(11) Application No. AU 2001263689 B2

(19) AUSTRALIAN PATENT OFFICE

(54) Title

Modified bovine adenovirus having altered tropism

(51)[6] International Patent Classification(s)

C12N 15/09 (2006.01) 20060101ALI2005122 0BMJP **A61P** 37/02
A61K 35/76 (2006.01) 20060101ALI2005122 0BMJP **C12N** 1/15
A61K 39/00 (2006.01) 20060101ALI2005122 0BMJP **C12N** 1/19
A61K 48/00 (2006.01) 20060101ALI2005122 0BMJP **C12N** 1/21
A61P 31/00 (2006.01) 20060101ALI2005122 0BMJP **C12N** 5/10
A61P 37/02 (2006.01) 20060101ALI2005122 0BMJP **C12N** 7/00
C12N 1/15 (2006.01) 20060101ALI2005122 0BMJP **C12N** 7/04
C12N 1/19 (2006.01) 20060101ALI2005100 8BMEP **C12N** 15/861
C12N 1/21 (2006.01) 20060101ALI2005100 8BMEP
C12N 5/10 (2006.01) PCT/CA01/00798
C12N 7/00 (2006.01)
C12N 7/04 (2006.01)
C12N 15/861 (2006.01)
C12N 15/09 20060101AFI2005122 0BMJP **A61K** 35/76 20060101ALI2005122 0BMJP **A61K** 39/00 20060101ALI2005122 0BMJP **A61K** 48/00 20060101ALI2005122 0BMJP **A61P** 31/00

(21) Application No: 2001263689 (22) Application Date: 2001.05.31

(87) WIPO No: WO01/92547

(30) Priority Data

(31) Number	(32) Date	(33) Country
60/208,678	2000.05.31	US

(43) Publication Date : 2001.12.11

(43) Publication Journal Date : 2002.02.28

(71) Applicant(s)

University of Saskatchewan

(72) Inventor(s)

Tikoo, Suresh K.

(74) Agent/Attorney

Blake Dawson Waldron Patent Services, Level 39 101 Collins Street, Melbourne, VIC, 3000

(56) Related Art

Xu, Z. Z. and Both, G. W., 1998, Virology, 248: 156-163

Douglas, J. T. et al., 1996, Nature Biotechnology, 14(1): 1574-1578

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/092547 A3

(51) International Patent Classification[7]: C12N 15/86, A61K 48/00

(21) International Application Number: PCT/CA01/00798

(22) International Filing Date: 31 May 2001 (31.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/208,678 31 May 2000 (31.05.2000) US

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF SASKATCHEWAN** [CA/CA]; 120 Veterinary Road, Saskatoon, Saskatchewan S7N 5E3 (CA).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **TIKOO, Suresh, K.** [CA/CA]; 302-102 Edinburgh Place, Saskatoon, Saskatchewan S7H 5J7 (CA).

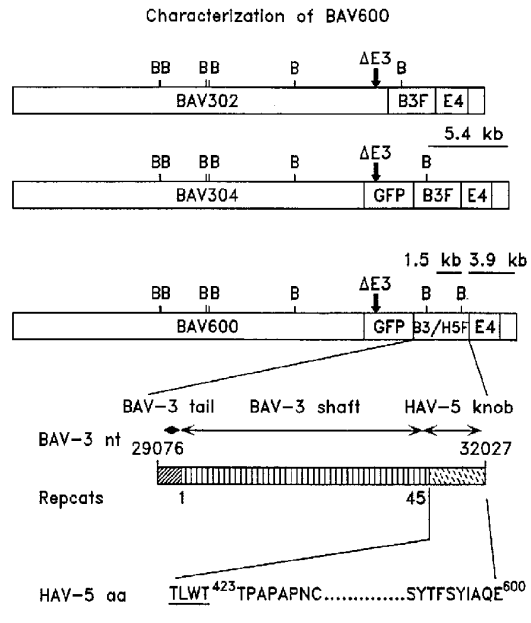
(74) Agents: **MARSMAN, Kathleen** et al.; Borden Ladner Gervais LLP, 1000-60 Queen Street, Ottawa, Ontario K1P 5Y7 (CA).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,

[Continued on next page]

(54) Title: MODIFIED BOVINE ADENOVIRUS HAVING ALTERED TROPISM



(57) Abstract: The present invention provides modified bovine adenoviruses comprising a modification in a capsid protein wherein said protein is associated with adenovirus tropism and wherein said modification is associated with altered tropism. The present invention provides adenovirus vectors and host cells comprising such vectors. The present invention also provides methods of making and using such adenoviruses.

WO 01/092547 A3

IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
8 August 2002

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

MODIFIED BOVINE ADENOVIRUS HAVING ALTERED TROPISM

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 60/208,678, filed May 31, 2000.

TECHNICAL FIELD

This invention relates to bovine adenoviruses comprising a modification in a capsid protein and which exhibit altered tropism. The present invention also relates to methods of making and using bovine adenoviruses having altered tropism.

BACKGROUND ART

The adenoviruses cause enteric or respiratory infection in humans as well as in domestic and laboratory animals. The bovine adenoviruses (BAV) comprise at least nine serotypes divided into two subgroups. These subgroups have been characterized based on enzyme-linked immunoassays (ELISA), serologic studies with immunofluorescence assays, virus-neutralization tests, immunoelectron microscopy, by their host specificity and clinical syndromes. Subgroup 1 viruses include BAV 1, 2, 3 and 9 and grow relatively well in established bovine cells compared to subgroup 2 which includes BAV 4, 5, 6, 7 and 8.

BAV3 was first isolated in 1965 and is the best characterized of the BAV genotypes, containing a genome of approximately 35 kb (Kurokawa et al (1978) *J. Virol.* 28:212-218). Reddy et al. (1998, *Journal of Virology*, 72:1394) disclose nucleotide sequence, genome organization, and transcription map of BAV3. Reddy et al. (1999, *Journal of Virology*, 73: 9137) disclose a replication-defective BAV3 as an expression vector. BAV3, a representative of subgroup 1 of BAVs (Bartha (1969) *Acta Vet. Acad. Sci. Hung.* 19:319-321), is a common pathogen of cattle usually resulting in subclinical infection (Darbyshire *et al.* (1965). *J. Comp. Pathol.* 75:327-330), though occasionally associated with a more serious respiratory tract infection (Darbyshire et al., 1966 *Res. Vet. Sci.* 7:81-93; Mattson et al., 1988 *J. Vet Res* 49:67-69). Like other adenoviruses, BAV3 is a non-enveloped icosahedral particle of 75 nm in diameter (Niiyama *et al.* (1975) *J. Virol.* 16:621-633) containing a linear double-stranded DNA molecule. BAV3 can produce

tumors when injected into hamsters (Darbyshire, 1966 *Nature* 211:102) and viral DNA can efficiently effect morphological transformation of mouse, hamster or rat cells in culture (Tsukamoto and Sugino, 1972 *J. Virol.* 9:465-473; Motoi et al., 1972 *Gann* 63:415-418). Cross hybridization was observed between BAV3 and human adenovirus type 2 (HAd2) (Hu et al., 1984 *J. Virol.* 49:604-608) in most regions of the genome including some regions near but not at the left end of the genome.

Porcine adenovirus (PAV) infection has been associated with encephalitis, pneumonia, kidney lesions and diarrhea. See Derbyshire (1992) In: "Diseases of Swine" (ed. Leman *et al.*), 7th edition, Iowa State University Press, Ames, IA. pp. 225-227. It has been shown that PAV is capable of stimulating both humoral response and a mucosal antibody responses in the intestine of infected piglets. Tuboly *et al.* (1993) *Res. in Vet. Sci.* 54:345-350. Cross-neutralization studies have indicated the existence of at least five serotypes of PAV. See Derbyshire *et al.* (1975) *J. Comp. Pathol.* 85:437-443; and Hirahara *et al.* (1990) *Jpn. J. Vet. Sci.* 52:407-409. Previous studies of the PAV genome have included the determination of restriction maps for PAV Type 3 (PAV-3) and cloning of restriction fragments representing the complete genome of PAV-3. See Reddy *et al.* (1993) *Intervirology* 36:161-168. In addition, restriction maps for PAV-1 and PAV-2 have been determined. See Reddy *et al.* (1995*b*) *Arch. Virol.* 140:195-200.

Nucleotide sequences have been determined for segments of the genome of various PAV serotypes. Sequences of the E3, pVIII and fiber genes of PAV-3 were determined by Reddy *et al.* (1995) *Virus Res.* 36:97-106. The E3, pVIII and fiber genes of PAV-1 and PAV-2 were sequenced by Reddy *et al.* (1996) *Virus Res.* 43:99-109, while the PAV-4 E3, pVIII and fiber gene sequences were determined by Kleiboeker (1994) *Virus Res.* 31:17-25. The PAV-4 fiber gene sequence was determined by Kleiboeker (1995) *Virus Res.* 39:299-309. Inverted terminal repeat (ITR) sequences for all five PAV serotypes (PAV-1 through PAV-5) were determined by Reddy *et al.* (1995) *Virology* 212:237-239. The PAV-3 penton sequence was determined by McCoy *et al.* (1996) *Arch. Virol.* 141:1367-1375. The nucleotide sequence of the E1 region of PAV-4 was determined by Kleiboeker (1995) *Virus Res.* 36:259-268. The sequence of the protease (23K) gene of PAV-3 was determined by McCoy *et al.* (1996) *DNA Seq.* 6:251-254. The sequence of the PAV-3 hexon gene (and the 14 N-terminal codons of the 23K protease gene) has been deposited in the GenBank database under accession No. U34592. The sequence of the PAV-3 100K

gene has been deposited in the GenBank database under accession No. U82628. The sequence of the PAV-3 E4 region has been determined by Reddy *et al.* (1997) *Virus Genes* 15:87-90. Vrati *et al.* (1995, *Virology*, 209:400-408) disclose sequences for ovine adenovirus.

At least 47 serotypes of human adenoviruses have been described. Reviews of the most common serotypes associated with particular diseases have been published. See for example, Foy H.M. (1989) *Adenoviruses* In Evans AS (ed). Viral Infections of Humans. New York, Plenum Publishing, pp 77-89 and Rubin B.A. (1993) *Clinical picture and epidemiology of adenovirus infections*, Acta Microbiol. Hung 40:303-323. The capsid of a human adenovirus demonstrates icosahedral symmetry and contains 252 capsomers. The capsomers consist of 240 hexons and 12 pentons with a projecting fiber on each of the pentons. The pentons and hexons are each derived from different viral polypeptides. The fibers, which are responsible for type-specific antibodies, vary in length among human strains. The hexons are group specific complement-fixing antibodies, whereas the pentons are especially active in hemgglutination (Plotkin and Orenstein, Vaccines, 3rd edition, W.B. Saunders Company Philadelphia, pp609-623). The fiber region assumes a homotrimeric conformation which is necessary for association of the mature fiber protein with the penton base in the formation of the adenovirus capsid. Fiber associates with penton base by virtue of non-covalent interactions between the amino terminus of the fiber trimer and a conserved domain within the penton base. It has been shown that the globular carboxyterminal knob domain of the adenovirus fiber protein is the ligand for attachment to the adenovirus primary cellular receptor (Krasnykh et al. (1996) Journal of Virology, 70:6839.). The distal, C-terminal domain of the trimeric fiber molecule terminates in a knob which binds with high affinity to a specific primary receptor. After binding, Arg-Gly-Asp (RGD) motifs in the penton base interact with cellular integrins of the $\alpha v\beta 3$ and $\alpha v\beta 5$ types which function as secondary receptors. This interaction triggers cellular internalization whereby the virion resides within the endosome. The endosome membrane is lysed in a process mediated by the penton base, releasing the contents of the endosome to the cytoplasm. During these processes, the virion is gradually uncoated and the adenovirus DNA is transported into the nucleus (Shayakhmetov *et al.* (2000) *Journal of Virology* 74:2567-2583).

For general background references regarding adenovirus and development of adenoviral vector systems, see Graham *et al.* (1973) *Virology* 52:456-467; Takiff *et al.* (1981) *Lancet* 11:832-834; Berkner *et al.* (1983) *Nucleic Acid Research* 11: 6003-6020; Graham (1984) *EMBO J* 3:2917-2922; Bett *et al.* (1993) *J. Virology* 67:5911-5921; and Bett *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:8802-8806.

Adenoviruses generally undergo a lytic replication cycle following infection of a host cell. In addition to lysing the infected cell, the replicative process of adenovirus blocks the transport and translation host cell mRNA, thus inhibiting cellular protein synthesis. For a review of adenoviruses and adenovirus replication, see Shenk, T. and Horwitz, M.S., *Virology*, third edition, Fields, B.N. *et al.*, eds., Raven Press Limited, New York (1996), Chapters 67 and 68, respectively.

The application of genetic engineering has resulted in several attempts to prepare adenovirus expression systems for obtaining vaccines. Examples of such research include the disclosures in U.S. Patent 4,510,245 of an adenovirus major late promoter for expression in a yeast host; U.S. Patent 4,920,209 on a live recombinant adenovirus type 7 with a gene coding for hepatitis-B surface antigen located at a deleted early region 3; European Patent 389 286 on a non-defective human adenovirus 5 recombinant expression system in human cells for HCMV major envelope glycoprotein; WO 91/11525 on live non-pathogenic immunogenic viable canine adenovirus in a cell expressing E1A proteins; and French Patent 2 642 767 on vectors containing a leader and/or promoter from the E3 region of adenovirus 2. United States Patent Numbers 6,001,591 and 5,820,868 and International Publication Number WO 95/16048 disclose recombinant protein production in bovine adenovirus expression vector systems. United States Patent Number 5,922,576 discloses systems for generating recombinant adenoviruses.

Krasnykh *et al.* (1996, *Journal of Virology*, 70:6839), Zabner *et al.* (1999) *Journal of Virology* ,73:8689), and Shayakhmetov *et al. supra* report generation of human adenovirus vectors with modified fiber regions. Xu *et al.* (1998, *Virology*, 248:156-163) disclose an ovine adenovirus carrying the fiber protein cell binding domain of human Adenovirus Type 5.

The disclosure of all patents and publications cited herein are incorporated by reference in their entirety.

DISCLOSURE OF THE INVENTION

The present invention provides adenoviruses, preferably bovine adenoviruses, comprising a modification in a polynucleotide encoding a capsid protein, or fragment thereof, wherein said protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism. The present invention further provides host cells and methods comprising the modified adenoviruses. Accordingly, the present invention provides bovine adenovirus vectors comprising a modification in a polynucleotide encoding a capsid protein, or fragment thereof, wherein said protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism. In some embodiments, the polynucleotide encoding a capsid protein, or fragment thereof, is replaced with a polynucleotide encoding a heterologous mammalian capsid protein, or fragment thereof. The capsid protein, or fragment thereof, includes adenovirus penton, hexon or fiber proteins, or fragments thereof. In some embodiments, the modification is in a polynucleotide encoding the knob region of a fiber protein. In other embodiments, a polynucleotide encoding a bovine adenovirus penton, hexon and/or fiber protein(s) is replaced with at least one polynucleotide encoding a heterologous mammalian adenovirus penton, hexon and/or fiber protein(s), respectively. In additional embodiments, a polynucleotide encoding a bovine adenovirus penton protein, or fragment thereof, is replaced with at least one polynucleotide encoding a heterologous mammalian adenovirus penton protein, or fragment thereof; a polynucleotide encoding a bovine adenovirus hexon protein, or fragment thereof, is replaced with at least one polynucleotide encoding a heterologous mammalian adenovirus hexon protein, or fragment thereof; or a polynucleotide encoding a bovine adenovirus fiber protein, or fragment thereof, such as a knob region, is replaced with at least one polynucleotide encoding a heterologous mammalian adenovirus fiber protein, or fragment thereof, such as a heterologous knob region of a fiber protein.

In further embodiments, heterologous mammalian adenoviruses include bovine, porcine, ovine, canine or human adenovirus. In additional embodiments, bovine adenoviruses include sub-type 1 adenovirus, and in particular BAV3, or sub-type 2 adenovirus. In other embodiments, the bovine adenovirus vector further comprises a polynucleotide encoding a heterologous protein. In some embodiments, the heterologous

protein is a therapeutic protein. In other embodiments, the heterologous protein includes cytokines; lymphokines; membrane receptors recognized by pathogenic organisms, dystrophins; insulin; proteins participating in cellular ion channels; antisense RNAs; proteins capable of inhibiting the activity of a protein produced by a pathogenic gene, a protein inhibiting an enzyme activity, protein variants of pathogenic proteins; antigenic epitopes; major histocompatibility complex classes I and II proteins; antibodies; immunotoxins; toxins; growth factors or growth hormones; cell receptors or their ligands; tumor suppressors; cellular enzymes; or suicide genes. In yet other embodiments, an adenovirus vector lacks E1 function. In additional embodiments, an adenovirus vector has a deletion in part or all of the E1 gene region. In further embodiments, the adenovirus vector has a deletion of part or all of the E3 gene region. In yet further embodiments, a polynucleotide encoding a heterologous protein is inserted in the adenovirus E1 gene region. In other embodiments, a polynucleotide encoding a heterologous protein is inserted in the adenovirus E3 gene region. In further embodiments, an adenovirus vector is replication-defective, and in yet further embodiments, an adenovirus vector is replication-competent. The present invention also encompasses host cells comprising a bovine adenovirus vector having a modification in a polynucleotide encoding a capsid protein, or fragment thereof.

The present invention also provides methods of producing a recombinant bovine adenovirus vector comprising a modification in a polynucleotide encoding a capsid protein, or fragment thereof, comprising the steps of, obtaining a bovine adenovirus vector; and introducing a modification into a polynucleotide encoding a capsid protein, or fragment thereof, wherein said capsid protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism. In some embodiments, the modification is a replacement of at least one polynucleotide encoding a bovine adenovirus penton, hexon and/or fiber protein, or fragment thereof, with a heterologous mammalian penton, hexon and/or fiber protein, or fragment thereof. In other embodiments, the modification is a replacement of a polynucleotide encoding a knob region of a fiber protein. In further embodiments, the adenovirus vector further comprises a polynucleotide encoding a heterologous protein.

The present invention further provides recombinant bovine adenoviruses comprising a modification in a capsid protein, or fragment thereof, wherein said capsid

protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism. In further embodiments, recombinant adenoviruses comprise polynucleotides encoding a heterologous protein. In further embodiments, a polynucleotide encoding a heterologous protein is inserted in the adenovirus E1 gene region; in yet further embodiments, a polynucleotide encoding a heterologous protein is inserted in the adenovirus E3 gene region. In some embodiments, a recombinant adenovirus is replication-competent and in other embodiments, a recombinant adenovirus is replication-defective. In some embodiments, a recombinant adenovirus comprises a replacement of at least one polynucleotide encoding a bovine adenovirus penton, hexon and/or fiber protein(s), or fragment thereof, with a heterologous mammalian penton, hexon and/or fiber protein(s), or fragment thereof. In yet further embodiments, a recombinant adenovirus comprises a modification in a knob region of a fiber protein.

The present invention also provides immunogenic compositions comprising a bovine adenovirus wherein said adenovirus comprises a polynucleotide encoding a modification in a capsid protein, or fragment thereof, and wherein said protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism In some embodiments, the capsid protein, or fragment thereof, includes penton, hexon or fiber protein(s), or a fragment thereof, of an adenovirus. In some embodiments of immunogenic compositions, the modification comprises a replacement of a polynucleotide encoding a bovine capsid protein, or fragment thereof, with a polynucleotide encoding a heterologous mammalian adenovirus capsid protein, or fragment thereof. In other embodiments of immunogenic compositions, the modification comprises a replacement of a polynucleotide encoding a bovine knob region of a fiber protein with a polynucleotide encoding a heterologous mammalian adenovirus knob region of a fiber protein. In other embodiments, the bovine adenovirus is a sub-type 1 adenovirus, in particular, BAV3, or a sub-type 2 adenovirus. In additional embodiments, immunogenic compositions comprise a bovine adenovirus comprises a polynucleotide encoding a heterologous protein. In other embodiments, immunogenic compositions comprise a bovine adenovirus comprising a polynucleotide encoding cytokines; lymphokines; membrane receptors recognized by pathogenic organisms, dystrophins; insulin; proteins participating in cellular ion channels; antisense RNAs; proteins capable of inhibiting the activity of a protein produced by a pathogenic gene, a protein inhibiting an enzyme activity, protein variants of pathogenic

proteins; antigenic epitopes; major histocompatibility complex classes I and II proteins; antibodies; immunotoxins; toxins; growth factors or growth hormones; cell receptors or their ligands; tumor suppressors; cellular enzymes; or suicide genes.

The present invention also encompasses pharmaceutical compositions capable of inducing an immune response in a mammalian subject. In some embodiments, pharmaceutical compositions comprise an immunogenic composition comprising a bovine adenovirus having a modified capsid protein, or fragment thereof, wherein the protein, or fragment thereof, is associated with tropism and wherein the modification is associated with altered tropism. In some embodiments of the pharmaceutical compositions, immunogenic compositions comprise bovine adenovirus vectors comprising a polynucleotide encoding a heterologous protein. In some embodiments, the heterologous protein is a therapeutic protein. In other embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

The present invention also provides methods for eliciting an immune response in a mammalian host to protect against infection, the method comprising administering a pharmaceutical composition of the present invention to a mammalian host in need. The present invention also provides methods of gene delivery in a mammalian host, the methods comprising administering to the host a bovine adenovirus vector comprising a polynucleotide encoding a modified capsid protein, or fragment thereof, wherein the protein is associated with tropism and wherein the modification is associated with altered tropism and wherein the adenovirus vector further comprises a polynucleotide encoding a heterologous protein. In some embodiments, the heterologous polynucleotide encodes a therapeutic protein

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A-1S shows the complete nucleotide sequence of the BAV3 genome. In the polynucleotide sequence for BAV3, the penton regions starts at 12919 and ends at 14367; the hexon region starts at 17809 and ends at 20517; the fiber region starts at 27968 and ends at 30898. The knob domain of the fiber region starts after the 4 residues, TLWT, as shown in Figure 4.

Figure 2 shows a transcriptional map of the BAV3 genome, derived from transcriptional mapping of mRNAs and sequencing of cDNA clones.

Figure 3 illustrates the construction of BAV600 that expresses the HAV-5 fiber knob protein.

Figure 4 illustrates the characterization of BAV600.

Figures 5A-5B shows the analysis of BAV600 by Restriction Enzyme *BglII* digestion. Figure 5A depicts a gel electrophoresis and Figure 5B depicts a Southern Blot.

Figure 6 shows the expression of HAV-5 fiber Knob by BAV600.

Figures 7A-7B show the transduction of Human cell lines by BAV600. Figure 7A show results of an MOI of 1 whereas Figure 7B shows results of an MOI of 5.

Figure 8 shows a FACS analysis of BAV304 and BAV600 transduction of Human cells.

Figure 9 shows the expression of early and late BAV-3 proteins in human cell lines, HeLa, HEp-2, A549, 293 and MDBK.

Figure 10 illustrates BAV3 replication in human cells.

Figure 11 shows the neutralization of BAV600 by a monoclonal antibody specific for HAV-5 fiber knob region.

Figure 12 depicts the amino acid sequence for Human adenovirus 5 (HAV-5) fiber protein.

Figure 13 depicts the amino acid sequence for the Bovine Adenovirus-3 (BAV-3) fiber protein.

Figure 14 depicts the amino acid sequence of Ovine Adenovirus 287 fiber protein.

Figure 15 shows the amino acid sequence of Porcine Adenovirus-3 (PAV-3) fiber protein.

Figure 16 shows the amino acid sequence of Canine Adenovirus -2 (CAV-2) fiber protein.

Figures 17A-17G depicts an amino acid alignment of various mammalian adenovirus fiber regions using the clustal method of the Multialign program.

BEST MODE FOR CARRYING OUT THE INVENTION

We have discovered and constructed improved adenovirus vectors, in particular improved bovine adenovirus vectors, having altered tropism. The bovine adenovirus vectors of the present invention comprise a modification in a polynucleotide encoding at

least one capsid protein, wherein the protein, or fragment thereof, is associated with tropism and wherein the modification is associated with altered tropism.

Capsid proteins include penton, hexon and fiber proteins. In one embodiment illustrated herein, a BAV3 adenovirus vector was constructed, BAV600, which comprised a replacement of the BAV3 fiber knob region with a human adenovirus (Ad5) fiber knob region. BAV600 demonstrated increased transduction in human cell lines as compared to a control adenovirus.

The present invention encompasses bovine adenovirus vectors comprising a replacement of a capsid protein, or fragment thereof, with a heterologous mammalian capsid protein, or fragment thereof, as long as the protein is associated with tropism and the replacement is associated with altered tropism. For example, in one embodiment, a bovine knob domain of a fiber protein is replaced with a porcine or ovine knob region of a fiber protein in order to alter species tropism. Such a bovine adenovirus vector can be used as an immunogen to boost immunity in a porcine or ovine mammal that has been primed with a porcine or ovine adenovirus, respectively. In such an immunization protocol, a boost immunization is achieved by administration of the bovine adenovirus having species specificity for the porcine or ovine mammal, while avoiding the affect of any neutralizing antibodies against the porcine or ovine mammal produced as a result of the priming immunization. Alternatively, in another embodiment, a bovine fiber protein, or fragment thereof, such as the knob region, is replaced with a heterologous bovine fiber protein, or fragment thereof, such as a knob region of a fiber protein in order to alter bovine cell specificity. For one example, a bovine adenovirus sub-type 1 fiber region, or fragment thereof, such as a knob domain, is replaced with a bovine adenovirus sub-type 2 fiber region, or fragment thereof, such as a knob domain, in order to alter bovine cell-type specificity. Such a bovine adenovirus vector can be used as an immunogen to target specific cells or tissues.

The invention also encompasses the use of a bovine adenovirus comprising a replacement of a bovine capsid protein, or fragment thereof, with a human adenovirus capsid protein, or fragment thereof, such that the modified bovine adenovirus has species specificity for humans. Such bovine adenoviruses can be used in human immunization protocols, where preexisting neutralizing antibodies against human adenovirus -5 (HAV-5) in clinical patients may present an obstacle for efficient use of HAV-5.

10

Additionally, to provide a therapeutic effect to target cells, one or more heterologous therapeutic proteins may be present in the adenovirus vector.

Definitions

In describing the present invention, the following terminology, as defined below, will be used.

An "adenovirus vector" or "adenoviral vector" (used interchangeably) comprises a polynucleotide construct of the invention. A polynucleotide construct of this invention may be in any of several forms, including, but not limited to, DNA, DNA encapsulated in an adenovirus coat, DNA packaged in another viral or viral-like form (such as herpes simplex, and AAV), DNA encapsulated in liposomes, DNA complexed with polylysine, complexed with synthetic polycationic molecules, conjugated with transferrin, and complexed with compounds such as PEG to immunologically "mask" the molecule and/or increase half-life, and conjugated to a nonviral protein. Preferably, the polynucleotide is DNA. As used herein, "DNA" includes not only bases A, T, C, and G, but also includes any of their analogs or modified forms of these bases, such as methylated nucleotides, internucleotide modifications such as uncharged linkages and thioates, use of sugar analogs, and modified and/or alternative backbone structures, such as polyamides. Adenovirus vectors may be replication-competent or replication-defective in a target cell.

As used herein, the term "altered tropism" refers to changing the specificity of an adenovirus. The term "altered tropism" encompasses changing species specificity as well as changing tissue or cell specificity of an adenovirus. In embodiments illustrated herein, species specificity is altered by producing modifications in a capsid protein(s), or fragment thereof, such as the fiber protein, and in particular the knob region of a fiber protein.

A "capsid protein" as used herein includes penton, hexon and fiber regions of an adenovirus. A capsid protein is associated with tropism if it directly or indirectly affects adenovirus tropism. A "modification of a capsid protein associated with altered tropism" as used herein refers to producing an alteration of a polynucleotide encoding a capsid protein, ie, a penton, hexon or fiber protein region, or fragment thereof, such as the knob domain of the fiber region such that specificity is altered. "Associated with" means that the modification contributes to the altered tropism either directly or indirectly. In embodiments illustrated herein, the modification is a replacement of bovine capsid protein regions with a heterologous mammalian capsid protein region in order to produce species

specificity in the adenovirus. Replacement of one species capsid protein region with a heterologous capsid protein region may also produce altered tissue or cell specificity.

