



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.	
	<i>A61K 8/65</i>	(2006.01)
	<i>A61K 8/64</i>	(2006.01)
	<i>A61Q 19/08</i>	(2006.01)
(52)	U.S. Cl.	
	CPC	<i>A61K 8/65</i> (2013.01); <i>A61K 8/64</i> (2013.01); <i>A61Q 19/08</i> (2013.01); <i>A61K 2800/86</i> (2013.01); <i>A61K 2800/91</i> (2013.01); <i>C12N 2710/16621</i> (2013.01); <i>C12N 2710/16631</i> (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.

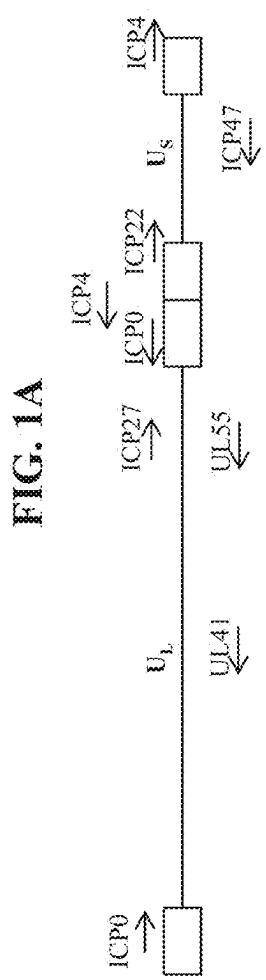


FIG. 1B

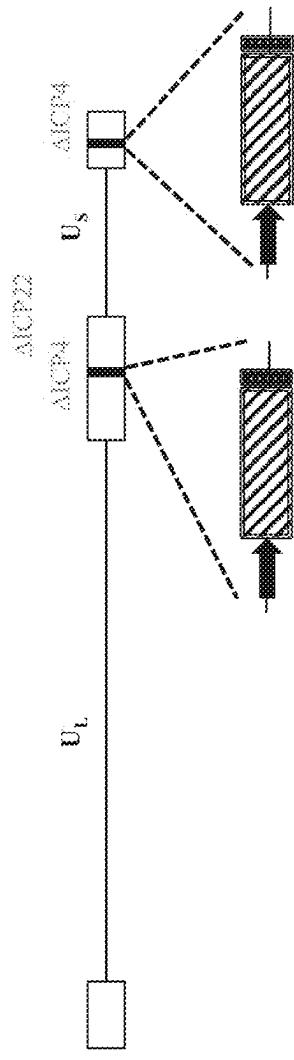
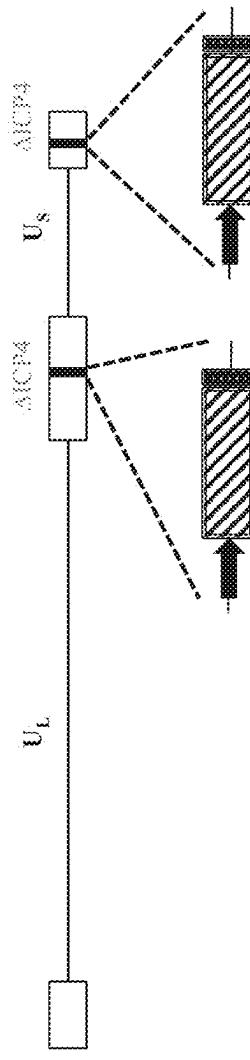


FIG. 1C



→ = heterologous promoter



= 1st human collagen polypeptide coding sequence



= regulatory elements

FIG. 1D

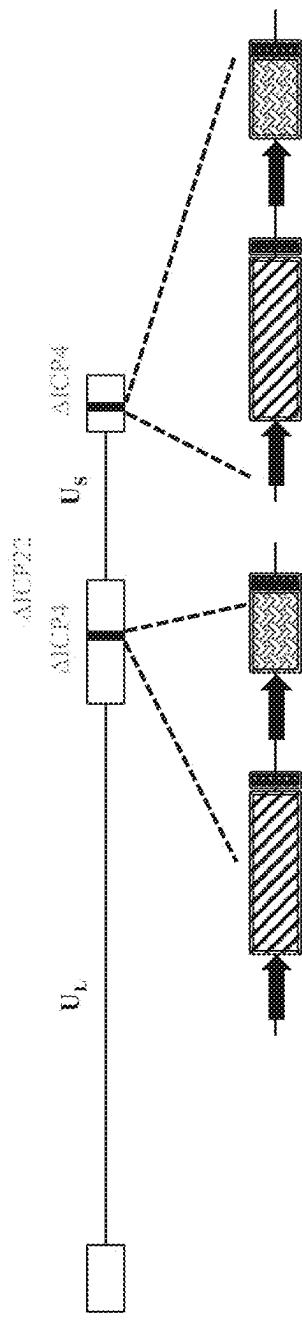


FIG. 1E

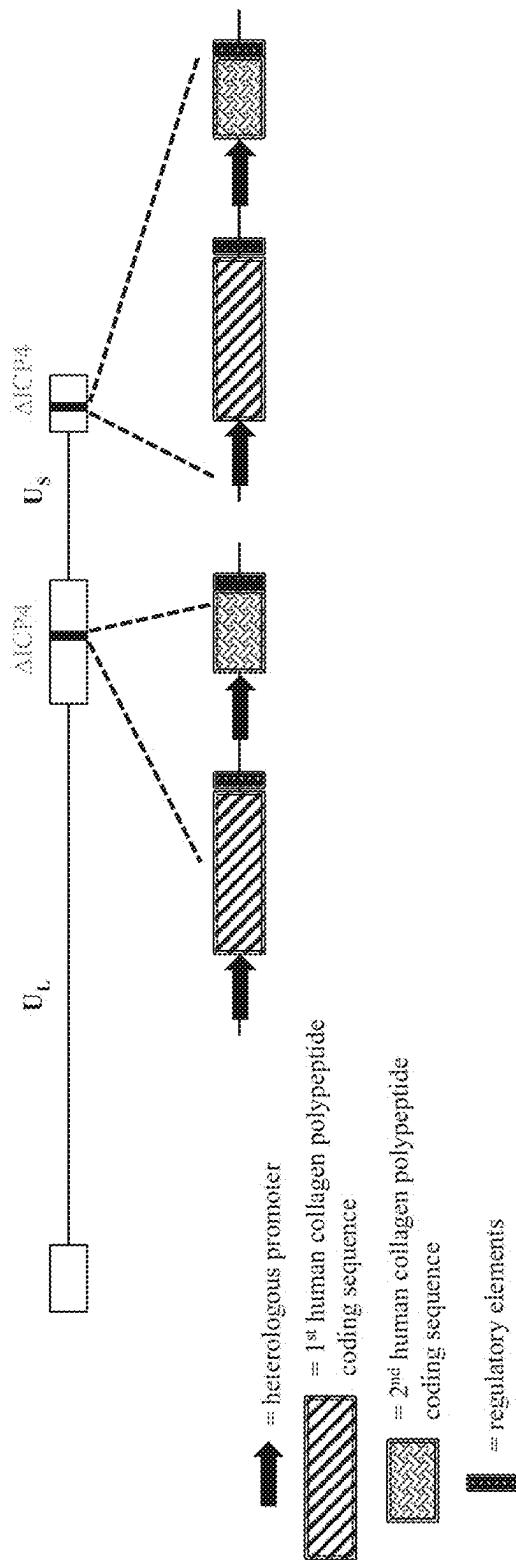


FIG. 1F

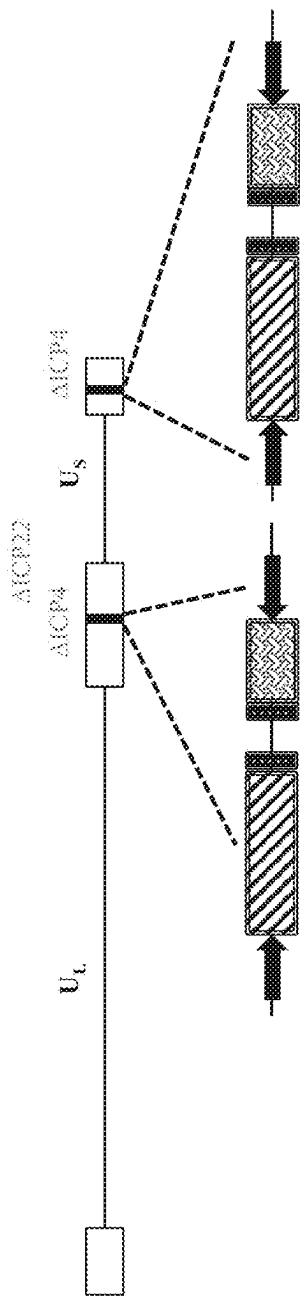


FIG. 1G

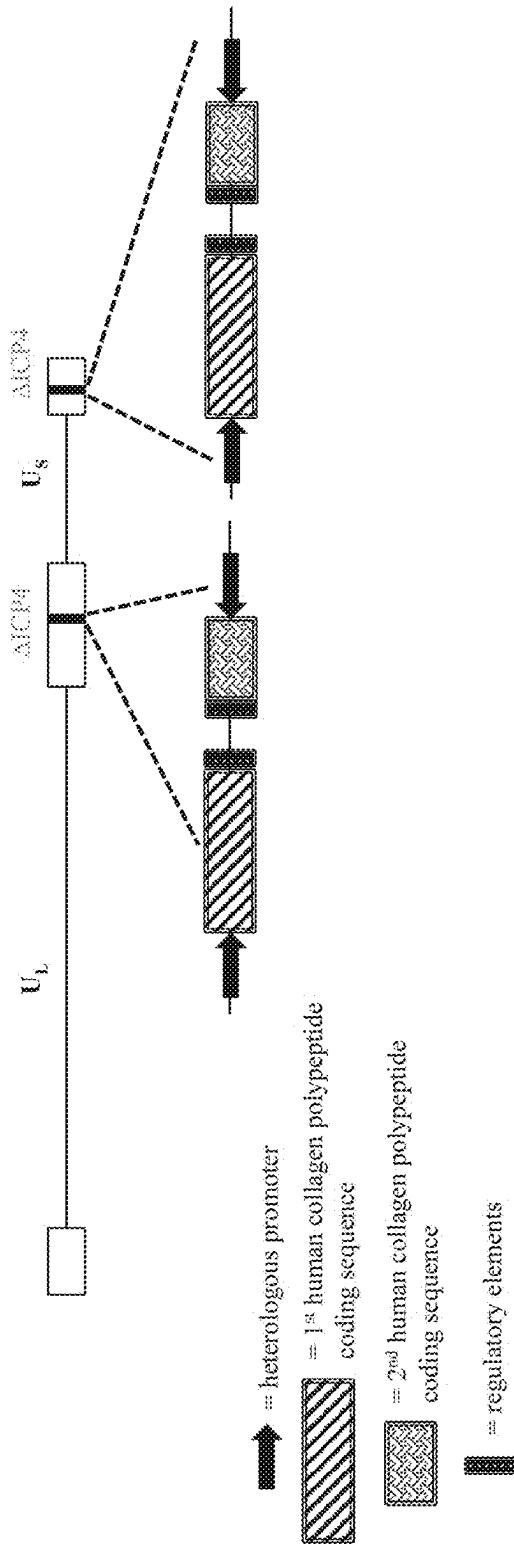


FIG. 1H

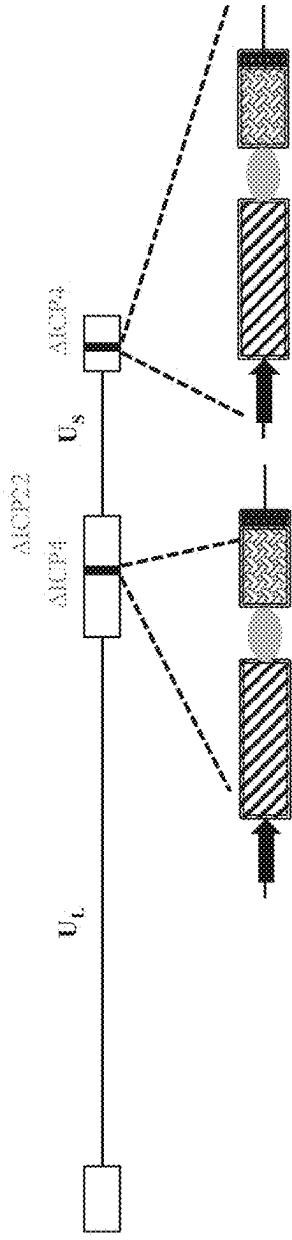


FIG. 1I

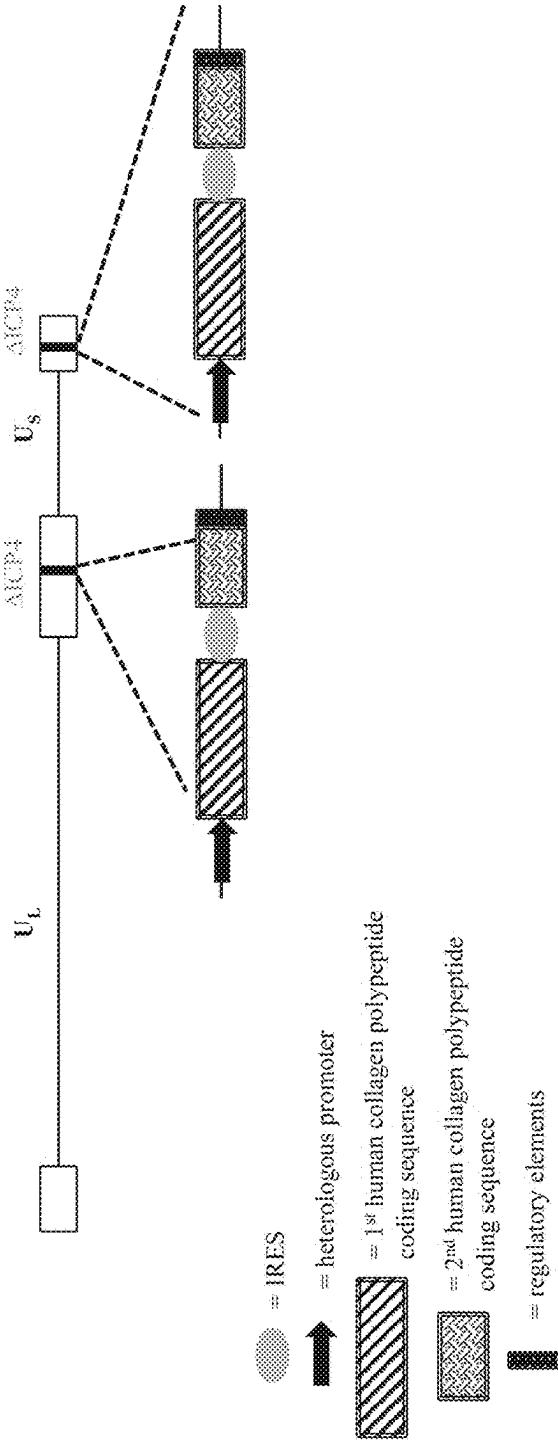


FIG. 1J

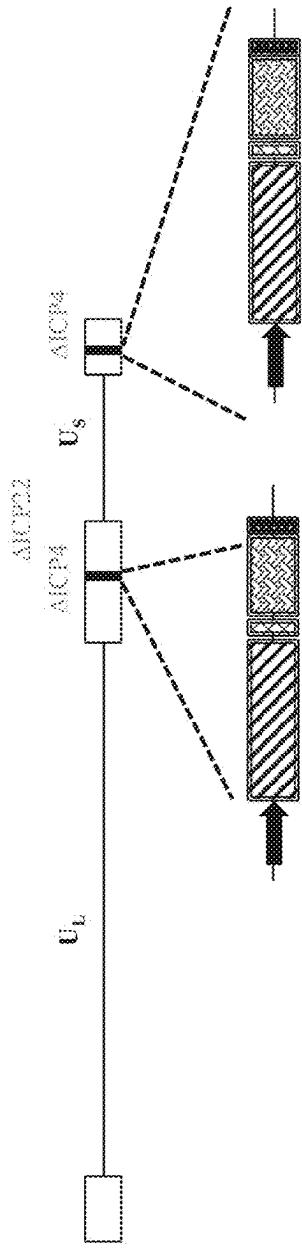


FIG. 1K

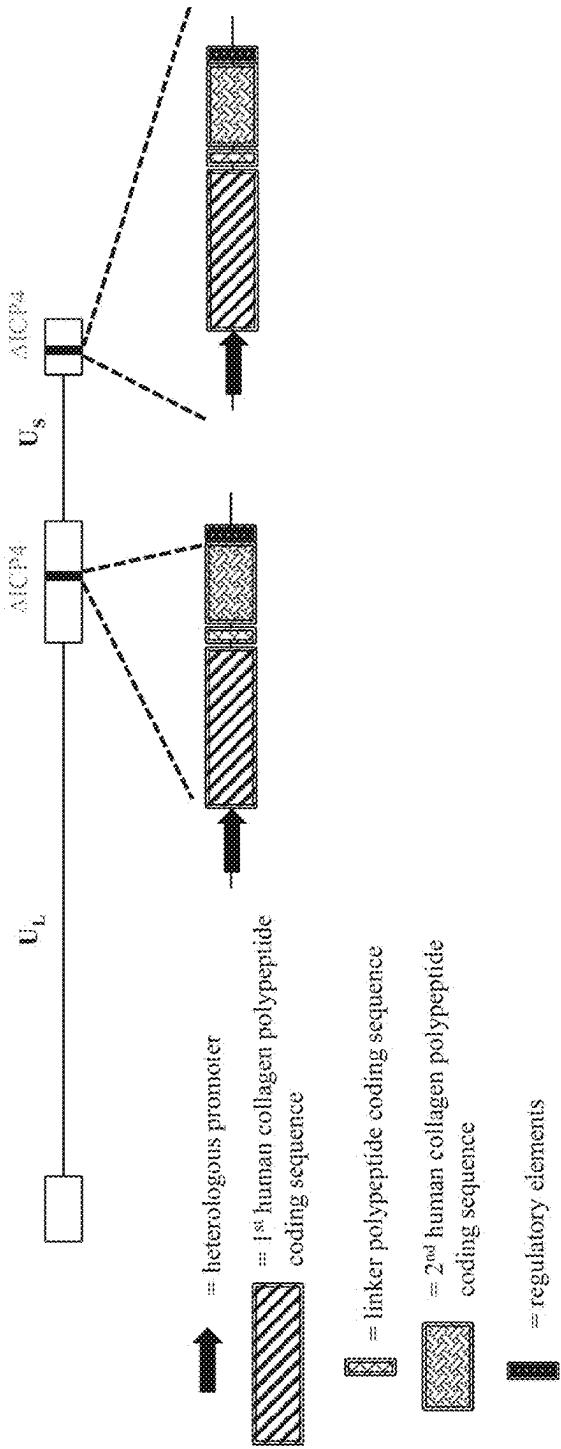


FIG. 11.

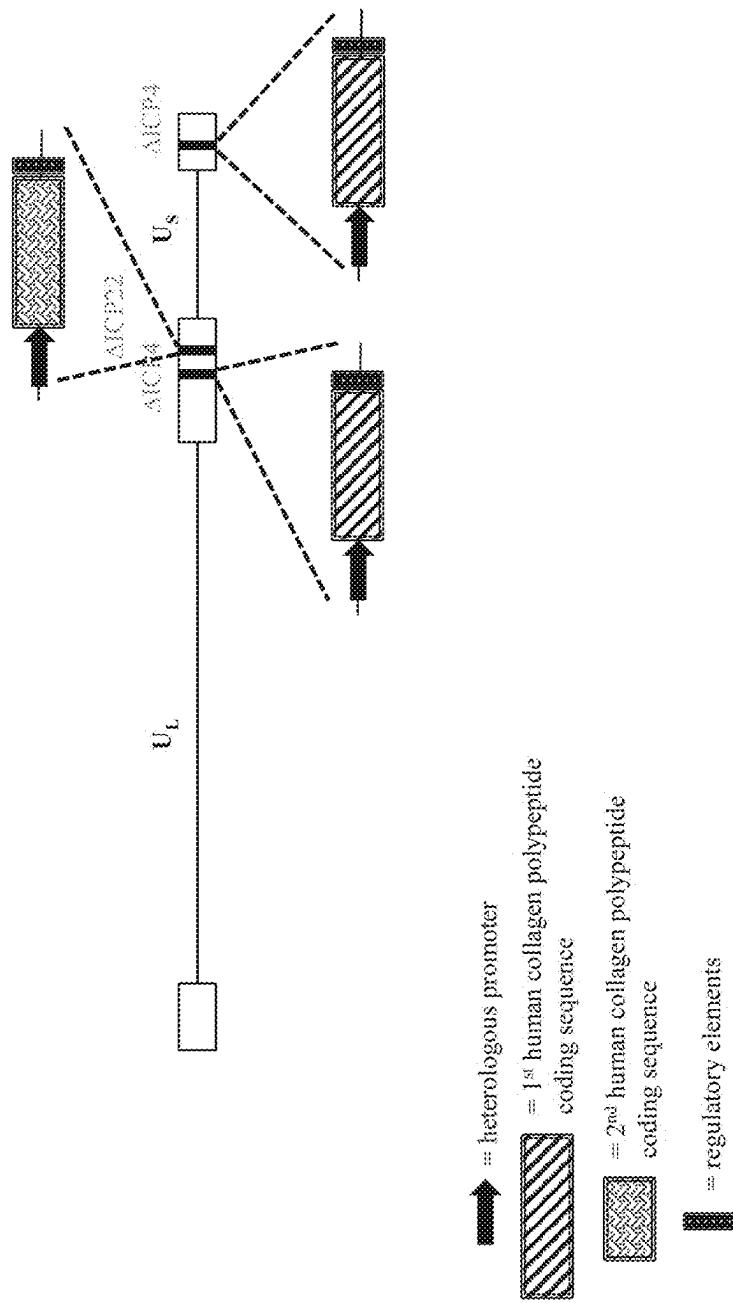


FIG. 1M

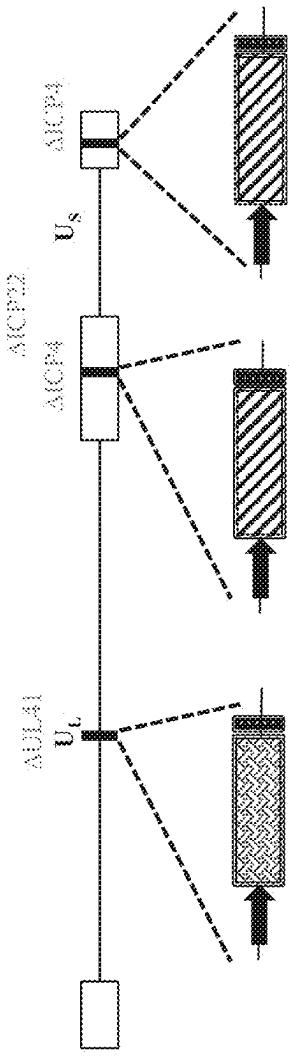


FIG. 1N

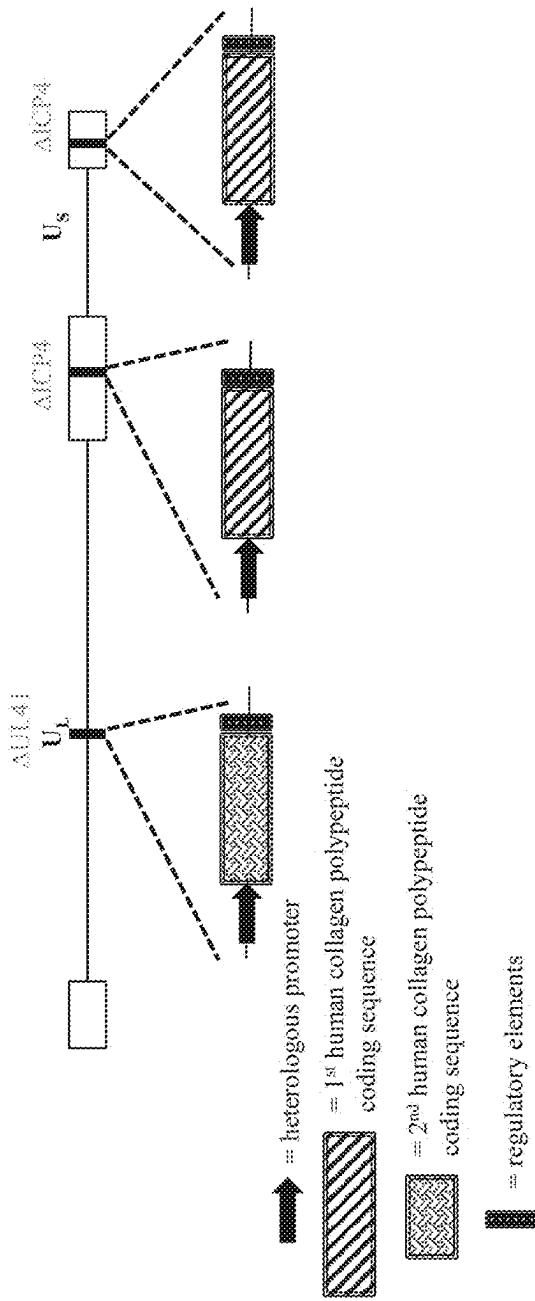


FIG. 2A

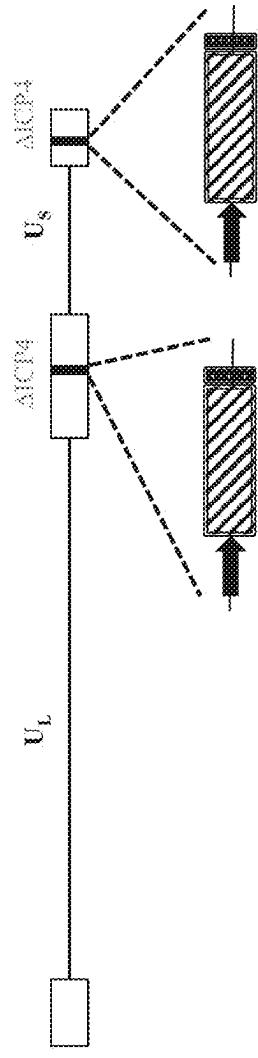


FIG. 2B

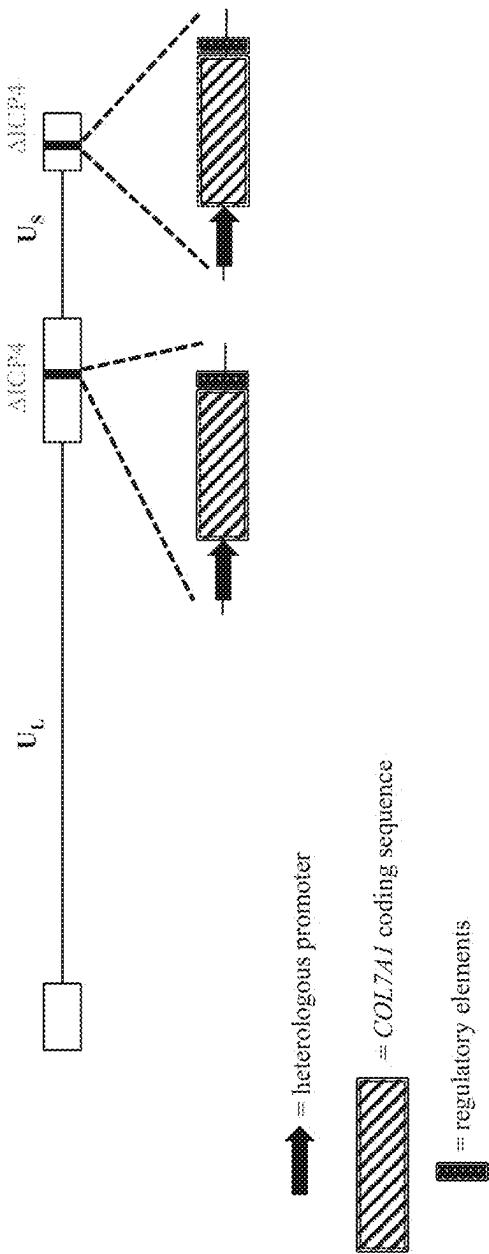


FIG. 3A

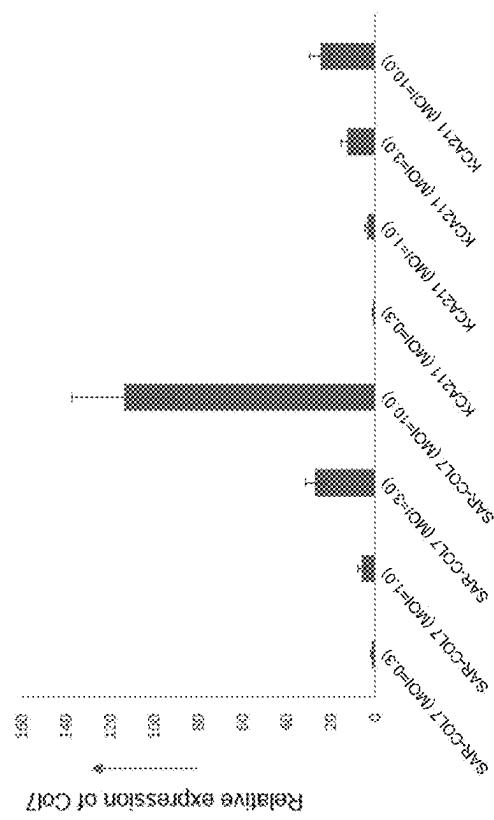


FIG. 3B

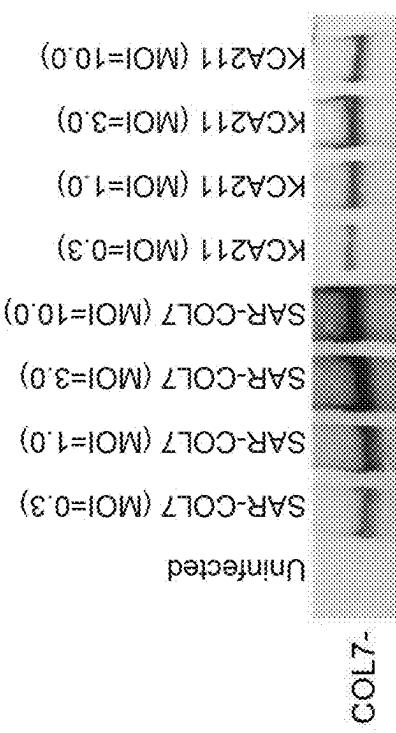


FIG. 4A

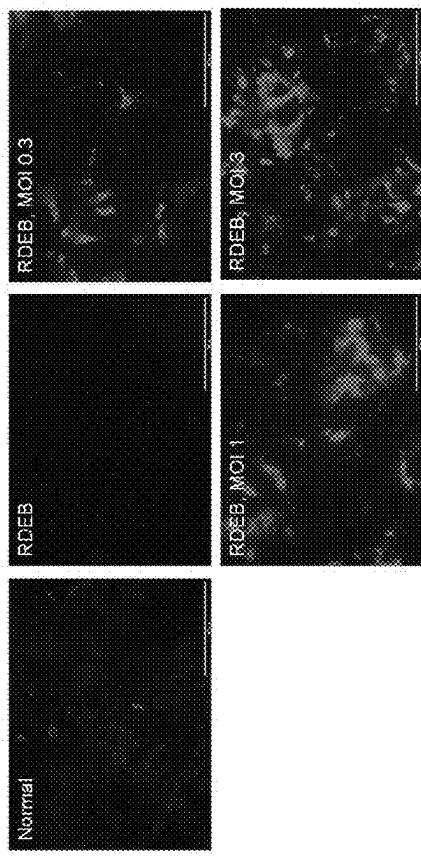


FIG. 4B

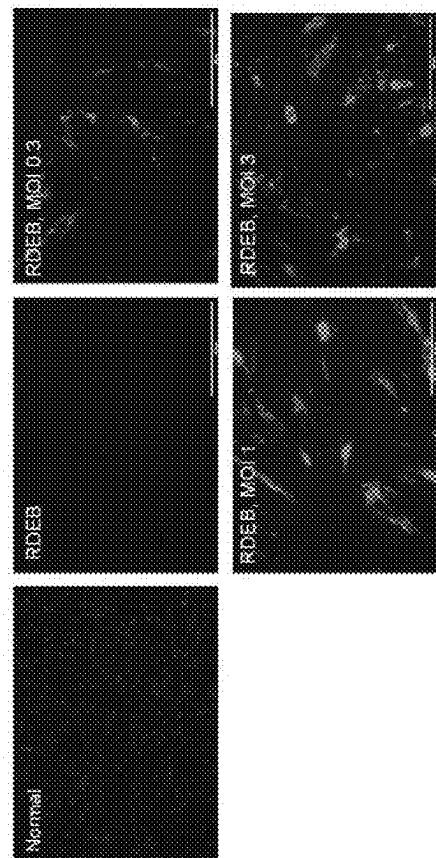


FIG. 5A

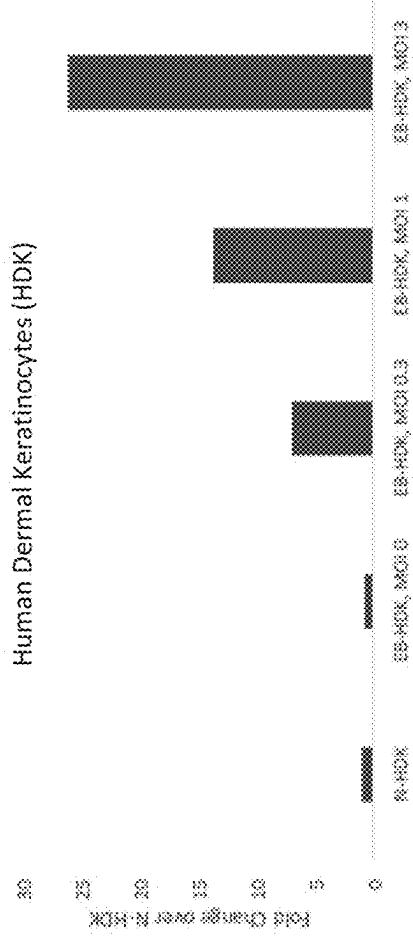


FIG. 5B



FIG. 6A
Adhesion to Collagen I

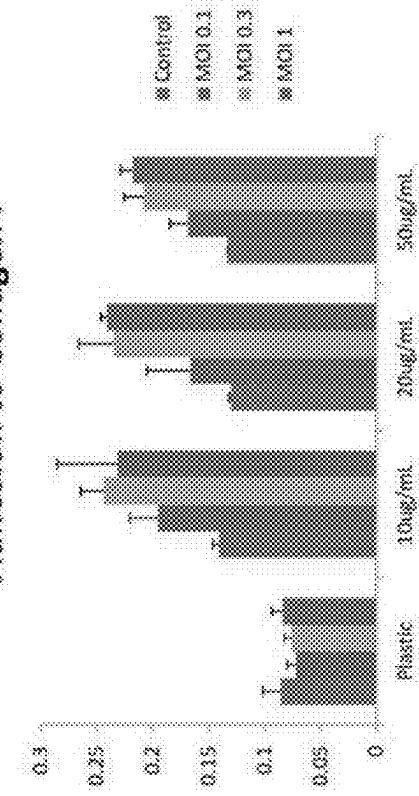


FIG. 6B
Adhesion to Fibronectin

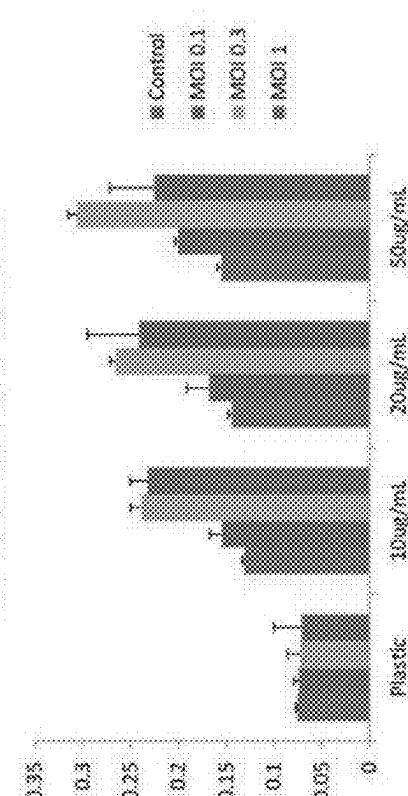


FIG. 7

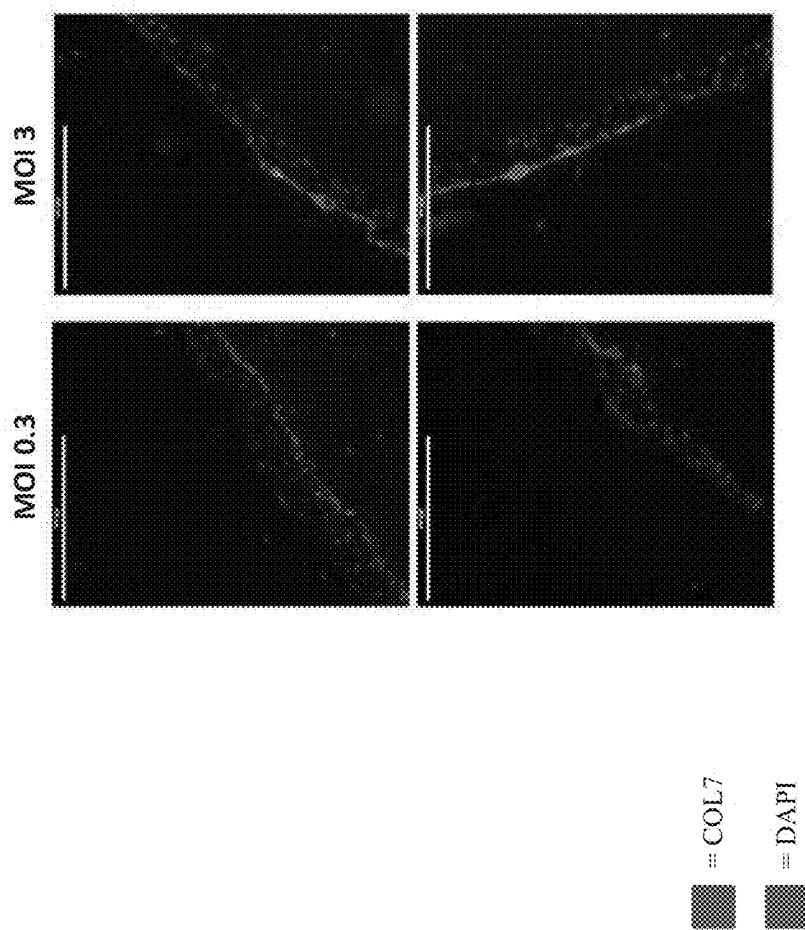


FIG. 8A

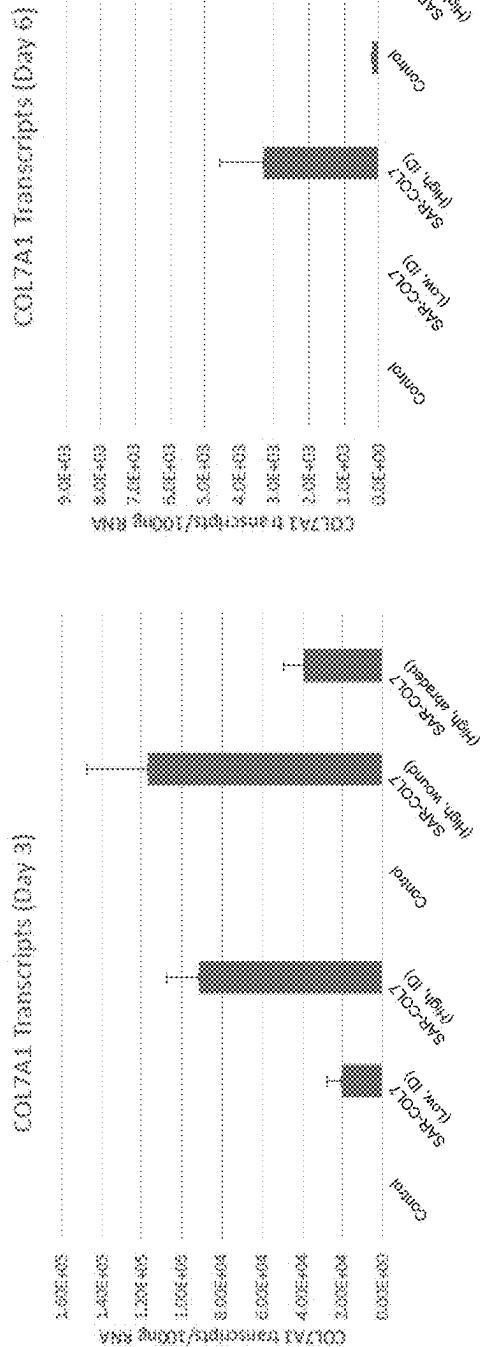


FIG. 8C

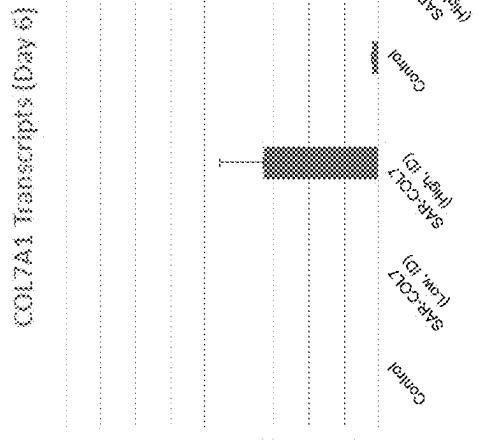


FIG. 8B

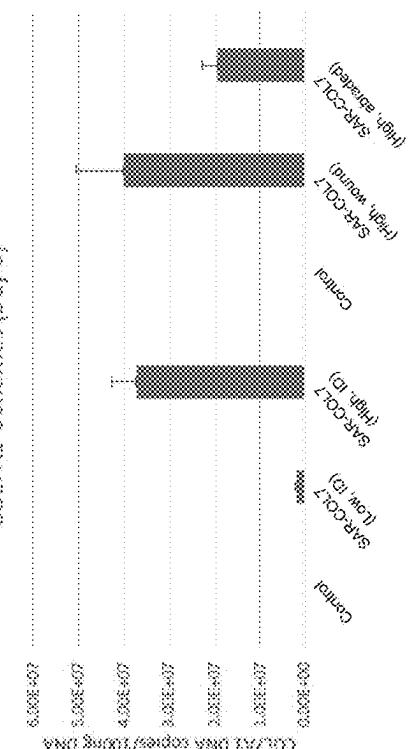


FIG. 8D

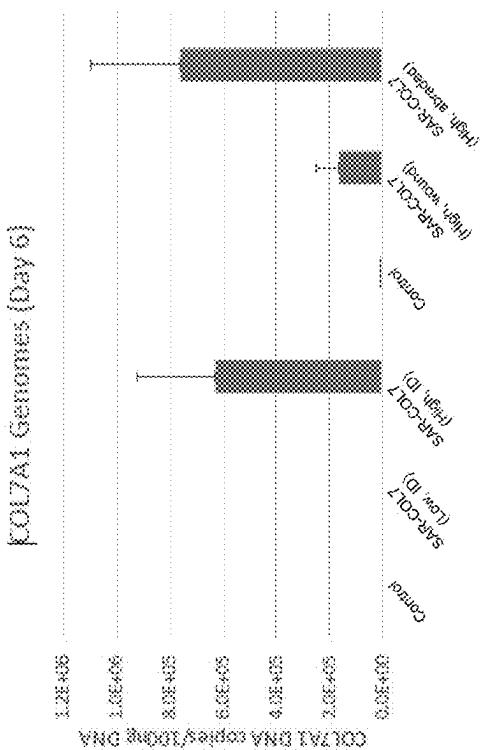


FIG. 9A
FIG. 9B

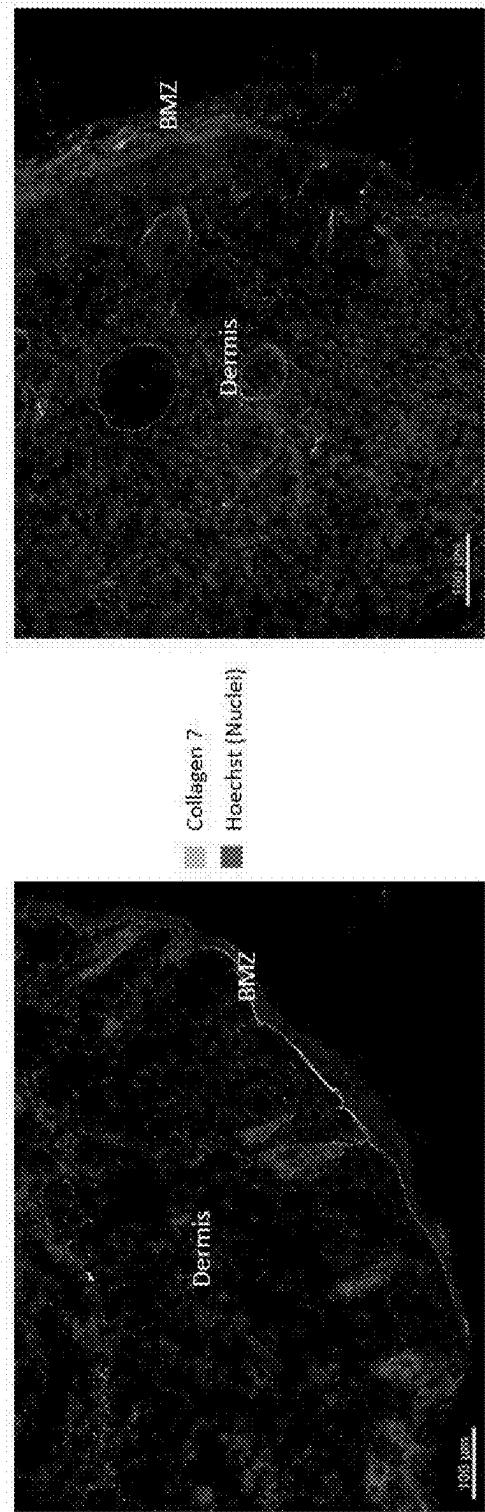


FIG. 10A

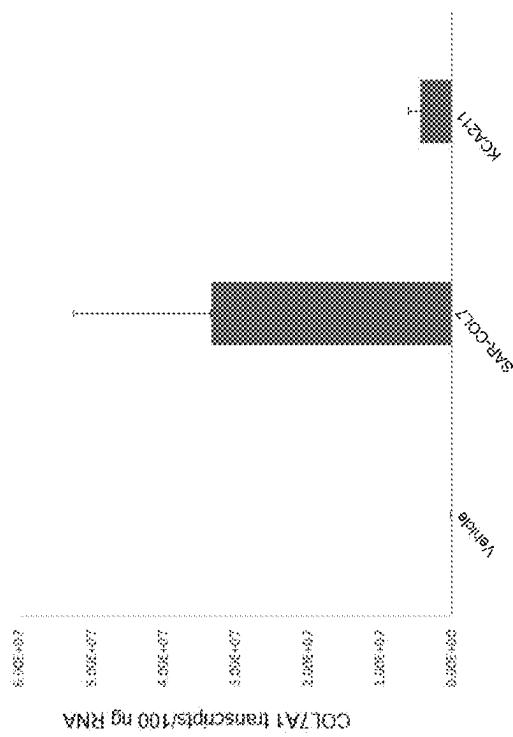


FIG. 10B

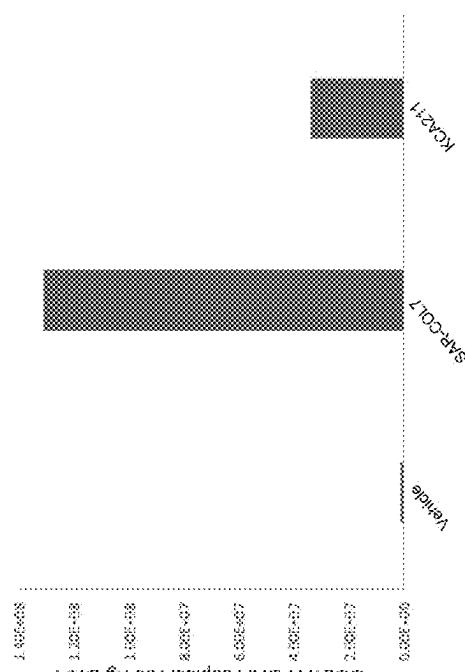
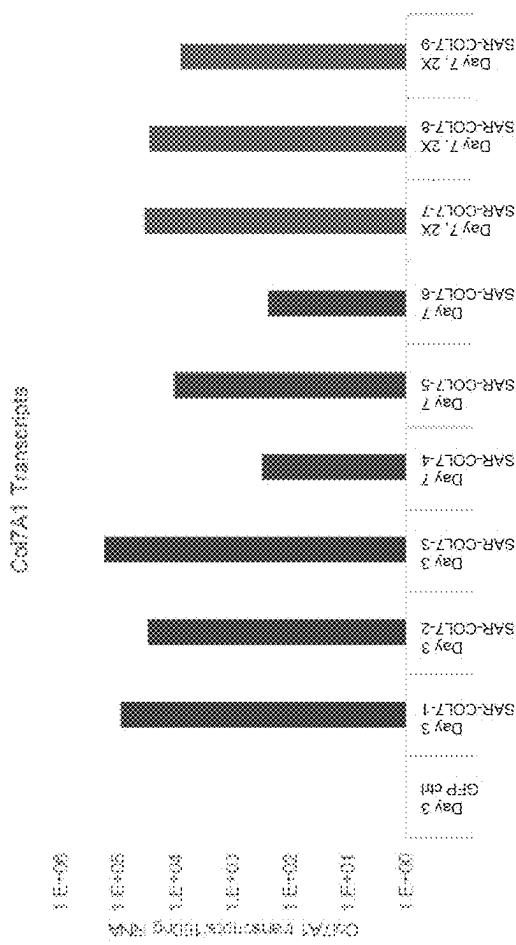


