号；MillerおよびRosman, BioTechniques (1989) 7:980-990；Miller, A. D., Human Gene Therapy (1990) 1:5-14；Scarpaら、Virology (1991) 180:849-852；Burnsら、Proc. Natl. Acad. Sci. USA (1993) 90:8033-8037；ならびにBoris-LawrieおよびTemin, Cur. Opin. Genet. Develop. (1993) 3:102-109）。

20

【0201】

多数のアデノウイルスベクターがまた記載されている。宿主ゲノムに組み込まれるレトロウイルスとは異なり、アデノウイルスは、挿入変異に関連する危険を染色体外的に最小化し続ける（Haj-AhmadおよびGraham, J. Virol. (1986) 57:267-274；Bettrら、J. Virol. (1993) 67:5911-5921；Mitterederら、Human Gene Therapy (1994) 5:717-729；Sethら、J. Virol. (1994) 68:933-940；Barrら、Gene Therapy (1994) 1:51-58；Berkner, K. L. BioTechniques (1988) 6:616-629；ならびにRichら、Human Gene Therapy (1993) 4:461-476）。

30

【0202】

さらに、種々のアデノ随伴ウイルス(AAV)ベクター系は、遺伝子送達のために開発されている。AAVベクターは、当該分野で周知の技術を使用して容易に構築され得る。例えば、米国特許第5,173,414号および同第5,139,941号；国際公開番号WO 92/01070 (1992年1月23日公開)およびWO 93/03769 (1993年3月4日公開)；Lebkowskiら、Molec. Cell. Biol. (1988) 8:3988-3996；Vincentら、Vaccines 90 (1990) (Cold Spring Harbor Laboratory Press)；Carter, B. J. Current Opinion in Biotechnology (1992) 3:533-539；Muzyczka, N. Current Topics in Microbial, and Immunol. (1992) 158:97-129；Kotin, R. M. Human Gene Therapy (1994) 5:793-801；ShellingおよびSmith, Gene Therapy (1994) 1:165-169；ならびにZhouら、J. Exp. Med. (1994) 179:1867-1875を参照のこと。

40

50

【0203】

本発明のポリヌクレオチドを送達するために有用な別のベクター系は、Small, Jr., P. A. ら (1997年10月14日に発行された、米国特許第5,676,950号)に記載される、腸管投与される組換えポックスウイルスワクチンである。

【0204】

目的の抗原をコードする核酸分子を送達するための用途を見出すさらなるウイルスベクターとしては、ワクシニアウイルスおよびトリポックスウイルスを含む、ポックスファミリーのウイルス由来のベクターである。例として、遺伝子を発現するワクシニア組換え体は、以下のように構築され得る。特定の合成HIVサブタイプCポリペプチドコード配列をコードするDNAは、最初に適切なベクターに挿入され、その結果、そのDNAは、ワクシニアプロモーターおよび隣接ワクシニアDNA配列 (例えば、チミジンキナーゼ (TK) をコードする配列) に隣接する。次いで、このベクターを使用して、細胞をトランスフェクトし、この細胞をワクシニアで同時に感染させる。相同組換えは、ワクシニアプロモーターおよび目的のコード配列をコードする遺伝子を、ウイルスゲノムに挿入するのに役立つ。得られるTK⁻組換え体は、細胞を5-ブromoデオキシウリジンの存在下で培養し、そしてそれに対して抵抗性のウイルスプラークを摘み上げることによって選択され得る。

10

【0205】

あるいは、アビポックスウイルス (例えば、鶏痘ウイルスおよびカナリア痘ウイルス) はまた、遺伝子を送達するために使用され得る。組換えアビポックスウイルス (哺乳動物病原体から免疫原を発現する) は、非トリ種に投与される場合、保護免疫を与えることが知られている。アビポックスベクターの使用は、ヒトおよび他の哺乳動物種において特に望ましい。なぜなら、アビポックス属のメンバーは、感受性のトリ種においてのみ産生的に複製し得、従って、哺乳動物細胞に感染しないからである。組換えアビポックスウイルスを産生するための方法は、当該分野において公知であり、そしてワクシニアウイルスの産生に関して上記のように、遺伝子組換えを使用する。例えば、WO 91/12882; WO 89/03429; およびWO 92/03545を参照のこと。

20

【0206】

分子結合ベクター (例えば、Michaelら、J. Biol. Chem. (1993) 268:6866-6869およびWagnerら、Proc. Natl. Acad. Sci. USA (1992) 89:6099-6103に記載されるアデノウイルスキメラベクター) はまた、遺伝子送達のために使用され得る。

30

【0207】

アルファウイルス属のメンバー (例えば、限定しないが、Sindbisウイルス、Semliki Forestウイルス、およびVenezuelan Equine Encephalitisウイルスに由来するベクター) はまた、本発明のポリヌクレオチドを送達するためのウイルスベクターとしての使用を見出す (例えば、発現カセットをコードする合成Gag-ポリヌクレオチド)。本発明の方法の実行に有用なSindbisウイルス由来ベクターの記載について、Dubenskyら、J. Virol. (1996) 70:508-519; ならびに国際公開番号WO 95/07995およびWO 96/17072; ならびに、Dubensky, Jr., T. W. ら、米国特許第5,843,723号 (1998年12月1日発行)、およびDubensky, Jr., T. W., 米国特許第5,789,245号 (1998年8月4日発行) を参照のこと。

40

【0208】

ワクシニアに基づく感染/トランスフェクション系は、宿主細胞中の目的のコード配列の誘導性一過性発現を提供するために便利に使用され得る。この系において、細胞は、最初に、バクテリオファージT7 RNAポリメラーゼをコードするワクシニアウイルス組換え体を用いてインビトロで感染される。このポリメラーゼは、それがT7プロモーターを有するテンプレートのみを転写する点で、すばらしい特異性を示す。感染に続いて、細胞を、T7プロモーターによって駆動される、目的のポリヌクレオチドを用いてトランスフ

50

エクトする。ワクシニアウイルス組換え体から細胞質において発現されるポリメラーゼは、トランスフェクトされたDNAをRNAに転写し、これは、次いで、宿主翻訳機構によってタンパク質に翻訳される。この方法は、大量のRNAおよびその翻訳産物の高いレベルの一過性細胞質産生を提供する。例えば、Elroy-SteinおよびMoss, Proc. Natl. Acad. Sci. USA (1990) 87: 6743-6747; Fuerstら、Proc. Natl. Acad. Sci. USA (1986) 83: 8122-8126を参照のこと。

【0209】

ワクシニアウイルス組換え体またはアビボックスウイルス組換え体を用いた感染、あるいは他のウイルスベクターを使用する遺伝子の送達に対する代替のアプローチとして、宿主細胞への導入に続く高レベルの発現を導く増幅系が使用され得る。具体的には、T7 RNAポリメラーゼについてのコード領域に先行するT7 RNAポリメラーゼプロモーターが操作され得る。このテンプレートから誘導されるRNAの翻訳は、T7 RNAポリメラーゼを産生し、これは、次いで、より多くのテンプレートを転写する。同時に、発現がT7プロモーターの制御下にあるcDNAが存在する。このように、増幅テンプレートRNAの翻訳から生成するT7 RNAポリメラーゼのいくつかは、所望の遺伝子の転写を導く。いくつかのT7 RNAポリメラーゼが、増幅を開始するために必要とされるので、T7 RNAポリメラーゼは、転写反応をプライムするために、テンプレートに沿って細胞に導入され得る。ポリメラーゼは、タンパク質として導入され得るか、またはRNAポリメラーゼをコードするプラスミド上に導入され得る。T7系および細胞を形質転換するためのそれらの使用についてのさらなる考察について、例えば、国際公開番号WO 94/26911; StudierおよびMoffatt, J. Mol. Biol. (1986) 189: 113-130; DengおよびWolff, Gene (1994) 143: 245-249; Gaoら、Biochem. Biophys. Res. Commun. (1994) 200: 1201-1206; GaoおよびHuang, Nuc. Acids Res. (1993) 21: 2867-2872; Chenら、Nuc. Acids Res. (1994) 22: 2114-2120; ならびに米国特許第5, 135, 855号を参照のこと。

【0210】

目的の合成発現カセットはまた、ウイルスベクターなしで送達され得る。例えば、合成発現カセットは、被験体への送達または被験体から誘導される細胞への送達の前に、リボソーム中にパッケージングされ得る。脂質カプセル化は、一般的に核酸に安定に結合し得るか、または核酸を捕捉および維持し得るリボソームを用いて達成される。縮合DNAの脂質調製物に対する比は、さまざまであり得るが、一般的に約1:1 (mg DNA: マイクロモル脂質) であるか、またはより多くの脂質である。核酸の送達のためのキャリアとしてのリボソーム使用の概説については、HugおよびSleight, Biochim. Biophys. Acta. (1991) 1097: 1-17; Straubingerら、Methods of Enzymology (1983) 101: 512-527頁を参照のこと。

【0211】

本発明における使用のためのリボソーム調製物としては、カチオン性（正に荷電した）、アニオン性（負に荷電した）および中性の調製物が挙げられ、カチオン性リボソームが特に好ましい。カチオン性リボソームは、機能的な形態で、プラスミドDNA (Felgnerら、Proc. Natl. Acad. Sci. USA (1987) 84: 7413-7416); mRNA (Maloneら、Proc. Natl. Acad. Sci. USA (1989) 86: 6077-6081); および精製した転写因子 (Debsら、J. Biol. Chem. (1990) 265: 10189-10192) の細胞内送達を媒介することが示されている。

【0212】

カチオン性リボソームは容易に入手可能である。例えば、N [1-2, 3-ジオレイルオ

10

20

30

40

50

キシ) プロピル] -N, N, N-トリエチルアンモニウム (DOTMA) リポソームは、GIBCO BRL、Grand Island、NYからの商標リポフェクチン (Lipofectin) の下で入手可能である (Felgnerら、Proc. Natl. Acad. Sci. USA (1987) 84: 7413-7416もまた参照のこと)。他の市販されている脂質としては、(DDAB/DOPE) およびDOTAP/DOPE (Boerhinger) が挙げられる。他のカチオン性リポソームは、当該分野で周知の技法を使用して容易に利用可能な物質から調製され得る。例えば、DOTAP (1, 2-ビス (オレイルオキシ) -3- (トリメチルアンモニオ) プロパン) リポソームの合成の記載について、Szokaら、Proc. Natl. Acad. Sci. USA (1978) 75: 4194-4198; PCT公開番号WO90/11092を参照のこと。

10

【0213】

同様に、アニオン性および中性リポソームは、例えば、Avanti Polar Lipids (Birmingham, AL) から容易に入手可能であるか、または容易に入手可能な物質を使用してたやすく調製され得る。このような物質としては、とりわけ、ホスファチジルコリン、コレステロール、ホスファチジルエタノールアミン、ジオレオイルホスファチジルコリン (DOPC)、ジオレオイルホスファチジルグリセロール (DOPG)、ジオレオイルホスファチジルエタノールアミン (DOPE) が挙げられる。これらの物質はまた、適切な比率のDOTMAおよびDOTAPの出発物質と混合され得る。これらの物質を使用してリポソームを作製する方法は、当該分野で周知である。

20

【0214】

このリポソームは、多重膜のベシクル (MLV)、小さな単一膜ベシクル (SUV)、または大きな単一膜のベシクル (LUV) を含み得る。種々のリポソーム-核酸複合体は当該分野で公知の方法を使用して調製され得る。例えば、Straubingerら、METHODS OF IMMUNOLOGY (1983) 101巻: 512-527頁; Szokaら、Proc. Natl. Acad. Sci. USA (1978) 75: 4194-4198; Papahadjopoulosら、Biochim. Biophys. Acta (1975) 394: 483; Wilsonら、Cell (1979) 17: 77; DeamerおよびBangham、Biochim. Biophys. Acta (1976) 443: 629; Ostroら、Biochem. Biophys. Res. Commun. (1977) 76: 836; Fraleyら、Proc. Natl. Acad. Sci. USA (1979) 76: 3348; EnochおよびStrittmatter、Proc. Natl. Acad. Sci. USA (1979) 76: 145; Fraleyら、J. Biol. Chem. (1980) 255: 10431; SzokaおよびPapahadjopoulos、Proc. Natl. Acad. Sci. USA (1978) 75: 145; ならびにSchaefer-Ridderら、Science (1982) 215: 166を参照のこと。

30

【0215】

DNAおよび/またはタンパク質抗原はまた、Papahadjopoulosら、Biochem. Biophys. Acta. (1975) 394: 483-491によって記載される組成物に類似の蝸牛殻 (cochleate) 脂質組成物で送達され得る。米国特許第4, 663, 161号および同第4, 871, 488号を参照のこと。

40

【0216】

目的の合成発現カセットはまた、粒子状のキャリアでカプセル化され得るか、粒子状のキャリアに吸収され得るか、または粒子状のキャリアに会合 (associate) し得る。このようなキャリアは、選択された抗原の複数コピーを、免疫系に提示し、そして局所的リンパ節における抗原の捕捉および保持を促進する。粒子は、マクロファージによって食作用を受け得、そしてサイトカイン放出を通じた抗原提示を促進し得る。特定のキャリアの例としては、ポリメチルメタクリレートポリマーから誘導されるキャリア、ならびにポリ (ラクチド) およびポリ (ラクチド-co-グリコリド) (PLGとして公知) から誘導される微粒子が挙げられ得る。例えば、Jefferyら、Pharm. Res. (

50

1993) 10:362-368; McGee JPら、J Microencapsul. 14(2) 197-210, 1997; O'Hagan DTら、Vaccine 11(2):149-54, 1993を参照のこと。適切な微粒子はまた、荷電した界面活性剤（例えば、アニオン性界面活性剤またはカチオン性界面活性剤）の存在下で作製されて、正味負の電荷または正味正の電荷を有する表面を有する微粒子を生じ得る。例えば、アニオン性界面活性剤（例えば、ヘキサデシルトリメチルアンモニウムブロミド（CTAB）とともに製造された微粒子（すなわち、CTAB-PLG微粒子）は、負に帯電した高分子（例えば、DNA）を吸収する（例えば、国際出願番号PCT/US99/17308を参照のこと）。

【0217】

さらに、他の粒子系およびポリマーは、目的の遺伝子のインビボまたはエキソビボ送達のために使用され得る。例えば、ポリリジン、ポリアルギニン、ポリオルニチン、スペルミン、スペルミジンのようなポリマー、ならびにこれらの分子の結合体は、目的の核酸を移すために有用である。同様に、DEAEデキストラン媒介トランスフェクション、リン酸カルシウム沈降または他の不溶性無機塩（例えば、リン酸ストロンチウム、ケイ酸アルミニウム（ベントナイトおよびカオリンを含む）、酸化クロム、ケイ酸マグネシウム、滑石など）は、本発明との使用を見出す。例えば、遺伝子移入に有用な送達系の総説について、Felgner, P. L., Advanced Drug Delivery Reviews (1990) 5:163-187を参照のこと。Peptoids (Zuckerman, R. N. ら、1998年11月3日に発行された、米国特許第5,831,005号)もまた、本発明の構築物の送達のために使用され得る。

【0218】

さらに、金およびタングステンのような特定のキャリアを使用する微粒子銃（biolistic）送達系は、本発明の合成発現カセットを送達するために特に有用である。この粒子は、送達される合成発現カセットでコーティングされ、そして「遺伝子銃」からの銃粉末放出を使用して、一般的に減圧大気下で高速に加速される。このような技術および従って有用な装置の説明について、例えば、米国特許第4,945,050号；同第5,036,006号；同第5,100,792号；同第5,179,022号；同第5,371,015号；および同第5,478,744号を参照のこと。また、ニードルレス注入系が使用され得る（Davis, H. L. ら、Vaccine 12:1503-1509, 1994; Bioject, Inc., Portland, OR）。

【0219】

本発明の合成発現カセットを運ぶ組換えベクターは、脊椎動物被験体への送達のための組成物に処方される。これらの組成物は、予防的（感染を予防するため）または治療的（感染の後に疾患を処置するため）のいずれかであり得る。組成物は、「治療的有効量」の目的の遺伝子を含み、その結果、免疫応答が、投与される個体において生成されるように、一定量の抗原が、インビボで産生され得る。必要な実際の量は、とりわけ、処置される被験体；処置される被験体の年齢および一般的状態；抗体を合成する、被験体の免疫系の能力；所望される保護の程度；処置される状態の重篤度；選択される特定の抗原およびその投与の様式に依存して、変動する。適切な有効量は、当業者によって容易に決定され得る。従って、「治療的有効量」は、慣用的な試行によって決定され得る比較的広い範囲である。

【0220】

組成物は、一般的に、1つ以上の「薬学的に受容可能な賦形剤またはビヒクル」（例えば、水、生理食塩水、グリセロール、ポリエチレングリコール、ヒアルロン酸、エタノールなど）を含む。さらに、補助物質（例えば、湿潤剤または乳化剤、pH緩衝化物質など）は、このようなビヒクル中に存在し得る。核酸取り込みおよび／または発現の特定の促進因子（facilitator）はまた、組成物中に含まれ得るか、または例えば、限定しないが、ブピバカイン、心臓毒およびスクロースと同時投与され得る。

【0221】

10

20

30

40

50

一旦処方されると、本発明の組成物は、被験体に直接投与され得る（例えば、上記のように）か、あるいは上記のような方法を使用して、被験体から誘導される細胞にエキソビボで送達され得る。例えば、エキソビボ送達および形質転換された細胞の被験体への再移植のための方法は、当該分野において公知であり、そして、これらの方法としては、例えば、デキストラン媒介トランスフェクション、リン酸カルシウム沈降、ポリブレン媒介トランスフェクション、リポフェクトアミンおよびL T-1媒介トランスフェクション、プロトプラスト融合、エレクトロポレーション、リポソーム中でのポリヌクレオチドのカプセル化（対応する抗原を含むかまたは含まない）、および核へのDNAの直接的な微量注入が挙げられ得る。

【0222】

合成発現カセット組成物のインビボでの直接的な送達は、一般的に、Accell（登録商標）遺伝子送達系（Powderject Technologies, Inc., Oxford, England）のような、従来のシリンジまたは遺伝子銃のいずれかを使用する注入によって、上記のように、ウイルスベクターを用いてまたは用いないで、達成される。構築物は、経鼻的、直腸的、および腔的、腹腔内的、静脈内的、経口的、または筋肉内的のような、皮下的、上皮的、皮内的、粘膜内的のいずれかで、注入され得る。表皮の細胞へのDNAの送達は、特に好ましい。なぜなら、この投与の様式は、皮膚関連リンパ球へのアクセスを提供し、そしてレシピエントにおけるDNAの一過性の存在を提供するからである。投与の他の様式としては、経口投与または肺投与、坐剤、ニードルレス注入、経皮的（transcutaneous）および経皮的（transdermal）適用が挙げられる。投薬処置は、単回用量スケジュールまたは複数用量スケジュールであり得る。核酸の投与はまた、ペプチドまたは他の物質の投与と組み合わせられ得る。

【0223】

（2.4.2 本発明の合成発現カセットのエキソビボ送達）

1つの実施形態において、T細胞、および関連する細胞型（限定しないが、抗原提示細胞（例えば、マクロファージ、単球、リンパ球、樹状細胞、B細胞、T細胞、幹細胞およびそれらの前駆細胞）が挙げられる）は、本発明の合成発現カセットのエキソビボ送達のために使用され得る。T細胞は、当業者に公知の種々の手順によって、末梢血リンパ球（PBL）から単離され得る。例えば、T細胞集団は、補助細胞およびB細胞の除去によってPBLの集団から「濃縮」され得る。特に、T細胞濃縮は、抗MHCI Iモノクローナル抗体を使用して、非T細胞の除去によって達成され得る。同様に、他の抗体は、非T細胞の特定の集団を枯渇させるために使用され得る。例えば、抗Ig抗体分子は、B細胞を枯渇させるために使用され得、そして抗MacI抗体分子は、マクロファージを枯渇させるために使用され得る。

【0224】

T細胞は、当業者に公知の技術によって、多くの異なる亜集団にさらに分別され得る。二つの大きな亜集団は、細胞表面マーカーCD4およびCD8の差示的発現に基づいて、単離され得る。例えば、上記のようなT細胞の濃縮に続いて、CD4⁺細胞は、CD4に特異的な抗体を使用して濃縮され得る（Coliganら、上記を参照のこと）。抗体は、磁気ビーズのような固体支持体に連結され得る。逆に、CD8⁺細胞は、CD4に特異的な抗体の使用（CD4⁺細胞を除くため）によって濃縮され得るか、または固体支持体に連結したCD8抗体の使用によって単離され得る。HIV-1感染患者由来のCD4リンパ球は、Wilsonら、（1995）J. Infect. Dis. 172: 88によって記載されるように、形質導入の前または後に、エキソビボで増殖され得る。

【0225】

T細胞の精製後、当業者に公知の遺伝子改変の種々の方法が、本明細書中に記載されるようにして構築された非ウイルスベースまたはウイルスベースの遺伝子移入ベクターを用いて実施され得る。例えば、1つのこのようなアプローチは、ベクター産生細胞に由来する培養物のベクター含有上清による精製T細胞集団の形質導入を含む。第2のアプローチは、照射した単層のベクター産生細胞と精製T細胞との共培養を含む。第3のアプローチは

10

20

30

40

50

、類似する共培養アプローチであるが；精製T細胞が種々のサイトカインで前刺激され、そして照射したベクター産生細胞との共培養の前に48時間培養される共培養アプローチを含む。このような形質導入前の前刺激は、効果的な遺伝子移入を増加させる（Noltaら（1992）Exp. Hematol. 20:1065）。これらの培養物を増殖させる刺激もまた、患者への再注入のための増加した細胞集団を提供する。共培養に続いて、T細胞は、ベクター産生細胞単層から回収され、展開され、そして液体窒素中で凍結される。

【0226】

1つ以上の本発明の合成発現カセット（単離されたT細胞への送達のための適切な制御エレメントを伴う）を含む遺伝子移入ベクターは、公知の方法を用いて構築され得る。

10

【0227】

選択マーカーはまた、遺伝子移入ベクターの構築において使用され得る。例えば、遺伝子転移ベクターで形質導入された哺乳動物細胞に、細胞毒性因子に対する耐性を与えるマーカーが、用いられ得る。細胞毒性因子は、ネオマイシン、アミノグリコシド、テトラサイクリン、クロラムフェニコール、スルホンアミド、アクチノマイシン、ネトロプシン、ジスタマイシンA、アントラサイクリン、またはピラジニアミドであり得るがこれらに限定されない。例えば、ネオマイシンホスホトランスフェラーゼIIは、ネオマイシンアナログであるジェネティシン（G418）に対する耐性を与える。

【0228】

T細胞はまた、選択される前に、少なくとも1つの型の増殖因子を含む培地中で維持され得る。特定の細胞型の増殖を維持する種々の増殖因子が当該分野で公知である。このような増殖因子の例は、リンパ球の増殖および活性化を促進するサイトカインマイトジェン（例えば、rIL-2、IL-10、IL-12、およびIL-15）である。特定の型の細胞は、ホルモン（ヒト絨毛性ゴナドトロピン（hCG）およびヒト成長ホルモンを含む）のような他の増殖因子により刺激される。特定の細胞集団について適切な増殖因子を選択することは、当業者により容易に達成される。

20

【0229】

例えば、白血球（例えば、分化した前駆細胞および幹細胞）は、種々の増殖因子により刺激される。より詳細には、活性化T_Hおよび活性化マクロファージにより産生される、IL-3、IL-4、IL-5、IL-6、IL-9、GM-CSF、M-CSF、およびG-CSFは、骨髄幹細胞を刺激し、骨髄幹細胞は、次いで、多能性幹細胞、顆粒球-単球前駆体、好酸球前駆体、好塩基球、巨核球、および赤血球前駆体に分化する。分化は、GM-CSF、IL-3、IL-6、IL-11、およびEPOのような増殖因子により調節される。

30

【0230】

次いで、多能性幹細胞は、リンパ系幹細胞、骨髄間質細胞、T細胞前駆体、B細胞前駆体、胸腺細胞、T_H細胞、T_C細胞、およびB細胞に分化する。この分化は、IL-3、IL-4、IL-6、IL-7、GM-CSF、M-CSF、G-CSF、IL-2、およびIL-5のような増殖因子により調節される。

【0231】

顆粒球-単球前駆体は、単球、マクロファージ、および好中球に分化する。このような分化は、GM-CSF、M-CSF、およびIL-8のような増殖因子により調節される。好酸球前駆体は、好酸球に分化する。このプロセスは、GM-CSFおよびIL-5により調節される。

40

【0232】

好塩基球前駆体のマスト細胞および好塩基球への分化は、GM-CSF、IL-4、およびIL-9により調節される。巨核球は、GM-CSF、EPO、およびIL-6にตอบสนองして血小板を産生する。赤血球前駆細胞は、EPOにตอบสนองして赤血球に分化する。

【0233】

従って、CD3結合因子による活性化の間、T細胞はまた、マイトジェン（例えば、IL

50

ー2のようなサイトカイン)と接触され得る。特に好ましい実施形態において、IL-2は、約50~100 $\mu\text{g}/\text{ml}$ の濃度でT細胞の集団に添加される。CD3結合因子による活性化は、2~4日間で実施され得る。

【0234】

一旦、適切に活性化されると、T細胞は、一般的に、T細胞へのベクターのトランスフェクションを可能にする条件下で、適切な遺伝子移入ベクターと接触することにより改変される。遺伝子改変は、T細胞集団の細胞密度が、約 0.1×10^6 と 5×10^6 との間、好ましくは、約 0.5×10^6 と 2×10^6 との間の場合に実施される。多くの適切なウイルススペースおよび非ウイルススペースの遺伝子移入ベクターが、本明細書中の使用のために記載されている。

10

【0235】

形質導入後、形質導入細胞は、公知の技術を用いて、非形質導入細胞から選択される。例えば、形質導入に使用される遺伝子移入ベクターが、細胞毒性因子に対する耐性を与える選択マーカーを含む場合、この細胞は、適切な細胞毒性因子と接触され得、それによって、非形質導入細胞が、形質導入細胞からネガティブ選択され得る。選択マーカーが細胞表面マーカーの場合、この細胞は、特定の細胞表面マーカーに特異的な結合因子と接触され得、それによって、形質導入細胞が、その集団からポジティブ選択され得る。この選択工程はまた、蛍光活性化細胞選別(FACS)技術を伴う(例えば、FACSが特定の細胞表面マーカーを含む集団から細胞を選択するために用いられる場合)か、またはこの選択工程は、標的細胞の捕捉および/またはバックグラウンドの除去のための取り出し可能な支持体として磁氣的に応答する粒子の使用を伴う。

20

【0236】

より詳細には、形質導入細胞のポジティブ選択は、選択可能な細胞表面マーカーを発現する形質導入細胞を選別および回収するために、FACSセルソーター(例えば、FACS VantageTM Cell Sorter、Becton Dickinson Immunocytometry Systems、San Jose、CA)を用いて実施され得る。形質導入後、この細胞は、特定の細胞表面マーカーに対する蛍光標識された抗体分子を用いて染色される。各々の細胞における結合する抗体の量は、細胞を含む液滴を、セルソーターを通過させることにより測定され得る。染色された細胞を含む液滴に電磁的電荷を与えることにより、形質導入細胞は、他の細胞から分離され得る。次いで、ポジティブ選択細胞は、滅菌回収容器中に収集され得る。これらの細胞選別手順は、例えば、FACS VantageTM Training Manual(特に、セクション3-11~3-28および10-1~10-17を参照のこと)に詳細に記載されている。

30

【0237】

形質導入細胞のポジティブ選択はまた、発現または特定の細胞表面マーカーに基づく細胞の磁氣的分離を用いて実施され得る。このような分離技術において、ポジティブ選択される細胞は、最初に、特定の結合因子(例えば、その細胞表面マーカーと特異的に相互作用する抗体または試薬)と接触される。次いで、この細胞は、(ポジティブ細胞に結合している)特定の結合因子に結合する試薬と結合した取り出し可能な粒子(例えば、磁気応答性粒子)と接触される。次いで、この細胞結合因子-粒子の複合体は、例えば、磁場を用いて非標識細胞から物理的に分離され得る。磁気応答性粒子を用いる場合、標識された細胞は、磁場を用いて容器中に維持され得るが、ネガティブ細胞は除去される。これらおよび同様の分離手順は、当業者に公知である。

40

【0238】

選択された形質導入細胞におけるベクターの発現は、当業者に公知の多くのアッセイにより評価され得る。例えば、ウェスタンブロットまたはノーザン分析は、目的の挿入されたヌクレオチド配列の性質に基づき用いられ得る。一旦、発現が確立され、そして形質転換T細胞が選択された合成発現カセットの存在について試験されると、それらは、末梢血流を介した患者への注入のために準備される。

50

【0239】

本発明は、初代哺乳動物細胞のエキソビボ集団の遺伝子改変のためのキットを含む。このキットは、代表的に、1つ以上の容器に含まれる少なくとも1つの選択マーカーおよび少なくとも1つの合成発現カセットをコードする遺伝子移入ベクター、補助的な試薬、またはハードウェア、およびこのキットの使用のための指示書を含む。

【0240】

(2. 4. 3 さらなる送達レジーム)

本明細書中に記載される(上記の任意の方法により送達される)任意のポリヌクレオチド(例えば、発現カセット)またはポリペプチドはまた、他のDNA送達系および/またはタンパク質送達系と組み合わせて使用され得る。非限定的な例として、例えば、1つ以上の分子が「プライミング工程」において送達され、その後1つ以上の分子が「ブースト工程」において送達される、プライム-ブースト法における、これらの分子の同時投与が挙げられる。特定の実施形態において、1つ以上の核酸含有組成物が送達され、その後、1つ以上の核酸含有組成物および/または1つ以上のポリペプチド含有組成物(例えば、HIV抗原を含むポリペプチド)が送達される。他の実施形態において、複数の核酸が(同じかまたは異なる核酸分子を)「プライム」し、その後、複数のポリペプチドが(同じかまたは異なるポリペプチドを)「ブースト」する。他の実施例として、複数の核酸の投与および複数のポリペプチドの投与が挙げられる。

【0241】

同時投与を含む任意の方法において、種々の組成物は、任意の順序で送達され得る。従って、複数の異なる組成物または分子の送達を含む実施形態において、これらの核酸は、ポリペプチドよりも前に全て送達される必要があるわけではない。例えば、プライミング工程は、1つ以上のポリペプチドの送達を含み得、そしてブースト工程は、1つ以上の核酸および/または1つ以上のポリペプチドの送達を含み得る。複数のポリペプチドが投与され、その後、複数の核酸が投与され得るか、またはポリペプチドおよび核酸の投与が任意の順序で実施され得る。本明細書中に記載される任意の実施形態において、核酸分子は、ポリペプチドの全てまたは一部分をコードし得るか、あるいは全くコードし得ない。従って、本明細書中に記載される1つ以上の核酸分子(例えば、発現カセット)および/または本明細書中に記載される1つ以上のポリペプチドは、任意の順序で、かつ任意の投与経路で同時投与され得る。従って、本明細書中に記載されるポリヌクレオチドおよび/またはポリペプチドの任意の組み合わせが、免疫反応を誘発するために使用され得る。

【0242】

(実験)

本発明を実施するための特定の実施形態の例を以下に示す。これらの実施例は、例示の目的のみのために提供され、いかなる手段によっても本発明の範囲を限定することを意図しない。

【0243】

使用される数値(例えば、量、温度など)に関する正確性を確証するための努力が成されているが、ある程度の実験的誤差および偏差が、当然、許容されるべきである。

【0244】

(実施例1)

(合成発現カセットの作製)

(A. HIV-1 Env、Gag、Pol核酸コード配列の改変)

Polコード配列を、C型株AF110975から選択した。Gagコード配列を、C型株AF110965およびAF110967から選択した。Envコード配列を、C型株AF110968およびAF110975から選択した。これらの配列を操作し、それらの遺伝子産物の発現を最大化した。

【0245】

第1に、HIV-1のコドン使用パターンを、得られる核酸コード配列が高度に発現されるヒト遺伝子において見出されるコドン使用に相当するように改変した。HIVコドン使

10

20

30

40

50

用は、高含有量のヌクレオチドAまたはTのコントリプレットを反映する。H I V-1 コドン使用の効果は、D N A 配列における高いA T 含有量であり、これにより、m R N A の翻訳能力が減少し、不安定になる。対照的に、高度に発現するヒトコドンは、ヌクレオチドGまたはCを好む。このコード領域を、高度に発現するヒト遺伝子において見出されるコドン使用に相当するように改変した。

【0246】

第2に、G a g のコード配列およびG a g プロテアーゼコード配列中に位置する阻害性（または不安定）エレメント（I N S）が存在する（S c h n e i d e r R ら、J. V i r o l. 71（7）：4892-4903、1997）。R R E は、H I V によりコードされるR e v タンパク質と相互作用し、I N S の発現ダウンレギュレーション効果に打ち勝つR N A 二次構造である。R R E およびR e v の翻訳後活性化機構に打ち勝つために、不安定エレメントは、コードされるタンパク質のリーディングフレームを変化しない複数の点変異を導入することにより、不活性化される。図5および6（配列番号3、4、20、および21）は、A F 110965株およびA F 110967株から獲得された合成配列におけるいくつかの残存するI N S の位置を示す。これらの配列に成された変化は、図において囲み表示する。図5および6において、上列は、示される株由来のG a g ポリペプチドの改変された配列を示す。囲まれた領域の列の下に示されるヌクレオチドは、I N S をさらに除去するために成された変化を示す。従って、囲み領域に示されるこれらの変化が成される場合、得られる配列は、それぞれ図1および2に示される配列に対応する。

【0247】

合成コード配列は、当該分野で公知の方法、例えば、M i d l a n d C e r t i f i e d R e a g e n t C o m p a n y（M i d l a n d、T e x a s）のような企業により構築される。

【0248】

本発明の1つの実施形態において、P o l ポリペプチドをコードする配列が、合成性の改変されたG a g / E n v 発現カセットにより発現されるウイルス様粒子のエピトープの数を増加させるために、合成G a g またはE n v 配列と共に含まれる。合成H I V-1 P o l は、（構造的タンパク質およびプロテアーゼに加えて）機能性酵素である逆転写酵素（R T）およびインテグラーゼ（I N T）を発現するので、R T およびI N T 機能を不活性化することは、いくつかの例において助けとなり得る。R T およびI N T コード領域におけるいくつかの欠失または変異が、それらのR T およびI N T 活性に関して触媒的に非機能的な酵素を達成するために作製され得る {J a y. A. L e v y（編集者）（1995）The R e t r o v i r i d a e, P l e n u m P r e s s, N e w Y o r k. I S B N 0-306-45033X. 215-20頁；G r i m i s o n, B. およびL a u r e n c e, J.（1995）, J o u r n a l O f A c q u i r e d I m m u n e D e f i c i e n c y S y n d r o m e s a n d H u m a n R e t r o v i r o l o g y 9（1）：58-68；W a k e f i e l d, J. K. ら、（1992）J o u r n a l O f V i r o l o g y 66（11）：6806-6812；E s n o u f, R. ら、（1995）N a t u r e S t r u c t u r a l B i o l o g y 2（4）：303-308；M a i g n a n, S. ら、（1998）J o u r n a l O f M o l e c u l a r B i o l o g y 282（2）：359-368；K a t z, R. A. およびS k a l k a, A. M.（1994）A n n u a l R e v i e w O f B i o c h e m i s t r y 73（1994）；J a c o b o - M o l i n a, A. ら、（1993）P r o c e e d i n g s O f t h e N a t i o n a l A c a d e m y O f S c i e n c e s O f t h e U n i t e d S t a t e s O f A m e r i c a 90（13）：6320-6324；H i c k m a n, A. B. ら、（1994）J o u r n a l O f B i o l o g i c a l C h e m i s t r y 269（46）：29279-29287；G o l d g u r, Y. ら、（1998）P r o c e e d i n g s O f t h e N a t i o n a l A c a d e m y O f S c i e n c e s O f t h e U n i t e d S t a t e s O f A m e r i c a 95

(16) : 9150-9154 ; Goette, M. ら, (1998) Journal Of Biological Chemistry 273 (17) : 10139-10146 ; Gorton, J. L. ら, (1998) Journal of Virology 72 (6) : 5046-5055 ; Engelmann, A. ら, (1997) Journal Of Virology 71 (5) : 3507-3514 ; Dyda, F. ら, Science 266 (5193) : 1981-1986 ; Davies, J. F. ら, (1991) Science 252 (5002) : 88-95 ; Bujacz, G. ら, (1996) Febs Letters 398 (2-3) : 175-178 ; Beard, W. A. ら, (1996) Journal Of Biological Chemistry 271 (21) : 12213-12220 ; Kohlstaedt, L. A. ら, (1992) Science 256 (5065) : 1783-1790 ; Krug, M. S. および Berger, S. L. (1991) Biochemistry 30 (44) : 10614-10623 ; Mazumder, A. ら, (1996) Molecular Pharmacology 49 (4) : 621-628 ; Palaniappan, C. ら, (1997) Journal Of Biological Chemistry 272 (17) : 11157-11164 ; Rodgers, D. W. ら, (1995) Proceedings Of the National Academy Of Sciences Of the United States Of America 92 (4) : 1222-1226 ; Sheng, N. および Dennis, D. (1993) Biochemistry 32 (18) : 4938-4942 ; Spence, R. A. ら, (1995) Science 267 (5200) : 988-993}。

【0249】

さらに、選択されたB細胞エピトープおよび／またはT細胞エピトープは、RTおよびINTの機能的改変により欠失される任意のエピトープを置換および増強するために、Pol構築物に加えられる得る（例えば、短縮INTの3'末端またはRTおよびINTコード配列の欠失物内に）。あるいは、RTおよびINTからの選択されたB細胞エピトープおよびT細胞エピトープ（CTLエピトープを含む）は、上記の合成Gagカセットまたは合成Polカセットの発現により形成される最小VLPに含まれる得る（既知のHIB B細胞エピトープおよびT細胞エピトープの説明については、HIV Molecular Immunology Database CTL Search Interface ; Los Alamos Sequence Compendia、1987-1997 ; インターネットアドレス : <http://hiv-web.lanl.gov/immunology/index.html> を参照のこと）。

【0250】

得られる改変コード領域配列を、合成Env発現カセット ; 合成Gag発現カセット ; 合成Pol発現カセットとして示す。共通のGag領域 (Gag-common) は、AF110965に対してヌクレオチド844位～903位 (配列番号1) (または配列番号17のアミノ酸残基約282位～301位) に延び、そしてAF110967に対してヌクレオチド841位～900位 (配列番号2) (または配列番号22のアミノ酸残基約281位～300位) に延びる。共通のEnv領域 (Env-common) は、AF110968に対してヌクレオチド1213位～1353位 (配列番号5) および配列番号23のアミノ酸405位～451位に延び、そしてAF110975に対してヌクレオチド1210位～1353位 (配列番号11) およびアミノ酸404位～451位 (配列番号24) に延びる。

【0251】

Pol、Gag、およびEnvの合成DNAフラグメントは、以下の真核生物発現ベクター中にクローニングされる : 一過性発現アッセイおよびDNA免疫研究のためのpCMV Km2、pCMV Km2ベクターは、pCMV6a (Chapmanら、Nuc, Acids. Res. (1991) 19 : 3979-3986) に由来し、そしてカナマイシン

選択マーカー、C o l E 1 複製起点、C M V プロモーターエンハンサー、およびイントロンAを含み、その後ろに以下に記載される合成配列のための挿入部位を含み、その後ろにウシ成長ホルモン由来のポリアデニル化シグナルを含むー p C M V K m 2 ベクターは、p C M V - l i n k ベクターと、ポリリンカー部位が p C M V K m 2 に挿入されて p C M V - l i n k が作製されるという点でのみ異なる；チャイニーズハムスター卵巣（C H O）細胞における発現のための p E S N 2 d h f r および p C M V P L E d h f r；および p A c C 1 3、バキュロウイルス発現系において使用するためのシャトルベクター（p A c C 1 3 は、Munemitsu S. ら、Mol Cell Biol. 10 (11) : 5977-5982、1990に記載される p A c C 1 2 に由来する）。

【0252】

10

簡単に述べると、p C M V P L E d h f r の構築は、以下の通りである。

【0253】

DHFR カセットを構築するために、EMCV IRES（内部リボソーム侵入部位）リーダーを、p C i t e - 4 a + (Novagen, Inc., Milwaukee, WI) から PCR 増幅し、そして X b a - N c o フラグメントとして p E T 2 3 d (Novagen, Inc., Milwaukee, WI) に挿入し、p E T - E M C V を獲得した。d h f r 遺伝子を、p E S N 2 d h r f から PCR 増幅して翻訳終止コドンの代わりに G l y - G l y - G l y - S e r スペーサーを有する産物を獲得し、そして N c o - B a m H 1 フラグメントとして挿入し、p E T - E - D H F R を獲得した。次に、希釈した n e o 遺伝子を p S V 2 N e o (Clontech, Palo Alto, CA) 誘導体から PCR 増幅し、そして p E T - E - D H F R の特有の B a m H 1 部位に挿入し、p E T - E - D H F R / N e o (m 2) を獲得した。最後に、p C D N A 3 (Invitrogen, Inc., Carlsbad, CA) 由来のウシ成長ホルモントーミネーターを、n e o 遺伝子の下流に挿入し、p E T - E - D H F R / N e o (m 2) B G H t を獲得した。EMCV - d h f r / n e o 選択マーカーカセットフラグメントを、p E T - E - D H F R / N e o (m 2) B G H t の切断により調製した。

20

【0254】

CMV エンハンサー／プロモーターおよびイントロンAを、p C M V 6 a (Chapman ら、Nuc. Acids. Res. (1991) 19 : 3979-3986) から H i n d I I I - S a l I フラグメントとして p U C 1 9 (New England Bio 30 labs, Inc., Beverly, MA) に移した。N d e I 部位から S a l I 部位まで、p U C 1 9 のベクター骨格を欠失させた。上記の DHFR カセットを、EMCV IRES が、CMV プロモーターの後ろにくるように、この構築物に加えた。このベクターはまた、amp^r 遺伝子および S V 4 0 複製起点を含んだ。

【0255】

(B. H I V - 1 p 5 5 G a g の主要相同性領域 (MHR) の規定)
H I V - 1 p 5 5 (G a g) の主要相同性領域 (MHR) は、G a g の p 2 4 - C A 配列中に位置する。それは、およそ 20 アミノ酸の保存されたストレッチである。野生型 A F 1 1 0 9 6 5 G a g タンパク質における位置は、282～301であり（配列番号 25）、そして G a g D N A 配列については、844～903の領域にまたがる（配列番号 26）。合成 G a g タンパク質における位置もまた、282～301であり（配列番号 25）、そして合成 G a g D N A 配列については、844～903の領域にまたがる（配列番号 1）。野生型および合成 A F 1 1 0 9 6 7 G a g タンパク質における位置は、281～300であり（配列番号 27）、そして改変された G a g D N A 配列については、841～900の領域にまたがる（配列番号 2）。MHR における変異または欠失は、粒子の産生を著しく減少させ得る (Borsetti ら、J. Virol. 72 (11) : 9313-9317、1988; Mammano, F. ら、J. Virol. 68 (8) : 4927-4936、1994)。

40

【0256】

この配列に対するパーセント同一性は、例えば、S m i t h - W a t e r m a n サーチア 50

ルゴリズム (Time Logic、Incline Village、NV) を以下の例示的なパラメータと共に用いて決定され得る：ウェイトマトリックス = nuc 4 x 4 hb；ギャップオープンペナルティー = 20；ギャップエクステンションペナルティー = 5。

【0257】

(C. HIV-1 Envの共通配列領域の定義)

HIV-1 Envの共通配列領域 (CSR) は、EnvのC4配列中に位置する。これは、約47アミノ酸の保存されたストレッチである。野生型および合成のAF110968 Envタンパク質のこの位置は、およそアミノ酸残基405～451 (配列番号28) であり、そしてEnv DNA配列に関して1213～1353 (配列番号5) の領域にわたる。野生型およびおおよび合成のAF110975 Envタンパク質のこの位置は、およそアミノ酸残基404～451 (配列番号29) であり、そしてEnv DNA配列に関して1210～1353 (配列番号11) の領域にわたる。

【0258】

この配列に対する%同一性を、例えば、Smith-Waterman検索アルゴリズム (Time Logic、Incline Village、NV) を、以下の例示的なパラメーター：重量マトリックス = nuc 4 x 4 hb；ギャップオープニングペナルティー = 20、ギャップ伸長ペナルティー = 5と共に使用して、決定し得る。

【0259】

本明細書中に記載される本発明の異なる実施形態の種々の形態を、組み合わせ得る。

【0260】

(D. 南アフリカC型HIV系統由来の例示的なHIV配列)

新規C型単離体のHIVコード配列を獲得した。ポリペプチドコード配列を、その遺伝子産物の発現を最大化するために操作した。

【0261】

上で考察したように、得られた核酸コード配列が、高度に発現されたヒト遺伝子において見出されるコドンの用法に匹敵するように、HIV-1コドンの用法パターンを改変した。このHIVコドンの用法は、コドン三つ組のヌクレオチドAまたはTの高い含量を反映する。HIV-1コドンの用法の効果は、減少した翻訳能力およびmRNAの不安定性を生じる、DNA配列中の高いAT含量である。相対的に、高く発現されたヒトのコドンは、ヌクレオチドGまたはCを好む。これらのコード配列を、高く発現されたヒト遺伝子において見出されるコドンの用法に匹敵するように、改変した。

【0262】

以下の表Cに、新規の南アフリカC型HIV単離体 (クローン8__5__TV1__C. ZA.) 由来の例示的な野生型および合成の配列を示す。表Dは、新規の南アフリカC型HIV単離体 (クローン8__2__TV1__C. ZA.) 由来の例示的な合成Env配列を示す。表Eは、南アフリカC型HIV系統 (12-5__1__TV2__C. ZA.) 由来の野生型および合成の配列を示す。

【0263】

【表3】

表 C

名前	配列 番号	説明
C4_Env_TV1_C_ZA_opt short	46	短い Env「共通領域」の合成配列
C4_Env_TV1_C_ZA_opt	47	Env「共通領域」の合成配列
C4_Env_TV1_C_ZA_wt	48	野生型 8_5_TV1_C.ZA Env 配列
Envgp160_TV1_C_ZAopt	49	合成 Env gp160
Envgp160_TV1_C_ZAwt	50	野生型 8_5_TV1_C.ZA Env gp160 配列
Gag_TV1_C_ZAopt	51	Gag の合成配列
Gag_TV1_C_ZAwt	52	野生型 8_5_TV1_C.ZA Gag 配列
Gag_TV1_ZA_MHROpt	53	Gag 主要相同性領域の合成配列
Gag_TV1_ZA_MHRwt	54	野生型 8_5_TV1_C.ZA Gag 主要相同性領域 配列
Nef_TV1_C_ZAopt	55	Nef の合成配列
Nef_TV1_C_ZAwt	56	野生型 8_5_TV1_C.ZA Nef 配列
NefD125G_TV1_C_ZAopt	57	Nef の合成配列, 非機能的遺伝子産物を 生じる、125位での変異を含む
p15RNaseH_TV1_C_ZAopt	58	RNaseH の合成配列 (Pol の p15)
p15RNaseH_TV1_C_ZAwt	59	野生型 8_5_TV1_C.ZA RNaseH 配列
p31Int_TV1_C_ZAopt	60	インテグラーゼの合成配列 (Pol の p31)
p31Int_TV1_C_ZAwt	61	野生型 8_5_TV1_C.ZA インテグラーゼ配列
Pol_TV1_C_ZAopt	62	Pol の合成配列
Pol_TV1_C_ZAwt	63	野生型 8_5_TV1_C.ZA Pol 配列
Prot_TV1_C_ZAopt	64	Prot の合成配列
Prot_TV1_C_ZAwt	65	野生型 8_5_TV1_C.ZA Prot 配列
Protina_TV1_C_ZAopt	66	プロテアーゼの不活化を生じる変異を含む Prot の合成配列

10

20

30

(表C 続き)

Protina_TV1_C_ZAwt	67	野生型 8_5_TV1_C.ZA Prot 配列 , プロテアーゼの不活性化を生じる 変異を含む。
ProtinaRTmut_TV1_C_ZAopt	68	Prot および逆転写酵素(RT)の合成配列 , プロテアーゼの不活化を生じる変異およびR Tの不活化を生じる変異を含む。
ProtinaRTmut_TV1_C_ZAwt	69	野生型 8_5_TV1_C.ZA Prot および RT, プロテアーゼの不活化を生じる変異およびR Tの不活化を生じる変異。
ProtwtRTwt_TV1_C_ZAopt	70	Prot および RT の合成配列
ProtwtRTwt_TV1_C_ZAwt	71	野生型 8_5_TV1_C.ZA Prot および RT
RevExon1_TV1_C_ZAopt	72	Rev のエキソン 1 の合成配列
RevExon1_TV1_C_ZAwt	73	Rev のエキソン 1 の野生型 8_5_TV1_C.ZA
RevExon2_TV1_C_ZAopt-2	74	Rev のエキソン 2 の合成配列
RevExon2_TV1_C_ZAwt	75	Rev のエキソン 2 の野生型 8_5_TV1_C.ZA
RT_TV1_C_ZAopt	76	RT の合成配列
RT_TV1_C_ZAwt	77	野生型 8_5_TV1_C.ZA RT
RTmut_TV1_C_ZAopt	78	RT の合成配列 , R T の不活化を生じる 変異を含む。
RTmut_TV1_C_ZAwt	79	野生型 8_5_TV1_C.ZA RT, R T の不活化を生じる 変異を含む。
TatC22Exon1_TV1_C_ZAopt	80	Tat のエキソン 1 の合成配列, 非機能的 T a t 遺伝子産物を生じる変異を 含む。
TatExon1_TV1_C_ZAopt	81	Tat のエキソン 1 の合成配列
TatExon1_TV1_C_ZAwt	82	Tat のエキソン 1 の野生型 8_5_TV1_C.ZA
TatExon2_TV1_C_ZAopt	83	Tat のエキソン 2 の合成配列
TatExon2_TV1_C_ZAwt	84	Tat のエキソン 2 の野生型 8_5_TV1_C.ZA
Vif_TV1_C_ZAopt	85	Vif の合成配列
Vif_TV1_C_ZAwt	86	野生型 8_5_TV1_C.ZA Vif
Vpr_TV1_C_ZAopt	87	Vpr の合成配列

10

20

30

40

(表C 続き)

Vpr_TV1_C_ZAwt	88	野生型 8_5_TV1_C.ZA Vpr
Vpu_TV1_C_ZAopt	89	Vpu の合成配列
Vpu_TV1_C_ZAwt	90	野生型 8_5_TV1_C.ZA Vpu
revexon1_2 TV1 C ZAopt	91	Rev のエキソン 1 および 2 の合成配列
RevExon1_2_TV1_C_ZAwt	92	野生型 8_5_TV1_C.ZA Rev (エキソン1および2)
TatC22Exon1_2_TV1_C_ZAopt	93	Tat のエキソン 1 および 2 の合成配列, 非機能的な T a t 遺伝子産物を生じる、エキソン 1 における変異を含む。
TatExon1_2_TV1_C_ZAopt	94	Tat のエキソン 1 および 2 の合成配列
TatExon1_2_TV1_C_ZAwt	95	野生型 8_5_TV1_C.ZA Tat (エキソン1および2)
NefD125G- Myr_TV1_C_ZAopt	96	Nef の合成配列, ミリスティル化部位を排除する 変異を含む。

10

【 0 2 6 4 】

【 表 4 】

20

表 D

名前	配列番号	説明
gp120mod.TV1.delV2	119	Env gp120 の合成配列, V2 欠失および野生型 8_2_TV1_C.ZA 配列に由来する改変されたリーダー配列を含む。
gp140mod.TV1.delV2	120	Env gp140 の合成配列, V2 欠失および野生型 8_2_TV1_C.ZA 配列に由来する改変されたリーダー配列を含む。
gp140mod.TV1.mut7.delV2	121	Env gp140 の合成配列, V2 欠失ならびに切断部位における変異および野生型 8_2_TV1_C.ZA 配列に由来する改変されたリーダー配列を含む。
gp160mod.TV1.delV1V2	122	Env gp160 の合成配列, V1/V2 欠失および野生型 8_2_TV1_C.ZA 配列に由来する改変されたリーダー配列を含む。
gp160mod.TV1.delV2	123	Env gp160 の合成配列, V2 欠失および野生型 8_2_TV1_C.ZA 配列に由来する改変されたリーダー配列を含む。
gp160mod.TV1.mut7.delV2	124	Env gp160 の合成配列, V2 欠失; 切断部位における変異および野生型 8_2_TV1_C.ZA 配列に由来する改変されたリーダー配列を含む。
gp160mod.TV1.tpa1	125	Env gp160 の合成配列, TPA1リーダー
gp160mod.TV1	126	Env gp160 の合成配列, 野生型(8_2_TV1_C.ZA)配列に由来する改変されたリーダー配列を含む。
gp160mod.TV1.wtlnative	127	Env gp160 の合成配列, 野生型 8_2_TV1_C.ZA (未改変)リーダーを含む。
gp140mod.TV1.tpa1	131	Env gp140 の合成配列, TPA1リーダー
gp140mod.TV1	132	Env gp140 の合成配列, 野生型 8_2_TV1_C.ZA 配列に由来する改変されたリーダー配列を含む。
gp140mod.TV1.wtlnative	133	Env gp120 の合成配列, 野生型 8_2_TV1_C.ZA (未改変)リーダー配列を含む。

上記のように、Envコード構築物を、全長の gp160 構築物のいずれかを使用して、調製し得る。例えば、gp140 形態（配列番号 132）を、gp160（配列番号 126）を、ヌクレオチド 2064 で短縮することによって作製し；gp120 を、gp160（配列番号 126）を、ヌクレオチド 1551 で短縮することによって作製した（配列番号 126）。さらなる gp140 形態および gp120 形態を、本明細書中に記載の方法を使用して、作製し得る。1 以上の終止コドン、代表的に付加する（例えば、配列番号 126 のヌクレオチド 2608～2610）。さらに、野生型のリーダー配列を、改変および／または他のリーダー配列（例えば、TPA1リーダー配列）で置換し得る。

【0265】

従って、ポリペプチド gp160 は、gp120 および gp41 のコード配列を含む。ポリペプチド gp41 は、オリゴマー化ドメイン（OD）および膜貫通ドメイン（TM）を

10

20

30

40

50

含むいくつかのドメインからなる。ネイティブのエンベロープにおいて、このオリゴマー化ドメインは、三量体構造を形成する3つのgp41ポリペプチドの非共有的な会合に必要である：gp41三量体（およびそれ自体）との非共有的な相互作用によって、gp120ポリペプチドもまた、三量体構造に組織される。切断部位は、ほぼ、gp120のポリペプチド配列とgp41に対応するポリペプチド配列との間に存在する。この切断部位を、この部位での切断を防止するために変異し得る。得られたgp140ポリペプチドは、gp160の短縮形態に対応し、ここで、gp41の膜貫通ドメインは、欠失されている。このgp140ポリペプチドは、gp41部分のオリゴマー化ドメインの存在によって、単量体形態およびオリゴマー（すなわち、三量体）形態の両方で存在し得る。切断を防止するためにこの切断部位が変異されており、そしてgp41の膜貫通部分が欠失されている状況において、得られたポリペプチド産物は、「変異」gp140（例えば、gp140.mut）と称される。当業者に明らかなように、この切断部位は、種々の方法によって変異され得る。本明細書中に記載される例示的な構築物（例えば、配列番号121および配列番号124）において、gp120/gp41切断部位における変異は、野生型のアミノ酸配列KR RVVQREKR（配列番号129）を、ISSVVQSEKS（配列番号130）に変化させる。

10

【0266】

なお別の実施形態において、超可変領域が欠失され、N-グリコシル化部位が除去され、そして／または切断部位が変異された。可変領域の欠失（V1および／またはV2）を有する例示的な構築物において、V2欠失を、（配列番号128に対して）約499～約593のヌクレオチドを欠失することによって構築し、そしてV1/V2欠失を、（配列番号128に対して）約375～約602のヌクレオチドを欠失することによって構築した。V1領域および／またはV2領域の相対的な位置もまた、表Aに示される領域に対するアラインメントによって、容易に決定し得る。表Eは、南アフリカC型HIV系統（12-5__1__TV2__C.ZA.）由来の野生型および合成の配列を示す。

20

【0267】**【表5】**

表 E

名前	配列番号	説明
Envgp160_TV2_C_ZAopt	97	Env gp160 の合成配列
Envgp160_TV2_C_ZAwt	98	野生型 12-5_1_TV2_C.ZA Env gp160.
Gag_TV2_C_ZAopt	99	Gag の合成配列
Gag_TV2_C_ZAwt	100	野生型 12-5_1_TV2_C.ZA Gag
Nef_TV2_C_ZAopt	101	Nef の合成配列
Nef_TV2_C_ZAwt	102	野生型 12-5_1_TV2_C.ZA Nef
Pol_TV2_C_ZAopt	103	Pol の合成配列
Pol_TV2_C_ZAwt	104	Pol の野生型 12-5_1_TV2_C.ZA
RevExon1_TV2_C_ZAopt	105	Rev のエキソン 1 の合成配列
RevExon1_TV2_C_ZAwt	106	Rev のエキソン 1 の野生型 12-5_1_TV2_C.ZA
RevExon2_TV2_C_ZAopt	107	Rev のエキソン 2 の合成配列
RevExon2_TV2_C_ZAwt	108	Rev のエキソン 2 の野生型 12-5_1_TV2_C.ZA
TatExon1_TV2_C_ZAopt	109	Tat のエキソン 1 の合成配列
TatExon1_TV2_C_ZAwt	110	Tat のエキソン 1 の野生型 12-5_1_TV2_C.ZA
TatExon2_TV2_C_ZAopt	111	Tat のエキソン 2 の合成配列
TatExon2_TV2_C_ZAwt	112	Tat のエキソン 2 の野生型 12-5_1_TV2_C.ZA
Vif_TV2_C_ZAopt	113	Vif の合成配列
Vif_TV2_C_ZAwt	114	Vif の野生型 12-5_1_TV2_C.ZA
Vpr_TV2_C_ZAopt	115	Vpr の合成配列
Vpr_TV2_C_ZAwt	116	Vpr の野生型 12-5_1_TV2_C.ZA
Vpu_TV2_C_ZAopt	117	Vpu の合成配列
Vpu_TV2_C_ZAwt	118	Vpu の野生型 12-5_1_TV2_C.ZA

10

20

30

40

50

任意の C 型 H I V 系統由来の配列またはクローンが、この系統における所望の改変を達成するために、本明細書中に記載のように改変され得ることは、容易に明らかである。さらに、ポリタンパク質を、ポリペプチドまたはペプチド産物をコードする 2 つ以上のポリヌクレオチド配列をインフレームで融合させることによって、構築し得る。さらに、ポリシストロン性コード配列を、互いに隣接するポリペプチド産物をコードする、代表的には、1 つのプロモーターの制御下にある、2 つ以上のポリヌクレオチド配列を配置することによって生成し得、ここで、各ポリペプチドコード配列を、内部リボソーム結合部位の配列を含むように改変し得る。

【0268】

本発明の配列（例えば、H I V ポリペプチドをコードする、改変された（合成）ポリヌクレオチド配列）を、欠失、点変異、置換、フレームシフト、および／またはさらなる遺伝的改変（例えば、ポリペプチドに関連する活性の不活化を生じる変異（例えば、プロテアーゼ活性、tat 活性、または逆転写酵素活性を不活化する変異））によって、改変し得る。このような改変は、当該分野において一般に教示され、そして本発明の教示に関して適用され得る。例えば、表 A に列挙された「H I V ゲノム領域」に対応する部位を、所望の改変を達成するために、本明細書中で開示される新規配列の対応する領域において改変

し得る。さらに、本発明の改変された（合成）ポリヌクレオチド配列を、（例えば、被験体において免疫応答を生成するための組成物における）使用のために、以下の方法を含むが、これらに限定されない種々の方法において、組み合わせ得る：各々が本発明の1つのポリヌクレオチド配列を含む複数の個々の発現カセット（例えば、gag発現カセット、env発現カセットおよびrev発現カセット、またはpol発現カセット、vif発現カセットおよびvpr発現カセットなど）；本発明の複数のポリヌクレオチドのインフレーション融合によって生成されたポリタンパク質、ならびに本発明の複数のポリヌクレオチドを使用して生成されたポリシストロン性のポリヌクレオチド。

【0269】

（実施例2）

（合成コード配列についての発現アッセイ）

（A．C型HIVコード配列）

野生型のCサブタイプのHIVコード配列（例えば、AF110965、AF110967、AF110968、AF110975ならびに新規の南アフリカ系統8__5__TV1__C．ZA、8__2__TV1__C．ZAおよび12-5__1__TV2__C．ZA由来）を、合成配列がクローニングされたベクターと同じ特徴を有する発現ベクターにクローニングした。

【0270】

野生型および合成の配列を保有する種々のベクターについての発現効率を、以下のように評価する。いくつかの哺乳動物細胞株（293、RD、COS-7およびCHO；全ては、American Type Culture Collection、10801 University Boulevard、Manassas、VA 20110-2209から獲得した）由来の細胞を、トランスフェクション試薬LT1（PanVera Corporation、545 Science Dr．、Madison、WI）中の2μgのDNAを用いてトランスフェクションする。細胞を、減少血清培地（Opti-MEM、Gibco-BRL、Gaithersburg、MD）中で5時間インキュベートする。次いで、培地を、以下のような正常な培地で置き換える：293細胞、IMDM、10%ウシ胎仔血清、2%グルタミン（BioWhittaker、Walkersville、MD）；RDおよびCOS-7細胞、DMEM、10%ウシ胎仔血清、2%グルタミン（Opti-MEM、Gibco-BRL、Gaithersburg、MD）；およびCHO細胞、Ham's F-12、10%ウシ胎仔血清、2%グルタミン（Opti-MEM、Gibco-BRL、Gaithersburg、MD）。細胞を、48時間かまたは60時間のいずれかでインキュベートする。細胞溶解物を、以下の実施例3に記載のように収集する。上清を回収し、そして0.45μmシリンジフィルターを介して濾過する。上清を、HIV抗原に対するマウスモノクローナル抗体で被覆した96ウェルのプレート（例えば、Coulter p24アッセイ（Coulter Corporation、Hialeah、FL、US）を使用して、評価する。HIV-1抗原は、被覆されたウェルに結合する。HIVに対するビオチン化抗体は、結合した抗原を認識する。結合体化されたストレプトアビジン-西洋ワサビペルオキシダーゼは、ビオチンと反応する。色は、TMB基質とのペルオキシダーゼの反応から発色する。この反応を、4N H₂SO₄の添加によって終了させる。色の強度は、サンプル中のHIV抗原の量と直接比例する。

【0271】

合成C型HIV発現カセットは、種々の細胞株において発現された場合に、ネイティブ（野生型のサブタイプC）の配列と比較して、そのタンパク質産物の生成における劇的な増大を提供する。

【0272】

（B．シグナルペプチドリーダー配列）

種々のリーダー配列が発現を駆動する能力を、異なるリーダー配列に作動可能に連結された、野生型または合成のEnvコード発現カセットを細胞にトランスフェクションし、そ

10

20

30

40

50

してE L I S AまたはウェスタンブロットによってE n vポリペプチドの発現を評価することによって、試験した。種々のシグナルペプチドリーダー配列のアミノ酸配列およびヌクレオチド配列を、表4に示す。

【0273】

【表6】

表4

リーダー	アミノ酸配列	DNA 配列
WTnative (8_2_TV 1_C.ZA)	MRVMGTQKNCQQWWIWGI LGFWMLMIC	ATGAGAGTGATGGGGACACAGA AGAATTGTCAACAATGGTGGATA TGGGGCATCTTAGGCTTCTGGAT GCTAATGATTGT
WTmod (8_2_TV 1_C.ZA)	MRVMGTQKNCQQWWIWGI LGFWMLMIC	ATGCGCGTGATGGGCACCCAGAA GAACTGCCAGCAGTGGTGGATCT GGGGCATCCTGGGCTTCTGGATG CTGATGATCTGC
Tpa1	MDAMKRGLCCVLLCGAVFVSPS AS	ATGGATGCAATGAAGAGAGGGC TCTGCTGTGTGCTGCTGCTGTGTG GAGCAGTCTTCGTTTCGCCCAGC GCCAGC
Tpa2	MDAMKRGLCCVLLCGAVFVSPS	ATGGATGCAATGAAGAGAGGGC TCTGCTGTGTGCTGCTGCTGTGTG GAGCAGTCTTCGTTTCGCCCAGC

293細胞を、8_2_TV1_C.ZA(TV1c8.2)エンベロープの、gp120形態およびgp140形態をコードする、ネイティブの構築物および配列改変された構築物を用いて、標準的な方法を使用して一過的にトランスフェクトした。Envタンパク質を、細胞溶解物および上清中で、内部Env捕捉ELISAを使用して、測定した。結果は、以下の表5に示され、そしてTV1c8.2の野生型のシグナルペプチドリーダー配列が、異種tpaリーダー配列を使用して観察されたレベルよりも良いかまたはこのレベルに匹敵するレベルで、コードされたエンベロープタンパク質を効果的に発現するために使用され得ることを示す。さらに、このTV1c8.2リーダーは、ネイティブの形態または配列改変された形態で働き、そしてネイティブのenv遺伝子または配列改変されたenv遺伝子と共に使用され得る。全ての構築物を、pCMVlink発現ベクターのEcoRIおよびXhoI部位への遺伝子カセットのクローニング後に、試験した。

【0274】

【表7】

表5

TV1c8.2 構築物	上清 (ng)	溶解物 (ng)	合計 (ng)
gp140nat.wtL	532	149	681
gp140nat.tpa1	250	20	270
gp140nat.tpa2	192	34	226
gp120mod.wtLmod	6186	4576	10762
gp120mod.tpa1	6932	3808	10740
gp120mod.wtLnat	6680	4174	10854
gp140mod.wtLmod	1844	8507	10351
gp140mod.tpa1	1854	2925	4779
gp140mod.wtLnat	1532	3015	4547

配列改変されたTV1c8.2エンベロープ改変体遺伝子カセットを、安定な哺乳動物細胞株の誘導のために、Chiron pCMV発現ベクターにサブクローニングした。T

V1c8. 2エンベロープタンパク質を発現する安定なCHO細胞株を、トランスフェクション、メトトレキサート増幅およびスクリーニングの標準的な方法を使用して、誘導した。これらの細胞株は、tpaリーダー配列を使用して発現されたタンパク質について観察されたレベルと匹敵するレベルのエンベロープタンパク質を分泌することが見出された。代表的な結果を、TV1c8. 2 gp120を発現する2つの細胞株クローンについて、表6に示す；これらを、類似の様式であるが、tpaリーダーを使用して誘導されたSF162サブタイプB gp120を発現する2つの参照クローンと比較する。タンパク質濃度を、半精製されたタンパク質のゲルを走査する濃度測定の後決定した。検量線を、SF2 gp120タンパク質の、高度に精製されかつ十分特徴付けられた調製物を使用して作製し、そして試験タンパク質の濃度を決定した。

10

【0275】

【表8】

表6

CHO細胞株	クローン番号	発現 (ng/ml)
gp120 SF162	クローン 65	921
	クローン 71	972
gp120TV1.C8.2	クローン 159	1977
	クローン 210	1920

20

これらの結果をまた、本質的に実施例3に記載のように、ウェスタンブロット分析によって確認した。

【0276】

(実施例3)

(発現のウェスタンブロット分析)

(A. HIV C型コード配列)

ヒト293細胞を、実施例2に記載されるように、天然または合成のHIV C型発現カセットを含むpCMVベースのベクターでトランスフェクトする。細胞を、トランスフェクト後の60時間、培養する。上清を記載のように調製する。細胞溶解物を、以下のように調製する。この細胞を、リン酸緩衝生理食塩水で1回洗浄し、界面活性剤(0.1M Tris-HCl, pH7.5中に1% NP40 (Sigma Chemical Co., St. Louis, MO)を含む)で溶解し、そしてこの溶解物を、新しいチューブに移す。SDS-ポリアクリルアミドゲル(あらかじめ成型した8-16%; Novex, San Diego, CA)に、20μlの上清または12.5μlの細胞溶解物を充填する。タンパク質スタンダード(5μl, 広範なサイズ範囲のスタンダード; BioRad Laboratories, Hercules, CA)もまた、充填する。電気泳動を実施し、タンパク質を、BioRad Transfer Chamber (BioRad Laboratories, Hercules, CA)を用いて、製造者の推薦の転写緩衝液(Millipore)を使用して、Immobilon P膜(Millipore Corp., Bedford, MA)に転写する。ここで転写は、100ボルトで90分間実施する。この膜を、HIV-1陽性ヒト患者の血清に曝露し、そしてオーフェニレンジアミン二塩化水素化物(OPD; Sigma)を用いて免疫染色する。

30

40

【0277】

免疫ブロッティング分析は、合成発現カセットを含む細胞が、ネイティブな発現カセットを含む細胞よりも細胞あたりより高い濃度で、予期したタンパク質を産生することを示す。このタンパク質は細胞溶解物および上清の両方に見られる。産生のレベルは、本発明の合成発現カセットでトランスフェクトした細胞の細胞上清において有意に高い。

【0278】

さらに、トランスフェクトした293細胞由来の上清を、ショ糖勾配で分画する。この上

50

清のアリコート、20% (wt/wt) ショ糖溶液が下層に入ったPolyclearTM超遠心チューブ(Beckman Instruments, Columbia, MD)に移し、そしてBeckman SW28ローターで、28,000rpmで2時間の遠心分離に供する。生じるペレットをPBSに懸濁し、20~60% (wt/wt) ショ糖勾配上に重層し、そしてBeckman SW41tiローターで、40,000rpmで2時間の遠心分離に供する。

【0279】

次いで、勾配をおよそ10×1mlのアリコートに分画した(勾配の最高点で開始し、20%で終了)。サンプルを1~9分画から回収し、そして8~16%SDSポリアクリルアミドゲル上で電気泳動に供する。293/合成細胞由来の上清は、293/ネイティブな細胞由来の上清よりも非常に強いバンドを生じる。 10

【0280】

(実施例4)

(合成HIV C型発現カセットのインビボでの免疫原性)

(A. 免疫)

合成HIV C型発現カセットのおそらく改善される免疫原性を評価するために、マウス研究を実施する。プラスミドDNA(合成Gag発現カセットを有するpCMV KM2)を、100μlの総注射容積に、以下の最終濃度で希釈する: 20μg、2μg、0.2μg、0.02μgおよび0.002μg。希釈したDNAの考えられる陰性の希釈効果に克服するために、各サンプルにおける総DNA濃度を、ベクター(pCMV KM2)のみを用いて20μgにする。コントロールとして、ネイティブのGag発現カセットのプラスミドDNAを、同様の様式で取り扱う。12の群(4~10匹)のBalb/cマウス(Charles River, Boston, MA)を、表1におけるスケジュールに従って、筋内で免疫した(1下肢あたり50μl、前脛骨筋中に筋内注射)。 20

【0281】

【表9】

表 1

群	Gag または Env の 発現カセット	Gag または Env プラスミド DNA の濃度 (μg)	免疫した時間 (週):
1	合成	20	0 ¹ , 4
2	合成	2	0, 4
3	合成	0.2	0, 4
4	合成	0.02	0, 4
5	合成	0.002	0, 4
6	合成	20	0
7	合成	2	0
8	合成	0.2	0
9	合成	0.02	0
10	合成	0.002	0
11	ネイティブ	20	0, 4
12	ネイティブ	2	0, 4
13	ネイティブ	0.2	0, 4
14	ネイティブ	0.02	0, 4
15	ネイティブ	0.002	0, 4
16	ネイティブ	20	0
17	ネイティブ	2	0
18	ネイティブ	0.2	0
19	ネイティブ	0.02	0
20	ネイティブ	0.002	0

10

20

30

40

50

1 = 「0 週間」で最初の免疫

1～5 群および 11～15 群を、0 週間（免疫前）、4 週間、6 週間、8 週間、および 12 週間で出血させる。6～20 群および 16～20 群を、0 週間（免疫前）および 4 週間で出血させる。

【0282】

（B. 体液性免疫応答）

体液性免疫応答を、免疫の 0 週間後および 4 週間後（5～12 群）、さらに免疫の 6 週間後および 8 週間後（それぞれ 2 回目の免疫の 2 週間後および 4 週間後）（1～4 群）に、マウス血清の抗 HIV 抗体 ELISA 法（固相酵素免疫検定法）を用いて確認する。

【0283】

血清の抗体価を、適切な抗 HIV ポリペプチド（例えば、抗-Pol、抗-Gag、抗-Env、抗-Vif、抗-Vpu など）抗体 ELISA を用いて決定する。簡単に言えば、免疫したマウス由来の血清を、HIV タンパク質（例えば、p55 Gag タンパク質、Env タンパク質（例えば、gp160 もしくは gp120）または Pol タンパク質（例えば、p6、p67 もしくは RT）など）に対して指向された抗体によりスクリーニングする。ELISA マイクロタイタープレートを、1 ウェルあたり 0.2 μg の HIV タンパク質でオーバーナイトでコートし、そして 4 回洗浄する；次に、ブロッキングを、PBS-0.2% Tween (Sigma) を用いて 2 時間行う。ブロッキング溶液の除去後、100 μl の希釈したマウス血清を添加する。その後、血清を 1/25 希釈で試験し、そして連続する 3 倍希釈で試験する。マイクロタイタープレートを、4 回洗浄し

、そして第二のペルオキシダーゼ結合抗マウスIgG抗体(Pierce, Rockford, IL)とともにインキュベートする。ELISAプレートを洗浄し、1ウェルあたり100 μ lの3, 3', 5, 5'-テトラメチルベンジジン(TMB; Pierce)を添加する。各々のウェルの光学密度を15分後に測定する。この報告される力価は、半分の最大光学密度(O. D.)を得る血清の希釈の逆数である。

【0284】

合成発現カセットは、ネイティブの発現カセットに対して免疫原性の明らかな改善をもたらす。

【0285】

(C. 細胞性免疫反応)

特異的細胞傷害性Tリンパ球(CTL)の頻度を、ペプチドパルスしたマウス(Balb/c, C57BL/6および/またはC3H)CD4細胞の標準的なクロマチン放出アッセイにより評価した。HIVポリペプチド(例えば、Pol, GagまたはEnv)発現ワクシニアウイルスに感染したCD8細胞を、陽性コントロールとして使用した。要するに、上記のように免疫したマウスから得た脾臓細胞(エフェクター細胞、E)を培養し、再刺激し、そしてGagペプチドパルスした標的細胞に対するCTL活性についてアッセイした(Doe, B., およびWalker, C. M., ADIS 10(7): 793~794, 1996に記載のように)。細胞傷害性活性を、標準的⁵¹Cr放出アッセイで測定する。標的(T)細胞を、種々のE:T比率でエフェクター(E)細胞と共に4時間培養し、そして二つ組のウェルからの平均cpmを用いて、⁵¹Cr放出比の割合を計算する。

10

20

【0286】

細胞傷害性T細胞(CTL)活性を、HIV GagまたはEnv DNAで免疫したマウスから回収する脾臓細胞において測定する。GagまたはEnv DNAで免疫した動物由来のエフェクター細胞は、HIVポリペプチドパルスしたSV-BALB(MHC適合する)標的細胞(CTL反応を示す)の溶解比を示す。ペプチドパルスされ、そしてMHC非適合マウス系(MC57)由来である標的細胞は、溶解しない。

【0287】

従って、合成発現カセットは、DNA免疫化による細胞傷害性Tリンパ球(CTL)反応の誘導能力の増強を示す。

30

【0288】

(実施例5)

(合成HIV C型発現カセットを用いる非ヒト霊長類のDNA免疫化)

非ヒト霊長類を、皮内に、粘膜にまたは両側の四頭筋に、筋内に種々の用量(例えば、1~5mg)および合成HIV C型プラスミドの種々の組み合わせを用いて、複数回(例えば、週に0、4、8および24回)免疫する。この動物を、各々の免疫の2週間後に出血させ、そしてELISAを単離した血漿を用いて実施する。このELISAを、二次抗体の結合が抗ヒトIgG(g鎖特異的)ペルオキシダーゼ結合であり(Sigma Chemical Co., St. Louis, MD 63178)、これを1:500の希釈で用いたことを除いて、本質的に実施例4に記載のように実施する。50 μ g/mlの酵母抽出物を、血漿サンプルの希釈物に添加し、そして抗体を結合させて、非ヒト霊長類において予め存在する酵母抗体に起因して非特異的バックグラウンドを減少させる。

40

【0289】

さらに、抗原に対するリンパ増殖性反応をまた、Tヘルパー細胞機能の誘導を示す免疫後に評価し得る。

【0290】

合成プラスミドDNAは、非ヒト霊長類において免疫原性になることが予測される。

【0291】

(実施例6)

(合成HIV C型発現カセットを含む組換えシンドビスのRNAおよびDNAのインビ

50

トロでの発現)

αウイルスベクターにおける合成P o l、E v nおよびG a g発現カセットの発現効率を評価するために、この選択した合成発現カセットを、プラスミドDNAベースのシンドビスウイルスベクターおよび組換えベクター粒子ベースのシンドビスウイルスベクターの両方にサブクロニングする。具体的には、シンドビスウイルスRNAベクターレプリコンのインビトロでの転写のためのcDNAベクター構築物(pRSIN-luc; Dubenskyら、J. Viol. 70:508~519, 1996)を、プラスミドの直線化のためのPme I部位および異種遺伝子の挿入のためのポリリンカーを含むように改変する。ポリリンカーは、Xho I部位、Pml I部位、Apa I部位、Nar I部位、Xba I部位およびNot I部位(XPANXNFおよびXPNXNR)を含む2つのオリゴヌクレオチドを用いて形成する。 10

【0292】

プラスミドpRSIN-luc(Dubenskyら、前出)を、Xho IおよびNot Iを用いて消化し、ルシフェラーゼ遺伝子挿入物を取り除き、KlenowおよびdNTPを用いて平滑末端化し、そしてGeneClean II(Bio01, Vista, CA)を用いてアガロースゲルから精製する。このオリゴヌクレオチドを、互いにアニールさせ、そしてプラスミドと結合させる。生じる構築物を、Not IおよびSac Iで消化し、最小のシンドビス3'末端配列およびA₄₀トラクトを取り除き、そしてPKSSIN1-BV(WO97/38087)由来の約0.4kbpのフラグメントと連結させる。この0.4kbpのフラグメントを、Not IおよびSac Iを用いるpKSSIN1 20-BVの消化、そしてアガロースゲルからのサイズ分画後の精製によって得る。このフラグメントは、完全なシンドビスウイルスの3'末端およびA₄₀トラクトならびに直線化のためのPme I部位を含む。この新しいベクター構築物をSINBVEと命名する。

【0293】

この合成HIVコード配列を、EcoRIで消化し、KlenowおよびdNTPを用いて平滑末端化し、GeneClean IIで精製し、Sal Iで消化し、アガロースゲル上でサイズ分画し、そしてGeneClean IIを用いてアガロースゲルから精製することによって親プラスミドから得る。この合成HIVポリペプチドコードフラグメントを、Xho IおよびPmt Iで消化したSINBVEベクターに連結させる。生じるベクターを、GeneClean IIを用いて精製し、そしてSINBVGagで消化する。ベ 30クターRNAレプリコンを、インビトロでSINBVGagから転写させ(Dubenskyら、前出)、そして細胞のトランスフェクションに直接使用し得る。あるいは、このレプリコンを、不完全なヘルパーRNAとともに同時トランスフェクションするか、またはαウイルスパッケージング細胞株を用いることによって、組換えベクター粒子中にパッケージし得る。