A "replicon" is any genetic element (e.g., plasmid, chromosome, virus) that functions as an autonomous unit of DNA replication *in vivo*; i.e., is capable of replication under its own control.

As used herein, the term "vector" refers to a polynucleotide construct designed for transduction/transfection of one or more cell types. Vectors may be, for example, "cloning vectors" which are designed for isolation, propagation and replication of inserted nucleotides, "expression vectors" which are designed for expression of a nucleotide sequence in a host cell, or a "viral vector" which is designed to result in the production of a recombinant virus or virus-like particle, or "shuttle vectors", which comprise the attributes of more than one type of vector.

By "live virus" is meant, in contradistinction to "killed" virus, a virus which is capable of producing identical progeny in tissue culture and inoculated animals.

A "helper-free virus vector" is a vector that does not require a second virus or a cell line to supply something defective in the vector.

A "double-stranded DNA molecule" refers to the polymeric form of deoxyribonucleotides (adenine, guanine, thymine, or cytosine) in its normal, double-stranded helix. This term refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary forms. Thus, this term includes double-stranded DNA found, inter alia, in linear DNA molecules (e.g., restriction fragments of DNA from viruses, plasmids, and chromosomes). In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5' to 3' direction along the nontranscribed strand of DNA (i.e., the strand having the sequence homologous to the mRNA).

A DNA "coding sequence" is a DNA sequence which is transcribed and translated into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at the 3' (carboxy) terminus. A coding sequence can include, but is not limited to, procaryotic sequences, cDNA from eucaryotic mRNA, genomic DNA sequences from eucaryotic (e.g., mammalian) DNA, viral DNA,

and even synthetic DNA sequences. A polyadenylation signal and transcription termination sequence will usually be located 3' to the coding sequence.

A "transcriptional promoter sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. For purposes of defining the present invention, the promoter sequence is bound at the 3' terminus by the translation start codon (ATG) of a coding sequence and extends upstream (5' direction) to include the minimum number of bases or elements necessary to initiate transcription at levels detectable above background. Within the promoter sequence will be found a transcription initiation site (conveniently defined by mapping with nuclease S1), as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase. Eucaryotic promoters will often, but not always, contain "TATA" boxes and "CAAT" boxes. Procaryotic promoters contain Shine-Dalgarno sequences in addition to the -10 and -35 consensus sequences.

DNA "control sequences" refer collectively to promoter sequences, ribosome binding sites, splicing signals, polyadenylation signals, transcription termination sequences, upstream regulatory domains, enhancers, translational termination sequences and the like, which collectively provide for the transcription and translation of a coding sequence in a host cell.

A coding sequence or sequence encoding a protein is "operably linked to" or "under the control of" control sequences in a cell when RNA polymerase will bind the promoter sequence and transcribe the coding sequence into mRNA, which is then translated into the polypeptide encoded by the coding sequence.

A "host cell" is a cell which has been transformed, or is capable of transformation, by an exogenous DNA sequence.

A cell has been "transformed" by exogenous DNA when such exogenous DNA has been introduced inside the cell membrane. Exogenous DNA may or may not be integrated (covalently linked) to chromosomal DNA making up the genome of the cell. In procaryotes and yeasts, for example, the exogenous DNA may be maintained on an episomal element, such as a plasmid. A stably transformed cell is one in which the exogenous DNA has become integrated into the chromosome so that it is inherited by daughter cells through chromosome replication. For mammalian cells, this stability is

demonstrated by the ability of the cell to establish cell lines or clones comprised of a population of daughter cell containing the exogenous DNA.

A "clone" is a population of daughter cells derived from a single cell or common ancestor. A "cell line" is a clone of a primary cell that is capable of stable growth *in vitro* for many generations.

A "heterologous" region of a DNA construct is an identifiable segment of DNA within or attached to another DNA molecule that is not found in association with the other molecule in nature. Thus, when the heterologous region encodes a viral gene, the gene will usually be flanked by DNA that does not flank the viral gene in the genome of the source virus or virus-infected cells. Another example of the heterologous coding sequence is a construct wherein the coding sequence itself is not found in nature (e.g., synthetic sequences having codons different from the native gene). Allelic variation or naturally occurring mutational events do not give rise to a heterologous region of DNA, as used herein. As used herein in describing adenovirus vectors, "heterologous mammalian capsid region" means that the capsid region is obtainable from another mammalian species of adenovirus or is obtainable from the same species mammal but from a different type or sub-type adenovirus. For example "heterologous mammalian capsid protein" encompasses replacement of one sub-type bovine adenovirus capsid protein with another sub-type bovine adenovirus capsid protein as well as replacement of a bovine adenovirus capsid protein with another species capsid protein, such as a human capsid protein, as well as replacement of bovine adenovirus capsid proteins regions with another serotype bovine adenovirus capsid protein.

"Bovine host" refers to cattle of any breed, adult or infant.

The term "protein" is used herein to designate a polypeptide or glycosylated polypeptide, respectively, unless otherwise noted. The term "polypeptide" is used in its broadest sense, i.e., any polymer of amino acids (dipeptide or greater) linked through peptide bonds. Thus, the term "polypeptide" includes proteins, oligopeptides, protein fragments, analogs, muteins, fusion proteins and the like.

"Native" proteins or polypeptides refer to proteins or polypeptides recovered from adenovirus or adenovirus-infected cells. Thus, the term "native BAV polypeptide" would include naturally occurring BAV proteins and fragments thereof. "Non-native" polypeptides refer to polypeptides that have been produced by recombinant DNA methods

or by direct synthesis. "Recombinant" polypeptides refers to polypeptides produced by recombinant DNA techniques; i.e., produced from cells transformed by an exogenous DNA construct encoding the desired polypeptide.

A "substantially pure" protein will be free of other proteins, preferably at least 10% homogeneous, more preferably 60% homogeneous, and most preferably 95% homogeneous.

An "antigen" refers to a molecule containing one or more epitopes that will stimulate a host's immune system to make a humoral and/or cellular antigen-specific response. The term is also used interchangeably with "immunogen."

A "hapten" is a molecule containing one or more epitopes that does not stimulate a host's immune system to make a humoral or cellular response unless linked to a carrier.

The term "epitope" refers to the site on an antigen or hapten to which a specific antibody molecule binds or is recognized by T cells. The term is also used interchangeably with "antigenic determinant" or "antigenic determinant site."

An "immunological response" to a composition or vaccine is the development in the host of a cellular and/or antibody-mediated immune response to the composition or vaccine of interest. Usually, such a response consists of the subject producing antibodies, B cells, helper T cells, suppressor T cells, and/or cytotoxic T cells directed specifically to an antigen or antigens included in the composition or vaccine of interest.

The terms "immunogenic polypeptide" and "immunogenic amino acid sequence" and "immunogen" refer to a polypeptide or amino acid sequence, respectively, which elicit antibodies that neutralize viral infectivity, and/or mediate antibody-complement or antibody-dependent cell cytotoxicity to provide protection of an immunized host. An "immunogenic polypeptide" as used herein, includes the full length (or near full length) sequence of the desired protein or an immunogenic fragment thereof.

By "immunogenic fragment" is meant a fragment of a polypeptide which includes one or more epitopes and thus elicits antibodies that neutralize viral infectivity, and/or mediates antibody-complement or antibody-dependent cell cytotoxicity to provide protection of an immunized host. Such fragments will usually be at least about 5 amino acids in length, and preferably at least about 10 to 15 amino acids in length. There is no critical upper limit to the length of the fragment, which could comprise nearly the full length of the protein sequence, or even a fusion protein comprising fragments of two or

more of the antigens. The term "treatment" as used herein refers to treatment of a mammal, such as bovine or human or other mammal, either (i) the prevention of infection or reinfection (prophylaxis), or (ii) the reduction or elimination of symptoms of an infection. The vaccine comprises the recombinant BAV itself or recombinant antigen produced by recombinant BAV.

By "infectious" is meant having the capacity to deliver the viral genome into cells.

The terms "polynucleotide" and "nucleic acid", used interchangeably herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxyribonucleotides. These terms include a single-, double- or triple-stranded DNA, genomic DNA, cDNA, RNA, DNA-RNA hybrid, or a polymer comprising purine and pyrimidine bases, or other natural, chemically, biochemically modified, non-natural or derivatized nucleotide bases. The backbone of the polynucleotide can comprise sugars and phosphate groups (as may typically be found in RNA or DNA), or modified or substituted sugar or phosphate groups. Alternatively, the backbone of the polynucleotide can comprise a polymer of synthetic subunits such as phosphoramidates and thus can be a oligodeoxynucleoside phosphoramidate (P-NH2) or a mixed phosphoramidate-phosphodiester oligomer. Peyrottes et al. (1996) *Nucleic Acids Res.* 24: 1841-8; Chaturvedi et al. (1996) *Nucleic Acids Res.* 24: 2318-23; Schultz et al. (1996) *Nucleic Acids Res.* 24: 2966-73. A phosphorothioate linkage can be used in place of a phosphodiester linkage. Braun et al. (1988) *J. Immunol.* 141: 2084-9; Latimer et al. (1995) *Molec. Immunol.* 32: 1057-1064. In addition, a double-stranded polynucleotide can be obtained from the single stranded polynucleotide product of chemical synthesis either by synthesizing the complementary strand and annealing the strands under appropriate conditions, or by synthesizing the complementary strand de novo using a DNA polymerase with an appropriate primer. Reference to a polynucleotide sequence (such as referring to a SEQ ID NO) also includes the complement sequence.

The following are non-limiting examples of polynucleotides: a gene or gene fragment, exons, introns, mRNA, tRNA, rRNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs, uracyl, other sugars and linking groups such as fluororibose and

thioate, and nucleotide branches. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Other types of modifications included in this definition are caps, substitution of one or more of the naturally occurring nucleotides with an analog, and introduction of means for attaching the polynucleotide to proteins, metal ions, labeling components, other polynucleotides, or a solid support. Preferably, the polynucleotide is DNA. As used herein, "DNA" includes not only bases A, T, C, and G, but also includes any of their analogs or modified forms of these bases, such as methylated nucleotides, internucleotide modifications such as uncharged linkages and thioates, use of sugar analogs, and modified and/or alternative backbone structures, such as polyamides.

A polynucleotide or polynucleotide region has a certain percentage (for example, 80%, 85%, 90%, or 95%) of "sequence identity" to another sequence means that, when aligned, that percentage of bases are the same in comparing the two sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art, for example those described in *Current Protocols in Molecular Biology* (F.M. Ausubel et al., eds., 1987) Supplement 30, section 7.7.18, Table 7.7.1. A preferred alignment program is ALIGN Plus (Scientific and Educational Software, Pennsylvania), preferably using default parameters, which are as follows: mismatch = 2; open gap = 0; extend gap = 2.

"Under transcriptional control" is a term well understood in the art and indicates that transcription of a polynucleotide sequence, usually a DNA sequence, depends on its being operably (operatively) linked to an element which contributes to the initiation of, or promotes, transcription. "Operably linked" refers to a juxtaposition wherein the elements are in an arrangement allowing them to function.

adenovirus. Preferably, the transgene will also not be expressed or present in the target cell prior to introduction by the adenovirus vector.

In the context of adenovirus, a "heterologous" promoter or enhancer is one which is not associated with or derived from an adenovirus gene.

In the context of adenovirus, an "endogenous" promoter, enhancer, or control region is native to or derived from adenovirus.

A "host cell" includes an individual cell or cell culture which can be or has been a recipient of an adenoviral vector(s) of this invention. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology or in total DNA complement) to the original parent cell due to natural, accidental, or deliberate mutation and/or change. A host cell includes cells transfected or infected *in vivo* or *in vitro* with an adenoviral vector of this invention.

"Replication" and "propagation" are used interchangeably and refer to the ability of an adenovirus vector of the invention to reproduce or proliferate. These terms are well understood in the art. For purposes of this invention, replication involves production of adenovirus proteins and is generally directed to reproduction of adenovirus. Replication can be measured using assays standard in the art and described herein, such as a burst assay or plaque assay. "Replication" and "propagation" include any activity directly or indirectly involved in the process of virus manufacture, including, but not limited to, viral gene expression; production of viral proteins, nucleic acids or other components; packaging of viral components into complete viruses; and cell lysis.

A polynucleotide sequence that is "depicted in"a SEQ ID NO means that the sequence is present as an identical contiguous sequence in the SEQ ID NO. The term encompasses portions, or regions of the SEQ ID NO as well as the entire sequence contained within the SEQ ID NO.

A "biological sample" encompasses a variety of sample types obtained from an individual and can be used in a diagnostic or monitoring assay. The definition encompasses blood and other liquid samples of biological origin, solid tissue samples such as a biopsy specimen or tissue cultures or cells derived therefrom, and the progeny thereof. The definition also includes samples that have been manipulated in any way after their procurement, such as by treatment with reagents, solubilization, or enrichment for certain components, such as proteins or polynucleotides. The term "biological sample"

encompasses a clinical sample, and also includes cells in culture, cell supernatants, cell lysates, serum, plasma, biological fluid, and tissue samples.

An "individual" or "mammalian subject" is a vertebrate, preferably a mammal, more preferably a human. Mammals include, but are not limited to, farm animals, sport animals, rodents, primates, and pets.

An "effective amount" is an amount sufficient to effect beneficial or desired results, including clinical results. An effective amount can be administered in one or more administrations. For purposes of this invention, an effective amount of an adenoviral vector is an amount that is sufficient to palliate, ameliorate, stabilize, reverse, slow or delay the progression of the disease state.

"Expression" includes transcription and/or translation.

As used herein, the term "comprising" and its cognates are used in their inclusive sense; that is, equivalent to the term "including" and its corresponding cognates.

"A," "an" and "the" include plural references unless the context clearly dictates otherwise.

Detailed Description

The present invention identifies capsid proteins associated with tropism and provides methods of constructing adenovirus vectors and recombinant adenoviruses having altered tropism. In preferred embodiments, the adenovirus is a bovine adenovirus, such as a sub-type 1 adenovirus, in particular BAV3, or a sub-type 2 adenovirus. In illustrative embodiments, part or all of a bovine capsid protein encoding polynucleotide sequence associated with tropism is deleted and replaced with part or all of a heterologous mammalian capsid protein encoding polynucleotide sequence which alters adenovirus tropism. In a particular embodiment disclosed herein, the knob region of a bovine fiber protein is replaced with a human knob region of a fiber protein. The present invention also encompasses adenoviruses comprising the replacement of one bovine serotype adenovirus capsid protein associated with tropism with a heterologous bovine serotype adenovirus capsid protein associated with tropism in order to alter cell specificity.

The complete nucleotide sequence of the BAV3 genome is disclosed herein. *See* Figure 1 (SEQ ID NO 1). A transcriptional map of the BAV3 genome, derived from transcriptional mapping of mRNAs and sequencing of cDNA clones, is presented in Figure 2. Although the size (34,446 bp) and the overall organization of the BAV3 genome appear

to be similar to that of HAVs, there are certain differences. Reddy *et al.* (1998) *supra*. One of the distinctive features of the BAV3 genome is the relatively small size of the E3 coding region (1517 bp). Mittal *et al.* (1992) *J. Gen. Virol.* **73**:3295-3300; Mittal *et al.* (1993). *J. Gen. Virol.* **74**:2825; and Reddy *et al.* (1998) *supra*. Analysis of the sequence of the BAV3 E3 region and its RNA transcripts suggests that BAV3 E3 may encode at least four proteins, one of which (121R) exhibits limited homology with the 14.7 kDa protein of HAV5. Idamakanti (1998) "Molecular characterization of E3 region of bovine adenovirus-3," M.Sc. thesis, University of Saskatchewan, Saskatoon, Saskatchewan.

Reddy *et al.* (1998) *Journal of Virology* 72:1394) disclose nucleotide sequences for BAV3. In the polynucleotide sequence for BAV3, the penton regions starts at 12919 and ends at 14367; the hexon region starts at 17809 and ends at 20517; the fiber region starts at 27968 and ends at 30898. The knob region (or domain) of the fiber protein starts after the residues TLWT motif as shown in Figure 4. The fiber protein also contains shaft and tail regions (or domains).

Human adenoviruses Ad3, Ad4, Ad5, Ad9 and Ad35 are available from the American Tissue Culture Collection ATCC). The National Center for Biotechnology Information GenBank accession number for Ad5 is M73260/M29978; for Ad9 X74659; and for Ad35, U10272. Chow *et al.* (1977, *Cell* 12:1-8) disclose human adenovirus 2 sequences; Davison *et al.* (1993, *J. Mole. Biol.* 234:1308-1316) disclose the DNA sequence of human adenovirus type 40; Sprengel *et al.* (1994, *J. Virol.* 68:379-389) disclose the DNA sequence for human adenovirus type 12 DNA; Vrati *et al.* (1995, *Virology*, 209:400-408) disclose sequences for ovine adenovirus; Morrison *et al.* (1997, *J. Gen. Virol.* 78:873-878) disclose canine adenovirus type 1 DNA sequence; and Reddy *et al.* (1998, *Virology*, 251:414) disclose DNA sequences for porcine adenovirus.

Shayakhmetov *et al., supra*, provide PCR primers for human Ad9 and human Ad35 fiber regions. The HAV-5 fiber protein is depicted in Figure 12; Figure 13 depicts the amino acid sequence for the Bovine Adenovirus-3 (BAV-3) fiber protein; Figure 14 depicts the amino acid sequence of Ovine Adenovirus 287 fiber protein; Figure 15 depicts the amino acid sequence of Porcine Adenovirus-3 (PAV-3)fiber protein; Figure 16 depicts the amino acid sequence of Canine Adenovirus -2 (CAV-2) fiber protein; and Figures 17A-17G depicts an amino acid alignment of mammalian adenovirus fiber regions using the clustal method of the multialign program. The knob domain of the fiber regions typically

starts after the amino acid residue motif TLWT (hinge region), see Figure 4 (one exception is the ovine adenovirus fiber region).

Adenovirus vector constructs can then undergo recombination *in vitro* or *in vivo*, with a BAV genome either before or after transformation or transfection of an appropriate host cell.

Suitable host cells include any cell that will support recombination between a BAV genome and a plasmid containing BAV sequences, or between two or more plasmids, each containing BAV sequences. Recombination is generally performed in procaryotic cells, such as *E. coli*, while transfection of a plasmid containing a viral genome, to generate virus particles, is conducted in eukaryotic cells, preferably mammalian cells, more preferably bovine cell cultures, most preferably MDBK or PFBR cells, and their equivalents. The growth of bacterial cell cultures, as well as culture and maintenance of eukaryotic cells and mammalian cell lines are procedures which are well-known to those of skill in the art.

One or more heterologous polynucleotide sequences can be inserted into one or more regions of the BAV genome to generate a recombinant BAV, limited only by the insertion capacity of the BAV genome and ability of the recombinant BAV to express the inserted heterologous sequences. In general, adenovirus genomes can accept inserts of approximately 5% of genome length and remain capable of being packaged into virus particles. The insertion capacity can be increased by deletion of non-essential regions and/or deletion of essential regions, such as, for example, E1 function, whose function is provided by a helper cell line, such as one providing E1 function. In some embodiments, a heterologous polynucleotide encoding a protein is inserted into an adenovirus E1 gene region. In some embodiments, an adenovirus has a deletion of part or all of the E1 gene region and is propagated in a helper cell line providing E1 function. In yet other embodiments, a heterologous polynucleotide encoding a protein is inserted into an adenovirus E3 gene region. In other embodiments, an adenovirus has a deletion of part or all of the E3 region.

In one embodiment of the invention, insertion can be achieved by constructing a plasmid containing the region of the BAV genome into which insertion is desired, such as a polynucleotide encoding a capsid protein. Additionally, a polynucleotide encoding a desired therapeutic protein can be inserted into the bovine adenovirus. The plasmid is then digested with a restriction enzyme having a recognition sequence in the BAV portion of the

plasmid, and a heterologous polynucleotide sequence is inserted at the site of restriction digestion. The plasmid, containing a portion of the BAV genome with an inserted heterologous sequence, is co-transformed, along with a BAV genome or a linearized plasmid containing a BAV genome, into a bacterial cell (such as, for example, *E. coli*), wherein the BAV genome can be a full-length genome or can contain one or more deletions. Homologous recombination between the plasmids generates a recombinant BAV genome containing inserted heterologous sequences.

Deletion of BAV sequences, to provide a site for insertion of heterologous sequences or to provide additional capacity for insertion at a different site, can be accomplished by methods well-known to those of skill in the art. For example, for BAV sequences cloned in a plasmid, digestion with one or more restriction enzymes (with at least one recognition sequence in the BAV insert) followed by ligation will, in some cases, result in deletion of sequences between the restriction enzyme recognition sites. Alternatively, digestion at a single restriction enzyme recognition site within the BAV insert, followed by exonuclease treatment, followed by ligation will result in deletion of BAV sequences adjacent to the restriction site. A plasmid containing one or more portions of the BAV genome with one or more deletions, constructed as described above, can be co-transfected into a bacterial cell along with a BAV genome (full-length or deleted) or a plasmid containing either a full-length or a deleted BAV genome to generate, by homologous recombination, a plasmid containing a recombinant BAV genome with a deletion at one or more specific sites. BAV virions containing the deletion can then be obtained by transfection of mammalian cells (including, but not limited to, MDBK or PFBR cells and their equivalents) with the plasmid containing the recombinant BAV genome.

In one embodiment of the invention, insertion sites are adjacent to and downstream (in the transcriptional sense) of BAV promoters. Locations of BAV promoters, and restriction enzyme recognition sequences downstream of BAV promoters, for use as insertion sites, can be easily determined by one of skill in the art from the BAV nucleotide sequence provided herein. Alternatively, various *in vitro* techniques can be used for insertion of a restriction enzyme recognition sequence at a particular site, or for insertion of heterologous sequences at a site that does not contain a restriction enzyme recognition sequence. Such methods include, but are not limited to, oligonucleotide-mediated

heteroduplex formation for insertion of one or more restriction enzyme recognition sequences (*see*, for example, Zoller *et al.* (1982) *Nucleic Acids Res.* **10:**6487-6500; Brennan *et al.* (1990) *Roux's Arch. Dev. Biol.* **199:**89-96; and Kunkel *et al.* (1987) *Meth. Enzymology* **154:**367-382) and PCR-mediated methods for insertion of longer sequences. *See*, for example, Zheng *et al.* (1994) *Virus Research* **31:**163-186.

It is also possible to obtain expression of a heterologous sequence inserted at a site that is not downstream from a BAV promoter, if the heterologous sequence additionally comprises transcriptional regulatory sequences that are active in eukaryotic cells. Such transcriptional regulatory sequences can include cellular promoters such as, for example, the bovine hsp70 promoter and viral promoters such as, for example, herpesvirus, adenovirus and papovavirus promoters and DNA copies of retroviral long terminal repeat (LTR) sequences.

In another embodiment, homologous recombination in a procaryotic cell can be used to generate a cloned BAV genome; and the cloned BAV genome can be propagated as a plasmid. See for example, U.S. patent 5,922,576. Infectious virus can be obtained by transfection of mammalian cells with the cloned BAV genome rescued from plasmid-containing cells.

The invention also provides BAV regulatory sequences which can be used to regulate the expression of heterologous genes. A regulatory sequence can be, for example, a transcriptional regulatory sequence, a promoter, an enhancer, an upstream regulatory domain, a splicing signal, a polyadenylation signal, a transcriptional termination sequence, a translational regulatory sequence, a ribosome binding site and a translational termination sequence.

In another embodiment, the cloned BAV genome can be propagated as a plasmid and infectious virus can be rescued from plasmid-containing cells.

The presence of viral nucleic acids can be detected by techniques known to one of skill in the art including, but not limited to, hybridization assays, polymerase chain reaction, and other types of amplification reactions. Similarly, methods for detection of proteins are well-known to those of skill in the art and include, but are not limited to, various types of immunoassay, ELISA, Western blotting, enzymatic assay, immunohistochemistry, *etc.* Diagnostic kits comprising the nucleotide sequences of the invention may also contain reagents for cell disruption and nucleic acid purification, as well

as buffers and solvents for the formation, selection and detection of hybrids. Diagnostic kits comprising the polypeptides or amino acid sequences of the invention may also comprise reagents for protein isolation and for the formation, isolation, purification and/or detection of immune complexes.

Various foreign genes or nucleotide sequences or coding sequences (prokaryotic, and eukaryotic) can be inserted in the bovine adenovirus nucleotide sequence, e.g., DNA, in accordance with the present invention, particularly to provide protection against a wide range of diseases and many such genes are already known in the art. The problem heretofore has been to provide a safe, convenient and effective vaccine vector for the genes or sequences, as well as safe, effective means for gene transfer to be used in various gene therapy applications.

An exogenous (*i.e.*, foreign) nucleotide sequence can consist of one or more gene(s) of interest, and preferably of therapeutic interest. In the context of the present invention, a gene of interest can code either for an antisense RNA, a ribozyme or for an mRNA which will then be translated into a protein of interest. A gene of interest can be of genomic type, of complementary DNA (cDNA) type or of mixed type (minigene, in which at least one intron is deleted). It can code for a mature protein, a precursor of a mature protein, in particular a precursor intended to he secreted and accordingly comprising a signal peptide, a chimeric protein originating from the fusion of sequences of diverse origins, or a mutant of a natural protein displaying improved or modified biological properties. Such a mutant may be obtained by, deletion, substitution and/or addition of one or more nucleotide(s) of the gene coding for the natural protein, or any other type of change in the sequence encoding the natural protein, such as, for example, transposition or inversion.

A gene of interest may be placed under the control of elements (DNA control sequences) suitable for its expression in a host cell. Suitable DNA control sequences are understood to mean the set of elements needed for transcription of a gene into RNA (antisense RNA or mRNA) and for the translation of an mRNA into protein. Among the elements needed for transcription, the promotcr assumes special importance. It can be a constitutive promoter or a regulatable promoter, and can he isolated from any gene of eukaryotic, prokaryotic or viral origin, and even adenoviral origin. Alternatively, it can be the natural promoter of the gene of interest. Generally speaking, a promoter used in the present invention may be modified so as to contain regulatory sequences. As examples, a

gene of interest in use in the present invention is placed under the control of the promoter of the immunoglobulin genes when it is desired to target its transfer to lymphocytic host cells. There may also be mentioned the HSV-1 TK (herpesvirus type 1 thymidine kinase) gene promoter, the adenoviral MLP (major late promoter), in particular of human adenovirus type 2, the RSV (Rous Sarcoma Virus) LTR (long terminal repeat), the CMV (Cytomegalovirus) early promoter, and the PGK (phosphoglycerate kinase) gene promoter, for example, permitting expression in a large number of cell types.

As disclosed herein altering species tropism is demonstrated in BAV by replacement of the native fiber protein region with a heterologous mammalian fiber protein region. The present invention also encompasses replacement of one bovine serotype adenovirus fiber region with another bovine serotype adenovirus fiber region wherein said replacement is associated with altered bovine cell specificity. Alternatively, targeting of a recombinant BAV vector to a particular cell type can be achieved by constructing recombinant hexon and/or fiber genes. The protein products of these genes are involved in host cell recognition; therefore, the genes can be modified to contain peptide sequences that will allow the virus to recognize alternative host cells.

