

US 20220062375A1

# (19) United States (12) Patent Application Publication (10) Pub. No.: US 2022/0062375 A1 STINCHCOMB et al.

## Mar. 3, 2022 (43) **Pub. Date:**

### (54) COMPOSITIONS AND METHODS FOR **DENGUE VIRUS CHIMERIC** CONSTRUCTIONS IN VACCINES

- (71) Applicants: TAKEDA VACCINES, INC., Cambridge, MA (US); THE **GOVERNMENT OF THE UNITED** STATES OF AMERICA AS **REPRESENTED BY THE** SECRETARY OF THE DEPARTMENT, Atlanta, GA (US)
- (72) Inventors: Dan T. STINCHCOMB, Enumclaw, WA (US); Claire KINNEY, Fort Collins, CO (US); Richard M. KINNEY, Fort Collins, CO (US); Jill A. LIVENGOOD, Fort Collins, CO (US)
- (73) Assignees: TAKEDA VACCINES, INC., Cambridge, MA (US); THE **GOVERNMENT OF THE UNITED** STATES OF AMERICA AS **REPRESENTED BY THE** SECRETARY OF THE **DEPARTMENT**, Atlanta, GA (US)
- (21) Appl. No.: 17/478,537
- (22) Filed: Sep. 17, 2021

#### **Related U.S. Application Data**

Continuation of application No. 16/561,755, filed on (60)Sep. 5, 2019, now abandoned, which is a continuation of application No. 15/492,981, filed on Apr. 20, 2017, now Pat. No. 10,449,231, which is a division of application No. 14/209,808, filed on Mar. 13, 2014, now Pat. No. 9,783,579.

(60)Provisional application No. 61/800,204, filed on Mar. 15, 2013.

#### **Publication Classification**

(51) Int. Cl.

A61K 38/16	(2006.01)
A61K 31/7048	(2006.01)
A61K 39/12	(2006.01)
A61K 31/713	(2006.01)
C07K 14/005	(2006.01)
C07K 14/18	(2006.01)
C07K 19/00	(2006.01)
C12N 15/09	(2006.01)
C12N 15/861	(2006.01)

(52) U.S. Cl.

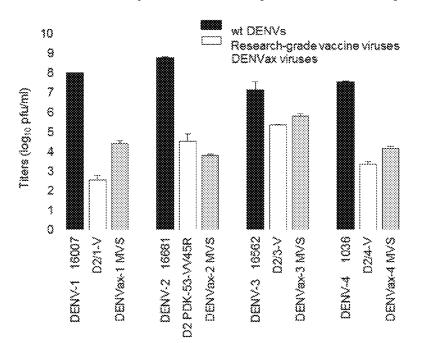
(57)

CPC ...... A61K 38/162 (2013.01); A61K 31/7048 (2013.01); A61K 39/12 (2013.01); A61K 31/713 (2013.01); A61K 2039/5254 (2013.01); C07K 14/1825 (2013.01); C07K 19/00 (2013.01); C12N 15/09 (2013.01); C12N 15/8613 (2013.01); C07K 14/005 (2013.01)

#### ABSTRACT

Embodiments herein report compositions, uses and manufacturing of dengue virus constructs and live attenuated dengue viruses. Some embodiments concern a composition that includes, but is not limited to, a tetravalent dengue virus composition. In certain embodiments, compositions can include constructs of one or more serotypes of dengue virus, such as dengue-1 (DEN-1) virus, dengue-2 (DEN-2) virus, dengue-3 (DEN-3) or dengue-4 (DEN-4) virus constructs. In other embodiments, constructs disclosed herein can be combined in a composition to generate a vaccine against more one or more dengue virus constructs that may or may not be subsequently passaged in mammalian cells.

#### Specification includes a Sequence Listing.



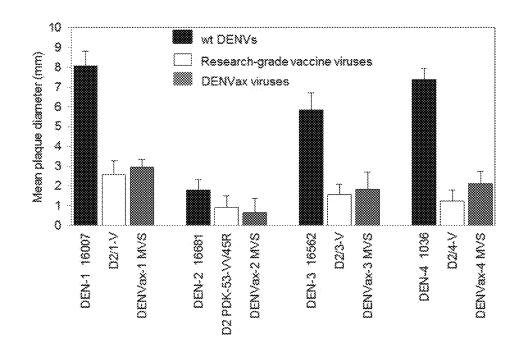
Genetic variations among	ons amo		chimer	as (com	pared to	wt D2 1(	D2/4 chimeras (compared to wt D2 16681 and D4-1036)	04-40	9
Genome		D2		junction		Ó	D4		junction
Genes	NCR		C//	prW					
Mutation types*	PDK-53	seed	Eng	Mul	seed	Marker	Eng	Eng	NgoMIV
Genome NT	C57T	A225T	A396C	A453G	C647G/C	A1401G	C2027T	A2275	TG2380/1CC
Protein-AA position NCR	NCR	C-silent	C-R100S	prM- silant	Prin- Tract	E-silent	E-A364V	Ц МИД 71	E-V482A
D7_16681		A	A(R)						TGW
D2-PDK-53	> +	¢ .		5					
D4-1036			· · · ·		G(T)	A	C(A)	A(M)	
Cloned D2/4-V1	L.	- / -	.c(s)	ő	1	6	(V)	(C(L)	CC(A)
(pD2/4-VP1)									
DENVax-4 (MVS)	1	t /	C(S)	g	-	g	T(V)	C(L)	CC(A)
<sup>"</sup> .": same as wt D2 16681 or D4 1036	6681 or D4	1036; sme	all nt letter:	silent muta	; small nt letter: silent mutation in open reading region	reading re(	gion		
*: PDK-53: D2-PDK-53 specific genotype (VS 16681), Italics: major attenuation PDK-53 loci; Seed: mutations found only	53 specific (	jenotype (	VS 16681)	, Italics: m	ajor attenu	ation PDK.	53 loci, Se	ed: mutat	ons found only
in specified virus seed and not in the original clone; Eng: Engineered mutations for the D2/4 clones; Miu and NgoMV:	seed and no	t in the on	ginal clone	; Eng: Engi	neered mut	ations for th	le D2/4 clon	es; Mlu a	nd NgoMV:
**: C8571T (PDK-53 silent mutation)	silent mutat	sites, fion) was r	not included	l in most D'	was not included in most D2/4 chimeric clones	clones			

FIG. 1

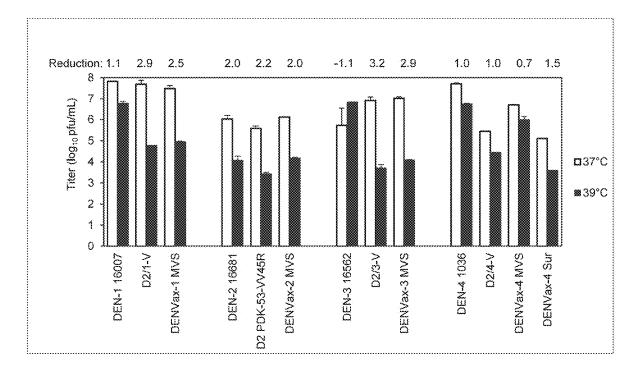
		77	A9750C			
	35	seed	A97	NS5- silent	A -	, 0,
	NS5	PDK53	C8571T**	NS5-silent		
	NS4B	seed	G6599C 17026C/T	NS4B- silent	- -	- ctt
	NS4A	PDK-53	G6599C	NS4A- G75A	G(G) C(A)	C(A) C(A)
	$\mathbf{N}$	seed	T5547C C6437T	NS4A- A21V	C(A)	- I(V)
2///		PDK-53 seed		NS3- silent	L C	<b>9 9</b>
/ D2	NS3	R. N	C5391T	NS3- silent	- -	
		PDK-53	A5270T	NS3- E250V	A(E) T(V)	TCN
		PDK-53	C4018T	NS2A- L181F	C(L) T(F)	.T(F) .T(F)
	/ NS2A/		32579A A3674G A3773A/G	NS2A- K99K/R	A(K)	- A/G(K/R)
		seed	A3674G	NS2A- D66G	A(D)	- (G(G)
	NS1		G2579A	NS1- G53D	G(G0 A(D)	(D) (D)

FIG. 1 (continued)

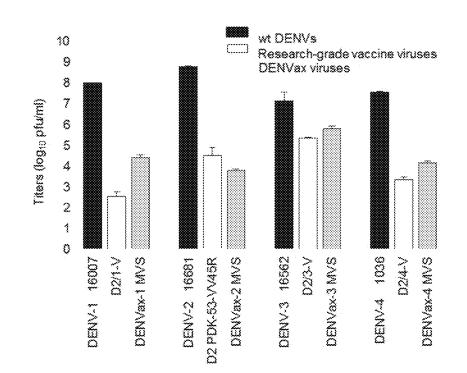
# Fig. 2



# Fig. 3







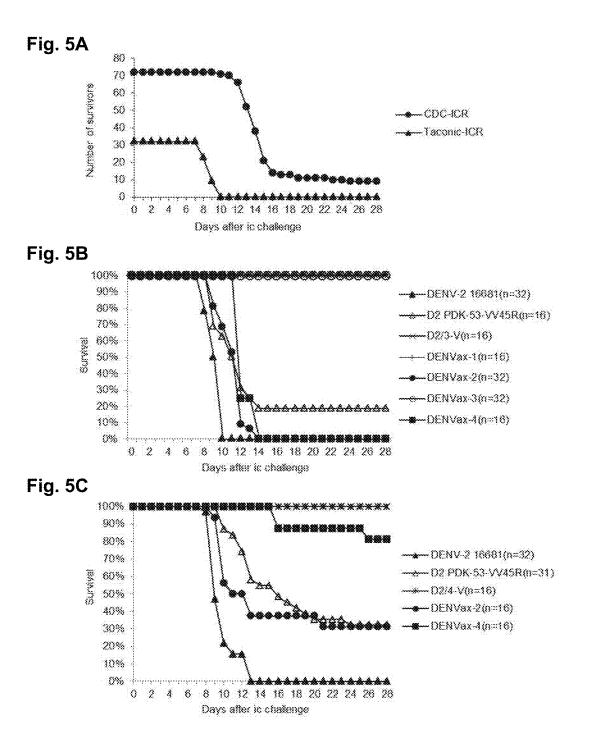


Fig. 6

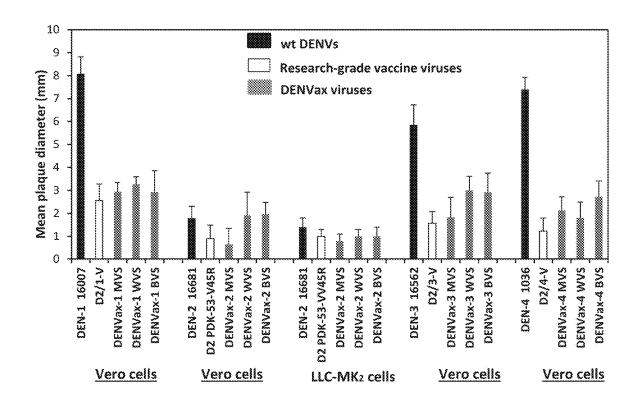


Fig. 7

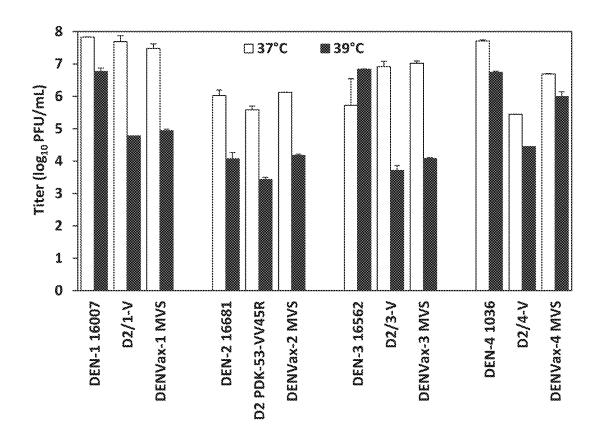
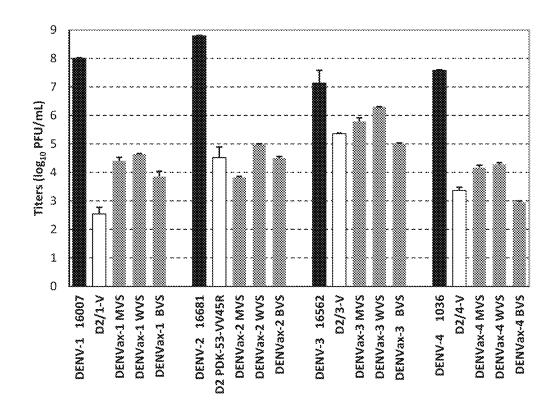
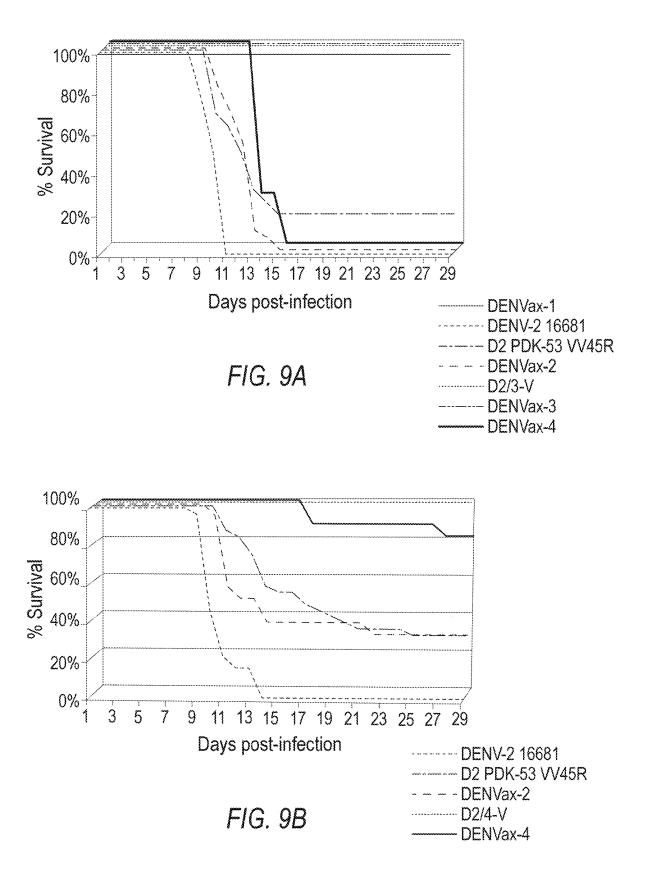


Fig. 8





Serntvne	Strain	virus origin	<u>C57T</u> 5'NCR	A524T	T900C <sup>a</sup> M (silent)
	18804			· · · · · · · · · · · · · · · · · · ·	T
	10001		د 	¢	-
	PDK-53	PDK cell pass of 16681	┣	<b>-</b>	Ŗ
	PDK-53-V(VV45R)	Recombinant PDK-53-V	┣	┢╼╾	U
	PDK-53-E(VE48R)	Recombinant PDK-53-E	┣	<u>+-</u>	U
<u>Underline</u>	<u>d</u> Mutations: the 3 mc	Underlined Mutations: the 3 most important attenuation loci of PDK-53	K-53		
Italics form	t: PDK-53 specific set	Italics font: PDK-53 specific sequence (change from 16681)			
<b>Bold font</b>	<b>Bold font: Different nt sequen</b>	nce between PDK-53 and clone-derived V or E virus	terived V or	E virus	
<sup>a</sup> Engineer	ed silent clone marke	<sup>a</sup> Engineered silent clone marker to differenciate original PDK-53 and recombinant (clone-derived) viruses	ind recombin:	ant (clone-deni	ved) viruses

FIG. 10

r.	3,	2022	Sheet	12	of	Ľ

					]
C8571T NS5 (silent	0	مۇمىرى -	U	U	
T5547C G6599C C8571T NS3 (silent) NS4A-G75A NS5 (silent)	ග	C	S	S	*************************
T5547C NS3 (silent)		c	c	0	
C4018T <u>A5270T</u> NS2A-L181 <u>NS3-E250V</u>	A	T/A mix	<b>j</b>	Z	
C4018T NS2A-L181	c	<b>I</b>	<b></b>	┣	
<u>G2579A</u> NS1-G53D	9	A	A	٨	******
C2055T E (silent)	0	*~~	-tour	-tour	******

FIG. 10 (continued)

#### COMPOSITIONS AND METHODS FOR DENGUE VIRUS CHIMERIC CONSTRUCTIONS IN VACCINES

#### CROSS REFERENCE TO PRIOR APPLICATIONS

**[0001]** This patent application is a continuation of U.S. patent application Ser. No. 16/561,755, filed Sep. 5, 2019, which is a U.S. continuation of U.S. patent application Ser. No. 15/492,981, filed on Apr. 20, 2017, which is a U.S. divisional application that claims priority to U.S. patent application Ser. No. 14/209,808, filed Mar. 13, 2014, which claims the benefit of U.S. Provisional Patent Application No. 61/800,204, filed Mar. 15, 2013, the disclosures of which are all incorporated herein by reference in their entireties.

#### STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

**[0002]** This invention was made with government support under Grant No. R43 AI084291-01 awarded by the National Institutes of Health. The Government has certain rights in the invention.

#### FIELD

[0003] Embodiments herein report compositions, methods, uses and manufacturing procedures for dengue virus constructs and vaccine compositions thereof. Some embodiments concern a composition that includes, but is not limited to, chimeric flavivirus virus constructs that alone or in combination with other constructs can be used in a vaccine composition. In certain embodiments, compositions can include constructs of more than one serotypes of dengue virus, such as dengue-1 (DEN-1) virus, dengue-2 (DEN-2) virus, dengue-3 (DEN-3) virus and/or dengue-4 (DEN-4) virus. In other embodiments, manufacturing strategy that can improve the safety and genetic stability of recombinant live-attenuated chimeric dengue vaccine (DENVax) viruses. Certain embodiments include at least one live, attenuated dengue virus in combination with dengue virus chimeric constructs identified to be both safe and effective in vaccine compositions where the constructs have undergone additional passages in cell cultures.

#### BACKGROUND

**[0004]** Infection with dengue virus can lead to a painful fever of varying severity. To date, four serotypes of dengue virus have been identified: dengue-1 (DEN-1), dengue-2 (DEN-2), or dengue-3 (DEN-3) in combination with dengue-4 (DEN-4). Dengue fever is caused by infection of a dengue virus. Other subtypes may be discovered in the future (e.g. DEN-5). Dengue virus serotypes 1-4 can also cause dengue hemorrhagic fever (DHF), and dengue shock syndrome (DSS). The most severe consequences of infection, DHF and DSS, can be life threatening. Dengue viruses cause 50-100 million cases of debilitating dengue fever, 500,000 cases of DHF/DSS, and more than 20,000 deaths each year. To date, there is no

effective vaccine to protect against dengue fever and no drug treatment for the disease. Mosquito control efforts have been ineffective in preventing dengue outbreaks in endemic areas or in preventing further geographic spread of the disease. It is estimated that 3.5 billion people are threatened by infection with dengue virus. In addition, dengue virus is a leading cause of fever in travelers to endemic areas, such as Asia. Central and South America, and the Caribbean.

[0005] All four dengue virus serotypes are endemic throughout the tropical regions of the world and constitute the most significant mosquito-borne viral threat to humans in tropical regions, worldwide. Dengue viruses are transmitted to humans primarily by Aedes aegypti mosquitoes. Infection with one dengue virus serotype results in life-long protection from re-infection by that serotype, but does not prevent secondary infection by one of the other three dengue virus serotypes. In fact, previous infection with one dengue virus serotype leads to an increased risk of severe disease (DHF/DSS) upon secondary infection with a different serotype. The development of an effective vaccine represents an important approach to the prevention and control of this global emerging disease. Multiple immunizations make complete vaccine coverage difficult both for public health efforts in dengue virus endemic countries as well as travelers.

#### SUMMARY

[0006] Embodiments herein concern compositions, methods and uses of chimeric dengue virus constructs. In some embodiments, a composition can include chimeric dengue virus constructs having an attenuated dengue virus backbone with structural genes from at least one other dengue virus serotype. Other embodiments concern at least one live, attenuated virus in combination with one or more chimeric dengue viruses. Other embodiments can include a composition of chimeric dengue viruses having a modified DEN-2 backbone (e.g. PDK-53 as a starting backbone in P1 (passage-1) and passage variability (after passage and growth in vitro on a permissive cell line) as indicated for P2, P3, ... P8... P10 etc.) and one or more structural components of DEN-1, DEN-2, DEN-3 or DEN-4. In other embodiments, an immunogenic composition is generated where when introduced to a subject, the composition produces an immune response to one or more dengue viruses in the subject. Therefore, constructs contemplated herein can be generated and passaged in vitro, and each of the passages provides an attenuated dengue virus contemplated of use in a pharmaceutically acceptable vaccine composition. In certain embodiments a live, attenuated virus can be a live, attenuated dengue-2 virus alone or in combination with one or more chimeric dengue viruses.

[0007] In certain examples, chimeric dengue virus constructs of dengue virus serotypes can include passage 7 (P7) live, attenuated viruses or chimeric viruses having nucleic acid sequences identified by SEQ ID NOS: 1, 4, 7 and 10 or polypeptide sequences indicated by SEQ ID NOS: 2, 3, 5, 6, 8, 9, 11 and 12. It is contemplated herein that any of the passages for any of the live, attenuated viruses described herein can be used in an immunogenic composition to induce immune responses to the represented dengue viruses (e.g. serotypes 1-4). In accordance with these embodiments, an immunogenic composition that includes a P-8 isolated live, attenuated virus can be administered to a subject to induce an immunogenic response against one or more dengue virus serotypes depending on the construct selected. In addition, a live, attenuated virus can be combined with one or more of these chimeric viruses. This is contemplated for each of the live, attenuated viruses isolated/produced in each subsequent cell passages (e.g. African Green Monkey Vero cell production, hereinafter: Vero cells). It is contemplated herein that any cell line (e.g. GMP-produced cell bank. FDA or EMA-approved) capable of producing dengue viruses is of use to passage any of the viral constructs at a manufacturing scale or as appropriate contemplated herein for subsequent use in a vaccine or immunogenic composition against Dengue virus.

**[0008]** In other embodiments, compositions contemplated herein can be combined with other immunogenic compositions against other Flaviviruses such as West Nile virus, Japanese encephalitis or any other flavivirus chimeric construct and/or live, attenuated virus. In certain embodiments, a single composition can be used against multiple flaviviruses.

**[0009]** In certain embodiments, an immunogenic composition of the present invention can include chimeric dengue viruses against one or more of DEN-1, DEN-2, DEN-3 and/or DEN-4, alone or in combination with a live, attenuated dengue virus composition.

**[0010]** In other embodiments, a construct can include a construct having adaptive mutations in the structural or non-structural regions of the virus that increase growth or production without affecting attenuation or safety of the virus when introduced to a subject. In certain embodiments, any of the contemplated chimeric dengue virus constructs can include a live, attenuated DEN-2 virus having specific mutations used as a backbone where the live attenuated DEN-2 PDK virus further includes structural proteins of one or more of prM (premembrane) and E (envelope) structural proteins of the other dengue virus serotypes. In addition, a DEN-2 backbone can include additional mutations in order to increase production of or enhance the immune response to a predetermine composition in a subject upon administration (e.g. chimeric Dengue virus 2/1, 2/3 or 2/4).

[0011] In some embodiments, structural protein genes can include prM and E genes of DEN-1, DEN-2, DEN-3 or DEN-4 on a DEN-2 backbone having one or two mutations that are part of a live, attenuated dengue virus. For example, a dengue construct, in certain embodiments can include those constructs termed DENVax-1-A, DENVax-2-F, DEN-Vax-3-F, and DENVax-4-F (see Example section) where the DEN-2 backbone has one or more mutations (e.g. not found in the P1 or other previous passaged virus or PDK-53) from the DEN-2 live, attenuated virus previously demonstrated to be safe and effective to induce an immune response. The DEN-2 live, attenuated virus of the instant application is an improved version of the originally used DEN-2 live, attenuated virus. A chimeric construct of the instant invention can include a modified attenuated DEN-2 PDK-53 backbone, having one or more structural proteins of the second dengue virus serotype wherein the structural proteins can include additional mutations to increase an immunogenic response to the chimeric construct. In some embodiments, certain mutations acquired by attenuated DEN-2 PDK-53 can produce a conservative amino acid change or not in a constructs different from the P1 construct which can result in desirable traits for production etc.

**[0012]** In other embodiments, a live, attenuated DEN-2 genome can be used to generate constructs of dengue virus serotype 1 (DEN-1) and dengue virus serotype 3 (DEN-3), dengue virus serotype 4 (DEN-4) where one or more structural protein genes of the DEN-2 viral genome can be replaced by one or more structural protein genes of DEN-1, DEN-3 or DEN-4, respectively. In some embodiments, a structural protein can be the C, prM or E protein of a second

dengue virus. In certain embodiments, structural protein genes include the prM and E genes of DEN-1, DEN-3 or DEN-4. These hybrid viruses express the surface antigens of DEN-1. DEN-3 or DEN-4 while retaining the attenuation phenotypes of the parent attenuated DEN-2.

**[0013]** Constructs disclosed herein can include chimeric constructs of DEN-4, DEN-2, DEN-1, and DEN-3 expressing surface antigens of DEN-1, DEN-3 and DEN-4 using attenuated DEN-2 virus as a backbone.

[0014] In certain embodiments, compositions of the instant invention can include a composition that comprises a single chimeric dengue virus construct disclosed herein and a pharmaceutically acceptable carrier or excipient. Alternatively, compositions of the instant invention can include a composition that comprises two or more, or three or more chimeric dengue virus constructs disclosed herein, and a pharmaceutically acceptable carrier or excipient. In accordance with these embodiments, a one or more dengue virus chimeric constructs contemplated herein can be combined with one or more, live attenuated dengue viruses. In certain embodiments, a live, attenuated virus can be a live, attenuated DEN-2 virus wherein additional mutations in the NCR, NS1 regions or other regions increase the immune response, increase viral growth or other improvement for an improved live, attenuated dengue virus.

**[0015]** In certain embodiments, the attenuation loci, nucleotide 5'NCR-57-T, NS1-53-Asp. and NS3-250-Val, of the DENV-2 vaccine have been previously determined, and all of these changes are shared by the common PDK-53 virus-specific genetic background of the four DENVax viruses. The genetic sequence of the three attenuation loci as well as the previously established in vitro and in vivo attenuation phenotypes of these vaccine candidates were carefully monitored for the cGMP-manufactured DENVax seeds. This report describes strategies used to generate master virus seeds (MVS) as well as their genetic and phenotypic characterization of use in the manufacture of dengue virus vaccine compositions. These MVS can be used for manufacture of clinical materials and ultimately commercial vaccine supplies.

#### BRIEF DESCRIPTION OF THE DRAWINGS

**[0016]** The following drawings form part of the present specification and are included to further demonstrate certain embodiments. Some embodiments may be better understood by reference to one or more of these drawings alone or in combination with the detailed description of specific embodiments presented.

**[0017]** FIG. 1 represents an exemplary chart reflecting an exemplary chimeric construct of the instant invention, DEN-2/DEN-4 compared to previously generated constructs and wild type dengue viruses.

**[0018]** FIG. **2** represents an exemplary histogram plot comparing various responses using a live, attenuated DEN-2 backbone (with additional mutations) and a second dengue virus serotype as structural components substituted for the dengue-2 structural components (e.g. DENVax-1 MVS). This plot illustrates plaque sizes of the DENVax MVS. Wild-type Dengue viruses and previously published research-grade vaccine candidate viruses were included for control and comparison. This plot illustrates improved production of the dengue virus constructs compared to control dengue virus chimeric constructs. **[0019]** FIG. **3** represents an exemplary histogram plot that represents temperature sensitivities of DENVax MVS (Master Virus Seed). Wild type dengue viruses and previously published research-grade vaccine candidate viruses were included for comparison with the MVS grade.

**[0020]** FIG. **4** represents an exemplary histogram plot that represents viral growth of DENVax MVS in C6/36 cells compared to controls. Wild-type dengue viruses and research-grade vaccine candidate viruses were included for comparison with the DENVax MVS.

**[0021]** FIGS. **5A-5**C represent exemplary plots of neurovirulence in newborn mice. FIG. **5A** shows pooled results of several experiments summarizing the neurovirulence of wt DENV-2 16681 virus in CDC-ICR (n=72) and Taconic-ICR (n=32) newborn mice challenged ic with  $10^4$  pfu of the virus. FIG. **5**B shows neurovirulence of DENVax MVS tested in Taconic-ICR mice with a dose of  $10^4$  pfu. FIG. **5**C shows neurovirulence of DENVax MVS tested in Taconic-ICR mice with a dose of  $10^3$  pfu. The numbers of animals tested per group in one experiment (n=16) or two pooled experiments (n=31 or 32) are indicated.

**[0022]** FIG. 6 represents an exemplary histogram illustrating plaque size of the DENVax MVS, WVS, and BVS. Mean plaque diameters $\pm$ SD (error bars) of the virus plaques in Vero or LLC-MK<sub>2</sub> cells under agarose overlay measured on day 9 pi. Wild type DENVs and previously published research-grade vaccine candidate viruses were included for control and comparison.

**[0023]** FIG. **7** represents an exemplary histogram plot illustrating growth of DENVax MSV, WVS, and BVS in C6/36 cells at two incubation temperatures to verify their retention of this in vitro attenuation marker after large scale manufacturing.

**[0024]** FIG. 8 represents an exemplary histogram plotting restricted growth of DENVax MVS, WVS, and BVS in C6/36 cells. Mean titers±SD (error bars) of the viruses replicated in C6/36 cells 7 days pi. The wt Dengue viruses and previously published research-grade vaccine candidate viruses were included for comparison.

**[0025]** FIGS. **9A-9**B represent exemplary graphs of data of neurovirulence of DENVax MVS in newborn ICR mice. FIG. **9**A shows IC inoculations of the virus at dose of 10<sup>4</sup> PFU. FIG. **9**B shows IC inoculation of the virus at dose of 10<sup>3</sup> PFU.

**[0026]** FIG. **10** represents an exemplary chart comparing new live, attenuated viruses to previously generated live, attenuated dengue viruses.

#### DEFINITIONS

**[0027]** As used herein, "a" or "an" may mean one or more than one of an item.

**[0028]** As used herein the specification, "subject" or "subjects" may include, but are not limited to, mammals such as humans or mammals, domesticated or wild, for example dogs, cats, other household pets (e. g. hamster, guinea pig, mouse, rat), ferrets, rabbits, pigs, horses, cattle, prairie dogs, wild rodents, or zoo animals.

**[0029]** As used herein, the terms "virus chimera," "chimeric virus," "flavivirus chimera" and "chimeric flavivirus" can mean a construct comprising a portion of the nucleotide sequence of a dengue-2 virus and further nucleotide sequence that is not from dengue-2 virus or is from a different flavivirus. A "dengue chimera" comprises at least two different dengue virus serotypes but not a different

flavivirus. Thus, examples of other dengue viruses or flaviviruses include, but are not limited to, sequences from dengue-1 virus, dengue-3 virus, dengue-4 virus, West Nile virus, Japanese encephalitis virus, St. Louis encephalitis virus, tick-borne encephalitis virus, yellow fever virus and any combination thereof.

**[0030]** As used herein. "nucleic acid chimera" can mean a construct of the invention comprising nucleic acid comprising a portion of the nucleotide sequence of a dengue-2 virus and further nucleotide sequence that is not of the same origin as the nucleotide sequence of the dengue-2 virus. Correspondingly, any chimeric flavivirus or flavivirus chimera disclosed herein can be recognized as an example of a nucleic acid chimera.

**[0031]** As used herein, "a live, attenuated virus" can mean a wild-type virus, mutated or selected for traits of use in vaccine or other immunogenic compositions wherein some traits can include reduced virulence, safety, efficacy or improved growth etc.

#### DESCRIPTION

**[0032]** In the following sections, various exemplary compositions and methods are described in order to detail various embodiments. It will be obvious to one skilled in the art that practicing the various embodiments does not require the employment of all or even some of the specific details outlined herein, but rather that concentrations, times and other specific details may be modified through routine experimentation. In some cases, well-known methods or components have not been included in the description.

**[0033]** In accordance with embodiments of the present invention, there may be employed conventional molecular biology, protein chemistry, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Animal Cell Culture, R. I. Freshney, ed., 1986).

**[0034]** Embodiments herein concern compositions, methods and uses for inducing immune responses against one or more dengue virus serotypes in a subject, individually or simultaneously. In accordance with these embodiments, attenuated dengue viruses and nucleic acid chimeras are generated and used in vaccine compositions disclosed herein. Some embodiments concern modified or mutated dengue constructs or chimeras. Other embodiments concern introducing mutations to modify the amino acid sequences of structural proteins of dengue viruses wherein the mutation increase immunogenicity to the virus.

**[0035]** Live, attenuated dengue viruses of all four serotypes have been developed by passaging wild-type viruses in cell culture. These are some of the most promising live, attenuated vaccine candidates for immunization against flavivirus and in particular dengue virus infection and/or disease. These vaccine candidates have been designated by a combination of their dengue serotype, the cell line through which they were passaged and the number of times they were passaged. Thus, a dengue serotype 1 wild-type virus passaged in PDK cells 13 times is designated as DEN-1 PDK-13 virus. Other vaccine candidates are DEN-2 PDK-53, DEN-3 PGMK-30/FRhL-3 (e.g. thirty passages in primary green monkey kidney cells, followed by three passages in fetal rhesus lung cells and DEN-4 PDK-48). These four candidate vaccine viruses were derived by tissue culture passage of wild-type parental DEN-1 16007, DEN-2 16681, DEN-3 16562 and DEN-4 1036 viruses, respectively.

[0036] In certain embodiments, live, attenuated dengue-2 PDK-53 vaccine virus contained a mixture of viruses, with the population containing varying nucleotide differences. After genetic characterization of the attenuating mutations, certain attenuating characteristics were outlined and engineered into a cDNA infectious clone. RNA was transcribed from this infectious clone and introduced into Vero cells as a passage 1 of the newly characterized and derived PDK-53-Vero-DEN-2-P 1 virus (see for example, Table 1). This attenuated virus was created for each DEN serotype, but for DEN-1, DEN-3 and DEN-4, the prM and E genes were engineered into 3 separate cDNA infectious clones, thus generating four separate PDK-53-Vero viruses (termed herein as: PDK-53-Vero-DEN-2-P 1, PDK-53-Vero-DEN-1-P 1. PDK-53-Vero-DEN-3-P 1. and PDK-53-Vero-DEN-4-P 1). These attenuated vaccine virus strains were passaged in Vero cells 10 times (Table 1), and each separate lineage acquired mutations upon their adaptation to grow in Vero cells (Table 3). Certain embodiments here are directed to derivation and uses for these live, attenuated dengue viruses. [0037] Previous human clinical trials with these attenuated viruses have indicated that DEN-2 PDK-53 has the lowest infectious dose (50% minimal infectious dose of 5 plaque forming units or PFU) in humans, is strongly immunogenic, and produces no apparent safety concerns. The DEN-1 PDK-13, DEN-3 PGMK-30/FRhL-3 and DEN-4 PDK-48 vaccine virus candidates have higher 50% minimal infectious doses of 10,000, 3500, and 150 PFU, respectively, in humans. Although only one immunization with monovalent DEN-2 PDK-53 virus or DEN-4 PDK-48 virus was required to achieve 100% seroconversion in human subjects, a booster was needed to achieve the same seroconversion rate for DEN-1 PDK-13 and DEN-3 PGMK-30/FRhL-3 viruses, which have the two highest infectious doses for humans.

**[0038]** DEN-2 PDK-53 virus vaccine candidate, also abbreviated PDK-53, has several measurable biological markers associated with attenuation, including temperature sensitivity, small plaque size, decreased replication in mosquito C6136 cell culture, decreased replication in intact mosquitoes, loss of neurovirulence for suckling mice and decreased incidence of viremia in monkeys. Clinical trials of the candidate PDK-53 vaccine have demonstrated its safety and immunogenicity in humans. Furthermore, the PDK-53 vaccine induces dengue virus-specific T-cell memory responses in human vaccine recipients. Some embodiments herein describe an improvement on the DEN-2 PDK-53 used in chimeric constructs disclosed herein.

**[0039]** Immunogenic flavivirus chimeras having a dengue-2 virus backbone and at least one structural protein of another dengue virus serotype can be used for preparing the dengue virus chimeras and methods for producing the dengue virus chimeras are described. The immunogenic dengue virus chimeras are provided, alone or in combination, in a pharmaceutically acceptable carrier as immunogenic compositions to minimize, inhibit, or immunize individuals against infection by one or more serotypes, such as dengue virus serotypes DEN-1, DEN-2, DEN-3 and DEN-4, alone or in combination. When combined, the immunogenic dengue virus chimeras may be used as multivalent vaccines (e.g. bi-, tri- and tetravalent) to confer simultaneous protection against infection by more than one species or strain of flavivirus. In certain embodiments, the dengue virus chimeras are combined in an immunogenic composition useful as a bivalent, trivalent or tetravalent vaccine against the known dengue virus serotypes or confer immunity to other pathogenic flaviviruses by including nucleic acids encoding one or more proteins from a different flavivirus.

[0040] In some embodiments, avirulent, immunogenic dengue virus chimeras provided herein contain the nonstructural protein genes of the attenuated dengue-2 virus (e.g. PDK-53), or the equivalent thereof, and one or more of the structural protein genes or immunogenic portions thereof of the flavivirus against which immunogenicity is to be induced in a subject. For example, some embodiments concern a chimera having attenuated dengue-2 virus PDK-53 genome as the viral backbone, and one or more structural protein genes encoding capsid, premembrane/membrane, or envelope of the PDK-53 genome, or combinations thereof, replaced with one or more corresponding structural protein genes from DEN-1. DEN-3 or DEN-4 or other flavivirus to be protected against, such as a different flavivirus or a different dengue virus serotype. In accordance with these embodiments, a nucleic acid chimera disclosed herein can have functional properties of the attenuated dengue-2 virus and is avirulent, but expresses antigenic epitopes of the structural gene products of DEN-1. DEN-3 or DEN-4 in addition to other flaviviruses and is immunogenic (e.g. induces an immune response to the gene products in a subject). Then, these DNA constructs are used to transcribe RNA from an infectious clone, this RNA is introduced into Vero cells again producing a new progeny virus at P1. These new progeny viruses are distinguishable from PDK-53. (See e.g. P1-P10).

**[0041]** In another embodiment, a nucleic acid chimera can be a nucleic acid chimera having, but not limited to, a first nucleotide sequence encoding nonstructural proteins from an attenuated dengue-2 virus, and a second nucleotide sequence encoding a structural protein from dengue-4 virus alone or in combination with another flavivirus. In other embodiments, the attenuated dengue-2 virus can be vaccine strain PDK-53 having one or more mutated amino acids (see Examples). These additional mutations confer desirable traits of use as live, attenuated dengue-2 or as chimeric constructs described herein. Some embodiments include structural proteins of one or more of C, prM or E protein of a second dengue virus.

**[0042]** Other aspects include that chimeric viruses can include nucleotide and amino acid substitutions, deletions or insertions for example, in the control PDK-53 dengue-2 genome to reduce interference with immunogenicity responses to a targeted dengue virus serotype. These modifications can be made in structural and nonstructural proteins alone or in combination with the example modifications disclosed herein and can be generated by passaging the attenuated virus and obtaining an improved composition for inducing an immune response against one or more dengue virus serotypes.

**[0043]** Certain embodiments disclosed herein provide for method for making the chimeric viruses of this invention using recombinant techniques, by inserting the required substitutions into the appropriate backbone genome. Other embodiments herein concern passaging a confirmed (e.g. safe and effective) live, attenuated chimeric virus for additional improvements. In certain embodiments, a dengue-2 backbone used herein can include one or more mutations presented in Table 3. In other embodiments, a denguedengue chimera of the instant application can include one or more mutations as presented in Table 3. In yet other embodiments, a dengue-dengue chimera can include all of the mutations for each chimera as represented in Table 3 for Den-2/Den-1, Den-2/Den-3 or Den-2/Den-4. Pharmaceutical compositions that include a live, attenuated virus represented by the constructs of Table 3 are contemplated. For example, mono-, di-, tri- or tetravalent compositions are contemplated of use herein using chimeras and live, attenuated dengue-2 viruses as presented in Table 3.

**[0044]** In certain embodiments, a live, attenuated DEN-2 variant contemplated herein can be formulated into a pharmaceutical composition wherein the pharmaceutical composition can be administered alone or in combination with dengue-dengue chimeras or dengue-flavivirus chimeras. In certain embodiments, a bi-, tri or tetravalent compositions can be administered in a single application or in multiple applications to a subject.

#### Flavivirus Chimeras

[0045] Dengue virus types 1-4 (DEN-1 to DEN-4) are mosquito-borne flavivirus pathogens. The flavivirus genome contains a 5'-noncoding region (5'-NC), followed by a capsid protein (C) encoding region, followed by a premembrane/membrane protein (prM) encoding region, followed by an envelope protein (E) encoding region, followed by the region encoding the nonstructural proteins (NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5) and finally a 3' noncoding region (3'NC). The viral structural proteins are C, prM and E, and the nonstructural proteins are NS1-NS5. The structural and nonstructural proteins are translated as a single polyprotein and processed by cellular and viral protesases. [0046] Flavivirus chimeras can be constructs formed by fusing non-structural protein genes from one type, or serotype, of dengue virus or virus species of the flaviviridae, with protein genes, for example, structural protein genes, from a different type, or serotype, of dengue virus or virus species of the flaviviridae. Alternatively, a flavivirus chimera of the invention is a construct formed by fusing nonstructural protein genes from one type, or serotype, of dengue virus or virus species of the flaviviridae, with further nucleotide sequences that direct the synthesis of polypeptides or proteins selected from other dengue virus serotypes or other viruses of the flaviviridae.

[0047] In other embodiments, avirulent, immunogenic flavivirus chimeras provided herein contain the nonstructural protein genes of the attenuated dengue-2 virus, or the equivalent thereof, and one or more of the structural protein genes, or antigenic portions thereof, of the flavivirus against which immunogenicity is to be conferred. Suitable flaviviruses include, but are not limited to those listed in Table 1. [0048] Other suitable dengue viruses for use in constructing the chimeras can be wild-type, virulent DEN-1 16007. DEN-2 16681. DEN-3 16562 and DEN-4 1036 and attenuated, vaccine-strain DEN-1 PDK-13, DEN-2 PDK-53, DEN-3 PMK-30/FRhL-3 and DEN-4 PDK-48. Genetic differences between the DEN-1, DEN-2, DEN-3 and DEN-4 wild type/attenuated virus pairs are contemplated along with changes in the amino acid sequences encoded by the viral genomes.

**[0049]** Sequence listings for DEN-2 PDK-53 correspond to the DEN-2 PDK-53-V variant, wherein genome nucleo-tide position 5270 is mutated from an A to a T and amino

acid position 1725 of the polyprotein or amino acid position 250 of the NS3 protein contains a valine residue. The DEN-2 PDK-53 variant without this nucleotide mutation, DEN-2 PDK-53-E, differs from PDK-53-V only in this one position. DEN-2 PDK-53-E has an A at nucleotide position 5270 and a glutamate at polyprotein amino acid position 1725, NS3 protein amino acid position 250. It is understood that embodiments herein include modified PDK 53 that include one or more passages in a separate host cell (e.g. Vero cells, see Table 1) where desirable traits of use in vaccine compositions contemplated herein are generated.

[0050] In certain embodiments, designations of the chimeras can be based on the DEN-2 virus-specific infectious clone modified backbones and structural genes (prM-E or C-prM-E) insert of other dengue viruses or other flaviviruses. DEN-2 for the dengue-2 backbone, followed by the strain from which the structural genes are inserted. One DEN-2 backbone variant is reflected in the next letter after the number designation. One particular DEN-2 backbone variant from which the chimera was constructed is indicated by the following letter placed after a hyphen, parent 16681 (P). PDK-53-E (E), or PDK-53-V (V); the last letter indicates the C-prM-E structural genes from the parental (P) strain or its vaccine derivative (V) or the prM-E structural genes from the parental (P) or its vaccine derivative (V1). For example; DEN-2/1-VP denotes the chimera comprising the attenuated DEN-2 PDK-53V backbone comprising a valine at NS3-250 and the C-prM-E genes from wild-type DEN-1 16007; DEN-2/1-VV denotes the DEN-2 PDK-53V backbone with the vaccine strain of dengue-1, DEN-1 PDK-13; DEN-2/1-VP1 denotes the DEN-2 PDK-53V backbone and the prM-E genes from wild-type DEN-1 16007; DEN-2/3-VP1 denotes the DEN-2 PDK-53V backbone and the prM-E genes from wild-type DEN-3 0.16562; DEN-2/4VP1 denotes the DEN-2 PDK-53V backbone and the prM-E genes from wild-type DEN-4 1036. Other chimeras disclosed herein are indicated by the same manner.

**[0051]** In one embodiment, chimeras disclosed herein contain attenuated dengue-2 virus PDK-53 genome as the viral backbone, in which the structural protein genes encoding C, prM and E proteins of the PDK-53 genome, or combinations thereof, can be replaced with the corresponding structural protein genes from dengue-1, dengue-3 or dengue-4 virus and optionally, another flavivirus to be protected against, such as a different flavivirus or a different dengue virus strain.

**[0052]** In the nonstructural protein regions, a Gly-to-Asp (wild type-to-PDK-53) mutation was discovered at nonstructural protein NS1-53 (genome nucleotide position 2579); a Leu-to-Phe (wild type-to-PDK-53) mutation was discovered at nonstructural protein NS2A-181 (genome nucleotide position 4018); a Glu-to-Val (wild type-to-PDK-53) mutation was discovered at nonstructural protein NS3-250 (genome nucleotide position 5270); and a Gly-to-Ala mutation (wild type-to-PDK-53) was discovered at nonstructural protein NS4A-75 (genome nucleotide position 6599). The live, attenuated DEN-2 virus of the instant invention further includes mutations as presented in any chimera or live, attenuated dengue-2 virus of Table 3.

**[0053]** PDK-53 virus strain has a mixed genotype at genome nucleotide 5270. A significant portion (approximately 29%) of the virus population encodes the non-mutated NS3-250-Glu that is present in the wild type DEN-2 16681 virus rather than the NS3-250-Val mutation. As both

genetic variants are avirulent, this mutation may not be necessary in an avirulent chimera.

[0054] Previously, it was discovered that avirulence of the attenuated PDK-53 virus strain can be attributed to mutations in the nucleotide sequence encoding nonstructural proteins and in the 5' noncoding region. For example, a single mutation at NS1-53, a double mutation at NS1-53 and at 5'NC-57, a double mutation at NS1-53 and at NS3-250 and a triple mutation at NS1-53, at 5'NC-57 and at NS3-250, result in attenuation of the DEN-2 virus. Therefore, the genome of any dengue-2 virus containing such non-conservative amino acid substitutions or nucleotide substitutions at these loci can be used as a base sequence for deriving the modified PDK-53 viruses disclosed herein. Another mutation in the stem of the stem/loop structure in the 5 noncoding region will provide additional avirulent phenotype stability, if desired. Mutations to this region disrupt potential secondary structures important for viral replication. A single mutation in this short (only 6 nucleotide residues in length) stem structure in both DEN and Venezuelan equine encephalitis viruses disrupts the formation of the hairpin structure. Further mutations in this stem structure decrease the possibility of reversion at this locus, while maintaining virus viability. [0055] Mutations disclosed herein can be achieved by any method known in the art including, but not limited to, naturally-occurring or selected clones having additional features once passaged in a cell line of interest (e.g. Vero cells). It is understood by those skilled in the art that the virulence screening assays, as described herein and as are well known in the art, can be used to distinguish between virulent and avirulent backbone structures.

#### Construction of Flavivirus Chimeras

**[0056]** Flavivirus chimeras described herein can be produced by splicing one or more of the structural protein genes of the flavivirus against which immunity is desired into a PDK-53 dengue virus genome backbone, or other methods known in the art, using recombinant engineering to remove the corresponding PDK-53 gene and replace it with a dengue-1, dengue-3 or dengue-4 virus gene or other gene known in the art.

**[0057]** Alternatively, using the sequences provided in the sequence listing, the nucleic acid molecules encoding the flavivirus proteins may be synthesized using known nucleic acid synthesis techniques and inserted into an appropriate vector. Avirulent, immunogenic virus is therefore produced using recombinant engineering techniques known to those skilled in the art.

**[0058]** A target gene can be inserted into the backbone that encodes a flavivirus structural protein of interest for DEN-1, DEN-3, DEN-4 or other flavivirus. A flavivirus gene to be inserted can be a gene encoding a C protein, a PrM protein and/or an E protein. The sequence inserted into the dengue-2 backbone can encode both PrM and E structural proteins. The sequence inserted into the dengue-2 backbone can encode all or one of C, prM and E structural proteins.

**[0059]** Suitable chimeric viruses or nucleic acid chimeras containing nucleotide sequences encoding structural proteins of other flaviviruses or dengue virus serotypes can be evaluated for usefulness as vaccines by screening them for the foregoing phenotypic markers of attenuation that indicate avirulence and by screening them for immunogenicity. Antigenicity and immunogenicity can be evaluated using in vitro or in vivo reactivity with flavivirus antibodies or

immunoreactive serum using routine screening procedures known to those skilled in the art.

#### Dengue Virus Vaccines

**[0060]** In certain embodiments, chimeric viruses and nucleic acid chimeras can provide live, attenuated viruses useful as immunogens or vaccines. Some embodiments include chimeras that exhibit high immunogenicity to dengue-4 virus while producing no dangerous pathogenic or lethal effects.

[0061] To reduce occurrence of DHF/DSS in subjects, a tetravalent vaccine is needed to provide simultaneous immunity for all four serotypes of the virus. A tetravalent vaccine is produced by combining a live, attenuated dengue-2 virus of the instant application with dengue-2/1, dengue-2/3, and dengue-2/4 chimeras described above in a suitable pharmaceutical carrier for administration as a multivalent vaccine. [0062] The chimeric viruses or nucleic acid chimeras of this invention can include structural genes of either wild-type or live, attenuated virus in a virulent or an attenuated DEN-2 virus backbone. For example, the chimera may express the structural protein genes of wild-type DEN-4 1036 virus, its candidate vaccine derivative in either DEN-2 backgrounds.

**[0063]** Viruses used in the chimeras described herein can be grown using techniques known in the art. Virus plaque titrations are then performed and plaques counted in order to assess the viability and phenotypic characteristics of the growing cultures. Wild type viruses can be passaged through cultured cell lines to derive attenuated candidate starting materials.

**[0064]** Chimeric infectious clones can be constructed from the various dengue serotype clones available. The cloning of virus-specific cDNA fragments can also be accomplished, if desired. The cDNA fragments containing the structural protein or nonstructural protein genes are amplified by reverse transcriptase-polymerase chain reaction (RT-PCR) from dengue virus RNA with various primers. Amplified fragments are cloned into the cleavage sites of other intermediate clones. Intermediate, chimeric dengue virus clones are then sequenced to verify the accuracy of the inserted dengue virus-specific cDNA.

**[0065]** Full genome-length chimeric plasmids constructed by inserting the structural protein and/or nonstructural protein gene region of dengue serotype viruses into vectors are obtainable using recombinant techniques well known to those skilled in the art.

#### Nucleotide and Amino Acid Analysis

**[0066]** The NS1-53 mutation in the DEN-2 PDK-53 vaccine virus is significant for the attenuated phenotype of this virus, because the NS1-53-Gly of the DEN-2 16681 virus is conserved in nearly all flaviviruses, including the tick-borne viruses, sequenced to date. DEN-4 vaccine virus can also contain an amino acid mutation in the NS1 protein at position 253. This locus, which is a Gln-to-His mutation in DEN-4 PDK-48 vaccine virus, is Gin in all four wild serotypes of dengue virus. This Gin residue is unique to the dengue viruses within the flavivirus genus. The NS1 protein is a glycoprotein that is secreted from flavivirus-infected cells. It is present on the surface of the infected cell and NS1-specific antibodies are present in the serum of virus-infected individuals. Protection of animals immunized with

NS1 protein or passively with NS1-specific antibody has been reported. The NS1 protein appears to participate in early viral RNA replication.

**[0067]** The mutations that occurred in the NS2A, NS2B, NS4A, and NS4B proteins of the DEN-1, -2, -3 and -4 attenuated strains are conservative in nature. The NS4A-75 and NS4A-95 mutations of DEN-2 and DEN-4 vaccine viruses, respectively, occurred at sites of amino acid conservation among dengue viruses, but not among flaviviruses in general.

[0068] The flaviviral NS3 protein possesses at least two recognized functions: the viral proteinase and RNA helicase/ NTPase. The 698-aa long (DEN-2 virus) NS3 protein contains an amino-terminal serine protease domain (NS3-51-His, -75-Asp, -135-Ser catalytic triad) that is followed by sequence motifs for RNA helicase/NTPase functions (NS3-196-GAGKT (SEQ ID NO:147), -284-DEAH, -459-GRIGR). None of the mutations in the NS3 proteins of DEN-1, DEN-2, or DEN-3 virus occurred within a recognized motif. The NS3-510 Tyr-to-Phe mutation in DEN-1 PDK-13 virus was conservative. Since the wild-type DEN-2, -3 and -4 viruses contain Phe at this position, it is unlikely that the Tyr-to-Phe mutation plays a role in the attenuation of DEN-1 virus. The NS3-182 Glu-to-Lys mutation in DEN-1 PDK-13 virus occurred at a position that is conserved as Asp or Glu in most mosquito-borne flaviviruses and it may play some role in attenuation. This mutation was located 15 amino acid residues upstream of the GAGKT helicase motif. As noted in previous reports, the NS3-250-Glu in DEN-2 16681 virus is conserved in all mosquitoborne flaviviruses except for yellow fever virus.

[0069] Nucleic acid probes selectively hybridize with nucleic acid molecules encoding the DEN-1, DEN-3 and DEN-4 viruses or complementary sequences thereof. By "selective" or "selectively" is meant a sequence which does not hybridize with other nucleic acids to prevent adequate detection of the dengue virus. Therefore, in the design of hybridizing nucleic acids, selectivity will depend upon the other components present in a sample. The hybridizing nucleic acid should have at least 70% complementarity with the segment of the nucleic acid to which it hybridizes. As used herein to describe nucleic acids, the term "selectively hybridizes" excludes the occasional randomly hybridizing nucleic acids, and thus, has the same meaning as "specifically hybridizing." The selectively hybridizing nucleic acid of this invention can have at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, and 99% complementarity with the segment of the sequence to which it hybridizes, preferably 85% or more.

**[0070]** Sequences, probes and primers which selectively hybridize to the encoding nucleic acid or the complementary, or opposite, strand of the nucleic acid are contemplated. Specific hybridization with nucleic acid can occur with minor modifications or substitutions in the nucleic acid, so long as functional species-specific hybridization capability is maintained. By "probe" is meant nucleic acid sequences that can be used as probes or primers for selective hybridization with complementary nucleic acid sequences for their detection or amplification, which probes can vary in length from about 5 to 100 nucleotides, or preferably from about 18-24 nucleotides.

**[0071]** if used as primers, the composition preferably includes at least two nucleic acid molecules which hybridize

to different regions of the target molecule so as to amplify a desired region. Depending on the length of the probe or primer, the target region can range between 70% complementary bases and full complementarity and still hybridize under stringent conditions. For example, for the purpose of detecting the presence of the dengue virus, the degree of complementarity between the hybridizing nucleic acid (probe or primer) and the sequence to which it hybridizes is at least enough to distinguish hybridization with a nucleic acid from other organisms.

**[0072]** Nucleic acid sequences encoding the DEN-4, DEN-3 or DEN-1 virus (e.g. structural elements) can be inserted into a vector, such as a plasmid, and recombinantly expressed in a living organism (e.g. into a dengue-2 backbone) to produce recombinant dengue virus peptides and/or polypeptides and/or viruses.

#### Nucleic Acid Detection Methods

**[0073]** A rapid genetic test that is diagnostic for each of the vaccine viruses described herein is provided by the current invention. This embodiment of the invention enhances analyses of viruses isolated from the serum of vaccinated humans who developed a viremia, as well as enhancing characterization of viremia in nonhuman primates immunized with the candidate vaccine viruses.

**[0074]** These sequences include a diagnostic TaqMan probe that serves to report the detection of the cDNA amplicon amplified from the viral genomic RNA template by using a reverse-transciptase/polymerase chain reaction (RT/PCR), as well as the forward and reverse amplimers that are designed to amplify the cDNA amplicon, as described below. In certain instances, one of the amplimers has been designed to contain a vaccine virus-specific mutation at the 3'-terminal end of the amplimer, which effectively makes the test even more specific for the vaccine strain because extension of the primer at the target site, and consequently amplification, will occur only if the viral RNA template contains that specific mutation.

**[0075]** Automated PCR-based nucleic acid sequence detection system can be used, or other known technology for nucleic acid detection. The TaqMan assay is a highly specific and sensitive assay that permits automated, real time visualization and quantitation of PCR-generated amplicons from a sample nucleic acid template. TaqMan can determine the presence or absence of a specific sequence. In this assay, a forward and a reverse primer are designed to anneal upstream and downstream of the target mutation site, respectively. A specific detector probe, which is designed to have a melting temperature of about 10.degree. C. higher than either of the amplimers and containing the vaccine virus-specific nucleotide mutation or its complement (depending on the strand of RT/PCR amplicon that is being detected), constitutes the third primer component of this assay.

**[0076]** A probe designed to specifically detect a mutated locus in one of the vaccine viral genomes will contain the vaccine-specific nucleotide in the middle of the probe. This probe will result in detectable fluorescence in the TaqMan assay if the viral RNA template is vaccine virus-specific. However, genomic RNA templates from wild-type DEN viruses will have decreased efficiency of probe hybridization because of the single nucleotide mismatch (in the case of the parental viruses DEN viruses) or possibly more than one mismatch (as may occur in other wild-type DEN viruses) and will not result in significant fluorescence. The DNA

polymerase is more likely to displace a mismatched probe from the RT/PCR amplicon template than to cleave the mismatched probe to release the reporter dye (TaqMan Allelic Discrimination assay, Applied Biosystems).

**[0077]** One strategy for diagnostic genetic testing makes use of molecular beacons. The molecular beacon strategy also utilizes primers for RT/PCR amplification of amplicons, and detection of a specific sequence within the amplicon by a probe containing reporter and quencher dyes at the probe termini. In this assay, the probe forms a stem-loop structure. The molecular beacons assay employs quencher and reporter dyes that differ from those used in the TaqMan assay.

#### Pharmaceutical Compositions

[0078] Embodiments herein provide for administration of compositions to subjects in a biologically compatible form suitable for pharmaceutical administration in vivo. By "biologically compatible form suitable for administration in vivo" is meant a form of the active agent (e.g. pharmaceutical chemical, protein, gene, of the embodiments) to be administered in which any toxic effects are outweighed by the therapeutic effects of the active agent. Administration of a therapeutically active amount of the therapeutic compositions is defined as an amount effective, at dosages and for periods of time necessary to achieve the desired result. For example, a therapeutically active amount of a compound may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of antibody to elicit a desired response in the individual. Dosage regima may be adjusted to provide the optimum therapeutic response.

[0079] In one embodiment, the compound (e.g. pharmaceutical chemical, protein, peptide etc. of the embodiments) may be administered in a convenient manner, for example, subcutaneous, intravenous, by oral administration, inhalation, intradermal, transdermal application, intravaginal application, topical application, intranasal or rectal administration. Depending on the route of administration, the active compound may be contained in a protective buffer (e.g. FTA, F127/trehalose/albumin). In one embodiment, a composition may be orally administered. In another embodiment, the composition may be administered intravenously. In one embodiment, the composition may be administered intranasally, such as inhalation. In yet another embodiment, the composition may be administered intradermally using a needle-free system (e.g. Pharmajet®) or other intradermal administration system.

**[0080]** A composition may be administered to a subject in an appropriate carrier or diluent, co-administered with enzyme inhibitors or in an appropriate carrier such as liposomes. The term "pharmaceutically acceptable carrier" as used herein is intended to include diluents such as saline and aqueous buffer solutions. It may be necessary to coat the compound with, or co-administer the compound with, a material to prevent its inactivation. The active agent may also be administered parenterally, or intraperitoneally. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations may contain a preservative to prevent the growth of microorganisms or other stabilizing formulation (e.g. FTA).

**[0081]** Pharmaceutical compositions suitable for injectable use may be administered by means known in the art. For example, sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion may be used. In all cases, the composition can be sterile and can be fluid to the extent that easy syringability exists. It might be stable under the conditions of manufacture and storage and may be preserved against the contaminating action of microorganisms such as bacteria and fungi. The pharmaceutically acceptable carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of microorganisms can be achieved by heating, exposing the agent to detergent, irradiation or adding various antibacterial or antifungal agents.

**[0082]** Sterile injectable solutions can be prepared by incorporating active compound (e.g. a compound that induces an immune response to one or more dengue virus serotypes) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization.

**[0083]** Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms, such as the type of injectable solutions described above. It is contemplated that compositions are especially suitable for intramuscular, subcutaneous, intradermal, intranasal and intraperitoneal administration. A particular ratio may be sought such as a 1:1, 1:2 or other ratio (e.g. PFUs of a given dengue virus serotype)

**[0084]** The active therapeutic agents may be formulated within a mixture predetermined ratios. Single dose or multiple doses can also be administered on an appropriate schedule for a given situation (e.g. prior to travel, outbreak of dengue fever).

**[0085]** In another embodiment, nasal solutions or sprays, aerosols or inhalants may be used to deliver the compound of interest. Additional formulations that are suitable for other modes of administration include suppositories and pessaries.

**[0086]** Certain formulations can include excipients, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate and the like.

[0087] A pharmaceutical composition may be prepared with carriers that protect active ingredients against rapid elimination from the body, such as time-release formulations or coatings. Such carriers include controlled release formulations, such as, but not limited to, microencapsulated delivery systems, and biodegradable, biocompatible polymers, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, polyorthoesters, polylactic acid and others are known. [0088] Pharmaceutical compositions are administered in an amount, and with a frequency, that is effective to inhibit or alleviate side effects of a transplant and/or to reduce or prevent rejection. The precise dosage and duration of treatment may be determined empirically using known testing protocols or by testing the compositions in model systems known in the art and extrapolating therefrom. Dosages may also vary with the severity of the condition. A pharmaceutical composition is generally formulated and administered to exert a therapeutically useful effect while minimizing undesirable side effects. In general, dose ranges from about  $10^2$  to  $10^6$  PFU can be administered initially and optionally, followed by a second administration within 30 days or up to 180 days later, as needed. In certain embodiments, a subject can receive dual administration of a mono, bi-, tri or tetravalent composition disclosed herein wherein the composition is a single composition mixture or has predetermined compositions of different dengue virus serotypes. In some embodiments, a DEN2/4 chimera can be present in higher concentrations than other dengue virus serotypes such as a live, attenuated dengue-1.

**[0089]** It will be apparent that, for any particular subject, specific dosage regimens may be adjusted over time according to the individual need.

**[0090]** In one embodiment, a composition disclosed herein can be administered to a subject subcutaneously or intradermally.

**[0091]** The pharmaceutical compositions containing live, attenuated dengue viruses may be administered to individuals, particularly humans, for example by subcutaneously, intramuscularly, intranasally, orally, topically, transdermally, parenterally, gastrointestinally, transbronchially and transal-veolarly. Topical administration is accomplished via a topically applied cream, gel, rinse, etc. containing therapeutically effective amounts of inhibitors of serine proteases. Transdermal administration is accomplished by application of a cream, rinse, gel, etc. capable of allowing the inhibitors of serine proteases to penetrate the skin and enter the blood stream. In addition, osmotic pumps may be used for administration. The necessary dosage will vary with the particular condition being treated, method of administration and rate of clearance of the molecule from the body.

**[0092]** In certain embodiments of the methods of the present invention, the subject may be a mammal such as a human or a veterinary and/or a domesticated animal or livestock or wild animal.

#### Therapeutic Methods

**[0093]** In one embodiment of the present invention, methods provide for inducing an immune response to dengue virus serotype(s) using a mono, bi-, tri or tetravalent formulation of live, attenuated and/or chimeric viral constructs contemplated herein.

**[0094]** Embodiments of the present invention is further illustrated by the following non-limiting examples, which are not to be construed in any way as imposing limitations upon the scope thereof. On the contrary, it is to be clearly understood that resort may be had to various other embodiments, modifications, and equivalents thereof which, after reading the description herein, may suggest themselves to those skilled in the art without departing from the spirit of the present invention or the scope of the appended claims.

#### EXAMPLES

**[0095]** The following examples are included to demonstrate certain embodiments presented herein. It should be appreciated by those of skill in the art that the techniques disclosed in the Examples which follow represent techniques discovered to function well in the practices disclosed herein, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in particular embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope herein.

### Example 1

**[0096]** In some exemplary methods, compositions used to generate as referred to herein as "master virus seeds (MVS)" are disclosed. These compositions may be derived from one or more live, attenuated dengue viruses, such as DEN-1, DEN-2, DEN-3, and DEN-4. In certain methods, compositions may be derived from one or more live attenuated Dengue viruses that include but are not limited to, specific constructs disclosed herein referred to as DENVax-1, DEN-Vax-2, DENVax-3, and DENVax-4. In other exemplary methods, strategies used to generate and characterize these compositions are provided. In yet other embodiments, tetravalent dengue virus formulations and genetic and phenotypic characterization of these formulations are provided.

#### Production and Analysis of Pre-Raster DENVax Viruses

[0097] Certain procedures were performed to generate pre-master dengue virus seeds, such as serial amplification and purification of dengue viruses (e.g. DENVax). First, DENVax viruses were re-derived by transfection of viral RNA transcribed from the full-length recombinant DENVax cDNA into production-certified cells (e.g. Vero cells), resulting in P1 (passage 1) virus seed. The four P1 viruses from each of dengue-1 to dengue-4 were then amplified and plaque purified to obtain the candidate pre-master vaccine P7 seeds (see Table 1). Certain tests were performed to analyze passages of dengue viruses. For example, full-length genome sequencing demonstrated that all four of the P2 (passage 2) seed viruses were genetically identical to their homologous progenitor, research-derived, research-grade candidate vaccine virus. The original plaque phenotypes were also retained in the P2 viruses. Six plaque purified viruses (P3 A-F) were isolated for each serotype of dengue virus (e.g. DENVax1-4) from the P2 seeds, and each isolated plaque was directly plaque purified two more times. The third plaque purification (P5) of each virus was amplified twice (P6 A-F and P7 A-F) in Vero cells to produce the potential pre-master P7 DENVax seeds (Table 1).

TABLE 1

	Example of a cGMP Rederivation of DEI WCB-Vero Cells	tvax virases in
Passage	Seed Production/Purification	Characterizations
P1	Transfect WCB-Vero with transcribed viral RNAs	Plaque titrate
P2	Amplify P1 virus	Full genome sequence
Р3	Pick 6 plaques (A-F)/serotype from P2 plaque assay	Plaque purification
P4	Pick plaques A-F from P3 plaque assay	Plaque purification
P5	Pick plaques A-F from P4 plaque assay	Plaque purification
P6	Amplify P5 A-F plaques	Plaque titrate
Ρ7	Pre-master seeds: Amplify P6 A-F	Full genome sequence TaqMAMA, Plaque phenotypes
P8*	MVS: Amplify selected P7 virus seed	Full genetic and phenotypic characterization

TABLE	1-continued
-------	-------------

1	Example of a cGMP Rederivation o WCB-Vero Cell	
Passage	Seed Production/Purification	Characterizations
Р9	WVS: Amplify P8 Master Seed viruses	Full genome sequence, TagMAMA
P10	BVS: Amplify P9 Working Seed viruses	Full genome sequence, TaqMAMA

\*One optimal P7 seed (A, B, C, D, E, or F) was selected based on the genetic and plaque analysis to make P8 MVS

[0098] Some tests were further performed to characterize P7 DENVax seeds, such as analysis of genome sequences and plaque phenotypes of the P7 seeds, and comparison with P2 seeds (Table 2). Plaque phenotypes of the P7 viruses were generally similar to those of the P2 seeds. In some exemplary experiments, virus titers were monitored. Virus titers reached over 6.0 log pfu/ml for most of the P7 seeds, except for 5 viruses. Genome sequencing of more than 60 candidate vaccine virus seeds after 10 or more serial passages in Vero cells identified no reversion event at NS1-53 and NS3-250 of the three major attenuation determinants of the DENV-2 PDK-53 genetic vector, suggesting that these 2 loci are quite stable in candidate vaccine virus seeds. All sequence chromatograms of the 24 candidate strains generated from both forward and reverse sequencing for these two sites were homogenous without any minor nucleotide populations evident at the NS1-53 and NS3-250 genetic loci. In contrast to the NS1 and NS3 sites, different levels of reversions at the 5'NCR-57 attenuation locus were identified from multiple serially passaged research grade vaccine viruses, suggesting this locus might not be as stable as NS1 and NS3 after multiple passages in cell culture. Therefore, a sensitive mismatch amplification assay (TaqMAMA) was developed to accurately measure the reversion rate at the 5'NCR-57 locus by real-time RT-PCR. In some studies, the 5'NCR-57 reversion rates of all 24 of the P7 seeds were measured by the TaqMAMA. Depending on the concentration of the input viral RNA for each virus in the assay, the sensitivity limit of the TaqMAMA ranged between 0.01% and 0.07% reversion, which is much more sensitive than the 10-30% reversion sensitivity limit detectable by consensus genome sequence analysis. The resulting data illustrates that 15 of the 24 P7 viruses had minimal or undetectable reversion (<0.07%), one virus (DENVax-3-D) had almost 100% reversion, and 8 viruses (1 DENVax-1, 1 DENVax-2, 2 DENVax-3, and 4 DENVax-4) had partial reversion ranging from 0.08% to 12.85% (Table 2). Full-length genome sequencing was conducted for 16 of the 24 P7 viruses with low levels of 5'NCR57 reversion as measured by TaqMAMA. All the sequenced viruses maintained the other two DENVax attenuation determinants (NS1-53. NS3-250), and all had acquired additional mutations that were not present in the original, engineered recombinant cDNA clones (Table 2). In one exemplary target vaccine composition, DENVax-1-A, DEN-Vax-2-F, DENVax-3-F, and DENVax-4-F were selected as target pre-master seed for each serotype because their genotypes and plaque phenotypes most closely resembled those of the originally designed vaccine recombinants. The DEN-Vax-1-A, DENVax-2-F, and DENVax-4-F had two nonsynonymous mutations, and the DENVax-3-F had one. The evidence suggests these additional mutations observed in these 4 pre-master seeds do not cause safety concerns or immunogenicity alterations for the viruses. These pre-master seeds were further amplified to generate the MVS (master seed, designated as P7, Table 1).

[0099] Exemplary methods provided herein used purified in-vitro transcribed viral RNA from cloned cDNA plasmid as the pure source to transfect vaccine-certified Vero cells to generate vaccine virus. Serial plaque purifications and fullgenome sequence analyses were incorporated into the manufacturing procedures to ensure manufactured vaccine seeds with optimal purity and genetic stability. Six cloned viruses were prepared as potential pre-master seeds for each serotype of DENVax. Through genomic analysis, including TaqMAMA and complete genomic sequencing, as well as characterization of viral plaque phenotypes, pre-master seeds were chosen to advance to master virus seeds production for each serotype (serotypes 1-4). The selected premaster seeds had undetectable reversions (<0.01% or <0.07%) at the 5'NCR-57 locus, with 1 or 2 amino acid substitutions in their genomes, and retained the small plaque phenotypes previously observed.

TABLE 2

		Charac	terization	ns of pre-	master (P7) seeds
Virus	Clone <sup>a</sup>	TaqMAMA $^b$	Log <sub>10</sub> pfu/ml	Plaque <sup>c</sup>	Mutations identified in genome <sup><math>d</math></sup>
DENVax-1	Α	**	6.85	P2	NS2A-116 I-L, NS2B-92 E-D, one silent
	В	*	6.93	P2	$\mathrm{nd}^e$
	С	*	6.93	D	nd
	D	**	7.02	D	C-67 K-A; one silent
	Е	0.57%	7.28	P2	nd
	F	**	7.18	P2	E473 T-M; one silent
DENVax-2	Α	0.03%	6.33	P2	NS1-341 K-N
	В	*	6.33	P2	E-305 K-T, two silent
	С	*	5.84	L	NS4A-18 T-A, four silent
	D	0.08%	6.20	P2	NS2B-99 I-L, one 3'NCR
	Е	0.03%	6.31	P2	prM-52 K-E, NS5-412 I-V, two silent
	F	**	6.15	P2	prM-52 K-E, NS5-412 I-V
DENVax-3	Α	*	6.00	P2	NS5-200 K-N, one silent, one 3'NCR
	В	0.05%	6.27	P2	NS2A-33 I-T NS2A-59 MT
	С	0.30%	6.25	P2	nd
	D	100.00%	6.27	P2	nd
	Е	0.31%	6.00	P2	nd
	F	**	6.30	P2	E-223 TS, one silent

		Charac	terizatio	ns of pre-	master (P7) seeds					
Virus	Clone <sup>a</sup>	TaqMAMA <sup>b</sup>	Log <sub>10</sub> pfu/ml	Plaque <sup>c</sup>	Mutations identified in genome <sup>d</sup>					
DENVax-4	А	0.47%	5.60	P2	E323 K-R/K, NS2B-21 L-F/L, NS2B-39 T-S, one silent					
	В	*	5.65	D	NS2A-126 A-V; NS4A-5 N-D; NS5-383 K-R, one silent					
	С	4.50%	5.90	P2	nd					
	D	12.85%	5.97	D	nd					
	Е	0.52%	6.85	S	prM-85 E-D, NS2B-45 T-A, NS5-320 M-T, NS5-551 E-G, two silent					
	F	0.02%	6.93	$\mathbf{S}$	NS2A-66 D-G, NS4A-21 A-V, four silent					

TABLE 2-continued

<sup>a</sup>Cloned viruses (by serial plaque purifications) selected for further development of MVS are designated bold.

<sup>b</sup>\*: Reversion rate < 0.07% (detection limit). \*\*: Reversion rate < 0.01% (detection limit).

• Reversion rate < 0.67 // (detection mind). • Reversion rate < 0.67 // (detection mind).</p>
• Plaque phenotypes: P2: similar P2 virus; L = larger than P2 virus, D = similar size, but appear somewhat different in clearness of the plaques; S = smaller than P2.
• Substitutions differing from the engineered DENVax cDNA clones. Amino acid mutations are listed with residue position of the virus protein and the changes (wt-mutation). Total number of silent mutations in structural and non-structural genes of each seed is listed. Mutations at non-coding region (NCR) are also noted.
• \*nd = Not done. These clones had higher 5'NCR-57 reversion rates (by TaqMAMA) than other clones, so were excluded from further sequence analyzis.

from further sequence analysis.

#### Example 2

[0100] In some exemplary methods, compositions of master virus seeds, working virus seeds and bulk virus seeds as well as their genetic and phenotypic characterization are described. These compositions are provided for manufacture of clinical materials and ultimately commercial vaccine supplies. Serial plaque purifications and full-genome sequence analyses were incorporated into the manufacturing process to ensure compositions of vaccine seeds with optimal safety and genetic stability for manufacture of clinical trial materials.

Production and Manufacturing Quality Controls for MVS, WVS, and BVS

[0101] In some studies, MVS of the 4 DENVax were produced by amplifying the pre-master P7 seed in certified Vero cells. In other studies. MVS were used to make large amount of WVS in cell factories. Further, the BVS stocks of DENVax were amplified from the WVS and were formulated into tetravalent drug product mixtures to be used for human clinic trials. Quality controls for product release were performed in some exemplary methods, including, but not limited to, testing all of the MVS, WVS, and BVS for identity, infectious titer, sterility, mycoplasma, and in vitro and in vivo adventitious agents. All seeds passed the virus identity test using serotype-specific RT-PCR assays, which showed positive amplification corresponding to its serotype and negative for heterologous serotypes (data not shown). No detectable mycoplasma or adventitious agents were detected in the MVS, WVS, or BVS stocks.

Genetic Analysis of the MVS, WVS, and BVS

[0102] In certain exemplary methods, after generation of MVS from the selected pre-MVS (P7) strains selected above were produced and the respective viral RNA was sequenced again. Full-length genome sequencing revealed that the MVS for DENVax-1 was identical to its pre-master seed, while the WVS and subsequent BVS acquired 2 additional substitutions at E-483 and NS4B-108 (see Tables 2 and 3). The Ala substitution at E-483 represented part of the genotype in the MVS, but became the dominant genotype in BVS. DENVax-2 and DENVax-3 were identical to their respective pre-master seeds (Table 2 and 3). The DENVax-2 MVS was identical to its pre-master seed, and the WVS and BVS had 2 additional mutations at NS4A-36 and NS4B-111. Both mutations were partial in WVS and were the major genotype in the BVS. The MVS of DENVax-3 was again identical to the pre-master seed, but the WVS and BVS contained an additional aa substitution at NS4A-23. The DENVax-4 MVS acquired an additional amino acid mutation, at locus NS2A-99 (from Lys to Lys/Arg mixed genotype) during production of the MVS (Table 3). Its WVS and BVS retained the NS2A-99 Lys/Arg mixed genotype, and the BVS had an extra NS4B-2384 Scr/Phe mixed genotype. Consensus sequence results also confirmed that MVS, WVS as well as BV retained the three genetic determinants of attenuation at the 5'NCR-57, NS1-53, and NS3-250 loci. Analysis of the least stable attenuating locus by TaqMAMA demonstrated that the 5'NCR-57 reversion rate between <0.7% to and 0.13% among MVS, 50.07% among WVS, and between <0.07 and 0.21% among BVS. A 3% reversion at the 5'NCR-57 locus was considered the maximum permissible rate for acceptance of a vaccine lot (Table 3).