【0294】

このDNAベースのシンドビスウイルスベクターpDCMV SIN-β-gal(Dubenskyら、J. Viol. 70:508~519, 1996)を、Sal IおよびXba Iで消化し、β-ガラクトシダーゼ遺伝子挿入物を取り除き、アガロースゲルのサイズ分画後、GeneClean IIを用いて精製する。Sal IおよびXho IでSIN 40BVGagを消化すること、アガロースゲルのサイズ分画後のGag含有フラグメントのGeneClean IIを用いた精製および連結によって、HIV GagまたはEnv遺伝子を、pDCMV SIN-β-galに挿入させる。生じる構築物を、pDSIN-Gagと呼び、これは本明細書中に記載の任意の方法を使用して、インビボでの投与または処方のため直接的に使用し得る。

【0295】

BHK細胞および293細胞を、それぞれ組換えシンドビスRNAおよびDNAでトランスフェクトする。この上清および細胞溶解物を、Coulter capture ELISAを用いて試験する(実施例2)。

【0296】

BHK細胞を、組換えシンドビスRNAを用いてエレクトロポレーションによってトランスフェクトする。

【0297】

293細胞を、組換えシンドビスDNAを用いてLT-1を使用してトランスフェクトする（実施例2）。合成Gag含有プラスミドおよび／またはEnv含有プラスミドを、陽性コントロールとして使用する。上清および溶解物をトランスフェクションの48時間後に収集する。

【0298】

C型 HIVタンパク質を、合成発現カセットを用いて、DNAおよびRNAベースのシンドビスベクターシステムの両方から効率よく発現させ得る。

10

【0299】

（実施例7）

（合成Pol、Gagおよび／またはEnv発現カセットを含む組換えシンドビスレプリコンベクターのインビボでの免疫原性）

（A. 免疫）

シンドビスレプリコンにおける組換え合成HIV C型発現カセットの免疫原性を評価するために、マウスの研究を実施する。この合成発現カセットを有するシンドビスウイルスDNAベクター（実施例6）を、総注射容積（100 μ l）に以下の最終濃度で希釈する：20 μ g、2 μ g、0.2 μ g、0.02 μ gおよび0.002 μ g。希釈したDNAの考えられる陰性の希釈効果を克服するために、各々のサンプルの総DNA濃度を、シンドビスレプリコンベクターDNAのみを用いて20 μ gにする。12群（4～10匹）のBalb/cマウス（Charles River, Boston, MA）を、表2のスケジュールに従って、筋内で免疫した（1下肢あたり50 μ l、前脛骨筋中に筋内注射）。あるいは、シンドビスウイルス粒子を、表3に示すように、100 μ l中に以下の用量で調整する： 10^3 pfu、 10^5 pfuおよび 10^7 pfu。シンドビスHIVポリペプチド粒子調整物を、マウスに筋内経路および皮下経路を用いて投与する（1部位あたり50 μ l）。

20

【0300】

【表10】

表 2

30

群	Gag または Env の 発現カセット	Gag または Env DNA の濃度 (μ g)	免疫した時間 (週):
1	合成	20	0 ¹ , 4
2	合成	2	0, 4
3	合成	0.2	0, 4
4	合成	0.02	0, 4
5	合成	0.002	0, 4
6	合成	20	0
7	合成	2	0
8	合成	0.2	0
9	合成	0.02	0
10	合成	0.002	0

40

1 = 「0 週間」で最初の免疫

【0301】

【表11】

50

表 3

群	Gag または Env 配列	ウイルス粒子の濃度 (pfu)	免疫した時間 (週):
1	合成	10^3	0 ¹ , 4
2	合成	10^5	0, 4
3	合成	10^7	0, 4
8	合成	10^3	0
9	合成	10^5	0
10	合成	10^7	0

10

1 = 「0 週間」で最初の免疫

群を出血させ、そして体液性および細胞性（例えば、特異的 C T L の頻度）の両方の評価を、本質的に実施例 4 に記載のように実施する。

【0302】

（実施例 8）

（新規の H I V C 型改変体の同定および配列決定）

H I V C 型をコードする 8 _ 5 _ T V 1 _ C . Z A と呼ばれる全長クローンを、単離し
そして配列決定した。手短に言えば、H I V - 1 C 亜型に感染した南アフリカ人の患者
由来のゲノム DNA を、アルカリ性溶解および陰イオン交換カラム (Q u i a g e n) に
より P B M C (末梢血単核細胞) から単離した。全長クローンのゲノムを得るために、二
等分を増幅した。後にこれらを、両方のフラグメントにおける特有な S a l I 部位を用い
て P o l 領域内で、インフレームで一緒に結合させる。増幅のために、業者のプロトコル
の後に、200 ~ 800 n g のゲノム DNA を、緩衝液および E x p a n d L o n g
T e m p l a t e P C R S y s t e m (B o e h r i n g e r M a n n h a e i m) の酵素混合物に添加した。このプライマーを既知の全長配列の整列の後に設計した。5
' 半分のために、20 位置にチミジン (S 1 F C S a c T A 5 ' - G T T T C T T G A
G C T C T G G A A G G G T T A A T T T A C T C C A A G A A - 3 '、配列番号 38)
またはシトシン (S 1 F T S a c T A 5 ' - G T T T C T T G A G C T C T G G A A G
G G T T A A T T T A C T C T A A G A A、配列番号 39) いずれか、および S a l I 部
位を含む 2 つの正方向プライマーのプライマー混合物を用いた。逆方向プライマーもまた
、13 位置にチミジンまたはシトシンのいずれかを含む 2 つのプライマー (S 1 4 5 R T
S a l T A 5 ' - G T T T C T T G T C G A C T T G T C C A T G T A T G G C T T C
C C C T - 3 '、配列番号 40 および S 1 4 5 R T S a l T A 5 ' - G T T T C T T G
T C G A C T T G T C C A T G C A T G G C T T C C C T - 3 ' 配列番号 41) の混合物
であり、そして S a l I 部位を含む。3 ' 半分についての正方向プライマーはまた、S a
l I 部位および 12 位置にアデニンまたはグアニンを有する 2 つのプライマー (S 2 4 F
A S a l T A 5 ' - G T T T C T T G T C G A C T G T A G T C C A G G a A T A T G
G C A A T T A G - 3 ' 配列番号 42 および S 2 4 F G S a l T A 5 ' - G T T T C T
T G T C G A C T G T A G T C C A G G g A T A T G G C A A T T A G - 3 ' 配列番号 4
3) の混合物である。逆方向プライマーは N o t I 部位を有する (S 2 _ F u l l N o t
T A 5 ' - G T T T C T T G C G G C C G C T G C T A G A G A T T T T C C A C A C
T A C C A - 3 ' 配列番号 44)。増幅後の P C R 産物を 1 % アガロースゲルを用いて精
製し、T A クローニング (I b v i t r o g e n) によって p C R - X L - T O P O ベク
ターにクローニングした。コロニーを制限分析によって確認し、配列を実証した。全長配
列のために、5 ' 半分および 3 ' 半分の配列を結合させた。この配列を配列番号 33 に示
す。さらに、重要なドメインを表 A に示す。

【0303】

50

1 2 - 5 _ 1 _ T V 2 _ C . Z A と呼ばれる別のクローンもまた、配列決定され、配列番号 4 5 に示す。このドメインは本明細書の技術から見て容易に決定され得る。例えば、表 A に示されるこれらの配列を整列されることによって、クローン 1 2 - 5 _ 1 _ T V 2 _ C . Z A の対応領域を発見する。

【0304】

上記のように（実施例 1、表 C）、合成発現カセットを、8 _ 5 _ T V 1 _ C . Z A または 1 2 - 5 _ 1 _ T V 2 _ C . Z A から得た 1 つ以上のポリヌクレオチド配列を用いて生成した。

【0305】

本明細書中に記載のポリヌクレオチドは、全て Chiron Corporation Emeryville, CA に寄託している。 10

【0306】

本発明の好ましい実施形態はいくらか詳細に記載されているが、明らかな改変が添付の請求項により規定される本発明の精神および範囲から逸脱することなしになされ得ることが理解される。

【図面の簡単な説明】

【図 1】

図 1（配列番号 3）は、合成 Gag ポリペプチドをコードするポリヌクレオチドのヌクレオチド配列を示す。示されるヌクレオチド配列は、C 型 A F 1 1 0 9 6 5 株を改変することによって得られ、INS のさらなる改変を含む。 20

【図 2】

図 2（配列番号 4）は、合成 Gag ポリペプチドをコードするポリヌクレオチドのヌクレオチド配列を示す。示されるヌクレオチド配列は、C 型 A F 1 1 0 9 6 7 株を改変することによって得られ、INS のさらなる改変を含む。

【図 3】

図 3（配列番号 9）は、合成 Env ポリペプチドをコードするポリヌクレオチドのヌクレオチド配列を示す。このヌクレオチド配列は、g p 1 6 0（シグナルペプチドを含む）を示し、C 型 A F 1 1 0 9 6 8 株を改変することによって得られた。矢印は、種々の領域のポリヌクレオチドの位置を示しており、これらとしては、以下をコードする配列が挙げられる：シグナルペプチド（ヌクレオチド 1～81）（配列番号 18）、g p 1 2 0 ポリペ 30
プチド（ヌクレオチド 82～1512）（配列番号 6）、g p 4 1 ポリペプチド（ヌクレオチド 1513～2547）（配列番号 10）、g p 1 4 0 ポリペプチド（ヌクレオチド 82～2025）（配列番号 7）および g p 1 6 0 ポリペプチド（ヌクレオチド 82～2547）（配列番号 8）。シグナルペプチドをコードするコドン、ネイティブの HIV-1 シグナル配列から（本明細書中に記載されるように）改変される。

【図 4】

図 4（配列番号 15）は、合成 Env ポリペプチドをコードするポリヌクレオチドのヌクレオチド配列を示す。このヌクレオチド配列は、g p 1 6 0（シグナルペプチドを含む）を示し、C 型 A F 1 1 0 9 7 5 株を改変することによって得られた。矢印は、種々の領域のポリヌクレオチドの位置を示しており、これらとしては、以下をコードする配列が挙げ 40
られる：シグナルペプチド（ヌクレオチド 1～72）（配列番号 19）、g p 1 2 0 ポリペプチド（ヌクレオチド 73～1509）（配列番号 12）、g p 4 1 ポリペプチド（ヌクレオチド 1510～2565）（配列番号 16）、g p 1 4 0 ポリペプチド（ヌクレオチド 73～2022）（配列番号 13）および g p 1 6 0 ポリペプチド（ヌクレオチド 73～2565）（配列番号 14）。シグナルペプチドをコードするコドン、ネイティブの HIV-1 シグナル配列から（本明細書中に記載されるように）改変される。

【図 5】

図 5 は、A F 1 1 0 9 6 5 から誘導された合成 Gag 配列における、いくつかの残存する INS の位置を示している。これらの配列に対してなされた変化が、図中で囲まれている。上の列は、示された株由来の Gag ポリペプチドの、コドンを改変された配列（配列番 50

号20)を示している。囲んだ領域中の列の下に現れるヌクレオチドは、さらなるINSを取り除かれ、かつ図1に示される配列(配列番号3)に対応するようになされた変化を示している。

【図6】

図6は、AF110967から誘導された合成Gag配列における、いくつかの残存するINSの位置を示している。これらの配列に対してなされた変化が、図中で囲まれている。上の列は、示された株由来のGagポリペプチドの改変された配列(配列番号21)を示している。囲んだ領域中の系列の下に現れるヌクレオチドは、さらなるINSを取り除かれ、かつ図2に示される配列(配列番号4)に対応するようになされた変化を示している。

10

【図7】

図7は、HIVのPol領域における、選択されたドメインを示した概略図である。

【図8】

図8(配列番号30)は、PR975(+)と表される合成構築物のヌクレオチド配列を示している。「(+)」は、逆転写が機能的であることを示している。この構築物としては、以下由来の配列が挙げられる:p2(配列番号30のヌクレオチド16~54);p7(配列番号30のヌクレオチド55~219);p1/p6(配列番号30のヌクレオチド220~375);prot(配列番号30のヌクレオチド376~672)、逆転写酵素(配列番号30のヌクレオチド673~2352);および図7に示される6のアミノ酸のインテグラーゼ(配列番号30のヌクレオチド2353~2370)。さらに、この構築物は、導入遺伝子およびYMDDエピトープカセット(配列番号30のヌクレオチド2371~2424)を挿入するための複数のクローニング部位(MCS、配列番号30のヌクレオチド2425~2463)を含む。

20

【図9】

図9(配列番号31)は、PR975YMと表される合成構築物のヌクレオチド配列を示している。図7に示されるように、RT領域は、触媒中心における変異(mut. center)を含んでいる。「YM」とは、ヌクレオチドがこの領域において、YMDDではなくアミノ酸APをコードする構築物をいう。逆転写酵素は、この構築物において機能的ではない。この構築物としては、p2(配列番号31のヌクレオチド16~54);p7(配列番号31のヌクレオチド55~219);p1/p6(配列番号31のヌクレオチド220~375);prot(配列番号31のヌクレオチド376~672)、および図7に示される逆転写酵素(配列番号31のヌクレオチド673~2346)由来の配列が挙げられるが、逆転写酵素タンパク質は、機能的ではない。さらに、この構築物は、導入遺伝子およびYMDDエピトープカセット(配列番号31のヌクレオチド2365~2418)を挿入するための複数のクローニング部位(MCS、配列番号31のヌクレオチド2419~2457)を含む。

30

【図10】

図10(配列番号32)は、PR975YMWと表される合成構築物のヌクレオチド配列を示している。「YM」とは、ヌクレオチドがこの領域において、YMDDではなくアミノ酸APをコードする構築物をいう。「WM」とは、ヌクレオチドがこの領域において、WMGYではなくアミノ酸PIをコードする構築物をいう。この構築物としては、p2(配列番号32のヌクレオチド16~54);p7(配列番号32のヌクレオチド55~219);p1/p6(配列番号32のヌクレオチド220~375);prot(配列番号32のヌクレオチド376~672)、および図7に示される逆転写酵素(配列番号32のヌクレオチド673~2340)由来の配列が挙げられるが、逆転写酵素タンパク質は、機能的ではない。さらに、この構築物は、導入遺伝子およびYMDDエピトープカセット(配列番号32のヌクレオチド2359~2412)を挿入するための複数のクローニング部位(MCS、配列番号32のヌクレオチド2413~2451)を含む。

40

【図11】

図11(配列番号33)は、8__5__TV1__C.ZAのヌクレオチド配列を示す。種々

50

の領域を、表 A に示す。

【図 1 2】

図 1 2 (配列番号 3 4) は、p 2 g a g から p 7 g a g までの、A F 1 1 0 9 7 5 P o l の野生型ヌクレオチド配列を示す。

【図 1 3】

図 1 3 (配列番号 3 5) は、p 1 からインテグラーゼタンパク質の最初の 6 アミノ酸の、A F 1 1 0 9 7 5 P o l の野生型ヌクレオチド配列を示す。

【図 1 4】

図 1 4 (配列番号 3 6) は、逆転写酵素の I l e 1 7 8 ~ セリン 1 9 1 をコードするカセットのヌクレオチド配列を示す。

10

【図 1 5】

図 1 5 (配列番号 3 7) は、逆転写酵素タンパク質の触媒中心領域中のエピトープを含むアミノ酸配列を示す。

【図 1 6】

図 1 6 (配列番号 4 5) は、1 2 - 5 _ 1 _ T V 2 _ C . Z A のヌクレオチド配列を示す。

【図 1 7】

図 1 7 (配列番号 4 6) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 E n v コードポリヌクレオチドのヌクレオチド配列を示す。この配列は、短い (9 7 塩基対) 共通領域に対応する。

20

【図 1 8】

図 1 8 (配列番号 4 7) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 E n v コードポリヌクレオチドのヌクレオチド配列を示す。この配列は、E n v 中の共通領域に対応する。

【図 1 9】

図 1 9 (配列番号 4 8) は、8 _ 5 _ T V 1 _ C . Z A E n v の野生型ヌクレオチド配列を示す。

【図 2 0】

図 2 0 (配列番号 4 9) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 E n v g p 1 6 0 コードポリヌクレオチドのヌクレオチド配列を示す。

【図 2 1】

図 2 1 (配列番号 5 0) は、8 _ 5 _ T V 1 _ C . Z A E n v g p 1 6 0 の野生型ヌクレオチド配列を示す。

30

【図 2 2】

図 2 2 (配列番号 5 1) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 G a g コードポリヌクレオチドのヌクレオチド配列を示す。

【図 2 3】

図 2 3 (配列番号 5 2) は、8 _ 5 _ T V 1 _ C . Z A G a g の野生型ヌクレオチド配列を示す。

【図 2 4】

図 2 4 (配列番号 5 3) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 G a g コードポリヌクレオチド (主要な相同性領域) のヌクレオチド配列を示す。

40

【図 2 5】

図 2 5 (配列番号 5 4) は、8 _ 5 _ T V 1 _ C . Z A G a g 主要相同性領域の野生型ヌクレオチド配列を示す。

【図 2 6】

図 2 6 (配列番号 5 5) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 N e f コードポリヌクレオチドのヌクレオチド配列を示す。

【図 2 7】

図 2 7 (配列番号 5 6) は、8 _ 5 _ T V 1 _ C . Z A N e f の野生型ヌクレオチド配列を示す。

50

【図 28】

図 28 (配列番号 57) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 N e f コードポリヌクレオチドのヌクレオチド配列を示す。この配列は、非機能遺伝子産物を生じる 1 2 5 位における変異を含む。

【図 29】

図 29 (配列番号 58) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 R N A s e H コードポリヌクレオチドのヌクレオチド配列を示す。R n a s e H は、P o l 遺伝子の機能的ドメインであり、p 1 5 に対応する (表 A)。

【図 30】

図 30 (配列番号 59) は、8 _ 5 _ T V 1 _ C . Z A R N A s e H の野生型ヌクレオチド配列を示す。 10

【図 31】

図 31 (配列番号 60) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成インテグラーゼ (I n t) コードポリヌクレオチドのヌクレオチド配列を示す。I n t は、P o l 遺伝子の機能的ドメインであり、p 3 1 に対応する (表 A)。

【図 32】

図 32 (配列番号 61) は、8 _ 5 _ T V 1 _ C . Z A I n t の野生型ヌクレオチド配列を示す。

【図 33】

図 33 (配列番号 62) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 P o l コードポリヌクレオチドのヌクレオチド配列を示す。 20

【図 34】

図 34 (配列番号 63) は、8 _ 5 _ T V 1 _ C . Z A P o l の野生型ヌクレオチド配列を示す。

【図 35】

図 35 (配列番号 64) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成プロテアーゼ (p r o t) コードポリヌクレオチドのヌクレオチド配列を示す。

【図 36】

図 36 (配列番号 65) は、8 _ 5 _ T V 1 _ C . Z A P r o t の野生型ヌクレオチド配列を示す。 30

【図 37】

図 37 (配列番号 66) は、プロテアーゼの不活性化を生じる変異を含む、8 _ 5 _ T V 1 _ C . Z A 由来の合成プロテアーゼ (p r o t) コードポリヌクレオチドのヌクレオチド配列を示す。

【図 38】

図 38 (配列番号 67) は、8 _ 5 _ T V 1 _ C . Z A 不活性化 P r o t の野生型ヌクレオチド配列を示す。

【図 39】

図 39 (配列番号 68) は、共に 8 _ 5 _ T V 1 _ C . Z A 由来である、合成プロテアーゼ (p r o t) コードポリヌクレオチドおよび合成逆転写酵素 (R T) コードポリヌクレオチドのヌクレオチド配列を示す。P r o t 配列および R T 配列は、共に、遺伝子産物の不活性化を生じる変異を含む。 40

【図 40】

図 40 (配列番号 69) は、8 _ 5 _ T V 1 _ C . Z A 不活性化 P r o t / 変異 R T の野生型ヌクレオチド配列を示す。

【図 41】

図 41 (配列番号 70) は、共に 8 _ 5 _ T V 1 _ C . Z A 由来である、合成プロテアーゼ (p r o t) コードポリヌクレオチドおよび合成逆転写酵素 (R T) コードポリヌクレオチドのヌクレオチド配列を示す。

【図 42】

図 4 2 (配列番号 7 1) は、8 _ 5 _ T V 1 _ C . Z A P r o t および R T の野生型ヌクレオチド配列を示す。

【図 4 3】

図 4 3 (配列番号 7 2) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 r e v コードポリヌクレオチドのヌクレオチド配列を示す。この示した合成配列は、r e v のエキソン 1 に対応する。野生型 r e v は、2 つのエキソンを有する。

【図 4 4】

図 4 4 (配列番号 7 3) は、R e v の 8 _ 5 _ T V 1 _ C . Z A エキソン 1 の野生型ヌクレオチド配列を示す。

【図 4 5】

図 4 5 (配列番号 7 4) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 r e v コードポリヌクレオチドのヌクレオチド配列を示す。この示した合成配列は、r e v のエキソン 2 に対応する。

【図 4 6】

図 4 6 (配列番号 7 5) は、R e v の 8 _ 5 _ T V 1 _ C . Z A エキソン 2 の野生型ヌクレオチド配列を示す。

【図 4 7】

図 4 7 (配列番号 7 6) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 R T コードポリヌクレオチドのヌクレオチド配列を示す。

【図 4 8】

図 4 8 (配列番号 7 7) は、8 _ 5 _ T V 1 _ C . Z A R T の野生型ヌクレオチド配列を示す。

【図 4 9】

図 4 9 (配列番号 7 8) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 R T コードポリヌクレオチドのヌクレオチド配列を示す。この合成ポリヌクレオチドは、遺伝子産物を不活性にする R T コード配列中の変異を含む。

【図 5 0】

図 5 0 (配列番号 7 9) は、R T 遺伝子産物を不活性化する変異を含む、8 _ 5 _ T V 1 _ C . Z A R T の野生型ヌクレオチド配列を示す。

【図 5 1】

図 5 1 (配列番号 8 0) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 T a t コードポリヌクレオチドのヌクレオチド配列を示す。この示した合成配列は、T a t のエキソン 1 に対応し、さらに、T a t 遺伝子産物を非機能的にする変異を含む。野生型 T a t は、2 つのエキソンを有する。

【図 5 2】

図 5 2 (配列番号 8 1) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 T a t コードポリヌクレオチドのヌクレオチド配列を示す。示した合成配列は、T a t のエキソン 1 に対応する。

【図 5 3】

図 5 3 (配列番号 8 2) は、T a t の 8 _ 5 _ T V 1 _ C . Z A エキソン 1 の野生型ヌクレオチド配列を示す。

【図 5 4】

図 5 4 (配列番号 8 3) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 T a t コードポリヌクレオチドのヌクレオチド配列を示す。示した合成配列は、T a t のエキソン 2 に対応する。

【図 5 5】

図 5 5 (配列番号 8 4) は、T a t の 8 _ 5 _ T V 1 _ C . Z A エキソン 2 の野生型ヌクレオチド配列を示す。

【図 5 6】

図 5 6 (配列番号 8 5) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 V i f コードポリヌク

10

20

30

40

50

レオチドのヌクレオチド配列を示す。

【図 5 7】

図 5 7 (配列番号 8 6) は、8 _ 5 _ T V 1 _ C . Z A V i f の野生型ヌクレオチド配列を示す。

【図 5 8】

図 5 8 (配列番号 8 7) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 V p r コードポリヌクレオチドのヌクレオチド配列を示す。

【図 5 9】

図 5 9 (配列番号 8 8) は、8 _ 5 _ T V 1 _ C . Z A V p r の野生型ヌクレオチド配列を示す。

10

【図 6 0】

図 6 0 (配列番号 8 9) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 V p u コードポリヌクレオチドのヌクレオチド配列を示す。

【図 6 1】

図 6 1 (配列番号 9 0) は、8 _ 5 _ T V 1 _ C . Z A V p u の野生型ヌクレオチド配列を示す。

【図 6 2】

図 6 2 (配列番号 9 1) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 r e v コードポリヌクレオチドのヌクレオチド配列を示す。この示した合成配列は、r e v のエキソン 1 および 2 に対応する。

20

【図 6 3】

図 6 3 (配列番号 9 2) は、8 _ 5 _ T V 1 _ C . Z A 由来の r e v のエキソン 1 および 2 の野生型ヌクレオチド配列を示す。

【図 6 4】

図 6 4 (配列番号 9 3) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 T a t コードポリヌクレオチドのヌクレオチド配列を示す。この合成ポリヌクレオチドは、T a t のエキソン 1 および 2 の両方を含み、そしてさらに、この遺伝子産物を非機能的にするエキソン 1 中の変異を含む。

【図 6 5】

図 6 5 (配列番号 9 4) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 T a t コードポリヌクレオチドのヌクレオチド配列を示す。この合成ポリヌクレオチドは、T a t のエキソン 1 および 2 の両方を含む。

30

【図 6 6】

図 6 6 (配列番号 9 5) は、8 _ 5 _ T V 1 _ C . Z A 由来の T a t のエキソン 1 および 2 の野生型ヌクレオチド配列を示す。

【図 6 7】

図 6 7 (配列番号 9 6) は、8 _ 5 _ T V 1 _ C . Z A 由来の合成 N e f コードポリヌクレオチドのヌクレオチド配列を示す。この配列は、非機能的遺伝子産物を生じる 1 2 5 位の変異および N e f 遺伝子産物のミリストイル化を排除する変異を含む。

【図 6 8】

図 6 8 (配列番号 9 7) は、1 2 - 5 _ 1 _ T V 2 _ C . Z A 由来の合成 E n v g p 1 6 0 コードポリヌクレオチドのヌクレオチド配列を示す。

40

【図 6 9】

図 6 9 (配列番号 9 8) は、1 2 - 5 _ 1 _ T V 2 _ C . Z A 由来の E n v g p 1 6 0 の野生型ヌクレオチド配列を示す。

【図 7 0】

図 7 0 (配列番号 9 9) は、1 2 - 5 _ 1 _ T V 2 _ C . Z A 由来の合成 G a g コードポリヌクレオチドのヌクレオチド配列を示す。

【図 7 1】

図 7 1 (配列番号 1 0 0) は、1 2 - 5 _ 1 _ T V 2 _ C . Z A 由来の G a g の野生型ヌ

50

クレオチド配列を示す。

【図 7 2】

図 7 2 (配列番号 1 0 1) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の合成 N e f コードポリヌクレオチドのヌクレオチド配列を示す。

【図 7 3】

図 7 3 (配列番号 1 0 2) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の N e f の野生型ヌクレオチド配列を示す。

【図 7 4】

図 7 4 (配列番号 1 0 3) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の合成 P o l コードポリヌクレオチドのヌクレオチド配列を示す。

10

【図 7 5】

図 7 5 (配列番号 1 0 4) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の P o l の野生型ヌクレオチド配列を示す。

【図 7 6】

図 7 6 (配列番号 1 0 5) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の R e v のエキソン 1 由来の合成 R e v コードポリヌクレオチドのヌクレオチド配列を示す。

【図 7 7】

図 7 7 (配列番号 1 0 6) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の R e v のエキソン 1 の野生型ヌクレオチド配列を示す。

【図 7 8】

20

図 7 8 (配列番号 1 0 7) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の R e v のエキソン 2 由来の合成 R e v コードポリヌクレオチドのヌクレオチド配列を示す。

【図 7 9】

図 7 9 (配列番号 1 0 8) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の R e v のエキソン 2 の野生型ヌクレオチド配列を示す。

【図 8 0】

図 8 0 (配列番号 1 0 9) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の T a t エキソン 1 由来の、合成 T a t コードポリヌクレオチドのヌクレオチド配列を示す。

【図 8 1】

図 8 1 (配列番号 1 1 0) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の T a t のエキソン 1 の野生型ヌクレオチド配列を示す。

30

【図 8 2】

図 8 2 (配列番号 1 1 1) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の T a t エキソン 2 由来の、合成 T a t コードポリヌクレオチドのヌクレオチド配列を示す。

【図 8 3】

図 8 3 (配列番号 1 1 2) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の T a t のエキソン 2 の野生型ヌクレオチド配列を示す。

【図 8 4】

図 8 4 (配列番号 1 1 3) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の合成 V i f コードポリヌクレオチドのヌクレオチド配列を示す。

40

【図 8 5】

図 8 5 (配列番号 1 1 4) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の V i f の野生型ヌクレオチド配列を示す。

【図 8 6】

図 8 6 (配列番号 1 1 5) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の合成 V p r コードポリヌクレオチドのヌクレオチド配列を示す。

【図 8 7】

図 8 7 (配列番号 1 1 6) は、1 2 - 5 _ 1 _ T V 2 _ C. Z A 由来の V p r の野生型ヌクレオチド配列を示す。

【図 8 8】

50

図 88 (配列番号 117) は、12-5__1__TV2__C. ZA 由来の合成 Vpu コードポリヌクレオチドのヌクレオチド配列を示す。

【図 89】

図 89 (配列番号 118) は、12-5__1__TV2__C. ZA 由来の Vpu の野生型ヌクレオチド配列を示す。

【図 90】

図 90 (配列番号 119) は、8__2__TV1__C. ZA 由来の合成 Env gp120 コードポリヌクレオチドのヌクレオチド配列を示す。V2 領域は、欠失している。この配列は、EcoRI 制限部位 (ヌクレオチド 1~6) ; コドン改変されたシグナルペプチドリーダー配列 (ヌクレオチド 7~87) ; gp120 コード配列 (ヌクレオチド 88~1464) ; 終止コドン (ヌクレオチド 1465~1467) ; XhoI 制限部位 (ヌクレオチド 1468~1473) を含む。 10

【図 91】

図 91 (配列番号 120) は、8__2__TV1__C. ZA 由来の合成 Env gp140 コードポリヌクレオチドのヌクレオチド配列を示す。V2 領域は、欠失している。この配列は、EcoRI 制限部位 (ヌクレオチド 1~6) ; 改変シグナルペプチドリーダー配列 (ヌクレオチド 7~87) ; gp140 コード配列 (ヌクレオチド 88~1977) ; 終止コドン (ヌクレオチド 1978~1980) ; XhoI 制限部位 (ヌクレオチド 1981~1986) を含む。

【図 92】

図 92 (配列番号 121) は、8__2__TV1__C. ZA 由来の合成 Env gp140 コードポリヌクレオチドのヌクレオチド配列を示す。V2 領域は、欠失し、この配列は、gp140 ポリペプチドの gp120 ポリペプチドおよび gp41 ポリペプチドへの切断を予防する切断部位中の変異を含む。この配列は、EcoRI 制限部位 (ヌクレオチド 1~6) ; 改変シグナルペプチドリーダー配列 (ヌクレオチド 7~87) ; gp140 コード配列 (ヌクレオチド 88~1977) ; 終止コドン (ヌクレオチド 1978~1980) ; XhoI 制限部位 (ヌクレオチド 1981~1986) を含む。 20

【図 93】

図 93 (配列番号 122) は、8__2__TV1__C. ZA 由来の合成 Env gp160 コードポリヌクレオチドのヌクレオチド配列を示す。V1/V2 領域は、欠失している。この配列は、EcoRI 制限部位 (ヌクレオチド 1~6) ; 改変シグナルペプチドリーダー配列 (ヌクレオチド 7~87) ; gp160 コード配列 (ヌクレオチド 88~2388) ; 終止コドン (ヌクレオチド 2389~2391) ; XhoI 制限部位 (ヌクレオチド 2392~2397) を含む。 30

【図 94】

図 94 (配列番号 123) は、8__2__TV1__C. ZA 由来の合成 Env gp160 コードポリヌクレオチドのヌクレオチド配列を示す。V2 領域は、欠失している。この配列は、EcoRI 制限部位 (ヌクレオチド 1~6) ; 改変シグナルペプチドリーダー配列 (ヌクレオチド 7~87) ; gp160 コード配列 (ヌクレオチド 88~2520) ; 終止コドン (ヌクレオチド 2521~2523) ; XhoI 制限部位 (ヌクレオチド 2524~2529) を含む。 40

【図 95】

図 95 (配列番号 124) は、8__2__TV1__C. ZA 由来の合成 Env gp160 コードポリヌクレオチドのヌクレオチド配列を示す。V2 領域は、欠失し、そして切断部位が変異している。この配列は、EcoRI 制限部位 (ヌクレオチド 1~6) ; 改変シグナルペプチドリーダー配列 (ヌクレオチド 7~87) ; gp160 コード配列 (ヌクレオチド 88~2520) ; 終止コドン (ヌクレオチド 2521~2523) ; XhoI 制限部位 (ヌクレオチド 2524~2529) を含む。

【図 96】

図 96 (配列番号 125) は、8__2__TV1__C. ZA 由来の合成 Env gp160 50

コードポリヌクレオチドのヌクレオチド配列を示す。このヌクレオチド配列は、T P A 1 リーダー配列を含む。この配列は、S a l I 制限部位（ヌクレオチド 1 ～ 6）；コザック配列（ヌクレオチド 7 ～ 1 2）；T P A 1 シグナルペプチドリーダー配列（ヌクレオチド 1 3 ～ 8 7）；g p 1 6 0 コード配列（ヌクレオチド 8 8 ～ 2 6 0 4）；終止コドン（ヌクレオチド 2 6 0 5 ～ 2 6 0 7）；X h o I 制限部位（ヌクレオチド 2 6 0 8 ～ 2 6 1 3）を含む。

【図 9 7】

図 9 7（配列番号 1 2 6）、8 __ 2 __ T V 1 __ C. Z A 由来の合成 E n v g p 1 6 0 コードポリヌクレオチドのヌクレオチド配列を示す。この配列は、E c o R I 制限部位（ヌクレオチド 1 ～ 6）；改変シグナルペプチドリーダー配列（ヌクレオチド 7 ～ 8 7）；g p 1 6 0 コード配列（ヌクレオチド 8 ～ 2 6 0 7）；終止コドン（ヌクレオチド 2 6 0 8 ～ 2 6 1 0）；X h o I 制限部位（ヌクレオチド 2 6 1 1 ～ 2 6 1 6）を含む。

【図 9 8】

図 9 8（配列番号 1 2 7）は、8 __ 2 __ T V 1 __ C. Z A 由来の合成 E n v g p 1 6 0 コードポリヌクレオチドのヌクレオチド配列を示す。このヌクレオチド配列は、野生型リーダー配列を含む。この配列は、E c o R I 制限部位（ヌクレオチド 1 ～ 6）；ネイティブ（未改変）シグナルペプチドリーダー配列（ヌクレオチド 7 ～ 8 7）；g p 1 6 0 コード配列（ヌクレオチド 8 8 ～ 2 6 0 7）；終止コドン（ヌクレオチド 2 6 0 8 ～ 2 6 1 0）；X h o I 制限部位（ヌクレオチド 2 6 1 1 ～ 2 6 1 6）を含む。

【図 9 9】

図 9 9（配列番号 1 2 8）は、8 __ 2 __ T V 1 __ C. Z A. 由来の野生型 g p 1 6 0 のヌクレオチド配列を示す。

【図 1 0 0】

図 1 0 0（配列番号 1 3 1）は、8 __ 2 __ T V 1 __ C. Z A. 由来の合成 E n v g p 1 4 0 コードポリヌクレオチドのヌクレオチド配列を示す。このヌクレオチド配列は、T P A 1 リーダー配列（ヌクレオチド 1 ～ 7 5）；g p 1 4 0 コード配列（ヌクレオチド 7 6 ～ 2 0 4 9）；終止コドン（ヌクレオチド 2 0 5 0 ～ 2 0 5 2）；を含む。

【図 1 0 1】

図 1 0 1（配列番号 1 3 2）は、8 __ 2 __ T V 1 __ C. Z A. 由来の合成 g p 1 4 0 コードポリヌクレオチドのヌクレオチド配列を示す。このヌクレオチド配列は、E c o R I 制限部位（ヌクレオチド 1 ～ 6）；T V 1 __ C. Z A 野生型リーダー配列から改変されたリーダー配列（ヌクレオチド 7 ～ 8 7）；g p 1 4 0 コード配列（ヌクレオチド 8 8 ～ 2 0 6 4）；終止コドン（ヌクレオチド 2 0 6 5 ～ 2 0 6 7）；X h o I 制限部位（ヌクレオチド 2 0 6 8 ～ 2 0 7 3）；を含む。

【図 1 0 2】

図 1 0 2（配列番号 1 3 3）は、8 __ 2 __ T V 1 __ C. Z A. 由来の合成 g p 1 4 0 コードポリヌクレオチドのヌクレオチド配列を示す。このヌクレオチド配列は、野生型 T V 1 __ C. Z A 非改変リーダー配列を含む。このヌクレオチド配列は、制限部位（ヌクレオチド 1 ～ 6）；野生型リーダー配列（ヌクレオチド 7 ～ 8 7）；g p 1 4 0 コード配列（ヌクレオチド 8 8 ～ 2 0 6 4）；終止コドン（ヌクレオチド 2 0 6 5 ～ 2 0 6 7）；X h o I 制限部位（ヌクレオチド 2 0 6 8 ～ 2 0 7 3）；を含む。

【図 1 0 3】

図 1 0 3（配列番号 1 3 4）は、1 2 - 5 __ 1 __ T V 2 __ C. Z A. 由来の合成 N e f コードポリヌクレオチドのヌクレオチド配列を示す。この配列は、非機能的遺伝子産物を生じる位置 1 2 5 の変異を含む。

【図 1 0 4】

図 1 0 4（配列番号 1 3 5）は、1 2 - 5 __ 1 __ T V 2 __ C. Z A. 由来の合成 N e f コードポリヌクレオチドのヌクレオチド配列を示す。この合成ポリヌクレオチドは、N e f 遺伝子産物のミリストイル化部位を排除する変異を含む。

【図 1 0 5 A】

図 105 は、種々の H1V 単離体由来の Env ポリペプチドの整列を示す。矢印の間の領域は、欠失および／または短縮され得る β および／または架橋シート領域における (TV1 クローンおよび TV2 クローンの) 領域を示す。「*」は、(TV1 クローンおよび TV2 クローンの) N 連結グリコシル化部位を意味する。その 1 つ以上が改変 (例えば、欠失および／または変異) され得る。

10

【図 3】

Fig. 3

Nav_AF110968_C_BW_opt
 -->
 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103

【図 4】

Fig. 4

[illegible]

【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/04493 A2(51) International Patent Classification: **C07K 14/155**

(21) International Application Number: PCT/US01/21241

(22) International Filing Date: 5 July 2001 (05.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09610.315 5 July 2000 (05.07.2000) US

(71) Applicants (for all designated States except US): CHIRON CORPORATION (US/US); 4560 Horton Street, Emeryville, CA 94608 (US); THE UNIVERSITY OF STELLERBOSCH (ZA/ZA); P.O. Box 19063, 7505 Tygerberg (ZA).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ZUR MEDEDE, Jan (DE/US); c/o Chiron Corporation, 4560 Horton Street, R-440, Emeryville, CA 94608 (US); BARNETT, Susan, W. (US/US); c/o Chiron Corporation, 4560 Horton Street, R-440, Emeryville, CA 94608 (US); ENGELBRECHT, Susan (ZA/ZA); c/o University of Stellenbosch, P.O. Box 19063, 7505 Tygerberg (ZA); VAN RENSBURG, Estrelita, Janse (ZA/ZA); c/o University of Stellenbosch, P.O. Box 19063, 7505 Tygerberg (ZA).

(74) Agents: DOLLARD, Anne, S. et al., Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662 8097 (US).

(81) Designated States (national): AF, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GG, GL, GM, HR, HU, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LI, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NG, NZ, PA, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW); Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM); European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LI, MC, NL, PT, SE, TR); OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette



WO 02/04493 A2

(54) Title: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

(57) Abstract: The present invention relates to polynucleotides encoding immunogenic HIV type C polypeptides. Uses of the polynucleotides in applications including DNA immunization, generation of packaging cell lines, and production of HIV Type C proteins are also described.

WO 02/04493

PCT/US01/21241

**POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C POLYPEPTIDES,
POLYPEPTIDES AND USES THEREOF**

TECHNICAL FIELD

5 Polynucleotides encoding antigenic Type C HIV polypeptides (e.g., Gag, pol, vif, vpr, tat, rev, vpu, env, and nef) are described, as are uses of these polynucleotides and polypeptide products in immunogenic compositions. Also described are polynucleotide sequences from South African variants of HIV Type C.

10 **BACKGROUND OF THE INVENTION**

Acquired immune deficiency syndrome (AIDS) is recognized as one of the greatest health threats facing modern medicine. There is, as yet, no cure for this disease. In 1983-1984, three groups independently identified the suspected etiological agent of AIDS. See, e.g., Barre-Sinoussi et al. (1983) *Science* 220:868-871; Montagnier et al., in *Human T-Cell Leukemia Viruses* (Gallo, Essex & Gross, eds., 1984); Vilmer et al. (1984) *The Lancet* 1:753; Popovic et al. (1984) *Science* 224:497-500; Levy et al. (1984) *Science* 225:840-842. These isolates were variously called lymphadenopathy-associated virus (LAV), human T-cell lymphotropic virus type III (HTLV-III), or AIDS-associated retrovirus (ARV). All of these isolates are strains of the same virus, and were later collectively named Human Immunodeficiency Virus (HIV). With the isolation of a related AIDS-causing virus, the strains originally called HIV are now termed HIV-1 and the related virus is called HIV-2 See, e.g., Guyader et al. (1987) *Nature* 326:662-669; Brun-Vezinet et al. (1986) *Science* 233:343-346; Clavel et al. (1986) *Nature* 324:691-695.

A great deal of information has been gathered about the HIV virus, however, to date an effective vaccine has not been identified. Several targets for vaccine development have been examined including the *env* and *Gag* gene products encoded by HIV. *Gag* gene products include, but are not limited to, Gag-polymerase and Gag-protease. *Env* gene products include, but are not limited to, monomeric gp120 polypeptides, oligomeric gp140 polypeptides and gp160 polypeptides.

30 Haas, et al., (*Current Biology* 6(3):315-324, 1996) suggested that selective codon usage by HIV-1 appeared to account for a substantial fraction of the inefficiency of viral protein synthesis. Andre, et al., (*J. Virol.* 72(2):1497-1503, 1998) described an increased

WO 02/04493

PCT/US01/21241

immune response elicited by DNA vaccination employing a synthetic gp120 sequence with modified codon usage. Schneider, et al., (*J Virol.* 71(7):4892-4903, 1997) discuss inactivation of inhibitory (or instability) elements (INS) located within the coding sequences of the Gag and Gag-protease coding sequences.

- 5 The *Gag* proteins of HIV-1 are necessary for the assembly of virus-like particles. HIV-1 *Gag* proteins are involved in many stages of the life cycle of the virus including, assembly, virion maturation after particle release, and early post-entry steps in virus replication. The roles of HIV-1 *Gag* proteins are numerous and complex (Freed, E.O., *Virology* 251:1-15, 1998).
- 10 Wolf, et al., (PCT International Application, WO 96/30523, published 3 October 1996; European Patent Application, Publication No. 0 449 116 A1, published 2 October 1991) have described the use of altered p55 *Gag* of HIV-1 to act as a non-infectious retroviral-like particulate carrier, in particular, for the presentation of immunologically important epitopes. Wang, et al., (*Virology* 200:524-534, 1994) describe a system to study
- 15 assembly of HIV *Gag*- β -galactosidase fusion proteins into virions. They describe the construction of sequences encoding HIV *Gag*- β -galactosidase fusion proteins, the expression of such sequences in the presence of HIV *Gag* proteins, and assembly of these proteins into virus particles.
- 20 Shiver, et al., (PCT International Application, WO 98/34640, published 13 August 1998) described altering HIV-1 (CAM1) *Gag* coding sequences to produce synthetic DNA molecules encoding HIV *Gag* and modifications of HIV *Gag*. The codons of the synthetic molecules were codons preferred by a projected host cell.
- 25 Recently, use of HIV *Env* polypeptides in immunogenic compositions has been described. (see, U.S. Patent No. 5,846,546 to Hurwitz et al., issued December 8, 1998, describing immunogenic compositions comprising a mixture of at least four different recombinant virus that each express a different HIV *env* variant; and U.S. Patent No. 5,840,313 to Vahne et al., issued November 24, 1998, describing peptides which correspond to epitopes of the HIV-1 gp120 protein). In addition, U.S. Patent No. 5,876,731 to Sia et al., issued March 2, 1999 describes candidate vaccines against HIV comprising an amino acid
- 30 sequence of a T-cell epitope of *Gag* linked directly to an amino acid sequence of a B-cell epitope of the V3 loop protein of an HIV-1 isolate containing the sequence GPGR. There remains a need for antigenic HIV polypeptides, particularly Type C isolates.

WO 02/04493

PCT/US01/21241

SUMMARY OF THE INVENTION

Described herein are novel Type C HIV sequences, for example, 8_5_TV1_C.ZA, 8_2_TV1_C.ZA and 12-5_1_TV2_C.ZA, polypeptides encoded by these novel sequences, and synthetic expression cassettes generated from these and other Type C HIV sequences.

In certain embodiments, the present invention relates synthetic expression cassettes encoding HIV Type C polypeptides, including Env, Gag, Pol, Prot, Vpr, Vpu, Vif, Nef, Tat, Rev and/or fragments thereof. In addition, the present invention also relates to improved expression of HIV Type C polypeptides and production of virus-like particles. Synthetic expression cassettes encoding the HIV polypeptides (*e.g.*, Gag-, pol-, protease (prot)-, reverse transcriptase, integrase, RNaseH, Tat, Rev, Nef, Vpr, Vpu, Vif and/or Env- containing polypeptides) are described, as are uses of the expression cassettes.

Thus, one aspect of the present invention relates to expression cassettes and polynucleotides contained therein. The expression cassettes typically include an HIV- polypeptide encoding sequence inserted into an expression vector backbone. In one embodiment, an expression cassette comprises a polynucleotide sequence encoding one or more *Pol*-containing polypeptides, wherein the polynucleotide sequence comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and more preferably about 98% sequence (and any integers between these values) identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Pol*-containing polypeptides include, but are not limited to, those shown in SEQ ID NO:30, SEQ ID NO:31; SEQ ID NO:32; SEQ ID NO:62; SEQ ID NO:103; SEQ ID NO:58; SEQ ID NO:60; SEQ ID NO:64; SEQ ID NO:66; SEQ ID NO:68; SEQ ID NO:70; SEQ ID NO:76; and SEQ ID NO:78.

The polynucleotides encoding the HIV polypeptides of the present invention may also include sequences encoding additional polypeptides. Such additional polynucleotides encoding polypeptides may include, for example, coding sequences for other viral proteins (*e.g.*, hepatitis B or C or other HIV proteins, such as, polynucleotide sequences encoding an HIV *Gag* polypeptide, polynucleotide sequences encoding an HIV *Env* polypeptide and/or polynucleotides encoding one or more of vif, vpr, tat, rev, vpu and nef); cytokines or other transgenes. In one embodiment, the sequence encoding the HIV *Pol* polypeptide(s) can be modified by deletions of coding regions corresponding to reverse transcriptase and integrase.

WO 02/04493

PCT/US01/21241

Such deletions in the polymerase polypeptide can also be made such that the polynucleotide sequence preserves T-helper cell and CTL epitopes. Other antigens of interest may be inserted into the polymerase as well.

In another embodiment, an expression cassette comprises a polynucleotide sequence encoding a polypeptide including an HIV *Gag*-containing polypeptide, wherein the polynucleotide sequence encoding the *Gag* polypeptide comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Gag*-containing polypeptides include, but are not limited to, the following polynucleotides: nucleotides 844-903 of Figure 1 (a *Gag* major homology region) (SEQ ID NO:1); nucleotides 841-900 of Figure 2 (a *Gag* major homology region) (SEQ ID NO:2); Figure 24 (SEQ ID NO:53, a *Gag* major homology region); the sequence presented as Figure 1 (SEQ ID NO:3); the sequence presented as Figure 22 (SEQ ID NO:51); the sequence presented as Figure 70 (SEQ ID NO:99); and the sequence presented as Figure 2 (SEQ ID NO:4). As noted above, the polynucleotides encoding the *Gag*-containing polypeptides of the present invention may also include sequences encoding additional polypeptides.

In another embodiment, an expression cassette comprises a polynucleotide sequence encoding a polypeptide including an HIV *Env*-containing polypeptide, wherein the polynucleotide sequence encoding the *Env* polypeptide comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Env*-containing polypeptides include, but are not limited to, the following polynucleotides: nucleotides 1213-1353 of Figure 3 (SEQ ID NO:5) (encoding an *Env* common region); the sequence presented as Figure 17 (SEQ ID NO:46) (encoding a 97 nucleotide long *Env* common region); SEQ ID NO:47 (encoding a 144 nucleotide long *Env* common region); nucleotides 82-1512 of Figure 3 (SEQ ID NO:6) (encoding a gp120 polypeptide); nucleotides 82-2025 of Figure 3 (SEQ ID NO:7) (encoding a gp140 polypeptide); nucleotides 82-2547 of Figure 3 (SEQ ID NO:8) (encoding a gp160 polypeptide); SEQ ID NO:49 (encoding a gp160 polypeptide); nucleotides 1-2547 of Figure 3 (SEQ ID NO:9) (encoding a gp160 polypeptide with signal sequence); nucleotides 1513-2547 of Figure 3 (SEQ ID NO:10) (encoding a gp41 polypeptide); nucleotides 1210-1353 of

WO 02/04493

PCT/US01/21241

Figure 4 (SEQ ID NO:11) (encoding an Env common region); nucleotides 73-1509 of Figure 4 (SEQ ID NO:12) (encoding a gp120 polypeptide); nucleotides 73-2022 of Figure 4 (SEQ ID NO:13) (encoding a gp140 polypeptide); nucleotides 73-2565 of Figure 4 (SEQ ID NO:14) (encoding a gp160 polypeptide); nucleotides 1-2565 of Figure 4 (SEQ ID NO:15) (encoding a gp160 polypeptide with signal sequence); the sequence presented as Figure 20 (SEQ ID NO:49) (encoding a gp160 polypeptide); the sequence presented as Figure 68 (SEQ ID NO:97) (encoding a gp160 polypeptide); nucleotides 1510-2565 of Figure 4 (SEQ ID NO:16) (encoding a gp41 polypeptide); nucleotides 7 to 1464 of Figure 90 (SEQ ID NO:119) (encoding a gp120 polypeptide with modified wild type signal sequence); nucleotides 7 to 1977 of Figure 91 (SEQ ID NO:120) (encoding a gp140 polypeptide including signal sequence modified from wild-type 8_2_TV1_C.ZA (*e.g.*, "modified wild type leader sequence")); nucleotides 7 to 1977 of Figure 92 (SEQ ID NO:121) (encoding a gp140 polypeptide with modified wild type 8_2_TV1_C.ZA signal sequence); nucleotides 7 to 2388 of Figure 93 (SEQ ID NO:122) (encoding a gp160 polypeptide with modified wild type signal sequence); nucleotides 7 to 2520 of Figure 94 (SEQ ID NO:123) (encoding a gp160 polypeptide with modified wild type 8_2_TV1_C.ZA signal sequence); nucleotides 7 to 2520 of Figure 95 (SEQ ID NO:124) (encoding a gp160 polypeptide with modified wild type 8_2_TV1_C.ZA signal sequence); nucleotides 13 to 2604 of Figure 96 (SEQ ID NO:125) (encoding a gp160 polypeptide with TPA1 signal sequence); nucleotides 7 to 2607 of Figure 97 (SEQ ID NO:126) (encoding a gp160 polypeptide with modified wild type 8_2_TV1_C.ZA signal sequence); nucleotides 1 to 2049 of Figure 100 (SEQ ID NO:131) (encoding a gp140 polypeptide with TPA1 signal sequence); nucleotides 7 to 1607 of Figure 98 (SEQ ID NO:126) (encoding a gp160 polypeptide with wild type 8_2_TV1_C.ZA signal sequence); nucleotides 7 to 2064 of SEQ ID NO:132 (encoding a gp140 polypeptide with modified wild-type 8_2_TV1_C.ZA leader sequence); and nucleotides 7 to 2064 of SEQ ID NO:133 (encoding a gp140 polypeptide with wild-type 8_2_TV1_C.ZA leader sequence).

In certain embodiments, the Env-encoding sequences will contain further modifications, for instance mutation of the cleavage site to prevent the cleavage of a gp140 polypeptide into a gp120 polypeptide and a gp41 polypeptide (SEQ ID NO:121 and SEQ ID NO:124) or deletion of variable regions V1 and/or V2 (SEQ ID NO:119; SEQ ID NO:120; SEQ ID NO:121; SEQ ID NO:122; SEQ ID NO:123; and SEQ ID NO:124).

WO 02/04493

PCT/US01/21241

In another embodiment, an expression cassette comprises a polynucleotide sequence encoding a polypeptide including an HIV *Nef*-containing polypeptide, wherein the polynucleotide sequence encoding the *Nef* polypeptide comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Nef*-containing polypeptides include, but are not limited to, the following polynucleotides: the sequence presented in Figure 26 (SEQ ID NO:55); the sequence presented in Figure 72 (SEQ ID NO:101); the sequence presented in Figure 28 (SEQ ID NO:57); the sequence presented in Figure 67 (SEQ ID NO:96); the sequence presented in Figure 103 (SEQ ID NO:134); and the sequence presented in Figure 104 (SEQ ID NO:135).

In another embodiment, an expression cassette comprises a polynucleotide sequence encoding a polypeptide including an HIV *Rev*-containing polypeptide, wherein the polynucleotide sequence encoding the *Rev* polypeptide comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Rev*-containing polypeptides include, but are not limited to, the following polynucleotides: the sequence presented in Figure 43 (SEQ ID NO:72); the sequence presented in Figure 76 (SEQ ID NO:105); the sequence presented in Figure 45 (SEQ ID NO:74); the sequence presented in Figure 78 (SEQ ID NO:107); and the sequence presented in Figure 62 (SEQ ID NO:91).

In another embodiment, an expression cassette comprises a polynucleotide sequence encoding a polypeptide including an HIV *Tat*-containing polypeptide, wherein the polynucleotide sequence encoding the *Tat* polypeptide comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Tat*-containing polypeptides include, but are not limited to, the following polynucleotides: the sequence presented in Figure 51 (SEQ ID NO:80); the sequence presented in Figure 80 (SEQ ID NO:109); the sequence presented in Figure 52 (SEQ ID NO:81); the sequence presented in Figure 54 (SEQ ID NO:83); and the sequence presented in Figure 82 (SEQ ID NO:111).

WO 02/04493

PCT/US01/21241

In another embodiment, an expression cassette comprises a polynucleotide sequence encoding a polypeptide including an HIV *Vif*-containing polypeptide, wherein the polynucleotide sequence encoding the *Vif* polypeptide comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Vif*-containing polypeptides include, but are not limited to, the following polynucleotides: the sequence presented in Figure 56 (SEQ ID NO:85); and the sequence presented in Figure 84 (SEQ ID NO:113).

In another embodiment, an expression cassette comprises a polynucleotide sequence encoding a polypeptide including an HIV *Vpr*-containing polypeptide, wherein the polynucleotide sequence encoding the *Vpr* polypeptide comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Vpr*-containing polypeptides include, but are not limited to, the following polynucleotides: the sequence presented in Figure 58 (SEQ ID NO:87); and the sequence presented in Figure 86 (SEQ ID NO:115).

In another embodiment, an expression cassette comprises a polynucleotide sequence encoding a polypeptide including an HIV *Vpu*-containing polypeptide, wherein the polynucleotide sequence encoding the *Vpu* polypeptide comprises a sequence having at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught in the present specification. The polynucleotide sequences encoding *Vpu*-containing polypeptides include, but are not limited to, the following polynucleotides: the sequence presented in Figure 60 (SEQ ID NO:89); and the sequence presented in Figure 88 (SEQ ID NO:117).

Further embodiments of the present invention include purified polynucleotides of any of the sequences described herein. Exemplary polynucleotide sequences encoding *Gag*-containing polypeptides include, but are not limited to, the following polynucleotides: nucleotides 844-903 of Figure 1 (SEQ ID NO:1) (a *Gag* major homology region); nucleotides 841-900 of Figure 2 (SEQ ID NO:2) (a *Gag* major homology region); the sequence presented as Figure 1 (SEQ ID NO:3); the sequence presented as Figure 2 (SEQ ID NO:4); the sequence presented as Figure 22 (SEQ ID NO:51); the sequence presented as Figure 70 (SEQ

WO 02/04493

PCT/US01/21241

ID NO:99); and the sequence presented as Figure 24 (SEQ ID NO:53) (a Gag major homology region).

Exemplary polynucleotide sequences encoding *Env*-containing polypeptides include, but are not limited to, the following polynucleotides: nucleotides 1213-1353 of Figure 3 (SEQ ID NO:5) (encoding an *Env* common region); the sequence presented as Figure 17 (SEQ ID NO:46) (encoding a 97 nucleotide long *Env* common region); SEQ ID NO:47 (encoding a 144 nucleotide long *Env* common region); nucleotides 82-1512 of Figure 3 (SEQ ID NO:6) (encoding a gp120 polypeptide); nucleotides 82-2025 of Figure 3 (SEQ ID NO:7) (encoding a gp140 polypeptide); nucleotides 82-2547 of Figure 3 (SEQ ID NO:8) (encoding a gp160 polypeptide); SEQ ID NO:49 (encoding a gp160 polypeptide); nucleotides 1-2547 of Figure 3 (SEQ ID NO:9) (encoding a gp160 polypeptide with signal sequence); nucleotides 1513-2547 of Figure 3 (SEQ ID NO:10) (encoding a gp41 polypeptide); nucleotides 1210-1353 of Figure 4 (SEQ ID NO:11) (encoding an *Env* common region); nucleotides 73-1509 of Figure 4 (SEQ ID NO:12) (encoding a gp120 polypeptide); nucleotides 73-2022 of Figure 4 (SEQ ID NO:13) (encoding a gp140 polypeptide); nucleotides 73-2565 of Figure 4 (SEQ ID NO:14) (encoding a gp160 polypeptide); nucleotides 1-2565 of Figure 4 (SEQ ID NO:15) (encoding a gp160 polypeptide with signal sequence); the sequence presented as Figure 20 (SEQ ID NO:49) (encoding a gp160 polypeptide); the sequence presented as Figure 68 (SEQ ID NO:97) (encoding a gp160 polypeptide); nucleotides 1510-2565 of Figure 4 (SEQ ID NO:16) (encoding a gp41 polypeptide); nucleotides 7 to 1464 of Figure 90 (SEQ ID NO:119) (encoding a gp120 polypeptide with modified wild type signal sequence); nucleotides 7 to 1977 of Figure 91 (SEQ ID NO:120) (encoding a gp140 polypeptide including signal sequence modified from wild-type 8_2_TV1_C.ZA (*e.g.*, "modified wild type leader sequence")); nucleotides 7 to 1977 of Figure 92 (SEQ ID NO:121) (encoding a gp140 polypeptide with modified wild type 8_2_TV1_C.ZA signal sequence); nucleotides 7 to 2388 of Figure 93 (SEQ ID NO:122) (encoding a gp160 polypeptide with modified wild type signal sequence); nucleotides 7 to 2520 of Figure 94 (SEQ ID NO:123) (encoding a gp160 polypeptide with modified wild type 8_2_TV1_C.ZA signal sequence); nucleotides 7 to 2520 of Figure 95 (SEQ ID NO:124) (encoding a gp160 polypeptide with modified wild type 8_2_TV1_C.ZA signal sequence); nucleotides 13 to 2604 of Figure 96 (SEQ ID NO:125) (encoding a gp160 polypeptide with TPA1 signal sequence); nucleotides 7 to 2607 of Figure 97 (SEQ ID NO:126) (encoding a gp160 polypeptide with modified wild type

WO 02/04493

PCT/US01/21241

8_2_TV1_C.ZA signal sequence); nucleotides 1 to 2049 of Figure 100 (SEQ ID NO:131) (encoding a gp140 polypeptide with TPA1 signal sequence); nucleotides 7 to 1607 of Figure 98 (SEQ ID NO:126) (encoding a gp160 polypeptide with wild type 8_2_TV1_C.ZA signal sequence); nucleotides 7 to 2064 of SEQ ID NO:132 (encoding a gp140 polypeptide with modified wild-type 8_2_TV1_C.ZA leader sequence); and nucleotides 7 to 2064 of SEQ ID NO:133 (encoding a gp140 polypeptide with wild-type 8_2_TV1_C.ZA leader sequence).

Exemplary purified polynucleotides encoding additional HIV polynucleotides include: Pol-encoding polynucleotides (e.g., SEQ ID NO:30; SEQ ID NO:31; SEQ ID NO:32; SEQ ID NO:62; SEQ ID NO:103; SEQ ID NO:58; SEQ ID NO:60; SEQ ID NO:64; SEQ ID NO:66; SEQ ID NO:68; SEQ ID NO:70; SEQ ID NO:76; and SEQ ID NO:78); Nef-encoding polynucleotides (e.g., SEQ ID NO:55; SEQ ID NO:101; SEQ ID NO:57; SEQ ID NO:96); Rev-encoding polynucleotides (e.g., SEQ ID NO:72; SEQ ID NO:105; SEQ ID NO:74; SEQ ID NO:107; SEQ ID NO:91); Tat-encoding polynucleotides (e.g., SEQ ID NO:80; SEQ ID NO:109; SEQ ID NO:81; SEQ ID NO:83; SEQ ID NO:111); Vif-encoding polynucleotides (e.g., SEQ ID NO:85; SEQ ID NO:113); and Vpr-encoding polynucleotides (e.g., SEQ ID NO:87; SEQ ID NO:115); Vpu-encoding polynucleotides (e.g., SEQ ID NO:89; SEQ ID NO:117).

In other embodiments, the present invention relates to native HIV polypeptide-encoding sequences obtained from novel Type C strains; fragments of those native sequences; expression cassettes containing those wild-type sequences; and uses of these sequences, fragments and expression cassettes. Exemplary full length sequences are shown in SEQ ID NO:33 and SEQ ID NO:45. Exemplary fragments coding for various HIV gene products include: the sequence presented in Figure 19 (SEQ ID NO:48) (an Env-encoding sequence); the sequence presented in Figure 69 (SEQ ID NO:98) (an Env-encoding sequence); the sequence presented in Figure 21 (SEQ ID NO:50) (a gp160 polypeptide); the sequence presented in Figure 23 (SEQ ID NO:52) (a Gag polypeptide); the sequence presented in Figure 71 (SEQ ID NO:100) (a Gag polypeptide); the sequence presented in Figure 25 (SEQ ID NO:54) (a Gag polypeptide); the sequence presented in Figure 27 (SEQ ID NO:56) (a Nef polypeptide); the sequence presented in Figure 73 (SEQ ID NO:102) (a Nef polypeptide); the sequence presented in Figure 30 (SEQ ID NO:59) (a p15RNAseH polypeptide); the sequence presented in Figure 32 (SEQ ID NO:61) (a p31Integrase polypeptide); the sequence presented in Figure 34 (SEQ ID NO:63) (a Pol polypeptide); the sequence presented in Figure 75 (SEQ

WO 02/04493

PCT/US01/21241

ID NO:104) (a Pol polypeptide); the sequence presented in Figure 36 (SEQ ID NO:65) (a Prot polypeptide); the sequence presented in Figure 38 (SEQ ID NO:67) (a inactivated Prot polypeptide); the sequence presented in Figure 40 (SEQ ID NO:69) (an inactivated Prot and RT polypeptide); the sequence presented in Figure 42 (SEQ ID NO:71) (a Prot and RT polypeptide); the sequence presented in Figure 44 (SEQ ID NO:73) (a Rev polypeptide); the sequence presented in Figure 77 (SEQ ID NO:106) (a Rev polypeptide); the sequence presented in Figure 46 (SEQ ID NO:75) (a Rev polypeptide); the sequence presented in Figure 79 (SEQ ID NO:108) (a Rev polypeptide); the sequence presented in Figure 48 (SEQ ID NO:77) (an RT polypeptide); the sequence presented in Figure 50 (SEQ ID NO:79) (a mutated RT polypeptide); the sequence presented in Figure 53 (SEQ ID NO:82) (a Tat polypeptide); the sequence presented in Figure 81 (SEQ ID NO:110) (a Tat polypeptide); the sequence presented in Figure 55 (SEQ ID NO:84) (a Tat polypeptide); the sequence presented in Figure 83 (SEQ ID NO:112) (a Tat polypeptide); the sequence presented in Figure 57 (SEQ ID NO:86) (a Vif polypeptide); the sequence presented in Figure 85 (SEQ ID NO:114) (a Vif polypeptide); the sequence presented in Figure 59 (SEQ ID NO:88) (a Vpr polypeptide); the sequence presented in Figure 82 (SEQ ID NO:116) (a Vpr polypeptide); the sequence presented in Figure 61 (SEQ ID NO:90) (a Vpu polypeptide); the sequence presented in Figure 89 (SEQ ID NO:118) (a Vpu polypeptide); the sequence presented in Figure 63 (SEQ ID NO:92) (a Rev polypeptide); and the sequence presented in Figure 66 (SEQ ID NO:95) (a Tat polypeptide).

The native and synthetic polynucleotide sequences encoding the HIV polypeptides of the present invention typically have at least about 85%, preferably about 90%, more preferably about 95%, and most preferably about 98% sequence identity to the sequences taught herein. Further, in certain embodiments, the polynucleotide sequences encoding the HIV polypeptides of the invention will exhibit 100% sequence identity to the sequences taught herein.

The polynucleotides of the present invention can be produced by recombinant techniques, synthetic techniques, or combinations thereof.

The present invention further includes recombinant expression systems for use in selected host cells, wherein the recombinant expression systems employ one or more of the polynucleotides and expression cassettes of the present invention. In such systems, the polynucleotide sequences are operably linked to control elements compatible with expression

WO 02/04493

PCT/US01/21241

in the selected host cell. Numerous expression control elements are known to those in the art, including, but not limited to, the following: transcription promoters, transcription enhancer elements, transcription termination signals, polyadenylation sequences, sequences for optimization of initiation of translation, and translation termination sequences. Exemplary transcription promoters include, but are not limited to those derived from CMV, CMV+intron A, SV40, RSV, HIV-Ltr, MMLV-Ltr, and metallothionein.

In another aspect the invention includes cells comprising one or more of the expression cassettes of the present invention where the polynucleotide sequences are operably linked to control elements compatible with expression in the selected cell. In one embodiment such cells are mammalian cells. Exemplary mammalian cells include, but are not limited to, BHK, VERO, HT1080, 293, RD, COS-7, and CHO cells. Other cells, cell types, tissue types, etc., that may be useful in the practice of the present invention include, but are not limited to, those obtained from the following: insects (e.g., *Trichoplusia ni* (Tn5) and Sf9), bacteria, yeast, plants, antigen presenting cells (e.g., macrophage, monocytes, dendritic cells, B-cells, T-cells, stem cells, and progenitor cells thereof), primary cells, immortalized cells, tumor-derived cells.

In a further aspect, the present invention includes compositions for generating an immunological response, where the composition typically comprises at least one of the expression cassettes of the present invention and may, for example, contain combinations of expression cassettes (such as one or more expression cassettes carrying a Pol-polypeptide-encoding polynucleotide, one or more expression cassettes carrying a Gag-polypeptide-encoding polynucleotide, one or more expression cassettes carrying accessory polypeptide-encoding polynucleotides (e.g., native or synthetic vpu, vpr, nef, vif, tat, rev), and/or one or more expression cassettes carrying an Env-polypeptide-encoding polynucleotide). Such compositions may further contain an adjuvant or adjuvants. The compositions may also contain one or more Type C HIV polypeptides. The Type C HIV polypeptides may correspond to the polypeptides encoded by the expression cassette(s) in the composition, or may be different from those encoded by the expression cassettes. An example of the polynucleotide in the expression cassette encoding the same polypeptide as is being provided in the composition is as follows: the polynucleotide in the expression cassette encodes the Gag-polypeptide of Figure 1 (SEQ ID NO:3), and the polypeptide (SEQ ID NO:17) is the polypeptide encoded by the sequence shown in Figure 1. An example of the polynucleotide in

WO 02/04493

PCT/US01/21241

the expression cassette encoding a different polypeptide as is being provided in the composition is as follows: an expression cassette having a polynucleotide encoding a Gag-polymerase polypeptide, and the polypeptide provided in the composition may be a Gag and/or Gag-protease polypeptide. In compositions containing both expression cassettes (or polynucleotides of the present invention) and polypeptides, various expression cassettes of the present invention can be mixed and/or matched with various Type C HIV polypeptides described herein.

In another aspect the present invention includes methods of immunization of a subject. In the method any of the above described compositions are into the subject under conditions that are compatible with expression of the expression cassette(s) in the subject. In one embodiment, the expression cassettes (or polynucleotides of the present invention) can be introduced using a gene delivery vector. The gene delivery vector can, for example, be a non-viral vector or a viral vector. Exemplary viral vectors include, but are not limited to Sindbis-virus derived vectors, retroviral vectors, and lentiviral vectors. Compositions useful for generating an immunological response can also be delivered using a particulate carrier. Further, such compositions can be coated on, for example, gold or tungsten particles and the coated particles delivered to the subject using, for example, a gene gun. The compositions can also be formulated as liposomes. In one embodiment of this method, the subject is a mammal and can, for example, be a human.

In a further aspect, the invention includes methods of generating an immune response in a subject. Any of the expression cassettes described herein can be expressed in a suitable cell to provide for the expression of the Type C HIV polypeptides encoded by the polynucleotides of the present invention. The polypeptide(s) are then isolated (e.g., substantially purified) and administered to the subject in an amount sufficient to elicit an immune response. In certain embodiments, the methods comprise administration of one or more of the expression cassettes or polynucleotides of the present invention, using any of the gene delivery techniques described herein. In other embodiments, the methods comprise co-administration of one or more of the expression cassettes or polynucleotides of the present invention and one or more polypeptides, wherein the polypeptides can be expressed from these polynucleotides or can be other subtype C HIV polypeptides. In other embodiments, the methods comprise co-administration of multiple expression cassettes or polynucleotides of the present invention. In still further embodiments, the methods comprise co-

WO 02/04493

PCT/US01/21241

administration of multiple polypeptides, for example polypeptides expressed from the polynucleotides of the present invention and/or other subtype C HIV polypeptides.

The invention further includes methods of generating an immune response in a subject, where cells of a subject are transfected with any of the above-described expression cassettes or polynucleotides of the present invention, under conditions that permit the expression of a selected polynucleotide and production of a polypeptide of interest (e.g., encoded by any expression cassette of the present invention). By this method an immunological response to the polypeptide is elicited in the subject. Transfection of the cells may be performed *ex vivo* and the transfected cells are reintroduced into the subject.

Alternately, or in addition, the cells may be transfected *in vivo* in the subject. The immune response may be humoral and/or cell-mediated (cellular). In a further embodiment, this method may also include administration of an Type C HIV polypeptides before, concurrently with, and/or after introduction of the expression cassette into the subject.

These and other embodiments of the present invention will readily occur to those of ordinary skill in the art in view of the disclosure herein.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 (SEQ ID NO:3) shows the nucleotide sequence of a polynucleotide encoding a synthetic Gag polypeptide. The nucleotide sequence shown was obtained by modifying type C strain AF110965 and include further modifications of INS.

Figure 2 (SEQ ID NO: 4) shows the nucleotide sequence of a polynucleotide encoding a synthetic Gag polypeptide. The nucleotide sequence shown was obtained by modifying type C strain AF110967 and include further modifications of INS.

Figure 3 (SEQ ID NO:9) shows the nucleotide sequence of a polynucleotide encoding a synthetic Env polypeptide. The nucleotide sequence depicts gp160 (including a signal peptide) and was obtained by modifying type C strain AF110968. The arrows indicate the positions of various regions of the polynucleotide, including the sequence encoding a signal peptide (nucleotides 1-81) (SEQ ID NO:18), a gp120 polypeptide (nucleotides 82-1512) (SEQ ID NO:6), a gp41 polypeptide (nucleotides 1513-2547) (SEQ ID NO:10), a gp140 polypeptide (nucleotides 82-2025) (SEQ ID NO:7) and a gp160 polypeptide (nucleotides 82-2547) (SEQ ID NO:8). The codons encoding the signal peptide are modified (as described herein) from the native HIV-1 signal sequence.

WO 02/04493

PCT/US01/21241

Figure 4 (SEQ ID NO:15) shows the nucleotide sequence of a polynucleotide encoding a synthetic Env polypeptide. The nucleotide sequence depicts gp160 (including a signal peptide) and was obtained by modifying type C strain AF110975. The arrows indicate the positions of various regions of the polynucleotide, including the sequence encoding a signal peptide (nucleotides 1-72) (SEQ ID NO:19), a gp120 polypeptide (nucleotides 73-1509) (SEQ ID NO:12), a gp41 polypeptide (nucleotides 1510-2565) (SEQ ID NO:16), a gp140 polypeptide (nucleotides 73-2022) (SEQ ID NO:13), and a gp160 polypeptide (nucleotides 73-2565) (SEQ ID NO:14). The codons encoding the signal peptide are modified (as described herein) from the native HIV-1 signal sequence.

Figure 5 shows the location of some remaining INS in synthetic Gag sequences derived from AF110965. The changes made to these sequences are boxed in the Figures. The top line depicts a codon modified sequence of Gag polypeptides from the indicated strains (SEQ ID NO:20). The nucleotide(s) appearing below the line in the boxed region(s) depicts changes made to remove further INS and correspond to the sequence depicted in Figure 1 (SEQ ID NO:3).

Figure 6 shows the location of some remaining INS in synthetic Gag sequences derived from AF110967. The changes made to these sequences are boxed in the Figures. The top line depicts a modified sequence of Gag polypeptides from the indicated strains (SEQ ID NO:21). The nucleotide(s) appearing below the line in the boxed region(s) depicts changes made to remove further INS and correspond to the sequence depicted in Figure 2 (SEQ ID NO:4).

Figure 7 is a schematic depicting the selected domains in the *Pol* region of HIV.

Figure 8 (SEQ ID NO:30) depicts the nucleotide sequence of the synthetic construct designated PR975(+). "+" indicates that the reverse transcriptase is functional. This construct includes sequence from p2 (nucleotides 16 to 54 of SEQ ID NO:30); p7 (nucleotides 55 to 219 of SEQ ID NO:30); p1/p6 (nucleotides 220-375 of SEQ ID NO:30); prot (nucleotides 376 to 672 of SEQ ID NO:30), reverse transcriptase (nucleotides 673 to 2352 of SEQ ID NO:30); and 6 amino acids of integrase shown in Figure 7 (nucleotides 2353 to 2370 of SEQ ID NO:30). In addition, the construct contains a multiple cloning site (MCS, nucleotides 2425 to 2463 of SEQ ID NO:30) for insertion of a transgene and a YMDD epitope cassette (nucleotides 2371 to 2424 of SEQ ID NO:30).

WO 02/04493

PCT/US01/21241

Figure 9 (SEQ ID NO:31) depicts the nucleotide sequence of the synthetic construct designated PR975YM. As illustrated in Figure 7, the RT region includes a mutation in the catalytic center (mut. cat. center). "YM" refers to constructs in which the nucleotides encode the amino acids AP instead of YMDD in this region. Reverse transcriptase is not functional in this construct. This construct includes sequence from the p2 (nucleotides 16 to 54 of SEQ ID NO:31); p7 (nucleotides 55 to 219 of SEQ ID NO:31); p1/p6 (nucleotides 220 to 375 of SEQ ID NO:31); prot (nucleotides 376 to 672 of SEQ ID NO:31); and reverse transcriptase (nucleotides 673 to 2346 of SEQ ID NO:31) shown in Figure 7, although the reverse transcriptase protein is not functional. In addition, the construct contains a multiple cloning site (MCS, nucleotides 2419 to 2457 of SEQ ID NO:31) for insertion of a transgene and a YMDD epitope cassette (nucleotides 2365 to 2418 of SEQ ID NO:31).

Figure 10 (SEQ ID NO:32) depicts the nucleotide sequence of the synthetic construct designated PR975YMWM. "YM" refers to constructs in which the nucleotides encode the amino acids AP instead of YMDD in this region. "WM" refers to constructs in which the nucleotides encode amino acids PI instead of WMGY in this region. This construct includes sequence from the p2 (nucleotides 16 to 54 of SEQ ID NO:32); p7 (nucleotides 55 to 219 of SEQ ID NO:32); p1/p6 (nucleotides 220 to 375 of SEQ ID NO:32); prot (nucleotides 376 to 672 of SEQ ID NO:32); and reverse transcriptase (nucleotides 673 to 2340 of SEQ ID NO:32) shown in Figure 7, although the reverse transcriptase protein is not functional. In addition, the construct contains a multiple cloning site (MCS, nucleotides 2413 to 2451 of SEQ ID NO:32) for insertion of a transgene and a YMDD epitope cassette (nucleotides 2359 to 2412 of SEQ ID NO:32).

Figure 11 (SEQ ID NO:33) depicts the nucleotide sequence of 8_5_TV1_C.ZA. Various regions are shown in Table A.

Figure 12 (SEQ ID NO:34) depicts the wild type nucleotide sequence of AF110975 Pol from p2gag until p7gag.

Figure 13 (SEQ ID NO:35) depicts the wild type nucleotide sequence of AF110975 Pol from p1 through the first 6 amino acids of the integrase protein.

Figure 14 (SEQ ID NO:36) depicts the nucleotide sequence of a cassette encoding Ile178 through Serine 191 of reverse transcriptase.

Figure 15 (SEQ ID NO:37) shows amino acid sequence which includes an epitope in the region of the catalytic center of the reverse transcriptase protein.

WO 02/04493

PCT/US01/21241

Figure 16 (SEQ ID NO:45) depicts the nucleotide sequence of 12-5_1_TV2_C.ZA.

Figure 17 (SEQ ID NO:46) depicts the nucleotide sequence of a synthetic Env-encoding polynucleotide derived from 8_5_TV1_C.ZA. The sequence corresponds to a short (97 base pair) common region.

5 Figure 18 (SEQ ID NO:47) depicts the nucleotide sequence of a synthetic Env-encoding polynucleotide derived from 8_5_TV1_C.ZA. The sequence corresponds to a common region in Env.

Figure 19 (SEQ ID NO:48) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Env.

10 Figure 20 (SEQ ID NO:49) depicts the nucleotide sequence of a synthetic Env gp160-encoding polynucleotide derived from 8_5_TV1_C.ZA.

Figure 21 (SEQ ID NO:50) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Env gp160.

15 Figure 22 (SEQ ID NO:51) depicts the nucleotide sequence of a synthetic Gag-encoding polynucleotide derived from 8_5_TV1_C.ZA.

Figure 23 (SEQ ID NO:52) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Gag.

Figure 24 (SEQ ID NO:53) depicts the nucleotide sequence of a synthetic Gag-encoding polynucleotide (major homology region) derived from 8_5_TV1_C.ZA.

20 Figure 25 (SEQ ID NO:54) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Gag major homology region.

Figure 26 (SEQ ID NO:55) depicts the nucleotide sequence of a synthetic Nef-encoding polynucleotide derived from 8_5_TV1_C.ZA.

25 Figure 27 (SEQ ID NO:56) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Nef.

Figure 28 (SEQ ID NO:57) depicts the nucleotide sequence of a synthetic Nef-encoding polynucleotide derived from 8_5_TV1_C.ZA. The sequence includes a mutation at position 125 which results in a non-functional gene product.

30 Figure 29 (SEQ ID NO:58) depicts the nucleotide sequence of a synthetic RNaseH-encoding polynucleotide derived from 8_5_TV1_C.ZA. RNaseH is a functional domain of the Pol gene, corresponding to p15 (Table A).

WO 02/04493

PCT/US01/21241

Figure 30 (SEQ ID NO:59) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA RNaseH.

Figure 31 (SEQ ID NO:60) depicts the nucleotide sequence of a synthetic integrase (Int)-encoding polynucleotide derived from 8_5_TV1_C.ZA. Int is a functional domain of the Pol gene, corresponding to p31 (Table A).

Figure 32 (SEQ ID NO:61) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Int.

Figure 33 (SEQ ID NO:62) depicts the nucleotide sequence of a synthetic Pol-encoding polynucleotide derived from 8_5_TV1_C.ZA.