Among genes of interest which are useable in the context of the present invention, there may be mentioned:

- genes coding for cytokines such as interferons and interleukins;
- genes encoding lymphokines;
- genes coding for membrane receptors such as the receptors recognized by pathogenic organisms (viruses, bacteria or parasites), preferably by the HIV virus (human immunodeficiency virus);
- genes coding for coagulation factors such as factor VIII and factor IX;
- genes coding for dystrophins;
- genes coding for insulin;
- genes coding for proteins participating directly or indirectly in cellular ion channels, such as the CFTR (cystic fibrosis transmembrane conductance regulator) protein;
- genes coding for antisense RNAs, or proteins capable of inhibiting the activity of a protein produced by a pathogenic gene which is present in the genome of a pathogenic organism, or proteins (or genes encoding them) capable of inhibiting the activity of a cellular gene whose expression is deregulated, for example an oncogene;

25

- genes coding for a protein inhibiting an enzyme activity, such as α_1-antitrypsin or a viral protease inhibitor, for example;

- genes coding for variants of pathogenic proteins which have been mutated so as to impair their biological function, such as, for example, trans-dominant variants of the *tat* protein of the HIV virus which are capable of competing with the natural protein for binding to the target sequence, thereby preventing the activation of HIV;

- genes coding for antigenic epitopes in order to increase the host cell's immunity;

- genes coding for major histocompatibility complex classes I and II proteins, as well as the genes coding for the proteins which are inducers of these genes;

- genes coding for antibodies;

- genes coding for immunotoxins;

- genes encoding toxins;

- genes encoding growth factors or growth hormones;

- genes encoding cell receptors and their ligands;

- genes encoding tumor suppressors;

- genes involved in cardiovascular disease including, but not limited to, oncogenes; genes encoding growth factors including, but not limited to, fibroblast growth factor (FGF), vascular endothelial growth factor (VEGF), and nerve growth factor (NGF); e-*nos*, tumor suppressor genes including, but not limited to, the Rb (retinoblastoma) gene; lipoprotein lipase; superoxide dismutase (SOD); catalase; oxygen and free radical scavengers; apolipoproteins; and *pai*-1 (plasminogen activator inhibitor-1);

- genes coding for cellular enzymes or those produced by pathogenic organisms; and

- suicide genes. The HSV-1 TK suicide gene may be mentioned as an example. This viral TK enzyme displays markedly greater affinity compared to the cellular TK enzyme for certain nucleoside analogues (such as acyclovir or gancyclovir). It converts them to monophosphorylated molecules, which can themselves be converted by cellular enzymes to nucleotide precursors, which are toxic. These nucleotide analogues can be incorporated into replicating DNA molecules, hence incorporation occurs chiefly in the DNA of dividing cells. This incorporation can result in specific destruction of dividing cells such as cancer cells.

This list is not restrictive, and other genes of interest may be used in the context of the present invention.

It is also possible that only fragments of nucleotide sequences of genes can be used (where these are sufficient to generate a protective immune response or a specific biological effect) rather than the complete sequence as found in the wild-type organism. Where available, synthetic genes or fragments thereof can also be used. However, the present invention can be used with a wide variety of genes, fragments and the like, and is not limited to those set out above.

In some cases the gene for a particular antigen can contain a large number of introns or can be from an RNA virus, in these cases a complementary DNA copy (cDNA) can be used.

In order for successful expression of the gene to occur, it can be inserted into an expression vector together with a suitable promoter including enhancer elements and polyadenylation sequences. A number of eucaryotic promoter and polyadenylation sequences which provide successful expression of foreign genes in mammalian cells and construction of expression cassettes, are known in the art, for example in U.S. Patent 5,151,267, the disclosures of which are incorporated herein by reference. The promoter is selected to give optimal expression of immunogenic protein which in turn satisfactorily leads to humoral, cell mediated and mucosal immune responses according to known criteria.

The foreign protein produced by expression *in vivo* in a recombinant virus-infected cell may be itself immunogenic. More than one foreign gene can be inserted into the viral genome to obtain successful production of more than one effective protein.

Thus with the recombinant viruses of the present invention, it is possible to provide protection against a wide variety of diseases affecting cattle, humans and other mammals. Any of the recombinant antigenic determinants or recombinant live viruses of the invention can be formulated and used in substantially the same manner as described for antigenic determinant vaccines or live vaccine vectors.

The present invention also includes pharmaceutical compositions comprising a therapeutically effective amount of a recombinant adenovirus vector, recombinant adenovirus or recombinant protein, prepared according to the methods of the invention, in combination with a pharmaceutically acceptable vehicle and/or an adjuvant. Such a

pharmaceutical composition can be prepared and dosages determined according to techniques that are well-known in the art. The pharmaceutical compositions of the invention can be administered by any known administration route including, but not limited to, systemically (for example, intravenously, intratracheally, intravascularly, intrapulmonarilly, intraperitoneally, intranasally, parenterally, enterically, intramuscularly, subcutaneously, intratumorally or intracranially) or by aerosolization or intrapulmonary instillation. Administration can take place in a single dose or in doses repeated one or more times after certain time intervals. The appropriate administration route and dosage will vary in accordance with the situation (for example, the individual being treated, the disorder to be treated or the gene or polypeptide of interest), but can be determined by one of skill in the art.

The invention also encompasses a method of treatment, according to which a therapeutically effective amount of a BAV vector, recombinant BAV, or host cell of the invention is administered to a mammalian subject requiring treatment.

The antigens used in the present invention can be either native or recombinant antigenic polypeptides or fragments. They can be partial sequences, full-length sequences, or even fusions (e.g., having appropriate leader sequences for the recombinant host, or with an additional antigen sequence for another pathogen). The preferred antigenic polypeptide to be expressed by the virus systems of the present invention contain full-length (or near full-length) sequences encoding antigens. Alternatively, shorter sequences that are antigenic (i.e., encode one or more epitopes) can be used. The shorter sequence can encode a "neutralizing epitope," which is defined as an epitope capable of eliciting antibodies that neutralize virus infectivity in an *in vitro* assay. Preferably the peptide should encode a "protective epitope" that is capable of raising in the host a "protective immune response;" i.e., an antibody- and/or a cell-mediated immune response that protects an immunized host from infection.

The antigens used in the present invention, particularly when comprised of short oligopeptides, can be conjugated to a vaccine carrier. Vaccine carriers are well known in the art: for example, bovine serum albumin (BSA), human serum albumin (HSA) and keyhole limpet hemocyanin (KLH). A preferred carrier protein, rotavirus VP6, is disclosed in EPO Pub. No. 0259149, the disclosure of which is incorporated by reference herein.

Genes for desired antigens or coding sequences thereof which can be inserted include those of organisms which cause disease in mammals, particularly bovine pathogens such as bovine rotavirus, bovine coronavirus, bovine herpes virus type 1, bovine respiratory syncytial virus, bovine parainfluenza virus type 3 (BPI-3), bovine diarrhea virus, *Pasteurella haemolytica, Haemophilus somnus* and the like. Genes encoding antigens of human pathogens also useful in the practice of the invention. The vaccines of the invention carrying foreign genes or fragments can also be orally administered in a suitable oral carrier, such as in an enteric-coated dosage form. Oral formulations include such normally-employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin cellulose, magnesium carbonate, and the like. Oral vaccine compositions may be taken in the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations, or powders, containing from about 10% to about 95% of the active ingredient, preferably about 25% to about 70%. Oral and/or intranasal vaccination may be preferable to raise mucosal immunity (which plays an important role in protection against pathogens infecting the respiratory and gastrointestinal tracts) in combination with systemic immunity.

In addition, the vaccine can be formulated into a suppository. For suppositories, the vaccine composition will include traditional binders and carriers, such as polyalkaline glycols or triglycerides. Such suppositories may be formed from mixtures containing the active ingredient in the range of about 0.5% to about 10% (w/w), preferably about 1% to about 2%.

Protocols for administering to animals the vaccine composition(s) of the present invention are within the skill of the art in view of the present disclosure. Those skilled in the art will select a concentration of the vaccine composition in a dose effective to elicit an antibody and/or T-cell mediated immune response to the antigenic fragment. Within wide limits, the dosage is not believed to be critical. Typically, the vaccine composition is administered in a manner which will deliver between about 1 to about 1,000 micrograms of the subunit antigen in a convenient volume of vehicle, e.g., about 1-10 cc. Preferably, the dosage in a single immunization will deliver from about 1 to about 500 micrograms of subunit antigen, more preferably about 5-10 to about 100-200 micrograms (e.g., 5-200 micrograms).

The timing of administration may also be important. For example, a primary inoculation preferably may be followed by subsequent booster inoculations if needed. It may also be preferred, although optional, to administer a second, booster immunization to the animal several weeks to several months after the initial immunization. To insure sustained high levels of protection against disease, it may be helpful to readminister a booster immunization to the animals at regular intervals, for example once every several years. Alternatively, an initial dose may be administered orally followed by later inoculations, or vice versa. Preferred vaccination protocols can be established through routine vaccination protocol experiments.

The dosage for all routes of administration of *in vivo* recombinant virus vaccine depends on various factors including, the size of patient, nature of infection against which protection is needed, carrier and the like and can readily be determined by those of skill in the art. By way of non-limiting example, a dosage of between 10^3 pfu and 10^{15} pfu, preferably between 10^5 and 10^{13} pfu, more preferably between 10^6 to 10^{11} pfu and the like can be used. As with *in vitro* subunit vaccines, additional dosages can be given as determined by the clinical factors involved.

In some embodiments of the invention, recombinant cell lines are produced by constructing an expression cassette comprising the BAV E1 region, and/or other essential gene region and transforming host cells therewith to provide complementing cell lines or cultures expressing the E1 proteins for use with replication-defective bovine adenoviruses modified to have altered tropism and lacking E1 function. These recombinant complementing cell lines are capable of allowing a defective recombinant BAV with deleted E1 sequences to replicate and express a desired foreign gene or fragment thereof which is optionally encoded within the recombinant BAV. These cell lines are also extremely useful in generating recombinant BAV, having an E3 gene deletion replaced by heterologous nucleotide sequence encoding for a foreign gene or fragment, by *in vivo* recombination following DNA-mediated cotransfection. More generally, defective recombinant BAV vectors, lacking one or more essential functions encoded by the BAV genome, can be propagated in appropriate complementing cell lines, wherein a particular complementing cell line provides a function or functions that is (are) lacking in a particular defective recombinant BAV vector. Complementing cell lines can provide viral functions through, for example, co-infection with a helper virus, or by integrating or otherwise

maintaining in stable form a fragment of a viral genome encoding a particular viral function.

In one embodiment of the invention, the recombinant expression cassette can be obtained by cleaving a BAV genome with an appropriate restriction enzyme to produce a DNA fragment representing the left end or the right end of the genome comprising E1 or E3 gene region sequences, respectively and inserting the left or right end fragment into a cloning vehicle, such as a plasmid, and thereafter inserting at least one heterologous DNA sequence into the E1 or E3 deletion with or without the control of an exogenous promoter. The recombinant expression cassette is contacted with a BAV genome within an appropriate cell and, through homologous recombination or other conventional genetic engineering method, a recombinant BAV genome is obtained. Appropriate cells include both prokaryotic cells, such as, for example, *E. coli*, and eukaryotic cells. Examples of suitable eukaryotic cells include, but are not limited to, MDBK cells, MDBK cells expressing adenovirus E1 function, primary fetal bovine retina cells, and cells expressing functions that are equivalent to those of the previously-recited cells. Restriction fragments of the BAV genome other than those comprising the E1 or E3 regions are also useful in the practice of the invention and can be inserted into a cloning vehicle such that heterologous sequences may be inserted into non-E1 and E3 BAV sequences. These DNA constructs can then undergo recombination *in vitro* or *in vivo*, with a BAV genome, either before or after transformation or transfection of a suitable host cell as described above. For the purposes of the present invention, a BAV genome can be either a full-length genome or a genome containing a deletion in a region other than that deleted in the fragment with which it recombines, as long as the resulting recombinant BAV genome contains BAV sequences required for replication and packaging. Methods for transfection, cell culture and recombination in procaryotic and eukaryotic cells such as those described above are well-known to those of skill in the art.

In another embodiment of the invention, the function of any viral region which may be mutated or deleted in any particular viral vector can be supplied (to provide a complementing cell line) by co-infection of cells with a virus which expresses the function that the vector lacks.

If an insertion is made in a gene essential for viral replication, the adenovirus must be grown in an appropriate complementing cell line (*i.e.*, a helper cell line). In human

adenoviruses, certain open reading frames in the E4 region (ORF 3 and ORF 6/7) are essential for viral replication. Deletions in analogous open reading frames in the E4 region of BAV-3 could necessitate the use of a helper cell line for growth of the viral vector.

The BAV E1 gene products of the adenovirus of the invention transactivate most of the cellular genes, and therefore, cell lines which constitutively express E1 proteins can express cellular polypeptides at a higher level than normal cell lines. The recombinant mammalian, particularly bovine, cell lines of the invention can be used to prepare and isolate polypeptides, including those such as (a) proteins associated with adenovirus E1A proteins: e.g. p300, retinoblastoma (Rb) protein, cyclins, kinases and the like; (b) proteins associated with adenovirus E1B protein: e.g. p53 and the like; (c) growth factors, such as epidermal growth factor (EGF), transforming growth factor (TGF) and the like; (d) receptors such as epidermal growth factor receptor (EGF-R), fibroblast growth factor receptor (FGF-R), tumor necrosis factor receptor (TNF-R), insulin-like growth factor receptor (IGF-R), major histocompatibility complex class I receptor and the like; (e) proteins encoded by proto-oncogenes such as protein kinases (tyrosine-specific protein kinases and protein kinases specific for serine or threonine), p21 proteins (guanine nucleotide-binding proteins with GTPase activity) and the like; (f) other cellular proteins such as actins, collagens, fibronectins, integrins, phosphoproteins, proteoglycans, histones and the like, and (g) proteins involved in regulation of transcription such as TATA-box-binding protein (TBP), TBP-associated factors (TAFs), Sp1 binding protein and the like.

The invention also includes a method for providing gene delivery to a mammal, such as a bovine or a human or other mammal in need thereof, to control a gene deficiency, to provide a therapeutic gene or nucleotide sequence and/or to induce or correct a gene mutation. The method can be used, for example, in the treatment of conditions including, but not limited to hereditary disease, infectious disease, cardiovascular disease, and viral infection. The method comprises administering to said mammal a live recombinant bovine adenovirus comprising a modification in a capsid protein, or fragment thereof, wherein said capsid protein is associated with tropism and said modification is associated with altered tropism and wherein said adenovirus vector further comprises a foreign polynucleotide sequence encoding a non-defective form of said gene under conditions wherein the recombinant virus vector genome is incorporated into said mammalian genome or is maintained independently and extrachromosomally to provide expression of the required

gene in the target organ or tissue. These kinds of techniques are currently being used by those of skill in the art for the treatment of a variety of disease conditions, non-limiting examples of which are provided above. Examples of foreign genes, nucleotide sequences or portions thereof that can be incorporated for use in a conventional gene therapy include, cystic fibrosis transmembrane conductance regulator gene, human minidystrophin gene, alpha-1-antitrypsin gene, genes involved in cardiovascular disease, and the like.

In particular, the practice of the present invention in regard to gene delivery in humans is intended for the prevention or treatment of diseases including, but not limited to, genetic diseases (for example, hemophilia, thalassemias, emphysema, Gaucher's disease, cystic fibrosis, Duchenne muscular dystrophy, Duchenne's or Becker's myopathy, *etc.*), cancers, viral diseases (for example, AIDS, herpesvirus infection, cytomegalovirus infection and papillomavirus infection), cardiovascular diseases, and the like. For the purposes of the present invention, the vectors, cells and viral particles prepared by the methods of the invention may be introduced into a subject either *ex vivo*, (*i.e.*, in a cell or cells removed from the patient) or directly *in vivo* into the body to be treated.

The following examples are provided to illustrate but not limit the invention.

EXAMPLES

Example 1: Construction of BAV600 containing a human fiber gene

To generate an BAV-3 vector with an altered tropism, the chimeric fiber gene construct containing the HAV-5 fiber knob fused to the BAV-3 tail and shaft was incorporated into the BAV-3 genome of BAV304, described in Reddy *et al.*, *supra* 1999 (Fig. 3). For the precise replacement of the wild-type BAV-3 fiber gene, a previously made plasmid pBAV301.gfp (Reddy *et al.*, 1999) was used for modification of BAV-3 fiber. The resulting transfer vector pBAV-301.G5FK contained a CMV promoter driven green fluorescent protein (GFP) expression cassette inserted into the E3 region, the chimeric BAV-3/HAV-5 fiber gene, and E4. This transfer vector was used for incorporation of GFP cassette and modified fiber gene into the backbone of an E3 deleted BAV-3 infectious plasmid, p.FBAV302 (Zakhartchouk *et al.*, 1998), via homologous recombination in *E. coli* BJ5183 (Chartier et al., 1996), creating plasmid pFBAV-600. The viral genome was released from the plasmid by PacI digestion and used to transfect cell line ATCC accession

number PTA156, fetal bovine retinal cells expressing E1 protein (see Reddy et al. 1999, *supra*). The corresponding chimeric virus BAV600 was produced 21 days following transfection.

Example 2: Characterization of BAV600

BAV600 obtained from the transfection of fetal bovine retinal cells expressing E1 protein, ATCC accession number PTA156, was amplified in MDBK cells, and the viral DNA was extracted from infected cells. The DNA was analyzed after digestion with restriction enzyme *Bg*/II and agarose gel electrophoresis (Figure 5A). As shown in Figures 5A-5B, both the parental BAV302 and BAV304 had *Bg*/II fragment of 5.4 kb at the right end of viral genome. The HAV-5 fiber knob region introduces an additional *Bg*/II restriction enzyme site within the BAV600 genome. Therefore, diagnostic 1.5 and 3.9 kb fragments were found after *Bg*/II digestion. Southern blot analysis with the HAV-5 fiber knob probe demonstrated the expected hybridization pattern for *Bg*/II-digested BAV600 (Figure 5B).

Expression and assembly of the chimeric BAV-3 and HAV-5 fiber protein by recombinant BAV600 were examined by immunoprecipitation assay. Metabolically radiolabeled immunoprecipitates from the parental (BAV304; Reddy *et al.*, 1999, *supra*) and chimeric (BAV600) viruses-infected MDBK cell lysates were subjected to SDS-PAGE under denaturing conditions. A wild-type HAV-5 containing a full-length fiber was also analyzed. Immunoprecipitation assay was carried out with a rabbit polyclonal antibody specific for the BAV3 fiber knob and an antifiber monoclonal antibody, ID6.14. The ID6.14 antibody recognizes a trimerized HAV-5 fiber knob and neutralizes HAV-5 through binding to knob domain (Douglas *et al.*, 1996). As shown in Figure 6, the BAV-3 and BAV304 viruses contain fiber proteins with sizes of approximately 100 kDa which react with the rabbit polyclonal antibody specific for the BAV3 fiber knob, while the HAV-5 contains a fiber protein with a size of approximately 64 kDa. The presence of the HAV-5 fiber knob within the BAV600 chimeric virus was confirmed by immunoprecipitation analysis with the monoclonal antibody ID6.14 specific for the HAV-5 knob.

The biological titer of the fiber chimeric virus BAV600 was compared with the BAV-3 and parental virus BAV304. Biological titers determined with MDBK cell monolayers indicated maximum plaque-forming titers of 10^8, 10^6, and 10^5PFU/ml for the

BAV-3, BAV304, and BAV600, respectively. The result suggested that the fiber modification and GFP insertion in E3 region significantly alter the cellular production of the virus.

Example 3: Transduction of human cell lines by BAV600

To characterize the transduction efficiency of BAV304 and BAV600 in different human cell lines, FACS analysis was performed to determine the percentage of transduction of each cell line at different virus input (Fig 7A). Cells grown in T25 flasks were infected at an MOI of 1 and 5 with either BAV304 or BAV600. Forty-eight hours after infection, the percentage of GFP-fluorescence positive cells were determined by flow cytometry. The percentage of transduction of each cell line was quantitated, and the fraction of dose is shown in Figure 7B. 293 cells were equally susceptible to transduction with both viruses (indicating that both the HAV-5 and BAV-3 receptors are present on the cell surface.) The transduction of HeLa and HEp-2 cells with BAV304 is dose dependent, with about 6% and 1% respectively at an MOI of 1 and about 25% and 5% respectively at an MOI of 5. Both cells were efficiently transduced with BAV600. The percentage of transduction with BAV600 reaches maximum level even at an MOI of 1 (94% and 93% for HeLa and Hep-2 respectively). In contrast less-efficient transduction of A549 cells with BAV600 was observed. These data taken together demonstrate that the BAV600 containing HAV-5 fiber knob was clearly superior to the BAV304 vector in transduction of human cell lines.

Example 4: HAV–5 and BAV–3 neutralizing antibodies in human serum

Preexisting neutralizing antibodies against HAV–5 in clinical patients represent a major obstacle for efficient use of HAV–5 in human gene therapy protocol. In order to explore the possibility for use of BAV–3 as an alternative vector to HAV–5-derived vectors, it was determined whether preexisting anti-HAV–5 neutralizing antibodies were also cross-reactive with BAV–3. 105 random samples of human sera from clinical patients were tested. Three (#50, 97, and 102) were found containing high titer of HAV–5 neutralizing antibodies ranging between 1:800 to 1,6000. These sera were tested for their ability to inhibit BAV–3-induced plaque formation on MDBK cells. Our data

demonstrated that none of these HAV–5 positive sera showed effect on BAV–3-induced plaque formation at a dilution of 1/50.

Example 5: Replication of BAV-3 in human cell lines

Virus production and the time course of virus infection were studied in different human cell lines to determine their degree of permissivity for BAV-3 growth. Confluent monolayer cultures of each cell line (HeLa, HEp-2, 293 and A549) were infected-with BAV–3 at an MOI of 10 and virus production at different time intervals after infection was assayed by titration of the cell lysates on MDBK cell monolayers. Virus growth in permissive MDBK cells resulted in, as expected, maximum yields of 10^8 pfu/ml by 48 hours after infection. In contrast, the level of BAV–3 production in all four human cell lines was constantly diminished, suggesting that there is a complete absence of viral replication in these human cell lines.

Example 6: Expression of early and later BAV-3 proteins in human cell lines

Viral proteins include early proteins (E1B small and single-stranded DNA binding protein [DBP])and late proteins (penton base and fiber). To identify the expression of early and late viral proteins in human cell lines, viral protein production was analyzed by Western immunoblotting. Cultures were infected with BAV–3 at an MOI of 10. At intervals after infection, cell extracts were prepared from each culture, separated on 10% SDS-PAGE, and transferred to nitrocellulose. Antigens immobilized on the nitrocellulose sheets were probed by reaction with rabbit polyclonal antibodies against E1B small, DBP, penton base, and fiber respectively. As expected, the E1B small and DBP antisera reacted with bands in 19 and 50 kDa, respectively, from BAV–3-infected MDBK cells. In contrast, all human cell lines except 293 cell lines showed no positive reactions with anti-E1B small or DBP polyclonal antibodies. No structural proteins were detected from BAV–3- infected human cell lines. These results indicated that the replication of BAV–3 in the majority of human cells tested in this study was blocked at E1B small level.

Example 7: Neutralization of BAV600 by an monoclonal antibody specific for HAV-5 fiber knob

It was hypothesized that BAV600 carrying the HAV–5 fiber knob should be neutralized by an antibody specific for HAV–5 knob. To confirm this, duplicate aliquots containing 100 pfu of BAV–3 or BAV600 were incubated at room temperature for two hours with serial twofold dilutions of a rabbit polyclonal antibody specific for the BAV3 fiber knob or a monoclonal antibody, 1D6.14, against HAV-5 fiber knob domain. MDBK cells were then infected with pre-incubated BAV–3 or BAV600 virus. Cells were incubated for 14 days to allow a complete CPE to develop. The data show that that none of the viruses were neutralized by serum from normal rabbit serum or a control monoclonal antibody 2C8 specific for bovine herpesvirus gD protein. BAV–3 and BAV600 were each neutralized by a rabbit polyclonal antibody specific for the BAV3 fiber knob (1:800) and ID6.14 (1:3,200), respectively. However, neither virus was neutralized by the reciprocal antiserum even at a dilution of 1:50. This further confirmed that BAV600 carried the HAV–5 fiber knob.

CLAIMS

1. A bovine adenovirus vector comprising a modification in a polynucleotide encoding a capsid protein, or fragment thereof, wherein said capsid protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism.

2. The adenovirus vector of claim 1 wherein said polynucleotide encoding a capsid protein, or fragment thereof, is replaced with a polynucleotide encoding a heterologous mammalian capsid protein, or fragment thereof.

3. The adenovirus vector of claim 1 wherein said capsid protein, or fragment thereof, is a penton protein, or fragment thereof.

4. The adenovirus vector of claim 1 wherein said capsid protein, or fragment thereof, is a hexon protein, or fragment thereof.

5. The adenovirus vector of claim 1 wherein said capsid protein, or fragment thereof, is a fiber protein, or fragment thereof.

6. The adenovirus vector of claim 5 wherein the modification is in the knob region of a fiber protein.

7. The adenovirus vector of claim 3 wherein said bovine adenovirus penton region, or fragment thereof, is replaced with at least one heterologous mammalian penton adenovirus region, or fragment thereof.

8. The adenovirus vector of claim 4 wherein said bovine adenovirus hexon region, or fragment thereof, is replaced with at least one heterologous mammalian adenovirus hexon region, or fragment thereof.

9. The adenovirus vector of claim 5 wherein said bovine adenovirus fiber region, or fragment thereof, is replaced with at least one heterologous mammalian adenovirus fiber region or fragment thereof.

10. The adenovirus vector of claim 2 wherein said heterologous mammalian adenovirus capsid protein, or fragment thereof, includes porcine, ovine, canine or human adenovirus capsid protein, or fragment thereof.

11. The adenovirus vector of claim 10 wherein said heterologous mammalian adenovirus capsid protein, or fragment thereof, is a human adenovirus capsid protein, or fragment thereof.

12. The adenovirus vector of claim 1 wherein said adenovirus is a sub-type 1 adenovirus.

13. The adenovirus vector of claim 1 wherein said adenovirus is a sub-type 2 adenovirus.

14. The adenovirus vector of claim 12 wherein said adenovirus vector is BAV3.

15. The adenovirus vector of claim 14 wherein said modification is a replacement of BAV3 fiber protein, or fragment thereof, with a heterologous mammalian adenovirus fiber protein, or fragment thereof.

16. The adenovirus vector of claim 15 wherein said mammalian adenovirus fiber protein includes bovine, porcine, ovine, canine or human adenovirus fiber protein.

17. The adenovirus vector of claim 16 wherein said mammalian adenovirus fiber protein is a human adenovirus fiber protein.

18. The adenovirus vector of claim 1 wherein said vector lacks E1 function.

19. The adenovirus vector of claim 18 wherein said vector has a deletion of part or all of the E1 gene region.

20. The adenovirus vector of claim 1 wherein said vector has a deletion of part or all of the E3 gene region.

21. The adenovirus vector of claim 1 wherein said vector further comprises a polynucleotide encoding a heterologous protein.

22. The adenovirus vector of claim 21 wherein said heterologous protein includes cytokines; lymphokines; membrane receptors recognized by pathogenic organisms, dystrophins; insulin; proteins participating in cellular ion channels; antisense RNAs; proteins capable of inhibiting the activity of a protein produced by a pathogenic gene, a protein inhibiting an enzyme activity, protein variants of pathogenic proteins; antigenic epitopes; major histocompatibility complex classes I and II proteins; antibodies; immunotoxins; toxins; growth factors or growth hormones; cell receptors or their ligands; tumor suppressors; cellular enzymes; or suicide genes.

23. The adenovirus of claim 22 wherein said polynucleotide encoding said heterologous protein is inserted in the adenovirus E1 gene region.

24. The adenovirus of claim 22 wherein said polynucleotide encoding said heterologous protein is inserted in the adenovirus E3 gene region.

25. The adenovirus vector of claim 1 wherein said vector is replication-competent.

26. The adenovirus vector of claim 1 wherein said vector is replication-defective.

27 A host cell comprising the bovine adenovirus vector of claim 1.

28. A host cell comprising the bovine adenovirus vector of claim 21.

29. A method of producing a recombinant bovine adenovirus vector comprising a modification in a polynucleotide encoding a capsid protein, or a fragment thereof, comprising the steps of, obtaining a bovine adenovirus vector; and introducing a

modification into a polynucleotide encoding a capsid protein, or fragment thereof, wherein said capsid protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism.

30. The method of claim 29 wherein said capsid protein, or fragment thereof, is a penton protein, or fragment thereof.

31. The method of claim 29 wherein said capsid protein, or fragment thereof, is a hexon protein, or fragment thereof.

32. The method of claim 29 wherein said capsid protein, or fragment thereof, is a fiber protein, or fragment thereof.

33. The method of claim 29 wherein said adenovirus vector further comprises a polynucleotide encoding a heterologous protein.

