



US 20250255797A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0255797 A1**
Krishnan et al. (43) **Pub. Date:** **Aug. 14, 2025**

(54) **RECOMBINANT NUCLEIC ACIDS
ENCODING COSMETIC PROTEIN(S) FOR
AESTHETIC APPLICATIONS**

(71) Applicant: **Krystal Biotech, Inc.**, Pittsburgh, PA (US)

(72) Inventors: **Suma Krishnan**, Pittsburgh, PA (US);
Trevor Parry, Pittsburgh, PA (US);
Pooja Agarwal, Pittsburgh, PA (US)

(21) Appl. No.: **18/896,155**

(22) Filed: **Sep. 25, 2024**

Related U.S. Application Data

- (63) Continuation of application No. 16/915,695, filed on Jun. 29, 2020, now Pat. No. 12,128,122, which is a continuation of application No. 16/395,896, filed on Apr. 26, 2019, now Pat. No. 10,786,438.
(60) Provisional application No. 62/663,476, filed on Apr. 27, 2018.

Publication Classification

- (51) **Int. Cl.**
A61K 8/65 (2006.01)
A61K 8/64 (2006.01)
A61Q 19/08 (2006.01)
(52) **U.S. Cl.**
CPC *A61K 8/65* (2013.01); *A61K 8/64* (2013.01); *A61Q 19/08* (2013.01); *A61K 2800/86* (2013.01); *A61K 2800/91* (2013.01); *C12N 2710/16621* (2013.01); *C12N 2710/16631* (2013.01)

ABSTRACT

The present disclosure provides recombinant nucleic acids comprising one or more polynucleotides encoding one or more cosmetic proteins (e.g., one or more human collagen proteins); viruses comprising the recombinant nucleic acids; compositions (e.g., cosmetic formulations) comprising the recombinant nucleic acids and/or viruses; methods of their use; and articles of manufacture or kits thereof.

Specification includes a Sequence Listing.

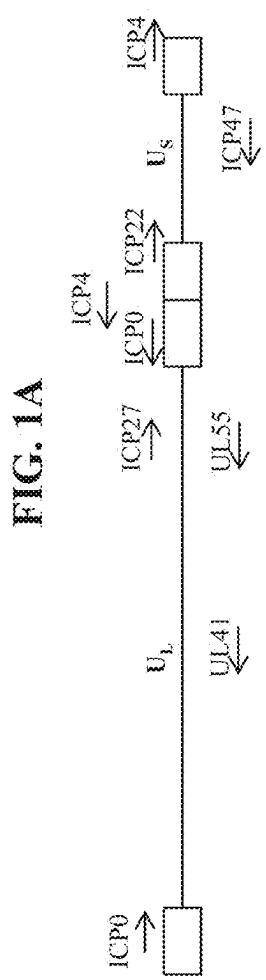


FIG. 1B

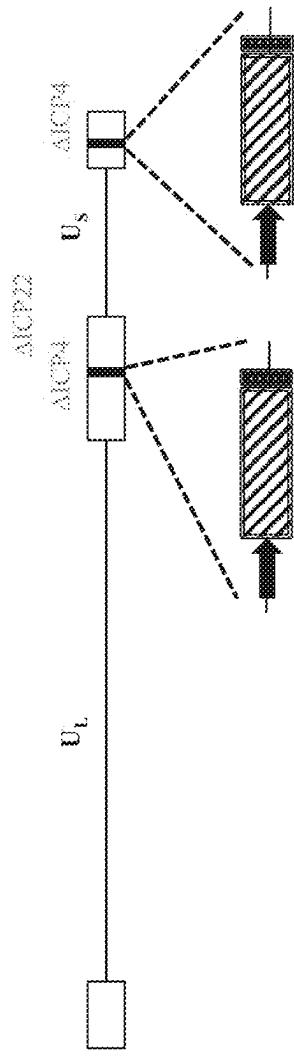
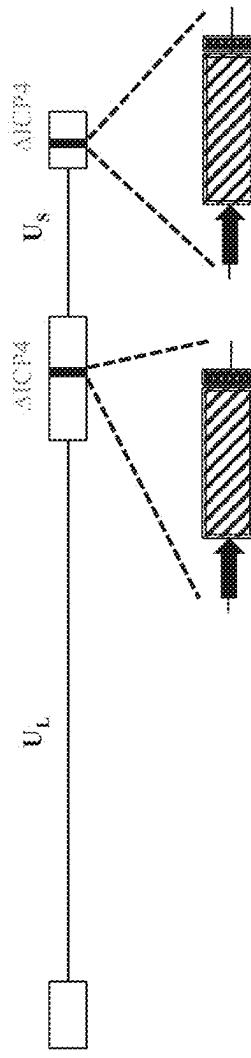


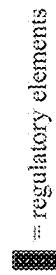
FIG. 1C



→ = heterologous promoter



→ = 1st human collagen polypeptide coding sequence



→ = regulatory elements

FIG. 1D

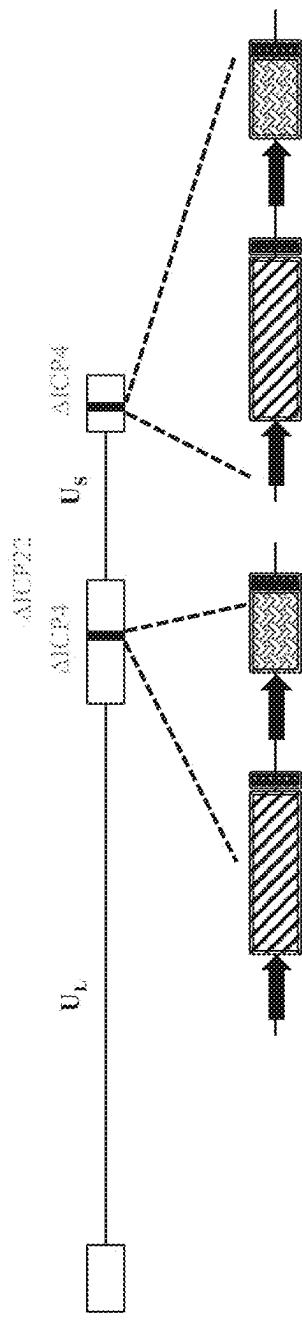


FIG. 1E

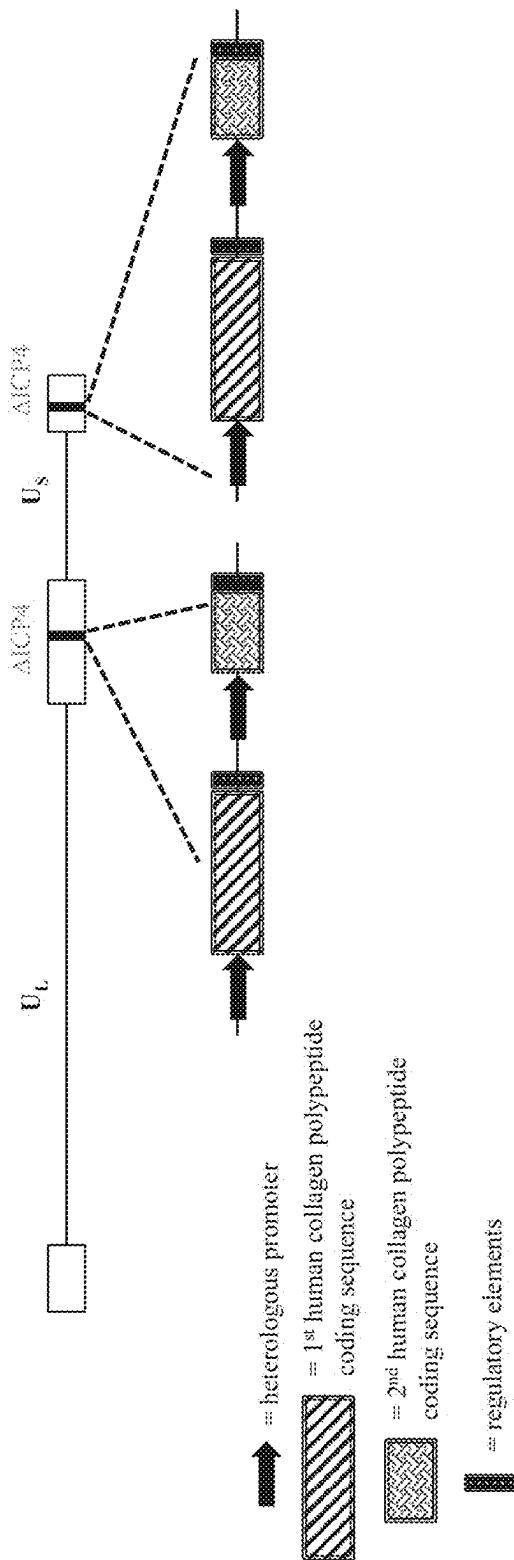


FIG. 1F

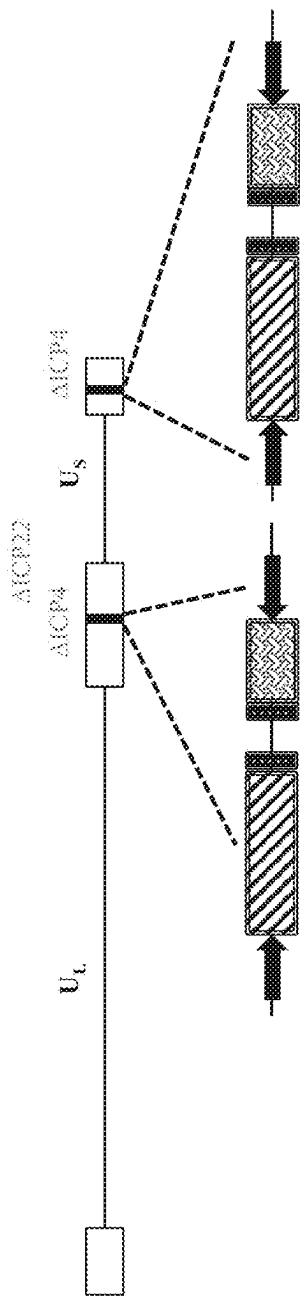


FIG. 1G

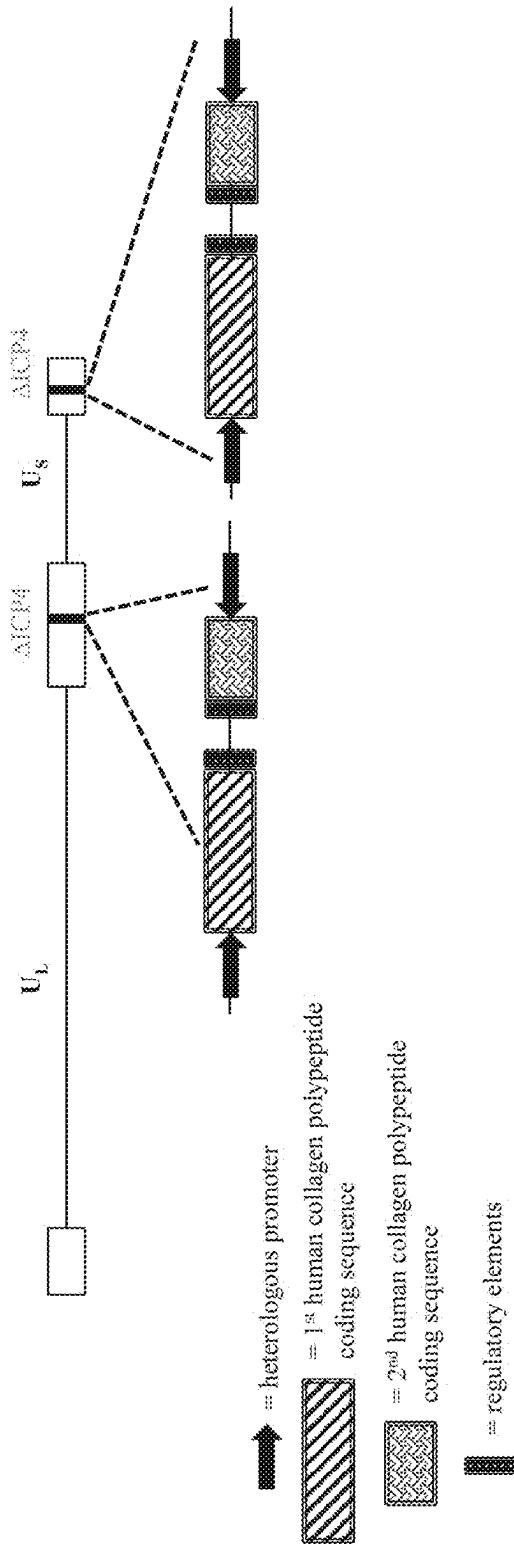


FIG. 1H

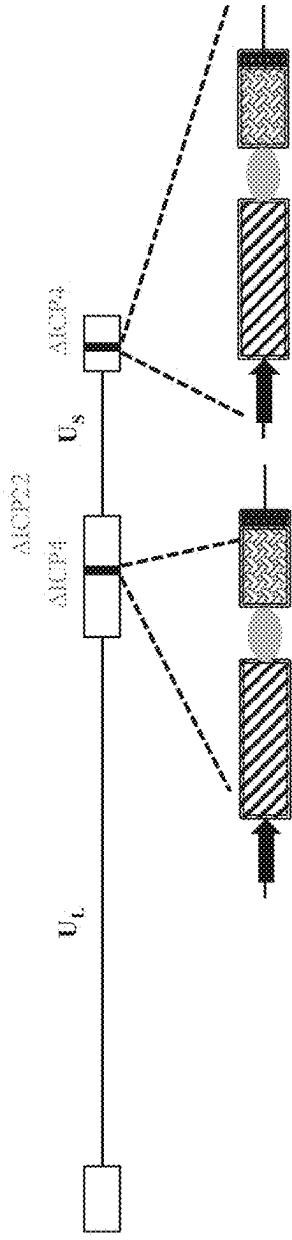


FIG. 1I

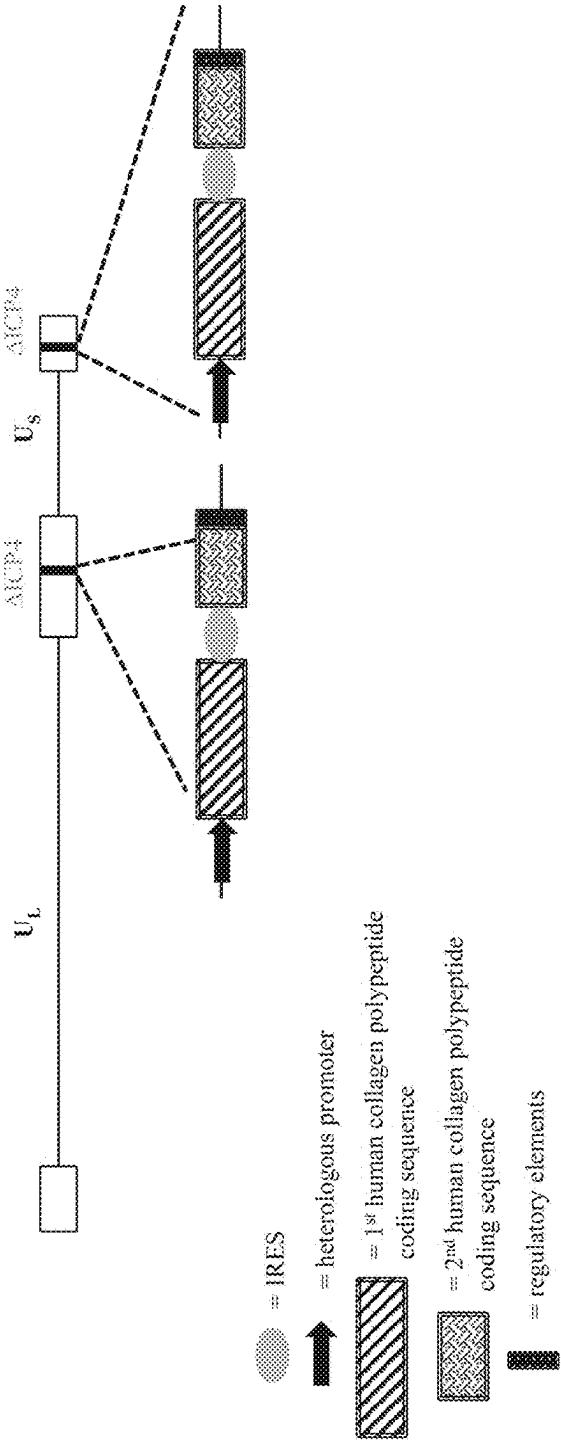


FIG. 1J

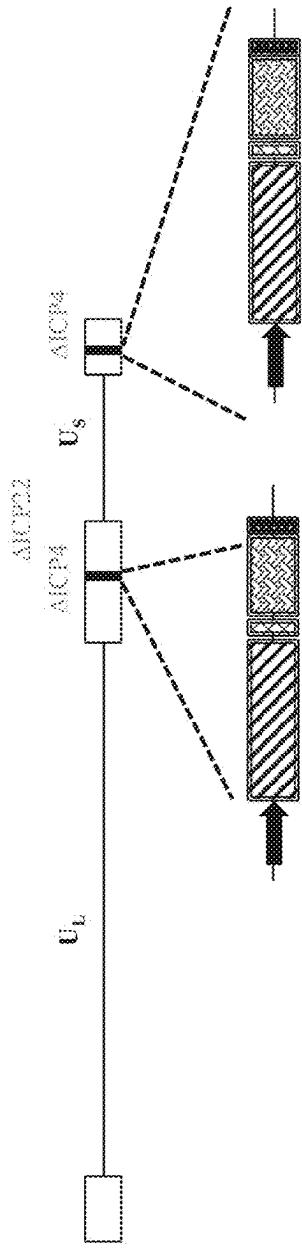


FIG. 1K

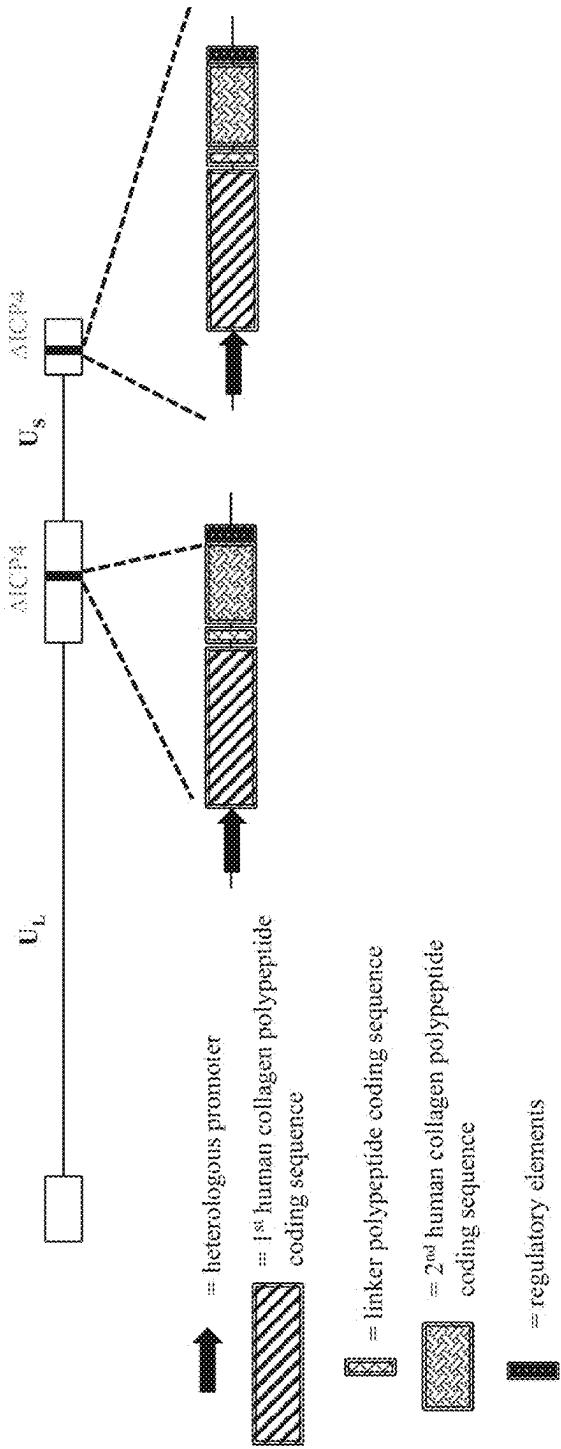


FIG. 11.

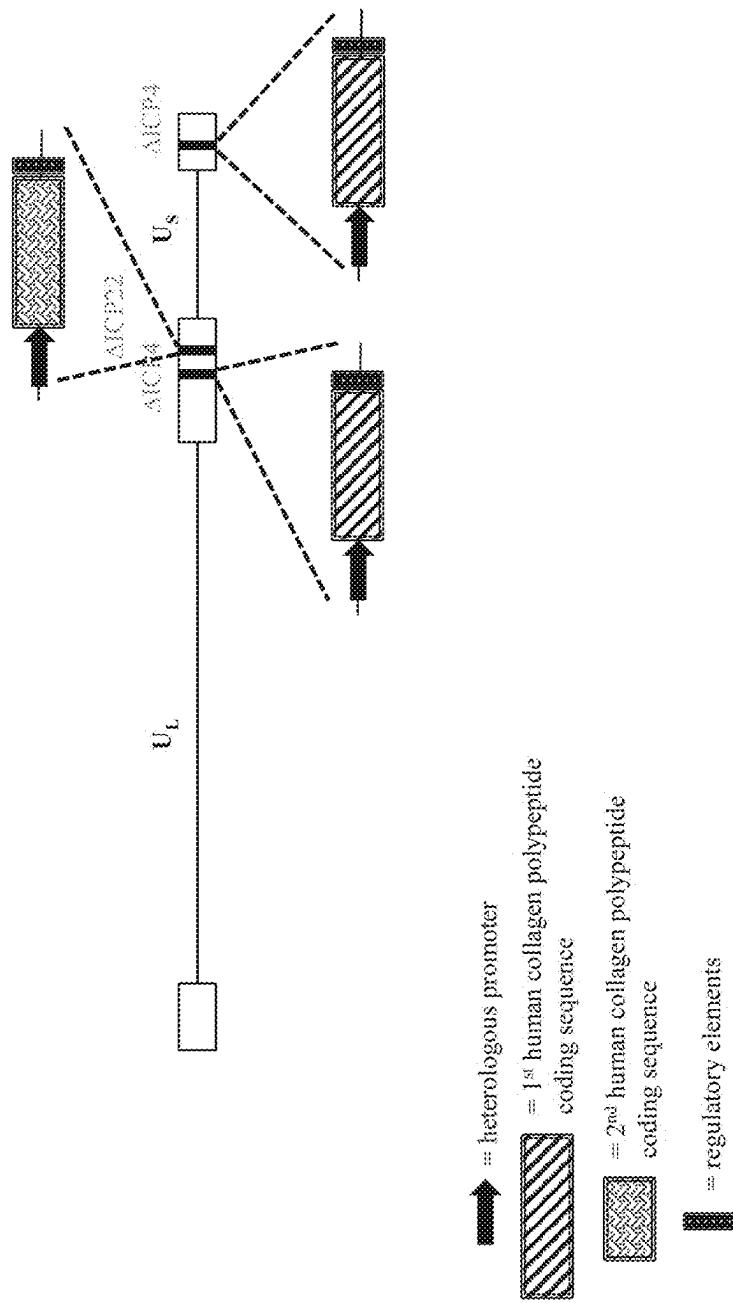


FIG. 1M

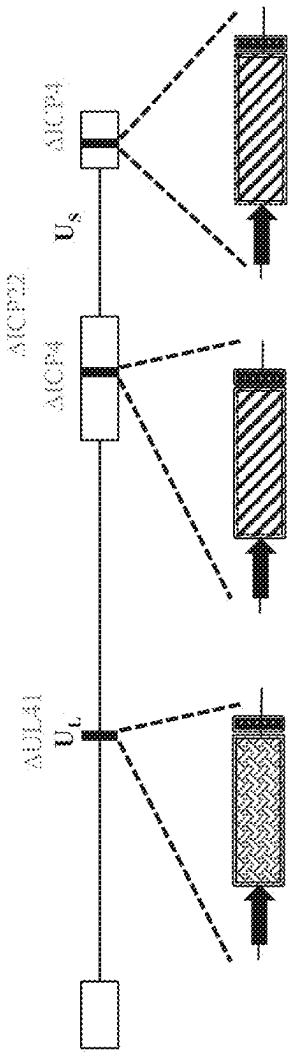


FIG. 1N

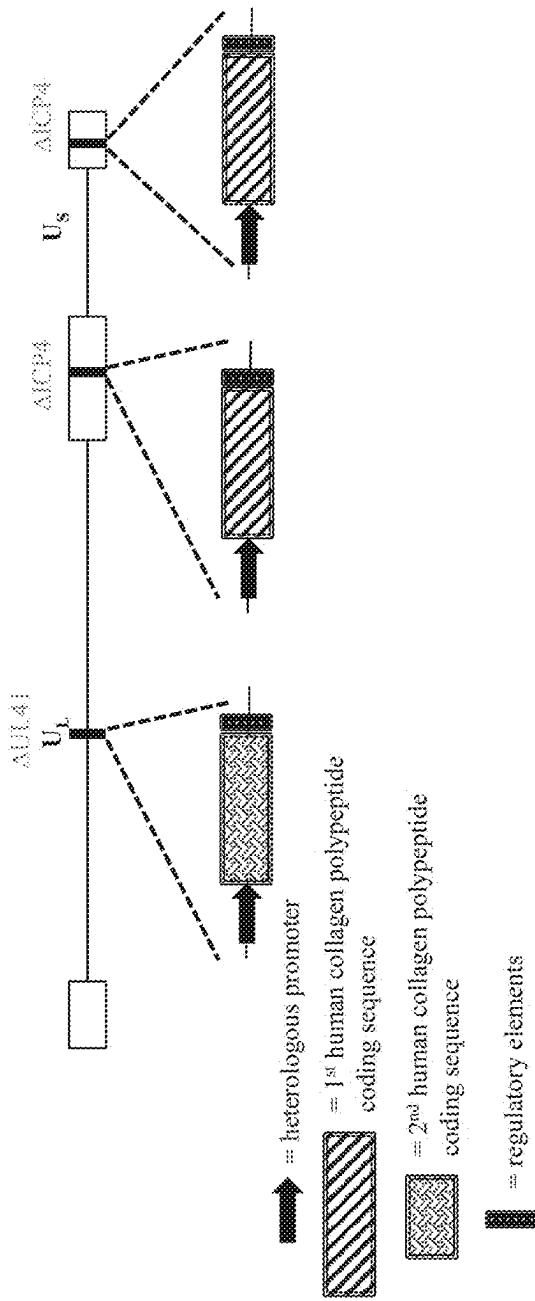


FIG. 2A

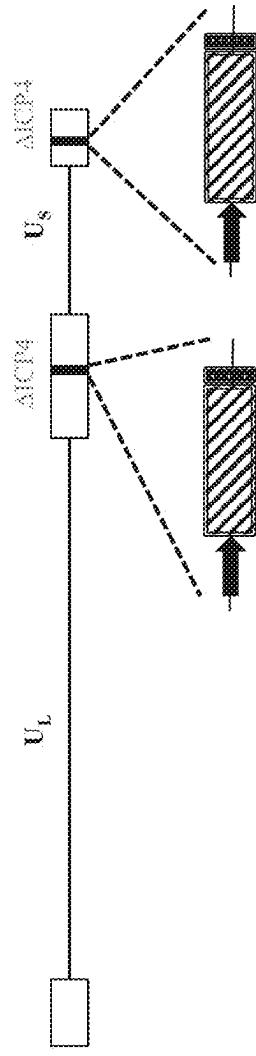


FIG. 2B

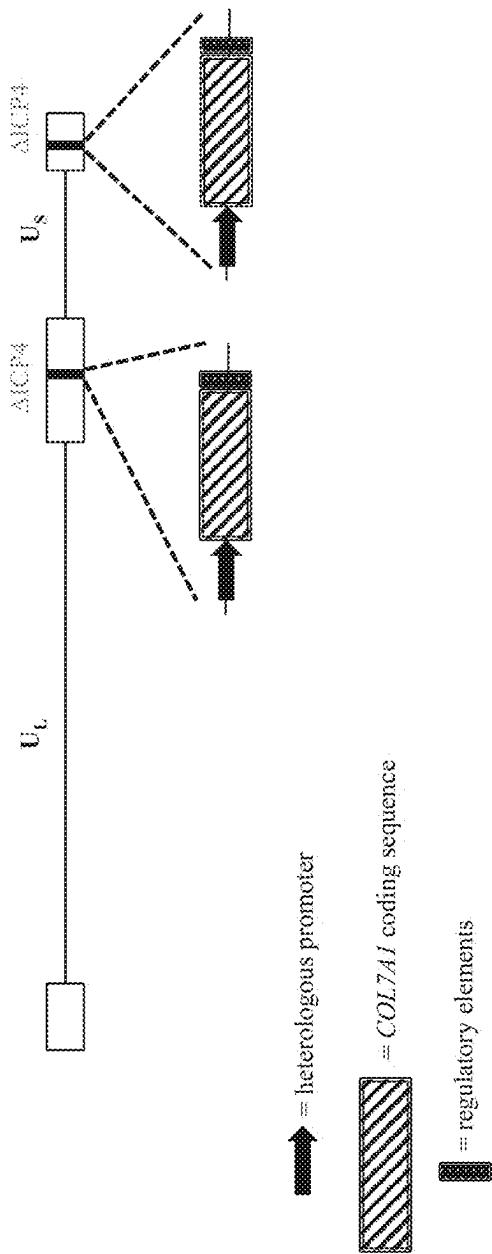


FIG. 3A

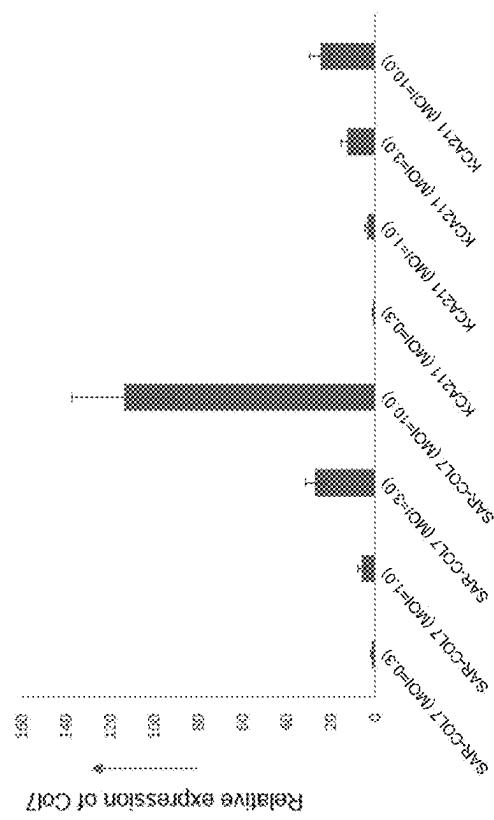


FIG. 3B

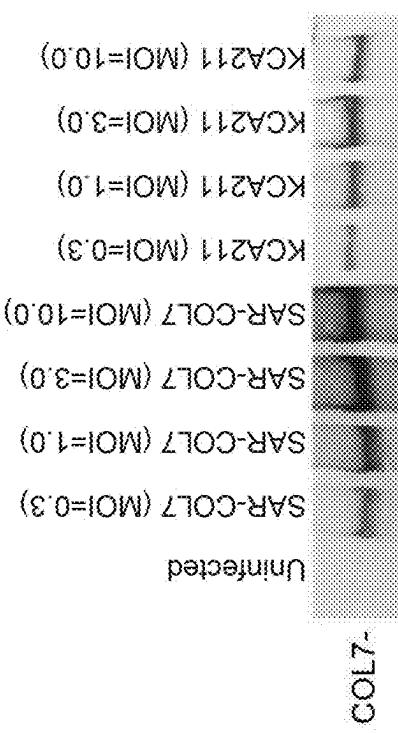


FIG. 4A

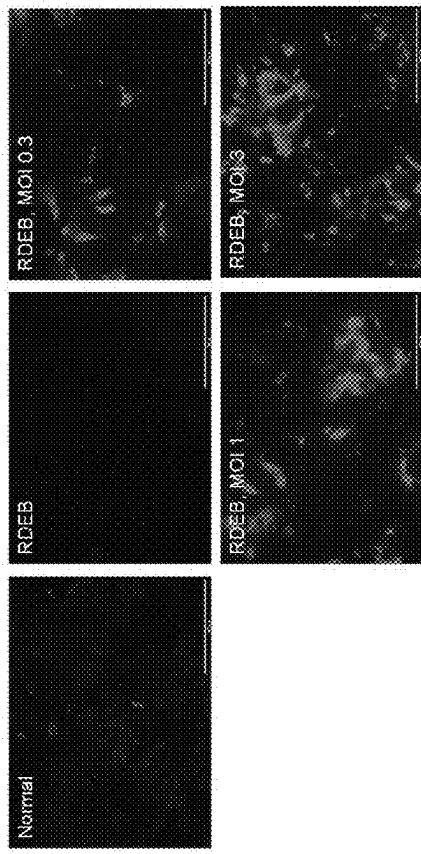


FIG. 4B

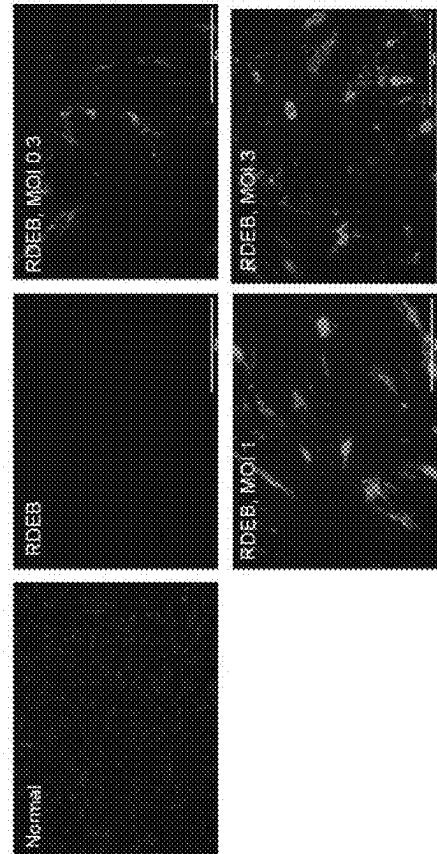


FIG. 5A

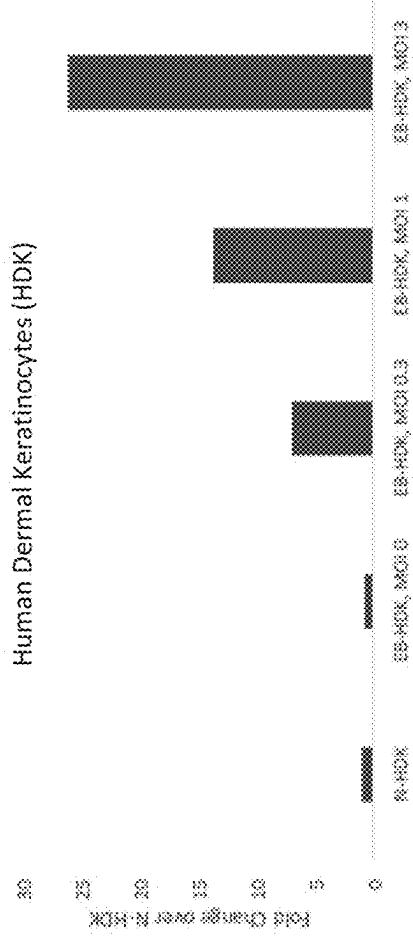


FIG. 5B

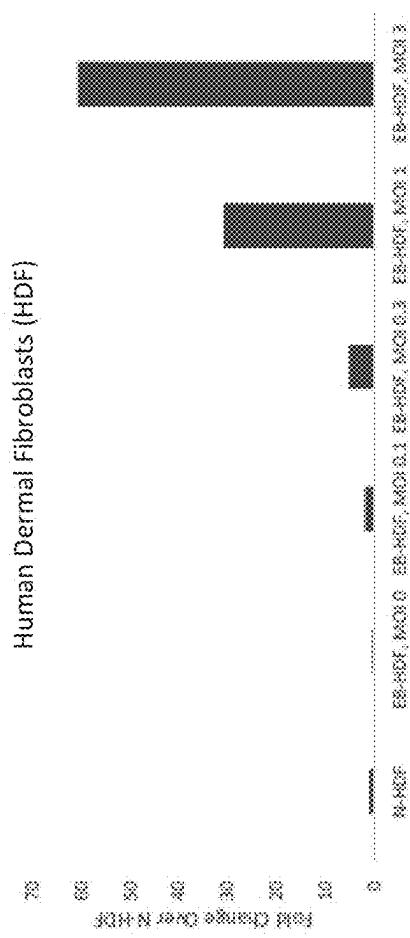


FIG. 6A
Adhesion to Collagen I

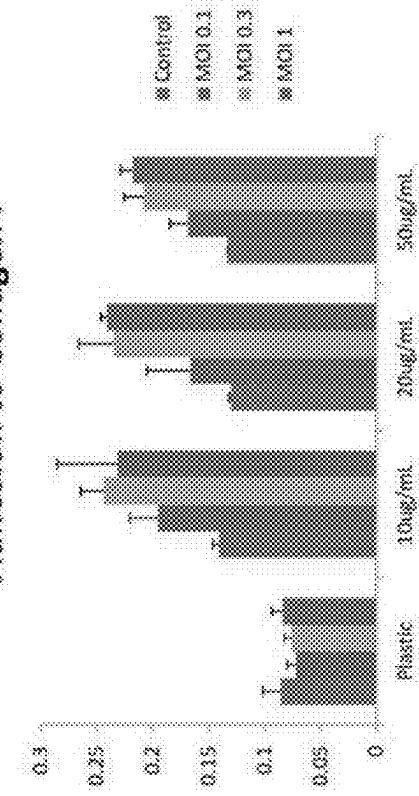


FIG. 6B
Adhesion to Fibronectin

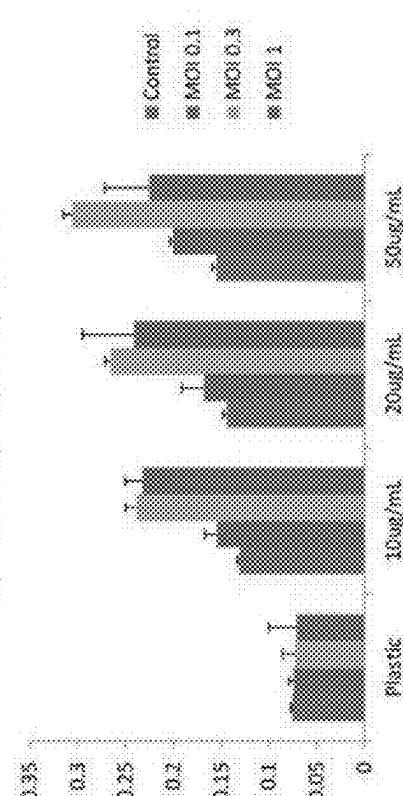


FIG. 7

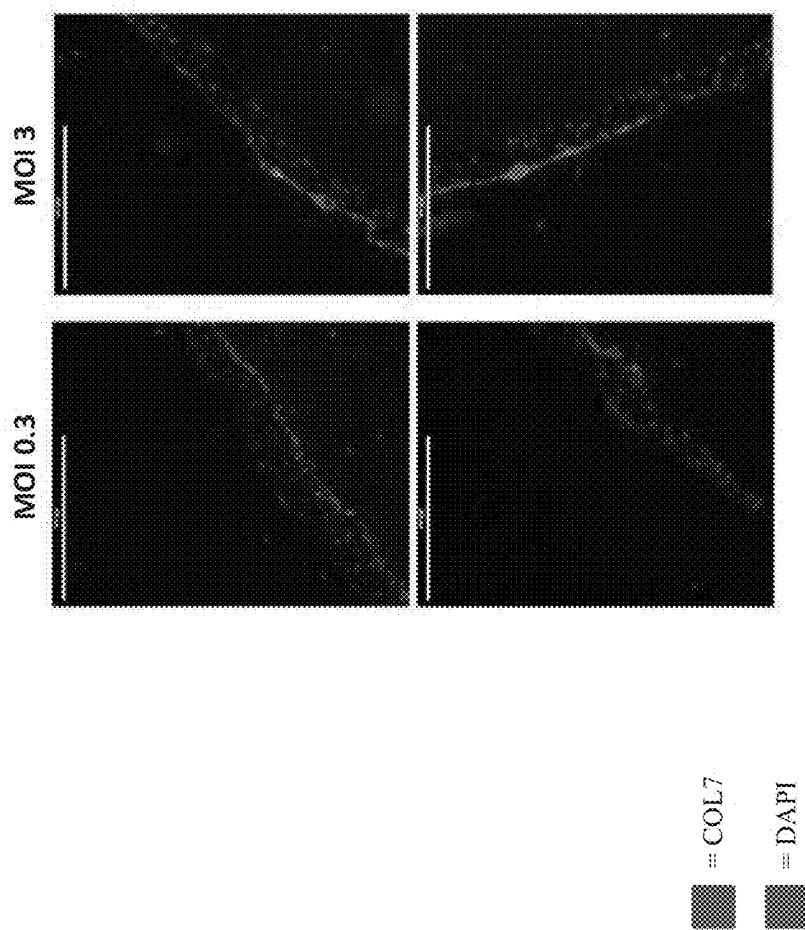


FIG. 8A

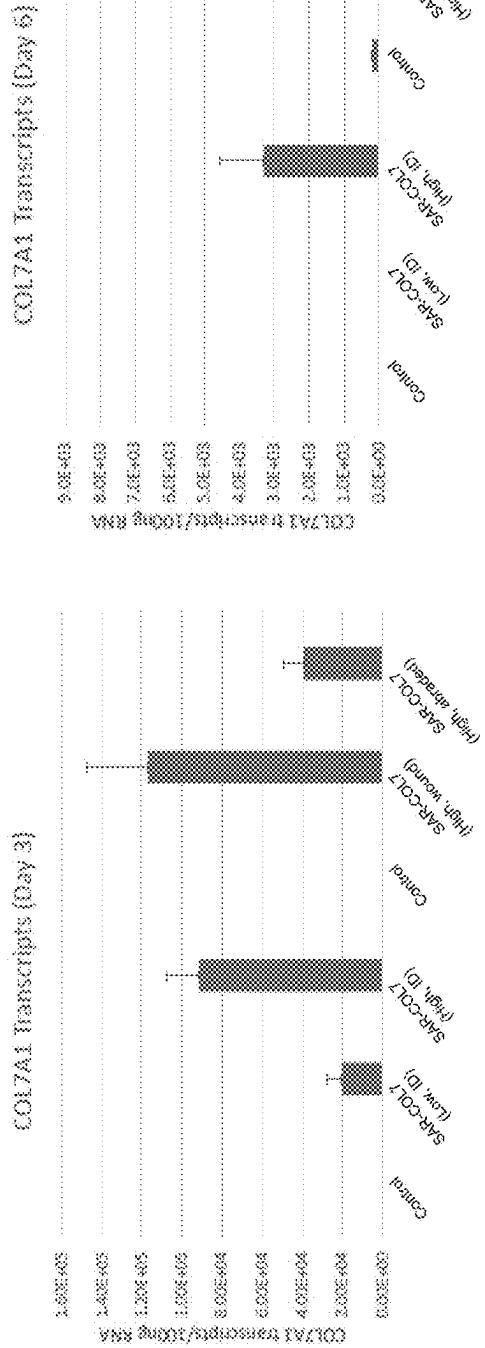


FIG. 8B

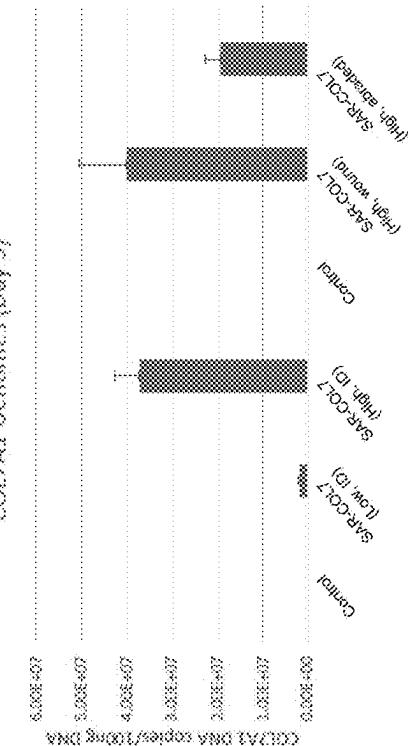


FIG. 8C

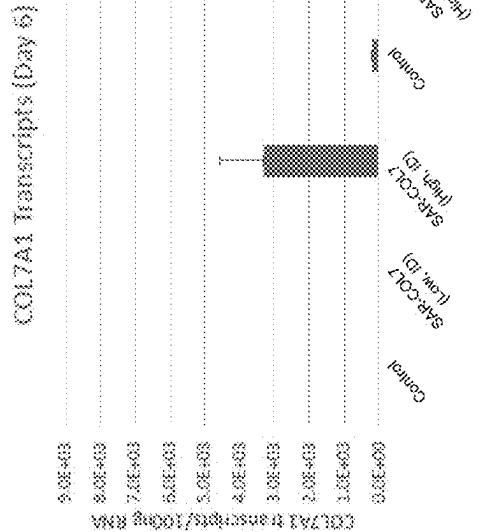


FIG. 8D

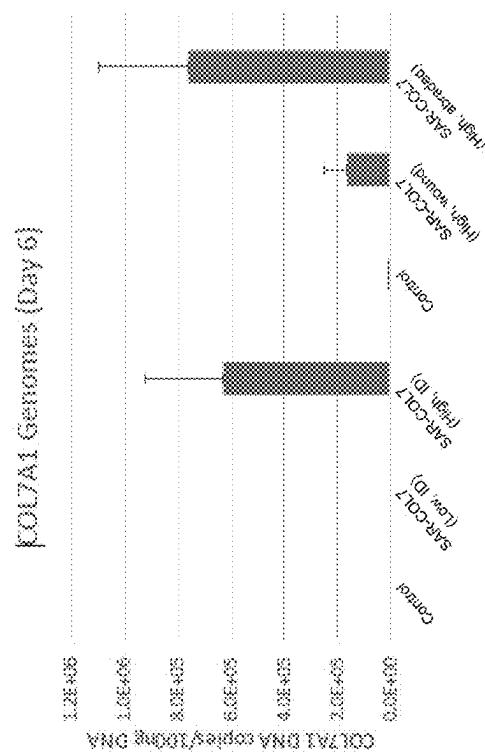


FIG. 9A
FIG. 9B

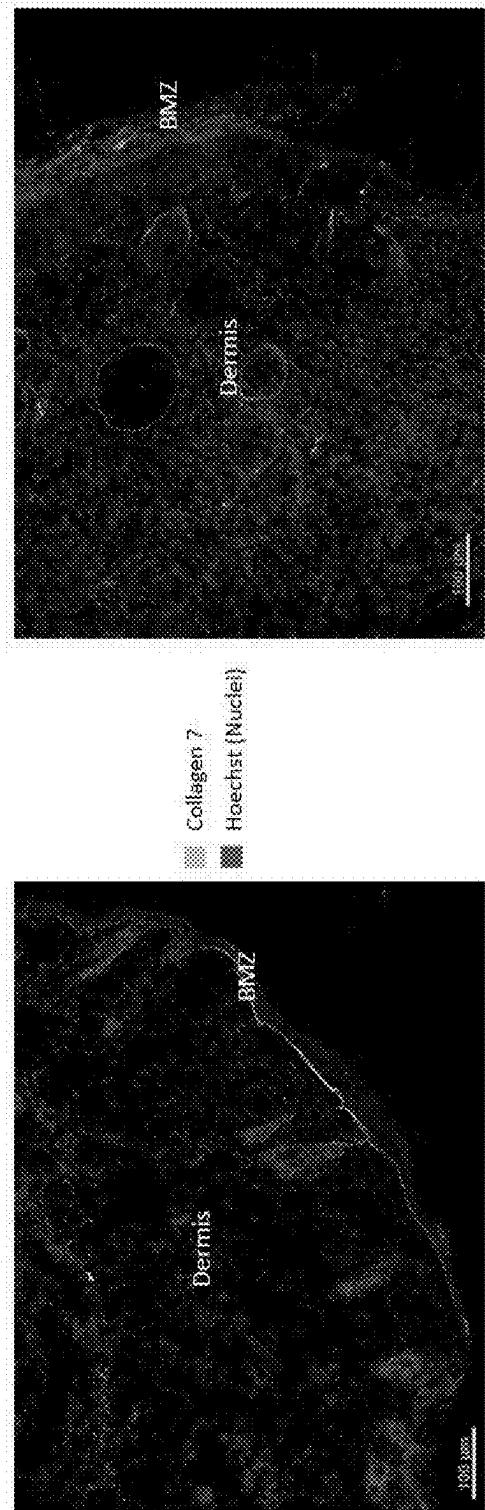


FIG. 10A

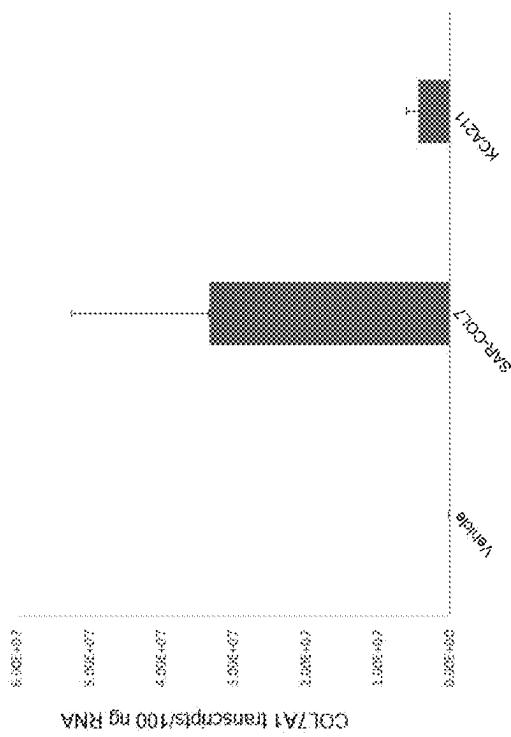


FIG. 10B

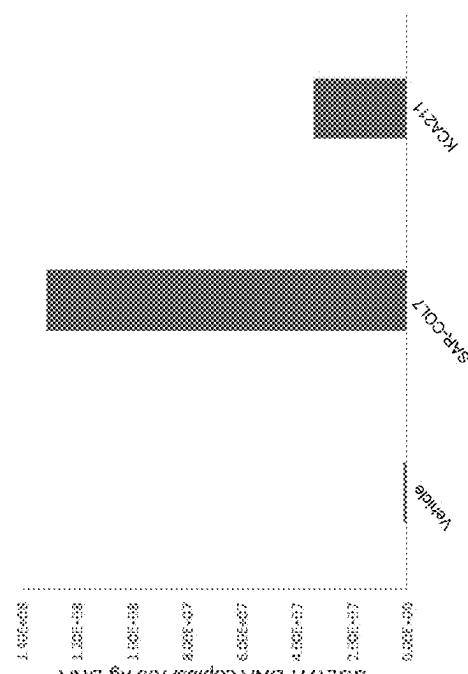


FIG. 11A

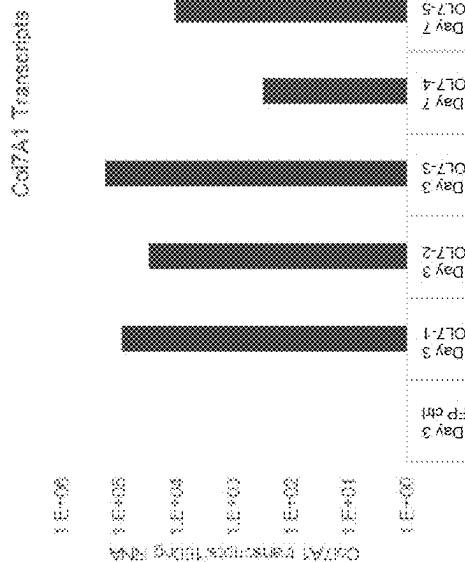


FIG. 11B

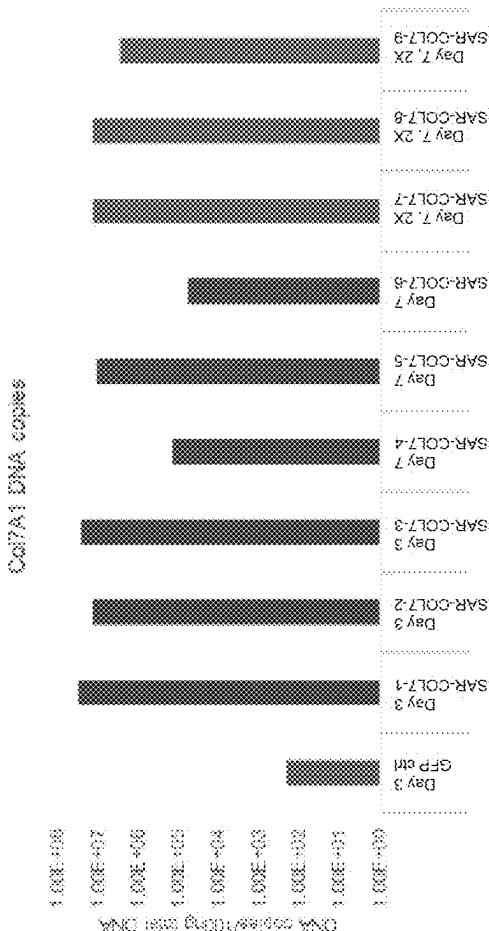


FIG. 12A

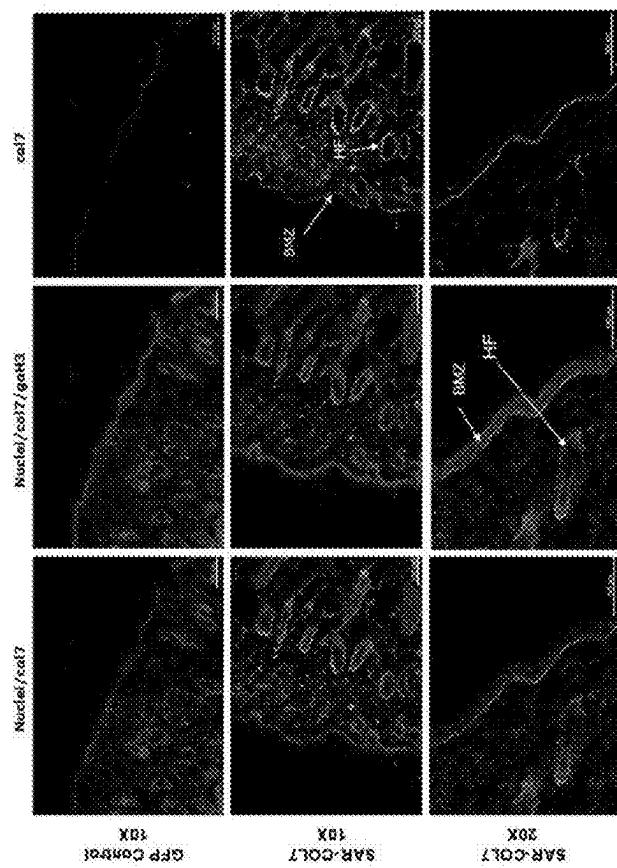


FIG. 12B

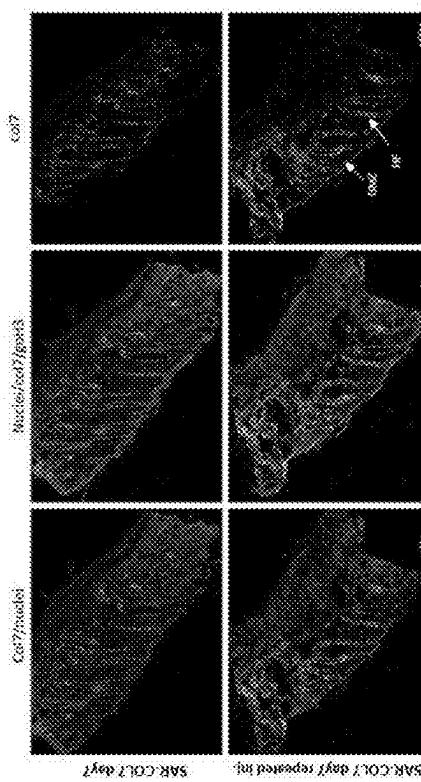


FIG. 13

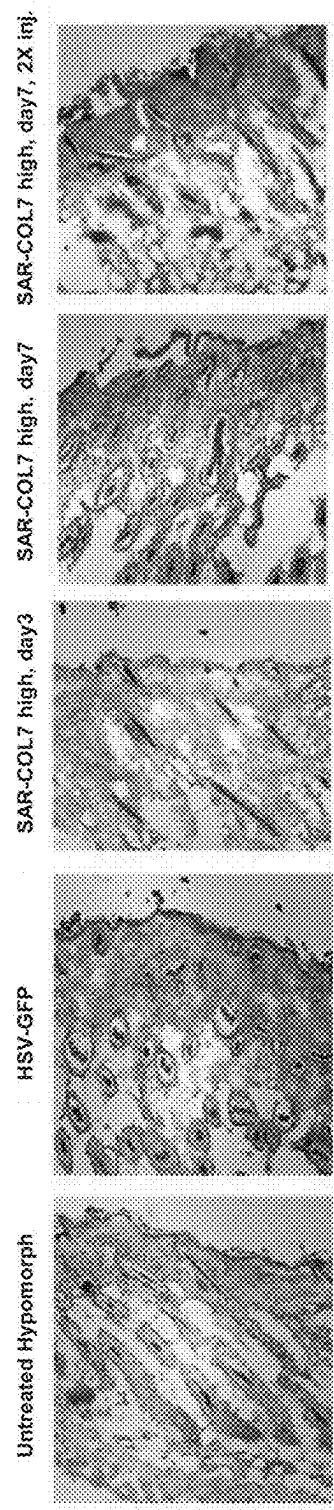


FIG. 14A

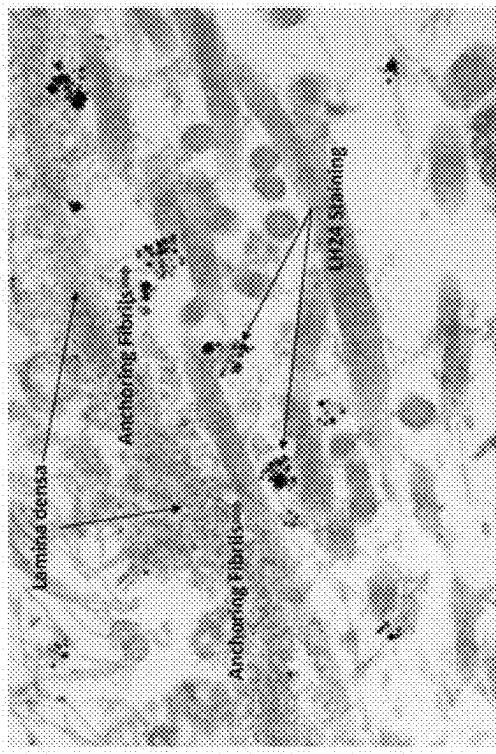


FIG. 14B

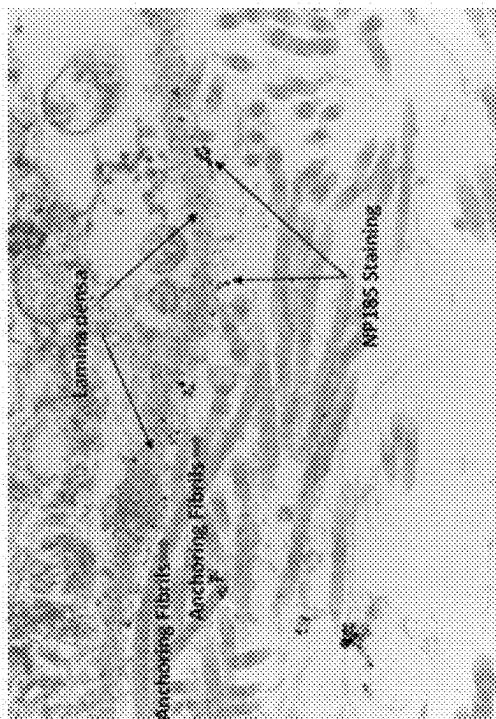


FIG. 15A

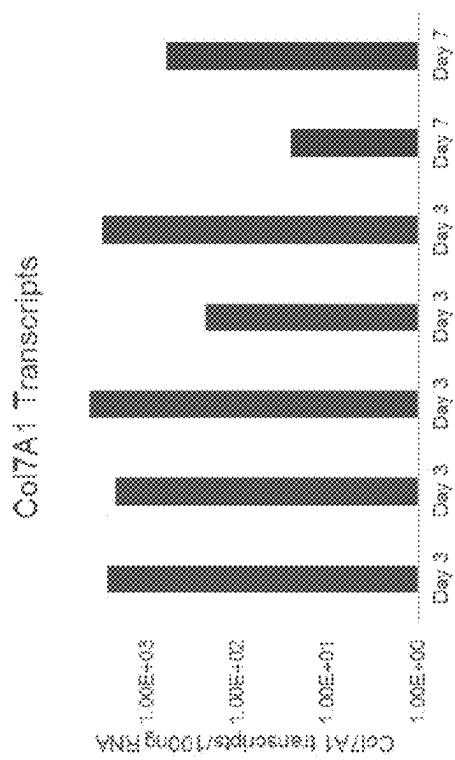
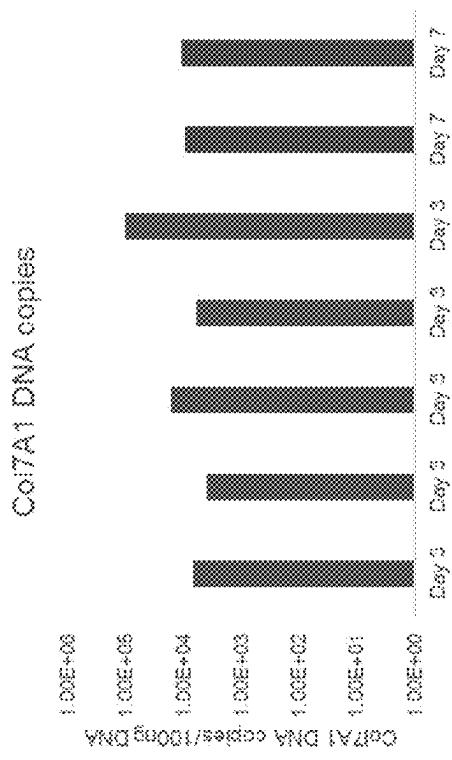


FIG. 15B



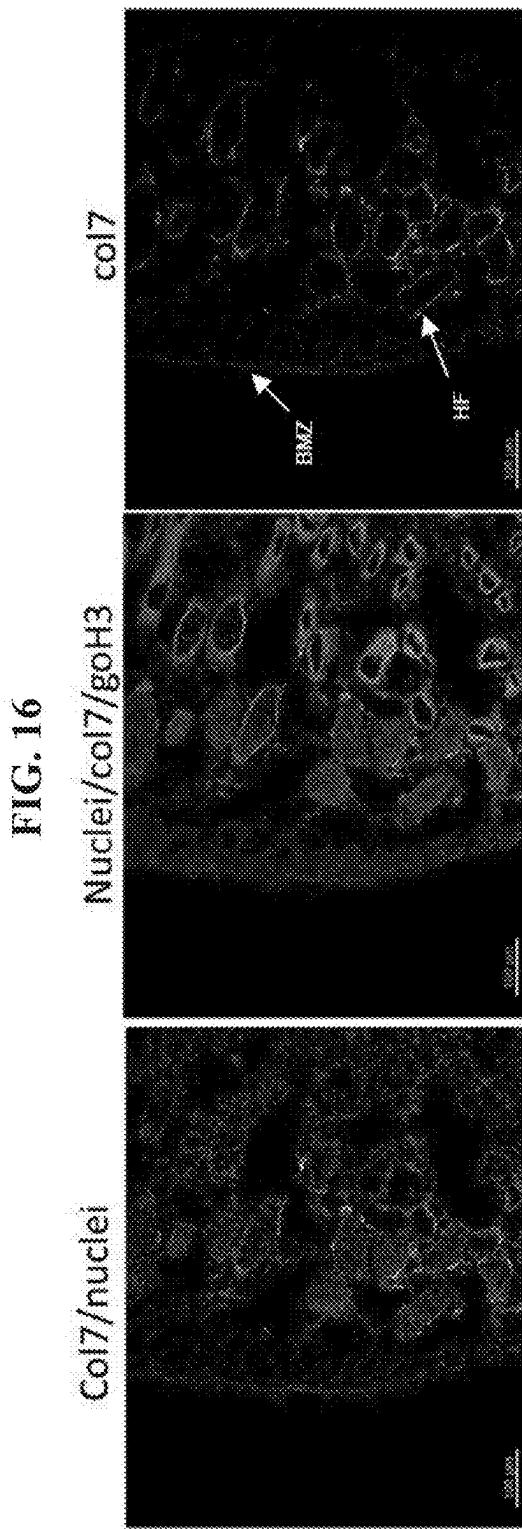


FIG. 17C

COL1A1 Transcripts

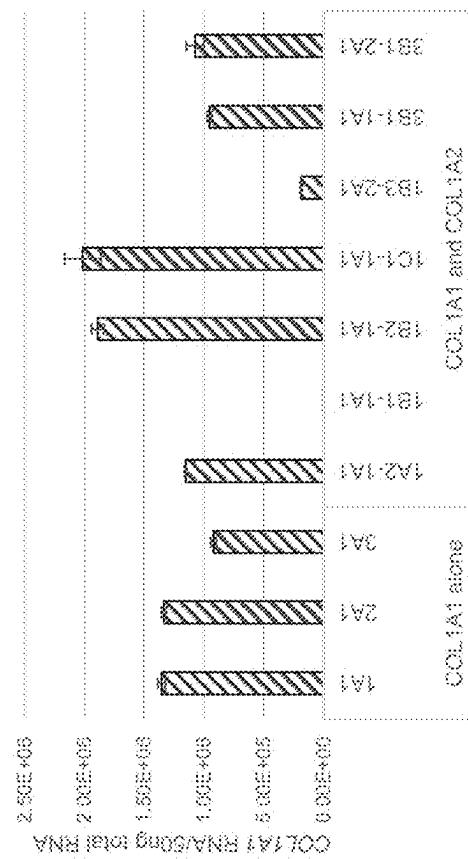


FIG. 17B
COL1A2 Transcripts

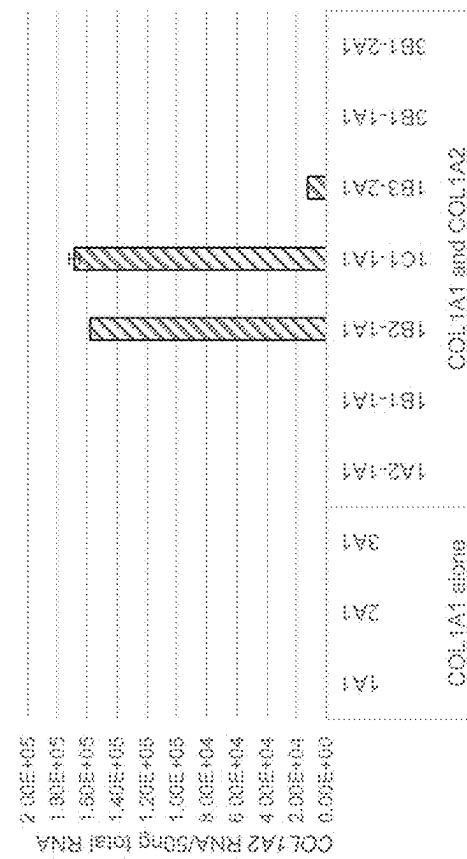


FIG. 18

No insertion
IRES-isolate 6

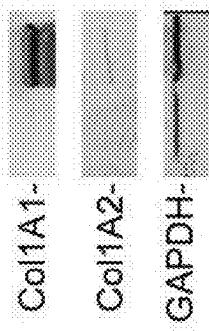


FIG. 19A

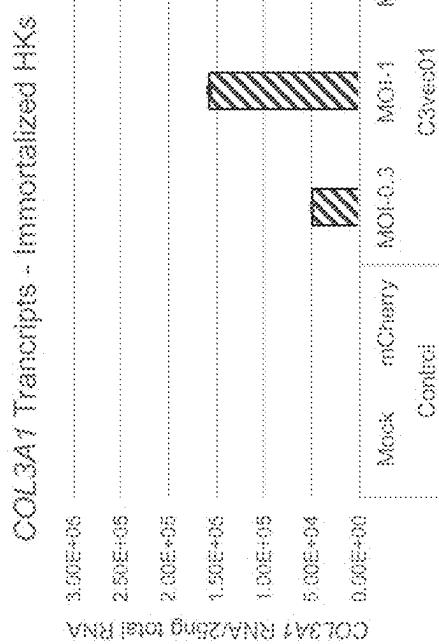


FIG. 19B

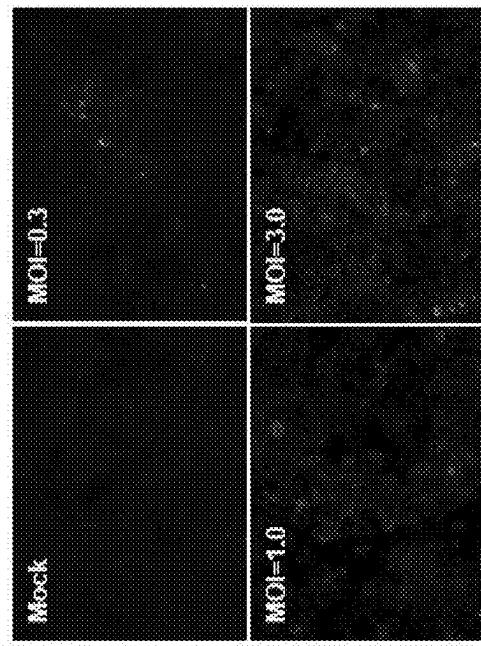


FIG. 20A

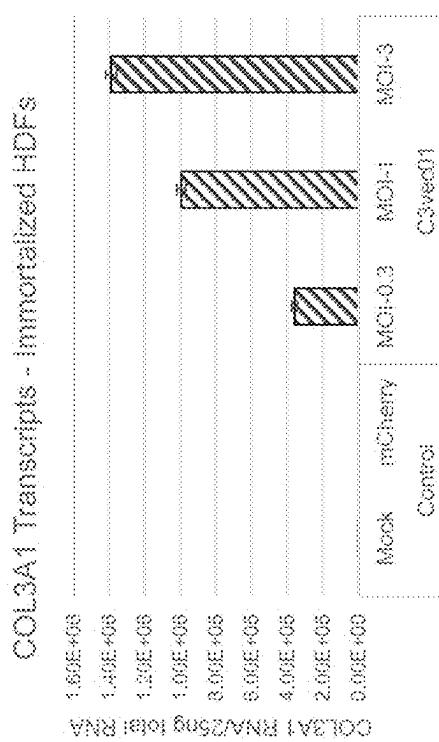


FIG. 20B

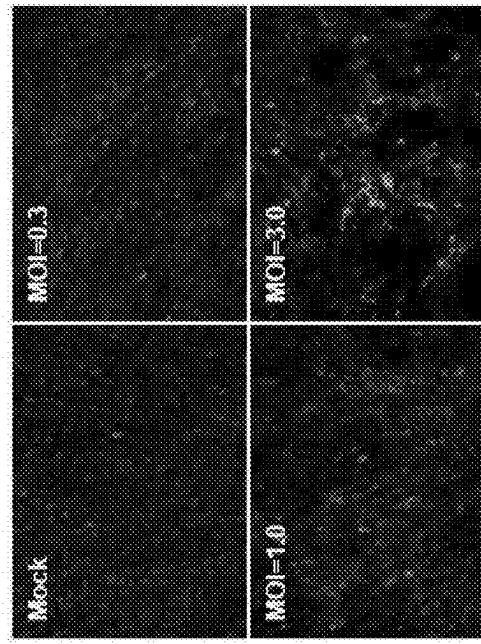


FIG. 21A

COL3A1 Transcripts - Primary HDFs

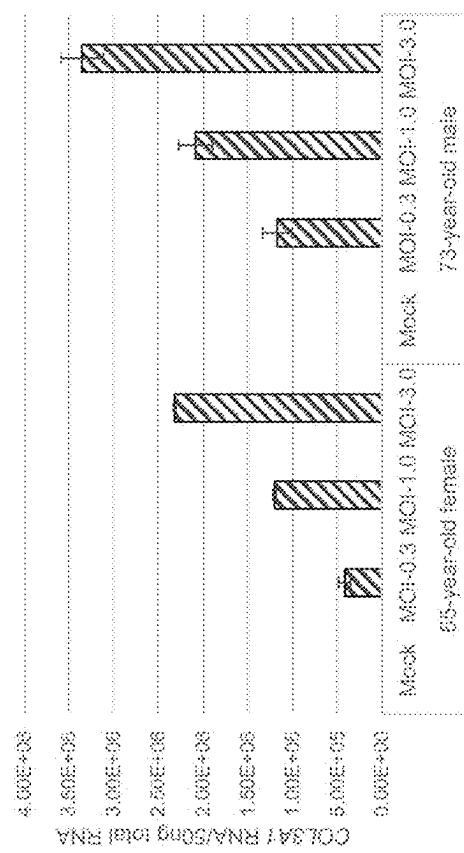


FIG. 21B

COL3A1 Transcripts - Primary HDFs

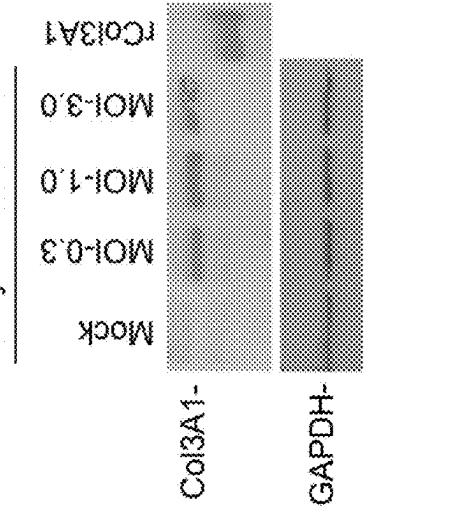


FIG. 21C

COL3A1 Transcripts - Primary HDFs

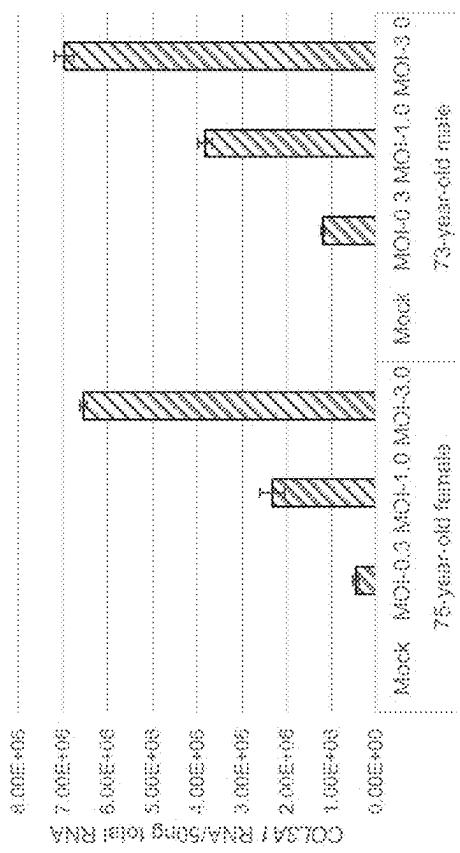


FIG. 21D

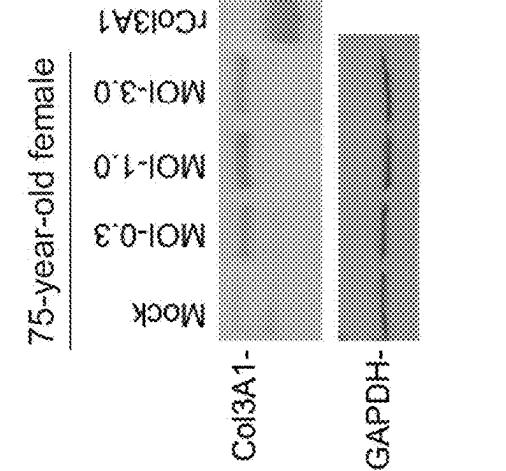


FIG. 22A
Secreted Human COL3

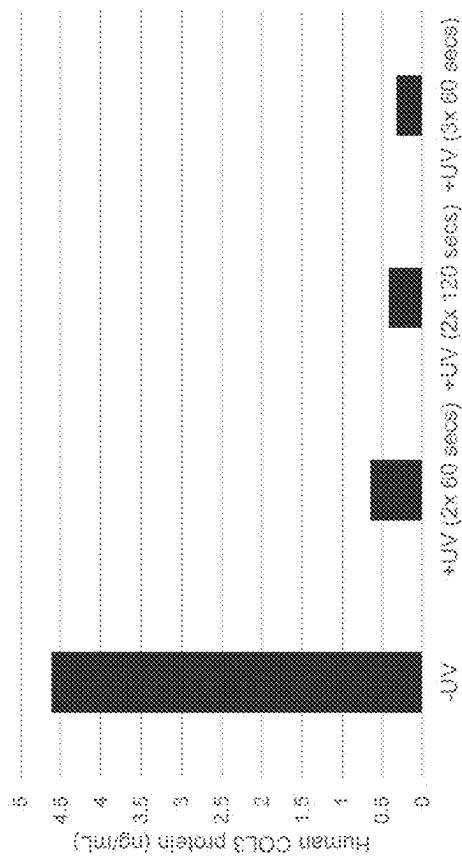


FIG. 22B
COL3A1 Transcripts - UV-damaged HUFs



FIG. 23A

Human COL3A1 DNA

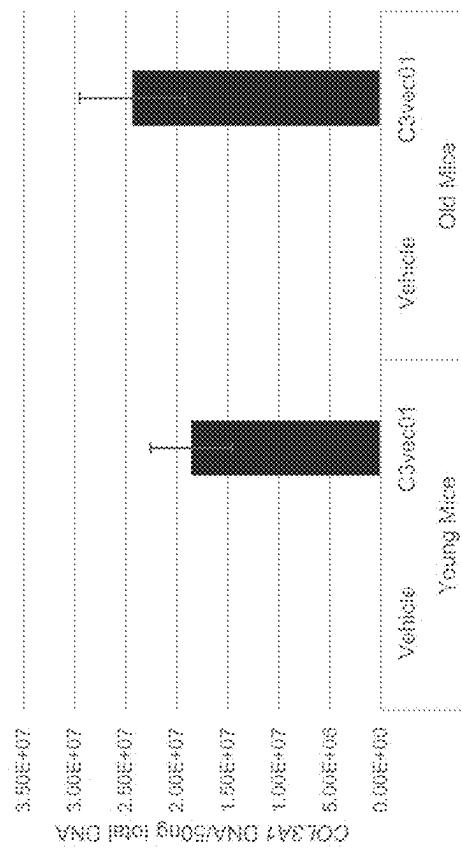


FIG. 23B

Human COL3A1 Transcripts

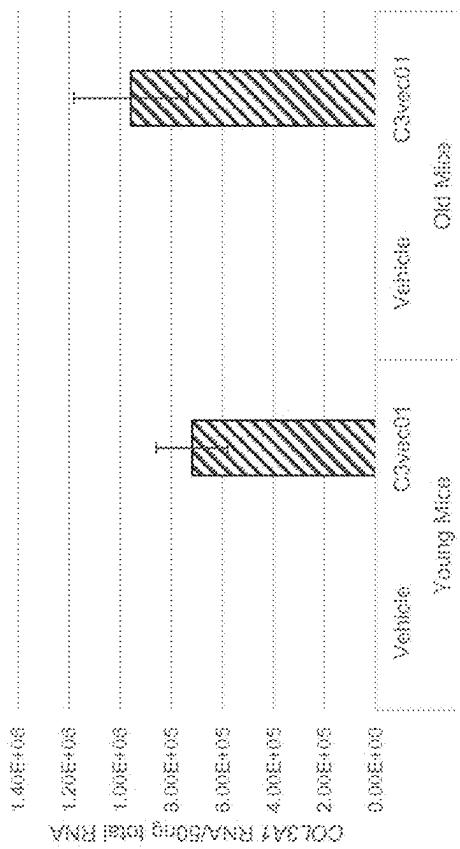


FIG. 23C

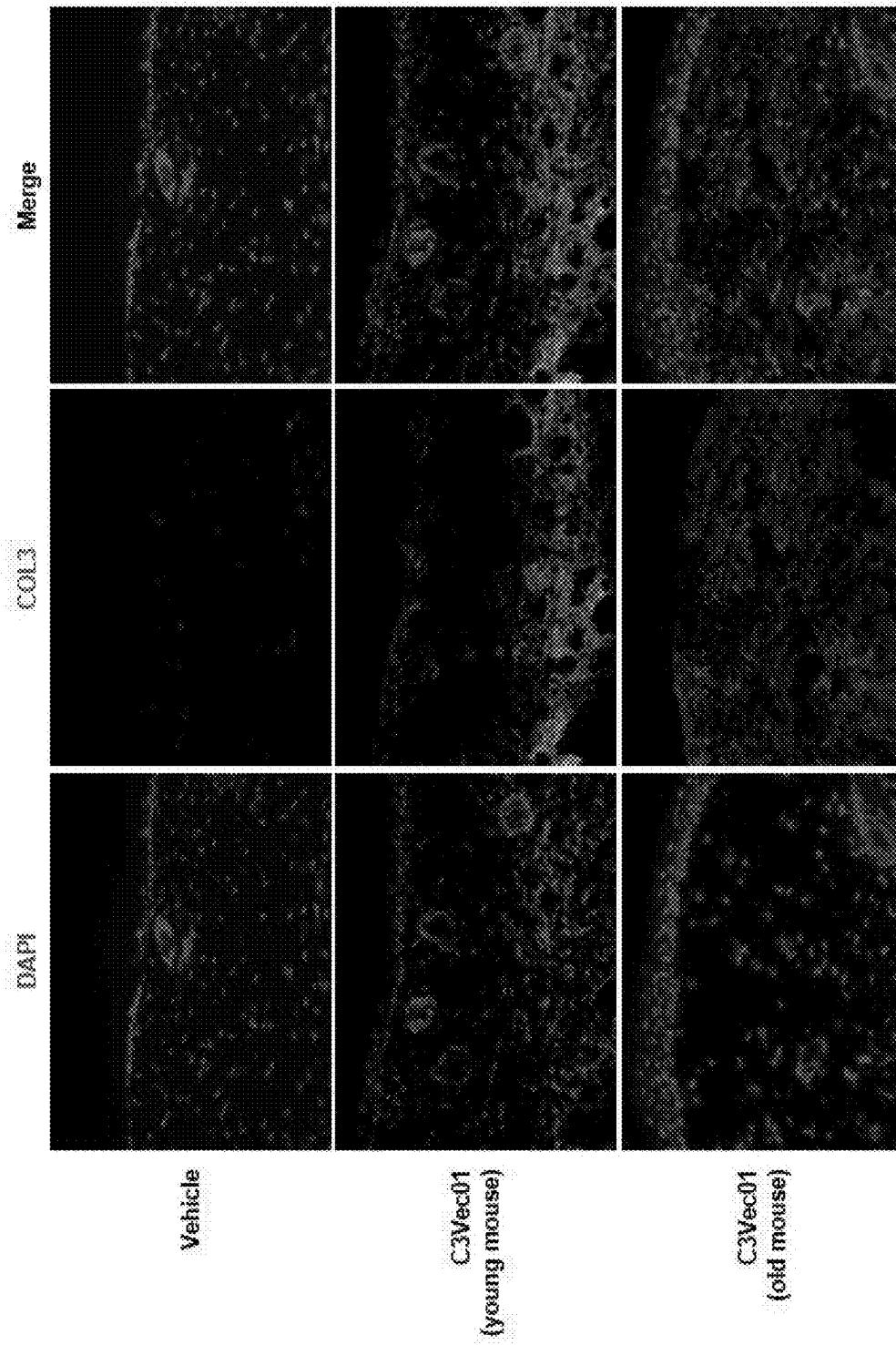


FIG. 24A

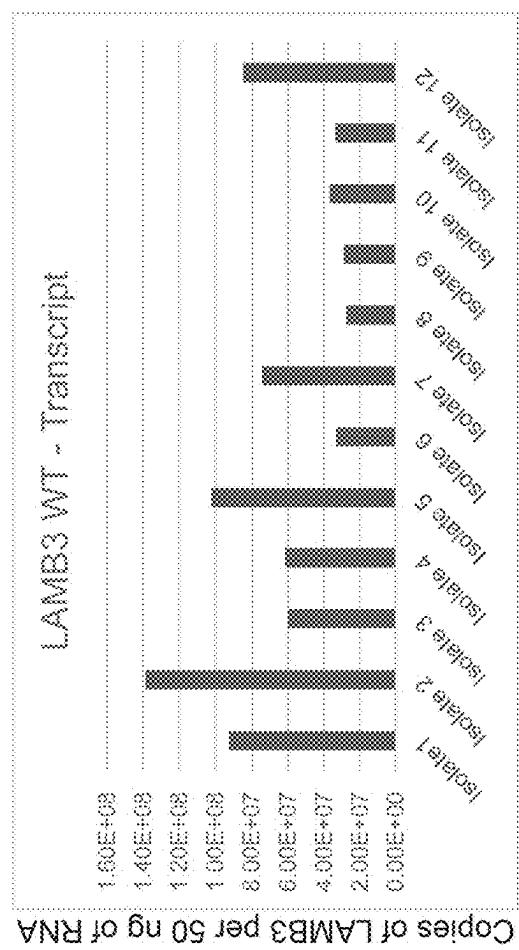


FIG. 24B

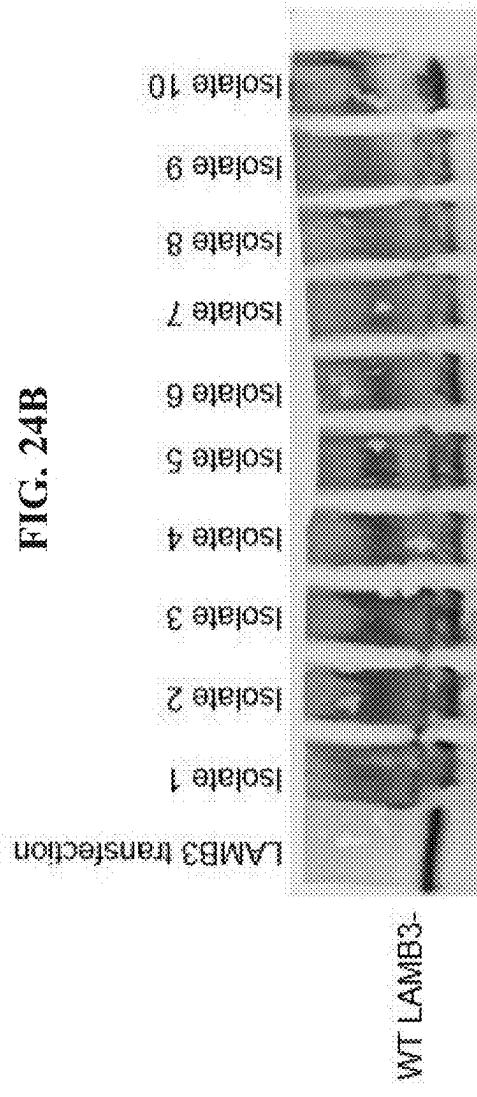


FIG. 25

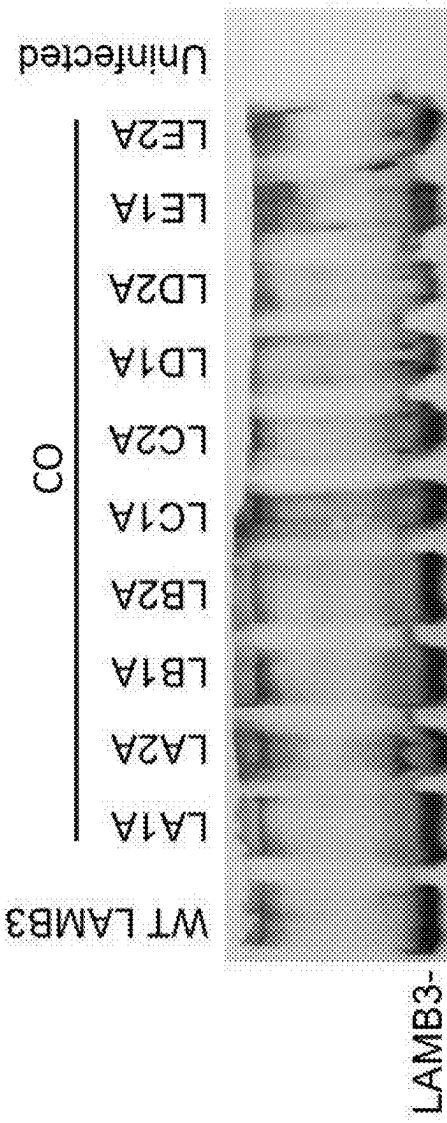


FIG. 26

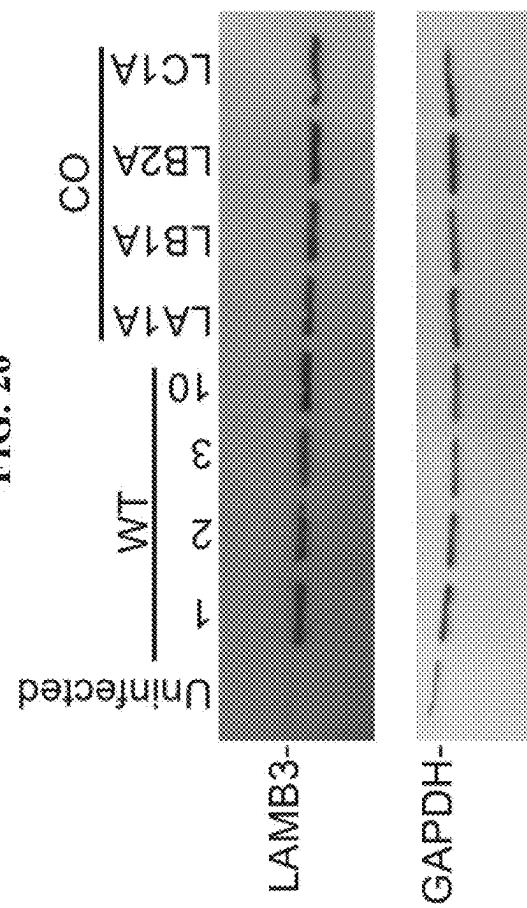


FIG. 27A

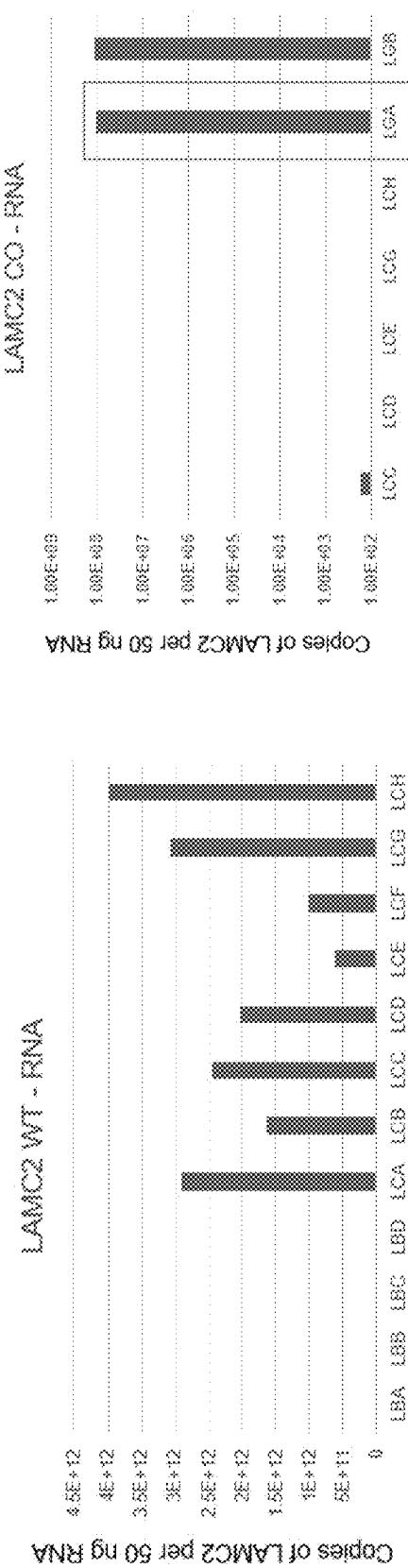


FIG. 27B

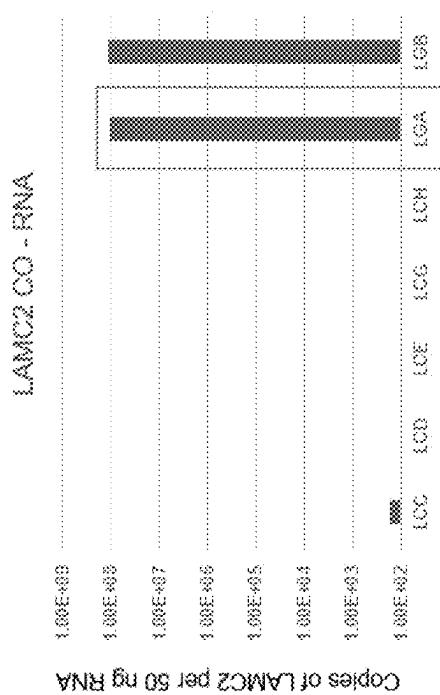


FIG. 27C



FIG. 28A

LAMC2 DNA

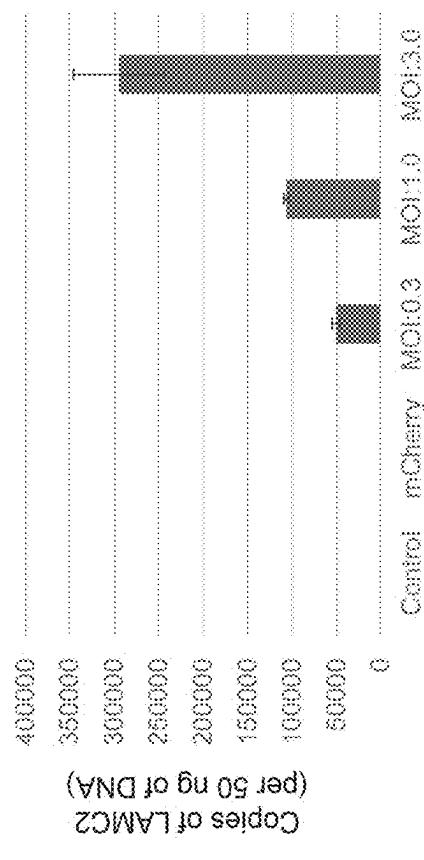
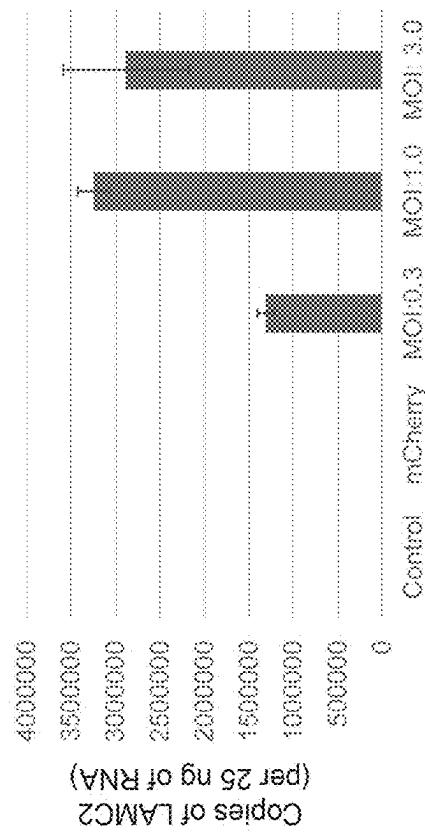


FIG. 28B

LAMC2 transcript



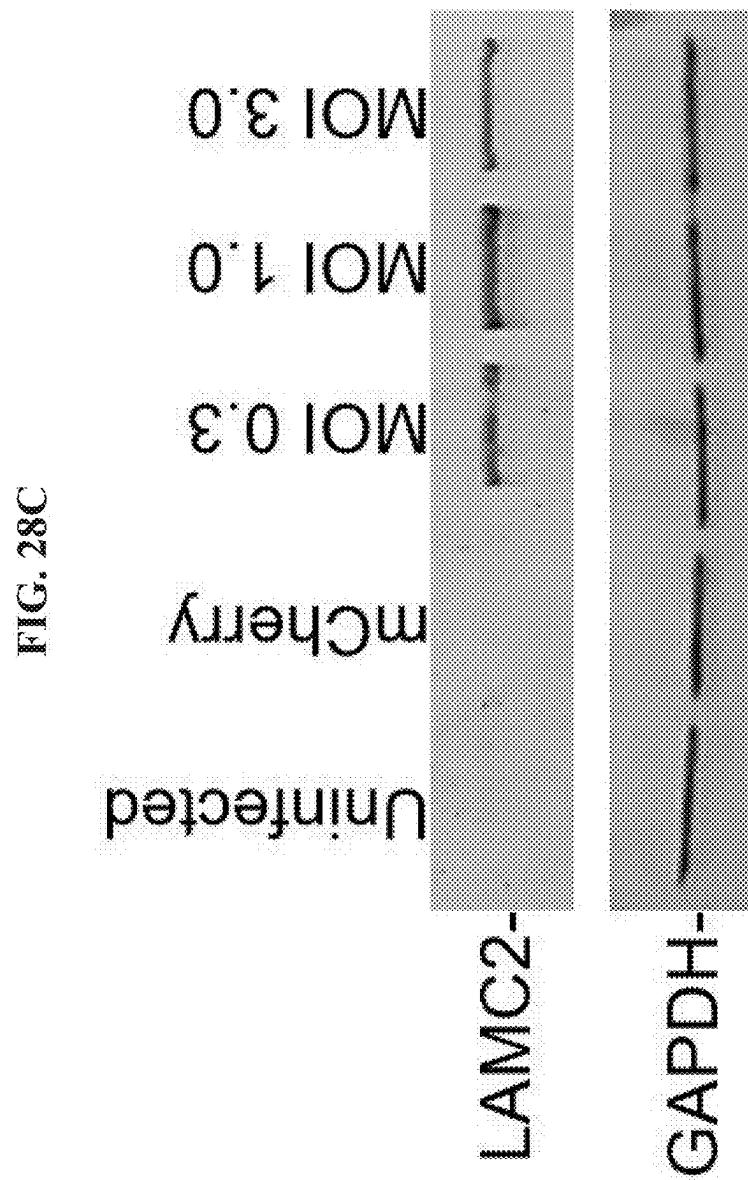


FIG. 29A
Mouse #1
Mouse #2

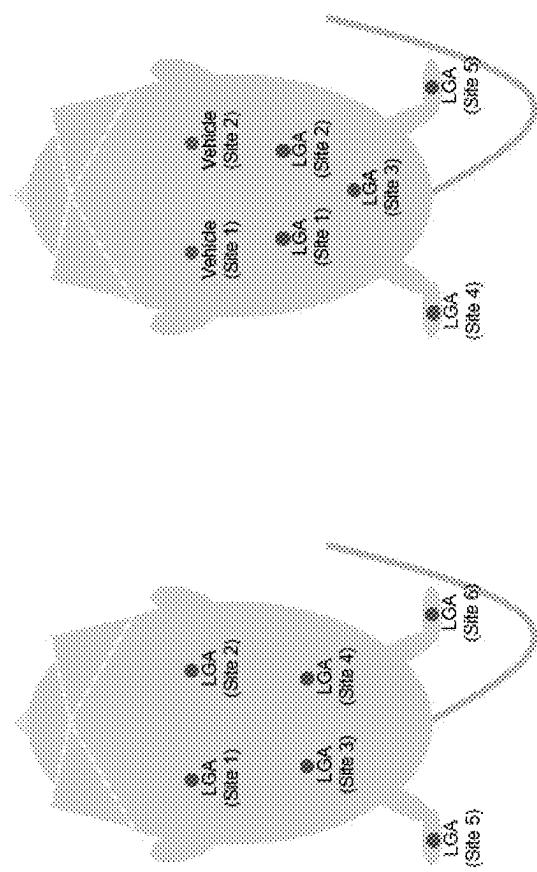


FIG. 29B

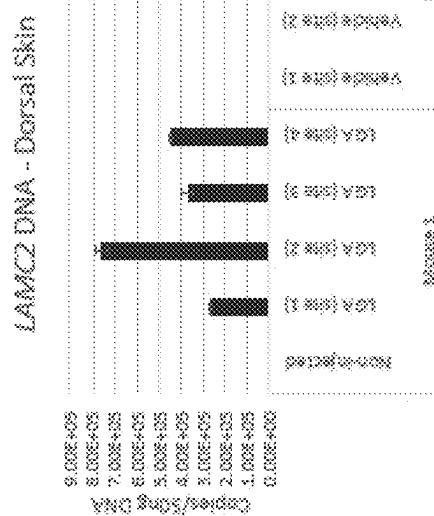
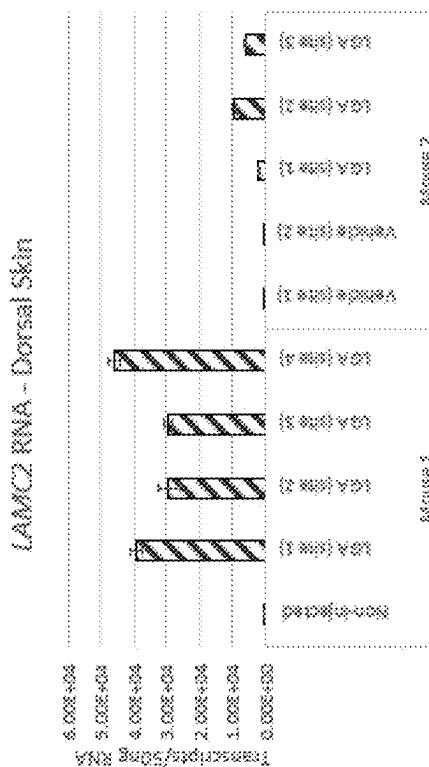
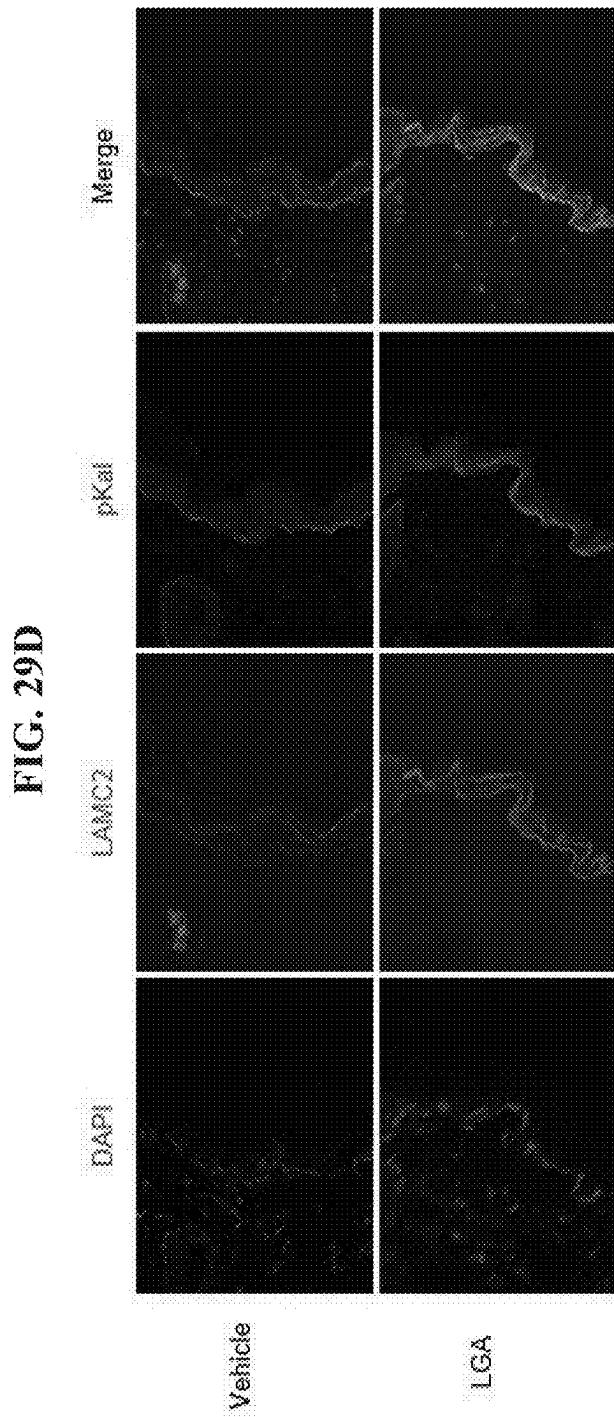


FIG. 29C





RECOMBINANT NUCLEIC ACIDS ENCODING COSMETIC PROTEIN(S) FOR AESTHETIC APPLICATIONS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation of U.S. patent application Ser. No. 16/915,695, filed Jun. 29, 2020, which is a continuation of U.S. patent application Ser. No. 16/395,896, filed Apr. 26, 2019, now issued as U.S. Pat. No. 10,786,438, which claims the priority benefit of U.S. Provisional Application Ser. No. 62/663,476, filed Apr. 27, 2018, each of which is incorporated herein by reference in its entirety.