FIG. 1A



13
14

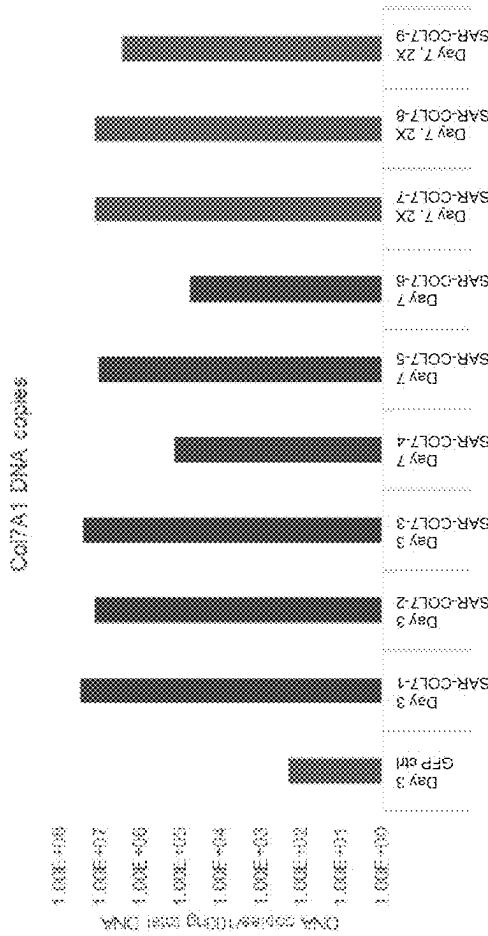


FIG. 12A

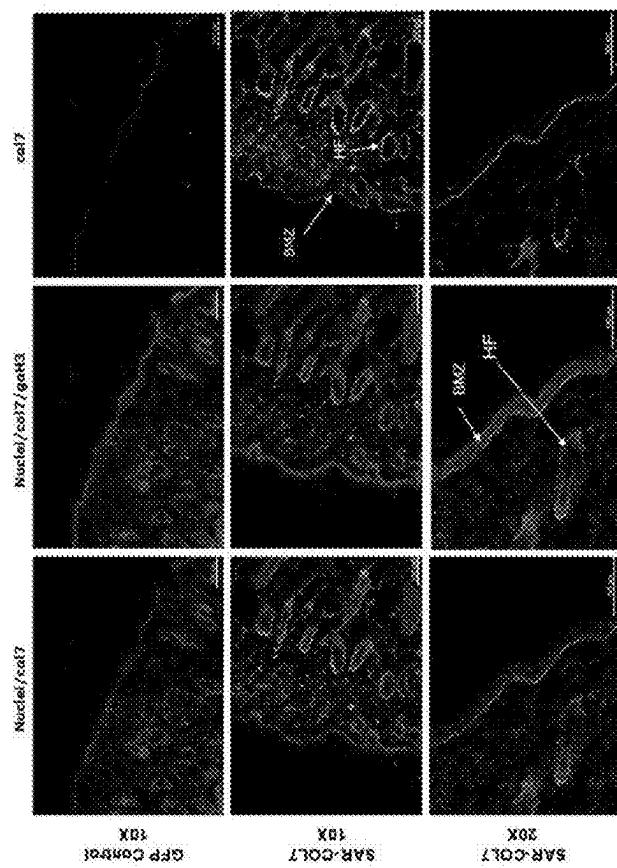
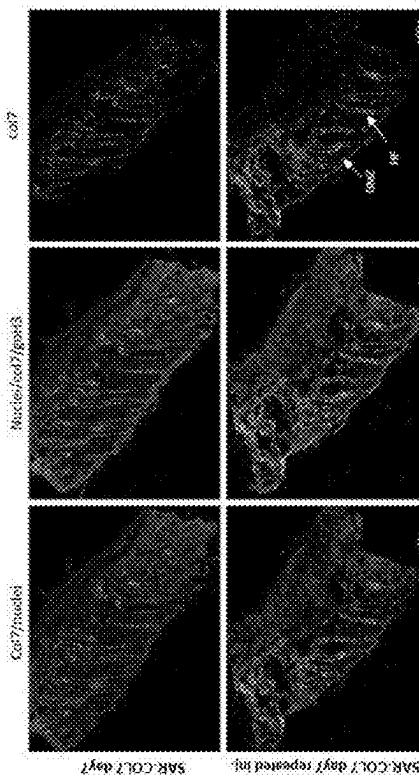


FIG. 12B



१३

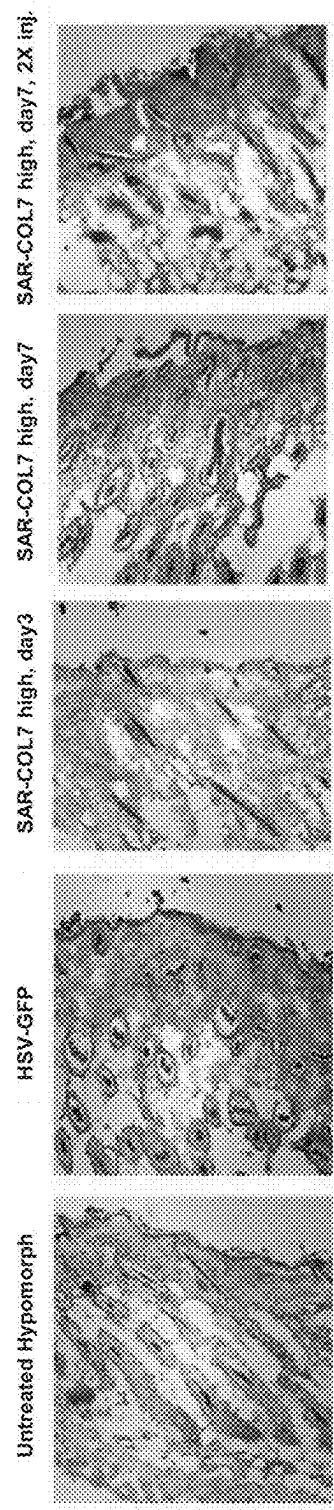


FIG. 14A
FIG. 14B

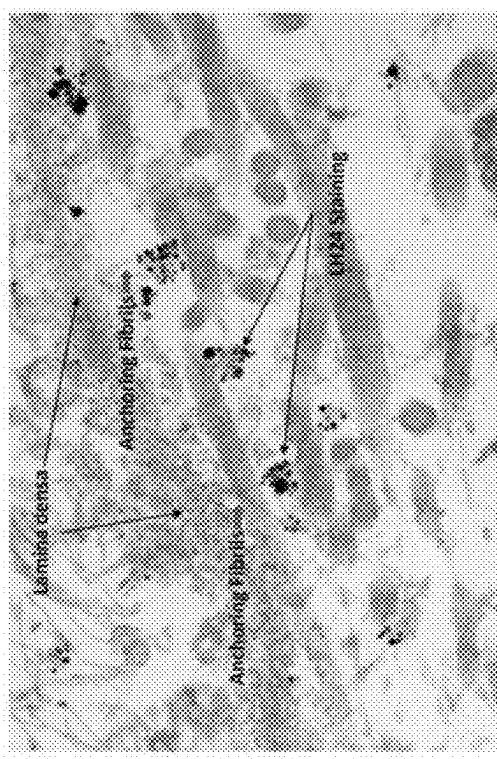


FIG. 15A

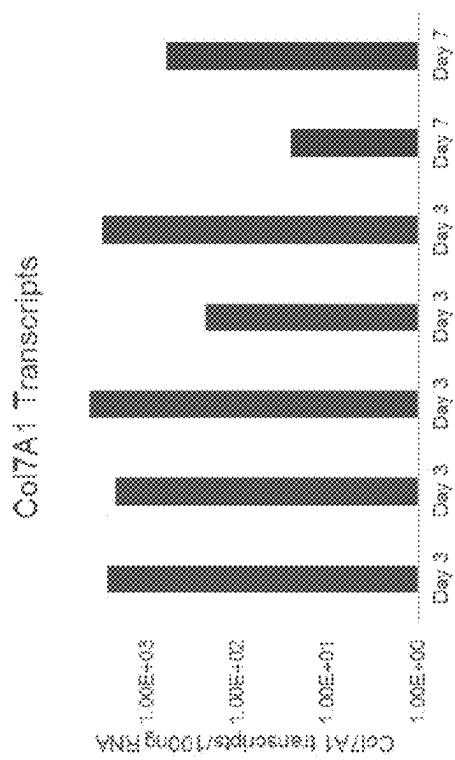
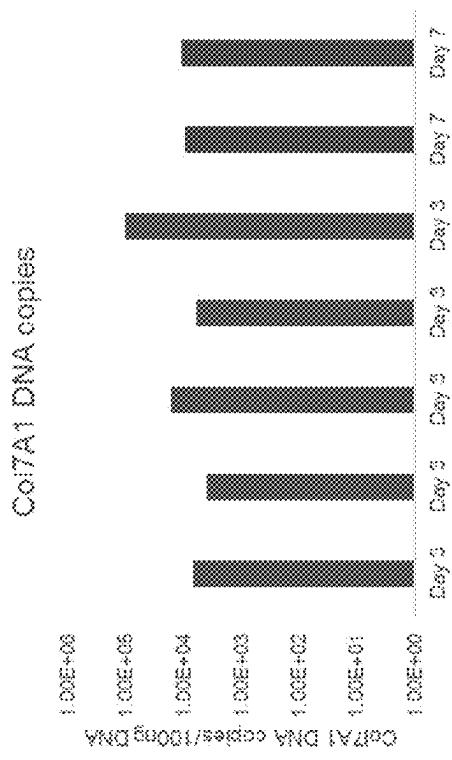


FIG. 15B



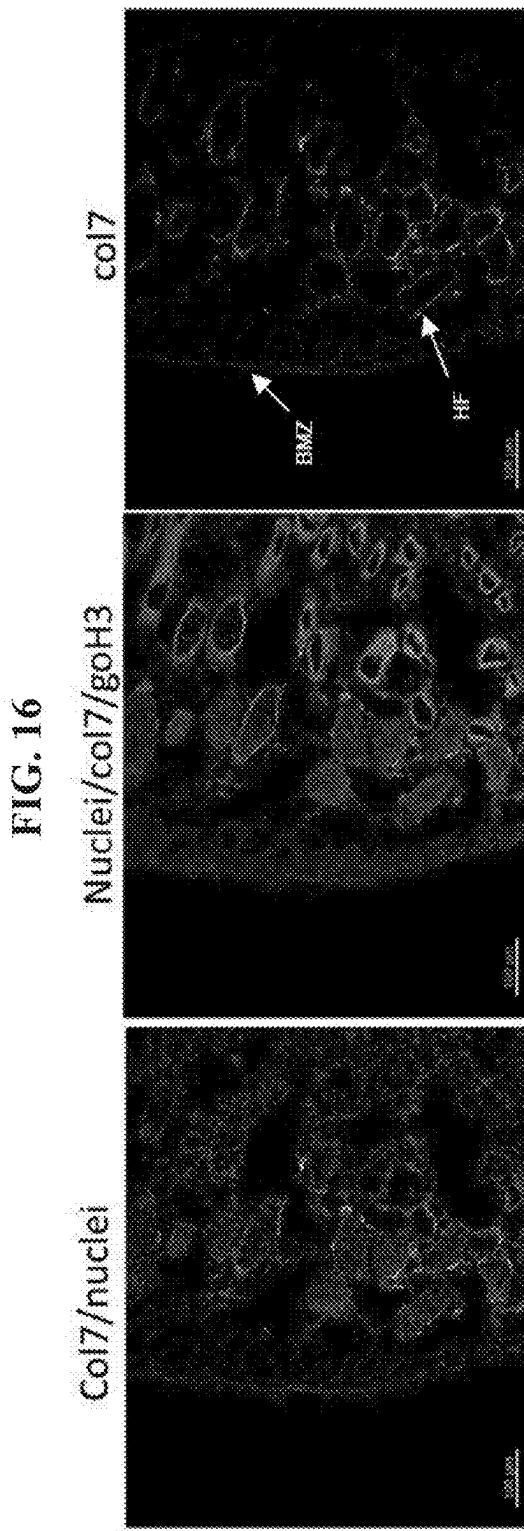


FIG. 17C

COL1A1 Transcripts

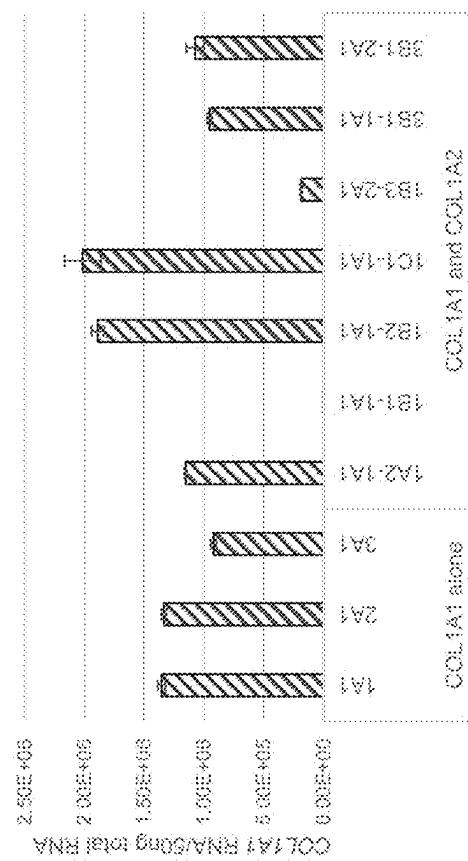


FIG. 17B

COL1A2 Transcripts

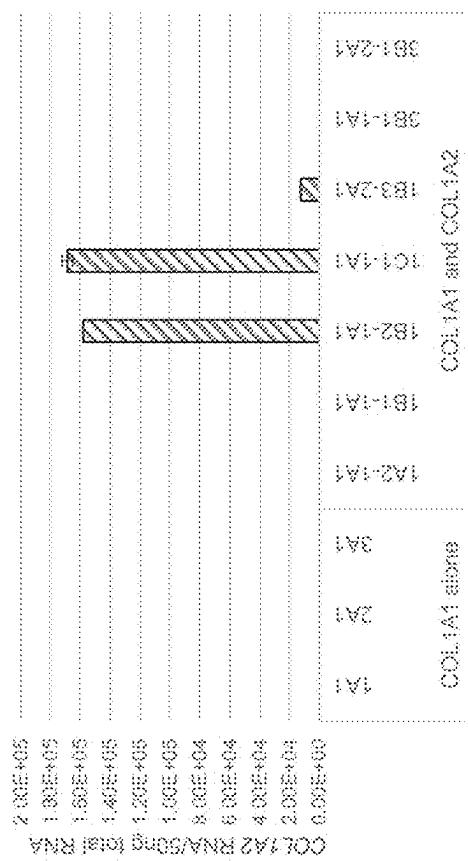
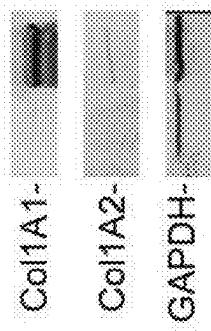


FIG. 18

No insertion



IRES-isolate 6

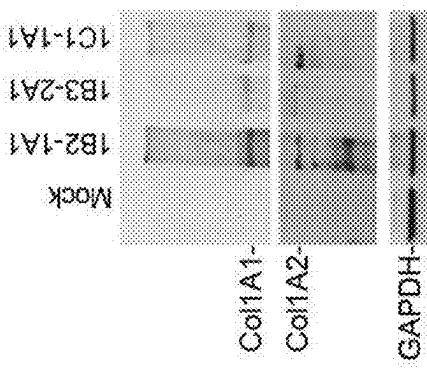


FIG. 19A

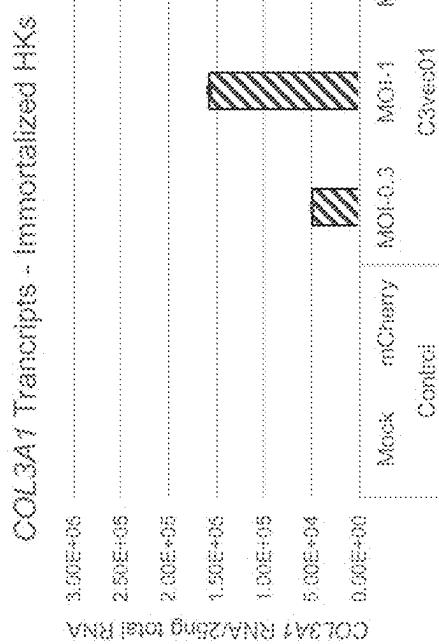


FIG. 19B

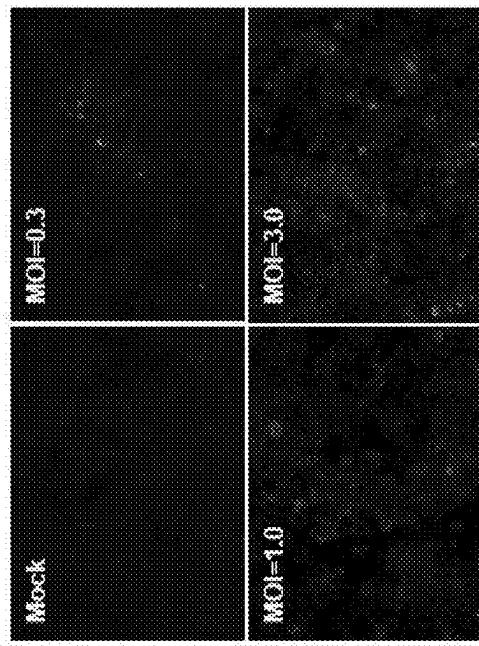


FIG. 20A

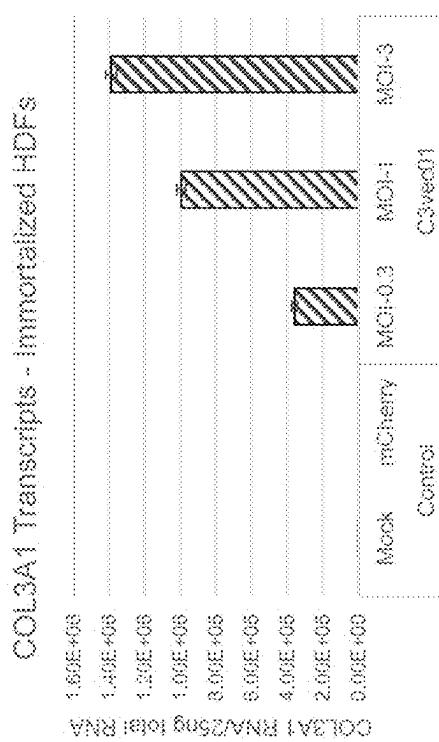


FIG. 20B

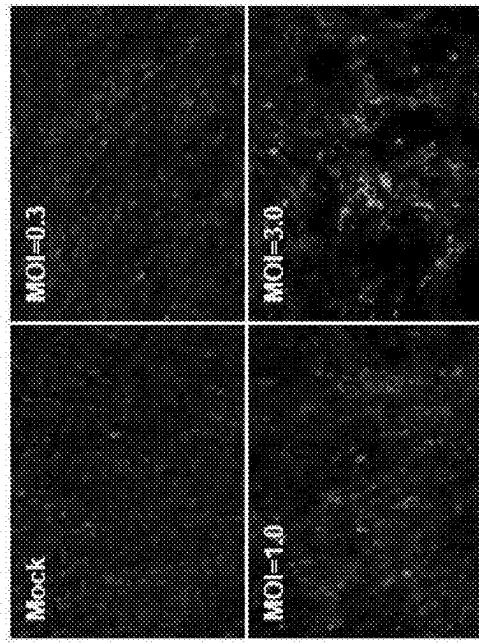


FIG. 21A

COL3A1 Transcripts - Primary HDFs

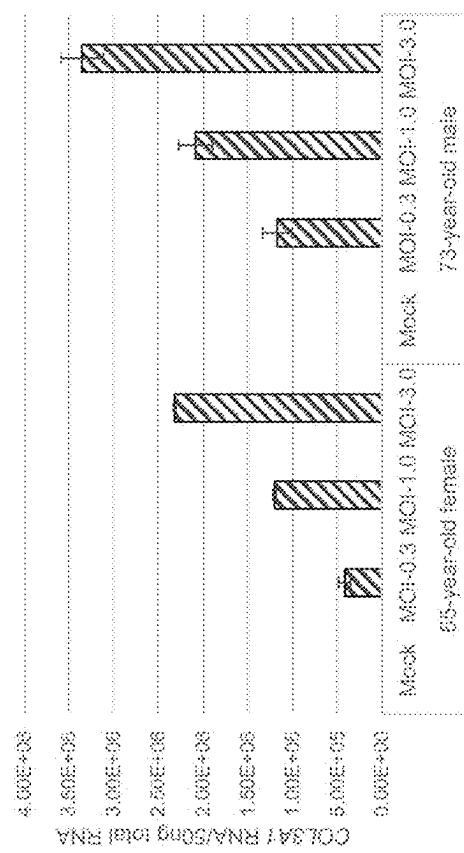


FIG. 21B

COL3A1 Transcripts - Primary HDFs

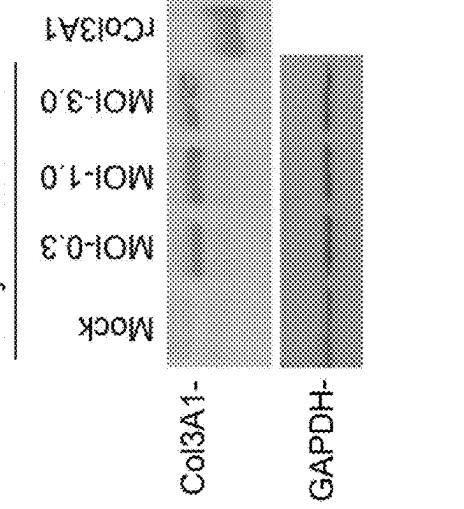


FIG. 21C

COL3A1 Transcripts - Primary HDFs

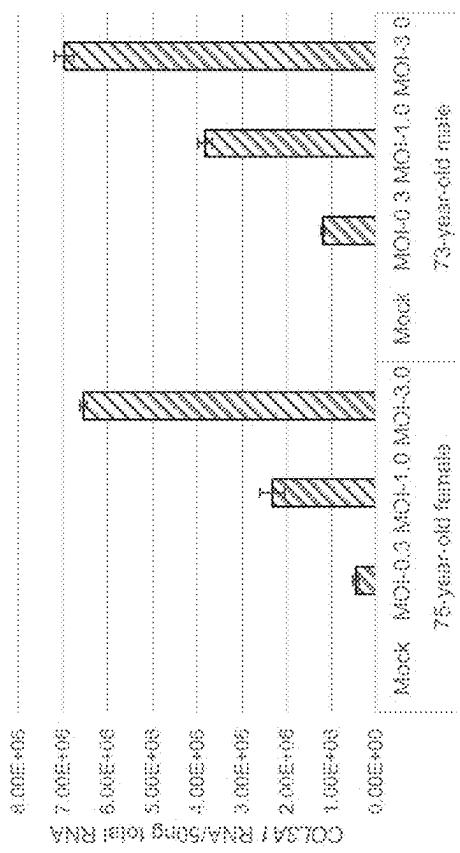


FIG. 21D

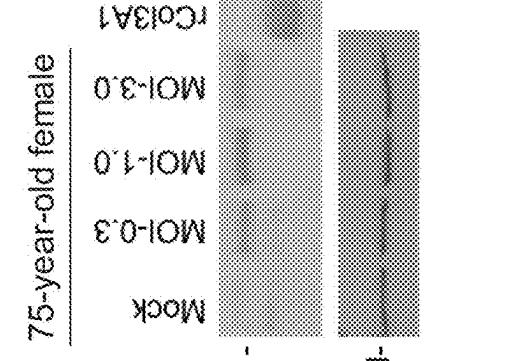


FIG. 22A
Secreted Human COL3

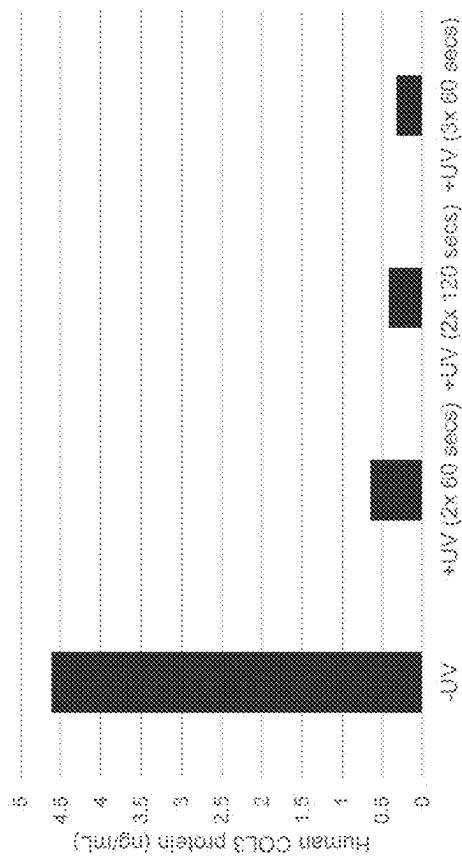


FIG. 22B
COL3A1 Transcripts - UV-damaged HUFs



FIG. 23A

Human COL3A1 DNA

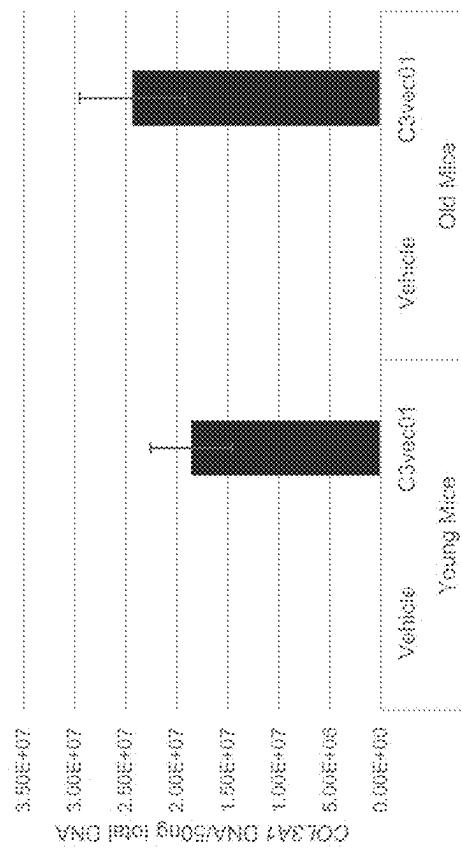


FIG. 23B

Human COL3A1 Transcripts

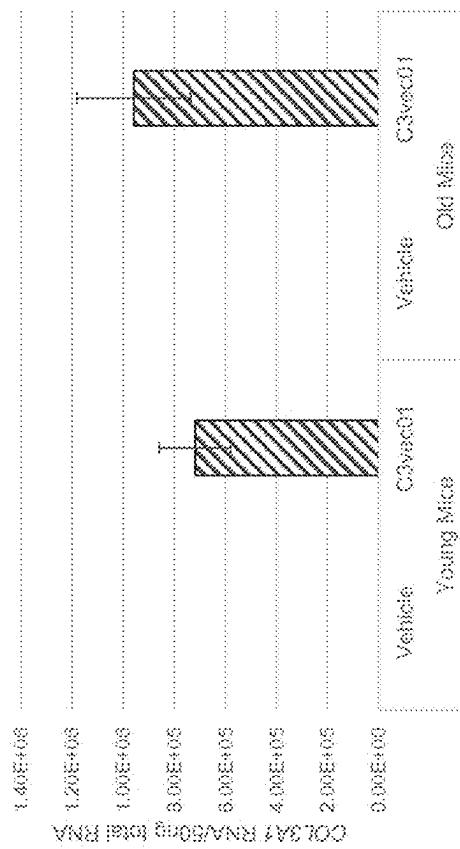


FIG. 23C

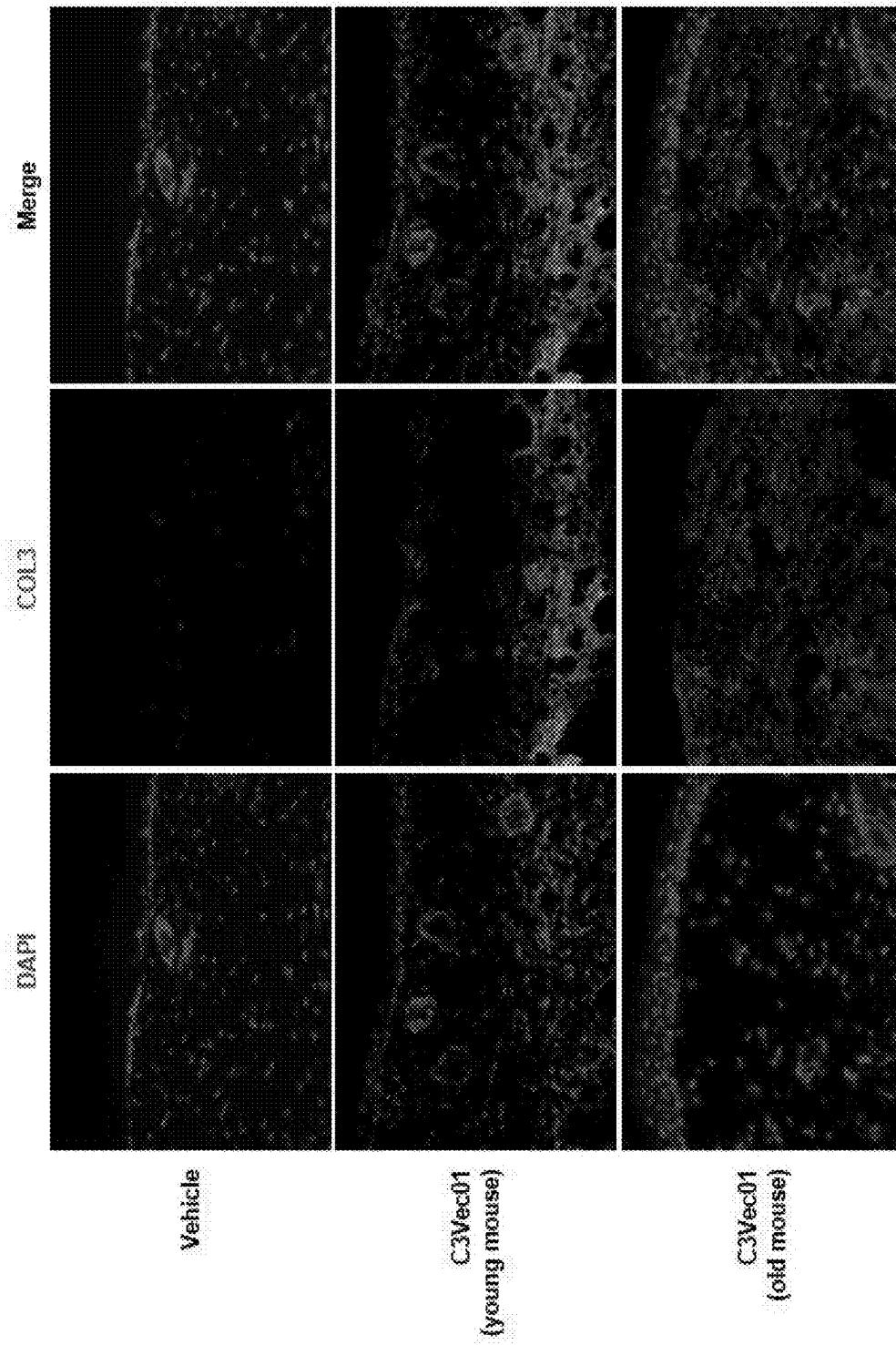


FIG. 24A

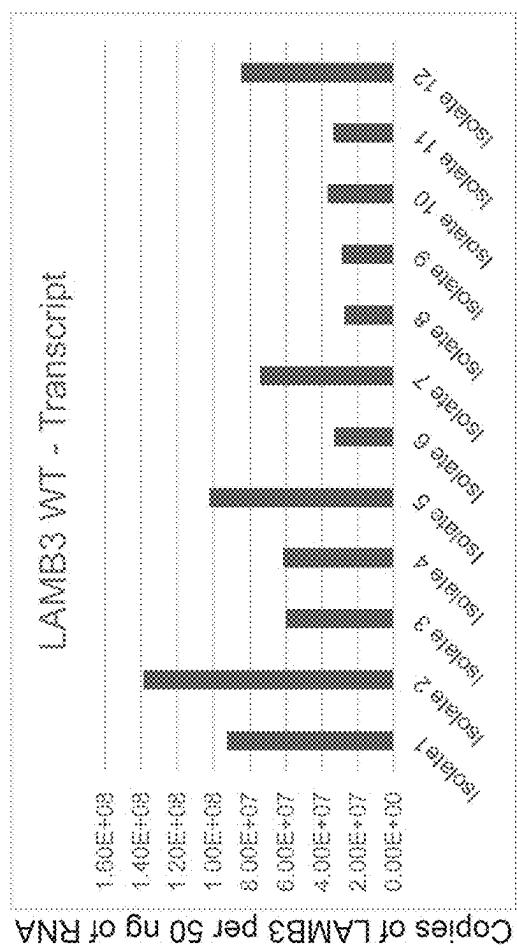


FIG. 24B

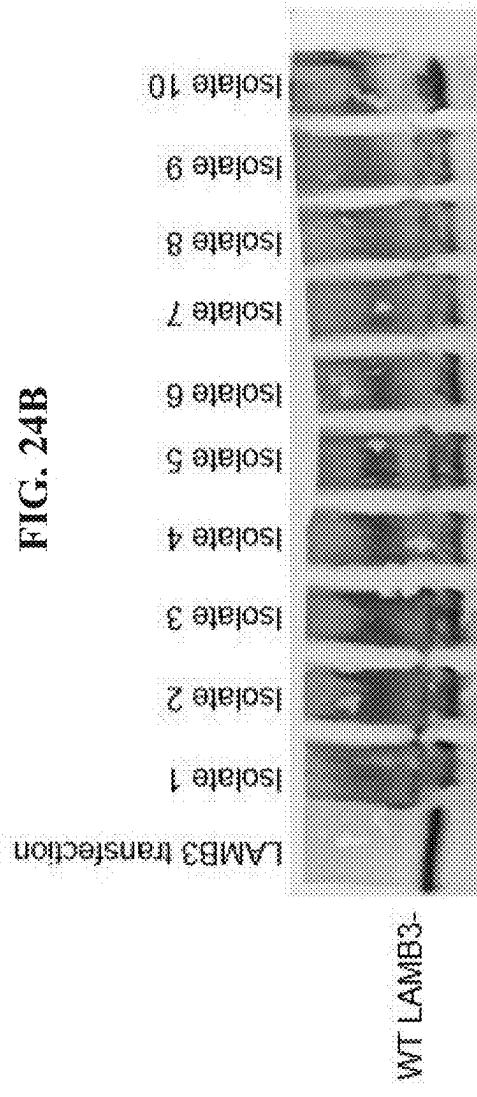


FIG. 25

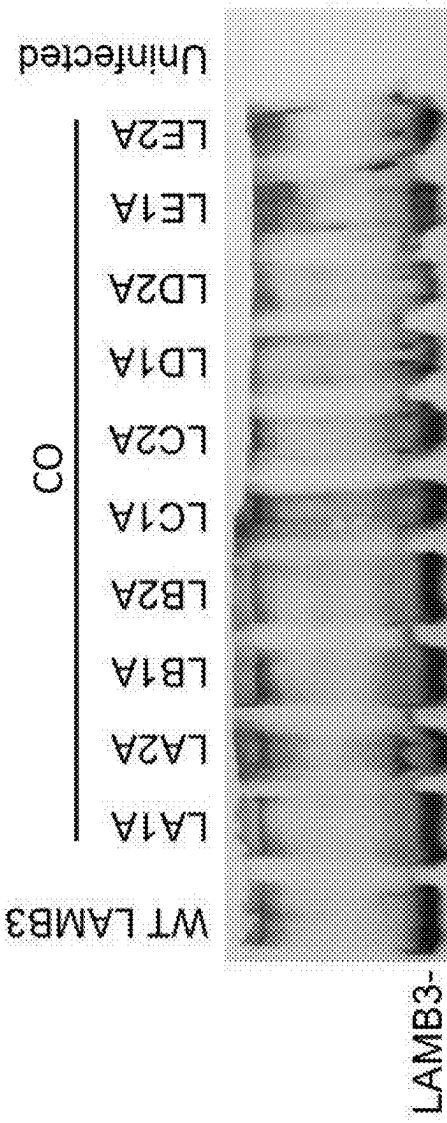


FIG. 26

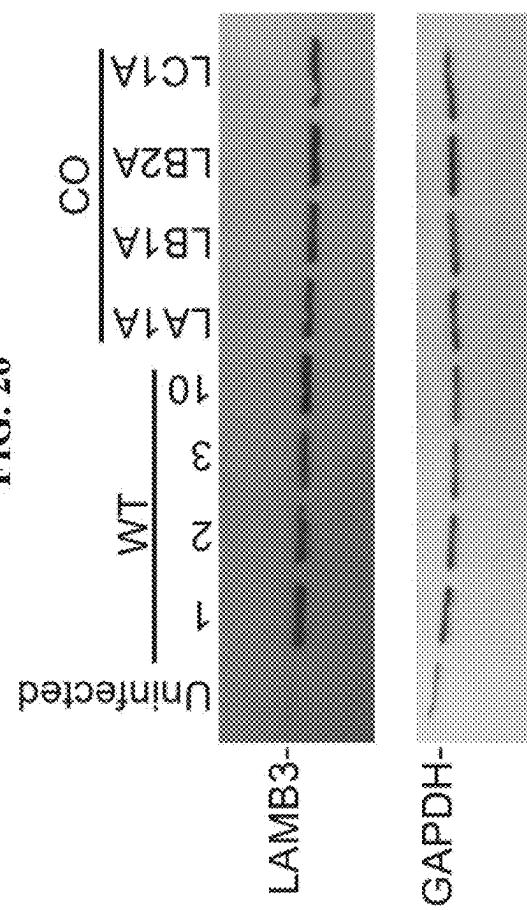


FIG. 27A

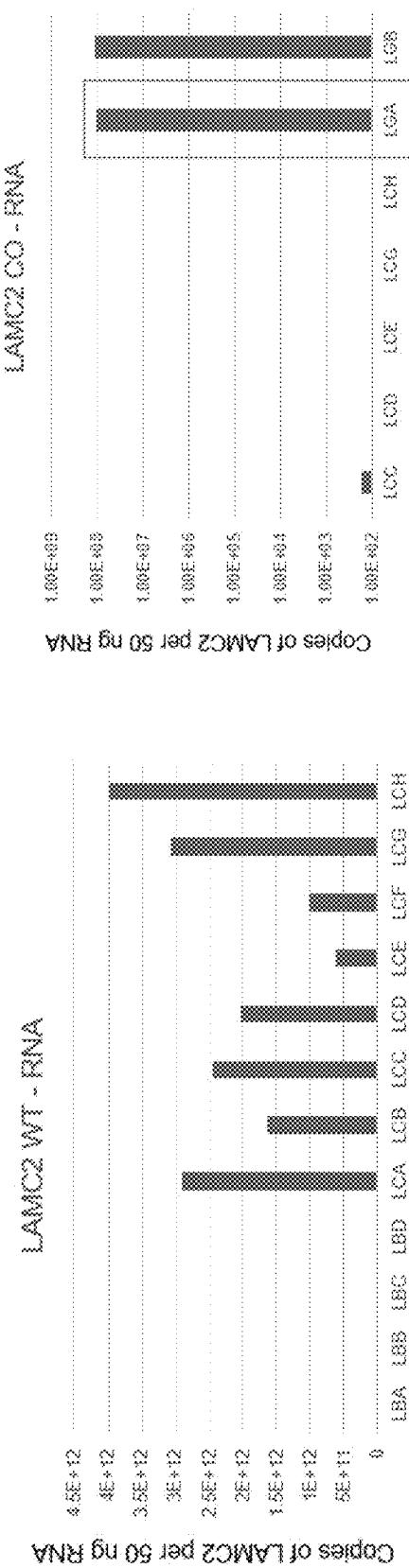


FIG. 27B

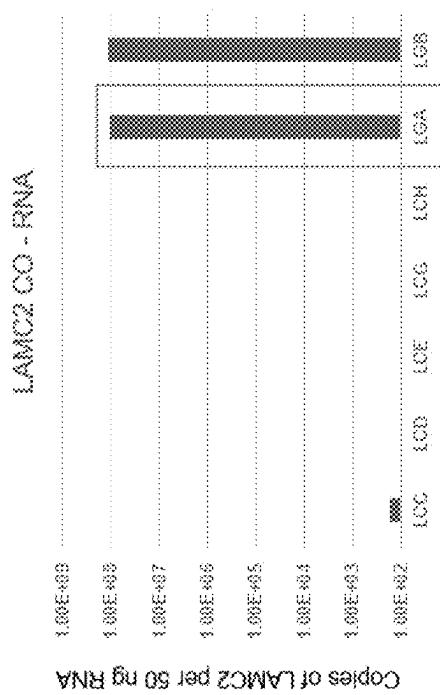


FIG. 27C



FIG. 28A

LAMC2 DNA

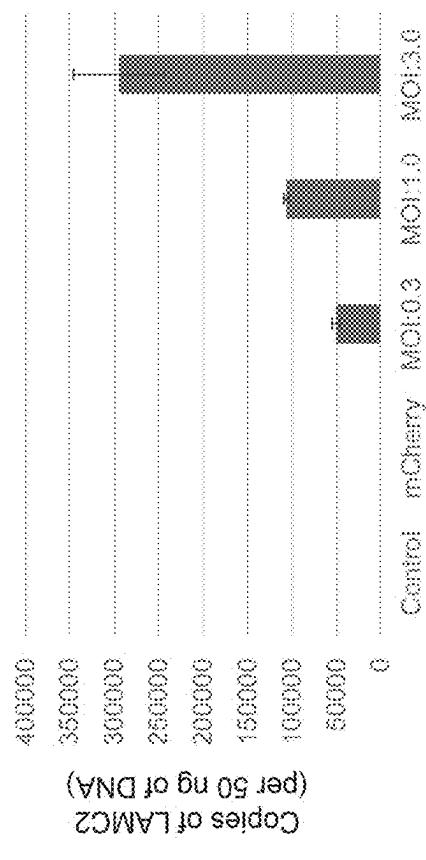
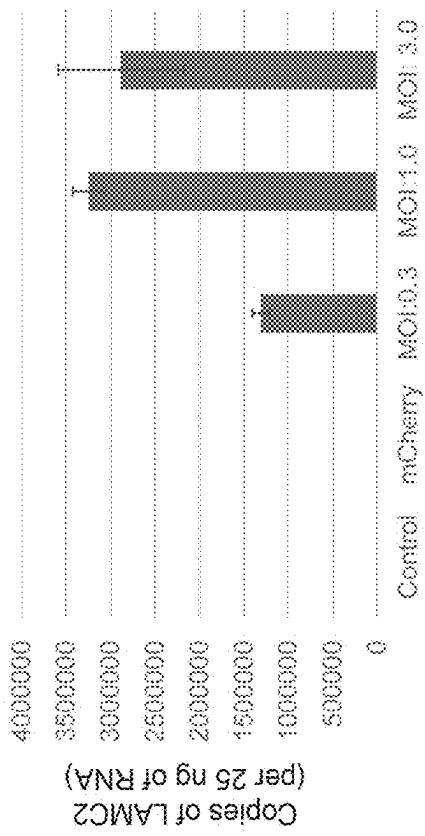