34. The method of claim 29 wherein said bovine adenovirus is a sub-type 1 bovine adenovirus.

35. A recombinant bovine adenovirus comprising a modification in a polynucleotide encoding a capsid protein, or fragment thereof, wherein said capsid protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism.

36. The recombinant adenovirus of claim 35 further comprising a polynucleotide encoding a heterologous protein.

37. The recombinant adenovirus of claim 36 wherein said polynucleotide encoding said heterologous protein is inserted in the adenovirus E1 gene region.

38. The recombinant adenovirus of claim 36 wherein said polynucleotide encoding said heterologous protein is inserted in the adenovirus E3 gene region.

39. The recombinant adenovirus of claim 35 wherein said capsid protein, or fragment thereof, is a penton protein, or fragment thereof.

40. The recombinant adenovirus of claim 35 wherein said capsid protein, or fragment thereof, is a hexon protein, or fragment thereof.

41. The recombinant adenovirus of claim 35 wherein said capsid protein, or fragment thereof, is a fiber protein, or fragment thereof.

42. The recombinant adenovirus of claim 41 wherein the modification is in the knob region of a fiber protein.

43. An immunogenic composition comprising a bovine adenovirus wherein said adenovirus comprises a polynucleotide encoding a modification in a capsid protein, or fragment thereof, and wherein said protein, or fragment thereof, is associated with tropism and wherein said modification is associated with altered tropism.

44. The immunogenic composition of claim 43 wherein said capsid protein is a penton protein, or fragment thereof.

45. The immunogenic composition of claim 43 wherein said capsid protein is a hexon protein, or fragment thereof.

46. The immunogenic composition of claim 43 wherein said capsid protein is a fiber protein, or fragment thereof.

47. The immunogenic composition of claim 46 wherein said capsid protein, or fragment thereof, is a knob domain of a fiber protein.

48. The immunogenic composition of claim 43 wherein said modification is a replacement of a bovine fiber protein, or fragment thereof, with a mammalian adenovirus fiber protein, or fragment thereof.

49. The immunogenic composition of claim 48 wherein said mammalian fiber protein is a human adenovirus fiber protein.

50. The immunogenic composition of claim 43 wherein said bovine adenovirus is a sub-type 1 adenovirus.

51. The immunogenic composition of claim 50 wherein said bovine adenovirus is BAV3.

52. The immunogenic composition of claim 43 wherein said bovine adenovirus comprises a polynucleotide encoding a heterologous protein.

53. A pharmaceutical composition capable of inducing an immune response in a mammalian subject, said composition comprising the immunogenic composition of claim 52.

54. The pharmaceutical composition of claim 53 further comprising a pharmaceutically acceptable excipient.

55. A method for eliciting an immune response in a mammalian host to protect against infection, the method comprising administration of the pharmaceutical composition of claim 54.

56. The method of claim 55 wherein said protein includes cytokines; lymphokines; membrane receptors recognized by pathogenic organisms, dystrophins; insulin; proteins participating in cellular ion channels; antisense RNAs; proteins capable of inhibiting the activity of a protein produced by a pathogenic gene, a protein inhibiting an enzyme activity, protein variants of pathogenic proteins; antigenic epitopes; major histocompatibility complex classes I and II proteins; antibodies; immunotoxins; toxins; growth factors or

growth hormones; cell receptors or their ligands; tumor suppressors; cellular enzymes; or suicide genes.

57. A method of gene delivery in a mammalian host, the method comprising administering to the host a bovine adenovirus vector comprising a polynucleotide encoding a modified capsid protein, or fragment thereof, wherein the protein is associated with tropism and wherein the modification is associated with altered tropism and wherein the adenovirus vector further comprises a polynucleotide encoding a heterologous protein.

58. The method of claim 57 wherein said heterologous polynucleotide encodes a therapeutic protein.

59. The method of claim 57 wherein said capsid protein, or fragment thereof, is a penton protein, or fragment thereof.

60. The method of claim 57 wherein said capsid protein, or fragment thereof, is a hexon protein, or fragment thereof.

61. The method of claim 57 wherein said capsid protein, or fragment thereof, is a fiber protein, or fragment thereof.

62. The method of claim 61 wherein the modification is in the knob region of a fiber protein.

63. The method of claim 57 wherein said mammalian host is human and said modification is a replacement of a bovine adenovirus fiber protein, or fragment thereof, with a human fiber protein, or fragment thereof.

FIGURE 1A

```
CATCATCAAT AATCTACAGT ACACTGATGG CAGCGGTCCA ACTGCCAATC ATTTTTGCCA    60
CGTCATTTAT GACGCAACGA CGGCGAGCGT GGCGTGCTGA CGTAACTGTG GGGCGGAGCG   120
CGTCGCGGAG GCGGCGGCGC TGGGCGGGGC TGAGGGCGGC GGGGGCGGCG CGCGGGGCGG   180
CGCGCGGGGC GGGGCGAGGG GCGGAGTTCC GCACCCGCTA CGTCATTTTC AGACATTTTT   240
TAGCAAATTT GCGCCTTTTG CAAGCATTTT TCTCACATTT CAGGTATTTA GAGGGCGGAT   300
TTTTGGTGTT CGTACTTCCG TGTCACATAG TTCACTGTCA ATCTTCATTA CGGCTTAGAC   360
AAATTTTCGG CGTCTTTTCC GGGTTTATGT CCCCGGTCAC CTTTATGACT GTGTGAAACA   420
CACCTGCCCA TTGTTTACCC TTGGTCAGTT TTTTCGTCTC CTAGGGTGGG AACATCAAGA   480
ACAAATTTGC CGAGTAATTG TGCACCTTTT TCCGCGTTAG GACTGCGTTT CACACGTAGA   540
CAGACTTTTT CTCATTTTCT CACACTCCGT CGTCCGCTTC AGAGCTCTGC GTCTTCGCTG   600
CCACCATGAA GTACCTGGTC CTCGTTCTCA ACGACGGCAT GAGTCGAATT GAAAAAGCTC   660
TCCTGTGCAG CGATGGTGAG GTGGATTTAG AGTGTCATGA GGTACTTCCC CCTTCTCCCG   720
CGCCTGTCCC CGCTTCTGTG TCACCCGTGA GGAGTCCTCC TCCTCTGTCT CCGGTGTTTC   780
CTCCGTCTCC GCCAGCCCCG CTTGTGAATC CAGAGGCGAG TTCGCTGCTG CAGCAGTATC   840
GGAGAGAGCT GTTAGAGAGG AGCCTGCTCC GAACGGCCGA AGGTCAGCAG CGTGCAGTGT   900
GTCCATGTGA GCGGTTGCCC GTGGAAGAGG ATGAGTGTCT GAATGCCGTA AATTTGCTGT   960
TTCCTGATCC CTGGCTAAAT GCAGCTGAAA ATGGGGGTGA TATTTTTAAG TCTCCGGCTA  1020
TGTCTCCAGA ACCGTGGATA GATTTGTCTA GCTACGATAG CGATGTAGAA GAGGTGACTA  1080
GTCACTTTTT TCTGGATTGC CCTGAAGACC CCAGTCGGGA GTGTTCATCT TGTGGGTTTC  1140
ATCAGGCTCA AAGCGGAATT CCAGGCATTA TGTGCAGTTT GTGCTACATG CGCCAAACCT  1200
ACCATTGCAT CTATAGTAAG TACATTCTGT AAAAGAACAT CTTGGTGATT TCTAGGTATT  1260
GTTTAGGGAT TAACTGGGTG GAGTGATCTT AATCCGGCAT AACCAAATAC ATGTTTTCAC  1320
AGGTCCAGTT TCTGAAGAGG AAATGTGAGT CATGTTGACT TTGGCGCGCA AGAGGAAATG  1380
TGAGTCATGT TGACTTTGGC GCGCCCTACG GTGACTTTAA AGCAATTTGA GGATCACTTT  1440
TTTGTTAGTC GCTATAAAGT AGTCACGGAG TCTTCATGGA TCACTTAAGC GTTCTTTTGG  1500
ATTTGAAGCT GCTTCGCTCT ATCGTAGCGG GGGCTTCAAA TCGCACTGGA GTGTGGAAGA  1560
GGCGGCTGTG GCTGGGACGC CTGACTCAAC TGGTCCATGA TACCTGCGTA GAGAACGAGA  1620
GCATATTTCT CAATTCTCTG CCAGGGAATG AAGCTTTTTT AAGGTTGCTT CGGAGCGGCT  1680
ATTTTGAAGT GTTTGACGTG TTTGTGGTGC CTGAGCTGCA TCTGGACACT CCGGGTCGAG  1740
TGGTCGCCGC TCTTGCTCTG CTGGTGTTCA TCCTCAACGA TTTAGACGCT AATTCTGCTT  1800
CTTCAGGCTT TGATTCAGGT TTTCTCGTGG ACCGTCTCTG CGTGCCGCTA TGGCTGAAGG  1860
```

FIGURE 1B

CCAGGGCGTT CAAGATCACC CAGAGCTCCA GGAGCACTTC GCAGCCTTCC TCGTCGCCCG 1920

ACAAGACGAC CCAGACTACC AGCCAGTAGA CGGGGACAGC CCACCCCGGG CTAGCCTGGA 1980

GGAGGCTGAA CAGAGCAGCA CTCGTTTCGA GCACATCAGT TACCGAGACG TGGTGGATGA 2040

CTTCAATAGA TGCCATGATG TTTTTTATGA GAGGTACAGT TTTGAGGACA TAAAGAGCTA 2100

CGAGGCTTTG CCTGAGGACA ATTTGGAGCA GCTCATAGCT ATGCATGCTA AAATCAAGCT 2160

GCTGCCCGGT CGGGAGTATG AGTTGACTCA ACCTTTGAAC ATAACATCTT GCGCCTATGT 2220

GCTCGGAAAT GGGGCTACTA TTAGGGTAAC AGGGGAAGCC TCCCCGGCTA TTAGAGTGGG 2280

GGCCATGGCC GTGGGTCCGT GTGTAACAGG AATGACTGGG GTGACTTTTG TGAATTGTAG 2340

GTTTGAGAGA GAGTCAACAA TTAGGGGGTC CCTGATACGA GCTTCAACTC ACGTGCTGTT 2400

TCATGGCTGT TATTTTATGG GAATTATGGG CACTTGTATT GAGGTGGGGG CGGGAGCTTA 2460

CATTCGGGGT TGTGAGTTTG TGGGCTGTTA CCGGGGAATC TGTTCTACTT CTAACAGAGA 2520

TATTAAGGTG AGGCAGTGCA ACTTTGACAA ATGCTTACTG GGTATTACTT GTAAGGGGGA 2580

CTATCGTCTT TCGGGAAATG TGTGTTCTGA GACTTTCTGC TTTGCTCATT TAGAGGGAGA 2640

GGGTTTGGTT AAAAACAACA CAGTCAAGTC CCCTAGTCGC TGGACCAGCG AGTCTGGCTT 2700

TTCCATGATA ACTTGTGCAG ACGGCAGGGT TACGCCTTTG GGTTCCCTCC ACATTGTGGG 2760

CAACCGTTGT AGGCGTTGGC CAACCATGCA GGGGAATGTG TTTATCATGT CTAAACTGTA 2820

TCTGGGCAAC AGAATAGGGA CTGTAGCCCT GCCCCAGTGT GCTTTCTACA AGTCCAGCAT 2880

TTGTTTGGAG GAGAGGGCGA CAAACAAGCT GGTCTTGGCT TGTGCTTTTG AGAATAATGT 2940

ACTGGTGTAC AAAGTGCTGA GACGGGAGAG TCCCTCAACC GTGAAAATGT GTGTTTGTGG 3000

GACTTCTCAT TATGCAAAGC CTTTGACACT GGCAATTATT TCTTCAGATA TTCGGGCTAA 3060

TCGATACATG TACACTGTGG ACTCAACAGA GTTCACTTCT GACGAGGATT AAAAGTGGGC 3120

GGGGCCAAGA GGGGTATAAA TAGGTGGGGA GGTTGAGGGG AGCCGTAGTT TCTGTTTTTC 3180

CCAGACTGGG GGGGACAACA TGGCCGAGGA AGGGCGCATT TATGTGCCTT ATGTAACTGC 3240

CCGCCTGCCC AAGTGGTCGG GTTCGGTGCA GGATAAGACG GGCTCGAACA TGTTGGGGGG 3300

TGTGGTACTC CCTCCTAATT CACAGGCGCA CCGGACGGAG ACCGTGGGCA CTGAGGCCAC 3360

CAGAGACAAC CTGCACGCCG AGGGAGCGCG TCGTCCTGAG GATCAGACGC CCTACATGAT 3420

CTTGGTGGAG GACTCTCTGG GAGGTTTGAA GAGGCGAATG GACTTGCTGG AAGAATCTAA 3480

TCAGCAGCTG CTGGCAACTC TCAACCGTCT CCGTACAGGA CTCGCTGCCT ATGTGCAGGC 3540

TAACCTTGTG GGCGGCCAAG TTAACCCCTT TGTTTAAATA AAAATACACT CATACAGTTT 3600

ATTATGCTGT CAATAAAATT CTTTATTTTT CCTGTGATAA TACCGTGTCC AGCGTGCTCT 3660

FIGURE 1C

GTCAATAAGG GTCCTATGCA TCCTGAGAAG GGCCTCATAT ACCATGGCAT GAATATTAAG 3720

ATACATGGGC ATAAGGCCCT CAGAAGGGTT GAGGTAGAGC CACTGCAGAC TTTCGTGGGG 3780

AGGTAAGGTG TTGTAAATAA TCCAGTCATA CTGACTGTGC TGGGCGTGGA AGGAAAAGAT 3840

GTCTTTTAGA AGAAGGGTGA TTGGCAAAGG GAGGCTCTTA GTGTAGGTAT TGATAAATCT 3900

GTTCAGTTGG GAGGGATGCA TTCGGGGGCT AATAAGGTGG AGTTTAGCCT GAATCTTAAG 3960

GTTGGCAATG TTGCCCCCTA GGTCTTTGCG AGGATTCATG TTGTGCAGTA CCACAAAAAC 4020

AGAGTAGCCT GTGCATTTGG GGAATTTATC ATGAAGCTTG GAGGGGAAGG CATGAAAAAA 4080

TTTTGAGATG GCTTTATGGC GCCCCAGGTC TTCCATGCAT TCGTCCATAA TAATAGCAAT 4140

AGGCCCGGTT TTGGCTGCCT GGGCAAACAC GTTCTGAGGG TGGGCGACAT CATAGTTGTA 4200

GTCCATGGTC AGGTCTTCAT AGGACATGAT CTTAAAGGCA GGTTTTAGGG TGCTGCTTTG 4260

AGGAACCAGA GTTCCTGTGG GGCCGGGGGT GTAGTTCCCT TCACAGATTT GGGTCTCCCA 4320

AGCAAGCAGT TCTTGCGGGG GTATCATGTC AACTTGGGGG ACTATAAAAA AAACAGTTTC 4380

GGGAGGTGGT TGAATGAGGC CCGTAGACAT AAGGTTTCTG AGGAGCTGGG ATTTTCCACA 4440

ACCGGTTGGT CCGTAGACCA CCCCAATAAC GGGTTGCATG GTAAAGTTTA AAGATTTGCA 4500

TGAACCGTCA GGGCCCAGAT ATGGCATGGT GGCATTCATG GCATCTCTTA TCGCCTGATT 4560

ATAGTCTGAG AGGGCATTGA GTAGGGTGGC GCCCCCCATA GCCAGTAGCT CGTCCAAGGA 4620

AGAAAAGTGT CTAAGAGGTT TGAGGCCTTC AGCCATGGGC ATGGACTCTA AGCACTGTTG 4680

CATGAGAGCA CATTTGTCCC AAAGCTCAGA GACGTGGTCT AGTACATCTC CATCCAGCAT 4740

AGCTCTTTGT TTCTTGGGTT GGGGTGGCTG TTGCTGTAGG GGGCGAGACG GTGACGGTCG 4800

ATGGCCCGCA GGGTGCGGTC TTTCCAGGGC CTGAGCGTCC TCGCCAGGGT CGTCTCGGTG 4860

ACCGTGAAGG GCTGCTGATG CGTCTGTCTG CTGACCAGCG AGCGCCTCAG GCTGAGCCTG 4920

CTGGTGCCGA ACTTTTCGTC GCCTAGCTGT TCAGTGGAAT AATAACAAGT CACCAGAAGG 4980

TCGTAGGAGA GTTGTGAGGT GGCATGGCCT TTGCTCGAAG TTTGCCAGAA CTCTCGGCGG 5040

CGGCAGCTTG GGCAGTAGAT GTTTTTAAGG GCATATAGTT TGGGGGCTAA GAAGACAGAT 5100

TCCTGGCTGT GGGCGTCTCC GTGGCAGCGG GGGCACTGGG TCTCGCATTC CACAAGCCAA 5160

GTCAGCTGAG GGTTGGTGGG ATCAAAGACC AGAGGACGGT TATTACCTTT CAGGCGGTGC 5220

TTGCCTCGGG TGTCCATGAG TTCCTTTCCC CTTTGGGTGA GAAACATGCT GTCCGTGTCT 5280

CCGTAGACAA ATTTGAGAAT CCGGTCTTCT AGGGGAGTGC CTCTGTCTTC TAAATAGAGG 5340

ATGTCTGCCC ATTCAGAGAC AAAGGCTCTA GTCCACGCGA GGACAAATGA AGCTATGTGT 5400

GAGGGGTATC TGTTATTAAA TATGAGAGAG GATTTTTTTT GCAAAGTATG CAGGCACAGG 5460

GCTGAGTCAT CAGCTTCCAG AAAGGTGATT GGTTTGTAAG TGTATGTCAC GTGATGGTTC 5520

FIGURE 1D

```
TGGGGGTCTC CCAGGGTATA AAAGGGGGCG TCTTCGTCTG AGGAGCTATT GCTAGTGGGT   5580
GTGCACTGAC GGTGCTTCCG CGTGGCATCC GTTTGCTGCT TGACGGGTGA GTAGGTGATT   5640
TTTAGCTCTG CCATGACAGA GGAGCTCAGG TTGTGAGTTT CCACGAAGGC GGTGCTTTTG   5700
ATGTCGTAGG TGCCGTCTGA AATGCCTCTA ACATATTTGT CTTCCATTTG GTCAGAAAAG   5760
ACAGTGACTC TGTTGTCTAG CTTAGTGGCA AAGCTGCCAT ACAGGGCATT GGACAGCAGT   5820
TTGGCAATGC TTCTGAGAGT TTGGTTTTTC TCTTTATCCG CCCTTTCCTT GGGCGCAATG   5880
TTAAGTTGCA CGTAGTCTCT AGCCAGACAC TCCCACTGGG GAAATACTGT GGTGCGGGGG   5940
TCGTTGAGAA TTTGGACTCT CCAGCCGCGG TTATGAAGCG TGATGGCATC CAAACAAGTT   6000
ACCACTTCCC CCCGTAGTGT CTCGTTGGTC CAGCAGAGGC GACCTCCTTT TCTGGAGCAG   6060
AAGGGCGGTA TAACGTCCAA GAATGCTTCT GGGGGTGGGT CTGCATCAAT GGTGAATATC   6120
GCGGGCAGTA GGGTGCGATC AAAATAGTCA ATGGGTCTGT GCAACTGGGT TAGGCGGTCT   6180
TGCCAGTTTT TAATTGCAAG CGCTCGATCA AAGGGGTTCA AAGGTTTTCC CGCTGGGAAA   6240
GGATGGGTGA GGGCGCTGGC ATACATGCCG CAGATGTCAT ACACATAGAT GGCTTCTGTT   6300
AGGACGCCTA TGTAGGTAGG ATAGCATCGG CCGCCCCGAA TACTTTCTCT AACGTAATCA   6360
TACATTTCAT TGGAAGGGGC TAGTAGAAAG TTGCCCAGAG AGCTCCTGTT GGGACGCTGG   6420
GATCGGTAGA CTACCTGTCT GAAGATGGCA TGGGAATTGG AGCTGATGGT GGGCCTTTGG   6480
AGGACATTGA AATTGCAGTG GGGCAGCCCC ACTGACGTGT GAACAAAGTC CAAATAAGAT   6540
GCTTGGAGTT TTTTAACCAA TTCGGCCGTA ACCAGCACGT CCATAGCACA GTAGTCCAAG   6600
GTGCGTTGCA CAATATCATA GGCACCTGAA TTCTCTTGCA GCCAGAGACT CTTATTGAGA   6660
AGGTACTCCT CGTCGCTGGA CCAGTAGTCC CTCTGAGGAA AAGAATCTGC GTCGGTTCGG   6720
TAGGTACCTA ACATGTAAAA TTCATTTACA GCTTTGTAAG GGCAGCAGCC TTTTTCCACG   6780
GGTAAAGCGT AAGCGGCAGC TGCGTTCCTG AGACTCGTGT GCGTGAGAGC AAAGGTATCT   6840
CGGACCATGA ACTTCACAAA CTGAAATTTA TAGTCTGCTG AGGTGGGAGT GCCTTCCTCC   6900
CAGTCTTTGA AGTCTTTTCG AGCAGCATGT GTGGGGTTAG GCAGAGCAAA AGTTAAGTCA   6960
TTGAAAAGAA TCTTGCCACA ACGAGGCATG AAATTTCTAC TGACTTTAAA AGCAGCTGGA   7020
ATACCTTGTT TGTTGTTAAT GACTTGTGCG GCTAGAACAA TCTCATCAAA GCCGTTTATG   7080
TTGTGCCCTA CGACATAGAC TTCCAAGAAA GTCGGTTGCC CTTTGAGTTC AAGCGTACAC   7140
AGTTCCTCGA AAGGAATGTC GCTGGCATGG ACATAGCCCA GTTTGAGGCA GAGGTTTTCT   7200
AAGCACGGAT TATCTGCCAG GAACTGGCGC CAAAGCAAAG TGCTGGCAGC TTCTTGAAGG   7260
GCATCCCGAT ACTGTTTAAA CAAGCTGCCT ACTTTGTTTC TTTGCGGGTT GAGGTAGTAG   7320
```

