



US012384823B2

(12) **United States Patent**
Lin

(10) **Patent No.:** US 12,384,823 B2
(45) **Date of Patent:** Aug. 12, 2025

(54) **VGLL4 WITH UCP-1 CIS-REGULATORY ELEMENT AND METHOD OF USE THEREOF**

(71) Applicant: **MASONIC MEDICAL RESEARCH LABORATORY**, Utica, NY (US)

(72) Inventor: **Zhiqiang Lin**, Utica, NY (US)

(73) Assignee: **MASONIC MEDICAL RESEARCH LABORATORY**, Utica, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 761 days.

(21) Appl. No.: **17/731,201**

(22) Filed: **Apr. 27, 2022**

(65) **Prior Publication Data**

US 2022/0259276 A1 Aug. 18, 2022

Related U.S. Application Data

(63) Continuation of application No. 16/925,632, filed on Jul. 10, 2020, now Pat. No. 11,319,354.

(60) Provisional application No. 62/872,624, filed on Jul. 10, 2019.

(51) **Int. Cl.**

C07K 14/47 (2006.01)
C12N 15/86 (2006.01)

(52) **U.S. Cl.**

CPC **C07K 14/4705** (2013.01); **C12N 15/86** (2013.01); **C12N 2750/14143** (2013.01); **C12N 2830/001** (2013.01)

(58) **Field of Classification Search**

CPC C07K 14/4705; C07K 14/4702; C12N 15/86; C12N 2750/14143; C12N 2830/001; C12N 2830/008; A61K 48/0058; A01K 2207/25; A01K 2227/105; A01K 2267/0362

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,139,941 A	8/1992	Muzyczka et al.
5,436,146 A	7/1995	Shenk et al.
5,446,143 A	8/1995	Simpson et al.
5,464,758 A	11/1995	Gossen et al.
5,478,745 A	12/1995	Samulski et al.
5,741,683 A	4/1998	Zhou et al.
6,001,650 A	12/1999	Coloski
6,057,152 A	5/2000	Samulski et al.
6,136,597 A	10/2000	Hope et al.
6,156,303 A	12/2000	Russell et al.
6,165,782 A	12/2000	Naldini et al.
6,204,059 B1	3/2001	Samulski et al.
6,207,455 B1	3/2001	Chang
6,218,181 B1	4/2001	Verma et al.
6,268,213 B1	7/2001	Samulski et al.
6,277,633 B1	8/2001	Olsen
6,323,031 B1	11/2001	Cichutek
6,358,732 B1	3/2002	Sedlacek et al.

6,432,705 B1	8/2002	Yee et al.
6,491,907 B1	12/2002	Rabinowitz et al.
6,531,456 B1	3/2003	Kurtzman et al.
6,596,535 B1	7/2003	Carter
6,660,514 B1	12/2003	Zolotukhin et al.
6,951,753 B2	10/2005	Shenk et al.
7,056,502 B2	6/2006	Hildinger et al.
7,094,604 B2	8/2006	Snyder et al.
7,125,717 B2	10/2006	Carter
7,172,893 B2	2/2007	Rabinowitz et al.
7,198,951 B2	4/2007	Gao et al.
7,201,898 B2	4/2007	Monahan et al.
7,220,577 B2	5/2007	Zolotukhin
7,229,823 B2	6/2007	Samulski et al.
7,235,393 B2	6/2007	Gao et al.
7,282,199 B2	10/2007	Gao et al.
7,319,002 B2	1/2008	Wilson et al.
7,439,065 B2	10/2008	Ferrari et al.
7,456,683 B2	11/2008	Takano et al.
7,790,449 B2	9/2010	Gao et al.
8,455,204 B2	6/2013	Boss et al.
8,852,939 B2	10/2014	Hall et al.
9,034,839 B2	5/2015	Thibonnier
2002/0065239 A1	5/2002	Caplan et al.
2002/0076395 A1	6/2002	Crystal et al.
2003/0138772 A1	7/2003	Gao et al.
2003/0148506 A1	8/2003	Kotin et al.
2003/0219409 A1	11/2003	Coffin et al.

(Continued)

FOREIGN PATENT DOCUMENTS

CN	104548131 B	4/2019
EP	1995309 A1	11/2008

(Continued)

OTHER PUBLICATIONS

Zhang, Y., et al., "A growing role for the Hippo signaling pathway in the heart: Hippo pathway function in the heart," Journal of Molecular Medicine, vol. 95, No. 5, pp. 465-472 (2018).

Zhang et al., "The TEA domain family transcription factor TEAD4 represses murine adipogenesis by recruiting the cofactors VGLL4 and CtBP2 into a transcriptional complex", Journal of Biological Chemistry, vol. 293, No. 44, pp. 17119-17134, Sep. 12, 2018.

Cassard, A., et al., "Human Uncoupling Protein Gene: Structure, Comparison With Rat Gene, and Assignment to the Long Arm of Chromosome 4", Journal of Cellular Biochemistry, vol. 43, pp. 255-264 (1990).

Jiao, S., et al., "A Peptide Mimicking VGLL4 Function Acts as a YAP Antagonist Therapy against Gastric Cancer", Cancer Cell, vol. 25, pp. 166-180 (2014).

Gonzalez-Barroso, M., et al., "Transcriptional Activation of the Human ucp1 Gene in a Rodent Cell Line", The Journal of Biological Chemistry, vol. 275, No. 41, pp. 31722-31732 (2000).

(Continued)

Primary Examiner — Paul J Holland

(74) Attorney, Agent, or Firm — Heslin Rothenberg Farley & Mesiti P.C.

(57) **ABSTRACT**

Provided is a polynucleotide, including a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein, wherein the cis-regulatory element includes an uncoupling protein 1 enhancer and an uncoupling protein 1 promoter. Also provided is a viral vector including said polynucleotide. Also provided is a method of transfecting a cell or a subject with said polynucleotide or said viral vector.

20 Claims, 24 Drawing Sheets

Specification includes a Sequence Listing.

(56)

References Cited**U.S. PATENT DOCUMENTS**

2004/0055023	A1	7/2004	Gao et al.
2005/0187154	A1	8/2005	Kahn et al.
2008/0075740	A1	3/2008	Gao et al.
2010/0216709	A1	8/2010	Scheule et al.
2010/0240029	A1	9/2010	Guarente et al.
2011/0166210	A1	7/2011	Felber et al.
2012/0040401	A1	2/2012	Ellis et al.
2016/0319303	A1	11/2016	Jiménez Cenzano et al.
2017/0290926	A1	10/2017	Smith et al.
2021/0009646	A1	1/2021	Lin
2021/0017606	A1	1/2021	Li et al.

FOREIGN PATENT DOCUMENTS

EP	2394667	A1	12/2011
EP	2492347	A1	8/2012
EP	3101125	A1	12/2016
WO	1996040954	A1	12/1996
WO	199749827	A2	12/1997
WO	1997049827	A2	12/1997
WO	1998009524	A1	3/1998
WO	1998011244	A2	3/1998
WO	1999061601	A2	12/1999
WO	2000028061	A2	5/2000
WO	2001083692	A2	11/2001
WO	2001091803	A2	12/2001
WO	2001094605	A1	12/2001
WO	2003014367	A1	2/2003
WO	2003042397	A3	5/2003
WO	2003052051	A2	6/2003
WO	2003052052	A2	6/2003
WO	2006110689	A2	10/2006
WO	2007127264	A2	11/2007
WO	2008021290	A2	2/2008
WO	2008103755	A1	8/2008
WO	2010010887	A1	1/2010
WO	2011154520	A1	12/2011
WO	2012007458	A1	1/2012
WO	2013063379	A1	5/2013
WO	2014023808	A2	2/2014
WO	2015173308	A1	11/2015
WO	2016193431	A1	12/2016
WO	2017127750	A1	7/2017
WO	2018060097	A1	4/2018
WO	2018215613	A1	11/2018

OTHER PUBLICATIONS

- Deng, X., et al., "VGLL4 is a transcriptional cofactor acting as a novel tumor suppressor via interacting with TEADs", *Am J Cancer Res.*, vol. 8, No. 6, pp. 932-943 (2018).
- "Minutes of the 48th General Assembly of the European Association for the Study of Diabetes", *Diabetologia*, vol. 56, Suppl 1, S1 and S319 (2013).
- Cassard-Doulcier, A., et al., "A 211-bp enhancer of the rat uncoupling protein-1 (UCP-1) gene controls specific and regulated expression in brown adipose tissue", *Biochem. J.*, vol. 333, pp. 243-246 (1998).
- Chen, H., et al., "Vgl-4, a Novel Member of the Vestigial-like Family of Transcription Cofactors, Regulates α 1-Adrenergic Activation of Gene Expression in Cardiac Myocytes", *The Journal of Biological Chemistry*, vol. 279, No. 29, pp. 30800-30806 (2004).
- Kozak, U.C., et al., "An Upstream Enhancer Regulating Brown-Fat-Specific Expression of the Mitochondrial Uncoupling Protein Gene", *Molecular and Cellular Biology*, vol. 14, No. 1, pp. 59-67 (1994).
- Larose, M., et al., "Essential cis-Acting Elements in Rat Uncoupling Protein Gene Are in an Enhancer Containing a Complex Retinoic Acid Response Domain", *vol. 271, No. 49, pp. 31533-31542 (1996).*
- Rim, J.S., et al., "Regulatory Motifs for CREB-binding Protein and Nfe212 Transcription Factors in the Upstream Enhancer of the Mitochondrial Uncoupling Protein 1 Gene", *The Journal of Biological Chemistry*, vol. 277, No. 37, pp. 34589-34600 (2002).
- Shore, A., et al., "Role of Ucp1 enhancer methylation and chromatin remodelling in the control of Ucp1 expression in murine adipose tissue", *Diabetologia*, vol. 53, pp. 1164-1173 (2010).
- Skarnes, W.C., et al., "A conditional knockout resource for the genome-wide study of mouse gene function", *Nature*, vol. 474, pp. 337-344 (2011).
- Gao et al., "Computational insights into the interaction mechanism of transcription cofactor vestigial-like protein 4 binding to TEA domain transcription factor 4 by molecular dynamics simulation and molecular mechanics generalized Born/surface area) calculation", *Journal of Biomolecular Structure & Dynamics*, vol. 37, No. 10, pp. 2538-2545, Nov. 9, 2018.
- Jiao et al. (2014), A Peptide Mimicking VGLL4 Function Acts as a YAP Antagonist Therapy against Gastric Cancer, *Cancer Cell*, 25:166-180.
- Monahan et al., "AAV vectors: is clinical success on the horizon?" *Gene Therapy*, 2000, vol. 7: pp. 24-30.
- Gonzalez-Barroso et al., Transcriptional Activation of the Human ucp 1 Gene in a rodent cell line: synergism of retinoid, isoprotenerol and thiazolidinedione is mediated by a multipartite response element, *The Journal of Biological Chemistry*, 2000, vol. 275(41): pp. 31722-31732.
- Ahi et al., "Adenoviral Vector Immunity: its implication and circumvention strategies," *Current Gene Therapy*, 2011, vol. 11(4): pp. 307-320.
- Donello et al., "Woodchuck Hepatitis Virus Contains a Tripartite Postranscriptional Regulatory Element," *Journal of Virology*, 1998, vol. 72(6): pp. 5085-5092.
- Al-Dosari et al., "Evaluation of viral and mammalian promoters for driving transgene expression in mouse liver," *Biochemical and Biophysical Research Communication*, 2006, vol. 339(2): pp. 673-678.
- Zhang et al., "Celastrol enhances AAV1-mediated gene expression in mice adipose tissues," *Gene Therapy*, 2011, vol. 18(2): pp. 128-134.
- Gao et al., Clades of Adeno-Associated Viruses Are Widely Disseminated in Human Tissues. *Journal of Virology*, 2004 vol. 78: pp. 6381-6388.
- Graves et al., Analysis of a Tissue-Specific Enhancer: ARF6 Regulates Adipogenic Gene Expression. *Molecular and Cellular Biology*, 1992, vol. 12(3):pp. 202-1208.
- Amado et al., Lentiviral Vectors—the Promise of Gene Therapy Within Reach?, *Science*, 1999, vol. 285: pp. 674-676.
- Graves et al., "Identification of a potent adipocyte-specific enhancer: involvement of an NF-1-like factor," *Genes & Development*, 1991, vol. 5(3): pp. 428-437.
- Mizukami et al., "Adipose Tissue as a Novel Target for In Vivo Gene Transfer by Adeno-Associated Viral Vectors," *Human Gene Therapy*, 2006, vol. 17: pp. 921-928.
- Apparailly et al., "Adeno-Associated Virus Pseudotype 5 Vector Improves Gene Transfer in Arthritic Joints," *Human Gene Therapy*, 2005, vol. 16(4): pp. 426-434.
- Ayuso et al., Production, Purification and Characterization of Adeno-Associated Vectors, *Current Gene Therapy*, 2010, vol. 10: pp. 423-436.
- Kitajima et al., "Persistent liver expression of murine apoA-I using vectors based on adeno-associated viral vectors serotypes 5 and 1," *Atherosclerosis*, 2006, vol. 186: pp. 65-73.
- Kozak et al., An Upstream Enhancer Regulating Brown-Fat Specific Expression of Mitochondrial Uncoupling Protein Gene, *Molecular and Cellular Biology*, 1994, vol. 14: pp. 59-67.
- Leherz et al., "Novel AAV serotypes for improved ocular gene transfer," *The Journal of Gene Medicine*, 2008, vol. 10: pp. 375-382.
- Lee et al. "Optimizing regulatable gene expression using adenoviral vectors," *Experimental Physiology*, 2004, vol. 90: pp. 33-37.
- Liu et al., "Promoter effects of adeno-associated viral vector for transgene expression in the cochlea in vivo," *Experimental and Molecular Medicine*, 2007, vol. 39: pp. 170-175, 2007.

