PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/00, 15/12, 15/63
C12N 15/64, 15/85, C12M 3/00
C12M 3/02, 3/04, A61K 31/70

A1

(11) International Publication Number: WO 93/22431

(43) International Publication Date: 11 November 1993 (11.11.93)

(21) International Application Number: PCT/US93/03993

(22) International Filing Date: 28 April 1993 (28.04.93)

(30) Priority data:
07/876,286 30 April 1992 (30.04.92) US

(71) Applicant: BAYLOR COLLEGE OF MEDICINE [US/US]; One Baylor Plaza, Houston, TX 77030 (US).

(72) Inventors: ROOP, Dennis, R. ; 2331 Bellefontaine, Houston, TX 77030 (US). ROTHNAGEL, Joseph, A. ; 5427 Queensloch Drive, Houston, TX 77096 (US). GREENHALGH, David, A. ; 7423 Brompton, Houston, TX 77025 (US).

(74) Agent: PAUL, Thomas, D.; Fulbright & Jaworski, 1301 McKinney, Suite 5100, Houston, TX 77010-3095 (US).

(81) Designated States: AT, AU, BB, BG, BR, CA, CH, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, PL, RO, RU, SD, SE, UA, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: CONSTITUTIVE AND INDUCIBLE EPIDERMAL VECTOR SYSTEMS

(57) Abstract

A loricrin constitutive vector for efficient expression of a nucleic acid sequence in epidermal cells comprising the 5′ flanking region of the loricrin gene, said flanking region containing a TATA box, a cap site and a first intron and intron/exon boundary all in appropriate sequential and positional relationship for expression of a nucleic acid cassette, a 3′ flanking sequence of the loricrin gene and a linker containing a unique restriction endonuclease site at the location of the start and stop codon. Said linker connecting the 5′ flanking region to the 3′ flanking sequence and said linker further providing a position for inserting the cassette. The cassette contains the specific nucleic acid sequence to be expressed. Also, there is a keratin K6 inducible vector for regulating expression of a nucleic acid sequence in epidermal cells comprising the 5′ flanking region of the keratin K6 gene, said flanking region including the TATA box, a cap site and the first intron and intron/exon boundary all in sequential and positional relationship for expression of a nucleic acid cassette, a 3′ flanking sequence of the keratin K6 gene, and a polylinker having a plurality of restriction endonuclease sites. The polylinker connects the 5′ flanking region to the 3′ flanking sequence and further provides a position for insertion of the cassette. The keratin K6 and loricrin vectors can be further regulated by the addition of a Vitamin D regulatory element. The vectors can be used in a bioreactor for generating a variety of products including proteins, polypeptides or antisense RNAs. The vectors can also be used for gene therapy in treatment of a variety of diseases in animals and humans including wound healing, surgical incisions, skin ulcers, psoriasis and skin cancer, and in vaccination.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**MR**	Mauritania
AU	Australia	**GA**	Gabon	**MW**	Malawi
BB	Barbados	**GB**	United Kingdom	**NL**	Netherlands
BE	Belgium	**GN**	Guinea	**NO**	Norway
BF	Burkina Faso	**GR**	Greece	**NZ**	New Zealand
BG	Bulgaria	**HU**	Hungary	**PL**	Poland
BJ	Benin	**IE**	Ireland	**PT**	Portugal
BR	Brazil	**IT**	Italy	**RO**	Romania
CA	Canada	**JP**	Japan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SK**	Slovak Republic
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SN**	Senegal
CM	Cameroon	**LI**	Liechtenstein	**SU**	Soviet Union
CS	Czechoslovakia	**LK**	Sri Lanka	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**MC**	Monaco	**UA**	Ukraine
DK	Denmark	**MG**	Madagascar	**US**	United States of America
ES	Spain	**ML**	Mali	**VN**	Viet Nam
FI	Finland	**MN**	Mongolia		

CONSTITUTIVE AND INDUCIBLE EPIDERMAL VECTOR SYSTEMS

The invention was partially supported by a grant from the United States government under AR40240 awarded by the National Institutes of Health. The government has certain rights in this invention.

FIELD OF THE INVENTION

The present invention relates generally to expression vectors for use in expressing proteins and polypeptides in epidermal cells. More particularly it relates to a constitutive vector consisting of the loricrin gene promoter, its 5′ flanking region, its 5′ transcribed but untranslated region, its intron, its 3′ transcribed but untranslated region, its contiguous non-coding DNA containing the gene's natural transcriptional termination region and its 3′ flanking region. It further relates to an inducible vector consisting of the K6 keratin gene promoter, its 5′ flanking region, its 5′ transcribed but untranslated region, its first intron, its 3′ transcribed but untranslated region, its contiguous non-coding DNA containing the gene's natural transcriptional termination region and its 3′ flanking region. Additionally it relates to the treatment of disease using the constitutive and inducible vectors.

BACKGROUND OF THE INVENTION

The skin is the largest organ in the human body and, due to its accessibility, it is an attractive target for gene therapy. The outer layer of the skin is called the epidermis, and it is particularly attractive since epidermal cells can be grown *in vitro* from normal and affected patients,

are easily transformed genetically by vectors, and can be readily reintroduced by autografting. Previous studies investigating the feasibility of using epidermal cells for gene therapy have only considered this *ex vivo* approach. These investigations utilized retroviral vectors and their promoters to introduce and express foreign genetic material in epidermal cells. Even though the epidermis is avascular, these studies demonstrated that proteins expressed in the epidermis were able to traverse the epidermal-dermal barrier and achieve systemic distribution (Morgan et al., Science, Vol. 237, pp. 1476-1479, (1987); Fenjves et al., PNAS USA, Vol. 86, pp. 8803-8807, (1989); Garlick et al., J. Invest. Dermatol., Vol. 97, pp. 824-829, (1991)). The accessibility of the epidermis makes it suitable for other routes of vector delivery that do not require an *ex vivo* approach, *e.g.*, a gene gun (Sanford et al., Techniques, Vol. 3, pp. 3-16, (1991); Williams et al., PNAS USA, Vol. 88, pp. 2726-2730, (1991); Johnstone et al., In Vitro Cell Dev. Biol., Vol. 27, pp. 11-14, (1991)). In addition, novel vector systems derived from genes normally expressed at high levels in epidermal cells, could prove optimal for achieving efficient, as well as regulated, expression of exogenous DNA. These vector systems are the subject of this invention.

The epidermis is a continuously regenerating stratified squamous epithelium. Differentiated epidermal cells are the progeny of proliferative cells located in the basal cell layer and there is substantial evidence suggesting that the regeneration process occurs in proliferative units composed of slowly cycling, self-renewing stem cells, proliferative but non-renewing transit amplifying cells, and post-mitotic maturing epidermal cells (Iversen, et al., Cell Tissue Kinet., Vol. 1, pp. 351-367, (1968); MacKenzie, et al., Nature, Vol. 226, pp. 653-655, (1970); Christophers, et al., J. Invest. Dermatol., Vol. 56, pp. 165-170, (1971); Potten, In Stem Cells: Their Identification and Characterization, pp. 200-232, (1983); Cotsarelis, et al., Cell, Vol. 61, pp. 1329-1337, (1990)). The maturation

process (terminal differentiation) is initiated when epidermal cells withdraw from the cell cycle and migrate from the basal layer into the spinous layer. Maturation continues as spinous cells migrate into the granular layer and terminates with the formation of the stratum corneum. Morphological and biochemical studies have shown that terminal differentiation occurs in stages. (Matoltsy, J. Invest. Dermatol., Vol. 65, pp. 127-142, (1975)). Keratins K5 and K14 are major products of basal epidermal cells (Woodcock-Mitchell, et al., J. Cell Biol., Vol. 95, pp. 580-588, (1982)). These proteins assemble into 10 nm filaments (intermediate filaments [IF]) and, together with microtubules (tubulin) and microfilaments (actin), comprise the cytoskeleton of epidermal cells (Steinert, P.M., et al., Cell, Vol. 42, pp. 411-419, (1985)). One of the earliest changes associated with the commitment to differentiation and migration into the spinous layer is the induction of another differentiation-specific pair of keratins (K1 and K10). IF containing K1 and K10 replace those containing K5 and K14 as the major products of cells in the spinous layer (Woodcock-Mitchell, et al., J. Cell Biol., Vol. 95, pp. 580-588, (1982); Roop, et al., Proc. Natl. Acad. Sci., USA, Vol. 80, pp. 716-720, (1983); Schweizer, et al, Cell, Vol. 37, pp. 159-170, (1984)). The keratin IF formed by these proteins assemble into bundles. In the granular layer, another high molecular weight non-IF protein is synthesized, which is processed into filaggrin, and is thought to promote keratin filament aggregation and disulfide-bond formation (Dale, B.A., et al., Nature, Vol. 276, pp. 729-731, (1978); Harding, C.R., et al., J. Mol. Biol., Vol. 170, pp. 651-673, (1983)). In the final stage of epidermal cell maturation, transglutaminase catalyzes the crosslinking of involucrin and loricrin, by the formation of (γ-glutamyl) lysine isopeptides, into a highly insoluble cornified envelope which is located just beneath the plasma membrane (Rice and Green, Cell, Vol. 11, pp. 417-422, (1977); Mehrel, et al., Cell, Vol. 61, pp. 1103-1112, (1990)).

Genes or cDNAs encoding the major keratins expressed in epidermal cells have now been cloned: K5 (Lersch, et al., Mol. and Cell Biol., Vol. 8, pp. 486-493, (1988), K14 (Marchuk, et al., Proc. Natl. Acad. Sci, USA, Vol. 82, pp. 1609-1613, (1985); Knapp, et al., J. Biol. Chem, Vol. 262, pp. 938-945, (1987); Roop, et al., Cancer Res., Vol. 48, pp. 3245-3252, (1988), K1 (Steinert, et al., J. Biol. Chem., Vol. 260, pp. 7142-7149, (1985) and K10 (Krieg, et al., J. Biol. Chem., Vol. 260, pp. 5867-5870, (1985)). Northern blot analysis and *in situ* hybridization studies suggest that keratin genes K5 and K14 are predominantly transcribed in the proliferating basal layer and transcription of keratin genes K1 and K10 is induced as cells migrate into the spinous layer (Lersch, et al., Mol. and Cell Biol., Vol. 8, pp. 486-493, (1988); Knapp, et al., J. Biol. Chem, Vol. 262, pp. 938-945, (1987); Roop, et al., Cancer Res., Vol. 48, pp. 3245-3252, (1988)). Genes encoding rat (Haydock, et al., J. Biol. Chem., Vol. 261, pp. 12520-12525, (1986)) and mouse (Rothnagel, et al., J. Biol. Chem., Vol. 262, pp. 15643-15648, (1987)) filaggrin have now been identified and *in situ* hybridization experiments have confirmed that transcription of this gene is restricted to the granular layer (Rothnagel, et al, J. Biol. Chem., Vol. 262, pp. 15643-15648, (1987); Fisher, et al. J. Invest. Dermatol., Vol. 88, pp. 661-664, (1987)). To date, loricrin is the only gene encoding a component of the cornified envelope to be studied at the molecular level by *in situ* hybridization and transcripts of this gene are restricted to the granular layer (Mehrel, et al., Cell, Vol. 61, pp. 1103-1112, (1990)).

Since the genes encoding the structured proteins described above are expressed at very high levels, i.e. their individual transcripts represent 5-10% of the total messenger RNA in epidermal cells, their regulatory regions could be utilized in the construction of vectors to direct efficient expression of exogenous DNA in epidermal cells. In particular, efforts have focused on the gene encoding loricrin, a major keratinocyte cell envelope protein (Mehrel et al., Cell, Vol. 61, pp. 1103-1112, (1990)).

Although this gene is normally only expressed in the most differentiated layers of the epidermis, the present invention demonstrates that it possible to remove sequences that normally restrict expression of the loricrin gene in undifferentiated cells and achieve high levels of expression in undifferentiated epidermal cells (greater than the viral promoter of SV40). Thus, this vector is constitutively expressed in epidermal cells at all differentiation states.

In addition to the constitutive vector, the present invention takes advantage of the expression characteristics of another gene encoding the K6 keratin to construct an inducible vector. The K6 gene is normally never expressed in the epidermis, but it can be induced under hyperproliferative conditions such as wound healing (Weiss, et al., J. Cell Biol., Vol. 98, pp. 1397-1406, (1984); Nakazawa, et al., J. Cell Biol., Vol. 103, pp. 561a, (1986); Stoler, et al., J. Cell Biol., Vol. 107, pp. 427-446, (1988) and topical application of retinoic acid (Rosenthal et al., J. Invest. Dermatol., Vol. 95, pp. 510-515, (1990)).

SUMMARY OF THE INVENTION

An object of the present invention is a loricrin constitutive vector for efficient expression of nucleic acid sequences in epidermal cells.

An additional object of the present invention is a keratin K6 inducible vector for regulated expression of nucleic acid sequences in epidermal cells.

Another object of the present invention is an *in vivo* method of transducing epidermal cells with a constitutive or inducible vector.

A further object of the present invention is a bioreactor for producing proteins and polypeptides.

An additional object of the present invention is an enhanced method of wound healing or healing of surgical incisions.

Another object of the present invention is a method of treating skin ulcers.

An additional object of the present invention is a method of treating psoriasis.

A further object of the present invention is a method of treating cancer.

Thus, in accomplishing the foregoing objects, there is provided in accordance with one aspect of the present invention, a loricrin constitutive vector for efficient expression of nucleic acid sequences in epidermal cells, comprising a 5′ flanking region of the loricrin gene, said flanking region including a TATA box, a cap site and a first intron and an intron/exon boundary, all in appropriate sequential and positional relationship for expression of a nucleic acid cassette; a 3′ flanking sequence of the loricrin gene; and a linker having a unique restriction endonuclease site at the location of the start and stop codon, said linker connecting the 5′ flanking region to the 3′ flanking sequence and said linker further providing a position for inserting the nucleic acid cassette which includes the specific nucleic acid sequence to be expressed.

In specific embodiments of the present invention, the loricrin constitutive vector has a 5′ flanking region of approximately 1.5 kb, an intron of approximately 1.1 kb and a 3′ flanking sequence of approximately 2.1 kb. In specific embodiments of the present invention, the loricrin constitutive vector also includes a poly-linker.

An alternative embodiment of the present invention is a keratin K6 inducible vector for regulated expression of a nucleic acid sequence in epidermal cells, comprising a 5′ flanking region of the keratin K6 gene, said flanking region including a TATA box, a cap site, a first intron and an intron/exon boundary, all in sequential and positional relationship for expression of a nucleic acid cassette; a 3′ flanking sequence of the keratin K6 gene; and a poly-linker having a plurality of restriction endonuclease

sites, said poly-linker connecting the 5′ flanking region to the 3′ flanking sequence and further providing a position for insertion of the nucleic acid cassette which includes the specific nucleic acid sequence to be expressed.

In specific embodiments of the present invention, the keratin K6 inducible vector, 5′ flanking region of approximately 8.0 kb, an intron and intron/exon boundary of approximately 0.56 kb and the 3′ flanking sequence of approximately 1.2 kb.

In the present invention, the restriction endonuclease sites in the linker or poly-linker are selected from the group consisting of Cla I, Not I, Xma I, Bgl II, Pac I, Xho I, Nhe I and Sfi I.