FIGURE 1E

AAGGTATTTG CTTGCTTTGG CCAGCTTGAC CACTTTTGCT TTTTAGCTAT GTTAACAGCC 7380

TGTTCGCATA GCTGCGCGTC ACCAAACAAA GTAAACACGA GCATAAAAGG CATGAGTTGC 7440

TTGCCAAAGC TACCGTGCCA AGTGTATGTT TCCACATCAT AGACGACAAA GAGGCGCCGG 7500

GTGTCGGGGT GAGCGGCCCA GGGGAAAAAC TTTATTTCTT CCCACCAGTC CGAAGATTGG 7560

GTGTTTATGT GGTGAAAGTA AAAGTCCCGG CGGCGAGTGC TGCAGGTGTG CGTCTGCTTA 7620

AAATACGAAC CGCAGTCGGC ACATCGCTGG ACCTCTGCGA TGGTGTCTAT GAGATAGAGC 7680

TTTCTCTTGT GAATAAGAAA GTTGAGGGGG AAGGGAAGGC GCGGCCTGTC AGCGCGGGCC 7740

GGGATGCTTG TAATTTTCAG CTTCCCCTTG TATGTTTTGT AAACGCACAT ATTTGCGTTG 7800

CAGAACCGGA CGAGCGTGTC TTGGAATGAA AGGATATTTT CTGGTTTTAA ATCAAATGGG 7860

CAGTGCTCCA AGTGCAGTTC AAAAAGGTTT CGGAGACTGC TGGAAACGTC TGCGTGATAC 7920

TTGACTTCCA GGGTGGTCCC GTCTTCAGTC TGACCGTGCA GCCGTAGGGT ACTGCGTTTG 7980

GCGACCAGGG GCCCCCTTGG GGCTTTCTTT AAAGGGGACG TCGAGGGCCG AGGGGCGGCC 8040

TTTGCCTTTC GGGCCTGAGG GGCGGTAGCT GGACCGGATC GTTGAGTTCG GGCATGGGTT 8100

GCAGCTGTTG GCGCAGGTCT GATGCGTGCT GCACGACTCT GCGGTTGATT CTCTGAATCT 8160

CCGGGTGTTG GGTGAATGCT ACTGGCCCCG TCACTTTGAA CCTGAAAGAG AGGTCGACAG 8220

AGTTAATAGA TGCATCGTTA AGCTCCGCCT GTCTAATAAT TTCTTCCACG TCACCGCTGT 8280

GGTCTCGGTA AGCAATGTCT GTCATAAACC GTTCGATCTC TTCCTCGTCC AGTTCTCCGC 8340

GACCAGCTCG GTGGACCGTG GCTGCCAAGT CCGTGCTAAT GCGTCGCATG AGCTGGGAAA 8400

AGGCATTGGT TCCCGGTTCA TTCCACACTC TGCTGTATAT AACAGCGCCA TCTTCGTCTC 8460

GGGCTCGCAT GACCACCTGG CCCAAGTTTA GCTCCACGTG GCGAGCAAAG ACGGGGCTGA 8520

GGCGGAGGTG GTGGTGCAGA TAATTGAGAG TGGTGGCTAT GTGCTCCACG ATGAAGAAGT 8580

AGATGACCCA TCTGCGGATG GTCGACTCGT TAATGTTGCC CTCTCGCTCC AGCATGTTTA 8640

TGGCTTCGTA AAAGTCCACA GCGAAGTTAA AAAACTGCTC GTTGCGGGCG GAGACTGTCA 8700

GCTCTTCTTG CAGGAGACGA ATGACTTCGG CTACGGCGGC GCGGACTTCT TCGGCAAAGG 8760

AGCGCGGCGG CACGTCCTCC TCCTCCTCTT CTTCCCCCTC CAGCGGGGGC ATCTCCAGCT 8820

CTACCGGTTC CGGGCTGGGG GACAGGGAAG GCGGTGCGGG CCGAACGACC CGTCGGCGTC 8880

GGGTGGGCAA GGGGAGACTC TCTATGAATC GCTGCACCAT CTCGCCCCGG CGTATCCGCA 8940

TCTCCTGGGT AACGGCACGC CCGTGTTCTC GGGGTCGGAG CTCAAAAGCT CCGCCCCGCA 9000

GTTCGGTCAG AGGCCGCGCC GCGGGCTGGG GCAGGCTGAG TGCGTCAATA ACATGCGCCA 9060

CCACTCTCTC CGTAGAGGCG GCTGTTTCGA ACCGAAGAGA CTGAGCATCC ACGGGATCGC 9120

TGAAGCGTTG CACAAAAGCT TCTAACCAGT CGCAGTCACA AGGTAGGCTG AGCATAGGTG 9180

FIGURE 1F

AGGCTCGCTC GGTGTTGTTT CTGTTTGGCG GCGGGTGGCT GAGGAGAAAA TTAAAGTACG 9240

CGCACCGCAG GCGCCGGATG GTTGTCAGTA TGATGAGATC CCTGCGACCC GCTTGTTGGA 9300

TTCTGATGCG GTTTGCAAAG CCCCAGGCTT GGTCTTGGCA TCGCCCAGGT TCATGCACTG 9360

TTCTTGGAGG AATCTCTCTA CGGGCACGTT GCGGCGCTGC GGGGGCAGGG TCAGCCATTT 9420

CGGTGCGTCC AAACCCACGC AATGGTTGGA TGAGAGCCAA GTCCGCTACT ACGCGCTCTG 9480

CTAGGACGGC TTGCTGGATC TGCCGCAGCG TTTCATCAAA GTTTTCCAAG TCAATGAAGC 9540

GGTCGTAGGG GCCCGCGTTT ATGGTGTAGG AGCAGTTTGC CATGGTGGAC CAGTCCACAA 9600

TCTGCTGATC TACCCGCACC GTTTCTCGGT ACACCAGTCG GCTATAGGCT CGCGTCTCGA 9660

AAACATAGTC GTTGCAAACG CGCACCACGT ATTGGTAGCC GATTAGGAAG TGCGGCGGCG 9720

GGTATAAGTA GAGCGGCCAG TTTTGCGTGG CCGGCTGTCT GGCGCCCAGA TTCCGTAGCA 9780

TGAGTGTGGG GTATCGGTAC ACGTGACGCG ACATCCAGGA GATGCCCGCG GCCGAAATGG 9840

CGGCCCTGGC GTACTCCCGG GCCCGGTTCC ATATATTCCT GAGAGGACGA AAGATTCCAT 9900

GGTGTGCAGG GTCTGCCCCG TAAGACGCGC GCAATCTCTC GCGCTCTGCA AAAAACATAC 9960

AGATGAAACA TTTTTGGGGC TTTTCAGATG ATGCATCCCG CTTTACGGCA AATGAAGCCC 10020

AGATCCGCGG CAGTGGCGGG GGTTCCTGCT GCGGCCGCCG GCGCGAGCGT TGACTCAGGC 10080

GGTACTACCG CGCCCCCTGG TGTCGAGTGC GGCGAGGGGG AAGGGTTAGC TCGGCTGTAC 10140

GCGCACCCGG ACACACACCC GCGCGTGTGC GTGAAGCGCG ATGCGGCGGA GGCGTACGTT 10200

CCCCGGGAGA ACTTATTCCG CGACCGCAGC GGGGAGGAAC CCGAAGGGAG CCGAGACCTA 10260

AAGTACAAGG CCGGTCGGCA GTTGCGCGCC GGCATGCCCC GAAAGCGGGT GCTGACCGAA 10320

GGGGACTTTG AGGTGGATGA GCGCACTGGC ATCAGCTCAG CCAAAGCCCA CATGGAGGCG 10380

GCCGATCTAG TGCGGGCTTA CGAGCAAACG GTGAAGCAAG AGGCTAATTT TCAAAAGTCA 10440

TTTAATAACC ACGTGCGGAC ACTGATCTCC CGCGAGGAGA CCACCCTGGG TTTGATGCAC 10500

TTGTGGGACT TTGCGGAGGC ATACGCGCAG AACCCCGGCA GCAAGACCCT TACGGCCCAA 10560

GTCTTTCTCA TCGTGCAGCA CTTGCAAGAT GAGGGCATTT TTGGGGAAGC TTTCTTAAGC 10620

ATAGCAGAGC CCGAGGGACG ATGGATGCTA GATCTGCTAA ACATATTGCA GTCCATTGTG 10680

GTGCAAGAGC GCCAGCTTTC GCTATCTGAA AAGGTAGCCG CGGTGAACTA CTCCGTAGTT 10740

ACCCTGGGCA AACATTATGC CCGCAAGATC TTTAAGAGCC CCTTTGTGCC GCTTGACAAG 10800

GAGGTGAAGA TCAGTACATT TTATATGCGC GCGGTGCTTA AGGTCCTGGG TCTAAGTCAC 10860

GACCTGGGCA TGTACAGAAA CGAAAAGGTG GAGAAGCTAG CTAGCATAGG CAGGCGTTCG 10920

GGAGATGAGC GACGCGGAGC TGCTGTTCAA CCTCCGCCGC GCACTAACCA CTGGCGATTC 10980

FIGURE 1G

TGAAGCATTC GATGAAGGCG GGGACTTTAC CTGGGCTCCG CCAACTCGCG CGACCGCGGC 11040

GGCCGCTTTG CCGGGGCCCG AGTTTGAGAG TGAAGAGACG GACGATGAAG TCGACGAATG 11100

AGTGATGCGG ACCCCCGTAT CTTTCAGCTG GTCAGTCGGC AAGAGACCGT AGCCATGGCG 11160

GAAGCGCCCC GAAGCCTGGG CCCCGCCCCT TCCAATCCTA GTTTGCAGGC TTTATTCCAA 11220

AGCCAGCCCA GCGCCGAGCA GGAGTGGCAC GGCGTGCTGG AGAGAGTCAT GGCCCTTAAC 11280

AAAAATGGAG ACTTTGGCTC GCAGCCCCAG GCGAACCGGT TTGGAGCCAT CCTCGAAGCC 11340

GTGGTGCCCC CGCGCTCCGA TCCCACCCAT GAAAAAGTGC TAGCTATTGT GAATGCGCTC 11400

TTGGAGACTC AGGCCATCCG TCGCGATGAG GCCGGACAGA TGTACACCGC GCTGTTGCAG 11460

CGGGTGGCCA GATACAACAG TGTGAATGTG CAGGGCAATT TGGACAGGCT GATTCAGGAC 11520

GTGAAGGAGG CTCTGGCGCA GCGCGAGCGC ACCGGGCCGG GGGCCGGCCT AGGGTCTGTG 11580

GTAGCCTTGA ATGCCTTCCT GAGCACACAG CCAGCGGTGG TGGAGAGGGG CCAGGAGAAC 11640

TATGTGGCCT TTGTGAGCGC CTTAAAACTC ATGGTGACCG AGGCGCCGCA GTCTGAGGTT 11700

TACCAGGCCG GACCTAGTTT CTTTTTTCAA ACCAGCCGGC ACGGTTCGCA GACGGTAAAC 11760

CTCAGTCAGG CCTTTGATAA CTTGCGACCC CTCTGGGGCG TGCGCGCGCC AGTACACGAG 11820

CGTACTACCA TCTCCTCTCT GCTCACACCA AACACCCGCT TGCTCTTGCT CCTCATTGCG 11880

CCCTTTACGG ACAGCGTGGG CATATCCCGG GACAGTTACC TGGGGCATCT GCTGACCCTT 11940

TACCGGGAGA CCATAGGTAA CACTCGAGTT GATGAGACCA CGTACAACGA GATCACGGAA 12000

GTGAGTCGGG CCCTGGGCGC CGAAGACGCG TCTAACTTGC AAGCCACTCT CAACTACTTA 12060

CTCACAAATA AGCAGAGCAA GTTGCCACAG GAGTTTTCTC TGAGTCCCGA AGAGGAGCGG 12120

GTGCTGCGCT ACGTGCAGCA ATCTGTCAGT TTATTTTTAA TGCAGGATGG ACACACGGCC 12180

ACCACTGCTC TAGATCAGGC TGCGGCCAAC ATAGCGCCCT CGTTTTACGC GTCCCACCGC 12240

GACTTTATAA ACCGACTGAT GGACTATTTC CAGCGAGCTG CGGCTATGGC CCCTGACTAC 12300

TTTTTACAGG CTGTTATGAA TCCCCACTGG CTCCCGCCGC CGGGTTTCTT TACTCAGGAG 12360

TTTGACTTTC CGGAGCCCAA CGGAGGCTTC CTGTGGGATG ATTTGGACAG CGCGCTCCTA 12420

CGCGCGCACG TAAAAGAAGA GGAGGATCAA GGAGCTGTGG GCGGCACGCC GGCGGCTTCG 12480

GCGCCCGCGT CTCGCGCGCA CACACCACCG CCGCCGCCCG GTGCCGCGGA CCTCTTTGCT 12540

CCTAACGCCT TCCGCAATGT GCAAAATAAC GGCGTGGATG AACTTATTGA CGGCTTAAGC 12600

AGATGGAAGA CTTACGCCCA GGAGAGGCAG GAAGTCGTTG AGCGGCACAG GCGCAGAGAG 12660

GCGCGTCGCC GGGCGCCCCA GGCGCGTCTA GAGTCGAGCG ATGATGACGA CAGCGACCTA 12720

GGGCCGTTTC TACGGGGCAC GGGGCACCTC GTTCACAACC AGTTTATGCA TCTGAAGCCC 12780

CGGGGTCCCC GCCAGTTTTG GTAACCGCAC TGTATTAAGC TGTAAGTCCT CTCATTTGAC 12840

FIGURE 1H

```
ACTTACCAAA GCCATGGTCT TGCTTCGCCT CTGACACTTT CTCTCCCCCC ACACGCGGCA  12900
CCCTACAGCC TAGGGGCGAT GCTCCAGCCC GAACTGCAGC CAATTCCGCT GTCCCGCCGC  12960
CGGCTTATGA GGCGGTGGTG GCTGGGGCCT TCCAGACGCT TTCTCTTCGA CGAGATCCAC  13020
GTCCCGCCGC GATATGCTGC CGCGTCTGCG GGGAGAAACA GTATCCGTTA TTCCATGCTG  13080
CCCCCGTTGT ATGACACCAC GAAGATATAC CTTATCGACA ACAAATCTTC AGACATCCAA  13140
ACTCTGAATT ACCAAAACGA CCACTCAGAT TACCTCACTA CCATCGTGCA GAACAGCGAC  13200
TTCACGCCCC TGGAGGCTAG CAACCACAGC ATCGAGCTAG ACGAGCGGTC CCGCTGGGGC  13260
GGAAACCTTA AAACCATCCT TTATACAAAC CTGCCTAATA TCACCCAGCA CATGTTTTCT  13320
AACTCTTTTC GGGTAAAGAT GATGGCCTCA AAAAAAGACG GCGTGCCCCA GTACGAGTGG  13380
TTCCCCCTAA GGCTGCCCGA GGGTAACTTT TCTGAGACTA TGGTCATTGA CCTCATGAAC  13440
AATGCCATCG TAGAGCTGTA CTTGGCTTTG GGGCGCCAGG AGGGCGTGAA GGAAGAGGAC  13500
ATCGGGGTAA AGATCGATAC GCGCAACTTT AGTCTGGGCT ATGACCCGCA GACCCAGTTA  13560
GTGACGCCCG GCGTATACAC CAATGAAGCT ATGCATGCGG ACATCGTGTT GCTGCCGGGC  13620
TGTGCTATAG ACTTTACGCA CTCCCGATTA AACAACCTCT TGGGCATACG CAAGCGTTTT  13680
CCGTACCAAG AGGGCTTCGT CATCTCCTAT GAGGACCTTA AGGGGGGTAA CATCCCCGCT  13740
TTGATGGACG TGGAGGAGTT TAACAAGAGC AAGACGGTTC GAGCTTTGCG GGAGGACCCC  13800
AAGGGGCGCA GTTATCACGT GGGCGAAGAC CCAGAAGCCA GAGAAAACGA AACCGCCTAC  13860
CGCAGCTGGT ACCTGGCTTA CAATTACGGG GACCCAGAAA AAGGGGTGCG GGCCACCACA  13920
CTGCTGACTA CCGGCGACGT GACCTGCGGG GTGGAACAGA TCTACTGGAG CTTGCCGGAC  13980
ATGGCACTGG ACCCAGTCAC TTTCAAGGCT TCGCTGAAAA CTAGCAATTA CCCCGTGGTG  14040
GGCACAGAAC TTTTGCCACT GGTGCCGCGT AGCTTTTATA ACGCTCAGGC TGTGTACTCA  14100
CAGTGGATAC AAGAAAAAAC TAACCAGACC CACGTTTTCA ATCGCTTTCC CGAAAATCAG  14160
ATCTTGGTGC GGCCCCCTGC GCCTACCATC ACGTCCATAA GTGAAAATAA GCCCAGCTTG  14220
ACAGATCACG GAATCGTGCC GCTCCGGAAC CGCTTGGGGG GCGTGCAACG TGTGACTTTG  14280
ACTGACGCGC GGCGAAGATC CTGCCCCTAC GTCTACAAGA GCTTAGGCAT TGTGACGCCG  14340
CAAGTGCTAT CTAGCCGCAC GTTTTAAGCA GACAGGGGCA CAGCAGCCGT TTTTTTTTTT  14400
TTTTTTTCGC TCCACCAGGG ACTGTCAGGA ACATGGCCAT TCTAATCTCT CCTAGCAATA  14460
ACACGGGCTG GGGCCTGGGA TGCAATAAGA TGTACGGGGG CGCTCGCATA CGTTCAGACT  14520
TGCATCCAGT GAAGGTGCGG TCGCATTATC GGGCCGCCTG GGGCAGCCGC ACCGGTCGGG  14580
TGGGTCGCCG CGCAACCGCA GCTTTAGCCG ATGCCGTCGC GGCCACCGGT GATCCGGTGG  14640
```