(56)

References Cited**OTHER PUBLICATIONS**

- Lock et al., "Characterization of a Recombinant Adeno-Associated Virus Type 2 Reference Standard Material," *Human Gene Therapy*, 2010, vol. 21: pp. 1273-1285.
- Mori et al. "Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsid protein," *Virology*, 2004, vol. 330: pp. 375-383.
- Muisse et al., "Adipose Fibroblast Growth Factor 21 is Up-Regulated by Peroxisome Proliferator-Activated Receptor gamma and Altered Metabolic States," *Molecular Pharmacology*, 2008, vol. 74: pp. 403-412.
- Okada et al., Scalable Purification of Adeno-Associated Virus Serotype 1 (AAV1) and AAV8 Vectors, Using Dual Ion-Exchange Adsorptive Membranes, *Human Gene Therapy*, 2009, vol. 20: pp. 1013-1021.
- Rabinowitz et al., "Cross-Packaging of a Single Adeno-Associated Virus (AAV) Type 2 Vector Genome into Multiple AAV Serotypes Enables Transduction with Broad Specificity," *2002*, vol. 76(2): pp. 791-801.
- Bish et al., "Adeno-Associated Virus (AAV) Serotype 9 Provides Global Cardiac Gene Transfer Superior to AAV1, AAV6, AAV7, and AAV8 in the Mouse and Rat," *Human Gene Therapy*, 2008, vol. 19(12): pp. 1359-1368.
- Rival et al., "Human Adipocyte Fatty Acid-Binding Protein (aP2) Gene Promoter-Driven Reporter Assay Discriminates Nonlipogenic Peroxisomes Proliferator-Activated Receptor gamma Ligands," *The Journal of Pharmacology and Experimental Therapeutics*, 2004, vol. 311(2): pp. 467-475.
- Ross et al., "A fat-specific enhancer is the primary determinant of gene expression for adipocyte P2 in vivo," *Proceedings of the National Academy of the Sciences of the United States of America*, 1990, vol. 87: pp. 9590-9594.
- Seale et al., "Prdm 16 determines the thermogenic program of subcutaneous white adipose tissue in mice," *The Journal of Clinical Investigation*, 2011, vol. 121(1): pp. 96-105.
- Boshart et al., "A Very Strong Enhancer is Located Upstream of an Immediate Early Gene of Human Cytomegalovirus," *Cell*, 1985, vol. 41: pp. 521-530.
- Boulous et al., "Assessment of CMV, RSV and SYN1 promoters and the woodchuck post-transcriptional regulatory element in adenovirus vectors for transgene expression in cortical neuronal cultures," *Brain Research*, 2006, vol. 1102: pp. 27-38.
- Boyer et al., "The Mitochondrial Uncoupling Protein Gene in Brown Fat: Correlation between DNase 1 Hypersensitivity and Expression in Transgenic Mice," *Molecular and Cellular Biology*, 1991, vol. 11: pp. 4147-4156.
- Taymans et al., "Comparative Analysis of Adeno-Associated Viral Vector Serotypes 1, 2, 5, 7 and 8 in Mouse Brain," *Human Gene Therapy*, 2007, vol. 18: pp. 195-206.
- Broekman et al., "Adeno-Associated Virus Vectors Serotypes With AAV8 Capsid Are More Efficient Than AAV1 Or -2 Serotypes for Widespread Gene Delivery to the Neonatal Mouse Brain," *Neuroscience*, 2006, vol. 138(2): pp. 501-510.
- Urabe et al., "A novel dicistronic AAV vector using a short IRES segment derived from hepatitis C virus genome," *Gene*, 1997, vol. 200: pp. 157-162.
- Wang et al., "Improved neuronal transgene expression from an AAV-2 vector with a hybrid CMV enhancer/PDGF-B promoter," *The Journal of Gene Medicine*, 2005, vol. 7: pp. 945-955.
- Wu et al., "Effect of Genome Size on AAV Vector Packaging," *Molecular Therapy*, 2010, vol. 18(1): pp. 80-86.
- Cao, "Adipose tissue angiogenesis as a therapeutic target for obesity and metabolic diseases," *Nature Reviews*, 2010, vol. 9: pp. 107-115.
- Zarrin et al., "Comparison of CMV, RSV, SV40 viral and Lambda 1 cellular promoters in B and T lymphoid and non-lymphoid cell lines," *Biochimica et Physica Acta*, 1999, vol. 1446: pp. 135-139.
- Zincarelli et al., Analysis of AAV Serotypes 1-9 Mediated Gene Expression and Tropism in Mice After Systemic Injection. *Molecular Therapy*, 2008, vol. 16(6): pp. 1073-1080.
- Zufferey et al., "Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element Enhances Expression of Transgenes Delivered by Retroviral Vectors," *Journal of Virology*, 1999, vol. 73: pp. 2886-2892.
- Cassard-Doulcier et al., "A 211-bp enhancer of the rat uncoupling protein-1 (UCP-1) gene controls specific and regulated expression in brown adipose tissue," *Biochemical Journal*, 1998, vol. 333: pp. 243-246.
- Casteilla Louise et al., "Virus-based gene transfer approaches and adipose tissue biology," *Current Gene Therapy*, 2008, vol. 8(2): pp. 79-87.
- Ayuso et al., "High AAV vector purity results in serotype and tissue independent enhancement of transduction efficiency," *Gene Therapy*, 2010, vol. 17: pp. 503-510.
- Chao et al., Several Log Increase in Therapeutic Transgene Delivery by Distinct Adeno-Associated Viral Serotype Vectors, *Molecular Therapy*, 2000, vol. 2(6): pp. 19-623.
- Chlorini et al., "Cloning of Adeno-Associated Virus type 4 (AAV4) and Generation of Recombinant AAV4 Particles," *Journal of Virology*, 1997, vol. 71: pp. 6823-6833.
- Daya et al., "Gene Therapy Using Adeno-Associated Virus Vectors," *Clinical Microbiology Reviews*, 2008, vol. 21(4): pp. 583-593.
- Dressman, "AAV-Mediated Gene Transfer to Models of Muscular Dystrophy: Insights Into Assembly of Multi-Subunit Membrane Proteins," University of Pittsburgh, Graduate Faculty of the School of Medicine, Dept. of Biochemistry and Molecular Genetics in partial fulfillment of the requirements for the degree of Doctor of Philosophy. 1997,183 pages.
- Elais et al., Adipose Tissue Overexpression of Vascular Endothelial Growth Factor Protects Against Diet-Induced Obesity and Insulin Resistance, *Diabetes*, 2012, vol. 61: pp. 1801-1813.
- Harms et al., "Brown and beige fat: development, function and therapeutic potential," *Nature Medicine*, 2013, vol. 19 (10): pp. 1252-1263.
- Gustafson et al., Restricted Adipogenesis in Hypertrophic Obesity: The Role of WISP2, Wnt and BMP4, *Diabetes*, 2013, vol. 62(9): pp. 2297-3004.
- Loeb et al., "Enhanced Expression of Transgenes from Adeno-Associated Virus Vectors with the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element: Implications for Gene Therapy," *Human Gene Therapy*, 1999, vol. 10: pp. 2295-2305.
- Hoffman et al., "Increased BMP4 improves insulin sensitivity and increases beige/brown adipogenesis in adult mice. *Diabetologica*," vol. 57, Supp , S1-S366, Sep. 26, 2014.
- Ichimura et al., "Dysfunction of lipid sensor GPR120 leads to obesity in both mouse and human," *Nature*, 2012, vol. 483: pp. 350-354.
- Jimenez et al., "In Vivo Adeno-Associated Viral Vector Mediated Genetic Engineering of White and Brown Adipose Tissue in Adult Mice," *Diabetes*, 2013, vol. 62: pp. 4012-4022.
- Kahn et al., "Mechanism linking obesity to insulin resistance and type 2 diabetes," *Nature*, 2006, vol. 444: pp. 840-846.
- Yee et al., "Subcutaneous adipose tissue fatty acid desaturation in adults with and without rare adipose disorders," *Lipids in Health and Disease*, 2012, vol. 11: pp. 19-30.
- Yu et al., "Protein deacetylation by SIRT 1: an emerging key post-translational modification in metabolic regulation," *Pharmacological Research*, 2010, vol. 62(1): pp. 35-41.
- Qiao et al., "Liver-specific microRNA-122 target sequences incorporated in AAV vectors efficiently inhibits transgene expression in the liver," *Gene Therapy*, 2011, vol. 18: pp. 403-410.
- Kelly et al., "Attenuation of Vesicular Stomatitis Virus Encephalitis through MicroRNA targeting," *Journal of Virology*, 2010, vol. 84(3): pp. 1550-1562.
- Chen et al., "The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation," *Nature Genetics*, 2006, vol. 38: pp. 228-233.
- Zhang et al., "Adipose Tissue Transduction Using AAV8-based Vectors: Inadvertent Gene Transfer into Liver," *Molecular Therapy*, 2005, vol. 11, supp 1, article 862.

(56)

References Cited

OTHER PUBLICATIONS

Yan et al., "Inverted Terminal Repeat Sequences are Important for Intermolecular Recombination and Circularization of Adeno-Associated Virus Genomes," *Journal of Virology*, 2005, vol. 79: pp. 364-379.

Card et al., MicroRNA Silencing Improves the Tumor Specificity of Adenoviral Transgene Expression, 2012, vol. 19: pp. 451-459.