TA	DT	$\mathbf{D}$	2
ТA	BL	Æ	- 5

Nucleotide and amino acid substitutions in DENVax seeds							
DENVax	Nucleotides	Amino Acids	Pre-master	MVS <sup>a</sup>	WVS <sup>a</sup>	BVS <sup>a</sup>	
DENVax-1	2384 G-C	E-483 Gly-Ala	-	-	Gly/Ala	Ala	
	3823 A-C	NS2A-116 Ile-Leu	Leu	Leu	Leu	Leu	
	4407 A-T	NS2B-92 Glu-Asp	Asp	Asp	Asp	Asp	
	7148 C-T	NS4B-108 Thr-Ile	-	-	Ile	Ile	
	7311 A-G	silent	G	G	G	G	
	TaqMAMA	5'NCR-57 reversion % <sup>b</sup>		-	-	-	

TABLE 3-continued

	Nucleotide and amino acid substitutions in DENVax seeds									
DENVax	Nucleotides	Amino Acids	Pre-master	$MVS^{a}$	WVS <sup>a</sup>	$BVS^a$				
DENVax-2	592 A-G	prM-52 Lys-Glu	Glu	Glu	Glu	Glu				
	6481 G-C	NS4A-36 Ala-Pro	-	-	Ala/Pro	Pro				
	7156 C-T	NS4B-111 Leu-Phe	-	-	Leu/Phe	Phe				
	8803 A-G	NS5-412 Ile-Val	Val	Val	Val	Val				
	TaqMAMA	5'NCR-57 reversion % <sup>b</sup>		-	0.07%	0.21%				
DENVax-3	1603 A-T	E-223 Thr-Ser	Ser	Ser	Ser	Ser				
	6436 G-A	NS4A-23 Asp-Asn	-	-	Asn	Asn				
	7620 A-G	silent	G	G	G	G				
	TaqMAMA	5'NCR-57 reversion % <sup>b</sup>		-	-	-				
DENVax-4	225 A-T	silent	Т	Т	Т	Т				
	3674 A-G	NS2A-66 Asp-Gly	Gly	Gly	Gly	Gly				
	3773 A-A/G	NS2A-99 Lys-Lys/Arg	-	Lys/Arg	Lys/Arg	Lys/Arg				
	5391 C-T	silent	Т	Τ ¯	Τ	T				
	6437 C-T	NS4A-21 Ala-Val	Val	Val	Val	Val				
	7026 T-C	silent	T/C	T/C	T/C	T/C				
	7538 C-C/T	NS4B-238 Ser-Ser/Phe	-	-	Ser/Phe	Ser/Phe				
	9750 A-C	silent	С	С	С	С				
	TaqMAMA	5'NCR-57 reversion $\%^b$	-	0.13%	-	-				

<sup>a</sup>Bold: Changes started at MVS stocks.

b"--" indicates reversion rate <0.01% (detection limit), "-" indicates reversion rate <0.07% (detection limit)

[0103] Full-genome sequence analysis revealed that an additional amino acid mutation developed in the DENVax-4 MVS, while the other three DENVax MVS lots retained the consensus genome sequence of their pre-master seeds. Overall, from deriving of the P1 seeds to the pre-master (P7) weeds, only 1 or 2 non-synonymous mutations occurred in a given seed. From P1 to MVS (P8) seeds, 2 to 7 nucleotide substitutions were identified in any given DENVax seed and only 2 to 3 of these substitutions resulted in amino acid changes. Thus, minor changes occurred. RNA viruses are error-prone in their genome replication, so genetic substitutions in flavivirus genome during cell passages are not unexpected. None of the silent mutations in the MVS were within the 5' or 3'NCR that may affect virus replication. Only the change in prM-52 Lys-Glu of the DENVax-2, and the substitution in NS2A-66 Asp-Gly of DENVax-4 are not conservative changes. The NS2A-66 mutation of the DEN-Vax-4 is in the nonstructural backbone part of the DENV-2 PDK-53. Although NS2A-66 locus is usually Asp among various strains of DENV-2, it is usually Gly for DENV-4. It is possible that the Asp to Gly change in the DENVax-4 is relevant for fitness of the DENVax-4 in Vero cells. The DENVax-2 prM-52 mutation resides in the C-terminal portion of the prM that is cleaved out from the mature virus particles. In some exemplary methods, phenotypic characterization was performed to confirm that none of the mutations in the MVS seeds significantly altered the attenuation phenotypes of the vaccine.

**[0104]** The DENVax viruses demonstrated high genetic stability during the manufacturing process. The three defined DENV-2 PDK-53 attenuation loci located in 5'NCR, NS1-53, and NS3-250 remained stable in the consensus genome sequence upon serial passage of the DENVax from pre-Master strains to bulk vaccine preparations. The highly sensitive TaqMAMA of the 5'NCR-57 locus demonstrated minimal or undetectable reversion in the MVS, WVS (P9/Working), and BVS (Bulk Virus Seed for vaccines) of dengue virus serotypes. The 5'NCR-57 reversion rates of the DENVax BVS preparations (P10-equivalent) were significantly lower than the 5'NCR-57 reversion rates that evolved in research-grade vaccine candidates after 10-serial passages

in Vero cells (4-74% reversion). The strategy for large-scale manufacturing of the DENVax seeds provided herein resulted in a genetically stable vaccine seed which retained the attenuation markers in the candidate vaccine viruses.

#### Plaque Phenotype of DENVax MVS

**[0105]** In one exemplary method, plaque phenotypes of the DENVax MVS were compared with wild type Dengue viruses and their homologous research-grade chimeric viruses in Vero cells (FIG. 2). All of the MVS of DENVax-1, -2, and -3 produced plaques that were significantly smaller than their wild type homologs and very similar (within 0.4-mm differences) to their homologous research-grade viruses in Vero cells. DENVax-4 MVS was also significantly smaller than the wild type DENV-4, but was slightly larger (0.9 mm difference) than the original lab derived D2/4-V chimera.

**[0106]** FIG. **2** represents an exemplary histogram illustrating plaque sizes of the DENVax MVS in contrast with control wild type viruses and research-grade vaccine candidate viruses. Mean plaque diameters (mm)±SD (error bars) of the virus plaques in Vero cells under agarose overlay measured on day 9 pi. The wild type DEN viruses, represented by black bars, and previously published research-grade vaccine candidate viruses, represented by white bars, were included for control and comparison to the DENVax master vaccine seeds represented by grey bars.

#### Temperature Sensitivity of DENVax MVS

**[0107]** In another exemplary method, temperature sensitivity was tested in Vero cells for the DENVax MVS and compared with their homologous wild type and the original research-grade chimeric vaccine virus. The wild type (wt) DENV-3 16562 was not temperature sensitive. The wt dengue virus serotype 1 and dengue virus serotype-4 were moderately temperature sensitive at 39° C. (titers were approximately 1.0 log<sub>10</sub> pfu/ml lower at 39° C. than at 37° C., FIG. **3**). Wt Dengue virus serotype-2 16681 was the most temperature sensitive of the wt Dengue viruses tested, and resulted in a 100-fold titer drop at 39° C. DENVax-1, -2, and -3 were as temperature sensitive as their original homologous research-grade chimeric vaccine viruses (FIG. 2). Titers at 39° C. dropped between 2.0 and 3.0 log<sub>10</sub> pfu/ml for these DENVax strains. DENVax-4 also was temperature sensitive, demonstrating a 5-fold reduction in titer. However, the original research-grade D2/4-V demonstrated about a 10-fold reduction in titer. The final stabilized DENVax-4 MVS contained F127 (and other agents known to stabilize these formulations (FTA)), which was shown to enhance thermal stability of the Dengue viruses. The presence of the F127 in DENVax-4 MVS likely contributed to the less pronounced temperature sensitivity of the virus in the Vero culture assay. In a separate experiment, temperature sensitivity of an MSV-derived DENVax-4 strain in the absence of F127 was further evaluated. To remove the F127 from the strain, viral RNA was isolated from a DENVax-4 bulk virus preparation and was transfected into Vero cells. This DEN-Vax-4 virus appeared to be as temperature sensitive as the D2/4 V research strain (titer reduced 1.5  $\log_{10}$  pfu/ml) on day 3 pi in the absence of F127 (FIG. 3).

**[0108]** FIG. **3** illustrates an exemplary histogram illustrating temperature sensitivities of DENVax MVS. The wild type Dengue viruses and previously published researchgrade vaccine candidate viruses were included for comparison. The DENVax-4 MVS contains additional F-127 that can mask the temperature sensitivity results of the virus in this assay. A separate experiment analyzing a surrogate DEN-Vax-4 in the absence of F127 was also included. Mean titers±SD (error bars) of the viruses replicated in Vero cells at 37° C. or 39° C.

#### DENVax MVS Replication in Mosquito C636 Cells

**[0109]** In some exemplary methods, the DENVax MVS were grown in C6/36 cells to verify their retention of the in vitro attenuation phenotype, with the knowledge that the research-grade chimeric vaccine viruses retained the attenuation phenotype of the backbone DENV-2 PDK53 virus in these mosquito cells. Compared to the wt Dengue viruses, DENVax-1. DENVax-2 and DENVax-4 MVS showed significant growth reduction (at least  $3 \log_{10}$  pfu/ml reduction) in C6/36 cells on day 6 pi (FIG. 4). The DENVax-3 MSV also exhibited reduced growth compared to the wt DENV-3 16562, but the reduction was not as marked (1-2  $\log_{10}$  pfu/mL reduction). However, the C6/36 titers of the DEN-Vax-3 seed lots were similar (within 1  $\log_{10}$  pfu/ml difference) to the C6/36 titer of the original research-grade chimeric D2/3-V vaccine virus.

**[0110]** FIG. **4** illustrates an exemplary histogram plotting restricted growth of DENVax MVS (grey bars) in C6/36 cells in comparison with wt Dengue viruses (black bars) and research-grade vaccine viruses (white bars). Mean titers±SD (error bars) of the viruses replicated in C6/36 cells 6 days pi.

Virus Infection, Dissemination, and Transmission Rates in Whole Mosquitoes

**[0111]** In some exemplary methods, the infection and dissemination rates of the DENVax were compared with their parental wt Dengue viruses. In certain exemplary experiments, oral infection experiments were conducted in

Ae. aegypti mosquitoes. Infectious blood meals were backtitrated to measure the virus titers and only the experiments with similar virus titers in the blood meal (less than  $1 \log_{10}$ pfu/mi differences) between parental Dengue viruses and DENVax for each serotype were included for comparisons in Table 4. DENVax-1, DENVax-2, and research-grade D2 PDK-53-VV45R did not infect mosquitoes through oral feeding, which is significantly different (p<0.0001) from their parental viruses, DENV-1 16007 (44% infection) and DENV-2 16681 (43.3% infection). Because no mosquito was infected by DENVax-1 and -2, there was little to no dissemination concern for these two vaccine viruses. While DENVax-4 did infect some mosquitoes through oral feeding (2 out of 55), the infection rate was significantly lower (p<0.05) than its parental wt virus, DENV-4 1036 (8 out of 50). DENVax-3 did not infect any mosquitoes in two experiments with blood meal viral titers of 5.2±0.02 log<sub>10</sub> pfu/mil (Table 4), and in a separate experiment with blood meal viral titer of 6.0 log<sub>10</sub> pfu/ml, only 1 out of 30 mosquitoes became infected (data not shown). However, wt Dengue virus-3 16562 also had a very low infection rate (8%) at 5.2 logo pfu/ml, and the rate did not increase in a separate experiment with a higher blood meal viral titer at  $6.2 \log_{10} \text{ pfu/ml}$  (3%, 1 positive out of 30 mosquitoes, data not shown). Although the wild type (wt) Dengue virus-3 and Dengue virus-4 had significantly lower infection rates than the wt Dengue virus-1 and Dengue virus-2, the mean virus titers in the infected mosquitoes were similar (3.1 to 3.9 log10 pfu/ mosquito). In contrast, the DENVax-4 titers from the two infected mosquitoes were both minimal (0.7 log10 pfu/ mosquito), which was 1,000-fold lower than the titer from the mosquitoes infected by wt Dengue virus serotype-4 1036 (3.9 f 1.5 pfu/mosquito).

[0112] For those mosquitoes that were infected, dissemination out of the midgut could be assessed by determining whether virus was present in the legs. The four parental DENVs resulted in dissemination rates ranging between 36.3% and 62.5%, and their mean virus titers (in  $\log_{10}$  pfu) from the legs were between 0.9±0.3 and 2.2±0.7 (excluding negative samples). Neither of the two DENVax-4 infected mosquitoes resulted in virus dissemination to the legs (Table 4). While disseminated virus was detectable in the legs, none of the four wt Dengue viruses was detectable in saliva of orally infected mosquitoes, suggesting that oral feeding conditions may not be sufficiently sensitive to measure the transmission rate of these DENVs. Therefore, in other exemplary methods, highly stringent artificial mosquito infections by direct IT inoculation were subsequently performed (Table 4). Except for DENVax-4, all viruses (wt and DENVax) achieved 100% infection of the IT inoculated Ae. aegypti. The DENVax-4 inoculum had a slightly lower viral titer than the other three viral inocula, but it still successfully infected 70% of the inoculated mosquitoes. Despite the high body infection rates achieved by IT inoculation, all four DENVax viruses exhibited significantly lower (p<0.005) or non-detectable transmission rates (0-10%) compared to the wt Dengue viruses (43-87%, Table 4). The DENVax viruses demonstrated little to no infection and dissemination after oral feeding, and the highly stringent IT results affirmed the minimal transmission capacity of these DENVax viruses in Ae. aegypti.

TA	BI	E.	4

14

			Oral Fe	ed	IT inoculation					
Virus	Blood Meal <sup>a</sup> Mean ± SD	Infection <sup>b</sup> % (P/N)	Body Titer <sup>e</sup> Mean ± SD	$p^d$	Dissemination <sup>e</sup> % (P/N) <sup>f</sup>	Inoculum pfu/dose	Infection <sup>b</sup> % (P/N)	Body Titer <sup>c</sup> Mean ± SD	Saliva <sup>f</sup> % (P/N)	p <sup>d</sup>
DENV-1	6.6	44.0%	3.6 ± 1.5		36.3%	53.9	100%	4.7 ± 0.48	43%	
16007		(11/25)			(4/11)		(30/30)		(13/30)	
DENVax-1	6.9	0%	NA	< 0.0010	NA	67.8	100%	$3.4 \pm 0.39$	10%	< 0.005
		(0/30)					(30/30)		(3/30)	
DENV-2	6.6	43.3%	$3.1 \pm 1.5$		38.5%	67.8	100%	$5.2 \pm 0.34$	87%	
16681	<i>.</i>	(13/30)	37.1	-0.0001	(5/13)		(30/30)		(26/30)	-0.000
D2 PDK53-	6.4	0%	NA	< 0.0001	NA	56.4	100%	$4.0 \pm 0.20$	0%	< 0.0001
VV45R	6.4	(0/30) 0%	NT A	< 0.0001	NT A	50.7	(30/30) 100%	$3.5 \pm 0.27$	(0/30) 7%	<0.0001
DENVax-2	6.4	(0/30)	NA	<0.0001	NA	52.7	(30/30)	$3.5 \pm 0.27$	(2/30)	< 0.0001
DENV-3	5.2	8%	$3.8 \pm 0.2$		50%	34.0	(30/30)	$4.2 \pm 0.50$	(2/30) 67%	
16562	5.2	(2/25)	J.0 ± 0.2		(1/23)	54.0	(30/30)	4.2 ± 0.50	(20/30)	
DENVax-3	$5.2 \pm 0.02$	0%	NA	0.108	NA	37.3	100%	$3.3 \pm 0.36$	3%	< 0.0001
DERIVARS	5.2 = 0.02	(0/50)	1.11	01100	2.12.1	57.5	(30/30)	515 = 0150	(1/30)	0.000.
DENV-4	$5.8 \pm 0.5$	16%	3.9 ± 1.5		62.5%	69.4	100%	$5.2 \pm 0.45$	70%	
1036		(8/50)			(5/8)		(30/30)		(21/30)	
DENVax-4	$5.4 \pm 0.4$	3.6%	$0.7 \pm 0.0$	0.033	0%	11.8	70%	$1.1 \pm 0.46$	0%	< 0.0001
		(2/55)			(0/2)		(21/30)		(0/21)	

<sup>a</sup>Virus titers or Mean ± standard deviation if from more than 1 experiment in blood meal (log<sub>10</sub> pfu/ml) by back titration

<sup>b</sup>Rate of virus detected in mosquito bodies. P/N = positive/total mosquitoes

"Mean virus titers ± standard deviation (log10 pfu/mosquito) in mosquito body, only positive sample are included for calculation

dStatistic analysis of the differences between wt DENV and DENVax by Fisher Exact probability

Rate of virus detected in legs of the positively infected mosquitoes

<sup>f</sup>Rate of virus detected in saliva of the positively infected mosquitoes. Used to measure transmission efficiency

[0113] Vector competence is an important safety component for live-attenuated flavivirus vaccine viruses. Previously, the research-grade DENV-2 PDK-53-VV45R virus and wt derivatives were tested in Ae. aegypti, and found that the NS1-53-Asp attenuating mutation was the dominant determinant for impaired mosquito replication. The other two major attenuation loci of the DENV-2 PDK-53 vaccine, nucleotide 5'NCR-57-T and NS3-250-Vail, also exhibited some inhibiting effect on replication in mosquitoes-, thus providing additional, redundant restrictions for mosquito vector competence. Some exemplary methods described herein were used to test the mosquito oral and IT infection and replication for all four DENVax strains. DENVax-1, -2, and -3 did not infect any Ae. aegypti mosquitoes through oral infection (Table 4). The DENVax-4 infected only 3.6% of orally exposed mosquitoes, a level significantly lower than that of the wt DENV-4 with a replicative mean titer in the mosquito bodies lower than that of wt DENV-4 infected mosquitoes. Surprisingly, DENVax-4 was detected in the legs of the infected mosquitoes, suggesting that DENVax-4 was not able to disseminate from the mosquito midgut following oral infection. The infection rates for the DEN-Vax-1, -2, and -4 were all significantly less than their wt counterparts, but the difference was not significant between DENVax-3 and wt DENV-3 16562 due to the very low infection rates for both viruses. Compared to other wt strains of DENV assessed in Ae. aegypti collected from the same Mae Sot Province, Thailand, the parental wt Dengue virus strains used for engineering DENVax appeared to have lower infectious and dissemination rates by oral infection. The wt DENV-1 PUO359, DENV-2 PUO218, DENV-3 PaH881/88, and DENV-4 1288 used for engineering the Yellow Fever (YF) 17D vaccine-based ChimeriVax-DEN vaccines had infection rates ranging 47-77%. In contrast, the YF 17D vaccine cannot infect *Ae. aegypti*. Although the ChimeriVax strains contained the prM-E from these highly infectious wt DENV, the ChimeriVax retain the mosquito attenuation phenotype of their YF 17D replicative backbone. Results provided herein also indicated that the mosquito attenuation of DENV-2 PDK-53 backbone was maintained in the DENVax strains. In addition, using the wt Dengue virus strains with lower mosquito-infectivity in constructs included in compositions described herein provides an additional safety feature.

[0114] The oral infection results illustrate that the DEN-Vax had minimum mosquito infectivity and dissemination capacity. In addition, the more sensitive and stringent IT infection experiments were performed to further analyze the potential of DENVax to be transmitted by Ae. aegypti. The IT results demonstrated that all four DENVax viruses had non-detectable or minimal mosquito transmission potential compared to their wt counterparts. DENVax transmission could only theoretically occur if (1) vector feeds on a vaccine with a sufficient viremia titer to infect mosquito midgut, (2) the virus is capable of replicating in the midgut epithelium and able to subsequently disseminate out of the midgut, and (3) the disseminated virus can replicate in salivary gland and expectorate sufficient virus in saliva for transmission. The threshold of human viremia required to infect mosquitoes has not been established adequately, but human viremia can be  $10^{6}$ -10 mosquito infectious dose<sub>50</sub> (MID<sub>50</sub>)/ml after natural wt DENV infection. This MID<sub>50</sub> was based on direct IT inoculation of mosquitoes with diluted human plasma. Analysis of DENVax in nonhuman primates indicated that viremia titers following DENVax immunization were very low (less than 2.4 log10 pfu/ml) and lasted for 2-7 days. Given the low viremia levels and the low mosquito infection, dissemination, and transmission capacity of DENVax, it is unlikely that these vaccine viruses could be transmitted by mosquitoes in nature or cause viremia.

**[0115]** Therefore, it is proposed that any of the passages of any of the serotypes (P1-P10) could be used in a composition to generate a safe and effective vaccine against one, two, three or all four dengue virus serotypes.

#### Neurovirulence in Suckling Mice

[0116] The original research-grade vaccine viruses were highly attenuated for neurovirulence in newborn ICR mice maintained in-house at DVBD/CDC. All of these mice survived ic (intracerebral) challenge with 10<sup>4</sup> pfu of each vaccine virus. The wt Dengue virus serotype-2 16681 virus, on the other hand, resulted in 62.5%-100% mortality in these CDC-LCR mice in various experiments. In some experiments, commercial ICR mice obtained from Taconic Labs (Taconic-ICR) were used to study neurovirulence in newborn mice. It was observed that newborn Taconic-ICR mice were significantly more susceptible to Dengue virus serotype-2 infection than the previous CDC-LCR mice. FIG. 5A summarizes the neurovirulence of wt Dengue virus serotype-2 16681 in CDC-ICR colony and Taconic-ICR, newborn mice challenged ic with  $10^4$  pfu of the virus. The Taconic-ICR mice (100% mortality in 32 mice, average survival time of 8.3±0.5 days) were more susceptible to ic Dengue virus serotype-2 16681 challenge than the previous CDC-ICR mice (91% fatalities in 72 mice, average survival time of  $14.6 \pm 2.3$  days).

[0117] In other exemplary methods, in order to evaluate neurovirulence of the DENVax MVS, the Taconic-ICR mice initially were challenged ie (intracerebrally) with a dose of approximately 10<sup>4</sup> pfu of wt Dengue virus serotype-2 16681, D2 PDK-53 VV45R, D2/3-V, or DENVax 1-4 virus in one (n=16) or two (n=31-32) experiments (FIG. 5B). At this dose, D2/3-V research grade virus, as well as DENVax-1, and DENVax-3 MVS exhibited fully attenuated neurovirulence phenotypes (no illness or mortality). As expected, wt Dengue virus serotype-2 was found to be "fatal", with average mouse survival time (AST) of 8.3±0.8 days. In these Dengue virus serotype-2-sensitive Taconic-ICR mice, the D2 PDK-53-VV45R research grade virus resulted in 81.3% mortality. The DENVax-2 MVS and DENVax-4 MVS were uniformly fatal in the Taconic-ICR, showing AST values of 9.8±1.7, 10.2±1.4, and 11.3±0.4 days, respectively.

[0118] In some exemplary methods, the neurovirulence of wt Dengue virus serotype-2 16681 virus was compared with that of D2 PDK-53 VV45R, DENVax-2 MVS and DEN-Vax-4 MVS, as well as D2/4-V research grade virus, at a 10-fold lower dose (10-pfu, FIG. 5C). The wt Dengue virus serotype-2 retained a uniformly fatal neurovirulent phenotype, with AST of 9.0±1.4 days, at this lower challenge dose. The other 4 viruses exhibited intermediate neurovirulence phenotypes, and the degree of neurovirulence was serotypespecific. The D2 PDK-53-VV45R virus and its DENVax-2 MVS cognate showed significant attenuation (32.3% survival with AST of 13.1±3.8 days and 31.2% survival with AST of 10.5±3.4 days, respectively). Both the DENVax-4 MVS and the research grade D2/4-V virus were highly attenuated for neurovirulence (81.3% survival with AST of 18.8±5.8 days and 100% survival, respectively). The results suggested that MVS of DENVax-1 and -3 exhibited complete attenuation of neurovirulence, while DENVax-2 and -4 MVS lots retained attenuation phenotypes that closely resembled their homologous research-grade virus vaccine candidates.

**[0119]** FIGS. **5**A-**5**C represent exemplary graphs illustrating neurovirulence in newborn mice tested with various compositions including wt Dengue virus serotype-2 and different attenuated Dengue viruses. Pooled results of numerous experiments summarizing the neurovirulence of wt Dengue virus serotype-2 16681 virus in CDC-ICR (n=72) and Taconic-ICR (n=32) newborn mice challenged ic with  $10^4$  pfu of the virus (A). Neurovirulence of DENVax MVS tested in Taconic-ICR mice with a dose of  $10^4$  pfu (B) or  $10^3$  pfu (C). The numbers of animals tested per group in one experiment (n=16) or two pooled experiments (n=31 or 32) are indicated.

#### Plaque Phenotype of WVS, and BVS

[0120] Certain studies were performed to compare plaque phenotypes of WVS and BVS with MVS, wt Dengue viruses and their homologous lab derived, research-grade chimeras in Vero cells (FIG. 6). Mean plaque sizes were calculated from 10 plaques for each vaccine virus, but from reduced numbers of wt DENV-1, -3, and -4. All of the MVS viruses of DENVax-1, -2, and -3 produced plaques that were significantly smaller than their wt homologs and very similar (within 0.4-mm differences) to their homologous researchgrade viruses in Vero cells. DENVax-4 MVS was also significantly smaller than the wt DENV-4, but was slightly (0.9 mm) larger than the original lab derived D2/4-V chimera. With the exception of the DENVax-2, all of the WVS and BVS of the DENVax-1, -3, -4 retained significantly smaller plaque sizes than those produced from their wt homologs. The DENVax-2 WVS and BVS produced plaques that were similar to the plaques of wt DENV-2 virus in Vero cells, but when tested in LLC-MK<sub>2</sub> cells all of the DEN-Vax-2 manufactured seeds produced plaques that were somewhat smaller than those of the wt DENV-2 (1.4 t 0.4)and similar to the lab derived D2 PDK-53-VV45R (1.0±0.3) (FIG. 6).

**[0121]** Evaluation of the phenotypic markers of viral attenuation, including small plaque phenotype, temperature sensitivity, reduced replication in mosquito cells, reduced infection/dissemination/transmission by mosquitoes, and reduced neurovirulence in newborn ICR mice, were assessed for the compositions of MVS stocks. Results indicated that all of the DENVax retained the expected attenuation phenotypes similar to the original research-grade vaccine viruses. Given the mutations responsible for attenuation are conserved in all MVS, WVS and BV, it can be expected the attenuated phenotypes to be retained in the material manufactured for human clinical testing.

**[0122]** FIG. 6 represents an exemplary histogram illustrating plaque size of the DENVax MVS, WVS, and BVS. Mean plaque diameters $\pm$ SD (error bars) of the virus plaques in Vero or LLC-MK<sub>2</sub> cells under agarose overlay measured on day 9 pi. The wt DENVs and previously published research-grade vaccine candidate viruses were included for control and comparison.

#### Virus Replication in Mosquito C6/36 Cells

**[0123]** Previous studies demonstrated that the researchgrade PDK-53-based chimeric vaccine viruses retained the attenuation phenotype of the backbone DENV-2 PDK53 virus in C6/36 cells. In some exemplary methods, the DENVax MSV, WVS, and BVS were grown in C6/36 cells to verify their retention of this in vitro attenuation marker after large scale manufacturing. Compared to the wt Dengue viruses, except for DENVax-3, the manufactured seeds showed marked growth reduction (at least 3  $\log_{10}$  PFU/ml reduction) in C6/36 cells on day 6 pi (FIG. 7). The DEN-Vax-3 seeds also exhibited reduced growth compared to the wt DENV-3 16562, but the reduction was not as marked (1-2  $\log_{10}$  PFU/ml reduction). However, the titers of the DEN-Vax-3 seed lots were similar (within 1  $\log_{10}$  PFU/ml difference) to the original research-grade chimeric D2/3-V vaccine virus.

**[0124]** FIG. 8 represents an exemplary histogram plotting restricted growth of DENVax MVS, WVS, and BVS in C6/36 cells. Mean titers $\pm$ SD (error bars) of the viruses replicated in C6/36 cells 7 days pi. The wt Dengue viruses and previously published research-grade vaccine candidate viruses were included for comparison.

#### Neurovirulence in Suckling Mice

[0125] Additional experiments were performed to analyze neurovirulence in newborn ICR mice. At an intracranial dose of 10<sup>4</sup> PFU, the survival rates for wt DENV-2 16681 and the D2 PDK-53-VV45R were 0% and 18.8%, respectively (FIG. 9A) in the ICR mice, but were about 20% for wt DENV-2 16681 and 100% for the D2 PDK-53-VV45R in the CDC ICR mice. In this study, DENVax-1 and DENVax-3 MVS were attenuated (100% survival) for the mice at a dose of 10<sup>4</sup> PFU, but the MVS of DENVax-2 and DENVax-4 caused 100% mortality at the dose of over  $10^4$  PFU (FIG. 5A). However, when tested at a dose of 10, PFU of virus, the DENVax-2 (31.3% survival) and DENVax-4 (81.3% survival) showed reduced neurovirulence relative to wt Dengue virus serotype-2 16681 (0% survival), and their survival rates were similar to those of the research-grade vaccine candidates D2 PKD-53-VV45R (32.3%) and D2/4-V (100%), respectively (FIG. 9B). Although, wt DENV-1, -3, or -4 were not included for comparison in this study, previous work demonstrated that wt DENV-1 16007 was attenuated in the CDC-ICR mice by the ic route, while both wt DENV-3 16562 and DENV-4 1036 were highly virulent (0% survival) for the CDC-ICR mice. It is likely that these 3 wt DENV would exhibit similar or greater virulence in the more susceptible Taconic ICR mice. Therefore, inclusion of these wt Dengue viruses for comparison with their homologous DENVax MVSs was considered to be uninformative. This study indicated that all 4 DENVax MVSs and original laboratory derived candidate vaccine viruses exhibit comparable mouse attenuation phenotypes relative to the wt DENV-2 16681.

**[0126]** FIGS. **9A-9**B represent exemplary graphs of data of neurovirulence of DENVax MVS in newborn ICR mice. (A) IC inoculations of the virus at dose of  $10^4$  PFU. (B) IC inoculation of the virus at dose of  $10^3$  PFU

**[0127]** All seed lots of the DENVax were tested for the identity, sterility, and freedom from undesirable agents. Full-genome sequence analysis revealed that one extra amino acid mutation evolved in the DENVax-4 MVS, while the other 3 DENVax MVSs retained the consensus genome sequence of their pre-master seeds. In WVS lots, the DEN-Vax-3 acquired an extra amino acid mutation and the other 3 serotypes accumulated 2 extra amino acid substitutions, relative to their pre-master seeds. Genome sequences of all

the 4 BVS lots were identical to their WVS lots. Overall from the P2 seeds to the pre-master (P7) seeds, only 1 or 2 non-silent mutations occurred in a given seed. Between pre-master and BCS (P10) seeds, only 1 to 2 nucleotide substitutions were observed, all of which occurred in NS2A, 4A, or 4B, with the exception of single nucleotide change resulting in a conserved glycine and alanine at residue E-483. From P2 to BVS (P10) seeds, total 3 to 8 nucleotide substitutions were identified in any given DENVax seed, and only 2 to 4 of these substitutions resulted in amino acid changes. None of the silent mutations in the BVS were within the 5'- or 3'-NCR region which may affects virus replication. These results suggest that the DENVax viruses were genetically highly stable during manufacture. The three defined DENV-2 PDK-53 attenuation loci located in 5'NCR, NS1-53, and NS3-250 remained unchanged in the consensus genome sequence upon serial passage of the DENVax to generate BVS stocks. The highly sensitive TaqMAMA of the 5'-NCR-57 locus showed minimal or undetectable reversion in the MVS, WVS, and BVS of DENVax. The highest reversion rate of 0.21% was identified in the DENVax-2 BVS. The reversion rates of the P10equivalent BVS (<0.07% to 0.21%) were significantly lower than the reversion rates that evolved in other vaccine candidates after serial passages in Vero cells (4-74% reversion by P10). This suggests that this strategy for large scale manufacturing of the DENVax seeds is successful, regarding maintaining genetic stability and retention of attenuation markers in the candidate vaccine viruses.

[0128] Since MVS stocks disclosed herein will be used for future manufacturing of WVS and BVS lots, full panels of virus attenuation phenotype evaluations, including small plaque phenotype, temperature sensitivity, reduced replication in mosquito cells, reduced infection/dissemination/ transmission in whole mosquitoes, and reduced neurovirulence in newborn ICR mice, were conducted for all MVS or their equivalent surrogate stocks. For the WVS and BVS stocks, plaque size, infectivity in mosquito cells, were also performed to confirm their attenuations. Results indicated that all the MVS stocks of the 4 serotypes of DENVax retained the expected attenuation phenotypes, such as small plaques, reduced replication in C6/36 cells, and reduced mouse neurovirulence, similar to the original lab-derived vaccine viruses (FIGS. 6, 8, and 9). Except for the DENVax-4, all other 3 MVS stocks of DENVax were TS at 39° C. as shown in FIGS. 3 and 7.

**[0129]** For the WVS and BVS stocks, two attenuation phenotypes, small plaques and restricted replication in C6/36 cells, were analyzed and confirmed. Since there are very little genetic changes between the MVS and BVS, it was expected that they would retain the attenuation phenotypes as MVS. In addition to the experiments described in this report, safety and immunogenicity of the manufactured DENVax in Ag129 mice and nonhuman primate have been tested.

**[0130]** Exemplary methods are provided herein to demonstrate manufacture of DENVax MVS, WVS, and BVS stocks under cGMP. The BVS stocks were used to formulate the tetravalent DENVax currently in human clinical trial evaluations. A unique manufacture strategy to optimize the genetic stability and safety of the manufactured MVS was provided in some exemplary methods. Since the main attenuation loci of the DENVax have been well characterized previously and a highly sensitive and quantifiable SNP assay, TaqMAMA was developed to integrate genome sequence and the TaqMAMA to identify optimal pre-master seeds for making the MVS. The genetic and phenotypic characterizations of the MVS were fully analyzed to confirm that these viruses retained desirable attenuations for safety of the vaccine. This may be the only live, attenuated viral vaccine that can be efficiently analyzed for all the major attenuation genetic loci during manufacturing from premaster all the way to BVS stocks. Results provided herein exemplified the advantage of strategically designed liveattenuated vaccines in vaccine safety.

**[0131]** FIG. **10** represents an exemplary table comparing new live, attenuated viruses to previously generated live, attenuated dengue viruses. Mutations are indicated where different from a control virus (e.g. 16681), or other live, attenuated dengue-2 viruses.

#### Materials and Methods

#### Viruses and Cells

**[0132]** DENV-1 16007, DENV-2 16681, DENV-3 16562, and DENV-4 1034 served as wild-type (wt) DENV controls, and they were the parental genotype viruses for the four recombinant DENVax vaccine candidates. DENVax progenitor research-grade viruses, designated as D2/1-V, D2 PDK-53-VV45R, D2/3-V, and D2/4-V, were prepared and characterized previously. Vero (African green monkey kidney) cells used for making the master and working cell banks for vaccine production were originated from the American Type Culture Collection (ATCC) CCL81 cell line that has been characterized by the World Health Organization (WHO) for vaccine manufacture (WCB-Vero cells).

Derivation of Live Recombinant DENVax Viruses from cDNA Clones

[0133] To re-derive the candidate vaccine viruses under cGMP manufacturing conditions, the previously engineered DENV infectious cDNA clones, pD2-PDK-53-VV45R, pD2/1-V, pD2/4-V, and in vitro-ligated pD2/3-V containing the fill genome-length viral cDNAs were used to make fresh viral RNA transcripts by in vitro transcription as described previously. Briefly, XbaI-linearized DENV genomic cDNAs were treated with proteinase K, extracted with phenol/ chloroform and precipitated in ethanol to remove any residual proteins, and then suspended in RNase-free Tris-EDTA buffer prior to transcription. The in vitro transcription was conducted using the AmpliScribe T7 High Yield Transcription kit (Epicentre Technologies) following the manufacturer's recommended protocol. The RNA A-cap analog, m7G(5')ppp(5')A (New England BioLabs), was incorporated during the 2-hr transcription reaction to add the 5'-terminal A-cap to the RNA transcript. The samples were then treated with DNase I to digest the template cDNA, followed by low pH phenol/chloroform extraction and ethanol precipitation to remove residual DNA and proteins. The purified RNA transcripts, suspended in RNase-free water, were distributed in 20-µl aliquots and stored at -80° C. until ready for transfection of cells. The integrity and concentration of the RNA transcripts were analyzed by agarose gel electrophoresis. Each 20-µl aliquot was estimated to contain sufficient genome-length viral RNA to permit transfection of 0.4-1×  $10^7$  production-certified Vero cells by electroporation.

**[0134]** Transfection of each RNA transcript into WCB-Vero cells was performed in the cGMP facility at Shantha Biotechnics. DENVax RNA transcripts were thawed, mixed with 400 µl of the Vero cell suspension  $(1 \times 10^7 \text{ cells/ml})$ , and transferred to a pre-chilled sterile electroporation cuvette (4-mm gap) for electroporation by a Gene Pulser Xcell total system (BioRad Laboratories). Each sample was pulsed once at 250V/ $\infty$  Ohms/500 µF, incubated for 10-15 min at room temperature, transferred to a 75-cm<sup>2</sup> flask containing 30 ml of cell growth medium (MEM with 10% FBS), and incubated at 36° C.±PC, 5% CO<sub>2</sub> for 6 to 11 days. The culture medium was harvested, clarified by centrifugation, stabilized, and stored in small aliquots below -60° C. The viral titers of candidate vaccine stocks (termed P1 for passage level 1) resulting from transfection were determined by plaque titration assay in Vero cells and used for further propagation of the DENVax seeds.

#### Manufacture of DENVax Virus Seeds

[0135] P1 virus seeds were used to propagate DENVax pre-master, master, working, and bulk virus seed lots through a strategy designed to ensure the optimal genetic stability and safety of the manufactured lots. This strategy included three serial plaque purifications, as well as genetic analyses of viruses at various passage levels to select the optimal clonal virus population for continued seed production (Table 1). Briefly, the P1 seeds harvested from transfected cells were amplified once by infection of Vero cells at a MOI of 0.001 to generate the P2 seeds. Aliquots of the P2 seed stocks were evaluated by plaque morphology and complete viral genomic sequencing. The genetically confirmed P2 stocks were plated on Vero cell monolayers with overlay medium as described in the plaque titration section below to generate well-isolated plaques. After visualization with neutral red, six individual plaques from each of the 4 serotypes of vaccine viruses were isolated (plaque clones A to F) and mixed into 0.5 ml of culture medium (passage P3). Each of the six plaque suspensions was subjected to two additional rounds of plaque purification, resulting in twiceand thrice-plaque purified virus seeds at passages P4 and P5, respectively. The P5 viruses were amplified through two sequential Vero passages to produce P7 seed stocks.

[0136] Genetic analysis of the three major DENVax attenuation loci using spot sequencing and/or Taqman-based mismatched amplification mutation assay (TaqMAMA) as previously disclosed, and plaque phenotype analysis were conducted to screen all 24 P7 seeds. Seeds possessing appropriate initial characteristics were then further characterized by full genomic sequencing. As a result of these analyses, one of the 6 (clone A-F) P7 seeds of each DENVax serotype was selected to be the pre-master seed, based on the presence of the DENV-2 PDK-53 attenuating mutations, minimal genomic sequence alterations, and expected plaque phenotype. Each selected pre-master seed was expanded to master virus seed (MVS or P8) by a one-time passage of the virus at MOI of 0.001 in multiple 175 cm<sup>2</sup> flasks of Vero cells. Except for the DENVax-4 MVS, the master virus seeds were harvested at 8-10 days post infection (pi). The MVS stocks were harvested at 6-10 days post infection (pi), clarified by centrifugation, stabilized by the addition of sucrose/phosphate/glutamate solution (final concentration 7.5% sucrose, 3.4 mM potassium dihydrogen phosphate, 7.2 mM dipotassium hydrogen phosphate, 5.4 mM monosodium glutamate, respectively) and 0.95 to 1.90% FBS (final concentration). DENVax-4 MVS was prepared differently to optimize its yield. Briefly, multiple flasks of cells were infected with DENVax-4 pre-master seed at a MOI of 0.001

in the presence of 0.1% F-127T<sup>TM</sup>, poloxamer 407, (other EO-PO block copolymers have been assessed and may substitute here, see issued patent) that have been demonstrated to enhance DENV virus thermal stability. Infectious media was harvested days 6-10 pi, and stabilized with 17% FBS (final concentration), pooled, and frozen. All four DENVax MVS stocks were stored as 1-mi aliquots below  $-60^{\circ}$  C.

[0137] The DENVax working virus seeds (WVS) were prepared by one-time passage in Vero cell culture of the MVS at a MOI of 0.001. The procedures were similar to the production of MVS, except they were cultured in multiplelayer cell factories (6360 cm<sup>2</sup>). The WVS stocks were filtered through 10 µM and 0.45 µM filters, stabilized with the same stabilizers used for the MVS, aliquoted into 30 ml PETG bottles or 2.0 ml cryovials, and stored below-60° C. [0138] In certain methods, bulk virus seeds (BVS) were produced by infecting multiple cell factories (6360 cm<sup>2</sup> each) of confluent Vero cells with 90 mL of diluted WVS to attain a MOI of 0.001. A media used for dilution of the WVS inocula contained 0.1% F-127™ without scrum. After 1.5 hr adsorption, cells were washed 3 times with PBS, and 800 ml of serum-free DMEM medium was added to each cell factory, and the factories were incubated at 36(±1°) C in  $5(\pm 0.5)\%$  CO<sub>2</sub>. After incubation for four days, small aliquots of medium were collected for sterility testing. Viruses were harvested between day 5 and day 10 pi, and immediately clarified by filtration through a 0.45 um pore size filter, and 1 L of each clarified virus pool was stabilized by addition of 500 ml of 3×FTA buffer (final concentrations of 15% trehalose, 1.0% Pluronic® F-127<sup>™</sup> poloxamer 407, 0.1% human albumin USP in PBS, pi 7.4). The stabilized virus was distributed into I-L PETG bottles and stored frozen below -60° C. for subsequent pooling and quality control testing. All stabilized virus harvests with a virus titer above 10<sup>5</sup> PFU/ml and an acceptable level of residual DNA were rapidly thawed in a water bath at 32° C., then aseptically pooled and mixed. Each pooled monovalent BVS was distributed into labeled PETG containers and stored at below -60° C. until further use.

#### Manufacture Product Quality Controls

**[0139]** The MVS, WVS, and BVS seeds were tested for identity, sterility, and detectable adventitious agents. The identity of each vaccine stock was confirmed by RT-PCR with DENVax serotype-specific primers. The amplified cDNA fragments contained the E/NS1 chimeric junction site to permit identification of each of the four DENVax serotypes. Each seed was tested in all 4 serotype-specific RT-PCR reactions to confirm viral identity and freedom from cross contamination with heterologous DENVax serotypes. Sterility testing was performed in accordance with USP 71 (United States Pharmacopeia, section 71). Mycoplasma testing was performed.

**[0140]** The following in vitro and in vivo tests for viral contamination were all performed using unclarified, unstabilized DENVax harvests collected during manufacture of the seeds. Harvested infectious media were first neutralized with DENV rabbit polyclonal antiserum (Inviragen) at  $36\pm1^{\circ}$  C. for 1 hr to inactivate the DENV. For in vitro test, the neutralized seeds were inoculated into three indicator cells lines, MRC5. VERO and MA 104, in 25 cm<sup>2</sup> flasks. Echo virus (CPE control) or mumps virus (hemadsorption control) were used as positive CPE or hemadsorption con-

trol, respectively. All cells were monitored daily for CPE for a total of 14 days. At the end of 14 days, the culture supernatant was removed and replaced with 10 mL of a guinea pig red blood cell (RBC) solution (3 mL of 0.5% guinea pig RBC in phosphate buffered saline, made up to 10 mL with cell growth medium). The flasks were then incubated at  $5\pm3^{\circ}$  C. for 30 minutes followed by incubation at room temperature for 30 minutes. The monolayers were washed with PBS and observed under 10x magnification for the presence of any star-shaped clumps of RBCs for hemadsorption.

**[0141]** In vivo tests for adventitious agents were performed in suckling mice, post-weaning mice and guinea pigs. Suckling mice were inoculated with 0.1 ml or 0.01 ml (10 mice in each dose group) of the DENV-antiserum neutralized seed sample through intraperitoneal (ip) injection. Similarly, 10 post-weaning mice were each inoculated ip with 0.5 ml or 0.03 ml of the sample. Guinea pigs (5/group) were each inoculated ip with 5.0 mL. Suckling mice were observed daily for morbidity and mortality for a total of 14 days following inoculation. Post-weaning mice were observed for a total of 28 days, and guinea pigs were observed for a total of 42 days following inoculation. The test articles met the acceptance criterion if  $\geq$ 80% of the inoculated animals remained healthy throughout the observation period.

[0142] The in vivo testing for contaminants was also performed in embryonated chicken eggs and was conducted. For every sample, 10 embryonated hen eggs (9 days old) were each inoculated with 0.5 mL of the DENV antiserumneutralized sample into the allantoic fluid and incubated at 35° C. for 3 days. The allantoic fluids from these 10 eggs were harvested, pooled and passaged into the allantoic fluid of 10 fresh embryonated eggs (10-11 days old; 0.5 mL/egg) and incubated at 35° C. for a further 3 days. Similarly, for each sample, 10 embryonated eggs (6-7 days old) were each inoculated with 0.5 mL per egg (DENVax-2 monovalent BVS) or 0.25 mL per egg (DENVax-1, DENVax-3 and DENVax-4 BVS) by injection into the yolk sac and incubated at 35° C. for 9 days. The yolk sacs from these 10 eggs were harvested and pooled, and a 10% suspension was passaged into the yolk sacs of 10 fresh embryonated eggs (6-7 days old; 0.5 mL/egg) and incubated at 35° C. for a further 9 days. Eggs inoculated into the allantoic fluid (both initial and passage inoculations) were observed for viability after 3 days incubation. Both pools of allantoic fluid were tested for hemagglutination activity using chicken, guinea pig and human type O erythrocytes at 4° C. and 25° C. Eggs inoculated into the volk sack (both initial and passage inoculations) were observed for viability after 9 days of incubation.

#### Virus Plaque Assay and Immunofocus Assay

**[0143]** Virus titers were measured by plaque assay or immunofocus assay using Vero cells. Plaque assays were performed in double agarose overlays in six-well plates of confluent Vero cells as previously described, and they were also used to evaluate the plaque phenotypes of the DENVax seeds. For accurate comparison, plaque sizes of all viruses were measured and compared in the same experiment. After visualization with neutral red on day 9 pi, up to 10 well isolated plaques for each virus were measured for mean plaque size calculation. Fewer plaques were measured for wt

DENV-1, -3, and -4, whose larger plaque sizes often did not permit measurement of 10 well-separated plaques.

[0144] Because tetravalent DENVax contains all four DENV serotypes, a DENV serotype-specific immunofocus assay was developed to quantitate each DENVax component in the tetravalent formulations. Immunofocus assays of each individual DENVax MVS were compared with the plaque assays to ensure virus titration results were comparable between the two assays. The immunofocus assay was conducted in 6-well plates of confluent Vero cells infected with serially diluted viruses. Cells were overlayed with a balanced salt medium (BSS/YE-LAH medium) containing 0.7% high viscosity carboxymethyl cellulose (Sigma) and incubated for 7 days at 37° C. with 5% CO<sub>2</sub>. After removal of overlays, cell sheets were washed 3 times with PBS, fixed with cold 80% acetone for 30 min at -20'C, washed once with PBS, and blocked with a blocking buffer containing 2.5% (w/v) nonfat dry milk, 0.5% Triton X-100, 0.05% Tween-20 in PBS at 37° C. for 30 min. Blocked cells were incubated with diluted DENV serotype-specific MAbs, IF1 (DENV-1), 3H5 (DENV-2), 8A-1 (DENV-3), or 1H10 (DENV-4) in blocking buffer at 37° C. for 1 hour or 4° C. overnight, washed 3 times with washing buffer (0.05% Tween-20 in PBS), and incubated with alkaline phosphatase- or horse radish peroxidase (HRP)-conjugated affinity-pure goat anti-mouse IgG (Jackson Immuno Research Laboratories) at 37° C. for 45-60 min. Plates were washed 3 times before the appropriate substrate, 1-Step NBT/BCIP plus suppressor (Pierce) for alkaline phosphatase or Vector-VIP kit (Vector Labs) for HRP, was added for color development. Color development was stopped by rinsing with water when the foci were fully developed. Stained immunofoci were directly visualized and counted on a light box.

#### Genetic Sequence

**[0145]** Full length genomes of the MVS and WVS were sequenced (see below). Briefly, viral RNA was extracted from DENVax seeds by using the QIAamp viral RNA kit (Qiagen), and overlapping cDNA fragments covering the entire genome were amplified using the Titan One Tube RT-PCR kit (Roche Applied Science, Inc.). The amplified cDNA fragments were gel purified before sequencing with both forward and reverse primers using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems). Sequence reactions were cleaned using the BigDye XTerminator Purification kit (Applied Biosystems), and run on the 3130x1 Genetic analyzer (Applied Biosystems) at DVBD/CDC. The Lasergene SeqMan software (DNAStar, Inc) was used for genome analysis and comparison.

Taqman-Based Mismatch Amplification Nutation Assay (TaqMAMA)

**[0146]** TaqMAMA is a sensitive, quantitative single nucleotide polymorphism assay developed to permit finer assessment of the level of reversion at the 5'NC-57 locus of attenuation, and was further optimized for this study. Extracted viral RNA from MVS and WVS were analyzed by the TaqMAMA with both sets of primers/Tagman probe that are specific to wt or the vaccine 5'NC-57 region. The forward primers used to detect DENV-2 wt and vaccine sequences were D2-41-GC and D2-40-TT, respectively. The 3'-terminal nucleotide of each forward primer matched the specific 5'NCR-57 nucleotide for each virus, while the

nucleotide adjacent to the 3'-terminal nucleotide in each primer differed from the DENV-2 viral genomic sequence to enhance the mismatch effect. The reverse primer, CD-207, and the Taqman probe, CD-169F, for both wt and vaccine sets were identical. Sequences of the primers and probe as well as cycling conditions were described previously. The real time RT-PCR was performed with the iQ5 or CFX-95 system (BioRad), using a BioRad iScript RT-PCR (for probes) kit, in a 25-µl reaction containing 5 µl of viral RNA template, 0.4 uM of each primer, and 0.2 uM of the probe. Triplicate reactions for each wt- and vaccine-specific assay were conducted for each sample. Genome copy numbers were determined relative to a standard curve prepared for each viral genotype, where the RNA standards were transcripts derived from plasmids containing nt 1-2670 of each genotype-specific cDNA. In addition, the specificity of the assay was confirmed by testing each RNA standard with the heterologous genotype primer/probe sets to ensure minimum cross-reactivity in every experiment. The results were reported as the percentage of viral genomes showing reversion. Previously, due to higher cross-reactive backgrounds that limited the input RNA levels for this assay, the original detection sensitivity was about 0.1% reversion (discrimination power). Since then, the assay has been further optimized using improved real-time PCR equipment and reaction kits, and the cross-reactive background was decreased considerably at much high levels (7-8  $\log_{10}$  copies) of RNA template input. This optimization resulted in significant improvement of the detection sensitivity, down to 0.01-0.07% reversion.

Virus Replication in Mosquito C6/36 Cells and Temperature Sensitivity in Mammalian Vero Cells

**[0147]** The replication phenotypes of the four DENVax MVS stocks and wt DENV-1, -2, -3, and -4 viruses were evaluated in C6/36 mosquito cells (*Aedes albopictus*). C6/36 cells grown in 6-well plates were infected in duplicate with each virus at a MOI of 0.001 and incubated with 4 ml/well of DMEM medium containing 2% FBS in a 5% CO<sub>2</sub> incubator at 28° C. Small aliquots of the culture supernatant were collected for each virus on day 6 pi, mixed with an equal volume of medium containing 40% FBS, and stored at -80° C. until ready for virus plaque titration.

**[0148]** Temperature sensitivity was conducted by comparing viral growth at  $39^{\circ}$  C. versus growth at  $37^{\circ}$  C. at five days pi of Vero cells in 6-well plates. Cells were infected in quadruplicate with each virus at a MOI of 0.001 at  $37^{\circ}$  C. Following adsorption of virus, the infected cultures were incubated with 4 mL/well of DMEM medium containing 2% FBS in 2 separate 5% CO<sub>2</sub> incubators, one set (duplicate plates) at  $37^{\circ}$  C. and the other at  $39^{\circ}$  C. Aliquots (50-µl) of the culture supernatant were collected on day 5 pi, mixed with an equal volume of DMEM containing 40% of FBS, and stored at  $-80^{\circ}$  C. until ready for virus plaque titration. Incubator temperatures were calibrated with NIST-traceable factory-calibrated thermometers (-1 to 51° C.; ERTCO).

Mosquito Infection, Dissemination, and Transmission

**[0149]** Aedes aegypti mosquitoes used for the study were from a colony established in 2002 from a village near Mac Sot (16' N, 33' E), Thailand. After emerging from larvae, adult mosquitoes were maintained at 28° C. at a 16:8 (light:dark) photoperiod with 10% sucrose solution provided ad libitum. Five-to-seven day old female mosquitoes were used for infectious blood meal feeding or intrathoracic (IT) inoculations. Aliquots of freshly cultured DENVax and wt DENV were used immediately upon harvest (without any freeze-thaw cycle) to make virus blood meals as indicated below for oral infection. Remaining virus supernatants were supplemented with FBS to a final concentration of 20%, and aliquots were stored at  $-80^{\circ}$  C. for future virus plaque titration and IT inoculation experiments. The freshly prepared DENVax seeds for these experiments were amplified from the pre-master seeds in Vero cells, and were considered DENVax MVS equivalents.

[0150] Infectious blood meals were prepared by mixing fresh virus at a ratio of 1:1 with defribrinated chicken blood (Colorado Serum Company) on the day of oral infection. Mosquitoes were sugar-starved overnight and then offered the virus:blood mixture for 1 hour using a Hemotek membrane feeding system (Discovery Workshops). A 50-µl aliquot of the blood meal was retained at -80° C. for backtitration of virus doses. Fully-engorged females were sorted under cold anesthesia and placed into cartons with 10% sucrose solution provided ad libitum. Cartons were placed at 28° C. with a photoperiod of 16:8 h (light:dark). After 14 days, 25-30 mosquitoes from each virus group were anesthetized via exposure to triethylamine (Flynap®, Carolina Biological Supply Company) and one hind leg was removed and placed in 0.5 ml of DMEM with 10% FBS and 5% penicillin/streptomycin (100 U/ml and 100 µg/ml respectively). Saliva was collected by inserting the proboscis of the anesthetized mosquito into a capillary tube containing 2.5% FBS and 25% sucrose solution. Mosquitoes were allowed to salivate for at least 15 minutes and then capillary tubes and bodies were placed into separate tubes containing DMEM. Mosquito bodies, legs and saliva were stored at -80° C. until they were triturated and assayed for infectious virus. For IT inoculation, mosquitoes were cold-anesthetized and inoculated with approximately 50 pfu of virus in 0.34 µl inoculum. Inoculated mosquitoes were kept for 7 days in the same conditions as described above. Mosquitoes were then anesthetized, and their saliva and bodies were collected as described above. Samples were stored at -80° C. until further processing.

**[0151]** To process the samples for virus titration, body and leg samples were homogenized with copper coated BBs (Crossman Corporation, N.Y.) at 24 cycles/second for 4 min using a mixer mill, and then clarified by centrifuging at  $3,000\times g$  for 3 min. Saliva samples were centrifuged at  $3.000\times g$  for 3 minute, to expel fluid from capillary tubes. Ten-fold dilutions of the body and leg homogenates and saliva samples were tested for presence of infectious virus by plaque assay. Results from bodies, legs, and saliva were used for determining the infection, dissemination, and transmission rates, respectively.

#### Mouse Neurovirulence

**[0152]** Timed pregnant female ICR mice were obtained from Taconic Labs, and monitored several times each day to determine approximate birth times of pup litters. In a given

experiment, approximately 12-24 hours after birth, two litters of eight pups per virus (n=16), was challenged with  $10^3$  to  $10^4$  pfu of virus in 20 µl of diluent by intracranial (ic) inoculation using a 30-gauge needle. Animals were monitored at least 3 times daily for at least 32 days following challenge. At the first sign of illness (rough fur, hunched back, weight loss, abnormal movement, paralysis, or lethargy) animals were euthanized by lethal anesthetization with isoflurane gas, followed by cervical dislocation. The postinfection day of euthanasia represented the "time to illness/ morbidity" or "survival time" for the animal. The animal experiments were conducted following a DVBD/CDC IACUC-approved animal protocol.

#### Derivation of Master Seed Viruses

#### DENvax-1 Master Virus Seed (MVS)

**[0153]** Nucleotide sequence of the chimeric viral genome and deduced amino acid sequence of the translated protein are provided herein. Most of the prM-E gene (nt 457 to -2379, underlined) is wild-type (wt) DEN-1 16007 virus specific; the remaining genome is DEN-2 PDK-53 virus specific. All engineered substitutions differ from wt virus (D1 16007 or D2 16681), as well as extra mutations (changes from engineered cDNA clone) detected in the MVS are marked.

[0154] Substitutions Included in the Genome and Protein: [0155] Junction sites between D1 (prM-E) and D2 backbone:

- [0156] a. MluI (nt 451-456): engineered silent mutation, nt-453 A-to-G
- [0157] b. NgoMIV (nt 2380-2385): engineered mutations, nt-2381/2382 TG-to-CC (resulted in E-482 Valto-Ala change)

**[0158]** D2 PDK-53 virus backbone (change from wt D2 16681): all in bold

- **[0159]** a. 5'-noncoding region(NCR)-57 (nt-57 C-to-T): major attenuation locus (in red)
- [0160] b. NS1-53 Gly-to-Asp (nt-2579 G-to-A): major attenuation locus (in red)
- [0161] c. NS2A-181 Leu-to-Phe (nt-4018 C-to-T)
- [0162] d. NS3-250 Glu-to-Val (nt-5270 A-to-T): major attenuation locus (in red)
- [0163] e. nt-5547 (NS3 gene) T-to-C silent mutation
- [0164] f. NS4A-75 Gly-to-Ala (nt-6599 G-to-C)

[0165] nt-8571 C-to-T silent mutation of PDK-53 is not engineered in the vaccine virus

[0166] DEN-1 prM-E (change from wt D1 16007)

[0167] a. Engineered nt-1575 T-to-C silent mutation to remove native XbaI site

**[0168]** Additional substitutions found in vaccine seed (0.03% nt different from original clone)

- [0169] a. NS2A-116 Ile-to-Leu (nt-3823 A-to-C, in bold)
- [0170] b. NS2B-92 Glu-to-Asp (nt-4407 A-to-T, in bold)
- [0171] c. nt-7311 A-to-G silent mutation (in bold)

			SCR-	57-T, D2	PDR-53 att	enqation le	cus (wt D	2 16681: 0)	•
>5'-Noncoding 1	Region				ţ				>C
10 ACTOTEASTORACS	20 DTOGROSOAD	9.0 Aradacarta	40 POTETGAGOG	50 NGCTANCCT	CO CAATSUAGTIS	70 DTMAGAGETEN	08 Aqaftaate	90 Sagoagaetete	001 Astaste 8 N
110 ATAACQAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA			140 CRATATCOTO N M L	east the contract of the	160 AGAAACCSCH R N R V	170 POTOSACTOTO S T V		190 ACABAGAGATI F K R F	200 FCFCACT S L
210 POSATOCTSCADO G M 5 0 0	226 1400406400 8 6 9		230 TECATOOCOO F M A L			270 DOTBACARTIC L P I F		236 DAGAGART 6 I 6	360 Jaagaga X R
310 TSOSGAACAATTAA. W G T I R	320 Xaattaaaa	330	340 TTTTGAGAGG	350 OffCAGGAA	060 Agagattogai	379 AGGATGOTGAJ R M L N	088 ARDTTCAOS	ore. Decagaecan	400
410	426	430	> <b>prM</b> 400	45.0		ning of D1 478		uence 496	500
CAGOCATGATOATTI O M I I I I		CA&CACTOAT 7 V M	CGCETFCCAT A T B	TTAACOSSE L T T	<u>айососко</u> в <u>о</u> е в	<u>rococcaitate</u> P H M	<u>ataottado.</u> I V S I	<u>naociacioaaa</u> 8. q. e. r.	<u>g k</u> G K
				Engin	eered MiuI	splicing s	ite (nt-4	53 <b>A-to-G</b> s	silent}
510 <u>SICACITITOTIA.</u> SILEPK	520 NGACCIOIGE T S A	S36 AGERERAA G V R	540 <u>ATOTSCA020</u> N C T L	550 TCATTOCA 1 A M	550 TSOATTIOSO/ D L O	S70 NGAGTINIGII E L C F	683 <u>ACCACACEAS</u> 8 T G C	500 Ehocitadam T y p	а 66 02020017 Ч
610	620	630	C40	650	600	670	680	690	700
CECATCACTOAGEC: RITEA >M		<u>98(197116AC)</u> D V D C		<u>THOCACHGA</u> A T E	<u>CACATGSGTG/</u> I W V I	<u>acotategaas.</u> f x q t	<u>етвететеа.</u> С 6 ў		<u>daoceaci</u> 1 R R
710 <u>SAGACARACOPTCO</u> D R R S	720 <u>FICSCATTSC</u> 7 N L A	730 <u>1906262999</u> V B V			760 <u>AGAGCCGAAA(</u> R	*********		750 <u>303750AAAC/</u> X W E Q	896 <u>NSATACA</u> 1 0
810 810	820 820	830 830	840 970 N. Thurston		960 960	870 10075/15/15/15/15	880 280	890 2008-02020	2 - 2 009 TTADDAE
X V B T W	A 5 8	X 2 G	F T V X	A 2. F	1. A 8	.8 1 6 1	S I T	0 X 6	2 2
910	920	930	> <b>£</b> 940	950	940	970	980	990	1090
FICATITICCICATI FILL I. I.	L V P	<u>ocatotales</u> P_S_M_A	. 2 R C	<u>06766844</u> 1 V Q I	C N R I	BACITICATORI D P V E	G L S	G A T V	1996-199 V V D
0101 SASETTATESTITA 1 E J V V	1020 CATEGAAGET 1 G S C	1030 900300080 90030080 90030080		1050 <u>8aca&amp;&amp;Cca</u> N 8 P	1000 <u>ACNOTOGACA</u> T E D I	1670 <u>7168&amp;070176</u> E E E	0800 0800 08000 0800 8	1999 STORORANCO 7 T N P	0011 <u>11062010</u> V A
1010 REFERENCES L R K L C	1120 <u>SCATISAASC</u> I E A	0190 <u>TAAAATATCA</u> E I S	1140 AACACCACCA N T T T	1150 <u>CCGATICGA</u> D S R	1160 <u>SATOTOCAACI</u> O P T	1179 <u>MCAASGAGAAG</u> Q G E <i>P</i>	1180 <u>XCACACTSO</u> X 1 1 V	1190 10088688688 E E O	1200 <u>9090A97</u> A G
1210	1220	1230	0421	1250	1250	1270	1280	1230	1300
MOTIFICIOIOCOC	<u>NOGRAOGIMO</u>	IS RESERVACIACIÓN	ARDER REPORT	insection (	-schattossa	SANGS SEARSTON	OCODAACAAC	<u>ngociaagtitti</u>	<u>valerore</u>

NEVCRRTTV	O R G N G N G	C G L F C K G S I	I T C A K F K C V
3330 1320 3	.330 1360 3	350 1360 1370	1380 1390 1400
IGACAAAACTAGAAGAAAAATAGTIC			OADAEIAAADDEIEDADAOGADAE
a ki seki ve	YENLKI	S V I V T V B T G	DQAQYGNBT
		480 1480 1470	1480 1490 1500
		<u> 1995 aataorgo tokoogactaoggaa</u>	
		5 8 I Q L T 8 M 6 T 1550 - 1560 - 1870	1 L T L D C C P R 1580 1530 1600
ACASESCHAERTITTERACEBORTS			
T G L D F N B M V	LLTMEEX	SWLVBKØWP	LDLPLP%YS
Engineered silent waterios	(nu-1675 T-uo-C): re	save the native OEN-1 virus	+ specific xbal site
1610 1626 1 CTOCOCOTTOAACATCOCCAACAGACTT		.880 1.680 1.570 Wycapatyttangapacyteratogramac	0001 0000 1500 0800 האמרההיההיהיההההההההההההההה
	INRQDIL		
1710 1720 1	.730 1.740 1	750 1700 1770	1780 1790 1800
- NORACIAGOSCONATOCACIOCOCT		NANCOTOLOGANCOLOLAGANTETIO.	
Q E G A M H T A L	TGATEIQ	/ 7 8 G T 7 T I F A	GHLECRLK
1816 1820 1	850 1840 1	.856 1860 1870	1886 1896 1969
ATGORCARACTANITTEARAAGOGAEG			
MDRLTLKOM	8 Y V 8 C 3 6	6 F K 5 E K 5 V A	B X Q H G X V L V
		.950 1960 1970	1980 1930 2050
		TCGACCCAAGATGAGAAAGGAGCAACC E T O O E K G A T	
A U T & A X A V Q	i s Ç Kırsır	S T Q U S K G A T	Q 10 15 11 15 11 1 18 10
		050 2050 2070	2080 2030 2100
		<u>VACCOTTICGTORCACCIACATOSTOC</u> PPPGB841VVV	
	y 28 1 15 24 15 11		
		150 2100 2170	2180 2198 2298
CTAAGCEGGTTCAAGAAAGGAAGCAGC 2 6 70 7 8 2 9 3 6		TAROAR AMAGAMATAN TAROAR BMA	
2210 2220 2	250 2240 2	250 2260 2276	2280 2290 2300
		<u>CAGUTTITTEGEBACTSCATAHGEBETT</u> OVFGTAYSV	
9 S I G G V P T S	LMGKLVR	Ő A S C X Y A C A	LFSGVSWTM
		m	

### End of D1 16007 sequence

### Engineered NGGMIV splicing site, E-482 Val-to-Ala (nt-2381/2382 TG-to-CC)

 2510
 2520
 2530
 2550
 2560
 2570
 2590
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 2690
 <th

#### D2 HDR-53 N61-53-aup attanuation locue (wt 02 16681; Gly, st-2879-0)

 2710
 2720
 2750
 2750
 2750
 2760
 2760
 2780
 2780
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 2800
 <th

2830 2826 2830 2840 2850 2850 2870 2880 2830 2330

AGTOTOATBACCAGACCTTTCTCATTGATGECCCCCEAAGACCCCCAACACACACACACCTCGAACTCCTTEGAAGTTEBAGACTATEGCTT S H N G T F L I G S F E T A B C F N T N R A W N S L E V B G Y G P

2910 2920 2930 2940 2950 2950 2970 2980 2970 3000 TGGARTATICACCECCATADATGCTBABATICBABGABABACAGEATGTATICTGCEACCABBACTCATRICACGECCABBABABACAGEAGACC Q V P T T N I Q L P L K E R Q D V P C D S K L K S A A I K D N R A

3030 3040 30.80 V B A B M G K W 1 E S A B A B X W K I E K A S P I B V K N C E W P E

33.86 ANTCACACING CONTRACTOR OF A CONTRACTOR OF A CONTRACTANT OF A CONTRACT AND A CONTR SHILWSNOVLESEMIIPPNLAGEVSOHNZRPO

H T Q I T G F W H L G K L E M D F D F C C G G T V V V T E D C G N

133.75 R O P 3 L R T T T A 3 O K L I T K W C C R 3 C T L P P L R 7 K 5 E

#### >NS2A

3450 3450 3470 35.06 A 7060 TO CONTROLO AND A TO A CONTROL OF TAKEN AND A CONTROL AND A CONTROL AND A CONTROL AND A CONTROL AND A CO G C W Y G M E X Y E K E K E E N E V N S E V T A G H G O V D B P

3590 3980 3576 3590 3520 3530 353.0 TTOACTAGOASTCRIGGSAARGGOATTOTTOTTOTGCAGAANGGOTTAGOAGCCGAGTAGJAAGGAAACARGGAATACTACTAGTAGTGCAGTFTTTTTOTG S L G V L G M A L P L E H L K T R V G T K H A I L L V A V S P V

TAPPPARTNESSED A FOR A STATASTAR CARDER AND EVEN AND TO A STATE COULD TO A CARDEN A CARDEN A CARDEN A CARDEN A ? L I T G N M 3 P E D L G K V & V M V G A ? & 7 D I G M G V T Y L

373.6 A L L A A F R V R P T F A A Q L L B K K L T S K B L M M T T I Q I

388.0 TO TACTOR TRANSPORTED TO TRADE OF TRADE OF THE TOTAL TO TRADE AND A TATE OF THE TACADA A TATE OF TACADA A TATE O V L L S Q G T **L** P E T I L E L F D A L A L O M M V L K M V R N A E

#### Additional NS2A-116 Ile-to-Leu (nt3823 A-to-C) mutation in master and pre-master seed

19950 X460 3496 19960 99940 3990 3930 3946 9.85.6 AXOFNYCAXYFGGCASYSXCIATCATGCCTATC TIGTGCCTCCCAAACGCCATGARATACGCATGGBAAGTGACTGACAAGACTGCCACAAGATTGCCACAAGACTGGC К 7 <u>6</u> L A V T I 8 A I L C V P 8 A V I L <u>6</u> N A W K V 8 П Т I L A V V

410.0 

#### D2 PDK-53 specific NS2A-181-Phe (wt D2 16681: Leu, nt-4018-C)

4530 4540

4.526

A\$ 20

>NS3

>N\$2B

4130 4130 4140 4150 4160 4170 4180 ARAKTOTOTOTAROADOSTAROADOSTAROADOSTAROADOSTAROASTAROASTAROASTAROASTAROASTAROASTAROASTAROASTAROASTAROASTAROAST

TTLSRTSRXRSWPLNEAIMAVGMVSILASSLLX

4550 4560 4570 4580

 $A \le Q \oplus$ 

a 6.0 0

OVKWREQAEI3GSSPILSITISEDOSMSIKNER

SEARGET CALACTERIZACIAN CALERATING A CAGGAT FET DECISION OF THE CONTRACT AND A CARACTERIZACIAN CALERACIAN AND A R D Q T L T I L I R T G L L Y I S G L F F Y S I P I T A A A Y Y

TO A A TRANSPORTED AND A DET STORAGE AND A DESCRIPTION OF A DESCRIPTION OF A DESCRIPTION OF A DESCRIPTION OF A

435.0

44.90

Additional NS2B-92 Glu-to-Asp (nt-4407 A-to-T) mutation (in master and pre-master seed)

WEVKKORAGVIWDVF3PFFMOKABIEDGAXBIKO

ABIG 4819 4839 4840 6850 4860 4870 4880 4890 4990 4990 CANGECOLGOTATEGEOCOTGOALAGAATCOALCOCOTECTATECALACOALCOCOCALTAGAGOTOCECTET E V G V L & L E P G K N P R A V G T K P G L P K T N & G T I G A V G

D2 PDX-53 NE3-230-Val attenuation locus (wt D2 19681: Glu, st-3270-A)

5510 5520 5530 5540 5550 5550 5570 5580 5530 500 AGRECARTGACCARTCATAGATUARGAGAGAGACCCTGRACCCTGRACCCTGCAGATCACGGACGATGTTARAGGGAAGACTGTTA 3 N A F I I G E E R E I P E R & W N 8 O R E W V T D P K G E T V W

#### D2 PDX-53 silent mutation nt-5547-C (wt D2 16681: T)

5710 5720 5780 5760 5760 5760 5770 5780 5780 5800 TREDICARGACTAGARGEARTEGOCACTECOTOGETACARCEGACACTECAGARTEGOCTOCAGACGOSTINACACCCCAGAC X V R T R T R D W B P V V T T D I S E M O A N P R A E R V I B P R R

 5810
 5820
 5830
 5860
 5870
 5880
 5870
 5970
 5870
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 5970
 <th

 8010
 6020
 6020
 5040
 6050
 6070
 5070
 6080
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 6090
 <th

>NS4A

 6510
 5320
 6350
 6350
 6350
 6370
 6380
 6400

 AGADEDITGERIGERIGERATURACEACTOGEOUTARREARDERTERSEACTORESEARSPACESCHERACEACTORESEARSPECIESARSPACESCHERACEACTORESEARSPECIESARSPEC

#### D2 PDK-53 specific NS4A-75-Ala (wt D2 16681: Gly, nt-6599-G)

**>N54B** 6810 6820 6850 6850 6850 6850 6870 6880 6890 6900 CACAGEGOCGCAACCACGGGAACGAGAEGGOTTCCTAGAAAAGAEGAECTOGGAEGGAEGGAECGCAACCCGAGAGC T Y A A T M A N E M G P L B P T K P O L G B G S I A T Q Q P E G

6910 6920 6330 6940 6950 6950 6970 6980 5930 7050 ANTATOTEGEATACCATCENCOLOGIATEGEOCOTTEACCOLOGICACCATECTEACACATECCACETECACACETEC N I E D I D E R P A S A W I E Y A V A T I P V I P M I R E S I E N S

7216 7220 7250 7240 7250 7260 7270 7280 7290 7390 CABECARAGURACCAGAGAGUTURGARABURGCAUCECCGUUCATCATURARAGUUCATCHATCURATCGARTARCAUTURTUGACTERCEATURATACU 5 A K A T R R A O F R A A A C I M R N F T V D C I T V I O I D F I F

#### Additional silent mutation (nt-7311 A-to-G, in master and pre-master seed)

 7410
 7420
 7430
 7430
 7450
 7470
 7480
 7490
 7500

 SHOTTTAACCTACCOSCICUTOTCACADTPEGGAAGAAACCTAGCGACGTTTTCGAACACTACCATCCCGOTCTCACADTCCTACCCATTCCCGOTCTCACADTCCTACCACADTCCTACCCATTCCCGOTCTCACADTCCCATTCCCGOTCTCACADTCCTACCCATTCCCGOTCTCACADTCCTACCCATTCCCGOTCTCACADTCCTACCACADTCCTACCCATTCCCGOTCTCACADTCCTACCACADTCCTACCCATTCCTACADTCCTADTCCTACADTCCTACADTCCTADTCCTACADTCCTACADTCCTADTCCTAC

> NS5

7610 7520 7630 7640 7550 7560 7676 7689 7540 7790 AMAATOGAAAAGCOBATTGAGGAAAAAATTGAGAATTGAGAACTGAAAAAAGTGGAATCGAGAACTGAAACAAAGAAGG K W K S R L N A L O K S E F 9 I Y K K O G I Q E V D R T L A K E G

7610 7920 7830 7860 7850 7860 7870 7880 7890 7960 DISCACCICSOITSPOSCACADDADECTODICALACTATIONSCACTATAAAAGCACADEACCAEGACACC V D L O C S R O S W S Y Y C O S L R N V R R V R O L T R O S P O H E

7630 7920 7330 7940 7950 7950 7970 7980 7990 8000 Abshacclatoccateficallatagosficadescetticarastegasficacofficiencatoccoccasaarotocscatetart D P 1 P 26 2 7 7 6 W N L V R L Q 2 6 V 6 V F P 1 P P X K 6 6 T L L

8010 8020 8030 8040 8050 8050 8070 8080 8090 8090 8090 Geographicateaccanatoccacateganscangaccarcactergangecotranotegananategotignacancarcactean C D I O K 3 S P N P T V E A O R F L K V L N L V E N W L N N N T Q 85.90

8118

8266

80.90

83.60

83.26

8180

8130 8140 8150

P C I K V E N P Y M P S V I E E M E A E G R E Y S O A L V P N P E E 8220 8230 8240 8256 8260 8276 8285 8285 8218 8360 CACOBAACTCCACACATORCATOTACTGGOTATOCCAATOCTTCCGOGAACATASTGTCBTCAGTOBACATOATTTCAASOGATOTTCCAACAGATTTCC R N S T H E M Y W V S N A S O N I V S S V N M I S R M L I N B P 838.6 8360 8376 8.826 8340 8380 8.3.36 2315 8330 8660 

B710 B720 B730 B740 B756 E750 B770 B780 B796 B50 GCACAOTGGCTTTGGCAAGAATAGCAAGAACAACACCAGGCACCAGGAGAAGAATGCAAGAAAGGCTGGGGGG A E W L W K E L 5 K E K T P R M C T R E B F T R E V R 5 N A A L 5 A

5010 0820 8530 8040 8050 8850 8070 5880 8970 800 CCATATICACTERTERACATEGRACTCOCACCITERCEACCATEGRACESCOPECITICACTERCEARSCATECTCCATCITER I P T D E N K N X Š A E E A V X D S E I W E L V D K E R B L N L E

 3310
 9320
 9360
 9360
 9360
 9370
 9380
 9390
 9400

 GTOCOMORACIA
 GTOCOM

9810 9820 9830 9850 9850 9850 9870 9880 9890 9950 AGGOLOFOTTOCOGRAGOOGICTUTTGGGAAGTCTGAOGICGAARTGTGAOGICTGACGCGGAART G W G L R D T A C L G K G Y A Q M W S L M Y F H X R D L R L A A N

												-											•							-	)									
	3.0	110		3	01	20			1.01	130			33	14	0			iO	180	)		$\sum_{i=1}^{n}$	) 1 C	0		3	0.2	20			10:	180	)		10	190			3.0.	20
A.A.M																																								
0	2	0	3	8	Ti-	1	(		( i		3	ĩ	2	A	4		88 88	A	×	5	8	Ŷ	Q		R,	Y	1		N	Q	Ŷ.	P.	<	S	1.	Ţ	0	N	;	16
																													>	.3.	~N(	onc	bot	in	3 F	eqi	.on			
	3.6.	210		)	02	20			1.03	230			13	24	Q.			10	250	)		23	226	0		3	62	26			1.60	280	2		10	290			1.0	(90)
NAA T	$A \cap A$	CAG	<b>N</b> T T	8.CP			811) 	:CA	393	K.A.P	26	<i>\$</i> 13	зCУ	(GA	<u> 4</u> 6	340	AА	634	964	AC)	A.A.	90)	NG0	AG	9726	777 G	710	972	A.GA	A3	900	0 N.A	SAC	TX	N. (2)	TGA	AA.	(TAA	66	0T
S - X	T.	15	Х	3e	3	£1	3	$\otimes$	3	K	R	<u>)</u> ;	3	Ş	R	R		2	32	)i		K,	0	V	X		80	*												
	3.0	310		3	0.3	20			503	330			3.0	334	8			10	350	;		1.1	539	0		3	03	70			3.03	580	3		10	390			3.0	40
3 <b>4</b> 46	TCA	997)	xo	823	ЪА	600	<u>(</u> )	(A.)	EA(	396	8A	5.A3	NA(	:32A	are)	C?	AC	CT:	01YC	9AG	ecc	cĠ	310	C35.	A(K	340	2937	TR	8A8	्र	AG:	202	seg	00	ATC.	A92		TGT	CA	ΤA
	1.0	<b>8</b> 3.0		3	0.4	20			3.64	630			11	584	0			10	490	)		3 :	345	ö		3	.04	26			10	<b>19</b> (	;		3.0	490			3.0	50
P76	NGT.	AAA	ATC	9QC	24G	ca	ren	20	C.37	202	00	res	NO7	W.S	GJ	G3.	XA.	AA.	AA0	900	SG	ÇA	260	ΩŊ	Ŵ	NA(	NOA:	36	0A9	950		PAC	250	Alt	9GC	532	.GT	692	C.S.	ΑÇ
	3.0	S:1.0		3	05	26			16	590			23	54	0			10	95.	;		3	32.6	0		2	.05	76			1.03	586	3		3.0	590			3.6	60
GCPR.	AGA	GOA	SAC	aac	TPC	ce	P 19	<u> </u>	AA'	rce	23	367	M(	YA A	TC	60	03	ec:	CA3	106	00	A03	495	λA	9C:	101	3403	3°C	301	300	'GG3	<u>(</u> )	:OA	$c\pi$	407	.072	Ϋ́Υ	OAC	NUM	СA
	20	610		7	$6 \leq$	20			i ne	ase			20	152	n.			3.0	846			1 :	56	a		3	0.61	76			3.64	681	;		10	5.95			10	76

10710 10720 Otoctottaatcaacasoftet [0172] DENvax-2 Master Virus Seed (MVS)

**[0173]** Nucleotide sequence of the recombinant viral genome and deduced amino acid sequence of the translated protein are provided herein. The engineered virus is based on D2 PDK-53 virus. All engineered substitutions that are different from wild-type DEN-2 16681 virus (also the parental virus for PDK-53), as well as extra mutations (changes from engineered cDNA clone) detected in the MVS are marked.