FIG. 28B

LAMC2 transcript



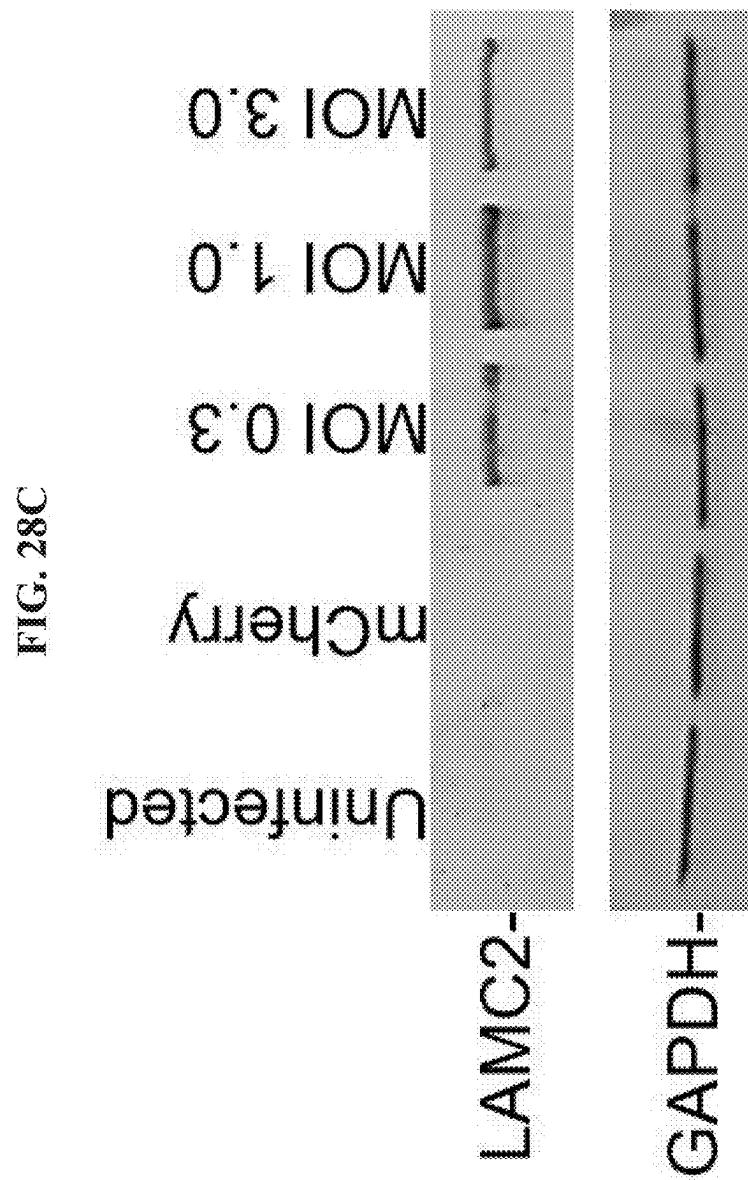


FIG. 29A
Mouse #1

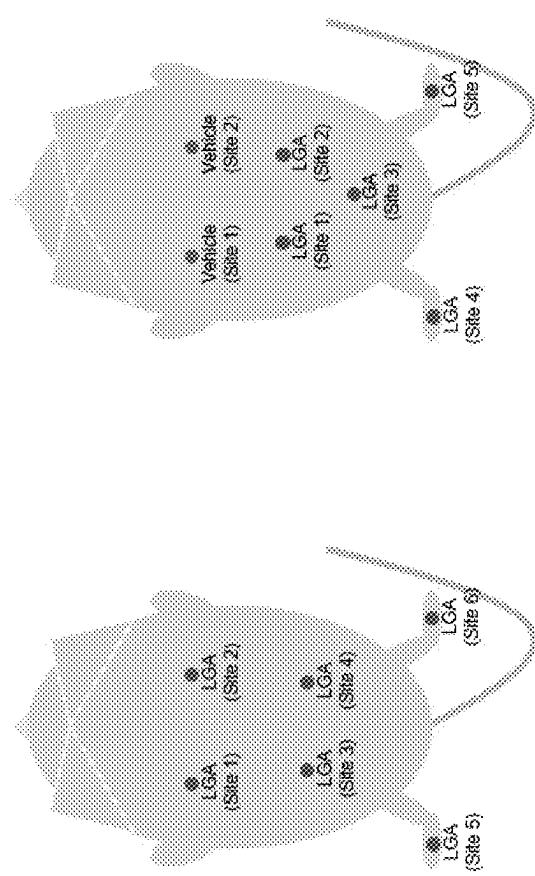


FIG. 29B

LAMC2 DNA - Dorsal Skin

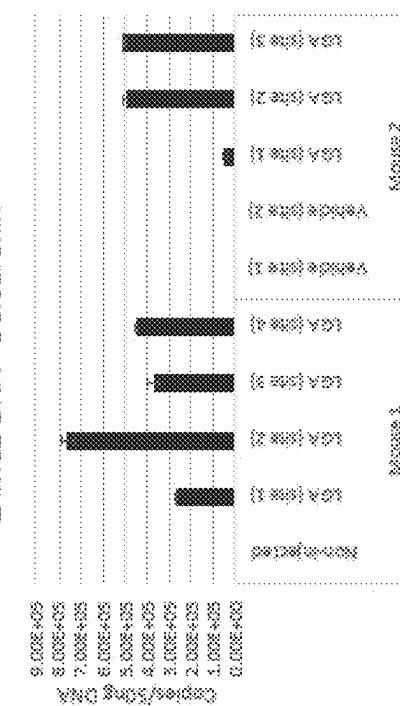


FIG. 29A
Mouse #2

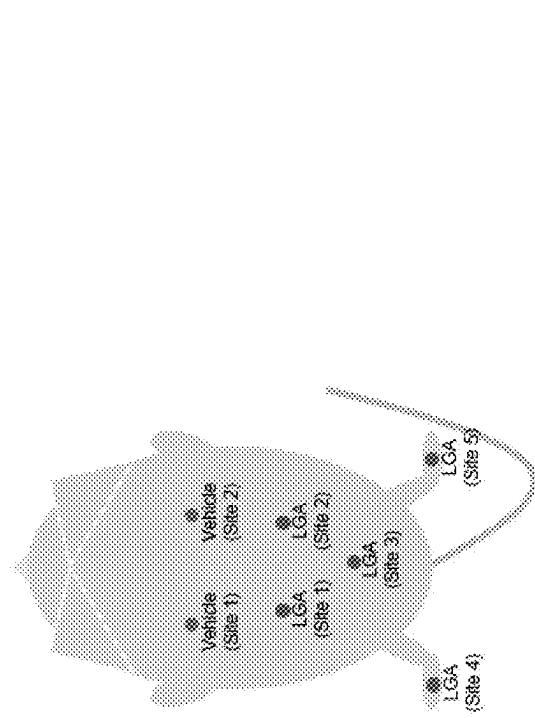
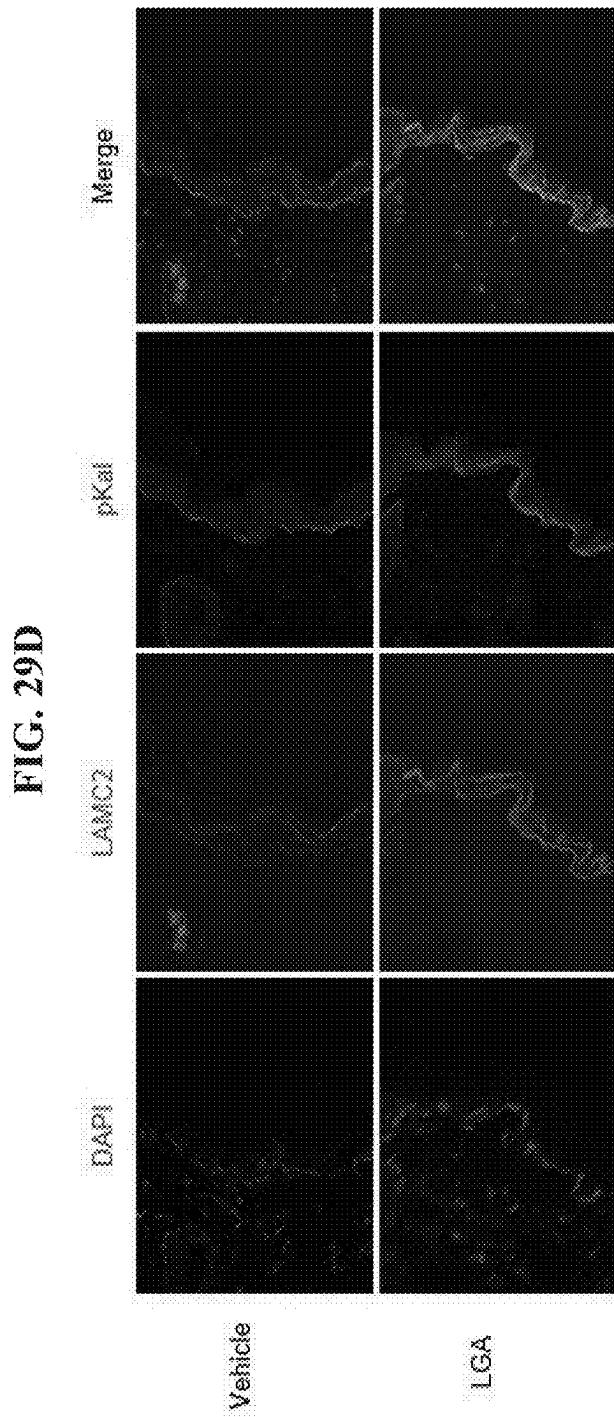


FIG. 29C

LAMC2 RNA - Dorsal Skin





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

-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 ggcttagata	ccacgaccgg gatgtgtggaa agcccgagcc ttgcagaatc 180
tgcgtgtgc acaatggcaa	ggtgcgtgtc gacgacgtca tctgcgacga gacaaagaat 240
tgcctgtggc cggaaagtgc	tgaggcgaa tggtgtctgt tggtgcctga tggcagcggag 300
agccccacac atcaagagac	aacaggcgtq gaaggccccca agggcgatac aggacctaga 360
ggtcctagag gacgtggcgg	accttctggc agagatggaa ttctggaca gcctggactg 420
cccgacccat ctggacccct	aggggcttca ggttcacccg gactcgagg aaattttgcc 480
ccacagctga gtcacggcta	cgacggaaa agcacaggcg gcatctgtt gcctggaccc 540
atggacccctt ctggcccaag	aggactcttctt gtgccttccgg acctcaggga 600
tttcaaggac caccaggcga	acctggcga ccaggcgctt gtggtcaat gggaccaaga 660
ggcccttcgtt ggcacccagg	caaaaatggc gacgatggcg aagccggaaa gcctggaaagg 720
cctggcggaa gaggccggcc	aggggccgca ggcgcgttagg ggttgcctgg aactgcacgg 780
ctgcctggca tgaaggggcca	cgagggtttt tctggactgg atggcgctaa gggcgacgct 840
ggaccacgac gacctaagg	cgacgcttgg tctctggcg agaardggc acctggacag 900
atgggtccca gaggattgcc	aggcgagaga ggttagaccc ggcgtccggg accagccgg 960
gtcagagggaa atgtggcgc	aacagggtgt gtggggcctc ctggaccaac cggacccgt 1020
ggccacccctt ggtttccagg	cgctgttggc gcaaaaggccg aaggccggcc acaaggaccc 1080
agggggatctg aagggtctca	ggggcggttta gggcgagccg ggccacccctgg gcctggccgt 1140
gcagctggac ctgctggaaa	ccctgggtgt gatggacacg cagggtccca aggtgtcta 1200
ggggccggcc gaaatggccg	cgctggccggc caaggaggcc atctggaccc 1260
caaggcccaag goggacccctt	gggttcttaag ggaataggcg gagagccagg cgcttctggg 1320
agtaaaaggcg atactggcgc	aaaaggcgaa cccggacccctt tggggagtca aggacccctt 1380
ggaccacgctg gcgaggaaagg	caaaagaggo gctagggggg aaccaggacc aacagggttc 1440
cctggccaccc ctggcgacgg	cgggacccctt ggtatcgatg gattccctgg cgcaatggc 1500
gtggccggac caaaaaggacc	tgcggcggaa agggggatcac cagggtctgc aggcccta 1560
ggttctccag gogaggctgg	cagacccggd gaagctggc tcccagggtc taaggactg 1620
acaggctcac eaggatctcc	cgggaccagac gaaaaaacag gaccccttccagg accggcagg 1680
caaggatgttta gacccggatc	tcctggcggcc cctgggtgtca gaggacaacg ttggcgtgtat 1740
gggtttcttg gacccggaaa	ggatggccgg gaaacccggg aacggggcga gagggggat 1800
cccgacccctt eagggtctgt	tggaccccttccgg gggggatccg gggggatccg 1860
cctccaggccg eacccggacc	agccggcggag agaggcgaac aaggaccagg cggatctcca 1920
ggatttcagg gactgtccagg	gcctgttggc cccgcctggc aggccaggaa gccaggcgaa 1980
cagggtgttc tggcgatct	tggggcccttggc ggttcttagcg gagettagagg cgaaggaggaa 2040
tttcctggccg aaaggggccgt	tcagggttca ccggggccagg ctggacccaa ggggtgc 2100
ggtgccccag gcaatgacgg	tgcctaaaggc gacgcaggcc ccccgagggtc tcctggatct 2160
caaggcgcac ctggacttca	ggggaaatggc ggcgaacccgg gactgtctgg acttcccggt 2220
ccaaaaggcg ataggggaga	tgtgttctt aaggggcgttctt atggcttcc tggaaaggat 2280
ggcgtcagag gctgtacagg	cccaattggc cctccgggac tgcgtggcgc tccaggccgt 2340
aaggggcaat ctggacccat	tggacccctt ggttcttacag gtgttagggg agcccaaggc 2400
gaccgggggg accctgtgtcc	accaggacccctt gtcggatggt ctggacccccc ttggcgctgt 2460
ggtcaacctg tgctaaaggg	cgagccggg gacgctgttgc caaaggcga cgctggtcca 2520
cctggacccgg ccggacccct	tggggccca qgaccttattt gaaatgttgg tgcccttgc 2580
gccaaaggcg caagaggatc	tgcggccca ccaggcgctt caggattccc aggtggccgt 2640
ggaaggatgtt gaccacccgg	gccaatgttgc aatgtgttgc caccggggacc gccaggacca 2700
gccccggcaag aagggtggaaa	aggcccttgg ggcggaaactg gecctgcagg cggcccaaggc 2760
gaagtggggcc ctccaggacc	tccggggcc gccggcggaa aaggatccccc aggcggcagat 2820
ggaccacccgg cgctcccccgg	aaacccaggatcc ccaacaggaa ttgtgttgc aaggggagat 2880
gtcggtctgc caggacacgg	ggggagagaga gttttccatg gactccctgg gccaaggccg 2940
gaacactggca aacaggaccc	aatgtgttgc acggggggg gaggccacc accggaccaatg 3000
ggtcctccatg gatggcagg	ggcttcttgc gaaatcttgc gagaagggtc tccaggccgc 3060
gaggggatctc ctggacgttgc	tgggttcttgc gggcccaagg ggcgtatgggg cgaaaaggcc 3120
ccagctggac ctccaggccgc	accggccgtt ccaggccgcac caggaccccttgc tggcccttgc 3180
ggaaaatctt gogacaggag	cgaaaacttgc cccggcggacc caggccggacc ttgtggaccc 3240
gtgggtgttca gaggacccgc	tggaccacaa ggttcttagcg ggcggaaagg cgaaacaggc 3300
gagcaaggccg acaggacccat	caaggccat gacggatcc ggcggacttca gggaccacca 3360
ggggccgtctg aaaggcccttgg	cgagoaggga ccaaggccggat ctgtgttgc cggccggacc 3420
agaggaccac ctgggttctgc	tgggtgcaccc gggaaaggccg gactgtatgg gctcccccgg 3480
cctattggcc cacctggacc	tagggaaaga acaggccgtt caggccggacttggaccat 3540
ggggccacccctg gaccggcccttgg	tcctcccttgc cccggatccgc ctccatgttgc 3600
ctgcctcaggc ctccctcaaga	gaaggccat gacggccggca gatattacag agccggaccc 3660
gccaacgtcg tgcgggacag	agatctggaa gtggacccca cactgttgc cctgtcttgc 3720
cagatcgaga acatcagaag	ccccggggc agcggaaaga accctgtccag aacctgtcgg 3780
gacctgaaga tggccacag	cgattggaaat tctggcgactt actggatccg ccccaaccat 3840
ggctgtcaacc tggatgttgc	caagggttgc tgcacatgg aaacccggca gacatgtcgt 3900
taccccccacac agccatctgt	ggctcagaagaa aactgttaca tcagcaagaa ccccaaggac 3960

-continued

aaggccgcacg	tttgggttcgg	cgagagcatg	accgtatggct	tccagtttga	gtatggccgc	4020
caggggcttcgt	acccttcggcc	tgttgcattc	cactgtggct	tcttcggcggt	gatgttcata	4080
gaggcccgcc	agaacatcac	cttccacttcg	aagaacacgg	tgccctatcat	ggatcagcag	4140
acccggcaacc	tgaagaaggc	actgtgtcgtt	caggggcagca	acaggatcgta	gatcagagcc	4200
gagggcacaaca	gcccgttcc	ctacagcggt	acagtggatg	gctgcaccag	ccatacaggc	4260
gctttggggca	agaccgtgtt	cgagtttttttt	accaccaaga	ccacggact	ccccatccat	4320
gtatggccca	ctctggatgt	tttttttttttt	gatcaagat	tccggcttcga	tgtttttttttt	4380
gttgtttttcc	tgttaa					4395

SEQ ID NO: 3 moltype = DNA length = 4101
 FEATURE Location/Qualifiers
 source 1..4101
 mol_type = unassigned DNA
 organism = Homo sapiens
 SEQUENCE: 3
 atgctcagat ttgtggatac gggacttgc ttgtcgcttg cagtaacccat atgccttagca 60
 acatggcaat ctttacaaga ggaaactgtt agaaaaggcc cagccggaga tagaggacca 120
 cggtggaaaa ggggtccacc agggccccca ggcagatgt gtgaagatgg tcccacaggc 180
 cttccctggtc cacctggccc tcctggcccc cctggctctcg gtggaaacct tgctgctcg 240
 tatgtggaaa aaggagtgg acttggccct ggaccaatgg cttaatgg acctagaggc 300
 ccacatgggt cagctggggc cccaggccct caaggttcc aaggacatgc ttgtggccct 360
 ggtgaactct gtcaaaactt tcctgggggt gctcggtgtc cagctggccc tccggcaag 420
 gctgggtaaat atggtcaccc tggaaaccc ggacgacccgt gtgagagagg agtttgttgg 480
 ccacagggtt ctcgtgggtt ccctggaaact cctggggctt ctggggccaa aggccattgg 540
 ggacaaatgtt ctgtggatgg attgtggaaa cagccgggtt ctctgggtt ggaaagggtgaa 600
 ctgggtggcc tcgtggaaaaa tggaaacttcc ggtcaaacat gagcccggtt gttttctgg 660
 gagagaggac gtgttgtgtc ccctggggca gctgggtccc gtggcgtgtt tggaaagtgtg 720
 ggttccctgg tgcctgtgtt tcccttggg tctgtgtccc ctccagggtt cccagggtcc 780
 cctggggccca agggttggaaat tggagctttt gtaacgtgt gtctgtgtt tccggccgg 840
 cccctgggtt aagtgggtt tccggccctt tccggggccctt tggggccctt tggtaatcc 900
 ggagcaaaac gccttacttgg tgccaagggt gctgtggcc ttccgggtt tgggggggtt 960
 cccggccctcc tcggggcccg cggttattcc tggcccttgg tgcgtggccgg tgtaactgtt 1020
 gcccaggaggat ttgttgtgtt gtcctgggtt gctgggttccaaaggagatgg cggtaacaagg 1080
 ggttgccggcc gctctgtgtt gcccggccat ggttccctgg tggcccttgg tccggccgg 1140
 agggccctta atggggaaat tggatgtcc ggccctccat gacccttccgg tggagggatgg 1200
 agtcctgggtt tcgtgggtt tcctgggggt gatggcagag ctgggtgtcat ggcccttcc 1260
 ggtatgtgtt gtgcgaatggg ccctgtgtt gtcggggatcc ctaatggaa tgcgtggccgg 1320
 cttttggggcc tcgtgttccat ggccggccat ggttccctgg tggcccttgg aaatatccgg 1380
 cccctggaa aagaagggtt tgcggccctt cctggccatcc accggccggcc tggcccaatt 1440
 gggccagctg gagaaggagg agagccgttcc aacattggat tccctggacc caaggcccc 1500
 actgggtatcc tcggcaaaaaa cggtgtataaa ggtcatgtt gtcgtggccgg tgcgtgggg 1560
 getccagggtt ctgtgttccat caatgggttcc cggggacttcc ctggaccata ggggttcc 1620
 ggtggaaaatg gtgaacagggtt tccccctggg tccctgggtt cccagggttcc gtcggccccc 1680
 tcagggtcccg ctgggttccat tggcaaaaaa ggagaaagggtt gtcctcatgg tgagttgtt 1740
 ctccctgggtt ctgtgtgtcc aagggggggat cccgggtccccc cagggtggatgg tgggtgtcc 1800
 ggttacttgcgtt gtccttattgg aaggccgggtt ctttctggatcc ccccaagggtt tgatggaaa 1860
 aagggttggac tcgtgtgtt gttgtgtgtt ggcactgtt gtcctatgg ttcattgttgg 1920
 ctcccaggagat aggggggtgc tgctggcata cctggggatcc agggggaaaaa ggggttggacc 1980
 ggtctcaaggat gtggaaattttt taaccctggc agatgtgttcc tgcgtgtgtcc tccctgtgtt 2040
 gttaggtgtcc tcgtgttccat tggagccaca ggttgcgggg cggatgtgtt ggttgcgtcc 2100
 ctctgtgttcc tcgtgtgttcc tcggggggat cctgtgttccat ggttgcgtgtt cggccctgtt 2160
 gggcccaatgtt gatttgctgg tcctgtgtt gtcgtgttcc aacccgggttcc taaaaggaaaa 2220
 agaggaggccaa aaggggcttaa ggggttggaaaat ggtgttggatcc gtcctacagg ccccttgg 2280
 gtcgtgtccat caggctgtgttcc aatgtgttcc cccgggttccat tcggggatgg tgggtatgtt 2340
 gggcccccctgtt gtatgtacttgg ttccctgtgtt gtcgtgttccat gggactgttcc cccaggacc 2400
 tcgtgttatttc tcggcccttcc tcggcccccgtt gtcgtgttccat gggactgttcc cccaggacc 2460
 ctcgtgtgtt accaagggttcc agttggccgtt actggggaaat taggtgttccat tgggtccctt 2520
 ggcttccgtt gttggaaagggtt tccttctggatcc ggggttggatcc tgcgtggccat tccctggcc 2580
 ccagggttccat aggggttccat tcgggttccat ggttattttgg ttccttccat tccggaggtt 2640
 gaaacgtgttcc taccgggttcc tcgtgtgttcc tcggggatcc tccctgttccat tgggttccat 2700
 gggcccttccat gggccgggttcc tccttctggatcc gtcgtgttccat gtcctggatcc caacccgttcc 2760
 ctcgtgtgttcc gttggatccat gggactgttccat gtcctggatcc tccctggatcc tccctggatcc 2820
 caacccggccat acaaggggatcc ggcgggttccat cctggccataa ttggggccctt tgggtgtcc 2880
 ggttgcacccat tccttccatgg cccctgggttccat tcgggttccat tggatgttcc tccctggatcc 2940
 atgggttccat tcgggttccatgg ttccttccat ggttgcgtgttccat gggactgttccat tccctggatcc 3000
 ccacaaggccat tcgtgtgttcc taaggggatcc cccgggttccat gggactgttccat tccctggatcc 3060
 ggcttaaaggat gacacaatggtccat gggactgttccat tcgtgtgttccat tccatggatcc 3120
 caagggtgttccat tcgggttccat gggactgttccat tcgtgtgttccat tccatggatcc 3180
 ctcgtgtgttccat gggactgttccat tcgtgtgttccat tcgtgtgttccat tccatggatcc 3240
 aagatgtgttccat cactggatccat tcgtgtgttccat tcgtgtgttccat tccatggatcc 3300
 gggcccttccat gtcctggatccat tccttccat gggactgttccat tcgtgtgttccat tccatggatcc 3360
 ccagggtgttccat gggactgttccat tcgtgtgttccat tcgtgtgttccat tccatggatcc 3420
 caggcctgttccat caggccatccat tccttccat gggactgttccat tcgtgtgttccat tccatggatcc 3480
 ttccttccat gggactgttccat tcgtgtgttccat tcgtgtgttccat tccatggatcc 3540
 cgcacatgttccat gttggatgttccat tccttccat gggactgttccat tcgtgtgttccat tccatggatcc 3600
 gaaacctgttccat aagggtgttccat tccttccat gggactgttccat tcgtgtgttccat tccatggatcc 3660

-continued

aaggacaaga	aacacgtctg	gcttaggagaa	actatcaatg	ctggcagcca	gtttaaatat	3720
aatgtagaag	gagtgacttc	caaggaaatg	gctacccaac	ttgccttcat	ggccctgtcg	3780
gccaacttgc	ccttcattggaa	catcacctac	cactcgaaa	acagcattgc	atacatggat	3840
gaggagactg	gcaacactgaa	aaagggtgtc	attctacagg	gtcttataatg	tgttgaaactt	3900
gttgctgagg	gcaacacgca	gttcaattac	actgttcttg	tagatggctg	ctctaaaaag	3960
acaatgaat	ggggggaaagac	ataattatggaa	tacaaaacaa	ataaaggccatc	acggccgtggcc	4020
ttcccttgcata	ttgcacccctt	ggacatcggt	ggtgtctgacc	aggaatttctt	tgtggacatt	4080
ggccccatct	gtttcaataa	a				4101

-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	tggcttcat	gagactgtg	3780
gccaattacg	ccagccagaa	catcaccta	caactgcaaga	acagcattgc	ctatcatggac	3840
gaggaaaaccg	gcaacctgaa	gaaagccgtg	atccgtcagg	gctctaa	cgtggaaactg	3900
gtggcccgagg	gcaacagcg	attcaccta	accgtgtctg	tggacccgtg	cagcaaaaaag	3960
accaacgagt	ggggcaagac	catcatcgag	tataagacc	acaagcccg	cagactgccc	4020
ttctctggata	tcgccccact	ggatattgga	ggcgcgacc	aagagtctt	tgtggacatc	4080
ggcccccgtgt	gcttc	aatgt	ta	4101		
SEQ ID NO: 5		moltype = DNA	length = 4401			
FEATURE		Location/Qualifiers				
source		1..4401				
		mol_type = unassigned DNA				
		organism = Homo sapiens				
SEQUENCE: 5						
atgtatggact	tttgcaaaa	ggggagctgg	ctacttctcg	ctctgttca	tcccactatt	60
attttgcac	aacaggaa	tggttgaagg	ggatgttccc	atcttggtca	gtcttatgcg	120
gatagagatg	tctggaa	gcca	agaaccatgc	caaata	atgtg	180
ctctgcgatg	acataatatg	tgacgat	caa	gaatttagact	gccccaa	240
tttggagaat	gttgc	caact	gttgc	acttgc	cccttctta	300
ggtcaaggac	ctcaagg	ccagg	ggat	ccggcc	cttgc	360
ggtgaccctg	gtat	ccagg	ggat	ccggcc	cccttgc	420
tgtgaatcat	gccc	acttgc	ttc	tccat	tttgc	480
aagtctggag	tagc	acttgc	ttc	tccat	tttgc	540
cctcccggt	ccct	gttgc	atc	tgc	tttgc	600
ggacccccc	gtgaa	cttgc	tttgc	tttgc	tttgc	660
ataggccat	ctgg	tttgc	tttgc	tttgc	tttgc	720
gagcggaggat	tgcc	tttgc	tttgc	tttgc	tttgc	780
gttatggaa	gac	tttgc	tttgc	tttgc	tttgc	840
cctggattaa	agggtgaa	tttgc	tttgc	tttgc	tttgc	900
ccaa	gggg	tttgc	tttgc	tttgc	tttgc	960
ggttaatgac	gtg	tttgc	tttgc	tttgc	tttgc	1020
gcccggattcc	ctgg	tttgc	tttgc	tttgc	tttgc	1080
tcaa	atgtgt	tttgc	tttgc	tttgc	tttgc	1140
ggtctcttgc	gccc	tttgc	tttgc	tttgc	tttgc	1200
gctggcattc	ctgg	tttgc	tttgc	tttgc	tttgc	1260
gtaatgttgc	ctcc	tttgc	tttgc	tttgc	tttgc	1320
ggagagccg	gacc	tttgc	tttgc	tttgc	tttgc	1380
aaaggcga	atgg	tttgc	tttgc	tttgc	tttgc	1440
gctgcaggag	aaagg	tttgc	tttgc	tttgc	tttgc	1500
ggagaaaa	gtcc	tttgc	tttgc	tttgc	tttgc	1560
gctgggaaac	ctgg	tttgc	tttgc	tttgc	tttgc	1620
agtccaggag	gacc	tttgc	tttgc	tttgc	tttgc	1680
ggtcgaccag	gtcc	tttgc	tttgc	tttgc	tttgc	1740
cccggttca	aaagg	tttgc	tttgc	tttgc	tttgc	1800
ggacttggcc	ctcagg	tttgc	tttgc	tttgc	tttgc	1860
gggcctactg	ggc	tttgc	tttgc	tttgc	tttgc	1920
caaggcgtc	ctgg	tttgc	tttgc	tttgc	tttgc	1980
caaaagggtt	atgg	tttgc	tttgc	tttgc	tttgc	2040
ggtgaacgtg	gacc	tttgc	tttgc	tttgc	tttgc	2100
cctggtcc	aaagg	tttgc	tttgc	tttgc	tttgc	2160
actcctggc	tgca	tttgc	tttgc	tttgc	tttgc	2220
ggtgacaagg	gtg	tttgc	tttgc	tttgc	tttgc	2280
aggggttca	ctgg	tttgc	tttgc	tttgc	tttgc	2340
gaaagggtgt	ccc	tttgc	tttgc	tttgc	tttgc	2400
ggtgaaactg	gccc	tttgc	tttgc	tttgc	tttgc	2460
cctggtgtt	aaagg	tttgc	tttgc	tttgc	tttgc	2520
gttgacggcc	ccc	tttgc	tttgc	tttgc	tttgc	2580
ggtgacgtg	gacc	tttgc	tttgc	tttgc	tttgc	2640
cctggtctc	ctgg	tttgc	tttgc	tttgc	tttgc	2700
aaggatggc	cccc	tttgc	tttgc	tttgc	tttgc	2760
ggacc	aaagg	tttgc	tttgc	tttgc	tttgc	2820
ccaggagtc	cagg	tttgc	tttgc	tttgc	tttgc	2880
ccaccaggca	tgcc	tttgc	tttgc	tttgc	tttgc	2940
ggaaaaccag	gag	tttgc	tttgc	tttgc	tttgc	3000
cctggtctg	ctgg	tttgc	tttgc	tttgc	tttgc	3060
cttccaggcc	gag	tttgc	tttgc	tttgc	tttgc	3120
gggtccctg	gctg	tttgc	tttgc	tttgc	tttgc	3180
agtggtaca	gagg	tttgc	tttgc	tttgc	tttgc	3240
tcctggatgt	ctcc	tttgc	tttgc	tttgc	tttgc	3300
ggagctgtgt	gcat	tttgc	tttgc	tttgc	tttgc	3360
ccaggccctg	ctgg	tttgc	tttgc	tttgc	tttgc	3420
cctgttggac	ccag	tttgc	tttgc	tttgc	tttgc	3480
ggaccaccag	ggcc	tttgc	tttgc	tttgc	tttgc	3540