In one embodiment of the present invention, the nucleic acid cassette, of the constitutive or inducible vectors, contains a sequence coding for a protein, polypeptide or antisense RNA.

In specific embodiments of the present invention, there is a bioreactor comprising transduced epidermal cells including either the loricrin constitutive or keratin K6 inducible vectors. The bioreactor can produce a variety of compounds selected from proteins, polypeptides, antisense RNA.

In specific embodiments of the present invention, the loricrin constitutive or keratin K6 inducible vectors are used for the treatment of wounds, surgical incisions, psoriasis, skin ulcers and cancer.

The method of the present invention can also be used for vaccination by transducing epidermal cells with a loricrin constitutive or keratin K6 inducible vector having proteins or polypeptides which induce an immunological response.

Another embodiment of the present invention is the nucleotide sequences for the loricrin gene and loricrin constitutive vector.

Another embodiment of the present invention is the nucleotide sequences for the keratin K6 gene and keratin K6 inducible vector.

Other and further objects, features and advantages will be apparent from the following description of the presently preferred embodiments of invention which are given for the purposes of disclosure when taken in conjunction with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic drawing of the mouse loricrin gene and the constitutive epidermal vector derived from its regulatory sequences.

Figure 2 shows the expression characteristics of the constitutive epidermal vector in undifferentiated and differentiated epidermal cells utilizing a reporter gene encoding chloramphenicol acetyl transferase (CAT).

Figure 3 shows the expression characteristics of the constitutive epidermal vector *in vivo* utilizing a reporter gene encoding E. coli β-galactosidase.

Figure 4 demonstrates the suppression by Vitamin D_3 of a novel negative regulatory element from the human K1 keratin gene (HK1.NRE).

Figure 5 is a schematic representative of the constitutive epidermal vector which can be suppressed by Vitamin D_3 via insertion of the HK1.NRE.

Figure 6 is a schematic drawing of a derivative of the mouse K6 keratin gene (BCM-MK6(A)-HK1).

Figure 7 shows the expression characteristics of BCM-MK6(A)-HK1 in transgenic animals.

Figure 8 is a schematic drawing of the mouse K6 keratin gene and the proposed construction of an inducible epidermal vector from its regulatory sequences.

Figure 9 is a schematic representative of the inducible epidermal vector which can be suppressed by Vitamin D_3 via insertion of HK1.NRE.

The drawings are not necessarily to scale, and certain features of the invention may be exaggerated in scale and shown in schematic form in the interest of clarity and conciseness.

DETAILED DESCRIPTION OF THE INVENTION

It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention.

The term "transformed" as used herein refers to the process or mechanism of inducing changes in the characteristics (expressed phenotype) of a cell by the mechanism of gene transfer whereby DNA is introduced into a cell in a form where it expresses a specific gene product or alters expression of endogenous gene products.

The term "transduction" as used herein refers to the process of introducing a DNA expression vector into a cell. Various methods of transduction are possible, including microinjection, $CaPO_4$, lipofection (lysosome fusion), use of a gene gun and DNA vector transporter.

The loricrin constitutive vector and the keratin K6 inducible vector can be transduced into the squamous epithelia cells by any of the variety of ways described above. The types of epithelia cells include epidermis, oral, esophageal, vaginal, tracheal, corneal and other squamous epithelia. They are transduced by contacting the vector with the cells. In the preferred embodiment this includes using a gene gun or DNA vector transporter.

The term "DNA vector transporter" as used herein refers to those molecules which bind to DNA vectors and are capable of being taken up by epidermal cells. DNA transporter is a molecular complex capable of non-covalent binding to DNA and efficiently transporting the DNA through the cell membrane. Although not necessary, it is preferable that the transporter also transport the DNA through the nuclear membrane.

The term "nucleic acid cassette" as used herein refers to the genetic material of interest which can express a protein polypeptide or RNA and which is capable of being incorporated into the epidermal cells. The nucleic acid cassette is positionally and sequentially oriented within the keratin K6 inducible vector or the loricrin constitutive vector such that the nucleic acid in the cassette can be transcribed into RNA or antisense RNA and, when necessary, translated into proteins or polypeptides in the transformed epidermal cells. A variety of proteins and polypeptides can be expressed by the sequence in the nucleic acid cassette in the transformed epidermal cells. These proteins or polypeptides which can be expressed include hormones, growth factors, enzymes, clotting factors, apolipoproteins, receptors, drugs, tumor antigens, viral antigens, parasitic antigens and bacterial antigens. Specific examples of these compounds include proinsulin, insulin, growth hormone, insulin-like growth factor I, insulin-like growth factor II, insulin growth factor binding protein, epidermal growth factor TGF-α, dermal growth factor PDGF, angiogenesis factors, e.g., acid fibroblast growth factor, basic fibroblast growth factor and angiogenin for instance, matrix proteins such as Type IV collagen, Type VII collagen, laminin and proteins from viral, bacterial and parasitic organisms which can be used to induce immunologic response.

The genetic material which is incorporated into the epidermal cells using the loricrin constitutive vector or the keratin K6 inducible vector includes DNA not normally found in epidermal cells, DNA which is normally found in epidermal cells but not expressed at physiological significant levels, DNA normally found in epidermal cells and normally expressed at physiological desired levels, any other DNA which can be modified for expression in epidermal cells, and any combination of the above.

The term "loricrin constitutive vector" as used herein refers to a vector which can be inserted into epidermal cells and which once inserted,

will express a constitutive (*i.e.*, a constant level) of protein, polypeptide or antisense RNA from the nucleic acid cassette which is part of the loricrin constitutive vector. The loricrin constitutive vector is used for efficient expression of a nucleic acid sequence in epidermal cells and is comprised of a 5′ flanking region of the loricrin gene, said flanking region including a TATA box, a cap site and a first intron and an intron/exon boundary all in appropriate sequential and positional relationship for expression of a nucleic acid cassette; a 3′ flanking sequence of the loricrin gene; and a linker having a unique restriction endonuclease site at the location of the start and stop codon, said linker connecting the 5′ flanking region to the 3′ flanking sequence and said linker further providing a position for inserting the nucleic acid cassette.

The sequence for the loricrin gene which is used for preparing the loricrin constitutive vector is shown in SEQ. ID No. 1. The loricrin constitutive vector has a 5′ flanking region comprising nucleotides 1 to 1540 of SEQ. ID. No. 1; an intron and intron/exon boundary comprising nucleotides 1587 to 2677 of SEQ. ID. No. 1, a 3′ flanking region comprising nucleotides 4384 to 6530 of SEQ. ID. No. 1; and a linker to be inserted at the unique Cla I site at nucleotides 2700 to 2705 of SEQ. ID. No. 2. The loricrin constitutive vector has a 5′ flanking region of approximately 1.5 kb, an intron of approximately 1.1 kb and a 3′ flanking sequence of approximately 2.1 kb. The linker of the loricrin constitutive vector can be a poly-linker. The poly-linker includes a plurality of restriction endonuclease sites.

The term "keratin K6 inducible vector" as used herein is a vector which is useful for regulated expression of a nucleic acid sequence in epidermal cells. The keratin K6 inducible vector comprises a 5′ flanking region of the keratin K6 gene, said flanking region including a TATA box, a cap site, a first intron and an intron/exon boundary all in sequential and positional relationship for the expression of a nucleic acid cassette; a 3′

flanking sequence of a keratin K6 gene; and a poly-linker. The poly-linker includes a plurality of restriction endonuclease sites, connects the 5′ flanking region to the 3′ flanking sequence and further provides a position for insertion of the nucleic acid cassette.

The partial sequence for the keratin K6 gene which is used for preparing the keratin K6 inducible vector is shown in schematic form in Figure 8 and the sequence is shown in SEQ. ID No. 3. The keratin inducible vector has a 5′ flanking region which extends from a unique 5′ Xho I site up to nucleotide 360 of SEQ. ID. No. 3; an intron and intron/exon boundary comprising nucleotides 928 to 1494 of SEQ. ID. No. 3; a 3′ flanking region which extends from nucleotide 4740 of SEQ. ID. No. 3 to a unique 3′ Xho I site; and a poly linker inserted between nucleotides 1504 to 1509 of SEQ. ID. No. 3.

The keratin K6 inducible vector has a 5′ flanking region of approximately 8.0 kb, an intron and intron/exon boundary of approximately 0.56 kb and a 3′ flanking sequence of approximately 1.2 kb. The restriction endonuclease sites found in the linker and poly-linker of the loricrin and keratin K6 vectors can be any restriction endonucleases which will allow insertion of the nucleic acid cassette. In the preferred embodiment they are usually selected from the group consisting of Cla I, Not I, Xma I, Bgl II, Pac I, Xho I Nhe I and Sfi I.

One skilled in the art will readily recognize that there are a variety of ways to introduce the loricrin constitutive vector or the keratin K6 inducible vector into epidermal cells. The vectors can be inserted either *in vivo* or *ex vivo*. The mode of insertion will, to a certain degree, determine the available methods for the insertion.

One embodiment of the present invention includes a bioreactor. A bioreactor is comprised of transformed epidermal cells which contain the loricrin constitutive vector or contain the keratin K6 inducible vector. Once the vector is inserted in the epidermal cells, the epidermal cells will

express the nucleic cassette and produce the protein, polypeptide or antisense RNA of interest. This can be done either *in vivo* or *ex vivo*. Any compound which can be encoded in, and expressed by, the nucleic acid cassette can be produced by the bioreactor.

One method for *ex vivo* introduction of the loricrin constitutive vector or the keratin K6 inducible vector into epidermal cells includes a cotransfection of the vector with a selectable marker. The selectable marker is used to select those cells which have become transformed. The cells can then be used in any of the methods described in the present invention.

One specific embodiment of the present invention is a method for the enhanced healing of a wound or surgical incision. This method comprises the *in vivo* transduction of epidermal cells with a loricrin constitutive vector or a keratin K6 inducible vector. In either case, the nucleic acid cassette of said vector contains a nucleic acid sequence for a growth factor.

In the preferred embodiment for the treatment of wounds or surgical incisions, a plurality of vectors are introduced into the epidermal cells. In the plurality of vectors, the cassette of at least one vector contains a nucleic acid sequence for an epidermal growth factor (TGF-α), the cassette of at least one vector contains a dermal growth factor (PDGF), a cassette of at least one vector contains a nucleic acid sequence for a matrix protein to anchor the epidermis to the dermis, and a cassette of at least one vector contains a nucleic acid sequence for an angiogenesis factor. The sequence for matrix proteins can be selected from any sequences useful for the anchoring of the epidermis to the dermis but are usually selected from the group consisting of Type IV collagen, laminin, nidogen, and Type VII collagen. The angiogenesis factor is usually selected from the group consisting of acid fibroblast growth factor, basic fibroblast growth factor and angiogenin. The combination of the vectors

provides all of the necessary elements for quick and rapid enhancement of healing of wounds or surgical incisions. This procedure is very helpful in the case of plastic or reconstructive surgery. Furthermore, skin ulcers can be treated by following similar procedures as described for wound healing or surgical incision. These procedures are useful in animals and humans.

In the *ex vivo* approach for treating or healing wounds, surgical incisions and skin lesions, the vectors are first transduced into the epidermal cells *ex vivo*. The transformed epidermal cells are transplanted onto the animal or human to be treated.

Another embodiment of the present invention is a method for treating psoriasis. In this method, epidermal cells are transduced in *vivo* with a loricrin constitutive vector or a keratin K6 inducible vector. A nucleic acid cassette in said vector contains a nucleic acid sequence for a protein or polypeptide selected from the group consisting of TGF-β, a soluble form of cytokine receptor, and an antisense RNA. The cytokine receptor can be selected from the group consisting of IL-1, IL-6 and IL-8. The antisense RNA sequence is selected from the group consisting of TGF-α, IL-1, IL-6 and IL-8.

In another embodiment of the present invention there is a method of treating cancer. This method comprises the steps of *in vivo* transduction of epidermal cells with a loricrin constitutive vector or a keratin K6 inducible vector into epidermal cells. The nucleic acid cassette of either vector contains the nucleic acid sequence coding for antisense RNA for the E6 or E7 gene of the human papilloma virus or coding for the normal p53 protein. Although the example given is for skin cancer, this same approach is used for cancers occurring in other squamous epithelial, since the constitutive and inducible vectors will also function in these tissue types.

It has been found that either the keratin K6 inducible vector or the loricrin constitutive vector can be further regulated by introducing the Vitamin D regulatory element into the vector. The Vitamin D regulatory element is usually introduced into the 3' flanking sequence. In the present invention, the Vitamin D regulatory element is from the human K1 keratin gene. With the Vitamin D regulatory element in the vector, the expression of the nucleic acid cassettes can be suppressed by Vitamin D, a commonly used substance in animals and humans.

An additional embodiment of the present invention is a method for vaccination comprising the step of *in vivo* introduction of a loricrin constitutive vector into epidermal cells. The nucleic acid cassette in the vectors usually codes for a polypeptide which induces an immunological response. An example of this is the viral capsid protein from the human papilloma virus. One skilled in the art will readily recognize that any other variety of proteins can be used to generate a immunologic response and thus produce antibodies for vaccination.

The following examples are offered by way of illustration and are not intended to limit the invention in any manner.

EXAMPLE 1

Isolation of the Mouse Loricrin Gene

Although it is a major keratinocyte cell envelope protein, loricrin was not identified until 1990 (Mehrel, et al., Cell, Vol. 61, pp. 1103-1112, (1990)). The primary sequence of the loricrin protein was deduced from the overlapping cDNA clones described in Mehrel, *id.* To obtain the full gene, the cDNA clones were used to screen an EMBL-3 Balb/c mouse genomic library. The gene encoding loricrin was located within two Bam HI fragments of 3.4 and 3.1 kb. The coding sequence within this genomic fragment is identical to the cDNA sequences and is not interrupted by introns. There is, however, an intron in the 5' non-coding region that is

approximately 1.1 kb in length. In addition to the intron and coding sequence, there is approximately 1.5 kb of 5′ flanking sequence and 2.1 kb of 3′ flanking sequence.

EXAMPLE 2

Construction and characterization of a constitutive epidermal expression vector from the mouse loricrin gene

Although all of the regulatory elements of the loricrin gene have not been identified, a functional loricrin constitutive expression construct was designed as follows. Briefly, polymerase chain reaction (PCR) technology was used to delete the loricrin coding region, leaving the 5′ and 3′ flanking regions, 5′ and 3′ non-coding regions and the intron (Figure 1). A unique Cla I restriction site was engineered at the start (ATG) and stop (TAA) codons to allow easy insertion of exogenous gene cassettes. To assess the expression characteristics of this vector, a reporter gene, the bacterial gene encoding chloramphenicol acetyl transferase (CAT), was inserted into the Cla I site. The expression vector was analyzed by transient transfection into primary mouse epidermal cells. Positive (pSV2.CAT, lane 1) and negative (pA10.CAT, lane 2) control vectors were included in the assay (Figure 2). The loricrin expression vector had high activity in undifferentiated (low Ca^{2+} medium, lane 3) and differentiated (high Ca^{2+} medium, lane 4) epidermal cells, surpassing levels obtained with the strong promoter of the virus SV40. This result was unexpected, since previous *in vivo* studies had demonstrated that the loricrin gene was only expressed at a late stage of epidermal differentiation (Mehrel, et al., Cell, Vol. 61, pp. 1103-1112, (1990)), and indicates that additional flanking sequences are required to suppress loricrin expression in undifferentiated epidermal cells.