[0174] Substitutions Included in the Genome and Protein:

**[0175]** D2 PDK-53 virus backbone (change from wt D2 16681): all in bold

**[0176]** a. 5-noncoding region (NCR)-57 (nt-57 C-to-T): major attenuation locus (in red)

[0177] b. prM-29 Asp-to-Val (nt-524 A-to-T)

[0178] c. nt-2055 C-to-T (E gene) silent mutation

[0179] d. NS1-53 Gly-to-Asp (nt-2579 G-to-A): major attenuation locus (in red)

[0180] e. NS2A-181 Leu-to-Phe (nt-4018 C-to-T)

[0181] f. NS3-250 Glu-to-Val (nt-5270 A-to-T): major attenuation locus (in red)

[0182] g. nt-5547 (NS3 gene) T-to-C silent mutation

[0183] h. NS4A-75 Gly-to-Ala (nt-6599 G-to-C)

**[0184]** \* nt-8571 C-to-T silent mutation of PDK-53 is not engineered in the vaccine virus

[0185] Engineered clone marker (silent mutation):

[0186] a. nt-900 T-to-C silent mutation: infectious clone marker

**[0187]** Additional substitutions found in vaccine seed (0.02% nt different from original clone)

[0188] a. prM-52 Lys-to-Glu (nt-592 A-to-G), in bold [0189] b. NS5-412 Ile-to-Val (nt-8803 A-to-G), in bold

130 AAAAACACGCCTT KNTPF											
250 GCCCTGGTGGCGT	260	270	280	290	300	310	320	330	340	350	1
A L V A F							KSK.				
370 SGAAGGATGCTGA	380 ACATCTTGAAT	390 AGGAGACGCA	400 GATCTGCAGG	410 CATGATCATT#	420 ATGCTGATTC	430 CAACAGTGAT	>prM 440 GGCGTTCCAT	450 TTAACCACAC	460 GTAACGGAG	470 AACCACACAT	GATCO
5 R M L N 490	I L N 1 500	R <u>R</u> RR 510	S A G 520	M I I M 530	4 <u>L</u> I F 540	т V М 550	А F H 560	L <u>T</u> T F 570	NGE 580	Р Н М 590	I
AGCAGACAAGAGA SRQEK	AAGGGAAAAGT	CTTCTGTTTA	AAACAGAGG <b>T</b>	TGGCGTGAAC	ATGTGTACCC	TCATGGCCAT	GGACCTTGGT	GAATTGTGTG	GAAGACACAA	TCAC GTAC <b>G</b> A	
02 PDK-53 spe	cific prM-2	9 Val (wt 1	D2 16681 A	sp, nt-524-	-A) Addi	tional prM	-52 Lys-to	-Glu mutat	ion (nt-5		
610 TTCTCAGGCAGA	620 ATGAGCCAGAA	630 GACATAGACT	640 GT TGGTGCAA	650 CTCTACGTCCA	660 ACGTGGGTAA	670 CTTATGGGAC	680 GTGTACCACC	690 ATGGGAGAAC	700 ATAGAAGAG	<b>&gt;M</b> 710 AAAAAAGATC <i>i</i>	
. L R Q N 730	E P E I 740	D <u>I</u> D C 750	W C N 760	S T S 1 770	с <u>w</u> V I 780	Y G T 790	СТТ. 800	м <u>с</u> ен 810	1 R R E 820	K R S 830	V
TCGTTCCACATG V P H V	rgggaatggga	CTGGAGACAC	GAACTGAAAC	ATGGATGTCA	rcagaagggg	CCTGGAAACA	TGTCCAGAGA	ATTGAAACTI	GGATCTTGA	GACATCCAGG	CTTC
850	860	870	880	890	900	910	920	930	> <b>E</b> 940	950	
TGATGGCAGCAA IMAAI											
0.50	020	0.00	1000			clone mar	ker: nt-90 1040				1
970 GAGACTTTGTGG. . D F V E							GATGGCAAAA				
1090 AAGCCAAACAGC A K Q P	CTGCCACCCTA		GTATAGAGGC		ACACAACAA	CAGAATCTCG	CTGCCCAACA	CAAGGGGAAC			
1210 GGTTCGTCTGCA . F V C K	1220 AACACTCCATG	1230 GTAGACAGAG	1240 GA TGGGGAAA	1250 TGGATGTGGA0	1260 CTATTTGGAA	1270 AGGGAGGCAT	1280 TGTGACCTGT	1290 GCTATGTTCA	1300 GATGCAAAA	1310 AGAACATGGAA	1 AGG <i>P</i>
1330 TTGTGCAACCAG. VQPE	1340 AAAACTTGGAA	1350 FACACCATTG	1360 IGATAACACC	1370 TCACTCAGGGG	1380 GAAGAGCATG	1390 CAGTCGGAAA	1400 TGACACAGGA		1420 AGGAAATCA	1430 AAATAACACC <i>i</i>	: ACAC
1450 CCATCACAGAAG	1460 CAGAATTGACA	1470 GGTTATGGCA	1480 CTGTCACAAT	1490 GGAGTGCTCTC	1500 CCAAGAACGG	1510 GCCTCGACTT	1 520 CAATGAGATG	1530 GTGTTGCTGC	1540 AGATGGAAA	1550 ATAAAGCTTG(	: GCT0
I T E A 1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	-
ACAGGCAATGGT R Q W F											
1690 ATGTTGTTGTTT V V V L				ACTTACAGGG		TCCAAATGTC					
1810 TGGACAAGCTAC.	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	-
скь Q 1930											
ACGGCTCTCCAT G S P C	GCAAGATCCCT	PTTGAGATAA'	rg gatttgga	AAAAAGACATO	STCTTAGGTC	GCCTGATTAC.	AGTCAACCCA	ATTGTGACAG	AAAAAGATA	GCCCAGTCAA	CATA
2050 CAGAACCTCCAT E P P F											
2 PDK-53 nt-	I	-			-			-			
2170 CGAAGAGAATGG KRMA											
2290 AGTGGGGTTTCAT											
GVSW	т м к : >NS:		VII	TWIO	g m n s	R S T	SLS	VTLV	'LVG	I V T	L
	- 40.	-						2490			

2530 2540 2550 2560 2570 2580 2590 2600 2610 2620 2630 2640

атослассковатосостискавающостиску посковкающих состоящих саконстрику саконстраний саконаловоса. С С Р R & F S K U & R R I Q K & B R B I C & I P S V I B \_ S V U M K Q I T P }

B2 908-88 851-53-Rep attanuation Locus (at 02 16601: 63y, st-2279-6)

U CONSTRUCTION CONTRACTOR CONT 2.91.0 31.90 32.10 PEZC 31.59 33.90 3399 3319 

3870 3580 3594 3530 3546 \$599 2.87.55 CATURD CARGE CUALINATETET CALLED AGENCIATION DA FORCATEUR POOTGRABURANT OCTA REGALLONAGEN CARACTER ETACTAC TA CELEVICA, CARACTER ETACTAC 12 - 28 - 15 - 15 A L <u>S</u> V L G N A L V L S <u>H</u> A L P 7 X V G 7 K S A T L L V A V 3755 3765 3770 3780 3790 3800 3810 3820 3830 38.69 

#### D2 PDE-53 specific NS2A-181-Phe (wt D2 16681: Leu, nt-4018-C)

42.45 42.30 433.0  $\sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \sum_{j=1}^{N-1}$ 4360 4370 4380 4390 44(*A*) 4416 4420 **EEM<** 6836 6268 6836 0.5.2011 \$520 48.30 8 800 8.2622 актереруйское атайлогетельтерельным мал 4720 4730 4740 4750 4766 4776 4386 CORECCUENTETRADECAALASIANDA FEGRALASIANDA FEGRALASIANDA CANANTÉ DACAALUS D'EDUTATA FEGRACAALUS EDAGAALASIA SUBACCAALASI R G X V L H N K G <u>B</u> R I E P R H A D V <u>E</u> E D L I S Y C G G <u>M</u> R L R G R K K G B 0.8850 \$ 8 5 9 0,888 8.000 0.030 6.99.6 5090 5100 5110 5120 5330 5140 3150 5160 UM NOCOLOGINA CALORIANTY I COLOR MARAKA ALTING CATORI CONTORIO CALORIZATIO CALARIA CANARA ALTINI COLORIA DI PORC D N F E F & D D T Y E N X & L 7 I N D L N P G A C X 7 K & Y L P A I V R E A F K

>NS4A

1

# 

32

#### DE ERR-63 Web-186-Wel ettempetion Locus (D2 18661: Glu, et-8270-8)

#### D2 PDK-53 silent mutation nt-5547-C (D2 16681: T)

5700 8700 8800 8810 5828 5835 5046 5050 5050 5070 5330 NAEBY LEVRECHKEV LLIDGEERYILKOPH VIN 85 3 8 A 5920 5320 \$340 \$355 \$360 8970  $f_{i} \otimes f_{i} \oplus f_{i}$ 5.92.0 Steel. +12(2)(2) + 16020 6930 805675 \$5.83 64.86 \$255.737 \$ 6080 2000 21.00 6330 P & <u>3</u> 7 7 9 8 M 7 8 F 8 <u>8</u> 6 M 7 6 8 7 8 8  $-\mathcal{B}=\underline{Y}=\mathcal{B}=L_{1}$ RIMPRODE XVMLETEVAL<u>E</u>GINYAOPRW<u>G</u>E J T Q V K N V Q T L 6270 6280 6200 6300 \$310 \$320 \$330 \$330 6268 6350 \$250 6360 

#### D2 PDK-53 specific NS4A-75-Ala (wt D2 16681: Gly, nt-6599-G)

6670 6630 6630 6766 863.6 6626 \$833 \$630 6850 6880 ) X X 8 F >NS4B 6720 C748 6780 8760 8760 6770 6780 6730 6876 6810 6820 6830 6848 スプムロジアログスのひつえんなななななななななないのつんなないないですないのためつうかなりないのようなりつうないのかはかいたいかいかいないのかいかいがいかい 1 9 1 1 1 9 8 5 1 9 0 8 7 9 0 5 9 0 1 9 1 7 7 7 7 7 8 5 1 7 7 2 8 7 8 8 8 8 8 9 9 \$870 \$886 6886 6366 6980 6360 6916 6930 \$940 2000 \$990 7010 7020 7830 7840 7850 7866  $\gamma_{2}\gamma_{3}\phi_{2}\phi_{3}$ 8976 1980 3970 7330 71.89 Ν. 7396 2820 7530 7530 7880 2480 7000 7810 2550 2.650 78.80 7540 7580 EANYBAARCCAREURGETTITEEANDWETACEATTUUGETGICAATGUTTAACAETTUTKEAGGCAETTWETGGECEEMUTGGGCTTUTUTTFCCAATAETUAGAACAERAGCARGOCAAC E & N P & X P W B <u>T</u> T I A V S M A N I <u>F</u> X G S Y L X B A G L X F S I M K G T P N >NS5 3570 7570 7589 7600 7610 7620 7630 7660 7650 7660 7670 7680

7280 7750 7750 7778 3740 7795 2720 7830 7840 7880 7869 7900 7910 7926 2876 2880 2896 5820 OTORACCE CONTOTOR COMMENCE CONTRACTATE COMPONENCES SET OTAGECEROTOMACCCE ENGLACES SONOCACORCES CONTRACTOR CONT C L K N X L B Y X C L 7 K C Q P C S \$ \$ ¥ 3 Y Y C G 8 8 8 7 9 M = S7930 7949 1950 2950 2910 7980 7990 8000 8010 8020 8035 8046 ACRIATOGOTOCIALTIAGTOCOPTITICOAAAACTICOCOCCIALAAACTICOSTAACATALTORSTGACATACCOCACTATICOACCOTOLAA すりと別のし、VALQとCUD VPをLP<u>P</u>をよりつうしたとし<u>え</u>Cにたまたのかり、 なりためのし、VALQとCUD VPをLP<u>P</u>をよりつうしたこと<u>え</u>Cにたまたのかりでし 8675 0556 0558 0558 6058 6953 6853 6853 8253 0558 0558 0558 0558 0558 DIALASSANTATARAADADATARAADADATTORATICARBANTATARAADADTORATARAADADTORAADADTORAADADTORAADADTORAADA R 3 1 M R V 3 C V 1 <u>8</u> D 3 K C V W Y R <u>3</u> R Y 3 M R 3 1 Y R <u>4</u> V J K D Y R 4 G \$256.23 \$300 8310 8326 \$330 8.54.5 8345 8360 8326 8380 0.000 RTUT TURFORACEUR TURACEATURUR TEORERERENT FOR CONCURRENT CONTRACTOR CONCURRENT CONTRACTOR CONCURRENT AUXILIARIE 8450 8480 8564 \$426 8430 8440 8470 8680 8636 8510 8530 8540 8550 8550 8550 8570 8580 8580 8600 9610 9620 8630 5440 Tunnun sedula usa kenta unipertalia ando peruatu perfecti and perta kan secal tan inclusion and the secola seco Q. 8710 8886 8 6 2 8 6838 8690 8700 8726 8746 8759 8800 8810 8829 8830 8830 8856 8730 3335 ани и на каказы на каказыкана каказыкана каказыкана каказыкана каказыкана каказыкана каказыкана каказыкана кака Каказыкана к Каказыкана к Additional NS5-412 Ile-to-Val (nt-8803 A-to-G) mutation in master and pre-master seed 8320 8930 8940 8978 8366 8910 835512 staw. 8586 CALCENTITICAL CETTURAGEARAN STUTANECH DESTUTANECAL DESTUTATOGRAFIAS ANAVARAK CARCYTANGGEART COURT AND CARAFTSCACCASASUCATANGGEART E B C H L B C K Ç B T C V Y N H M G E B B K K L G E F G X A K G E B N I N Z N L G A R F L E F Z A Z G F L B R D H ¥ F S R A U S L S C <u>Y</u> B C E S U R Z G Y 9180 93.60 9170 9200 01.00 6276 4220 \$224 61.84 44 S.N 30.00

 93.50
 91.40
 91.50
 91.60
 91.80
 91.80
 91.80
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 92.00
 <th

о каса блас Хорнасли и и Убор кой кой кологой раз 9520 3500 953.0 9530 9540 95.50 95.95 \$538 9686 3540 36.00 9650 228.0 2250 377.0 9788 \$7.90 057.6 9910 32.40 44.97 38.56 3846 ĸŢġĸĨĸĬŀŴŮŔŦŎŔĊġŇĸŢŦĊŊĸŢŦġŦŢĊĊĸŢĸŦŦĸġĸĸĔĊĔĸŀĸĸŶŎġĸĊŦĠĸĊĊġĊĸĹĸŀĸĸŦŎġĊĿĊĬĸŀĸĿġĸġĊġĸĠĿĊĿŦŢŎĊĿſĸŔŴĬĸĊġġĊĊĸŦŢŦŢĬġġġĸĸĊĊŢ NEDGEATAADCEESGDEPLGBFHIQGGFOMSTEDDIGCCCCC 9880 \$870 9889 9896 9966 9926 9920 9930 9940 9950 9960 1990 TACKORGENGATURACTICACATANCECACATANCESCECCAMUSCICACUTIC TOTAC ACCATEGATURACACATANCACATANCACATANCACACUTINACACC Y A ζ H & B L B Y <u>§</u> H R R D L R L A B <u>N</u> A I C B B Y P B B <u>B</u> Y P I S R I I **B** S I Y A ζ H & B L B Y <u>§</u> H R R D L R L A B <u>N</u> A I C B B Y P B B <u>B</u> Y P I S R I I B 9990 10000 10010 10020 10030 10050 10050 10050 10075 10085 8.576 9985 CATHER A REAL AND A REAL AND A REAL AND THE DESTROYMENT OF A THE ARRAGE THE ARDER AND A REAL AND A REAL AND A R Y K S S M X Z L B D W P L A M S K A M L O S S A M R B D K L D A S S M S S L D 10300 30150 10160 16136 0014-10180 19110 16120

**38-**" ⊓⊂ D260° D260 D260 D260 C8301 C8301 D250 D2501 D250

10090 10100 10100 10100 10100 10100 10100 10100 10100 10100 10100 10100 10100 10100 10100 10100 10100 10100 1010 TATTTELEENAMULACHETANEROLOGISTITUSTENTANERISTINGEN ANTALESTITUSTEN ANTALESTINGEN ANTALESTINGEN ANTALESTINGEN A Y E O K R & D Q W G O Z L I O Z T O R & T W K K H X Q A & G K & C R & L I & S X аттот**хэ**ав<u>жжевже</u>ве<u>ее</u>усж

**[0190]** Nucleotide sequence of the chimeric viral genome and deduced amino acid sequence of the translated protein are provided herein. Most of the prM-E gene (nt-457 to -2373, underlined) is wild-type (wt) DEN-3 16562 virusspecific; the remaining nucleotide sequence is DEN-2 PDK-53 virus-specific. The E protein of DEN-3 virus has two fewer amino acids than the E protein of DEN-2. Therefore, nt position starting from NgoMIV is 6 nt less than the original DEN-2 PDK-53 nt position. All engineered substitutions differ from wt virus (DEN-3 16562 or DEN-2 16681), as well as extra mutations (changes from engineered cDNA clone) are marked.

Substitutions Included in the Genome and Protein

[0191] Junction sites:

- [0192] a. MluI (nt 451-456): engineered silent mutation, nt-453 A-to-G
- [0193] b. NgoMIV (nt 2374-2379): engineered mutations, nt-2375/2376 TG-to-CC (resulted in E-480 Valto-Ala change)
- [0194] D2 PDK-53 virus backbone (change from wt D2 16681): in bold
  - **[0195]** a. 5'-noncoding region(NCR)-57 (nt-57 C-to-T): major attenuation locus (in red)

- [0196] b. NS1-53 Gly-to-Asp (nt-2573 G-to-A): major attenuation locus (in red)
- [0197] c. NS2A-181 Leu-to-Phe (nt-4012 C-to-T)
- [0198] d. NS3-250 Glu-to-Val (nt-5264 A-to-T): major attenuation locus (in red)
- [0199] e. nt-5541 (NS3 gene) T-to-C silent mutation
- [0200] f. NS4A-75 Gly-to-Ala (nt-6593 G-to-C)

**[0201]** \* nt-8565 C-to-T silent mutation of PDK-53 is not engineered in the vaccine virus

**[0202]** Engineered mutation in DEN-3 prM-E (change from wt D3 16562)

- [0203] a. Engineered nt-552 C-to-T silent mutation: clone marker
- [0204] b. Engineered E-345 His-to-Leu (nt-1970 A-to-T) for efficient replication in cultures

**[0205]** Additional substitutions found in vaccine seed (0.02% nt different from original clone)

- [0206] a. E-223 Thr-to-Ser mutation (nt-1603 A-to-T, in bold)
- [0207] b. nt-7620 A-to-G silent mutation (in bold)

			NCR-57-27	D2 PDR-53	attemati	on Loous (	wt D2 1068	1:0)		
> 5'-	Noncoding	Region				1				> C
	10	2.0	8.9	4 Q	5.0	60	2.0	80	0.6	100
AGTTÉ	MITAGROTACO	TEGACOGAC	AAAGACAGATT	CTTTGLOGG	AGGTAAGGTG	AATOTAGTTC	TAACAGETET	ttaattagag	agcacatcy:	Iteatga
										M N

330 140 150 1.60 376 180 550 205 110 3.20 N Q P K K & K N T P P N K L K R E K N R Y S T V Q Q L T K R P S L

326 330 340 195.0 946 370 300 396 400 TOSOGRACRACTRARRARTCARRAGCTATTARTOTTTTOADAGOSTTCARGARADARTTCOARGATOCTOARCRYTTTOARTAGGRORGCACATCTG **WEELKKEKAENVLREFREEEERMINIINRERREA** 

> prM Beginning of D3 16562 sequence 440 450 1460 470 480 490 

#### Engineered Miul splicing site (nt-453 A-to-G silent mutation)

540 64.0 55 65 (C 8.7.0 513 526 \$30 589 €90 К Ť A X  $\mathbb{N}$ Μ X I. I. A MDLGEMC O. Ð T. Y. 2 1

#### Silent C-to-T nt mutation as clone marker

660 680 870 €90 610 620 630 840 690 700 CECATIZACESAESTOSAESTAESAESTESAESTOSAESTOSAESTOSAESTOSAESTAESAESTOSAESTAESAESTOSAESTAESAESTAESAESTAESAESTAESAES R I Y E V E F E D I D C M C N L Y S T V W Y S T C R Q A G E H R P

720 
 719
 720
 730
 740
 750
 770
 780
 796
 800

 SOURTABUSCIONSTINUTIONALITICOLATION COLSPANDING CONTINUATION COLSPANDING 830 840 896 8.50 810 890 8.80 87.8 999 90.0 CARGOTACADACATOSOCICTADSCACCERGOGITIRCCATACTACCCCTATITCICCCCATACATACSCACTICCICACCACACAGOTOSITATI R V E T 20 A L P E R O P T E L A L P L A B I I G T S L T Q K V V I

> E 930 340 950 950 976 980 93.9 920 9.50 TITATACT&TTAATOCTSCITACCCCATCAATGACA&TGAG&TGTAGG&GCA&GGAGAGAGAGAGAGAGCCTATCGGGAGCTACSTGGGATTG I L L M L V T P S M T M R C V G V G N R D P V B G L S G A T N V

1630 1070 1100 10.20 3040 2050 3880 2080 1090 1010 ACCIDENTICAL CALCALOR CALCAL T Q L A E A

1200 1320 1130 3340 3360 1120 13.36 13.90 1333 OTTARGEBASCITATEURTTCAGESARARBITACCABURTRACHACCEACTURAEATSTCCCACOCHAEGECARGCOATTTTRCCTERGEAUCASCRCCAU 

1230 1240 125.0 3,260 1.270 1280 1290 3.23.0 3.226 1.300 raciacelololaascalacalacologocacacacelisocaaacelistestistischasceaacelisocscaaacelisaasel N Y V C K H I I V E R G W G N G C G L F G K G S L V T C A K E Q C L

1390 1320 1536 13.40 1340 5 3 400 1380 1496 1310  $\frac{1}{1} \frac{1}{1} \frac{1}$ 

14.20 14.30 1480 3480 1470 1480 1410 3,4,4,0 1490 1800 

	1.	510		1	520			153	90		15	s <b>4</b> 0			1550	)		3.54	50		1	570			1586	3		156	30		34	50.0
11136	ATT	n (XA	CEGA	AAT	(JA T	0796	[X,T]	GAC	XAN	r GAG	NGA3	NCA.	(ACC	$(\mathbf{A}^{*})$	SGA	Neci	TAC	A TA	VACO	λ Arr	een	nch	776	$\mathbb{A}(\mathbb{C})$	$(A_{i})$	an a	eace	$(\mathbf{R},\mathbf{V})$	33AC	:Arte	1AG9	$ \lambda $
X. D	F	$\mathcal{K}$	35	14	χ	S	х.	3	М	ĸ	N	2	А	Ŵ	.03	V	В	R	Q	8	R	F	0	χ,	P	E2	8	\$\$	2	£	9	8
	3	610		3	626			161	30		] (	54.0			1853	)		3.8	50		3	670			1686	)		3.63	30		2.3	70 C
<u>0770</u>	AGCO	rgað	A A	CCA	8/CT	26K	SA&C	38.34 <u>)</u>	3A3	V684	Ett.	1021	GT(	SAC/	8.T.T	Γ&Α	SAB	79C)	ACA:	19C)	$\delta \Delta \delta$	AAA	6-CA3	<u> 80 8</u>	re de la compañía de	<b>NG 20</b>	IGT/	1011	19497	0003	NC&8	NG S
8	$\mathbb{A}$	X	З	P	2	N	55	191	X	H:	17	.t.	V	3	P	Ж	N	A	81	Α	25	Ж.	Q	E.	v	V	V	1.	G	$\Sigma_{-}$	$\odot$	8
1																																
Addi	tio	nal	E-2	33	Thr	-te	>−Se	er n	nute	itic	m	(wt	D3	16!	562	: n	t-1	603	A)	т£	ťла	ste	r a	nd :	pre-	-mais	ste:	r \$6	eed			
																									-							

1710	1720	1730	1740	3780	1760	1770	1780	1790	1808
<u>urateraatee</u>	CACTACTOR	CAGEAGCTAC	AGAGAIRCEAA	accittaggas	iotacaastat	CEPTIOSESSO	Cacettaaaat	GTA9ACTCAA	<u>BARGAC</u>

О А И В Т А Ь Т 9 В Т Е І О Т 5 6 0 Т 3 І Г А 6 В Ь К С Я Ь И D 18.20 1830 1880 1870 1985 1.8.90 ġĸĸŢŀġĨĸĊŗĊĸŔġġŎĨĸŤġĸĠŢĸĨĔĊĸĔſĠŢŀĊŢĬĠĸŨĸĠĊŢĨŶġŦĠŢĬĊŗĸŔŎĸŔĸŎĸŎŗĊŢĊġĸĸĔĠĠĸĊĸĊġĸĔġĸĸĸĬĊĸĊĔĿĔĸŎġ 1950 1960 1970 1980 1,290 THEASTAC MARGEMAGATCA TO THE INCATTOCT TO TO ACCESSION TRANSPORTATION CACAGO TRADECIAST BY KOEDA PC KIPFSTEDOOG KALNORLITAN P9

#### Engineered E-345 His-to-Leu (wt D3 16562: nt-1970-A) for efficient growth

2070 2080 205.0 SUTACOAR FRASSACHUSCOLUTOARCAIDASCAISACOLOCITIZGOASAAASTASTAACAATASTAATASTAATASTAATASTAATASTAATASTAATASTAATASTA

23.20 95.96 TSO FATAAAAAAAAAATTTSAFTSGFAAGAFGTTICGASGF0ACTSCC40665T5CA6665CCAFCGTIG5GAGACACACACCCF66GACTFF5GAF S S I C K M P BATARGARKMAILG B 17 A 18 0 2260 2270 2280 2290 2230 2240 2250 

TICTEGATGATEAAAAT S W M M M 1 

#### End of D3 16562 sequence

3 M S 8 - 3 Q L Y L Y L Q K K

#### Engineered NgoMIV splicing site, E-480 Val-to-Ala (nt-2375/2376 TG-to-CC)

> NS1 28.50 

25.80 25.99 E Q Y E E Q P E S P S K L A S A I Q R A H E E D I C S I R S V T K

#### D2 REX-53 ME1-53-Aep attenuation locus (wt D2 16681: Gly, nt-2373-6)

2630 2640 2650 2650 2670 2680 2690 I B N I M N K O I T E B N N I B S B N B V B B T I M T S D I K S

223.0 7.7.2.0I M Q & G X R S 5 R P Q P T E 5 R X S W R X W G R & X M 5 S T E S R

A FRACCAGACCT TTOTCATTOR TOSOCOCOGRAGACKOCRGRAFACCCCCARCACRATER SECOTO SERTOPTOS CARGET SAGACCACTATOSCT TOSCAP N Q 7 F L I D 5 P E T A E C P N 7 N K A W N 3 L X V E D Y 6 P 6

9020 3030 3040 9050 3020 3070 3030 303.0 9090 3100 on can be a considered and the second and the second and a second the second the second se

A D M G Y W Y B G A L N D T W F I E F A S F I E V F N C H N P K S H 3140 3150 

SCACOCIONOCACIÓN DA CACADACIÓN CONCECTA CACADO CACIDA CACIACIÓN CACIACIÓN CONCECTACIÓN CONCECTACIÓN CONCECTACIÓN TL B S N S V G B S E M I I P K B G A G P V S Q E N S P S X E T

32.96 3.85.0 

335.0 COUPCENT SASAACABUDAD TYOUT CESSABAD TUBE BACADBAL OF TSOCCATELE SCALAR FROUD COUPLAGE ACAGAGO FOR SOUT PSKRTTIASGKLITEWCCESCTLPEKYEGRDGC

						>	NS2A		
3418	3420	3430	3440	348.0	3460	3470	3480	3490	95,6.0

CONFERENCE OF A CONTRACT AND A CONTRACT OF A CONT G M & I R P L X E K X E R L V N S L V T A G B O G V D N F B L

O V L O M A L E E E M L E T R V G T E H A E L L V A V E F V T E

3-51.0 TO DEPT OF TARGET TARGET TARGET AND A DESCRIPTION OF TARGET AND A DEPT A DEPT AND A DEPT I T G N M S F R D L G R V M V M V G A T M T D D E S M G V T Z L A L

39.0.9 TADTACCAGECTTCAAACTCAGACCCAACTTTTECCAGCTEESACTACTCCTGAGAAAGOTGACCTGCCAAEGAATTCATEATGACTACTATAGESACTEETACT 

38.66 9.000 L S Q S T I P E T I L E L T D A L A L G M M V L K M V K N M S K Y

395.0 3.370 CAAPTGGUAGPGACTATCATGGUTATCTTGTGGGTUGUAAAGGGAGTGATATPACAAAAGGGATGGAAAGTGAGTGGUGUAATATPGGCAGTGGUGUG Q E A V T E M A E L C V E N A V E E Q N A W K V S C T E E A V V S V

43.06 

#### 02 PDK-53 specific NS2A-181-Phe (wt DZ 16681: Leu, nt-4012-C)

> NS2BCOTOTCAAGAACCABCAAGRAABEGAECT GECCRTTBAATGAEGTPA TCATEGCAGT DEGCATEGTGAEGATTTTTAGCCAGTTOTOTOCTAABABATGAC L S R T S K K R S W P L N R A I W A V G M V S I L A S S L L K N D

ATROCARS REPORT FACTOR OF A CONSTRUCTION OF A CO I P M T S P L V A G S L L V C 7 V L T G R S A S L E L E R A A S V

4380 4390 \$350 4.800 TORATO SOLA DE ALCARCA DE ALCARCA DE ALCARCA DE ALCARCE DE ALCARCE DE ALCARCA DE A X N E O Q A E I S G S E I G S I T I S E O G S M S I B N E E E E

4420 4430 4440 4450 4450 4470 4480 4490 ACABAGET FECTATES CATEGO AND A CONTRACT STATES AND A CONTRACT AND G T L T I L I N T C L L V I S C L F F V S I P I T A A A W I L W H

> NS3 808.0 44.50 88.7.0 45.80 

4510 4620 4€30 4540 4650 4650 4€70 4580 DEALST STRUCT OF ACTIVE OF A CARACTER ACATED OF TACK AND A CARACTER DE THE ACTIVE OF ACTIVE A I L G Y G Q I G A G V I K E G I F H I M W H V T R G A V L M H K G

473.0 ARAGAGINA DE CANDRAD TA KARADARADA TA CANDA CANDA A TA CANDA A CANDA A CANDA CANDA CANDA CANDA A CANDA A CANDA C ERIEPSWAOVERDLISYOOGWELEOEWEEGEEV

48.30 

\$95.6 A 4 4 6 PSPGTSGSPIIDXKÇKVVĢIYÇNGVVFRSGAT

\$040 5.05-0 \$-0.8.0 

\$110 \$140 \$3,60 \$1.80 ST 90 

GCASCEGGANAGACGANGAGATACTTTCCSGCCATAGTUAGAGAASCTATAAAACGGCGTTGAGAAAACTAATCTTGCCCCCCACTAGAGTTGTGGCAG G A G K T E R Y L P A I 9 R E A I K R G L R T L I L A P T R V V A A

\$220 5230 5240 5250 5260 5270 5280 \$2.943 5305 5210 ACTORE ACTOBASE ACTOBASE ACTOBASE ACTIVATION ACTORES ACTORES ACTORATED ACTORATION ACTORATI EMBEALBGLPIRIOTPAIRAVHTGEEIVOLGCH

#### D2 PDR-53 N83-286-Val attemustion logus (D2 16681: Glu, nt-8270-A)

5340 5950 \$370 5360 9396 5383 5313 8826 83.06 9.8.6.0 

8420 8430 5440 5480 95.96 545.0 5450 5470 5430 8410 Y I S T P V E M G E A & G I P M T A T P P G G K P P P Q G N 8 8 6

5808 3520 5530 55-50 5580 5590 5540 5550 5576 953.0 A P I I P E R B I P E R S W N 3 G H E W V P D F K G K T V W F V

#### D2 PDK-53 silent mutation nt-5541-C (D2 16681: T)

5640 5650 5560 5670 5480 5690 5620 5630 5610 9766 0% exters to 200 a carbardaria a trada a carbardara carbara carbardi do eta carbardaria e carbardaria e carbard PSIKAGNDIA & CLRENGKEVICLSRET # DSEX

9360 9739 5740 5760 5770 5750 8780 5710 5726 8766 ADOTOCIAR ASSOCIAR K T R T N S W D F V Y T T D I S B K G A N F K A E R V I D F R P C M

5840 5866 - 5870 5880 5830 5850 9900 5820 5830 5810 TEABASCORGTORTAS TRACKORATESTERAGEGOGITERTTOTGCORGEGOGOCIPATECIONACTOTECTOTECTOREGEORARAEEASGEGOGEGEATARG E P V I L T D G E E R V I L & O P M P V T H G S A & Q R R G X I G

5930 5940 5950 5960 5,970 5920 5980 5990 6000 Sana RNFENENDQYIIMOBPLBNDBDCABWKBAKMЬE

<del>6</del>676 66.30 5048 8080 6363 \$080 801.0 68.63 6899 6108 CATAACATCAACACCCCACAAGCAAFCATCCTACCATCTCCAACCACCACCACGCCGATCCCATGATCCCCATGATCCCCATGATCCCCTCAGAGCAACGCCATCCCATGATCCCATGATCCCCATGATCCCCATGATCCCACACGCCAACGCCAATGATCCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCCACGCCAATGATCCACGCCAATGATCCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCACGCCAATGATCCAATGATCAATGATCCACGCAATGATCCACGCCAATGATCAA

23.2.0 6330 <0.8 C2 \$156 81.60 **41 8**3 6333 8126 &1 QG RKTF90LMKRODLP9WLXIRVAAEGINYAORR#

\$220 6230 82.80 6250 6250 8270 6280 F O C V K N N Q I L E E N V E V E I W Y K E C E K K K K K P R W

> NS4A 6370 - 6380 6950 6350 \$320 6330 6340 33.36 \$310 8466 LOARTYSOPIALXEPXEPAAGRKSLTINLITEMG

8470 5450 8.590 6430 8440 6450 64.80 5.4.90 8426 OUTADDAR, AT DOUGLARD DE TORODAGED ADAD COTTRECT THAT AND A DAT DUALARAM. DUARD THAT THE AT TO CAN DUT COARD C REFTFMTQKARDALDSLAVERTAEAGGRAYNBA

6930 6546 8550 6380 6576 8510 \$520 6530 62.60 6800 TUTCRUTGASCTECCOSGAGACOCTEGASSCATTECTTTACTOACACOTICTOECTACACTOACGEGAGESSTCTTTTTATTCTTOATEACCOCCAGESSC L S E L P E T L X T L L X T V T G G I P L F L Ø S **A**R S

#### D2 PDK-53 specific NS4A-75-Ala (wt D2 16681: Gly, nt-6599-G)

6010 6020 650 650 650 6560 6560 6560 6570 6580 6590 700 RTROFCRAGETSCOCOPERATORISTIC TROFCOLOGICACORDINAL TROFCRACTICACOPTICAL AND TROFCOLOGICACORDINAL AND TROFCOLOGICACO I S E M I E O M O C I I E A E I C A E I C O O M O J T M E A E I L

6780 6760 8770 6710

8730 6740 \$796 6860 \$726 6780

TO SECTIFIC TO A TAGET TO TO TRATO CRORE AS CASE AND A CRORE AS CASES AND A CASES AND A CASES AND A CASES AND A

B F T E X V X E I P E P X K Q R T P Q G N Q E F R V V 3 3 2 2 2

> NS48 6820 6830 66622 5840 \$350 6860 \$187.0 6880 62.00 42900 

VAATMANEMGELEKTEKDEGIGSIATQOPESNI

## CARGERALAGENERATORSCATCESCATCHENESCECTOTESCONCENTRESECTATORSCATEGENERALAGENESCECTOTESCATEGENERALAGENESCECTOTESC L G I B L R F A S A M T I Y A V A T T F V F H L R H B I E N B S A V S S A V S S A V S S A V S S A V S S A V S S

7030 7040 7050 7050 7070 7080 NYSETAIANGATVEMOEGYBESKWOIGVPEE

73.60

ADEAD OF DESCRIPTION OF A A I S C Z S Q V N P I T L P A L P A L V A H Y A I I S P G L Q A

12 (2 12 C ALASCENCE AND A CONSTRUCTION OF A CONSTRUCTICONO KATHEACKRAARCIMENFIVOGITVIDEPIFYD

248.0 A&CONTAGOTACCOGGCCCATCTCCACATTETEGGCASGCARATCCACGCCGGTTTTGGCACCACT&CCATTGCGCTSTCAALCGCTAACGCCTACCATTTTTACACGC TEATUPIER LWEGNPGEPWNTTIAVS MANIPRG

#### > NS5

EYLACAGLIIGIM RNTINTRROTORISBILGBKW

7640 7650 7650 7670 7680 7630 7700 7€30 E G K L N A L G E E F Q E Y E K S G I Q E V G R Y L A K E G E E

#### Additional nt-7260 A-to-G silent mutation in master and pre-master seeds

7720 7730 7789 7776 7786 7740 7750 A REPERTIONATION OF THE ACCOUNT OF THE ACTOR AND A COUNT OF THE ACTOR AND THE ACTOR AND A COUNT ACTOR AND A 

I G C G R G G W S Y Y C G G I K N V R B V K G B T K G G F G B E B P

7930 7940 7980 7960 7970 7980 7996 ARTOTO TATA ASSOCIATE A CARACTERICAL AND A CONTRACTOR CONTRACTOR OF A CONTRACTOR CONTRACT IPMSTYSWNLVRLOSGVDVFFIPFBKCDTLLCD

8040 8050 8660 8070 8080 8020 8030 CATAGONI CAT I S B S P N P F V E A S P T L P V L N L V E N N N N R F O P C

82.76 83.20 83.80 83.90 A TRANSCE TO A COLORGE A COLORGE A CARACTER A REAL AND A CARACTER A COLORGE A REAL AND A CALL AND A C

I F V I N P Y M P S V I F K M F & I Q P K Y S G A I V X N F L S R 8

82.90 a circla catgaga to tacteo taccaste circla a catga catga catga catga catga catga catga catga catga cata cacatta caatga c

STHEMYWYSNASGNIV38VNMISRMLINRFTMR 838.6 

ATACANGANAGCOACTFACGAGCCOTATETFGACCFC5GA8GCGTACCCCFFACAFC5GGATGAGACTTSAGAACTAGAAATAAATFGGAAA Y K K A T T E E D Y O L O E C T R N I O I E S E I P E L O I I C K

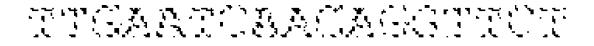
85.40 85.80 85.3.0 28.00 TARADADAS PARTICIDAD STOLEMENT AND STOLEMENT AND A STOLEMENT AND STOLEMENT AND STOLEMENT AND A STOLEMENT AND A

T G B A B M V N G V V R L L T E P W D V V P W V T G M A M T O T . 8630 8640 8650

8€10 TPEGQURVPKBEVDTETQEPKEGTKELMEIT& B

8720 8730 8740 8750 8760 8770 8780 8790 8850 

nonga	0108 AACATO					AAAAI		GAAG	8AGC				CALOG		an de la cale		97CAR		racas	iono:	
C E	a c	V	X .9	M	≪ S	£ 9	ξĒ	2.13	X I	. G	3	e c	A B.	ĸ	G S	R J	N I	W 1	Y X	20	Î.
NINGCA	S X COCIAC 8010					CERCE		AATC:	ANGR	ana c	acca.		NGAQA		100001		BAGEG		a gaz	0333	
CRAGO	9310 TAGGTI G X					GCAAC		AGGG.	AGGB				anow.		(ACIGA)		CASIDA		$A \cap A$	TAGA	
	9210 AAATGA	NGAJ	9226 AAC9337	aaca	9230 MCCA		924 MANGO			22.0 AGAR	ACTN	9256 AFOD6		9270 TETEK		92.80 (AACO)		.ee Wara			9300 3052
1. R	50 E.	Я;	M V	T I	× 8	M )	6	8	н н	C R	E. 1	A B	A 3	35	K E	.H. 1	2 Q	N B	< 9	V	R
ACEN	9310 Agacca R P					Arggi		AEATO	CGRÓ				3672AG		ASGE		.CTAT		rėbag	aca	
XIMATA	9410 TGSAG					TGGA		AAGO.	Aerc				NGCAC		MATC)				POTTON	(AAA)	
etrago	9510 AAGACH R V	996				aaga7	08000		asts		rear		FGTGA		12:13:63		NGGAT		49030	is rio	
erera.	9610 MATGAG N D			ATTA		GACA		CAAT	0930Å		TCAN		PGCAR		IGGACJ		rocee		389003	CAC	
CCA16	9710 AGTTAA	CA:	9720 198886	ACUG	9730 10606	TACIC	974 975 S7076	0 77200	3 8090	1750 IAGA&	AGCA	9750 89750	RACTO	977( ARTSK	) Firstan	978) 9006)	) ArtCit	37: COCM	70 NGGAC	) SCAGI	9890 9670
	L I	3.		19							12					988)		984 984			
nchan	9010 606068 R B					laren:		CCAA:	APG3				orroc		0.0003	vo en ca	NGGOT	0.000	SCAA7	ureo:	
00700	9910 GCAGTA X 9					ACAN		ACAN	CCTG				NAACA		13081(		2002aa		rocre	ACA:	
GANCA	0010 000707 V N	'GGA		asas/		GGARC		ACLAR	8ACC		TGOX		BRORG		KCCC&C		9909.A		SAAG	sacos	
rreces	0110 CTCAPI 3 L	6203		AACAJ		96072		ĠGOA	a.		COAA		NATAA		ATTEN				sÅtter	AGN/	
													>	3'-No	naodi	ing R	agion				
VCASAT	0210 7ACAR6 Y M	eca/		дада		AGAAC		<b>38</b> 08	8.A.GA		68A8		7:3017A	16270 GBAAG		1028) \07A3\		1030 (81)88			9396 4870
	0310 GATTAA		19320 Ataqta		10230 AAAAA		1034 90980			1359 2006		10360 Aggao		10970 AAGAI		1038) 9900A		103) XATGO			9490 PGAG
	0413 ATOCAS		19426 978607		16430 210A0		1044 657A3			456 XGGAG		10450 CAAAC		1647( A&GO)		0.04.8: 20470-		104: STSO			asac PTAC
	0510 COCCTC		10520 Facraa		10530 NGCAA		1054 33936			ISSO IGAGA		10560 20767:		1057( 06076		1058) ACTAS		1059 XAGAC			0600 0000
	0510 AAAAAC		19520 Atatto		19630 19995		1064 XCAGA			isso TGTC		10660 DAGCA		10676 CCAGC		1949) MACGO		106) AAATN			9790 9076



DENvax-4 Master Virus Seed (MVS)

**[0208]** Nucleotide sequence of the chimeric viral genome and deduced amino acid sequence of the translated protein. Most of the prM-E gene (nt-457 to -2379, underlined) is wild-type (wt) DEN-4 1036 virus-specific; the remaining nucleotide sequence is DEN-2 PDK-53 virus-specific. All engineered substitutions differ from wt virus (DEN-3 16562 or DEN-2 16681), as well as extra mutations (changes from engineered cDNA clone) are marked.

Substitutions Included in the Genome and Protein:

[0209] Junction sites:

- [0210] a. MluI (nt 451-456): engineered silent mutation, nt-453 A-to-G
- [0211] b. NgoMIV (nt 2380-2385): engineered mutations, nt-2381/2382 TG-to-CC (resulted in E-482 Valto-Ala change)
- [0212] D2 PDK-53 virus backbone (change from wt D2 16681)
  - **[0213]** a. 5'-noncoding region(NCR)-57 (nt-57 C-to-T): major attenuation locus (in red)
  - [0214] b. NS1-53 Gly-to-Asp (nt-2579 G-to-A): major attenuation locus (in red)
  - [0215] c. NS2A-181 Leu-to-Phe (nt-4018 C-to-T, in bold)
  - [0216] d. NS3-250 Glu-to-Val (nt-5270 A-to-T): major attenuation locus (in red)

- [0217] e. nt-5547 (NS3 gene) T-to-C silent mutation (in bold)
- [0218] f. NS4A-75 Gly-to-Ala (nt-6599 G-to-C, in bold)

**[0219]** \* nt-8571 C-to-T silent mutation of PDK-53 is not engineered in the vaccine virus

- [0220] Engineered substitutions in cDNA clone
- **[0221]** a. Engineered C-100 Arg-to-Ser (nt-396 A-to-C): may improve viral replication in culture
- [0222] b. Engineered nt-1401 A-to-G silent mutation
- **[0223]** c. Engineered E-364 Ala-to-Val (nt-2027 C-to-T): may improve viral replication in culture
- **[0224]** d. Engineered E-447 Met-to-Leu (nt-2275 A-to-C): may improve viral replication in culture
- **[0225]** Additional substitutions found in vaccine seed (0.06% nt different from original clone)
  - [0226] a. nt-225 (C gene) A-to-T silent mutation (in bold)
  - [0227] b. NS2A-66 Asp-to-Gly (nt-3674 A-to-G) mutation (in bold)
  - [0228] c. NS2A-99 Lys-to-Lys/Arg mix (nt-3773 A-to-A/G mix, in bold)
  - [0229] d. nt-5391 C-to-T (NS3 gene) silent mutation (in bold)
  - [0230] e. NS4A-21 Ala-to-Val (nt-6437 C-to-T, in bold)
  - [0231] f. nt-7026 T-to-C/T mix silent mutation (in bold)
  - [0232] g. nt-9750 A-to-C silent mutation (in bold)

				NCR-	57-1, D2	PDE-53 atte	mustion lo	was (wi Di	16681: C	}
> 51-	Noncoding	Region				ŝ				>C
	1.0	20	30	40	50	60	70	80	96	3.00
AGTES	FTAGTCTAC	aroosoorac	BAAGACADAI	TTTTERGG	ACCEAACCE	UNARCTACITY	TANCAGUTU	TTRATTRGAS	ROCKGATCE	ARTARES N M
	3.1.0	120	130	140	2.5.0	1.60	370	180	1.80	200

TOPOTOTOGRAPHAGESCEPTICALARSECTOTOTOGRAPHAGESCEPTICESC

 210
 220
 230
 250
 250
 270
 280
 290
 300

 TGCARTOCTCUNGGGACSAGGACCTTTTAAAACTEPTTTCUNGGCOTECTTCUTTCUTTCUTTCUTTCUATAACCEACGACAGESATATEGAAGAAG
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0
 0

#### Additional nt-225 A-to-T silent mutation in master and pre-master seeds

 310
 320
 350
 360
 370
 380
 390
 400

 PEGERAACABTIAANABAICANABCOINTTEATOTITIGACAGEGICAGEBAAGAGAITEGAAEGAICEINTEEADATOTICAAINEGAGACOCHCENTEE
 W
 G
 I
 K
 K
 K
 I
 N
 V
 L
 E
 F
 K
 E
 I
 C
 R
 M
 L
 R
 R
 P
 S
 S

Engineered C-100 Arg-to-Ser (nt 3%6 A-to-C)

 > prM
 Beginning of D4 1036 sequence

 616
 420
 420
 640
 950
 1660
 470
 980
 490
 500

 CAOSCATGAPTATOCTGAPTOCTACATGATAGOTOCCATTPAACCACGOSCAATACCCCCCAGATAGOTOCCATGATAGOTOCCATTPAACCACGOSCAATACCCCCCCAGATAGOTOCCAATAGOTOCCATGATAGOTOCCATGATAGOTOCCATGATAGOTOCCATGATAGOTOCCATGATAGOTOCCATGATAGOTOCCATGATAGOTOCCATGATAGOTOCCATGATAGOTOCCAGATAGOTOCCATGATAGOTOCCATGATAGOTOCCATGATAGOTOCCAGATAGOTOCCATGATAGOTOCCAGATGOTOCCAGATAGOTOCA

#### Engineered MluI splicing site (nt-453 A-to-G silent)

530 550 560 870 580 520 540 4.96 \$6.0 51.0 TBGINKC T L I A M D L G B M C V L L E K  $\mathcal{T}$  $B=D=\mathbb{P}$ T X X 580 620 630 640 85.0 660 670 630 TTACTNET/CANTACCEARCETSACACETECTECTECTECTECANTCHCACETCACETCECCATETATESCACETECTECCAGECEGACARCEGECANCEGECATETATESCACETECTECCAGECEGECANCEGECANCEGECATETATESCACETECTECCAGECEGECANCEGECCANCEGECCANCEGECANCEGECANCEGECANCEGECANCEGECANCEGECANCEGECANCEGECANCEGEC Q S G E P 2

> M 720 739 740 756 769 770 780 73.0 790 -860 880 8.60 370 830 840 850 9.0*P* 820 830 810 

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 CARTACATTERCERATATERCENTINGARAACTAGERATEREDACATERCERATERCENTER 2 Y I C R R 5 V V 5 & G W G N G C G L F G X G G V V F C A R F S C S

#### Silent nt-1401 A-to-G mutation in engineered clone

1.540 3,550 1510 1820 1530 1560 1970 1580 1890 1600 TITE CALE TO A CONTRACT TO A CONTRACT OF A CALE AND A C 1600 1640 1.650 1.560 1.690 1700 1620 1670 1680 1.610X V 1760 1750 1770 1780 1790 1720 1740 1860 1710 3730 

1810	1820	1830	1840	1850	1860	1870	1880	1890 1900
M E K L F		M S Y	T M C S		S I D K			CATGGGACAACAGTGG
			1 11 0 1	5 6 11 1	0 1 0 1		11 1 y	
1910	1920	1930	1940	1950	1960	1970	1980	1990 2000
TGAAAGTCAAGT		GGAGCTCCGI	GTAAAGTCCC	5 611111611611111	AGAGATGTGAA	ornroornnnn		GTATCATCTCATCCAC
K V K Y	EGA	GAPO	K V P	IEI	RDVN	KEK	VVGR	E I I S S T
2010	2020	2030	2040	2050	2060	2070	2080	2090 2100
								AAACAGTGCATTAACA
PLAE	NTNS		IEL	EPPE		YIV	G V G	NSALT
		I						
	Engineered	E-364 Ala	u-to-Val (r	nt-2027 C-t	co-T) to imp	rove vira	l growth in	. culture
2110	2120	2130	2140	2150	2160	2170	2180	2190 2200
								GAAACAGCTTGGGATT
	R K G S	S I G		ESTY	R G A K			ETAWDF
2210	2220	2230	2240	2250	2260	2270	2280	2290 2300
								GAGTCTCATGGATGAT
G S V G	GLF	TSLO	K V V	ΗQV	FGSV	Ϋ́Τ	L F G G	G V S W M I
				<b>F</b> ==		+ += T ===	 (~~~ 2275 )	
				Fugineer	eu 12-44/ Me	r-ro-rea	(IIL-22/5 A-	to-C) mutation
					End of D	4 1036 sea	quence	
2310	2320	2330	2340	2350	2360	2370	- 	2390 2400

 231C
 2320
 2330
 2340
 2350
 236C
 2370
 I
 2390
 2400

 TAGAATCCTAATTGGGTTCCTACTGTTGTGGATTGCCACGAACTCAAGGAACTCCAATGGCCTATGACGTGCATAGCT
 2390
 2400
 2400

 R
 I
 L
 I
 G
 F
 L
 V
 L
 W
 I
 G
 T
 N
 S
 R
 N
 T
 S
 M
 A
 M
 T
 C
 I
 A
 A
 G
 I
 V
 T
 L
 Y

#### Engineered NgoMTV splicing site, E-482 Val-to-Ala (nt-2381/2382 TG-to-CC)

 241C
 2420
 2430
 2440
 2450
 246C
 2470
 2480
 2490
 2500

 TTGGGAGCTCATGGTGCAGGCCGATAGTGGTTGCGTTGGGAGCTGGAAAACAAAGAACTGAAATGTGGCAGTGGGATTTTCATCACAGACAACGTGCAAA
 L
 V
 M
 V
 Q
 A
 D
 S
 G
 V
 V
 S
 W
 K
 N
 K
 E
 L
 K
 C
 G
 S
 G
 I
 F
 I
 T
 D
 N
 V
 H
 T

D2 PDE-53 N31-53-Asp attenuation locus (wt D2 16681: GIy, st-2579-6)

 291C
 2920
 2930
 2940
 2950
 296C
 2970
 2980
 2990
 3000

 TGGAGTATTCACCACCAACTATATGGCTAAAATGGAAAACAGGATGTATTCTGCGACCAAAACTCATGTCAGCGGCCAATAAAGACACGAGCC
 G
 V
 F
 T
 N
 I
 K
 L
 K
 E
 X
 Q
 D
 V
 F
 C
 D
 S
 K
 L
 K
 D
 N
 R
 A

301C 3020 3030 3040 3050 3060 3070 3080 3090 3100 GTCCATGCCGATATGGGTTATTGGATAGAAAGTGCACTCAATGACAATGGAAGATAGAGAAAGCCTCTTTCATGAAGTTAAAAACTGCCACTGGCCAA V H A D M G V W L F S A L N D T W K T F K A S F L F V K N C F W P K

VHADMGYWIESALNDTWKIEKASFIEVKNCHWPK

321C 322O 323O 324O 325O 326C 327O 328O 329O 330O CCATACACACAGAGGACCATGGCATCTAGGTAAGCTTGAGATGGACTTTCTGTGATGGAACAACAGTGGTAGTGACTCAGGACTGCGGAAAT

H T Q I T G P W H L G K L E M D F D F C D G T T V V V T E D C G N

331C 3320 3330 3340 3350 336C 3370 3380 3390 3400 AGAGGGACCCTCTTTJAGGAACAACCACTGCCTCTGGAAAACTCATAACAGAATGGTGCTGCCGATCTTGCACATTACCACCGCTAAGATACAGAGGTGAGG R G P S L R T T T A S G K L I T E W C C R S C T L P P L R Y R G E D

							> NS2A	7	
3410	3420	3430	3440	3450	3460	3470	3480	3490	3500

APSOFFOTSOTSOTSOTSOTAATUSAATUSAACAATUSAAGAAAAGAAAATTOGTUAACTOCTTUGTUAUSOTSGACATOGGUAUGTUGACAACTT 9 C W Y G M E I E F E K E K E K N E V N 3 E V T K G E 4 Q V U N F

2510 2520 2530 2540 2550 2540 2550 2550 2580 2590 2600 TECRETAGRAFTETESGARAFGCATEGEAGGARAFGCEEAGGACCCCAGEAGGARACAFGCAREACEEGAGEEGAGEEGAGGAAAFGCEETETETES S E G V E G M A E E L E B M E R E R V G E K E B I E L V A V S F V

#### Additional NS2A-66 Asp-to-Gly (nt-3674 A-to-G mutation) in master and pre-master seeds

#### Additional NS2A-99 K to R/K (mix) (nt-3773 A-to-G/A) mutation in master seed

 3810
 3830
 3840
 3850
 3870
 3880
 3890

 TOTRETOCTECCOAGAGEACORTECTHAGETGAUGESTGA

 4010
 4020
 6020
 4040
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 4060
 <th

#### D2 PDK-53 specific NS2A-181-Phe (wt D2 16681: Leu, nt-4018-C)

> NS2B

4310 4520 4330 5340 4350 4560 4370 4380 4590 4400 Cosmistiannissianscastasiasintasionasiticanticiteitaatakanatkaanatkasia D V K W E D O A E E 3 O E 3 P I L 3 I T I S E D O E M 3 I K N E E

> SR3 
6510 4520 4530 4540 4550 4560 4580 4580 4600
1005AADTGAADAADAGADGADGATGTTCOTCCACCCACCCATGGGAAGATGAACTGAGATGAACCCTATAGAATTAGGC
1005AADTGAACAACACCGCCACCCACCCCACCCCACGGAAGATGAACTGAGATGAACTAGGAATTAGGC
1005AADTGAACAACACCGCCACCCACCCCACGCGAAGATGAACTGAGATGAACCCCAATTAGGAATTAGGC
1005AADTGAACAACACCGCCACCCACCCCACGCGAAGATGAACTGAGATGAACTGAACCCCAATTAGGAATTAGGCCCAATTAGGAATTAGATTAGAATTAGAATTAGAATTAGATTAGATTAGATTAGATTAGATTAGATTAGAT

4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 GRAGTOTAGGENTEGERTEGERGENTEGERATAGENCERCEGERGENTETTECERACECEGERECERTEGERT E V O V L A E E P O K N P P A V O T K P O L F K T N A O T I O A V S

 4910
 4920
 4930
 5940
 4950
 4960
 4970
 4980
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 5900
 <th

 5210
 5220
 5230
 5240
 5250
 5270
 5280
 5290
 5300

 INSCRIPTIONALISERSEANSCOTTAGASSITUTCCALTACEASACCCCAGOCENTCASACCTEREACC

02 PDE-83 NB3-250-Val attenuation lonus (D2 16681: Olu, nt-5270-A)

:

 5310
 5320
 5340
 5350
 5370
 5380
 5890

 GIGTOADOCRAPTIASCASEGOCOCTAICACEGOTACEACEACEACCICATEATOEACEAAGCOCATTEACEACEAEECAEECEAECCAEECEAECCAEECEAECCAECEAEECEAECCAEECEAECCAECEAEECEAECCAECEAEECEAECCAECEAEECEAECEAEECEAECEAEECEAECEAEECEAECEAEECEAECEAEECEAECEAECEAEECEAECEAEECEAE

#### Additional nt-5391 C-to-T silent mutation in mater and pre-master seeds

5510 5520 5539 5560 5550 5560 5570 5580 5590 5600 AGAGGAALGGALCAALGAAGAAGAAGAGAAGAGGACCILGGGAALGGGALGAGGGALGAGGGALGAGGGALGAGGGALGAGGGALGAGGGALGAGGGALGAGGGALGAGG S N & P E I D E E R E E P E K S K S G E E W V P 5 E E G R P V M

#### D2 PDK-53 specific silent mutation nt-5547-C (D2 16681: E)

5910 5920 5930 5940 5950 5960 5970 5980 8990 (000 BATASGARBAATCABABAATCBCAGTACCAGTECTATACATCEGCGABACCTGTEGARAATCBTSGACCTGTEGARAACBACCTGTBCAGACCTGCABAGT I G R N P K N B N 5 Q X I Y M G E F L E N D E 0 C A E W F E A K M

6110 6120 5130 6140 5150 5160 8170 6180 5190 5200 GREARGCARGGAAAAOCTTTGTREACTARTRAGRAGAGGGAGAOCTRECAGTTFGTTEGOCTAGAGAGFGGCBACGGURTCARCTREGCRGAGAG B A R R T P V D L M R R G D L P V N L A Y R V A A E G I N Y A D R

> NS4A

C310 5320 5330 E340 5350 5350 5370 E380 5380 5400 AGATGGTTGGATGTTGGACCCACIGGCSCTAAAGAATTARGGAATTGCGGCGGGAGBABGTCTCTGACCCTGACCTAATCACAG R N E D A R I X S D P E A E K E K B E F A & G R B G L T E N E I T E

#### Additional NS4A-21 Ala-to-Val (nt-6437 C-to-T) mutation in mater and pre-master seeds

#### D2 FDK-53 specific NS4A-75-Ala (wt D2 16681: Gly, nt-6599-G)

ILEFFLIVLLIPEFEKQRTPQDNQLTYVIAIL

#### > NS4B

 6810
 6820
 6830
 6810
 6820
 6820
 6820
 6820
 6820
 6900

 Characteresconcerascarescerasseries
 Characterescencerascares
 Characterescencerascares
 6820
 6820
 6900

 F V V A A T M A N E M G F L E K T X K D L G L G S I A T Q Q P E S

5210 5220 5330 5340 5350 5360 5360 5370 5280 5390 7000 MACATCOPSUACATMANCTACATCOPSCATCASCATUSACCOPSCACCACATUSATACCATUSACATUSACATUSACATUSACATUSACATUSACATUSACATUSACATUS N I L D I D L R P A S A W P L X A V A P P F V P F M L R H B I E N B

7015 7020 7039 7040 7050 7050 7070 7080 7090 7100 CCTCAGTSAAISTFTCCCTAACASCCATAGCCAACCAAGCCACTSTTAATGGGTCTSGGGAAAGCATGGCCATTFTCAAAGATGGACATCGGAGTCCC S V N V S E 2 A 3 B N Q A 2 V 2 M G 2 G 8 0 W P L S N M D 3 G V F

#### Additional nt-7026 T-to-C/T mix silent mutation in master and pre-master seeds

7410742074307460745574607470748074907550GUCTTTABCCTESCOGCOCCALTICCCALATEGEGABGGABGGABGCARTCCACGCGAGGTTTTEGGABCGACCTCCACGCGAGGTTTEGGABCGACCTCCACGCGAGGTTTEGGABCGACCTCCACTECACTECCACTECACTECCACTECACTECCACTEC

#### > NSS

7610 7620 7530 7640 7650 7660 7670 7680 7690 7750 Debaatgabaaccoebiigabccatiggababgigabiicagbigtofbcabgbabgggabiicagbagtgalagbaccitac/babagbagg K W K G K E N A E G K S B P Q 3 X K X G E Q B V G K I L A K E G

7910 7820 7830 7340 7850 7860 7870 7880 7890 7890 7900 GIGGEOTOCGITUTGCAMBAGGAGGITGTIGTACIATISTIGGAGGACTBARGASUSTRAGASUSASGICCIABLAGAGGACUNGGACUNG V D L G C G R G G W S I Y C G G L F N V R B V K G L T K G G F G B E

8510 8520 8530 8540 8550 8560 8570 8580 8590 8600 CAANACAGACTEGATCATCETCETCETCACAGAGTGCTCACAGAAACTETCEGATCTCCTCCTCACAGAGTGCTACTEGATGAC R C T C S A S S M V R C V V R L L T K F W D V V D V V T C M A M T

8810 8820 8830 8840 8850 8860 8870 8880 8990 8990 8990 CULTATICACINATIANANANINGTONICACINGASIOUTICARIANINGPUNTINGANINGTONIALINAANINAACINATICUTUATI E Y L E N K W K S A R B A V E D E R E W E I V L X B P N I B L E

9010 9020 9030 9050 9050 9050 9070 9030 9030 9100 TOGOTTOGREGACOCITOTIAGRETICARCOCITAGENTOTIARITGARCACICCOTGACAGACOCOTGACIOGACOCAGAGACA W L G & P F L B P E & L G F L X E D B W F S E E N S L S G V E G E G

9210 9220 9230 9240 9250 9260 9270 9280 9290 9300 MTRACACCTRARARATCARCTRATERACCTRCTTCCARCCARCERCOCCTTTTCCARACCTRCTRATCARCCARCARCESCT B D L K N E B M V T N B M B G E H K K L A B A I F K L T X Q N K V

9610 9620 9639 9640 9650 9660 9660 9680 9680 9690 9700 TIBACASCITTAATSECATSGGAAGATTEGGAAGECKTACAECAATGGGAACTTCAEGAGEETEGAETGATTSGACACAAGTOCOCTTUIGITUSC L T & L N D M G K I R K D I Q Q W E F 8 8 G W R G W T O V F F C 5 8

9710 9720 9730 9740 9750 9760 9770 9780 9790 9800 ACCATETCCATEGEGITARTCATEGEAGOGICSCSTACTOGITYTTCCCTYTASAAACCAAGATERACTSATEGECAGOSCCCSAATCICCCAAGSASC BFHELIMELGEVLVVPCBNODELICEARTECCCA

#### Additional nt-9750 A-to-C silent mutation in master and pre-master seeds

10010 10020 10020 10040 10050 10060 10070 10080 10090 10100 CASTOTIGABCAGGGTOTOCATCARGABCCCATEGATGEBAGACBABACTOCAGTOGABTCATEGGAGEBABACCCATEGGABABAGAGAAA V W N R V W I Q E N P W M R D K I P V B 8 W B B I F Y L G K R R D

10110 10120 10130 10140 10150 10160 10170 10180 10190 10200 CCAATGEGGEGCTCATEGEGTIACCAAGGGCGEGCGEGCGAAGAACACCGAAGGAAGAACAACGAAGTAAGGCAAGGAAGAAGAA Q X C G S E I S E T S R A T N A K N I Q A A I N Q V R E I I G N B

#### > 3'-Noncoding Region

10510 10320 10550 10540 10350 10360 10370 10580 10390 10400 CANSTCADEPOSATERADCORESTACCEREDACCECECCOECCECCECEREDACCEEERAASAN PERSOCCECECEREDACCEEERA

10410	10420	10430	10440	10450	10460	10470	10480	10490	10500
CTTGAGTAAACTA	TGCAGCCTG	AGCTCCACCI	GAGAAGGTG	AAAAAATCC	GGGAGGCCACA	AACCATGGAA	AGCTGTACGCA	TGGCGTAGT	GGACTAGC
10510	10500	10500	10540	10550	1.05.60	10570	10500	10500	10000
10510	10520	1 ()())()	10540	10550	10560	10570	10580	10590	10600
GGTTAGAGGAGAC	LCCTCCCTT4	CAAATCGCAU	JCAACAATGGU	GGUUUAAGGU	GAGATGAAGC	TGTAGTCTU	JUTGGAAGGAU	TAGAGGTTAU	JAGGAGAC
10610	10620	10630	10640	10650	10660	10670	10680	10690	10700
CCCCCCGAAACAA	AAAACAGCA	ATTGACGCTO	GGAAAGACCA	GAGATCCTG	CTGTCTCCTCZ	AGCATCATTCO	CAGGCACAGA	ACGCCAGAAAA	A-GGAATG

10710 10720 GTGCTGTTGAATCAACAGGTTCT SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 16 <210> SEQ ID NO 1 <211> LENGTH: 10723 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Denque virus serotype 1, BVS <400> SEQUENCE: 1 60 agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta 120 gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg 180 aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg 240 300 gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt 360 ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat cattatgctg 420 attccaacag tgatggcgtt ccatttaacc acgcgtgggg gagagccgca tatgatagtt 480 agcaagcagg aaagaggaaa gtcacttttg ttcaagacct ctgcaggtgt caacatgtgc 540 acceteattg egatggattt gggagagttg tgtgaggaea egatgaeeta eaaatgeeee 600 cggatcactg aggcggaacc agatgacgtt gactgttggt gcaatgccac ggacacatgg 660 gtgacctatg gaacgtgctc tcaaactggc gaacaccgac gagacaaacg ttccgtcgca 720 780 ttggccccac acgtggggct tggcctagaa acaagagccg aaacgtggat gtcctctgaa ggtgcttgga aacagataca aaaagtagag acttgggctc tgagacatcc aggattcacg 840 gtgatageee tttttetage acatgeeata ggaacateea teaceeagaa agggateatt 900 ttcattttgc tgatgctggt aacaccatct atggccatgc gatgcgtggg aataggcaac 960 agagacttcg tggaaggact gtcaggagca acatgggtgg atgtggtact ggagcatgga 1020 agttgcgtca ccaccatggc aaaaaacaaa ccaacactgg acattgaact cttgaagacg 1080 gaggtcacaa accctgcagt tctgcgtaaa ttgtgcattg aagctaaaat atcaaacacc 1140 accaccgatt cgagatgtcc aacacaagga gaagccacac tggtggaaga acaagacgcg 1200 aactttgtgt geegaegaae gttegtggae agaggetggg geaatggetg tgggetatte 1260 ggaaaaggta gtctaataac gtgtgccaag tttaagtgtg tgacaaaact agaaggaaag 1320 1380 atagttcaat atgaaaacct aaaatattca gtgatagtca ccgtccacac tggagatcag caccaggtgg gaaatgagac tacagaacat ggaacaactg caaccataac acctcaagct 1440 cctacqtcqq aaatacaqct qaccqactac qqaaccctta cattaqattq ttcacctaqq 1500 acagggctag attttaacga gatggtgttg ctgacaatga aagaaagatc atggcttgtc 1560 1620 cacaaacaat ggttcctaga cttaccactg ccttggacct ctggggcttc aacatcccaa gagacttgga acagacaaga tttactggtc acatttaaga cagctcatgc aaagaagcag 1680 gaagtagtcg tactaggatc acaagaagga gcaatgcaca ctgcgctgac tggagcgaca 1740 gaaatccaaa cgtcaggaac gacaacaatt ttcgcaggac acctaaaatg cagactaaaa 1800 atggacaaac taactttaaa agggatgtca tatgtgatgt gcacaggctc attcaagtta 1860 gagaaagaag tggctgagac ccagcatgga actgttctgg tgcaggttaa atatgaagga 1920

-continued	
acagacgcac catgcaagat teeettteg acceaagatg agaaaggage aacceagaat	1980
gggagattaa taacagccaa ccccatagtc actgacaaag aaaaaccagt caatattgag	2040
gcagaaccac cctttggtga gagctacatc gtggtaggag caggtgaaaa agctttgaaa	2100
ctaagctggt tcaagaaagg aagcagcata gggaaaatgt ttgaagcaac tgcccgagga	2160
gcacgaagga tggccattct gggagacacc gcatgggact tcggttctat aggaggagtg	2220
ttcacgtcta tgggaaaact ggtacaccag gtttttggaa ctgcatatgg agttttgttt	2280
agcggagttt cttggaccat gaaaatagga atagggattc tgctgacatg gctaggatta	2340
aattcaagga acacgtccct ttcgatgatg tgcatcgcag ccgccattgt gacactgtat	2400
ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg	2460
aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag	2520
ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac	2580
atttgtggaa teegeteagt aacaagaetg gagaatetga tgtggaaaca aataacaeca	2640
gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc	2700
aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat	2760
tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg	2880
gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa	2940
aaacaggatg tattetgega etcaaaaete atgteagegg eeataaaaga eaacagagee	3000
gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag	3060
aaageetett teattgaagt taaaaaetge eaetggeeaa aateacaeae eetetggage	3120
aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa	3180
cacaactata gaccaggeta ceataeacaa ataacaggae catggeatet aggtaagett	3240
gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat	3300
agaggaccet etttgagaae aaceaetgee tetggaaaae teataaeaga atggtgetge	3360
cgatettgea cattaceace getaagatae agaggtgagg atgggtgetg gtaegggatg	3420
gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga	3480
catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa	3540
atgettagga eeegagtagg aacgaaacat geaataetae tagttgeagt ttettttgtg	3600
acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc	3660
gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagcette	3720
aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg	3780
atgatgacta ctataggaat tgtactcctc tcccagagca ccctaccaga gaccattctt	3840
gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa	3900
aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta	3960
caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc	4020
ttaacateet cacageaaaa aacagattgg ataceattag cattgaegat caaaggtete	4080
aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca	4140
ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa	4200

	-continued	
aatgatattc ccatgacagg accattagtg gc	tggagggc teeteactgt gtgetaegtg	4260
ctcactggac gatcggccga tttggaactg ga	gagagcag ccgatgtcaa atgggaagac	4320
caggcagaga tatcaggaag cagtccaatc ct	gtcaataa caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga ggaagatcaa ac	actgacca tactcattag aacaggattg	4440
ctggtgatct caggactttt tcctgtatca at	accaatca cggcagcagc atggtacctg	4500
tgggaagtga agaaacaacg ggccggagta tt	gtgggatg tteetteace eccaeceatg	4560
ggaaaggctg aactggaaga tggagcctat ag	aattaagc aaaaagggat tettggatat	4620
tcccagatcg gagccggagt ttacaaagaa gg	aacattcc atacaatgtg gcatgtcaca	4680
cgtggcgctg ttctaatgca taaaggaaag ag	gattgaac catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg aggaggctgg aa	gttagaag gagaatggaa ggaaggagaa	4800
gaagteeagg tattggeaet ggageetgga aa	aaatccaa gagccgtcca aacgaaacct	4860
ggtcttttca aaaccaacgc cggaacaata gg	tgetgtat etetggaett tteteetgga	4920
acgtcaggat ctccaattat cgacaaaaaa gg	aaaagttg tgggtcttta tggtaatggt	4980
gttgttacaa ggagtggagc atatgtgagt gc	tatagccc agactgaaaa aagcattgaa	5040
gacaacccag agatcgaaga tgacattttc cg	aaagagaa gactgaccat catggacctc	5100
cacccaggag cgggaaagac gaagagatac ct	tccggcca tagtcagaga agctataaaa	5160
cggggtttga gaacattaat cttggccccc ac	tagagttg tggcagctga aatggaggaa	5220
gcccttagag gacttccaat aagataccag ac	ecccageca teagagetgt geacaeeggg	5280
cgggagattg tggacctaat gtgtcatgcc ac	atttacca tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct gattatcatg ga	cgaagccc atttcacaga cccagcaagt	5400
atagcagcta gaggatacat ctcaactcga gt	ggagatgg gtgaggcagc tgggattttt	5460
atgacageca etcecceggg aageagagae ee	attteete agageaatge accaateata	5520
gatgaagaaa gagaaatccc tgaacgctcg tg	gaatteeg gacatgaatg ggteaeggat	5580
tttaaaggga agactgtttg gttcgttcca ag	tataaaag caggaaatga tatagcagct	5640
tgcctgagga aaaatggaaa gaaagtgata ca	actcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac tt	cgtggtta caactgacat ttcagaaatg	5760
ggtgccaatt tcaaggctga gagggttata ga	ecccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattctg gc	aggaccta tgccagtgac ccactctagt	5880
gcagcacaaa gaagagggag aataggaaga aa	tccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aaatgatgaa ga	ctgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga at	catteeta geatgttega accagagegt	6060
gaaaaggtgg atgccattga tggcgaatac cg	cttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gt	ctggttgg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aaggtggtgt tt	tgatggag tcaagaacaa ccaaatccta	6240
gaagaaaacg tggaagttga aatctggaca aa	agaagggg aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat ctattctgac cc	actggcgc taaaagaatt taaggaattt	6360
gcagccggaa gaaagtctct gaccctgaac ct	aatcacag aaatgggtag gctcccaacc	6420
ttcatgactc agaaggcaag agacgcactg ga	caacttag cagtgctgca cacggctgag	6480

-continued	
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc	6900
aacateetgg acatagatet aegteetgea teageatgga egetgtatge egtggeeaca	6960
acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta	7020
acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca toggagttoo oottotogoo attggatgot actoacaagt caacocoata	7140
actctcatag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca	7260
actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc gaagtttgaa	7320
aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt	7500
agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga	7740
ggeteageaa aaetgagatg gttegttgag agaaacatgg teacaeeaga agggaaagta	7800
gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa	8040
gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa	8100
ttttgcataa aggttctcaa cccatatatg ccctcagtca tagaaaaaat ggaagcacta	8160
caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag	8220
atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg	8280
atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac	8340
ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt	8400
gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac	8460
cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca	8520
tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg	8580
gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag	8640
aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa	8760

				-001011			 
gaattcacaa	gaaaggtgag	aagcaatgca	gccttggggg	ccatattcac	tgatgagaac	8820	
aagtggaagt	cggcacgtga	ggctgttgaa	gatagtaggt	tttgggagct	ggttgacaag	8880	
gaaaggaatc	tccatcttga	aggaaagtgt	gaaacatgtg	tgtacaacat	gatgggaaaa	8940	
agagagaaga	agctagggga	attcggcaag	gcaaaaggca	gcagagccat	atggtacatg	9000	
tggcttggag	cacgcttctt	agagtttgaa	gccctaggat	tcttaaatga	agatcactgg	9060	
ttctccagag	agaactccct	gagtggagtg	gaaggagaag	ggctgcacaa	gctaggttac	9120	
attctaagag	acgtgagcaa	gaaagaggga	ggagcaatgt	atgccgatga	caccgcagga	9180	
tgggatacaa	gaatcacact	agaagaccta	aaaaatgaag	aaatggtaac	aaaccacatg	9240	
gaaggagaac	acaagaaact	agccgaggcc	attttcaaac	taacgtacca	aaacaaggtg	9300	
gtgcgtgtgc	aaagaccaac	accaagaggc	acagtaatgg	acatcatatc	gagaagagac	9360	
caaagaggta	gtggacaagt	tggcacctat	ggactcaata	ctttcaccaa	tatggaagcc	9420	
caactaatca	gacagatgga	gggagaagga	gtctttaaaa	gcattcagca	cctaacaatc	9480	
acagaagaaa	tcgctgtgca	aaactggtta	gcaagagtgg	ggcgcgaaag	gttatcaaga	9540	
atggccatca	gtggagatga	ttgtgttgtg	aaacctttag	atgacaggtt	cgcaagcgct	9600	
ttaacagctc	taaatgacat	gggaaagatt	aggaaagaca	tacaacaatg	ggaaccttca	9660	
agaggatgga	atgattggac	acaagtgccc	ttctgttcac	accatttcca	tgagttaatc	9720	
atgaaagacg	gtcgcgtact	cgttgttcca	tgtagaaacc	aagatgaact	gattggcaga	9780	
gcccgaatct	cccaaggagc	agggtggtct	ttgcgggaga	cggcctgttt	ggggaagtet	9840	
tacgcccaaa	tgtggagctt	gatgtacttc	cacagacgcg	acctcaggct	ggcggcaaat	9900	
gctatttgct	cggcagtacc	atcacattgg	gttccaacaa	gtcgaacaac	ctggtccata	9960	
catgctaaac	atgaatggat	gacaacggaa	gacatgctga	cagtctggaa	cagggtgtgg	10020	
attcaagaaa	acccatggat	ggaagacaaa	actccagtgg	aatcatggga	ggaaatccca	10080	
tacttgggga	aaagagaaga	ccaatggtgc	ggctcattga	ttgggttaac	aagcagggcc	10140	
acctgggcaa	agaacatcca	agcagcaata	aatcaagtta	gatcccttat	aggcaatgaa	10200	
gaatacacag	attacatgcc	atccatgaaa	agattcagaa	gagaagagga	agaagcagga	10260	
gttctgtggt	agaaagcaaa	actaacatga	aacaaggcta	gaagtcaggt	cggattaagc	10320	
catagtacgg	aaaaaactat	gctacctgtg	ageccegtee	aaggacgtta	aaagaagtca	10380	
ggccatcata	aatgccatag	cttgagtaaa	ctatgcagcc	tgtagctcca	cctgagaagg	10440	
tgtaaaaaat	ccgggaggcc	acaaaccatg	gaagctgtac	gcatggcgta	gtggactagc	10500	
ggttagagga	gacccctccc	ttacaaatcg	cagcaacaat	gggggcccaa	ggcgagatga	10560	
agctgtagtc	tcgctggaag	gactagaggt	tagaggagac	ccccccgaaa	caaaaaacag	10620	
catattgacg	ctgggaaaga	ccagagatcc	tgctgtctcc	tcagcatcat	tccaggcaca	10680	
gaacgccaga	aaatggaatg	gtgctgttga	atcaacaggt	tct		10723	

<210> SEQ ID NO 2 <211> LENGTH: 3391 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Dengue virus serotype 1, BVS

<400> SEQUENCE: 2

Met 1	Asn	Asn	Gln	Arg 5	Lys	Lys	Ala	Lys	Asn 10	Thr	Pro	Phe	Asn	Met 15	Leu
Lys	Arg	Glu	Arg 20	Asn	Arg	Val	Ser	Thr 25	Val	Gln	Gln	Leu	Thr 30	Lys	Arg
Phe	Ser	Leu 35	Gly	Met	Leu	Gln	Gly 40	Arg	Gly	Pro	Leu	Lys 45	Leu	Phe	Met
Ala	Leu 50	Val	Ala	Phe	Leu	Arg 55	Phe	Leu	Thr	Ile	Pro 60	Pro	Thr	Ala	Gly
Ile 65	Leu	Lys	Arg	Trp	Gly 70	Thr	Ile	Lys	Lys	Ser 75	Lys	Ala	Ile	Asn	Val 80
Leu	Arg	Gly	Phe	Arg 85	Lya	Glu	Ile	Gly	Arg 90	Met	Leu	Asn	Ile	Leu 95	Asn
Arg	Arg	Arg	Arg 100	Ser	Ala	Gly	Met	Ile 105	Ile	Met	Leu	Ile	Pro 110	Thr	Val
Met	Ala	Phe 115	His	Leu	Thr	Thr	Arg 120	Gly	Gly	Glu	Pro	His 125	Met	Ile	Val
Ser	Lys 130	Gln	Glu	Arg	Gly	Lys 135	Ser	Leu	Leu	Phe	Lys 140	Thr	Ser	Ala	Gly
Val 145	Asn	Met	Сүз	Thr	Leu 150	Ile	Ala	Met	Asp	Leu 155	Gly	Glu	Leu	Суз	Glu 160
	Thr	Met	Thr	Tyr 165	Lys	Сүз	Pro	Arg	Ile 170		Glu	Ala	Glu	Pro 175	
Asp	Val	Asp	Cys 180		Сув	Asn	Ala	Thr 185		Thr	Trp	Val	Thr 190		Gly
Thr	Суз	Ser 195		Thr	Gly	Glu	His 200		Arg	Asp	Lys	Arg 205		Val	Ala
Leu	Ala 210		His	Val	Gly	Leu 215		Leu	Glu	Thr	Arg 220		Glu	Thr	Trp
Met 225		Ser	Glu	Gly	Ala 230		Lys	Gln	Ile	Gln 235		Val	Glu	Thr	Trp 240
	Leu	Arg	His	Pro 245	Gly	Phe	Thr	Val	Ile 250		Leu	Phe	Leu	Ala 255	
Ala	Ile	Gly	Thr 260		Ile	Thr	Gln	Lys 265		Ile	Ile	Phe			Leu
Met	Leu			Pro	Ser	Met			Arg	Cys	Val	-	270 Ile	Gly	Asn
Arg		275 Phe	Val	Glu	Gly		280 Ser	Gly	Ala	Thr		285 Val	Asp	Val	Val
	290 Glu	His	Gly	Ser	Суа	295 Val	Thr	Thr	Met		700 200	Asn	Lys	Pro	
305 Leu	Asp	Ile	Glu	Leu	310 Leu	Lys	Thr	Glu	Val	315 Thr	Asn	Pro	Ala	Val	320 Leu
	-			325	Glu	-			330					335	
			340					345					350		
-	-	355			Gly		360					365		-	
Asn	Phe 370	Val	Суз	Arg	Arg	Thr 375	Phe	Val	Asp	Arg	Gly 380	Trp	Gly	Asn	Gly
Cys 385	Gly	Leu	Phe	Gly	Lys 390	Gly	Ser	Leu	Ile	Thr 395	Суз	Ala	Lys	Phe	Lys 400
Суа	Val	Thr	Lys	Leu	Glu	Gly	ГЛа	Ile	Val	Gln	Tyr	Glu	Asn	Leu	rÀa