Figure 34 (SEQ ID NO:63) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Pol.

Figure 35 (SEQ ID NO:64) depicts the nucleotide sequence of a synthetic protease (prot)-encoding polynucleotide derived from 8_5_TV1_C.ZA.

Figure 36 (SEQ ID NO:65) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Prot.

Figure 37 (SEQ ID NO:66) depicts the nucleotide sequence of a synthetic protease (prot)-encoding polynucleotide derived from 8_5_TV1_C.ZA containing a mutation in which results in inactivation of the protease.

Figure 38 (SEQ ID NO:67) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA inactivated Prot.

Figure 39 (SEQ ID NO:68) depicts the nucleotide sequence of a synthetic protease (prot)-encoding polynucleotide and a synthetic reverse transcriptase (RT)-encoding polynucleotide, both derived from 8_5_TV1_C.ZA. The Prot and RT sequences both contain a mutation which results in inactivation of the gene product.

Figure 40 (SEQ ID NO:69) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA inactivated Prot/mutated RT.

Figure 41 (SEQ ID NO:70) depicts the nucleotide sequence of a synthetic protease (prot)-encoding polynucleotide and a synthetic reverse transcriptase (RT)-encoding polynucleotide, both derived from 8_5_TV1_C.ZA.

Figure 42 (SEQ ID NO:71) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Prot and RT.

WO 02/04493

PCT/US01/21241

Figure 43 (SEQ ID NO:72) depicts the nucleotide sequence of a synthetic rev-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic sequence depicted corresponds to exon 1 of rev. Wild-type rev has two exons.

Figure 44 (SEQ ID NO:73) depicts the wild-type nucleotide sequence of
5 8_5_TV1_C.ZA exon 1 of Rev.

Figure 45 (SEQ ID NO:74) depicts the nucleotide sequence of a synthetic rev-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic sequence depicted corresponds to exon 2 of rev.

Figure 46 (SEQ ID NO:75) depicts the wild-type nucleotide sequence of
10 8_5_TV1_C.ZA exon 2 of Rev.

Figure 47 (SEQ ID NO:76) depicts the nucleotide sequence of a synthetic RT-encoding polynucleotide derived from 8_5_TV1_C.ZA.

Figure 48 (SEQ ID NO:77) depicts the wild-type nucleotide sequence of
8_5_TV1_C.ZA RT.

Figure 49 (SEQ ID NO:78) depicts the nucleotide sequence of a synthetic RT-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic polynucleotide includes a mutation in the RT coding sequence which renders the gene product inactive.

Figure 50 (SEQ ID NO:79) depicts the wild-type nucleotide sequence of
8_5_TV1_C.ZA RT including a mutation which inactivates the RT gene product.

Figure 51 (SEQ ID NO:80) depicts the nucleotide sequence of a synthetic Tat-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic sequence depicted corresponds to exon 1 of Tat and further includes a mutation that renders the Tat gene product non-functional. Wild-type Tat has two exons.

Figure 52 (SEQ ID NO:81) depicts the nucleotide sequence of a synthetic Tat-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic sequence depicted
25 corresponds to exon 1 of Tat.

Figure 53 (SEQ ID NO:82) depicts the wild-type nucleotide sequence of
8_5_TV1_C.ZA exon 1 of Tat.

Figure 54 (SEQ ID NO:83) depicts the nucleotide sequence of a synthetic Tat-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic sequence depicted
30 corresponds to exon 2 of Tat.

WO 02/04493

PCT/US01/21241

Figure 55 (SEQ ID NO:84) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA exon 2 of Tat.

Figure 56 (SEQ ID NO:85) depicts the nucleotide sequence of a synthetic Vif-encoding polynucleotide derived from 8_5_TV1_C.ZA.

5 Figure 57 (SEQ ID NO:86) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Vif.

Figure 58 (SEQ ID NO:87) depicts the nucleotide sequence of a synthetic Vpr-encoding polynucleotide derived from 8_5_TV1_C.ZA.

10 Figure 59 (SEQ ID NO:88) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Vpr.

Figure 60 (SEQ ID NO:89) depicts the nucleotide sequence of a synthetic Vpu-encoding polynucleotide derived from 8_5_TV1_C.ZA.

Figure 61 (SEQ ID NO:90) depicts the wild-type nucleotide sequence of 8_5_TV1_C.ZA Vpu.

15 Figure 62 (SEQ ID NO:91) depicts the nucleotide sequence of a synthetic rev-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic sequence depicted corresponds to exons 1 and 2 of rev.

Figure 63 (SEQ ID NO:92) depicts the wild-type nucleotide sequence of exons 1 and 2 of rev derived from 8_5_TV1_C.ZA.

20 Figure 64 (SEQ ID NO:93) depicts the nucleotide sequence of a synthetic Tat-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic polynucleotide includes both exons 1 and 2 of Tat and further includes a mutation in exon 1 which renders the gene product non-functional.

25 Figure 65 (SEQ ID NO:94) depicts the nucleotide sequence of a synthetic Tat-encoding polynucleotide derived from 8_5_TV1_C.ZA. The synthetic polynucleotide includes both exons 1 and 2 of Tat.

Figure 66 (SEQ ID NO:95) depicts the wild-type nucleotide sequence of exons 1 and 2 of Tat derived from 8_5_TV1_C.ZA.

30 Figure 67 (SEQ ID NO:96) depicts the nucleotide sequence of a synthetic Nef-encoding polynucleotide derived from 8_5_TV1_C.ZA. The sequence includes a mutation at position 125 which results in a non-functional gene product and a mutation that eliminates the myristoylation site of the Nef gene product.

WO 02/04493

PCT/US01/21241

Figure 68 (SEQ ID NO:97) depicts the nucleotide sequence of a synthetic Env gp160-encoding polynucleotide derived from 12-5_1_TV2_C.ZA.

Figure 69 (SEQ ID NO:98) depicts the wild-type nucleotide sequence of Env gp160 derived from 12-5_1_TV2_C.ZA.

5 Figure 70 (SEQ ID NO:99) depicts the nucleotide sequence of a synthetic Gag-encoding polynucleotide derived from 12-5_1_TV2_C.ZA.

Figure 71 (SEQ ID NO:100) depicts the wild-type nucleotide sequence of Gag derived from 12-5_1_TV2_C.ZA.

10 Figure 72 (SEQ ID NO:101) depicts the nucleotide sequence of a synthetic Nef-encoding polynucleotide derived from 12-5_1_TV2_C.ZA.

Figure 73 (SEQ ID NO:102) depicts the wild-type nucleotide sequence of Nef derived from 12-5_1_TV2_C.ZA.

Figure 74 (SEQ ID NO:103) depicts the nucleotide sequence of a synthetic Pol-encoding polynucleotide derived from 12-5_1_TV2_C.ZA.

15 Figure 75 (SEQ ID NO:104) depicts the wild-type nucleotide sequence of Pol derived from 12-5_1_TV2_C.ZA.

Figure 76 (SEQ ID NO:105) depicts the nucleotide sequence of a synthetic Rev-encoding polynucleotide derived from exon 1 of Rev from 12-5_1_TV2_C.ZA.

20 Figure 77 (SEQ ID NO:106) depicts the wild-type nucleotide sequence of exon 1 of Rev derived from 12-5_1_TV2_C.ZA.

Figure 78 (SEQ ID NO:107) depicts the nucleotide sequence of a synthetic Rev-encoding polynucleotide derived from exon 2 of Rev from 12-5_1_TV2_C.ZA.

Figure 79 (SEQ ID NO:108) depicts the wild-type nucleotide sequence of exon 2 of Rev derived from 12-5_1_TV2_C.ZA.

25 Figure 80 (SEQ ID NO:109) depicts the nucleotide sequence of a synthetic Tat-encoding polynucleotide derived from exon 1 of Tat from 12-5_1_TV2_C.ZA.

Figure 81 (SEQ ID NO:110) depicts the wild-type nucleotide sequence of exon 1 of Tat derived from 12-5_1_TV2_C.ZA.

30 Figure 82 (SEQ ID NO:111) depicts the nucleotide sequence of a synthetic Tat-encoding polynucleotide derived from exon 2 of Tat from 12-5_1_TV2_C.ZA.

Figure 83 (SEQ ID NO:112) depicts the wild-type nucleotide sequence of exon 2 of Tat derived from 12-5_1_TV2_C.ZA.

WO 02/04493

PCT/US01/21241

Figure 84 (SEQ ID NO:113) depicts the nucleotide sequence of a synthetic Vif-encoding polynucleotide derived from 12-5_1_TV2_C.ZA.

Figure 85 (SEQ ID NO:114) depicts the wild-type nucleotide sequence of Vif derived from 12-5_1_TV2_C.ZA.

5 Figure 86 (SEQ ID NO:115) depicts the nucleotide sequence of a synthetic Vpr-encoding polynucleotide derived from 12-5_1_TV2_C.ZA.

Figure 87 (SEQ ID NO:116) depicts the wild-type nucleotide sequence of Vpr derived from 12-5_1_TV2_C.ZA.

10 Figure 88 (SEQ ID NO:117) depicts the nucleotide sequence of a synthetic Vpu-encoding polynucleotide derived from 12-5_1_TV2_C.ZA.

Figure 89 (SEQ ID NO:118) depicts the wild-type nucleotide sequence of Vpu derived from 12-5_1_TV2_C.ZA.

Figure 90 (SEQ ID NO:119) depicts the nucleotide sequence of a synthetic Env gp120-encoding polynucleotide derived from 8_2_TV1_C.ZA. The V2 region is deleted.

15 The sequence includes: an EcoRI restriction site (nucleotides 1 to 6); a codon modified signal peptide leader sequence (nucleotides 7 to 87); a gp120 coding sequence (nucleotides 88 to 1464); a stop codon (nucleotides 1465 to 1467); an XhoI restriction site (nucleotides 1468 to 1473).

Figure 91 (SEQ ID NO:120) depicts the nucleotide sequence of a synthetic Env gp140-encoding polynucleotide derived from 8_2_TV1_C.ZA. The V2 region is deleted.

20 The sequence includes: an EcoRI restriction site (nucleotides 1 to 6); a modified signal peptide leader sequence (nucleotides 7 to 87); a gp140 coding sequence (nucleotides 88 to 1977); a stop codon (nucleotides 1978 to 1980); an XhoI restriction site (nucleotides 1981 to 1986).

25 Figure 92 (SEQ ID NO:121) depicts the nucleotide sequence of a synthetic Env gp140-encoding polynucleotide derived from 8_2_TV1_C.ZA. The V2 region is deleted and the sequence includes mutations in the cleavage site that prevent the cleavage of a gp140 polypeptide into a gp120 polypeptide and a gp41 polypeptide. The sequence includes: an

30 EcoRI restriction site (nucleotides 1 to 6); a modified signal peptide leader sequence (nucleotides 7 to 87); gp140 coding sequence (nucleotides 88 to 1977); a stop codon (nucleotides 1978 to 1980); an XhoI restriction site (nucleotides 1981 to 1986).

WO 02/04493

PCT/US01/21241

Figure 93 (SEQ ID NO:122) depicts the nucleotide sequence of a synthetic Env gp160-encoding polynucleotide derived from 8_2_TV1_C.ZA. The V1/V2 regions are deleted. The sequence includes: an EcoRI restriction site (nucleotides 1 to 6); a modified signal peptide leader sequence (nucleotides 7 to 87); gp160 coding sequence (nucleotides 88 to 2388); a stop codon (nucleotides 2389 to 2391); an XhoI restriction site (nucleotides 2392 to 2397).

Figure 94 (SEQ ID NO:123) depicts the nucleotide sequence of a synthetic Env gp160-encoding polynucleotide derived from 8_2_TV1_C.ZA. The V2 region is deleted. The sequence includes: an EcoRI restriction site (nucleotides 1 to 6); a modified signal peptide leader sequence (nucleotides 7 to 87); a gp160 coding sequence (nucleotides 88 to 2520); a stop codon (nucleotides 2521 to 2523); an XhoI restriction site (nucleotides 2524 to 2529).

Figure 95 (SEQ ID NO:124) depicts the nucleotide sequence of a synthetic Env gp160-encoding polynucleotide derived from 8_2_TV1_C.ZA. The V2 region is deleted and the cleavage site is mutated. The sequence includes: an EcoRI restriction site (nucleotides 1 to 6); a modified signal peptide leader sequence (nucleotides 7 to 87); a gp160 coding sequence (nucleotides 88 to 2520); a stop codon (nucleotides 2521 to 2523); an XhoI restriction site (nucleotides 2524 to 2529).

Figure 96 (SEQ ID NO:125) depicts the nucleotide sequence of a synthetic Env gp160-encoding polynucleotide derived from 8_2_TV1_C.ZA. The nucleotide sequence includes a TPA1 leader sequence. The sequence includes: a SalI restriction site (nucleotides 1 to 6); a Kozak sequence (nucleotides 7 to 12); a TPA1 signal peptide leader sequence (nucleotides 13 to 87); a gp160 coding sequence (nucleotides 88 to 2604); a stop codon (nucleotides 2605 to 2607); an XhoI restriction site (nucleotides 2608 to 2613).

Figure 97 (SEQ ID NO:126) depicts the nucleotide sequence of a synthetic Env gp160-encoding polynucleotide derived from 8_2_TV1_C.ZA. The sequence includes: an EcoRI restriction site (nucleotides 1 to 6); a modified signal peptide leader sequence (nucleotides 7 to 87); a gp160 coding sequence (nucleotides 88 to 2607); a stop codon (nucleotides 2608 to 2610); an XhoI restriction site (nucleotides 2611 to 2616).

Figure 98 (SEQ ID NO:127) depicts the nucleotide sequence of a synthetic Env gp160-encoding polynucleotide derived from 8_2_TV1_C.ZA. The nucleotide sequence includes a wild type leader sequence. The sequence includes: an EcoRI restriction site

WO 02/04493

PCT/US01/21241

(nucleotides 1 to 6); a native (unmodified) signal peptide leader sequence (nucleotides 7 to 87); a gp160 coding sequence (nucleotides 88 to 2607); a stop codon (nucleotides 2608 to 2610); an XhoI restriction site (nucleotides 2611 to 2616).

Figure 99 (SEQ ID NO:128) depicts the nucleotide sequence of wild type gp160 derived from 8_2_TV1_C.ZA.

Figure 100 (SEQ ID NO:131) depicts the nucleotide sequence of a synthetic Env gp140-encoding polynucleotide derived from 8_2_TV1_C.ZA. The nucleotide sequence includes a TPA1 leader sequence (nucleotides 1-75); a gp140 coding sequence (nucleotides 76 to 2049); a stop codon (nucleotides 2050 to 2052).

Figure 101 (SEQ ID NO:132) depicts the nucleotide sequence of a synthetic gp140-encoding polynucleotide derived from 8_2_TV1_C.ZA. The nucleotide sequence includes an EcoRI restriction site (nucleotides 1 to 6); a leader sequence modified from the TV1_C.ZA wild-type leader sequence (nucleotides 7 to 87); a gp140 coding sequence (nucleotides 88 to 2064); a stop codon (nucleotides 2065 to 2067); a XhoI restriction site (nucleotides 2068 to 2073).

Figure 102 (SEQ ID NO:133) depicts the nucleotide sequence of a synthetic gp140-encoding polynucleotide derived from 8_2_TV1_C.ZA. The nucleotide sequence includes wild-type TV1_C.ZA unmodified leader sequence. The nucleotide sequence includes a restriction site (nucleotides 1 to 6); a wild type leader sequence (nucleotides 7 to 87); a gp140 coding sequence (nucleotides 88 to 2064); a stop codon (nucleotides 2065 to 2067); a XhoI restriction site (nucleotides 2068-2073).

Figure 103 (SEQ ID NO:134) depicts the nucleotide sequence of a synthetic Nef-encoding polynucleotide derived from 12-5_1_TV2_C.ZA. The sequence includes a mutation at position 125 which results in a non-functional gene product.

Figure 104 (SEQ ID NO:135) depicts the nucleotide sequence of a synthetic Nef-encoding polynucleotide derived from 12-5_1_TV2_C.ZA. The synthetic polynucleotide includes a mutation that eliminates the myristoylation site of the Nef gene product.

Figure 105 depicts an alignment of Env polypeptides from various HIV isolates. The regions between the arrows indicate regions (of TV1 and TV2 clones) in the beta and/or bridging sheet region(s) that can be deleted and/or truncated. The "*" denotes N-linked glycosylation sites (of TV1 and TV2 clones), one or more of which can be modified (e.g., deleted and/or mutated).

WO 02/04493

PCT/US01/21241

DETAILED DESCRIPTION OF THE INVENTION

The practice of the present invention will employ, unless otherwise indicated, conventional methods of chemistry, biochemistry, molecular biology, immunology and pharmacology, within the skill of the art. Such techniques are explained fully in the literature. See, e.g., *Remington's Pharmaceutical Sciences*, 18th Edition (Easton, Pennsylvania: Mack Publishing Company, 1990); *Methods In Enzymology* (S. Colowick and N. Kaplan, eds., Academic Press, Inc.); and *Handbook of Experimental Immunology*, Vols. I-IV (D.M. Weir and C.C. Blackwell, eds., 1986, Blackwell Scientific Publications); Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); *Short Protocols in Molecular Biology*, 4th ed. (Ausubel et al. eds., 1999, John Wiley & Sons); *Molecular Biology Techniques: An Intensive Laboratory Course*, (Rcam et al., eds., 1998, Academic Press); *PCR (Introduction to Biotechniques Series)*, 2nd ed. (Newton & Graham eds., 1997, Springer Verlag).

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise. Thus, for example, reference to "an antigen" includes a mixture of two or more such agents.

1. DEFINITIONS

In describing the present invention, the following terms will be employed, and are intended to be defined as indicated below.

"Synthetic" sequences, as used herein, refers to Type C HIV polypeptide-encoding polynucleotides whose expression has been modified as described herein, for example, by codon substitution and inactivation of inhibitory sequences. "Wild-type" or "native" sequences, as used herein, refers to polypeptide encoding sequences that are essentially as they are found in nature, e.g., Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env and/or Nef encoding sequences as found in Type C isolates, e.g., AF110965, AF110967, AF110968, AF110975, 8_5_TV1_C.ZA, 8_2_TV1_C.ZA or 12-5_1_TV2_C.ZA. The various regions of the HIV genome are shown in Table A, with numbering relative to 8_5_TV1_C.ZA (SEQ ID NO:33). Thus, the term "Pol" refers to one or more of the following polypeptides: polymerase (p6Pol); protease (prot); reverse transcriptase (p66RT or RT); RNaseH (p15RNaseH); and/or integrase (p31Int or Int).

WO 02/04493

PCT/US01/21241

As used herein, the term "virus-like particle" or "VLP" refers to a nonreplicating, viral shell, derived from any of several viruses discussed further below. VLPs are generally composed of one or more viral proteins, such as, but not limited to those proteins referred to as capsid, coat, shell, surface and/or envelope proteins, or particle-forming polypeptides derived from these proteins. VLPs can form spontaneously upon recombinant expression of the protein in an appropriate expression system. Methods for producing particular VLPs are known in the art and discussed more fully below. The presence of VLPs following recombinant expression of viral proteins can be detected using conventional techniques known in the art, such as by electron microscopy, X-ray crystallography, and the like. See, e.g., Baker et al., *Biophys. J.* (1991) 60:1445-1456; Hagensee et al., *J. Virol.* (1994) 68:4503-4505. For example, VLPs can be isolated by density gradient centrifugation and/or identified by characteristic density banding. Alternatively, cryoelectron microscopy can be performed on vitrified aqueous samples of the VLP preparation in question, and images recorded under appropriate exposure conditions.

By "particle-forming polypeptide" derived from a particular viral protein is meant a full-length or near full-length viral protein, as well as a fragment thereof, or a viral protein with internal deletions, which has the ability to form VLPs under conditions that favor VLP formation. Accordingly, the polypeptide may comprise the full-length sequence, fragments, truncated and partial sequences, as well as analogs and precursor forms of the reference molecule. The term therefore intends deletions, additions and substitutions to the sequence, so long as the polypeptide retains the ability to form a VLP. Thus, the term includes natural variations of the specified polypeptide since variations in coat proteins often occur between viral isolates. The term also includes deletions, additions and substitutions that do not naturally occur in the reference protein, so long as the protein retains the ability to form a VLP. Preferred substitutions are those which are conservative in nature, i.e., those substitutions that take place within a family of amino acids that are related in their side chains. Specifically, amino acids are generally divided into four families: (1) acidic -- aspartate and glutamate; (2) basic -- lysine, arginine, histidine; (3) non-polar -- alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar -- glycine, asparagine, glutamine, cystine, serine, threonine, tyrosine. Phenylalanine, tryptophan, and tyrosine are sometimes classified as aromatic amino acids.

WO 02/04493

PCT/US01/21241

An "antigen" refers to a molecule containing one or more epitopes (either linear, conformational or both) that will stimulate a host's immune system to make a humoral and/or cellular antigen-specific response. The term is used interchangeably with the term "immunogen." Normally, a B-cell epitope will include at least about 5 amino acids but can be as small as 3-4 amino acids. A T-cell epitope, such as a CTL epitope, will include at least about 7-9 amino acids, and a helper T-cell epitope at least about 12-20 amino acids. Normally, an epitope will include between about 7 and 15 amino acids, such as, 9, 10, 12 or 15 amino acids. The term "antigen" denotes both subunit antigens, (i.e., antigens which are separate and discrete from a whole organism with which the antigen is associated in nature), as well as, killed, attenuated or inactivated bacteria, viruses, fungi, parasites or other microbes. Antibodies such as anti-idiotypic antibodies, or fragments thereof, and synthetic peptide mimotopes, which can mimic an antigen or antigenic determinant, are also captured under the definition of antigen as used herein. Similarly, an oligonucleotide or polynucleotide which expresses an antigen or antigenic determinant *in vivo*, such as in gene therapy and DNA immunization applications, is also included in the definition of antigen herein.

For purposes of the present invention, antigens can be derived from any of several known viruses, bacteria, parasites and fungi, as described more fully below. The term also intends any of the various tumor antigens. Furthermore, for purposes of the present invention, an "antigen" refers to a protein which includes modifications, such as deletions, additions and substitutions (generally conservative in nature), to the native sequence, so long as the protein maintains the ability to elicit an immunological response, as defined herein. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through mutations of hosts which produce the antigens.

An "immunological response" to an antigen or composition is the development in a subject of a humoral and/or a cellular immune response to an antigen present in the composition of interest. For purposes of the present invention, a "humoral immune response" refers to an immune response mediated by antibody molecules, while a "cellular immune response" is one mediated by T-lymphocytes and/or other white blood cells. One important aspect of cellular immunity involves an antigen-specific response by cytolytic T-cells ("CTL"s). CTLs have specificity for peptide antigens that are presented in association with proteins encoded by the major histocompatibility complex (MHC) and expressed on the

WO 02/04493

PCT/US01/21241

surfaces of cells. CTLs help induce and promote the destruction of intracellular microbes, or the lysis of cells infected with such microbes. Another aspect of cellular immunity involves an antigen-specific response by helper T-cells. Helper T-cells act to help stimulate the function, and focus the activity of, nonspecific effector cells against cells displaying peptide antigens in association with MHC molecules on their surface. A "cellular immune response" also refers to the production of cytokines, chemokines and other such molecules produced by activated T-cells and/or other white blood cells, including those derived from CD4+ and CD8+ T-cells.

A composition or vaccine that elicits a cellular immune response may serve to sensitize a vertebrate subject by the presentation of antigen in association with MHC molecules at the cell surface. The cell-mediated immune response is directed at, or near, cells presenting antigen at their surface. In addition, antigen-specific T-lymphocytes can be generated to allow for the future protection of an immunized host.

The ability of a particular antigen to stimulate a cell-mediated immunological response may be determined by a number of assays, such as by lymphoproliferation (lymphocyte activation) assays, CTL cytotoxic cell assays, or by assaying for T-lymphocytes specific for the antigen in a sensitized subject. Such assays are well known in the art. See, e.g., Erickson et al., *J. Immunol.* (1993) 151:4189-4199; Doe et al., *Eur. J. Immunol.* (1994) 24:2369-2376. Recent methods of measuring cell-mediated immune response include measurement of intracellular cytokines or cytokine secretion by T-cell populations, or by measurement of epitope specific T-cells (e.g., by the tetramer technique)(reviewed by McMichael, A.J., and O'Callaghan, C.A., *J. Exp. Med.* 187(9)1367-1371, 1998; Mohezyer-Williams, M.G., et al, *Immunol. Rev.* 150:5-21, 1996; Lalvani, A., et al, *J. Exp. Med.* 186:859-865, 1997).

Thus, an immunological response as used herein may be one which stimulates the production of CTLs, and/or the production or activation of helper T- cells. The antigen of interest may also elicit an antibody-mediated immune response. Hence, an immunological response may include one or more of the following effects: the production of antibodies by B-cells; and/or the activation of suppressor T-cells and/or $\gamma\delta$ T-cells directed specifically to an antigen or antigens present in the composition or vaccine of interest. These responses may serve to neutralize infectivity, and/or mediate antibody-complement, or antibody dependent

WO 02/04493

PCT/US01/21241

cell cytotoxicity (ADCC) to provide protection to an immunized host. Such responses can be determined using standard immunoassays and neutralization assays, well known in the art.

An "immunogenic composition" is a composition that comprises an antigenic molecule where administration of the composition to a subject results in the development in the subject of a humoral and/or a cellular immune response to the antigenic molecule of interest. The immunogenic composition can be introduced directly into a recipient subject, such as by injection, inhalation, oral, intranasal and mucosal (e.g., intra-rectally or intra-vaginally) administration.

By "subunit vaccine" is meant a vaccine composition which includes one or more selected antigens but not all antigens, derived from or homologous to, an antigen from a pathogen of interest such as from a virus, bacterium, parasite or fungus. Such a composition is substantially free of intact pathogen cells or pathogenic particles, or the lysate of such cells or particles. Thus, a "subunit vaccine" can be prepared from at least partially purified (preferably substantially purified) immunogenic polypeptides from the pathogen, or analogs thereof. The method of obtaining an antigen included in the subunit vaccine can thus include standard purification techniques, recombinant production, or synthetic production.

"Substantially purified" general refers to isolation of a substance (compound, polynucleotide, protein, polypeptide, polypeptide composition) such that the substance comprises the majority percent of the sample in which it resides. Typically in a sample a substantially purified component comprises 50%, preferably 80%-85%, more preferably 90-95% of the sample. Techniques for purifying polynucleotides and polypeptides of interest are well-known in the art and include, for example, ion-exchange chromatography, affinity chromatography and sedimentation according to density.

A "coding sequence" or a sequence which "encodes" a selected polypeptide, is a nucleic acid molecule which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences (or "control elements"). The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to, cDNA from viral, procaryotic or eucaryotic mRNA, genomic DNA sequences from viral or procaryotic DNA, and even synthetic DNA sequences. A transcription termination sequence such as a stop codon may be located 3' to the coding sequence.

WO 02/04493

PCT/US01/21241

Typical "control elements", include, but are not limited to, transcription promoters, transcription enhancer elements, transcription termination signals, polyadenylation sequences (located 3' to the translation stop codon), sequences for optimization of initiation of translation (located 5' to the coding sequence), and translation termination sequences.

5 A "polynucleotide coding sequence" or a sequence which "encodes" a selected polypeptide, is a nucleic acid molecule which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences (or "control elements"). The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. Exemplary coding sequences are the modified viral polypeptide-coding sequences of the present invention. A transcription termination sequence may be located 3' to the coding sequence. Typical "control elements", include, but are not limited to, transcription regulators, such as promoters, transcription enhancer elements, transcription termination signals, and polyadenylation sequences; and translation regulators, 15 such as sequences for optimization of initiation of translation, *e.g.*, Shine-Dalgarno (ribosome binding site) sequences, Kozak sequences (*i.e.*, sequences for the optimization of translation, located, for example, 5' to the coding sequence), leader sequences, translation initiation codon (*e.g.*, ATG), and translation termination sequences. In certain embodiments, one or more translation regulation or initiation sequences (*e.g.*, the leader sequence) are derived from wild-type translation initiation sequences, *i.e.*, sequences that regulate translation of the coding region in their native state. Wild-type leader sequences that have been modified, using the methods described herein, also find use in the present invention. Promoters can include inducible promoters (where expression of a polynucleotide sequence operably linked to the promoter is induced by an analyte, cofactor, regulatory protein, etc.), repressible 20 promoters (where expression of a polynucleotide sequence operably linked to the promoter is induced by an analyte, cofactor, regulatory protein, etc.), and constitutive promoters.

A "nucleic acid" molecule can include, but is not limited to, procaryotic sequences, eucaryotic mRNA, cDNA from eucaryotic mRNA, genomic DNA sequences from eucaryotic (*e.g.*, mammalian) DNA, and even synthetic DNA sequences. The term also captures 30 sequences that include any of the known base analogs of DNA and RNA.

"Operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, a given promoter

WO 02/04493

PCT/US01/21241

operably linked to a coding sequence is capable of effecting the expression of the coding sequence when the proper enzymes are present. The promoter need not be contiguous with the coding sequence, so long as it functions to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between the promoter sequence and the coding sequence and the promoter sequence can still be considered "operably linked" to the coding sequence.

"Recombinant" as used herein to describe a nucleic acid molecule means a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which, by virtue of its origin or manipulation: (1) is not associated with all or a portion of the polynucleotide with which it is associated in nature; and/or (2) is linked to a polynucleotide other than that to which it is linked in nature. The term "recombinant" as used with respect to a protein or polypeptide means a polypeptide produced by expression of a recombinant polynucleotide. "Recombinant host cells," "host cells," "cells," "cell lines," "cell cultures," and other such terms denoting procaryotic microorganisms or eucaryotic cell lines cultured as unicellular entities, are used interchangeably, and refer to cells which can be, or have been, used as recipients for recombinant vectors or other transfer DNA, and include the progeny of the original cell which has been transfected. It is understood that the progeny of a single parental cell may not necessarily be completely identical in morphology or in genomic or total DNA complement to the original parent, due to accidental or deliberate mutation. Progeny of the parental cell which are sufficiently similar to the parent to be characterized by the relevant property, such as the presence of a nucleotide sequence encoding a desired peptide, are included in the progeny intended by this definition, and are covered by the above terms.

Techniques for determining amino acid sequence "similarity" are well known in the art. In general, "similarity" means the exact amino acid to amino acid comparison of two or more polypeptides at the appropriate place, where amino acids are identical or possess similar chemical and/or physical properties such as charge or hydrophobicity. A so-termed "percent similarity" then can be determined between the compared polypeptide sequences.

Techniques for determining nucleic acid and amino acid sequence identity also are well known in the art and include determining the nucleotide sequence of the mRNA for that gene (usually via a cDNA intermediate) and determining the amino acid sequence encoded thereby, and comparing this to a second amino acid sequence. In general, "identity" refers to

WO 02/04493

PCT/US01/21241

an exact nucleotide to nucleotide or amino acid to amino acid correspondence of two polynucleotides or polypeptide sequences, respectively.

Two or more polynucleotide sequences can be compared by determining their "percent identity." Two or more amino acid sequences likewise can be compared by determining their "percent identity." The percent identity of two sequences, whether nucleic acid or peptide sequences, is generally described as the number of exact matches between two aligned sequences divided by the length of the shorter sequence and multiplied by 100. An approximate alignment for nucleic acid sequences is provided by the local homology algorithm of Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981).

This algorithm can be extended to use with peptide sequences using the scoring matrix developed by Dayhoff, *Atlas of Protein Sequences and Structure*, M.O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA, and normalized by Gribakov, *Nucl. Acids Res.* 14(6):6745-6763 (1986). An implementation of this algorithm for nucleic acid and peptide sequences is provided by the Genetics Computer Group (Madison, WI) in their BestFit utility application. The default parameters for this method are described in the *Wisconsin Sequence Analysis Package Program Manual*, Version 8 (1995) (available from Genetics Computer Group, Madison, WI). Other equally suitable programs for calculating the percent identity or similarity between sequences are generally known in the art.

For example, percent identity of a particular nucleotide sequence to a reference sequence can be determined using the homology algorithm of Smith and Waterman with a default scoring table and a gap penalty of six nucleotide positions. Another method of establishing percent identity in the context of the present invention is to use the MPSRCH package of programs copyrighted by the University of Edinburgh, developed by John F. Collins and Shane S. Sturrol, and distributed by IntelliGenetics, Inc. (Mountain View, CA). From this suite of packages, the Smith-Waterman algorithm can be employed where default parameters are used for the scoring table (for example, gap open penalty of 12, gap extension penalty of one, and a gap of six). From the data generated, the "Match" value reflects "sequence identity." Other suitable programs for calculating the percent identity or similarity between sequences are generally known in the art, such as the alignment program BLAST, which can also be used with default parameters. For example, BLASTN and BLASTP can be used with the following default parameters: genetic code = standard; filter = none; strand =

WO 02/04493

PCT/US01/21241

both; cutoff = 60; expect = 10; Matrix = BLOSUM62; Descriptions = 50 sequences; sort by = HIGH SCORE; Databases = non-redundant, GenBank + EMBL + DDBJ + PDB + GenBank CDS translations + Swiss protein + Spupdate + PIR. Details of these programs can be found at the following internet address: <http://www.ncbi.nlm.gov/cgi-bin/BLAST>.

5 One of skill in the art can readily determine the proper search parameters to use for a given sequence, exemplary preferred Smith Waterman based parameters are presented above. For example, the search parameters may vary based on the size of the sequence in question. Thus, for the polynucleotide sequences of the present invention the length of the polynucleotide sequence disclosed herein is searched against a selected database and
10 compared to sequences of essentially the same length to determine percent identity. For example, a representative embodiment of the present invention would include an isolated polynucleotide having X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least about a selected level of percent identity relative to Y contiguous nucleotides of the sequences described herein, and (ii) for search purposes X equals Y, wherein Y is a
15 selected reference polynucleotide of defined length.

The sequences of the present invention can include fragments of the sequences, for example, from about 15 nucleotides up to the number of nucleotides present in the full-length sequences described herein (e.g., see the Sequence Listing, Figures, and claims), including all integer values falling within the above-described range. For example, fragments of the
20 polynucleotide sequences of the present invention may be 30-60 nucleotides, 60-120 nucleotides, 120-240 nucleotides, 240-480 nucleotides, 480-1000 nucleotides, and all integer values therebetween.

The synthetic expression cassettes (and purified polynucleotides) of the present invention include related polynucleotide sequences having about 80% to 100%, greater than
25 80-85%, preferably greater than 90-92%, more preferably greater than 95%, and most preferably greater than 98% up to 100% (including all integer values falling within these described ranges) sequence identity to the synthetic expression cassette (and purified polynucleotide) sequences disclosed herein (for example, to the claimed sequences or other sequences of the present invention) when the sequences of the present invention are used as
30 the query sequence against, for example, a database of sequences.

Two nucleic acid fragments are considered to "selectively hybridize" as described herein. The degree of sequence identity between two nucleic acid molecules affects the

WO 02/04493

PCT/US01/21241

efficiency and strength of hybridization events between such molecules. A partially identical nucleic acid sequence will at least partially inhibit a completely identical sequence from hybridizing to a target molecule. Inhibition of hybridization of the completely identical sequence can be assessed using hybridization assays that are well known in the art (e.g.,

5 Southern blot, Northern blot, solution hybridization, or the like, see Sambrook, et al., *supra* or Ausubel et al., *supra*). Such assays can be conducted using varying degrees of selectivity, for example, using conditions varying from low to high stringency. If conditions of low stringency are employed, the absence of non-specific binding can be assessed using a secondary probe that lacks even a partial degree of sequence identity (for example, a probe

10 having less than about 30% sequence identity with the target molecule), such that, in the absence of non-specific binding events, the secondary probe will not hybridize to the target.

When utilizing a hybridization-based detection system, a nucleic acid probe is chosen that is complementary to a target nucleic acid sequence, and then by selection of appropriate conditions the probe and the target sequence "selectively hybridize," or bind, to each other to

15 form a hybrid molecule. A nucleic acid molecule that is capable of hybridizing selectively to a target sequence under "moderately stringent" typically hybridizes under conditions that allow detection of a target nucleic acid sequence of at least about 10-14 nucleotides in length having at least approximately 70% sequence identity with the sequence of the selected nucleic acid probe. Stringent hybridization conditions typically allow detection of target

20 nucleic acid sequences of at least about 10-14 nucleotides in length having a sequence identity of greater than about 90-95% with the sequence of the selected nucleic acid probe. Hybridization conditions useful for probe/target hybridization where the probe and target have a specific degree of sequence identity, can be determined as is known in the art (see, for example, *Nucleic Acid Hybridization: A Practical Approach*, editors B.D. Barnes and S.J.

25 Higgins, (1985) Oxford; Washington, DC; IRL Press).

With respect to stringency conditions for hybridization, it is well known in the art that numerous equivalent conditions can be employed to establish a particular stringency by varying, for example, the following factors: the length and nature of probe and target sequences, base composition of the various sequences, concentrations of salts and other

30 hybridization solution components, the presence or absence of blocking agents in the hybridization solutions (e.g., formamide, dextran sulfate, and polyethylene glycol), hybridization reaction temperature and time parameters, as well as, varying wash conditions.

WO 02/04493

PCT/US01/21241

The selection of a particular set of hybridization conditions is selected following standard methods in the art (see, for example, Sambrook, et al., *supra* or Ausubel et al., *supra*).

A first polynucleotide is "derived from" second polynucleotide if it has the same or substantially the same basepair sequence as a region of the second polynucleotide, its cDNA, complements thereof, or if it displays sequence identity as described above.

A first polypeptide is "derived from" a second polypeptide if it is (i) encoded by a first polynucleotide derived from a second polynucleotide, or (ii) displays sequence identity to the second polypeptides as described above.

Generally, a viral polypeptide is "derived from" a particular polypeptide of a virus (viral polypeptide) if it is (i) encoded by an open reading frame of a polynucleotide of that virus (viral polynucleotide), or (ii) displays sequence identity to polypeptides of that virus as described above.

"Encoded by" refers to a nucleic acid sequence which codes for a polypeptide sequence, wherein the polypeptide sequence or a portion thereof contains an amino acid sequence of at least 3 to 5 amino acids, more preferably at least 8 to 10 amino acids, and even more preferably at least 15 to 20 amino acids from a polypeptide encoded by the nucleic acid sequence. Also encompassed are polypeptide sequences which are immunologically identifiable with a polypeptide encoded by the sequence. Further, polyproteins can be constructed by fusing in-frame two or more polynucleotide sequences encoding polypeptide or peptide products. Further, polycistronic coding sequences may be produced by placing two or more polynucleotide sequences encoding polypeptide products adjacent each other, typically under the control of one promoter, wherein each polypeptide coding sequence may be modified to include sequences for internal ribosome binding sites.

"Purified polynucleotide" refers to a polynucleotide of interest or fragment thereof which is essentially free, e.g., contains less than about 50%, preferably less than about 70%, and more preferably less than about 90%, of the protein with which the polynucleotide is naturally associated. Techniques for purifying polynucleotides of interest are well-known in the art and include, for example, disruption of the cell containing the polynucleotide with a chaotropic agent and separation of the polynucleotide(s) and proteins by ion-exchange chromatography, affinity chromatography and sedimentation according to density.

By "nucleic acid immunization" is meant the introduction of a nucleic acid molecule encoding one or more selected antigens into a host cell, for the *in vivo* expression of an

WO 02/04493

PCT/US01/21241

antigen, antigens, an epitope, or epitopes. The nucleic acid molecule can be introduced directly into a recipient subject, such as by injection, inhalation, oral, intranasal and mucosal administration, or the like, or can be introduced *ex vivo*, into cells which have been removed from the host. In the latter case, the transformed cells are reintroduced into the subject where an immune response can be mounted against the antigen encoded by the nucleic acid molecule.

"Gene transfer" or "gene delivery" refers to methods or systems for reliably inserting DNA of interest into a host cell. Such methods can result in transient expression of non-integrated transferred DNA, extrachromosomal replication and expression of transferred replicons (e.g., episomes), or integration of transferred genetic material into the genomic DNA of host cells. Gene delivery expression vectors include, but are not limited to, vectors derived from alphaviruses, pox viruses and vaccinia viruses. When used for immunization, such gene delivery expression vectors may be referred to as vaccines or vaccine vectors.

"T lymphocytes" or "T cells" are non-antibody producing lymphocytes that constitute a part of the cell-mediated arm of the immune system. T cells arise from immature lymphocytes that migrate from the bone marrow to the thymus, where they undergo a maturation process under the direction of thymic hormones. Here, the mature lymphocytes rapidly divide increasing to very large numbers. The maturing T cells become immunocompetent based on their ability to recognize and bind a specific antigen. Activation of immunocompetent T cells is triggered when an antigen binds to the lymphocyte's surface receptors.

The term "transfection" is used to refer to the uptake of foreign DNA by a cell. A cell has been "transfected" when exogenous DNA has been introduced inside the cell membrane. A number of transfection techniques are generally known in the art. See, e.g., Graham et al. (1973) *Virology*, 52:456, Sambrook et al. (1989) *Molecular Cloning, a laboratory manual*, Cold Spring Harbor Laboratories, New York, Davis et al. (1986) *Basic Methods in Molecular Biology*, Elsevier, and Chu et al. (1981) *Gene* 13:197. Such techniques can be used to introduce one or more exogenous DNA moieties into suitable host cells. The term refers to both stable and transient uptake of the genetic material, and includes uptake of peptide- or antibody-linked DNAs.

A "vector" is capable of transferring gene sequences to target cells (e.g., viral vectors, non-viral vectors, particulate carriers, and liposomes). Typically, "vector construct,"

WO 02/04493

PCT/US01/21241

"expression vector," and "gene transfer vector," mean any nucleic acid construct capable of directing the expression of a gene of interest and which can transfer gene sequences to target cells. Thus, the term includes cloning and expression vehicles, as well as viral vectors.

Transfer of a "suicide gene" (e.g., a drug-susceptibility gene) to a target cell renders the cell sensitive to compounds or compositions that are relatively nontoxic to normal cells. Moolten, F.L. (1994) *Cancer Gene Ther.* 1:279-287. Examples of suicide genes are thymidine kinase of herpes simplex virus (HSV-tk), cytochrome P450 (Manome et al. (1996) *Gene Therapy* 3:513-520), human deoxycytidine kinase (Manome et al. (1996) *Nature Medicine* 2(5):567-573) and the bacterial enzyme cytosine deaminase (Dong et al. (1996) *Human Gene Therapy* 7:713-720). Cells which express these genes are rendered sensitive to the effects of the relatively nontoxic prodrugs ganciclovir (HSV-tk), cyclophosphamide (cytochrome P450 2B1), cytosine arabinoside (human deoxycytidine kinase) or 5-fluorocytosine (bacterial cytosine deaminase). Culver et al. (1992) *Science* 256:1550-1552, Huber et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:8302-8306.

A "selectable marker" or "reporter marker" refers to a nucleotide sequence included in a gene transfer vector that has no therapeutic activity, but rather is included to allow for simpler preparation, manufacturing, characterization or testing of the gene transfer vector.

A "specific binding agent" refers to a member of a specific binding pair of molecules wherein one of the molecules specifically binds to the second molecule through chemical and/or physical means. One example of a specific binding agent is an antibody directed against a selected antigen.

By "subject" is meant any member of the subphylum chordata, including, without limitation, humans and other primates, including non-human primates such as chimpanzees and other apes and monkey species; farm animals such as cattle, sheep, pigs, goats and horses; domestic mammals such as dogs and cats; laboratory animals including rodents such as mice, rats and guinea pigs; birds, including domestic, wild and game birds such as chickens, turkeys and other gallinaceous birds, ducks, geese, and the like. The term does not denote a particular age. Thus, both adult and newborn individuals are intended to be covered. The system described above is intended for use in any of the above vertebrate species, since the immune systems of all of these vertebrates operate similarly.

By "pharmaceutically acceptable" or "pharmacologically acceptable" is meant a material which is not biologically or otherwise undesirable, i.e., the material may be

WO 02/04493

PCT/US01/21241

administered to an individual in a formulation or composition without causing any undesirable biological effects or interacting in a deleterious manner with any of the components of the composition in which it is contained.

By "physiological pH" or a "pH in the physiological range" is meant a pH in the range of approximately 7.2 to 8.0 inclusive, more typically in the range of approximately 7.2 to 7.6 inclusive.

As used herein, "treatment" refers to any of (i) the prevention of infection or reinfection, as in a traditional vaccine, (ii) the reduction or elimination of symptoms, and (iii) the substantial or complete elimination of the pathogen in question. Treatment may be effected prophylactically (prior to infection) or therapeutically (following infection).

By "co-administration" is meant administration of more than one composition or molecule. Thus, co-administration includes concurrent administration or sequentially administration (in any order), via the same or different routes of administration. Non-limiting examples of co-administration regimes include, co-administration of nucleic acid and polypeptide; co-administration of different nucleic acids (e.g., different expression cassettes as described herein and/or different gene delivery vectors); and co-administration of different polypeptides (e.g., different HIV polypeptides and/or different adjuvants). The term also encompasses multiple administrations of one of the co-administered molecules or compositions (e.g., multiple administrations of one or more of the expression cassettes described herein followed by one or more administrations of a polypeptide-containing composition). In cases where the molecules or compositions are delivered sequentially, the time between each administration can be readily determined by one of skill in the art in view of the teachings herein.

"Lentiviral vector", and "recombinant lentiviral vector" refer to a nucleic acid construct which carries, and within certain embodiments, is capable of directing the expression of a nucleic acid molecule of interest. The lentiviral vector include at least one transcriptional promoter/enhancer or locus defining element(s), or other elements which control gene expression by other means such as alternate splicing, nuclear RNA export, post-translational modification of messenger, or post-transcriptional modification of protein. Such vector constructs must also include a packaging signal, long terminal repeats (LTRS) or portion thereof, and positive and negative strand primer binding sites appropriate to the retrovirus used (if these are not already present in the retroviral vector). Optionally, the

WO 02/04493

PCT/US01/21241

recombinant lentiviral vector may also include a signal which directs polyadenylation, selectable markers such as Neo, TK, hygromycin, phleomycin, histidinol, or DHFR, as well as one or more restriction sites and a translation termination sequence. By way of example, such vectors typically include a 5' LTR, a tRNA binding site, a packaging signal, an origin of second strand DNA synthesis, and a 3' LTR or a portion thereof

"Lentiviral vector particle" as utilized within the present invention refers to a lentivirus which carries at least one gene of interest. The retrovirus may also contain a selectable marker. The recombinant lentivirus is capable of reverse transcribing its genetic material (RNA) into DNA and incorporating this genetic material into a host cell's DNA upon infection. Lentiviral vector particles may have a lentiviral envelope, a non-lentiviral envelope (e.g., an amphi or VSV-G envelope), or a chimeric envelope.

"Nucleic acid expression vector" or "Expression cassette" refers to an assembly which is capable of directing the expression of a sequence or gene of interest. The nucleic acid expression vector includes a promoter which is operably linked to the sequences or gene(s) of interest. Other control elements may be present as well. Expression cassettes described herein may be contained within a plasmid construct. In addition to the components of the expression cassette, the plasmid construct may also include a bacterial origin of replication, one or more selectable markers, a signal which allows the plasmid construct to exist as single-stranded DNA (e.g., a M13 origin of replication), a multiple cloning site, and a "mammalian" origin of replication (e.g., a SV40 or adenovirus origin of replication).

"Packaging cell" refers to a cell which contains those elements necessary for production of infectious recombinant retrovirus which are lacking in a recombinant retroviral vector. Typically, such packaging cells contain one or more expression cassettes which are capable of expressing proteins which encode *Gag*, *pol* and *env* proteins.

"Producer cell" or "vector producing cell" refers to a cell which contains all elements necessary for production of recombinant retroviral vector particles.

2. MODES OF CARRYING OUT THE INVENTION

Before describing the present invention in detail, it is to be understood that this invention is not limited to particular formulations or process parameters as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments of the invention only, and is not intended to be limiting.

WO 02/04493

PCT/US01/21241

Although a number of methods and materials similar or equivalent to those described herein can be used in the practice of the present invention, the preferred materials and methods are described herein.

5 2.1. THE HIV GENOME

The HIV genome and various polypeptide-encoding regions are shown in Table A. The nucleotide positions are given relative to 8_5_TV1_C.ZA (SEQ ID NO:33, Figure 11). However, it will be readily apparent to one of ordinary skill in the art in view of the teachings of the present disclosure how to determine corresponding regions in other HIV strains or
 10 variants (e.g., isolates HIV_{HTB}, HIV_{SF2}, HIV-1_{SF162}, HIV-1_{SF170}, HIV_{LAIV}, HIV_{LAJ}, HIV_{ABN}, HIV-1_{CM2353}, HIV-1_{US4}, other HIV-1 strains from diverse subtypes (e.g., subtypes A through G, and O), HIV-2 strains and diverse subtypes (e.g., HIV-2_{UC1} and HIV-2_{UC2}), and simian immunodeficiency virus (SIV). (See, e.g., Virology, 3rd Edition (W.K. Joklik ed. 1988); *Fundamental Virology*, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); *Virology*, 3rd
 15 Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA; for a description of these and other related viruses), using for example, sequence comparison programs (e.g., BLAST and others described herein) or identification and alignment of structural features (e.g., a program such as the "ALB" program described herein that can identify the various regions).

WO 02/04493

PCT/US01/21241

Table A: Regions of the HIV Genome relative to 8_5_TV1_C.ZA

	Region	Position in nucleotide sequence
	5'LTR	1-636
	U3	1-457
5	R	458-553
	U5	554-636
	NFkB II	340-348
	NFkB I	354-362
	Sp1 UI	379-388
10	Sp1 II	390-398
	Sp1 J	400-410
	TATA Box	429-433
	TAR	474-499
	Poly A signal	529-534
15	PBS	638-655
	p7 binding region, packaging signal	685-791
20	Gag:	792-2285
	p17	792-1178
	p24	1179-1871
	Cyclophilin A bdg.	1395-1505
	MHR	1632-1694
25	p2	1872-1907
	p7	1908-2072
	Frameshift slip	2072-2078
	p1	2073-2120
	p6Gag	2121-2285
30	Zn-motif I	1950-1991
	Zn-motif II	2013-2054

WO 02/04493

PCT/US01/21241

	Pol:	2072-5086
	p6Pol	2072-2245
	Prot	2246-2542
	p66RT	2543-4210
5	p15RNaseH'	3857-4210
	p31Int	4211-5086
	Vif:	5034-5612
	Hydrophilic region	5292-5315
10	Vpr:	5552-5839
	Oligomerization	5552-5677
	Amphipathic α -helix	5597-5653
15	Tat:	5823-6038 and 8417-8509
	Tat-1 exon	5823-6038
	Tat-2 exon	8417-8509
	N-terminal domain	5823-5885
	Trans-activation domain	5886-5933
20	Transduction domain	5961-5993
	Rev:	5962-6037 and 8416-8663
	Rev-1 exon	5962-6037
	Rev-2 exon	8416-8663
25	High-affinity bdg. site	8439-8486
	Leu-rich effector domain	8562-8588
	Vpu:	6060-6326
	Transmembrane domain	6060-6161
30	Cytoplasmic domain	6162-6326

WO 02/04493

PCT/US01/21241

	Env (gp160):	6244-8853
	Signal peptide	6244-6324
	gp120	6325-7794
	V1	6628-6729
5	V2	6727-6852
	V3	7150-7254
	V4	7411-7506
	V5	7663-7674
	C1	6325-6627
10	C2	6853-7149
	C3	7255-7410
	C4	7507-7662
	C5	7675-7794
	CD4 binding	7540-7566
15	gp41	7795-8853
	Fusion peptide	7789-7842
	Oligomerization domain	7924-7959
	N-terminal heptad repeat	7921-8028
	C-terminal heptad repeat	8173-8280
20	Immunodominant region	8023-8076
	Nef:	8855-9478
	Myristoylation	8858-8875
	SH3 binding	9062-9091
25	Polypurine tract	9128-9154
	SH3 binding	9296-9307

It will be readily apparent that one of skill in the art can readily align any sequence to that shown in Table A to determine relative locations of any particular HIV gene. For example, using one of the alignment programs described herein (e.g., BLAST), other HIV Type C sequences can be aligned with S_5_TV1_C.ZA (Table A) and locations of genes determined.

Polypeptide sequences can be similarly aligned. For example, Figure 103 shows the alignment of Env polypeptide sequences from various strains, relative to SF-162. As described in detail in co-owned WO/99303, Env polypeptides (e.g., gp120, gp140 and gp160) include a "bridging sheet" comprised of 4 anti-parallel β -strands (β -2, β -3, β -20 and β -21) that form a β -sheet. Extruding from one pair of the β -strands (β -2 and β -3) are two loops, V1

WO 02/04493

PCT/US01/21241

and V2. The β -2 sheet occurs at approximately amino acid residue 113 (Cys) to amino acid residue 117 (Thr) while β -3 occurs at approximately amino acid residue 192 (Ser) to amino acid residue 194 (Ile), relative to SF-162 (see, Figure 103). The "V1/V2 region" occurs at approximately amino acid positions 120 (Cys) to residue 189 (Cys), relative to SF-162.

- 5 Extruding from the second pair of β -strands (β -20 and β -21) is a "small-loop" structure, also referred to herein as "the bridging sheet small loop." The locations of both the small loop and bridging sheet small loop can be determined relative to HXB-2 following the teachings herein and in WO/99303. Also shown by arrows in Figure 103A-C are approximate sites for deletions sequence from the beta sheet region. The "*" denotes N-glycosylation sites that can
10 be mutated following the teachings of the present specification.

2.2 SYNTHETIC EXPRESSION CASSETTES

2.2.1 MODIFICATION OF HIV-1-TYPE C *POL*, *PROT*, *RT*, *INT*, *GAG*, *ENV*, *TAT*, *REV*, *NEF*, *RNAseH*, *VIF*, *VPR*, AND *VPU* NUCLEIC ACID CODING SEQUENCES

- 15 One aspect of the present invention is the generation of HIV-1 type C coding sequences, and related sequences, having improved expression relative to the corresponding wild-type sequences.

2.2.1.1. MODIFICATION OF *GAG* NUCLEIC ACID CODING SEQUENCES

- 20 An exemplary embodiment of the present invention is illustrated herein by modifying the Gag protein wild-type sequences obtained from the AF110965 and AF110967 strains of HIV-1, subtype C. (see, for example, Korber et al. (1998) *Human Retroviruses and Aids*, Los Alamos, New Mexico: Los Alamos National Laboratory; Novitsky et al. (1999) *J. Virol.* 73(5):4427-4432, for molecular cloning of various subtype C clones from Botswana). Also illustrated herein is the modification of wild-type sequences
25 from novel isolates 8_5_TV1_C.ZA (also called TV001 or TV1) and 12-5_1_TV2_C.ZA (also called TV002 or TV2). SEQ ID NO:52 shows the wild-type sequence of Gag from 8_5_TV1_C.ZA and SEQ ID NO:54 shows the wild-type sequence of the major homology region of Gag (nucleotides 1632-1694 of Table A) of the same strain. SEQ ID NO:100
30 shows the wild-type sequence of Gag of 12-5_1_TV2_C.ZA.

WO 02/04493

PCT/US01/21241

Gag sequence obtained from other Type C HIV-1 variants may be manipulated in similar fashion following the teachings of the present specification. Such other variants include, but are not limited to, Gag protein encoding sequences obtained from the isolates of HIV-1 Type C, for example as described in Novitsky et al., (1999), *supra*; Myers et al., *infra*; 5 Virology, 3rd Edition (W.K. Joklik ed. 1988); *Fundamental Virology*, 2nd Edition (B.N. Fields and D.M. Knipe, eds. 1991); *Virology*, 3rd Edition (Fields, BN, DM Knipe, PM Howley, Editors, 1996, Lippincott-Raven, Philadelphia, PA and on the World Wide Web (Internet), for example at <http://hiv-web.lanl.gov/cgi-bin/hivDB3/public/wdb/ssampublic> and <http://hiv-web.lanl.gov>.

10 First, the HIV-1 codon usage pattern was modified so that the resulting nucleic acid coding sequence was comparable to codon usage found in highly expressed human genes (Example 1). The HIV codon usage reflects a high content of the nucleotides A or T of the codon-triplet. The effect of the HIV-1 codon usage is a high AT content in the DNA sequence that results in a decreased translation ability and instability of the mRNA. In 15 comparison, highly expressed human codons prefer the nucleotides G or C. The Gag coding sequences were modified to be comparable to codon usage found in highly expressed human genes.

Second, there are inhibitory (or instability) elements (INS) located within the coding sequences of the Gag coding sequences. The RRE is a secondary RNA structure that 20 interacts with the HIV encoded Rev-protein to overcome the expression down-regulating effects of the INS. To overcome the post-transcriptional activating mechanisms of RRE and Rev, the instability elements can be inactivated by introducing multiple point mutations that do not alter the reading frame of the encoded proteins.

Subtype C Gag-encoding sequences having inactivated RRE sites are shown, for example, in 25 Figures 1 (SEQ ID NO:3), 2 (SEQ ID NO:4), 5 (SEQ ID NO:20) and 6 (SEQ ID NO:26). Similarly, other synthetic polynucleotides derived from other Subtype C strains can be modified to inactivate the RRE sites.

Modification of the Gag polypeptide coding sequences results in improved expression 30 relative to the wild-type coding sequences in a number of mammalian cell lines (as well as other types of cell lines, including, but not limited to, insect cells). Further, expression of the sequences results in production of virus-like particles (VLPs) by these cell lines (see below).

WO 02/04493

PCT/US01/21241

2.2.1.2 MODIFICATION OF ENV NUCLEIC ACID CODING SEQUENCES

Similarly, the present invention also includes synthetic Env-encoding polynucleotides and modified Env proteins. Wild-type Env sequences are obtained from the AF110968 and AF110975 strains as well as novel strains 8_5_TV1_C.ZA (SEQ ID NO:33) and 12-5_1_TV2_C.ZA (SEQ ID NO:45) of HIV-1, type C. (see, for example, Novitsky et al. (1999) *J. Virol.* 73(5):4427-4432, for molecular cloning of various subtype C clones from Botswana). Wild-type Env sequences of 8_5_TV1_C.ZA are shown, for example, in SEQ ID NO:48 (wild-type Env common region, nucleotides 7486-7629 as shown in Table A); and SEQ ID NO:50 (wild type gp160, nucleotides 6244-8853 as shown in Table A). Wild-type Env gp160 of 12-5_1_TV2_C.ZA is shown in SEQ ID NO:98. It will be readily apparent from the disclosure herein that polynucleotides encoding fragments of Env gp160 (e.g., gp120, gp41, gp140) can be readily obtained from the larger, full-length sequences disclosed herein. It will also be readily apparent that other modifications can be made, for example deletion of regions such as the V1 and/or V2 region; mutation of the cleavage site and the like (see, Example 1). Exemplary sequences of such modification as shown in SEQ ID NO:119 through 127.

Further, Env sequences obtained from other Type C HIV-1 variants may be manipulated in similar fashion following the teachings of the present specification. Such other variants include, but are not limited to, Env protein encoding sequences obtained from the isolates of HIV-1 Type C, described above.

The codon usage pattern for Env was modified as described above for Gag so that the resulting nucleic acid coding sequence was comparable to codon usage found in highly expressed human genes. Experiments performed in support of the present invention show that the synthetic Env sequences were capable of higher level of protein production relative to the native Env sequences.

Modification of the Env polypeptide coding sequences results in improved expression relative to the wild-type coding sequences in a number of mammalian cell lines (as well as other types of cell lines, including, but not limited to, insect cells). Similar Env polypeptide coding sequences can be obtained, modified and tested for improved expression from a variety of isolates, including those described above for Gag.

Further modifications of Env include, but are not limited to, generating polynucleotides that encode Env polypeptides having mutations and/or deletions therein. For

WO 02/04493

PCT/US01/21241

instance, the hypervariable regions, V1 and/or V2, can be deleted as described herein. Additionally, other modifications, for example to the bridging sheet region and/or to N-glycosylation sites within Env can also be performed following the teachings of the present specification. (see, Figure 103A-C and WO/39303). Various combinations of these
 5 modifications can be employed to generate synthetic expression cassettes as described herein.

2.2.1.3 MODIFICATION OF SEQUENCES INCLUDING HIV-1 POL NUCLEIC ACID CODING SEQUENCES

The present invention also includes expression cassettes which include synthetic Pol sequences. As noted above, "Pol" includes, but is not limited to, the protein-encoding regions shown in Figure 7, for example polymerase, protease, reverse transcriptase and/or integrase-containing sequences. The regions shown in Figure 7 are described, for example, in Wan et al (1996) *Biochem. J.* 316:569-573; Kohl et al. (1988) *PNAS USA* 85:4686-4690; Krausslich et al. (1988) *J. Virol.* 62:4393-4397; Coffin, "Retroviridae and their Replication" in *Virology*, pp1437-1500 (Raven, New York, 1990); Patel et. al. (1995) *Biochemistry*
 15 34:5351-5363. Thus, the synthetic expression cassettes exemplified herein include one or more of these regions and one or more changes to the resulting amino acid sequences.

Wild type Pol sequences were obtained from the AF110975, 8_5_TV1_C.ZA and 12-5_1_TV2_C.ZA strains of HIV-1, type C. (see, for example, Novitsky et al. (1999) *J. Virol.* 73(5):4427-4432, for molecular cloning of various subtype C clones from Botswana). SEQ ID NO:34 shows the wild type sequence of AF110975 from the p2 through p7 region of Pol (see, Figure 7 and Table A). SEQ ID NO:35 shows the wild type sequence of AF110975 from p1 through the first 6 amino acids of integrase (see, Figure 7 and Table A). SEQ ID NO:63 and SEQ ID NO:104 show wild-type sequences of Pol from 8_5_TV1_C.ZA and 12-5_1_TV2_C.ZA, respectively (see, also, Table A).
 25

Sequence obtained from other Type C HIV-1 variants may be manipulated in similar fashion following the teachings of the present specification. Such other variants include, but are not limited to, Pol protein encoding sequences obtained from the isolates of HIV-1 Type C described herein.

30 The codon usage pattern for Pol was modified as described above for Gag and Env so that the resulting nucleic acid coding sequence was comparable to codon usage found in highly expressed human genes.

WO 02/04493

PCT/US01/21241

Table B shows the nucleotide positions of various regions found in the Pol constructs exemplified herein (e.g., SEQ ID NOs: 30-32).

Table B

Region	Position in nucleotide sequence in construct		
	PR975(+)	PR975YM	PR975(+) YMWM
	Seq Id No:30	Seq Id No:31	Seq Id No:32
Sal I restriction site	1-6	1-6	1-6
Kozak start codon	7-16	7-16	7-16
p2	16-54	16-54	16-54
p7	55-219	55-219	55-219
p1/p6 pol	220-375	220-375	220-375
Insertion mutation for in frame	225	225	225
p10Protease	376-672	376-672	376-672
p6GRT	673-2352	673-2346	673-2340
p5IRT	673-1992	673-1986	673-1980
p15RNaseH	1993-2352	1993-2346	1993-2340
catalytic center region (YMDD)	1219-1230	1219-1224	1219-1224
primer grip region (WMGY)	1357-1368	1351-1362	1351-1356
Gaa Integrase	2353-2370	2347-2364	2341-2358
YMDD epitope cassette (incl. 5'+3'Gly)	2371-2424	2365-2418	2359-2412
MCS (multiple cloning site)	2425-2463	2419-2457	2413-2451
EcoR I restriction site	2464-2469	2458-2463	2452-2457

As shown in Table B, exemplary constructs were modified in various ways. For example, the expression constructs exemplified herein include sequence that encodes the first 6 amino acids of the integrase polypeptide. This 6 amino acid region is believed to provide a cleavage recognition site recognized by HIV protease (see, e.g., McCormack et al. (1997) *FEBS Letts* 414:84-88). As noted above, certain constructs exemplified herein include a multiple cloning site (MCS) for insertion of one or more transgenes, typically at the 3' end of the construct. In addition, a cassette encoding a catalytic center epitope derived from the catalytic center in RT is typically included 3' of the sequence encoding 6 amino acids of integrase. This cassette (SEQ ID NO:36) encodes Ile178 through Serine 191 of RT (amino acids 3 through 16 of SEQ ID NO:37) and was added to keep this well conserved region as a possible CTL epitope. Further, the constructs contain an insertion mutations (position 225 of SEQ ID NOs:30 to 32) to preserve the reading frame. (see, e.g., Park et al. (1991) *J. Virol.* 65:5111).

WO 02/04493

PCT/US01/21241

In certain embodiments, the catalytic center and/or primer grip region of RT are modified. The catalytic center and primer grip regions of RT are described, for example, in Patel et al. (1995) *Biochem.* 34:5351 and Palaniappan et al. (1997) *J. Biol. Chem.* 272(17):11157. For example, in the construct designated PR975YM (SEQ ID NO:31), wild type sequence encoding the amino acids YMDD at positions 183-185 of p66 RT, numbered relative to AF110975, are replaced with sequence encoding the amino acids "AP". In the construct designated PR975YMWM (SEQ ID NO:32), the same mutation in YMDD is made and, in addition, the primer grip region (amino acids WMGY, residues 229-232 of p66RT, numbered relative to AF110975) are replaced with sequence encoding the amino acids "PI."

For the Pol sequence, the changes in codon usage are typically restricted to the regions up to the -1 frameshift and starting again at the end of the Gag reading frame; however, regions within the frameshift translation region can be modified as well. Finally, inhibitory (or instability) elements (INS) located within the coding sequences of the protease polypeptide coding sequence can be altered as well.

Experiments can be performed in support of the present invention to show that the synthetic Pol sequences were capable of higher level of protein production relative to the native Pol sequences. Modification of the Pol polypeptide coding sequences results in improved expression relative to the wild-type coding sequences in a number of mammalian cell lines (as well as other types of cell lines, including, but not limited to, insect cells). Similar Pol polypeptide coding sequences can be obtained, modified and tested for improved expression from a variety of isolates, including those described above for Gag and Env.

2.2.1.4 MODIFICATION OF OTHER HIV SEQUENCES

The present invention also includes expression cassettes which include synthetic HIV Type C sequences derived HIV genes other than Gag, Env and Pol, including but not limited to, regions within Gag, Env, Pol, as well as, *vif*, *vpr*, *tat*, *rev*, *vpu*, and *nef*, for example from 8_5_TV1_C.ZA (SEQ ID NO:33) or 12-5_1_TV2_C.ZA (SEQ ID NO:45). Sequences obtained from other strains can be manipulated in similar fashion following the teachings of the present specification.

As noted above, the codon usage pattern is modified as described above for Gag, Env and Pol so that the resulting nucleic acid coding sequence is comparable to codon usage found in highly expressed human genes. Experiments can be performed in support of the present invention to show that these synthetic sequences were capable of higher level of protein production relative to the native sequences and that modification of the wild-type polypeptide coding sequences results in improved expression relative to the wild-type coding

WO 02/04493

PCT/US01/21241

sequences in a number of mammalian cell lines (as well as other types of cell lines, including, but not limited to, insect cells). Furthermore, the nucleic acid sequence can also be modified to introduce mutations into one or more regions of the gene, for instance to render the gene product non-functional and/or to eliminate the myristoylation site in Nef.

5 Synthetic expression cassettes exemplified herein include SEQ ID NO:49 and SEQ ID NO:97 (Env gp160-encoding sequences, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:51 and SEQ ID NO:99 (Gag-encoding sequences modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:53 (Gag major homology region, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:55 and SEQ ID NO:101 (Nef-encoding sequences, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:57 and SEQ ID NO:134 (Nef-encoding sequences with a mutation at position 125 resulting in a non-functional gene product, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA, respectively); SEQ ID NO:58 (RNaseH-encoding sequences, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:60 (Integrase-encoding sequences, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:62 and SEQ ID NO:103 (Pol-encoding sequences, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:64 (Protease-encoding sequences, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:66 (inactivated protease-encoding sequences, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:68 (inactivated protease and RT mutated sequences, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:70 (protease and reverse-transcriptase-encoding sequences, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:72 and SEQ ID NO:105 (exon 1 of Rev, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:74 and SEQ ID NO:107 (exon 2 of Rev, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:76 (reverse transcriptase-encoding sequences, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:78 (mutated reverse-transcriptase, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:80 (exon 1 of Tat including a mutation that results in non-functional Tat, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:81 and SEQ ID NO:109 (exon 1 of Tat, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:83 and SEQ ID NO:111 (exon 2 of Tat, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:85 and SEQ ID NO:113 (Vif-encoding sequences, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ

WO 02/04493

PCT/US01/21241

ID NO:87 and SEQ ID NO:115 (Vpr-encoding sequences, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:89 and SEQ ID NO:117 (Vpu-encoding sequences, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA wild-type, respectively); SEQ ID NO:91 (sequences of exons 1 and 2 of Rev, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:93 (sequences of mutated exon 1 of Tat and exon 2 of Tat, where mutation of exon 1 results in non-functional Tat, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:94 (sequences of exons 1 and 2 of Tat, modified based on 8_5_TV1_C.ZA wild type); SEQ ID NO:96 and SEQ ID NO:135 (Nef-encoding sequences including a mutation to eliminate myristoylation site, modified based on 8_5_TV1_C.ZA wild type and 12-5_1_TV2_C.ZA, respectively).

2.2.1.5 FURTHER MODIFICATION OF SEQUENCES INCLUDING HIV-1 NUCLEIC ACID CODING SEQUENCES

The Type C HIV polypeptide-encoding expression cassettes described herein may also contain one or more further sequences encoding, for example, one or more transgenes. Further sequences (e.g., transgenes) useful in the practice of the present invention include, but are not limited to, further sequences are those encoding further viral epitopes/antigens (including but not limited to, HCV antigens (e.g., E1, E2; Houghton, M., et al., U.S. Patent No. 5,714,596, issued February 3, 1998; Houghton, M., et al., U.S. Patent No. 5,712,088, issued January 27, 1998; Houghton, M., et al., U.S. Patent No. 5,683,864, issued November 4, 1997; Weiner, A.J., et al., U.S. Patent No. 5,728,520, issued March 17, 1998; Weiner, A.J., et al., U.S. Patent No. 5,766,845, issued June 16, 1998; Weiner, A.J., et al., U.S. Patent No. 5,670,152, issued September 23, 1997), HIV antigens (e.g., derived from *tat*, *rev*, *nef* and/or *env*); and sequences encoding tumor antigens/epitopes. Further sequences may also be derived from non-viral sources, for instance, sequences encoding cytokines such interleukin-2 (IL-2), stem cell factor (SCF), interleukin 3 (IL-3), interleukin 6 (IL-6), interleukin 12 (IL-12), G-CSF, granulocyte macrophage-colony stimulating factor (GM-CSF), interleukin-1 alpha (IL-1 α), interleukin-11 (IL-11), MIP-1 α , tumor necrosis factor (TNF), leukemia inhibitory factor (LIF), c-kit ligand, thrombopoietin (TPO) and flt3 ligand, commercially available from several vendors such as, for example, Genzyme (Framingham, MA), Genentech (South San Francisco, CA), Amgen (Thousand Oaks, CA), R&D Systems and ImmuneX (Seattle, WA). Additional sequences are described below, for example in Section 2.3. Also, variations on the orientation of the Gag and other coding sequences, relative to each other, are described below.

WO 02/04493

PCT/US01/21241

HIV polypeptide coding sequences can be obtained from other Type C HIV isolates, see, e.g., Myers et al. Los Alamos Database, Los Alamos National Laboratory, Los Alamos, New Mexico (1992); Myers et al., *Human Retroviruses and Aids*, 1997, Los Alamos, New Mexico: Los Alamos National Laboratory. Synthetic expression cassettes can be generated using such coding sequences as starting material by following the teachings of the present specification (e.g., see Example 1).

Further, the synthetic expression cassettes of the present invention include related polypeptide sequences having greater than 85%, preferably greater than 90%, more preferably greater than 95%, and most preferably greater than 98% sequence identity to the synthetic expression cassette sequences disclosed herein (for example, (SEQ ID NOs:30-32; SEQ ID NOs: 3, 4, 20, and 21 and SEQ ID NOs:5-17). Various coding regions are indicated in Figures 3 and 4, for example in Figure 3 (AF110968), nucleotides 1-81 (SEQ ID NO:18); nucleotides 82-1512 (SEQ ID NO:6) encode a gp120 polypeptide, nucleotides 1513 to 2547 (SEQ ID NO:10) encode a gp41 polypeptide, nucleotides 82-2025 (SEQ ID NO:7) encode a gp140 polypeptide and nucleotides 82-2547 (SEQ ID NO:8) encode a gp160 polypeptide. Similarly, in Figure 98 (SEQ ID NO:127, strain 8_2_TV1_C.ZA), nucleotides 1-6 are an EcoRI restriction site; nucleotides 7-87 a encode a wild-type (from 8_2_TV1_C.ZA) leader signal peptide; nucleotides 88 to 1563 encode a gp120 polypeptide; nucleotides 88 to 2064 encode a gp140 polypeptide; nucleotides 88 to 2607 encode a gp160 polypeptide.

2.2.3 EXPRESSION OF SYNTHETIC SEQUENCES ENCODING HIV-1 SUBTYPE C AND RELATED POLYPEPTIDES

Synthetic HIV-encoding sequences (expression cassettes) of the present invention can be cloned into a number of different expression vectors to evaluate levels of expression and, in the case of Gag, production of VLPs. The synthetic DNA fragments for HIV polypeptides can be cloned into eucaryotic expression vectors, including, a transient expression vector, CMV-promoter-based mammalian vectors, and a shuttle vector for use in baculovirus expression systems. Corresponding wild-type sequences can also be cloned into the same vectors.

These vectors can then be transfected into a several different cell types, including a variety of mammalian cell lines (293, RD, COS-7, and CHO, cell lines available, for example, from the A.T.C.C.). The cell lines are then cultured under appropriate conditions and the levels of any appropriate polypeptide product can be evaluated in supernatants. (see, Table A and Example 2). For example, p24 can be used to evaluate Gag expression; gp160, gp140 or gp120 can be used to evaluate Env expression; p66pol can be used to evaluate Pol

WO 02/04493

PCT/US01/21241

expression; prot can be used to evaluate protease; p15 for RNaseH; p31 for Integrase; and other appropriate polypeptides for Vif, Vpr, Tat, Rev, Vpu and Nef. Further, modified polypeptides can also be used, for example, other Env polypeptides include, but are not limited to, for example, native gp160, oligomeric gp140, monomeric gp120 as well as modified and/or synthetic sequences of these polypeptides. The results of these assays demonstrate that expression of synthetic HIV polypeptide-encoding sequences are significantly higher than corresponding wild-type sequences.

Further, Western Blot analysis can be used to show that cells containing the synthetic expression cassette produce the expected protein at higher per-cell concentrations than cells containing the native expression cassette. The HIV proteins can be seen in both cell lysates and supernatants. The levels of production are significantly higher in cell supernatants for cells transfected with the synthetic expression cassettes of the present invention.

Fractionation of the supernatants from mammalian cells transfected with the synthetic expression cassette can be used to show that the cassettes provide superior production of HIV proteins and, in the case of Gag, VLPs, relative to the wild-type sequences.

Efficient expression of these HIV-containing polypeptides in mammalian cell lines provides the following benefits: the polypeptides are free of baculovirus contaminants; production by established methods approved by the FDA; increased purity; greater yields (relative to native coding sequences); and a novel method of producing the Subtype C HIV-containing polypeptides in CHO cells which is not feasible in the absence of the increased expression obtained using the constructs of the present invention. Exemplary Mammalian cell lines include, but are not limited to, BHK, VBRO, IIT1080, 293, 293T, RD, COS-7, CHO, Jurkat, HUT, SUPT, C8166, MOLT4/clone8, MT-2, MT-4, II9, PM1, CEM, and CBMX174, such cell lines are available, for example, from the A.T.C.C.).

A synthetic Gag expression cassette of the present invention will also exhibit high levels of expression and VLP production when transfected into insect cells. Synthetic expression cassettes described herein also demonstrate high levels of expression in insect cells. Further, in addition to a higher total protein yield, the final product from the synthetic polypeptides consistently contains lower amounts of contaminating baculovirus proteins than the final product from the native Type C sequences.

Further, synthetic expression cassettes of the present invention can also be introduced into yeast vectors which, in turn, can be transformed into and efficiently expressed by yeast cells (*Saccharomyces cerevisiae*; using vectors as described in Rosenberg, S. and Tekamp-Olson, P., U.S. Patent No. RE35,749, issued, March 17, 1998).

WO 02/04493

PCT/US01/21241

In addition to the mammalian and insect vectors, the synthetic expression cassettes of the present invention can be incorporated into a variety of expression vectors using selected expression control elements. Appropriate vectors and control elements for any given cell type can be selected by one having ordinary skill in the art in view of the teachings of the present specification and information known in the art about expression vectors.

For example, a synthetic expression cassette can be inserted into a vector which includes control elements operably linked to the desired coding sequence, which allow for the expression of the gene in a selected cell-type. For example, typical promoters for mammalian cell expression include the SV40 early promoter, a CMV promoter such as the CMV immediate early promoter (a CMV promoter can include intron A), RSV, HIV-Ltr, the mouse mammary tumor virus LTR promoter (MMTV-Ltr), the adenovirus major late promoter (Ad MLP), and the herpes simplex virus promoter, among others. Other nonviral promoters, such as a promoter derived from the murine metallothionein gene, will also find use for mammalian expression. Typically, transcription termination and polyadenylation sequences will also be present, located 3' to the translation stop codon. Preferably, a sequence for optimization of initiation of translation, located 5' to the coding sequence, is also present. Examples of transcription terminator/polyadenylation signals include those derived from SV40, as described in Sambrook, et al., *supra*, as well as a bovine growth hormone terminator sequence. Introns, containing splice donor and acceptor sites, may also be designed into the constructs for use with the present invention (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986).

Enhancer elements may also be used herein to increase expression levels of the mammalian constructs. Examples include the SV40 early gene enhancer, as described in Dijkema et al., *EMBO J.* (1985) 4:761, the enhancer/promoter derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus, as described in Gorman et al., *Proc. Natl. Acad. Sci. USA* (1982b) 79:6777 and elements derived from human CMV, as described in Boshart et al., *Cell* (1985) 41:521, such as elements included in the CMV intron A sequence (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986).

The desired synthetic polypeptide encoding sequences can be cloned into any number of commercially available vectors to generate expression of the polypeptide in an appropriate host system. These systems include, but are not limited to, the following: baculovirus expression (Reilly, P.R., et al., BACULOVIRUS EXPRESSION VECTORS: A LABORATORY MANUAL (1992); Beames, et al., *Biotechniques* 11:378 (1991); Pharmingen; Clontech, Palo Alto, CA)), vaccinia expression (Earl, P. L., et al., "Expression of proteins in mammalian cells using vaccinia" In *Current Protocols in Molecular Biology* (F. M. Ausubel, et al. Eds.),

WO 02/04493

PCT/US01/21241

- Greene Publishing Associates & Wiley Interscience, New York (1991); Moss, B., *et al.*, U.S. Patent Number 5,135,855, issued 4 August 1992), expression in bacteria {Ausubel, F.M., *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley and Sons, Inc., Media PA; Clontech}, expression in yeast {Rosenberg, S. and Tekamp-Olson, P., U.S. Patent No. RE35,749, issued, March 17, 1998; Shuster, J.R., U.S. Patent No. 5,629,203, issued May 13, 1997; Gellissen, G., *et al.*, *Antonie Van Leeuwenhoek*, 62(1-2):79-93 (1992); Romanos, M.A., *et al.*, *Yeast* 8(6):423-488 (1992); Goeddel, D.V., *Methods in Enzymology* 185 (1990); Guthrie, C., and G.R. Fink, *Methods in Enzymology* 194 (1991)}, expression in mammalian cells {Clontech; Gibco-BRL, Grand Island, NY; e.g., Chinese hamster ovary (CHO) cell lines {Haynes, J., *et al.*, *Nuc. Acid. Res.* 11:687-706 (1983); 1983, Lau, Y.F., *et al.*, *Mol. Cell Biol.* 4:1469-1475 (1984); Kaufman, R. J., "Selection and coamplification of heterologous genes in mammalian cells," in *Methods in Enzymology*, vol. 185, pp537-566. Academic Press, Inc., San Diego CA (1991)}, and expression in plant cells {plant cloning vectors, Clontech Laboratories, Inc., Palo Alto, CA, and Pharmacia LKB Biotechnology, Inc., Piscataway, NJ; Hood, E., *et al.*, *J. Bacteriol.* 168:1291-1301 (1986); Nagel, R., *et al.*, *FRMS Microbiol. Lett.* 67:325 (1990); An, *et al.*, "Binary Vectors", and others in *Plant Molecular Biology Manual* A3:1-19 (1988); Miki, B.L.A., *et al.*, pp.249-265, and others in *Plant DNA Infectious Agents* (Hohn, T., *et al.*, eds.) Springer-Verlag, Wien, Austria, (1987); *Plant Molecular Biology: Essential Techniques*, P.G. Jones and J.M. Sutton, New York, J. Wiley, 1997; Miglani, Gurbachan *Dictionary of Plant Genetics and Molecular Biology*, New York, Food Products Press, 1998; Henry, R. J., *Practical Applications of Plant Molecular Biology*, New York, Chapman & Hall, 1997}.