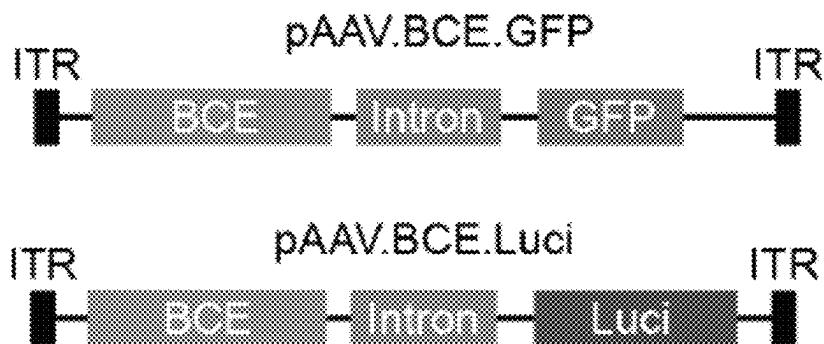


FIG. 1

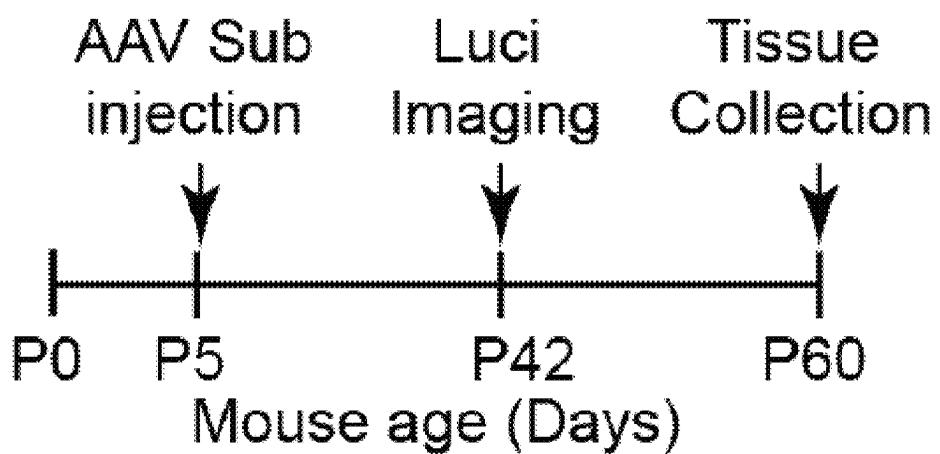


FIG. 2

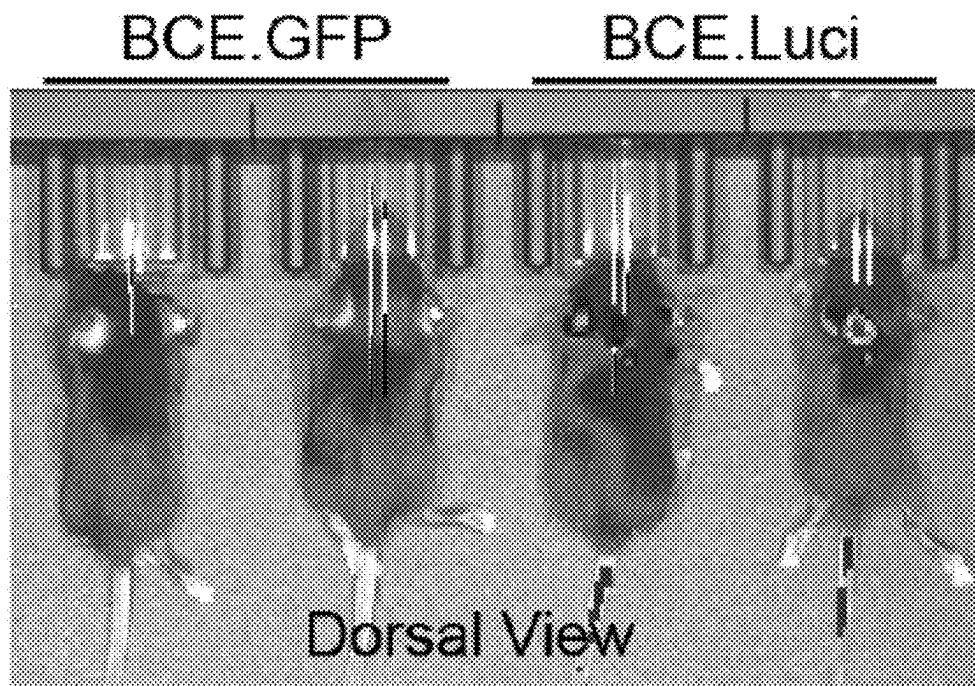


FIG. 3

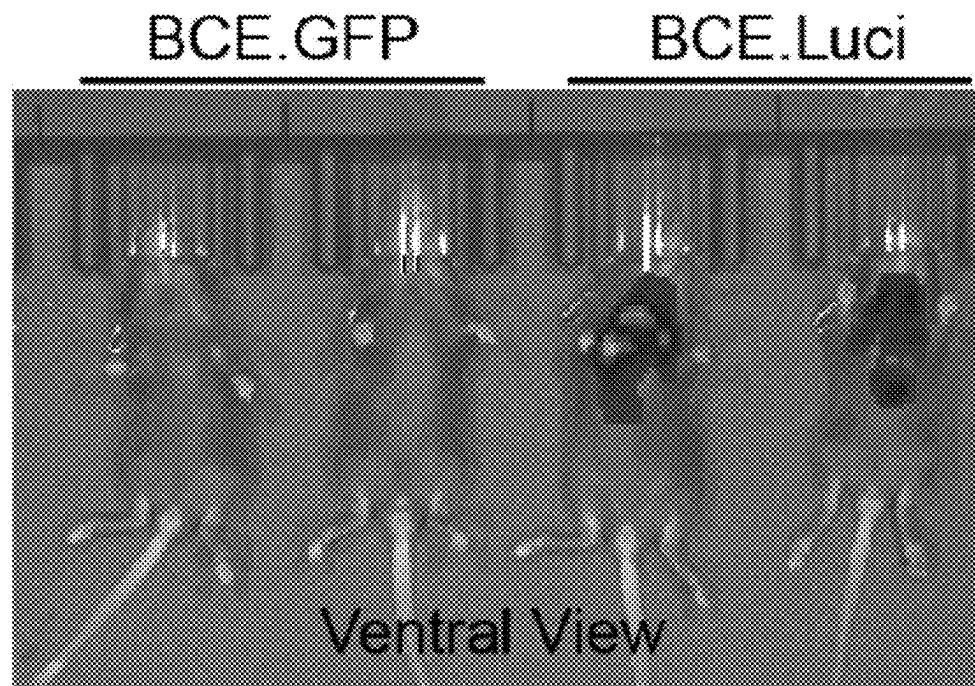


FIG. 4

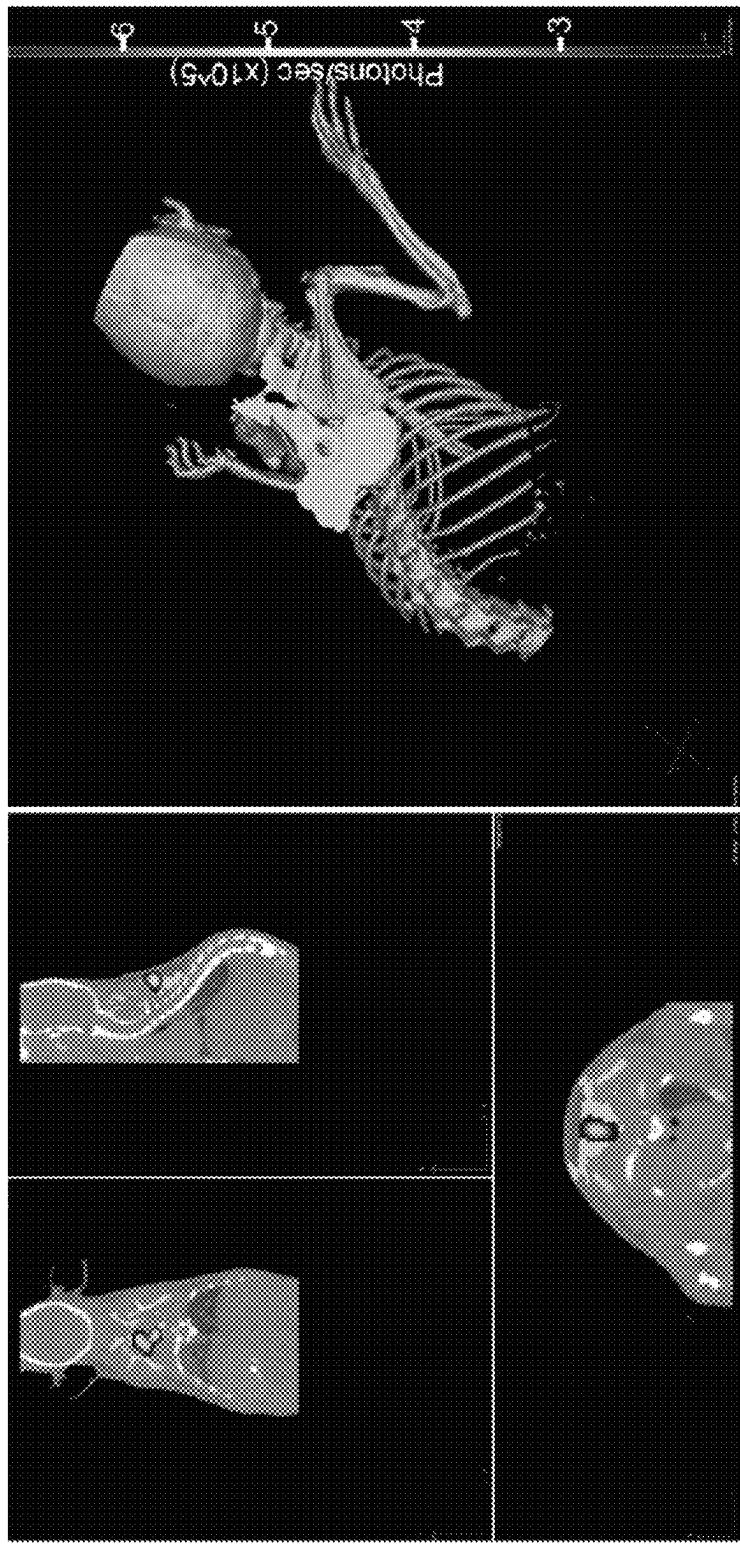


FIG. 5

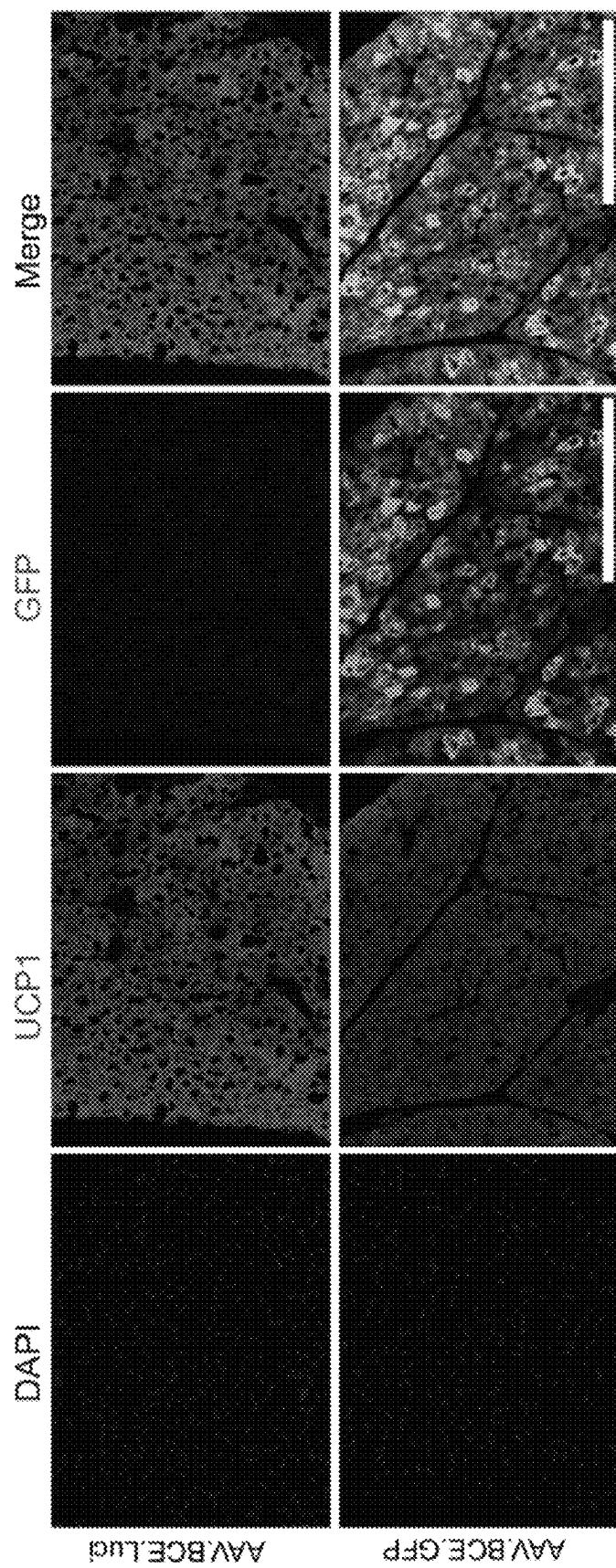


FIG. 6

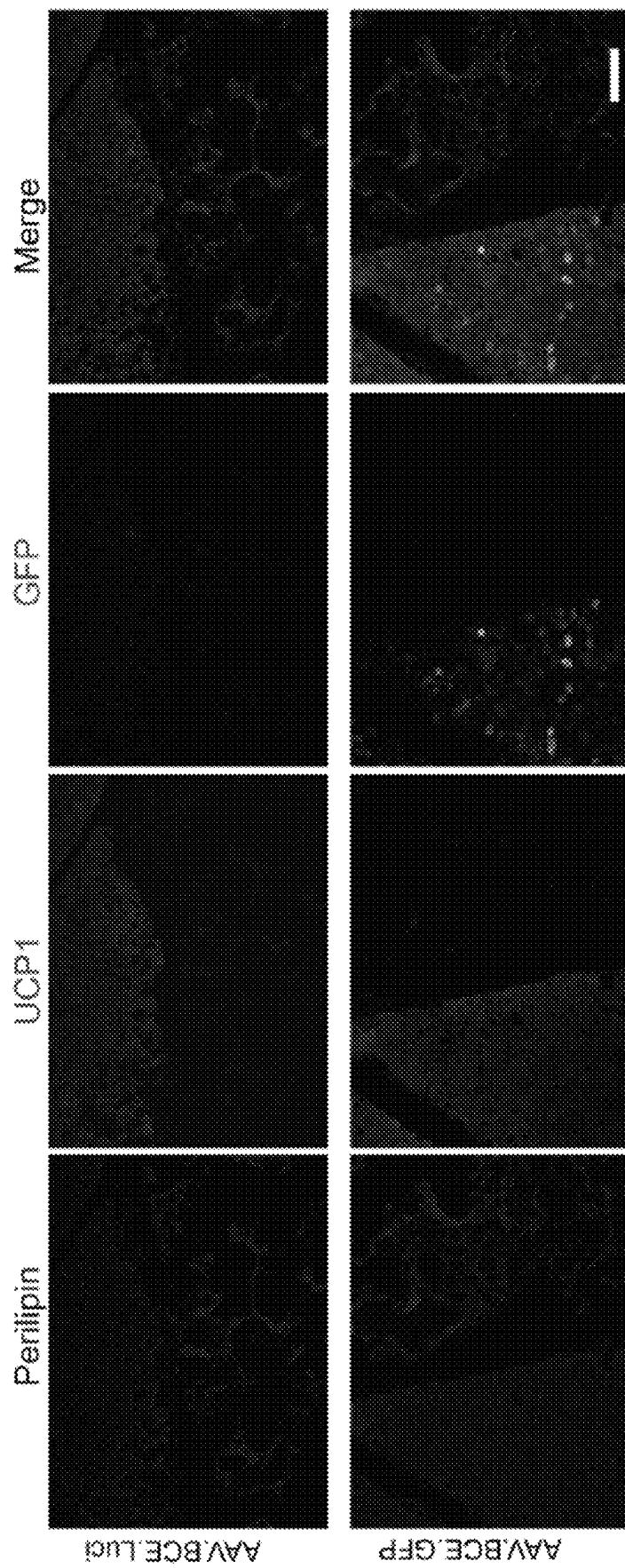


FIG. 7

pAAV.BCE.VgII4-GFP



FIG. 8

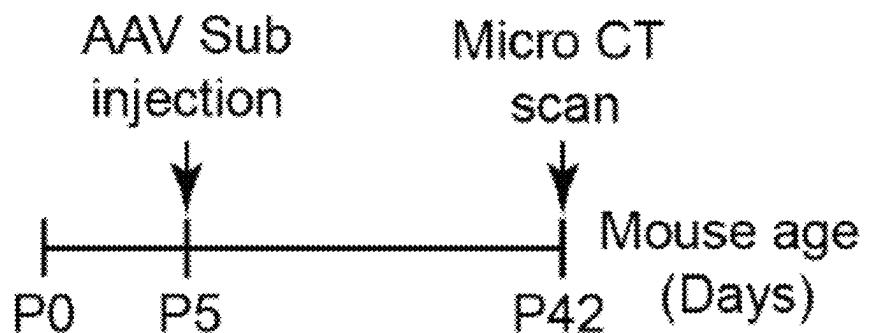