To analyze the expression characteristics of the loricrin vector *in vivo*, the bacterial gene encoding β-galactosidase was inserted into the Cla

I site. The β-galactosidase gene has frequently been used as a reporter gene to assess targeting specificity (MacGregor, et al., In: Methods in Molecular Biology, Vol. 7, pp. 217-235, (1991)). This construct was designated pML-β-gal and was used in the production of transgenic mice. This construct was digested with Apa I and subjected to preparative agarose gel electrophoresis to purify the pML-β-gal expression construct away from plasmid sequences (pGEM72) which might interfere with expression. The separated expression construct sequences were purified and recovered using NA 45 DEAE membrane (Schleicher & Schuell). DNA was precipitated and resuspended at 1-3 ng/ul. ICR outbred female mice (Sasco) were given PMS and HCG to stimulate superovulation, mated to FVB males (Taconic) and resulting one-cell fertilized embryos were collected from the oviducts. DNA was micro-injected into the pronuclei and the embryos were surgically transferred to pseudopregnant recipient females (the result of mating ICR females with vasectomized $B_6D_2F_1$ males (Taconic). Normal gestation and birth was allowed to continue and at approximately three weeks of age the pups were screened for evidence of the transgene using total genomic DNA extracted from the tail.

PCR analysis was performed on the extracted tail using oligo primers specific for β-galactosidase. Animals positive for the transgene were further analyzed to assess the expression characteristics of pML-β-gal. This was done by removing part of the ear and incubating the tissue in a staining solution containing X-gal. This was done by removing part of the ear and incubating the tissue in a staining solution containing X-gal. Typical results are seen in Figure 3 where a PCR positive animal expressed high levels of β-galactosidase in the epidermis (Figure 3b) while a PCR negative animal shows no such staining (Figure 3a) indicating that endogenous murine β-galactosidase is not expressed at sufficient levels in the epidermis to cause false positives in this assay. Intense X-gal staining

was detected in the basal compartment as well as the suprabasal, more differentiated layers.

To analyze the expression characteristics of the loricrin vector *in vivo*, the bacterial gene encoding β-galactosidase was inserted into the Cla I site. This data is shown in Figure 3. This observation indicates that the loricrin expression vector is useful as a constitutive vector to direct the efficient expression of exogenous DNA in both the undifferentiated and differentiated compartments of the epidermis.

EXAMPLE 3

Utilization of a novel Vitamin D_3 responsive element to modulate expression levels in the epidermis

This example demonstrates that a novel negative regulatory element from the human K1 keratin gene (HK1.NRE) is able to suppress a heterologous promoter in response to Vitamin D_3. The HK1.NRE is 70 nucleotides in length (see Figure 4). PCR technology was used to generate Bam HI and Bgl II sites at opposite ends of this fragment. This facilitates generating multiple copies of this fragment since ligation and digestion with Bam HI and Bgl II will select for oligomers which have ligated head to tail. Four tandem copies of the HK1.NRE were inserted into the Bgl II cloning site of pA10.CAT. In the absence of Vitamin D_3 this construct is highly expressed when transfected into primary mouse epidermal cells (Figure 4). The addition of increasing concentrations of Vitamin D_3 to the culture medium completely suppresses transcription of this heterologous promoter. Thus, by using Vitamin D_3, the activity of the expression vector is modulated. Figure 5 shows a schematic representative of a derivative of the loricrin constitutive epidermal vector which contains the HK1.NRE in its 3′ flanking region. The activity of this vector within epidermal cells

can be suppressed by topical application of Vitamin D_3, or an analogue, to the skin.

EXAMPLE 4

Isolation and characterization of a Mouse K6 Keratin Gene

Several laboratories have reported that keratin K6 is not expressed in normal epidermis, but is expressed under hyperproliferative conditions such as wounding (Weiss, et al., J. Cell Biol., Vol. 98, pp. 1397-1406, (1984); Nakazawa, et al., J. Cell Biol., Vol. 103, pp. 561a (1986); Stoler, et al., J. Cell Biol., Vol. 107, pp. 427-446, (1988)) or topical application of retinoic acid (Rosenthal, et al., J. Invest. Dermatol., Vol. 95, pp. 510-515, (1990). Although K6 expression does not occur in interfollicular epidermis, it does occur in hair follicles (Nakazawa, et al., J. Cell Biol., Vol. 103, pp. 561a, (1986)). Recent results indicate that there are two K6 cDNAs that differ in sequence in only a few nucleotides. These cDNA clones have been used to differentially screen a EMBL 3 Balb/c mouse genomic library and isolate two distinct K6 genes. These genes are closely linked within genomic DNA, i.e., arranged in tandem. They have almost identical 3′ halves, including identical 3′ non-coding and flanking regions. Interestingly, the 5′ halves of the 2 genes differ greatly in their restriction fragment patterns. Sequence analysis of the region near the ATG shows many differences between the two genes. The sequence of one of these genes, designated BCM-MK6(A), is shown in SEQ. ID. No. 3. To determine the expression characteristics of this gene *in vivo* in transgenic mice, PCR technology was used to modify a 13.5 kb Xho I fragment containing BCM-MK6(A). Nucleotides encoding the C-terminal region of the K6 protein were deleted and nucleotides encoding the amino acid sequence SEQ. ID. No. 4 were inserted. These amino acids are at the C-terminal of human keratin K1 (Johnson, et al., PNAS, USA, Vol. 82, pp. 1896-1900, (1985)). A schematic representative of this derivative of the

mouse K6 gene (BCM-MK6(A)-HK1) is shown in Figure 6. Antisera have previously been generated against the HK1 C-terminal peptide (Rosenthal, et al., J. Invest. Dermatol., Vol. 95, pp. 510-515, (1990)). These antibodies are monospecific for this human K1 peptide and allow expression of the derivatized BCM-MK6(A)-HK1 transgene to be followed against the expression pattern of the endogenous mouse K6 genes.

The derivatized mouse K6 transgene shown in Figure 6 was used in the production of transgenic mice as outlined in Example 2. Mice resulting from the initial injections were screened by PCR analysis for presence of the BCM-MK6(A)-HK1 transgene. Positive founders were initially analyzed for transgene expression as follows. A small ear biopsy was taken and after 48 hours a second biopsy was taken at the same site to score for expression during wound healing. Transgene expression was limited to hair follicles in the initial biopsy and was not present in interfollicular epidermis. Transgene expression was observed in the epidermis in the 48 hour biopsies, but only at the site of wounding. To further confirm the inducibility of the BCM-MK6(A)-HK1 transgene under hyperproliferative conditions, F1 generation offspring from the initial founders were treated topically with the hyperplasiogenic agent 12-O-tetradecanoylphorbol-13-acetate. Biopsies were taken before and 48 hours after topical application of this agent. Immunofluorescence was performed on frozen sections of these biopsies with antisera specific for the HK1 peptide. No expression was observed prior to the induction of hyperplasia, however, the BCM-MK6(A)-HK1 protein was expressed at very high levels in all layers of the epidermis 48 hours after hyperplasia was induced (Figure 7).

EXAMPLE 5

Construction of an inducible epidermal expression vector from the mouse K6 gene (BCM-MK6(A)

Results obtained with the derivative of BCM-MK6(A) (Figure 7) indicate that all of the regulatory sequences required to suppress expression of this gene in normal epidermis and activate its expression under hyperproliferative conditions, such as in wounding healing or experimentally induced hyperplasia, are located within the 13.5 kb Xho I fragment (Figure 6). Therefore, an inducible vector was developed from this fragment. This vector is very useful in gene therapy applications where dosage of pharmaceuticals needs to be regulated. In addition, this vector is ideally suited for wound healing applications since it is induced during the wound healing process but suppressed after healing has occurred. Figure 8 illustrates how a vector is constructed from the BCM-MK6(A) gene. The vector is derived from the 13.5 kb Xho I fragment which contains the entire K6 gene. The same general strategy used in construction of the constitutive epidermal vector (Figure 1) is followed. The expression vector retains all of the 5′ flanking sequences, the 5′ non-coding sequences up to but not including the ATG, the first intron including the splice-sites of the intron-exon boundary and all of the 3′ non-coding and flanking sequences after the TAA codon. A polylinker is engineered 3′ of the first intron to allow easy insertion of exogenous DNA cassettes. These manipulations are performed through the use of PCR technology. Unique Xho I sites are conserved at the ends of the vector to allow easy amplification in pGEM vectors and excision for purification from plasmid sequences. Recent *in vivo* results indicate that the endogenous human K6 gene is inducible after topical application of all-trans retinoic acid. Further, *in vivo* mouse experiments suggest that the vector shown in Figure 9 is inducible by topical application of retinoic acid, or an analogue, to the skin.

EXAMPLE 6

Construction of a derivative of the inducible epidermal vector which could be suppressed by Vitamin D_3

Even though the inducible epidermal vector depicted in Figure 8 is suppressed or silent in normal epidermis, it can be accidently induced as a result of injury. Therefore, it is desirable to have an additional suppressor engineered into this construct. In addition, this suppressor is used to more tightly regulate pharmaceutical delivery. This is achieved by insertion of the HK1.NRE described in Figure 4. Figure 9 shows a schematic representative of a derivative of the K6 inducible epidermal vector which contains the HK1.NRE in its 3' flanking region. The activity of this vector within epidermal cells is suppressed by topical application of Vitamin D_3, or an analogue, to the skin.

EXAMPLE 7

Utilization of the inducible epidermal vector in wound healing

Greater than 3.5 million individuals develop skin ulcers. During normal healing, epidermal cells produce growth factors which affect not only epidermal cells but also cells within the dermis. In addition, epidermal cells synthesize several matrix proteins which provide an anchor to the underlying dermis. Many skin ulcers occur in patients with disorders such as circulatory problems and diabetes, and the normal healing process in impaired. The inducible epidermal vector is used to target the combined expression of growth factors, to accelerate growth of cells in both the epidermal and dermal compartments; matrix proteins, to increase tensile strength; and angiogenesis factors, to improve circulation, in an attempt to improve healing these patients.

EXAMPLE 8

Utilization of the constitutive epidermal vector in gene therapy approaches to cancer

Skin cancer is by far the most common form of cancer with greater than 600,000 new cases reported each year. Several genes have been implicated in causing skin cancer, including loss or mutation of the host cell tumor suppressor gene, p 53 and expression of the E6 and E7 transforming genes of human papilloma virus (HPV). *In vitro* studies suggest that the normal or wild type p53 gene can revert the phenotype of malignant cells or induce programmed cell death. The constitutive epidermal vector is used to target expression of the normal p53 gene to cause reversion to a non-malignant phenotype or induction of programmed death *in vivo*. In cancers where HPV is suspected of being the etiological agent, the constitutive vector is used to target expression of antisense RNA specific for the E6 and E7 genes of HPV.

EXAMPLE 9

Utilization of the epidermal vector systems in gene therapy approaches to psoriasis

Psoriasis is a common inherited skin disease which affects approximately 4 million individuals in the U.S., 20 million world-wide. It is characterized by the presence of inflamed scaly skin. Although the specific defect for psoriasis is not known, inappropriate expression of growth factors, and cytokines appears to be responsible for its pathogenesis. Epidermal vectors are used to inhibit the mitogenic effects of positive growth factors produced in psoriatic lesions by expressing negative growth factors which induce growth arrest of epidermal cells.

The inflammation observed in psoriasis most likely results from inappropriate expression of cytokines. Targeted expression of soluble cytokine receptors prevents stimulation of an inflammatory infiltrate in

this disease. In another approach, antisense RNA is directed against transcripts of positive growth factors or cytokines. These approaches have therapeutic potential for other dermatoses resulting from inflammation.

All patents and publications mentioned in this specification are indicative of the levels of those skilled in the art to which the invention pertains. All patents and publications which are incorporated herein by reference are incorporated to the same extent as if each individual publication was specifically and individually indicated to be incorporated by reference.

One skilled in the art will readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The bioreactors, nucleic acid sequences, transformed epidermal cells, loricrin constitutive vector and keratin K6 inducible vector, along with the methods, procedures, treatments, molecules of specific compounds described herein are presently representative of preferred embodiments, are exemplary and are not intended as limitations on the scope of the invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention as defined by the scope of the claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Roop, Dennis R.
Rothnagel, Joseph A.
Greenhalgh, David A.

(ii) TITLE OF INVENTION: CONSTITUTIVE AND INDUCIBLE EPIDERMAL VECTOR SYSTEMS

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fulbright & Jaworski
(B) STREET: 1301 McKinney, Suite 5100
(C) CITY: Houston
(D) STATE: Texas
(E) COUNTRY: U.S.A.
(F) ZIP: 77010-3095

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Paul, Thomas D.
(B) REGISTRATION NUMBER: 32,714
(C) REFERENCE/DOCKET NUMBER: D-5405

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 713/651-5325
(B) TELEFAX: 713/651-5246
(C) TELEX: 762829

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6530 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGATCCTGAT ATAGCTGTCT CTTTTGAGAC TATCCCGGGG CCTAGCAAAC ACAGAAGTGG 60