-continued

				405					410					415	
Tyr	Ser	Val	Ile 420	Val	Thr	Val	His	Thr 425	Gly	Asp	Gln	His	Gln 430	Val	Gly
Asr	Glu	Thr 435	Thr	Glu	His	Gly	Thr 440	Thr	Ala	Thr	Ile	Thr 445	Pro	Gln	Ala
Pro	Thr 450	Ser	Glu	Ile	Gln	Leu 455	Thr	Asp	Tyr	Gly	Thr 460	Leu	Thr	Leu	Asp
Суа 465	Ser	Pro	Arg	Thr	Gly 470	Leu	Asp	Phe	Asn	Glu 475	Met	Val	Leu	Leu	Thr 480
Met	ГЛа	Glu	Arg	Ser 485	Trp	Leu	Val	His	Lys 490	Gln	Trp	Phe	Leu	Asp 495	Leu
Pro	Leu	Pro	Trp 500	Thr	Ser	Gly	Ala	Ser 505	Thr	Ser	Gln	Glu	Thr 510	Trp	Asn
Arg	Gln	Asp 515	Leu	Leu	Val	Thr	Phe 520	Lys	Thr	Ala	His	Ala 525	Lys	Lys	Gln
Glu	. Val 530	Val	Val	Leu	Gly	Ser 535	Gln	Glu	Gly	Ala	Met 540	His	Thr	Ala	Leu
Thr 545	Gly	Ala	Thr	Glu	Ile 550	Gln	Thr	Ser	Gly	Thr 555	Thr	Thr	Ile	Phe	Ala 560
Gly	His	Leu	Lys	Сув 565	Arg	Leu	Lys	Met	Asp 570	Lys	Leu	Thr	Leu	Lys 575	Gly
Met	Ser	Tyr	Val 580	Met	Сүв	Thr	Gly	Ser 585	Phe	Lys	Leu	Glu	Lys 590	Glu	Val
Ala	. Glu	Thr 595	Gln	His	Gly	Thr	Val 600	Leu	Val	Gln	Val	Lys 605	Tyr	Glu	Gly
Thr	Asp 610	Ala	Pro	Суз	Lys	Ile 615	Pro	Phe	Ser	Thr	Gln 620	Asp	Glu	Lys	Gly
Ala 625	Thr	Gln	Asn	Gly	Arg 630	Leu	Ile	Thr	Ala	Asn 635	Pro	Ile	Val	Thr	Asp 640
Lys	Glu	Lys	Pro	Val 645	Asn	Ile	Glu	Ala	Glu 650	Pro	Pro	Phe	Gly	Glu 655	Ser
Tyr	Ile	Val	Val 660	Gly	Ala	Gly	Glu	Lys 665	Ala	Leu	Lys	Leu	Ser 670	Trp	Phe
Lys	Lys	Gly 675	Ser	Ser	Ile	Gly	Lys 680	Met	Phe	Glu	Ala	Thr 685	Ala	Arg	Gly
Ala	Arg 690	Arg	Met	Ala	Ile	Leu 695	Gly	Asp	Thr	Ala	Trp 700	Asp	Phe	Gly	Ser
Il∈ 705	Gly	Gly	Val	Phe	Thr 710	Ser	Met	Gly	Lys	Leu 715	Val	His	Gln	Val	Phe 720
Gly	Thr	Ala	Tyr	Gly 725	Val	Leu	Phe	Ser	Gly 730	Val	Ser	Trp	Thr	Met 735	Lys
Ile	Gly	Ile	Gly 740	Ile	Leu	Leu	Thr	Trp 745	Leu	Gly	Leu	Asn	Ser 750	Arg	Asn
Thr	Ser	Leu 755	Ser	Met	Met	Суз	Ile 760	Ala	Ala	Ala	Ile	Val 765	Thr	Leu	Tyr
Leu	. Gly 770	Val	Met	Val	Gln	Ala 775	Asp	Ser	Gly	Суз	Val 780	Val	Ser	Trp	Lys
Asr 785	Lys	Glu	Leu	Lys	Cys 790	Gly	Ser	Gly	Ile	Phe 795	Ile	Thr	Asp	Asn	Val 800
	Thr	Trp	Thr			Tyr	Lys	Phe			Glu	Ser	Pro		
				805					810					815	

-	СО	nt	in	ue	d

Leu	Ala	Ser	Ala 820	Ile	Gln	Lys	Ala	His 825	Glu	Glu	Asp	Ile	Сув 830	Gly	Ile
Arg	Ser	Val 835	Thr	Arg	Leu	Glu	Asn 840	Leu	Met	Trp	ГЛЗ	Gln 845	Ile	Thr	Pro
Glu	Leu 850	Asn	His	Ile	Leu	Ser 855	Glu	Asn	Glu	Val	Lys 860	Leu	Thr	Ile	Met
Thr 865	Gly	Aap	Ile	ГЛа	Gly 870	Ile	Met	Gln	Ala	Gly 875	ГЛа	Arg	Ser	Leu	Arg 880
Pro	Gln	Pro	Thr	Glu 885	Leu	Lys	Tyr	Ser	Trp 890	Lys	Thr	Trp	Gly	Lys 895	Ala
Lys	Met	Leu	Ser 900	Thr	Glu	Ser	His	Asn 905	Gln	Thr	Phe	Leu	Ile 910	Asp	Gly
Pro	Glu	Thr 915	Ala	Glu	САа	Pro	Asn 920	Thr	Asn	Arg	Ala	Trp 925	Asn	Ser	Leu
Glu	Val 930	Glu	Asp	Tyr	Gly	Phe 935	Gly	Val	Phe	Thr	Thr 940	Asn	Ile	Trp	Leu
Lys 945	Leu	Lys	Glu	Lys	Gln 950	Asp	Val	Phe	Сүз	Asp 955	Ser	Lys	Leu	Met	Ser 960
Ala	Ala	Ile	Lys	Asp 965	Asn	Arg	Ala	Val	His 970	Ala	Asp	Met	Gly	Tyr 975	Trp
Ile	Glu	Ser	Ala 980	Leu	Asn	Asp	Thr	Trp 985	Гла	Ile	Glu	Lys	Ala 990	Ser	Phe
Ile	Glu	Val 995	Lys	Asn	Сүз	His	Trp 100		э Цу:	s Sei	r Hi:	s Th: 10		eu T:	rp Ser
	Gly 1010		l Lei	ı Glu	ı Ser	Glu 101		et I	le I	le Pi		ys 1 020	Asn I	Leu J	Ala
Gly	Pro 1025		l Se:	r Glı	n His	Asr 103		yr A	rg Pi	ro G		yr 1 035	His 1	Thr (	Gln
Ile	Thr 1040		y Pro	o Trj	p His	5 Lei 104		ly Ly	ys Le	eu Gi		et 1 050	Asp I	Phe J	Asp
Phe	Cys 1055		o Gly	y Th:	r Thr	: Val 100		al Va	al Tł	nr G		ap ( 065	Cys (	Gly J	Asn
Arg	Gly 1070		Se:	r Leı	ı Arç	j Thi 107		hr Tì	nr Al	la Se		ly 1 080	Lys I	Leu :	Ile
Thr	Glu 1085		p Cy	а Су:	s Arg	9 Sei 109		ys Tł	nr Le	eu Pi		ro 1 095	Leu A	Arg '	Tyr
Arg	Gly 1100	Glu )	ı Asj	p Gly	ү Суа	5 Tr <u>p</u> 11(	)5 T	yr G	ly Me	et G	lu I: 1:	le 2 110	Arg I	Pro I	Leu
Lys	Glu 1115		s Glı	ı Glı	ı Asr	1 Leu 112		al A:	sn Se	er Le		al ' 125	Thr A	Ala (	Gly
His	Gly 1130		n Val	l Asj	o Asr	n Phe 113		er Le	eu G	ly Va		∋u ( 140	Gly N	Met i	Ala
Leu	Phe 1145		ı Glı	ı Glı	ı Met	: Leu 115		rg Tì	hr Ai	rg Vá		ly ' 155	Thr I	ràa j	His
Ala	Ile 1160		ı Leı	ı Val	l Ala	116		er Pl	ne Va	al Tł		eu : 170	Ile ?	Thr (	Gly
Asn	Met 1175		r Phe	e Arg	g Asp	) Leu 118		ly A:	rg Va	al Me		al 1 185	Met \	Val (	Gly
Ala	Thr 1190		: Th:	r Asj	o Asp	) Ile 119		ly Me	et Gi	ly Vá		nr ' 200	Tyr I	Leu J	Ala

											- COI	ncir	luec	1	
Leu	Leu 1205		Ala	Phe	Lys	Val 1210		Pro	Thr	Phe	Ala 1215	Ala	Gly	Leu	
Leu	Leu 1220	-	Lys	Leu	Thr	Ser 1225	-	Glu	Leu	Met	Met 1230	Thr	Thr	Ile	
Gly	Ile 1235		Leu	Leu	Ser	Gln 1240		Thr	Leu	Pro	Glu 1245	Thr	Ile	Leu	
Glu	Leu 1250		Asp	Ala	Leu	Ala 1255		Gly	Met	Met	Val 1260	Leu	Lys	Met	
Val	Arg 1265		Met	Glu	Гла	Tyr 1270		Leu	Ala	Val	Thr 1275	Ile	Met	Ala	
Ile	Leu 1280	-	Val	Pro	Asn	Ala 1285		Ile	Leu	Gln	Asn 1290	Ala	Trp	ГÀа	
Val	Ser 1295	-	Thr	Ile	Leu	Ala 1300		Val	Ser	Val	Ser 1305	Pro	Leu	Phe	
Leu	Thr 1310		Ser	Gln	Gln	Lys 1315		_	Trp		Pro 1320	Leu	Ala	Leu	
Thr	Ile 1325	-		Leu		Pro 1330		Ala	Ile	Phe	Leu 1335	Thr	Thr	Leu	
Ser	Arg 1340			Lys		Arg 1345		Trp			Asn 1350	Glu	Ala	Ile	
Met	Ala 1355			Met		Ser 1360		Leu	Ala	Ser	Ser 1365	Leu	Leu	Lys	
Asn	Asp 1370		Pro	Met	Thr	Gly 1375		Leu	Val	Ala	Gly 1380	Gly	Leu	Leu	
Thr	Val 1385			Val		Thr 1390					Asp 1395	Leu	Glu	Leu	
Glu	Arg 1400			Asp		Lys 1405			Aap		Ala 1410	Glu	Ile	Ser	
Gly	Ser 1415		Pro	Ile	Leu	Ser 1420		Thr	Ile	Ser	Glu 1425	Asp	Gly	Ser	
Met	Ser 1430			Asn		Glu 1435				Thr		Thr	Ile	Leu	
Ile	Arg 1445		Gly	Leu	Leu	Val 1450			Gly		Phe 1455	Pro	Val	Ser	
Ile	Pro 1460	Ile	Thr	Ala	Ala	Ala 1465	-	Tyr	Leu	Trp	Glu 1470	Val	Lys	Lys	
Gln	Arg 1475	Ala	Gly	Val	Leu	Trp 1480		Val	Pro	Ser	Pro 1485	Pro	Pro	Met	
Gly	Lys 1490	Ala	Glu	Leu	Glu	Asp 1495		Ala	Tyr	Arg	Ile 1500	Lys	Gln	Lys	
Gly	Ile 1505	Leu	Gly	Tyr	Ser	Gln 1510		Gly	Ala	Gly	Val 1515	Tyr	Lys	Glu	
Gly	Thr 1520	Phe	His	Thr	Met	Trp 1525		Val	Thr	Arg	Gly 1530	Ala	Val	Leu	
Met	His 1535	Гла	Gly	Гла	Arg	Ile 1540		Pro	Ser	Trp	Ala 1545	Asp	Val	Гла	
Lys	Asp 1550		Ile	Ser	Tyr	Gly 1555	-	Gly	Trp	Lys	Leu 1560	Glu	Gly	Glu	
Trp			Gly	Glu	Glu	Val 1570		Val	Leu	Ala		Glu	Pro	Gly	
Lys		Pro	Arg	Ala	Val	Gln	Thr	Lys	Pro	Gly		Phe	Гла	Thr	

-continued

	1580					1585					1590			
Asn	Ala 1595	Gly	Thr	Ile	Gly	Ala 1600	Val	Ser	Leu	Asp	Phe 1605	Ser	Pro	Gly
Thr	Ser 1610	Gly	Ser	Pro	Ile	Ile 1615	Asp	Lys	Lys	Gly	Lys 1620	Val	Val	Gly
Leu	Tyr 1625	Gly	Asn	Gly	Val	Val 1630		Arg	Ser	Gly	Ala 1635	Tyr	Val	Ser
Ala	Ile 1640	Ala	Gln	Thr	Glu	Lys 1645		Ile	Glu	Asp	Asn 1650	Pro	Glu	Ile
Glu	Asp 1655	Asp	Ile	Phe	Arg	Lys 1660	Arg	Arg	Leu	Thr	Ile 1665	Met	Asp	Leu
His	Pro 1670	Gly	Ala	Gly	Lys	Thr 1675	Lys	Arg	Tyr	Leu	Pro 1680	Ala	Ile	Val
Arg	Glu 1685	Ala	Ile	Lys	Arg	Gly 1690	Leu	Arg	Thr	Leu	Ile 1695	Leu	Ala	Pro
Thr	Arg 1700	Val	Val	Ala	Ala	Glu 1705	Met	Glu	Glu	Ala	Leu 1710	Arg	Gly	Leu
Pro	Ile 1715	Arg	Tyr	Gln	Thr	Pro 1720	Ala	Ile	Arg	Ala	Val 1725	His	Thr	Gly
Arg	Glu 1730	Ile	Val	Asp	Leu	Met 1735	Сув	His	Ala	Thr	Phe 1740	Thr	Met	Arg
Leu	Leu 1745	Ser	Pro	Val	Arg	Val 1750	Pro	Asn	Tyr	Asn	Leu 1755	Ile	Ile	Met
Asp	Glu 1760	Ala	His	Phe	Thr	Asp 1765	Pro	Ala	Ser	Ile	Ala 1770	Ala	Arg	Gly
Tyr	Ile 1775	Ser	Thr	Arg	Val	Glu 1780	Met	Gly	Glu	Ala	Ala 1785	Gly	Ile	Phe
Met	Thr 1790	Ala	Thr	Pro	Pro	Gly 1795	Ser	Arg	Asp	Pro	Phe 1800	Pro	Gln	Ser
Asn	Ala 1805	Pro	Ile	Ile	Asp	Glu 1810	Glu	Arg	Glu	Ile	Pro 1815	Glu	Arg	Ser
Trp	Asn 1820	Ser	Gly	His	Glu	Trp 1825	Val	Thr	Asp	Phe	Lys 1830	Gly	Lys	Thr
Val	Trp 1835	Phe	Val	Pro	Ser	Ile 1840	Lys	Ala	Gly	Asn	Asp 1845	Ile	Ala	Ala
Суз	Leu 1850	Arg	ГÀа	Asn	Gly	Lys 1855		Val	Ile	Gln	Leu 1860	Ser	Arg	Lys
Thr	Phe 1865	Asp	Ser	Glu	Tyr	Val 1870	Lys	Thr	Arg	Thr	Asn 1875	Asp	Trp	Asp
Phe	Val 1880	Val	Thr	Thr	Asp	Ile 1885	Ser	Glu	Met	Gly	Ala 1890	Asn	Phe	Lys
Ala	Glu 1895	Arg	Val	Ile	Asp	Pro 1900	Arg	Arg	Cys	Met	Lys 1905	Pro	Val	Ile
Leu	Thr 1910	Asp	Gly	Glu	Glu	Arg 1915	Val	Ile	Leu	Ala	Gly 1920	Pro	Met	Pro
Val	Thr 1925	His	Ser	Ser	Ala	Ala 1930		Arg	Arg	Gly	Arg 1935	Ile	Gly	Arg
Asn	Pro 1940	Lys	Asn	Glu	Asn	Asp 1945	Gln	Tyr	Ile	Tyr	Met 1950	Gly	Glu	Pro
Leu	Glu 1955	Asn	Asp	Glu	Asp	Cys 1960	Ala	His	Trp	Lys	Glu 1965	Ala	Lys	Met

Leu	Leu 1970		Asn	Ile	Asn	Thr 1975		Glu	Gly	Ile	Ile 1980		Ser	Met
Phe	Glu 1985		Glu	Arg	Glu	Lys 1990		Asp	Ala	Ile	Asp 1995		Glu	Tyr
Arg	Leu 2000		Gly	Glu	Ala	Arg 2005		Thr	Phe	Val	Asp 2010		Met	Arg
Arg	Gly 2015		Leu	Pro	Val	Trp 2020		Ala	Tyr	Arg	Val 2025		Ala	Glu
Gly	Ile 2030		Tyr	Ala	Asp	Arg 2035		Trp	Cys	Phe	Asp 2040		Val	Lys
Asn	Asn 2045		Ile	Leu	Glu	Glu 2050		Val	Glu	Val	Glu 2055		Trp	Thr
ГЛа	Glu 2060		Glu	Arg	ГÀа	Lys 2065		Lys	Pro	Arg	Trp 2070		Asp	Ala
Arg	Ile 2075		Ser	Asp	Pro	Leu 2080		Leu	Lys	Glu	Phe 2085		Glu	Phe
Ala	Ala 2090	Gly	Arg	ГЛЗ	Ser	Leu 2095		Leu	Asn	Leu	Ile 2100		Glu	Met
Gly	Arg 2105		Pro	Thr	Phe	Met 2110		Gln	Lys	Ala	Arg 2115	Asp	Ala	Leu
Asp	Asn 2120		Ala	Val	Leu	His 2125		Ala	Glu	Ala	Gly 2130		Arg	Ala
Tyr	Asn 2135		Ala	Leu	Ser	Glu 2140		Pro	Glu	Thr	Leu 2145		Thr	Leu
Leu	Leu 2150		Thr	Leu	Leu	Ala 2155		Val	Thr	Gly	Gly 2160		Phe	Leu
Phe	Leu 2165		Ser	Ala	Arg	Gly 2170		Gly	Lys	Met	Thr 2175		Gly	Met
Сүз	Cys 2180	Ile	Ile	Thr	Ala	Ser 2185		Leu	Leu	Trp	Tyr 2190		Gln	Ile
Gln	Pro 2195	His	Trp	Ile	Ala	Ala 2200		Ile	Ile	Leu	Glu 2205		Phe	Leu
Ile	Val 2210	Leu	Leu	Ile	Pro	Glu 2215		Glu	Lys	Gln	Arg 2220		Pro	Gln
Asp	Asn 2225	Gln	Leu	Thr	Tyr	Val 2230		Ile	Ala	Ile	Leu 2235	Thr	Val	Val
Ala	Ala 2240	Thr	Met	Ala	Asn	Glu 2245	Met	Gly	Phe	Leu	Glu 2250	Lys	Thr	Lya
ГЛа	Asp 2255	Leu	Gly	Leu	Gly	Ser 2260		Ala	Thr	Gln	Gln 2265	Pro	Glu	Ser
Asn	Ile 2270	Leu	Asp	Ile	Asp	Leu 2275	Arg	Pro	Ala	Ser	Ala 2280	Trp	Thr	Leu
Tyr	Ala 2285	Val	Ala	Thr	Thr	Phe 2290	Val	Thr	Pro	Met	Leu 2295	Arg	His	Ser
Ile	Glu 2300	Asn	Ser	Ser	Val	Asn 2305	Val	Ser	Leu	Thr	Ala 2310	Ile	Ala	Asn
Gln	Ala 2315	Thr	Val	Leu	Met	Gly 2320		Gly	Lys	Gly	Trp 2325	Pro	Leu	Ser
ГЛа	Met 2330	Asp	Ile	Gly	Val	Pro 2335		Leu	Ala	Ile	Gly 2340	Суз	Tyr	Ser

-	COI	nti	nu	ed

											- COI	nt 11	nuec	1
Gln	Val 2345		Pro	Ile	Thr	Leu 2350		Ala	Ala	Leu	Phe 2355		Leu	Val
Ala	His 2360		Ala	Ile	Ile	Gly 2365		Gly	Leu	Gln	Ala 2370		Ala	Thr
Arg	Glu 2375	Ala	Gln	Гла	Arg	Ala 2380		Ala	Gly	Ile	Met 2385	_	Asn	Pro
Thr	Val 2390		Gly	Ile	Thr	Val 2395			Leu		Pro 2400	Ile	Pro	Tyr
Asp	Pro 2405	-	Phe	Glu	Lys	Gln 2410		Gly	Gln	Val	Met 2415	Leu	Leu	Val
Leu	Cys 2420	Val	Thr	Gln	Val	Leu 2425			Arg		Thr 2430	_	Ala	Leu
Суз	Glu 2435	Ala	Leu	Thr	Leu	Ala 2440		Gly	Pro	Ile	Ser 2445	Thr	Leu	Trp
Glu	Gly 2450	Asn		Gly		Phe 2455					Ile 2460	Ala	Val	Ser
Met	Ala 2465		Ile	Phe	Arg	Gly 2470		Tyr	Leu	Ala	Gly 2475	Ala	Gly	Leu
Leu			Ile	Met	Lys		Thr	Thr	Asn	Thr	Arg 2490	-	Gly	Thr
Gly				Glu			Gly		Lys		Lys 2505		Arg	Leu
Asn		Leu	Gly		Ser		Phe				Lys 2520	Lys	Ser	Gly
Ile		Glu	Val		Arg		Leu				Gly 2535	Ile	Lys	Arg
Gly		Thr	Asp				Val	Ser		Gly	Ser 2550	Ala	Lys	Leu
Arg	Trp	Phe			Arg	Asn	Met				Glu	Gly	ГЛЗ	Val
Val		Leu					Gly				2565 Tyr		Суз	Gly
Gly		Lys	Asn		Arg		Val	Lys	Gly	Leu	2580 Thr		Gly	Gly
Pro				Glu			Pro	Met			2595 Tyr	Gly	Trp	Asn
Leu	2600 Val	Arg	Leu	Gln	Ser	2605 Gly		Asp	Val	Phe	2610 Phe	Ile	Pro	Pro
Glu	2615 Lys	Суз	Asp	Thr	Leu	2620 Leu		Asp	Ile	Gly	2625 Glu		Ser	Pro
	2630	-	-			2635	-	-		-	2640 Val			
	2645					2650	-			-	2655			
	2660		-			2665					Cys 2670		-	
Leu	Asn 2675	Pro	Tyr	Met	Pro	Ser 2680		Ile	Glu	Lys	Met 2685		Ala	Leu
Gln	Arg 2690	-	Tyr	Gly	Gly	Ala 2695		Val	Arg	Asn	Pro 2700	Leu	Ser	Arg
Asn	Ser 2705	Thr	His	Glu	Met	Tyr 2710	-	Val	Ser	Asn	Ala 2715	Ser	Gly	Asn
Ile	Val	Ser	Ser	Val	Asn	Met	Ile	Ser	Arg	Met	Leu	Ile	Asn	Arg

-continued

	2720					2725					2730			
Phe	Thr 2735	Met	Arg	Tyr	Гла	Lys 2740	Ala	Thr	Tyr	Glu	Pro 2745	Asp	Val	Asp
Leu	Gly 2750	Ser	Gly	Thr	Arg	Asn 2755	Ile	Gly	Ile	Glu	Ser 2760	Glu	Ile	Pro
Asn	Leu 2765	Asp	Ile	Ile	Gly	Lys 2770	Arg	Ile	Glu	Гла	Ile 2775	Lys	Gln	Glu
His	Glu 2780	Thr	Ser	Trp	His	Tyr 2785	Asp	Gln	Asp	His	Pro 2790	_	Lys	Thr
Trp	Ala 2795	Tyr	His	Gly	Ser	Tyr 2800	Glu	Thr	Lys	Gln	Thr 2805	Gly	Ser	Ala
Ser	Ser 2810	Met	Val	Asn	Gly	Val 2815	Val	Arg	Leu	Leu	Thr 2820	Lys	Pro	Trp
Asp	Val 2825	Val	Pro	Met	Val	Thr 2830	Gln	Met	Ala	Met	Thr 2835	Asp	Thr	Thr
Pro	Phe 2840	Gly	Gln	Gln	Arg	Val 2845	Phe	Lys	Glu	ГЛа	Val 2850	Asp	Thr	Arg
Thr	Gln 2855	Glu	Pro	Гла	Glu	Gly 2860	Thr	Lys	Lys	Leu	Met 2865	Lys	Ile	Thr
Ala	Glu 2870	Trp	Leu	Trp	Lys	Glu 2875	Leu	Gly	Lys	Гла	Lys 2880	Thr	Pro	Arg
Met	Сув 2885	Thr	Arg	Glu	Glu	Phe 2890	Thr	Arg	Lys	Val	Arg 2895	Ser	Asn	Ala
Ala	Leu 2900	Gly	Ala	Ile	Phe	Thr 2905	Asp	Glu	Asn	Гла	Trp 2910	Lys	Ser	Ala
Arg	Glu 2915	Ala	Val	Glu	Asp	Ser 2920	Arg	Phe	Trp	Glu	Leu 2925	Val	Asp	Lys
Glu	Arg 2930	Asn	Leu	His	Leu	Glu 2935	Gly	Lys	Сүз	Glu	Thr 2940	Сүз	Val	Tyr
Asn	Met 2945	Met	Gly	Lys	Arg	Glu 2950	Lys	Lys	Leu	Gly	Glu 2955	Phe	Gly	Lys
Ala	Lys 2960	Gly	Ser	Arg	Ala	Ile 2965	Trp	Tyr	Met	Trp	Leu 2970	Gly	Ala	Arg
Phe	Leu 2975	Glu	Phe	Glu	Ala	Leu 2980	Gly	Phe	Leu	Asn	Glu 2985	Asp	His	Trp
Phe	Ser 2990	Arg	Glu	Asn	Ser	Leu 2995	Ser	Gly	Val	Glu	Gly 3000	Glu	Gly	Leu
His	Lуя 3005	Leu	Gly	Tyr	Ile	Leu 3010	Arg	Asp	Val	Ser	Lуя 3015	Lys	Glu	Gly
Gly	Ala 3020	Met	Tyr	Ala	Asp	Asp 3025	Thr	Ala	Gly	Trp	Aap 3030	Thr	Arg	Ile
Thr	Leu 3035	Glu	Asp	Leu	Lys	Asn 3040	Glu	Glu	Met	Val	Thr 3045	Asn	His	Met
Glu	Gly 3050	Glu	His	Lys	Lys	Leu 3055	Ala	Glu	Ala	Ile	Phe 3060	Lys	Leu	Thr
Tyr	Gln 3065	Asn	Lys	Val	Val	Arg 3070	Val	Gln	Arg	Pro	Thr 3075	Pro	Arg	Gly
Thr	Val 3080	Met	Asp	Ile	Ile	Ser 3085	Arg	Arg	Asp	Gln	Arg 3090	Gly	Ser	Gly
Gln	Val 3095	Gly	Thr	Tyr	Gly	Leu 3100	Asn	Thr	Phe	Thr	Asn 3105	Met	Glu	Ala

Gln	Leu 3110		Arg	Gln	Met	Glu 3115		Glu	Gly	Val	Phe 3120		Ser	Ile
Gln	His 3125		Thr	Ile	Thr	Glu 3130		Ile	Ala	Val	Gln 3135	Asn	Trp	Leu
Ala	Arg 3140		Gly	Arg	Glu	Arg 3145	Leu	Ser	Arg	Met	Ala 3150	Ile	Ser	Gly
Asp	Asp 3155		Val	Val	Гла	Pro 3160	Leu	Aap	Asp	Arg	Phe 3165	Ala	Ser	Ala
Leu	Thr 3170		Leu	Asn	Asp	Met 3175	Gly	Lys	Ile	Arg	Lys 3180	Asp	Ile	Gln
Gln	Trp 3185		Pro	Ser	Arg	Gly 3190	Trp	Asn	Asp	Trp	Thr 3195	Gln	Val	Pro
Phe	Cys 3200		His	His	Phe	His 3205		Leu	Ile	Met	Lys 3210	Asp	Gly	Arg
Val	Leu 3215		Val	Pro	Сүз	Arg 3220	Asn	Gln	Aap	Glu	Leu 3225	Ile	Gly	Arg
Ala	Arg 3230		Ser	Gln	Gly	Ala 3235	Gly	Trp	Ser	Leu	Arg 3240	Glu	Thr	Ala
Cys	Leu 3245	Gly	Lys	Ser		Ala 3250		Met	Trp	Ser	Leu 3255	Met	Tyr	Phe
His	Arg 3260		Asp	Leu	Arg	Leu 3265	Ala	Ala	Asn	Ala	Ile 3270	Суз	Ser	Ala
Val	Pro 3275	Ser	His	Trp	Val	Pro 3280	Thr	Ser	Arg	Thr	Thr 3285	Trp	Ser	Ile
His	Ala 3290		His	Glu	Trp	Met 3295	Thr	Thr	Glu	Asp	Met 3300	Leu	Thr	Val
Trp	Asn 3305	Arg	Val	Trp	Ile	Gln 3310		Asn	Pro	Trp	Met 3315	Glu	Asp	Гла
Thr	Pro 3320		Glu	Ser	Trp	Glu 3325	Glu	Ile	Pro	Tyr	Leu 3330	Gly	Lys	Arg
Glu	Asp 3335	Gln	Trp	Сүз	Gly	Ser 3340	Leu	Ile	Gly	Leu	Thr 3345	Ser	Arg	Ala
Thr	Trp 3350		Lys	Asn	Ile	Gln 3355	Ala	Ala	Ile	Asn	Gln 3360	Val	Arg	Ser
Leu	Ile 3365	Gly	Asn	Glu	Glu	Tyr 3370	Thr	Asp	Tyr	Met	Pro 3375	Ser	Met	Lys
Arg	Phe 3380	Arg	Arg	Glu	Glu	Glu 3385	Glu	Ala	Gly	Val	Leu 3390	Trp		
<211 <212 <213 <220	)> FEA	NGTH PE: 1 GANIS ATURI	: 339 PRT SM: <i>1</i> E:	91 Artif		al Sec Dengu	-		sero	otype	≥ 1, M	IVS		
<400	)> SE(	QUEN	CE: 3	3										
Met 1	Asn i	Asn (		Arg I 5	'Àa I	Jya Al	la Ly	/s As 10		nr Pi	ro Phe	e Asr	n Met 15	: Leu
Lys	Arg (		Arg <i>1</i> 20	Asn A	\rg \	/al Se	er Th 25		al GI	ln G	ln Leu	1 Thi 30	c Ly:	s Arg
Phe		Leu ( 35	Gly M	4et I	Jeu (	3ln Gl 40	-	rg Gl	Ly Pi	ro Le	eu Ly: 45	s Leu	ı Phe	e Met

-continued		

Ala	Leu 50	Val	Ala	Phe	Leu	Arg 55	Phe	Leu	Thr	Ile	Pro 60	Pro	Thr	Ala	Gly
Ile 65	Leu	Lys	Arg	Trp	Gly 70	Thr	Ile	Lys	Lys	Ser 75	Lys	Ala	Ile	Asn	Val 80
Leu	Arg	Gly	Phe	Arg 85	Lys	Glu	Ile	Gly	Arg 90	Met	Leu	Asn	Ile	Leu 95	Asn
Arg	Arg	Arg	Arg 100	Ser	Ala	Gly	Met	Ile 105	Ile	Met	Leu	Ile	Pro 110	Thr	Val
Met	Ala	Phe 115	His	Leu	Thr	Thr	Arg 120	Gly	Gly	Glu	Pro	His 125	Met	Ile	Val
Ser	Lys 130	Gln	Glu	Arg	Gly	Lys 135	Ser	Leu	Leu	Phe	Lys 140	Thr	Ser	Ala	Gly
Val 145	Asn	Met	Cys	Thr	Leu 150	Ile	Ala	Met	Asp	Leu 155	Gly	Glu	Leu	Суз	Glu 160
Asp	Thr	Met	Thr	Tyr 165	LYa	Сүз	Pro	Arg	Ile 170	Thr	Glu	Ala	Glu	Pro 175	Asp
Asp	Val	Aab	Cys 180	Trp	Сүз	Asn	Ala	Thr 185	Asp	Thr	Trp	Val	Thr 190	Tyr	Gly
Thr	Суз	Ser 195	Gln	Thr	Gly	Glu	His 200	Arg	Arg	Asp	ГЛа	Arg 205	Ser	Val	Ala
Leu	Ala 210	Pro	His	Val	Gly	Leu 215	Gly	Leu	Glu	Thr	Arg 220	Ala	Glu	Thr	Trp
Met 225	Ser	Ser	Glu	Gly	Ala 230	Trp	Lys	Gln	Ile	Gln 235	ГЛа	Val	Glu	Thr	Trp 240
Ala	Leu	Arg	His	Pro 245	Gly	Phe	Thr	Val	Ile 250	Ala	Leu	Phe	Leu	Ala 255	His
Ala	Ile	Gly	Thr 260	Ser	Ile	Thr	Gln	Lys 265	Gly	Ile	Ile	Phe	Ile 270	Leu	Leu
Met	Leu	Val 275	Thr	Pro	Ser	Met	Ala 280	Met	Arg	Сув	Val	Gly 285	Ile	Gly	Asn
Arg	Asp 290	Phe	Val	Glu	Gly	Leu 295	Ser	Gly	Ala	Thr	Trp 300	Val	Asp	Val	Val
Leu 305	Glu	His	Gly	Ser	Cys 310	Val	Thr	Thr	Met	Ala 315	Lys	Asn	Lys	Pro	Thr 320
Leu	Asp	Ile	Glu	Leu 325	Leu	Lys	Thr	Glu	Val 330	Thr	Asn	Pro	Ala	Val 335	Leu
Arg	Lys	Leu	Cys 340	Ile	Glu	Ala	Lys	Ile 345	Ser	Asn	Thr	Thr	Thr 350	Asp	Ser
Arg	Суз	Pro 355	Thr	Gln	Gly	Glu	Ala 360	Thr	Leu	Val	Glu	Glu 365	Gln	Asp	Ala
Asn	Phe 370	Val	Сув	Arg	Arg	Thr 375	Phe	Val	Asp	Arg	Gly 380	Trp	Gly	Asn	Gly
Суя 385	Gly	Leu	Phe	Gly	Lys 390	Gly	Ser	Leu	Ile	Thr 395	Сүз	Ala	Lys	Phe	Lys 400
Суз	Val	Thr	Lys	Leu 405	Glu	Gly	Lys	Ile	Val 410	Gln	Tyr	Glu	Asn	Leu 415	Lys
Tyr	Ser	Val	Ile 420	Val	Thr	Val	His	Thr 425	Gly	Asp	Gln	His	Gln 430	Val	Gly
Asn	Glu	Thr 435	Thr	Glu	His	Gly	Thr 440	Thr	Ala	Thr	Ile	Thr 445	Pro	Gln	Ala

											-	con	tin	ued	
Pro	Thr 450	Ser	Glu	Ile	Gln	Leu 455	Thr	Asp	Tyr	Gly	Thr 460	Leu	Thr	Leu	Asp
Cys 465	Ser	Pro	Arg	Thr	Gly 470	Leu	Asp	Phe	Asn	Glu 475	Met	Val	Leu	Leu	Thr 480
Met	Lys	Glu	Arg	Ser 485	Trp	Leu	Val	His	Lys 490	Gln	Trp	Phe	Leu	Asp 495	Leu
Pro	Leu	Pro	Trp 500	Thr	Ser	Gly	Ala	Ser 505	Thr	Ser	Gln	Glu	Thr 510	Trp	Asn
Arg	Gln	Asp 515	Leu	Leu	Val	Thr	Phe 520	Lys	Thr	Ala	His	Ala 525	Lys	Lys	Gln
Glu	Val 530	Val	Val	Leu	Gly	Ser 535	Gln	Glu	Gly	Ala	Met 540	His	Thr	Ala	Leu
Thr 545		Ala	Thr	Glu	Ile 550	Gln	Thr	Ser	Gly	Thr 555	Thr	Thr	Ile	Phe	Ala 560
Gly	His	Leu	Lys	Суз 565	Arg	Leu	Lys	Met	Asp 570	Lys	Leu	Thr	Leu	Lys 575	Gly
Met	Ser	Tyr	Val 580		Суз	Thr	Gly	Ser 585		Lys	Leu	Glu	Lys 590		Val
Ala	Glu	Thr 595		His	Gly	Thr	Val 600		Val	Gln	Val	Lys 605		Glu	Gly
Thr	Asp 610		Pro	СЛа	Lys	Ile 615		Phe	Ser	Thr	Gln 620		Glu	Lys	Gly
Ala 625		Gln	Asn	Gly	Arg 630		Ile	Thr	Ala	Asn 635		Ile	Val	Thr	Asp 640
	Glu	Lys	Pro	Val 645	Asn	Ile	Glu	Ala	Glu 650		Pro	Phe	Gly	Glu 655	
Tyr	Ile	Val	Val 660		Ala	Gly	Glu	Lys 665		Leu	Гла	Leu	Ser 670		Phe
Lys	Lys	Gly 675		Ser	Ile	Gly	Lys 680		Phe	Glu	Ala	Thr 685		Arg	Gly
Ala	-		Met	Ala	Ile			Asp	Thr	Ala	_		Phe	Gly	Ser
	690 Gly	Gly	Val	Phe	Thr	695 Ser	Met	Gly	ГÀа		700 Val	His	Gln	Val	
705 Gly	Thr	Ala	Tyr	-	710 Val	Leu	Phe	Ser	-		Ser	Trp	Thr		720 Lys
Ile	Gly	Ile	Gly	725 Ile	Leu	Leu	Thr	Trp	730 Leu		Leu	Asn	Ser	735 Arg	Asn
			740		Met			745					750		
		755				-	760			-		765			-
цец	770	vaı	met	var	Gln	A1a 775	Чар	ser	σтλ	сув	780	vaı	ser	ττþ	пув
Asn 785		Glu	Leu	Lys	Cys 790	Gly	Ser	Gly	Ile	Phe 795	Ile	Thr	Aab	Asn	Val 800
His	Thr	Trp	Thr	Glu 805	Gln	Tyr	Lys	Phe	Gln 810	Pro	Glu	Ser	Pro	Ser 815	Lys
Leu	Ala	Ser	Ala 820	Ile	Gln	Гла	Ala	His 825	Glu	Glu	Asp	Ile	Сув 830	Gly	Ile
Arg	Ser	Val 835	Thr	Arg	Leu	Glu	Asn 840	Leu	Met	Trp	ГЛа	Gln 845	Ile	Thr	Pro
Glu	Leu	Asn	His	Ile	Leu	Ser	Glu	Asn	Glu	Val	Lys	Leu	Thr	Ile	Met

850 855 860
Thr Gly Asp Ile Lys Gly Ile Met Gln Ala Gly Lys Arg Ser Leu Arg865870875880
Pro Gln Pro Thr Glu Leu Lys Tyr Ser Trp Lys Thr Trp Gly Lys Ala 885 890 895
Lys Met Leu Ser Thr Glu Ser His Asn Gln Thr Phe Leu Ile Asp Gly 900 905 910
Pro Glu Thr Ala Glu Cys Pro Asn Thr Asn Arg Ala Trp Asn Ser Leu 915 920 925
Glu Val Glu Asp Tyr Gly Phe Gly Val Phe Thr Thr Asn Ile Trp Leu 930 935 940
Lys Leu Lys Glu Lys Gln Asp Val Phe Cys Asp Ser Lys Leu Met Ser 945 950 955 960
Ala Ala Ile Lys Asp Asn Arg Ala Val His Ala Asp Met Gly Tyr Trp 965 970 975
Ile Glu Ser Ala Leu Asn Asp Thr Trp Lys Ile Glu Lys Ala Ser Phe 980 985 990
Ile Glu Val Lys Asn Cys His Trp Pro Lys Ser His Thr Leu Trp Ser 995 1000 1005
Asn Gly Val Leu Glu Ser Glu Met Ile Ile Pro Lys Asn Leu Ala 1010 1015 1020
Gly Pro Val Ser Gln His Asn Tyr Arg Pro Gly Tyr His Thr Gln 1025 1030 1035
Ile Thr Gly Pro Trp His Leu Gly Lys Leu Glu Met Asp Phe Asp         1040       1045
Phe Cys Asp Gly Thr Thr Val Val Val Thr Glu Asp Cys Gly Asn 1055 1060 1065
Arg GlyPro Ser Leu Arg ThrThr Thr Ala Ser GlyLysLeu Ile107010751080
Thr GluTrp Cys Cys Arg SerCys Thr Leu ProProLeu Arg Tyr108510901095
Arg Gly Glu Asp Gly Cys TrpTyr Gly Met Glu IleArg Pro Leu110011051110
Lys Glu Lys Glu Glu Asn Leu Val Asn Ser Leu Val Thr Ala Gly 1115 1120 1125
His Gly Gln Val Asp Asn Phe Ser Leu Gly Val Leu Gly Met Ala 1130 1135 1140
Leu Phe Leu Glu Glu Met Leu Arg Thr Arg Val Gly Thr Lys His 1145 1150 1155
Ala Ile Leu Leu Val Ala Val Ser Phe Val Thr Leu Ile Thr Gly 1160 1165 1170
Asn Met Ser Phe Arg Asp Leu Gly Arg Val Met Val Met Val Gly 1175 1180 1185
Ala Thr Met Thr Asp Asp Ile Gly Met Gly Val Thr Tyr Leu Ala 1190 1195 1200
Leu Leu Ala Ala Phe Lys Val Arg Pro Thr Phe Ala Ala Gly Leu 1205 1210 1215
Leu Leu Arg Lys Leu Thr Ser Lys Glu Leu Met Met Thr Thr Ile 1220 1225 1230
Gly Ile Val Leu Leu Ser Gln Ser Thr Leu Pro Glu Thr Ile Leu 1235 1240 1245

Glu	Leu 1250	Thr	Asp	Ala	Leu	Ala 1255	Leu	Gly	Met	Met	Val 1260	Leu	Lys	Met
Val	Arg 1265	Asn	Met	Glu	Lys	Tyr 1270	Gln	Leu	Ala	Val	Thr 1275	Ile	Met	Ala
Ile	Leu 1280	CAa	Val	Pro	Asn	Ala 1285	Val	Ile	Leu	Gln	Asn 1290	Ala	Trp	Lys
Val	Ser 1295	Сүз	Thr	Ile	Leu	Ala 1300		Val	Ser	Val	Ser 1305	Pro	Leu	Phe
Leu	Thr 1310	Ser	Ser	Gln	Gln	Lys 1315	Thr	Aab	Trp	Ile	Pro 1320	Leu	Ala	Leu
Thr	Ile 1325	Lys	Gly	Leu	Asn	Pro 1330	Thr	Ala	Ile	Phe	Leu 1335	Thr	Thr	Leu
Ser	Arg 1340	Thr	Ser	Lys	Lys	Arg 1345	Ser	Trp	Pro	Leu	Asn 1350	Glu	Ala	Ile
Met	Ala 1355	Val	Gly	Met	Val	Ser 1360	Ile	Leu	Ala	Ser	Ser 1365	Leu	Leu	Lys
Asn	Asp 1370	Ile	Pro	Met	Thr	Gly 1375	Pro	Leu	Val	Ala	Gly 1380	Gly	Leu	Leu
Thr	Val 1385	Cys	Tyr	Val	Leu	Thr 1390		Arg	Ser	Ala	Asp 1395	Leu	Glu	Leu
Glu	Arg 1400	Ala	Ala	Asp	Val	Lys 1405	Trp	Glu	Aap	Gln	Ala 1410	Glu	Ile	Ser
Gly	Ser 1415	Ser	Pro	Ile	Leu	Ser 1420		Thr	Ile	Ser	Glu 1425	Asp	Gly	Ser
Met	Ser 1430	Ile	Lys	Asn	Glu	Glu 1435	Glu	Aap	Gln	Thr	Leu 1440	Thr	Ile	Leu
Ile	Arg 1445	Thr	Gly	Leu	Leu	Val 1450		Ser	Gly	Leu	Phe 1455	Pro	Val	Ser
Ile	Pro 1460	Ile	Thr	Ala	Ala	Ala 1465	Trp	Tyr	Leu	Trp	Glu 1470	Val	Lys	Гла
Gln	Arg 1475	Ala	Gly	Val	Leu	Trp 1480	Asp	Val	Pro	Ser	Pro 1485	Pro	Pro	Met
Gly	Lys 1490	Ala	Glu	Leu	Glu	Asp 1495	Gly	Ala	Tyr	Arg	Ile 1500	Lys	Gln	Гла
Gly	Ile 1505	Leu	Gly	Tyr	Ser	Gln 1510	Ile	Gly	Ala	Gly	Val 1515	Tyr	ГЛЗ	Glu
Gly	Thr 1520	Phe	His	Thr	Met	Trp 1525	His	Val	Thr	Arg	Gly 1530	Ala	Val	Leu
Met	His 1535	ГÀа	Gly	Lys	Arg	Ile 1540	Glu	Pro	Ser	Trp	Ala 1545	Asp	Val	Гла
Lys	Asp 1550	Leu	Ile	Ser	Tyr	Gly 1555	Gly	Gly	Trp	Lys	Leu 1560	Glu	Gly	Glu
Trp	Lys 1565	Glu	Gly	Glu	Glu	Val 1570	Gln	Val	Leu	Ala	Leu 1575	Glu	Pro	Gly
Lys	Asn 1580	Pro	Arg	Ala	Val	Gln 1585	Thr	Lys	Pro	Gly	Leu 1590	Phe	Lys	Thr
Asn	Ala 1595	Gly	Thr	Ile	Gly	Ala 1600	Val	Ser	Leu	Asp	Phe 1605	Ser	Pro	Gly
Thr	Ser 1610	Gly	Ser	Pro	Ile	Ile 1615	Asp	Lys	Lys	Gly	Lys 1620	Val	Val	Gly

											- COI	ntır	nued	1
Leu	Tyr 1625	Gly	Asn	Gly	Val	Val 1630		Arg	Ser	Gly	Ala 1635		Val	Ser
Ala	Ile 1640	Ala	Gln	Thr	Glu	Lys 1645		Ile	Glu	Asp	Asn 1650	Pro	Glu	Ile
Glu	Asp 1655	-	Ile	Phe	Arg	Lys 1660		Arg	Leu	Thr	Ile 1665	Met	Asp	Leu
His	Pro 1670		Ala	Gly	Lys	Thr 1675			Tyr		Pro 1680	Ala	Ile	Val
Arg	Glu 1685	Ala	Ile	Lys	Arg	Gly 1690			Thr		Ile 1695		Ala	Pro
Thr	Arg 1700		Val	Ala	Ala	Glu 1705		Glu	Glu	Ala	Leu 1710	-	Gly	Leu
Pro	Ile 1715	-	Tyr	Gln	Thr	Pro 1720			Arg		Val 1725	His	Thr	Gly
Arg	Glu 1730	Ile		Asp		Met 1735		His	Ala	Thr	Phe 1740	Thr	Met	Arg
Leu	Leu 1745	Ser	Pro	Val	Arg	Val 1750			Tyr		Leu 1755	Ile	Ile	Met
Asp		Ala	His	Phe	Thr	Asp 1765	Pro					Ala	Arg	Gly
Tyr		Ser		Arg		Glu 1780	Met	Gly		Ala		-	Ile	Phe
Met		Ala				Gly	Ser	Arg	Asp		Phe	Pro	Gln	Ser
Asn	Ala		Ile	Ile	Asp	1795 Glu	Glu		Glu	Ile			Arg	Ser
Trp		Ser	Gly	His	Glu	1810 Trp	Val		Asp	Phe			Lys	Thr
Val	1820 Trp		Val	Pro	Ser	1825 Ile		Ala	Gly	Asn		Ile	Ala	Ala
Cys	1835 Leu		Lys	Asn	Gly	1840 Lys					1845 Leu		Arg	Lys
Thr	1850 Phe			Glu		1855 Val		Thr	Arq	Thr	1860 Asn		Trp	- Asp
	1865	-			•	1870 Ile	-		5		1875	-	-	-
	1880				-	1885				-	1890			-
	1895					Pro 1900					1905			
Leu	Thr 1910	Aap	Gly	Glu	Glu	Arg 1915		Ile	Leu	Ala	Gly 1920		Met	Pro
Val	Thr 1925	His	Ser	Ser	Ala	Ala 1930		Arg	Arg	Gly	Arg 1935		Gly	Arg
Asn	Pro 1940	Lys	Asn	Glu	Asn	Asp 1945		Tyr	Ile	Tyr	Met 1950	-	Glu	Pro
Leu	Glu 1955	Asn	Asp	Glu	Asp	Суз 1960		His	Trp	Lys	Glu 1965	Ala	Lys	Met
Leu	Leu 1970	Asp	Asn	Ile	Asn	Thr 1975		Glu	Gly	Ile	Ile 1980	Pro	Ser	Met
Phe	Glu 1985	Pro	Glu	Arg	Glu	Lys 1990		Asp	Ala	Ile	Asp 1995	Gly	Glu	Tyr
Arg		Arg	Gly	Glu	Ala	Arg		Thr	Phe	Val		Leu	Met	Arg

-continued

	2000					2005					2010			
Arg	Gly 2015	Asp	Leu	Pro	Val	Trp 2020		Ala	Tyr	Arg	Val 2025	Ala	Ala	Glu
Gly	Ile 2030	Asn	Tyr	Ala	Asp	Arg 2035	Arg	Trp	Сүз	Phe	Asp 2040		Val	Lys
Asn	Asn 2045	Gln	Ile	Leu	Glu	Glu 2050		Val	Glu	Val	Glu 2055		Trp	Thr
ГЛЗ	Glu 2060	Gly	Glu	Arg	ГЛЗ	Lуя 2065	Leu	Lys	Pro	Arg	Trp 2070		Asp	Ala
-	Ile 2075	-		-		2080			-		Phe 2085	-		
	Ala 2090					2095					2100			
	Arg 2105					2110					2115			
	Asn 2120					2125					2130			
-	Asn 2135 Leu					2140					2145			
	Leu 2150 Leu					2155					2160			
	2165 Cys				-	2170		-	-		2175 Tyr		-	
-	2180 Pro					2185				-	2190			
	2195 Val		-			2200					2205			
Asp	2210 Asn	Gln	Leu	Thr	Tyr	2215 Val	Val	Ile	Ala	Ile	2220 Leu	Thr	Val	Val
Ala	2225 Ala	Thr	Met	Ala	Asn			Gly	Phe	Leu			Thr	Lys
Lys	2240 Asp	Leu	Gly	Leu	Gly			Ala	Thr	Gln			Glu	Ser
Asn	2255 Ile 2270		Asp	Ile	Asp	2260 Leu 2275	Arg	Pro	Ala	Ser	2265 Ala 2280	_	Thr	Leu
Tyr	Ala 2285		Ala	Thr	Thr		Val	Thr	Pro	Met			His	Ser
Ile	Glu 2300	Asn	Ser	Ser	Val	Asn 2305	Val	Ser	Leu	Thr	Ala 2310	Ile	Ala	Asn
Gln	Ala 2315	Thr	Val	Leu	Met	Gly 2320	Leu	Gly	Гла	Gly	Trp 2325	Pro	Leu	Ser
ГÀа	Met 2330	Asp	Ile	Gly	Val	Pro 2335	Leu	Leu	Ala	Ile	Gly 2340	-	Tyr	Ser
Gln	Val 2345	Asn	Pro	Ile	Thr	Leu 2350		Ala	Ala	Leu	Phe 2355		Leu	Val
Ala	His 2360	Tyr	Ala	Ile	Ile	Gly 2365	Pro	Gly	Leu	Gln	Ala 2370	-	Ala	Thr
Arg	Glu 2375	Ala	Gln	Lys	Arg		Ala	Ala	Gly	Ile	Met 2385		Asn	Pro

Thr	Val 2390		Gly	Ile	Thr	Val 2395		Asp	Leu	Asp	Pro 2400	Ile	Pro	Tyr
Asp	Pro 2405		Phe	Glu		Gln 2410		Gly	Gln	Val	Met 2415	Leu	Leu	Val
Leu	Cys 2420	Val	Thr	Gln	Val	Leu 2425		Met	Arg	Thr	Thr 2430	Trp	Ala	Leu
Сүз	Glu 2435		Leu	Thr	Leu	Ala 2440		Gly	Pro	Ile	Ser 2445	Thr	Leu	Trp
Glu	Gly 2450	Asn	Pro	Gly	Arg	Phe 2455	-	Asn	Thr	Thr	Ile 2460	Ala	Val	Ser
Met	Ala 2465	Asn	Ile	Phe	Arg	Gly 2470		Tyr	Leu	Ala	Gly 2475	Ala	Gly	Leu
Leu	Phe 2480		Ile	Met	Lys	Asn 2485		Thr	Asn	Thr	Arg 2490	Arg	Gly	Thr
Gly	Asn 2495		Gly	Glu	Thr	Leu 2500		Glu	Lys	Trp	Lys 2505	Ser	Arg	Leu
Asn	Ala 2510	Leu	Gly	Lys	Ser	Glu 2515		Gln	Ile	Tyr	Lys 2520	Lys	Ser	Gly
Ile	Gln 2525	Glu	Val	Asp	Arg	Thr 2530		Ala	ГÀа	Glu	Gly 2535	Ile	ГЛа	Arg
Gly	Glu 2540	Thr	Asp	His	His	Ala 2545		Ser	Arg	Gly	Ser 2550	Ala	Lys	Leu
Arg	Trp 2555	Phe	Val	Glu	Arg	Asn 2560		Val	Thr	Pro	Glu 2565	Gly	ГЛа	Val
Val	Asp 2570	Leu	Gly	Суз	Gly	Arg 2575		Gly	Trp	Ser	Tyr 2580	Tyr	Суз	Gly
Gly	Leu 2585	ГЛа	Asn	Val	Arg	Glu 2590		Lys	Gly	Leu	Thr 2595	Lys	Gly	Gly
Pro	Gly 2600	His	Glu	Glu	Pro	Ile 2605		Met	Ser	Thr	Tyr 2610	Gly	Trp	Asn
Leu	Val 2615	Arg	Leu	Gln	Ser	Gly 2620		Asp	Val	Phe	Phe 2625	Ile	Pro	Pro
Glu	Lys 2630	Сүз	Asp	Thr	Leu	Leu 2635		Asp	Ile	Gly	Glu 2640	Ser	Ser	Pro
Asn	Pro 2645	Thr	Val	Glu	Ala	Gly 2650		Thr	Leu	Arg	Val 2655	Leu	Asn	Leu
Val	Glu 2660	Asn	Trp	Leu	Asn	Asn 2665	Asn	Thr	Gln	Phe	Cys 2670	Ile	ГЛа	Val
Leu	Asn 2675	Pro	Tyr	Met	Pro	Ser 2680		Ile	Glu	ГЛа	Met 2685	Glu	Ala	Leu
Gln	Arg 2690	ГЛа	Tyr	Gly	Gly	Ala 2695	Leu	Val	Arg	Asn	Pro 2700	Leu	Ser	Arg
Asn	Ser 2705	Thr	His	Glu	Met	Tyr 2710	Trp	Val	Ser	Asn	Ala 2715	Ser	Gly	Asn
Ile	Val 2720	Ser	Ser	Val	Asn	Met 2725	Ile	Ser	Arg	Met	Leu 2730	Ile	Asn	Arg
Phe	Thr 2735	Met	Arg	Tyr	Lys	Lys 2740		Thr	Tyr	Glu	Pro 2745	Asp	Val	Asp
Leu	Gly 2750	Ser	Gly	Thr	Arg	Asn 2755		Gly	Ile	Glu	Ser 2760	Glu	Ile	Pro

-continued
------------

											- COI	ntir	nuec	1
Asn	Leu 2765		Ile	Ile	Gly	Lys 2770		Ile	Glu	Lys	Ile 2775		Gln	Glu
His	Glu 2780	Thr	Ser	Trp	His	Tyr 2785		Gln	Asp	His	Pro 2790		Lys	Thr
Trp	Ala 2795		His	Gly	Ser	Tyr 2800		Thr	Lys	Gln	Thr 2805		Ser	Ala
Ser	Ser 2810	Met	Val	Asn	Gly	Val 2815		Arg	Leu	Leu	Thr 2820		Pro	Trp
Asp	Val 2825		Pro	Met	Val	Thr 2830		Met	Ala	Met	Thr 2835	-	Thr	Thr
Pro	Phe 2840	Gly	Gln	Gln	Arg	Val 2845		Lys	Glu	LYa	Val 2850	-	Thr	Arg
Thr	Gln 2855	Glu	Pro	Lys	Glu	Gly 2860			Lys		Met 2865		Ile	Thr
Ala	Glu 2870		Leu	Trp	Lys	Glu 2875			Lys		Lys 2880		Pro	Arg
Met	Cys 2885		Arg	Glu	Glu	Phe 2890		Arg		Val	Arg 2895		Asn	Ala
Ala			Ala	Ile	Phe	Thr 2905	Asp					Lys	Ser	Ala
Arg		Ala	Val	Glu	Asp	Ser 2920	Arg		Trp				Asp	Lys
Glu			Leu	His	Leu	2920 Glu 2935	Gly	Lys					Val	Tyr
Asn	Met					Glu	Lys			Gly	Glu		Gly	Lys
Ala				Arg		2950 Ile	Trp						Ala	Arg
Phe			Phe	Glu	Ala	2965 Leu	Gly			Asn	Glu		His	Trp
Phe	2975 Ser	Arg	Glu	Asn	Ser	2980 Leu		Gly	Val	Glu	2985 Gly	Glu	Gly	Leu
His	2990 Lys		Gly	Tyr	Ile	2995 Leu		Asp	Val	Ser	3000 Lys		Glu	Gly
	3005		-	-		3010 Asp	-	-			3015	-		-
-	3020		-		-	3025			-	-	3030		-	
	3035					Asn 3040					3045			
	3050			-	-	Leu 3055					3060	-		
Tyr	Gln 3065	Asn	Lys	Val	Val	Arg 3070		Gln	Arg	Pro	Thr 3075	Pro	Arg	Gly
Thr	Val 3080	Met	Asp	Ile	Ile	Ser 3085	-	Arg	Asp	Gln	Arg 3090	-	Ser	Gly
Gln	Val 3095	-	Thr	Tyr	Gly	Leu 3100		Thr	Phe	Thr	Asn 3105		Glu	Ala
Gln	Leu 3110	Ile	Arg	Gln	Met	Glu 3115	-	Glu	Gly	Val	Phe 3120	-	Ser	Ile
Gln	His 3125	Leu	Thr	Ile	Thr	Glu 3130		Ile	Ala	Val	Gln 3135	Asn	Trp	Leu
Ala		Val	Gly	Arg	Glu	Arg		Ser	Arg	Met		Ile	Ser	Gly

-continued

3140 3145	3150
Asp Asp Cys Val Val Lys Pro Leu Asp Asp Arg	Phe Ala Ser Ala
3155 3160	3165
Leu Thr Ala Leu Asn Asp Met Gly Lys Ile Arg	Lys Asp Ile Gln
3170 3175	3180
Gln Trp Glu Pro Ser Arg Gly Trp Asn Asp Trp	Thr Gln Val Pro
3185 3190	3195
Phe Cys Ser His His Phe His Glu Leu Ile Met	Lys Asp Gly Arg
3200 3205	3210
Val Leu Val Val Pro Cys Arg Asn Gln Asp Glu	Leu Ile Gly Arg
3215 3220	3225
Ala Arg Ile Ser Gln Gly Ala Gly Trp Ser Leu	Arg Glu Thr Ala
3230 3235	3240
Cys Leu Gly Lys Ser Tyr Ala Gln Met Trp Ser	Leu Met Tyr Phe
3245 3250	3255
His Arg Arg Asp Leu Arg Leu Ala Ala Asn Ala	Ile Cys Ser Ala
3260 3265	3270
Val Pro Ser His Trp Val Pro Thr Ser Arg Thr	Thr Trp Ser Ile
3275 3280	3285
His Ala Lys His Glu Trp Met Thr Thr Glu Asp	Met Leu Thr Val
3290 3295	3300
Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp	Met Glu Asp Lys
3305 3310	3315
Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr	Leu Gly Lys Arg
3320 3325	3330
Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu	Thr Ser Arg Ala
3335 3340	3345
Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn	Gln Val Arg Ser
3350 3355	3360
Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met	Pro Ser Met Lys
3365 3370	3375
Arg Phe Arg Arg Glu Glu Glu Glu Ala Gly Val	Leu Trp
3380 3385	3390
<pre>&lt;210&gt; SEQ ID NO 4 &lt;211&gt; LENGTH: 10723 &lt;212&gt; TYPE: DNA &lt;211&gt; ORGANISM: Artificial Sequence &lt;220&gt; FEATURE: &lt;223&gt; OTHER INFORMATION: Dengue virus serotype &lt;400&gt; SEQUENCE: 4</pre>	e 2, BVS
agttgttagt ctacgtggac cgacaaagac agattctttg ag	gggagctaa gctcaatgta 60
gttctaacag ttttttaatt agagagcaga tctctgatga a	taaccaacg gaaaaaggcg 120
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc g	cgtgtcgac tgtgcaacag 180
ctgacaaaga gattctcact tggaatgctg cagggacgag g	accattaaa actgttcatg 240
gccctggtgg cgttccttcg tttcctaaca atcccaccaa c	agcagggat attgaagaga 300
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga g	agggttcag gaaagagatt 360
ggaaggatgc tgaacatctt gaataggaga cgcagatctg c	aggcatgat cattatgctg 420
attocaacag tgatggogtt ocatttaaco acaogtaaog g	agaaccaca catgatcgtc 480

				- conc 11			
agcagacaag	agaaagggaa	aagtcttctg	tttaaaacag	aggttggcgt	gaacatgtgt	540	
accctcatgg	ccatggacct	tggtgaattg	tgtgaagaca	caatcacgta	cgagtgtccc	600	
cttctcaggc	agaatgagcc	agaagacata	gactgttggt	gcaactctac	gtccacgtgg	660	
gtaacttatg	ggacgtgtac	caccatggga	gaacatagaa	gagaaaaaag	atcagtggca	720	
ctcgttccac	atgtgggaat	gggactggag	acacgaactg	aaacatggat	gtcatcagaa	780	
ggggcctgga	aacatgtcca	gagaattgaa	acttggatct	tgagacatcc	aggetteace	840	
atgatggcag	caatcctggc	atacaccata	ggaacgacac	atttccaaag	agccctgatc	900	
ttcatcttac	tgacagctgt	cactccttca	atgacaatgc	gttgcatagg	aatgtcaaat	960	
agagactttg	tggaaggggt	ttcaggagga	agctgggttg	acatagtett	agaacatgga	1020	
agctgtgtga	cgacgatggc	aaaaaacaaa	ccaacattgg	attttgaact	gataaaaaca	1080	
gaagccaaac	agcctgccac	cctaaggaag	tactgtatag	aggcaaagct	aaccaacaca	1140	
acaacagaat	ctcgctgccc	aacacaaggg	gaacccagcc	taaatgaaga	gcaggacaaa	1200	
aggttcgtct	gcaaacactc	catggtagac	agaggatggg	gaaatggatg	tggactattt	1260	
ggaaagggag	gcattgtgac	ctgtgctatg	ttcagatgca	aaaagaacat	ggaaggaaaa	1320	
gttgtgcaac	cagaaaactt	ggaatacacc	attgtgataa	cacctcactc	aggggaagag	1380	
catgcagtcg	gaaatgacac	aggaaaacat	ggcaaggaaa	tcaaaataac	accacagagt	1440	
tccatcacag	aagcagaatt	gacaggttat	ggcactgtca	caatggagtg	ctctccaaga	1500	
acgggcctcg	acttcaatga	gatggtgttg	ctgcagatgg	aaaataaagc	ttggctggtg	1560	
cacaggcaat	ggttcctaga	cctgccgtta	ccatggttgc	ccggagcgga	cacacaaggg	1620	
tcaaattgga	tacagaaaga	gacattggtc	actttcaaaa	atccccatgc	gaagaaacag	1680	
gatgttgttg	ttttaggatc	ccaagaaggg	gccatgcaca	cagcacttac	aggggccaca	1740	
gaaatccaaa	tgtcatcagg	aaacttactc	ttcacaggac	atctcaagtg	caggctgaga	1800	
atggacaagc	tacagctcaa	aggaatgtca	tactctatgt	gcacaggaaa	gtttaaagtt	1860	
gtgaaggaaa	tagcagaaac	acaacatgga	acaatagtta	tcagagtgca	atatgaaggg	1920	
gacggctctc	catgcaagat	cccttttgag	ataatggatt	tggaaaaaag	acatgtctta	1980	
ggtcgcctga	ttacagtcaa	cccaattgtg	acagaaaaag	atagcccagt	caacatagaa	2040	
gcagaacctc	catttggaga	cagctacatc	atcataggag	tagagccggg	acaactgaag	2100	
ctcaactggt	ttaagaaagg	aagttctatc	ggccaaatgt	ttgagacaac	aatgaggggg	2160	
gcgaagagaa	tggccatttt	aggtgacaca	gcctgggatt	ttggatcctt	gggaggagtg	2220	
tttacatcta	taggaaaggc	tctccaccaa	gtctttggag	caatctatgg	agctgccttc	2280	
agtggggttt	catggactat	gaaaatcctc	ataggagtca	ttatcacatg	gataggaatg	2340	
aattcacgca	gcacctcact	gtctgtgaca	ctagtattgg	tgggaattgt	gacactgtat	2400	
ttgggagtca	tggtgcaggc	cgatagtggt	tgcgttgtga	gctggaaaaa	caaagaactg	2460	
aaatgtggca	gtgggatttt	catcacagac	aacgtgcaca	catggacaga	acaatacaag	2520	
ttccaaccag	aatccccttc	aaaactagct	tcagctatcc	agaaagccca	tgaagaggac	2580	
atttgtggaa	tccgctcagt	aacaagactg	gagaatctga	tgtggaaaca	aataacacca	2640	
gaattgaatc	acattctatc	agaaaatgag	gtgaagttaa	ctattatgac	aggagacatc	2700	
aaaggaatca	tgcaggcagg	aaaacgatct	ctgcggcctc	agcccactga	gctgaagtat	2760	
		5		- 0			

-continued	
tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg	2880
gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa	2940
aaacaggatg tattetgega etcaaaaete atgteagegg eeataaaaga eaacagagee	3000
gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag	3060
aaagcetett teattgaagt taaaaaetge eaetggeeaa aateaeaeae eetetggage	3120
aatggagtgc tagaaagtga gatgataatt ccaaagaatc tcgctggacc agtgtctcaa	3180
cacaactata gaccaggcta ccatacacaa ataacaggac catggcatct aggtaagctt	3240
gagatggact ttgatttctg tgatggaaca acagtggtag tgactgagga ctgcggaaat	3300
agaggaccct ctttgagaac aaccactgcc tctggaaaac tcataacaga atggtgctgc	3360
cgatettgea cattaceace getaagatae agaggtgagg atgggtgetg gtaegggatg	3420
gaaatcagac cattgaagga gaaagaagag aatttggtca actccttggt cacagctgga	3480
catgggcagg tcgacaactt ttcactagga gtcttgggaa tggcattgtt cctggaggaa	3540
atgettagga eeegagtagg aacgaaacat geaataetae tagttgeagt ttettttgtg	3600
acattgatca cagggaacat gtcctttaga gacctgggaa gagtgatggt tatggtaggc	3660
gccactatga cggatgacat aggtatgggc gtgacttatc ttgccctact agcagccttc	3720
aaagtcagac caacttttgc agctggacta ctcttgagaa agctgacctc caaggaattg	3780
atgatgacta ctataggaat tgtactcctc tcccagagca ccataccaga gaccattctt	3840
gagttgactg atgcgttagc cttaggcatg atggtcctca aaatggtgag aaatatggaa	3900
aagtatcaat tggcagtgac tatcatggct atcttgtgcg tcccaaacgc agtgatatta	3960
caaaacgcat ggaaagtgag ttgcacaata ttggcagtgg tgtccgtttc cccactgttc	4020
ttaacateet cacageaaaa aacagattgg ataeeattag cattgaegat caaaggtete	4080
aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca	4140
ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa	4200
aatgatattc ccatgacagg accattagtg gctggagggc teeteactgt gtgetaegtg	4260
ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac	4320
caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg	4440
ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg	4500
tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg	4560
ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat	4620
teecagateg gageeggagt ttacaaagaa ggaacattee atacaatgtg geatgteaca	4680
cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa	4800
gaagteeagg tattggeact ggageetgga aaaaateeaa gageegteea aaegaaaeet	4860
ggtettttea aaaceaaege eggaacaata ggtgetgtat etetggaett tteteetgga	4920
acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt	4980
gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa	5040

				-contir		
gacaacccag	agatcgaaga	tgacattttc	cgaaagagaa	gactgaccat	catggacctc	5100
cacccaggag	cgggaaagac	gaagagatac	cttccggcca	tagtcagaga	agctataaaa	5160
cggggtttga	gaacattaat	cttggccccc	actagagttg	tggcagctga	aatggaggaa	5220
gcccttagag	gacttccaat	aagataccag	accccagcca	tcagagctgt	gcacaccggg	5280
cgggagattg	tggacctaat	gtgtcatgcc	acatttacca	tgaggctgct	atcaccagtt	5340
agagtgccaa	actacaacct	gattatcatg	gacgaagccc	atttcacaga	cccagcaagt	5400
atagcagcta	gaggatacat	ctcaactcga	gtggagatgg	gtgaggcagc	tgggattttt	5460
atgacagcca	ctcccccggg	aagcagagac	ccatttcctc	agagcaatgc	accaatcata	5520
gatgaagaaa	gagaaatccc	tgaacgctcg	tggaattccg	gacatgaatg	ggtcacggat	5580
tttaaaggga	agactgtttg	gttcgttcca	agtataaaag	caggaaatga	tatagcagct	5640
tgcctgagga	aaaatggaaa	gaaagtgata	caactcagta	ggaagacctt	tgattctgag	5700
tatgtcaaga	ctagaaccaa	tgattgggac	ttcgtggtta	caactgacat	ttcagaaatg	5760
ggtgccaatt	tcaaggctga	gagggttata	gaccccagac	gctgcatgaa	accagtcata	5820
ctaacagatg	gtgaagagcg	ggtgattctg	gcaggaccta	tgccagtgac	ccactctagt	5880
gcagcacaaa	gaagagggag	aataggaaga	aatccaaaaa	atgagaatga	ccagtacata	5940
tacatggggg	aacctctgga	aaatgatgaa	gactgtgcac	actggaaaga	agctaaaatg	6000
ctcctagata	acatcaacac	gccagaagga	atcattccta	gcatgttcga	accagagcgt	6060
gaaaaggtgg	atgccattga	tggcgaatac	cgcttgagag	gagaagcaag	gaaaaccttt	6120
gtagacttaa	tgagaagagg	agacctacca	gtctggttgg	cctacagagt	ggcagctgaa	6180
ggcatcaact	acgcagacag	aaggtggtgt	tttgatggag	tcaagaacaa	ccaaatccta	6240
gaagaaaacg	tggaagttga	aatctggaca	aaagaagggg	aaaggaagaa	attgaaaccc	6300
agatggttgg	atgctaggat	ctattctgac	ccactggcgc	taaaagaatt	taaggaattt	6360
gcagccggaa	gaaagtctct	gaccctgaac	ctaatcacag	aaatgggtag	gctcccaacc	6420
ttcatgactc	agaaggcaag	agacgcactg	gacaacttag	cagtgctgca	cacggctgag	6480
ccaggtggaa	gggcgtacaa	ccatgctctc	agtgaactgc	cggagaccct	ggagacattg	6540
cttttactga	cacttctggc	tacagtcacg	ggagggatct	ttttattctt	gatgagcgca	6600
aggggcatag	ggaagatgac	cctgggaatg	tgctgcataa	tcacggctag	catcctccta	6660
tggtacgcac	aaatacagcc	acactggata	gcagcttcaa	taatactgga	gttttttctc	6720
atagttttgc	ttattccaga	acctgaaaaa	cagagaacac	cccaagacaa	ccaactgacc	6780
tacgttgtca	tagccatcct	cacagtggtg	gccgcaacca	tggcaaacga	gatgggtttc	6840
ctagaaaaaa	cgaagaaaga	tctcggattg	ggaagcattg	caacccagca	acccgagagc	6900
aacatcctgg	acatagatct	acgtcctgca	tcagcatgga	cgctgtatgc	cgtggccaca	6960
acatttgtta	caccaatgtt	gagacatagc	attgaaaatt	cctcagtgaa	tgtgtcccta	7020
acagctatag	ccaaccaagc	cacagtgtta	atgggtctcg	ggaaaggatg	gccattgtca	7080
aagatggaca	tcggagttcc	ccttctcgcc	attggatgct	actcacaagt	caaccccata	7140
actctcacag	cagctttttt	cttattggta	gcacattatg	ccatcatagg	gccaggactc	7200
caagcaaaag	caaccagaga	agctcagaaa	agagcagcgg	cgggcatcat	gaaaaaccca	7260
actgtcgatg	gaataacagt	gattgaccta	gatccaatac	cttatgatcc	aaagtttgaa	7320

		-concinued	
aagcagttgg gacaagtaat	gctcctagtc ctctgcgtga	ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga	ggctttaacc ttagctaccg	ggcccatctc cacattgtgg	7440
gaaggaaatc cagggaggtt	ttggaacact accattgcgg	tgtcaatggc taacattttt	7500
agagggagtt acttggccgg	agctggactt ctcttttcta	ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa	. cataggagag acgcttggag	agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga	. attccagatc tacaagaaaa	gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg	cattaaaaga ggagaaacgg	accatcacgc tgtgtcgcga	7740
ggctcagcaa aactgagatg	gttcgttgag agaaacatgg	tcacaccaga agggaaagta	7800
gtggacctcg gttgtggcag	aggaggctgg tcatactatt	gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac	aaaaggagga ccaggacacg	aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt	gcgtcttcaa agtggagttg	acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt	gtgtgacata ggggagtcat	caccaaatcc cacagtggaa	8040
gcaggacgaa cactcagagt	ccttaactta gtagaaaatt	ggttgaacaa caacactcaa	8100
ttttgcataa aggttctcaa	. cccatatatg ccctcagtca	tagaaaaaat ggaagcacta	8160
caaaggaaat atggaggagc	cttagtgagg aatccactct	cacgaaactc cacacatgag	8220
atgtactggg tatccaatgo	ttccgggaac atagtgtcat	cagtgaacat gatttcaagg	8280
atgttgatca acagatttac	aatgagatac aagaaagcca	cttacgagcc ggatgttgac	8340
ctcggaagcg gaacccgtaa	. catcgggatt gaaagtgaga	taccaaacct agatataatt	8400
gggaaaagaa tagaaaaaat	aaagcaagag catgaaacat	catggcacta tgaccaagac	8460
cacccataca aaacgtgggc	ataccatggt agctatgaaa	caaaacagac tggatcagca	8520
tcatccatgg tcaacggagt	ggtcaggctg ctgacaaaac	cttgggacgt cgtccccatg	8580
gtgacacaga tggcaatgac	agacacgact ccatttggac	aacagcgcgt ttttaaagag	8640
aaagtggaca cgagaaccca	. agaaccgaaa gaaggcacga	agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga	. attagggaag aaaaagacac	ccaggatgtg caccagagaa	8760
gaattcacaa gaaaggtgag	aagcaatgca gccttggggg	ccgtattcac tgatgagaac	8820
aagtggaagt cggcacgtga	. ggctgttgaa gatagtaggt	tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga	. aggaaagtgt gaaacatgtg	tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga	. attcggcaag gcaaaaggca	gcagagccat atggtacatg	9000
tggcttggag cacgcttctt	agagtttgaa geeetaggat	tcttaaatga agatcactgg	9060
ttetecagag agaacteect	gagtggagtg gaaggagaag	ggctgcacaa gctaggttac	9120
attctaagag acgtgagcaa	. gaaagaggga ggagcaatgt	atgeegatga caeegeagga	9180
tgggatacaa gaatcacact	agaagaccta aaaaatgaag	aaatggtaac aaaccacatg	9240
gaaggagaac acaagaaact	agccgaggcc attttcaaac	taacgtacca aaacaaggtg	9300
gtgcgtgtgc aaagaccaac	accaagaggc acagtaatgg	acatcatatc gagaagagac	9360
caaagaggta gtggacaagt	tggcacctat ggactcaata	ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga	. gggagaagga gtctttaaaa	gcattcagca cctaacaatc	9480
acagaagaaa tcgctgtgca	. aaactggtta gcaagagtgg	ggcgcgaaag gttatcaaga	9540
atggccatca gtggagatga	. ttgtgttgtg aaacctttag	atgacaggtt cgcaagcgct	9600
0	J	0	