-continued

```

caggggcaac caggccctcc tggaeccctc ggtgecccgt gtccttgcgt tggtgggttt 3600
ggagccgctg ccattgctgg gattggaggt gaaaaagctg gcggtttgc cccgtattat 3660
ggagatggac caatggatt caaaaatcaac accgtatgaga ttatgacttc actaaagtct 3720
gttaatggac aaatagaagaat cctcatttagt cctgtatgggtt ctgcgtaaaaa ccccgctaga 3780
aactcgagag acctgaaatt ctggccatcc gaaactcaaga gtggggaaaata ctgggttgac 3840
cctaaccaaag gatgcaaaatt ggatgtatc aagggtttctt gtatatatggaa aactggggaa 3900
acatgcataa gtgccaatcc tttgaatgtt ccacggaaac actgggtggac agattctagt 3960
gtctggaaaga aacacgttgc gtttggagag tccatgtatgtt gtgggtttca gtttagtac 4020
ggcaacttcgt aacttccatgt aagatgtccctt gatgtcgacg tggccatcttc tgcacttctc 4080
ttccaggcgag ctccggagaa catcacatcatgt cttccggagaa atagatgtcc atatcgat 4140
caggccgtg gaaatgtaaa gaaggccctg aaggtgtatgg ggtcaatgt aaggtaattc 4200
aaggctgaag gaaatagcaa attcacctac acagttctgg aggatgggtt cacgaaacac 4260
actggggatgg gtagccaaac agtcttggaa tatcgaaacac gcaaggctgt gagactacct 4320
atgttagata ttgcacccctt tgacatgggtt ggtcctgtatc aagaatttgg tggacgttt 4380
ggcccttctttt qctttttata 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
 atgatgcgact tcgtgcagaa aggccagctgg ctgctgtcg cactgctgc ccctacaatc 60
 attctggccc agcaagaggc cgtggaaaggc ggatgttctc acctgggaca gagetacgc 120
 gagacggatg tggtaaaaggc tgagcccttg cagatctcg tggtagatag cggcagctg 180
 ctgtgcgacg acatcatctg cgacgaccaa gagctggact gccccaaatcc tgagatccc 240
 ttccggcgagt gctgtggccgt ttgtccccaa cctccataccg ctctccatccg accttcataat 300
 ggacaggaggc ctccaggggc taaaaggcgat cctggacccct ctggaatccc cggcagaaaat 360
 ggcgcattccg gcatcttggc acagcttgc tctcttggaa tggccaggcc acctggcattc 420
 tgcgagact gtcctacagg ccctcagaac tacagccccac agtacgacagc ctacgacgtg 480
 aagtctggc tggccgttgg aggactggct ggttattccag gacctgctgg accaccaggg 540
 cctccaggcg cgccttgcgaa acatggacat ccaggatctc cccggggactt cgggtttatcag 600
 ggaccacccgt gogaaccttgg acaggccagga ccttagtggc ctctccgttc accagggtgcc 660
 atttggccatct ctggacccgtc cggggaaaggat ggcaaatctg gcacccaggd cagaacttgc 720
 gagagaggat tgcctgggtcc tccgggtatt aaggccccag ccggaaattcc aggattcccc 780
 ggaatggatgg gccacacaggat ctgtcgatggc cggatggccg agaaaggccgaa aacagggtcc 840
 ctggactgt aaggccgaaaaa tggactggc cggcggaaatg ggcggccccc accatgtggaa 900
 ccttagaggtt ctccaggcgc aagaggaaagg ccaggacttc ctgggtgtc agggcctaga 960
 ggcaatgtatg ggcggcaggagg atctgtatggc caacctggc cgcggaggacc tccaggcaca 1020
 gtcgttccctt cggcgatctcc tggtgcaaaag ggcaatggg gaccaggctg aaggccctggaa 1080
 tctaattgttgc ccccgaggaca gaggggggaa ccaggaccac aaggacatgc tgggtctca 1140
 ggcctccggg gtccctccagg gatcaatggc tctccaggccg cggccggcga tatggccct 1200
 gctggaaattc ctggcgctcc aggttattatg ggagccagag ggccacctgg accaggcaggc 1260
 gcaaatggcc ctccctgact tagaggccggc gcaggccggc ctggggaaaaaa tggccaaaaaa 1320
 ggcgagccgg gaccaaggggg agaaaggccg gaaatgttgc ttccagggtt ccctgggtcc 1380
 aaaggccgggg acggccaaaggaa tggtagtgcct ggccggccgg ccggccatgg actccctggc 1440
 gcaatggccg aaaggggggac acctgggttt agaggaccgg ccggaccta tggatttcc 1500
 ggcggaaaaaaat gtccaggccgg cgaggcggtt gtcctccggc ctggccggccc aaggagggt 1560
 gtcggcgccggc caggccgggg tgggtgtccca ggccggaccgg ggtatggagg catggccaggc 1620
 tcacccggccg gacccgttgc tgatggaaaaa cccggggctc ctgggtatggc gggccaaatcc 1680
 ggttagacccgg gacccctggg accatccggg ccaaggaggac aacccggcgt gatgggttcc 1740
 cctggaccta aggggaaatga cggcgctccccc ggaaaggaaacg ccggccggccgg tggccctggc 1800
 ggttccgggtt ctcacccggggcc acaggccggaa aacccggccggaa ccggcccccaggccaggatc 1860
 ggacccatccacg gacccctgggg agatccggc gatccatggc caccaggacc tcaggccgggt 1920
 caaggactgc ctggaaacttgg cggaccccttcc gggccggccatg gaaaaccagg ccggccaggaa 1980
 cctaaggccggc acgtcggtgc acctggccgc ccaggccggaa agggccggccggc agggcgttcc 2040
 ggcggaggaggc gacccctccagg atttgggttgc gtcctccagg tggaggccgg agtccgttgc 2100
 ccaggacttgc aagggtggaaa aagggttgcggc ggaccccttgc ggcggccctgg cgtgttgcgg 2160
 actcccaaggac ttcaagggtt gcctggccaa cgagggtggc ttggagcccc aggacccaaaa 2220
 ggcgataagg ggcgaaacctgg cggacccgggt gcagatgggtt ttccggggaaa agatggacca 2280
 cggggcccccggc caggacttat aggccttcca gggccggccggc gacccggggccggccaaaggcc 2340
 gaagggtggcc ccccttgggttgc gctggaaatttgc ttccgggttca gagggttacc tggccggccgg 2400
 ggagaaaaacgg ccccttggggcc acggccggggaa ttcccccggc ctctccggccca aatggccgg 2460
 cctggggggaa aaggccggaaag aggtgcacccg ggcggaaaaag ggcggccgggg acctccctgg 2520
 gttgtcgaccc ctccctgggg atctggggccca gctggggctc ctgggttctca aggtgtttaag 2580
 ggcggaaaaagg gcttcctccagg cggacccgggt gctgtggat ttccggccgc aaggaggatgg 2640
 cccggccaccgg caggcttccatggcataccca ggtcccttgc gacccatggc ctctccctggc 2700
 aaaaatggccg caccggccggc acggccggaaatg actgggttgc ctggatcacc tggccgttgc 2760
 ggaccggaaaaa ggcgacccggc acaaccaggc gaaaaggat ctctccggc tcaaggccct 2820
 cctggccggc ac ctggccattt gggatttgc ggaattacag gtgcggagg cctgggttgc 2880
 ccacccctggaa tgcctggggcc aaggaggtagc cccggggactt aaggccgttaccc aaggccatcc 2940
 gggaaacccggc ggcggccaaacgg actggccggaa gaaaggaggac ctccgggttcc acaaggccct 3000
 ctggatgggg tggggccggc tggggccggc aaggccgttaccc gcaatccctgg ctctccgttgc 3060
 ttggccggggaa gggatggccccc cccagggttgg aaggccggatc gccggggaaaaa ttggtagccca 3120

-continued

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

-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
tctcccgga	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	acaagggtc	aaagggtgag	gtgggttcc	caggattag	cggggccca	3600
gaaatctct	gatccaaagg	agagcaagg	tccatgggt	ctccggggcc	ccagggacag	3660
ccggggttac	cgggatccc	aggccatggc	acggggggc	ceaaaggaga	ccggggact	3720
caaggccagc	ctggcttgc	aggacttcc	ggacccatgg	ggcccttcc	gtttcttgg	3780
attgtggag	ttaaagggt	caaaggaaat	ccaggctgg	caggagcacc	cgggtgtccca	3840
gggccccaaagg	gagaccctgg	atccaggat	atgcctgtt	ttgggtggct	tccaggatc	3900
acaggctcta	agggtgat	ggggccttca	ggagttccag	gatttcaagg	tccaaaagg	3960
cttcctggcc	tccaggat	taaaagggt	caaggcgttca	aaggcgtccc	ggggagctaaa	4020
ggtctcccg	gtccctctgg	ccccccagg	ccttacgaca	tcatcaaagg	ggagccccgg	4080
ctccctggc	ctgaggggcc	ccggggctg	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	gtggccctgt	4680
cacagccaga	ccatccat	ccccatgt	ccacgggggt	gttcctgt	gtggatccgc	4740
taacttcttgc	tgatgcac	cagegctgt	gcagaaggct	ctggcca	cctggcttgc	4800
cccggtcttgc	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 gtggctgtc ttcttcctg ccgtctgtc gtcacgcag 60
gaacatttca gagcccccgc taaaaggcgga tggccggat ctggatggg caagtgtat 120
tgtcacggc tgaaggggca gaaggcgaa agaggacttc ctggactgca gggcggtatc 180
ggatttctg gaatgcaagg acctgaggga ccacaggac ctccaggaca aaaaggcgat 240
acaggcgaac ctggcttgc tggcacaagg ggaacaagg gaccacccgg cgcctcttgc 300
tatccctggaa atccaggact ggctggaaatc ccaggacagg atggacttcc ttgttcttct 360
ggcatcttcgt gtgttataatgg cacaaggcgg qagaggggtt cactcgacc accaggact 420
ccaggatttgc ttggaaaccc tggccacact ggattgtctg gcatgaagggg cgatccttgc 480
gaaatttctgg gacacgttcc cggcatgtcg ctgaaaggcg aaaggggtt tcccgaaat 540
cctggcacac ctggttccac aggettgcac ggacttcaag gaccgttgg cccaccttgc 600
tttacagggtc ctccaggatc acggggctt ccaggccgc caggcgaaaa aggacaatg 660
ggactgtatc ttcaaggggcc caaggcgac aaaggcgatc aagggtttag tggcccttca 720
ggtgttcttcg gacaggccca ggttcaaggaa aaggcgatc tgccacccaa gggcgagaag 780
ggacagaaaaag gcaacccagg ctttcagggg atgccttggc tggggagaaaa gggcgaaaccc 840
ggaaaacccgt gaccttgcgg aaaaaacccgc aaggatggcc ataaggccga gaaaggatct 900
cctggattcc ctggcgaaacc tggatatect ggactgtatc gtagacaggg ccctcaaggc 960
gagaaggccg aagctggacc gcctggacca cccggaaatg tgatttggaaac agggcccttc 1020
ggagaaaaaaag gogagccagg atccctggc actccctggc ctggggaga accaggaccc 1080
aaagggtttc cggactgtcc aggacaaacc gggcccttcgt ggttctctgt gcttggacaa 1140
gtctgggtc tggactgtcc cggccaaatgt ggcggaaaagg gcgatagagg ttttcttgc 1200
acaaggctgc ctggcccttc tggagagat ggacttccgg ggcacccagg ttctcccgga 1260
ccacctggac agcctggcta taccatggc atcgtcgatgt ctaaccctgg tccacccgtc 1320
gatcaaggac ctccctggat accaggccgg cctggcttta tcggcgagat tggagaaaa 1380
gggcggaaaagg gggagagctg cctgtatgc gacatcgatg gatacgagg cccttcttgc 1440
cctcaaggccc caccaggcga gatagtttt ccaggccac tcggccaaatggcggacccgt 1500
ggactccctg ttagagatgg ttttgcgttgc gttccaggac cacaaggccac ccctggactt 1560
attggacacgc cagggtcataa gggcgagcc ggcgagttt acttcgcact gagacttaaa 1620
ggcgacaaagg gcgaccctgg ctttcttgc caacctggaa tgcccggtt agctggatca 1680
ccaggccgtg atggacatcc agggttggcc gggacaaaaag gtcctccagg atctgtggc 1740
ctcaaaggccg aaagggccccc tccaggccgg gttggatttc ctggcttgc aggcgtatct 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	ttctctggc	tcctggtcc	aaaagggttt	2100
atggactgc	cggcgat	gggcccac	ggaacaccc	gcagaccc	ctttaatgg	2160
ttgc	ccggaa	atcccg	ccaaaggcc	aaaggcg	ccgttgttgg	2220
cttaaaggac	ttccaggc	gcccgcata	cctgggac	ctgggaaa	gggatctat	2280
ggagtctg	cgctgc	cgaacatgt	gcaattgg	cacctgg	gcaaggcatt	2340
agaggc	ac	ccggcctc	aggactcc	ggtctgtt	gaagtc	2400
atgggac	cac	gggtgttag	gggacact	ggcc	cgccccgg	2460
ttgggac	cc	ggggacac	gtc	ccac	attgttgg	2520
ccac	ctgg	tcaaaaggc	gaaaggct	ccccgtt	ccggccttga	2580
cctaaaggc	g	atagggtc	ccaggc	cctgg	tatgcctgg	2640
gggc	ctcc	gacaacagg	tgctc	ttcccgat	taaggcgaa	2700
atgggc	gtgg	tggatcacc	tgg	ggatcacc	gacctgttgg	2760
ttggc	ccgg	aaaaaggc	ccac	ccac	ggggacaa	2820
ggattg	aaag	gcgataagg	cgac	cttcc	aaccagg	2880
gtgg	cat	gttcatgg	ggg	ccgg	gatcg	2940
cctat	cgcc	aaaaagg	gggg	ccgg	ccggat	3000
ggaca	ac	gccaac	gcac	ccaa	ggcgtt	3060
ccagg	gtc	caggact	cg	ggat	tttgg	3120
acacc	ccgc	agaaagg	tc	cttcc	cttcc	3180
ggog	gacaa	agtg	gggg	ccgg	ccggat	3240
ggactt	agg	gagaga	gggg	ccgg	ccggat	3300
aaaggc	aaa	aaaggctat	cg	ccat	ccagg	3360
tcac	ctgg	gctgtgg	tcc	ggatc	ccat	3420
gtact	gtct	gtctgttgg	cata	cttcc	gggg	3480
cctgg	gacc	c	gggg	ccgg	ccggat	3540
tctg	ctgg	aaaaaggc	ggcc	ccgg	ccggat	3600
gcaa	aggcg	ataagg	taagg	gtcg	ccgg	3660
ggcat	cccg	cc	gggg	ccgg	ccggat	3720
ccagg	gttgc	ctgg	gggg	ccgg	ccggat	3780
ca	ggg	gac	gggg	ccgg	ccggat	3840
attg	atgg	tcaagg	caagg	ccagg	ccggat	3900
gttcc	aaaagg	gogat	gggg	ccgg	ccggat	3960
actgg	aa	gacat	gggg	ccgg	ccggat	4020
ttgg	cagg	gtaa	gggg	ccgg	ccggat	4080
ggatt	gtct	ggcc	acc	ccgg	ccggat	4140
ctgc	ctgg	ccgg	ccgg	ccgg	ccggat	4200
ggac	agc	gtgt	ccgg	ccgg	ccggat	4260
gatgg	gtc	ccgg	aggc	ccgg	ccggat	4320
ttccc	ccgg	cc	gggg	ccgg	ccggat	4380
ccctg	ggat	ccgg	ccgg	ccgg	ccggat	4440
cctag	ccgg	ccaa	gggg	ccgg	ccggat	4500
ccttt	ttgt	caac	gggg	ccgg	ccggat	4560
tact	gggt	ctac	gggg	ccgg	ccggat	4620
atc	gggg	atcg	gggg	ccgg	ccggat	4680
cact	ccgg	ccat	gggg	ccgg	ccggat	4740
tact	ccgg	ttgt	gggg	ccgg	ccggat	4800
cccg	gtct	gtgg	gggg	ccgg	ccggat	4860
acct	gtta	actac	gggg	ccgg	ccggat	4920
atgt	caag	gggg	ccgg	ccgg	ccggat	4980
agat	gtca	ttct	gggg	ccgg	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	gtctgtct	gtctgtcagg	ctgtgtggac	agccgcgcag	60
gatgagccgg	agac	cccg	ggccgtggcc	ttccaggact	gccccgtgga	120
gtctgtggaca	cct	ctgtcgag	cgtggccctg	aggctgtggac	cctacggggc	180
aaagtcaagt	c	ttcc	cacca	ggttttccatc	gacaacctgt	240
gaccggaaac	cc	ttgt	gggg	ggggacaggta	ctaccgtgt	300
caaggcctca	cg	cg	ccgg	ctggggccgc	gacgcactca	360
aagtactttg	gg	aa	gggg	gacgcactca	aaagcagcgt	420
ctctgtgggg	gt	cc	gggg	tgcgctatca	ggaggggt	480
ccctgtgggg	gt	cc	gggg	tgcgctatca	ggaggggt	540
caccttggcc	tc	aa	gggg	tgcgctatca	accatcgg	600
agcatcatcg	cc	ac	gggg	tgcgctatca	ccggggctga	660
agccgcgacg	c	ac	gggg	tgcgctatca	ctggggccag	720
ggagaagccg	g	ag	gggg	tgcgctatca	ccatcgaca	780
aataacgtgg	ag	ca	gggg	tgcgctatca	ccatcgaca	840
ctccggggcc	ac	cc	gggg	tgcgctatca	ccatcgaca	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	gcagttacac	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	tcttcctccc	ggctgttcat	accatcttc	tggacggctc	cgccagcggt	2520
ggcagccaca	attttgacac	caccaagcgc	ttcgccaaacg	gcctggccga	gcgttctcc	2580
acagcggggca	ggacggaccc	cgccccacgc	gtgcgggtgg	cggtgttgc	gtacagccgc	2640
acggggccacg	agcgcggcaga	gccccggctg	ctgcagttcc	tcgcagaaacta	cacggccctg	2700
gccagtgccg	tccatgttccat	ggacttatac	aacgacgcaca	ccgacgtcaa	cgatggctcg	2760
ggttatgttgc	cccggttctt	cccgaggccc	tggtccggcc	ctggcaagaa	gaggctgtcg	2820
ctcttcctcg	atggcaactc	gcaggggcc	acgcccgttgc	ccatcgagaa	ggccgtgcag	2880
gaagccccccgg	gggcaggcat	cgagatctt	tggtgtggcc	ttggccgc	ggtaatggag	2940
ccccacatcc	gggttcttgc	caccgcacgg	acggcccgat	acgacgtggc	ctacggccgg	3000
agccacacgtt	tccgtgtccc	cagctaccgg	ggccgttcc	gggtgttctt	ccaccagaca	3060
gttcccaaggg	aggtggcgct	gggtctag				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	gccccgttac	gacaacctgc	gccccggta	ctacagatgc	240
gacagaaaaac tctgtgtggaa	ccgcggcc	ctgcactact	ctgacgagggt	ggaaatcatc	300
caggggcttgc ccagaatggcc	tgccggaaaga	gatgccttgc	agttctagct	ggacggccgt	360
aagtactttt gcaaggggcac	cttacaccgc	tgccgcatac	aaagaaggctt	ggaaacagctg	420
ctcgtcgccg cgacccatct	gaaagagaac	aagtacctga	tcgtgtttac	cgacggacac	480
cctctggaaat gtatcaaaa	acccctgttgc	ggacttggaa	atgcctgtaa	cgaggccaaa	540
caccttggccg tgaagggttt	cagegtggcc	atcacaccccg	accacctgaa	acctcggtcg	600
agcatcatcg ccacccgacca	caccatccgg	cggttattca	cagccgttgc	ttggggccag	660
agcagggatg ccgaagaggc	catcagccag	accatcgaca	ccatctgttgc	catgtatcaag	720
aacaacgtgg aacaatgttgc	ctgcagttcc	gagtggccac	ctgtctagg	acccctcttgc	780
cttagaggccg accctggctt	tgaggccgg	agagggaaaac	ctggacttgc	ccggcgaaaa	840
ggcggaaatgttgcgatcc	tagacccgttgc	gatctggggac	ctgtggggctt	tcagggtat	900
aaggggcgaga aaggctcccg	ggggagagaag	ggatcttagag	gcccataagg	ctacaagggc	960
aaaaaaaggca agagggggcat	cgatggcggt	gacggcgctt	aaaggcgagat	gggatatactt	1020
ggactccctgtgtttaagg	cggccctgttgc	ttcgatgttgc	tccaggggacc	tccaggacc	1080
aaggggcgatc caggccgcctt	tggactgaaa	ggcgaaaaagg	gcaaggacc	tgccgtatggc	1140
gaaggcggca gacctgttgc	ttctggccctt	acggcgatgt	aaggacacc	ttggcgaaat	1200
ggaccacccgt gcgagaagggg	cgaacccggc	gacgaggggaa	atccaggacc	agatgggtct	1260
cctggcgaaa gaggttggacc	aggcgaaagg	ggaccttagag	gaacaccagg	cacaaggaggc	1320
ccaaaggccgc atccggccg	ggctggaccc	caaggcgatc	agggaagaga	aggaccatgt	1380
ggagggttctg ggcacccagg	cgaaggccg	cctatcgcc	ctaaaggata	tagaggccac	1440
gaaggcccttc ctggatctgc	agggtctaga	ggccgttcc	gtcccgagg	ccctccaggc	1500
gaccggccgc ttatgggaga	acggggagaa	gatggcccttgc	ccggcaatgg	cacagaggc	1560
tttccaggcttccgttccgt	ccccggaaat	agaggcgctc	ctggatcaa	ccggccacca	1620
ggatataccatggcttccgt	ggctcaaagg	cgacgaaagg	gaggccaggc	atccaggggta	1680
gatatacgcccttccgttccgt	gaaaggccgc	aaaggctata	gaggaccaga	gggaccacaa	1740
ggcccaatctgttccatcaagg	gtcatcaagg	ggccaccagg	cctgcacgtt	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	ccctcgat	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
cccatcac	tttagcagccc	tgcgcacatc	accatctgc	ttgatggc	tcgtacgtgt	2520
ggcagccaca	acttcgacac	caccaagaga	ttcgccaage	ggctggccga	gagatttctg	2580
acagccggca	gaacccatgt	tgctcagat	gtgcgagtgg	ccgtgttcca	gtatcttgc	2640
acaggccagc	aaagacccga	gagacgttct	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
cctcacatc	gagtgctgtt	cacccgcaag	accgcggatg	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

atgacgcgtgc	ggcttcttgtt	ggccgcgtcc	tgcgcgggaa	ttctggcaga	ggcgccccga	60
gtgcgagccc	agcacacggg	gagagtgacc	tgcacgcggc	tttacgcgc	tgacatttgt	120
ttttaactgg	atggcttcc	atccattggc	cgacgaaat	tccgcgagg	ccgcagctt	180
ctcgaaggcc	tttgtgtgtc	tttctcttga	gcagccatg	cacagggtgt	gctttgc	240
acagtgcagt	acagcgatg	cccacggaca	gagttcgcc	tggatgcact	tggctcttgg	300
gttgatgtga	tccgcggat	ccgttgaggat	agttacaagg	ggggcaaac	tgcacacagg	360
gctgcaattc	tccatgtggc	tgaccatgtc	ttcttgcccc	agctggcc	acctgggtgc	420
cccaaggctt	gatcttgc	cacagacggg	aagtccagg	acctgggtga	cacagctgc	480
caaaggctg	aggggcagg	ggtcaagct	tttgtgtgg	ggatcaagaa	tgctgacc	540
gaggagctg	agcggatgt	ctcacaggcc	accagtact	tcttcttct	cgtaatgac	600
ttcagcatct	tgaggacact	actgccttc	tgcgcgggaa	gagtgtgcac	gactgtgtt	660
ggcgtgcctg	tgacccgacc	tccggatgac	tccaccttgc	ctccacgaga	cctgggtgt	720
tctgagccaa	cgagccaaatc	tttggagatg	cagtggacag	cgccca	ggccatgt	780
ggctacaa	tccatgtac	tcctgcgacg	gggctgggac	agccactg	gggtgagg	840
caggaggctg	atgttccac	ttgtggagacc	atgtgtcgcc	tgcgggg	ccggccact	900
accggatacc	aagtgtact	gattgcctc	tacgccaaca	gatcgggg	ggctgtgac	960
gggacagctc	ggaccactc	ctttagaagg	ccggaaatg	ccatccagaa	taccacagcc	1020
cacagctcc	ttgtgtggc	ccggatgttg	ccaggatgtt	ccgttgcac	tgtgacatgg	1080
cggttctca	ttgtgtggcc	cacacagcg	caggatgtt	cccttgggca	gggttca	1140
ttgtgtgtgt	acttggatc	tggacggac	tatggatgt	ccgttgcac	cctattttgc	1200
cgcagtgtgg	ggccgcac	tttctgtat	gctcgact	acgttctgt	tgagcagacc	1260
ctgcgcggcc	tcatecttgg	ccccacate	atcttctt	cttggaaactt	ggtgc	1320
gcccgtgtt	accgggttga	atggggcg	gagactgtt	tggacccacc	gcagaagg	1380
gtactgcct	ctgtatgtac	ccgttacca	tttggatggc	tgcacccgg	cactgagat	1440
cgcctcac	tcttacact	gttggagg	cacggatgtt	ccacccctgc	aaccgtgtt	1500
cccaactggc	caagatgtcc	tgtgtggcc	gttacagacc	tgcacccac	cgagctgccc	1560
gggcgcgg	tgccgtgtc	ctggacccca	gtccctgg	ccacccat	ccgcacat	1620
gtgcgcacg	ccccagggtt	tgacccggacc	ctgggttcc	ctggggatca	gacagat	1680
gacttggat	acgttccagg	ttggatgtt	tacactgttgc	gggtgttgc	tcgagtgg	1740
cccccgtgg	caatgtccat	gttccgttgc	gtccgcggg	agccggaa	tccacttgt	1800
gttccatggc	ttgtgtgtt	gggttgc	gttacatgt	tgatgggtgc	ctggggacc	1860
gtccctgtgg	ccatgtgtt	tcggatgtt	tgaggacac	cgatgtgttcc	ggatccac	1920
cacacactc	ccccagactc	tactgcaca	gatcatac	ggctgcagcc	tggaaacc	1980
taccagggtt	ttgtgtcggt	actgcgagg	agagaggagg	ccctgttgc	agtcatgt	2040
gtctcgaccc	cccaactgg	cccaatgtt	acgttccat	tgcacccat	ccggccat	2100
tctgtca	ttatctggac	caggatgtt	ggcgcac	gatacagggt	ttctggc	2160
tcagccac	gcccacac	atcccaatgt	gttcttgg	aggccacgtt	ggctgtgt	2220
gatggactgg	agccagatac	tgtatgtac	gtgcgttgc	ggggccatgt	ggctgggt	2280
gatggggccc	ctgcctctgt	gggttgc	actgccttgc	agcgttgc	ttctgtgt	2340
aggctcgaga	tttcaatgc	ttccagc	tttcttgc	tttcttgc	tttcttgc	2400
ggagccac	tttacacact	ggcctgggg	ccggatgtt	ggggccat	ggggccat	2460
atactccac	gaaacacaga	ctctcgat	atccgggtt	tgcacccat	tttcttgc	2520
tcagtgtcg	tgactgtact	tttggatgtt	ccggacccgg	ccatgttgc	tattttgc	2580
actacgcgc	ctgaggctt	ggccaccc	gggacgc	tttcttgc	tttcttgc	2640
cactcgctg	ggctcgatgt	ggagccgg	cccagacgc	agggttcc	tctgcact	2700
caacactg	gtggccagg	acagtcccg	tttggatgtt	ccgacgtt	tgcacat	2760
ctggacggcc	ttggaccc	gacacatgt	ccgtgttgc	tttcttgc	tttcttgc	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	aggcggcc	ggccggcc	ggaatcttgg	gacccttgg	3960
gccccctggcc	taaagggtctc	tccagggtt	cctggccctt	tggggacccc	gggagacgca	4020
ggacctcgag	gccccaaagg	ggacccgggg	gtcccccgg	aagtcatcg	aggtgaagga	4080
cctgggttcc	ctggggggaa	aggggaccc	ggaccatcg	gcccccttgg	acccctgtt	4140
caactggggg	accggggacc	ccgtggcc	ccagggttcc	ctggaaacagc	catgaagggt	4200
gacaaaggccg	atcggtggga	ggggggcc	cctggaggac	gtgaagggtt	cattgtctt	4260
ggggagctgt	ggctgcgggg	tcttccgg	agccctggac	cccaaggccc	cgtggccccc	4320
cctggaaaga	aaggagaaaa	aggtgactt	gaggatgg	ctccaggctt	cccaggacaa	4380
cctgggttcc	cggtgtggaca	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	cggttgggt	4740
cttcctggag	acccttggcc	caaggagac	ccttggagacc	ggggccctt	tggcttact	4800
ggcagagcag	gaccccccagg	tgactcagg	ccttgggtgg	agaaggggaa	ccctggccgg	4860
cctggggccc	caggacgtt	tggccccc	ggacggatgt	gtgaagggtt	agagaaaagg	4920
gacgggggtc	tcctgggtga	ccccgggtt	ccttggaaa	caggccggcg	tggccttcgg	4980
ggggcacctg	gatgttgggg	gcctgttgg	gaaaagggg	accaggggaa	tcctggagag	5040
gatggacgaa	atggcagccc	tggatcat	ggacccaa	gtgaccgtt	ggagccgggt	5100
ccccccggac	ccccggggac	gtgtgttgc	acaggactt	gaggccagaa	gaagggag	5160
cctggggacc	cgccgacaa	gggttctcg	ggggccaa	gtatcttgg	cctcccttgg	5220
gccccctgggg	aaaggggcat	tgaagggtt	cgggggcccc	caggccccaa	gggggaccca	5280
ggtgtccgag	gccccagg	agaaaagggt	gacccgggtc	ccccctgggt	ggatggccgg	5340
agcggactgg	atggggaaacc	aggacccgt	ggggcccttc	ggccgaatgg	tgcttgaggc	5400
aaagctgggg	accggggag	agacgggtt	ccaggccctc	gtggagaaca	ggggccctt	5460
ggccccccttgc	gtccccctgg	attaccgg	aaaggcagg	aggatggcc	accttggctt	5520
aatggaaaaa	acggagaacc	tggggacc	ggagaagac	ggagggaaagg	agagaaaagg	5580
gattcaggcg	ctctggggat	agaaggctgt	gtatggccca	agggttgc	tggagctt	5640
gttatcttgc	gacccgggg	gtccctgg	gtccctgg	cagtggggcc	tcctggccag	5700
ggtttcttgc	gtgttccagg	aggacgggg	cccaagggtt	accgtgggg	gacttgatcc	5760
aaaggggagg	aggggctccc	tggagagcg	ggcttgc	gagagcttgg	aagtgtggc	5820
aatgtggatc	gtgtgttgc	aactgttgc	atcaaggat	ctggcttgc	ggagatctgt	5880
gagacgggg	atggagatc	tggtatgttgc	ctggcttgc	cgaaacggcg	tgcaggcccc	5940
aagggggact	caggcgaaca	ggggccccc	ggcaaggagg	ccccatccg	tttcttgg	6000
gaacgcgggc	tgaagggcga	cgctggagac	cctggccct	aggggcacc	tggcttggcc	6060
cttggggaga	ggggcccccc	cgggcttcc	ggcttgc	gggagcttgg	aaaggcttgg	6120
attcccccggc	tcccaggcg	ggctgggggt	gtggggagg	caggaaagg	aggagagg	6180
ggagaacccgg	gagagaaaagg	agaacgttgc	ggacaggcc	gagatggccc	tcctggactc	6240
cctggaaacc	ctggggcccc	cgggcccc	ggcccaagg	tgtcttgg	tgagccagg	6300
cctggactct	ctggagaaca	ggggccccc	ggactcaagg	gtgtcaagg	ggagccgggc	6360
agoaatgttgc	accaaggatc	caaaggac	aggggtgtgc	caggcatca	aggagacccg	6420
ggagagccgttgc	gagggggggg	tcaggacgg	acccgggttgc	taccggaga	tcgtgttat	6480
gctgggttgc	aaagggaaggcc	gggttgc	ggtccaaagg	gcccccttgg	cccagggtt	6540
ggtcatggag	acccttggacc	accttggtgc	ccgggttgc	ctggcccttc	aggaccccaa	6600
ggaccccttc	gtcttgcgg	ggagacccgt	ctccaggac	gggcctgtact	6660	
ggacacttgc	gagcttgcgg	acttcttgc	ccccccgg	cttcaggctt	tgtgggttca	6720
cagggttgc	cagggttgc	tggacaatgt	ggggagac	ggagccccc	agccccagg	6780
cgagatgtgt	ccagggttgc	agatggagac	agggggac	ctgtgttgc	agggttccacca	6840
ggtctgttgc	gcccgttgc	acctaaagg	gaaacccgttgc	ccacggggc	ccctggacag	6900
gtctgttgc	ggtcttgcgg	agcaaggagg	gagaaggagg	ccccctgg	ccttgcgttgc	6960
gaccccttgc	gtgttgcgg	agccaaagggt	gaccggagg	tgccaggggc	gcgaggccg	7020
aagggttgc	ctggccgtgc	agggggaccc	ggagacccct	ggaaatgg	tcaaaaagg	7080
gctccaggac	ccaaagggtt	caagggttgc	ccaggatgt	gggtcccccgg	ctcccttgg	7140
cctcccttgc	ctccagggttgc	gaagggttgc	ctggcttgc	gggttcccccgg	ccctggactt	7200
ggtgttgc	ggttcccccgg	tcagacagg	cctcgttgc	ggatgggttca	gccaggccct	7260
agtggagac	ggggtcttgc	agggggaccc	ggggagaa	gaatcccagg	acccttgggg	7320
ccaccccttgc	cacccgggg	atgtggggacca	cctggggcc	ctggactcaa	aggagacaag	7380

-continued

ggagaccctg	gagttagggc	gcctggggcc	cgaggcgagc	gtggggagcc	aggcatccgg	7440
ggtaaagatg	ggccggccccc	ccaggaggga	ccccggagac	tcacggggccc	cccttgcacg	7500
agggggagac	gtggggagaa	gggtgtatgt	gggagtgcag	gactaaaggg	tgacaagggg	7560
gactcactg	tgatcttggg	gccttcaggag	ccacgggggt	ccaaggggga	catgggttga	7620
cgagggcctc	ggggcttgg	tggtgacaaa	ggacctcggt	gagacaatgg	ggacccctgtt	7680
gacaaggggc	gcaaggggaga	gcctgttgac	aagggtctcg	ccgggttgc	aggactcgct	7740
ggactcttgt	gaccgggggg	tcacacttgg	gcacgggg	tcctttgtgt	ccggggatcc	7800
ccaggaaaaag	atgggatgtcc	tggatctccca	ggaaaaaaag	gagatgttgg	cttcatgggtt	7860
ccccggggcc	tcaagggtga	acggggagtg	aaggggact	gtggccttga	tggagagaag	7920
ggagacaagg	gagaagcttgg	tccccccaggc	cgccccccgg	tggcaggaca	caaaggagag	7980
atggggggatc	ctgggtgtcc	ggggccatcg	ggggccctcg	gcggaggagg	cctgtatcggt	8040
cccaagggttgc	accggggcttgc	tgacggggcc	ccaggccccca	agggttgacca	ggggcggaaaa	8100
ggggagccgg	gaaccccaagg	aatttgggggc	ttcccaaggcc	ccagtggaaa	tgatggctct	8160
getgggtcccc	caggggccacc	tggcaagtgg	gttcccaagg	gcccccaagg	acttcaggggc	8220
cagaagggttgc	aggcgagggttcc	ccccggggag	agagtgttgg	ggggcttctgg	gggtccctgg	8280
getcttgcgc	agagagggggaa	gcggggggcc	ccaggggctcg	ccgggttctcg	aggcgagaag	8340
ggagaqaactg	cactgacgg	ggatgcata	ccggggctttg	tgcccaaga	gatgagtctg	8400
cactgtgcct	gccaggggca	gttcatcgca	tctggatcac	gaccctcccc	tagttatgt	8460
gcagacactg	ccgggtccccc	gctccatgt	gtggcttgc	tccggcttc	tcatcgacag	8520
gaggaaggagc	gggtatcccc	ttggatgtat	gagttactctg	aatactccga	gtatctgtgt	8580
gaggagttacc	aggaccctg	agcttcttgg	gatagtgtat	accctgttc	cctggccactg	8640
gatgagggttgc	cctgcactgc	ctacaccctg	cgctgttacc	atcgggtctgt	gacaggcagc	8700
acagggggct	gtcacccctt	tgtctatgtt	ggctgtggag	ggaatgccaa	ccgtttttggg	8760
acccctgtgg	cctgcgtggc	ccgctgccca	ccccgggttg	tccagagcca	ggggacagggt	8820
actggccagg	actgt					8835

SEQ ID NO: 12 moltype = DNA length = 8835
 FEATURE Location/Qualifiers
 misc_feature 1..8835 note = Synthetic Construct
 source 1..8835 mol_type = other DNA
 organism = synthetic construct
 SEQUENCE: 12
 atgaccctga gactgttgtt ggctgcctg tggctggaa ttctggctga ggctcctaga 60
 gtggggccc agcatagaga aagagtgacc tgtacacggc tgtacccgc cgatatcg 120
 tttctgtgg atggcagcag cagcatcgcc cggttcaatt tcagagaatg cgccgggttt 180
 ctggaaaggcc tggtgtgtcc ttttctggc gctgttcgtt cccaggccg catatggcc 240
 acgtgtcagt acagcagcaca ccctagaacaat ggttggcc tggatgtctt ccggcagggc 300
 ggagatgtga tcagagccat cagagatgt agtacaaaag gccggcaatac cagaacaggc 360
 gccgttattc tgcaegtgc cgaccatgtt ttctggccc aactgtgtg accccgggtg 420
 ccaaagggtg gcatctgtt cagacaggc aagggccagg acctgtgttacccgcgt 480
 cagagactga aaggcaggg cgtttaactt ttcggctgg gatcaaaaggccgcattt 540
 gaggaactga agagagtggc cagccagcc accagcgatt tcttttctt cgtgaacgac 600
 ttcagcatcc tgccggaccct gctgtccctg gtgtttagaa ggttggcc tttttttttt 660
 ggatgtgtcc tgaccaggac tctgtatgtt agccacaaggcc cccctagaga tctgtgtgt 720
 tctggccaa cggccaggg tctgtatgtt cagtgccatg cttgtggccggg cccctgtt 780
 ggctacaagg tgcgtatatac acctctgaca ggctggcc agcctctgc ttctggaga 840
 caagaagtga acgtccccagc cggcgaaaca tctgtcagac tgagggact gggccctgt 900
 accggacttca aagtgtacatg gatccgtcc ttcggccat tctatggccggccgtt 960
 ggcacagcca gaacaacaggcc tctggaaatg ccagactgtg ccatccggaa cacaacaggc 1020
 cattttctgc tgggtgttcc gatgttgcgtt cttggccca ccggctatag agtgcattgg 1080
 agagtttgcg gccggaggccc cacacaggc caagaacttg gacctggaca gggctccgt 1140
 ctggctgtggg atcttcggcc tggcacccat tacaatgtg ccgtgtctca cttttttttt 1200
 agatctgtgg gacccgtccaa acgttgcgtt gtcggccat tccatggccggccgtt 1260
 ctggggccgtt tgatgttggg ccctacatct atctgtgtt gttggaaatcc tttttttttt 1320
 gccagaggct atagacttggatggggaga gaaaccggcc tggaaacctcc tcagaaatgt 1380
 gtctctgttca gcgacgttgcg aagatccacg ctggatggac tgcaggcccg caccggat 1440
 agactgtacatc tttttttttt gtcggggccg cacaaggatgttccatggccgtt 1500
 cttttttttt cttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 1560
 ggcacaggatgttccatggccgtt gttttttttt gttttttttt gttttttttt gttttttttt 1620
 gtggccgttca cacaggccgtt gggatggccatc tttttttttt gttttttttt gttttttttt 1680
 gatctgtggat atgtgtggcc cggactgtcc tacacatgtt ggttggcc tttttttttt 1740
 ctttagagaatgttccatggccgtt gttttttttt gttttttttt gttttttttt gttttttttt 1800
 gatctgtggccgtt gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 1860
 gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 1920
 cttttttttt gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 1980
 tttttttttt gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 2040
 gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 2100
 acgtgtccatcc tttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 2160
 totgttcaccc gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 2220
 gacggactggaaatccatgttccatggccgtt gttttttttt gttttttttt gttttttttt 2280
 gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 2340
 gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt gttttttttt 2400
 gggcccaacatgttccatggccgtt gttttttttt gttttttttt gttttttttt gttttttttt 2460
 atctctgttccatggccgtt gttttttttt gttttttttt gttttttttt gttttttttt 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	gg	gg	ccctgqcg	3540
caggccacg	gcctgtatgt	ttt	gg	gg	gttatct	3600
agaagggtt	ccccctggat	ggatagcgt	ccac	ttt	ggcgtgtt	3660
tctctggatc	aggctgtgtc	tggactgt	acagactgt	ggc	ccatccac	3720
cagecttagac	ctgagcttg	tccagtgtac	tgcc	ttt	gacaa	3780
gaatggac	tttagaggcca	agtgggacca	cctggc	ttt	acttgc	3840
gtgtgcctcg	gaccccaagg	tccttctgtt	agtg	cc	ttt	ttt
ccttgcgttgc	atggcagacc	aggatctct	gttag	cc	ttt	3960
gcaccaggac	tcaaaggatc	tccaggact	cctgg	ttt	ttt	4020
ggaccaagag	gtccaaaagg	cgaa	ttt	ttt	ttt	4080
cccgattgc	cggggaaagg	aggcgacc	ttt	ttt	ttt	4140
ccttggggat	atcccggtcc	ttt	ttt	ttt	ttt	4200
gacaaaggcg	ataggggaga	ttt	ttt	ttt	ttt	4260
ggcgaaccag	gtttgcctgg	actccctgc	ttt	ttt	ttt	4320
cctggaaaga	aaggcgaaaa	gg	ttt	ttt	ttt	4380
ccaggcttc	caggcgaaca	ttt	ttt	ttt	ttt	4440
gacagagggt	ttcccg	ttt	ttt	ttt	ttt	4500
ggacctgcg	gatctgatgg	ttt	ttt	ttt	ttt	4560
gaaggccac	ttt	ttt	ttt	ttt	ttt	4620
gatctctgt	ttt	ttt	ttt	ttt	ttt	4680
cctgctggac	caaggggag	ttt	ttt	ttt	ttt	4740
ttt	ttt	ttt	ttt	ttt	ttt	4800
ggcagagctg	ggcc	ttt	ttt	ttt	ttt	4860
cctggac	ttt	ttt	ttt	ttt	ttt	4920
gacagaggc	ttcc	ttt	ttt	ttt	ttt	4980
ggcgctctg	gttttagagg	ttt	ttt	ttt	ttt	5040
gatggaaag	atggccccc	ttt	ttt	ttt	ttt	5100
ccacgggtc	caccaggc	ttt	ttt	ttt	ttt	5160
cccg	ttt	ttt	ttt	ttt	ttt	5220
gcaccaggcg	agcgaggat	ttt	ttt	ttt	ttt	5280
gttgtcagag	gac	ttt	ttt	ttt	ttt	5340
tccggactt	atggaaaacc	ttt	ttt	ttt	ttt	5400
aaagctggc	acc	ttt	ttt	ttt	ttt	5460
ggac	ttt	ttt	ttt	ttt	ttt	5520
aacggccaaa	acggc	ttt	ttt	ttt	ttt	5580
gacagcggag	cctt	ttt	ttt	ttt	ttt	5640
ggaattttt	gaccccaagg	ttt	ttt	ttt	ttt	5700
ggtttctg	gtgtcccagg	ttt	ttt	ttt	ttt	5760
aagg	ttt	ttt	ttt	ttt	ttt	5820
aatgtggaca	gactcttgg	ttt	ttt	ttt	ttt	5880
gaaacccctgg	atagagac	ttt	ttt	ttt	ttt	5940
aaaggcgatt	ctggc	ttt	ttt	ttt	ttt	6000
gacgggg	ttaaaggc	ttt	ttt	ttt	ttt	6060
cttggagaa	gggg	ttt	ttt	ttt	ttt	6120
ataccggac	ttcccg	ttt	ttt	ttt	ttt	6180
ggagaaaagg	gaga	ttt	ttt	ttt	ttt	6240
cctggaaactc	ctgg	ttt	ttt	ttt	ttt	6300
ccaggctta	gggg	ttt	ttt	ttt	ttt	6360
tcaaatggc	agg	ttt	ttt	ttt	ttt	6420
ggcgaaccag	gac	ttt	ttt	ttt	ttt	6480
gcccggac	agg	ttt	ttt	ttt	ttt	6540
ggacatggc	acc	ttt	ttt	ttt	ttt	6600
ggaccatctc	gact	ttt	ttt	ttt	ttt	6660
ggacatctc	gact	ttt	ttt	ttt	ttt	6720
caagg	ttt	ttt	ttt	ttt	ttt	6780
cgagatggcg	cat	ttt	ttt	ttt	ttt	6840
ggggctcc	gac	ttt	ttt	ttt	ttt	6900
gcagtgtt	gat	ttt	ttt	ttt	ttt	6960
gatctgtt	gaga	ttt	ttt	ttt	ttt	7020
aaaggcga	ctgg	ttt	ttt	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	ctgtggcgaaa	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	ggggcaaqag	ggcacaaagg	cgacccaggc	7680
gataagggt	ctaaaggcga	aecggcgat	aagggtctg	ctggactgc	aggacttcga	7740
ggacttctgg	gtccacaagg	ccaaacctgg	ggcccgaggca	ttcccgccga	tccgggtct	7800
cctggcaaaag	atgggttcc	tggcatcagg	ggcgaaaagg	gcegactcg	ttttatgggc	7860
cetcgcggat	tgaaaaggcg	gaggggtt	aaggcgcc	gtggacttga	cgcgaaaaaa	7920
ggcgcatagg	tgcaagctgg	acccctggc	agacctgg	tggctggata	caaggcgcag	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	gaccagaagg	ctgtcgagg	8220
caaaaaggcg	agcgcggacc	accaggcgaa	agagtgttgg	gagcccccgg	cggtcccggt	8280
gctccggcg	aaagaggcga	acaaggcaga	cctggaccag	ctggccctag	aggcgaaaaaa	8340
ggggaaagccg	ctctgaccga	ggcgcataat	agaggctttt	tgccgcaaga	gtgagccag	8400
cattgtaccc	gtcaggggc	gtttatcg	agegggttca	gacctctgg	tagctatgc	8460
gctgtatccg	cggatctca	gtctgtatgt	gtgcctgtc	ttagagtgtc	tcacggccag	8520
gaagaggaaa	gagtccctcc	agaggacgac	gagtaactcc	agtatagcga	gtactctgt	8580
gaagagttatc	aggaccccga	ggctcttgg	gatagcga	atccatgtc	tctgccact	8640
gtatggggca	gtgtgttgc	ctacactgt	aggtggat	acagagocgt	gaccggaa	8700
accggaggct	gccccat	tgtttatgg	ggctgcggcg	gcaacgccaa	tagatttgg	8760
acaagagagg	cctgcgagcg	gagatgcct	ccaaagatgg	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	cggtgggaga	aacaaagct	gactcatggc	180
agcagcgct	acataaaactc	aacttgcgg	acacggggcc	atgcctccac	ctctgttac	240
aggagggtctc	actcacctgc	cttccactctg	cccaactccc	caggctcaac	ctttgaaagg	300
aaaactcactc	tatccggcc	tgccgtatgg	ggggagcttca	gtggcaactc	ttctccggag	360
taccctcgga	aggaatttgc	atcttctca	accaggagg	gggtcaaaac	acgagagagt	420
gaaatttgcg	ttcgcactca	gagtgcgcc	ccatccaccc	gtatggacaa	attggatgt	480
gttaagcggt	tgctcaaggg	gagtcgtat	gcaaggtgt	gccccacccg	gaattctcc	540
aacacactcc	ccatccccc	gaaaggcact	gtggagacca	aaatgtgtac	agcgagctcc	600
cagtctgggt	caggcaccta	cgatgcac	atcttgcgt	ccaaaccttcc	ctcccatgt	660
tggtcctcca	ccctgcggc	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	cctccggc	ctgcaccaca	agtgtgcaga	gcgtatgt	tttgcacaag	1020
gactgcaga	tctctgtat	agagaaaaaa	aacacaccc	ccaaaggaga	gtggagct	1080
ctctatcatga	ccaaggacag	cgggagggt	tttacagcc	ccctgcac	catcgctc	1140
acttctttt	cagaagacac	cctaaaaaaa	gaaaaggca	ctgcctacaa	tgttgcacat	1200
ggctctaaag	ccgaagctaa	tggagacgt	aagactgtgt	ccacaaagg	caagaccacc	1260
actgcagata	tccacactga	cggcggcgt	ggtgtgtgg	gcagtgg	agggtggccgt	1320
gttgtggccg	ctgggggggg	ccctttgggg	ccaggcc	ctgggtgccc	ctcgccgtcc	1380
tgtgtcgact	gtgtggaaat	gtctgtggc	tttgcgtat	ccctggctgt	actcttgggg	1440
ctgtctttcg	gcctcatgtc	tctggggag	gaggtggag	agctgtggc	gcgtgtggat	1500
gagctggaga	ggatcaggag	gagcataact	ccctatgggg	acagcatgga	tagataagaa	1560
aaaggccggc	tccaggat	ggcaccggcg	ggggggagcc	acctggacaa	aattgggtcg	1620
cacagtgc	gcccaggag	gtctgtat	ttcgtgg	ggaga	gttggacac	1680
gaaaatggaa	atctccgg	aaggccctgc	cctaaagggt	acatggga	tccaggccct	1740
aaaggagatc	gagggttccc	tggacttca	ggtatccctg	ggcccttggg	ccacccagg	1800
ccacaaggac	caaagggtca	aaaaggcago	gtggggagatc	ctggcatgga	aggccccat	1860
ggccagagag	ggcgagaagg	ccccatggga	cctctgtgt	aggcaggggcc	tcctgtatct	1920
ggagagaaag	ggggaaaggagg	ggctgtgt	ggaccagg	ttcatggccc	acctgggtgc	1980
ccagggttct	tgggtccaa	agttccago	ggctctctgg	gccccagg	ccctccagg	2040
cctgttagtc	tccaaagggtc	ccggaggt	gttggacttc	ctgggtgtca	aggtgacaaa	2100
ggaccaatgg	gaccaccagg	accctaaagg	gaccagggt	agaaagg	tcgaggctc	2160
acaggcgac	ctggcatgaa	agggttgc	ggtgtgttgc	gtgagccccc	ggctaaagg	2220
gcaatgggtc	ctgtggccc	agacggacac	caaggccaa	gagggtaa	acgttctact	2280
ggatctgg	gatctgg	ccaccagg	ccttctgg	accctagg	ggccaggatc	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	gccccaggagg	acaccaaggc	gagaaggc	tcccagggtt	ctcaacccctca	2820
gggtccagggt	tttcggact	caacccatcg	ggaccaccag	gcccacctgg	ccccaggga	2880
ccaaaagggtg	acaaaaggta	tccagggtt	ccagggttcc	tttgcattcc	tagtggtct	2940
tctgaaggggg	gatcatcaag	tccatgtac	tggtcaggcc	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	tggtcagtt	gttctcgcc	3240
tccatctttt	cttgaagacat	tctggatqtg	ctgcageggg	atgacgttgcg	tcagtactca	3300
cgtcagtaact	tgatggggcc	tcggggtccg	ccagggccac	caggagccag	tggagatggg	3360
tcctctctgt	ctttggacta	tgcaagatgt	agtatgtcc	ttctcagcta	catgtcgagt	3420
tctgggatca	gcatgggtt	tcctggtccc	ccggggccc	ctgggttgc	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	ttgtgttgc	tggtgttgc	gcaggagaca	gggggttcccta	4020
atcggccccc	gggggggttca	ttggggcagea	gcagaaggcc	gcatgtatgc	tggcaatggc	4080
ggacttattgg	gagctgtactt	tgctggatgt	ctggattaca	atgagctggc	tgtgagggtt	4140
tcagagagca	tgcaagttca	ggggctactg	caaggatgtt	cctacactgt	ccagggccca	4200
ccaggccatc	ctggggccacc	tggcccttca	ggacacaagg	gagggggccc	tgctggggca	4380
gaccaagtct	atgtctggccg	gagaaggaga	agaagtattt	ctgtcaagcc	gtga	4440
						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	aatcgatgacc	60
gaaaccgtga	ccaccagact	gaccgcgtt	cctccctaaag	ggggcacctc	taatggctac	120
gccaagagac	tttctctcg	cgaggcagc	agactggaaa	agcagttct	gacacacggc	180
agcagcgctt	acatcaatag	cacggctt	acaaaggccc	acggccagcac	aagcagctac	240
agaagggtctc	acagcccttc	cageacactg	cctaatacgcc	caggcggcac	cttcgagaga	300
aaagccccacg	tgaccagaca	cgccctacag	ggctctatca	ggggcaatag	cagccctggag	360
taccccccagg	aagagtgtcc	cagcggcgg	accaggaggc	gaaggccagc	acgggaaagc	420
gagatcagag	tgccgttgc	gtctgttgc	ccttagccac	gatggaccga	gtctggacgc	480
gtgaagggc	tgctgttgc	aagcagatcc	ggccagctgt	cccttaccag	aaacagcagc	540
aacacccttc	ctatttctaa	gaaaggccac	gtcgagacta	agatcgatc	agccagcagc	600
cagagctgtt	ccggccacata	cgatggccaca	atcctggacg	ccaaacctgg	tgcctatgt	660
tggtccttca	cactgttgc	cggaaggcago	atggccaccc	accacaaacaa	catgaccacaa	720
cagagcagca	gcctgttgc	caccaatgc	tactctggcc	gtccctgtt	cgccgttccca	780
aacaatattgg	ccagctgttgc	ccctctggcc	tgagcacaag	ctccctccgt	840	
tttggatgc	agaacaatct	ggccccttgc	ctgacacaacc	tgagccacgg	aaacaccaccc	900
accagcacat	cttacggcggt	ggaaaaaaac	atgccttgc	ctccagccgc	cgtgaacacaa	960
ggcggttagca	catgttgcgc	ctgttccaca	agcgatgc	gacgacgtt	tgcacaaag	1020
gactgtcaatgtt	ttctgttgc	ggaaaaggac	aacacggccc	ccaagaaga	gatggacttgc	1080
ctgtatcgat	ccaaaggacat	cggaagggtt	tttacccgtt	ctccagatct	tatcgccgc	1140
acaagcttta	cgcgaggac	ccttggatgtt	ggaaaaggac	ccgcgttacaa	cgccgactct	1200
ggactgttgc	ccggggccaa	cgccgttgc	aaaaacccgtt	cttaccaagg	caagaccacc	1260
accggccaca	tccacatgtt	tggatcttgc	ggaggccgg	gatctggatgg	tggccggccga	1320
gttggagggtt	ctggccgggg	accttgggg	cctgttgc	tttgggttcc	tttgggttcc	1380
tgttgcgtt	gggtggaaatgt	gtctgttgc	cttgggttgc	tcttctgttgc	1440	
ctgtgttgc	ggctgttgc	ccttgggttgc	ggatgttgc	agctgttgc	1500	
gagctgttgc	ggatcaggcg	gagatcgttgc	ccttacccgtt	acagcatgga	ccggatcgag	1560
aaggacacat	tgcaaggatgtt	ggcttccatgtt	gctggccgg	atctggat	gatcgacttgc	1620
cacagcgaca	gccaaggatgtt	actctgttgc	ttcgttgc	agaaactgt	gatggaaacaa	1680
gagaacggca	acctgttgc	cggcccttgc	cttaaggccgc	atatggaa	cccaggacca	1740
aaaggccaca	gaggcttcc	tggccacatgtt	ggcatttgc	gacatctgttgc	1800	
cctcaggccat	ctaaaggccat	gaaaggcttgc	gtggggatc	ccggcatgttgc	aggccctatgt	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
gctatggac	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
ggcggccagg	caggactcc	agggccat	aaagtgt	acctgcagg	gcctcctgg	2580
ccgcctggac	caagagggcc	accagggca	tctattcc	gaccaccag	tccttagagg	2640
cetccggc	aaaggatggc	aggctctt	ggggcacct	gcagctt	gagaacacag	2700
gagacatcc	tgagcggccc	tcaggacat	cctggacc	caggacat	aggcgat	2760
ggacctccag	gaccgagagg	acatcaggc	gaacaaggac	tgccctgc	tagcacaag	2820
ggcagctcta	gttccggct	caatctgc	ggtccaccag	ggccaccagg	acctaagg	2880
cctaaaggcg	ataaaggcga	tccaggc	ccagggc	tgggtat	ttctggacca	2940
tctgaaggcg	gtctccatcg	cactatgt	gtgtcagg	caccgg	acctggacc	3000
ccaggaccac	ctggatctat	ctctagctt	ggccaa	tccagc	catcagc	3060
tacatgcagt	ctgacagcat	ccggcttac	ctgtctgg	ttcaagg	accggac	3120
ccgggac	ctggac	tacaaca	accggc	gata	tttgcacta	3180
gectctc	tggtgttca	tctgaga	acgggtat	atgtgt	cccttgc	3240
agcatc	ccggaggacat	ctggcag	ctggcagg	atgac	gtcg	3300
agacagtatc	tgatggggcc	cagaggc	cctgg	caggc	cgatgt	3360
tctctgt	gcttgattt	cgccg	agcagc	gata	tctgacta	3420
ccoggcatct	ccatggg	gctctgg	ccagg	tttgc	tttgc	3480
tacgaggaa	tgctgttcc	gttgggg	agcagg	tttgc	tttgc	3540
ggggcc	gttccac	tatcc	aatgtgt	cttgc	tttgc	3600
ctgagc	acttgc	acac	acggcc	atgtt	tttgc	3660
gggcctc	ctgtcc	ggcc	accgtt	tttgc	tttgc	3720
gagaac	actcttc	tgatc	tttgc	tttgc	tttgc	3780
agtttatc	tttgc	tttgc	tttgc	tttgc	tttgc	3840
ctgctg	gagca	ccac	tttgc	tttgc	tttgc	3900
gttgc	ggagac	gttcc	tttgc	tttgc	tttgc	3960
cttgg	gatct	tttgc	tttgc	tttgc	tttgc	4020
attgg	gggttgc	tttgc	tttgc	tttgc	tttgc	4080
ggact	gttgc	tttgc	tttgc	tttgc	tttgc	4140
tccg	ggac	tttgc	tttgc	tttgc	tttgc	4200
ccagg	acgc	tttgc	tttgc	tttgc	tttgc	4260
aacgt	ggactt	tttgc	tttgc	tttgc	tttgc	4320
ggccaa	aaagg	tttgc	tttgc	tttgc	tttgc	4380
cctgg	gatc	tttgc	tttgc	tttgc	tttgc	4440
gacca	acttgc	tttgc	tttgc	tttgc	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	LLAATALLT	HGQEEGQVEG	QDEDIPPI	TVCNGLRYHDR	DVWKPEPCRI	60
CVDNGKVLC	DDVICDET	KCPGAEVPEGE	CCPVCPDGSE	SPTDQETTV	EGPKGDTGPR	120
GPRGPAGPPG	RDGIPQPGPL	PGPPGPPGPP	GPPGLGGNFNA	PQLSYGYDEK	STGGISVPGP	180
MGSFSGPRLP	GPPGAPGPQG	FQGPPGEPGE	PGASGPMGPR	GPPGPPGKNG	DDGEAGKPGR	240
PGERGPPGQ	GARGLPGTAG	LPGMKGHRGF	SGLDGAKGDA	GPAGPKGEPG	SPGENGAPGQ	300
MGPRGLPER	GRPGAPGPAG	ARGNNDGATGA	AGPPGPTGPA	GPPGFPGAVG	AKGEAGPQGP	360
RGSEGPGQVR	GEPPGPPGAG	AAGPAGNPGA	DGQPAGAKG	GAPGIAGAPG	FPGARGPSGP	420
QGPGGPPGPK	GNSGEPGAPG	SKGDTGAKG	PGFVGVQGP	GPAGEEGKRG	ARGEPGPTGL	480
PGPPGERGPP	GSRGFPGADG	VAGKPGPAGE	PGFVGVQGP	GPAGEEGKRG	ARGEPGPTGL	540
TGSPGSPGPD	GKTGPPGPAG	QDGRGPQGP	PGARGQAVM	GFPGPKGAAG	EPGKAGERGV	600
PGPPGAVGPA	GKDGEAGAQQ	PPGPAGPAGE	RGEQGPAGSP	GFQGLPGPAG	PPGEAKPGE	660
QGVPGLQGLP	GPSGARGER	FPGERGVQGP	PGPAGPRGAN	GAPGNDGAKG	DAGAPGAPGS	720
QGAPGLQGMP	GERGAAGLPG	PKGDRGDAGP	KGADGSPGK	GVRLTGPIG	PPGPAGAPGD	780
KGESGPGSPA	GPTGARGAPG	DRGEPPGPPG	AGFAGPPGAD	GQPGAKGEPG	DAGAKGDAGP	840
PGFAGPAGPP	GPIGNVGAPG	AKGARGSAGP	PGATGFPGA	GRVGPQPSG	NAGPPGPPG	900
AGKEGGKGP	GETGPAGRP	EVGPPGPPGP	AGEKGSPGAD	GPAGAPGTPG	PQGIAGQRGV	960
VGLPGQRGER	GFPGLGPBGS	EPGKQGPSPA	SGERGPPGPM	GPPGLAGPPG	ESGRREGAPGA	1020
EGSPGRDGP	GAKGDRGET	PAGPPGAPGA	PGAPGPVGPA	GKSGDRGET	PAGPAGPVGP	1080
VGARGPAGPQ	GPRGDKGET	EQGDRGKIGH	RGFSGLQGP	GPPGSPGEQG	PSGASGPAGP	1140
RGPGPSAGAP	GKDGLNGLPG	PIGPPGPRGR	TGDAGPVGP	GPPGPPGPPG	PPSAGFDTSF	1200
LPQPQKEAH	DGGRYYYRADD	ANVVRDRDLE	VDTTLKSLSQ	QIENIRSP	EGSRKNPARTCR	1260
DLKMCHSDW	SGEYWIDPNQ	GCNLDAIKVF	CINMETGETCV	YPTQPSVAQK	NWYISKNPKD	1320
KRHVWFGE	TDGFQFEYGG	QGSDPADVAI	QLTFLRLMST	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 TINAGSQFPEY 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
SNGAPQRGGE PGQPGHAGAQ GPPGPPGNG AGIPGAPGLM GARGPPGPAG 420
ANGAPGLRRG AGEPKNGKA GEPGPRGERG EAIGPVGPGA KGEDGKDGP GEPGANGLP 480
AAGERGAPGF RGPAGPNIGP GEKGPGAGERG APGPAGPRGA AGEPGRDGPV GGPGRMRGMPG 540
SPGPGSDGK PGPPGSQGES GRPGPPGPGP PRQGPQVMGF PGPKGNDGAP GKNGERGGPG 600
GPGPQGPQGP NGETGPQGP GPTGPQGDKG DTGPGPQGPQQL QGLPGTGGPP GENKGKPEPG 660
PKGDAGAPGA PGKGDAGAP GERGPGLAG APGLRGAGP PGPEGGKGAAG GPPGPAGAAG 720
TPGLQGMPGE RGGLGSPGPK GDKGEPPGPG ADGVPGKDPG RGPTGPIGPP GPAGQPGDKG 780
EGGAPGLPGI AGPRGPSPGER GETGPQGPAG FPGAPQGNGE PGKGGERGP GEKGEQGPPG 840
VAGPPGPGVK AGPPGPQGVK GERGSPGPGP AAGFPGARGL PGPPGSNGNP GPPGPSPG 900
KDGPGPAGN TGAPGSPGVG GPKGDAGQGP 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 GAPGPCCGG 1200
GAAAIAIGIGG EKAGGFAPYY GDEPMDFKIN TDEIMTSLSK VNGQIESLIS PDGSRKNPAR 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 GLPGMKGDG EILGHVPGML LKGERGFPGI 180
PGTPGPQGPCLP GLQGPVGPQGP FTGPGPQGP PGPPGEKQGM GLSFQGPQKGD KGDQGVSGPP 240
GVPGQAQVQE KGDFATKGEK GQKGEPEFQG MPGVGKGEKGP GKPQPRGKPG KGDQKGEKGS 300
PGFPGEFQGP GLIGRQGPQG EKGEAGPPGP PGIVIGTGP GEKGERGYPG TPGPRGEPPG 360
KGFPGLPQGP GPPGPVPGQ AGAPGFPGER GEKGDRGFPQG TSLPGPSPGRD GLPQGPQGP 420