SUBMISSION OF SEQUENCE LISTING ON ASCII TEXT FILE

[0002] The Sequence Listing associated with this application is filed in electronic format and is hereby incorporated by reference into the specification in its entirety. The name of the XML file containing the Sequence Listing is 2407502.xml. The size of the XML file is 258,534 bytes and the XML file was created on Sep. 19, 2024.

FIELD OF THE INVENTION

[0003] The present disclosure relates, in part, to recombinant nucleic acids comprising one or more polynucleotides encoding one or more cosmetic proteins (e.g., one or more human collagen proteins); to viruses comprising the recombinant nucleic acids; to compositions (e.g., cosmetic formulations) comprising the recombinant nucleic acids and/or viruses; to methods of their use; and to articles of manufacture or kits thereof.

BACKGROUND

[0004] Skin, like all organs in the human body, undergoes sequential and often cumulative alterations with the passage of time. Aging of the skin occurs as the result of numerous factors, including inherent changes within the skin, the effects of gravity and facial muscles acting on the skin, soft tissue loss or shift, and loss of tissue elasticity. Interestingly, the “aged” phenotype of skin may be accelerated by environmental factors, most notably, chronic exposure to ultraviolet irradiation (e.g., from the sun). Clinically, the aged phenotype of skin may be described as wrinkled, sagging, and/or generally less elastic and resilient than its youthful counterpart, although variations within this phenotype exist between natural, chronological aging and photoaging.

[0005] The dermal extracellular matrix (ECM) comprises the bulk of skin and confers both strength and resiliency. Collagen, a major component of the connective tissue providing support to the skin, decreases as a person ages. In aged skin, collagen fibrils display high levels of degradation and fragmentation, and are replenished by dermal fibroblasts at diminishing rates. These degraded and fragmented collagen bundles become looser and lose strength (disrupting the structural organization of the dermal ECM), and inextricably leads to an “aged” manifestation of the skin.

[0006] Numerous skincare products have been developed for improving the appearance of human skin. Wrinkles and skin folds are commonly treated with dermal and subdermal injections of aesthetic facial fillers; however, such a superficial approach does not address the structural changes

underlying skin aging, in particular, the damage or loss of collagen. Thus, there exists a clear need for alternative strategies to supplement, strengthen, or replace dermal ECM components (e.g., human collagen), in individuals desiring to combat or reverse the physiological effects of skin aging.

[0007] All references cited herein, including patent applications, patent publications, non-patent literature, and NCBI/UniProtKB/Swiss-Prot Accession numbers are herein incorporated by reference in their entirety, as if each individual reference were specifically and individually indicated to be incorporated by reference.

BRIEF SUMMARY

[0008] In order to meet these and other needs, provided herein are recombinant nucleic acids (e.g., recombinant herpes viral genomes) encoding one or more cosmetic proteins for use in viruses (e.g., herpes viruses), compositions, formulations, medicaments, and/or methods for aesthetic/cosmetic applications (e.g., treating wrinkles). The present inventors have shown that the recombinant, attenuated viruses described herein were capable of 1) effectively transducing human epidermal/dermal cells, and 2) successfully expressing the encoded exogenous human collagen (mRNA and protein), where the protein could then localize to the appropriate region in skin-equivalent organotypic cultures (see e.g., Example 2). Moreover, the present inventors have shown that the viruses described herein may be successfully administered either topically or intradermally without significant host cell cytotoxicity, allowing for the human collagen expressed from these viruses to localize to the appropriate region of the dermal ECM after in vivo administration without observable damage to the skin (see e.g., Examples 3 and 7). In addition, the present inventors have shown that multiple different HSV backbones can be used to construct viruses expressing human collagens (see e.g., Example 2), that multiple strategies can be employed to successfully express more than one human collagen protein from a single recombinant genome (see e.g., Example 5), and that candidate viruses can successfully express human collagen proteins in multiple relevant in vitro and in vivo models of chronological or UV-induced skin aging (see e.g., Examples 6 and 7). Furthermore, the present inventors have shown that the viruses described herein can be successfully engineered to express other cosmetic proteins (e.g., human laminins) both in vitro and in vivo, where these proteins localize to the appropriate region of the dermal ECM (see e.g., Example 8). Without wishing to be bound by theory, the data described herein provides strong evidence that the recombinant nucleic acids and/or viruses of the present disclosure may constitute a novel means for delivering cosmetic proteins (e.g., human collagen proteins, such as human Collagen 1 and human Collagen 3), and in particular, to supplement or replace natural human dermal ECM proteins in aesthetic applications (e.g., to reduce the appearance of age or photo-induced wrinkles).

[0009] Accordingly, certain aspects of the present disclosure relate to a recombinant herpes virus genome comprising a first polynucleotide encoding a first polypeptide comprising a first cosmetic protein. In some embodiments, the recombinant herpes virus genome comprises two or more copies of the first polynucleotide. In some embodiments, the recombinant herpes virus genome is replication competent. In some embodiments, the recombinant herpes virus genome is replication defective. In some embodiments that may be

combined with any of the preceding embodiments, the recombinant herpes virus genome is selected from a recombinant herpes simplex virus genome, a recombinant varicella zoster virus genome, a recombinant human cytomegalovirus genome, a recombinant herpesvirus 6A genome, a recombinant herpesvirus 6B genome, a recombinant herpesvirus 7 genome, a recombinant Kaposi's sarcoma-associated herpesvirus genome, and any derivatives thereof. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes virus genome is a recombinant herpes simplex virus genome. In some embodiments, the recombinant herpes simplex virus genome is a recombinant type 1 herpes simplex virus (HSV-1) genome, a recombinant type 2 herpes simplex virus (HSV-2) genome, or any derivatives thereof.

[0010] In some embodiments, the recombinant herpes simplex virus genome is a recombinant type 1 herpes simplex virus (HSV-1) genome. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation. In some embodiments that may be combined with any of the preceding embodiments, the inactivating mutation is in a herpes simplex virus gene. In some embodiments, the inactivating mutation is a deletion of the coding sequence of the herpes simplex virus gene. In some embodiments, the herpes simplex virus gene is selected from Infected Cell Protein (ICP) 0, ICP4, ICP22, ICP27, ICP47, thymidine kinase (tk), Long Unique Region (UL) 41, and UL55. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in one or both copies of the ICP4 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP22 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the UL41 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in one or both copies of the ICP0 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP27 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the UL55 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the Joint region. In some embodiments, the recombinant herpes simplex virus genome comprises a deletion of the Joint region. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both of the ICP4 viral gene loci.

[0011] In some embodiments that may be combined with any of the preceding embodiments, the first cosmetic protein is selected from a first collagen protein, a first fibronectin protein, a first elastin protein, a first lumican protein, a first vitronectin protein, a first vitronectin receptor protein, a first laminin protein, a first neuromodulator protein, and a first

fibrillin protein. In some embodiments that may be combined with any of the preceding embodiments, the first cosmetic protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21 and 53-64. In some embodiments, the first cosmetic protein is a structural extracellular matrix protein (e.g., a collagen protein, an elastin protein, a fibronectin protein, a laminin protein, a fibrillin protein, etc.). In some embodiments, the first cosmetic protein is a collagen protein, an elastin protein, a fibronectin protein, or a laminin protein (e.g., a human collagen protein, a human elastin protein, a human fibronectin protein, or a human laminin protein). In some embodiments that may be combined with any of the preceding embodiments, the first collagen protein is a human collagen protein. In some embodiments that may be combined with any of the preceding embodiments, the first collagen protein is selected from a Collagen alpha-1(I) chain polypeptide (COL1-1), Collagen alpha-2(I) chain polypeptide (COL1-2), a Collagen alpha-1(II) chain polypeptide (COL2), a Collagen alpha-1(III) chain polypeptide (COL3), a Collagen alpha-1(IV) chain polypeptide (COL4-1), a Collagen alpha-2(IV) chain polypeptide (COL4-2), a Collagen alpha-3(IV) chain polypeptide (COL4-3), a Collagen alpha-4(IV) chain polypeptide (COL4-4), a Collagen alpha-5(IV) chain polypeptide (COL4-5), a Collagen alpha-6(IV) chain polypeptide (COL4-6), a Collagen alpha-1(V) chain polypeptide (COL5-1), a Collagen alpha-2(V) chain polypeptide (COL5-2), a Collagen alpha-3(V) chain polypeptide (COL5-3), a Collagen alpha-1(VI) chain polypeptide (COL6-1), a Collagen alpha-2(VI) chain polypeptide (COL6-2), a Collagen alpha-3(VI) chain polypeptide (COL6-3), a Collagen alpha-4(VI) chain polypeptide (COL6-4), a Collagen alpha-5(VI) chain polypeptide (COL6-5), a Collagen alpha-6(VI) chain polypeptide (COL6-6), a Collagen alpha-1(VIII) chain polypeptide (COL8), a Collagen alpha-1(IX) chain polypeptide (COL9-1), a Collagen alpha-2(IX) chain polypeptide (COL9-2), a Collagen alpha-3(IX) chain polypeptide (COL9-3), a Collagen alpha-1(X) chain polypeptide (COL10), a Collagen alpha-1(XI) chain polypeptide (COL11-1), a Collagen alpha-2(XI) chain polypeptide (COL11-2), a Collagen alpha-1(XII) chain polypeptide (COL12), a Collagen alpha-1(XIII) chain polypeptide (COL13), a Collagen alpha-1(XIV) chain polypeptide (COL14), a Collagen alpha-1(XV) chain polypeptide (COL15), a Collagen alpha-1(XVI) chain polypeptide (COL16), a Collagen alpha-1(XVII) chain polypeptide (COL17), a Collagen alpha-1(XVIII) chain polypeptide (COL18), a Collagen alpha-1(XIX) chain polypeptide (COL19), a Collagen alpha-1(XX) chain polypeptide (COL20), a Collagen alpha-1(XXI) chain polypeptide (COL21), a Collagen alpha-1(XXII) chain polypeptide (COL22), a Collagen alpha-1(XXIII) chain polypeptide (COL23), a Collagen alpha-1(XXIV) chain polypeptide (COL24), a Collagen alpha-1(XXV) chain polypeptide (COL25), a Collagen alpha-1(XXVI) chain polypeptide (COL26), a Collagen alpha-1(XXVII) chain polypeptide (COL27), and a Collagen alpha-1(XXVIII) chain polypeptide (COL28). In some embodiments that may be combined with any of the preceding embodiments, the first collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, and COL17. In some embodiments that

may be combined with any of the preceding embodiments, the first collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments that may be combined with any of the preceding embodiments, the first collagen protein is COL3. In some embodiments that may be combined with any of the preceding embodiments, the first collagen protein is human COL3. In some embodiments that may be combined with any of the preceding embodiments, the first collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the amino acid sequence of SEQ ID NO: 17. In some embodiments that may be combined with any of the preceding embodiments, the first cosmetic protein is not a Collagen alpha-1(VII) chain polypeptide (COL7).

[0012] In some embodiments, the first polypeptide consists essentially of the first cosmetic protein. In some embodiments, the first polypeptide consists of the first cosmetic protein. In some embodiments, the first polypeptide comprises: (a) the first cosmetic protein; (b) a further cosmetic protein; and (c) a linker polypeptide linking (a) to (b). In some embodiments, the further cosmetic protein is selected from a collagen protein, a fibronectin protein, a elastin protein, a lumican protein, a vitronectin protein, a vitronectin receptor protein, a laminin protein, a neuromodulator protein, and a fibrillin protein. In some embodiments, the further cosmetic protein is a structural extracellular matrix protein (e.g., a collagen protein, an elastin protein, a fibronectin protein, a laminin protein, a fibrillin protein, etc.). In some embodiments, the further cosmetic protein is a collagen protein, an elastin protein, a fibronectin protein, or a laminin protein (e.g., a human collagen protein, a human elastin protein, a human fibronectin protein, or a human laminin protein). In some embodiments, the further collagen protein (e.g., a further human collagen protein) is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the further collagen protein (e.g., a further human collagen protein) is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, COL7, and COL17. In some embodiments, the first cosmetic protein and the further cosmetic protein are different. In some embodiments, the first cosmetic protein is COL1-1 (e.g., human COL1-1) and the further cosmetic protein is COL1-2 (e.g., human COL1-2). In some embodiments, the linker polypeptide is a cleavable linker polypeptide. In some embodiments, the linker polypeptide comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected SEQ ID NOS: 28-31.

[0013] In some embodiments that may be combined with any of the preceding embodiments, the first polynucleotide

encodes a polycistronic mRNA comprising: (a) a first open reading frame (ORF) encoding the first polypeptide; (b) a second ORF encoding an additional cosmetic protein; and (c) an internal ribosomal entry site (IRES) separating (a) and (b). In some embodiments, the additional cosmetic protein is selected from a collagen protein, a fibronectin protein, a elastin protein, a lumican protein, a vitronectin protein, a vitronectin receptor protein, a laminin protein, a neuromodulator protein, and a fibrillin protein. In some embodiments, the additional cosmetic protein is a structural extracellular matrix protein (e.g., a collagen protein, an elastin protein, a fibronectin protein, a laminin protein, a fibrillin protein, etc.). In some embodiments, the additional cosmetic protein is a collagen protein, an elastin protein, a fibronectin protein, or a laminin protein (e.g., a human collagen protein, a human elastin protein, a human fibronectin protein, or a human laminin protein). In some embodiments, the additional collagen protein (e.g., an additional human collagen protein) is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the additional collagen protein (e.g., an additional human collagen protein) is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, COL7, and COL17. In some embodiments, the first cosmetic protein and the additional cosmetic protein are different. In some embodiments, the first cosmetic protein is COL1-1 (e.g., human COL1-1) and the additional cosmetic protein is COL1-2 (e.g., human COL1-2). In some embodiments, the nucleic acid sequence encoding the IRES has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NO: 22 or SEQ ID NO: 23.

[0014] In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes virus genome further comprises a second polynucleotide encoding a second cosmetic protein. In some embodiments, the second cosmetic protein is selected from a collagen protein, a fibronectin protein, a elastin protein, a lumican protein, a vitronectin protein, a vitronectin receptor protein, a laminin protein, a neuromodulator protein, and a fibrillin protein. In some embodiments, the second cosmetic protein is a structural extracellular matrix protein (e.g., a collagen protein, an elastin protein, a fibronectin protein, a laminin protein, a fibrillin protein, etc.). In some embodiments, the second cosmetic protein is a collagen protein, an elastin protein, a fibronectin protein, or a laminin protein (e.g., a human collagen protein, a human elastin protein, a human fibronectin protein, or a human laminin protein). In some embodiments, the second collagen protein (e.g., a second human collagen protein) is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the second

collagen protein (e.g., a second human collagen protein) is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, COL7, and COL17. In some embodiments, the first and second cosmetic proteins are different. In some embodiments, the first cosmetic protein is COL1-1 (e.g., human COL1-1) and the second cosmetic protein is COL1-2 (e.g., human COL1-2). In some embodiments, the first cosmetic protein is COL1-1 (e.g., human COL1-1) and the second cosmetic protein is COL3 (e.g., human COL3).

[0015] In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes virus genome has reduced cytotoxicity when introduced into a target cell, as compared to a corresponding wild-type herpes virus genome. In some embodiments, the target cell is a cell of the epidermis and/or dermis. In some embodiments, the target cell is a human cell. In some embodiments, the target cell is a fibroblast.

[0016] Other aspects of the present disclosure relate to a herpes virus comprising any of the recombinant herpes virus genomes described herein. In some embodiments, the herpes virus is replication competent. In some embodiments, the herpes virus is replication defective. In some embodiments, the herpes virus is attenuated. In some embodiments that may be combined with any of the preceding embodiments, the herpes virus has reduced cytotoxicity as compared to a corresponding wild-type herpes virus. In some embodiments that may be combined with any of the preceding embodiments, the herpes virus is selected from a herpes simplex virus, a varicella zoster virus, a human cytomegalovirus, a herpesvirus 6A, a herpesvirus 6B, a herpesvirus 7, and a Kaposi's sarcoma-associated herpesvirus. In some embodiments that may be combined with any of the preceding embodiments, the herpes virus is a herpes simplex virus. In some embodiments, the herpes simplex virus is a type 1 herpes simplex virus (HSV-1), a type 2 herpes simplex virus (HSV-2), or any derivatives thereof. In some embodiments, the herpes simplex virus is a type 1 herpes simplex virus (HSV-1).

[0017] Other aspects of the present disclosure relate to a composition comprising: (a) any of the recombinant herpes virus genomes described herein and/or any of the herpes viruses described herein; and (b) an excipient. In some embodiments, the composition is sterile. In some embodiments that may be combined with any of the preceding embodiments, the composition is suitable for topical, transdermal, subcutaneous, intradermal, oral, intranasal, intratracheal, sublingual, buccal, rectal, vaginal, inhaled, intravenous, intraarterial, intramuscular, intracardiac, intraosseous, intraperitoneal, transmucosal, intravitreal, subretinal, intraarticular, peri-articular, local, or epicutaneous administration. In some embodiments that may be combined with any of the preceding embodiments, the composition is suitable for intradermal administration. In some embodiments that may be combined with any of the preceding embodiments, the composition is suitable for superficial injection. In some embodiments that may be combined with any of the preceding embodiments, the composition is a cosmetic composition. In some embodiments that may be combined with any of the preceding embodiments, the composition is a skin care product.

[0018] Other aspects of the present disclosure relate to the use of any of the recombinant herpes virus genomes described herein and/or any of the herpes viruses described herein as a medicament (e.g., for an aesthetic indication).

[0019] Other aspects of the present disclosure relate to the use of any of the recombinant herpes virus genomes described herein and/or any of the herpes viruses described herein as a therapy (e.g., as an aesthetic or cosmetic therapy).

[0020] Other aspects of the present disclosure relate to the use of any of the recombinant herpes virus genomes described herein and/or any of the herpes viruses described herein in the manufacture of a medicament useful for treating one or more signs or symptoms of dermatological aging.

[0021] Other aspects of the present disclosure relate to a method of enhancing, increasing, augmenting, and/or supplementing the levels of one or more dermal extracellular matrix proteins in a subject, the method comprising administering to the subject an effective amount of any of the herpes viruses described herein and/or any of the compositions described herein.

[0022] Other aspects of the present disclosure relate to a method of enhancing, increasing, augmenting, and/or supplementing the levels of one or more collagen proteins in a subject, the method comprising administering to the subject an effective amount of any of the herpes viruses described herein and/or any of the compositions described herein. In some embodiments, the one or more collagen proteins are collagen 3. In some embodiments, the levels of endogenous collagen 3 are reduced as a result of chronological or photo-aging.

[0023] Other aspects of the present disclosure relate to a method of enhancing, increasing, augmenting, and/or supplementing the soft tissue of a subject, the method comprising administering to the subject an effective amount of any of the herpes viruses described herein and/or any of the compositions described herein. In some embodiments, the composition is injected into the soft tissue of the subject.

[0024] Other aspects of the present disclosure relate to a method of improving skin condition, quality, and/or appearance in a subject in need thereof, the method comprising administering to the subject an effective amount of any of the herpes viruses described herein and/or any of the compositions described herein. In some embodiments, the composition is administered to one or more sites of sun damage or other UV exposure, rough texture, skin sagging, wrinkles, or any combinations thereof.

[0025] Other aspects of the present disclosure relate to a method of reducing the appearance of one or more superficial depressions in the skin of a subject in need thereof, the method comprising administering to the subject an effective amount of any of the herpes viruses described herein and/or any of the compositions described herein. In some embodiments, the one or more superficial depressions in the skin are selected from the group consisting of nasolabial folds, crows' feet, frown lines, worry lines, scars, glabellar lines, brow ptosis, tear troughs, nasojugal lines, bunny lines, cheek/mid-face ptosis, marionette lines, poppy dimpling, smile lines, laugh lines, chin creases, neck lines, platysma bands, and any combinations thereof.

[0026] Other aspects of the present disclosure relate to a method of increasing and/or improving at least one of texture, smoothness, elasticity, or tension of the skin of a subject in need thereof, the method comprising administering to the subject an effective amount of any of the herpes viruses described herein and/or any of the compositions described herein.

[0027] In some embodiments that may be combined with any of the preceding embodiments, the skin of the subject is aging skin. In some embodiments that may be combined with any of the preceding embodiments, the skin of the subject has been damaged due to exposure to ultraviolet light. In some embodiments that may be combined with any of the preceding embodiments, the skin of the subject is wrinkled.

[0028] Other aspects of the present disclosure relate to a method of diminishing one or more dermatological signs of aging in a subject in need thereof, the method comprising administering to the subject an effective amount of any of the herpes viruses described herein and/or any of the compositions described herein. In some embodiments, the diminishing of one or more dermatological signs of aging is indicated by the: (a) treatment, reduction, and/or prevention of fine lines and/or wrinkles; (b) reduction of skin pore size; (c) improvement in skin thickness, plumpness, and/or tautness; (d) improvement in skin smoothness, suppleness, and/or softness; (e) improvement in skin tone, radiance, and/or clarity; (f) improvement in procollagen and/or collagen production; (g) improvement in skin texture and/or promotion of retexturization; (h) improvement in appearance of skin contours; (i) restoration of skin luster and/or brightness; (j) improvement of skin appearance decreased by aging and/or menopause; (k) improvement in skin moisturization; (l) increase in skin elasticity and/or resiliency; (m) treatment, reduction, and/or prevention of skin sagging; (n) improvement in skin firmness; (o) reduction of pigment spots, mottled skin, and/or scars (such as acne scars); (p) improvement of optical properties of skin by light diffraction or reflection; or (q) any combinations thereof.

[0029] In some embodiments that may be combined with any of the preceding embodiments, the subject is a human. In some embodiments that may be combined with any of the preceding embodiments, the herpes virus or composition is administered topically, transdermally, subcutaneously, epicutaneously, intradermally, orally, sublingually, buccally, rectally, vaginally, intravenously, intraarterially, intramuscularly, intraosseously, intracardially, intraperitoneally, transmucosally, intravitreally, subretinally, intraarticularly, peripherally, locally, or via inhalation to the subject. In some embodiments that may be combined with any of the preceding embodiments, the herpes virus or composition is administered intradermally to the subject. In some embodiments that may be combined with any of the preceding embodiments, the herpes virus or composition is administered by superficial injection.

[0030] Other aspects of the present disclosure relate to a composition comprising: a herpes simplex virus (HSV) comprising a recombinant nucleic acid, wherein the recombinant nucleic acid comprises a first polynucleotide encoding a first polypeptide comprising a first human collagen protein, and an excipient. In some embodiments, the recombinant nucleic acid comprises two or more copies of the first polynucleotide. In some embodiments that may be combined with any of the preceding embodiments, the HSV is replication-defective. In some embodiments that may be combined with any of the preceding embodiments, the HSV is replication-competent. In some embodiments that may be combined with any of the preceding embodiments, the HSV is a herpes simplex type 1 virus, a herpes simplex type 2 virus, or any derivatives thereof.

[0031] In some embodiments, the recombinant nucleic acid is a herpes simplex virus amplicon. In some embodiments, the herpes simplex virus amplicon is an HSV-1 amplicon or an HSV-1 hybrid amplicon. In some embodiments, the HSV-1 hybrid amplicon is an HSV/AAV hybrid amplicon, an HSV/EBV hybrid amplicon, and HSV/EBV/RV hybrid amplicon, or an HSV/Sleeping Beauty hybrid amplicon.

[0032] In some embodiments, the recombinant nucleic acid is a recombinant herpes simplex virus genome. In some embodiments, the recombinant herpes simplex virus genome is a recombinant HSV-1 genome, a recombinant HSV-2 genome, or any derivatives thereof. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in a herpes simplex virus gene. In some embodiments, the herpes simplex virus gene is selected from the group consisting of Infected Cell Protein (ICP) 0, ICP4, ICP22, ICP27, ICP47, thymidine kinase (tk), Long Unique Region (UL) 41, and UL55. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in one or both copies of the ICP4 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP22 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in the UL41 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP0 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP27 gene. In some embodiments that may be combined with any of the preceding embodiments, the inactivating mutation is a deletion of the coding sequence of the gene(s).

[0033] In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within a viral gene locus. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within the ICP22 viral gene locus. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within the UL41 viral gene locus. In some embodiments that may be combined with any of the preceding embodiments, the HSV has reduced cytotoxicity as compared to a wild-type herpes simplex virus.

[0034] In some embodiments that may be combined with any of the preceding embodiments, the first human collagen protein is selected from Collagen alpha-1(I) chain polypeptide (COL1-1), Collagen alpha-2(I) chain polypeptide (COL1-2), a Collagen alpha-1(II) chain polypeptide (COL2), a Collagen alpha-1(III) chain polypeptide (COL3), a Collagen alpha-1(IV) chain polypeptide (COL4-1), a Col-

lagen alpha-2(IV) chain polypeptide (COL4-2), a Collagen alpha-3(IV) chain polypeptide (COL4-3), a Collagen alpha-4(IV) chain polypeptide (COL4-4), a Collagen alpha-5(IV) chain polypeptide (COL4-5), a Collagen alpha-6(IV) chain polypeptide (COL4-6), a Collagen alpha-1(V) chain polypeptide (COL5-1), a Collagen alpha-2(V) chain polypeptide (COL5-2), a Collagen alpha-3(V) chain polypeptide (COL5-3), a Collagen alpha-1(VI) chain polypeptide (COL6-1), a Collagen alpha-2(VI) chain polypeptide (COL6-2), a Collagen alpha-3(VI) chain polypeptide (COL6-3), a Collagen alpha-4(VI) chain polypeptide (COL6-4), a Collagen alpha-5(VI) chain polypeptide (COL6-5), a Collagen alpha-6(VI) chain polypeptide (COL6-6), a Collagen alpha-1(VII) chain polypeptide (COL7), a Collagen alpha-1(VIII) chain polypeptide (COL8), a Collagen alpha-1(IX) chain polypeptide (COL9-1), a Collagen alpha-2(IX) chain polypeptide (COL9-2), a Collagen alpha-3(IX) chain polypeptide (COL9-3), a Collagen alpha-1(X) chain polypeptide (COL10), a Collagen alpha-1(XI) chain polypeptide (COL11-1), a Collagen alpha-2(XI) chain polypeptide (COL11-2), a Collagen alpha-1(XII) chain polypeptide (COL12), a Collagen alpha-1(XIII) chain polypeptide (COL13), a Collagen alpha-1(XIV) chain polypeptide (COL14), a Collagen alpha-1(XV) chain polypeptide (COL15), a Collagen alpha-1(XVI) chain polypeptide (COL16), a Collagen alpha-1(XVII) chain polypeptide (COL17), a Collagen alpha-1(XVIII) chain polypeptide (COL18), a Collagen alpha-1(XIX) chain polypeptide (COL19), a Collagen alpha-1(XX) chain polypeptide (COL20), a Collagen alpha-1(XXI) chain polypeptide (COL21), a Collagen alpha-1(XXII) chain polypeptide (COL22), a Collagen alpha-1(XXIII) chain polypeptide (COL23), a Collagen alpha-1(XXIV) chain polypeptide (COL24), a Collagen alpha-1(XXV) chain polypeptide (COL25), a Collagen alpha-1(XXVI) chain polypeptide (COL26), a Collagen alpha-1(XXVII) chain polypeptide (COL27), and a Collagen alpha-1(XXVIII) chain polypeptide (COL28). In some embodiments that may be combined with any of the preceding embodiments, the first human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17. In some embodiments that may be combined with any of the preceding embodiments, the nucleic acid sequence encoding the first human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments that may be combined with any of the preceding embodiments, the first human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments that may be combined with any of the preceding embodiments, the first human collagen protein is not COL7.

[0035] In some embodiments that may be combined with any of the preceding embodiments, the first polypeptide comprises: (a) the first human collagen protein; (b) a further human collagen protein; and (c) a linker polypeptide linking (a) to (b). In some embodiments, the linker polypeptide is a cleavable linker polypeptide. In some embodiments, the linker polypeptide comprises a sequence having at least

80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 28-31. In some embodiments, the further human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the further human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17. In some embodiments, the nucleic acid sequence encoding the further human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments, the further human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments, the first human collagen protein and the further human collagen protein are different.

[0036] In some embodiments that may be combined with any of the preceding embodiments, the first polynucleotide encodes a polycistronic mRNA comprising: (a) a first open reading frame (ORF) encoding the first polypeptide; (b) a second ORF encoding an additional human collagen protein; and (c) an internal ribosomal entry site (IRES) separating (a) and (b). In some embodiments, the nucleic acid sequence encoding the IRES has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NO: 22 or SEQ ID NO: 23. In some embodiments, the additional human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the additional human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17. In some embodiments, the nucleic acid sequence encoding the additional human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments, the additional human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments,

the first human collagen protein and the additional human collagen protein are different.

[0037] In some embodiments that may be combined with any of the preceding embodiments, the recombinant nucleic acid further comprises a second polynucleotide encoding a second human collagen protein. In some embodiments, the recombinant nucleic acid comprises two or more copies of the second polynucleotide. In some embodiments, the second human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the second human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17. In some embodiments, the nucleic acid sequence encoding the second human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments, the second human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments, the first and second human collagen proteins are different.

[0038] In some embodiments that may be combined with any of the preceding embodiments, the recombinant nucleic acid is a recombinant herpes simplex virus genome, and wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within a viral gene locus. In some embodiments, the recombinant herpes simplex virus genome comprises the second polynucleotide within one or both copies of the ICP4 viral gene loci. In some embodiments, the recombinant herpes simplex virus genome comprises the second polynucleotide within the ICP22 viral gene locus. In some embodiments, the recombinant herpes simplex virus genome comprises the second polynucleotide within the UL41 viral gene locus. In some embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci and the second polynucleotide within the ICP22 viral gene locus. In some embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci and the second polynucleotide within the UL41 viral gene locus.

[0039] In some embodiments that may be combined with any of the preceding embodiments, the excipient is adapted for cutaneous (systemic or topical), transdermal, subcutaneous, and/or intradermal administration. In some embodiments that may be combined with any of the preceding embodiments, the excipient comprises a hydroxypropyl methylcellulose gel. In some embodiments that may be combined with any of the preceding embodiments, the excipient is adapted for intradermal administration. In some embodiments that may be combined with any of the preceding embodiments, the excipient comprises a phosphate

buffer. In some embodiments that may be combined with any of the preceding embodiments, the excipient comprises glycerol. In some embodiments that may be combined with any of the preceding embodiments, the excipient comprises a lipid carrier. In some embodiments that may be combined with any of the preceding embodiments, the excipient comprises a nanoparticle carrier.

[0040] In some embodiments that may be combined with any of the preceding embodiments, the composition is a cosmetic composition. In some embodiments, the cosmetic composition is a skin care product.

[0041] Other aspects of the present disclosure relate to a kit comprising any of the compositions described herein and instructions for administering the composition.

[0042] Other aspects of the present disclosure relate to a method of enhancing, increasing, augmenting, and/or supplementing the levels of one or more human collagen proteins in a subject, the method comprising administering to the subject an effective amount of any of the compositions described herein.

[0043] Other aspects of the present disclosure relate to a method of enhancing, increasing, augmenting, and/or supplementing soft tissue of a subject, the method comprising administering to the subject an effective amount of any of the compositions described herein. In some embodiments, the composition is injected into a soft tissue of the subject.

[0044] Other aspects of the present disclosure relate to a method of improving skin quality, condition and/or appearance in a subject in need thereof, the method comprising administering to the subject an effective amount of any of the compositions described herein. In some embodiments, the condition is selected from sun damage, aging, UV exposure, rough texture, skin sagging, wrinkles, and any combinations thereof.

[0045] Other aspects of the present disclosure relate to a method of reducing the appearance of one or more superficial depressions in the skin of a subject in need thereof, the method comprising administering to the subject an effective amount of any of the compositions described herein. In some embodiments, administration of the composition reduces the appearance of the one or more superficial depressions in the skin of the subject for at least about three months, at least about six months, at least about nine months, or at least about 12 months. In some embodiments, the appearance of the one or more superficial depressions in the skin of the subject is reduced after administration of the composition, as compared to the appearance of the one or more superficial depression in the skin of the subject prior to administration of the composition.

[0046] Other aspects of the present disclosure relate to a method of increasing and/or improving at least one of texture, smoothness, elasticity, or tension of the skin of a subject in need thereof, the method comprising administering to the subject an effective amount of any of the compositions described herein. In some embodiments, the skin of the subject maintains at least one of an increased and/or improved texture, smoothness, elasticity, or tension for at least about three months, at least about six months, at least about nine months, or at least about 12 months after administration of the composition. In some embodiments, at least one of texture, smoothness, elasticity, or tension of the skin of the subject is increased and/or improved after administration of the composition, as compared to the texture,

smoothness, elasticity, or tension of the skin of the subject prior to administration of the composition.

[0047] In some embodiments that may be combined with any of the preceding embodiments, the skin of the subject is aging skin. In some embodiments that may be combined with any of the preceding embodiments, the skin of the subject has been damaged due to exposure to ultraviolet light. In some embodiments that may be combined with any of the preceding embodiments, the skin of the subject is wrinkled.

[0048] Other aspects of the present disclosure relate to a method of diminishing one or more dermatological signs of aging in a subject in need thereof, the method comprising administering to the subject an effective amount of any of the compositions described herein. In some embodiments, the diminishing of one or more dermatological signs of aging is selected from: (a) treatment, reduction, and/or prevention of fine lines and/or wrinkles; (b) reduction of skin pore size; (c) improvement in skin thickness, plumpness, and/or tautness; (d) improvement in skin smoothness, suppleness, and/or softness; (e) improvement in skin tone, radiance, and/or clarity; (f) improvement in procollagen and/or collagen production; (g) improvement in skin texture and or promotion of retexturization; (h) improvement in appearance of skin contours; (i) restoration of skin luster and/or brightness; (j) improvement of skin appearance decreased by aging and/or menopause; (k) improvement in skin moisturization; (l) increase in skin elasticity and/or resiliency; (m) treatment, reduction, and/or prevention or skin sagging; (n) improvement in skin firmness; (o) reduction of pigment spots, mottled skin, and/or acne scars; (p) improvement of optical properties of skin by light diffraction or reflection; and (q) any combinations thereof. In some embodiments, the one or more dermatological signs of aging in the subject is diminished after administration of the composition, as compared to the one or more dermatological signs of aging in the subject prior to administration of the composition.

[0049] In some embodiments that may be combined with any of the preceding embodiments, the subject is a human. In some embodiments that may be combined with any of the preceding embodiments, the composition is administered cutaneously (systemically or topically), transdermally, subcutaneously, or intradermally to the subject. In some embodiments, the composition is administered by superficial injection. In some embodiments, the composition is administered intradermally to the subject. In some embodiments, the composition is administered once to the subject. In some embodiments, the composition is administered at least twice to the subject. In some embodiments, at least about 15, at least about 30, at least about 60, at least about 90, or at least about 120 days passes between administrations. In some embodiments that may be combined with any of the preceding embodiments, the composition is administered to one or more affected and/or unaffected areas of the subject. In some embodiments that may be combined with any of the preceding embodiments, the skin of the is abraded prior to administration.

[0050] Other aspects of the present disclosure relate to a recombinant nucleic acid comprising a first polynucleotide encoding a first polypeptide comprising a first human collagen protein, wherein the recombinant nucleic acid is a recombinant herpes simplex virus genome. In some embodiments, the recombinant nucleic acid comprises two or more

copies of the first polynucleotide. In some embodiments, the recombinant herpes simplex virus genome is a recombinant HSV-1 genome, a recombinant HSV-2 genome, or any derivatives thereof.

[0051] In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in a herpes simplex virus gene. In some embodiments, the herpes simplex virus gene is selected from the group consisting of Infected Cell Protein (ICP) 0, ICP4, ICP22, ICP27, ICP47, thymidine kinase (tk), Long Unique Region (UL) 41, and UL55. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in one or both copies of the ICP4 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP22 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in the UL41 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP0 gene. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP27 gene. In some embodiments that may be combined with any of the preceding embodiments, the inactivating mutation is a deletion of the coding sequence of the gene(s).

[0052] In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within a viral gene locus. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within the ICP22 viral gene locus. In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within the UL41 viral gene locus. In some embodiments that may be combined with any of the preceding embodiments, the HSV has reduced cytotoxicity as compared to a wild-type herpes simplex virus.

[0053] In some embodiments that may be combined with any of the preceding embodiments, the first human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL2. In some embodiments that may be combined with any of the preceding embodiments, the first human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17. In some embodiments that may be combined with any of the preceding embodiments, the nucleic acid sequence encoding the first

human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments that may be combined with any of the preceding embodiments, the first human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments that may be combined with any of the preceding embodiments, the first human collagen protein is not COL7.

[0054] In some embodiments that may be combined with any of the preceding embodiments, the first polypeptide comprises: (a) the first human collagen protein; (b) a further human collagen protein; and (c) a linker polypeptide linking (a) to (b). In some embodiments, the linker polypeptide is a cleavable linker polypeptide. In some embodiments, the linker polypeptide comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 28-31. In some embodiments, the further human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the further human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17. In some embodiments, the nucleic acid sequence encoding the further human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments, the further human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments, the first human collagen protein and the further human collagen protein are different.

[0055] In some embodiments that may be combined with any of the preceding embodiments, the first polynucleotide encodes a polycistronic mRNA comprising: (a) a first open reading frame (ORF) encoding the first polypeptide; (b) a second ORF encoding an additional human collagen protein; and (c) an internal ribosomal entry site (IRES) separating (a) and (b). In some embodiments, the nucleic acid sequence encoding the IRES has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NO: 22 or SEQ ID NO: 23. In some embodiments, the additional human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1,

COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the additional human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17. In some embodiments, the nucleic acid sequence encoding the additional human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments, the additional human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments, the first human collagen protein and the additional human collagen protein are different.

[0056] In some embodiments that may be combined with any of the preceding embodiments, the recombinant nucleic acid further comprises a second polynucleotide encoding a second human collagen protein. In some embodiments, the recombinant nucleic acid comprises two or more copies of the second polynucleotide. In some embodiments, the second human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28. In some embodiments, the second human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17. In some embodiments, the nucleic acid sequence encoding the second human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments, the second human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments, the first and second human collagen proteins are different.

[0057] In some embodiments that may be combined with any of the preceding embodiments, the recombinant herpes simplex virus genome comprises the second polynucleotide within a viral gene locus. In some embodiments, the recombinant herpes simplex virus genome comprises the second polynucleotide within one or both copies of the ICP4 viral gene loci. In some embodiments, the recombinant herpes simplex virus genome comprises the second polynucleotide within the ICP22 viral gene locus. In some embodiments, the recombinant herpes simplex virus genome comprises the second polynucleotide within the UL41 viral gene locus. In

some embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci and the second polynucleotide within the ICP22 viral gene locus. In some embodiments, the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci and the second polynucleotide within the UL41 viral gene locus.

[0058] Other aspects of the present disclosure relate to a host cell comprising any of the recombinant nucleic acids described herein. In some embodiments, the host cell is a eukaryotic cell. In some embodiments, the host cell is a mammalian cell. In some embodiments, the host cell is a human cell or a non-human primate cell. In some embodiments, the host cell is a Vero cell. In some embodiments, the host cell is a complementing host cell.

[0059] Other aspects of the present disclosure relate to a method of collecting a herpes simplex virus, the method comprising: (a) contacting a complementing host cell with any of the recombinant nucleic acids described herein; and (b) collecting the herpes simplex virus generated by the complementing host cell.

[0060] Other aspects of the present disclosure relate to a method of collecting a herpes simplex virus, the method comprising: (a) culturing a host cell comprising any of the recombinant nucleic acids described herein; and (b) collecting the herpes simplex virus generated by the host cell.

BRIEF DESCRIPTION OF THE DRAWINGS

[0061] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings will be provided by the Office upon request and payment of the necessary fee.

[0062] FIGS. 1A-N show schematics of wild-type and modified herpes simplex virus genomes. FIG. 1A shows a wild-type herpes simplex virus genome. FIG. 1B shows a modified herpes simplex virus genome comprising deletions of the coding sequences of ICP4 (both copies) and ICP22, with a polynucleotide containing the coding sequence of a first human collagen polypeptide operably linked to a heterologous promoter integrated at each of the ICP4 loci. FIG. 1C shows a modified herpes simplex virus genome comprising deletions of the coding sequence of ICP4 (both copies), with a polynucleotide containing the coding sequence of a first human collagen polypeptide operably linked to a heterologous promoter integrated at each of the ICP4 loci. FIG. 1D shows a modified herpes simplex virus genome comprising deletions of the coding sequences of ICP4 (both copies) and ICP22, with a polynucleotide containing 1) the coding sequence of a first human collagen polypeptide operably linked to a first heterologous promoter, and 2) the coding sequence of a second human collagen polypeptide operably linked to a second heterologous promoter, integrated at each of the ICP4 loci. Both the first and second human collagen polypeptides are encoded on the same strand of DNA. FIG. 1E shows a modified herpes simplex virus genome comprising deletions of the coding sequence of ICP4 (both copies), with a polynucleotide containing 1) the coding sequence of a first human collagen polypeptide operably linked to a first heterologous promoter, and 2) the coding sequence of a second human collagen polypeptide operably linked to a second heterologous promoter, integrated at each of the ICP4 loci. Both the first and second human collagen polypeptides are encoded on the

same strand of DNA. FIG. 1F shows a modified herpes simplex virus genome comprising deletions of the coding sequences of ICP4 (both copies) and ICP22, with a polynucleotide containing 1) the coding sequence of a first human collagen polypeptide operably linked to a first heterologous promoter, and 2) the coding sequence of a second human collagen polypeptide operably linked to a second heterologous promoter, integrated at each of the ICP4 loci. The first and second human collagen polypeptides are encoded on opposite strands of DNA. FIG. 1G shows a modified herpes simplex virus genome comprising deletions of the coding sequence of ICP4 (both copies), with a polynucleotide containing 1) the coding sequence of a first human collagen polypeptide operably linked to a first heterologous promoter, and 2) the coding sequence of a second human collagen polypeptide operably linked to a second heterologous promoter, integrated at each of the ICP4 loci. The first and second human collagen polypeptides are encoded on opposite strands of DNA. FIG. 1H shows a modified herpes simplex virus genome comprising deletions of the coding sequences of ICP4 (both copies) and ICP22, with a polynucleotide encoding a polycistronic mRNA operably linked to a heterologous promoter integrated at each of the ICP4 loci. The polycistronic mRNA contains the coding sequence of a first human collagen polypeptide and a second human collagen polypeptide separated by an internal ribosomal entry site (IRES). FIG. 1I shows a modified herpes simplex virus genome comprising deletions of the coding sequence of ICP4 (both copies), with a polynucleotide encoding a polycistronic mRNA operably linked to a heterologous promoter integrated at each of the ICP4 loci. The polycistronic mRNA contains the coding sequence of a first human collagen polypeptide and a second human collagen polypeptide separated by an internal ribosomal entry site (IRES). FIG. 1J shows a modified herpes simplex virus genome comprising deletions of the coding sequences of ICP4 (both copies) and ICP22, with a polynucleotide containing the coding sequence of a chimeric polypeptide operably linked to a heterologous promoter integrated at each of the ICP4 loci. The chimeric polypeptide comprises the amino acid sequence of a first human collagen polypeptide and second human collagen polypeptide separated by a cleavable linker. FIG. 1K shows a modified herpes simplex virus genome comprising deletions of the coding sequence of ICP4 (both copies), with a polynucleotide containing the coding sequence of a chimeric polypeptide operably linked to a heterologous promoter integrated at each of the ICP4 loci. The chimeric polypeptide comprises the amino acid sequence of a first human collagen polypeptide and second human collagen polypeptide separated by a cleavable linker. FIG. 1L shows a modified herpes simplex virus genome comprising deletions of the coding sequences of ICP4 (both copies) and ICP22, with a first polynucleotide containing the coding sequence of a first human collagen polypeptide operably linked to a heterologous promoter integrated at each of the ICP4 loci, and a second polynucleotide containing the coding sequence of a second human collagen polypeptide operably linked to a heterologous promoter integrated at the ICP22 locus. FIG. 1M shows a modified herpes simplex virus genome comprising deletions of the coding sequences of ICP4 (both copies), ICP22, and UL41, with a first polynucleotide containing the coding sequence of a first human collagen polypeptide operably linked to a heterologous promoter integrated at each of the ICP4 loci, and a

second polynucleotide containing the coding sequence of a second human collagen polypeptide operably linked to a heterologous promoter integrated at the UL41 locus. FIG. 1N shows a modified herpes simplex virus genome comprising deletions of the coding sequences of ICP4 (both copies) and UL41, with a first polynucleotide containing the coding sequence of a first human collagen polypeptide operably linked to a heterologous promoter integrated at each of the ICP4 loci, and a second polynucleotide containing the coding sequence of a second human collagen polypeptide operably linked to a heterologous promoter integrated at the UL41 locus.

[0063] FIGS. 2A-B show schematics of replication-defective herpes simplex type-1 viruses carrying human collagen 7 (COL7) expression cassettes. FIG. 2A shows a schematic of the virus "KCA211". FIG. 2B shows a schematic of the virus "SAR-COL7".

[0064] FIGS. 3A-B show human COL7 expression in HaCaT cells infected with KCA211 or SAR-COL7 at the indicated MOIs. FIG. 3A shows human COL7 expression in HaCaT cells infected with KCA211 or SAR-COL7 at the indicated MOIs, as assessed by qPCR. Data is shown as fold change relative to SAR-COL7, after normalization to GAPDH. FIG. 3B shows human COL7 expression in uninfected HaCaT cells, or HaCaT cells infected with KCA211 or SAR-COL7 at the indicated MOIs, as assessed by western blot analysis.

[0065] FIGS. 4A-B show immunofluorescence images of human COL7 expression in mock infected primary human cells isolated from a healthy patient (Normal), and mock or SAR-COL7 infected primary human cells isolated from a patient suffering from recessive dystrophic epidermolysis bullosa (RDEB). FIG. 4A shows human COL7 expression in mock infected wild-type and RDEB primary human keratinocytes, or in RDEB primary human keratinocytes infected with SAR-COL7 at the indicated multiplicity of infections (MOIs). FIG. 4B shows human COL7 expression in mock infected wild-type and RDEB primary human fibroblasts, or in RDEB primary human fibroblasts infected with SAR-COL7 at the indicated MOIs.

[0066] FIGS. 5A-B show quantitative PCR analysis of human COL7 expression in mock infected primary human cells isolated from a healthy patient, and mock or SAR-COL7 infected primary human cells isolated from a patient suffering from recessive dystrophic epidermolysis bullosa (EB). FIG. 5A shows human COL7 expression in mock infected wild-type (N-HDK) and RDEB (EB-HDK) primary human keratinocytes, or in RDEB primary human keratinocytes infected with SAR-COL7 at the indicated MOIs. COL7 expression is shown as the relative fold change over mock infected wild-type primary human keratinocytes. FIG. 5B shows human COL7 expression in mock infected wild-type (N-HDF) and RDEB (EB-HDF) primary human fibroblasts, or in RDEB primary human fibroblasts infected with SAR-COL7 at the indicated MOIs. COL7 expression is shown as the relative fold change over mock infected wild-type primary human fibroblasts.

[0067] FIGS. 6A-B show cellular adhesion of uninfected (control) or SAR-COL7 infected RDEB primary human keratinocytes to untreated (plastic) or treated wells of a microwell plate. FIG. 6A shows cellular adhesion to untreated wells (plastic), or wells treated with increasing concentrations of rat tail Collagen 1. FIG. 6B shows cellular

adhesion to untreated wells (plastic), or wells treated with increasing concentrations of human plasma fibronectin.

[0068] FIG. 7 show representative immunofluorescence images of human COL7 expression and deposition at the basement membrane zone (BMZ) at day 5 in organotypic cultures constructed with SAR-COL7 infected RDEB primary human keratinocytes and fibroblasts. Both keratinocytes and fibroblasts were infected *in situ* at the indicated MOI after culture construction.

[0069] FIGS. 8A-D show human COL7A1 transcript and genome levels observed in uninfected mouse skin (control), or in mouse skin after topical or intradermal delivery of SAR-COL7, as assessed by qPCR. Error bars represent SEM. FIG. 8A shows human COL7A1 transcripts levels/100 ng total RNA in mouse skin at day 3 after infection. FIG. 8B shows copy number of human COL7A1 DNA/100 ng total DNA in mouse skin at day 3 after infection. FIG. 8C shows human COL7A1 transcripts levels/100 ng total RNA in mouse skin at day 6 after infection. FIG. 8D shows copy number of human COL7A1 DNA/100 ng total DNA in mouse skin at day 6 after infection.

[0070] FIGS. 9A-B show representative immunofluorescence images of human COL7 expression in mouse skin after delivery of SAR-COL7. FIG. 9A shows a representative immunofluorescent image of human COL7 expression in mouse skin after intradermal delivery of SAR-COL7. FIG. 9B shows a representative immunofluorescent image of human COL7 expression in mouse skin after topical delivery of SAR-COL7.

[0071] FIGS. 10A-B show human COL7A1 transcript and genome levels observed in BALB/c mouse skin after intradermal delivery of vehicle, SAR-COL7, or KCA211, as assessed by qPCR. FIG. 10A shows human COL7A1 transcripts levels/100 ng total RNA in BALB/c mouse skin. FIG. 10B shows copy number of human COL7A1 DNA/100 ng total DNA in BALB/c mouse skin.

[0072] FIGS. 11A-B show human COL7A1 transcript and genome levels observed at each injection site in hypomorph mouse skin after high-dose intradermal delivery of HSV-GFP (GFP ctrl) or SAR-COL7, as assessed by qPCR. Each bar represents a single sample at the indicated time point. FIG. 11A shows human COL7A1 transcripts levels/100 ng total RNA in hypomorph mouse skin. FIG. 11B shows copy number of human COL7A1 DNA/100 ng total DNA in hypomorph mouse skin.

[0073] FIGS. 12A-B show representative immunofluorescence images of human COL7 expression in hypomorph mouse skin after high-dose intradermal delivery of HSV-GFP (GFP Control) or SAR-COL7. FIG. 12A shows control (GFP) and SAR-COL7 immunofluorescence imaging from hypomorph mouse 1 (harvested at day 3) at 10 and 20 \times magnification. FIG. 12B shows SAR-COL7 immunofluorescence imaging from hypomorph mouse 2 and hypomorph mouse 3 (harvested at day 7). The figure represents a tiled image of 16 fields acquired with a 10 \times lens, capturing the entire skin section.

[0074] FIG. 13 shows H&E stained samples from hypomorph mouse 1, 2, and 3 (harvested at days 3 and 3). The samples were taken from untreated hypomorph mouse skin, and hypomorph mouse skin after intradermal delivery of HSV-GFP or SAR-COL7.

[0075] FIGS. 14A-B show representative electron micrograph images of human COL7 expression in hypomorph mouse skin after intradermal delivery of SAR-COL7. The

lamina densa is the dark band indicated through the middle of the images; the black dots are the stained NC domains of human COL7; the blue arrows indicate the formation of anchoring fibrils. FIG. 14A shows electron micrograph images of infected hypomorph mouse skin stained with an antibody specific to the NC2 domain of human COL7 (LH24). FIG. 14B shows electron micrograph images of infected hypomorph mouse skin stained with an antibody specific to the NC1 domain of human COL7 (NP185).

[0076] FIGS. 15A-B show human COL7A1 transcript and genome levels observed at each injection site in hypomorph mouse skin after low-dose intradermal delivery of SAR-COL7, as assessed by qPCR. Each bar represents a single sample at the indicated time point. FIG. 15A shows human COL7A1 transcripts levels/100 ng total RNA in hypomorph mouse skin. FIG. 15B shows copy number of human COL7A1 DNA/100 ng total DNA in hypomorph mouse skin.

[0077] FIG. 16 shows representative immunofluorescence images of human COL7 expression in hypomorph mouse skin (from mouse 1) after low-dose intradermal delivery of SAR-COL7.

[0078] FIGS. 17A-C show human COL1A1 and COL1A2 nucleic acid and protein analyses in Vero cells infected with the indicated clones of HSV encoding COL1A1 alone (inserted into the ICP4 loci) or COL1A1 and COL1A2 (inserted into the ICP4 and ICP22 loci, respectively). FIG. 17A shows the levels of human COL1A1 transcripts present in Vero cells 5 days after infection with the indicated HSV clones, as determined by qRT-PCR analysis. Data is presented for two replicates \pm SEM. FIG. 17B shows the levels of human COL1A2 transcripts present in Vero cells 5 days after infection with the indicated HSV clones, as determined by qRT-PCR analysis. Data is presented for two replicates \pm SEM. FIG. 17C shows western blot analysis of human COL1A1 and COL1A2 protein expression in Vero cells 5 days after infection with the indicated COL1A1/COL1A2 positive clones, as determined by qRT-PCR. Uninfected (mock) Vero cells were used as a negative control; GAPDH was used as a loading control.

[0079] FIG. 18 shows western blot analysis of human COL1A1 and COL1A2 protein expression in Vero cells 5 days after infection with an HSV isolate encoding a COL1A1-IRES-COL1A2 sequence (IRES-Isolate 6) inserted into the ICP4 loci. Infection with an isolate that does not contain the IRES construct (no insertion) was used as a negative control; GAPDH was used as a loading control.

[0080] FIGS. 19A-B show human COL3 nucleic acid and protein analyses in immortalized human keratinocytes (Ha-CaTs) infected with C3vec01. FIG. 19A shows the levels of human COL3A1 transcripts present in immortalized human keratinocytes (HKs) after infection with C3vec01 at the indicated MOIs. Uninfected (mock) and HSV-mCherry-infected (mCherry) cells were used as negative controls. Data is presented for two replicates \pm SEM. FIG. 19B shows representative immunofluorescence images of human COL3 protein expression in immortalized human keratinocytes 48 hours after infection with C3vec01 at the indicated MOIs. Uninfected (mock) cells were used as negative controls.

[0081] FIGS. 20A-B show human COL3 nucleic acid and protein analyses in immortalized human dermal fibroblasts (HDFs) infected with C3vec01. FIG. 20A shows the levels of human COL3A1 transcripts present in immortalized human dermal fibroblasts (HDFs) after infection with

C3vec01 at the indicated MOIs. Uninfected (mock) and HSV-mCherry-infected (mCherry) cells were used as negative controls. Data is presented for two replicates \pm SEM. FIG. 20B shows representative immunofluorescence images of human COL3 protein expression in immortalized human dermal fibroblasts 48 hours after infection with C3vec01 at the indicated MOIs. Uninfected (mock) cells were used as negative controls.

[0082] FIGS. 21A-D show human COL3 nucleic acid and protein analyses in aged primary human fibroblasts (HDFs), sourced from two different vendors, infected with C3vec01 at the indicated MOIs. FIG. 21A shows the levels of human COL3A1 transcripts present in primary HDFs harvested from either a 65-year-old female patient or a 73-year-old male patient (vendor 1) after infection with C3vec01 at the indicated MOIs. Uninfected (mock) cells were used as a negative control. Data is presented for two replicates \pm SEM. FIG. 21B shows western blot analysis of human COL3A1 protein expression in primary HDFs harvested from a 73-year-old male patient (vendor 1) after infection with C3vec01 at the indicated MOIs. Uninfected (mock) cells were used as a negative control; recombinant human COL3A1 (rCOL3A1) was used as a positive control; GAPDH was used as a loading control. FIG. 21C shows the levels of human COL3A1 transcripts present in primary HDFs harvested from either a 75-year-old female patient or a 73-year-old male patient (vendor 2) after infection with C3vec01 at the indicated MOIs. Uninfected (mock) cells were used as a negative control. Data is presented for two replicates \pm SEM. FIG. 21D shows western blot analysis of human COL3A1 protein expression in primary HDFs harvested from a 75-year-old female patient (vendor 2) after infection with C3vec01 at the indicated MOIs. Uninfected (mock) cells were used as a negative control; recombinant human COL3A1 (rCOL3A1) was used as a positive control; GAPDH was used as a loading control.

[0083] FIGS. 22A-B show human COL3 nucleic acid and protein analyses in immortalized human dermal fibroblasts (HDFs) upon UV exposure. FIG. 22A shows the concentration of COL3 secreted into the supernatant of cultured HDFs 24 hours after exposure to various dosages and times of UV light, as assessed by ELISA. Supernatant collected from non-UV exposed (-UV) HDFs cultured in parallel was used as a control. FIG. 22B shows the levels of human COL3A1 transcripts present in UV-exposed immortalized human dermal fibroblasts (HDFs) after infection with C3vec01 at the indicated MOIs. Uninfected (mock) and HSV-mCherry-infected (mCherry) cells were used as negative controls. Data is presented for two replicates \pm SEM.

[0084] FIGS. 23A-C show COL3 nucleic acid and protein analyses of skin biopsies taken from control- or C3vec01-treated young (6-8-week-old) and old (~13-months-old) C57BL/6 mice 48 hours after intradermal application. FIG. 23A shows the levels of human COL3A1 DNA present in skin biopsies taken from young and old mice 48 hours after being intradermally administered either C3vec01 or vehicle control, as assessed by qPCR analysis. FIG. 23B shows the levels of human COL3A1 transcripts present in skin biopsies taken from young and old mice 48 hours after being intradermally administered either C3vec01 or vehicle control, as assessed by qRT-PCR analysis. For each condition in the qPCR and qRT-PCR analysis, data is presented as the average of four tissue samples (two replicates/tissue sample) \pm SEM. FIG. 23C shows representative immunofluorescence

images of human COL3 expression in skin biopsies taken from young and old mice 48 hours after being intradermally administered C3vec01. A young mouse intradermally administered vehicle alone was used as a negative control. DAPI staining was used to visualize nuclei.

[0085] FIGS. 24A-B show expression of wild-type (WT) human LamB3 in Vero cells infected with the indicated viral isolates. FIG. 24A shows expression of wild-type human LAMB3 in infected Vero cells, as assessed by qPCR analysis. FIG. 24B shows expression of wild-type human LamB3 protein in infected Vero cells, as assessed by western blot.

[0086] FIG. 25 shows expression of wild-type (WT) or codon-optimized (CO) human LamB3 protein in Vero cells infected with the indicated viral isolates, as assessed by western blot. Uninfected Vero cells were used as a negative control.

[0087] FIG. 26 shows expression of wild-type (WT) or codon-optimized (CO) human LamB3 protein in primary human keratinocytes infected with the indicated viral isolates, as assessed by western blot. Uninfected primary keratinocytes were used as a negative control.

[0088] FIGS. 27A-C show expression of wild-type (WT) and codon-optimized (CO) human LamC2 in Vero cells infected with the indicated viral isolates. FIG. 27A shows expression of wild-type human LAMC2 in infected Vero cells, as assessed by qPCR analysis. FIG. 27B shows expression of codon-optimized human LAMC2 in infected Vero cells, as assessed by qPCR analysis. FIG. 27C shows expression of wild-type and codon-optimized human LamC2 protein in infected Vero cells, as assessed by western blot. The boxed viral isolate "LGA" expressing codon-optimized LamC2 was selected for additional experimentation.

[0089] FIGS. 28A-C show human LAMC2 expressed from viral isolate "LGA" in immortalized primary human keratinocytes infected at the indicated multiplicities of infection (MOIs). FIG. 28A shows the viral genome copy number in primary immortalized human keratinocytes after infection with viral isolate "LGA" at the indicated MOIs. FIG. 28B shows the transcript level of codon-optimized LAMC2 expressed in primary immortalized human keratinocytes after infection with viral isolate "LGA" at the indicated MOIs. FIG. 28C shows expression of human LamC2 protein in primary immortalized human keratinocytes after infection with viral isolate "LGA" at the indicated MOIs, as assessed by western blot.

[0090] FIGS. 29A-D show LAMC2 nucleic acid and protein analysis of skin biopsies taken from control (vehicle)- or HSV isolate "LGA"-treated mice 72 hours after intradermal application. FIG. 29A shows a schematic of the intradermal injection sites on the treated animals. FIG. 29B shows the levels of human LAMC2 DNA present in skin biopsies taken from mice 72 hours after being intradermally administered either HSV isolate LGA or vehicle control, as assessed by qPCR analysis. FIG. 29C shows the levels of human LAMC2 transcripts present in skin biopsies taken from mice 72 hours after being intradermally administered either HSV isolate LGA or vehicle control, as assessed by qRT-PCR analysis. For each condition in the qPCR and qRT-PCR analysis, data is presented as the average of two replicates \pm SEM. FIG. 29D shows representative immunofluorescence images of human LAMC2 expression in skin biopsies taken from mice 72 hours after being intradermally administered HSV isolate LGA. A site that was intradermally administered vehicle alone was used as a negative

control. DAPI staining was used to visualize nuclei; pKal staining was used to visualize mouse laminin-332.

DETAILED DESCRIPTION

[0091] In some embodiments, the present disclosure relates to recombinant nucleic acids (e.g., recombinant herpes viral genomes) encoding one or more cosmetic proteins, and to uses of these recombinant nucleic acids in viruses (e.g., in a herpes virus), compositions, formulations, medicaments, and/or methods for delivering one or more cosmetic proteins to the skin, such as onto, into, and/or through the skin (e.g., to the dermal ECM). In some embodiments, the present disclosure relates to recombinant nucleic acids (e.g., recombinant herpes viral genomes) encoding one or more cosmetic proteins, and to uses of these recombinant nucleic acids in viruses (e.g., in a herpes virus), compositions, formulations, medicaments, and/or methods in order to increase, augment, and/or supplement one or more dermal ECM proteins (e.g., one or more collagen proteins). In some embodiments, the present disclosure relates to recombinant nucleic acids (e.g., recombinant herpes viral genomes) encoding one or more cosmetic proteins, and to uses of these recombinant nucleic acids in viruses (e.g., in a herpes virus), compositions, formulations, medicaments, and/or methods in the aesthetic context (e.g., to reduce one or more dermatological signs of aging). In some embodiments, the present disclosure relates to compositions comprising a recombinant herpes viral vector and methods comprising the delivery of the recombinant herpes viral vector onto, into, and/or through the skin of a mammal, wherein the recombinant herpes viral vector comprises a promoter operable in a mammalian cell and a heterologous nucleic acid which is expressed to achieve a cosmetic effect in mammalian skin. The heterologous nucleic acid may be delivered to a mammalian target skin cell of a mammal, comprising contacting the epidermis, dermis, or subcutaneous tissue of the mammal with the composition comprising the recombinant herpes viral vector, under conditions whereby the recombinant herpes viral vector is transported onto, into, and/or through the epidermis, dermis or subcutaneous tissue and introduced into the target skin cell, where it is expressed. Without wishing to be bound by theory, it is believed that administering one or more of the recombinant nucleic acids, viruses, and/or formulations described herein to an individual will allow for increased production of functional dermal ECM proteins (e.g., human collagen) in the individual. Furthermore, without wishing to be bound by theory, it is believed that increasing, augmenting, and/or supplementing the levels of cosmetic proteins in an individual by administering one or more of the recombinant nucleic acids, viruses, and/or formulations described herein will lead to at least one of: 1) the enhancement, augmentation, and/or supplementation of soft tissue; 2) the improvement of skin quality, condition, and/or appearance; 3) the reduction of one or more superficial depressions in the skin (e.g., wrinkles); 4) the improvement of texture, smoothness, elasticity, and/or tension of the skin; and/or 5) the reduction of one or more dermatological signs of aging. Ultimately, without wishing to be bound by theory, it is believed that the recombinant nucleic acids, viruses, compositions, and methods described herein provide a novel strategy for delivering functional cosmetic proteins in aesthetic settings.

[0092] The following description sets forth exemplary methods, parameters, and the like. It should be recognized,

however, that such a description is not intended as a limitation on the scope of the present disclosure but is instead provided as a description of exemplary embodiments.

I. General Techniques

[0093] The techniques and procedures described or referenced herein are generally well understood and commonly employed using conventional methodology by those skilled in the art, such as, for example, the widely utilized methodologies described in Sambrook et al., *Molecular Cloning: A Laboratory Manual* 3d edition (2001) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; *Current Protocols in Molecular Biology* (F. M. Ausubel, et al. eds., (2003)); the series *Methods in Enzymology* (Academic Press, Inc.); *PCR 2: A Practical Approach* (M. J. MacPherson, B. D. Hames and G. R. Taylor eds. (1995)), Harlow and Lane, eds. (1988); *Oligonucleotide Synthesis* (M. J. Gait, ed., 1984); *Methods in Molecular Biology*, Humana Press; *Cell Biology: A Laboratory Notebook* (J. E. Cellis, ed., 1998) Academic Press; *Animal Cell Culture* (R. I. Freshney), ed., 1987); *Introduction to Cell and Tissue Culture* (J. P. Mather and P. E. Roberts, 1998) Plenum Press; *Cell and Tissue Culture: Laboratory Procedures* (A. Doyle, J. B. Griffiths, and D. G. Newell, eds., 1993-8) J. Wiley and Sons; *Gene Transfer Vectors for Mammalian Cells* (J. M. Miller and M. P. Calos, eds., 1987); *PCR: The Polymerase Chain Reaction*, (Mullis et al., eds., 1994); *Short Protocols in Molecular Biology* (Wiley and Sons, 1999).

II. Definitions

[0094] Before describing the present disclosure in detail, it is to be understood that the present disclosure is not limited to particular compositions or biological systems, which can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

[0095] As used herein, the singular forms “a” “an” and “the” include plural referents unless the content clearly dictates otherwise. Thus, for example, reference to “a molecule” optionally includes a combination of two or more such molecules, and the like.

[0096] As used herein, the term “and/or” may include any and all combinations of one or more of the associated listed items. For example, the term “a and/or b” may refer to “a alone”, “b alone”, “a or b”, or “a and b”; the term “a, b, and/or c” may refer to “a alone”, “b alone”, “c alone”, “a or b”, “a or c”, “b or c”, “a, b, or c”, “a and b”, “a and c”, “b and c”, or “a, b, and c”; etc.

[0097] As used herein, the term “about” refers to the usual error range for the respective value readily known to the skilled person in this technical field. Reference to “about” a value or parameter herein includes (and describes) embodiments that are directed to that value or parameter per se.

[0098] It is understood that aspects and embodiments of the present disclosure include “comprising”, “consisting”, and “consisting essentially of” aspects and embodiments.

[0099] As used herein, the terms “polynucleotide”, “nucleic acid sequence”, “nucleic acid”, and variations thereof shall be generic to polydeoxyribonucleotides (containing 2-deoxy-D-ribose), to polyribonucleotides (containing D-ribose), to any other type of polynucleotide that is an N-glycoside of a purine or pyrimidine base, and to other polymers containing non-nucleotidic backbones, provided

that the polymers contain nucleobases in a configuration that allows for base pairing and base stacking, as found in DNA and RNA. Thus, these terms include known types of nucleic acid sequence modifications, for example, substitution of one or more of the naturally occurring nucleotides with an analog, and inter-nucleotide modifications.

[0100] As used herein, a nucleic acid is “operatively linked” or “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For example, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, “operatively linked” or “operably linked” means that the DNA sequences being linked are contiguous.

[0101] As used herein, the term “vector” refers to discrete elements that are used to introduce heterologous nucleic acids into cells for either expression or replication thereof. An expression vector includes vectors capable of expressing nucleic acids that are operatively linked with regulatory sequences, such as promoter regions, that are capable of effecting expression of such nucleic acids. Thus, an expression vector may refer to a DNA or RNA construct, such as a plasmid, a phage, recombinant virus or other vector that, upon introduction into an appropriate host cell, results in expression of the nucleic acids. Appropriate expression vectors are well known to those of skill in the art and include those that are replicable in eukaryotic cells, and those that remain episomal or those which integrate into the host cell genome.

[0102] As used herein, an “open reading frame” or “ORF” refers to a continuous stretch of nucleic acids, either DNA or RNA, that encode a protein or polypeptide. Typically, the nucleic acids comprise a translation start signal or initiation codon, such as ATG or AUG, and a termination codon.

[0103] As used herein, an “untranslated region” or “UTR” refers to untranslated nucleic acids at the 5' and/or 3' ends of an open reading frame. The inclusion of one or more UTRs in a polynucleotide may affect post-transcriptional regulation, mRNA stability, and/or translation of the polynucleotide.

[0104] As used herein, the term “transgene” refers to a polynucleotide that is capable of being transcribed into RNA and translated and/or expressed under appropriate conditions, after being introduced into a cell. In some aspects, it confers a desired property to a cell into which it was introduced, or otherwise leads to a desired cosmetic, therapeutic, or diagnostic outcome.

[0105] As used herein, the terms “polypeptide,” “protein,” and “peptide” are used interchangeably and may refer to a polymer of two or more amino acids.

[0106] As used herein, a “subject”, “host”, or an “individual” refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, horses, cats, cows, as well as animals used in research, such as mice, rats, hamsters, rabbits, and non-human primates, etc. In some embodiments, the mammal is human.

[0107] As used herein, the terms “pharmaceutical formulation” or “pharmaceutical composition” refer to a preparation which is in such a form as to permit the biological activity of the active ingredient(s) to be effective, and which contains no additional components which are unacceptably toxic to a subject to which the composition or formulation

would be administered. "Pharmaceutically acceptable" excipients (e.g., vehicles, additives) are those which can reasonably be administered to a subject to provide an effective dose of the active ingredient(s) employed.

[0108] As used herein, "cutaneous administration" or "cutaneously administering" refers to the delivery of a composition to a subject by contacting, directly or otherwise, a formulation comprising the composition to all ("systemic") or a portion ("topical") of the skin of a subject. The term encompasses several routes of administration including, but not limited to, topical and transdermal. Topical administration may be used as a means to deliver a composition to the epidermis or dermis of a subject, or to specific strata thereof.

[0109] As used herein, "treatment" refers to clinical intervention designed to alter the natural course of the individual or cell being treated during the course of clinical pathology. Desirable effects of treatment include decreasing the rate of disease/disorder/defect progression, ameliorating or palliating the disease/disorder/defect state, and remission or improved prognosis. For example, an individual is successfully "treated" if one or more symptoms associated with dermatological aging are reduced, mitigated, or eliminated, including the reduction or elimination of wrinkles.

[0110] As used herein, the term "delaying progression of" a disease/disorder/defect refers to deferring, hindering, slowing, retarding, stabilizing, and/or postponing development of the disease/disorder/defect (e.g., skin wrinkles). This delay can be of varying lengths or time, depending on the history of the disease/disorder/defect and/or the individual being treated. As is evident to one of ordinary skill in the art, a sufficient or significant delay can, in effect, encompass prevention, in that the individual does not develop the disease/disorder/defect.

III. Recombinant Nucleic Acids

[0111] Certain aspects of the present disclosure relate to recombinant nucleic acids (e.g., isolated recombinant nucleic acids) comprising one or more polynucleotides (e.g., one or more, two or more, three or more, four or more, five or more, ten or more, etc.) encoding a cosmetic protein. Any suitable cosmetic protein described herein or known in the art may be encoded by the polynucleotides of the present disclosure, including, for example, collagen proteins, fibronectins, elastins, lumicans, vitronectins/vitronectin receptors, laminins, neuromodulators, fibrillins, additional dermal ECM proteins, etc. In some embodiments, the cosmetic protein is a structural extracellular matrix protein (e.g., a collagen, elastin, fibronectin, laminin, fibrillin, etc.). In some embodiments, the cosmetic protein is a collagen, elastin, fibronectin, or laminin protein (e.g., a human collagen, elastin, fibronectin, or laminin protein).

[0112] In some embodiments, the present disclosure relates to recombinant nucleic acids (e.g., isolated recombinant nucleic acids) comprising one or more polynucleotides (e.g., one or more, two or more, three or more, four or more, five or more, ten or more, etc.) encoding a collagen protein. In some embodiments, the collagen protein is a human collagen protein. In some embodiments, the present disclosure relates to recombinant nucleic acids comprising one or more polynucleotides encoding a homotrimeric collagen (e.g., a heterotrimeric human collagen, such as human Collagen 3 (e.g., comprising three COL3A1 (COL3) polypeptides) or human Collagen 7 (e.g., comprising three

COL7A1 (COL7) polypeptides). In some embodiments, the present disclosure relates to recombinant nucleic acids comprising one or more polynucleotides encoding a heterotrimeric collagen (e.g., a heterotrimeric human collagen, such as human Collagen 1 (e.g., comprising two COL1A1 (COL1-1) polypeptides and one COL1A2 (COL1-2) polypeptide) or human Collagen 4 (e.g., comprising two COL4A1 (COL4-1) polypeptides and one COL4A2 (COL4-2) polypeptide). In some embodiments, the present disclosure relates to recombinant nucleic acids comprising one or more polynucleotides encoding a homotrimeric collagen and a heterotrimeric collagen (e.g., a recombinant nucleic acid comprising one or more polynucleotides encoding a human Collagen 1 and a human Collagen 3). In some embodiments, the present disclosure relates to recombinant nucleic acids comprising one or more polynucleotides encoding human Collagen 1. In some embodiments, the present disclosure relates to recombinant nucleic acids comprising one or more polynucleotides encoding human Collagen 3.

[0113] In some embodiments, the present disclosure relates to recombinant nucleic acids comprising a first polynucleotide encoding a first polypeptide comprising a first cosmetic protein (e.g., a first human collagen protein). In some embodiments, the first polypeptide consists essentially of or consists of the first cosmetic protein (e.g., consists essentially of or consists of a first human collagen protein). In some embodiments, the present disclosure relates to recombinant nucleic acids comprising a first polynucleotide encoding a first polypeptide comprising a first cosmetic protein (e.g., a first human collagen protein), a linker polypeptide, and a further cosmetic protein (e.g., a further human collagen protein). In some embodiments, the first and further cosmetic proteins (e.g., the first and further human collagen proteins) are the same. In some embodiments, the first and further cosmetic proteins (e.g., the first and further human collagen proteins) are different. In some embodiments, the linker polypeptide is a cleavable linker polypeptide.

[0114] In some embodiments, the present disclosure relates to recombinant nucleic acids comprising a first polynucleotide encoding a first polypeptide comprising a first cosmetic protein (e.g., a first human collagen protein), wherein the first polynucleotide encodes a polycistronic mRNA comprising a first open reading frame (ORF) encoding the first polypeptide, an internal ribosomal entry site (IRES), and a second ORF encoding an additional cosmetic protein (e.g., an additional human collagen protein). In some embodiments, the first and additional cosmetic proteins (e.g., the first and additional human collagen proteins) are the same. In some embodiments, the first and additional cosmetic proteins (e.g., the first and additional human collagen proteins) are different.

[0115] In some embodiments, the present disclosure relates to recombinant nucleic acids comprising a first polynucleotide encoding a first polypeptide comprising a first cosmetic protein (e.g., a first human collagen protein), and a second polynucleotide encoding a second cosmetic protein (e.g., a second human collagen protein). In some embodiments, the first and second cosmetic proteins (e.g., the first and second human collagen proteins) are the same. In some embodiments, the first and second cosmetic proteins (e.g., the first and second human collagen proteins) are different.

[0116] In some embodiments, the recombinant nucleic acid is a vector. In some embodiments, the recombinant nucleic acid is a viral vector. In some embodiments, the recombinant nucleic acid is a herpes viral vector. In some embodiments, the recombinant nucleic acid is a herpes simplex virus amplicon. In some embodiments, the recombinant nucleic acid is a recombinant herpes virus genome. In some embodiments, the recombinant nucleic acid is a recombinant herpes simplex virus genome. In some embodiments, the recombinant herpes simplex virus genome is a recombinant type 1 herpes simplex virus (HSV-1) genome.

Polynucleotides Encoding Cosmetic Proteins

Polynucleotides Encoding Collagen Proteins

[0117] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising one or more polynucleotides comprising the coding sequence of a collagen gene. The coding sequence of any collagen gene (including any isoform thereof) from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, human collagen genes (see e.g., NCBI Gene IDs: 1277, 1278, 1281, 1282, 1284, 1291, 1294, 1308, etc.), mouse collagen genes (see, e.g., NCBI Gene IDs: 12842, 12843, 12825, 12826, 12827, 12833, 12836, 12821, etc.), chimpanzee collagen genes (see e.g., NCBI Gene IDs: 104001053, 455117, 459815, 452689, 452661, 450204, 101056895, 101058306, etc.), rat collagen genes (see e.g., NCBI Gene IDs: 29393, 84352, 84032, 290905, 306628, 294337, 301012, 294027, etc.), rabbit collagen genes (see e.g., NCBI Gene IDs: 100347598, 100008997, 100009177, 100358256, 100358522, 100343947, 100356561, 100339335, etc.) etc. Methods of identifying collagen gene homologs/orthologs from additional species are known to one of ordinary skill in the art, including, for example, using a nucleic acid sequence alignment program such as the BLAST® blastn suite. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the collagen genes (and/or coding sequences thereof) described herein or known in the art.

[0118] In some embodiments, a polynucleotide of the present disclosure comprises a codon-optimized variant of the coding sequence of any of the collagen genes described herein or known in the art. In some embodiments, use of a codon-optimized variant of the coding sequence of a collagen gene increases stability and/or yield of heterologous expression (RNA and/or protein) of the encoded collagen protein in a target cell (such as a cell of the epidermis and/or dermis), as compared to the stability and/or yield of heterologous expression of a corresponding, non-codon-optimized, wild-type sequence. Any suitable method known in the art for performing codon optimization of a sequence for expression in one or more target cells (e.g., one or more human cells) may be used, including, for example, by the methods described by Fath et al. (PLoS One. 2011 Mar. 3; 6(3): e17596).

[0119] In some embodiments, the present disclosure relates to one or more polynucleotides (i.e., one or more first polynucleotides and/or one or more second polynucleotides)

comprising the coding sequence of a human collagen gene. Any suitable human collagen gene (including any isoform thereof) known in the art may be encoded by a nucleic acid of the present disclosure, including, for example, a COL1A1 gene (see e.g., NCBI Gene ID: 1277; SEQ ID NO: 1), a COL1A2 gene (see e.g., NCBI Gene ID: 1278; SEQ ID NO: 3), a COL2A1 gene (see e.g., NCBI Gene ID: 1280), a COL3A1 gene (see e.g., NCBI Gene ID: 1281; SEQ ID NO: 5), a COL4A1 gene (see e.g., NCBI Gene ID: 1282; SEQ ID NO: 7), a COL4A2 gene (see e.g., NCBI Gene ID: 1284), a COL4A3 gene (see e.g., NCBI Gene ID: 1285), a COL4A4 gene (see e.g., NCBI Gene ID: 1286), a COL4A5 gene (see e.g., NCBI Gene ID: 1287), a COL4A6 gene (see e.g., NCBI Gene ID: 1288), a COL5A1 gene (see e.g., NCBI Gene ID: 1289), a COL5A2 gene (see e.g., NCBI Gene ID: 1290), a COL5A3 gene (see e.g., NCBI Gene ID: 50509), a COL6A1 gene (see e.g., NCBI Gene ID: 1291; SEQ ID NO: 9), a COL6A2 gene (see e.g., NCBI Gene ID: 1292), a COL6A3 gene (see e.g., NCBI Gene ID: 1293), a COL6A4 gene (see e.g., NCBI Gene ID: 344875), a COL6A5 gene (see e.g., NCBI Gene ID: 256076), a COL6A6 gene (see e.g., NCBI Gene ID: 131873), a COL7A1 gene (see e.g., NCBI Gene ID: 1294; SEQ ID NO: 10), a COL8A1 gene (see e.g., NCBI Gene ID: 1295), a COL9A1 gene (see e.g., NCBI Gene ID: 1297), a COL9A2 gene (see e.g., NCBI Gene ID: 1298), a COL9A3 gene (see e.g., NCBI Gene ID: 1299), a COL10A1 gene (see e.g., NCBI Gene ID: 1300), a COL11A1 gene (see e.g., NCBI Gene ID: 1301), a COL11A2 gene (see e.g., NCBI Gene ID: 1302), a COL12A1 gene (see e.g., NCBI Gene ID: 1303), a COL13A1 gene (see e.g., NCBI Gene ID: 1305), a COL14A1 gene (see e.g., NCBI Gene ID: 7373), a COL15A1 gene (see e.g., NCBI Gene ID: 1306), a COL16A1 gene (see e.g., NCBI Gene ID: 1307), a COL17A1 gene (see e.g., NCBI Gene ID: 1308; SEQ ID NO: 12), a COL18A1 gene (see e.g., NCBI Gene ID: 80781), a COL19A1 gene (see e.g., NCBI Gene ID: 1310), a COL20A1 gene (see e.g., NCBI Gene ID: 57642), a COL21A1 gene (see e.g., NCBI Gene ID: 81578), a COL22A1 gene (see e.g., NCBI Gene ID: 169044), a COL23A1 gene (see e.g., NCBI Gene ID: 91522), a COL24A1 gene (see e.g., NCBI Gene ID: 255631), a COL25A1 gene (see e.g., NCBI Gene ID: 84570), a COL26A1 gene (see e.g., NCBI Gene ID: 136227), a COL27A1 gene (see e.g., NCBI Gene ID: 85301), a COL28A1 gene (see e.g., NCBI Gene ID: 340267), etc. In some embodiments, a polynucleotide (i.e., one or more first polynucleotides and/or one or more second polynucleotides) of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the human collagen genes (and/or coding sequences thereof) described herein or known in the art.

[0120] In some embodiments, a polynucleotide (i.e., one or more first polynucleotides and/or one or more second polynucleotides) of the present disclosure comprises a codon-optimized variant of any of the human collagen genes described herein. In some embodiments, use of a codon-optimized variant of a human collagen gene increases stability and/or yield of heterologous expression (RNA and/or protein) of the human collagen in a target cell (such as a human keratinocyte or fibroblast), as compared to the sta-

[0135] In some embodiments, a polynucleotide of the present disclosure encoding one or more human collagen proteins (e.g., a first human collagen protein, a further human collagen protein, an additional human collagen protein, and/or a second human collagen protein) has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14. In some embodiments, a polynucleotide of the present disclosure encoding one or more human collagen proteins (e.g., a first human collagen protein, a further human collagen protein, an additional human collagen protein, and/or a second human collagen protein) comprises a sequence selected from SEQ ID NOS: 1-14.

Polynucleotides Encoding Fibronectin Proteins

[0136] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising one or more polynucleotides comprising the coding sequence of a fibronectin gene. The coding sequence of any fibronectin gene (including any isoform thereof) from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human fibronectin gene (see e.g., NCBI Gene ID: 2335), a mouse fibronectin gene (see, e.g., NCBI Gene ID: 14268), a chimpanzee fibronectin gene (see e.g., NCBI Gene ID: 459926), a rat fibronectin gene (see e.g., NCBI Gene ID: 25661), a rabbit fibronectin gene (see e.g., NCBI Gene ID: 100328589), etc. Methods of identifying fibronectin gene homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the fibronectin genes (and/or coding sequences thereof) described herein or known in the art. In some embodiments, a polynucleotide of the present disclosure comprises a codon-optimized variant of any of the fibronectin genes (and/or coding sequences thereof) described herein or known in the art.

[0137] In some embodiments, the present disclosure relates to one or more polynucleotides (i.e., one or more first polynucleotides and/or one or more second polynucleotides) comprising the coding sequence of a human fibronectin gene. In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the human FN1 gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 35 or SEQ ID NO: 36. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 35 or SEQ ID NO: 36.

[0138] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 35 or SEQ ID

NO: 36. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 35 or SEQ ID NO: 36 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2500, at least 3000, at least 3500, at least 4000, at least about 4500, at least about 5000, at least about 5500, at least about 6000, at least about 6500, at least about 7000, but fewer than 7434, consecutive nucleotides of SEQ ID NO: 35 or SEQ ID NO: 36. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-7431 of SEQ ID NO: 35 or SEQ ID NO: 36. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-7431 of SEQ ID NO: 35 or SEQ ID NO: 36.

Polynucleotides Encoding Elastin Proteins

[0139] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising one or more polynucleotides comprising the coding sequence of an elastin gene. The coding sequence of any elastin gene (including any isoform thereof) from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human elastin gene (see e.g., NCBI Gene ID: 2006), a mouse elastin gene (see, e.g., NCBI Gene ID: 13717), a chimpanzee elastin gene (see e.g., NCBI Gene ID: 463943), a rat elastin gene (see e.g., NCBI Gene ID: 25043), a rabbit elastin gene (see e.g., NCBI Gene ID: 100344271), etc. Methods of identifying elastin gene homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the elastin genes (and/or coding sequences thereof) described herein or known in the art. In some embodiments, a polynucleotide of the present disclosure comprises a codon-optimized variant of any of the elastin genes (and/or coding sequences thereof) described herein or known in the art.

[0140] In some embodiments, the present disclosure relates to one or more polynucleotides (i.e., one or more first polynucleotides and/or one or more second polynucleotides) comprising the coding sequence of a human elastin gene. In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the human ELN gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 37 or SEQ ID NO: 38. In some embodiments, a

polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 37 or SEQ ID NO: 38.

[0141] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 37 or SEQ ID NO: 38. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 37 or SEQ ID NO: 38 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2250, but fewer than 2361, consecutive nucleotides of SEQ ID NO: 37 or SEQ ID NO: 38. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-2358 of SEQ ID NO: 37 or SEQ ID NO: 38. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-2358 of SEQ ID NO: 37 or SEQ ID NO: 38.

Polynucleotides Encoding Lumican Proteins

[0142] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising one or more polynucleotides comprising the coding sequence of a lumican gene. The coding sequence of any lumican gene (including any isoform thereof) from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human lumican gene (see e.g., NCBI Gene ID: 4060), a mouse lumican gene (see, e.g., NCBI Gene ID: 17022), a chimpanzee lumican gene (see e.g., NCBI Gene ID: 452119), a rat lumican gene (see e.g., NCBI Gene ID: 81682), a rabbit lumican gene (see e.g., NCBI Gene ID: 100008665), etc. Methods of identifying lumican gene homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the lumican genes (and/or coding sequences thereof) described herein or known in the art. In some embodiments, a polynucleotide of the present disclosure comprises a codon-optimized variant of any of the lumican genes (and/or coding sequences thereof) described herein or known in the art.

[0143] In some embodiments, the present disclosure relates to one or more polynucleotides (i.e., one or more first polynucleotides and/or one or more second polynucleotides) comprising the coding sequence of a human lumican gene. In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the human LUM gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least

98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 39 or SEQ ID NO: 40. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 39 or SEQ ID NO: 40.

[0144] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 39 or SEQ ID NO: 40. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 39 or SEQ ID NO: 40 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, but fewer than 1017, consecutive nucleotides of SEQ ID NO: 39 or SEQ ID NO: 40. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-1014 of SEQ ID NO: 39 or SEQ ID NO: 40. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-1014 of SEQ ID NO: 39 or SEQ ID NO: 40.

Polynucleotides Encoding Vitronectin and Vitronectin Receptor Proteins

[0145] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising one or more polynucleotides comprising the coding sequence of a vitronectin or vitronectin receptor gene. The coding sequence of any vitronectin or vitronectin receptor gene (including any isoform thereof) from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human vitronectin or vitronectin receptor gene (see e.g., NCBI Gene IDs: 7448 and 3685), a mouse vitronectin or vitronectin receptor gene (see e.g., NCBI Gene IDs: 22370 and 16410), a chimpanzee vitronectin or vitronectin receptor gene (see e.g., NCBI Gene IDs: 738261 and 459807), a rat vitronectin or vitronectin receptor gene (see e.g., NCBI Gene IDs: 29169 and 257645), a rabbit vitronectin or vitronectin receptor gene (see e.g., NCBI Gene IDs: 100009128 and 100008956), etc. Methods of identifying vitronectin or vitronectin receptor gene homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the vitronectin or vitronectin receptor genes (and/or coding sequences thereof) described herein or known in the art. In some embodiments, a polynucleotide of the present disclosure comprises a codon-optimized variant of any of the vitronectin or vitronectin receptor genes (and/or coding sequences thereof) described herein or known in the art.

[0146] In some embodiments, the present disclosure relates to one or more polynucleotides (i.e., one or more first polynucleotides and/or one or more second polynucleotides)

comprising the coding sequence of a human vitronectin or vitronectin receptor gene. In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the human VTN gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 41 or SEQ ID NO: 42. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 41 or SEQ ID NO: 42.

[0147] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 41 or SEQ ID NO: 42. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 41 or SEQ ID NO: 42 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, at least about 1250, but fewer than 1437, consecutive nucleotides of SEQ ID NO: 41 or SEQ ID NO: 42. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-1034 of SEQ ID NO: 41 or SEQ ID NO: 42. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-1034 of SEQ ID NO: 41 or SEQ ID NO: 42.

Polynucleotides Encoding Laminin Proteins

[0148] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising one or more polynucleotides comprising the coding sequence of a laminin gene. The coding sequence of any laminin gene (including any isoform thereof) from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, human laminin genes (see e.g., NCBI Gene IDs: 284217, 3908, 3909, 3910, 3911, 3912, 3913, 3914, 3915, 3918, and 10319), mouse laminin genes (see e.g., NCBI Gene IDs: 16774, 16780, and 16782), chimpanzee laminin genes (see e.g., NCBI Gene IDs: 455339, 469668, and 457571), rat laminin genes (see e.g., NCBI Gene IDs: 307582, 305078, and 192362), rabbit laminin genes (see e.g., NCBI Gene IDs: 100346886 and 100342905), etc. Methods of identifying laminin gene homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the laminin genes (and/or coding sequences thereof) described herein or known in the art. In some embodiments, a polynucleotide of the present disclosure comprises a

codon-optimized variant of any of the laminin genes (and/or coding sequences thereof) described herein or known in the art.

[0149] In some embodiments, the present disclosure relates to one or more polynucleotides (i.e., one or more first polynucleotides and/or one or more second polynucleotides) comprising the coding sequence of a human laminin gene, such as a human LAMA1 gene (see e.g., NCBI Gene ID: 284217), a human LAMA2 gene (see e.g., NCBI Gene ID: 3908), a human LAMA3 gene (see e.g., NCBI Gene ID: 3909), a human LAMA4 gene (see e.g., NCBI Gene ID: 3910), a human LAMA5 gene (see e.g., NCBI Gene ID: 3911), a human LAMB1 gene (see e.g., NCBI Gene ID: 3912), a human LAMB2 gene (see e.g., NCBI Gene ID: 3913), a human LAMB3 gene (see e.g., NCBI Gene ID: 3914), a human LAMC1 gene (see e.g., NCBI Gene ID: 3915), a human LAMC2 gene (see e.g., NCBI Gene ID: 3918), or a human LAMC3 gene (see e.g., NCBI Gene ID: 10319).

[0150] In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the human LAMA3 gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 43 or SEQ ID NO: 44. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 43 or SEQ ID NO: 44.

[0151] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 43 or SEQ ID NO: 44. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 43 or SEQ ID NO: 44 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2500, at least 3000, at least 3500, at least 4000, at least about 4500, at least about 5000, but fewer than 5175, consecutive nucleotides of SEQ ID NO: 43 or SEQ ID NO: 44. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-5172 of SEQ ID NO: 43 or SEQ ID NO: 44. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-5172 of SEQ ID NO: 43 or SEQ ID NO: 44.

[0152] In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the human LAMB3 gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the

sequence of SEQ ID NO: 45 or SEQ ID NO: 46. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 45 or SEQ ID NO: 46.

[0153] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 45 or SEQ ID NO: 46. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 45 or SEQ ID NO: 46 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2500, at least 3000, at least 3500, but fewer than 3519, consecutive nucleotides of SEQ ID NO: 45 or SEQ ID NO: 46. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-3516 of SEQ ID NO: 45 or SEQ ID NO: 46. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-3516 of SEQ ID NO: 45 or SEQ ID NO: 46.

[0154] In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the human LAMC2 gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 47 or SEQ ID NO: 48. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 47 or SEQ ID NO: 48.

[0155] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 47 or SEQ ID NO: 48. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 47 or SEQ ID NO: 48 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2500, at least 3000, at least 3500, but fewer than 3582, consecutive nucleotides of SEQ ID NO: 47 or SEQ ID NO: 48. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-3579 of SEQ ID NO: 47 or SEQ ID NO: 48. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-3579 of SEQ ID NO: 47 or SEQ ID NO: 48.