												<u> </u>	<u></u>			 	 	
ttaacag	ctc t	caaat	gaca	at go	ggaaa	agatt	: ago	gaaag	jaca	taca	aacaa	atg g	ggaad	ccttca	9660			
agaggat	gga a	atgat	tgga	ac ac	caagt	geed	tt:	ctgtt	cac	acca	attto	cca t	gagt	taatc	9720			
atgaaag	acg q	gtcgo	gtad	ct co	gttgt	tcca	a tgi	tagaa	acc	aaga	atgaa	act ç	gatto	ggcaga	9780			
gcccgaa	tct d	cccaa	aggag	gc ag	gggtg	ggtct	t tt	geggg	jaga	cggo	cctgt	ttt ç	adda	agtct	9840			
tacgccc	aaa t	gtgg	gaget	t ga	atgta	actto	c ca	cagac	gcg	acct	cago	get g	ggcgg	gcaaat	9900			
gctattt	gct (	cggca	agtad	cc at	ccaca	attg	g gti	tccaa	icaa	gtco	gaaca	aac d	stggt	ccata	9960			
catgcta	aac a	atgaa	atgga	at ga	acaad	cggaa	a gao	catgo	tga	cagt	cctgg	gaa d	aggg	gtgtgg	10020			
attcaag	aaa a	accca	atgga	at go	gaaga	acaaa	a act	tccag	ıtgg	aato	catgo	gga g	ggaaa	atccca	10080			
tacttgg	gga a	aaaga	agaag	ga co	caato	ggtgo	g g g	ctcat	tga	ttgg	ggtta	aac a	aagca	agggcc	10140			
acctggg	caa a	agaad	catco	ca aç	gcago	caata	a aat	tcaag	gtta	gato	ccctt	tat a	aggca	aatgaa	10200			
gaataca	.cag a	attad	catgo	cc at	cccat	cgaaa	a aga	attca	igaa	gaga	aagaq	gga a	agaag	gcagga	10260			
gttctgt	ggt a	agaaa	agcaa	aa ad	ctaad	catga	a aa	caagg	jcta	gaaq	gtcaç	ggt d	eggat	taagc	10320			
catagta	.cgg a	aaaaa	acta	at go	ctaco	ctgtg	g ago	acace	gtcc	aago	gacgt	tta a	aaaga	aagtca	10380			
ggccato	ata a	aatgo	ccata	ag ct	tgag	gtaaa	a cta	atgca	ıgcc	tgta	ageto	cca d	cctga	agaagg	10440			
tgtaaaa	aat d	ccggg	gaggo	cc ad	caaad	ccato	g gaa	agctg	gtac	gcat	ggcę	gta ç	gtgga	actagc	10500			
ggttaga	.gga q	gacco	cctco	cc tt	cacaa	aatco	g caq	gcaac	aat	gggg	ggcco	caa ç	ggcga	agatga	10560			
agctgta	.gtc t	cget	ggaa	ag ga	actaç	gaggt	t taq	gagga	igac	ccc	cccga	aaa d	caaaa	aacag	10620			
catattg	acg d	ctggg	gaaag	ga co	cagag	gatco	tgo	ctgto	tcc	tcaç	gcato	cat t	ccag	ggcaca	10680			
gaacgcc	aga a	aaato	ggaat	cg gt	gete	gttga	a ato	caaca	ıggt	tct					10723			
<211> L <212> T <213> C <220> F <223> C	YPE : RGAN EATUI	PRT ISM: RE:	Arti					<b>1</b> 8 ຣ∈	eroty	/pe 2	2, B1	/S						
<400> S	EQUEI	ICE :	5															
Met Asn 1	Asn	Gln	Arg 5	Lys	Lys	Ala	Lys	Asn 10	Thr	Pro	Phe	Asn	Met 15	Leu				
Lys Arg	Glu	-	Asn	Arg	Val	Ser		Val	Gln	Gln	Leu		Lys	Arg				
Phe Ser	In	20 Glv	Mot	Lor	<u>61</u> ~	c1	25 Arg	G1	Dro	Lor	Lare	30 Leu	Dh a	Mo+				
rne ser	35	σιγ	met	цец	GTU	40 40	нrg	σтγ	PTO	цец	Lys 45	цец	File	Met				
Ala Leu 50	. Val	Ala	Phe	Leu	Arg 55	Phe	Leu	Thr	Ile	Pro 60	Pro	Thr	Ala	Gly				
Ile Leu 65	. Lys	Arg	Trp	Gly 70	Thr	Ile	Lys	Lys	Ser 75	Lys	Ala	Ile	Asn					
	<b>C</b> 1	D'	<b>n</b>		<b>a</b> 1	<b>T</b> 7	<b>a</b> 7	7		T -	<b>n</b>	<b>T</b> 7	T -	80				
Leu Arg	GIY	гhе	Arg 85	гда	GIU	тте	σту	Arg 90	met	ьeu	Asn	тте	Leu 95	Asn				
Arg Arg	Arg	Arg 100	Ser	Ala	Gly	Met	Ile 105	Ile	Met	Leu	Ile	Pro 110	Thr	Val				
Met Ala	. Phe 115	His	Leu	Thr	Thr	Arg 120	Asn	Gly	Glu	Pro	His 125	Met	Ile	Val				
Ser Arg	Gln	Glu	Lvs	Gly	<b>T</b>		• • • •	T	Dlas	Lare	Thr	Glu	Val	Gly				
-			цур	Gry	-	Ser	Leu	ьец	Pne	-	TIIT	Gru	var	Gry				
130 Val Asr			-	-	135					140				-				

-continued

145					150					155					160
Asp	Thr	Ile	Thr	Tyr 165	Glu	Сув	Pro	Leu	Leu 170	Arg	Gln	Asn	Glu	Pro 175	Glu
Asp	Ile	Asp	Cys 180	Trp	Суз	Asn	Ser	Thr 185	Ser	Thr	Trp	Val	Thr 190	Tyr	Gly
Thr	Суз	Thr 195	Thr	Met	Gly	Glu	His 200	Arg	Arg	Glu	Lys	Arg 205	Ser	Val	Ala
Leu	Val 210	Pro	His	Val	Gly	Met 215	Gly	Leu	Glu	Thr	Arg 220	Thr	Glu	Thr	Trp
Met 225	Ser	Ser	Glu	Gly	Ala 230	Trp	Lys	His	Val	Gln 235	Arg	Ile	Glu	Thr	Trp 240
Ile	Leu	Arg	His	Pro 245	Gly	Phe	Thr	Met	Met 250	Ala	Ala	Ile	Leu	Ala 255	Tyr
Thr	Ile	Gly	Thr 260	Thr	His	Phe	Gln	Arg 265	Ala	Leu	Ile	Phe	Ile 270	Leu	Leu
Thr	Ala	Val 275	Thr	Pro	Ser	Met	Thr 280	Met	Arg	Суа	Ile	Gly 285	Met	Ser	Asn
Arg	Asp 290		Val	Glu	Gly	Val 295		Gly	Gly	Ser	Trp 300		Asp	Ile	Val
Leu 305		His	Gly	Ser	Суз 310		Thr	Thr	Met	Ala 315		Asn	Lys	Pro	Thr 320
	Asp	Phe	Glu	Leu 325	Ile	Lys	Thr	Glu	Ala 330		Gln	Pro	Ala	Thr 335	
Arg	Lys	Tyr			Glu	Ala	Lys			Asn	Thr	Thr			Ser
Arg	Суз		340 Thr	Gln	Gly	Glu		345 Ser	Leu	Asn	Glu		350 Gln	Asp	Lys
Arg		355 Val	Суз	Lys	His	Ser	360 Met	Val	Asp	Arg	Gly	365 Trp	Gly	Asn	Gly
Cys	370 Gly	Leu	Phe	Gly	Lys	375 Gly	Gly	Ile	Val	Thr	380 Cys	Ala	Met	Phe	Arg
385 Cvs	Lvs	Lvs	Asn	Met	390 Glu	Glv	Lvs	Val	Val	395 Gln	Pro	Glu	Asn	Leu	400 Glu
-	-	-		405	Thr	-	-		410					415	
-			420					425	-				430		-
	-	435	-	-	His	-	440			-		445			
Ser	Ile 450	Thr	Glu	Ala	Glu	Leu 455		Gly	Tyr	Gly	Thr 460	Val	Thr	Met	Glu
Cys 465	Ser	Pro	Arg	Thr	Gly 470	Leu	Asp	Phe	Asn	Glu 475	Met	Val	Leu	Leu	Gln 480
Met	Glu	Asn	Гла	Ala 485	Trp	Leu	Val	His	Arg 490	Gln	Trp	Phe	Leu	Asp 495	Leu
Pro	Leu	Pro	Trp 500	Leu	Pro	Gly	Ala	Asp 505	Thr	Gln	Gly	Ser	Asn 510	Trp	Ile
Gln	Lys	Glu 515	Thr	Leu	Val	Thr	Phe 520	Lys	Asn	Pro	His	Ala 525	Lys	Lys	Gln
Asp	Val 530	Val	Val	Leu	Gly	Ser 535	Gln	Glu	Gly	Ala	Met 540	His	Thr	Ala	Leu
		Ala	Thr	Glu	Ile		Met	Ser	Ser	Gly 555		Leu	Leu	Phe	
545					550					555					560

\_\_\_\_\_

-continued
------------

Gly	His	Leu	Lys	Cys 565	Arg	Leu	Arg	Met	Asp 570	Lys	Leu	Gln	Leu	Lys 575	Gly
Met	Ser	Tyr	Ser 580	Met	Cya	Thr	Gly	Lys 585	Phe	Lys	Val	Val	Lys 590	Glu	Ile
Ala	Glu	Thr 595	Gln	His	Gly	Thr	Ile 600	Val	Ile	Arg	Val	Gln 605	Tyr	Glu	Gly
Asp	Gly 610	Ser	Pro	Суз	LÀa	Ile 615	Pro	Phe	Glu	Ile	Met 620	Asp	Leu	Glu	Lys
Arg 625	His	Val	Leu	Gly	Arg 630	Leu	Ile	Thr	Val	Asn 635	Pro	Ile	Val	Thr	Glu 640
Lys	Asp	Ser	Pro	Val 645	Asn	Ile	Glu	Ala	Glu 650	Pro	Pro	Phe	Gly	Asp 655	Ser
Tyr	Ile	Ile	Ile 660	Gly	Val	Glu	Pro	Gly 665	Gln	Leu	Lys	Leu	Asn 670	Trp	Phe
L'Aa	Lys	Gly 675	Ser	Ser	Ile	Gly	Gln 680	Met	Phe	Glu	Thr	Thr 685	Met	Arg	Gly
Ala	Lys 690	Arg	Met	Ala	Ile	Leu 695	Gly	Aap	Thr	Ala	Trp 700	Aap	Phe	Gly	Ser
Leu 705	Gly	Gly	Val	Phe	Thr 710	Ser	Ile	Gly	Lys	Ala 715	Leu	His	Gln	Val	Phe 720
Gly	Ala	Ile	Tyr	Gly 725	Ala	Ala	Phe	Ser	Gly 730	Val	Ser	Trp	Thr	Met 735	Lya
Ile	Leu	Ile	Gly 740	Val	Ile	Ile	Thr	Trp 745	Ile	Gly	Met	Asn	Ser 750	Arg	Ser
Thr	Ser	Leu 755	Ser	Val	Thr	Leu	Val 760	Leu	Val	Gly	Ile	Val 765	Thr	Leu	Tyr
Leu	Gly 770	Val	Met	Val	Gln	Ala 775	Asp	Ser	Gly	Суз	Val 780	Val	Ser	Trp	Lys
Asn 785	Lys	Glu	Leu	Lys	Cys 790	Gly	Ser	Gly	Ile	Phe 795	Ile	Thr	Aab	Asn	Val 800
His	Thr	Trp	Thr	Glu 805	Gln	Tyr	Lys	Phe	Gln 810	Pro	Glu	Ser	Pro	Ser 815	ГЛа
Leu	Ala	Ser	Ala 820	Ile	Gln	Lys	Ala	His 825	Glu	Glu	Asp	Ile	Cys 830	Gly	Ile
Arg	Ser	Val 835	Thr	Arg	Leu	Glu	Asn 840	Leu	Met	Trp	Lys	Gln 845	Ile	Thr	Pro
Glu	Leu 850	Asn	His	Ile	Leu	Ser 855	Glu	Asn	Glu	Val	Lys 860	Leu	Thr	Ile	Met
Thr 865	Gly	Asp	Ile	Lys	Gly 870	Ile	Met	Gln	Ala	Gly 875	Lys	Arg	Ser	Leu	Arg 880
Pro	Gln	Pro	Thr	Glu 885	Leu	Lys	Tyr	Ser	Trp 890	Lys	Thr	Trp	Gly	Lys 895	Ala
Lys	Met	Leu	Ser 900	Thr	Glu	Ser	His	Asn 905	Gln	Thr	Phe	Leu	Ile 910	Asb	Gly
Pro	Glu	Thr 915	Ala	Glu	Cys	Pro	Asn 920	Thr	Asn	Arg	Ala	Trp 925	Asn	Ser	Leu
Glu	Val 930	Glu	Aap	Tyr	Gly	Phe 935	Gly	Val	Phe	Thr	Thr 940	Asn	Ile	Trp	Leu
Lys 945	Leu	Lys	Glu	Lys	Gln 950	Asp	Val	Phe	Cys	Asp 955	Ser	Lys	Leu	Met	Ser 960

con			

												001	10 11		<i>а</i>
Ala	Ala	Ile	Lys	Asp 965	Asn	Arg	Ala	Val	Hi: 970		La A	sp Met	Gl	у Ту: 97!	-
Ile	Glu	Ser	Ala 980	Leu	Asn	Asp	Thr	Trp 985	Гλε	; I]	Le G	lu Ly:	8 Al. 99		r Phe
Ile	Glu	Val 995	Lys	Asn	Cys	His	Trp 1000		о Цу	s S	Ser		nr 1 )05	Leu '	Trp Sei
Asn	Gly 1010		L Lei	ı Glu	ı Sei	r Glu 101		et I	le I	le	Pro	Lys 1020	Asn	Leu	Ala
Gly	Pro 1025		L Sei	r Glı	n His	s Asr 103	-	yr A	rg I	ro	Gly	Tyr 1035	His	Thr	Gln
Ile	Thr 1040	-	/ Pro	o Trj	p His	s Leu 104		ly L	ys I	eu	Glu	Met 1050	Asp	Phe	Asp
Phe	Cys 1059		Gl	y Thi	r Thi	r Va] 106		al V	al T	'hr	Glu	Asp 1065	Суз	Gly	Asn
Arg	Gly 1070		Sei	r Leu	ı Arç	g Thi 107		nr T	nr A	la	Ser	Gly 1080	ГÀа	Leu	Ile
Thr	Glu 1089	_	o Cyr	а Суя	a Arç	g Sei 109		ys Tl	nr I	eu	Pro	Pro 1095	Leu	Arg	Tyr
Arg	Gly 1100		ı Asl	p Gly	7 Cys	s Trp 110		yr G	ly M	let	Glu	Ile 1110	Arg	Pro	Leu
Lys	Glu 1119	-	s Glu	ı Glu	ı Ası	n Leu 112		al A	sn S	er	Leu	Val 1125	Thr	Ala	Gly
His	Gly 1130		n Val	l Asp	o Ası	n Phe 113		∋r L	eu G	ly	Val	Leu 1140	Gly	Met	Ala
Leu	Phe 1149		ı Glu	ı Glu	ı Met	: Leu 115		rg T	nr A	rg	Val	Gly 1155	Thr	Lys	His
Ala	Ile 1160		ı Leı	ı Val	L Alá	a Val 116		er Pl	he V	al	Thr	Leu 1170	Ile	Thr	Gly
Asn	Met 1179		c Phe	e Arg	g Ası	p Leu 118		ly A	rg V	al	Met	Val 1185	Met	Val	Gly
Ala	Thr 1190		: Thi	r Ası	o Asi	9 Ile 119		ly M	et G	ly	Val	Thr 1200	Tyr	Leu	Ala
Leu	Leu 1209		a Ala	a Phe	e Ly:	s Val 121		rg P	ro I	'hr	Phe	Ala 1215	Ala	Gly	Leu
Leu	Leu 1220	-	g Ly:	s Leu	ı Thi	r Sei 122	-	ys G	lu I	eu	Met	Met 1230	Thr	Thr	Ile
Gly	Ile 1239		L Lei	ı Leı	ı Sei	r Glr 124		er Ti	hr I	le	Pro	Glu 1245	Thr	Ile	Leu
Glu	Leu 1250		r Asl	ọ Ala	a Leu	ı Ala 125		eu G	ly M	let	Met	Val 1260	Leu	Lys	Met
Val	Arg 1269		n Met	t Glu	ı Lys	s Tyi 127		ln L	eu A	la	Val	Thr 1275	Ile	Met	Ala
Ile	Leu 1280	-	9 Val	l Pro	o Ası	n Ala 128		al I	le I	eu	Gln	Asn 1290	Ala	Trp	ГЛа
Val	Ser 1295	-	5 Thi	r Ile	e Lei	ı Ala 130		al V	al S	er	Val	Ser 1305	Pro	Leu	Phe
Leu	Thr 1310		: Sei	r Glr	ı Glr	n Lys 131		nr A	ab 1	rp	Ile	Pro 1320	Leu	Ala	Leu
Thr	Ile 1325	-	a Gly	y Lei	ı Ası	n Pro 133		nr A	la I	le	Phe	Leu 1335	Thr	Thr	Leu
Ser	Arg	Thi	: Sei	r Ly:	a Pà	s Arç	g Se	er T	rp F	ro	Leu	Asn	Glu	Ala	Ile

-continued

	1340					1345					1350			
Met	Ala 1355	Val	Gly	Met	Val		Ile	Leu	Ala	Ser		Leu	Leu	Гла
Asn	Asp 1370	Ile	Pro	Met	Thr	Gly 1375	Pro	Leu	Val	Ala	Gly 1380	Gly	Leu	Leu
Thr	Val 1385	Сүз	Tyr	Val	Leu	Thr 1390	Gly	Arg	Ser	Ala	Asp 1395		Glu	Leu
Glu	Arg 1400	Ala	Ala	Asp	Val	Lys 1405	-	Glu	Asp	Gln	Ala 1410		Ile	Ser
Gly	Ser 1415	Ser	Pro	Ile	Leu	Ser 1420	Ile	Thr	Ile	Ser	Glu 1425	Asp	Gly	Ser
Met	Ser 1430	Ile	Lys	Asn	Glu	Glu 1435	Glu	Glu	Gln	Thr	Leu 1440		Ile	Leu
Ile	Arg 1445	Thr	Gly	Leu	Leu	Val 1450	Ile	Ser	Gly	Leu	Phe 1455		Val	Ser
Ile	Pro 1460	Ile	Thr	Ala	Ala	Ala 1465		Tyr	Leu	Trp	Glu 1470	Val	Гла	Lys
Gln	Arg 1475	Ala	Gly	Val	Leu	Trp 1480	Asp	Val	Pro	Ser	Pro 1485		Pro	Met
Gly	Lys 1490	Ala	Glu	Leu	Glu	Asp 1495	Gly	Ala	Tyr	Arg	Ile 1500		Gln	Lys
Gly	Ile 1505	Leu	Gly	Tyr	Ser	Gln 1510	Ile	Gly	Ala	Gly	Val 1515		Lys	Glu
Gly	Thr 1520	Phe	His	Thr	Met	Trp 1525	His	Val	Thr	Arg	Gly 1530	Ala	Val	Leu
Met	His 1535	Lys	Gly	Lys	Arg	Ile 1540	Glu	Pro	Ser	Trp	Ala 1545	Asp	Val	Lys
Lys	Asp 1550	Leu	Ile	Ser	Tyr	Gly 1555	Gly	Gly	Trp	ГЛЗ	Leu 1560	Glu	Gly	Glu
Trp	Lys 1565	Glu	Gly	Glu	Glu	Val 1570	Gln	Val	Leu	Ala	Leu 1575	Glu	Pro	Gly
Lys	Asn 1580	Pro	Arg	Ala	Val	Gln 1585	Thr	Lys	Pro	Gly	Leu 1590	Phe	ГЛЗ	Thr
Asn	Ala 1595	Gly	Thr	Ile	Gly	Ala 1600	Val	Ser	Leu	Asp	Phe 1605	Ser	Pro	Gly
Thr	Ser 1610	Gly	Ser	Pro	Ile	Ile 1615	Asp	Lys	Lys	Gly	Lys 1620	Val	Val	Gly
Leu	Tyr 1625	Gly	Asn	Gly	Val	Val 1630	Thr	Arg	Ser	Gly	Ala 1635		Val	Ser
Ala	Ile 1640	Ala	Gln	Thr	Glu	Lys 1645	Ser	Ile	Glu	Asp	Asn 1650	Pro	Glu	Ile
Glu	Asp 1655	Asp	Ile	Phe	Arg	Lys 1660	Arg	Arg	Leu	Thr	Ile 1665	Met	Asp	Leu
His	Pro 1670	Gly	Ala	Gly	Lys	Thr 1675	Lys	Arg	Tyr	Leu	Pro 1680		Ile	Val
Arg	Glu 1685	Ala	Ile	Lys	Arg	Gly 1690	Leu	Arg	Thr	Leu	Ile 1695		Ala	Pro
Thr	Arg 1700	Val	Val	Ala	Ala	Glu 1705	Met	Glu	Glu	Ala	Leu 1710	Arg	Gly	Leu
Pro	Ile 1715	Arg	Tyr	Gln	Thr	Pro 1720	Ala	Ile	Arg	Ala	Val 1725		Thr	Gly

Arg	Glu 1730		Val	Asp	Leu	Met 1735	-	His	Ala	Thr	Phe 1740		Met	Arg
Leu	Leu 1745		Pro	Val	Arg	Val 1750		Asn	Tyr	Asn	Leu 1755		Ile	Met
Asp	Glu 1760		His	Phe	Thr	Asp 1765		Ala	Ser	Ile	Ala 1770	Ala	Arg	Gly
Tyr	Ile 1775		Thr	Arg	Val	Glu 1780		Gly	Glu	Ala	Ala 1785	Gly	Ile	Phe
Met	Thr 1790	Ala	Thr	Pro	Pro	Gly 1795		Arg	Asp	Pro	Phe 1800	Pro	Gln	Ser
Asn	Ala 1805		Ile	Ile	Asp	Glu 1810		Arg	Glu	Ile	Pro 1815		Arg	Ser
Trp	Asn 1820		Gly	His	Glu	Trp 1825		Thr	Asp	Phe	Lys 1830	Gly	ГÀа	Thr
Val	Trp 1835		Val	Pro	Ser	Ile 1840		Ala	Gly	Asn	Asp 1845		Ala	Ala
СЛа	Leu 1850	Arg	Гла	Asn	Gly	Lys 1855		Val	Ile	Gln	Leu 1860		Arg	LYa
Thr	Phe 1865	Asp	Ser	Glu	Tyr	Val 1870		Thr	Arg	Thr	Asn 1875		Trp	Asp
Phe	Val 1880	Val	Thr	Thr	Asp	Ile 1885	Ser	Glu	Met	Gly	Ala 1890		Phe	Lys
Ala	Glu 1895		Val	Ile	Asp	Pro 1900		Arg	Сүз	Met	Lys 1905	Pro	Val	Ile
Leu	Thr 1910	Asp	Gly	Glu	Glu	Arg 1915		Ile	Leu	Ala	Gly 1920	Pro	Met	Pro
Val	Thr 1925	His	Ser	Ser	Ala	Ala 1930		Arg	Arg	Gly	Arg 1935	Ile	Gly	Arg
Asn	Pro 1940		Asn	Glu	Asn	Asp 1945		Tyr	Ile	Tyr	Met 1950	Gly	Glu	Pro
Leu	Glu 1955	Asn	Asp	Glu	Asp	Cys 1960		His	Trp	Lys	Glu 1965	Ala	Гла	Met
Leu	Leu 1970	Asp	Asn	Ile	Asn	Thr 1975		Glu	Gly	Ile	Ile 1980	Pro	Ser	Met
Phe	Glu 1985	Pro	Glu	Arg	Glu	Lys 1990		Asp	Ala	Ile	Asp 1995	Gly	Glu	Tyr
Arg	Leu 2000	Arg	Gly	Glu	Ala	Arg 2005	Lys	Thr	Phe	Val	Asp 2010	Leu	Met	Arg
Arg	Gly 2015	Asp	Leu	Pro	Val	Trp 2020	Leu	Ala	Tyr	Arg	Val 2025	Ala	Ala	Glu
Gly	Ile 2030	Asn	Tyr	Ala	Asp	Arg 2035	Arg	Trp	Сув	Phe	Asp 2040	Gly	Val	Lys
Asn	Asn 2045	Gln	Ile	Leu	Glu	Glu 2050	Asn	Val	Glu	Val	Glu 2055	Ile	Trp	Thr
Lys	Glu 2060	Gly	Glu	Arg	Lys	Lys 2065	Leu	Lys	Pro	Arg	Trp 2070	Leu	Asp	Ala
Arg	Ile 2075	Tyr	Ser	Asp	Pro	Leu 2080		Leu	Lys	Glu	Phe 2085	Lys	Glu	Phe
Ala	Ala 2090	Gly	Arg	Гла	Ser	Leu 2095	Thr	Leu	Asn	Leu	Ile 2100	Thr	Glu	Met

Gly	Arg 2105	Leu	Pro	Thr	Phe	Met 2110	Thr	Gln	L'Aa	Ala	Arg 2115	Asp	Ala	Leu
Asp	Asn 2120		Ala	Val	Leu	His 2125	Thr	Ala	Glu	Pro	Gly 2130	Gly	Arg	Ala
Tyr	Asn 2135		Ala	Leu	Ser	Glu 2140	Leu	Pro	Glu	Thr	Leu 2145	Glu	Thr	Leu
Leu	Leu 2150	Leu	Thr	Leu	Leu	Ala 2155	Thr	Val	Thr	Gly	Gly 2160	Ile	Phe	Leu
Phe	Leu 2165		Ser	Ala	Arg	Gly 2170	Ile	Gly	Lys	Met	Thr 2175	Leu	Gly	Met
Суа	Cys 2180		Ile	Thr	Ala	Ser 2185	Ile	Leu	Leu	Trp	Tyr 2190	Ala	Gln	Ile
Gln	Pro 2195	His	Trp	Ile	Ala	Ala 2200	Ser	Ile	Ile	Leu	Glu 2205	Phe	Phe	Leu
Ile	Val 2210	Leu	Leu	Ile	Pro	Glu 2215	Pro	Glu	Lys	Gln	Arg 2220		Pro	Gln
Asp	Asn 2225	Gln	Leu	Thr	Tyr	Val 2230	Val	Ile	Ala	Ile	Leu 2235	Thr	Val	Val
Ala	Ala 2240	Thr	Met	Ala	Asn	Glu 2245	Met	Gly	Phe	Leu	Glu 2250	Lys	Thr	ГЛа
Lys	Asp 2255	Leu	Gly	Leu	Gly	Ser 2260	Ile	Ala	Thr	Gln	Gln 2265	Pro	Glu	Ser
Asn	Ile 2270	Leu	Asp	Ile	Asp	Leu 2275	Arg	Pro	Ala	Ser	Ala 2280	Trp	Thr	Leu
Tyr	Ala 2285	Val	Ala	Thr	Thr	Phe 2290	Val	Thr	Pro	Met	Leu 2295	Arg	His	Ser
Ile	Glu 2300	Asn	Ser	Ser	Val	Asn 2305	Val	Ser	Leu	Thr	Ala 2310	Ile	Ala	Asn
Gln	Ala 2315	Thr	Val	Leu	Met	Gly 2320	Leu	Gly	Lys	Gly	Trp 2325	Pro	Leu	Ser
Lys	Met 2330	Asp	Ile	Gly	Val	Pro 2335	Leu	Leu	Ala	Ile	Gly 2340	Суз	Tyr	Ser
Gln	Val 2345	Asn	Pro	Ile	Thr	Leu 2350	Thr	Ala	Ala	Phe	Phe 2355	Leu	Leu	Val
Ala	His 2360	Tyr	Ala	Ile	Ile	Gly 2365	Pro	Gly	Leu	Gln	Ala 2370	ГЛЗ	Ala	Thr
Arg	Glu 2375	Ala	Gln	Lys	Arg	Ala 2380	Ala	Ala	Gly	Ile	Met 2385	ГЛЗ	Asn	Pro
Thr	Val 2390	_	Gly	Ile	Thr	Val 2395	Ile	Asp	Leu	Asp	Pro 2400	Ile	Pro	Tyr
Asp	Pro 2405		Phe	Glu	Lys	Gln 2410	Leu	Gly	Gln	Val	Met 2415	Leu	Leu	Val
Leu	Cys 2420		Thr	Gln	Val	Leu 2425	Met	Met	Arg	Thr	Thr 2430	Trp	Ala	Leu
Сүз			Leu	Thr	Leu	Ala 2440	Thr	Gly	Pro	Ile	Ser 2445	Thr	Leu	Trp
Glu		Asn	Pro	Gly	Arg	2440 Phe 2455	Trp	Asn	Thr	Thr			Val	Ser
Met	Ala	Asn	Ile	Phe	Arg	Gly	Ser	Tyr	Leu	Ala	Gly		Gly	Leu
Leu	2465 Phe	Ser	Ile	Met	Lys	2470 Asn	Thr	Thr	Asn	Thr	2475 Arg	Arg	Gly	Thr

-continued

	2480					2485					2490			
Gly	Asn 2495	Ile	Gly	Glu	Thr	Leu 2500	Gly	Glu	Lys	Trp	Lys 2505	Ser	Arg	Leu
Asn	Ala 2510	Leu	Gly	Lys	Ser	Glu 2515	Phe	Gln	Ile	Tyr	Lys 2520	Lys	Ser	Gly
Ile	Gln 2525	Glu	Val	Asp	Arg	Thr 2530	Leu	Ala	Lys	Glu	Gly 2535	Ile	Lys	Arg
Gly	Glu 2540	Thr	Asp	His	His	Ala 2545	Val	Ser	Arg	Gly	Ser 2550	Ala	ГЛа	Leu
Arg	Trp 2555	Phe	Val	Glu	Arg	Asn 2560	Met	Val	Thr	Pro	Glu 2565	Gly	ГЛа	Val
Val	Asp 2570	Leu	Gly	Суз	Gly	Arg 2575	Gly	Gly	Trp	Ser	Tyr 2580		Суз	Gly
Gly	Leu 2585	ГЛЗ	Asn	Val	Arg	Glu 2590	Val	Lys	Gly	Leu	Thr 2595	Lys	Gly	Gly
Pro	Gly 2600	His	Glu	Glu	Pro	Ile 2605	Pro	Met	Ser	Thr	Tyr 2610	Gly	Trp	Asn
Leu	Val 2615	Arg	Leu	Gln	Ser	Gly 2620	Val	Asp	Val	Phe	Phe 2625	Ile	Pro	Pro
Glu	Lys 2630	Сув	Asp	Thr	Leu	Leu 2635	Сув	Asp	Ile	Gly	Glu 2640	Ser	Ser	Pro
Asn	Pro 2645	Thr	Val	Glu	Ala	Gly 2650	Arg	Thr	Leu	Arg	Val 2655	Leu	Asn	Leu
Val	Glu 2660	Asn	Trp	Leu	Asn	Asn 2665	Asn	Thr	Gln	Phe	Сув 2670	Ile	Lys	Val
Leu	Asn 2675	Pro	Tyr	Met	Pro	Ser 2680	Val	Ile	Glu	Lys	Met 2685	Glu	Ala	Leu
Gln	Arg 2690	Lys	Tyr	Gly	Gly	Ala 2695	Leu	Val	Arg	Asn	Pro 2700	Leu	Ser	Arg
Asn	Ser 2705	Thr	His	Glu	Met	Tyr 2710	Trp	Val	Ser	Asn	Ala 2715	Ser	Gly	Asn
Ile	Val 2720	Ser	Ser	Val	Asn	Met 2725	Ile	Ser	Arg	Met	Leu 2730	Ile	Asn	Arg
Phe	Thr 2735	Met	Arg	Tyr	Lys	Lys 2740	Ala	Thr	Tyr	Glu	Pro 2745	Asp	Val	Asp
Leu	Gly 2750	Ser	Gly	Thr	Arg	Asn 2755	Ile	Gly	Ile	Glu	Ser 2760	Glu	Ile	Pro
Asn	Leu 2765	Asp	Ile	Ile	Gly	Lys 2770	Arg	Ile	Glu	Lys	Ile 2775	Lys	Gln	Glu
His	Glu 2780	Thr	Ser	Trp	His	Tyr 2785	Asp	Gln	Asp	His	Pro 2790		Lys	Thr
Trp	Ala 2795	Tyr	His	Gly	Ser	Tyr 2800	Glu	Thr	Lys	Gln	Thr 2805	Gly	Ser	Ala
Ser	Ser 2810	Met	Val	Asn	Gly	Val 2815	Val	Arg	Leu	Leu	Thr 2820	ГЛа	Pro	Trp
Asp	Val 2825	Val	Pro	Met	Val	Thr 2830	Gln	Met	Ala	Met	Thr 2835	Asp	Thr	Thr
Pro	Phe 2840	Gly	Gln	Gln	Arg	Val 2845	Phe	Lys	Glu	Lys	Val 2850	Asp	Thr	Arg
Thr	Gln 2855	Glu	Pro	Lys	Glu	Gly 2860	Thr	Lys	Lys	Leu	Met 2865	Lys	Ile	Thr

Ala	Glu 2870		Leu	Trp	Lys	Glu 2875		Gly	Lys	Lys	Lys 2880		Pro	Arg
Met	Сув 2885		Arg	Glu	Glu	Phe 2890		Arg	Lys	Val	Arg 2895		Asn	Ala
Ala	Leu 2900		Ala	Val	Phe	Thr 2905		Glu	Asn	Lys	Trp 2910	-	Ser	Ala
Arg	Glu 2915		Val	Glu	Asp	Ser 2920		Phe	Trp	Glu	Leu 2925		Asp	Lya
Glu	Arg 2930		Leu	His	Leu	Glu 2935		Lys	Суз	Glu	Thr 2940		Val	Tyr
Asn	Met 2945	Met	Gly	Lys	Arg	Glu 2950		Lys	Leu	Gly	Glu 2955		Gly	LÀa
Ala	Lys 2960		Ser	Arg	Ala	Ile 2965		Tyr	Met	Trp	Leu 2970		Ala	Arg
Phe	Leu 2975	Glu	Phe	Glu	Ala	Leu 2980		Phe	Leu	Asn	Glu 2985		His	Trp
Phe	Ser 2990		Glu	Asn	Ser	Leu 2995		Gly	Val	Glu	Gly 3000		Gly	Leu
His	Lуя 3005		Gly	Tyr	Ile	Leu 3010		Asp	Val	Ser	Lуя 3015		Glu	Gly
Gly	Ala 3020		Tyr	Ala	Asp	Asp 3025		Ala	Gly	Trp	Aap 3030		Arg	Ile
Thr	Leu 3035		Asp	Leu	Lys	Asn 3040		Glu	Met	Val	Thr 3045		His	Met
Glu	Gly 3050		His	Гла	Lys	Leu 3055		Glu	Ala	Ile	Phe 3060		Leu	Thr
Tyr	Gln 3065	Asn	Lys	Val	Val	Arg 3070		Gln	Arg	Pro	Thr 3075		Arg	Gly
Thr	Val 3080	Met	Asp	Ile	Ile	Ser 3085		Arg	Asp	Gln	Arg 3090		Ser	Gly
Gln	Val 3095	Gly	Thr	Tyr	Gly	Leu 3100		Thr	Phe	Thr	Asn 3105		Glu	Ala
Gln	Leu 3110		Arg	Gln	Met	Glu 3115		Glu	Gly	Val	Phe 3120		Ser	Ile
Gln	His 3125	Leu	Thr	Ile	Thr	Glu 3130		Ile	Ala	Val	Gln 3135	Asn	Trp	Leu
Ala	Arg 3140	Val	Gly	Arg	Glu	Arg 3145	Leu	Ser	Arg	Met	Ala 3150		Ser	Gly
Asp	Asp 3155		Val	Val	Lys	Pro 3160		Asp	Asp	Arg	Phe 3165	Ala	Ser	Ala
Leu	Thr 3170		Leu	Asn	Asp	Met 3175	Gly	Lys	Ile	Arg	Lys 3180	Asp	Ile	Gln
Gln	Trp 3185	Glu	Pro	Ser	Arg	Gly 3190	Trp	Asn	Asp	Trp	Thr 3195	Gln	Val	Pro
Phe	Сув 3200	Ser	His	His	Phe	His 3205	Glu	Leu	Ile	Met	Lys 3210	Asp	Gly	Arg
Val	Leu 3215	Val	Val	Pro	Суз	Arg 3220		Gln	Asp	Glu	Leu 3225	Ile	Gly	Arg
Ala	Arg 3230	Ile	Ser	Gln	Gly	Ala 3235	Gly	Trp	Ser	Leu	Arg 3240		Thr	Ala

Суз	Leu 3245		Lys	s Ser	тул	r Ala 325		ln M	let	Trp	Ser	Leu 325		Met	Tyr	Phe
His	Arg 3260	-	Asp	) Leu	ı Arg	g Lei 326		la A	Ala	Asn	Ala	Ile 327		Суз	Ser	Ala
Val	Pro 3275		His	s Trp	Va]	L Pro 328		hr S	Ser	Arg	Thr	Thr 328		Irp	Ser	Ile
His	Ala 3290	-	His	s Glu	ı Tr <u>p</u>	Met 329		hr 1	[hr	Glu	Asp	Met 330		Ĺeu	Thr	Val
Trp	Asn 3305	-	Val	. Trp	) Ile	e Glr 331		lu A	Asn	Pro	Trp	Met 331		Glu	Asp	Lys
Thr	Pro 3320		Glu	ı Ser	Tr <u>p</u>	9 Glu 332		lu I	le	Pro	Tyr	Leu 333		Gly	Lya	Arg
Glu	Asp 3335		Trp	суз	; Gl	7 Sei 334		eu I	le	Gly	Leu	Thr 334		Ser	Arg	Ala
Thr	Trp 3350		Lys	a Asr	n Il€	e Glr 335		la A	Ala	Ile	Asn	Gln 336		Val	Arg	Ser
Leu	Ile 3365	-	Asr	ı Glu	ı Glu	1 Tyj 337		hr A	/ab	Tyr	Met	Prc 337		Ser	Met	Lya
Arg	Phe 3380	-	Arg	g Glu	ı Glı	1 Gli 338		lu A	Ala	Gly	Val	Leu 339		Irp		
<40	3> OT 0> SE Asn	QUEN	CE :	6			-			n Tì					ı Met 15	: Leu
Lys	Arg		Arg 20	Asn	Arg	Val	Ser	Thr 25	: Va	1 G	ln G	ln L	Jeu	Thr 30	Lys	Arg
Phe	Ser	Leu 35	Gly	Met	Leu	Gln	Gly 40	Arg	g Gl	y P:	ro L		.уя 15	Leu	. Phe	e Met
Ala	Leu 50	Val	Ala	Phe	Leu	Arg 55	Phe	Leu	ı Th	r I	Le P 6		ro	Thr	Ala	u Gly
Ile 65	Leu	Lys	Arg	Trp	Gly 70	Thr	Ile	Lys	s Ly	rs Se 7!		ys A	Ala	Ile	e Asr	n Val 80
Leu	Arg	Gly	Phe	Arg 85	Lys	Glu	Ile	Glγ	7 Ar 90	-	et L	eu A	lsn	Ile	e Leu 95	ı Asn
Arg	Arg	-	Arg 100	Ser	Ala	Gly	Met	Ile 105		e Me	et L	eu I	le	Prc 110		Val
Met	Ala	Phe 115	His	Leu	Thr	Thr	Arg 120		n Gl	y G	Lu P		lis 25	Met	Ile	e Val
Ser	Arg 130	Gln	Glu	LÀa	Gly	Lys 135	Ser	Leu	ı Le	u Pl		ys 1 40	hr	Glu	l Val	. Gly
Val 145	Asn	Met	Cys	Thr	Leu 150	Met	Ala	Met	: As	-	eu G	ly G	Ju	Leu	. Суа	Glu 160
Asp	Thr	Ile	Thr	Tyr 165	Glu	Суз	Pro	Leu	ı Le 17		rg G	ln A	Asn	Glu	1 Pro	
-	Thr Ile	Asp		165		-			17 : Se	0	-				175 Тул	5

												con	tin	ued	
Thr	Суз	Thr 195	Thr	Met	Gly	Glu	His 200	Arg	Arg	Glu	LÀa	Arg 205	Ser	Val	Ala
Leu	Val 210	Pro	His	Val	Gly	Met 215	Gly	Leu	Glu	Thr	Arg 220	Thr	Glu	Thr	Trp
Met 225	Ser	Ser	Glu	Gly	Ala 230	Trp	Lys	His	Val	Gln 235	Arg	Ile	Glu	Thr	Trp 240
Ile	Leu	Arg	His	Pro 245	Gly	Phe	Thr	Met	Met 250	Ala	Ala	Ile	Leu	Ala 255	Tyr
Thr	Ile	Gly	Thr 260	Thr	His	Phe	Gln	Arg 265	Ala	Leu	Ile	Phe	Ile 270	Leu	Leu
Thr	Ala	Val 275	Thr	Pro	Ser	Met	Thr 280	Met	Arg	Сув	Ile	Gly 285	Met	Ser	Asn
Arg	Asp 290	Phe	Val	Glu	Gly	Val 295	Ser	Gly	Gly	Ser	Trp 300	Val	Asp	Ile	Val
Leu 305		His	Gly	Ser	Cys 310	Val	Thr	Thr	Met	Ala 315		Asn	Lys	Pro	Thr 320
	Asp	Phe	Glu	Leu 325	Ile		Thr	Glu	Ala 330		Gln	Pro	Ala	Thr 335	
Arg	Lys	Tyr	-		Glu	Ala	Lys			Asn	Thr	Thr			Ser
Arg	Сув		340 Thr	Gln	Gly	Glu		345 Ser	Leu	Asn	Glu		350 Gln	Asp	Lys
Arg		355 Val	Cys	Lys	His		360 Met	Val	Asp	Arg	-	365 Trp	Gly	Asn	Gly
	370 Gly	Leu	Phe	Gly	Lys		Gly	Ile	Val		380 CYa	Ala	Met	Phe	
385 Cys	Lys	Lys	Asn	Met	390 Glu		Lys	Val	Val	395 Gln	Pro	Glu	Asn	Leu	400 Glu
Tvr	Thr	Ile	Val	405 Ile	Thr	Pro	His	Ser	410 Glv	Glu	Glu	His	Ala	415 Val	Glv
			420		His			425					430		
	-	435	-	-		-	440			-		445			
	450				Glu	455					460				
465			-		Gly 470		-			475					480
Met	Glu	Asn	Lys	Ala 485	Trp	Leu	Val	His	Arg 490	Gln	Trp	Phe	Leu	Asp 495	Leu
Pro	Leu	Pro	Trp 500	Leu	Pro	Gly	Ala	Asp 505	Thr	Gln	Gly	Ser	Asn 510	Trp	Ile
Gln	Lys	Glu 515	Thr	Leu	Val	Thr	Phe 520	Lys	Asn	Pro	His	Ala 525	Lys	Lys	Gln
Asp	Val 530	Val	Val	Leu	Gly	Ser 535	Gln	Glu	Gly	Ala	Met 540	His	Thr	Ala	Leu
Thr 545	Gly	Ala	Thr	Glu	Ile 550	Gln	Met	Ser	Ser	Gly 555	Asn	Leu	Leu	Phe	Thr 560
Gly	His	Leu	Lys	Суя 565	Arg	Leu	Arg	Met	Asp 570	Lys	Leu	Gln	Leu	Lys 575	Gly
Met	Ser	Tyr			Суз	Thr	Gly	-		Lys	Val	Val	-		Ile
Ala	Glu	Thr	580 Gln	His	Gly	Thr	Ile	585 Val	Ile	Arg	Val	Gln	590 Tyr	Glu	Gly

-continued

		595					600					605			
Asp	Gly 610	Ser	Pro	Cys	Lys	Ile 615	Pro	Phe	Glu	Ile	Met 620	Asp	Leu	Glu	Lys
Arg 625	His	Val	Leu	Gly	Arg 630	Leu	Ile	Thr	Val	Asn 635	Pro	Ile	Val	Thr	Glu 640
Lys	Asp	Ser	Pro	Val 645	Asn	Ile	Glu	Ala	Glu 650	Pro	Pro	Phe	Gly	Asp 655	Ser
Tyr	Ile	Ile	Ile 660	Gly	Val	Glu	Pro	Gly 665	Gln	Leu	ГÀа	Leu	Asn 670	Trp	Phe
ГЛа	Lys	Gly 675	Ser	Ser	Ile	Gly	Gln 680	Met	Phe	Glu	Thr	Thr 685	Met	Arg	Gly
Ala	Lys 690	Arg	Met	Ala	Ile	Leu 695	Gly	Asp	Thr	Ala	Trp 700	Asp	Phe	Gly	Ser
Leu 705	Gly	Gly	Val	Phe	Thr 710	Ser	Ile	Gly	Lys	Ala 715	Leu	His	Gln	Val	Phe 720
Gly	Ala	Ile	Tyr	Gly 725	Ala	Ala	Phe	Ser	Gly 730	Val	Ser	Trp	Thr	Met 735	Lys
Ile	Leu	Ile	Gly 740	Val	Ile	Ile	Thr	Trp 745	Ile	Gly	Met	Asn	Ser 750	Arg	Ser
Thr	Ser	Leu 755	Ser	Val	Thr	Leu	Val 760	Leu	Val	Gly	Ile	Val 765	Thr	Leu	Tyr
Leu	Gly 770	Val	Met	Val	Gln	Ala 775	Asp	Ser	Gly	Сүв	Val 780	Val	Ser	Trp	Lys
Asn 785	Lys	Glu	Leu	Lys	Cys 790	Gly	Ser	Gly	Ile	Phe 795	Ile	Thr	Asp	Asn	Val 800
His	Thr	Trp	Thr	Glu 805	Gln	Tyr	Lys	Phe	Gln 810	Pro	Glu	Ser	Pro	Ser 815	Lys
Leu	Ala	Ser	Ala 820	Ile	Gln	ГЛЗ	Ala	His 825	Glu	Glu	Asp	Ile	Cys 830	Gly	Ile
Arg	Ser	Val 835	Thr	Arg	Leu	Glu	Asn 840	Leu	Met	Trp	Lys	Gln 845	Ile	Thr	Pro
Glu	Leu 850	Asn	His	Ile	Leu	Ser 855	Glu	Asn	Glu	Val	Lys 860	Leu	Thr	Ile	Met
Thr 865	Gly	Asp	Ile	Lys	Gly 870	Ile	Met	Gln	Ala	Gly 875	Lys	Arg	Ser	Leu	Arg 880
Pro	Gln	Pro	Thr	Glu 885	Leu	Гла	Tyr	Ser	Trp 890	Lys	Thr	Trp	Gly	Lys 895	Ala
Lys	Met	Leu	Ser 900	Thr	Glu	Ser	His	Asn 905	Gln	Thr	Phe	Leu	Ile 910	Asp	Gly
Pro	Glu	Thr 915	Ala	Glu	Сүз	Pro	Asn 920	Thr	Asn	Arg	Ala	Trp 925	Asn	Ser	Leu
Glu	Val 930	Glu	Asp	Tyr	Gly	Phe 935	Gly	Val	Phe	Thr	Thr 940	Asn	Ile	Trp	Leu
Lys 945	Leu	LÀa	Glu	ГЛа	Gln 950	Asp	Val	Phe	Суз	Asp 955	Ser	ГЛа	Leu	Met	Ser 960
Ala	Ala	Ile	Lys	Asp 965	Asn	Arg	Ala	Val	His 970	Ala	Asp	Met	Gly	Tyr 975	Trp
Ile	Glu	Ser	Ala 980	Leu	Asn	Asp	Thr	Trp 985	Lys	Ile	Glu	ГЛа	Ala 990	Ser	Phe
Ile	Glu	Val 995	Lys	Asn	Суз	His	Trp 1000		о Буя	s Se:	r Hi	s Th: 100		eu T:	rp Ser

Asn	Gly 1010		Leu	Glu	Ser	Glu 1015		Ile	Ile	Pro	Lys 1020		Leu	Ala
Gly	Pro 1025		Ser	Gln	His	Asn 1030		Arg	Pro	Gly	Tyr 1035		Thr	Gln
Ile	Thr 1040		Pro	Trp	His	Leu 1045		Lys	Leu	Glu	Met 1050		Phe	Aap
Phe	Суз 1055	Asp	Gly	Thr	Thr	Val 1060		Val	Thr	Glu	Asp 1065		Gly	Asn
Arg	Gly 1070		Ser	Leu	Arg	Thr 1075		Thr	Ala	Ser	Gly 1080	-	Leu	Ile
Thr	Glu 1085		Суз	Сүз	Arg	Ser 1090		Thr	Leu	Pro	Pro 1095		Arg	Tyr
Arg	Gly 1100		Asp	Gly	Суз	Trp 1105		Gly	Met	Glu	Ile 1110		Pro	Leu
ГЛа	Glu 1115		Glu	Glu	Asn	Leu 1120		Asn	Ser	Leu	Val 1125		Ala	Gly
His	Gly 1130		Val	Asp	Asn	Phe 1135		Leu	Gly	Val	Leu 1140	_	Met	Ala
Leu	Phe 1145		Glu	Glu	Met	Leu 1150		Thr	Arg	Val	Gly 1155		ГЛа	His
Ala	Ile 1160		Leu	Val	Ala	Val 1165		Phe	Val	Thr	Leu 1170		Thr	Gly
Asn	Met 1175	Ser	Phe	Arg	Asp	Leu 1180		Arg	Val	Met	Val 1185		Val	Gly
Ala	Thr 1190		Thr	Asp	Asp	Ile 1195		Met	Gly	Val	Thr 1200		Leu	Ala
Leu	Leu 1205	Ala	Ala	Phe	Lys	Val 1210		Pro	Thr	Phe	Ala 1215		Gly	Leu
Leu	Leu 1220	Arg	Lys	Leu	Thr	Ser 1225		Glu	Leu	Met	Met 1230		Thr	Ile
Gly	Ile 1235		Leu	Leu	Ser	Gln 1240		Thr	Ile	Pro	Glu 1245		Ile	Leu
Glu	Leu 1250		Asp	Ala	Leu	Ala 1255		Gly	Met	Met	Val 1260		ГЛа	Met
Val	Arg 1265	Asn	Met	Glu	Lys	Tyr 1270		Leu	Ala	Val	Thr 1275		Met	Ala
Ile	Leu 1280	Сүз	Val	Pro	Asn	Ala 1285		Ile	Leu	Gln	Asn 1290	Ala	Trp	LYa
Val	Ser 1295	-	Thr	Ile	Leu	Ala 1300		Val	Ser	Val	Ser 1305	Pro	Leu	Phe
Leu	Thr 1310		Ser	Gln	Gln	Lys 1315		Aab	Trp	Ile	Pro 1320	Leu	Ala	Leu
Thr	Ile 1325	ГЛа	Gly	Leu	Asn	Pro 1330	Thr	Ala	Ile	Phe	Leu 1335	Thr	Thr	Leu
Ser	Arg 1340	Thr	Ser	Гла	Гла	Arg 1345	Ser	Trp	Pro	Leu	Asn 1350	Glu	Ala	Ile
Met	Ala 1355	Val	Gly	Met	Val	Ser 1360		Leu	Ala	Ser	Ser 1365	Leu	Leu	Lya
Asn	Asp 1370		Pro	Met	Thr	Gly 1375		Leu	Val	Ala	Gly 1380	Gly	Leu	Leu

											- COI	1111	nuec	1
Thr	Val 1385		Tyr	Val	Leu	Thr 1390		Arg	Ser	Ala	Asp 1395	Leu	Glu	Leu
Glu	Arg 1400	Ala	Ala	Asp	Val	Lys 1405	-	Glu	Asp	Gln	Ala 1410	Glu	Ile	Ser
Gly	Ser 1415	Ser	Pro	Ile	Leu	Ser 1420	Ile	Thr	Ile	Ser	Glu 1425	Asp	Gly	Ser
Met	Ser 1430		Lys	Asn	Glu	Glu 1435	Glu	Glu	Gln	Thr	Leu 1440	Thr	Ile	Leu
Ile	Arg 1445		Gly	Leu	Leu	Val 1450		Ser	Gly	Leu	Phe 1455	Pro	Val	Ser
Ile	Pro 1460	Ile	Thr	Ala	Ala	Ala 1465	-	Tyr	Leu	Trp	Glu 1470	Val	ГЛЗ	Lys
Gln	Arg 1475	Ala	Gly	Val	Leu	Trp 1480		Val	Pro	Ser	Pro 1485	Pro	Pro	Met
Gly	Lys 1490		Glu	Leu	Glu	Asp 1495			Tyr		Ile 1500		Gln	LYa
Gly	Ile 1505			Tyr		Gln 1510		Gly	Ala	Gly	Val 1515		Lys	Glu
Gly		Phe				Trp 1525		Val	Thr	Arg			Val	Leu
Met		Lys	-	Lys	Arg	Ile 1540		Pro	Ser	Trp		Asp	Val	Lys
Lys		Leu			Tyr	Gly 1555	Gly		Trp			Glu	Gly	Glu
Trp	Lys	Glu		Glu		1555 Val 1570	Gln					Glu	Pro	Gly
Lys		Pro	Arg	Ala	Val	Gln	Thr				Leu	Phe	Lys	Thr
Asn		Gly		Ile		1585 Ala	Val		Leu			Ser	Pro	Gly
Thr		Gly		Pro	Ile	1600 Ile	Asp					Val	Val	Gly
Leu		Gly	Asn			1615 Val	Thr		Ser	Gly			Val	Ser
Ala	1625 Ile			Thr		1630 Lys		Ile			1635 Asn		Glu	Ile
	1640					1645 Lys				-	1650			
	1655					1660 Thr					1665			
	1670	-		-	-	1675	-	-	-		1680			
Arg	Glu 1685	Ala	Ile	Lys	Arg	Gly 1690		Arg	Thr	Leu	Ile 1695		Ala	Pro
Thr	Arg 1700	Val	Val	Ala	Ala	Glu 1705		Glu	Glu	Ala	Leu 1710	-	Gly	Leu
Pro	Ile 1715		Tyr	Gln	Thr	Pro 1720		Ile	Arg	Ala	Val 1725		Thr	Gly
Arg	Glu 1730		Val	Asp	Leu	Met 1735	-	His	Ala	Thr	Phe 1740	Thr	Met	Arg
Leu	Leu 1745	Ser	Pro	Val	Arg	Val 1750	Pro	Asn	Tyr	Asn	Leu 1755	Ile	Ile	Met
Asp		Ala	His	Phe	Thr	Asp	Pro	Ala	Ser	Ile		Ala	Arg	Gly

-continued

	1760					1765					1770			
Tyr	Ile 1775	Ser	Thr	Arg	Val	Glu 1780	Met	Gly	Glu	Ala	Ala 1785	Gly	Ile	Phe
Met	Thr 1790	Ala	Thr	Pro	Pro	Gly 1795	Ser	Arg	Asp	Pro	Phe 1800	Pro	Gln	Ser
Asn	Ala 1805	Pro	Ile	Ile	Asp	Glu 1810	Glu	Arg	Glu	Ile	Pro 1815	Glu	Arg	Ser
Trp	Asn 1820	Ser	Gly	His	Glu	Trp 1825	Val	Thr	Asp	Phe	Lys 1830	Gly	ГЛа	Thr
Val	Trp 1835	Phe	Val	Pro	Ser	Ile 1840	Lys	Ala	Gly	Asn	Asp 1845	Ile	Ala	Ala
Сүз	Leu 1850	Arg	Lys	Asn	Gly	Lys 1855	Lys	Val	Ile	Gln	Leu 1860	Ser	Arg	Гла
Thr	Phe 1865	Asp	Ser	Glu	Tyr	Val 1870	Lys	Thr	Arg	Thr	Asn 1875	Asp	Trp	Asp
Phe	Val 1880	Val	Thr	Thr	Asp	Ile 1885	Ser	Glu	Met	Gly	Ala 1890	Asn	Phe	Lya
Ala	Glu 1895	Arg	Val	Ile	Asp	Pro 1900	Arg	Arg	Сүв	Met	Lys 1905	Pro	Val	Ile
Leu	Thr 1910	Asb	Gly	Glu	Glu	Arg 1915	Val	Ile	Leu	Ala	Gly 1920	Pro	Met	Pro
Val	Thr 1925	His	Ser	Ser	Ala	Ala 1930	Gln	Arg	Arg	Gly	Arg 1935	Ile	Gly	Arg
Asn	Pro 1940	Lys	Asn	Glu	Asn	Asp 1945	Gln	Tyr	Ile	Tyr	Met 1950	Gly	Glu	Pro
Leu	Glu 1955	Asn	Asp	Glu	Asp	Cys 1960	Ala	His	Trp	Lys	Glu 1965	Ala	Lys	Met
Leu	Leu 1970	Asp	Asn	Ile	Asn	Thr 1975	Pro	Glu	Gly	Ile	Ile 1980	Pro	Ser	Met
Phe	Glu 1985	Pro	Glu	Arg	Glu	Lys 1990	Val	Asp	Ala	Ile	Asp 1995	Gly	Glu	Tyr
Arg	Leu 2000	Arg	Gly	Glu	Ala	Arg 2005	Lys	Thr	Phe	Val	Asp 2010	Leu	Met	Arg
Arg	Gly 2015	Asp	Leu	Pro	Val	Trp 2020	Leu	Ala	Tyr	Arg	Val 2025	Ala	Ala	Glu
Gly	Ile 2030	Asn	Tyr	Ala	Asp	Arg 2035	Arg	Trp	Суа	Phe	Asp 2040	Gly	Val	Lys
Asn	Asn 2045	Gln	Ile	Leu	Glu	Glu 2050	Asn	Val	Glu	Val	Glu 2055	Ile	Trp	Thr
ГЛа	Glu 2060	Gly	Glu	Arg	Lys	Lys 2065	Leu	Lys	Pro	Arg	Trp 2070	Leu	Asp	Ala
Arg	Ile 2075	Tyr	Ser	Asp	Pro	Leu 2080	Ala	Leu	Lys	Glu	Phe 2085	Lys	Glu	Phe
Ala	Ala 2090	Gly	Arg	ГЛа	Ser	Leu 2095	Thr	Leu	Asn	Leu	Ile 2100	Thr	Glu	Met
Gly	Arg 2105	Leu	Pro	Thr	Phe	Met 2110	Thr	Gln	Lya	Ala	Arg 2115	Asp	Ala	Leu
Asp	Asn 2120	Leu	Ala	Val	Leu	His 2125	Thr	Ala	Glu	Ala	Gly 2130	Gly	Arg	Ala
Tyr	Asn 2135	His	Ala	Leu	Ser	Glu 2140	Leu	Pro	Glu	Thr	Leu 2145	Glu	Thr	Leu

-	СО	nt	in	ue	d

Leu	Leu 2150		Thr	Leu	Leu	Ala 2155	Thr	Val	Thr	Gly	Gly 2160	Ile	Phe	Leu
Phe	Leu 2165		Ser	Ala	Arg	Gly 2170		Gly	Lys	Met	Thr 2175	Leu	Gly	Met
Суз	Cys 2180		Ile	Thr	Ala	Ser 2185	Ile	Leu	Leu	Trp	Tyr 2190	Ala	Gln	Ile
Gln	Pro 2195	His	Trp	Ile	Ala	Ala 2200	Ser	Ile	Ile	Leu	Glu 2205	Phe	Phe	Leu
Ile	Val 2210		Leu	Ile	Pro	Glu 2215	Pro	Glu	Lys	Gln	Arg 2220	Thr	Pro	Gln
Asp	Asn 2225		Leu	Thr	Tyr	Val 2230		Ile	Ala	Ile	Leu 2235	Thr	Val	Val
Ala	Ala 2240		Met	Ala	Asn	Glu 2245	Met	Gly	Phe	Leu	Glu 2250	Lys	Thr	Lys
ГЛа	Asp 2255		Gly	Leu	Gly	Ser 2260	Ile	Ala	Thr	Gln	Gln 2265	Pro	Glu	Ser
Asn	Ile 2270		Asp	Ile	Asp	Leu 2275	Arg	Pro	Ala	Ser	Ala 2280	Trp	Thr	Leu
Tyr	Ala 2285		Ala	Thr	Thr	Phe 2290	Val	Thr	Pro	Met	Leu 2295	Arg	His	Ser
Ile	Glu 2300		Ser	Ser	Val	Asn 2305	Val	Ser	Leu	Thr	Ala 2310	Ile	Ala	Asn
Gln	Ala 2315		Val	Leu	Met	Gly 2320	Leu	Gly	Lys	Gly	Trp 2325	Pro	Leu	Ser
Lys	Met 2330		Ile	Gly	Val	Pro 2335	Leu	Leu	Ala	Ile	Gly 2340	Суз	Tyr	Ser
Gln	Val 2345		Pro	Ile	Thr	Leu 2350	Thr	Ala	Ala	Leu	Phe 2355	Leu	Leu	Val
Ala	His 2360		Ala	Ile	Ile	Gly 2365	Pro	Gly	Leu	Gln	Ala 2370	Lys	Ala	Thr
Arg	Glu 2375		Gln	Lys	Arg	Ala 2380	Ala	Ala	Gly	Ile	Met 2385	Lys	Asn	Pro
Thr	Val 2390		Gly	Ile	Thr	Val 2395	Ile	Asp	Leu	Asp	Pro 2400	Ile	Pro	Tyr
Asp	Pro 2405	Lys	Phe	Glu	Lys	Gln 2410	Leu	Gly	Gln	Val	Met 2415	Leu	Leu	Val
Leu	Cys 2420	Val	Thr	Gln	Val	Leu 2425	Met	Met	Arg	Thr	Thr 2430	Trp	Ala	Leu
Суз	Glu 2435	Ala	Leu	Thr	Leu	Ala 2440	Thr	Gly	Pro	Ile	Ser 2445	Thr	Leu	Trp
Glu	Gly 2450	Asn	Pro	Gly	Arg	Phe 2455	Trp	Asn	Thr	Thr	Ile 2460	Ala	Val	Ser
Met	Ala 2465	Asn	Ile	Phe	Arg	Gly 2470	Ser	Tyr	Leu	Ala	Gly 2475	Ala	Gly	Leu
Leu	Phe 2480	Ser	Ile	Met	Lys	Asn 2485	Thr	Thr	Asn	Thr	Arg 2490	Arg	Gly	Thr
Gly	Asn 2495	Ile	Gly	Glu	Thr	Leu 2500	Gly	Glu	Гла	Trp	Lys 2505	Ser	Arg	Leu
Asn	Ala 2510	Leu	Gly	Гла	Ser	Glu 2515	Phe	Gln	Ile	Tyr	Lys 2520	Гуз	Ser	Gly

-continued

													iuct	~
Ile	Gln 2525	Glu	Val	Asp	Arg	Thr 2530		Ala	Lys	Glu	Gly 2535	Ile	ГЛа	Arg
Gly	Glu 2540	Thr	Asp	His	His	Ala 2545		Ser	Arg	Gly	Ser 2550	Ala	ГЛа	Leu
Arg	Trp 2555	Phe	Val	Glu	Arg	Asn 2560		Val	Thr	Pro	Glu 2565	Gly	Lys	Val
Val	Asp 2570		Gly	Сүз	Gly	Arg 2575		Gly	Trp	Ser	Tyr 2580		Сүз	Gly
Gly	Leu 2585	-	Asn	Val	Arg	Glu 2590		Lys	Gly	Leu	Thr 2595	Lys	Gly	Gly
Pro	Gly 2600	His	Glu	Glu	Pro	Ile 2605		Met	Ser	Thr	Tyr 2610	Gly	Trp	Asn
Leu	Val 2615	-	Leu	Gln	Ser	Gly 2620		Aap	Val	Phe	Phe 2625	Ile	Pro	Pro
Glu		Суз	Asp	Thr	Leu	Leu 2635	Cys	-	Ile	-		Ser	Ser	Pro
Asn			Val	Glu	Ala	Gly 2650	Arg			Arg		Leu	Asn	Leu
Val		Asn	Trp	Leu	Asn	2650 Asn 2665	Asn	Thr	Gln			Ile	ГЛа	Val
Leu	Asn	Pro	Tyr	Met	Pro	Ser	Val	Ile	Glu	Lys	Met	Glu	Ala	Leu
Gln			Tyr	Gly	Gly	2680 Ala	Leu					Leu	Ser	Arg
Asn			His	Glu	Met	2695 Tyr	Trp		Ser	Asn		Ser	Gly	Asn
Ile			Ser	Val	Asn	2710 Met	Ile			Met		Ile	Asn	Arg
Phe	2720 Thr		Arg	Tyr	Lys	2725 Lys					2730 Pro	Asp	Val	Asp
	2735					2740 Asn					2745			
	2750		-		-	2755 Lys		-			2760			
	2765					2770					2775			
	2780			_		Tyr 2785 -	-		-		2790	-	-	
Trp	Ala 2795	Tyr	His	Gly	Ser	Tyr 2800	Glu	Thr	Lys	Gln	Thr 2805	Gly	Ser	Ala
Ser	Ser 2810	Met	Val	Asn	Gly	Val 2815		Arg	Leu	Leu	Thr 2820	Lys	Pro	Trp
Asb	Val 2825	Val	Pro	Met	Val	Thr 2830		Met	Ala	Met	Thr 2835	Asp	Thr	Thr
Pro	Phe 2840	Gly	Gln	Gln	Arg	Val 2845		Lys	Glu	Lys	Val 2850	Asp	Thr	Arg
Thr	Gln 2855	Glu	Pro	Lys	Glu	Gly 2860		Lys	Lys	Leu	Met 2865	Lys	Ile	Thr
Ala	Glu 2870	Trp	Leu	Trp	Lys	Glu 2875	Leu	Gly	Lys	Lys	Lys 2880	Thr	Pro	Arg
Met	Cys	Thr	Arg	Glu	Glu	Phe	Thr	Arg	Lys	Val	Arg	Ser	Asn	Ala
Ala	2885 Leu	Gly	Ala	Val	Phe	2890 Thr	Asp	Glu	Asn	Lys	2895 Trp	Lys	Ser	Ala

-continued

	2900					2905					2910			
Arg	Glu 2915	Ala	Val	Glu	Asp	Ser 2920	Arg	Phe	Trp	Glu	Leu 2925	Val	Asp	Lys
Glu	Arg 2930	Asn	Leu	His	Leu	Glu 2935	Gly	Lys	Суз	Glu	Thr 2940	Суз	Val	Tyr
Asn	Met 2945	Met	Gly	Lys	Arg	Glu 2950	Lys	Lys	Leu	Gly	Glu 2955	Phe	Gly	Lys
Ala	Lys 2960	Gly	Ser	Arg	Ala	Ile 2965	Trp	Tyr	Met	Trp	Leu 2970	Gly	Ala	Arg
Phe	Leu 2975	Glu	Phe	Glu	Ala	Leu 2980	Gly	Phe	Leu	Asn	Glu 2985	Asp	His	Trp
Phe	Ser 2990	Arg	Glu	Asn	Ser	Leu 2995	Ser	Gly	Val	Glu	Gly 3000	Glu	Gly	Leu
His	Lys 3005	Leu	Gly	Tyr	Ile	Leu 3010	Arg	Asp	Val	Ser	Lys 3015	Lys	Glu	Gly
Gly	Ala 3020	Met	Tyr	Ala	Asp	Asp 3025	Thr	Ala	Gly	Trp	Aap 3030	Thr	Arg	Ile
Thr	Leu 3035	Glu	Asp	Leu	Lys	Asn 3040	Glu	Glu	Met	Val	Thr 3045	Asn	His	Met
Glu	Gly 3050	Glu	His	Lys	Lys	Leu 3055	Ala	Glu	Ala	Ile	Phe 3060	Lys	Leu	Thr
Tyr	Gln 3065	Asn	Гла	Val	Val	Arg 3070	Val	Gln	Arg	Pro	Thr 3075	Pro	Arg	Gly
Thr	Val 3080	Met	Asp	Ile	Ile	Ser 3085	Arg	Arg	Asp	Gln	Arg 3090	Gly	Ser	Gly
Gln	Val 3095	Gly	Thr	Tyr	Gly	Leu 3100	Asn	Thr	Phe	Thr	Asn 3105	Met	Glu	Ala
Gln	Leu 3110	Ile	Arg	Gln	Met	Glu 3115	Gly	Glu	Gly	Val	Phe 3120	Lys	Ser	Ile
Gln	His 3125	Leu	Thr	Ile	Thr	Glu 3130	Glu	Ile	Ala	Val	Gln 3135	Asn	Trp	Leu
Ala	Arg 3140	Val	Gly	Arg	Glu	Arg 3145	Leu	Ser	Arg	Met	Ala 3150	Ile	Ser	Gly
Asp	Asp 3155	Сүз	Val	Val	Lys	Pro 3160	Leu	Asp	Asp	Arg	Phe 3165	Ala	Ser	Ala
Leu	Thr 3170	Ala	Leu	Asn	Asp	Met 3175	Gly	Lys	Ile	Arg	Lys 3180	Asp	Ile	Gln
Gln	Trp 3185	Glu	Pro	Ser	Arg	Gly 3190		Asn	Asp	Trp	Thr 3195	Gln	Val	Pro
Phe	Суз 3200	Ser	His	His	Phe	His 3205	Glu	Leu	Ile	Met	Lys 3210	Asp	Gly	Arg
Val	Leu 3215	Val	Val	Pro	Cys	Arg 3220	Asn	Gln	Asp	Glu	Leu 3225	Ile	Gly	Arg
Ala	Arg 3230		Ser	Gln	Gly	Ala 3235	Gly	Trp	Ser	Leu	Arg 3240	Glu	Thr	Ala
Суз	Leu 3245		Гла	Ser	Tyr	Ala 3250		Met	Trp	Ser	Leu 3255	Met	Tyr	Phe
His	Arg 3260	Arg	Asp	Leu	Arg	Leu 3265	Ala	Ala	Asn	Ala	Ile 3270	Cys	Ser	Ala
Val	Pro 3275	Ser	His	Trp	Val	Pro 3280	Thr	Ser	Arg	Thr	Thr 3285	Trp	Ser	Ile

His Ala Lys His Glu Trp Met Thr Thr Glu Asp Met Leu 3290 3295 3300	Thr Val										
Trp Asn Arg Val Trp Ile Gln Glu Asn Pro Trp Met Glu 3305 3310 3315	Asp Lys										
Thr Pro Val Glu Ser Trp Glu Glu Ile Pro Tyr Leu Gly 3320 3325 3330	Lys Arg										
Glu Asp Gln Trp Cys Gly Ser Leu Ile Gly Leu Thr Ser 3335 3340 3345	Arg Ala										
Thr Trp Ala Lys Asn Ile Gln Ala Ala Ile Asn Gln Val 3350 3355 3360	Arg Ser										
Leu Ile Gly Asn Glu Glu Tyr Thr Asp Tyr Met Pro Ser 3365 3370 3375	Met Lys										
Arg PheArg Arg Glu Glu Glu Glu Ala Gly Val LeuTrp338033853390											
<210> SEQ ID NO 7 <211> LENGTH: 10717 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Dengue virus serotype 3, BVS											
<400> SEQUENCE: 7											
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa											
gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg											
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac											
ctgacaaaga gatteteact tggaatgetg cagggaegag gaecattaaa											
gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat											
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag	5 5 5										
ggaaggatgc tgaacatctt gaataggaga cgcagatctg caggcatgat attccaacag tgatggcgtt ccatttaacc acgcgtgatg gagagccgcg											
gggaagaatg aaagaggaaa atccctactt ttcaagacag cctctggaat											
acactcataq ctatqqatct qqqaqaqatq tqtqatqaca cqqtcactta	5.5										
cacattaccq aaqtqqaqcc tqaaqacatt qactqctqqt qcaaccttac	5										
qtqacttatq qaacatqcaa tcaaqctqqa qaqcataqac qcqataaqaq	0 00										
ttagctcccc atgttggcat gggactggac acacgcactc aaacctggat											
ggagettgga gacaagtega gaaggtagag acatgggeee ttaggeacee											
atactagece tatteettge ceattacata ggeactteet tgacceagaa											
tttatactat taatgctggt taccccatcc atgacaatga gatgtgtagg											
agagattttg tggaaggcct atcgggagct acgtgggttg acgtggtgct											
gggtgtgtgtga ctaccatggc taagaacaag cccacgctgg acatagagct	tcagaagacc 1080										
gaggccaccc aactggcgac cctaaggaag ctatgcattg agggaaaaat											
acaaccgact caagatgtcc cacccaaggg gaagcgattt tacctgagga											
aactacgtgt gtaagcatac atacgtggac agaggctggg gaaacggttg											
ggcaagggaa gcttggtgac atgcgcgaaa tttcaatgtt tagaatcaat											
yyuuuyyyaa yeelyylyae alyeyeyaaa eelealyee tayddleddl	ugugggaaaa 1320										

		-continued	
gtggtgcaac atgagaacct	caaatacacc gtcatcatca	cagtgcacac aggagaccaa	1380
caccaggtgg gaaatgaaac	gcagggagtc acggctgaga	taacacccca ggcatcaacc	1440
gctgaagcca ttttacctga	atatggaacc ctcgggctag	aatgeteace aeggaeaggt	1500
ttggatttca atgaaatgat	ctcattgaca atgaagaaca	aagcatggat ggtacataga	1560
caatggttct ttgacttacc	cctaccatgg acatcaggag	cttcagcaga aacaccaact	1620
tggaacagga aagagcttct	tgtgacattt aaaaatgcac	atgcaaaaaa gcaagaagta	1680
gttgttcttg gatcacaaga	gggagcaatg catacagcac	tgacaggagc tacagagatc	1740
caaacctcag gaggcacaag	tatctttgcg gggcacttaa	aatgtagact caagatggac	1800
aaattggaac tcaaggggat	gagctatgca atgtgcttga	gtagctttgt gttgaagaaa	1860
gaagteteeg aaaegeagea	tgggacaata ctcattaagg	ttgagtacaa aggggaagat	1920
gcaccctgca agattccttt	ctccacggag gatggacaag	gaaaagctct caatggcaga	1980
ctgatcacag ccaatccagt	ggtgaccaag aaggaggagc	ctgtcaacat tgaggctgaa	2040
cctccttttg gagaaagtaa	catagtaatt ggaattggag	acaaagccct gaaaatcaac	2100
tggtacaaga agggaagctc	gattgggaag atgttcgagg	ccactgccag aggtgcaagg	2160
cgcatggcca tcttgggaga	cacagcctgg gactttggat	cagtgggtgg tgttttgaat	2220
tcattaggga aaatggtcca	ccaaatattt gggagtgctt	acacagccct atttggtgga	2280
gtctcctgga tgatgaaaat	tggaataggt gtcctcttaa	cctggatagg gttgaactca	2340
aaaaatactt ctatgtcatt	ttcatgcatc gcggccggca	ttgtgacact gtatttggga	2400
gtcatggtgc aggccgatag	tggttgcgtt gtgagctgga	aaaacaaaga actgaaatgt	2460
ggcagtggga ttttcatcac	agacaacgtg cacacatgga	cagaacaata caagttccaa	2520
ccagaatccc cttcaaaact	agcttcagct atccagaaag	cccatgaaga ggacatttgt	2580
ggaatccgct cagtaacaag	actggagaat ctgatgtgga	aacaaataac accagaattg	2640
aatcacattc tatcagaaaa	tgaggtgaag ttaactatta	tgacaggaga catcaaagga	2700
atcatgcagg caggaaaacg	atctctgcgg cctcagccca	ctgagctgaa gtattcatgg	2760
aaaacatggg gcaaagcaaa	aatgetetet acagagtete	ataaccagac ctttctcatt	2820
gatggccccg aaacagcaga	atgccccaac acaaatagag	cttggaattc gttggaagtt	2880
gaagactatg gctttggagt	attcaccacc aatatatggc	taaaattgaa agaaaaacag	2940
gatgtattct gcgactcaaa	actcatgtca gcggccataa	aagacaacag agccgtccat	3000
geegatatgg gttattggat	agaaagtgca ctcaatgaca	catggaagat agagaaagcc	3060
tctttcattg aagttaaaaa	ctgccactgg ccaaaatcac	acaccctctg gagcaatgga	3120
gtgctagaaa gtgagatgat	aattccaaag aatctcgctg	gaccagtgtc tcaacacaac	3180
tatagaccag gctaccatac	acaaataaca ggaccatggc	atctaggtaa gcttgagatg	3240
gactttgatt tctgtgatgg	aacaacagtg gtagtgactg	aggactgcgg aaatagagga	3300
ccctctttga gaacaaccac	tgcctctgga aaactcataa	cagaatggtg ctgccgatct	3360
tgcacattac caccgctaag	atacagaggt gaggatgggt	gctggtacgg gatggaaatc	3420
agaccattga aggagaaaga	agagaatttg gtcaactcct	tggtcacagc tggacatggg	3480
caggtcgaca acttttcact	aggagtettg ggaatggeat	tgttcctgga ggaaatgctt	3540
		cagtttcttt tgtgacattg	3600
		_ 00 0	

100

-continued	
atcacaggga acatgtcctt tagagacctg ggaagagtga tggttatggt aggcgccact	3660
atgacggatg acataggtat gggcgtgact tatcttgccc tactagcagc cttcaaagtc	3720
agaccaactt ttgcagctgg actactcttg agaaagctga cctccaagga attgatgatg	3780
actactatag gaattgtact cctctcccag agcaccatac cagagaccat tcttgagttg	3840
actgatgcgt tagccttagg catgatggtc ctcaaaatgg tgagaaatat ggaaaagtat	3900
caattggcag tgactatcat ggctatcttg tgcgtcccaa acgcagtgat attacaaaac	3960
gcatggaaag tgagttgcac aatattggca gtggtgtccg tttccccact gttcttaaca	4020
teeteacage aaaaaacaga ttggatacca ttagcattga egateaaagg teteaateea	4080
acagctattt ttctaacaac cctctcaaga accagcaaga aaaggagctg gccattaaat	4140
gaggctatca tggcagtcgg gatggtgagc attttagcca gttctctcct aaaaaatgat	4200
atteccatga caggaccatt agtggetgga gggeteetea etgtgtgeta egtgeteaet	4260
ggacgatcgg ccgatttgga actggagaga gcagccgatg tcaaatggga agaccaggca	4320
gagatatcag gaagcagtcc aatcctgtca ataacaatat cagaagatgg tagcatgtcg	4380
ataaaaaatg aagaggaaga acaaacactg accatactca ttagaacagg attgctggtg	4440
atctcaggac tttttcctgt atcaatacca atcacggcag cagcatggta cctgtgggaa	4500
gtgaagaaac aacgggccgg agtattgtgg gatgttcctt cacccccacc catgggaaag	4560
gctgaactgg aagatggagc ctatagaatt aagcaaaaag ggattcttgg atattcccag	4620
atcggagccg gagtttacaa agaaggaaca ttccatacaa tgtggcatgt cacacgtggc	4680
gctgttctaa tgcataaagg aaagaggatt gaaccatcat gggcggacgt caagaaagac	4740
ctaatatcat atggaggagg ctggaagtta gaaggagaat ggaaggaagg agaagaagtc	4800
caggtattgg cactggagcc tggaaaaaat ccaagagccg tccaaacgaa acctggtctt	4860
ttcaaaacca acgccggaac aataggtgct gtatctctgg acttttctcc tggaacgtca	4920
ggateteeaa ttategacaa aaaaggaaaa gttgtgggte tttatggtaa tggtgttgtt	4980
acaaggagtg gagcatatgt gagtgctata gcccagactg aaaaaagcat tgaagacaac	5040
ccagagatcg aagatgacat tttccgaaag agaagactga ccatcatgga cctccaccca	5100
ggagcgggaa agacgaagag ataccttccg gccatagtca gagaagctat aaaacggggt	5160
ttgagaacat taatcttggc ccccactaga gttgtggcag ctgaaatgga ggaagccctt	5220
agaggactte caataagata ccagaceeca gecateagag etgtgeacae egggegggag	5280
attgtggacc taatgtgtca tgccacattt accatgaggc tgctatcacc agttagagtg	5340
ccaaactaca acctgattat catggacgaa gcccatttca cagacccagc aagtatagca	5400
gctagaggat acatctcaac tcgagtggag atgggtgagg cagctgggat ttttatgaca	5460
gccactcccc cgggaagcag agacccattt cctcagagca atgcaccaat catagatgaa	5520
gaaagagaaa teeetgaacg etegtggaat teeggacatg aatgggteae ggattttaaa	5580
gggaagactg tttggttcgt tccaagtata aaagcaggaa atgatatagc agcttgcctg	5640
aggaaaaatg gaaagaaagt gatacaactc agtaggaaga cctttgattc tgagtatgtc	5700
aagactagaa ccaatgattg ggacttcgtg gttacaactg acatttcaga aatgggtgcc	5760
aatttcaagg ctgagagggt tatagacccc agacgctgca tgaaaccagt catactaaca	5820
gatggtgaag agogggtgat totggoagga ootatgooag tgaccoacto tagtgoagca	5880
Jan	