```

- continued

PPGQPGYNTNG	IVECQPGPPG	DQGPPGIPGQ	PGFIGEIGEK	GQKGESCLIC	DIDGYRGPPG	480
PQGPPGIEGF	PQGPKAKGKD	GLPGRDGVAG	VPGPQGTPGL	IQGPGAKGEP	GEFYFDLRLK	540
GDKGDPGPFG	QPGMPGRAGS	PGRDGHGPGLP	GPKGSPGSVG	LKGERGPPGG	VGFPGSRGDT	600
GPPGPPGYGP	AGPIGDKGQA	GFPGPGPGSPG	LPGPKGEPK	IVPLPGPPGA	EGLPGSPGFP	660
GPQGDRGFPG	TPGRPGLPGE	KGAVGQPGIG	FPGPPGPKGV	DGLPGDMGPP	GTPGRGPFNG	720
LPGNPGVQGQ	KGEPEVGLPG	LKGLPGLPGI	PCTPGEKEKSI	GVPGVPGEHG	AIGPPGLQGI	780
RGEPEPGPPGLP	GSGVSPGVPG	IGPPGARGPP	GGQGPPGLSG	PPGIKGEKGF	PGFPGDMPG	840
PKGDKGQAGL	PGITQSGSLP	GLPGQQGAPG	IPGFPGSKGE	MGVMGTPGQP	GSPGPVGAPG	900
LPGEKGHDHG	PGSSGPRGDP	GLKGDKGDV	LPGKPGSMDK	VDMGSMKGQK	GQDQEKGQIG	960
PIGEKGSRGD	PTGPGVPGKD	QGAGPQGPQG	PKGDPGIGST	PGAPGLPGBK	GSGVGMGLPG	1020
TPGEKGVPGI	PQPGQSGLKG	CDKAKGEKG	QAGPPGIGIP	GLRGEKGDQG	IAGPPGSPGE	1080
KGEKGSGTIP	CMPGSPGLKG	SPGSVGVPGS	PGLPGEKDK	GLPGLDGIPG	VKEAGEGLPGT	1140
PGPTGPAGQK	GEPGSDGIPG	SAGEKEPGL	PGRGFPGPFG	AKGDKGSKGE	VGFPGLAGSP	1200
GIPGSKGEQG	FMGPGPQGQ	PGLPGSPGHA	TEGPKGDGRP	QGQGPLPGLP	GPMGPGLPG	1260
IDGVKGDKGN	PGWPGAPGV	GPKGDGPQG	MPIGIGGSPGI	TGSKGDGMGPP	GVPGFQGPKG	1320
LPGLOQIGKD	QGDQGVPGA	GLPGLPGPPG	PYDIIKGEPG	LPGEPGPGL	KGLQGLPGPK	1380
GQQGVTGLVG	IPGPGPIPFG	DGAPGQKJEM	GPAGPTGPRG	FPGPPGPGL	PGSMGPPGTP	1440
SVDHGFLVTR	HSQTIDDPQC	PSGTKILYHG	YSSLVYQCN	RAHGQDLGTA	GSCLRKFSTM	1500
PFLFCNCINNV	CNFASRNNDY	YWLSTPEMP	MSMAPITGEN	IRPFISRCAV	CEAPAMVMAV	1560
HSQTIQIPPC	PSGWSSLWIG	YSFVMMHTSAG	AEGSGQALAS	PGSCLEERS	APFIECHGRG	1620
TNCYYANAYS	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	CAIKKGLEQI	LVGGSKLTK	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	PRGDPGEAIP	QGDQGREGP	GVPGDPGEAG	PIGPKGYRGD	480
EGPPGSEGAR	GAPGPAGPPG	DPLGLMGERGE	DGPAGNGTEG	FPGFPGPYGN	RGAPGINGTK	540
GYPGLKGDEG	EAGDPGDDNN	DIAPRGVKGA	KGYRGPEGPQ	GPPGHQGPQ	PDECEILDII	600
MKMCSCCECK	CGPIDLFLV	DSSEISGLQN	FEIAKDFVVK	VIDRLSRDEL	VKFEFGQSYA	660
GUUVQYSHSQM	QEVHVSLSRSPS	IRNVQELKEA	IQLSLQWMAGG	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	EFGLDALGSG	GDVIRAIREL	SYKGGNTRTG	120
AAILHVADHV	FLPQLARPGV	PVKCILITDG	KSQDLVTDAA	QRLKGQGVKL	FAVGIKNADP	180
EELKRVASQ	TSDFFFFVN	FSILRLLPL	VSRVRCVTTAG	GVPVTRPDD	STSAPRDLVL	240
SEPSSQSRLR	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	DTGLQPGIT	660
YQVAVSVLRL	REGPEAAVIV	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	SSQRTGLEP	1020
GVSYIFSLTP	VLDVGRGPEA	SVTQTPVCP	GLADVVFLPH	ATQDNRAA	ATRRVLERLV	1080
LALGPLGPQA	VQVGLLSSYH	RPSPLFLPLNG	SHDLGIILQR	IRDMPYMDPS	GNNLGTA	1140
AHYRMLAPDA	PGRRQHVGPG	MVLLVDEPLR	GDIFSPIRE	QASGLNVML	GMAGADPEQL	1200
RRLAPGMDSV	QTFFAVDDGP	SLDQAVSGLA	TALCQASFTT	QPRPEPCPVY	CPKGQKGEPG	1260
EMGLRGQVGP	PGDPGLPGR	GAPGPQGPPG	SATAKGERGF	PGADGRPGSP	GRAGNPCTPG	1320
APGLKGSPGL	PGPRGDPGER	GPRGPKGEPG	APGQVIGEGG	PGLPGRKGDP	GPSGPGPGR	1380

-continued

PLGDPGPRGP	PGLPGTAMKG	DKGDRGERGP	PGPGEGGIAP	GEPLGPLGP	SPGPQGPVGP	1440
PGKKGKEDGS	EDGAPGLPGD	PGSPGEQQPR	GPPGAIGPKG	DRGFPGLGE	AGEKGERGPP	1500
GPAGRSRGLPG	VAGRPAKGP	EGPPGPTGRQ	GEKGEPPRPG	DPAVVGPAVA	GPKGEKGDVG	1560
PAGPRGATGV	QGERGPGLV	LPGDPGPKD	PGDRGPIGLT	GRAGPPGDSG	PPGEKDPGR	1620
PGPPGPVGPR	GRDGEVGEKG	DEGPPGDPGL	PGKAGERGLR	GAPGVVRGPVG	EKGDQGDPGE	1680
DGRNGSPGSS	GPKGDRGEPE	PPGPPGRLVD	TGPGAREKEE	PGDRGQEGPR	GPKGDPGLP	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	IKAASALREIV	ETWDESSGS	LPVPERRRGP	1980
KGDSCGEQCPP	KGEKPTGFP	ERGLKDRGD	PGPQGPPGLA	LGERGPPGPs	GLAGEPKPG	2040
IPGLPGRAGG	VGEAGRPGF	GERGEKGERG	EQQRDGPGL	PGTPGPVPPP	GPKVSVDEPG	2100
PGLSGEQGPP	GLKGAKGEPE	SNGDQGPKD	RGVPGIKGRD	GEPPGPRQGD	NPGLPGERGM	2160
AGFEGKPLQ	CPRGPGPVGP	GHGDPGPPGA	PGLAGPAGPQ	GPSGLKCEPG	ETGPPGRGLT	2220
GPTGAVGLPG	PGPSPGLVGP	QGSPGLPGQV	GETGKPGPAG	RDGASGKDG	RGSPGVPGSP	2280
GLPVGPKG	EPGPTGAPGQ	AVVGLPGAKG	EKGAPGGLAG	DLVGEPKAGK	DRGLPGRPRGE	2340
KGEAGRAGEP	GDPGEDQKG	APGPKGFKD	PGVGVPGSPG	PPGPPGVKD	LGLPGLPGAP	2400
GVVGFPCQTG	PRGEMCOPGP	SGERGLAGPP	GREGIPGPLG	PPGPPGSVGP	PGASGLKGDK	2460
GDGPVGLPGP	RGERGEPEGR	GEDGRPGQEG	PRLTGPPGS	RGERGEKGDV	GSAGLKGDKG	2520
DSAVILPGP	PRGAKGDMGE	RGPGRGLDGD	GPRGDNGDPG	DKGSKEPGD	KGSAGLPGCLR	2580
GLLGPQGQPG	AAGIPGDPGS	PGKDGVPGIR	GEKGDVFGM	PRGLKGERGV	KGACGLDGEK	2640
GDKGEAGC	PGPGLAGHKGE	MGEPGVPGQS	GAPGKEGLIG	PKGDRGFDGQ	PGPKGDQGEK	2700
GERGTPGIGG	PGPSPGNDGS	AGPPGPGV	GPRGPGLQ	QKGERGPPE	RVVGAPVPG	2760
APGERGEQGR	PGPAGPRGEK	GEAALTEDSV	RGFVRQEMSQ	HACACQOFIA	SGSRPLPSYA	2820
ADTAGSQLHA	VPVLRVSHAE	EEERVPPEDD	EYSEYSEYSV	EEYQDPEAPW	DSDDPCSLPL	2880
DEGSCTAYTL	RWYHRARTGS	TEACHPFVY	GCGGNANRFG	TREACERRCP	PRVVQSQGTG	2940
TAQD						2944

SEQ ID NO: 21	moltype = AA	length = 1497
FEATURE		
source		
1..1497		
mol_type = protein		
organism = Homo sapiens		
SEQUENCE: 21		
MDVTKKNKR DGETVERITV ETVTTRLTS L PPKGGTSNGY AKTASLGGS RLEKQSLTHG 60		
SSGYINSTGS TRGHASTSSY RRRAHSPTSL PNPGSTFER KTHVTRHAYE GSSSGNSSPE 120		
YPRKEFASSS TRGRSQTRS EIRVRLQSAS PSTRWTLEDD VKRLLKGSR ASVSPTRNS 180		
NTLPPIPCKGT VETKIVTASS QSVSGTYDATT IL DANLP SHV WSSTLPAGSS MGTYHNNMTT 240		
QSSSLLNNTNA YSAGSVGPV NNMPQSPAAVNT NMSLNTN QMSPAAVNT SVQSDLLHK DCKFLILEKD NTPAKKEMEL 300		
TSTAYGVKVN MPQSPAAVNT GSLSVAACTT TSFSAACTT SVQSDLLHK DCKFLILEKD NTPAKKEMEL 360		
LIMTKDSGKV FTASPIASIAA TSFASDETLKK EKQAYANADS GLKAEANGDL KTVSTKGKT 420		
TADIHSYGS CGGGSGGGGG VGGAGGGPWG PAPAWCPCGS CCSWWKLILG LLLTWLLL LG 480		
LIFGLTALAE EVRKLLKARVD ELERIRSL PYGDSDMIDI KDRLQGMAPA AGADLDKIGL 540		
HSDSQEELWM PFKLKLMMEQ ENGNLRLGSPG PKGDMGSPGP KGDRGFPGT GIPGPLHGP 600		
PQGPKGQKGS VGDPMEGPM QRQREGPMG PRGEAGPPGS GEKGERGAAG EPGPHGPV 660		
PGSVPGPKGSS GSPGPQGPPG PVGLQGLRGE VGLPGVKGDK GPMGPPGPKG DQEKEKPRGL 720		
TGPGRPMGRGL GAVGEGPAK AMGPGAPGDGH QGPRGEQGLT GMPGIRGP GP SGDPGKPG 780		
TGPQGPQGLP TGPGRPGIKG EPAGPKIVT SEGSSMLTV P GPPGPPGAMG PPGPPGAPGP 840		
AGPAGLPGHQ EVLNLQGPPG PPGRGPQGP S IGPQGPGRG PPGEGLPQPP GPPGFLSNS 900		
ETFLSGPPG PGPPGPQGDQ GPPGPGRHQ EQGLPGFSTS GSSSFGLNLQ GPPGPPGPQG 960		
PKGDKGDPV PGALGIPSPG SEGGSSSTMV VSGPPGPQGP PGPPGSISSS QOEIQQYISE 1020		
YMQSDSIRSY LSGVQDGPV GP GPPGPVTTI TGETFDYSEL ASHVSYLRT SGYGVSLFSS 1080		
SISSESDILAV LQRDDVRQYL RQYLMGPGRG PGPGASGDG SLLSLDYAEL SSRILSYMSS 1140		
SGISIGLPGP PGPPGLPGTS YEELLSLLRG SEFRGIVGP GPPGPPGIPG NVWSSISVED 1200		
LSSYLYTHAGL SFIPGPGP GPPGPGRGP VSGALATYAA ENSDSFRSEL ISYLTSPDVR 1260		
SFIVGPPGP EA GQGPQGPDSR LLSTDASHSR GSSSSHHSS VRRGSSYSSS MSTGGGGAGS 1320		
LGAGGAFGEA AGDRGPYGTG IGPQGGYGA AEGGMYA CNG GLLGADFAGL LDYNELAVRV 1380		
SESMQRQGLL QGMAYTVQGP PGQPGPQGP GISKVFSAYS NVTADMDF QTYGAIQGPP 1440		
GQKGEMGTPG PKGDRGPAGP PGHPGPPGP GHKGEKGDKG DQVYAGR RRR RSAVKP 1497		

SEQ ID NO: 22	moltype = DNA	length = 148
FEATURE		
misc_feature		
1..148		
note = Synthetic Construct		
source		
1..148		
mol_type = other DNA		
organism = synthetic construct		
SEQUENCE: 22		
gaatttcgtcg attcggttgc agcatttaaaa gccgttacata actttaaaag aaggaaaaag 60		
aagggttgaag aaaagggtgt agtaagtaat tataagtaca gaccggagaa gtacccgggt 120		
ctcgatttcgat ttaatttgcgaa aagaagaaa 148		
SEQ ID NO: 23	moltype = DNA	length = 491
FEATURE		
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 gtgcagaaac gcctgagata ccacgaccgg gatgtgtgg agcccgagcc ttgcagaatc 180 tgcgtgtgcg acaaattggca ggtgtgtgc gacgacgtga tctgcgacga gacaaagaat 240 tgcctggccg cggaaatggcc tgaggggccg tgggtgtccg tggccctgg tggcagcgg 300 agccccacag atcaagagac aacaggccgtg gaaggccccca agggcgatac aggacataga 360 ggtcctagag gacctgcccgg acctctggc agagatggaa ttctggaca gcctggactg 420 cccgaggacc acggacccctcc agggcctcca ggtccaccag gactcgagg aaattttgcc 480 ccacagctga gotacggcta cgacggagaa agcacaggcc gcatctctgt gcctggaccc 540 atggggaccc ttggcccaag aggacttccctt ggtcctccgt gtgcctccgg accctcaggga 600 tttcaaggac caccaggcga acctggcga ccaggcgcta ttggcccaat gggaccaaga 660 ggccctctgg ggccaccagg caaaaatggc gacgatggc aagccggaaa gcctggaaagg 720 cttggcggaaa gaggccggccg aggacgcga ggccgttagag ggttgcctgg aactgcagg 780 ctgcctggcc tgaaggggcca gagggcttt tctggactgg atggcgctaa gggcgacgct 840 ggaccaggac gacctaagg cgacgctgga tctccctggc agaatggtc acctggacag 900 atggggccca gaggattggc aggccggaga ggttagacctg gcgtccggg accagccgg 960 gtcagggaaa atgatggcgc aacaggctg gtcggccctt ctggaccaac cggaccagct 1020 ggccacccctg gatttccagg cgctgttgg acaaaaggccg aacggccccc acaaggaccc 1080 aggggatcttca aaggctctca ggccgtttaga ggccggccg ggccacccctgg gcctggccgt 1140 gcagctggac ctgctggaaa cctctggctg gatggacage cagggtccca aggtgctaat 1200 ggcgccccctg gatttggccg cgctccaggat tttccctggc caaggaggcc atctggaccc 1260 caaggccccag ggccacccctcc gggtccttaag ggaatagccg gagggccagg cgctctgg 1320 agttaaaggccg atactggcgc aaaaaggccaa cccggacccctt ttggagttca aggacccct 1380 ggaccaggctg gcgaggaaagg caaaaggaggc gctaggggag aaccaggacc aacagggctc 1440 ctctggccac ctggcggagcc cggggacccctt ggtcgttagag gatccctgg cgccatggc 1500 gtggccggac caaaaaggacc tgcaggccggaa agggggatcc cagggtctgc agggccctaag 1560 ggttctccag gcgaggctgg cagacccggc gaaatggcc tcccagggtc taaggactg 1620 acaggctcac caggatctcc cggaccaggac gggaaaacag gaccccttccagg accggccagg 1680 caggatgtta gacccggctt ccgttgacccg cctggtccaa gaggacaagc tggcgatgt 1740 ggctttctgg gacccaaagg tgcaggccggaa gaaacctggaa aaggaggccgaa gaggggagtt 1800 cccgacccctt cagggtctgt tggacccctt gggaaaaggat ggcggatcttgc tgcacaaagg 1860 cctccaggccg cagccggacc agccggccgg agggccgaaac aaggaccagg cggatctcca 1920 ggatttcagg gactggcagg gcctgttggc ccgcctggc aggccaggaa gcccaggccaa 1980 cagggtgttcc tggccgatctt tggcccttagc gaggatgggg cgaaaaggagga 2040 tttccctggccg aaaaaggccgt tggaggccgtt ccggggacccat ctggccggcc gggatgttcaat 2100 ggtgccccccg gcaatgtacgg tgcttaaggcc gggccggccg ccccgatggc tcctggatct 2160 caaggccgcac ctggacttca gggaaatggc tggccaaacgggg gaggtgttgg acttccgg 2220 ccaaaaggccg ataggggaga tgctggctt aaggccgtgg atggcttcc ttggaaaggat 2280 ggcgcaagat ggcgtacagg cccaaattggc cctccgggac ccgtggccgc tccaggccat 2340 aaggggcaat ctggacccat tggacccctt ggtcgttagag gttgttgggg agcccccaggc 2400 gaccggggag accctggcc accaggaccc tggccgttcc ctggccggcc tggccgttat 2460 ggtcacccctt gtgctaaaggcc gggccggccg gacccggccgg taaaaggccg cggatcttca 2520 cctggaccgg ccggaccccttc tggccggccca ggacccatgg gaaatgttgg tggccctggc 2580 gccaaggccg caagaggatc tggccggccca ccaggccgtt caggattccc aggtggccgt 2640 ggaagagttt gaccaccggg gccaagggttga aatgtggac caccggggacc gccaggacca 2700 ggccggccaaag aagggtggaaa aggcccttagg ggcggaaaatgg ccctggaggc caggccaggc 2760	

-continued

gaagtggccc	ctccaggacc	tccggggct	gccggcgaaa	aaggatctcc	aggcgcgat	2820
ggacccgcag	gogctcccg	aacacagggt	ccacaggaa	ttgtctggaca	aaggggagtt	2880
gtcggcgtgc	caggacagag	gggagagaga	ggtttccag	gactccctgg	gccaagcgga	2940
gaacctggca	aacaggggacc	atcttgtcc	agcgagaga	gagggccacc	aggaccaatg	3000
ggtcctccag	gattggcagg	gectctggc	aatctgtta	gagaagggtc	tccaggcgcc	3060
gagggatctc	ctggacgtg	tggttctct	ggcgccaaagg	gcatagagg	cgaaaacaggc	3120
ccagctggac	ctccaggcgc	acccggcgct	ccaggcgac	caggactgt	tgccctgtct	3180
gaaaaatctt	gogacagagg	cgaaaactgg	cccgccaggac	cagccggacc	tgttggacct	3240
gtgggtgtct	gaggaccggc	tggaccacaa	ggtccttaga	gogacaagg	cgaaaacaggc	3300
gaccaaggcc	acagaggcat	caaggacac	agaggattca	cgccgactca	gggaccacca	3360
ggcccgccgt	gaaatcccg	cgacaggaa	ccaaaggcgaa	ctatggtcc	cgccggaccc	3420
agaggaccac	ctgggtctgc	tggtgccacc	ggaaaggacg	gactgtatgg	gtccccggga	3480
cctattgggc	cacccggacc	tagaggaaga	acaggcgac	caggaccatg	tggaccaccc	3540
ggggccaccc	gaccggctgg	tcctcttcg	cctcccttc	ccggattcga	cttcagcttcc	3600
ctgcctcage	ctcccaaga	gaaggccat	gacggcgccg	gatatacag	agccgacgac	3660
gccaacgtc	tccggggacag	agatctggaa	gtggacacca	cactgaatgc	cctgttccag	3720
cagatcgaga	acatcagaag	ccccgaggc	agcagaaaga	accctgcac	aacctgtcgg	3780
gacctgaga	tgtggccac	cgatggaaag	tctggcgagt	actgtatcga	ccccaaacccag	3840
ggctgcaacc	ttggatggat	caagggttcc	tgcaacatgg	aaaccggcga	gatcatgcgt	3900
taccccaac	acccatctgt	ggctcagaag	aactgttaca	tcaacaaaggaa	ccccaaaggac	3960
aagccggcac	tttgggtccg	cgagagcatg	accgtatgc	tccagtttg	gtatggcgcc	4020
caggggcttc	accctggcc	tgttgctat	caagtcaccc	tcctgcggct	gatgtctaca	4080
gaggccagcc	agaacatcac	cttacactg	aagaacagc	tggcttacat	ggatcagcg	4140
accggcaacc	tgaagaaggc	actgtctt	cgaggcgac	acgagatcga	gtatggcgcc	4200
gagggcaaca	gcccgttac	ctacagcg	acagtggatg	gtgcacccag	ccatccaggg	4260
gcttggggca	agaccgtat	cgatgtacag	accaccaaa	ccagcgacat	gccccatcatc	4320
atgtgtggcc	ctctgtatgt	tggggcaccc	gtcaagatg	togggttca	tgtggggccca	4380
gtgtgtctcc	tgaggcccaa	ggggggcago	ggcgaggggca	ggggcagct	gtctgaccctc	4440
ggcgacgtgg	aggagaaccc	cgccccctcg	agttctgtgg	acacccaaac	actgtctgt	4500
ctggccgtga	cactgtgt	ggccacttgc	caagggctgc	aagggaaac	agtgcggaaa	4560
ggacccgtccg	cgatgtatgg	acccatgggg	gaaagggatgc	ctccctggacc	tcctgtttaga	4620
atggcggagg	atggacccat	aggaccatc	gttccaccc	gaccccttcc	gctcttggcc	4680
cttggaggaa	attttggcc	tcagttacat	ggcaaaaggcg	tcggacttgg	cccttgaccc	4740
atgggactta	tggggccaa	aggaccacca	ggtgctcg	gctgtccagg	accacaagga	4800
tttcaaggac	cagctggcc	gcctggccaa	cctggacaaa	cagggttcc	tggtgctttag	4860
ggaccaggcg	ggccaccc	aaaaggctgg	gaagatggc	acccttggaa	gcctgtttaga	4920
ccccggcaaa	gggggttgc	tggacccat	ggggccacag	gttttccctt	aacacccctt	4980
ctgcctgtgt	tcaaggccat	cagaggccac	aatggccctgg	acggactgaa	aggacaaccc	5040
gtgtgtcttc	gggtgtatgg	cgaaacccgg	gcacccgtgg	aaaatggccac	accaggacaa	5100
accggcgtaa	gaggacttcc	ttggcgaga	ggaagatgtt	gagcccccgg	tccaggacgg	5160
gcacggat	ctgtatgtat	tgtggaccc	gttggccctt	ccggacccat	tggatgtct	5220
ggccctctgt	gatttctgg	cgccaccc	ccaaaggccg	aaatgggac	tgtggaaac	5280
gcccggcc	caggcccg	tggaccaagg	ggagaagggt	tattgcctgg	actjagcg	5340
ccagttggcc	caccaggaa	tcctgttcc	aatggactg	caggcgctaa	aggtcgacgt	5400
ggccctccat	gogttggcc	tgccaccc	ctggccaggac	caagggat	ccctgtgtct	5460
gttggagctg	ctggcgat	gggttgc	ggacttgc	gagaacctgg	gccagccgga	5520
tctaaggccg	actgtgtat	caaggccgg	ccaggatct	ctggccacca	aggcccccct	5580
ggacccatcg	cgcaagaagg	caaaggccg	cctaattggcc	aaggccgtat	tggccggcc	5640
cctggaccac	caggctttag	aggatcttc	ggctctat	gattgcctgg	cgctgtatgt	5700
agacgaggcg	ttatgggtcc	acctggatca	agggcgct	ctggcccttgc	tggcgat	5760
gttccaaatg	gogccgtct	cagccaggac	gagcccggtt	ttatggggcc	tagagggtt	5820
cctggaaagcc	ctggcaat	cgccccc	ggaaaaggaa	gcccctgttt	actccctggc	5880
atcgacggcc	gacctgtat	aatcgaccc	gcaggcgct	ggggaggaccc	tggaaatatt	5940
ggcttccctg	ggcctaaagg	ccccaccc	gatctggaa	agaacccgca	taaggccat	6000
gtctggactcg	ctgggtatgg	ggggccaccc	ggacccgt	aaaacaatgg	tgtctcaagg	6060
ccgcctggcc	cacaagggt	tcaagggtt	aaaggccgac	aggcccccac	tgggccttca	6120
ggcttccaa	gacttcccg	accatctgg	ccaggcgcc	aagttggaaa	gcctggccaa	6180
agaggactgc	acggcgat	tggcttgc	ggtcttgc	gtccacccgg	agagagac	6240
cctccaggcc	aatctggcc	cgcggaccc	actggccctt	tcggaaaccc	aggacccat	6300
ggacccctcc	gacccatgtt	caacaaaggc	gaaccttgg	ttgtggccgc	tgtggaaaca	6360
gtctggaccc	cttggcttcc	tggatgtt	ggcgacccgg	gaggatgtt	tattccctgt	6420
ggccaaaggcc	aaaaggccg	gcctggact	agggcgac	tccggcaatcc	cggacccat	6480
ggcgcttag	gogccccc	tgcagttgt	gccccggac	ctgtgtccgc	aacaggccac	6540
agaggccagg	ctgggtcc	tggcttgc	ggggccaccc	gttccat	ccggccgg	6600
gagggccagg	aatgtggacc	cgctggacc	aatggat	ttggccccc	tggcgat	6660
ggtcaaccc	gogccaaagg	cgaggccgg	gtaaaggcc	ctaaaggcc	gaatggcg	6720
gtgggtgtct	ctggaccat	ggggccat	ggggccac	gtccctaa	accacccctt	6780
ccagctgggt	ctagaggcc	cgccggacc	cctggat	cagggtttcc	aggcccccgt	6840
ggaagaacac	gtccctcc	accatctgg	atctctgg	caccaggccc	acccgttcc	6900
gctggaaaag	aaggactg	aggccctt	ggcgat	ttccat	ccggccgg	6960
gaagtccgg	ctgtccgg	accagg	ttttt	ggggccctt	cggagaat	7020
ggaactgcac	gaccggccgg	aactcccg	cctcaaggat	tgcttggcc	ccctggaaat	7080
ctgggactgc	ccggtagt	cgggagaac	ggactcc	gttgc	ccggccgg	7140
gaaccgggac	cacttggat	tgctggat	cctggat	gaggtcc	acc	7200
ggaagtctg	gcgttaacgg	tgctcc	ggac	gagatggaaa	tcccgcaat	7260
gtggccccc	ctggggagaga	tggacage	tttcc	ggacataagg	gcccgtcg	7320

-continued

aatattggac	ctgtcgccgc	agccgggtgct	cccggacactc	atggtccagt	cggtccagcc	7380
gggaaggcacg	gaaatagggg	agaaaacaggaa	ccctccggtc	ctgtggccc	agctggcgca	7440
gttggaccAA	gaggccatc	cggaacctcg	ggaatcccg	gagataagggg	cgaacctggc	7500
gagaaggggac	ctagaggact	gcctgggtg	aaaggccata	acggactgca	aggccgtcca	7560
ggcattgtcg	gccatcatgg	cgatcaagg	gcacccggta	gtgtgggtcc	cgccggacgg	7620
aggggtcccg	ctgggtccatc	tggacccggd	ggaaaagatg	cgagaacagg	acatctggc	7680
acagtggggc	ctgcccgaat	tagaggccca	caggacatc	aaggccccgc	tggggcccca	7740
ggacctccgg	gaccggcagg	gccaccaggg	gttagtgccg	gaggatcaga	tttcggctac	7800
gacggcgaact	tctacagagc	cgaccagct	agatctgcc	ctagctgag	gcctaaggac	7860
taacaaatgg	acggccacact	gaatgtccctg	aacaacccaga	tgcgacact	gtgaccct	7920
gaggggcagca	gaaagaaacc	tgcaccaacc	tgcacccggacc	tgtagactgtc	tcacccggaa	7980
tggtcctccg	gtactactg	gategacccc	aatcagggtc	gcacatggaa	cgccatcaag	8040
gtgtactgcg	acttcagcac	cggegagaca	tgcatcagag	cccagectga	gaacatcccc	8100
gccaagaact	ggtcagaag	cagaacggg	aagaaacacg	tgtggctggg	cgagacatac	8160
aacccggcct	gccaggctga	gtacaacgtg	gaaggcgtga	ccagcaaaaga	gttggccaca	8220
caqctggctt	tcatgagact	gctggccat	tacggccacg	aaacatcac	ctaccactc	8280
aagaacacga	ttgcctacat	ggacgaggaa	accggcaacc	tgaagaaagc	cgtgtaccc	8340
cagggtccca	acggcgtgg	actgtggcc	gagggcaaca	cgagattcac	ctacaccgt	8400
ctgggtggacg	gctgcagca	aaagaccaac	gatgtggggca	agaccatcat	cgagtataaag	8460
accaacaacg	ccaggcagact	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	tttcgagaatc	180
tgcgtgtcg	acaatggcaa	ggtgcgtgtc	gacgacgtga	tctgcgacga	gacaaagaat	240
tgcctctggc	cggaaatgtcc	tggggccgg	tgtttgtctt	tgtgccttgc	tggcagcgag	300
agccccacag	atcaagagac	aaacggcggt	gaaggccccca	agggcgtatc	aggacataga	360
ggtcctagag	gacctgcccgg	acccctctgg	agagatggaa	ttccctggaca	gcctggactg	420
cccccggaccac	ctggacccctc	agggcctcca	ggtccaccag	gactcgagg	aaatttgcc	480
ccacacgtgt	gatcgtggca	cgacgagaaa	agcacaggcc	gcatctctgt	gcctggaccc	540
atggggacctt	ctggcccaag	aggactctt	ggtcctcttg	gtgtccagg	acctcgagg	600
tttcaaggac	caccaggcga	acctggcga	ccaggcgctt	tggtgtccat	gggaccaaga	660
ggccctctgt	ggccaccagg	caaaaatggc	gacgtggcc	aagccggaaa	gcctggaaagg	720
cctggcggaaa	gaggccggcc	aggacgcga	ggggcttagag	ggtgtccctgg	aactgcagg	780
ctgcctggcc	tgaaggggca	cgagggttt	tctggactgg	atggcgctaa	gggcgcacgt	840
ggaccacgt	gacctaaagg	cgagctgtt	tctccctggc	agaatgttgc	acctggacag	900
atgggtccca	gaggattggc	aggcgagaga	ggttagacctg	gcgtccggg	accagccgg	960
gtctagggaa	atgtggcgcc	aacaggtgt	gtggggcttc	ctggaccaa	cggaccag	1020
ggccacacgt	gattttccgg	cgctgttgg	gcaaaaggccg	aagcaggccc	acaaggacac	1080
agggggatctt	aagggtctca	gggggttta	ggcgagccag	ggccaccccttgg	gcctggccgt	1140
gcagctggac	ctgtgtggaa	ccctgggtct	gatggacagg	cagggtccaa	agggtctaat	1200
ggcgccccgt	gtatttgcgg	cgctccaggat	tttccctggc	caagaggccc	atctggaccc	1260
caaggccccag	ggggacccctc	gggttctta	ggaaatagcg	gagacccagg	cgcttctgg	1320
agtaaaaggcg	atactggcgc	aaaaggcga	cccgaccccttgg	tggggttca	aggacccctt	1380
ggaccacgt	gctggggacgg	caaaagaggg	gctaggggg	aaccaggacc	aacagggttc	1440
cctggccacac	ctggcggacgg	cgggggaccc	ggatcttagag	gatcccttgg	cgccagatgg	1500
gtggccggac	aaaaggacc	tgcggggaa	agggggatcac	cagggtctgc	agggccctaag	1560
ggttctccgg	gogaggctgg	cagacccggc	gaaagctggac	tcccagggtc	taagggactg	1620
acaggctcac	caggatctcc	cgggaccagac	ggaaaaaacag	gaccccttgg	accggccagg	1680
caggatgtt	gacccgggtcc	tcctggaccc	cctggtcaaa	gaggacaagc	tggtgtatg	1740
ggtttccctgt	gacccaaagg	tgcggccgg	gaaacctggaa	aagcggccga	gagggggatgt	1800
cccgacccatc	cagggtctgt	tggaccccttgg	ggaaaagatg	cgcaacgttt	tgccacaagg	1860
cctccaggcc	cacccggacc	agccggcgag	agggcgacac	aaggaccagg	cggtatctca	1920
ggatttcagg	gactggcagg	gcctgttgc	ccgccttggc	aggcaggggaa	gccaggcgaa	1980
cagggtgttc	tcggcgatct	tggccatcc	gggttcttag	gaggttccat	cgaaaaggaga	2040
tttccctggcc	aaaggccatgt	tggggccat	ccggggaccc	ctggggacaa	gggtgtcaat	2100
ggtgcccccc	gcaatgtacgg	tgcttaaggcc	tgccggccgg	ccccagggtc	tcctggatct	2160
caaggccgcac	ctggacttca	ggggatgcct	ggcgaacccgg	gaggtgttgc	acttcccggt	2220
ccaaaaggcg	ataggggaga	tgctggctt	aggggcgct	atggcttcc	tggaaaggat	2280
ggcggtcgat	tcggcgatct	tggccatcc	gggttcttag	gaggttccat	cgaaaaggaga	2340
aaggggcaat	ctggacccat	tggaccccttgg	gggttcttag	ggtgttccat	cgaaaaggaga	2400
gaccggggag	accctgttcc	accaggaccc	tggggatcc	ctggggatcc	ttggcgctat	2460
ggtcaacccgt	gtgtcaagg	cgacccggc	gacgttgg	aaaaaggcga	cgctggtcca	2520
cctggaccgg	ccggacccctc	tggccggcca	ggacccatgg	gaaatgttgg	ttggcccttgc	2580
gccaaggcg	caaggatgtt	tgctggccca	ccaggccatgg	agggttcc	ggccaggacca	2640
ggaagagatgtt	gaccacccgg	gccaaggatgg	aatgttggac	caccggggac	ggccaggacca	2700
ggccggcaaa	aaagggtggaaa	aggcccttgg	ggcggaaaatg	ccctgtcagg	caggccaggc	2760