Also included in the invention is an expression vector, containing coding sequences and expression control elements which allow expression of the coding regions in a suitable host. The control elements generally include a promoter, translation initiation codon, and translation and transcription termination sequences, and an insertion site for introducing the insert into the vector. Translational control elements have been reviewed by M. Kozak (e.g., Kozak, M., *Mamm. Genome* 7(8):563-574, 1996; Kozak, M., *Biochimie* 76(9):815-821, 1994; Kozak, M., *J Cell Biol* 108(2):229-241, 1989; Kozak, M., and Shatkin, A.J., *Methods Enzymol* 60:360-375, 1979).

Expression in yeast systems has the advantage of commercial production. Recombinant protein production by vaccinia and CHO cell line have the advantage of being mammalian expression systems. Further, vaccinia virus expression has several advantages including the following: (i) its wide host range; (ii) faithful post-transcriptional modification, processing, folding, transport, secretion, and assembly of recombinant proteins; (iii) high

WO 02/04493

PCT/US01/21241

level expression of relatively soluble recombinant proteins; and (iv) a large capacity to accommodate foreign DNA.

The recombinantly expressed polypeptides from synthetic HIV polypeptide-encoding expression cassettes are typically isolated from lysed cells or culture media. Purification can be carried out by methods known in the art including salt fractionation, ion exchange chromatography, gel filtration, size-exclusion chromatography, size-fractionation, and affinity chromatography. Immunoaffinity chromatography can be employed using antibodies generated based on, for example, HIV antigens.

Advantages of expressing the proteins of the present invention using mammalian cells include, but are not limited to, the following: well-established protocols for scale-up production; the ability to produce VLPs; cell lines are suitable to meet good manufacturing process (GMP) standards; culture conditions for mammalian cells are known in the art.

Various forms of the different embodiments of the invention, described herein, may be combined.

2.3 PRODUCTION OF VIRUS-LIKE PARTICLES AND USE OF THE CONSTRUCTS OF THE PRESENT INVENTION TO CREATE PACKAGING CELL LINES.

The group-specific antigens (Gag) of human immunodeficiency virus type-1 (HIV-1) self-assemble into noninfectious virus-like particles (VLP) that are released from various eucaryotic cells by budding (reviewed by Freed, E.O., *Virology* 251:1-15, 1998). The synthetic expression cassettes of the present invention provide efficient means for the production of HIV-Gag virus-like particles (VLPs) using a variety of different cell types, including, but not limited to, mammalian cells.

Viral particles can be used as a matrix for the proper presentation of an antigen entrapped or associated therewith to the immune system of the host.

2.3.1 VLP PRODUCTION USING THE SYNTHETIC EXPRESSION CASSETTES OF THE PRESENT INVENTION

Experiments can be performed in support of the present invention to demonstrate that the synthetic expression cassettes of the present invention provide superior production of both Gag proteins and VLPs, relative to native Gag coding sequences. Further, electron microscopic evaluation of VLP production can show that free and budding immature virus particles of the expected size are produced by cells containing the synthetic expression cassettes.

WO 02/04493

PCT/US01/21241

Using the synthetic expression cassettes of the present invention, rather than native Gag coding sequences, for the production of virus-like particles provide several advantages. First, VLPs can be produced in enhanced quantity making isolation and purification of the VLPs easier. Second, VLPs can be produced in a variety of cell types using the synthetic expression cassettes, in particular, mammalian cell lines can be used for VLP production, for example, CHO cells. Production using CHO cells provides (i) VLP formation; (ii) correct myristoylation and budding; (iii) absence of non-mammalian cell contaminants (e.g., insect viruses and/or cells); and (iv) ease of purification. The synthetic expression cassettes of the present invention are also useful for enhanced expression in cell-types other than mammalian cell lines. For example, infection of insect cells with baculovirus vectors encoding the synthetic expression cassettes results in higher levels of total Gag protein yield and higher levels of VLP production (relative to wild-type coding sequences). Further, the final product from insect cells infected with the baculovirus-Gag synthetic expression cassettes consistently contains lower amounts of contaminating insect proteins than the final product when wild-type coding sequences are used.

VLPs can spontaneously form when the particle-forming polypeptide of interest is recombinantly expressed in an appropriate host cell. Thus, the VLPs produced using the synthetic expression cassettes of the present invention are conveniently prepared using recombinant techniques. As discussed below, the Gag polypeptide encoding synthetic expression cassettes of the present invention can include other polypeptide coding sequences of interest (for example, HIV protease, HIV polymerase, HCV core, Env, synthetic Env; see, Example 1). Expression of such synthetic expression cassettes yields VLPs comprising the Gag polypeptide, as well as, the polypeptide of interest.

Once coding sequences for the desired particle-forming polypeptides have been isolated or synthesized, they can be cloned into any suitable vector or replicon for expression. Numerous cloning vectors are known to those of skill in the art, and the selection of an appropriate cloning vector is a matter of choice. See, generally, Sambrook et al, *supra*. The vector is then used to transform an appropriate host cell. Suitable recombinant expression systems include, but are not limited to, bacterial, mammalian, baculovirus/insect, vaccinia, Semliki Forest virus (SFV), Alphaviruses (such as, Sindbis, Venezuelan Equine Encephalitis (VEE)), mammalian, yeast and *Xenopus* expression systems, well known in the art. Particularly preferred expression systems are mammalian cell lines, vaccinia, Sindbis, insect and yeast systems.

WO 02/04493

PCT/US01/21241

For example, a number of mammalian cell lines are known in the art and include immortalized cell lines available from the American Type Culture Collection (A.T.C.C.), such as, but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), as well as others. Similarly, bacterial hosts such as *E. coli*, *Bacillus subtilis*, and *Streptococcus spp.*, will find use with the present expression constructs. Yeast hosts useful in the present invention include *inter alia*, *Saccharomyces cerevisiae*, *Candida albicans*, *Candida maltosa*, *Hansenula polymorpha*, *Kluyveromyces fragilis*, *Kluyveromyces lactis*, *Pichia guilliermondii*, *Pichia pastoris*, *Schizosaccharomyces pombe* and *Yarrowia lipolytica*. Insect cells for use with baculovirus expression vectors include, *inter alia*, *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni*. See, e.g., Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987).

Viral vectors can be used for the production of particles in eucaryotic cells, such as those derived from the pox family of viruses, including vaccinia virus and avian poxvirus. Additionally, a vaccinia based infection/transfection system, as described in Tomei et al., *J. Virol.* (1993) 67:4017-4026 and Selby et al., *J. Gen. Virol.* (1993) 74:1103-1113, will also find use with the present invention. In this system, cells are first infected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the DNA of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into protein by the host translational machinery. Alternately, T7 can be added as a purified protein or enzyme as in the "Progenitor" system (Studier and Moffatt, *J. Mol. Biol.* (1986) 182:113-130). The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation product(s).

Depending on the expression system and host selected, the VLPS are produced by growing host cells transformed by an expression vector under conditions whereby the particle-forming polypeptide is expressed and VLPs can be formed. The selection of the appropriate growth conditions is within the skill of the art. If the VLPs are formed intracellularly, the cells are then disrupted, using chemical, physical or mechanical means, which lyse the cells yet keep the VLPs substantially intact. Such methods are known to those of skill in the art and are described in, e.g., *Protein Purification Applications: A Practical Approach*, (E.L.V. Harris and S. Angal, Eds., 1990).

WO 02/04493

PCT/US01/21241

The particles are then isolated (or substantially purified) using methods that preserve the integrity thereof, such as, by gradient centrifugation, e.g., cesium chloride (CsCl) sucrose gradients, pelleting and the like (see, e.g., Kimbauer et al. *J. Virol.* (1993) 67:6929-6936), as well as standard purification techniques including, e.g., ion exchange and gel filtration chromatography.

VLPs produced by cells containing the synthetic expression cassettes of the present invention can be used to elicit an immune response when administered to a subject. One advantage of the present invention is that VLPs can be produced by mammalian cells carrying the synthetic expression cassettes at levels previously not possible. As discussed above, the VLPs can comprise a variety of antigens in addition to the Gag polypeptide (e.g., Gag-protease, Gag-polymerase, Env, synthetic Env, etc.). Purified VLPs, produced using the synthetic expression cassettes of the present invention, can be administered to a vertebrate subject, usually in the form of vaccine compositions. Combination vaccines may also be used, where such vaccines contain, for example, an adjuvant subunit protein (e.g., Env). Administration can take place using the VLPs formulated alone or formulated with other antigens. Further, the VLPs can be administered prior to, concurrent with, or subsequent to, delivery of the synthetic expression cassettes for DNA immunization (see below) and/or delivery of other vaccines. Also, the site of VLP administration may be the same or different as other vaccine compositions that are being administered. Gene delivery can be accomplished by a number of methods including, but are not limited to, immunization with DNA, alphavirus vectors, pox virus vectors, and vaccinia virus vectors.

VLP immuno-stimulating (or vaccine) compositions can include various excipients, adjuvants, carriers, auxiliary substances, modulating agents, and the like. The immune stimulating compositions will include an amount of the VLP/antigen sufficient to mount an immunological response. An appropriate effective amount can be determined by one of skill in the art. Such an amount will fall in a relatively broad range that can be determined through routine trials and will generally be an amount on the order of about 0.1 μ g to about 1000 μ g, more preferably about 1 μ g to about 300 μ g, of VLP/antigen.

A carrier is optionally present which is a molecule that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or liposomes), and inactive virus particles. Examples of particulate carriers include those derived from polymethyl methacrylate polymers, as well as microparticles derived from poly(lactides) and poly(lactide-co-

WO 02/04493

PCT/US01/21241

glycolides), known as PLG. See, e.g., Jeffery et al., *Pharm. Res.* (1993) 10:362-368; McGee JP, et al., *J Microencapsul.* 14(2):197-210, 1997; O'Hagan DT, et al., *Vaccine* 11(2):149-54, 1993. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the antigen may be conjugated to a bacterial toxoid, such as toxoid from diphtheria, tetanus, cholera, etc., as well as toxins derived from *E. coli*.

Adjuvants may also be used to enhance the effectiveness of the compositions. Such adjuvants include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc.; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (International Publication No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalene, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) RibiTM adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphoryl lipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (DetoxTM); (3) saponin adjuvants, such as StimulonTM (Cambridge Bioscience, Worcester, MA) may be used or particle generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (IL-1, IL-2, etc.), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc.; (6) oligonucleotides or polymeric molecules encoding immunostimulatory CpG motifs (Davis, H.L., et al., *J. Immunology* 160:870-876, 1998; Sato, Y. et al., *Science* 273:352-354, 1996) or complexes of antigens/oligonucleotides (Polymeric molecules include double and single stranded RNA and DNA, and backbone modifications thereof, for example, methylphosphonate linkages; or (7) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63 (where lysine is substituted for the wild-type amino acid at position 63) LT-R72 (where arginine is substituted for the wild-type amino acid at position 72), CT-S109 (where serine is substituted for the wild-type amino acid at position 109), and PT-K9/G129 (where lysine is substituted for the wild-type amino acid at position 9 and glycine substituted

WO 02/04493

PCT/US01/21241

at position 129) (see, e.g., International Publication Nos. W093/13202 and W092/19265); and (8) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Further, such polymeric molecules include alternative polymer backbone structures such as, but not limited to, polyvinyl backbones (Pitha, *Biochem Biophys Acta*, 204:39, 1970a; Pitha, *Biopolymers*, 2:965, 1970b), and morpholine backbones (Summerton, J., *et al.*, U.S. Patent No. 5,142,047, issued 08/25/92; Summerton, J., *et al.*, U.S. Patent No. 5,185,444 issued 02/09/93). A variety of other charged and uncharged polynucleotide analogs have been reported. Numerous backbone modifications are known in the art, including, but not limited to, uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, and carbamates) and charged linkages (e.g., phosphorothioates and phosphorodithioates); and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the VLP immune-stimulating (or vaccine) composition. Alum, CpG oligonucleotides, and MF59 are preferred.

Muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamate (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

Dosage treatment with the VLP composition may be a single dose schedule or a multiple dose schedule. A multiple dose schedule is one in which a primary course of vaccination may be with 1-10 separate doses, followed by other doses given at subsequent time intervals, chosen to maintain and/or reinforce the immune response, for example at 1-4 months for a second dose, and if needed, a subsequent dose(s) after several months. The dosage regimen will also, at least in part, be determined by the need of the subject and be dependent on the judgment of the practitioner.

If prevention of disease is desired, the antigen carrying VLPs are generally administered prior to primary infection with the pathogen of interest. If treatment is desired, e.g., the reduction of symptoms or recurrences, the VLP compositions are generally administered subsequent to primary infection.

2.3.2 USING THE SYNTHETIC EXPRESSION CASSETTES OF THE PRESENT INVENTION TO CREATE PACKAGING CELL LINES

A number of viral based systems have been developed for use as gene transfer vectors for mammalian host cells. For example, retroviruses (in particular, lentiviral vectors) provide a convenient platform for gene delivery systems. A coding sequence of interest (for example, a sequence useful for gene therapy applications) can be inserted into a gene delivery vector

WO 02/04493

PCT/US01/21241

- and packaged in retroviral particles using techniques known in the art. Recombinant virus can then be isolated and delivered to cells of the subject either *in vivo* or *ex vivo*. A number of retroviral systems have been described, including, for example, the following: (U.S. Patent No. 5,219,740; Miller et al. (1989) *BioTechniques* 7:980; Miller, A.D. (1990) *Human Gene Therapy* 1:5; Scarpa et al. (1991) *Virology* 180:849; Burns et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:8033; Boris-Lawrie et al. (1993) *Cur. Opin. Genet. Develop.* 3:102; GB 2200651; EP 0415731; EP 0345242; WO 89/02468; WO 89/05349; WO 89/09271; WO 90/02806; WO 90/07936; WO 90/07936; WO 94/03622; WO 93/25698; WO 93/25234; WO 93/11230; WO 93/10218; WO 91/02805; in U.S. 5,219,740; U.S. 4,405,712; U.S. 4,861,719; U.S. 4,980,289 and U.S. 4,777,127; in U.S. Serial No. 07/800,921; and in Vile (1993) *Cancer Res* 53:3860-3864; Vile (1993) *Cancer Res* 53:962-967; Ram (1993) *Cancer Res* 53:83-88; Takamiya (1992) *J Neurosci Res* 33:493-503; Baba (1993) *J Neurosurg* 79:729-735; Mann (1983) *Cell* 33:153; Cane (1984) *Proc Natl Acad Sci USA* 81:6349; and Miller (1990) *Human Gene Therapy* 1.
- 15 In other embodiments, gene transfer vectors can be constructed to encode a cytokine or other immunomodulatory molecule. For example, nucleic acid sequences encoding native IL-2 and gamma-interferon can be obtained as described in US Patent Nos. 4,738,927 and 5,326,859, respectively, while useful mutants of these proteins can be obtained as described in U.S. Patent No. 4,853,332. Nucleic acid sequences encoding the short and long forms of mCSF can be obtained as described in US Patent Nos. 4,847,201 and 4,879,227, respectively.
- 20 In particular aspects of the invention, retroviral vectors expressing cytokine or immunomodulatory genes can be produced as described herein (for example, employing the packaging cell lines of the present invention) and in International Application No. PCT US 94/02951, entitled "Compositions and Methods for Cancer Immunotherapy."
- 25 Examples of suitable immunomodulatory molecules for use herein include the following: IL-1 and IL-2 (Karupiah et al. (1990) *J. Immunology* 144:290-298, Weber et al. (1987) *J. Exp. Med.* 166:1716-1733, Gausbacher et al. (1990) *J. Exp. Med.* 172:1217-1224, and U.S. Patent No. 4,738,927); IL-3 and IL-4 (Tepper et al. (1989) *Cell* 57:503-512, Golumbok et al. (1991) *Science* 254:713-716, and U.S. Patent No. 5,017,691); IL-5 and IL-6 (Brakenhof et al. (1987) *J. Immunol.* 139:4116-4121, and International Publication No. WO 90/06370); IL-7 (U.S. Patent No. 4,965,195); IL-8, IL-9, IL-10, IL-11, IL-12, and IL-13 (*Cytokine Bulletin*, Summer 1994); IL-14 and IL-15; alpha interferon (Finter et al. (1991) *Drugs* 42:749-765, U.S. Patent Nos. 4,892,743 and 4,966,843, International Publication No. WO 85/02862, Nagata et al. (1980) *Nature* 284:316-320, Familletti et al. (1981) *Methods in*
- 35 *Enz.* 78:387-394, Twu et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:2046-2050, and Faktor et

WO 02/04493

PCT/US01/21241

al. (1990) *Oncogene* 5:867-872; beta-interferon (Seif et al. (1991) *J. Virol.* 65:664-671); gamma-interferons (Radford et al. (1991) *The American Society of Hepatology* 20082015, Watanabe et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:9456-9460, Gansbacher et al. (1990) *Cancer Research* 50:7820-7825, Maio et al. (1989) *Can. Immunol. Immunother.* 30:34-42, and U.S. Patent Nos. 4,762,791 and 4,727,138; G-CSF (U.S. Patent Nos. 4,999,291 and 4,810,643); GM-CSF (International Publication No. WO 85/04188).

Immunomodulatory factors may also be agonists, antagonists, or ligands for these molecules. For example, soluble forms of receptors can often behave as antagonists for these types of factors, as can mutated forms of the factors themselves.

Nucleic acid molecules that encode the above-described substances, as well as other nucleic acid molecules that are advantageous for use within the present invention, may be readily obtained from a variety of sources, including, for example, depositories such as the American Type Culture Collection, or from commercial sources such as British Biotechnology Limited (Cowley, Oxford England). Representative examples include BBG 12 (containing the GM-CSF gene coding for the mature protein of 127 amino acids), BBG 6 (which contains sequences encoding gamma interferon), A.T.C.C. Deposit No. 39656 (which contains sequences encoding TNF), A.T.C.C. Deposit No. 20663 (which contains sequences encoding alpha-interferon), A.T.C.C. Deposit Nos. 31902, 31902 and 39517 (which contain sequences encoding beta-interferon), A.T.C.C. Deposit No. 67024 (which contains a sequence which encodes Interleukin-1b), A.T.C.C. Deposit Nos. 39405, 39452, 39516, 39626 and 39673 (which contain sequences encoding Interleukin-2), A.T.C.C. Deposit Nos. 59399, 59398, and 67326 (which contain sequences encoding Interleukin-3), A.T.C.C. Deposit No. 57592 (which contains sequences encoding Interleukin-4), A.T.C.C. Deposit Nos. 59394 and 59395 (which contain sequences encoding Interleukin-5), and A.T.C.C. Deposit No. 67153 (which contains sequences encoding Interleukin-6).

Plasmids containing cytokine genes or immunomodulatory genes (International Publication Nos. WO 94/02951 and WO 96/21015) can be digested with appropriate restriction enzymes, and DNA fragments containing the particular gene of interest can be inserted into a gene transfer vector using standard molecular biology techniques. (See, e.g., Sambrook et al., *supra.*, or Ausbel et al. (eds) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience).

Polynucleotide sequences coding for the above-described molecules can be obtained using recombinant methods, such as by screening cDNA and genomic libraries from cells expressing the gene, or by deriving the gene from a vector known to include the same. For example, plasmids which contain sequences that encode altered cellular products may be

WO 02/04493

PCT/US01/21241

obtained from a depository such as the A.T.C.C., or from commercial sources. Plasmids containing the nucleotide sequences of interest can be digested with appropriate restriction enzymes, and DNA fragments containing the nucleotide sequences can be inserted into a gene transfer vector using standard molecular biology techniques.

5 Alternatively, cDNA sequences for use with the present invention may be obtained from cells which express or contain the sequences, using standard techniques, such as phenol extraction and PCR of cDNA or genomic DNA. See, e.g., Sambrook et al., *supra*, for a description of techniques used to obtain and isolate DNA. Briefly, mRNA from a cell which expresses the gene of interest can be reverse transcribed with reverse transcriptase using
10 oligo-dT or random primers. The single stranded cDNA may then be amplified by PCR (see U.S. Patent Nos. 4,683,202, 4,683,195 and 4,800,159, see also *PCR Technology: Principles and Applications for DNA Amplification*, Erlich (ed.), Stockton Press, 1989)) using oligonucleotide primers complementary to sequences on either side of desired sequences.

The nucleotide sequence of interest can also be produced synthetically, rather than
15 cloned, using a DNA synthesizer (e.g., an Applied Biosystems Model 392 DNA Synthesizer, available from ABI, Foster City, California). The nucleotide sequence can be designed with the appropriate codons for the expression product desired. The complete sequence is assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge (1981) *Nature* 292:756; Nambair et al.
20 (1984) *Science* 223:1299; Jay et al. (1984) *J. Biol. Chem.* 259:6311.

The synthetic expression cassettes of the present invention can be employed in the construction of packaging cell lines for use with retroviral vectors.

One type of retrovirus, the murine leukemia virus, or "MLV", has been widely utilized for gene therapy applications (see generally Mann et al. (*Cell* 33:153, 1993), Cane
25 and Mulligan (*Proc. Nat'l. Acad. Sci. USA* 81:6349, 1984), and Miller et al., *Human Gene Therapy* 1:5-14, 1990).

Lentiviral vectors typically, comprise a 5' lentiviral LTR, a tRNA binding site, a packaging signal, a promoter operably linked to one or more genes of interest, an origin of second strand DNA synthesis and a 3' lentiviral LTR, wherein the lentiviral vector contains a
30 nuclear transport element. The nuclear transport element may be located either upstream (5') or downstream (3') of a coding sequence of interest (for example, a synthetic Gag or Env expression cassette of the present invention). Within certain embodiments, the nuclear transport element is not RRE. Within one embodiment the packaging signal is an extended packaging signal. Within other embodiments the promoter is a tissue specific promoter, or,

WO 02/04493

PCT/US01/21241

alternatively, a promoter such as CMV. Within other embodiments, the lentiviral vector further comprises an internal ribosome entry site.

A wide variety of lentiviruses may be utilized within the context of the present invention, including for example, lentiviruses selected from the group consisting of HIV, HIV-1, HIV-2, FIV and SIV.

In one embodiment of the present invention synthetic Gag-polymerase expression cassettes are provided comprising a promoter and a sequence encoding synthetic Gag-polymerase and at least one of vpr, vpu, nef or vif, wherein the promoter is operably linked to Gag-polymerase and vpr, vpu, nef or vif.

Within yet another aspect of the invention, host cells (e.g., packaging cell lines) are provided which contain any of the expression cassettes described herein. For example, within one aspect packaging cell line are provided comprising an expression cassette that comprises a sequence encoding synthetic Gag-polymerase, and a nuclear transport element, wherein the promoter is operably linked to the sequence encoding Gag-polymerase. Packaging cell lines may further comprise a promoter and a sequence encoding tat, rev, or an envelope, wherein the promoter is operably linked to the sequence encoding tat, rev, Env or sequences encoding modified versions of these proteins. The packaging cell line may further comprise a sequence encoding any one or more of nef, vif, vpu or vpr (wild-type or synthetic).

In one embodiment, the expression cassette (carrying, for example, the synthetic Gag-polymerase) is stably integrated. The packaging cell line, upon introduction of a lentiviral vector, typically produces particles. The promoter regulating expression of the synthetic expression cassette may be inducible. Typically, the packaging cell line, upon introduction of a lentiviral vector, produces particles that are essentially free of replication competent virus.

Packaging cell lines are provided comprising an expression cassette which directs the expression of a synthetic *Gag-polymerase* gene or comprising an expression cassette which directs the expression of a synthetic Env genes described herein. (See, also, Andre, S., et al., *Journal of Virology* 72(2):1497-1503, 1998; Haas, J., et al., *Current Biology* 6(3):315-324, 1996) for a description of other modified Env sequences). A lentiviral vector is introduced into the packaging cell line to produce a vector producing cell line.

As noted above, lentiviral vectors can be designed to carry or express a selected gene(s) or sequences of interest. Lentiviral vectors may be readily constructed from a wide variety of lentiviruses (see RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985). Representative examples of lentiviruses included HTV, HIV-1, HIV-2, FIV and SIV. Such lentiviruses may either be obtained from patient isolates, or, more

WO 02/04493

PCT/US01/21241

preferably, from depositories or collections such as the American Type Culture Collection, or isolated from known sources using available techniques.

Portions of the lentiviral gene delivery vectors (or vehicles) may be derived from different viruses. For example, in a given recombinant lentiviral vector, LTRs may be derived from an HIV, a packaging signal from SIV, and an origin of second strand synthesis from Hrv-2. Lentiviral vector constructs may comprise a 5' lentiviral LTR, a tRNA binding site, a packaging signal, one or more heterologous sequences, an origin of second strand DNA synthesis and a 3' LTR, wherein said lentiviral vector contains a nuclear transport element that is not RRE.

Briefly, Long Terminal Repeats ("LTRs") are subdivided into three elements, designated U5, R and U3. These elements contain a variety of signals which are responsible for the biological activity of a retrovirus, including for example, promoter and enhancer elements which are located within U3. LTRs may be readily identified in the provirus (integrated DNA form) due to their precise duplication at either end of the genome. As utilized herein, a 5' LTR should be understood to include a 5' promoter element and sufficient LTR sequence to allow reverse transcription and integration of the DNA form of the vector. The 3' LTR should be understood to include a polyadenylation signal, and sufficient LTR sequence to allow reverse transcription and integration of the DNA form of the vector.

The tRNA binding site and origin of second strand DNA synthesis are also important for a retrovirus to be biologically active, and may be readily identified by one of skill in the art. For example, retroviral tRNA binds to a tRNA binding site by Watson-Crick base pairing, and is carried with the retrovirus genome into a viral particle. The tRNA is then utilized as a primer for DNA synthesis by reverse transcriptase. The tRNA binding site may be readily identified based upon its location just downstream from the 5'LTR. Similarly, the origin of second strand DNA synthesis is, as its name implies, important for the second strand DNA synthesis of a retrovirus. This region, which is also referred to as the poly-purine tract, is located just upstream of the 3'LTR.

In addition to a 5' and 3' LTR, tRNA binding site, and origin of second strand DNA synthesis, recombinant retroviral vector constructs may also comprise a packaging signal, as well as one or more genes or coding sequences of interest. In addition, the lentiviral vectors have a nuclear transport element which, in preferred embodiments is not RRE. Representative examples of suitable nuclear transport elements include the element in Rous sarcoma virus (Ogert, et al., *J Virol* 70, 3834-3843, 1996), the element in Rous sarcoma virus (Liu & Mertz, *Genes & Dev.* 9, 1766-1789, 1995) and the element in the genome of simian retrovirus type I (Zolotukhin, et al., *J Virol.* 68, 7944-7952, 1994). Other potential elements

WO 02/04493

PCT/US01/21241

include the elements in the histone gene (Kedes, *Annu. Rev. Biochem.* 48, 837-870, 1970), the α -interferon gene (Nagata et al., *Nature* 287, 401-408, 1980), the β -adrenergic receptor gene (Koilka, et al., *Nature* 329, 75-79, 1987), and the c-Jun gene (Hattori, et al., *Proc. Natl. Acad. Sci. USA* 85, 9148-9152, 1988).

5 Recombinant lentiviral vector constructs typically lack both *Gag-polymerase* and *Env* coding sequences. Recombinant lentiviral vector typically contain less than 20, preferably 15, more preferably 10, and most preferably 8 consecutive nucleotides found in *Gag-polymerase* and *Env* genes. One advantage of the present invention is that the synthetic *Gag-polymerase* expression cassettes, which can be used to construct packaging cell lines for the recombina-
10 nt retroviral vector constructs, have little homology to wild-type *Gag-polymerase* sequences and thus considerably reduce or eliminate the possibility of homologous recombination between the synthetic and wild-type sequences.

Lentiviral vectors may also include tissue-specific promoters to drive expression of one or more genes or sequences of interest.

15 Lentiviral vector constructs may be generated such that more than one gene of interest is expressed. This may be accomplished through the use of di- or oligo-cistronic cassettes (e.g., where the coding regions are separated by 80 nucleotides or less, see generally Levin et al., *Gene* 108:167-174, 1991), or through the use of Internal Ribosome Entry Sites ("IRES").

Packaging cell lines suitable for use with the above described recombinant retroviral
20 vector constructs may be readily prepared given the disclosure provided herein. Briefly, the parent cell line from which the packaging cell line is derived can be selected from a variety of mammalian cell lines, including for example, 293, RD, COS-7, CHO, BHK, VERO, HT1080, and myeloma cells.

After selection of a suitable host cell for the generation of a packaging cell line, one or
25 more expression cassettes are introduced into the cell line in order to complement or supply *in trans* components of the vector which have been deleted.

Representative examples of suitable expression cassettes have been described herein and include synthetic *Env*, synthetic *Gag*, synthetic *Gag-protease*, and synthetic *Gag-polymerase* expression cassettes, which comprise a promoter and a sequence encoding, e.g.,
30 *Gag-polymerase* and at least one of *vpr*, *vpu*, *nef* or *vif*, wherein the promoter is operably linked to *Gag-polymerase* and *vpr*, *vpu*, *nef* or *vif*. As described above, the native and/or synthetic coding sequences may also be utilized in these expression cassettes.

Utilizing the above-described expression cassettes, a wide variety of packaging cell lines can be generated. For example, within one aspect packaging cell line are provided
35 comprising an expression cassette that comprises a sequence encoding synthetic *Gag*-

WO 02/04493

PCT/US01/21241

polymerase, and a nuclear transport element, wherein the promoter is operably linked to the sequence encoding Gag-polymerase. Within other aspects, packaging cell lines are provided comprising a promoter and a sequence encoding tat, rev, Env, or other HIV antigens or epitopes derived therefrom, wherein the promoter is operably linked to the sequence encoding tat, rev, Env, or the HIV antigen or epitope. Within further embodiments, the packaging cell line may comprise a sequence encoding any one or more of nef, vif, vpr or vpr. For example, the packaging cell line may contain only nef, vif, vpr, or vpr alone, nef and vif, nef and vpr, nef and vpr, vif and vpr, vif and vpr, vpr and vpr, nef vif and vpr, nef vif and vpr, nef vpr and vpr, vpr and vpr, or, all four of nef, vif, vpr, and vpr.

In one embodiment, the expression cassette is stably integrated. Within another embodiment, the packaging cell line, upon introduction of a lentiviral vector, produces particles. Within further embodiments the promoter is inducible. Within certain preferred embodiments of the invention, the packaging cell line, upon introduction of a lentiviral vector, produces particles that are free of replication competent virus.

The synthetic cassettes containing modified coding sequences are transfected into a selected cell line. Transfected cells are selected that (i) carry, typically, integrated, stable copies of the HIV coding sequences, and (ii) are expressing acceptable levels of these polypeptides (expression can be evaluated by methods known in the prior art, e.g., see Examples 1-4). The ability of the cell line to produce VLPs may also be verified.

A sequence of interest is constructed into a suitable viral vector as discussed above. This defective virus is then transfected into the packaging cell line. The packaging cell line provides the viral functions necessary for producing virus-like particles into which the defective viral genome, containing the sequence of interest, are packaged. These VLPs are then isolated and can be used, for example, in gene delivery or gene therapy.

Further, such packaging cell lines can also be used to produce VLPs alone, which can, for example, be used as adjuvants for administration with other antigens or in vaccine compositions. Also, co-expression of a selected sequence of interest encoding a polypeptide (for example, an antigen) in the packaging cell line can also result in the entrapment and/or association of the selected polypeptide in/with the VLPs.

Various forms of the different embodiments of the present invention (e.g., constructs) may be combined.

2.4 DNA IMMUNIZATION AND GENE DELIVERY

A variety of HIV polypeptide antigens, particularly Type C HIV antigens, can be used in the practice of the present invention. HIV antigens can be included in DNA immunization

WO 02/04493

PCT/US01/21241

constructs containing, for example, a synthetic Gag expression cassette fused in-frame to a coding sequence for the polypeptide antigen (synthetic or wild-type), where expression of the construct results in VLPs presenting the antigen of interest.

HIV antigens of particular interest to be used in the practice of the present invention include tat, rev, nef, vif, vpr, vpu, and other HIV antigens or epitopes derived therefrom. These antigens may be synthetic (as described herein) or wild-type. Further, the packaging cell line may contain only nef, and HIV-1 (also known as HTLV-III, LAV, ARV, etc.), including, but not limited to, antigens such as gp120, gp41, gp160 (both native and modified), Gag, and pol from a variety of isolates including, but not limited to, HIV_{BR}, HIV_{ST2}, HIV-1_{EF162}, HIV-1_{SP170}, HIV_{LAV}, HIV_{LAI}, HIV_{M89}, HIV-1_{OM235}, HIV-1_{VE4}, other HIV-1 strains from diverse subtypes (e.g., subtypes A through G, and O), HIV-2 strains and diverse subtypes (e.g., HIV-2_{UC1} and HIV-2_{UC2}). See, e.g., Myers, et al., Los Alamos Database, Los Alamos National Laboratory, Los Alamos, New Mexico; Myers, et al., *Human Retroviruses and Aids*, 1990, Los Alamos, New Mexico: Los Alamos National Laboratory.

To evaluate efficacy, DNA immunization using synthetic expression cassettes of the present invention can be performed, for instance as described in Example 4. Mice are immunized with both the Gag (and/or Env) synthetic expression cassette and the Gag (and/or Env) wild type expression cassette. Mouse immunizations with plasmid-DNAs will show that the synthetic expression cassettes provide a clear improvement of immunogenicity relative to the native expression cassettes. Also, the second boost immunization will induce a secondary immune response, for example, after approximately two weeks. Further, the results of CTL assays will show increased potency of synthetic Gag (and/or Env) expression cassettes for induction of cytotoxic T-lymphocyte (CTL) responses by DNA immunization.

It is readily apparent that the subject invention can be used to mount an immune response to a wide variety of antigens and hence to treat or prevent a HIV infection, particularly Type C HIV infection.

2.4.1 DELIVERY OF THE SYNTHETIC EXPRESSION CASSETTES OF THE PRESENT INVENTION

Polynucleotide sequences coding for the above-described molecules can be obtained using recombinant methods, such as by screening cDNA and genomic libraries from cells expressing the gene, or by deriving the gene from a vector known to include the same. Furthermore, the desired gene can be isolated directly from cells and tissues containing the same, using standard techniques, such as phenol extraction and PCR of cDNA or genomic DNA. See, e.g., Sambrook et al., *supra*, for a description of techniques used to obtain and

WO 02/04493

PCT/US01/21241

isolate DNA. The gene of interest can also be produced synthetically, rather than cloned. The nucleotide sequence can be designed with the appropriate codons for the particular amino acid sequence desired. In general, one will select preferred codons for the intended host in which the sequence will be expressed. The complete sequence is assembled from overlapping oligonucleotides prepared by standard methods and assembled into a complete coding sequence. See, e.g., Edge, *Nature* (1981) 292:756; Nambair et al., *Science* (1984) 223:1299; Jay et al., *J. Biol. Chem.* (1984) 259:6311; Stemmer, W.P.C., (1995) *Gene* 164:49-53.

Next, the gene sequence encoding the desired antigen can be inserted into a vector containing a synthetic expression cassette of the present invention. In certain embodiments, the antigen is inserted into the synthetic Gag coding sequence such that when the combined sequence is expressed it results in the production of VLPs comprising the Gag polypeptide and the antigen of interest, e.g., Env (native or modified) or other antigen(s) (native or modified) derived from HIV. Insertions can be made within the coding sequence or at either end of the coding sequence (5' amino terminus of the expressed Gag polypeptide; or 3' carboxy terminus of the expressed Gag polypeptide) (Wagner, R., et al., *Arch Virol.* 127:117-137, 1992; Wagner, R., et al., *Virology* 200:162-175, 1994; Wu, X., et al., *J. Virol.* 69(6):3389-3398, 1995; Wang, C.-T., et al., *Virology* 200:524-534, 1994; Chazal, N., et al., *Virology* 68(1):111-122, 1994; Griffiths, J.C., et al., *J. Virol.* 67(6):3191-3198, 1993; Reicin, A.S., et al., *J. Virol.* 69(2):642-650, 1995).

Up to 50% of the coding sequences of p55Gag can be deleted without affecting the assembly to virus-like particles and expression efficiency (Borsetti, A., et al., *J. Virol.* 72(11):9313-9317, 1998; Gamier, L., et al., *J. Virol.* 72(6):4667-4677, 1998; Zhang, Y., et al., *J. Virol.* 72(3):1782-1789, 1998; Wang, C., et al., *J. Virol.* 72(10):7950-7959, 1998). In one embodiment of the present invention, immunogenicity of the high level expressing synthetic Gag expression cassettes can be increased by the insertion of different structural or non-structural HIV antigens, multi-epitope cassettes, or cytokine sequences into deleted regions of Gag sequence. Such deletions may be generated following the teachings of the present invention and information available to one of ordinary skill in the art. One possible advantage of this approach, relative to using full-length sequences fused to heterologous polypeptides, can be higher expression/secretion efficiency of the expression product.

When sequences are added to the amino terminal end of Gag, the polynucleotide can contain coding sequences at the 5' end that encode a signal for addition of a myristic moiety to the Gag-containing polypeptide (e.g., sequences that encode Met-Gly).

WO 02/04493

PCT/US01/21241

The ability of Gag-containing polypeptide constructs to form VLPs can be empirically determined following the teachings of the present specification.

The synthetic expression cassettes can also include control elements operably linked to the coding sequence, which allow for the expression of the gene *in vivo* in the subject species. For example, typical promoters for mammalian cell expression include the SV40 early promoter, a CMV promoter such as the CMV immediate early promoter, the mouse mammary tumor virus LTR promoter, the adenovirus major late promoter (Ad MLP), and the herpes simplex virus promoter, among others. Other nonviral promoters, such as a promoter derived from the murine metallothionein gene, will also find use for mammalian expression. Typically, transcription termination and polyadenylation sequences will also be present, located 3' to the translation stop codon. Preferably, a sequence for optimization of initiation of translation, located 5' to the coding sequence, is also present. Examples of transcription terminator/polyadenylation signals include those derived from SV40, as described in Sambrook et al., *supra*, as well as a bovine growth hormone terminator sequence.

Enhancer elements may also be used herein to increase expression levels of the mammalian constructs. Examples include the SV40 early gene enhancer, as described in Dijkema et al., *EMBO J.* (1985) 4:761, the enhancer/promoter derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus, as described in Gorman et al., *Proc. Natl. Acad. Sci. USA* (1982b) 79:6777 and elements derived from human CMV, as described in Boshart et al., *Cell* (1985) 41:521, such as elements included in the CMV intron A sequence.

Furthermore, plasmids can be constructed which include a chimeric antigen-coding gene sequences, encoding, e.g., multiple antigens/epitopes of interest, for example derived from more than one viral isolate.

Typically the antigen coding sequences precede or follow the synthetic coding sequence and the chimeric transcription unit will have a single open reading frame encoding both the antigen of interest and the synthetic coding sequences. Alternatively, multi-cistronic cassettes (e.g., bi-cistronic cassettes) can be constructed allowing expression of multiple antigens from a single mRNA using the EMCV IRES, or the like.

Once complete, the constructs are used for nucleic acid immunization using standard gene delivery protocols. Methods for gene delivery are known in the art. See, e.g., U.S. Patent Nos. 5,399,346, 5,580,859, 5,589,466. Genes can be delivered either directly to the vertebrate subject or, alternatively, delivered *ex vivo*, to cells derived from the subject and the cells reimplanted in the subject.

A number of viral based systems have been developed for gene transfer into mammalian cells. For example, retroviruses provide a convenient platform for gene delivery

WO 02/04493

PCT/US01/21241

systems. Selected sequences can be inserted into a vector and packaged in retroviral particles using techniques known in the art. The recombinant virus can then be isolated and delivered to cells of the subject either *in vivo* or *ex vivo*. A number of retroviral systems have been described (U.S. Patent No. 5,219,740; Miller and Rosman, *BioTechniques* (1989) 7:980-990; Miller, A.D., *Human Gene Therapy* (1990) 1:5-14; Scapa et al., *Virology* (1991) 180:849-852; Burns et al., *Proc. Natl. Acad. Sci. USA* (1993) 90:8033-8037; and Boris-Lawrie and Temin, *Curr. Opin. Genet. Develop.* (1993) 3:102-109.

A number of adenovirus vectors have also been described. Unlike retroviruses which integrate into the host genome, adenoviruses persist extrachromosomally thus minimizing the risks associated with insertional mutagenesis (Haj-Ahmad and Graham, *J. Virol.* (1986) 57:267-274; Bett et al., *J. Virol.* (1993) 67:5911-5921; Mittereder et al., *Human Gene Therapy* (1994) 5:717-729; Seth et al., *J. Virol.* (1994) 68:933-940; Barr et al., *Gene Therapy* (1994) 1:51-58; Berkner, K.L. *BioTechniques* (1988) 6:616-629; and Rich et al., *Human Gene Therapy* (1993) 4:461-476).

Additionally, various adeno-associated virus (AAV) vector systems have been developed for gene delivery. AAV vectors can be readily constructed using techniques well known in the art. See, e.g., U.S. Patent Nos. 5,173,414 and 5,139,941; International Publication Nos. WO 92/01070 (published 23 January 1992) and WO 93/03769 (published 4 March 1993); Lebkowski et al., *Molec. Cell. Biol.* (1988) 8:3988-3996; Vincent et al., *Vaccines 90* (1990) (Cold Spring Harbor Laboratory Press); Carter, B.J. *Current Opinion in Biotechnology* (1992) 3:533-539; Muzyczka, N. *Current Topics in Microbiol. and Immunol.* (1992) 158:97-129; Kotin, R.M. *Human Gene Therapy* (1994) 5:793-801; Shelling and Smith, *Gene Therapy* (1994) 1:165-169; and Zhou et al., *J. Exp. Med.* (1994) 179:1867-1875.

Another vector system useful for delivering the polynucleotides of the present invention is the enterically administered recombinant poxvirus vaccines described by Small, Jr., P.A., et al. (U.S. Patent No. 5,676,950, issued October 14, 1997).

Additional viral vectors which will find use for delivering the nucleic acid molecules encoding the antigens of interest include those derived from the pox family of viruses, including vaccinia virus and avian poxvirus. By way of example, vaccinia virus recombinants expressing the genes can be constructed as follows. The DNA encoding the particular synthetic HIV subtype C polypeptide coding sequence is first inserted into an appropriate vector so that it is adjacent to a vaccinia promoter and flanking vaccinia DNA sequences, such as the sequence encoding thymidine kinase (TK). This vector is then used to transfect cells which are simultaneously infected with vaccinia. Homologous recombination serves to insert the vaccinia promoter plus the gene encoding the coding sequences of interest

WO 02/04493

PCT/US01/21241

into the viral genome. The resulting TK recombinant can be selected by culturing the cells in the presence of 5-bromodeoxyuridine and picking viral plaques resistant thereto.

Alternatively, avipoxviruses, such as the fowlpox and canarypox viruses, can also be used to deliver the genes. Recombinant avipox viruses, expressing immunogens from mammalian pathogens, are known to confer protective immunity when administered to non-avian species. The use of an avipox vector is particularly desirable in human and other mammalian species since members of the avipox genus can only productively replicate in susceptible avian species and therefore are not infective in mammalian cells. Methods for producing recombinant avipoxviruses are known in the art and employ genetic recombination, as described above with respect to the production of vaccinia viruses. See, e.g., WO 91/12882; WO 89/03429; and WO 92/03545.

Molecular conjugate vectors, such as the adenovirus chimeric vectors described in Michael et al., *J. Biol. Chem.* (1993) 268:6866-6869 and Wagner et al., *Proc. Natl. Acad. Sci. USA* (1992) 89:6099-6103, can also be used for gene delivery.

Members of the Alphavirus genus, such as, but not limited to, vectors derived from the Sindbis, Semliki Forest, and Venezuelan Equine Encephalitis viruses, will also find use as viral vectors for delivering the polynucleotides of the present invention (for example, a synthetic Gag-polypeptide encoding expression cassette). For a description of Sindbis-virus derived vectors useful for the practice of the instant methods, see, Dubensky et al., *J. Virol.* (1996) 70:508-519; and International Publication Nos. WO 95/07995 and WO 96/17072; as well as, Dubensky, Jr., T.W., et al., U.S. Patent No. 5,343,723, issued December 1, 1998, and Dubensky, Jr., T.W., U.S. Patent No. 5,789,245, issued August 4, 1998.

A vaccinia based infection/transfection system can be conveniently used to provide for inducible, transient expression of the coding sequences of interest in a host cell. In this system, cells are first infected *in vitro* with a vaccinia virus recombinant that encodes the bacteriophage T7 RNA polymerase. This polymerase displays exquisite specificity in that it only transcribes templates bearing T7 promoters. Following infection, cells are transfected with the polynucleotide of interest, driven by a T7 promoter. The polymerase expressed in the cytoplasm from the vaccinia virus recombinant transcribes the transfected DNA into RNA which is then translated into protein by the host translational machinery. The method provides for high level, transient, cytoplasmic production of large quantities of RNA and its translation products. See, e.g., Elroy-Stein and Moss, *Proc. Natl. Acad. Sci. USA* (1990) 87:6743-6747; Fuerst et al., *Proc. Natl. Acad. Sci. USA* (1986) 83:8122-8126.

As an alternative approach to infection with vaccinia or avipox virus recombinants, or to the delivery of genes using other viral vectors, an amplification system can be used that

WO 02/04493

PCT/US01/21241

will lead to high level expression following introduction into host cells. Specifically, a T7 RNA polymerase promoter preceding the coding region for T7 RNA polymerase can be engineered. Translation of RNA derived from this template will generate T7 RNA polymerase which in turn will transcribe more template. Concomitantly, there will be a cDNA whose expression is under the control of the T7 promoter. Thus, some of the T7 RNA polymerase generated from translation of the amplification template RNA will lead to transcription of the desired gene. Because some T7 RNA polymerase is required to initiate the amplification, T7 RNA polymerase can be introduced into cells along with the template(s) to prime the transcription reaction. The polymerase can be introduced as a protein or on a plasmid encoding the RNA polymerase. For a further discussion of T7 systems and their use for transforming cells, see, e.g., International Publication No. WO 94/26911; Studier and Moffatt, *J. Mol. Biol.* (1986) 189:113-130; Deng and Wolff, *Gene* (1994) 143:245-249; Gao et al., *Biochem. Biophys. Res. Commun.* (1994) 200:1201-1206; Gao and Huang, *Nuc. Acids Res.* (1993) 21:2867-2872; Chen et al., *Nuc. Acids Res.* (1994) 22:2114-2120; and U.S. Patent No. 5,135,855.

Synthetic expression cassettes of interest can also be delivered without a viral vector. For example, the synthetic expression cassette can be packaged in liposomes prior to delivery to the subject or to cells derived therefrom. Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed DNA to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight, *Biochim. Biophys. Acta.* (1991) 1097:1-17; Straubinger et al., in *Methods of Enzymology* (1983), Vol. 101, pp. 512-527.

Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations, with cationic liposomes particularly preferred. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner et al., *Proc. Natl. Acad. Sci. USA* (1987) 84:7413-7416); mRNA (Malone et al., *Proc. Natl. Acad. Sci. USA* (1989) 86:6077-6081); and purified transcription factors (Debs et al., *J. Biol. Chem.* (1990) 265:10189-10192), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-di(olcyl)oxy]propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner et al., *Proc. Natl. Acad. Sci. USA* (1987) 84:7413-7416). Other commercially available lipids include (DDAB/DOPE) and DOTAP/DOPE (Boehringer). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g., Szoka et al., *Proc. Natl.*

WO 02/04493

PCT/US01/21241

Acad. Sci. USA (1978) 75:4194-4198; PCT Publication No. WO 90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as, from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials.

5 Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

10 The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See, e.g., Straubinger et al., in METHODS OF IMMUNOLOGY (1983), Vol. 101, pp. 512-527; Szoka et al., *Proc. Natl. Acad. Sci. USA* (1978) 75:4194-4198; Papahadjopoulos et al., *Biochim. Biophys. Acta* (1975) 394:483; Wilson et al., *Cell* (1979) 17:77; Deamer and Bangham, *Biochim. Biophys. Acta* (1976) 443:629; Ostro et al., *Biochem. Biophys. Res. Commun.* (1977) 76:836; Fraley et al., *Proc. Natl. Acad. Sci. USA* (1979) 76:3348; Enoch and Strittmatter, *Proc. Natl. Acad. Sci. USA* (1979) 76:145; Fraley et al., *J. Biol. Chem.* (1980) 255:10431; Szoka and Papahadjopoulos, *Proc. Natl. Acad. Sci. USA* (1978) 75:145; and Schaefer-Ridder et al., *Science* (1982) 215:166.

The DNA and/or protein antigen(s) can also be delivered in cocholate lipid compositions similar to those described by Papahadjopoulos et al., *Biochem. Biophys. Acta* (1975) 394:483-491. See, also, U.S. Patent Nos. 4,663,161 and 4,871,488.

25 The synthetic expression cassette of interest may also be encapsulated, adsorbed to, or associated with, particulate carriers. Such carriers present multiple copies of a selected antigen to the immune system and promote trapping and retention of antigens in local lymph nodes. The particles can be phagocytosed by macrophages and can enhance antigen presentation through cytokine release. Examples of particulate carriers include those derived from polymethyl methacrylate polymers, as well as microparticles derived from poly(lactides) and poly(lactide-co-glycolides), known as PLG. See, e.g., Jeffery et al., *Pharm. Res.* (1993) 10:362-368; McGee JP, et al., *J. Microencapsul.* 14(2):197-210, 1997; O'Hagan DT, et al., *Vaccine* 11(2):149-54, 1993. Suitable microparticles may also be manufactured in the presence of charged detergents, such as anionic or cationic detergents, to yield microparticles with a surface having a net negative or a net positive charge. For
35 example, microparticles manufactured with anionic detergents, such as

WO 02/04493

PCT/US01/21241

hexadecyltrimethylammonium bromide (CTAB), i.e. CTAB-PLG microparticles, adsorb negatively charged macromolecules, such as DNA. (see, e.g., Int'l Application Number PCT/US99/17308).

Furthermore, other particulate systems and polymers can be used for the *in vivo* or *ex vivo* delivery of the gene of interest. For example, polymers such as polylysine, polyarginine, polyornithine, spermine, spermidine, as well as conjugates of these molecules, are useful for transferring a nucleic acid of interest. Similarly, DEAE dextran-mediated transfection, calcium phosphate precipitation or precipitation using other insoluble inorganic salts, such as strontium phosphate, aluminum silicates including bentonite and kaolin, chromic oxide, magnesium silicate, talc, and the like, will find use with the present methods. See, e.g., Felgner, P.L., *Advanced Drug Delivery Reviews* (1990) 5:163-187, for a review of delivery systems useful for gene transfer. Peptoids (Zuckerman, R.N., et al., U.S. Patent No. 5,831,005, issued November 3, 1998) may also be used for delivery of a construct of the present invention.

Additionally, biolistic delivery systems employing particulate carriers such as gold and tungsten, are especially useful for delivering synthetic expression cassettes of the present invention. The particles are coated with the synthetic expression cassette(s) to be delivered and accelerated to high velocity, generally under a reduced atmosphere, using a gun powder discharge from a "gene gun." For a description of such techniques, and apparatuses useful therefore, see, e.g., U.S. Patent Nos. 4,945,050; 5,036,006; 5,100,792; 5,179,022; 5,371,015; and 5,478,744. Also, needle-less injection systems can be used (Davis, H.L., et al., *Vaccine* 12:1503-1509, 1994; Bioject, Inc., Portland, OR).

Recombinant vectors carrying a synthetic expression cassette of the present invention are formulated into compositions for delivery to the vertebrate subject. These compositions may either be prophylactic (to prevent infection) or therapeutic (to treat disease after infection). The compositions will comprise a "therapeutically effective amount" of the gene of interest such that an amount of the antigen can be produced *in vivo* so that an immune response is generated in the individual to which it is administered. The exact amount necessary will vary depending on the subject being treated; the age and general condition of the subject to be treated; the capacity of the subject's immune system to synthesize antibodies; the degree of protection desired; the severity of the condition being treated; the particular antigen selected and its mode of administration, among other factors. An appropriate effective amount can be readily determined by one of skill in the art. Thus, a "therapeutically effective amount" will fall in a relatively broad range that can be determined through routine trials.

WO 02/04493

PCT/US01/21241

The compositions will generally include one or more "pharmaceutically acceptable excipients or vehicles" such as water, saline, glycerol, polyethyleneglycol, hyaluronic acid, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Certain facilitators of nucleic acid uptake and/or expression can also be included in the compositions or coadministered, such as, but not limited to, bupivacaine, cardiotoxin and sucrose.

Once formulated, the compositions of the invention can be administered directly to the subject (e.g., as described above) or, alternatively, delivered *ex vivo*, to cells derived from the subject, using methods such as those described above. For example, methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and can include, e.g., dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, lipofectamine and LT-1 mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) (with or without the corresponding antigen) in liposomes, and direct microinjection of the DNA into nuclei.

Direct delivery of synthetic expression cassette compositions *in vivo* will generally be accomplished with or without viral vectors, as described above, by injection using either a conventional syringe or a gene gun, such as the Accell® gene delivery system (PowderJect Technologies, Inc., Oxford, England). The constructs can be injected either subcutaneously, epidermally, intradermally, intramucosally such as nasally, rectally and vaginally, intraperitoneally, intravenously, orally or intramuscularly. Delivery of DNA into cells of the epidermis is particularly preferred as this mode of administration provides access to skin-associated lymphoid cells and provides for a transient presence of DNA in the recipient. Other modes of administration include oral and pulmonary administration, suppositories, needle-less injection, transcutaneous and transdermal applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. Administration of nucleic acids may also be combined with administration of peptides or other substances.

2.4.2 EX VIVO DELIVERY OF THE SYNTHETIC EXPRESSION CASSETTES OF THE PRESENT INVENTION

In one embodiment, T cells, and related cell types (including but not limited to antigen presenting cells, such as, macrophage, monocytes, lymphoid cells, dendritic cells, B-cells, T-cells, stem cells, and progenitor cells thereof), can be used for *ex vivo* delivery of the synthetic expression cassettes of the present invention. T cells can be isolated from peripheral blood lymphocytes (PBLs) by a variety of procedures known to those skilled in the art. For example, T cell populations can be "enriched" from a population of PBLs through

WO 02/04493

PCT/US01/21241

the removal of accessory and B cells. In particular, T cell enrichment can be accomplished by the elimination of non-T cells using anti-MHC class II monoclonal antibodies. Similarly, other antibodies can be used to deplete specific populations of non-T cells. For example, anti-Ig antibody molecules can be used to deplete B cells and anti-MacI antibody molecules can be used to deplete macrophages.

T cells can be further fractionated into a number of different subpopulations by techniques known to those skilled in the art. Two major subpopulations can be isolated based on their differential expression of the cell surface markers CD4 and CD8. For example, following the enrichment of T cells as described above, CD4⁺ cells can be enriched using antibodies specific for CD4 (see Coligan et al., *supra*). The antibodies may be coupled to a solid support such as magnetic beads. Conversely, CD8⁺ cells can be enriched through the use of antibodies specific for CD4 (to remove CD4⁺ cells), or can be isolated by the use of CD8 antibodies coupled to a solid support. CD4 lymphocytes from HIV-1 infected patients can be expanded *ex vivo*, before or after transduction as described by Wilson et. al. (1995) *J. Infect. Dis.* 172:38.

Following purification of T cells, a variety of methods of genetic modification known to those skilled in the art can be performed using non-viral or viral-based gene transfer vectors constructed as described herein. For example, one such approach involves transduction of the purified T cell population with vector-containing supernatant of cultures derived from vector producing cells. A second approach involves co-cultivation of an irradiated monolayer of vector-producing cells with the purified T cells. A third approach involves a similar co-cultivation approach; however, the purified T cells are pre-stimulated with various cytokines and cultured 48 hours prior to the co-cultivation with the irradiated vector producing cells. Pre-stimulation prior to such transduction increases effective gene transfer (Nolta et al. (1992) *Exp. Hematol.* 20:1065). Stimulation of these cultures to proliferate also provides increased cell populations for re-infusion into the patient. Subsequent to co-cultivation, T cells are collected from the vector producing cell monolayer, expanded, and frozen in liquid nitrogen.

Gene transfer vectors, containing one or more synthetic expression cassette of the present invention (associated with appropriate control elements for delivery to the isolated T cells) can be assembled using known methods.

Selectable markers can also be used in the construction of gene transfer vectors. For example, a marker can be used which imparts to a mammalian cell transduced with the gene transfer vector resistance to a cytotoxic agent. The cytotoxic agent can be, but is not limited to, neomycin, aminoglycoside, tetracycline, chloramphenicol, sulfonamide, actinomycin,

WO 02/04493

PCT/US01/21241

netropsin, distamycin A, anthracycline, or pyrazinamide. For example, neomycin phosphotransferase II imparts resistance to the neomycin analogue geneticin (G418).

The T cells can also be maintained in a medium containing at least one type of growth factor prior to being selected. A variety of growth factors are known in the art which sustain the growth of a particular cell type. Examples of such growth factors are cytokine mitogens such as rIL-2, IL-10, IL-12, and IL-15, which promote growth and activation of lymphocytes. Certain types of cells are stimulated by other growth factors such as hormones, including human chorionic gonadotropin (hCG) and human growth hormone. The selection of an appropriate growth factor for a particular cell population is readily accomplished by one of skill in the art.

For example, white blood cells such as differentiated progenitor and stem cells are stimulated by a variety of growth factors. More particularly, IL-3, IL-4, IL-5, IL-6, IL-9, GM-CSF, M-CSF, and G-CSF, produced by activated T_H and activated macrophages, stimulate myeloid stem cells, which then differentiate into pluripotent stem cells, granulocyte-monocyte progenitors, eosinophil progenitors, basophil progenitors, megakaryocytes, and erythroid progenitors. Differentiation is modulated by growth factors such as GM-CSF, IL-3, IL-6, IL-11, and EPO.

Pluripotent stem cells then differentiate into lymphoid stem cells, bone marrow stromal cells, T cell progenitors, B cell progenitors, thymocytes, T_H cells, T_C cells, and B cells. This differentiation is modulated by growth factors such as IL-3, IL-4, IL-6, IL-7, GM-CSF, M-CSF, G-CSF, IL-2, and IL-5.

Granulocyte-monocyte progenitors differentiate to monocytes, macrophages, and neutrophils. Such differentiation is modulated by the growth factors GM-CSF, M-CSF, and IL-8. Eosinophil progenitors differentiate into eosinophils. This process is modulated by GM-CSF and IL-5.

The differentiation of basophil progenitors into mast cells and basophils is modulated by GM-CSF, IL-4, and IL-9. Megakaryocytes produce platelets in response to GM-CSF, EPO, and IL-6. Erythroid progenitor cells differentiate into red blood cells in response to EPO.

Thus, during activation by the CD3-binding agent, T cells can also be contacted with a mitogen, for example a cytokine such as IL-2. In particularly preferred embodiments, the IL-2 is added to the population of T cells at a concentration of about 50 to 100 $\mu\text{g/ml}$. Activation with the CD3-binding agent can be carried out for 2 to 4 days.

Once suitably activated, the T cells are genetically modified by contacting the same with a suitable gene transfer vector under conditions that allow for transfection of the vectors

WO 02/04493

PCT/US01/21241

into the T cells. Genetic modification is carried out when the cell density of the T cell population is between about 0.1×10^6 and 5×10^6 , preferably between about 0.5×10^6 and 2×10^6 . A number of suitable viral and nonviral-based gene transfer vectors have been described for use herein.

5 After transduction, transduced cells are selected away from non-transduced cells using known techniques. For example, if the gene transfer vector used in the transduction includes a selectable marker which confers resistance to a cytotoxic agent, the cells can be contacted with the appropriate cytotoxic agent, whereby non-transduced cells can be negatively selected away from the transduced cells. If the selectable marker is a cell surface marker, the cells can
10 be contacted with a binding agent specific for the particular cell surface marker, whereby the transduced cells can be positively selected away from the population. The selection step can also entail fluorescence-activated cell sorting (FACS) techniques, such as where FACS is used to select cells from the population containing a particular surface marker, or the selection step can entail the use of magnetically responsive particles as retrievable supports for target cell capture and/or background removal.

15 More particularly, positive selection of the transduced cells can be performed using a FACS cell sorter (e.g. a FACS Vantage™ Cell Sorter, Becton Dickinson Immunocytometry Systems, San Jose, CA) to sort and collect transduced cells expressing a selectable cell surface marker. Following transduction, the cells are stained with fluorescent-labeled
20 antibody molecules directed against the particular cell surface marker. The amount of bound antibody on each cell can be measured by passing droplets containing the cells through the cell sorter. By imparting an electromagnetic charge to droplets containing the stained cells, the transduced cells can be separated from other cells. The positively selected cells are then harvested in sterile collection vessels. These cell sorting procedures are described in detail,
25 for example, in the FACS Vantage™ Training Manual, with particular reference to sections 3-11 to 3-28 and 10-1 to 10-17.

Positive selection of the transduced cells can also be performed using magnetic separation of cells based on expression of a particular cell surface marker. In such separation techniques, cells to be positively selected are first contacted with specific binding agent (e.g.,
30 an antibody or reagent that interacts specifically with the cell surface marker). The cells are then contacted with retrievable particles (e.g., magnetically responsive particles) which are coupled with a reagent that binds the specific binding agent (that has bound to the positive cells). The cell-binding agent-particle complex can then be physically separated from non-labeled cells, for example using a magnetic field. When using magnetically responsive
35 particles, the labeled cells can be retained in a container using a magnetic field while the

WO 02/04493

PCT/US01/21241

negative cells are removed. These and similar separation procedures are known to those of ordinary skill in the art.

Expression of the vector in the selected transduced cells can be assessed by a number of assays known to those skilled in the art. For example, Western blot or Northern analysis
5 can be employed depending on the nature of the inserted nucleotide sequence of interest. Once expression has been established and the transformed T cells have been tested for the presence of the selected synthetic expression cassette, they are ready for infusion into a patient via the peripheral blood stream.

The invention includes a kit for genetic modification of an *ex vivo* population of
10 primary mammalian cells. The kit typically contains a gene transfer vector coding for at least one selectable marker and at least one synthetic expression cassette contained in one or more containers, ancillary reagents or hardware, and instructions for use of the kit.

2.4.3 FURTHER DELIVERY REGIMES

Any of the polynucleotides (e.g., expression cassettes) or polypeptides described herein (delivered by any of the methods described above) can also be used in combination with other DNA delivery systems and/or protein delivery systems. Non-limiting examples include co-administration of these molecules, for example, in prime-boost methods where one or more molecules are delivered in a "priming" step and, subsequently, one or more
20 molecules are delivered in a "boosting" step. In certain embodiments, the delivery of one or more nucleic acid-containing compositions and is followed by delivery of one or more nucleic acid-containing compositions and/or one or more polypeptide-containing compositions (e.g., polypeptides comprising HIV antigens). In other embodiments, multiple nucleic acid "primes" (of the same or different nucleic acid molecules) can be followed by
25 multiple polypeptide "boosts" (of the same or different polypeptides). Other examples include multiple nucleic acid administrations and multiple polypeptide administrations.

In any method involving co-administration, the various compositions can be delivered in any order. Thus, in embodiments including delivery of multiple different compositions or molecules, the nucleic acids need not be all delivered before the
30 polypeptides. For example, the priming step may include delivery of one or more polypeptides and the boosting comprises delivery of one or more nucleic acids and/or one or more polypeptides. Multiple polypeptide administrations can be followed by multiple nucleic acid administrations or polypeptide and nucleic acid administrations can be performed in any order. In any of the embodiments described herein, the nucleic acid
35 molecules can encode all, some or none of the polypeptides. Thus, one or more of the nucleic

WO 02/04493

PCT/US01/21241

acid molecules (e.g., expression cassettes) described herein and/or one or more of the polypeptides described herein can be co-administered in any order and via any administration routes. Therefore, any combination of polynucleotides and/or polypeptides described herein can be used to generate elicit an immune reaction.

5 EXPERIMENTAL

Below are examples of specific embodiments for carrying out the present invention. The examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

Efforts have been made to ensure accuracy with respect to numbers used (e.g.,
10 amounts, temperatures, etc.), but some experimental error and deviation should, of course, be allowed for.

Example 1

Generation of Synthetic Expression Cassettes

15 A. Modification of HIV-1 *Env*, *Gag*, *Pol* Nucleic Acid Coding Sequences

The *Pol* coding sequences were selected from Type C strain AF110975. The *Gag* coding sequences were selected from the Type C strains AF110965 and AF110967. The *Env* coding sequences were selected from Type C strains AF110968 and AF110975. These sequences were manipulated to maximize expression of their gene products.

20 First, the HIV-1 codon usage pattern was modified so that the resulting nucleic acid coding sequence was comparable to codon usage found in highly expressed human genes. The HIV codon usage reflects a high content of the nucleotides A or T of the codon-triplet. The effect of the HIV-1 codon usage is a high AT content in the DNA sequence that results in a decreased translation ability and instability of the mRNA. In comparison, highly expressed
25 human codons prefer the nucleotides G or C. The coding sequences were modified to be comparable to codon usage found in highly expressed human genes.

Second, there are inhibitory (or instability) elements (INS) located within the coding sequences of the *Gag* and *Gag-protease* coding sequences (Schneider R, et al., *J Virol.* 71(7):4892-4903, 1997). RRE is a secondary RNA structure that interacts with the HIV
30 encoded Rev-protein to overcome the expression down-regulating effects of the INS. To overcome the post-transcriptional activating mechanisms of RRE and Rev, the instability elements are inactivated by introducing multiple point mutations that do not alter the reading frame of the encoded proteins. Figures 5 and 6 (SEQ ID Nos: 3, 4, 20 and 21) show the location of some remaining INS in synthetic sequences derived from strains AF110965 and
35 AF110967. The changes made to these sequences are boxed in the Figures. In Figures 5 and

WO 02/04493

PCT/US01/21241

6, the top line depicts a modified sequence of Gag polypeptides from the indicated strains. The nucleotide(s) appearing below the line in the boxed region(s) depicts changes made to further remove INS. Thus, when the changes indicated in the boxed regions are made, the resulting sequences correspond to the sequences depicted in Figures 1 and 2, respectively.

- 5 The synthetic coding sequences are assembled by methods known in the art, for example by companies such as the Midland Certified Reagent Company (Midland, Texas).
- In one embodiment of the invention, sequences encoding Pol-polypeptides are included with the synthetic Gag or Env sequences in order to increase the number of epitopes for virus-like particles expressed by the synthetic, modified Gag/Env expression cassette.
- 10 Because synthetic HIV-1 Pol expresses the functional enzymes reverse transcriptase (RT) and integrase (INT) (in addition to the structural proteins and protease), it may be helpful in some instances to inactivate RT and INT functions. Several deletions or mutations in the RT and INT coding regions can be made to achieve catalytic nonfunctional enzymes with respect to their RT and INT activity. {Jay. A. Levy (Editor) (1995) *The Retroviridae*, Plenum Press, New York. ISBN 0-306-45033X. Pages 215-20; Grimson, B. and Laurence, J. (1995), *Journal Of Acquired Immune Deficiency Syndromes and Human Retrovirology* 9(1):58-68; Wakefield, J. K., et al., (1992) *Journal Of Virology* 66(11):6806-6812; Esnouf, R., et al., (1995) *Nature Structural Biology* 2(4):303-308; Maignan, S., et al., (1998) *Journal Of Molecular Biology* 282(2):359-368; Katz, R. A. and Skalka, A. M. (1994) *Annual Review Of Biochemistry* 73 (1994); Jacobo-Molina, A., et al., (1993) *Proceedings Of the National Academy Of Sciences Of the United States Of America* 90(13):6320-6324; Hickman, A. B., et al., (1994) *Journal Of Biological Chemistry* 269(46):29279-29287; Goldgur, Y., et al., (1998) *Proceedings Of the National Academy Of Sciences Of the United States Of America* 95(16):9150-9154; Goette, M., et al., (1998) *Journal Of Biological Chemistry* 273(17):10139-10146; Gorton, J. L., et al., (1998) *Journal of Virology* 72(6):5046-5055; Engelman, A., et al., (1997) *Journal Of Virology* 71(5):3507-3514; Dyda, F., et al., *Science* 266(5193):1981-1986; Davies, J. F., et al., (1991) *Science* 252(5002):88-95; Bujacz, G., et al., (1996) *Febs Letters* 398(2-3):175-178; Beard, W. A., et al., (1996) *Journal Of Biological Chemistry* 271(21):12213-12220; Kohlstaedt, L. A., et al., (1992) *Science* 256(5065):1783-1790; Krug, M. S. and Berger, S. I., (1991) *Biochemistry* 30(44):10614-10623; Mazumder, A., et al., (1996) *Molecular Pharmacology* 49(4):621-628; Palaniappan, C., et al., (1997) *Journal Of Biological Chemistry* 272(17):11157-11164; Rodgers, D. W., et al., (1995) *Proceedings Of the National Academy Of Sciences Of the United States Of America* 92(4):1222-1226; Sheng, N. and Dennis, D. (1993) *Biochemistry* 32(18):4938-4942; Spence, R. A., et al., (1995) *Science* 267(5200):988-993.}
- 35

WO 02/04493

PCT/US01/21241

Furthermore selected B- and/or T-cell epitopes can be added to the Pol constructs (e.g., 3' of the truncated INT or within the deletions of the RT- and INT-coding sequence) to replace and augment any epitopes deleted by the functional modifications of RT and INT. Alternately, selected B- and T-cell epitopes (including CTL epitopes) from RT and INT can be included in a minimal VLP formed by expression of the synthetic Gag or synthetic Pol cassette, described above. (For descriptions of known HIV B- and T-cell epitopes see, HIV Molecular Immunology Database CTL Search Interface; Los Alamos Sequence Compendia, 1987-1997; Internet address: <http://hiv-web.lanl.gov/immunology/index.html>.)

The resulting modified coding sequences are presented as a synthetic Env expression cassette; a synthetic Gag expression cassette; a synthetic Pol expression cassette. A common Gag region (Gag-common) extends from nucleotide position 844 to position 903 (SEQ ID NO:1), relative to AF110965 (or from approximately amino acid residues 282 to 301 of SEQ ID NO:17) and from nucleotide position 841 to position 900 (SEQ ID NO:2), relative to AF110967 (or from approximately amino acid residues 281 to 300 of SEQ ID NO:22). A common Env region (Env-common) extends from nucleotide position 1213 to position 1353 (SEQ ID NO:5) and amino acid positions 405 to 451 of SEQ ID NO:23, relative to AF110968 and from nucleotide position 1210 to position 1353 (SEQ ID NO:11) and amino acid positions 404-451 (SEQ ID NO:24), relative to AF110975.

The synthetic DNA fragments for Pol, Gag and Env are cloned into the following eucaryotic expression vectors: pCMVKm2, for transient expression assays and DNA immunization studies, the pCMVKm2 vector is derived from pCMV6a (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986) and comprises a kanamycin selectable marker, a ColE1 origin of replication, a CMV promoter enhancer and Intron A, followed by an insertion site for the synthetic sequences described below followed by a polyadenylation signal derived from bovine growth hormone -- the pCMVKm2 vector differs from the pCMV-link vector only in that a polylinker site is inserted into pCMVKm2 to generate pCMV-link; pESN2dhfr and pCMVPLEdhfr, for expression in Chinese Hamster Ovary (CHO) cells; and, pAcC13, a shuttle vector for use in the Baculovirus expression system (pAcC13, is derived from pAcC12 which is described by Munemitsu S., et al., *Mol Cell Biol.* 10(11):5977-5982, 1990).

Briefly, construction of pCMVPLEdhfr was as follows.

To construct a DHFR cassette, the EMCV IRES (internal ribosome entry site) leader was PCR-amplified from pCite-4a+ (Novagen, Inc., Milwaukee, WI) and inserted into pET-23d (Novagen, Inc., Milwaukee, WI) as an *Xba*-*Nco* fragment to give pET-EMCV. The *dhfr* gene was PCR-amplified from pESN2dhfr to give a product with a Gly-Gly-Gly-Ser spacer in place of the translation stop codon and inserted as an *Nco*-*Bam*HI fragment to give pET-E-

WO 02/04493

PCT/US01/21241

DHFR. Next, the attenuated *neo* gene was PCR amplified from a pSV2Neo (Clontech, Palo Alto, CA) derivative and inserted into the unique *Bam*HI site of pBT-E-DHFR to give pET-E-DHFR/Neo_(neo). Finally the bovine growth hormone terminator from pCDNA3 (Invitrogen, Inc., Carlsbad, CA) was inserted downstream of the *neo* gene to give pET-E-DHFR/Neo_(neo)BGHT. The EMCV-*dhfr*/*neo* selectable marker cassette fragment was prepared by cleavage of pET-E-DHFR/Neo_(neo)BGHT.

The CMV enhancer/promoter plus Intron A was transferred from pCMV6a (Chapman et al., *Nuc. Acids Res.* (1991) 19:3979-3986) as a *Hind*III-*Sal*I fragment into pUC19 (New England Biolabs, Inc., Beverly, MA). The vector backbone of pUC19 was deleted from the NdeI to the SapI sites. The above described DHFR cassette was added to the construct such that the EMCV IRES followed the CMV promoter. The vector also contained an *amp*^r gene and an SV40 origin of replication.

B. Defining of the Major Homology Region (MHR) of HIV-1 p55Gag

The Major Homology Region (MHR) of HIV-1 p55 (Gag) is located in the p24-CA sequence of Gag. It is a conserved stretch of approximately 20 amino acids. The position in the wild type AF110965 Gag protein is from 282-301 (SEQ ID NO:25) and spans a region from 844-903 (SEQ ID NO:26) for the Gag DNA-sequence. The position in the synthetic Gag protein is also from 282-301 (SEQ ID NO:25) and spans a region from 844-903 (SEQ ID NO:1) for the synthetic Gag DNA-sequence. The position in the wild type and synthetic AF110967 Gag protein is from 281-300 (SEQ ID NO:27) and spans a region from 841-900 (SEQ ID NO:2) for the modified Gag DNA-sequence. Mutations or deletions in the MHR can severely impair particle production (Borsetti, A., et al., *J. Virol.* 72(11):9313-9317, 1998; Mammano, F., et al., *J. Virol* 68(8):4927-4936, 1994).

Percent identity to this sequence can be determined, for example, using the Smith-Waterman search algorithm (Time Logic, Incline Village, NV), with the following exemplary parameters: weight matrix = nuc4x4hb; gap opening penalty = 20, gap extension penalty = 5.

C. Defining of the Common Sequence Region of HIV-1 Env

The common sequence region (CSR) of HIV-1 Env is located in the C4 sequence of Env. It is a conserved stretch of approximately 47 amino acids. The position in the wild type and synthetic AF110968 Env protein is from approximately amino acid residue 405 to 451 (SEQ ID NO:28) and spans a region from 1213 to 1353 (SEQ ID NO:5) for the Env DNA-sequence. The position in the wild type and synthetic AF110975 Env protein is

WO 02/04493

PCT/US01/21241

from approximately amino acid residue 404 to 451 (SEQ ID NO:29) and spans a region from 1210 to 1353 (SEQ ID NO:11) for the Env DNA-sequence.

Percent identity to this sequence can be determined, for example, using the Smith-Waterman search algorithm (Time Logic, Incline Village, NV), with the following exemplary
 5 parameters: weight matrix = nuc4x4hb; gap opening penalty = 20, gap extension penalty = 5.

Various forms of the different embodiments of the invention, described herein, may be combined.

D. Exemplary HIV Sequences Derived from South African HIV Type C Strains

10 HIV coding sequences of novel Type C isolates were obtained. Polypeptide-coding sequences were manipulated to maximize expression of their gene products.

As described above, the HIV-1 codon usage pattern was modified so that the resulting nucleic acid coding sequence was comparable to codon usage found in highly expressed human genes. The HIV codon usage reflects a high content of the nucleotides A or T of the
 15 codon-triplet. The effect of the HIV-1 codon usage is a high AT content in the DNA sequence that results in a decreased translation ability and instability of the mRNA. In comparison, highly expressed human codons prefer the nucleotides G or C. The coding sequences were modified to be comparable to codon usage found in highly expressed human genes.

20 Shown below in Table C are exemplary wild-type and synthetic sequences derived from a novel South African HIV Type C isolate, clone 8_5_TV1_C.ZA. Table D shows exemplary synthetic Env sequences derived from a novel South African HIV Type C isolate, clone 8_2_TV1_C.ZA. Table E shows wild-type and synthetic sequences derived from South African HIV Type C strain 12-5_1_TV2_C.ZA.