-continued	
caaagaagag ggagaatagg aagaaatcca aaaaatgaga atgaccagta catatacatg	5940
ggggaacctc tggaaaatga tgaagactgt gcacactgga aagaagctaa aatgctccta	6000
gataacatca acacgccaga aggaatcatt cctagcatgt tcgaaccaga gcgtgaaaag	6060
gtggatgcca ttgatggcga ataccgcttg agaggagaag caaggaaaac ctttgtagac	6120
ttaatgagaa gaggagacct accagtctgg ttggcctaca gagtggcagc tgaaggcatc	6180
aactacgcag acagaaggtg gtgttttgat ggagtcaaga acaaccaaat cctagaagaa	6240
aacgtggaag ttgaaatctg gacaaaagaa ggggaaagga agaaattgaa acccagatgg	6300
ttggatgcta ggatctattc tgacccactg gcgctaaaag aatttaagga atttgcagcc	6360
ggaagaaagt ctctgaccct gaacctaatc acagaaatgg gtaggctccc aaccttcatg	6420
actcagaagg caagaaacgc actggacaac ttagcagtgc tgcacacggc tgaggcaggt	6480
ggaagggggt acaaccatgc tctcagtgaa ctgccggaga ccctggagac attgctttta	6540
ctgacacttc tggctacagt cacgggaggg atcttttat tcttgatgag cgcaaggggc	6600
atagggaaga tgaccctggg aatgtgctgc ataatcacgg ctagcatcct cctatggtac	6660
gcacaaatac agccacactg gatagcagct tcaataatac tggagttttt tctcatagtt	6720
ttgcttattc cagaacctga aaaacagaga acaccccaag acaaccaact gacctacgtt	6780
gtcatagcca tcctcacagt ggtggccgca accatggcaa acgagatggg tttcctagaa	6840
aaaacgaaga aagatctcgg attgggaagc attgcaaccc agcaacccga gagcaacatc	6900
ctggacatag atctacgtcc tgcatcagca tggacgctgt atgccgtggc cacaacattt	6960
gttacaccaa tgttgagaca tagcattgaa aatteeteag tgaatgtgte eetaacaget	7020
atagccaacc aagccacagt gttaatgggt ctcgggaaag gatggccatt gtcaaagatg	7080
gacatoggag tteecettet egecattgga tgetaeteac aagteaacee cataaetete	7140
acagcagete ttttettatt ggtageaeat tatgeeatea tagggeeagg acteeaagea	7200
aaagcaacca gagaagctca gaaaagagca gcggcgggca tcatgaaaaa cccaactgtc	7260
gatggaataa cagtgattga cctagatcca ataccttatg atccaaagtt tgaaaagcag	7320
ttgggacaag taatgeteet agteetetge gtgaeteaag tattgatgat gaggaetaea	7380
tgggetetgt gtgaggettt aacettaget acegggeeea tetecaeatt gtgggaagga	7440
aatccaggga ggttttggaa cactaccatt gcggtgtcaa tggctaacat ttttagaggg	7500
agttacttgg ccggagctgg acttctcttt tctattatga agaacacaac caacacaaga	7560
aggggaactg gcaacatagg agagacgctt ggagagaaat ggaaaagccg attgaacgcg	7620
ttgggaaaaa gtgaattcca gatctacaag aaaagtggaa tccaggaagt ggatagaacc	7680
ttagcaaaag aaggcattaa aagaggagaa acggaccatc acgctgtgtc gcgaggctca	7740
gcaaaactga gatggttcgt tgagagaaac atggtcacac cagaagggaa agtagtggac	7800
ctcggttgtg gcagaggagg ctggtcatac tattgtggag gactaaagaa tgtaagagaa	7860
gtcaaaggcc taacaaaagg aggaccagga cacgaagaac ccatccccat gtcaacatat	7920
gggtggaatc tagtgcgtct tcaaagtgga gttgacgttt tcttcatccc gccagaaaag	7980
tgtgacacat tattgtgtga cataggggag tcatcaccaa atcccacagt ggaagcagga	8040
cgaacactca gagtccttaa cttagtagaa aattggttga acaacaacac tcaattttgc	8100
ataaaggtte teaacceata tatgeeetea gteatagaaa aaatggaage actacaaagg	8160

				- conc 11	Iueu		
aaatatggag	gagccttagt	gaggaatcca	ctctcacgaa	actccacaca	tgagatgtac	8220	
tgggtatcca	atgcttccgg	gaacatagtg	tcatcagtga	acatgatttc	aaggatgttg	8280	
atcaacagat	ttacaatgag	atacaagaaa	gccacttacg	agccggatgt	tgacctcgga	8340	
agcggaaccc	gtaacatcgg	gattgaaagt	gagataccaa	acctagatat	aattgggaaa	8400	
agaatagaaa	aaataaagca	agagcatgaa	acatcatggc	actatgacca	agaccaccca	8460	
tacaaaacgt	gggcatacca	tggtagctat	gaaacaaaac	agactggatc	agcatcatcc	8520	
atggtcaacg	gagtggtcag	gctgctgaca	aaaccttggg	acgtcgtccc	catggtgaca	8580	
cagatggcaa	tgacagacac	gactccattt	ggacaacagc	gcgtttttaa	agagaaagtg	8640	
gacacgagaa	cccaagaacc	gaaagaaggc	acgaagaaac	taatgaaaat	aacagcagag	8700	
tggctttgga	aagaattagg	gaagaaaaag	acacccagga	tgtgcaccag	agaagaattc	8760	
acaagaaagg	tgagaagcaa	tgcagccttg	ggggccatat	tcactgatga	gaacaagtgg	8820	
aagtcggcac	gtgaggctgt	tgaagatagt	aggttttggg	agctggttga	caaggaaagg	8880	
aatctccatc	ttgaaggaaa	gtgtgaaaca	tgtgtgtaca	acatgatggg	aaaaagagag	8940	
aagaagctag	gggaattcgg	caaggcaaaa	ggcagcagag	ccatatggta	catgtggctt	9000	
ggagcacgct	tcttagagtt	tgaagcccta	ggattcttaa	atgaagatca	ctggttctcc	9060	
agagagaact	ccctgagtgg	agtggaagga	gaagggctgc	acaagctagg	ttacattcta	9120	
agagacgtga	gcaagaaaga	gggaggagca	atgtatgccg	atgacaccgc	aggatgggat	9180	
acaagaatca	cactagaaga	cctaaaaaat	gaagaaatgg	taacaaacca	catggaagga	9240	
gaacacaaga	aactagccga	ggccattttc	aaactaacgt	accaaaacaa	ggtggtgcgt	9300	
gtgcaaagac	caacaccaag	aggcacagta	atggacatca	tatcgagaag	agaccaaaga	9360	
ggtagtggac	aagttggcac	ctatggactc	aatactttca	ccaatatgga	agcccaacta	9420	
atcagacaga	tggagggaga	aggagtcttt	aaaagcattc	agcacctaac	aatcacagaa	9480	
gaaatcgctg	tgcaaaactg	gttagcaaga	gtggggcgcg	aaaggttatc	aagaatggcc	9540	
atcagtggag	atgattgtgt	tgtgaaacct	ttagatgaca	ggttcgcaag	cgctttaaca	9600	
gctctaaatg	acatgggaaa	gattaggaaa	gacatacaac	aatgggaacc	ttcaagagga	9660	
tggaatgatt	ggacacaagt	gcccttctgt	tcacaccatt	tccatgagtt	aatcatgaaa	9720	
gacggtcgcg	tactcgttgt	tccatgtaga	aaccaagatg	aactgattgg	cagagcccga	9780	
atctcccaag	gagcagggtg	gtctttgcgg	gagacggcct	gtttggggaa	gtcttacgcc	9840	
caaatgtgga	gcttgatgta	cttccacaga	cgcgacctca	ggctggcggc	aaatgctatt	9900	
tgctcggcag	taccatcaca	ttgggttcca	acaagtcgaa	caacctggtc	catacatgct	9960	
aaacatgaat	ggatgacaac	ggaagacatg	ctgacagtct	ggaacagggt	gtggattcaa	10020	
gaaaacccat	ggatggaaga	caaaactcca	gtggaatcat	gggaggaaat	cccatacttg	10080	
gggaaaagag	aagaccaatg	gtgcggctca	ttgattgggt	taacaagcag	ggccacctgg	10140	
gcaaagaaca	tccaagcagc	aataaatcaa	gttagatccc	ttataggcaa	tgaagaatac	10200	
acagattaca	tgccatccat	gaaaagattc	agaagagaag	aggaagaagc	aggagttctg	10260	
tggtagaaag	caaaactaac	atgaaacaag	gctagaagtc	aggtcggatt	aagccatagt	10320	
acggaaaaaa	ctatgctacc	tgtgagcccc	gtccaaggac	gttaaaagaa	gtcaggccat	10380	
cataaatgcc	atagettgag	taaactatgc	agcctgtagc	tccacctgag	aaggtgtaaa	10440	

aaat	ccgé	gga g	ggcc	acaa	ac c	atgga	aagci	t gta	acge	atgg	cgta	agtg	gac '	tage	ggttag	10500
agga	igaco	ccc f	tccc	ttac	aa a	tege	agcaa	a caa	atgg	gggc	ccaa	agge	gag .	atga	agctgt	10560
agto	tcg	ctg 🤉	gaag	gact	ag a	ggtta	agago	g aga	accc	cccc	gaaa	acaa	aaa	acago	catatt	10620
gaco	gctg	gga a	aaga	ccag	ag a	teet	gctgi	c cto	cctca	agca	tca	ttcc	agg	cacaç	gaacgc	10680
caga	aaat	gg a	aatg	gtgc	tg t	tgaa	tcaa	c ago	gttc	5						10717
<211 <212 <213 <220	.> LI :> T :> OH :> OH	EATU	H: 3 PRT ISM: RE:	389 Art		ial : : Dem	-		18 50	eroty	ype :	3, В'	vs			
<400	)> SI	EQUEI	NCE :	8												
Met 1	Asn	Asn	Gln	Arg 5	Lys	Lys	Ala	Lys	Asn 10	Thr	Pro	Phe	Asn	Met 15	Leu	
Lys	Arg	Glu	Arg 20	Asn	Arg	Val	Ser	Thr 25	Val	Gln	Gln	Leu	Thr 30	Lys	Arg	
Phe	Ser	Leu 35	Gly	Met	Leu	Gln	Gly 40	Arg	Gly	Pro	Leu	Lys 45	Leu	Phe	Met	
Ala	Leu 50	Val	Ala	Phe	Leu	Arg 55	Phe	Leu	Thr	Ile	Pro 60	Pro	Thr	Ala	Gly	
Ile 65	Leu	LÀa	Arg	Trp	Gly 70	Thr	Ile	Lys	ГÀа	Ser 75	ГЛа	Ala	Ile	Asn	Val 80	
Leu	Arg	Gly	Phe	Arg 85	LÀa	Glu	Ile	Gly	Arg 90	Met	Leu	Asn	Ile	Leu 95	Asn	
Arg	Arg	Arg	Arg 100		Ala	Gly	Met	Ile 105	Ile	Met	Leu	Ile	Pro 110	Thr	Val	
Met	Ala	Phe 115	His	Leu	Thr	Thr	Arg 120	Asp	Gly	Glu	Pro	Arg 125	Met	Ile	Val	
Gly	Lys 130	Asn	Glu	Arg	Gly	Lys 135	Ser	Leu	Leu	Phe	Lys 140	Thr	Ala	Ser	Gly	
Ile 145	Asn	Met	Суз	Thr	Leu 150	Ile	Ala	Met	Asp	Leu 155	Gly	Glu	Met	Суз	Asp 160	
Asp	Thr	Val	Thr	Tyr 165	Lys	Сүз	Pro	His	Ile 170	Thr	Glu	Val	Glu	Pro 175	Glu	
Asp	Ile	Asp	Cys 180	-	САа	Asn	Leu	Thr 185	Ser	Thr	Trp	Val	Thr 190	Tyr	Gly	
Thr	Суа	Asn 195	Gln	Ala	Gly	Glu	His 200	Arg	Arg	Asp	ГЛа	Arg 205	Ser	Val	Ala	
Leu	Ala 210	Pro	His	Val	Gly	Met 215	Gly	Leu	Asp	Thr	Arg 220	Thr	Gln	Thr	Trp	
Met 225	Ser	Ala	Glu	Gly	Ala 230	Trp	Arg	Gln	Val	Glu 235	ГÀа	Val	Glu	Thr	Trp 240	
Ala	Leu	Arg	His	Pro 245	-	Phe	Thr	Ile	Leu 250	Ala	Leu	Phe	Leu	Ala 255	His	
Tyr	Ile	Gly	Thr 260		Leu	Thr	Gln	Lys 265	Val	Val	Ile	Phe	Ile 270	Leu	Leu	
Met	Leu	Val 275	Thr	Pro	Ser	Met	Thr 280	Met	Arg	Суз	Val	Gly 285	Val	Gly	Asn	
Arg	Asp 290	Phe	Val	Glu	Gly	Leu 295	Ser	Gly	Ala	Thr	Trp 300	Val	Asp	Val	Val	

Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Ile Glu Leu Gln Lys Thr Glu Ala Thr Gln Leu Ala Thr Leu Arg Lys Leu Cys Ile Glu Gly Lys Ile Thr Asn Ile Thr Thr Asp Ser Arg Cys Pro Thr Gln Gly Glu Ala Ile Leu Pro Glu Glu Gln Asp Gln Asn Tyr Val Cys Lys His Thr Tyr Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser Leu Val Thr Cys Ala Lys Phe Gln Cys Leu Glu Ser Ile Glu Gly Lys Val Val Gln His Glu Asn Leu Lys 405 410 Tyr Thr Val Ile Ile Thr Val His Thr Gly Asp Gln His Gln Val Gly Asn Glu Thr Gln Gly Val Thr Ala Glu Ile Thr Pro Gln Ala Ser Thr Ala Glu Ala Ile Leu Pro Glu Tyr Gly Thr Leu Gly Leu Glu Cys Ser Pro Arg Thr Gly Leu Asp Phe Asn Glu Met Ile Ser Leu Thr Met Lys Asn Lys Ala Trp Met Val His Arg Gln Trp Phe Phe Asp Leu Pro Leu Pro Trp Thr Ser Gly Ala Ser Ala Glu Thr Pro Thr Trp Asn Arg Lys Glu Leu Leu Val Thr Phe Lys Asn Ala His Ala Lys Lys Gln Glu Val Val Val Leu Gly Ser Gln Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Thr Ser Gly Gly Thr Ser Ile Phe Ala Gly His Leu Lys Cys Arg Leu Lys Met Asp Lys Leu Glu Leu Lys Gly Met Ser Tyr Ala Met Cys Leu Ser Ser Phe Val Leu Lys Lys Glu Val Ser Glu Thr Gln His Gly Thr Ile Leu Ile Lys Val Glu Tyr Lys Gly Glu Asp Ala Pro Cys Lys Ile Pro Phe Ser Thr Glu Asp Gly Gln Gly Lys Ala Leu Asn Gly Arg Leu Ile Thr Ala Asn Pro Val Val Thr Lys Lys Glu Glu Pro Val Asn Ile Glu Ala Glu Pro Pro Phe Gly Glu Ser Asn Ile Val Ile Gly Ile Gly Asp Lys Ala Leu Lys Ile Asn Trp Tyr Lys Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ala Thr Ala Arg Gly Ala Arg Arg Met Ala Ile Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly 

-continued

	1100					1105					1110			
Lys	Glu 1115	Glu	Asn	Leu	Val	Asn 1120		Leu	Val	Thr	Ala 1125	Gly	His	Gly
Gln	Val 1130	Asp	Asn	Phe	Ser	Leu 1135	Gly	Val	Leu	Gly	Met 1140	Ala	Leu	Phe
Leu	Glu 1145	Glu	Met	Leu	Arg	Thr 1150		Val	Gly	Thr	Lys 1155	His	Ala	Ile
Leu	Leu 1160	Val	Ala	Val	Ser	Phe 1165	Val	Thr	Leu	Ile	Thr 1170	Gly	Asn	Met
Ser	Phe 1175	Arg	Asp	Leu	Gly	Arg 1180	Val	Met	Val	Met	Val 1185	Gly	Ala	Thr
Met	Thr 1190	Aab	Asp	Ile	Gly	Met 1195	Gly	Val	Thr	Tyr	Leu 1200	Ala	Leu	Leu
Ala	Ala 1205	Phe	ГЛа	Val	Arg	Pro 1210		Phe	Ala	Ala	Gly 1215	Leu	Leu	Leu
Arg	Lys 1220	Leu	Thr	Ser	Гла	Glu 1225	Leu	Met	Met	Thr	Thr 1230	Ile	Gly	Ile
Val	Leu 1235	Leu	Ser	Gln	Ser	Thr 1240	Ile	Pro	Glu	Thr	Ile 1245	Leu	Glu	Leu
Thr	Asp 1250	Ala	Leu	Ala	Leu	Gly 1255	Met	Met	Val	Leu	Lys 1260	Met	Val	Arg
Asn	Met 1265	Glu	Lys	Tyr	Gln	Leu 1270	Ala	Val	Thr	Ile	Met 1275	Ala	Ile	Leu
Суз	Val 1280	Pro	Asn	Ala	Val	Ile 1285	Leu	Gln	Asn	Ala	Trp 1290	Lys	Val	Ser
Сүз	Thr 1295	Ile	Leu	Ala	Val	Val 1300	Ser	Val	Ser	Pro	Leu 1305	Phe	Leu	Thr
Ser	Ser 1310	Gln	Gln	Lys	Thr	Asp 1315	Trp	Ile	Pro	Leu	Ala 1320	Leu	Thr	Ile
ГЛЗ	Gly 1325	Leu	Asn	Pro	Thr	Ala 1330	Ile	Phe	Leu	Thr	Thr 1335	Leu	Ser	Arg
Thr	Ser 1340	Lys	Lys	Arg	Ser	Trp 1345	Pro	Leu	Asn	Glu	Ala 1350	Ile	Met	Ala
Val	Gly 1355	Met	Val	Ser	Ile	Leu 1360	Ala	Ser	Ser	Leu	Leu 1365	Lys	Asn	Asp
Ile	Pro 1370	Met	Thr	Gly	Pro	Leu 1375	Val	Ala	Gly	Gly	Leu 1380	Leu	Thr	Val
Сүз	Tyr 1385	Val	Leu	Thr	Gly	Arg 1390	Ser	Ala	Asp	Leu	Glu 1395	Leu	Glu	Arg
Ala	Ala 1400	Asp	Val	Lys	Trp	Glu 1405	Asp	Gln	Ala	Glu	Ile 1410	Ser	Gly	Ser
Ser	Pro 1415	Ile	Leu	Ser	Ile	Thr 1420	Ile	Ser	Glu	Asb	Gly 1425	Ser	Met	Ser
Ile	Lys 1430	Asn	Glu	Glu	Glu	Glu 1435	Gln	Thr	Leu	Thr	Ile 1440	Leu	Ile	Arg
Thr	Gly 1445	Leu	Leu	Val	Ile	Ser 1450	-	Leu	Phe	Pro	Val 1455	Ser	Ile	Pro
Ile	Thr 1460	Ala	Ala	Ala	Trp	Tyr 1465		Trp	Glu	Val	Lys 1470	Lys	Gln	Arg
Ala	Gly 1475	Val	Leu	Trp	Asp	Val 1480	Pro	Ser	Pro	Pro	Pro 1485	Met	Gly	Lys

-	CO	nt	iı	าน	eċ	l

-														
Ala	Glu 1490	Leu	Glu	Asp	Gly	Ala 1495		Arg	Ile	Гла	Gln 1500		Gly	Ile
Leu	Gly 1505		Ser	Gln	Ile	Gly 1510		Gly	Val	Tyr	Lys 1515	Glu	Gly	Thr
Phe	His 1520	Thr	Met	Trp	His	Val 1525	Thr	Arg	Gly	Ala	Val 1530	Leu	Met	His
Lys	Gly 1535	Lys	Arg	Ile	Glu	Pro 1540	Ser	Trp	Ala	Asp	Val 1545	Lys	Lys	Aap
Leu	Ile 1550	Ser	Tyr	Gly	Gly	Gly 1555	-	Lys	Leu	Glu	Gly 1560	Glu	Trp	Lys
Glu	Gly 1565	Glu	Glu	Val	Gln	Val 1570		Ala	Leu	Glu	Pro 1575	Gly	ГЛа	Asn
Pro	Arg 1580	Ala	Val	Gln	Thr	Lys 1585	Pro	Gly	Leu	Phe	Lys 1590	Thr	Asn	Ala
Gly	Thr 1595	Ile	Gly	Ala	Val	Ser 1600	Leu	Asp	Phe	Ser	Pro 1605	Gly	Thr	Ser
Gly	Ser 1610	Pro	Ile	Ile	Asp	Lys 1615	Lys	Gly	Lys	Val	Val 1620	Gly	Leu	Tyr
Gly	Asn 1625	Gly	Val	Val	Thr	Arg 1630		Gly	Ala	Tyr	Val 1635	Ser	Ala	Ile
Ala	Gln 1640	Thr	Glu	Lys	Ser	Ile 1645	Glu	Asp	Asn	Pro	Glu 1650	Ile	Glu	Aap
Asp	Ile 1655	Phe	Arg	Lys	Arg	Arg 1660		Thr	Ile	Met	Asp 1665	Leu	His	Pro
Gly	Ala 1670	Gly	Lys	Thr	Lys	Arg 1675		Leu	Pro	Ala	Ile 1680	Val	Arg	Glu
Ala	Ile 1685		Arg	Gly	Leu	Arg 1690		Leu	Ile	Leu	Ala 1695	Pro	Thr	Arg
Val	Val 1700	Ala	Ala	Glu	Met	Glu 1705	Glu	Ala	Leu	Arg	Gly 1710	Leu	Pro	Ile
Arg	Tyr 1715	Gln	Thr	Pro	Ala	Ile 1720	Arg	Ala	Val	His	Thr 1725	Gly	Arg	Glu
Ile	Val 1730	Asp	Leu	Met	Суз	His 1735	Ala	Thr	Phe	Thr	Met 1740	Arg	Leu	Leu
Ser	Pro 1745	Val	Arg	Val	Pro	Asn 1750		Asn	Leu	Ile	Ile 1755	Met	Asp	Glu
Ala	His 1760	Phe	Thr	Asp	Pro	Ala 1765	Ser	Ile	Ala	Ala	Arg 1770	Gly	Tyr	Ile
Ser	Thr 1775	Arg	Val	Glu	Met	Gly 1780		Ala	Ala	Gly	Ile 1785	Phe	Met	Thr
Ala	Thr 1790	Pro	Pro	Gly	Ser	Arg 1795	Asp	Pro	Phe	Pro	Gln 1800	Ser	Asn	Ala
Pro	Ile 1805	Ile	Asp	Glu	Glu	Arg 1810		Ile	Pro	Glu	Arg 1815	Ser	Trp	Asn
Ser	Gly 1820	His	Glu	Trp	Val	Thr 1825	Asp	Phe	Гла	Gly	Lys 1830	Thr	Val	Trp
Phe	Val 1835	Pro	Ser	Ile	Lys	Ala 1840	Gly	Asn	Asp	Ile	Ala 1845	Ala	Сүз	Leu
Arg	Lys 1850		Gly	Гла	Lys	Val 1855		Gln	Leu	Ser	Arg 1860		Thr	Phe

-	cont	cin	ued

											-001	.1011	iuec	л 
Asp	Ser 1865	Glu	Tyr	Val	Lys	Thr 1870		Thr	Asn	Asp	Trp 1875	Asp	Phe	Val
Val	Thr 1880		Asp	Ile	Ser	Glu 1885		Gly	Ala	Asn	Phe 1890	Lys	Ala	Glu
Arg	Val 1895		Asp	Pro	Arg	Arg 1900		Met	Lys	Pro	Val 1905	Ile	Leu	Thr
Asp	Gly 1910		Glu	Arg	Val	Ile 1915		Ala	Gly	Pro	Met 1920	Pro	Val	Thr
His	Ser 1925		Ala	Ala	Gln	Arg 1930					Gly 1935	Arg	Asn	Pro
Lys	Asn 1940		Asn	Asp	Gln	Tyr 1945	Ile	Tyr	Met		Glu 1950	Pro	Leu	Glu
Asn	Asp 1955		Asp	Сүз	Ala	His 1960	Trp	Lys	Glu	Ala	Lys 1965	Met	Leu	Leu
Asp	Asn 1970		Asn	Thr	Pro	Glu 1975	Gly	Ile	Ile		Ser 1980	Met	Phe	Glu
Pro		Arg		-	Val	Asp 1990	Ala			-		-	Arg	Leu
Arg		Glu	Ala	Arg		Thr 2005	Phe			Leu			Arg	Gly
Asp		Pro	Val	Trp	Leu	Ala 2020	Tyr	Arg	Val	Ala		Glu	Gly	Ile
Asn		Ala	Asp	Arg	Arg	Trp	Cys	Phe	Asp	Gly			Asn	Asn
Gln	Ile	Leu				Val	Glu				Trp		Гла	Glu
Gly		Arg				2050 Lys	Pro					Ala	Arg	Ile
Tyr		Asp	Pro	Leu	Ala	2065 Leu	Lys	Glu	Phe	Lys		Phe	Ala	Ala
Gly		Lys	Ser	Leu		Leu	Asn	Leu	Ile			Met	Gly	Arg
Leu	2090 Pro				Thr	2095 Gln				Asn	2100 Ala	Leu	Asp	Asn
Leu	2105 Ala		Leu	His	Thr	2110 Ala					2115 Arg	Ala	- Tyr	Asn
	2120					2125 Pro			-	-	2130		-	
	2135					2140 Val					2145			
	2150					2155		-	-		2160			
	2165		-	-		Gly 2170	-				2175		-	-
Ile	Ile 2180	Thr	Ala	Ser	Ile	Leu 2185		Trp	Tyr	Ala	Gln 2190	Ile	Gln	Pro
His	Trp 2195	Ile	Ala	Ala	Ser	Ile 2200		Leu	Glu	Phe	Phe 2205	Leu	Ile	Val
Leu	Leu 2210	Ile	Pro	Glu	Pro	Glu 2215	-	Gln	Arg	Thr	Pro 2220	Gln	Asp	Asn
Gln	Leu 2225		Tyr	Val	Val	Ile 2230		Ile	Leu	Thr	Val 2235	Val	Ala	Ala
Thr	Met	Ala	Asn	Glu	Met	Gly	Phe	Leu	Glu	Lya	Thr	Lys	ГЛа	Aap

-continued

	2240					2245					2250			
Leu	Gly 2255	Leu	Gly	Ser	Ile	Ala 2260		Gln	Gln	Pro	Glu 2265	Ser	Asn	Ile
Leu	Asp 2270	Ile	Asp	Leu	Arg	Pro 2275	Ala	Ser	Ala	Trp	Thr 2280	Leu	Tyr	Ala
Val	Ala 2285	Thr	Thr	Phe	Val	Thr 2290	Pro	Met	Leu	Arg	His 2295	Ser	Ile	Glu
Asn	Ser 2300	Ser	Val	Asn	Val	Ser 2305	Leu	Thr	Ala	Ile	Ala 2310	Asn	Gln	Ala
Thr	Val 2315	Leu	Met	Gly	Leu	Gly 2320	Lys	Gly	Trp	Pro	Leu 2325	Ser	ГЛа	Met
Asp	Ile 2330	Gly	Val	Pro	Leu	Leu 2335	Ala	Ile	Gly	СЛа	Tyr 2340	Ser	Gln	Val
Asn	Pro 2345	Ile	Thr	Leu	Thr	Ala 2350	Ala	Leu	Phe	Leu	Leu 2355	Val	Ala	His
Tyr	Ala 2360	Ile	Ile	Gly	Pro	Gly 2365	Leu	Gln	Ala	ГЛа	Ala 2370	Thr	Arg	Glu
Ala	Gln 2375	ГЛа	Arg	Ala	Ala	Ala 2380	Gly	Ile	Met	ГЛа	Asn 2385	Pro	Thr	Val
Asp	Gly 2390	Ile	Thr	Val	Ile	Asp 2395	Leu	Asp	Pro	Ile	Pro 2400	Tyr	Asp	Pro
Lys	Phe 2405	Glu	Гла	Gln	Leu	Gly 2410	Gln	Val	Met	Leu	Leu 2415	Val	Leu	СЛа
Val	Thr 2420	Gln	Val	Leu	Met	Met 2425	Arg	Thr	Thr	Trp	Ala 2430	Leu	Суз	Glu
Ala	Leu 2435	Thr	Leu	Ala	Thr	Gly 2440	Pro	Ile	Ser	Thr	Leu 2445	Trp	Glu	Gly
Asn	Pro 2450	Gly	Arg	Phe	Trp	Asn 2455	Thr	Thr	Ile	Ala	Val 2460	Ser	Met	Ala
Asn	Ile 2465	Phe	Arg	Gly	Ser	Tyr 2470	Leu	Ala	Gly	Ala	Gly 2475	Leu	Leu	Phe
Ser	Ile 2480	Met	ГЛЗ	Asn	Thr	Thr 2485	Asn	Thr	Arg	Arg	Gly 2490	Thr	Gly	Asn
Ile	Gly 2495	Glu	Thr	Leu	Gly	Glu 2500	Lys	Trp	Lys	Ser	Arg 2505	Leu	Asn	Ala
Leu	Gly 2510	ГЛа	Ser	Glu	Phe	Gln 2515		Tyr	Lys	ГЛа	Ser 2520	Gly	Ile	Gln
Glu	Val 2525	Asp	Arg	Thr	Leu	Ala 2530	Lys	Glu	Gly	Ile	Lys 2535	Arg	Gly	Glu
Thr	Asp 2540	His	His	Ala	Val	Ser 2545	Arg	Gly	Ser	Ala	Lys 2550	Leu	Arg	Trp
Phe	Val 2555	Glu	Arg	Asn	Met	Val 2560	Thr	Pro	Glu	Gly	Lys 2565	Val	Val	Asp
Leu	Gly 2570	Суа	Gly	Arg	Gly	Gly 2575	Trp	Ser	Tyr	Tyr	Cys 2580	-	Gly	Leu
ГЛа	Asn 2585	Val	Arg	Glu	Val	Lys 2590	Gly	Leu	Thr	ГЛа	Gly 2595	Gly	Pro	Gly
His	Glu 2600	Glu	Pro	Ile	Pro	Met 2605	Ser	Thr	Tyr	Gly	Trp 2610	Asn	Leu	Val
Arg	Leu 2615	Gln	Ser	Gly	Val	Asp 2620	Val	Phe	Phe	Ile	Pro 2625	Pro	Glu	Lys

-	CO	nt	ir	ıu	e	d
---	----	----	----	----	---	---

Сүз	Asp 2630	Thr	Leu	Leu	Суз	Asp 2635	Ile	Gly	Glu	Ser	Ser 2640	Pro	Asn	Pro
Thr	Val 2645	Glu	Ala	Gly	Arg	Thr 2650	Leu	Arg	Val	Leu	Asn 2655	Leu	Val	Glu
Asn	Trp 2660	Leu	Asn	Asn	Asn	Thr 2665	Gln	Phe	Суз	Ile	Lys 2670	Val	Leu	Asn
Pro	Tyr 2675	Met	Pro	Ser	Val	Ile 2680	Glu	Lys	Met	Glu	Ala 2685	Leu	Gln	Arg
Lys	Tyr 2690	Gly	Gly	Ala	Leu	Val 2695	Arg	Asn	Pro	Leu	Ser 2700	Arg	Asn	Ser
Thr	His 2705	Glu	Met	Tyr	Trp	Val 2710	Ser	Asn	Ala	Ser	Gly 2715	Asn	Ile	Val
Ser	Ser 2720	Val	Asn	Met	Ile	Ser 2725	Arg	Met	Leu	Ile	Asn 2730	Arg	Phe	Thr
Met	Arg 2735	Tyr	Lys	Lys	Ala	Thr 2740	Tyr	Glu	Pro	Asp	Val 2745	Asp	Leu	Gly
Ser	Gly 2750	Thr	Arg	Asn	Ile	Gly 2755	Ile	Glu	Ser	Glu	Ile 2760	Pro	Asn	Leu
Asp	Ile 2765	Ile	Gly	Lys	Arg	Ile 2770	Glu	Lys	Ile	Lys	Gln 2775	Glu	His	Glu
Thr	Ser 2780	Trp	His	Tyr	Asp	Gln 2785	Asp	His	Pro	Tyr	Lys 2790	Thr	Trp	Ala
Tyr	His 2795	Gly	Ser	Tyr	Glu	Thr 2800	Lys	Gln	Thr	Gly	Ser 2805	Ala	Ser	Ser
Met	Val 2810	Asn	Gly	Val	Val	Arg 2815	Leu	Leu	Thr	Lys	Pro 2820	Trp	Asp	Val
Val	Pro 2825	Met	Val	Thr	Gln	Met 2830	Ala	Met	Thr	Asp	Thr 2835	Thr	Pro	Phe
Gly	Gln 2840	Gln	Arg	Val	Phe	Lys 2845	Glu	Lys	Val	Asp	Thr 2850	Arg	Thr	Gln
Glu	Pro 2855	Гла	Glu	Gly	Thr	Lys 2860	Lys	Leu	Met	Гла	Ile 2865	Thr	Ala	Glu
Trp	Leu 2870	Trp	Lys	Glu	Leu	Gly 2875	Lys	Lys	Lys	Thr	Pro 2880	Arg	Met	СЛа
Thr	Arg 2885	Glu	Glu	Phe	Thr	Arg 2890	Lys	Val	Arg	Ser	Asn 2895	Ala	Ala	Leu
Gly	Ala 2900	Ile	Phe	Thr	Asp	Glu 2905	Asn	Lys	Trp	Гла	Ser 2910	Ala	Arg	Glu
Ala	Val 2915	Glu	Asp	Ser	Arg	Phe 2920	Trp	Glu	Leu	Val	Asp 2925	Lys	Glu	Arg
Asn	Leu 2930	His	Leu	Glu	Gly	Lys 2935	Cys	Glu	Thr	Суз	Val 2940	Tyr	Asn	Met
Met	Gly 2945	Lys	Arg	Glu	Lys	Lys 2950	Leu	Gly	Glu	Phe	Gly 2955	Lys	Ala	Lys
Gly	Ser 2960	Arg	Ala	Ile	Trp	Tyr 2965	Met	Trp	Leu	Gly	Ala 2970	Arg	Phe	Leu
Glu	Phe 2975	Glu	Ala	Leu	Gly	Phe 2980	Leu	Asn	Glu	Aap	His 2985	Trp	Phe	Ser
Arg	Glu 2990	Asn	Ser	Leu	Ser	Gly 2995	Val	Glu	Gly	Glu	Gly 3000	Leu	His	Lys

											-00		iue	л 
Leu	Gly 3005		Ile	Leu	Arg	Asp 3010		Ser	Lys	Lys	Glu 3015		Gly	Ala
Met	Tyr 3020			Asp	Thr	Ala 3025		Trp	Asp		Arg 3030	Ile	Thr	Leu
Glu	Asp 3035				Glu	Glu 3040		Val	Thr	Asn	His 3045	Met	Glu	Gly
Glu	His 3050				Ala	Glu 3055		Ile	Phe	Lys	Leu 3060	Thr	Tyr	Gln
Asn	Lys 3065					Gln 3070				Pro	Arg 3075		Thr	Val
Met	Asp 3080		Ile	Ser		Arg 3085						Gly	Gln	Val
Gly		-	-			Thr 3100						Ala	Gln	Leu
Ile	Arg 3110		Met	Glu		Glu 3115						Ile	Gln	His
Leu	Thr 3125		Thr	Glu	Glu	Ile 3130					_	Leu	Ala	Arg
Val	Gly	Arg				Ser 3145	Arg	Met	Ala	Ile	Ser	Gly	Asp	Asp
Сүз		Val	Lys	Pro		Asp 3160	Asp	Arg	Phe	Ala	Ser	Ala	Leu	Thr
Ala		Asn	Asp	Met		Lys 3175	Ile	Arg	Lys	Asp	Ile	Gln	Gln	Trp
Glu		Ser	Arg	Gly	Trp	Asn 3190	Asp	Trp	Thr	Gln		Pro	Phe	Суа
Ser	His	His				Leu	Ile	Met	Lys	Asp			Val	Leu
Val		Pro				Gln	Asp	Glu	Leu		Gly		Ala	Arg
Ile		Gln	Gly	Ala	Gly	3220 Trp	Ser	Leu	Arg			Ala	Суз	Leu
Gly	3230 Lys					3235 Met						Phe	His	Arg
						3250 Ala						Ala	Val	Pro
-	3260		-			3265 Ser				-	3270			
	3275	-				3280 Thr	-			-	3285			
-	3290					3295		-			3300		_	
Arg	Val 3305	Trp	тте	GIN	GLU	Asn 3310		Trp	Met	GIU	Asp 3315		Thr	Pro
Val	Glu 3320	Ser	Trp	Glu	Glu	Ile 3325		Tyr	Leu	Gly	Lуз 3330		Glu	Asp
Gln	Trp 3335	СЛа	Gly	Ser	Leu	Ile 3340	-	Leu	Thr	Ser	Arg 3345	Ala	Thr	Trp
Ala	Lys 3350	Asn	Ile	Gln	Ala	Ala 3355		Asn	Gln	Val	Arg 3360		Leu	Ile
Gly	Asn 3365	Glu	Glu	Tyr	Thr	Asp 3370	-	Met	Pro	Ser	Met 3375	-	Arg	Phe
Arg	Arg	Glu	Glu	Glu	Glu	Ala	Gly	Val	Leu	Trp				

	3380	)				338	85								
						201	-								
<213 <213 <213 <220	)> FI	ENGTI (PE : RGAN] EATUI	H: 3 PRT ISM: RE:	389 Art:	ific: TION		-		us se	erot	ype :	з, м	VS		
<400	)> SI	EQUEI	ICE :	9											
Met 1	Asn	Asn	Gln	Arg 5	Lys	Lys	Ala	Lys	Asn 10	Thr	Pro	Phe	Asn	Met 15	Leu
ГЛа	Arg	Glu	Arg 20	Asn	Arg	Val	Ser	Thr 25	Val	Gln	Gln	Leu	Thr 30	Lys	Arg
Phe	Ser	Leu 35	Gly	Met	Leu	Gln	Gly 40	Arg	Gly	Pro	Leu	Lys 45	Leu	Phe	Met
Ala	Leu 50	Val	Ala	Phe	Leu	Arg 55	Phe	Leu	Thr	Ile	Pro 60	Pro	Thr	Ala	Gly
Ile 65	Leu	Lys	Arg	Trp	Gly 70	Thr	Ile	Lys	Lys	Ser 75	Lys	Ala	Ile	Asn	Val 80
Leu	Arg	Gly	Phe	Arg 85	Lys	Glu	Ile	Gly	Arg 90	Met	Leu	Asn	Ile	Leu 95	Asn
Arg	Arg	Arg	Arg 100	Ser	Ala	Gly	Met	Ile 105	Ile	Met	Leu	Ile	Pro 110	Thr	Val
Met	Ala	Phe 115	His	Leu	Thr	Thr	Arg 120	Asp	Gly	Glu	Pro	Arg 125	Met	Ile	Val
Gly	Lys 130	Asn	Glu	Arg	Gly	Lys 135	Ser	Leu	Leu	Phe	Lys 140	Thr	Ala	Ser	Gly
Ile 145	Asn	Met	СЛа	Thr	Leu 150	Ile	Ala	Met	Asp	Leu 155	Gly	Glu	Met	Сүз	Asp 160
Asp	Thr	Val	Thr	Tyr 165	ГЛа	СЛа	Pro	His	Ile 170	Thr	Glu	Val	Glu	Pro 175	Glu
Aap	Ile	Asp	Cys 180	Trp	Суз	Asn	Leu	Thr 185	Ser	Thr	Trp	Val	Thr 190	Tyr	Gly
	-	195			-		200	-	-	-	-	205		Val	
	210				-	215	-		-		220			Thr	_
225				-	230	-	5			235	-			Thr	240
		-		245	-				250					Ala 255	
Tyr	Ile	Gly	Thr 260	Ser	Leu	Thr	Gln	Lys 265		Val	Ile	Phe	Ile 270	Leu	Leu
Met	Leu	Val 275	Thr	Pro	Ser	Met	Thr 280	Met	Arg	САа	Val	Gly 285	Val	Gly	Asn
Arg	Asp 290	Phe	Val	Glu	Gly	Leu 295	Ser	Gly	Ala	Thr	Trp 300	Val	Asp	Val	Val
Leu 305	Glu	His	Gly	Gly	Суз 310	Val	Thr	Thr	Met	Ala 315	ГЛа	Asn	ГÀа	Pro	Thr 320
Leu	Asp	Ile	Glu	Leu 325	Gln	ГЛЗ	Thr	Glu	Ala 330	Thr	Gln	Leu	Ala	Thr 335	Leu
Arg	Lys	Leu	Суз	Ile	Glu	Gly	Lys	Ile	Thr	Asn	Ile	Thr	Thr	Asp	Ser

-	CC	nt	in	ued

-				340					345					350		
				340					545					350		
A	rg	Cys	Pro 355	Thr	Gln	Gly	Glu	Ala 360	Ile	Leu	Pro	Glu	Glu 365	Gln	Asb	Gln
A	sn	Tyr 370	Val	Суз	Lys	His	Thr 375	Tyr	Val	Asp	Arg	Gly 380	Trp	Gly	Asn	Gly
	ys 85	Gly	Leu	Phe	Gly	Lys 390	Gly	Ser	Leu	Val	Thr 395	Суз	Ala	Lys	Phe	Gln 400
С	ya	Leu	Glu	Ser	Ile 405	Glu	Gly	Lys	Val	Val 410	Gln	His	Glu	Asn	Leu 415	Lys
Т	'yr	Thr	Val	Ile 420	Ile	Thr	Val	His	Thr 425	Gly	Asp	Gln	His	Gln 430	Val	Gly
A	sn	Glu	Thr 435	Gln	Gly	Val	Thr	Ala 440	Glu	Ile	Thr	Pro	Gln 445	Ala	Ser	Thr
A	la	Glu 450	Ala	Ile	Leu	Pro	Glu 455	Tyr	Gly	Thr	Leu	Gly 460	Leu	Glu	Суз	Ser
	ro 65	Arg	Thr	Gly	Leu	Asp 470	Phe	Asn	Glu	Met	Ile 475	Ser	Leu	Thr	Met	Lys 480
A	sn	Lys	Ala	Trp	Met 485	Val	His	Arg	Gln	Trp 490	Phe	Phe	Asp	Leu	Pro 495	Leu
F	ro	Trp	Thr	Ser 500	Gly	Ala	Ser	Ala	Glu 505	Thr	Pro	Thr	Trp	Asn 510	Arg	Гла
Ģ	lu	Leu	Leu 515	Val	Thr	Phe	Lys	Asn 520	Ala	His	Ala	Lys	Lys 525	Gln	Glu	Val
V	al	Val 530	Leu	Gly	Ser	Gln	Glu 535	Gly	Ala	Met	His	Thr 540	Ala	Leu	Thr	Gly
	la 45	Thr	Glu	Ile	Gln	Thr 550	Ser	Gly	Gly	Thr	Ser 555	Ile	Phe	Ala	Gly	His 560
L	eu	Lys	Cys	Arg	Leu 565	Lys	Met	Asp	ГЛЗ	Leu 570	Glu	Leu	Lys	Gly	Met 575	Ser
Т	'yr	Ala	Met	Cys 580	Leu	Ser	Ser	Phe	Val 585	Leu	Lys	ГЛЗ	Glu	Val 590	Ser	Glu
Т	'hr	Gln	His 595	Gly	Thr	Ile	Leu	Ile 600	Lys	Val	Glu	Tyr	Lys 605	Gly	Glu	Asp
A	la	Pro 610	Суз	Lys	Ile	Pro	Phe 615	Ser	Thr	Glu	Asp	Gly 620	Gln	Gly	Lys	Ala
	eu 25	Asn	Gly	Arg	Leu	Ile 630	Thr	Ala	Asn	Pro	Val 635	Val	Thr	Lys	Lys	Glu 640
G	lu	Pro	Val	Asn	Ile 645	Glu	Ala	Glu	Pro	Pro 650	Phe	Gly	Glu	Ser	Asn 655	Ile
V	al	Ile	Gly	Ile 660	Gly	Asp	Lys	Ala	Leu 665	Гла	Ile	Asn	Trp	Tyr 670	Lys	Гла
G	ly	Ser	Ser 675	Ile	Gly	Lys	Met	Phe 680	Glu	Ala	Thr	Ala	Arg 685	Gly	Ala	Arg
A	rg	Met 690	Ala	Ile	Leu	Gly	Asp 695	Thr	Ala	Trp	Asp	Phe 700	Gly	Ser	Val	Gly
	1y 05	Val	Leu	Asn	Ser	Leu 710	Gly	Lys	Met	Val	His 715	Gln	Ile	Phe	Gly	Ser 720
A	la	Tyr	Thr	Ala	Leu 725	Phe	Gly	Gly	Val	Ser 730	Trp	Met	Met	Lys	Ile 735	Gly
I	le	Gly	Val	Leu 740	Leu	Thr	Trp	Ile	Gly 745	Leu	Asn	Ser	Lys	Asn 750	Thr	Ser

-	СО	nt	in	ue	d

Met	Ser	Phe 755	Ser	Сув	Ile	Ala	Ala 760	Gly	Ile	Val	Thr	Leu 765	Tyr	Leu	Gly
Val	Met 770	Val	Gln	Ala	Asp	Ser 775	Gly	Сүз	Val	Val	Ser 780	Trp	Lys	Asn	Lys
Glu 785	Leu	Lys	Суз	Gly	Ser 790	Gly	Ile	Phe	Ile	Thr 795	Asp	Asn	Val	His	Thr 800
Trp	Thr	Glu	Gln	Tyr 805	LÀa	Phe	Gln	Pro	Glu 810	Ser	Pro	Ser	Lys	Leu 815	Ala
Ser	Ala	Ile	Gln 820	Lys	Ala	His	Glu	Glu 825	Asp	Ile	СЛа	Gly	Ile 830	Arg	Ser
Val	Thr	Arg 835	Leu	Glu	Asn	Leu	Met 840	Trp	Lys	Gln	Ile	Thr 845	Pro	Glu	Leu
Asn	His 850	Ile	Leu	Ser	Glu	Asn 855	Glu	Val	Lys	Leu	Thr 860	Ile	Met	Thr	Gly
Asp 865	Ile	Lys	Gly	Ile	Met 870	Gln	Ala	Gly	Lys	Arg 875	Ser	Leu	Arg	Pro	Gln 880
Pro	Thr	Glu	Leu	Lys 885	Tyr	Ser	Trp	Lys	Thr 890	Trp	Gly	Lys	Ala	Lys 895	Met
Leu	Ser	Thr	Glu 900	Ser	His	Asn	Gln	Thr 905	Phe	Leu	Ile	Asp	Gly 910	Pro	Glu
Thr	Ala	Glu 915	Cys	Pro	Asn	Thr	Asn 920	Arg	Ala	Trp	Asn	Ser 925	Leu	Glu	Val
Glu	Asp 930	Tyr	Gly	Phe	Gly	Val 935	Phe	Thr	Thr	Asn	Ile 940	Trp	Leu	Lys	Leu
Lys 945	Glu	Lys	Gln	Asp	Val 950	Phe	Сүз	Asp	Ser	Lys 955	Leu	Met	Ser	Ala	Ala 960
Ile	Lys	Asp	Asn	Arg 965	Ala	Val	His	Ala	Asp 970	Met	Gly	Tyr	Trp	Ile 975	Glu
Ser	Ala	Leu	Asn 980	Asp	Thr	Trp	Гла	Ile 985	Glu	Lys	Ala	Ser	Phe 990	Ile	Glu
Val	Lys	Asn 995	Суз	His	Trp	Pro	Lys 100		r Hi	s Thi	r Le	u Tr <u>:</u> 10		er A	sn Gly
Val	Leu 1010		ı Sei	r Glı	ı Met	: Ile 10:		le Pi	ro L <u>i</u>	ys A:		eu . 020	Ala	Gly	Pro
Val	Ser 1025		n Hi:	s Ası	а Туз	r Arg 103		ro G	ly T	yr H:		hr 035	Gln	Ile	Thr
Gly	Pro 1040		) Hi:	: Lei	ı Gly	7 Ly: 104		eu G	lu Me	et A:		he . 050	Asp	Phe	Суз
Asp	Gly 1055		r Thi	r Val	l Val	L Va 100		hr G	lu A	ab CJ		ly . 065	Asn .	Arg	Gly
Pro	Ser 1070		ı Arç	g Th:	r Thi	7 Th: 107		la Se	er G	ly Ly	•	eu 080	Ile	Thr	Glu
Trp	Cys 1089	-	a Arç	g Se:	r Cys	5 Th: 109		eu Pi	ro P:	ro Le		rg 095	Tyr .	Arg	Gly
Glu	Asp 1100		y Cys	s Trj	р Туз	f Gly 110		et Gi	lu I	le An		ro 110	Leu	Lys -	Glu
ГЛа	Glu 1119		ı Ası	ı Leı	ı Val	l Ası 112		er Le	eu Va	al Tł		la 125	Gly :	His	Gly
Gln	Val 1130		) Ası	n Phe	e Sei	: Le: 113		ly Va	al Le	eu Gi		et . 140	Ala	Leu	Phe

-	CC	nt	in	ued

											-00	.1011	iuec	л 
Leu	Glu 1145	Glu	Met	Leu	Arg	Thr 1150		Val	Gly	Thr	Lys 1155		Ala	Ile
Leu	Leu 1160		Ala	Val		Phe 1165		Thr	Leu	Ile	Thr 1170	-	Asn	Met
Ser	Phe 1175	-	_	Leu		Arg 1180		Met	Val	Met	Val 1185		Ala	Thr
Met	Thr 1190	-	_	Ile		Met 1195	-	Val	Thr	-	Leu 1200	Ala	Leu	Leu
Ala	Ala 1205					Pro 1210		Phe	Ala	Ala	Gly 1215	Leu	Leu	Leu
Arg	Lys 1220		Thr	Ser	Lys	Glu 1225		Met	Met	Thr	Thr 1230	Ile	Gly	Ile
Val	Leu 1235					Thr 1240		Pro	Glu		Ile 1245	Leu	Glu	Leu
Thr	Asp 1250		Leu	Ala		Gly 1255		Met	Val		Lys 1260	Met	Val	Arg
Asn	Met 1265		-	-	Gln	Leu 1270		Val	Thr	Ile	Met 1275	Ala	Ile	Leu
Суз		Pro			Val	Ile 1285	Leu	Gln	Asn	Ala		-	Val	Ser
Сув		Ile				Val	Ser	Val	Ser	Pro			Leu	Thr
Ser		Gln				1300 Asp 1315	Trp					Leu	Thr	Ile
	Gly	Leu				Ala	Ile				Thr	Leu	Ser	Arg
		Lys	Lys	Arg	Ser	1330 Trp	Pro	Leu	Asn	Glu		Ile	Met	Ala
Val	-	Met	Val	Ser	Ile	1345 Leu	Ala	Ser	Ser	Leu		Lys	Asn	Asp
Ile		Met	Thr	Gly	Pro	1360 Leu	Val					Leu	Thr	Val
Cys	1370 Tyr			Thr		1375 Arg					1380 Glu	Leu	Glu	Arg
-	1385				-	1390 Glu			-		1395			-
	1400	-		-	-	1405	-				1410		-	
	1415					Thr 1420					1425			
Ile	Lys 1430	Asn	Glu	Glu	Glu	Glu 1435		Thr	Leu	Thr	Ile 1440		Ile	Arg
Thr	Gly 1445	Leu	Leu	Val	Ile	Ser 1450		Leu	Phe	Pro	Val 1455	Ser	Ile	Pro
Ile	Thr 1460	Ala	Ala	Ala	Trp	Tyr 1465		Trp	Glu	Val	Lys 1470		Gln	Arg
Ala	Gly 1475	Val	Leu	Trp	Asp	Val 1480		Ser	Pro	Pro	Pro 1485	Met	Gly	ГЛа
Ala	Glu 1490	Leu	Glu	Asp	Gly	Ala 1495	-	Arg	Ile	Lys	Gln 1500	Lys	Gly	Ile
Leu	Gly 1505	-	Ser	Gln	Ile	Gly 1510		Gly	Val	Tyr	Lys 1515	Glu	Gly	Thr
Phe			Met	Trp	His	Val		Arg	Gly	Ala		Leu	Met	His

-continued

	1520					1525					1530			
Lys	Gly 1535	Lys	Arg	Ile	Glu	Pro 1540		Trp	Ala	Asp	Val 1545	Lys	Lys	Asp
Leu	Ile 1550	Ser	Tyr	Gly	Gly	Gly 1555		Lys	Leu	Glu	Gly 1560	Glu	Trp	Lys
Glu	Gly 1565	Glu	Glu	Val	Gln	Val 1570		Ala	Leu	Glu	Pro 1575	Gly	Lys	Asn
Pro	Arg 1580	Ala	Val	Gln	Thr	Lys 1585		Gly	Leu	Phe	Lys 1590	Thr	Asn	Ala
Gly	Thr 1595	Ile	Gly	Ala	Val	Ser 1600		Asp	Phe	Ser	Pro 1605	Gly	Thr	Ser
Gly	Ser 1610	Pro	Ile	Ile	Asp	Lys 1615		Gly	Lys	Val	Val 1620	Gly	Leu	Tyr
Gly	Asn 1625	Gly	Val	Val	Thr	Arg 1630		Gly	Ala	Tyr	Val 1635	Ser	Ala	Ile
Ala	Gln 1640	Thr	Glu	Lys	Ser	Ile 1645	Glu	Aab	Asn	Pro	Glu 1650	Ile	Glu	Aap
Asp	Ile 1655	Phe	Arg	Гла	Arg	Arg 1660		Thr	Ile	Met	Asp 1665	Leu	His	Pro
Gly	Ala 1670	Gly	ГЛа	Thr	ГЛЗ	Arg 1675		Leu	Pro	Ala	Ile 1680	Val	Arg	Glu
Ala	Ile 1685	Lys	Arg	Gly	Leu	Arg 1690		Leu	Ile	Leu	Ala 1695	Pro	Thr	Arg
Val	Val 1700	Ala	Ala	Glu	Met	Glu 1705	Glu	Ala	Leu	Arg	Gly 1710	Leu	Pro	Ile
Arg	Tyr 1715	Gln	Thr	Pro	Ala	Ile 1720	-	Ala	Val	His	Thr 1725	Gly	Arg	Glu
Ile	Val 1730	Asp	Leu	Met	Суз	His 1735		Thr	Phe	Thr	Met 1740	Arg	Leu	Leu
Ser	Pro 1745	Val	Arg	Val	Pro	Asn 1750		Asn	Leu	Ile	Ile 1755	Met	Asp	Glu
Ala	His 1760	Phe	Thr	Asp	Pro	Ala 1765	Ser	Ile	Ala	Ala	Arg 1770	Gly	Tyr	Ile
Ser	Thr 1775	Arg	Val	Glu	Met	Gly 1780	Glu	Ala	Ala	Gly	Ile 1785	Phe	Met	Thr
Ala	Thr 1790	Pro	Pro	Gly	Ser	Arg 1795		Pro	Phe	Pro	Gln 1800	Ser	Asn	Ala
Pro	Ile 1805	Ile	Asp	Glu	Glu	Arg 1810	Glu	Ile	Pro	Glu	Arg 1815	Ser	Trp	Asn
Ser	Gly 1820	His	Glu	Trp	Val	Thr 1825	Asp	Phe	Lys	Gly	Lys 1830	Thr	Val	Trp
Phe	Val 1835	Pro	Ser	Ile	Lys	Ala 1840	-	Asn	Asp	Ile	Ala 1845	Ala	Суз	Leu
Arg	Lys 1850	Asn	Gly	Lys	Lys	Val 1855		Gln	Leu	Ser	Arg 1860	Lys	Thr	Phe
Asp	Ser 1865	Glu	Tyr	Val	Lys	Thr 1870	Arg	Thr	Asn	Asp	Trp 1875	Asp	Phe	Val
Val	Thr 1880	Thr	Asp	Ile	Ser	Glu 1885	Met	Gly	Ala	Asn	Phe 1890	Lys	Ala	Glu
Arg	Val 1895	Ile	Asp	Pro	Arg	Arg 1900	Суз	Met	Lys	Pro	Val 1905	Ile	Leu	Thr

-cont	inued

Asp	Gly 1910	Glu	Glu	Arg	Val	Ile 1915	Leu	Ala	Gly	Pro	Met 1920	Pro	Val	Thr
His	Ser 1925	Ser	Ala	Ala	Gln	Arg 1930	Arg	Gly	Arg	Ile	Gly 1935	Arg	Asn	Pro
Lys	Asn 1940	Glu	Asn	Asp	Gln	Tyr 1945	Ile	Tyr	Met	Gly	Glu 1950	Pro	Leu	Glu
Asn	Asp 1955	Glu	Asp	Сүз	Ala	His 1960	Trp	Lys	Glu	Ala	Lys 1965	Met	Leu	Leu
Aap	Asn 1970	Ile	Asn	Thr	Pro	Glu 1975	Gly	Ile	Ile	Pro	Ser 1980	Met	Phe	Glu
Pro	Glu 1985	Arg	Glu	Lys	Val	Asp 1990	Ala	Ile	Asp	Gly	Glu 1995	Tyr	Arg	Leu
Arg	Gly 2000	Glu	Ala	Arg	Lys	Thr 2005	Phe	Val	Aap	Leu	Met 2010	Arg	Arg	Gly
Asp	Leu 2015	Pro	Val	Trp	Leu	Ala 2020	Tyr	Arg	Val	Ala	Ala 2025	Glu	Gly	Ile
Asn	Tyr 2030	Ala	Asp	Arg	Arg	Trp 2035	Сув	Phe	Aap	Gly	Val 2040	Lys	Asn	Asn
Gln	Ile 2045	Leu	Glu	Glu	Asn	Val 2050	Glu	Val	Glu	Ile	Trp 2055	Thr	Lys	Glu
Gly	Glu 2060	Arg	Lys	Lys	Leu	Lys 2065	Pro	Arg	Trp	Leu	Asp 2070	Ala	Arg	Ile
Tyr	Ser 2075	Asp	Pro	Leu	Ala	Leu 2080	Lys	Glu	Phe	Lys	Glu 2085	Phe	Ala	Ala
Gly	Arg 2090	Lys	Ser	Leu	Thr	Leu 2095	Asn	Leu	Ile	Thr	Glu 2100	Met	Gly	Arg
Leu	Pro 2105	Thr	Phe	Met	Thr	Gln 2110	Lys	Ala	Arg	Asp	Ala 2115	Leu	Asp	Asn
Leu	Ala 2120	Val	Leu	His	Thr	Ala 2125	Glu	Ala	Gly	Gly	Arg 2130	Ala	Tyr	Asn
His	Ala 2135	Leu	Ser	Glu	Leu	Pro 2140	Glu	Thr	Leu	Glu	Thr 2145	Leu	Leu	Leu
Leu	Thr 2150	Leu	Leu	Ala	Thr	Val 2155	Thr	Gly	Gly	Ile	Phe 2160	Leu	Phe	Leu
Met	Ser 2165	Ala	Arg	Gly	Ile	Gly 2170	Lys	Met	Thr	Leu	Gly 2175	Met	Cys	Cys
Ile	Ile 2180	Thr	Ala	Ser	Ile	Leu 2185	Leu	Trp	Tyr	Ala	Gln 2190	Ile	Gln	Pro
His	Trp 2195	Ile	Ala	Ala	Ser	Ile 2200	Ile	Leu	Glu	Phe	Phe 2205	Leu	Ile	Val
Leu	Leu 2210	Ile	Pro	Glu	Pro	Glu 2215	Lys	Gln	Arg	Thr	Pro 2220	Gln	Asp	Asn
Gln	Leu 2225	Thr	Tyr	Val	Val	Ile 2230	Ala	Ile	Leu	Thr	Val 2235	Val	Ala	Ala
Thr	Met 2240	Ala	Asn	Glu	Met	Gly 2245	Phe	Leu	Glu	ГЛа	Thr 2250	ГÀа	ГЛа	Asp
Leu	Gly 2255	Leu	Gly	Ser	Ile	Ala 2260	Thr	Gln	Gln	Pro	Glu 2265	Ser	Asn	Ile
Leu	Asp 2270	Ile	Asp	Leu	Arg	Pro 2275	Ala	Ser	Ala	Trp	Thr 2280		Tyr	Ala

												.1011	Iuec	
Val	Ala 2285		Thr	Phe	Val	Thr 2290		Met	Leu	Arg	His 2295		Ile	Glu
Asn	Ser 2300		Val	Asn	Val	Ser 2305		Thr	Ala	Ile	Ala 2310	Asn	Gln	Ala
Thr	Val 2315		Met	Gly	Leu	Gly 2320	-	Gly	Trp	Pro	Leu 2325	Ser	Lys	Met
Asp	Ile 2330	-	Val	Pro	Leu	Leu 2335		Ile	Gly	Суз	Tyr 2340	Ser	Gln	Val
Asn	Pro 2345		Thr	Leu	Thr	Ala 2350		Leu	Phe	Leu	Leu 2355	Val	Ala	His
Tyr	Ala 2360		Ile	Gly		Gly 2365		Gln	Ala	Lys	Ala 2370	Thr	Arg	Glu
Ala	Gln 2375		Arg	Ala	Ala	Ala 2380					Asn 2385	Pro	Thr	Val
Asp	Gly 2390		Thr	Val	Ile	Asp 2395					Pro 2400		Asp	Pro
Lys	Phe 2405		Lys	Gln	Leu	Gly 2410		Val	Met	Leu	Leu 2415	Val	Leu	Суз
Val	Thr 2420		Val	Leu	Met	Met 2425					Ala 2430	Leu	Сүа	Glu
Ala		Thr	Leu	Ala	Thr		Pro				Leu 2445		Glu	Gly
Asn		Gly	-		Trp		Thr	Thr	Ile	Ala	Val 2460		Met	Ala
Asn		Phe			Ser		Leu	Ala	Gly	Ala	Gly 2475	Leu	Leu	Phe
Ser		Met	Гла	Asn	Thr		Asn	Thr	Arg	Arg	Gly 2490	Thr	Gly	Asn
Ile	Gly	Glu	Thr	Leu	Gly	Glu	Lys	Trp	Lys	Ser	Arg	Leu	Asn	Ala
Leu	-	Lys	Ser	Glu	Phe		Ile				2505 Ser	Gly	Ile	Gln
Glu		Asp					Lys	Glu		Ile	2520 Lys		Gly	Glu
Thr		His					Arg		Ser		2535 Lys		Arg	Trp
Phe	2540 Val		Arg	Asn	Met	2545 Val		Pro	Glu	Gly	2550 Lys	Val	Val	Asp
	2555					2560					2565 Сув			
	2570	-	-	-	-	2575	-		-	-	2580 Gly	-	-	
-	2585		-			2590	-			-	2595 Trp	-		-
	2600					2605			-	-	2610			
Arg	Leu 2615	Gln	Ser	Gly	Val	Asp 2620		Phe	Phe	Ile	Pro 2625	Pro	Glu	ГЛа
Суз	Aap 2630		Leu	Leu	Сув	Asp 2635		Gly	Glu	Ser	Ser 2640	Pro	Asn	Pro
Thr	Val 2645		Ala	Gly	Arg	Thr 2650		Arg	Val	Leu	Asn 2655	Leu	Val	Glu
Asn	Trp	Leu	Asn	Asn	Asn	Thr	Gln	Phe	Суа	Ile	Lys	Val	Leu	Asn

-continued

	2660					2665					2670			
Pro	Tyr 2675	Met	Pro	Ser	Val	Ile 2680	Glu	Lys	Met	Glu	Ala 2685	Leu	Gln	Arg
Lys	Tyr 2690	Gly	Gly	Ala	Leu	Val 2695	Arg	Asn	Pro	Leu	Ser 2700	Arg	Asn	Ser
Thr	His 2705	Glu	Met	Tyr	Trp	Val 2710	Ser	Asn	Ala	Ser	Gly 2715	Asn	Ile	Val
Ser	Ser 2720	Val	Asn	Met	Ile	Ser 2725	Arg	Met	Leu	Ile	Asn 2730	-	Phe	Thr
Met	Arg 2735	Tyr	Lys	Lys	Ala	Thr 2740		Glu	Pro	Asp	Val 2745	Asp	Leu	Gly
Ser	Gly 2750	Thr	Arg	Asn	Ile	Gly 2755	Ile	Glu	Ser	Glu	Ile 2760		Asn	Leu
Asp	Ile 2765	Ile	Gly	Lys	Arg	Ile 2770	Glu	Lys	Ile	Lys	Gln 2775	Glu	His	Glu
Thr	Ser 2780	Trp	His	Tyr	Asp	Gln 2785	Aab	His	Pro	Tyr	Lys 2790		Trp	Ala
Tyr	His 2795	Gly	Ser	Tyr	Glu	Thr 2800	Lys	Gln	Thr	Gly	Ser 2805	Ala	Ser	Ser
Met	Val 2810	Asn	Gly	Val	Val	Arg 2815	Leu	Leu	Thr	ГÀа	Pro 2820		Asp	Val
Val	Pro 2825	Met	Val	Thr	Gln	Met 2830	Ala	Met	Thr	Asp	Thr 2835		Pro	Phe
Gly	Gln 2840	Gln	Arg	Val	Phe	Lys 2845	Glu	Lys	Val	Asp	Thr 2850		Thr	Gln
Glu	Pro 2855	ГЛЗ	Glu	Gly	Thr	Lys 2860	Lys	Leu	Met	ГЛЗ	Ile 2865		Ala	Glu
Trp	Leu 2870	Trp	Lys	Glu	Leu	Gly 2875	Lys	Lys	Lys	Thr	Pro 2880		Met	Суз
Thr	Arg 2885	Glu	Glu	Phe	Thr	Arg 2890	Lys	Val	Arg	Ser	Asn 2895		Ala	Leu
Gly	Ala 2900	Ile	Phe	Thr	Asp	Glu 2905	Asn	Lys	Trp	ГЛЗ	Ser 2910		Arg	Glu
Ala	Val 2915	Glu	Asp	Ser	Arg	Phe 2920	Trp	Glu	Leu	Val	Asp 2925		Glu	Arg
Asn	Leu 2930	His	Leu	Glu	Gly	Lys 2935		Glu	Thr	Сүз	Val 2940		Asn	Met
Met	Gly 2945	Lys	Arg	Glu	Lys	Lys 2950	Leu	Gly	Glu	Phe	Gly 2955		Ala	Lys
Gly	Ser 2960	Arg	Ala	Ile	Trp	Tyr 2965	Met	Trp	Leu	Gly	Ala 2970		Phe	Leu
Glu	Phe 2975	Glu	Ala	Leu	Gly	Phe 2980	Leu	Asn	Glu	Asp	His 2985	Trp	Phe	Ser
Arg	Glu 2990	Asn	Ser	Leu	Ser	Gly 2995	Val	Glu	Gly	Glu	Gly 3000	Leu	His	Lys
Leu	Gly 3005	Tyr	Ile	Leu	Arg	Asp 3010	Val	Ser	Lys	Lys	Glu 3015	_	Gly	Ala
Met	Tyr 3020	Ala	Asp	Asp	Thr	Ala 3025	Gly	Trp	Asp	Thr	Arg 3030		Thr	Leu
Glu	Asp 3035	Leu	Lys	Asn	Glu	Glu 3040	Met	Val	Thr	Asn	His 3045	Met	Glu	Gly

-	con	tir	lued	

Glu	His 3050	Гла	Гла	Leu	Ala	Glu 3055	Ala	Ile	Phe	Гла	Leu 3060	Thr	Tyr	Gln
Asn	Lys 3065	Val	Val	Arg	Val	Gln 3070	Arg	Pro	Thr	Pro	Arg 3075	Gly	Thr	Val
Met	Asp 3080	Ile	Ile	Ser	Arg	Arg 3085	Asp	Gln	Arg	Gly	Ser 3090	Gly	Gln	Val
Gly	Thr 3095	Tyr	Gly	Leu	Asn	Thr 3100	Phe	Thr	Asn	Met	Glu 3105	Ala	Gln	Leu
Ile	Arg 3110	Gln	Met	Glu	Gly	Glu 3115	Gly	Val	Phe	Lys	Ser 3120	Ile	Gln	His
Leu	Thr 3125	Ile	Thr	Glu	Glu	Ile 3130	Ala	Val	Gln	Asn	Trp 3135	Leu	Ala	Arg
Val	Gly 3140	Arg	Glu	Arg	Leu	Ser 3145	Arg	Met	Ala	Ile	Ser 3150	Gly	Asp	Aap
Cys	Val 3155	Val	Lys	Pro	Leu	Asp 3160	Aab	Arg	Phe	Ala	Ser 3165	Ala	Leu	Thr
Ala	Leu 3170	Asn	Asp	Met	Gly	Lys 3175	Ile	Arg	ГÀа	Asp	Ile 3180	Gln	Gln	Trp
Glu	Pro 3185	Ser	Arg	Gly	Trp	Asn 3190	Asp	Trp	Thr	Gln	Val 3195	Pro	Phe	Сув
Ser	His 3200	His	Phe	His	Glu	Leu 3205	Ile	Met	Lys	Asp	Gly 3210	Arg	Val	Leu
Val	Val 3215	Pro	Суз	Arg	Asn	Gln 3220	Asp	Glu	Leu	Ile	Gly 3225	Arg	Ala	Arg
Ile	Ser 3230	Gln	Gly	Ala	Gly	Trp 3235	Ser	Leu	Arg	Glu	Thr 3240	Ala	Суз	Leu
Gly	Lys 3245	Ser	Tyr	Ala	Gln	Met 3250	Trp	Ser	Leu	Met	Tyr 3255	Phe	His	Arg
Arg	Asp 3260	Leu	Arg	Leu	Ala	Ala 3265	Asn	Ala	Ile	Суз	Ser 3270	Ala	Val	Pro
Ser	His 3275	Trp	Val	Pro	Thr	Ser 3280	Arg	Thr	Thr	Trp	Ser 3285	Ile	His	Ala
Lys	His 3290	Glu	Trp	Met	Thr	Thr 3295	Glu	Asp	Met	Leu	Thr 3300	Val	Trp	Asn
Arg	Val 3305	Trp	Ile	Gln	Glu	Asn 3310	Pro	Trp	Met	Glu	Asp 3315	Lys	Thr	Pro
Val	Glu 3320	Ser	Trp	Glu	Glu	Ile 3325	Pro	Tyr	Leu	Gly	Lуя 3330		Glu	Asp
Gln	Trp 3335	Суз	Gly	Ser	Leu	Ile 3340	Gly	Leu	Thr	Ser	Arg 3345	Ala	Thr	Trp
Ala	Lys 3350		Ile	Gln	Ala	Ala 3355		Asn	Gln	Val	Arg 3360		Leu	Ile
Gly	Asn 3365	Glu	Glu	Tyr	Thr	Asp 3370	-	Met	Pro	Ser	Met 3375	-	Arg	Phe
Arg	Arg 3380	Glu	Glu	Glu	Glu	Ala 3385	Gly	Val	Leu	Trp				
<210	)> SE(		NO 1											

<210> SEQ ID NO 10 <211> LENGTH: 10723 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE:

121

-continued

				-contir	nued		
<pre>&lt;220&gt; FEAT &lt;221&gt; NAME, &lt;222&gt; LOCA' &lt;223&gt; OTHEI &lt;220&gt; FEAT &lt;221&gt; NAME, &lt;222&gt; LOCA' &lt;223&gt; OTHEI &lt;220&gt; FEAT &lt;222&gt; LOCA' &lt;221&gt; NAME, &lt;221&gt; NAME,</pre>	JRE: /KEY: n TION: (3773) R INFORMATI( JRE: /KEY: n TION: (7026) R INFORMATI( JRE:	)(3773) DN: a or g )(7026) DN: t or c )(7538)	virus seroty	ype 4, BVS			
<400> SEQUI	ENCE: 10						
agttgttagt	ctacgtggac	cgacaaagac	agattctttg	agggagctaa	gctcaatgta	60	
gttctaacag	tttttaatt	agagagcaga	tctctgatga	ataaccaacg	gaaaaaggcg	120	
aaaaacacgc	ctttcaatat	gctgaaacgc	gagagaaacc	gcgtgtcgac	tgtgcaacag	180	
ctgacaaaga	gattctcact	tggaatgctg	cagggacgag	gacctttaaa	actgttcatg	240	
gccctggtgg	cgttccttcg	tttcctaaca	atcccaccaa	cagcagggat	attgaagaga	300	
tggggaacaa	ttaaaaaatc	aaaagctatt	aatgttttga	gagggttcag	gaaagagatt	360	
ggaaggatgc	tgaacatctt	gaataggaga	cgcagctctg	caggcatgat	cattatgctg	420	
attccaacag	tgatggcgtt	ccatttaacc	acgcgtgatg	gcgaacccct	catgatagtg	480	
gcaaaacatg	aaaggggggag	acctctcttg	tttaagacaa	cagaggggat	caacaaatgc	540	
actctcattg	ccatggactt	gggtgaaatg	tgtgaggaca	ctgtcacgta	taaatgcccc	600	
ttactggtca	ataccgaacc	tgaagacatt	gattgctggt	gcaatctcac	gtctacctgg	660	
gtcatgtatg	ggacatgcac	ccagagcgga	gaacggagac	gagagaagcg	ctcagtagct	720	
ttaacaccac	attcaggaat	gggattggaa	acaagagctg	agacatggat	gtcatcggaa	780	
ggggcttgga	agcatgctca	gagagtagag	agctggatac	tcagaaaccc	aggattcgcg	840	
ctcttggcag	gatttatggc	ttatatgatt	gggcaaacag	gaatccagcg	aactgtcttc	900	
tttgtcctaa	tgatgctggt	cgccccatcc	tacggaatgc	gatgcgtagg	agtaggaaac	960	
agagactttg	tggaaggagt	ctcaggtgga	gcatgggtcg	atctggtgct	agaacatgga	1020	
ggatgcgtca	caaccatggc	ccagggaaaa	ccaaccttgg	attttgaact	gactaagaca	1080	
acagccaagg	aagtggetet	gttaagaacc	tattgcattg	aagceteaat	atcaaacata	1140	
accacggcaa	caagatgtcc	aacgcaagga	gagccttatc	taaaagagga	acaagaccaa	1200	
cagtacattt	gccggagaga	tgtggtagac	agagggtggg	gcaatggctg	tggcttgttt	1260	
ggaaaaggag	gagttgtgac	atgtgcgaag	ttttcatgtt	cggggaagat	aacaggcaat	1320	
ttggtccaaa	ttgagaacct	tgaatacaca	gtggttgtaa	cagtccacaa	tggagacacc	1380	
catgcagtag	gaaatgacac	gtccaatcat	ggagttacag	ccacgataac	tcccaggtca	1440	
ccatcggtgg	aagtcaaatt	gccggactat	ggagaactaa	cactcgattg	tgaacccagg	1500	
tctggaattg	actttaatga	gatgattctg	atgaaaatga	aaaagaaaac	atggcttgtg	1560	
cataagcaat	ggtttttgga	tctacctcta	ccatggacag	caggagcaga	cacatcagag	1620	
gttcactgga	attacaaaga	gagaatggtg	acatttaagg	ttcctcatgc	caagagacag	1680	
gatgtgacag	tgctgggatc	tcaggaagga	gccatgcatt	ctgccctcgc	tggagccaca	1740	
gaagtggact	ccggtgatgg	aaatcacatg	tttgcaggac	atctcaagtg	caaagtccgt	1800	

				-cont II	lucu	
atggagaaat	tgagaatcaa	gggaatgtca	tacacgatgt	gttcaggaaa	gttctcaatt	1860
gacaaagaga	tggcagaaac	acagcatggg	acaacagtgg	tgaaagtcaa	gtatgaaggt	1920
gctggagctc	cgtgtaaagt	ccccatagag	ataagagatg	tgaacaagga	aaaagtggtt	1980
gggcgtatca	tctcatccac	ccctttggct	gagaatacca	acagtgtaac	caacatagag	2040
ttagaacccc	cctttgggga	cagctacata	gtgataggtg	ttggaaacag	tgcattaaca	2100
ctccattggt	tcaggaaagg	gagttccatt	ggcaagatgt	ttgagtccac	atacagaggt	2160
gcaaaacgaa	tggccattct	aggtgaaaca	gcttgggatt	ttggttccgt	tggtggactg	2220
ttcacatcat	tgggaaaggc	tgtgcaccag	gtttttggaa	gtgtgtatac	aaccctgttt	2280
ggaggagtct	catggatgat	tagaateeta	attgggttcc	tagtgttgtg	gattggcacg	2340
aactcaagga	acacttcaat	ggctatgacg	tgcatagctg	ccggcattgt	gacactgtat	2400
ttgggagtca	tggtgcaggc	cgatagtggt	tgcgttgtga	gctggaaaaa	caaagaactg	2460
aaatgtggca	gtgggatttt	catcacagac	aacgtgcaca	catggacaga	acaatacaag	2520
ttccaaccag	aatccccttc	aaaactagct	tcagctatcc	agaaagccca	tgaagaggac	2580
atttgtggaa	tccgctcagt	aacaagactg	gagaatctga	tgtggaaaca	aataacacca	2640
gaattgaatc	acattctatc	agaaaatgag	gtgaagttaa	ctattatgac	aggagacatc	2700
aaaggaatca	tgcaggcagg	aaaacgatct	ctgcggcctc	agcccactga	gctgaagtat	2760
tcatggaaaa	catggggcaa	agcaaaaatg	ctctctacag	agtctcataa	ccagaccttt	2820
ctcattgatg	gccccgaaac	agcagaatgc	cccaacacaa	atagagcttg	gaattcgttg	2880
gaagttgaag	actatggctt	tggagtattc	accaccaata	tatggctaaa	attgaaagaa	2940
aaacaggatg	tattctgcga	ctcaaaactc	atgtcagcgg	ccataaaaga	caacagagcc	3000
gtccatgccg	atatgggtta	ttggatagaa	agtgcactca	atgacacatg	gaagatagag	3060
aaagcetett	tcattgaagt	taaaaactgc	cactggccaa	aatcacacac	cctctggagc	3120
aatggagtgc	tagaaagtga	gatgataatt	ccaaagaatc	tcgctggacc	agtgtctcaa	3180
cacaactata	gaccaggcta	ccatacacaa	ataacaggac	catggcatct	aggtaagctt	3240
gagatggact	ttgatttctg	tgatggaaca	acagtggtag	tgactgagga	ctgcggaaat	3300
agaggaccct	ctttgagaac	aaccactgcc	tctggaaaac	tcataacaga	atggtgctgc	3360
cgatcttgca	cattaccacc	gctaagatac	agaggtgagg	atgggtgctg	gtacgggatg	3420
gaaatcagac	cattgaagga	gaaagaagag	aatttggtca	actccttggt	cacagctgga	3480
catgggcagg	tcgacaactt	ttcactagga	gtcttgggaa	tggcattgtt	cctggaggaa	3540
atgettagga	cccgagtagg	aacgaaacat	gcaatactac	tagttgcagt	ttcttttgtg	3600
acattgatca	cagggaacat	gtcctttaga	gacctgggaa	gagtgatggt	tatggtaggc	3660
gccactatga	cgggtgacat	aggtatgggc	gtgacttatc	ttgccctact	agcagccttc	3720
aaagtcagac	caacttttgc	agctggacta	ctcttgagaa	agctgacctc	canggaattg	3780
atgatgacta	ctataggaat	tgtactcctc	tcccagagca	ccataccaga	gaccattett	3840
gagttgactg	atgcgttagc	cttaggcatg	atggtcctca	aaatggtgag	aaatatggaa	3900
aagtatcaat	tggcagtgac	tatcatggct	atcttgtgcg	tcccaaacgc	agtgatatta	3960
caaaacgcat	ggaaagtgag	ttgcacaata	ttggcagtgg	tgtccgtttc	cccactgttc	4020
ttaacatcct	cacagcaaaa	aacagattgg	ataccattag	cattgacgat	caaaggtctc	4080