FIG. 9

AAV9.BCE.GFP

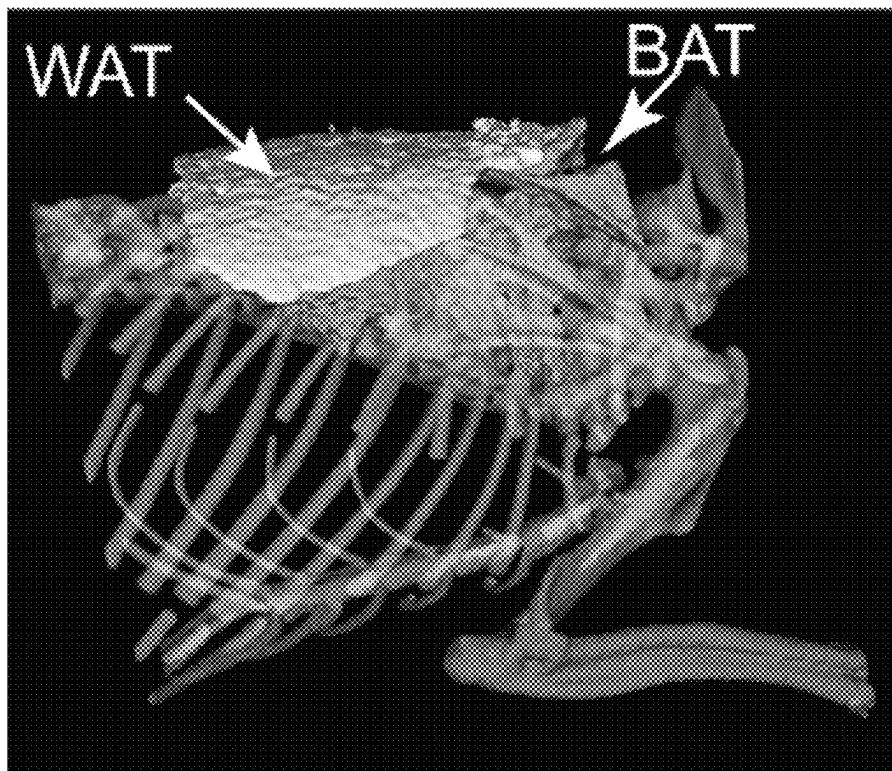


FIG. 10

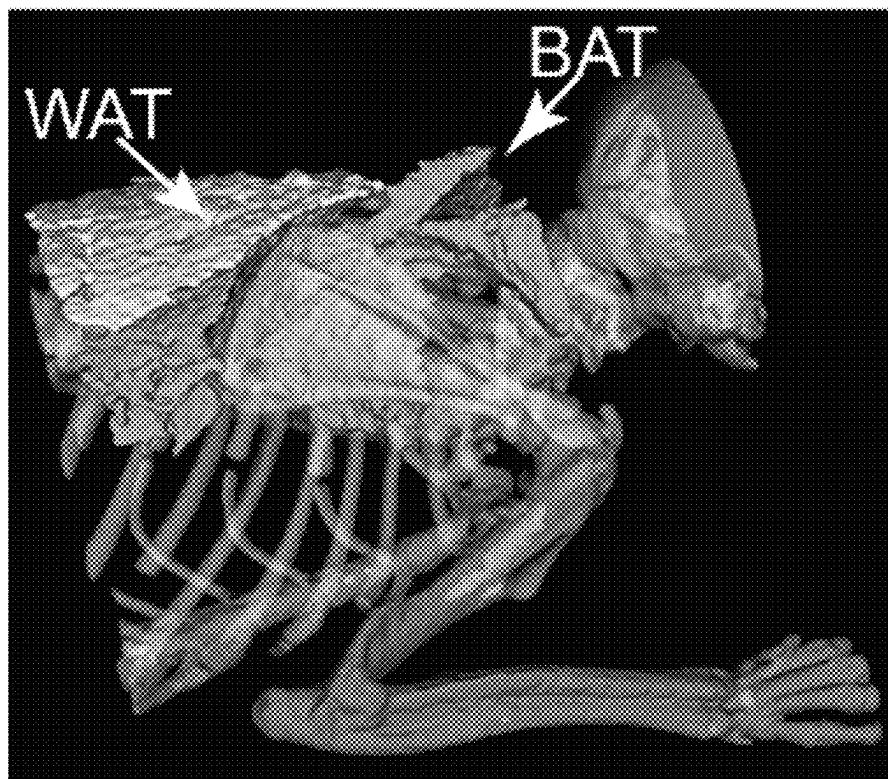
AAV9.BCE.Vgll4-GFP

FIG. 11

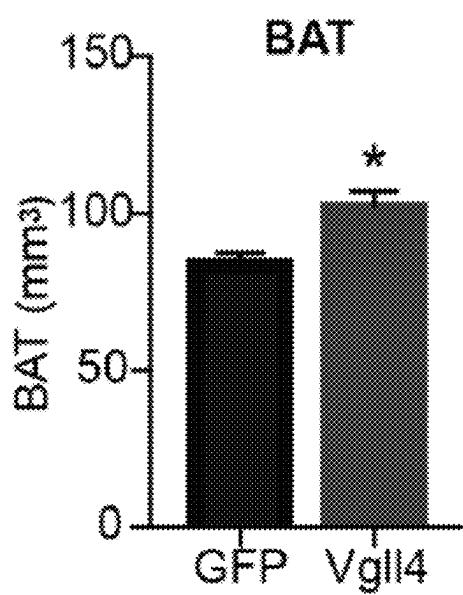


FIG. 12

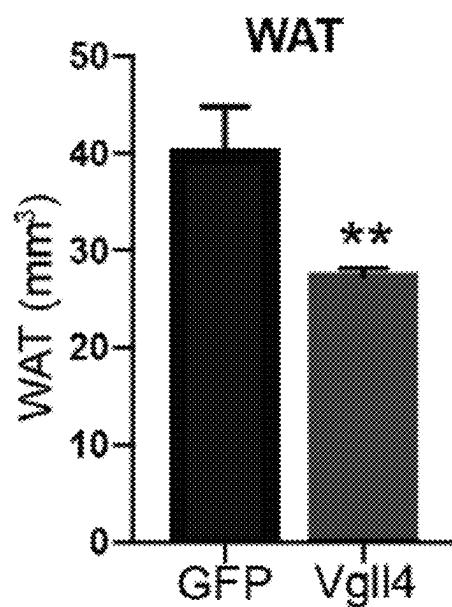


FIG. 13

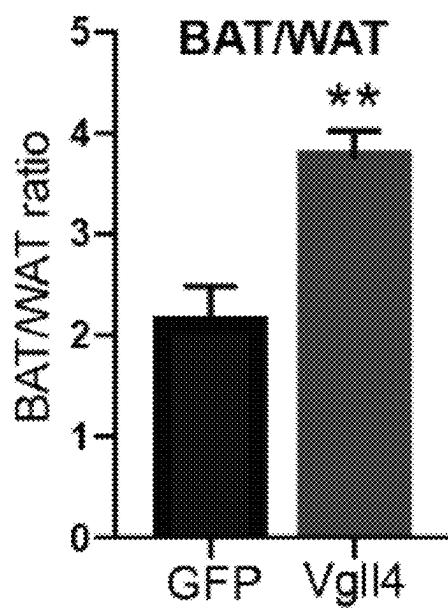


FIG. 14

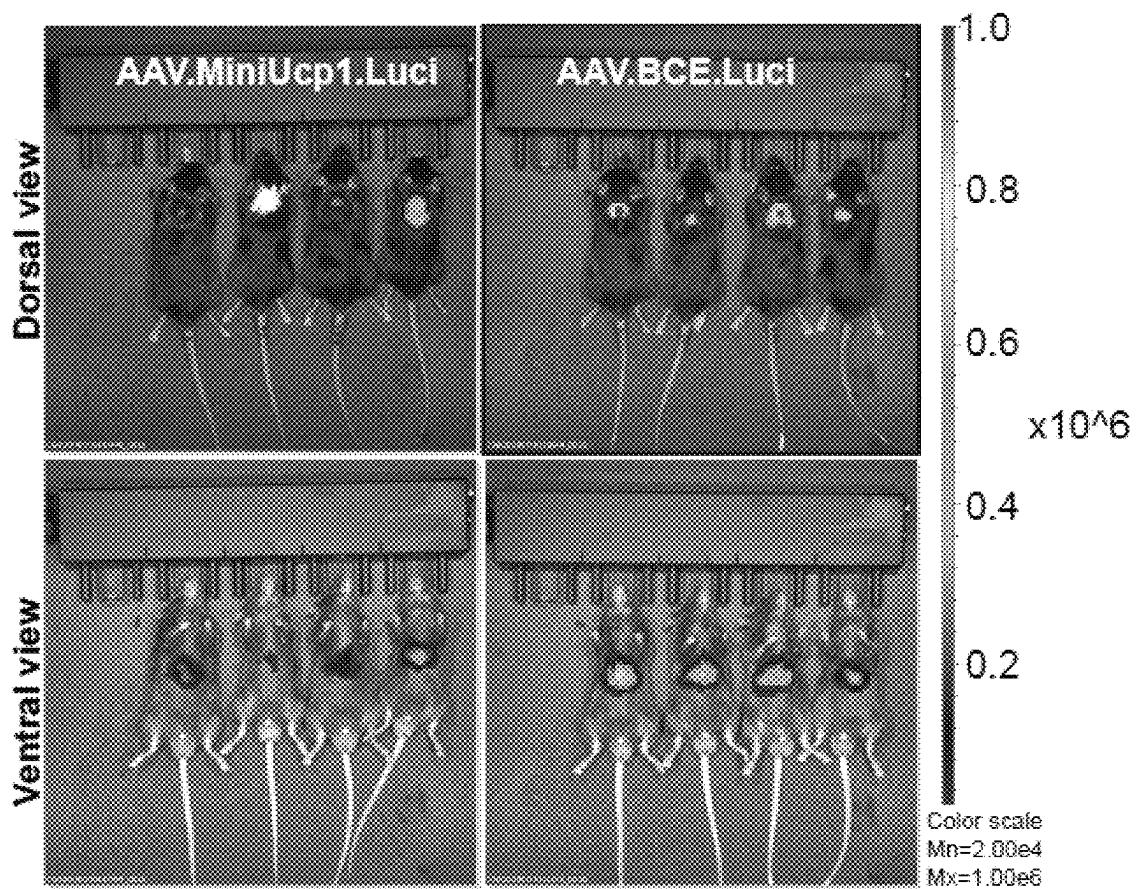


FIG 15

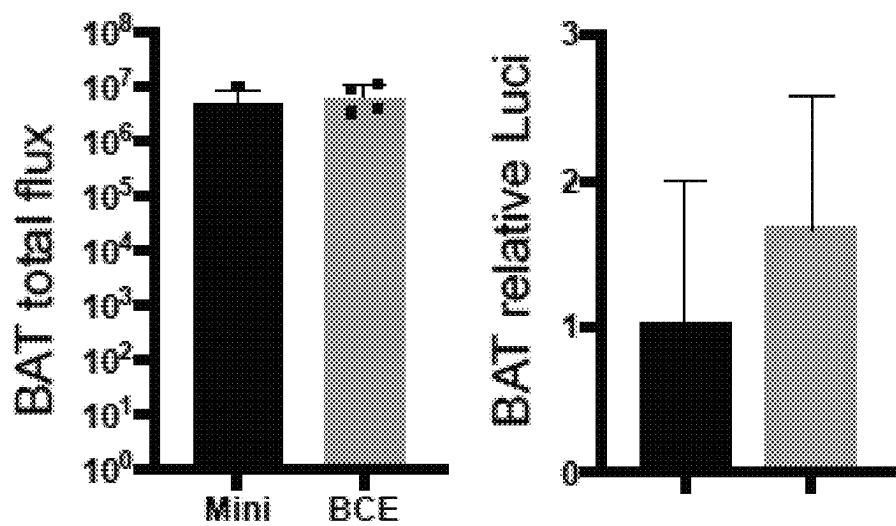


FIG. 16