25

WO 02/04493

PCT/US01/21241

Table C

Name	SEQ ID	Description
C4_Env_TV1_C_ZA_opt short	46	synthetic sequence of short Env "common region"
5 C4_Env_TV1_C_ZA_opt	47	synthetic sequence of Env "common region"
C4_Env_TV1_C_ZA_wt	48	wild type 8_5_TV1_C.ZA Env sequence
Envgp160_TV1_C_ZAopt	49	synthetic Env gp160
Envgp160_TV1_C_ZAwt	50	wild type 8_5_TV1_C.ZA Env gp160 sequence
Gag_TV1_C_ZAopt	51	synthetic sequence of Gag
10 Gag_TV1_C_ZAwt	52	wild type 8_5_TV1_C.ZA Gag sequence
Gag_TV1_ZA_MHRopt	53	synthetic sequence of Gag major homology region
Gag_TV1_ZA_MHRwt	54	wild type 8_5_TV1_C.ZA Gag major homology region sequence
Nef_TV1_C_ZAopt	55	synthetic sequence of Nef
Nef_TV1_C_ZAwt	56	wild type 8_5_TV1_C.ZA Nef sequence
15 NefD125G_TV1_C_ZAopt	57	synthetic sequence of Nef _i including mutation at position 125 resulting in non-functional gene product
p15RNaseH_TV1_C_ZAopt	58	synthetic sequence of RNaseH (p15 of Pol)
p15RNaseH_TV1_C_ZAwt	59	wild type 8_5_TV1_C.ZA RNaseH sequence
p31Int_TV1_C_ZAopt	60	synthetic sequence of Integrase (p31 of Pol)
p31Int_TV1_C_ZAwt	61	wild type 8_5_TV1_C.ZA Integrase sequence
20 Pol_TV1_C_ZAopt	62	synthetic sequence of Pol
Pol_TV1_C_ZAwt	63	wild type 8_5_TV1_C.ZA Pol sequence
Prot_TV1_C_ZAopt	64	synthetic sequence of Prot
Prot_TV1_C_ZAwt	65	wild type 8_5_TV1_C.ZA Prot sequence
Protina_TV1_C_ZAopt	66	synthetic sequence of Prot including mutation resulting in inactivation of protease

WO 02/04493

PCT/US01/21241

5	Protina_TV1_C_ZAwt	67	wild type 8_5_TV1_C.ZA Prot sequence, including mutation resulting in inactivation of protease.
	ProtinaRTmut_TV1_C_ZAopt	68	synthetic sequence of Prot and reverse transcriptase (RT), including mutation resulting in inactivation of protease and mutation resulting in inactivation of RT.
	ProtinaRTmut_TV1_C_ZAwt	69	wild type 8_5_TV1_C.ZA Prot and RT, mutation resulting in inactivation of protease and mutation resulting in inactivation of RT.
	ProtwtRTwt_TV1_C_ZAopt	70	synthetic sequences of Prot and RT
	ProtwtRTwt_TV1_C_ZAwt	71	wild type 8_5_TV1_C.ZA Prot and RT
10	RevExon1_TV1_C_ZAopt	72	synthetic sequence of exon 1 of Rev
	RevExon1_TV1_C_ZAwt	73	wild type 8_5_TV1_C.ZA of exon 1 of Rev
	RevExon2_TV1_C_ZAopt-2	74	synthetic sequence of exon 2 of Rev
	RevExon2_TV1_C_ZAwt	75	wild type 8_5_TV1_C.ZA of exon 2 of Rev
	RT_TV1_C_ZAopt	76	synthetic sequence of RT
15	RT_TV1_C_ZAwt	77	wild type 8_5_TV1_C.ZA RT
	RTmut_TV1_C_ZAopt	78	synthetic sequence of RT, including mutation resulting in inactivation of RT
	RTmut_TV1_C_ZAwt	79	wild type 8_5_TV1_C.ZA RT, including mutation resulting in inactivation of RT
	TatC22Exon1_TV1_C_ZAopt	80	synthetic sequence of exon 1 of Tat, including mutation resulting in non-functional Tat gene product
	TatExon1_TV1_C_ZAopt	81	synthetic sequence of exon 1 of Tat
20	TatExon1_TV1_C_ZAwt	82	wild type 8_5_TV1_C.ZA exon 1 of Tat
	TatExon2_TV1_C_ZAopt	83	synthetic sequence of exon 2 of Tat
	TatExon2_TV1_C_ZAwt	84	wild type 8_5_TV1_C.ZA exon 2 of Tat
	Vif_TV1_C_ZAopt	85	synthetic sequence of Vif
	Vif_TV1_C_ZAwt	86	wild type 8_5_TV1_C.ZA Vif
	Vpr_TV1_C_ZAopt	87	synthetic sequence of Vpr

WO 02/04493

PCT/US01/21241

5	Vpr_TV1_C_ZAwt	88	wild type 8_5_TV1_C.ZA Vpr
	Vpu_TV1_C_ZAopt	89	synthetic sequence of Vpu
	Vpu_TV1_C_ZAwt	90	wild type 8_5_TV1_C.ZA Vpu
	revexon1_2_TV1_C_ZAopt	91	synthetic sequence of exons 1 and 2 of Rev
	RevExon1_2_TV1_C_ZAwt	92	wild type 8_5_TV1_C.ZA Rev (exons 1 and 2)
10	TatC22Exon1_2_TV1_C_ZAopt	93	synthetic sequence of exons 1 and 2 of Tat, including mutation in exon 1 resulting in non-functional Tat gene product
	TatExon1_2_TV1_C_ZAopt	94	synthetic sequence of exons 1 and 2 of Tat
	TatExon1_2_TV1_C_ZAwt	95	wild type 8_5_TV1_C.ZA Tat (exons 1 and 2)
	NefD125G-Myr_TV1_C_ZAopt	96	synthetic sequence of Nef, including mutation eliminating myristoylation site.

WO 02/04493

PCT/US01/21241

Table D

Name	Seq ID	Description
gp120mod.TV1.delV2	119	synthetic sequence of Env gp120, including V2 deletion and modified leader sequences derived from wild-type 8_2_TV1_C.ZA sequences
gp140mod.TV1.delV2	120	synthetic sequence of Env gp140, including V2 deletion and modified leader sequences derived from wild-type 8_2_TV1_C.ZA sequences
5 gp140mod.TV1.mut7.delV2	121	synthetic sequence of Env gp140, including V2 deletion and mutation in cleavage site and modified leader sequences derived from wild-type 8_2_TV1_C.ZA sequences
gp160mod.TV1.delV1V2	122	synthetic sequence of Env gp160, including V1/V2 deletion and modified leader derived from wild-type 8_2_TV1_C.ZA sequences
gp160mod.TV1.delV2	123	synthetic sequence of Env gp160, including V2 deletion and modified leader sequences derived from wild-type 8_2_TV1_C.ZA sequences
gp160mod.TV1.mut7.delV2	124	synthetic sequence of Env gp160, including V2 deletion; a mutation in cleavage site; and modified leader sequences derived from wild-type 8_2_TV1_C.ZA sequences
10 gp160mod.TV1.tpa1	125	synthetic sequence of Env gp160, TPA1 leader
gp160mod.TV1	126	synthetic sequence of Env gp160, including modified leader sequences derived from wild-type (8_2_TV1_C.ZA) sequences
gp160mod.TV1.wtNative	127	synthetic sequence of Env gp160, including wild type 8_2_TV1_C.ZA (unmodified) leader
gp140mod.TV1.tpa1	131	synthetic sequence of Env gp140, TPA1 leader
gp140mod.TV1	132	synthetic sequence of Env gp140, including modified leader sequences derived from wild-type 8_2_TV1_C.ZA sequences
15 gp140mod.TV1.wtNative	133	synthetic sequence of Env gp120, including wild type 8_2_TV1_C.ZA (unmodified) leader sequence.

As noted above, Env-encoding constructs can be prepared using any of the full-length of gp160 constructs. For example, a gp140 form (SEQ ID NO:132) was made by truncating gp160 (SEQ ID NO:126) at nucleotide 2064; gp120 was made by truncating gp160 (SEQ ID

WO 02/04493

PCT/US01/21241

NO:126) at nucleotide 1551 (SEQ ID NO:126). Additional gp140 and gp120 forms can be made using the methods described herein. One or more stop codons are typically added (e.g., nucleotides 2608 to 2610 of SEQ ID NO:126). Further, the wild-type leader sequence can be modified and/or replaced with other leader sequences (e.g., TPA1 leader sequences).

Thus, the polypeptide gp160 includes the coding sequences for gp120 and gp41. The polypeptide gp41 is comprised of several domains including an oligomerization domain (OD) and a transmembrane spanning domain (TM). In the native envelope, the oligomerization domain is required for the non-covalent association of three gp41 polypeptides to form a trimeric structure: through non-covalent interactions with the gp41 trimer (and itself), the gp120 polypeptides are also organized in a trimeric structure. A cleavage site (or cleavage sites) exists approximately between the polypeptide sequences for gp120 and the polypeptide sequences corresponding to gp41. This cleavage site(s) can be mutated to prevent cleavage at the site. The resulting gp140 polypeptide corresponds to a truncated form of gp160 where the transmembrane spanning domain of gp41 has been deleted. This gp140 polypeptide can exist in both monomeric and oligomeric (*i.e.* trimeric) forms by virtue of the presence of the oligomerization domain in the gp41 moiety. In the situation where the cleavage site has been mutated to prevent cleavage and the transmembrane portion of gp41 has been deleted the resulting polypeptide product is designated "mutated" gp140 (e.g., gp140.mut). As will be apparent to those in the field, the cleavage site can be mutated in a variety of ways. In the exemplary constructs described herein (e.g., SEQ ID NO:121 and SEQ ID NO:124), the mutation in the gp120/gp41 cleavage site changes the wild-type amino acid sequence KRRVVQREKR (SEQ ID NO:129) to ISSVVQSEKS (SEQ ID NO:130).

In yet other embodiments, hypervariable region(s) were deleted, N-glycosylation sites were removed and/or cleavage sites mutated. Exemplary constructs having variable region deletions (V1 and/or V2), V2 deletes were constructed by deleting nucleotides from approximately 499 to approximately 593 (relative to SEQ ID NO:128) and V1/V2 deletes were constructed by deleting nucleotides from approximately 375 to approximately 602 (relative to SEQ ID NO:128). The relative locations of V1 and/or V2 regions can also be readily determined by alignment to the regions shown in Table A. Table E shows wild-type and synthetic sequences derived from South African HIV Type C strain 12-S_1_TV2_C.ZA.

Table E

Name	SEQ ID	Description
Envgp160_TV2_C_ZAopt	97	synthetic sequence of Env gp160

WO 02/04493

PCT/US01/21241

	Envgp160_TV2_C_ZAwt	98	wild type 12-5_1_TV2_C.ZA Env gp160.
	Gag_TV2_C_ZAopt	99	synthetic sequence of Gag
	Gag_TV2_C_ZAwt	100	wild type 12-5_1_TV2_C.ZA Gag
	Nef_TV2_C_ZAopt	101	synthetic sequence of Nef
5	Nef_TV2_C_ZAwt	102	wild type 12-5_1_TV2_C.ZA Nef
	Pol_TV2_C_ZAopt	103	synthetic sequence of Pol
	Pol_TV2_C_ZAwt	104	wild type 12-5_1_TV2_C.ZA of Pol
	RevExon1_TV2_C_ZAopt	105	synthetic sequence of exon 1 of Rev
	RevExon1_TV2_C_ZAwt	106	wild type 12-5_1_TV2_C.ZA of exon 1 of Rev
10	RevExon2_TV2_C_ZAopt	107	synthetic sequence of exon 2 of Rev
	RevExon2_TV2_C_ZAwt	108	wild type 12-5_1_TV2_C.ZA of exon 2 of Rev
	TatExon1_TV2_C_ZAopt	109	synthetic sequence of exon 1 of Tat
	TatExon1_TV2_C_ZAwt	110	wild type 12-5_1_TV2_C.ZA of exon 1 of Tat
	TatExon2_TV2_C_ZAopt	111	synthetic sequence of exon 2 of Tat
15	TatExon2_TV2_C_ZAwt	112	wild type 12-5_1_TV2_C.ZA of exon 2 of Tat
	Vif_TV2_C_ZAopt	113	synthetic sequence of Vif
	Vif_TV2_C_ZAwt	114	wild type 12-5_1_TV2_C.ZA of Vif
	Vpr_TV2_C_ZAopt	115	synthetic sequence of Vpr
	Vpr_TV2_C_ZAwt	116	wild type 12-5_1_TV2_C.ZA of Vpr
20	Vpu_TV2_C_ZAopt	117	synthetic sequence of Vpu
	Vpu_TV2_C_ZAwt	118	wild type 12-5_1_TV2_C.ZA of Vpu

It will be readily apparent that sequences derived from any HIV type C strain or clone can be modified as described herein in order to achieve desirable modifications in that strain.

- 25 Additionally, polyproteins can be constructed by fusing in-frame two or more polynucleotide sequences encoding polypeptide or peptide products. Further, polycistronic coding sequences may be produced by placing two or more polynucleotide sequences encoding polypeptide products adjacent each other, typically under the control of one promoter, wherein each polypeptide coding sequence may be modified to include sequences for internal ribosome
- 30 binding sites.

WO 02/04493

PCT/US01/21241

The sequences of the present invention, for example, the modified (synthetic) polynucleotide sequences encoding HIV polypeptides, may be modified by deletions, point mutations, substitutions, frame-shifts, and/or further genetic modifications (for example, mutations leading to inactivation of an activity associated with a polypeptide, e.g., mutations that inactivate protease, tat, or reverse transcriptase activity). Such modifications are taught generally in the art and may be applied in the context of the teachings of the present invention. For example, sites corresponding to the "Regions of the HIV Genome" listed in Table A may be modified in the corresponding regions of the novel sequences disclosed herein in order to achieve desirable modifications. Further, the modified (synthetic) polynucleotide sequences of the present invention can be combined for use, e.g., in an composition for generating an immune response in a subject, in a variety of ways, including but not limited to the following ways: multiple individual expression cassettes each comprising one polynucleotide sequence of the present invention (e.g., a gag-expression cassette, an env expression cassette, and a rev expression cassette, or a pol-expression cassette, a vif expression cassette, and a vpr expression cassette, etc.); polyproteins produced by in-frame fusions of multiple polynucleotides of the present invention, and polycistronic polynucleotides produced using multiple polynucleotides of the present invention.

Example 2

Expression Assays for the Synthetic Coding Sequences

A. Type C HIV Coding Sequences

The wild-type Subtype C HIV coding (for example from AF110965, AF110967, AF110968, AF110975, as well as novel South African strains 8_5_TV1_C.ZA, 8_2_TV1_C.ZA and 12-5_1_TV2_C.ZA) sequences are cloned into expression vectors having the same features as the vectors into which the synthetic sequences are cloned.

Expression efficiencies for various vectors carrying the wild-type and synthetic sequences are evaluated as follows. Cells from several mammalian cell lines (293, RD, COS-7, and CHO; all obtained from the American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209) are transfected with 2 µg of DNA in transfection reagent LT1 (PanVera Corporation, 545 Science Dr., Madison, WI). The cells are incubated for 5 hours in reduced serum medium (Opti-MEM, Gibco-BRL, Gaithersburg, MD). The medium is then replaced with normal medium as follows: 293 cells, IMDM, 10% fetal calf

WO 02/04493

PCT/US01/21241

serum, 2% glutamine (BioWhittaker, Walkersville, MD); RD and COS-7 cells, D-MEM, 10% fetal calf serum, 2% glutamine (Opti-MEM, Gibco-BRL, Gaithersburg, MD); and CHO cells, Ham's F-12, 10% fetal calf serum, 2% glutamine (Opti-MEM, Gibco-BRL, Gaithersburg, MD). The cells are incubated for either 48 or 60 hours. Cell lysates are collected as described below in Example 3. Supernatants are harvested and filtered through 0.45 μ m syringe filters. Supernatants are evaluated using the using 96-well plates coated with a murine monoclonal antibody directed against HIV antigen, for example a Coulter p24-assay (Coulter Corporation, Hialeah, FL, US). The HIV-1 antigen binds to the coated wells. Biotinylated antibodies against HIV recognize the bound antigen. Conjugated streptavidin-horseradish peroxidase reacts with the biotin. Color develops from the reaction of peroxidase with TMB substrate. The reaction is terminated by addition of 4N H₂SO₄. The intensity of the color is directly proportional to the amount of HIV antigen in a sample.

Synthetic HIV Type C expression cassettes provides dramatic increases in production of their protein products, relative to the native (wild-type Subtype C) sequences, when expressed in a variety of cell lines.

B. Signal Peptide Leader Sequences

The ability of various leader sequences to drive expression was tested by transfecting cells with wild type or synthetic Env-encoding expression cassettes operably linked to different leader sequences and evaluating expression of Env polypeptide by ELISA or Western Blot. The amino acid and nucleotide sequence of various signal peptide leader sequences are shown in Table 4.

Table 4

Leader	Amino acid sequence	DNA sequence
WTnative (8 ₂ _TV 1_C.ZA)	MRVMGTQKNCQQWWIWGI LGFWMMLMIC	ATGAGAGTGATGGGGACACAGA AGAATTGTCAACAATGGTGGATA TGGGGCATCTTAGGCTTCTGGAT GCTAATGATTGT
WTmod (8 ₂ _TV 1_C.ZA)	MRVMGTQKNCQQWWIWGI LGFWMMLMIC	ATGCGCGTGATGGGCACCCAGAA GAACTGCCAGCAGTGGTGGATCT GGGGCATCCTGGGCTTCTGGATG CTGATGATCTGC
Tpa1	MDAMKRGLCVLLCGAVFVSPS AS	ATGGATGCAATGAAGAGAGGGC TCTGCTGTGCTGCTGCTGTGTG

WO 02/04493

PCT/US01/21241

		GAGCAGICTTCGTTTCGCCAGC GCCAGC
Tpa2	MDAMERGLCEVLLCGAVFVSPS	ATGGATGCAATGAAGAGAGGGC TCTGCTGTGTGCTGCTGCTGTGTG GAGCAGICTTCGTTTCGCCAGC

35 293 cells were transiently transfected using standard methods with native and
sequence-modified constructs encoding the gp120 and gp140 forms of the 8_2_TV1_CZA
(TV1c8.2) envelope. Env protein was measure in cell lysates and supernatants using an in-
house Env capture ELISA. Results are shown in Table 5 below and indicate that the wild-
type signal peptide leader sequence of the TV1c8.2 can be used to efficiently express the
40 encoded envelope protein to levels that are better or comparable to those observed using the
heterologous tpa leader sequences. Furthermore, the TV1c8.2 leader works in its native or
sequence-modified forms and can be used with native or sequence-modified env genes. All
constructs were tested after cloning of the gene cassettes into the EcoRI and XhoI sites of the
pCMVlink expression vector.

45

WO 02/04493

PCT/US01/21241

Table 5

	TV1c8.2 construct	Supernatant (ng)	Lysate (ng)	Total (ng)
5	gp140nat.wtL	532	149	681
	gp140nat.tpa1	250	20	270
	gp140nat.tpa2	192	34	226
	gp120mod.wtLmod	6186	4576	10762
	gp120mod.tpa1	6932	3808	10740
10	gp120mod.wtLnat	6680	4174	10854
	gp140mod.wtLmod	1844	8507	10351
	gp140mod.tpa1	1854	2925	4779
	gp140mod.wtLnat	1532	3015	4547

The sequence-modified TV1c8.2 envelope variant gene cassettes were subcloned into a Chiron pCMV expression vector for the derivation of stable mammalian cell lines. Stable CHO cell lines expressing the TV1c8.2 envelope proteins were derived using standard methods of transfection, methotrexate amplification, and screening. These cell lines were found to secrete levels of envelope protein that were comparable to those observed for proteins expressed using the tpa leader sequences. Representative results are shown in Table 6 for two cell line clones expressing the TV1c8.2 gp120; they are compared to two reference clones expressing SF162 subtype B gp120 derived in a similar fashion but using the tpa leader. Protein concentrations were determined following densitometry of scanned gels of semi-purified proteins. Standard curves were generated using a highly purified and well-characterized preparation of SF2 gp120 protein and the concentrations of the test proteins were determined.

Table 6

CHO cell line	Clone #	Expression (ng/ml)
gp120 SF162	Clone 65	921
	Clone 71	972
gp120TV1.C8.2	Clone 159	1977
	Clone 210	1920

The results were also confirmed by Western Blot Analysis, essentially as described in Example 3.

WO 02/04493

PCT/US01/21241

Example 3Western Blot Analysis of ExpressionA. HIV Type C Coding Sequences

Human 293 cells are transfected as described in Example 2 with pCMV-based vectors containing native or synthetic HIV Type C expression cassettes. Cells are cultivated for 60 hours post-transfection. Supernatants are prepared as described. Cell lysates are prepared as follows. The cells are washed once with phosphate-buffered saline, lysed with detergent [1% NP40 (Sigma Chemical Co., St. Louis, MO) in 0.1 M Tris-HCl, pH 7.5], and the lysate transferred into fresh tubes. SDS-polyacrylamide gels (pre-cast 8-16%; Novex, San Diego, CA) are loaded with 20 μ l of supernatant or 12.5 μ l of cell lysate. A protein standard is also loaded (5 μ l, broad size range standard; BioRad Laboratories, Hercules, CA). Electrophoresis is carried out and the proteins are transferred using a BioRad Transfer Chamber (BioRad Laboratories, Hercules, CA) to Immobilon P membranes (Millipore Corp., Bedford, MA) using the transfer buffer recommended by the manufacturer (Millipore), where the transfer is performed at 100 volts for 90 minutes. The membranes are exposed to HIV-1-positive human patient serum and immunostained using o-phenylenediamine dihydrochloride (OPD; Sigma).

Immunoblotting analysis shows that cells containing the synthetic expression cassette produce the expected protein at higher per-cell concentrations than cells containing the native expression cassette. The proteins are seen in both cell lysates and supernatants. The levels of production are significantly higher in cell supernatants for cells transfected with the synthetic expression cassettes of the present invention.

In addition, supernatants from the transfected 293 cells are fractionated on sucrose gradients. Aliquots of the supernatant are transferred to Polyclear™ ultra-centrifuge tubes (Beckman Instruments, Columbia, MD), under-laid with a solution of 20% (wt/wt) sucrose, and subjected to 2 hours centrifugation at 28,000 rpm in a Beckman SW28 rotor. The resulting pellet is suspended in PBS and layered onto a 20-60% (wt/wt) sucrose gradient and subjected to 2 hours centrifugation at 40,000 rpm in a Beckman SW41ti rotor.

The gradient is then fractionated into approximately 10 x 1 ml aliquots (starting at the top, 20%-end, of the gradient). Samples are taken from fractions 1-9 and are electrophoresed on 8-16% SDS polyacrylamide gels. The supernatants from 293/synthetic cells give much stronger bands than supernatants from 293/native cells.

WO 02/04493

PCT/US01/21241

Example 4In Vivo Immunogenicity of Synthetic HIV Type C Expression CassettesA. Immunization

To evaluate the possibly improved immunogenicity of the synthetic HIV Type C expression cassettes, a mouse study is performed. The plasmid DNA, pCMVCM2 carrying the synthetic Gag expression cassette, is diluted to the following final concentrations in a total injection volume of 100 μ l: 20 μ g, 2 μ g, 0.2 μ g, 0.02 and 0.002 μ g. To overcome possible negative dilution effects of the diluted DNA, the total DNA concentration in each sample is brought up to 20 μ g using the vector (pCMVCM2) alone. As a control, plasmid DNA of the native Gag expression cassette is handled in the same manner. Twelve groups of four to ten Balb/c mice (Charles River, Boston, MA) are intramuscularly immunized (50 μ l per leg, intramuscular injection into the *tibialis anterior*) according to the schedule in Table 1.

WO 02/04493

PCT/US01/21241

Table 1

Group	Gag or Env Expression Cassette	Concentration of Gag or Env plasmid DNA (μ g)	Immunized at time (weeks):
1	Synthetic	20	0 ¹ , 4
2	Synthetic	2	0, 4
3	Synthetic	0.2	0, 4
4	Synthetic	0.02	0, 4
5	Synthetic	0.002	0, 4
6	Synthetic	20	0
7	Synthetic	2	0
8	Synthetic	0.2	0
9	Synthetic	0.02	0
10	Synthetic	0.002	0
11	Native	20	0, 4
12	Native	2	0, 4
13	Native	0.2	0, 4
14	Native	0.02	0, 4
15	Native	0.002	0, 4
16	Native	20	0
17	Native	2	0
18	Native	0.2	0
19	Native	0.02	0
20	Native	0.002	0

1 = initial immunization at "week 0"

Groups 1-5 and 11-15 are bled at week 0 (before immunization), week 4, week 6, week 8, and week 12. Groups 6-20 and 16-20 are bled at week 0 (before immunization) and at week 4.

B. Humoral Immune Response

The humoral immune response is checked with an anti-HIV antibody ELISAs (enzyme-linked immunosorbent assays) of the mice sera 0 and 4 weeks post immunization (groups 5-12) and, in addition, 6 and 8 weeks post immunization, respectively, 2 and 4 weeks post second immunization (groups 1-4).

WO 02/04493

PCT/US01/21241

The antibody titers of the sera are determined by using the appropriate anti-HIV polypeptide (e.g., anti-Pol, anti-Gag, anti-Env, anti-Vif, anti-Vpu, etc.) antibody ELISA. Briefly, sera from immunized mice are screened for antibodies directed against the HIV proteins (e.g., p55 Gag protein, an Env protein, e.g., gp160 or gp120 or a Pol protein, e.g., p6, 5 prot or RT, etc). ELISA microtiter plates are coated with 0.2 µg of HIV protein per well overnight and washed four times; subsequently, blocking is done with FBS-0.2% Tween (Sigma) for 2 hours. After removal of the blocking solution, 100 µl of diluted mouse serum is added. Sera are tested at 1/25 dilutions and by serial 3-fold dilutions, thereafter. Microtiter 10 plates are washed four times and incubated with a secondary, peroxidase-coupled anti-mouse IgG antibody (Pierce, Rockford, IL). ELISA plates are washed and 100 µl of 3, 3', 5, 5'-tetramethyl benzidine (TMB; Pierce) is added per well. The optical density of each well is measured after 15 minutes. The titers reported are the reciprocal of the dilution of serum that gave a half-maximum optical density (O.D.).

15 Synthetic expression cassettes will provide a clear improvement of immunogenicity relative to the native expression cassettes.

C. Cellular Immune Response

The frequency of specific cytotoxic T-lymphocytes (CTL) is evaluated by a standard chromium release assay of peptide pulsed mouse (Balb/c, CB6F1 and/or C3H) CD4 cells. 20 HIV polypeptide (e.g., Pol, Gag or Env) expressing vaccinia virus infected CD-8 cells are used as a positive control. Briefly, spleen cells (Effector cells, E) are obtained from the mice immunized as described above are cultured, restimulated, and assayed for CTL activity against Gag peptide-pulsed target cells as described (Doe, B., and Walker, C.M., *AIDS* 10(7):793-794, 1996). Cytotoxic activity is measured in a standard ⁵¹Cr release assay. Target 25 (T) cells are cultured with effector (E) cells at various E:T ratios for 4 hours and the average cpm from duplicate wells are used to calculate percent specific ⁵¹Cr release.

Cytotoxic T-cell (CTL) activity is measured in splenocytes recovered from the mice immunized with HIV Gag or Env DNA. Effector cells from the Gag or Env DNA-immunized animals exhibit specific lysis of HIV polypeptide-pulsed SV-BALB (MHC 30 matched) targets cells, indicative of a CTL response. Target cells that are peptide-pulsed and derived from an MHC-unmatched mouse strain (MCS7) are not lysed.

WO 02/04493

PCT/US01/21241

Thus, synthetic expression cassettes exhibit increased potency for induction of cytotoxic T-lymphocyte (CTL) responses by DNA immunization.

Example 5

DNA-immunization of Non-Human Primates Using a Synthetic HIV Type C Expression Cassette

Non-human primates are immunized multiple times (*e.g.*, weeks 0, 4, 8 and 24) intradermally, mucosally or bilaterally, intramuscular, into the quadriceps using various doses (*e.g.*, 1-5 mg) and various combinations of synthetic HIV Type C plasmids. The animals are bled two weeks after each immunization and ELISA is performed with isolated plasma. The ELISA is performed essentially as described in Example 4 except the second antibody-conjugate is an anti-human IgG, g-chain specific, peroxidase conjugate (Sigma Chemical Co., St. Louis, MD 63178) used at a dilution of 1:500. Fifty µg/ml yeast extract is added to the dilutions of plasma samples and antibody conjugate to reduce non-specific background due to preexisting yeast antibodies in the non-human primates.

Further, lymphoproliferative responses to antigen can also be evaluated post-immunization, indicative of induction of T-helper cell functions.

Synthetic plasmid DNA are expected to be immunogenic in non-human primates.

Example 6

In vitro expression of recombinant Sindbis RNA and DNA containing the synthetic HIV Type C expression cassette

To evaluate the expression efficiency of the synthetic Pol, Env and Gag expression cassette in Alphavirus vectors, the selected synthetic expression cassette is subcloned into both plasmid DNA-based and recombinant vector particle-based Sindbis virus vectors. Specifically, a cDNA vector construct for *in vitro* transcription of Sindbis virus RNA vector replicons (pRSIN-luc; Dubensky, et al., *J Virol.* 70:508-519, 1996) is modified to contain a *PmeI* site for plasmid linearization and a polylinker for insertion of heterologous genes. A polylinker is generated using two oligonucleotides that contain the sites *XhoI*, *FmII*, *ApaI*, *NarI*, *XbaI*, and *NotI* (XPANXNF, and XPANXNR).

The plasmid pRSIN-luc (Dubensky et al., *supra*) is digested with *XhoI* and *NotI* to remove the luciferase gene insert, blunt-ended using Klenow and dNTPs, and purified from

WO 02/04493

PCT/US01/21241

an agarose gel using GeneCleanII (BioLol, Vista, CA). The oligonucleotides are annealed to each other and ligated into the plasmid. The resulting construct is digested with *NotI* and *SacI* to remove the minimal Sindbis 3'-end sequence and A_{10} tract, and ligated with an approximately 0.4 kbp fragment from PKSSIN1-BV (WO 97/38087). This 0.4 kbp fragment is obtained by digestion of PKSSIN1-BV with *NotI* and *SacI*, and purification after size fractionation from an agarose gel. The fragment contains the complete Sindbis virus 3'-end, an A_{10} tract and a *PmeI* site for linearization. This new vector construct is designated SINBVE.

The synthetic HIV coding sequences are obtained from the parental plasmid by digestion with *EcoRI*, blunt-ending with Klenow and dNTPs, purification with GeneCleanII, digestion with *SaI*, size fractionation on an agarose gel, and purification from the agarose gel using GeneCleanII. The synthetic HIV polypeptide-coding fragment is ligated into the SINBVE vector that is digested with *XhoI* and *PmlI*. The resulting vector is purified using GeneCleanII and is designated SINBVGag. Vector RNA replicons may be transcribed *in vitro* (Dubensky et al., *supra*) from SINBVGag and used directly for transfection of cells. Alternatively, the replicons may be packaged into recombinant vector particles by co-transfection with defective helper RNAs or using an alphavirus packaging cell line.

The DNA-based Sindbis virus vector pDCMV SIN-beta-gal (Dubensky, et al., *J Virol.* 70:508-519, 1996) is digested with *SaI* and *XbaI*, to remove the beta-galactosidase gene insert, and purified using GeneCleanII after agarose gel size fractionation. The HIV Gag or Env gene is inserted into the pDCMV SIN-beta-gal by digestion of SINBVGag with *SaI* and *XhoI*, purification using GeneCleanII of the Gag-containing fragment after agarose gel size fractionation, and ligation. The resulting construct is designated pDSIN-Gag, and may be used directly for *in vivo* administration or formulated using any of the methods described herein.

BHK and 293 cells are transfected with recombinant Sindbis RNA and DNA, respectively. The supernatants and cell lysates are tested with the Coulter capture ELISA (Example 2).

BHK cells are transfected by electroporation with recombinant Sindbis RNA.

293 cells are transfected using LT-1 (Example 2) with recombinant Sindbis DNA. Synthetic Gag- and/or Env-containing plasmids are used as positive controls. Supernatants and lysates are collected 48h post transfection.

WO 02/04493

PCT/US01/21241

Type C HIV proteins can be efficiently expressed from both DNA and RNA-based Sindbis vector systems using the synthetic expression cassettes.

Example 7

5 *In Vivo* Immunogenicity of recombinant Sindbis Replicon Vectors
containing synthetic Pol, Gag and/or Env Expression Cassettes

A. Immunization

To evaluate the immunogenicity of recombinant synthetic HIV Type C expression cassettes in Sindbis replicons, a mouse study is performed. The Sindbis virus DNA vector carrying synthetic expression cassettes (Example 6), is diluted to the following final
10 concentrations in a total injection volume of 100 μ l: 20 μ g, 2 μ g, 0.2 μ g, 0.02 and 0.002 μ g. To overcome possible negative dilution effects of the diluted DNA, the total DNA concentration in each sample is brought up to 20 μ g using the Sindbis replicon vector DNA alone. Twelve groups of four to ten Balb/c mice (Charles River, Boston, MA) are
15 intramuscularly immunized (50 μ l per leg, intramuscular injection into the *tibialis anterior*) according to the schedule in Table 2. Alternatively, Sindbis viral particles are prepared at the following doses: 10^3 pfu, 10^5 pfu and 10^7 pfu in 100 μ l, as shown in Table 3. Sindbis HIV polypeptide particle preparations are administered to mice using intramuscular and
20 subcutaneous routes (50 μ l per site).

WO 02/04493

PCT/US01/21241

Table 2

Group	Gag or Env Expression Cassette	Concentration of Gag or Env DNA (μ g)	Immunized at time (weeks):
1	Synthetic	20	0 ¹ , 4
2	Synthetic	2	0, 4
3	Synthetic	0.2	0, 4
4	Synthetic	0.02	0, 4
5	Synthetic	0.002	0, 4
6	Synthetic	20	0
7	Synthetic	2	0
8	Synthetic	0.2	0
9	Synthetic	0.02	0
10	Synthetic	0.002	0

1 = initial immunization at "week 0"

Table 3

Group	Gag or Env sequence	Concentration of viral particle (p.f.u.)	Immunized at time (weeks):
1	Synthetic	10^5	0 ¹ , 4
2	Synthetic	10^5	0, 4
3	Synthetic	10^7	0, 4
8	Synthetic	10^5	0
9	Synthetic	10^5	0
10	Synthetic	10^7	0

1 = initial immunization at "week 0"

Groups are bled and assessment of both humoral and cellular (e.g., frequency of specific CTLs) is performed, essentially as described in Example 4.

Example 8

Identification and Sequencing of a Novel HIV Type C Variants

WO 02/04493

PCT/US01/21241

A full-length clone, called 8_5_TV1_C.ZA, encoding an HIV Type C was isolated and sequenced. Briefly, genomic DNA from HIV-1 subtype C infected South African patients was isolated from PBMC (peripheral blood mononuclear cells) by alkaline lysis and anion-exchange columns (Qiagen). To get the genome of full-length clones two halves were amplified, that could later be joined together in frame within the Pol region using an unique Sal I site in both fragments. For the amplification, 200-800 ng of genomic DNA were added to the buffer and enzyme mix of the Expand Long Template PCR System after the protocol of the manufacturer (Boehringer Mannheim). The primer were designed after alignments of known full length sequences. For the 5'half a primer mix of 2 forward primers containing either thymidine (S1FCSacTA 5'-GTTTCTTGAGCTCTGGAAGGGTTAATTAC TCCAAGAA-3', SEQ ID NO:38) or cytosine on position 20 (S1FTSacTA 5'-GTTTCTTGAGCTCTGGAAGGGTTAATTACTCTAAGAA, SEQ ID NO:39) plus Sal I site, were used. The reverse primer were also a mix of two primers with either thymidine or cytosine on position 13 (S145RTSaTA 5'-GTTTCTTGTCGACTTGTCATGTATGGCTTCCCC T-3', SEQ ID NO:40 and S145RCSaTA 5'-GTTTCTTGTCGACTTGTCATGCATGGCTTCCCT-3' SEQ ID NO:41) and contained a Sal I site. The forward primer for the 3'half was also a mixture of two primers (S245FASaTA 5'-GTTTCTTGTCGACTGTAGTCCAGGaATATGGCAAT TAG-3' SEQ ID NO:42 and S245FGSaTA 5'-GTTTCTTGTCGACTGTAGTCCAGGaATATG CCAA TTAG-3' SEQ ID NO:43) with Sal I site and adenine or guanine on position 12. The reverse primer had a Not I site (S2_FullNotTA 5'-GTTTCTTGCGGCCGCTGCTAGA GATTTCCACACTACCA-3' SEQ ID NO:44). After amplification the PCR products were purified using a 1% agarose gel and cloned into the pCR-XL-TOPO vector via TA cloning (Invitrogen). Colonies were checked by restriction analysis and sequence verified. For the full length sequence the sequences of the 5'- and 3'half were combined. The sequence is shown in SEQ ID NO:33. Furthermore, important domains are shown in Table A.

Another clone, designated 12-5_1_TV2_C.ZA was also sequenced and is shown in SEQ ID NO:45. The domains can be readily determined in view of the teachings of the specification, for example by aligning the sequence to those shown in Table A to find the corresponding regions in clone 12-5_1_TV2_C.ZA.

WO 02/04493

PCT/US01/21241

As described above (Example 1, Table C), synthetic expression cassettes were generated using one or more polynucleotide sequences obtained from 8_5_TV1_C.ZA or 12-5_1_TV2_C.ZA.

5 The polynucleotides described herein have all been deposited at Chiron Corporation, Emeryville, CA.

Although preferred embodiments of the subject invention have been described in some detail, it is understood that obvious variations can be made without departing from the spirit and the scope of the invention as defined by the appended claims.

WO 02/04493

PCT/US01/21241

Claims

1. An expression cassette comprising
a polynucleotide sequence encoding a polypeptide including an HIV *Pol* polypeptide,
5 wherein the polynucleotide sequence encoding said *Pol* polypeptide comprises a sequence
having at least 90% sequence identity to the sequence presented of Figure 8 (SEQ ID NO:30);
Figure 9 (SEQ ID NO:31) or Figure 10 (SEQ ID NO:32).
2. An expression cassette comprising
10 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:46,
(ii) X equals Y, and (iii) Y is at least 97.
3. The expression cassette of claim 2, comprising
15 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:47,
(ii) X equals Y, and (iii) Y is at least 144.
4. The expression cassette of claim 3, comprising
20 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:49
or SEQ ID NO:97, (ii) X equals Y, and (iii) Y is at least 300.
5. The expression cassette of claim 4, comprising
25 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:49,
(ii) X equals Y, and (iii) Y is 2610.
6. The expression cassette of claim 4, comprising
30 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:97,
(ii) X equals Y, and (iii) Y is 2565.

WO 02/04493

PCT/US01/21241

7. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:51
5 (ii) X equals Y, and (iii) Y is 1494.

8. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:99,
10 (ii) X equals Y, and (iii) Y is 1491.

9. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:55;
15 SEQ ID NO:57; SEQ ID NO:101; SEQ ID NO:96; SEQ ID NO:134 or SEQ ID NO:135, (ii)
X equals Y, and (iii) Y is at least 60.

10. The expression cassette of claim 9, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
20 nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:55,
SEQ ID NO:57; SEQ ID NO:101; SEQ ID NO:96; SEQ ID NO:134 or SEQ ID NO:135, (ii)
X equals Y, and (iii) Y is 624.

11. An expression cassette comprising
25 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:58;
(ii) X equals Y, and (iii) Y is 354.

12. An expression cassette comprising
30 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:60;
(ii) X equals Y, and (iii) Y is 876.

WO 02/04493

PCT/US01/21241

13. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:62;
(ii) X equals Y, and (iii) Y is 3015.

5

14. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID
NO:103; (ii) X equals Y, and (iii) Y is 3009.

10

15. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:64
or SEQ ID NO:66; (ii) X equals Y, and (iii) Y is 297.

15

16. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:68,
(ii) X equals Y, and (iii) Y is 1965.

20

17. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:70;
(ii) X equals Y, and (iii) Y is 1977.

25

18. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:72
or SEQ ID NO:105, (ii) X equals Y, and (iii) Y is at least 30.

30

19. The expression cassette of claim 18, comprising

WO 02/04493

PCT/US01/21241

a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:72 or SEQ ID NO:105; (ii) X equals Y, and (iii) Y is 75.

5 20. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:74 or SEQ ID NO:107; (ii) X equals Y, and (iii) Y is at least 30.

10 21. The expression cassette of claim 20, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:74 or SEQ ID NO:107; (ii) X equals Y, and (iii) Y is 246.

15 22. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:76; (ii) X equals Y, and (iii) Y is 1680.

20 23. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:78; (ii) X equals Y, and (iii) Y is 1668.

25 24. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:80, SEQ ID NO:81 or SEQ ID NO:109; (ii) X equals Y, and (iii) Y is 216.

30 25. An expression cassette comprising

WO 02/04493

PCT/US01/21241

a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:83; (ii) X equals Y, and (iii) Y is 93.

- 5 26. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:111; (ii) X equals Y, and (iii) Y is 90.
- 10 27. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:85, or SEQ ID NO:113; (ii) X equals Y, and (iii) Y is 579.
- 15 28. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:87; (ii) X equals Y, and (iii) Y is 288.
- 20 29. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:115; (ii) X equals Y, and (iii) Y is 287.
- 25 30. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:89 or SEQ ID NO:117; (ii) X equals Y, and (iii) Y is at least 30.
- 30 31. The expression cassette of claim 30 comprising

WO 02/04493

PCT/US01/21241

a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:89; (ii) X equals Y, and (iii) Y is 267.

- 5 32. The expression cassette of claim 30 comprising
 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:117; (ii) X equals Y, and (iii) Y is 261.
- 10 33. An expression cassette comprising
 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:91; (ii) X equals Y, and (iii) Y is at least 30.
- 15 34. The expression cassette of claim 33 comprising
 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:91; (ii) X equals Y, and (iii) Y is 321.
- 20 35. An expression cassette comprising
 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:93 or SEQ ID NO:94; (ii) X equals Y, and (iii) Y is 309.
- 25 36. An expression cassette comprising
 a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:96; (ii) X equals Y, and (iii) Y is at least 60.
- 30 37. The expression cassette of claim 36 comprising

WO 02/04493

PCT/US01/21241

a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:96; (ii) X equals Y, and (iii) Y is 624.

- 5 38. An expression cassette comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:119; SEQ ID NO:120; SEQ ID NO:121; SEQ ID NO:122; SEQ ID NO:123; SEQ ID NO:124; SEQ ID NO:125; SEQ ID NO:126; SEQ ID NO:127; SEQ ID NO:131; SEQ ID NO:132 or SEQ ID NO:133, (ii) X equals Y, and (iii) Y is at least 60.

39. The expression cassette of claim 38, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:119; SEQ ID NO:120; SEQ ID NO:121; SEQ ID NO:122; SEQ ID NO:123; SEQ ID NO:124; SEQ ID NO:125; SEQ ID NO:126; SEQ ID NO:127; SEQ ID NO:131; SEQ ID NO:132 or SEQ ID NO:133, (ii) X equals Y, and (iii) Y is at least 300.

40. The expression cassette of claim 39, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:123 or SEQ ID NO:124, (ii) X equals Y, and (iii) Y is 2433.

41. The expression cassette of claim 39, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:122, (ii) X equals Y, and (iii) Y is 2301.

42. The expression cassette of claim 39, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID NO:125; (ii) X equals Y, and (iii) Y is 2517.

WO 02/04493

PCT/US01/21241

43. The expression cassette of claim 39, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID
NO:126 or SEQ ID NO:127, (ii) X equals Y, and (iii) Y is 2520.

5

44. The expression cassette of claim 39, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID
NO:119, (ii) X equals Y, and (iii) Y is 1377.

10

45. The expression cassette of claim 39, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID
NO:120 or SEQ ID NO:121, (ii) X equals Y, and (iii) Y is 1839.

15

46. The expression cassette of claim 39, comprising
a polynucleotide comprising X contiguous nucleotides, wherein (i) the X contiguous
nucleotides have at least 90% percent identity to Y contiguous nucleotides of SEQ ID
NO:132 or SEQ ID NO:133, (ii) X equals Y, and (iii) Y is 1890.

20

47. A polynucleotide comprising the sequence depicted in SEQ ID NO:33 or
fragments derived therefrom.

48. The polynucleotide of claim 47, wherein said fragments comprise coding
sequence for the gene products selected from the group consisting of Gag, Pol, Vif, Vpr, Tat,
Rev, Vpu, Env and Nef.

25

49. The polynucleotide of claim 48, wherein the fragment comprises a Gag gene
product.

30

50. The polynucleotide of claim 48, wherein the fragment comprises an Env gene
product.

WO 02/04493

PCT/US01/21241

51. The polynucleotide of claim 50, wherein the Env gene product is gp160, gp140 or gp120.

52. A polynucleotide comprising the sequence depicted in SEQ ID NO:45 or
5 fragments derived therefrom.

53. The polynucleotide of claim 52, wherein said fragments comprise coding
sequence for the gene products selected from the group consisting of Gag, Pol, Vif, Vpr, Tat,
Rev, Vpu, Env and Nef.

54. The polynucleotide of claim 53, wherein the fragment comprises a Gag gene
product.

55. The polynucleotide of claim 53, wherein the fragment comprises an Env gene
15 product.

56. The polynucleotide of claim 55, wherein the Env gene product is gp160, gp140 or
gp120.

57. A polynucleotide comprising the sequence depicted in SEQ ID NO:128 or
20 fragments derived therefrom.

58. The polynucleotide of claim 57, wherein the fragments comprise coding sequence
for Env gene products gp160, gp140 or gp120.

59. The expression cassette of any of claims 1 to 46, further comprising one or more
25 nucleic acids encoding one or more viral polypeptides or antigens.

60. The expression cassette of claim 59, wherein the viral polypeptide or antigen is
30 selected from the group consisting of Gag, Env, vif, vpr, tat, rev, vpu, nef and combinations
thereof.

WO 02/04493

PCT/US01/21241

61. The expression cassette of any of claims 1 to 46, further comprising one or more nucleic acids encoding one or more cytokines.

5 62. A recombinant expression system for use in a selected host cell, comprising, an expression cassette of any of claims 1 to 46, and wherein said polynucleotide sequence further comprises control elements capable of driving expression in the selected host cell.

63. The recombinant expression system of claim 62, wherein said control elements are selected from the group consisting of a transcription promoter, a transcription enhancer
10 element, a transcription termination signal, polyadenylation sequences, sequences for optimization of initiation of translation, and translation termination sequences.

64. The recombinant expression system of claim 62 wherein said transcription promoter is selected from the group consisting of CMV, CMV+intron A, SV40, RSV, HIV-
15 Ltr, MMTV-Ltr, and metallothionein.

65. A cell comprising an expression cassette of any of claims 1 to 46, and wherein said polynucleotide sequence further comprises control elements compatible with expression in the selected cell.
20

66. The cell of claim 65, wherein the cell is selected from the group consisting of a mammalian cell, an insect cell, a bacterial cell, a yeast cell, a plant, an antigen presenting cell, a primary cell, an immortalized cell, and a tumor derived cell.

25 67. The cell of claim 66, wherein the cell is selected from the group consisting of BHK, VERO, HT1080, 293, RD, COS-7, and CHO cells.

68. The cell of claim 67, wherein said cell is a CHO cell.

30 69. The cell of claim 66, wherein the cell is either *Trichoplusia ni* (Tn5) or Sf9 insect cells.

WO 02/04493

PCT/US01/21241

70. The cell of claim 66, wherein the antigen presenting cell is a lymphoid cell selected from the group consisting of macrophage, monocytes, dendritic cells, B-cells, T-cells, stem cells, and progenitor cells thereof.

5 71. A composition for generating an immunological response, comprising an expression cassette of any of claims 1 to 46.

72. The composition of claim 71, further comprising one or more *Pot* polypeptides.

10 73. The composition of claim 72, further comprising an adjuvant.

74. A composition for generating an immunological response, comprising an expression cassette of claim 52.

15 75. The composition of claim 74, further comprising a *Pot* polypeptide.

76. The composition of claim 74, further comprising one or more polypeptides encoded by the nucleic acid molecules of claim 60.

20 77. The composition of claim 76, further comprising an adjuvant.

78. A method of immunization of a subject, comprising,
introducing a composition of claim 71 into said subject under conditions that are compatible with expression of said expression cassette in said subject.

25 79. The method of claim 78, wherein said expression cassette is introduced using a gene delivery vector.

80. The method of claim 79, wherein the gene delivery vector is a non-viral vector.

30 81. The method of claim 79, wherein said gene delivery vector is a viral vector.

WO 02/04493

PCT/US01/21241

82. The method of claim 79, wherein said gene delivery vector is selected from the group consisting of an adenoviral vector, a vaccinia viral vector, an AAV vector, a retroviral vector, a lentiviral vector and an alphaviral vector.

5 83. The method of claim 82, wherein said gene delivery vector is a Sindbis-virus derived vector.

84. The method of claim 82, wherein said gene delivery vector is a cDNA vector.

10 85. The method of claim 82, wherein said gene delivery vector is a eukaryotic layered viral initiation system (ELVIS).

86. The method of claim 79, wherein said composition delivered using a particulate carrier.

15 87. The method of claim 79, wherein said composition is coated on a gold or tungsten particle and said coated particle is delivered to said subject using a gene gun.

20 88. The method of claim 79, wherein said composition is encapsulated in a liposome preparation.

89. The method of claim 79, wherein said subject is a mammal.

90. The method of claim 89, wherein said mammal is a human.

25 91. A method of generating an immune response in a subject, comprising:
providing an expression cassette of any of claims 1 to 46,

expressing said polypeptide in a suitable host cell,

isolating said polypeptide, and

30 administering said polypeptide to the subject in an amount sufficient to elicit an immune response.

WO 02/04493

PCT/US01/21241

92. A method of generating an immune response in a subject, comprising
introducing into cells of said subject an expression cassette of any one of claims 1 to
46, under conditions that permit the expression of said polynucleotide and production of said
polypeptide, thereby eliciting an immunological response to said polypeptide.

5

93. The method of claim 92, where the method further comprises co-administration
of an HIV polypeptide.

94. The method of claim 93, wherein co-administration of the polypeptide to the
subject is carried out before introducing said expression cassette.

10

95. The method of claim 93, wherein co-administration of the polypeptide to the
subject is carried out concurrently with introducing said expression cassette.

15

96. The method of claim 93, wherein co-administration of the polypeptide to the
subject is carried out after introducing said expression cassette.

20

97. The expression cassette of claim 59, wherein the viral polypeptide or antigen is
selected from the group consisting of polypeptides derived from hepatitis B, hepatitis C and
combinations thereof.

25

WO 02/04493

1/114

PCT/US01/21241

Gag_AFI10965_BW_mod

ATGGGCGCCCGGCCAGCATCCTGCGGGGGCCAGCTGGGACGCTGGGAGCGCATCCGCC
TGCCGCCCGCGGCCAAGAAGTGTACATGATGAAGCACCTGGTGTGGGCCAGCGCGAGCT
GGAGAGTTTCGCCCTGAAACCCCGGCTGTCTGGAGACCAAGCGAGGGCTGCAAGCAGATCATC
CGCCAGCTGCACCCCGCCCTGCAGACCGGCAGCGAGGAGCTGAGAGCCCTGTTCAACACCG
TGGCCACCCCTGTACTGCGTGCAAGAGAGATCGAGGTCTGGGACACCAAGAGAGGCCCTGGA
CAAGATCGAGGAGGAGCAGAACAGTGCAGCAGAAATCCAGCAGGCCGAGGCCCGCGAGC
AAGGGCAGGCTGAGCCAGAACTACCCCATCGTGCAAGAACCTGCAGGGCCAGATGCTGCACC
AGGCCATCAGCCCGCCGACCTGAAACGCTGGGTGAAGGTGATCGAGGAGAGAGGCCCTCAG
CCCCGAGGTGATCCCATGTTTCAACGCGCTGAGGAGGCGGCCACCCCGCAGGACCTGAAAC
ACGATGTTTGAACACCGTGGGGGGCCACAGGCGCGCATGCAAGTGTGAAGGACACCATCA
ACGAGGAGGCGCGGAGTGGGACCGCTGCAACCCCGTGCAAGCGGCCCATCGCCCGGG
CCAGATGGGAGAGCCCGCGGACGACATCGCGCGCACCAACAGCACCTGCAAGGAGCAG
ATCGCTGGATGACAGCAACCCCGCATCCCGTGGGGGACATCTACAGCGGTGGATCA
TCTTGGGCTGAAACAGATCGTGCGGATGTACAGCCCGGTGAGCATCTGTGACATCAAGCA
GGGCCCAAGGAGCCCTTCGGGACTACGTGGAACGCTTCTTCAAGACCTGCGCGCGAG
CAGAGCACCCAGGAGGTGAAGAACTGGATGACCGACACCCCTGCTGGTGAGAAAGCCAAAC
CCGACTGCAGACCATCTCGCGCTCTCGGCCCGCGCCAGCCTGGAGGAGATGATGAC
CGCTTGCCAGGGCGTGGGCGGCCAGCCACAGGCCCGCGCTGCTGGCCGAGCGATGAGC
CAGGCGACACCCAGCGTGATGATGACAGAGAGCACTTCAGGGGCCCGCGCGCATCTCA
AGTGTCTTACTGCGGCAAGGAGGGCCACATCCCGCGCACTGCGCGCCCGCCGCAAGAA
GGGCTGCTGAGAGTGGGCGAGGAGGGCCACCAAGATGAAGGACTGCACCGAGCGCCAGGGC
AACTTCTGGGCAAGATCTGGGCGACCAAGGGCGCGCCCGCAACTTCTGCAAGAGGC
GCCCCGAGCCCAACCGCCCCCGCGCGAGAGCTTCCGCTTCGAGGAGACCAACCCCGGCCA
GAAGCAGGAGAGCAAGGACCGGAGACCCCTGACACCTGAAAGAGCCTGTTGGGCAACGAC
CCCTGAGCCAGTAA

Figure 1

WO 02/04493

2/114

PCT/US01/21241

Gag_AF110967_BW_med.

ATGGGCGCCCGCCAGCATCTGCGCGGCGAGAGCTGGACAAAGTGGGAGAAAGATCCGGC
TGGCUCUCCGCGCGCAAGCACTACATGCTGAAGCACCCTGGTGTGGGCCAGCCGCGAGCT
GGAGGGCTTCGCCCTGAACCCCGGCTGCTGGAGACCGCCGAGGGCTGCAAGCAGATCATG
AAGCAGCTGCAGCCCGCCCTGCAGACCGGCAACGAGGAGCTGGCGAGCTGTATACACACCG
TGGCCACCTGTACTGCTGCACCGCGCATCGAGGTCCGCGACACCAAGGAGGCGCTGGA
CAAGATCGAGGAGGAGCAGAAAGTCCAGCAGAGAGCCACGAGGCCAAGGAGGCGGAC
GGCAAGGTGAGCCAGAACTACCCCATCTGTCAAGAACTGCAAGGCCAGATGGTGGACACAG
CCATCAGCCCGCCGACCCCTGAACGCTGGGTGAAGGTGATCGAGGAGAGGCTTCAGCCG
CGAGGTGATCCCATGTTCACCGGCTGAGCGAGGGCGCCACCCCGAGGACCTGAACACG
ATGTTGAACACCGTGGCGGGCCACAGGCGGCGCATGCAGATGCTGAGGAGACCATCAAG
AGGAGGCGCGGAGTGGGACCGCTGCACCCCGTGCAGGCGGCGCCCGTGGCCCGCGGCA
GATGCGCGACCCCGCGCGAGGACATCGCGCGGCGCCACGACACCTGCAGGAGCAGATC
GCTGATGACCAAGACCCCGCGTGGCGGCGGCGACATCTACAGCGGTGATCATTC
TGGGCTGAACAGATCGTGGGATCTACAGCCCGTGAACATCTGGACATCCGCGAGGG
CCCGAGGAGCCCTTCGCGACTACGTGACCGCTTCTTCAAGACCTGCGGCGCGAGCAG
GCCACCCAGGACGTGAAGAACTGGATGACCGAGACCTGCTGGTGCAGAACGCCAACCCG
ACTGCAAGACCATCTGCGCGCTCTGGCGCGGCGCCACCTGAGGAGATGATGACCGC
CTGCCAGGGGTGGCGCGCGCGCCACAGGCGCGGTGCTGGCGAGGCGATGAGCCAG
GCCACAGCGTGAACATCTGATGCAGAGGCAACTTCAAGGCGCCCGCGGCAACGTCA
AGTGTCTCACTGCGGCAAGGAGGGCCACATCGCCAGAACTGCGCGCGCCCGCAAGAA
GGGCTGCTGGAGTGGCGGCAAGGAGGGCCACAGATGAAGGACTGCACCGAGCGCCAGGC
AACTTCTTGGGCAAGATCTGGCCAGCCACAGGGCGCGCGCGCAACTTCTGCAAGAAC
GCAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
GGAGACCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
ACCGCGCTGGCGAGCCTGTTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG

Figure 2

WO 02/04493

3/114

PCT/US01/21241

Fig. 3

Env_AF110968_C_BW_opt

```

--> signal peptide (1-81)
ATGCCGCTGCTGTCGACATCTGAGAACTACCAAGCAGTGGTGGATGTGGGGCATCTGGGCTTCTGGATGCTGATCA
V--> gp120/140/160 (82)
TCAGCAAGTGTGGGCAACCTGTGGGTGACCGTGTACTACGCGCTGACCGTGTGGAGGAGGCCAAGACCAACCT
GTTCTGCACACGACGCGAAGSCTACGAGACGAGGTGCACACGTTGTGGCCACCCACGCTGCGTGGCCACCC
GACCCCAACCCAGGAGATCGTCTGGAGAACCTGACCGAGACTTCACACATGTGGAGAGAGGACATGGTGGAGCC
AGATGACACGAGGACATCATCAGCTGTGGGACACAGCCTGAAGCCCTGGCTGAAGCTGACCCCTGTGGCTGAC
CCTGAATGCTGCCAAGCTGAACGCCACCAACACATCAACAGCATGATCGACACAGCAACAGGGGAGAGTGAAG
AACTGCAGCTTCACCTGACACCGAGCTGCGGACCGCAAGCAGGAGGTGACCGCCCTGTCTACCGCTGGACG
TGGTGGCCCTGCAAGGCAACACAGCAGAGTACCGCTGATCAACTGCAACACGAGGCGCATCACCCAGGCTG
CCCAAGCTGAGGTTGAGACCCATCCCATCTCACTACTGCAACCCCGCGGCTACGCGCATCTGAAGTGCACACAC
CAGACCTTCAACGGCACCGCCCTGCAACACGTTGAGCAGCGTGGAGTGGCCACCGGCATCAAGCCGCTGGTGA
GCACCGAGCTCTGCTGACGGCAGCTGTGGCAAGGGCGAGATCATCATCCGACCGAGAACCTGGCTCAACACG
CAAGTATCATCATCTGTCAGCTGACACAGCCCGTGAAGATGTGTGCTGGCCCAACCAACACCCGCAAGAG
CTGCGCATCTGGCCCGCCGAGACCTTCTAGCCACCGCGGAGATCATCTGGCGACATCCGCGAGGCTACTGCTCA
TCACACAGACCGAGTGAACACGACCCCTGCAAGGGCTGAGCAGAGCTGGAGGAGCTTCAAGCAAGAGCCAT
CAGGTTCGAGCCACGACGCGCTGGCGACTTGGAGATCACCACTCACAGCTTCACTGCGCGGCGAGTTCTTCTAC
TGCGACACCGCCAGCGCTGTTCACAGCACCTACAGCCCGCTTCAACGGCACCTGAACACAGCTGAACGGCACCA
TCACCATCACTGCGCATCAAGCAGATCATCAACATGTGGCAGAGGCTGGCCGCGCATGTACGCGCCCGCAT
CGCGGCAACCTGACCTGCGAGAGCAATCAACCGGCTGCTGCTGACCCGCGCGCGGCGAGACCGCGCCCAAC
GACACCGAGATCTTCCGCGCGCGCGCGCGCATGTGGCGACACCTGGCGCAACGAGCTGTACAAAGTCAAGCTGG
TGGAGATCAAGCCCTGGGCTGGCCCGCCACCGAGGCGCAAGGCGCGGCTGGTGGGCTGGAGAGCCGCTGGTGG
CATCGGCGGCTGTCTGCTGGCTTCTTGGGCGCGCGCGGCGAGCAACATGGGCGCGCGCAGCATCACCTGACCGTG
CAGCGCGCTGCTGCTGAGCGGCTGCTGCGAGCAGACACACCTGCTGGCGGCTATCGAGGCGCGAGCAGCACC
TGCTGCACTGACCTGTGGGCTACAGCAGCTGCAAGCCCGCATCTTGGCTGGAGGCTACCTGAGAGACCA
GGAGCTGCTGGGCTGTGGGCTGCGAGCGGCAAGCTGATCTGCAACACCGGCTGGCTGGAGAGCAGCTGAGC
AACCGCAGCCACAGAGATCTGGACACATGCTGGATGCACTGGAGCGAGCGAGATCAACAACTACACGAGCA
CCATCTACCGCTGCTGGAGAGAGCCACAGCAGAGAGAGACAGAGAGAGCTGCTGGCCCTGGAGAGCTG
GCAGAACCTGTGAACGCTGCTGAGCATCAACACTGCTGTGTACATCAAGATCTTCACTATGATCTGGGCGGCG
CTGATCGGCTGCGCATCATCTTCCGCGTGTGAGCATCTGTAACCGGCTGGCGAGGCTACAGCCCTTGCCT
TCCAGACCTGACCCCAACCCCGAGAGCCGCGCTGGGCGCATGAGGAGAGGGGCGGAGAGGAGCCG
CGGCGCAGCATCGGCTGCTGAGCGCTTCTTGGCCCTGGCTGGAGACCTGCGAGGCTGTGCTGTTCAGC
TACCAACCGCTGCGGACTTCATCTGATGCTGCGCGGCTGCTGGAGCTGCTGGCGAGCGGCTGGAGGCGCC
TGAATGATCTGGGAGCCCTGGTGGAGTCTGGGCTGGAGCTGAAGAGAGCGCATCAGCTGCTGGAGACCAT
CGCATCGGCTGGCGGAGCGCGAGCCGATCATGAGTTCATCCAGGCGATCTGCGCGCATCTCGAGACATC
gp160, gp41 (8547) <--\
CCCCCGCATCCGACAGGCTTCAGAGCGCGCTGCACTRA

```


WO 02/04493

4/114

PCT/US01/21241

Fig. 4

Env_AF110975_C.BW_opt

```

--> signal peptide (1-72)
ATGCGCGTGGGGGATCTTGGGCGAGCTGGCAGAGTGGTGGATCTGGGGCATCTGGGCTTCTGGATCTGGAGCG
gp120(140/160) (72)
GCTTGGGCACTCTGCTGGTGGAGCGGTGTACGACGGGTGGGCGGTGGTGGCGCGAGGCGCAGACCACTGTTCTGGCG
CAGCGAGCGCCAGGCTACGAGAGAGAGGTGCACAACTGTGGGCGCAGCGCTGGCTGGCCAGCGACCGCAAC
CCCCAGGAGTCTGGAGCTGGACAACTGTACCGAGACTTCAACATGTGGAGAACGACATGGTGGAGCAGATGCAGC
AGGACATCATCTAGCTGTGGGACCGAGAGCTGAAGCGCGGTGGAGCTGACCGCTGGCTGGAGCTGGAGTGAAGT
CACCACTACAGGACCACTACAGCAACCAATGAGCGCGACCGACTACACAGCAGAGACCGAGGAGATCAAG
AGCTGACCTTCACAGAGACCGAGAGCTGGGCGACAGAGAGCAGAGGTGTACGCGCTGTCTTCAGAGCTGGACA
TCTGCGCTGTACAGCAGACAGCGAGTACCGAGTACCGCTGTACTGCAACACAGCGCGCTACCGAGCGCTGGCG
CAAGGTGAGCTTGGAGCCATCCCATCTCACTAGCGCGCGCGCGCTAGCGCATCTGAGTGAAGTGAAGACAGC
ACAGCTAGCGGACCGCGCGCTGGCGAGAGCTGAGCAGCGCTGAGTGCACCCAGCGCATCAAGCGCGTGGTGAAG
CCCCCTGCTGTGAGAGCGAGCTGGCGAGGCGCGAGATCATCTCCGAGCAGAGACCTGAGCAACAGCGC
CTACAGCTCATCTGTGCACTGAGCAGAGCTGGAGATCTGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG
ATCGCATCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
TCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
CGAGTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
TGCTACAGCAGCAGCTGTGTACAGCAGCTACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
TCACCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
CGAGGAGCAGCTACCTGCGAGCAGCAGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
ACCGAGATCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
AGATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
TGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
GCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
AAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
TCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
CAACCTGTGGAGCTGTCTACAGTACAGCACTGGCTGTGGTACATCAATCATGATGTGGCGCGCGCG
ATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
CGCGAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
CAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
CGCGCTGTGGAGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
CGCGCTTCTGCAACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG

```

Figure 5

Figure 6

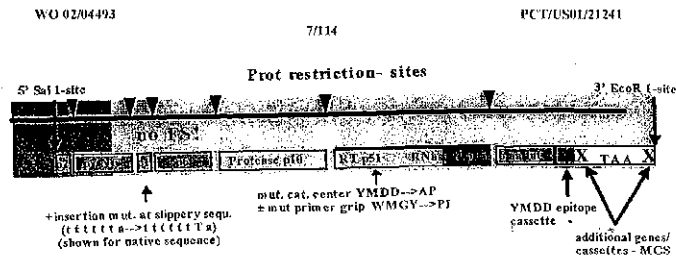


FIGURE 7

WO 02/04493

8/114

PCT/US01/21241

PR975(+) (SEQ ID NO:30)

GTGCAAGCCACCATGCGGAGGCCATGAGCCAGGCCACAGCGCCAACTCCTGAT
 GCAGCGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTTAACTGCGGCAA
 GGAGGGCCACATGCGCGCAACTGCGCGCCCCCGCAAGAAAGGGCTGCTGGAAGT
 GCGGCAAGGAGGGCCACCATGAGTGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTC
 CGCGAGGACCTGGCTTCCCCAGGGCAAGGCCCGGAGTTCCCGAGCGAGCAGAA
 CCGCGCAACAGCCCAACAGCCCGAGCGGAGCTGCAGGTGCGCGGCGCAACCCCGCA
 GCGAGCGCGCGCGAGCGCGAGGGCACTTGAAGTTCCCGAGATCACTCTGTGGC
 AGCGCCCCGTGTGAGCATCAAGGTGGCGGCGAGATCAAGGAGGGCCCTGCTGGAC
 ACCGGCCCGACGACACCGTGTCTGGAGAGATGAGCTGCGCGCAAGTGAAGCC
 CAAGATGATCGCGGGCATCGCGGCTTCATCAAGGTGCGCGAGTACGAGCCAGATCCT
 GATCGAGATCTGCGGCAAGAAAGGCCATCGCGCACCGTGTGATCGCGCCCAACCCCGT
 GAACATCATGCGCGCAACATGCTGACCGAGCTGGGCTGCACTGAACTTCCCAT
 CAGCCCCATCGAGACCGTGGCGTGAAGCTGAAGCGCGCATGGAAGGCCCAAG
 TGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAG
 GAGATGAGAGAGAGGGCAAGATCACCAAGATCGGCCCCGAGAAACCTTACCAACAC
 CCGCGTGTTCGCCATCAAGAAAGAGGACAGCACCAAGTGGCGCAAGCTGTGGACT
 TCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGCACTGGGCATCCCC
 AGCCCGCGCGCTGAAGAAAGAAAGAGCGTGAACGTGCTGGAAGTGGCGAGCGC
 TACTTCAGGTGCGCTGGAGGAGACTTCCGCAAGTACACCGCTTACCATCCCC
 AGCATCAACAAAGAGACCCCGCATCGCTACCAAGTACCAAGTCTGCGCCAGGGC
 TGGAAAGGGCAGCGCCAGCATCTTCAAGAGCAGCATGACCAAGATCTGGAGCCCTTC
 CGCGCCCGCAACCCCGAGATCGTGTATCTTACAGTACATGGACGACCTGTACGTGGC
 AGCGACCTGGAGATCGCGCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCACT
 GCTGCGCTGGGCTTACCAACCCCGCAAGAAAGCACCAGAAAGGAGCCCTTCTCT
 GTGGATGGGCTACGAGCTGCAACCCGCAAGTGGACCGTGCAGCCCATCGAGCTGCC
 CGAGAGGAGAGCTGGACCGTGAACGACATCCAGAAAGTGGTGGCAAGCTGAACT
 GGGCCAGCCAGATCTACCCCGCATCAAGGTGCGCGAGCTGTGCAAGCTGTGCGCG
 GCGCCAGGCGCTGACCGACATCGTGGCGCTGACCGAGGAGCGCGAGCTGGAGCTG
 GCGGAGAACCGCGAGATCTGCGCGAGCGCTGCACGGCTGTACTACGACCCAG
 CAAGGACCTGGTGGCGAGATCCAGAGCAAGGCGCAGCAAGTGGACCTACCGA
 TCTACAGGAGCCCTTCAAGAACTGAAGACCGGCAAGTACGCCAAGATGGCAAC
 GCGCACCAAGAGCTGAAGCAGCTGACCGAGGCGCTGCAAGATGGCCATGGA
 GAGCATCGTGTATCTGGGGCAAGACCCCAAGTTCCCGCTGCCATCCAGAAAGGAG
 CTGGGAGACTGTGGACCGACTACTGGCAGGCCACTGGATCCCGAGTGGGAGTT
 CGTGAACACCCCGCTGTGTAAGCTGTGTACCAGCTGGAGAAAGGAGCCATCAT
 CGGCGCGAGACCTTCTACGTGGACGCGCGCGCAACCGGAGACCAAGATCGGCA
 AGCGCGCTACGTGAACCGAGCGCGCGCGCAGAAAGATCGTGAAGCTGACCGAGAC
 ACCAAACGAGAGACCGAGCTGCAGGCCATCCAGCTGGCGCTGCAGGACAGCGGCA
 CGAGGTGAACATCGTGACCGACAGCGATACCGCTGGGCATCATCCAGGCCAGCC
 CGACAAGAGCGAGAGCGAGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAAAG
 AGAAGGTGTACCTGAGCTGGTGGCGCGCCACAAGGGCATCGCGCGCAACGAGCAG
 ATCGACAAGCTGGTGAAGCAAGGGCATCGCAAGGTGCTGTCTGGAAGCGCATCGAT
 GCGCGCATCGTGTATCTACCAAGTACATGGAGCAGCTGTACGTGGGAGCGCGCGCC
 AGGATCGATTAAAGGCTTCCCGGGCTAGCACCGTGAATTC

FIGURE 8

WO 02/04493

9/114

PCT/US01/21241

PR975YM (SEQ ID NO:31)

GTCGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCAGGCCAACATCCTGAT
 GCAGCCGACCACTTCAAGGGGCCAAGGCGATCATCAAGTGCTTCAACTGCGGCAA
 GGAGGGCCACATCGCCCGCAACTGCCCGCCCCCGCAAGAGGGCTGCTGGAAGT
 GCGGCAAGGAGGGCCACCAATGAAGGACTGCACCGAGGCCAGGCCAACTTCTTC
 CGCGAGGACCTGGCCCTTCCCCAGGGCAAGCGCCGCGAGTTCCCGCAGGACAGAA
 CCGCGCCAACAGCCCCACCGCCGAGCTGCAGGTGCGGGCGGCAACCCCGCA
 GCGAGGCCGCGCGAGGCCAGGGCACCTGAACTTCCCGCAGATCACCTGTGGC
 AGCGCCCTGTGTGACATCAAGGTGGCGGCCAGATCAAGGAGGGCCCTGTGGAC
 ACCGGCCCGACGACACCGTGTGGAGGAGATGAGCCTGCCCGCAAGTGGAAAGCC-
 CAAGATGATGCGGCGCATGGCGGCTTCATCAAGGTGCGCCAGTACGACCAAGATCCT
 GATCGAGATCTGGCGCAAGAAGGCCATCGGCAACGTGTGATGGGCCACCCCGT
 GAAACATCATGGCCGCAACATGCTGACCCAGCTGGGTGACCCCTGAACTTCCCAT
 CAGCCCATCTGAGACCGTGGCGTGAAGCTGAAGCCGGCATGGACGGGCCCAAGG
 TGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCGAG
 GAGATGGAGAAGGAGGCCAAGATCACCAAGATGGCCCGAGAAACCCCTACAACAC
 CCCCCTGTTCGATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT
 TCCCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGCATCCCC
 ACCCGCGCGCTGAAGAAGAAGAAGAGCGTGAACGTGTGGACGTGGCGACGCC
 TACTTCAGCGTCCCTTGGACGAGGACTTCGCAAGTACACCGCTTCAACATCCCC
 AGCATCAACAACGAGACCCCGGCATCCGCTACCAAGTACAACTGTGCTGCCAGGGC
 TGGAGGGCAGCCCGACATCTTCCAGAGCAGCATGAACCAAGATCTTGGAGCCCTTC
 CGCGCCCGCAACCCCGAGATCTGTGATCTACCAAGGCCCTGTACGTGGGACGCGAC
 CTGGAGATGGGCCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCACCTGTGCG
 CTGGGCTTCAACACCCCGCAAGAAGCACCAAGAGAGCCCGCTTCTGTGGAT
 GGCTACGAGCTGCAACCCGACAAAGTGGACCGTGCAGGCCATCGAGCTGCCGAGA
 AGGAGAGCTGGAACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCC
 AGCCAGATCTACCCCGCATCAAGGTGGCCAGCTGTGCAAGCTGTGCGCGCGCC
 AAGGCCCTGACCGACATCGTGCCTGACCGAGGAGGCCGAGCTGGAGCTGGCCGA
 GAACCGCGAGATCTTGGCGAGCCCGTGCACGGCGTGTACTACGACCCAGCAAGG
 ACCTGCTGGCGAGATCCAGAAGCAGGGCCACGACCAAGTGGACCTACCAAGATCTAC
 CAGGAGCCCTTCAAGAACCTGAAGAAGCGCAAGTACGCCAAGATGGCGACCCGCCA
 CACCAACGACGTGAAGCAGCTGAACGAGGCCGTGCAGAAGATGGCCATGGAGAGCA
 TCGTGTATCTGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAAGGAGACCTGGG
 AGACCTGTGTGACCGACTACTGCGAGGCCACCTGGATCCCGAGTGGGAGTTCTGTGA
 ACAACCCCGCTGTGGAAGCTGTGTACCAAGCTGGAGAAGGAGCCATCATCGGCG
 CCGAGACCTTCTACCTGGAAGGGCGGCCAACCGCGAGACCAAGATCGGCAAGGCC
 GGCTACGTGACCGACCGGGCGCGGCAAGATCGTGAGCCTGACCGAGACCAACCA
 CCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGACGAGG
 TGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCAGCCGACA
 AGAGCGAGAGCGAGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAA
 GTGTACCTGAGCTGGTGGCCCGCCACAAGGGCATGGGGCAACGAGCAGATCGA
 CAAGCTGGTGAAGCAAGGGCATCGCAAGGTGTCTTCTGGACGGCATCGATGGCG
 GCATCGTGTATCTACCAAGTACATGGACGACCTGTACGTGGGACGGGGCGCTAGGA
 TCGATTAAAGCTTCCCGGGCTAGCACCGGTGAATT

FIGURE 9

WO 02/04493

10/114

PCT/US01/21241

PR975YMWM (SEQ ID NO:32)

GTGACGCCACCATGGCCGAGGCCATGAGCCAGGCCACCGGCCAACATCCTGAT
 GCAGCGCAGCAACTTCAAGGGCCCAAGGCCATCATCAAGTGCTTCAACTGGGGCAA
 GGAGGGCCACATCCCGGCAACTGCGCGCCCGCGCAAGAGGGCTGCTGGAAGT
 GCGGCAAGGAGGGCCACAGATGAAGGACTGCAACCGAGGCGCAGGCCAACTTCTT
 CGCGAGGAOCTGGCCTTCCCCAGGGCAAGGCCCGGAGTTCCCGAGGAGCAGAA
 CCGCGCCAAACAGCCCAACAGCGCGGAGCTGCAAGGTGGGCGGACAAACCCCGCA
 GCGAGCGCGGCGCGAGCGCCAGGGCAACCTGAACTTCCCGCAGATCAACCTTGTGGC
 AGCGCCOCTGTGTGAGCATCAAGGTGGGCGGCGCAGATCAAGGAGGCCCTGTGGAC
 ACGGCGCGGACGACACCGTGCTGGAGGAGATGAGCCTGCCCGGCAAGTGGAGGCC
 CAAGATGATCGGGCGCATGGGCGGCTTCATCAAGGTGGCGCAGTACGACCCAGATCCT
 GATCGAGATCTGGCGCAAGAGGCCATCGGCACCGTGCTGATGGGCGCCACCCCGT
 GAACATCATGGCGCAACATGCTGACCCAGCTGGGCTGCAACCGGCGATGGAGCGCCAAAG
 CAGCCCATCGAGACCGTGCGGTGAAGCTGAAGCCCGGCGATGGAGCGCCAAAG
 TGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGGAG
 GAGATGGAGAAGGAGGGCAAGATCAACCAAGATCGGCCCGGAGAACCCCTACAAC
 CCGCGTGTTCGCGATCAAGAGAGGACACGACCAAGTGGCGCAAGCTGGTGGACT
 TCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC
 ACCCGCGCGGCTGAAGAGAGAGAGAGCGTGACCGTGCTGGAGCTGGGCGACGCC
 TACTTACGCTGCCCTGGAGGAGGACTTCGCAAGTACACCGCTTACCATCCCC
 AGCATCAACAACGAGACCCCGGCGATCGCTACAGTACAACGTGCTGCCCGAGGGC
 TGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGAACCAAGATCTGGAGCCCTTC
 CGGCGCCGCAACCCCGAGATCGTGATCTACAGGCCCGCTGTACGTGGGCGAGCGAC
 CTGGAGATCGGGCAGCACCGCGCAAGATCGAGGAGCTGGCGAAGCACCTGTGCG
 CTGGGGCTTACACCCCGGACAGAGCAACAGCACCGAGCCATCGAGCTGCCCGAGAAGGAG
 CGAGCTGCACCCCGACAAGTGGACCGTGCGAGCCATCGAGCTGCCCGAGAAGGAG
 GCTGGACCGTGAACGACATCCAGAGCTGGTGGGCAAGCTGAAGTGGGCGAGCCAG
 ATCTACCCCGCATCAAGGTGGCGCAGCTGTGCAAGCTGCTGCCGCGCGCAAGGCC
 CTGACCGACATCGTGCCCTGACCGAGGAGGCGGAGCTGGAGCTGGCGGAGAACCG
 CGAGATCTGGCGGAGCCCGTGACCGCGGTGTACTACGACCCAGCAAGGACCTGGT
 GGCGGAGATCCAGAGCAGGGCCACGACCAAGTGGACCTACCAAGATCTACAGGAGC
 CCTTCAAGAACCTGAAGACCGGCAAGTACGCAAGATGGCGACCGCCACACCAAC
 GACGTGAAGCAGCTGACCGAGGCGGTGCAGAGATCGCCATGGAGAGCATCGTGAT
 CTGGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAGGAGACCTGGGAGACCT
 GGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTCGTGAACACCC
 CCCCCCTGGTGAAGCTGTGGTACCAAGCTGGAGAAGGAGGCCATCATGGGCGCGAG
 ACCTTCTACGTGGAGCGGCGCGGCAACCGCGAGACCAAGATCGGCAAGGCCGCTA
 CGTGACCGACCGGGCGCGGCAAGATCGTGAGCTGACCGAGACCAACCAACAG
 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGCGAGGAGGTGAAC
 ATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCAGCCGACCAAGAG
 CGAGAGCGAGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGT
 ACCTGAGCTGGTGGCGCCGCAAGGGCATCGGCGGCAACGAGCAGATCGACAAG
 CTGGTGAGCAAGGGCATCGGCAAGGTGCTGTCTTCTGGACGGCATCGATGGCGGCATC
 GTGATCTACCAAGTACATGGACGAGCTGTACGTGGGCGAGCGCGGCCCTAGGATCGAT
 TAAAAGCTTCCCGGGCTAGCACCGGTGAATTC

FIGURE 10

WO 02/04493

11/114

PCT/US01/21241

8_5_2A (SEQ ID NO:33)