ATGCTCACAG TCAGGGATTG GGTGAATCAC AGGGCCCCCA ATGTTGGAGC TAGAGAAAGA 120

ACCCAGGGAG CTGGGGGGAT CACCTGAGTT CATACTGTCC AAACTGAAAC AAGTGGCACA 180

AGTTTCTGAG AGCCAAAGTC TAATCAGGAT CGTTTAGATC ATTAATGCTC CCCCATAATT 240

AAGACAATTT CTGATTAGAA TTATTCTTTC AACACAGCTG GGTGGAACAA GGTTCAACAG 300

TGGTATCTTA ATAGCAACTG AGTTCCAATG ATGAAAGAAA GGAAAAACAC TATGTTCTTC 360

ATACACAGAG GGGGGCTGCT CTTGGCCCTA GGGTCATCAG AGAACTGAGT AAATCTTATA 420

GGAAAATAGT TAAGATGTCT TCACACACCT CCTTTCCAAT AGGGTTCAAG GGCAGGCATG 480

ATTGGAAGGA AAAGTGTTCT GTCATGTGAG AAAAGAGCAA AAGTATTAAT ATCACATACT 540

ATGTAGTACA TTCATATTTC ATAACTTCCA TTTTCATGTT TCTGTGAAAT AAATTATAGG 600

ATTCCTGCTT GGTAGACCAA ATGGGGATCA GACAGCTCAA CAATGAACAA GTACTCAGTA 660

ACTGCCCTGT TGGTGGCATT GCATGAACTA CTGTGCTTTG CCCATGGTGA CATAGCTTGA 720

AATAGTAATG GAAGACCTGA ACCCAACTGA GATCTCTAAG TACATTCCAC TCTATGGTGG 780

CATCTCAGAG GTCAGAGTCA CTGTGCAGCG CCATAGGACA TCAGAATCAA AGGGTCATGG 840

TGAAAAGGCT GCCAGGGTCT GTCTTGTTAG TTCTCACCTT TGTAAGTAAA GTCAGTAGTC 900

AGTAACAAAG ATCAAAACAC CTGCTCTCAC AAGGAATAAC TTAAAGTAGA CTAAAGTCAT 960

GCTAGTTACA GTGCTGTCTT TTCCGTGGTA CCATCCCAAA CTGGGAGCTG GGGACTCACG 1020

AACTCTCACA ACCAATAAAG TAAGCAGAAC AGAAGCAACC CAATGAAGTG TTCATGAAAC 1080

TGGAATGGAG AAATTGTGGC ATAAGAGATG GATTCTAAAA TTTTGAGAAT TTCCAAGATA 1140

ATGAAATTAA AACCAAACAT CAAAATTGGA AAGATACAAC TGAACTAGCT TCTATGTCTT 1200

AGACAATGTC TTAGATCTCT AGATTCCGTA AGGCTGCTTC ACAAGTCTGC AACCTAGTCC 1260

TCTAGAATAG CCCTCTGGTT ATGGCACGCA ACCTATACAG AAGTTTTGAA AACAATTTCT 1320

GCCATCCACA CTGCTGGCCA TCTCTAATGA CCAACCTGCT CACTGTTACA TCAGAGAAGT 1380

GGCCAGTCAT ACACCAAACT GCCTATCCCT ATCCCAAGAA TTTGAAATCT TCATGAATGG 1440

GTCAATCCTT CCCCTGCAAT CACAGGGAGG AGGTGCCTGA TCAATAGATG AGTCAGAGCA 1500

GGACAAGAGT ATAAAACACA GGAGCACCAG TGTCCCTCAC ATCAGCATCA CCTCCTTCCC 1560

TCACTCATCT TCCCTGGTGC TTCAGGTAAG TGTGGGCTCT CCTGGCTGTC TGGTCTCTCC 1620

AGTTGGCCTT GCTCAGCTTG CAGAGAGGTT AAGGAACAGA GCCTTTCTCC CCTTTGGAAG 1680

GTACTCTGTT CAAATTGAGA AGGGCTTTAG GAAAGCACTG GGAGAGTGGT AAGCTGGTGC 1740

TGGGCAGATG ATGTGTCTGG TCTTCTGGGC AGAATGTTAA AACTTCACAA AGATATGACT 1800

ATCTCCTACT TCTCTGGCAC CCTGGGAGCT GAGGGTTAGA ATACTGGATG ACTGCAGTGG 1860

CAGGCCTCCA TGGGCTGGAT GAACCTTTTG AACCTGCCAG AAGTGGCTGA ATACACTATC 1920

AGGAAGGGAG AGGGACGATA AGTCATAGAA TGGTGCTGAT GGGAGATTTG AGAAGCCACA 1980

AAAACCCAAG CTCTGCTTTA TGAGGGCAGA TGTTCTGACA GATAAATGAC TTGTGAGGTG 2040

CTGAACTACA CAGCTTCCTA TTAGCTACAG CTAATTGGAG TCTACCAAAT TTAGACTCCT 2100

GCATATCTCA AAAAGATGTC TACTTTCTTC TGGTTAGATG TACTGGTCCA AAAGGTTCAG 2160

AGTTCTTCCA TTTGTTTGCA GACAGGACCA CAGTAGAGCT GTCTTGTCTA ATAATTGGCC 2220

CTTGGAGGAT ATCTCACTCA ATAGGACAGA TCAAGAGTTT AAACTAAGGA CTTTATACAG 2280

GAAATGCTAA TGTCCAAACA AATCTTTTCT TATTGTGCTG GGAGTGGATA AAATCCACGT 2340

GGAATTTTTG CAACTTTCTA CTGAATTTAA AGAATCAGCA CTGGGACTTG GGAGCACCCT 2400

TAGACATGGA GTGTTTATTA ATGTAAGATC AAAAGCAGGT GGGAATGTGG GGGTTCTGCT 2460

TCCCAAATCA CATAGTAGAA GAAAGGCAGA GTTGAGGGAA AAGGGGGTCA CTATTAACGG 2520

GACTTTTGAA GAGCTAACCA GTCCAGGAAT GGAGTCCAGA CACCTAGTCT GCATAAAGCT 2580

AGGAGTCAGA AGTATGTTGG CATGGATGCA TCTGCCACCT TCACAGCGTC CTCTTGCTGC 2640

TGTTGGTCTA ATGTTGCTCT TCTGCTCTTC TTCCAGGGTT CCCCTTCTCC TTAAACAAGA 2700

TGTCTCACCA GAAAAAGCAG CCCACTCCCT GCCCTCCTGT GGGTTGTGGA AAGACCTCTG 2760

GTGGAGGAGG AGGCGGCGGC GGCTATTATA GCGGTGGCGG CTCTGGCTGC GGAGGCGGCT 2820

CATCTGGAGG AGGCTCTAGC TGTGGAGGCG GAGGCGGTGG TTCCTATGGA GGTGGTTCCA 2880

GCTGCGGCGG TGGAGGCGGC TCCGGTGGGG GCGTCAAGTA CTCCGGAGGC GGCGGTGGCT 2940

CTAGCTGCGG CGGCGGCTAC TCCGGAGGCG GTGGTGGCTC TAGCTGCGGC GGTGGCTACT 3000

CTGGGGGCGG CGGCGGCTCC AGCTGCGGAG GTGGCTACTC CGGAGGCGGC GGCGGCTCCA 3060

GCTGCGGCGG CGGCAGCTAC TCCGGGGGTG GCTCCAGCTG TGGAGGCGGT GGCGGCTCTG 3120

GTGGGGGCGT CAAGTACTCC GGAGGTGGTG GCGGCGGCGG CTCTAGCTGC GGCGGCGGCT 3180

CCTCCGGGGG CGGCGGCGGC GGCTCCAGCT GCGGAGGCGG ATCAGGAGGC GGCGGCTCCT 3240

ACTGCGGAGG CTCCTCTGGA GGCGGCAGCT CCGGTGGCTG CGGCGGCGGT TCCGGAGGCG 3300

GCAAGTACTC TGGTGGCGGC GGTGGCTCCA GCTGCGGAGG CGGCTATTCC GGCGGCGGTG 3360

GAAGCAGCGG CGGCTCTAGC TGTGGCGGCG GCTACTCAGG TGGCGGTGGA TCCAGCTGCG 3420

GCGGCGGCGG CGGCTATTCC GGTGGCGGCG GCACGAGCTG CGGAGGTGGT TCCTCCGGTG 3480

GCGGCGGCGG CGGATCGTCC CAACAGTATC AGTGCCAGAG CTACGGAGGC GGTTCTAGCG 3540

GTGGCTCCAG CTGCGGCGGC GGCTACTCCG GGGGCGGAGG CTCCAGCTGC GGTGGCGGCT 3600

ACTCCGGGGG CGGAGGCTCT AGCTGCGGAG GCGGCTCCTC TGGTGGTGGC TCCAGTTGCG 3660

GCGGCAGCGG CGGCGGCGGC TATTCCGGTG GTGGCGGTGG CAGCTGCGGC GGCGGCTCCT 3720

CTGGCGGCGG AGGGGGCTAT TACTCCTCTC AGCAGACCAG TCAGACCTCC TGCGCCCCCC 3780

AGCAGAGCTA CGGAGGGGGC TCTTCCGGAG GAGGTGGTAG CTGTGGAGGT GGCTCCTCTG 3840

GCGGCGGTGG CGGCGGTGGC TGCTACTCCA GCGGTGGTGG CGGCAGCAGC GGTGGCTGCG 3900

GTGGAGGCTA CTCCGGAGGC GGCGGTGGCT GTGGCGGCGG CTCTTCCGGG GGCAGCGGCG 3960

GTGGCTGCGG AGGTGGCTCT TCCGGAGGCA GCGGCGGTGG CTGCGGAGGA GGCTACTCCG 4020

GAGGCGGAGG CGGTGGCTCC AGCTGCGGAG GCGGCTCCTC TGGTGGCGGC TCTGGAGGTG 4080

GCAAGGGTGT GCCAGTCTGC CACCAGACCC AGCAGAAGCA GGCGCCTACC TGGCCGTGCA 4140

AGTAAGGTCA CCGGGTTGCA ACGGAGACAA CAGAGCTGGA AGAGTTCTCC GTGGGCGCCG 4200

ATGGGCTTAA CTTTCTCATG AATTTGCCTG AGGTTTCCAA ACCCTTCACA TTTTAAGCGC 4260

CCCTTCCCCC AGAAGAAGCC ATTGAGTCGC TCAAGGTGTA TCCTGTTCTG CAGATTTTTC 4320

ATCTTGGTTT CTGAATGACT ACCTCCCAAT TCTAGTGTCT CCTCAGTCAA TAAATTTGCT 4380

ATTCATGAGA ATCTCTGAGT TTGCTGTAGT CTTTGTAGCT TGCAAATTTA CTCAGTTCAT 4440

TCTGTGTTTG CTTTTTCCAT TCATTAGTTC ACATTTAAAT TCACTGAACA AGTGTTCTAT 4500

CCCAAGGTGG GGGAGTAGAT AGATGGAATG GGGCAAAGGA TGACCAAGGT TGTGAACAGT 4560

CTGGGGTGTG GCTTAAAAAT CATGAGATGG TCCTCAAACA CCAAGAAAAG TCTTCACTGG 4620

ACATCCTACA CATCACTGAA ATTGGGCCTG CGCAGGCAAT TTCTAGCAGT GCAGAGTTCA 4680

CTCTCCAAGT TCTGGAAGCA GGATGGCTCT CAGATTAGGT TAGCTACCAG AGGTCCAAGT 4740

CCACTGACAT GTTCTGACCT AAGAAGAAGG ACATTCACCC CTGAACAAAA GACCCCTGCC 4800

CATGCGATCT TCCGGAACAC TATAACTACT TTCCTTACTC ATGACCCATG ATAGAGCTTT 4860

GAGGCAAAGA TACAAACCCT CTATGTCTTC TCAAGATTGC CAGTTCTTCA TTAAGCCTGA 4920

TACCTTCTTA CCAGCGCACG TCTCCTGAAT ACTGATAAAG TCTGGTTTTG TTAGTCTGTT 4980

AGAAAAATAT TATATCAGAT AATCAAGATC CTCTACAGTG TGTGAGACAG TTTACTGAGC 5040

ATCTATAGAG ATAGAAGGCA GCCCTCTTGA AGGATTGAAC GCGTACGTTT CGTCCAATTT 5100

GAGAAGGTAC ATCGTAAGTA TTTAAGATGC TTAACATCAG TATCACAGAG GTCACTGGAA 5160

ACATTAGGGG CCTCCTGATT AGCAAGCATA AAGCTAGAGT TGCTCAAAGG CATGTGTAAC 5220

AACCATCCCC TGGCCAGATC CTGTTTTACA GTCAGATTTT ATCAGCTTTA GGTAAATGCT 5280

AACTTACTGA CTTACTCAAG TTAATTTTGC TATACTAAAA AGCCAATGTG CCTTCCTACA 5340

TTTAGCTAAT GATAGAAATA AAAAGATTTC ATCTCACTCT TCCATTTGGA GTCATCACTA 5400

CCTTCATCAT TTGCATCAGA GATAGAGCAT GCCAAGTAGC AACCTCAGTG ACACAGTAGT 5460

CTTACCACCA CATTTTTATG GATTAAATGT ATTTTTTTTA GCATGGTTAT ATGTGCATAT 5520

AATACACTCT GATTACTCAC TTCCCTATCC TTTCTTACTC CTCCCCATCC CAACCTGTAT 5580

CAATCCTTAC CTTCCCTACA ATTCCCTTTA CCATGTTTTT GTTAGTTTTG TTGGTTTGTT 5640

TTGTGACCCA CTGAGCTAAC CAGGGCCATC TGTATGACCA TGGGTTTGGA TTCTGATGGA 5700

ATCCCACTGG GTACACAACT GAAACTAGTG ACTCCCCTTC ACAGAATCTA TCAGTAGACA 5760

ATAATTCAAC AGGGAATGGT GGGGCTCTCT CCATCCTTGG CTAACTGTTG ACAGGACAGT 5820

CTTGTGCAGG CCTAGTGCAG ACAACCATAG TTGCTGTGAG CTCATGTTTG CAATGGCTGT 5880

GTTATACATA GGAGATAGTA TTTTGGAGCC ATTATCCATG TCTGGCTCTT ATATTCCACC 5940

TTCTCTTTTA GGATGTTCCT TGAGTCTTTG AGGAATGTTT TGGTTAGAAC CGAGTGCTCA 6000

GTTGTCATTT ATTTTCAGAA TCTTGAGCAT CAAAGGATAC ATAAGATATT ATATTATAGG 6060

ATACTAAATT TTTGTACAGA TTTTTCATAT ACCCTTCATA TTGGTTAACC ATAATCCCCA 6120

ATTTTTCTCT CCTCTAACAC TCCACTGCTC CCATACCAGA TGAAACCTTT CAACTCCATG 6180

TATTTTCCCT CTTTGCTTTC ATTTTATCTA TATTGTATGA TCTCAACTCC CTTAATCTAT 6240

CTCACTACCA ATAACCCTTT TCTAAACTGG TAGCCTACAA CTTTAGTTCC AGTACTTGAT 6300

GCAGAAGTAG ATGGAGCAAT GTGAACTCAT GCTCAGCCTG GTCTATGGAA TGGGTTACAA 6360

GCCAGCCCGG ACTATGTAAT AGGACCCTGT CTCAAAAACA ACTAAACCAA ACAAACAAAC 6420

AAACAAAGAA CAAACAAACA AACAAACCAA AAATCTCAAC CATTTCTAGT TTTTCTAGTT 6480

TTTACTTGAA CATCAAGTTA AGCATAACTA AAGTTTCAAA AATAGGATCC 6530

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5092 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GGATCCTGAT ATAGCTGTCT CTTTTGAGAC TATCCCGGGG CCTAGCAAAC ACAGAAGTGG 60