Polynucleotides Encoding Neuromodulator Proteins

[0156] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising one or more polynucleotides comprising the coding sequence of a neuromodulator gene. The coding sequence of any neuromodulator gene (including any isoform thereof) from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a *Clostridium botulinum* neuromodulator gene (see e.g., NCBI Gene IDs: 5185061 and 39483740), etc. Methods of identifying neuromodulator gene homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the neuromodulator genes (and/or coding sequences thereof) described herein or known in the art. In some embodiments, a polynucleotide of the present disclosure comprises a codon-optimized variant of any of the neuromodulator genes (and/or coding sequences thereof) described herein or known in the art.

[0157] In some embodiments, the present disclosure relates to one or more polynucleotides (i.e., one or more first polynucleotides and/or one or more second polynucleotides) comprising the coding sequence of a *Clostridium botulinum* neuromodulator gene.

[0158] In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the *Clostridium botulinum* botA gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 49 or SEQ ID NO: 50. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 49 or SEQ ID NO: 50.

[0159] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 49 or SEQ ID NO: 50. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 49 or SEQ ID NO: 50 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, at least about 1250, at least 1500, at least 1750, at least 2000, at least 2500, at least 3000, at least 3500, but fewer than 3891, consecutive nucleotides of SEQ ID NO: 49 or SEQ ID NO: 50. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-3888 of SEQ ID NO: 49 or SEQ ID NO: 50. In some embodiments, a

polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-3888 of SEQ ID NO: 49 or SEQ ID NO: 50.

[0160] In some embodiments, a polynucleotide of the present disclosure comprises the coding sequence of the *Clostridium botulinum* botB gene (or a codon-optimized variant thereof). In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 51 or SEQ ID NO: 52. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of SEQ ID NO: 51 or SEQ ID NO: 52.

[0161] In some embodiments, a polynucleotide of the present disclosure comprises a 5' truncation, a 3' truncation, or a fragment of the sequence of SEQ ID NO: 51 or SEQ ID NO: 52. In some embodiments, the 5' truncation, 3' truncation, or fragment of the sequence of SEQ ID NO: 51 or SEQ ID NO: 52 is a polynucleotide that has at least 25, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, or at least 350, at least 400, at least 450, at least 500, at least 750, at least 1000, at least about 1250, at least 1500, at least 1750, at least 2000, at least 2500, at least 3000, at least 3500, but fewer than 3876, consecutive nucleotides of SEQ ID NO: 51 or SEQ ID NO: 52. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of nucleic acids 1-3873 of SEQ ID NO: 51 or SEQ ID NO: 52. In some embodiments, a polynucleotide of the present disclosure comprises the sequence of nucleic acids 1-3873 of SEQ ID NO: 51 or SEQ ID NO: 52.

Polynucleotides Encoding Fibrillin Proteins

[0162] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising one or more polynucleotides comprising the coding sequence of a fibrillin gene. The coding sequence of any fibrillin gene (including any isoform thereof) from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, human fibrillin genes (see e.g., NCBI Gene IDs: 2200, 2201, and 84467), mouse fibrillin genes (see e.g., NCBI Gene IDs: 14118 and 14119), chimpanzee fibrillin genes (see e.g., NCBI Gene IDs: 453411, 471621, and 455669), rat fibrillin genes (see e.g., NCBI Gene IDs: 83727 and 689008), rabbit fibrillin genes (see e.g., NCBI Gene IDs: 100350931, 100357126, and 100359336), etc. Methods of identifying fibrillin gene homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the fibrillin genes (and/or coding sequences thereof)

described herein or known in the art. In some embodiments, a polynucleotide of the present disclosure comprises a codon-optimized variant of any of the fibrillin genes (and/or coding sequences thereof) described herein or known in the art.

[0163] In some embodiments, the present disclosure relates to one or more polynucleotides (i.e., one or more first polynucleotides and/or one or more second polynucleotides) comprising the coding sequence of a human fibrillin gene, such as a human FBN1 gene (see e.g., NCBI Gene ID: 2200), a human FBN2 gene (see e.g., NCBI Gene ID: 2201), or a human FBN3 gene (see e.g., NCBI Gene ID: 84467).

Exemplary Polynucleotides

[0164] In some embodiments, a polynucleotide of the present disclosure encoding one or more cosmetic proteins (e.g., a first cosmetic protein, a further cosmetic protein, an additional cosmetic protein, and/or a second cosmetic protein) has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14 or 35-52. In some embodiments, a polynucleotide of the present disclosure encoding one or more cosmetic proteins (e.g., a first cosmetic protein, a further cosmetic protein, an additional cosmetic protein, and/or a second cosmetic protein) comprises a sequence selected from SEQ ID NOS: 1-14 or 35-52.

[0165] In some embodiments, a polynucleotide of the present disclosure encoding one or more cosmetic proteins (e.g., a first cosmetic protein, a further cosmetic protein, an additional cosmetic protein, and/or a second cosmetic protein) has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NOS: 1-14, 35-38, or 43-48. In some embodiments, a polynucleotide of the present disclosure encoding one or more cosmetic proteins (e.g., a first cosmetic protein, a further cosmetic protein, an additional cosmetic protein, and/or a second cosmetic protein) comprises a sequence selected from SEQ ID NOS: 1-14, 35-38, or 43-48.

[0166] A polynucleotide of the present disclosure encoding a cosmetic protein (e.g., a human collagen protein) may further encode additional coding and non-coding sequences. Examples of additional coding and non-coding sequences may include, but are not limited to, sequences encoding additional polypeptide tags (e.g., encoded in-frame with the cosmetic protein in order to produce a fusion protein), introns (e.g., native, modified, or heterologous introns), 5' and/or 3' UTRs (e.g., native, modified, or heterologous 5' and/or 3' UTRs), and the like. Examples of suitable polypeptide tags may include, but are not limited, to any combination of purification tags, such as his-tags, flag-tags, maltose binding protein and glutathione-S-transferase tags, detection tags, such as tags that may be detected photometrically (e.g., green fluorescent protein, red fluorescent protein, etc.) and tags that have a detectable enzymatic activity (e.g., alkaline phosphatase, etc.), tags containing secretory sequences, signal sequences, leader sequences, and/or stabilizing sequences, protease cleavage sites (e.g., furin cleavage sites, TEV cleavage sites, Thrombin cleavage sites, etc.), and the like. In some embodiments, the 5' and/or 3'UTRs increase the stability, localization, and/or translational effi-

ciency of the polynucleotides. In some embodiments, the 5' and/or 3'UTRs improve the level and/or duration of protein expression. In some embodiments, the 5' and/or 3'UTRs include elements (e.g., one or more miRNA binding sites, etc.) that may block or reduce off-target expression (e.g., inhibiting expression in specific cell types (e.g., neuronal cells), at specific times in the cell cycle, at specific developmental stages, etc.). In some embodiments, the 5' and/or 3'UTRs include elements (e.g., one or more miRNA binding sites, etc.) that may enhance cosmetic protein expression in specific cell types (such as human keratinocytes and/or fibroblasts).

[0167] In some embodiments, a polynucleotide of the present disclosure encoding a cosmetic protein (e.g., a human collagen protein) is operably linked to one or more (e.g., one or more, two or more, three or more, four or more, five or more, ten or more, etc.) regulatory sequences. The term "regulatory sequence" may include enhancers, insulators, promoters, and other expression control elements (e.g., polyadenylation signals). Any suitable enhancer(s) known in the art may be used, including, for example, enhancer sequences from mammalian genes (such as globin, elastase, albumin, α -fetoprotein, insulin and the like), enhancer sequences from a eukaryotic cell virus (such as SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, adenovirus enhancers, and the like), and any combinations thereof. Any suitable insulator(s) known in the art may be used, including, for example, HSV chromatin boundary (CTRL/CTCF-binding/insulator) elements CTRL1 and/or CTRL2, chicken hypersensitive site 4 insulator (cHS4), human HNRPA2B1-CBX3 ubiquitous chromatin opening element (UCOE), the scaffold/matrix attachment region (S/MAR) from the human interferon beta gene (IFNB1), and any combinations thereof. Any suitable promoter (e.g., suitable for transcription in mammalian host cells) known in the art may be used, including, for example, promoters obtained from the genomes of viruses (such as polyoma virus, fowlpox virus, adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus, Simian Virus 40 (SV40), and the like), promoters from heterologous mammalian genes (such as the actin promoter (e.g., the 3-actin promoter), a ubiquitin promoter (e.g., a ubiquitin C (UbC) promoter), a phosphoglycerate kinase (PGK) promoter, an immunoglobulin promoter, from heat-shock promoters, and the like), promoters from homologous mammalian genes (e.g., native human collagen, fibronectin, elastin, lumican, vitronectin, laminin, and/or fibrillin promoters), synthetic promoters (such as the CAGG promoter), and any combinations thereof, provided such promoters are compatible with the host cells. Regulatory sequences may include those which direct constitutive expression of a nucleic acid, as well as tissue-specific regulatory and/or inducible or repressible sequences.

[0168] In some embodiments, a polynucleotide of the present disclosure encoding cosmetic protein (e.g., a human collagen protein) is operably linked to one or more heterologous promoters. In some embodiments, the one or more heterologous promoters are one or more of constitutive promoters, tissue-specific promoters, temporal promoters, spatial promoters, inducible promoters and repressible promoters. In some embodiments, the one or more heterologous promoters are one or more of the human cytomegalovirus

(HCMV) immediate early promoter, the human elongation factor-1 (EF1) promoter, the human β -actin promoter, the human UbC promoter, the human PGF promoter, the synthetic CAGG promoter, and any combinations thereof. In some embodiments, a polynucleotide of the present disclosure encoding a cosmetic protein (e.g., a human collagen protein) is operably linked to an HCMV promoter.

[0169] In some embodiments, a polynucleotide of the present disclosure does not comprise the coding sequence of (e.g., a transgene encoding) a Collagen alpha-1 (VII) chain polypeptide (COL7). In some embodiments, a polynucleotide of the present disclosure does not comprise the coding sequence of (e.g., a transgene encoding) a Lysyl hydroxylase 3 polypeptide (LH3). In some embodiments, a polynucleotide of the present disclosure does not comprise the coding sequence of (e.g., a transgene encoding) a Keratin type I cytoskeletal 17 polypeptide (KRT17). In some embodiments, a polynucleotide of the present disclosure does not comprise the coding sequence of (e.g., a transgene encoding) a transglutaminase (TGM) polypeptide (e.g., a human transglutaminase polypeptide such as a human TGM1 polypeptide). In some embodiments, a polynucleotide of the present disclosure does not comprise the coding sequence of (e.g., a transgene encoding) a laminin subunit beta-3 polypeptide (LAMB3). In some embodiments, a polynucleotide of the present disclosure does not comprise the coding sequence of (e.g., a transgene encoding) a Collagen alpha-1 (VII) chain polypeptide, a Lysyl hydroxylase 3 polypeptide, a Keratin type I cytoskeletal 17 polypeptide, and/or any chimeric polypeptides thereof. In some embodiments, a polynucleotide of the present disclosure does not comprise the coding sequence of (e.g., a transgene encoding) a Collagen alpha-1 (VII) chain polypeptide, a Lysyl hydroxylase 3 polypeptide, a Keratin type I cytoskeletal 17 polypeptide, a transglutaminase (TGM) polypeptide (e.g., a human transglutaminase polypeptide such as a human TGM1 polypeptide), a laminin subunit beta-3 (LAMB3) polypeptide (e.g., a human LamB3 polypeptide) and/or any chimeric polypeptides thereof.

Cosmetic Proteins

Collagen Proteins

[0170] In some embodiments, the present disclosure relates to one or more polynucleotides encoding a full-length collagen protein or any isoforms or portions thereof. Any collagen protein from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, human collagen proteins (see e.g., UniProt accession numbers P02452, P08123, P02461, P02462, P08572, P12109, Q02388, Q9UMD9, etc.), mouse collagen proteins (see, e.g., UniProt accession numbers P11087, Q01149, P08121, P02463, P08122, Q04857, Q63870, Q07563, etc.), chimpanzee collagen proteins (see e.g., UniProt accession numbers A0A2I3SM98, A0A2J8L483, H2QJ46, K7C8P4, K7C8W0, A0A2J8M8U9, H2QMJ5, H2Q2J4, etc.), rat collagen proteins (see e.g., UniProt accession numbers P02454, P02466, P13941, P02466, F1M6Q3, D3ZUL3, D3ZE04, D3ZE04, etc.), rabbit collagen proteins (see e.g., UniProt accession numbers G1T4A5, Q28668, G1T8J0, G1U9R7, G1T548, G1T380, G1T548, etc.) etc. Methods of identifying collagen protein homologs/orthologs from additional species are known to one of ordinary skill in the art, including, for example, using an amino acid sequence alignment program such as the

BLAST® blastp suite or OrthoDB. In some embodiments, a collagen polypeptide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the collagen polypeptides described herein or known in the art.

[0171] In some embodiments, the present disclosure relates to one or more polynucleotides encoding a human collagen protein. Any suitable human collagen protein known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a Collagen alpha-1(I) chain polypeptide (COL1-1) (see e.g., UniProt accession number P02452; SEQ ID NO: 15), a Collagen alpha-2(I) chain polypeptide (COL1-2) (see e.g., UniProt accession number P08123; SEQ ID NO: 16), a Collagen alpha-1(II) chain polypeptide (COL2) (see e.g., UniProt accession number P02458), a Collagen alpha-1(III) chain polypeptide (COL3) (see e.g., UniProt accession number P2461; SEQ ID NO: 17), a Collagen alpha-1(IV) chain polypeptide (COL4-1) (see e.g., UniProt accession number P02462; SEQ ID NO: 18), a Collagen alpha-2(IV) chain polypeptide (COL4-2) (see e.g., UniProt accession number P08572), a Collagen alpha-3(IV) chain polypeptide (COL4-3) (see e.g., UniProt accession number Q01955), a Collagen alpha-4(IV) chain polypeptide (COL4-4) (see e.g., UniProt accession number P53420), a Collagen alpha-5(IV) chain polypeptide (COL4-5) (see e.g., UniProt accession number 29400), a Collagen alpha-6(IV) chain polypeptide (COL4-6) (see e.g., UniProt accession number Q14031), a Collagen alpha-1(V) chain polypeptide (COL5-1) (see e.g., UniProt accession number P20908), a Collagen alpha-2(V) chain polypeptide (COL5-2) (see e.g., UniProt accession number P05997), a Collagen alpha-3(V) chain polypeptide (COL5-3) (see e.g., UniProt accession number P25940), a Collagen alpha-1(VI) chain polypeptide (COL6-1) (see e.g., UniProt accession number P12109; SEQ ID NO: 19), a Collagen alpha-2(VI) chain polypeptide (COL6-2) (see e.g., UniProt accession number P12110), a Collagen alpha-3(VI) chain polypeptide (COL6-3) (see e.g., UniProt accession number P12111), a Collagen alpha-4(VI) chain polypeptide (COL6-4), a Collagen alpha-5(VI) chain polypeptide (COL6-5) (see e.g., UniProt accession number A8TX70), a Collagen alpha-6(VI) chain polypeptide (COL6-6) (see e.g., UniProt accession number A6NMZ7), a Collagen alpha-1(VII) chain polypeptide (COL7) (see e.g., UniProt accession number Q02388; SEQ ID NO: 20), a Collagen alpha-1(VIII) chain polypeptide (COL8) (see e.g., UniProt accession number P27658), a Collagen alpha-1(IX) chain polypeptide (COL9-1) (see e.g., UniProt accession number P20849), a Collagen alpha-2(IX) chain polypeptide (COL9-2) (see e.g., UniProt accession number Q14055), a Collagen alpha-3(IX) chain polypeptide (COL9-3) (see e.g., UniProt accession number Q14050), a Collagen alpha-1(X) chain polypeptide (COL10) (see e.g., UniProt accession number Q03692), a Collagen alpha-1(XI) chain polypeptide (COL11-1) (see e.g., UniProt accession number P12107), a Collagen alpha-2(XI) chain polypeptide (COL11-2) (see e.g., UniProt accession number P13942), a Collagen alpha-1(XII) chain polypeptide (COL12) (see e.g., UniProt accession number Q99715), a Collagen alpha-1(XIII) chain polypeptide (COL13) (see e.g., UniProt accession number Q5TAT6), a Collagen alpha-

1(XIV) chain polypeptide (COL14) (see e.g., UniProt accession number Q05707), a Collagen alpha-1(XV) chain polypeptide (COL15) (see e.g., UniProt accession number P39059), a Collagen alpha-1(XVI) chain polypeptide (COL16) (see e.g., UniProt accession number Q07092), a Collagen alpha-1(XVII) chain polypeptide (COL17) (see e.g., UniProt accession number Q9UMD9; SEQ ID NO: 21), a Collagen alpha-1(XVIII) chain polypeptide (COL18) (see e.g., UniProt accession number P39060), a Collagen alpha-1(XIX) chain polypeptide (COL19) (see e.g., UniProt accession number Q14993), a Collagen alpha-1(XX) chain polypeptide (COL20) (see e.g., UniProt accession number Q9P218), a Collagen alpha-1(XXI) chain polypeptide (COL21) (see e.g., UniProt accession number Q96P44), a Collagen alpha-1(XXII) chain polypeptide (COL22) (see e.g., UniProt accession number Q8NFW1), a Collagen alpha-1(XXIII) chain polypeptide (COL23) (see e.g., UniProt accession number Q86Y22), a Collagen alpha-1(XXIV) chain polypeptide (COL24) (see e.g., UniProt accession number Q17RW2), a Collagen alpha-1(XXV) chain polypeptide (COL25) (see e.g., UniProt accession number Q9BXS0), a Collagen alpha-1(XXVI) chain polypeptide (COL26) (see e.g., UniProt accession number Q96A83), a Collagen alpha-1(XXVII) chain polypeptide (COL27) (see e.g., UniProt accession number Q8IZC6), a Collagen alpha-1(XXVIII) chain polypeptide (COL28) (see e.g., UniProt accession number Q2UY09), etc. In some embodiments, a polynucleotide of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a sequence encoding any of the human collagen polypeptides described herein or known in the art. Methods of identifying additional human collagen or collagen-like polypeptide homologs/orthologs are known to one of ordinary skill in the art, including, for example, using an amino acid sequence alignment program such as the BLAST® blastp suite or OrthoDB.

[0172] In some embodiments, a polynucleotide of the present disclosure encodes a human COL1-1 protein. In some embodiments, a polynucleotide encoding a COL1-1 protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 15. In some embodiments, a polynucleotide encoding a human COL1-1 protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 15.

[0173] In some embodiments, a polynucleotide encoding a COL1-1 protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 15. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, at least 1200, at least 1300, at least 1400, but fewer than 1464, consecutive amino acids of SEQ ID NO: 15.

[0184] In some embodiments, a polynucleotide of the present disclosure encodes a human COL17 protein. In some embodiments, a polynucleotide encoding a COL17 protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 21. In some embodiments, a polynucleotide encoding a human COL17 protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 21.

[0185] In some embodiments, a polynucleotide encoding a COL17 protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 21. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, at least 1200, at least 1300, at least 1400, but fewer than 1497, consecutive amino acids of SEQ ID NO: 21.

[0186] In some embodiments, one or more human collagen proteins of the present disclosure (e.g., a first human collagen protein, a further human collagen protein, an additional human collagen protein, and/or a second human collagen protein) comprise an amino acid sequence comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21. In some embodiments, one or more human collagen proteins of the present disclosure (e.g., a first human collagen protein, a further human collagen protein, an additional human collagen protein, and/or a second human collagen protein) comprise a sequence selected from SEQ ID NOS: 15-21.

[0187] In some embodiments, one or more human collagen proteins of the present disclosure (e.g., a first human collagen protein, a further human collagen protein, an additional human collagen protein, and/or a second human collagen protein) comprise an amino acid sequence comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-17. In some embodiments, one or more human collagen proteins of the present disclosure (e.g., a first human collagen protein, a further human collagen protein, an additional human collagen protein, and/or a second human collagen protein) comprise a sequence selected from SEQ ID NOS: 15-17.

Fibronectin Proteins

[0188] In some embodiments, the present disclosure relates to one or more polynucleotides encoding a full-length fibronectin protein or any isoforms or portions thereof. Any fibronectin protein from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human fibronectin protein (see e.g., UniProt accession number P02751), a

mouse fibronectin protein (see, e.g., UniProt accession number P11276), a chimpanzee fibronectin protein (see e.g., UniProt accession number P11276), a rat fibronectin protein (see e.g., UniProt accession number P04937), a rabbit fibronectin protein (see e.g., UniProt accession number P04937), etc. Methods of identifying fibronectin protein homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a fibronectin protein of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the fibronectin proteins described herein or known in the art.

[0189] In some embodiments, a polynucleotide of the present disclosure encodes a human fibronectin protein. In some embodiments, a polynucleotide encoding a human fibronectin protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 53. In some embodiments, a polynucleotide encoding a human fibronectin protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 53.

[0190] In some embodiments, a polynucleotide encoding a human fibronectin protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 53. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, at least 1200, at least 1300, at least 1400, at least 1500, at least 1600, at least 1700, at least 1800, at least 1900, at least 2000, at least 2100, at least 2200, at least 2300, at least 2400, but fewer than 2477, consecutive amino acids of SEQ ID NO: 53.

Elastin and Associated Proteins

[0191] Elastic fibers in the extracellular matrix give elastic properties to the tissue. The elastic fibers generally contain two morphologically distinct components—the mature elastin fibers, and the micro-fibrils which mainly contain fibrillin and are associated with further proteins such as the micro-fibrils associated glycoproteins (MAGPs), fibulin, and the elastin-micro-fibrils-interface localized proteins (EMILIN). Elastin and its soluble precursor tropoelastin belong to the major structural proteins of the body.

[0192] In some embodiments, the present disclosure relates to one or more polynucleotides encoding an elastin or elastin-associated protein, including a tropoelastin, a fibrillin, a micro-fibrils associated glycoprotein, a fibulin, or an elastin-micro-fibrils-interface localized protein. In some embodiments, the present disclosure relates to one or more polynucleotides encoding a full-length elastin protein or any isoforms or portions thereof. Any elastin protein from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human elastin protein (see e.g., UniProt acces-

sion number P15502), a mouse elastin protein (see, e.g., UniProt accession number P15502), a chimpanzee elastin protein (see e.g., UniProt accession number H2QUQ6), a rat elastin protein (see e.g., UniProt accession number Q99372), etc. Methods of identifying elastin protein homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, an elastin protein of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the elastin proteins described herein or known in the art.

[0193] In some embodiments, a polynucleotide of the present disclosure encodes a human elastin protein. In some embodiments, a polynucleotide encoding a human elastin protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 54. In some embodiments, a polynucleotide encoding a human elastin protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 54.

[0194] In some embodiments, a polynucleotide encoding a human elastin protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 54. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, but fewer than 786, consecutive amino acids of SEQ ID NO: 54.

Lumican Proteins

[0195] In some embodiments, the present disclosure relates to one or more polynucleotides encoding a full-length lumican protein or any isoforms or portions thereof. Any lumican protein from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human lumican protein (see e.g., UniProt accession number P51884), a mouse lumican protein (see, e.g., UniProt accession number P51885), a chimpanzee lumican protein (see e.g., UniProt accession number H2Q6L3), a rat lumican protein (see e.g., UniProt accession number H2Q6L3), a rabbit lumican protein (see e.g., UniProt accession number 046379), etc. Methods of identifying lumican protein homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a lumican protein of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the lumican proteins described herein or known in the art.

[0196] In some embodiments, a polynucleotide of the present disclosure encodes a human lumican protein. In some embodiments, a polynucleotide encoding a human

lumican protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 55. In some embodiments, a polynucleotide encoding a human lumican protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 55.

[0197] In some embodiments, a polynucleotide encoding a human lumican protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 55. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, but fewer than 338, consecutive amino acids of SEQ ID NO: 55.

Vitronectin and Vitronectin Receptor Proteins

[0198] In some embodiments, the present disclosure relates to one or more polynucleotides encoding a full-length vitronectin or vitronectin receptor protein or any isoforms or portions thereof. Any vitronectin or vitronectin receptor protein from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human vitronectin or vitronectin receptor protein (see e.g., UniProt accession numbers P04004 and P06756), a mouse vitronectin or vitronectin receptor protein (see, e.g., UniProt accession numbers P29788 and P43406), a chimpanzee vitronectin or vitronectin receptor protein (see e.g., UniProt accession numbers H2QCH3 and H2R6C3), a rat vitronectin or vitronectin receptor protein (see e.g., UniProt accession number Q7TQ11), a rabbit vitronectin or vitronectin receptor protein (see e.g., UniProt accession number P22458), etc. Methods of identifying vitronectin or vitronectin receptor protein homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a vitronectin or vitronectin receptor protein of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the vitronectin or vitronectin receptor proteins described herein or known in the art.

[0199] In some embodiments, a polynucleotide of the present disclosure encodes a human vitronectin protein. In some embodiments, a polynucleotide encoding a human vitronectin protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 56. In some embodiments, a polynucleotide encoding a human vitronectin protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 56.

[0200] In some embodiments, a polynucleotide encoding a human vitronectin protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a

fragment of the amino acid sequence of SEQ ID NO: 56. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, but fewer than 478, consecutive amino acids of SEQ ID NO: 56.

Laminin Proteins

[0201] In some embodiments, the present disclosure relates to one or more polynucleotides encoding a full-length laminin protein or any isoforms or portions thereof. Any laminin protein from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human laminin protein (see e.g., UniProt accession numbers P25391, P24043, Q16787, Q16363, 015230, P07942, P55268, Q13751, P11047, Q13753, and Q9Y6N6), a mouse laminin protein (see e.g., UniProt accession numbers Q61789, Q61087, and Q61092), a chimpanzee laminin protein (see e.g., UniProt accession numbers H2QEC7, H2R041, and H2Q0R2), a rat laminin protein (see e.g., UniProt accession numbers D3ZN05, F1LP15, and F1LRH4), a rabbit laminin protein (see e.g., UniProt accession numbers G1SY40 and A0A0B5JSH0), etc. Methods of identifying laminin protein homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a laminin protein of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the laminin proteins described herein or known in the art.

[0202] In some embodiments, a polynucleotide of the present disclosure encodes a human laminin protein, such as a human Laminin subunit alpha-1 (LamA1) polypeptide (see e.g., UniProt accession number P25391), a human Laminin subunit alpha-2 (LamA2) polypeptide (see e.g., UniProt accession number P24043), a human Laminin subunit alpha-3 (LamA3) polypeptide (see e.g., UniProt accession number Q16787), a human Laminin subunit alpha-4 (LamA4) polypeptide (see e.g., UniProt accession number Q16363), a human Laminin subunit alpha-5 (LamA5) polypeptide (see e.g., UniProt accession number 015230), a human Laminin subunit beta-1 (LamB1) polypeptide (see e.g., UniProt accession number P07942), a human Laminin subunit beta-2 (LamB2) polypeptide (see e.g., UniProt accession number P55268), a human Laminin subunit beta-3 (LamB3) polypeptide (see e.g., UniProt accession number Q13751), a human Laminin subunit gamma-1 (LamC1) polypeptide (see e.g., UniProt accession number P11047), a human Laminin subunit gamma-2 (LamC2) polypeptide (see e.g., UniProt accession number Q13753), a human Laminin subunit gamma-3 (LamC3) polypeptide (see e.g., UniProt accession number Q9Y6N6), etc.

[0203] In some embodiments, a polynucleotide of the present disclosure encodes a human LamA3 polypeptide. In some embodiments, a polynucleotide encoding a human LamA3 polypeptide is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least

96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 57. In some embodiments, a polynucleotide encoding a human LamA3 polypeptide is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 57.

[0204] In some embodiments, a polynucleotide encoding a human LamA3 polypeptide is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 57. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2250, at least 2500, at least 2750, at least 3000, at least 3250, but fewer than 3333, consecutive amino acids of SEQ ID NO: 57.

[0205] In some embodiments, a polynucleotide of the present disclosure encodes a human LamB3 polypeptide. In some embodiments, a polynucleotide encoding a human LamB3 polypeptide is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 58. In some embodiments, a polynucleotide encoding a human LamB3 polypeptide is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 58.

[0206] In some embodiments, a polynucleotide encoding a LamB3 polypeptide is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 58. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, but fewer than 1172, consecutive amino acids of SEQ ID NO: 58.

[0207] In some embodiments, a polynucleotide of the present disclosure encodes a human LamC2 polypeptide. In some embodiments, a polynucleotide encoding a human LamC2 polypeptide is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 59. In some embodiments, a polynucleotide encoding a human LamC2 polypeptide is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 59.

[0208] In some embodiments, a polynucleotide encoding a LamC2 polypeptide is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 59. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at

least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, but fewer than 1193, consecutive amino acids of SEQ ID NO: 59.

Neuromodulator Proteins

[0209] In some embodiments, the present disclosure relates to one or more polynucleotides encoding a full-length neuromodulator protein or any isoforms or portions thereof. Any neuromodulator protein from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a *Clostridium botulinum* protein (see e.g., UniProt accession numbers P0DPI0, Q45894, P0DPI1, P10844, and B1INP5), etc. Methods of identifying neuromodulator protein homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a neuromodulator protein of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the neuromodulator proteins described herein or known in the art.

[0210] In some embodiments, a polynucleotide of the present disclosure encodes a *Clostridium botulinum* neuromodulator protein.

[0211] In some embodiments, a polynucleotide of the present disclosure encodes a *Clostridium botulinum* neurotoxin type A protein. In some embodiments, a polynucleotide encoding a *Clostridium botulinum* neurotoxin type A protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 60. In some embodiments, a polynucleotide encoding a *Clostridium botulinum* neurotoxin type A protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 60. In some embodiments, a *Clostridium botulinum* neurotoxin type A protein of the present disclosure comprises an alanine to valine mutation at a position corresponding to position 27 of SEQ ID NO: 60.

[0212] In some embodiments, a polynucleotide encoding a *Clostridium botulinum* neurotoxin type A protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 60. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, at least 1200, but fewer than 1296, consecutive amino acids of SEQ ID NO: 60.

[0213] In some embodiments, a polynucleotide of the present disclosure encodes a *Clostridium botulinum* neurotoxin type B protein. In some embodiments, a polynucleotide encoding a *Clostridium botulinum* neurotoxin type B protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at

least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 61. In some embodiments, a polynucleotide encoding a *Clostridium botulinum* neurotoxin type B protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 61.

[0214] In some embodiments, a polynucleotide encoding a *Clostridium botulinum* neurotoxin type B protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 61. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, at least 1200, but fewer than 1291, consecutive amino acids of SEQ ID NO: 61.

Fibrillin Proteins

[0215] In some embodiments, the present disclosure relates to one or more polynucleotides encoding a full-length fibrillin protein or any isoforms or portions thereof. Any fibrillin protein from any suitable species known in the art may be encoded by a polynucleotide of the present disclosure, including, for example, a human fibrillin protein (see e.g., UniProt accession numbers P35555, P35556, and Q75N90), a mouse fibrillin protein (see, e.g., UniProt accession numbers Q61554 and Q61555), a chimpanzee fibrillin protein (see e.g., UniProt accession numbers A0A2I3RTE4 and K7CZX0), a rat fibrillin protein (see e.g., UniProt accession number G3V9M6 and F1M5Q4), a rabbit fibrillin protein (see e.g., UniProt accession number G1SKM2, G1SUS5, and G1T1H4), etc. Methods of identifying fibrillin protein homologs/orthologs from additional species are known to one of ordinary skill in the art. In some embodiments, a fibrillin protein of the present disclosure comprises a sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of any of the fibrillin proteins described herein or known in the art.

[0216] In some embodiments, a polynucleotide of the present disclosure encodes a human fibrillin protein.

[0217] In some embodiments, a polynucleotide of the present disclosure encodes a human fibrillin-1 protein. In some embodiments, a polynucleotide encoding a human fibrillin-1 protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 62. In some embodiments, a polynucleotide encoding a human fibrillin-1 protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 62.

[0218] In some embodiments, a polynucleotide encoding a human fibrillin-1 protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a frag-

ment of the amino acid sequence of SEQ ID NO: 62. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2250, at least 2500, at least 2750, but fewer than 2871, consecutive amino acids of SEQ ID NO: 62.

[0219] In some embodiments, a polynucleotide of the present disclosure encodes a human fibrillin-2 protein. In some embodiments, a polynucleotide encoding a human fibrillin-2 protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 63. In some embodiments, a polynucleotide encoding a human fibrillin-2 protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 63.

[0220] In some embodiments, a polynucleotide encoding a human fibrillin-2 protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 63. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2250, at least 2500, at least 2750, but fewer than 2912, consecutive amino acids of SEQ ID NO: 63.

[0221] In some embodiments, a polynucleotide of the present disclosure encodes a human fibrillin-3 protein. In some embodiments, a polynucleotide encoding a human fibrillin-3 protein is a polynucleotide that encodes a polypeptide comprising an amino acid sequence having at least 75%, at least 80%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to the sequence of SEQ ID NO: 64. In some embodiments, a polynucleotide encoding a human fibrillin-3 protein is a polynucleotide that encodes a polypeptide comprising the amino acid sequence of SEQ ID NO: 64.

[0222] In some embodiments, a polynucleotide encoding a human fibrillin-3 protein is a polynucleotide that encodes an N-terminal truncation, a C-terminal truncation, or a fragment of the amino acid sequence of SEQ ID NO: 64. N-terminal truncations, C-terminal truncations, or fragments may comprise at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 30, at least 40, at least 50, at least 75, at least 100, at least 200, at least 300, at least 400, at least 500, at least 750, at least 1000, at least 1250, at least 1500, at least 1750, at least 2000, at least 2250, at least 2500, at least 2750, but fewer than 2809, consecutive amino acids of SEQ ID NO: 64.

Exemplary Cosmetic Polypeptides

[0223] In some embodiments, one or more cosmetic proteins of the present disclosure (e.g., a first cosmetic protein, a further cosmetic protein, an additional cosmetic protein,

and/or a second cosmetic protein) comprise an amino acid sequence comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21 or 53-64. In some embodiments, one or more cosmetic proteins of the present disclosure (e.g., a first cosmetic protein, a further cosmetic protein, an additional cosmetic protein, and/or a second cosmetic protein) comprises a sequence selected from SEQ ID NOS: 15-21 or 53-64.

[0224] In some embodiments, one or more cosmetic proteins of the present disclosure (e.g., a first cosmetic protein, a further cosmetic protein, an additional cosmetic protein, and/or a second cosmetic protein) comprise an amino acid sequence comprising at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 15-21, 53-54, or 57-59. In some embodiments, one or more cosmetic proteins of the present disclosure (e.g., a first cosmetic protein, a further cosmetic protein, an additional cosmetic protein, and/or a second cosmetic protein) comprise a sequence selected from SEQ ID NOS: 15-21, 53-54, or 57-59.

First Polynucleotides

[0225] In some embodiments, the present disclosure relates to a recombinant nucleic acid comprising a first polynucleotide encoding a first polypeptide comprising a first cosmetic protein. The first cosmetic protein may be any of the cosmetic proteins described herein or known in the art, including, for example a collagen protein, a fibronectin, an elastin, a lumican, a vitronectin/vitronectin receptor, a laminin, a neuromodulator, a fibrillin, etc. In some embodiments, the first cosmetic protein is a structural extracellular matrix protein (e.g., a collagen, elastin, fibronectin, laminin, fibrillin, etc.). In some embodiments, the first cosmetic protein is a collagen, elastin, fibronectin, or laminin protein (e.g., a human collagen, elastin, fibronectin, or laminin protein).

[0226] In some embodiments, a recombinant nucleic acid of the present disclosure comprises one copy of the first polynucleotide. In some embodiments, a recombinant nucleic acid of the present disclosure comprises two or more (e.g., two or more, three or more, four or more, five or more, ten or more, etc.) copies of the first polynucleotide. In some embodiments, a recombinant nucleic acid of the present disclosure comprises two copies of the first polynucleotide.

[0227] In some embodiments, the first cosmetic protein is a first human collagen protein. The first human collagen protein may be any of the human collagen proteins described herein or known in the art. In some embodiments, the first human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, or COL28. In some embodiments, the first human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL1-1.

In some embodiments, the first human collagen protein is COL1-2. In some embodiments, the first human collagen protein is COL3. In some embodiments, the first human collagen protein is COL4-1. In some embodiments, the first human collagen protein is COL4-2. In some embodiments, the first human collagen protein is COL6-1. In some embodiments, the first human collagen protein is COL7. In some embodiments, the first human collagen protein is not COL7. In some embodiments, the first human collagen protein is COL17.

[0228] In some embodiments, the first polypeptide consists essentially of the first cosmetic protein. In some embodiments, the first polypeptide consists of the first cosmetic protein. In some embodiments, the first polypeptide is the first cosmetic protein.

Chimeric Polypeptides

[0229] In some embodiments, the first polypeptide is a chimeric polypeptide comprising the first cosmetic protein. In some embodiments, the first polypeptide is a chimeric polypeptide comprising the first cosmetic protein and a further cosmetic protein. In some embodiments, the chimeric polypeptide comprises a linker polypeptide linking the first cosmetic protein and the further cosmetic protein. In some embodiments, the chimeric polypeptide comprises, from n-terminus to c-terminus, the first cosmetic protein—the linker polypeptide—the further cosmetic protein. The first and/or further cosmetic proteins may be any of the cosmetic proteins described herein or known in the art, including, for example a collagen protein, a fibronectin, an elastin, a lumican, a vitronectin/vitronectin receptor, a laminin, a neuromodulator, a fibrillin, etc. In some embodiments, the first and/or further cosmetic protein is a structural extracellular matrix protein (e.g., a collagen, elastin, fibronectin, laminin, fibrillin, etc.). In some embodiments, the first and/or further cosmetic protein is a collagen, elastin, fibronectin, or laminin protein (e.g., a human collagen, elastin, fibronectin, or laminin protein). In some embodiments, the first and further cosmetic proteins are the same. In some embodiments, the first and further cosmetic proteins are different.

[0230] In some embodiments, the linker polypeptide is a cleavable linker polypeptide. Any cleavable linker polypeptide known in the art may be used in the chimeric polypeptides of the present disclosure, including, for example, a T2A linker, a P2A linker, an E2A linker, and F2A linker, etc. In some embodiments, the linker polypeptide is a T2A linker polypeptide. An exemplary nucleic acid sequence encoding a T2A linker polypeptide is provided as SEQ ID NO: 24. An exemplary amino acid sequence of a T2A linker polypeptide is provided as SEQ ID NO: 28. In some embodiments, the linker polypeptide is a P2A linker polypeptide. An exemplary nucleic acid sequence encoding a P2A linker polypeptide is provided as SEQ ID NO: 25. An exemplary amino acid sequence of a P2A linker polypeptide is provided as SEQ ID NO: 29. In some embodiments, the linker polypeptide is an E2A linker polypeptide. An exemplary nucleic acid sequence encoding an E2A linker polypeptide is provided as SEQ ID NO: 26. An exemplary amino acid sequence of an E2A linker polypeptide is provided as SEQ ID NO: 30. In some embodiments, the linker polypeptide is an F2A linker polypeptide. An exemplary nucleic acid sequence encoding an F2A linker polypeptide is provided as SEQ ID NO: 27.

An exemplary amino acid sequence of an F2A linker polypeptide is provided as SEQ ID NO: 31.

[0231] In some embodiments, the linker polypeptide comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from SEQ ID NOS: 28-31. In some embodiments, the linker polypeptide comprises a sequence selected from SEQ ID NOS: 28-31.

[0232] In some embodiments, the first cosmetic protein is a first collagen protein (e.g., a first human collagen protein), and the further cosmetic protein is a further collagen protein (e.g., a further human collagen protein). An exemplary nucleic acid sequence encoding a chimeric polypeptide comprising a first human collagen protein, a linker polypeptide, and a further human collagen protein is provided as SEQ ID NO: 32.

[0233] In some embodiments, the first cosmetic protein is a first human collagen protein, and the further cosmetic protein is a further human collagen protein. The further human collagen protein may be any of the human collagen proteins described herein or known in the art. In some embodiments, the further human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, or COL28. In some embodiments, the further human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the further human collagen protein is COL1-1. In some embodiments, the further human collagen protein is COL1-2. In some embodiments, the further human collagen protein is COL3. In some embodiments, the further human collagen protein is COL4-1. In some embodiments, the further human collagen protein is COL4-2. In some embodiments, the further human collagen protein is COL6-1. In some embodiments, the further human collagen protein is COL7. In some embodiments, the further human collagen protein is not COL7. In some embodiments, the further human collagen protein is COL17. In some embodiments, the first human collagen protein and the further human collagen protein are the same. In some embodiments, the first human collagen protein and the further human collagen protein are different.

[0234] In some embodiments, the first human collagen protein is COL1-1, and the further human collagen protein is selected from COL1-2, COL3, COL4-1, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL1-1, and the further human collagen protein is COL1-2. In some embodiments, the first human collagen protein is COL1-1, and the further human collagen protein is COL3.

[0235] In some embodiments, the first human collagen protein is COL1-2, and the further human collagen protein is COL1-1, COL3, COL4-1, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL1-2, and the further human collagen protein is COL1-1.

[0236] In some embodiments, the first human collagen protein is COL3, and the further human collagen protein is selected from COL1-1, COL1-2, COL4-1, COL4-2, COL6-1, COL7, or COL17.

[0237] In some embodiments, the first human collagen protein is COL4-1, and the further human collagen protein is COL1-1, COL1-2, COL3, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL4-1, and the further human collagen protein is COL4-2.

[0238] In some embodiments, the first human collagen protein is COL6-1, and the further human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL7, or COL17.

[0239] In some embodiments, the first human collagen protein is COL7, and the further human collagen protein is COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, or COL17.

[0240] In some embodiments, the first human collagen protein is COL17, and the further human collagen protein is COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, or COL7.

[0241] In some embodiments, the first cosmetic protein is a first laminin protein (e.g., a first human laminin protein), and the further cosmetic protein is a further laminin protein (e.g., a further human laminin protein). In some embodiments, the first cosmetic protein is a first human laminin protein, and the further cosmetic protein is a further human laminin protein. The further human laminin protein may be any of the human laminin proteins described herein or known in the art. In some embodiments, the first human laminin protein is a human LamA3 polypeptide and the further human laminin protein is a human LamB3 polypeptide. In some embodiments, the first human laminin protein is a human LamA3 polypeptide and the further human laminin protein is a human LamC2 polypeptide. In some embodiments, the first human laminin protein is a human LamB3 polypeptide and the further human laminin protein is a human LamC2 polypeptide.

[0242] In some embodiments, the first polynucleotide encodes a monocistronic mRNA. In some embodiments, the monocistronic mRNA comprises an open reading frame (ORF) encoding the first polypeptide.

[0243] In some embodiments, the first polynucleotide encodes a polycistronic mRNA. In some embodiments, the polycistronic mRNA comprises an open reading frame (ORF) encoding the first polypeptide.

Polycistronic mRNA

[0244] In some embodiments, the first polynucleotide encodes a polycistronic mRNA. In some embodiments, the polycistronic mRNA comprises an open reading frame (ORF) encoding the first polypeptide. In some embodiments, the first polynucleotide encodes a polycistronic mRNA comprising: 1) a first open reading frame (ORF) encoding the first polypeptide, and 2) a second open reading frame (ORF) encoding an additional cosmetic protein. In some embodiments, the polycistronic mRNA further comprises an internal ribosomal entry site (IRES) separating the first ORF and the second ORF. In some embodiments, the polycistronic mRNA comprises, from 5' to 3', the first ORF encoding the first polypeptide—the IRES—the second ORF encoding the additional cosmetic protein. The first polypeptide may be any of the first polypeptides described herein. The additional cosmetic protein may be any of the cosmetic

proteins described herein or known in the art, including, for example a collagen protein, a fibronectin, an elastin, a lumican, a vitronectin/vitronectin receptor, a laminin, a neuromodulator, a fibrillin, etc. In some embodiments, the additional cosmetic protein is a structural extracellular matrix protein (e.g., a collagen, elastin, fibronectin, laminin, fibrillin, etc.). In some embodiments, the additional cosmetic protein is a collagen, elastin, fibronectin, or laminin protein (e.g., a human collagen, elastin, fibronectin, or laminin protein).

[0245] Any suitable IRES known in the art may be used in the polycistronic mRNAs of the present disclosure, including, for example, a virally-derived IRES (e.g. an IRES derived from a poliovirus, rhinovirus, encephalomyocarditis virus (EMCV), foot-and-mouth disease virus, hepatitis C virus, classic swine fever virus, rous sarcoma virus, human immunodeficiency virus, cricket paralysis virus, Kaposi's sarcoma-associated herpesvirus, etc.), a cellular mRNA-derived IRES (e.g. an IRES derived from growth factor mRNAs, such as fibroblast growth factor 2, platelet-derived growth factor B, and vascular endothelial growth factor; an IRES derived from transcription factor mRNAs, such as antennapedia, ultrabithorax, and NF- κ B repressing factor; an IRES derived from oncogene mRNAs, such as c-myc, pim-1, and protein kinase p58^{PTTSLRE}, etc.), a synthetic IRES (e.g., a CP148 IRES), and others (see e.g., Mokrejs et al. (2007) A Bioinformatical Approach to the Analysis of Viral and Cellular Internal Ribosome Entry Sites. Columbus F editors. New Messenger RNA Research Communications. Hauppauge, NY: Nova Science Publishers; pp. 133-166). In some embodiments, the IRES is a CP148 IRES. An exemplary nucleic acid sequence encoding a CP148 IRES is provided as SEQ ID NO: 22. In some embodiments, the IRES is an EMCV IRES. An exemplary nucleic acid sequence encoding an EMCV IRES is provided as SEQ ID NO: 23.

[0246] In some embodiments, the nucleic acid sequence encoding the IRES comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NO: 22 or SEQ ID NO: 23. In some embodiments, the nucleic acid sequence encoding the IRES comprises the sequence of SEQ ID NO: 22 or SEQ ID NO: 23.

[0247] In some embodiments, the first polypeptide is a first collagen protein (e.g., a first human collagen protein), and the additional cosmetic protein is an additional collagen protein (e.g., an additional human collagen protein). An exemplary nucleic acid encoding a polycistronic mRNA comprising a first ORF, an IRES, and second ORF is provided as SEQ ID NO: 33 or SEQ ID NO: 34. The additional human collagen protein may be any of the human collagen proteins described herein. In some embodiments, the additional human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, or COL28. In some embodiments, the additional human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2,

COL6-1, COL7, or COL17. In some embodiments, the additional human collagen protein is COL1-1. In some embodiments, the additional human collagen protein is COL1-2. In some embodiments, the additional human collagen protein is COL3. In some embodiments, the additional human collagen protein is COL4-1. In some embodiments, the additional human collagen protein is COL4-2. In some embodiments, the additional human collagen protein is COL6-1. In some embodiments, the additional human collagen protein is COL7. In some embodiments, the additional human collagen protein is not COL7. In some embodiments, the additional human collagen protein is COL17. In some embodiments, the first human collagen protein and the additional human collagen protein are the same. In some embodiments, the first human collagen protein and the additional human collagen protein are different.

[0248] In some embodiments, the first human collagen protein is COL1-1, and the additional human collagen protein is selected from COL1-2, COL3, COL4-1, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL1-1, and the additional human collagen protein is COL1-2. In some embodiments, the first human collagen protein is COL1-1, and the additional human collagen protein is COL3.

[0249] In some embodiments, the first human collagen protein is COL1-2, and the additional human collagen protein is selected from COL1-1, COL3, COL4-1, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL1-2, and the additional human collagen protein is COL1-1.

[0250] In some embodiments, the first human collagen protein is COL3, and the additional human collagen protein is selected from COL1-1, COL1-2, COL4-1, COL4-2, COL6-1, COL7, or COL17.

[0251] In some embodiments, the first human collagen protein is COL4-1, and the additional human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL4-1, and the additional human collagen protein is COL4-2.

[0252] In some embodiments, the first human collagen protein is COL6-1, and the additional human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL7, or COL17.

[0253] In some embodiments, the first human collagen protein is COL7, and the additional human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, or COL17.

[0254] In some embodiments, the first human collagen protein is COL17, and the additional human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, or COL7.

[0255] In some embodiments, the first polypeptide is a first collagen protein (e.g., a first human collagen protein), and the additional cosmetic protein is an additional collagen protein (e.g., an additional human collagen protein).

[0256] In some embodiments, the first polypeptide is a first laminin protein (e.g., a first human laminin protein), and the additional cosmetic protein is an additional laminin protein (e.g., an additional human laminin protein). In some embodiments, the first polypeptide is a first human laminin protein, and the additional cosmetic protein is an additional human laminin protein. The additional human laminin protein may be any of the human laminin proteins described

herein or known in the art. In some embodiments, the first human laminin protein is a human LamA3 polypeptide and the additional human laminin protein is a human LamB3 polypeptide. In some embodiments, the first human laminin protein is a human LamA3 polypeptide and the additional human laminin protein is a human LamC2 polypeptide. In some embodiments, the first human laminin protein is a human LamB3 polypeptide and the additional human laminin protein is a human LamC2 polypeptide.

Second Polynucleotides

[0257] In some embodiments, the present disclosure relates to a recombinant nucleic acid further comprising a second polynucleotide encoding a second cosmetic protein. The second cosmetic protein may be any of the cosmetic proteins described herein or known in the art, including, for example a collagen protein, a fibronectin, an elastin, a lumican, a vitronectin/vitronectin receptor, a laminin, a neuromodulator, a fibrillin, etc. In some embodiments, the second cosmetic protein is a structural extracellular matrix protein (e.g., a collagen, elastin, fibronectin, laminin, fibrillin, etc.). In some embodiments, the second cosmetic protein is a collagen, elastin, fibronectin, or laminin protein (e.g., a human collagen, elastin, fibronectin, or laminin protein). In some embodiments, the first and second cosmetic proteins are the same. In some embodiments, the first and second cosmetic proteins are different. In some embodiments, the recombinant nucleic acid comprises one copy of the second polynucleotide. In some embodiments, the recombinant nucleic acid comprises two or more (e.g., two or more, three or more, four or more, five or more, ten or more, etc.) copies of the second polynucleotide. In some embodiments, the recombinant nucleic acid comprises two copies of the second polynucleotide.

[0258] In some embodiments, the second cosmetic protein is a collagen protein. In some embodiments, the second cosmetic protein is a second human collagen protein. The second human collagen protein may be any of the human collagen proteins described herein. In some embodiments, the second human collagen protein is selected from COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, or COL28. In some embodiments, the second human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, or COL17. In some embodiments, the second human collagen protein is COL1-1. In some embodiments, the second human collagen protein is COL1-2. In some embodiments, the second human collagen protein is COL3. In some embodiments, the second human collagen protein is COL4-1. In some embodiments, the second human collagen protein is COL4-2. In some embodiments, the second human collagen protein is COL6-1. In some embodiments, the second human collagen protein is COL7. In some embodiments, the second human collagen protein is not COL7. In some embodiments, the second human collagen protein is COL17.

[0259] In some embodiments, the first polynucleotide encodes a first collagen protein and the second polynucleotide encodes a second collagen protein. In some embodiments, the first polynucleotide encodes a first human colla-

gen protein and the second polynucleotide encodes a second human collagen protein. In some embodiments, the first human collagen protein (encoded by the first polynucleotide) and the second human collagen protein (encoded by the second polynucleotide) are the same. In some embodiments, the first human collagen protein (encoded by the first polynucleotide) and the second human collagen protein (encoded by the second polynucleotide) are different.

[0260] In some embodiments, the first human collagen protein is COL1-1, and the second human collagen protein is selected from COL1-2, COL3, COL4-1, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL1-1, and the second human collagen protein is COL1-2. In some embodiments, the first human collagen protein is COL1-1, and the second human collagen protein is COL3.

[0261] In some embodiments, the first human collagen protein is COL1-2, and the second human collagen protein is selected from COL1-1, COL3, COL4-1, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL1-2, and the second human collagen protein is COL1-1.

[0262] In some embodiments, the first human collagen protein is COL3, and the second human collagen protein is selected from COL1-1, COL1-2, COL4-1, COL4-2, COL6-1, COL7, or COL17.

[0263] In some embodiments, the first human collagen protein is COL4-1, and the second human collagen protein is selected from COL1-2, COL1-2, COL3, COL4-2, COL6-1, COL7, or COL17. In some embodiments, the first human collagen protein is COL4-1, and the second human collagen protein is COL4-2.

[0264] In some embodiments, the first human collagen protein is COL6-1, and the second human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL7, or COL17.

[0265] In some embodiments, the first human collagen protein is COL7, and the second human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, or COL17.

[0266] In some embodiments, the first human collagen protein is COL17, and the second human collagen protein is selected from COL1-1, COL1-2, COL3, COL4-1, COL4-2, COL6-1, or COL7.

[0267] In some embodiments, the first polynucleotide encodes a first laminin protein (e.g., a first human laminin protein), and the second polynucleotide encodes a second laminin protein (e.g., a second human laminin protein). In some embodiments, the first polynucleotide encodes a first human laminin polypeptide and the second polynucleotide encodes a second human laminin protein. The second human laminin protein may be any of the human laminin proteins described herein or known in the art. In some embodiments, the first human laminin protein is a human LamA3 polypeptide and the second human laminin protein is a human LamB3 polypeptide. In some embodiments, the first human laminin protein is a human LamA3 polypeptide and the second human laminin protein is a human LamC2 polypeptide. In some embodiments, the first human laminin protein is a human LamB3 polypeptide and the second human laminin protein is a human LamC2 polypeptide.

Recombinant Nucleic Acids

[0268] In some embodiments, the present disclosure relates to recombinant nucleic acids comprising any one or more of the polynucleotides described herein. In some embodiments, the recombinant nucleic acid comprises one copy of the first polynucleotide. In some embodiments, the recombinant nucleic acid comprises two copies of the first polynucleotide. In some embodiments, the recombinant nucleic acid comprises one copy of the first polynucleotide and one copy of the second polynucleotide. In some embodiments, the recombinant nucleic acid comprises one copy of the first polynucleotide and two copies of the second polynucleotide. In some embodiments, the recombinant nucleic acid comprises two copies of the first polynucleotide and one copy of the second polynucleotide. In some embodiments, the recombinant nucleic acid comprises two copies of the first polynucleotide and two copies of the second polynucleotide.

[0269] In some embodiments, the recombinant nucleic acid is a vector (e.g., an expression vector, a display vector, etc.). In some embodiments, the vector is a DNA vector or an RNA vector. Generally, vectors suitable to maintain, propagate, and/or express polynucleotides to produce one or more polypeptides in a subject may be used. Examples of suitable vectors may include, for example, plasmids, cosmids, episomes, transposons, and viral vectors (e.g., adenoviral vectors, adeno-associated viral vectors, vaccinia viral vectors, Sindbis-viral vectors, measles vectors, herpes viral vectors, lentiviral vectors, retroviral vectors, etc.). In some embodiments, the vector is a herpes viral vector. In some embodiments, the vector is capable of autonomous replication in a host cell. In some embodiments, the vector is incapable of autonomous replication in a host cell. In some embodiments, the vector can integrate into a host DNA. In some embodiments, the vector cannot integrate into a host DNA (e.g., is episomal). Methods of making vectors containing one or more polynucleotides of interest are well known to one of ordinary skill in the art, including, for example, by chemical synthesis, or by artificial manipulation of isolated segments of nucleic acids (e.g., by genetic engineering techniques).

[0270] In some embodiments, a recombinant nucleic acid of the present disclosure is a herpes simplex virus (HSV) amplicon. Herpes virus amplicons, including the structural features and methods of making the same, are generally known to one of ordinary skill in the art (see e.g., de Silva S. and Bowers W. "Herpes Virus Amplicon Vectors". *Viruses* 2009, 1, 594-629). In some embodiments, the herpes simplex virus amplicon is an HSV-1 amplicon. In some embodiments, the herpes simplex virus amplicon is an HSV-1 hybrid amplicon. Examples of HSV-1 hybrid amplicons may include, but are not limited to, HSV/AAV hybrid amplicons, HSV/EBV hybrid amplicons, HSV/EBV/RV hybrid amplicons, and/or HSV/Sleeping Beauty hybrid amplicons. In some embodiments, the amplicon is an HSV/AAV hybrid amplicon. In some embodiments, the amplicon is an HSV/Sleeping Beauty hybrid amplicon.

[0271] In some embodiments, a recombinant nucleic acid of the present disclosure is a recombinant herpes virus genome. The recombinant herpes virus genome may be a recombinant genome from any member of the Herpesviridae family of DNA viruses known in the art, including, for example, a recombinant herpes simplex virus genome, a recombinant varicella zoster virus genome, a recombinant

human cytomegalovirus genome, a recombinant herpesvirus 6A genome, a recombinant herpesvirus 6B genome, a recombinant herpesvirus 7 genome, a recombinant Kaposi's sarcoma-associated herpesvirus genome, and any combinations or any derivatives thereof. In some embodiments, the recombinant herpes virus genome comprises one or more (e.g., one or more, two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, ten or more, etc.) inactivating mutations. In some embodiments, the one or more inactivating mutations are in one or more (e.g., one or more, two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, ten or more, etc.) herpes virus genes. In some embodiments, the recombinant herpes virus genome is attenuated (e.g., as compared to a corresponding, wild-type herpes virus genome). In some embodiments, the recombinant herpes virus genome is replication-competent. In some embodiments, the recombinant herpes virus genome is replication-defective.

[0272] In some embodiments, the recombinant nucleic acid is a recombinant herpes simplex virus (HSV) genome. In some embodiments, the recombinant herpes simplex virus genome is a recombinant type 1 herpes simplex virus (HSV-1) genome, a recombinant type 2 herpes simplex virus (HSV-2) genome, or any derivatives thereof. In some embodiments, the recombinant herpes simplex virus genome is a recombinant HSV-1 genome. In some embodiments, the recombinant herpes simplex virus genome is replication-competent. In some embodiments, the recombinant herpes simplex virus genome is replication-defective. In some embodiments, the recombinant herpes simplex virus genome comprises one or more (e.g., one or more, two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, ten or more, etc.) inactivating mutations. In some embodiments, the one or more inactivating mutations are in one or more (e.g., one or more, two or more, three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, ten or more, etc.) herpes simplex virus genes. As used herein, an "inactivating mutation" may refer to any mutation that results in a gene or regulon product (RNA or protein) having reduced, undetectable, or eliminated quantity and/or function (e.g., as compared to a corresponding sequence lacking the inactivating mutation). Examples of inactivating mutations may include, but are not limited to, deletions, insertions, point mutations, and rearrangements in transcriptional control sequences (promoters, enhancers, insulators, etc.) and/or coding sequences of a given gene or regulon. Any suitable method of measuring the quantity of a gene or regulon product known in the art may be used, including, for example, qPCR, Northern blots, RNAseq, western blots, ELISAs, etc.

[0273] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in at least one, at least two, at least three, at least four, at least five, at least six, at least seven, or all eight of the Infected Cell Protein (or Infected Cell Polypeptide) (ICP) 0, ICP4, ICP22, ICP27, ICP47, thymidine kinase (tk), Long Unique Region (UL) 41 and/or UL55 herpes simplex virus genes. In some embodiments, the recombinant herpes simplex virus genome does not comprise an inactivating mutation in the ICP34.5 and/or ICP47 herpes simplex virus genes (e.g., to avoid production of an immune-stimulating virus). In some embodiments, the recombinant herpes simplex virus genome

does not comprise an inactivating mutation in the ICP34.5 herpes simplex virus gene (one or both copies). In some embodiments, the recombinant herpes simplex virus genome does not comprise an inactivating mutation in the ICP47 herpes simplex virus gene. In some embodiments, the recombinant herpes simplex virus genome does not comprise an inactivating mutation in the ICP34.5 (one or both copies) and ICP47 herpes simplex virus genes. In some embodiments, the recombinant herpes simplex virus genome is not oncolytic.

[0274] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies). In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies), and further comprises an initiating mutation in the ICP4 (one or both copies) ICP22, ICP27, ICP47, UL41, and/or UL55 genes. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies), and an inactivating mutation in the ICP4 gene (one or both copies). In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies), and an inactivating mutation in the ICP22 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies), and an inactivating mutation in the UL41 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies), an inactivating mutation in the ICP4 gene (one or both copies), and an inactivating mutation in the ICP22 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies), an inactivating mutation in the ICP4 gene (one or both copies), and an inactivating mutation in the UL41 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies), an inactivating mutation in the ICP22 gene, and an inactivating mutation in the UL41 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 gene (one or both copies), an inactivating mutation in the ICP4 gene (one or both copies), an inactivating mutation in the ICP22 gene, and an inactivating mutation in the UL41 gene. In some embodiments, the inactivating mutation is a deletion of the coding sequence of the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, and/or UL41 genes. In some embodiments, the recombinant herpes simplex virus genome further comprises an inactivating mutation in the ICP27, ICP47, and/or UL55 genes.

[0275] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP4 gene (one or both copies). In some embodiments, the recombinant herpes complex virus genome comprises an inactivating mutation in the ICP4 (one or both copies), and further comprises an inactivating mutation in the ICP0 (one or both copies), ICP22, ICP27, ICP47, UL41, and/or UL55 genes. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP4 gene (one or both copies), and an inactivating mutation in the ICP22 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP4 gene (one or both copies), and an inactivating mutation in the ICP27, ICP47, and/or UL55 genes.

inactivating mutation in the UL41 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP4 gene (one or both copies), an inactivating mutation in the ICP22 gene, and an inactivating mutation in the UL41 gene. In some embodiments, the inactivating mutation is a deletion of the coding sequence of the ICP4 (one or both copies), ICP22, and/or UL41 genes. In some embodiments, the recombinant herpes simplex virus genome further comprises an inactivating mutation in the ICP0, ICP27, ICP47, and/or UL55 genes.

[0276] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP22 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP22 gene, and further comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP27, ICP47, UL41, and/or UL55 genes. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP22 gene, and an inactivating mutation UL41 gene. In some embodiments, the inactivating mutation is a deletion of the coding sequence of the ICP22 and/or UL41 genes. In some embodiments, the recombinant herpes simplex virus genome further comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP27, ICP47, and/or UL55 genes.

[0277] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP27 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP27 gene, and further comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, ICP47, UL41, and/or UL55 genes. In some embodiments, the inactivating mutation is a deletion of the coding sequence of the ICP27 gene.

[0278] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP47 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP47 gene, and further comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, ICP27, UL41, and/or UL55 genes. In some embodiments, the inactivating mutation is a deletion of the coding sequence of the ICP47 gene.

[0279] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the UL41 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the UL41 gene, and further comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, ICP27, ICP47, and/or UL55 genes. In some embodiments, the inactivating mutation is a deletion of the coding sequence of the UL41 gene.

[0280] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the UL55 gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the UL55 gene, and further comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, ICP27, ICP47, and/or UL41 genes. In some embodiments, the inactivating mutation is a deletion of the coding sequence of the UL55 gene.

[0281] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in

(e.g., a deletion of) the internal repeat (Joint) region comprising the internal repeat long (IRL) and internal repeat short (IRS) regions. In some embodiments, inactivation (e.g., deletion) of the Joint region eliminates one copy each of the ICP4 and ICP0 genes. In some embodiments, inactivation (e.g., deletion) of the Joint region further inactivates (e.g., deletes) the promoter for the ICP22 and ICP47 genes. If desired, expression of one or both of these genes can be restored by insertion of an immediate early promoter into the recombinant herpes simplex virus genome (see e.g., Hill et al. (1995). *Nature* 375(6530): 411-415; Goldsmith et al. (1998). *J Exp Med* 187(3): 341-348). Without wishing to be bound by theory, it is believed that inactivating (e.g., deleting) the Joint region may contribute to the stability of the recombinant herpes simplex virus genome and/or allow for the recombinant herpes simplex virus genome to accommodate more and/or larger transgenes.

[0282] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP4 (one or both copies), ICP22, and ICP27 genes. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP4 (one or both copies), ICP27, and UL55 genes. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP4 (one or both copies), ICP22, ICP27, ICP47, and UL55 genes. In some embodiments, the inactivating mutation in the ICP4 (one or both copies), ICP27, and/or UL55 genes is a deletion of the coding sequence of the ICP4 (one or both copies), ICP27, and/or UL55 genes. In some embodiments, the inactivating mutation in the ICP22 and ICP47 genes is a deletion in the promoter region of the ICP22 and ICP47 genes (e.g., the ICP22 and ICP47 coding sequences are intact but are not transcriptionally active). In some embodiments, the recombinant herpes simplex virus genome comprises a deletion in the coding sequence of the ICP4 (one or both copies), ICP27, and UL55 genes, and a deletion in the promoter region of the ICP22 and ICP47 genes. In some embodiments, the recombinant herpes simplex virus genome further comprises an inactivating mutation in the ICP0 and/or UL41 genes.

[0283] In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 (one or both copies) gene. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 (one or both copies) and ICP4 (one or both copies) genes. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), and ICP22 genes. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, and ICP27 genes. In some embodiments, the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, ICP27 and UL55 genes. In some embodiments, the inactivating mutation in the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, ICP27 and/or UL55 genes comprises a deletion of the coding sequence of the ICP0, ICP4 (one or both copies), ICP22, ICP27 and/or UL55 genes. In some embodiments, the recombinant herpes simplex virus genome further comprises an inactivating mutation in the ICP47 and/or the UL41 genes.

[0284] In some embodiments, a recombinant herpes simplex virus genome comprises one or more polynucleotides of the present disclosure within one, two, three, four, five, six, seven or more viral gene loci. Examples of suitable viral loci may include, without limitation, the ICP0 (one or both copies), ICP4 (one or both copies), ICP22, ICP27, ICP47, tk, UL41 and UL55 herpes simplex viral gene loci. In some embodiments, a recombinant herpes simplex virus genome comprises one or more polynucleotides of the present disclosure within one or both of the viral ICP4 gene loci (e.g., a recombinant virus carrying a first polynucleotide encoding a first human collagen protein in one or both of the ICP4 loci; a recombinant virus carrying a second polynucleotide encoding a second human collagen protein in one or both of the ICP4 loci; etc.). In some embodiments, a recombinant herpes simplex virus genome comprises one or more polynucleotide of the present disclosure within the viral ICP22 gene locus (e.g., a recombinant virus carrying a first polynucleotide encoding a first human collagen protein in the ICP22 locus; a recombinant virus carrying a second polynucleotide encoding a second human collagen protein in the ICP22 locus; etc.). In some embodiments, a recombinant herpes simplex virus genome comprises one or more polynucleotide of the present disclosure within the viral UL41 gene locus (e.g., a recombinant virus carrying a first polynucleotide encoding a first human collagen protein in the UL41 locus; a recombinant virus carrying a second polynucleotide encoding a second human collagen protein in the UL41 locus; etc.). In some embodiments, a recombinant herpes simplex virus genome comprises one or more polynucleotides of the present disclosure within one or both of the viral ICP4 gene loci, and one or more polynucleotides of the present disclosure within the viral ICP22 locus (e.g., a recombinant virus carrying a first polynucleotide encoding a first human collagen protein in one or both of the ICP4 loci and a second polynucleotide encoding a second human collagen protein in the ICP22 locus; a recombinant virus carrying a second polynucleotide encoding a second human collagen protein in one or both of the ICP4 loci and a first polynucleotide encoding a first human collagen protein in the ICP22 locus; etc.). In some embodiments, a recombinant herpes simplex virus genome comprises one or more polynucleotides of the present disclosure within one or both of the viral ICP4 gene loci, and one or more polynucleotides of the present disclosure within the viral UL41 locus (e.g., a recombinant virus carrying a first polynucleotide encoding a first human collagen protein in one or both of the ICP4 loci and a second polynucleotide encoding a second human collagen protein in the UL41 locus; a recombinant virus carrying a second polynucleotide encoding a second human collagen protein in one or both of the ICP4 loci and a first polynucleotide encoding a first human collagen protein in the UL41 locus; etc.). In some embodiments, a recombinant herpes simplex virus genome comprises one or more polynucleotides of the present disclosure within one or both of the viral ICP4 gene loci, one or more polynucleotides of the present disclosure within the viral ICP22 locus, and one or more polynucleotides of the present disclosure within the viral UL41 locus (e.g., a recombinant virus carrying a first polynucleotide encoding a first human collagen protein in one or both of the ICP4 loci and a second polynucleotide encoding a second human collagen protein in the ICP22 and UL41 loci; a recombinant virus carrying a second polynucleotide encoding a second human collagen protein in one or

both of the ICP4 loci and a first polynucleotide encoding a first human collagen protein in the ICP22 and UL41 loci; etc.).

[0285] In some embodiments, the recombinant herpes virus genome (e.g., a recombinant herpes simplex virus genome) has been engineered to decrease or eliminate expression of one or more toxic herpes simplex genes (such as one or both copies of the HSV ICP0 gene, one or both copied of the HSV ICP4 gene, the ICP22 gene, and/or the UL41 gene). In some embodiments, the recombinant herpes virus genome (e.g., a recombinant herpes simplex virus genome) has been engineered to reduce cytotoxicity of the recombinant genome (e.g., when introduced into a target cell) as compared to a corresponding wild-type herpes virus genome (e.g., a wild-type herpes simplex virus genome). In some embodiments, cytotoxicity (e.g., in human keratinocytes and/or fibroblast cells) of the recombinant virus genome (e.g., a recombinant herpes simplex virus genome) is reduced by at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% as compared to a corresponding wild-type herpes virus genome (e.g., measuring the relative cytotoxicity of a recombinant ΔICP4 (one or both copies) herpes simplex virus genome vs. a wild-type herpes simplex virus genome in human keratinocytes or fibroblasts (primary cells or cell lines); measuring the relative cytotoxicity of a recombinant ΔICP4 (one or both copies)/ΔICP22 herpes simplex virus genome vs. a wild-type herpes simplex virus genome in human keratinocytes or fibroblasts (primary cells or cell lines); etc.). In some embodiments, cytotoxicity (e.g., in human keratinocytes and/or fibroblast cells) of the recombinant herpes virus genome (e.g., a recombinant herpes simplex virus genome) is reduced by at least about 1.5-fold, at least about 2-fold, at least about 3-fold, at least about 4-fold, at least about 5-fold, at least about 6-fold, at least about 7-fold, at least about 8-fold, at least about 9-fold, at least about 10-fold, at least about 15-fold, at least about 20-fold, at least about 25-fold, at least about 50-fold, at least about 75-fold, at least about 100-fold, at least about 250-fold, at least about 500-fold, at least about 750-fold, at least about 1000-fold, or more as compared to a corresponding wild-type herpes virus genome (e.g., measuring the relative cytotoxicity of a recombinant ΔICP4 (one or both copies) herpes simplex virus genome vs. a wild-type herpes simplex virus genome in human keratinocytes or fibroblasts (primary cells or cell lines); measuring the relative cytotoxicity of a recombinant ΔICP4 (one or both copies)/ΔICP22 herpes simplex virus genome vs. a wild-type herpes simplex virus genome in human keratinocytes or fibroblasts (primary cells or cell lines); etc.). Methods of measuring cytotoxicity are known to one of ordinary skill in the art, including, for example, through the use of vital dyes (formazan dyes), protease biomarkers, an MTT assay (or an assay using related tetrazolium salts such as XTT, MTS, water-soluble tetrazolium salts, etc.), measuring ATP content, etc.

[0286] In some embodiments, the recombinant herpes virus genome (e.g., a recombinant herpes simplex virus genome) has been engineered to reduce its impact on host cell proliferation after exposure of the target cell to the

recombinant genome, as compared to a corresponding wild-type herpes virus genome (e.g., a wild-type herpes simplex virus genome). In some embodiments, the target cell is a human cell. In some embodiments, the target cell is a cell of the epidermis and/or dermis. In some embodiments, the target cell is a keratinocyte and/or fibroblast. In some embodiments, host cell proliferation (e.g., human keratinocytes and/or fibroblast cells) after exposure to the recombinant genome is at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% faster as compared to host cell proliferation after exposure to a corresponding wild-type herpes virus genome (e.g., measuring the relative cellular proliferation after exposure to a recombinant ΔICP4 (one or both copies) herpes simplex virus genome vs. cellular proliferation after exposure to a wild-type herpes simplex virus genome in human keratinocytes or fibroblasts (primary cells or cell lines); measuring the relative cellular proliferation after exposure to a recombinant ΔICP4 (one or both copies)/ΔICP22 herpes simplex virus genome vs. cellular proliferation after exposure to a wild-type herpes simplex virus genome in human keratinocytes or fibroblasts (primary cells or cell lines); etc.). In some embodiments, host cell proliferation (e.g., human keratinocytes and/or fibroblast cells) after exposure to the recombinant genome is at least about 1.5-fold, at least about 2-fold, at least about 3-fold, at least about 4-fold, at least about 5-fold, at least about 6-fold, at least about 7-fold, at least about 8-fold, at least about 9-fold, at least about 10-fold, at least about 15-fold, at least about 20-fold, at least about 25-fold, at least about 50-fold, at least about 75-fold, at least about 100-fold, at least about 250-fold, at least about 500-fold, at least about 750-fold, or at least about 1000-fold faster as compared to host cell proliferation after exposure to a corresponding wild-type herpes virus genome (e.g., measuring the relative cellular proliferation after exposure to a recombinant ΔICP4 (one or both copies) herpes simplex virus genome vs. cellular proliferation after exposure to a wild-type herpes simplex virus genome in human keratinocytes or fibroblasts (primary cells or cell lines); measuring the relative cellular proliferation after exposure to a recombinant ΔICP4 (one or both copies)/ΔICP22 herpes simplex virus genome vs. cellular proliferation after exposure to a wild-type herpes simplex virus genome in human keratinocytes or fibroblasts (primary cells or cell lines); etc.). Methods of measuring cellular proliferation are known to one of ordinary skill in the art, including, for example, through the use of a Ki67 cell proliferation assay, a BrdU cell proliferation assay, etc.

[0287] A vector (e.g., herpes viral vector) may include one or more polynucleotides of the present disclosure in a form suitable for expression of the polynucleotide in a host cell. Vectors may include one or more regulatory sequences operatively linked to the polynucleotide to be expressed (e.g., as described above).

[0288] In some embodiments, a recombinant nucleic acid of the present disclosure (e.g., a recombinant herpes simplex virus genome) comprises one or more of the polynucleotides described herein inserted in any orientation in the recombinant nucleic acid. If the recombinant nucleic acid comprises

two or more polynucleotides described herein (e.g., two or more, three or more, etc.), the polynucleotides may be inserted in the same orientation or opposite orientations to one another. Without wishing to be bound by theory, incorporating two polynucleotides (e.g., two transgenes) into a recombinant nucleic acid (e.g., a vector) in an antisense orientation may help to avoid read-through and ensure proper expression of each polynucleotide.

IV. Viruses

[0289] Certain aspects of the present disclosure relate to viruses comprising any of the polynucleotides and/or recombinant nucleic acids described herein. In some embodiments, the virus is capable of infecting one or more target cells of a subject (e.g., a human). In some embodiments, the virus is suitable for delivering the polynucleotides and/or recombinant nucleic acids into one or more target cells of a subject (e.g., a human subject). In some embodiments, the one or more target cell are one or more human cells. In some embodiments, the one or more target cells are one or more cells of the skin (e.g., one or more cells of the epidermis, dermis, and/or subcutis). In some embodiments, the one or more cells are selected from keratinocytes, melanocytes, Langerhans cells, Merkel cells, mast cells, fibroblasts, and/or adipocytes. In some embodiments, the one or more cells are keratinocytes. In some embodiments, the one or more cells reside in the stratum corneum, stratum granulosum, stratum spinulosum, stratum basale, and/or basement membrane. In some embodiments, the one or more target cells are one or more epidermal cells.

[0290] Any suitable virus known in the art may be used, including, for example, adenovirus, adeno-associated virus, retrovirus, lentivirus, sendai virus, herpes virus (e.g., a herpes simplex virus), vaccinia virus, and/or any hybrid virus thereof. In some embodiments, the virus is attenuated. In some embodiments, the virus is replication-defective. In some embodiments, the virus is replication-competent. In some embodiments, the virus has been modified to alter its tissue tropism relative to the tissue tropism of an unmodified, wild-type virus. In some embodiments, the virus has reduced cytotoxicity as compared to a corresponding wild-type virus. Methods for producing a virus comprising recombinant nucleic acids are well known to one of ordinary skill in the art.

[0291] In some embodiments, the virus is a member of the Herpesviridae family of DNA viruses, including, for example, a herpes simplex virus, a varicella zoster virus, a human cytomegalovirus, a herpesvirus 6A, a herpesvirus 6B, a herpesvirus 7, and a Kaposi's sarcoma-associated herpesvirus, etc. In some embodiments, the herpes virus is attenuated. In some embodiments, the herpes virus is replication-defective. In some embodiments, the herpes virus is replication-competent. In some embodiments, the herpes virus has reduced cytotoxicity as compared to a corresponding wild-type herpes virus. In some embodiments, the herpes virus is not oncolytic.

[0292] In some embodiments, the virus is a herpes simplex virus. Herpes simplex viruses comprising recombinant nucleic acids may be produced by a process disclosed, for example, in WO2015/009952 and/or WO2017/176336. In some embodiments, the herpes simplex virus is attenuated. In some embodiments, the herpes simplex virus is replication-competent. In some embodiments, the herpes simplex virus is replication-defective. In some embodiments, the

herpes simplex virus is a herpes simplex type 1 virus (HSV-1), a herpes simplex type 2 virus (HSV-2), or any derivatives thereof. In some embodiments, the herpes simplex virus is a herpes simplex type 1 virus (HSV-1). In some embodiments, the HSV-1 is attenuated. In some embodiments, the HSV-1 has reduced cytotoxicity as compared to a corresponding wild-type HSV-1. In some embodiments, the HSV-1 is not oncolytic.

[0293] In some embodiments, the herpes simplex virus has been modified to alter its tissue tropism relative to the tissue tropism of an unmodified, wild-type herpes simplex virus. In some embodiments, the herpes simplex virus comprises a modified envelope. In some embodiments, the modified envelope comprises one or more (e.g., one or more, two or more, three or more, four or more, etc.) mutant herpes simplex virus glycoproteins. Examples of herpes simplex virus glycoproteins may include, but are not limited to, the glycoproteins gB, gC, gD, gH, and gL. In some embodiments, the modified envelope alters the herpes simplex virus tissue tropism relative to a wild-type herpes simplex virus.

[0294] In some embodiments, the transduction efficiency (in vitro and/or in vivo) of a virus of the present disclosure (e.g., a herpes virus) for one or more target cells (e.g., one or more human keratinocytes and/or fibroblasts) is at least about 25%. For example, the transduction efficiency of the virus for one or more target cells may be at least about 25%, at least about 30%, at least about 35%, at least about 40%, at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 99%, at least about 99.5%, or more. In some embodiments, the virus is a herpes simplex virus and the transduction efficiency of the virus for one or more target cells (e.g., one or more human keratinocytes and/or fibroblasts) is about 85% to about 100%. In some embodiments, the virus is a herpes simplex virus and the transduction efficiency of the virus for one or more target cells (e.g., one or more human keratinocytes and/or fibroblasts) is at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100%. Methods of measuring viral transduction efficiency in vitro or in vivo are well known to one of ordinary skill in the art, including, for example, qPCR analysis, deep sequencing, western blotting, fluorometric analysis (such as fluorescent in situ hybridization (FISH), fluorescent reporter gene expression, immunofluorescence, FACS), etc.

V. Compositions and Formulations

[0295] Certain aspects of the present disclosure relate to compositions and formulations (e.g., pharmaceutical compositions and formulations) comprising any of the recombinant nucleic acids (e.g., a recombinant herpes virus genome) and/or viruses (e.g., a herpes virus comprising a recombinant genome described herein (such as a herpes simplex virus comprising a recombinant herpes simplex virus genome), and an excipient or carrier (e.g., a pharmaceutically acceptable excipient or carrier). In some embodiments, the composition or formulation is a cosmetic composition or formulation (e.g., a skin care product).

[0296] In some embodiments, the composition or formulation comprises any one or more of the viruses (e.g., herpes viruses) described herein. In some embodiments, the composition or formulation comprises from about 10^4 to about 10^{12} plaque forming units (PFU)/mL of the virus. For example, the composition or formulation may comprise from about 10^4 to about 10^{12} , about 10^5 to about 10^{12} , about 10^6 to about 10^{12} , about 10^7 to about 10^{12} , about 10^8 to about 10^{12} , about 10^9 to about 10^{12} , about 10^{10} to about 10^{12} , about 10^{11} to about 10^{12} , about 10^4 to about 10^{11} , about 10^5 to about 10^{11} , about 10^6 to about 10^{11} , about 10^7 to about 10^{11} , about 10^8 to about 10^{11} , about 10^9 to about 10^{11} , about 10^{10} to about 10^{11} , about 10^4 to about 10^{10} , about 10^5 to about 10^{10} , about 10^6 to about 10^{10} , about 10^7 to about 10^{10} , about 10^8 to about 10^{10} , about 10^9 to about 10^{10} , about 10^{10} to about 10^{10} , about 10^4 to about 10^9 , about 10^5 to about 10^9 , about 10^6 to about 10^9 , about 10^7 to about 10^9 , about 10^8 to about 10^9 , about 10^9 to about 10^9 , about 10^{10} to about 10^9 , about 10^4 to about 10^8 , about 10^5 to about 10^8 , about 10^6 to about 10^8 , about 10^7 to about 10^8 , about 10^8 to about 10^8 , about 10^9 to about 10^8 , about 10^{10} to about 10^8 , about 10^4 to about 10^7 , about 10^5 to about 10^7 , about 10^6 to about 10^7 , about 10^7 to about 10^7 , about 10^8 to about 10^7 , about 10^9 to about 10^7 , about 10^{10} to about 10^7 , about 10^4 to about 10^6 , or about 10^4 to about 10^5 PFU/mL of the virus. In some embodiments, the composition or formulation comprises about 10^4 , about 10^5 , about 10^6 , about 10^7 , about 10^8 , about 10^9 , about 10^{10} , about 10^{11} , or about 10^{12} PFU/mL of the virus.

[0297] Compositions and formulations (e.g., pharmaceutical compositions and formulations) as described herein can be prepared by mixing the active ingredient(s) (such as a recombinant nucleic acid or a virus) having the desired degree of purity with one or more acceptable carriers or excipients. Acceptable carriers or excipients (e.g., pharmaceutically acceptable carriers or excipients) are generally nontoxic to recipients at the dosages and concentrations employed, and may include, but are not limited to: buffers (such as phosphate, citrate, acetate, and other organic acids); antioxidants (such as ascorbic acid and methionine); preservatives (such as octadecyltrimethylbenzyl ammonium chloride, benzalkonium chloride, benzethonium chloride, phenol, butyl or benzyl alcohol, alkyl parabens, catechol, resorcinol, cyclohexanol, 3-pentanol, and m-cresol); amino acids (such as glycine, glutamine, asparagine, histidine, arginine, or lysine); low molecular weight (less than about 10 residues) polypeptides; proteins (such as serum albumin, gelatin, or immunoglobulins); polyols (such as glycerol, e.g., formulations including 10% glycerol); hydrophilic polymers (such as polyvinylpyrrolidone); monosaccharides, disaccharides, and other carbohydrates (including glucose, mannose, or dextrans); chelating agents (such as EDTA); sugars (such as sucrose, mannitol, trehalose, or sorbitol); salt-forming counter-ions (such as sodium); metal complexes (such as Zn-protein complexes); liposomes (e.g., cationic lipids); nanoparticle carriers; and/or non-ionic surfactants (such as polyethylene glycol (PEG)). A thorough discussion of carriers is available in REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991).

[0298] In some embodiments, the composition or formulation comprises one or more lipid (e.g., cationic lipid) carriers. In some embodiments, the composition or formulation comprises one or more nanoparticle carriers. Nanoparticles are submicron (less than about 1000 nm) sized drug delivery vehicles that can carry encapsulated drugs (such as synthetic small molecules, proteins, peptides, cells, viruses, and nucleic acid-based biotherapeutics for rapid or

controlled release. A variety of molecules (e.g., proteins, peptides, recombinant nucleic acids, etc.) can be efficiently encapsulated in nanoparticles using processes well known in the art. In some embodiments, a molecule “encapsulated” in a nanoparticle may refer to a molecule (such as a virus) that is contained within the nanoparticle or attached to and/or associated with the surface of the nanoparticle, or any combination thereof. Nanoparticles for use in the compositions or formulations described herein may be any type of biocompatible nanoparticle known in the art, including, for example, nanoparticles comprising poly(lactic acid), poly(glycolic acid), PLGA, PLA, PGA, and any combinations thereof (see e.g., Vauthier et al. *Adv Drug Del Rev.* (2003) 55: 519-48; US2007/0148074; US2007/0092575; US2006/0246139; U.S. Pat. Nos. 5,753,234; 7,081,483; and WO2006/052285).

[0299] In some embodiments, the carrier or excipient (e.g., a pharmaceutically acceptable carrier or excipient) may be adapted for or suitable for any administration route known in the art, including, for example, intravenous, intramuscular, subcutaneous, cutaneous, intranasal, intratracheal, sublingual, buccal, topical, oral, transdermal, intradermal, intraperitoneal, intraorbital, intravitreal, subretinal, transmucosal, intraarticular, by superficial injection, by implantation, by inhalation, intrathecal, intraventricular, and/or intranasal administration. In some embodiments, the carrier or excipient (e.g., pharmaceutically acceptable carrier or excipient) is adapted for or suitable for topical, transdermal, subcutaneous, and/or intradermal administration. In some embodiments, the carrier or excipient is adapted for or suitable for topical, transdermal, and/or intradermal administration. In some embodiments, the carrier or excipient is adapted for or suitable for superficial injection.

[0300] Examples of carriers or excipients adapted for or suitable for use in a topical, transdermal, subcutaneous, superficial, and/or intradermal application/administration may include, but are not limited to, ointments, oils, pastes, creams, aerosols, suspensions, emulsions, fatty ointments, gels, powders, liquids, lotions, solutions, sprays, patches (e.g., transdermal patches or microneedle patches), adhesive strips, a microneedle or microneedle arrays, and inhalants. In some embodiments, the carrier or excipient (e.g., the pharmaceutically acceptable carrier or excipient) comprises one or more (e.g., one or more, two or more, three or more, four or more, five or more, etc.) of an ointment, oil, paste, cream, aerosol, suspension, emulsion, fatty ointment, gel, powder, liquid lotion, solution, spray, adhesive strip, and an inhalant. In some embodiments, the carrier comprises a patch (e.g., a patch that adheres to the skin), such as a transdermal patch or a microneedle patch. In some embodiments, the carrier comprises a microneedle or microneedle array. Methods for making and using microneedle arrays suitable for composition delivery are generally known in the art (Kim Y. et al. “Microneedles for drug and vaccine delivery”. *Advanced Drug Delivery Reviews* 2012, 64 (14): 1547-68).

[0301] In some embodiments, the composition or formulation (e.g., the pharmaceutical composition or formulation) is adapted for or suitable for any administration route known in the art, including, for example, intravenous, intramuscular, subcutaneous, cutaneous, oral, intranasal, intratracheal, sublingual, buccal, topical, transdermal, intradermal, intraperitoneal, intraorbital, intravitreal, subretinal, transmucosal, intraarticular, by superficial injection, by implanta-

tion, by inhalation, intrathecal, intraventricular, and/or intranasal administration. In some embodiments, the composition or formulation is adapted for or suitable for cutaneous, topical, transdermal, subcutaneous, and/or intradermal administration. In some embodiments, the pharmaceutical composition or formulation is adapted for or suitable for topical, transdermal, and/or intradermal administration. In some embodiments, the composition or formulation is adapted for or suitable for intradermal administration. In some embodiments, the composition of formulation is adapted for or suitable for superficial injection.

[0302] In some embodiments, the composition or formulation (e.g., pharmaceutical composition or formulation) further comprises one or more additional components. Examples of additional components may include, but are not limited to, binding agents (e.g., pregelatinized maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose, etc.); fillers (e.g., lactose and other sugars, microcrystalline cellulose, pectin, gelatin, calcium sulfate, ethyl cellulose, polyacrylates or calcium hydrogen phosphate, etc.); lubricants (e.g., magnesium stearate, talc, silica, colloidal silicon dioxide, stearic acid, metallic stearates, hydrogenated vegetable oils, corn starch, polyethylene glycols, sodium benzoate, sodium acetate, etc.); disintegrants (e.g., starch, sodium starch glycolate, etc.); wetting agents (e.g., sodium lauryl sulphate, etc.); salt solutions; alcohols; polyethylene glycols; gelatin; lactose; amylase; magnesium stearate; talc; silicic acid; viscous paraffin; hydroxymethylcellulose; polyvinylpyrrolidone; sweetenings; flavorings; perfuming agents; colorants; moisturizers; sunscreens; antibacterial agents; agents able to stabilize polynucleotides or prevent their degradation, and the like. In some embodiments, the composition or formulation comprises a hydroxypropyl methylcellulose gel. In some embodiments, the composition or formulation comprises a phosphate buffer. In some embodiments, the composition or formulation comprises glycerol (e.g., at about 1%, about 2%, about 3%, about 4%, about 5%, about 6%, about 7%, about 8%, about 9%, about 10%, about 15%, etc.).

[0303] Compositions and formulations (e.g., pharmaceutical compositions and formulations) to be used for in vivo administration are generally sterile. Sterility may be readily accomplished, e.g., by filtration through sterile filtration membranes.

[0304] In some embodiments, any of the recombinant nucleic acids, viruses, and/or compositions or formulations described herein may be used to deliver one or more polynucleotides encoding a collagen protein (e.g., a human collagen protein such as Collagen 3) into one or more cells of a subject (e.g., one or more collagen-deficient cells). In some embodiments, any of the recombinant nucleic acids, viruses, and/or compositions or formulations described herein may be used in a therapy. In some embodiments, any of the recombinant nucleic acids, viruses, and/or compositions or formulations described herein may be used in the treatment of a cosmetic or aesthetic condition that would benefit from the expression of a collagen polypeptide (e.g., a cosmetic or aesthetic condition associated with a collagen deficiency (such as aged and/or UV-damaged skin)). In some embodiments, any of the recombinant nucleic acids, viruses, and/or compositions or formulations described herein may be used in the treatment of dermatological aging (e.g., as described below).

[0305] In some embodiments, any of the recombinant nucleic acids, viruses, and/or compositions or formulations described herein may be used in the preparation or manufacture of a medicament. In some embodiments, any of the recombinant nucleic acids, viruses, and/or compositions or formulations described herein may be used in the preparation or manufacture of a medicament useful for delivering one or more polynucleotides encoding a collagen protein (e.g., a human collagen protein such as Collagen 3) into one or more cells of a subject (e.g., one or more collagen-deficient cells). In some embodiments, any of the recombinant nucleic acids, viruses, and/or compositions or formulations described herein may be used in the preparation or manufacture of a medicament useful for the treatment of a cosmetic or aesthetic condition that would benefit from the expression of a collagen polypeptide (e.g., a cosmetic or aesthetic condition associated with a collagen deficiency (such as aged and/or UV-damaged skin)). In some embodiments, any of the recombinant nucleic acids, viruses, and/or compositions or formulations described herein may be used in the preparation or manufacture of a medicament useful for the treatment of dermatological aging (e.g., as described below).

VI. Methods

[0306] Certain aspects of the present disclosure relate to a method of enhancing, increasing, augmenting, and/or supplementing the levels of one or more dermal extracellular matrix proteins in a subject (e.g., in one or more cells of a subject) comprising administering to the subject any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. In some embodiments, the subject is a human.

[0307] Other aspects of the present disclosure relate to method of stabilizing or improving the structure and/or organization of the dermal extracellular matrix in a subject comprising administering to the subject any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. In some embodiments, the subject is a human.

[0308] Other aspects of the present disclosure relate to a method of enhancing, increasing, augmenting, and/or supplementing the levels of one or more human collagen proteins in a subject (e.g., in one or more cells of a subject) comprising administering to the subject any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. In some embodiments, the subject is a human.

[0309] In some embodiments, administration of the recombinant nucleic acid, virus, medicament, and/or composition to the subject increases collagen (e.g., COL1-1; COL1-2; COL3; COL1-1 and COL1-2; COL1-1 and COL3; etc.) levels (transcript or protein levels) in one or more cells of the subject by at least about 10%, as compared to the endogenous levels of the collagen(s) in one or more corresponding untreated cells (e.g., one or more cells prior to treatment, one or more uninfected cells during treatment, etc.) of the subject. For example, administration of the recombinant nucleic acid, virus, medicament, and/or composition may increase collagen levels (transcript or protein levels) in one or more cells of the subject by at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about

80%, at least about 90%, at least about 95%, at least about 99%, or more, as compared to the endogenous levels of the collagen(s) in one or more corresponding untreated cells of the subject. In some embodiments, administration of the recombinant nucleic acid, virus, medicament, and/or composition to the subject increases collagen levels (transcript or protein levels) in one or more cells of the subject by at least about 2-fold, as compared to the endogenous levels of the collagen(s) in one or more corresponding untreated cells (e.g., one or more cells prior to treatment, one or more uninfected cells during treatment, etc.) of the subject. For example, administration of the recombinant nucleic acid, virus, medicament, and/or composition may increase collagen levels (transcript or protein levels) in one or more cells of the subject by at least about 2-fold, at least about 3-fold, at least about 4-fold, at least about 5-fold, at least about 6-fold, at least about 7-fold, at least about 8-fold, at least about 9-fold, at least about 10-fold, at least about 15-fold, at least about 20-fold, at least about 25-fold, at least about 50-fold, at least about 75-fold, at least about 100-fold, at least about 250-fold, at least about 500-fold, at least about 750-fold, at least about 1000-fold, or more, as compared to the endogenous levels of the collagen(s) in one or more corresponding untreated cells of the subject. Methods of measuring transcript or protein levels from a sample are well known to one of ordinary skill in the art, including, for example, by qPCR, RNAseq, ELISA, western blot, mass spectrometry, etc.

[0310] Other aspects of the present disclosure relate to a method of enhancing, increasing, augmenting, and/or supplementing the soft tissue of a subject comprising administering to the subject any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. In some embodiments, the subject is a human. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are injected into a soft tissue of the subject. In some embodiments, the skin of the subject is aging skin. In some embodiments, the skin of the subject has been damaged due to exposure to ultraviolet light (e.g., from the sun, from a tanning bed, etc.). In some embodiments, the skin of the subject is wrinkled.

[0311] In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions may be used in a method to repair and/or augment the soft tissue of a subject. In some embodiments, "tissue repair" refers to the restoration of tissue architecture and/or function and encompasses tissue regeneration and replacement. In some embodiments, repair or augmentation of the soft tissue refers to procedures that are used to restore the youthful appearance of skin (e.g., as compared to "aged" skin whose appearance is due to defects resulting from chronological aging or other physical, chemical, or UV damage). In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are useful in cosmetic soft tissue applications, such as to fill wrinkles, lines, folds, scars, and to enhance dermal tissue (e.g., plump thin lips, fill in sunken eyes and/or shallow cheeks, etc.).

[0312] Other aspects of the present disclosure relate to a method of improving skin quality, condition, and/or appearance in a subject in need thereof comprising administering to the subject any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. In some embodiments, the subject is a human. In some embodiments, the skin condition is one or more of sun damage,

aging, UV exposure, rough texture, skin sagging, and/or wrinkles. Improvement of skin quality, condition, and/or appearance (e.g., as compared to before treatment) may be assessed using any appropriate method or scale known in the art, including, for example, FACE Q, GAIS, etc. In some embodiments, the skin of the subject is aging skin. In some embodiments, the skin of the subject has been damaged due to exposure to ultraviolet light (e.g., from the sun, from a tanning bed, etc.). In some embodiments, the skin of the subject is wrinkled.

[0313] Other aspects of the present disclosure relate to a method of reducing the appearance of one or more superficial depressions in the skin of a subject in need thereof comprising administering to the subject any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. In some embodiments, the subject is a human. In some embodiments, administration of the recombinant nucleic acid, virus, medicaments, and/or composition reduces the appearance of one or more superficial depressions in the skin of the subject for at least about three months, at least about six months, at least about nine months, or at least about 12 months. In some embodiments, the appearance of one or more superficial depressions in the skin of the subject is reduced after administration of the composition, as compared to the appearance of the one or more superficial depression in the skin of the subject prior to administration of the composition. In some embodiments, the one or more superficial depressions in the skin are one or more of fine lines and wrinkles (e.g., forehead wrinkles, "crow's feet", wrinkles at the edges of the eye or mouth, etc.). In some embodiments, the treatment of one or more superficial skin depressions is measured by an improvement in skin texture or skin quality, such as smoothness, hydration, and elasticity, as compared to non-treated skin. In some embodiments, the treatment of one or more superficial skin depressions is measured by a reduction in the severity (e.g., depth) of the superficial depressions and/or a reduction in the number of fine lines or wrinkles in a given area of skin. In some embodiments, the skin of the subject is aging skin. In some embodiments, the skin of the subject has been damaged due to exposure to ultraviolet light (e.g., from the sun, from a tanning bed, etc.). In some embodiments, the skin of the subject is wrinkled.

[0314] Other aspects of the present disclosure relate to a method of increasing and/or improving at least one of texture, smoothness, elasticity, and/or tension of the skin of a subject comprising administering to the subject any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. In some embodiments, the subject is a human. In some embodiments, the skin of the subject maintains at least one of an increased and/or improved texture, smoothness, elasticity, or tension for at least about three months, at least about six months, at least about nine months, or at least about 12 months after administration of the composition. In some embodiments, at least one of texture, smoothness, elasticity, or tension of the skin of the subject is increased and/or improved after administration of the composition, as compared to the texture, smoothness, elasticity, or tension of the skin of the subject prior to administration of the composition. Methods of measuring texture, smoothness, elasticity, and/or tensions of the skin are known to one of skill in the art. In some embodiments, the skin of the subject is aging skin. In some embodiments, the skin of the subject has been damaged due

to exposure to ultraviolet light (e.g., from the sun, from a tanning bed, etc.). In some embodiments, the skin of the subject is wrinkled.

[0315] Other aspects of the present disclosure relate to a method of diminishing one or more dermatological signs of aging in a subject in need thereof comprising administering to the subject an effective amount of any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. In some embodiments, the subject is a human. In some embodiments, diminishing one or more dermatological signs of aging include any one of more of the following: treatment, reduction, and/or prevention of fine lines and/or wrinkles; reduction of skin pore size; improvement in skin thickness, plumpness, and/or tautness; improvement in skin smoothness, suppleness, and/or softness; improvement in skin tone, radiance, and/or clarity; improvement in procollagen and/or collagen production; improvement in skin texture and or promotion of retexturization; improvement in appearance of skin contours; restoration of skin luster and/or brightness; improvement of skin appearance decreased by aging and/or menopause; improvement in skin moisturization; increase in skin elasticity and/or resiliency; treatment, reduction, and/or prevention or skin sagging; improvement in skin firmness; reduction of pigment spots, mottled skin, and/or scars (such as acne scars); and/or improvement of optical properties of skin by light diffraction or reflection. In some embodiments, the one or more dermatological signs of aging in the subject is diminished after administration of the composition, as compared to the one or more dermatological signs of aging in the subject prior to administration of the composition. Any suitable method for measuring one or more signs of dermatological aging known in the art may be used.