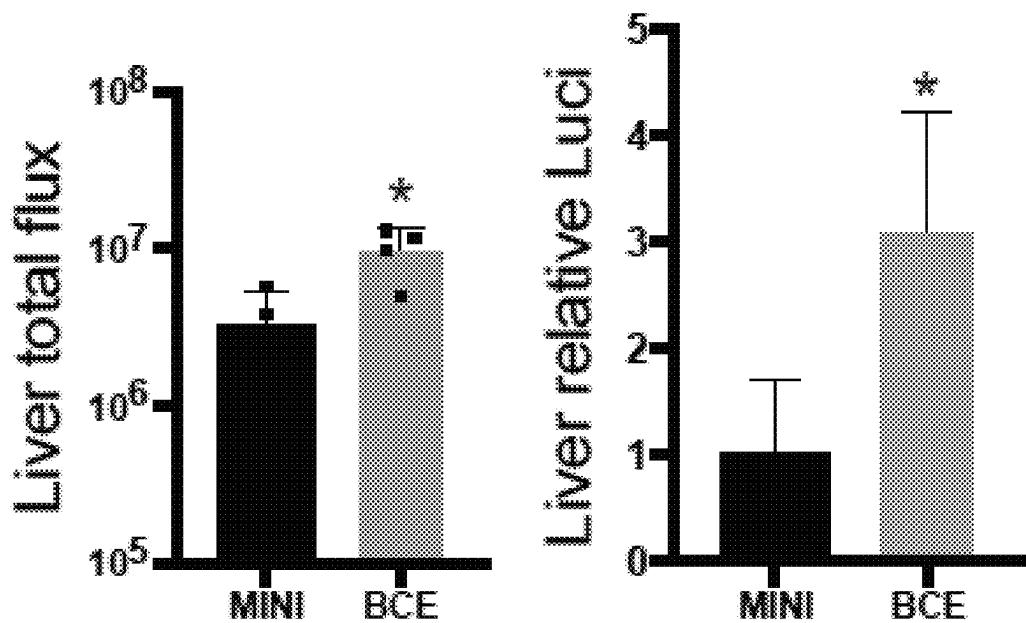


FIG. 17

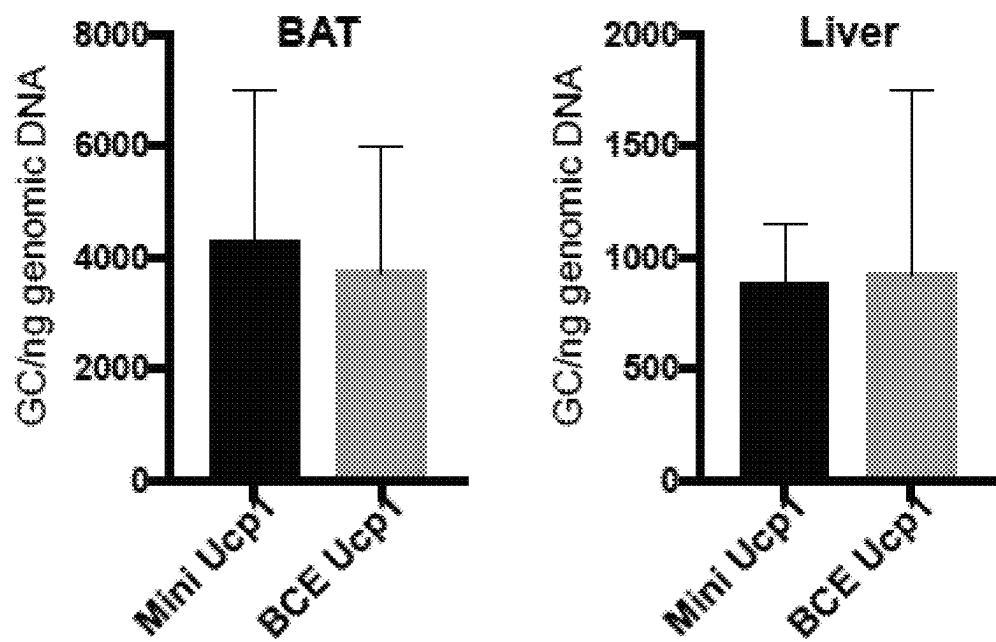


FIG. 18

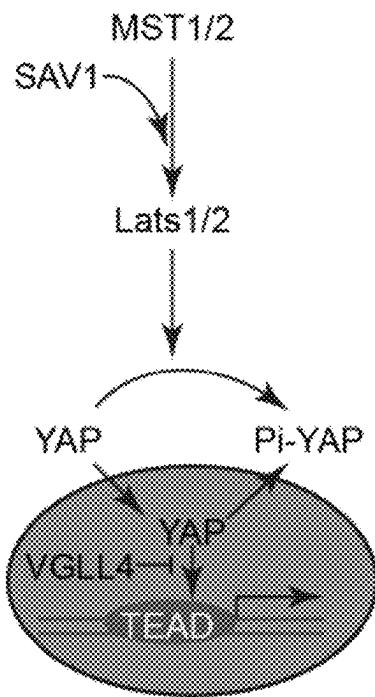
Hippo-YAP pathway

FIG. 19

	TDU_1	TDU_2
hVGLL4	DPVVVEHFRRSLGKNY . . .	TGSVDDHFAKALGDTW
hVGLL4 ^{HF4A}	DPVVVEEAARRSLGKNY . . .	TGSVDDAAAKALGDTW

hVGLL4 TDU_1: SEQ ID NO: 41
hVGLL4 TDU_2: SEQ ID NO: 42
hVGLL4-HF4A TDU_1: SEQ ID NO: 43
hVGLL4-HF4A TDU_2: SEQ ID NO: 44