-continued

gaagtggccc	ctccaggacc	tccggggct	gccccgc当地	aaggatctcc	aggcgc当地	2820
ggacccgc当地	gogctccccc当地	aacccagggt	ccacaggaa当地	ttgtctggaca当地	aaggggaggt	2880
gtccggctgc当地	caggacagag	gggagagaga当地	gtttttccag	gactccctgg当地	gccaagcgga当地	2940
gaacctggca当地	aacaggggacc当地	atctgggtcc当地	agcggagaga当地	gagggccacc当地	aggaccaatg当地	3000
ggtcctccag当地	gattggcagg当地	gectcttgc当地	aatctgtta当地	gagaagggtc当地	tccaggcgcc当地	3060
gaggggatctc当地	ctggacgtg当地	tggttctctc当地	ggcgccaaagg当地	gcatagagg当地	cgaaaacaggc当地	3120
ccagctggac当地	ctccaggcgc当地	acccggcgct当地	ccaggcgca当地	caggacctgt当地	tgcccctgt当地	3180
gaaaaatctc当地	gogacagagg当地	cgaaaactgg当地	cccgccaggac当地	cagccggacc当地	tgtttggacc当地	3240
gtgggtgtct当地	gaggaccggc当地	tggaccacaa当地	ggtcttagag当地	gogacaaggg当地	cgaaaacaggc当地	3300
gaccaaggcc当地	acagaggcat当地	caaggggacac当地	agaggattca当地	cgccgactgca当地	gggaccacca当地	3360
ggcccgccctg当地	gaaatccccc当地	cgacaggagg当地	cgacaggggca当地	cgacggggca当地	cgacggggca当地	3420
agaggaccac当地	ctgggtctgc当地	tggtgaccac当地	ggaaaggagg当地	gactgtatgg当地	gctccccc当地	3480
cctattgggc当地	cacctggacc当地	tagaggaaga当地	acaggcgacg当地	caggaccagt当地	tggaccacct当地	3540
ggggccacccgt当地	gaccggctgg当地	tcctcccttg当地	cccccattc当地	ccggatttgc当地	cttcagctt当地	3600
ctgcctcage当地	cttcctaaga当地	gaaggcccat当地	gacggggcga当地	gatatacag当地	agccgacgac当地	3660
gccaacgtcg当地	tccggggacag当地	agatgttgc当地	gttggacacc当地	cactgaatgc当地	cctgttctag当地	3720
cagatcgaga当地	acatcagaag当地	ccccgagggc当地	agcggaaagg当地	accctgtccag当地	aacctgtcgg当地	3780
gaccttgaaga当地	tgtggccacag当地	cgatggaaagg当地	tctggcgag当地	actggatcga当地	ccccaaacccag当地	3840
ggctgc当地	ttggatgtt当地	tgcaacatgg当地	aaaccggcga当地	gacatgcgt当地	3900	
taccccaacac当地	acccatctgt当地	ggctc当地	aactgttaca当地	tcagcaaggaa当地	ccccaaaggac当地	3960
aagccggc当地	tttgggtc当地	cgagagcatg当地	accgtatggc当地	tcctgttgc当地	gtatggccg当地	4020
caggggcttc当地	accctggcc当地	tgttgtctc当地	caagctgaccc当地	tcctgtccgt当地	gatgtctaca当地	4080
gaggccaggc当地	agaacatcac当地	cttccactgc当地	aagaacaggc当地	ttggcttacat当地	gttgc当地	4140
accggcaacc当地	tgaagaaggc当地	actgtgtt当地	cgaggccaggc当地	acgagatcga当地	gtatggccg当地	4200
gagggcaaca当地	gocgggtc当地	cttccactgc当地	acagtggtat当地	gttgc当地	ccatccaggc当地	4260
gcttggggca当地	agaccgtgt当地	cgatgttaca当地	accaccaagg当地	ccagcaact当地	gcccatcatc当地	4320
gtatgtggccc当地	ctctgtatgt当地	ttggggcacc当地	gatcaaggat当地	tcctgttgc当地	gttggcccca当地	4380
gttgtcttcc当地	tgttaaaggat当地	cgtcgat当地	ttttggccgt当地	ttaaaggcgt当地	tgacaactt当地	4440
aaaaaaatgg当地	aaaaaaatgg当地	tgaagaaaag当地	gggtgttagta当地	gtaagtataa当地	gtacagaccg当地	4500
gagaagtacg当地	cggtgttctg当地	ttcggtttaat当地	tttggaaagg当地	aaaatgttgc当地	gtttcggtt当地	4560
caccaggaaaca当地	ctgtgttgc当地	ttggccgtglocal	actgtgttgc当地	gcccacttgc当地	agagctgtc当地	4620
agagggaaaca当地	gtggccggaaag当地	gacccgttgc当地	cgatgttgc当地	ccctaggggc当地	aaaggagggtc当地	4680
tccctggact当地	cttggtagag当地	atggccggg当地	tggaccctaca当地	ggaccacctg当地	gttccaccagg当地	4740
acctccaggc当地	cctccctggcc当地	ttggaggaaa当地	ttttggccgt当地	cagtacgt当地	gcaaaggcgt当地	4800
cggaacttggc当地	cctggaccatc当地	ttggacttctg当地	ggggcccaagg当地	ggaccacccag当地	gttgc当地	4860
cgttccaggc当地	ccacaaggat当地	ttcaaggacc当地	agctggcgag当地	cttggccgac当地	ctggacaaac当地	4920
agggtccctgt当地	gtgtctgtc当地	gaccaggccg当地	ccacccttgc当地	aaagctgtgc当地	aagatgggc当地	4980
ccctggaaag当地	ccttggtagac当地	ccggccggaaag当地	gggtgttgc当地	ggacccctcaag当地	gccccaggagg当地	5040
cttcctctgg当地	acacccctggc当地	tgcctgggt当地	caaggccatc当地	agacggccaca当地	atggcttgc当地	5100
cggactgttgc当地	ggacaatctgc当地	gtgtctctgg当地	cgtaaaaggc当地	gaaccaggcg当地	caccctggc当地	5160
aaatggcaca当地	ccaggacaaa当地	ccggccgcaag当地	aggacttccct当地	ggcgagagag当地	gaagagtgt当地	5220
agccccaggat当地	ccaggcaggcg当地	cacggaggatc当地	tgatgttgc当地	gttgggactt当地	ttggccctgc当地	5280
cggacccatt当地	ggaagtgtcc当地	gccccctctgg当地	atttccctggc当地	gcacccggac当地	caaaggccgca当地	5340
aatttggaget当地	gtggggaaacg当地	ccggccggat当地	aggccctgc当地	ggacccaggat当地	ggggaggat当地	5400
atttgc当地	cttggccggc当地	cgttggggcc当地	accagggtat当地	cttgggtc当地	atggactgtc当地	5460
aggccgtaaa当地	gggtgc当地	gccttccagg当地	cggttgc当地	gcaccaggac当地	tgccaggacc当地	5520
aaggaggatc当地	cttgggtctg当地	tttgggtc当地	ttgggtgtc当地	gggtccaggag当地	gacttgggt当地	5580
agaaacctgg当地	caaggccgtat当地	cttgggtc当地	ttgggtgtc当地	gggtccaggag当地	gacttgggt当地	5640
tggtccacaa当地	ggccccc当地	gaccatcagg当地	cgaaagaagg当地	aaacggaggc当地	ctaattggc当地	5700
agccggtagt当地	cgccggcc当地	cttggaccacc当地	aggcccttgc当地	ggatcttctg当地	gtctctaggag当地	5760
atttgc当地	cttggccggc当地	gttggacttgc当地	gaggccgtat当地	tatgggttca当地	cttggatcaa当地	5820
tggccctgt当地	ggccgttagag当地	gttccaaatgc当地	cgacggccgt当地	agaccaggcg当地	agccgggtt当地	5880
tatggggct当地	agagggttgc当地	cttggaccacc当地	ttggccatatc当地	ggcccaaggcc当地	gaaaaggagg当地	5940
ccctgttgg当地	ctccctggca当地	tcgacggtag当地	acctggaccac当地	atcgccaccg当地	caggccgttag当地	6000
ggggagagct当地	ggaaatattt当地	gttccctctgg当地	gcttaaaggc当地	ccccacaggc当地	atccctggaaa当地	6060
gaacccggcat当地	aaggccgtat当地	cttggacttgc当地	ttttggccgt当地	ggggcccttgc当地	gacccgtacgg当地	6120
aaacaatctgt当地	gttcaaggcc当地	ccgttccggcc当地	acaagggttgc当地	caagggttgc当地	aaggccgac当地	6180
ggggccaccc当地	ggccctccag当地	gttcccaagg当地	acttcccgcc当地	ccatctggc当地	caggccggc当地	6240
agttggaaag当地	cttggccgaaag当地	gaggacttgc当地	ccggccgtat当地	ggcttgc当地	gttccctgg当地	6300
tccacgggg当地	gagagaggcc当地	cttccaggcc当地	atctggccgc当地	ggccggacc当地	tcggccctt当地	6360
cggaaaggc当地	ggacccatgt当地	gaccctccagg当地	acctgtatgc当地	aaacccaggc当地	aacccgttgc当地	6420
tgtggggc当地	gtggggacag当地	cttggaccctt当地	ttttggccgt当地	ggatttgc当地	ggccggccgg当地	6480
agcagctgt当地	atttctgttgc当地	gcaaaaggcc当地	aaaggccggag当地	ccttggactc当地	gaggccggag当地	6540
cgccaatcc当地	ggacccatgt当地	cgccctccagg当地	cgccctccagg当地	gttccctggc当地	ccccccggacc当地	6600
tgttggccgca当地	acaggccgaca当地	gaggccggat当地	ttttggccgt当地	ggccggccgt当地	ggccggccgg当地	6660
tccctaggat当地	agttccaggcc当地	aggcccttgc当地	gttcccttgc当地	ggccggccgt当地	ggccggccgg当地	6720
tggggccctg当地	ggccgtctglocal	gttccaaatgc当地	cgccaaaggc当地	gagccggggag当地	cttaagggtt当地	6780
taaaggccgag当地	aatggccgtc当地	ttggcccttac当地	ttggccatgt当地	ggagccggcc当地	ggccccc当地	6840
tccctaaccgg当地	ccaccctggc当地	caggccgggtc当地	tttggccggcc当地	ggccggaccgc当地	cttggaaatgc当地	6900
agggttttcc当地	ggccggccgt当地	gagaacccagg当地	ttttggccgt当地	ccatctggc当地	tctctgggtt当地	6960
accaggccgca当地	ccttggctctg当地	ctggaaaaggat当地	aggactgttgc当地	ggcccttgc当地	ggccggccgg当地	7020
tccagttgg当地	agaacccggcc当地	aaggccgtat当地	ttttggccgt当地	ccagggtttt当地	ccggccggcc当地	7080
aggcccttagc当地	ggagaagctg当地	gaactgttgc当地	accggccggcc当地	acttcccggtc当地	cttcaaggatt当地	7140
gttttggccgcc当地	ccttggaaattc当地	ttttggacttgc当地	ccggccgtat当地	ggagaacccgtg当地	gactcccgacc当地	7200
tgttgc当地	ccggccggcc当地	aaaggccggcc当地	ttttggccgt当地	gttccggacc当地	cttggccgcaag当地	7260
aggttccaccct当地	gggtccatgt当地	ttttggccgt当地	ccggccgtat当地	gttccacccggc当地	aaggccggcc当地	7320

-continued

agatggaaat	ccggcaatg	atggcccgc	tggagagat	ggacagccgt	gacataaggg	7380
c gagcaggc	tacccaggc	atatggacc	tgtcgccgca	gccggtgctc	ccggacactca	7440
tggccagtc	ggtcagccg	ggaagcacgg	aatagggga	gaaacaggac	cctccgggtcc	7500
tgttggcca	gtggcgca	ttggaccaag	aggccatcc	ggacctcagg	gaatcccg	7560
agataaggc	gaacctggcg	agaagggacc	tagaggact	ctgggtctga	aaggccataa	7620
cggactgcaa	ggccctggccag	gcatgtcg	ccatcatggc	gatcaagggt	caccggtag	7680
tgtgggtccc	gccccggca	ggggtccccc	tggccatct	ggacccgcg	gaaaagatgg	7740
cagaacagga	catcttggca	cagtggggcc	tgcggaaat	agagggccac	agggacatca	7800
aggccccgt	ggccggccag	gacccgggg	accggcagg	ccaccaggcg	ttatggccgg	7860
aggatatacgat	tccgtctacg	acggcactt	ctacagagcc	gaccagccta	gatctgcccc	7920
tagctgtgg	cataaggcg	acgaatgg	cgcacactg	aagtccctga	acaaccagat	7980
c gagacact	ctgacccctg	agggcagcg	aaaaaaccc	gccagaaact	gcaggaccc	8040
gagactgtct	caccccgaa	ggteetccgg	ctactactgg	atcgacccca	atcagggtcg	8100
caccatggac	gocatcaagg	tgtactgca	cttcgaccc	ggcgagacat	gcatcagac	8160
ccagcctgt	aacatcccg	ccaagaact	gtacagaa	agcaaggaca	agaaacacgt	8220
gtgtgtggc	gagacaatc	acggcccg	ccagttcg	tacaacgtgg	aaggcgtgac	8280
cagcaaagag	atggccacac	agctggctt	catgagact	ctggccatt	acggccagcc	8340
gaacatcacc	taccactgca	agaacagat	tgcctatcat	gacgaggaaa	ccggcaaccc	8400
gaagaaaagcc	gtgatccctgc	agggcttaa	cgacgtggaa	ctgggtggcc	agggcaacag	8460
casattacc	taacccgtgc	tgggtggcc	ctgcagcaaa	aagaccaacg	agtggggccaa	8520
gaccatcatc	gagtataaga	ccaacaagec	cagcagact	cccttcttg	atatcgcccc	8580
actggatatt	ggaggcggcc	accaagagtt	cttggac	atcgcccccg	tgtgtttcaa	8640
gtga						8644

SEQ ID NO: 34	moltype = 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	ctgactgt	ctgtcttgg	ctgtctacagc	cctgtctgaca	60
cacggacaag	aggaaggc	ggtcgaagga	caggacgagg	acatccccc	tatcacctgt	120
gtgcagaa	gcctgtggat	ccacggaccgg	gatgtgtgg	agcccgaggcc	ttgcagaatc	180
tgcgtgtcg	aatatggca	ggtgtgtgc	gacgacgtg	tctgcgacga	gacaaagaat	240
tgccttggc	ccggaaatggc	tgaggccgaa	tgttgtctt	tgtgccttgc	ttggcagcgag	300
agccccacag	atcaagagac	aacaggcg	gaaggccca	agggcgat	aggacactag	360
gtgtccatag	gacccgtccg	acccctctgc	agagatggaa	ttccctggaca	gccttgactg	420
cccgaggacc	ctggacatcc	agggcctca	gggtccaccag	gactcgagg	aaattttgccc	480
ccacagctg	gatccggct	cgacggaaa	agcagggcc	gcatcttgc	gccttgaccc	540
atgggaccc	ctggcccaag	aggacttct	ggtcccttgc	gtgtccagg	acccctagg	600
tttcaaggac	caccggcg	acccggcg	ccaggcgat	gtgggtccat	gggaccaaga	660
ggccctctg	ggccaccagg	ccaaaatgg	gacgatggc	aaaggccggaa	gcctggaa	720
cctggc	gaggcccgcc	aggacccaa	ggcgctag	ggtgtgcctt	aactcgagg	780
ctgccttgc	tgaaggcc	cagaggctt	tctgtactgg	atggcgctaa	gggcgacgc	840
ggaccaggac	gacccatgg	cgacgttgc	tctcttggcc	agaatgttgc	accttgacag	900
atgggtccca	gaggattgg	aggcgagaa	ggtagac	ggctccggg	accaggccg	960
gtctagggaa	atgtggc	aacaggct	gtctggctt	ctggaccaac	cgaggacgt	1020
ggcccac	ctggatcc	cgctgttgc	gcaaaaggcc	aaggcaggcc	acaaggaccc	1080
agggggat	aaatgttca	ggggcttgc	ggcgacgg	ggccaccc	gccttgccgt	1140
gcagctgg	ctgtggaa	ccctgggt	gtatggac	cagggtccaa	agggtctaat	1200
ggcgcccc	gatgtggcg	cgctccagg	ttttccggc	caagggggcc	atctggaccc	1260
caaggccc	cgccac	gggtctca	ggaaatagc	gagacccgg	cgctcttgg	1320
agtaaaaggc	atactggc	aaaaggcg	cccgaccc	tggggatc	aggacccctt	1380
ggaccacgt	gcaaggagg	caaaaggagg	gctagggggg	aaccaggacc	aacagggtc	1440
cctgggtcc	ctggcgac	cgaggac	ggatctag	gattcccttgc	cgccatggc	1500
gtggccgg	caaaaggacc	tgcaggcg	agggatc	caggcttgc	aggccctaa	1560
ggttcttcc	cgagggtcg	caggccgg	gaagctgt	tcccagggtc	taaggactg	1620
acagggtcc	caggatcc	cgggacc	ggaaaacac	gacccggcc	accggccagg	1680
caaggatgt	gaccgggtcc	tcctgtgg	gggtccat	gaggacaa	ttggctgtat	1740
ggcttcc	gacccaaagg	tgcagccgc	gaaatgttgc	aaggcaggcc	gagggggat	1800
cccgaccc	cagggtctgt	tggaccc	ggaaaatgt	gcgaagcttgc	tgcacaagg	1860
cctccagg	caggccggacc	agccggcc	agaggcg	aaggaccgc	cggtatccca	1920
ggatcc	cgactggc	gctgtct	cccccggcc	ggcaggggaa	cgccaggcc	1980
caaaaaggcg	ataggggaga	tgcgtgttct	gggtcttgc	gaggatagg	cgaaaaggag	2040
ggcgctac	gctgtac	ttggggcc	tcggaccc	ctggacca	gggtgtcaat	2100
gggtcccc	gcaatgtac	tgctaa	gacgcaggcg	ccccagggtc	tcctggatct	2160
caaggcgc	ctggacttca	gggaatgc	ggcgaacccgg	gagctgttgc	acttcccggt	2220
caaaaaggcg	ataggggaga	tgcgtgttct	aaaggcgct	atggcttcc	tggaaaggat	2280
ggcgctac	gctgtac	ttggggcc	tcctccggac	ctgctggcc	tccaggccat	2340
aggggcga	ctggaccc	tggaccc	ggtcttac	gtgttgggg	agccccaggc	2400
gaccgggg	agcctggtcc	accaggac	gctggatttgc	tggaccc	tggcgctgtat	2460
ggtcacac	gtgtcaagg	cgagccagg	caaaaggccg	cgctggcc	2520	
cctggaccc	ccggac	tggccggcc	ggaccttattgc	tggcccttgg	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	agagttccgg	cgagcaggga	ccaagcgag	ctagtggcc	cgccggaccc	3420
agaggaccac	ctgggtctcg	tgggtcaccc	ggaaaggagg	gactgaatgg	gtccccggga	3480
cctattgggc	cacccgtgg	tagaggaga	acaggcgac	caggaccat	tggaccaccc	3540
ggggccacctg	gaccgctgg	tcctcttgg	cctcccttgc	ccggatttca	cttgcaccc	3600
ctgcctcage	ctccctaaga	gaaggccat	gacggggca	gatattacag	agccgacac	3660
gccaacgtcg	tgccggacag	agatctggaa	gtggacacca	cactgaatgc	cctgtctcag	3720
cagatcgaga	acatcaga	ccccgggg	agcagaaa	accctgcac	aaccgtgtcg	3780
gaccttgaaga	tgtgcaccc	cgatggaa	tctggcgat	actggatcga	ccccaaaccc	3840
ggctgtcaacc	ttggatggccat	caagggttcc	tgcaacatgg	aaaccggcga	gacatcgctg	3900
taccccaac	agccatctgt	ggctcagaag	aactggata	tcagcaagaa	ccccaaaggac	3960
aagccggcacc	tttgggttcc	cgagagat	accgtatgg	tccatggta	gtatggcgcc	4020
cagggctctg	accctggcc	tgggtctatc	cgactgtatc	tcttgcgggt	gtatgtctaca	4080
gaggccagcc	agaacatcac	cttccactgc	aagaacagc	tggcttacat	ggatcagcag	4140
acccggcaacc	tgaagaaggc	actgtgttcc	caggccagca	acgagatcga	gatcagagcc	4200
gagggcaaca	gcgcgttcc	ctacaggctg	acagtggat	gtgtcaccag	ccatcacaggc	4260
gcttggggca	agaccgtat	cgatgtccaa	accaccaaga	ccagcagact	gcccatcatc	4320
gatgtggccc	ctctgtatgt	tggggcaccc	gatcaagat	tccggttgc	tgtgggccc	4380
gtgtgttcc	tgtaaatgttat	tttccacat	attggcgat	ttttggcaat	tgaggggcccg	4440
gaaaaccttgc	cttgcgttct	tgaegacat	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
ggggctgtca	gatgcccaga	aggtatccca	ttgtatggg	tctgtatgg	ggcctcggt	4800
cacatgttt	acatgtgtt	agtcgagg	aaaaaggcgt	taggc	ccccccca	4860
gacgtgggtt	tcctttgaaa	aacacgatgc	tgagcttgc	ggacaccaga	acactgtgc	4920
tgtgtggccgt	gacactgtgc	ctggcactt	gtcagagcgt	gcaagaggaa	acagtgcgg	4980
aaggacgtgc	cgccgtataga	ggacatagag	cgcaagatgg	gcacccctgg	cctcttgcgt	5040
gagatggcga	gatgtggac	acaggaccac	ctggccacc	aggaccccca	gggcctctg	5100
gccttggagg	aaatttgccc	gctcgtac	atggcaaa	cgtcggactt	ggccctggac	5160
ctatggact	tatggggccca	agaggaccac	cagggtctgc	aggcgttcc	ggaccacaa	5220
gatttcaagg	accaggttgc	gagcttgg	aaccttggca	aaacgggtct	gtctgtgtca	5280
gaggaccagg	cgggccaccc	ggaaaatgt	cgcaagatgg	gcacccctgg	aagcttgcgt	5340
gaccggcga	aaagggtgtt	tttggaccc	aaggcgccag	aggcttctc	ggaacacactg	5400
gactgtgttcc	cttcaaggcc	atcaggagg	acaatggct	ggacggactg	aaaggacaa	5460
ctgggtgtcc	ttggctggaaa	ggcgaaccc	cgccaccc	cgaaaatgtc	acaccaggac	5520
aaaccggcgc	aaggaggactt	cctggcgaga	gaggaaatgt	tggagccca	ggtcaggac	5580
gcmcacgg	atctgtatgg	tctgtggac	ctgttggcc	tgccggact	attgttgcgt	5640
ctggccctcc	ttggatgttcc	ggccgttcc	ggccaaagg	cgaaaatgg	gtctgtggaa	5700
acgcggacc	tgccaggccca	ttggaccc	ggggagaatgt	tggattgc	ggactgtac	5760
gaccaggttt	gcacccagg	aatcttgcgt	ccaaatggact	gacaggccgt	aaagggtcgt	5820
ctggccttcc	aggcgttgc	ggtgcacc	gactgtcc	accaggagg	atccctgtc	5880
ctgttggagc	tgctggcgct	acgggttgc	gaggactgt	tggagaaact	gggcacccgg	5940
gatctaaagg	cgagtgttgc	aacaagggt	agccaggat	tgtgttgc	caaggccgc	6000
ctggaccatc	aggcgttgc	ggcaaa	cccttgcgt	cgaaaggccgt	agtgtccggc	6060
ctctctggacc	accaggcc	agaggat	ctggcttgc	aggattgc	ggcgttgc	6120
gtatgttgc	cgatgttgc	ccatgttgc	caagaggcgt	ttcttgcgt	gctggcgtt	6180
gagggttccaa	ttggcgatgt	ggcgttgc	ggggccgg	tcttgcgt	tcttgcgt	6240
tgccctggaa	cccttgcgt	atcggccat	ggccaaagg	aggccctgtt	ggacttccct	6300
gcatcgacgg	tagatgttgc	ccaaatggac	ccgcaggcgt	tagggagag	ccttggaaata	6360
ttggcttccc	ttggcttcaa	ggccccacag	gcatgttgc	aaagaacggc	gataagg	6420
atgtgttgc	cgatgttgc	aggggaggat	ctggacc	ccggaaatgt	gtgttgc	6480
ggccgcgttgg	cccaaaatgt	gttcaagg	ggaaaggcg	gcaggccca	ccttggcctc	6540
caggcttcc	aggacttcc	ggaccat	ggccgttgc	cgaaatgttgc	cccttgcgt	6600
aaagaggact	gcacggcg	tttggcttgc	cggttgc	cggttccac	ggagagag	6660
gccctccagg	cgatgttgc	ctacttgc	tatcgg	aaaggttgc	agaggacta	6720
gttgcgttcc	aggacttgc	ggcaacaa	cgcaaccc	tgttgcgt	gtgttgc	6780
cagcttggacc	tttggatgt	ctggatgt	ccggcgac	cggttgc	ggatgttgc	6840
gttgcgttcc	cgaaaatgg	cgatgttgc	gtatgttgc	cccttgcgt	ggatgttgc	6900
atggcgat	aggcgttgc	gggttgc	gttgcgttgc	ccgttgc	ggatgttgc	6960
acagaggcgt	ggcttgc	gcttgc	ccggccac	cggttgc	ggatgttgc	7020
gcttgcgttcc	tttttttttt	ccggccac	cggttgc	ccggccac	ggatgttgc	7080
ctgttgcgttcc	tttttttttt	ccggccac	cggttgc	ccggccac	ggatgttgc	7140
tcgttggccccc	tacttgc	gttggatgt	ccggccac	cggttgc	ccggccac	7200

-continued

gaccagctgg	gtcttagaggc	gacggcgccgac	cgccttggaaat	gacaggtttt	ccaggcgccg	7260
ctggaagaac	aggctctcca	ggaccatctg	gcattcttg	tccaccaggg	ccaccttgg	7320
ctgtggaa	agaaggactg	agaggcccta	ggggogatca	gggttcagg	ggaagaaccc	7380
cgagaatcg	agctgtcgcc	ccaccaagg	ttgcggcg	aaaaggccct	acggcggaa	7440
ctggaactgc	aggaccgcgc	ggaactcccg	gtctcaagg	attgttggc	gccccggaa	7500
ttctgggact	gcccggtagt	cgcggagaa	gtggactccc	aggtgttgc	ggccggctcg	7560
gagaacccgg	accacttgg	atgtgttgc	caccgttgc	aaaggttca	cctgttgc	7620
ttggaagtcc	ttggcgttaac	ggtgc	gccaaggccg	cagatgttga	aatccggca	7680
atgtatggcc	gcctgggaga	gatggacac	ctggacataa	gggcgagcga	ggcttacccag	7740
gcaatattgg	acctgtcg	gcagccgt	ctccggacc	tcatgttca	gtcggtcc	7800
ccggggaa	cggaaaatagg	gggaaacac	gaccctccgg	tccgttggc	ccatgttgc	7860
cattttggacc	aaggccca	tccggac	agggatcc	cgggatataa	ggcaaccc	7920
gcgagaaggg	acctaaggaga	ctggcttggc	tgaaggccca	taacggactg	caaggcc	7980
caggcattgc	ttggccatcat	ggcgatea	gtgcacccgg	tagtgtgg	cccgccggac	8040
cgagggggtc	cgctgttca	ttgtggacc	ccggaaaaa	ttggcaca	ggacatcc	8100
gcacatgggg	gcctggccg	atttagaggc	caacggaca	taaggcccc	gttggggccg	8160
caggac	gggacccca	ggggcacc	gcgttagtgg	cgaggat	gatttgcgt	8220
acgacggcg	cttctacaga	gccgacc	ctagatctc	ccctagc	aggcttaagg	8280
actacgaatg	ggacggccca	ctgaatgg	tgaacaacc	gatcggagaca	ctgtcgacc	8340
ctggggcc	cagaagaa	cctggccgaa	cctgggg	cctggacty	tctacc	8400
aatggtctc	cggctactac	tggatcg	ccatccatgg	tcgcacat	gaccatca	8460
aggtgtactg	cgacttc	accggcgaga	catgc	agccca	gagaacatcc	8520
ccgccaagaa	ctggatcaga	agcgcaagg	aca	aaacaa	cgttgttgc	8580
tcacatggcc	cggccatgt	tgatcaca	ttggaa	ggcg	gagatggcc	8640
cacatgttgc	tttcatgaga	ctgtggcc	attaccc	ccaa	caacatc	8700
geaagaacag	cattgttac	atggacgagg	aaacccgg	cctgaa	ggcggtatcc	8760
tgccagg	taacgacgt	gaacttgg	ccgggg	cacgac	atccatcaccc	8820
tgctgttgc	cggctgc	aaaaagacca	acggatgg	caagacc	atcgagttata	8880
agaccaacaa	gcccac	atcgcc	tggatatc	ccact	ccatggat	8940
ccgaccaaga	gttcttgc	gacatcg	ccgtgtt	caatgt	ggcg	8987

-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	caacctgggt	ttcagtataa	catcaactatc	2640
tatgctgtgg	aagaaaatca	agaaaatgtca	cctgtgtc	ttaacaaca	aaaccacttgc	2700
acccccacgct	cagatacagt	gecctctcc	aggggacctgc	agtttgcga	agtgcacag	2760
gtgaagggtca	ccatcatgtg	gacacgcgt	gagagtgc	tgaccggct	ccgtgtggat	2820
gtgtatccccg	tcaacctgdc	tggegagac	gggcagagg	tgccccatc	caggaacacc	2880
tttgcagaag	tcacccgggt	gtccccctgg	gtcacctat	acttcaaaatg	ctttgcagt	2940
agccatcggtt	ggggagggaa	gecctgtact	gtcaaacaga	caaccaaact	ggatgtccc	3000
actaacctcc	agtttgtca	tgaaactgtat	tctactgttc	ttgtgtgatg	gactccac	3060
cggggcccaga	taacaggata	ccgactgacc	gtgggccta	cccgaaagagg	acagccagg	3120
cagtacatgt	tgggtccctc	tggttccca	tacccactgt	ggaatctgc	gcctgcat	3180
gagttacaccg	tatccctgt	ggccataaa	ggcaaccaaa	agggccccaa	agccactgt	3240
gttcttacca	cactgcagcc	tgggactct	atccaccc	acaacaccg	ggtactqag	3300
accaccattg	tgatcacatg	gacgcctgt	ccaagaatttgc	gttttaagct	gggtgtacga	3360
ccaaaggccagg	gaggagaggc	accacaggaa	gtgacttc	actcaggaa	categttg	3420
tccggcttgc	tccggaggat	agaatacgtc	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
cttggagata	atgtcagtgt	tttactatgc	aaggatgaca	aggaaatgt	cccttatct	3780
gataccatca	tcccagaggt	gccccactc	actgaccta	gtttgttga	tataaccat	3840
taaagcatcg	gtctgggttgc	gaccccgct	aactcttcc	ccatttattgg	gtacccatc	3900
acagtagttg	cgcccgaggaga	aggtatccat	attttgcga	attttgcga	cttccca	3960
ggatactaca	cagtacagg	gtctggagcc	ggcattgtact	atgatatacg	cgttatca	4020
cttattaatg	ggggcgagag	tgcctctact	actatgcac	aacaaacggc	tgttctct	4080
cccaactgacc	tgcgatttac	caacattgtt	ccagacacca	tgcgtgtc	ctgggcttca	4140
cccccattatc	ttttttatca	caacttgcgt	gtgcgttact	cacctgtgaa	aatatggaa	4200
gtatgttgc	agttgtcaat	ttctcttca	gacaatgc	tgttcttac	aatatctct	4260
ccttggtagac	aatatgttagt	gagttgtctc	agttgttca	aaaaatgc	gagcacacc	4320
cttagagaa	gacagaaaaac	agggtttat	tccccaactg	gcattgtactt	ttctgtat	4380
actggccat	cttttactgt	gtactggat	gtcttgc	ccacccatc	ttggctacagg	4440
atccggccat	atcccgagca	tttcgttggg	agacccatc	ttggcttgc	gccccact	4500
cggaatttca	tcacccctac	caacccatc	ccaggcagac	agttatgtgt	cagcatcg	4560
gctcttaatg	gcagagagga	aagttccat	tttgcgttgc	aaaaatcaac	agtttctgt	4620
gttcccgagg	acctgttgc	ttgttgcgt	accccccacca	gcctactgt	cagctgggat	4680
gctctctgt	gttgcgttgc	attatcagg	atccatc	gagagacagg	aggaaatag	4740
cctgtccagg	agtttactgt	gtcctgggg	aaatgttac	ctaccatc	cgccctttaa	4800
cctggagttg	attataccat	cactgtgtat	gtctgtact	gtcgttgc	cgccgttgc	4860
agdagcaacg	caatttccat	taatttccca	acagaaatttgc	aaaaaccatc	ccagatcg	4920
gtgaccgtat	ttttagggca	cagtttgc	gtcaatgtgc	tgcgttca	ttccctgtt	4980
actgggttca	gagtttaccc	cactccaaa	aatggccac	gaccaacaaa	aactaaat	5040
gcagggtccag	atcaaaacaga	aatgtactt	gaaggcttc	agccccacgt	ggagtatgt	5100
gttagtgcgt	atgtcgtat	tcccaatgc	gttgcgttgc	gttgcgttgc	gactgcgt	5160
accacccatc	atgccttca	aggactgttgc	tttactgtat	tttgcgttgc	tttgcgttgc	5220
attgttgcgg	aaagccccac	ggggcaatgtt	tccaggatca	gggtgtact	ctcgagcc	5280
gaggatggaa	tccatgtat	atccctgc	cctgtatgttgc	aagaagac	tgcagatgt	5340
caaggccatca	gacccgggtt	tgatcaca	gtcgttgc	ttgccttgc	cgatgtat	5400
gagggccagg	ccatgtatgg	aaccctgtt	acagcttac	ctgcacca	tgatgttgc	5460
tttacttgc	tcacccccc	aacctgttgc	gccccatgtt	caccacccaa	ttgttgcgt	5520
actggatata	gagtgcgggt	gaccccaag	gagaagac	gaccaatgaa	agaaaatcaac	5580
cttgcctctg	acagcttccat	cgtgttgc	tcaggacttgc	ttgtggccac	caaataatgaa	5640
gtgatgttgc	atgtctttaa	ggatgttgc	acaaagac	cagettcagg	atgttgc	5700
actctggaga	atgtcagcc	accaagaagg	gtctgttgc	cagatgttgc	tgagaccacc	5760
atcaccatca	gttgcgttgc	caagactgtat	acgtatc	gttgcgttgc	tgatgttgc	5820
ccagccaaat	gcacacttgc	aatccagaga	accatcaac	cagatgttgc	aaatgttgc	5880
atcacaggat	tacaaccagg	cactgtacttgc	aaatgttgc	tgttgcac	aatatgttgc	5940
gtctccggat	ccccctgtgt	tcacccatc	tcacccatc	tttgcgttgc	atccaaact	6000
cgtttccctgg	ccaccacacc	caatttcttgc	ctgttgc	ggcagccgg	acgtgtcc	6060
attaccggct	acatcatca	gtatgttgc	cctgggttgc	ttccca	agggttgc	6120
cgcccccgg	ctgggttgc	agggcttgc	attactgttgc	tttgcgttgc	tttgcgttgc	6180
acaattttat	tcatttgcct	gaagaataat	cagaagac	agccccatgt	ttttgcgttgc	6240
aagacagac	agctttccca	actgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6300
atcttggat	tttgcgttgc	agttaaaatg	accccttgc	tttgcgttgc	tttgcgttgc	6360
actggaaatg	gtatttgcgt	tcctggtact	tctggtgc	tttgcgttgc	tttgcgttgc	6420
atgtatgttgc	aggaacatgg	tttttagggc	accacccatc	ccacaac	tttgcgttgc	6480
aggcatggc	caagaccata	cccgccgtat	gtatgttgc	aaatccaaat	tttgcgttgc	6540
cccaaggaaat	atgtatgttgc	tcacccatc	ccacac	tttgcgttgc	tttgcgttgc	6600
tctacaggac	aagaagatgt	tccttgc	accatcttgc	tttgcgttgc	tttgcgttgc	6660
tctgagtttca	tcatttgcgt	tcatttgcgt	tttgcgttgc	tttgcgttgc	tttgcgttgc	6720
gttccctggaa	tttgcgttgc	acaggcttgc	ccaggggttgc	cacccatc	tttgcgttgc	6780
atcatatgttgc	aggcacttgc	agaccac	tttgcgttgc	tttgcgttgc	tttgcgttgc	6840
gttggccaaat	ctgtcaac	aggcttgc	caacccatc	tttgcgttgc	tttgcgttgc	6900

-continued

-continued

gaaaaggatg	ccctatcgt	gaacaagggtg	gtcacccctc	tgagcccacc	aacaaacctg	3540
cacctggaa	ccaatctcg	tacccgcgtg	ctgactgtgt	cctggaaag	aagcaccaca	3600
cctgacatta	ccggctatcg	gatcaccaca	acacccacca	atggccagca	gggcaactcc	3660
cttgaagagg	tggtgcatgc	cgatcgtcc	agctgtactt	tcgacaact	gagccctggc	3720
ctcgagtaca	atgtgtccgt	gtacacagt	aaggacgaca	aagaaaagct	gcccatctcc	3780
gacacaatca	tccctgaggt	gccccagctg	accgacccgt	gcttcgttga	tatcaccgac	3840
agcagcatcg	gcctggaggt	gacacctctg	aactctctg	ccatcatcg	atacagaatc	3900
accgtgggtt	ccgtggcgta	gggatccca	atcttcgagg	actttgttga	cagcagctgt	3960
ggctactaca	ccgtgacttgg	actggaaacc	ggcatcgact	acgacatcg	cgtcatcacc	4020
ctgtatcaatg	ggggcggagat	cgccccatca	acactgacac	aacagactgc	cgtgcttct	4080
cataccgatc	tccgggttac	aaatactggc	ccggacacc	tgagagtgac	ttgggtctt	4140
ccaccaagta	tccgacccgt	caacttcc	gtcggtaca	gccccgtgaa	gaacgaggaa	4200
gatgtggccg	agctgtccat	ctctcccago	gataatgcgg	tggtgtgtac	caatctgtc	4260
ccccggaaac	agtagctgtt	gtccgttago	agcgtgtac	aacagacaca	gagcacaccc	4320
ctgcggggca	gacaaaaaaac	aggttggat	ageccccacc	gaatcgact	cagcgatatc	4380
acagccaaac	gttaccctgt	gcattggatc	gccccctaa	ccacccat	cggatata	4440
atccggcata	accccgagca	cttcagccgc	agacctagag	aagatagagt	gccccactcc	4500
agaaacacgca	taccccttcac	caatctcg	cccgccaccc	aatatgttgt	gtccatctgt	4560
gcccgttac	gcagggaaaga	gtctccctcg	ctgatcgcc	agcagacac	agtgtctgac	4620
gtgcccagat	atctggaaat	ggttgcgtcc	acacatccaa	gctctgtat	tagctggat	4680
gtctctgtgt	tgacagtgcg	gtattaccgg	atcacctac	gctgagactgg	cgccaaactct	4740
cccggtcaag	agtttacatgt	gccttgcgc	aagagcaccc	ctaccatctc	tggactcaag	4800
ccaggcgttc	actacaccat	tacccgtgt	gctgtgacc	gcaggggcga	ttctctgtcc	4860
tcttctaacc	ctatcgtat	caactatccg	accggatcg	acaagccaa	ccagatgcaa	4920
gtgacagatg	tgcaggacaa	cageatctcc	gtgaagtggc	tgcttagcag	ctctccatgt	4980
accgggtaca	gagtgtaccac	cacaccaaa	aacggccctg	gacctacca	gaccaagacc	5040
gtctggatcg	atcagacccg	gtatggaccat	gaggggctgc	agcctaccgt	cgagatgtc	5100
gtgtctgtgt	acggcccaaa	tccttgcgg	gagtctcgc	ctcttgcga	gaccggctg	5160
accaacatcg	acagacactaa	aggccctggc	ttcacccgac	tggaatgtga	ctctatcaag	5220
atccgcctgg	agttccctca	ggggcaatgt	tccagatata	gagtgtacta	cagctccct	5280
gaggacggca	tccacacgt	gtttccacgg	ccagacccgg	aagaggatac	cgccgaactg	5340
caaggactcg	ggccctggctc	cgatataca	gtcttcgttgc	tggccctgc	cgacgacatg	5400
gaaaggccac	ctctgtatcg	aacccatgtc	accgctattc	ccgctcttac	agacctgaa	5460
tttaccccaag	tgaccccaac	cagctgtgc	gcacaatgg	ctccctccaa	cgtccatgt	5520
actgggttata	gagtgtcg	gacacccaaa	gaaaagactg	gccccatggaa	ggaaatcaat	5580
ctggccctgt	actccagtc	cgtgggttgc	tctgttctta	tgttggatc	caaatacggag	5640
gtttccgtgt	atgcccgtaa	ggacacccctg	accccttgcac	ctgcacagg	cgttgtgacc	5700
acactggaaa	acgtgtcccc	acccctggaga	gccagagtga	cagacgcac	cgaaaccaca	5760
atcaccat	cttggggggc	caagacacag	acaatccac	gattccagg	cgaccccgctg	5820
cctgccaatg	gacagacacc	tatccggcgg	accatcaac	ctgacgtgc	gagctacaca	5880
atcacagggc	tgcaacccgt	caccgactc	aaatctacc	tgtacaccc	gaacgacaa	5940
gcccgtct	gccccctgg	tatcgatgc	tctaccgc	tcgacgc	ccggccatcg	6000
agatttctgg	ccacacactt	caacgtgt	ctctgtgtt	ggcagccccc	tccggccaga	6060
atcaactgtgt	atcatctaa	gtacggaaag	ccaggcagcc	ctcttagaga	gggtgtccct	6120
agacctagac	cttgggttcac	agggccaca	attacccggc	tcgagccccc	cactgagatc	6180
acaatctacg	tgatgtccct	gaagaacaa	cagaatgtcc	agccactgt	tggccggaa	6240
aaaacccgac	acgtgtctca	gtctgttgc	ctgctctatc	caaatactgc	cgccccggag	6300
atccctggat	tcccatctac	ctgtggat	accccatttgc	tgacacaccc	cggttacgac	6360
acccggaaac	gtatttacgt	gtctggaaac	tctgggcac	agccttctgt	gggacacccg	6420
atgatcttt	aggaacacgg	tttccggcgg	accacaccc	ctacaac	cacaccaatc	6480
aggcacaac	cccgccctta	tccttctaa	gttggcgaag	agatccat	cgggacacatc	6540
ccctggatgg	atgtcgacta	ccacctgtt	cctcacccgg	caggactgaa	ccctatgtcc	6600
acgacacggac	aaagggccct	gagccagacc	actatccgt	gggcttccat	ccaggacacc	6660
tccgagatca	tcatcttctt	tcacccctgt	ggcaccgac	aggaacacc	gcaattcaga	6720
gttccggcga	cttcattac	ccggccacact	acaggacttg	ctagaggcgc	cacctataac	6780
atcatctgtt	aaagccctgg	agaccac	ccggcaca	tcgatgttgc	gggtgtact	6840
gttggcaat	ccgttacac	ggggctgtat	caagccaccc	acgacacgt	ctttgacccc	6900
tacacagtgt	cccaactatgc	cgtggggat	gagtgggaa	gcatgtccg	gagccgttcc	6960
aaactgtct	gtcgtgtcc	ccgggtttcc	tcggccact	tcagatgtc	tagtccagg	7020
tggtgcac	ataacccgt	gaacttacaa	atcgagag	atggggac	acagggcag	7080
aaacggccaga	tgtatgtct	ccatctctc	ggaaatgtt	ggggcaggat	caagtgcgc	7140
cctcaca	ccacccgtt	cgacgttgc	aaacccat	acgtgggg	gcagtggcag	7200
aaagagtacc	tggggccat	ctgttagct	acatgttgc	tcgggcac	cggtggcgc	7260
tgtgtat	gcgtgtacc	ccatctcc	agggaacaa	aggccat	7320	
tacaaccat	acagccac	ataccacc	cgccacca	ccatgtt	ctgcccattc	7380
gagtgttca	tccggccat	cgccggcc	gacacggaa	attcttagaga	gtga	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	ccggagatcc	tcctgttct	gctgtccatc	60
ctccacccct	ctcggccat	aggggccat	ggggccat	ctgggtggat	tcctggagga	120
gttctttatc	caggggctgg	tctggagac	cttggagg	gagcgttgg	gcctggaggc	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	tcctcggtgt	tggaggggct	840
ggtgttcctg	ggtgtctgg	ggcaattctt	ggaaatggag	gcatcgagg	cgttggact	900
ccagctcgag	ctgcagctgc	agcagcagec	gctaaggcg	ccaagttatgg	agctgctgca	960
ggcttagtgc	ctgggtggcc	aggctttgc	ccggggatag	ttgggtgtcc	aggagctggc	1020
gttccagggt	ttgggtgcc	aggagatgg	atcccgatgg	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	gggttcccg	gagttggcat	tcccccggaa	1320
gctcaggcag	cagtcggcc	caaggctcc	aagtacgggt	ctgcaggagc	aggagtgtc	1380
ggtgggtctag	tgccagggtc	cccaggcgc	gtccccagggt	tgccggggac	gggaggagtg	1440
ccaggagttgg	ggaccccaagg	agctcgatgt	gctaaaggcg	ccgccaaggc	cgcccaagg	1500
gggttagtgt	ctgggtgtgg	cgtggctctt	ggagttggcg	tgcgtcttgg	tgcgtgtgt	1560
gtccctggag	ttgggttggc	tcctggggat	ggcggtggct	ctggagttgg	tgtggctcc	1620
ggcggtggcg	ttggctccgg	cattggccct	ggtggagttg	cagtcgcagc	aaaatccgt	1680
gccaagggtgg	ctgccaatagg	ccagtcggca	gctgcagtc	ggcttgggtc	tggcatcc	1740
ggactctggag	ttgggtgtgg	cgtccctgtt	cttggagttt	tggtctgtgt	tcctggactt	1800
ggagttttgt	ttgggtgtcc	tgggttgggg	cgagggtcg	ataggggagt	taggcggagc	1860
ctgtcccttg	agctcaggga	aggagatccc	tccctcttcc	agcacccccc	cagcacccccc	1920
tcatcaccca	gggttacctgg	agccctggct	gcccgtaaag	cagccaaata	tggagcagca	1980
gtgcctgggg	ttcttgggg	gtctgggggt	ctcggtggag	taggcattcc	aggcggtgt	2040
gtggggagcc	gaccggccgc	cgccgtcgcc	gcagccaaag	ctgcgtccaa	agccggccag	2100
tttggcttag	ttggggccgc	tgggttcgga	ggactcgagg	tcggagggtt	tggagttcca	2160
gggtgttgggg	gccttggagg	tataacctca	gctgcagcc	ctaaaggcagc	taaatacgg	2220
gctgctggcc	ttggagggtt	cctatgggggt	gcggggcagt	tcccacttgg	aggagtggca	2280
gcaagacgtt	gtctccatt	ttcccaagggt	ggccgtcgct	ggggaaagct	a	2340
tgtggccgga	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	tgtgtctgtc	cctgagcatt	60
ctgcaccctt	ctagaccagg	cggagtgc	cttgcatttc	ctggggaggt	tccgggggca	120
gtttttatc	ctgggtgtgg	actggggcc	cttggcgag	gtgtcttgg	acctgggtgg	180
aaggctctga	aacctgtgcc	tggggactt	gctgggtcg	gtcttggagc	cggaacttgg	240
gtttttcccg	ccgtgacatt	tcctggcgct	cttgggtcgat	gtggcggtgg	agatgtgtcc	300
gctgttata	aggctggcaa	agccggcgct	ggactcgagg	gtgttcttgg	tgttggaggc	360
ctggggattt	ctgctgtgtt	tgtgttctt	caacctggcc	cttgggtttaa	gcctggcaaa	420
gttccaggcg	ttggactgcc	tgggtttac	cctgggttgtt	tttccttgg	cgtagattc	480
cctggcggtt	gagttttcc	tggcggtcc	acagggtccg	gctgtaaacc	taaagctcc	540
gggtgtcgcc	gagcccttgc	tggaaatacc	ggcggtgggg	cttttgggg	acctcaacca	600
gggtgtccctc	ttgggtatcc	tataaggcc	ctaaactgc	caggcggtca	cgccgtcc	660
tacacaactg	gcaagctgc	ctatggctat	ggacctggcc	gtgttcttgg	cgccgtcgga	720
aaagctgtat	accctactgg	aacaggcg	ggccctcaag	cagcagtc	cgccaggagct	780
aaaggcccg	ctaaatttgg	agctggcg	gcaggcg	tgcctgggt	tggcgagca	840
ggcgctccagg	gtgtttccagg	gtcaacttcc	ggaaatcgag	gaatttggcc	cgctcgaa	900
ccagctcgag	cagccgcagc	cgccgtcgcc	gcaaaagct	ctaaatatgg	cgccgtcgca	960
ggtttgggtc	caggcgccacc	tggatttgg	ccagggttgtt	ttggagtggc	aggcgctggc	1020
gtcccccggc	ttggggatcc	tgggttggcc	atccctgtt	ttccctggcc	ttgtatttcc	1080
ggcgctgtgt	ttccagggtt	tccgtcttcc	gaagccgtgt	ccaaggccgc	tgctaaaggca	1140
gtctaaatcc	gtggccggacc	aggcggtgg	gttggcgaa	tcccaacata	tgggtgtcg	1200
ggcgccggat	tcccaggatt	tggagttgg	gtcgaggc	tcccgggtgt	tgcaaggcgtt	1260
ccaggcgctg	ttggggatcc	tggcggttgt	gggtgttcc	gctgtggaaat	ttctcttgc	1320
gtctcaggcc	ctggcgctgc	caaaggcgcc	aaatatggcc	ctggcggtgc	ttggcggttc	1380
ggaggattgg	ttccgggtgc	tccagggtgt	gttccgggg	ttcccgaaac	ttggcggtgt	1440
cctgggtgtcg	tgacaccaggc	cgctcgac	gcaaaaggct	ctgtcaac	cgctcgagg	1500
ggactgtgtt	ctggcggtgg	agtggccaca	ggtgggtgg	ttgcaccc	cggtggcg	1560
ggccctggcg	ttgggttgc	tcctgggttt	gggtgttccc	cagggtcg	agtcgtctcc	1620
gggtgtcggt	tcgcacccgtt	tattggctt	ggtggcggtt	ccgcagtc	caaatctgt	1680
gcaaaagggtgg	ccgccaataggc	acagctgaga	gctgtcc	gttttggcc	tggaaatcc	1740
ggactcggtgt	ttggggatcc	gttgggttgtt	gggtgttccc	gagggtcg	gccccggactc	1800