[0316] In some embodiments, one or more portions of the skin of the subject is abraded or made more permeable prior to treatment with an effective amount of any of the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein. Any suitable method of abrading the skin or increasing skin permeability known in the art may be used, including, for example, use of a dermal roller, repeated use of adhesive strips to remove layers of skin cells (tape stripping), scraping with a scalpel or blade, use of sandpaper, use of chemical permeation enhancers or electrical energy, use of sonic or ultrasonic energy, use of light (e.g., laser) energy, use of micron-sized needles or blades with a length suitable to pierce but not completely pass through the epidermis, etc.

[0317] In some embodiments, the methods of the present disclosure are for cosmetic applications, such as to reduce or eliminate one or more superficial depressions in the skin, to reduce or eliminate one or more wrinkles, and/or to prevent the occurrence or reoccurrence of one or more wrinkles. In some embodiments, the one or more superficial depressions in the skin or wrinkles are selected from nasolabial folds, crows' feet, frown lines, worry lines, scars, glabellar lines, brow ptosis, tear troughs, nasojugal lines, bunny lines, cheek/mid-face ptosis, marionette lines, poppy dimpling, smile lines, laugh lines, chin creases, neck lines, platysma bands, and any combinations thereof.

[0318] In some embodiments, an "effective amount" is at least the minimum amount required to affect a measurable improvement in or prevention of one or more signs or symptoms of a particular condition (e.g., a cosmetic condition such as skin aging). An "effective amount" may vary

according to factors such as the age, sex, and weight of the patient. An effective amount is also one in which any toxic or detrimental effects of the treatment are outweighed by the beneficial effects. An effective amount can be administered in one or more administrations. For purposes of the present disclosure, an effective amount of a recombinant nucleic acid, virus, medicament, and/or composition is an amount sufficient to accomplish a measurable improvement either directly or indirectly. As is understood in the clinical context, an effective amount of a recombinant nucleic acid, virus, medicament, and/or composition may or may not be achieved in conjunction with another drug, compound, or composition. Thus, an "effective amount" may be considered in the context of administering one or more agents, and a single agent may be considered to be given in an effective amount if, in conjunction with one or more other agents, a desirable result may be or is achieved.

[0319] In some embodiments, the recombinant nucleic acid, virus, medicament, and/or composition is administered once to the subject. In some embodiments, the recombinant nucleic acid, virus, medicament, and/or composition is administered at least twice (e.g., at least 2 times, at least 3 times, at least 4 times, at least 5 times, at least 10 times, etc.) to the subject. In some embodiments, at least about 15 days (e.g., at least about 15 days, at least about 20 days, at least about 30 days, at least about 40 days, at least about 50 days, at least about 60 days, at least about 70 days, at least about 80 days, at least about 90 days, at least about 100 days, at least about 120 days, etc.) pass between administrations (e.g., between the first and second administrations, between the second and third administrations, etc.).

[0320] The recombinant nucleic acids, viruses, medicaments, and/or compositions or formulations described herein may be administered by any suitable method or route known in the art, including, without limitation, by oral administration, sublingual administration, buccal administration, topical administration, rectal administration, via inhalation, transdermal administration, subcutaneous injection, intradermal injection, superficial injection, intravenous (IV) injection, intra-arterial injection, intramuscular injection, intracardiac injection, intraosseous injection, intraperitoneal injection, transmucosal administration, vaginal administration, intraurethral administration, intravitreal administration, intraorbital administration, subretinal administration, intra-articular administration, peri-articular administration, local administration, epicutaneous administration, or any combinations thereof. The present disclosure thus encompasses methods of delivering any of the recombinant nucleic acids, viruses, medicaments, and/or compositions or formulations described herein to an individual.

[0321] In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions or formulations are administered cutaneously, topically, transdermally, subcutaneously, or intradermally to the subject. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions or formulations are administered intradermally and/or subcutaneously. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions or formulations are administered via superficial injection. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions or formulations are administered topically and/or transdermally. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or composi-

tions or formulations are administered one, two, three, four, five or more times per day. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions or formulations are administered to one or more affected (e.g., wrinkled) areas of an individual. In some embodiments, the composition is administered to one or more unaffected areas of the individual.

[0322] In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered at a superficial depth in the skin. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are introduced into the skin at a depth of about 2000 microns or less. For example, the recombinant nucleic acids, viruses, medicaments, and/or compositions may be administered into the skin at a depth of about 2000 microns or less, at about 1750 microns or less, at about 1500 microns or less, at about 1250 microns or less, at about 1000 microns or less, at about 900 microns or less, at about 800 microns or less, at about 700 microns or less, at about 600 microns or less, or at about 500 microns or less. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are introduced at an injection depth between about 0.5 mm and 5.0 mm. For example, the recombinant nucleic acids, viruses, medicaments, and/or compositions may be introduced at an injection depth of between about 0.5 mm and 5.0 mm, about 0.5 mm and 4.5 mm, about 0.5 mm and 4.0 mm, about 0.5 mm and 3.5 mm, about 0.5 mm and 3.0 mm, about 0.5 mm and 2.5 mm, about 0.5 mm and 2.0 mm, about 0.5 mm and 1.5 mm, about 0.5 mm and 1.0 mm, 1.0 mm and 5.0 mm, about 1.0 mm and 4.5 mm, about 1.0 mm and 4.0 mm, about 1.0 mm and 3.5 mm, about 1.0 mm and 3.0 mm, about 1.0 mm and 2.5 mm, about 1.0 mm and 2.0 mm, about 1.0 mm and 1.5 mm, 1.5 mm and 5.0 mm, about 1.5 mm and 4.5 mm, about 1.5 mm and 4.0 mm, about 1.5 mm and 3.5 mm, about 1.5 mm and 3.0 mm, about 1.5 mm and 2.5 mm, about 1.5 mm and 2.0 mm, etc.

[0323] In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered by injections spaced apart by a distance of between about 1 mm to about 30 mm. For example, the recombinant nucleic acids, viruses, medicaments, and/or compositions may be administered by injections spaced apart by a distance of between about 1 mm and 30 mm, 2 mm and 30 mm, 5 mm and 30 mm, 10 mm and 30 mm, 15 mm and 30 mm, 1 mm and 20 mm, 2 mm and 20 mm, 5 mm and 20 mm, 10 mm and 20 mm, 15 mm and 20 mm, 1 mm and 15 mm, 2 mm and 15 mm, 5 mm and 15 mm, 10 mm and 15 mm, 1 mm and 10 mm, 2 mm and 10 mm, 5 mm and 10 mm, 1 mm and 5 mm, 2 mm and 5 mm, etc. In some embodiments, the injections are spaced apart by a distance of at least about 1 mm, at least about 2 mm, at least about 5 mm, at least about 10 mm, at least about 15 mm, at least about 20 mm, at least about 30 mm or more.

[0324] Numerous areas of the body may be treated with the recombinant nucleic acids, viruses, medicaments, and/or compositions described herein, including, for example, the face, forehead, lips, scalp, neck, arms, hands, legs, knees, feet, chest, back, groin, buttocks, thighs, etc. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to the face. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more nasolabial folds. In some embodiments, the

recombinant nucleic acids, viruses, medicaments, and/or compositions are administered around one or both eyes. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more crows' feet. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more frown lines. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more scars (e.g., acne scars). In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more glabellar lines. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more brow ptosis. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more deep tear troughs. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more nasojugal lines. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more bunny lines. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more marionette lines. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or sites of poppy dimpling. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more chin creases. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more neck lines. In some embodiments, the recombinant nucleic acids, viruses, medicaments, and/or compositions are administered to one or more platysma bands.

[0325] In some embodiments, the recombinant nucleic acid expresses the cosmetic protein(s) (e.g., human collagens) when the recombinant nucleic acid is delivered into one or more target cells of a subject. In some embodiments, expression of the cosmetic protein(s) (e.g., human collagens) enhances, increases, augments, and/or supplements the levels of human collagen in one or more target cells. In some embodiments, expression of the cosmetic protein(s) (e.g., human collagens) enhances, increases, augments, and/or supplements the levels of human collagen secreted by one or more target cells. In some embodiments, expression of the cosmetic protein(s) (e.g., human collagens) enhances, increases, augments, and/or supplements the levels of human collagen in the extracellular matrix. In some embodiments, expression of the cosmetic protein(s) (e.g., human collagens) enhances, increases, augments, and/or supplements the stability of the extracellular matrix in the subject. In some embodiments, expression of the cosmetic protein(s) (e.g., human collagens) enhances, augments, and/or supplements the soft tissue of the subject. In some embodiments, expression of the cosmetic protein(s) (e.g., human collagens) improves the skin quality, condition, and/or appearance of the individual. In some embodiments, expression of the cosmetic protein(s) (e.g., human collagens) reduces one or more superficial depressions (e.g., wrinkles) in the skin of the subject. In some embodiments, expression of the cos-

metic protein(s) (e.g., human collagens) improves the texture, smoothness, elasticity, and/or tension of the skin of the subject. In some embodiments, expression of the cosmetic protein(s) (e.g., human collagens) reduces one or more dermatological signs of aging in the subject.

VII. Host Cells

[0326] Certain aspects of the present disclosure relate to one or more host cells comprising any of the recombinant nucleic acids described herein. Any suitable host cell (prokaryotic or eukaryotic) known in the art may be used, including, for example: prokaryotic cells including eubacteria, such as Gram-negative or Gram-positive organisms, for example Enterobacteriaceae such as *Escherichia* (e.g., *E. coli*), *Enterobacter*, *Erminia*, *Klebsiella*, *Proteus*, *Salmonella* (e.g., *S. typhimurium*), *Serratia* (e.g., *S. marcescans*), and *Shigella*, as well as Bacilli such as *B. subtilis* and *B. licheniformis*; fungal cells (e.g., *S. cerevisiae*); insect cells (e.g., S2 cells, etc.); and mammalian cells, including monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651), human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture), baby hamster kidney cells (BHK, ATCC CCL 10), mouse Sertoli cells (TM4), monkey kidney cells (CV1 ATCC CCL 70), African green monkey kidney cells (VERO-76, ATCC CRL-1587), human cervical carcinoma cells (HELA, ATCC CCL 2), canine kidney cells (MDCK, ATCC CCL 34), buffalo rat liver cells (BRL 3A, ATCC CRL 1442), human lung cells (W138, ATCC CCL 75), human liver cells (Hep G2, HB 8065), mouse mammary tumor (MMT 060562, ATCC CCL51), TRI cells, MRC 5 cells, FS4 cells, human hepatoma line (Hep G2), Chinese hamster ovary (CHO) cells, including DHFR^r CHO cells, and myeloma cell lines such as NS0 and Sp2/0. In some embodiments, the host cell is a human or non-human primate cell. In some embodiments, the host cell is a Vero cell. In some embodiments, the host cell is a complementing host cell. In some embodiments, the host cell (e.g., the Vero cell) expresses one or more herpes simplex virus genes (e.g., an ICP4 gene). In some embodiments, the host cells are cells from a cell line. Examples of suitable host cells or cell lines may include, but are not limited to, 293, HeLa, SH-Sy5y, Hep G2, CACO-2, A549, L929, 3T3, K562, CHO-K1, MDCK, HUVEC, Vero, N20, COS-7, PSN1, VCaP, CHO cells, and the like.

[0327] In some embodiments, the recombinant nucleic acid is a herpes simplex viral vector. In some embodiments, the recombinant nucleic acid is a herpes simplex virus amplicon. In some embodiments, the recombinant nucleic acid is an HSV-1 amplicon or HSV-1 hybrid amplicon. In some embodiments, a host cell comprising a helper virus is contacted with an HSV-1 amplicon or HSV-1 hybrid amplicon described herein, resulting in the production of a virus comprising one or more recombinant nucleic acids described herein. In some embodiments, the virus is collected from the supernatant of the contacted host cell. Methods of generating virus by contacting host cells comprising a helper virus with an HSV-1 amplicon or HSV-1/hybrid amplicon are known in the art.

[0328] In some embodiments, the host cell is a complementing host cell. In some embodiments, the complementing host cell expresses one or more genes that are inactivated in any of the viral vectors described herein. In some embodiments, the complementing host cell is contacted with a recombinant herpes virus genome (e.g., a recombinant her-

pes simplex virus genome) described herein. In some embodiments, contacting a complementing host cell with a recombinant herpes virus genome results in the production of a herpes virus comprising one or more recombinant nucleic acids described herein. In some embodiments, the virus is collected from the supernatant of the contacted host cell. Methods of generating virus by contacting complementing host cells with a recombinant herpes simplex virus are generally described in WO2015/009952 and/or WO2017/176336.

VIII. Articles of Manufacture or Kits

[0329] Certain aspects of the present disclosure relate to an article of manufacture or a kit comprising any of the recombinant nucleic acids, viruses, medicaments and/or compositions or formulations (e.g., pharmaceutical compositions or formulations) described herein. In some embodiments, the article of manufacture or kit comprises a package insert comprising instructions for administering the recombinant nucleic acid, virus, medicament, and/or composition or formulation (e.g., to treat a dermal extracellular matrix protein (e.g., collagen) deficiency and/or to correct one or more dermatological signs of aging).

[0330] Suitable containers for the recombinant nucleic acids, viruses, medicaments and/or compositions or formulations may include, for example, bottles, vials, bags, tubes, and syringes. The container may be formed from a variety of materials such as glass, plastic (such as polyvinyl chloride or polyolefin), or metal alloy (such as stainless steel or hastelloy). In some embodiments, the container comprises a label on, or associated with the container, wherein the label indicates directions for use. The article of manufacture or kit may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, syringes, package inserts, and the like.

IX. Enumerated Embodiments

[0331] Embodiment 1: a composition comprising: (a) a herpes simplex virus (HSV) comprising a recombinant nucleic acid, wherein the recombinant nucleic acid comprises a first polynucleotide encoding a first polypeptide comprising a first human collagen protein, and (b) an excipient.

[0332] Embodiment 2: the composition of embodiment 1, wherein the recombinant nucleic acid comprises two or more copies of the first polynucleotide.

[0333] Embodiment 3: the composition of embodiment 1 or 2, wherein the HSV is replication-defective.

[0334] Embodiment 4: the composition of embodiment 1 or 2, wherein the HSV is replication-competent.

[0335] Embodiment 5: the composition of any one of embodiments 1-4, wherein the HSV is a herpes simplex type 1 virus, a herpes simplex type 2 virus, or any derivatives thereof.

[0336] Embodiment 6: the composition of any one of embodiments 1-5, wherein the recombinant nucleic acid is a herpes simplex virus amplicon.

[0337] Embodiment 7: the composition of embodiment 6, wherein the herpes simplex virus amplicon is an HSV-1 amplicon or an HSV-1 hybrid amplicon.

[0338] Embodiment 8: the composition of embodiment 7, wherein the HSV-1 hybrid amplicon is an HSV-AAV hybrid

amplicon, an HSV/EBV hybrid amplicon, and HSV/EBV/RV hybrid amplicon, or an HSV/Sleeping Beauty hybrid amplicon.

[0339] Embodiment 9: the composition of any one of embodiments 1-5, wherein the recombinant nucleic acid is a recombinant herpes simplex virus genome.

[0340] Embodiment 10: the composition of embodiment 9, wherein the recombinant herpes simplex virus genome is a recombinant HSV-1 genome, a recombinant HSV-2 genome, or any derivatives thereof.

[0341] Embodiment 11: the composition of embodiment 9 or 10, wherein the recombinant herpes simplex virus genome comprises an inactivating mutation in a herpes simplex virus gene.

[0342] Embodiment 12: the composition of embodiment 11, wherein the herpes simplex virus gene is selected from the group consisting of Infected Cell Protein (ICP) 0, ICP4, ICP22, ICP27, ICP47, thymidine kinase (tk), Long Unique Region (UL) 41, and UL55.

[0343] Embodiment 13: the composition of embodiment 12, wherein the recombinant herpes simplex virus genome comprises an inactivation mutation in one or both copies of the ICP4 gene.

[0344] Embodiment 14: the composition of embodiment 12 or 13, wherein the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP22 gene.

[0345] Embodiment 15: the composition of any one of embodiments 12-14, wherein the recombinant herpes simplex virus genome comprises an inactivation mutation in the UL41 gene.

[0346] Embodiment 16: the composition of any one of embodiments 12-15, wherein the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP0 gene.

[0347] Embodiment 17: the composition of any one of embodiments 12-16, wherein the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP27 gene.

[0348] Embodiment 18: the composition of any one of embodiments 11-17, wherein the inactivating mutation is a deletion of the coding sequence of the gene(s).

[0349] Embodiment 19: the composition of any one of embodiments 9-18, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within a viral gene locus.

[0350] Embodiment 20: The composition of any one of embodiments 9-19, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci.

[0351] Embodiment 21: the composition of any one of embodiments 9-20, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within the ICP22 viral gene locus.

[0352] Embodiment 22: the composition of any one of embodiment 9-21, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within the UL41 viral gene locus.

[0353] Embodiment 23: the composition of any one of embodiments 1-22, wherein the HSV has reduced cytotoxicity as compared to a wild-type herpes simplex virus.

[0354] Embodiment 24: the composition of any one of embodiments 1-23, wherein the first human collagen protein is selected from the group consisting of Collagen alpha-1(I)

chain polypeptide (COL1-1), Collagen alpha-2(I) chain polypeptide (COL1-2), a Collagen alpha-1(II) chain polypeptide (COL2), a Collagen alpha-1(III) chain polypeptide (COL3), a Collagen alpha-1(IV) chain polypeptide (COL4-1), a Collagen alpha-2(IV) chain polypeptide (COL4-2), a Collagen alpha-3(IV) chain polypeptide (COL4-3), a Collagen alpha-4(IV) chain polypeptide (COL4-4), a Collagen alpha-5(IV) chain polypeptide (COL4-5), a Collagen alpha-6(IV) chain polypeptide (COL4-6), a Collagen alpha-1(V) chain polypeptide (COL5-1), a Collagen alpha-2(V) chain polypeptide (COL5-2), a Collagen alpha-3(V) chain polypeptide (COL5-3), a Collagen alpha-1(VI) chain polypeptide (COL6-1), a Collagen alpha-2(VI) chain polypeptide (COL6-2), a Collagen alpha-3(VI) chain polypeptide (COL6-3), a Collagen alpha-4(VI) chain polypeptide (COL6-4), a Collagen alpha-5(VI) chain polypeptide (COL6-5), a Collagen alpha-6(VI) chain polypeptide (COL6-6), a Collagen alpha-1(VII) chain polypeptide (COL7), a Collagen alpha-1(VIII) chain polypeptide (COL8), a Collagen alpha-1(IX) chain polypeptide (COL9-1), a Collagen alpha-2(IX) chain polypeptide (COL9-2), a Collagen alpha-3(IX) chain polypeptide (COL9-3), a Collagen alpha-1(X) chain polypeptide (COL10), a Collagen alpha-1(XI) chain polypeptide (COL11-1), a Collagen alpha-2(XI) chain polypeptide (COL11-2), a Collagen alpha-1(XII) chain polypeptide (COL12), a Collagen alpha-1(XIII) chain polypeptide (COL13), a Collagen alpha-1(XIV) chain polypeptide (COL14), a Collagen alpha-1(XV) chain polypeptide (COL15), a Collagen alpha-1(XVI) chain polypeptide (COL16), a Collagen alpha-1(XVII) chain polypeptide (COL17), a Collagen alpha-1(XVIII) chain polypeptide (COL18), a Collagen alpha-1(XIX) chain polypeptide (COL19), a Collagen alpha-1(XX) chain polypeptide (COL20), a Collagen alpha-1(XXI) chain polypeptide (COL21), a Collagen alpha-1(XXII) chain polypeptide (COL22), a Collagen alpha-1(XXIII) chain polypeptide (COL23), a Collagen alpha-1(XXIV) chain polypeptide (COL24), a Collagen alpha-1(XXV) chain polypeptide (COL25), a Collagen alpha-1(XXVI) chain polypeptide (COL26), a Collagen alpha-1(XXVII) chain polypeptide (COL27), and a Collagen alpha-1(XXVIII) chain polypeptide (COL28).

[0355] Embodiment 25: the composition of any one of embodiments 1-24, wherein the first human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17.

[0356] Embodiment 26: the composition of any one of embodiments 1-25, wherein the nucleic acid sequence encoding the first human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1-14.

[0357] Embodiment 27: the composition of any one of embodiments 1-26, wherein the first human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 15-21.

[0358] Embodiment 28: the composition of any one of embodiments 1-27, wherein the first human collagen protein is not COL7.

[0359] Embodiment 29: the composition of any one of embodiments 1-28, wherein the first polypeptide comprises: (a) the first human collagen protein; (b) a further human collagen protein; and (c) a linker polypeptide linking (a) to (b).

[0360] Embodiment 30: the composition of embodiment 29, wherein the linker polypeptide is a cleavable linker polypeptide.

[0361] Embodiment 31: the composition of embodiment 29 or 30, wherein the linker polypeptide comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 28-31.

[0362] Embodiment 32: the composition of any one of embodiments 29-31, wherein the further human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28.

[0363] Embodiment 33: the composition of any one of embodiments 29-32, wherein the further human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17.

[0364] Embodiment 34: the composition of any one of embodiments 29-33, wherein the nucleic acid sequence encoding the further human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1-14.

[0365] Embodiment 35: the composition of any one of embodiments 29-34, wherein the further human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 15-21.

[0366] Embodiment 36: the composition of any one of embodiments 29-35, wherein the first human collagen protein and the further human collagen protein are different.

[0367] Embodiment 37: the composition of any one of embodiments 1-36, wherein the first polynucleotide encodes a polycistronic mRNA comprising: (a) a first open reading frame (ORF) encoding the first polypeptide; (b) a second ORF encoding an additional human collagen protein; and (c) an internal ribosomal entry site (IRES) separating (a) and (b).

[0368] Embodiment 38: the composition of embodiment 37, wherein the nucleic acid sequence encoding the IRES has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100%

sequence identity to a nucleic acid sequence selected from SEQ ID NO: 22 or SEQ ID NO: 23.

[0369] Embodiment 39: the composition of embodiment 37 or 38, wherein the additional human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28.

[0370] Embodiment 40: the composition of any one of embodiments 37-39, wherein the additional human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17.

[0371] Embodiment 41: the composition of any one of embodiments 37-40, wherein the nucleic acid sequence encoding the additional human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1-14.

[0372] Embodiment 42: the composition of any one of embodiments 37-41, wherein the additional human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 15-21.

[0373] Embodiment 43: the composition of any one of embodiments 37-42, wherein the first human collagen protein and the additional human collagen protein are different.

[0374] Embodiment 44: the composition of any one of embodiments 1-43, wherein the recombinant nucleic acid further comprises a second polynucleotide encoding a second human collagen protein.

[0375] Embodiment 45: the composition of embodiment 44, wherein the recombinant nucleic acid comprises two or more copies of the second polynucleotide.

[0376] Embodiment 46, the composition of embodiment 44 or 45, wherein the second human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28.

[0377] Embodiment 47: the composition of any one of embodiments 44-46, wherein the second human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17.

[0378] Embodiment 48: the composition of any one of embodiments 44-48, wherein the nucleic acid sequence encoding the second human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1-14.

[0379] Embodiment 49: the composition of any one of embodiments 44-48, wherein the second human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 15-21.

[0380] Embodiment 50: the composition of any one of embodiments 44-49, wherein the first and second human collagen proteins are different.

[0381] Embodiment 51: the composition of any one of embodiments 44-50, wherein the recombinant nucleic acid is a recombinant herpes simplex virus genome, and wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within a viral gene locus.

[0382] Embodiment 52: the composition of embodiment 51, wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within one or both copies of the ICP4 viral gene loci.

[0383] Embodiment 53: the composition of embodiment 51 or 52, wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within the ICP22 viral gene locus.

[0384] Embodiment 54: the composition of any one of embodiments 51-53, wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within the UL41 viral gene locus.

[0385] Embodiment 55: the composition of any one of embodiments 51-54, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci and the second polynucleotide within the ICP22 viral gene locus.

[0386] Embodiment 56: the composition of any one of embodiments 51-54, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci and the second polynucleotide within the UL41 viral gene locus.

[0387] Embodiment 57: the composition of any one of embodiments 1-56, wherein the excipient is adapted for cutaneous (systemic or topical), transdermal, subcutaneous, and/or intradermal administration.

[0388] Embodiment 58: the composition of any one of embodiments 1-57, wherein the excipient comprises a hydroxypropyl methylcellulose gel.

[0389] Embodiment 59: the composition of any one of embodiments 1-58, wherein the excipient is adapted for intradermal administration.

[0390] Embodiment 60: the composition of any one of embodiments 1-59, wherein the excipient comprises a phosphate buffer.

[0391] Embodiment 61: the composition of any one of embodiments 1-60, wherein the excipient comprises glycerol.

[0392] Embodiment 62: the composition of any one of embodiments 1-61, wherein the excipient comprises a lipid carrier.

[0393] Embodiment 63: the composition of any one of embodiments 1-62, wherein the excipient comprises a nanoparticle carrier.

[0394] Embodiment 64: the composition of any one of embodiments 1-63, wherein the composition is a cosmetic composition.

[0395] Embodiment 65: the composition of embodiment 64, wherein the cosmetic composition is a skin care product.

[0396] Embodiment 66: a kit comprising: (a) the composition of any one of embodiment 1-65; and (b) instructions for administering the composition.

[0397] Embodiment 67: a method of enhancing, increasing, augmenting, and/or supplementing the levels of one or more human collagen proteins in a subject, the method comprising administering to the subject an effective amount of the composition of any one of embodiments 1-65.

[0398] Embodiment 68: a method of enhancing, increasing, augmenting, and/or supplementing soft tissue of a subject, the method comprising administering to the subject an effective amount of the composition of any one of embodiments 1-65.

[0399] Embodiment 69: the method of embodiment 68, wherein the composition is injected into a soft tissue of the subject.

[0400] Embodiment 70: a method of improving skin quality, condition and/or appearance in a subject in need thereof, the method comprising administering to the subject an effective amount of the composition of any one of embodiments 1-65.

[0401] Embodiment 71: the method of embodiment 70, wherein the condition is selected from the group consisting of sun damage, aging, UV exposure, rough texture, skin sagging, wrinkles, and any combinations thereof.

[0402] Embodiment 72: a method of reducing the appearance of one or more superficial depressions in the skin of a subject in need thereof, the method comprising administering to the subject an effective amount of the composition of any one of embodiments 1-65.

[0403] Embodiment 73: the method of embodiment 72, wherein administration of the composition reduces the appearance of the one or more superficial depressions in the skin of the subject for at least about three months, at least about six months, at least about nine months, or at least about 12 months.

[0404] Embodiment 74: the method of embodiment 72 or 73, wherein the appearance of the one or more superficial depressions in the skin of the subject is reduced after administration of the composition, as compared to the appearance of the one or more superficial depression in the skin of the subject prior to administration of the composition.

[0405] Embodiment 75: a method of increasing and/or improving at least one of texture, smoothness, elasticity, or tension of the skin of a subject in need thereof, the method comprising administering to the subject an effective amount of the composition of any one of embodiments 1-65.

[0406] Embodiment 76: the method of embodiment 75, wherein the skin of the subject maintains at least one of an increased and/or improved texture, smoothness, elasticity, or tension for at least about three months, at least about six months, at least about nine months, or at least about 12 months after administration of the composition.

[0407] Embodiment 77: the method of embodiment 75 or 76, wherein at least one of texture, smoothness, elasticity, or tension of the skin of the subject is increased and/or improved after administration of the composition, as compared to the texture, smoothness, elasticity, or tension of the skin of the subject prior to administration of the composition.

[0408] Embodiment 78: the method of any one of embodiments 70-77, wherein the skin of the subject is aging skin.

[0409] Embodiment 79: the method of any one of embodiments 70-78, wherein the skin of the subject has been damaged due to exposure to ultraviolet light.

[0410] Embodiment 80: the method of any one of embodiments 70-79, wherein the skin of the subject is wrinkled.

[0411] Embodiment 81: a method of diminishing one or more dermatological signs of aging in a subject in need thereof, the method comprising administering to the subject an effective amount of the composition of any one of embodiments 1-65.

[0412] Embodiment 82: the method of embodiment 81, wherein the diminishing of one or more dermatological signs of aging is selected from the group consisting of: (a) treatment, reduction, and/or prevention of fine lines and/or wrinkles; (b) reduction of skin pore size; (c) improvement in skin thickness, plumpness, and/or tautness; (d) improvement in skin smoothness, suppleness, and/or softness; (e) improvement in skin tone, radiance and/or clarity; (f) improvement in procollagen and/or collagen production; (g) improvement in skin texture and or promotion of retexturization; (h) improvement in appearance of skin contours; (i) restoration of skin luster and/or brightness; (j) improvement of skin appearance decreased by aging and/or menopause; (k) improvement in skin moisturization; (l) increase in skin elasticity and/or resiliency; (m) treatment, reduction, and/or prevention or skin sagging; (n) improvement in skin firmness; (o) reduction of pigment spots, mottled skin, and/or acne scars; (p) improvement of optical properties of skin by light diffraction or reflection; and (q) any combinations thereof.

[0413] Embodiment 83: the method of embodiment 81 or 82, wherein the one or more dermatological signs of aging in the subject is diminished after administration of the composition, as compared to the one or more dermatological signs of aging in the subject prior to administration of the composition.

[0414] Embodiment 84: the method of any one of embodiments 67-83, wherein the subject is a human.

[0415] Embodiment 85: the method of any one of embodiments 67-84, wherein the composition is administered cutaneously (systemically or topically), transdermally, subcutaneously, or intradermally to the subject.

[0416] Embodiment 86: the method of any one of embodiments 67-85, wherein the composition is administered by superficial injection.

[0417] Embodiment 87: the method of any one of embodiments 67-85, wherein the composition is administered intra-dermally to the subject.

[0418] Embodiment 88: the method of any one of embodiments 67-87, wherein the composition is administered once to the subject.

[0419] Embodiment 89: the method of any one of embodiments 67-87, wherein the composition is administered at least twice to the subject.

[0420] Embodiment 90: the method of embodiment 89, wherein at least about 15, at least about 30, at least about 60, at least about 90, or at least about 120 days passes between administrations.

[0421] Embodiment 91: the method of any one of embodiments 67-90, wherein the composition is administered to one or more affected and/or unaffected areas of the subject.

[0422] Embodiment 92: the method of any one of embodiments 67-91, wherein the skin of the subject is abraded prior to administration.

[0423] Embodiment 93: a recombinant nucleic acid comprising a first polynucleotide encoding a first polypeptide comprising a first human collagen protein, wherein the recombinant nucleic acid is a recombinant herpes simplex virus genome.

[0424] Embodiment 94: the recombinant nucleic acid of embodiment 93, wherein the recombinant nucleic acid comprises two or more copies of the first polynucleotide.

[0425] Embodiment 95: the recombinant nucleic acid of embodiment 93 or 94, wherein the recombinant herpes simplex virus genome is a recombinant HSV-1 genome, a recombinant HSV-2 genome, or any derivatives thereof.

[0426] Embodiment 96: the recombinant nucleic acid of any one of embodiments 93-95, wherein the recombinant herpes simplex virus genome comprises an inactivating mutation in a herpes simplex virus gene.

[0427] Embodiment 97: the recombinant nucleic acid of embodiment 96, wherein the herpes simplex virus gene is selected from the group consisting of ICP0, ICP4, ICP22, ICP27, ICP47, tk, UL41, and UL55.

[0428] Embodiment 98: the recombinant nucleic acid of embodiment 97, wherein the recombinant herpes simplex virus genome comprises an inactivation mutation in one or both copies of the ICP4 gene.

[0429] Embodiment 99: the recombinant nucleic acid of embodiment 97 or 98, wherein the recombinant herpes simplex virus genome comprises an inactivating mutation in the ICP22 gene.

[0430] Embodiment 100: the recombinant nucleic acid of any one of embodiments 97-99, wherein the recombinant herpes simplex virus genome comprises an inactivation mutation in the UL41 gene.

[0431] Embodiment 101: the recombinant nucleic acid of any one of embodiments 97-100, wherein the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP0 gene.

[0432] Embodiment 102: the recombinant nucleic acid of any one of embodiments 97-101, wherein the recombinant herpes simplex virus genome comprises an inactivation mutation in the ICP27 gene.

[0433] Embodiment 103: the recombinant nucleic acid of any one of embodiments 96-102, wherein the inactivating mutation is a deletion of the coding sequence of the gene(s).

[0434] Embodiment 104: the recombinant nucleic acid of any one of embodiments 93-103, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within a viral gene locus.

[0435] Embodiment 105: the recombinant nucleic acid of any one of embodiments 93-104, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci.

[0436] Embodiment 106: the recombinant nucleic acid of any one of embodiments 93-105, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within the ICP22 viral gene locus.

[0437] Embodiment 107: the recombinant nucleic acid of any one of 93-106, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within the UL41 viral gene locus.

[0438] Embodiment 108: the recombinant nucleic acid of any one of embodiments 93-107, wherein the first human

collagen protein is selected from the group consisting of COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28.

[0439] Embodiment 109: the recombinant nucleic acid of any one of embodiments 93-108, wherein the first human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17.

[0440] Embodiment 110: the recombinant nucleic acid of any one of embodiments 93-109, wherein the nucleic acid sequence encoding the first human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1-14.

[0441] Embodiment 111: the recombinant nucleic acid of any one of embodiments 93-110, wherein the first human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 15-21.

[0442] Embodiment 112: the recombinant nucleic acid of any one of embodiments 93-111, wherein the first human collagen protein is not COL7.

[0443] Embodiment 113: the recombinant nucleic acid of any one of embodiments 93-112, wherein the first polypeptide comprises: (a) the first human collagen protein; (b) a further human collagen protein; and (c) a linker polypeptide linking (a) to (b).

[0444] Embodiment 114: The recombinant nucleic acid of embodiment 113, wherein the linker polypeptide is a cleavable linker polypeptide.

[0445] Embodiment 115: the recombinant nucleic acid of embodiment 113 or 114, wherein the linker polypeptide comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 28-31.

[0446] Embodiment 116: the recombinant nucleic acid of any one of embodiments 113-115, wherein the further human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28.

[0447] Embodiment 117: the recombinant nucleic acid of any one of embodiments 113-116, wherein the further human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17.

[0448] Embodiment 118: the recombinant nucleic acid of any one of embodiments 113-117, wherein the nucleic acid sequence encoding the further human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1-14.

[0449] Embodiment 119: the recombinant nucleic acid of any one of embodiments 113-118, wherein the further human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 15-21.

[0450] Embodiment 120: the recombinant nucleic acid of any one of embodiments 113-119, wherein the first human collagen protein and the further human collagen protein are different.

[0451] Embodiment 121: the recombinant nucleic acid of any one of embodiments 93-120, wherein the first polynucleotide encodes a polycistronic mRNA comprising: (a) a first open reading frame (ORF) encoding the first polypeptide; (b) a second ORF encoding an additional human collagen protein; and (c) an internal ribosomal entry site (IRES) separating (a) and (b).

[0452] Embodiment 122: the recombinant nucleic acid of embodiment 121, wherein the nucleic acid sequence encoding the IRES has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from SEQ ID NO: 22 or SEQ ID NO: 23.

[0453] Embodiment 123: the recombinant nucleic acid of embodiment 121 or 122, wherein the additional human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28.

[0454] Embodiment 124: the recombinant nucleic acid of any one of embodiments 121-123, wherein the additional human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17.

[0455] Embodiment 125: the recombinant nucleic acid of any one of embodiments 121-124, wherein the nucleic acid sequence encoding the additional human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1-14.

[0456] Embodiment 126: the recombinant nucleic acid of any one of embodiments 121-125, wherein the additional human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least

97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 15-21.

[0457] Embodiment 127: the recombinant nucleic acid of any one of embodiments 121-126, wherein the first human collagen protein and the additional human collagen protein are different.

[0458] Embodiment 128: the recombinant nucleic acid of any one of embodiments 93-127, wherein the recombinant nucleic acid further comprises a second polynucleotide encoding a second human collagen protein.

[0459] Embodiment 129: the recombinant nucleic acid of embodiment 128, wherein the recombinant nucleic acid comprises two or more copies of the second polynucleotide.

[0460] Embodiment 130: the recombinant nucleic acid of embodiment 128 or 129, wherein the second human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL2, COL3, COL4-1, COL4-2, COL4-3, COL4-4, COL4-5, COL4-6, COL5-1, COL5-2, COL5-3, COL6-1, COL6-2, COL6-3, COL6-4, COL6-5, COL6-6, COL7, COL8, COL9-1, COL9-2, COL9-3, COL10, COL11-1, COL11-2, COL12, COL13, COL14, COL15, COL16, COL17, COL18, COL19, COL20, COL21, COL22, COL23, COL24, COL25, COL26, COL27, and COL28.

[0461] Embodiment 131: the recombinant nucleic acid of any one of embodiments 128-130, wherein the second human collagen protein is selected from the group consisting of COL1-1, COL1-2, COL3, COL4-1, COL6-1, COL7, and COL17.

[0462] Embodiment 132: the recombinant nucleic acid of any one of embodiments 128-131, wherein the nucleic acid sequence encoding the second human collagen protein has at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NOS: 1-14.

[0463] Embodiment 133: the recombinant nucleic acid of any one of embodiments 128-132, wherein the second human collagen protein comprises a sequence having at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NOS: 15-21.

[0464] Embodiment 134: the recombinant nucleic acid of any one of embodiments 128-133, wherein the first and second human collagen proteins are different.

[0465] Embodiment 135: the recombinant nucleic acid of any one of embodiments 128-134, wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within a viral gene locus.

[0466] Embodiment 136: the recombinant nucleic acid of embodiment 135, wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within one or both copies of the ICP4 viral gene loci.

[0467] Embodiment 137: the recombinant nucleic acid of embodiment 135 or 136, wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within the ICP22 viral gene locus.

[0468] Embodiment 138: the recombinant nucleic acid of any one of embodiments 135-137, wherein the recombinant herpes simplex virus genome comprises the second polynucleotide within the UL41 viral gene locus.

[0469] Embodiment 139: the recombinant nucleic acid of any one of embodiments 135-138, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci and the second polynucleotide within the ICP22 viral gene locus.

[0470] Embodiment 140: the recombinant nucleic acid of any one of embodiments 135-138, wherein the recombinant herpes simplex virus genome comprises the first polynucleotide within one or both copies of the ICP4 viral gene loci and the second polynucleotide within the UL41 viral gene locus.

[0471] Embodiment 141: a host cell comprising the recombinant nucleic acid of any one embodiments 93-140.

[0472] Embodiment 142: the host cell of embodiment 141, wherein the host cell is a eukaryotic cell.

[0473] Embodiment 143: the host cell of embodiment 141 or 142, wherein the host cell is a mammalian cell.

[0474] Embodiment 144: the host cell of any one of embodiments 141-143, wherein the host cell is a human cell or a non-human primate cell.

[0475] Embodiment 145: the host cell of any one of embodiments 141-144, wherein the host cell is a Vero cell.

[0476] Embodiment 146: the host cell of any one of embodiments 141-145, wherein the host cell is a complementing host cell.

[0477] Embodiment 147: a method of collecting a herpes simplex virus, the method comprising: (a) contacting a complementing host cell with the recombinant nucleic acid of any one of embodiments 93-140; and (b) collecting the herpes simplex virus generated by the complementing host cell.

[0478] Embodiment 148: a method of collecting a herpes simplex virus, the method comprising: (a) culturing the host cell of any one of embodiments 141-146; and (b) collecting the herpes simplex virus generated by the host cell.

[0479] The specification is considered to be sufficient to enable one skilled in the art to practice the present disclosure. Various modifications of the present disclosure in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

EXAMPLES

[0480] The present disclosure will be more fully understood by reference to the following examples. It should not, however, be construed as limiting the scope of the present disclosure. It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art, and are to be included within the spirit and purview of this application and scope of the appended claims.

Example 1: Modified Herpes Simplex Virus Vectors Encoding Human Collagen Protein(s)

[0481] To make modified herpes simplex virus genome vectors capable of expressing human collagen protein(s) in a target mammalian cell (such as a human keratinocyte or fibroblast), a herpes simplex virus genome (FIG. 1A) is first modified to inactivate one or more herpes simplex virus genes. Such modifications may decrease the toxicity of the genome in mammalian cells. Next, variants of these modi-

fied/attenuated recombinant viral constructs are generated such that they carry one or more polynucleotides encoding human collagen protein(s). These variants include: 1) a recombinant ΔICP4/ΔICP22-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a heterologous promoter integrated at each ICP4 locus (FIG. 1B); 2) a recombinant ΔICP4-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a heterologous promoter integrated at each ICP4 locus (FIG. 1C); 3) a recombinant ΔICP4/ΔICP22-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a first heterologous promoter and the coding sequence of a second human collagen protein under the control of a second heterologous promoter on the same strand of DNA integrated at each ICP4 locus (FIG. 1D); 4) a recombinant ΔICP4-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a first heterologous promoter and the coding sequence of a second human collagen protein under the control of a second heterologous promoter on the same strand of DNA integrated at each ICP4 locus (FIG. 1E); 5) a recombinant ΔICP4/ΔICP22-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a first heterologous promoter and the coding sequence of a second human collagen protein under the control of a second heterologous promoter on opposite strands of DNA integrated at each ICP4 locus (FIG. 1F); 6) a recombinant ΔICP4-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a first heterologous promoter and the coding sequence of a second human collagen protein under the control of a second heterologous promoter on opposite strands of DNA integrated at each ICP4 locus (FIG. 1G); 7) a recombinant ΔICP4/ΔICP22-modified HSV-1 genome comprising expression cassettes encoding a polycistronic mRNA under the control of a heterologous promoter integrated at each of the ICP4 loci, where the polycistronic mRNA contains the coding sequence of a first human collagen protein and the coding sequence of a second human collagen protein separated by an internal ribosomal entry site (IRES) (FIG. 1H); 8) a recombinant ΔICP4-modified HSV-1 genome comprising expression cassettes encoding a polycistronic mRNA under the control of a heterologous promoter integrated at each of the ICP4 loci, where the polycistronic mRNA contains the coding sequence of a first human collagen protein and the coding sequence of a second human collagen protein separated by an internal ribosomal entry site (IRES) (FIG. 1I); 9) a recombinant ΔICP4/ΔICP22-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a chimeric protein under the control of a heterologous promoter integrated at each of the ICP4 loci, where the chimeric protein contains the amino acid sequence of a first human collagen protein and the amino acid sequence of a second human collagen protein separated by the amino acid sequence of a linker polypeptide (FIG. 1J); 10) a recombinant ΔICP4-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a chimeric protein under the control of a heterologous

promoter integrated at each of the ICP4 loci, where the chimeric protein contains the amino acid sequence of a first human collagen protein and the amino acid sequence of a second human collagen protein separated by the amino acid sequence of a linker polypeptide (FIG. 1K); 11) a recombinant ΔICP4/ΔICP22-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a heterologous promoter integrated at each of the ICP4 loci, and an expression cassette containing the coding sequence of a second human collagen protein under the control of a heterologous promoter integrated at the ICP22 locus (FIG. 1L); 12) a recombinant ΔICP4/ΔICP22/ΔUL41-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a heterologous promoter integrated at each of the ICP4 loci, and an expression cassette containing the coding sequence of a second human collagen protein under the control of a heterologous promoter integrated at the UL41 locus (FIG. 1M); and 13) a recombinant ΔICP4/ΔUL41-modified HSV-1 genome comprising expression cassettes containing the coding sequence of a first human collagen protein under the control of a heterologous promoter integrated at each of the ICP4 loci, and an expression cassette containing the coding sequence of a second human collagen protein under the control of a heterologous promoter integrated at the UL41 locus (FIG. 1N).

[0482] These modified herpes simplex virus genome vectors are transfected into engineered Vero cells that are modified to express one or more herpes virus genes. These engineered Vero cells secrete replication-defective herpes simplex virus with the modified genomes packaged therein into the supernatant. The supernatant is then collected, concentrated, and sterile filtered through a 5 µm filter.

Example 2: Construction and in Vitro Analysis of HSV Candidates Encoding Human COL7

[0483] Collagen alpha-1(VII) chain protein (COL7) functions to strengthen and stabilize the skin. Briefly, COL7A1 transcripts are translated, the resulting COL7 peptides are post-translationally modified by hydroxylation and glycosylation, and glycosylated COL7 tri-peptides form a triple helix known as pro-collagen, which is secreted from the cell. The pro-collagen associates into higher-order structures upon secretion, forming anchoring fibrils, which are then available to help organize, stabilize, and aid in the adherence of the epithelial basement membrane. The epithelial basement membrane is responsible for anchoring the epithelium to the underlying loose connective tissue and is essential for dermal-epidermal stability (dermo-epidermal junction integrity). Dystrophic epidermolysis bullosa is an inherited genetic condition caused by mutations in the COL7A1 gene; mutations in this gene impair the ability of COL7 to properly connect the epidermis to the dermis in dystrophic epidermolysis bullosa patients, leading to fragile skin. Recessive dystrophic epidermolysis bullosa (RDEB), the most severe form of epidermolysis bullosa, is most often characterized by extensive blistering and scarring of the skin and mucosal membranes.

[0484] The following example describes the construction of a recombinant herpes simplex type-1 viruses modified to express human Collagen alpha-1(VII) chain polypeptide (COL7), and further provides experiments showing that the recombinant HSV was capable of expressing functional

human collagen in vitro in primary human keratinocytes and fibroblasts from healthy and RDEB patients.

Materials and Methods

Virus Construction

[0485] The “KCA211” viral vector (FIG. 2A) was generated as follows: a wild-type herpes simplex virus genome was first modified by deleting the coding sequence of both copies of the viral ICP4 gene (ΔICP4). The ΔICP4-modified viral genome was also engineered to contain an mCherry expression cassette in each of the ICP4 loci. The viral genome was then further modified to encode wild-type human COL7. Briefly, a plasmid containing the coding sequence for wild-type COL7 (under control of the hCMV promoter) flanked by the upstream (US) and downstream (DS) regions of ICP4 was transfected into Vero cells modified to express the herpes virus ICP4 gene. These transfected cells were then infected with the modified ΔICP4 mCherry-expressing virus described above. The US and DS ICP4 regions flanking COL7 allowed for a double crossover and replacement of each of the mCherry loci. Visual screening for the absence of mCherry fluorescence was then used to identify cells containing recombined virus.

[0486] The “SAR-COL7” viral vector (FIG. 2B) was generated as follows: a wild-type herpes simplex virus genome was first modified by deleting the coding sequence of both copies of the viral ICP4 gene as well as the single copy ICP22 gene (ΔICP4/ΔICP22). The ΔICP4/ΔICP22-modified viral genome was also engineered to contain an mCherry expression cassette in each of the ICP4 loci. The viral genome was then further modified to encode wild-type human COL7. Briefly, a plasmid containing the coding sequence for wild-type COL7 (under control of the hCMV promoter) flanked by the upstream (US) and downstream (DS) regions of ICP4 was transfected into Vero cells modified to express the herpes virus ICP4 gene. These transfected cells were then infected with the modified ΔICP4/ΔICP22 mCherry-expressing virus described above. The US and DS ICP4 regions flanking COL7 allowed for a double crossover and replacement of each of the mCherry loci. Visual screening for the absence of mCherry fluorescence was then used to identify cells containing recombined virus.

Cell Culture

[0487] Cells were previously isolated from skin biopsies taken as part of routine surgical or diagnostic procedures. Informed written consent was obtained from each patient, or in the case of children, from the parent or legal guardian. This study was performed in accordance with the Helsinki declaration. All cells were cultured at 37° C. in 5% CO₂. Human fibroblasts were grown in Dulbecco's modified essential medium supplemented with 10% fetal bovine serum (PEAK® serum, cat. no. PS-FB1). RDEB and normal keratinocytes were cultured in DMEM/Ham's F12 medium (3:1) supplemented with 10% FBS, 10 ng/mL epidermal growth factor, 10-10 cholera toxin, 0.4 µg/mL hydrocortisone, 5 µg/mL transferrin, 5 µg/mL insulin, and 5 µg/mL liothyronine. All media contained ascorbic acid (150 µM). Keratinocytes were grown in the presence of a mitotically activated feeder layer of 3T3 cells and in the presence of 10 µM of the Rho-kinase inhibitor Y-27632.

Virus Infections

[0488] Viral aliquots were stored at -80° C., and were left to defrost under the tissue culture laminar flow hood before use. Target cell number was determined prior to infection after detaching cells from the tissue culture plastic and counting on a hemocytometer. Multiplicity of infection (MOI) was calculated from the virus titer and target cell number, and the appropriate volume of virus stock was diluted in 10% serum containing DMEM and incubated with the target cells for 2 hours at 37° C. Virus was then removed, and fresh media was supplied to target cells after washing twice with pre-warmed media.

Western Blots

[0489] Keratinocytes were plated in a 100 mm dish at 8×10⁵ to achieve 70-80% confluence the following day. 48 hours after infection, cells were lysed with radioimmuno-precipitation assay buffer. Lysate was placed in a centrifuge for 5 minutes at 4° C., and the supernatant was mixed with a 6x Laemmli loading buffer. Before loading onto SDS-PAGE, the samples were boiled for 5 minutes at 95° C. For COL7 detection, 5-30 µg of protein was loaded on a 6% acrylamide gel. The primary antibody used for COL7 detection was a rabbit antibody (Sigma, cat. no. HPA042420). Resolved proteins were transferred onto a nitrocellulose membrane, blocked in PBS-0.1% Tween with 5% milk or 5% BSA according to requirements of the primary antibody, and incubated overnight with the primary antibody. After incubation with IgG-HRP conjugated secondary antibody (Santa Cruz Biotechnology), the membrane was incubated with western blotting substrate (ThermoFisher Scientific, cat. no. 32106) and exposed to film (ThermoFisher Scientific, cat. no. 34090).

qRT-PCR

[0490] RNA was isolated using RNeasy® Mini Kit (Qiagen) according to the manufacturer's protocol. RNA extractions were quantified using a NanoDrop™ spectrophotometer (Fisher Scientific), and 1.5 µg RNA was used for cDNA synthesis using a SuperScript III First-Strand Synthesis system (Invitrogen). For qPCR, SYBR Select Master mix (Life Technologies) was used, and cDNA samples were diluted 1:25 to serve as template. Experiments were performed in triplicate.

Adhesion Assay

[0491] 96-well plates were left uncoated, or were coated with 10, 20, or 50 µg/mL rat tail collagen 1 (BD Biosciences) or human fibronectin (Millipore) in 100 µL reaction volume at 4° C. overnight, then washed with PBS, and blocked with PBS+0.1% BSA for 1 hour at 37° C. Mock (control) or SAR-COL7 infected RDEB keratinocytes (2.4×10⁴ cells in 100 µL of DMEM/HamF12+0.1% BSA) were added to the plates and incubated at 37° C. for 90 minutes. Wells were washed three times with PBS to remove any unbound cells, and adherent cells were fixed with PFE for 20 minutes. The fixed cells were then treated with 70% ethanol, stained with crystal violet, resolved in 100% ethanol, and quantified by measuring absorbance at 630 nM with a Flex Station 3 plate reader (Molecular Devices).

Organotypic Skin Equivalents

[0492] Bovine fibrinogen (90% clottable, MP Biomedicals) was dissolved in 1.1% NaCl at 37° C. for 4 hours and

then filtered with a 0.45 µm nylon membrane filter. Fibroblasts were collected with the use of trypsin and centrifugation and were resuspended in media to a final concentration of 2×10⁶ cells/mL. 150 µL of the cell suspension was mixed with 1 mL of thrombin (3 IU—Sigma Aldrich), and the cell/thrombin mix was added to fibrinogen at a ratio of 1:1. The mixture was quickly but gently distributed at 1 mL/well into a 12-well plate and incubated at 37° C. After 20 minutes, medium supplemented with ascorbic acid and aprotinin (Sigma) at a final concentration of 10 µg/mL was added. The matrices were left to mature for 5-7 days while medium was changed every other day. Keratinocytes were plated on top at 2×10⁶ cells/well, and on the next day the culture was raised to the air-liquid interface on a metal grid, and treatment with amlexanox was started. Medium was changed every other day with fresh drug, ascorbic acid and aprotinin. Cultures were collected at one or two weeks of treatment and frozen with OCT in liquid nitrogen-cooled isopentane. 8 µm sections were cut using a cryostat (Avantek QS11) and immunostained with polyclonal anti-COL7 antibody at a dilution of 1:800. Nuclei were counterstained with DAPI (Invitrogen).

Results

[0493] First, COL7 expression from modified HSV was assessed by qPCR and western blot analyses in the HaCaT human keratinocyte cell line to determine whether the modified HSVs were capable of expressing their cargo. HaCaT cells were transduced with either SAR-COL7 or KCA211 at MOIs ranging from 0.3-10. 48 hours after infection, cells were collected and processed for either qPCR (FIG. 3A) or western blot (FIG. 3B) analyses. The results demonstrated that full length COL7 was expressed in a dose-dependent manner from human keratinocytes infected with either modified HSV.

[0494] Next, immunofluorescence experiments were conducted to visualize COL7 expression in primary RDEB keratinocytes or RDEB fibroblasts infected for 24-48 hours with SAR-COL7 at various MOIs (ranging from 0.1 to 10). A strong COL7 signal was observed at all doses of SAR-COL7 tested for both keratinocytes (FIG. 4A) and fibroblasts (FIG. 4B), as compared to uninfected normal and RDEB keratinocytes and fibroblasts. The infection efficiencies of SAR-COL7 at MOIs of 0.1-1 in fibroblasts ranged from 16-36%. The infection efficiencies of SAR-COL7 at MOIs of 3.0 and above in fibroblasts were ≥90%.

[0495] RDEB patient-derived human dermal keratinocytes (HDKs) and fibroblasts (HDFs) were infected with SAR-COL7 at varying MOIs in order to evaluate COL7A1 RNA expression. Normal HDKs and HDFs, as well as mock infected RDEB HDKs and HDFs, were used as negative controls. Dose-dependent increases in COL7A1 transcripts were observed in both HDKs (FIG. 5A) and HDFs (FIG. 5B) infected with SAR-COL7. The relative fold change in COL7A1 transcript expression after infection vs. uninfected healthy HDKs (Table 1A) or uninfected healthy HDFs (Table 1B) was also calculated. The use of COL7A1-encoding HSV was capable of increasing COL7A1 transcript expression by approximately 26-fold in RDEB HDKs and 60-fold in RDEB HDFs relative to wild-type COL7A1 transcript levels in healthy HDKs and HDFs at an MOI of 3.

TABLE 1A

COL7A1 expression in HDKs

Cell type:	MOI:	Fold change over N-HDK:
Normal HDKs (N-HDK)	0	1.00
RDEB HDKs (EB-HDK)	0	0.79
RDEB HDKs (EB-HDK)	0.3	7.00
RDEB HDKs (EB-HDK)	1.0	13.73
RDEB HDKs (EB-HDK)	3.0	26.25

TABLE 1B

COL7A1 expression in HDFs

Cell type:	MOI:	Fold change over N-HDF:
Normal HDFs (N-HDF)	0	1.000
RDEB HDFs (EB-HDF)	0	0.340
RDEB HDFs (EB-HDF)	0.1	1.805
RDEB HDFs (EB-HDF)	0.3	5.134
RDEB HDFs (EB-HDF)	1.0	30.788
RDEB HDFs (EB-HDF)	3.0	60.571

[0496] Next, the functionality of human COL7 expressed from SAR-COL7 was tested by a cell adhesion assay. The ability of uninfected RDEB keratinocytes, and RDEB keratinocytes infected with SAR-COL7 at varying MOIs, to adhere to wells treated with Collagen 1 or fibronectin was studied. Interestingly, RDEB keratinocytes infected with SAR-COL7 showed increased adhesion to Collagen 1 (FIG. 6A) and fibronectin (FIG. 6B) in dose-dependent manner using a plate-based adhesion assay.

[0497] Finally, a skin equivalent (SE) organotypic culture composed of RDEB fibroblasts and keratinocytes was used to evaluate the expression of COL7 from SAR-COL7 at the Basement Membrane Zone (BMZ). Organ cultures were constructed with RDEB or normal fibroblasts and keratinocytes. RDEB cells were either infected with SAR-COL7 prior to organ culture construction (data not shown), or SAR-COL7 was added drop-wise onto the cultures prior to raising at the air-liquid interface (FIG. 7). The resulting skin equivalents (SEs) were isolated, sectioned, and stained for immunofluorescence to detect COL7 protein expression. COL7 was detected in these organotypic cultures from cells infected with SAR-COL7, and the initiation of COL7 protein deposition at the BMZ was observed. This data suggested that not only could SAR-COL7 deliver COL7A1 and express COL7 protein efficiently, but the COL7 protein began to organize in organotypic cultures similar to the pattern of organization expected for COL7 protein in vivo.

[0498] Taken together, the data provided herein indicated that replication-defective HSV may be employed as a vehicle for effectively delivering and expressing high levels of functional human collagen in wild-type primary human cells, as well as primary human cells isolated from patients suffering from a collagen deficiency, without any obvious toxicity in either 2D or 3D culture systems.

Example 3: In Vivo Analysis of an HSV Candidate Encoding Human COL7 in Wild-Type Animals

[0499] The following example describes experiments showing that recombinant viruses constructed and validated in vitro in human cells (see Example 2 above) were capable of expressing the encoded human collagen in vivo in wild-

type animals. The purpose of the study was, in part, to evaluate the skin biodistribution of HSV-mediated collagen expression in healthy immunocompetent animals.

Materials and Methods

Test Article

[0500] The active ingredient in the formulations administered to mice was the modified herpes simplex virus SAR-COL7 or KCA211 (see Example 2 above) at a titer of 4.8×10^8 plaque forming units (PFU)/mL formulated in PBS+ 10% glycerol. The vehicle used for intradermal administration was Dulbecco's phosphate-buffered saline (DPBS)+ 10% glycerol. The vehicle used for topical administration was 3% hydroxypropyl methylcellulose (HPMC) gel formulated in sterile double distilled water.

Animals

[0501] Healthy male BALB/c mice between 6 and 10 weeks of age were used. All procedures used in the protocol were in compliance with applicable animal welfare acts and were approved by the local Institutional Animal Care and Use Committee (IACUC).

Intradermal Injections

[0502] Prior to and during test article administration, mice were anesthetized using a cocktail of Dexdomitor (30 μ L or 0.05 mg/mL) and Telzol (50 μ L of 10 mg/mL). The sedative was reversed with Antisedan (50 μ L of 0.5 mg/mL). The back and flank areas were shaved using an electrical pet clipper, and the area was wiped with an alcohol wipe. Intradermal injections were performed using the Mantoux technique with a syringe and 27G needle (VWR, cat. no. BD305620), ensuring creation of a superficial “bleb” at each site. The virus was kept on dry ice, was thawed at room temperature, and was administered within 30 minutes of thawing. Two intradermal injections were administered to the back of each mouse, and the edges of the “bleb” were marked with a permanent marker.

Topical Application

[0503] Topical administration was conducted either on open wounds or abraded or scarified skin. The back and flank regions were shaved using electrical clippers. Scarification was performed by gently abrading the skin with a mechanical Dremel followed by superficial perforation with a 22G needle. For creation of a wound, a 5-6 mm diameter biopsy of the skin was removed using sharp scissors. In order to contain the topical formulation to the abraded or wounded site, a well was created from cut and autoclaved tops of 1.5 mL micro-centrifuge tubes. The lids were retained, and the cut side was covered with transparent adhesive dressing. The “wells” were adhered to the abraded/wounded region using surgical glue with the lid side down. 100 μ L of SAR-COL7 was mixed with 20 μ L of topical vehicle and was applied to the wound site by injection through the transparent adhesive in order to contain the topical gel on the wound and prevent leakage.

Tissue Collection

[0504] At the indicated time points following SAR-COL7 administration, the mice were euthanized, and the injection

site was removed using an 8 mm biopsy punch. One half of the biopsy was quick-frozen using liquid nitrogen, while the other half was embedded in OCT and cryopreserved for immunofluorescence staining.

Real Time Quantitative PCR

[0505] Quick-frozen biopsy halves were stored at -80° C. until analysis. For processing and analysis, samples were resuspended in 350 µL RLT buffer prepared with fresh DTT following the manufacturer's protocol (Qiagen). The sample were sonicated 3 times at 25% amplitude with intermittent incubation for 1 minute on ice, and DNA and RNA extractions were performed according to the manufacturer's protocol (Qiagen AllPrep DNA and RNA extraction kit). Both RNA and DNA samples were resuspended in distilled, deionized RNase free water and quantified spectrophotometrically on a Take3 microplate reader (BioRad).

[0506] Absolute quantification of COL7A1 DNA copies and RNA transcripts was performed by Tagman Real Time PCR analysis using a custom primer/probe assay that spanned the 3' end of the human COL7A1 open reading frame and the 3' UTR, specifically detecting the COL7A1 transgene. 100 ng of DNA and RNA was used for the qPCR and qRT-PCR assays respectively, and a plasmid standard containing the region to be amplified was prepared in 100 ng mouse genomic or RNA matrix. GAPDH was used as the control for both analyses.

Immunofluorescent Staining

[0507] OCT frozen tissue was sectioned at 5-8 µm and left to air dry for up to 1 hour. The slides were dipped in 100% methanol for 10 minutes at -20° C. and left to air dry. The methanol-fixed sections were rehydrated through 3 washes in PBS (5 minutes each) at room temperature. The sections were incubated with a blocking solution composed of 10% serum (mixed species) for 1 hour at room temperature in a humid chamber. The excess blocking solution was removed and a drop of primary antibody (anti-human collagen 7, Sigma, cat. no. HPA042420; anti-integrin alpha 6 (clone goH3), BD Biosciences, cat. no. 555734) solution, prepared in 5% blocking solution, was applied on each section (30-50 µL/section). The sections were incubated with the primary antibody for 16 hours at 4° C., washed 3 time in PBS for 5 minutes each at room temperature, and secondary antibody (anti-rabbit AF 647, Invitrogen, cat. no. A21244; anti-rat AF-594, Invitrogen, cat. no. A11007) was applied at a 1:400 dilution in PBS for 1 hour at room temperature in a humid chamber. The 3 times PBS wash was repeated, then slides were immersed in Hoechst solution (1:1000) for 5 minutes at room temperature. The 3 times PBS wash was repeated, and the stained sections were mounted with mounting media (Fluorometer G, Southern Biotech, cat. no. 0100-01) and covered with a coverslip. The sections were imaged after dehydration (approximately 24 hours) using a Widefield Fluorescence Microscope.

Results

[0508] A total of 30 male BALB/c mice divided into 6 groups were used for this study. SAR-COL7 was administered either by intradermal injection or topical application on day 1, and a subset of mice were harvested on day 3, and the remaining mice on day 6. Animals in group 2 received a low dose of SAR-COL7 (4.8×10^6 pfu/site) in the same volume

by intradermal injection, and group 1 served as a control for the intradermal cohorts. In groups 4, 5, and 6, the topical vehicle (group 4) or SAR-COL7 in topical gel (groups 5 and 6) was applied either to a wounded (groups 4 and 6) or abraded (group 5) area in a total volume of 120 µL. Tissues were harvested and processed for qPCR and immunofluorescence analysis as described above.

[0509] Post-sacrifice qPCR analysis was undertaken. COL7A1 transcripts and DNA levels were detected in all cohorts that received SAR-COL7 either by intradermal or topical application (FIGS. 8A-D), and a clear dose response was observed. Transcript and DNA levels were comparable between intradermal and topical cohorts (SAR-COL7 high ID, high wound, and high abraded), suggesting that topical application was as efficient at delivering COL7A1 as intradermal injection. Overall, the DNA and RNA levels were lower in the day 6 samples (FIGS. 8C-D) than the day 3 samples (FIGS. 8A-B), which was not unexpected since SAR-COL7 is a non-integrating vector that remains episomal (which would be expected to clear over time).

[0510] Next, immunofluorescence experiments were conducted to visualize COL7 expression after SAR-COL7 infection in vivo. As observed in the representative images provided in FIGS. 9A-B, COL7 was detected in most of the animal cohorts at both time points examined. Many samples showed correct localization of COL7 at the BMZ and around the hair follicles. In some instances, specifically with intradermal application, strong COL7 expression was observed in deeper layers of the skin, closer to the underlying fascia, possibly due to the injection being subcutaneous rather than intradermal. Similarly, in many of the abraded skin samples where the BMZ was likely removed during abrasion (as suggested by the lack of representative goH3 staining), the strong COL7 staining was limited to the skin surface. Overall, the presence of COL7 in the immunofluorescence samples strongly supported the robust efficacy of the modified HSV SAR-COL7.

[0511] In addition, the ability of KCA211 to express human COL7 was tested in vivo and compared to SAR-COL7 administration. KCA211 was found to also express the COL7 transgene in vivo in immunocompetent mice (FIGS. 10A-B).

[0512] Taken together, the data indicated that modified HSV was capable of delivering and expressing a human collagen protein in vivo in healthy, immunocompetent animals after topical or intradermal administration, and further, that collagen expression from topically administered virus into compromised skin or open wounds was comparable to intradermal administration into intact skin.

Example 4: In Vivo Analysis of Low and High Doses of an HSV Candidate Encoding Human COL7 in Hypomorphic Animals

[0513] The following example describes experiments showing that recombinant viruses constructed and validated in vitro in human cells (see Example 2 above) and in vivo in wild-type mice (see Example 3 above) were capable of expressing functional human collagen in vivo in COL7 hypomorphic mice. The purpose of the study was, in part, to evaluate the skin biodistribution of HSV-mediated collagen expression in COL7-deficient, immunocompetent animals.

Materials and Methods

[0514] Unless indicated to the contrary, experiments were conducted as described in Example 3 above.

Hypomorphic Mice

[0515] The COL7 hypomorphic mouse model (Fritsch et al. J Clin Invest. 2008 May; 118(5):1669-79) was used in this study. This hypomorphic mouse model is an immunocompetent animal model for dystrophic epidermolysis bullosa (DEB) in which the mice express about 10% of normal levels of COL7. Their phenotype closely resembles characteristics of severe human DEB, including mucocutaneous blistering, nail dystrophy, and mitten deformities of the extremities.

[0516] The mice were generated by flp/flrt-mediated removal of exon 2 of mouse COL7A1. Animals lacking both functional copies of COL7A1 (Col7a1flNew/flNeo), referred to as COL7 "hypomorphic mice", expressed about 10% of normal levels of COL7. From a total of 15 breeding pairs, 58 pups were obtained, with the litters ranging from 2-7 mice/litter. Out of these 58 pups, 6 were genotyped to be hypomorphs. Mice were genotyped with DNA extracted from an ear punch tissue sample. PCR analysis detected the presence of a loxP site upstream of exon 2 of COL7A1. Wild-type (WT) mice showed a band at 269 base pairs (bp), hypomorph mice showed a band at 435 bp, and heterozygous mice showed both bands. All procedures were in compliance with applicable animal welfare acts and were approved by the local Institutional Animal Care and Use Committee (IACUC).

Intradermal Injections

[0517] Prior to and during test article administration, mice were maintained under inhalation anesthesia using 2% isoflurane. Eye ointment (Puralube® Vet) was applied on the eyes to prevent dryness. Intradermal injections were performed using the Mantoux technique with a 31G needle. Up to four intradermal injections were administered to the back of each mouse at the specified doses.

Tissue Collection

[0518] Prior to tissue collection, animals were euthanized by CO₂ inhalation followed by cervical dislocation. The injection sites were biopsied using sharp scissors.

Hematoxylin and Eosin (H&E) Staining

[0519] Cryopreserved tissues were sectioned at a thickness of 5-8 µm and left to air dry for up to one hour. The slides were dipped in 100% methanol for ten minutes at -20° C. and left to air dry. Methanol fixed sections were rehydrated in PBS for 5 minutes at room temperature. The sections were incubated in hematoxylin (Weigert's modified hematoxylin) for 5-10 minutes at room temperature, followed by a wash in PBS for 15 minutes at room temperature. The sections were then rinsed in Eosin (Eosin Y solution, cat. no. HT110116) 3 times followed by one rinse in water.

[0520] The sections were gradually dehydrated with ethanol by dipping in 70% ethyl alcohol 10 times, 95% ethyl alcohol ten time, and 100% ethyl alcohol ten times. The sections were set to dry, mounted with mounting media (Fischer Scientific, cat. no. SPF15-100) and covered with a coverslip. The sections were imaged after dehydration (approximately 24 hours) using a bright field microscope.

Electron Microscopy

[0521] Skin was prepared for electron microscopy by immersion in 1.5% glutaraldehyde/1.5% paraformaldehyde in Dulbecco's serum free media (SFM) containing 0.05% tannic acid for a minimum of one hour, followed by an extensive rinse in SFM, and a post-fixation step in 1% OsO₄ for 60 minutes. The samples were washed in SFM then dehydrated in a graded series of ethanol to 100%, rinsed in propylene oxide, and infiltrated in Spurr's epoxy over a total time of two hours, accelerated via microwave energy. Samples were polymerized at 70° C. over 18 hours. Additional samples were prepared by extensively rinsing in SFM then immersing in mouse IgM LH24 antibody or mouse IgG NP185 antibody diluted 1:5 in SFM overnight at 4° C. The samples were then rinsed extensively in SFM, exposed to gold enhancement solution (Nanoprobes) for 15 minutes on ice, then rapidly warmed to 25° C. and incubated an additional 5 minutes. The samples were rinsed with ice cold SFM, fixed, and embedded.

Results

[0522] Three hypomorphic mice were used for the high-dose SAR-COL7 study. All mice received a dose of 4.6×10⁷ PFU/50 µL/injection site by intradermal injection on day 1 (Table 2). Each animal was shaved and injected at 4 sites on the back, including 1 control injection and 3 SAR-COL7 injections. One animal (mouse 3) received a second injection at the same 4 sites on day 3. One mouse (mouse 1) was sacrificed on day 3, while mouse 2 and mouse 3 were sacrificed on day 7.

TABLE 2

study design for intradermal injection of high-dose SAR-COL7				
Sample:	Mouse:	Treatment (Day 1):	Treatment (Day 3):	Day of Sacrifice:
1	1	HSV-GFP	PBS	Day 3
2		SAR-COL7	—	
3		SAR-COL7	—	
4		SAR-COL7	—	
5	2	PBS	PBS	Day 7
6		SAR-COL7	—	
7		SAR-COL7	—	
8		SAR-COL7	—	
9	3	PBS	PBS	Day 7
10		SAR-COL7	SAR-COL7	
11		SAR-COL7	SAR-COL7	
12		SAR-COL7	SAR-COL7	

[0523] Post-sacrifice qPCR analysis was undertaken. COL7A1 transcripts (FIG. 11A) and DNA levels (FIG. 11B) were detected at each viral injection site in all three mice. Transcript levels in all of the control samples (PBS or HSV-GFP) were at or below the level of detection in the assay, so only 1 control (day 3, sample 1, HSV-GFP) was included for comparison in the graphs. Some decrease in DNA and transcript levels by day 7 after single administration of SAR-COL7 (mouse 2) was observed; however, DNA and transcript levels increased upon re-administration of SAR-COL7 (mouse 3).

[0524] Next, immunofluorescence (IF) experiments were conducted to visualize COL7 expression in hypomorphic mice after SAR-COL7 infection *in vivo* (FIGS. 12A-B). The IF experiments demonstrated that robust and widespread

COL7 protein expression was observed in the BMZ, as well as around the hair follicles (HF), at both the day 3 and day 7 timepoints. No negative impact on skin morphology (even after repeat administration) was observed, as the SAR-COL7 treated samples showed a normal skin morphology with no obvious signs of fibrosis or acute inflammation (FIG. 13). Overall, the presence of COL7 in the immunofluorescence samples strongly supported the robust efficacy of SAR-COL7 in delivering human collagen capable of being secreted and appropriately organized in the underlying skin substructures.

[0525] Day 3 biopsies were also evaluated for anchoring fibril formation by electron microscopy. In appropriately structured anchoring fibrils, the NC1 domain of COL7 (which is stained with the NP158 antibody) aligns towards the lamina densa, while the NC2 domain of COL7 (which is stained with the LH24 antibody) aligns away from the lamina densa. Biopsies from SAR-COL7-injected mice showed COL7 staining with both the LH24 (FIG. 14A) and NP185 (FIG. 14B) antibodies, and importantly, the electron microscopy (EM) images revealed the formation of anchoring fibrils. The lamina densa was observed as a dark band through the middle of the EM images. The NC2 domains of the exogenous human COL7 were positioned away from the lamina densa, while the NC1 domains of the exogenous human COL7 were positioned along the lamina densa, as would be expected in properly formed anchoring fibrils. This data indicated that SAR-COL7 could not only express an encoded human COL7 that was capable of being secreted and appropriately organized at the BMZ, but that the secreted COL7 was functional and properly positioned in the resulting anchoring fibrils, supporting the skin tissue of the hypomorphic mice.

[0526] Three additional hypomorphic mice were used for the low-dose SAR-COL7 study. All mice received a dose of 6.4×10^6 PFU/50 μ L/injection site of SAR-COL7 in 3 (mouse 1) or 2 (mouse 2 and 3) sites by intradermal injection on day 1 (Table 3).

TABLE 3

study design for intradermal injection of low-dose SAR-COL7				
Mouse:	Treatment Day:	No. of Injection Sites:	Termination Day:	
1	1	3	3	
2	1	2	3	
3	1	2	7	

[0527] Post-sacrifice qPCR analysis was undertaken. COL7A1 transcripts (FIG. 15A) and DNA levels (FIG. 15B) were detected at each viral injection site in all three mice. A dose response was observed. While both DNA and transcript levels of COL7A1 were lower in these low-dose samples relative to those observed in high-dose study, COL7A1 expression was still detectable. Despite an approximate 2 log reduction in DNA or RNA levels, COL7 protein was still detectable in the skin of low-dose SAR-COL7 injected animals (FIG. 16). The COL7 protein was observed in both the BMZ and around the hair follicles. The data indicated that even at lower dosages, the modified HSV was still effective at delivering and expressing the encoded human collagen protein, which was then secreted and localized to the appropriate region of the skin.

[0528] Taken together, the data provided in the examples indicated that the modified HSVs described herein were able to: 1) transduce skin cells to produce functional human collagen in vitro and in organotypic culture (see Example 2); 2) transduce skin and deliver human collagen in vivo in healthy, immunocompetent animals after both topical and intradermal applications (see Example 3); and 3) transduce skin, deliver human collagen, and initiate the formation of anchoring fibrils in vivo in collagen-deficient animals (see Example 4). Without wishing to be bound by theory, it is believed that the recombinant vectors described herein provide a desirable strategy for delivering functional human collagen to the skin, allowing for novel approaches to support the dermis and/or epidermis in cosmetic settings.

Example 5: Construction and Validation of Multiple Engineered HSVs Encoding Human Collagen 1

[0529] The following example describes two different approaches for engineering a recombinant HSV-1 to successfully express human Collagen 1. Because human Collagen 1 is a heterotrimeric macromolecule comprising two copies of a COL1A1 polypeptide and one copy of a COL1A2 polypeptide, two distinct strategies were undertaken to generate a single HSV genome encoding both human COL1A1 and COL1A2; these two approaches are generally depicted in FIGS. 1H-1I and 1L, and described in detail below.

[0530] As a first approach, a recombinant HSV-1 was initially engineered to incorporate a human COL1A1 expression cassette, containing a heterologous promoter and polyA sequence, into each of the ICP4 loci as described in Example 2 above. Multiple plaques of viruses putatively containing the human COL1A1 cassette were picked and screened by infection in Vero cells to test for COL1A1 expression (data not shown). At least three isolates (clones 1A1, 2A1, and 3A1) were positive for human COL1A1 transgene expression. Next, the human COL1A1-positive clones 1A1 and 2A1 were then engineered to incorporate a human COL1A2 expression cassette, containing a different heterologous promoter than the COL1A1 cassette and its own polyA sequence, inserted into the ICP22 locus. Seven isolates (clones 1A2-1A1, 1B2-1A1, 1C1-1A1, 1B3-2A1, 3B1-1A1, and 3B1-2A1) were then screened to identify attenuated HSV-1 viruses capable of expressing both human COL1A1 and COL1A2, and thus, full-length human Collagen 1.

[0531] Vero cells were infected with one of the seven putative COL1A1-COL1A2-positive isolates or one of three of the COL1A1-alone isolates (as a negative control for COL1A2 expression) in order to identify recombinant HSV-1 clones capable of expressing both transgenes. Infections of the Vero cells were allowed to proceed for five days, cells were harvested by gentle scraping, and cell pellets were collected by centrifugation at 6000xg for five minutes. Half of each pellet was harvested for qRT-PCR analysis, while the other half was processed for western blotting. Absolute quantification of COL1A1 or COL1A2 RNA copies was performed by Tagman Real-Time PCR analysis using custom primer/probe assays that span the 3' end of the human COL1A1 or COL1A2 open reading frame (ORF) and the 3' untranslated region (UTR), and are specific to the transgene. RNA extracted from cells infected with all three COL1A1-alone viruses, and cells infected with 6/7 of the COL1A1-COL1A2 viruses, were positive for human COL1A1 trans-

gene expression (FIG. 17A). Interestingly, only 3/7 of the COL1A1-COL1A2 viruses (clones 1B2-1A1, 1C1-1A1, and 1B3-2A1) also expressed detectable levels of human COL1A2 (FIG. 17B). Western blotting of COL1A1 and COL1A2 protein expression in cell lysates infected with these three isolates was performed to ensure that the transgenes were expressed at both the nucleic acid and protein levels. Briefly, cell pellets were resuspended in RIPA buffer containing Halt protease inhibitor cocktail, incubated at 4° C. for ten minutes, and were then treated with benzonase for 10 minutes at room temperature. Samples were then diluted with 4×LDS sample buffer containing 5% 2-mercaptoethanol, and resolved on Tris-Glycine gels. Mock-infected Vero cell lysate was loaded on the gel as a negative control. Proteins were transferred to PVDF, blocked, and then stained overnight with primary antibody (anti-human COL1A1, ThermoFisher cat. no. PA5-29569; anti-human COL1A2, Abcam cat. no. ab96723; or anti-GAPDH, Abcam cat. no. ab9485). Stained membranes were then washed, incubated with secondary antibody, then developed. In agreement with the qRT-PCR analysis, all three clones were capable of expressing both human COL1A1 and COL1A2 proteins at detectable levels (FIG. 17C).

[0532] A second approach was pursued in parallel to generate COL1A1-COL1A2-positive recombinant, attenuated HSV vectors. Here, a polynucleotide construct was generated which encoded, from 5' to 3', a human COL1A1 ORF, a synthetic IRES, and a human COL1A2 ORF. This IRES-based construct (encoded in an expression cassette also containing a heterologous promoter and a polyA sequence) was then inserted into each of the ICP4 loci, as described in Example 2 above. Multiple plaques were picked and screened to identify vectors with correctly inserted IRES constructs. Briefly, Vero cells were infected with putative IRES viral isolates, infections were allowed to proceed for five days, and cells were harvested and processed for qPCR and qRT-PCR analysis. Absolute quantification of insert DNA and RNA copies was performed by Taqman Real-Time PCR analysis using custom primer/probe assays that recognize the synthetic IRES. 13 viral isolates were screened by qPCR and qRT-PCR analysis; only one of these isolates (isolate 6) tested positive for the IRES DNA and RNA transcripts comprising the IRES sequence (data not shown). To confirm that this isolate was capable of expressing both human COL1A1 and COL1A2 at the protein level, infected Vero cells were processed for western blotting as described above. A viral isolate (isolate 1) that showed no expression cassette incorporation by qPCR and transgene expression by qRT-PCR was used as a negative control. Paralleling the qPCR/qRT-PCR data, viral isolate 6 was capable of expressing both human COL1A1 and COL1A2 protein after infection (FIG. 18).

[0533] Taken together, the data presented in this sample indicate that: (1) multiple recombinant HSV-1 vectors were successfully constructed that were proficient in expressing both human COL1A1 and COL1A2 after infecting targeted cells; (2) vectors can be engineered to express heterotrimeric human collagen proteins; and (3) multiple different approaches can be taken to express multiple proteins from a single recombinant genome. Without wishing to be bound by theory, it is believed that successful expression of human Collagen 1 from a recombinant HSV-1 genome provides support for the use of engineered HSV to express any heterotrimeric collagen protein (e.g., human Collagen 4).

Example 6: Construction, Validation, and in Vitro Characterization of an Engineered HSV Encoding Human Collagen 3

[0534] The following example describes the engineering of a recombinant HSV-1 that successfully expressed human Collagen 3 (termed C3vec01). In addition, the following example describes in vitro experiments establishing multiple relevant 2D cell culture model systems suitable for characterizing the efficacy of C3vec01, including the use of immortalized human keratinocytes and fibroblasts in dose-ranging studies, the use of primary human dermal fibroblasts biopsied from multiple aged human patients as a model of C3vec01-mediated Collagen 3 rescue in older patients, and the use of in vitro UV-irradiated immortalized human fibroblasts as a model for sun exposure/skin aging.

[0535] Human skin is largely composed of collagen-rich connective tissue which is produced, organized, and maintained by dermal fibroblasts. Dermal collagen represents >90% (dry weight) of human skin and is composed primarily of COL1 and COL3 fibrils at a typical ratio of about 85:15. These fibrils provide strength to the skin and are critical for the maintenance of skin tissue architecture.

[0536] Skin aging characteristics are largely due to aberrant collagen homeostasis, resulting in a net collagen deficiency; biosynthesis of collagen is reduced, collagen fibril fragmentation is increased, and there is a progressive loss of dermal collagen, all of which contribute to the aged phenotype. Skin aging is influenced by a combination of both internal and external factors: intrinsic factors—the passage of time, genetics, cellular metabolism, hormones, etc.; and extrinsic—chronic light exposure, pollution, ionizing radiation, etc. These factors together lead to cumulative structural and physiological alterations to the skin, ultimately leading to the appearance of, and worsening in, skin wrinkles.

[0537] Skin rejuvenation, the process of reversing or repairing irregularities in the skin (such as wrinkles), is achieved, in part, by the synthesis of new collagen (neocollagenesis). In the skin, neocollagenesis is affected by the deposition of, and complex interactions between, collagens 1 and 3. COL3 appears early during collagen fibrillogenesis, and the subsequent replacement of this COL3 by COL1 is a critical step for collagen fibril maturation and extracellular matrix reorganization (Wang, et al., 2018, Journal of the Chinese Medical Association, 81(2), pp. 94-101). In addition, COL3 both regulates the dimensions of COL1 fibers (Liu, et al., 1997, Proc Natl Acad Sci USA, 94(5), pp. 1853-6) and enhances COL1 elasticity (Asgari, et al., 2017, Sci Rep, 7(1), p. 1392). As such, the appearance of early COL3 expression, and ensuing replacement with COL1, has been used as a marker of efficacy for injectable facial fillers in humans (Yutskovskaya, et al., 2014, J Drugs Dermatol, 13(9), pp. 1047-52).

[0538] All experiments were conducted as described above unless noted otherwise.

[0539] To begin, a recombinant HSV-1 was engineered to incorporate a human COL3A1 expression cassette, containing a heterologous promoter and polyA sequence, into each of the ICP4 loci. Multiple plaques of viruses putatively containing the human COL3A1 cassette were picked and screened by infection in Vero cells to test for COL3A1 expression (data not shown). One of the high expressing clones, termed C3vec01, was subsequently selected for additional in vitro (described below) and in vivo (Example 7) analyses.

[0540] First, a dose-ranging study was conducted to determine the efficacy of C3vec01-mediated delivery of its encoded human cargo in both immortalized human keratinocytes (FIG. 19) and immortalized human dermal fibroblasts (FIG. 20). The immortalized cells were infected for 48 hours at various multiplicities of infection (MOI) ranging from 0.3 to 3, and human Collagen 3 expression was quantitatively and qualitatively measured via multiple assays. Dose-dependent increases were observed in both effector DNA by qPCR analysis (data not shown) and effector transcript levels by qRT-PCR analysis in the immortalized HKs and HDFs (FIGS. 19A and 20A, respectively). Mock infected cells, and cells infected with a virus containing the same HSV-1 backbone as C3vec01 but instead encoding an mCherry effector, were used as negative controls. Paralleling these results, a dose-dependent increase in COL3 protein expression after C3vec01 infection was observed by immunofluorescence in immortalized HKs (FIG. 19B) and HDFs (FIG. 20B). While the immortalized HDFs expressed endogenous human COL3 prior to infection (as expected), a significant increase in COL3 expression was observed after infection with C3vec01, even at a low dose (primary anti-COL3 antibody, Abcam cat. no. ab7778). Little-to-no detectable endogenous COL3 was observed in the uninfected immortalized keratinocytes. Importantly, no significant effect on cell morphology or viability was observed in immortalized keratinocytes or fibroblasts infected with C3vec01, even at high doses.

[0541] As the skin ages, resident dermal fibroblasts produce less Collagen 3, and the ratio of COL1:COL3 in the skin skews towards Collagen 1. In order to provide skin rejuvenation through the synthesis of new Collagen 3, it was important to understand whether C3vec01 would be capable of effectively infecting aged dermal fibroblasts and robustly express its encoded human Collagen 3. As such, the ability of C3vec01 to infect aged primary dermal fibroblasts and express exogenous Collagen 3 at multiple MOIs was tested in cells sourced from two different vendors. Table 4 below provides donor information for the four primary HDF samples used in this study.

TABLE 4

primary human dermal fibroblast donors						
Age	Sex	Race	Tissue	Cat. No.	Lot No.	Company
73	M	Caucasian	Skin/ eyelid	C-12302	435Z009.2	PromoCell
65	F	Caucasian	Skin/ eyelid	C-12302	417Z010.2	PromoCell
73	M	Caucasian	Left lower back	CC-2511	0000633428	Lonza
75	F	Caucasian	Back	CC-2511	18TL057585	Lonza

[0542] As compared to the immortalized HDFs, a similar dose-dependent increase with comparable or higher COL3 transcript levels was observed after C3vec01 infection of the primary HDFs from vendor 1 (FIG. 21A) and vendor 2 (FIG. 21C). A sample of representative primary cells from each vendor was also tested for COL3 expression by western blot analysis (primary anti-COL3 antibody, Abcam cat. no. ab7778). C3vec01 was capable of rescuing high levels of human Collagen 3 expression in primary HDFs biopsied

from a 73-year-old male patient (FIR. 21B) and a 75-year-old female patient (FIG. 21D), even at the lowest MOI tested.

[0543] Finally, the ability of C3vec01 to induce robust Collagen 3 expression in UV-exposed immortalized human fibroblasts was tested. Sun exposure, and the corresponding UV damage, is known to be the single largest extrinsic contributor to the aged skin phenotype, and multiple groups have employed an *in vitro* skin fibroblast UV-exposure system to model certain aspects of photoaging (see e.g., Qin et al. 2018, *Cell Physiol Biochem* 46(5):1849-1860). To confirm that UV exposure caused human dermal fibroblasts to secrete less Collagen 3 (as would be expected given the phenotype of photo-aged skin), COL3 secretion into the supernatants of cultured immortalized human dermal fibroblasts was measured by ELISA before and after three different levels of UV exposure (FIG. 22A). Indeed, UV irradiation of cultured fibroblasts significantly reduced endogenous COL3 expression. COL1 expression were monitored in parallel in this experiment, and were not significantly affected by UV irradiation, indicating that UV exposure induced specific repression of COL3, as opposed to global suppression of protein synthesis. Next, the ability of C3vec01 to infect UV-irradiated immortalized HDFs and express exogenous COL3 was tested. Here, immortalized HDFs were exposed to UV-irradiation, and then allowed to recover for 24 hours prior to infection with C3vec01 at an MOI of 0.3 or 1. 48 hours after infection, exogenous human COL3 was assessed by qRT-PCR analysis. Strong COL3 expression was detected in C3vec01-infected, UV-irradiated HDFs at both tested MOIs (FIG. 22B), indicating that C3vec01 efficiently transduced photo-damaged cells and delivered its encoded cargo. Mock infected cells, and cells infected with a virus containing the same HSV-1 backbone as C3vec01 but instead encoding an mCherry effector, were used as negative controls to ensure specificity of transgene detection.

[0544] Taken together, the data presented in this example indicates that the recombinant HSV-1 vector C3vec01 efficiently transduces multiple human skin cell types, is capable of rescuing Collagen 3 expression in aged primary fibroblasts harvested from old patients, and is capable of salvaging Collagen 3 expression from UV-damaged HDFs. Without wishing to be bound by theory, it is believed that the data supports the use of a recombinant HSV encoding human Collagen 3 to correct the collagen defects of aged skin.

Example 7: In Vivo Characterization of Intradermally Administered C3Vec01

[0545] The following example described *in vivo* experiments establishing methods of intradermally administering C3vec01 in young and old healthy immunocompetent animals.

[0546] All experiments were conducted as described above unless noted otherwise.

[0547] All procedures conducted in this example were in compliance with applicable animal welfare acts and were approved by the local Institutional Animal Care and Use Committee (IACUC).

[0548] The backs of mice were shaved before further manipulations. C3vec01 (or vehicle control) was then injected intradermally to four sites in the backs of the mice. After infection and the subsequent recovery period, the animals were euthanized, and the treatment sites were removed using an 8 mm punch biopsy. One half of each

biopsy was quick-frozen in liquid nitrogen for qPCR/qRT-PCR analysis, while the other half was processed for immunofluorescence analysis.

[0549] Tissue samples were processed for nucleic acid and protein analysis as described above. For COL3 immunofluorescence staining, a rabbit anti-human Collagen 3 primary antibody (Abcam, cat. no. ab7778), and an Alexa Fluor® 488-conjugated secondary antibody were used. Tissue samples were mounted in mounting media containing DAPI to visualize nuclei.

[0550] An in vivo pharmacology study was conducted in young (6-8-week-old) and old (approximately 13-month-old) C57BL/6 mice to evaluate C3vec01-mediated expression of human COL3 in immunocompetent animals upon intradermal administration of the vector. A total of 10 animals were used for this study. The back of each mouse was first shaved and then intradermally injected with 2×10^8 PFU/site of C3vec01 (or vehicle control) at 4 sites/animal. Injected sites were biopsied at either 48-hours or 1-week post-dosing, and were evaluated for human COL3 expression by qPCR and immunofluorescence. Table 5 below provides a synopsis of the experimental design.

TABLE 5

study design and test article administration						
Group No.	N	Test Article	Mouse	Route of Administration	Volume of Test Article	Location, No. of Sites
1	1	Vehicle	Young	Intradermal	100 μ L	Back, 4
2	2	C3vec01	Young	Intradermal	100 μ L	Back, 4
3	2	C3vec01	Young	Intradermal	100 μ L	Back, 4
4	1	Vehicle	Old	Intradermal	100 μ L	Back, 4
5	2	C3vec01	Old	Intradermal	100 μ L	Back, 4
6	2	C3vec01	Old	Intradermal	100 μ L	Back, 4

[0551] Intradermal delivery of C3vec01 led to high levels of transduced vector genomes detected in skin biopsies harvested 48-hours post-injection (FIG. 23A), as well as yielded high levels of human Collagen 3 transcripts in both young and old mice (FIG. 23B). Immunofluorescence-based detection also showed visibly increased levels of human COL3 throughout the dermis in C3vec01-treated skin relative to vehicle-treated skin (FIG. 23C), correlating with transcript levels.

[0552] Taken together, the data provided in this example indicates that can efficiently transduce skin and express human Collagen 3 in vivo after intradermal injection. Without wishing to be bound by theory, it is believed that the in vivo study presented here lends further support for the use of HSV-1 as a novel gene therapy to delivery human Collagen 3 in the aesthetic setting.

Example 8: Generation and Validation of Modified Herpes Simplex Virus Vectors Encoding Human Laminins

[0553] To begin, recombinant herpes virus vectors were engineered to incorporate either wild-type or codon-opti-

mized variants of two human laminin proteins, LAMB3 or LAMC2, as described in Example 3 above. A number of isolates were picked for each type of virus. To test whether certain isolates were capable of expressing the encoded wild-type human LAMB3 protein, ICP4-complementing Vero cells were plated in 6-well plates and were infected with 12 untitered viral isolates of wild-type LamB3-encoding viruses until completion of infection. After infection, RNA was harvested, cDNA was generated, and expression of wild-type LamB3 from each isolate was determined by qPCR (FIG. 24A). All 12 isolates were capable of expressing wild-type human LamB3 in the transduced Vero cells at varying levels. The ability of 10 of these isolates to express human LamB3 was also tested by western blot. ICP4-complementing Vero cells were plated in 6-well plates and were infected with 10 untitered viral isolates of wild-type LamB3 expressing viruses until completion of infection. A well of Vero cells was transfected with a LamB3 expression plasmid as a positive control. After infection, the cells were collected by gentle scraping, centrifuged to collect cell pellets, culture medium was aspirated, and the cell pellets were washed once with PBS. Following washing, each cell pellet was resuspended in 200 μ L RIPA buffer containing protease inhibitors, and the suspensions were incubated at 4° C. for 20 minutes with gentle agitation every 5 minutes. After incubation, the samples were centrifuged at 17,000 \times g for 5 minutes, the supernatant was removed, and 4xLDS reducing sample buffer containing 5% 2-mercaptopethanol was added to each clarified supernatant. The samples were then boiled for 10 minutes before loading on a 4-20% Tris-Glycine polyacrylamide gel. After electrophoresis, the protein was transferred to a PVDF membrane, and the membrane was blocked for 30 minutes in 5% milk/TBS. Primary rabbit anti-LamB3 antibody (Abcam, cat. No. ab128864) was then added to the PVDF membrane at 1:1000 dilution in 5% milk/TBS and incubated overnight at RT° C. (~16 hours). The blots were then washed 3x for 5 minutes each with TBS, and then stained with an AP-conjugated goat anti-rabbit IgG antibody (Sigma, cat. No. A3687) in 5% milk/TBS for 1 hour at RT° C. The membranes were then washed 3x for 5 minutes each with TBS, BCIP/NBT was added, and the blots were developed for ~10 minutes at RT° C. In agreement with the qPCR data, all 10 viral isolates were capable of expressing the encoded wild-type human LamB3 at varying levels (FIG. 24B).

[0554] Viruses encoding codon-optimized variants of human LamB3 were also tested for their ability to express their cargo in Vero cells by western blot analysis. Briefly, 10 untitered viral isolates of codon-optimized (CO) LamB3-encoding viruses were used to infect Vero cells, cell pellets were collected, each pellet was resuspended in RIPA buffer containing protease inhibitors, and western blots were conducted using these cell lysates, as described above. All 10 viral isolates were capable of expressing the encoded codon-optimized human LamB3 in Vero cells (FIG. 25).

[0555] Next, four viral isolates encoding either wild-type or codon-optimized LAMB3 were tested for their capacity to transduce primary human cells and express their cargo. Immortalized primary normal keratinocytes were infected at a multiplicity of infection (MOI) of 1.0 for 48 hours. Uninfected cells were used as a negative control. Expression of LamB3 in the infected human keratinocytes was then examined by western blot. Western blots were carried out as described above using a primary rabbit anti-LamB3 anti-

body (Abcam, cat. No. ab 128864). In line with the data generated using Vero cells, the viral isolates expressing either wild-type and codon-optimized LamB3 were confirmed to effectively transduce primary human keratinocytes and express their encoded construct at suitable levels (FIG. 26).

[0556] To test whether LamC2-containing isolates were capable of expressing the encoded wild-type or codon-optimized human LAMC2, ICP4-complementing Vero cells were plated in 6-well plates and were infected with a number of untitered wild-type or codon-optimized LamC2-expressing viral isolates until completion of infection. After infection, RNA was harvested, cDNA was generated, and expression of wild-type LamC2 (FIG. 27A) or codon-optimized LamC2 (FIG. 27B) from each isolate was determined by qPCR. 8/12 isolates were capable of expressing wild-type human LamC2 in the transduced Vero cells at varying levels, while 3/7 isolates were capable of expressing codon-optimized human LamC2. The ability of certain wild-type and codon-optimized isolates to express human LamC2 was next tested by western blot. ICP4-complementing Vero cells were plated in 6-well plates and were infected with untitered viral isolates until completion of infection. A well of Vero cells was left uninfected as a negative control. After infection, the cells were collected by gentle scraping, centrifuged to collect cell pellets, culture medium was aspirated, and the cell pellets were washed once with PBS. Following washing, each cell pellet was resuspended in 200 µL RIPA buffer containing protease inhibitors, and the resuspensions were incubated at 4° C. for 20 minutes with gentle agitation every 5 minutes. After incubation, the samples were centrifuged at 17,000xg for 5 minutes, the supernatant was removed, and 4xLDS reducing sample buffer containing 5% 2-mercaptoethanol was added to each clarified supernatant. The samples were then boiled for 10 minutes before loading on a 4-20% Tris-Glycine polyacrylamide gel. After electrophoresis, the protein was transferred to a PVDF membrane, and the membrane was blocked for 30 minutes in 5% milk/TBS. Primary rabbit anti-LamC2 antibody (Abcam, cat. No. ab96327) was then added to the PVDF membrane at 1:1000 dilution in 5% milk/TBS and incubated overnight at RT° C. (~16 hours). The blots were then washed 3x for 5 minutes each with TBS, and then stained with an AP-conjugated goat anti-rabbit IgG antibody (Sigma, cat. No. A3687) in 5% milk/TBS for 1 hour at RT° C. The membranes were then washed 3x for 5 minutes each with TBS, BCIP/NBT was added, and the blots were developed for ~10 minutes at RT° C. In agreement with the qPCR data, all 9 of the tested viral isolates were also able to express the encoded human LamC2 (FIG. 27C).

[0557] The codon-optimized LamC2-expressing viral isolate "LGA" was selected for further testing in human cells. Immortalized primary normal keratinocytes were infected with the LGA isolate at a multiplicity of infection (MOI) of 0.3, 1.0, or 3.0 for 48 hours. Uninfected (control) and mCherry-expressing virus infected cells were used as a

negative control. DNA and RNA were extracted from the immortalized keratinocytes after 48 hours of infection, and qPCR/qRT-PCR was performed (FIGS. 28A-B). A good dose-response was observed for the LGA isolate in the immortalized keratinocytes, as assessed by viral genome copies detected per 50 ng of DNA (FIG. 28A). Interestingly, while a dose response was observed at the transcript level when increasing the MOI from 0.3 to 1.0, no additional increase in transcript levels were observed when increasing from an MOI of 1.0 to 3.0 (FIG. 28B). Expression of LamC2 in the infected human keratinocytes were also examined by western blot. Western blots were carried out as described above (primary rabbit anti-LamC2 antibody (Abcam, cat. No. ab96327) was used). In line with the transcript analysis, a dose response was observed at the protein level when increasing MOI from 0.3 to 1.0, but not from 1.0 to 3.0 (FIG. 28C).