ATGCTCACAG TCAGGGATTG GGTGAATCAC AGGGCCCCCA ATGTTGGAGC TAGAGAAAGA 120

ACCCAGGGAG CTGGGGGGAT CACCTGAGTT CATACTGTCC AAACTGAAAC AAGTGGCACA 180

AGTTTCTGAG AGCCAAAGTC TAATCAGGAT CGTTTAGATC ATTAATGCTC CCCCATAATT 240

AAGACAATTT CTGATTAGAA TTATTCTTTC AACACAGCTG GGTGGAACAA GGTTCAACAG 300

TGGTATCTTA ATAGCAACTG AGTTCCAATG ATGAAAGAAA GGAAAAACAC TATGTTCTTC 360

ATACACAGAG GGGGGCTGCT CTTGGCCCTA GGGTCATCAG AGAACTGAGT AAATCTTATA 420

GGAAAATAGT TAAGATGTCT TCACACACCT CCTTTCCAAT AGGGTTCAAG GGCAGGCATG 480

ATTGGAAGGA AAAGTGTTCT GTCATGTGAG AAAAGAGCAA AAGTATTAAT ATCACATACT 540

ATGTAGTACA TTCATATTTC ATAACTTCCA TTTTCATGTT TCTGTGAAAT AAATTATAGG 600

ATTCCTGCTT GGTAGACCAA ATGGGGATCA GACAGCTCAA CAATGAACAA GTACTCAGTA 660

ACTGCCCTGT TGGTGGCATT GCATGAACTA CTGTGCTTTG CCCATGGTGA CATAGCTTGA 720

AATAGTAATG GAAGACCTGA ACCCAACTGA GATCTCTAAG TACATTCCAC TCTATGGTGG 780

CATCTCAGAG GTCAGAGTCA CTGTGCAGCG CCATAGGACA TCAGAATCAA AGGGTCATGG 840

TGAAAAGGCT GCCAGGGTCT GTCTTGTTAG TTCTCACCTT TGTAAGTAAA GTCAGTAGTC 900

AGTAACAAAG ATCAAAACAC CTGCTCTCAC AAGGAATAAC TTAAAGTAGA CTAAAGTCAT 960

GCTAGTTACA GTGCTGTCTT TTCCGTGGTA CCATCCCAAA CTGGGAGCTG GGGACTCACG 1020

AACTCTCACA ACCAATAAAG TAAGCAGAAC AGAAGCAACC CAATGAAGTG TTCATGAAAC 1080

TGGAATGGAG AAATTGTGGC ATAAGAGATG GATTCTAAAA TTTTGAGAAT TTCCAAGATA 1140

ATGAAATTAA AACCAAACAT CAAAATTGGA AAGATACAAC TGAACTAGCT TCTATGTCTT 1200

AGACAATGTC TTAGATCTCT AGATTCCGTA AGGCTGCTTC ACAAGTCTGC AACCTAGTCC 1260

TCTAGAATAG CCCTCTGGTT ATGGCACGCA ACCTATACAG AAGTTTTGAA AACAATTTCT 1320

GCCATCCACA CTGCTGGCCA TCTCTAATGA CCAACCTGCT CACTGTTACA TCAGAGAAGT 1380

GGCCAGTCAT ACACCAAACT GCCTATCCCT ATCCCAAGAA TTTGAAATCT TCATGAATGG 1440

GTCAATCCTT CCCCTGCAAT CACAGGGAGG AGGTGCCTGA TCAATAGATG AGTCAGAGCA 1500

GGACAAGAGT ATAAAACACA GGAGCACCAG TGTCCCTCAC ATCAGCATCA CCTCCTTCCC 1560

TCACTCATCT TCCCTGGTGC TTCAGGTAAG TGTGGGCTCT CCTGGCTGTC TGGTCTCTCC 1620

AGTTGGCCTT GCTCAGCTTG CAGAGAGGTT AAGGAACAGA GCCTTTCTCC CCTTTGGAAG 1680

GTACTCTGTT CAAATTGAGA AGGGCTTTAG GAAAGCACTG GGAGAGTGGT AAGCTGGTGC 1740

TGGGCAGATG ATGTGTCTGG TCTTCTGGGC AGAATGTTAA AACTTCACAA AGATATGACT 1800

ATCTCCTACT TCTCTGGCAC CCTGGGAGCT GAGGGTTAGA ATACTGGATG ACTGCAGTGG 1860

CAGGCCTCCA TGGGCTGGAT GAACCTTTTG AACCTGCCAG AAGTGGCTGA ATACACTATC 1920

AGGAAGGGAG AGGGACGATA AGTCATAGAA TGGTGCTGAT GGGAGATTTG AGAAGCCACA 1980

AAAACCCAAG CTCTGCTTTA TGAGGGCAGA TGTTCTGACA GATAAATGAC TTGTGAGGTG 2040

CTGAACTACA CAGCTTCCTA TTAGCTACAG CTAATTGGAG TCTACCAAAT TTAGACTCCT 2100

GCATATCTCA AAAAGATGTC TACTTTCTTC TGGTTAGATG TACTGGTCCA AAAGGTTCAG 2160

AGTTCTTCCA TTTGTTTGCA GACAGGACCA CAGTAGAGCT GTCTTGTCTA ATAATTGGCC 2220

CTTGGAGGAT ATCTCACTCA ATAGGACAGA TCAAGAGTTT AAACTAAGGA CTTTATACAG 2280

GAAATGCTAA TGTCCAAACA AATCTTTTCT TATTGTGCTG GGAGTGGATA AAATCCACGT 2340

GGAATTTTTG CAACTTTCTA CTGAATTTAA AGAATCAGCA CTGGGACTTG GGAGCACCCT 2400

TAGACATGGA GTGTTTATTA ATGTAAGATC AAAAGCAGGT GGGAATGTGG GGGTTCTGCT 2460

TCCCAAATCA CATAGTAGAA GAAAGGCAGA GTTGAGGGAA AAGGGGGTCA CTATTAACGG 2520

GACTTTTGAA GAGCTAACCA GTCCAGGAAT GGAGTCCAGA CACCTAGTCT GCATAAAGCT 2580

AGGAGTCAGA AGTATGTTGG CATGGATGCA TCTGCCACCT TCACAGCGTC CTCTTGCTGC 2640

TGTTGGTCTA ATGTTGCTCT TCTGCTCTTC TTCCAGGGTT CCCCTTCTCC TTAAACAACA 2700

TCGATAAGGT CACCGGGTTG CAACGGAGAC AACAGAGCTG GAAGAGTTCT CCGTGGGCGC 2760

CGATGGGCTT AACTTTCTCA TGAATTTGCC TGAGGTTTCC AAACCCTTCA CATTTTAAGC 2820

GCCCCTTCCC CCAGAAGAAG CCATTGAGTC GCTCAAGGTG TATCCTGTTC TGCAGATTTT 2880

TCATCTTGGT TTCTGAATGA CTACCTCCCA ATTCTAGTGT CTCCTCAGTC AATAAATTTG 2940

CTATTCATGA GAATCTCTGA GTTTGCTGTA GTCTTTGTAG CTTGCAAATT TACTCAGTTC 3000

ATTCTGTGTT TGCTTTTTCC ATTCATTAGT TCACATTTAA ATTCACTGAA CAAGTGTTCT 3060

ATCCCAAGGT GGGGGAGTAG ATAGATGGAA TGGGGCAAAG GATGACCAAG GTTGTGAACA 3120

GTCTGGGGTG TGGCTTAAAA ATCATGAGAT GGTCCTCAAA CACCAAGAAA AGTCTTCACT 3180

GGACATCCTA CACATCACTG AAATTGGGCC TGCGCAGGCA ATTTCTAGCA GTGCAGAGTT 3240

CACTCTCCAA GTTCTGGAAG CAGGATGGCT CTCAGATTAG GTTAGCTACC AGAGGTCCAA 3300

GTCCACTGAC ATGTTCTGAC CTAAGAAGAA GGACATTCAC CCCTGAACAA AAGACCCCTG 3360

CCCATGCGAT CTTCCGGAAC ACTATAACTA CTTTCCTTAC TCATGACCCA TGATAGAGCT 3420

TTGAGGCAAA GATACAAACC CTCTATGTCT TCTCAAGATT GCCAGTTCTT CATTAAGCCT 3480

GATACCTTCT TACCAGCGCA CGTCTCCTGA ATACTGATAA AGTCTGGTTT TGTTAGTCTG 3540

TTAGAAAAAT ATTATATCAG ATAATCAAGA TCCTCTACAG TGTGTGAGAC AGTTTACTGA 3600

GCATCTATAG AGATAGAAGG CAGCCCTCTT GAAGGATTGA ACGCGTACGT TTCGTCCAAT 3660

TTGAGAAGGT ACATCGTAAG TATTTAAGAT GCTTAACATC AGTATCACAG AGGTCACTGG 3720

AAACATTAGG GGCCTCCTGA TTAGCAAGCA TAAAGCTAGA GTTGCTCAAA GGCATGTGTA 3780

ACAACCATCC CCTGGCCAGA TCCTGTTTTA CAGTCAGATT TTATCAGCTT TAGGTAAATG 3840

CTAACTTACT GACTTACTCA AGTTAATTTT GCTATACTAA AAAGCCAATG TGCCTTCCTA 3900

CATTTAGCTA ATGATAGAAA TAAAAAGATT TCATCTCACT CTTCCATTTG GAGTCATCAC 3960

TACCTTCATC ATTTGCATCA GAGATAGAGC ATGCCAAGTA GCAACCTCAG TGACACAGTA 4020

GTCTTACCAC CACATTTTTA TGGATTAAAT GTATTTTTTT TAGCATGGTT ATATGTGCAT 4080

ATAATACACT CTGATTACTC ACTTCCCTAT CCTTTCTTAC TCCTCCCCAT CCCAACCTGT 4140

ATCAATCCTT ACCTTCCCTA CAATTCCCTT TACCATGTTT TTGTTAGTTT TGTTGGTTTG 4200

TTTTGTGACC CACTGAGCTA ACCAGGGCCA TCTGTATGAC CATGGGTTTG GATTCTGATG 4260

GAATCCCACT GGGTACACAA CTGAAACTAG TGACTCCCCT TCACAGAATC TATCAGTAGA 4320

CAATAATTCA ACAGGGAATG GTGGGGCTCT CTCCATCCTT GGCTAACTGT TGACAGGACA 4380

GTCTTGTGCA GGCCTAGTGC AGACAACCAT AGTTGCTGTG AGCTCATGTT TGCAATGGCT 4440

GTGTTATACA TAGGAGATAG TATTTTGGAG CCATTATCCA TGTCTGGCTC TTATATTCCA 4500

CCTTCTCTTT TAGGATGTTC CTTGAGTCTT TGAGGAATGT TTTGGTTAGA ACCGAGTGCT 4560

CAGTTGTCAT TTATTTTCAG AATCTTGAGC ATCAAAGGAT ACATAAGATA TTATATTATA 4620

GGATACTAAA TTTTTGTACA GATTTTTCAT ATACCCTTCA TATTGGTTAA CCATAATCCC 4680

CAATTTTTCT CTCCTCTAAC ACTCCACTGC TCCCATACCA GATGAAACCT TTCAACTCCA 4740

TGTATTTTCC CTCTTTGCTT TCATTTTATC TATATTGTAT GATCTCAACT CCCTTAATCT 4800

ATCTCACTAC CAATAACCCT TTTCTAAACT GGTAGCCTAC AACTTTAGTT CCAGTACTTG 4860

ATGCAGAAGT AGATGGAGCA ATGTGAACTC ATGCTCAGCC TGGTCTATGG AATGGGTTAC 4920

AAGCCAGCCC GGACTATGTA ATAGGACCCT GTCTCAAAAA CAACTAAACC AAACAAACAA 4980

ACAAACAAAG AACAAACAAA CAAACAAACC AAAAATCTCA ACCATTTCTA GTTTTTCTAG 5040

TTTTTACTTG AACATCAAGT TAAGCATAAC TAAAGTTTCA AAAATAGGAT CC 5092

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5159 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGGAAAACCT GTGTGGTGAG GGGGCACACA GGGAGTGTCT ACATGGGGCA AGAAGGAAAG 60