1 TGAAGGGTT AATTACTCC AAGAAAGGC AAGAAATCCT TGATTGTGG GTCTATCACA
 61 CACAGGGCTT CTTCCTGAT TGGCAAACT ACACCCGGG GCCAGGGTC AGATATCCAC
 121 TGACCTTTGG ATGGTGTAC AAGCTATGC CAGTTGACC CTATGAGCA ACATGGAGCA GAGGATGAG
 181 ACGGAGGAGA AGACAACTGT TTGCTACAC GCTTCTAGC AGGCAAGAC ATGGCCCGG
 241 ATAGAGAGGT ATTAAAGTG AAGTTT9ACA GCTTCTAGC AGGCAAGAC ATGGCCCGG
 301 AGCTACATCC GGAGTATTAC AAGAGCTGCT GACACGAGG GGCCTTTCCG CTTGGGACTT
 361 TCCTAGTGGG CGTCCGGGA GGTGTGGCT GGCCTTTCCG TGCGAGTGGT CACCCCTCAG
 421 ATGCTGCATA TAAGCAGCTG CTTTTCGCT GTACTGGTC TCTCTGGTA GACGAGTCT
 481 GAGCTTGGGA GCGCTCTGGC TATCTAGGA ACCACTGCT TTAGCCTCAA TAAAGCTTC
 541 CTTGAGTCTT TTAGTAGTG TGTGCCATC TGTGTGTGA CCGCTTCTG CCGCCGACA GAGGACGAA
 601 TCAGACCTTT TGTGTAGTG TGGAAATCT CTAGCAGTG CCGCTTCTG CCGCCGACA GAGGACGAA
 661 AGTGAAGTG AGACGAGAG AGATCTCTG AGCAGGACT CCGCTTCTG CCGCCGACA GAGGACGAA
 721 CGCAGAGGC GAGGAGGAG GCTGTGAGT AGCCCAATT TACTTGAAT GCGAGGACTA
 781 GAGGAGAGA GATGGTGTG AGAGCTCTA TATTAGGAG CCGAAATTA GATAAATGG
 841 AAGCAATTAG GTTAAGGCA GGGGCAAGA AACCTTATAT GTTAAACAT CTAGTATGG
 901 CAGCAGGGA GCTGGAAGA TTGTGACTA ACCCTGAGT GTTAAACAT CTAGTATGG
 961 GTAAACAAAT AATAAAGAG CTACACAGG CTCTTCAGC AGGACAGAG GAATTTAGT
 1021 CATTATTCAG CACATAGCA ACTCTCTAT GTGTACATA AGGATAGAG GTAGGAGCA
 1081 CCAAGGAGC CTTAGAGAG ATAGAGGAG AACAAACAA ATGTACAGA AAGGACACG
 1141 AGCAGAGAG AGCTAGAGA AAGGCTAGT AATATTATCC TATATAGAG AATGCCAGG
 1201 AGCAAAAGT ACACCAAGT ATATCACTA GAACATTGAA TGCAATGATA AAGTATAG
 1261 AGCAAAAGC TTTCATCCA GAGGAATAC CCAATTTTAC AGCATTATCA GAGGAGCCA
 1321 CCGCAAGAG TTAAACACA ATGTAAATA CAGTGGGGG ACATCAAGCA GCCATGCAA
 1381 TGTAAAGTA TACCATCAAT GAGGAGGCT GAGATGGGA TAGGACACAT CCAATACAT
 1441 CAGGCTCTG TGCAACAGG CAGATGAGG AACCAAGGC AAGTACATA GAGGAGCTA
 1501 CTGCTCTCT TCAGGACAA ATAGCATGGA TGCAAGTAA TCCACTATT CCAATAGAG
 1561 ACATCTATAA AAGATGATA ATTCTGGGT TAAATAAAT AGTAAAGAT TATAGCCCTG
 1621 TTAGCTATT GGCATATAA CAGGAGGCA AAGACCCCT TAGGAGTAT GTAGCCGGT
 1681 TCTTAAAC CTTAGAGCT GAACAGCTA CAGAGATCT AAGAAATGG ATGACAGCA
 1741 CTTCTCTCT CCAAAATCCG AACCCAGAT GTAGAGCAT TTAGAGGCA TTAGAGCCG
 1801 GGGCTCTAT AAGAGAGAT ATGACAGCT GTAGAGCAT GGGAGGACT AGCCATTAAG
 1861 CAGAGCTTT GCTGAGGCA ATGACAGCT CAAACAGTA CTAATAGTG CAGAGAGCA
 1921 ATTAAAGG CTTACAGCA ATTATTAAAT GTTCAACTG TGGCAAGTA GGGACATAG
 1981 CAGGAGGCT CAGGAGGCT AAGGAAAGG GCTGTGGGA ATGTGAGAG GAGGAGACC
 2041 AAGTAAAGA CTGTACTAG AGGCAAGCA GACCAAGGC AACAGCCCA CAGGAGAAC
 2101 AGGGAGGCC AGGGAATTC CTCCAGAAC GACCAAGGC AACAGCCCA CAGGAGAAC
 2161 CAGGAGGCC ACCAGCAGG AGCTTCAAGT AATCACTCT TGGCAGGAG CCGTGTCTC
 2221 AAGAAAGAG GGAACCTTTA ACTTCCCTC TCTCTTAGC AGGAGAGAG ATGATAGCT
 2281 AATAAAGTA GAGGAGGCA TAAAGGAGG AATCACTCT TGGCAGGAG CCGTGTCTC
 2341 ATTAGAGAA ATAGATTTC CAGGAGAGT GAAACCAAA ATGTAGGGG GAATAGGAG
 2401 TTTATCAAA GTAGACAGT ATGATCAAT ACTTATAGA ACTGTAGAG TAAATTAAG
 2461 AGGTACAGTA TTAGTAGGC CTACACAGT CAGCATATG ACTGTAGAG TAAATTAAG
 2521 GCTTGGATG ACATTAATT TTCCATATG TCTATTTGA ACTGTAGAG TAAATTAAG
 2581 ACCAGGAGG GATGGCCCA AGGTCAACA ATGGCCATT GACAGAGAA AATTAAGG
 2641 ATTACAGCA ATTTGTAGG AATGAGCAA GAGGAGAAA ATTACAGAA TTGGGCTGT
 2701 TATCCATAT AACACTCCG TATTTGCTT AAAAAGAGG GACATATG ACTGTAGAG
 2761 ATTATAGAT TTGAGGAGC TCAATTAAG AACPCAGAC TTTTGGAGG TCAATTAAG
 2821 AATACACAC CAGCAGGAT TAAAAAGAA AATATCAGT ACAGTGTAG ATGTGGGGA
 2881 TGCAATTTT TCAATCTCT TAGATGAAG CTTCAGGAA TACTGTGAT TCACATACC

FIGURE 11

WO 02/04493

12/114

PCT/US01/21241

2941 TATGTATATAC AATGAACAC CAGGGAITAG ATATCAATAT AATGTGCTGC CACAGGGATG
 3001 GAAAGGATCA CCAGCAATAT TCCAGAGTAG CATGACAAA ATCTTAGAGC CCTTAGAGGC
 3061 AAAAATATCCA GACATAGTTA TCTATCAATA TATGGATGAC TGTATATGAG GATCTGACTT
 3121 AGAATATGGG CAACATAGAG CAATAATAGA AGAGTTAGG CAACATTTAT TGAATGGGG
 3181 ATTTACACCA CCAGACAGA AACATCAAAA AGAACCCCA TTTCTTTGGA TGGGTATGTA
 3241 ACTCCATCCT GACAAATGGA CAGTACAACC TATACTGCTG CCAGAAAAGG ATAGTTGGAC
 3301 TGTCAATGAT ATACAGAGGT TAGTGGGAAA ATTAAACTGG GCAAGTCAGA TTTACCCAGG
 3361 GATTAAGATA AGGCACTCTT GTTAACTCCT CAGGGGGGCC AAAGCACTAA CAGACATAGT
 3421 ACCACTACTT GAGAGAGCAG AATTAGAAAT GCGAGAGAAC AGGGAAATTT TAAGAGAAC
 3481 AGTACTATGA GTATATTATG ATCCATCAAA AGACTTGATA GCTGAATAC AGAAACAGGG
 3541 GATGTACAAA TGGACATATC AAATTATATCA AGAACCAATT AAAAATCTGA AAACAGGGAA
 3601 GTATGTACAAA ATGAGTACTA CCCACACTAA TGTATGTAAA CAGTTACAG AGGCAGTGCA
 3661 AAAAATATGCC ATGGAAGACA TAGTAAATATG GGGAAAGACT CCTAAATTTA GACTACCCAT
 3721 CCAAAAAGAA ACATGGGAGA CATGTGGGAC AGACTATTGG CAGCCACCTT GATCCCTGA
 3781 GTGGAGATTI GTTAATACCC CTCCCTAGT AAATATATGG TACCAACTGA AAAAAGATCC
 3841 CATAGCAGGA GTAGAAACTT TCTATGTAGA TGGAGCAACT AATAGGGAG CTAATATAGG
 3901 AAAAGCAGGG TATGTTACTG ACAGAGGAAAG GCAAAAATTT GTTACTCTAA CTAACACAC
 3961 AATACAGAGG ACTGAGTTAC AAGCAATTCA GCTAGCTCTG CAGGATTCAG GATCAGAGT
 4021 AACTATAGTA ACAGACTCAC AGTATGCAAT AGGAATCTAT CAGACACAC CAGATAGAG
 4081 TGACTCAGAG ATATTTAACC AAATAATAGA ACAGTAAATA AACAGGAAA GATTTACCT
 4141 GTCATGGGTA CCAGCACATA AAGGAATGG GGGAAATGAA CAGTAGATA AATTAGTAG
 4201 TAGGGGAAAT AGCAAAATGT TGTTCCTAGA TGGAATAGAT AAGCTCAGG AAGAGCATGA
 4261 AAGGTACCAC AGCAATTTGA GAGCAATGGC TAATGAGTTT AATCTGCCAC CCAATAGTAG
 4321 AAAAAGAAATA GTAGCTAGCT GTGATAATG TCAGCTAATA GGGGAAGCCA TACATGAGCA
 4381 AGTCAGCTGT AGTCAGGGA TATGGCAATT AGATTTGACC CATTTAGAGG GAAAAATCAT
 4441 CCTGTAGACA GTCCATGTAG CTAGTGGCTA CATGGAGACA GAGTTATATC CAGCAGAAC
 4501 AGGACAGAAA ACAGCATATT TTATATTAAA ATTAGCAGGA AGATGGCCAG TCAAAATAT
 4561 ACATACAGAC AATGGCAGTA ATTTTACCAG TACTGCCAGT AAGGCAGGCT GTTGGTGGCT
 4621 AGGTATCCAA CAGGAATTTG GAATTCCTTA CACTGCCAAA AGTCAGGGAG TGGTAGATC
 4681 CATGAATATA CAATTAAGGA AATTAATAGG ACAAGTAAGA GATCAGCTG AGCACTTAA
 4741 GACAGCAGTA CAATGGCAG TATTCAATCA CAATTTTAAA AGAAAGGGGG GAATTTGGGG
 4801 GTACAGTGCA GGGGAAGAAA TATAGACAT AATAGCAACA GACATACAAA CTAAAGAAAT
 4861 ACAAAACAAA ATTATAAGAA TTCAAAATTT TCGGTTTAT TACAGAGACA GCGAGAGCCC
 4921 TATTTGGAAA GGACCAAGCG AACTACTCTG GAAAGGTGAA GGGGTAGTAG TAATAGAGA
 4981 TAGAGGTGAC ATAAAGGTAG TACCAGGGAG GAAGGCAAAA ATCTTAGAG ATTAGGAAA
 5041 ACAGATGSCA GGTGCTGATT GTGTGGCAGG TGGACAGGAT GAAGATTAGA GCAATGAAATA
 5101 GTTTAGTAAA GCACTATATG TATATATCAA GGAGAGCTAG TGGATGGGTC TACAGACATC
 5161 ATTTTGAAG CAGACATCCA AANGTAAGTT CAGAGTACA TATCCCATTA GCGGATGCTA
 5221 GATTAGTAAT AAAACATAT TGGGGTTTGC AGCAGGAGA AAGAGATTGG CATTTGGCTC
 5281 ATGGAGTCTC CATAGATGG AGACTGAGAG AATACAGCAC ACAAGTAGAC CTTAGCTGG
 5341 CAGACCACTT AATTCACATG CATTAATTTG ATTTGTTTAC AGAATCTGCC ATAGACAGAG
 5401 CCATATTAGG ACACATAGTT TTTCTAGGT GTGACTATCA AGCAGGACAT AAGAGGTAG
 5461 GATCTCTGCA ATACTTGGCA CTGACAGCAT TGTAAAAACC AAAAAGAGA AAGCCACCTC
 5521 TGCCATGTGT TAGAAATTA GTAGAGGATA GATGAAACGA CCCCAGAG ACCAGGGGCC
 5581 GACAGAGGAA CCATACATAG AATGACACT AGAGATTCTA GAAGACTCA AGCAGGAGC
 5641 TGTAGAGCAC TTTCTAGAG CATGGCTCCA TAGCTTAGGA CAATATATCT ATGAACCTA
 5701 TGGGATACT TGGACGGAG TTGAAGCTAT AATAGAGTA CTGACACAC TACTTTTCAT
 5761 TCAATTCAGA ATTGGATGCC AACTAGCAG AATAGGCATC TTGACAGCA GAACAGCAG
 5821 AATGGAGGC AGTAGATCCT AACTAAAGC CCTCGAACCA TCCAGGAGC CAACCTAAA
 5881 CAGCTGTAA TAATGTCTTT TGCAACACT GTAGCTATCA TTGCTAGTT TGGTTTCAGA

FIGURE 11

WO 02/04493

13/114

PCT/US01/21241

5941 CAAAGGTTT AGGCAITTC TATGGCAGG AGAGCGGAG ACAGCGACGA AGCGCTCCTC
 6001 CAACTGCTGA AGATCATCAA AATCTCTAT CAAAGCAGTA AGTACACATA GTAGATGTAA
 6061 TGGTAAGTTT AAGTTTATTT AAAAGGATAG ATTATAGATT AGGAGTAGGA GCATTGATAG
 6121 TAGCCTCAAT CATAGCAATA ATAGTGTGGA CCATAGCATA TATAGAAATAT AGGAAATGG
 6181 TAAGACAAA GAAATAGAG TGGTTAATTA AAGAAATTAG GGAAAGAGCA GAAGCAGTGG
 6241 GCATGAGAG TGAATGGGAC ACAGAAAGAT TGTCAATCAAT GGTGGATATG GGGCATCTTA
 6301 GGCTTCTGGA TGCTAATGAT TGTTAACAGG GAGGACTTGT GGTTCACAGT CTACTATGGG
 6361 GTACCTGTGT GGAAGGAGGC AAAAAGCTACT CTATTCTGTG CATCAGATGC TAAGGCATAT
 6421 GAGACAGAG TGCATAATGT CTGGGCTACA CATGCTGTGG TACCCACAGA CCCCACCCCA
 6481 CAAGAAATAG TTTTGGGAAA TGTACAGAAA AATTTTAAATA TGTGGAAAAA TAACATGGCA
 6541 GATCAGATGC ATGAGGATAT ATCAGATTTA TGGGATCAAA GCCTAAAGCC ATGTGTAAAG
 6601 TTGACCCGAC TCTGTGTAC TTTAACTGT ACAGATCAAA ATGTACAGG TAATAGAACT
 6661 GTTACAGGTA ATACGAATGA TACCAATATT GCAAAATGCTA CATATAAGTA TGAGAAATG
 6721 AUAATTTGCT CTITCAATGC AACCAAGAAA TTAAGAGATA AGAAACATAA AGAGTATGCA
 6781 CTCTTTTATA AACTTGATAT AGTACCACTT AATGAAAAA GTAACTACTT TACATATAGA
 6841 TTAATAAATT GCAATACCTC AACCAATACA CAAGCCTGTG CAAAGGTCTC TTTTGGACCG
 6901 ATTCCATATC ATTACTGTGC TCCAGTGTAT TATGCGATTC TAAAGTGTAA TAATAGACA
 6961 TTCAATGGGA CAGGACCAAT TTATATGTCT AGCAGCAGTAC AATGTACACA TGGAAATTAAG
 7021 CCAATGTGAT CAATCAACT ACTGTTAAAT GGTACTCTTG CAGAAAGAGG GATATATAAT
 7081 AGATCTGAAA ATTGACAGA GAATACCAA ACAATTAATG TACTCTTAA TGAATCTGTA
 7141 GAGATTAAAT GTACAAGGCC CAACATAAT ACAAGGAAAA GTGTAAAGAT AGGACAGGA
 7201 CAAGCATTTCT ATGCAACAAA TGACCTAATA GGAACATTA GACAGGACCA TTGTAAACAT
 7261 AGTACAGATA GATGGAATAA AACTTTACAA CAGGTAAATG AUAATTAGG AGAGCAATTC
 7321 CCTAATAAAA CAATTAATTT TGAACCAAT CCAAGGAGGG ATCTAGAAT TACATGTCAAT
 7381 AGCTTTAAT GTAGAGGAGA ATTTTCTAT TGCATACAT CAAAGCTGTT TAATGTACA
 7441 TACTACCTAC AGAATGGTAC ATCAATATC AATGTAAAT CAAGCTTACC CATGCACTC
 7501 CAATGCCAAA TAAACAAAT TGTACGCTG TGGCAAGGG TACGACAGC AATGTATGCC
 7561 CTTCCCTGTG CAGGAAACAT AACATGTAGA TCAACATCA CAGGAATACT ATTGACAGT
 7621 GATGGGGAT TTAACAAAC AACCAACGAC ACAGAGGAGA CTTCAAGCC TGGAGAGGA
 7681 GATATAGGG ATAACTGGAG AACTCAATTA TATAATATAA AAGTGTAGA AATTAAGCCA
 7741 TTGGGAATAG CACCACTAA GGCABAAAGA AGATGTGTC AGAGAAAAA AAGAGCAGTG
 7801 GGAATAGGAG CTGTGTTCTT TGGGTTCTTG GGAGCAGCAG GAGGACATAT GGGGCGAGCG
 7861 TCAATACGC TGACGATACA GGCAGACAA CTGTTGTCTG GTATAGTGA CAAGCAAGGC
 7921 AATTTGCTGA AGGCTATAGA GGCACACAG CATATGTTGC AACTCACAGT CTGGGCAATT
 7981 AAGCACTGCC AGGCGAGAGT CTTGCTTATA GAAGATACC TAAAGGATCA ACAGTCTCTA
 8041 GGGATTGGG GCTGCTCTGG AGAGCTCATC TGCAACACTG CTGTGCTTGG GAATCCAGT
 8101 TGGAGTAAAT AATCTGAGC AGATATTGAG GATAACATGA CTGGAATGCA GTGGGATAGA
 8161 GUAATTAAAT ATTACACAGA AACATATTTC AGGTGCTTG AAGACTGCA AACACAGCAG
 8221 GAAAGGAATG AAAAAATTT ATTAGAATTG GACCAAGTGA ATATCTGTG GAATTTGTTT
 8281 GACATATCAA ACTGCTGTG GATATATAAA ATATTCTATA TGAATAGTAG AGGCTTGATA
 8341 GGTTTAAGAA TATTTTTCG TGTGCTCTCT ATAGTGAATA GAGTTAGGCA GGGTACTCA
 8401 CTTTGTCTAT TTCAAGCCCT TACCCCAAGC CCGAGGGGAC TCGACAGGCT CCGAGGAATC
 8461 GAAGAGAGAG GTGAGAGGCA AGACAGAGAC AGATCCATAC GATTCGTGAG CCGATTTCTG
 8521 TCGCTTGCTT GGGAGATCT GCGAGGCTG TGCTCTTCA GCTACCAAG CTTCAGAGAC
 8581 TTCAATTTAA TTGCACTGAG GGCAGTGAAT CTTCGTGGAAC ACAGCAGTCT CAGGAGACTA
 8641 CAGAGGGGCT GGGAGATCTT TAAGTATCTG GGAAGTCTTG TGCAATATT GGGTCTAGAG
 8701 CTAAAGAAAG GTGCTATTAG TCCGCTTGT ACCATAGCAA TAGCAGTAGC TGAAGGAACA
 8761 GATAGGATTA TAGAATTGGT ACRAAGAAAT TGTAGAGCTA TCTCAACAT ACCTAGGAGA
 8821 ATAAGCAGG GCTTTGAAG AGCTTTGCTA TAAATGGGA GGCAGTGTGT CAAAGCAGG
 8881 CATAGTTGGA TGGCTTGAG TAAAGAAAG AATGAGAAAG ACTGAGCCAG CAGCAGAGGG
 8941 AGTAGGACCA GCCTCTCAG ACTTAGATAG ACATGGGCGA CTTAAGACA GCAACACACC

FIGURE 11

WO 02/04493

14/114

PCT/US01/21241

9001 TGCTACTAAT GAGCTTCTG CCGGCTGCA AGCACAAGAG GAGGACGGAG ATGTAGGCTT
9061 TCCAGTCAGA CTCAGGTAC CTTTAAAGACC AATGACTTAT AAGAGTGCAG TAGATCTCAG
9121 CTTCTTTTTA AAGAAAAAG GGGGACTGGA AGGCTTAAIT TACTCTAGGA AAAGCCAGA
9181 AATCCTTGAT TTGTGGGTCT ATACACACAG AGGCTTCTTC CCGATTTGGC AAACCTACAC
9241 ATCGGGGCCA GGGTCCGAT TCCACATGAC CTTTGGATCG TGCTCTAAGC TAGTACCGAT
9301 TGACCCAGAG GAGGTGAAG AGGCCAATGA AGGAGAGGAC AACTGTTTGC TAGACCCAT
9361 GAGCCACAT GAGGCAGAG ATGAGAGTAG AGAAGTATTA AACTGGAGT TTGACAGCT
9421 TCTAGCACAC AGACACATGG CCGCGAGCT ACATCCGGAG TATTACAGAG ACTGCTGACA
9481 CAGAGGGGAC TTTCGCTCG GAGCTTTCOA CTGGGGCGTT CCGGAGGTTG TGGTCTGGC
9541 GGGACTTGGG AGTGGTCACC CTCAGATGCT GCATATAGC AGCTGCTTTT CGCTTGTACT
9601 GGGTCTCTCT CGGTAGACCA GATCTGAGCC TGGGAGCTCT CTGGCTATCT AGGAAACCA
9661 CTGCTTAGGC CTCAATAAAG CTTGCTTGA GTGCTCTAAG TAGTGTGTGC CCACTCTTTC
9721 TGTGACTCTG GTAACCTAGG ATCCCTCAGA CCCTTGTGG TAGTGTGGAA AATCTCTAGC
9781 A

FIGURE 11

WO 02/04493

15/114

PCT/US01/21241

SEQ ID NO:34

GCTGAGGCAATGAGCCAAGCAACCAGCGCAACATACTGATGCAGAGAAGCAATTT
CAAAGGCCCTAAAAGAATTATTAAATGTTCACTGTGGCAAGGAAGGGCACATAG
CTAGAAATTGTAGGGCCCTAGGAAAAAGGCTGTTGGAAATGTGGAAAGGAAGGA
CACCAAATGAAAGACTGTACTGAGAGGCAGGCTAA

FIGURE 12

WO 02/04493

16/114

PCT/US01/21241

975Pol wt until 6aa Int: (SEQ ID NO:35)
TTTTTATGGGAAGATTGGGCTTCCACACAGGGAAGGCCAGGGAATTTCTTCAGAA
CAGAACAGAGCCCAACAGCCCCACCAGCAGAGAGCTTCAAGTTGGAGGAGACAACCC
CCGCTCCGAAGCAGGAGCCGAAAGACAGGGAACDCTTAATTTCCCTCAATCACTCT
TTGGCAGGGACCCCTTGTCTCAATAAAAGTAGGGGTCAAATAAAGGAGGCTCTCTT
AGACACAGGAGCTGATGATACAGTATTAGAAGAAATGAGTTGCCAGGAAATGGA
AACCAAAAATGATAGGAGGAATTGGAGGTTTATCAAAGTAAGACAGTATGATCAA
ATACTTATAGAAAATTTGGGAAAAAGGCTATAGGTACAGTATTAATAGGACCTACA
CCTGTCAACATAATTGGAAGGAATATGTTGACTCAGCTTGGATGCACACTAAATTTT
CCAAATTAGTCCCAATTGAAACTGTGCCAGTAAATTAAGGCCAGGAATGGATGGCCCA
AAGGTTAAACAATGGCCATTGACAGAGAGAAAAATAAAGCATTAACAGCAATTTG
TGAAGAAAATGGAGAAAAGAGGAAAAATTACAAAAATTGGGCTGAAAAATCCATATA
ACACTCCAGTATTTGCCATAAAAAAGAGACAGTACTAAGTGGAGAAAAGTTAGTA
GATTTAGGGAACTTAATAAAGAACTCAAGACTTTTGGGAAGTTCAATTAGGAATA
CCACACCCAGCAGGTTAAAAAGAAAAATCAGTGACAGTACTGGATGTGGGGGA
TGCATATTTTTCAGTTCTTTAGATGAGGACTTCAGGAAATAACTGCATTCACCATA
CCTAGTATAAACAATGAAACACCCAGGGATTAGATATCAATAATGTGCTTCCACAG
GGATGGAAAGGATCAACATCAATATTCAGAGTAGCATGACAAAAATCTTAGAGCC
CTTAGAGCAAGAAATCCAGAAATAGTCATCTATCAATATATGGATGACTTGTATGT
AGGATCTGACTTAGAAAATAGGGCAACATAGAGCAAAAAATAGAGGAGTTAAGAAAAAC
ATCTGTTAAGGTGGGGATTTACCACCCGACAGAAACATCAGAAAGAACCCCA
TTTCTTTGGATGGGTATGAACTCCATCCTGACAAATGGACAGTACAGCCTATAGAG
TTGCCAGAAAAGGAAAGCTGGACTGTCAATGATATACAGAAAGTTAGTGGGAAATTT
AAATTTGGGCCAGTCAGATTTACCCAGGAATTAAGTAAGGCAACTTTGTAAATCTCT
TAGGGGGGCCAAGCACTAACAGATATAGTACCACTAACTGAAGAAGCAGAAATTAG
AATTTGGCAGAGAACAGGGAATTTCTAAGAGAACAGTACATGGAGTATATTATGAC
CCATCAAAAGACITGGTAGCTGAAATACAGAAACAGGGGCATGACCAATGGACATA
TCAAAATTAACAAAGAACCAATTCAAAACCTGAAACAGGGGAAGTATGCAAAAATGA
GGACTGCCCACTAATGATGTAAACAGTTAAACAGAGGCAGTCAAAAAATAGCT
ATGGAAGCATAGTAATATGGGAAAGACTCTCTAAATTTAGACTACCCATCCAAAA
AGAAACATGGGAGACATGGTGGACAGACTATTGGCAAGCCACCTGGATCTCTGAGT
GGGAGTTTGTATACCCCTCCCTTAGTAAATTTATGGTACCAGCTAGAGAAAGAAC
CCATAATAGGAGCAGAAACITTTCTATGTAGATGGAGCAGCTAATAGGGAACATAAA
ATAGGAAAAGCAGGCTATGTTACTGACAGAGGAAGGCAGAAATTTGTTCTCTAAC
AGAAACAACAATCAGAGAGCTGAATTACAAGCAATTCAGCTAGCTTTGCAAGATT
AGGATCAGAGTAACATAGTAACAGACTCAGATATGCATTAGGAATCATTCAG
CACAACAGATAAGAGTGAATCAGAGTTAGTCAACCAATAATAGAAATTAATA
AAAAAGGAAAAGGTCTACCTGTATGGGTACCAGCACATAAAGGAATTTGGAGGAAA
TGAACAAATAGATAAATTAGTAAAGTAAGGGAATCAGGAAAAGTGTGTTTCTAGATG
GAATAGAT

FIGURE 13

WO 02/04493

17/114

PCT/US01/21241

SEQ ID NO:36

GGGGGCATCGTGATCTACCACTACATGGACGACCTGTACGTGGGCAGGGCG
GC

FIGURE 14

WO 02/04493

18/114

PCT/US01/21241

SEQ ID NO: 37

GGIVTYQYMDDLTVGSGG

FIGURE 15

WO 02/04493

19/114

PCT/US01/21241

12_5/1ZA (SEQ ID NO:45)

TGGAAGGGTTAATTACTCCAGGAAAAGGCAAGAGATCCTTGATTATGGGTCTATC
ACACACAAGGCTACTTCCCTGATTGGCAAACTACACACCGGGACCAGGGTTCAGA
TATCCACTGACCTTTGGATGGTGTCTCAAGCTAGTGCCAGTTGACCCAAAGGAAGTA
GAAGAGGCCAACGGAGGAGAGACAACCTGTTTGTCTACACCCCTATGAGCCAGTATGG
AATGGATGATGAACACAAAGAAAGTGTACAGTGGAAAGTTTGACAGCAGCCTAGCAC
GCAGACACCTGGCCCGGAGCTACATCCGGATTATTACAAAGACTGCTGAACACAGA
AGGGACTTTCCGCTGGGACTTTCCACTGGGGCGTTCCAGGGGAGTGGTCTGGGCG
GGACTGGGAGTGGCCAGCCCTCAGATGCTGCATATAAGCAGCGGCTTTCCGCTGT
CTGGTCTCTCTAGGTAGACAGATCCGAGCCTGGGAGCTCTCTGTCTATCTGGGGA
ACCCACTGCTTAGGCCCTCAATAAAGCTTGCCTTGAGTGCTCTAAGTAGTGTGTGCCC
ATCTGTGTGTGACTCTGTGAACCTCTGGTAACTAGAGATCCCTCAGACCCCTTTGTGGT
AGTGTGGAAAATCTCTAGCAGTGGCGCCGAAACAGGGACTTGAAAGCGAAAGTGAG
ACCAAGAGAGATCTCTCGACGCAGGACTCGGCTTGTGAAGTGCACTCGGCAAGAG
GCGAGGGGGGCGACTGGTGTGAGTACGCCAAATTTTTTTGACTAGCGGAGGCTAGA
AGGAGAGAGATGGGTGCGAGAGCGTCAATATTAAGAGGGGAAAAATTAGACAAAT
GGGAAAAAATTAGGTTACGCCCAGGGGGGAGAAAAACACTATATGCTAAACACCTA
GTATGGGCAAGCAGAGAGCTGGAAAGATTTCAGTTAACCCCTGGGCTTTTAGAGAC
ATCAGACGGATGTAGAC AAATAATAAAACAGCTACAACCAGCTCTTCAGA
CAGGAACAGAGGAAATTAGATCATTATTTAACACAGTAGCAACTCTCTATTGTGTAC
ATAAAGGGATAGATGTACGAGACACCAAGGAAGCCTTAGACAAAGATAGAGGAGGA
ACAAACAAATGTACGCAAAAAACACAGCAGGCGGGAAGCGGCTGACAAAAGGTC
AGTCAAAATTATCCTATAGTGCAGAACCTCCAAGGGCAATGGTACACCAAGGCCAT
ATCACTAGAACCTTGAATGCATGGGTAAAAGTAATAGAGGAGAGGCTTTTAGCC
CAGAGGTAATACCCATGTTTACAGCATTATCAGAAGGAGCCACCCCAAGATTTA
AACACCATGTTAAATACAGTGGGGGACATCAAGCAGCCATGCAAATGTTAAAAG
ATAACCATCAATGAGGAGGCTGCAGAAATGGGATAGGTTACATCCAGTACATGCAGGG
CCTGTTGCACACAGGCCAGATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACCTA
CTAGTACCCCTTCAAGAACAAATAGCATGGATGACAAGTAACCCACCTATCCCACTA
GGGGACATCTATAAAAGGTGGATAATTCTGGGGTTAAATAAAATAGTAAGAATGTA
CAGCCCTGTACGACTTTTAGACATAAAACAAGGACCAAGGAACCCCTTTAGAGACT
ATGTAGACCGGTTCTTCAAACTTTAAGAGCTGAACATCTACACAAGAGGTAAAA
AATTGGATGACAGACACCTTGTAGTCCAAAATGCGAACCCAGATTGTAAGACCAT
TTAAGAGCATTAGGACCAGGGCTTCATTAGAAGAAATGATGACAGCATGTCAGGG
AGTGGGAGGACCTAGCCACAAAGCAAGAGTTTGGCTGAGGCAATGAGCCAAGCAA
ACAATACAAGTGTAAATGATACAGAAAAGCAATTTTAAAGGCCCTAGAAGAGCTGTT
AAATGTTTCAACTGTGCGCAGGAAAGGGCAGATAGCCAGGAATTGACGGGCCCTAG
GAAAAGGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAAATGAAAGACTGTACT
GAGAGGCAGGCTAATTTTTTAGGGAAAATTTGGCCTTCCACAAGGGGAGGCCAGG
GAATTTCTTCAGAGCAGACAGGCCAACAGCCCACTAGAACCAACAGCCC
CACCAGCAGAGAGCTTCAAGTTCAAGGAGACTCCGAAGCAGGAGCCGAAAGACAG
GGAACCTTTAACTTCCCTCAAATCACTCTTTGGCAGGACCCCTGTCTCAATAAAA

FIGURE 16

WO 02/04493

20/114

PCT/US01/21241

GTAGCGGGCCAAACAAAGGAGGCTCTTTAGATACAGGAGCAGATGATACAGTACT
 AGAAGAAATAAACTTGCCAGGAAATGGAACCAAAATGATAGGAGGAATTGGA
 GGTTTTATCAAAAGTAAGACAGTATGATCAAAATACCTTATAGAAATTTGTGGAAAAGG
 GCTATAGTACAGTATTAGTAGGACCTACACCTGTCAACATAAATGGAAAGAAATCTG
 TTGACTCAGCTTGGATGCACACTAAATTTCCAATTAGCCCCATTGAAACTGTACCA
 GTAAATTTAAAGCCAGGAATGGATGGCCAAAGGTTAAACAATGGCCATTGACAGA
 AGAAAAATAAAAGCATTAAACAGAAATTTGTGAGGAAATGGAGAAGGAAGGAAAA
 ATTACAAAAATTGGGCGCTGAAATCCATATAACACTCCAGTATTGGCCATAAGAAAG
 AAGGACAGTACAAAGTGGAGAAAATTAGTAGATTTTCAAGGAACTCAATAAAGAAC
 TCAAGACTTTTGGGAAGTCCAATTAGGAATACCACACCCAGCAGGTTTAAAAAAGA
 AAAAATCAGTGACAGTACTGGATGTGGAGATGCATATTTTCAGTCCCTTAGATG
 AGAGCTTCAGAAAAATATCTGCATTCCACATACTAGTATAACAATGAAACACCA
 GGGATTAGATAATCAATATAATGTTCTTCCACAGGGATGGAAAGGATCACCAGCAA
 TATTCCAGAGTAGCATGACAAGAACTTAGAGCCCTTAGAACACAAAAACCCAGAA
 GTAGTTATCTCAATATATGGATGACTTATATGTAGGATCTGACTTAGAAATAGGG
 CAACATAGAGCAAAAAATAGAGGAGTTAAGAGGACACCTATTGAAATGGGGATTAC
 CACACCAAGCAAGAAACATCAGAAAGAACCCCACTTTCTTTGGATGGGGTATGAAC
 TCCATCCTGACAAATGGACAGTACAGCCTATACAGCTGCCAGAAAGGAGAGCTGG
 ACTGTCAATGATATACAGAAAGTTAGTGGGAAAGTTAACTGGGCAAGTCAGATTTA
 CCCAGGGATTAAAGTAAGGCAACTGTGTAACCTCCTTAGGGGAGCCAAAGCACTAA
 CAGACATAGTGCCACTGACTGAAAGAGCAGAATTAGAATTGGCTGAGAACAGGGA
 AATTCTAAAAGAACAGTACATGGAGTATATTATGACCCATCAAAAGATTTAATAG
 CTGAAATACAGAAACAGGGGAATGACCAATGGACATATCAAAATTTACCAAGAAC
 ATTTAAAAATCTGAGAACAGGAAAGTATGCAAAAATGAGGACTGCCCACTAATG
 ATGTGAAACAGTTAGCAGAGGCAAGTGCAAAAGATAACCCAGGAAAGCATAAGTAATA
 TGGGGAAAAACTCCTAAATTTAGACTACCCATCCCAAAAGAACATGGGAGACATG
 GTGGTCAGACTATTGGCAAGCCACCTGGATTCTGAGTGGGAGTTTGTCAATACCCC
 TCCCCTAGTAAATTTGTGTACCAAGCTGGAAAAAGAACCCATAGTAGGGGCAGAAA
 CTTTCTATGTAGATGGAGCAGCAATAGGGAACTAAATAGGAAAGCAGGGTAT
 GTCAGTACAAAAGGAAGGCAGAAAGTTGTTTCCCTCACTGAAACACAAATCAGAA
 GACTGAATTACAAGCAATTCAGCTAGCTTTGCAGGATTGAGGGCCAGAAAGTAAACA
 TAGTAACAGACTCACAGTATGCATTAGGAATCATTCAAGCACAACAGATAAGAGT
 GAATCAGAATTAGTCAGTCAATAATAGAACAGTTGATAAAAAAGGAAAAAGTCTA
 CCTATCATGGGTACCAGCACATAAAGGAATTGGAGGAAATGAACAAGTAGACAAAT
 TAGTAAGTAGTGGAAATCAGAAAAGTACTGTTTCTAGATGGAATAGATAAAGCTCAA
 GAAGAGCATGAAAAATATCACAGCAATTGGAGAGCAATGGCTAGTGAATTTAATCT
 GCCACCCATAGTAGCAAGGAAATAGTAGCCAGCTGTGATAAATGTAGCTAAAG
 GGGAAAGCCATGCATGGACAAGTCGACTGTAGTCCAGGAATATGGCAATTAGACTGT
 ACACATTTAGAAAGGAAAAATCATCTCTAGTAGCAGTCCATGTAGCCAGTGGCTACAT
 GGAAGCAGAGGTTATCCAGCAGAAACAGGACAAAGAAACAGCATACTTTATACTAA
 AATTAGCAGGAAGATGGCCAGTCAAAGTAATACATACAGATAATGGCAGTAATTTT
 ACCAGTACCGCAGTTAAGCAGCCTGTTGGTGGGCAGATATCCAACGGGAATTTGG
 AATTCCTACAATCCCCAAAGTCAAGGAGTAGTAGAATCCATGAATAAAGAAATTAA

FIGURE 16

WO 02/04493

21/114

PCT/US01/21241

AGAAAAATCATAGGGCAAGTAAGAGATCAAGCTGAGCACCTTAAGACAGCAGTACAA
 ATGGCAGTATTCATTCAAAATTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGC
 AGGGGAGAGAATAATAGACATAATAGCATCAGACATACAACTAAAGAATTACAAA
 AACAAATTATAAAAATTCAAAATTTTCGGGTTTATTACAGAGACAGCAGAGCCCTA
 TTTGGAAAGGACCAAGCCAACTACTCTGGAAAGGTGAAGGGGCAGTAGTAATACAA
 GATAATAGTGATATAAAGGTAGTACCAAGAAAGGAAAGCAAAAATCATTAAAGGACTA
 TGGAAAACAGATGGCAGGTGCTGATTGTGTGGCAGGTAGACAGGATGAAGATTAGA
 ACATGGCACAGTTTAGTAAAGCACCATATGTATGTTTCGAGGAGAGCTGATGGATGG
 TTCTACAGACATCATTATGAAAGCAGACACCCAAAAGTAAGTTACAGAAATACACAT
 CCCATTAGGAGATGCCAGGTAGTAATAAAAACATATTGGGGTCTGCAGACAGGAG
 AAAGAGCTTGGCATTGGGTACGGAGTCTCCATAGAAATGGAGATTGAGAAGATAT
 AGCACACAAGTAGACCTGACCTGACAGACCAACTAATTTCATATGCATTATTTGAT
 TGTTTTGCAGAAATCTGCCATAAGGAAAGCCATACTAGGACAGATAGTTAGCCCTAA
 GTGTGACTATCAAGCAGGACATAACAAGGTAGGATCTCTACAATACTTGGCACTGA
 CAGCATTGATAAAACCAAAAAAGATAAAGCCACCTCTGGCTAGTGTAGGAAATTA
 GTAGAGGATAGATGGAACAAGCCCAAGAGACCAAGGGGCGCAGAGGGAACCAT
 CAATGAATGGACACTAGAGCTTTTGAAGAACTCAAGCAGGAAGCTGTCAGACACT
 TTCTAGACCAATGGCTCCATAACTTAGGACAACATATCTATGAAACCTTATGGAGATA
 CTGGACAGGAGTTGAAGCAATAATAAGAAATCCTGCAACAATTACTGTTTATTCATT
 TCAGGATTGGGTGCCATATAGCAGAAATAGGCATTTTGGACAGAGAAGAGCAAGA
 AATGGAGCCAATAGATCCTAACCTAGAACCTTGAAACCATCCAGGAAGTCAGCCTA
 AAACGTCTGTGAATGGGTGTTACTGTAAACGTTGCAGCTATCATTGTCTAGTTTGTCT
 TCAGAAAAAAGGCTTAGGCATTTACTATGGCAGGAAGAAAGCGAGACAGCGACGAA
 GCGCTCCTCCAAAGCAATAAAGATCATCAAGATCCTTACCAAAAGCAGTAAGTACCG
 AATAGTATATGTAATGTTAGATTAACTGCAAGAATAGATTCTAGATTAGGAATAGG
 AGCATTTGATAGTAGCACTAATCATAGCAATAATAGTGTGGACCATAGTATATATAG
 AATATAGGAAATTTGTAAGGCAAGGAAAAATAGACTGGTTAGTTAAAAGGATTAGG
 GAAAGAGCAGAAGACAGTGGCAATGAGAGCGAGGGGGATCTGAGAATTATCGA
 CACTGGTGGATATGGGGCACTTAGGCTTTTGGATGCTAATGATGTGTAATGTGAA
 GGGCTTGTGGGTCACAGTCTACTACGGGTACCTGTGGGGAGAGAAGCAAAAACCT
 ACTCTATTTTGTGATCAGATGCTAAAGCATATGAGAAAGAAGTGCATAATGTCTG
 GGCTACACATGCTGTGTACCCACAGACCCCAACCCACAAGAGTGAATTTGGGC
 AATGTAAACAGAAAATTTTAAACATGTGGAAAAATGACATGTGTGGATCAGATGCAGG
 AAGATATAATCAGTTTATGGGATCAAGCCCTTAAGCCATGTGTAATAATGACCCCA
 CTCTGTGTCACCTTAAACGTGTACAAATGCAACTGTAACTACAATAATACCTCTAAA
 GACATGAAAAATGCTCTTTCTATGTAACACAGAATTAAGAGATAAGAAAAAGAA
 AGAAAAATGCACCTTTTATAGACTTGATATAGTACCCTTAATAATAGGAAGAATGG
 GAATATTAACAACCTATAGATTAAATAATTGTAATACCTCAGCCATAACACAAGCCTG
 TCCAAAAGTCTGTTTGAACCAATTCCTATACATTATTGTGCTCCAGCTGGTTATGG
 CCTCTAAAATGTAATAATAAGAAATTCATGGAATAGGACCATGCGATAATGTGAG
 CACAGTACAAATGTACACATGAATTAAGCCAGTGTATCAACTCAATTACTGTTAAA
 TGGTAGCCTAGCAGAAGAAGAGATAATAATTAGATCTGAAAACTGACAAAACAATG
 TCAAAAACAATAATAGTACATCTTAATGAATCTATAGAGATTAAATGTACAAGACC

FIGURE 16

WO 02/04493

22/114

PCT/US01/21241

TGGCAATAATACAAGAAAGAGTGTGAGAATAGGACCAGGACAAAGCATTTCTATGCA
 ACAGGAGACATAATAGGAGATATAAGACAAGCACATTTGTAACATTAGTAAAAATGA
 ATGGAAATACAACCTTTACAAAGGGTAAGTCAAAAAATTACAAGAACTCTTCCCTAATA
 GTACAGGGATAAAAATTTGCACCACACTCAGGAGGGGACCTAGAAAATTACTACACAT
 AGCTTTAATTTGGGAGGAGAATTTTCTATTGCAATACAACAGACCTGTTTAATAGT
 ACATACAGTAATGGTACATGCACTAATGGTACATGCTCTAATAATACAGAGCG
 CATCACACTCCAATGCAGAATAAAACAAATTATAAACATGTGGCAGGAGGTAGGAC
 GAGCAATGTATGCCCTCCCTATTGCAGGAACATAACATGTAGATCAAAATATTACA
 GGACTACTATTAACACGTGATGGAGGAGATAATAACTGAAACAGAGACATTTCAG
 ACCTGGAGGAGGAGACATGAGGGACAATTGGAGAAGTGAATTATATAAATACAAG
 GTGCTAGAAATTAACCATTAGGACTAGCACCCACTGTCTGCAAAAAGGAGAGTGGT
 GGAGAGAGAAAAAGAGCAGTAGGAATAGGAGCTGTGTTCCCTGGGTCTCTGGGAG
 CAGCAGGAAGCACTATGGGCGCAGCATCAATAACGCTGACGGTACAGGCCAGACAA
 TTATTGTCTGGTATAGTGCAACAGCAAAAGTAATTTGCTGAGGGCTATAGAGGCGCAA
 CAGCAATGTGCAACTCAGGTCTGGGGCAATTAAGCAGCTCCAGGCAAGAGTCCCTG
 GCTATAGAGAGATACCTACAGGATCAACAGCTCCTAGGACTGTGGGGCTGCTCTGG
 AAAACTCATCTGCACCACTAATGTGCTTTGGAACTCTAGTTGGAGTAATAAACTCA
 AAGTGATATTGGGATAACATGACCTGGATGCAAGGGATAGGAAATTAAGTAAT
 ACACAAACACAAATATACAGGTGTCTGAAGACTCGCAAAAGCCAGCAGGAAAGAAA
 TGAAAAAGATTTACTAGCAATTGGACAGGTGGAAACAATCTGTGGAATTGGTTTAGCAT
 AACAAATTTGGCTGTGTATATAAAATATTCATAATGATAGTAGGAGGCTTGATAG
 GTTAAAGAATAATTTTGTCTGTCTCTCTAGTAAATAGAGTTAGGCAGGGATACT
 CACCCCTTGTCAATTGCAGACCTTATCCCAAACCCGAGGGGACCCGACAGGCTCGGA
 GGAATCGAAGAGAAGGTGGAGAGCAAGACAGCAGCAGATCCATTGATAGTGA
 GCGGATTTCTGACACTTGGCTGGGACGACCTACGAAGCCTGTGCCCTCTCTGCTACC
 ACCGATTGAGAGACTTCATATTAATTGTAGTGAGAGCAGTGAACCTCTGGGACAC
 AGTAGTCTCAGGGGACTGCAGAGGGGTGGGGAACCCCTAAGTATTTGGGGAGTCT
 TGTGCAATATTGGGGTCTAGAGTTAAAAAAGAGTGTATTAATCTGCTTGATACTAT
 AGCAATAGCAGTAGCTGAAGGAACAGATAGGATTCTAGAATTCATACAAAACCTTT
 GTAGAGGTATCCGCAACGTACCTAGAAAGATAAGACAGGGCTTCGAAGCAGCTTTG
 CAATAAAATGGGGGGCAAGTGGTCAAAAAGCAGTATAATTGGATGGCCTGAAGTAA
 GAGAAAGAATCAGACGAACTAGGTCAGCAGCAGAGGGAGTAGGATCAGCGTCTCA
 AGACTTAGAGAAACATGGGGCACTTACAACCAGCAACACAGCCACAACAATGCTG
 CTTCGCCCTGGCTGGAAGGCAAGAGGAGGAAGGAGAAGTAGGCTTTCCAGTCAGA
 CCTCAGGTACCTTTAAGACCAATGACTTATAAAGCAGCAATAGATCTCAGCTTCTTT
 TTAAAGAAAAGGGGGGACTGGAAGGGTTAATTTACTCCAAGAAAAGGCAAGAGAT
 CCTTGAATTTGGGGTTTATAACACACAAGGCTTCTTCCCTGATTGGCAAAACTACAC
 ACCGGGACCCAGGGGTAGATTTCCACTGACCTTTGGATGGTACTTCAAGCTAGAGCC
 AGTCGATCCAAGGGAAGTAGAAGAGGCCAATGAAGGAGAAAACAATGTTTACTAC
 ACCCTATGAGCCAGCATGGAATGGAGGATGAAGACAGAGAAATTAAGATGGAAG
 TTTGACAGTACGCTAGCAGCAGACATGGCCCGCAGCTACATCCGGAGTATTAC
 AAAGACTGCTGACACAGAAGGGACTTTCCGCTGGGACTTCCACTGGGGCGTTCCAG
 GAGGTGTGGTCTGGGCGGACAGGGGAGTGGTCAGCCCTGAGATGCTGCATATAAG
 CAGCTGCTTTTCGCTGTACTGGGTCTCTAGGTAGACCAGATCTGAGCCCCGGAG

FIGURE 16

WO 02/04493

23/114

PCT/US01/21241

CTCTCTGGCTATCTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTG
CCTTGAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGA
CCACTTGTGGTAGTGTGGAAAATCTCTAGCA

FIGURE 16

WO 02/04493

24/114

PCT/US01/21241

>C4_Env_TV1_C_ZA_opt_short (SEQ ID NO:46)

CATCACCCCTGCAGTGCAGATCAAGCAGATCGTGCGCATGTGGCAGGGCGTGGCCAGGCGCATGTACGCCGCCCCCATCG
CCGGCAACATCACTGC

FIGURE 17

WO 02/04493

25/114

PCT/US01/21241

>C4_Rev_TV1_C_3A_opt (SEQ ID NO:47)

CTGCCCATCACCCTGCAGTGCAGATCAAGCAGATCGTGGCATGTGGCAGGGCGTGGGCCAGGCCATGTACGCCCCCCC
CATCGCCGCGACATCACCTGCCGAGCAACATCACCGGCATCTGCTGACCGGAGCGGGGGC

FIGURE 18

WO 02/04493

26/114

PCT/US01/21241

>C4_Env_TV3_C_ZA_wt (SEQ ID NO:48)

TTACCCATCACAATGCAAAATAAACAAATTGTACGCATGTGGCAAGGGGTAGGACAAGCAATGTATGCCCCCTCC
CATGTGAGGAAACATAACATGTAGATCAACATCAGGGAATACTATTGACACCTGATGGGGGA

FIGURE 19

PCIT/US01/21241

>ENVGP160 TV1_C_ZAopt (SEQ ID NO:49)

[illegible]

FIGURE 20

WO 02/04493

28/114

PCT/US01/21241

>E000p160_TV1_C_ZAwT (SEQ ID NO:501)

ATGAGAGTGTATGGGGACACAGAAAGAAATTGTCAACAATGGTGGATATGGGGCATCTTAGGCTTCGGATGCTAATGATTTG
 TAACACGGGAGGACTTGTGGGTCACAGTCTACTATGGGGTACCTGTGTGGAGAGCAAAACTACTCTATTTCTGTGCAAT
 CAGATGCTAAAGCATATGAGACAGAGAGTGCATTAATGTCTGGGCTACACATGCTTGTGTACCCACAGACCCCAACCCACAA
 GAAATAGTTTGGGAAATGTAAACAGAAATTTAATATGTGGGAAATTAACATGGGAGATCAGATGCATGAGGATATTAAT
 CAGTTTATGGGATCAAGCCATAAGCCATGTGTAAAGTTGACCCCATCTGTGTGTCACCTTTAAACTGTACAGATACAAATG
 TTACAGGCTATAGAACTGTACAGGTAATACAAATGATACCAATATTGCAAAATGCTACATATAAGTATGAGAAATGAAA
 AATTGCTCTTTCAATGCAACACAGAAATTAGAGATAGAAACATAAAGAAATATGCACCTCTTTATAAACTTGATATAGT
 ACCACTTAATGAAATAGTAACTACTTTACATATAGATTAAATAAATGGCAATACCTCAACCAATACCAAACTGTGTCAR
 AGGTCTCTTTGACCCGATTCCTATACATTTACGTGTGCTCCAGCTGATTATGCCATTCTAAAGTGTAAATATAGCAATTC
 AATGCGACAGGACCATGTGTATATATGTGACGACATGACAAATGTACACATGGAAATTAAGCCAGTGTGATCAACTCAACTACT
 GTTAAATGTAGTCTAGCAGAAAGAGGGATATTAATAGATCTGAAATTTGACAGGATACCAAAACATAAATAGTATC
 ATCTTAATGATCTGTAGAGATTAAATGTACAAAGGCCCAACAAATATACAAAGGAAAGTATACCAAAACATAAATAGTATC
 GCATCTATGTGACAAATAGCAATATAGGAAACATAGACAAAGCACATTTGAAATTTAGTACACATAGATGGATATAAAC
 TTTACAAACGGTAATGAAATAATAGGAGAGCATTTCCCTAAATAAAGCAATTAATTTGAACCCATTCGAGGAGGGATTC
 TAGAAATTACAATGCATAGCTTTAATTGTAGAGGAGAAATTTCTATTTGCAATACATCAAACTGTATTATAGTAAATAG
 TACCTTAAGAAATGTTACATACAAATACAAATGTTAAATTTCAATTTGCAATACATCAAACTGTATTATAGTAAATAG
 ACAGATGTGGCAGGGGTAGGACAAAGCAATGTATGTCCCTCCATTTGCAAGAAACATACATTTAGATCAAAACATAAATAGT
 GAATACTATTGACAGCTGTATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 ATGAGGATTAATGAGAGAGTGTGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 AAAAGAGAGAGTGTGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 GCATTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 TTGCTTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 GGTATTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 TGTCTTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 AGAATTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 TAGTATTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 TTGCTTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 CAGACAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 ACCACCGCTTGAAGAGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 AGCGGTGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 GCTTATTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT
 TCAATATTAATGAGGAGTTTAAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT

FIGURE 21

PCY/US01/21241

[illegible]

FIGURE 22

WO 02/04493

30/114

PCT/US01/21241

>Gas_TV1_C_ZAw1 (58Q ID NO:52)

ATGGGTGCGAGAGCGTCATATTAAACCGCGGAAATTAGATAAATGGGAAGAATTAGGTTAAGGCCAGGGGGAAGAA
ACATTTATATGTTAAACATCTAGTATGGGCAAGCGGGAGCTGGAAAGATTTCACCTTAACCTTGGCTTGTAGAAACAT
CAGGAGGCTGTAAACAAATATATAAACAGCTACCAACAGCTCTTCAGACAGGACAGAGAGACTTAGATCATTTATCAAC
ACAGTAGCCTCTCTATTTGTGTACATAAAGGGATAGAGGTACAGACACCAAGGAGGCTTAGACATAGATAGAGGAGA
ACAAAACAAATGTCTAGCAAAAGCCACACAGGCAAAAGCGAGCTGACGAAAGGTCAGTCAAAATTTATCCTATAGTACAGA
ATGCCCAAGGGCAATGGTACACCAAGCTATATCACCTAGACATTTGAATGCCATGGAATAAAGTATATAGAGGAAAGGCT
TTCAATCCAGAGGAATGCCATGTTTACAGCATTTATCAGAAAGGAGCCACCCACAGATTTAAACACAAATGTTAAATAC
AGTGGGGGACATCAGGAGCCATGCAATGTTAAAGATACCATCAATGAGGAGGCTGAGAAATGGGATAGGACACATC
CAGTACATGTCAGGGCTGTGCAACAGGCGAGATGAGAGAACCAAGGGGAGTGACATAGCAGGAACTACTAGTACCCCTT
CAGGAACAAATAGCATGGATGACCAAGTATCCACCTATTCAGTAGAAGACATCTATTAAGATGGATATTCTGAGGTT
AAATAAATAGTAGAATGTATAGCCCTGTAGCATTTTGGACATAAACAGGCGCAAAAGAACCCCTTAGAGACTATC
TAGACCGGTTCTTTAAACCTTAGAGCTGAACAGCTTACCAAGATGTAAAGATTTGGATGACAGACCCCTTGTGGTC
CAAAATGOGAACCCAGATTTGTAAGACCTTTTAAAGCATTAGGACCCAGGGGCTCATTAGAGAGAAATGATGACAGCATG
TCAGGAGGTGGAGGACCTAGCCATAGGCAAGAGTGTGCTGAGGCAATGAGCCAGCAGACAGTAACTACTAGTGC
AGGAGAGCAATTTTAAAGGCTCTAACAGATTTTAAATGTTTCAACTGTGGCAAGTAGGGCACATAGCCAGAAATTGC
AGGGCCCTTAGGAAAGAGGCTGTGGAATGTGGACAGGAAGGACACCAATATGAAGAGCTGTACTGAGAGGCGAGGCTAA
TTTTTTAGGGAATTTGGCTTCCCAAGGGGAGGCGAGGAATTTCTTCCAGAACAGACAGAGCCAAACAGCCCCAC
CAGCAAAACCAAGCCCCAGCGAGAGAGCTTCAGGTTCCAGGAGACACCCCCGTGCGAGGAAGGAGAAAGAGAGG
GAGCTTTAACTTCCCTCAATCACTCTTTGGCAGCGACCCCTTGTCTCAATAA

FIGURE 23

WO 02/04493

31/114

PCT/US01/21241

>Gag_TV1_2A_MIRopt (SEQ ID NO:53)

GACATCAAGCAGGGCCCCAAGGAGCCCTTCGGGACTACGTGGACCGCTTCTTCAAGACC

FIGURE 24

WO 02/04493
32/114
PCT/US01/21241
>Gag_TV1_ZA_NHRwt (SEQ ID NO:54)
GACATAAAACAAGGGCCAAAGAACCCCTTAGAGACTATGTAGACCGTTCTTAAACC

FIGURE 25

WO 02/04493

33/114

PCT/US01/21241

>Nef_TV1_C_2Aopt (SEQ ID NO:55)

ATGGGCGGCAAGTGGAGCAAGCCAGCATCTGGGCTGSCCCGCGGTGCGCGAGGCGATGCGCGCACCGAGCCCGCCGC
CGAGGGCGTGGGCGCCGCCAGCCAGGACCTGGACCGCCACCGCGCCCTGACCGAGCAACACCCCGCCACCAACGAGG
CCTGCGCTTGGCTGCGAGCCAGGAGGAGGAGCGGAGGTGGGCTTCCCCGTGGCGCCCGAGGTGCCCCGCGCCCGTGG
ACCTACAAGAGCGCCCTGGACCTGAGCTTCTTCCCTGAAGGAGAAGGGCGGCGCTGGAGGCGCTGATCTACAGCGCCGAGCG
CCAGGAGATCCCTGAGCTTGGGTGTACACACCCAGGGCTTCTTCCCGGACTGGCGAGAACTACACCGAGCGGCCCGGCG
TGGGCTTCCCGCTGACCTTGGCTGGTGTCTCAAGCTGGTGGCGCCCGCGAGGTGAAGGAGGCGCAACGAGGGC
GAGGACAACCTGCTTGCACCCCATGAGCCAGCACGGCGCCGAGGACCGAGGACCGGAGGTGCTGAAGTGAAGTTCGA
CAGCCTGCTGGCCACCGCCACATGGCCCGGAGCTGCACCCGAGTACTACAGGAATGCTGA

FIGURE 26

WO 02/04493

34/114

PCT/US01/21241

>Ref_TV1_C_ZAwc (SEQ ID NO:56)

ATGGGAGGCAGTGTGTCAAAACGCAGCATACTTGGATGGGCTGCGAGTAAGAGAAAGATGAGAAGAACTGAGCCAGCCAGC
AGAGGGAGTAGGAGCAGCCCTCTCAAGACTTAGATAGACATGGGGCACTTACAGCCAGCAACACACCTGCTACTAATGAAAG
CTTGTGCTGGCTGCAAGCACAAGAGGAGGACGAGATGTAGGCTTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATG
ACTTATAAGAGTGCAGTAGATCTCAGCTTCTTTTAAAGAAAAAGGGGGGACTGGAAAGGTTAATTTACTCTAGGAAAG
GCAAGAAATCCTTGATTGTGGCTCTATAACACACAGGGCTTCTTCCCTGATTGGCAAACTACACATCGGGGCCAGGGG
TCCGATTCCCACTGACCTTTGGATGGTGCTTCAAGCTAGTACCAGTTGACCCAGGGAGGTGAAGAGGGCCCAATGAGGA
GAAGACAACCTGTTTGCTACACCTATGAGCCACATGGAGCAGAGGATGAAGATAGAGAACTATTAAAGTGGAAAGTTTGA
CAGGCTTCTAGCACACAGACACATGGCCCGCAGCTACATCCGGAGTATTACAAAGACTGCTGA

FIGURE 27

WO 02/04493

35/114

PCT/US01/21241

>RefD125G_TV1_C_ZAopt (SEQ ID NO:57)

ATGGGCGGCAAGTGGAGCAAGCGCAGCATCGTGGGCTGGCCCGCCCTGGCGGAGCGCATGGCCGACCCGAGCCCGCCG
CGAGGCGCTGGCGCGCGCGAGCGAGGACCTGGACCGCCACGGCGCCCTGACCGAGCGCAACACCCCGCCACCAACGAGG
CCTGCGCCTGGCTGCGAGGCCAGGAGGAGGAGGAGGACCTGGGCTTCCCGTGGCGCCCGCAGGTGCGCCCTGGCGCCCATG
ACCTACAAAGAGCGCCCTGGACCTGAGCTTCTTCCCTGAAGGAGAGGGCGCGCTGGAGGGCTGATCTACAGCGCGAAGCG
CCAGGAGATCCTGGACCTTGGGTGTACAACACCCAGGGCTTCTTCCCGGCTGGCGAGAACTACACAGCGCGCCCGCG
TGCGCTTCCCGCTGACCTTGGCTGTGCTTCAAGCTGGGTGCCCGTGGACCCCGCGAGGTGAAGGAGGCCAACGAGGGC
GAGGACAACTGCGCTGCTGCACCCCATGAGCCAGCACGGCGCGGAGGACGAGACCCCGAGGTGCTGAAGTGGAACTTCCA
CAGCCTGCTGGCCACCGCCACATGGCCCGGAGCTGCCACCCCGAGTACTACAAAGGACTGCTGA

FIGURE 28

WO 02/04493

36/114

PCT/US01/21241

>p15RnaseH_TV1_C_2hapt (SEQ ID NO:58)

ACCTTCFACGTGGACGGCGCCACCAACCGCGAGGCCAAGATCGGCAAGGCCGGCTACCTGACCSACCCGCGGCCAAGAA
GATCCTGACCCCTGACCAACACCAACCAACCAAGACCGGAGCTGCAGGCCATCCAGCTGACCCCTGCAGGACAGCGGCAGCG
AGGTGAACATCGTGACCSACAGCCAGTACCGCCCTGGGCATCATCCAGGCCACCCCGACAGAGCCGACAGCGAGATCTTC
AACCAAGATCATCGAGAGCTGATCAACAAAGGAGCGCATCTACCTGAGCTGGGTGCCCGCCACAGGGGCATCGCGGCAA
CGAGCAGGTGGACAGCTGGTGAGCAAGGGCATC

FIGURE 29

WO 02/04493

37/114

PCT/US01/21241

>p15RNAseH_TV1_C_Zawt. (SEQ ID NO:59)

ACTTTCCTAGTAGATGGAGCACTAATAGCGAAGCTAAATAGGAAAGCAGGGTATGTTACTGACAGAGGAAGGCAGAA
AATTGTTACTCTAATACACACAAATCAGAGAGCTGAGTTACAGCAATTCAGCTAGCTCTGCAGGATTCAAGATCAG
AAGTAAACATAGTACAGACTCAGATATGCATTAGGAATCATTCAGGACACACAGATAAGAGTGAGCTCAGAGATATTT
AACCAAAATATAGACAGTTTATAAACAGGAAAGAACTACCTGTCATGGTACCAGCACATAAGGAATTGGGGGAAA
TGAACAGTACATAAATTACTAAGTAAGGGAAAT

FIGURE 30

WO 02/04493

38/114

PCT/US01/21241

ap31Int_TV1_C_Zaopt (SEQ ID NO:60)

CGCAGGTTGCTGTTCCTGGACGGCATCGACAAAGGCCAGGAGGACCGAGCGTACACAGCAACTGAGCGCCCATGGC
CAACGAGTTTCAACCTGCCCCCATCGTGGCCAAAGGAGATCGTGGCCAGCTGGGACAAAGTGGCAGCTGAAGGCGAGGCGCA
TCCACGGCCAGGTGGACTGCAAGCCCCGGCATCTGGCAGCTGGACTGCCACCCACCTGGAGGGCAGGATATCTTGGTGGCC
GTGCACGTGGCCAGCGGCTACATGGAGGCCGAGGTGATCCCGCCGAGACCCGCCAGGAGACCCCTTACTTCACTCTGAA
GCTGGCCGGCCGCTGGCCCGTGAAGGTGATCCACACCGACAAAGGCGAGCACTTCACGAGCCAGCCCTGCTGAGAGCATGAAACAG
GCTGGTGGCCGGCATCCAGCAGGAGTTGGGATCCCTTACACCCCGAGAGCCAGGCGCTGCTGAGAGCATGAAACAG
GAGCTGAAGAGATCATCGCCAGGTGCGGACCGAGCCAGCCACTGAAGAGCCCGGCTGAGAGTGGCCCTTTTCATCCA
CAACTTCAAGCGCAAGGCGCGCATCGGCGCTACAGCGCGCGGAGCGCATCATCGAGCATCATGCGCCAGCATCCAG
CCAGGAGCTGCAGAGCAGATCATCCGCATCCAGACTTCCGCGTGTACTACCGGAGACAGCCCGAGCCCATCTGGAG
GGCCCCCGGAGCTGCTGTGAAGGGCGAGGCCGTGGTGGTATCGAGGACAGGCGAGCATCAAGGTGGTGGCCCGG
CAAGGCCAAGATCATCCGCACTACGCGCAGCAGATGGCCGGCCGAGCTGGTGGCCGGCCAGGAGGAG

FIGURE 31

PCY/US01/21241

[illegible]

FIGURE 32

WO 02/04493

40/114

PCT/US01/21241

>Pol_Tv1_c_2Aopt (SEQ ID NO:62)

TTCTTCGCGAGAACTTGGCTTCCCTCCAGGCGAGGCCCCGAGTTCCCTCCCGAGCAGACCGCGCCCAACAGCCAC
 CAGCGCACCAACAGGCCCAACAGCGCGAGCTGAGGTGGCGGGGACAAACCCCGCGCCAGGAGGGCGAGCGCGAGG
 GCACCTTCACACTTCCCTCAGATCACTCTGCGAGCGCCCTTGGTGGACATCAAGGTGGAGGGCCAGATCAAGGAGGCG
 CTGCTGGACACCGCGCGGACACCGTGTGAGAGGAGATCGACCTGCCCGGCAAGTGGAGGCCAAGATGATCGCGG
 CATCGGCGCTTCATCAAGGTGGCGCTGATGAGACAGATCTGATCGAGATCTGCGGCAAGAGCCCATCGGCACTGCTGC
 TGGTGGGCCCCACCCCGGTGACATCATCGCGCGCAACCTGCTGACCGAGCTGGGCTGACCTGCACTTCCCATCAGC
 CCGATCGAGACCGTGGCGGTGAAGCTGAAGCCCGCATGGACCGCCCAAGTGAAGAGTGGCGCTGACCTGACCTGACCTGAC
 GATCAAGGSCCTTGACCGCATCAAGAGAGAGGACAGCAACAGTGGCGAGGCTGATGAGATCAACAGATGGCGCCCGACAA
 ACCCGCGGTGTGCGCATCAAGAGAGAGGACAGCAACAGTGGCGAGGCTGATGAGATCTCGCGAGCTGAGCAAGGCG
 ACCAGGACTTCTGGAGGTGCGCTGGGCTCCCGACCCCGCGCTGAGAGAGAGAGAGAGCGTGAACCTGCTGGA
 CGTGGCGACGCTTACTTCAGCGTGGCTTGGACGAGGCTTCCGCAAGTACCGCGCTTCACCATCCCGAGCATCAACA
 AGAGAGACCCCGGCTACCGCTACCGTACAGCTGCGCGAGGCTGGAGAGGCGACCCCGCATCTTCAGAGCGAGC
 ATGACCAAGATCTGGAGGCTTCCCGCGCAAGACCCCGACATCGTGAATCAACAGTACATGAGACCTGTAAGTGGG
 CAGCGACCTGGAGATCGCGCGACCGCGCGAGATCGAGAGGCTGGCGAGGACCTGCTGAGTGGGCTTCAACACCC
 CCGCAAGAGAGACCAAGAGAGCGCGCGCTTCTGCTGATGGCTACGAGCTGCGACCCCGCAAGTGGAGCTGCGAGCC
 ATCTGCTGCGCGAGAGAGGACAGCTGGACCGTGAACGACATCCAGAGCTGGTGGGCAAGCTGACTGAGCTGGGCGAGGAT
 TACCCCGCGCATCAAGGTGGCGCGAGTGTGAGAGCTGCTGCGCGCGCGCAAGGCGCTGACCGGACCTGCTGCGCTGAGCG
 AGAGGCGCGAGCTGGAGCTGGCGCGAGAGCGCGAGATCTGCGCGCGCGCGCAAGGCGCTGAGCGCTGAGCGCTGAGCG
 GACCTGATGGCGCGAGCTCCAGAGAGCGCGCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 GAGCGCGAGAGTACCGCAAGAGGCGCGCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 TGGAGAGCATCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 GAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 GAGAGAGCGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 AGCTGAGCGAGCGCGCGCGCGCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 CTGGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 CCGCGCGAGAGCGCGCGCGCGCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 GCGCTGAGCGCGCGCGCGCGCGCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 CATCTGCGCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 ATGAGAGCGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 GAGAGTGTACCGCGCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 AGAGCTTGGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 CATCGCGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 CATCGCGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 TCACTCGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 AAGCGAGAGCGCGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT
 CTACCGCGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT

FIGURE 33

WO 02/04493

41/114

PCT/US01/21241

>Pol_TV1_C_ZAME (SEQ ID NO:63)

TTTTITAGGGAAATTTGGGCTTCCCAAGGGGAGGCCAGGGAATTTCCCTCCAGAACAGACAGACCCACAGCCCCAC
 CAGCAGAACCAACAGCCCCACAGCAGAGAGGCTTCAGGTTCCAGGGAGACAAACCCGTCGCGAGGAAGGAGAAAGAGGG
 GAACCTTTACTTTCCCTCAATCACTCTTTGGCAGGGACCCCTTGTCTCAATAAAGTAGAGGGCCAGATAAAGGAGGGCT
 CTCTTAGACACAGGAGCAGATGATACAGTATTAGAGAAATAGATTTCACAGGAAATGGAAACCAAAATGATAGGGGG
 AATTGAGGCTTTTATCAAGTAGAGACAGTATGATCAATACTTATAGAAATTTGTGGAAAAAGGCTATAGGATACAGTAT
 TAGTAGGGCTTACCCAGTCAACATAATTGGAGGAATCTGTAACTCAGCTTGGATGACACATAAATTTCCCAATTAGT
 CCTATTGAAATGTACCGTAAATTAAGAACAGGAATGGATGGCCCAAGGTCAAAACATGGCCATTGACAGAGAAAA
 AATAAAGCAATTACAGCATTTTGTAGGAATGGAGAGGAGGAGGAAAAATACAAAAATTTGGGCTTGTATATCCATATA
 ACCTCCAGTATTTCGCATTAAGAGAGGACAGTACTAAGTGGAGAAAAATAGTAGATTTTCAGGGAACTCAATAAAGA
 ACTCAAGACTTTTGGGAAGTTCAATTAGGAATACCAACCCAGCAGGATTAAGAAAGAAAAATCAGTGAACAGTGTAGA
 TGTGGGAGATGCAATTTTCAGGTTCTTTAGATGAAAGCTTCAGGAAATATACTGCAATCACCATACCTAGTATAAACA
 ATGAACACACAGGATTTAGATATCAATATATTTGCTGACAGGGATTCGAAAGCATCCACAGCAATATCCAGAGTAGC
 ATGACAAAAATCTTAGAGCCCTTCAGAGCAAAATCCAGCATATCTTCTATCAATAATATGATGATGATGATGATGATG
 ATCTGACTTAGAATATAGGCTACATAGAGCAAAATAGAGAGTTTACGGAGACTTTATGGAATGGGATTTACACAC
 CAGACAGAAACATCAAAAGAAACCCCTATTTCTTGGATGGGTATGAACTCCATCTGACAAATGGAGCTACACCT
 ATACTGCTGCCAGAAAGAGTACTTGGACTGTCAATGATATACAGAAATTACTGGGAAATTAACCTGGGCAAGTCAACCT
 TTACCCAGGCAATTAAAGTAGGCAACTCTGTAACTCTCAGCGGGGCCAAGCACTAACAGACATAGTACCACTAATCTG
 AAGAGCAGATATAGAAATGGCAGAACAGGAAATTTTAGAGAAACAGTACATGGATATTTATGATCCATCAAAA
 GACTGTGATAGTGAATACCAAGACAGGCGCATGACCAATGACATATCAATTTATCAAGAACCTTTAAAAATCTGAA
 AAGAGGAACTATGAAAAATGAGACTACCCACACTATGATGATTAAGACAGTTAACAGAGGCAATGAAAAATAGGCA
 TGGAAAGCATATGATATAGGGAAGACTCTCTAATTTTAACTACCCATCCAAAAAGAAATAGGAGCATGGTGGAGC
 GCTATTTGACAGGACACTGTGATCTCTGATGGGAGTTTGTATTAACCCCTCCCTAGTAAATATGATACCACTAG
 AAGAGATCCCATAGCAGGAGTAGAATCTTCTATGTAGATGGAGCACTAATAGGAGGCTAAAAATAGGAAAGCAGGCT
 ATGTACTGACAGGAGAGGCGAGAAATTTTACTCTACTACCAACAAATCAGAGAGTGTAGTACAAAGCAATTCAG
 CTAGCTCTGACAGGATTCAGGATCAGAAGTAAACATAGTAACAGACTACAGATATGCTTAGGAATCATTCAGCACAAC
 AGATAAGAGTGACTCAGAGATATTTAACCAATATAGAACAGTTAATAACAGGAAGAAATCTACCTGTCTAGGATC
 CAGCACAATAAGGATTTGGGGGAATGAACAGTATGATTAATTTAGTAAAGGAATTAGGAAAGGCTTTTCTCTAGAT
 GGAATAGTAAAGCTCAGAGAGAGCATGAAGGTAACCAAGCAATTTGGAGAGCAATGGCTATAGGTTTATCTGCCACC
 CATAGTAGCAAGAAAGATATAGTATGATACCTTTAGAGGGAATAATCTCTGATAGCAGTCCATGATAGTGTGCTAC
 ATGGAGCAGAGGTTATCCAGCAGAACAGGACAGAGAACAGCATATTTATATTAAATTTAGCAGGAGATGGCCAGT
 CAAAGTAAATACATACAGCAATGGCATATTTTACCACTACTGCACTTAAGGCAGCCTGTTGTGGGCAAGTATCCAAC
 AGGAATTTGGAATTCCTTACATCCCAAGTCCAGGAGTGGTAGAATCCATGATTAAGAAATTAAGAAATAATAGGA
 CAGATAGAGATCAAGCTGAGCACCCTTAAGACAGCAGTACAAATGGAGTATTCATTCACAAATTTTAAGAAAGGGGG
 AATTGGGGGTACAGTGCAGGGGAAAGAAATATAGACATATAGCAACAGACATACAACTAAAGAAATTAAGAAACAAA
 TTATAAGAAATTCAAATTTTCGGGTTATTTACAGACACAGCAGAGACCTATTTGGAAAGGACAGCCCACTACTCTGG
 AAAGGTGAAGGGTGTATTAATAGAGATTAAGGTGACATAAAGGTATACCAAGGAGAAAGCAAAATCATTAGAGA
 TTATGGAACACAGATGACAGTCTGATTTGTGGGAGGTGGACAGGATGAAGT

FIGURE 34

WO 02/04493

42/114

PCT/US01/21241

>Prot_TV1_C_Zrept (SEQ ID NO:64)

CCCCAGATCACCCCTGTGGCAGCGCCCCCTGGTGGCATCAAGGTGGAGGGCCAGATCAAGGAGGCCCTGCTGGACACCGG
CGCCGACGACACCGTGTGGAGGAGATCGACCTGCCCGCAAGTGGAGGCCAAGATGATCGGCGCATCGGCGGCTTCA
TCAGGTGGCCCGTACGACCCAGATCCGTGATCGAGATCTGGGGCAAGAGGCCATCGGCAACCTGCTGGGCCCCACC
CCCTGGACATCATCGGCCCAACCTGTGACCCAGCTGGGCTGCACCCCTGAACCTC

FIGURE 35

WO 02/04493

43/114

PCT/US01/21241

>Prot_TV1_C_Zawt (SEQ ID NO:65)

CCTCAATCACTCTTTGGCGCCGCCCTTGTCTCAATAAAAGTAGAGGGCCAGATPAGGAGGCTCTCTTAGACACAGG
AGCAGATGATACAGTATTAGAGAAATAGATTGGCAGGGAAATGGAAACCAAAATGATAGGGGGAATTGGAGGTTTTA
TCGAAGTAAGACAGTATGATCAAACTTATAGAAATTGTGGAAAAAGCTATAGGTACAGTATTAGTAGGGCCTACA
CCAGTCAACATAATTGGAAGAAATCTGTTAATCTGCTTGGATGCCACTAAATTTT

FIGURE 36

WO 02/04493

44/114

PCT/US01/21241

>Protina_TV1_C_2Rept (SEQ ID NO:66)

CCCCAGATCACCCTGTGGCAGCGCCCCCTGGTGAGCATCAGGTGGAGGGCCAGATCAGGAGGCCCCCTGTGGCCACCGG
CGCCGACGACACCGTGTGGAGGAGATCGACCTGCCCCGCAAGTGGAAAGCCAGATGATCGGCGGCATCGCGGGCTTCA
TCAGGTGCGCCGTTACGACAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCCGTGCTGTGGGCCCCACC
CCCGTGAAATCATCGGCGCCAGACCTGCTGACCCAGCTGGGCTGCACCTGAACTTC

FIGURE 37

WO 02/04493

45/114

PCT/US01/21241

>Protina_TV1_C_ZAWt (SEQ ID NO:67)

CCTCAAAATCACTCTTTGGCAGCGACCCCTTGTCTCAATAAAAGTAGAGGGCCAGATAAAGGAGGCTCTCTTAGCCACAGG
AGCAGATGATACAGTATTAGAAGAAATAGATTTGCCAGGGAAATGGAAACCAAAAATGATAGGGGAATTGGAGGTTTAA
TCRAAGTARGACAGTATGATCATAATCTATAGAAATTTGTGGAAAAAGGCTATAGGTACAGTATTAGTAGGGCTTACA
CCAGTCACATATATTGGAGAAATCTGTAACTCAGCTTGGATGCACACTAAATTTT

FIGURE 38

WO 02/04493

46/114

PCT/US01/21241

>ProtinaRtmnt_TV1_C_ZAopt (SEQ ID NO:68)

CCCCAGATCAACCTGTGTCAGCGCCCTTGGTGAACATCAAGGTGGAGGCGCAGATCAAGGAGGCCCTGTGSCCAACGG
 CGCCGAGGACACCGTGTCTGGAGGAGATCGAAGCTGCGCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCA
 TCAAGGTGCGCGCATACGACCAAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCAACCGTGTGTGCGCGCCCAAC
 CCGGTGAACATCATCGCGCAACCTGCTGAACCCAGCTGGGCTGCAACCTGAACTTCCCATCGCGCCATCGAGACCGT
 GCGCGTGAAGCTGAAGCCCGCATGGACGGCCCAAGGTGAAGCATGTGCGCCCTGACCGAGGAGAGATCAAGGCGCTGA
 CCGCCATCTGCGAGGAGATGGAGAGGAGGCGAGATCAACAGATCGGCGCCGACCAACCCCTACAGACCCCGCTTTC
 GGCATCAAGAGAGAGAGAGACAGACCAAGTGGCGAGCTGGTGGACTTCCGCGAGCTGACAGAGCGACCGAGGACTTCTG
 GAGGCTGAGCTGGGCGATCCCGACCCCGCGGCTGAAGAGAGAGAGAGAGCTGACCGGTGCTGACCTGGCGAGCGCGCT
 ACTTCAGCGTGCCTCGAGAGAGAGCTTCCGCAAGTACACCGCTTCAACATCCCGAGCATCAACAGAGAGCCCGCGC
 ATCCGCTACCAAGTACAGCTGCTGCGCCAGGCTGGAGGGGACCCCGCATCTTCAGAGAGCGCATGACCAAGATCGT
 GAGGCGCTTCCGCGCAAGACCCGACATCGTGAATCAAGCGCCCGCTTACGTGGGAGCGAGCTGGAGATCGGCG
 AGCACCGCGCTAGAGTGAAGGAGCTGCGGAGCGCTGCTGAAGTGGGCTTCAACCGCGGACAGAGAGCGAGAG
 GAGCGCCCGCTTCCGCGCATCGAGCTGACCGCGAGAGTGAAGCGTGAAGCGCATCTGCTGCGCGGAGAGAGAGAGCTG
 GAGCGTGAAGAGATCCAGAGCTGCTGGTGGCGAGCTGAAGTGGCGAGAGTCAACCGCGCATCAAGGTGCGCGAGC
 TGTGCAAGCTGCTGCGCGCGCAAGGCGCTGAAGAGAGTGTGCGCGCTGACCGAGGAGGCGAGCTGAGCTGCGCGAG
 AACCGAGAGTCTGCGCGAGCGCGTGCAGCGGTGTACTACAGCGCGAGAGAGAGCTGATGCGCGAGATCCAGAGGCA
 GCGCGAGAGAGAGAGAGTACAGATCTACAGGAGCGCTTCAAGAGAGTGAAGAGCGAGTACAGAGAGAGAGAGAG
 CCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 ACGCGAGAGTTCGCGCTGCGCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 CCGGTGGAGAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 CTTCTACGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 ATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 ACCAGATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