[0558] Finally, to test whether isolate "LGA" was capable of expressing its human laminin when delivered in vivo, human LAMC2 expression was assessed by qPCR, qRT-PCR, and immunofluorescence after intradermal injection in animals. 1×10⁸ PFUs of LGA formulated in PBS+10% glycerol (vehicle) was intradermally injected into the dorsal skin and footpads of two mice. An equivalent volume of vehicle alone was intradermally administered to the dorsal skin of one mouse to act as a negative control. A schematic of the injection sites for the two animals treated in this study is provided in FIG. 29A.

[0559] 72 hours post-administration, a full thickness 8 mm biopsy was taken from each treatment site and split in half. One half of each section was flash frozen in liquid nitrogen and subsequently processed for qPCR and qRT-PCR analysis in order to quantify LAMC2 DNA copy numbers (FIG. 29B) and transcript levels (FIG. 29C) in the dorsal skin. The remaining half of each biopsy was embedded in OCT for immunofluorescence (IF). LAMC2 expression in cryosections was determined by immunofluorescent analysis using an anti-human LAMC2 antibody (Abcam, cat. no. ab96327). To confirm that the human LAMC2 expressed from LGA was correctly localized to the region of the skin where native laminin-332 is found, the dorsal skin samples were also counterstained for mouse laminin-332 (pKal). Intradermal administration of LGA led to successful transduction of mouse skin, with robust expression of the encoded human transgene in the correct layer of the epidermis. Histological evaluation showed no inflammatory infiltration at the treated site, demonstrating the safety of this therapy.

[0560] Taken together, the data presented in this example demonstrates that (1) HSV vectors could be successfully engineered to robustly express human laminin proteins (specifically, the R or γ subunit of human laminin-332), (2) an HSV-based vector (LGA) could successfully deliver a human laminin-332 subunit in vivo, and (3) the recombinant human laminin-332 subunit was expressed in, and localized to, the appropriate region of the epidermis of treated animals.

SEQUENCE LISTING

Sequence total quantity: 64
SEQ ID NO: 1 moltype = DNA length = 4395
FEATURE Location/Qualifiers
source 1..4395

-continued

mol_type = unassigned DNA
organism = Homo sapiens

SEQUENCE: 1

atgttcagct ttgtggacct ccggctcctg ctcccttag cggccaccgc cctcctgacg 60
cacggccaag aggaaggcca agtcgagggc caagacgaag acatcccacc aatcacctgc 120
gtacagaacg goctcaggta ccatgaccgc gacgtgtgga aacccgaggcc ctgcgggatc 180
tgcgtctgcg acaacggcaa ggtgttgcgatgacgt tctgtgacga gccaagaac 240
tgccccggcg cccaaatccc cgaggcggag tgcgttcggc tctgtcccgca cggctcagag 300
tcaccaccc accaagaaac caccggcgta gaggggacca aaggagacac tgcccccg 360
ggcccaaggg gaccccgagg cccccctggc cgagatggca tccctggaca gcctggact 420
ccccggaccgc cccggaccgc cggacccctgg gggacccctgg gcctggaggaa aacttttgt 480
ccccagctgtt atatggcta tgatgagaaa tcaaccggagaa gaatttcctgt gcctggcccc 540
atgggttcctt ctgggtctcg tgggttcctc ggccccctgg gtgcacctgg tcccaagggc 600
ttccaagggtc cccctggta gcctggcggag ctggagactt caggteccat gggcccccgg 660
ggtcccccccg gtccccctgg aaagaatggaa gatgatgggggaa aactgttgt 720
cctgggtgacg tgccggctcc tggggctca ggtgctcgag gattggcccg aacagctggc 780
ctccctggaa tgaagggaca cagagggttc agtgggttgg atgggtccaa gggagatgt 840
ggtgcgtctgtgtccctggaa tgaggcttgc agccctgggaaaatggagc tcctggctcg 900
atggggccccctggccctggc tgggtgagaaa ggtcccccggagggccctgg ccctgtctgg 960
gctcgtggaa atatgggtc tactgtgtgc cccggccccccctggcccccac cggccccgg 1020
ggtccctctgg tcttgggtgt gctaagggtt gaaatgggttcccaaggggcccc 1080
cgaggctctgg aagggtccccca ggtgtgtcgt ggtgagccctgg ccctggctgtt 1140
gctgctctggct ctgggtggaaa ccctgggtgt gatggacaggccctgggtcaat 1200
ggtgctctgt gtattctgg tgcgttcggc tttccctggg cccggcccccctggcc 1260
caggggcccccccgccggccctcc tgggtcccaag ggttaacagcg gttgaacctgg tgcgtctggc 1320
agcaaaaggag acactgtgtc taaggggagag cctggccctgg ttgggtgttca aggaccct 1380
ggccctgtgtg agagggggaa aaaggagaggag gttgggggttgc aaccgggaccactggctc 1440
ccccggacccttccggcggcgtc tgggtggactt ggtggccctgg gttggccctgg ccggagatgtt 1500
gttgctgtcc ccaagggtcc cgttgggtgaa cgttgggttccctggccctggcc 1560
ggatctctgggtgtcggatggcgtggatctgg tgggtccgggttggatggcttcc 1620
actggaaaccctggccggccctgggtctgtgtggcccaatgggttccctggcc 1680
caagatgggtc cccggggccca cccaggggccca cttgggtggcc tgggttggc tgggtgtatg 1740
ggattccctgt gacccaaagg tgcgtctggaa gagccggccca aggctggaga gcgagggtt 1800
ccccggacccttccggcgttgcgtgttccctggccctggccctggccctggcc 1860
ccccctggccctggccctggccctggccctggccctggccctggccctggcc 1920
ggatccagggtctggccctgg tcttgggtgttccctggccctggccctggcc 1980
cagggtgttccctggccctggccctggccctggccctggccctggccctggcc 2040
ttccctggcg acgtgtgtgt gcaagggtccc cttgggtccctggccctggccctggccctggccctggcc 2100
gttgctctggcc gaaacatgttgc tgcgttgggggatggccctggccctggccctggccctggcc 2160
ctggccctggccctggccctggccctggccctggccctggccctggccctggcc 2220
cctaagggttccctggccctggccctggccctggccctggccctggccctggcc 2280
ggcgccctggccctggccctggccctggccctggccctggccctggccctggcc 2340
aagggttggaaatggggccctggccctggccctggccctggccctggccctggcc 2400
gaccgtgttgc acgtgttgc cccggccctggccctggccctggccctggccctggccctggcc 2460
ggcccaacccctgtgtcaagggttgc cggaaatgttgc tggatggccctggccctggccctggcc 2520
cctggccctggccctggccctggccctggccctggccctggccctggccctggcc 2580
gccaagggttccctggccctggccctggccctggccctggccctggccctggcc 2640
ggccggatgtcgtccctggccctggccctggccctggccctggccctggccctggcc 2700
gttggcaagaaaggccggccctggccctggccctggccctggccctggccctggcc 2760
gaagttggtccctggccctggccctggccctggccctggccctggccctggcc 2820
gttgctctgtgtgttggccctggccctggccctggccctggccctggccctggcc 2880
gttggccctggccctggccctggccctggccctggccctggccctggccctggcc 2940
gaaacctggca aacaagggttccctggccctggccctggccctggccctggcc 3000
ggccccctggccctggccctggccctggccctggccctggccctggccctggcc 3060
gaagggttccctggccctggccctggccctggccctggccctggccctggcc 3120
ccccctggccctggccctggccctggccctggccctggccctggccctggcc 3180
ggcaagggttccctggccctggccctggccctggccctggccctggccctggcc 3240
gttggccctggccctggccctggccctggccctggccctggccctggccctggcc 3300
gaacaggccctggccctggccctggccctggccctggccctggccctggcc 3360
gcccctggccctggccctggccctggccctggccctggccctggccctggcc 3420
cgagggtccctggccctggccctggccctggccctggccctggccctggcc 3480
ccccatggccctggccctggccctggccctggccctggccctggccctggcc 3540
ggccctggccctggccctggccctggccctggccctggccctggccctggcc 3600
ctggccctggccctggccctggccctggccctggccctggccctggccctggcc 3660
gccaatgttggccctggccctggccctggccctggccctggccctggcc 3720
cagatcggataacatccggag cccggccctggccctggccctggccctggcc 3780
gacctcaaga tggccactc tggactggaaatggatggactggatggatgg 3840
gggtgtcaacc tggatggccat caaagtcttgc tggatggatggactggatgg 3900
taccccaactc acggccatgttgc ggttggatggccctggccctggccctggcc 3960
aaggaggatgttgc tggatggccctggccctggccctggccctggccctggcc 4020
cagggttccctggccctggccctggccctggccctggccctggccctggcc 4080
gaggccctccc agaacatcac ctaccactgca aagaacagcg tggccatcat ggaccagcc 4140
actggcaacc tcaagaaggc cctgttccctggccctggccctggccctggcc 4200
gaggccaaaca ggcgcgttccac ctacagctgc actgtgtgtatggccctggccctggcc 4260
gcctggggca accacagtgtat tgaatacaaa accaccaaga cccctggccctggccctggcc 4320
gtgtggccctggccctggccctggccctggccctggccctggccctggcc 4380

-continued

gtctgcttcc tgtaa	4395
SEQ ID NO: 2	moltype = DNA length = 4395
FEATURE	Location/Qualifiers
misc_feature	1..4395
note = Synthetic Construct	
source	1..4395
mol_type = other DNA	
organism = synthetic construct	
SEQUENCE: 2	
atgttcagct tcgtggacct gagactgctg ctgcttctgg ctgtacaga cctgtgaca	60
cacggacaag aggaaggcca ggcgaaggaa caggacgagg acatccccc tatcacctgt	120
gtcagaacg ccgttagata ccacgaccgg gatgtgtggaa agcccgagcc ttgcagaatc	180
tgcgtgtgc acaatggcaa ggtgtgtgc gacgacgtca tctgcgacga gacaaagaat	240
tgcgtgtgc cggaaagtcc tgaggcgaa tgggtgtcttg tggccctga tggcagcggag	300
agccccacac atcaagagac aacaggcgtq gaaggccccca agggcgatac aggacctaga	360
ggtcctagag gacgtggccgg accttcttgc agagatggaa ttctggaca gcctggactg	420
cccgacccat ctggacccctt aggggcttca ggttcacccag gactcgagg aaattttgcc	480
ccacagctga gtcacggcta cgacggaaa agcacaggcg gcatctgtt gcctggaccc	540
atggacccctt ctggcccaag aggacttctt tggctccagg acctcaggga	600
tttcaaggac caccaggcga acctggcgaa ccaggcgctt gtggtcaat gggaccaaga	660
ggcccttcgtt ggccacccagg caaaaatggc gacgatggcg aagccggaaa gcctggagg	720
cctggcggaa gaggccggcc aggacccgca ggcgcgttagag ggttgcctgg aactgcagg	780
ctgcctggca tgaaggggcca cagaggcttt tctggactgg atggcgttaa gggcgacgct	840
ggaccacgac gacctaaggg cgacgcttgg tctctggcg agaardggc acctggacag	900
atgggtccca gaggattgcc aggccgagaa ggttagaccc ggcgtccggg accagccgg	960
gtcagagggaa atgtggcgc aacaggctgt gtcggggcctt ctggaccaac cggaccgt	1020
ggccacccctt ctggggccagg cctgttggaa gcaaaaggccg aaggccggcc acaaggaccc	1080
agggggatctg aagggtctca ggggggtttaga ggcgcggccg ggccacccctt ggctggccgt	1140
gcagctggac ctgctggaaa ccctgggtctt gatggacacg cagggtccca aggtgtcta	1200
ggggccggcc gaaattggcc cgctggccagg tttccggccg caaggaggcc atctggaccc	1260
caaggcccaag gggggccatc ggggttcaat ggaataggcg gagaggccgg cgcttctgg	1320
agtaaaggcg atactggcgc aaaaggcgaa cccggacccctt tggggagtca aggacccctt	1380
ggaccacgctg ggcggaggaa caaaagaggg ctagggggag aaccaggacc aacagggttc	1440
cctggccaccc ctggccggcc cggggaccc ggtatctgg gattccctgg cgcaatggc	1500
gtggccggac caaaaggacc tgcggcgaa agggggatcac cagggtctgc aggcccta	1560
ggttctccag gggggccatc caggactgtt cagacccggg gaagctggac tcccagggtc taaggactg	1620
acaggctcac caggatctt cggaccacac gaaaaaacag gaccccttccagg accggccagga	1680
caaggatgtt gacccatgtt tcctggccgg cctgggtcaaa gaggacaacg tggcgtgtat	1740
gggtttcttg gacccatgtt tgcggccgg gaaacccggaa aacggccggaa gaggggagtt	1800
cccgacccctt cagggtctgt tggacccgtt ggaaaggatg ggcggactgg tgccacaagg	1860
cctccaggcc gacccggacc acggccggag agggccgac aaggaccacg cggatcttca	1920
ggatttcagg gactgtccgg cctggcttggc cccgcctggcc aggcggggaa gccaggccaa	1980
cagggtgttc tggccatct tggggccctt ggttcttagcg gagettagagg cgaaggaggaa	2040
tttcctggcc aaaggggccgt tcagggttca ccggggccag ctggacccaa ggggtgcata	2100
ggtgccccag gcaatgacgg tgcttcaaggc gacgcggccg ccccgagggtc tctggatct	2160
caaggccgac ctggacttca gggaaatgtt ggcggaaacgg gactgtgtgg acttcccggt	2220
ccaaaaggcg ataggggaga tgggttctt aaggggccgtt atggcttcc tggaaaggat	2280
ggcgccatagg gctgtacagg cccaaattggc cttccggggac ctgtggccgc tccaggccat	2340
aaggccgaat ctggacccat tggacccgtt ggttcttacag gtgttaggg agcccaaggc	2400
gacccggggg accctgtgtcc accggaccc tgggttcttcc ctggacccccc tgccgtgtat	2460
ggtcacactg tgcttcaaggg cgacggccgg gacgctgtttt caaaggccg cgcgttcca	2520
cctggccatcc cccggaccc tggggccca qgaccttattt gaaatgtttt tgcccttgc	2580
gccaaggccg caagaggatc tgctggccca ccaggcgctt caggattccc aggtggccgt	2640
ggaaggatgtt gacccatggcc gccaatgtt aatgtgttcc acggccggcc gccaggacca	2700
gccccggcaag aagggtggaaa agggccatcc ggcggaaactg gccctgcagg cggcccaagg	2760
gaagtggggcc ctccaggacc tccggggcc gccggccaaa aaggatccc aggcggccat	2820
ggaccacccgg cgcttccccc aacaccaggat ccacaggaa ttgtggaca aaggggagtt	2880
gtcgccatgc caggacccat gggggaggaa gttttccat gactccctgg gccaaggccg	2940
gaacactggca aacaggccatc atctgggtcc acggccggaa gaggcccaacc agggaccaat	3000
ggtcctccatg gatggccagg cccttcttgc gatattgtt gagaagggtc tccaggccgc	3060
gagggtatctc ctggacgtt ggggttcttcc ggcggccagg gcatatgggg cggaaacaggc	3120
ccagctggac ctccaggccgc accccggccgtt ccaggccgcac caggacccctt tggccctgt	3180
ggaaaatctt gacccatgtt cggccatggc cccggccatcc cggccggacc tgggtggaccc	3240
gtgggtgtttt gaggacccgc tggaccacaa ggttcttagcg gacggccagg cggaaacaggc	3300
gagcaaggccg acaggccatc caaggccatcc agggatcc ggcggacttca gggccaccca	3360
ggggccgtt gaaatggccg cgacggccgg ccaaggccggat ctgtggtcc cggccggaccc	3420
agaggaccac ctgggttcttcc tgggtgcaccc gggaaaggccg gactgtatgg gctcccccgg	3480
cctattggcc caccctggacc tagggaaaga acaggccgac caggacccat tggaccaccc	3540
ggggccatctg gacccatggcc tcccttcttgc cccggatccgat cttccatgtt	3600
ctgcctccatc ctccatgttca gaaaggccat gacggccgcata gatattacag agccggccat	3660
gccaacgtcg tgcggggaccc agatctggaa gtggacccat cactgttcc gctgttccat	3720
cagatcgaga acatcgaga ccccgaggcc agcggaaaga accctgtccag aacctgtccg	3780
gacctgaaga tggccacccat cgttggaaatg tctggccatgtt actggatccg ccccaaccat	3840
ggctgtcaacc tggatgttccat caagggttccat tgcaacatgg aaacccggccgacatgtccg	3900
taccccccacac agccatctgtt ggtcagaaggaa aactgttccatc tcaacatggaa ccccaaggac	3960

-continued

aaggccgacg	tttgggtcg	cgagagcatg	accgatggct	ccagttga	gtatggccgc	4020
cagggtctcg	accctcgca	tgtgtctatc	cagotgaccc	tcctegggct	gatgtctaca	4080
gaggccagcc	agaacatcac	ctaccactgc	aagaacagcg	tgccctacat	ggatcagcag	4140
accggcaacc	tgaagaaggc	actgtgttt	cagggcagca	acgagatcga	gatcagagcc	4200
gaggggcaaca	gcccgttac	ctacagcgtg	acagtggatg	gctgcaccag	ccatacaggc	4260
gcttggggca	agaccgttat	cgagtacaag	accaccaaga	ccagcagact	gcccatcata	4320
gatgtggccc	ctctgtatgt	tggggcaccc	gatcaagagt	tccggttcga	tgtgggcca	4380
gtgtgtttcc	tgtaa					4395

```

SEQ ID NO: 3          moltype = DNA  length = 4101
FEATURE
source
        Location/Qualifiers
        1..4101
        mol_type = unassigned DNA
        organism = Homo sapiens
SEQUENCE: 3
atctcaact ttgtggatac gcggactttg ttgctgtt cagtaaccc ttgccttagca 60
acatgccaat ctttacaaga gggaaactgt aaaaaggccc cagccggaga tagaggacca 120
cgtggagaaa ggggtccccc aggccccccca ggcagagatgg gtgaagatgg tcccacaggc 180
cctccctgtc cacctgtcc tcctggcccc cctggcttcg tgggaaacct tgctgtctcg 240
tatgtatggaa aaggagttgg acttggccct ggaccaatgg gcttaatggg acctagaggc 300
ccacctgtgtc cagctggagc cccaggccct caaggtttcc aaggacctgc tggtaggcct 360
ggtagaacctg gtcaaaactgg tcctcgagggt gtcgtgttgc cagctggccc tcctggcaag 420
gtctggtaga atggtcaccc tggaaaaccc ggacgacactg gtgagagagg agttgttgg 480
ccacagggtt ctctgtgttt ccctggaaact cctggacttc tcggcttcaa aggatttagg 540
ggacacaatg gtctggatgg attgaaggga cagccgggtg tcctgggtgt gaagggtgaa 600
cctgggtgccc ctggtagaaaa tggaaactca ggtcaaaacag gagccggcgg gcttccttgt 660
gagagaggac gtgttgggtc ccctggccca gttggtgccg gtggcagtgta tggaaagtgg 720
ggccccctgg tccctctggg tctgtgttgc tcctcaggctt cccagggtgcc 780
cctggccca aagggtgaaat tggagctgtt ggttaacgtgt tcctctgtgg tcccccgggt 840
ccccctgtgtc aagtgggtt tccaggccctc tccggccccc ttggacccccc tggtaatcc 900
ggagacaaacg gtctgtactgg tgccaaagggt gtcgtgtggc ttccggcgggt tgctggggct 960
cccccgttcc ctggaccccg cggatattcc tcggccgttgc tggtgtccgg tgctacttgt 1020
ggccagaggac ttgttggta gccttggtca gctggcttca aaggagagag cggtaaacaaag 1080
ggtgagcccg gtcctgttgc gccccaaagggt ctcctgttgc ccagtgtgtg agaaggaaag 1140
agaggcccta atggggaaaggc tggatctggc gggccctccgg gaccccttgg gtcgagaggt 1200
agtctctgttgc ctctgtgttgc tcctgtgttgc gatggcagagc tcggcgtcat gggccctcc 1260
ggtagtgcgtg gtcaaatgg ccctgtgttgc tgccggaggac ctaatggaga tgctgtgtgc 1320
cctggggagc tggtgttcat gggacccaga ggtttccctgg gttcccttgg aaatatccgc 1380
cccgctggaa aagaaggtcc tgctggccct ctcggcatcg acggcagggcc tggcccaatt 1440
ggcccaggtg gaggacatggc aacattggat tccttggacc caaaggcccc 1500
actgggtatc ctggcaaaaaa cggtgataaa ggtcatgtg tcctctgtgg tgctgggggt 1560
gctccaggtc ctgtatggaaa caatgggtct cagggacccct tcggaccacca ggggttccaa 1620
ggtagaaaggt tgtaacagggt tcccccgggtc ctcagggtctc ggcttggccccc 1680
tcagggtcccg ctggtaggt tgccaaacca gggaaaagggt gtcctcatgg tgagtttgg 1740
ctccctgtgtc ctctgtgttgc aagaggggaa cggggcccccc caggtgtggg tggtgtgtcc 1800
ggtctactgt gtcttattgg aagccgaggt ctcctgtggc ccccccggggc tgatggaaac 1860
aagggtgaaat ctgggtgtgt tggtgtctgt tggtactgtc tcctctgtgg 1920
cttccaggag aagggggtgtc tgctggatca cttggaggca aggggaaaaa ggggtgaacct 1980
ggtagtgcagg tgtagaaattgg taaccctggc agagatgttgc tcctgtgtgc tcctgtgtct 2040
gttaggtggcc ctgggtctgc tggagccaca ggtgaccgggg gegaagctgg ggctgtgtgt 2100
ctctgtgttgc ctctgtgttgc tggggggaa ctcgggttgc acgtgtgtgg cggctctgt 2160
ggcccccaatg gatttgggtg tcctgtgttgc tggtgtggc aacatgggtc taaaggagaa 2220
agaggagccaa aagggtctaa gggtagaaaaa ggtgttgggtt gtcctccagg ccccttggg 2280
gctgtggcc cagctggtcc aaatggtccc cccggctctg ctggaaactg tggtgtatgg 2340
ggccccctgtgtatgtactgg tttccctggt gtcgtgttgc ggactgttcc cccaggaccc 2400
tctgggtattt ctggccctcc tgggtttggt ggtgtgggtt gggaaaagggt gttctgtgtt 2460
cttcctgtgttgc accaaagggtc agttggccca aactggagaa taggtgtca gtcctccct 2520
ggttctgtgtc gtgagaagggt tcctctgttgc gaggctgttgc tgctgttgc tcctggcaact 2580
ccagggttcccttcttcttgc tggtgttgc tgctccctgg ctggagagggt 2640
gaacgtgttgc taccagggtgt tgctgggtgt tggtgtggaaac ctgggttcttct tggtgtgtcc 2700
ggccctctgtgttgc tccctgtgttgc tggtgtgggtt gtcctgtgttgc caacgggtgtt 2760
cctgggtgaag ctgggtgttgc tggcaaccctt gggaaacgtatg tgccccccagg tcggcgtgtt 2820
caaccggac acaaggggaga ggcgggttac ctcggcaata ttggccctgg tggtgtgtca 2880
ggtagcacttgc tccctcatgg cccctgggtt ctcgggttgc aacatggaaa ccgtgggtaa 2940
actgggttcccttcttcttcttgc tggtgttgc tggtgtggt gcccaggagg ttcttagtggc 3000
ccacaaggccaa ttcctgtgttgc taaggggagag cccctggaa aggggccccccagg aggtttccct 3060
ggatggaaagg gacacaatgg attgtcaagggt ctgggttgc tggtgtgtca ccatgggtat 3120
caagggtgtc ctgggtccgtt ggggttctgttgc ggttcttgggtt gcccctgtgg tgcttctggc 3180
cctgggttgc aagatgttgc cactggatgt ctcgggttgc tggtgtgtcc 3240
ggccctctgtgttgc tccctgtgttgc cccctgggtt cccctggccccc tgctggaccc 3300
ccagggttac ggggtgttgc ttatgttgc ggttacgtatg gagacttctca cagggtgtac 3360
cagccctgtgtc cagcacccctc ttcctgtgttgc aaggactatg aagtgtatgc tactctgtt 3420
tctctcaaca accaggatgtt gacccttcttacttcttgc tggtgtgtca gtcctgtgtt 3480
cgacatgtcc gttgtgttgc actcagccac ccaggtgttgc tggtgtgttgc ctactgtt 3540
gaccctaacc aaggatgttgc tatgtgtatgtt atcaaaaggatgtt actgtgttgc ctctactgtt 3600
gaaacactgttca tccggccca acctgttgc aagactgttgc taggtgttgc 3660

```

-continued

aaggacaaga	aacacgtctg	gctaggagaa	actatcaatg	ctggcagcca	gtttaaatat	3720
aatgttaaag	gagtgaacttc	caaggaaatg	gctacccaac	ttgccttcat	gwgctgctg	3780
gccaactatg	cctctcagaa	catcaccta	cactgcaaga	acagcattgc	atacatggat	3840
gaggagactg	gcaacctgaa	aaaggcttc	attctacagg	gctctaata	tgttgaactt	3900
gttgctgagg	gcaacagcag	gttcaatttc	actgttctt	tagatggctg	ctctaaaaag	3960
acaatgtaa	ggggaaagac	aatcatgaa	tacaaaacaa	ataaggccatc	acgcctgcc	4020
ttccttgata	ttgcacctt	ggacatcggt	ggtgctgacc	aggaattctt	tgtggacatt	4080
ggcccgatct	gtttcaaata	a				4101

```

SEQ ID NO: 4      moltype = DNA  length = 4101
FEATURE          Location/Qualifiers
misc_feature     1..4101
                  note = Synthetic Construct
source           1..4101
                  mol_type = other DNA
                  organism = synthetic construct
SEQUENCE: 4
atgctgacgt tcgtggacac cagaacactg ctgtgtctgg ccgtgacact gtgtctggcc 60
acttgtcaga gcctgcaaga gggaaacagtg cggaaaggac ctggccggca tagaggacct 120
agaggcggaa gaggtctcc tggaccttc ggttagagatg gcgaggatgg acctacagga 180
ccacctggtc caccaggacc tccaggccc cctggcctt gaggaaattt tgccgcttag 240
tacgatggca aaggcgctcg acttggccc ggacatctgg gacttatggg cccaaaggaa 300
ccaccagggt ctgcaggcgc tccaggacca caaggatttc aaggaccagc ttggcgaccct 360
ggcgaacctg gacaaacaggc tcctgtctgg gcttagaggac cagccggggc acctggaaaa 420
gttggcgaag atggcaccct tggaaagct ggttagaccgg gggaaagggg ttttttggaa 480
cctcaaggcg ccagaggctt tcctgaaaca cctggacttc ctggcttcaa gggcatcaga 540
ggccacaatg gcttgacccg actggaaaggc caacctgggt ctccctggcgt gaaaggcgaa 600
ccaggcgcac ctggcgaaaa tggccaccca ggacaaacccg ggcgaaggagg acttccctggc 660
gagagaggaa gagtttgagc cccagggtca gcaaggcgcac gaggatctga ttggatctgt 720
ggacctgtt gcccctggcg acctatttgg agtgctggcc tcctggatt tcctggcgca 780
cccgaggccaa aggccggaaat tggagctgtg gggaaacggcc gacctgcagg cccagctgg 840
ccaaaggggaa aagtggatt gcctggactg agcggacccg ttggggccacc agggaaatcc 900
gttgccaaatg gactgacagg cgctttaagggt gcagctggcc ttccaggcgt tgccggtgca 960
ccaggactgc caggaccaag aggtatccct ggtctctgtt gagctgtctgg cgctacgggt 1020
gcccaggacg ttgttggaga accttggggca gccggatcta agggcgagtc ttggaaacaag 1080
ggcggagccag gatctgttgc tcaccaaggc cccgctggac catcaggcga agaaggccaa 1140
cgaggcccta atggcgaagc cggtatgtcc gggctcttgc gaccacccgg cctttagagga 1200
tctctggct etagaggatt gccaggcgt gatggtagag caggcggtat gggtccaccc 1260
ggatcaagag gogctctgg ccctgtctgg gtttagagggt caaatggcga cgctggcaga 1320
ccaggcgcgc cccgttctat ggggttctaga ggggttctgg gaagggcttgc caatatccgc 1380
ccagccggaa aagaaggccc ttgttgactc cttggcatcg acggtagatc ttggaccatc 1440
ggaccggcag ggcgttgggg agagccttggg aatattggct tccttggccc taaaggcccc 1500
acaggcgtc ctggaaagaa cggcgataag ggccatcggt gactcgctgg tgcaaggggg 1560
gcacactggac ctgacggaaa caatgtgtc caaggccgcgcttggccaca aggtgttcaa 1620
gttggaaaaag gogagctaggg cccacactggg cttccagggt tccaaaggact tccctggacca 1680
tctggcccgag caggcgaagt tggaaagctt ggcgaaagag gactgcacgg cgagtttggc 1740
ttggccgggtc ttggcgttgc acggggggatc agaggccctc caggcgaatc ttggccggcga 1800
ggacactactg gcccatttcgg aaggcaggaa cttatgtggac ctccaggacc tgatggcaac 1860
aaaggcgaac ttgggttgtt gggcgctgtg ggaacagctg gaccttctgg tccttctgg 1920
ttgcccggcg agcgcggagc agctggattt cttgggtggca aaggcgaaaa gggcgagcc 1980
ggactcaagag gogagatcgg caatccggca cgagatggcg cttagaggccg cccagggtgca 2040
gttggtggccc cggggacatc tggcgcacca ggcgacacag gcgaggcttg tgccgtctgg 2100
cctggccggc caggcgttgc tagaggaaatg ccaggcggaa gggggcaatg gggaccctgt 2160
ggacccaatg gatttgcggc gctgtgttc aacctggcgc caaaggcgg 2220
cgggggagctt aaggcttcaaa aggccggaaa ggctgtgttgc gcccctactgg accagtggga 2280
gcagcagggtcc cccggcgttgc taacggacca cttggaccac ctgggtctat aggcgcggc 2340
ggaccggctgt gaatgcaggc ttccggggcc ggcgttggaa gaacagggttcc tccaggacca 2400
tctggcatct ctggtcacc acggccaccc ggtctctgtt gaaaagaagg actgagaggc 2460
cttagggccg atcagggtcc agtttgcggg accggcggaa tggggatgtt cggcccccacca 2520
gttggggccg cgccggggaa ccctgtggaa gaagctggaa ctggcggacc gccggggaaat 2580
cccggttctc aaggatgttgc tggggccctt ggaatttctgg gactggccccc tagtgcggga 2640
gaacgttggac tcccagggtt tggctggccg gtcggagaac cggggaccat ttggattgtt 2700
ggaccacccg gtgcgaagg tccacctggt gcagttggaa gtccctggcgt taacgggtct 2760
ccaggcgaag cccggcggaaa tggaaatcc ggcataatgg gcccggctgg gagaatggg 2820
cagectggac aaggatggca gcgaggcttc ccaggcgttgc ttggacccgtt cggccggcc 2880
gttgctcccg gacctcatgg tccctgtgtt ccaggcggggaa agcacggaaa tagggggaaa 2940
acaggacccctt cccgggttgc tggccacatc ggccgttgc gaccaaggagg cccatccgg 3000
cctcaggaa tccggccgaa taaggccggaa cctggcggaga agggacccat agggactgcct 3060
gggcttggaa gocataacgg actgtcaaggc ctggcggccca tcatggcgat 3120
caagggtgcac cccgttgcgtt gggcccccgg gggccgggggg gtccctgttgg tccatctgg 3180
cccgccggaa aagatggcag aacaggatcatc cttggcggacat tggggccctgc cggatattaa 3240
ggcccccagg gacatcaagg ccccggttgc cccggggggcc gccaggggcc 3300
ccaggcgttac ttggccgggg atacgatttc ggctacggc ggcacttcata cagaggccac 3360
cagccttagat ctggcccttag cctggggctt aaggactacg aagtggacgc cacactgaag 3420
tccctgaaca accagatcga gacactgtcg accccctgagg gcagcggaaa gaaccctgccc 3480
agaacactgtca gggacactgtt gactgtctca cccgaatgtt cttccggcata ctactggatc 3540

```

-continued

gaccccaatc	agggtgcac	catggacgc	atcaagggtg	actgcgactt	cagcacccgc	3600
gagacatgc	tcagagcc	gcctgagaa	atccccccca	agaacttgta	cagaacgcgc	3660
aaggacaaga	aacacgtgt	gtctggcgag	acaatcaacg	ccggcagc	gttcgagttac	3720
aacgtggaa	gogtgaccag	caaagagatg	gccacacagc	tggcttccat	gagactgtg	3780
gccaattacg	ccagccagaa	catcaccta	caactgcaaga	acagcattgc	ctatcatggac	3840
gaggaaaaccg	gcaacctgaa	gaaagccgtg	atccgtcagg	gctctaaacga	cgtggaaactg	3900
gtggcccgagg	gcaacagcg	attcaccta	accgtgtctg	tggacccgtg	cagcaaaaaag	3960
accaacgagt	ggggcaagac	catcatcgag	tataagacca	acaagcccg	cagactgccc	4020
ttctctggata	tcgccccact	ggatattgga	ggcgcgacc	aagagtctt	tgtggacatc	4080
ggcccccgtgt	gcttcagaatg	a				4101
 SEQ ID NO: 5			moltotype = DNA	length = 4401		
FEATURE			Location/Qualifiers			
source			1..4401			
			mol_type = unassigned DNA			
			organism = Homo sapiens			
SEQUENCE: 5						
atgtatggact	tttgtaaaaa	ggggagctgg	ctacttctcg	ctctgtttca	tcccactatt	60
atttttgcac	aacaggaa	tggttgaagg	ggatgttccc	atcttggtca	gtcttatgcg	120
gatagagatg	tctggaa	gcca	agaaccatgc	caaataatgtg	tctgtgactc	180
ctctgcgatg	acataatatg	tgacgatca	gaatttagact	gccccaa	ccca agaaaattcca	240
tttggagaat	gttgcgact	ttgc	ccacac	acttgc	ccctccat	300
ggtcaaggac	ctcaaggccc	caaggagat	ccaggccctc	tgggtattcc	ttggagaaat	360
ggtgaccctg	gtatcccgagg	acaaccagg	tcccctgtt	ctcc	tgccc	420
tgtgaatcat	gccc	tactgg	tcctcagaad	tattctcccc	agatgatc	480
aagtctggag	tagcgttgg	aggactcgc	ggttatectg	gaccagctgg	ccccccaggc	540
cctcccccgt	ccctcggtac	atctgg	cttgc	cttgc	ccggatctcc	600
ggaccccccgt	tgtaa	ccctgtgt	gc	aaagg	ttcc	660
ataggccat	ctgg	cttgc	ttggaa	atcg	gttagacccgg	720
gagcggaggat	tgcc	ttcc	ttccaggat	aaagg	ttcc	780
gttgcgttgc	ggat	ttcc	ttccaggat	ttccaggat	ttcc	840
cctggattaa	agggttgc	ttgtt	ttcc	ttcc	ttcc	900
ccaa	gggg	ctcc	cttgc	ttcc	ttcc	960
ggtaatgc	gtgt	cttgc	ttcc	ttcc	ttcc	1020
ggccggattcc	ctgg	ttcc	ttcc	ttcc	ttcc	1080
tcaa	atgtgt	ccccc	ttcc	ttcc	ttcc	1140
ggtctcttgc	gccc	ttcc	ttcc	ttcc	ttcc	1200
gctggcattc	tcgg	ttcc	ttcc	ttcc	ttcc	1260
gtaatgttgc	tc	ttcc	ttcc	ttcc	ttcc	1320
ggagagcccg	gacc	ttcc	ttcc	ttcc	ttcc	1380
aaaggcga	atggc	ttcc	ttcc	ttcc	ttcc	1440
gctgcaggag	aaagggttgc	ccct	ttcc	ttcc	ttcc	1500
ggagaaaa	gtcc	ttcc	ttcc	ttcc	ttcc	1560
gctgggaaac	tcgg	ttcc	ttcc	ttcc	ttcc	1620
agtccaggag	gacc	ttcc	ttcc	ttcc	ttcc	1680
ggtcgaccag	gtcc	ttcc	ttcc	ttcc	ttcc	1740
cccggttca	aaagg	ttcc	ttcc	ttcc	ttcc	1800
ggacttggcc	tc	ttcc	ttcc	ttcc	ttcc	1860
gggcctactg	ggc	ttcc	ttcc	ttcc	ttcc	1920
caaggcgtc	ctgg	ttcc	ttcc	ttcc	ttcc	1980
caaaagggttgc	atgg	ttcc	ttcc	ttcc	ttcc	2040
ggtgaacgttgc	gacc	ttcc	ttcc	ttcc	ttcc	2100
cctggtcccg	aaagg	ttcc	ttcc	ttcc	ttcc	2160
actcctggtc	tgca	ttcc	ttcc	ttcc	ttcc	2220
ggtgacaagg	gtg	ttcc	ttcc	ttcc	ttcc	2280
aggggttca	ctgg	ttcc	ttcc	ttcc	ttcc	2340
gaaagggttgc	ccc	ttcc	ttcc	ttcc	ttcc	2400
ggtgaaactgc	gccc	ttcc	ttcc	ttcc	ttcc	2460
cctggtgc	ttgg	ttcc	ttcc	ttcc	ttcc	2520
gttgacggat	cc	ttcc	ttcc	ttcc	ttcc	2580
ggtgacgttgc	g	ttcc	ttcc	ttcc	ttcc	2640
cctggtcttc	ttgg	ttcc	ttcc	ttcc	ttcc	2700
aaggatgggc	cccc	ttcc	ttcc	ttcc	ttcc	2760
ggacc	ttcc	ttcc	ttcc	ttcc	ttcc	2820
ccaggagatc	cagg	ttcc	ttcc	ttcc	ttcc	2880
ccaccaggc	tcgg	ttcc	ttcc	ttcc	ttcc	2940
ggaaaaccag	gag	ttcc	ttcc	ttcc	ttcc	3000
cctggtctgg	ttgg	ttcc	ttcc	ttcc	ttcc	3060
cttccaggccc	gagatggatc	ttcc	ttcc	ttcc	ttcc	3120
gggtcccttg	gtgt	ttcc	ttcc	ttcc	ttcc	3180
agtggtgac	gagg	ttcc	ttcc	ttcc	ttcc	3240
tc	ttcc	ttcc	ttcc	ttcc	ttcc	3300
ggagctgtgt	gcat	ttcc	ttcc	ttcc	ttcc	3360
ccaggccctg	ctgg	ttcc	ttcc	ttcc	ttcc	3420
cctgttggac	cc	ttcc	ttcc	ttcc	ttcc	3480
ggaccaccag	ggcc	ttcc	ttcc	ttcc	ttcc	3540

-continued

ccaggcgaac	cagggccctcc	tggacetcct	ggtgcccctg	gtccttgcgt	tgggttgtt	3600
ggagccgtc	ccatgtcg	gattggaggt	aaaaaagctg	gcgggtttgc	ccctatattat	3660
ggagatgaaac	caatggattt	caaaaatcaac	accgatgaga	ttatgacttc	actcaagtct	3720
gttaatggac	aatatagaaag	cctcattatgt	cctgatggt	ctcgtaaaaaa	ccccctgatag	3780
aactgcagag	acctgaaattt	ctgcccattct	gaactcaaga	gtggagaata	ctgggttgac	3840
cctaaccagg	gatgcaattt	ggatgtctatc	aaggatatttc	gtaatatgga	aactggggaa	3900
acatgcataa	gtgccaatcc	tttgaatgtt	ccacggaaac	actgggtggac	agattctatg	3960
gtctgagaaga	aacacgtttt	gtttggagag	tccatggat	gtggttttca	gttttagctac	4020
ggcaatctg	aacttcctga	agatgtcctt	gatgtgcage	ttggcattct	tcgacttctc	4080
tcctaggcgg	cttcaggagaa	catcacat	cactgcaaaa	atagcatgca	atacatggat	4140
caggccatgt	gaaatgtaaa	gaaggccctg	aagctgtatg	gttcaatgtg	agggtgaattc	4200
aaggctgaaag	gaaatagcaa	attaccctac	acagttctgg	aggatgttg	cacgaaacac	4260
actggggaaat	ggagcaaaaac	agtctttgaa	tatcgaacac	gcaaggctgt	gagactaacct	4320
tttgcataata	ttgcaccctt	tgacattgtt	ggtcctgatc	aagaatttgg	tgtggacgtt	4380
ggccctgttt	gtctttata	a				4401

SEQ ID NO: 6 moltype = DNA length = 4401
 FEATURE Location/Qualifiers
 misc_feature 1..4401
 note = Synthetic Construct
 source 1..4401
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 6

atgtatgtatgt	tcgtgcagaa	aggcagctgg	ctgctgtcg	cactgtcgca	ccctacaatc	60
attctggccc	agcaagaggc	cgtggaaaggc	ggatgttctc	acctgggaca	gagctacgccc	120
gacaggggat	tgtggaaaggc	tgagocctgc	catatcgcg	tgtgtatag	ccgcgcgcgt	180
ctgtgcacg	acatcatcg	ccgcgcacca	gagctggact	cccccaatcc	tgagatccat	240
tttcggcagt	gtctgtccgt	tttgcacccaa	cctccatccac	ctccatccac	acctctctat	300
ggacaggggac	ctcaggggccc	taaaggcgtat	cctggaccc	ctggaaatccc	ccggcagaata	360
ggcgcgcgcgg	gcattccatgg	acagcttgcgg	tctccatggaa	gttccaggttc	acctggcatc	420
tgcgcgcgcgt	gtctctacagg	ccctcragaa	tcacggccac	agttacgacag	ctacgcgcgt	480
aagtctggcg	ttggccgttgg	aggactggct	ggttatccat	gacctgttgg	accaccagggg	540
cctccaggac	cgcccttggaa	aagtggacat	ccaggatctc	ccggcgacttc	cggttatccat	600
ggaccacccgt	cgccatgttgg	acaggcaggaa	cctactgttgc	ctccctggttc	accagggtgcc	660
atttgaccat	ctggacatcg	ccggaaaggat	ggcgaatctg	gcagacccagg	cagacccatgg	720
gagagaggat	tgccttgtcc	tccgggtatt	aaggggccat	ccggcatttc	aggattcccc	780
ggaatgaagg	gccacacagg	cttcgtatgg	cggaatggcg	agaaggccga	aacagggtgcc	840
cctggactga	aaggcgaaaaa	tggactgtca	ggcgaccccg	acctatggga	900	
ccttagaggt	ctccaggcg	aaggaggaa	ccaggacttc	ctgggtgtgc	aggcgctaga	960
ggcaatgtatgt	gogccaggagg	atctgtatgg	caacctggcc	cgccaggacc	tccaggcaca	1020
gctggcttcc	ccggatctcc	tggtgcataa	ggcgaatgtt	gaccagctgg	aaggcccttgg	1080
tctaattgttgc	ccccaggacca	gaggggggaa	ccaggacccac	aaggacatgc	tggtgctcaat	1140
ggccctccgg	gttccatcg	gatcaatgtt	tccctcaggcg	gcaaaaggcg	gtatggccct	1200
gtctggatcc	ctggcgctcc	aggctttatg	ggagccaggat	ggccacccctt	accaggcaggc	1260
gcaaatggcg	ctctctggact	tagaggcgga	gcaggcgcgc	ctggaaaaaaa	tggcgcacaaa	1320
ggcgccgcgc	gacccagggg	agaaaaaggcg	gaatctgttg	ttccagggtt	cccttggtgcc	1380
aaaggccgg	aggccaaaggaa	tggatgttct	ggcgaggcc	ggccaaatgg	actccctggc	1440
gcacgttgc	aaaggggggc	acctggttt	agaggaccc	ccggacattaa	tggcattttcc	1500
ggcgaaaaaa	gttccaggcc	cgagctgtgt	gtccctggac	ctgcaggccc	aagagggtgt	1560
gtctggacac	caggcagaga	tgtgttccca	ggcgacccgg	ggatgtatgg	catcccgaggc	1620
tcaccctggcg	gaccgttgc	tgtatggaaa	cccgccgtt	ctggtagcc	gggcaatcc	1680
gttagacccgg	gacccctgggg	accatcgaa	ccaagaggac	aacctggccgt	gtatggccct	1740
cctggaccta	aggggaaatgt	cggtgttcc	ggaaagaac	gcgaacgcgg	tggcccttggc	1800
gttcccggtc	ctcaaggggc	accaggccaa	aacggcgaaa	ccgttccaca	aggaccacat	1860
ggacactacag	gaccgttgggg	agataaggcc	gatacaggatc	caccaggacc	tcaagggtact	1920
caaggactgc	ctggacactgg	cggtgttcc	ggcgagaatg	aaaaacccagg	cgaacccagg	1980
cctaaggggc	acgttgcgtc	acctggccca	ccaggccgaa	agggcgacgc	aggcgcttca	2040
ggcgagagg	gaccgttccgg	attggctgtt	gttccaggct	ttagaggccgg	agctgtgtcc	2100
ccaggacactcg	aaagggtggaa	agggtgttca	ggacacttctg	ggccacccctt	cgctgttgg	2160
actccaggac	ttcaatggat	ggcttggccaa	cgagggttgc	ttggaaaggccc	aggacaaaaa	2220
ggcgataagg	gogaacactgg	cggttgggg	cgatgttgc	ttccctggaaa	agatggacca	2280
cgggggccca	caggacatctat	aggcccttca	ggggccagca	gacagccggg	cgacaaaggc	2340
gaagggtggc	ccccctggctt	gtctggatca	ttagggatcc	ggatggccgg	agctgtgtcc	2400
ggagaaacac	gcctcttgg	accgggggaa	tttccctgggt	ctccctggca	aaatggcgag	2460
cctggcgcaa	aaggcgaaaa	agggtgcac	ggcgaaaaaa	gcgaaggccgg	accttcttgg	2520
gttgcgttgc	ctccctggcg	atctggatca	gttggccctt	ctggtcttca	aggtgttaag	2580
ggcgaaaag	gtctccagg	cggtgttcc	gtgtgttgc	ttccctggc	aaggaggattt	2640
ccccggaccac	caggcttcaa	tggcaatcca	ggtccttc	gacctagccg	ctcttc	2700
aaagatggcc	caccaggacc	agccggaaaat	actgggtgttc	ctggatcc	tggcgtgtcc	2760
ggaccggaaa	gogacggccgg	acaaccaggc	aaaaaaggat	ctccctggcc	tcaaggccct	2820
cctggcgac	ctggcgatcc	gggttggcc	ggttattcc	gttggccat	ccctggcttgc	2880
ccacctggaa	tgccttggcc	aagggttgc	ccgggaccc	aaggcgatc	aggcgatct	2940
ggaaaggccct	gcccggccacgg	actggatcc	ggaaaggaggac	ctccagggttcc	acaaggcttgc	3000
cctggatttt	ctggaaacacgc	tggcgaacat	gaaagagatg	gcaatcttgc	ctctgtatgg	3060
ttgcgggggaa	gagatggctc	cccaagggttgc	aaggggcgatc	ggcgagaaaa	tggtagccca	3120

-continued

ggcgctcccg	gctccaggc	acaccaggc	ccacccgtgc	cagtcggacc	tgcgtggaaag	3180
tctggcgata	gagggcgatc	tggaccgc	ggtcccgctg	gcgccccagg	gcctgcccga	3240
tctaggggag	ccccctggacc	gcaaggact	aggggagaca	aaggcgagac	tggcgaaacgc	3300
ggagccgtc	gaatcaaagg	ccatagagg	ttccccaggca	accctggtgc	accggatca	3360
ccaggacccg	caggacaaca	aggcgctatt	ggcagtcggc	ggctgtctgg	gccgagagga	3420
ccagttggac	ctagtggacc	accggggcaaa	gatggaaaca	gcccggacccc	tggactatc	3480
ggaccaccag	gaccttagagg	caataggg	gagagaggac	gcgggggatc	tcccgagacat	3540
cctggacaac	ccgggttcacc	ggggccacca	ggccggccagg	ggccatgttg	tggcgaggt	3600
ggagctgtc	ccattgtctgg	aatcgggga	gagaaagccg	gcggatttgc	cccttattac	3660
ggcgacgacg	ccatgttatt	caatgtca	accggacggaa	tcatgaccag	cctgaatgtcc	3720
gtgaaacggcc	agatcgagag	cctgtacggc	cggacggggc	gcagaaaaga	ccccggccaga	3780
aactgcgcgc	acctgttgtt	ctgttacccc	gagctgtaaa	ggggcgagta	ctgggttgc	3840
cccaaccagg	gctgttaact	ggacgcatt	aagggttct	gcaatgtga	aaccggcgag	3900
acatgcatca	ggcgcaatcc	tctgttgtc	ccccggaaac	actgggtggc	agatagcagc	3960
ggcgagaaga	agcacgttgc	gttccggag	agcatggacg	gggggttcca	gttcttctac	4020
ggcaatcccg	agctgtccgc	ggacgtgtgt	gtatgtca	ttggcccttct	gagactgtcg	4080
ageagccgcg	ccagccagaa	tatcacctac	caactgtca	acagcattgc	ctatgtggat	4140
caggccacgc	gcaacgttgc	gaaagccctg	aagctgttgt	gcacgttgc	ggggcgagtt	4200
aaggccgggg	ggcacatcc	gttccatctt	accgtgtctgg	aagatggctc	caccaagcac	4260
acagggcgat	gttccaaagac	cgttgtc	tacccggacaa	gaaaggccgt	gggggtgcct	4320
atctgtggata	tcggccctta	cgatatcgga	ggccccgtatc	aagagtccg	cggtgacgtg	4380
ggccctgtgt	gtttctgtat	a				4401

SEQ ID NO: 7 moltype = DNA length = 5010
 FEATURE Location/Qualifiers
 source 1..5010
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 7

atggggcccc	ggctcagcgt	ctggctgtcg	ctgtgtccgg	ccggcccttct	gtccacggag	60
gagcacaaggc	ggggcgctgc	gaagggtggc	tgtgttgtgc	ctggctgtgg	caaagtgtac	120
tgccatggag	tgaaggggaca	aaagggtgaa	agaggccctc	cgggggtaca	agggtgttac	180
gggtttctgt	aatgtcaagg	acctgggggg	ccacaggggc	caccaggaca	aaagggtgt	240
actggagaac	caggactacc	tggaaacaaa	gggacaagag	gacccctggg	agcatctggc	300
taccctggaa	accggactgt	tccgggaa	cctggccaa	acggcccccgg	aggcccccca	360
ggtattcccg	gatgtcaatgg	cacaaggggg	gagagggggc	cgctcgcccc	tcctggcttg	420
cctgggttccg	ctggaaatcc	cgggaccaca	ggcttacca	ggatgtgggg	tgatccagg	480
gagataacttgc	ccatgtgtcc	cgggatgtcg	ttgaaagggt	aaagaggatt	tcccggaatc	540
ccaggccgtc	caggccccc	aggactgtcc	gggggttca	gttctgttgg	gcctccagg	600
tttaccggac	caccagggtc	cccaggccct	ccggccccc	cagggtaaaa	gggacaatgt	660
ggcttaaagg	ttaaagggtc	aaaagggtac	aagggttacc	aagggttca	tgggtcttca	720
ggagtaccag	gacaagctca	agttcaagaa	aaaggagact	tcgcccacaa	ggggagaaaag	780
ggccaaaaaa	gttgcacgttgc	atttcagggg	atggccagggg	tcggagagaa	agggtgtaccc	840
ggaaaacccag	gaccggagg	caaaccggg	aaagatgggt	acaaaagggg	aaaaggggat	900
cccggttttc	ctgggtgtacc	cggttacca	ggactcatag	ggccggagg	cccgccagg	960
gaaaagggttgc	aaggcagggtc	tcctggccca	cctggaaattt	ttataggcac	aggaccttgc	1020
ggagaaaaaa	gagagggggc	cttccctgg	actccggggc	caaggaggaa	ggccaggccca	1080
aaagggtttcc	caggactacc	aggccaaacc	ggaccccttgc	gcctccctgt	acctggggc	1140
gttgggttcc	ctgggttcc	tgggtaaa	ggagaaaaaa	gtgaccagg	atttctgtgt	1200
acatctctgc	caggaccaag	tggaaagat	gggctccccc	gttctcttgg	ttcccttgg	1260
ccccctggcc	acccgttgc	cacaatgttgc	atgggtttgc	gtcagccccc	acctccagg	1320
gaccagggtc	ctccgttgc	tccagggtat	tcaggggc	ccaggattt	taggcaaaa	1380
ggtcggaaagg	gagaggttgc	cctcatctgt	gatatacgtc	gatatacg	ttccgggg	1440
ccacaggcgc	ccccgggaga	aatagggttc	ccaggccgc	caggggccaa	ggggcgacaga	1500
gttggcttgc	gttgcgttgc	tggtcagg	gttgcaggcc	ctcaagggtac	accagggtcg	1560
ataggccatgc	caggggccaa	ggggggggct	gggtgtgttt	atttcgactt	gggggttcaaa	1620
gggtgacaaagg	gagaccgggg	ctttccggaa	caaggccgg	tcggcagggg	agccgggttct	1680
cctggaaagg	atggccatcc	gggttttctt	ggccccaagg	gttgcgggg	ttctgttag	1740
tgtggaaagg	acgcgtggcc	ccctggaggaa	gttgcgttgc	caggcgttgc	tggtgacacc	1800
ggcccccctg	ggcccccctg	atatggcttgc	gtgtgttcc	ttggtgacaa	aggacaaacg	1860
gggtttctgt	ggggccctgg	atccggggcc	caaggccgg	caagggttgc	accaggaaaa	1920
attgttcc	taccaggccc	ccctggggca	gaaggactgc	cggggttccc	aggcttcca	1980
gggtccccaa	gagaccgggg	ctttccggaa	accccaggaa	ggccaggcct	gccaggagag	2040
aaagggtcg	ttggggccagg	aggccatgttgc	ttttccagg	ccccggccc	caaagggtgt	2100
gacggcgttac	ctggggatcc	ggggccaccc	ggggactcc	gttgcgggg	atttaatggc	2160
ttaacctggga	acccggatgt	gcaggggcc	aaaggggagg	tcggagtttgc	tctaccgg	2220
ctcaaaagggtt	tgccagggtt	tccggcatt	cctggcacac	ccggggagaa	ggggggat	2280
gggggttcc	gggttccctgg	agaacatgttgc	ggcgttccgt	ttcagggttgc	caaggagaa	2340
agaggtgtac	cgggggactcc	tgggttgtcc	gggttcccg	ttttccagg	ttttccagg	2400
ataggcccccc	ctggggatgt	gggttcccttgc	ggggactcc	gttgcgggg	ttttccagg	2460
cctccctggaa	taaaagggttgc	ggggggcc	ggggggcc	ggggggcc	ttttccagg	2520
cctaaagggtt	ataaaagggttgc	tcaaggactc	cctggcataaa	cgggacatgc	gggggttccc	2580
ggcccccctg	gacggcgttgc	gggttccctgg	atttccgttgc	ttccagggttgc	caaggagaa	2640
atggggcgatc	tggggacccc	cgggcgttgc	gggttcccttgc	ggggggcc	ttttccagg	2700
ttaccgggttgc	aaaaagggttgc	ccatgggttcc	ccggggcc	ggggggcc	ttttccagg	2760
gggttgcgttgc	gtgtgtatgg	gggttgcgttgc	ttccgttgc	ttccagggttgc	caaggagaa	2820

-continued

gtggacatgg	gcagcatgaa	gggccagaaa	ggagaccaag	gagagaaaagg	acaaaattgga	2880
caatttggt	agaagggtc	ccgaggagac	cctgggacc	caggagtgc	tggaaaggac	2940
ggcaggcag	gacagcctgg	gcagccagga	cctaaagggt	atccaggat	aagtggaaacc	3000
ccaggtgtc	caggacttc	gggacaaaaa	ggatctttg	gtggaatggg	cttgccagga	3060
acacctggg	agaaaagggt	gcctggcata	cctggccac	aaggttcac	tggcttacct	3120
ggagacaaag	gtgcaaaaagg	agagaaaaggg	caggcaggcc	cacctggcat	aggcatccca	3180
gggtctcgag	gtgaaaagg	agatcaagg	atagcgggt	teccaggaa	ccctggagag	3240
aaggagaaa	aaggaaagcat	tggatccca	ggaatgcag	gttccccagg	ccttaaaggg	3300
tctccggga	gtgtggcta	tccaggaat	cctgggtac	ctggagaaaa	aggtgacaaa	3360
ggcctcccg	gatggatgg	catecctgt	gtcaaaggag	aaggaggct	tcctggact	3420
cetggcccca	caggcccag	tggccagaaa	ggggagccag	gcagtgtat	aatccgggg	3480
tcagcaggag	agaagggtg	accaggctca	ccaggaagag	gattccagg	gtttccagg	3540
gccaaaggag	acaaaagggt	aaagggtgag	gtgggttcc	caggattag	cggggccca	3600
gaaatctct	gatccaaagg	agagcaagg	tccatgggt	ctccggggcc	ccagggacag	3660
ccggggttac	cgggatccc	aggccatggc	acggaggggc	ceaaaggaga	ccggggact	3720
caaggccagc	ctggcttgc	aggacttcc	ggacccatgg	ggcccttcc	gtttcttgg	3780
attgtggag	ttaaagggt	caaaggaaat	ccaggctgg	caggagcacc	cgggtgtccca	3840
gggccccaaagg	gagacccctgg	atccaggat	atgcctgtt	ttgggtggct	tccaggatc	3900
acaggctcta	agggtgat	ggggccttca	ggagttccag	gatttcaagg	tccaaaagg	3960
cttcctggcc	tccaggat	taaaagggt	caaggcgttca	aaggcgtccc	gggagctaaa	4020
ggtctccogg	gtccctctgg	ccccccagg	ccttacgaca	tcatcaaagg	ggagccccgg	4080
ctccctggc	ctgaggggcc	ccggggctgt	aaagggttcc	agggactg	aggccggaaa	4140
ggccagcag	gtgttacagg	attgggtgg	ataccctgg	ctccaggat	tcctgggtt	4200
gacgggtccc	ctggccagaa	aggagagatg	ggacctgtcc	ggctacttgg	tccaaagagga	4260
tttccaggtc	eaccaggccc	cgatgggtt	ccaggatcca	tggggcccc	aggcacccca	4320
tctgttgc	acgggttct	tgtgaccagg	catatcaaa	caatagatg	cccacagt	4380
ccttcgggaa	ccaaaattct	ttaccacggg	tactcttgc	tatcgtca	aggcaatgaa	4440
cggggcccat	ggcaggactt	gggcacggcc	ggcagctgc	tgccaaatgt	cagcaatgt	4500
ccttcctgt	tctgaat	taacaacgt	tgcacat	catcaacaaa	tgactactg	4560
taactgtgt	ccaaaaatgt	gccccatgg	atgtcaatg	cacccatcac	ggggggaaac	4620
ataagacat	tatttagt	gtgtgtgt	tgtgaggcgc	ctggcatgtt	gtggccgtg	4680
cacagccaga	ccatccat	ccccatgt	ccacgggggt	gttccctgt	gtggatccgc	4740
tactctttt	tgatgcac	cagggctgt	gcagaaggct	ctggccaag	cctggcttgc	4800
cccggtct	gcctggagga	gtttagaagt	gcccattca	tgcagtgt	ccggccgtgg	4860
acctgcattt	actacgcaaa	cgcttacago	ttttggctc	ccaccataga	gaggagcgc	4920
atgttcaaga	agcctacggc	gtccaccc	aaggcagggg	agtcgtgcac	gcacgtcag	4980
cgtgtccaa	tctgtatgag	aagaacataa				5010

SEQ ID NO: 8	moltype = DNA	length = 5010				
FEATURE	Location/Qualifiers					
misc_feature	1..5010					
	note = Synthetic Construct					
source	1..5010					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 8						
atggcccta	gactgtctgt	gtggctgt	tttcttcctg	ccgctctgt	gctgcacag	60
gaacatttca	gagcccccgc	taaaggcgg	tgtccggat	ctggatgg	caagtgtat	120
tgtcacggc	tgaaggggca	gaaggcgaa	agaggacttc	ctggactgca	gggcgtgtac	180
ggatttctg	gaatgcaagg	acctgaggga	ccacaggac	ctccaggaca	aaaaggcgat	240
acaggcgaac	ctggcttgc	tggcacaagg	ggacaacagg	gaccacctgg	ccgcctctgg	300
tatcctggaa	atccaggact	gcctggaaatc	ccaggacagg	atggacttcc	tggctcttct	360
ggcattctgt	gtgttaatgg	caccaagg	qagaggggtc	cactcgacc	accaggact	420
ccaggatttgc	ctggaaaccc	tggccac	ggattgtctgt	gcatgaagg	cgatccttgc	480
gaaaatcttgc	gacacgttcc	cgccatgt	ctgaaaggcg	aaaggggtt	tcccgaaat	540
cctggcacac	ctggttccac	aggettgc	ggacttcaag	gaccgttgg	cccaccttgg	600
tttacagggtc	ctccaggatc	acccgggtt	ccaggccgc	caggcgaaaa	aggacaatg	660
ggactgtat	ttcaggggcc	caaggcgac	aaaggcgatc	aaagggtttag	tggcccttca	720
ggtgttctgt	gacaggccca	ggttaaagg	aaaggcgat	tgcacccaa	gggcgagaag	780
ggacagaaagg	gcaacccagg	cttccagg	atgccttgg	tggggaaaa	gggcgaaaccc	840
ggaaaacccgt	gacctagg	caaaaaacgg	aaaggatggc	ataaggcg	gaaaggat	900
cctggattcc	ctggcgaaacc	tggatatect	ggactgtat	gtagacagg	ccctcaagg	960
gagaaggcg	aaagctggacc	gcctggacca	ccggaaatag	tgatttgg	aggcccttc	1020
ggagaaaaaa	gggggggggg	atccctgg	actccctgg	ctaggggaga	accaggac	1080
aaagggttcc	ccggacttgc	aggacaaac	gggcctctgt	gggttctgt	gcctggacaa	1140
gtctgggtc	caggatttcc	cgccgaaat	ggcggaaaagg	gctgatagg	ttttcttgg	1200
acaaggctgc	ctggcccttc	tggagagat	ggacttccgg	ggccacccgg	tttcccccgg	1260
ccacctggac	agcctggct	taccaatgg	atcgtcgat	gtcaacccgg	tccacccgg	1320
gatcaaggac	ctccctggat	accaggccag	cttggcttta	tcggcgat	tggagaaaa	1380
gggcggaaaagg	gggggggggg	cttccctgg	cttccctgg	gatcagg	cccttcttgg	1440
cctcaaggcc	caccaggcga	gatagttt	ccaggccac	tcggccaa	aggcgaccc	1500
ggactccctg	gttagagatgg	tgttgcgt	gttccaggac	cacaagg	ccctggact	1560
attggacacg	cagggtcaaa	gggcgagcc	ggcgagt	tcttcgac	gagactt	1620
ggcgacaaagg	gcaacccagg	cttccctgg	caacctggaa	tgcctgg	agctggat	1680
ccaggccgt	atggacatcc	agggttgc	ggacaaaaaa	gttcccaagg	atctgtgg	1740
ctcaaaggcg	aaagggcccc	tccaggcg	gttggattt	ctggcttgc	aggcgata	1800

-continued

ggccaccac	gtccacactgg	atatggacact	gctggcccta	ttggagataa	gggccaagca	1860
ggatcccc	goggaccggc	ttcccgaggc	c当地ggggtc	ctaaaggcga	gccttgaaaa	1920
atcgttcac	tgcctggacc	tcaggcgctg	gaaggattgc	ctggatctcc	cggttttcc	1980
ggaccgc	caag	gogatagagg	atccccgg	acaccggta	gaccaggcct	2040
aaagggtct	tggtcaacc	tggatcg	ttccctgg	tcctgg	aaaagggttt	2100
gatggactgc	cgccgat	gggcccac	ggaacacc	gcagaccc	ctttaatgg	2160
ttgc	ccggaa	atcccg	ccaaaggcc	aaaggcg	ccgttgtt	2220
cttaaaggac	ttccaggc	gcccgcata	cctgg	ctgg	aaaaggatctat	2280
ggagtctg	cgctgc	cgaacatgt	gcaat	tcac	ctgggtt	2340
agaggc	ac	ccggcctc	aggact	ggtctgtt	gaatgc	2400
atgggac	ccgg	ccgg	ccgg	ccgg	ccgtccgg	2460
atgggac	ccgg	ccgg	ccgg	ccgg	ccgtccgg	2520
cacactgg	tcaaa	ggcga	gaaaggctc	ccccgc	ccggcctt	2580
cctaaaggc	atagg	gtgc	ccaggc	cctgg	ctggaca	2640
gggc	ccgg	gaca	ccgg	ttcc	ccggat	2700
atggg	ccgg	ccgg	ccgg	ccgg	taagg	2760
ttccccc	gg	aaaaggc	ccac	ccac	ggatc	2820
ggattg	aaag	gcgata	cgac	cttcc	ggc	2880
gtgg	cat	gttca	gggca	ggc	ccagat	2940
cctatcg	cc	aaaagg	ccgg	ccgg	ccggat	3000
ggaca	ccgg	ccgg	ccgg	ccgg	ccggat	3060
ccagg	gtc	ca	ccgg	ccgg	ccggat	3120
acacc	ccgg	ccgg	ccgg	ccgg	ccggat	3180
ggog	gacaa	ag	ccgg	ccgg	ccggat	3240
ggactt	agg	gaga	ccgg	ccgg	ccggat	3300
aaaggc	aaa	aggct	ccgg	ccgg	ccggat	3360
tcac	ctgg	ccgtgg	tc	ccgg	ccggat	3420
gact	gtct	gttctgtt	cata	ccgg	ccggat	3480
cctgg	gac	ccgg	ccgg	ccgg	ccggat	3540
tctg	ctgg	aaaaaggc	ggcc	ccgg	ccggat	3600
gcaa	agg	atagg	ccgg	ccgg	ccggat	3660
ggcat	cccg	ccgg	ccgg	ccgg	ccggat	3720
ccagg	gtt	ccgg	ccgg	ccgg	ccggat	3780
ca	ggg	ccgg	ccgg	ccgg	ccggat	3840
attgat	ggc	caagg	ccgg	ccgg	ccggat	3900
gttcc	aaa	aggcc	ccgg	ccgg	ccggat	3960
actgg	aa	ggc	ccgg	ccgg	ccggat	4020
ttcc	cagg	ccgg	ccgg	ccgg	ccggat	4080
ggatt	gtct	ggcc	ccgg	ccgg	ccggat	4140
ctgc	ctgg	ccgg	ccgg	ccgg	ccggat	4200
ggac	agg	ccgg	ccgg	ccgg	ccggat	4260
ttccc	ccgg	ccgg	ccgg	ccgg	ccggat	4320
cccg	gtt	ccgg	ccgg	ccgg	ccggat	4380
cctag	ccgg	ccgg	ccgg	ccgg	ccggat	4440
agag	ccac	ccgg	ccgg	ccgg	ccggat	4500
ccttt	ctgt	ccac	ccac	ccac	ccggat	4560
tact	ggct	ccat	ccat	ccat	ccggat	4620
atc	ccat	ccat	ccat	ccat	ccggat	4680
cact	ccc	ccat	ccat	ccat	ccggat	4740
tact	cttt	tgat	tgat	tgat	ccggat	4800
cccg	gtct	ccat	ccat	ccat	ccggat	4860
acct	gtta	acta	acta	acta	ccggat	4920
atgtt	caag	agcc	ccat	ccat	ccggat	4980
agat	gtca	ccat	ccat	ccat	ccggat	5010

SEQ ID NO: 9 moltype = DNA length = 3087
 FEATURE
 source Location/Qualifiers
 1..3087

mol_type = unassigned DNA

organism = Homo sapiens

SEQUENCE: 9

atgagggcgg	cccg	gtct	ccgg	ccgg	ccgg	60
gatgagccgg	ccgg	ccgg	ccgg	ccgg	ccgg	120
gtctggaca	cct	ctgg	ccgg	ccgg	ccgg	180
aaagtcaagt	cctt	caccaa	ccgg	ccgg	ccgg	240
gaccggaaac	tgg	tgg	ccgg	ccgg	ccgg	300
caaggcctca	cg	cg	ccgg	ccgg	ccgg	360
aagtacttgc	gg	gg	ccgg	ccgg	ccgg	420
ctcg	gggg	gggg	ccgg	ccgg	ccgg	480
ccct	gggg	gggg	ccgg	ccgg	ccgg	540
cac	ctgg	ctgg	ccgg	ccgg	ccgg	600
agatcatcg	ccac	ccac	ccgg	ccgg	ccgg	660
agccgcgac	ccac	ccac	ccgg	ccgg	ccgg	720
aataacgtgg	ccac	ccac	ccgg	ccgg	ccgg	780
ctccggggcc	cc	cc	ccgg	ccgg	ccgg	840
ggagaagccg	ccat	ccat	ccgg	ccgg	ccgg	900

-continued

aaggggagaaa	aaggggagccg	tggggagaag	ggctccaggg	gacccaaggg	ctacaaggga	960
gagaaggggca	agcgtggcat	cgcgccggtg	gacggcgtga	agggggagat	gggttaccca	1020
ggcgtggccag	gctgcagggg	ctcgccccgg	tttgacggca	ttcaaggacc	ccctggcccc	1080
aaggggagacc	cgggtgcctt	tggactgaaa	ggagaaaaagg	gcgagcttgg	agctgacggg	1140
gaggcggggaa	gaccaggggag	ctcgggacca	tctggagacg	agggccagcc	gggagagcct	1200
ggggccccccgg	gagagaaaagg	agaggccggg	gacgaggggg	acccaggacc	tgacgggtgcc	1260
ccccggggagc	gggggtggccc	tggagagaga	ggaccacggg	ggaccggagg	cacgogggg	1320
ccaaaggagg	accctgggtg	agctggccgg	cagggtgatc	agggaaagaga	aggccccgtt	1380
ggtgtccctg	gagaccgggg	cgaggctgg	cctatcgac	ctaaaggcta	ccgaggcgat	1440
gagggtcccccc	cagggtccgg	gggttccca	ggagccccca	gacctgccc	acccccttgg	1500
gaccggggggc	tgtatgggtg	aaggggagaa	gacggggccc	ctggaaatgg	caccgagggc	1560
ttccccggct	tcccccggta	tccgggcaac	agggggcgctc	ccgggataaa	ccggacgaaag	1620
ggctaccctg	gcctcaaggg	ggacgaggg	gaagccgggg	accccgaga	cgataaacac	1680
gacattgcac	cccgaggagt	caaaggacg	aaagggttacc	gggggtccca	gggccccccag	1740
ggaccccccgg	gacaccaagg	accgcctgg	ccggacgaaat	gcegagattt	ggacatcattc	1800
ataaaaatgt	gtcttgcgt	tgaatgcag	tgccggccca	tcgacccctt	gttctgtctg	1860
gacagctcg	agagcattgg	cctgcagaac	ttcgagattt	ccaaggactt	cgttgtcaag	1920
gtcategacc	ggctgagccg	ggacgaggt	gtcaagttcg	agccaggcc	gtctgtacgg	1980
ggtgtgggtc	agtacagccca	cagccatcg	caggacacg	tgagctgtcg	cagccccagc	2040
atccggaaacg	tgcaggagct	caagggaccc	atcaagggcc	tcgactgtat	ggccggccgg	2100
accttcacgg	gggaggccct	gcagttacacg	cgggaccagg	tgctggggcc	cagccggaaac	2160
aaccgcacatcg	ccctgttcat	caactgacggg	cgctcagaca	ctcagaggaa	caccacacccg	2220
ctcaacatgc	tctcgagccc	cggeatccg	tggtgtcccg	ttggcatcaa	agacgtttt	2280
gacttcatcc	caggctcaga	cacgtcaat	gtcattttt	gccaaggctt	ggcaccatcc	2340
caggggccgg	ccggctctc	gctggtcaag	gagaactatg	cagactgtgt	ggaggatgcc	2400
ttctctaaaga	atgttacccgc	ccagatctg	atagacaaga	agtgttccaga	ttacacctgc	2460
cccatcaacgt	tcttcctcccc	ggctgttccat	accatcttc	tggacggctc	cgccagcgctg	2520
ggcagccaca	attttgacac	caccaagcgc	ttcgccaaacg	gcctggccga	gcgttctcc	2580
acagcggggca	ggacggaccc	cgccccacgc	gtgcgggtgg	cggtgttgc	gtacagccgc	2640
acggggccagc	agcgcggcaga	gccccggctcg	ctgcagttcc	tcgcagaaacta	cacggccctg	2700
gccagtgccg	tccatgttccat	ggacttatac	aacgacgcaca	ccgacgtcaa	cgatggccctg	2760
ggttatgttgc	cccggttctt	cccgaggccc	tggtccggcc	ctggcaagaa	gaggctgtcg	2820
ctcttcctcg	atggcaactc	gcaggggcc	acgcccgttgc	ccatcgagaa	ggccgtgcag	2880
gaagccccccgg	gggcaggcat	cgagatctt	tggtgtggcc	ttggccgc	ggtaatgg	2940
ccccacatcc	gggttccctgg	caccgcacgg	acggccggat	acgacgtggc	ctacggccgg	3000
agccacacgtt	tccgtgtccc	cagctaccgg	ggccgttcc	gggtgttcc	ccaccagaca	3060
gttcccaaggg	aggtggcgct	gggttag				3087

SEQ ID NO: 10	moltype = DNA	length = 3087			
FEATURE	Location/Qualifiers				
misc_feature	1..3087				
	note = Synthetic Construct				
source	1..3087				
	mol_type = other DNA				
	organism = synthetic construct				
SEQUENCE: 10					
atgagagccg ctagagccct	tctgcctctg	ctgtgtcaag	cctgttggac	agccgcttag	60
atgtggccgg aaacacctag	agccgtggca	tttcaggact	gccccgtgga	tctgttctt	120
gtgtgtggata ccagcgagag	cgtggccctg	agactgaac	cttgcgcgc	cctgtggac	180
aaagtgaagt ctttcaccaa	cggttcatc	gacaacctgc	gacgacggta	ctacagatgc	240
gacagaaaaac ttgtgtggaa	cgcggccgc	ctgcactact	ctgacgaggt	ggaaatcatc	300
caggggctga ccagaatggc	tgccggaaaga	gatgcctgta	agttctagct	ggacggccgt	360
aagtactttt gcaaggggac	cttacaccgc	tgcccatca	aaagaaggctt	ggaacagctg	420
ctcgtcgccg cgagccatct	gaaagagaac	aagtacctga	tcgtgtttac	cgacggacac	480
cctctggaaat gtatcaaaa	acccctgttgc	ggacttggaa	atgcctgtaa	cgaggccaaa	540
caccttggccg tgaagggttt	cagegtggcc	atcacaccccg	accacctgaa	acctcggtcg	600
agcatcatcg ccacccgacca	caccatccgg	cggttattca	cagccgttga	ttggggccag	660
agcagggatg ccgaagaggc	catcagccag	accatcgaca	ccatctgtga	catgtcaag	720
aacaacgtgg aacaatgtt	ctgcagtc	gatgtccggc	ctgtctagg	acccctcttgg	780
cttagaggccg accctggctt	tgaggccgg	agagggaaaac	ctggacttgc	ccggcgaaaa	840
ggcgtggatgtt gogatcttgg	tagacccgtt	gatctggggac	ctgtgggtca	tcagggatgt	900
aaggggcgaga aaggctcccg	ggggagagaag	ggatcttagag	gcccataagg	ctacaagggc	960
aaaaaaaggca agagggggat	cgatggcggt	gacggcgta	aaaggcgagat	gggatatactt	1020
ggactccctgtgttggaaa	cggccctgttgc	ttcgatgtggaa	tccaggggacc	tccaggacc	1080
aaggggcgatc caggccgcctt	tggactgaaa	ggcgaaaaagg	gcaacccagg	tgccgtatggc	1140
gaaggcggca gacctgttgc	ttctggccctt	acggcgatgt	aaggacaggcc	ttggcgaaacct	1200
ggaccacactg gcgagaagg	cgaacccggc	gacgaggggaa	atccaggacc	agatgggtct	1260
cctggcgaaa gaggttggacc	aggcgaaagg	ggaccttagag	gaacaccagg	cacaaggaggc	1320
ccaaaggccgc atccggccg	ggctggacatc	caaggcgatc	agggaagaga	aggaccatgt	1380
ggagtccctggcgcacc	cgacggccagg	cgtatccggc	ctaaaggata	tagaggcgac	1440
gaaggcccttc ctggatctg	agggtctaga	ggcgcacccg	gtccagccgg	ccctccaggc	1500
gaccggccgc ttatgggaga	acggggagaa	gatggccctg	ccggcaatgg	cacagagggc	1560
tttccaggctt tcctggctt	ccccggaaat	agaggcgctc	ctggatcaa	ccggccacaa	1620
ggatataccatggcgttcaagg	cgacgaaaggc	gaggcaggcc	atccaggggaa	tgacaacaac	1680
gatatacgcccttccatggcgttcaagg	gaaaggcgcc	aaaggctata	gaggaccaga	gggaccacaa	1740
ggcccaacccgttgcataagg	gtcatcaagg	ggccaccagg	cctgcacgat	ggacatcatt	1800

-continued

atgaagatgt	gcagctgctg	cgagtgcaga	tgcggcccta	tcatgtgtc	gtttgtgtg	1860
gacagctccg	agagcatcg	cctcgaga	ttcgagatcg	ccaaggactt	cgtggtaaaa	1920
gtgatcgaca	gactgagcc	ggacgagctg	gtcaagtttg	agectggcca	gtcttatgcc	1980
ggcgtggc	agtagaccca	cagccatgt	caagagcacg	tgtccctgag	aagccccagc	2040
atcagaaac	tcgaagagct	gaaagaagcc	atcaagtccc	tgcagtggat	ggctggccga	2100
acctttaactg	gogaggccct	gcagatcac	agagatcaac	tgctgecc	ttctcttaac	2160
aaccggattg	ccctcgtat	caccgacgg	agaagcgaca	cccagagaga	caccacac	2220
ctgaacgtgc	tgtcagcccc	cggattcag	gtggtgtctg	tgggcatcaa	ggacgtgtt	2280
gacttcatcc	ccggcagcga	ccagctgaac	gtatcttctt	gtcaaggact	ggcccttagc	2340
caaggcagac	caggactgtc	tctgcataaa	gagaactacg	cgagactgt	cgaggacgcc	2400
ttcctgaaga	atgtgacagc	ccagatctgc	atgcacaaga	agtgc	cccgatcacatgc	2460
cccatcacct	tttagcagccc	tgcgcacatc	accatctgc	ttgatggc	tcgtacgtgt	2520
ggcagccaca	acttcgacac	caccaagaga	ttcgccaage	ggctggccga	gagatttctg	2580
acagccggca	gaacccatgt	tgctcagat	gtgcgagtgg	ccgtgttcca	gtatcttgc	2640
acaggcagc	aaagacccga	gagacgtctt	ctgcagt	tgcagaacta	cactgcctg	2700
gcctctgc	ttggatgcata	ggactttata	aacgcaccca	ccgacgtgaa	cgacgcctg	2760
ggctatgtga	cccggtttt	cagagaagcc	tctagcggag	cegccaagaa	gagactgtg	2820
ctgttcacg	acggcaactc	ccaaatgtgt	acaccagccg	ccatttggaa	ggccgtgca	2880
gaagctcaga	gagccggcat	cgagatctt	gtgggtgtcg	tgggcagaca	agtgaacgag	2940
cctcacatca	gagtgctgtt	cacccgcaag	accgcggat	acgtatgtc	ttatggcgag	3000
agccacatgt	tcaagatgtc	cagetatcg	gctctgtga	gaggcgtt	ccaccagacc	3060
gtgtctagaa	agggtggccct	gggatga				3087

SEQ ID NO: 11 moltype = DNA length = 8835
 FEATURE Location/Qualifiers
 source 1..8835
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 11

atgacgctgc	ggcttcttgtt	ggccgcgc	tgcgcgggaa	ttctggcaga	ggcgccccga	60
gtgcgagccc	agcacacggg	gagagtgacc	tgcacgcggc	tttacgcgc	tgacatttgt	120
ttttaactgg	atggcttc	atccattggc	cgacaa	tccgcgagg	ccgcagctt	180
ctcgaaggcc	tttgtgtgc	tttctctgaa	gcagccatg	cacagggtgt	gcfgttgc	240
acagtgcagt	acagcgat	cccacggaca	gagttcgcc	tggatgcact	tggctctgg	300
gttgatgtga	tccgcggat	ccgtgagat	agttacaagg	ggggcaaac	tgcacacagg	360
gctgcaattc	tccatgtggc	tgaccatgtc	ttctgtcccc	agctggcc	acctgggtgc	420
ccaaaggctt	gatcttgc	cacagacggg	aagtccagg	acctgggtga	cacagctgc	480
caaaggctga	aggggcagg	ggtcaagct	tttgtgtgg	ggatcaagaa	tgctgacc	540
gaggagctg	acgcgatgtc	ctcagacgg	accagtact	tcttcttct	cgtaatgac	600
ttcagcatct	tgaggacact	actgc	tttccggga	gagtgtgcac	gactgtgtt	660
ggcgtgcctg	tgacccgacc	tccggat	tgcaccttgc	ctccacgaga	cctgggtgt	720
tctgagccaa	cgagccaaatc	tttggagat	cagtggacag	cgccca	ggccatgt	780
ggctacaa	ttccatgtac	tcctgtgc	gggctgggac	agccactg	gggtgagg	840
caggagg	atgtccac	ttgtggag	atgtgtggc	tgcgggg	ccggccact	900
accgagatcc	aagtactgt	gattgc	tacgcca	gatcgggg	ggctgtgac	960
gggacagctc	ggaccactc	ctttagaagg	ccggaa	ccatcc	taccacagcc	1020
cacagccccc	ttgtgtggc	ggggatgt	ccaggatgt	ccgttaccc	tgtacatgg	1080
cggttctca	ttgtgtggc	cacacag	caggatgt	gccttgggg	gggttca	1140
tttgtgtgt	acttggat	tggacgg	tatggat	ccgtgac	cctatttgc	1200
cgcagtgtgg	ggccgc	tttctgtat	gctcgact	acgttctgt	tgagcagacc	1260
ctgcgcggcc	tcatecttgg	ccccacat	atcttctt	cttggaa	cttggctg	1320
gcccgtgtt	accgggttga	atggggcg	gagactgt	tggacccacc	gcagaagg	1380
gtactgcct	ctgtatgtac	ccgttac	tttggatggc	tgcacccgg	cactgagat	1440
cgcctcac	tcttacatct	gttggagg	cacggatgt	ccacccctgc	aaccgtgtt	1500
cccaactggc	caagatgtcc	tgtgtggc	gttacagac	tgcacccac	cgagctgccc	1560
gggcagccgg	tgccgtgtc	ctggac	gtccctgg	ccacccat	ccgcac	1620
gtgcgcac	ccccagggtt	tgaggcgg	ctgggtt	ctggggat	tgacat	1680
gacttggat	acgttca	ttggat	tacactgt	gggtgtctc	tgcagtgg	1740
cccccgtgg	caatgtccat	gttgc	ccgggg	agccggaa	tccacttgt	1800
gttccatggc	ttgggtgtt	gggtt	gttgc	tgagggt	ccggggacc	1860
gtccctgtgg	caatgtgtt	ttggat	tgaggc	ccatgttcc	gggttcc	1920
cacacactc	ccccagactc	tactg	gacat	ggctgc	tggaa	1980
taccagg	ttgtgtcggt	actgc	gggtt	gggtgtctc	ttgtgtgt	2040
gttcgcac	cccaactgg	cccaatgt	acgttccat	tgactc	ccggccat	2100
tctgtca	ttatctgg	caggatgt	ggc	ccacccat	tgatc	2160
tcagccac	gcccac	atccatgt	gttcttgg	aggccac	ggctgac	2220
gatggactt	agccagata	tgtat	gtgc	ccatgttcc	gggttcc	2280
gatggggccc	ctgcctctgt	gggtt	actgc	ccatgtt	tgatgtgt	2340
aggctcaga	tttcaatgc	ttccag	tttca	ccatgtt	gggggtact	2400
ggagccac	tttacatgt	ggccttgg	ccggatgt	ggggcc	gaggac	2460
atactccca	gaaacacaga	ctctcg	atccgggtt	tgc	ggatgttcc	2520
tcagtgc	gatgtgt	tttgc	tggtt	ccatgtt	ttgtgtt	2580
actacgcgc	ctgagg	tttgc	ccatgtt	ccatgtt	ccatgtt	2640
cactcg	ggctgcgt	ggagccgt	cccagac	ccatgtt	ccatgtt	2700
caacatcg	gtggccagg	acatccc	gttctgg	ccatgtt	ccatgtt	2760
ctggacggc	ttggaccc	gacacatgt	ccgcgt	tgatgtt	ccatgtt	2820

-continued

ggagaagggc	cctctgcaga	ggtaactgctg	cgcaactgatgt	cacccgtgt	tccaaagcatt	2880
gaactacgtg	tggtggacac	ctcgatcgac	tcgggtactt	tggcctggac	tccagtgtcc	2940
agggcatacca	gtacatccct	atccctggcg	ccactcaga	gcccctggca	ggaagtgcct	3000
gggtccccgc	agacacttcc	aggatcta	agtcctccagc	gggtgacagg	gttagaggct	3060
ggcgctcttt	acatttctc	cctgacgct	gtcttggatg	gtgtgegggg	tcctgaggca	3120
tctgtcacac	agacgcacgt	gtgecccccgt	ggcttggcg	atgtgggtt	cctaccacat	3180
gccactcaag	acaatgtca	cegtgcgg	gtacgagg	gggtcttgtt	gctgttgtt	3240
ttggacttgc	ggccctttgg	gcccaggca	gttcagggtt	gtctgtgtc	ttacagtat	3300
cggccctccc	oactgttcc	actgaatgg	tcctatgacc	ttggcattat	cttgc当地	3360
atccgtgaca	tgccctatcat	ggaccaagt	ggaaacaaacc	ttggcacagc	cgtgttca	3420
getcacagat	acatgttgc	accatgt	cctggcgcc	gcaacacgt	accagggtt	3480
atggtttgc	tagtggatg	acccttgg	gggtgacat	tcagccccat	ccgtgaggcc	3540
caggcttgc	ggcttaatgt	ggtgatgtt	ggaatggct	gagcggaccc	agacgactg	3600
cgtcgcttgc	ccccgggtat	ggatctgtt	caagacccat	tcgcctgtt	tgatgggca	3660
agectggacc	aggcagtc	tggttgc	acagccctgt	gtcaggcata	cttactact	3720
cagccccctg	cagacccctg	cccaatgtat	tgtccaaagg	gcaacagggg	ggaaccttgg	3780
gagatggggc	tgagaggaca	agttgggc	cctggcgacc	ctggccccc	gggcaggacc	3840
gggtgcctccg	gccccccagg	gcccccttgg	agtgccact	ccaaggccga	gaggggttcc	3900
cctggagcag	atggggctcc	aggcagcc	ggccgcgc	ggaatcttgg	gacccttgg	3960
gccccctggcc	taaagggtctc	tccagggtt	cctggccctt	tggggacccc	gggagacgca	4020
ggacctcgag	gccccaaagg	ggacccgggg	gtcccccgg	aagtcatcgg	aggtgaagga	4080
cctgggttcc	ctggggggaa	aggggaccc	ggaccatcg	gcccccttgg	acctcggtt	4140
caactggggg	accaggacc	ccgtggcc	ccagggttcc	ctggaaacagc	catgaagggt	4200
gacaaaggccg	atcggtggga	ggggggcc	cctggaggac	gtgaagggtt	cattgtctt	4260
ggggagctgt	ggctgcgggg	tcttccgg	agccctggac	cccaaggccc	cgtggcccc	4320
cctggaaaga	aaggagaaaa	aggtgactt	gaggatgg	ctccaggctt	cccaggacaa	4380
cctgggttcc	cggtgtgaca	ggggccacgg	ggacccctcg	gagctatttg	ccccaaagg	4440
gaccggggctt	tctccaggccc	ccttgggtgg	gctggagaga	aggcgacacg	tggaccccca	4500
ggcccgccgg	gatccccggg	gctgocagg	gttgctgg	gtcttggagc	caagggttct	4560
gaaggggccac	caggacccac	tggccggcc	ggagagaagg	gggagcttgg	tcgccttgg	4620
gaccctggcg	tggtggggac	tgctgttgc	ggacccaaag	gagaaaagg	agatgtgggg	4680
cctgggttgc	ccaggagac	taccggatc	caaggggaa	ggggcccccc	cggttggtt	4740
cttcctggag	acccttggcc	caaggagac	ccttggagacc	ggggccctt	tggcttact	4800
ggcagagcag	gaccccccagg	tgactcagg	ccttgggtgg	agaaggggaa	cccttggccgg	4860
cctggggccc	caggactgt	tggccccc	ggacggatgt	gtgaagggtt	agagaaaagg	4920
gacggagggtc	tcctgggtga	ccttgggtt	ccttggaaa	caggcgacgg	tggccttcgg	4980
ggggcacctg	gatgttgggg	gcctgttgg	gaaaagggg	accaggggaa	tcctggagag	5040
gatggacgaa	atggcagccc	tggatcat	ggacccaa	gtgaccgtt	ggagccgggt	5100
ccccccggac	ccccccggac	gtgtgttgc	acaggactt	gaggccagaa	gaagggag	5160
cctggggacc	cgccgacaa	gggttctcg	ggggccaa	gtatcttgg	cctcccttgg	5220
gccccctgggg	aaaggggcat	tgaagggtt	cgggggaccc	caggccccaa	gggggaccca	5280
ggtgtccgag	gccccagg	agaaaagggt	gacccgggtc	ccccctgggt	ggatggccgg	5340
agcggactgg	atgggaaacc	aggacccgt	ggggcccttc	ggccgaatgg	tgcttgaggc	5400
aaagctgggg	accaggggag	agacgggtt	ccaggccctc	gtggagaaca	ggggccctt	5460
ggccccccttgc	gtccccctgg	attaccgg	aaaggcagg	aggatggcc	accttggctg	5520
aatggaaaaa	acggagaacc	tggggacc	ggagaagac	ggagggaaagg	agagaaaagg	5580
gattcaggcg	ctctggggat	agaaggctgt	gtatggccca	agggttgcgt	ttggagctt	5640
gttatcttgc	gaccccgagg	gttccagg	tcctccagg	cagtggggcc	tccttggccag	5700
ggtttcttgc	gtgttccagg	aggacccgg	cccaagggtt	accgtgggg	gacttgatcc	5760
aaaggggagg	aggggctccc	tggagagcg	ggcttgc	gagagcttgg	aagtgtggc	5820
aatgtggatc	gtgtgttgc	aactgttgc	atcaaggat	ctgccttgc	ggagatctgt	5880
gagacggatgg	atagagatc	tggtatgttgc	ctgccttgc	cgaaacggcg	tcgaggcccc	5940
aagggggact	caggcgacaa	ggggccccc	ggcaaggagg	ccccatcgg	ttttcttgg	6000
gaacgcggcc	tgaagggcga	cgctggagac	cctggccct	aggggcaccc	tggcttggcc	6060
cttggggaga	ggggcccccc	cgggcttcc	ggcttgc	gggagcttgg	aaaggcttgg	6120
attcccccggc	tcccaggcg	ggctgggggt	gtggggagg	caggaaagg	aggagagg	6180
ggagaacccgg	gagagaaaagg	agaacgttgc	ggacaggc	gatgttccc	tccttggactc	6240
cctggaaacc	ctggggcccc	cggacccct	ggcccaagg	tgtcttgc	tgagccagg	6300
cctggactct	ctggagaaca	ggggccccc	ggactcaagg	gtgtcaagg	ggagccgggc	6360
agoaatgttgc	accaaggatc	caaaggac	aggggtgtgc	caggcatca	aggagacccg	6420
ggagagccgttgc	gaggaggggg	tcaggacgg	acccgggttcc	taccaggaga	tcgttgtat	6480
gctgggttgc	aaagggaaggcc	gggttgc	ggtccaaagg	gcccccttgg	ccccatgggt	6540
ggtcatggag	acccttggacc	accttggtgc	ccgggttgc	ctggcccttc	aggaccccaa	6600
ggaccccttc	gtcttgcgg	ggagacccgt	ctccaggac	gggcctgtact	6660	
ggacacttgc	gagcttgcgg	acttcttgc	ccccccggcc	cttcaggctt	tgtgggttca	6720
cagggttgc	cagggttgc	tggacaatgt	ggggagac	ggagccccc	agccccagg	6780
cgagatgtgt	ccaggatggaa	agatggac	agggggac	ctgtgttgc	agggttccacca	6840
ggtctgttgc	ggccctgtcg	acctaaagg	gaaaccttgc	ccacggggc	cccttggacag	6900
gtctgttgc	ggcttgcgg	agcaaggagg	gagaaggagg	cccccttgg	ccttgcgttgc	6960
gaccccttgc	gttcccccgg	agccaaagggt	gaccggagg	tgccaggggc	gcgaggccg	7020
aagggttgc	ctggccgtgc	agggggaccc	ggagacccct	gggaatgttgc	tcaaaaagg	7080
gctccaggac	ccaaagggttgc	caagggttgc	ccaggatgt	gggttcccccgg	ctcccttgg	7140
cctcccttgc	ctccagggttgc	gaaggagat	ctggcccttgc	ctggcccttgc	cggtgttct	7200
ggtgttgttgc	gttcccccgg	tcagacagg	cctcgttgc	ggatgggttca	gccaggccct	7260
agtggagac	ggggtcttgc	agggggaccc	ggggagaga	gaatcccagg	acccttgggg	7320
ccaccccttgc	cacccggggc	agttggggacca	ccttggggcc	ctggactca	aggagacaag	7380

-continued

ggagaccctg	gagttagggct	gcctggggccc	cgaggcgagc	gtggggagcc	aggcatccgg	7440
gtgtgaatgat	gocgcggccgg	ccaggaggaga	ccccggaggac	tcacggggcc	cccttggcagc	7500
aggggagagc	gtggggagaa	gggtgtatgtt	gggagtgcag	gactaaaggg	tgacaaggga	7560
gactcagotg	tgtatcttggg	gcctccaggg	ccacgggggt	ccaaggggga	catgggtgaa	7620
cgagggectc	ggggcttggg	tggtgcacaaa	ggacacctggg	gagacaatgg	ggaccccttgtt	7680
gacaaggggca	gcaaggggaga	gcctgggtgac	aagggtctca	cggggttggc	aggactgcgt	7740
ggactcttgg	gaccccaggg	teaaacctgtt	cgacgcaggga	teccctgttga	cccgggatcc	7800
ccaggaaagg	atggagtgcc	tggtatccga	ggagaaaaaa	gagatgttgg	tttcatgttgg	7860
cccccggggcc	tcaagggtga	acggggagtg	aaggggagct	gtggccttga	tggagagaag	7920
ggagacaaagg	gagaagggtg	tcccggggc	cgccccggc	tggcaggaca	caaaggagag	7980
atggggggatc	otgggtgtcc	ttcccccagg	cgccccggc	ttcccccagg	ttcccccagg	8040
cacaagggttgc	accgaggctt	tgacggggcag	ccaggccccca	agggtgacca	gggcgagaaaa	8100
ggggagcggg	gaaaaaccagg	aatttggggc	ttcccccagg	ccagtggaaa	tgatggctct	8160
gtctggcccc	caggggccccc	tggcgttgc	gggtttccaa	gcccccaagg	acttcagggc	8220
cagaagggttgc	acggaggatcc	ccccggagag	agagtgttgg	gggttcttgg	ggttcccttgg	8280
gtctctggcc	aqagaggggga	gcagggggg	ccagggtcttgc	ccgggtcttc	aggcgagaaag	8340
ggagaagctg	cactgacgga	ggatgacatc	cggggttttgc	tgcccaaga	gatgagtcag	8400
caactgtgcct	gcccggccca	gttcatcgca	tctggatcac	gacccttccc	tagttatgtt	8460
gcagacatc	cgggtccca	gttccatgt	gtgcctgtgc	teccgttctc	tcatgcagag	8520
gaggaagagc	gggttacccccc	tgaggatgt	gagactcttg	aatactccga	gtattctgtt	8580
gaggagtacc	aggaccctga	agcttcttgc	gatagtgtatg	accctgttc	cctgcactg	8640
gatgaggggct	cttgacttgc	cttaccccttgc	cgctggtaacc	atcgggctgt	gacaggcagc	8700
acagaggctc	gtcaccccttgc	tgtctatgtt	ggctgttggag	ggaatgcocaa	ccgtttttggg	8760
accctgttgc	cttgcgttgc	ccgtgtccca	ccccgggttgc	tccagagcaca	ggggacaggt	8820
actgcccagg	actga					8835

SEQ ID NO: 12	moltype = DNA	length = 8835
FEATURE	Location/Qualifiers	
misc_feature	1..8835	
source	note = Synthetic Construct	
	1..8835	
	mol_type = other DNA	
	organism = synthetic construct	