FIG. 20

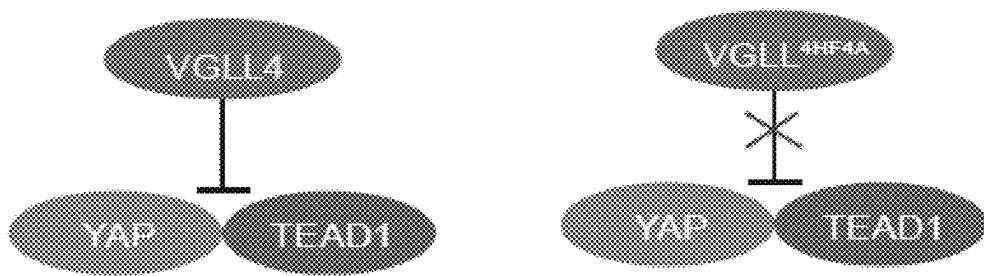


FIG. 21

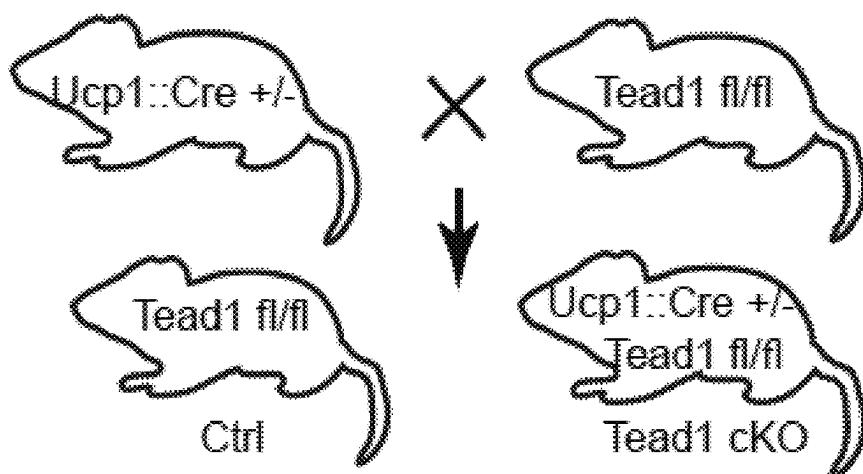


FIG. 22

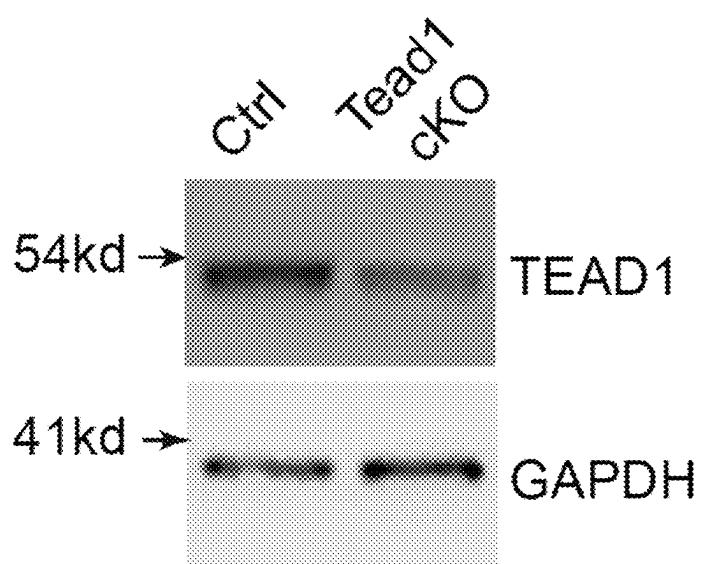


FIG. 23

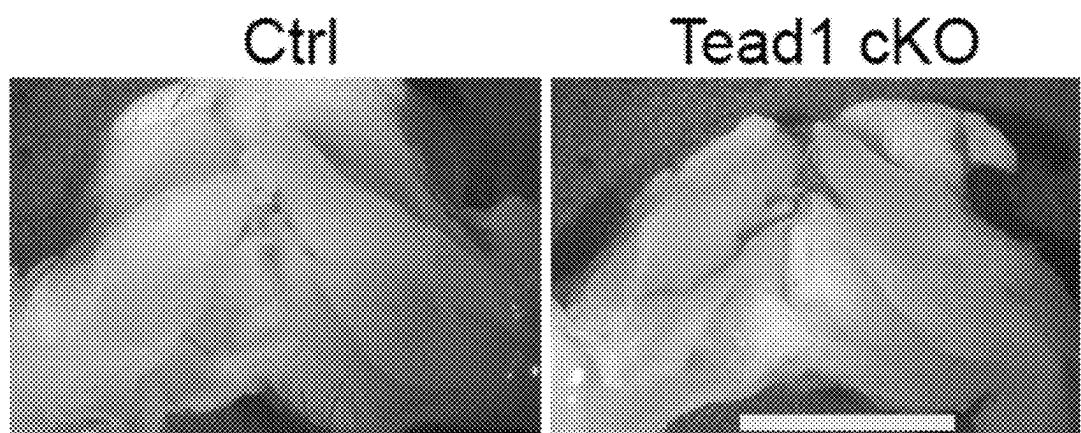


FIG. 24

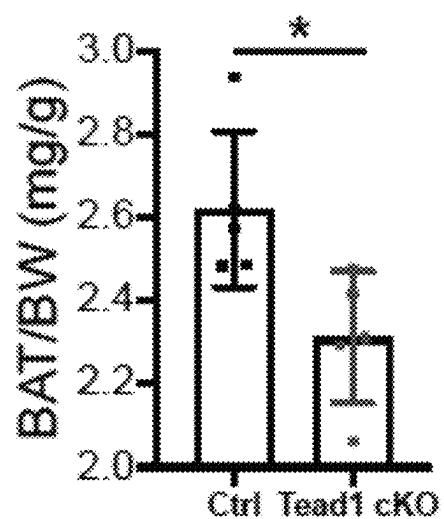


FIG. 25

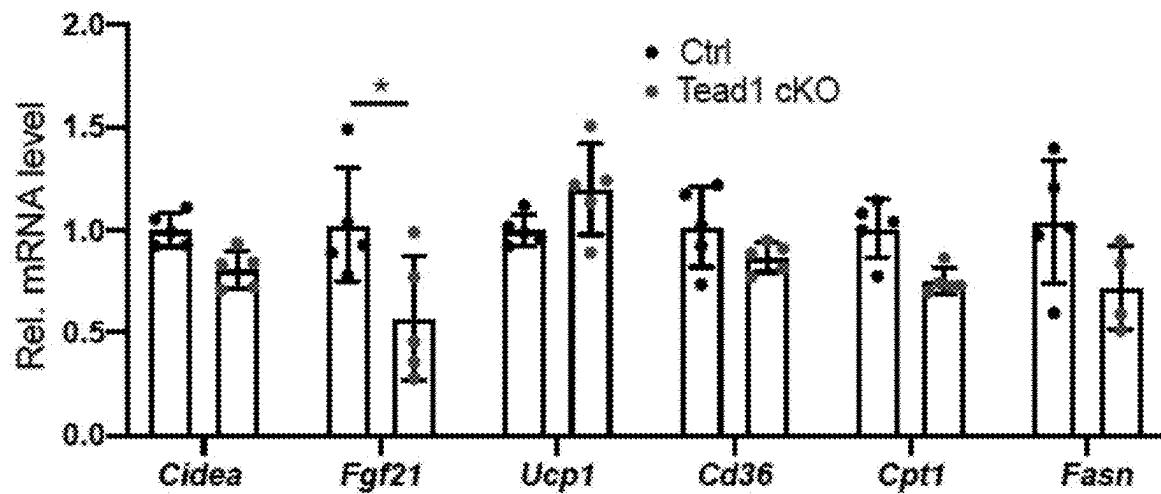


FIG. 26

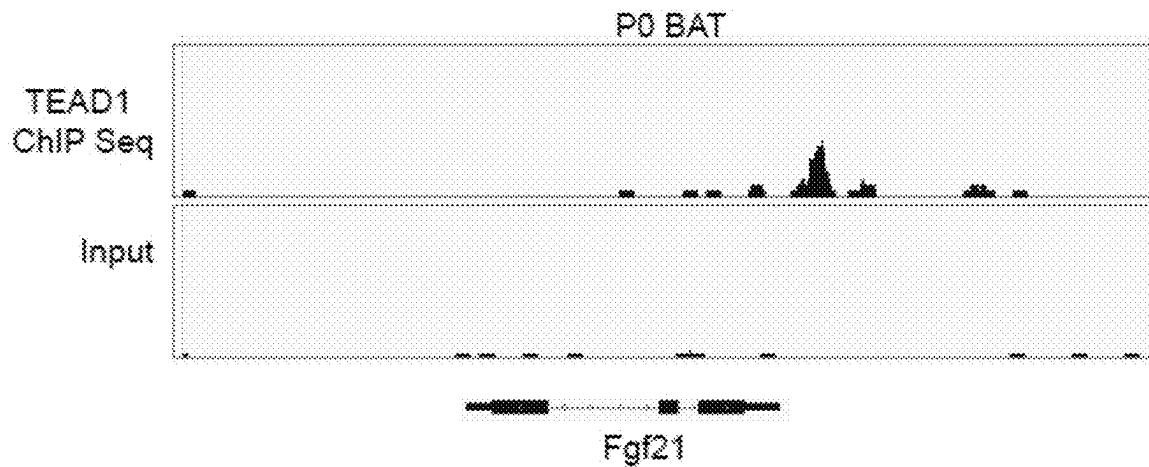


FIG. 27

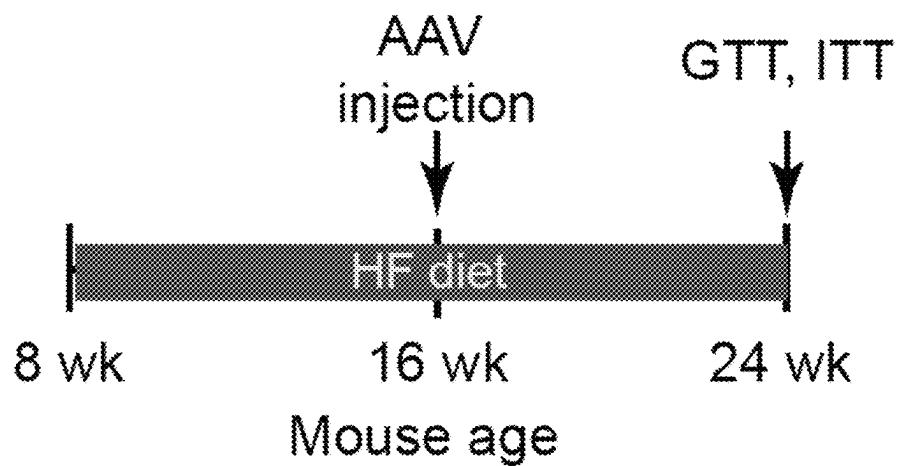


FIG. 28

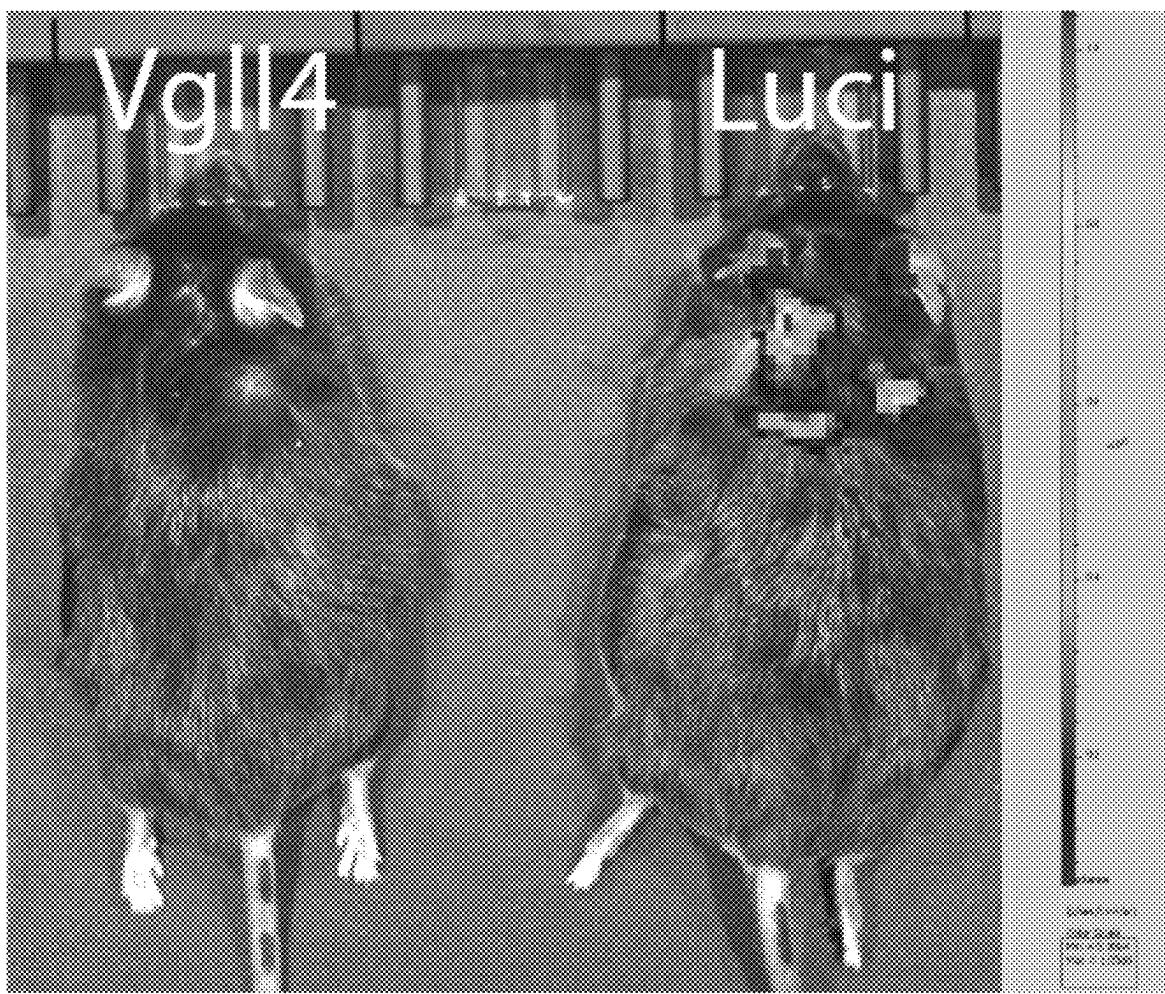


FIG. 29

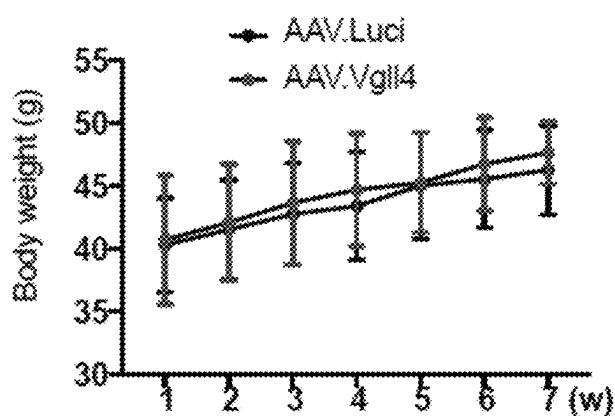


FIG. 30

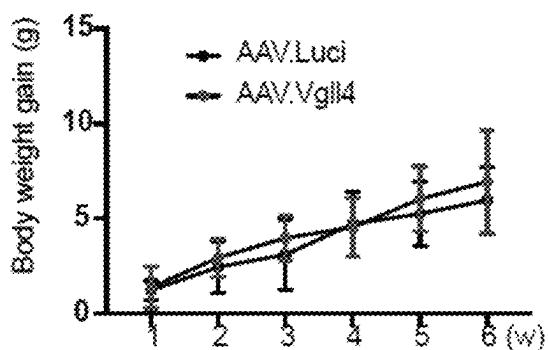


FIG. 31

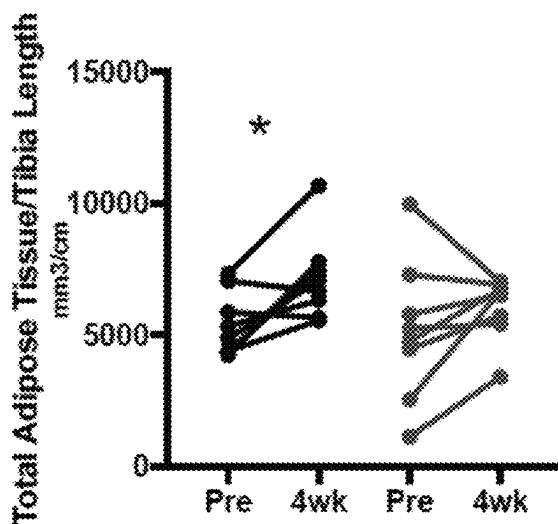


FIG. 32A

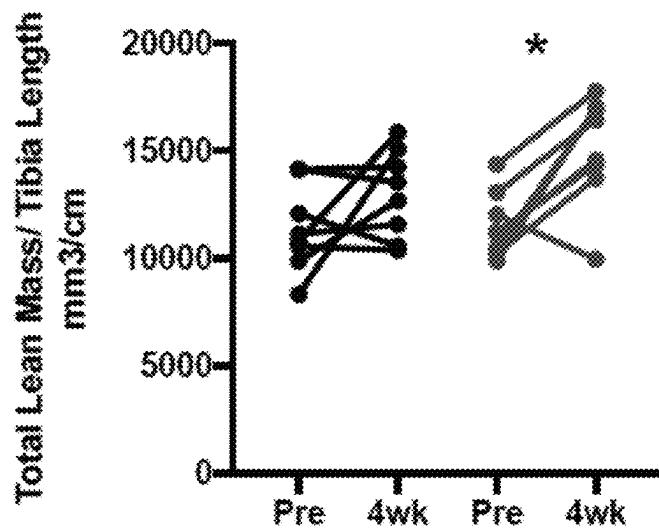


FIG. 32B

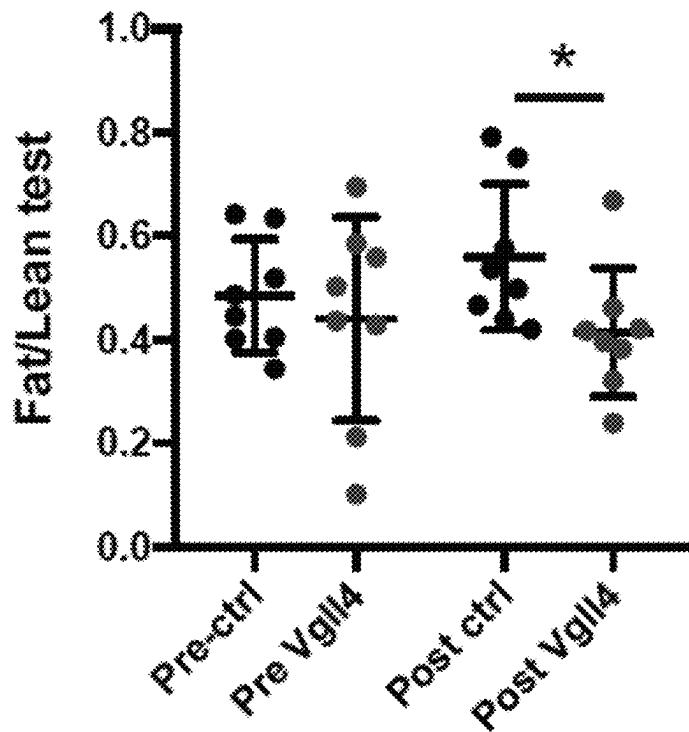


FIG. 32C

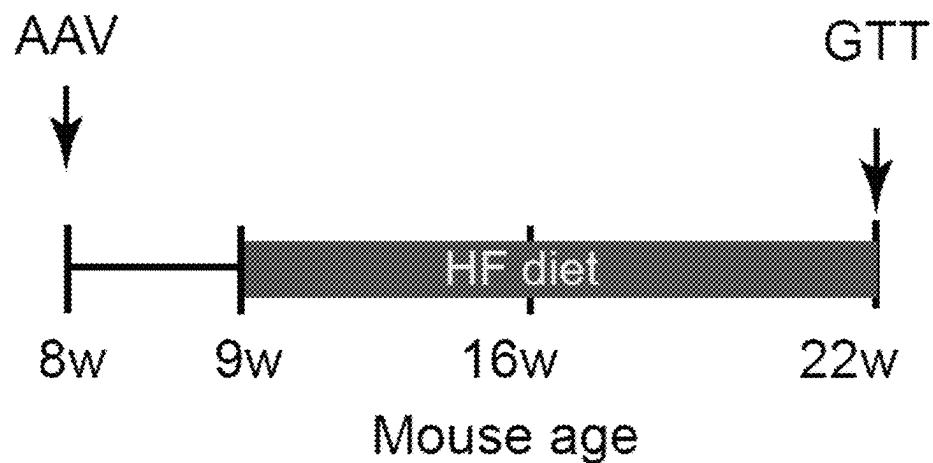


FIG. 33A

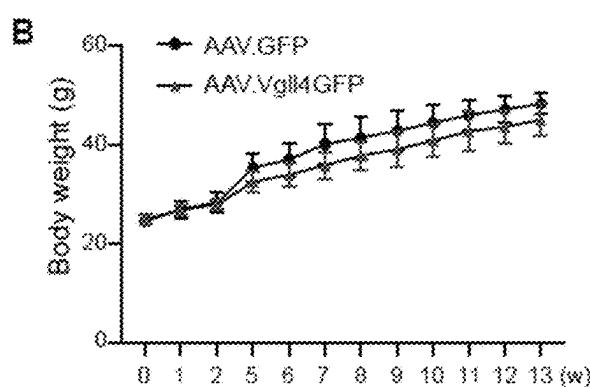


FIG. 33B

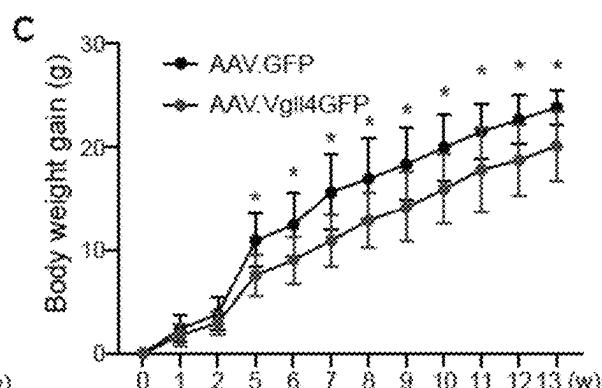


FIG. 33C

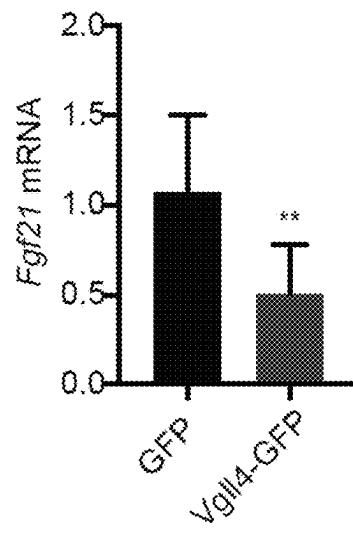


FIG. 33D

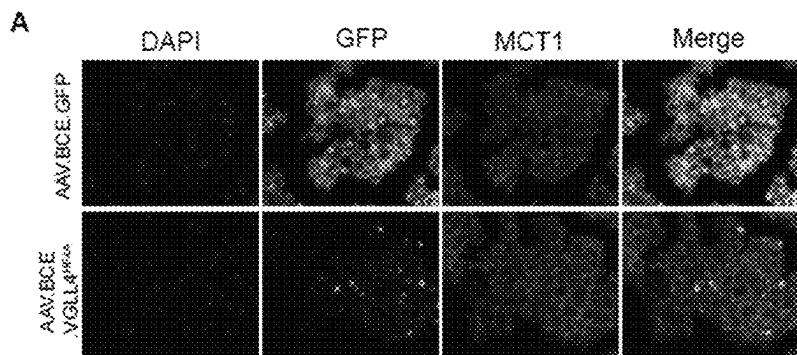


FIG. 34A

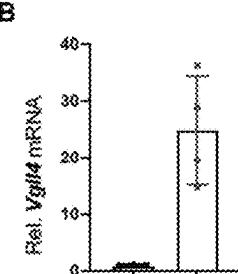


FIG. 34B

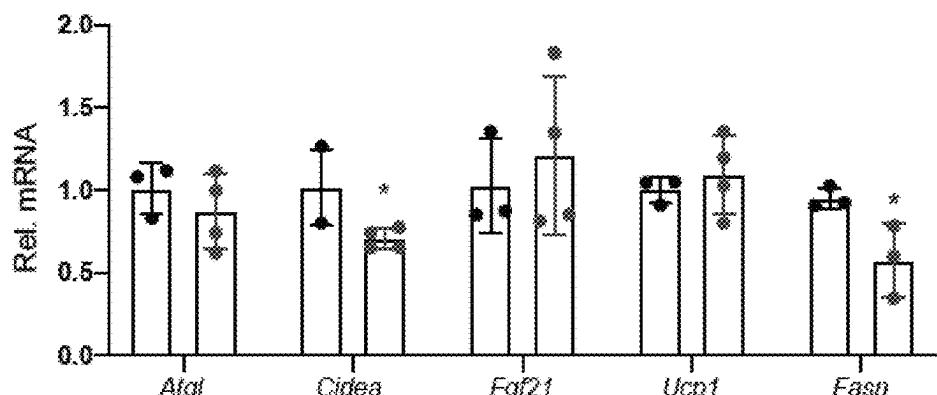


FIG. 34C

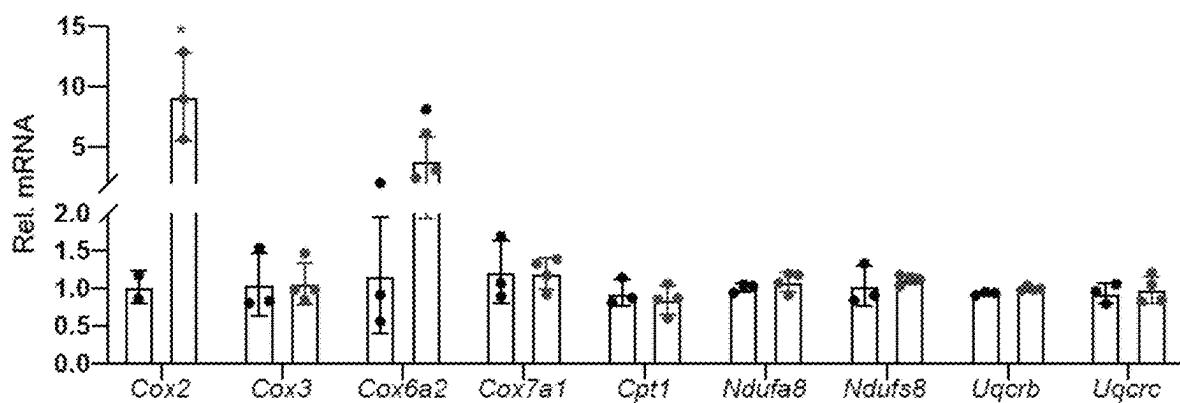


FIG. 34D

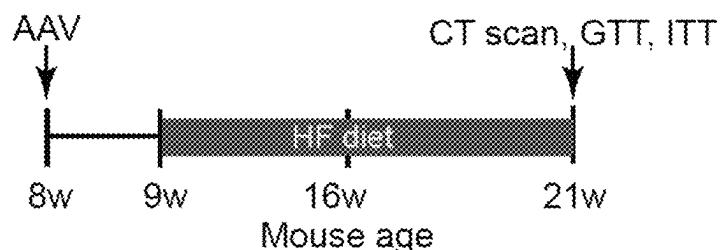


FIG. 35A

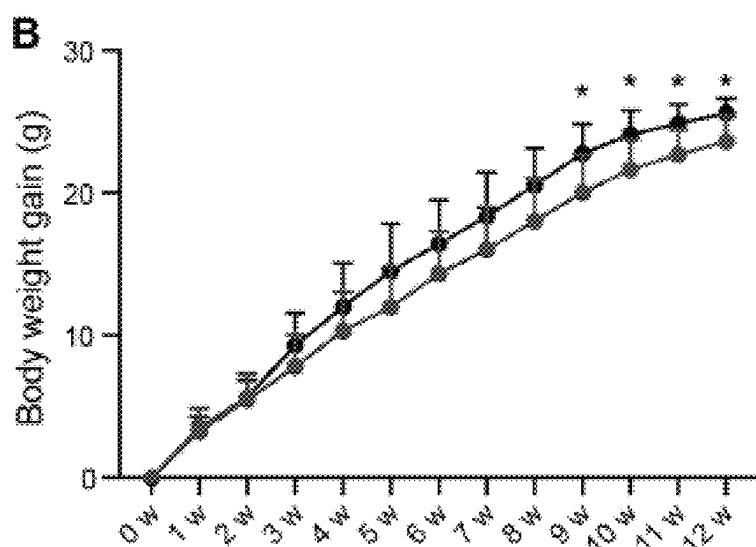


FIG. 35B

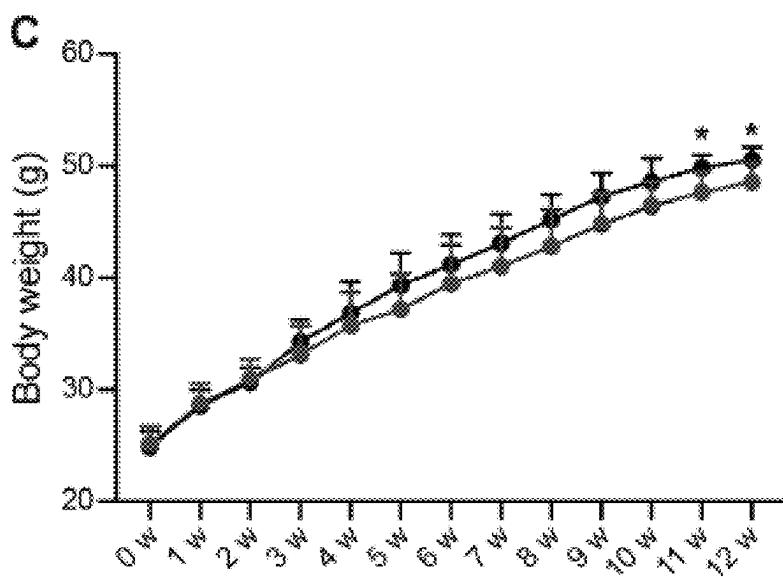


FIG. 35C

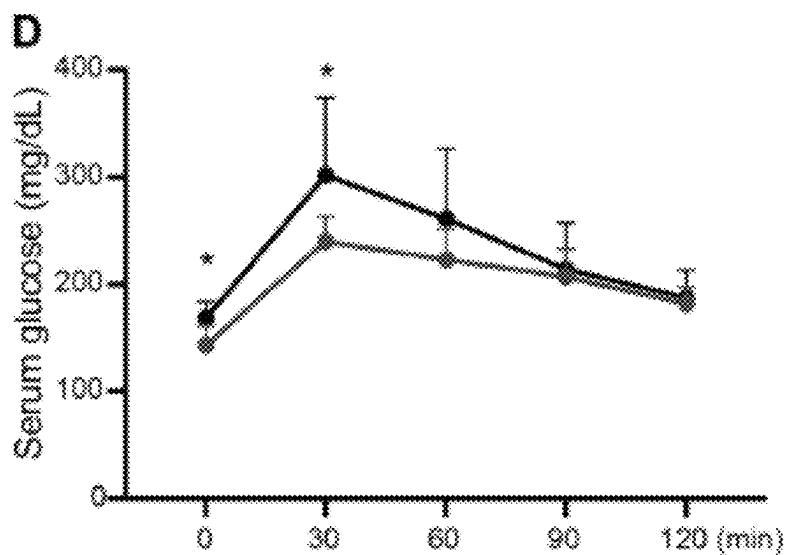


FIG. 35D

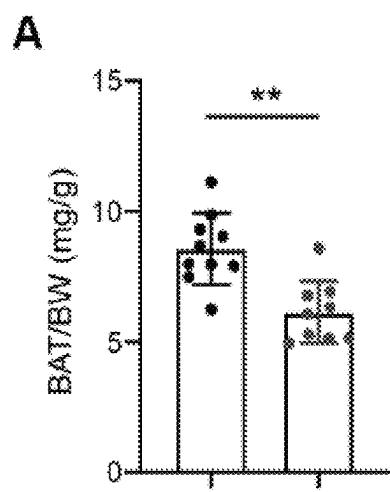


FIG. 36A

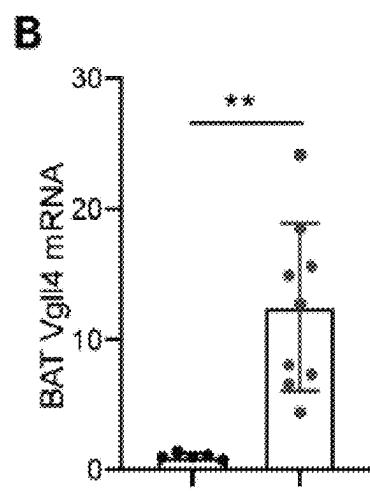


FIG. 36 B

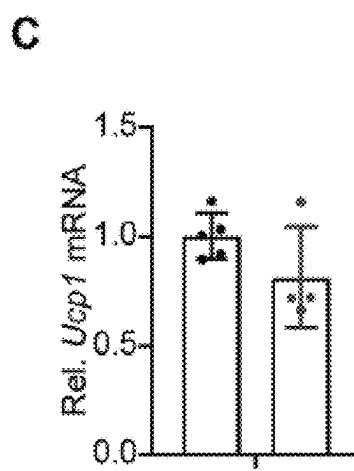


FIG. 36C

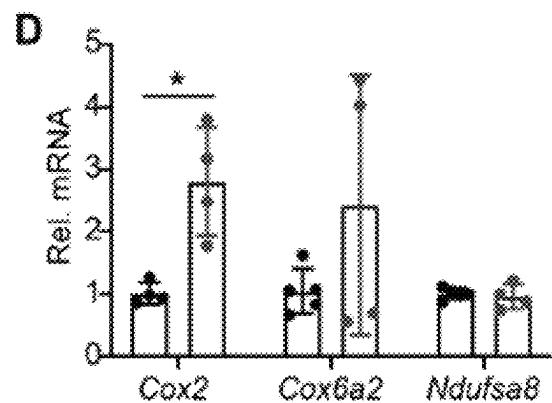


FIG. 36D

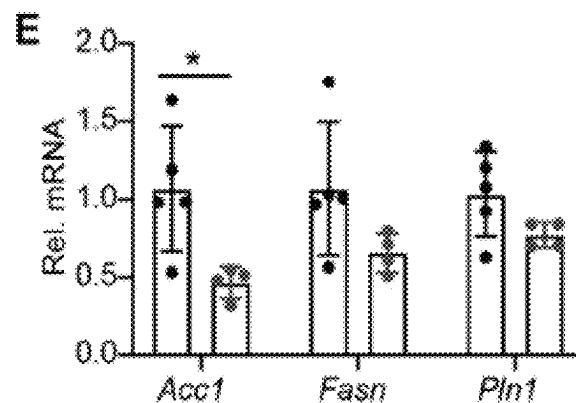


FIG. 36E

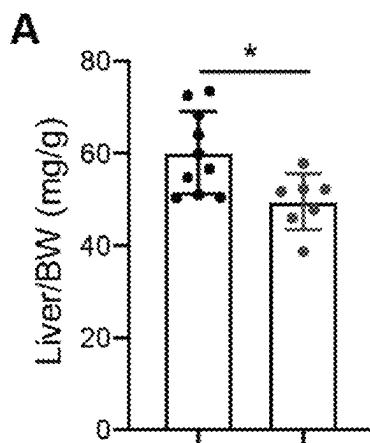


FIG. 37A

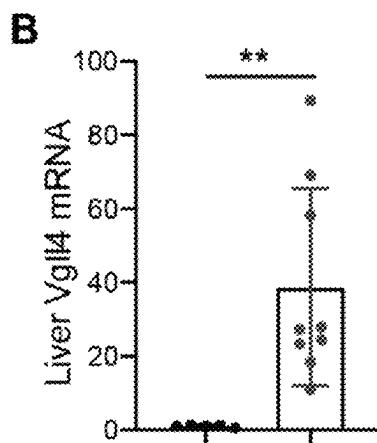


FIG. 37B

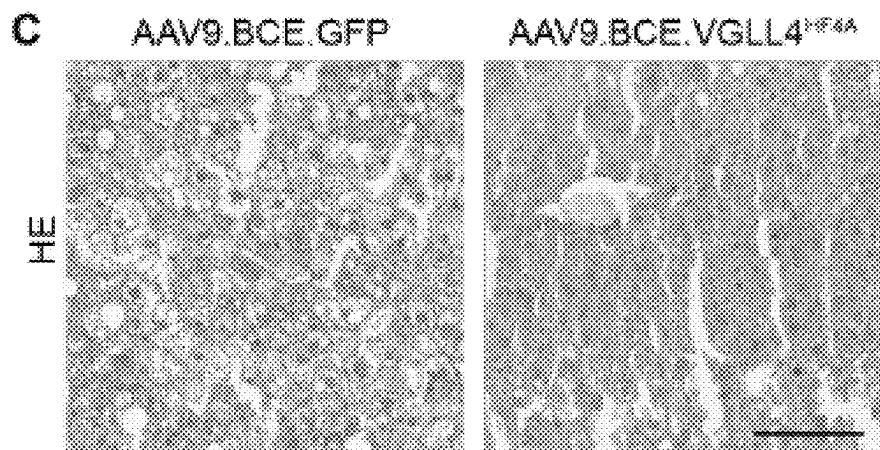


FIG. 37C

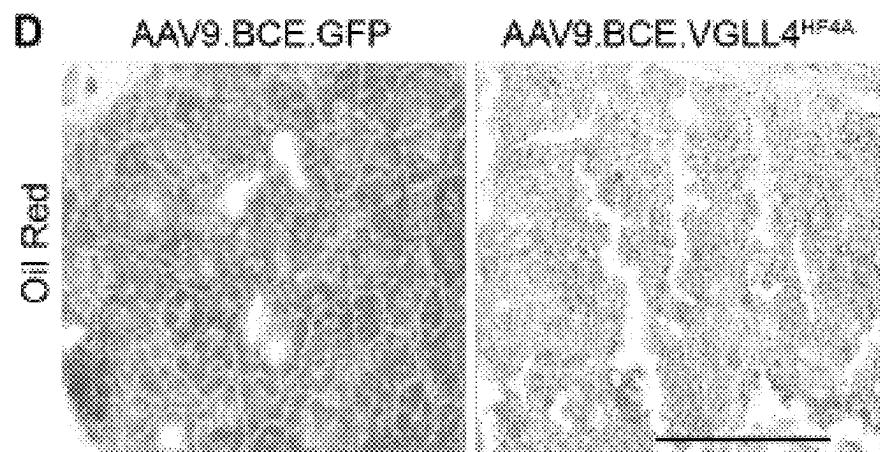


FIG. 37D

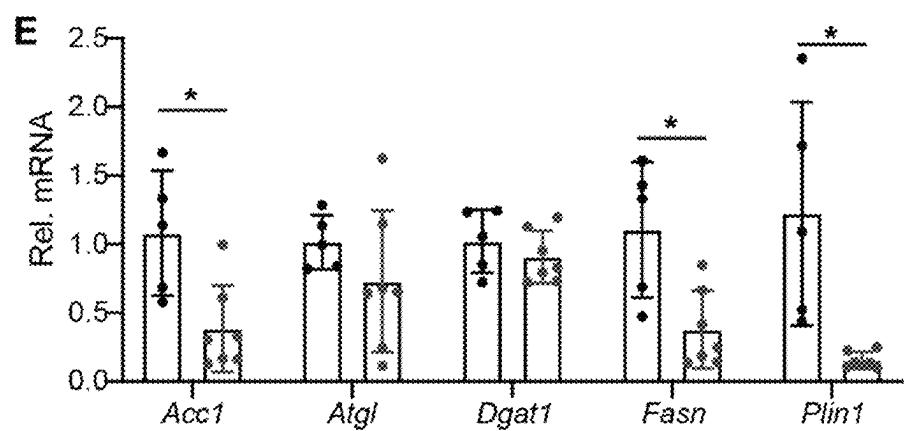


FIG. 37E

VGLL4 WITH UCP-1 CIS-REGULATORY ELEMENT AND METHOD OF USE THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation of U.S. patent application Ser. No. 16/925,632 which was filed on Jul. 10, 2022, published as US 2021/0009646 A1 on Jan. 14, 2021, and claims priority to U.S. Provisional Application No. 62/872,624, filed Jul. 10, 2019, the entire contents of which are incorporated herein in their entirieties.

GOVERNMENT RIGHTS STATEMENT

This invention was made with Government support under grant number HL138454 awarded by the National Institutes of Health. The Government has certain rights in the invention.