-continued

ggagtccggag	ctgggttccc	aggtttgga	gctgggtca	atgaaggcgt	gcggagatct	1860
ctgagccctg	actgtccat	gggcatatcc	agctctatcc	agcatatccc	tagcacac	1920
agcagcccta	gagtccttgg	ggcttgcgt	gcagccaa	ccgctaa	tggggctgt	1980
gtccccgtt	ttcttggagg	acttggatca	ctggcgag	tgggaattcc	aggcggtgc	2040
gttgggtc	gacactgtc	tgccggcga	gctgcaaa	cagcagctaa	ggcagccag	2100
tttggcttgc	ttggagccgc	tggacttgg	ggctctgg	ttgggtgt	tggtgtcc	2160
gggtgttggc	gacttggagg	aattccacca	gccgcag	ccaaggctgc	taaataccgc	2220
gtgtcaggac	tttggcggt	tcttggcg	gtggacat	tccacttgg	aggcggtgc	2280
gccagacgt	gttttgcctt	gtcttctt	tttccctgg	gctgtgt	ggccaaagcc	2340
tcggcgacaa	aa	aa	aa	aa	aa	2361

SEQ ID NO: 39	moltype = DNA length = 1017					
FEATURE	Location/Qualifiers					
source	1..1017					
	mol_type = unassigned DNA					
	organism = Homo sapiens					
SEQUENCE: 39						
atgagtcata	gtgcatttac	tcttttctgt	gcatttgcatt	gtggtaacc	tggccat	60
atgattat	atttccct	atcaatttat	gggcaatcat	caccaact	tgcacc	120
tgttaactgc	ctgaaagcta	cccaagtcc	atgtactgt	atgagctgaa	attgaaa	180
gttaccaatgg	tgccctctgg	aatcaagtat	cttacat	ggaataacca	gattgacc	240
attgtat	aaaaggcttgc	gaatgtat	gatctgc	ggcttattt	agatcac	300
cttctat	acttccat	aaaagggaga	gttttcttca	aatttcaaca	actgaaga	360
ctgcatat	accacaa	cctgacag	tctgtggcc	cacttccca	atctctgg	420
gatctgc	ttactcata	caagatcaca	aagctggc	cttttgaagg	attgttac	480
ctgacat	tccatcttca	gcacaatcg	ctgaaagagg	atgtgtt	agctgtt	540
aaaggcttca	atacactcg	atacatgt	tttgcgttca	atcagat	cgactgc	600
tcttgcgttcc	ctgtctct	tctacttct	tacttgc	acaat	cagcaacat	660
cctgtat	atttcaagcg	ttttaatgc	tttgcgtt	tgcgtt	tcaacac	720
ctggctgata	gtggaaattt	ttaatgt	catccctgt	tgactgt	gtt	780
ctgtccata	acaagat	aaacatcca	actgtca	aaaac	tttcaacttac	840
ctggagggtca	atcaacttgc	aaatggat	tctgcgtt	tttgcgtt	ccctgggg	900
ttatccat	ccaatgttca	gcatttgc	tttgcgtt	atcgtat	agaaa	960
cttccacccg	atatgtat	atgttac	gttgcgtt	atgttac	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						
atgagcctga	gcgccttca	actgtttctg	gcctgtat	ggggcaca	cgccag	60
tacgactac	atttccact	gagcatc	ggccag	ccctaa	tgc	120
tgcaattgc	ccgagat	tccttag	atgtact	acgag	gtc	180
gtgcctatgg	tgccctctgg	catcaat	ctgttactt	ggaaac	atcg	240
atcgacg	aggcccttgc	aaacgt	gac	tttgc	atct	300
ctgctggaa	acagca	aa	ggcc	gttgc	atc	360
ctgcacat	accacaa	cctgac	agcgt	ggcc	cttgc	420
gatctgc	tgacccat	caat	actgtgg	ggcc	cctcgt	480
ctgacat	tccatcttca	gcacaac	tttgcgtt	atgcgtt	ccccc	540
aaggcc	agatgttca	atacc	ctgat	tttgcgtt	atc	600
agccgc	ctgttctt	gttgc	actgt	tac	ccat	660
cccgac	acttcaagcg	gttcaac	ctgcgtt	tgcgtt	ccacaac	720
ctggcc	ctggcat	ccgcaac	ttcaat	ccat	ccat	780
ctgtccat	acaagat	aaacat	accgtt	tttgcgtt	ggaaact	840
ctggaa	accgttca	aaatggat	tttgcgtt	atcaag	tttgcgtt	900
ctgcct	atatgtat	gttgcgtt	gttgcgtt	atgttac	aaatggat	960
						1017

SEQ ID NO: 41	moltype = DNA length = 1437					
FEATURE	Location/Qualifiers					
source	1..1437					
	mol_type = unassigned DNA					
	organism = Homo sapiens					
SEQUENCE: 41						
atggcacccc	ttagacccct	tctcata	gcctgttgc	atgggttgc	tctgg	60
caagagat	gcaaggcc	ctgcact	ggcttca	ttgg	caac	120
gacgagat	ctgttacta	ccagat	tgac	atac	ggcc	180
caagt	gcggggat	gttact	ccggagg	atgtac	gttgc	240
ggcgagg	aaaacat	cactgtt	tttgcgtt	tttgcgtt	tttgcgtt	300
gac	cccaat	tttgcgtt	tttgcgtt	tttgcgtt	tttgcgtt	360
gaggcc	cccaat	tttgcgtt	tttgcgtt	tttgcgtt	tttgcgtt	420
acc	tttgcgtt	tttgcgtt	tttgcgtt	tttgcgtt	tttgcgtt	480
ttc	tttgcgtt	tttgcgtt	tttgcgtt	tttgcgtt	tttgcgtt	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 agtacaccgtt gtacgacgac 240
ggcgaggaaa agaacaacgc caccgtgcac gggcaatctt gggcggccatc tctgtgggg 300
gatctgcagg ctccatggcgtt gggcaatctt gggcggccatc tctgtgggg 360
gaaggccctt ctccatggcgtt gggcaatctt aacgctgtggat gcaatggccatc cagacccgg 420
aacatgtcatc caggcggacc tcacgttctt gggcggccatc aactgtgttag cggcaaccc 480
ttcgacgcctt tcacccgacccatgg gaaacccggc agcctgttcg ccttttaggg ccagtactgc 540
tacgagatgtt acggaaaggc cgtggggctt ggatatcttca agtgcgttca agatgtgtgg 600
ggatcgaggcc gcccatttcgc cccggccatcc accagaatccatc actgtcagggg caagacccatc 660
ctgttcaagg gggccatggcgtt ttggatggatcc gggccatggcgtt tgactacccc 720
agaatataatc gggccatggcgtt cggccatccatgg gggccatggcgtt tgactacccc 780
cctgctcaca gttactctgg cggccatggcgtt ttaaggggaa acagtactgg 840
gagtaccatg tccagcacca gcccggccatgg gggccatggcgtt tctgtgggg 900
gtgtttggcc acttgcattt gatgcagccg gacatgtgggg aatattttt cgagctgtgt 960
ttctggggca gaaccacggcc cggccatggcgtt cggccatggcgtt tgactacccc 1020
ggcggtccatgg gacaatgtggc tggccatggcgtt gggccatggcgtt tgactacccc 1080
cttagccatgg gggccatggcgtt tccggccatgg gggccatggcgtt tgactacccc 1140
agccagagatggcc gggccatggcgtt cggccatggcgtt tgactacccc 1200
ttggctgtcc ttgtttagccatgg gggccatggcgtt cggccatggcgtt tgactacccc 1260
cgatggattt ggcttgcc tggccatggcgtt gggccatggcgtt tgactacccc 1320
ggcgacaaatg attaccggatggcc gggccatggcgtt tggacactgtt ggaccctccc 1380
tacccatggcgtt tccatggcgtt gggccatggcgtt tgactacccc 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 ggtgtggat ctttggggca 60
ggccctggggc atgtgttggg ctacatggcgtt cggccatggcgtt tccatggcgtt 120
cccggtcaatggcgtt gtcacatggcgtt agcggatgtt gttggatggatcc gggccatggcgtt 180
ctggatcatggcgtt taaaggccatggcgtt tataccggatgg ggtgtggatcc ctgcacatggcgtt 240
aacggacatggcgtt caaatggcgtt ccaggatggcgtt tcaggccatgg tggatggatcc gggccatggcgtt 300
accggccatggcgtt aacatggcgtt ccaggatggcgtt tcaggccatgg tggatggatcc gggccatggcgtt 360
tcctggccatggcgtt tccatggcgtt aacatggcgtt ccaggatggcgtt tggatggatcc 420
ggggggatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 480
gcacccggatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 540
aacacggatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 600
cccaaaatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt tggatggatcc 660
ctgaacccatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 720
ctgatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 780
ctggatggatcc tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 840
ggccctggggc atgtgttggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 900
caatggatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 960
gcaaaatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 1020
cagatcttcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 1080
tggctgtccatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 1140
agagaacggatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 1200
caaaaaatggcgtt tccatggcgtt gggccatggcgtt tggatggatcc gggccatggcgtt 1260

```

-continued

gaatacgaag	ccaaactcgag	tgaccttcgt	gctcggctgc	aggaggcgc	tgcggcaagcc	1320
aagcaggcaa	atggcttcaa	ccaaagaaaaac	gagagagcgtt	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	aagagaaaac	ccagcgggaa	tgagctgg	1680
cgtgtgtgt	ttgatgccgc	cacccctac	gagaacatcc	tcaatgcacat	caaagcggcc	1740
gaggacgcag	ccaacaggc	tgccagtca	tctgaatctg	ccctccagac	agtgataaag	1800
gaagatctgc	caagaaaagg	taaaaccctg	agttccaaac	gtgataaact	gttaaatgaa	1860
gccaagatgc	cacaaaaaaa	gctaagcaa	gaagtca	cagetc	caacccatcag	1920
caaacccctga	atattgtac	agttcagaaa	gaagtatag	acaccaatct	cacaactctc	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	acctgtatct	2220
tggcgcaga	ttgaaaagtat	caaccaacag	ctgttgcctt	ttggaaacat	ctctgacaa	2280
atggacagaa	tacgagaact	aattcagcag	gcccagat	ctgcaatgaa	ggtgtctgtc	2340
cccatgggt	ttcatggtaa	atctggatgt	gaaatccgc	tgccaaatgaa	cctggaaagat	2400
ttgaaaggat	atacatctct	gtcttgc	ctccaaaggc	ccaactca	agaaaaatgg	2460
ggtactgaga	atatgttgc	gatgtaccc	ggaataaaag	atgcctcc	ggactacatc	2520
ggcatggcag	tttgtggatgg	ccagtcata	tgtgttaca	acctgggg	ccgtgaggct	2580
gaacttccaa	ttggaggat	cttgacca	agtggacta	aggggcact	tatggatcg	2640
gtgaaatttca	agagaattt	tcagttca	aggcttaat	acaccaagg	agccacatcc	2700
agtaaaccag	aaacaccccg	agtctatgc	atggatgtt	gaaatagc	tacactc	2760
aatttggatc	ctgaaaatgt	tgtatttat	gttggggat	acccaccc	ttttaaactt	2820
cccaatcgac	taatgttccc	tccatcaaaa	ggttattt	aatttagat	cctcaatgaa	2880
aatgttctga	gcttgcataa	cttcaaaaaa	atccatca	tcaacaaac	tgaatgg	2940
ccttgttagaa	ggggaaagga	agagtcagac	aaaaattttt	ttgaaggat	gggtatgt	3000
cgagttccaa	ttcaaccaca	tgctccatc	ccaaaccc	tttgc	gacagacaat	3060
gtggatagag	gttgcgtgtt	cttgcgaa	aacggggatc	gcttcata	tctaaatata	3120
gaagatggca	agctcatgtt	gagataaaaa	ctgaattc	actccat	agagagagga	3180
gttggagacg	ccataaaacaa	cgccagagac	cattcgattc	agatcaaaat	ttggaaactc	3240
caaaggcgt	tgtggataaa	tgtggacgtt	caaaccacta	taattgtat	tgaatgtt	3300
gatttcagca	catattatct	ggggatgtt	ccaaatgc	tcagggaa	atttaaat	3360
tctacgcct	cttccgagg	ctgtatcata	aatttga	aaacccat	tttgcgtt	3420
ttgaatgata	ctgtggggat	aaccaaaaaa	tgctcgaa	acttggaa	tgtgcgtat	3480
gcctcattct	ccagaggagg	acaatttgc	tttactgtt	ttggcttacc	acctactgc	3540
caacctccagg	cctcattttg	atccatcgac	tttcaaccc	gtggcatatt	attatgc	3600
cagacatgg	caaggaaac	gcaggatct	tcggaa	atggatgtt	attggcacc	3660
agcgcatacg	ggggccaaat	ttttaaatct	ccacagat	atatggat	tttgcgtat	3720
tatgtatctg	taataagcga	caactctg	ctacggcttc	tcatcgat	ccagcttct	3780
agaaatagca	aaaggcttca	acatccatca	agttccccc	agttctgtc	tctgggg	3840
agcaattttgc	agggttgc	tatcaatgtt	tttgc	gggttact	gatgttgc	3900
gttcttagat	tgaccgtaa	ctcttc	agagatgtt	ccctgggg	ctgcgtt	3960
aacaaaccac	cttttcaat	gttgc	tttgc	ggtttacca	gaccaagact	4020
tttcgtatca	accatgttgc	gcaggacaca	ccatggcc	cccaacgg	ctgtgaatgt	4080
ttggcaatgt	tttgcgttacc	acttccaa	accaggc	atcatgg	cctccat	4140
ggggacatcc	ccacccatca	cttgcattt	aaatcttcc	aggagctgt	gaaacccagg	4200
tcacagtttgc	ttgtggacat	gcagacaaca	tcctcc	aggactgtt	tcacacgg	4260
actaaacact	cttttatggc	tctttatc	tcaaaaggac	gtctgtt	tgcactgg	4320
acagatgg	aaaatttgc	gatcaaaac	aggaggat	gcaatgtat	ggaaatggc	4380
acgggtgttgc	ttggccat	ttggggaa	ggggctt	tttgcgtt	actgggg	4440
cgggagggg	gttgccttgc	aaactccacc	atcagcat	gagcgc	ccat	4500
tcacccat	caggaaacc	aaagagcc	cccaacaa	gtttgtt	atgcctg	4560
aacttc	cgttgcataa	acccttgc	atccatgtt	caagctt	gggttctt	4620
tgttgcgttgc	ttcccttgc	gaaaggat	tatttctt	agaaggagg	tcatgtc	4680
ttggctact	ctgtatgtt	ggggccat	ttaatgtt	tttgc	ccgcca	4740
agtctactgt	ggatctat	acacatcg	agtc	aggacactt	atgtgtt	4800
ctggaggc	gaaaggatc	ggcctat	gacgttgg	cagggtgg	ctcaacgt	4860
gtcacaccaa	aggatcttct	gttgc	ctgttgc	ccgtggc	aggat	4920
caacacatcc	tgcacccat	actgg	acacat	gacatgtt	acacagctt	4980
ttcccacat	ccagcact	agaggcact	caccc	tttgcgtt	caatgtac	5040
acactgg	ccctctgtgt	gaaatccat	tttgcgtt	tgaggaaat	tcatgt	5100
cacatccctg	ccctctgtc	tgaaccc	tttgcgtt	ggccgtt	ccatgt	5160
tgccctgacc	aggat	tttgcgtt	tttgcgtt	ggccgtt	ccatgt	5175

```

SEQ ID NO: 44      moltype = DNA  length = 5175
FEATURE
misc_feature      Location/Qualifiers
source            1..5175
note = Synthetic Construct
mol_type = other DNA
organism = synthetic construct
SEQUENCE: 44
atgcctcctg ctgtgcggag aagcgcctgt tctatggat ggctgtggat ctttggcgcc 60

```

-continued

gctctggac	agtgtctggg	ctactttct	cagcagcagc	gggtgcatt	tctgcagcca	120
cctggacagt	cttcagtgca	ggccagctac	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	tttggcgct	ctgtccagcc	ttgtccctgc	540
aattctaatt	gccagctggg	ctcttgcac	cctctgacc	gctgactcat	caatcaagag	600
cctaaggaca	gcaagccctgc	cgaggatgc	gacgatttgc	atagtcgt	gtgacccctg	660
ctgaacgaca	ttggccaaat	gggagaacag	ctggggctgg	ttaagagcc	gtctcaggga	720
ctgtctgcct	ctgctgact	gttgcacag	atgcggcaca	tggaaaccca	ggccaaggac	780
ctgagaaacc	agctgtgaa	ctacagaacg	gccccatctca	accacccgg	caagatcgaa	840
ggcctggaaa	gagagctgca	cgccatgtac	caagaggctc	agacactgc	agagaaggcc	900
caagtgaaca	gcccggaggc	ccagactctg	aacaacaacg	tgaacccggc	cacacagtcc	960
gccaaaggaaac	tggacgtgaa	gatcaagaa	gttgcacccg	acgtgcacat	cctgtcaag	1020
cagatcagcg	gcacagatgg	cgaggccaa	aatgtgccta	ggggcgactt	tagcagagag	1080
tgggcccgg	cttcagcgat	gatgcggag	ctgcggaa	ggaaaccttgc	caaggcacctg	1140
agagaaggcc	aggccgacaa	gagagagac	caactgtgc	tcaaccggat	cagaaccctgg	1200
cagaaaaacc	accaggccgg	gaaaacccg	ctggccaa	gcatcagaga	cagctgtaa	1260
gagtaacgagg	ccaagctgag	cgatctgcgg	gcccggactt	aagaagctgc	cgtcaggccc	1320
aagcaggcoca	acggccctaa	tcaagagaa	gagagagccc	tggggccat	ccagagacaa	1380
gtgaaagaga	tcaacagcc	gcaagcgtac	tttaccaatg	accttgcac	cgccgatagc	1440
agccctgtgc	agccatataat	cccccgtac	cttgcggaaa	agagccggaa	agatcagaa	1500
aagctggccg	ccagccgtaa	cgaggccagg	caagaactgt	ctgacaaatgt	gcgcggactg	1560
agcagatctgc	ccggcaagac	atcttgcgt	gaagaggccc	agaagcacgc	cagatctctg	1620
caagagctgg	ccaaacagct	ggaagagat	aagcggaa	ccagcggcga	cgaactctgc	1680
agatgtcgac	tggatcgccg	caccgcctac	gagaacatcc	tgaatgcct	caaggccgccc	1740
gaggacgccc	ctaatagagc	cgcttctgt	tctgagttgc	ccctgcagac	cgttgcataaa	1800
gaggacctgc	ctagaaaggc	caagacacty	agcagcaaca	gcgcacaaact	gctgaatgt	1860
gccaaggatgc	acttgcggaa	acttgcggaa	gagggttccc	ctgcactgaa	caaccctgcag	1920
cagacccctg	acatctgtac	ctgtggaaa	gaagtgtatc	acaccaact	gacaaccctg	1980
agagatggcc	tgcacccgaa	ccagagac	ccatgtcag	cagcgcacaa	2040	
agcatgggtc	gaaaaggccaa	cgacatcaco	gacgagggtc	ttggacggct	gaatccatc	2100
cagacccgac	tggacccggat	caaggacacc	tacggccagaa	cccagaacga	ggatttcaag	2160
aaggccctg	ccgacccgca	caactccgt	aacaagctg	ccaacaagct	gcccgcattc	2220
tggccggaa	tccggagcat	caaccagaa	ctgtccctc	tggcaacat	cagcgcacac	2280
atggacagaa	tccggaaact	gatccagcg	gccagagatg	ccgcctccaa	agtggctgt	2340
cccatagat	tcaacggccaa	gaggcgatgt	gaagtgcggc	tgccccaa	tctgaaagat	2400
ctgaagggtc	ataccagcc	gagctgttgc	ctgcggatgc	ccaacagcg	agagaatggc	2460
ggcaccggaa	atatgttcgt	gatgtacgt	ggaacacaag	acgcacccgg	ggacttatgc	2520
ggaatggccg	ttgtggacgg	ccagctgacc	tgcgtgtaca	acctgggaga	cagagaagct	2580
gaactcgagg	tccggacat	cctgaccaat	agcggagacaa	aaagggccgt	gtggacacaa	2640
gtgaagttcc	acgggatcta	ccatgtcgac	cggtgtaaat	acaccaaggc	gcgcacaa	2700
agaaggcc	aaacacactgg	cgtgtacac	atgcgcggc	ggaaacttca	cactctgt	2760
aatctggacc	ccgagaacat	ggtgtttac	gtggcggt	accctctgt	cttcaagctg	2820
cctagcagac	ttagctcc	acccatcaac	ggctgtatcg	agctgtatgt	cctgaacgaa	2880
aacgtgtcg	cctctgtacaa	cttcaaaa	acccatcaac	tgaacaccac	cgagggtggaa	2940
ccttcgcagg	ccatggaa	ggatccgac	aagaactact	tgcggccac	cggtcagcc	3000
agagtgcata	cacaacactca	cgctccatt	cctacccctc	gccagaccat	ccagacaacc	3060
gtggatagag	gtctgtgtt	cttcggcc	aacggcgac	gattcatc	cctgaatata	3120
gaggatggca	agctgtgtt	cgatgtac	ctgatgtac	agctgtccaa	agaaaaggcc	3180
gttggcggac	ccatcaacaa	cggtggat	cacagcatc	agatcaagat	cgccaaactg	3240
cagaaacgg	tgtggatcaa	cgtggacgt	cagaacacca	tcatcgac	cgagggtttc	3300
gacttcgcagg	cctactatct	cgccggtac	cctatcgcca	tcaagagac	gttcaatata	3360
agccccctg	ccttcgggg	ctgcgtac	aaactgtaaa	agaccaggc	cgtcgtgcgg	3420
ctgaatgata	cagtggcg	gaccaaaag	tgcggcgg	actggaaatgt	tgtgcggagc	3480
ccgagtttt	ctagaggcgg	acagctgac	tttaccgacc	tgggacttcc	tcctaccat	3540
catctgcagg	caagcttcgg	ttccagccaa	ggggaaatct	gctggacccac	3600	
cagacccctg	ccgaaaccc	gcaagtggac	tccggatgt	gctatcatcg	actggacacc	3660
agcgcactctg	ccggccctat	ccttacagaa	cttcagaccat	acatggatgg	gctgtgcac	3720
tacgtgtcc	tgatcagcga	taacagccgc	ctgagactgc	tgatcagcga	ccagctctg	3780
cggaacacgca	agccgctgaa	gcacatctcc	agcggcagac	agagctgag	actccggcc	3840
agcaatttc	agggtgtat	cagcaacat	tgcgtgcac	ccctgcagg	3900	
gtgtgtggacc	tggaccat	tagctgtac	agggtgtgt	cttcggccgg	ctgtccctg	3960
aacaaaccc	cttccctgtat	gctgtgtat	ggcagaccat	ggttcaacaa	gaccaagacc	4020
tttcggatca	atcagctgt	ccaggacaco	cctgtggct	gcccataa	cgtgaaatgt	4080
tggcaggacg	cctgcgtcc	cctgcctaaa	acacaggcca	atcaggggc	tctgcagg	4140
ggcgatatcc	ccacaaggcc	tctgtgtt	aagctgccc	aaagactgt	caagctcg	4200
agccagttcg	ctgtggat	cgagaccac	tcccccac	gacttgggtt	tcacaccggc	4260
accaagaaca	gttcatggc	cctgtacat	agcaaaaggc	ggctgggtt	tgcctgggg	4320
accgacggaa	agaaaactgcg	gatcaagac	aaagagaatgt	gcaacgcac	caagtggcc	4380
accgtgttgt	tccggacacga	tggcgagaaa	ggcagactcg	tggtggatgg	cctgagagcc	4440
agagagggat	ctctgcctgg	caactccatca	gagccctgt	gtatctggc	4500	
agccctctca	gcggaaagcc	taagacccctg	cctaccaact	ccttcgtgg	ctgtctgaag	4560
aactttcagc	tggacagacaa	gcctctgtac	acccttagca	gcagctttgg	cgtgtccct	4620

-continued

tgtctcgag	gcctctgga	aaaggcata	tacttctctg	aggaaggcgg	ccacgttgc	4680
ctcgcttatt	ctgttctgt	ggggcccgg	tcaagctgg	tgttcttat	ccggcctaga	4740
agoctgaccg	gcatctgtat	teacatcggo	agccagcttg	ggaagcacct	gtgtgttat	4800
ctcgaggccg	gaaaaatgac	cgcacatgt	gattctggg	ctggcggcac	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	tcttcttctt	gtgtttgcc	ctgcctggcc	tcttgcacgc	ccaacaaggc	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	agtccctggc	tcagaggcct	540
aatgcacccg	taatgggggg	gaaggctcaa	cttaaccta	tggatttagt	gtctgggat	600
ccagcaactc	aaagtcaaaa	aatggcaagag	gtgggggaga	tcacaaactt	gagagtcaat	660
ttcaccaggc	tgccccctgt	gccccaaagg	ggtaccacc	ctcccaagcgc	ctactatgt	720
gtgtccccc	tcgcgtctgc	ggggagctg	tctgtcaccg	gccatgtca	tcgctgcgca	780
cccaaccctg	ggccctctgc	aggggccatc	acggctgtgc	aggtccacga	tgctgtgtc	840
tgcagcaca	acactggccg	cccaattgt	gagggctgt	cacccttca	caacaacccg	900
ccttggagac	ccggcgagggg	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	acccgcagg	ctggccacc	tgtgactgca	acatctggg	gtccggagg	1320
gacatcgct	tgacggaga	gatggggcgc	tgccttgc	tgcccaactg	gttgggtccc	1380
aaatgtgacc	atgtgtctc	ctaccaactg	aaactggcc	gtggccagg	ctgtgaacccg	1440
tgtgcctcgc	acccgcacaa	ctccctcago	ccacagtgc	accagtac	agggcagtgc	1500
ccttgcgggg	aaaggcttgg	tggctgtat	tgccgtgt	caggcgtcc	ccagtgtcc	1560
gaccggacat	atggagacgt	ggccacagg	tgccgagcct	gtgactgt	tttccggg	1620
acagaggggc	ccggctgcga	caaggatca	ggccgtgc	tctggccccc	tggctgtacc	1680
ggggccccc	gtgaccagt	ccagcgagg	tactgtatc	gttcccggt	gtgcgtggcc	1740
tgcaccatccc	gttccagat	ctatgtatc	gacccctgg	agcaggccct	gcgttttgt	1800
agactccgc	atggccacgc	caggctgg	teaggccgt	ggctggagga	ccgtggct	1860
gcctcccgga	tcctagatgc	aaagtagta	attgagcaga	tccgagcagt	tctcagcagc	1920
cccgcaigtca	cagacgagg	ggtggtctag	gtggccagt	ccatctctc	cctcaggcga	1980
actctccgg	gcgtcagct	ggatgtccc	ctggaggagg	agacgttgc	ccttccgaga	2040
gaccctggaga	tgcttgacag	aagettcaat	gttcttctt	ctatgtatca	gaggaagagg	2100
gagcagtttgc	aaaaataag	cagtgtat	ctttcaggag	ctttccggat	gtctgacaca	2160
gcctacgagc	agtccagcc	ggctgtctag	caggctcc	acagctcg	cctttggac	2220
cagctcagg	acagccggag	agaggcagag	aggctgtgc	ggcaggccgg	aggaggagga	2280
ggcacccggca	gccccaaatg	tgtggccctg	aggctgg	tgttcttgc	gtctgacatg	2340
acacccac	tcaacaatgc	ctgtggca	tccaggcaga	tggctgcac	cccaatata	2400
tgccttgg	agctatgtcc	ccaagacaat	ggcacagcc	gtggctcc	ctgcagggt	2460
gtcttccca	ggggcgttgg	ggcccttctt	atggcggggc	aggtggatg	gcagctgcgg	2520
ggcttcaatg	cccaatgtca	gcccggagg	catatgttca	gggcacccga	ggaatctgc	2580
tcacatcgat	aatccatgtc	ccagcgctt	gacacccagg	tgatgc	ccgcctcc	2640
atggagggaa	atgtcagacg	cacacggc	ctaatcc	aggccgg	cttcttaaca	2700
gaccggac	ctgatgc	cactatcc	gaggctcag	aggccgt	ggccctgtgg	2760
ctgcccacac	actcaatgt	tgttctgt	aaatgtat	gatccaggc	cattgcagcc	2820
aggetccccc	acgtgttgc	gggtgttgc	catccaac	aggatccat	gcgtggcc	2880
cgggttgc	ctgaggctg	ggaaggcc	agccgaccc	atgcgttgc	ggccagggt	2940
gaagatgtgg	ttggaaacct	gccccgggg	acgtggac	tgcaggaa	tcaggacacc	3000
atgcacggca	ccagccgtc	ccttcgg	atccaggac	gggttgc	gtttcagcag	3060
gtactggcc	cagcaaaaa	gtctgtaca	agcatgac	agcagctgg	tgacttctgg	3120
acacggatgg	aggagctcc	ccaccaagcc	ccgcagcagg	gggcagaggc	agtcaggcc	3180
cagcagcttgc	cgaaagggtc	cagcagcag	catgttgc	cccaagagg	atttgcag	3240
ataaaacaaa	aatatgt	gttgcaggac	cggttgg	ccatgtat	gtctgggt	3300
cagggtgtcc	ggatccagag	tgtgcagaca	gaggcagagg	agctgttgg	ggagaccat	3360
gagatgtgg	acaggatgaa	agacatgg	ttggagctgc	tgccggccag	ccaggccatc	3420
atgtctgc	cagccgac	gacaggact	gagaacgtg	tggagcagat	ccgtgaccac	3480
atacatgggc	gcgtgtctca	ctatgc	tgcaagt			3519

-continued

-continued

cccacccatcca ggagggaaat ctgtgatgc aatggaaatg ccaggcagtgc tatcttttat
 cgggaaacttc acagacaatac tggttaatggg ttccgcgtcc tcaactgca tgacaacact 180
 gatggcatttc actcgccgaa gtgcgaagat ggcttttacc ggcacagaga aaggggccgc 240
 tggggccct gcaattgtta cttccaaagggt tcttttagtg ctcgcgtgtg caactccggg 300
 cggtgacgatc gtaaaaccagggt tggtcacaggg gccagatgcg accgatgtt qccagggttc 360
 cacatgtcga cggatcgccc gtgcaccggaa gaccagagac tgctgatgactc caagtgtgac 420
 tgggaccccg ctggcatcgc aggcccctgt gacggggcc gctgtgtgtc caaggccgat 480
 gtcactggaa aacgctgtga taggtgtcga tcagggtact ataactctggg tggggggaaac 540
 ctggggcgt gtaaccagggt tttctgtcat gggcatctcg ccaactggccg cagctgtca 600
 gaatacagtg tccataagat cacctctacc tttcatcaag atggttgcgtt ctggaaaggct 660
 gtcccaacgaa atgggtctcc tgcaaaagtc caatgtgtc acggccatca agatgtgttt 720
 agetcagcccc aacgactgaa ccctgtctat ttgtgtgtc ctgcggaaattt tcttgggaaat 780
 caaacatgttgc gtcataatgcgaa aagegtgtc ttgtacttgc ttgtggacag aggaggcaga 840
 caccatctg cccatgtatg gattctggaa ggtgtggtc tacggatcac agtccccttg 900
 atgcccattt gcaagacact gccttgcgtt ctcaccaaga ctacacatc caagttaaat 960
 gaggatccaa gcaataatgg gggccccccat ctgagttact ttgatgtatc aaggttactt 1020
 cggaaatctca cggccctccg catccggatc acatattggg aatacactac tgggttacatt 1080
 gacaatgtga ccctgtattt agccccccct gtctgtggag ccccaagcacc ctgggttggaa 1140
 cagtgtatata gtcctgtttt gtacaagggg caattctgcg aggatgtgc ttctgggtac 1200
 aagagagatc cagcgagact gggccctttt ggcacccgttata ttctgtgtt ctgtcaagggg 1260
 gggggggccct gtgtcccaaa cacaggatgt ttgttattcag gggatgagaaat tctgttacatt 1320
 gagtgtgtctg actgccccat tggtttctac aacgatccgc acgacccccc cagctgcaag 1380
 ccatgtccct gtcataacgg gttcagctgc tcagtgatgc cggagacggg ggaggttgggt 1440
 tgcacaataactt gcccctccccc ggtcaacccgtt gcccggctgtg agctctgtgc tgatgtgttca 1500
 ttggggaccc ctttgggtga acatggccca gtgaggccctt gtccggcttgc tcaatgcac 1560
 aacaatgtgg accccatgtc ctctggaaat tggtacccggc tgacaggcag gtgtttgaag 1620
 tgtatccaca acacagccgg catctactgc gaccagtgca aacgaggctta cttccggggac 1680
 ccattgtggccctt ccaacccccc agacaaatgtt cgagctgtca actgttacccc catgggttca 1740
 gaggctgttag gatgtcgaaat tgatggcacc ttgtgttgcg agccggattt tggggcccc 1800
 aactgtggaccc atggggcattt cagctgtccca gctgtgtata caactgtaaa gattcagatg 1860
 gatcagtttca tgcagcgttgc tcagagaatg gaggcccttgc ttcaaaaggc tcagggttgggt 1920
 gatgggatgt tacctgtatc agaggtgtttt ggcaggatgtc agccggcttgc gcaggccctt 1980
 caggacatttc tgagagatgc ccaggatttca gaagggtgtca cggatccct tgggttccgg 2040
 ttggccaaagg tgaggggccca agagaaacccatc ttagggggccgc gcttggatgaa cttcaaaatgtt 2100
 actgtggaaa gagtttgggc tcttggaaatg cgttacccggc accggatgtcg ggataacttc 2160
 aggttccatca ctcagatgttgc gtcgtggcttgc gcaaaatgttgc aactgttccctt gggaaacactt 2220
 aacattcttgc ctcagacca ctaatggggcc ccaaatgttgc taaaatgttgc ggttccgggg 2280
 ggcaccaatgt tagccaaatgg ccacgttggatc tcacgcgttgc acatggggcc actggccaaagg 2340
 gaaactgtggg actattccaa acaaggccctc tcactgtgtc gcaaggccctt gcatgtggaa 2400
 gtccggaaatgg gaagegggttgc cccggggccgtt gctgtgtgtc aagggttgcgggatggaaaattt 2460
 gagaaaaacccatc agtccctgtcc gcaatggggcc ccaactcaaggc gggaaatgttgc 2520
 gcaatgttgc cttatcggatc cagttccggc cttctggatc tgggttccgc gttccgggg 2580
 gtcagtgtatc agtcccttgc ggtggaaatgg gcaaaaggatgg taaaatgttgc ggttccgggg 2640
 ctcttcggccatc tggttaaccatc gcatatgttgc gatgttgcggat gacatggggcc gatgttccgggg 2700
 aactgtggaaa aagaacccatc gcaatgttgc cagaatggggcc aagatgggggg agggaaatgttgc 2760
 gatcgtgtcc tttccggccatc caatgttgc taaaatgttgc gcaacggatgg cacaacggatgg 2820
 ggcaccaatgttgc ctttttataatgg agttggggcc gatgttgcggat gtcgtgttgc gttccgggg 2880
 cagggtggaca acagaaaaatgg agaaggtgttgc gaaaggccatc agagacttcc tcacatccggc 2940
 cagaatgttgc cttatcggatc tggtggaaatgg gcaacggatgg cacaacggatgg gggggggccgtt 3000
 gtcgtgtatc gtcgtgttgc gaaaggatgg gggggggccgtt cttccgggg 3060
 atttggggccatc agtccctgtcc tttccggccatc gatgttgcggat gtcgtgttgc gttccgggg 3120
 gccatggaaa agggacttgc ctctgttgc gatgttgcggat gggaaatgttgc gggggggccgtt 3180
 gaaaggaaatgg agtccctgtcc tttccggccatc gatgttgcggat gtcgtgttgc gttccgggg 3240
 gcccggccatc tggtttccggatc gggggggccgtt gggggggccgtt gggggggccgtt 3300
 acatggggccatc gtcgtgttgc gttccggggccgtt gggggggccgtt gggggggccgtt 3360
 gtcttactgttgc agcagaatgttgc tttccggccatc gatgttgcggat gtcgtgttgc gttccgggg 3420
 atgtgttgc gtcgtgttgc gttccggggccgtt gggggggccgtt gggggggccgtt 3480
 aacagatgttgc tttccggccatc gatgttgcggat gtcgtgttgc gttccggggccgtt 3540
 ccccccggccatc gtcgtgttgc gttccggggccgtt gggggggccgtt gggggggccgtt 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 constru
```

SEQUENCE:	48	Organism = Synthetic construct				
atgcctgtct	tgtggctggg	ctgctgcctg	tgttttagtc	tgctgtgcc	agccgccaga	60
gccccatcta	gaagagaagt	gtgcgactgc	aacggcaaga	gcccgcagtgc	catttcgac	120
agagagctgc	acagacagac	cggcaacggc	ttcagatgcc	tgaactgcaa	cgacaacacc	180
gacggcatcc	actgcgagaa	gtgcagaac	ggcttctacc	ggcacccgca	gagggataga	240
tgcctgcct	gcaactgcata	cttccaagggc	agctgcgacgc	ccagatgcga	caatagccgc	300
agatgtatgc	gcaaggctgg	cgtgcacggc	getatagtcgc	atatagttct	ccccggcttc	360
ccatctgtca	ccgtgcggcc	atgttacccgg	gaccagagac	tgctggacag	caagtgcgt	420
tgcgaccctg	ccggaaattgc	cggacattgt	gatggccggaa	gatgcgtgtc	taaacctggc	480