SEQUENCE: 12						
atgacccttgc	gactgttgtt	ggctggccctg	tgtgtctggaa	ttcttggctgta	ggctcttaga	60
gtgcggggccc	acatagaga	aaagtgacc	tgtacacggc	tgtacccgc	cgatatctgt	120
tttctgttgg	atggcagcag	cagcatcgcc	cggttcaatt	tcaagagaatg	cgccggatctt	180
ctggaaaggcc	ttggtgcgtcc	tttttcttgc	gctgtttctg	cccaagggtt	cagatttgc	240
accgttgcgt	acagcgacca	ccctagaaca	gagtttggcc	tggatgtct	cggcagccgc	300
ggagatgttgc	tcaagccat	caagagctg	agctacaagg	cgcccaatac	cagaacaggc	360
gcccgttattc	tgcactgtgc	cgaccatgtt	tttctggccc	aactggctag	accccgctgt	420
ccaaaagtgt	ycatcttgc	cacagacggc	aagagccagg	acctgggttgc	tacagccgt	480
caagagactga	aaggacagg	cgtgttgcgtt	tgatcaaggaa	cgcccgatcc	540	
gaggaacttgc	agagatggc	cagccagct	accagcgat	tettctttt	cgtgaacaggc	600
ttcagcatcc	tgccggccct	gctgttgc	tggtcttgc	tggtcttgc	aacagccggc	660
ggagtgcgttgc	tgaccagacc	tcctgtatgt	agcacaaggc	cccttgcata	tctgtgttgc	720
tctggccca	tgccggccat	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	780
ggttacaagg	tgcaatatac	acctctgaca	ggctggggcc	acccctgtcc	ttcttgagaga	840
caagaatgttgc	acgtccccc	cgccggaaaca	tctgtcaac	tgagaggact	ggggccctgt	900
accggatgttgc	aaatgttgc	gatccgttgc	tacgcaatt	ctatccggc	ggccgtgtct	960
ggcacacoca	gaacaacacg	tctggaaagg	ccagagctg	ccatccggaa	cacaacagcc	1020
cattctctgc	ttgggttgc	gagatctgt	cctggccca	ccggctatag	atgtacttgc	1080
agagtttgc	tgccggccat	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	1140
ctgtgttgc	tgccggatgt	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	1200
agatctgttgc	gacccgttgc	aaggctgttgc	ggccggatgt	atggccatgt	ggacacagacc	1260
ctggggccgttgc	tgatgttgc	ccctgttgc	atccctgttgc	gttggatct	ggtgcggcag	1320
ggccggatgttgc	atagacttgc	atggccggat	gaaaccggcc	tggaaatccc	tcagaaatgt	1380
gtctctgttgc	tgccgttgc	aaagataccag	ctggatgttgc	tgccggatgt	ccctgttgc	1440
agactgacac	tgtatatact	gtctggggc	cacggatgttgc	ccacaccaggc	tacagtgttgc	1500
ccaaacaggcc	ctggatgttgc	tgtgttgc	gtgactgtatc	tgccggccac	agaatctggc	1560
ggccggatgttgc	tgccggatgt	ttgggttgc	ttgggttgc	ttgggttgc	ttgggttgc	1620
gttccctgttgc	ttgggttgc	taaatcgc	tggtcttgc	tggtcttgc	tggtcttgc	1680
gatctgttgc	tgccgttgc	aaatgttgc	tgccggatgt	tgccggatgt	ccctgttgc	1740
cttggatgttgc	tgccggatgt	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	1800
gttccctgttgc	ttgggttgc	ttgggttgc	ttgggttgc	ttgggttgc	ttgggttgc	1860
gttccctgttgc	ttgggttgc	ttgggttgc	ttgggttgc	ttgggttgc	ttgggttgc	1920
cagacactgc	ctcccgatag	caccggccac	gatattacag	ggctgcagcc	tggaaatctacc	1980
taccaggatgc	ccgttgcgttgc	gtctggatgttgc	tgccggatgt	tgccggatgt	ccctgttgc	2040
gtctggatgc	tgccggatgt	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	2100
agactgacac	atccctgttgc	acccctgttgc	acccctgttgc	acccctgttgc	acccctgttgc	2160
agcgttgc	tgccggatgt	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	2220
gacggacttgc	aaatgttgc	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	2280
gatggacatc	ctgcttgc	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	2340
agactgacac	ttctgttgc	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	2400
ggccggatgttgc	ttgggttgc	ttgggttgc	ttgggttgc	ttgggttgc	ttgggttgc	2460
atccctgttgc	tgccggatgt	tgccggatgt	tgccggatgt	tgccggatgt	ccctgttgc	2520

-continued

agcgtagatg	tgcacagccct	cgtggggcat	agagaaggca	cacctgtgtc	catcggtgc	2580
accacaccc	cagaagctcc	tccagctctg	ggaacactgc	atgtgggtca	gagaggcgag	2640
cactctcg	gactgatgt	ggaacactgt	ccacgcgc	ccacgcgc	aggaggatct	2700
caaccagaag	goggacaaga	gcagagcaga	gtgctggac	cagaactgag	cagtcaccac	2760
ctcgatggac	ttgagctgc	cactcagtat	agagtca	gac	tgtcagtgt	2820
ggcgaaggac	cttctgctg	agtgacaget	agaaccgc	gat	ctcccagagt	2880
gagctgagag	tcgtggatac	ctccatcgat	agcgtgacac	tg	gctggac	2940
agagccacgt	tttacatccct	gtcttgcgg	cctctcaqag	gaccaggcca	agaagtccct	3000
ggatctccac	agacactccc	aggcattgt	agtcccaga	gagtga	acttgaacca	3060
ggcgtgtct	acatcttcag	cctgacac	gtgctggac	gctgg	actctgaa	3120
tetgtgaccc	agacttcgt	gtgttgcata	ggactggcc	atgtgg	gtt cctgctc	3180
gtcacacagg	acaatgccc	tagagccag	gccaca	aa	gagactgtgtt	3240
cttgccttg	gaccactggg	accc	ttt	gttca	actgtgtc	3300
agacccttc	cactgttccc	tctgaa	ccgc	ttt	tcacgcac	3360
atccgggaca	tgccctcat	ggacccat	ggcaaca	ttt	ggcgcac	3420
gccccacat	atatgtggc	ccc	ttt	ttt	gttgcact	3480
atggtgtgc	ttgtcgatg	acc	ttt	gg	ccctgqcg	3540
caggccacg	gcctgtatgt	ttt	gg	gg	gtatct	3600
agaagggtt	ccccctggat	ggatagcgt	ccac	ttt	ttt	3660
tctctggatc	aggctgtgc	tggactgt	ttt	gg	gttgcact	3720
cagecttagac	ctgagcttg	tccagtgtac	ttt	gg	ttt	3780
gaatgtggac	tttagaggcca	agtgggacca	ttt	gg	ttt	3840
gtgtgcctcg	gaccccaagg	ttcttctgtt	ttt	gg	ttt	3900
ccttgcgtcg	atggcagacc	aggatctct	ttt	gg	ttt	3960
gcaccaggac	tcaaaggatc	tccaggact	ttt	gg	ttt	4020
ggaccaagag	gtccaaaagg	cgaa	ttt	gg	ttt	4080
cccgattgc	cggggaaagg	aggccac	ttt	gg	ttt	4140
ccttggggat	atcccggtcc	ttt	gg	ttt	ttt	4200
gacaaaggcg	ataggggaga	ttt	gg	ttt	ttt	4260
ggcgaaccag	gttgccttgg	actccctgg	ttt	gg	ttt	4320
cctggaaaga	aaggcgaaaa	gg	ttt	gg	ttt	4380
ccaggcttc	caggcgaaca	aggcc	ttt	gg	ttt	4440
gacagagggt	ttcccg	gg	ttt	gg	ttt	4500
ggacctgcg	gatctgatgg	acttcc	ttt	gg	ttt	4560
gaaggcccc	ccatc	gg	ttt	gg	ttt	4620
gatccctgtcg	tttgcggacc	ttt	gg	ttt	ttt	4680
cctgtggac	caagggggac	ttt	gg	ttt	ttt	4740
ttgcggggc	atcccgacc	taaggc	ttt	gg	ttt	4800
ggcagagctg	ggcc	ttt	gg	ttt	ttt	4860
cctggac	ccttggc	ttt	gg	ttt	ttt	4920
gacagaggc	ttcc	ttt	gg	ttt	ttt	4980
ggcgctcg	gtgttagagg	ccctgtt	ttt	gg	ttt	5040
gatggaaag	atggccccc	ttt	gg	ttt	ttt	5100
ccacgggtc	caccaggc	ttt	gg	ttt	ttt	5160
cccg	gttggac	ttt	gg	ttt	ttt	5220
gcaccaggcg	agcgaggat	ttt	gg	ttt	ttt	5280
gttgtcagag	gac	ttt	gg	ttt	ttt	5340
tccggactt	atggaaaacc	ttt	gg	ttt	ttt	5400
aaagctggc	acc	ttt	gg	ttt	ttt	5460
ggac	ttt	gg	ttt	ttt	ttt	5520
aacggccaaa	acggc	ttt	gg	ttt	ttt	5580
gacagcgg	ctt	ttt	gg	ttt	ttt	5640
ggaattttt	gaccccaagg	ttt	gg	ttt	ttt	5700
ggtttctcg	gtgtcc	ttt	gg	ttt	ttt	5760
aaggcgacg	aaggacttcc	ttt	gg	ttt	ttt	5820
aatgtggaca	gactcttgg	ttt	gg	ttt	ttt	5880
gaaacccctgg	atagagac	ttt	gg	ttt	ttt	5940
aaaggcgatt	ctggc	ttt	gg	ttt	ttt	6000
gacggggc	ttaaaggc	ttt	gg	ttt	ttt	6060
cttggagaa	gggg	ttt	gg	ttt	ttt	6120
ataccggac	ttcccg	ttt	gg	ttt	ttt	6180
ggagaaaagg	gaga	ttt	gg	ttt	ttt	6240
cctggaaactc	ctgg	ttt	gg	ttt	ttt	6300
ccaggctt	gggg	ttt	gg	ttt	ttt	6360
tcaaatggc	agg	ttt	gg	ttt	ttt	6420
ggcgaacc	gac	ttt	gg	ttt	ttt	6480
gcccggac	agg	ttt	gg	ttt	ttt	6540
ggacatggc	acc	ttt	gg	ttt	ttt	6600
ggaccatctc	gact	ttt	gg	ttt	ttt	6660
ggacatctc	gact	ttt	gg	ttt	ttt	6720
caaggcttac	ctgg	ttt	gg	ttt	ttt	6780
cgagatggcg	cat	ttt	gg	ttt	ttt	6840
gggctcc	gac	ttt	gg	ttt	ttt	6900
gcagtgtt	gat	ttt	gg	ttt	ttt	6960
gatctgtt	gaga	ttt	gg	ttt	ttt	7020
aaaggcga	ctgg	ttt	gg	ttt	ttt	7080

-continued

ccccctggc	ctaaaggctt	taagggcat	cctgggtcg	gcgtcccaagg	tttccttgg	7140
ccacctggc	cgccaggcg	taagggcat	ttggacttc	ctgggtgcc	tgccgctcc	7200
ggcggttgg	gatttccagg	acaaacttgc	cctcggggag	aatgggtca	acctgggcct	7260
agoggagaaa	gaggcattgc	cggacctct	ggaagagagg	gaataccctgg	accacttgg	7320
ccgcctggc	ctccaggttc	tgtggacct	cctggcgott	caggattgaa	ggcgataag	7380
ggcgaccctg	gtgttggact	tcccggtct	cggtggcgaaa	ggggagaacc	aggcattaga	7440
ggcgaagatg	gacggcctgg	acaagaggga	cctagaggcc	ttacaggccc	accaggetct	7500
agaggcgaaa	ggggcgagaa	aggegacgtg	ggttctcgac	gacttaagggg	cgacaagggc	7560
gatccgctg	tgattttggg	accaccggg	ccaagaggcg	ctaaaggcga	tatgggcag	7620
agaggccta	gaggcctgg	tgggcataaa	ggggcaaaagg	ggcacaaagg	cgacccaggc	7680
gataagggt	ctaaaggcga	aecggcgat	aagggtctg	etggactggc	aggacttcga	7740
ggacttctgg	gtccacaagg	ccaaacctgg	gcccggcgca	ttcccgccga	tccgggtct	7800
cctggcaaaag	atgggttcc	tggcatcagg	ggcgaaaagg	gcegactcg	ttttatggc	7860
cetcgcggat	tgaaaaggcg	gaggggtt	aaggcgccct	gtggacttga	cgcgaaaaaa	7920
ggcgcatagg	tgcaagctgg	accttgcgc	agaccttgg	tggctggata	caaggcggag	7980
atggagagc	ccgggttcc	tggacaaaatg	ggcgccccctg	gaaaagaagg	actgtatcg	8040
ccaaaggccg	accggcgatt	tgatggccaa	ccgggtctca	aaggcgatca	gggcgagaaa	8100
ggcgaaaaag	gcaactctgg	catecgagg	tttccaggac	caagggccaa	tgatggatct	8160
gctggccgc	cagggtctcc	tggatctgtt	ggtccaaagg	gaccggaaagg	ctgtcgagg	8220
caaaaaggcg	agcgcggacc	accaggcgaa	agagtgttgg	gagcccccgg	cggtcccggt	8280
gctccggcg	aaagaggcga	acaaggcaga	cctggaccag	ctggccctag	aggcgaaaaaa	8340
ggggaaagccg	ctctgaccga	ggacgatata	agaggctttt	tgccgcaaga	gatgagccag	8400
cattgtaccc	gtcaggggc	gtttatcg	agegggttca	gacctctggc	tagctatg	8460
gctgtatccg	cggatctca	gtctgtatgt	gtgcctgtc	ttagagtgtc	tcacggccag	8520
gaagaggaaa	gagtccctcc	agaggacgac	gagtaactcc	agtatagcga	gtactctgt	8580
gaagagttatc	aggaccccga	ggctcttgg	gatagcga	atccatgtc	tctgccact	8640
gatggaggca	gtgttacccg	ctacacact	aggtggatata	acagagocgt	gaccggaa	8700
accggaggct	gccccatctt	tgtttatgg	ggctgcggcg	gcaacgccaa	tagatttgg	8760
acaagagagg	cctgcgagcg	gagatgcct	ccaaagactgg	ttcgtctca	aggcaccggc	8820
actgcccagg	actaa					8835

SEQ ID NO: 13 moltype = DNA length = 4494
 FEATURE Location/Qualifiers
 source 1..4494
 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 13

atggatgtaa	ccaagaaaaa	caaacgagat	ggaactgaag	tcactgagag	aattgtcact	60
gaaacagtaa	ccacaagact	tacatcccta	ccacccaaaag	gggggaccag	caatggctat	120
gctaaaacag	cctcttctgg	tggggggagc	cggtggggag	aacaaagct	gactcatggc	180
agcagcgct	acataaaactc	aacttgcggc	acacggggcc	atgcctccac	ctctgttac	240
aggagggtctc	actcacctgc	cttccactctg	cccaactccc	caggctcaac	ctttgaaagg	300
aaaactcactc	tatccggcc	tgccgtatgg	ggggagcttca	gtggcaactc	ttctccggag	360
taccctcgga	aggaatttgc	atcttctca	accaggaggc	gggtcaaaac	acgagagagt	420
gaaatttccg	tgcactgc	gagtgcgcc	ccatccaccc	gatggacaga	attggatgt	480
gttaagcggt	tgctcaaggg	gagtcgtatc	gcaaggtgt	gccccacccg	gaattctcc	540
aacacactcc	ccatccccc	gaaaggcact	gtggagacca	aaatgtgtac	agcgagctcc	600
cagtctgggt	caggcaccta	cgatgcac	atcttggat	ccaaaccttcc	ctcccatgt	660
tggtcctcca	ccctgcccgc	ggggtctcc	atggggactt	atcacaacaa	catgacaacc	720
cagagctcat	ccctctctca	caccaatgc	tactctcg	gtcgttgc	cgaggatcca	780
aacaacatgg	cgctcgctc	accctacttgc	caccctggac	tcagcacatc	ctccctca	840
tttggcatgc	agaacaatct	ggccccccag	ttgaccaccc	tgttccatgg	caccaccc	900
acttccacag	catatgggg	gaaaaaaaaaa	atgccccaga	gtccctggc	tgttgcacat	960
ggcgtttcca	cctccggcgc	ctgcaccaca	agtgtgcaga	gcgtatgc	tttgcacaag	1020
gactgcaga	tctctgtatcc	agagaaaaaa	aacacacactt	ccaaagaa	gatggagct	1080
ctctatcatga	ccaaggacag	cgggagggtc	tttacagcc	ccctgc	catcgctgca	1140
acttctttt	cagaagacac	cctaaaaaaa	gaaaaggca	ctgcctacaa	tgttgcacat	1200
ggctctaaag	ccgaagctaa	tggagac	ttgtgtgt	ccacaaagg	caagaccacc	1260
actgcagata	tccacacgt	cggcacgt	gggtgggt	gcagtg	ggggaggg	1320
gttggggcgc	ctgggggggg	ccctttgggg	ccaggcc	ctgggtgccc	ctcgccgtcc	1380
tgttgtcgat	gggtggaaatg	gtgtgttgg	ctgtgttca	cttgggtgt	actcttgggg	1440
ctgtctttcg	gcctcatgtc	tctggggag	gaggtggag	agctgttgc	gctgtgttgg	1500
gagctggaga	ggatcaggag	gagcataact	ccctatgggg	acagcatg	gtaataagaa	1560
aaaggccaggc	tccacggat	ggcaccggcg	ggggggagcc	acctggacaa	aattgggtcg	1620
cacagtgc	gcccaggag	gtcttggat	ttcgttgg	ggag	ggggccgt	1680
gaaaatggaa	atctccgg	aaggccctgc	cctaaagggt	acatggga	tccaggccct	1740
aaaggagatc	gagggttccc	tggacttca	ggtatccctg	ggcccttggg	ccacccagg	1800
ccacaaggac	caaagggtca	aaaaggcago	gtggggagatc	ctggcatg	aggccccat	1860
ggccagagag	ggcgagaagg	ccccatgg	cctctgtgt	aggcagg	ggggcc	1920
ggagagaaag	ggggaaagagg	ggctgtgt	ggacccagg	ctcatggccc	acctgggtgc	1980
ccagggttgc	tgggtccaa	agttccago	ggctcttgc	gcccacagg	ccctccagg	2040
cctgttagtc	tccaaagggt	ccggaggt	gttggacttc	ctgggttca	aggtgacaaa	2100
ggaccaatgg	gaccaccagg	accctaaagg	gaccagg	ttgtgttgc	tcgagg	2160
acaggcgac	ctggcatg	agggttgc	gggtgtgttgc	gtgagcc	ggggccgg	2220
gcaatgggtc	ctgtggccc	agacggacac	caaggccaa	gagggt	aaaca	2280
ggatctgg	gatctgg	ccaccagg	ccttctgg	accctagg	ggccagg	2340

-continued

acaggacccc	agggacactca	gggacttccc	ggtacccctg	gcccaccagg	aataaaaggt	2400
gaaccaggag	ctccaggcaa	gatecgact	tcggagggggt	catcgatgt	cactgtccca	2460
ggccccccag	gacctcctgg	aggcatggga	ccccaggac	ctccagggtc	cccaggccct	2520
gcggggcccg	ctggttccc	aggacatcaa	gaagttcta	atttacaagg	tcccccaggc	2580
ccaceccggcc	cacgcgggcc	accagggcct	tccattccag	gcccaccagg	accccgaggc	2640
caaccagggg	agggttgtcc	aggeccacca	ggeccaccac	gatcggtct	gtccaactca	2700
aaaacccctcc	tctccggccc	cccaggccca	cctggggccc	caggteccaa	gggagaccaa	2760
gggtccccccag	gccccagagg	acaccaaggc	gagaaggc	tcccagggtt	ctcaacccctca	2820
gggtccagggt	tttcggact	caacccatcg	ggaccaccag	gcccacctgg	ccccaggga	2880
ccaaaagggt	acaaaaggta	tccagggtt	ccagggttcc	tttgcattcc	tagtggtct	2940
tctgaagggg	gatcatcaag	tccatgtac	ttgtcaggcc	cgccaggggc	ccctggggcc	3000
cctgggcttc	cgggctctat	cagcagctt	ggccaggaga	ttcagcagta	catactctgag	3060
tacatgcaga	gtgacagtat	tagatcttac	ctatccggag	ttcagggtcc	cccaggccca	3120
cctgggtcccc	caggacgtt	caccacccat	acaggcggaga	cttgcacta	ctcagagctg	3180
gcaaggccacg	ttgtgactca	cttacggact	tcggggatc	ttgtcagtt	gttctcgcc	3240
tccatcttt	cttgaqaacat	tctggatqtg	ctgcagcggg	atgacgttgcg	tcagatacta	3300
cgtcagtaact	tgatgggccc	tcggggtccg	ccagggccac	caggagccag	tggagatggg	3360
tcctctctgt	ctttgacta	tcgcagatgt	agtatgtcc	ttctcagcta	catgtcgagt	3420
tctgggatca	gcatgggtt	tcctggtccc	ccggggccc	ctgggttccc	ggggacccctcc	3480
tatgaggagc	tcctctctt	gctgcgggg	tctgaaatca	gaggcatcgt	tggaccccca	3540
ggtccccccgg	gtccaccagg	gatcccgagg	aatgtgtggt	ccagcatcag	cgtggaggac	3600
cttcgttatt	atttacatac	tccgggttcc	tcattcatcc	caggccctcc	aggaccttct	3660
ggtccccccag	ggcctcgagg	gccccgggtt	gtctcaggag	ccctggcaac	ctatgcaggt	3720
aaaaacagcg	acagttccg	gagcagatgt	atcagttacc	tcacaaggcc	tgatgtgcgc	3780
agtttcatgtt	ttggcccccc	agggcctctt	ggggccgggg	gacccttgg	ggacagcgcg	3840
cttcgttcca	cggtatccctc	ccacagtcgg	ggtagcagct	cctccctaca	cagtcatct	3900
gtcaggccgg	gcaagtccta	cagetttcc	atgagcacag	gaggagggtt	tgcaaggctcc	3960
ctgggttccag	gggggttccctt	ttgttgcgt	ttgttgcgt	ttggacttca	ttggcaactgac	4020
atcggccccc	gggggggttca	ttggggcagea	ttggggcagea	ttggggcagea	ttggcaatggc	4080
ggacttattgg	gagctgactt	ttgttgcgt	ttgttgcgt	ttgttgcgt	ttgttgcgt	4140
tcagagagca	tcgcagtcgt	ggggctactg	ttgttgcgt	ttgttgcgt	ttgttgcgt	4200
ccaggccatc	ctggggccacc	ttggccctca	ttggccctca	ttggccctca	ttggccctca	4260
ccagggttccat	atgtggggc	ttggccctca	ttggccctca	ttggccctca	ttggccctca	4320
gagaaggatggg	ttggccctca	ttggccctca	ttggccctca	ttggccctca	ttggccctca	4380
gagaaggatggg	ttggccctca	ttggccctca	ttggccctca	ttggccctca	ttggccctca	4440
gagaaggatgg	ttggccctca	ttggccctca	ttggccctca	ttggccctca	ttggccctca	4494

SEQ ID NO: 14	moltype = DNA	length = 4494				
FEATURE	Location/Qualifiers					
misc_feature	1..4494	note = Synthetic Construct				
source	1..4494	mol_type = other DNA				
	organism = synthetic construct					
SEQUENCE: 14						
atggacgtga	ccaagaagaa	caagcgccac	ggcaccgaag	tgaccgagag	aatcgatcacc	60
gaaaccgtga	ccaccagact	gaccgcgtt	cctccctaaag	ggggcacctc	taatggctac	120
gccaagagac	tttctctcg	cgaggcagc	agactggaaa	agcagttct	gacacacggc	180
agcagcgctt	acatcaatag	cacggctt	acaaagggcc	acggccagcac	aagcagctac	240
agaagggtctc	acagcccttc	cageacactg	cctaatacgcc	caggcagcac	cttcgagaga	300
aaagccccacg	tgaccagaca	cgccctacag	gggtctacgtca	ggggcaatag	cagccctcgag	360
taccccccaga	aagagtgtcc	cagcagcagc	accaggaggca	gaaggccagac	acgggaaagc	420
gagatcagag	tgccgtgtca	gtctgtca	ccttagccacca	gtatggacca	gtctggacgc	480
gtgaagggc	tgctgaaggg	aagcagatcc	ggccagctgt	cccttaccag	aaacagcagc	540
aacacccttc	ctatttctaa	gaaaggccac	gtcgagacta	agatcgatc	agccagcagc	600
cagagcgtgt	ccggccacata	cgatggccaca	atcctggacg	ccaaacctgg	ttggcatatgt	660
ttggccctta	cactgttcc	ttggccacgt	ttggccacgt	ttggccacgt	ttggccacgt	720
cagagcagca	ccctgtgtaa	caccaatggc	tactctggcc	ttggccacgt	ttggccacgt	780
aacaatattgg	ccagctgtca	ccctctggcc	ttggccacgt	ttggccacgt	ttggccacgt	840
tttggatgc	agaacaatct	ggccccttgc	ttggccacgt	ttggccacgt	ttggccacgt	900
accacgacat	cttacggcg	ttggccacgt	ttggccacgt	ttggccacgt	ttggccacgt	960
ggcggttagca	catctggccgc	ttggccacgt	ttggccacgt	ttggccacgt	ttggccacgt	1020
gactgtcaatgt	ttctgtatcc	ttggccacgt	ttggccacgt	ttggccacgt	ttggccacgt	1080
ctgtatcatgt	ccaaaggacat	ccctctggcc	ttggccacgt	ttggccacgt	ttggccacgt	1140
acaaggcttta	ttggccacgt	ttggccacgt	ttggccacgt	ttggccacgt	ttggccacgt	1200
ggactgttgg	ccggggccaa	ttggccacgt	ttggccacgt	ttggccacgt	ttggccacgt	1260
accggccaca	ttccacagacta	ttggccacgt	ttggccacgt	ttggccacgt	ttggccacgt	1320
gttggagggt	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1380
tgttgcgt	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1440
ctgtctgttttgc	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1500
gagctgttgg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1560
aaggacacatgt	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1620
cacagcgaca	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1680
gagaacggca	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1740
aaaggcgaca	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1800
cctcaggggcc	ttaaaggccca	ttggccggagg	ttggccggagg	ttggccggagg	ttggccggagg	1860

-continued

ggacagagag	gtagagaagg	accatgggc	cctagaggcg	aagctggacc	tcctggatct	1920
ggcggagaaag	gogaaaaaggg	agctgcggc	gaaccaggac	cacatggacc	tccagggttt	1980
cccgatctg	tggccctaa	gggatctt	gttccccctg	gaccacaagg	cccacctgg	2040
cctgttgac	tccaaggact	gagggcgaa	gttggactgc	caggcgtcaa	gggcgacaa	2100
ggtccaatgg	gaccacctgg	tccaaagggc	gatcaggcg	aaaaggacc	tagaggact	2160
acaggcggc	ccggaaatgag	aggacttctt	gttgctgtt	gagagccgg	cgctaaagg	2220
gctatgggc	ctgcaggccc	cgacggacat	caaggacat	ggggagaaca	gggcctgacc	2280
ggcatggct	gaatttagagg	tcctcttgg	ccttccggc	atccctggca	accaggatt	2340
actggaccc	agggaccaca	gggactgc	ggaacaccag	gcagacctg	aatcaaagg	2400
gaacacctgg	ctccggcga	gatttgaca	tctgagggg	gtccatgt	gaccgtgc	2460
gttccacctg	ggcccttccagg	cgccatgg	cctccggc	caccagg	tccaggacca	2520
gcggccagg	caggactcc	agggccat	aaagtgt	acctgcagg	gcctcctgg	2580
ccgcctggc	caagagggcc	accagggca	tctattcc	gaccaccag	tccttagagg	2640
cetccggcc	aaaggatggc	aggctctt	ggggcacct	gcagctt	gagaacacag	2700
gagacatcc	tgagcggccc	tcaggacat	cctggacc	caggacat	aggcgat	2760
ggacctccg	gaccgagagg	acatcaggc	gaacaaggac	tgccctgc	tagcacaag	2820
gcaagctcta	gttccggct	caatctgc	ggtccaccag	ggccaccagg	acctaagg	2880
cctaaaggcg	ataaaggcga	tccaggc	ccagggc	tgggtat	ttctggacca	2940
tctgaaggcg	gtctccatcg	cactatgt	gtgtcagg	caccgg	acctggacc	3000
ccaggaccat	ctggatctat	ctctagctt	ggccaa	tccagc	catcagc	3060
tacatgcagt	ctgacagcat	ccggcttac	ctgtctgg	ttcaagg	accggac	3120
ccgggcctc	ctggacctgt	tacaaca	acccggc	ggttgcact	cagcgac	3180
gectctcacg	ttgtgttctt	tctgaga	acgggtat	ggtgttcc	tttcgc	3240
agcatcagt	ccgaggacat	ctggcag	ctgcagg	atgacgt	cgact	3300
agacagtatc	tgatggggcc	cagaggcc	cctgg	caggcgt	cggagat	3360
tctctgtga	gtctggatta	cgccgag	tgac	tgact	catgac	3420
ccoggcatct	ccatggggtt	gcctggac	ccagg	tcggac	ttggacaa	3480
tacgaggaa	tgctgttcc	gtctgggg	agcagg	tttgc	ggggac	3540
ggggccaccag	gtccac	tatccctg	aatgtgt	ctagcat	cgtggaa	3600
ctgagcagct	actgcac	acccggc	agcttatt	caggacc	aggccg	3660
gggcctcctg	gtctccgggg	accctgg	gttagcgg	cacttgc	atatgc	3720
gagaacacg	actcttc	aagg	atcttctt	tgcact	cgatgt	3780
agtttatc	ttggccgc	ttggctt	ggac	cttgc	cgattct	3840
ctgctgagca	catatgc	ccac	ggcag	cttgc	ctcc	3900
gtcgccag	gtctccatgt	cagc	ctggact	atgtca	gttgc	3960
cttggagat	tgccgtt	tggagaa	gtctgg	gttgc	tttgc	4020
attggacc	gggggttata	tggcgtt	gtctgg	gttgc	tttgc	4080
ggactgtgg	ggccgat	tgctgg	ctggact	atgac	gttgc	4140
tccgagac	tgcaaa	gggggtt	caagg	cttgc	tttgc	4200
ccaggac	ctggcc	ggggacc	gtaatc	gttgc	tttgc	4260
aacgtgac	ccgacc	ggactt	caac	gttgc	tttgc	4320
ggccaaaagg	gcgaaatgg	tacac	ccgaaagg	accgagg	tttgc	4380
cctggacatc	ccggcc	aggc	tttgc	ggac	tttgc	4440
gaccaagtct	acggcg	ggaga	agatcc	ccgt	tttgc	4494

SEQ ID NO: 15 moltype = AA length = 1464
 FEATURE Location/Qualifiers
 source 1..1464
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 15

```

MFSFVDLRL LLLAATALLT HQQEQQVEQ QDEDIPPI TC VQNGLRYHDR DVWKPEPCRI 60
CVCDNGKVLC DDVICDETKN CPGAEVPGE CCPVCPDGSE SPTDQETTVG EGPKGDTGPR 120
GPRGPAGPPG RDGIPQPGPL PGPPGPPGPP GPPGLGGNFNA PQLSYGYDEK STGGISVPGP 180
MGSFSGPRLP CGPGAPGPQG FQGPPGEPGE PGASGPMGPR GPPGPPGKNG DDGEAKGPGR 240
PGERGPPGQ PGERGPPGQ TAGLPGKGRF SGLDGAKGDA GPAGPKGEPG SPGENGAPGQ 300
MGPRLGLPER GRPGAPGPAG ARGNNDGATGA AGPPGPTGPA GPPGFPGAVG AKGEAGPQGP 360
RGSEGPGQVR GEPGPPGPAG AAGPAGNPGA DGQPGAKGAN GAPGIAGAPG FPGARGPSGP 420
QGPGGPPGPK CNSGEPGPAG SKGDTGAKGE PGPVGVQGP GPAGEEGKRG ARGEPEGTGL 480
PGPPGERGPP GSRGFPGADG VAGKPGPAGE PGATGFPGP GSPGAKGEPG EAFLPGAKGL 540
TGSPPGSPGD CKTPGPPGPAG QDGRPGPQ PGARGQAVM GPPGPKGAAG EPGKAGERGV 600
PGPPGAVGPA GKDGEEAGAQG PPGPAGPAGE RGEQGPAGSP GFQGLPGPAG PPGEAKPGE 660
QGVPGLQGLP GPGSGLPER FPGRGVPQGP PGPAGPRGAN GAPGNDGAKG DAGAPGAPGS 720
QGAPGLQGMP GERGAAGLPG PKGDRGDAGP KGADGSPGK GVGRGLTGPIG PPGPAGAPGD 780
KGESGPGSPA GPTGARGAP DRGEPPGPAG AGFAGPPGPAG QGPGAKGEPG DAGAKGDAGP 840
PGPAGPAGP GPIGNVGAPG AKGARGSAGP PGATGFPGA GRVGPQPSG NAGPPGPAGP 900
AGKEGGKGP GETGPAGRPG EVGPPGPPGP AGEKGSAGP GPAGAPGTPG PQGIAGQRGV 960
VGLPGQRGER GPPGLGPAG EPGKQGPAGA SGERGPPGP GPPGLAGPPG ESGREGAPGA 1020
EGSPGRDGP GAKGDRGETC PAGPPGAGPA PGAPGPVGPKA GKSGDRGETG PAGPAGPVGP 1080
VGARGPAGPQ GPRGDKGETG EQGDRGKIGH RGFSGLQGP GPPGSPGEQG PSGASGPAGP 1140
RGPGPSAGAP GKDGLNGLPG PIGPPGPRGR TGADAGPVGPP GPPGPPGPAG PPSAGFDPSF 1200
LPQPQPKAH DGGRYYYRADD ANVVRDRDLE VDTTLKSLSQ QIENIRSPG SRKNPARTCR 1260
DLKMCHSDWK SGEYWIDPNQ GCNLDAIKVF CNMETGETCV YPTQPSVAQK NWYISKNPKD 1320
KRHWWFGESM TDGFQFEYGG QGSDPADVAI QLTFLRMLST EASQNITYHC KNSVAYMDQQ 1380
TGNLKKALL QGSNEIEIRA EGNSRFTYSV TVDGCTSHTG AWGKTVIEYK TTKTSRLPII 1440
DVAPLDVGAP DQEFGFDVGP VCFL 1464

```

-continued

```

SEQ ID NO: 16          moltype = AA  length = 1366
FEATURE             Location/Qualifiers
source              1..1366
                     mol_type = protein
                     organism = Homo sapiens
SEQUENCE: 16
MSLFVDTRTL LLLAVTLCIA TCQSLQEETV RKGPGDRGP RGERGPPGPP GRDGEDGPTG 60
PPGPPGPPGP PGLGGNFAAQ YDGKGVGLGP GPMGLMGPGRG PPGAAGAPGP QGFQGPAGEP 120
GEPGQTPGAG ARGPAGPPGK AGEDGHPGKP GRPGERGVVG PQGARGFPGT PGLPFGKGR 180
GHNGLDGLKG QPGAPGVKGE PGAPCNGTP QGTGARGLPG ERGRVAGAPG AGARGSDGSV 240
GPVGPAGPIG SAGPPGFPGA PGKGEIGAV GNAGPAGPAG PRGEVGLPGL SGPVGPPGNP 300
GANGLTGAKG AAGLPGVAGA PGLPGPRGIP GPVGAAGATG ARGLVGEPPG AGSKGESGNK 360
GEPGSAGPQG PPGPSGEEEGP RGPNGEAGSA GPPGPPGLRG SPGSRGLPGA DGRAGVMGPP 420
GSRGASGPAG VRGPNGDAGR PGEPELMGRP GLPGSPGNIG PAGKEGPVGL PGIDGRPGLP 480
GPAGARGEPEP NIGFPGPKGP TGDPGKNGDK GHAGLAGARG APGPDGNNGA QGPPGPPQGVQ 540
GGKGEQGPPG PPGFQGLPGRP SGPAGEVGKGP GERGLHGEFG LPGPAGPRGE RGPPGESGAA 600
GPTGPPIGSRG PSGPVGPDGN KGEPGVVGAV GTAGPSPGSG LPGERAAGI PGGKGEKGEP 660
GLRGEIGNPG RDGARGAPGA VGAPGPGATG DRGEAGAAAG PAGPGRPRS PGERGEVGP 720
GPNGFAGPAG AAGQPGAKGE RGAKGPKGEN VGVGPTGPVG AAGPAGPNGP PGPAGSRGDG 780
GPPGMGTGFP AAGRTPGPPG SGISGPPGPP GPAGKEGLRG PRGDQGPVGR TGEVGAvgpp 840
GFAGEKGKPSG EAGTAGPPGT PGPCQGLLGPAG GILGLPGRSG ERGLPGVAGA VGECPGLGIA 900
GPPGARGPPG AVGSPGVNGA PGEAGRDGPN GNDGPPGRDQ QPGHKGERGY PGNIIPVGAA 960
GAPGPHGPPV PAGKHGNRGE TGPSGPVGPAGA GAVGPPGPPSG PQGIRGDKGE PGEKGPRLP 1020
GLKGHNGLQG LPGIAGHHGD QGAPGSVGPAGA GPPGPPGPPSG PAGKDRTGH PGTVPAGIR 1080
GPQGHQGPAG PPGPPGPPGP PGVGSQGYDF DYQDFYRAD QPRAPSRLP KDYEVDATLK 1140
SINNQIETLL TPEGSRKNPA RTCRDLRLSH PEWSSGYWI DPNQGCTMDA IKVYCDSTG 1200
ETCIRAOOPEN IPAKNWYRSS KDKKHWLSE TINAGSQFEY NVEGVTSKEM ATQLAFMRL 1260
ANYASQNITY HCKNSIAYMD EETGNLKKAV ILQGSNDVEL VAEGNSRFTY TVLVDGCSKK 1320
TNEWGKTIIE YTKTNKPSRLP FLDIAPLDIG GADQEFFVDI GPVCFK 1366

SEQ ID NO: 17          moltype = AA  length = 1466
FEATURE             Location/Qualifiers
source              1..1466
                     mol_type = protein
                     organism = Homo sapiens
SEQUENCE: 17
MSMFVQKGSW LLLALLHPTI ILAQQEAEVG GCSHLGQSVA DRDVWKPEPC QICVCDGSV 60
LCDDIICDDQ ELDCPNNEPI FGECAVCPQ PPTAPTRPPN QGQGPQPKGD PGPPGIPGRN 120
GDPGIPQPGP SPGSPGPPGI CESCSTGPQN YSPQYDSYDV KSGVAVGGLA GYPGPAGPPG 180
PPGPPGTSGH PGSPGSPGQY GPPGEPPQAG PSPPGPPGPA IGPSGPAGKD GESGRPGRPG 240
ERGLPGPPGI KGPAIGPGFP GMKGRHRFDG RNKEKGETGA PGLKGENGELP GENGAAPGPMG 300
PRGAPGBERGR PGLPGAAAGR GNDGARGSDG QPGPPGPGRP AGFPGSPGAK GEVGPAGSPG 360
SNGAPQGRGE PGQPGHAGAQ GPPGPPGNG AGIPGAPGLM GARGPPGPAG 420
ANGAPGLRRG AGEPKNGKA GEPGPRGERG EAIGPVGPGA KGEDGKDGP GEPGANGLP 480
AAGERGAPGF RGPAGPNIGIP GEKGPGAGERG APGPAGPRGA AGEPGRDGPV GGPGRMRGMPG 540
SPGPGSDGK PGPPGSGQES GRPGPPGPSS PRQGPQVMGF PGPKGNDGAP GKNGERGGPG 600
GPGPQGPQGP NGETGPQGP GPTGPQGDKG DTGPPGPQGL QGLPGTGGPP GENKGKPEPG 660
PKGDAGAPGA PGKGDAGAP GERGPGLAG APGLRGAGP PGPEGGKGAAG GPPGPAGAAG 720
TPGLQGMPGE RGGLGSPGPK GDKGEPPGPG ADGVPKGDP RGPTGPIGPP GPAGQPGDKG 780
EGGAPGLPGI AGPRGPSPGER GETGPQGPAG FPGAPQGNGE PGKGGERGP GEKGEQGPPG 840
VAGPPGPGVK AGPPGPQGVK GERGSPGPGV AAGFPGARGL PGPPGSNGNP GPPGPSSPG 900
KDGPGPAGN TGAPGSPGVG GPKGDAGQPG EKGSQGAQGP PGAPGPLGIA GITGARGLAG 960
PPGMPGPRGS PGPQGVKGS GKPQGNGLS ERGPPGPQGL PGLAGTAGEP GRDGNPGSDG 1020
LPRGDGSPGG KGDRGENGP GAPGPAGHPG PPGPVGPAKG SGDGRGESGPA GPAGAPGPAG 1080
SRGAPGPQGP RGDKGETGER GAAGIKGHRP FPQGPQGPAGS PGPAQQGAI GSPGPAGPRG 1140
PVPGSPGPPG DGTSGHGPPI GPPGPQGPNRG ERGSEGPSPH PGQGPQGPAGP GAPGPCCGGV 1200
GAAAIAIGIGG EKAGGFAPYY GDEPMDFKIN TDEIMTSLSK VNGQIESLIS PDGSRKNP 1260
NCRDLKFCHP ELKSGEYWD PNQGCKLDAI KVPCNMETGE TCISANPLNV PRKHWWTDSS 1320
AEKKHVWFGE SMDGGFQFSY GNPELPEDVL DVQLAFLRLL SSRASQNITY HCKNSIAYMD 1380
QASGNVKKAL KLMGSNEGEF KAEGNSKFTY TVLEDGCTKH TGEWSKTVFE YRTRKAVRLP 1440
IVDIAPYDIG GPDQEFGVDV GPVCFL 1466

SEQ ID NO: 18          moltype = AA  length = 1669
FEATURE             Location/Qualifiers
source              1..1669
                     mol_type = protein
                     organism = Homo sapiens
SEQUENCE: 18
MGPRLSVWL LLLPAALLHE EHSRAAAKGG CAGSGCGKCD CHGVKGKGE RGLPGLQGVI 60
GFFGMQGPQEG PQGPPQKGD TGEPGLPGTK GTRGPPGASG YPGNPGPLGI PGQDGPQGP 120
GIPGCNGTK ERGPLGPGLP PGFAGNPQPP GLPGMKGDPG EILGHVPGML LKGERGFPGI 180
PGTPGPQGPCLP GLQGPVGPQPG FTGPPGPQGP PGPPGEKQGM GLSFQGPQKGD KGDQGVSGPP 240
GVPGQAQVQE KGDFATKGEK GQKGEPEFQG MPGVGKGEKGP GKPQPRGKPG KGDQKGEKGS 300
PGFPGEPPGP GLIGRQGPQG EKGEAGPPGP PGIVIGTGP GEKGERGYPG TPGPRGEPPG 360
KGFPGLPQGP GPPGPVPGQ AGAPGFPGER GEKGDRGFPQG TSLPGPSPGRD GLPQGPQGP 420

```

- continued

PPGQPGYNTNG	IVECQPGPPG	DQGPPGIPGQ	PGFIGEIGEK	GQKGESCLIC	DIDGYRGPPG	480
PQGPPGIEGF	PQGPAKGDR	GLPGRDGVAG	VPGPQGTPGL	IQGPGAKGEF	GEFYFDLRLK	540
GDKGDPGPFG	QPGMPGRAGS	PGRDGHGPGLP	GPKGSPGSVG	LKGERGPPGG	VGFPGRSGDT	600
GPPGPPGYGP	AGPIGDKGQA	GFPGPGPGSPG	LPGPKGEPK	IVPLPGPPGA	EGLPGSPGFP	660
GPQGDRGFPG	TGPRPGLPGE	KGAVGQPGIG	FPGPPGPKGV	DGLPGDMGPP	GTPGRGPFNG	720
LPGNPGVQGQ	KGEPEVGLPG	LKGLPGLPGI	PCTPGEKGSII	GVPGVPGEHG	AIGPPGLQGI	780
RGEPEPGPPGLP	GSGVSPGVPG	IGPPGARGPP	GGQGPPGLSG	PPGIKGEKGF	PGFPGDMPG	840
PKGDKGQAGL	PGITQOSGLP	GLPGQQGAPG	IPGFPGSKGE	MGVMGTPGQP	GSPGPVGAPG	900
LPGEKGHDHG	PGSSGPRGDP	GLKGDKGDV	LPGKPGSMDK	VDMGSMKGQK	GQGEKEQIG	960
PGEKGSRGD	PCTPGEVPGKD	QGAGPQGPQG	PKGDPGIGST	PGAPGLPGPK	GSGVGMGLPG	1020
TPGEKGVPGI	PQPGQSPGLKG	CDKAKGEKG	QAGPPGIGIP	GLRGEKGDQG	IAGPPGSPGE	1080
KGEKGSGTIP	CMPGSPGLKG	SPGSVGVPGS	PGLPGEKDK	GLPGLDGIPG	VKEAGEGLPGT	1140
PGPTGPAGQK	GEPGSDGIPG	SAGEKEPGL	PGRGFPGPFG	AKGDKGSKGE	VGFPGLAGSP	1200
GIPGSKGEQG	FMGPPGPQGQ	PGLPGSPGHA	TEGPKGDGPQ	QGQGPPGLPGLP	GPMGPGLPG	1260
IDGVKGDKGN	PWPGQAVPGV	GPKGDGPQG	MPIGIGGSPGI	TGSKGDGMGPP	GVPGFQGPKG	1320
LPGLOQIGKD	QGDQGVPGA	GLPQGPGLPPG	PYDIKGPB	LPGEPGPGL	KGLQGLPGLPK	1380
GQQGVTGLVG	IPGPPGIPGF	DGAPGQKJEM	GPAGPTGPRG	FPGPPGPGL	PGSMGPPGTP	1440
SVDHGFLVTR	HSQTIDDPQC	PSGTKILYHG	YSLLYVQCN	RAHGQDLGTA	GSCLRKFSTM	1500
PFLFCNCINNV	CNFASRNNDY	YWLSTPEMP	MSMAPITGEN	IRPFISRCAV	CEAPAMVMAV	1560
HSQTIQIPPC	PSGWSSLWIG	YSFVMMHTSAG	AEGSGQALAS	PGSCLEEFRS	APFIECHGRG	1620
TCNYYANAYS	FWLATIERSE	MFKKPTPSTL	KAGELRTHVS	RCQVCMRRRT		1669

SEQ ID NO: 19	moltype = AA	length = 1028				
FEATURE	Location/Qualifiers					
source	1..1028					
	mol_type = protein					
	organism = Homo sapiens					
SEQUENCE: 19						
MRAARALLPL	LLQACWTAAQ	DEPETPRAVA	FQDCPVDLFF	VLDTSSEVAL	RLKPYGALVD	60
KVKSFTKRFI	DNLDRYYRC	DRNLVWNAGA	LHYSDEVEII	QGLTRMPGGR	DALKSSVDAV	120
KYFGKGTYTD	CAIKKGELEQ	LVGGSKL	KYLIVVTDHG	PLEGYKEPCG	GLEDAVNEAK	180
HILGVKVFSVA	ITPDHLEPRL	SIIATDHTYR	RNPTAADWQG	SRDAEEAISQ	TIDTIVDMIK	240
NNVEQVCCSF	ECQPARPGPPG	LRGDPGFEGE	RGKPGLPGEK	GEAGDPGRPG	DLGPVGYQGM	300
KGEKGSRGEK	CSRGPKGYKG	EKGKRGIDGV	DGVKGEMLGP	GLPGCKGSPG	FDGIQGPGP	360
KDGPQAFGLK	GEKGEPEGADG	EAGRPGSSGP	SGDQGQREGP	GPPGDPGEAG	DEGNPGPDGA	420
PGERGGPGER	CPRGTPGTRG	PRDPGAEAGP	QGDQGREGP	GVPGPGEAG	PIGPKGYRGD	480
EGPPGSEGAR	GAPGPAGPPG	DPLGLMGERGE	DGPAGNGTEG	FPGFPGPYGN	RGAPGINGTK	540
GYPGLKGDEG	EAGDPGDDNN	DIAPRGVKGA	KGYRGPEGPQ	GPPGHQGPQ	PDECEILDII	600
MKMCSCCECK	CGPIDLFLV	DSSEISGLQN	FEIAKDFVVK	VIDRLSRDEL	VKFEPPGQSYA	660
GUUVQYSHSQM	QEVHVSLSRSPS	IRNVQELKEA	IQLSLOWMAGG	TFTGEALQYT	RDQLLPPSPN	720
NRIALVITDQ	RSQDTQRDTTP	LNVLCSPGIQ	VVSVGKIDVF	DFIPGSDQLN	VISCGQLAPS	780
QGRPGLSLVK	ENYAEELLED	FLKNVTAQIC	IDKKCPDYTC	PITFSSPADI	TILLDGSASV	840
GSHNFDTTKR	FAKRLAERFL	TAGRTDPAHD	VRVAVVQYSSG	TGQQRPERAS	LQFLQNYTAL	900
ASAVIDAMDFI	NADATDVNDAL	GYVTRFYREA	SSGAACKRLL	LFSDGNSQGA	TPAAIEKAVQ	960
EAQRAGIEIF	VVVVGRQVNE	PHIRVLVTGK	TAEYDVAYGE	SHLFRVPSYQ	ALLRGVFHQT	1020
VSRKVALG						1028

SEQ ID NO: 20	moltype = AA	length = 2944				
FEATURE	Location/Qualifiers					
source	1..2944					
	mol_type = protein					
	organism = Homo sapiens					
SEQUENCE: 20						
MTLRLLLAAL	CAGILAEAPR	VRAQHRRERVT	CTRLYAADV	FLLDGSSSI	RSNFREVRSF	60
LEGLVLPFSG	AASAQGVRFA	TVQYSSDDPRT	EFLGLDALGSG	GDVIRAIREL	SYKGGNTRTG	120
AAILHVADHV	FLPQLARPGV	PVKCILITDG	KSQDLVTDAA	QRLKGQGVKL	FAVGIKNADP	180
EELKRVASQ	TSDFFFFVN	FSILRLLPL	VSRRVCTTAG	GVPVTRRPDD	STSAPRDLVL	240
SEPSSQSLSRV	QWTAASGPVT	GYKVQYTPLT	GLGQPLPSER	QEVNVPAGET	SVRLRGLRPL	300
TEYQVTIAL	YANISGEAVS	GTARTTALEG	PELTINQNTA	HSLLVAWRSV	PGATGYRVWT	360
RVLSGGTQQ	QELGPQGGSV	LLRDLEPGTD	YEVTVSTLFG	RSGVGPATSLM	ARTDASVEQT	420
LRPVILGPTS	ILLSWNLVP	ARGYRLEWR	ETGLEPPQKV	VLPSPDVTRYQ	LDGLQPGTEY	480
RLTLYTLLEG	HEVATPATVV	PTGPELPVSP	VTDLQATEL	GQRVRVWSWP	VPGATQYRII	540
VRSTQGVERT	LVLPGSQTAF	DLDVVQAGLS	YTVRVSRARV	PREGSASVLT	VRREPETPLA	600
VPGLRVVVS	ATRVRVAVG	VPGASGPFR	WSTGSGPESS	QTLPPDSTAT	DTGLQPGTT	660
YQVAVSVLRL	REGPAAVIV	ARTDPLGPVR	TVHVTQASS	SVTITWTRVP	GATGYRVWSH	720
SAHGPEKSQL	VSGEATVAEL	DGLEPDTEYT	VHVRRAHVG	DGPPASVVR	TAPEPVGRVS	780
RLQILNASSD	VLRITWVGVT	GATAYRLAWG	RSEGGPMRHQ	ILPGNTDSAE	IRGLEGGVSY	840
SVRVTALVGD	REGTPVSI	TTPEAPPAL	GTLHVVQRGE	HSLRLRWEPV	PRAQGFLHW	900
QPEGGQEQCSR	VLGPELSSYH	LDGLEPATQY	RVRLSVLGP	GEGPSAEVTA	RTESPRVPsi	960
ELRIVVDTSID	SVTLAWTPVS	RASSYILSWR	PLRGPQEV	GSPQTLPGIS	SSQRVTGLEP	1020
GVSYIFSLTP	VLDVGRGPEA	SVTQTPVCP	GLADVVFLPH	ATQDNRAA	ATRRVLERLV	1080
LALGPLGPQA	VQVGLLSYSH	RPSPLFLPLNG	SHDLGIILQ	IRDMPYMDPS	GNNLGTA	1140
AHYRMLAPDA	PGRRQHVG	MVLLVDEPLR	GDIIFSPIRE	QASGLNVML	GMAGADPEQL	1200
RRLAPGMDSV	QTFFAVDDGP	SLDQAVSGLA	TALCQASFTT	QPRPEPCPVY	CPKGQKGEPG	1260
EMGLRGQVGP	PQDPGLPGR	GAPGPQGPPG	SATAKGERGF	PGADGRPGSP	GRAGNPCTPG	1320
APGLKGSPGL	PGPRGDPGER	GPRGPKGEPG	APGQVIGEGG	PGLPGRKGDP	GPSGPGPGR	1380

-continued

PLGDPGPRGP	PGLPGTAMKG	DKGDRGERGP	PGPGEGGIAP	GEPLGPLGP	SPGPQGPVGP	1440
PGKKGKEKGS	EDGAPGLPGD	PGSPGEQQGP	GPPGAIGPKG	DRGFPGLGE	AGEKGERGP	1500
GPAGRSRGLPG	VAGRPGAKGP	EGPPGPTGRQ	GEKGEPPRG	DPAVVGPAVA	GPKGEKGDVG	1560
PAGPRGATGV	QGERGPGLV	LPGDPGPKGD	PGDRGPIGLT	GRAGPPGDSG	PPGEKDPGR	1620
PGPPGPVGPR	GRDGEVGEKG	DEGPPGDPGL	PGKAGERGLR	GAPGVVRGPVG	EKGDQGDPGE	1680
DGRNGSPGSS	GPKGDRGEPG	PPGPPGRLVD	TGPGAREKGE	PGDRGQEGPR	GPKGDPGLPG	1740
APGERGIEGF	RGPPGPQGDP	VGVRGPAGEKG	DRGPPGLDGR	SGLDGKPGAA	GPSGPNGAAG	1800
KAGDPGRDGL	PGLRGEQGLP	GPSGPGLPG	KPGEDGKPGL	NGKNGEPPGD	GEDGRKGEKG	1860
DSGASGREGR	DGPKGERGAP	GILGPQGPPG	LPGPVGPPGQ	GFPGVPGGTG	PKGDRGETGS	1920
KGEQGLPGER	CLRGEEPVS	NVDRRLETAG	IKASALREIV	ETWDESSGS	LPVPERRRGP	1980
KGDSCGEQCPP	KGEKPTGFP	ERGLKDRGD	PGPQGPPGLA	LGERGPPGPs	GLAGEPKPG	2040
IPGLPGRAGG	VGEAGRPGF	GERGEKGERG	EQQRDGPPGL	PGTPGPPGPP	GPKVSVDEPG	2100
PGLSGEQGPP	GLKGAKGEPG	SNGDQGPKD	RGVPGIKGD	GEPPGPRQGD	NPGLPGERGM	2160
AGFEGKPLQ	CPRGPGPVPG	GHGDPGPPGA	PGLAGPAGPQ	GPSGLKCEPG	ETGPPGRGLT	2220
GPTGAVGLPG	PGPSPGLVGP	QGSPGLPGQV	GETGKPGPAG	RDGASGKDG	RGSPGVPGSP	2280
GLPVGPKG	EPGPTGAPGQ	AVVGLPGAKG	EKGAPGGLAG	DLVGEPKAGK	DRGLPGRGE	2340
KGEAGRAGEP	GDPGEDQKG	APGPKGFKD	PGVGVPGSPG	PPGPPGVKD	LGLPGLPGAP	2400
GVVGFPGQTG	PRGEMCOPGP	SGERGLAGPP	GREGIPGPLG	PPGPPGSVGP	PGASGLKGDK	2460
GDGPVGLPGP	RGERGEPEGR	GEDGRPGQEG	PRLTGPPGS	RGERGEKGDV	GSAGLKGDKG	2520
DSAVILPGP	PRGAKGDMGE	RGPGLDGD	GPRGDNGDPG	DKGSKEPGD	KGSAGLPGLR	2580
GLLGPQGQPG	AAGIPGDPGS	PGKDGVPGIR	GEKGDVGFMG	PRGLKGERGV	KGACGLDGEK	2640
GKGEAGCOPG	RPGLAGHKGE	MGEPPVPGQS	GAPGKEGLIG	PKGDRGFDGQ	PGPKGDQGEK	2700
GERGTPGIGG	FGPPGSGNDGS	AGPPGPPGS	GPRGPGLQG	QKGERGPGE	RVVGAPVPG	2760
APGERGEQGR	PGPAGPRGEK	GEAALTEDSV	RGFVRQEMSQ	HACACQOFIA	SGSRPLPSYA	2820
ADTAGSQLHA	VPVLRVSHAE	EEERVPPEDD	EYSEYSEYSV	EEYQDPEAPW	DSDDPCSLPL	2880
DEGSCTAYTL	RWYHRARTGS	TEACHPFVYG	GCGGNANRFG	TREACERRCP	PRVVQSQGTG	2940
TAQD						2944

SEQ ID NO: 21	moltype = AA	length = 1497				
FEATURE	Location/Qualifiers					
source	1..1497					
	mol_type = protein					
	organism = Homo sapiens					
SEQUENCE: 21						
MDVTKKNKR	GTEVTERIVT	ETVTTRLTS	PPKGCTSNGY	AKTASLGGS	RLEKQSLTHG	60
SSGYINSTGS	TRGHASTSSY	RRAHSPASTL	PNSPGSTFER	KTHVTRHAYE	GSSSGNSSPE	120
YPRKEFASS	TRGRSQTRES	EIRVRLQSAS	PSTRWTLEDD	VKRLLKGSRS	ASVSPTRNSS	180
NTLPPIPCKGT	VTKIVTASS	QSVSGTYD	ILDANLPShV	WSSTLPAGSS	MGTYHNMNTT	240
QSSSLLNNTNA	YSAHSVGP	NNMSPATL	HPGLSTVSS	FGMQNLLAPS	LTTLSHGT	300
TSTAYGVKNN	MPQSPAAVNT	VGTSAACTT	SVQSDLLHK	DCKFLILEKD	NTPAKKEMEL	360
LIMTKDSGKV	FTASPASIAA	TSFSTEDTLKK	EKOQAYNADS	GLKAEANGDL	KTVSTKGKTT	420
TADIHSYGS	GGGGSGGGGG	VGGAGGGPWG	PAPAWCPCGS	CCSWWKWLIG	LLLTWLLLIG	480
LIFGLITLAE	EVRKLIKARVD	ELERIRRSL	PYGDSDMIDI	KDRLQGMAPA	AGADLDKIGL	540
HSDSQEELWM	FPVKLMMEQ	ENGNLRGSPG	PKGDMGSPGP	KDRGFPGLP	GIPGPLGHPG	600
PQGPKGQKGS	VGDPMEGPM	QQRREGPML	PRGEAGPPGS	GEKGERGAAG	EPGPHPGPPV	660
PGSVPGPKGSS	GSPPGPQGPPG	PVGLQGLRGE	VGLPGVKGDK	GPMGPPGPKG	DQGEKGPRL	720
TGEPGPGRMLG	GAVGEPGAKG	AMGPAGPDGH	QGPRGEQGLT	GMPGIRGPGP	PSGDPGKPG	780
TGPQGPQGLP	PTPGRPGIKG	EPGAPGKIVT	SEGSSMLTV	GPPGPPGAMG	PPGPPGAPGP	840
AGPAGLPHQ	EVLNLQGPPG	PPGPRGPGP	SIPGPPGPRG	PPGEGLPGLP	GPPGSFLNS	900
ETFLSGPPGP	PGPPGPKGQ	GPPGPGRHQ	EQGLPGFSTS	GSSSFGLNLQ	GPPGPGPQG	960
PKGDKGDPV	PGALGIPSGP	SEGGSSSTMV	VSGPPGPGP	PGPPGSISSS	GQEIQYISE	1020
YMQSDSIRSY	LSGVQDVGPPG	PPGPPVTI	TGETFDYSEL	ASHVSYLRT	SGYGVSLFSS	1080
SISSEEDILAV	LQRDDVRQYL	RQYLMGPRGP	PGPGASGDDG	SLLSLDYAEL	SSRLISYMS	1140
SGISIGLPGP	PGPPGLPGTS	YEELLSLLRG	SEFRGIVGP	GPPGPPGIPG	NVWSSISVED	1200
LSSYLYTHAGL	SFIPGPGPGP	GPPGPGRGP	VSGALATYAA	ENSDSFRSEL	ISYLTSPDVR	1260
SFIVGPPGP	GPQGPQGDSR	LLSTDASHSR	GSSSSSHSS	VRRGSSYSS	MSTGGGGAGS	1320
LGAGGAFGEA	AGDRGPYGTG	IGPGGGYGA	AEGGMYACNG	GLLGADFAGL	LDYNELAVRV	1380
SESMQRQGLL	QGMAYTVQGP	PQGPQGPQGP	GISKVFSAYS	NVTADMDF	QTYGAIQGPP	1440
GQKGEMGTPG	PKGDRGPAGP	PGHPGPPGP	GHKGEKGDKG	DQVYAGR	RRRSIAVKP	1497

SEQ ID NO: 22	moltype = DNA	length = 148				
FEATURE	Location/Qualifiers					
misc_feature	1..148					
	note = Synthetic Construct					
source	1..148					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 22						
gaatttcgtcg	attcggttgc	agcatttaaa	gccccgttacata	actttaaaag	aaggaaaaag	60
aagggttgc	aaaagggttgt	atgttgcgtt	tataagtaca	gaccggagaa	gtacccgggt	120
cctcgattcgt	ttaatttgcgtt	aaatgttttt				148
SEQ ID NO: 23	moltype = DNA	length = 491				
FEATURE	Location/Qualifiers					
misc_feature	1..491					
	note = Synthetic Construct					

-continued

```

source          1..491
               mol_type = other DNA
               organism = synthetic construct

SEQUENCE: 23
gttattttcc accatattgc cgtttttgg caatgtgagg gcccggaaac ctggccctgt 60
cttccttgacg agcattccta ggggtcttgc cccctctcgcc aaaggaatgc aaggctctgtt 120
gaatgtcgta aaggaagcag ttccctcttgc agcttcttgc agacaaacaa cgtctgttagc 180
gacccttgc aggcaagcggg acccccccaco tggcgacagg tgccctctggc gccaaaagcc 240
acgtgtataa gatacacctg caaaggcgcc acaaccccaag tgccacgttg tgagtggat 300
agtgtgtggaa aagactaaatggccttgc aagcttgcattt aacaaggggc tgaaggatgc 360
ccagaaggta ccccatgttgc tggatctga tctggggctt cggtgcacat gctttacatg 420
tgtttagtgc aggttaaaaaa gcgtttaggc ccccgaaacc acggggacgt ggttttcctt 480
tggaaaaaacac g 491

SEQ ID NO: 24      moltype = DNA length = 75
FEATURE           Location/Qualifiers
misc_feature      1..75
note = Synthetic Construct
source            1..75
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 24
aggggccaaga ggggcagcgg cgagggcagg ggcagcctgc tgacctgcgg cgacgtggag 60
gagaaccccg gcccc 75

SEQ ID NO: 25      moltype = DNA length = 66
FEATURE           Location/Qualifiers
misc_feature      1..66
note = Synthetic Construct
source            1..66
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 25
ggaagcggag ctactaactt cagctgtctg aagcaggctg gagacgtgga ggagaacctt 60
ggacct 66

SEQ ID NO: 26      moltype = DNA length = 69
FEATURE           Location/Qualifiers
misc_feature      1..69
note = Synthetic Construct
source            1..69
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 26
ggaagcggac aagtgtactaa ttatgctctc ttgaaattgg ctggagatgt tgagagcaac 60
cctggacct 69

SEQ ID NO: 27      moltype = DNA length = 75
FEATURE           Location/Qualifiers
misc_feature      1..75
note = Synthetic Construct
source            1..75
mol_type = other DNA
organism = synthetic construct

SEQUENCE: 27
ggaagcggag tgaaacagac tttgaatttt gacttctca agttggcgaa agacgtggag 60
tccaacctt gacct 75

SEQ ID NO: 28      moltype = AA length = 25
FEATURE           Location/Qualifiers
REGION            1..25
note = Synthetic Construct
source            1..25
mol_type = protein
organism = synthetic construct

SEQUENCE: 28
RAKRGSGEGR GSLLTCGDVE ENPGP 25

SEQ ID NO: 29      moltype = AA length = 22
FEATURE           Location/Qualifiers
REGION            1..22
note = Synthetic Construct
source            1..22
mol_type = protein
organism = synthetic construct

SEQUENCE: 29

```

-continued

GSGATNFSL KQAGDV EENP GP	22
SEQ ID NO: 30 moltype = AA length = 23 FEATURE Location/Qualifiers REGION 1..23 note = Synthetic Construct source 1..23 mol_type = protein organism = synthetic construct	
SEQUENCE: 30 GSGQCTNYAL LKLAGDV ESN PGP	23
SEQ ID NO: 31 moltype = AA length = 25 FEATURE Location/Qualifiers REGION 1..25 note = Synthetic Construct source 1..25 mol_type = protein organism = synthetic construct	
SEQUENCE: 31 GSGVKQTLNF DLLKLAGDVE SNP PGP	25
SEQ ID NO: 32 moltype = DNA length = 8565 FEATURE Location/Qualifiers misc_feature 1..8565 note = Synthetic Construct source 1..8565 mol_type = other DNA organism = synthetic construct	
SEQUENCE: 32 atgttcagct tcgtggaccc gagaactgctg ctgtccctgg ctgctcacagc cctgtcgaca 60 cacggacaag aggaaggcca ggtcgaaggaa caggacgagg acatccccc tatcacctgt 120 gtgcagaaacg gcctgagata ccacgaccgg gatgtgtgg aagcccgagcc ttgcagaatc 180 tgcgtgtgcg acaaattggca ggtgtgtgc gacgacgtga tctgcgacga gacaaagaat 240 tgcctggccg cggaaatggcc tgaggggccg tgggtgtccgg tggccctgg tgccagcgg 300 agccccacag atcaagagac aacaggccgtg aaaggccccca agggcgatac aggacataga 360 ggtcctagag gacctggccgg acctctggc agagatggaa ttctggaca gcctggactg 420 cccgaggacc acggacccctcc agggcctcca ggtccaccag gactcgagg aaattttgcc 480 ccacagctga gotacggcta cgacggagaa agcacaggcc gcatctctgt gcctggaccc 540 atggggaccc ttggcccaag aggacttccctt ggtccctccgg gtgcctccgg accctcaggga 600 tttcaaggac caccaggcga acctggcga ccaggcgcta tggtcccaat gggaccaaga 660 ggccctctgg ggccaccagg caaaaatggc gacgatggcg aagccggaaa gcctggaaagg 720 cttggcggaaa gaggccggccg aggacgcga ggccgttagag ggttgcctgg aacttgagg 780 ctgcctggcc tgaaggggcca gagggcttt tctggactgg atggcgctaa gggcgacgct 840 ggaccaggac gacctaagg cgacgctgg tctccctggcc agaatggtc acctggacag 900 atgggtccca gaggattggc aggccggaga ggttagacctg gcgtccggg accagccgg 960 gtcagggaaa atgatggcgc aacaggctg gtcggccctt ctggaccaac cggaccagct 1020 ggccacccctg gatttccagg cgctgtgg acaaaaggccg aacggccccc acaaggaccc 1080 aggggatcttca aaggctctca ggccgtttaa ggccggccg ggccacccctgg gcctggccgt 1140 gcagctggac ctgctggaaa cctctggctg gatggacagg cagggtccca aggtgctaat 1200 ggcgccccctg gaatttggccg cgctccaggat tttccctggcc caaggaggcc atctggaccc 1260 caaggccccag ggccacccctcc gggtccttaag ggaatagccg gagggccagg cgctctgg 1320 agttaaaggccg atactggcgc aaaaaggccaa cccggacccctgg gaggatgtca aggacccct 1380 ggaccaggctg gcgaggaaagg caaaaggaggc gctaggggag aaccaggacc aacagggtctc 1440 ctctggccac ctggcggagcc cggggacccctt ggttgcctgg cgcacatggc 1500 gtggccggac caaaaggacc tgcaggccgaa agggggatcc cagggtctgc agggccctaag 1560 ggttctccag gcgaggctgg cagacccggc gaaatggcc tcccagggtc taaggactg 1620 acaggctcac caggatctcc cggaccaggac gggaaaacag gaccccttgg accggccagg 1680 caggatgtta gacccggctcc tcctggaccc cctggtccaa gaggacaagc tggcgatgt 1740 ggctttctgg gacccaaagg tgcaggccggaa gaaacctggaa aaggaggccgaa gaggggagtt 1800 cccgacccctt cagggtctgt tggaccccttcc gggaaaaggatc ggcggatctgg tgcacaaagg 1860 cctccaggccg cagccggacc agccggccgg agggccgaaac aaggaccagg cggatctcca 1920 ggatttcagg gactggcagg gcctgtggc cccgttggcc aggcaggggaa gccaggccaa 1980 cagggtgttc ctggccatctt tggccatccatggtccatggccatccatggccatccatggccat 2040 tttccctggccg aaaaaggccgt tggccatccatggccatccatggccatccatggccatccatggccat 2100 ggtgccccccg gcaatgtacgg tgcttaaggcc gggcaacccggg gaggtgtccatggccatccatggccat 2160 caaggccgcac ctggacttca gggaaatggcc tggccaaacccggg gaggtgtccatggccatccatggccat 2220 ccaaaaggccg ataggggaga tgctggccatccatggccatccatggccatccatggccatccatggccat 2280 ggcgcaagat ggcgtacagg cccaaatggcc cctccggccatccatggccatccatggccatccatggccat 2340 aaggggcaat ctggccatccatggccatccatggccatccatggccatccatggccatccatggccat 2400 gaccgggggg accctggccatccatggccatccatggccatccatggccatccatggccatccatggccat 2460 ggtcacccctt gtgcttaaggcc gggccatccatggccatccatggccatccatggccatccatggccat 2520 cctggaccgg cccggacccctcc tggccatccatggccatccatggccatccatggccatccatggccat 2580 gccaaggccg caagaggatc tggccatccatggccatccatggccatccatggccatccatggccat 2640 ggaagagatgg tggccatccatggccatccatggccatccatggccatccatggccatccatggccat 2700 ggccggccaaag aagggtggaaa aggcccttagg ggcggaaaatccatggccatccatggccatccatggccat 2760	

-continued

gaagtggccc	ctccaggacc	tccggggct	gccggcgaaa	aaggatctcc	aggcgcgat	2820
ggacccgcag	gogctcccg	aacacagggt	ccacaggaa	ttgtctggaca	aaggggagtt	2880
gtcggcgtgc	caggacagag	gggagagaga	ggtttccag	gactccctgg	gccaaagcgga	2940
gaacctggca	aacaggggacc	atcttgtcc	agcgagaga	gagggccacc	aggaccaatg	3000
ggtcctccag	gattggcagg	gectctggc	aatctgtta	gagaagggtc	tccaggcgcc	3060
gagggatctc	ctggacgtg	tggttctct	ggcgccaagg	gcatagagg	cgaaaacaggc	3120
ccagctggac	ctccaggcgc	acccggcgct	ccaggcgac	caggactgt	tgccctgtct	3180
gaaaatctt	gogacagagg	cgaaaactgg	cccgccagg	cagccggacc	tgttggacct	3240
gtgggtgtct	gaggacccgc	tggaccacaa	ggtccttag	gogacaagg	cgaaaacaggc	3300
gaccaaggcc	acagaggcat	caaggacac	agaggattca	cgccgactca	gggaccacca	3360
ggcccgccgt	gaaatcccg	cgacaggaa	ccaaaggcg	etagggtcc	cgccggaccc	3420
agaggaccac	ctgggtctgc	tggtgccacc	ggaaaggac	gactgtatgg	gtccccggga	3480
cctattgggc	cacccggacc	tagaggaaga	acaggcgac	caggaccat	tggaccaccc	3540
ggggccaccc	gaccggctgg	tcctcttcg	cctcccttc	ccggattcga	cttcagcttc	3600
ctgcctcage	ctccctaaaga	gaaggccat	gacggcgcc	gatatacag	agccgacgac	3660
gccaacgtc	tccggggacag	agatctggaa	gtggacacca	cactgaatgc	cctgttccag	3720
cagatcgaga	acatcagaag	ccccgaggc	agcagaaaga	accctgcac	aacctgtcgg	3780
gacctgaga	tgtggccac	cgatggaa	tctggcgagt	actggatcga	ccccaaacccag	3840
ggctgcaacc	ttggatggccat	caagggtt	tgcaacatgg	aaaccggcga	gatcatgcgt	3900
taccccaac	acccatctgt	ggctcagaag	aactgttaca	tcaacaaaggaa	ccccaaaggac	3960
aagccggcac	tttgggtcgg	cgagagcat	accgtatgc	tccagtttg	gtatggcgcc	4020
caggggcttc	accctggcc	tgttgctat	caagtcac	tcctgcggct	gatgtctaca	4080
gaggccagcc	agaacatcac	cttcaactgc	aagaacag	tggcttacat	ggatcgacg	4140
accggcaacc	tgaagaaggc	actgtgtt	cgaggcgac	acgagatcga	gtatggcgcc	4200
gagggcaaca	gcccgttac	ctacagcg	acagtggatg	gtgcacccag	ccatccaggg	4260
gcttggggca	agaccgtgt	cgatgtac	accaccaaa	ccagcgacat	gccccatcatc	4320
atgtgtggcc	ctctgtatgt	tggggcac	gtcaagatg	togggttca	tgtggggccca	4380
gtgtgtctcc	tggggcccaa	ggggggcago	ggcgaggggca	ggggcagct	gtctgaccctc	4440
ggcgacgtgg	aggagaaccc	cgccccctcg	agttctgtgg	acacccaaac	actgtgtctg	4500
ctggccgtga	cactgtgt	ggccacttgc	caagggctgc	aagggaaac	agtgcggaaa	4560
ggacccgtccg	cgatgtgggg	acccatgttgc	gaaagggatgg	ctccctggacc	tcctgttgc	4620
atggcggagg	atggacccat	aggacccat	gttccaccc	gaccccttcc	gctcttggcc	4680
cttggaggaa	attttggccgc	tcagatcgt	ggcaaaaggcg	tcggacttgg	cccttgaccc	4740
atgggactta	tggggccaa	aggaccacca	ggtgctcg	gctgtccagg	accacaagga	4800
tttcaaggac	cagctggccg	gcctggccaa	cctggacaaa	cagggttgc	tggtgctaga	4860
ggaccaggcg	ggccaccc	aaaaggctgg	gaagatggc	acccttggaa	gcctgttgc	4920
ccccggcaaa	gggggttgcgt	tggaccc	ggggccac	gttttcccttgg	aacaccccttgc	4980
ctgcctgtgt	tcaaggccat	cagaggcc	aatggccctgg	acggactgaa	aggacaaccc	5040
gtgtgtcttc	gggtgttgc	cgaaacccgg	gcacccggcc	aaaatggccac	accaggacaa	5100
accggcgc	gaggacttcc	tggcg	gagaagatgg	ggccccccagg	tccaggacgg	5160
gcacggat	ctgtatgtac	tgtggaccc	gttggcccttgc	ccggacccat	tggatgtgt	5220
ggccctctgt	gatttgcgg	cgccaccc	ccaaaggccg	aaatgggac	tgtggaaac	5280
gcccggcc	caggcccg	tggcca	ggagaagg	gttgccttgg	actjagcg	5340
ccagttggcc	caccaggaa	tcctgtcc	aatggactg	caggcgctaa	aggtcgacgt	5400
ggccctccag	gogttggcc	tgccaccc	ctggccagg	caagggat	ccctgtgtcc	5460
gttggagctg	ctggcgat	gggttgc	ggacttgc	gagaacctgg	gccagccgga	5520
tctaaggccg	actgtgttgc	caaggcc	ccaggatct	ctggccacca	aggccccct	5580
ggacccatcg	cgcaagaagg	caaaggcc	cttaatggcc	aaggccgt	tgcggggcc	5640
cctggaccac	caggcttgc	aggatcttc	ggctctag	gattgc	cgctgtatgt	5700
agacgaggcg	ttatgggtcc	acctggat	agggcgct	ctggccctgc	tggcgat	5760
gttccaaatg	cgccatgttgc	cagccagg	gagcccggt	ttatggggcc	tagagggtt	5820
cctggaaagcc	ctggcaat	cgccaccc	ggaaaaga	gcccgttgc	actccctggc	5880
atcgacggcc	gacctgttgc	aatcgaccc	gcaggcgct	ggggaggac	tggaaat	5940
ggcttccctg	ggcctaaagg	ccccac	gatctggaa	agaacggc	taaggccat	6000
gtctggactcg	ctgggtca	ggggccac	ggacctgt	gaaacaatgg	tgtcaagg	6060
ccgcctggcc	cacaagggt	tcaagg	aaaggccg	aggccccacc	tgggccttca	6120
ggcttccaa	gacttcccg	accatctgg	ccaggcg	aagttgaaa	gcctggcc	6180
agaggactgc	acggcgat	tggcttgc	ggtcttgc	gtccacgg	agagagac	6240
cctccaggcc	aatctggcc	cgcgac	actggccctt	tcggaa	aggacccat	6300
ggacccctcc	gacccatgttgc	caacaa	gaaccttgc	ttgtgggc	tgtggaaac	6360
gtctggaccc	cttggcttc	tggat	ggcgac	cgacgttgc	tattccctgt	6420
ggccaaaggcc	aaaaggccg	gcctggact	agggcg	tccgaat	cgacgg	6480
ggcgcttag	cgccccc	tgcgttgc	gccccggac	ctgtggcc	aacaggcg	6540
agaggccagg	ctgggtcc	tggcttgc	ggggccac	gttgc	ccggcc	6600
gagggccagg	aatgttgc	cgcttgc	aatggat	ttggcc	cgccgt	6660
ggtcaccc	cgcccaagg	cgacgg	gtaaagg	ctaaagg	gaatggc	6720
gtggccctt	ctggatgttgc	gggg	ggcgac	ggccgt	tattccctgt	6780
ccagctgggt	ctagaggc	cgccgg	cctggat	gtcc	accacccgt	6840
ggaaagac	gtccccc	accatctgc	atctgt	ccac	aggccgt	6900
gctggaa	aaggactgt	aggcccttgc	ggcgat	ttgg	ccggcc	6960
gaatgtcg	ctgtcg	accagg	tttt	ggcc	ccggcc	7020
ggaactgtcg	gaccggcc	aactcc	cttc	ggcc	cccttgc	7080
ctgggactgc	ccggatgt	cg	ggact	cc	ttcc	7140
gaaccgggac	cacttggat	tgctgg	ggac	ttgg	ccgg	7200
gaaagtctg	cgcttaac	tgc	ggat	ggat	ccgg	7260
gtggcc	ctggatgt	ggac	ggact	cc	ccgg	7320

-continued

aatattggac	ctgtcgccgc	agccgggtgct	cccggacactc	atggtccagt	cggtccagcc	7380
gggaaggcacg	gaaatagggg	agaaaacaggaa	ccctccggtc	ctgtggccc	agctggcgca	7440
gttggaccacaa	gaggccccatc	cggaacctcg	ggaatcccgcg	gagataagggg	cgaacctggc	7500
gagaaggggac	ctagaggact	gcctgggtg	aaaggccata	acggactgca	aggccgtcca	7560
ggcattgtcg	gccatcatgg	cgatcaaggt	gcacccggta	gtgtgggtcc	cgccggacgg	7620
aggggtcccg	ctgggtccatc	tggacccggd	ggaaaagatg	cgagaacagg	acatcttggc	7680
acagtggggg	ctgcccggat	tagaggccca	caggacatc	aaggccccgc	tggggcccca	7740
ggacctccgg	gaccggccagg	gccaccaggg	gttagtggcg	gaggatacga	tttccggctac	7800
gacggcgact	tctacagagc	cgaccagct	agatctgcc	ctagctgag	gcctaaggac	7860
taacaaatgg	acggccacact	gaatgtccctg	aacaacccaga	tgcgacact	gtgaccct	7920
gaggggcagca	gaaagaaaccc	tgcaccaacc	tgcaggggacc	tgagactgtc	tcacccggaa	7980
tggtcctccg	gtactactg	gategacccc	aatcaggggct	gcacatggaa	cgccatcaag	8040
gtgtactgcg	acttcagcac	cggegagaca	tgcatcagag	cccagectga	gaacatcccc	8100
gccaagaact	ggtcagaag	cagaacggag	aagaaacacg	tgtggctggg	cgagacatac	8160
aacccggcctt	cccgatcgaa	gtacaacgtg	gaaggcgtga	ccagcaaaaga	gttggccaca	8220
caqctggctt	tcatgagact	gctggccat	tacggccacg	aaacatcac	ctaccactgc	8280
aagaacacga	ttgcctacat	ggacgaggaa	accggcaacc	tgaagaaagc	cgtgtaccc	8340
cagggtccca	acggactgtgg	actgtggcc	gagggcaaca	cgagattcac	ctacaccgt	8400
ctgggtggacg	gctgcagcaa	aaagaccaac	gatgtggggca	agaccatcat	cgagtataaag	8460
accaacaacg	ccaggact	gcccctctg	gatatccccc	cactggat	tggaggcgcc	8520
gaccaagagt	tctttgtgga	categgcccc	gtgtgcttca	agtga		8565
SEQ ID NO: 33		moltype = DNA	length = 8644			
FEATURE		Location/Qualifiers				
misc_feature		1..8644				
source		note = Synthetic Construct				
		1..8644				
		mol_type = other DNA				
		organism = synthetic construct				
SEQUENCE: 33						
atgttcagct	tcgtggacct	gagactgctg	ctgtccctgg	ctgctacagc	cctgtgtaca	60
cacggacaag	aggaaaggca	ggtcgaaggaa	caggacgagg	acatccccc	tatcacctgt	120
gtcagaaacg	gctctgagata	ccacgaccgg	gatgtgtgg	agcccgagcc	tttcagaatc	180
tgcgtgtcg	acaatggcaa	ggtgcgtgt	gacgacgtga	tctgcgacga	gacaaagaat	240
tgcctggcc	cggaaatgtcc	tggggccgg	tgttgtccctg	tgtgccttgc	tggcagcgg	300
agccccacag	atcaagagac	aaacggcggt	gaaggccccca	aggcgatc	aggacataga	360
ggtcctagag	gacctggccgg	acccctctgg	agagatggaa	ttccctggaca	gcctggactg	420
cccgaccac	ctggacccctc	agggcctcca	ggtccaccag	gactcgagg	aaatttgcc	480
ccacagctga	gotacggcta	cgacggagaa	agcacaggcg	gcatctctgt	gcctggaccc	540
atgggacctt	ctggcccaag	aggactctct	ggtccctctg	gtgcctccgg	acctcgagg	600
tttcaaggac	caccaggcga	acctggcga	ccaggcgct	ttgggtccaa	gggaccaaga	660
ggccctctgt	ggccaccagg	caaaaatggc	gacgatggc	aaacccggaaa	gcctggaaagg	720
cctggcggaa	gaggccggcc	aggacggca	ggggcttagag	gggtgcctgg	aactgcagg	780
ctgcctggcc	tgaaggggca	cgagggttt	tctggactgg	atggcgctaa	ggggcgacgt	840
ggaccacgac	gacctaaagg	cgagctgtg	tctccctggc	agaatgtgtc	acctggacag	900
atgggtccca	gaggattggc	aggcgagaga	ggtagacctg	gcgtccggg	accagccgg	960
gtctagggaa	atgtggcgcc	aacaggtgt	gtggggcttc	ctggaccaa	cggaccag	1020
ggccacacgt	gattttcagg	cgctgttgc	gcaaaaggccg	aagcaggccc	acaaggacac	1080
aggggatctc	aagggtctca	gggggttgc	ggcgagccg	ggccacccctt	gcctggccgt	1140
gcagctggac	ctgtgtggaa	ccctgggtct	gatggacage	cagggtccaa	aggtgtctaa	1200
ggcgccccgt	gtattggccg	cgctccaggat	tttccctggc	caagaggccc	atctggaccc	1260
caaggccccag	ggggacccctc	gggttcttaag	ggaatatacg	gagacccagg	cgctcttgg	1320
agtaaaaggcg	atactggcgc	aaaaggcga	cccgacccctg	ttgggtgtca	aggacccct	1380
ggaccacgt	gctggggatc	caaaaggggg	gctaggggg	aaccaggacc	aacagggtcc	1440
cctggccac	ctggcgacgc	cgggggatc	ggatcttagag	gatcccttgg	cgccagatggc	1500
gtggccggac	aaaaggacc	tgcggcgaa	agggggatc	cagggtctgc	aggccctaag	1560
ggttctccag	gogaggctgg	cagacccggc	gaaggtggac	tcccagggtc	taaggactg	1620
acaggctcac	caggatctcc	cgggaccagac	ggaaaaaac	gaccccttgg	accggcgagg	1680
caggatgtt	gacccggatc	ccctggaccc	cctggtca	gaggacaagc	tggtgtatg	1740
ggtttccctgt	gacccaaagg	tgcaggccgg	gaacctggaa	aagcaggcga	gagggggat	1800
cccgacccat	cagggtctgt	tggacccatc	ggaaaaggat	cgcaacgtgg	tgccacaagg	1860
cctccaggcc	cacccggacc	agccggcgag	agaggcgaac	aaggaccagc	cggtatctca	1920
ggatttcagg	gactggcagg	gcctgttgc	ccgcctggcc	aggcaggggaa	gccaggcgaa	1980
cagggtgttc	ctggcgatct	tggccatcc	gggtccatag	gaggttaggg	cgaaaaggaga	2040
tttccctggcc	aaaggccatgt	tggggatcc	ccggggacc	ctggggccaa	gggtgtcaat	2100
ggtgcccccc	gcaatgtacgg	tgtctaaaggc	tgccggccgg	ccccagggtc	tcctggatct	2160
caaggcgccac	ctggacttca	ggggatgcct	ggcgaacccgg	gaggtgtctgg	acttcccggt	2220
ccaaaaggcg	ataggggaga	tgctggctt	aggggcgct	atggctctcc	tggaaaggat	2280
ggcgatgtat	gctgtatgtt	gggtccatcc	gggtccatag	gggtccatcc	gggtccatcc	2340
aaggggcaat	ctggacccat	tggacccat	gggtccatag	gtgttaggg	agccccaggc	2400
gaccggggag	accctgttcc	accaggaccc	gtggatcc	ctggacccctt	tggcgctgt	2460
ggtcaacccgt	gtgtcaagg	cgacccggc	gacgtggat	aaaaggcga	cgctggtcca	2520
cctggaccgg	ccggacccatc	tggccggcca	ggacccat	gaaatgttgg	tggcccttgg	2580
gccaaggcg	caaggatgtt	tgctggccca	ccaggccat	agggtggcc	2640	
ggaagagatgtt	gaccaccggg	gccaaggatgtt	tgctggccat	caccggggacc	ggcaggacca	2700
ggccggaaatgg	aaagggtggaaa	aggcccttgg	ggcgaatactg	ggcccttgcgg	caggccaggc	2760

-continued

aaaggatggggcc ctccaggacc tcgggggccc gccggcggaaa aaggatctcc aggccgcagat
ggaccggcag gcgtccccgg aacaccagggt ccacaggaga ttgctggaca aaggggagtt 2880
gtcgccgtc cggacagag gggagagaga gggtttcccg gactccctgg gccaagcagt 2940
gaacccggca aacaggggat atctggtgc acggggagaga gagggccacc aggaccaatg 3000
ggtcctccag gattggcagg gcctccggc gaatttggta gaaagggtgc tccaggcgcc 3060
gagggatctc ctggacgtg tggttctct ggcccaagg gcatagagg cgaaacaggc 3120
ccagctggac ctccaggcgc accccggcgt ccaggccac caggacctg tggccctgt 3180
ggaaaatctc gcgacagagg cggaaactgg cccggaggac cggccggacc tggtggacc 3240
tggtggctca gaggaccggc gggcaccata ggtcttagag gggacaaaggc gaaacaggc 3300
gagcaaggcg acagaggcat caaggcacac agaggattca gggactgtca gggaccacca 3360
ggggccctcg gaagtccccg cgaggcaggga ccaaggcgag ctatggtcc cgccggacc 3420
agaggaccac ctgggttctc tggtgcaccc gggaaaggcag gactgtatgg gctcccccgg 3480
cttattggc caccctggacc tagaggaga acaggcgac cggaccatgt tgaccaccc 3540
ggggccactcg gaccgcttg tcctccgtt cctcttctg cccggattcgat cttcagttc 3600
ctgcctcagc ctccctaaga gaaggccat gacggccgca gatattacag aggcagcagc 3660
gccaacgtcg tggggggacag agatctggaa gtggacacca cactgtggc cctgtctcg 3720
cagatcgaga acatcagaag cccggaggc acgagaaaga accctggccag aacctgtcg 3780
gacctgaaga tggccacac cgattggaa tggccgtact atggatcgaa ccccaaccc 3840
ggctgcaacc tggatggcat caagggttgc tgcaacatgg aaaccggcga gacatgcgtg 3900
taccggccac aggccatctg ggctcggaa aactgtgtaca tcggaaaggc aaaaaaggcg 3960
aaggccgacg ttgggttcc cgagagactt accgtatgtt ccaggatgg tgatggggcc 4020
cagggttctg accctggcca tggtgttata cagtgaccc tccctgggt tatgtctaca 4080
gaggccagcc agaacatcac ctaccactgc aagaacacgc tggcttacat ggatcagcag 4140
accggcaacc tgaagaaggc actgtgtttt cggggccggc acgagatcgaa gatcagcggc 4200
gaggccgcaaa gccgggttccat ctacagctgt acagtggatg gtcggaccgg ccataccatc 4260
ggttggggca agaccgtgtt ccggatccaa accaccaaga ccacggactt gcccattatc 4320
gatgtggccc ctctggatgt tggtggacccc gatcaagagt tggcttgc tggggccca 4380
gtgtgttcc tggtaagaatt cgtcgatcg gttcagcat ttaaaggcggt tgacaactt 4440
aaaagaaggaa aaaagaaggat tggaaagaaagg ggtgtatgg tgaatataa gtacagaccc 4500
gagaatgtac ccgggttctgat ttcgtttaa tggaaagaa aaaaatgtca gtttctgttga 4560
caccagaaca ctgtgtgtc tggtgggttgc actgtgtctg gccacttgtc agagcgtcga 4620
agaggaaaca gtgcggaaag gacctggccg cgatagaggaa cctagaggcg aaaggaggcc 4680
tcctggacat cctgttggatg atggggggaa tggttttttttggggccatccgttccgg 4740
acccctgggg acccttggggcc ttggggaaaatttttccggccatccgttccgg 4800
cgggacttggc ctttggacttca ttggggacttat gggcccaaga ggaccacccat gtttctgttga 4860
cgctccaggaa ccacaaggat ttcaaggacc agctggccgg ctttggccggaaatccgttccgg 4920
agggttctgtt ggtgttggatg gaccggccgg gccacttggaa aaaaatgtca gtttctgttga 4980
cccttggggaaatccgttccggccatccgttccgg 5040
cttccctgggaaatccgttccggccatccgttccgg 5100
cgggacttggc ctttggacttca ttggggacttat gggcccaaga ggaccacccat gtttctgttga 5160
aaatggccaca ccggacaaaatccgttccggccatccgttccgg 5220
ageccccaggat ccaggccggc acggaggat ttgttggatgtt gttggggacttggccctgt 5280
cgccggatatttccgttccggccatccgttccgg 5340
aatttggatgtt gttggggaaatccgttccggccatccgttccgg 5400
atttccctgggaaatccgttccggccatccgttccgg 5460
aggggcttggc ggttggacttca ttggggacttat gggcccaaga ggaccacccat gtttctgttga 5520
aagggttataatccgttccggccatccgttccgg 5580
cttccctgggaaatccgttccggccatccgttccgg 5640
aaggccctgggaaatccgttccggccatccgttccgg 5700
ttgggtccaaatccgttccggccatccgttccgg 5760
aaggccctgggaaatccgttccggccatccgttccgg 5820
ttggccctgtt ggttggacttca ttggggacttat gggcccaaga ggaccacccat gtttctgttga 5880
tatggggccatccgttccggccatccgttccgg 5940
cccttggggccatccgttccggccatccgttccgg 6000
ggggagccatccgttccggccatccgttccgg 6060
aaacccggccatccgttccggccatccgttccgg 6120
aaacaaatgtt gttggggacttca ttggggacttat gggcccaaga ggaccacccat gtttctgttga 6180
ggggccctggccatccgttccggccatccgttccgg 6240
ggggccctggccatccgttccggccatccgttccgg 6300
ttccctggggccatccgttccggccatccgttccgg 6360
cggaaggccatccgttccggccatccgttccgg 6420
ttggggccctggccatccgttccggccatccgttccgg 6480
agccgttggatccgttccggccatccgttccgg 6540
cggttggccatccgttccggccatccgttccgg 6600
ggggccctggccatccgttccggccatccgttccgg 6660
ttccctggggccatccgttccggccatccgttccgg 6720
ttggggccctggccatccgttccggccatccgttccgg 6780
taaaggccatccgttccggccatccgttccgg 6840
tccctggggccatccgttccggccatccgttccgg 6900
ggggccctggccatccgttccggccatccgttccgg 6960
ttccctggggccatccgttccggccatccgttccgg 7020
ttggggccctggccatccgttccggccatccgttccgg 7080
ttccctggggccatccgttccggccatccgttccgg 7140
ggggccctggccatccgttccggccatccgttccgg 7200
ttccctggggccatccgttccggccatccgttccgg 7260
ttggggccctggccatccgttccggccatccgttccgg 7320

-continued

agatggaaat	ccggcaatg	atggcccgc	tggagagat	ggacagccgt	gacataaggg	7380
c gagcaggc	tacccaggc	atatggacc	tgtcgccgc	gccggtgctc	ccggacactca	7440
tggccagtc	ggtcagcc	ggaagcacgg	aatagggga	gaaacaggac	cctccgggtcc	7500
tgttggcca	gtggcgca	ttggaccaag	aggccatcc	ggacctcagg	gaatcccg	7560
agataagggc	gaacctggcg	agaagggacc	tagaggactg	ctgggctga	aaggccataa	7620
cggactgcaa	ggccctggcag	gcatgtgg	ccatcatggc	gatcaagggt	caccggtag	7680
tgtgggtccc	gccccggca	ggggtccccc	tggccatct	ggacccgc	gaaaagatgg	7740
cagaacagga	catcttggc	cagtggggc	tgccggaa	agagggccac	agggacatca	7800
aggccccgt	ggccggccag	gacccgggg	accggcagg	ccaccaggcg	ttatggccgg	7860
aggatatacgat	tccgtgtacg	acggcactt	ctacagagcc	gaccagccta	gatctgcccc	7920
tagctgtgg	cataaggcg	acgaatgg	cgcacactg	aagtccctga	acaaccagat	7980
c gagacactg	ctggccctg	agggcagcg	aaaaaaccc	gccagaaact	gcaggaccc	8040
gagactgtct	caccccgaa	ggteetccgg	ctactactgg	atcgacccca	atcagggtcg	8100
caccatggac	gocatcaagg	tgtactgca	cttcagcacc	ggcgagacat	gcatcagac	8160
ccagcctgt	aacatcccg	ccaagaactg	gtacagaa	agcaaggaca	agaaacacgt	8220
gtgtgtggc	gagacaatc	acggcccg	ccagttcgag	tacaacgtgg	aaggcgtgac	8280
cagcaaagag	atggccacac	agctggctt	catgagactg	ctggccaa	acgcccggca	8340
gaacatcacc	taccactgca	agaacagat	tgccatcat	gacgaggaaa	ccggcaaccc	8400
gaagaaaagcc	gtgatccgtc	agggcttaa	cgacgtggaa	ctgggtggcc	agggcaacag	8460
casattacc	taacccgtc	tgggtggacgg	ctgcagcaaa	aagacaaacg	agtggggccaa	8520
gaccatcatc	gagtataaga	ccaacaagcc	cagcagactg	cccttctgg	atatcgcccc	8580
actggatatt	ggaggcggcc	accaagagtt	cttggac	atcgccccgg	tgtgtttcaa	8640
gtga						8644
SEQ ID NO: 34			molttype = DNA	length = 8987		
FEATURE			Location/Qualifiers			
misc_feature		1..8987				
		note = Synthetic Construct				
source		1..8987				
		mol_type = other DNA				
		organism = synthetic construct				
SEQUENCE: 34						
atgttcagct	tctgtggac	ctgactgtcg	ctgtctctgg	ctgtctacagc	cctgttgtaca	60
cacggacaag	aggaaggcc	ggtcgaagga	caggacgagg	acatccctcc	tatcacctgt	120
gtgcagaa	gcctgtggat	ccacggaccgg	gatgtgtgg	agcccgaggcc	ttgcagaatc	180
tgcgtgtgc	aatatggca	ggtgcgtgtc	gacgacgtga	tctgcgacga	gacaaagaat	240
tgccttggc	ccggaaatggc	tgaggccgaa	tgttgtctt	tgtgccttgc	ttggcagcgg	300
agccccacag	atcaagagac	aacaggcg	gaaggcccca	agggcgat	aggacactaga	360
gtgtccatag	gacccgtccg	acccctctgg	agagatggaa	ttccctggaca	gccttgactg	420
cccgaggacc	ctggacatcc	agggcctca	gggtccaccag	gactcgagg	aaattttggcc	480
ccacagctga	gotacggcta	cgacgagaaa	agcagaggcg	gcatcttgt	gccttgaccc	540
atgggaccc	ctggcccaag	aggacttctt	ggtcctctgg	gtgtccagg	acccctcagg	600
tttcaaggac	caccggcg	acccggcg	ccaggcgat	gtgggtccat	gggaccaaga	660
ggccctctgt	ggccaccagg	ccaaaatgg	gacgatggcg	aaggccggaa	gcctggaaagg	720
cctggcggaa	gaggcccgcc	aggacccaa	ggcgctagag	ggtgtccctt	aactcgagg	780
ctgccttggc	tgaaggccca	cagaggctt	tctgtgtgg	atggcgctaa	gggcgacgc	840
ggaccaggac	gacccatgg	cgacgtctgg	tctctctggcc	agaatgtggc	accttgacag	900
atgggtccca	gaggattgg	aggcgagaga	gttagacctg	ggtgtccgg	accaggccgt	960
gttagagggaa	atgtggcgc	aacaggctgt	gtgtggctt	ctggaccaac	cggaccagct	1020
ggcccacctg	gatttccagg	cgctgttgg	gcaaaaggccg	aaggccggcc	acaaggaccc	1080
agggggatcg	aaaggcttca	ggcggttga	ggcgacccgg	ggccacccgg	gcctggccgt	1140
gcagctggac	ctgtgtggaa	ccctgggtgt	gtatggac	cagggtccaa	agggtctaat	1200
ggcgccccctg	gatattggccg	cgctccagg	ttttccggcg	caagggggcc	atctggaccc	1260
caaggcccc	ggggaccc	gggtctcaag	ggaatagcg	gagacccagg	cgctcttggg	1320
agtaaaggcg	atactggc	aaaaggcgaa	cccgaccc	tggggatgtca	aggacccctt	1380
ggaccacgt	gcgaggaa	caaaaggagg	gctagggggg	aaccaggacc	aacagggctc	1440
cctgggtccac	ctggcgac	cgggggac	ggatctagag	gattccctgg	cgccatggc	1500
gtggccggac	caaaaggacc	tgcaggcgaa	aggggatc	caggctctgc	aggccctaag	1560
ggttctccag	cgagggtcg	caggccgg	gaagctggac	tcccagggtc	taaggactg	1620
acagggtcac	caggatcc	cgggaccag	ggaaaaacac	gacccggagg	accggccagg	1680
caaggatgtt	gaccgggtcc	tcctgtgg	cctgggtca	gaggacaa	ttggctgtat	1740
ggcttccctg	gacccaaagg	tgcagccgc	gaaacctggaa	aaggcaggcga	gagggggat	1800
cccgaccc	cagggtctgt	tggaccc	ggaaaagatg	gcgaagctgg	tgcacaagg	1860
cctccaggcc	caggccggacc	agccggcc	agaggcg	aaggaccgc	cggtatccca	1920
ggatttccagg	gactgtcc	ccggcc	ggggcgg	ggggcgg	ggcaggccgaa	1980
caagggtgtc	ctggcgatct	tggggccctt	gggtcttgc	gaggatagg	cgaaaaggag	2040
tttcccttggc	aaaggggcg	tcagggtca	ccggggac	ctggacca	gggtgtcaat	2100
ggtgccccag	gcaatgacgg	tgctaaaggc	gacgcaggcg	ccccagggtc	tcctggatct	2160
caaggcgac	ctggactca	gggaatgc	ggcgaacccgg	gagctgtgg	acttcccggt	2220
caaaaaggcg	ataggggaga	tgctggct	aaggccgt	atggctctcc	tggaaaggat	2280
ggcgctacag	gcctgtacgg	cccaatttgc	cctccggac	ctgctggc	tccaggccat	2340
aaggcgaat	ctggacatcg	tggacccgt	ggtcttacag	gtgttgggg	agccccaggc	2400
gaccggggag	agcctggtcc	accaggac	gctggattt	ctggaccc	tggcgctgtat	2460
ggtcacac	gtgtcaagg	cgagccagg	gacgctgggt	aaaaaggcga	cgctggccca	2520
cctggaccc	ccggacatcg	tggccggcc	ggaccttattt	gaaatgttgg	tcccccgtgc	2580
ccaaaggcg	caaggaggatc	tgctggccca	ccaggcgct	caggattccc	aggtgcgcgt	2640