SEQUENCE LISTING

The instant application contains a Sequence Listing, created on Apr. 27, 2022; the file, in ASCII format, is designated H2341109 and is 65.9 KB in size. The file is hereby incorporated by reference in its entirety into the instant application.

BACKGROUND

Obesity is a global epidemic that plagues the human society, threatening the health of both adult and children. Effective pharmacological therapies for obesity are urgently needed. Obesity-related pathologies include, among others, diabetes and liver disease. Adipose tissue overgrowth is the root of obesity, with deleterious health effects. Adipose tissue is composed of white and brown adipose tissue (BAT). White adipose tissue (WAT) stores triglycerides in adipocytes, and BAT burns triglycerides and glucose for generating heat. The development of obesity depends not only on the balance between food intake and caloric utilization but also on the balance between BAT and WAT. Higher BAT is correlated with leanness in the adult and greater muscle volume in children, indicating that functional BAT benefits both energy homeostasis and muscle growth. In humans, BAT is abundant in infants, and decreases with age. Recently, the discovery of functional BAT in adult individuals raised the possibility of treating obesity by activating BAT. However, compositions and methods for increasing BAT are lacking. The present disclosure is directed to overcoming these and other deficiencies in the art.

SUMMARY

The following disclosure includes improvements over such shortcomings.

In an aspect, provided is a polynucleotide, including a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein, wherein the cis-regulatory element includes an uncoupling protein 1 enhancer and an uncoupling protein 1 promoter. In an example, the uncoupling protein 1 enhancer has at least 90% identity with a sequence selected from SEQ ID NO: 1, SEQ ID NO 4, and SEQ ID NO: 7. In another example, the uncoupling protein 1 enhancer is selected from SEQ ID NO: 1, SEQ ID NO 4, and SEQ ID NO: 7. In another example, the uncoupling protein

1 promotor has at least 90% identity with a sequence selected from SEQ ID NO: 2, SEQ ID NO 5, and SEQ ID NO: 8. In still another example, the uncoupling protein 1 promotor is selected from SEQ ID NO: 2, SEQ ID NO 5, and