FIGURE 39

WO 02/04493

47/114

PCT/US01/21241

>ProteinAEMut_TV1_C_ZAw1 (SEQ ID NO:69)

CCTCAATCCTCTTTGGCAGCGACCCCTTGTCTCAATANAAGTAGAGGGCCAGATANAAGSAGCTCTCTTAGCCACAGG
AGCAGATGATACAGTATTAGAGAGAAATAGATTGGCAGGGAAATGGAAACCAAAATGATAGGGGGAATTGGAGGTTTDA
TCAGAGTAGACAGTATGATCAAACTATTATAGAAATTTGGGAGAAAGGCTATAGGTACAGTATTAGTAGGGGCTTACA
CCAGTCAACATAATTGGAGAGAAATCTGTAACTCAGCTTGGATGCCACTAAATTTTCCAAATTAGTCCATTGAAACTGT
ACCACTAAATTAAGCCAGGAATGGATGGCCAAAGGTCAAAACAATGGCCATTGACAGAGAGAAAAATTAAGCAATTAA
CAGCAATTTGTGAGGAATGGAGAGGAGGAGAAATTAAGAAATTTGGGCTGATATCCATATAACACTCCAGTATT
GCCATAAAAAAGAGGACAGTACTAGTGGAGAAATTAAGTAGATTTGAGGGAATCAATAAAGAACTCAAGACTTTG
GAGAGTTCAATTAGGAAATACCAACCCAGCAGGATTAAGAAAGAAAAATCAGTGACAGTGTCTAGATGGGGGATGCAT
ATTTTCAGTCTCTTTAGATGAAGCTTCAGGAAATATATCTGCAATTCACCATACCTAGTATTAACAATGAACACACAGG
ATTAGATATCAATATAAATGTCTGCCACAGGGATGGAAAGGATCAACAGCAATATCCAGAGTAGCATGACAAATCTCT
AGAGCCCTTCAGAGCAAAATCCAGACATGPTTATCTCAAGCCCCCTGTATGTAGGATCTGACTTAGAATACGGC
AACTAGAGCAAAAAAGAGAGTTAAGGGAACTTTATTAAGATGGGGATTTACAACACCCAGACAGAGAACATCAAAA
GACCCCTCTCTTCCCTCAGACTCCCTCTGACAAATGGACAGTACAACTATCTGCTGCCAGAAAGGATAGTTG
GACTGTCAATGATATACAGAGTTATGGGAAATTAAGCTGGGCAAGTCAAGATTTACCCAGGCAITTAAGTAGAGCAAC
TCTGTAACTCTCAGGGGGCCAGAGCTACAGACATAGTACCACTAAGTGAAGAGCAGAAITAGAAITGGCGAG
ACAGGGGAATTTTAGAGAGAACCACTACATGAGTATATATGATCCCTCAAGGACTTGATAGCTGAATACAGAAACA
GGGCAATGAACAATGGACATATCAATTTATCAGAGACATTAAGAAATCTGAAGACAGGCAAGTATGCCAAATGAGGA
CTACCCACACTAATGATGTAAACAGTTAAACAGAGGACAGTGAAGAAATAGCCATGAGAAACATAGTAATGGGAAAG
ACTCTTAATTTAGACTACCCCTCCAAAGAGACATGCAACATGCTGCAAGACTATTTGCAAGCCACCTGGATCCC
TGACTGGAGTTTCTATATACCCCTCCCTAGTAAATTAAGTACCACTAGAAAGATAGAAAGATCCCAAGAGAGGCAAAA
CTTTCTATGTAGATGGAGCACTAATAGGAACTAAGATAGAGAAAGCAGGCTATGTTACTGACAGAGAGGCAAAA
ATTGTTACTCTAATTAACACACAAATCAGAGACTGAGTTACAGAGCAATTCAGCTAGCTCTGACAGATTCAGGATCAG
AGTAAACATAGTAACAGACTACAGATATGATAGGAATCATTCAGCAACACAGATAGAGTGACTCAGAGATATTA
ACCAATATAGACAGTTTATAAACAAGGAAGAACTTACCTGTCTATGCTACCAACACATTAAGGAATTTGGGGAAAT
GACCAAGTAGATAAATAGTAACTAGGAAATTAAGGAATGTTG

FIGURE 40

PCY/US01/21241

[illegible]

•

WO 02/04493

49/114

PCT/US01/21241

>Protwt:RTwt_TV1_C_2AwT (SEQ ID NO:71)

CCTCAATCACTCTTTGGCAGCGACCCCTTGTCTCAATAAAGTAGAGGGCCAGATAAAGGAGGCTCTCTTAGACACAGG
 AGCAGATGATACAGTATTAGAGCAAAATAGATTTCAGGGGAATGGAAACCAAAATGATAGGGGAATTGGAGGTTTFA
 TCRAAGTAAACAGTATGATCAAAATCTTATAGAAATTTGGGAAAAAAGGCTATAGGTACAGTATTASTAGGSOCTACR
 CCAGTCAACATAATTGGAGGAATCTGTTRACTCAGCTTGGATGCCACATTAATTTCCAAATTAGTCCATTGGAACTGT
 ACCAGTAAATTAAGAACAGGAATGGATGGCCCAAGGTCAAAACAATGGCCATTGACAGAGAGAAAAATAAAAGCAATTA
 CAGCAATTTGTGAGAAATGGAGAGGAAGGAAAAATTACAAAAATTTGGCCCTGATATCCATATAACACTCCAGTATTT
 GCCATRAAAAGAGGACAGTACTAAGTGGGAAAAATTAGTAGATTTCAGGGAACTCAATAAAGAACCTCAAGACTTTTG
 GGAAGTTCAATTAGGANTACCAACCCCGCAGGATTAAAAAGAAAAAATCAGTGACAGTGTCTAGATGTGGGGGATGCAT
 ATTTTTCAGTTCTTTAGATGAAGCTTCAGGAAATATCTGCAATTCACCATACCTAGTATAACAAATGAACCAACAGGG
 ATTAGATATCAATATAATGTGCTGCCAGGGATGGAAAGGATCACCGCAATATTCAGAGTAGCATGACAAAAATCTT
 AGGGCCCTTCAGAGCAAAAAATCCAGACATAGTTATCTTCATATATGATGACTTGTATGTAGCATCTGACTTAGAAN
 TAGGGCAACATAGAGCAAAAAATAGAAAGTTAAGGGAACTTTATTGAAATGGGGATTACACACCCAGACAGAAACAT
 CAAAAAGAACCCCTTTCTTGGATGGGTATGAACTCCATCTCTGACAAATGGACATACACCTATCTGCTGCCAGA
 AAAGGATAGTTGGACTGTCAATGATATACAGAACTTAGTGGGAAAAATTAACCTGGGCAAGTCAAGTTTACCCAGGGATT
 AAGTAGGCAACTCTGTAACTCTTCAGGGGGCCAGAGCTAAACAGACATAGTACCACTAAGTGAAGAGCAGAAATTA
 GATTGGCAGAGACAGGGGAATTTTAGAGGAACCAAGTTCATGGAGTATATTATGATCCATCAAAAGACTTGATAGCTGA
 AATACAGAAACAGGGGCAATGAACATGGACATATCAATTTATCAAGAACCTTTAAAAATCTGAAACAGGGGAAGTAG
 CAAAAATAGAGACTCCAGACTTATGATGATGATTAACAGTTACAGAGGCACTGCAAAAAATAGCCNTGGAAAGCATAGTA
 ATATGGGAAAGACTCCCTAATTATAGACTACCCCTCCAAAAAGAAACATGGAGACATGTTGGACAGACTATTGGCAGC
 CAGCTGCTATCCCTGAGTGGAGTTTCTTAATACCCCTCCCTAGTAAATTAATGGTACCACTAGAAAAAGATCCCATAG
 CAGGATAGAAACTTTCTATGTGATGGAAGCACTAATAGGAAAGCTAAATAGGAAAGCAAGGATATCTCTACAGGA
 GGAAGGCAAGAAATTTGTATCTTAACCTAACCAACAAATCAGAAAGCTGATACAGCAATTCAGTACTCTACAGGA
 TTCCAGATCAGAACTAACATACATACCACTCAAGTATGCAATAGGAATCATTACAGCAACCCAGTACAGTACT
 CAGAGTATTTAACCAATATATAGAACAGTTAATRAACAGGAACAATCTACTCTCATCGGTACAGCACTAAGGA
 ATGGGGGAATGAACAGTAGATTAATTAGTAATAGGGAATTAGGAAATGTTG

FIGURE 42

WO 02/04493

50/114

PCT/US01/21241

>RevExon1_TV1_C_ZAopt (SEQ ID NO:72)

ATGGCCGGCCGACGCGGCCACGCCACGAGGCCCTGCTGCAGGTGGTGAAGATCATCAAGATCCTGTACCCAGAGC

FIGURE 43

WO 02/04493

51/114

PCT/US01/21241

>RevExon1_TV1_C_2Awt (SEQ ID NO:73)

ATGGCAGGAAAGAGCCGAGACAGCGACGAAGCGCTCCTCCAAGTGGTGAAGATCATCAAAATCCTCTATCAAGCA

FIGURE 44

WO 02/04493

52/114

PCT/US01/21241

>RevExon2_TV1_C_2Aopt-2 (SEQ ID NO: 74)

CCCTACCCGAGGCCCGAGGCCCGCCGCAACCGCCGCCGCCGTGGCGGCCCGCAGCGCCAGATCCA
CACCATCGCGAGCGCATCTGTGTGCTGCTGGGCCCGAGCCCGAGCGCGTGGCCCTGCAGCTGCCGCCCTGGAGC
GCTGACATCAACTGCGAGGGCGAGCGCACCAAGCGCAACCGAGAGCCAGGCCACCGAGCGGTGGGCGAC
CCCTAA

FIGURE 45

WO 02/04493

53/114

PCT/US01/21241

>Exon2_TV1_C_ZAWt (SEQ ID NO:75)

ACCCCTACCCCAAGCCCGAGGGGACTCGACAGGCTCCGAGGAATCGAAGAGAGGTGAGAGCAGACAGAGACAGATC
CATAGGATTGGTGAGCGGATTCTTGTGCTTGCTTGGGACGATCTCCGGAGCCTGTGCTTTCAGCTACCAACCGCTTGA
GAGACTTCATATTATTCAGTGAAGGCGAGTGGAACTTCTCGGACACAGCAGTCTCAGGGGACTACAGAGGGGTTGGAG
ATCCTTAA

FIGURE 46

WO 02/04493

54/114

PCT/US01/21241

RT_TV1_c_ZAopt (SEQ ID NO:76)

CCCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCCA
AGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCATCTGCG
AGGAGATGGAGAAGGAGGGCAAGATCACCAAGATCGGCCCGGACAACCCCTACAACA
CCCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTT
CCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCGAGCTGGGCATCCCCAC
CCGCGCGGCTGAAGAAGAAGAAGAGCGTGAACCGTGCTGGACGTGGCGGACGCCTAC
TTCAGCGTGCCCTGGACGAGAGCTTCGCAAGTACACCGCCTTCAACATCCCCAGCA
TCAACAACGAGAACCCCGGCATCGCTACCAAGTACAACGTGCTGCCCGAGGGCTGAA
GGGACGCCCGCCATCTTCAGAGCAGCATGACCAAGATCTGGAGCCCTTCGCGGCC
AAGAACCCTGACATCTGTGATCTACCAAGTACATGGACGACCTGTACGTGGGCAGCGACC
TGGAGATCGGCCAGCACCGGCCCAAGATCGAGGAGCTGCGCGAGCACCTGCTGAAGT
GGGGCTTCAACACCCCGACAAGAAGCACCAAGAAGGAGCCCGCTTCTGTGGATGGG
CTACGAGCTGCACCCCGACAAGTGAACCGTGCAGCCCATCTGCTGCCCGAGAAGGAC
AGCTGGACCGTGAAACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCCAGGCCAG
ATCTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGCGCGGCCCAAGGCC
TGACCGACATCTGTGCCCTGACCGAGGAGGCGGAGCTGGAGCTGGCGGAGAACCGCG
AGATCTGCGCGAGCCCGTGACCGGCGTGTACTACGACCCAGCAAGGACCTGATCGC
CGAGATCCAGAAGCAGGGCCACGAGCAGTGGACCTACCAATCTACCAAGAGCCCTT
CAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCCACCCACCAACGACGT
GAAGCAGCTGACCGGAGCGCGTGCAAGATCGCCATGGAGAGCATCGTGATCTGGGG
CAAGACCCCAAGTTCCGCTGCCATCCAGAAGGAGACCTGGGAGACCTGGTGGACC
GACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCTGTGAACACCCCGCCCTGG
TGAAGCTGTGGTACAGCTGGAGAAGGACCCCATCGCCGGCTGGAGACCTTCTACGT
GGACGGCGCCACCAACCGCGAGGCCAAGATCGGCAAGCGCGGTACGTGACCGACCG
CGGCCGCCAGAAGATCGTGACCTGACCAACACCAACCAAGAGACCGAGCTGCA
GGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGGTGAACATCGTGACCGACAG
CCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAAGAGCGACAGCGAGATCTTC
AACCAGATCATCGAGCAGCTGATCAACAAGGAGCGCATCTACCTGAGCTGGGTGCCCG
CCACAAGGCATCGCGGCCAACGAGCAGGTGGACAAGCTGGTGAGCAAGGGCATCC
GCAAGGTGCTG

FIGURE 47

WO 02/04493

55/114

PCT/US01/21241

>RT_TV1_C_ZAwE (SEQ ID NO: 77)

CCAAATTGCTCTATTGAAACTGTACCACTAAATTTAAACCCAGGAATGGATGGCCCAAAGTCAAAACAAAGGCCATTGAC
AGAGGAAAAATATAAGCATTAAACAGCAATTTGTAGGAAATGGAGAGGAAAGGAAAAATACAAAAATTTGGCCCTGATA
ATCCATATACCACTCCAGTATTGTCCATAAAAAAGAGGACAGTACTAAGTGGAGAAAAATTAGTAGAATTTCAAGGAACTC
AATTAAGAAGACTCAAGACTTTTGGGAAGTTCAATTTAGGAATACCAACCCAGCAGGAATTAAGAAAGAAAAATCAGTGAC
AGTGTAGATGTGGGGAGTCATAATTTTCAGTTCCCTTTAGATGAAAGCTTCAGGAAATATACCTGCATTCAOCATACCTTA
GTATATAACAATGAAACACCCAGGGATTAGATATCAATATATGTGTGCTGCCACAGGGATGGAAGGGATCCAGCCATATTC
CAGAGTAGCATGACAAAAATCTTAGAGCCCTTCAGAGCCAAAAATCCAGACATAGTTATCTATCAATATATGGATGACTT
GTATCTAGGATCTGACTTAGAATTAGGGCCACATAGAGCAAAAAATAGAAGAGTTAAGGGACATTTATTGAAATGGGAT
TTACAAACCCAGACAGGAACATCAAAAGGACCCCCATTTCTTTGGATGGGGTATGAACCTCATCTTGACAAATGGACA
AGTCAAGTTTACCCAGGATTAAAGTAAGGCAACTCTGTAACTCTCAGGGGGCCCAAGCACTAACAGACATAGTAC
CACTAACTGAAGAACAGGAATTAGAATTGGCAGAGAACAGGGAATTTTAGAGAACCCAGTCATGGAGTATATATGAT
CCATCAAGAGACTTGTAGATGAATACAGAAACAGGGGCATGAACAAATGGACATATCAATTTATCAGGACCATTTAA
AAATCTGAAGAACAGGCAAGTATGCAAAATGAGGACTCCCACTAATGATGTAAACAGTTAAACAGAGGCAGTGCAAA
AAATAGCCATGGAAGCATATGATATGGGGAAGACTCCCTAAATTTAGACTACCCATCCAAAGAAACATGGGAGACA
TGSTGGACAGACTATTGGCAACCCACCTGGATCCCTGAGTGGGAGTTTGTAAATACCCCTCCCTAGTAAATTTATGGTA
CCAACTAGAAAGAGTCCCATAGCAGGATAGAACTTTCTATGTAGATGGAGCACTAATAGGGAAGCTAAATAGGAA
AAGCAGGGTATGTACTGACAGAGGAGGAGCAAGAAATTTTACTCTAATACACACAAATCAGAAAGCTGAAGTTACAA
GCAATTCAGCTAGCTCTGAGGATTCAGGATCAGAGTAAACATAGTACAGCTCAGAGTATGATTAGGAATGATCA
AGCACACCCAGATAGAGTGAATCAGAGATATTACCAAAATATAGACAGTTTATAAACAGGAAAGAACTTACCTGT
CATGGGTACCGCACATAAGGATTTGGGGGAATGAACAAAGTAGATAAATTAGTAAAGGAATTAGGAAAGTGTG

FIGURE 48

PCY/US01/21241

[illegible]

FIGURE 49

WO 02/04493

57/114

PCT/US01/21241

>RTmut_TV1_C_2Awt (SEQ ID NO:79)

CCGATTAGTCTCTATTGAAACTGTACCAGTAAATTAAGCCAGGAATGGATGGCCCAAGGTCAAACAATGGCCATTGAC
AGGAGAAAAATAAAGCATTACAGCAATTTGTGAGGAAATGAGAAAGGAAGGAAAAATTAAGAAAAATGGCCCTGATA
ATCCATATAACACTCCAGTATTTGCCATAAAAAGAGGACAGTACTAGTGGAGAAAAATAGTAGATTTCAGGGGACTC
AATAAAGAACTCAGAGCTTTGGGAAGTTCRAATTAGGAATACCACACCCAGCAGGATTAAAAAGAAAAATCAGTAGAC
AGTGCTAGATGTGGGGATGCAATTTTTCAGTTCTTTAGATGAAGCTTCAGGAAATATATGCAATTCACCATACCTA
GTATAACCAATGAACACAGGGATTAGATATCAATATAATGTGCTGCCACAGGGATGGAAGGATCACCGCAATATTC
CAGAGTAGCATGACAAAAATCTTAGAGCCCTTCAGAGCAAAAAATCCAGACATAGTTATCTATCAAGCCGGTTGTATGT
AGGATCTTACTTAAGAAATAGGGGACATAGAGCAAAATAGAAAGTTAAGGGAACTTTATGAAATGGGGTTTACAA
CACCGACAAAGAACATCAAAAGAACCCCATTTCTTCCATCGAACTCCATCTGACAAATGACAGTACAACCTATA
CTGCTGCCAGAAAGGATAGTTGGACTGTCAATGATATACAGAAAGTTAGTGGGAAAAATTAAGCTGGGCAAGTCAAGTTTA
CCCAAGGATTAAGTAAGGCACTCTGTAACTCCCTCAGGGGGCCAAAGCACTAACAGACATAGTACCACTAAGTGAAG
AAGCAGAAATTGAATTTGGCAGAGAACAGGGAAATTTAAGAGAACCAATACATGGAGTATATATGATCCATCAAAAGAC
TTGATAGCTGAATACAGAAACAGGGGCAATGAACATAGGACATATCAATTTATCAGAAACCAATTAAAGAAATCTGAAAC
AGGGAAAGTATGCAAAATGAGGACTACCCACACTAATGATGTAAACAGTTACAGAGGCACTGCAAAATAGCCATGC
AAAGCATAGTATATGGGAAAGACTCCCAAAATTTAGACTACCCATCCAAAGAAACATGGGAGACATGGTGGACAGAC
TATTGGCAAGCCACCTGGATCCCTGACTCCGAGTTTGTATAACCCCTCCCTGTAAATTTATCGTACCAACTAGAAAC
AGATCCCATAGCAGGGTAGAACTTTCTATGTGATGGAGCACTAATGGGAAAGCTAAATAGGAAAGCAGGGTATG
TTACTACAGAGAGGAGGAGAAATTTGTTACTCTAATCAACACAAATCAGAAAGCTGAGTTACAGCAATTCAGCTA
GCTCTCAGAGATTCAAGATCAGAGTTAACTACTACAGACTCACAGTATGCTTGGAAATCAATCAGGCAACACAG
TAGAGTGACTCAGAGTATTTTACCAATATAGAACAGTTAATAAACAGGAAGAACTTACTCTATGGGTACAG
CACATTAAGGAATTGGGGAATGNACAAATAGATAAATAGTAAAGTAAAGGAATTAGGAAGTCTTG

FIGURE 50

WO 02/04493

58/114

PCT/US01/21241

>TatC22Exon1_TV1_C_2Aopt (SEQ ID NO:89)

ATGAGCCCGTGGACCCAGCTGAAGCCCTGGAAACACCCCGGCAAGCCAGCCAGACCCGGCAACAACTGCTTCG
CAAGCACTGCAAGCTACCTGCTGCTGCTTCAGAACCAAGGCTGGGCTCAGCTACGGCGCAAGAACGGCGCC
AGCGCCGCAAGCCCGCCCGGAGGACCAACAGAACCCCTGAGCAAGCAG

FIGURE 51

WO 02/04493

59/114

PCT/US01/21241

>TatScon1_TV3_C_ZAopt (SEQ ID NO:81)

ATGAGCCCCGTGGACCCCAAGCTGAAGCCCTGGAACACCCCGGCGCCAGCCAGACCCGCTGCAACCACTGCTTCTG
CAAGCACTGCAAGCTACCACTGCTGCTGCTTCCAGAACCAAGGCCCTGGGCATCAGCTACGGGCGCAAGAACCGCGGCC
AGCCCGCGAGGCCGCCGCCAGCGCGAGGACCAACGAAACCCCTGAGCAAGCG

FIGURE 52

WO 02/04493

60/114

PCT/US01/21241

>TatExon1_TV1_C_2Pwt (SEQ ID NO: 82)

ATGGAGCCAGTAGATCCTAAACTAAGCCCTGGAAACCATCCAGGAAGCCAACTAAAACAGCTTGTAAATAATTGCTTTG
CAAACTCTAGCTATCATTTGCTTAGTTTCTTCAGACAAAAGTTTAGGCATTTCCATATGSCAGGAAGAAAGCGAGAC
AGCGACGAGAGCGCTCCTCCAGTGTGTAGATCATCAAAATCCTCTATCAAGCGG

FIGURE 53

WO 02/04493

61/114

PCT/US01/21241

>TatExon2_TV1_C_ZAopt (SRQ ID NO:83)

CCCCTGCCTCCAGCCCGGGGACACACCGGCACCGAGGAGGCACGAGAGGTGGAGACACGACCGACCC
CTACGACTGTGTGA

FIGURE 54

WO 02/04493

62/114

PCT/US01/21241

>TatExon2_TV1_C_2Awt (SEQ ID NO: 84)

CCCTTACCCCTAGCCGAGGGGACTCGACAGGCTCGGAGGAATCGAAGAGAGGTGGAGAGCAAGACAGAGACAGATCC
ATACGATTGGTGA

FIGURE 55

WO 02/04493

63/114

PCT/US01/21241

>Vif_TV1_C_ZAept.[SEQ ID NO:85]

ATGGAGAACCGCTGGCAGGTGCTGATCGTGTGGCAGGTGGACCGCATGAAAGATCCGGGCTGGAAACAGCCTGGTGAAGCA
CCACATGTACATCAGCCGCGGCCAGCGGCTGGGTGTACCGGCCACCACTTCGAGAGCCGCCACCCCAAGGTGAGCAGCG
ABGTGCACATCCCTGGGCGACGCCCGCCTGGTGATCAAGACCTACTGGGGCCTGCGAGACCGGGCGAGCGGACTGGCAC
CTGGGCCACCGCGTGAACATCGAGTGGCGCCTGGGCGGAGTACAGCACCCAGGTGGACCCGACCTGGCGGACCACTGGAT
CCACATGCACACTTTCGACTGCTTCACCCAGAGCGCCATCCGCCAGGCCATCTGGGGCCACATCGTGTTCGCCCGCTGGG
ACTACAGGCCGCCACAGAGGTGGGCAAGCTGCAGTACCTGGCCCTGACCGCCCTGATCAAGCCCAAGAGGCGCAAG
CCCCCCTGCCCAGCGTGGCAAGCTGGTGGAGGACCGCTGGACGACCCCCAGAAGACCCGCGCCGCCGCGCAACCA
CAACATGAAAGGCCACTAG

FIGURE 56

WO 02/04493

64/114

PCT/US01/21241

->vif_TV1_C_ZAwE(SRQ ID NO:86)

ATGGGAAAACAGATGGCAGGTGCTGATTGTGTGGCAGGTGGACAGGATGAAGATTAGAGCATGGAATAGTTTAGTAAAGCA
CCATATGTATATATCAAGGAGGAGCTAGTGGATGGGTCTACAGACATCATTTTGAAGCAGACATCCAAAGTAAGTTCAAG
AAGTACATATCCCATAGGAGATGCTAGATTAGTAAATAAAACATATTGGGTTTGCAGACAGGAGAAAGAGATTGGCAT
TTGGGTCTATGGAGTCTCCATAGAAATGGAGACTGAGAGAAATACAGCACACAAGTAGACCCCTGACCTGGCAGACCCAGCTAAT
TCAATGCAATTAATTTGATTGTTTACAGAACTGCGCATAGACAAAGCCATATTAGGACACATAGTTTTCCTAGGTGTG
ACTATCAAGCAGGACATAAGAGGTAGGATCTCTGCATATCTTGGCACTGACAGCATTTGATAAAACCAAAAGAGAAAG
CCACTCTGCTAGTGTAGAAAAATTAGTAGAGGATAGTGGAAACGACCCCCAGAGAACAGGGGCCGAGAGGGAACCA
TACAATGAATGGACACTAG

FIGURE 57

WO 02/04493

65/114

PCT/US01/21241

>Vpr_TV1_C_ZAopt (SRQ ID NO: 67)

ATGGAGGCGCCCCCGAGGACCAAGGCCCCCAGCGGAGCCCTACAAACGAGTGGACCCCTGGAGATCCCTGGAGAGCTGAA
GCAGGAAGCCCTGCGGCACTTCCCCGCCCCCTGGCTGCAAGCCCTGGGCGAGTACATCTACGAGACCTAAGGCGACACCT
GGACCGGCGTGGAGCCATCATCCGCGTCTGCAAGCAGCTGCTGTTCTATCCACTTCCGCACTGGGCTGCGAGCAGCCGCG
ATCGGCATCTTGGCGAGCGCGCGCGCGCAACGGCGCCAGCCGCGAGC

FIGURE 58

WO 02/04493

66/114

PCT/US01/21241

>vpf_tv1_c_2Awc (SEQ ID NO:88)

ATGGAAAGACCCACAGACCCAGGGGCCGCGRGGGAACCATACAAATGAATGGACACTAGAGATTCTAGAAAGAACTCAA
GCAGGAAGCTGTAGACACITTCCTAGACCATGGCTCCATAGCTTAGGACCAATATATCTATGAAACCTATGGGGATCTT
GCACGGAGATTGAAGCTATATAGAGTACTGCAACAACTACTGTTCAATTCATTTTCAGATTGGATGCCACATAGCAGA
ATAGGCATCTTGCGACAGAGAGGCCAAGAAATGGAGCCATAGATCC

FIGURE 59

WO 02/04493

67/114

PCT/US01/21241

>Vpu_TV1_C_ZAopt (SEQ ID NO: 89)

ATGGTGAACCTGACCTGTTCAAGGGCGTGGACTACCGCCTGGGCTGGGGCCCTGATCGTGGCCCTGATCATCGCCAT
CATCGTGTGGACCATCGCCTACATCGAGTACCGCAAGCTGGTGGCCAGAGAGATCGACTGGCTGATCAAGCGCATCC
GCGAGCGCGCCGAGGACAGCGGCAACGAGAGCGGCGGAGAACCGAGGAGCTGAGCACCATGCTGGACATGGGCCAAGCTG
CGCCTGCTGGACGCGCAACGACCTGTAA

FIGURE 60

WO 02/04493

68/114

PCT/US01/21241

>Vpu_TV1_C_2Awt (SEQ ID NO:90)

ATGGTAGTTTAACTTTATTTAAAGGAGTAGATTATAGATTAGGAGTAGGAGCATTGATAGTAGCACTANTCATAGCAAT
AATAGTGTGGACCATAGCATATATAGATATAGGAAATTGCTAAGACAAAGAAAATAGACTGGTTAATTAAAGAAATTA
GGGAAAGAGCAGACAGTGGCAATGAGAGTGTGGGGACACAGAGATTGTCAACAATGGTGGATATGGGCATCTT
AGGCTTCTGGATGCTAATGATTTGTA

FIGURE 61

WO 02/04493

69/114

PCT/US01/21241

dna revexon1_2TV1_C_ZAop (SEQ ID NO:91)

ATGGCCGGCCGCGAGCGGCGACAGCGACGAGGCCCTGCTGCAGGTGGTGAAGATCATC
AAGATCCTGTACAGAGCCCTACCCCAAGCCCGAGGGCACCCGCCAGGCCCGCCGCA
ACCGCCGCGCCGCTGGCGCGCCGCGCAGCGCCAGATCCACCATCGGCGAGCGCAT
CCTGGTGGCCTGCCTGGGCGCGAGCGCCGAGCCGTGCCCTGCAGCTGCCCCCCTG
GAGCGCTGCACATCAACTGCAGCGAGGGCAGCGGCACCAGCGGCACCCAGCAGAGC
CAGGGCACCAACGAGGGCGTGGGCGACCCCTAA

FIGURE 62

WO 02/04493

70/114

PCT/US01/21241

dna Revexon1_2_TV1_C_ZAw1 (SEQ ID NO:92)

ATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCTCCAAGTGGTGAAGATCATC
AAAATCCTCTATCAAAGCAACCTTACCCCAAGCCGAGGGGACTCGACAGGCTCGGA
GGAATCGAAGAAGAAGGTGGAGAGCAAGACAGAGACAGATCCATACGATTGGTGAGC
GGATTCTTGTGCTTGCCTGGGACGATCTGCGGAGGCTGTGCCTCTTCAGCTACCAACG
CTTGAGAGACTTCATATTAATTGCAGTGAGGGCAGTGGAACCTTCTGGGACACAGCAGT
CTCAGGGGACTACAGAGGGGTGGGAGATCCTTAA

FIGURE 63

WO 02/04493

71/114

PCT/US01/21241

dna TatC22Exon1_2_TV1_C_ZAopt (SEQ ID NO:93)

ATGGAGCCCGTGGACCCCAAGCTGAAGCCCTGGAACCAACCCCGGCAGCCAGCCCAAG
ACCGCCGGCAACAACCTGCTTCTGCAAGCACTGCAGCTACCACTGCCTGGTGTGCTTCC
AGACCAAGGGCTGGGCATCAGCTACGGCCGCAAGAAGCCCGCCAGCCGCCGAGCG
CCCCCCCCAGCGGCGAGGACCAACCAGAACCCCTGAGCAAGCAGCCCTGCCCCAGGC
CCGCGGCGACAGCACCGGCAGCGAGGAGCAAGAAGAAGGTGGAGAGCAAGACCG
AGACCGAACCCTACGACTGGTGA

FIGURE 64

WO 02/04493

72/114

PCT/US01/21241

dna TatExon1_2_TV1_C_ZAopt (SEQ ID NO:94)

ATGGAGCCCGTGGACCCCAAGCTGAAGCCCTGGAACCAACCCGGCAGCCAGCCCAAG
ACCGCCTGCAACAACCTGCTTCTGCAAGCACTGCAGCTACCACTGCCTGGTGTGCTTCCA
GACCAAGGGCCTGGGCATCAGCTACGGCCGCAAGAAGCGCCGAGCGCCGAGCGCC
CCCCCAGCGGCGAGGACCACCAGAACCCCTGAGCAAGCAGCCCTGCCCCAGGCCC
GCGGCGACAGCACCGGCAGCGAGGAGAGCAAGAAGAAGGTGGAGAGCAAGACCGAG
ACCGACCCCTACGACTGGTGA

FIGURE 65

WO 02/04493

73/114

PCT/US01/21241

dna TatExon1_2_TV1_C_ZAwt (SEQ ID NO:95)

```
ATGGAGCCAGTAGATCCTAAACTAAAGCCCTGGAACCATCCAGGAAGCCAACCTAAA  
ACAGCTTGTAATAATGCTTTTGCAAACACTGTAGCTATCATTGTCTAGTTTGCTTCA  
GACAAAAGGTTTAGGCATTTCTATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGC  
TCCTCCAAGTGGTGAAGATCATCAAAATCCTCTATCAAAGCAGCCCTTACCCCAAGCC  
CGAGGGGACTCGACAGGCTCGGAGGAATCGAAGAAGAAGGTGGAGAGCAAGACAGA  
GACAGATCCATACGATTGGTGA
```

FIGURE 66

WO 02/04493

74/114

PCT/US01/21241

NefD125G-Myr_TV1_C_ZAopt (SEQ ID NO:96)

ATGGCCGGCAAGTGGAGCAAGCGCAGCATCGTGGGCTGGCCCGCGTGC
GAGCGCATGCGCCGCACCGAGCCCGCCGCGAGGGCGTGGGCGCCGCAAGC
CAGGACCTGGACCGCCACGCGCCCTGACCAGCAGCAACACCCCGCCACCA
ACGAGGCTGCGCTGGCTGCAGGCCCAGGAGGAGGACGGCGACGTGGGCT
TCCCGTGCGCCCCAGGTGCCCTGCGCCCATGACCTACAAGAGCGCGCT
GGACCTGAGCTTCTTCTGAAGGAGAAGGGCGGCTGGAGGGCTGATCTAC
AGCCGCAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGCT
TCTTCCCGGCTGGCAGAACTACACGAGCGCCCGCGGTGCGCTTCCGCTG
ACCTTGGGTGGTGTCTCAAGCTGGTGCCTGGACCCCGCGAGGTGAAGG
AGGCCAACGAGGGCGAGGACAACCTGCTGCTGCACCCCATGAGCCAGCAG
GCGCCGAGGACGAGGACCGCGAGGTGCTGAAGTGGAAAGTTCGACAGCCTGC
TGGCCACCGCCACATGGCCCGGAGCTGCACCCCGAGTACTACAAGGACTG
CTGA

FIGURE 67

WO 02/04493

PCT/US01/21241

75/114

ATGGCGCCCGCGGCATCCTGAAGAACTACCGCCACTGGTGGATCTGGGGCATCCT
 GGGCTTCTGGATGCTGATGATGTGCAACGTGAAGGGCTGTGGGTGACCGTGTACTA
 CGCGGTGCCCGTGGGCCGCGAGGCCAAGACCAACCTGTTCTGGGCCAGCGACGCCA
 AGGCCTACGAGAAGGAGGTGCACAACGTGTGGGCCACCCACGCTGCGTGGCCACC
 GACCCCAACCCCAAGGAGGTGATCCTGGGCAACGTGACCGAGAAGCTTCAACATGTG
 GAAGAACGACATGGTGGACCAGATGCAGGAGGACATCATCAGCCTGTGGGACCAGA
 GCCTGAAGCCCTGCGTGAAGCTGACCCCTGTGCGTGACCTGAACTGCACCAACG
 CCACCGTGAATACAAACAACACAGCAAGGACATGAAGAACTGCAGCTTCTACGTG
 ACCACGAGCTGCGCGGACAAAGAAGAAGAGGAGAAGCGCCTGTTCTACCGCCTGGA
 CATCGTGCCCTGAACAACCGCAAGAACGGCAACATCAACAACCTACCGCTGTATCA
 ACTGCAACACCAGCGCCATCACCCAGGCTGCCCAAGGTGAGCTTGGACCCCATCC
 CCAITCCACTACTGGGCCCGCGCGGCTACGCCCTGAAAGTGCAACAACAAGAAG
 TTCAACGGCATCGGCCCTGCGACAACGTGAGCACCGTGCAGTGCACCCACGGCAT
 CAAGCCCGTGTGTGAGCACCCAGCTGTCTGCTGAACGGCAGCCTGGCCGAGGAGGAGA
 TCATCATCCGCGAGCGAGAACCTGACCAACAACGTGAAGACCATCATCGTGACCTG
 AACGAGAGCATCGAGATCAAGTGACCCCGCGGCAACAACACCGCAAGAGCGT
 GCGCATCGGCCCGCGCCAGGCTTCTACGCCACCGGCGACATCATCGGCGACATCC
 GCCAGGCCCATGCAACATCAGCAAGAAGAGTGGAAACACCACCTGACGCGCGTG
 AGCCAGAAAGCTGCGAGGAGCTGTTCCTCAACAGCACCGGCATCAAGTTCGCCCTCA
 CAGCGCGCGCAGCTGGAGATCACCCACACAGCTTCAACTGCGGCGCGGAGTTC
 TCTACTCAACACCACCGACCTGTTCAACAGCACCTACAGCAAGGCGACCTGCACCA
 ACGGCACCTGCATGAGCAACAACCGAGGCGATCACCTGCAGTGCAGCATCAAG
 CAGATCATCAACATGTGGCAGGAGTGGGCCGCGCATGTACGCCCTCCCATCGC
 CGGCAACATCACCTGCCCGAGCAACATCACCGGCTGTCTGCTGACCCCGGACCGCG
 GCGACAACAACACCGAGACCGAGACCTTCCGCCCGGCGGCGGACATGCGCGAC
 AACTGGCGCAGCGAGCTGTACAAGTACAAGGTGGTGGAGATCAAGCCCTGGGCGT
 GGCCCCACCGCGGCAAGCGCGCTGTGTGGAGCGCGAGAAGCGCGCGTGGGCA
 TCGGCGCGCTGTCTGGCTTCTTGGGCGCGCGCGCAGCACCATGGGCGCGGCCA
 GCATCAACCTGACCGTGCAAGGCCCGGCACTGTGTGAGCGCATCGTGACAGCAGCAG
 AGCAACCTGCTGCGCGCCATCGAGGCCAGCAGCACATGTGACGCTGACCGTGTG
 GGGCATCAAGCAGCTGCAGGCCCGCTGTGTGGCATGGAGCGCTACCTGCAGGACC
 AGCAGCTGTGGGCTGTGGGCTGCAGCGGCAAGCTGATCTGCACCAACCAACGTG
 CTGTGGAACAGCAGCTGGAGCAACAAGACCCAGAGCGACATCTGGGACAACATGAC
 CTGGATGCACTGGGAACCGGAGATCAGCAACTACACCAACACCATCTACCGCTGC
 TGGAGGACAGCCAGAGCCAGCAGGAGCGCAACGAGAAGGACCTGCTGCCCTGGA
 CCGCTGGAAACAACCTGTGGAACCTGGTTACGATCAACCAACTGGCTGTGGTACATCAA
 GATCTTCAATCATGATCGTGGGCGGCTGATCGGCTGCGCATCATCTTCGCCGTGCT
 GAGCTGTGTGAACCGGCTGCCAGGGCTACAGCCCTGAGCTGCAGACCTGA
 TCCCCAACCCCGCGGCCCGACCGCTGGCGGCGATCGAGGAGGAGGGCGGCGAG
 CAGGACAGCAGCCGACGATCCGCTGTGTGAGCGCTTCCTGACCTGCCCTGGGA
 CGACCTGCGCAGCCTGTGCTGTCTGTCTACCAACCGCTGCGGAGCTTCACTCTGAT
 CGTGTGCGCGCGTGGAGCTGTGGGCCACAGCAGCTGCGCGGCTGCAGGCGG
 GCTGGGCGACCTGAAGTACCTGGCAGCCTGGTGCAGTACTGGGCTGTGAGCTG
 AAGAAGAGCGCCATCAACCTGCTGGACACCATCGCCATCGCGTGGCGAGGGCAC
 CGACCGCATCTGGAGTTCATCCAGAACCTGTGCCGCGCATCCGCAACGTGCCCG
 CCGCATCCGCGAGGGCTTCGAGGCGCGCTGCAGTAA

FIGURE 68

WO 02/04493

76/114

PCT/US01/21241

ATGAGAGCGAGGGGGATACTGAAGAATTATCGACACTGGTGGATATGGGGCATCTT
 AGGCTTTTGGATGCTAATGATGTGTAATGTGAAGGCTTGTGGGTCACAGTCTACTA
 CGGGGTACCTGTGGGGAGAGAAGCAAAACTACTCTATTTTGTGCATCAGATGCTA
 AAGCATATGAGAAAGAAAGTGCATAATGTCTGGGCTACACATGGCTGTGTACCCACA
 GACCCCAACCCACAAGAAGTGAATTTGGGCAATGTAACAGAAAATTTAACATGTG
 GAAAAATGACATGGTGGATCAGATGCAGGAAGATATAATCAGTTTATGGGATCAAA
 GCCTTAAGCCATGTGTAAAATTGACCCCACTCTGTGFCACCTTAAACTGTACAAAATG
 CAACTGTTAACTACAATAATACCTCTAAAAGACATGAAAAATTGCTCTTTCTATGTAA
 CCACAGAAATTAAGAGATAAGAAAAAGAAAGAAAATGCACCTTTTATAGACTTGAT
 ATAGTACCACTTAATAATAGGAAGAATGGGAATATTAACAACATATAGATTAATAAA
 TTGTAATACCTCAGCCATAACACAAGCCTGTCCAAAAGTCTCGTTTGACCCAATTCC
 TATACATTATTGTGCTCCAGCTGGTTATGCGCCTCTAAAATGTAATAATAAGAAATT
 CAATGGAATAGGACCATGCGATAATGTGAGCACAGTACAATGTACACATGGAATTA
 AGCCAGTGGTATCAACTCAATTACTGTTAAATGGTAGCCTAGCAGAAGAAGAGATA
 ATAAATTAGATCTGAAAAATCTGACAAACAATGTCAAAACAATAATAGTACATCTTAAT
 GAATCTATAGAGATTAAATGTACAAGACCTGGCAATAATACAAGAAAGAGTGTGAG
 AATAGGACCAAGACAAGCATTCTATGCAACAGGAGACATAATAGGAGATATAAGAC
 AAGCACATTGTAACATTAGTAAAAATGAATGGAATACAACTTTACAAAAGGTAAGT
 CAAAAATTACAAAGAACTCTCCCTAATAGTACAGGGATAAAAATTTGCACACACTCA
 GGAGGGGACCTAGAAAATTACTACACATAGCTTTAATTGTGGAGGAGAATTTTCTAT
 TGCATACACAGACCTGTTAATAGTACATACAGTAATGGTACATGCCTAATGGT
 ACATGCTATGCTTAATAATACAGAGCGCATCACACTCCAATGCAGATAAAACAAAAT
 TATAAACATGTGGCAGGAGGTAGGACGAGCAATGTATGCCCTCCCATTTGCAGGAA
 ACATAACATGTAGATCAAAATATTACAGGACTACTAATAACACGTGATGGAGGAGAT
 AATAATACTGAAACAGAGACATTACAGCCTGGAGGAGGAGACATGAGGGACAATTG
 GAGAAGTGAAATATATAAATACAAGGTGGTAGAAAATTAACCATIAGGAGTAGCAC
 CCACTGCTGCAAAAAGGAGAGTGGTGGAGAGAGAAAAAGAGCAGTAGGAATAGG
 AGCTGTGTTCTTGGGTCTTGGGAGCAGCAGGAAGCACTATGGGCGCAGCATCAAT
 AACGCTGACGGTACAGGCCAGACAATTATTGTCTGTTATAGTGCAACAGCAAAAGTA
 ATTTGCTGAGGGCTATAGAGGCGCAACAGCATATGTTGCAACTACGGTCTGGGGC
 ATTAAGCAGCTCCAGGCAAGAGTCTGGCTATAGAGAGATACCTACAGGATCAACA
 GCTCCTAGGACTGTGGGGCTGCTCTGGAAAACCTCATCTGCACCACTAATGTGCTTTG
 GAACTCTAGTTGGAGTAATAAACTCAAAGTGATATTGGGATAACATGACCTGGAT
 GCAGTGGGATAGGGAATTAAGTAATTACACAAACACAATATACAGGTGCTTGAAG
 ACTCGCAAAGCCAGCAGGAAAGAAATGAAAAAGATTACTAGCATTTGACAGGTGG
 AACAACTCTGTGGAATTGGTTTAGCATAACAAATTGGCTGTGGTATATAAAAAATTC
 ATAATGATAGTAGGAGGCTTGATAGGTTTAAGAATAATTTTGTGCTGCTCTCTA
 GTAAATAGAGTTAGGCAGGGATACTCACCTTGTCAATTGCAGACCTTATCCCAAAC
 CCGAGGGGACCCACAGGCTCGGAGGAATCGAAGAAGAAGGTGGAGAGCAAGACA
 GCAGCAGATCCATTGATTAGTGAGCGGATTCTTGACACTTGCTGGGACGACCTAC
 GAAGCCTGTGCTCTTCTGCTACCAACCGATTGAGAGACTTCATATTAATTGTAGTGA
 GAGCAGTGGAACCTCTGGGACACAGTAGTCTCAGGGGACTGCAGAGGGGGTGGGGA
 ACCCTTAAGTATTTGGGGAGTCTTGTGCAATATTGGGCTAGAGTTAAAAAGAGT
 GCTATTAATCTGCTTGATACTATAGCAATAGCAGTAGCTGAAGGAACAGATAGGATT
 CTAGAATTACAAAAACCTTTGTAGAGGTATCCGCAACGTACCTAGAAGAATAAG
 ACAGGGCTTCGAAGCAGCTTTGCAATAA

FIGURE 69

WO 02/04493

77/114

PCT/US01/21241

Gag_TV2_C_ZAopt (SEQ ID NO:99)

ATGGGCGCCGCGCCAGCATCCTGCGCGGCGGCAAGCTGGACAAGTGGGAG
AAGATCCGCTGCGCCCGGCGGCGCAAGCACTACATGCTGAAGCACCTGG
TGTGGGCCAGCCGCGAGCTGGAGCGCTTGGCCGTGAACCCCGGCTGCTGGA
GACCAGCGACGGCTGCCGCCAGATCATCAAGCAGCTGCAGCCCGCCCTGCAG
ACCGGCACCGAGGAGATCCGCAGCCTGTCAACACCGTGGCCACCCGTGACT
GCGTGACAAAGGGCATCGACGTGCGGACACCAAGGAGGCCCTGGACAAGA
TCGAGGAGGAGCAGAACAAGTGCCAGCAGAAGACCCAGCAGGCCGAGGCCG
CCGACAAGAAGGTGAGCCAGAACTACCCCATCGTGACAGAACCTGCAGGGCC
AGATGGTGCACCAAGCCATCAGCCCCGACCCCTGAACGCGCTGGGTGAAGGT
GATCGAGGAGAAGGCCATCAGCCCGAGGTGATCCCATGTTCACCGCCCTG
AGCGAGGGCGCCACCCCCAGGACCTGAACACCATGCTGAACACCGTGGGC
GGCCACCAAGGCCCATGTCAGATGCTGAAGGACACCATCAACGAGGAGGCC
GCCGAGTGGGACCCCTGCACCCCGTGACGCCGCCCGGTGGCCCCCGGCC
AGATGCGCGAGCCCCGCGGACGACATCGCCCGCACCAACAGCACCCCTGCA
GGAGCAGATCGCTGGATGACCAACAACCCCCCATCCCGTGGGCGACATC
TACAAGCGCTGGATCATCCTGGGCGTGAACAAGATCGTGCGCATGTACAGCC
CCGTGAGCATCCTGGACATCAAGCAGGGCCCCAAGGAGCCCTTCCGCGACTA
CGTGGACCCCTTCTTCAAGACCCCTGCGCGCGGAGCAGACCCAGGAGGTG
AAGAAGTGGATGACCGACACCCTGCTGGTGCAAGGCCAACCCGACTGCA
AGACCATCCTGCGCGCCCTGGGCCCGCGCGCAGCCTGGAGGAGATGATGAC
CGCCTGCCAGGGCGTGGGCGGCCCGAGCCACAAGGCCCGCGTCTGGCCGAG
GCCATGAGCCAGGCCAACAAACACCAAGCGTGATGATCCAGAAAGCAACTTC
AAGGGCCCCCGCGCGCGTGAAGTGCTTCAACTGCGGCCGCGAGGCCACA
TCGCCCGCAACTGCCGCGCCCCCGCAAGCGCGGCTGCTGGAAGTGCGGCAA
GGAGGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCCTGGG
CAAGATCTGGCCAGCCACAAGGGCCGCCCGGCAACTTCCTGCAGAGCGG
CCCGAGCCCAACCGCCCCCCCCTGGAGCCCAACCGCCCCCGCGGAGAGCT
TCAAGTTCAAGGAGACCCCAAGCAGGAGCCCAAGGACCGCGAGCCCTGA
CCAGCCTGAAGACCTGTTGGCAGCGACCCCTGAGCCAGTAA

FIGURE 70

WO 02/04493

78/114

PCT/US01/21241

Gag_TV2_C_ZAwt (SEQ ID NO:100)

ATGGGTGCGAGAGCGTCAATATTAAGAGGGGGAAAAATTAGACAAATGGGAA
AAAATTAGGTTACGGCCAGGGGGGAGAAAAACACTATATGCTAAAAACACCTA
GTATGGGCAAGCAGAGAGCTGGAAAGATTTGCAGTTAACCTGGCCTTTTAG
AGACATCAGACGGATGTAGACAAATAATAAAACAGCTACAACCAGCTCTTCA
GACAGGAACAGAGGAAATTAGATCATTATTTAACACAGTAGCAACTCTCTAT
TGTGTACATAAAGGGATAGATGTACGAGACACCAAGGAAGCCTTAGACAAG
ATAGAGGAGGAACAAAACAAATGTCAGCAAAAAACACAGCAGGCGGAAGCG
GCTGACAAAAAGTCAAGTCAAAATTATCCTATAGTGCAGAACCTCCAGGGC
AAATGGTACACCAGGCCATATCACCTAGAACCTTGAATGCATGGGTAAAAGT
AATAGAGGAGAAGGCTTTTAGCCAGAGGTAATACCCATGTTACAGCATT
TCAGAAAGGAGCCACCCACAAAGATTTAAACACCATGTTAATACAGTGGGG
GACATCAAGCAGCCATGCAAAATGTTAAAGATACCATCAATGAGGAGGCTGC
AGAAATGGGATAGGTTACATCCAGTACATGCAGGGCCTGTTGACCCAGGCCAG
ATGAGAGAACCAAGGGGAAGTGACATAGCAGGAACCTACTAGTACCTTCAA
GAACAAATAGCATGGATGACAAGTAACCCACCTATCCCAGTAGGGACATCT
ATAAAAGGTGGATAATCTGGGGTTAAATAAAATAGTAAGAAATGTACAGCCC
TGTCAGCATTTTAGACATAAAACAAGGACCAAAAGGAACCTTTAGAGACTAT
GTAGACCGGTTCTTCAAACTTTAAGAGCTGAACAATCTACACAAGAGGTAA
AAAATTGGATGACAGACACCTTGTAGTCCAAATGCGAACCCAGATTGTAA
GACCATTTTAAAGAGCATTAGGACCAGGGGCTTCATTAGAAGAAATGATGACA
GCATGTCAGGGAGTGGGAGGACCTAGCCACAAAGCAAGAGTTTGGCTGAG
GCAATGAGCCAAGCAAACAATACAAGTGTAATGATACAGAAAAGCAATTTA
AAGGCCCTAGAAGAGCTGTTAAATGTTTCACTGTGGCAGGGAAGGGCACAT
AGCCAGGAATTGCAGGGCCCTAGGAAAAGGGGCTGTTGGAATGTGGA
GGAAAGGACACCAATGAAAGACTGTACTGAGAGGCCAGGCTAATTTTATAGG
AAAATTGGCTTCCACAAAGGGGAGGCCAGGGAATTTCTTCAGAGCAGAC
CAGAGCCAACAGCCCCCCTAGAACCAACAGCCCCACAGCAGAGAGCT
TCAAGTTCAAGGAGACTCCGAAGCAGGAGCCGAAAGACAGGGAACCTTTAA
CTTCCCTCAATCACTCTTTGGCAGCGACCCCTGTCTCAATAA

FIGURE 71

WO 02/04493

79/114

PCT/US01/21241

Nef_TV2_C_ZAopt (SEQ ID NO:101)

ATGGGCGGCAAGTGGAGCAAGAGCAGCATCATCGGCTGGCCCGAGGTGCGC
GAGCGCATCCGCCGACCCGCAGCGCGCCGAGGGCGTGGGCAGCGCCAGC
CAGGACCTGGAGAAGCACGGCGCCCTGACCAACAGCAACACCGCCACAAC
AACGCCGCTGCGCCTGGCTGGAGGCCAGGAGGAGGAGGGCGAGGTGGGC
TTCGCCGTGCGCCCGAGGTGCCCTGCGCCCATGACCTACAAGGCCGCAT
CGACCTGAGCTTCTTCTGAAGGAGAAGGGCGGCTGGAGGGCTGATCTAC
AGCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGC
TTCTTCCCGACTGGCAGAACTACACCCCGGCCCGGGCTGCGCTTCCCGCT
GACCTTCGGCTGGTACTTCAAGCTGGAGCCCGTGACCCCGCGAGGTGGAG
GAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCATGAGCCAGCAC
GGCATGGAGGACGAGGACCGCGAGGTGCTGCCGTGGAAATTTCGACAGCAC
CTGGCCCGCGCCACATGGCCCGCGAGCTGCACCCGAGTACTACAAGGACT
GCTGA

FIGURE 72

WO 02/04493

80/114

PCT/US01/21241

Nef_TV2_C_Zawt (SEQ ID NO:102)

ATGGGGGGCAAGTGGTCAAAAAGCAGTATAATTGGATGGCCTGAAGTAAGA
GAAAGAATCAGACGAAC TAGGTCAGCAGCAGAGGGAGTAGGATCAGCGTCT
CAAGACTTAGAGAAACATGGGGCAGTTACAACCAGCAACACAGCCCACAAC
AATGCTGCTTGCGCCTGGCTGGAAGCGCAAGAGGAGGAAGGAGAAGTAGGC
TTTCCAGTCAGACCTCAGGTACCTTTAAGACCAATGACTTATAAAGCAGCAAT
AGATCTCAGCTTCTTTTAAAGAAAAGGGGGGACTGGAAGGGTTAATTAC
TCCAAGAAAAGGCAAGAGATCCTTGATTGTGGGTTTATAACACACAAGGCT
TCTTCCCTGATTGGCAAACTACACACCGGGACCAGGGGTCAGATTTCCTCT
GACCTTTGGATGGTACTTCAAGCTAGAGCCAGTCGATCCAAGGGAAGTAGAA
GAGGCCAATGAAGGAGAAAACAAC TTTTACTACCCCTATGAGCCAGCATG
GAATGGAGGATGAAGACAGAGAAAGTATTAAGATGGAAGTTTGACAGTACGC
TAGCACGCAGACACATGGCCGCGAGCTACATCCGGAGTATTACAAAGACTG
CTGA

FIGURE 73

WO 02/04493

81/114

PCT/US01/21241

Pol_TV2_C_ZAopt (SEQ ID NO:103)

TTCTTCCGCGAGAACTGGGCTTCCCCAGGGCGAGGCCCGCGAGTTCCCGAGCGAGCAGACC
 CGCCCAACAGCCCAACCCGCAACCAACAGCCCAACAGCCCGAGCTGCAGGTGCAGGG
 CGACACGCGAGGCGCGGCCGAGCGGCAAGGCACTTCAACTTCCCGCAGATCACTCTGTGGC
 AGCGCCCGCTGGTGAAGATCAAGGTGGCGGCGCAGACCAAGGAGGCGCTGTGGACACCGGC
 CGCGAGCGACACCGTGTGGAGGAGATCAACTTCCCGGCGCAAGTGGAGCGCAAGATGATCGG
 CGGATCGGCGGCTTATCAAGGTGGCGCAGTACGACCGAGATCTGATGAGATCTGGCGCA
 AGCGCGCATCGGCACCGTGTGTGGGCGCCACCCCGTGAACATCATCGGCCCAACCTGTC
 TGAACCGAGCTGGGCTGCAACCTGAACCTTCCCAATCAGCCCATCGAGACCGTGGCGGTAAAGC
 TGAAGCCCGCATGGACGGCCCAAGGTGAAGCAGTGGCGCTGACCGAGGAGAAATCAAG
 GCGCTGACCGAGATCTGGAGGAGATGGAGAGGAGGGCAAGATCAACCAAGATCGGCCCGG
 AGAACCCCTACAAACCCCGCTGTTCGCGATCAAGAGAAAGGACAGCAACCAAGTGGCGCAAG
 CTGGTGGACTTCCGCGAGCTGAACAAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGCAAT
 CCGCCACCGCGCGCTGAAGAAAGAAAGAGCGTGACCGTGTGGACGTGGGCGAGCGCT
 ACTTCAGCGTGGCGCTGGACGAGAGCTTCCGCAAGTACCGGCTTCAACCATCCCGAGCATCA
 ACAACGAGACCCCGGCTACCGTACAGTACACGTGCTGCGCCAGGCGTGAAGGCGAGC
 CCGCGCATCTTCAGAGCAGCATGACCGCATCTCGGAGCGCTTCCGCAACCGAGAACCCCGAG
 GTGGTATCTACCACTACATGGACCGCTGTACGTGGGCGAGCGCTGTGAGATCGGCCAGCA
 CCGCGCAAGATCGAGGAGCTGGCGGCCACCTGCTGAAGTGGGCTTCAACCCCGCGACA
 AGAAGCAACAGAGAGCGCCCTCTCTGTGGATGGGCTACGAGCTGCACCCGACAAAGTGG
 ACGTGCAGCGCATTCAGCTGCGGAGAGAGGAGAGCTGGAACGTGAACGACATCCAGAAAGCT
 GGTGGCGAAGCTGAACCTGGGCGCGAGATCTACCCCGCATCAAGGTGGCGAGCTGTGCA
 AGCTGTGGCGGCGCAAGGCGCTGACCGCATCTGTGGCGCTGACCGAGGAGGCGAGCTG
 GAGCTGGCGGAGAACCGCGAGATCTGGAAGGAGCGCTGCGACGGGTGTACTACGACCCAG
 CAAGGACCTGTATCGCGAGATCCAGAGCAGGCGAACGACCAAGTGGACCTACAGATCTACC
 AGGAGCCCTTCAAGAACCTGGCGACCGGCGAGTACGCGCAAGATGGCGACCGCCACACCAAC
 GACGTGAAGCAGCTGGCGGAGCGCTGCAGAGATCAACCGAGGAGAGCATCTGTATCTGGGG
 CAAGACCCCAAGTTCCGCTGCGCATCCCAAGGAGACCTGGGAGACCTGTGTGGAGCGACT
 ACTGGCAGGCGCACTGGATCCCGAGTGGGAGTTCTGTGAACACCCCGCGCTGTGTGAAGCTGT
 GGTACAGCTGGAGAGAGCGCATCTGTGGCGCGCGAGACCTTCTACGTGGAGCGCGCGCGC
 AACCGCGAGACCAAGATCGGCAAGGCGCGCTACGTGACCGACAAGGCGCGCGCAAGAGTGG
 TGAAGTTCAACGAGACCAACCAAGAGAGCGAGCTGACGGCATCCAGCTGGCGCTGCAAG
 GACAGCGCGCGCGAGGTGAACATCTGTGACCGACAGCGCATCGCGCTGGGCTATCCAGGC
 CCAGCGCGACAAGAGCGAGAGCGAGCTGGTGAAGCAGATCATGAGCGAGCTGATCAAGAG
 GAGAAAGGTGTACCTGAGCTGGTGGCGCGCGCAAGAGGCGATCGCGCGCAACGAGCAAGTGA
 CAAGCTGGTGAAGCGCGCATCCGCAAGGTGTCTTCTGTGACCGCATCGACAAGGCGCGAGG
 AGGAGCAGAGAAAGTACCAAGCAACTGGCGCGCATGGCGAGCGAGTTCAACCTGCCCGCC
 ATCTGTGGCAAGGAGATCTGTGGCGAGCTGGCAAGTGGCGAGCTGAAGGCGGAGCGCATGCA
 CGCGCAAGTGGACTGCGAGCGCGCGCATCTGGCAGCTGGACTGCACCCAGCTGGAGGCGCAAG
 TCATCTGTGTGGCGTGCACGTGGCGAGCGGCTACATGGAGGCGCGAGTGTATCCCGCGCGAG
 ACCGCGCAGGAGACCGCTACTTCACTGAAAGCTGGCGCGCGCTGGCGCGTGAAGGTGATC
 CACACCGACAACCGCGAGCACTTCAACGACCGCGCTGAAGCGCGCTGTGTGTGGCGCGA
 CATOCAGCGCGAGTTTCGGCATCCCTACAAACCCCGAGAGCGAGGCGTGTGGAGAGCATGA
 ACAAGGAGCTGAAGAGATCATCGGCCAGGTGGCGGACAGGCGCGAGCACTGAAGACCGCG
 GTGCAGATGGCGGTGTTTATCCCAACTTCAAGCGCAAGGCGCGCATGGCGCGCTACAGCGC
 CGCGAGCGCATCATCGACATCATCGCGAGCGCATCCAGACCAAGGAGCTGCAAGAGCAGA
 TCATCAAGATCCAGAACTTCCGCGTGTACTACCGCGACAGCGCGACCCCATCTGGAAGGCGC
 CGCGCAAGCTGTGTGGAAGGCGGAGGCGCGCTGTGTGTATCCAGGACAAACAGCGCATCAAG
 GTGTGCCCCCGCGCAAGGCGCAAGATCATCAAGGACTACGGCAAGCAGATGGCGCGCGCGCA
 CTGGCTGGCGCGCGCGAGGAGGAG

FIGURE 74

WO 02/04493

82/114

PCT/US01/21241

Pol_TV2_C_ZAwI (SEQ ID NO:104)

TTTTTATGGGAAAAATTTGGCCCTCCCAAGGGGAGGOCAGGGAATTTCTTCAGAGCAGACC
AGAGCCCAACAGCCCCACCACTAGAACCAACAGCCCCCAGCAGAGAGCTTCAAGTTCAAGG
AGACTCCGAAAGCAGGAAGCGAAAGACAGGGGACCTTTAACTTCCCTCAAAITCACTCTTTGGCA
GCGACCCCTTGTCTCAATAAAAGTAGGGGGCCAAACAAAGGAGGCTCTTTTAGATACAGGAG
CAGATGATACAGTACTAGAGAAATAAACTTGCCAGGAAAAATGGAAACCAAAAAATGATAGG
AGGAATTTGGAGGTTTTATCAAGTAAGACAGTATGATCAAACTTATAGAAATTTTGGGAAA
AAGGGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCAACATAATTGGAGGAAATCTGTT
GACTCAGCTTGGATGACACTAAATTTTCAATTAGCCCCATTGAAACTGTACCAAGTAAATTT
AAAGCCAGGAATTTGGATGGCCAAAGGTTAAACAAATGGCCATTGACAGAGAAAAATAAA
GCATTAAACAGAAATTTGTAGGAAATGGAGAGGAAAGGAAAAATACAAAAATTTGGGCTGA
AAATCCATATAACACTCCAGTATTGGCCATAAGAAAGAGGACAGTACAAAGTGGAGAAAT
TAGTAGATTTTCAAGGAACTCAATAAAAGAACTCAAAOACTTTTGGGAAGTCCAATTAGGAATA
CCACACCCAGCAGGGTTAAAAAGAAAAAATCAGTGACAGTACTGGATGTGGGAGATGCATA
TTTTTCAGTCCCTTTAGATGAGAGCTTCAAGAAAAATATAGTGCATTACCATACCTAGTATAAAC
AATGAAACACCCAGGGATTAGATATCAATATAATGTCTTCCACAGGGATGAAAGGATCAACC
AGCAATATTTCCAGAGTAGCATGACAGAAATCTTAGAGCCCTTTAGAAACAAAAACCCAGAAAG
TAGTTATCTATCAATAATATGAGTAGCTTATATGTAGGATCTGACTTAGAAATAGGGCAACATA
GAGCAAAAAATAGAGGAGTTAAGAGGACACCTATTGAAAATGGGGATTACCAACACCAAGACAG
AAACATCAGAAAGAAACCCCAATTTCTTGGATGGGGTATGAATCCATCTGACAAATGGACA
GTACAGCTATACAGCTGCCAGAAAAAGGAGAGCTGGAGCTGTCAATGATATACAGAAAGTTAGT
GGGAAAGTTAAACTGGGCAAGTCAGATTTACCCAGGGATTAAAGTAAAGGCAACTGTGTAAC
TCCCTTAGGGGAGCCAAAGCACTAACAGACATAGTCCCACTGACTGAAAGAAAGCAGAAATAGAA
TTGGCTGAGAACAGGGAAATTTCTAAAAAGAACAGTACATGAGTATATTATGAGCCATCAAA
AGATTTAATAGCTGAAATACAGAAACAGGGGAAATGACCAATGGACATATCAAAATTTACCAAG
AACCATTTAAAAATTTGAGAACAGGAAAGTATGCAAAAAATGAGGACTGCCACACTATGAT
GTGAAACAGTATAGCAGAGCTGCAAAAGATTAACCAAGAAAGCATAGTATATGGGAA
AACTCTTAATTTAGACTACCCATCCAAAGAAACATGGGAGACATGGTGGTCAGACTATT
GGCAAGCCACCTGGATTTCTGAGTGGGAGTTTGTCAATACCCCTCCCTAGTAAAAATTTGGT
ACCACTGGAAAAAGAACCCATAGTAGGGGCAAAACTTTCTATGATAGTGGAGCAGCCAAT
AGGAAACTAAAAATAGGAAAAAGCAGGGTATGTCACTGACAAAGGAAGGCAGAAAGTTGTTTC
CTTCACTGAACAAACAAATCAGAGACTGAATTAACAGCAATTCAGTACCTTTGCAAGGATTC
AGGGCCAGAAAGTAAACATAGTAACAGACTCAGATGCAATTAGGAATCATTCAAGCACAAC
CAGATAAGAGTGAATCAGAAATAGTCAGTCAAAATATAGAACAGTTGATAAAAAGGAAAAA
GTCTACCTATCTGGGTACCAAGCACATAAAGGAATTGGAGGAAATGAACAAAGTACACAAATT
AGTAAGTAGGGAATCAGAAAAAGTACTGTTTCTAGATGGAATAGATAAGCTCAAGAAAGGC
ATGAAAAAATACACAGCAATTGGAGAGCAATGGCTAGTGAATTTAACTTGCACCCATAGTA
GCAAAAGGAATAGTAGCCAGCTGTGATAATGTCAAGCTAAAGGGGAAAGCCATGCAATGACA
AGTCGACTGTAGTCCAGGAATATGGCAATTAGACTGTACACATTTAGAAAGGAAAAATCATCTCT
AGTAGCAGTCCATGTAGCCATGGCTACATGGAAGCAGAGGTTATCCAGCAGAAACAGGAC
AAGAAACAGCATACTTTACTATAAATTAGCAGGAAGATGGCCAGTCAAGTAATACATACA
GATTAATGGAGTAATTTACCAAGTACCCAGTAAAGGAGCTGTTGGTGGGAGATATCCAA
CGGGAATTTGGAAATTCCTACCAATCCCAAGTCAAGGAGTAGTAGAATCCATGAATAAAGA
ATTAAGAAATCATAGGCAAGTAAGAGATCAAGCTGAGCACTTAAGACAGCAGTACAAA
TGGCAGTATTTCATTCAAAATTTTAAAGAAAGGGGGGATTGGGGGTACAGTGCAGGGGAG
AGATAATAGACATAATAGCATCAGACATAAACTAAAGAAATTAACAAAAACAAATTTAA
AATTCAAAATTTTCGGGTTTATTACAGAGACAGCAGAGACCTATTGGAAAGGACGACCA
ACTACTCTGGAAAGGTGAAGGGGAGTATATACAGATAATAGTATATAAGGTAGTAC
CAAGTAAGAAAGCAAAAAATCATTAAAGGACTATGAAAAACAGATGGCAGTCTGATTGTGTG
GCAGGTACAGAGGATGAAGAT

FIGURE 75

WO 02/04493

83/114

PCT/US01/21241

RevExon1_TV2_C_ZAopt (SEQ ID NO:105)

ATGGCCGGCGCAGCGGGACAGCGACGAGGCCCTGCTGCAGGCCATCAAG
ATCATCAAGATCCTGTACCAGAGC

FIGURE 76

WO 02/04493

84/114

PCT/US01/21241

RevExon1_TV2_C_ZAw1 (SEQ ID NO:106)

ATGGCAGGAAGAAGCGGAGACAGCGACGAAGCGCTCCTCCAAGCAATAAAG
ATCATCAAGATCCTCTACCAAAGCA

FIGURE 77

WO 02/04493

85/114

PCT/US01/21241

RevExon2_TV2_C_ZAopt (SEQ ID NO:107)

CCCTACCCCAAGCCCGAGGGCACCCGCCAGGCCCGCCGCAACCGCCGCGCC
GCTGGCGCGCCCGCCAGCAGCAGATCCACAGCATCAGCGAGCGCATCCTGGA
CACCTGCGCTGGCGCGCCCGCCACCAAGCCCGTGCCCGTGCTGCTGCCCCCA1CG
AGCGCTGCACATCAACTGCAGCGAGAGCAGCGGCACCAGCGGCACCCAGT
AGAGCCAGGGCACCGCCGAGGGCGTGGGCAACCCCTAA

FIGURE 78

WO 02/04493

86/114

PCT/US01/21241

RevExon2_TV2_C_ZAw1 (SEQ ID NO:108)

ACCCCTTATCCCAAACCCGAGGGGACCCGACAGGCTCGGAGGAATCGAAGAA
GAAGGTGGAGAGCAAGACAGCAGCAGATCCATTGATTAGTGAGCGGATTCT
TGACACTTGCCCTGGGACGACCTACGAAGCCTGTGCCTCTTCTGCTACCACCGA
TTGAGAGACTTCATATTAAATTGTAGTGAGAGCAGTGGAACTTCTGGGACACA
GTAGTCTCAGGGGACTGCAGAGGGGTGGGGAACCCTTAA

FIGURE 79

WO 02/04493

87/114

PCT/US01/21241

TatExon1_TV2_C_ZAopt (SEQ ID NO:109)

ATGGAGCCCATCGACCCCAACCTGGAGCCCTGGAACCAACCCGGCAGCCAGC
CCAAGACCGCCTGCAACGGCTGCTACTGCAAGCGCTGCAGTACCACTGCCT
GGTGTGCTTCCAGAAGAAGGGCTGGGCATCTACTACGGCCGCAAGAAGCGC
CGCCAGCGCCGAGCGCCCCCCAGCAACAAGGACCACCAGGACCCCTGC
CCAAGCAG

FIGURE 80

WO 02/04493

88/114

PCT/US01/21241

TatExon1_TV2_C_ZAwt (SEQ ID NO:110)

ATGGAGCCAATAGATCCTAACCTAGAACCCCTGGAACCATCCAGGAAGTCAGC
CTAAACTGCTTGTAAATGGGTGTTACTGTAAACGTTGCAGCTATCATTGTCTA
GTTTGCTTTCAGAAAAAAGGCTTAGGCATTTACTATGGCAGGAAGAAGCGGA
GACAGCGACGAAGCGCTCCTCCAAGCAATAAAGATCATCAAGATCCTCTACC
AAAGCAG

FIGURE 81

WO 02/04493

89/114

PCT/US01/21241

TatExon2_TV2_C_ZAopt (SEQ ID NO:111)

CCCCTGAGCCAGACCCGCGCGACCCACCGGCAGGAGAGCAAGAAG
AAGGTGGAGAGCAAGACCGCGCGACCCCTCGACTAG

FIGURE 82

WO 02/04493

90/114

PCT/US01/21241

TatExon2_TV2_C_ZAw1 (SEQ ID NO:112)

CCCTTATCCCAAACCCGAGGGGACCCGACAGGCTCGGAGGAATCGAAGAAG
AAGGTGGAGAGCAAGACAGCAGCAGATCCATTCGATTAG

FIGURE 83

WO 02/04493

91/114

PCT/US01/21241

Vif_TV2_C_ZAopt (SEQ ID NO:113)

ATGGAGAACCGCTGGCAGGTGCTGATCGTGTGGCAGGTGGACCGCATGAAGA
TCCGCACCTGGCACAGCCTGGTGAAGCACCACATGTACGTGAGCCGCCGCGC
CGACGGCTGGTTCTAOCGCCACCACTACGAGAGCCGCCACCCCAAGGTGAGC
AGCGAGGTGCACATCCCCCTGGGCGACGCCCGCCTGGTGATCAAGACCTACT
GGGGCTGCAGACCGGCGAGGCGCCTGGCACCTGGGCCACGGCGTGAGCA
TCGAGTGGCGCTGCGCCGCTACAGCACCCAGGTGGACCCGACCTGAQCGA
CCAGCTGATCCACATGCACTACTTGGACTGCTTCGCCGAGAGCGCCATCCGC -
AAGGCCATCCTGGGCCAGATCGTGAGCCCCAAGTGCCTACTCCAGGCCGGCC
ACAACAAGGTGGGCAGCCTGCAGTACCTGGCCCTGACCGCCTGATCAAGCC
CAAGAAGATCAAGCCCCCTGCCAGCGTCCGCAAGCTGGTGAGGACCGC
TGGAAACAAGCCCAGAAAGACCCGCGCCGCCGCGCAACCACACCATGAAC
GGCCACTAG

FIGURE 84

WO 02/04493

92/114

PCT/US01/21241

Vif_TV2_C_ZAwt (SEQ ID NO:114)

ATGGAAAACAGATGGCAGGTGCTGATTGTGTGGCAGGTAGACAGGATGAAG
ATTAGAACATGGCACAGTTTAGTAAAGCACCATATGTATGTTTCGAGGAGAG
CTGATGGATGGTTCTACAGACATCATTATGAAAGCAGACACCCAAAAGTAAG
TTCAGAAGTACACATCCCATTAGGAGATGCCAGGTTAGTAATAAAAACATAT
TGGGGTCTGCAGACAGGAGAAAAGAGCTTGGCATTITGGGTCACGGAGTCTCCA
TAGAATGGAGATTGAGAAGATATAGCACACAAGTAGACCTGACCTGAÇAG
ACCAACTAATTGATATGCATTATTTGATTGTTTGCAGAATCTGCCATAAGG
AAAGCCATACTAGGACAGATAGTTAGCCCTAAGTGTGACTATCAAGCAGGAC
ATAACAAGGTAGGATCTCTACAATACTTGGCACTGACAGCATTGATAAAACC
AAAAAAGATAAAGCCACCTCTGCCTAGTGTTAGGAAATTAGTAGAGGATAGA
TGGAACAAGCCCCAGAGAAGCAGGGGCCGAGAGGGAACCATAACAATGAAT
GGACACTAG

FIGURE 85

WO 02/04493

93/114

PCT/US01/21241

Vpr_TV2_C_ZAopt (SEQ ID NO:115)

ATGGAGCAGCCCCCGAGGACAGGGCCCCAGCGGAGCCCTACAACGAG
TGGACCCCTGGAGCTGCTGGAGGAGCTGAAGCAGGAGGCGGTGCGCCACTTCC
CCCGCCCTGGCTGCACAACCTGGGCCAGCACATCTACGAGACCTACGGCGA
CACCTGGACCGGCGTGGAGGCCATCATCCGCATCCTGCAGCAGCTGCTGTTT
ATCCACTTCCGCATCGGCTGCCACCACAGCCGCATCGGCATCCTGCGCCAGC
GCCGCGCCCGCAACGGCGCCAACCGCAGC

FIGURE 86

WO 02/04493

94/114

PCT/US01/21241

Vpr_TV2_C_ZAwt (SEQ ID NO:116)

ATGGAACAAGCCCCAGAAGACCAGGGGCCGAGAGGGAACCATACAATGAA
TGGACACTAGAGCTTTTAGAAGAACTCAAGCAGGAAGCTGTCAGACACTTC
CTAGACCATGGCTCCATAACTTAGGACAACATATCTATGAAACCTATGGAGA
TACTTGGACAGGAGTTGAAGCAATAATAAGAATCCTGCAACAATTACTGTTT
ATTCATTTCAAGGATTGGGTGCCATCATAGCAGAATAGGCATTTTCCGACAGA
GAAGAGCAAGAAATGGAGCCAATAGATCC

FIGURE 87

WO 02/04493

95/114

PCT/US01/21241

Vpu_TV2_C_ZAopt (SEQ ID NO:117)

ATGCTGGACCTGACCGCCCGCATCGACAGCCGCTGGGCATGGCGCCCTGA
TCGTGGCCCTGATCATCGCCATCATCGTGTGGACCATCGTGTACATCGAGTAC
CGCAAGCTGGTGCGCCAGCGCAAGATCGACTGGCTGGTGAAGCGCATCCGCG
AGCGCGCCGAGGACAGCGGCAACGAGAGCGAGGGCGACACCGAGGAGCTGA
GCACCTGGTGGACATGGGCCACCTGGGCTGCTGGACGCCAACGACGTGTA
A

FIGURE 88

WO 02/04493

96/114

PCT/US01/21241

Vpu_TV2_C_ZAw1 (SEQ ID NO:118)

ATGTTAGATTAACTGCAAGAATAGATTCTAGATTAGGAATAGGAGCATTGA
TAGTAGCACTAATCATAGCAATAATAGTGIGGACCATAGTATATATAGAATA
TAGGAAATTGGTAAGGCAAAGGAAAATAGACTGGTTAGTTAAAAGGATTAG
GGAAAGAGCAGAAAGACAGTGGCAATGAGAGCGAGGGGGATACTGAAGAATT
ATCGACACTGGTGGATATGGGGCATCTTAGGCITTTGGATGCTAATGATGTGT
AA