-continued

ggaagagttg	gaccaccggg	gccaaagtgg	aatgctggc	caccgggacc	gccaggacca	2700
gcccggcaaaag	aagggtggaaa	aggccctagg	ggcgaaaactg	gccctgcagg	cagggcaggc	2760
gaagtggggcc	ctccaggacc	tcggggcct	ggccggcgaaa	aaggatctcc	aggcgcagat	2820
ggaccccgacg	gogctccgg	aacaccagg	ccacaggaa	ttgctggaca	aaggggagtt	2880
gtcggcctgc	caggacagag	gggagagaga	ggttttccag	gactccctgg	gccaagcgga	2940
gaacctggca	aacaggggacc	atctggtgcc	ageggagaga	gaggggcacc	aggaccaatg	3000
ggtctccag	gattggcagg	gectcttgg	gaatcttgg	gagaagggtgc	tccagggcgcc	3060
gagggttctc	ctggacgtga	tggttctct	ggcgccaa	cgatagagg	cgaaaacagge	3120
ccagctggac	ctccaggcgc	acccggcgct	ccaggcgac	caggaectgt	tgccctgtct	3180
gaaaatcttgc	gogacaggag	cgaaaactgt	cccgccaggac	cagggcaggc	tgttgaccc	3240
gtgggtgtca	gaggaccggc	tggtccatgg	ggtcttagag	gogacaagg	cgaaaacaggc	3300
gagcaaggcg	acagaggcat	caaggggac	agaggattca	cgggactgca	gggaccacca	3360
ggggcgeectg	gaggtcccg	cgagcaggga	ccaagcgag	ctagtggcc	cgccggaccc	3420
agaggaccac	ctgggtctcg	tgggtcaccc	ggaaaggagg	gactgaatgg	gtccccggga	3480
cctattggc	cacccgtgg	tagaggaga	acaggcgac	caggaccat	tggaccaccc	3540
ggggccacctg	gaccgctgg	tcctcttgg	cctcccttgc	ccggattcga	cttgcaccc	3600
ctgcctcage	ctccctaaga	gaaggccat	gacggggca	gatattacag	agccgacac	3660
gccaacgtcg	tgccggacag	agatctggaa	gtggacacca	cactgaatgc	cctgtctcag	3720
cagatcgaga	acatcaga	ccccgggg	agcagaaa	accctgcac	aaccgtgtcg	3780
gacccgttga	tgtgcaccc	cgatggaa	tctggcgat	actggatcga	ccccaaaccc	3840
ggctgtcaacc	ttggatgcac	caagggttcc	tgcaacatgg	aaaccggcga	gacatcggt	3900
taccccaac	agccatctcg	ggctcagaag	aactggata	tcagcaagaa	ccccaaaggac	3960
aagccggac	tttgggtcg	cgagatgg	accgtatgg	tccatggta	gtatggccgc	4020
cagggtctcg	accctgcga	tgggtctatc	cgactgtcc	tcttgcgggt	gtatgtctaca	4080
gaggccagcc	agaacatcac	cttccactgc	aagaacagc	tggcttacat	ggatcagcag	4140
acccggcaacc	tgaagaaggc	actgtgtt	caggccagca	acgagatcga	gatcagagcc	4200
gaggccaaaca	gcgcgttcac	ctacaggctg	acagtggat	gtgtcaccag	ccatcacaggc	4260
gcttggggca	agaccgtat	cgatgtccat	accaccaaa	ccagcagact	gccccatcatc	4320
gatgtggccc	ctctgtatgt	tggggcaccc	gatcaagat	tccggttgc	tgtgggccc	4380
gtgtgttcc	tgtaaatgtt	tttccacat	attggcgct	tttggcaatg	tgaggggcccg	4440
gaaaaccttgc	cttgcgttct	tgaagacat	tccttgcgt	tttcccctc	tcgccaagg	4500
aatgtcaagg	ctgttgaatg	tcgttgaagg	agcagtcc	ttggaaatgtt	tttggaaatgg	4560
aacaacgtct	gttagcgtacc	tttgcaggca	gcggaacccc	ccacctggcg	acaggtgcct	4620
ctgcggccaa	aagccacgt	tataagatac	acctgcaaa	cgccgcacaa	cccagtggca	4680
cgttgtgatg	tgtatgttgc	tggaaatgt	caaatggctc	acctcaaggc	tatcaacaa	4740
ggggctgtaa	gatgcccaga	aggtatccca	ttgtatggg	tctgtatgg	ggcctcggt	4800
cacatgttt	acatgtgtt	agtcgagg	aaaaaggcgt	taggc	cccccccc	4860
gacgtgggtt	tcctttgaaa	aacacgatgc	tgagcttgc	ggacaccaga	acactgtgc	4920
tgtgtggcgt	gacactgtgc	ctggcactt	gtcagagcgt	gcaagaggaa	acagtgcgg	4980
aaggacgtc	cgccgtataga	ggacatggag	tccttgc	cccttgcgt	5040	
gagatggcga	gatgtggac	acaggacc	ctgttccacc	aggacccca	gggcctctg	5100
gccttggagg	aaatttgccc	gctcgtac	atggcaaa	cgtcggactt	ggccctggac	5160
ctatggact	tatggggccca	agaggacc	cagggtctgc	aggcgttcc	ggaccacaa	5220
gatttcaagg	accaggttgc	gagcttgg	aaccttggca	aaacgg	tatccgtt	5280
gaggaccagg	cgggccaccc	ggaaaatgt	cgcaagatgg	gcacccctgg	aaagcttgcgt	5340
gaccggcga	aaagggtgtt	tttggaccc	aaggcgcc	aggcttgc	ggaacacctg	5400
gactgtgttcc	cttcaaggcc	atcaggagg	acaatggct	ggacggact	aaaggacaa	5460
ctgggtgtcc	ttggctggaaa	ggcgaaccc	cgccaccc	cgaaaatggc	acaccaggac	5520
aaaccggcgc	aaggaggactt	cctggcgaga	gaggaaatgt	tggagccca	ggtcaggac	5580
gcmcacgg	atctgtatgg	tctgtggac	ctgttggcc	tgccggact	attgttgc	5640
ctggccctcc	ttggatgttcc	ggccgttcc	ggccaaagg	cgaaaatgg	gtgttggaa	5700
acgcggacc	tgccaggccca	ttggaccc	ggggagaatgt	tggattgc	ggactgtac	5760
gaccaggttt	gcacccagg	aatcttgcgt	ccaaatggact	gacaggccgt	aaagggtcag	5820
ctggccttcc	aggcgttgc	ggtgcacc	gactgtcc	accaggatgt	atccctgtc	5880
ctgttggagc	tgctggcgct	acgggttgc	gaggactgt	tggagaaat	gggcacccgg	5940
gatctaaagg	cgagtgttgc	aacaagggt	agccaggat	tgtgttgc	caaggccgc	6000
ctggaccatc	aggcgttgc	ggcaaaac	gccttgcgt	cgaaaggccgt	agtgcgggc	6060
cttcttgcacc	accaggcc	agaggatc	ctggcttgc	aggattgc	ggcgttgc	6120
gtatgttgc	cgatgttgc	tttgcgttgc	ccatgttgc	gacaggccgt	aaagggtcag	6180
gagggttccaa	ttggcgatgt	ggcgttgc	ggggccgtt	tttgcgttgc	tttgcgttgc	6240
tgccctggaa	cccttgcgtt	atccggcc	ggccaaatgg	aggccctgtt	ggactccctg	6300
gatctgttgc	ttggatgttgc	ccaaatggac	cgccgttgc	cgccgttgc	tatccgtt	6360
ttggcttccc	ttggcttccaa	ggccccac	gcatgttgc	ggccgttgc	tttgcgttgc	6420
atgtgttgc	cgatgttgc	aggggaggat	ctggatgttgc	ccggaaatgt	tttgcgttgc	6480
ggccgttcc	ggccaaatgt	tttgcgttgc	ggggccgtt	ggccgttgc	tttgcgttgc	6540
caggcttcc	aggacttcc	ggaccatgt	ggccgttgc	cgaaatgttgc	tttgcgttgc	6600
aaagaggact	gcacggcg	tttgcgttgc	cggttgc	cggttgc	tttgcgttgc	6660
gccctccagg	cgatcttgc	gtacttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6720
gttgcgttcc	aggacttgc	ggcaacaa	cgcaaccc	tttgcgttgc	tttgcgttgc	6780
cagcttgc	tttgcgttcc	tttgcgttgc	cgccgttgc	cgccgttgc	tttgcgttgc	6840
gttgcgttcc	tttgcgttcc	tttgcgttgc	cgccgttgc	cgccgttgc	tttgcgttgc	6900
atggcgctag	aggcgttcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6960
acagaggcga	ggcttgcgttcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	7020
ggcggagggg	cgaaatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	7080
ctgttgcgttcc	tttgcgttcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	7140
tcgttggcc	tacttgcgttcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	7200

-continued

```

gaccagctgg gtcttagaggc gacggcgac cgccctggaaat gacagggttt ccaggcgccg 7260
cttgaagaac aaggcttcga ggaccatctg gcatctctgg tccaccaggg ccacctggc 7320
ctgctggaaa aagaaggactg agagggccta gggcgcataa gggtcccgat ggaagaacccg 7380
gccaagtcgg agctgtcgcc ccaccaagggtt ttgcccggcga aaaaggccct akgcgagaaag 7440
cttggaaactgc aggaccggcg ggaactcccg gtcctcaagg attgcttggc gcccctggaa 7500
ttctggact gcccggtagt cgcggagaaat gtggacttccc aggtgttgc ggcgcgcgtcg 7560
gagaaccggg accacttggaa attgtcgac cacctgggtc aagagggtccaa ctgttgccag 7620
tttggaaatgcc ttggcgtaac ggtgtcccg gccaagtcggg cagatggaa aatcccgca 7680
atgtatggcccc ttgggggaga gatggacago ctggacataa gggcgagcgaa ggttaccatcg 7740
gcaatattgg acctgtcgcc gcaacccggg cttccggaccc tcatggtccca gtcgggtccag 7800
ccggggaaacg ggaaatccggg gacccctccggg tcccttggc ccagctggcg 7860
cagttggacc aagaggccca tccggaccc agggaaatccg cggagataaag ggcgaactctg 7920
gggagaagggg accttagagga ctgcctggcc tggaaaggccca taacggactg caaggccctgc 7980
caggcatttg ttggccatcat ggccatcaaa ggtgtcccg tagtgttgcgtt cccggccggac 8040
cgagggggtcc ctgttggccaa ttggggggcc cccggaaaaga tgccgagaaca ggacatctg 8100
gacacagtggg gctgtccggg attagaggcc cacaaggacata tcaaggcccc gctggccgc 8160
caggaccccg gggaccggca gggccaccagg gctttagtgg cggaggatcc gatcccgct 8220
acgcacggcga ctttcatacgcc ggcggaccgc cttagatctgc cccttagcccg aggcttaagg 8280
actacaaatgg gacgcgcaca ctgaaatcccg tggaaacccca gatcgagaca ctgtgacc 8340
cttgaggcccg cggaaatggc cctccggacaa cctggcgccg gctggactc tcttcccccc 8400
aatgttctcc cggctactac tggatcgacc ccaatcagggg ctgcaccatcg gacccatca 8460
agggtgtactg cggacttccgc accccggcaga catgcatacg agcccaagctt gagaacatcc 8520
ccggccaaaggaa ctggtagaca gacggccaaagg acaagaaatgg cgtgtggcg ggcgagacaa 8580
tcaacccggg cggccatcc gggatccaaatgg tggaaaggccg gaccggccaa gagatggcc 8640
cacagctggc ttccatgaga ctgtggccaa attacggccg cccggccaaatcc accttccact 8700
gcaagaaacac cattccctac atggacgggg aaaccggcaaa ccttggaaatgg gccgtatcc 8760
tggggggtcc taacgacgtg gaaatggcg cccggggccaa cccggatcc accttccact 8820
tgcttggggcc cggccgtccccc aaaaacccca gacgtgtgggg ccggccatcc atcgatata 8880
agaccaaaaccc gcccggccaa ctgccttcc tggatccgacc cccactggat attggaggcg 8940
ccggccaaaggaa gtttttttgtt gacatcgacc cccggccgtt ccgtgtccgtt caatgt 8987

```

```

SEQ ID NO: 35 moltype = DNA length = 7434
FEATURE Location/Qualifiers
source 1..7434
mol_type = unassigned DNA
organism = Homo sapiens

```

```

SEQUENCE: 35
atgtcttaggg gtccggggcc cggggctgtg ctgttggccg tccagtgcct gggggacacgc 60
gtggcccttcga cggggggatccaa gaagggcgaa aggccggctc acaatggatgt tcagccccc 120
tcccccgggtgg ctgtcgtca aagaagcccc ggttggttatg acaatggaaa acactatcc 180
ataaatcaac aatggggggcg gacccatctca ggcataatggc tggggatgtac ttgttattgg 240
ggaaagccggag gttttaaactg cgagatggaaatgg ccttgcggatcg aagagactcg ttggaca 300
tacactggaaac acttccatcccg agtgggtgtcc gttccataaaga ctccatgtatcc 360
tggggacttca ctgtcatacgcc ggtctggccg gggggatggaaatgg ctttgcgttcc 420
tgccatcaatgg ggggttccgtcc tccaaatgggatcc ggttgcaccc gggggggccatcc 480
ggttgggttaca ttggatgtgg tggatgttcc ggttgcgttcc aaggatggcc gaccgtcc 540
cccatatgtggaaatggatgtt tgatgtatccg ggttgcgttcc cttatgtgtt ccggggaaacg 600
ttggggggggcc ttggatgtggcc ggttgcgttcc ttggatgttcc gggggggccatcc 660
ggggccatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 720
agaattgggg acacccggcc caagaaggat aatccggatcc acctgtccca gtcgtccatcc 780
acccggccatcc gggggggccatcc ggttgcgttcc ttggatgtggcc gggggggccatcc 840
acggggatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 900
cggcccttccttccatggcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 960
cagtggccatcc gggggccatcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1020
agctggccatcc gggggccatcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1080
gttttccatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1140
ggccggatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1200
tgcacacccatcc acactgtttt ggttgcgttcc ttggatgtggcc gggggggccatcc 1260
caatccatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1320
gacaacatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1380
tgcccccgttccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1440
ggggatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1500
aatgggtccatcc gggggatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1560
gttggccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1620
cttggccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1680
tgccatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1740
catgggtccatcc gggggatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1800
ccttggccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1860
agtcggccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1920
tacatttccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 1980
ggccacttccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 2040
cttccatccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 2100
agcaccggccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 2160
cttggccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 2220
gttggccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 2280
gttggccatcc tggatgttcc ggttgcgttcc ttggatgtggcc gggggggccatcc 2340

```

-continued

cttcctggcc	gaaaatacat	tgtaaatgtc	tatcagatata	ctgaggatgg	ggagcagagt	2400
ttgatcttgt	ctacttcaca	aacacacgg	cctgtatgcc	ctccctgacc	gactgtggac	2460
caagttgtat	acacccat	ttgttgcgc	tggagcagac	cccaggctcc	catcacagg	2520
tacagaatag	tctattcgcc	atcagtagaa	ggtagcagca	cagaactcaa	ccttcctgaa	2580
actgcaaaact	ccgtcaccct	cagtgacttg	caacctggtg	ttcagtataa	catcaactatc	2640
tatgctgtgg	aagaaaatca	agaaaatgtca	cctgtgtc	ttaacaaca	aaaccacttgc	2700
acccccacgct	cagatacagt	geccetctcc	aggggacctgc	agtttgcga	agtgcacag	2760
gtgaagggtca	ccatcatgtg	gacacgcgc	gagagtgcag	tgaccggcta	ccgtgtggat	2820
gtgtatccccg	tcaacctgdc	tggegagac	gggcagagg	tgccccatcg	caggaacacc	2880
tttgcagaag	tcacccgggt	gtccccctgg	gtcaccttat	acttcaaaatg	ctttgcagt	2940
agccatcgaaa	ggggagagca	gectctgact	gtcaaacaga	caaccaaact	ggatgtccc	3000
actaacctcc	agtttgcata	tgaaactgtat	tctactgtcc	ttgttgatgt	gactccac	3060
cggggcccaga	taacaggata	ccgactgacc	gtgggccta	cccgaaagagg	acagccagg	3120
cagtacatgt	tgggtccctc	tggttccca	tacccactgt	ggaatctgca	gcctgcat	3180
gagttacaccg	tatccctgt	ggccataaa	ggcaaccaaa	agggccccaa	agccactgt	3240
gtctttacca	cactgcagcc	tgggactct	atccaccc	acaacaccg	ggtactqag	3300
accaccattg	tgatcacatg	gacgcctgt	ccaagaatttgc	gttttaagct	gggtgtacga	3360
ccaaaggccagg	gaggagaggc	accacaggaa	gtgacttca	actcaggaa	categttg	3420
tccggcttgc	tcccgaggat	agaatactgc	tacccatcc	aagtctgt	agatggacag	3480
gaaaagatgt	cgccaaatgt	aaacaaatgt	gtgacccat	tgttccccc	aacaaacttgc	3540
catctggagg	caaaccctga	cactggatgt	ctcacagtc	cttgggagag	gagcaccacc	3600
ccagacattt	ctgggtatag	aatttaccata	acggccagca	ggggaaatttgc	3660	
tttggaaag	ttgttccatgc	tgatcagat	ttctgcactt	ttgtataact	gatgtccggc	3720
cttggatgtca	atgtcagtgt	tttactgtc	aaggatgaca	aggaaatgt	ccctatctt	3780
gataccatca	tcccagaggt	gccccactc	actgaccta	gtttgttga	tataaccat	3840
taaagcatcg	gtctgtgggt	gaccccgct	aactcttcc	ccatttattgg	gtacccatc	3900
acagtagttg	cgcccgaggaa	aggtatccat	attttgcga	attttgcga	cttccca	3960
ggatactaca	cagtacagg	gtctggagcc	ggcattgtact	atgatatacg	cgttatca	4020
cttattaatg	ggggcgagag	tgcctctact	acactgac	aaacaaacggc	tggttccct	4080
cccaactgacc	tgcgatttac	caacattgtt	ccagacacca	tgcgtgtc	ctgggcttca	4140
ccccccatc	ttttttatca	caacttgc	gtgcgttact	cacctgtgaa	aatatggaa	4200
gtatgttgc	agtttgcata	ttctcttca	gacaatgc	ttgttcaatc	aaatctct	4260
ccttggatag	aatatgtatgt	gagtttcttgc	atgttctac	aaacaaatgt	gagcacacc	4320
cttagagaa	gacagaaaaac	agggtttgtat	tccccaactg	gcattgtactt	ttctgtat	4380
actggccatc	cttttactgt	gtactgggt	gtcttctcg	ccacccatc	ttggctacagg	4440
atccggccatc	atcccgagca	tttcgttggg	agacccatcg	aatatgggt	gccccactt	4500
cggaatttca	tcacccctac	caacccatc	ccaggcagac	agtatgtgt	cagcatcg	4560
gtctttaatg	gcagagagga	aagtccctt	tttgcgttgc	aaacaatca	agtttctgt	4620
gttcccgagg	acctgtggat	ttgttgcgt	accccccacca	gcctactgt	cagctgggt	4680
gtctctgtgt	atattacagg	atattacagg	tgcgttgc	gagagacagg	aggaaatagc	4740
cctgtccagg	agtttactgt	gtcctgggg	aaatgttac	ctaccatca	cggtttaaa	4800
cctggatgt	attataccat	cactgtgtat	gtctgtact	gcccgtgg	agccccgc	4860
agdagcaacg	caatttccat	taatttccca	acagaaatttgc	acaaacccatc	ccagatcg	4920
gtgaccgtat	ttttagggca	cagtttgc	gtcaatgtgc	tgcgttca	ttccctgtt	4980
actgggttca	gagtttaccc	cactccaaa	aatggccac	gaccaacaaa	aactaaat	5040
gcagggtccag	atcaaaacaga	aatgtactt	gaaggctgc	agccccacgt	ggagtatgt	5100
gttagtgcgt	atgtcaggat	tcccaatgg	gaggttc	ctctgggtca	gactgc	5160
accaacatttgc	atgccttca	aggactgtca	tttactgtat	tttgcgttgc	ttccatca	5220
attgttgcgg	aaaggccac	ggggcaatgtt	tccaggatca	gggtgtacca	ctcgaccc	5280
gaggatggaa	tccatgtat	atccctgc	cctgtatgtt	aagaagac	tgcagatgt	5340
caaggccatca	gacccgggtt	tgatcaca	gtcgttgc	ttgccttgc	cgatgtat	5400
gagggccac	ccatgtatgg	aaccctgtt	acagcttac	ctgcacca	tgatgttgc	5460
tttacttgc	tcacccacc	aagcttgc	gccccatgtt	caccacccaa	ttgttgc	5520
actggatata	gagtgcgggt	gacccca	gagaagac	gaccaatgt	agaaaatca	5580
cttgcctcttgc	acagcttac	cgtgttgc	tcaggactt	ttgtggcc	caaataat	5640
gtgtatgtat	atgtctttaa	ggacatgtt	ggacatgtt	ttgtatgtt	aatccaa	5700
actctggatgt	atgtcaggcc	accaaaatgt	gtcgtgt	catgttca	tgagacc	5760
atcaccatca	gtctggatca	caagactgt	acgtact	gtttcaat	tgatgtcc	5820
ccagccaaatgc	gocagactt	aatccatgg	accatca	tcagatgt	aaatgttgc	5880
atcacaggat	tacaaccagg	cactgtact	aaatgttac	tgttgcac	aatatgttgc	5940
gtctccggatct	ccccctgtgt	tcacccatc	tcacccatc	ttgtatgtt	atccaa	6000
cgtttccctgg	ccaccacacc	caatttcttgc	ctggatcat	ggcagccgg	acgtgtcc	6060
attaccggct	acatcatca	gtatgttgc	cctgggttgc	ctccca	agttgtcc	6120
cgcccccgg	ctgggttgc	agaggcttac	attactgttgc	ttgtgggg	aaatgttgc	6180
acaattttatgc	tcatttgcct	gaagaataat	cagaagac	agccccatgt	ttgtggaaa	6240
aagacagac	agctttccca	actgttgc	cttccacacc	ccaatcttca	ttggccac	6300
atcttggatgt	ttccttccac	agtgttgc	acccttgc	tcacccatcc	ttgggtat	6360
actggaaatgt	gtatttgc	tcctggact	tctggtgc	aaatgttgc	ttggcaacaa	6420
atgtatgttgc	aggaacatgg	tttttaggg	accacacc	ccacaacgg	cacccatca	6480
aggcatggc	caagaccata	ccggccgtat	gtatgttgc	aaatgttac	tgttgcac	6540
cccaaggaaatgt	atgtatgttgc	tcacccatc	ccacacgttgc	cgggacttca	tccaaatgt	6600
tctacaggac	aagaatgttgc	accatcttca	ggggccat	ccaggacact	ttttgttgc	6660
tctgagtttca	tcatttgc	tcatttgc	ggcactgtat	ttttgttgc	ttttgttgc	6720
gttccctggaa	tttcttccat	tgccacttgc	acaggcttca	ccaggggtgc	cacccatca	6780
atcatatgttgc	aggcacttgc	agaccac	aggcata	tttgcgttgc	ttttgttgc	6840
gttggccaaatgt	ctgtca	aggcttgc	caacccatc	ttttgttgc	ttttgttgc	6900

-continued

tacacagttt cccattatgc cgftggagat gagtggaaac gaatgtctga atcaggcttt	6960
aaactgttgtt gocagtgctt aggctttgga agtggcatc tcatgttgc ttcatctaga	7020
tggtgccatg acaatgggtt gaactacaag attggagaga agtggaccc tcagggagaa	7080
aatggccaga ttagttagctg cacatgtttt gggAACGGAA aaggagaatt caagtgtac	7140
cctcatgagg caacgttta tgatgtggg aagacatacc acgttaggaga acatggcag	7200
aaggaatata tcgggtccat ttgtcttcgc acatgtttt gaggccagcg gggctggcgc	7260
tgtgacaact gcccagacc tgggggtgaa cccagttccc aaggactac tggccagtc	7320
tacaaccagt attctcagag ataccatcg agaacaaca acaaataat ttgccttata	7380
gagtgcttca tgcctttaga tgtacaggct gacagagaag attccgagaa gtaa	7434
 SEQ ID NO: 36	moltyppe = DNA length = 7434
FEATURE	Location/Qualifiers
misc_feature	1..7434
	note = Synthetic Construct
source	1..7434
	mol_type = other DNA
	organism = synthetic construct
 SEQUENCE: 36	
atgctgagag gacctggacc aggactgctg ctgtggctt ttcagtgtct gggAACAGCC	60
gtgccttaga cagggccag caagtctaaa agacaggccc agcagatggt gcagccttag	120
tctcctgtgg ccgtgtctca gtctaaagccc ggctgctacg acaacggcaa gcactaccag	180
atcaaccagg agtgggagag aacatccatcg ggcaacgccc tcgtgtgtac atgttacggc	240
ggcagcagag gtttcaactcg cgagtctaaa cccggggcccg aggaaacccgt ctgcacaaag	300
tacacccggca acacccatcg agtggccac acatccatcg gggccaaagga cagcatgatc	360
tgggactgca catgtatcg agccggcaga ggcagaatca gtcgcaaat cgccaacccg	420
tgtcacaaag ggccggcagag ctataagatc ggcgacactt ggagaaggcc ccacggagaca	480
gggggttaca tgctggaaat cggtgtctt gggcaacggca aaggcgatg gacgtgcag	540
cctatcgccg aagaatgtttt cgatcatggc gcccggaaata gtcgtctgt gggcgagaca	600
tgggagaagg ctttacaaagg ctggatgtatg gtggactgtt cttgccttgg cgaaggcagc	660
ggaaagaatca ctgttaccatcg ccggaaaccgg tgcaacgacc agatattccat aaccatgtac	720
cggatccggcg atacccgttca caaaggaaac aacccggggca acctgtgtca gtgcattgt	780
accggaaatgtt gaaggaggca atggaaatgtt gggggccca caagcgatc gacaacatct	840
tctggcagcg gccccttttac cgatgtgcgg gctgttattt atcagccccca gcctcatctt	900
cagcctccac cttatggcca ctgcgtgaca gatagccggc tggtgtatag cgtggccat	960
caatgggtcg aaaccccgagg caaacaggcc atgctgttgc catgccttgg caatggccgt	1020
tccctgccaag agacacccgt gacacgacc tatggccggca atagcaatgg cgagccttgc	1080
gtgtgtccctt tccatcttca tggccggacc ttctacatcg gcaccaccc gggtagacag	1140
gacggacacc tgggtgttag caccaccagg aactacgacc agggaccggaa gtacaccc	1200
tgcacccgacc acaccgtgtt ggtggaaatcc agaggccggca atttcaaccc cgccctgtt	1260
cacttccctt ttctgttcaaa caaccacccatcg tacaccgttgc gcacccttgg aaggcagacgg	1320
gacaacatgaa atggttgcgg cacaaccccg aactacgacc ccgatcggaa gttcgcttc	1380
tgtcccatgg ccgtctacca gggaaatctgt accaccaacg aaggcgatgt gtacacaaatc	1440
ggcgccaggat gggaaacccgca gcacccatcg gggccatcg tgccgtgtac ctgcgttggc	1500
aatgggtcgcg gagatggac atgcattccatcg tacaggccatcg tgccgttgc gtcgtcg	1560
gacggatccatca cttatggcca gaaatggccatcg ttccacaaacg ggcacaaaga gggacacat	1620
ctgaactgtca ctgtttcccg ccaaggccatcg gggccgttgg aatgtgtatcc tggatcg	1680
tgcacccggca gogaaaccccg caccatccatcg cagatccggat actccctggaa gaaatgttgc	1740
cacggcgatcg ctatccatcg ggcacccatcg ttggcgatgg gcaatgtcg 1800	
ccctctcgatca cttatggcca ctgcgttgcgg tggttcatcac cgacccatcc	1860
agccatcgatca actctccatcc catccatcg aatgccttccatcg acatccatcgca catcgatca	1920
taatccatcgta gatggggccatcg caaactgtca ctgcgttgcgttccatcg gggccatcc	1980
ggccacccatcg acatccatcg catcaaggaa ctgacccatcg ggcgttgcgttccatcg	2040
ctgatcttca tccatcgatca cggccatccatcg gaaatgttgc gatccatcgatccatcg	2100
tccaccatcg caccatcgatca cggccatccatcg gtcacccatcg agacaacccc tttcttccatcg	2160
ctgtgttgcata caatggatcgatcg ctgcgttgcgg attaccatcgatca gcaatgttgcgttgc	2220
gtgtgtccatcg cgcataatcgatcg ctgcgttgcgg ataggatcgatca gcaatgttgcgttgc	2280
gacggccccc agatccatcgatcg tccatcgatcg ggcacccatcg ggcgttgcgttccatcg	2340
ctgccccatcg gaaatgttgcatcg ctgcgttgcgttccatcg taccatcgatccatcg	2400
ctgatcttca gacccatcgatcg ctgcgttgcgttccatcg taccatcgatccatcg	2460
cacggatccatcg ataccatcgatcg ctgcgttgcgg tggttcaatcg cccggatccatcgatcg	2520
taatccatcgatcg tgcgttgcgttccatcg taccatcgatcg taccatcgatccatcg	2580
acacccatcgatcg tgcgttgcgttccatcg taccatcgatcg taccatcgatccatcg	2640
tacccatcgatcg aagatccatcgatcg tgcgttgcgttccatcg taccatcgatccatcg	2700
actccatcgatcg tccatcgatcg taccatcgatcg taccatcgatccatcg	2760
gtgtgtccatcg caatccatcgatcg gacccatcgatcg gacccatcgatccatcg	2820
gtgtgtccatcg tccatcgatcg taccatcgatcg taccatcgatccatcg	2880
tttgcgttccatcg tgcgttgcgttccatcg taccatcgatcg taccatcgatccatcg	2940
tccacccatcg gaaatgttgcatcg ctgcgttgcgttccatcg taccatcgatccatcg	3000
acccatcgatcg atccatcgatcg taccatcgatcg taccatcgatccatcg	3060
agacccatcgatcg tccatcgatcg taccatcgatcg taccatcgatccatcg	3120
cgtatccatcg tccatcgatcg taccatcgatcg taccatcgatccatcg	3180
gacccatcgatcg tccatcgatcg taccatcgatcg taccatcgatccatcg	3240
gtgtgtccatcg tccatcgatcg taccatcgatcg taccatcgatccatcg	3300
accacccatcgatcg tccatcgatcg taccatcgatcg taccatcgatccatcg	3360
ccttctcgatcg tccatcgatcg taccatcgatcg taccatcgatccatcg	3420
tctggactgtca caccaggatcg gaaatgttgcatcg taccatcgatccatcg	3480

-continued

gaaaaggatg	ccctatcgt	gaacaagggtg	gtcacccctc	tgagcccacc	aacaaacctg	3540
cacctggaa	ccaatctcg	tacccgcgtg	ctgactgtgt	cctggaaag	aagcaccaca	3600
cctgacatta	ccggctatcg	gatcaccaca	acacccacca	atggccaga	gggcaactcc	3660
cttgaagagg	tggtgatgc	cgatcgtcc	agctgtactt	tcgacaact	gagccctggc	3720
ctcgagtaca	atgtgtccgt	gtacacagt	aaggacgaca	aagaaaagct	gcccatctcc	3780
gacacaatca	tccctgaggt	gccccagctg	accgacccgt	gcttcgttga	tatcaccgac	3840
agcagcatcg	gcctggaggt	gacacctctg	aactctctg	ccatcatcg	atacagaaatc	3900
accgtgggtt	ccgtggcg	ggcatccca	atcttcgagg	actttgttga	cagcagctgt	3960
ggctactaca	ccgtgacttg	actggaaacc	ggcatcgact	acgacatcg	cgtcatcacc	4020
ctgatcaatg	ggggcgagag	cgccccata	acactgacac	aacagactgc	cgtgctccct	4080
cetacccgt	tcgggttac	aaatactggc	ccggacacca	tgagagtgac	ttgggtctct	4140
ccaccaagta	tccgacatcg	caacttcc	gtcggtaca	gccccgtgaa	gaacgaggaa	4200
gatgtggccg	agctgtccat	ctctcccago	gataatgcgg	tggtgtgtac	caatctgtc	4260
ccccggaaac	agtagctgtt	gtccgttago	agcgtgtac	aacagacaca	gagcacaccc	4320
ctgcggggca	gacaaaaaaac	aggectggat	ageccccacc	gaatcgact	cagcgatata	4380
acagccaaac	gcttaccctgt	gcattggatc	gccccctaa	ccacccat	cacccgtataga	4440
atccggcata	accccgagca	cttcagcggc	agacctagag	aagatagagt	gccccactcc	4500
agaaacacgca	tacccatctac	caatctcg	cccgccaccc	aatatgttgt	gtccatctgt	4560
gcccgttac	gcagggaga	gtctccatcg	ctgatcgcc	agcagacac	agtgtctgac	4620
gtgcccagag	atctggaaat	ggtggcttgc	acacatccaa	gctctgtat	tagctggat	4680
gtctctgtgt	tgacagtgc	gtattaccgg	atcacatcg	gctgagactgg	cgccaaactct	4740
cccggtcaag	agtttacatgt	gccttggcago	aagagcaccc	ctaccatctc	tggactcaag	4800
ccaggcgttc	actacaccat	tacccgtgt	gctgtgacc	gcggggcga	ttctctgtcc	4860
tcttctaagc	ctatcagcat	caactatccgg	accggatcg	acaagccaa	ccagatgcaa	4920
gtgacagatg	tgcaaggacaa	cageatctcc	gtgaagtggc	tgcttagcag	ctctccatgt	4980
acccggataca	gagtgtaccac	cacaccaaa	aacggccctg	gacctaccaa	gaccaagacc	5040
gtctggatcg	atcagacccg	gatgttgcatt	gaggggctgc	agccttacccgt	cgagttatgtc	5100
gtgtctgtgt	acggcccaaa	tccttgcgg	gagtctc	tccttgc	gaccggctg	5160
accaacatcg	acagacactaa	aggccctggc	ttcacccgac	tggaatgttgc	ctctatcaag	5220
atccgcctgg	agtccctcta	ggggcaagtg	tccagatata	gagtgtacca	cagctccct	5280
gaggacggca	tccacagatc	gtttccatcg	ccagacccgg	aagaggatac	cgccgaactg	5340
caaggactcg	ggccctggctc	cgatataca	gtcttcgttgc	tggccctgc	cgacgacatg	5400
gaaaggccac	ctctgtatcg	aacccatcg	accgctattc	ccgctcttac	agacctgaa	5460
tttaccccaag	tgaccccaac	cagctgttgc	gcacaatgg	ctccctccaa	cgtccatgt	5520
actgggttata	gagtgtcg	gacacccaaa	gaaaagactg	gccccatggaa	ggaaatcaat	5580
ctggccctgt	actccagtc	cgtgggttgc	tctgttctta	tgtgtgttac	caaatacggag	5640
gtttccgtgt	atgcccgtaa	ggacacccctg	acccatccac	ctgcacagg	cggtgttgc	5700
acactggaaa	acgtgtcccc	acccctggaga	gccagagtga	cagacgcac	cgaaaccaca	5760
atcaccat	cttggggggc	caagacagac	acaatccac	gattccatgt	cgacccgtg	5820
cctgccaatg	gacagacacc	tatccatcg	accatcaac	ctgacgttgc	gagctacaca	5880
atcacagggc	tgcaacatcg	caccgactc	aaatctacc	tgtacaccc	gaacgacaa	5940
gcccgtct	gccccgtgg	tatcgatgc	tctaccgc	tcgacgc	ccccgcac	6000
agattttctgg	ccacaaatcg	caacatgttgc	ctctgttgc	ggcagccccc	tccggccaga	6060
atcaactgtgt	atcatctaa	gtacggaa	ccaggccac	ctcttagaga	gggtgtccct	6120
agacctagac	cttgggttc	agggccca	attacccgac	tcgagccccc	cactgagatc	6180
acaatctacg	tgatgtccct	gaagaacaa	cagaatgtcc	agccactgt	tggccggaa	6240
aaaacccgac	acgtgtctca	gtctgttgc	ctgcttcatc	caaatactgc	cgccccggag	6300
atccctggat	tcccatctac	ctgttgcggaa	accccatttgc	tgacacaccc	cggttacgac	6360
acccggaaac	gtatttgcgt	gtctggaa	tctttggc	agccttctgt	ggggacacag	6420
atgatcttt	aggaacacgg	tttccggcgg	accacaccc	ctacaac	cacaccaatc	6480
aggcacaac	cccgccctta	tccttctta	gttggcga	agatccatgt	cgggacacatc	6540
ccctggatgg	atgtcgacta	ccacatgttgc	ccttcacggc	caggactgaa	ccctatgtcc	6600
acgacacggac	aaagggccct	gagccagacc	actatcgat	gggcttccat	ccaggacacc	6660
tccgagatca	tcatcttcttgc	tcacccgt	ggcaccgac	aggaacacc	gcaattcaga	6720
gttccggc	cttcattac	ccggccacact	acaggacttgc	ctagaggcgc	cacctataac	6780
atcatctgttgc	aaagccctgg	agaccac	ccggcaca	tcgcgcga	gggtgttact	6840
gttggcaat	ccgttacac	ggggcttgc	caagccaccc	acgacatgt	cttttgc	6900
tacacagtgt	cccaatatgc	ctgttgcggat	gatgtggaa	gcatgttgc	gagcggcttc	6960
aaactgtct	gtcgtgttgc	ccggccact	tcagatgttgc	tagtccatgg	7020	
ttgttccac	ataacccgt	gaacttacaa	atcgagag	atggggac	acagggcag	7080
aaacggccaga	tgtatgttgc	ccatcttc	ggaaatgttgc	aggggcagg	caagtgcgc	7140
cctcaca	ccacatgttgc	cgacatgttgc	aagacatcc	acgtgggaga	cgagtggcag	7200
aaagagtacc	ttgggc	ccatgttgc	acatgttgc	tcgggc	ccggcttgc	7260
ttgtgtat	ccgttacac	ccatgttgc	ccatgttgc	ccatgttgc	ccatgttgc	7320
tacaaccat	acagccat	ataccac	ccgacacaa	ccatgttgc	ccatgttgc	7380
gagtgttca	tgccccctgg	cttggc	gagacggaa	attcttagaga	gttgc	7434

SEQ ID NO: 37 moltype = DNA length = 2361

FEATURE Location/Qualifiers

source 1..2361

mol_type = unassigned DNA

organism = Homo sapiens

SEQUENCE: 37

atggcggtc	tgacggccggc	ggcccccgggg	ccggagttcc	tcctgttct	gctgtccatc	60
ctccacccct	cttggcctgg	aggggccatcc	ggggccatcc	ctgggttgc	ttctggat	120
gttctttatc	caggggctgg	tcttggc	ggggcgggg	ggggcgggg	ccctgggg	180

-continued

aaaccttta	agccagtcc	cggaggcctt	cggggtctg	gccttgggc	agggctggc	240
gccttcccc	cagtaccct	tccggggct	ctgggtctt	tgggagtggc	tgacgctgtc	300
gcagccata	aagctctaa	ggctggcgct	gggcttggtg	gtgtcccagg	agttgggtggc	360
ttaggagtgt	ctgcaggtc	ggtggtctt	cagoctggag	ccggagtgaa	gcctggaaa	420
gtccgggtg	tggggtgcc	agggtatac	ccaggtggcg	tgctcccagg	agctcggttc	480
cccggtgtgg	gggtgtccc	tgggatccc	actggagcac	gagttaaaggc	caaggctcca	540
ggtgttaggt	gagctttgc	tggaaatccc	ggagttggac	ctttggggg	accgcaaccc	600
ggagttccac	ttgggtatcc	catcaaggcc	cccaagtc	ctgggtggct	tggactggcc	660
tacaccacag	ggaaactgcc	ctatggctat	ggggccggag	gagtggctgg	tgcaagggggc	720
aaggctgtt	acccaaacagg	gacaggggtt	ggccccagg	cagcagcagc	agcggcagct	780
aaagcagcag	aaagtcgg	tgcgtggaca	gcccggatcc	tcctctgggt	tggaggggct	840
ggtgttcctg	ggtgtctgg	ggcaattctt	ggaaatggag	gcatcgagg	cgttggact	900
ccagctcgag	ctgcagctgc	agcagcagec	gctaaggcg	ccaagtatgg	agctgctgca	960
ggcttagtgc	ctgggtggcc	aggctttgc	ccggggatag	ttgggtgtcc	aggagctggc	1020
gttccagggt	ttgggtgcc	aggagatgg	atcccagtgg	teccagggtc	tgggatccca	1080
ggtgtctggc	ttcccaagggt	tgtgtcaca	gaagcagtc	ctaaggcagc	tgcaaaaggca	1140
gccaataacg	ggcccgaggcc	cgggatcgga	gttggaggca	ttccctactt	cgggggttgg	1200
gtctggggct	ttcccccgtt	tgggtgtcga	gtcggggat	tccctggagt	cgcagggtgtc	1260
cctgggtgtcg	gagggtgtcc	cgggatcgga	gggtgtccgg	gagttggcat	tcccccggaa	1320
gctcaggcag	cagtcggcc	caaggctcc	aagtacgggt	ctgcaggagc	aggagtgtc	1380
ggtgggtctag	tgccagggtc	cccaggcgc	gtccccagggt	tgccggggac	gggaggagtg	1440
ccaggagttgg	ggaccccaagg	agctcgatgt	gctaaaggcg	ccgc当地	cgccaggat	1500
gggttagtgc	ctgggtgtcg	cgtggctctt	ggagttggcg	tgcgtctgg	tgcgtgttg	1560
gtccctggag	ttgggttggc	tcctggggat	ggcgtgggtc	ctggagttgg	tgtggctcc	1620
ggcggtggcg	ttggctccgg	cattggccct	ggtggagttg	cagtcgcagc	aaaatccgt	1680
gccaagggtgg	ctgccaataagc	ccagtcggca	gctgcagtc	ggcttgggtc	tggcatcc	1740
ggactctggag	ttgggtgtcg	cgtccctgtt	cttggagttt	tgctgtgtt	tcctggactt	1800
ggagttgtgt	ttgggtgtcc	tgggttgggg	gcagggtcg	atggggatgt	taggcggagc	1860
ctgtccctgt	agctcaggga	aggagatccc	tccctcttc	agcacccccc	cagcacccccc	1920
tcatcaccca	gggttacccgt	agccctggct	gcccgtaaag	cagccaaata	tggagcagca	1980
gtgcctgggg	ttcttgggg	gtctgggggt	ctcggtggag	taggcattcc	aggcggtgt	2040
gtggggagcc	gaccggccgc	cgccgtgtcc	gcagccaaag	ctgctccaa	agccgc当地	2100
tttggccctag	ttgggagccgc	tgggtcgg	ggactcgag	tcgggggat	tggagttcca	2160
gggtgttgggg	gccttggagg	tataacctca	gctgcagcc	ctaaaggcagc	taaatacgg	2220
gctgctggcc	ttggaggttgc	cctatggggat	ggccggggat	tcccacttgg	aggagtggca	2280
gcaagacgttgc	gtctccatt	ttcccaagggt	ggccgtgc	ggggaaagct	a	2340
tgtggccggaa	agagaaaaat					2361

SEQ ID NO: 38	moltype = DNA	length = 2361				
FEATURE	Location/Qualifiers					
misc_feature	1..2361					
	note = Synthetic Construct					
source	1..2361					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 38						
atggctggac	tgacagccgc	tgctctaga	cctgggtttc	tgtgtgtgt	cctgagcatt	60
ctgcaccctt	ctagaccagg	cggagtgc	cttgcatttc	ctggggagat	tccgggggaa	120
gtttttatc	ctgggtgtgg	actggggcc	cttggggag	gtgtcttgg	acctgggtgg	180
aaggctctga	aacctgtgcc	tggggactt	gctgggtcg	gtcttggagc	cggaacttgg	240
gtttttcccg	ccgtgacatt	tcctggcgct	cttgggtcgat	gtggcggtgg	agatgtgtcc	300
gctgttata	aggctggcaa	agccggcgct	ggactcggg	gtgttcttgg	tgttggaggc	360
ctggggatgtt	ctgctgtgtt	tgtgttctt	caacccgtgg	cttgggttta	gcctggccaaa	420
gttccaggcg	ttggactgcc	tgggtttac	cctgggttgtt	ttctccctgg	cgtagattc	480
cctggcggtt	gagttttcc	tggcggtcc	acagggtggcc	gctgtaaacc	taaagctcc	540
gggtgtcgcc	gagcccttc	tggaaatcc	ggcgtgggg	cttttgggg	acctcaacca	600
gggtgtccctc	ttgggtatcc	tataaggcc	cataaacttc	caggcggtca	cgccgtcc	660
tacacaactg	gcaagctgc	ctatggctat	ggacccgtgg	gtgttcttgg	cgccgtgtgg	720
aaagctgtat	accctactgg	aacaggcg	ggccctca	cagcagtc	cgccaggat	780
aaaggcccg	ctaaatctgg	agctggcg	gcaggcg	tggccgggt	tggcgagca	840
ggcgccccc	gtgttccagg	gtcaactt	ggaaatcc	gaattccgg	ctggcgact	900
ccagctcgag	cagccgcagc	cgccgtgtcc	gcaaaagct	ctaaatatgg	cgccgtgtca	960
gggttggcc	caggcggtcc	tggatggaa	ccagggttgtt	ttggagtggc	aggcgctggc	1020
gtcccccggc	ttggaggttgc	tgggtccgg	atccctgtt	ttccctggc	ttgttattcc	1080
ggcgctgtgt	ttccagggtt	tgggtcttcc	gaagccgtt	ccaaggccgc	tgctaaaggca	1140
gtctaaatcc	gtggccggcc	aggcggttgc	gttggcgaa	tcccaacata	tgggtgtgg	1200
ggcgccggat	tcccaggatt	tggagttgg	gtcgaggcc	tcccgggtgt	tgcaaggcgtt	1260
ccaggcggtc	ttggggatgtt	tgggttgggt	gggtgttcc	gctgtggaa	ttctccctgaa	1320
gtccaggcc	ctggccgtcc	caaaggcc	aaatatggc	ctggcggtgc	ttggcggttc	1380
ggaggattgg	ttccgggttgc	tccagggtt	gttccgggg	ttcccgaa	ttggcggtgt	1440
cctgggtgtcg	tgacaccagg	cgctgtcg	gcaaaaggct	ctgtca	aggcgttgc	1500
ggactgtgtt	ctggcggtcg	agtggcc	ggtgggtgg	ttgcaccc	ctggcggtgt	1560
ggccctggcc	ttgggttgc	tcctgggtt	gggtgttccc	cagggtgtgg	agtcgttccc	1620
gggtgtcggt	tcgcacccgtt	tattggtct	ggtggcggtt	ccgcagtc	caaatctgt	1680
gcaaaagggtgg	ccgccaataagc	acagctgaga	gctgtcc	gttccggcc	tggaaatcc	1740
ggactcggttgc	ttggcggtgg	agtgccagg	tttgggttgc	gagcaggcg	gccccggactc	1800

-continued

ggagtccggag	ctgggtgtccc	aggtttttggaa	gctgggtgcag	atgaaggcgt	gcggagatct	1860
ctgagccctg	agctgagaga	gggcgcacccct	agctcttagcc	agcacatctccc	tagcacacct	1920
agcagccccct	gagtcctccgg	ggctcttttgtt	gcagccccaa	ccgcattaaatg	tggggctgtct	1980
ttcccttggtg	ttcttggagg	actttggagca	ctcgccggag	tggaaattcc	aggcggtgtc	2040
gtttgttcgac	gacccgtctc	tgcccccggcc	gctgcaaaaag	cgcacgttaa	ggccaccccaag	2100
tttggccttg	ttggagccgc	tggacttttgtt	ggcctcgagg	ttgggtgtct	ttgggtgtcccc	2160
ggtgttggcg	gacttggagg	aattccacca	ggccgcacag	ccaaaggctgc	taataatccggc	2220
gctgcaggac	tttggcggtt	ttcttggggaa	gctggacagtt	ttccacttgg	aggcgttgc	2280
ggccacacctg	gottttggctt	gtctccattt	tttccctggcg	gctgttgtct	ggccaaaggcc	2340
tgccggcagaa	agcggaaagta	a				2361

```
SEQ ID NO: 39          moltype = DNA  length = 1017
FEATURE                Location/Qualifiers
source                 1..1017
                      mol_type = unassigned DNA
                      organism = Homo sapiens
```

```

SEQUENCE: 39
atgagtc当地 gtc当地 cttcttccctg gcattgattt gttggaccag tggccaggac 60
tatgtatc当地 atttccccct atcaattat gggcaatcat caccaactg tgccaggaa 120
tgttaactgc当地 ctgaaaagctt cccaaaggcc atgtactgtg atgagtc当地 attgaaaagg 180
gttccatgg当地 tgc当地 ctccctgg aatcaatgtt ctttacccctt ggaataaccacca gattggccat 240
attgtatc当地 aggcccttgc当地 gaatgtatc当地 gatctgc当地 ggctc当地 ttc当地 agatccaaac 300
cttcttagaaa当地 actccaaagat aaaaggggaga gtttttccctt当地 aattgaaaaggcc actgtatc当地 360
ctgc当地 atatccaaaacc accacaaccat cccatc当地 gagatccatc当地 tctgtggcc当地 cacttcccaa当地 atctctgg 420
gatctgc当地 cggccatccatc当地 ttactccatca当地 caagatccatc当地 aagctggccatc当地 cttttgc当地 aagg 480
ctgacccttca当地 tccatcttca当地 gcacaatccgg当地 ctgaaaggaggatc当地 atgctgttccatc当地 agtgc当地 tttt 540
aaaggcttca当地 aatcaactc当地 ataccccttgc当地 ttgaggccatc当地 atcagatccatc当地 cagactgc当地 600
tctgttccatc当地 ctgttccatc当地 ttcaacttccatc当地 tactttccatc当地 aacaaatccatc当地 cagaaacatc当地 660
cctgtatc当地 atttcaaggcc当地 tttaatccatc当地 ttgcaatccatc当地 tgc当地 gtgttccatc当地 tc当地 cacaacgaa 720
ctggctgatc当地 ttggaaatccatc当地 ttggaaatccatc当地 ttcaatccatc当地 catccctggatc当地 tgatc当地 gtgt 780
ctgttccatc当地 acaagcttca当地 aaacatccatc当地 actgtatc当地 aaaaccccttgc当地 aaactattac 840
ctggaggccatc当地 atcaacttgc当地 gaatgttgc当地 ataaaggccatc当地 tctgtatc当地 ggatccatc当地 cctggccatc当地 900
ttatccatc当地 ccaagatccatc当地 ggatccatc当地 ttggatccatc当地 atcgc当地 atctccatc当地 agaaaccatc当地 960
cttccaccqccatc当地 atatqtatca当地 atqtatccatc当地 qttqtaatccatc当地 aqtaatctctt当地 taattaa 1017

```

```
SEQ ID NO: 40          moltype = DNA    length = 1017
FEATURE                Location/Qualifiers
misc_feature           1..1017
note = Synthetic Construct
source                 1..1017
mol_type = other DNA
organism = synthetic construct
```

SEQUENCE: 40 Organism = synthetic construct
atgagcctga ggcgttccac actgtttctg gccctgtatcg gggcacaag cggccatgc 60
tacgactacg attccccact gagcatcac ggccaaaggac gccttcaattg cggcccttgag 120
tgcctatggc cggagacta tccttagccgc atgtatgcg acgagctgaa gctgaagtcc 180
gtgcctatgg tgccttcctgg catcaagtcg ctgtacccgtc gggaaacccaatgaccac 240
atcgacgaga aggccttga gaacgtgacc gacctgcagt ggctgtatctt ggaccacaaac 300
ctgctggaaa acagcaagat caaggggcgc gtgttcagca agctgaagca gctgaagaaaa 360
ctgcacatca accacacaaa cctgaccggg agcgtggggc ctctggctaaat gcttcgtggaa 420
gtatcgccgc tgaccaccaa caagatcacc aagctggggc gcttcggggg cctctgtgaac 480
ctgacacctca tccatctgca gcacaaaccc ctgaaaggagg atgcgttag cggccgttcc 540
aagggcctga agagtctgga atacccgtggac ctgagcttca atcagatcgc cagactgcct 600
agccgcctgc ctgtttctct gtcgtacactg tacctggaca acaaaaaatc cagaacacatc 660
cccacgactg acttcacggc ttcaacggc ctgcgtatcc tgagactgag ccacaaacgag 720
ctggccgatgtt ccggcatcccc cggcaacacg ttcaatgtt ccagccgttgg ggaactggac 780
ctgtccataca acaagctgaa aaacatcccc accgtgaacg agaacatcga gaactactac 840
ctgaaagtga accagcttgcg aaggtctcat atcaaaagct tctgttttttccctggggct 900
ctggactcatc ccaagatcaa gacccctggggcc ctggaaaggca accggatcgc cggaaacaaacg 960
cttcccttcgtt atatgttacca ttgttccatggat ttgttccaaacgg aatgttcccttggaaacttgc 1017

SEQ ID NO: 41 moltype = DNA length = 1437
FEATURE Location/Qualifiers
source 1..1437
mol_type = unassigned DNA

```

SEQUENCE: 41
organism = Homo sapiens
atggcacccc tgagacccct tctcatactg gcctgtgg catgggtgc tctggctgac 60
caagagtcat gcaaggcccg ctgcactgag ggcttcaacg tggacaagaa gtgcaggatgt 120
gacgagctct gctcttaacta ccagagctgc tgcacagact atacggctga gtgcagacc 180
caagtgaactc gcggggatgt gttcaactatg ccggaggatg agtacacggc ctagacgat 240
ggcgaggaga aaaacaatgc caactgtccat gaacaggatgg ggggccccctc cctgcaccc 300
gactctccagg cccaggccaa agggaaatctc gagcagacac ctgttctgaa accttgaggaa 360
gaggccctgtc cgccctgaggat gggccgcctc aagccctgagg ggatagactc aaggcttgag 420
accccttcattc caggaggaccc tcagggccca gcaggaggatc agctgtgcgg tggggagccc 480
ttcgacgcctc tcacccgacct caaaaacgggttcccttgcggatcccttgcgg 540

```

-continued

```
tatgaactgg acgaaaaggc agtggggct gggtaaaaaa agtcatacg agatgtctgg 600
ggatcgagg gcccattca tgccgcctt acccgatca actgtcaggaa gaagaccc 660
ctttcaagg gtagtcaga ctggcgtttt gaggatgggg tccggacc tgattacccc 720
cgaatatct ctgacggctt cgatggatcc cccgacaacg tggatgcagc cttggccctc 780
cctggccata gctacagtgg ccggagccgg gtctacttct tcaagggaa acagtactgg 840
gagtaccatg tccagcacca gcccaggcag gaggatgtgg aaggacgctt cctgtcggt 900
gtgtttgaac acttgcatt gatgcagccg gacatgtgg aggacatctt cgagtttcc 960
ttctggggca gaacctctgc ttgttccaga cagccccatg tcattacgg ggactggcac 1020
ggtgtgecag ggcaagtggc cgcagccatg gctggccgca tctatctc aggcacggca 1080
ccccggccctt cttggcccaaa gaaacaaagg tttaggcattt gcaacccaa aggttacccgt 1140
tcacaacggccgacccatggccgaaatcccgccatc gcaacccaaagg 1200
ttggctgtct ttgttccag tgaggagago aacttggggat ccaacaaacta tgatgactac 1260
aggatggact ggcttgcc tggccatctt gaaaccatcc agagtgtttt cttttctct 1320
ggagacaaatg actaccggat caatcttcgc acacggccgg tggacactgt ggaccctccc 1380
tacccacgtt ccatacgctca gtactggctg ggctggcccaatc tctgttag 1437
```

```
SEQ ID NO: 42 moltype = DNA length = 1437
FEATURE Location/Qualifiers
misc_feature 1..1437
note = Synthetic Construct
source 1..1437
mol_type = other DNA
organism = synthetic construct
```

```
SEQUENCE: 42
atggccctt ttagacccct gctgttctg gctctgtgg ctggggttgc cctggccat 60
caagagtctt gcaaggccag atgcaccatgg ggttcaacgg tggacaaagg atgcaggatgc 120
gacgagatgtt gcaatgttccatgg ccacggatgtt tgccaggatcc acacccgcga gtcaagggcc 180
caagtgcacaa gggccatggcgtt ttccatggatcc cctggggatcc agtacaccgt gtacgacgac 240
ggcgaggaaa agaacaacgc caccgtgcac gggcaatctt gggcggccatc tctgtgggg 300
gatctgcagg ctccatggcgtt gggcaatctt gggcggccatc tctgtgggg 360
gaaggccctt ctccatggcgtt gggcaatctt aacgctgtggat gcaatggccatc cagacccgg 420
aacatgtcatc caggcggacc tcacgttccatgg gggcggccatc tctgtgggg 480
ttcgacgcctt tcacccgacccatgg gaaacccggc agcctgttcg ccttttaggg ccagtactgc 540
tacgagatgtt acggaaaggc cgtggccatgg ggtatcttca acgtgtatcc agatgtgtgg 600
ggatcgaggcc gcccattccatgg ccggccatcc accagaatccatc actgtcagggg caagacccatc 660
ctgttcaagg gggccatggcgtt ttggatggatcc gggcggccatc tctgtgggg 720
agaatataatc gggccatggcgtt ccacggatccatgg tggatgttcg tctggccctc 780
cctgctcaca gttactctgg cggccatggcgtt ttaaggggaa acagtactgg 840
gagtaccatg tccagcacca gcccggccatgg gggccatggcgtt cttggccatc tctgtgggg 900
gtgtttggcc acttgcattt gatgcagccg gacatgtgggg aatattttt cgagctgtgt 960
ttctggggca gaaccacggcc cggccatggcgtt cccatggccatc tctatccatc 1020
ggcggtccatgg gacaatgtggc tggccatggcgtt gggccatggcgtt tttttttttt 1080
cttagccatgg gggccatggcgtt tttttttttt gggccatggcgtt tttttttttt 1140
agccagagatggcc gggccatggcgtt cccatggccatgg gggccatggcgtt tttttttttt 1200
ttggctgtcc ttgtttagccatgg cggccatggcgtt aacccatggccatc cggccatggcgtt tttttttttt 1260
cgatggattt ggcttgcc tggccatggcgtt gggccatggcgtt tttttttttt 1320
ggcgacaaatg attaccggatggcc gggccatggcgtt tttttttttt 1380
tacccatggcgtt ccatacgctca gtattggctg ggctgtccatgg cttctggccatgg 1437
```

```
SEQ ID NO: 43 moltype = DNA length = 5175
FEATURE Location/Qualifiers
source 1..5175
mol_type = unassigned DNA
organism = Homo sapiens
```

```
SEQUENCE: 43
atgcctccatggcgtt gtcagccatggcgtt agcatggatgg ggtctgtggat ctttggggca 60
ggccctggggc atgtgttggg ctacatggcgtt cccatggccatgg tccatggccatgg 120
cccggtcaaa gtcacatggcgtt agcggatgtt gtcacatggccatgg tccatggccatgg 180
ctggatcatatccatggccatgg taaaggatggcgtt tccatggccatgg ggtctgtggat 240
aacggacattt cccatggccatgg ccacggatgtt gtcacatggccatgg tccatggccatgg 300
accggccggatccatggccatgg acggatgtt gtcacatggccatgg tccatggccatgg 360
tcctggccatgg cccatggccatgg tccatggccatgg aacatgtggatggccatgg tccatggccatgg 420
ggggggggatccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 480
gcacccggatccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 540
aacacggatccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 600
cccaaaatggatccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 660
ctgaacccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 720
ctgatgtccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 780
ctggatgtccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 840
ggccctggggca gggccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 900
caagatgtccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 960
gcaaaatggatccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 1020
cagatgtccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 1080
tggctgtccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 1140
agagaacggatccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 1200
caaaaaatggatccatggccatgg tccatggccatgg ggttacacggatccatggccatgg tccatggccatgg 1260
```

-continued

gaatacgaag	ccaaactcgag	tgaccttcgt	gctcggctgc	aggaggcgc	tgcggcaagcc	1320
aagcaggcaa	atggcttcaa	ccaaagaaaaac	gagagagctt	ttggagocat	tcagagacaa	1380
gtgaaagaaa	taaattccct	gcagagtat	ttcaccaat	atctaaccac	tgcagactca	1440
tctttgtgc	aaacaacat	tgcgtcag	ctgatggaga	aaagccagaa	ggaatatgaa	1500
aaatttagct	ccagttaaa	tgaagcaaga	caagaactaa	gtgacaaatg	aagagaactt	1560
tccagatctg	ctggcaaaac	atccccgtg	gaggaggcag	aaaagcacgc	gcggccctta	1620
caagagctgg	caaaggcagct	ggaagagatc	aagagaaaaac	ccagcgggaa	tgagctgg	1680
cgtgtgtgt	ttgatgccgc	cacccctac	gagaacatcc	tcatgcacat	caaagcggcc	1740
gaggacgcag	ccaacaggc	tgccagtca	tctgaatctg	ccctccagac	agtgataaag	1800
gaagatctgc	caagaaaaag	taaaaccctg	agttccaaac	gtgataaact	gttaaatgaa	1860
gccaagatgc	cacaaaaagaa	gctaagcaa	gaagtca	cagetc	caacccatcag	1920
caaacccctga	atattgtac	agttcagaaa	gaagttagat	acaccaatct	cacaacttc	1980
cgagatggc	ttcatggat	acagagaggt	gatattgtat	ctatgtat	tagtgc	2040
agcatggtca	gaaaggccaa	cgacatcata	gatggggct	tggatgggt	caacccatc	2100
cagacagatg	tggaaaagaa	taaggacacc	tatggggaga	cacagaacga	agacttcaaa	2160
aaggctctg	ctgtatcaga	taactcggt	aataagtta	ccaacaaact	acctgtat	2220
tggcgcaga	ttgaaaagtat	caaccaacag	ctgtgcctt	ttggaaacat	ctctgacaa	2280
atggacagaa	tacgagaact	aattcagcag	gcccagat	ctgcaat	ggtgtctgtc	2340
cccatgggt	ttcatggtaa	atctggat	gaagtccgac	tgccaaat	cttggaaat	2400
ttgaaaggat	atacatctc	gtcctgtt	ctccaaaggc	ccaactca	agaaaaatgg	2460
ggtactgaga	atatgttgc	gatgtaccc	ggaataaaag	atgcctcc	ggactacatc	2520
ggcatggcag	tttgtggatgg	ccagtcata	tgtgtctaca	acctgggg	ccgtgaggct	2580
gaacttccaa	ttggaggat	cttgacca	agtggacta	aggggcact	tatggatcg	2640
gtgaaatttca	agagaattt	tcagttca	aggcttaat	acaccaagg	agccacatc	2700
agtaaaccag	aaacaccccg	agtctat	atggatgtt	gaaatagca	tacactc	2760
aatttggatc	ctgaaaatgt	tgtat	gttggaggat	acccaccc	ttttaaactt	2820
cccaatgcac	taatgttccc	tccatcaaa	ggttattt	aatttagat	cctcaat	2880
aatgttctga	ggttgcataa	cttcaaaaaa	atccat	tcacaaac	tgaatgg	2940
ccttgtagaa	ggggaaagga	agagtcagac	aaaaatttt	ttgaaggat	gggtatgt	3000
cgagttccaa	ttcaaccaca	tgctccatc	ccaaacctt	gacagacaat	tcagaccacc	3060
gtgatagag	gttgcgtgtt	cttgcgaa	aacggggatc	gcttcata	tctaaatata	3120
gaagatggca	agctcatgtt	gagataaaa	ctgaattca	actccat	agagagagga	3180
gttggagacg	ccataaaacaa	cgccagagac	cattcgattc	agatcaaaat	ttggaaactc	3240
caaaggcgt	tgtggataaa	tgtggacgtt	caaacaacta	taattgtat	tgaatgtt	3300
gatttcagca	catattatct	ggggat	ccaaatgc	tcagggaa	atttaaat	3360
tctacgcct	cttccgagg	ctgtat	aatttga	aaacccat	tttgcgtt	3420
ttgaatgata	ctgtgggat	aaccaaaaag	tgctcgaa	acttggaa	tgtgcgtat	3480
gcctcattct	ccagaggagg	acaattgat	tttactgtt	ttggcttacc	acctactgac	3540
caacctccagg	cctcatttt	atccat	tttcaaccc	gtggcat	attatgtat	3600
cagacatgg	caaggaacat	gcaggat	tcggat	tttgcgtt	attggcacc	3660
agcgatagc	ggggccaaat	ttttaaatct	ccacagat	atatgtat	tttgcgtat	3720
tatgtatctg	taataagcga	caactctg	ctacggcttc	tcatcgat	ccagcttct	3780
agaaatagca	aaaggctaa	acacat	agttccccc	agttctgtc	tctgggg	3840
agcaattttgc	agggttgcata	tgcaat	tttgc	gggttact	gatgttgc	3900
gttcttagat	tgaccat	ctcttc	agagatgtt	ccctgggg	ctgcgtt	3960
aacaaaccac	cttttcaat	gttgc	tttgc	tttgc	tttgc	4020
tttcgtatca	accatgttgc	tgccat	ccatggcc	cccaacgg	ctgtaaat	4080
ttggcaatgt	tttgc	acttccaa	accaggc	atcatgg	tttgc	4140
ggggacatc	ccacccatc	ttgttattc	aaatcttcc	aggagctgt	tttgc	4200
tcacagt	ttgtggat	tcacat	tttcc	aggactgt	tttgc	4260
actaaacat	cttttatgc	tcttattat	tcacaa	tttgc	tttgc	4320
acagatgg	aaaatttgc	gtcaat	tttgc	tttgc	tttgc	4380
acgggtgttgc	ttggccat	ttggggaa	tttgc	tttgc	tttgc	4440
cgggagg	tttgc	aaatccac	atcagat	tttgc	tttgc	4500
tcacccat	caggaaacc	aaagggc	cccaacaa	tttgc	tttgc	4560
aacttc	ttgtat	ccat	tttgc	tttgc	tttgc	4620
ttgttgcgttgc	ttcccttgc	tttgc	tttgc	tttgc	tttgc	4680
ttggctact	ttgtat	tttgc	tttgc	tttgc	tttgc	4740
agttctact	ttgtat	tttgc	tttgc	tttgc	tttgc	4800
ctggaggc	tttgc	tttgc	tttgc	tttgc	tttgc	4860
gtcacaccaa	tttgc	tttgc	tttgc	tttgc	tttgc	4920
caacacatcc	tttgc	tttgc	tttgc	tttgc	tttgc	4980
ttcccacat	tttgc	tttgc	tttgc	tttgc	tttgc	5040
acactgg	tttgc	tttgc	tttgc	tttgc	tttgc	5100
cacatcc	tttgc	tttgc	tttgc	tttgc	tttgc	5160
tttgc	tttgc	tttgc	tttgc	tttgc	tttgc	5175

```

SEQ ID NO: 44      moltype = DNA  length = 5175
FEATURE
misc_feature
source
SEQUENCE: 44
atgcctcctg ctgtgcggag aagcgcctgt tctatggat ggctgtggat ctttggcgcc 60

```

-continued

gctctggac	agtgtctggg	ctactttct	cagcagcagc	gggtgcatt	tctgcagcca	120
cctggacagt	cttcagtgca	ggccgcgtac	gttgaattca	gacctagcca	gggtgttagc	180
cccggtact	acagagatca	caaggccctg	tacaccggca	gatgcgtgcc	ctgcaactgt	240
aacggccaca	goaaccagt	tcaggacggc	tctggcatct	gctgtaaactg	ccagcataat	300
actgcggcg	agcaactgcga	gagatccaa	gagggtact	acggcaatgc	cgtgcattgc	360
agotgtcggg	cttgccttg	tcctcacaco	aacagcttgc	ccacccggctg	cgttgtgaac	420
ggcggagatg	ttcgggttgc	tttcaaggcc	ggctacacag	gcacacagt	cgaaaagatgt	480
geccctggct	actttggcaa	ccctcagaag	tttggcggt	ctgtccagcc	ttgtccctgc	540
aattctaatt	gccagctggg	ctcttgcac	cctctgacc	gctgactcat	caatcaagag	600
cctaaggaca	gcaagccctgc	cgaggatgc	gacgatttgc	atagtcgt	gtgacccctg	660
ctgaacgaca	ttggccaaat	gggagaacag	ctggggctgg	ttaagagcc	gttcagggaa	720
ctgtctgcct	ctgctgact	gttgcacag	atgcggcaca	tggaaaccca	ggccaaggac	780
ctgagaaacc	agctgtgaa	ctacagaagc	gcccattcc	accacggcag	caagatcgaa	840
ggcctggaaa	gagagctgca	cgccatgtac	caagagtgc	agacactgca	agagaaggcc	900
caagtgaaca	gccccaggc	ccagactctg	aacaacaacg	tgaacccggc	cacacagtcc	960
gccaaaggaaac	tggacgtaa	gatcaagaa	gttgcacca	acgtgcacat	cctgtcaag	1020
cagatcagcg	gcacagatgg	cgaggccaa	aatgtgccta	ggggcgactt	tagcagagag	1080
tgggcccgg	cttcagggat	gatgcggag	ctgcggaa	ggaaaccttgc	caaggcacctg	1140
agagaaggcc	aggccgacaa	gagagagac	caactgtgc	tcaaccggat	cagaaccctgg	1200
cagaaaaacc	accaggccga	gaaaacccg	ctggccaa	gcatcagaga	cagctgtaa	1260
gagtaacgagg	ccaaagctgag	cgatctgcgg	gcccagactt	aagaagctgc	cgtcaggccc	1320
aagcaggcoca	acggccctaa	tcaagagaa	gagagagccc	ttgggcggat	ccagagacaa	1380
gtgaaagaga	tcaacagcct	gcaagcgtac	tttacca	acttgcacac	cgccgatagc	1440
agccctgtgc	agccatataat	cccccgtac	cttgcgtaaa	agagccggaa	agatcagaa	1500
aagctggccg	ccagcctgaa	cgaggccagg	caagaactgt	ctgacaaatgt	gcgcggactg	1560
agcagatctgc	ccggcaagac	atcttgcgt	gaagaggccc	agaagcacgc	cagatctctg	1620
caagagctgg	ccaaacagct	ggaagagat	aagcggaa	ccagcggcga	cgaactctg	1680
agatgtcgac	tggatcgccg	caccgcctac	gagaacatcc	tgaatgcct	caaggccgccc	1740
gaggacgccc	ctaatagagc	cgcttctgt	tctggtctg	ccctgcagac	cgttgcataaa	1800
gaggacctgc	ctagaaaggc	caagacacty	agcagcaaca	gcgacaaact	gctgaatgt	1860
gccaaggatgc	acttgcggaa	acttgcggaa	gagggttccc	ctgcactgaa	caaccctgcag	1920
cagacccctg	acatctgtac	ctgtgcggaa	gaagtgtatc	acaccaact	gacaaccctg	1980
agagatggcc	tgcacggat	ccagagaggg	gacatcgac	ccatgtatc	cagcgcacaa	2040
agcatgggtc	gaaaaggccaa	cgacatcaco	gacgagggtc	ttggacggct	gaatccatc	2100
caagaccgc	tggacggat	caaggacacc	tacggccagaa	cccagaacga	ggatttcaag	2160
aaggcccgt	ccgacgccc	caactccgt	aacaagctg	ccaacaagct	gcccgcattc	2220
tggccggaa	tccggagcat	caaccagaa	ctgtccctc	ttggcaacat	cagcgcacac	2280
atggacagaa	tccgggaaact	gatccagcg	gccagagatg	ccgcctccaa	agtggctgt	2340
cccatagat	tcaacggccaa	gaggcgatgt	gaagtgcggc	tgccccaa	tctgaaagat	2400
ctgaagggtc	ataccagcct	gagctgttc	ctgcagggc	ccaacagcg	agagaatggc	2460
ggcaccggaa	atatgttcgt	gatgtacctg	ggaacacaag	acgcacccgg	ggacttatatc	2520
ggaatggccg	ttgtggacgg	ccagctgacc	tgctgttaca	acctgggaga	cagagaagct	2580
gaactcgagg	tccggacat	cctgtggat	agggagacaa	aaagggccgt	gtggacacaa	2640
gtgaagttcc	acgggatcta	ccatgttgc	cggtgttata	acaccaaggc	ggccacaa	2700
agaaggcc	aaacacactgg	cgtgtacac	atggacggcc	ggaaacttca	cacttgcgt	2760
aatctggacc	ccgagaacat	ggtgtttac	gttgcgggt	accctctgt	cttcaagctg	2820
cctagcagac	ttagcttccc	acccatcaaa	gggtgtatcg	agctgttgc	tctgaacgaa	2880
aacgtgtcg	ccctgttacaa	cttcaaaa	acccatcaac	tgaacaccac	cgagggtggaa	2940
cctgtcaggc	ccgaaaaaaa	ggatccac	aagaactact	tgcggacac	cggttgcac	3000
agagtgcata	cacaacactca	cgctccatt	cctacccctc	gccagaccat	ccagacaacc	3060
gtggatagag	gtctgtgtt	cttgcggcc	aacggcgcac	gattcatc	cctgttata	3120
gaggatggca	agctgtgtt	cgatgttac	ctgatagcc	agctgtccaa	agaaaaggcc	3180
gttggcggac	ccatcaacaa	cggtggat	cacagcatc	agatcaagat	cgccaaactg	3240
cagaaacgg	tgtgttacaa	ctgtggacgt	cagaacacca	tcatcgac	cgagggtttc	3300
gacttcagca	cctactatct	cggtggat	cctatcgcca	tcaagatgc	gttcaatatc	3360
agccccctg	ccttccgggg	ctgtatgc	aaactgtaaa	agaccaggc	cgtcgtgcgg	3420
ctgaatgata	cagtggccgt	gaccaaaag	tgcggacgg	actggaaatgt	tgtgcggagc	3480
ccgagttttt	ctagaggccg	acagctgac	tttaccgacc	ttggacttcc	tccttaccat	3540
catctgcagg	caagcttcgg	ttccagccaa	ggggaaatct	gctggacccac	3600	
cagacccctg	ccgaaaaacc	gcaagtggac	tccggatgt	gctatcatcg	actggacacc	3660
agcgactctg	ccggccctat	ccttacagat	ccctcagacc	acatggatgg	gctgtgcac	3720
tacgtgtccg	tgatcagcga	taacagccgc	ctgagactgc	tgatcagcga	ccagctctg	3780
cggaacacgca	agccggctgaa	gcacatctcc	agcagcagac	agagctgag	actcgccgc	3840
agcaatttcg	agggtgtat	cagcaacat	ttctgtgtt	aaagactgt	gtctccagaa	3900
gtgtgtggacc	tggaccat	tagctgtat	agggtgtgt	cttcggccgg	ctgtccctg	3960
aacaaaccc	cttccctgtat	gctgttacat	ggcagacc	ggttcaacaa	gaccaagacc	4020
tttccggatca	atcagctgt	ccaggacaco	cctgtggct	gcccataa	cgtgaaatgt	4080
tggcaggacg	cctgcgttcc	cctgcctaaa	acacaggcca	atcaggggc	tctgcgttcc	4140
ggcgatatcc	ccacaaggcc	tctgtgtt	aaagctggccc	aaagactgt	caagctcg	4200
agccatgttc	ctgtggat	cgagacc	tcctcccgat	gacttgggtt	tcacaccggc	4260
accaagaaca	gttcatggc	cctgttac	agcaaaaggca	ggctgggtt	tgccctgggc	4320
accgacggaa	agaaaactgcg	gatcaagac	aaagagaatgt	gcaacgcac	caagtggcc	4380
accgtgttgt	tccggacacg	tggcgagaaa	ggcagactcg	ttgtggatgg	cctgagagcc	4440
agagagggat	ctctgcctgg	caactccatca	gagccctgt	gtatctggc	4500	
agccctctca	gcggaaagcc	taagacgt	cctaccaact	ccttcgtgg	ctgttgcag	4560
aactttcagc	tggacagacaa	gcctctgtac	acccttagca	gcagtttgc	cgttgcctcc	4620

-continued

tgtctcgag	gcctctgga	aaaggcata	tacttctctg	aggaaggcgg	ccacgttgc	4680
ctcgccatt	ctgttcgtc	ggggcccgg	tccaagctgg	tgttcttat	ccggcctaga	4740
agoctgaccg	gcatctgtat	teacatcggo	agccagcttg	ggaagcacct	gtgtgttat	4800
ctcgaggccg	gaaaaatgac	cgcacatgt	gattctggtg	ctggcgccac	aagcacctcc	4860
gtgacaccta	agcagaccc	gtgtatggc	cagtggcaca	gtgtggccgt	gacaatcaag	4920
cagcacatc	tgcacatgg	actggacacc	gacagcagct	ataccgcgg	acagatccca	4980
tttcctccag	ccagcacaca	agagctctg	caccttggag	gegcctctgc	caatctgacc	5040
acactgaa	tccctgtgt	gaagtcctc	ttcggctgc	tgccggatat	ccatgtgaac	5100
cacatccag	tgccctgtac	agaggccctg	gaagtgcagg	gaccctgtc	tctgaatgga	5160
tgccccgatc	agtga					5175

SEQ ID NO: 45	moltype = DNA	length = 3519				
FEATURE	Location/Qualifiers					
source	1..3519					
	mol_type = unassigned DNA					
	organism = Homo sapiens					
SEQUENCE: 45						
atgagaccat	tcttctctt	gtgtttgcc	ctgcctggcc	tcttgcacgc	ccaaacaaggc	60
tgctccctgt	gggcgtcta	tccacctgtt	ggggacctgg	ttgttgggg	gaccgggttt	120
ctccgagtt	catctacctg	tggactgacc	aaggctgaga	cctactgcac	ccagttatggc	180
gagtggcaga	tgaaaatgt	caagtgtac	tccaggcgc	ctcacaacta	ctacagtac	240
cgagtagaga	atgtggctt	atccctccgg	cccatggcgt	ggtggcagtc	acagaatgtat	300
gtgaaccctg	tctctctgc	gctggaccc	gacaggat	tcacatgttca	agaagtcatg	360
atggagttcc	aggggcccat	ggccggccgg	atgctgatgg	agcgcttc	agacttcgg	420
aagacctggc	gagtgtacca	gtacctgggt	gccgactgca	cctccaccc	ccctcgggtc	480
cgcagggttc	ggccctcragag	ctggcaggat	gttcgggtcc	agtccctgcc	tcagaggcct	540
aatgcacccg	taatgggggg	gaaggctcaa	cttaacctt	tggatttagt	gtctgggat	600
ccagcaactc	aaagtcaaaa	aatggcaagag	gtggggggaa	tcacaaactt	gagagtcaat	660
ttcaccaggc	tgccccctgt	gccccaaagg	ggctaccacc	ctcccaagcgc	ctactatgt	720
gtgtccccc	tcgcgtctgc	ggggagctg	tctgtcaccg	gccatgtca	tcgctgcgca	780
cccaaccctg	ggccctctgc	aggggccatc	acggctgtgc	aggtccacga	tgctgtgtc	840
tgcagcaca	acactggccg	cccaattgt	gagggctgt	cacccttca	caacaacccg	900
ccctggagac	ccgcggagggg	ccaggacgc	catgaatgc	aaaggtgcga	ctgcaatggg	960
caactcagaga	catgtcactt	tgaccctcg	gtgttgc	ccagccagg	ggcataatgg	1020
gtgtgtgt	acaattgcgc	gggacaccc	gaaggcaga	actgtgaggc	gtgtcagctg	1080
caactatttcc	ggaaaccggc	ccggggagct	tccattcagg	agacatgc	ctcctgcgag	1140
tgtgatccgg	atggggcagt	gccagggggt	ccctgtgacc	cagtgacccg	gcagtgtgt	1200
tgcaaggagc	atgtgcaggg	agagcgtgt	gacatatgc	agccgggctt	cactggactc	1260
acctacgoca	acccgcgggg	ctggccacc	tgtgactgca	acatctggg	gtccgggagg	1320
gacatcgctg	tgacggagga	gatggggcgc	tgccttgc	tgcccaactg	gttgggtccc	1380
aaatgtgacc	atgtgtctc	ctacactgg	aaactggcca	gtggccagg	ctgtgaacccg	1440
tgtgcctcgc	acccgcacaa	ctccctcago	ccacagtgc	accagtac	agggcagtgc	1500
ccctgtcggg	aaaggcttgg	tggctgtat	tgccgtgt	caggcgtcc	ccagtgtcc	1560
gaccggacat	atggagacgt	ggccacagg	tgccgagcct	gtgactgt	tttccgggaa	1620
acagaggggc	ccggctgcga	caaggatca	ggccgtgc	tctggccccc	ttggctgtacc	1680
ggggccccc	gtgaccagtg	ccagcgagg	tactgtatc	gttcccggt	gtgcgtggcc	1740
tgcaccatccc	gttccagatc	ctatgtatgc	gacccctccgg	agcaggccct	ggcttgggt	1800
agactccgc	atggccacgc	caggctgg	teaggccgt	ggctggagga	ccgtggct	1860
gcctcccgaa	tcctagatgc	aaagtagta	attgagcaga	tccgagcagt	tctcagcagc	1920
cccgcaigtca	cagacccagg	ggtggtctag	gtggccagt	ccatctctc	cctcaggcga	1980
actctccgg	gcgtcgacgt	ggatctggc	ctggaggagg	agacgttgtc	ccttccggaa	2040
gaccctggaga	tgcttgacag	aaaccttcaat	gttcttgc	ctatgtatca	gaggaagagg	2100
gagcagtttgc	aaaaataaag	cagtgtat	ctttcaggag	ctttccggat	gtctgacaca	2160
gcctacgagc	agtccggcca	ggctgtctag	caggctccg	acagctcg	cctttggac	2220
cagctcgagg	acagccggag	agaggcagg	aggctgtgc	ggcaggccgg	aggaggagg	2280
ggcacccggca	gccccaaatg	tgtggccctg	aggctgg	tgttcttgc	gtctgacactg	2340
acacccaccc	tcaacaatgc	ctgtggca	tccaggcaga	tggctgcac	cccaatata	2400
tgccttggtg	agctatgtcc	ccaagacaat	ggcacagcc	gtggctcccg	ctgcagggg	2460
gtcttccca	ggccgggtgg	ggccctcttg	atggcggggc	aggtggatg	gcagctcg	2520
ggcttcaatg	cccgatctca	gcccggcagg	catatgtat	gggcacccga	ggaatctgc	2580
tcacatcgatc	aatccatgtc	ccagcgcttgc	gagacccagg	tgagccgc	ccgtccccc	2640
atggagggaa	atgtcagacg	cacacggctc	ctaattccac	aggccgg	cttcttaaca	2700
gaccggacata	ctgatcgac	cactatccag	gaggctcagc	aggccgtgt	ggccctgtgg	2760
ctgcccacac	actcaatct	tgttctcg	aaatgtat	gatccaggc	cattgcagcc	2820
aggetccccc	acgtgttccg	gggtgtgtc	ggatccagg	aggatccatgc	ccgtggcc	2880
cggttgcagg	ctgaggctga	ggaaggccagg	agccgagcc	atgcgtgt	ggccagggt	2940
gaagatgtgg	ttggaaacct	gccccgggg	acgtggcac	tgccggaa	tcaggacacc	3000
atgcacggca	ccagccgctc	ccttcggctt	atccaggaca	gggttgc	gttgcagcag	3060
gtactcgcc	cagcaaaaa	gtctgtaca	agcatgacca	agcagctgg	tgacttctgg	3120
acacggatgg	aggagctccg	ccaccaagg	cggcagcagg	gggcagaggc	agtcaggcc	3180
cagcagcttgc	cgaaagggtgc	cagcagcag	catgtgt	cccaaggagg	atttgcag	3240
ataaaacaaa	aatatgtat	gttgcaggac	cggttgggtc	agatccat	gtctgggt	3300
cagggtgtcc	ggatccagag	tgtgcagaca	gaggcagagg	agctgttgg	ggagaccat	3360
gagatgtgg	acaggatgaa	agacatggag	ttggagctgc	tgccggccag	ccaggccatc	3420
atgctgcgt	cagccgaccc	gacaggactg	gagaacgtgc	tggagcagat	ccgtgaccac	3480
atacatgggc	cgctgtctca	ctatgcacc	tgcaagt			3519

-continued

SEQ ID NO: 46 moltype = DNA length = 3519
 FEATURE Location/Qualifiers
 misc_feature 1..3519 note = Synthetic Construct
 source 1..3519 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 46
 atgaggccct tcttctgtct gtgccttgcc ctgcctggac tgctgcata tcagcaggct 60
 tgttagcaga ggcgcgtgta tccttcgtgt ggcgcatactgc ttgtggccag aaccaggatt 120
 ctgcggggca gcttcatatcg cggcgtgaca aagctgtgaa catactgcac ccacgtacggc 180
 gagttggcaaa tgaatgtgtc caagtggcgc agcagacagc cccacaacta ctacaggcc 240
 agatgtggaaa acgtggccatc cagcggcgc cctatgtatggat ggtggccagag ccagaacgac 300
 gtgaaccccg ttatgcgtca gctggacatcg gacagacgggt ttcaatgcata agaaatgtatg 360
 atggaatttc agggccccat gcctgcggc atgtgtatcg agagaacgac cgatttcggc 420
 aagacatgttggg ggtgttacat gtatgtggc ggcgcgtgca ccacgtacat cccacatgtt 480
 agacaggcga gaccccgatc ctggcggatc gttcgtgtc aatgtgtcc ccacggccgt 540
 aacgtatcagc tgaatggccg aaaggtgcgc ctcaatctga tggacatgtt gtctggcatc 600
 cctgcccacac agtccccagaa aatccaagaa gtggggcaga tcaccaatctt gagatgtgaa 660
 ttccatccggc tggccatccgtt tcctcgatc gatgtatcata ctccatgcgc ctatcagcc 720
 gtgttcgtatc ttagactgca gggcgtgttcc ttctgtcgtc ggcacatgtca tagatgtcc 780
 cttaaacctg gtgcctctgc cggacatctt acacgtgtc aatgtgcacga tggatgtgt 840
 tgccagcaca ataccggccg acctaactgc gagatgttgc ccccttctta caacaacccgg 900
 ccttggggcgc ctgcgttgcggc acaggatgtc cagatgtcc agatgtgcgc ctgcgttgcgc 960
 cacagcgaga catgcacatc tgacccatgc ttgtttccgc ttctcgaggc cgatgtatggc 1020
 ggcgtgtgtt acaactgtcag agatcacatc gagggttgcaga actgcgtatgc ctgtatgt 1080
 cactacttc ggaatagaag gccaggccgc aacatccaag agacatgtcat cagatgtcg 1140
 tgcgtatcccg atgtgtgttcc ttctgtggcgtt ccttgcgtc ctgtgtatgc ccatgtgt 1200
 tgtaaagaaac acgtgtgggg cggaaatgtc gacatgtcgc acatgtggctt taccggctt 1260
 acatccgcgc atccctcgaggc ctgcgttgcgc tttttgtgc acatctgggg ctagatgtggc 1320
 gacatgcctt gtgtatgaa gtcgttgcgc tgcgtgtgc tgcctaatgt cgtggggccc 1380
 aagtgtcgatc atgtgtgtccca atatcattgc aatgtgttcc ttgtgtccggc atgcgtatgtt 1440
 tttgtgtgtcgtc atccctccaa cggcgtgttcc ttgtgtgtccgtt ccacatgtca accatgtcc 1500
 ctttgcgtatc aagggttttgc cggcgtgtatc ttgtgtgtccgtt ccgtatgtatc acatgtgtcc 1560
 gatagaacat atggcgactt ggcacatggc tgcgtatgtt gcatgtgttgc ttccgggg 1620
 acagaaggcgc cccgtgtgcga taaggccgc ggaatgtc ttgtgtccggc tggactcaca 1680
 ggccccatgt gtgtatgttgc tggccggccg tactgcacaa gataatgttgc ttgtgtggcc 1740
 tggccatctt ctttgcgtatc ctttgcgtatc gatgtatgttgc ttgtgtgtggcc 1800
 agactgtggaa atgcgtatgttgc ctttgcgtatc gatgtatgttgc ttgtgtgtggcc 1860
 gcttccggaaatc ttttgcgtatc gatgtatgttgc ttgtgtgtggcc 1920
 ccgtatgtatc gatgtatgttgc ttgtgtgtggcc 1980
 atctgtcgatc gatgtatgttgc ttgtgtgtggcc 2040
 gatctggaaatc gatgtatgttgc ttgtgtgtggcc 2100
 gggccatgttgc gatgtatgttgc ttgtgtgtggcc 2160
 agactgtatgttgc ttgtgtgtggcc 2220
 gatctggatc gatgtatgttgc ttgtgtgtggcc 2280
 ggaacatgttgc gatgtatgttgc ttgtgtgtggcc 2340
 accccatctt gatgtatgttgc ttgtgtgtggcc 2400
 tttttgcgtatc gatgtatgttgc ttgtgtgtggcc 2460
 gatctggatc gatgtatgttgc ttgtgtgtggcc 2520
 gatctggatc gatgtatgttgc ttgtgtgtggcc 2580
 agccatgttgc gatgtatgttgc ttgtgtgtggcc 2640
 atggaaatgttgc gatgtatgttgc ttgtgtgtggcc 2700
 gatectgtatc gatgtatgttgc ttgtgtgtggcc 2760
 ctgcgtatccatc gatgtatgttgc ttgtgtgtggcc 2820
 agactgtatgttgc ttgtgtgtggcc 2880
 aggtgtatc gatgtatgttgc ttgtgtgtggcc 2940
 gggccatgttgc gatgtatgttgc ttgtgtgtggcc 3000
 atgcgtatgttgc ttgtgtgtggcc 3060
 ctttgcgtatc gatgtatgttgc ttgtgtgtggcc 3120
 acgtatgtatc gatgtatgttgc ttgtgtgtggcc 3180
 caacaaatgttgc gatgtatgttgc ttgtgtgtggcc 3240
 atcaacatgttgc gatgtatgttgc ttgtgtgtggcc 3300
 agtacgtatc gatgtatgttgc ttgtgtgtggcc 3360
 gaaatgtatgttgc gatgtatgttgc ttgtgtgtggcc 3420
 atgtgtatc gatgtatgttgc ttgtgtgtggcc 3480
 atcaacatgttgc gatgtatgttgc ttgtgtgtggcc 3519

SEQ ID NO: 47 moltype = DNA length = 3582
 FEATURE Location/Qualifiers
 source 1..3582 mol_type = unassigned DNA
 organism = Homo sapiens

SEQUENCE: 47
 atgcgtgtgc tcttgcgttgc ctgcgtgtatc ttcttcgtatc cggccggccgg 60

-continued

gccacacctca	ggagggaaagt	ctgtgattgc	aatgggaagt	ccaggcagtg	tatcttgat	120
cgggaacttc	acagacaaac	tggtaatggc	tcccgctgca	tcactgcgg	tgacaacact	180
gatggcatcc	actgcggaaa	gtgcaagaat	ggcttttacc	ggcacagaga	aagggacccg	240
tgtttgcct	goaattgtaa	ctccaaagggt	tctcttagtg	ctcgatgtg	caactccgg	300
cggtgcaagct	gttaaaccagg	tgtgacagga	gccagatgcg	accgatgtct	gccaggcttc	360
cacatgtc	cggtatgcggg	gtgcacccaa	gaccagagac	tgcttagactc	caagtgtgac	420
tgtgacccag	ctggcatacg	aggcccgt	gacgcggg	gctgtgtctg	caagccagct	480
gtcaactgg	aacgtgtgt	tagtgtgtca	tcagggtact	ataatctgga	tggggggaaac	540
cctgagggt	gtacccagtg	tttctgtat	gggcattca	ccagctgcg	cagctctgca	600
gaatacagt	tccataaagat	cacccatacc	tttcatacaat	atgttgcgg	ctggaaaggct	660
gtccaacgaa	atgggttcc	tgcacaaat	caatggtcac	agcgcacata	agatgtgtt	720
agotcagccc	aacgactaga	ccctgtctat	ttttgtggctc	ctgccaatt	tcttggaaat	780
caacagggt	gctatggtca	aagectgtcc	tttgactacc	gtgtggacag	aggaggcaga	840
caccatcttgc	ccccatgtgt	gattctggaa	ggtgctgttc	tacggatcac	agctcccttg	900
atgcacttgc	gtcaagacat	gccttgcgg	ctcacaaga	tttacacatt	caggtaaat	960
gagcatccaa	gtaataattg	gagccccca	ctgagttact	ttgagttatcg	aaggtaactg	1020
cggaatctca	cageccctcg	catccgagct	acatatggag	aatacagtac	tgggtacatt	1080
gacaatgtga	ccctgtatcc	agccccccct	gtctctggag	ccccagcacc	ctgggttgaa	1140
cagtgtatata	tcattctgtgg	gtacaagggg	caattctgc	aggattgtgc	ttctggctac	1200
aagagagat	cagegagact	ggggcccttt	ggcacctgt	ttccttggta	ctgtcaagg	1260
ggaggggcct	gtgatocaga	cacaggagat	tgttattca	gggatgagaa	tcctgcacatt	1320
gagtgtgtgc	actgcctaact	tggtttctca	aacgatccgc	acgacccccc	cagctgcag	1380
ccatgtccct	gtcataaacgg	gttcagctg	tcagtgtgc	cgaggacgga	gggggtgtg	1440
tgcataacttgc	gcctcccccgg	ggtcacccgt	gccccgtgt	agctgtgtgc	tgatggctac	1500
tttggggacc	ccttttgtga	acatggccca	gtgaggectt	gtcageccctg	tcaatgcac	1560
aacaatgtgg	accccaatgtgc	ctcttggaa	tgtgacccgc	tgacaggcag	gtgtttgaag	1620
tgtatccaca	acacagcccg	cattacttgc	gaccagtgca	aaggcggctca	tttggggac	1680
ccattgggtc	ccaaacccagg	agacaatgt	cgagcttgc	actgttaacc	catgggtctca	1740
gaggcctgt	gtatgtcaag	tgatggcacc	tgtgtttgc	agccaggatt	tggtgccccc	1800
aactgtgtgc	atggaggatt	cagctgtca	gcttgcata	atcaagtgaa	gattcagat	1860
gatcgttta	tgccagcagct	tcagagaatg	gaggccctt	tttcaaaaggc	tcaggggtgt	1920
gatgggtatgc	ttacgttata	agatgtggaa	ggcaggatgc	agcaggctgt	gcaggccctt	1980
caggacattc	tgagagatgc	ccagattca	gaaggtgct	gcagatccct	tggtctccag	2040
tttggccaagg	tgaggagcc	agagaacago	taccagacc	gcctggatg	cctcaagat	2100
actgtggaa	gagttccggg	tcttggaaat	cagtaccagg	accgagttcg	ggataactcac	2160
aggctcatca	ctcagatgc	gctggccctg	cgacaaatgt	aagcttccct	ggggaaacact	2220
aacattcttc	cctcaagacca	ctacgtgggg	ccaaatggct	ttaaaaatct	ggctcaggag	2280
gccacaagat	tagcagaaag	ccacgttgc	tcagccagta	acatggagca	actgacaagg	2340
gaaaactgggg	atcttccaa	acaacgcctt	tcactgtgtc	gcaaggccct	gcatgaaagg	2400
gtcgggaatgc	gtacggcg	cccgagccgt	gctgtgtgc	aaaggcttgc	ggaaaaattt	2460
gagaaaaacc	agtccctggc	ccacgtgt	acaaggggagg	ccactcaagc	ggaaatggaa	2520
gcagatagg	tttatcagca	cagtcctccgc	ctcttggatt	cagtgtctgc	gtttcaggga	2580
gtcagtgtac	acttcattca	gggttggaa	gcaaaaggaa	tcaaacaaaa	agcggatctca	2640
ctctcaagcc	ttgttaccag	gcatatggat	gagttca	gtacacagaa	gaatctggga	2700
aactggaa	agaagacaca	cgactctta	cagaatggaa	aaatgtgggg	agagaaatca	2760
gatcgttgc	tttccctgtc	caatcttgc	aaaagcagag	cacaagaagc	actgtgtat	2820
gcaatgtoca	ctttttatga	agtttgcgg	atctttaaaa	acctcagaga	gtttgcactg	2880
cagggtggaca	acaaaaaaagg	agaatgtggaa	gaagggatgc	agagactctc	ctacatcagc	2940
cagaagggtt	cagatgccc	tgacaaagacc	cggcaagc	aaagagocct	ggggagcgct	3000
gctgctgt	cacagaggcc	aaagaatggg	gggggggggg	ccctggaaat	cttcaggatgc	3060
attgttgcgg	atgttgcgg	tcttgcacttgc	gaaagccatgt	tgacagcaga	tggagccctt	3120
gcccattgg	aggggactggc	ctctctgt	agtttgcgg	ggggatgtgc	aggagatgt	3180
gaaaaggaa	agctgtggat	tgacacaaat	atgttgcgg	tgacatgtgt	gattacagaa	3240
gcccagaagg	ttgtataccag	agccaagaac	gctgggttca	caatccaa	cacactcaac	3300
acattgtacg	gtctgtatgc	tctgtatggc	cagcctctca	gtgttagatg	agaggggtct	3360
gtcttactgg	agcagaatgt	tttcccgatcc	aaagccatgg	tcaacacgca	actcgccccc	3420
atgtatgtc	agctgtggaa	gagggtacgt	cgacacgg	gcccacccca	tttgcgttgc	3480
acaagcatag	atgggattct	ggctgtatgc	aaagacttgg	agaacattag	ggacaactcg	3540
ccccccaggct	gttacaatac	ccaggcttgc	gagcaacatgt	ga		3582

```

SEQ ID NO: 48          moltype = DNA  length = 3582
FEATURE             Location/Qualifiers
misc_feature        1..3582
                      note = Synthetic Construct
source              1..3582
                      mol_type = other DNA
                      organism = synthetic construct