-continued	
aatccaacag ctatttttct aacaaccctc tcaagaacca gcaagaaaag gagctggcca	4140
ttaaatgagg ctatcatggc agtcgggatg gtgagcattt tagccagttc tctcctaaaa	4200
aatgatatte ceatgaeagg aceattagtg getggaggge teeteactgt gtgetaegtg	4260
ctcactggac gatcggccga tttggaactg gagagagcag ccgatgtcaa atgggaagac	4320
caggcagaga tatcaggaag cagtccaatc ctgtcaataa caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga ggaagaacaa acactgacca tactcattag aacaggattg	4440
ctggtgatct caggactttt tcctgtatca ataccaatca cggcagcagc atggtacctg	4500
tgggaagtga agaaacaacg ggccggagta ttgtgggatg ttccttcacc cccacccatg	4560
ggaaaggctg aactggaaga tggagcctat agaattaagc aaaaagggat tcttggatat	4620
teecagateg gageeggagt ttacaaagaa ggaacattee atacaatgtg geatgteaca	4680
cgtggcgctg ttctaatgca taaaggaaag aggattgaac catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg aggaggctgg aagttagaag gagaatggaa ggaaggagaa	4800
gaagteeagg tattggeact ggageetgga aaaaateeaa gageegteea aaegaaaeet	4860
ggtettttea aaaccaaege eggaacaata ggtgetgtat etetggaett tteteetgga	4920
acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt	4980
gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa	5040
gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc	5100
cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa	5160
cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa	5220
gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg	5280
cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga tccagcaagt	5400
atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt	5460
atgacageca eteceeegg aageagagae ceattteete agageaatge aceaateata	5520
gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct	5640
tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg	5760
ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattetg geaggaeeta tgeeagtgae ceaetetagt	5880
gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060
gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta	6240
gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt	6360
agaaggaagg acgeeaggae eeaceegae eeaceggege baaaagaact taaggaatt	

-continued	
gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc	6420
ttcatgactc agaaggtaag agacgcactg gacaacttag cagtgctgca cacggctgag	6480
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttetgge tacagteacg ggagggatet ttttattett gatgagegea	6600
agggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccateet cacagtggtg geegeaacea tggeaaaega gatgggttte	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc	6900
aacateetgg acatagatet acgteetgea teageatgga egetgtatge egtggeeaca	6960
acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta	7020
acagenatag ceaaceaage eacagtgtta atgggteteg ggaaaggatg geeattgtea	7080
aagatggaca teggagttee eettetegee attggatget acteacaagt caaceceata	7140
actctcacag cagctctttt cttattggta gcacattatg ccatcatagg gccaggactc	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca	7260
actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa	7320
aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt	7500
agagggagtt acttggccgg agctggactt ctcttttnta ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga	7740
ggctcagcaa aactgagatg gttcgttgag agaaacatgg tcacaccaga agggaaagta	7800
gtggaceteg gttgtggeag aggaggetgg teataetatt gtggaggaet aaagaatgta	7860
agagaagtca aaggeetaac aaaaggagga eeaggacaeg aagaaeeeat eeecatgtea	7920
acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa	8040
gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa	8100
ttttgcataa aggtteteaa eecatatatg eeeteagtea tagaaaaaat ggaageaeta	8160
caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag	8220
atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg	8280
atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac	8340
ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt	8400
gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac	8460
cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca	8520
tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg	8580
gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag	8640

		-concinued	
aaagtggaca cgagaaccca	agaaccgaaa gaaggcacga	agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga	attagggaag aaaaagacac	ccaggatgtg caccagagaa	8760
gaattcacaa gaaaggtgag	aagcaatgca gccttggggg	ccatattcac tgatgagaac	8820
aagtggaagt cggcacgtga	ggctgttgaa gatagtaggt	tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga	aggaaagtgt gaaacatgtg	tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga	attcggcaag gcaaaaggca	gcagagccat atggtacatg	9000
tggcttggag cacgcttctt	agagtttgaa gccctaggat	tcttaaatga agatcactgg	9060
ttctccagag agaactccct	gagtggagtg gaaggagaag	ggctgcacaa gctaggttac	9120
attctaagag acgtgagcaa	gaaagaggga ggagcaatgt	atgeegatga caeegeagga	9180
tgggatacaa gaatcacact	agaagaccta aaaaatgaag	aaatggtaac aaaccacatg	9240
gaaggagaac acaagaaact	agccgaggcc attttcaaac	taacgtacca aaacaaggtg	9300
gtgcgtgtgc aaagaccaac	accaagaggc acagtaatgg	acatcatatc gagaagagac	9360
caaagaggta gtggacaagt	tggcacctat ggactcaata	ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga	gggagaagga gtctttaaaa	gcattcagca cctaacaatc	9480
acagaagaaa tcgctgtgca	aaactggtta gcaagagtgg	ggcgcgaaag gttatcaaga	9540
atggccatca gtggagatga	ttgtgttgtg aaacctttag	atgacaggtt cgcaagcgct	9600
taacagctc taaatgacat	gggaaagatt aggaaagaca	tacaacaatg ggaaccttca	9660
agaggatgga atgattggac	acaagtgccc ttctgttcac	accatttcca tgagttaatc	9720
atgaaagacg gtcgcgtact	cgttgttccc tgtagaaacc	aagatgaact gattggcaga	9780
geeegaatet eecaaggage	agggtggtct ttgcgggaga	cggcctgttt ggggaagtct	9840
tacgcccaaa tgtggagctt	gatgtacttc cacagacgcg	acctcaggct ggcggcaaat	9900
gctatttgct cggcagtacc	atcacattgg gttccaacaa	gtcgaacaac ctggtccata	9960
catgctaaac atgaatggat	gacaacggaa gacatgctga	cagtctggaa cagggtgtgg	10020
attcaagaaa acccatggat	ggaagacaaa actccagtgg	aatcatggga ggaaatccca	10080
tacttgggga aaagagaaga	ccaatggtgc ggctcattga	ttgggttaac aagcagggcc	10140
acctgggcaa agaacatcca	agcagcaata aatcaagtta	gatcccttat aggcaatgaa	10200
gaatacacag attacatgcc	atccatgaaa agattcagaa	gagaagagga agaagcagga	10260
gttctgtggt agaaagcaaa	actaacatga aacaaggcta	gaagtcaggt cggattaagc	10320
catagtacgg aaaaaactat	gctacctgtg agccccgtcc	aaggacgtta aaagaagtca	10380
ggccatcata aatgccatag	cttgagtaaa ctatgcagcc	tgtagctcca cctgagaagg	10440
tgtaaaaaat ccgggaggcc	acaaaccatg gaagctgtac	gcatggcgta gtggactagc	10500
ggttagagga gacccctccc	ttacaaatcg cagcaacaat	ggggggcccaa ggcgagatga	10560
agctgtagtc tcgctggaag	gactagaggt tagaggagac	ccccccgaaa caaaaaacag	10620
catattgacg ctgggaaaga	ccagagatcc tgctgtctcc	tcagcatcat tccaggcaca	10680
gaacgccaga aaatggaatg	gtgctgttga atcaacaggt	tct	10723

<210> SEQ ID NO 11 <211> LENGTH: 3391 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE:

<223> OTHER INFORMATION: Dengue virus serotype 4, BVS <220> FEATURE: <221> NAME/KEY: Xaa <222> LOCATION: (1226)..(1226) <223> OTHER INFORMATION: Lys or Arg <220> FEATURE: <221> NAME/KEY: Xaa <222> LOCATION: (2481)..(2481) <223> OTHER INFORMATION: Ser or Phe <400> SEQUENCE: 11 Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly 50 55 60 Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val65707580 Leu Arg Gly Phe Arg Lys Glu Ile Gly Arg Met Leu Asn Ile Leu Asn Arg Arg Arg Ser Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Leu Met Ile Val Ala Lys His Glu Arg Gly Arg Pro Leu Leu Phe Lys Thr Thr Glu Gly Ile Asn Lys Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Glu Asp Thr Val Thr Tyr Lys Cys Pro Leu Leu Val Asn Thr Glu Pro Glu Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Met Tyr Gly Thr Cys Thr Gln Ser Gly Glu Arg Arg Arg Glu Lys Arg Ser Val Ala Leu Thr Pro His Ser Gly Met Gly Leu Glu Thr Arg Ala Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Ala Gln Arg Val Glu Ser Trp 225 230 235 240 Ile Leu Arg Asn Pro Gly Phe Ala Leu Leu Ala Gly Phe Met Ala Tyr Met Ile Gly Gln Thr Gly Ile Gln Arg Thr Val Phe Phe Val Leu Met Met Leu Val Ala Pro Ser Tyr Gly Met Arg Cys Val Gly Val Gly Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ala Trp Val Asp Leu Val Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Gln Gly Lys Pro Thr Leu Asp Phe Glu Leu Thr Lys Thr Thr Ala Lys Glu Val Ala Leu Leu Arg Thr Tyr Cys Ile Glu Ala Ser Ile Ser Asn Ile Thr Thr Ala Thr

-	cont	inu	ed

			340					345					350		
Arg	Сув	Pro 355	Thr	Gln	Gly	Glu	Pro 360	Tyr	Leu	Lys	Glu	Glu 365	Gln	Asp	Gln
Gln	Tyr 370	Ile	Суз	Arg	Arg	Asp 375	Val	Val	Asp	Arg	Gly 380	Trp	Gly	Asn	Gly
Cys 385	Gly	Leu	Phe	Gly	Lys 390	Gly	Gly	Val	Val	Thr 395	Суз	Ala	Lys	Phe	Ser 400
Суз	Ser	Gly	Lys	Ile 405	Thr	Gly	Asn	Leu	Val 410	Gln	Ile	Glu	Asn	Leu 415	Glu
Tyr	Thr	Val	Val 420	Val	Thr	Val	His	Asn 425	Gly	Asp	Thr	His	Ala 430	Val	Gly
Asn	Asp	Thr 435	Ser	Asn	His	Gly	Val 440	Thr	Ala	Thr	Ile	Thr 445	Pro	Arg	Ser
Pro	Ser 450	Val	Glu	Val	ГЛа	Leu 455	Pro	Asp	Tyr	Gly	Glu 460	Leu	Thr	Leu	Asp
Cys 465	Glu	Pro	Arg	Ser	Gly 470	Ile	Asp	Phe	Asn	Glu 475	Met	Ile	Leu	Met	Lуз 480
Met	Гла	Lys	ГЛа	Thr 485	Trp	Leu	Val	His	Lys 490	Gln	Trp	Phe	Leu	Asp 495	Leu
Pro	Leu	Pro	Trp 500	Thr	Ala	Gly	Ala	Asp 505	Thr	Ser	Glu	Val	His 510	Trp	Asn
Tyr	Lys	Glu 515	Arg	Met	Val	Thr	Phe 520	Lys	Val	Pro	His	Ala 525	Lys	Arg	Gln
Asp	Val 530	Thr	Val	Leu	Gly	Ser 535	Gln	Glu	Gly	Ala	Met 540	His	Ser	Ala	Leu
Ala 545	Gly	Ala	Thr	Glu	Val 550	Asp	Ser	Gly	Asp	Gly 555	Asn	His	Met	Phe	Ala 560
Gly	His	Leu	Lys	Cys 565	Lys	Val	Arg	Met	Glu 570	Lys	Leu	Arg	Ile	Lys 575	Gly
Met	Ser	Tyr	Thr 580	Met	Суз	Ser	Gly	Lys 585	Phe	Ser	Ile	Asp	Lys 590	Glu	Met
Ala	Glu	Thr 595	Gln	His	Gly	Thr	Thr 600	Val	Val	Lys	Val	Lys 605	Tyr	Glu	Gly
Ala	Gly 610	Ala	Pro	Суз	Lys	Val 615	Pro	Ile	Glu	Ile	Arg 620	Asp	Val	Asn	Lys
Glu 625	Lys	Val	Val	Gly	Arg 630	Ile	Ile	Ser	Ser	Thr 635	Pro	Leu	Ala	Glu	Asn 640
Thr	Asn	Ser	Val	Thr 645	Asn	Ile	Glu	Leu	Glu 650	Pro	Pro	Phe	Gly	Asp 655	Ser
Tyr	Ile	Val	Ile 660	Gly	Val	Gly	Asn	Ser 665	Ala	Leu	Thr	Leu	His 670	Trp	Phe
Arg	Lys	Gly 675	Ser	Ser	Ile	Gly	Lys 680	Met	Phe	Glu	Ser	Thr 685	Tyr	Arg	Gly
Ala	Lys 690	Arg	Met	Ala	Ile	Leu 695	Gly	Glu	Thr	Ala	Trp 700	Asp	Phe	Gly	Ser
Val 705	Gly	Gly	Leu	Phe	Thr 710	Ser	Leu	Gly	Lys	Ala 715	Val	His	Gln	Val	Phe 720
Gly	Ser	Val	Tyr	Thr 725	Thr	Leu	Phe	Gly	Gly 730	Val	Ser	Trp	Met	Ile 735	Arg
Ile	Leu	Ile	Gly 740	Phe	Leu	Val	Leu	Trp 745	Ile	Gly	Thr	Asn	Ser 750	Arg	Asn

Thr	Ser	Met 755	Ala	Met	Thr	Сүз	Ile 760	Ala	Ala	Gly	Ile	Val 765	Thr	Leu	Tyr
Leu	Gly 770	Val	Met	Val	Gln	Ala 775	Asp	Ser	Gly	Сүз	Val 780	Val	Ser	Trp	Lys
Asn 785	Lys	Glu	Leu	ГЛа	Cys 790	Gly	Ser	Gly	Ile	Phe 795	Ile	Thr	Asp	Asn	Val 800
His	Thr	Trp	Thr	Glu 805	Gln	Tyr	Гла	Phe	Gln 810	Pro	Glu	Ser	Pro	Ser 815	Lys
Leu	Ala	Ser	Ala 820	Ile	Gln	Lys	Ala	His 825	Glu	Glu	Asp	Ile	Суз 830	Gly	Ile
Arg	Ser	Val 835	Thr	Arg	Leu	Glu	Asn 840	Leu	Met	Trp	ГЛа	Gln 845	Ile	Thr	Pro
Glu	Leu 850	Asn	His	Ile	Leu	Ser 855	Glu	Asn	Glu	Val	Lүз 860	Leu	Thr	Ile	Met
Thr 865	Gly	Asp	Ile	ГЛа	Gly 870	Ile	Met	Gln	Ala	Gly 875	Lys	Arg	Ser	Leu	Arg 880
Pro	Gln	Pro	Thr	Glu 885	Leu	Lys	Tyr	Ser	Trp 890	Lys	Thr	Trp	Gly	Lys 895	Ala
Lys	Met	Leu	Ser 900	Thr	Glu	Ser	His	Asn 905	Gln	Thr	Phe	Leu	Ile 910	Asp	Gly
Pro	Glu	Thr 915	Ala	Glu	Сүз	Pro	Asn 920	Thr	Asn	Arg	Ala	Trp 925	Asn	Ser	Leu
Glu	Val 930	Glu	Asp	Tyr	Gly	Phe 935	Gly	Val	Phe	Thr	Thr 940	Asn	Ile	Trp	Leu
Lys 945	Leu	Lys	Glu	Lys	Gln 950	Asp	Val	Phe	Суз	Asp 955	Ser	Lys	Leu	Met	Ser 960
Ala	Ala	Ile	Lys	Asp 965	Asn	Arg	Ala	Val	His 970	Ala	Asp	Met	Gly	Tyr 975	Trp
Ile	Glu	Ser	Ala 980	Leu	Asn	Asp	Thr	Trp 985	Lys	Ile	Glu	Lys	Ala 990	Ser	Phe
Ile	Glu	Val 995	Lys	Asn	СЛа	His	Trp 1000		э Цу:	s Sei	r Hi	s Th: 10		eu T:	rp Ser
Asn	Gly 1010		l Leı	u Glı	ı Sei	c Glu 103		et I	le I	le Pi		ys 1 020	Asn I	Leu i	Ala
Gly	Pro 1025		l Se:	r Glı	n His	5 Ası 103		yr A:	rg Pi	ro G		yr 1 035	His 7	Thr (	Gln
Ile	Thr 1040		y Pro	o Trj	9 His	5 Lei 104	1 G. 15	ly Ly	ys L€	eu G	lu M 1	et 1 050	Asp I	Phe i	Aap
Phe	Cys 1059		o Gly	y Th:	r Thi	r Va 100		al Va	al Tł	nr G		ap ( 065	Cys (	Gly A	Asn
Arg	Gly 1070		o Se:	r Leı	ı Arç	g Th: 107		hr Tl	hr A	la Se		ly 1 080	Lys I	Leu :	Ile
Thr	Glu 1089	-	o Cyr	а Су:	s Arg	g Sei 109	-	ys Tl	hr Le	eu Pi		ro 1 095	Leu A	Arg	Fyr
Arg	Gly 1100		ı Asj	p Gly	y Cys	3 Tr] 110		yr G	ly Me	et Gi		le 1 110	Arg I	Pro 1	Leu
Lys	Glu 1119	-	s Glı	u Glı	ı Asr	n Leu 112		al A	sn Se	er Le		al ' 125	Thr A	Ala (	Gly
His	Gly 1130		n Vai	l Asj	o Asr	n Phe 113		er L	eu Gi	ly Vá		eu ( 140	Gly M	Met i	Ala

-continued
------------

											-00		iuet	л 
Leu	Phe 1145		Glu	Glu	Met	Leu 1150	-	Thr	Arg	Val	Gly 1155	Thr	Lys	His
Ala	Ile 1160		Leu	Val		Val 1165		Phe	Val	Thr	Leu 1170	Ile	Thr	Gly
	Met 1175		Phe	Arg	Asp	Leu 1180	-	-		Met	Val 1185	Met	Val	Gly
Ala	Thr 1190		Thr	Gly		Ile 1195					Thr 1200	-	Leu	Ala
Leu	Leu 1205					Val 1210				Phe	Ala 1215	Ala	Gly	Leu
Leu	Leu 1220	-	-		Thr	Ser 1225		Glu	Leu		Met 1230	Thr	Thr	Ile
-	Ile 1235		Leu	Leu		Gln 1240		Thr	Ile		Glu 1245	Thr	Ile	Leu
	Leu 1250		_			Ala 1255		-			Val 1260	Leu	Lys	Met
	Arg 1265					Tyr 1270		Leu	Ala	Val	Thr 1275	Ile	Met	Ala
	Leu 1280					Ala 1285		Ile	Leu	Gln	Asn 1290	Ala	Trp	ГЛа
	Ser 1295				Leu	Ala 1300					Ser 1305	Pro	Leu	Phe
	Thr 1310		Ser	Gln		Lys 1315					Pro 1320	Leu	Ala	Leu
	Ile 1325				Asn	Pro 1330	Thr	Ala	Ile	Phe	Leu 1335	Thr	Thr	Leu
Ser	Arg	Thr	Ser	Lys	Lys	Arg 1345		Trp		Leu	Asn 1350	Glu	Ala	Ile
	Ala 1355			Met		Ser 1360		Leu	Ala		Ser 1365	Leu	Leu	ГЛа
	Asp 1370	Ile			Thr	Gly 1375		Leu	Val		Gly 1380	Gly	Leu	Leu
					Leu	Thr 1390	Gly				Asp 1395	Leu	Glu	Leu
		Ala				Lys 1405	Trp					Glu	Ile	Ser
		Ser	Pro	Ile	Leu	Ser 1420	Ile	Thr	Ile	Ser		Asp	Gly	Ser
Met		Ile	Lys	Asn	Glu	Glu 1435	Glu	Glu	Gln	Thr		Thr	Ile	Leu
Ile		Thr	Gly	Leu	Leu	Val 1450	Ile	Ser	Gly	Leu		Pro	Val	Ser
Ile	Pro	Ile	Thr	Ala	Ala	Ala	Trp	Tyr	Leu	Trp	Glu	Val	Lys	Lys
	-		Gly	Val	Leu	1465 Trp	Asp	Val	Pro	Ser		Pro	Pro	Met
Gly	1475 Lys	Ala	Glu	Leu	Glu	1480 Asp		Ala	Tyr	Arg	1485 Ile	Lys	Gln	Lys
Glv	1490 Ile		Glv	Tvr	Ser	1495 Gln		Glv	Ala	Glv	1500 Val	Tvr	Lvs	G]11
-	1505		-	-		1510		-		-	1515	-	-	
сту	mr	гле	ніз	mr	Met	Trp	ніз	vai	mr	Arg	сту	АІА	vai	ьeu

-continued

_														
	1520					1525					1530			
Met	His 1535	Lys	Gly	Lys	Arg	Ile 1540	Glu	Pro	Ser	Trp	Ala 1545		Val	Lys
Lys	Asp 1550	Leu	Ile	Ser	Tyr	Gly 1555	Gly	Gly	Trp	Lys	Leu 1560		Gly	Glu
Trp	Lys 1565	Glu	Gly	Glu	Glu	Val 1570	Gln	Val	Leu	Ala	Leu 1575		Pro	Gly
Lys	Asn 1580	Pro	Arg	Ala	Val	Gln 1585	Thr	Lys	Pro	Gly	Leu 1590		ГЛа	Thr
Asn	Ala 1595	Gly	Thr	Ile	Gly	Ala 1600	Val	Ser	Leu	Asp	Phe 1605		Pro	Gly
Thr	Ser 1610	Gly	Ser	Pro	Ile	Ile 1615	Asp	Lys	Lys	Gly	Lys 1620		Val	Gly
Leu	Tyr 1625	Gly	Asn	Gly	Val	Val 1630	Thr	Arg	Ser	Gly	Ala 1635		Val	Ser
Ala	Ile 1640	Ala	Gln	Thr	Glu	Lys 1645	Ser	Ile	Glu	Asp	Asn 1650		Glu	Ile
Glu	Asp 1655	Asp	Ile	Phe	Arg	Lys 1660	Arg	Arg	Leu	Thr	Ile 1665		Asp	Leu
His	Pro 1670	Gly	Ala	Gly	Lys	Thr 1675	Lys	Arg	Tyr	Leu	Pro 1680		Ile	Val
Arg	Glu 1685	Ala	Ile	Lys	Arg	Gly 1690	Leu	Arg	Thr	Leu	Ile 1695		Ala	Pro
Thr	Arg 1700	Val	Val	Ala	Ala	Glu 1705	Met	Glu	Glu	Ala	Leu 1710		Gly	Leu
Pro	Ile 1715	Arg	Tyr	Gln	Thr	Pro 1720	Ala	Ile	Arg	Ala	Val 1725		Thr	Gly
Arg	Glu 1730	Ile	Val	Asp	Leu	Met 1735	Суз	His	Ala	Thr	Phe 1740		Met	Arg
Leu	Leu 1745	Ser	Pro	Val	Arg	Val 1750	Pro	Asn	Tyr	Asn	Leu 1755	Ile	Ile	Met
Asp	Glu 1760	Ala	His	Phe	Thr	Asp 1765	Pro	Ala	Ser	Ile	Ala 1770	Ala	Arg	Gly
Tyr	Ile 1775	Ser	Thr	Arg	Val	Glu 1780	Met	Gly	Glu	Ala	Ala 1785		Ile	Phe
Met	Thr 1790	Ala	Thr	Pro	Pro	Gly 1795	Ser	Arg	Asp	Pro	Phe 1800		Gln	Ser
Asn	Ala 1805	Pro	Ile	Ile	Asp	Glu 1810	Glu	Arg	Glu	Ile	Pro 1815		Arg	Ser
Trp	Asn 1820	Ser	Gly	His	Glu	Trp 1825	Val	Thr	Asp	Phe	Lуз 1830		ГЛЗ	Thr
Val	Trp 1835	Phe	Val	Pro	Ser	Ile 1840	Lys	Ala	Gly	Asn	Asp 1845	Ile	Ala	Ala
Суз	Leu 1850	Arg	Lys	Asn	Gly	Lys 1855	Lys	Val	Ile	Gln	Leu 1860		Arg	Lys
Thr	Phe 1865	Asp	Ser	Glu	Tyr	Val 1870	Lys	Thr	Arg	Thr	Asn 1875		Trp	Asp
Phe	Val 1880	Val	Thr	Thr	Asp	Ile 1885	Ser	Glu	Met	Gly	Ala 1890		Phe	Lys
Ala	Glu 1895	Arg	Val	Ile	Asp	Pro 1900	Arg	Arg	Суз	Met	Lys 1905		Val	Ile

-cont	inued	

Leu	Thr 1910	Asp	Gly	Glu	Glu	Arg 1915	Val	Ile	Leu	Ala	Gly 1920	Pro	Met	Pro	
Val	Thr 1925	His	Ser	Ser	Ala	Ala 1930	Gln	Arg	Arg	Gly	Arg 1935	Ile	Gly	Arg	
Asn	Pro 1940	Lys	Asn	Glu	Asn	Asp 1945	Gln	Tyr	Ile	Tyr	Met 1950	Gly	Glu	Pro	
Leu	Glu 1955	Asn	Asp	Glu	Asp	Cys 1960	Ala	His	Trp	Lys	Glu 1965	Ala	Lys	Met	
Leu	Leu 1970	Asp	Asn	Ile	Asn	Thr 1975	Pro	Glu	Gly	Ile	Ile 1980	Pro	Ser	Met	
Phe	Glu 1985	Pro	Glu	Arg	Glu	Lys 1990	Val	Asp	Ala	Ile	Asp 1995	Gly	Glu	Tyr	
Arg	Leu 2000	Arg	Gly	Glu	Ala	Arg 2005	Lys	Thr	Phe	Val	Asp 2010	Leu	Met	Arg	
Arg	Gly 2015	Asp	Leu	Pro	Val	Trp 2020	Leu	Ala	Tyr	Arg	Val 2025	Ala	Ala	Glu	
Gly	Ile 2030	Asn	Tyr	Ala	Asp	Arg 2035	Arg	Trp	Cys	Phe	Asp 2040	Gly	Val	ГЛа	
Asn	Asn 2045	Gln	Ile	Leu	Glu	Glu 2050	Asn	Val	Glu	Val	Glu 2055	Ile	Trp	Thr	
LÀa	Glu 2060	Gly	Glu	Arg	Lys	Lys 2065	Leu	ràa	Pro	Arg	Trp 2070	Leu	Asp	Ala	
Arg	Ile 2075	Tyr	Ser	Asp	Pro	Leu 2080	Ala	Leu	Lys	Glu	Phe 2085	Lys	Glu	Phe	
Ala	Ala 2090	Gly	Arg	Lys	Ser	Leu 2095	Thr	Leu	Asn	Leu	Ile 2100	Thr	Glu	Met	
Gly	Arg 2105	Leu	Pro	Thr	Phe	Met 2110	Thr	Gln	Lys	Val	Arg 2115	Asp	Ala	Leu	
Asp	Asn 2120	Leu	Ala	Val	Leu	His 2125	Thr	Ala	Glu	Ala	Gly 2130	Gly	Arg	Ala	
Tyr	Asn 2135	His	Ala	Leu	Ser	Glu 2140	Leu	Pro	Glu	Thr	Leu 2145	Glu	Thr	Leu	
Leu	Leu 2150	Leu	Thr	Leu	Leu	Ala 2155	Thr	Val	Thr	Gly	Gly 2160	Ile	Phe	Leu	
Phe	Leu 2165	Met	Ser	Ala	Arg	Gly 2170	Ile	Gly	Lys	Met	Thr 2175	Leu	Gly	Met	
Суз	Cys 2180	Ile	Ile	Thr	Ala	Ser 2185	Ile	Leu	Leu	Trp	Tyr 2190	Ala	Gln	Ile	
Gln	Pro 2195	His	Trp	Ile	Ala	Ala 2200	Ser	Ile	Ile	Leu	Glu 2205	Phe	Phe	Leu	
Ile	Val 2210	Leu	Leu	Ile	Pro	Glu 2215	Pro	Glu	ràa	Gln	Arg 2220	Thr	Pro	Gln	
Asp	Asn 2225	Gln	Leu	Thr	Tyr	Val 2230	Val	Ile	Ala	Ile	Leu 2235	Thr	Val	Val	
Ala	Ala 2240	Thr	Met	Ala	Asn	Glu 2245	Met	Gly	Phe	Leu	Glu 2250	Lys	Thr	Гла	
Lys	Asp 2255	Leu	Gly	Leu	Gly	Ser 2260	Ile	Ala	Thr	Gln	Gln 2265	Pro	Glu	Ser	
Asn	Ile 2270	Leu	Asp	Ile	Asp	Leu 2275	Arg	Pro	Ala	Ser	Ala 2280	Trp	Thr	Leu	

led
i u u

-continued

	2660					2665					2670			
Leu	Asn 2675	Pro	Tyr	Met	Pro	Ser 2680	Val	Ile	Glu	Lys	Met 2685	Glu	Ala	Leu
Gln	Arg 2690	Lys	Tyr	Gly	Gly	Ala 2695	Leu	Val	Arg	Asn	Pro 2700	Leu	Ser	Arg
Asn	Ser 2705	Thr	His	Glu	Met	Tyr 2710	Trp	Val	Ser	Asn	Ala 2715	Ser	Gly	Asn
Ile	Val 2720	Ser	Ser	Val	Asn	Met 2725	Ile	Ser	Arg	Met	Leu 2730	Ile	Asn	Arg
Phe	Thr 2735	Met	Arg	Tyr	Lys	Lys 2740	Ala	Thr	Tyr	Glu	Pro 2745	Asp	Val	Aap
Leu	Gly 2750	Ser	Gly	Thr	Arg	Asn 2755	Ile	Gly	Ile	Glu	Ser 2760	Glu	Ile	Pro
Asn	Leu 2765	Asp	Ile	Ile	Gly	Lys 2770	Arg	Ile	Glu	Lys	Ile 2775	Lys	Gln	Glu
His	Glu 2780	Thr	Ser	Trp	His	Tyr 2785	Aab	Gln	Asp	His	Pro 2790	Tyr	Lys	Thr
Trp	Ala 2795	Tyr	His	Gly	Ser	Tyr 2800	Glu	Thr	ГÀа	Gln	Thr 2805	Gly	Ser	Ala
Ser	Ser 2810	Met	Val	Asn	Gly	Val 2815	Val	Arg	Leu	Leu	Thr 2820	Lys	Pro	Trp
Asp	Val 2825	Val	Pro	Met	Val	Thr 2830	Gln	Met	Ala	Met	Thr 2835	Asp	Thr	Thr
Pro	Phe 2840	Gly	Gln	Gln	Arg	Val 2845	Phe	Lys	Glu	ГÀа	Val 2850	Asp	Thr	Arg
Thr	Gln 2855	Glu	Pro	Lys	Glu	Gly 2860	Thr	Lys	Гүз	Leu	Met 2865	Lys	Ile	Thr
Ala	Glu 2870	Trp	Leu	Trp	Lys	Glu 2875	Leu	Gly	Гүз	Lys	Lys 2880	Thr	Pro	Arg
Met	Cys 2885	Thr	Arg	Glu	Glu	Phe 2890	Thr	Arg	Lys	Val	Arg 2895	Ser	Asn	Ala
Ala	Leu 2900	Gly	Ala	Ile	Phe	Thr 2905	Asb	Glu	Asn	Lys	Trp 2910	Lys	Ser	Ala
Arg	Glu 2915	Ala	Val	Glu	Asp	Ser 2920	Arg	Phe	Trp	Glu	Leu 2925	Val	Asb	Lys
	2930					2935	-	-	-		Thr 2940	-		-
Asn	Met 2945	Met	Gly	Lys	Arg	Glu 2950	Lys	Lys	Leu	Gly	Glu 2955	Phe	Gly	Lys
Ala	Lys 2960	Gly	Ser	Arg	Ala	Ile 2965	Trp	Tyr	Met	Trp	Leu 2970	Gly	Ala	Arg
Phe	Leu 2975	Glu	Phe	Glu	Ala	Leu 2980	Gly	Phe	Leu	Asn	Glu 2985	Asp	His	Trp
Phe	Ser 2990	Arg	Glu	Asn	Ser	Leu 2995	Ser	Gly	Val	Glu	Gly 3000	Glu	Gly	Leu
His	Lys 3005	Leu	Gly	Tyr	Ile	Leu 3010	-	Asp	Val	Ser	Lys 3015	Lys	Glu	Gly
Gly	Ala 3020	Met	Tyr	Ala	Asp	Asp 3025	Thr	Ala	Gly	Trp	Aap 3030	Thr	Arg	Ile
Thr	Leu 3035	Glu	Asp	Leu	Lys	Asn 3040	Glu	Glu	Met	Val	Thr 3045	Asn	His	Met

Glu	Gly 3050	Glu	His	Lys	Lys	Leu 3055	Ala	Glu	Ala	Ile	Phe 3060		Leu	Thr
Tyr	Gln 3065	Asn	ГЛЗ	Val	Val	Arg 3070		Gln	Arg	Pro	Thr 3075	Pro	Arg	Gly
Thr	Val 3080	Met	Asp	Ile	Ile	Ser 3085	Arg	Arg	Asp	Gln	Arg 3090	Gly	Ser	Gly
Gln	Val 3095	Gly	Thr	Tyr	Gly	Leu 3100	Asn	Thr	Phe	Thr	Asn 3105	Met	Glu	Ala
Gln	Leu 3110	Ile	Arg	Gln	Met	Glu 3115	Gly	Glu	Gly	Val	Phe 3120	Lys	Ser	Ile
Gln	His 3125	Leu	Thr	Ile	Thr	Glu 3130		Ile	Ala	Val	Gln 3135	Asn	Trp	Leu
Ala	Arg 3140	Val	Gly	Arg	Glu	Arg 3145	Leu	Ser	Arg	Met	Ala 3150	Ile	Ser	Gly
Asp	Asp 3155	Сүз	Val	Val	Lys	Pro 3160	Leu	Aab	Asp	Arg	Phe 3165	Ala	Ser	Ala
Leu	Thr 3170	Ala	Leu	Asn	Asp	Met 3175	Gly	Lys	Ile	Arg	Lys 3180	Asp	Ile	Gln
Gln	Trp 3185	Glu	Pro	Ser	Arg	Gly 3190		Asn	Asp	Trp	Thr 3195	Gln	Val	Pro
Phe	Cys 3200	Ser	His	His	Phe	His 3205	Glu	Leu	Ile	Met	Lys 3210	Asp	Gly	Arg
Val	Leu 3215	Val	Val	Pro	Суз	Arg 3220	Asn	Gln	Asp	Glu	Leu 3225	Ile	Gly	Arg
Ala	Arg 3230	Ile	Ser	Gln	Gly	Ala 3235	Gly	Trp	Ser	Leu	Arg 3240	Glu	Thr	Ala
Сүз	Leu 3245	Gly	Lys	Ser	Tyr	Ala 3250	Gln	Met	Trp	Ser	Leu 3255	Met	Tyr	Phe
His	Arg 3260	Arg	Asp	Leu	Arg	Leu 3265	Ala	Ala	Asn	Ala	Ile 3270	Суз	Ser	Ala
Val	Pro 3275	Ser	His	Trp	Val	Pro 3280	Thr	Ser	Arg	Thr	Thr 3285	Trp	Ser	Ile
His	Ala 3290	LYa	His	Glu	Trp	Met 3295	Thr	Thr	Glu	Aap	Met 3300	Leu	Thr	Val
Trp	Asn 3305	Arg	Val	Trp	Ile	Gln 3310	Glu	Asn	Pro	Trp	Met 3315	Glu	Asp	Lys
Thr	Pro 3320	Val	Glu	Ser	Trp	Glu 3325	Glu	Ile	Pro	Tyr	Leu 3330	Gly	Lys	Arg
Glu	Asp 3335	Gln	Trp	Суз	Gly	Ser 3340	Leu	Ile	Gly	Leu	Thr 3345	Ser	Arg	Ala
Thr	Trp 3350	Ala	Lys	Asn	Ile	Gln 3355	Ala	Ala	Ile	Asn	Gln 3360	Val	Arg	Ser
Leu	Ile 3365	Gly	Asn	Glu	Glu	Tyr 3370		Asp	Tyr	Met	Pro 3375	Ser	Met	Lys
Arg	Phe 3380	Arg	Arg	Glu	Glu	Glu 3385	Glu	Ala	Gly	Val	Leu 3390	Trp		
	)> SE( L> LEI													

<210> SEQ ID NO 12 <211> LENGTH: 3391 <212> TYPE: PRT <213> ORGANISM: Artificial Sequence <220> FEATURE:

-continu	ed
----------	----

<223> OTHER INFORMATION: Dengue virus serotype 4, MVS <220> FEATURE: <221> NAME/KEY: Xaa <222> LOCATION: (1226)..(1226) <223> OTHER INFORMATION: Lys or Arg <400> SEQUENCE: 12 Met Asn Asn Gln Arg Lys Lys Ala Lys Asn Thr Pro Phe Asn Met Leu Lys Arg Glu Arg Asn Arg Val Ser Thr Val Gln Gln Leu Thr Lys Arg Phe Ser Leu Gly Met Leu Gln Gly Arg Gly Pro Leu Lys Leu Phe Met Ala Leu Val Ala Phe Leu Arg Phe Leu Thr Ile Pro Pro Thr Ala Gly Ile Leu Lys Arg Trp Gly Thr Ile Lys Lys Ser Lys Ala Ile Asn Val65707580 Leu  $\operatorname{Arg}$  Gly Phe  $\operatorname{Arg}$  Lys Glu Ile Gly  $\operatorname{Arg}$  Met Leu  $\operatorname{Asn}$  Ile Leu  $\operatorname{Asn}$ Arg Arg Arg Ser Ser Ala Gly Met Ile Ile Met Leu Ile Pro Thr Val Met Ala Phe His Leu Thr Thr Arg Asp Gly Glu Pro Leu Met Ile Val Ala Lys His Glu Arg Gly Arg Pro Leu Leu Phe Lys Thr Thr Glu Gly Ile Asn Lys Cys Thr Leu Ile Ala Met Asp Leu Gly Glu Met Cys Glu Asp Thr Val Thr Tyr Lys Cys Pro Leu Leu Val As<br/>n Thr Glu Pro Glu $\ensuremath{\mathsf{Pro}}$ Asp Ile Asp Cys Trp Cys Asn Leu Thr Ser Thr Trp Val Met Tyr Gly Thr Cys Thr Gln Ser Gly Glu Arg Arg Arg Glu Lys Arg Ser Val Ala Leu Thr Pro His Ser Gly Met Gly Leu Glu Thr Arg Ala Glu Thr Trp Met Ser Ser Glu Gly Ala Trp Lys His Ala Gln Arg Val Glu Ser Trp Ile Leu Arg Asn Pro Gly Phe Ala Leu Leu Ala Gly Phe Met Ala Tyr Met Ile Gly Gln Thr Gly Ile Gln Arg Thr Val Phe Phe Val Leu Met Met Leu Val Ala Pro Ser Tyr Gly Met Arg Cys Val Gly Val Gly Asn Arg Asp Phe Val Glu Gly Val Ser Gly Gly Ala Trp Val Asp Leu Val Leu Glu His Gly Gly Cys Val Thr Thr Met Ala Gln Gly Lys Pro Thr Leu Asp Phe Glu Leu Thr Lys Thr Thr Ala Lys Glu Val Ala Leu Leu Arg Thr Tyr Cys Ile Glu Ala Ser Ile Ser Asn Ile Thr Thr Ala Thr Arg Cys Pro Thr Gln Gly Glu Pro Tyr Leu Lys Glu Glu Gln Asp Gln 

Gln Tyr Ile Cys Arg Arg Asp Val Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Gly Val Val Thr Cys Ala Lys Phe Ser Cys Ser Gly Lys Ile Thr Gly Asn Leu Val Gln Ile Glu Asn Leu Glu Tyr Thr Val Val Val Thr Val His Asn Gly Asp Thr His Ala Val Gly Asn Asp Thr Ser Asn His Gly Val Thr Ala Thr Ile Thr Pro Arg Ser Pro Ser Val Glu Val Lys Leu Pro Asp Tyr Gly Glu Leu Thr Leu Asp Cys Glu Pro Arg Ser Gly Ile Asp Phe Asn Glu Met Ile Leu Met Lys Met Lys Lys Thr Trp Leu Val His Lys Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Thr Ala Gly Ala Asp Thr Ser Glu Val His Trp Asn Tyr Lys Glu Arg Met Val Thr Phe Lys Val Pro His Ala Lys Arg Gln Asp Val Thr Val Leu Gly Ser Gln Glu Gly Ala Met His Ser Ala Leu Ala Gly Ala Thr Glu Val Asp Ser Gly Asp Gly Asn His Met Phe Ala Gly His Leu Lys Cys Lys Val Arg Met Glu Lys Leu Arg Ile Lys Gly Met Ser Tyr Thr Met Cys Ser Gly Lys Phe Ser Ile Asp Lys Glu Met Ala Glu Thr Gln His Gly Thr Thr Val Val Lys Val Lys Tyr Glu Gly Ala Gly Ala Pro Cys Lys Val Pro Ile Glu Ile Arg Asp Val Asn Lys Glu Lys Val Val Gly Arg Ile Ile Ser Ser Thr Pro Leu Ala Glu Asn Thr Asn Ser Val Thr Asn Ile Glu Leu Glu Pro Pro Phe Gly Asp Ser Tyr Ile Val Ile Gly Val Gly Asn Ser Ala Leu Thr Leu His Trp Phe Arg Lys Gly Ser Ser Ile Gly Lys Met Phe Glu Ser Thr Tyr Arg Gly Ala Lys Arg Met Ala Ile Leu Gly Glu Thr Ala Trp Asp Phe Gly Ser Val Gly Gly Leu Phe Thr Ser Leu Gly Lys Ala Val His Gln Val Phe Gly Ser Val Tyr Thr Thr Leu Phe Gly Gly Val Ser Trp Met Ile Arg Ile Leu Ile Gly Phe Leu Val Leu  $\operatorname{Trp}$  Ile Gly Thr As<br/>n Ser Arg Asn Thr Ser Met Ala Met Thr Cys Ile Ala Ala Gly Ile Val Thr Leu Tyr 

_															.c mueu
ı	Gly 770	Val	Met	Val	Gln	Ala 775	Aap	Ser	Gly	сув	Val 780		Ser	Trp	) L
Asn 785	Lys	Glu	Leu	LÀa	Cys 790	-	Ser	Gly	Ile	Phe 795		Thr	Asp	) Asn	1 \ 8
His	Thr	Trp	Thr	Glu 805	Gln	Tyr	Lys	Phe	Glr. 810	ı Pro	Glu	Ser	Pro	Ser 815	-
Leu	Ala	Ser	Ala 820		Gln	Lys	Ala	His 825		ı Glu	Asp	Ile	Суа 830	-	/ Ile
Arg	Ser	Val 835	Thr	Arg	Leu	Glu	Asn 840		. Met	Trp	Lys	Gln 845		e Thr	r Pro
	Leu 850	Asn	His	Ile	Leu	Ser 855	Glu	Asn	Glu	ı Val	Lys 860		Thr	: Ile	e Met
Thr 865	Gly	Asp	Ile	Lys	Gly 870		Met	Gln	. Ala	Gly 875		Arg	Ser	: Leu	1 Arg 880
Pro	Gln	Pro	Thr	Glu 885	Leu	Lys	Tyr	Ser	Trp 890		Thr	Trp	Glγ	7 Lys 895	3 Ala 5
Lys	Met	Leu	Ser 900		Glu	Ser	His	Asn 905		1 Thr	Phe	Leu	11€ 910		Gly
Pro	Glu	Thr 915	Ala	Glu	Сүз	Pro	Asn 920		Asr	ı Arg	Ala	Trp 925		n Ser	r Leu
	Val 930	Glu	Asp	Tyr	Gly	Phe 935	Gly	Val	Phe	• Thr	Thr 940		Ile	e Trp	) Leu
Lys 945	Leu	Lys	Glu	Lys	Gln 950		Val	Phe	суз	Asp 955		Lys	Leu	ı Met	Ser 960
Ala	Ala	Ile	Lys	Asp 965		Arg	Ala	Val	His 970		Asp	Met	Gly	7 Tyr 975	r Trp
Ile	Glu	Ser	Ala 980	Leu	Asn	Asp	Thr	Trp 985		Ile	Glu	Lys	Al <i>a</i> 990		: Phe
Ile	Glu	Val 995	Lys	Asn	Суз	His	Trp 100		ю Бу	rs Se	r Hi	s Th 10		eu I	rp Se
	Gly 1010		l Le	u Glı	u Se:	r Gl 10		et I	le I	le P		уз 020	Asn	Leu	Ala
	Pro 1025		l Se:	r Gli	n Hi:	s As: 10		yr A	rg F	ro G		yr 035	His	Thr	Gln
	Thr 1040	-	y Pro	o Trj	p Hi	s Le 10		ly L	ys I	eu G		et 050	Asp	Phe	Asp
	Cys 1055		o Gl	y Th:	r Th:	r Va 10		al V	al T	'hr G		sp 065	Суз	Gly	Asn
	Gly 1070		Se:	r Le	u Arg	g Th: 10		hr T	hr A	la S		ly 080	Lys	Leu	Ile
	Glu 1085	-	o Cy	s Cy:	s Arg	g Se: 10:		уз Т	hr L	eu P		ro 095	Leu	Arg	Tyr
	Gly 1100		ı Asj	p Gl	у Су	s Trj 11	-	yr G	ly M	let G		le 110	Arg	Pro	Leu
-	Glu 1115	-	s Gl	u Glı	u Ası	n Le: 11:		al A	sn S	er L		al 125	Thr	Ala	Gly
	Gly 1130		n Va	l Asj	p Ası	n Ph		er L	eu G	sly V		eu 140	Gly	Met	Ala
Leu		Leu	ı Glı	u Glı	u Met		u A	rg T	hr A	rg V	al G		Thr	Lys	His
			ı Le	u Va	l Ala			er P	he V	'al T			Ile	Thr	Gly

-continued

		1160					1165					1170			
	Asn	Met 1175	Ser	Phe	Arg	Asp	Leu 1180	Gly	Arg	Val	Met	Val 1185	Met	Val	Gly
1	Ala	Thr 1190	Met	Thr	Gly	Asp	Ile 1195	Gly	Met	Gly	Val	Thr 1200		Leu	Ala
	Leu	Leu 1205	Ala	Ala	Phe	Lys	Val 1210		Pro	Thr	Phe	Ala 1215	Ala	Gly	Leu
	Leu	Leu 1220	Arg	ГÀа	Leu	Thr	Ser 1225	Хаа	Glu	Leu	Met	Met 1230	Thr	Thr	Ile
(	Gly	Ile 1235	Val	Leu	Leu	Ser	Gln 1240	Ser	Thr	Ile	Pro	Glu 1245	Thr	Ile	Leu
(	Glu	Leu 1250	Thr	Asp	Ala	Leu	Ala 1255	Leu	Gly	Met	Met	Val 1260	Leu	Lys	Met
1	Val	Arg 1265	Asn	Met	Glu	Lys	Tyr 1270	Gln	Leu	Ala	Val	Thr 1275	Ile	Met	Ala
	Ile	Leu 1280	Суз	Val	Pro	Asn	Ala 1285	Val	Ile	Leu	Gln	Asn 1290	Ala	Trp	Гла
	Val	Ser 1295	Сүз	Thr	Ile	Leu	Ala 1300	Val	Val	Ser	Val	Ser 1305	Pro	Leu	Phe
	Leu	Thr 1310	Ser	Ser	Gln	Gln	Lys 1315	Thr	Aab	Trp	Ile	Pro 1320	Leu	Ala	Leu
	Thr	Ile 1325	Γλa	Gly	Leu	Asn	Pro 1330	Thr	Ala	Ile	Phe	Leu 1335	Thr	Thr	Leu
		1340			-	-	1345		-			Asn 1350			
		1355		-			1360					Ser 1365			-
		1370					1375					Gly 1380			
		1385					1390					Asp 1395			
		1400			-		1405	-		-		Ala 1410			
(	Gly	Ser 1415					1420					Glu 1425	_	-	
		Ser 1430					1435					Leu 1440			
		1445		-			1450			_		Phe 1455			
	Ile	Pro 1460	Ile	Thr	Ala	Ala	Ala 1465	Trp	Tyr	Leu	Trp	Glu 1470	Val	Lys	Гла
(	Gln	Arg 1475	Ala	Gly	Val	Leu	Trp 1480	Asp	Val	Pro	Ser	Pro 1485	Pro	Pro	Met
(	Gly	Lys 1490	Ala	Glu	Leu	Glu	Asp 1495	Gly	Ala	Tyr	Arg	Ile 1500	-	Gln	Гла
(	Gly	Ile 1505	Leu	Gly	Tyr	Ser	Gln 1510	Ile	Gly	Ala	Gly	Val 1515	_	Lys	Glu
(	Gly	Thr 1520	Phe	His	Thr	Met	Trp 1525	His	Val	Thr	Arg	Gly 1530	Ala	Val	Leu
J	Met	His 1535	Lys	Gly	Lys	Arg	Ile 1540	Glu	Pro	Ser	Trp	Ala 1545	Asp	Val	Lys

-continued	

Lys	Asp 1550		Ile	Ser	Tyr	Gly 1555		Gly	Trp	Lys	Leu 1560		Gly	Glu
Trp	Lys 1565	Glu	Gly	Glu	Glu	Val 1570		Val	Leu	Ala	Leu 1575	Glu	Pro	Gly
Lys	Asn 1580	Pro	Arg	Ala	Val	Gln 1585		Lys	Pro	Gly	Leu 1590	Phe	Lys	Thr
Asn	Ala 1595	Gly	Thr	Ile	Gly	Ala 1600		Ser	Leu	Asp	Phe 1605	Ser	Pro	Gly
Thr	Ser 1610	Gly	Ser	Pro	Ile	Ile 1615	Aab	Lys	Lys	Gly	Lys 1620	Val	Val	Gly
Leu	Tyr 1625	Gly	Asn	Gly	Val	Val 1630		Arg	Ser	Gly	Ala 1635	Tyr	Val	Ser
Ala	Ile 1640	Ala	Gln	Thr	Glu	Lys 1645		Ile	Glu	Asp	Asn 1650	Pro	Glu	Ile
Glu	Asp 1655	Asp	Ile	Phe	Arg	Lys 1660		Arg	Leu	Thr	Ile 1665	Met	Asp	Leu
His	Pro 1670	Gly	Ala	Gly	Lys	Thr 1675	Lys	Arg	Tyr	Leu	Pro 1680	Ala	Ile	Val
Arg	Glu 1685	Ala	Ile	Lys	Arg	Gly 1690		Arg	Thr	Leu	Ile 1695	Leu	Ala	Pro
Thr	Arg 1700	Val	Val	Ala	Ala	Glu 1705	Met	Glu	Glu	Ala	Leu 1710	Arg	Gly	Leu
Pro	Ile 1715	Arg	Tyr	Gln	Thr	Pro 1720	Ala	Ile	Arg	Ala	Val 1725	His	Thr	Gly
Arg	Glu 1730	Ile	Val	Asp	Leu	Met 1735	Суз	His	Ala	Thr	Phe 1740	Thr	Met	Arg
Leu	Leu 1745	Ser	Pro	Val	Arg	Val 1750	Pro	Asn	Tyr	Asn	Leu 1755	Ile	Ile	Met
Asp	Glu 1760	Ala	His	Phe	Thr	Asp 1765	Pro	Ala	Ser	Ile	Ala 1770	Ala	Arg	Gly
Tyr	Ile 1775	Ser	Thr	Arg	Val	Glu 1780	Met	Gly	Glu	Ala	Ala 1785	Gly	Ile	Phe
Met	Thr 1790	Ala	Thr	Pro	Pro	Gly 1795	Ser	Arg	Asp	Pro	Phe 1800	Pro	Gln	Ser
Asn	Ala 1805	Pro	Ile	Ile	Asp	Glu 1810	Glu	Arg	Glu	Ile	Pro 1815	Glu	Arg	Ser
Trp	Asn 1820	Ser	Gly	His	Glu	Trp 1825	Val	Thr	Asp	Phe	Lys 1830	Gly	Lys	Thr
Val	Trp 1835	Phe	Val	Pro	Ser	Ile 1840	Lys	Ala	Gly	Asn	Asp 1845	Ile	Ala	Ala
Суз	Leu 1850	Arg	Lys	Asn	Gly	Lys 1855	Lys	Val	Ile	Gln	Leu 1860	Ser	Arg	Lys
Thr	Phe 1865	Asp	Ser	Glu	Tyr	Val 1870	-	Thr	Arg	Thr	Asn 1875	Asp	Trp	Aap
Phe	Val 1880	Val	Thr	Thr	Asp	Ile 1885		Glu	Met	Gly	Ala 1890	Asn	Phe	ГЛа
Ala	Glu 1895	Arg	Val	Ile	Asp	Pro 1900	Arg	Arg	Суз	Met	Lys 1905	Pro	Val	Ile
Leu	Thr 1910	Asp	Gly	Glu	Glu	Arg 1915	Val	Ile	Leu	Ala	Gly 1920	Pro	Met	Pro

													iuct	
Val	Thr 1925		Ser	Ser	Ala	Ala 1930		Arg	Arg	Gly	Arg 1935		Gly	Arg
Asn	Pro 1940	-	Asn	Glu	Asn	Asp 1945		Tyr	Ile	Tyr	Met 1950	Gly	Glu	Pro
Leu	Glu 1955		Asp	Glu	Asp	Cys 1960		His	Trp	Lys	Glu 1965	Ala	Lys	Met
Leu	Leu 1970	-	Asn	Ile	Asn	Thr 1975		Glu	Gly	Ile	Ile 1980	Pro	Ser	Met
Phe	Glu 1985		Glu	Arg	Glu	Lys 1990		Asp	Ala	Ile	Asp 1995	Gly	Glu	Tyr
Arg	Leu 2000	-	Gly	Glu	Ala	Arg 2005	-	Thr	Phe	Val	Asp 2010	Leu	Met	Arg
Arg	Gly 2015	-	Leu	Pro	Val	Trp 2020			Tyr	-	Val 2025	Ala	Ala	Glu
Gly	Ile 2030		Tyr	Ala	Asp	Arg 2035		Trp	Суа	Phe	Asp 2040	Gly	Val	Lys
Asn	Asn 2045		Ile	Leu	Glu	Glu 2050		Val	Glu	Val	Glu 2055	Ile	Trp	Thr
ГЛЗ	Glu 2060	-	Glu	Arg	Lys	Lys 2065		Lys	Pro	Arg	Trp 2070	Leu	Asp	Ala
Arg	Ile 2075			Asp		Leu 2080			Lys	Glu	Phe 2085	Lys	Glu	Phe
Ala	Ala 2090			Lys		Leu 2095		Leu	Asn		Ile 2100	Thr	Glu	Met
Gly	Arg 2105		Pro	Thr	Phe	Met 2110			Lys		Arg 2115		Ala	Leu
Asp	Asn 2120		Ala	Val	Leu	His 2125		Ala	Glu	Ala	Gly 2130	Gly	Arg	Ala
Tyr	Asn 2135		Ala	Leu	Ser	Glu 2140		Pro	Glu	Thr	Leu 2145	Glu	Thr	Leu
Leu	Leu 2150		Thr	Leu	Leu	Ala 2155					Gly 2160	Ile	Phe	Leu
Phe	Leu 2165		Ser	Ala	Arg						Thr 2175	Leu	Gly	Met
Суз	Cys 2180		Ile	Thr	Ala	Ser 2185					Tyr 2190	Ala	Gln	Ile
Gln	Pro 2195	His	Trp	Ile	Ala	Ala 2200	Ser	Ile	Ile	Leu	Glu 2205	Phe	Phe	Leu
Ile	Val 2210		Leu	Ile	Pro	Glu 2215		Glu	Lys	Gln	Arg 2220	Thr	Pro	Gln
Aap	Asn 2225	Gln	Leu	Thr	Tyr	Val 2230		Ile	Ala	Ile	Leu 2235	Thr	Val	Val
Ala	Ala 2240	Thr	Met	Ala	Asn	Glu 2245		Gly	Phe	Leu	Glu 2250	Lys	Thr	Lys
ГЛа	Asp 2255	Leu	Gly	Leu	Gly	Ser 2260		Ala	Thr	Gln	Gln 2265	Pro	Glu	Ser
Asn		Leu	Asp	Ile	Asp		Arg	Pro	Ala	Ser	Ala 2280	Trp	Thr	Leu
Tyr	Ala	Val	Ala	Thr	Thr	Phe	Val	Thr	Pro	Met	Leu	Arg	His	Ser
Ile	2285 Glu	Asn	Ser	Ser	Val	2290 Asn		Ser	Leu	Thr	2295 Ala	Ile	Ala	Asn

-continued

	2300					2305					2310			
Gln	Ala 2315	Thr	Val	Leu	Met	Gly 2320	Leu	Gly	Lys	Gly	Trp 2325	Pro	Leu	Ser
Lys	Met 2330	Asp	Ile	Gly	Val	Pro 2335		Leu	Ala	Ile	Gly 2340		Tyr	Ser
Gln	Val 2345	Asn	Pro	Ile	Thr	Leu 2350		Ala	Ala	Leu	Phe 2355		Leu	Val
Ala	His 2360		Ala	Ile	Ile	Gly 2365	Pro	Gly	Leu	Gln	Ala 2370		Ala	Thr
Arg	Glu 2375	Ala	Gln	Lys	Arg	Ala 2380	Ala	Ala	Gly	Ile	Met 2385		Asn	Pro
Thr	Val 2390	Asp	Gly	Ile	Thr	Val 2395		Asp	Leu	Asp	Pro 2400		Pro	Tyr
Asp	Pro 2405	ГЛа	Phe	Glu	ГЛа	Gln 2410	Leu	Gly	Gln	Val	Met 2415		Leu	Val
Leu	Сув 2420	Val	Thr	Gln	Val	Leu 2425	Met	Met	Arg	Thr	Thr 2430		Ala	Leu
Сув	Glu 2435	Ala	Leu	Thr	Leu	Ala 2440	Thr	Gly	Pro	Ile	Ser 2445		Leu	Trp
Glu	Gly 2450	Asn	Pro	Gly	Arg	Phe 2455	Trp	Asn	Thr	Thr	Ile 2460	Ala	Val	Ser
Met	Ala 2465	Asn	Ile	Phe	Arg	Gly 2470	Ser	Tyr	Leu	Ala	Gly 2475	Ala	Gly	Leu
Leu	Phe 2480	Ser	Ile	Met	Lys	Asn 2485	Thr	Thr	Asn	Thr	Arg 2490	Arg	Gly	Thr
Gly	Asn 2495	Ile	Gly	Glu	Thr	Leu 2500	Gly	Glu	Lys	Trp	Lys 2505		Arg	Leu
Asn	Ala 2510	Leu	Gly	Lys	Ser	Glu 2515	Phe	Gln	Ile	Tyr	Lys 2520		Ser	Gly
Ile	Gln 2525	Glu	Val	Asp	Arg	Thr 2530	Leu	Ala	Lys	Glu	Gly 2535	Ile	Lys	Arg
Gly	Glu 2540	Thr	Asp	His	His	Ala 2545	Val	Ser	Arg	Gly	Ser 2550		Lys	Leu
Arg	Trp 2555	Phe	Val	Glu	Arg	Asn 2560	Met	Val	Thr	Pro	Glu 2565	Gly	Lys	Val
Val	Asp 2570	Leu	Gly	Сүз	Gly	Arg 2575		Gly	Trp	Ser	Tyr 2580		СЛа	Gly
Gly	Leu 2585	Lys	Asn	Val	Arg	Glu 2590	Val	Lys	Gly	Leu	Thr 2595		Gly	Gly
Pro	Gly 2600	His	Glu	Glu	Pro	Ile 2605	Pro	Met	Ser	Thr	Tyr 2610	_	Trp	Asn
Leu	Val 2615	Arg	Leu	Gln	Ser	Gly 2620	Val	Asb	Val	Phe	Phe 2625	Ile	Pro	Pro
Glu	Lys 2630		Asp	Thr	Leu	Leu 2635		Asb	Ile	Gly	Glu 2640		Ser	Pro
Asn	Pro 2645	Thr	Val	Glu	Ala	Gly 2650	Arg	Thr	Leu	Arg	Val 2655		Asn	Leu
Val	Glu 2660	Asn	Trp	Leu	Asn	Asn 2665	Asn	Thr	Gln	Phe	Cys 2670	Ile	Lys	Val
Leu	Asn 2675	Pro	Tyr	Met	Pro	Ser 2680	Val	Ile	Glu	Lys	Met 2685	Glu	Ala	Leu

Gln	Arg 2690	Lys	Tyr	Gly	Gly	Ala 2695	Leu	Val	Arg	Asn	Pro 2700	Leu	Ser	Arg
Asn	Ser 2705	Thr	His	Glu	Met	Tyr 2710	Trp	Val	Ser	Asn	Ala 2715	Ser	Gly	Asn
Ile	Val 2720	Ser	Ser	Val	Asn	Met 2725	Ile	Ser	Arg	Met	Leu 2730	Ile	Asn	Arg
Phe	Thr 2735	Met	Arg	Tyr	Lys	Lys 2740	Ala	Thr	Tyr	Glu	Pro 2745	Asp	Val	Asp
Leu	Gly 2750	Ser	Gly	Thr	Arg	Asn 2755	Ile	Gly	Ile	Glu	Ser 2760	Glu	Ile	Pro
Asn	Leu 2765	Asp	Ile	Ile	Gly	Lys 2770	Arg	Ile	Glu	Lys	Ile 2775	Lys	Gln	Glu
His	Glu 2780	Thr	Ser	Trp	His	Tyr 2785	Asp	Gln	Asp	His	Pro 2790	Tyr	Lys	Thr
Trp	Ala 2795	Tyr	His	Gly	Ser	Tyr 2800	Glu	Thr	Lys	Gln	Thr 2805	Gly	Ser	Ala
Ser	Ser 2810	Met	Val	Asn	Gly	Val 2815	Val	Arg	Leu	Leu	Thr 2820	Lys	Pro	Trp
Asp	Val 2825	Val	Pro	Met	Val	Thr 2830	Gln	Met	Ala	Met	Thr 2835	Asp	Thr	Thr
Pro	Phe 2840	Gly	Gln	Gln	Arg	Val 2845	Phe	Lys	Glu	Lys	Val 2850	Asp	Thr	Arg
Thr	Gln 2855	Glu	Pro	Lys	Glu	Gly 2860	Thr	Lys	Lys	Leu	Met 2865	Lys	Ile	Thr
Ala	Glu 2870	Trp	Leu	Trp	Lys	Glu 2875	Leu	Gly	Lys	Lys	Lys 2880	Thr	Pro	Arg
Met	Cys 2885	Thr	Arg	Glu	Glu	Phe 2890	Thr	Arg	Lys	Val	Arg 2895	Ser	Asn	Ala
Ala	Leu 2900	Gly	Ala	Ile	Phe	Thr 2905	Aab	Glu	Asn	Lys	Trp 2910	Lys	Ser	Ala
Arg	Glu 2915	Ala	Val	Glu	Asp	Ser 2920	Arg	Phe	Trp	Glu	Leu 2925	Val	Asp	Lya
Glu	Arg 2930	Asn	Leu	His	Leu	Glu 2935	Gly	ГЛа	Суа	Glu	Thr 2940	Суз	Val	Tyr
Asn	Met 2945	Met	Gly	Lys	Arg	Glu 2950	Lys	Lys	Leu	Gly	Glu 2955	Phe	Gly	Lya
Ala	Lys 2960	Gly	Ser	Arg	Ala	Ile 2965	Trp	Tyr	Met	Trp	Leu 2970	Gly	Ala	Arg
Phe	Leu 2975	Glu	Phe	Glu	Ala	Leu 2980	Gly	Phe	Leu	Asn	Glu 2985	Asp	His	Trp
Phe	Ser 2990	Arg	Glu	Asn	Ser	Leu 2995	Ser	Gly	Val	Glu	Gly 3000	Glu	Gly	Leu
His	Lуя 3005	Leu	Gly	Tyr	Ile	Leu 3010	Arg	Aap	Val	Ser	Lys 3015	Lys	Glu	Gly
Gly	Ala 3020	Met	Tyr	Ala	Asp	Asp 3025	Thr	Ala	Gly	Trp	Asp 3030	Thr	Arg	Ile
Thr	Leu 3035	Glu	Asp	Leu	Lys	Asn 3040	Glu	Glu	Met	Val	Thr 3045	Asn	His	Met
Glu	Gly 3050	Glu	His	Lys	Lys	Leu 3055	Ala	Glu	Ala	Ile	Phe 3060	Lys	Leu	Thr

-continued	
------------	--

Tyr	Gln 3065		ГЛЗ	Val	Val	Arg 3070		Gln	Arg	Pro	Thr 3075	Pro	Arg	Gly		
Thr	Val 3080		Asp	Ile	Ile	Ser 3085		Arg	Asp	Gln	Arg 3090	Gly	Ser	Gly		
Gln	Val 3095	-		-	_	Leu 3100		Thr	Phe	Thr	Asn 3105	Met	Glu	Ala		
Gln	Leu 3110		Arg	Gln	Met	Glu 3115	-		Gly		Phe 3120	Lys	Ser	Ile		
Gln	His 3125		Thr	Ile		Glu 3130		Ile	Ala	Val	Gln 3135	Asn	Trp	Leu		
Ala	Arg 3140		-	-		Arg 3145			-		Ala 3150	Ile	Ser	Gly		
Asp	Asp	Суз	Val	Val	Lys	Pro 3160		_	_	-	Phe 3165	Ala	Ser	Ala		
Leu		Ala			Asp		Gly	Lys	Ile		Lys 3180	Asp	Ile	Gln		
Gln	Trp 3185		Pro	Ser	Arg		Trp	Asn	Asp		Thr 3195	Gln	Val	Pro		
Phe		Ser	His	His			Glu				Lys 3210	Asp	Gly	Arg		
Val	Leu 3215		Val	Pro	Cys		Asn				Leu 3225	Ile	Gly	Arg		
Ala		Ile	Ser	Gln			Gly	Trp	Ser	Leu	Arg 3240	Glu	Thr	Ala		
Суз		Gly			Tyr	Ala	Gln	Met	Trp	Ser	Leu 3255	Met	Tyr	Phe		
His		Arg	Asp	Leu	Arg		Ala	Ala	Asn	Ala	Ile 3270	Сув	Ser	Ala		
Val		Ser	His	Trp	Val		Thr				Thr 3285	Trp	Ser	Ile		
His	Ala	Lys	His	Glu	Trp		Thr				Met	Leu	Thr	Val		
Trp	Asn	Arg	Val	Trp	Ile		Glu	Asn	Pro	Trp	Met	Glu	Asp	Гла		
Thr	Pro					Glu					Leu	Gly	Гла	Arg		
Glu		Gln	Trp	Суз	Gly		Leu	Ile	Gly	Leu	3330 Thr	Ser	Arg	Ala		
Thr	Trp	Ala	ГЛа	Asn	Ile	Gln		Ala	Ile	Asn	Gln	Val	Arg	Ser		
Leu	Ile		Asn	Glu	Glu	Tyr	Thr	Asp	Tyr	Met	Pro	Ser	Met	Lys		
	3365		Arg	Glu	Glu	Glu		Ala	Gly	Val	Leu	Trp				
Thr	3335 Trp 3350	Ala Gly	Lys Asn	Asn Glu	Ile Glu	3340 Gln 3355 Tyr 3370	Ala Thr	Ala Asp	Ile Tyr	Asn Met	3345 Gln 3360 Pro 3375	Val Ser	Arg	Ser		

				- cont 11	lucu		
agttgttagt	ctacgtggac	cgacaaagac	agattctttg	agggagctaa	gctcaatgta	60	
gttctaacag	tttttaatt	agagagcaga	tctctgatga	ataaccaacg	gaaaaaggcg	120	
aaaaacacgc	ctttcaatat	gctgaaacgc	gagagaaacc	gcgtgtcgac	tgtgcaacag	180	
ctgacaaaga	gattctcact	tggaatgctg	cagggacgag	gaccattaaa	actgttcatg	240	
gccctggtgg	cgttccttcg	tttcctaaca	atcccaccaa	cagcagggat	attgaagaga	300	
tggggaacaa	ttaaaaaatc	aaaagctatt	aatgttttga	gagggttcag	gaaagagatt	360	
ggaaggatgc	tgaacatctt	gaataggaga	cgcagatctg	caggcatgat	cattatgctg	420	
attccaacag	tgatggcgtt	ccatttaacc	acgcgtgggg	gagagccgca	tatgatagtt	480	
agcaagcagg	aaagaggaaa	gtcacttttg	ttcaagacct	ctgcaggtgt	caacatgtgc	540	
accctcattg	cgatggattt	gggagagttg	tgtgaggaca	cgatgaccta	caaatgcccc	600	
cggatcactg	aggcggaacc	agatgacgtt	gactgttggt	gcaatgccac	ggacacatgg	660	
gtgacctatg	gaacgtgctc	tcaaactggc	gaacaccgac	gagacaaacg	ttccgtcgca	720	
ttggccccac	acgtggggct	tggcctagaa	acaagagccg	aaacgtggat	gtcctctgaa	780	
ggtgcttgga	aacagataca	aaaagtagag	acttgggctc	tgagacatcc	aggattcacg	840	
gtgatagccc	tttttctagc	acatgccata	ggaacatcca	tcacccagaa	agggatcatt	900	
ttcattttgc	tgatgctggt	aacaccatct	atggccatgc	gatgcgtggg	aataggcaac	960	
agagacttcg	tggaaggact	gtcaggagca	acatgggtgg	atgtggtact	ggagcatgga	1020	
agttgcgtca	ccaccatggc	aaaaaacaaa	ccaacactgg	acattgaact	cttgaagacg	1080	
gaggtcacaa	accctgcagt	tctgcgtaaa	ttgtgcattg	aagctaaaat	atcaaacacc	1140	
accaccgatt	cgagatgtcc	aacacaagga	gaagccacac	tggtggaaga	acaagacgcg	1200	
aactttgtgt	gccgacgaac	gttcgtggac	agaggctggg	gcaatggctg	tgggctattc	1260	
ggaaaaggta	gtctaataac	gtgtgccaag	tttaagtgtg	tgacaaaact	agaaggaaag	1320	
atagttcaat	atgaaaacct	aaaatattca	gtgatagtca	ccgtccacac	tggagatcag	1380	
caccaggtgg	gaaatgagac	tacagaacat	ggaacaactg	caaccataac	acctcaagct	1440	
cctacgtcgg	aaatacagct	gaccgactac	ggaaccctta	cattagattg	ttcacctagg	1500	
acagggctag	attttaacga	gatggtgttg	ctgacaatga	aagaaagatc	atggcttgtc	1560	
cacaaacaat	ggttcctaga	cttaccactg	ccttggacct	ctggggcttc	aacatcccaa	1620	
gagacttgga	acagacaaga	tttactggtc	acatttaaga	cageteatge	aaagaagcag	1680	
gaagtagtcg	tactaggatc	acaagaagga	gcaatgcaca	ctgcgctgac	tggagcgaca	1740	
gaaatccaaa	cgtcaggaac	gacaacaatt	ttcgcaggac	acctaaaatg	cagactaaaa	1800	
atggacaaac	taactttaaa	agggatgtca	tatgtgatgt	gcacaggete	attcaagtta	1860	
gagaaagaag	tggctgagac	ccagcatgga	actgttctgg	tgcaggttaa	atatgaagga	1920	
acagacgcac	catgcaagat	tecettteg	acccaagatg	agaaaggagc	aacccagaat	1980	
gggagattaa	taacagccaa	ccccatagtc	actgacaaag	aaaaaccagt	caatattgag	2040	
gcagaaccac	cctttggtga	gagctacatc	gtggtaggag	caggtgaaaa	agctttgaaa	2100	
ctaagctggt	tcaagaaagg	aagcagcata	gggaaaatgt	ttgaagcaac	tgcccgagga	2160	
gcacgaagga	tggccattct	gggagacacc	gcatgggact	tcggttctat	aggaggagtg	2220	
ttcacgtcta	tgggaaaact	ggtacaccag	gtttttggaa	ctgcatatgg	agttttgttt	2280	

		-concinued	
agcggagttt cttgga	accat gaaaatagga ataggga	ttc tgctgacatg gctaggatta	2340
aattcaagga acacgi	ccct ttcgatgatg tgcatcg	cag ccggcattgt gacactgtat	2400
ttgggagtca tggtgo	aggc cgatagtggt tgcgttg	tga gctggaaaaa caaagaactg	2460
aaatgtggca gtggga	atttt catcacagac aacgtgc	aca catggacaga acaatacaag	2520
ttccaaccag aatcc	ccttc aaaactagct tcagcta	tcc agaaagccca tgaagaggac	2580
atttgtggaa tccgci	ccagt aacaagactg gagaatc	tga tgtggaaaca aataacacca	2640
gaattgaatc acatto	tatc agaaaatgag gtgaagt	taa ctattatgac aggagacatc	2700
aaaggaatca tgcag	ycagg aaaacgatct ctgcggc	ctc agcccactga gctgaagtat	2760
tcatggaaaa catgg	ygcaa agcaaaaatg ctctcta	cag agtctcataa ccagaccttt	2820
ctcattgatg gcccco	yaaac agcagaatgc cccaaca	caa atagagcttg gaattcgttg	2880
gaagttgaag actat	ggett tggagtatte accacea	ata tatggctaaa attgaaagaa	2940
aaacaggatg tattc	gega etcaaaaete atgteag	cgg ccataaaaga caacagagcc	3000
gtccatgccg atatg	ygtta ttggatagaa agtgcac	tca atgacacatg gaagatagag	3060
aaagcetett teatte	yaagt taaaaactgc cactggc	caa aatcacacac cctctggagc	3120
aatggagtgc tagaaa	agtga gatgataatt ccaaaga	atc tcgctggacc agtgtctcaa	3180
cacaactata gaccaq	ggcta ccatacacaa ataacag	gac catggcatct aggtaagctt	3240
gagatggact ttgatt	tetg tgatggaaca acagtgg	tag tgactgagga ctgcggaaat	3300
agaggaccct ctttga	agaac aaccactgcc tctggaa	aac tcataacaga atggtgctgc	3360
cgatcttgca cattad	ccacc gctaagatac agaggtg	agg atgggtgctg gtacgggatg	3420
gaaatcagac cattga	aagga gaaagaagag aatttgg	tca actccttggt cacagctgga	3480
catgggcagg tcgaca	aactt ttcactagga gtcttgg	gaa tggcattgtt cctggaggaa	3540
atgettagga eeega	ytagg aacgaaacat gcaatac	tac tagttgcagt ttcttttgtg	3600
acattgatca caggga	aacat gtcctttaga gacctgg	gaa gagtgatggt tatggtaggc	3660
gccactatga cggato	gacat aggtatgggc gtgactt	atc ttgccctact agcagccttc	3720
aaagtcagac caacti	ttgc agctggacta ctcttga	gaa agctgacctc caaggaattg	3780
atgatgacta ctata	ggaat tgtactcctc tcccaga	gca ccctaccaga gaccattctt	3840
gagttgactg atgcg	tage ettaggeatg atggtee	tca aaatggtgag aaatatggaa	3900
aagtatcaat tggca	ytgac tatcatggct atcttgt	gcg tcccaaacgc agtgatatta	3960
caaaacgcat ggaaaq	ytgag ttgcacaata ttggcag	tgg tgtccgtttc cccactgttc	4020
ttaacateet cacago	caaaa aacagattgg ataccat	tag cattgacgat caaaggtctc	4080
aatccaacag ctatt	ttct aacaaccctc tcaagaa	cca gcaagaaaag gagctggcca	4140
ttaaatgagg ctatca	atggc agtcgggatg gtgagca	ttt tagccagttc tctcctaaaa	4200
aatgatattc ccatga	acagg accattagtg gctggag	ggc tcctcactgt gtgctacgtg	4260
ctcactggac gatcg	yccga tttggaactg gagagag	cag ccgatgtcaa atgggaagac	4320
caggcagaga tatcag	ggaag cagtccaatc ctgtcaa	taa caatatcaga agatggtagc	4380
atgtcgataa aaaat	yaaga ggaagatcaa acactga	cca tactcattag aacaggattg	4440
ctggtgatct caggad	etttt teetgtatea ataecaa	tca cggcagcagc atggtacctg	4500
tgggaagtga agaaa	aacg ggccggagta ttgtggg	atg tteetteace eccaeceatg	4560

		-concinded	
ggaaaggctg aactggaaga	. tggagcctat agaattaagc	aaaaagggat tcttggatat	4620
tcccagatcg gagccggagt	ttacaaagaa ggaacattcc	atacaatgtg gcatgtcaca	4680
cgtggcgctg ttctaatgca	taaaggaaag aggattgaac	catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgg	aggaggctgg aagttagaag	gagaatggaa ggaaggagaa	4800
gaagtccagg tattggcact	ggagcctgga aaaaatccaa	gagccgtcca aacgaaacct	4860
ggtcttttca aaaccaacgc	cggaacaata ggtgctgtat	ctctggactt ttctcctgga	4920
acgtcaggat ctccaattat	cgacaaaaaa ggaaaagttg	tgggtettta tggtaatggt	4980
gttgttacaa ggagtggagc	atatgtgagt gctatagccc	agactgaaaa aagcattgaa	5040
gacaacccag agatcgaaga	. tgacattttc cgaaagagaa	gactgaccat catggacctc	5100
cacccaggag cgggaaagac	gaagagatac cttccggcca	tagtcagaga agctataaaa	5160
cggggtttga gaacattaat	cttggccccc actagagttg	tggcagctga aatggaggaa	5220
gcccttagag gacttccaat	aagataccag accccagcca	tcagagctgt gcacaccggg	5280
cgggagattg tggacctaat	gtgtcatgcc acatttacca	tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct	gattatcatg gacgaagccc	atttcacaga cccagcaagt	5400
atagcagcta gaggatacat	ctcaactcga gtggagatgg	gtgaggcagc tgggattttt	5460
atgacagcca ctcccccggg	aagcagagac ccatttcctc	agagcaatgc accaatcata	5520
gatgaagaaa gagaaatccc	tgaacgctcg tggaattccg	gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg	gttcgttcca agtataaaag	caggaaatga tatagcagct	5640
tgcctgagga aaaatggaaa	. gaaagtgata caactcagta	ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa	tgattgggac ttcgtggtta	caactgacat ttcagaaatg	5760
ggtgccaatt tcaaggctga	. gagggttata gaccccagac	gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg	ggtgattetg geaggaeeta	tgccagtgac ccactctagt	5880
gcagcacaaa gaagagggag	aataggaaga aatccaaaaa	atgagaatga ccagtacata	5940
tacatggggg aacctctgga	. aaatgatgaa gactgtgcac	actggaaaga agctaaaatg	6000
ctcctagata acatcaacac	gccagaagga atcattccta	gcatgttcga accagagcgt	6060
gaaaaggtgg atgccattga	tggcgaatac cgcttgagag	gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg	agacctacca gtctggttgg	cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag	aaggtggtgt tttgatggag	tcaagaacaa ccaaatccta	6240
gaagaaaacg tggaagttga	. aatctggaca aaagaagggg	aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat	ctattctgac ccactggcgc	taaaagaatt taaggaattt	6360
gcagccggaa gaaagtctct	gaccetgaac etaateacag	aaatgggtag gctcccaacc	6420
ttcatgactc agaaggcaag	agacgcactg gacaacttag	cagtgctgca cacggctgag	6480
gcaggtggaa gggcgtacaa	. ccatgctctc agtgaactgc	cggagaccct ggagacattg	6540
cttttactga cacttctggc	tacagtcacg ggagggatct	ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac	cctgggaatg tgctgcataa	tcacggctag catcctccta	6660
tggtacgcac aaatacagcc	acactggata gcagcttcaa	taatactgga gttttttctc	6720
atagttttgc ttattccaga	. acctgaaaaa cagagaacac	cccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct	cacagtggtg gccgcaacca	tggcaaacga gatgggtttc	6840