FIGURE 89

WO 02/04493

97/114

PCT/US01/21241

gp120mod.TV1.delV2 (SEQ ID NO:119)

```

1  gaattcatgc gctgtaagg caccagaag aactgccagc agtggtagat ctggggcacc
61  ctgggcttct ggaatgat gatgtgcaac accgaggacc tggggtagc cgtgtactac
121  ggcgtgcccg tggggcgcga cggcaagacc accctgttct gggccagoga cggcaaggcc
181  tacgagaccg aggtgcacaa cgtgtgggccc acccagcctt gctgtgccac cgaaccacaac
241  cccagggaga tctgtctggg caactgtgac gagaactta acatgtgaa gaacgacatg
301  gccgaccaga tgcacgagga cgtgtacgc cgtgtggacc agagccgaa gccctgcgtg
361  aagctgaccc cctgtgtgt gaccctgaac tgcaccgaca caacgtgac cggcaaccgc
421  accgtgaccg gcaacagcac caacacacac aacggcaccg gcaatctaca catcgaggag
481  atgagaacat gcagcttcaa cgcggggccc gggccgccta tcaacttcaa caccagcaac
541  atcaccagc cctgcccaca ggtgagcttc gaccctacc ccatcacta ctggccccc
601  gccggctacg caatctgaa gtgcaacaac agacattca accgcaaccg cccctgtac
661  aacgtgagca cctgtcagtg caccacggc atcaagccc tggtagaac ctagctgtg
721  ctgaacggca gctggcgcga gggaggcacc atcaacgca gcgagaacct gaccgagaa
781  accaagacca tcatgtgca cctgaacgag agcgtggaga tcaactgca cggcccaac
841  aacaacacc gcaagagcgt ggcacgggc ccggccagg ccttctacg caaccaagac
901  gtatcggca acatcgcca gggccactgc aacatcaga ccgaccgtg gacaaagac
961  ctgcagcagg tgatgaaga gctggggcag cacttccca acaagacat ctagtcaag
1021  cccacagccc gggcgacat ggaatcacc atgcacagt tcaactgoc cggcgagt
1081  ttctactgca acaccagca cctgttcaa agcactacc acagaaaca cggacactac
1141  aagtacaacg gcaacagcag cagcccatc accctgcagt gcaagatcaa gcaatctg
1201  gcaatgtggc agggcggtgg ccaggccacc taccgcccc ccacggccg caacatcacc
1261  tggcgagca acatcaccg catctgtgt acccgagac ggggttcaa caccacaac
1321  aacaccgaga ccttcgccc cggcgccggc gacatggcg acaactggc cagcgagctg
1381  tacaagtaca aggtgtgga gatcaagccc ctgggcatg ccccaacca ggcgaaggc
1441  cgtgtgtg ncccgagaa ggccttaact gag

```

FIGURE 90

WO 02/04493

98/114

PCT/US01/21241

gp140mod.TV1.delV2 (SRQ ID NO:120)

```

1 gaattcatgc gogtcatggg caccacagaag aaotgcacag agtggtgatc ctggggcatc
61 ctgggcttct ggatgctgat gatctgcaac accgaggacc tggagggtgac cgtgtactac
121 ggcttgcccg tgtggcgaga cgcacagacc accctgttct gcgcacagga cgcacaggcc
181 tacgagacgg aggtgacaca cgtgtgggac acccaagcct gogtgccac cgaacccaac
241 cccagggaga tctgtgtggg caactgtgac gagaacttca acatgtggaa gaacgacatg
301 gcccacaga tgcacgagga cgtgatcaga ctgtgggacc agagcctgaa ggcctgcgtg
361 aagctgaccc cctgtgtcgt gaccctgacg tgcacggaca ccaactgtac tggcaccggc
421 accgtgacgg gcaacagcac caacaacacc aacggcaccg gcactacaa catcaggag
481 atgaagaact gcagcttcaa cgcggggccc ggcggcctga tcaactgcaa caccagcaac
541 atcaccagg cctgcctcaa ggtgagcttc gaccccatcc ccactcaata ctgggcccac
601 gcccgtacgg caatcctgaa gtagcaaac aagaacctcc accgacagg cccctgtcac
661 aacgtgagca cctgtgagtg caccacgggc atcaagcccg tggtaggac ccagctgctg
721 ctgaactggc cctgtgagtg ggggggac atcatccga gcgagaaact gacgagaaac
781 accaagacca tcatctgca cctgaaagg agcgtggaga tcaactgac cgcgcccaac
841 aacaaaccc gcaagagcgt ggcacatggc cccggccagg ccttctaaac caacacggac
901 gtgactggca acatccggca ggcacactgc aacatcagca ccgaccgtg gaacaagacc
961 ctgcagcagg tgatgaagaa gctggagcag cacttcccca caaagcccat ccanttcaag
1021 cccacggccg gggcgacat ggagatcacc atgcacagct tcaactgccc cggcgagttc
1081 ttctactgca acacagcaa cctgttcaac agccctacat acagcaacaa cggcactac
1141 aagtactaac gcaacagcag cagcccatc accctgcagt gcaagatcaa gcaatcgtg
1201 cgcattgtgg agggcgtggg ccaggccacc taagccccc ccactgcgg caacatcacc
1261 tgnccgagca acatcacgg cactctgtg acccgcgagc ggggttcaa caccaccaac
1321 aacacggaga ccttccgcc cggcggcggc gacatgagcg acaactggcg caggagctg
1381 tacaagtaca aggtgggtgga gatcaagccc ctgggcatcg cccccaacca ggcacagcg
1441 cgcgtgtgtg agcgcgagaa cgcgcgcgtg ggcactggcg cctgttctc gggcttctg
1501 gcccgcgcgg cgcgcacccat ggcgcgcgc agcatcacc ctgacctgca ggcgcgcag
1561 ctgtctgagc gcatcgtgca gcagcagagc aaactgtgga aggcactgga ggcacagag
1621 cacatgctgc agctgaccgt ggggggac aagcagctg aggcacggct gctggcctac
1681 gagcgtacc tgaaggacca gacgtgctg ggcactctgg cctgcacgg ccgctgac
1741 tgcacacacg cctgacccgt gaacagcagc tggagcaaca agagcagaa ggaactctg
1801 gacaacatga cctggatgca gtgggacggc gagatcagca actacacgg cctgactac
1861 aaactgtctg aggcacagca gaacagcag gagaagacg agaaggacct cctggagctg
1921 gcaagtgga acaactctg gaactgttc gacatcagca actggccctg gtacatctaa
1981 ctogag

```

FIGURE 91

WO 02/04493

99/114

PCT/US01/21241

gp140mod.TV1.mut7.delV2 (SEQ ID NO:121)

1 gaattcatgc gctgtatggg caaccagaag aactgccagc agtgggtgat ctggggcattc
 61 ctgggcttct ggatgtgtat gatctgaac accgaggacc tgtgggtgac cgtgtactac
 121 ggctgtcccg tgtgtcgcga cgcgaagacc accctgttct ggcgcagcga cgcgaaggcc
 181 taagagagcc aggtgcacaa tgtgtggggc acccagccct gctgtcccaac cgaacccaaac
 241 cccagagaga tgtgtctggg caactgtgac gagaacttca acatgtggaa gaagacatg
 301 gccgaccaga tgcacgagga cgtgtacagc ctgtggggacc agagccctgaa gccctgcgtg
 361 aagctgaccc cctgtgtcgt gaacctgaac tgcacggaca ccaactgtac cggcaaccgc
 421 accgtgaccc gcaacagcac caacaacac aacggcaacc gcaactcaaa catcgaggag
 481 atgaagaact gcagcttcaa cgcggggccc ggcggcctga tcaactgcaa caccagcacc
 541 atcaccagg ccgtgcccac ggtgtgcttc gaccccatcc ccatcacta ctgggcccc
 601 gcgggtatgc ccatctgaa gtgcacaaac aagacattca accgcaaccg cccctgtatc
 661 aagctgagca ccgtgcagtg caaccacgga atcaagcccg tgggtgagca ccaactgtctg
 721 ctgaacggca gccgtgcccga gtaggggcat atcaiccgca gcgagagacct gaccgagaa
 781 accaagacca tcatctgtga cctgaacgag agcgtgggaga tcaactgca cgcgcccaac
 841 aacaacacc gcaagagcgt ggcctcagc cccggccaagg cctctacgc caacaacgac
 901 gtgtcggca acatccgca ggccactg ccaatcagca ccgaccgtgt gaacagacc
 961 ctgcagcagg tgaagagaa gctgggcgag cacttcacca acaagaccat caagttaag
 1021 cccaccccg gcggcgacct ggaatcacc atgcacagct tcaactgccc cggcgagttc
 1081 ttctatgca acaccagcaa cctgttcaac agcaactacc acagcaacaa cggcaactac
 1141 aagtacaacg gcaacagcag cagcccccac accctgcagt gcaagataaa gcagatcgtg
 1201 cgtatgtggc agggcgtggg ccaggccacc taagccccc ccatcgcggg caacatcac
 1261 tgcgcagca acatccagg calctgtctg acccgcgagc ggcgttcaa caaccacac
 1321 aacaccaga ccttcgcgc cggccggcgc gacatggcg acaactggcg cagcgagctg
 1381 tacaagtaaa aggtgtgtga gatcaagccc ctgggcagc ccccaacaa ggccatcagc
 1441 agcgtgtgtc agagcgagaa gagcgccgtg ggcacggcg cgtgttctt ggccttcctg
 1501 ggccgtcccg gcagaccat gggcgccgc agcatcaccc tgaccgtgca ggcccgccag
 1561 ctgctgagcg gcatctgtca gcagcagagc aacctgtga aggcacatga ggccacgag
 1621 caatctgtgc agctgacct gtggggcatt aagcagctg aggcocogct gctggccatc
 1681 gagcgctacc tgaagagca gcagctgtc ggcatctgg gctgcagcg cggccgtatc
 1741 tgcaccacc cgtgtccctg gaacagcagc tggagcaaca agagcgagaa ggcactctg
 1801 gacaacatga cctggatga gttggacgc gacatcaga actacacgg cctgatctac
 1861 aacctgtcg aggacagca gaacagcag gagaagaac agaaaggact gctggagctg
 1921 gacaagtga acaacctgtg gaactgttc gacatcaga actggccctg gtacatcaa
 1981 ctgag

FIGURE 92

WO 02/04493

100/114

PCT/US01/21241

gp160mod.TV1 delV1V2 (SEQ ID NO:122)

1 gaattcatgc gctgtatggg caccagaag aactgccagc agtgggtgat ctggggcalt
 61 ctgggtctct ggtatctgat gatctgcaa acgaggagcc tgggggtgac cgtgtactac
 121 ggggtgcccc tgttgccgga cgcacaagacc accctgttct ggcgcagcga cgcacaaggcc
 181 tacgagaccc aggtgcacaa cgtgtggggc acccagcgc tgcgtgcccac cgaacccaac
 241 ccccgaggaga tctgtctggg caactgtacc gagaacttca acatgtggaa gaacgacatg
 301 gccgaccaga tgcacgagga cgtgtatcag ctgtgggacc agagccctga gcctgtcctg
 361 aagctgaccc cctgtgtcgt gggcgccggc aactgcaaca ccagcccat caaccagctc
 421 tgcocaaagg tgaacttga cccatcccc atccactact gcgccccgc cggctacgcc
 481 atctgaagt gcaacaaca gactttcaac ggcaccggcc cctgttaca cgtgagcacc
 541 gtgcagtga occacggcat caagccctg gtgagcacc agctgtgtct gaacggcagc
 601 ctggccgagg agggacatc catccgcagc gagaacctga ccgagaacac caagaccatc
 661 atcgtgcacc tgaacgagag cgttgagatc aactgcacc gcccaaca caacaccgc
 721 aagagcgtgc gcatggccc cggccaggcc ttctaccca ccaacgact gatcggaac
 781 atccgccagg ccacttcaa catcagacc gacccgtgga acaagacct gacgagggtg
 841 atgaagaagc tggcgagaca ctccccaac aagaccatcc agttcaagcc ccacgcgggc
 901 ggcgacctgg agatccat gcacagctc aactgccgc gcgagttct ctactgaac
 961 accagcaacc tgttaaacg caactccac agcaacaag gcaacttaca gtacaaggc
 1021 aacagcagca gcccatcac cctgagtc agatcaagc agatcgtgog catgtggcag
 1081 ggcgtgggcc aggcaccta cgcgccccc atcgccggca acatccctg ccgcagcacc
 1141 atcaccggca tctgtctgac ccgcgacggc ggttcaaca ccaccaaca caccgagacc
 1201 ttccgcccc ggcggcgga catcgcgac aactgggga gcgagctga caagtacaag
 1261 gtgtgggaga tcaagccct ggcctatgcc occaccaag ccaagcccg cgtgtgtcag
 1321 cgcgagaagc ggcctgtggg catcgcgcc cgtgtctgg gcttccctgg cgcgcggccg
 1381 agcaccatgg gcgcgcgag catcaccctg accgtgcagg ccgcccagct gctgagcggc
 1441 atcgtgcagc agcagagcaa cctgtgaag gcatcgagg ccagcagca catgtcgag
 1501 ctgaccgtgt ggggcaicaa gcagctgag gcccgctgc tggccatga gcgttaactg
 1561 aaggaccagc agctgtggg catctgggc tgcagcgcc gcctgactg caccacggc
 1621 gtgcccgtga acagcagctg gacgaacag agcgagagc acatctgga caactgacc
 1681 tggatgcagt gggaccgca gatcagcaa tacacggcc tgaactaca cctgtggag
 1741 gacagccaga accagcagga gaagaaagag aaggacctgc tggagctgga caagtggaac
 1801 aacctgtgga actgttoga catcagaac tggccctgt acatcaagat ctatcatg
 1861 atcgtgggog gcctgctgg cctgcgcat atcttcggg tgcgagcat cgtgaacgc
 1921 gtgcgccagg gctacagccc cctgagctt cagacctga cccccagcc ctcgggcctg
 1981 gaccgcctgg ggcgcctga ggaaggagg ggcgagcagg accgcgacc cagcatccgc
 2041 ctgtgtagcg gcttctgag cctgectag gacgacctgc gcaacctgt cctgtcagc
 2101 taccacggcc tgcgagact catctgato gcctgtcgog cctgtgagct gctggggcac
 2161 agcagccctg ggcgcctga gcgcggctgg ggaatccga agtacctgg cagcctgggt
 2221 cagtactggg cctgtgagct gaagaagag gccatcagc tgcgtgacac catcgccatc
 2281 accgtggcog agggcaocga ccgcatcat gactgtgtgc agcgcatct cgcggcatc
 2341 ctgaacatcc cccgcgcac ccgcagaggc ttgagggcg cctgtctga actcgag

FIGURE 93

WO 02/04493

101/114

PCT/US01/21241

gp160mod.TV1.delV2 (SEQ ID NO:123)

1 gaattcatgc gcgtgatggg caaccagaag aactgccagc agtgggtggat ctggggcattc
 61 ctgggtttct ggaigtgat gatgtcaac accgaggacc tgggtggagc cgtgtactac
 121 ggctgtcccg tgggtgcgga cgcacagacc accctgttct ggcacagca cgcacaggcc
 181 taaggagacc aggtgcacaa cgtgtgggac acccagcct ggtgtccac cgaacccaac
 241 cccagagaga tctgtctggg caactgtacc gagaacttca acatgtgaa gaacgacatg
 301 gccgaccaga tgcacggaga cgtgtatcag ctgtgggacc agagcctgaa ggcctgtgtg
 361 aagctgacc cctgtgtgt gacotgaa tgatccgaca ccaactgac cggcaaccgc
 421 accgtgacc gcaacagca caacaacac aacggcacc gcatctaca catcgaggag
 481 atgaagaact gcacttcaa cgcggcgcc ggccgcctga tcaactgaa caccagacc
 541 atcaccagg cctgtccaa ggtgtgtgt gacccatcc ccatcacta ctggccccc
 601 gccgtctac ccatctgaa gtgcacaa aagaacttca acggcaccgg cccctgttcc
 661 aactgtgaca cgtgtgact caccacagg atcaagccc tgggtgagca ccagctgtg
 721 ctgaacgga cctgtgtgt ggaaggatc atcatcgca ggcagacct gacggagac
 781 accaagaca tcatctgca cctgaacag agcgtggaga tcaactgac cgcgccaac
 841 aacaacacc gcaagagcgt ggcacaggc cccggccagg ccttctacc caccacagc
 901 ggtatcgga acatcgcca ggcacatgc aacatcaga ccgacgcgt gaacagacc
 961 ctgtgagcag tgaatgaaa gctgtggag cacttccca acaagacct ccagttaac
 1021 cccacgccc ggcgggacct gggatcacc atgcacagt tcaactgac cggcgagttc
 1081 ttctactga acacngcaa cctgttcaa agcaacttcc acagcaaca cggcaccac
 1141 aagtacaac gcaacagcag cagcccatc accctgagt gcaagataa gcagatgtg
 1201 cgtatgtgg aggtgtggg ccaggccac tacgcccac ccatgccc gcaacatcac
 1261 tgcgcagca acatcccg catctgtct acccgagag ggcgttcaa caccacaa
 1321 aacacgaga ccttccccc cggcgccgc gacatgcgc acaactggc cagcgagctg
 1381 taacagtaca aggtgtgtga gatcaagcc ctgtggcatg cccccacca ggcacagcc
 1441 cgcgtgtgc agcgcgaga ggcgcctgt ggcacggcg cgtgttctt ggtttctt
 1501 ggcgcgcgc gacgtacct ggcgcgcgc agcatcac tgaactgca ggcgcgcag
 1561 ctgtgagcg gcatctgca gcagcagcg aactgtga aggcacatg ggcgcgcag
 1621 cactgtgtg agctgacct gttgggcatc aagcagctg aggcgcgtg gctggccatc
 1681 gagcgtacc tgaagacca gcactgtct ggcacttgc gctgcagcg cgcctgac
 1741 tgcaccacc cgtgtccct gacacagc tggagcaaa agagcgaga ggcacttgg
 1801 gacaacatg cctggatga gttggaccg gacatcaga actacaccg ctgactac
 1861 aactgtgtg aggacagca gaacacagc gagaagaac agagggacct gctggagctg
 1921 gacaagtga acaactgtg gaotgttt gacatagca actgccccg ctacatcag
 1981 atcttcatc tgatgtggg cggcctgac ggcctgcca tcatctgc cgtgtgagc
 2041 atctgaac gctgtgcca ggcctacag cccctgagt tccagacct gacccagc
 2101 cccgcggcc tggaccgct ggcagagcat gaggaggag ggcgcgaga ggcgcgac
 2161 cgcagatcc gctgtgtgag cgtgttctg agcctgctt ggcagacct ggcacactg
 2221 tgcctgtca gctacaccg cctgcgcac tcatctga tgcctgtgc ggcgtggag
 2281 ctgtgtggc acagcagct ggcgcctg cagcgcgtt gggagatct gaagtacct
 2341 ggcagcttg tgcagtact ggcctgtgag ctgaagaaga ggcacatcag cctgtgtgac
 2401 accatgcca taccgtggc caggggcaco gacgcacat tgaagctgt gacgcgac
 2461 tgcgcgcca tctgaacat ccccgccgc atcgccagg gcttggagg cgcctgtgt
 2521 taactgag

FIGURE 94

WO 02/04493

102/114

PCT/US01/21241

gp160mod.TV1.mut7.delV2 (SEQ ID NO:124)

1 gaattatgc gctgtaggg caccagaag aactccagc agtggtagt ctggggcatt
 61 ctgggtctt ggtgctgat gatctgaac accgaggacc tgtgggtgac cgtgtatc
 121 ggctggccc tgtggcgga cgcacagacc accctgtctt ggcacagca cgcacagccc
 181 tacgagaccg aggtgcaca cgtgtgggccc acccagcctt gctgtccca cgcacccaa
 241 cccagggaga tctgtctgg caactgtacc gagaattca acatgggaa gaacgacatg
 301 gccgaccaga tgcacggaga ctgtatcagc ctgtgggacc agagcctgaa gctgtcgtg
 361 aagctgacc cctgtgctg gacctgaac tgcacgaca ccaactgac cggcaaccgc
 421 accgtgacc gcaacagca caacaacac accgacacc gcatctaca catcgaggag
 481 atgaagaact gacgttcaa cgcggcgccc ggccgcgiga tcaactgca caccgacc
 541 atcaaccagg cctgcccac ggtagcttc gacccatcc ccatccacta ctggccccc
 601 gccggttacc ccatctgaa gtcacacac aagacctta accgacacc cctgtctac
 661 aacgtgaga cctgtcagt caccacagg atcaagccc tggtagaac cagctgtctg
 721 ctgaacggca cctgtggcga ggagggcatt atcatcgca gcgagaact gacggagaa
 781 accaagacca tcatctgca cctgaacag agcgtggaga tcaactgac cgcgccac
 841 aacaacacc gcaagagctt ggcctatggc cccggccagg ccttaccg caccacagc
 901 gtgacggca acatcgcca ggccactgc aacatcgca cgcacgctg gaacaagacc
 961 ctgacgagg tgaagaaa gctggcgag cacttccca acaagacct cagtttaag
 1021 cccacgccc gctgggacct ggagatcac atgcacagt tcaactgag cggcgagttc
 1081 ttctatgca acaccagaa cctgttcaa agcactacc acagacaaa cggcactac
 1141 aagtacacg gcaacagca cagcccatc accctgagt gcaatgaa cagatctg
 1201 cgtatgtgg agggcgtgg ccaggccacc taccgccc ccatgccc ccatcacc
 1261 tgcgcagca acatcacc cactctgtg acccgagag cggcctcaa caaccacac
 1321 aacacgaga ccttccccc cggcggggg gacatggcg acaactggc cagcgagctg
 1381 tacaagtaca agtgggtga gatcaagccc ctgggcatg ccccaacca ggccatcag
 1441 agcgtgggc agagcgaga gagcgccgtt ggcacggcg cgtgttctt ggtttctg
 1501 ggcgccgag gacgacacc gggcgccgc agcatcac tgaactgca ggcgcgag
 1561 ctgtgtgag gcatctgca gacgagagc aactgtga aggcacaga ggccagcag
 1621 cactgtctg agctgacct ggggggcat aagcagctg aggcgcgct gctggccatc
 1681 gacgctacc tgaagacca cagctgtct ggcacatgg gctcagcgg cgcgctgac
 1741 tgcacaccg cgtgtccct gaacagcag tggagcaaa agagcgaga ggcacatgg
 1801 gacaacatg cctggatga gtcggaccg gacatgca actacacgg cctgatctac
 1861 aactgtctg aggacagca gaaccagag gagaagaag agaaggact gctggagctg
 1921 gacaagtga acaacctgt gaactgttc gacatgca actggccctg gtacataag
 1981 atcttcatc tgaatggg cggcctgac ggcctgca tcatctgc cgtgtgagc
 2041 atctgaacc gctgtgcca ggcctacag cccctgagt tccagacct gaaccacag
 2101 ccccgccgc tggacgctt gggcgccatc gaggagagg gcggcgagc gaaccgagc
 2161 cgcagcacc gctgtgtgag cggctctct agcctgctt gggagacct ggcacactg
 2221 tgcctgtca gctaccag cctgcgcag tcatctga tgcctgtg cgcgtggag
 2281 ctgtgtgac acagcagct ggcggcctg cagcgcgct gggagacct gaactacctg
 2341 ggcagcctg tgcagtact ggcctggag ctgaagaaga ggcacalag cctgtggac
 2401 accatgcca taccgtgg cagggcacc gaccgatca tgaactgtt gacgacatc
 2461 tgcgcgcca tctgaacat ccccgccgc atccgacag gcttggagg cgcctgtctg
 2521 taactagag

FIGURE 95

WO 02/04493

103/114

PCT/US01/21241

gp160mod.TV1.tpa1 (SEQ ID NO:125)

1 gtcgaagcca ccatgatgc aalgaagaga gggctctgct gttgtctgct gctgtgtgga
 61 gcagttcttg ttcccccag cgcagcacc gaggacctgt ggtgtgacct gtaatacgc
 121 gtgcccgtgt ggcgcgacgc caagaccacc ctgtctctgc ccagcgacgc caaggcctac
 181 gagaccgagg tgcacaactg gtgggcacc cagccttgcc tgcacacga ccccaacccc
 241 caggagatgc tgcgtggcaa cgtgaccgag aacticaa tgttgaagaa cgcataggcc
 301 gaccagatgc acgaggacgt gatcagcctg tgggaccaga gcttgaagcc ctgcgtgaag
 361 ctgacccccc tgtgctgtac cctgaactgc accgacaca acgtgacccc caaccgaccc
 421 gtgaccgcca acagaccaa caacaccaac ggcaccgcca tctacaacat cgaagagatg
 481 aagaactgca gttcaacgc caccacgag ctgcgcgaca agaagcaca ggaaglacgc
 541 ctgtttacc gcttgacat cgtgcctctg aacgagaaca ggcacaact cactacccc
 601 ctgataact gcaacaccg caactatccc caggcctgcc ccaaggtgag ttccgacccc
 661 atcccatccc actactgccc cccgcgcggc tagccatccc tgaagtgcac caacagacc
 721 ttaaacggca ccggcccttg ctacaactg agcaccgagg agtgcacca cggcatcaag
 781 cccgtgttga gacccagct gctgctgaac ggcagcctgg ccgaggagg catatcatc
 841 cgcagcaga acctgacga gaacccaag accatcatg tgcacctgaa cgagagcgtg
 901 gagatcaact gcaccggccc caacaacaa acccgcaaga gctgtcgcac cggcccccgc
 961 caggccttcc atgcaccaa cgaactgac ggcacatcc gccagcccca ctgcaacatc
 1021 agcaccgacc gctggaaca gacctgctg caggtgatga agaaactggg cgaagacttc
 1081 cccaacaga ccatcagtt caagccccc gccggcggcg acctggagat caccatgcac
 1141 agcttcaact gccggcgcca gttctctac tgaacacca gcaacctgtt caacagacc
 1201 taacacgca acaacggcac ctacaagtaac aacggcaaca gcagcagccc catnacctg
 1261 cagtgcaga tcaagcagat cgtgcctatg tggcaggggc tggccaggc caactacgccc
 1321 ccccccctgc ccggcaacat caactgccc agcaacatca ccggcatctc gctgacccc
 1381 gacggcggtc tcaacacac caacaacacc gagacttccc gcccccggcg ccggagacatg
 1441 cgcgacaact ggcgcagcga gctgtacaag taacaggtgg tggagatcaa gcccttgggc
 1501 atcgccccc ccaaggccaa ggcgcgctg gtcagcgcgc agaagcgcgc cgttgggcac
 1561 ggcgcgctgt tcttgggctt cctgggcgcc gccgcagca ccattggggc cgcagcacc
 1621 accctgacc tgcaggccc ccagctctg agcggcatcg tgcagcagca gacgaacctg
 1681 ctgaaggcca tgcaggcca gcagcagatg ctgcagctga ccgtgtggcg catcaagcag
 1741 ctgagggccc gctgtctggc catcgagcgc taactgaagg accagcagct gctgggcac
 1801 tggggctgca gctggcgcct gatctgcacc accgcctg cctgggacag cagctgggac
 1861 aacaagagcg agaaagacat ctgggacaac atgacctgga tgcagtggga ccggagatc
 1921 agcaactaca ccggcctgat ctacaactg ctggaggaca gccagaacca gcaggagaag
 1981 aacgagaagg acctgctgga gtcggacaag tggacaacc gtgtgaactg gttgacatc
 2041 agcnaactgc cctggtaact caagatctc atcatgatg tgggcggcct gatcgccctg
 2101 cgcatactc tgcggctgt ggcacatgta aaccgcgtgc gccagggcta cagccccc
 2161 agcttccaga cctgaacccc cagcccccgc ggcctggaac gcttggggcg catcaggag
 2221 gagggcgccg agcaggaacc gcaccgcaga atccgctgg tgcagcgtt cctgagcctg
 2281 gcttgggagc acctgcgcaa cctgtgcttg ttacgtacc accgcctg cgcacttcatc
 2341 ctgacgcgc tgcgcgcctg gtagctgctg ggcacagca gctgcgcgg cctgcagcgc
 2401 ggcctggaga tctgaagta cctgggcagc ctgtgtcagt actggggcct ggaactgaag
 2461 aagagcgcca taagctgctt ggaacatcc gcatcaccg tggccgggg caccgacccc
 2521 atcatcgagc tgggtcagcg catctgcgc gccatcttga acatccccc cgcataccgc
 2581 caggacttgc aggcggccct gctgtactc gag

FIGURE 96

WO 02/04493

104/114

PCT/US01/21241

1 gaattatgc gctgatggg caccagaag aactgcagc agtggggat ctggggcalt
 61 ctgggtctct ggtatgtgat gatcgaac accgaggacc tgggggtgac cgtgtactac
 121 ggcgtgccc gttgggcga gcccaagacc accctgtcti ggcgaagcga cggcaaggcc
 181 tacgagaccg aggtgcacaa cgtgtgggc acccaagcct ggtgcccac cgaaccacac
 241 cccagggaga tctgtctggg caactgtgac gagaactca acutgtggaa gaacgacatg
 301 ggcgaccaga tgcacgagga cgtgatcagc ctgtgggacc agagcctgaa gccctgctgt
 361 aagctgacc cctgtgctgt gacctgaac tgcaccgaca ccaactgaa cggcaacgac
 421 accgtgaccg gcaacagcac caacaacac aacggcacgc gaactacaa catctaggag
 481 atgaagaact gcagcttcaa cgcaccacc gagctgcgcg acaagaagca caaggaglac
 541 gccctgtct accgcttga calctgtccc ctgaacgaga acagcgacaa cttaacctac
 601 gccctgata actgaacac cagcaccac acccaggcct gccccaaggt gagcttggac
 661 cccatccca tcaactatg cgcctccgac ggtacgaca tctgaagtg caacaacag
 721 accctcaac gcaacggccc ctgtacaa gtagcaccg tgcagtgcac ccacggcalt
 781 aagcccggtg tgaacacca cgtgtgtctg aacggcagcc tggcctgaga gggcaltac
 841 atccgacgg agaactgat cgaacacac aagaccata tctgacact gaacggagac
 901 gtaggagata actgacccc cccaacac aacaccgca agagcgtgac catctggccc
 961 ggcagggcct tctatgcac caacgtgtgt atgggcaaa tccgctagc ccactgcaac
 1021 atcagaccg accgtgtgaa caagacctgt cagcaggtga tgaagagct gggcgagac
 1081 tccccaaca agacatcaa gttaagccc cagcctggcg ggcacttga gatcaccatg
 1141 cacagctta actgctggcg cgaattctt tactcaaa ccagcaact gttaacago
 1201 accatcaca gcaacacgg caactacag tactacggca acagcagcag cccatcac
 1261 ctgcagtga agatcaagg gatctgtgac atgtggcagg gctgtggcca ggcacctac
 1321 gcccccaca tgcctggcaa cactacctgc cgcagcaaa taccggcat ctgtgtgac
 1381 cgcgacggcg gcttaacac caacaacac accgagactt tccgcccgc cggcgcgac
 1441 atcgcgaca actgctggcg cgaactgtac aagtacaag tgggtgagat caagccctgt
 1501 ggcaltgccc ccaacaagg caagcctgc gttgtgacg gcgagaagcg cgcctgtggc
 1561 atcgcgccg tgttctggg ctctctggc ggcctggca gcaactggg cgcgccaac
 1621 atcaacctga cgtgtgagg cgcctagctg ctgagcgga tctgtgaca cgaagacac
 1681 ctgtgaagg ccatcgagg ccagcagcac atgtctgag tgaactgtg ggcaltcaag
 1741 cagctgcagg cccgctgtct ggcactgag cgtactcga agcaccaga cgtgtgggc
 1801 atctgggct gacggcgcg cctgactgc accaccgcg tgcctggaa cagcagctg
 1861 agcaacaaga gcgagaagg catctggac aactgacct ggaatgagtg ggcctggag
 1921 atcagcaact acaccggct gatcacaac ctgtgtgagg acagcgaga ccagcaggag
 1981 aagaacgaga aggaactgt ggaactgac aagtggaaca accgtggaa ctgtgtgac
 2041 atcagcaact ggcctgtga catcaagac ttaactgta tctgtggcg cctgtgtgac
 2101 ctgcgcatca tcttgcctg gctgagcat gtgaaccg tgcgcccagg ctacagccc
 2161 ctgagctcc agacctgac cccagcccc cgcgctctg accgcttgg cgcctggag
 2221 gaggagggcg gcgagcagg cgcgacgcg apcaltgccc tggtagcgg ctctctgag
 2281 ctggcctggg acgactgag caactgtgc ctgttaact accaccgct ggcgacttc
 2341 atctgatcg cgtgtgctg cgtgtgctg ctgtggcaca gcagcctgc cgcctgtgag
 2401 cgcgctggg agatctgaa gtacctggg agcctgtgc agtactggg cctgtgagctg
 2461 aagaagagcg ccatcagct gctgacac atcgcata cgtgtggcg ggcacccgac
 2521 cgcaltatg agctgtgta gcgaltctg cgcgcatcc tgaacatcc ccgctgac
 2581 cgcagggct tctggggcg cgtgtgaa ctgag

FIGURE 97

WO 02/04493

105/114

PCT/US01/21241

1 gaattcatga gaggatggg gacacagaag aattgtcaac aatgggtggat atggggcatc
 61 ttaggottct ggaigttaat gattgttaac accgaggacc tgggggtgac cgtgtactac
 121 ggctgtcccg tgggtgcgga cggcaagacc accctgttct ggccacgcca gccaaggccc
 181 taaggagacc aggtggcaca cgtgtggccc accacggcot gctgtgcccac cgaacccaac
 241 cccagagaga tctgtctggg caacgtgacc gagaactica acaigtggaa gaacgacatg
 301 ggcgaccaga tgcacgagga cgtgtacaga cgtgtggacc agagcctgaa gccctgcctg
 361 aagctgaccc cctgtgtgtt gaccctgaac tgcacgaca ccaacgtgaa cggcaaccgc
 421 acgtgacccg gcaacgacac caacacacc aacggcaccg gcattotaca catcgaggag
 481 atgaagaact gcagcttcaa cggcaccacc gactgtgcgg acagaagga caaggagtac
 541 gccctgttct accgcttggg calogtggcc ctgaacgaga acagcgacaa ctacactac
 601 cgtgtgatca actgcaacac cagcaccate acccaggcct gcccaagggt gactttgac
 661 cccatcccca tccactactg cggcccccgc ggtacgcca tctggaagt caacaacaag
 721 accttaacg gcaccggccc ctgtctaac gttagcaccg tgcagtgcac ccaaggcatc
 781 aagccctgga tggacacca gctgtgtgtg aacggcagcc tggccgagga gggtacatc
 841 atccgcagcg agaaccctgac cgggaacacc aagaccatca tctgtacact gaacgagagc
 901 gttgagatca actgacccc cccaacaac aacacggcca agaggtgtgc catcgccccc
 961 ggccaggcct tctacggac caacgacgtg atcgccaaca tccggcaggg ccaatgcaac
 1021 atcagacacg accgtgtgaa caagacgtg cagcaggtga tgaagagctt ggccgagcag
 1081 tcccacaaca agacacatca gttaacggcc cagccggcgg ggacactgga gataccatg
 1141 cacagcttca actgcccggg cgaattcttc tactgaaca ccagcaacct gttaacagc
 1201 accatccaca gcaacaacgg cactacaaag tacaaggcca acagcagcag cccatcacc
 1261 ctgcagtcca agatcaagca gatctggcgc atgtggcagg gctgtggcca ggcacactac
 1321 gcccccacca tgcggcgcaa catcacctgc cgcagcaaca tcaacggcat cctgtgtacc
 1381 cggcagcgcg gcttcaacac caacaacaac accgagacct tcccccggg cggcgggcag
 1441 atgcgcgaca actggcgacg cgaagctgac aagtaacagg tgggtggagat caagcccttg
 1501 ggcatcgccc caacaaggc caagcgccgc gttgtgtagc gggagaagcg cgccttggg
 1561 atggcgcccg tgttctggg ctctctgggc ggcggcgcca gcacatggg cgcggccaac
 1621 atcaacctga cgtgtcagcg ccggcagctg ctgagcgcca tctgtcagca gcaagcaaac
 1681 ctgtctgaag ccatcgagcg ccagcagcac atgtctcagc tgaactgtgt gggcatcaag
 1741 cagctgcagg ccccgctgct ggccatcgag cgtacactga aggaacagca gctgtcggc
 1801 atctgggctt gcagcgcccg cctgatctgc accacggcg tgcctggaa cagcagctgg
 1861 agcaacaaga gcgagaagga catctgggac aacatgaact ggaatcgatg ggaacggcag
 1921 atcagcaact acacggcctt gatctacaac ctgtgtggag acagccagaa ccagcaggag
 1981 aagaacgaga aggaactgtt ggaactggac aagtggaaca acctgtgaa ctgtgtgac
 2041 atcagcaact ggccctgga catcaagatc tcatcatga tctgtggcgg cctgacggc
 2101 ctgcacatca tcttgcctgt gctgagcatc gtaaacggc tggccagggg ctacagccc
 2161 ctgagcttcc agacctgac cccagcccc cgggctggg accgctggg cggcatcgag
 2221 gaggaggggg gcgagcagga ccggcagccc agcatcgcc tgggtgagcg ctctcgagc
 2281 ctggcctggg acgactggc caacctgtgc ctgtcagct accacggct ggcgacttc
 2341 atctgatcg cgtgtcgtgc cgttgagctg ctggccaca gcagcctgac cggcctgag
 2401 cgggctggg agatctgaa gtactgggc agcctgtgac agtactggg cctggagctg
 2461 aagaagagcg ccacagcct gctggacacc atcgccatca cgttgccga ggccacgac
 2521 cgcacatcag actgtgtgca ggcacatgc cgcgcacac tgaacatcc cgcgcacac
 2581 cggcaggcct tggaggccc cctgtgttaa ctggag

FIGURE 98

WO 02/04493

106/114

PCT/US01/21241

I atgagagtgta tggggacaca gaagaattgt caacaatggt ggatattggg catcttaggc
61 ttctggatgc taatgatttg taacacggag gacttggtgg tcacagtcta ctatgggta
121 cctgtgtgga gagacgcaaa aactactcta ttctgtgcat cagatgctaa agcatatgag
181 acagaagtgcc ataatgtctg ggctacaat gctgtgtac ccacagacco caaccocaaa
241 gaaatagttt tgggaatgt aacagaaat ttaatatgt ggaanaaiga catggcagat
301 cagatgcatg aggatgtaat cagtitaagg gatcaaaagcc taagccatg tetaaagtg
361 accccactct gtgtacttt aaactgtaca gatacaaatg ttacaggtaa tagaactgtt
421 acaggtaata gtaccaataa tacaatggt acaggtattt alaacttga agaatgaaa
481 aattgctctt tcaatgcac cacaanaa agagataaga aacataaaga gtatgcactc
541 tttatagac ttgalatagt accactaat gagaatagtg acaacttac atatagatfa
601 ataatatga atacttcaac cataacaaa gctgtgcaa aggtctcttt tgaccgatt
661 cctatacatl actgtgtccc agctgttat gctgtctaa agtgtatlaa taagacatc
721 aatgggacag gacctatgta taatgtcagc acagtacaat gtacacatgg aattaagcca
781 gtgtatcaa ctcaattact gtaaatggt agctagcag aagaagggat aataataga
841 tctgaaaatt tgacagagaa taccacaaa ataatagtac acctaatga atctgtagag
901 ataatatga caagcccaa caataatac agaaaagtg taaggatagg aaccggacaa
961 gcatctatg caacaaatga tetaatgga aacataagac aagccatg taacatng
1021 acagatagat ggaacaaaac ttacaacag gtaatganaa aattagagga gcatctcc
1081 aataaacaac tacaatttaa accacatga gaggggatc tagaattac aatgcatagc
1141 ttaattgta gaggagaatt ttctatgt aatacatcaa acctgttaa tagcacatc
1201 cactataata atggatata caataacat gtaattcaa gctacccat cacactcaa
1261 tetaaataa aacaaattgt acgatgtgg caagggtgag gacaagcaac gtatgccc
1321 ccaattgacg gaacataac atgtatga caacacacag gaatactatt gacatgtgat
1381 ggaggatita acaccacaa caacacagag acatcacag ctggaggagg agatagagg
1441 gatactgga gaagtgaatt atataaatt aaagtatga aattaagcc atgggaata
1501 gacccacta aggcanaag aagatgggtg cagagagaaa aaagagcagt gggaatagga
1561 gctgtgtccc ttgggtctt gggagcagca ggaagcacta tggggcagc gtcaataacg
1621 ctgaaggtag aggcagaca actgtgtctt ggtatagtc aacagcaaa caattgtctg
1681 aaggtatag aggcacaaa gcatatgt caactacag tctggggcat taagcagctc
1741 caggcgagag tctgtgctat agaaagatac ctaaaggatc aacagctctt agggattgg
1801 ggtgtctctg gaagactcat ctgacccat gctgtgctt ggaactocag ttggagtaat
1861 aatctgaaa aagatatttg ggaatacatg acttgatgc agtgggatag agaaattagt
1921 aattacacag gcttaalata caattgtct gaagactgc aaaaaccagca ggaanaagat
1981 gaaaagatt tattagaatt ggacaagtgg aacatctgt ggaattgatt tgcataata
2041 aactggcctg ggtatataa atattatca atgatagtg gaggctgat aggtttaaga
2101 ataatttg ctgtgttic latagtgaat agagttaggc agggatctc accttgca
2161 ttcaagccc ttacccaag ccggaggga ctgcagagc tggaggat cgaagaanga
2221 ggtggagagc aagacagaga cagatcata cgattggiga ggggallct gtgcctggc
2281 tgggagato tgggaacct gtgctcttc agctaccac gcttgagaga ctacataa
2341 attgcagta gggcagtgga actctggga cagacagtc tcaaggagct acagaggggg
2401 tgggaatcc taagtatct ggaagtctt gtcactatt ggggtctaga gtaaaaaag
2461 agtgcattta gtgtgttga taccataga ataacagtag ctgaaggac agataggatt
2521 atagaattag tacaagaal ttgtatagct atctcaaca taccagaag aataagacag
2581 ggtttgaag cagcttgct ataa

FIGURE 99

WO 02/04493

107/114

PCT/US01/21241

gp140mod.TV1.tpa1 (SEQ ID NO:131)

```

1 atggtatgcaa tgaagagagg gctctgtgt gtgtgtgtgt tttgttgagg agttcttggt
61 tggccacagg ccagacacga ggaactgttg gtgaccgtgt actacggcgt gccctgtgtg
121 cgcgacgcca agacacacct gtctgtggcc agcgacgcca aggcctacga gaccgaggtg
181 cacaacgtgt gggccaccca egcctgtgtg cccacggacc ccaaccccca gggatgtgtg
241 ctgggcaacg tgaccgagaa cttaacatg tggaaagaag acatggccga ccagatgcac
301 gaggacgtga tcagctgttg ggaacagagg ctgaagccct gctggaagct gaaccccttg
361 tgcgtgaccc tgaactgcac cgacacacac gtgaccggca accgacccgt gaacggcaac
421 agcaccaaca acaaccaagg caccggcacc tacaacatg aggaatgtaa gaactgcagg
481 ttaacggcca ccaccgagct ggcggacaag aagacaaggg agtaagccct gtctaacgca
541 ctggacatcg tgcctctgaa cgagaacagc gacaaactca cctacggcct gatcaactgc
601 aacaccagca ccaacaccca ggcctgcccc aaggtgagct tggaccccat ccccatccac
661 tacttgcccc ccggcgctca cgcactcttg aagtgcaaca acaagaccti caacggcaac
721 ggccctgtct acaacgtgag ccactgtcag tgcacccatg gcatcaagcc cgttgtgagg
781 acccagctgc tgcgtaacgg cagcctggcc gaggagggga tcatcatcgg cagcgagaaac
841 ctgaccgaga acaccaagac catcatcttg cactgaagc agagcgtgga gatcaactgc
901 acccggccca acaacaacac ccgcaagagc gtgcgcctcg gcccgggcca ggcctctac
961 gccaccaacg acgtgaltcg caacatccgc caggcccaat gcaacatcag caacgaccgc
1021 tggacaaga cctgtaagca ggtgatgaag aagctggcgc agcaatcccc caacaagacc
1081 atccattica agcccacgc cggcggggac ctggagatca ccatgcacag cttaactgc
1141 cgcggcgagt tctctatcg caacaccagc aacctgtica atagaaccti ccacagcaac
1201 aacggcacct acaagtacaa cggcaacagc agcagcccca tcacctgca gtgcaagatc
1261 aagcagatcg tgcgcatgtg gcaggcggtg ggcacggcca cctacggccc ccccatggcc
1321 ggcacatca cctgcgcgag caacatcacc ggcactctcg tgaccccgga cggcggttc
1381 aacaccacca acaacaccca gaccttcgc ccggcgggcg gcgacatcg cgacaactgg
1441 cgcagcgagc tglacaatga caagtggtg gagatcaagc cctggggcat cggcccccac
1501 aagcccaagc gcccggtggt gcagcgcgag aagcgcgccg tgggcacgg ccgcgtgttc
1561 ctgggcttcc tggcgccgcg cggcagcacc atggcgcccg ccagatcac cctgacgttg
1621 caggcccgcc agctgtgtg cggcatctgt cagcagcaga gcaacctgt gaaggccatc
1681 gaggcccgcc agcactgtct gcagctgacc gtgtggggca tcaagcagct gcaggccctgc
1741 gtgttgcca toagcgctca cctgaaggac cagcagctgc tgggcacttg ggcctgcagg
1801 ggcgcctgta tctgaacac cgcctgtccc tggacagca cctggagcaa caagagcgag
1861 aagacatct gggacaacat gacctggatg cagtgggacc gcgagatcag caactacac
1921 ggcctgatct acaactgtct ggaagacagc cagaaccagc aggaagaaga cgagaaggac
1981 ctgttgagc tggacaagtg gaacacccgt tggaaatggt toagatcag caactggccc
2041 tggtaact aa

```

FIGURE 100

WO 02/04493

108/114

PCT/US01/21241

gp140mod.TV1 (SEQ ID NO:132)

1 gaattctatgc gcgtgatggg caccagaag aactgcagc agtgggtgat ctggggcato
 61 ctgggcttct ggaigtgat gatgtgaac aocgaggacc tgggggtgac cgtgtactac
 121 gggtgtgocg tggggcgga cgcgaagacc accctgttct gcgcagcga cgcraaggcc
 181 taccagaccg aggtgcacaa cgtgtgggccc acccagcct gcgtgcacac cgaaccaaac
 241 cccagagaga tctgtctggg caactgtacc gagaactca acatgtggaa gaacgacatg
 301 gccgaacaga tgcagagga cgtgtatcagc ctgtgggacc agatgttgaa gccctgcctg
 361 aagctgaccc cctgtgtcgt gaacctgaac tgcacogaca ccaactgtac cagcaaccgc
 421 accgtgaccc gcaacagcac caacaacacc aacggcaccc gcactacaac catcgagtag
 481 atgaagaact gcagcttcaa cgcaccacc gagctggcgc acaagaagca caagggtatc
 541 gccctgttct accgcctgga catgtgtccc ctgaacgaga acagcgacaa ctacacctac
 601 cgtctgtatc actgcaacac cagcacctac acccaggtct gccccaaggt gacgtctgac
 661 cccatccca tccactatg cgcctccgcc ggtctagcca tctggaagtg caacaacaag
 721 accitcaacg gcaccggccc ctgtctaac gtaggcaccc tgcagtgcac ccacggcalt
 781 aagccctgtg tgaagccca cgtgtgtgtg aacggcagcc tggctcgagg gggcaltalc
 841 atcccgacgg agaactcgac cgaagaacac aagaccatca tctgtcactt gaacgagagc
 901 gtggagatca actgcacccg ccccaacaac aacaccgcca agagcgtgct catcggtccc
 961 gccaggtctt tctagccac caacgacgtg atcggcaaca tccgccaagg ccactgcaac
 1021 atcagcaccc accgtctgaa caagacctg cagcagtgta tgaagaagct gggcgagcac
 1081 ttccccaaca agaccatca gttcaagccc cagcctggcg gcgacctgga gatcaccatg
 1141 cacagcttca actgcggcgg cgaagtcttc tactgcaaca ccagcaactt gttcaacagt
 1201 acctaccaca gcaacaacgg caactacaag tacaacggca acagcagcag ccccatcacc
 1261 ctgcagtgtca agatcaagca gatctgtggc atgtgtcagg gcgtggacca ggcacactac
 1321 gcccccccca tgcctggcaa catcaccgtc cgcagcaaca tcaocggcat cctgtctgac
 1381 cgcgacggcg gcttcaaac caccacaac accgagacct tccgccccgg cggcgggagc
 1441 atgcgcgaca actggcgtag cgaactgtac aagtacaagg tggtagagat caagccctg
 1501 ggcacgtccc ccaacaagg caagcgccgc gtgtgtcagc gcgagaagcg cgcctgtggc
 1561 atcggcgccg tttctctggg ctctctggcg gccgcggcca gcaactatgg cgcctggcgc
 1621 atcaacctga cgtgtcagc cgcgcagctg ctgagcggca tctgtcagca ccaagacaa
 1681 ctgtgaaggc ccatcgaggc ccagcagcac atgtctcagc tgaactgtg gggcatcag
 1741 cagctgcagg cccgctgtct ggcacatgag cgtctacctg aggaaccaga gctgtctggc
 1801 atctggggct gcagcggcgc cctgactgac accaccgccc tgcctggaa cagcagctgg
 1861 agcaacaaga gcgagaagga catctgggac aacatgacct ggaatcagtg ggcacgggag
 1921 atcagcaact acacggcctt gacttacaac ctgtctggag acagccagaa ccagcaggag
 1981 aagaacgaga aggacctgct gagcgtggac aagtggaa accgtgtgaa ctgtctgac
 2041 atcagcaact ggcctgtgta catctaatc gag

FIGURE 101

WO 02/04493

109/114

PCT/US01/21241

gp140mod.TV1.wtLnative (SEQ ID NO:133)

1 gaaticatga gagtgatggg gacacagaag aattgtcaac aatgggtggaat atgggggcato
 61 ttaggctttt ggaftgtaat gatttgaac accgaggacc tgtgggtgac cgtgtactac
 121 ggcgtgcccg ttggcgccga cggcaagacc accctgtcti ggcgcagcga cggcaaggcc
 181 taogagaccg aggtgacaaa cgtgtggccc acccaagcct gctgtccac cgaacccaac
 241 cccagagaga tctgtctggg caacgtgacc gagaacttca acatgtggaa gaaagacalg
 301 gctgacwaga tgcacgagga cgtgtatcago ctgtgggacc agagcctgaa gcccgtgtg
 361 aagctgacc cccgtgtgtt gaccttgacc tgcaccgaca ccaacgtgac cggcaaccgc
 421 accgtgaccg gcaacagcac caacaacacc aacggcaacc gcatctaca catogagtgag
 481 atgaagaact gcagcttcaa cggcaaccacc gagctgcgcg acaagaagca caaggagtac
 541 gcccgtgtct accgcttga catogtgccc ctgaacgaga acagcgacaa cttaacctac
 601 cgcctgatca atgcaaacac cagcaaccac acccaggcct gcccaagggt gagcttgac
 661 cccatccca tccactactg cggcccgcgc ggctacgca tctgtgagtg caacaacaa
 721 accttaacg gcaaccggcc ctgtctaac gtagagaccg tgcagtgac ccaacggcac
 781 aagccgtgg tgaacacca gctgtgtctg aacggcagcc tggcctgaga gggcaltac
 841 atccgacagc agaactgac cgaagaacac aagaccatca tctgtacct gaacgagagc
 901 gtggagatca actgcaaccg ccccaacaa acaccccgca agagcgtgag catcgccccc
 961 ggcagagcct tctacgcaac caacgacgtg atcggcaca tccgcccagg ccaactgcaac
 1021 atcagcaccg accgttgaa caagaccctg cagcaggtga tgaagaagct gggcagagca
 1081 ttcccaaca agaccatca gttcaagccc cagcccgagg ggcacctgga gatcaccaatg
 1141 cacagcttca actgcccggc cgaattcttc tactgaaca ccagcaacct gttcaacagc
 1201 acctaccaca gcaacaacgg cactatacag tacaacggca acagcagcag cccatcacc
 1261 ctgcagtga agatcaagca gatcgtgcgc atgtggcagg gcgtggccca ggcacactac
 1321 gcccccccca tgcctggcaa catcacttgc cgcagcaaca tcaacggcat cctgtgtacc
 1381 cggcagggcg gcttaacac caccacaac accgagacct tccgcccagg cggcggcgac
 1441 atgcgcgaca actggcgtag ctagctgtac aagtacaagg ttgtggagat caagcccctg
 1501 ggcacgccc caccacaggc caagcgcgcg gtgtgtgagc gtagaagagc cgcctgtggc
 1561 atcggcgccg tttctctggg ctctctgggc ggcgcggca gcaacctggg cgcgcgacg
 1621 atcaccctga cctgtcaggc cgcgcagctg ctgacgggca tctgtcagca gcagagcaac
 1681 ctgttgagg ccatcgaggc ccagcagcac atgtctcagc tgaactgtg gggcatcaag
 1741 cagctgcagg cccgcgtgtt ggcctcgag cgtactga aggaccagca gctgtctggc
 1801 atctgggcti gcagcggccg cctgtatgac accacggccg tgccttgaa cagcagctg
 1861 agcaacaaga gcgagaagga catctgggac aacatgacct ggaatcagtg ggaacggag
 1921 atcagcaact acacggcct gactacaac ctgtggaggg acagccagaa ccagcaggag
 1981 aagaacgaga aggacctgtt gtagctgac aagtggaaca acctgtgaa ctgtgtcagc
 2041 atcagcaact ggccttgga catctaacg gag

FIGURE 102

WO 02/04493

110/114

PCT/US01/21241

NeD125G_TV2_C_ZAopt (SEQ ID NO:134)

ATGGGCGGCAAGTGGAGCAAGAGCAGCATCATCGGCTGGCCCGAGGTGCGC
GAGCGCATCCGCCGACCCGCGCGCCGCGAGGGCGTGGGCAGCGCCAGC
CAGGACCTGGAGAAGCACGGCGCCCTGACCACCAGCAACACCGCCACAAC
AACGCCGCTGCGCTGGCTGGAGGCCAGGAGGAGGAGGGCGAGGTGGGC
TTCCCCGTGCGCCCGAGGTGCCCCGTGCGCCCATGACCTACAAGGCCGCAT
CGACCTGAGCTTCTTCTGAAGGAGAAGGGCGGCTGGAGGGCTGATCTAC
AGCAAGAAGCGCCAGGAGATCCTGGACCTGTGGGTGTACAACACCCAGGGC
TTCTTCCCCGGCTGGCAGAACTACACCCCGGCCCGGCGTGGCTTCCCCCT
GACCTTCGGCTGGTACTTCAAGCTGGAGCCCGTGGACCCCGCGAGGTGGAG
GAGGCCAACGAGGGCGAGAACAACCTGCCTGCTGCACCCCATGAGCCAGCAC
GCCATGGAGGACGAGGACCGGAGGTGCTGCGCTGGAAGTTCGACAGCAC
CTGGCCCGCCGACATGGCCCGGAGCTGCACCCGAGTACTACAAGGACT
GCTGA

FIGURE 103

WO 02/04493

111/114

PCT/US01/21241

NeD125G-Myr_TV2_C_ZAopt (SEQ ID NO:135)

ATGGCCGGCAAGTGGAGCAAGAGCAGCATCATCGGCTGGCCCGAGGTGCGC
GAGCGCATCCGCCGCAACCCGACGCGCGCCGAGGGCGTGGGCAGCGCCAGC
CAGGACCTGGAGAAGCACGGCGCCCTGACCACCAGCAACACCGCCACAAC
AACGCCGCTGCGCTGGCTGGAGGCCAGGAGGAGGAGGGCGAGGTGGGC
TTCCCGTGCGCCCGCAGGTGCCCTGCGCCCATGACCTACAAGGCCGCAT
CGACCTGAGCTTCTTCTGAAGGAGAAAGGCGGCCTGGAGGGCTGATCTAC
AGCAAGAAGCGCCAGGAGATCTGGACCTGTGGGTGTACAACACCCAGGGC
TTCTTCCCGGCTGGCAGAACTACACCCCGGCCCGCGGTGCGCTTCCCCT
GACCTTCGGCTGGTACTTCAAGCTGGAGCCCGTGGACCCCGCGAGGTGGAG
GAGGCCAACGAGGGCGAGAACAACTGCCTGCTGCACCCCATGAGCCAGCAC
GGCATGGAGGACGAGGACCGCGAGGTGCTGCGCTGGAAGTTGACAGCAC
CTGGCCCGCCGCACATGGCCCGCGAGCTGCACCCGAGTACTACAAGGACT
GCTGA

FIGURE 104

WO 02/04493

112/114

PCT/US01/21241

↓: is the regions for β -sheet deletions

*: is the N-linked glycosylation sites for subtype C TV1 and TV2. Possible mutation (N→Q) or deletions can be performed.

		1		50
SF162	(1)	---MDAMKRGSLCCFLIIGCAVFFSPSAKELTATVWVQVYANPSTV		
TV1.8.2	(1)	MRVNGTQKNQCCQWINGILSPWMLMIONTEDEYLLVQVQVWREKAK		
TV1.8.5	(1)	MRVNGTQKNQCCQWINGILSPWMLMIONTEDEYLLVQVQVWREKAK		
TV2.12-5/1	(1)	MRVNGTQKNQCCQWINGILSPWMLMIONTEDEYLLVQVQVWREKAK		
Consensus	(1)	MRVNGTQKNQCCQWINGILSPWMLMIONTEDEYLLVQVQVWREKAK		
		51		100
SF162	(47)	PCASDAKAYETEVENVATHACVPTDENPQEVILGNVTENPMWKNKVD		
TV1.8.2	(51)	PCASDAKAYETEVENVATHACVPTDENPQEVILGNVTENPMWKNKVD		
TV1.8.5	(51)	PCASDAKAYETEVENVATHACVPTDENPQEVILGNVTENPMWKNKVD		
TV2.12-5/1	(51)	PCASDAKAYETEVENVATHACVPTDENPQEVILGNVTENPMWKNKVD		
Consensus	(51)	PCASDAKAYETEVENVATHACVPTDENPQEVILGNVTENPMWKNKVD		
		101		150
SF162	(97)	QMHEDIISLWDSLPCVKLTPLCVTLNCTINVTGNRTVIGNSNSN		
TV1.8.2	(101)	QMHEDIISLWDSLPCVKLTPLCVTLNCTINVTGNRTVIGNSNSN		
TV1.8.5	(101)	QMHEDIISLWDSLPCVKLTPLCVTLNCTINVTGNRTVIGNSNSN		
TV2.12-5/1	(101)	QMHEDIISLWDSLPCVKLTPLCVTLNCTINVTGNRTVIGNSNSN		
Consensus	(101)	QMHEDIISLWDSLPCVKLTPLCVTLNCTINVTGNRTVIGNSNSN		
		151		200
SF162	(139)	NKEMDRGELENSEKRRSTENKMKQKAYALFVGLVATLON---		
TV1.8.2	(151)	TSIYNLEEMKNCENATLELDKKRKHRYALFGLVATLON--ENS		
TV1.8.5	(151)	NATYKYEMKNCENATLELDKKRKHRYALFGLVATLON--ENS		
TV2.12-5/1	(141)	-----KQNKNGFNVTELEDKKRKHRYALFGLVATLON--ENS		
Consensus	(151)	A Y ENKNGGFNVTELEDKKRKHRYALFGLVATLON ENSNFTY		
		201		250
SF162	(185)	RLINCNSTITQACPKVSDPIPIHYCAPAGVALLKCNMTFNGTGPCYN		
TV1.8.2	(199)	RLINCNSTITQACPKVSDPIPIHYCAPAGVALLKCNMTFNGTGPCYN		
TV1.8.5	(199)	RLINCNSTITQACPKVSDPIPIHYCAPAGVALLKCNMTFNGTGPCYN		
TV2.12-5/1	(185)	RLINCNSTITQACPKVSDPIPIHYCAPAGVALLKCNMTFNGTGPCYN		
Consensus	(201)	RLINCNSTITQACPKVSDPIPIHYCAPAGVALLKCNMTFNGTGPCYN		
		251		300
SF162	(235)	VSTVQCTRGIKPVVSTQILLINGSLAREGIIIRSEMLTENTKTIIVH		
TV1.8.2	(249)	VSTVQCTRGIKPVVSTQILLINGSLAREGIIIRSEMLTENTKTIIVH		
TV1.8.5	(249)	VSTVQCTRGIKPVVSTQILLINGSLAREGIIIRSEMLTENTKTIIVH		
TV2.12-5/1	(235)	VSTVQCTRGIKPVVSTQILLINGSLAREGIIIRSEMLTENTKTIIVH		
Consensus	(251)	VSTVQCTRGIKPVVSTQILLINGSLAREGIIIRSEMLTENTKTIIVH		
		301		350
SF162	(285)	VEINCTRPWNTKESVRIGFGQAFYATNDIICNTROAHCTISTDRW		
TV1.8.2	(299)	VEINCTRPWNTKESVRIGFGQAFYATNDIICNTROAHCTISTDRW		
TV1.8.5	(299)	VEINCTRPWNTKESVRIGFGQAFYATNDIICNTROAHCTISTDRW		
TV2.12-5/1	(285)	VEINCTRPWNTKESVRIGFGQAFYATNDIICNTROAHCTISTDRW		
Consensus	(301)	VEINCTRPWNTKESVRIGFGQAFYATNDIICNTROAHCTISTDRW		

FIGURE 105A

WO 02/04493

113/114

PCT/US01/21241

```

351      *                               *                               400
SF162 (335) KQVTKLHACGPKT - EVKQSSGGQPELVDRNSGSEKMAEPAAG
TV1.8.2 (349) QQVMKLGSEHPNKT - YQKPHAGDLITMRSPNCRGSEFFYCSTNLEN
TV1.8.5 (349) QQVMKLGSEHPNKT - YQKPHAGDLITMRSPNCRGSEFFYCSTNLEN
TV2.12-5/1 (335) QRVQCGQLELNGGCTKAPHSGDLETTSEK - SLEKATNDRN
Consensus (351) QQVMKLGSEHPNKT IKKPHAGDLITMRSPNCRGSEFFYCSTNLEN

401      *                               *                               ↓ β20/β21 ↓ 450
SF162 (384) STYNN-----SIGPH - NNGTSTHPEYTGIDR - ESKMAEPAAG
TV1.8.2 (398) STYHS---NNGTYKYGNSSPITDCH - GQDMGQVQAMYPPIAG
TV1.8.5 (398) STYAP---KNQTYKYGNSSLPTDCH - GQDMGQVQAMYPPIAG
TV2.12-5/1 (385) STYENGTCINQCMNH - NDERYTLOHETKQ - DMGQVQAMYPPIAG
Consensus (401) STYNN NGTYKYGNS S PITLQCKIKQLIKMWQGVQAMYPPIAG

451      *                               *                               *                               500
SF162 (427) QRCSSNITGILLTRDGGKSLNT - TETFRPGGDMRDNRSELYKYK
TV1.8.2 (445) NITCRSNITGILLTRDGGKSLNT - TETFRPGGDMRDNRSELYKYK
TV1.8.5 (445) NITCRSNITGILLTRDGGKSLNT - TETFRPGGDMRDNRSELYKYK
TV2.12-5/1 (433) NITCRSNITGILLTRDGGKSLNT - TETFRPGGDMRDNRSELYKYK
Consensus (451) NITCRSNITGILLTRDGGKSLNT TETFRPGGDMRDNRSELYKYK

501      *                               *                               *                               550
SF162 (475) VEKPELQVAPKAKRERVQREKRAVGIGAVPLGFGAAGSTMGAASITLT
TV1.8.2 (493) VEKPELQVAPKAKRERVQREKRAVGIGAVPLGFGAAGSTMGAASITLT
TV1.8.5 (495) VEKPELQVAPKAKRERVQREKRAVGIGAVPLGFGAAGSTMGAASITLT
TV2.12-5/1 (480) VEKPELQVAPKAKRERVQREKRAVGIGAVPLGFGAAGSTMGAASITLT
Consensus (501) VEKPELQVAPKAKRERVQREKRAVGIGAVPLGFGAAGSTMGAASITLT

551      *                               *                               *                               600
SF162 (525) VQARQLSGIVQQQSNLLKAIKRAQQMLQLTVWGIKQLQARVLATERYLK
TV1.8.2 (543) VQARQLSGIVQQQSNLLKAIKRAQQMLQLTVWGIKQLQARVLATERYLK
TV1.8.5 (545) VQARQLSGIVQQQSNLLKAIKRAQQMLQLTVWGIKQLQARVLATERYLK
TV2.12-5/1 (530) VQARQLSGIVQQQSNLLKAIKRAQQMLQLTVWGIKQLQARVLATERYLK
Consensus (551) VQARQLSGIVQQQSNLLKAIKRAQQMLQLTVWGIKQLQARVLATERYLK

601      *                               *                               *                               *                               650
SF162 (575) DQQLLGIVGCGKLLCTTAVFWNSNSWNSKSEADINDGNTMWDREISNY
TV1.8.2 (593) DQQLLGIVGCGKLLCTTAVFWNSNSWNSKSEADINDGNTMWDREISNY
TV1.8.5 (595) DQQLLGIVGCGKLLCTTAVFWNSNSWNSKSEADINDGNTMWDREISNY
TV2.12-5/1 (580) DQQLLGIVGCGKLLCTTAVFWNSNSWNSKSEADINDGNTMWDREISNY
Consensus (601) DQQLLGIVGCGKLLCTTAVFWNSNSWNSKSEADINDGNTMWDREISNY

651      *                               *                               *                               700
SF162 (625) TETTYTLEKSNV - ESKCHN - QKACHN - ESKCHN - ESKCHN - ESKCHN
TV1.8.2 (643) TETTYTLEKSNV - ESKCHN - QKACHN - ESKCHN - ESKCHN - ESKCHN
TV1.8.5 (645) TETTYTLEKSNV - ESKCHN - QKACHN - ESKCHN - ESKCHN - ESKCHN
TV2.12-5/1 (630) TETTYTLEKSNV - ESKCHN - QKACHN - ESKCHN - ESKCHN - ESKCHN
Consensus (651) TETTYTLEKSNV ESKCHN QKACHN ESKCHN ESKCHN ESKCHN

701      *                               *                               *                               750
SF162 (675) VEGGLIGRLIFAVLSIVNEVRQGSPLSFQTLTPEPGPDRLGGLIEBGG
TV1.8.2 (693) VEGGLIGRLIFAVLSIVNEVRQGSPLSFQTLTPEPGPDRLGGLIEBGG
TV1.8.5 (695) VEGGLIGRLIFAVLSIVNEVRQGSPLSFQTLTPEPGPDRLGGLIEBGG
TV2.12-5/1 (680) VEGGLIGRLIFAVLSIVNEVRQGSPLSFQTLTPEPGPDRLGGLIEBGG
Consensus (701) VEGGLIGRLIFAVLSIVNEVRQGSPLSFQTLTPEPGPDRLGGLIEBGG

```

FIGURE 105B

WO 02/04493

114/114

PCT/US01/21241

```

751                                     800
SF162 (725) EQDRDSIRLVSGFLSLAWDDLRLCLFSTHRLDFILIAVRAVLLGHS
TV1.8_2 (743) EQDRDSIRLVSGFLSLAWDDLRLCLFSTHRLDFILIAVRAVLLGHS
TV1.8_5 (745) EQDRDSIRLVSGFLSLAWDDLRLCLFSTHRLDFILIAVRAVLLGHS
TV2.12-5/1 (730) EQDRDSIRLVSGFLSLAWDDLRLCLFSTHRLDFILIAVRAVLLGHS
Consensus (751) EQDRDSIRLVSGFLSLAWDDLRLCLFSTHRLDFILIAVRAVLLGHS
801                                     850
SF162 (774) -----EQDRDSIRLVSGFLSLAWDDLRLCLFSTHRLDFILIAVRAVLLGHS
TV1.8_2 (793) SLRGLQGRWILKYLGSVVOYMGLELKKSAISLLDTIALAVAGTDRIIE
TV1.8_5 (795) SLRGLQGRWILKYLGSVVOYMGLELKKSAISLLDTIALAVAGTDRIIE
TV2.12-5/1 (780) SLRGLQGRWILKYLGSVVOYMGLELKKSAISLLDTIALAVAGTDRIIE
Consensus (801) SLRGLQGRWILKYLGSVVOYMGLELKKSAISLLDTIALAVAGTDRIIE
851                                     876
SF162 (818) SADRIGRAFLHIFATYNSHTSAPL-
TV1.8_2 (843) LVQRICRAILNIPRRIRQGFRAALL
TV1.8_5 (845) LVQRICRAILNIPRRIRQGFRAALL
TV2.12-5/1 (830) FRSATCEIRKWFPHETLSPFAMIQ-
Consensus (851) LVQRICRAILNIPRRIRQGFRAALL

```

FIGURE 105C

【国際公開パンフレット（コレクトバージョン）】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/004493 A3(51) International Patent Classification: C07K 14/16
C12N 5/10, A61K 39/22(74) Agents: DOLLARD, Anne, S. et al.; Chiron Corporation,
Intellectual Property - R140, P.O. Box 8097, Emeryville,
CA 94662-8097 (US).

(21) International Application Number: PCT/US01/21241

(22) International Filing Date: 5 July 2001 (05.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 09/618,513 5 July 2000 (05.07.2000) US

(71) Applicants (for all designated States except US): CHIRON CORPORATION [US/US]; 4560 Horton Street,
Emeryville, CA 94608 (US); UNIVERSITY OF STELLENBOSCH [ZA/ZA]; P.O. Box 19063, 7505 Tygerberg
(ZA).(81) Designated States (national): AF, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CT,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GR,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LA,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MY, NZ, NG, NI, NO, NL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SI, SZ, TZ, UG, ZW), European
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Eurasian
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PL, SE, TR), OAPI patent (BF, BI, CE,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ZUR MEGEDE,
Jan [DE/US]; c/o Chiron Corporation, 4560 Horton Street,
Emeryville, CA 94608 (US); BARNETT, Susan,
W. [US/US]; c/o Chiron Corporation, 4560 Horton Street,
Emeryville, CA 94608 (US); ENGELBRECHT,
Susan [ZA/ZA]; c/o University of Stellenbosch, P.O.
Box 19063, 7505 Tygerberg (ZA); VAN RENSBURG,
Estrelita, June [ZA/ZA]; c/o University of Stellenbosch,
P.O. Box 19063, 7505 Tygerberg (ZA).

Published:

with international search report

(88) Date of publication of the international search report:
29 June 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/004493 A3

(54) Title: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

(57) Abstract: The present invention relates to polynucleotides encoding immunogenic HIV type C polypeptides. Uses of the polynucleotides in applications including DNA immunization, generation of packaging cell lines, and production of HIV Type C proteins are also described.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/21241
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/16 C12N5/10 A61K39/21		
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) IPC 7 C07K C12N A61K		
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 practical, search terms used)		
EMBL, BIOSIS, EP0-Internal, MPI Data, PAJ, CHEM ABS Data, MEDLINE, EMBASE, SCISEARCH		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WILLIAMSON C ET AL: "Designing HIV-1 subtype C vaccines for South Africa." SOUTH AFRICAN JOURNAL OF SCIENCE, vol. 96, no. 6, June 2000 (2000-06), pages 318-324, XP008689291 ISSN: 0038-2353 page 321-322 --- -/--	1,59-71, 78-84, 86-97
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents: (A) documents cited in the international search report which are not considered to be of particular relevance (B) earlier document but published on or after the international filing date (C) document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another document or other special reason (as specified) (D) document relating to actual disclosure, use, exhibition or other reasons (E) document published prior to the international filing date but later than the priority date claimed (F) later document published after the international filing date, or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention (G) 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 (H) document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other such documents, such combination being obvious to a person skilled in the art (I) document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
18 October 2002		31. 01. 2003
Name and mailing address of the ISA European Patent Office, P.O. Box 5010 Patentamt 2 NL - 2200 LV Rijswijk Tel: (+31-70) 340-2040, Tx: 31 651 epo nl, Fax: (+31-70) 340-3018		Authorized officer Brouns, G

Form PCT/ISA/210 (second sheet) July 1992

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/21241
C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	NOVITSKY VA ET AL: "Molecular cloning and phylogenetic analysis of human immunodeficiency virus type 1 subtype C: a set of 23 full-length clones from Botswana" JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 73, no. 5, May 1999 (1999-05), pages 4427-4432, XP002144689 ISSN: 0022-538X cited in the application	1,59-71, 78-84, 86-97
Y	the whole document -& DATABASE EMBL [Online] 28 February 1999 (1999-02-28) NOVITSKY VA ET AL.: "HIV-1 isolate C-96BN15C05 country Botswana, complete genome" retrieved from E.B.I., HINXTON, U.K. Database accession no. AF110975 XP002217412 cited in the application abstract	1
Y	WO 00 15819 A (CHILDRENS MEDICAL CENTER) 23 March 2000 (2000-03-23) claim 42	1,59-71, 78-84, 86-97
Y	MEGEDE J Z ET AL: "INCREASED EXPRESSION AND IMMUNOGENICITY OF SEQUENCE-MODIFIED HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 GAG GENE" JOURNAL OF VIROLOGY, THE AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 74, no. 6, March 2000 (2000-03), pages 2628-2635, XP000980650 ISSN: 0022-538X page 2631, column 2, paragraph 2; table 2	1,59-71, 78-84, 86-97
Y	KENT STEPHEN J ET AL: "A recombinant avipoxvirus HIV-1 vaccine expressing interferon-gamma is safe and immunogenic in macaques." VACCINE, vol. 18, no. 21, 28 April 2000 (2000-04-28), pages 2250-2256, XP004191009 ISSN: 0264-410X abstract	61,81

	-/-	

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 01/21241
C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE AIDSLINE [Online] National Library of Medicine; 11th Int. Conf. AIDS, 1996, July 7-12, 1996 HAMAJIMA ET AL.: "The combination of DNA and peptide vaccines induces strong immunities against HIV-1 in both humoral and CMI" retrieved from HTTP://WWW.AEGIS.COM/PUBS/AIDSLINE/1997/JAN/M9/164 Database accession no. JCA11/96920829 XP002217413 abstract	93-96
A	HAAS ET AL: "Cytotoxic T-cell responses to HIV-1 reverse transcriptase, integrase and protease" AIDS, LONDON, GB, vol. 12, no. 12, 1998, pages 1427-1436, XP002117622 ISSN: 0269-9370 the whole document	78-97

INTERNATIONAL SEARCH REPORT		International application No. PC1/US 01/21241		
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)				
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:				
1. ☒	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Although claims 78-97 are directed to a method of treatment of the human or animal body, the search has been carried out and based on the alleged effects of the composition.			
2. ☐	Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:			
3. ☐	Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).			
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)				
This International Searching Authority found multiple inventions in this International application, as follows:				
see additional sheet				
1. ☐	As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.			
2. ☐	As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.			
3. ☐	As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:			
4. ☒	No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1 in part, 59-97 in part as far as applicable			
<table style="width: 100%; border: none;"> <tr> <td style="width: 30%; vertical-align: top;">Remark on Protest</td> <td style="border-left: 1px solid black; padding-left: 10px;"> ☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees. </td> </tr> </table>			Remark on Protest	☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees.
Remark on Protest	☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees.			

International Application No. PCT/US 01/21241

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 218

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1 in part, 59-97 all in part, as far as applicable

Invention 1

An expression cassette comprising a polynucleotide sequence having at least 90% sequence identity to the sequence of SEQ ID NO:30, further comprising one or more nucleic acids encoding one or more viral polypeptides or antigens, or cytokines. Said expression cassette in a recombinant expression system for use in a selected host cell, a cell comprising said expression cassette, a composition for generating an immunological response, comprising said expression cassette and a method of immunization of a subject using said composition.

2. Claims: 2-46, 59-73, 75-97, all as far as applicable

Inventions 2-53

As for invention 1, wherein
invention 2 is limited to SEQ ID NO:31, invention 3 is limited to SEQ ID NO:32, invention 4 is limited to SEQ ID NO:46, invention 5 is limited to SEQ ID NO:47, invention 6 is limited to SEQ ID NO:49, invention 7 is limited to SEQ ID NO:97, invention 8 is limited to SEQ ID NO:51, invention 9 is limited to SEQ ID NO:99, invention 10 is limited to SEQ ID NO:55, invention 11 is limited to SEQ ID NO:57, invention 12 is limited to SEQ ID NO:101, invention 13 is limited to SEQ ID NO:96, invention 14 is limited to SEQ ID NO:134, invention 15 is limited to SEQ ID NO:135, invention 16 is limited to SEQ ID NO:58, invention 17 is limited to SEQ ID NO:60, invention 18 is limited to SEQ ID NO:62, invention 19 is limited to SEQ ID NO:103, invention 20 is limited to SEQ ID NO:64, invention 21 is limited to SEQ ID NO:66, invention 22 is limited to SEQ ID NO:68, invention 23 is limited to SEQ ID NO:70, invention 24 is limited to SEQ ID NO:72, invention 25 is limited to SEQ ID NO:74, invention 26 is limited to SEQ ID NO:76, invention 27 is limited to SEQ ID NO:78, invention 28 is limited to SEQ ID NO:80, invention 29 is limited to SEQ ID NO:81, invention 30 is limited to SEQ ID NO:109, invention 31 is limited to SEQ ID NO:83, invention 32 is limited to SEQ ID NO:111, invention 33 is limited to SEQ ID NO:85, invention 34 is limited to SEQ ID NO:113, invention 35 is limited to SEQ ID NO:87, invention 36 is limited to SEQ ID NO:115, invention 37 is limited to SEQ ID NO:89, invention 38 is limited to SEQ ID NO:117, invention 39 is limited to SEQ ID NO:91, invention 40 is limited to SEQ ID NO:93, invention 41 is limited to SEQ ID NO:94, invention 42 is limited to SEQ ID NO:119, invention 43 is limited to SEQ ID NO:120, invention 44 is limited to SEQ ID NO:121, invention 45 is limited to SEQ ID

International Application No. PCT/US 01/21241

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

NO:122, invention 46 is limited to SEQ ID NO:123, invention 47 is limited to SEQ ID NO:124, invention 48 is limited to SEQ ID NO:125, invention 49 is limited to SEQ ID NO:126, invention 50 is limited to SEQ ID NO:127, invention 51 is limited to SEQ ID NO:131, invention 52 is limited to SEQ ID NO:132, invention 53 is limited to SEQ ID NO:133,

3. Claims: 47-51

Invention 54

A polynucleotide comprising SEQ ID NO:33 or fragments derived therefrom

4. Claims: 52-56,74

Invention 55

A polynucleotide comprising SEQ ID NO:45 or fragments derived therefrom

5. Claims: 57,58

Invention 56

A polynucleotide comprising SEQ ID NO:128 or fragments derived therefrom

INTERNATIONAL SEARCH REPORT				International Application No. PCT/US 01/21241	
Information on patent family members					
Patent document cited in search report	Publication date	Patent family member(s)	Publication date		
WO 0015819	A	23-03-2000	AU 6139699 A	03-04-2000	
		WO 0015819 A1	23-03-2000		

Form PCT/ISA/210 (patent family member) (July 1992)

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
A 6 1 P 31/12	A 6 1 P 31/18	
A 6 1 P 31/18	C 1 2 N 1/15	
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 N 5/00	A
C 1 2 N 5/10	A 6 1 K 37/02	

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

- (72)発明者 ツア メゲーデ, ヤン
アメリカ合衆国 カリフォルニア 94608, エミリービル, ホートン ストリートアー
ル440 4560, カイロン コーポレイション気付
- (72)発明者 バーネット, スーザン ダブリュー.
アメリカ合衆国 カリフォルニア 94608, エミリービル, ホートン ストリートアー
ル440 4560, カイロン コーポレイション気付
- (72)発明者 エンジェルブレヒト, スーザン
南アフリカ 7505 タイガーバーグ, ピー. オー. ボックス 19063, ユニバーシ
ティ オブ ステレンボスク気付
- (72)発明者 ヴァン レンズバーグ, エストレーリタ ジャンセ
南アフリカ 7505 タイガーバーグ, ピー. オー. ボックス 19063, ユニバーシ
ティ オブ ステレンボスク気付

F ターム(参考) 4B024 AA01 BA35 CA04 CA07 DA01 DA02 DA05 DA11 EA01 EA02
EA03 EA04 FA02 GA11
4B065 AA01X AA57X AA87X AA95Y AB01 BA01 CA23 CA24 CA45
4C084 AA02 AA13 BA44 CA01 MA02 NA14 ZB332 ZC552
4C085 BA69 BB11 CC21
4C087 AA01 AA02 BC83 CA12 NA14 ZB33 ZC55