FIGURE 11

CCGACACAAT CGAGGCGGTG GTGGCTGACG CCCGCCAGTA CCGGCGCCGC AGACGGCGAG 14700

GGGTGCGCCG AGTCAGAAGG TTGCGTCGGA GCCCCCGCAC TGCCCTGCAG CGACGGGTTC 14760

GTAGCGTACG CCGACAAGTG GCGAGGGCCC GCAGGGTGGG CCGGCGCGCG GCCGCTATCG 14820

CAGCAGACGC GGCCATGGCC ATGGCGGCGC CAGCTCGGCG ACGCCGTAAC ATCTACTGGG 14880

TACGCGATGC GGCAACCGGA GCCCGCGTTC CGGTGACAAC CCGGCCTACG GTCAGCAACA 14940

CCGTTTGAAA TGTCTGCTAC TTTTTTTTGC TTCAATAAAA GCCCGCCGAC TGATCAGCCA 15000

CACCTTGTCA CGCAGAATTC TTTCAAACCA TTGCGCTCTC AGCGCGCGCG CCGATAAACC 15060

CACTGTGATG GCCTCCTCTC GGTTGATTAA AGAAGAAATG TTAGACATCG TGGCGCCTGA 15120

GATTTACAAG CGCAAACGGC CCAGGCGAGA ACGCGCAGCA CCGTATGCTG TGAAGCAGGA 15180

GGAGAAGCCT TTAGTAAAGG CGGAGCGCAA AATTAAGCGC GGCTCCAGAA AGCGGGCCTT 15240

GTCAGGCGTT GACGTTCCTC TGCCCGATGA CGGCTTTGAG GACGACGAGC CCCACATAGA 15300

ATTTGTGTCT GCGCCGCGTC GGCCCTACCA GTGGAAGGGC AGGCGGGTGC GCCGGGTTTT 15360

GCGTCCCGGC GTGGCCGTTA GTTTCACGCC CGGCGCGCGC TCCCTCCGTC CGAGTTCCAA 15420

GCGGGTGTAT GACGAGGTGT ACGCAGACGA CGACTTCTTA GAAGCGGCCG CGGCCCGTGA 15480

GGGGGAGTTT GCTTACGGAA AGCGGGGACG CGAGGCGGCC CAGGCCCAGC TGCTACCGGC 15540

TGTGGCCGTG CCGGAACCGA CTTACGTAGT TTTGGATGAG AGCAACCCCA CCCCGAGCTA 15600

CAAGCCTGTA ACCGAGCAGA AAGTTATTCT TTCCCGCAAG CGGGGTGTGG GGAAGGTAGA 15660

GCCTACCATC CAGGTTTTAG CTAGCAAGAA GCGGCGCATG GCCGAGAATG AGGATGACCG 15720

CGGGGCCGGC TCCGTGGCCG AAGTGCAGAT GCGAGAAGTT AAACCGGTAA CCGCTGCCTT 15780

GGGTATTCAG ACCGTGGATG TTAGCGTGCC CGACCACAGC ACTCCCATGG AGGTCGTGCA 15840

GAGTCTCAGT CGGGCGGCTC AAGTAGCTCA ACGCCTGACC CAACAACAGG TGCGGCCTTC 15900

GGCTAAGATT AAAGTGGAGG CCATGGATCT TTCTGCTCCC GTAGACGCAA AGCCTCTTGA 15960

CTTAAAACCC GTGGACGTAA AGCCGACCCC GACCTTCGTG CTTCCCAGCT TTCGTTCACT 16020

CAGCACCCAA ACTGACTCTT TGCCCGCGGC AGTGGTCGTG CCGCGCAAGC CCCGCGTGCA 16080

CCGTGCTACT AGGCGTACTG CGCGCGGCTT GCTGCCCTAT TACCGCCTGC ATCCTAGCAT 16140

CACGCCGACA CCGGGTTACC GAGGATCTGT CTACACGAGC TCGGGTGTGC GCCTGCCCGC 16200

CGTCCGGGCG CCGCCGTCGC CGCCGTACCC GCAGGGCGAC TCCCCGCCTC AGCGCTGCCG 16260

CGGCCGCGGC GCTGCTGCCC GGCGTGCGCT ATCACCCTAG CATCCGCCAA GCGGCCACAG 16320

TAACCCGGCT CCGCCGTTAA GCGCTGTGAA ACTGCAACAA CAACAACAAA AATAAAAAAA 16380

AGTCTCCGCT CCACTGTGCA CCGTTGTCCA TCGGCTAATA AAGTCCCGCT TTGTGCGCCG 16440

CAGGAACCAC TATCCGTAAC CTGCGAAAAT GAGTCCCCGC GGAAATCTGA CTTACAGACT 16500

FIGURE 1J

GAGAATACCG GTCGCCCTCA GTGGCCGGCG CCGGCGCCGA ACAGGCTTGC GAGGAGGGTC 16560

TGCGTACCTG CTCGGCCGCC GCAGAAGGCG CGCGGGCGGC GGCCGCCTGC GCGGGGGCTT 16620

CCTTCCCCTC CTGGCTCCCA TCATTGCAGC CGCCATCGGC GCAATCCCCG GCATCGCATC 16680

AGTGGCCATT CAGGCGGCCC ACAACAAATA GGGACAGTGT AAAGAAAGCT CAATCTCAAT 16740

AAAACAAACC GCTCGATGTG CATAACGCTC TCGGCCTGCA ACTTCTGCTG CTTACGTCTT 16800

TGACCAAAGT CACTACTGTT TTCCTTTTAC CCAGAGCCGG CGCCAGCCCC ACACAGCTTG 16860

TTAACACGCC ATGGACGAAT ACAATTACGC GGCTCTTGCT CCCCGGCAAG GCTCCCGACC 16920

CATGCTGAGC CAGTGGTCCG GCATCGGCAC GCACGAAATG CACGGCGGAC GTTTTAATCT 16980

GGGCAGTTTG TGGAGCGGGA TCAGGAATGT GGGCAGCGCG TTAAGAACTG GGGCTCTCGG 17040

GCCTGGCACA GCAATGCGGG CAAGCGTTGC GCGCCCAGCT GAAAAAGACG GGCTTGCAAG 17100

AAAAGATATT GAGGGCGTTA GCGCCGGTAT CCACGGAGCC GTGGATCTGG GCCGTCAGCA 17160

GCTAGAGAAA GCTATTGAGC AGCGCCTAGA GCGTCGCCCC ACCGCTGCCG GTGTGGAAGA 17220

CCTTCCGCTT CCCCCGGGAA CAGTCTTAGA AGCTGATCGT TTACCGCCCT CCTACGCCGA 17280

AGCGGTGGCT GAGCGCCCGC CGCCGGCTGA CGTTCTCCTG CCCGCATCCT CAAAGCCGCC 17340

GGTGGCGGTG GTGACCTTGC CCCCGAAAAA GAGAGTGTCT GAAGAGCCTG TGGAGGAAGT 17400

TGTGATTCGT TCCTCCGCAC CGCCGTCGTA CGACGAGGTT ATGGCACCGC AGCCGACTCT 17460

GGTAGCCGAG CAGGGCGCCA TGAAAGCAGT GCCCGTGATT AAGCCGGCTC AACCTTTTAC 17520

CCCAGCTGTG CACGAAACGC AACGCATAGT GACCAACTTG CCAATCACCA CAGCTGTGAC 17580

ACGGCGACGC GGGTGGCAGG GCACTCTGAA TGACATCGTG GGCCTCGGCG TTCGTACCGT 17640

GAAGCGCCGG CGGTGCTATT GAGGGGGCGC GCAGCGGTAA TAAAGAGAAC ATAAAAAAGC 17700

AGGATTGTGT TTTTGTTTA GCGGCCACTG ACTCTCCCTC TGTGTGACAC GTCCTCCGCC 17760

AGAGCGTGAT TGATTGACCG AGATGGCTAC CCCGTCGATG CTGCCGCAAT GGTCCTACTG 17820

CACATCGCCG GTCAGGACGC GTCCGAGTAC CTGTCCCCCG GCTTGGTGCA ATTCGCACAA 17880

GCCACCGAAT CCTACTTTAA CATTGGGAAC AAGTTTAGAA ACCCCACCGT CGCCCCGACG 17940

CACGATGTCA CCACGGAGCG TTCGCAGCGT CTGCAGCTCC GCTTCGTGCC CGTAGACCGG 18000

GAGGACACAC AGTACTCCTA CAAAACCCGC TTCCAGCTAG CCGTGGGCGA CAACCGGGTG 18060

CTGGACATGG CCAGCACGTA TTTTGACATC CGCGGTACGC TGGAGAGGGG CGCCAGTTTC 18120

AAGCCTTACA GCGGCACGGC CTACAACTCC TTTGCCCCCA ACAGTGCCCC TAACAATACG 18180

CAGTTTAGGC AGGCCAACAA CGGTCATCCT GCTCAGACCA TAGCTCAAGC TTCTTACGTG 18240

GCTACCATCG GCGGTGCCAA CAATGACTTG CAAATGGGTG TGGACGAGCG TCAGCAGCCG 18300

FIGURE 1K

GTGTATGCGA ACACTACGTA CCAGCCGGAA CCTCAGCTCG GCATTGAAGG TTGGACAGCT 18360

GGATCCATGG CGGTCATCGA TCAAGCAGGC GGGCGGGTTC TCAGGAACCC TACTCAAACT 18420

CCCTGCTACG GGTCCTATGC TAAGCCGACT AACGAGCACG GGGGCATTAC TAAAGCAAAC 18480

ACTCAGGTGG AGAAAAAGTA CTACAGAACA GGGGACAACG GTAACCCGGA AACAGTGTTT 18540

TATACTGAAG AGGCTGACGT GCTAACGCCC GACACCCACC TTGTTCACGC GGTACCGGCC 18600

GCGGATCGGG CAAAGGTGGA GGGGCTATCT CAGCACGCAG CTCCCAACAG GCCGAACTTT 18660

ATCGGCTTTC GGGACTGCTT TGTAGGCTTG ATGTATTATA ACAGCGGGGC CAACCTGGGC 18720

GTCTTAGCGG GTCAATCCTC TCAGCTGAAT GCCGTGGTAG ACCTGCAAGA CCGCAACACT 18780

GAGCTTTCCT ATCAGATGCT TCTTGCAAAC ACGACGGACA GATCCCGCTA TTTTAGCATG 18840

TGGAACCAAG CCATGGACTC GTACGACCCG GAGGTCAGGG TGATAGATAA CGTGGGCGTA 18900

GAGGACGAGA TGCCTAATTA CTGCTTTCCG TTGTCGGGGG TTCAGATTGG AAACCGTAGC 18960

CACGAGGTTC AAAGAAACCA ACAACAGTGG CAAAATGTAG CTAATAGTGA CAACAATTAC 19020

ATAGGCAAGG GGAACCTACC GGCCATGGAG ATAAATCTAG CGGCCAATCT CTGGCGTTCC 19080

TTTTTGTACA GTAATGTGGC GTTGTACTTG CCAGACAACC TTAAATTCAC CCCTCACAAC 19140

ATTCAACTCC CGCCTAACAC GAACACCTAC GAGTACATGA ACGGGCGAAT CCCCGTTAGC 19200

GGCCTTATTG ATACGTACGT AAATATAGGC ACGCGGTGGT CGCCCGATGT GATGGACAAC 19260

GTGAATCCCT TTAACCACCA CCGCAACTCG GGCCTGCGTT ACCGCTCCCA GCTGCTGGGC 19320

AACGGCCGCT TCTGCGACTT TCACATTCAG GTGCCACAAA AGTTTTTGC TATTCGAAAC 19380

CTGCTTCTCC TGCCCGGCAC GTACACTTAC GAGTGGTCCT TTAGAAAGGA CGTAAACATG 19440

ATCCTTCAGA GCACTCTGGG CAATGATCTG CGGGTCGATG GGGCCACTGT TAATATTACC 19500

AGCGTCAACC TCTACGCCAG CTTCTTTCCC ATGTCACATA ACACCGCTTC CACTTTGGAA 19560

GCTATGCTCC GCAACGACAC TAATGACCAG TCTTTTAATG ACTATCTCTC GGCGGCTAAC 19620

ATGTTGTATC CCATTCCGCC CAATGCCACC CAACTGCCCA TCCCCTCACG CAACTGGGCA 19680

GCGTTCCGTG GCTGGAGTCT CACCCGGCTA AAACAGAGGG AGACACCGGC GCTGGGGTCC 19740

CCGTTCGATC CCTATTTCAC CTATTCGGGC ACCATCCCGT ACCTGGACGG CACTTTTTAC 19800

CTCAGCCACA CCTTTCGCAA GGTGGCCATC CAGTTTGACT CTTCTGTGAC CTGGCCCGGC 19860

AATGACAGGC TTTTAACCCC TAACGAGTTC GAAATAAAAA TAAGTGTGGA CGGTGAAGGC 19920

TACAACGTGG CTCAGAGCAA TATGACTAAG GACTGGTTCC TGGTGCAGAT GCTAGCGAAT 19980

TACAACATAG GCTACCAGGG ATATCACCTG CCCCCGGACT ACAAGGACAG GACATTTTCC 20040

TTCCTGCATA ACTTCATACC CATGTGCCGA CAGGTTCCCA ACCCAGCAAC CGAGGGCTAC 20100

TTTGGACTAG GCATAGTGAA CCATAGAACA ACTCCGGCTT ATTGGTTTCG ATTCTGCCGC 20160

FIGURE 1L

GCTCCGCGCG AGGGCCACCC CTACCCCCAA CTGGCCTTAC CCCCTCATTG GGACCCACGC 20220

CATGCCCTCC GTGACCCAGA GAGAAAGTTT CTCTGCGACC GCACCCTCTG GCGAATCCCC 20280

TTCTCCTCGA ACTTCATGTC CATGGGGTCG CTCACAGATC TCGGACAGAA CCTACTGTAT 20340

GCCAATGCCG CGCATGCCCT AGACATGACT TTTGAGATGG ATCCCATCAA TGAGCCCACT 20400

CTGCTGTACG TTCTGTTTGA GGTGTTTGAC GTGGCCCGCG TTCACCAGCC CCACAGAGGC 20460

GTGATCGAAG TGGTGTACTT GAGAACGCCA TTCTCAGCCG GCAACGCTAC CACATAAGTG 20520

CCGGCTTCCC TCTCAGGCCC CGCGATGGGT TCTCGGGAAG AGGAGCTGAG ATTCATCCTT 20580

CACGATCTCG GTGTGGGGCC ATACTTCCTC GGCACTTTCG ATAAACACTT TCCGGGGTTC 20640

ATCTCCAAAG ACCGAATGAG CTGTGCCATA GTCAACACTG CCGGACGCGA AACCGGGGGC 20700

GTGCATTGGC TGGCCATGGC TTGGCACCCA GCCTCGCAGA CCTTTTACAT GTTTGACCCT 20760

TTCGGTTTCT CGGATCAAAA GCTAAAGCAA ATTTACAACT TTGAGTATCA GGGCCTCCTA 20820

AAGCGCAGCG CCCTGACTTC CACTGCTGAC CGCTGCCTGA CCCTTATTCA AAGCACTCAA 20880

TCTGTCCAGG GACCCAACAG CGCCGCCTGC GGTCTGTTCT GCTGCATGTT CCTCCACGCC 20940

TTTGTCCGCT GGCCGCTTAG GGCCATGGAC AACAATCCCA CCATGAACCT CATCCACGGA 21000

GTTCCCAACA ACATGTTGGA GAGCCCCAGC TCCCAAAATG TGTTTTTGAG AAACCAGCAA 21060

AATCTGTACC GTTTCCTAAG ACGCCACTCC CCCCATTTTG TTAAGCATGC GGCTCAAATT 21120

GAGGCTGACA CCGCCTTTGA TAAAATGTTA ACAAATTAGA CCGTGAGCCA TGATTGCAGA 21180

AGCATGTCAT TTTTTTTTTA TTGTTTAAAA TAAAAACAAC ACATAACATC TGCCGCCTGT 21240

CCTCCCGTGA TTTCTTCTGC TTTATTTGCA AATGGGGGGC ACCTTAAAAC AAAGAGTCAT 21300

CTGCATCGTA CTGATCGATG GGCAGAATAA CATTCTGATG CTGGTACTGC GGGTCCCAGC 21360

GGAATTCGGG AATGGTAATG GGGGGGCTCT GTTTAACCAG CGCGGACCAC ATCTGCTTAA 21420

CCAGCTGCAA GGCTGAAATC ATATCTGGAG CCGAAATCTT GAAATCGCAG TTTCGCTGGG 21480

CATTAGCCCG CGTCTGCCGG TACACAGGGT TACAGCACTG AAATACTAAC ACCGATGGGT 21540

GTTCTACGCT GGCCAGGAGT TTGGGATCTT CTACGAGGCT CTTATCTACC GCAGAGCCCG 21600

CGTTGATATT AAAGGGCGTT ATCTTGCATA CCTGACGGCC TAGGAGGGGC AATTGGGAGT 21660

GACCCCAGTT ACAATCACAC TTTAAAGGCA TAAGCAGATG AGTTCCGGCA CTTTGCATCC 21720

TGGGGTAACA GGCTTTCTGA AAGGTCATGA TCTGCCAGAA AGCCTGCAAA GCCTTGGGCC 21780

CCTCGCTGAA AAACATACCA CAAGACTTTG AGGTAAAGCT GCCGGCCGGC AAAGCGGCGT 21840

CAAAGTGACA GCAAGCCGCG TCTTCATTCT TTAGCTGCAC TACGTTCATA TTCCACCGGT 21900

TGGTGGTGAT CTTTGTCTTA TGCGGGGTCT CTTTTAAAGC CCGCTGCCCA TTTTCGCTGT 21960

FIGURE 1M

TCACATCCAT CTCTATCACT TGGTCTTTGG TAAGCATAGG CAGGCCATGC AGGCAGTGAA 22020

GGGCCCCGTC TCCCCCCTCG GTACACTGGT GGCGCCAGAC CACACAGCCC GTGGGGCTCC 22080

ACGAGGTCGT CCCCAGGCCT GCGACTTTTA ACACAAAATC ATACAAGAAG CGGCCCATAA 22140

TAGTTAGCAC GGTTTTCTGA GTACTGAAAG TAAGAGGCAG GTACACTTTA GACTCATTAA 22200

GCCAAGCTTG TGCAACCTTC CTAAAACACT CGAGCGTGCC AGTGTCGGGC AGCAAGGTTA 22260

AGTTTTTAAT ATCCACTTTC AAAGGCACAC ACAGCCCCAC TGCTAATTCC ATGGCCCGCT 22320

GCCAAGCAAC TTCGTCGGCT TCCAGCAAGG CCCGGCTGGC CGCCGGCAGG GCGGGAGCGG 22380

CGGCCTCAGC GGCTGGGGCT GAAGGTTTGA AAATCTTGGC GCGCTTAACG GCTGTGACAT 22440

CTTCGGCGGG GGGCTCAGCG ATCGGCGCGC GCCGTTTGCG GCTGACTTTT TTCCGGGGCG 22500

TCTCATCTAT CACTAAGGGG TTCTCGTCCC CGCTGCTGTC AGCCGAACTC GTGGCTCGCG 22560

TTAAGTCACC GCTGCGATTC ATTATTCTCT CCTAGATAAC GACAACAAAT GGCAGAGAAA 22620

GGCAGTGAAA ATCAGCGGCC AGAGAACGAC ACTGAGCTAG CAGCGGTTTC AGAAGCCCTA 22680

GGCGCGGCCG CTTCGGCCCC CTCACGTAAC TCCCCGACTG ACACGGATTC AGGGGTGGAA 22740

ATGACGCCCA CCAGCAGCCC CGAGCCGCCC GCCGCTCCCC CAAGTTCGCC TGCCGCAGCA 22800

CCTGCCCCTC AGAAGAACCA GGAGGAGCTC TCTTCCCCCG AGCCCGCGGT AGCAGCAGCG 22860

GAGCCAGAAG CCGCTTCGCG GCCCAGACCA CCCACACCCA CCGTTCAGGT CCCGCGGGAG 22920

CCGAGCGAGG ATCAACCTGA CGGACCCGCG ACGAGGCCTT CGTACGTGAG CGAGGATTGC 22980

CTCATCGGCC ATATCTCTCG CCAGGCTAAC ATTGTTAGAG ACAGCCTGGC AGACCGCTGG 23040

GAGTTAGAGC CCACCGTGTC GGCTCTCTCC GAGGCTTACG AAAAGCTCCT CTTTTGTCCC 23100

AAGGTACCAC CCAAGAAGCA AGAGAATGGC ACTTGCGAAC CTGAACCTCG CGTTAATTTT 23160

TTCCCCACCT TTGTAGTGCC CGAAACTTTA GCCACGTACC ACATCTTTTT CCAAAACCAA 23220

AAAATCCCCC TGTCTTGTCG CGCCAACCGC ACCCACACAG ACACCATCAT GCACCTCTAC 23280

TCGGGGGACT CCTTACCGTG CTTCCCCACG CTGCAGCTGG TCAACAAAAT CTTTGAAGGC 23340

TTGGGCTCAG AGGAGCGGCG CGCAGCCAAC TCGCTGAAAG ATCAAGAGGA TAACAGCGCG 23400

TTAGTTGAGC TCGAAGGGGA CAGTCCCCGA CTGGCTGTGG TTAAGCGCAC ACTGTCTTTG 23460

ACACATTTCG CCTACCCTGC CATAACACTA CCGCCTAAGG TGATGGCAGC TGTCACTGGC 23520

AGCCTCATTC ATGAATCAGC AGCGACCGCC GAACCGGAAG CTGAGGCGCT GCCAGAAGCC 23580

GAGGAGCCCG TGGTTAGTGA CCCTGAACTT GCTCGCTGGT TGGGGCTCAA CTTACAACAG 23640

GAGCCCGAGG CCACGGCCCA GGCTTTGGAA GAAAGACGCA AGATTATGTT GGCAGTATGC 23700

TTAGTCACAC TTCAGCTCGA GTGCCTGCAC AAGTTTTTTT CTTCAGAGGA TGTCATCAAA 23760

AAGCTGGGAG AGAGCCTCCA CTACGCCTTT CGCCACGGCT ACGTGCGCCA AGCCTGCTCC 23820

FIGURE 1N

ATTTCTAACG TGGAACTAAC GAACATCGTC TCATACCTGG GTATCTTGCA CGAAAACCGC 23880
TTGGGACAGA GTACCCTACA CGCCACCCTT AAAGACGAGA ACCGCAGAGA CTACATCAGA 23940
GACACAGTCT TTCTCTTTCT GGTTTATACT TGGCAGACTG CCATGGGCAT TTGGCAGCAG 24000
TGCCTCGAGA CTGAGAACGT AAAAGAACTT GAAAAGCTCT TGCAAAAAAG CAAGAGGGCT 24060
CTCTGGACGG GCTTCGACGA GCTCACCATA GCTCAAGACC TAGCTGACAT AGTGTTCCCC 24120
CCCAAATTCT TGCACACCTT GCAAGCCGGC CTGCCAGACC TTACATCCCA GAGTCTCCTT 24180
CACAACTTTC GCTCCTTCAT TTTCGAACGC TCGGGCATTC TACCCGCCAT GTGCAATGCA 24240
CTGCCCACCG ACTTCATCCC TATCAGCTAC CGGGAGTGCC CTCCAACTTT CTGGGCCTAC 24300
ACCTACCTCT TTAAACTGGC CAATTACCTC ATGTTTCACT CCGACATCGC TTACGATCGG 24360
AGCGGCCCCG GTCTCATGGA ATGCTACTGT CGCTGCAACC TGTGCAGTCC TCACCGCTGC 24420
TTGGCGACCA ACCCCGCCCT GCTCAGCGAG ACCCAAGTTA TCGGTACCTT CGAGATTCAG 24480
GGCCCTCCTG CTCAAGACGG ACAGCCGACC AAACCGCCCC TCAGGCTGAC TGCAGGTCTC 24540
TGGACTTCCG CCTACCTGCG CAAATTTGTA CCGCAAGACT TCAACGCCCA CAAAATAGCC 24600
TTCTACGAAG ACCAATCCAA AAAGCCGAAA GTGACCCCCA GCGCTTGTGT CATCACTGAA 24660
GAAAAAGTTT TAGCCCAATT GCATGAAATT AAAAAAGCGC GGGAAGACTT TCCTCTTAAA 24720
AAGGGGCACG GAGTGTATCT GGACCCTCAG ACCGGCGAGG AGCTGAACGG ACCCGCACCC 24780
TCCGCAGCTA GGAATGAAAC CCCGCAGCAT GTCGGCAGCC GGGCCTTCCG CGGCTCAGGC 24840
TTCGGAGGGC CAACAGCTGC CGCCACAGAC AGCGGGGCTG CAGCCGAGCA AGAGGGCTGT 24900
GAGGAAGGTA GTAGCTTCTC TGAATCCCAC CGCCGCCCTG GAAGACATAT CCGAGGGGGA 24960
GGAAGGCTTC CCCCTGACGG ACGAGGAAGA CGGGGACACC CTGGAGAGCG ATTTCAGCGA 25020
CTTCACGGAC GAAGACGTCG AGGAGGAGGA TATGATTTCG ATACCCCGCG ACCAGGGGCA 25080
CTCCGGCGAG CTCGAGGAGG GCGAAATTCC CGCAACGGTA GCGGCGACGG CGGTCAAGAA 25140
GGGCCAGGGC AAGAAGAGTA GGTGGGACCA GCAGGTCCGC TCCACAGCGC CTCTAAAGGG 25200
CGCTAGAGGT AAGAGGAGCT ACAGCTCCTG GAAACCCCTC AAGCCCACTA TCCTTTCATG 25260
CTTACTGCAG AGCTCCGGCA GCACTGCCTT CACTCGCCGC TATCTGCTTT TTCGCCATGG 25320
CGTGTCCGTT CCCTCCAGGG TAATTCATTA CTATAATTCT TACTGCAGAC CCGAAGCTGA 25380
CCAAAACCGC CACTCAGAGC AAAAAGAGCC GCCGGAGTGC CAGCGCGGCG CGCCCTCGCC 25440
CTCCTCCTCT TCCTCCCAAG CGTGCTCGGG CGCCCCGCCG CCCCAAAGGC CAGCGCCATC 25500
AGGCCGACGA CGCAAGCACC GAGGGCCGCG ACAAGCTTCG GGAGCTGATC TTTCCCACTC 25560
TCTATGCCAT ATTCCAACAA AGTCGCGCTC AGCGGTGTCA CCTCAAAGTG AAAAATAGAT 25620

FIGURE 1O

```
CCTTACGTTC ACTGACGCGC AGCTGCCTCT ACCACAACAA GGAGGAACAG CTCCAGCGAA  25680
CCCTAGCAGA CTCCGAGGCG CTTCTCAGTA AATACTGCTC TGCAGCTCCG ACACGATTCT  25740
CGCCGCCCTC TTATACCGAG TCTCCCGCCA AGGACGAATC CGGACCCGCC TAAACTCTCA  25800
GCATGAGCAA AGAAATTCCC ACACCTTATG TTTGGACCTT TCAACCTCAG ATGGGAGCGG  25860
CCGCAGGTGC CAGTCAAGAT TACTCGACCC GCATGAATTG GTTCAGCGCG GGACCTGATA  25920
TGATCCACGA CGTTAACAAC ATTCGTGACG CCCAAAACCG CATCCTTATG ACTCAGTCGG  25980
CCATTACCGC CACTCCCAGG AATCTGATTG ATCCCAGACA GTGGGCCGCC CACCTCATCA  26040
AACAACCGGT GGTGGGCACC ACCCACGTGG AAATGCCTCG CAACGAAGTC CTAGAACAAC  26100
ATCTGACCTC ACATGGCGCT CAAATCGCGG GCGGAGGCGC TGCGGGCGAT TACTTTAAAA  26160
GCCCCACTTC AGCTCGAACC CTTATCCCGC TCACCGCCTC CTGCTTAAGA CCAGATGGAG  26220
TCTTTCAACT AGGAGGAGGC TCGCGTTCAT CTTTCAACCC CCTGCAAACA GATTTTGCCT  26280
TCCACGCCCT GCCCTCCAGA CCGCGCCACG GGGGCATAGG ATCCAGGCAG TTTGTAGAGG  26340
AATTTGTGCC CGCCGTCTAC CTCAACCCCT ACTCGGGACC GCCGGACTCT TATCCGGACC  26400
AGTTTATACG CCACTACAAC GTGTACAGCA ACTCTGTGAG CGGTTATAGC TGAGATTGTA  26460
AGACTCTCCT ATCTGTCTCT GTGCTGCTTT TCCGCTTCAA GCCCCACAAG CATGAAGGGG  26520
TTTCTGCTCA TCTTCAGCCT GCTTGTGCAT TGTCCCCTAA TTCATGTTGG GACCATTAGC  26580
TTCTATGCTG CAAGGCCCGG GTCTGAGCCT AACGCGACTT ATGTTTGTGA CTATGGAAGC  26640
GAGTCAGATT ACAACCCCAC CACGGTTCTG TGGTTGGCTC GAGAGACCGA TGGCTCCTGG  26700
ATCTCTGTTC TTTTCCGTCA CAACGGCTCC TCAACTGCAG CCCCCGGGGT CGTCGCGCAC  26760
TTTACTGACC ACAACAGCAG CATTGTGGTG CCCCAGTATT ACCTCCTCAA CAACTCACTC  26820
TCTAAGCTCT GCTGCTCATA CCGGCACAAC GAGCGTTCTC AGTTTACCTG CAAACAAGCT  26880
GACGTCCCTA CCTGTCACGA GCCCGGCAAG CCGCTCACCC TCCGCGTCTC CCCCGCGCTG  26940
GGAACTGCCC ACCAAGCAGT CACTTGGTTT TTTCAAAATG TACCCATAGC TACTGTTTAC  27000
CGACCTTGGG GCAATGTAAC TTGGTTTTGT CCTCCCTTCA TGTGTACCTT TAATGTCAGC  27060
CTGAACTCCC TACTTATTTA CAACTTTTCT GACAAAACCG GGGGGCAATA CACAGCTCTC  27120
ATGCACTCCG GACCTGCTTC CCTCTTTCAG CTCTTTAAGC CAACGACTTG TGTCACCAAG  27180
GTGGAGGACC CGCCGTATGC CAACGACCCG GCCTCGCCTG TGTGGCGCCC ACTGCTTTTT  27240
GCCTTCGTCC TCTGCACCGC CTGCGCGGTG TTGTTAACCG CCTTCGGTCC ATCGATTCTA  27300
TCCGGTACCC GAAAGCTTAT CTCAGCCCGC TTTTGGAGTC CCGAGCCCTA TACCACCCTC  27360
CACTAACAGT CCCCCCATGG AGCCAGACGG AGTTCATGCC GAGCAGCAGT TTATCCTCAA  27420
TCAGATTTCC TGCGCCAACA CTGCCCTCCA GCGTCAAAGG GAGGAACTAG CTTCCCTTGT  27480
```

FIGURE 1P

CATGTTGCAT GCCTGTAAGC GTGGCCTCTT TTGTCCAGTC AAAACTTACA AGCTCAGCCT 27540

CAACGCCTCG GCCAGCGAGC ACAGCCTGCA CTTTGAAAAA AGTCCCTCCC GATTCACCCT 27600

GGTCAACACT CACGCCGGAG CTTCTGTGCG AGTGGCCCTA CACCACCAGG GAGCTTCCGG 27660

CAGCATCCGC TGTTCCTGTT CCCACGCCGA GTGCCTCCCC GTCCTCCTCA AGACCCTCTG 27720

TGCCTTTAAC TTTTTAGATT AGCTGAAAGC AAATATAAAA TGGTGTGCTT ACCGTAATTC 27780

TGTTTTGACT TGTGTGCTTG ATTTCTCCCC CTGCGCCGTA ATCCAGTGCC CCTCTTCAAA 27840

ACTCTCGTAC CCTATGCGAT TCGCATAGGC ATATTTTCTA AAAGCTCTGA AGTCAACATC 27900

ACTCTCAAAC ACTTCTCCGT TGTAGGTTAC TTTCATCTAC AGATAAAGTC ATCCACCGGT 27960

TAACATCATG AAGAGAAGTG TGCCCCAGGA CTTTAATCTT GTGTATCCGT ACAAGGCTAA 28020

GAGGCCCAAC ATCATGCCGC CCTTTTTTGA CCGCAATGGC TTTGTTGAAA ACCAAGAAGC 28080

CACGCTAGCC ATGCTTGTGG AAAAGCCGCT CACGTTCGAC AAGGAAGGTG CGCTGACCCT 28140

GGGCGTCGGA CGCGGCATCC GCATTAACCC CGCGGGGCTT CTGGAGACAA ACGACCTCGC 28200

GTCCGCTGTC TTCCCACCGC TGGCCTCCGA TGAGGCCGGC AACGTCACGC TCAACATGTC 28260

TGACGGGCTA TATACTAAGG ACAACAAGCT AGCTGTCAAA GTAGGTCCCG GGCTGTCCCT 28320

CGACTCCAAT AATGCTCTCC AGGTCCACAC AGGCGACGGG CTCACGGTAA CCGATGACAA 28380

GGTGTCTCTA AATACCCAAG CTCCCCTCTC GACCACCAGC GCGGGCCTCT CCCTACTTCT 28440

GGGTCCCAGC CTCCACTTAG GTGAGGAGGA ACGACTAACA GTAAACACCG GAGCGGGCCT 28500

CCAAATTAGC AATAACGCTC TGGCCGTAAA AGTAGGTTCA GGTATCACCG TAGATGCTCA 28560

AAACCAGCTC GCTGCATCCC TGGGGGACGG TCTAGAAAGC AGAGATAATA AAACTGTCGT 28620

TAAGGCTGGG CCCGGACTTA CAATAACTAA TCAAGCTCTT ACTGTTGCTA CCGGGAACGG 28680

CCTTCAGGTC AACCCGGAAG GGCAACTGCA GCTAAACATT ACTGCCGGTC AGGGCCTCAA 28740

CTTTGCAAAC AACAGCCTCG CCGTGGAGCT GGGCTCGGGC CTGCATTTTC CCCCTGGCCA 28800

AAACCAAGTA AGCCTTTATC CCGGAGATGG AATAGACATC CGAGATAATA GGGTGACTGT 28860

GCCCGCTGGG CCAGGCCTGA GAATGCTCAA CCACCAACTT GCCGTAGCTT CCGGAGACGG 28920

TTTAGAAGTC CACAGCGACA CCCTCCGGTT AAAGCTCTCC CACGGCCTGA CATTTGAAAA 28980

TGGCGCCGTA CGAGCAAAAC TAGGACCAGG ACTTGGCACA GACGACTCTG GTCGGTCCGT 29040

GGTTCGCACA GGTCGAGGAC TTAGAGTTGC AAACGGCCAA GTCCAGATCT TCAGCGGAAG 29100

AGGCACCGCC ATCGGCACTG ATAGCAGCCT CACTCTCAAC ATCCGGGCGC CCCTACAATT 29160

TTCTGGACCC GCCTTGACTG CTAGTTTGCA AGGCAGTGGT CCGATTACTT ACAACAGCAA 29220

CAATGGCACT TTCGGTCTCT CTATAGGCCC CGGAATGTGG GTAGACCAAA ACAGACTTCA 29280

FIGURE 1Q

GGTAAACCCA GGCGCTGGTT TAGTCTTCCA AGGAAACAAC CTTGTCCCAA ACCTTGCGGA 29340

TCCGCTGGCT ATTTCCGACA GCAAAATTAG TCTCAGTCTC GGTCCCGGCC TGACCCAAGC 29400

TTCCAACGCC CTGACTTTAA GTTTAGGAAA CGGGCTTGAA TTCTCCAATC AAGCCGTTGC 29460

TATAAAAGCG GGCCGGGGCT TACGCTTTGA GTCTTCCTCA CAAGCTTTAG AGAGCAGCCT 29520

CACAGTCGGA AATGGCTTAA CGCTTACCGA TACTGTGATC CGCCCCAACC TAGGGGACGG 29580

CCTAGAGGTC AGAGACAATA AAATCATTGT TAAGCTGGGC GCGAATCTTC GTTTTGAAAA 29640

CGGAGCCGTA ACCGCCGGCA CCGTTAACCC TTCTGCGCCC GAGGCACCAC CAACTCTCAC 29700

TGCAGAACCA CCCCTCCGAG CCTCCAACTC CCATCTTCAA CTGTCCCTAT CGGAGGGCTT 29760

GGTTGTGCAT AACAACGCCC TTGCTCTCCA ACTGGGAGAC GGCATGGAAG TAAATCAGCA 29820

CGGACTTACT TTAAGAGTAG GCTCGGGTTT GCAAATGCGT GACGGCATTT TAACAGTTAC 29880

ACCCAGCGGC ACTCCTATTG AGCCCAGACT GACTGCCCCA CTGACTCAGA CAGAGAATGG 29940

AATCGGGCTC GCTCTCGGCG CCGGCTTGGA ATTAGACGAG AGCGCGCTCC AAGTAAAAGG 30000

TGGGCCCGGC ATGCGCCTGA ACCCTGTAGA AAAGTATGTA ACCCTGCTCC TGGGTCCTGG 30060

CCTTAGTTTT GGGCAGCCGG CCAACAGGAC AAATTATGAT GTGCGCGTTT CTGTGGAGCC 30120

CCCCATGGTT TTCGGACAGC GTGGTCAGCT CACATTTTTA GTGGGTCACG GACTACACAT 30180

TCAAAATTCC AAACTTCAGC TCAATTTGGG ACAAGGCCTC AGAACTGACC CCGTCACCAA 30240

CCAGCTGGAA GTGCCCCTCG GTCAAGGTTT GGAAATTGCA GACGAATCCC AGGTTAGGGT 30300

TAAATTGGGC GATGGCCTGC AGTTTGATTC ACAAGCTCGC ATCACTACCG CTCCTAACAT 30360

GGTCACTGAA ACTCTGTGGA CCGGAACAGG CAGTAATGCT AATGTTACAT GGCGGGGCTA 30420

CACTGCCCCC GGCAGCAAAC TCTTTTTGAG TCTCACTCGG TTCAGCACTG GTCTAGTTTT 30480

AGGAAACATG ACTATTGACA GCAATGCATC CTTTGGGCAA TACATTAACG CGGGACACGA 30540

ACAGATCGAA TGCTTTATAT TGTTGGACAA TCAGGGTAAC CTAAAAGAAG GATCTAACTT 30600

GCAAGGCACT TGGGAAGTGA AGAACAACCC CTCTGCTTCC AAAGCTGCTT TTTGCCTTC 30660

CACCGCCCTA TACCCCATCC TCAACGAAAG CCGAGGGAGT CTTCCTGGAA AAAATCTTGT 30720

GGGCATGCAA GCCATACTGG GAGGCGGGGG CACTTGCACT GTGATAGCCA CCCTCAATGG 30780

CAGACGCAGC AACAACTATC CCGCGGGCCA GTCCATAATT TTCGTGTGGC AAGAATTCAA 30840

CACCATAGCC CGCCAACCTC TGAACCACTC TACACTTACT TTTTCTTACT GGACTTAAAT 30900

AAGTTGGAAA TAAAGAGTTA AACTGAATGT TTAAGTGCAA CAGACTTTTA TTGGTTTTGG 30960

CTCACAACAA ATTACAACAG CATAGACAAG TCATACCGGT CAAACAACAC AGGCTCTCGA 31020

AAACGGGCTA ACCGCTCCAA GAATCTGTCA CGCAGACGAG CAAGTCCTAA ATGTTTTTTC 31080

ACTCTCTTCG GGGCCAAGTT CAGCATGTAT CGGATTTTCT GCTTACACCT TTTTAGACAG 31140

FIGURE 1R

CAGTTTACAC TCATTTCCGT TAAAGGATTA CAACTGCGGC ATATGAGAAT TAAGTATATA 31200

CAACTATTGC CCTTTACCCA CAAACACTCC CCCCACGGGG TGCACCTGAT GTAGCTGCCC 31260

TCCTCAATCA TGAAAGTGCT ATTAAAGTAA ATTAAATGAA CATTATTCAC ATACACGCTT 31320

CCCACATAGG CCAAAAAAAC AGAGGACAAC TTTGACAGCT CCCGCCTGAA ATACCAATAC 31380

ACTCTATCAA ACTGCGCACC GTGCACGCAC TGCTTTACCA GGCCTTGAAA GTAAACAGCG 31440

GCGGACCGAC ACTGCAAGCT TCTAGGCTTT GGGCAGTGGC AGTGAATATA TAGCCACTCC 31500

TCCCCATGCA CGTAGTAGGA ACGCCGCTTC CCGGGAATCA CAAATGACAA GCAGTAGTCA 31560

CAGAGGCAAC TAGTCAAGTG AGCGTCCTCC TGAGGCATGA TTACCTTCCA TGGAATGGGC 31620

CAGTGAATCA TAGTGGCAAA GCCAGCTGCA TCTGGAGCGC TGCGAACCTT GGCTACATGT 31680

GGTGATTGGC GACGCAGATG GAGACAGGAC CTTGCATTCT GAAGACCACT GCAACAGCTT 31740

CTGCGTACGC TTGTATTTAC AGTACATAAA AAAGCACTTT TGCCACAGAG CGGTCTTACT 31800

CAACCGACAG CTTTTTTCTT TCTGACGCTG CCTTCTGCTA CTCAGGTAGT ACAAGTCCAA 31860

AAGAGCCAAA CGGACACTCA AATCCGGGTT ATCTCGATGC TGAAGCCAGA GTCCAAAAGT 31920

AACCACGCTA AAAGCCTGCA TCCATATTTT GTAACTGCTG TAACTCCATC CCAGAGCCGG 31980

GCACCGCACT TGGTCCACCA TAGCTGCAAA CAAACGGGAC AATTAAGGAA AGTAAAATGA 32040

GCGCTGGGGG CGGACTCTTC TCCCGTTCGT AGGAAACAGC CACGTATCAA ACACCCTTTT 32100

CAACACTGGC TCTCCAGCCG CTACTCGTTG AATTAATTTG TCCCTGTGCT CAAACAACCC 32160

ACACTGGTAA CGGTGGTCGC TAGGCAAACA TGTCAAATAG CACATAATCA TTTCCTTCAC 32220

TTTAAGCAAA CATCGACTAG CAGACACTTC ACTTAATTCA GCACAGTCAT AGCAAGGAAT 32280

GATTATACAC TTGTCATCTA ATCCACTGCC CATGTACACA TTGCCCCAGG CAAAAGTGGG 32340

CAGGGACTTT AAGAGCTGAT TGCTCGCCCC GACATAGTTG GTAAAATACA GCAAATGCAC 32400

CTTGTTAACA TACACACTCC CCACATAGTA AATATACCGA GTAGACAGCT TAGAAAGCTC 32460

CCTCCGAAAA AATGGGAACA TGGTATCAAA GGCAGTGCCC GCAACACACA TCTTGAACAG 32520

ATCCATCAGG ATAGTAGCTC GACACAGCCC CTGCAGACTT TGGTCAGCTT GCTTGCTGCA 32580

GCAGTACACT CTCCACGTAG CATCTCCGCT GATGAAGTAT TCGCTATCGC AGCGACCAAA 32640

AATACAGCAA TCACAAGGCA GACGCAACAG TCTTTCATCC AGACTGTTCA TGAGAGGCTT 32700

TAGAGGTATG GGAAAAAATC CAAAGTGCTC AAAATAAGCA GCGCTGGGCT CATTCTGACA 32760

TTCCCCCAAC ATGCTGAGTC GAACCATAGC ACAGTCATAC AAACTCAGCT GTCGGAATTG 32820

ATCTTCCATG ATTGAGTTTC TACTGAGATA TTATCTCAAA CTTAAAACTG TTGCTCACCA 32880

ACTCTATGCG AACTTGCTCA AGAAGCTCTT GGTTTAGGGC GACCTCTTCT GGTCGTCGGA 32940

FIGURE 1S

```
AGTTACTGAT GGAACAACAA GCGCCGCCCA ACTTCAAATT TCCAGCCGAC CCAATCCAGT  33000
GGTCTCTCAA CTCACGCGCA CAAGCTACTA TGCAGTCCTC ACTTTCGTCA AAGTCAGCAG  33060
CGCCTATAGA AATCAACACA CTGAGTCCAC CATCTTCAGC TTTTAAGGGA TAACAGCTGA  33120
TAGCAAACTG GTTCTGAGAC CACGGCAAAG CACGTAGGAA TTGCTGTTAA GTTAATTTCC  33180
AAACACCGCT GAAGCAGCTC TATGGTTGCT GGACATATGT CCTCTGCATA GAAGCTTTGA  33240
ACATAACTTA AGACAGGGCC GGGCACATGA AACACAAACA GAGAACTATA CACAATCTGG  33300
GCCATGATCA CTCACATTTA AATAGCAGCT GAAAAGTGGC TTTCTTCACT TGGGAGCAAA  33360
ATTAGCGAAG ACTGTGCCAG AATGCTCACG TCGAAAGGCG GTGGGTCTCG CAGAGGCAGG  33420
TTCGGAGCTC TAATTAAACA CAGGTGGGTA ATCCAGTCAA CGATGAGGAC CAGCTGAAAA  33480
GTGGCTTTCT TCACTTGGGA GCAAAATTAG CGAAGACTGT GCCAGAATGC TCACGTCGAA  33540
AGGCGGTGGG TCTCGCAGAG GCAGGTTCGG AGCTCTAATT AAACACAGGT GGGTAATCCA  33600
GTCAACGATG AGGACTTTTA AAAAACTGTC TAAAACTGAA GCAGTTAAGT TAGAGGCAGA  33660
CACAGAAAAA ACTACAGTTA AACTATCAGT TGCTGAAATT GAAAAGCACC CAATAATTAT  33720
GCGCGAGGGC ACAGGCAATA AAAGTGTTAG CCCCTCGGCT AACGCGTCAG CTAAAAAATC  33780
TTTAGCTAAA GTATCTACTG GCCGCGTGGT AAAAGTTTGA ATATAATTTA CGACAGGAGC  33840
TGGCAAGTGA AACTCCACAA AAAAAGTAAA TGGCTGCACA CACGCCATTA TTTTGAAAAT  33900
AAGAAGTACT CACAAAATCA GCTGGAGCTG CCGCAAGTGA AAAAGACCAG CTGAAGTCTT  33960
ATTTTAAACT GTAAAATATA AAAAAAAAAA TAGGGCGTGA ACAAAAATGA GAAAATAATA  34020
CCGGATATGA CTATTAAGGG CGTACACTGA AACTGGGTAA TATTTGAGAA AAAGATTAAG  34080
ATAATAGCTG AACAAATGTT GTGTGCAGAA CACGGAAQAA TGGTGGCGAA AAAAAAAAAC  34140
AGTGTAAGCA CATGGCGCGC ACGTACTTCC GTGAGAAAAA TTAAAAAAAT TTACCCAGTA  34200
TAAGGTGCGT CATTAGACCC GCTTTGTGGC GCGGTTGTAG CCCTGCCCTT TGCCCCGCCC  34260
CGCGCGCCGC CCCGCGCGCC GCCCCCGCCG CCCTCAGCCC CGCCCAGCGC CGCCGCCTCC  34320
GCGACGCGCT CCGCCCCACA GTTACGTCAG CACGCCACGC TCGCCGTCGT TGCGTCATAA  34380
ATGACGTGGC AAAAATGATT GGCAGTTGGA CCGCTGCCAT CAGTGTACTG TAGATTATTG  34440
ATGATG                                                             34446
```