-continued	
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc	6900
aacateetgg acatagatet acgteetgea teageatgga egetgtatge egtggeeaca	6960
acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta	7020
acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca toggagttoo oottotogoo attggatgot actoacaagt caacocoata	7140
acteteacag cagetetttt ettattggta geacattatg ceateatagg geeaggaete	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca	7260
actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc gaagtttgaa	7320
aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt	7500
agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg cattaaaaga ggagaaacgg accatcacgc tgtgtcgcga	7740
ggetcagcaa aactgagatg gttegttgag agaaacatgg teacaecaga agggaaagta	7800
gtggacctcg gttgtggcag aggaggctgg tcatactatt gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac aaaaggagga ccaggacacg aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt gcgtcttcaa agtggagttg acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt gtgtgacata ggggagtcat caccaaatcc cacagtggaa	8040
gcaggacgaa cactcagagt ccttaactta gtagaaaatt ggttgaacaa caacactcaa	8100
ttttgcataa aggtteteaa eecatatatg eeeteagtea tagaaaaaat ggaageaeta	8160
caaaggaaat atggaggagc cttagtgagg aatccactct cacgaaactc cacacatgag	8220
atgtactggg tatccaatgc ttccgggaac atagtgtcat cagtgaacat gatttcaagg	8280
atgttgatca acagatttac aatgagatac aagaaagcca cttacgagcc ggatgttgac	8340
ctcggaagcg gaacccgtaa catcgggatt gaaagtgaga taccaaacct agatataatt	8400
gggaaaagaa tagaaaaaat aaagcaagag catgaaacat catggcacta tgaccaagac	8460
cacccataca aaacgtgggc ataccatggt agctatgaaa caaaacagac tggatcagca	8520
tcatccatgg tcaacggagt ggtcaggctg ctgacaaaac cttgggacgt cgtccccatg	8580
gtgacacaga tggcaatgac agacacgact ccatttggac aacagcgcgt ttttaaagag	8640
aaagtggaca cgagaaccca agaaccgaaa gaaggcacga agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga attagggaag aaaaagacac ccaggatgtg caccagagaa	8760
gaattcacaa gaaaggtgag aagcaatgca gccttggggg ccatattcac tgatgagaac	8820
aagtggaagt cggcacgtga ggctgttgaa gatagtaggt tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga aggaaagtgt gaaacatgtg tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga attcggcaag gcaaaaggca gcagagccat atggtacatg	9000
tggettggag caegettett agagtttgaa geeetaggat tettaaatga agateaetgg	9060
tteteeagag agaacteest gagtggagtg gaaggagaag ggetgeacaa getaggttae	9120

-continued	
attctaagag acgtgagcaa gaaagaggga ggagcaatgt atgccgatga caccgcagga	9180
tgggatacaa gaatcacact agaagaccta aaaaatgaag aaatggtaac aaaccacatg	9240
gaaggagaac acaagaaact agccgaggcc attttcaaac taacgtacca aaacaaggtg	9300
gtgcgtgtgc aaagaccaac accaagaggc acagtaatgg acatcatatc gagaagagac	9360
caaagaggta gtggacaagt tggcacctat ggactcaata ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga gggagaagga gtctttaaaa gcattcagca cctaacaatc	9480
acagaagaaa tcgctgtgca aaactggtta gcaagagtgg ggcgcgaaag gttatcaaga	9540
atggccatca gtggagatga ttgtgttgtg aaacctttag atgacaggtt cgcaagcgct	9600
ttaacagctc taaatgacat gggaaagatt aggaaagaca tacaacaatg ggaaccttca	9660
agaggatgga atgattggac acaagtgccc ttctgttcac accatttcca tgagttaatc	9720
atgaaagacg gtcgcgtact cgttgttcca tgtagaaacc aagatgaact gattggcaga	9780
geeegaatet eeeaaggage agggtggtet ttgegggaga eggeetgttt ggggaagtet	9840
tacgcccaaa tgtggagett gatgtaette cacagaegeg aceteagget ggeggeaaat	9900
gctatttget eggeagtace ateacattgg gttecaacaa gtegaacaae etggteeata	9960
catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg	10020
attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca	10080
tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc	10140
acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa	10200
gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga	10260
gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc	10320
catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca	10380
ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg	10440
tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc	10500
ggttagagga gacccctccc ttacaaatcg cagcaacaat ggggggcccaa ggcgagatga	10560
agetgtagte tegetggaag gaetagaggt tagaggagae eeecegaaa caaaaaacag	10620
catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca	10680
gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct	10723
<210> SEQ ID NO 14 <211> LENGTH: 10723 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Dengue virus serotype 2, MVS	
<400> SEQUENCE: 14	
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta	60
gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg	120
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag	180
	240
ctgacaaaga gattctcact tggaatgctg cagggacgag gaccattaaa actgttcatg	
ctgacaaaga gatteteatt tggaatgetg cagggaegag gaecattaaa aetgtteatg geeetggtgg egtteetteg ttteetaaca ateeeaceaa cageagggat attgaagaga	300

				-contir	nuea		
ggaaggatgc	tgaacatctt	gaataggaga	cgcagatctg	caggcatgat	cattatgctg	420	
attccaacag	tgatggcgtt	ccatttaacc	acacgtaacg	gagaaccaca	catgatcgtc	480	
agcagacaag	agaaagggaa	aagtettetg	tttaaaacag	aggttggcgt	gaacatgtgt	540	
accctcatgg	ccatggacct	tggtgaattg	tgtgaagaca	caatcacgta	cgagtgtccc	600	
cttctcaggc	agaatgagcc	agaagacata	gactgttggt	gcaactctac	gtccacgtgg	660	
gtaacttatg	ggacgtgtac	caccatggga	gaacatagaa	gagaaaaaag	atcagtggca	720	
ctcgttccac	atgtgggaat	gggactggag	acacgaactg	aaacatggat	gtcatcagaa	780	
ggggcctgga	aacatgtcca	gagaattgaa	acttggatct	tgagacatcc	aggetteace	840	
atgatggcag	caatcctggc	atacaccata	ggaacgacac	atttccaaag	agccctgatc	900	
ttcatcttac	tgacagctgt	cactccttca	atgacaatgc	gttgcatagg	aatgtcaaat	960	
agagactttg	tggaaggggt	ttcaggagga	agctgggttg	acatagtctt	agaacatgga	1020	
agctgtgtga	cgacgatggc	aaaaacaaa	ccaacattgg	attttgaact	gataaaaaca	1080	
gaagccaaac	agcctgccac	cctaaggaag	tactgtatag	aggcaaagct	aaccaacaca	1140	
acaacagaat	ctcgctgccc	aacacaaggg	gaacccagcc	taaatgaaga	gcaggacaaa	1200	
aggttcgtct	gcaaacactc	catggtagac	agaggatggg	gaaatggatg	tggactattt	1260	
ggaaagggag	gcattgtgac	ctgtgctatg	ttcagatgca	aaaagaacat	ggaaggaaaa	1320	
gttgtgcaac	cagaaaactt	ggaatacacc	attgtgataa	cacctcactc	aggggaagag	1380	
catgcagtcg	gaaatgacac	aggaaaacat	ggcaaggaaa	tcaaaataac	accacagagt	1440	
tccatcacag	aagcagaatt	gacaggttat	ggcactgtca	caatggagtg	ctctccaaga	1500	
acgggcctcg	acttcaatga	gatggtgttg	ctgcagatgg	aaaataaagc	ttggctggtg	1560	
cacaggcaat	ggttcctaga	cctgccgtta	ccatggttgc	ccggagcgga	cacacaaggg	1620	
tcaaattgga	tacagaaaga	gacattggtc	actttcaaaa	atccccatgc	gaagaaacag	1680	
gatgttgttg	ttttaggatc	ccaagaaggg	gccatgcaca	cagcacttac	aggggccaca	1740	
gaaatccaaa	tgtcatcagg	aaacttactc	ttcacaggac	atctcaagtg	caggctgaga	1800	
atggacaagc	tacagctcaa	aggaatgtca	tactctatgt	gcacaggaaa	gtttaaagtt	1860	
gtgaaggaaa	tagcagaaac	acaacatgga	acaatagtta	tcagagtgca	atatgaaggg	1920	
gacggetete	catgcaagat	cccttttgag	ataatggatt	tggaaaaaag	acatgtetta	1980	
ggtcgcctga	ttacagtcaa	cccaattgtg	acagaaaaag	atagcccagt	caacatagaa	2040	
gcagaacctc	catttggaga	cagctacatc	atcataggag	tagagccggg	acaactgaag	2100	
ctcaactggt	ttaagaaagg	aagttctatc	ggccaaatgt	ttgagacaac	aatgaggggg	2160	
gcgaagagaa	tggccatttt	aggtgacaca	gcctgggatt	ttggatcctt	gggaggagtg	2220	
tttacatcta	taggaaaggc	tctccaccaa	gtctttggag	caatctatgg	agetgeette	2280	
agtggggttt	catggactat	gaaaatcctc	ataggagtca	ttatcacatg	gataggaatg	2340	
aattcacgca	gcacctcact	gtctgtgaca	ctagtattgg	tgggaattgt	gacactgtat	2400	
ttgggagtca	tggtgcaggc	cgatagtggt	tgcgttgtga	gctggaaaaa	caaagaactg	2460	
aaatgtggca	gtgggatttt	catcacagac	aacgtgcaca	catggacaga	acaatacaag	2520	
			tcagctatcc			2580	
			gagaatctga			2640	
ulligiyydd	llugulayt	uacaayacty	yayaattiya	cycyyddadd	uacaacaccd	2010	

		-concinued	
gaattgaatc acattctato	c agaaaatgag gtgaagttaa	ctattatgac aggagacatc	2700
aaaggaatca tgcaggcagg	g aaaacgatct ctgcggcctc	agcccactga gctgaagtat	2760
tcatggaaaa catggggcaa	a agcaaaaatg ctctctacag	agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaad	c agcagaatgc cccaacacaa	atagagcttg gaattcgttg	2880
gaagttgaag actatggctt	t tggagtattc accaccaata	tatggctaaa attgaaagaa	2940
aaacaggatg tattctgcga	a ctcaaaactc atgtcagcgg	ccataaaaga caacagagcc	3000
gtccatgccg atatgggtta	a ttggatagaa agtgcactca	atgacacatg gaagatagag	3060
aaagcetett teattgaagt	t taaaaactgc cactggccaa	aatcacacac cctctggagc	3120
aatggagtgc tagaaagtga	a gatgataatt ccaaagaatc	tcgctggacc agtgtctcaa	3180
cacaactata gaccaggcta	a ccatacacaa ataacaggac	catggcatct aggtaagctt	3240
gagatggact ttgatttctg	g tgatggaaca acagtggtag	tgactgagga ctgcggaaat	3300
agaggaccct ctttgagaad	c aaccactgcc tctggaaaac	tcataacaga atggtgctgc	3360
cgatcttgca cattaccaco	c gctaagatac agaggtgagg	atgggtgctg gtacgggatg	3420
gaaatcagac cattgaagga	a gaaagaagag aatttggtca	actccttggt cacagctgga	3480
catgggcagg tcgacaact	t ttcactagga gtcttgggaa	tggcattgtt cctggaggaa	3540
atgettagga eeegagtagg	g aacgaaacat gcaatactac	tagttgcagt ttcttttgtg	3600
acattgatca cagggaacat	t gteetttaga gaeetgggaa	gagtgatggt tatggtaggc	3660
gccactatga cggatgacat	z aggtatgggc gtgacttatc	ttgccctact agcagccttc	3720
aaagtcagac caacttttgo	c agctggacta ctcttgagaa	agctgacctc caaggaattg	3780
atgatgacta ctataggaat	t tgtactcctc tcccagagca	ccataccaga gaccattctt	3840
gagttgactg atgcgttage	c cttaggcatg atggtcctca	aaatggtgag aaatatggaa	3900
aagtatcaat tggcagtgad	c tatcatggct atcttgtgcg	tcccaaacgc agtgatatta	3960
caaaacgcat ggaaagtgaq	g ttgcacaata ttggcagtgg	tgtccgtttc cccactgttc	4020
ttaacateet cacageaaaa	a aacagattgg ataccattag	cattgacgat caaaggtete	4080
aatccaacag ctatttttct	t aacaaccete teaagaacea	gcaagaaaag gagctggcca	4140
ttaaatgagg ctatcatggo	c agtcgggatg gtgagcattt	tagccagttc tctcctaaaa	4200
aatgatattc ccatgacago	g accattagtg gctggagggc	teeteactgt gtgetaegtg	4260
ctcactggac gatcggccga	a tttggaactg gagagagcag	ccgatgtcaa atgggaagac	4320
caggcagaga tatcaggaaq	g cagtccaatc ctgtcaataa	caatatcaga agatggtagc	4380
atgtcgataa aaaatgaaga	a ggaagaacaa acactgacca	tactcattag aacaggattg	4440
ctggtgatct caggacttt	t teetgtatea ataceaatea	cggcagcagc atggtacctg	4500
tgggaagtga agaaacaaco	g ggccggagta ttgtgggatg	tteetteace cecaeceatg	4560
ggaaaggctg aactggaaga	a tggagcctat agaattaagc	aaaaagggat tettggatat	4620
tcccagatcg gagccggagt	t ttacaaagaa ggaacattcc	atacaatgtg gcatgtcaca	4680
cgtggcgctg ttctaatgca	a taaaggaaag aggattgaac	catcatgggc ggacgtcaag	4740
aaagacctaa tatcatatgo	g aggaggetgg aagttagaag	gagaatggaa ggaaggagaa	4800
gaagtccagg tattggcact	: ggagcctgga aaaaatccaa	gagccgtcca aacgaaacct	4860
ggtcttttca aaaccaacgo	c cggaacaata ggtgctgtat	ctctggactt ttctcctgga	4920

-continued	
acgtcaggat ctccaattat cgacaaaaaa ggaaaagttg tgggtcttta tggtaatggt	4980
gttgttacaa ggagtggagc atatgtgagt gctatagccc agactgaaaa aagcattgaa	5040
gacaacccag agatcgaaga tgacattttc cgaaagagaa gactgaccat catggacctc	5100
cacccaggag cgggaaagac gaagagatac cttccggcca tagtcagaga agctataaaa	5160
cggggtttga gaacattaat cttggccccc actagagttg tggcagctga aatggaggaa	5220
gcccttagag gacttccaat aagataccag accccagcca tcagagctgt gcacaccggg	5280
cgggagattg tggacctaat gtgtcatgcc acatttacca tgaggctgct atcaccagtt	5340
agagtgccaa actacaacct gattatcatg gacgaagccc atttcacaga cccagcaagt	5400
atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt	5460
atgacageca eteecceggg aageagagae ceatteete agageaatge aceaateata	5520
gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct	5640
tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg	5760
ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattetg geaggaeeta tgeeagtgae ceaetetagt	5880
gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060
gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta	6240
gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt	6360
gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc	6420
ttcatgactc agaaggcaag agacgcactg gacaacttag cagtgctgca cacggctgag	6480
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccateet cacagtggtg geogeaacea tggeaaaega gatgggttte	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc	6900
aacateetgg acatagatet aegteetgea teageatgga egetgtatge egtggeeaca	6960
acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta	7020
acagctatag ccaaccaagc cacagtgtta atgggtctcg ggaaaggatg gccattgtca	7080
aagatggaca teggagttee eettetegee attggatget acteacaagt caaceceata	7140
acteteacag cagetettt ettattggta geacattatg ceateatagg geeaggaete	7200
5 - 45	

		-continued	
caagcaaaag caaccagaga	. agctcagaaa agagcagcgg	cgggcatcat gaaaaaccca	7260
actgtcgatg gaataacagt	gattgaccta gatccaatac	cttatgatcc aaagtttgaa	7320
aagcagttgg gacaagtaat	gctcctagtc ctctgcgtga	ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga	ggctttaacc ttagctaccg	ggcccatctc cacattgtgg	7440
gaaggaaatc cagggaggtt	ttggaacact accattgcgg	tgtcaatggc taacattttt	7500
agagggagtt acttggccgg	agctggactt ctcttttcta	ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa	. cataggagag acgcttggag	agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga	. attccagatc tacaagaaaa	gtggaatcca ggaagtggat	7680
agaaccttag caaaagaagg	cattaaaaga ggagaaacgg	accatcacgc tgtgtcgcga	7740
ggctcagcaa aactgagate	gttcgttgag agaaacatgg	tcacaccaga agggaaagta	7800
gtggacctcg gttgtggcag	aggaggctgg tcatactatt	gtggaggact aaagaatgta	7860
agagaagtca aaggcctaac	aaaaggagga ccaggacacg	aagaacccat ccccatgtca	7920
acatatgggt ggaatctagt	gcgtcttcaa agtggagttg	acgttttctt catcccgcca	7980
gaaaagtgtg acacattatt	gtgtgacata ggggagtcat	caccaaatcc cacagtggaa	8040
gcaggacgaa cactcagagt	ccttaactta gtagaaaatt	ggttgaacaa caacactcaa	8100
ttttgcataa aggttctcaa	. cccatatatg ccctcagtca	tagaaaaaat ggaagcacta	8160
caaaggaaat atggaggago	cttagtgagg aatccactct	cacgaaactc cacacatgag	8220
atgtactggg tatccaatgo	ttccgggaac atagtgtcat	cagtgaacat gatttcaagg	8280
atgttgatca acagatttac	aatgagatac aagaaagcca	cttacgagcc ggatgttgac	8340
ctcggaagcg gaacccgtaa	. catcgggatt gaaagtgaga	taccaaacct agatataatt	8400
gggaaaagaa tagaaaaaat	aaagcaagag catgaaacat	catggcacta tgaccaagac	8460
cacccataca aaacgtgggc	ataccatggt agctatgaaa	caaaacagac tggatcagca	8520
tcatccatgg tcaacggagt	ggtcaggctg ctgacaaaac	cttgggacgt cgtccccatg	8580
gtgacacaga tggcaatgac	agacacgact ccatttggac	aacagcgcgt ttttaaagag	8640
aaagtggaca cgagaaccca	. agaaccgaaa gaaggcacga	agaaactaat gaaaataaca	8700
gcagagtggc tttggaaaga	. attagggaag aaaaagacac	ccaggatgtg caccagagaa	8760
gaattcacaa gaaaggtgag	aagcaatgca gccttggggg	ccgtattcac tgatgagaac	8820
aagtggaagt cggcacgtga	. ggctgttgaa gatagtaggt	tttgggagct ggttgacaag	8880
gaaaggaatc tccatcttga	. aggaaagtgt gaaacatgtg	tgtacaacat gatgggaaaa	8940
agagagaaga agctagggga	. attcggcaag gcaaaaggca	gcagagccat atggtacatg	9000
tggcttggag cacgcttctt	agagtttgaa gccctaggat	tcttaaatga agatcactgg	9060
ttctccagag agaactccct	gagtggagtg gaaggagaag	ggctgcacaa gctaggttac	9120
attctaagag acgtgagcaa	. gaaagaggga ggagcaatgt	atgccgatga caccgcagga	9180
tgggatacaa gaatcacact	agaagaccta aaaaatgaag	aaatggtaac aaaccacatg	9240
gaaggagaac acaagaaact	agccgaggcc attttcaaac	taacgtacca aaacaaggtg	9300
gtgcgtgtgc aaagaccaac	accaagaggc acagtaatgg	acatcatatc gagaagagac	9360
caaagaggta gtggacaagt	tggcacctat ggactcaata	ctttcaccaa tatggaagcc	9420
caactaatca gacagatgga	. gggagaagga gtctttaaaa	gcattcagca cctaacaatc	9480
_ 0 00			

				- conc 11			
acagaagaaa	tcgctgtgca	aaactggtta	gcaagagtgg	ggcgcgaaag	gttatcaaga	9540	
atggccatca	gtggagatga	ttgtgttgtg	aaacctttag	atgacaggtt	cgcaagcgct	9600	
ttaacagctc	taaatgacat	gggaaagatt	aggaaagaca	tacaacaatg	ggaaccttca	9660	
agaggatgga	atgattggac	acaagtgccc	ttctgttcac	accatttcca	tgagttaatc	9720	
atgaaagacg	gtcgcgtact	cgttgttcca	tgtagaaacc	aagatgaact	gattggcaga	9780	
gcccgaatct	cccaaggagc	agggtggtct	ttgcgggaga	cggcctgttt	ggggaagtct	9840	
tacgcccaaa	tgtggagctt	gatgtacttc	cacagacgcg	acctcaggct	ggcggcaaat	9900	
gctatttgct	cggcagtacc	atcacattgg	gttccaacaa	gtcgaacaac	ctggtccata	9960	
catgctaaac	atgaatggat	gacaacggaa	gacatgctga	cagtctggaa	cagggtgtgg	10020	
attcaagaaa	acccatggat	ggaagacaaa	actccagtgg	aatcatggga	ggaaatccca	10080	
tacttgggga	aaagagaaga	ccaatggtgc	ggctcattga	ttgggttaac	aagcagggcc	10140	
acctgggcaa	agaacatcca	agcagcaata	aatcaagtta	gatcccttat	aggcaatgaa	10200	
gaatacacag	attacatgcc	atccatgaaa	agattcagaa	gagaagagga	agaagcagga	10260	
gttctgtggt	agaaagcaaa	actaacatga	aacaaggcta	gaagtcaggt	cggattaagc	10320	
catagtacgg	aaaaaactat	gctacctgtg	agccccgtcc	aaggacgtta	aaagaagtca	10380	
ggccatcata	aatgccatag	cttgagtaaa	ctatgcagcc	tgtagctcca	cctgagaagg	10440	
tgtaaaaaat	ccgggaggcc	acaaaccatg	gaagctgtac	gcatggcgta	gtggactagc	10500	
ggttagagga	gacccctccc	ttacaaatcg	cagcaacaat	ggggggcccaa	ggcgagatga	10560	
agctgtagtc	tcgctggaag	gactagaggt	tagaggagac	ccccccgaaa	caaaaaacag	10620	
catattgacg	ctgggaaaga	ccagagatcc	tgctgtctcc	tcagcatcat	tccaggcaca	10680	
gaacgccaga	aaatggaatg	gtgctgttga	atcaacaggt	tct		10723	
<220> FEAT	TH: 10717 : DNA NISM: Artif:			Ype 3, MVS			
<400> SEQU	ENCE: 15						
agttgttagt	ctacgtggac	cgacaaagac	agattetttg	agggagctaa	gctcaatgta	60	
gttctaacag	tttttaatt	agagagcaga	tctctgatga	ataaccaacg	gaaaaaggcg	120	
aaaaacacgc	ctttcaatat	gctgaaacgc	gagagaaacc	gcgtgtcgac	tgtgcaacag	180	
ctgacaaaga	gatteteact	tggaatgctg	cagggacgag	gaccattaaa	actgttcatg	240	
gccctggtgg	cgttccttcg	tttcctaaca	atcccaccaa	cagcagggat	attgaagaga	300	
tggggaacaa	ttaaaaaatc	aaaagctatt	aatgttttga	gagggttcag	gaaagagatt	360	
ggaaggatgc	tgaacatctt	gaataggaga	cgcagatctg	caggcatgat	cattatgctg	420	
attccaacag	tgatggcgtt	ccatttaacc	acgcgtgatg	gagagccgcg	catgattgtg	480	
gggaagaatg	aaagaggaaa	atccctactt	ttcaagacag	cctctggaat	caacatgtgc	540	
acactcatag	ctatggatct	gggagagatg	tgtgatgaca	cggtcactta	caaatgcccc	600	
cacattaccg	aagtggagcc	tgaagacatt	gactgctggt	gcaaccttac	atcgacatgg	660	
	gaacatgcaa					720	
		5 55					

-continued	
ttagctcccc atgttggcat gggactggac acacgcactc aaacctggat gtcggctgaa	780
ggagettgga gacaagtega gaaggtagag acatgggeee ttaggeaeee agggtttaee	840
atactageee tatteettge ceattacata ggeaetteet tgaeeeagaa agtggttatt	900
tttatactat taatgctggt taccccatcc atgacaatga gatgtgtagg agtaggaaac	960
agagattttg tggaaggcct atcgggagct acgtgggttg acgtggtgct cgagcacggt	1020
gggtgtgtga ctaccatggc taagaacaag cccacgctgg acatagagct tcagaagacc	1080
gaggccaccc aactggcgac cctaaggaag ctatgcattg agggaaaaat taccaacata	1140
acaaccgact caagatgtcc cacccaaggg gaagcgattt tacctgagga gcaggaccag	1200
aactacgtgt gtaagcatac atacgtggac agaggctggg gaaacggttg tggtttgttt	1260
ggcaagggaa gcttggtgac atgcgcgaaa tttcaatgtt tagaatcaat agagggaaaa	1320
gtggtgcaac atgagaacct caaatacacc gtcatcatca cagtgcacac aggagaccaa	1380
caccaggtgg gaaatgaaac gcagggagtc acggctgaga taacacccca ggcatcaacc	1440
gctgaagcca ttttacctga atatggaacc ctcgggctag aatgctcacc acggacaggt	1500
ttggatttca atgaaatgat ctcattgaca atgaagaaca aagcatggat ggtacataga	1560
caatggttct ttgacttacc cctaccatgg acatcaggag cttcagcaga aacaccaact	1620
tggaacagga aagagcttct tgtgacattt aaaaatgcac atgcaaaaaa gcaagaagta	1680
gttgttcttg gatcacaaga gggagcaatg catacagcac tgacaggagc tacagagatc	1740
caaacctcag gaggcacaag tatctttgcg gggcacttaa aatgtagact caagatggac	1800
aaattggaac tcaaggggat gagctatgca atgtgcttga gtagctttgt gttgaagaaa	1860
gaagteteeg aaacgeagea tgggacaata eteattaagg ttgagtacaa aggggaagat	1920
gcaccctgca agattccttt ctccacggag gatggacaag gaaaagctct caatggcaga	1980
ctgatcacag ccaatccagt ggtgaccaag aaggaggagc ctgtcaacat tgaggctgaa	2040
cctccttttg gagaaagtaa catagtaatt ggaattggag acaaagccct gaaaatcaac	2100
tggtacaaga agggaagctc gattgggaag atgttcgagg ccactgccag aggtgcaagg	2160
cgcatggcca tcttgggaga cacagcctgg gactttggat cagtgggtgg tgttttgaat	2220
tcattaggga aaatggtcca ccaaatattt gggagtgctt acacagccct atttggtgga	2280
gtctcctgga tgatgaaaat tggaataggt gtcctcttaa cctggatagg gttgaactca	2340
aaaaatactt ctatgtcatt ttcatgcatc gcggccggca ttgtgacact gtatttggga	2400
gtcatggtgc aggccgatag tggttgcgtt gtgagctgga aaaacaaaga actgaaatgt	2460
ggcagtggga ttttcatcac agacaacgtg cacacatgga cagaacaata caagttccaa	2520
ccagaatccc cttcaaaact agcttcagct atccagaaag cccatgaaga ggacatttgt	2580
ggaatccgct cagtaacaag actggagaat ctgatgtgga aacaaataac accagaattg	2640
aatcacattc tatcagaaaa tgaggtgaag ttaactatta tgacaggaga catcaaagga	2700
atcatgcagg caggaaaacg atctctgcgg cctcagccca ctgagctgaa gtattcatgg	2760
aaaacatggg gcaaagcaaa aatgctctct acagagtctc ataaccagac ctttctcatt	2820
gatggccccg aaacagcaga atgccccaac acaaatagag cttggaattc gttggaagtt	2880
gaagactatg gctttggagt attcaccacc aatatatggc taaaattgaa agaaaaacag	2940
gatgtattet gegaeteaaa acteatgtea geggeeataa aagaeaacag ageegteeat	3000
5 5 5-5 <u>5-55</u> <u>5</u> <u>5</u>	

-continued	
gccgatatgg gttattggat agaaagtgca ctcaatgaca catggaagat agagaaagcc	3060
tettteattg aagttaaaaa etgecaetgg eeaaaateae acaeeetetg gageaatgga	3120
gtgctagaaa gtgagatgat aattccaaag aatctcgctg gaccagtgtc tcaacacaac	3180
tatagaccag gctaccatac acaaataaca ggaccatggc atctaggtaa gcttgagatg	3240
gactttgatt tctgtgatgg aacaacagtg gtagtgactg aggactgcgg aaatagagga	3300
ccctctttga gaacaaccac tgcctctgga aaactcataa cagaatggtg ctgccgatct	3360
tgcacattac caccgctaag atacagaggt gaggatgggt gctggtacgg gatggaaatc	3420
agaccattga aggagaaaga agagaatttg gtcaactcct tggtcacagc tggacatggg	3480
caggtcgaca acttttcact aggagtcttg ggaatggcat tgttcctgga ggaaatgctt	3540
aggacccgag taggaacgaa acatgcaata ctactagttg cagtttcttt tgtgacattg	3600
atcacaggga acatgtcctt tagagacctg ggaagagtga tggttatggt aggcgccact	3660
atgacggatg acataggtat gggcgtgact tatcttgccc tactagcagc cttcaaagtc	3720
agaccaactt ttgcagctgg actactcttg agaaagctga cctccaagga attgatgatg	3780
actactatag gaattgtact cctctcccag agcaccatac cagagaccat tcttgagttg	3840
actgatgcgt tagccttagg catgatggtc ctcaaaatgg tgagaaatat ggaaaagtat	3900
caattggcag tgactatcat ggctatcttg tgcgtcccaa acgcagtgat attacaaaac	3960
gcatggaaag tgagttgcac aatattggca gtggtgtccg tttccccact gttcttaaca	4020
tcctcacagc aaaaaacaga ttggatacca ttagcattga cgatcaaagg tctcaatcca	4080
acagctattt ttctaacaac cctctcaaga accagcaaga aaaggagctg gccattaaat	4140
gaggctatca tggcagtcgg gatggtgagc attttagcca gttctctcct aaaaaatgat	4200
atteccatga caggaccatt agtggetgga gggeteetea etgtgtgeta egtgeteaet	4260
ggacgatcgg ccgatttgga actggagaga gcagccgatg tcaaatggga agaccaggca	4320
gagatatcag gaagcagtcc aatcctgtca ataacaatat cagaagatgg tagcatgtcg	4380
ataaaaaatg aagaggaaga acaaacactg accatactca ttagaacagg attgctggtg	4440
atctcaggac tttttcctgt atcaatacca atcacggcag cagcatggta cctgtgggaa	4500
gtgaagaaac aacgggccgg agtattgtgg gatgtteett caececeace catgggaaag	4560
gctgaactgg aagatggagc ctatagaatt aagcaaaaag ggattcttgg atattcccag	4620
atcggagccg gagtttacaa agaaggaaca ttccatacaa tgtggcatgt cacacgtggc	4680
gctgttctaa tgcataaagg aaagaggatt gaaccatcat gggcggacgt caagaaagac	4740
ctaatatcat atggaggagg ctggaagtta gaaggagaat ggaaggaagg agaagaagtc	4800
caggtattgg cactggagcc tggaaaaaat ccaagagccg tccaaacgaa acctggtctt	4860
ttcaaaacca acgccggaac aataggtgct gtatctctgg acttttctcc tggaacgtca	4920
ggateteeaa ttategacaa aaaaggaaaa gttgtgggte tttatggtaa tggtgttgtt	4980
acaaggagtg gagcatatgt gagtgctata gcccagactg aaaaaagcat tgaagacaac	5040
ccagagatcg aagatgacat tttccgaaag agaagactga ccatcatgga cctccaccca	5100
ggagcgggaa agacgaagag ataccttccg gccatagtca gagaagctat aaaacggggt	5160
ttgagaacat taatcttggc ccccactaga gttgtggcag ctgaaatgga ggaagccctt	5220
agaggactte caataagata ecagaceeca geeateagag etgtgeacae egggegggag	5280
2 22 03330333443	

		-concinued	
attgtggacc taatgtgtca	tgccacattt accatgaggc	tgctatcacc agttagagtg	5340
ccaaactaca acctgattat	catggacgaa gcccatttca	cagacccagc aagtatagca	5400
gctagaggat acatctcaac	tcgagtggag atgggtgagg	cagctgggat ttttatgaca	5460
gccactcccc cgggaagcag	agacccattt cctcagagca	atgcaccaat catagatgaa	5520
gaaagagaaa tccctgaacg	ctcgtggaat tccggacatg	aatgggtcac ggattttaaa	5580
gggaagactg tttggttcgt	tccaagtata aaagcaggaa	atgatatagc agcttgcctg	5640
aggaaaaatg gaaagaaagt	gatacaactc agtaggaaga	cctttgattc tgagtatgtc	5700
aagactagaa ccaatgattg	ggacttcgtg gttacaactg	acatttcaga aatgggtgcc	5760
aatttcaagg ctgagagggt	tatagacccc agacgctgca	tgaaaccagt catactaaca	5820
gatggtgaag agcgggtgat	tctggcagga cctatgccag	tgacccactc tagtgcagca	5880
caaagaagag ggagaatagg	aagaaatcca aaaaatgaga	atgaccagta catatacatg	5940
ggggaacctc tggaaaatga	tgaagactgt gcacactgga	aagaagctaa aatgctccta	6000
gataacatca acacgccaga	aggaatcatt cctagcatgt	tcgaaccaga gcgtgaaaag	6060
gtggatgcca ttgatggcga	ataccgcttg agaggagaag	caaggaaaac ctttgtagac	6120
ttaatgagaa gaggagacct	accagtctgg ttggcctaca	gagtggcagc tgaaggcatc	6180
aactacgcag acagaaggtg	gtgttttgat ggagtcaaga	acaaccaaat cctagaagaa	6240
aacgtggaag ttgaaatctg	gacaaaagaa ggggaaagga	agaaattgaa acccagatgg	6300
ttggatgcta ggatctattc	tgacccactg gcgctaaaag	aatttaagga atttgcagcc	6360
ggaagaaagt ctctgaccct	gaacctaatc acagaaatgg	gtaggctccc aaccttcatg	6420
actcagaagg caagagacgc	actggacaac ttagcagtgc	tgcacacggc tgaggcaggt	6480
ggaagggcgt acaaccatgc	tctcagtgaa ctgccggaga	ccctggagac attgctttta	6540
ctgacacttc tggctacagt	cacgggaggg atcttttat	tcttgatgag cgcaaggggc	6600
atagggaaga tgaccctggg	aatgtgctgc ataatcacgg	ctagcatect ectatggtae	6660
gcacaaatac agccacactg	gatagcagct tcaataatac	tggagttttt tctcatagtt	6720
ttgcttattc cagaacctga	aaaacagaga acaccccaag	acaaccaact gacctacgtt	6780
gtcatagcca tcctcacagt	ggtggccgca accatggcaa	acgagatggg tttcctagaa	6840
aaaacgaaga aagatctcgg	attgggaagc attgcaaccc	agcaacccga gagcaacatc	6900
ctggacatag atctacgtcc	tgcatcagca tggacgctgt	atgeegtgge cacaacattt	6960
gttacaccaa tgttgagaca	tagcattgaa aatteeteag	tgaatgtgtc cctaacagct	7020
atagccaacc aagccacagt	gttaatgggt ctcgggaaag	gatggccatt gtcaaagatg	7080
gacatcggag ttccccttct	cgccattgga tgctactcac	aagtcaaccc cataactctc	7140
acagcagete ttttettatt	ggtagcacat tatgccatca	tagggccagg actccaagca	7200
aaagcaacca gagaagctca	gaaaagagca gcggcgggca	tcatgaaaaa cccaactgtc	7260
gatggaataa cagtgattga	cctagatcca ataccttatg	atccaaagtt tgaaaagcag	7320
ttgggacaag taatgctcct	agteetetge gtgaeteaag	tattgatgat gaggactaca	7380
tgggctctgt gtgaggcttt	aaccttagct accgggccca	tctccacatt gtgggaagga	7440
aatccaggga ggttttggaa	cactaccatt gcggtgtcaa	tggctaacat ttttagaggg	7500
agttacttgg ccggagctgg	acttctcttt tctattatga	agaacacaac caacacaaga	7560
	5	5	

				- conc 11	lued	
aggggaactg	gcaacatagg	agagacgctt	ggagagaaat	ggaaaagccg	attgaacgcg	7620
ttgggaaaaa	gtgaattcca	gatctacaag	aaaagtggaa	tccaggaagt	ggatagaacc	7680
ttagcaaaag	aaggcattaa	aagaggagaa	acggaccatc	acgctgtgtc	gcgaggctca	7740
gcaaaactga	gatggttcgt	tgagagaaac	atggtcacac	cagaagggaa	agtagtggac	7800
ctcggttgtg	gcagaggagg	ctggtcatac	tattgtggag	gactaaagaa	tgtaagagaa	7860
gtcaaaggcc	taacaaaagg	aggaccagga	cacgaagaac	ccatccccat	gtcaacatat	7920
gggtggaatc	tagtgcgtct	tcaaagtgga	gttgacgttt	tcttcatccc	gccagaaaag	7980
tgtgacacat	tattgtgtga	cataggggag	tcatcaccaa	atcccacagt	ggaagcagga	8040
cgaacactca	gagtccttaa	cttagtagaa	aattggttga	acaacaacac	tcaattttgc	8100
ataaaggttc	tcaacccata	tatgccctca	gtcatagaaa	aaatggaagc	actacaaagg	8160
aaatatggag	gagccttagt	gaggaatcca	ctctcacgaa	actccacaca	tgagatgtac	8220
tgggtatcca	atgcttccgg	gaacatagtg	tcatcagtga	acatgatttc	aaggatgttg	8280
atcaacagat	ttacaatgag	atacaagaaa	gccacttacg	agccggatgt	tgacctcgga	8340
agcggaaccc	gtaacatcgg	gattgaaagt	gagataccaa	acctagatat	aattgggaaa	8400
agaatagaaa	aaataaagca	agagcatgaa	acatcatggc	actatgacca	agaccaccca	8460
tacaaaacgt	gggcatacca	tggtagctat	gaaacaaaac	agactggatc	agcatcatcc	8520
atggtcaacg	gagtggtcag	gctgctgaca	aaaccttggg	acgtcgtccc	catggtgaca	8580
cagatggcaa	tgacagacac	gactccattt	ggacaacagc	gcgtttttaa	agagaaagtg	8640
gacacgagaa	cccaagaacc	gaaagaaggc	acgaagaaac	taatgaaaat	aacagcagag	8700
tggctttgga	aagaattagg	gaagaaaaag	acacccagga	tgtgcaccag	agaagaattc	8760
acaagaaagg	tgagaagcaa	tgcagccttg	ggggccatat	tcactgatga	gaacaagtgg	8820
aagtcggcac	gtgaggctgt	tgaagatagt	aggttttggg	agctggttga	caaggaaagg	8880
aatctccatc	ttgaaggaaa	gtgtgaaaca	tgtgtgtaca	acatgatggg	aaaaagagag	8940
aagaagctag	gggaattcgg	caaggcaaaa	ggcagcagag	ccatatggta	catgtggctt	9000
ggagcacgct	tcttagagtt	tgaagcccta	ggattettaa	atgaagatca	ctggttctcc	9060
agagagaact	ccctgagtgg	agtggaagga	gaagggctgc	acaagctagg	ttacattcta	9120
agagacgtga	gcaagaaaga	gggaggagca	atgtatgccg	atgacaccgc	aggatgggat	9180
acaagaatca	cactagaaga	cctaaaaaat	gaagaaatgg	taacaaacca	catggaagga	9240
gaacacaaga	aactagccga	ggccattttc	aaactaacgt	accaaaacaa	ggtggtgcgt	9300
gtgcaaagac	caacaccaag	aggcacagta	atggacatca	tatcgagaag	agaccaaaga	9360
ggtagtggac	aagttggcac	ctatggactc	aatactttca	ccaatatgga	agcccaacta	9420
atcagacaga	tggagggaga	aggagtcttt	aaaagcattc	agcacctaac	aatcacagaa	9480
gaaatcgctg	tgcaaaactg	gttagcaaga	gtggggcgcg	aaaggttatc	aagaatggcc	9540
atcagtggag	atgattgtgt	tgtgaaacct	ttagatgaca	ggttcgcaag	cgctttaaca	9600
gctctaaatg	acatgggaaa	gattaggaaa	gacatacaac	aatgggaacc	ttcaagagga	9660
tggaatgatt	ggacacaagt	gcccttctgt	tcacaccatt	tccatgagtt	aatcatgaaa	9720
gacggtcgcg	tactcgttgt	tccatgtaga	aaccaagatg	aactgattgg	cagagcccga	9780
atctcccaag	gagcagggtg	gtctttgcgg	gagacggcct	gtttggggaa	gtcttacgcc	9840
	0	0	-		-	

-cont	inued
-conc	Innea

-continued	
caaatgtgga gcttgatgta cttccacaga cgcgacctca ggctggcggc aaatgctatt	9900
tgctcggcag taccatcaca ttgggttcca acaagtcgaa caacctggtc catacatgct	9960
aaacatgaat ggatgacaac ggaagacatg ctgacagtct ggaacagggt gtggattcaa	10020
gaaaacccat ggatggaaga caaaactcca gtggaatcat gggaggaaat cccatacttg	10080
gggaaaagag aagaccaatg gtgcggctca ttgattgggt taacaagcag ggccacctgg	10140
gcaaagaaca tccaagcagc aataaatcaa gttagatccc ttataggcaa tgaagaatac	10200
acagattaca tgccatccat gaaaagattc agaagagaag	10260
tggtagaaag caaaactaac atgaaacaag gctagaagtc aggtcggatt aagccatagt	10320
acggaaaaaa ctatgctacc tgtgagcccc gtccaaggac gttaaaagaa gtcaggccat	10380
cataaatgcc atagcttgag taaactatgc agcctgtagc tccacctgag aaggtgtaaa	10440
aaatccggga ggccacaaac catggaagct gtacgcatgg cgtagtggac tagcggttag	10500
aggagacccc tcccttacaa atcgcagcaa caatggggggc ccaaggcgag atgaagctgt	10560
agtetegetg gaaggaetag aggttagagg agaeceeee gaaacaaaaa acageatatt	10620
gacgctggga aagaccagag atcctgctgt ctcctcagca tcattccagg cacagaacgc	10680
cagaaaatgg aatggtgctg ttgaatcaac aggttct	10717
<pre>&lt;220&gt; FEATURE: &lt;220&gt; FEATURE: &lt;220&gt; FEATURE: &lt;221&gt; NAME/KEY: n &lt;222&gt; LOCATION: (3773)(3773) &lt;222&gt; LOCATION: (3773)(3773) &lt;222&gt; FEATURE: &lt;221&gt; NAME/KEY: n &lt;221&gt; NAME/KEY: n &lt;222&gt; LOCATION: (7026)(7026) &lt;223&gt; OTHER INFORMATION: t or c</pre>	
<400> SEQUENCE: 16	
agttgttagt ctacgtggac cgacaaagac agattctttg agggagctaa gctcaatgta	60
gttctaacag ttttttaatt agagagcaga tctctgatga ataaccaacg gaaaaaggcg	120
aaaaacacgc ctttcaatat gctgaaacgc gagagaaacc gcgtgtcgac tgtgcaacag	180
ctgacaaaga gatteteact tggaatgetg cagggaegag gaeetttaaa aetgtteatg	240
gccctggtgg cgttccttcg tttcctaaca atcccaccaa cagcagggat attgaagaga	300
tggggaacaa ttaaaaaatc aaaagctatt aatgttttga gagggttcag gaaagagatt	360
ggaaggatgc tgaacatott gaataggaga cgcagototg caggoatgat cattatgotg	420
attecaacag tgatggegtt ecatttaace aegegtgatg gegaaceeet catgatagtg	480
gcaaaacatg aaagggggag acctctcttg tttaagacaa cagaggggat caacaaatgc	540
acteteattg ceatggaett gggtgaaatg tgtgaggaea etgteaegta taaatgeeee	600
ttactggtca ataccgaacc tgaagacatt gattgctggt gcaatctcac gtctacctgg	660
gtcatgtatg ggacatgcac ccagagcgga gaacggagac gagagaagcg ctcagtagct	
222 22222-222 2000220200 20202020	720
ttaacaccac attcaggaat gggattggaa acaagagctg agacatggat gtcatcggaa	720 780

-continued	
ctcttggcag gatttatggc ttatatgatt gggcaaacag gaatccagcg aactgtcttc	900
tttgtcctaa tgatgctggt cgccccatcc tacggaatgc gatgcgtagg agtaggaaac	960
agagactttg tggaaggagt ctcaggtgga gcatgggtcg atctggtgct agaacatgga	1020
ggatgcgtca caaccatggc ccagggaaaa ccaaccttgg attttgaact gactaagaca	1080
acagccaagg aagtggctct gttaagaacc tattgcattg aagcctcaat atcaaacata	1140
accacggcaa caagatgtcc aacgcaagga gagccttatc taaaagagga acaagaccaa	1200
cagtacattt gccggagaga tgtggtagac agagggtggg gcaatggctg tggcttgttt	1260
ggaaaaggag gagttgtgac atgtgcgaag ttttcatgtt cggggaagat aacaggcaat	1320
ttggtccaaa ttgagaacct tgaatacaca gtggttgtaa cagtccacaa tggagacacc	1380
catgcagtag gaaatgacac gtccaatcat ggagttacag ccacgataac tcccaggtca	1440
ccatcggtgg aagtcaaatt gccggactat ggagaactaa cactcgattg tgaacccagg	1500
totggaattg actttaatga gatgattotg atgaaaatga aaaagaaaac atggottgtg	1560
cataagcaat ggtttttgga tctacctcta ccatggacag caggagcaga cacatcagag	1620
gttcactgga attacaaaga gagaatggtg acatttaagg ttcctcatgc caagagacag	1680
gatgtgacag tgctgggatc tcaggaagga gccatgcatt ctgccctcgc tggagccaca	1740
gaagtggact ccggtgatgg aaatcacatg tttgcaggac atctcaagtg caaagtccgt	1800
atggagaaat tgagaatcaa gggaatgtca tacacgatgt gttcaggaaa gttctcaatt	1860
gacaaagaga tggcagaaac acagcatggg acaacagtgg tgaaagtcaa gtatgaaggt	1920
gctggagctc cgtgtaaagt ccccatagag ataagagatg tgaacaagga aaaagtggtt	1980
gggcgtatca teteateeae eeetttgget gagaataeea acagtgtaae caacatagag	2040
ttagaacccc cctttgggga cagctacata gtgataggtg ttggaaacag tgcattaaca	2100
ctccattggt tcaggaaagg gagttccatt ggcaagatgt ttgagtccac atacagaggt	2160
gcaaaacgaa tggccattct aggtgaaaca gcttgggatt ttggttccgt tggtggactg	2220
ttcacatcat tgggaaaggc tgtgcaccag gtttttggaa gtgtgtatac aaccctgttt	2280
ggaggagtet catggatgat tagaateeta attgggttee tagtgttgtg gattggeaeg	2340
aactcaagga acacttcaat ggctatgacg tgcatagctg ccggcattgt gacactgtat	2400
ttgggagtca tggtgcaggc cgatagtggt tgcgttgtga gctggaaaaa caaagaactg	2460
aaatgtggca gtgggatttt catcacagac aacgtgcaca catggacaga acaatacaag	2520
ttccaaccag aatccccttc aaaactagct tcagctatcc agaaagccca tgaagaggac	2580
atttgtggaa teegeteagt aacaagaetg gagaatetga tgtggaaaca aataacaeca	2640
gaattgaatc acattctatc agaaaatgag gtgaagttaa ctattatgac aggagacatc	2700
aaaggaatca tgcaggcagg aaaacgatct ctgcggcctc agcccactga gctgaagtat	2760
tcatggaaaa catggggcaa agcaaaaatg ctctctacag agtctcataa ccagaccttt	2820
ctcattgatg gccccgaaac agcagaatgc cccaacacaa atagagcttg gaattcgttg	2880
gaagttgaag actatggctt tggagtattc accaccaata tatggctaaa attgaaagaa	2940
aaacaggatg tattctgcga ctcaaaactc atgtcagcgg ccataaaaga caacagagcc	3000
gtccatgccg atatgggtta ttggatagaa agtgcactca atgacacatg gaagatagag	3060
aaagcetett teattgaagt taaaaaetge caetggeeaa aateaeaeae eetetggage	3120

				-cont II	Iueu	
aatggagtgc	tagaaagtga	gatgataatt	ccaaagaatc	tcgctggacc	agtgtctcaa	3180
cacaactata	gaccaggcta	ccatacacaa	ataacaggac	catggcatct	aggtaagctt	3240
gagatggact	ttgatttctg	tgatggaaca	acagtggtag	tgactgagga	ctgcggaaat	3300
agaggaccct	ctttgagaac	aaccactgcc	tctggaaaac	tcataacaga	atggtgctgc	3360
cgatcttgca	cattaccacc	gctaagatac	agaggtgagg	atgggtgctg	gtacgggatg	3420
gaaatcagac	cattgaagga	gaaagaagag	aatttggtca	actccttggt	cacagctgga	3480
catgggcagg	tcgacaactt	ttcactagga	gtcttgggaa	tggcattgtt	cctggaggaa	3540
atgcttagga	cccgagtagg	aacgaaacat	gcaatactac	tagttgcagt	ttcttttgtg	3600
acattgatca	cagggaacat	gtcctttaga	gacctgggaa	gagtgatggt	tatggtaggc	3660
gccactatga	cgggtgacat	aggtatgggc	gtgacttatc	ttgccctact	agcagcette	3720
aaagtcagac	caacttttgc	agctggacta	ctcttgagaa	agctgacctc	canggaattg	3780
atgatgacta	ctataggaat	tgtactcctc	tcccagagca	ccataccaga	gaccattctt	3840
gagttgactg	atgcgttagc	cttaggcatg	atggtcctca	aaatggtgag	aaatatggaa	3900
aagtatcaat	tggcagtgac	tatcatggct	atcttgtgcg	tcccaaacgc	agtgatatta	3960
caaaacgcat	ggaaagtgag	ttgcacaata	ttggcagtgg	tgtccgtttc	cccactgttc	4020
ttaacatcct	cacagcaaaa	aacagattgg	ataccattag	cattgacgat	caaaggtctc	4080
aatccaacag	ctatttttct	aacaaccctc	tcaagaacca	gcaagaaaag	gagctggcca	4140
ttaaatgagg	ctatcatggc	agtcgggatg	gtgagcattt	tagccagttc	tctcctaaaa	4200
aatgatattc	ccatgacagg	accattagtg	gctggagggc	tcctcactgt	gtgctacgtg	4260
ctcactggac	gatcggccga	tttggaactg	gagagagcag	ccgatgtcaa	atgggaagac	4320
caggcagaga	tatcaggaag	cagtccaatc	ctgtcaataa	caatatcaga	agatggtagc	4380
atgtcgataa	aaaatgaaga	ggaagaacaa	acactgacca	tactcattag	aacaggattg	4440
ctggtgatct	caggactttt	tcctgtatca	ataccaatca	cggcagcagc	atggtacctg	4500
tgggaagtga	agaaacaacg	ggccggagta	ttgtgggatg	ttccttcacc	cccacccatg	4560
ggaaaggctg	aactggaaga	tggagcctat	agaattaagc	aaaaagggat	tcttggatat	4620
tcccagatcg	gagccggagt	ttacaaagaa	ggaacattcc	atacaatgtg	gcatgtcaca	4680
cgtggcgctg	ttctaatgca	taaaggaaag	aggattgaac	catcatgggc	ggacgtcaag	4740
aaagacctaa	tatcatatgg	aggaggctgg	aagttagaag	gagaatggaa	ggaaggagaa	4800
gaagtccagg	tattggcact	ggagcctgga	aaaaatccaa	gagccgtcca	aacgaaacct	4860
ggtcttttca	aaaccaacgc	cggaacaata	ggtgctgtat	ctctggactt	tteteetgga	4920
acgtcaggat	ctccaattat	cgacaaaaaa	ggaaaagttg	tgggtcttta	tggtaatggt	4980
gttgttacaa	ggagtggagc	atatgtgagt	gctatagccc	agactgaaaa	aagcattgaa	5040
gacaacccag	agatcgaaga	tgacattttc	cgaaagagaa	gactgaccat	catggacctc	5100
cacccaggag	cgggaaagac	gaagagatac	cttccggcca	tagtcagaga	agctataaaa	5160
cggggtttga	gaacattaat	cttggccccc	actagagttg	tggcagctga	aatggaggaa	5220
gcccttagag	gacttccaat	aagataccag	accccagcca	tcagagctgt	gcacaccggg	5280
cgggagattg	tggacctaat	gtgtcatgcc	acatttacca	tgaggctgct	atcaccagtt	5340
agagtgccaa	actacaacct	gattatcatg	gacgaagccc	atttcacaga	tccagcaagt	5400

-continuea	
atagcagcta gaggatacat ctcaactcga gtggagatgg gtgaggcagc tgggattttt	5460
atgacagcca ctcccccggg aagcagagac ccatttcctc agagcaatgc accaatcata	5520
gatgaagaaa gagaaatccc tgaacgctcg tggaattccg gacatgaatg ggtcacggat	5580
tttaaaggga agactgtttg gttcgttcca agtataaaag caggaaatga tatagcagct	5640
tgcctgagga aaaatggaaa gaaagtgata caactcagta ggaagacctt tgattctgag	5700
tatgtcaaga ctagaaccaa tgattgggac ttcgtggtta caactgacat ttcagaaatg	5760
ggtgccaatt tcaaggctga gagggttata gaccccagac gctgcatgaa accagtcata	5820
ctaacagatg gtgaagagcg ggtgattetg geaggaeeta tgeeagtgae ceactetagt	5880
gcagcacaaa gaagagggag aataggaaga aatccaaaaa atgagaatga ccagtacata	5940
tacatggggg aacctctgga aaatgatgaa gactgtgcac actggaaaga agctaaaatg	6000
ctcctagata acatcaacac gccagaagga atcattccta gcatgttcga accagagcgt	6060
gaaaaggtgg atgccattga tggcgaatac cgcttgagag gagaagcaag gaaaaccttt	6120
gtagacttaa tgagaagagg agacctacca gtctggttgg cctacagagt ggcagctgaa	6180
ggcatcaact acgcagacag aaggtggtgt tttgatggag tcaagaacaa ccaaatccta	6240
gaagaaaacg tggaagttga aatctggaca aaagaagggg aaaggaagaa attgaaaccc	6300
agatggttgg atgctaggat ctattctgac ccactggcgc taaaagaatt taaggaattt	6360
gcagccggaa gaaagtctct gaccctgaac ctaatcacag aaatgggtag gctcccaacc	6420
ttcatgactc agaaggtaag agacgcactg gacaacttag cagtgctgca cacggctgag	6480
gcaggtggaa gggcgtacaa ccatgctctc agtgaactgc cggagaccct ggagacattg	6540
cttttactga cacttctggc tacagtcacg ggagggatct ttttattctt gatgagcgca	6600
aggggcatag ggaagatgac cctgggaatg tgctgcataa tcacggctag catcctccta	6660
tggtacgcac aaatacagcc acactggata gcagcttcaa taatactgga gttttttctc	6720
atagttttgc ttattccaga acctgaaaaa cagagaacac cccaagacaa ccaactgacc	6780
tacgttgtca tagccatcct cacagtggtg gccgcaacca tggcaaacga gatgggtttc	6840
ctagaaaaaa cgaagaaaga tctcggattg ggaagcattg caacccagca acccgagagc	6900
aacateetgg acatagatet aegteetgea teageatgga egetgtatge egtggeeaca	6960
acatttgtta caccaatgtt gagacatagc attgaaaatt cctcagtgaa tgtgtcccta	7020
acagenatag ceaaceaage cacagtgtta atgggteteg ggaaaggatg geeattgtea	7080
aagatggaca tcggagttcc ccttctcgcc attggatgct actcacaagt caaccccata	7140
acteteacag cagetetttt ettattggta geacattatg ceateatagg geeaggaete	7200
caagcaaaag caaccagaga agctcagaaa agagcagcgg cgggcatcat gaaaaaccca	7260
actgtcgatg gaataacagt gattgaccta gatccaatac cttatgatcc aaagtttgaa	7320
aagcagttgg gacaagtaat gctcctagtc ctctgcgtga ctcaagtatt gatgatgagg	7380
actacatggg ctctgtgtga ggctttaacc ttagctaccg ggcccatctc cacattgtgg	7440
gaaggaaatc cagggaggtt ttggaacact accattgcgg tgtcaatggc taacattttt	7500
agagggagtt acttggccgg agctggactt ctcttttcta ttatgaagaa cacaaccaac	7560
acaagaaggg gaactggcaa cataggagag acgcttggag agaaatggaa aagccgattg	7620
aacgcattgg gaaaaagtga attccagatc tacaagaaaa gtggaatcca ggaagtggat	7680
5 55 5	

				-001011	lucu	
agaaccttag	caaaagaagg	cattaaaaga	ggagaaacgg	accatcacgc	tgtgtcgcga	7740
ggctcagcaa	aactgagatg	gttcgttgag	agaaacatgg	tcacaccaga	agggaaagta	7800
gtggacctcg	gttgtggcag	aggaggctgg	tcatactatt	gtggaggact	aaagaatgta	7860
agagaagtca	aaggcctaac	aaaaggagga	ccaggacacg	aagaacccat	ccccatgtca	7920
acatatgggt	ggaatctagt	gcgtcttcaa	agtggagttg	acgttttctt	catcccgcca	7980
gaaaagtgtg	acacattatt	gtgtgacata	ggggagtcat	caccaaatcc	cacagtggaa	8040
gcaggacgaa	cactcagagt	ccttaactta	gtagaaaatt	ggttgaacaa	caacactcaa	8100
ttttgcataa	aggttctcaa	cccatatatg	ccctcagtca	tagaaaaaat	ggaagcacta	8160
caaaggaaat	atggaggagc	cttagtgagg	aatccactct	cacgaaactc	cacacatgag	8220
atgtactggg	tatccaatgc	ttccgggaac	atagtgtcat	cagtgaacat	gatttcaagg	8280
atgttgatca	acagatttac	aatgagatac	aagaaagcca	cttacgagcc	ggatgttgac	8340
ctcggaagcg	gaacccgtaa	catcgggatt	gaaagtgaga	taccaaacct	agatataatt	8400
gggaaaagaa	tagaaaaaat	aaagcaagag	catgaaacat	catggcacta	tgaccaagac	8460
cacccataca	aaacgtgggc	ataccatggt	agctatgaaa	caaaacagac	tggatcagca	8520
tcatccatgg	tcaacggagt	ggtcaggctg	ctgacaaaac	cttgggacgt	cgtccccatg	8580
gtgacacaga	tggcaatgac	agacacgact	ccatttggac	aacagcgcgt	ttttaaagag	8640
aaagtggaca	cgagaaccca	agaaccgaaa	gaaggcacga	agaaactaat	gaaaataaca	8700
gcagagtggc	tttggaaaga	attagggaag	aaaaagacac	ccaggatgtg	caccagagaa	8760
gaattcacaa	gaaaggtgag	aagcaatgca	gccttggggg	ccatattcac	tgatgagaac	8820
aagtggaagt	cggcacgtga	ggctgttgaa	gatagtaggt	tttgggagct	ggttgacaag	8880
gaaaggaatc	tccatcttga	aggaaagtgt	gaaacatgtg	tgtacaacat	gatgggaaaa	8940
agagagaaga	agctagggga	attcggcaag	gcaaaaggca	gcagagccat	atggtacatg	9000
tggcttggag	cacgcttctt	agagtttgaa	gccctaggat	tcttaaatga	agatcactgg	9060
ttctccagag	agaactccct	gagtggagtg	gaaggagaag	ggctgcacaa	gctaggttac	9120
attctaagag	acgtgagcaa	gaaagaggga	ggagcaatgt	atgccgatga	caccgcagga	9180
tgggatacaa	gaatcacact	agaagaccta	aaaaatgaag	aaatggtaac	aaaccacatg	9240
gaaggagaac	acaagaaact	agccgaggcc	attttcaaac	taacgtacca	aaacaaggtg	9300
gtgcgtgtgc	aaagaccaac	accaagaggc	acagtaatgg	acatcatatc	gagaagagac	9360
caaagaggta	gtggacaagt	tggcacctat	ggactcaata	ctttcaccaa	tatggaagcc	9420
caactaatca	gacagatgga	gggagaagga	gtctttaaaa	gcattcagca	cctaacaatc	9480
acagaagaaa	tcgctgtgca	aaactggtta	gcaagagtgg	ggcgcgaaag	gttatcaaga	9540
atggccatca	gtggagatga	ttgtgttgtg	aaacctttag	atgacaggtt	cgcaagcgct	9600
ttaacagete	taaatgacat	gggaaagatt	aggaaagaca	tacaacaatg	ggaaccttca	9660
agaggatgga	atgattggac	acaagtgccc	ttctgttcac	accatttcca	tgagttaatc	9720
atgaaagacg	gtcgcgtact	cgttgttccc	tgtagaaacc	aagatgaact	gattggcaga	9780
gcccgaatct	cccaaggagc	agggtggtct	ttgcgggaga	cggcctgttt	ggggaagtct	9840
tacgcccaaa	tgtggagctt	gatgtacttc	cacagacgcg	acctcaggct	ggcggcaaat	9900
gctatttgct	cggcagtacc	atcacattgg	gttccaacaa	gtcgaacaac	ctggtccata	9960

-cont	tin	ued
-------	-----	-----

catgctaaac atgaatggat gacaacggaa gacatgctga cagtctggaa cagggtgtgg 10020 attcaagaaa acccatggat ggaagacaaa actccagtgg aatcatggga ggaaatccca 10080 tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140 acctggggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200 gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260 gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320 catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380 ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440 tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
tacttgggga aaagagaaga ccaatggtgc ggctcattga ttgggttaac aagcagggcc 10140 acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200 gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260 gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320 catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380 ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440 tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
acctgggcaa agaacatcca agcagcaata aatcaagtta gatcccttat aggcaatgaa 10200 gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260 gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320 catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380 ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440 tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
gaatacacag attacatgcc atccatgaaa agattcagaa gagaagagga agaagcagga 10260 gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320 catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380 ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440 tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
gttctgtggt agaaagcaaa actaacatga aacaaggcta gaagtcaggt cggattaagc 10320 catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380 ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440 tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
catagtacgg aaaaaactat gctacctgtg agccccgtcc aaggacgtta aaagaagtca 10380 ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440 tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
ggccatcata aatgccatag cttgagtaaa ctatgcagcc tgtagctcca cctgagaagg 10440 tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
tgtaaaaaat ccgggaggcc acaaaccatg gaagctgtac gcatggcgta gtggactagc 10500
ggttagagga gacccctccc ttacaaatcg cagcaacaat gggggcccaa ggcgagatga 10560
agctgtagtc tcgctggaag gactagaggt tagaggagac ccccccgaaa caaaaaacag 10620
catattgacg ctgggaaaga ccagagatcc tgctgtctcc tcagcatcat tccaggcaca 10680
gaacgccaga aaatggaatg gtgctgttga atcaacaggt tct 10723

1. A modified live, attenuated dengue-2 virus strain PDK-53, wherein the modified live, attenuated dengue-2 virus strain PDK-53

- is represented by a polynucleotide molecule encoding a modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule,
- comprises an RNA transcribed from a cDNA comprising a polynucleotide molecule encoding the modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule,
- comprises one or more polypeptide molecules encoded by a polynucleotide molecule encoding the modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule, or
- is obtainable by a method for producing a modified live, attenuated dengue-2 virus strain PDK-53, the method comprising the following steps:
- a) transcribing an RNA from a cDNA comprising a polynucleotide molecule encoding the modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule, and
- b) introducing the RNA transcribed in step a) into cells for production of the modified live, attenuated dengue-2 virus strain PDK-53,

wherein the polynucleotide molecule encoding the modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule comprises at least one mutation, wherein the at least one mutation comprises:

- an adenine to guanine mutation at position 592 encoding a glutamic acid instead of a lysine in the polypeptide molecule at amino acid position 166 corresponding to prM-52, and
- an adenine to guanine mutation at position 8803 encoding a valine instead of an isoleucine in the polypeptide molecule at amino acid position 2903 corresponding to NS5-412.

**2**. The modified live, attenuated dengue-2 virus strain PDK-53 according to claim **1**, wherein the polynucleotide molecule encoding the modified live, attenuated dengue-2

virus strain PDK-53 polypeptide molecule further comprises at least one additional mutation of:

- a guanine to cytosine mutation at nucleic acid position 6481 encoding a proline instead of an alanine in the polypeptide molecule at amino acid position 2129 corresponding to NS4A-36, and
- a cytosine to thymine mutation at position 7156 encoding a phenylalanine instead of a leucine in the polypeptide molecule at amino acid position 2354 corresponding to NS4B-111.

**3**. The modified live, attenuated dengue-2 virus strain PDK-53 according to claim **1**, wherein the polynucleotide molecule encoding the modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule is represented by SEQ ID NO: 4 or SEQ ID NO: 14.

**4**. The modified live, attenuated dengue-2 virus strain PDK-53 according to claim **1**, wherein the modified, live attenuated dengue-2 virus strain PDK-53 polypeptide molecule is represented by SEQ ID NO: 5 or SEQ ID NO: 6.

**5**. A pharmaceutical composition comprising the modified live, attenuated dengue-2 virus strain PDK-53 according to claim **1**, and a pharmaceutically acceptable excipient.

**6**. A method for inducing an immune response against dengue virus in a subject, the method comprising administering to the subject a pharmaceutically acceptable amount of the composition of claim **5**.

7. A vector comprising a polynucleotide molecule encoding a modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule, wherein the polynucleotide molecule encoding the modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule comprises at least one mutation, wherein the at least one mutation comprises:

- an adenine to guanine mutation at position 592 encoding a glutamic acid instead of a lysine in the polypeptide molecule at amino acid position 166 corresponding to prM-52, and
- an adenine to guanine mutation at position 8803 encoding a valine instead of an isoleucine in the polypeptide molecule at amino acid position 2903 corresponding to NS5-412.

**8**. The vector according to claim **7**, wherein the vector is a plasmid vector.

**9**. The vector according to claim **7**, wherein the vector is a cDNA infectious clone.

- **10**. An isolated cell comprising
- the modified live, attenuated dengue-2 virus strain PDK-53 according to claim 1, or
- a vector comprising a polynucleotide molecule encoding a modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule, wherein the polynucleotide molecule encoding the modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule comprises at least one mutation, wherein the at least one mutation comprises:
- an adenine to guanine mutation at position 592 encoding a glutamic acid instead of a lysine in the polypeptide molecule at amino acid position 166 corresponding to prM-52, and
- an adenine to guanine mutation at position 8803 encoding a valine instead of an isoleucine in the polypeptide molecule at amino acid position 2903 corresponding to NS5-412.

**11**. A method for producing a modified live, attenuated dengue-2 virus strain PDK-53, the method comprising the following steps:

- a) transcribing an RNA from a cDNA comprising a polynucleotide molecule encoding a modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule, wherein the polynucleotide molecule encoding the modified live, attenuated dengue-2 virus strain PDK-53 polypeptide molecule comprises at least one mutation, wherein the at least one mutation comprises:
- an adenine to guanine mutation at position 592 encoding a glutamic acid instead of a lysine in the polypeptide molecule at amino acid position 166 corresponding to prM-52, and
- an adenine to guanine mutation at position 8803 encoding a valine instead of an isoleucine in the polypeptide molecule at amino acid position 2903 corresponding to NS5-412;
- and
  - b) introducing the RNA transcribed in step a) into cells for production of the modified live, attenuated dengue-2 virus strain PDK-53.

**12**. The method of claim **11**, further comprising the following step:

c) passaging the modified live, attenuated dengue-2 virus strain PDK-53 produced in step b) up to 10 times in cells.

13. The method of claim 11, wherein the cells are Vero cells.

14. An immunogenic composition comprising the modified live, attenuated dengue-2 virus strain PDK-53 according to claim 1 and a pharmaceutically acceptable carrier.

**15**. The immunogenic composition of claim **14**, further comprising a dengue-1/dengue-2 chimera, wherein the dengue-1/dengue-2 chimera

- is represented by a polynucleotide molecule encoding a dengue-1/dengue-2 polypeptide chimera,
- comprises an RNA transcribed from a cDNA comprising a polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera,

- comprises one or more polypeptide molecules encoded by a polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera, or
- is obtainable by a method for producing a dengue-1/ dengue-2 chimera, the method comprising the following steps:
- a) transcribing an RNA from a cDNA comprising a polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera, and
- b) introducing the RNA transcribed in step a) into cells for production of the dengue-1/dengue-2 chimera,

wherein the polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera comprises a first nucleotide sequence encoding nonstructural proteins from a modified live, attenuated dengue-2 virus strain PDK-53, a second nucleotide sequence encoding at least one structural protein from dengue-1, and at least one mutation, wherein the at least one mutation comprises one or more of:

- an adenine to cytosine mutation at position 3823 encoding a leucine instead of an isoleucine in the dengue-1/ dengue-2 polypeptide chimera at amino acid position 1243 corresponding to NS2A-116;
- an adenine to thymine mutation at position 4407 encoding an aspartic acid instead of a glutamic acid in the dengue-1/dengue-2 polypeptide chimera at amino acid position 1437 corresponding to NS2B-92; and
- an adenine to guanine mutation at position 7311.

**16**. The immunogenic composition according to claim **15**, wherein the polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera further comprises at least one additional mutation of:

- a cytosine to thymine mutation at position 7148 encoding an isoleucine instead of a threonine in the dengue-1/ dengue-2 polypeptide chimera at amino acid position 2351 corresponding to NS4B-108; and
- a guanine to cytosine mutation at position 2384 encoding an alanine instead of glycine in the dengue-1/dengue-2 polypeptide chimera at amino acid position 763 corresponding to E-483.

**17**. The immunogenic composition according to claim **15**, wherein the polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera is represented by SEQ ID NO: 1 or SEQ ID NO: 13.

**18**. The immunogenic composition according to claim **15**, wherein the dengue-1/dengue-2 polypeptide chimera is represented by SEQ ID NO: 2 or SEQ ID NO: 3.

**19**. The immunogenic composition of claim **15**, wherein the nonstructural proteins from the modified live, attenuated dengue-2 virus strain PDK-53 are selected from the group consisting of NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5.

**20**. The immunogenic composition of claim **15**, wherein the at least one structural protein from dengue-1 is selected from the group consisting of capsid protein (C), premembrane/membrane protein (prM) and envelope protein (E).

**21**. The immunogenic composition of claim **14**, further comprising a dengue-3/dengue-2 chimera, wherein the dengue-3/dengue-2 chimera

- is represented by a polynucleotide molecule encoding a dengue-3/dengue-2 polypeptide chimera,
- comprises an RNA transcribed from a cDNA comprising a polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera,

- comprises one or more polypeptide molecules encoded by a polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera, or
- is obtainable by a method for producing a dengue-3/ dengue-2 chimera, the method comprising the following steps:
- a) transcribing an RNA from a cDNA comprising a polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera, and
- b) introducing the RNA transcribed in step a) into cells for production of the dengue-3/dengue-2 chimera,

wherein the polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera comprises a first nucleotide sequence encoding nonstructural proteins from a modified live, attenuated dengue-2 virus strain PDK-53, a second nucleotide sequence encoding at least one structural protein from dengue-3, and at least one mutation, wherein the at least one mutation comprises one or more of:

an adenine to thymine mutation at position 1603 encoding a serine instead of a threonine in the dengue-3/dengue-2 polypeptide chimera at amino acid position 503 corresponding to E-223; and

an adenine to guanine mutation at position 7620.

**22**. The immunogenic composition of claim **21**, wherein the polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera further comprises a guanine to adenine mutation at position 6436 encoding an asparagine instead of an aspartic acid in the dengue-3/dengue-2 polypeptide chimera at amino acid position 2114 corresponding to NS4A-23.

**23**. The immunogenic composition of claim **21**, wherein the polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera is represented by SEQ ID NO: 7 or SEQ ID NO: 15.

**24**. The immunogenic composition of claim **21**, wherein the dengue-3/dengue-2 polypeptide chimera is represented by SEQ ID NO: 8 or SEQ ID NO: 9.

**25**. The immunogenic composition of claim **21**, wherein the nonstructural proteins from the modified live, attenuated dengue-2 virus strain PDK-53 are selected from the group consisting of NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5.

**26**. The immunogenic composition of claim **21**, wherein the at least one structural protein from dengue-3 is selected from the group consisting of capsid protein (C), premembrane/membrane protein (prM) and envelope protein (E).

**27**. The immunogenic composition of claim **14**, further comprising a dengue-4/dengue-2 chimera, wherein the dengue-4/dengue-2 chimera

- is represented by a polynucleotide molecule encoding a dengue-4/dengue-2 polypeptide chimera,
- comprises an RNA transcribed from a cDNA comprising a polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera,
- comprises one or more polypeptide molecules encoded by a polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera, or
- is obtainable by a method for producing a dengue-4/ dengue-2 chimera, the method comprising the following steps:
- a) transcribing an RNA from a cDNA comprising a polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera, and

b) introducing the RNA transcribed in step a) into cells for production of the dengue-4/dengue-2 chimera,

wherein the polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera comprises a first nucleotide sequence encoding nonstructural proteins from a modified live, attenuated dengue-2 virus strain PDK-53, a second nucleotide sequence encoding at least one structural protein from dengue-4, and at least one mutation, wherein the at least one mutation comprises one or more of:

an adenine to thymine mutation at position 225;

- an adenine to guanine mutation at position 3674 encoding a glycine instead of an aspartic acid in the dengue-4/ dengue-2 polypeptide chimera at amino acid position 1193 corresponding to NS2A-66;
- a cytosine to thymine mutation at position 5391;
- a cytosine to thymine mutation at position 6437 encoding a valine instead of an alanine in the dengue-4/dengue-2 polypeptide chimera at amino acid position 2114 corresponding to NS4A-21, and

an adenine to cytosine mutation at position 9750.

**28**. The immunogenic composition of claim **27**, wherein the polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera further comprises a thymine to cytosine mutation at position 7026.

**29**. The immunogenic composition of claim **27**, wherein the polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera further comprises an adenine to guanine mutation at position 3773 encoding an arginine instead of a lysine in the dengue-4/dengue-2 polypeptide chimera at amino acid position 1226 corresponding to NS2A-99.

**30**. The immunogenic composition of claim **27**, wherein the polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera further comprises a cytosine to thymine mutation at position 7538 encoding a phenylalanine instead of a serine in the dengue-4/dengue-2 polypeptide chimera at amino acid position 2481 corresponding to NS4B-238.

**31**. The immunogenic composition of claim **27**, wherein the polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera is represented by SEQ ID NO: 10 or SEQ ID NO: 16.

**32**. The immunogenic composition of claim **27**, wherein the dengue-4/dengue-2 polypeptide chimera is represented by SEQ ID NO: 11 or SEQ ID NO: 12.

**33**. The immunogenic composition of claim **27**, wherein the nonstructural proteins from the modified live, attenuated dengue-2 virus strain PDK-53 are selected from the group consisting of NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5.

**34**. The immunogenic composition of claim **27**, wherein the at least one structural protein from dengue-4 is selected from the group consisting of capsid protein (C), premembrane/membrane protein (prM) and envelope protein (E).

**35.** An immunogenic composition comprising the modified live, attenuated dengue-2 virus strain PDK-53 according to claim 1 and a pharmaceutically acceptable carrier, wherein the immunogenic composition is a tetravalent composition.

**36**. The immunogenic composition of claim **35**, further comprising one or more of:

i) a dengue-1/dengue-2 chimera, wherein the dengue-1/ dengue-2 chimera

- is represented by a polynucleotide molecule encoding a dengue-1/dengue-2 polypeptide chimera,
- comprises an RNA transcribed from a cDNA comprising a polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera,
- comprises one or more polypeptide molecules encoded by a polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera, or
- is obtainable by a method for producing a dengue-1/ dengue-2 chimera, the method comprising the following steps:
- a) transcribing an RNA from a cDNA comprising a polynucleotide molecule encoding the dengue-1/ dengue-2 polypeptide chimera, and
- b) introducing the RNA transcribed in step a) into cells for production of the dengue-1/dengue-2 chimera,
- wherein the polynucleotide molecule encoding the dengue-1/dengue-2 polypeptide chimera comprises a first nucleotide sequence encoding nonstructural proteins from a modified live, attenuated dengue-2 virus strain PDK-53, a second nucleotide sequence encoding at least one structural protein from dengue-1, and at least one mutation, wherein the at least one mutation comprises one or more of:
  - an adenine to cytosine mutation at position 3823 encoding a leucine instead of an isoleucine in the dengue-1/dengue-2 polypeptide chimera at amino acid position 1243 corresponding to NS2A-116;
  - an adenine to thymine mutation at position 4407 encoding an aspartic acid instead of a glutamic acid in the dengue-1/dengue-2 polypeptide chimera at amino acid position 1437 corresponding to NS2B-92; and an adenine to guanine mutation at position 7311;
- ii) a dengue-3/dengue-2 chimera, wherein the dengue-3/ dengue-2 chimera
  - is represented by a polynucleotide molecule encoding a dengue-3/dengue-2 polypeptide chimera,
  - comprises an RNA transcribed from a cDNA comprising a polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera,
  - comprises one or more polypeptide molecules encoded by a polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera, or
  - is obtainable by a method for producing a dengue-3/ dengue-2 chimera, the method comprising the following steps:
  - a) transcribing an RNA from a cDNA comprising a polynucleotide molecule encoding the dengue-3/ dengue-2 polypeptide chimera, and
  - b) introducing the RNA transcribed in step a) into cells for production of the dengue-3/dengue-2 chimera,
- wherein the polynucleotide molecule encoding the dengue-3/dengue-2 polypeptide chimera comprises a first nucleotide sequence encoding nonstructural proteins from a modified live, attenuated dengue-2 virus strain PDK-53, a second nucleotide sequence encoding at least one structural protein from dengue-3, and at least one mutation, wherein the at least one mutation comprises one or more of:
  - an adenine to thymine mutation at position 1603 encoding a serine instead of a threonine in the dengue-3/ dengue-2 polypeptide chimera at amino acid position 503 corresponding to E-223; and
  - an adenine to guanine mutation at position 7620; and

- iii) a dengue-4/dengue-2 chimera, wherein the dengue-4/ dengue-2 chimera
  - is represented by a polynucleotide molecule encoding a dengue-4/dengue-2 polypeptide chimera,
  - comprises an RNA transcribed from a cDNA comprising a polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera,
  - comprises one or more polypeptide molecules encoded by a polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera, or
  - is obtainable by a method for producing a dengue-4/ dengue-2 chimera, the method comprising the following steps:
  - a) transcribing an RNA from a cDNA comprising a polynucleotide molecule encoding the dengue-4/ dengue-2 polypeptide chimera, and
  - b) introducing the RNA transcribed in step a) into cells for production of the dengue-4/dengue-2 chimera,
- wherein the polynucleotide molecule encoding the dengue-4/dengue-2 polypeptide chimera comprises a first nucleotide sequence encoding nonstructural proteins from a modified live, attenuated dengue-2 virus strain PDK-53, a second nucleotide sequence encoding at least one structural protein from dengue-4, and at least one mutation, wherein the at least one mutation comprises one or more of:
  - an adenine to thymine mutation at position 225;
  - an adenine to guanine mutation at position 3674 encoding a glycine instead of an aspartic acid in the dengue-4/dengue-2 polypeptide chimera at amino acid position 1193 corresponding to NS2A-66;
  - a cytosine to thymine mutation at position 5391;
  - a cytosine to thymine mutation at position 6437 encoding a valine instead of an alanine in the dengue-4/ dengue-2 polypeptide chimera at amino acid position 2114 corresponding to NS4A-21, and
  - an adenine to cytosine mutation at position 9750.

**37**. The immunogenic composition of claim **35**, in combination with an immunogenic composition for a flavivirus selected from the group consisting of yellow fever virus, tick-borne encephalitis virus, Japanese encephalitis virus, West Nile virus, hepatitis C virus, and a combination of two or more thereof.

**38**. A method for inducing an immune response against all four dengue virus serotypes in a subject, the method comprising administering to the subject a pharmaceutically acceptable amount of the immunogenic composition of claim **35**.

**39**. A method for inducing an immune response against all four dengue virus serotypes in a subject, the method comprising administering to the subject a pharmaceutically acceptable amount of the immunogenic composition of claim **36**.

- **40**. A kit comprising
- at least one modified live, attenuated dengue-2 virus strain PDK-53 according to claim 1; or
- at least one composition comprising a modified live, attenuated dengue-2 virus strain PDK-53 according to

claim 1 and a pharmaceutically acceptable excipient; and a container.

\* \* \* \* \*