GGACAATTAT CACAGATCAG CTCCTTGTCT CTTTTGTTTG AGAAGATGAC TAACTCATGA 120

CTTAAGAGAA TTTACGTCCT GGCTCATTGT GTTCAGATCA AGTCAAGGCT GGAAGGCAGG 180

AGAATTTGCT CCGTGACTAA AGGAATCCAA AAGCAATCTT CATGTATCAT ACCTTTCTAG 240

AACTTGGGGG TGATCTCATT ATTTGTAAAG CCCTGCCCTA CCCACTCTGC AAGCTCACCA 300

TCAGGACCCA ACCCAGCCCA TCTGTACCAT ATATAAGCGG CTGCCCAGAG CTCAACACAC 360

TCATCTCTTC AGCTCTGCCC TGCCGTTTCT CTACTTCCCA GCCTTCTCAT CTCCAGGAAC 420

CATGTCTACC AAAACCACCA TCAAAAGTCA AACCAGCCAC CGTGGCTACA GTGCCAGCTC 480

AGCCAGAGTG CCTGGGCTCA ACCGCTCTGG CTTCAGCAGT GTGTCCGTGT GCCGCTCCCG 540

GGGCAGCGGT GGCTCCAGTG CAATGTGTGG AGGAGCTGGC TTTGGCAGCA GGAGCCTCTA 600

TGGTGTGGGG AGCTCCAAGA GGATCTCCAT CGGAGGGGGC AGCTGTGGCA TTGGAGGAGG 660

CTATGGCAGC CGATTTGGAG GAAGCTTCGG CATTGGAGGT GGAGCTGGTA GTGGCTTTGG 720

CTTCGGTGGT GGAGCTGGCT TTGGTGGTGG CTATGGGGGA GCTGGCTTCC CGGTGTGCCC 780

ACCTGGAGGC ATCCAAGAGG TCACCATCAA CCAGAGCCTC CTCACACCCC TGAACCTGCA 840

AATTGACCCC ACCATCCAGC GGGTCAGGAC TGAGGAGAGG GAGCAGATCA AGACCCTCAA 900

TAACAAGTTT GCCTCCTTCA TCGACAAGGT GAGACATGGT CCTCCCTAGA GCACCCTGTG 960

TGTCTACAGG GAATGCTGAA CAGAGGTGTA GGGAAGAGGC TTCAGTCTCA GCTCTGATAC 1020

TGCCTGTGTT GCTAGTTGAT GCTCTGTCCT GGTTTGTGTT CCTCTTCAGT TAGACTGGCA 1080

TCTGGAAATC AGGGTCAGCG TTCCTCTCCT CCAGAGGTTG CCCTATAAGG GTGTCTGGTC 1140

CCAGTGGACT GAGATGACTT AAAGACTCAC AAAACAGGCT TGTAGGGAAA TGGAAGATTA 1200

TAACTATGTA TAGTGCAGTT GGGAGGCATG CCAGCCTCAC TAAGCTGCAG CACACTTCAT 1260

CAAGCCATGG CTAACCTGCC AGTGCCCTAC ATGAGTTCTC TGCCCTCCTT AGAGAGGTGG 1320

CATTGGGTGC TTCAGTCTGG ACTGTTTCCC TCAGACCCAG GGTCAGGGTC TAACTACACT 1380

GAATGAGTTT AGTCAGACAG CCTGAGAGGG TACACACACT AGTGAAGTGT TCATAGAAGG 1440

ATGAAACCCA AACTTCTCCC CCTCATACTT GCCCCCCCGC CCCCACCAGG TGCGGTTCCT 1500

GGAGCAGCAG AACAAGGTCC TGGACACCAA GTGGGCCCTG CTGCAAGAGC AGGGCACCAA 1560

GACCGTGAGG CAGAACCTGG AGCCTATGTT TGAGCAGTAC ATCAGCAACC TCCGCAGACA 1620

GCTGGACAGC ATCATTGGAG AGAGGGGTCG CCTGGACTCA GAGCTGAGGA ACATGCAGGA 1680

CACAGTGGAG GACTACAAGA GCAAGTGAGT TACAAAGAAG GGAGAATCCA GTCTCCGGAC 1740

TTTATAAAAA TGGAAGCCCA AATCTAAACA AGGGCTCCAT GATGTAAGAA AGCTTGGTCA 1800

CATCTGGGAC AGAGGCTGCC ATTGATACCA TCCACCCCGT GGCTCCAATA TAGTGCACCT 1860

TTCCTCTTGT AGATATGAAG ATGAAATCAA CAAGCGCACA GCAGCAGAGA ATGAATTCGT 1920

GACCCTGAAG AAGGTGAGTT GACTAACCAC AAGGATGGGT TTCTCTGCGG AATGACATAA 1980

AAGGCCTTGT ATATCTGCGT CATTCCAGAG AAATGGTGGT TACAGGGAAA GAAGTGAACG 2040

GTCTGGGGAA GAGAGGTAAC CTGATTCCAT GTTCTTGATG GTTTTCTCAG GATGTAGATG 2100

CTGCCTACAT GAACAAAGTT GAACTGCAAG CCAAGGCAGA CAGTCTAACA GATGATATCA 2160

ACTTCTTGAG AGCTCTCTAT GAAGCAGTAA GCCCCCCTTG TCTTCTCTTC TCCTTTCCAT 2220

TCACCACTCC CTTTATTTTT TTCCCCCTGG GCAAAGTGTT TGACCTCTGC AGTTCTCAAA 2280

GACAAAGATG ACTATGGCTC TTTCTGTCCT GCAGGAACTG TCTCAGATGC AAACTCACAT 2340

CTCAGACACA TCTGTGGTCC TCTCCATGGA CAACAACCGT AGCCTGGACC TGGACAGCAT 2400

CATCGCTGAG GTCAAGGCCC AGTATGAGGA CATTGCTCAG AGAAGTCGGG CTGAAGCTGA 2460

GTCCTGGTAC CAGACTAAAG TGAGTATTGG GGTGGAGGCT GATGGGGATG CCTGGGGTCC 2520

ACCCTGAACT CCATGAGTCT CTGAGTTCAG TATTGGAGGC CCACTAAAAG AAATAGGGAT 2580

GTTGTCCCAG AAAATGCACT GTGCACATGT ACCATAGAAT AATGTTTTAC TCGAAGAGTA 2640

AAAGAACACA GAGGTAGATG CAAAGTTGCC ATAAATGGGG TCCATGCTCT TTGCTTGAGC 2700

TGTACTCTGA ACAATGATCC TCTTGAGAAA CTAGAGAACA TTTTCACTTC CTGAGGGAAC 2760

TATGGAGTCT GTGGTCTCCT AAAGCTTCTC TTGAGGAAAA GCCAGCACAT CCATGGAAGT 2820

GTGTGCCACT CAGAGGTGGG TTTCGTTCCG CATGTAACAA CTCACATAGA TGTCCTCTCT 2880

TTGATTGGCC TTCAGTATGA GGAGCTGCAG GTCACAGCTG GCAGACATGG GGACGACCTG 2940

CGCAACACCA AGCAGGAGAT TGCTGAGATC AACCGCATGA TCCAGAGGCT GAGATCTGAG 3000

ATCGACCACG TTAAGAAGCA GGTGGGGTAG ACAGAGAAAT GCATGGGTTG CGGGTTGTGT 3060

TTCCTGTCCT CTAACTCTTG CTCACCAGAA ACCATGGTCT GGGGCTCAGC CTCTGCAGAG 3120

ATGTACACTC CACGATTATT TTTGTTGCTC TCTCTGCCCA GTGTGCCAAC CTGCAAGCTG 3180

CTATTGCTGA TGCTGAGCAA CGTGGGGAGA TGGCCCTGAA GGATGCCAGG GGCAAGCTGG 3240

AAGGGCTGGA GGATGCCCTG CAGAAGGCCA AACAGGACAT GGCCAGGCTG CTGAAGGAGT 3300

ACCAGGAACT CATGAATGTC AAGCTGGCCC TGGATGTGGA AATTGCCACC TACAGGAAGC 3360

TGCTGGAAGG AGAGGAGTGC AGGTGGGTAA CTATATCCTC CAACCCCTGA GGACAGCTCC 3420

TTGGTGCAAG CACTGAGCAC AAGAAGGGAG CACTGACTAT GCCCACAATA GTCCCTTTAA 3480

GAAACTCCTT GCTGTGCTGG AGAGATGGCT CATTGTTTAA GAGCACTAAC TCCTCTTCCA 3540

GAGTTACTGA GTTTAATTCC CAGCAACCAC CTGGTGATTC ACAATCATCT CTATTGAGAT 3600

CCAGTGCCCC CTTCTGGTGT GTTTGAAAAC AGCTACAGTG AACTAAAATA CATATACTAA 3660

ATAAAGAATA TTTTTAAACA AACAAACAAA ACAAAACAAA CAAACAAACA ATCAACCCAA 3720

AACAAAACTC TAGTGGATTC TCTCTGAGCC TTCACTAGAT TGAGGCTTCC CATTCAGGCT 3780

GAAGTGATGG CTGCCTAGTT CTCACCTGTT GCTTTCCTCT TGTAGGTTGA ATGGTGAAGG 3840

TGTTGGACCA GTCAACATCT GTAAGTACTC TGCTTGTCCG AATCCCCTTC TCCTTACTTT 3900

GTGGCTTAAT TATCTGGTCA CAGTGGGCTG ACCATGTCTG TGGTGTCCTT TTCCTCCTTC 3960

ACAGCTGTGG TGCAGTCCAC CGTGTCCAGC GGCTATGGCA GTGCCGGGGG TGCCAGCAGC 4020

AGCTTAGGCC TGGGTGGAGG CAGCAGCTAC TCCTATAGCA GCAGCCATGG CCTTGGAGGT 4080

GGCTTCAGTG CTGGCAGTGG CAGAGCCATC GGAGGTGGCC TCAGCTCTTC TGGTGGCCTC 4140

AGCTCTTCTA CCATCAAATA CACCACCACC TCCTCCAGCA AGAAGAGCTA CAGGCAGTGA 4200

ATTCTGTCAC CAAGAGCTTG TCTCTGGTCC CAGATGTCAT GGCTGCAGAA TCCTGTGCTC 4260

AGAGCCCCGA GTTCAGGGGC TTCTCCTCCC TGGACCCCAC CTCTGCTCCC TTCTTGGGAC 4320

TGAGGAGGCT GTGTCATTTT GCTCATATTT CTGTCCCCAT GGGTCCCCAC TGCTCATCTC 4380

TTTATAGTCA TCCTGTGAGC TTACATCACA ATTCACTCAC ATTTGGTGCT TCATGTTGTA 4440

TTTGGTTGCC AGGCTCCTGC CTCCCTACCT CTGTCTCTGA GGCTGCCTGT GACAGGGTGT 4500

TTCCGACACC TTCATTTTTG AAATCATTGT CTGGGTCCTA CTCAAGTAAT GAGCAGCTCC 4560

CTGTGAGTTT CTAATGGCCT GAGAAACCCC ATCTCTCAAC ATCATAACCC TCCCTGTCAG 4620

TAACTGTGAC TGCCCCGTCA CTGGTCCTGT GATGTAAGTT TCTGCTCATG TGATGTCTTT 4680

GCTTTCCTTG ATGCTCTTGG CTTCCTTGTA ATTTCTAAAT AAAGCAGGTT TATACATAAT 4740

AAAATTTTCC ACGTGCATTT TTTGTTGCAA TGTTTTTAAT ATAGAAATTC TGTGGCCTTG 4800

CTAGACAAGG CATCATTACA GTTCCCTCTC CCAGGTCTAT ATGTCTTCAT CTGTTAGTAT 4860

ATAGTTTAAA TTTAAGTTCA CATTTTAAAT TAATTTCAAT AACTTTTTAA ATAAAATAGA 4920

ATTCCATCAA TTCCCCCCCC TTCATTTTTC ACCTGCCCAG ATGTCTTCAC TCCAAACCCT 4980

CACCTGTTTC TCCATTTTCA AATTGAGAGT CTTTTGAGGA AGCCTATATT TCCTTCATTT 5040

TCTTATAAAT AATTTTGTAA TGTATCCATT TCCCTTTCTT TAAAGATAAT CAACAGATGT 5100

CAGTTCAGCG TTCCTTCCCA CATGAATTGC CTTCCTGTCA GCAAGAACAT GATCTGCAG 5159

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Cys	Ser	Ser	Val	Lys	Phe	Val	Ser	Thr	Thr	Tyr	Ser	Gly	Val	Thr	Arg
1				5					10						15

5

CLAIMS

What we claim is:

1. A loricrin constitutive vector for efficient expression of a nucleic acid sequence in epidermal cells, comprising:

a 5′ flanking region of the loricrin gene, said flanking region including a TATA box, a cap site and a first intron and intron/exon boundary all in appropriate sequential and positional relationship for expression of a nucleic acid cassette;

a 3′ flanking sequence of the loricrin gene; and

a linker having a unique restriction endonuclease site at the location of the start and stop codon, said linker connecting the 5′ flanking region to the 3′ flanking sequence and said linker further providing a position for inserting the nucleic acid cassette.

2. The loricrin constitutive vector of claim 1, wherein the 5′ flanking region is approximately 1.5 kb, the intron is approximately 1.1 kb and the 3′ flanking sequence is approximately 2.1 kb.

3. The loricrin constitutive vector of claim 1, wherein the unique restriction site is selected from the group consisting of Cla I, Not I, Xma I and Bgl II, Pac I, Xho I, Nhe I and Sfi I.

4. The loricrin constitutive vector of claim 1, wherein the linker is a poly-linker, said poly-linker including a plurality of restriction endonuclease sites.

5. A keratin K6 inducible vector for regulated expression of a nucleic acid sequence in epidermal cells, comprising:

a 5′ flanking region of the keratin K6 gene, said flanking region including a TATA box, a cap site, a first intron and intron/exon boundary sequence all in sequential and positional relationship for expression of a nucleic acid cassette;

a 3′ flanking sequence of the keratin K6 gene; and

a poly-linker having a plurality of restriction endonuclease sites, said poly-linker connecting the 5′ flanking region to the 3′ flanking sequence and further providing a position for insertion of the nucleic acid cassette.

6. The keratin K6 inducible vector of claim 5, wherein the 5′ flanking region is approximately 8.0 kb and the intron and intron/exon boundary is approximately 0.56 kb and the 3′ flanking sequence is approximately 1.2 kb.

7. The vector according to claims 1, 4 or 5, wherein the cassette includes a nucleic acid sequence coding for a protein or polypeptide selected from the group consisting of a hormone, a growth factor, an enzyme, a clotting factor, an apolipoprotein, a receptor, a drug and a tumor antigen.

8. The vector according to claims 4 or 5, wherein the plurality of restriction endonuclease sites are selected from the group consisting of Cla I, Not I, Xma I, Bgl II, Pac I, Xho I, Nhe I and Sfi I.

9. A method for *in vivo* transduction of epidermal cells with a loricrin constitutive vector comprising the step of contacting the vector with epidermal cells for sufficient time to transfect the epidermal cells.

10. A method for *in vivo* transduction of epidermal cells with a keratin K6 inducible vector comprising the step of contacting the vector with epidermal cells for sufficient time to transfect the epidermal cells.

11. A bioreactor comprising transformed epidermal cells including the loricrin constitutive vector of claim 1.

12. A bioreactor comprising transformed epidermal cells including the keratin K6 inducible vector of claim 5.

13. The bioreactor according to claims 11 or 12 wherein the loricrin constitutive vector includes a cassette having a nucleic acid sequence coding for a protein or polypeptide selected from the group consisting of a hormone, a growth factor, an enzyme, a drug, a tumor suppressor,

a receptor, an apolipoprotein, a clotting factor a tumor antigen, a viral antigen, a bacterial antigen and a parasitic antigen.

14. The bioreactor of claim 13, wherein the nucleic acid sequence encodes proinsulin or insulin.

15. The bioreactor of claim 13, wherein the nucleic acid sequence encodes growth hormone.

16. The bioreactor of claim 13, wherein the nucleic acid sequence encodes insulin-like growth factor I, insulin-like growth factor II or insulin growth factor binding protein.

17. The bioreactor of claim 13, wherein the nucleic acid sequence encodes antihemophilic factor (Factor VIII), Christmas factor (Factor IX) or Factor VII.

18. The bioreactor of claim 13, wherein the nucleic acid sequence encodes an epidermal growth factor (TGF-α), a dermal growth factor (PDGF) or an angiogenesis factor.

19. The bioreactor of claim 13, wherein the nucleic acid sequence encodes Type IV collagen, laminin, nidogen, or Type VII collagen.

20. The bioreactor of claim 13 for vaccine production, wherein the cassette includes a protein which induces an immunological response.

21. A method for *ex vivo* introduction of a loricrin constitutive vector into epidermal cells comprising the steps of co-transfecting the vector with a selectable marker and selecting the transformed cells.

22. A method for *ex vivo* introduction of a keratin K6 inducible vector into epidermal cells comprising the steps of co-transfecting the vector with a selectable marker and selecting the transformed cells.

23. A loricrin gene of SEQ. ID. No. 1.

24. A loricrin constitutive vector having:

a 5′ flanking region comprising nucleotides 1 to 1540 of SEQ. ID. No. 1;

an intron and intron/exon boundary comprising nucleotides 1587 to 1679 of SEQ. ID. No. 1;

a 3′ flanking region comprising nucleotides 4384 to 6530 of SEQ. ID. No. 1; and

a linker to be inserted at the unique Cla I site at nucleotides 2700 to 2705 SEQ. ID. No. 2

25. A keratin K6 gene of SEQ. ID. No. 3.

26. A keratin K6 inducible vector having:

a 5′ flanking region which extends from a unique 5′ Xho I site up to nucleotide 360 of SEQ. ID. No. 3;

an intron and intron/exon boundary comprising nucleotides 928 to 1494 of SEQ. ID. No. 3.

a 3′ flanking region which extends from nucleotide 4740 of SEQ. ID. No. 3 to a unique 3′ Xho I site; and

a poly-linker inserted between nucleotides 1504 to 1509 of SEQ. ID. No. 3

27. A method for enhanced healing of a wound or surgical incision comprising the steps of *in vivo* transduction of epidermal cells with a loricrin constitutive vector, wherein said vector includes a nucleic acid cassette having nucleic acid sequence for a growth factor.

28. A method of enhanced healing of a wound or surgical incision comprising the step of *in vivo* transduction of epidermal cells with a keratin K6 inducible vector, wherein said vector includes a nucleic acid cassette having a nucleic acid sequence for a growth factor.

29. The method according to claims 27 or 28, wherein the epidermal cells are transduced with a plurality of vectors and wherein the cassette of at least one vector includes the nucleic acid sequence of epidermal growth factor (TGF-α), the cassette of at least one vector includes dermal growth factor (PDGF), the cassette of at least one vector includes the nucleic acid sequence for a matrix protein to anchor the

epidermis to the dermis and the cassette of at least one vector includes the nucleic acid sequence for an angiogenesis factor.

30. The method of claim 29, wherein the sequence for the matrix protein is selected from sequences coding for a protein selected from the group consisting of Type IV collagen, laminin, nidogen and Type VII collagen.

31. The method of claim 29, wherein the angiogenesis factor is selected from the group consisting of acid fibroblast growth factor, basic fibroblast growth factor and angiogenin.

32. A method of treating skin ulcers comprising the steps of *in vivo* transduction of epidermal cells with a loricrin constitutive vector, wherein said vector includes a nucleic acid cassette having a nucleic acid sequence for a growth factor.

33. A method of treating skin ulcers comprising the steps of *in vivo* transduction of epidermal cells with a keratin K6 inducible vector, wherein said vectors include a nucleic acid cassette having a nucleic acid sequence for a growth factor.

34. The method according to claims 32 or 33, wherein the epidermal cells are transduced with a plurality of vectors and wherein the cassette of at least one vector includes the nucleic acid sequence of epidermal growth factor (TGF-α), the cassette of at least one vector includes dermal growth factor (PDGF), the cassette of at least one vector includes the nucleic acid sequence for a matrix protein to anchor the epidermis to the dermis and the cassette of at least one vector includes the nucleic acid sequence for an angiogenesis factor.