```

```

SEQUENCE: 48
atgcctgtct tgggttgggg ctgtgcctg tgttttatgc tcgtgtgtcc agccggccaga 60
gccacatcta gaagagaatgt gtgcgactgc aacggcaaga gcccggcgtt cattttcgac 120
agagagctgc acagacagac ccggcaacggc ttccatgtcc tgaactgcgg cgacaacacc 180
gacggcatcc actgcggaaa gtgcaagaac ggcttctacc ggcacccggc gagggataga 240
tgccctgcctt gcaactgcgg ctccaaaggcc agcctgacgg ccagatgcgg caatagccgc 300
agatgttagct gcaaggctgg cgtgacaggc gcttagatgc atagatgtct ccccgcttc 360
cacatgtctgaa ccgtatgcgg atgttccatgc gaccagacac tgctggacag caagtgcgt 420
tgcgaccctgc ccggaaattgc cggacccctt gatgcgttgc taaacatgc 480

```

-continued

gtgaccggcg agagatgtga	cagatgtaga	agcggctact	acaacctgga	cggcgccaat	540	
cctgaaaggct	gcacccagg	cttgcgtac	ggccacacg	ccagctgtag	aagcagcgc	600
gaataactcg	tgccacaagat	caccagcacc	tccaccagg	atgtggacgg	atgaaaggcc	660
gtgcagagaa	atggctctcc	tgccaaagtg	cagtggccc	agagacacca	ggacgtttc	720
agcagcgc	tc	cccgtgtac	tttgtggccc	ctgccaagtt	cctggcaac	780
cagcaagtgt	cttacggcc	gagcgtgago	ttcgactaca	gagtggatag	aggcggcaga	840
cacccccc	ctcacatgt	gattttgaa	ggggccggac	tgccgtatcac	agcccccttt	900
atgcctctgg	gcaagaccc	gccttggc	ctgaccaaga	cctacaccc	ccggctgtat	960
gagcacccca	gcaacaactg	gtccccacag	ctgagctact	tcgagatcag	acggctgtcg	1020
cgaaacctga	cagccatgt	aatcagac	acccatcgcc	agtacacac	cggtatcatc	1080
gacaacgtga	cctgtatcg	cgccacact	gtttctgtt	ctctgtcc	ttgggtcgag	1140
cagtgtatct	gtccctgtgg	ctacaagg	cggttctgc	aggattgtgc	cagcggctac	1200
aagagagact	ctgcacact	ggggccctt	ggcacatgca	tccttggtaa	ttgtcaaggc	1260
ggcgagact	gogatccca	tacaggcgt	tgctacagcg	gogacgagaa	cccccgatatc	1320
gagtgcgc	attgtccat	cggttttac	aacgacccte	acgaccccg	atctgtcaag	1380
ccatgtctt	gocacaatgg	cttgcgtc	agcgtgtac	ccgaaacccga	agaggctgt	1440
tgcaacaatt	gcccaccagg	cgttacagg	gccagatgt	aactgtgtc	cgacggctac	1500
ttcggegat	cttttggaga	acacggcc	gtgcgac	ccctgc	tcagtgcaac	1560
aacaacgtgg	acccaagcgc	cageggca	tgcgatagac	tgacaggcag	atgtctgaag	1620
tgcatccaca	ataccggcc	gatctact	gaccagg	aggccggct	ttttggc	1680
cctctggc	ccaaatctgc	cgataagt	agagcctgca	actgtaaccc	tatggctct	1740
gaggcgttgg	gtcgacat	tgatgaa	tgctgtgc	agccagg	ttggcgac	1800
aatttgc	acggcc	tagtgc	gcctgtc	atcaagt	gatccagat	1860
gaccagg	tgcacat	gcagagg	gaagcc	tctctaa	ccaagg	1920
gatggcgttgg	tgccgtat	agagotgg	ggcaga	aatgc	acaggcc	1980
caggacat	tgagagat	ccagat	gagggc	ctaga	ggactgc	2040
ctggctaa	tgccggac	agagaac	taccag	gactgg	ctgt	2100
accgtggaa	gagt	cgac	cttgg	aggat	ggatacc	2160
cggctgtat	cc	ccatgt	actgt	cc	ggcaat	2220
aatattcc	cc	ctacgt	cccac	cc	ttaa	2280
gcaccc	cc	ccatgt	ccatgt	cc	ggct	2340
gaaaac	cc	actact	ccatgt	cc	ca	2400
gtcggc	cc	ccatgt	ccatgt	cc	gg	2460
gaaaag	cc	aatct	ccatgt	cc	aaaa	2520
gc	cc	ccatgt	ccatgt	cc	gg	2580
gttccg	cc	agat	ccatgt	cc	cc	2640
ctgagc	cc	ccatgt	ccatgt	cc	cc	2700
aactgg	cc	ccatgt	ccatgt	cc	cc	2760
gaccat	cc	ccatgt	ccatgt	cc	cc	2820
ggca	cc	ccatgt	ccatgt	cc	cc	2880
caatgt	cc	ccatgt	ccatgt	cc	cc	2940
cagaa	cc	ccatgt	ccatgt	cc	cc	3000
gcac	cc	ccatgt	ccatgt	cc	cc	3060
atcg	cc	ccatgt	ccatgt	cc	cc	3120
gccc	cc	ccatgt	ccatgt	cc	cc	3180
gaac	cc	ccatgt	ccatgt	cc	cc	3240
gccc	cc	ccatgt	ccatgt	cc	cc	3300
actctgg	cc	ccatgt	ccatgt	cc	cc	3360
gttctgttgg	cc	ccatgt	ccatgt	cc	cc	3420
atgatg	cc	ccatgt	ccatgt	cc	cc	3480
accagcat	cc	ccatgt	ccatgt	cc	cc	3540
ccac	cc	ccatgt	ccatgt	cc	cc	3582

SEQ ID NO: 49 moltype = DNA length = 3891
 FEATURE Location/Qualifiers
 source 1..3891
 mol_type = unassigned DNA
 organism = Clostridium botulinum

SEQUENCE: 49
 atgc

atgc	catttgc	tataaaaaca	at	ttttttat	aaagatc	ttttttgt	tgatattgt	60
at	at	aaaaaa	ttccaaatgc	agg	acaaat	caacc	gttacat	120
aa	aa	aaat	atgg	ttat	ccat	aaat	tttacaaatc	180
cc	cc	ccaa	aaac	aa	at	ccat	atgt	240
aa	aa	aaat	at	at	ttt	aaat	tttataat	300
cc	cc	at	at	at	ttt	aaat	tttataat	360
aa	aa	aaat	at	at	ttt	aaat	tttataat	420
cc	cc	at	at	at	ttt	aaat	tttataat	480
aa	aa	aaat	at	at	ttt	aaat	tttataat	540
cc	cc	at	at	at	ttt	aaat	tttataat	600
aa	aa	aaat	at	at	ttt	aaat	tttataat	660
cc	cc	at	at	at	ttt	aaat	tttataat	720
aa	aa	aaat	at	at	ttt	aaat	tttataat	780
cc	cc	at	at	at	ttt	aaat	tttataat	840
aa	aa	aaat	at	at	ttt	aaat	tttataat	900
cc	cc	at	at	at	ttt	aaat	tttataat	960
aa	aa	aaat	at	at	ttt	aaat	tttataat	1020

-continued

ttatacataaa	tgttaacaga	gatttacaca	gaggataatt	ttgttaaggt	ttttaaagta	1080
cttaacagaaa	aaacatattt	gaatttttat	aaagccgtat	ttaagataaa	tatagtacct	1140
aaggtaattt	acacaatata	tgtggatt	aatttaagaa	atacaattt	agcagcaac	1200
ttaatggtc	aaaatcaga	aattaataat	atgaatttt	ctaaactaaa	aaattttact	1260
ggattgttg	aattttataa	gttgctatgt	gtaagaggga	taataacttc	taaaaactaaa	1320
tcatttagata	aaggatacaa	taaggcatta	aatgatttat	gtatcaagt	taataattgg	1380
gacttgtttt	ttatgccttc	agaagataat	tttactaat	atcttaataa	aggagaagaa	1440
attacatctt	actaataat	agaagcgc	gaagaaaata	tttagtttaa	ttaatataca	1500
caatattatt	taacctttaa	ttttgataat	gaacctgaaa	atatttcaat	agaaaaatctt	1560
tcaagtgaca	ttatagggca	attagactt	atgcctaata	tagaaagatt	tcctaattgg	1620
aaaaggatag	agtttagataa	atatactatg	ttccattatc	ttcggtc	agaatttgaa	1680
catggtaat	ctaggattgc	ttaaacaat	tctgttaac	aagcattat	aaatcctagt	1740
cgtgtttata	cattttttc	ttcagactat	gtaaagaag	ttaataaagc	tacggaggca	1800
gctatgtttt	taggtgggt	agaacatata	gtatatgt	ttaccgtga	aactagcgaa	1860
gtaagtacta	cggataaaaat	tgccgatata	actataatta	ttccatata	aggacctgt	1920
ttaaatatag	gtaatatgtt	atataaagat	gattttgtat	gtgcattat	attttcaggaa	1980
gctgttattc	tgtttagaaatt	tataccagag	attgcaatac	ctgtattagg	tacttttgc	2040
cttgcattat	atatttgcga	taagggttca	accgttcaaa	caatagataa	tgctttaagt	2100
aaaagaaaatg	aaaatgggg	tgagggtt	aaatataat	taacaaat	gttagcaaaag	2160
gttaatacac	agattgtatc	aataagaaaa	aaaatggaa	aagctttaga	aaatcaagca	2220
gaagcaacaa	aggctataat	aaactatcg	tataatcaat	atactgagga	agagaaaaat	2280
aatattaaat	ttaatattgt	tgatttatagt	tcgaaactt	atgagtctat	aaataaaagct	2340
atgattaat	ttataataat	tttgaatca	tgtctgtt	cataattat	gaattctatg	2400
atcccctatg	gtgtttaaacg	ttttagaagat	ttttagtgc	gtctttaaaa	tgcattatta	2460
aagtatata	atgataatag	aggaacttta	attggtcaag	tagatagatt	aaaagataaa	2520
gttaataata	cacttagtac	agatatactt	tttcagttt	ccaaatactgt	agataatcaa	2580
agatttattat	ctacatttac	tgaatattat	aagaatattt	ttaatacttc	tatatttgaat	2640
ttaagatgt	aaagtaat	ttaatagac	ttatcttagt	atgcattcaaa	aataaattat	2700
ggttagtaaag	taaattttga	tccaatagat	aaaatcaaa	ttcaattatt	taatttagaa	2760
agtagtaaaa	ttgagggtat	ttttaaaaaat	gctattgtat	ataatagtat	gtatgaaaat	2820
tttagtacta	gcttttggat	aagaatttc	aagtatttt	acagtataag	tctaataat	2880
gaatataacaa	taataaattt	tgaggaaaat	aattcaggat	ggaaaatgtc	acttaattat	2940
ggtgaaataa	tctggactt	acaggatact	caggaaataa	aacaaagagt	agttttaaa	3000
tacagtcaaa	tgattaat	atcagattat	ataaacat	ggattttgt	aactatcact	3060
aataatagat	taaataact	taaaatttt	ataaattggaa	gattataga	tcaaaaacca	3120
atttcaatt	taggtat	tcatgtat	ataatataat	tgttttaatt	agatggtgt	3180
agagatacac	atagatata	ttggataaaa	tattnnaatc	tttttgat	ggaatttaat	3240
aaaaaaagaaa	tcaaagat	atatgataat	caatcaaatt	caggatattt	aaaagacttt	3300
tgggggtatt	atttacata	tgataaaac	tactatgt	taaatttata	tgatccaaat	3360
aaatatgtcg	atgttaaaa	tgtaggtat	agaggttata	tgtatcttaa	agggcctaga	3420
ggttagcgtaa	tgactacaaa	catttattt	aattcaagt	tgtatagggg	gacaaaattt	3480
attataaaaa	aatatgttc	tggaaataaa	gataatattt	tttagaaataa	tgatcggt	3540
tatattaaat	tagtagttaa	aaataaaagaa	tataggtt	ctactaatgc	atcagggca	3600
ggcgttagaa	aaatactaa	tgcattagaa	atacctgtat	taggaaatct	aagtcaagta	3660
gttagatgt	agtcaaaaa	tgatcaaggaa	ataacaaata	aatgcaat	gaatttacaa	3720
gataataatg	ggaatgtat	aggctttata	ggatttcatc	agtttataat	tatagctaaa	3780
ctagtagcaa	gtatgggt	taatagacaa	atagaaagat	ctagtaggac	tttgggttgc	3840
ctatggaaat	ttattccgt	agatgtat	tggggagaaa	ggccactgt	a	3891

SEQ ID NO: 50 moltype = DNA length = 3891
 FEATURE Location/Qualifiers
 misc_feature 1..3891
 note = Synthetic Construct
 source 1..3891
 mol_type = other DNA
 organism = synthetic construct

SEQUENCE: 50
 atgcccctcg tgaacaagca gttcaactac aaggccccg tgaacggcgt ggacattgcc 60
 tacatcaaga tcccccaacgc cggccagatc cggccatgtc agggcattaa gatccacaac 120
 aagatctggg tcataccatca gccgggacacc ttccacaatcc cccggaaagg cgacctgtat 180
 ctcaccatcg aggccaaaca ggttcccgtg ttctactacg acagcaacta cctgagcacc 240
 gacaacgaga aggacaacta cctgaagggc gtgaccaagc tgttcgagcg gatctacagc 300
 accgatctgg gcagaatgtc gctgacctt atcgtgcggg gcatccatt ttggggccgc 360
 agcacaatcg acacggactg gaaactgtatc gacaccaact gcatcaactgt gatccaggct 420
 gacggcagct acagaagcga ggaactgtatc ctggtcatac tggccccccat ccggcgcacatc 480
 atccaggatcg agtcaagatc ctccggccat gaggtgtatc acctgacccg aatatggctac 540
 ggcagcaccg agtacatccg gttcagcccc gatttcaccc tggcttgcg ggaatccctg 600
 gaagtggaca ccaatccctc gctcgagcc ggcaagttt ccacccatcc tgctgtgaca 660
 ctggccacag aactgtatca cggccggacac agactgtatc gatcgcctt caatcccaac 720
 cgggtgttca aagtgaacac caacgcctatc tatggatgtc gggactgtg agtgtccctt 780
 gaggaactgc ggacccctgg cggacacgc gccaagttt tcgacacgtc gcaagagaaac 840
 gagttccggc tgcataacta caacaagttc aaggatatcg ccagcacgtc gaacaaggcc 900
 aagagcatcg tgggcacaac agccggctgtc cagtagatca agaactgtt caaagagaaag 960
 tacctgtca gcgaggacac cagcggcaag ttctccgtgg acaagctgaa gttcgacaaag 1020
 ctgtacaaga tgctgaccga gatctacacc gaggacaact tcgtgaagtt cttcaaggtg 1080
 ctcaaccgcgca agacccatcttcaacttgcac aaggccgtgt tcaagatca catcgccc 1140

-continued

aaagtcaact	acaccatcta	cgacggcttc	aacctgcgca	acaccaaccc	ggccgcacac	1200
ttaaacccgc	agaacaccgc	gatcaacaaad	atgaacttca	ccaagctgaa	aaacttcacc	1260
ggcctgtcg	aggcttataa	getgtgtgc	gtgcgcggca	tcatcaccag	caagaccaag	1320
agcctggaca	agggtataa	caagggcttg	aacgaccctgt	gcatcaaaatg	taacaactgg	1380
gacctgtct	ttagccccag	cgaggataat	ttcaccaac	acctgaacaa	aggcgagggaa	1440
atcaccagcg	acaccaatat	cgaggccgc	gaggaaaaac	tcagctgga	cctgatccag	1500
cagtactacc	tgacccatca	tttcgataac	gagccccgaga	acatcagcat	cgagaactcg	1560
agcagcgcata	tcatecgaca	gctggaaactg	atgcccacaa	tcgagagatt	ccccaacggc	1620
aagaagtacg	agctggacaa	gtacaccatg	ttccactacc	tgccggccca	agagttcgag	1680
cacggcaatg	ctagaatcgc	cctgaccaaa	agcgtgaac	aggccctgt	gaaccccaac	1740
agagtgtaca	ccttcttcag	cagegatact	gtgaaagaaag	tcacaaaggc	tacagggcc	1800
gccccatgttcc	tcgggtgggt	tgagcagctg	gtgtacgact	tcacccgacga	gacaaggcgag	1860
gtgtccacca	ccgacaagat	cgccgatatac	accatcatca	tccttacat	cggccctgtct	1920
ctgaacatcg	gcaacatgt	gttacaggaa	gatttcgtgg	gcccctgtat	cttctctggc	1980
gccccgttgc	tgctcgagtt	cateccatgc	atcgctatcc	cctgtgtgg	cacatttgc	2040
ctgtgtttt	atactccaa	caaggtgt	acatgtcaga	ccatcgacaa	cgcctgtac	2100
aagcggaaacg	agaagtggga	cgaagtgtac	aagtacatcg	tgaccaactg	gctggccaaa	2160
gtgtaaatcc	agatcgac	gattcggaa	aagatgtgg	aaggccctcg	gaaccaggcc	2220
gaggccacaa	aggccatcat	caactaccag	tacaatcag	acacagagga	agagaagaaac	2280
aacatcaatt	tcaacatcg	cgacccgttc	tccaaagtc	acgagacat	caacaaagcc	2340
atgatcaata	tcaacaagtt	tctgaaccag	tgccgtgtc	gatcacatgt	gaacagcatg	2400
atccccatcg	ggtgtaaacg	gctggaaat	ttcgatgtcc	gcctgaagg	cgctctgt	2460
aagtacatcg	acgataacc	ggggccctcg	atcgccagg	ttggacagact	gaaggacaaa	2520
gtgtacacata	ccctgtccac	cgacatccg	tttcagctg	gcaaaatcgt	ggacaacccag	2580
agactgtga	gcacattc	cgagtatac	aagaatatac	tcaacaccc	catctgtac	2640
ctccgcgtac	agagaacca	cctgtatcg	ctgagcagat	acggccacaa	gatcaatatt	2700
gtgtctaaag	tgaacttgc	cccgatcg	aagaaccaga	tcacatgtt	caatctcgag	2760
tctagaaaga	tgaatgtat	cttgcggaa	gccatctgt	acaacttat	gtacgagaac	2820
tttccacca	gtttttggat	cagaatcccc	aagtacttca	acagcatctc	cctgaaacac	2880
gagttacacg	tcatcaatt	catggaaaac	aactccggct	ggaaagtgtc	cctgaaactac	2940
ggcggatgt	tctggacact	gcgggacac	caagagatca	acgagagatg	gtgttcaag	3000
tactctcaga	tgtacaaat	tagcactac	atcaaccgg	ggatcttgt	gaccatcacc	3060
aacaacccgc	tgaacaactc	caagatctac	atcaatggcc	ggctcatcg	ccagaaggct	3120
atcagcaacc	ttggaaacat	ccacgcctcc	aacaatatac	tggtcaag	ggacggctgc	3180
cggggacacc	accggatata	ctggatca	tacttttaacc	tggtcgacaa	agagctgaaac	3240
gagaagaa	ttaaggac	gtacgacaa	cagtccaa	ggggcatct	gaaggattc	3300
tggggcact	acctgcgt	tgacaaagcc	tactacatgc	tgaatctgt	cgaccccaac	3360
aaatatgtgg	acgtgaacaa	cgtggggatc	agaggctaca	tgtacctgaa	aggccccaga	3420
ggcggcgtga	tgaccggaa	catctaccc	aactccacg	tgtacagagg	caccaagttc	3480
atcatcaaga	atgtgc	cggtcaac	gacaacat	tgcggaaacaa	cgaccgggt	3540
tacattaaac	tgggtgtca	gaacaaagag	tacccgttg	ccaccaatgc	cttcaggc	3600
ggcgtggaaa	agatctgt	cgccctggaa	atcccgac	tgggcaatct	gtctcaggt	3660
gtctgtat	atgtcaagaa	cgccagggg	atcacaaatc	agtgcaagat	gaacctccag	3720
gacaacaaac	gcaacgc	cggttttac	ggcttccac	agtttaacaa	cattgccaag	3780
ctgtgtcoca	gcaactgtt	caacccgg	atcgagagaa	cgacgacaa	cctgggctgc	3840
agctggagat	ttatccctgt	ggatgtgg	tggggcgaa	gaccctgt	a	3891

SEQ ID NO: 51 moltype = DNA length = 3876
 FEATURE Location/Qualifiers
 source 1..3876
 mol_type = unassigned DNA
 organism = Clostridium botulinum

SEQUENCE: 51

atggccagtt	caataaataa	ttttaattat	aatgatccta	ttgataataa	taatattatt	60
atgtatggagc	ctccatttgc	gagaggatcg	gggagatatt	ataaagctt	taaaatcaca	120
gatcgatattt	ggataatacc	ggaaagatag	actttttgtat	ataaaacctga	ggatttttaat	180
aaaaggttcg	gtattttaa	tagatgtt	tgtgaatatt	atgatccaga	ttacttaaat	240
actaatgtata	aaaagaataat	attttacaa	acaatgtac	agttatttaa	tagatcaaa	300
tcaaaacccat	tgggtggaaa	gtttagatag	atgattatac	ttatcttgga	360	
gatagacgtg	ttccactcg	agagtttac	acaacat	ctgtgtac	tgttataaaa	420
tttaatcgata	atccaggaga	agtggagcga	aaaaagggtt	ttttcgcaaa	tttaataata	480
tttggacactg	ggccagttt	aatatggaaat	gagactata	atataggtat	acaaaatcat	540
tttgcataat	gggaaggcctt	cgggggatata	atgcaaata	agttttgtcc	agaatatgt	600
agcgttata	ataatgttca	agaaaaacaa	ggggcaatgt	tatttaaat	acgttgat	660
tttttcagatc	caggcttgat	attaatgc	gaaacttatac	atgttttata	tggattttat	720
ggcattaaat	tagatgtt	accatgtt	ccaaatgaa	aaaaattttt	tatgtcaatct	780
acagatgtca	tacaggcaga	agaactata	acattttgg	gacaagatcc	cagcatcata	840
actcccttca	cggtataaa	tatctatgt	aaagtttgc	aaaattttt	aggatagtt	900
gatagactta	acaagggttt	agtttgcata	tcagatccta	acattataat	taatataat	960
aaaaataataat	ttaaagatata	atataaattt	gttgaagat	ctgagggaaa	atatagtata	1020
gatgtgat	aaaaatgtt	attatataaa	agtttaatgt	ttggttttac	agaaactat	1080
atagcagaaa	attatataaa	aaaaactaga	gcttcttatt	ttatgtatcc	tttaccacca	1140
gtaaaaataa	aaaattttt	agataatgaa	atctatacta	tagaggaa	ttttaatata	1200
tctgataaa	atatggaaaa	agaatata	gttcagaata	aagctataaa	taaacaagct	1260
tatgaagaaa	ttagcaagga	gcattttgg	gtatataaga	tacaaatgt	taaaggatgtt	1320
aaagctccag	gaatatgtat	tgtatgtt	aatgaagat	tgttctt	atgtatataa	1380

-continued

aatagtttt cagatgattt atctaaaaac gaaagaatag aatataatac acagagtaat	1440
tatataaaaa atgacttccc tataaatgaa ttatTTTtaa atactgatt aataagtaaa	1500
atagaattac caagtggaaa tacagaatca cttaactgatt ttatgtaga tgTTCCAGTA	1560
tatgaaaaac aaccgcctat aaaaaaaatt ttacagatg aaaataccat ctTCAATAT	1620
ttataactctc agacatttc tctagatata agagatataa gtttaacatc ttCATTTGAT	1680
gtatgcattat tattttctaa caaagtttat tcattttttt ctatggatata tattaaaact	1740
gctaataaaag tggtagaagc aggattttt gcaggTTGGG tgaaacagat agtaaatgat	1800
tttgtaatcg aagctaataa aagaataact atggataaaa ttgcagatat atCTCTAATT	1860
gttccttata taggattagc tttaaatgta gggaaatgaaa cagctaaagg aaatTTGAA	1920
aatgcTTTT agattgcgg accggatTTT ctactgaat ttatACCGA actTTTATAA	1980
cctgtatTTG gagCCTTTT attagaatca tatattgaca ataaaaataa aattattaaa	2040
acaatagata atgCTTAAc taaaagaaa gaaaaatgga gtgatATGTA CGGATTAATA	2100
gtagcgaat ggCTCTCAAC agttaatact caattttata caataaaaAGA GGGATGTAT	2160
aaggCTTAA attatCAAGC acaAGCATTG gaagaaataa taaaatacag atataatata	2220
tattCTGAAA aagaaaaAGTC aaatTTAAc atcgatTTA atgatataaa ttCTAAACTT	2280
aatgaggggA ttaacAAcG tatagataat ataataataa ttataatgg atgttCTGTA	2340
tcataTTAA tgaaaaaaat gattccatta gctgttagaa aattactaga ctTTGATAAT	2400
actCTCAAA aaaaTTGTTT aaattatata gatggaaaata aattatattt gattggaaGT	2460
gcagaatATG aaaaATCAA AGTAAATAAA tacttgaaa ccattatGCC gTTTGATCTT	2520
tcaatataa ccaatgatAC aataataa gaaaATGTTA ataaatataa tagcgaatt	2580
ttaaataata ttatCTTAA tttaagatata aaggataata attaataga ttatTCAGGA	2640
tatggggAA aggttagatTTT atatgtatgg gtcgagCTTA atgataaaaaa TCAATTAAA	2700
ttaactagTTT cagcaatATG taatggatTA gtgactcaaa atcagaatAT CATAATTAT	2760
agtgtgttcc ttgatTTAG cgtagCTT ttgataagaa taccaataa taagaatGAT	2820
ggtatacAAA attatATCA taatgatata acaataatTA attgtatgaa aaataatTCG	2880
ggctggAAA tatCTTAGG gggtaatAGG ataataatgGA ctTTAATTGA tataatGGa	2940
aaaacccaat cggatTTTTT tgaatataaC ataaagagaa atataatCAGA gtatataaAT	3000
agatggTTT ttgtaactat tactaataaC ttgataACG ctaaaATTa tattatGGT	3060
aagcttagat caaaatCAGA tattaaAGAT ataaAGAGAAG ttattGCTAA tggtgaaATA	3120
atatttaat tagatGGTGA tatagatAGA acacaatTTA ttggatGAA atattTCAGT	3180
atTTTAAAtt cggatTAAG tcaatcaaa attgaagaaa gatataaaaat tcaatcatat	3240
aggcaatTTT taaaAGGTTT ttggggAAat cctttaatGt acataaAGA atattatATG	3300
ttaatGCGG ggaataaaaa ttcatatatt aaactaaAGA aagattCACC tGTTAGTGAA	3360
atTTTAAACAC gtagcaAAAt taatcaaaAT tctaaatata taaattatAG agatttat	3420
atTggagAAA aattttatTTT aagaAGAAAG tcaaaATCTC aatCTATAAA tgatGATATA	3480
gttagaaaaAG aagatttat atatGATAG ttttttaat taaaatCAGA tgGAGAGATA	3540
tatacctata aatTTTAA gaaAGAGGAA gaaaaATTG ttttagotCC tataatGAT	3600
tctgtatGAGT ttTACAATAC tatacaaATA aaAGAATATG atgaACAGCC aacatataG	3660
tgtcagtTGC ttTTTAAAGA agatGAGAAAG agtactGATG agatAGGATT gattGGTT	3720
catcgTTCT acgaatTCG aattgtattt ggagatGATA aagatTTTT ttgtataAGT	3780
aaatGTTACT taaaAGGAGGT aaaaAGGAAAC ccatataatt taaaattGGG atgtatTTG	3840
cgatTTATC CTTAAAGATGA aggGTGGACT gaataa	3876

```

SEQ ID NO: 52          moltype = DNA  length = 3876
FEATURE
misc_feature          Location/Qualifiers
1..3876
note = Synthetic Construct
source
1..3876
mol_type = other DNA
organism = synthetic construct

```

```

SEQUENCE: 52
atgcccgtga ccatcaacaa cttcaactac aacgacccccca tcgacaacaa caacatcatt 60
atgatggAAC cgcCTTcgc cagaggcaca ggccggTACT acaaggccTT caagatcacc 120
gaccggatCT ggatcatccc cgagagatac accttcggct acaaggcccga ggacttcaac 180
aagagcAGCg ccatTTCAAC cggggacgtg tgcgagtact acgacccccga ctacCTGAAC 240
accaacgaca agaagaacat ctteCTGCG accatgatca agctgttcaac ccggatcaag 300
agcaagcccc TTGGCGAGAA gctgtggaa atgatcatca acggcatccc ttacCTGGC 360
gacagaagAG tgccccgtGA agaggTCAAC accaataatCg ccacgcgtac cgtgaacaaG 420
ctgatCTCTA acccCGCGA ggtggAACGG aagaaggGGCA ttctegoccaa CCTGATCATC 480
ttcggCCCTG gacCTGTGCT gaacAGAGAAAG gagacaatCg acatCggcat ccagaaccac 540
ttcggCCAGCA gagaAGGCTT CGGCGCGATC atcgatGATG atgttCTCCC CGAGTACGTG 600
tccgtgttca acaacGTGCA agagaacaAG ggGCGCAGCA ttTTTAAACAG acggggGTAC 660
ttcagcgacc CGCGCTCTGAT tctgtatGCAc gagCTGATCC acgtGCTGCA CGGCCTGTAT 720
ggcatcaAGG tggacGACCTC GCCTATCGTG CCCAACGAGA agaaatttCTC catcgatGTC 780
accgacgCCA tccaggCGGA gggactGTAC acatTTGGCG gacaggacCC cgcataCAtC 840
acccCTAGCA cggacaAGAG catCTACGAC aaggGTGTCG agaaattCCG gggcatCGTG 900
gaccggCTGA acaaAGGTGCT CGTGTGCAc acgcatCCCA acatcaatAT caacatCTAC 960
aagaacaAGT tcaaggatTA GTACAAGTTC gtcgaggaca gCGAGGGCAA gtacAGCATC 1020
gacgtggAAA gCTCTGACAA gCTGTACAAG agcCTGATGT tcggCTCAC CGAGACAAAT 1080
atcgcGCCAGA actacaAGAT caagACCCGG gCCAGCTACT tCTCCGACTC tCTGCTCTCT 1140
gtGAAGATTA agAACCTGCT ggacaACCGAG atCTACACCA tcgAGGAAGG CTTCACATC 1200
agcgacaAGG acatGGAAAAA agagtACCGG ggCCAGAAcA aggCCATTAA caAGCAGGCC 1260
tacgaggAAA tcAGCAAGA acacCTGGCC gtgtacaAGA ttcaGATGTG caAGAGCGTG 1320
aaggCCCTG gcatCTGcat tgacgtggac aatgaggacc tgTTCTTAT CGCCGACAAg 1380
aacagCTTA CGCAGCAGCT gagaAGAAc gacggatCg agtacaACAC CCAGAGCAAC 1440
tacatcgaga acgacttccc catcaacgaa ctgatCCTGG acaccgacCT gatcagcaag 1500

```

-continued

atcgagctgc	ccagcgagaa	caccgagago	ctgaccgact	tcaatgtgga	cgtgcccgtg	1560
taacggaaagc	agcccgccat	caaaagatc	tttaccgac	agaataacat	cttccatgtac	1620
ctgtacagcc	agaccccttc	tctggacatc	cgggacatca	gcctgaccc	cagcttcgat	1680
gatgccctgc	tgttcagcaa	caaggctcac	agcttctca	gcatggacta	catcaagacc	1740
gccaacaagg	tggttggaaagc	cggcctgttt	gcccggctggg	ttaaggcagat	cgtgaacgat	1800
ttcgtgtatcg	aggccaaacaa	gtccaaacacc	atggacaaga	tcggccgat	atccctgtac	1860
gtgccttaca	tcggactggc	cctgaacgtg	ggaaacggaga	cagccaaagg	caactctcgag	1920
aatgccttcg	agatttgcgg	cgtctcgatc	ctgctcgatg	tcatccctga	gctgtgtatc	1980
cctgtctgg	gctgtttct	gctggaatcc	tacatcgata	acaaaaacaa	gatcatcaag	2040
acqatcaca	acggccctgac	caaggcggaa	gagaagtgg	gccccatgt	ccggactgtac	2100
gtggcccaatg	gtgttggac	cgttgcatacc	cagttttaca	ccatcaaaga	agggtatgtac	2160
aaggccctga	attaccaggc	tcaggctctg	gaagagatca	ttaagtaccc	ctacaatata	2220
taacagcgaga	aagagaagtc	taacatcaac	atcgacttca	acgacatcaa	cagcaagctc	2280
aacgaggccg	tcaaccaggc	cattgcacaa	attaacaatc	ttatcaacccg	ctgcagcgat	2340
tcctacatcg	tgaaggaaat	gatccctgt	gcccgtggaa	agctgtcgat	cttcgcacaaat	2400
acccctgaaga	agaacccctc	gaaacttacat	gacgagaa	acgttactatc	gatecgccac	2460
gcccggatcg	agaaaaagca	agtgaacaag	tacctcaaga	ccatcatgccc	cttgcacatc	2520
tccatctaca	caaaacgacac	cattctgtatc	gagatgttta	acaagatcaa	cagcgagatc	2580
ctgaacaata	tcatctctga	cctgcggatc	aaggacaa	atctgtatca	tctgaggccg	2640
tacggccoca	aggttggaaat	gtatgtatcc	gtggactgt	acgatggaa	tcagttcaag	2700
ctgaccggca	gcccggactc	caagatcaga	gtgaccggaa	accagaaacat	tatcttcaac	2760
agcgtgttcc	tggatcttc	cgtgtcttc	tggatcaagaa	tccccaatgt	caagaacggac	2820
gcoattccaga	actacatcca	caacggatc	acaatcatca	actgttatgaa	gaacaacacg	2880
ggcttggaa	ttagatcccg	ggcaacaga	atcatctgt	ccctgtatca	catcaatggc	2940
aagacaaaga	gcgtgttctt	cgagtataac	atccgcgagg	acatctccga	gtatcatcaac	3000
cgggttgttct	tcgttggacaa	cacaacaa	ctgaacaaac	ccaagatcta	catcaacccgg	3060
aagctcgaga	gcaacacccg	catcaaggat	atccggggaa	tgatcgocaa	ccggcgagatc	3120
atctttaa	tgccggccgca	catcgacccg	acacagtctc	tctggatgaa	gtacttttc	3180
atcttcaata	ccgagctgat	ccagagcaat	atcgaagaga	ggtacaatgat	ccagtcttac	3240
agcgagttacc	tgaaggactt	ctggggcaac	cctctgtatg	acaacaaaga	atattatcatg	3300
ttaaacggcc	gcaaaacaa	cttccatcata	aagcttgcac	aaagacgccc	cgtgggagaa	3360
atcttgcacca	ggtctaaatg	ccaaacaa	tccaaatgt	tcaactaccc	ggacctgtac	3420
atccggcaga	agttcatcat	ccggcgcaag	agaacacagcc	agtccatcaa	tgacgacatc	3480
gtgcggaaag	aggactatata	ctacccgtac	ttcttcaacc	tcaatcaaga	gtggcgctg	3540
tacaccatca	agttttttaa	gaaaaggagg	gagaagtgt	tcctggotcc	tatcagcgac	3600
tccgcacgat	tcttataatc	catccgcata	aaagatgtac	acgagcggcc	cacccatcagc	3660
tgcgcagctgc	tgttttaagaa	ggacggagaa	tccaccatgt	agatcgccct	gattggcatc	3720
caccgggttct	acgagagccg	catctgttcc	gaagagtaca	aggattactt	ctgcatac	3780
aagtggatc	tgaaaatggaa	gaaggcggaa	cccttataatc	tgaagctggg	ctgcaactgg	3840
catgttatacc	ccaaggacga	aggctggacc	gagtga			3876

SEQ ID NO: 53 moltype = AA length = 2477

FEATURE Location/Qualifiers

source 1..2477

mol_type = protein

organism = Homo sapiens

SEQUENCE: 53

MLRGPGPGLL	LLAVQCLGTA	VPSTGASKSK	RQAQQMVQPQ	SPVAVSQSKP	GCYDNGKHYQ	60
INQWQERTYI	GNALVCTCYG	GSRGFNCESK	PEAEETCPDK	YTGNTRYVRGD	TYERPKDSMI	120
WDCTCIGAGR	GRISCTIANR	CHEGGQSYKI	GDTWRRPHET	GGYMLECVCL	GNGKGEWTCK	180
PIAEKCFDHA	AGTSYVVGET	WEKPYQGWMM	VDCTCLGEBS	GRITCTSRNR	CNDQDTRTSY	240
RIGDTWSKKD	NRGNLQCIC	TGNGRSHKC	ERHTSVQTTT	SGSGPFTDVR	AAVYQPQPHP	300
QPQPPYGHCVT	DSGVVYSVGVM	QWLKTQGNKQ	MLCTCLGNVG	SQETATVQT	YGGNSNGEPC	360
VLPFTYNGRT	FYSCTTEGRQ	DGHLCWCSTTS	NYEQDQKYSF	CTDHVTLVQT	RGGNSNGALC	420
HFFFLYNNHH	YTDCTCTSEGRR	DNMWKWCGTQQ	NYDADQKPGF	CPMAAHEEIC	TTNEGVMYRI	480
GDQWDKQHDM	GHMMRCTCVG	NGRQLTCTVIA	YSQLRDRQCIC	DDITYNVNNDT	FHKRHEEGHM	540
LNCCTFCQGQR	GRWKCDPVQD	CQDSETGTTFY	QIGDSWEKVV	HGVRYQCYCY	GRGIGEWHHQ	600
PLQTYPSSSS	PVEVFITETP	SQPNHSPIQW	NAPQPSHISK	YLRLWRPKNS	VGRWKEATIP	660
GHLSNSYTIKG	LKPGVYYEGQ	LISIQYQGHQ	EVTRFDFTTT	STSTPVISNT	VTGETTPFSP	720
LVATSESVTE	TIASSFVVSW	VSASDVSFVSG	RVEYELSEEEG	DEPQYLDLPS	TATSVNIPDL	780
LPGRKYIVNV	YQISEDQEQS	LILSTSQTAA	PDAPPDPTVD	QVDDTSIVVR	WSRPQAPITG	840
YRIVYSPSVE	GSSTELNLPE	TANSVTLSDL	QPGVQYNITI	YAVEENQEST	PVVIQQETTG	900
TPRSDTVESP	RDLQFVEVTD	VKVTIMWTPP	ESAVTGYRV	VIPVNLPGEH	GQRLPISRNT	960
FAEVTVGLSPG	VTVYYPFKVFAV	SHGRESKPL	AQQTTKLDAP	TNLQFVNEDT	STVLVRWTPP	1020
RAQITGYRLT	VGLTRRQGPR	QYNVGSVSK	YPLRNLLQPA	EYTIVSLVAIK	GNQESPATG	1080
VFTTLQPGSS	IPPPYNTTEVTE	TTIVITWTPTA	PRIGFKLGV	PSQGGEAPRE	VTSDSGSIVV	1140
SGLTPGVEYV	YTIQVLRDQG	ERDAPIVNKV	VTPLSPPTNL	HLEANPDTGV	LTVSWERSTT	1200
PDTITGYRITT	TPTNGQQGNS	LEEVVHADQS	SCTFDNLSPG	LEYNVSVYTV	KDDKESVPIS	1260
DTTIPEVPQL	TDLSFVDITD	SSIGLRWTPL	NSSTIIGYRI	TVVAAGEGIP	IFEDFVDSVV	1320
GYYFTVTGLEP	GIDYDISVIT	LINGGESAPT	TLTQQTAVPP	PTDLRFTNIG	PDTMRVTWAP	1380
PPSIDDLTNFL	VRYSPVKNEE	DVAELSIIPS	DNAVVLTNLL	PGTEYVVS	SVYEQHESTP	1440
LRGRQKTTGLD	SPTGIDFSDI	TANSFTVHWI	APRATITGYR	IRHHPEHFSG	RPREDRVPHS	1500
RNSITLTNL	PGTEYVVSIV	ALNGREESPL	LIGQQSTVSD	VPRDLEVVA	TPTSSLISWD	1560
APAVTVRYYR	ITYGETGGNS	PVQEFTVPGS	KSTATISGLK	PGVDYITIVY	AVTGRGDSPA	1620
SSKPISINYR	TEIDKPSQMQ	VTVDQDNSIS	VKWLPSSSPV	TGYRVTTTPK	NGPGPTKTKT	1680
AGPDQTEMPI	EGLQOPTVEYV	VSVYQAQNPSG	ESQPLVQTA	TNIDRPKGLA	FTDVDSIK	1740

-continued

IAWESPQGV	SYRVTYSSP	EDGIHELPFA	PDGEEDTAEL	QGLRPGSEYT	VSVVALHDDM	1800
ESQPLIGTQS	TAIPAPTDLF	FTQVTPTSLS	AQWTPPNVQL	TGYRVRTPK	EKTGPMKEIN	1860
LAPDSSSVVV	SGLMVATKYE	VSVYALKDTL	TSRPAQGVVT	TLENVSPPRR	ARVTDATETT	1920
ITISWRKTKE	TITGFQVDAV	PANQCTPIQR	TIKPDVRSYT	ITGLQPGTDY	KIYLTLNDN	1980
ARSSPVVIDA	STAIDAPSNL	RFLATTPNSL	LVSWOPPRAR	ITGYIIKYEK	PGSPPREVVVP	2040
RPRPGVTEAT	ITGLEPGTEY	TIYVIALKNN	QKSEPLIGRK	KTDELQOLV	LPHPNLHGPE	2100
ILDVPSTVQK	TPFVTHPGYD	TGNGIQLPGT	SGQQPSVGQQ	MIFEEHGFR	TPPPPTATPI	2160
RHRPRPYPPN	VGEETIQIGHI	PREDVDYHLY	PHGPGLNPNIA	STGQEALSQT	TISWAPFQDT	2220
SEYIISCHPV	GTDEEPLQFR	VPGTSTSATL	TGLTRGATYN	IIVEALKDQQ	RHKVREEVVT	2280
VGNVSNEGLN	QPTDDSCFDP	YTVSHYAVGD	EWERMSESGF	KLLCQCLGFG	SGHFRCDSSR	2340
WCHDNGVNPK	IGEKWDROGE	NGQMMSTCL	GNGKGEFKCD	PHEATCYDDG	KTYHVGEQWQ	2400
KEYLGAICSC	TCFGGQGRWR	CNDNCRPGGE	PSPEGTTGQS	YNQYSQRYHQ	RTNTNVNCPI	2460
ECFMPLDVQA	DREDSRE					2477

SEQ ID NO: 54 moltype = AA length = 786
 FEATURE Location/Qualifiers
 source 1..786
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 54
 MAGLTAAAPR PGVLLLLSSI LHPSRPGGVP GAIPGGVPGG VFYPGAGLGA LGGGALGPGG 60
 KPLKPVPGGL AGAGLGAGLG AFPAVTFPGVA LVPGGVADAA AAYKAAKAGA GLGGVPGVGG 120
 LGVSAGAVVP QPGAGVKPGK PPGVGLPGVY PPGVLPGARF PGVGLPGVP TGAGVKPKAP 180
 GVGAGAFAGIP CGVPGFPGQP GPVPLGYPIKA PKLPGGYGLP YTTGKLPHYG GPGGVAGAAG 240
 KAGYPTGTGV GPQAAAAAAA KAAAKFGAGA AGVLPGVGGA GPVGPVGAIP GIGGIAGVGT 300
 PAAAAAAA AAKAAYGAAA GLVPGCPGPQ PGVVGVPAGG VPGVGVPAGG IPVVPGAGIP 360
 GAAVPGVSP BAAAAKAAA AKYCARPGVG VGGIPTYGVG AGGFPFGFVG VGGIPGVAGV 420
 PGVGGVPVG VGPVGIVSPE QAAAAAAAKAA KYAAGAGAVL GGLVPGAPGA VPGVPGTGGV 480
 PGVGTAAAAA AKAAAKAAQF GLVPGVGVAP GVGVAPGVG APGVGLAPGV GVAPGVGVAP 540
 GVGVAPGIGP CGVAAAAKSA AKVAAKAQLR AAAGLGAQIP GLGVGVGVPG LGVGAGVPG 600
 GVAGVPGFQG AGADEGVRRS LSPELREGDP SSSQHLPSTP SSPPRPGALA AAKAAYKGAA 660
 VPGVVLGLGA LGGVGIPGV VGAGPAAAAA AAKAAAKAQF FGLVGAAGLG GLGVGGLGVP 720
 GVGLGGLIPP AAAAKAAKYG AAGLGGVLGG AGQFPLGGV ARPGFGLSPI FPGGACLGKA 780
 CGRKRK 786

SEQ ID NO: 55 moltype = AA length = 338
 FEATURE Location/Qualifiers
 source 1..338
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 55
 MLSAFTLFL ALIGGTSGQY YDYDFPLSIY GQSSPNCAPE CNCPESYPSA MYCDELKLKS 60
 VPMVPPGIKY LYLRNNQIDH IDEKAFENVT DLQWLILDHN LLENSKIKGR VFSKLKQLKK 120
 LHNHNNLNLK SVGPLPKSL E DQLQHNLKIT KLGSFEGL VN LTIFIHLQHNR LKEDAVSAAF 180
 KGLKSLEYLD LSFNQIARLP SGLPVSLLLT YLDNNKLSNI PDEYFKRFNA LQYLRLSHNE 240
 LADSGIPGNS FNVSSLVELD LSYNKLNKIP TVNENLENYY LEVNQLEKFD IKSFKILGP 300
 LSYSKIKHLR LDGNRNISETS LPPDMYECLR VANEVTLN 338

SEQ ID NO: 56 moltype = AA length = 478
 FEATURE Location/Qualifiers
 source 1..478
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 56
 MAPLRPLIL ALLAWVALAD QESCKGRCTE GFNVDKKCQC DELCSYYQSC CTDYTAECKP 60
 QVTRGDFVFTM PEDEYTVYDD GEEKNNATVH EQVGGPSLTS DLQAQSKGNP EQTPVLPKEE 120
 EAAPAEVGAS KPEGIDS RPE TLHPGRPQPP AEEELCSGKP FDAFTDLKNG SLFAFRGQYC 180
 YELDEKAVRP GYPKLI RDVW GIEGPIDAAF TRINCQGKTY LFKGSQYWRF EDGVLDPDYP 240
 RNISDGFDGI PDNVDAALAL PAHSYSGRER VYFFKGKQYW EYQFOHQPSQ EECECGSSLSA 300
 VFEHFAMMQR DSWEDELFELL FWGRTSAGTR QPQFISRDWH GVPQVDAAM AGRIYISGMA 360
 PRPSLAKKQR FRHRNRKGYR SQRGHSGRGRN QNSRRPSRAT WLSLFSSEES NLGANNYDDY 420
 RMDWLVPATC EPIQS VFFF S GDKYYRVNLR T RRVDTVDP YPRSIAQYWL GCPAPGHL 478

SEQ ID NO: 57 moltype = AA length = 1724
 FEATURE Location/Qualifiers
 source 1..1724
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 57
 MPPAVRRSAC SMGWLWIFGA ALGQCLGYSS QQORVPFLQP PGSQLQASY VEFRPSQGCS 60
 PGYRYDHGKL YTGRVCPCNC NGHSNQCQDG SGICVNCQHIN TAGEH CERCQ EGYYGNAVHG 120
 SCRACPCPHT NSFATGCVN GGDVRCSCKA GYTGTQCERCQ APGYFGNPQK FGGSCQPCSC 180
 NSNGQLGSCH PLTGCDCINQE PKDSSPAEEC DDCDSCVMTL LNLDLATMGEQ LRLVKSQQLQG 240
 LSASAGLLEQ MRHMETQAKD LRNQLLNYSR A ISNHGSKIE GLEREI LTDLN QEFETLQEK 300
 QVNSRKAQTL NNNVNRATQS AKELDVKIKN VIRNVHILLK QISGTDGEQN NVPSGDFSRE 360
 WAEAQRMMRE LRNRNPGKHL REAEADKRES QLLNIRIWTW QKTHQGENNG LANSIRDSL 420

- continued

EYEAKLSDLR	ARLQEAAAQA	KQANGLNQEN	ERALGAIQRQ	VKEINSQSD	FTKYLTADS	480
SLLQTNTALQ	LMEKSQKEYI	KLAESTNEAR	QELSDKVREL	SRSAGKTSVL	EEAEKHARSL	540
QELAKQLEEI	KRNASGDELV	RCAVDAATAY	ENILNAIKAA	EDAANRAASA	SESALQTVIK	600
EDLPRKAKTL	SSNSDKLLNK	AKMTQKKLKQ	EVSPALNNLQ	QTLNIVITVQK	EVIDTNLTTL	660
RDGLHGIQRG	DIDAMISSAK	SMVRKANDIT	DEVLDGLNP1	QTDVERIKDT	YGRTOQNEDFK	720
KALTDAHSV	NKLTNKLPLD1	WRKIESINQQ	LLPLGNISDN	MDRIRELIQQ	ARDAASKVAV	780
PMRFNGKSGV	EVRLPNLDED	LKGYSLSLFL	LQRPNSRENG	GTENMFVMYL	GNKDASRDYI	840
GMAVVGDQLT	CVYNLGDREA	ELQVDQILTK	SETKEAVMDR	VKFQRIYQFA	RLNYTKGATS	900
SKPETPGYD	MDGRNSNTLL	NLDPENVVFY	VGGYPPDFKL	PSRLSFPYK	GCIELDDLNE	960
NVLSLYNFFK	TFNNTLITEV	PCRNRKEESD	KNYFEGTGYA	RVPTQPQAPI	PTFGOTIQTT	1020
VDRGLYNNFFAE	NGDRFISLNI	EDGKLMVRYK	LNSELPKERG	VGDAINNCRD	HSIQIKIGKL	1080
QKRMWINNDV	QNTIIDGEVF	DFSTYLYGGI	PIAIRERPN1	STPAFRGCMK	NLKKTSGVVR	1140
LNDTGVVTKK	CSEDWKLVRS	ASFSSRGQLS	FTDLGLPPTD	HLQASFGFQT	FQPSGILLDH	1200
QTWTTRNLQVT	LEDGYVIELST	SDSGPIFKS	PQTYMDGLHH	YVSVISDNG	LRLLIIDDQQL	1260
RNSKRLKHIS	SSRQLSLLGG	SNFEGCNSL	FVQRQLSLSPE	RDVSLGGCSL	1320	
NKPPFLMLK	GSTRFNKTKT	FRINOLLODT	PVASPRSVKV	WQDACSPLPK	TQANHGALQF	1380
GDIPTSHLLF	KLPQELLKPR	SQFAVDMQT	SSRGLVFTHG	TKNSFMALYL	SKGRLVFAFLG	1440
TDGKKLRIKS	KEKCNNDGKWH	TVVFGHDGK	GRLVVDGLRA	REGSLPGNST	ISIRAPVYLG	1500
SPPSGKPKSL	PTNSRDSKPL	TPSSSFVGCLK	NFQLDSSKPLY	CLGGPLEKGI	YFSEEGGHVV	1560
LAHSVLLGPE	FKLVFSIRPR	SLTGILIHIG	SQPGKHLCVY	LEAGKVTASM	DSGAGGTSTS	1620
VTPKQSLCDG	QWHSVAVTIK	QHILHLELDT	DSSYTAGQIP	FPPASTQEPL	HLGGAPANLT	1680
TLRIPVWKSF	FGCLRNIHVN	HIPPVTEAL	EVQGPVSLNG	CPDQ		1724

SEQ ID NO: 58 moltype = AA length = 1172
 FEATURE Location/Qualifiers
 source 1..1172
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 58						
MRPFPLLCFA	LPGLLHAQQA	CSRGCACYPPV	GDLLVGRTRF	LRASSTCGLT	KPETYCTQYG	60
EWQMKCCKCD	SRQPHNYYSH	REVNENASSG	PMRWWQSQND	VNPVSLQLDL	DRRFQQLQEVN	120
MERQGPMPAG	MLIERSSDFG	KTWRVYQYLA	ADCTSTFPRV	QOGRPQSWQD	VRCQSLPQRP	180
NARLNGGKVQ	LNLMDLVSGI	PATQSOKIQE	VGEITNLRVN	FTRLAPVQQR	GYHPSSAYYA	240
VSQLRLQGSC	FCHGHAUDRCA	PKPGASAGPS	TAQVHVDVCB	CQHNTAGPNC	ERCAFYNNR	300
PWRPAEQRGA	HECQNDCDNG	HSETCHFDPA	VFAASQGAYG	GVCDCNCRDHT	EGKNCERCQ	360
HYFRNRRPQA	SIQETCISCE	CDPDGAVPGA	PCDPVTGQCV	CHEHVQGERC	DLCKPGFTG	420
TYANPQGCHR	CDCNILGSRR	DMPCDEESGR	CLCLPNVVG	KCDQCAPYHW	KLASQGQCEP	480
CACDPHNSLS	PQCNQFTGQC	PCREGFGGLM	CSAAAIRQCP	DRTYGDVATG	CRACDCDFRG	540
TEGPGCDKAS	GRCLCRPGLT	GPRCRDQCQ	YCNRYPVCVA	CHPCFQTYDA	DLREQALRFG	600
RLRNATASLW	SGPGLEDRGL	ASRILDAKSK	IEQIRAVLSS	PATVTEQEVAQ	VASAILSLRR	660
TIQLQLQDLP	LEEETLSSLPR	DELESLDRSFN	GLLITMYQRKR	EQFEKISSAD	PSGAFRMLST	720
AYBQSAQAAQ	QVSDSSRLLD	QLRDRSREAB	RIVRQAGGGG	GTGSPKLVAL	RLEMSSLPDL	780
PTPTFNKLCGN	SRQMACTPIS	CPGELCPQDN	GTACGSRGRQ	VLPRAGGAFL	MAGQVAEQLR	840
GFNAQLQRTR	QMIRAAEESA	SQIQLSQAQL	ETQVSASRSQ	MEEDVRRTRL	LIQQVRDFLT	900
DPPDTAATIQ	EVSEAVLALW	LPTDSATVLLQ	KMNEIQIAIA	RLPNVDLVLS	QTQDQIARAR	960
RLQAEAEEAR	SRAHAVEGQV	EDVVGNLRLQ	TVALQEAQDT	MQGTSRSRLR	IQDRVAEVQQ	1020
VLRPAEKLVT	SMTKQLQDFW	TRMEELRHQA	RQOGAEAVQA	QQLAEGASEQ	ALSAQEGER	1080
IIQKYAELKD	RIGQSSMLGE	QGARIQSVKT	EAEELFGETM	EMMDRMKDME	LELLRGSQAI	1140
MLRSDLTGL	EKRVEQIRDH	INGRVLYYAT	CK			1172

SEQ ID NO: 59 moltype = AA length = 1193
 FEATURE Location/Qualifiers
 source 1..1193
 mol_type = protein
 organism = Homo sapiens

SEQUENCE: 59						
MPALWLGCCL	CFSLLLPAAR	ATSRREVCD	NGKSRQCIFD	RELRHRTGNG	FRCLNCNDNT	60
DGIHCEKCKN	GYFYRHRERDR	CLPCNCNSKG	SLSARCDNSG	RCSCPKGVTG	ARCDRCLPGF	120
HMLTDAGCTQ	DQRLLDKSCD	CDPAGIAGPC	DAGRVCVKPA	VTGERCDRC	SGYYNLDGGN	180
PEGCTQCFCY	GHSASCRSSA	EYSVYHAT	FHQDVWDGWKA	VQRNGSPAKL	QWSQRHQDVF	240
SSAQRLDPVY	FVAPAKFLGN	QVSYQGQSL	FDYRVDRGRR	HPSAHDVILE	GAGLRITAPL	300
MPLGKTLPCG	LTKTYTFRNL	EHPSPNNWSPQ	LSYFEYRRL	RNLTLALRIRA	TYGEYSTGYI	360
DNVTLISARP	VSGAPAPWVE	QCICPVGYKG	QFCODCAGSY	KRDSARLGPF	GTCIPCNCG	420
GGACDPDTG	CYSGDENPDI	ECADCPPIGFY	NDPHDPRSSCK	PCPCCHNGFSC	SVMPETEEVV	480
CNNCPPGVTG	ARCELCAADCY	FGDPFGHEGP	VRPCQPCQCN	NNVDPASGN	CDRLTGRCCLK	540
CIHNTAGIYC	DQCKAGYFGD	PLAPNPADKC	RACNCNPMGS	EPVGCRSDGT	CVCKPGFFGP	600
NCEHGAFCSCP	ACYNQVKIQM	DQFMQQLQRM	EALISKAQGG	DGVVPDTELE	GRMQQAEGQAL	660
QDILRDAQIS	EGASRSRSLGLQ	LAKVRSQENS	YQSRLLDDLKM	TVERVRALGS	QYQNRVRDTH	720
RLITQMQSL	AESEASLGN	NIPASDHVY	PNGFKSLAQE	ATRLAESHVE	SASNMEQLTR	780
ETEDYSKQAL	SLVRKALHEG	VGGSGSPDG	AVVQGLVEKL	EKTKSLAQQL	TREATQAEIE	840
ADRSYQHSLR	LLDSVSRQG	VSDQSFQVEE	AKRIKQKADS	LSSLVTRHML	EFKRTQKNLG	900
NWKEEAAQQLL	QNGKSGREKS	DQLLSRANLA	KSRAQEALSM	GNATFYEVES	ILKNLREFDL	960
QVDNRKAEEA	EAMKRLSYIS	QKVSADSKT	QQAERALGSA	AADAQRAKNG	AGEALEISSE	1020
IEQEIGSLNL	EANVTADGAL	AMEKGLASLK	SEMREVEGEL	ERKELEFDTN	MDAVQMVITE	1080
AQKVDTRAKN	AGVTIQDTLN	TLDGLLHLM	QPLSVDEEGL	VLLEQKLSRA	KTQINSQLRP	1140
MMSELEERAR	QQRGHLHLL	TSIDGILADV	KNLENIRDNL	PPGCYNTQAL	EQQ	1193

-continued

```

SEQ ID NO: 60          moltype = AA  length = 1296
FEATURE           Location/Qualifiers
source            1..1296
                   mol_type = protein
                   organism = Clostridium botulinum
SEQUENCE: 60
MPFVNQKFNY KDPVNGVDIA YIKIPNAGQM QPVKAFKIHN KIWVIPERDT FTNPPEEGDLN 60
PPPEAKQVPV SYDSTYLS DNEKDNYLKG VTKLFERIYS TDLGRMLLTS IVRGIPFWGG 120
STIDTELKV DTNCINVIQP DGSYRSEELN LVIIGPSADI IQFECKSFHG EVLNLTTRNGY 180
GSTQYIRFSP DFTFGFESL EVDTNPLGK GKFATDPAVT LAHELIHAGH RLYGIAINPN 240
RVPFKVNTNAY YEMSGLEVS EELRTFGGHD AKPIDSLQEN EFRRLYYNNKE KDIASTLNKA 300
KSIVGTTASL QYMKNVFKEK YLLSEDTSGK FSVDKLKFDFK LYKMLTEIYT EDNFVKFKV 360
LNRKTYLNFD KAVFKINIVP KVNVITYDGF NLRNTNLAAAN FNGQNTIEINN MNFTKLKNFT 420
GLLFEPYKLLC VRGIITSKTK SLDKGYNKAL NDLCIKVNNL SSDFSPSEDN FTNDLNKGEE 480
ITSDTNIEAA EENISLDLIQ QYLYTFNFDN EPENISIENL SSDIIQCLEL MPNIERFPNG 540
KKYELDKYTM PHYLRAQEFE HGKSRIALTN SVNEALLNPS RVYTFFSSDY VKVKVNKATEA 600
AMPLGIVBQL VYDFTDETSE VSTTDKIADI TIIIPYIGPA LNIGNMLYKD DFVGALIFSG 660
AVILLEPIPE IAIPVLUFTFA LVSYIANKL TVQTIDNAL KRNEKWDEVY KYIVTNWLAK 720
VNTOQIDLIRK MCKEALENQA EATKAIINYQ YNQYTEEENK NINFNIDLS SKLNESINKA 780
MININKFLNQ CSVSYLMNSM IPYGVKRLED FDASLKDALI KYIYDNRGLT IGQVDRLSDK 840
VNNTLSTDIP FQLSKYVVDNQ RLLSTFTVEI KNIINTSILN LRYESNHLD LSRYASKINI 900
GSKVNFDPQD KNQIQLFNLN SSKIEVILN AIVYNSMLYEN FSTSFWRIP KFYFNISLNN 960
EYTIINCENM NSGWKVSLNY GEIIWTLQDT QEIKQRVYVFK YSQMINISDY INRWIFVTIT 1020
NNRLNNNSKIY INGRLLIDQKP ISNLGNIHAS NNIMFKLDCG RDTHRYIWIK YFNLFDEKLN 1080
EKEIKDLYDN QSNSGILKDF WGDYLQYDKE YYMLNLYDPN KVVDVNNGVI RGYMYLKGPR 1140
GSVMTTNIYL NSSLYRGTKF IIKKYASGNK DNIVRNNDRV YINVVVKNKE YRLATNASQA 1200
GVEKILSALE IPDVGNLSQLV VVMKSKNNDQ ITNKCKMNLQ DNNGNDIGFI GFHQFNNIAK 1260
LVASNWYINRQ IERSSRTLGC SWEIFPVDDG WGERPL 1296

SEQ ID NO: 61          moltype = AA  length = 1291
FEATURE           Location/Qualifiers
source            1..1291
                   mol_type = protein
                   organism = Clostridium botulinum
SEQUENCE: 61
MVRTINNFNY NDPIDNNNII MMEPPFARGT GRYYKAFKIT DRIWIIPERY TFGYKPEDFN 60
KSSGIFNRDV CEEYDPPDYLQ TNDDKKNIFLQ TMIKLFNRK SKPLGKEKLLIE MIINGIPYLG 120
DRRVPLPEEFN TNIASTVTNQ LISPNPGEVER KKGIFANLII FGPGPVLNEN ETIDIGIQNH 180
FASREGFGGI MQMKFCPEYV SFVNNVQENK GASIFNRRGY FSDPALLMH ELIHVLHGLY 240
G1KVDDLPV PNEKKPFMQS TDAIQAEELY TFQQGDPSII TPSTDKSIYD KVLQNFRGIV 300
DRLNKVLVCI SDPNININII KNKFKDKYKF VEDSEGKYSI DVESFDKLYK SLMFGFTETN 360
IAENYKIKTR ASYFSDSLP VVIKNNLDE IYTIEEGPNI SDKDMEKERYR GQNKAINKQA 420
YEEISKEHLA VVYKIQMKSV KAPGICIDV NEDLFFIAKD NSFSDDLSKN ERIEYNTQSN 480
YIENDFPIE LIIDTDLISK IELPSENTE LTFDNVDPV YEKQPAIKI FTDEDNTIFQY 540
LYSQTFPLDI RDISLTSSFD DALLFSNKVY SFFSMMDYIKT ANKVVVEAGLF AGWVKQIVND 600
FVIEANKSNT MDKIADISLII VPYIGLALNV GNBTAKGNF ENAFEIAGASI LLEFIPPELLI 660
PVVGAFLLES YIDNKNKIIK TIDNALTKRN EWKSDMYGLI VAQWLSTVNT QFYTIKEGMY 720
KALNYQAOAL EEEIKYRYN YSEKEKSNIN IDPNNDINSKL NEGINQAIIDN INNFINGCSV 780
SYLMKKMPL AVEKLDDFDN TLKKNLLNYI DENKLYLIGS AEYEKSUNK YLKTIMEPFDL 840
SIYTNDTILI EMFNKYNSEI LNNNIIILNNRY KDNNLIDLSG YGAKVEVYDG VELNDKNQFK 900
LTSSANSKIR VTQNQNYIIFN SVFLDFSVS WIRIPKYKND GIQNYIHNEY TIINCMKNN 960
GWKISIRGNR IWTTLIDING KTKSVFFFYEN IRBDISEYIN RWFFFVITNN LNNAKIYING 1020
KLESNTDIKD IREVIANGEI IFKLDGDIDR TQFIWMKYF IFNTELSQSN IEERYKIQSY 1080
SEYLKDWFGN PLMYNKEYMM FNAGKNKNSYI KLKKDSPVGE ILTRSKYQNQ SKYINYRDLY 1140
IGEKFIIIRRK SNSQSINDDI VRKEDYRKK FFNLNQEWVR YTYKFKKE EKLFLAPISD 1200
SDEFYNTIQLI KEYDEOPTYS CQLLFFKDEE STDEIGLIGI HRFYESGIVF EYKDYFCIS 1260
KWLKEVKRK PYNLKLGNCNW QFIPKDEGWT E 1291

SEQ ID NO: 62          moltype = AA  length = 2871
FEATURE           Location/Qualifiers
source            1..2871
                   mol_type = protein
                   organism = Homo sapiens
SEQUENCE: 62
MRGRRLLEIA LGFTVLLASY TSHGADANLE AGNVKETRAS RAKRRGGGGH DALKGPNVCG 60
SRYNAYCCPG WKTLPGGNQC IVPICRHSCG DGFCSRPNMC TCPSGQIAPS CGRSRISQHCN 120
IRCMNGGSCS DDHCLCQKGY IGHCGQPVCE ESCGLNGGRC VAPNRCACTY GFTGPQCERD 180
YRTGPCFTVI SNQMCQGQLS GIVCTKTLCC ATVGRANGHP CEMCPAQPHP CRRGPINIR 240
TGACQDVDEC QAIPGLCQGG NCINTVGSFE CKCPAGHKLN EVSQKCEDID ECSTIPGICE 300
GGECTNTVSS YFCKCPGPFY TSPDGTRCID VRPGYCYTAL TNGRCSNQLP QSITKMQCCC 360
DAGRWCSPGV TVAPEMCPIR ATEDFNKLCS VPMVIPGRPE YPPPPLGPPIP PVLPVPPGFP 420
PGPQIPVPRP PVEYLIPSPRE PPRVLPVNVT DYCOLVRYLC QNGRCIPTPG SYRCECNKGF 480
QLDLRGECD VDECEKNPCA GGECCINNOQS YTQCRCAGYQ STLTRTECRD IDECLQNGRI 540
CNGRCINTD GSFHVCNAG FHVTRDGKNC EDMDECSIRN MCLNGMCINE DGSFKCICKP 600
GFQLASDGRY CKDINECETP GICMNGRCVN TDGSYRCECF PGLAVGLDGR VCVDTMRST 660

```

- continued

CYGGYKRGQC	I KPLFGAVTK	SECCC CASTEY	A FGEPCQPCP	A QNSAEYQAL	CSSGPGMTSA	720
GSDINECALD	P DIPCPNGICE	N LRGTYKIC	N SGYEV DSTG	K NCVDINECV	L NSLLCDNGQ	780
CRNTPGSFVC	T CPKGFIYKP	L DLKTCEDIDE	C ESSPCINGV	C KNSPGSFIC	E CSSESTLDP	840
TKTICIEITK	G TCTWQTVIDG	R CEININGAT	L KSQCCSSLG	A AWGSPCTLC	Q VDPICGKY	900
SRIKGTQCED	I DECEVFPGV	C CKNGLCVNTR	G GSFKCQCPSP	M TLDA GRIC	L LDIRETCFL	960
RYEDEECLP	I AGGRHMDAC	C CCSVGA AWG	T EECEECPMR	T PEYEELCPR	G GPFATKEIT	1020
NGKPFKDIN	E CKMIPS LCT	H GKCRNTIGS	F FKCRCDSGFA	L DLSEERNCTD	I DEC RIS PDL	1080
CRGRQCVNTP	G DFECKCDEC	Y EGFM MMK	C MD IDEC QRD	P PLLC RG GVC	H NTEG SYR CEC	1140
PPGHQLSPNI	S ACIDINECE	L SAHLC PGNR	C VN LIG KYQC	A CNPGY HSTP	D RLFCVD I	1200
CSIMNGCET	F CTN SEGSY	C CSCQ FALM	P DQRSCTD ID	E ECDNP NIC	G GGQCTN IPGE	1260
YRCLCYDGF	A SEDMKTCD	V NECDLN PN	C CLSGT CEN	G GSFKCH DMG	Y SGKK GTG	1320
TDINECEIGA	H NC GKH AVCT	N TAGSF KCS	S SPGWIGD	I CTDL DEC SNG	T THMC SQH ADC	1380
KNTMGSYRCL	C KEGYTGDGF	T CTDL DEC S	N NL CNG QNC	L NAPG GYR CE	D CMGF VP SAD	1440
GHAKEDIDE	S SLPNIC VFGT	C CHNL PGL F	E CEI GYEL DR	S GGN CTD VNE	C LDPTT C	1500
NCVNTPGSYI	D CDCP PDF L	E PTRV GVD	F SG NCYLD IRP	G RGD NGD TA	H NEIG VGV SKA	1560
SCCCSLGKAW	G GTP CEMCPA V	H NTSE YKIL CP	I GGE GFR PNP I	J TIVED I	K QEL PGL CQ	1620
KCINTFGFSQ	C CRC PTG YYL	N EDTRV CDD V	O EC ETPG IC GP	P GTC YNTV G	Q TCIC PP DYM	1680
VNGNNCMMD	R RSL CYRN Y	A ADN QT CDG	B LF NM TK KM CC	C CSY NI GR AW	D KPC ECP I	1740
TDEFATLCGS	Q QR PGF VD	I T GLP VD I	E RE IPG VCE	F WCN MVG SFR	G CEC PVG FF	1800
DKL LV CED ID	E CQ NGP V CQR	N A E C INT AGS	P YR DC K PG YR	R FT STG Q CND R	S NEC Q EIP N	1860
SHGQCIDTVG	S FY CLC HT GF	T KTN DD QT M	U DINE CER D	V GNG TCR NT	W SFN CR CN HG	1920
I LSH NN C D	V D ECA S GNG	W LCR N Q C	X VG SF Q C Q	Y CNE	Z GY EV A P D	1980
PRKCAPGTCQ	N LLD GS Y RC	M C	O PPG YL Q NEK	P C E D I	Q E C I P A	2040
PEGFSLSSSG	R R C Q D L R	S M S Y	T C Y A K P E	U G S P K S R	V H N S K	2100
TEPDEAFRQI	C PY GSG II	I V G	P D D S A V D M	O D C E P D V	N K H G	2160
GNECVDTDEC	S V G N P C G N G	T C K N V I G G	Q T C E E G F E	W P G M M T C	X E D I N E C	2220
CVNTYGSYEC	I K P C V G V L	R R M K C D E	S D E E G K H D	T C T K Q M E C K N	U L I G T Y M C I C	2280
QRRPDGECCV	D E N E C Q T K P G	I C E N G R C L	R G S Y T C E C	N D G F T A S P N	Q D E C L D N R E G Y	2340
TEVLQNMQCI	G S S N R N P V	T K	S E C C D G G	W G P H C E I C P	Q G T V A F K K L C	2400
ADIDECKVIH	D V C R N G E C V	R D G S Y H C I	T G Y T P D I	S C V D L N E C	N Q A P K C P N F I	2460
NTEGYSYQSC	P K G Y I L Q E D G	R S C K D L D E	R S C K D L D E	T C K Q H N C Q F L C	V N T I G G F T C K	2520
SCIDNNECTS	D I N L C G S K G I	C Q N T G S F T C	E C Q R G F E S L D Q	G T G S C E D V D E	C E G N H R C Q H G	2580
CQNIIGGYRC	S C P Q G Y L Q H Y	Q W N Q C V D	S L C A H I C G G A	S C H N T L G S Y K	C M C P A G F Q Y E	2640
QFSGGCQDIN	E C G S A Q A P C S	Y G C S N T E G	Y L C G C P P G Y F R	I G Q G H C V S G M	G M G R G N P E P P	2700
VSGEMDDNSL	S P E A C Y E C K I	N G Y P K R G R K R	R S T N E T D A S N	I E D Q S E T E A N	V S L A S W D V E K	2760
TAIFAFNISH	V S N K V R I L E	L P A L T L T L N H	N R Y L I E S G N E	D G F F K I N Q K E	G I S Y L H F T K K	2820
KPVAGTYSLQ	I S S T P L Y K K K	E L N Q L E D K Y D	K D Y L S G E L G D	N L K M K I Q V L L	H	2871

```
SEQ ID NO: 63          moltype = AA      length = 2912
FEATURE          Location/Qualifiers
source           1..2912
                 mol_type = protein
                 organism = Homo sapiens
```

```
SEQUENCE: 63
MGRRRLCLQ  LYFLWLGCVV  LWAQQTAGQP  QPPPQKPPR  QPPPQQV RSA  TAGSEGGFLA  60
PEYREEGAAV  ASRVRRRGQQ  DVLRGPNVCG  SRFHSYCCP  WKTLPGGNQC  IVPICRN SCG  120
DGFCSPRNMC  TCSSQQISST  CGSKSIQQCS  VRCMNGGTCA  DDHCQCQKGY  IGTYCGQPV C  180
ENG CQNGGRC  IGP NRCAVY  GFTGPQ CERD  YRTG PCFT QV  NNQMCQGQLT  GIVCTKTLCC  240
ATIGRAWHGP  CEMCPAQQPQ  CRRGFIPNIR  TGAC QDV D E C  QAI PGIC QGG  NCINTVGS F E  300
CRC PAGHKQS  ETTQ K CEDID  ECSII PGICE  TGECSNTVGS  YFCVCP RGYV  TSTDG SRC ID  360
QRTGMC F SGL  VNGR CAQ ELP  GRM T K MQ C C C  EPGRCW G I G T  IPEACPV RGS  E E Y R R L C M D G  420
LPMGGIP F SGL  GSRPGG TGGN  GFA PGS NGA  YGP GGTG FIP  I PGG NGF S P G  VGGAGVGAG G  480
QGP I IT GLT I  LNQ TIDICKH  HAN L C L N G R C  I P T V S S Y R C E  C N M G Y Q D A N  G D C I D V D E C T  540
SNPCTNGDCV  NTPGSY YCKC  HAG F Q R T P T K  Q A C I D I D E C I  Q N G V L C K N G R  C V N T D G S F Q C  600
I C N A G F E L T T  DGK N C V D H D E  C T T T M C L N G  M C I N E D G S F K  C I C K P G F V L A  P N G R Y C T D V D  660
E C Q T P G I C M N  G H C I N S E G S F  R C D C P G L A V G  G M D G R V C V D T  H M R T C Y G G I  K K G V C V R P F P  720
G A V T K S E C C C  A N P D Y G F G E P  C Q P C P A K N S A  E F H G L C S S G V  G I T V D G R D I N  E C A L D P D I C A  780
N G I C E N L R G S  Y R C N C N S G Y E  P D A S G R N C I D  I D E C L V N R L L  C D N G L C R N T P  G S Y S C T C P P G  840
Y V F R T E T E T  D E I N E C S N P  C V N G A C R N N L  G S P N C E C S P G  S K L S S T G L I C  I D S L K G T C W L  900
N I Q D S R C E V N  I N G A T L K S C E  C A T L Q G A A W S  P C E R C E L D T  C P R G L A R I K G  V T C E D V N E C E  960
V F F G V C P N G R  C V N S K G S F H C  E C P G E L T D G  T G R V C L D I R M  E Q C Y L K W D E D  E C I H V P G K F  1020
R M D A C C C A V G  A A W G T E C E E C  P K P G T K E Y E T  L C P R G A G F A N  R G D V L T G R P F  Y K D I N E C K A F  1080
P G M C T Y G K C R  N T I G S F K C R C  N S G F A L D M E E  R N C T D I D E C R  I S P D L C G S G I  C V N T P G S F E C  1140
E C F E G Y E S G F  M M M K N C M D I  E C E R N P L L C R  G G T C V N T E G S  F Q C D C P L G H G  L S P S R E D C V D  1200
I N E C S L D N L  C R N G K C V N M I  G T Y Q C S C N P G  Y Q A T P D R Q G C  T D I D E C M I M N  G G C D T Q C T N S  1260
E G S Y E C S C E  G Y A L M P D G R S  C A D I C E C N N  P D I C D G G Q C T  N I P G E Y R C L C  Y D G F M A S M D M  1320
K T C I D V N E C D  L N S N I C M F G E  C E N T K G S F I C  H C Q L G Y S V K K  G T T G C T D V D E  C E I G A H N C D M  1380
H A S C L N I P G S  F K C S C R E G W I  G N G I K C I D L D  E C S N G T H Q C S  I N A Q C V N T P G  S Y R C A C S E G F  1440
T G D G F T C S D V  D E C A E N I N L C  E N G Q C L N V P G  A Y R C E C E M G F  T P A S D S R S C Q  D I D E C S F Q N I  1500
C V F G T C N N L P  G M F H C I C D D G  Y E L D R T G G N C  T D I D E C A D P I  N C V N G L C V N T  P G R Y E C N C P P  1560
D F Q L N P T G V G  C V D N R V G N C Y  L K F G P R G D G S  L S C N T E I G V G  V S R S S C C S L  G K A W G N P C E T  1620
C P P V N S T E Y Y  T L C P G G E G F R  P N P I T I I L E D  I D E C Q E L P G L  C Q G G N C I N T F  G S F Q C E C P Q G  1680
Y Y L S E D T R I C  E D I D E C F A H P  G V C G P G T C Y N  T L C N Y T C I C P  P E Y M Q V N G G H  N C M D M R K S F C  1740
Y R S Y N G T T C E  N E L P F N V T K R  M C C C T Y N V G K  A W N K P C E P C P  T P G T A D F K T I  C G N I P G F T F D  1800
I H T G K A V D I D  E C K E I P G I C A  N G V C I N Q I G S  F R C E C P T G F S  Y N D L L V C E D  I D E C S N G D N L  1860
C Q R N A D C I N S  P G S Y R C E C A A  G F K L S P N G A C  V D R N E C L E I P  N V C S H G L C V D  L Q G S Y Q C I C H  1920
```

-continued

NGFKASQDQT	MCMVDVDECER	HPCGNGTCKN	TVGSYNCLCY	PGFELTHNNND	CLDIDECCSSF	1980
FGQVCRNGRC	FNEIGSFKCL	CNEGELYELTPD	GKNCIDTNEC	VALPGSCSPG	TCQNLEGSFR	2040
CICPPGYEVK	SENCIDINEC	DEDPNICLFG	SCTNTPGGFQ	CLCPPGFVLS	DNGRRCFDTR	2100
QSFCFTNFEN	GKCSVPKAFN	TTKAKCCCSK	MPGEGWGDPC	ELCPKDDEVA	FQDLCPYGHG	2160
TVPSSLHDTR	DVNECLESPG	ICSNQQCINT	DGSFRCECPM	GYNLDYTGVR	CVDTDCECSIG	2220
NPCGNGTCTN	VIGSFECNCN	EGFEPGPMMN	CEDINECAQN	PLLCAFRCMN	TFGSYECTCP	2280
IGYALREDQK	MCKDLDECAE	GLHDCESRGM	MCKNLIGTFM	CICPPGMARR	PDGEGCVDEN	2340
ECRTKPGICE	NRGCVNIIGS	YRCECNEFGQ	SSSSGTECLRD	NRQGLCFEAEV	LQTICQMASS	2400
SRNLVTKSEC	CCDGGRGWGH	QCECLPLPGT	AQYKKICPHG	PGYTTDGRDI	DECKVMPNLC	2460
TNGQCINTMG	SFRFCFKVGY	TTDISGTSCI	DLECSQS PK	PCNYICKNTE	GSYQCSCP R	2520
YVLQEDGKTC	KDLDECQTKQ	HNCQFLCVNT	LGGFTCKCP	GFTQHHTACI	DNNNECGSQPS	2580
LCGAKGICQN	TPGSFSCECQ	RGFSDATQ	NCEDVDECQD	NHRCQHGCQN	ILGGYRCGCP	2640
QGYIQHYQWN	QCVDENCSN	PNACGSASCY	NTLGSYKCAC	PSGFSFDQFS	SACHDVNECS	2700
SSKNPNCNYGC	SNTEGGYLCG	CPPGYYRVQG	GHCVSGMGFN	KGQYLSLDTE	VDEENALSPE	2760
ACYECKINGY	SKKDSRSQRKS	IHEPDPTAV E	QISLESVMDM	SPVNMKFNL	HLGSKEHILE	2820
LRPAIQPLNN	HIRYVISQGN	DDSVFRIHQ	NGLSYLHTAK	KKLMPGTYTL	EITSIPLYKK	2880
KEKKLEESN	EDDYLLGELG	EALRMRLQI	LY			2912

SEQ ID NO: 64	moltype = AA	length = 2809				
FEATURE	Location/Qualifiers					
source	1..2809					
	mol_type = protein					
	organism = Homo sapiens					
SEQUENCE: 64						
MTLEGLYLAR	GPLARLLLAW	SALLCMAGGQ	GRWDGALEAA	GPGRVRRRGS	PGILOGPNV	60
GSRFHAYCCP	GWRTFPGRSQ	CVVPICCRAC	GEFGCSQP	N CTCADGTLAP	SCGVSRGSGC	120
SVSCMNGTC	RGASCLCQKG	YTGTVCQPI	CDRGCHNGRR	C1GPNRCA	YGFMGQ	180
DYRTGPGFGQ	VGPEGCQHQL	TGLVTKALC	CATVGR	WGL PCELCPAQPH	PCRRGFIPN	240
HTGACQDVDE	CAQAVPLCQG	GSCVNMVGFS	HCRCPVGHRL	SDSSAAC	EDY RAGACFSVLF	300
GGRGAGDLAG	HYTRRQCCCD	RGRCWAAQPV	PELCPPRGSN	E FQQLCAQRL	PLLPGHPGL	360
PGLLGFNSNG	MGPPLGPARL	NPHGSDARGI	PSLGPGNSN	I GTATLNQ	TID ICRHFTNLCL	420
NGRCLPTPSS	YRCECNVGYT	QDVRGECIDV	DECTSSPCH	GDCVNIPGTY	HCRCY	480
TPTRQACVVD	DECIVSGGLC	HLGRCVNT	TEG SFQCVNAGF	ELSPDGKNCV	DHNECATSTM	540
CVNGVCLNED	GSFSCLCQKG	FLLAPGGH	YC MDI	DECQTPG	E SFRQCQCLG	600
GLAVGTDGRV	CVDTHRSTC	YGAIEKGSCA	RPFPGTVK	I ECCCANPDHG	FGEPCOLCPA	660
KDSAEEQALC	SSSLGITTDG	RDINECALDP	EVCANGVCEN	L RG SYRCV	CN LGYEAGASGK	720
DCTDVDECAL	NSLLCDNGWC	QNSPGSYSCS	CPPGFHF	WQD TEICKDV	DEC LSSPCVSGVC	780
RNLAGSTCKT	CGPGSRSLDP	GTFCLDSTKG	TWLKIQ	BESR CEVNLQGASL	RSECCATLGA	840
AWGSPCRCE	IDPACRGFA	RMTGTV	CDV DDV	NECESFPGV	C PNGRCVNTAG	900
MLDASGRLCV	DVRLEPCFLR	WDEDEC	GVTL PGKYRMDV	C C S I G A V W G V E	CEACDPESL	960
EFA	ASRDFL	SGL	RDY RY KDV	N EC KV FP GLC	L THG T C H G T K H	1020
AQERNCTDID	ECRISPDL	QG	T C V N T P G S	F D G E C I D V	Y C H D C I D V	1080
LCRGGTCTNT	DGSYKCQCP	G	H E L T A K G	C E D I D E C S L	D G L C P H G Q C V	1140
HAGFQSTPDR	QGCVDINCR	V Q N G G C D V H	I N T E G S Y R C S	C P Q D F E L N P S	G R A C A D V D E C	1200
EENPRVCDQG	HCTNMPGGHR	C L C Y D G F M A T	P D M R T C V D V D	E C D L N P H I C L	H G D C E N T K G S	1260
FVCHCQLGYM	VRKGATGCS	V D E C E V G G H N	C D S H A S C L N I	P G S F S C R C L P	G W V G D G F E C H	1320
DLDECVSQEH	RCS	P G D C L N	V P G S Y R C T C R	Q G F A G D G F F C	E D R D E C A E N V	1380
APG	GYCE	M G F D P T E D H	A C Q D V D E C A Q	G N L C A F G S C E	N L P G M F R C I C	1440
GNCTDINECA	DPVNCINGVC	I N T P G S Y C	C P Q D F E L N P S	G V G C V D T R A G	N C F L E T H D R G	1500
DSGISCSAEI	CGVGVTRASCC	C S L G R A W G N P	C E L C P M A N T	E Y R T L C P G G E	G F Q P N R I T V I	1560
LED	IDE	P G L C Q G G D C V	N T F G S F Q C E C	P P G Y H L S E H T	R I C E D I D E C S	1620
CYNTLGN	YTC	N T F G S F Q C E C	P G S F S C R C L P	R I C E D I D E C S	T H S G I C G P G T	1680
GQAWNRPCEA	CPTPISPDYQ	I L C Q N Q A P G F	L T D I H T G P K	D I D E C G E I P A	I C A N G I C I N Q	1740
IGSFRCECPA	GFNYNSILLA	C E D V D E C S R	E S P C Q Q N A D C	I N I P G S Y R C K	C T R G Y K L S P G	1800
GACVGRN	EIP	P N V C S H G D	C M D T E G S Y	L C H R G F Q A S A	Q T L C M D I D E	1860
CKNIIGSYNC	LC	C Q Q G D C V	T T L V G Q V C R F	G H C L N T A G S F	H C L C Q D G F E L	1920
TADGKNCVDT	NE	C L F P G F V V T H	T L P G T C Q N L E G	S F R C I C P P G F	Q V Q S D H C I D I	1980
LFGTCTNSP	G	N G D C V D F D E C	S F Q C L C P P G	D T R Q S F C F T R	F E A G K C S V P K	2040
CSKRPGEWG	W	S F Q C L C P P G	V L S D N G H R C F	A F N T K T K R C C	A H L C A H G E C I	2100
VNTDGSFRCE	C P F G Y S L D F T	K N T K G S F L C S	G I N C V D T D E C	A H L C A H G E C I	N S L G S F R C H C	2160
MMTCEDIDE	S L N P L L C A F R	C H N T E G S Y L C	S V G H P C G Q G T	Q A G Y T P D A T A	Q A G Y T P D A T A	2220
RGMECKN	L I G	T C P A E Y L Q V N	T C P A G Y T L R E	D G A M C R D V D E	C A D G Q Q D C H A	2280
EGFQPSPTLT	E C H D I R Q G P C	F A E V L Q T M C R	T D D N E C H A Q P	D L C V N G R C V N	T A G S F R C D C D	2340
LPGTSA	R K L	F D Q A L G G C Q E	G R D V D E C R M L	A H L C A H G E C I	N S L G S F R C H C	2400
TTC	L D M D E C S	Q P V K P C T F L C	K N T K G S F L C S	C P R G Y L L E D D	G T R C K D L D E C	2460
CVNTVGAF	T C	R C P P G F T Q H H	A Q C F D N E C S	G H C H N T P G S F R	C E C H Q G F T L V	2520
SSGHG	C E D V N	Q A C F D N E C S	A Q P G P C G A H G	S Q W A Q C V D E N	E C A L S P P T C G	2580
SASC	R N T L G G	E C D V C P S G F D	G C Q N Q L G G Y R	C S P C Q G F T Q H	D R E D V N C E A E	2640
FRAGQGHCVS	G L G F S P G P Q D	F D Q A L G G C Q E	V D E C A G R R G P	C S Y S C A N T P G	G F L C G C P Q G Y	2700
TLDSE	A L L T L	T P D K E E L L S S	E A C Y E C K I N G	L S P R D R P R S	A H R D H Q V N L A	2760
LQLGRRRPGP	G T Y R L E V V S H	M A G P W G V Q P E	G Q P G P W G Q A L	R L K V Q L Q L L		2809

1. A composition comprising:
 - (a) a herpes virus comprising a recombinant herpes virus genome, wherein the recombinant herpes virus genome comprises a first polynucleotide encoding a first collagen protein; and
 - (b) an excipient, wherein the recombinant herpes virus genome does not comprise a polynucleotide encoding a Collagen alpha-1(VII) chain (COL7) polypeptide.
2. The composition of claim 1, wherein the recombinant herpes virus genome is replication defective.
3. The composition of claim 1, wherein the recombinant herpes virus genome is a recombinant herpes simplex virus type 1 (HSV-1) genome.
4. The composition of claim 3, wherein the recombinant HSV-1 genome comprises an inactivating mutation in a herpes simplex virus gene selected from the group consisting of Infected Cell Protein (ICP) 0, ICP4, ICP22, ICP27, ICP47, thymidine kinase (tk), Long Unique Region (UL) 41, and UL55.
5. (canceled)
6. (canceled)
7. The composition of claim 1, wherein the first collagen protein is selected from the group consisting of a human Collagen alpha-1(I) chain polypeptide (COL1-1), a human Collagen alpha-2(I) chain polypeptide (COL1-2), a human Collagen alpha-1(II) chain polypeptide (COL2), a human Collagen alpha-1(III) chain polypeptide (COL3), a human Collagen alpha-1(IV) chain polypeptide (COL4-1), a human Collagen alpha-2(IV) chain polypeptide (COL4-2), a human Collagen alpha-3(IV) chain polypeptide (COL4-3), a human Collagen alpha-4(IV) chain polypeptide (COL4-4), a human Collagen alpha-5(IV) chain polypeptide (COL4-5), a human Collagen alpha-6(IV) chain polypeptide (COL4-6), a human Collagen alpha-1(V) chain polypeptide (COL5-1), a human Collagen alpha-2(V) chain polypeptide (COL5-2), a human Collagen alpha-3(V) chain polypeptide (COL5-3), a human Collagen alpha-1(VI) chain polypeptide (COL6-1), a human Collagen alpha-2(VI) chain polypeptide (COL6-2), a human Collagen alpha-3(VI) chain polypeptide (COL6-3), a human Collagen alpha-4(VI) chain polypeptide (COL6-4), a human Collagen alpha-5(VI) chain polypeptide (COL6-5), a human Collagen alpha-6(VI) chain polypeptide (COL6-6), a human Collagen alpha-1(VIII) chain polypeptide (COL8), a human Collagen alpha-1(IX) chain polypeptide (COL9-1), a human Collagen alpha-2(IX) chain polypeptide (COL9-2), a human Collagen alpha-3(IX) chain polypeptide (COL9-3), a human Collagen alpha-1(X) chain polypeptide (COL10), a human Collagen alpha-1(XI) chain polypeptide (COL11-1), a human Collagen alpha-2(XI) chain polypeptide (COL11-2), a human Collagen alpha-1(XII) chain polypeptide (COL12), a human Collagen alpha-1(XIII) chain polypeptide (COL13), a human Collagen alpha-1(XIV) chain polypeptide (COL14), a human Collagen alpha-1(XV) chain polypeptide (COL15), a human Collagen alpha-1(XVI) chain polypeptide (COL16), a human Collagen alpha-1(XVII) chain polypeptide (COL17), a human Collagen alpha-1(XVIII) chain polypeptide (COL18), a human Collagen alpha-1(XIX) chain polypeptide (COL19), a human Collagen alpha-1(XX) chain polypeptide (COL20), a human Collagen alpha-1(XXI) chain polypeptide (COL21), a human Collagen alpha-1(XXII) chain polypeptide (COL22), a human Collagen alpha-1(XXIII) chain polypeptide (COL23), a human Collagen alpha-1(XXIV) chain polypeptide (COL24), a human Collagen alpha-1(XXV) chain polypeptide (COL25), a human Collagen alpha-1(XXVI) chain polypeptide (COL26), a human Collagen alpha-1(XXVII) chain polypeptide (COL27), and a human Collagen alpha-1(XXVIII) chain polypeptide (COL28).
8. (canceled)
9. (canceled)
10. The composition of claim 1, wherein the recombinant herpes virus genome further comprises a second polynucleotide encoding a second protein.
11. (canceled)
12. (canceled)
13. The composition of claim 1, wherein the composition is suitable for intradermal administration or superficial injection.
14. (canceled)
15. (canceled)
16. A method of diminishing one or more dermatological signs of aging in a subject, the method comprising administering to the subject an effective amount of a composition comprising:
 - (a) a herpes virus comprising a recombinant herpes virus genome, wherein the recombinant herpes virus genome comprises a first polynucleotide encoding a first collagen protein; and
 - (b) an excipient.
17. (canceled)
18. The method of claim 16, wherein the subject is a human.
19. The method of claim 16, the composition is administered topically, transdermally, subcutaneously, epicutaneously, intradermally, orally, sublingually, buccally, rectally, vaginally, intraurethrally, intravenously, intraarterially, intramuscularly, intraosseously, intracardially, intraperitoneally, transmucosally, intravitrally, subretinally, intraarticularly, peri-articularly, locally, or via inhalation to the subject.
20. The method of claim 16, wherein the composition is administered intradermally or via superficial injection to the subject.
21. (canceled)
22. The method of claim 16, wherein the recombinant herpes virus genome is a recombinant herpes simplex virus type 1 (HSV-1) genome.
23. The method of claim 22, wherein the recombinant HSV-1 genome comprises an inactivating mutation in a herpes simplex virus gene selected from the group consisting of Infected Cell Protein (ICP) 0, ICP4, ICP22, ICP27, ICP47, thymidine kinase (tk), Long Unique Region (UL) 41, and UL55.
24. (canceled)
25. (canceled)
26. The method of claim 16, wherein the first collagen protein is selected from the group consisting of a human Collagen alpha-1(I) chain polypeptide (COL1-1), a human Collagen alpha-2(I) chain polypeptide (COL1-2), a human Collagen alpha-1(II) chain polypeptide (COL2), a human Collagen alpha-1(III) chain polypeptide (COL3), a human Collagen alpha-1(IV) chain polypeptide (COL4-1), a human Collagen alpha-2(IV) chain polypeptide (COL4-2), a human Collagen alpha-3(IV) chain polypeptide (COL4-3), a human Collagen alpha-4(IV) chain polypeptide (COL4-4), a human Collagen alpha-5(IV) chain polypeptide (COL4-5), a human Collagen alpha-6(IV) chain polypeptide (COL4-6), a human Collagen alpha-1(V) chain polypeptide (COL5-1), a human Collagen alpha-2(V) chain polypeptide (COL5-2), a human

Collagen alpha-3(V) chain polypeptide (COL5-3), a human Collagen alpha-1(VI) chain polypeptide (COL6-1), a human Collagen alpha-2(VI) chain polypeptide (COL6-2), a human Collagen alpha-3(VI) chain polypeptide (COL6-3), a human Collagen alpha-4(VI) chain polypeptide (COL6-4), a human Collagen alpha-5(VI) chain polypeptide (COL6-5), a human Collagen alpha-6(VI) chain polypeptide (COL6-6), a human Collagen alpha-1(VIII) chain polypeptide (COL8), a human Collagen alpha-1(IX) chain polypeptide (COL9-1), a human Collagen alpha-2(IX) chain polypeptide (COL9-2), a human Collagen alpha-3(IX) chain polypeptide (COL9-3), a human Collagen alpha-1(X) chain polypeptide (COL10), a human Collagen alpha-1(XI) chain polypeptide (COL11-1), a human Collagen alpha-2(XI) chain polypeptide (COL11-2), a human Collagen alpha-1(XII) chain polypeptide (COL12), a human Collagen alpha-1(XIII) chain polypeptide (COL13), a human Collagen alpha-1(XIV) chain polypeptide (COL14), a human Collagen alpha-1(XV) chain polypeptide (COL15), a human Collagen alpha-1(XVI) chain polypeptide (COL16), a human Collagen alpha-1(XVII) chain polypeptide (COL17), a human Collagen alpha-1(XVIII) chain polypeptide (COL18), a human Collagen alpha-1(XIX) chain polypeptide (COL19), a human Collagen alpha-1(XX) chain polypeptide (COL20), a human Collagen alpha-1(XXI) chain polypeptide (COL21), a human Collagen alpha-1(XXII) chain polypeptide (COL22), a human Collagen alpha-1(XXIII) chain polypeptide (COL23), a human Collagen alpha-1(XXIV) chain polypeptide (COL24), a human Collagen alpha-1(XXV) chain polypeptide (COL25), a human Collagen alpha-1(XXVI) chain polypeptide (COL26), a human Collagen alpha-1(XXVII) chain polypeptide (COL27), and a human Collagen alpha-1(XXVIII) chain polypeptide (COL28).

27. (canceled)

28. (canceled)

29. The method of claim 16, wherein the recombinant herpes virus genome further comprises a second polynucleotide encoding a second protein.

30. (canceled)

31. The composition of claim 10, wherein the second protein is selected from the group consisting of a fibronectin protein, an elastin protein, a lumican protein, a vitronectin protein, a vitronectin receptor protein, a laminin protein, a neuromodulator protein, and a fibrillin protein.

32. The method of claim 29, wherein the second protein is selected from the group consisting of a fibronectin protein, an elastin protein, a lumican protein, a vitronectin protein, a vitronectin receptor protein, a laminin protein, a neuromodulator protein, and a fibrillin protein.

33. The method of claim 16, wherein the diminishing of one or more dermatological signs of aging is indicated by the: (a) treatment, reduction, and/or prevention of fine lines and/or wrinkles; (b) reduction of skin pore size; (c) improvement in skin thickness, plumpness, and/or tautness; (d) improvement in skin smoothness, suppleness, and/or softness; (e) improvement in skin tone, radiance, and/or clarity; (f) improvement in procollagen and/or collagen production; (g) improvement in skin texture and or promotion of retextrurization; (h) improvement in appearance of skin contours; (i) restoration of skin luster and/or brightness; (j) improvement of skin appearance decreased by aging and/or menopause; (k) improvement in skin moisturization; (l) increase in skin elasticity and/or resiliency; (m) treatment, reduction, and/or prevention of skin sagging; (n) improvement in skin firmness; (o) reduction of pigment spots, mottled skin, and/or scars; (p) improvement of optical properties of skin by light diffraction or reflection; or (q) any combinations thereof.

34. The method of claim 16, wherein the recombinant herpes virus genome does not comprise a polynucleotide encoding a Collagen alpha-1(VII) chain polypeptide (COL7).

* * * * *