-continued

tgacccggcg agagatgtga cagatgtaga a诶cggctact acaacctgga cggccgcaat 540
 cctgaaggt gcacccagtg ctttgcgtac ggccacagcg ccagctgttag aagcagcgcc 600
 gaataccctc tgccacaagat caccaccc ttccacccagg atgtggacgg atgaaaggcc 660
 gtgcagagaa atggcttcc tgccaaagtg cagtggcccc agagaccca ggacgtgttc 720
 acaacgctc agagatgtga ccccgctgtac ttgtggcccc ctgccaagtt ctggcccaac 780
 cagcaagtgt cttacggca gagcctgace ttgcactaca gagtggatag aggccgcaga 840
 caccccccgc ctcaagatgt gattcttggaa ggccggccac tgccgtatcc agcccttctt 900
 atgccttgg gcaagacccgt gccttggcc ctgaccaga cctcacaccc cccgctgtaa 960
 gagcaccacca gcaaaactcg gtcccccaccc ctgactact tccagatcag acggctgtg 1020
 cggaaacctga cagccctgag aatcagagcc acctacccggc agtacagcac eggtacatc 1080
 gacaacgtga ccctgtatcag cgccacagcc ttttctgggt ctccgtctcc ttgggtcetag 1140
 cagtgtatct gttccctgtgg ctacaaggcc cagttcttgcg aggatgtgc cagggctcat 1200
 aagagactgt ctgcacagat gggcccccgt ggcacatcga tcctctgttgc tggtcaggcc 1260
 ggcggagccct gcgatcccgta tacaggccat tgctacccgt ggcacgagaaa ccccgatatac 1320
 gagtgccgccc attgtccccat cggcttttac aacgaccctc acgaccacccag atccgtcaag 1380
 ccatgttccct gccacaatgg cttagctgc a诶cgtgtatc ccggaaacccaa agaggtcg 1440
 tgcaaaatccat gcccacccagg cttagacagg ccggatgtg aactgtgtgc cggggctcat 1500
 ttccggatcc tttttggaga acacggccatc gtggccaccc ttggcccttgc tcagtgcaac 1560
 aacaacgtgg acccaacgcg cagcggcaac tgcgatagatc tgacaggccatc atgctgttgc 1620
 tgcattccaca ataccggccatc gatctactgt gaccatgtca agggccggta ttttggccatc 1680
 cctctggctc ccaatcttgc ctggatgtcc agacgttgc actgttaccc tatgggtctt 1740
 gagccgttgg gtcgtccatc tgatggaaacc tgctgtgtca agggccgtt ttggggccatc 1800
 aatttgtgaac acggcccccct tagctgcccc gctgtctaca atcaagtgtt gatccagatg 1860
 gaccatgttca tgcacgactgt gcaaggatgtt gaaggccctgtatccaaaggcc 1920
 gatggcgtgg tgcctgtatc agatgtggaa ggcagaatgtc agcaggccgc acaggccctt 1980
 caggacatttc tgagatgtcc ctggatgttcc gggggccctt ctggatgtt gggactgtc 2040
 ctggctaaag tgccggacca agagaacccgc taccagagca gactggacga cctgaagatg 2100
 acccggtggaa gagtcagatc cctggccgcg cagttccatc acagactgtcg ggataccac 2160
 cgggtgtatcc cccagatgtca atgttcttgc gccggagccg aaggccatgtt gggcaatacc 2220
 aattatcccg ccacgcacca ctactgtggcc cccaaaggctt ttggactgtt ggttcaagag 2280
 gccaccacccat tggccggaaatccatgtggaa agcgttccca acatggaaaca gctgaccggc 2340
 gaaaccggagg actactctaa gcaggccctg agctctgtca gaaaaggccct gcatggaggc 2400
 gtccggcggcc gctctggatcc ttctgtatgtt gctgtgtgc agggactctgtt gggaaaggcc 2460
 gaaaagacca aatcttgcg ccacgcgttcc acagagaaatcc ccacacccggc cgatgtcg 2520
 ggcggacacaa gctaccatgtcc tcaactgtggatcc ctgtgtgttccatgttactgtcc 2580
 gtgtccgacc agagttccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 2640
 ctggagccgcg tggttccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 2700
 aactgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 2760
 gaccatgtcc tgacccggatcc acactgtccatgtggaaatccatgtggaaatccatgtggaaatcc 2820
 ggcaacgccta ccttcttgcgttccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 2880
 caagtggccatc acagaaaggccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 2940
 cagaaaggatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3000
 gcaacggatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3060
 ctcacggatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3120
 atcgagacaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3180
 gccatggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3240
 gaacggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3300
 gcccggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3360
 tggccggccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3420
 atgtatgtccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3480
 accaggatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3540
 ccacccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatccatgtggaaatcc 3582

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

-continued

```
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
atggcccttcg tgaacaagca gttcaactac aaggaccccg tgaacggcgt ggacattgcc 60
tatcatcaaga tcccccaacgc cggccgatag cagccctgtga aggcccttaaa gatcccacaa 120
aatgatctggg tcatacccccga gggggacacat ttccacaaaatc ccggagaaagg cgacactgtaa 180
cctccacccgt aagccaaacaa ggtggccctgtg tcctactactcg acagacactca ctggagacc 240
gacaacgaga aggacaacta ccttgaaggcgt gtgaccaaaagc tggtccgagcg gatctacagc 300
accgatctgg gcagaaatgtc gctgacccctt atcgtgcggg gcatccatc ttggggccggc 360
agcacaatcg acaccggatc gaaagtgtac gacaccaact gcatcaacatcg gatccacgt 420
gacggccgatc agacaaggcga ggaactgtaa ctggtccatca tcggcccccgg ccggccacatc 480
atccagttcg agtgcacccatc cttccggccac gaggtgtctga acctgacccatc aaatggctac 540
ggcagcaccc agtacatccg gttcagcccc gatttaccc tccgggttccgaa ggaatccctg 600
gaagtggacca ccaatccctt gctccggaccc ggcacaaatgg ccacccatcc tgctgtccgaa 660
ctggcccccacg aactgtatcca cccggggacac agactgtacg gaatccgtccat caatcccaac 720
cgggtgttca aagtgcacccatc caacgccttccatc tatggatgaa ccggacttgaa agtgccttt 780
gaggaacttgc ggaccccttcgg cggacacccac gccaaatggtca tcgacagccatc gcaagagaa 840
gagttccggc tgtaactacta caacaaggatc aaggatatacg ccagcacgtt gaaacaaggcc 900
aagagcatcg tggggcacaac acggccatccgtc cagtagatcg aagaacgtttt caaaagagaa 960
tacccatgtggc gggggggacac cggccggccatc ttctccgtgg aacaacgttggaa gttccgacaa 1020
ctgtacaaaga tgctggccca gatctacccatc gagggacact tctgtgaaatgtt ccataagggtt 1080
ctcaaccggca agacacttccatc caacttcggc aaggccgttca tcgacatcaa catctgtggcc 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	aggcgagaaa	1440
atcaccagcg	acaccaatat	cgaggccgc	gaggaaaaac	tcagctgga	cctgatccag	1500
cagtactacc	tgacccatca	tttcgataac	gagccccgaga	acatcagcat	cgagaactcg	1560
agcagcgcata	tcatecgaca	gctggaaactg	atgcccaca	tcgagagatt	ccccaacggc	1620
aagaagtacg	agctggacaa	gtacaccatg	ttccactacc	tgccggccca	agagttcgag	1680
cacggcaatg	ctagaatcgc	cctgaccaaa	agcgtgaac	aggccctgt	gaaccccaac	1740
agagtgtaca	ccttcttcag	cagegtacat	gtgaagaaag	tcacaaaggc	tacagggcc	1800
gccccatgttcc	tcgggtgggt	tgagcagctg	gtgtacgact	tcacccgacga	gacaaggcgag	1860
gtgtccacca	ccgacaagat	cgccgatatac	accatcatca	tccttacat	cggccctgtct	1920
ctgaacatcg	gcaacatgt	gtataaggac	gatttcgtgg	gcccctgtat	cttctctggc	1980
gccccgttgc	tgtctcgagtt	cateccatgg	atcgctatcc	cctgtgtgg	cacatttgc	2040
ctgtgttctt	atactccaa	caaggtgt	acatgtcaga	ccatcgacaa	cgcctgtac	2100
aagcggaaacg	agaagtggga	cgaagtgtac	aagtacatcg	tgaccaactg	gctggccaaa	2160
gtgtaaatcc	agatcgac	gattcggaa	aagatgtgg	aaggccctga	gaaccaggcc	2220
gaggccacaa	aggccatcat	caactaccag	tacaatcag	acacagagga	agagaagaaac	2280
aacatcaatt	tcaacatcg	cgacccgttc	tccaaatcg	acgagacat	caacaaagcc	2340
atgatcaata	tcaacaagtt	tctgaaccag	tgccgtgtc	gatcacatgt	gaacagcatg	2400
atccccatcg	ggtgtaaacg	gctggaaatg	ttcgatgtca	gcctgaagg	cgctctgt	2460
aagtacatcg	acgataacc	ggggccctcg	atcgccagg	ttggacagact	gaaggacaaa	2520
gtgtacacata	ccctgtccac	cgacatccg	tttcagctg	gcaaatatcg	ggacaacccag	2580
agactgtga	gcacattc	cgagtatata	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
ttctccacca	gtttttggat	cagaatcccc	aagtacttca	acagcatctc	cctgaaacac	2880
gagttacacga	tcatcaattg	catggaaaac	aactccggct	ggaaagtgtc	cctgaaactac	2940
ggcggatggac	tctggacact	gcggacac	caagagatca	acgagagatg	gtgttcaag	3000
tactctcaga	tgtcaacat	tgccgtact	atcaaccgg	ggatcttgt	gaccatcacc	3060
aacaacccgc	tgaacaactc	caagatctac	atcaatggcc	ggctcatcg	ccagaaggct	3120
atcagcaacc	ttggaaacat	ccacgcctcc	aacaatatac	tggtcaatg	ggacggctgc	3180
cggggacacc	accggatata	ctggatca	tacttttaacc	tgccgtacaa	agagctgtac	3240
gagaagaa	ttaaggac	gtacgacaa	cagtccaa	ggccgtatct	gaaggattc	3300
tggggcact	acctgcgt	tgacaaagcc	tactatcg	tgaatctgt	cgaccccaac	3360
aaatatgtgg	acgtgtacaa	cgtggggatc	agaggctaca	tgtacatgaa	aggccccaga	3420
ggcagcgtga	tgaccacaa	catttcac	aactccacg	tgtacagagg	caccaagttc	3480
atcatcaaga	atgtgc	cggtacaa	gacaacat	tgcggaa	cgaccgggt	3540
tacattaaacg	tgggtgtca	gaacaaagag	tacccgttg	ccaccaatgc	cttcaggc	3600
ggcgtggaaa	agatctgt	cgccctggaa	atcccgac	tgggtcaatct	gtctcaggt	3660
gtctgtatgc	atgtcaagaa	cgccagg	atcacaaatc	agtgcaagat	gaacctccag	3720
gacaacaaacg	gcaacgc	cggttttac	ggcttccac	agtttaacaa	cattgcaag	3780
ctgtgtcoca	gcaactgtt	caacccgc	atcgagagaa	cgacgacaa	cctgggtc	3840
agctggagat	ttatccctgt	ggatgtgc	tggggc	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
 atggccaggta caataaataa tttaattat aatgatccta ttgataataa taatattatt 60
 atgatggagc ctccatttgc gagaggatcg gggagatatt ataaagctt taaaatcaca 120
 gatcgatattt ggataatacc gggaaatgat accttttggat ataaacctga ggattttat 180
 aaaagttccg gtattttaa tagatgtt tgtaatatt atgatccaga ttacttaat 240
 actaatgata aaaagaatata attttacaa acaatgtca agttatttaa tagatcaaa 300
 tcaaaacccat tgggtgaaaat gttttagag atgattatac atggatatac ttatcttgg 360
 gatagacgtg ttccactcg aaggtttac acaaacattt ctagtgtac tgtaatataa 420
 tttaatcgta atccaggaga agtggagcga aaaaaggta ttttcgcaaa tttaatataa 480
 tttggacctg ggccagttt aaatgaaaat gagactatac atataggtat acaaataatcat 540
 tttgcataca gggaaaggctt cgggggtata atgcaatgaa agttttgtcc agaatatgt 600
 agcgttataa ataatgttca agaaaacaaa ggccgtatcgtatataatg acgtgtat 660
 ttttcagatc cagccgtatc attatgtca gaaacttatac atgttttatac tggatttat 720
 ggcattaaag tagatgtt accaatgtt ccaaatgaa aaaaattttt tatgtatct 780
 acagatgtca tacaggcaga agaactatac acatgttggag gacaagatcc cagcatcata 840
 actcccttca cggataaaaag tatctatgt aaagtttgc aaaattttag agggatagtt 900
 gatagactta acaagggtt agttgtcata tcagatccta acatataat taatataat 960
 aaaaataataat taaaagatata atataaattt gttgaagatc ctgagggaaa atatagtata 1020
 gatgtgaaaat gtttgcata attatataaa agttaatgtt tggttttac agaaactat 1080
 atagcagaaa attataaaaat aaaaactaga gcttcttatt ttagtgcattt cttaccacca 1140
 gtaaaaataa aaaattttt agataatgaa atctatacta tagaggaaat gtttaatata 1200
 tctgataaaag atatggaaaaa agaataataga ggtcagaataa aagctataaa taaaacaaatgt 1260
 tatgaagaaa ttagcaagga gcattttggct gtatataaga tacaatgtg taaaatgtt 1320
 aaagctccag gaatatgtat tgatgttgc aatgaagatc tggttcttgc agctgtataa 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
gatgcattat tattttcaa caaagtttat tcattttttt ctatggatta tattaaaact	1740
gctaataaag tggtagaagc aggattttt gcaggTTGGG tgaaacagat agtaaatgat	1800
tttgtaatcg aagctaataa aagaataact atggataaaa ttgcagatat atCTCTAATT	1860
gttccttata taggattagc tttaaatgta gggaaatgaaa cagctaaagg aaatTTGAA	1920
aatgcTTTG agattgcgg agccggattt 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 aaatTTAAAC atcgatTTA atgatataaa ttCTAAACTT	2280
aatgaggggA ttaaccaAGC tatagataat ataataataat ttataatgg atgttCTGTA	2340
tcataTTAA tgaaaaaaat gattccatta gctgttagaa aattactaga ctTGTATAAT	2400
actCTCAAA aaaaTTGTTT aaattatata gatggaaaata aattatattt gattggaaGT	2460
gcagaatATG aaaaATCAAAGT agttaataaa tacttgaaa ccattatGCC gtttGATCTT	2520
tcaatataa ccaatgatac aataataa gaaaATGTTA ataaatataa tagcgaatt	2580
ttaaataata ttatCTTAA tttaagatata aaggataata attaataga ttatTCAGGA	2640
tatggggAA aggttaggat atatgtatgg gtcgagCTTA atgataaaaa tcaattttAA	2700
ttaactagTT cagcaatAGT taatggatTA gtgactcaaa atcagaatAT catatTTAT	2760
agtgtgttcc ttgatTTAG cgttagCTT tggtataAGA tacCTAAATA taagaatGAT	2820
ggtatacAAA attatATCA taatgatata acaataatTA attgtatgaa aaataatTCG	2880
ggctggAAA tatCTTAGG gggtaatAGG ataataatgGA ctTTAATTGA tataatGGa	2940
aaaacccaat cggatTTTT tgaatataaD ataagagaAG atataatCAGA gtatataaAT	3000
agatggTTT ttgtAACTAT tactaataaC ttgtataACG ctaaaATTa tattatGGT	3060
aagCTGAAT CAAATACAGA tattaaAGAT ataaGAGAAG ttattGCTAA tggtGAATA	3120
atattttAAAT tagatGGTGA tatagatAGA acacaatttA ttTGGGATGA atattTCAGT	3180
atTTTAAATA cggatTAAG tcaatcaaa atggaaAGA atggaaAGA atataatTCAT	3240
aggcaatTTT taaaAGATTG ttggggAAAT cctttatGt acataAAAG atattatATG	3300
ttaatGCGG ggaataaaaa ttcatatatt aaactaaAGA aagattcAcc tGTTAGTGAA	3360
atTTTAAACAC gtagcaAAAt taatcaAAAt tcttaatata taaattatAG agatttat	3420
atTGGGAGAAA aattttatTTT aagaAGAAAG tcaaaTTCTC aatCTATAAA tgatGATATA	3480
gttagaaaaAG aagatttat atatGATAG tttttatA taaaTCAGA gttggAGAGTA	3540
tatacCTATA aatTTTTAA gaaAGAGGAA gaaaaATTG tttttagtCC tataatGAT	3600
tctgtatGAGT tttacaatac tatacaAAAt aaAGAAATAG atgaACAGCC aacatataGT	3660
tgtcagtTGC ttTTTAAAGA agatGAGAAG agtactGATG agatAGGATT gattGGTT	3720
catcgTTCTC acgaatTCG aattgtATTt gaagAGTAA aagatTTAT ttgtataAGT	3780
aaatGgtACT taaaAGAGGT aaaaAGGAAa ccataATAA taaaATTGGG atgtatTTG	3840
cgatTTATC tcaaAGATG 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 TGGCGAGAA gctgtggaa atgatcatca acggcatccc ttacCTGGC 360
gacagaagAG tgccccTGGa agagttcaac accaataatcg ccagcgtgac cgtgaacaaG 420
ctgatCTCTA acccCGCGA ggtggAACGG aagaaggGCA tcttgcocca cctgtatcatc 480
ttcggCCCTG gacCTGTGCT gaacAGAGAAG gagacaatcg acatcggcat ccagaaccac 540
ttcggCCAGCA gagaAGGCTT CGGCGGATC atcgatGATG atgttCTCCC cgagtacGT 600
tccgtgttca acaacGTGCA agagaacaAG ggGCGAGCA tctttaacAG acggggCTAC 660
ttcagcgacc CGGTCTGTat tctgtatGCAc gagCTGATCC acgtGCTGCA cggcCTGTat 720
ggcatcaAGG tggacGACCT GcCTATCGTG cccaaACGAGA agaaATTCTT catcgatGTC 780
accgacGCA tccaggCGGA gcaactGTAC acatTTGGCG gacaggACCC cagcatcatc 840
acccCTAGCA cggacaAGAG catCTACGAC aaggGTGTCG agaaATTCCG gggcatCGTG 900
gaccGGCTGA acaaAGGTGT CTGTGTCATC acgcataCCCA acatcaatAT caacatCTAC 960
aagaacaAGT tcaaggATAA gtacaAGTT GTCGAGGACA GCGAGGGCAA gtacAGCATC 1020
gacgtggAAA gcttcgacAA gctgtacaAG agcCTGATGT tcggCTCAC CGAGACAAAT 1080
atcgCCGAGA actacaAGAT caagACCCGG GCGAGCTACT tctccGACTC tctgCTCTCT 1140
gtgaagATTA agaacCTGTCT ggacaACCGAG atctacACCA tcgAGGAAGG cttaacatc 1200
agcgacaAGG acatGGAAA agagtACCGG GGCGAGAAAC aGGCCATTAA caAGCAGGCC 1260
tacgaggAAA tcagcaAGA ACACCTGGCC GTGTACAAGA ttcaGATGTG caAGAGCGTG 1320
aaggCCCTG GcatCTGcat tgacgtggac aatgaggACc tGTTCTTAT CGCCGACAAg 1380
aacagCTTA GCGACGACCT gagaAGAAC gacggatCAG agtacaACAC CCAGAGCAAC 1440
tacatcgaga acgacttccc catcaacgaa ctgatCCTGG acaccGACCT gatcAGCAAG 1500

```

-continued

atcgagctgc	ccagcgagaa	caccgagago	ctgaccgact	tcaatgtgga	cgtgcccgtg	1560
taacggaaagc	agcccgccat	caaaagatc	tttaccgac	agaataaccat	cttccatgtac	1620
ctgtacagcc	agaccccttc	tctggacatc	cgggacatca	gcctgaccc	cagcttcgat	1680
gatgccctgc	tgttcagcaa	caaggctcac	agcttctca	gcatggacta	catcaagacc	1740
gccaacaagg	tggttggaaagc	cggcctgttt	gcccggctggg	ttaaggcagat	cgtgaacgat	1800
ttcgtatcg	aggccaaacaa	gtccaaacacc	atggacaaga	tcggccgat	atccctgtac	1860
gtgccttaca	tcggactggc	cctgaacgtg	ggaaacggaga	cagccaaagg	caactctcgag	1920
aatgccttc	agatttgcgg	cgtacgtac	ctgctcgat	tcatccctga	gctgtgtac	1980
cctgtctgg	gctgtttct	gctggaaatcc	tacatcgata	acaaaaacaa	gatcatcaag	2040
acqatcaca	acggccctgac	caaggcggaa	gagaagtgg	gccccatgt	ccggactgtac	2100
gtggcccaat	ggctgagcac	cgttaatacc	cagttttaca	ccatcaaaga	agggtatgtac	2160
aaggccctga	attaccaggc	tcaggctctg	gaagagatca	ttaagtaccc	ctacaatata	2220
taacagcgaga	aagagaagtc	taacatcaac	atcgacttca	acgacatcaa	cagcaagctc	2280
aacgaggcga	tcaaccaggc	cattgacata	attaacaatc	ttatcaacccg	ctgcagcgat	2340
tcctacatg	tgaaggaaat	gatccctgt	gcccgtggaa	agctgtcgat	cttcgacaaat	2400
acccctgaaga	agaacccctc	gaaacttacat	gacgagaa	acgtctactt	gatecgccac	2460
gcccggatcg	agaaaaagca	agtgaacaag	tacctcaaga	ccatcatgac	cttgcacat	2520
tccatctaca	caaaacgacac	catectgtac	gagatgttta	acaagatcaa	cagcgagatc	2580
ctgaacaata	tcatctctga	cctgcggat	aaggacaa	atctgtatca	tctgaggccg	2640
tacggccoca	aggttggaaat	gtatgtatcc	gtggactgt	acgatggaa	tcagttcaag	2700
ctgaccggca	gcccggactc	caagatcaga	gtgacccaga	accagaacat	tatcttcaac	2760
agcgtgttcc	ttggatcttc	cgtgtcttc	tggatcaagaa	tccccaatgt	caagaacggac	2820
gcoattccaga	actacatcca	caacggatca	acaatcatca	actgtatgaa	gaacaacacg	2880
ggcttggaa	ttagatcccg	ggcaacaga	atcatctga	ccctgtatca	catcaatggc	2940
aagacaaaga	gcgtgttctt	cgagtataac	atccgcgagg	acatctccg	gtatcatcaac	3000
cgggttgttct	ttgttggaaat	caccaacaa	ctgaacaaac	ccaagatcta	catcaacccgg	3060
aagctcgaga	gcaacacccg	catcaaggat	atccggggaa	tgatcgocaa	ccggcgagatc	3120
atctttaa	tgccggccgca	catcgacccg	acacagtca	tctggatgaa	gtacttttc	3180
atcttcaata	cccgactgt	ccagacaa	atcgaagaga	ggtacaatgt	ccagtcttac	3240
agcgagttacc	tgaaggactt	ctggggcaac	cctctgtat	acaacaaaga	atattatcat	3300
ttaaacggcc	gcaacaaagaa	cttccatatac	aagctgtatc	aaagacggccc	cgtgggagaa	3360
atcttgcacca	ggtctaaatg	ccaaacggaa	tccaaatgt	tcaactaccc	ggacctgtac	3420
atccggcaga	agttcatcat	ccggcgcaag	agaacacac	agtccatcaa	tgacgacatc	3480
gtgcggaaag	aggactatata	ctacccgtac	ttcttcaacc	tcaatcaaga	gtggcgctg	3540
tacaccatca	actgtttaa	gaaaaggagg	gagaagtgt	tcctggotcc	tatcagcgac	3600
tccgacgat	tctataatac	catccgatc	aaagatgtac	acgacggccc	cacccatcagc	3660
tgcagctgc	tgttttaagaa	ggacggagaa	tccaccatgt	agatcgccct	gattggatc	3720
caccgggttct	acgagagccg	catctgttcc	gaagagtaca	aggattactt	ctgcatac	3780
aagtggatc	tgaaaatggaa	gaagcggaa	ccctacaatc	tgaagctgg	ctgcaactgg	3840
catgttatcc	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
INQWQERTYL	GNALVCTCYG	GSRGFNCESK	PEAEETCPDK	YTGNTRYVRGD	TYERPKDSMI	120
WDCTCIGAGR	GRISCTIANR	CHEGGQSYKI	GDTWRRPHET	GGYMLECVCL	GNGKGEWTCK	180
PIAEKCFDHA	AGTSYVVGET	WEKPYQGWMM	VDCTCLGEBS	GRITCTSRNR	CNDQDTRTSY	240
RIGDTWSKKD	NRGNLLOCIC	TGNGRSHKC	ERHTSVQTTT	SGSGPFTDVR	AAVYQPQPHP	300
QPQPPYGHCVT	DSGVVYSVGVM	QWLKTQGNKQ	MLCTCLGNVG	SQETATVQT	YGGNSNGEPC	360
VLPFTYNGRT	FYSCTTEGRQ	DGHLWCSTTS	NYEQDQKYSF	CTDHTVLVQT	RGGNSNGALC	420
HFFFLYNNHH	YTDCTCTSEGRR	DNMWKWCGTQQ	NYDADQKPGF	CPMAAHEEIC	TTNEGVMYRI	480
GDQWDKQHDM	GHMMRCTCVG	NGRQLTCTVIA	YSQLRDRQCCT	DDITYNVNNDT	FHKRHEEGHM	540
LNCCTFCQGQR	GRWKCDPVQD	CQDSETGTTFY	QIGDSWEKVV	HGVRYQCYCY	GRGIGEWHHQ	600
PLQTYPSSSS	PVEVFITETP	SQPNSHPIQW	NAPQPSHISK	YILRWRPKNS	VGRWKEATIP	660
GHLSNSYTIKG	LKPGVYYEGQ	LISIQQYGHQ	EVTRFDFTTT	STSTPVTSNT	VTGETTPFSP	720
LVATSESVTE	TIASSFVVSW	VSASDVSFVSG	RVEYELSEEG	DEPQYLDLPS	TATSVNIPDL	780
LPGRKYIVNV	YQISEDQEQS	LILSTSQTAA	PDAPPDPTVD	QVDDTSIVVR	WSRPQAPITG	840
YRIVYSPSVE	GSSTELNLPE	TANSVTLSDL	QPGVQYNITI	YAVEENQEST	PVVIQQETTG	900
TPRSDFTVSP	RDLQFVEVTD	VKVTIMWTPP	ESAVTGYRV	VIPVNLPGEH	GQRLPISRNT	960
FAEVTVGLSPG	VTVYYPFKVFAV	SHGRESKPL	AQQTTKLDAP	TNLQFVNEDT	STVLRWTPP	1020
RAQITGYRLT	VGLTRRQGPR	QYNVGSVSK	YPLRNLLQPA	EYTIVSLVAIK	GNQESPATG	1080
VFTTLQPGSS	IPPPYNTETVTE	TTIVITWTPTA	PRIGFKLGV	PSQGGEAPRE	VTSDSGSIVV	1140
SGLTPGVEYV	YTIQVLRDQG	ERDAPIVNKV	VTPLSPPTNL	HLEANPDTGV	LTWSWERSTT	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	IRHHPEHFS	RPREQDVPHS	1500
RNSITLTNL	PGTEYVVSIV	ALNGREESPL	LIGQQSTVSD	VPRDLEVVA	TPTSSLISWD	1560
APAVTVRYYR	ITYGETGGNS	PVQEFTVPGS	KSTATISGLK	PGVDYTTIVY	AVTGRGDS	1620
SSKPISINYR	TEIDKPSQMQ	VTVDQDNSIS	VKWLPSSSPV	TGYRVTTTPK	NGPGPTKTKT	1680
AGPDQTEM	EGLQOPTVEYV	VSVYAQNPSG	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 LGGGALGP GG 60
 KPLKPVPGGL AGAGLGAGLG AFPAVTFPGVA LVPGGVADAA AAYKAAKAGA GLGGVPGVGG 120
 LGVSAGAVVP QPGAGVKPGK PPGVGLPGVY PPGVLPGARF PGVGLPGVP TGAGVKPKAP 180
 GVGAGAFAGIP CGVPGFPGQP GPVPLGYPPIKA PKLPGGGYGLP YTTGKLPLYGY GPGGVAGAAG 240
 KAGYPTGTGV GPQAAAAAAA KAAAKFGAGA AGVLPGVGGA GPVGPVGAIP GIGGIAGVGT 300
 PAAAAAAA AAKAAYGAAA GLVPGCPGPQ PGVVGVPAGG VPGVGVPAGG IPVVPAGIIP 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
 GVAGVPGFAG 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
 LHNHHNNLTL SVGPLPKSLE DLQLTHNNKIT 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 AEEELCSGKPF FDAFTDLKNG SLFAFRGQYC 180
 YELDEKAVRP GYPKLI RDVW GIEGPIDAAF TRINCQGKTY LFKGSQYWRF EDGVLDPDYP 240
 RNISDGFDGI PDNVDAALAL PAHSYSGRER VYFFKGKQYW EYQFOHQPSQ EECECGSSLSA 300
 VFEHFAMMQR DSWE DIFELL FWGRTSAGTR QPQFISRDWH GVPQVDAAM AGRIYISGMA 360
 PRPSLAKKQR FRHRNRKGYR SQRGHSGRGRN QNSRRPSRAT WLSLFSSSES 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 PLTGDCINQE PKDSSPAEEC DDCDSCVMTL LNLDLATMGEQ LRLVKSQQLQG 240
 LSASAGLLEQ MRHMETQAKD LRNQLLN YRS AISNHGSKIE GLERE LTL DNL QEFETLQEK 300
 QVNSRKAQTL NNNVNRATQS AKELDVKIKN VIRNVHILLK QISGTDGEQN NVPSGDFSRE 360
 WAEAQRMMRE LRNRNPGKHL REAEADKRES QLLN RIRIWT 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	RLNYTGATG	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	NFQLDKSPLY	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
EWQMKCCCKD	SRQPHNYYSH	REVNENSSG	PMRWWQSQND	VNPVSLQLDL	DRRFQQLQEV	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	QWSQRHQDV	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	NDPHDPRSCK	PCPCCHNGFSC	SVMPETE	480
CNNCPPGVTG	ARCELCAADCY	FGDPFGHEGP	VRPCQPCQCN	NNVDPASGN	CDRLTGRC	540
CIHNTAGIYC	DQCKAGYFGD	PLAPNPADKC	RACNCNPMSG	EPVGCRSDGT	CVCKPGFFGP	600
NCEHGAFCSCP	ACYNQVKIQM	DQFMQQLQRM	EALISKAQGG	DGVVPDTELE	GRMQQAEGQAL	660
QDILRDAQIS	EGASRSLGLQ	LAKVRSQENS	YQSRLLDDLM	TVERVRALGS	QYQNRVRDTH	720
RLITQMQSL	AESEASLGN	NIPASDHVY	PNGFKSLAQE	ATRLAESHVE	SASNMEQLTR	780
ETEDYSKQAL	SLVRKALHEG	VGSQSGSPDG	AVVQGLVEKL	EKTKSLAQQL	TREATQAEIE	840
ADRSYQHSLR	LLDSVSRQG	VSDQSFQVEE	AKRIKQKADS	LSSLVTRHML	EFKRTQKNLG	900
NWKEEAQQLL	QNGKSGREKS	DQLLSRANLA	KSRAQEALSM	GNATFYEVES	ILKNLREFDL	960
QVDNRKAEE	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
ITSSTDNIEAA EENISLDLIQ QYLYTFNFDN EPENISIENL SSDIIQCLEL MPNIERFPNG 540
KKYELDKYTM PHYLRAQEFE HGKSRIALTN SVNEALLNPS RVYTFFSSDY VKVKVNKATEA 600
AMPLGUVBQL VYDFDTDETSE 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 IPDVGNLSQV VVMKSKNNDQ ITNKCKMNLQ DNNGNDIGFI GFHQFNNIAK 1260
LVASNWNYRNQ 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 CEEYDPPDYL TNDKKNIFLQ TMIKLFNRK SKPLGKEKLL MIINGIPYLG 120
DRRVPLPEEFN TNIASTVTN LISNPGEVER KKGIFANLII FGPGPVLNEN ETIDIGIQNH 180
FASREGFGGI MQMKFCPEYV SFVNNVQENK GASIFNRRGY FSDPALILMH ELIHVLHGLY 240
G1KVDDLPV PNEKKPFMQS TDAIQAEELY TFQQDPSII TPSTDKSIYD KVLQNFRGIV 300
DRLNKVLVCI SDPNININII KNKFKDKYKF VEDSEGKYSI DVESFDKLYK SLMFGFTETN 360
IAENYKIKTR ASYFSDSLP VVIKNNLDE IYTIEEGPNI SDKDMEKERYR GQNKAINKQA 420
YEEISKEHLA VVYKIQMKSV KAPGICIDV NEDLFFIAKD NSFSDDLSSKN ERIEYNTQSN 480
YIENDFPIE LIIDTDLISK IELPSENTE LTFDNVDPV YEKQPAIKI FTDENTIFQY 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 YTCQCGRAGYQ 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
TKTICIETIK	G TCTWQTVIDG	R CEININGAT	L KSQCCSSLG	A AWGSPCTLC	Q VDPICGKY	900
SRIKGTQCED	I DECEVFPGV	C KNGLCVNTR	G SFKCQCPSP	M TL DATGRIC	L DIRLETCFL	960
RYEDEECLP	I AGHRHMDAC	C CSGVAAWGT	E EEECPCMTR	T PEYEELCPR	G PGPFAKET	1020
NGKPFKDIN	E CKMIPS LCT	H GKRNTIGS	F FKCRCDSGFA	L DSERNCTD	I DEC RIS PDL	1080
CRGRQCVNTP	G DFECKCDEC	Y EGFM MMK	C MD IDEC QRD	P LLRCGGVCH	N TE GSYRCEC	1140
PPGHQLSPNI	S ACIDINECE	L SAHLCPNGR	C VN LIGKYQC	A CNPGYHSTP	D RLFCVDIDE	1200
CSIMNGCET	F CTN SEGSY	C CSCQFCALM	P DQRSCTDID	E ECDNP NIC	G GGQCTNIPGE	1260
YRCLCYDGF	A SEDMKTCD	V NECDLNPNI	C LSGTCENTK	G SFHC DMD	Y SGKKGTG	1320
TDINECEIGA	H NC GKH AVCT	N TAGSFKSC	S PGWIGDGK	T CTDLDEC SNG	TH MCSQHAD	1380
KNTMGSYRL	C KEGYTGDGF	T CTDLDECSE	N NL CNGNC	L NAPGGYRCE	C DMGFVPSAD	1440
GHAKEDIDE	S SLPNICVFGT	C HNL PGLFRC	E C EIGYELDR	S GGNCTDVNE	C LDPTTCISG	1500
NCVNTPGSYI	D CDCPPDFELN	P TRVGVDT	S GN CYLD IRP	R GDNGD TACS	N EIGVGVSKA	1560
SCCCSLGKAW	G TPCEMCPAV	N TS EYKILCP	G GEGFRPNPI	T VILED IDEC	Q ELPGLCQGG	1620
KCINTFGSQ	C CRCPTGYYL	N EDTRV CDDVN	E ECETPGICGP	G TCYNTVGN	T CICPPDYM	1680
VNGNNCMMD	R RSLCYRNYY	A ADN QTC DGE	L LFNM TKKMC	C CSY NI GR	A WKN KPC ECP	1740
TDEFATLCGS	Q QR PGFV	I T GLPV DIDE	R EIPGV CENG	V CIN MVG SFR	C ECP VGF	1800
DKL LV CEDID	E CQ NGPVCQR	N AE CINTAGS	Y RDC DKPGY	F TSTG QCN DR	N EC QEP IPN	1860
SHGQCIDTVG	S FY CLCH TGF	K TNDD QT MCL	D INECER DAC	G NGT CRNTIG	S FNCR CNHGF	1920
I LSHNN DCID	V D ECA SGN	N LCRN QCN C	V GS FQ CQ CNE	G YEV A P D G	R T CVD IN ECL	1980
PRKCAPGTCQ	N L DGS YRC	C PPG YLS QNE	K C E D I D C V E	P E I C A L G T C S	N T E G S F K C L	2040
PEGFSLSSSG	R RC QD L R	S ZY AK PEGG K	S S P K S R N H S K	Q E C C C A L K G E	G W G D P C E L C P	2100
TEPDEAFRQI	C PY GSG II VG	P DD S A V D M D E	C K E P D V C K H G	Q C I N T D G S Y R	C E C P F G Y I L A	2160
GNECVDTDEC	S VGN PC GNG	T C K N V I G G F C	T C B E E G F E P G P	M M T C E D I N E C	A Q N P L L C A F R	2220
CVN TYG S Y E C	K C P V G Y V L R	D R R M C K D E D E	C E E G K H D C T E	K Q M E C K N L I G	T Y M C I C G P Y	2280
QRRPDGE GCV	D E N E C Q T K P G	I C EN G R C L N T	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 C F	2340
TEVLQNMQCI	G S S N R N P V T K	S E C C D G G R G	W G P H C E I C P F	Q G T V A F K K L C	P H G R G F M T N G	2400
ADIDECKVIH	D V C R N G E C V N	D R G S Y H C I C K	T G Y T P D I T G T	S C V D L N E C N Q	A P K C P N F I C K	2460
NTEG S Y Q C S C	R S C K D L D E C A	R S C K D L D E C A	T K Q H N C Q F L C	V N T I G G F T C K	C P P G F T Q H H T	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 E S L D Q	T G S 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 E N E	C L S 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 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	L P A L T L T 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 LY 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						
MGRRRRLCQ	L YFLWLGCVV	L W A Q Q T A G Q P	Q P P P P K P P R P	Q P P P Q Q V R S A	T A G S E G G F L A	60
PEYREEGAAV	A S R V R R R G Q Q	D V L R G P N V C G	S R F H S Y C C P G	W K T L P G G N Q C	I V P I C R N S C G	120
DGFCSPRNMC	T C S S Q C I S S T	C G S K S I Q Q C S	V R C M N G G T C A	D D H C Q C Q K G Y	I G T Y C Q Q P V C	180
ENG CQ NG G R C	I G P N R C A C V Y	G F T G P Q C E R D	Y R T G P C F T O V	N N Q M C Q Q Q L T	G I V C T K T L C C	240
ATIGRAWHGP	C E M C P A Q P Q P	C R R G F I P N I R	T G A C Q D V D E C	Q A I P G I C Q G G	N C I N T V G S F E	300
CRC PAGH K Q S	E T T Q K C E D I D	E C S I I P G I C E	T G E C S N T V G S	Y F C V C P R G Y V	T S T D G S R C I D	360
QRTG M C F S G L	V N G R C A Q E L P	G R M T K M Q C C C	E P G R C W G I G T	I P E A C P V R G S	E E Y R R L C M D G	420
LPMGGI P G S A	G S R P G G T G G N	G F A P G S A	Y G P G G T G F I P	I P G G N G F S P G	V G G A G V G A G G	480
QGP I IT GLT I	L N Q T I D I C K H	H A N 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
SNP CT N G D C V	N T P G S Y Y C K C	H A G 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	D G K 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 D C D P G L A V	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	720
G A V T K S E 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
NG 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	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 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	DEDPNCLFG	SCTNTPGGFQ	CLCPPGFVLS	DNGRRCFDTR	2100
QSFCFTNFEN	GKCSVPKAFN	TTKAKCCCSK	MPGEGWGDPC	ELCPKDDEVA	FQDLCPYGHG	2160
TVPSLHDTR	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	CCDGGRWGH	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	SALLCMAGQQ	GRWDGALEAA	GPGRVRRRGS	PGILOGPNV	60
GSRFHAYCCP	GWRTFPGRSQ	CVVPICCRAC	GEFGCSQP	N CTCADGTLAP	SCGVSRGSGC	120
SVSCMNGTC	RGASCLCQKG	YTGTVCQPI	CDRGCHNGRR	C1GPNRCA	YGFMGQ	180
DYRTGPGFGQ	VGPEGCQHQL	TGLVTKALC	CATVGR	AWGL PCELC	PCRRGFIPN	240
HTGACQDVDE	CAQAVPLCQG	GSCVNMVGFS	HCRCPVGHRL	SDSSAAC	EDY RAGACF	300
GGRGAGDLAG	HYTRRQCCCD	RGRCWAA	GPV	PELCPPRGSN	EFQQLCA	360
PGLLGFNSNG	MGPPLGPARL	NPHGSDARGI	PSLGPGNSN	GTATLNQ	TID ICRHFTNL	420
NGRCLPTPSS	YRCECNVGYT	QDVRGECIDV	DECTSSPCH	GDCVNIPGTY	HCRCY	480
TPTRQACVVD	DECIVSGGLC	HLGRCVNT	TEG SFQCV	CNAGF ELSPDGK	NCV DHNECAT	540
CVNGVCLNED	GSFSCLCQKG	FLLAPGGH	YC MDI	DECQTPG	EGSFRCQCL	600
GLAVGTDGRV	CVDTHRSTC	YGAIEKGSCA	RPFPGTVK	TS ECCC	CANPDHG FGEPCOLCP	660
KDSAEEQALC	SSSLGITTDG	RDINECALDP	EVCANGVC	EN LRGSY	RCVN LGYEA	720
DCTDVDECAL	NSLLCDNGWC	QNSPGSYSCS	CPPGFHF	WQD TEICKD	VDE LSSPCVSG	780
RNLAGSTCKT	CGPGSRSLDP	GTFCLDSTKG	TWLKI	QBSR CEVNLQGASL	RSECCATLGA	840
AWGSPCRCE	IDPACRGFA	RMTGTV	CDV DDV	NECESFP	GVC PNGRCVNTAG	900
MLDASGRLCV	DVRLEPCFLR	WDEDEC	GVTL PGKYR	MDV V	C D S F R C E P E G L	960
EFA	ASRDFL	LSG RPLF	YKDVNEC	KVFPGLC	HTH TCRN TGSF	1020
AQERNCTDID	ECRISPDL	QGTCV	NTPGS	FECECF	PGY H C I D V	1080
LCRGGTCTNT	DGSYKCQCP	GHELTAKG	TA CED	IDECS L	DGLCPHGC	1140
HAGFQSTPDR	QGCVDINCR	VQNGGCDV	H INTEG	SYRCS CGQG	YSLMPD GRACAD	1200
EENPRVCDQG	HCTNMPGGHR	CLCYDGF	MAT PDMR	TCDV DVD	PD ECDLNP	1260
FVCHCQLGYM	VRKGATGCS	VDECEVGGH	N CDSH	ASCLN PGFS	RCR CLP GWVGD	1320
DLDECVSQEH	RCS	PRGDCLN	VPGSY	RCR CTQ GFAGD	GFFC EDR	1380
APG	GYCE	MGFDPTEDH	ACQD	VDECAQ GNLCAF	GSC E NLPGM	1440
GNCTDINECA	DPVNCINGVC	INTPGS	CPQDFELN	PS	PSV GVGCDV	1500
DSGISCSAEI	CGVGVTRASCC	CSL RAWGNP	CELC	PMANTT YRTLCP	PGGE GFQPNR	1560
LED	IDE	QCGGDC	NTFGS	PPG YHLSE	HT RIC E D I E C S	1620
CYNTLGN	YTC	CPCE	FCE	TCR H	TH S G I C G P G T	1680
GQAWNR	CEA	PTPIS	QD	YNG	TCR H D R G G	1740
IGSF	RCECPA	GFNYSN	ILL CED	VQD E	GEIPA ICANGIC	1800
GACVGR	NECR	PSHGD	DEGS	QNAD	IC INIPGSYRCK	1860
CKNI	IGSYNC	PL	CMD	TEG	CTR GYKLSPG	1920
TADG	KNCVDT	LC	PGV	QSA	QDA	1980
LF	GTCTNSP	PGV	QD	QTL	CD M D I D E	2040
CSKR	PGEWG	CS	QD	LCR	DR Q PC G N G	2100
VNTDGS	FRCE	CPFGY	SLDFT	FG Q A S A	Q D T L C M D I D E	2160
MMTC	EIDEC	SLNPLL	CAFR	TC	CD R Q P E P G L	2220
RGME	CNKLIG	CHNTEG	GSYLC	PA G Y	T R E DG A M C R D V D E	2280
ECHDIR	QGP	RPL	PGS	G E G	C A D Q Q D C H A	2340
LPGT	SAYRKL	CPHGS	GYTAE	T D D	N E C S E A V T	2400
TTCL	DMDECS	QVP	KPCTFLC	N S L S S	R A E C C C G G R	2460
CVNTV	GAFTC	RCP	PGT	Q D S	Q W G P R C E L C P	2520
SSGH	GCEDVN	PGT	QHH	Q A C	Q G Y T P D A T A	2580
SASC	RNTLGG	RC	PPG	Q F T Q H	E C A L S P P T C G	2640
FRAG	QGHCVS	RC	PGQ	Q S W A	Q S Y C A N T P G	2700
TLDSE	ALLTL	GLG	FSPGPQD	T D Q A L	G F L C G C P Q G Y	2760
LQLGRR	RRGP	TPD	KEELLSS	E A C Y E C K I N G	L S P R D R P R S A H R D	2809
QL	RRGP	G	YRLEV	V V S H	H Q V N L A	
QL	QL	QL	QL	QL	QL	

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).

* * * * *