35. The method of claim 34, wherein the sequence for the matrix protein is selected from sequences coding for a protein selected from the group consisting of Type IV collagen, laminin, nidogen, and Type VII collagen.

36. The method of claim 34, wherein the angiogenesis factor is selected from the group consisting of acid fibroblast growth factor, basic fibroblast growth factor and angiogenin.

37. A method of enhanced healing of a wound, surgical incision or skin ulcers in humans and animals comprising the steps of:

ex vivo transduction of epidermal cells with a loricrin constitutive vector, wherein said vector includes a nucleic acid cassette a nucleic acid sequence for a growth factor; and

transplanting said transduced epidermal cells into the animal or human to be treated.

38. A method of enhanced healing of a wound, surgical incision or skin ulcers in humans and animals, comprising the steps of:

ex vivo transduction of epidermal cells with a keratin K6 inducible vector, wherein said vector includes a nucleic acid cassette having a nucleic acid sequence for a growth factor; and

transplanting said transduced epidermal cells into the animal or human to be treated.

39. The method according to claims 37 or 38, wherein the epidermal cells are transduced with a plurality of vectors and wherein the cassette of at least one vector includes the nucleic acid sequence of epidermal growth factor (TGF-α), the cassette of at least one vector includes dermal growth factor (PDGF), the cassette of at least one vector includes the nucleic acid sequence for a matrix protein to anchor the epidermis to the dermis and the cassette of at least one vector includes the nucleic acid sequence for an angiogenesis factor.

40. The method of claim 39, wherein the sequence for the matrix protein is selected from sequences coding for a protein selected from the group consisting of Type IV collagen, laminin, nidogen and Type VII collagen.

41. The method of claim 39, wherein the angiogenesis factor is selected from the group consisting of acid fibroblast growth factor, basic fibroblast growth factor and angiogenin.

42. A method for treating psoriasis comprising the step of *in vivo* transduction of epidermal cells with a loricrin constitutive vector, wherein said vector includes a nucleic acid cassette having a nucleic acid sequence coding for a protein or polypeptide selected from the group consisting of TGF-β, a soluble form of cytokine receptor, and an antisense RNA.

43. A method for treating psoriasis comprising the step of *in vivo* transduction of epidermal cells with a keratin K6 inducible vector, wherein said vector includes a nucleic acid cassette having a nucleic acid sequence coding for a protein or polypeptide selected from the group consisting of TGF-β, a soluble form of cytokine receptor, and an antisense RNA.

44. The method of claims 42 or 43 wherein the cassette contains the sequence for TGF-β.

45. The method of claims 42 or 43 wherein the cassette contains a soluble form of cytokine receptor selected from the group consisting of IL-1, IL-6 and IL-8.

46. The method of claims 42 or 43 wherein the cassette contains antisense RNA to TGF-α, IL-1, IL-6 or IL-8.

47. A method of treating cancer of squamous epithelia comprising the step of *in vivo* transduction of squamous epithelia cells with a loricrin constitutive vector or a keratin K6 vector, said vector includes a nucleic acid cassette having a nucleic acid sequence coding for an antisense RNA.

48. The method of claim 47 wherein the squamous epithelia cells are selected from the group of cells consisting of epidermis, oral, esophageal, vaginal, tracheal and corneal epithelia.

49. The method of claim 47 for treating skin cancer wherein transduction of epidermal cells is with a loricrin constitutive vector and said nucleic acid cassette has a nucleic acid sequence coding for an antisense RNA for the E6 or E7 gene of human papilloma virus.

50. The method of claim 47 for treating skin cancer wherein transduction of epidermal cells is with a loricrin constitutive vector and said nucleic acid cassette has a nucleic acid sequence coding for the normal p53 protein.

51. The method of claim 47 for treating skin cancer wherein transduction of epidermal cells is with a keratin K6 vector and said nucleic acid cassette has a nucleic acid sequence coding for an antisense RNA for the E6 or E7 gene of human papilloma virus.

52. The method of claim 47 for treating skin cancer wherein transduction of epidermal cells is with a keratin K6 vector and said nucleic acid cassette has a nucleic acid sequence coding for the normal p53 protein.

53. The vector according to claims 1, 4 or 5, further including a Vitamin D regulatory element.

54. The vector of claim 53, wherein the Vitamin D regulatory element is from the human K1 keratin gene.

55. A method for vaccination comprising the step of the *in vivo* transduction of epidermal cells with a loricrin constitutive vector or a keratin K6 inducible vector, wherein said vector includes a nucleic acid cassette having a nucleic acid sequence coding for a protein or polypeptide which induces an immunological response.

56. The method of claim 55, wherein the cassette includes a sequence for a viral capsid protein.

57. The method of claim 56, wherein the capsid protein is from the human papilloma virus.

58. A transgenic animal containing the vector of claims 1, 4 or 5 in its germ and somatic cells, wherein said vector was introduced into said

animal or an ancestor of said animal at an embryonic stage and the nucleic acid cassette of said vector is only expressed in squamous epithelia.

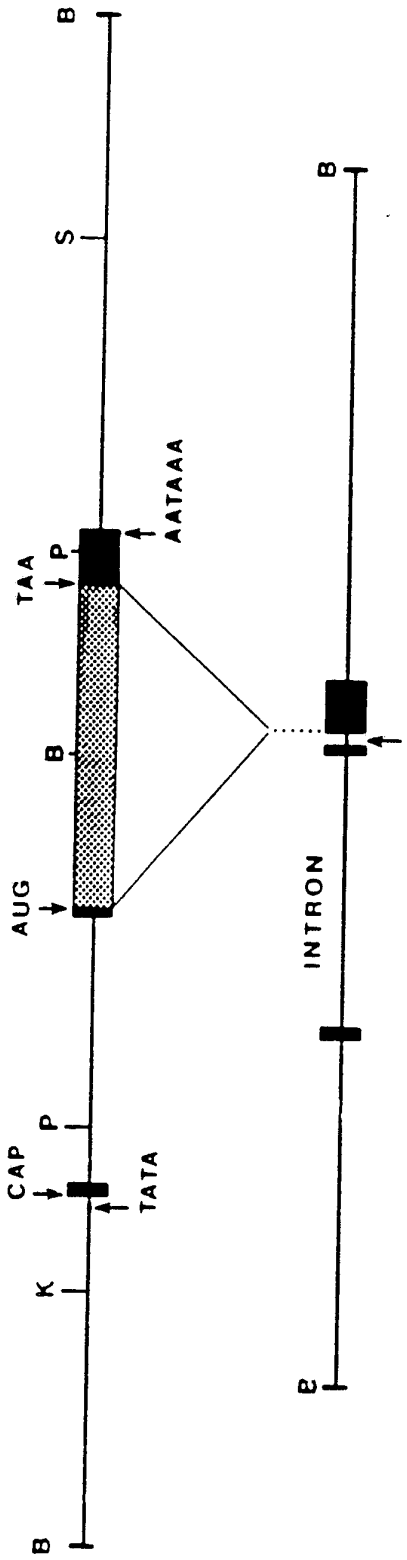


FIGURE 1

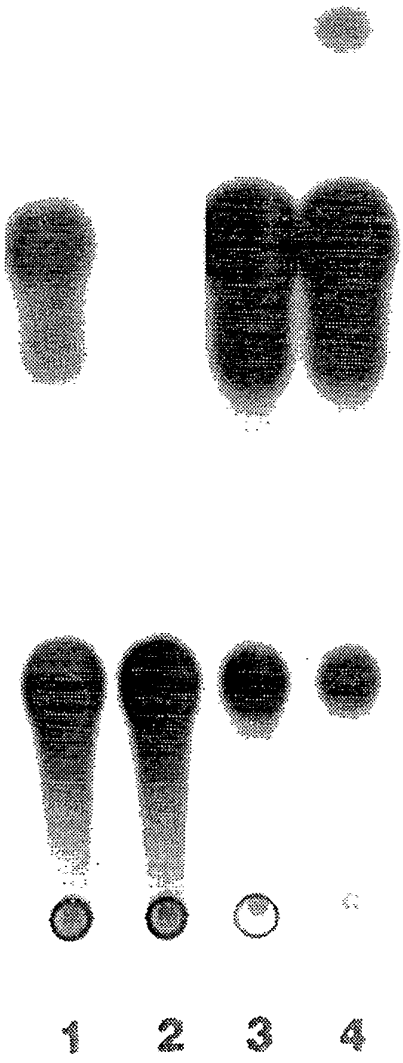


FIGURE 2

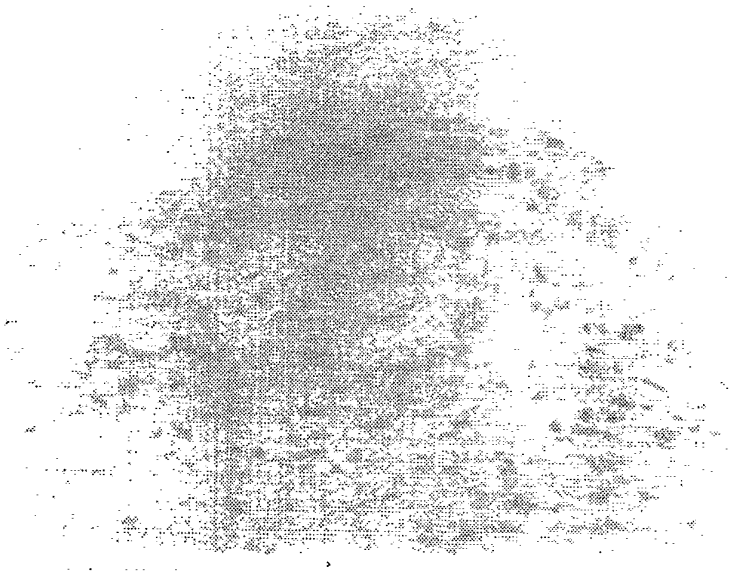
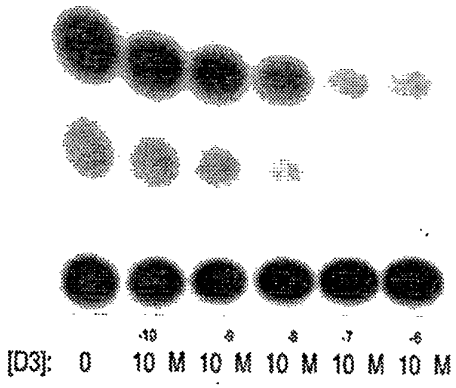