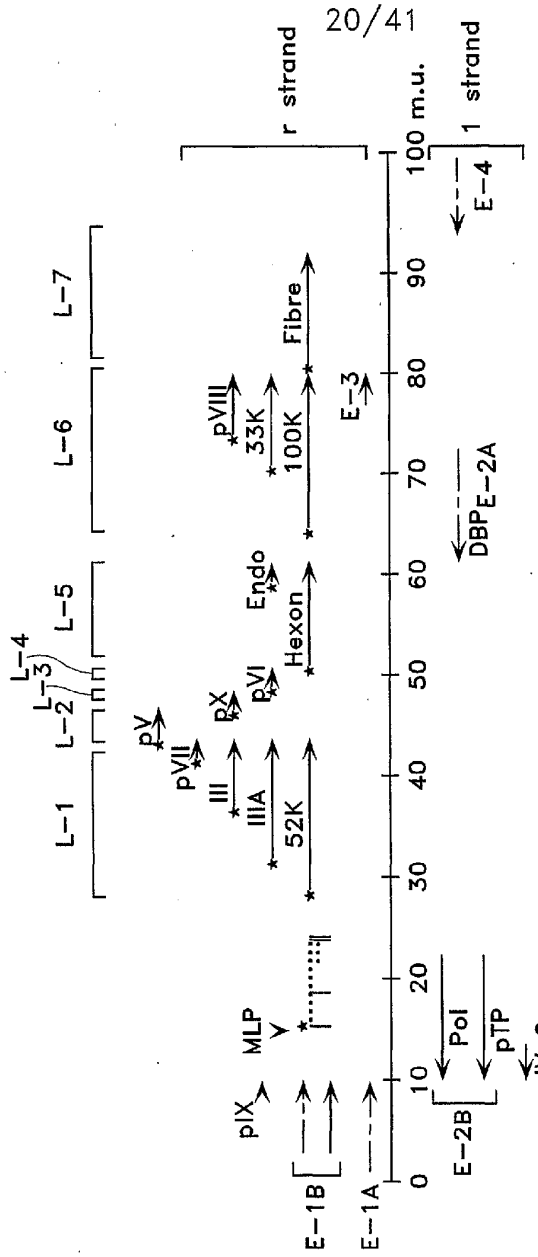

r strand
1 strand
100 m.u.
0
10
20
30
40
50
60
70
80
90
L-1
L-2
L-3
L-4
L-5
L-6
L-7
pIX
MLP
pV
pVII
III
IIIA
52K
pX
pVI
Hexon
Endo
pVIII
33K
100K
Fibre
E-1A
E-1B
E-2B
Pol
pTP
IVa2
E-3
DBP
E-2A
E-4

FIGURE 2

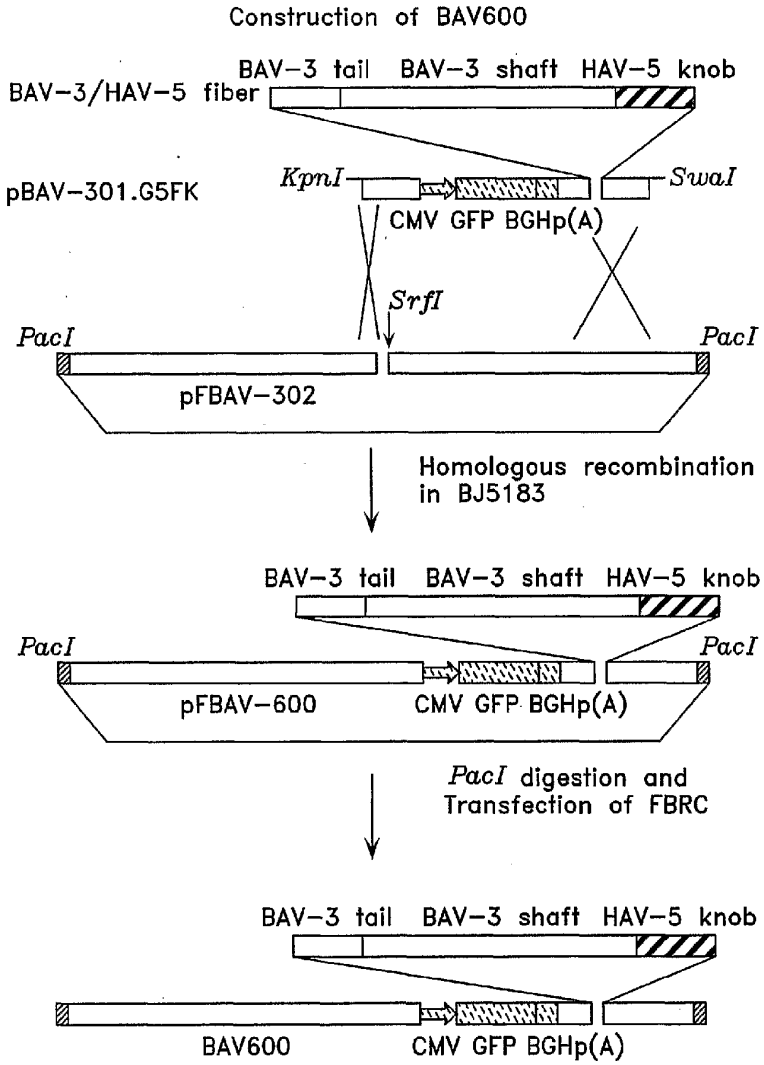

Construction of BAV600
BAV-3 tail BAV-3 shaft HAV-5 knob
BAV-3/HAV-5 fiber
pBAV-301.G5FK
KpnI
SwaI
CMV GFP BGHp(A)
SrfI
PacI
PacI
pFBAV-302
Homologous recombination in BJ5183
BAV-3 tail BAV-3 shaft HAV-5 knob
PacI
PacI
pFBAV-600
CMV GFP BGHp(A)
PacI digestion and Transfection of FBRC
BAV-3 tail BAV-3 shaft HAV-5 knob
BAV600
CMV GFP BGHp(A)

FIGURE 3

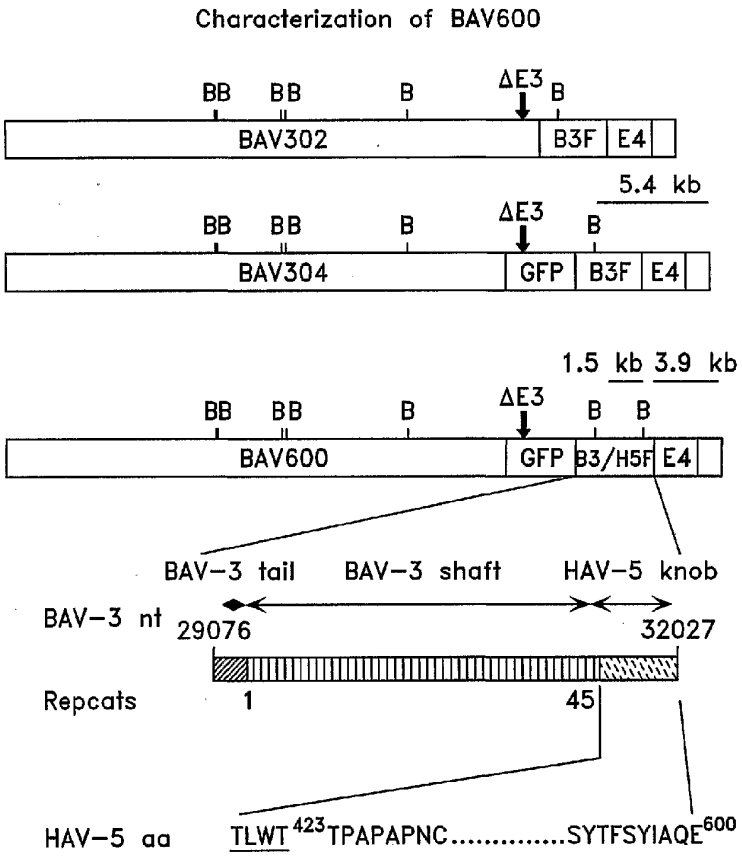
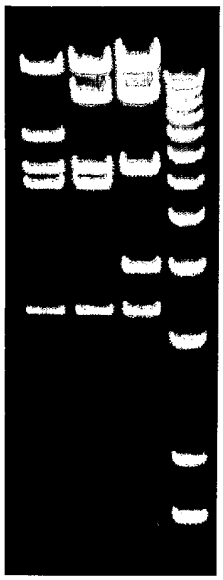

Characterization of BAV600
BB BB B ΔE3 B
BAV302
B3F E4
5.4 kb
BB BB B ΔE3 B
BAV304
GFP B3F E4
1.5 kb 3.9 kb
BB BB B ΔE3 B B
BAV600
GFP B3/H5F E4
BAV-3 tail
BAV-3 shaft
HAV-5 knob
BAV-3 nt
29076
32027
Repcats
1
45
HAV-5 aa
TLWT 423 TPAPAPNC.....SYTFSYIAQE 600

FIGURE 4

Analysis of BAV600 by Restriction Enzyme *Bgl* II Digestion

1 2 3



Lane 1. BAV302
2. BAV304
3. BAV600

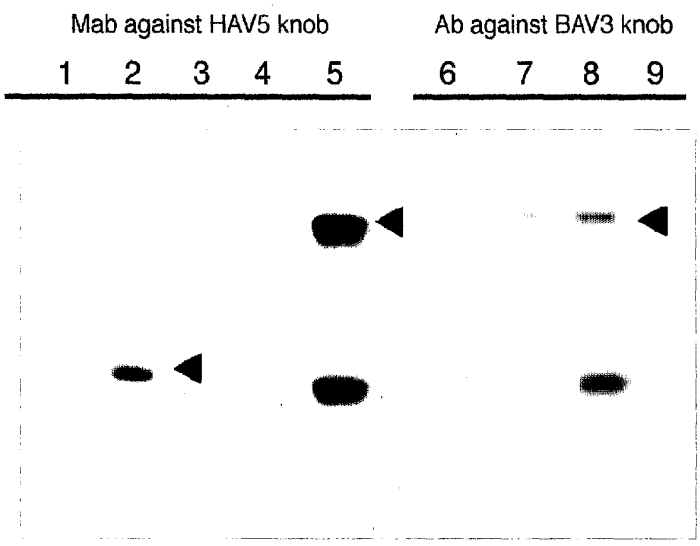
Figure 5A

1 2 3

Figure 5B

24/41

Expression of HAV-5 Fiber Knob by BAV600



Lane 1. Mock
2. HAV-5
3. BAV3
4. BAV304
5. BAV600
6. Mock
7. BAV3
8. BAV304
9. BAV600

Figure 6

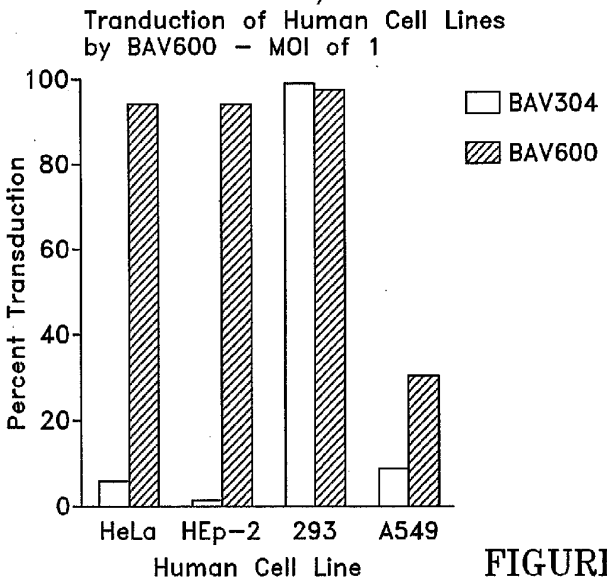


FIGURE 7A

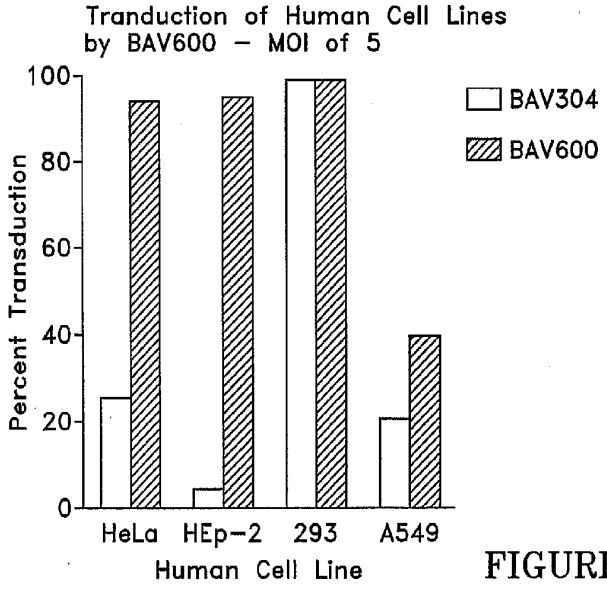


FIGURE 7B

FACS ANALYSIS OF BAV304 AND BAV600 TRANSDUCTION OF HUMAN CELLS

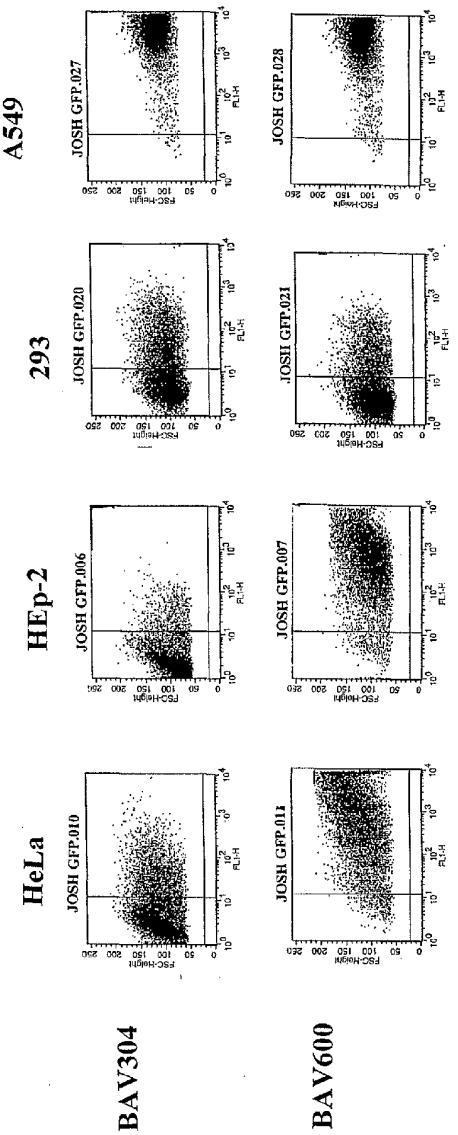


FIGURE 8

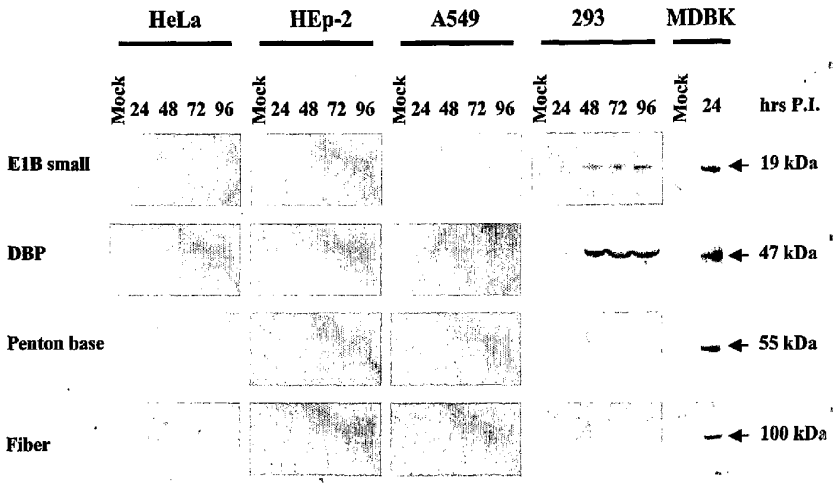

HeLa
HEp-2
A549
293
MDBK
Mock 24 48 72 96 Mock 24 48 72 96 Mock 24 48 72 96 Mock 24 48 72 96 Mock 24 hrs P.I.
E1B small
DBP
Penton base
Fiber
19 kDa
47 kDa
55 kDa
100 kDa

Figure 9

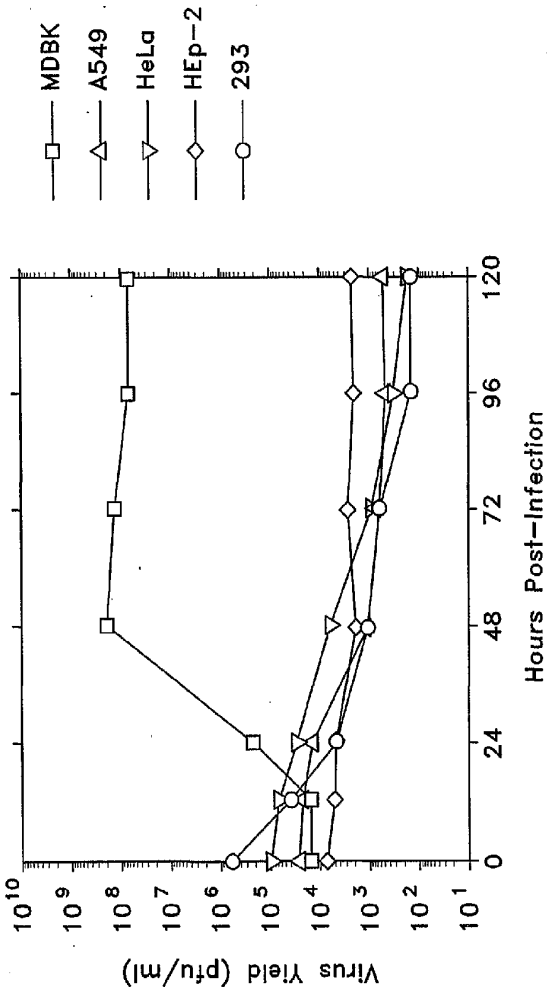

MDBK
A549
HeLa
HEp-2
293
Virus Yield (pfu/ml)
10^{10}
10^{9}
10^{8}
10^{7}
10^{6}
10^{5}
10^{4}
10^{3}
10^{2}
10^{1}
0
24
48
72
96
120
Hours Post-Infection

FIGURE 10

	Virus	
	BAV-3	BAV600
Normal Rabbit Serum	<1:50	<1:50
Rabbit Antiserum against BAV3 FK	1:800	<1:50
Monoclonal Ab against BHV gD (2C8)	<1:50	<1:50
Monoclonal Ab against HAd5 FK (1D6.14)	<1:50	1:3,200

FIGURE 11

FIGURE 12

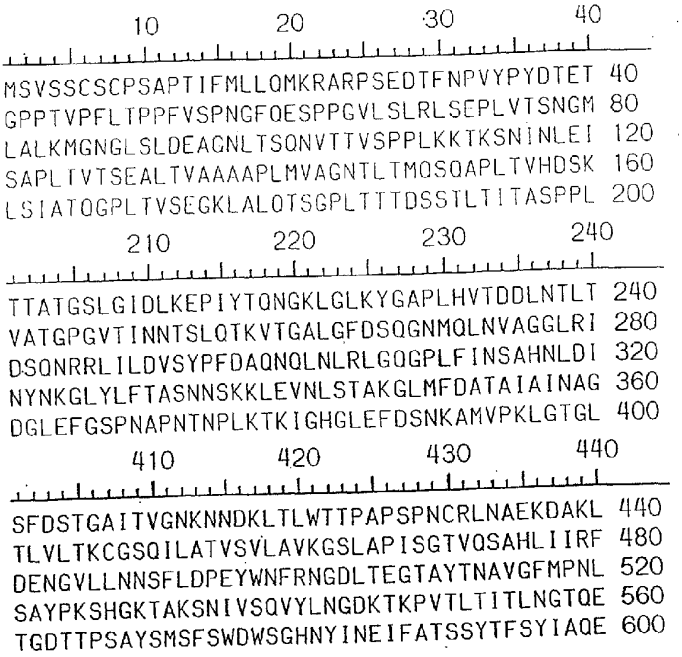

10 20 30 40
MSVSSCSCPSAPTIFMLLQMKRARPSEDTFNPVYPYDTET 40
GPPTVPFLTPPFVSPNGFQESPPGVLSLRLSEPLVTSNGM 80
LALKMGNGLSLDEAGNLTSQNVTTVSPPLKKTKSNINLEI 120
SAPLTVTSEALTVAAAAPLMVAGNTLTMQSQAPLTVHDSK 160
LSIATQGPLTVSEGKLALQTSGPLTTTDSSTLTITASPPL 200
210 220 230 240
TTATGSLGIDLKEPIYTQNGKLGLKYGAPLHVTDDLNTLT 240
VATGPGVTINNTSLQTKVTGALGFDSQGNMQLNVAGGLRI 280
DSQNRRLILDVSYPFDAQNQLNLRLGQGPLFINSAHNLDI 320
NYNKGLYLFTASNNSKKLEVNLSTAKGLMFDATAIAINAG 360
DGLEFGSPNAPNTNPLKTKIGHGLEFDSNKAMVPKLGTGL 400
410 420 430 440
SFDSTGAITVGNKNNDKLTLWTTPAPSPNCRLNAEKDAKL 440
TLVLTKCGSQILATVSVLAVKGSLAPISGTVQSAHLIIRF 480
DENGVLLNNSFLDPEYWNFRNGDLTEGTAYTNAVGFMPNL 520
SAYPKSHGKTAKSNIVSQVYLNGDKTKPVTLTITLNGTQE 560
TGDTTPSAYSMSFSWDWSGHNYINEIFATSSYTFSYIAQE 600

FIGURE 13

```
         10         20         30         40
MKRSVPQDFNLVYPYKAKRPNIMPPFFDRNGFVENQEATL 40
AMLVEKPLTFDKEGALTLQVGRGIRINPAGLLETNDLASA 80
VFPPLASDEAGNVTLNMSDGLYTKDNKLAVKVGPGLSLDS 120
NNALQVHTGDGLTVTDDKVSLNTQAPLSTTSAGLSLLLGP 160
SLHLGEEERLTVNTGAGLQISNNALAVKVGSGITYDAQNQ 200
         210        220        230        240
LAASLGDGLESRDNKTVVKAGPGLTITNQALTVATGNGLQ 240
VNPEGQLQLNITAGQGLNFANNSLAVELQSGLHFPPGQNQ 280
VSLYPGDGIDIRDNRVTVPAGPGLRMLNHQLAVASGDGLE 320
VHSDTLRLKLSHGLTFENGAVRAKLGPGLQTDDSGRSVVR 360
TGRGLRVANGQVQIFSGRGTAIGTDSSLTLNIRAPLQFSG 400
         410        420        430        440
PALTASLQGSGPITYNSNNGTFGLSIGPGMWVDQNRLQVN 440
PGAGLVFQGNNLVPNLADPLAISDSKISLSLGPGLTQASN 480
ALTLSLGNGLEFSNQAVAIKAGRGLRFESSSQALESSLTV 520
GNGLTLTDTVIRPNLGDGLEVRDNKIIVKLGANLRFENGA 560
VTAGTVNPSAPEAPPTLTAEPPLRASNSHLQLSLSEGLVV 600
         610        620        630        640
HNNALALQLGDGMEVNQHGLTLRVGSGLQMRDGILTVTPS 640
GTPIEPRLTAPLTQTENGIGLALGAGLELDESALQVKVGP 680
GMRLNPVEKYVTLLLGPGLSFGQPANRTNYDVRVSVEPPM 720
VFGQRGQLTFLVGHGLHIQNSKLQLNLGQGLRTDPVTNQL 760
EVPLGQGLEIADESQVRVKLGDGLQFDSQARITTAPNMVT 800
         810        820        830        840
ETLWTGTGSNANVTWRGYTAPGSKLFLSLTRFSTGLVLGN 840
MTIDSNASFGQYINAGHEQIECFILLDNQGNLKEGSNLQG 880
TWEVKNNPSASKAAFLPSTALYPILNESRGSLPGKNLVGM 920
QAILGGGGTCTVIATLNGRRSNNYPAGQSIIFVWQEFNTI 960
ARQPLNHSTLTFSYWT 976
```