FIGURE 3a

FIGURE 3b



Sequence of HK1.NRE

TGGCCTTGAG$_{10}$GGAGATGATT$_{20}$CACTCTCCTT$_{30}$CACAGAAGAG$_{40}$CTGACCTCTG$_{50}$GGGTCAACAG$_{60}$ATATAGCACC$_{70}$

FIGURE 4

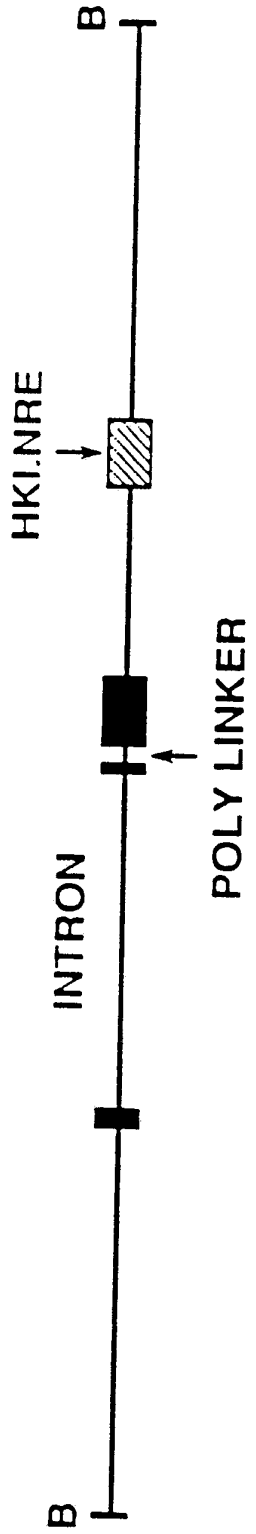


FIGURE 5

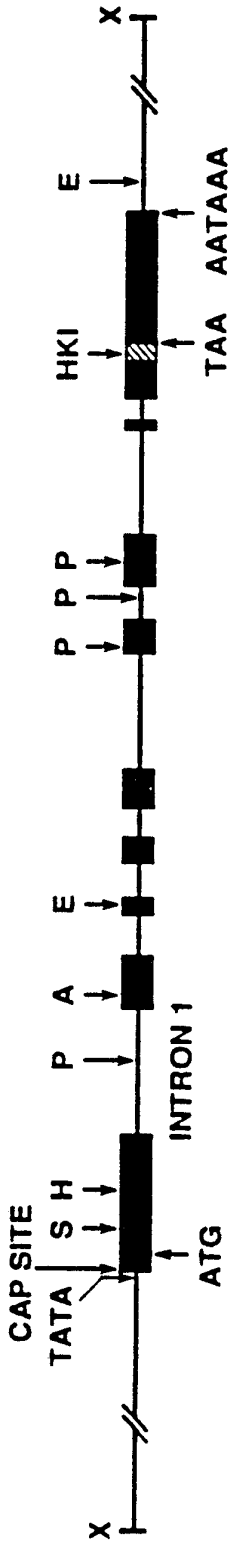


FIGURE 6

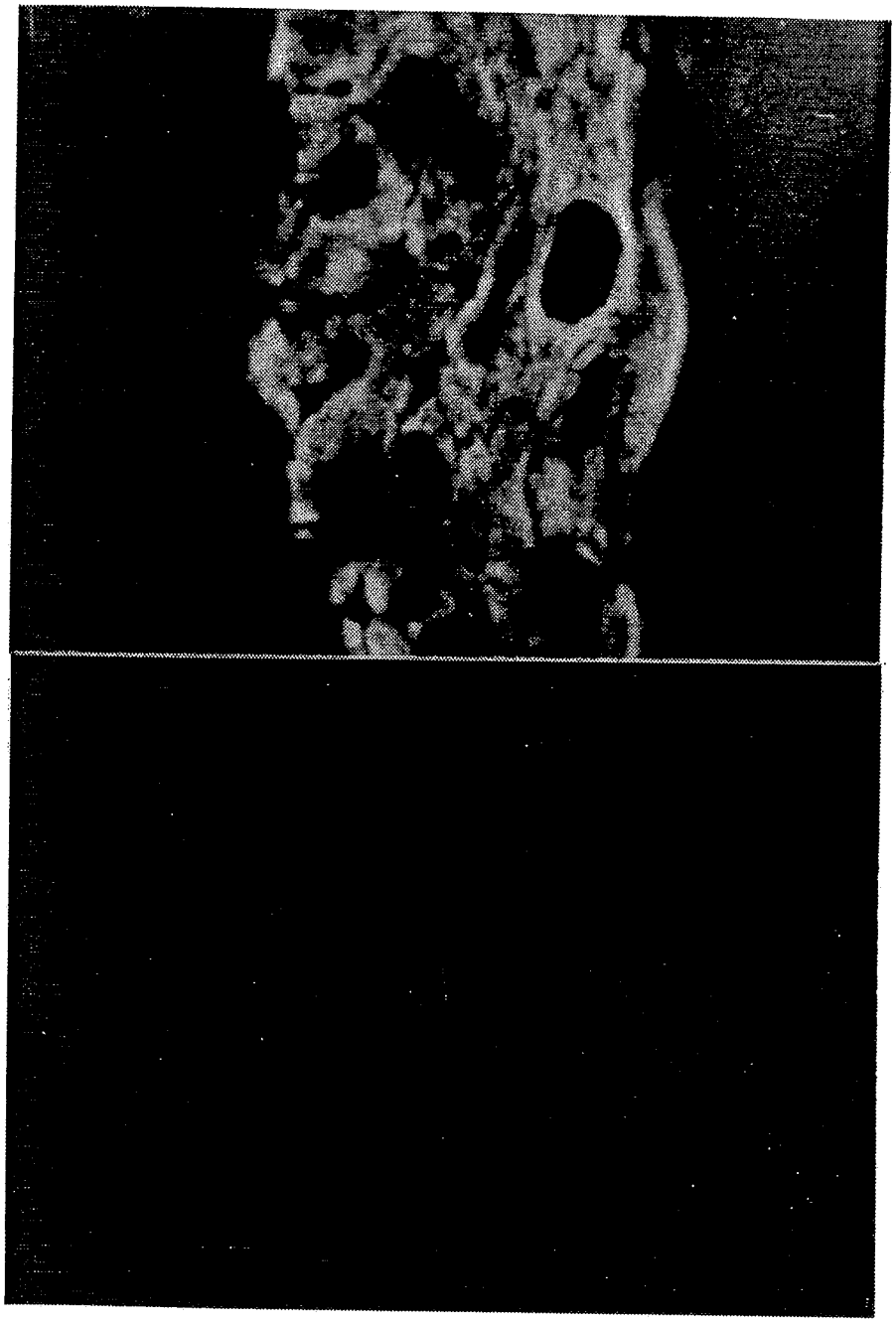


FIGURE 7

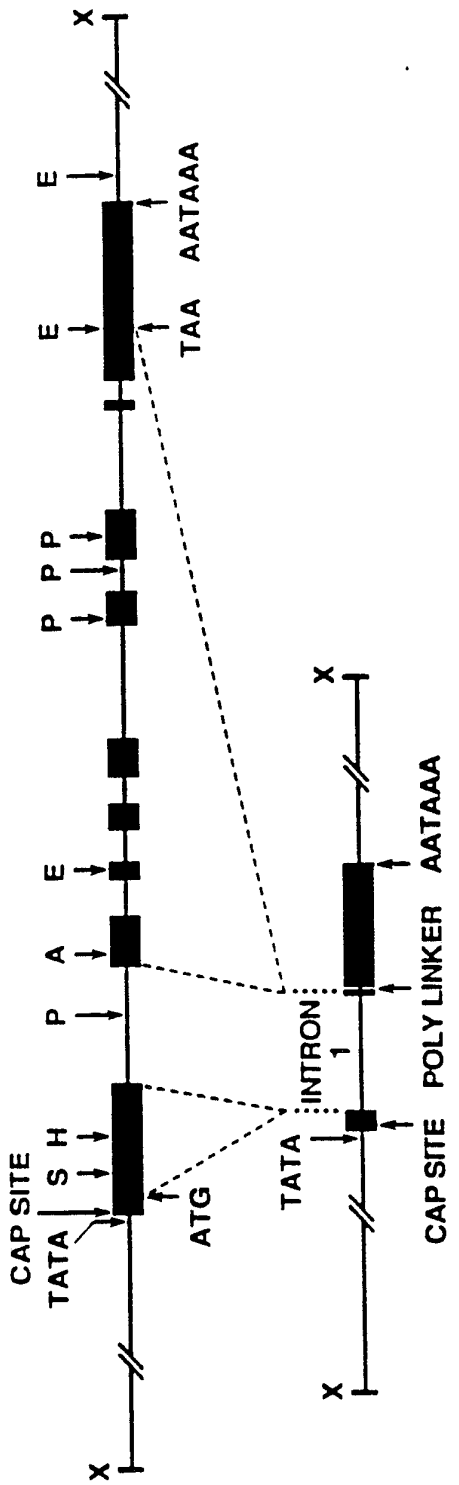

X
E
AATAAA
E
TAA
P P P
E
P A
P
S H
CAP SITE
TATA
ATG
INTRON 1
TATA
X
AATAAA
POLY LINKER
CAP SITE
X

FIGURE 8

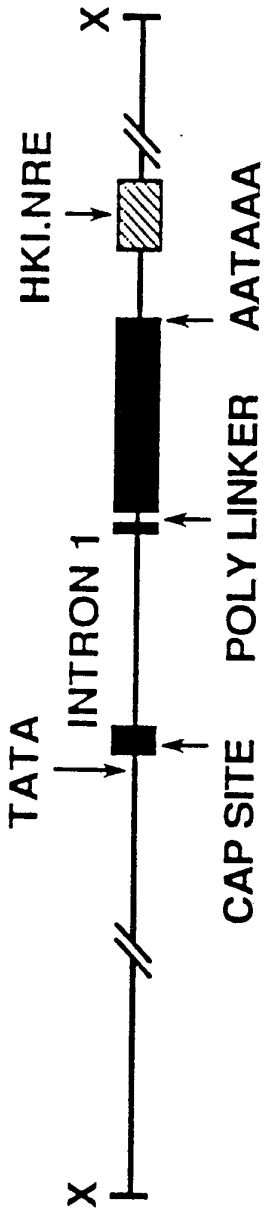


FIGURE 9

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/03993

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :Please See Extra Sheet.
US CL :435/320.1, 284, 285, 286; 424/93B; 536/23.5; 800/2
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, CAS, BIOSIS, WORLD PATENTS INDEX
search terms: loricrin, keratin K6, epidermal vector

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,087,617 (Smith) 11 February 1992, see col. 3, lines 31-38.	42-52
Y	US, A, 5,057,411 (Lancaster et al) 15 October 1991, see col. 3, lines 56-64.	55-57
Y	US, A, 5,008,240 (Bentz et al) 16 April 1991, see col. 16, lines 44-61.	42-46

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 29 JULY 1993	Date of mailing of the international search report AUG 17 1993
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer R. KEITH BAKER Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/03993

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,024,841 (Chu et al) 18 June 1991, see col. 1, lines 62-68.	27-41
Y	US, A, 4,863,899 (Todaro) 05 September 1989, see the Abstract.	27-41
Y	US, A, 4,816,464 (Gilman et al) 28 March 1989, see col. 1, lines 53-63.	42-46
Y	US, A, 5,057,428 (Mizutani et al) 15 October 1991, see the Abstract.	11-20
Y	Cell, Volume 61, issued 15 June 1990, T. Mehrel et al, "Identification of a Major Keratinocyte Cell Envelope Protein, Loricrin," pages 1103-1112, see entire document.	1-4, 7-9, 11, 13-20, 21, 23, 24, 27, 29-32, 34-37, 39-42, 44-50, 53-58
Y	Carcinogenesis, Volume 12, No. 8, issued August 1991, J. Finch et al, "Identification of a Cloned Sequence Activated During Multi-stage Carcinogenesis in Mouse Skin," pages 1519-1522, see Figure 1 and pages 1519-1520, bridging paragraph.	5-8, 10, 12, 22, 25, 26, 28-31, 33-36, 38-41, 44-48, 51-58
Y	Proceedings of the National Academy of Science USA, Volume 82, issued April 1985, LD Johnson et al, "Structure of a Gene for the Human Epidermal 67-kDa Keratin," pages 1896-1900, see p. 1898.	1-58
Y	Clinical Research, Volume 39, No. 2, issued April 1991, K Yoneda et al, "Structure of the Human Loricrin Gene: Linkage at 1q21 with Profilaggrin and Involucrin Genes," page 496A, see entire document.	1-4, 7-9, 11, 13-20, 21, 23, 24, 27, 29-32, 34-37, 39-42, 44-50, 53-58
Y	Clinical Research, Volume 38, No. 2, issued April 1990, JA Rothnagel et al, "Development of an Epidermal-specific Expression Vector for Targeting Gene Expression to the Epidermis of Transgenic Mice," page 688A, see entire document.	1-58

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/03993

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	The Journal of Investigative Dermatology, Volume 96, No. 4, issued April 1991, JA Rothnagel et al, "Use of Transgenic Mice to Identify Regulatory Sequences 5' to the Human K1 Keratin Gene," page 541, see entire document.	1-58
Y	Science, Volume 237, issued 18 September 1987, JR Morgan et al, "Expression of an Exogenous Growth Hormone Gene by Transplantable Human Epidermal Cells," pages 1476-1479, see the Abstract.	1-58
Y	The FASEB Journal, Volume 4, No. 14, issued November 1990, J Teumer et al, "Human Growth Hormone in the Blood of Athymic Mice Grafted with Cultures of Hormone-secreting Human Keratinocytes," pages 3245-3250, see the Abstract.	1-58
Y	Proceedings of the National Academy of Science USA, Volume 86, issued November 1989, ES Fenjves et al, "Systemic Distribution of Apolipoprotein E Secreted by Grafts of Epidermal Keratinocytes: Implications for Epidermal Function and Gene Therapy," pages 8803-8807, see entire document.	1-58
Y	The Journal of Investigative Dermatology, Volume 97, No. 5, issued November 1991, JA Garlick et al, "Retrovirus-mediated Transduction of Cultured Epidermal Cells," pages 824-829, see entire document.	1-58
Y	Proceedings of the National Academy of Science USA, Volume 88, issued April 1991, RS Williams et al, "Introduction of Foreign Genes into Tissues of Living Mice by DNA-coated Microprojectiles," pages 2726-2730, see the Abstract.	1-58
Y	Sambrook et al, "Molecular Cloning: A Laboratory Manual" published 1989 by Cold Spring Harbor Laboratory Press, pages 1613-1627, see entire document.	1-58

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/03993

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Cancer Research, Volume 50, issued 15 June 1990, MK Doeberitz et al, "Growth-regulating Functions of Human Papillomavirus Early Gene Products in CervicalCancer Cells Acting Dominant over Enhanced Epidermal Growth Factor Receptor Expression," pages 3730-3736, see page 3730.	47-52
Y	Cell Regulation, Volume 1, issued November 1990, M. Tomic et al, "Nuclear Receptors for Retinoic Acid and Thyroid Hormone Regulate Transcription of Keratin Genes," pages 965-973, see the Abstract.	53-54
Y	Cell, Volume 25, issued September 1981, E Fuchs et al, "Regulation of Terminal Differentiation of Cultured Human Keratinocytes by Vitamin A," pages 617-625, see the Abstract.	53-54
Y	Cell, Volume 62, issued 24 August 1990, B Bailleul et al, "Skin Hyperkeratosis and Papilloma Formation in Transgenic Mice Expressing a ras Oncogene from a Suprabasal Keratin Promoter," pages 697-708, see entire document.	58
Y	Proceedings of the National Academy of Science USA, Volume 86, issued March 1989, R Vassar et al, "Tissue-specific and Differentiation-specific Expression of a Human K14 Keratin Gene in Transgenic Mice," pages 1563-1567, see entire document.	58
Y	The EMBO Journal, Volume 4, No. 13A, issued 1985, JD Kelly et al, "The B Chain of PDGF Alone is Sufficient for Mitogenesis," pages 3399-3405, see the Abstract.	11-20
Y	Proceedings of the National Academy of Science USA, Volume 76, issued January 1979, DV Goeddel et al, "Expression in Escherichia coli of Chemically Synthesized Genes for Human Insulin," pages 106-110, see the Abstract.	11-20
Y	Biochemical and Biophysical Research Communications, Volume 138, No. 1, issued 16 July 1986, Y Itoh et al, "Expression of Hepatitis B Virus Surface aAntigen P31 in Yeast," pages 268-274, see the Abstract.	11-20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/03993

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Nature, Volume 312, issued 22 November 1984, WI Wood et al, "Expression of Active Human Factor VIII from Recombinant DNA Clones," pages 330-337, see entire document.	11-20
Y	Nature, Volume 306, issued 08 December 1983, M Jansen et al, "Sequence of cDNA Encoding Human Insulin-like Growth Factor I Precursor," pages 609-911, see Figure 2.	16
Y	American Journal of Human Genetics, issued June 1989, DR Olsen et al, "Human Nidogen: cDNA Cloning, Cellular Expression, and Mapping of the Gene to Chromosome Iq43," pages 876-885, see Figure 2.	19, 29-30, 34-35, 39-40

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/03993

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (5):

C12N 15/00, 15/12, 15/63, 15/64, 15/85; C12M 3/00, 3/02, 3/04; A61K 31/70

B. FIELDS SEARCHED
Minimum documentation searched
Classification System: U.S.

435/69.1, 172.2, 172.3, 240.2, 240.21, 284, 285, 286, 320.1; 424/93B, 93A, 88, 89, 92; 536/ 23.5, 23.51, 23.72; 514/2, 8, 12, 44; 800/2

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

I. Claims 1-4, 11, 21, 23 and 24, drawn to a loricrin vector, classified in Class 435, subclass 320.1, for example.

II. Claims 5-6, 12, 22, 25, and 26 drawn to a keratin K6 vector, classified in Class 435, subclass 320.1, for example.

III. Claim 9, drawn to in vivo transduction of cells with a loricrin vector, classified in Class 424, subclass 93B, for example.

IV. Claim 10, drawn to in vivo transduction of cells with a keratin K6 vector, classified in Class 424, subclass 93B, for example.

V. Claim 27, drawn to a method of enhancing wound healing with a loricirn vector, classified in Class 424, subclass 93B.

VI. Claim 28, drawn to a method of enhancing wound healing with a keratin K6 vector, classified in Class 424, subclass 93B.

VII. Claim 32, drawn to a method of treating skin ulcers with a loricrin vector, classified in Class 424, subclass 93B.

VIII. Claim 33, drawn to a method of treating skin ulcers with a keratin K6 vector, classified in Class 424, subclass 93B.

IX. Claim 37, drawn to methods of treament with a loricrin vector comprising ex vivo transduction and transplantation, classified in Class 424, subclass 93B.

X. Claim 38, drawn to methods of treament with a keratin K6 vector comprising ex vivo transduction and transplantation, classified in Class 424, subclass 93B.

XI. Claim 42, drawn to an in vivo method of treating psoriasis with a loricrin vector, classified in Class 424, subclass 93B.

XII. Claim 43, drawn to an in vivo method of treating psoriasis with a keratin K6 vector, classified in Class 424, subclass 93B.

XIII. Claims 47-52, drawn to in vivo methods of treating cancer, classified in Class 424, subclass 93B.

XIV. Claims 55-57, drawn to in vivo methods of vaccination, classified in Class 424, subclass 93B.

XV. Claim 58 drawn to a transgenic animal, classified in Class 800, subclass 2.

Claims 7, 8, 13-20 and 53-54 link inventions I and II.
Claims 29-31 link inventions V and VI and will be examined if the required fees are paid for either of these groups.
Claims 34-36 link inventions VII and VIII and will be examined if the required fees are paid for either of these groups.
Claims 39-41 link inventions IX and X and will be examined if the required fees are paid for either of these groups.

Claims 44-46 link inventions XI and XII and will be examined if the required fees are paid for either of these groups.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/03993

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
(Form PCT/ISA/206 Previously Mailed.)
Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.