FIGURE 14

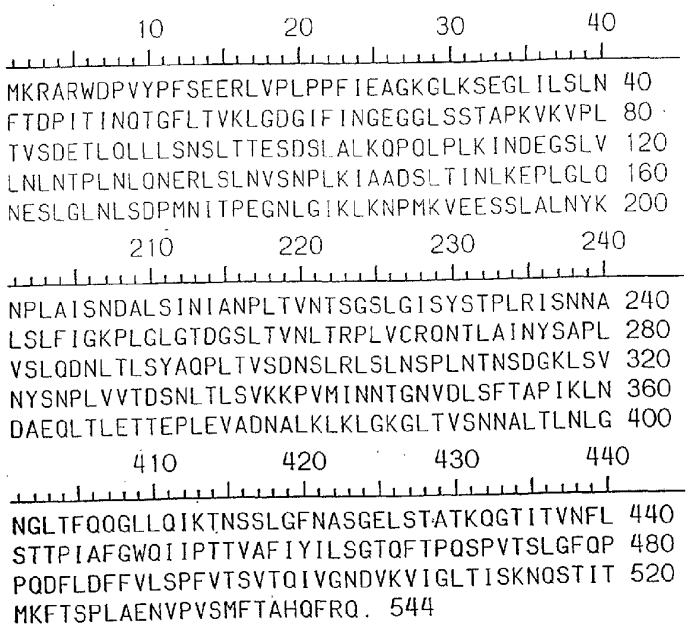

10 20 30 40
MKRARWDPVYPFSEERLVPLPPFIEAGKGLKSEGLILSLN 40
FTDPITINQTGFLTVKLGDGIFINGEGGLSSTAPKVKVPL 80
TVSDETLQLLLSNSLTTESDSLALKQPQLPLKINDEGSLV 120
LNLNTPLNLQNERLSLNVSNPLKIAADSLTINLKEPLGLQ 160
NESLGLNLSDPMNITPEGNLGIKLKNPMKVEESSLALNYK 200
210 220 230 240
NPLAISNDALSINIANPLTVNTSGSLGISYSTPLRISNNA 240
LSLFIGKPLGLGTDGSLTVNLTRPLVCRQNTLAINYSAPL 280
VSLQDNLTLSYAQPLTVSDNSLRLSLNSPLNTNSDGKLSV 320
NYSNPLVVTDSNLTLSVKKPVMINNTGNVDLSFTAPIKLN 360
DAEQLTLETTEPLEVADNALKLKLGKGLTVSNNALTLNLG 400
410 420 430 440
NGLTFQQGLLQIKTNSSLGFNASGELSTATKQGTITVNFL 440
STTPIAFGWQIIPTTVAFIYILSGTQFTPQSPVTSLGFQP 480
PQDFLDFFVLSPFVTSVTQIVGNDVKVIGLTISKNQSTIT 520
MKFTSPLAENVPVSMFTAHQFRQ. 544

FIGURE 15

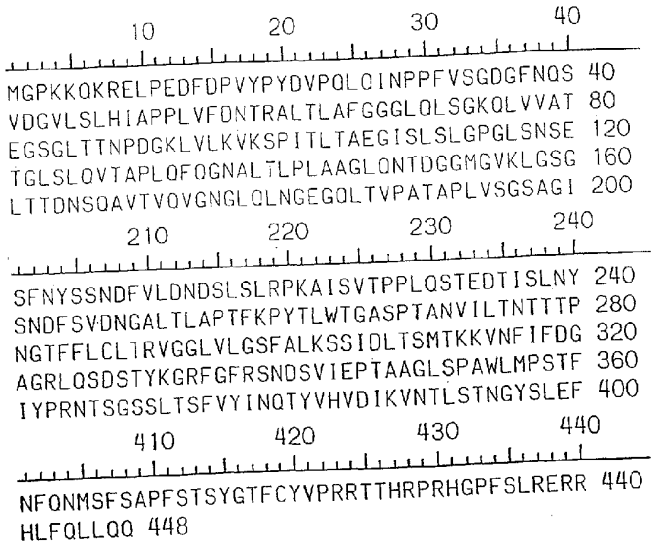
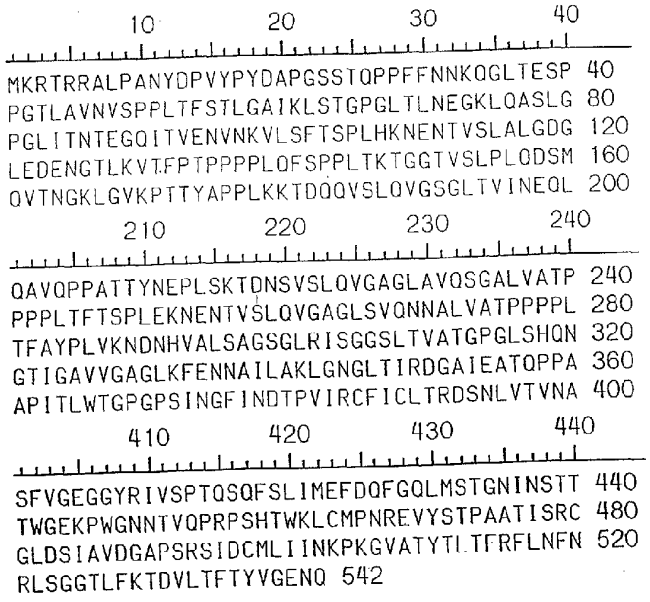
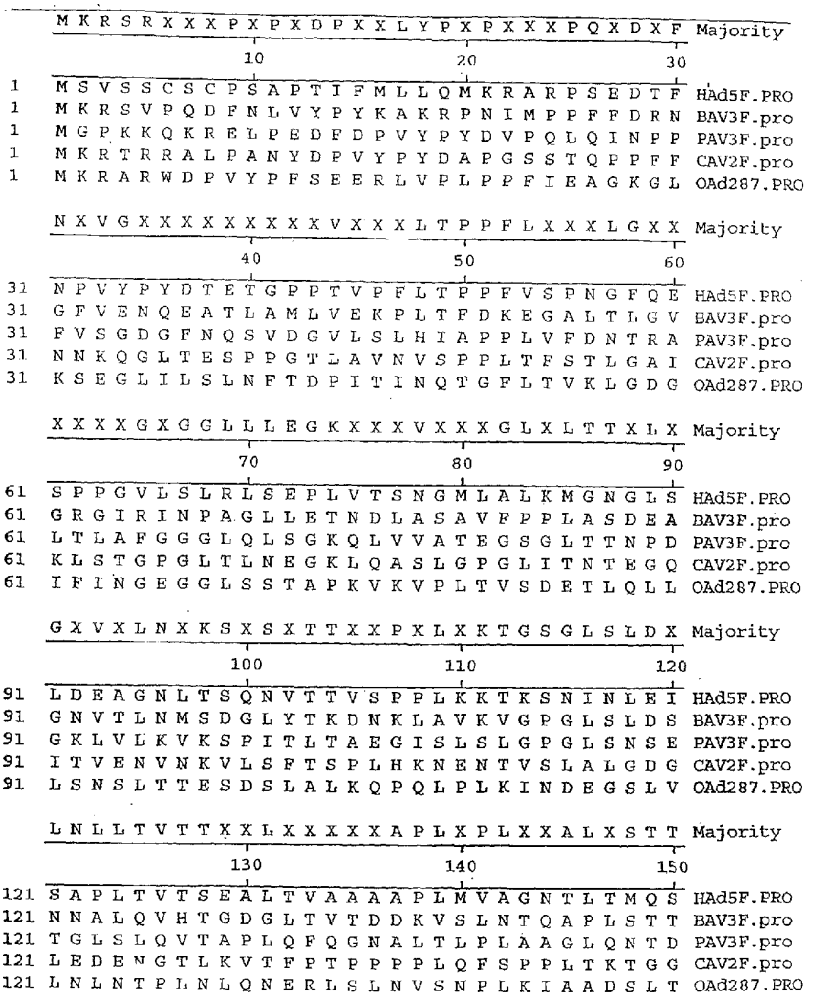

10 20 30 40
MGPKKQKRELPEDFDPVYPYDVPQLCINPPFVSGDGFNQS 40
VDGVLSLHIAPPLVFDNTRALTLAFGGGLQLSGKQLVVAT 80
EGSGLTTNPDGKLVLKVKSPITLTAEGISLSLGPGLSNSE 120
TGLSLQVTAPLQFQGNALTLPLAAGLQNTDGGMGVKLGSG 160
LTTDNSQAVTVQVGNGLQLNGEGQLTVPATAPLVSGSAGI 200
210 220 230 240
SFNYSSNDFVLDNDSLSLRPKAISVTPPLQSTEDTISLNY 240
SNDFSVDNGALTLAPTFKPYTLWTGASPTANVILTNTTTP 280
NGTFFLCLTRVGGLVLGSFALKSSIDLTSMTKKVNFIFDG 320
AGRLQSDSTYKGRFGFRSNDSVIEPTAAGLSPAWLMPSTF 360
IYPRNTSGSSLTSFVYINQTYVHVDIKVNTLSTNGYSLEF 400
410 420 430 440
NFQNMSFSAPFSTSYGTFCYVPRRTTHRPRHGPFSLRERR 440
HLFQLLQQ 448

FIGURE 16



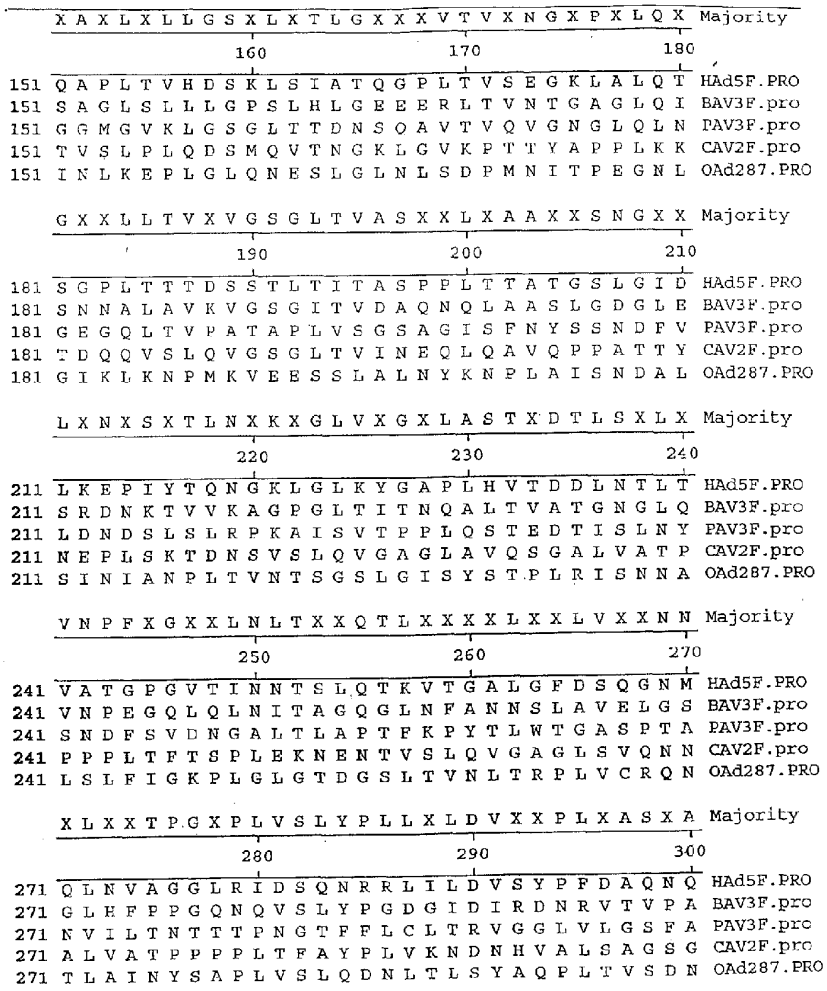
SUBSTITUTE SHEET (RULE 26)

FIGURE 17A



SUBSTITUTE SHEET (RULE 26)

FIGURE 17B



SUBSTITUTE SHEET (RULE 26)

FIGURE 17C

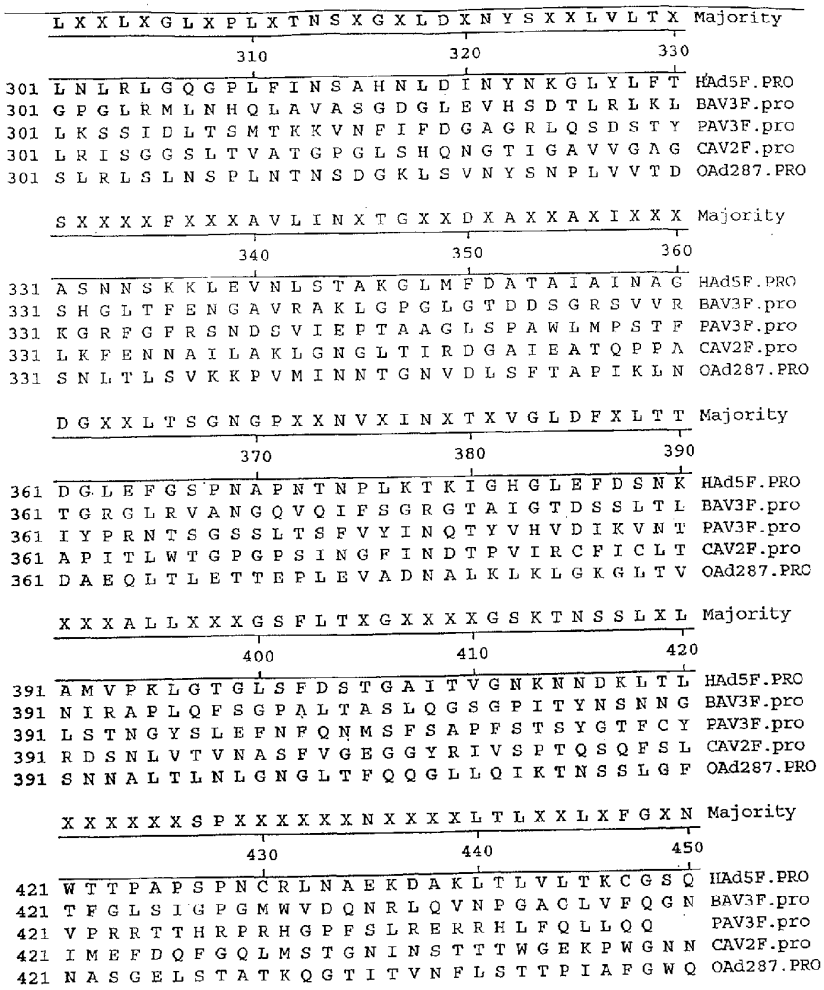
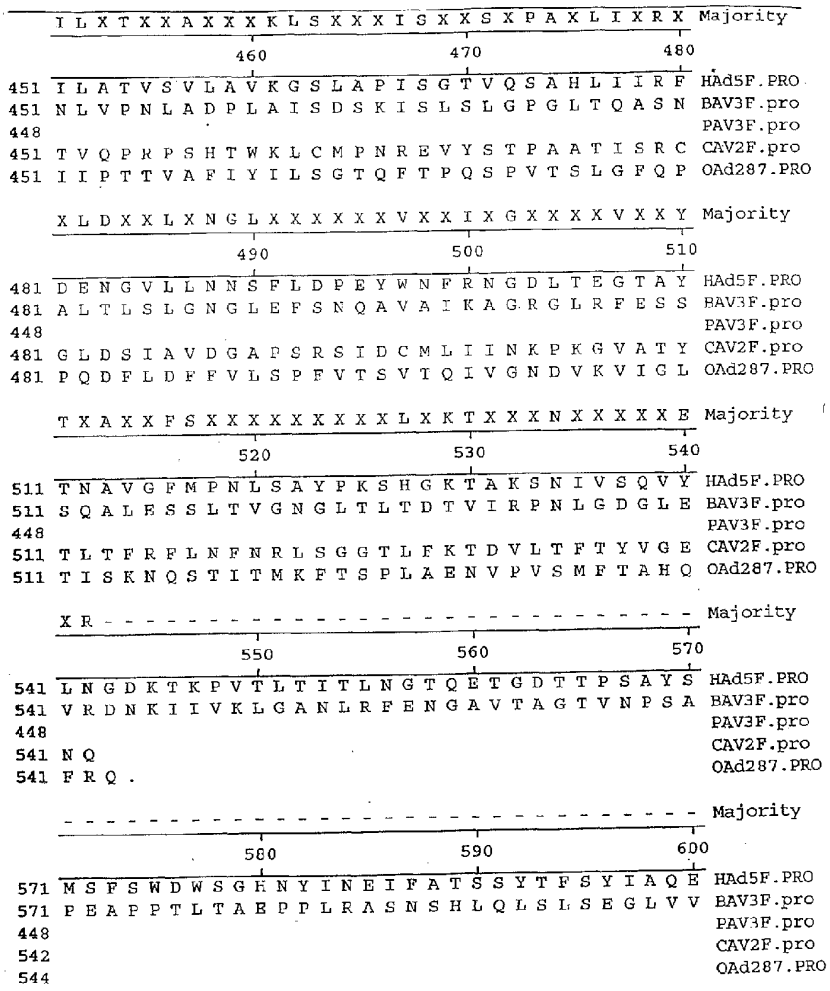

L X X L X G L X P L X T N S X G X L D X N Y S X X L V L T X Majority
310 320 330
301 L N L R L G Q G P L F I N S A H N L D I N Y N K G L Y L F T HAd5F.PRO
301 G P G L R M L N H Q L A V A S G D G L E V H S D T L R L K L BAV3F.pro
301 L K S S I D L T S M T K K V N F I F D G A G R L Q S D S T Y PAV3F.pro
301 L R I S G G S L T V A T G P G L S H Q N G T I G A V V G A G CAV2F.pro
301 S L R L S L N S P L N T N S D G K L S V N Y S N P L V V T D OAd287.PRO
S X X X X F X X X A V L I N X T G X X D X A X X A X I X X X Majority
340 350 360
331 A S N N S K K L E V N L S T A K G L M F D A T A I A I N A G HAd5F.PRO
331 S H G L T F E N G A V R A K L G P G L G T D D S G R S V V R BAV3F.pro
331 K G R F G F R S N D S V I E P T A A G L S P A W L M P S T F PAV3F.pro
331 L K F E N N A I L A K L G N G L T I R D G A I E A T Q P P A CAV2F.pro
331 S N L T L S V K K P V M I N N T G N V D L S F T A P I K L N OAd287.PRO
D G X X L T S G N G P X X N V X I N X T X V G L D F X L T T Majority
370 380 390
361 D G L E F G S P N A P N T N P L K T K I G H G L E F D S N K HAd5F.PRO
361 T G R G L R V A N G Q V Q I F S G R G T A I G T D S S L T L BAV3F.pro
361 I Y P R N T S G S S L T S F V Y I N Q T Y V H V D I K V N T PAV3F.pro
361 A P I T L W T G P G P S I N G F I N D T P V I R C F I C L T CAV2F.pro
361 D A E Q L T L E T T E P L E V A D N A L K L K L G K G L T V OAd287.PRO
X X X A L L X X X G S F L T X G X X X X G S K T N S S L X L Majority
400 410 420
391 A M V P K L G T G L S F D S T G A I T V G N K N N D K L T L HAd5F.PRO
391 N I R A P L Q F S G P A L T A S L Q G S G P I T Y N S N N G BAV3F.pro
391 L S T N G Y S L E F N F Q N M S F S A P F S T S Y G T F C Y PAV3F.pro
391 R D S N L V T V N A S F V G E G G Y R I V S P T Q S Q F S L CAV2F.pro
391 S N N A L T L N L G N G L T F Q Q G L L Q I K T N S S L G F OAd287.PRO
X X X X X X S P X X X X X X N X X X X L T L X X L X F G X N Majority
430 440 450
421 W T T P A P S P N C R L N A E K D A K L T L V L T K C G S Q HAd5F.PRO
421 T F G L S I G P G M W V D Q N R L Q V N P G A C L V F Q G N BAV3F.pro
421 V P R R T T H R P R H G P F S L R E R R H L F Q L L Q Q PAV3F.pro
421 I M E F D Q F G Q L M S T G N I N S T T T W G E K P W G N N CAV2F.pro
421 N A S G E L S T A T K Q G T I T V N F L S T T P I A F G W Q OAd287.PRO

FIGURE 17 D



SUBSTITUTE SHEET (RULE 26)

FIGURE 17E

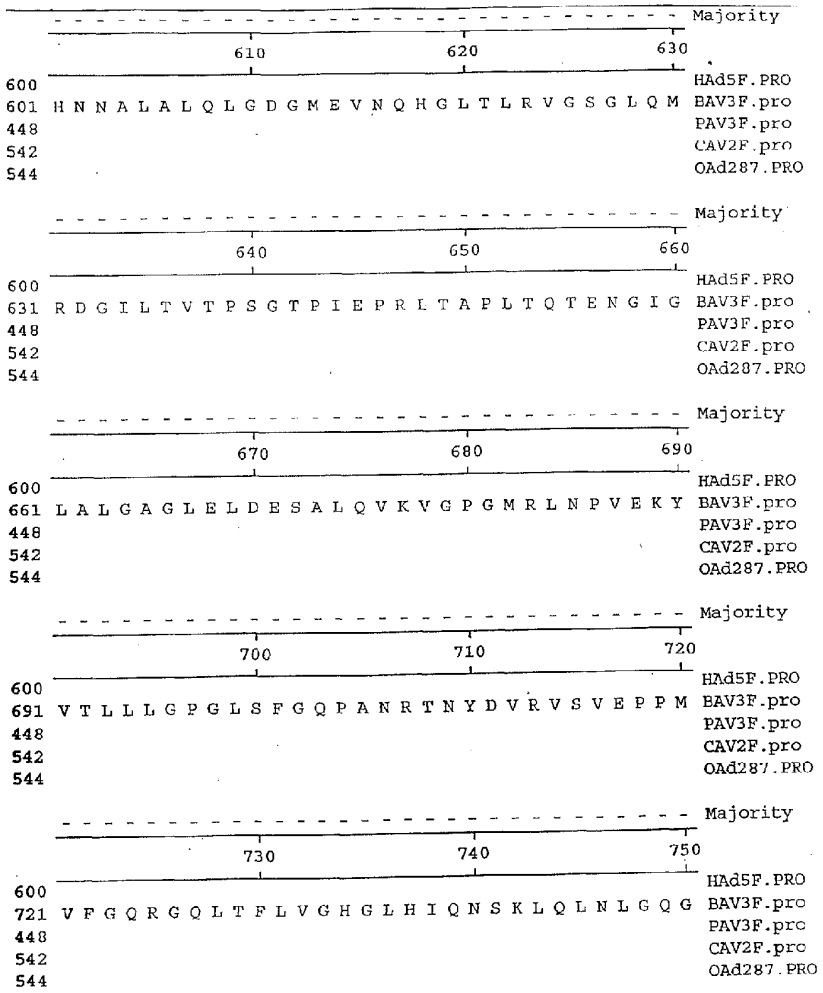

Majority
610
620
630
600
601 H N N A L A L Q L G D G M E V N Q H G L T L R V G S G L Q M
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
640
650
660
600
631 R D G I L T V T P S G T P I E P R L T A P L T Q T E N G I G
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
670
680
690
600
661 L A L G A G L E L D E S A L Q V K V G P G M R L N P V E K Y
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
700
710
720
600
691 V T L L L G P G L S F G Q P A N R T N Y D V R V S V E P P M
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
730
740
750
600
721 V F G Q R G Q L T F L V G H G L H I Q N S K L Q L N L G Q G
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO

FIGURE 17F

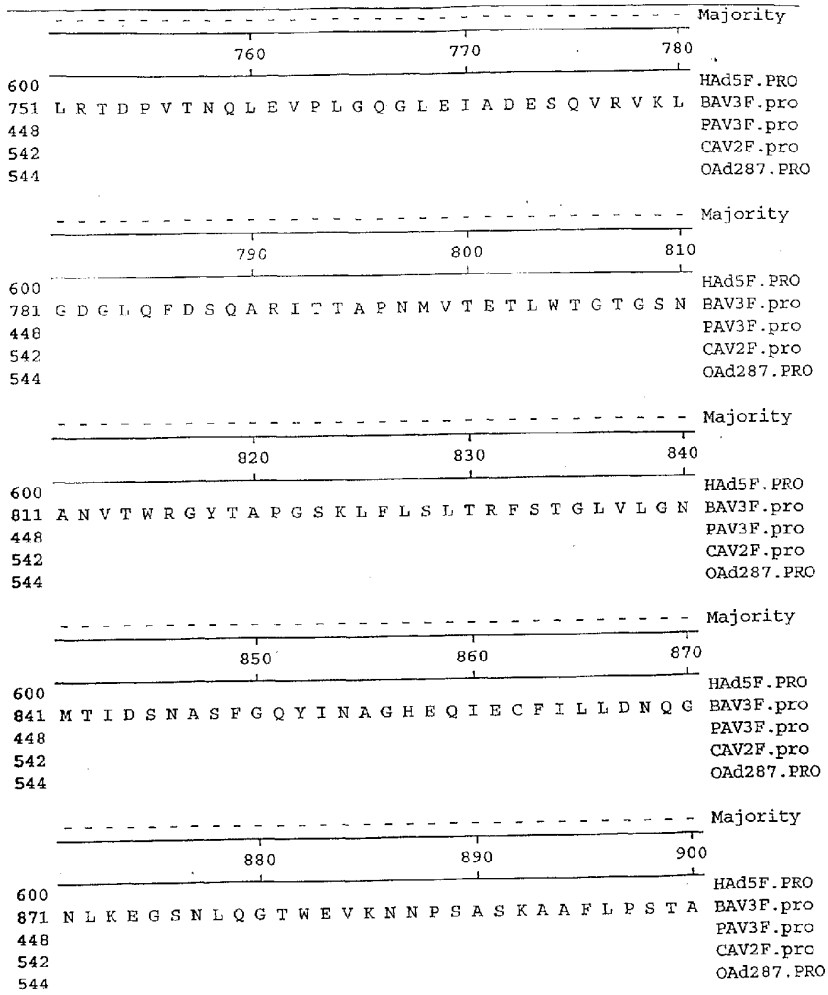
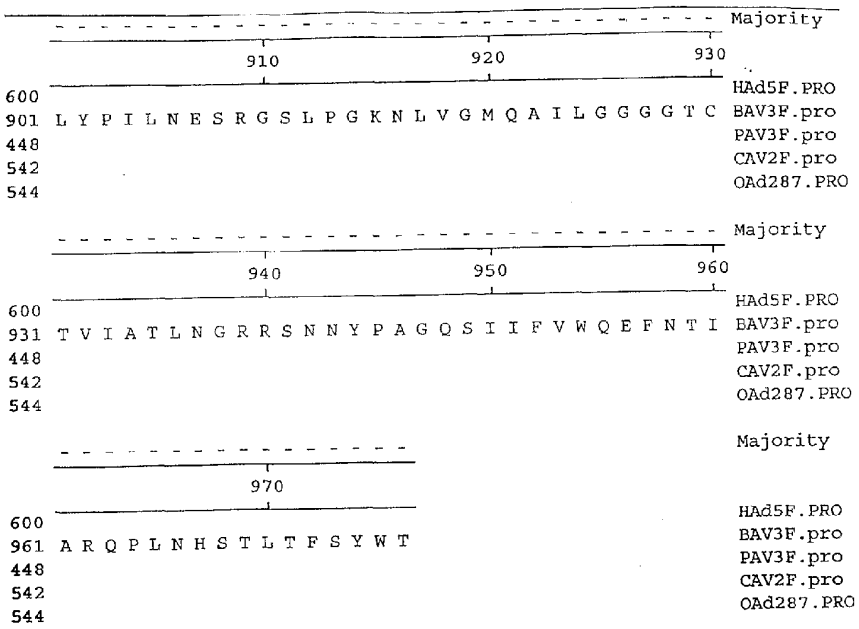

Majority
760
770
780
600
751 L R T D P V T N Q L E V P L G Q G L E I A D E S Q V R V K L
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
790
800
810
600
781 G D G L Q F D S Q A R I T T A P N M V T E T L W T G T G S N
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
820
830
840
600
811 A N V T W R G Y T A P G S K L F L S L T R F S T G L V L G N
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
850
860
870
600
841 M T I D S N A S F G Q Y I N A G H E Q I E C F I L L D N Q G
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
880
890
900
600
871 N L K E G S N L Q G T W E V K N N P S A S K A A F L P S T A
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO

FIGURE 17G


Majority
910
920
930
600
901 L Y P I L N E S R G S L P G K N L V G M Q A I L G G G G T C
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
940
950
960
600
931 T V I A T L N G R R S N N Y P A G Q S I I F V W Q E F N T I
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO
Majority
970
600
961 A R Q P L N H S T L T F S Y W T
448
542
544
HAd5F.PRO
BAV3F.pro
PAV3F.pro
CAV2F.pro
OAd287.PRO