5 SEQ ID NO: 8. In yet another example, the cis-regulatory element has at least 90% homology with a sequence selected from SEQ ID NO: 3, SEQ ID NO: 6, and SEQ ID NO: 9. In a further example, the cis-regulatory element is selected from SEQ ID NO: 3, SEQ ID NO: 6, and SEQ ID NO: 9.

10 In another example, the vestigial like 4 protein has at least 90% homology with a sequence selected from SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO:

15 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, and SEQ ID NO: 33. In still another example, the vestigial like 4 protein is selected from SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27,

20 29, SEQ ID NO: 31, and SEQ ID NO: 33. In yet another example, the sequence encoding a vestigial like 4 protein has at least 90% identity with a sequence selected from SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14,

25 26, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, and SEQ ID NO: 32. In a further example, the sequence encoding a vestigial like 4 protein is selected from SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14,

30 26, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, and SEQ ID NO: 32.

In another example, the vestigial like 4 protein has from 0 to 3 substitutions to a sequence selected from SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ

35 35 ID NO: 19, and SEQ ID NO: 21, wherein the substitutions are not in a TDU domain. In still another example, the vestigial like 4 protein has from 0 to 3 substitutions to a sequence selected from SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, and SEQ ID NO: 33, wherein the substitutions are not in a TDU domain.

Another example further includes an intron between the cis-regulatory element and the nucleotide sequence encoding a vestigial like 4 protein. In another example, the intron has at least 90% homology with SEQ ID NO: 34. In still another example, the intron is SEQ ID NO: 34.

45 Another example includes a nucleotide sequence having at least 90% homology with SEQ ID NO: 35. An example includes a nucleotide sequence of SEQ ID NO: 35. Another example includes a nucleotide sequence having at least 90% homology with SEQ ID NO: 36. An example includes a nucleotide sequence of SEQ ID NO: 36.

50 Another example further includes a nucleotide sequence encoding a reporter protein. In another example, the reporter protein is selected from a green fluorescent protein, a yellow fluorescent protein, a red fluorescent protein, a blue fluorescent protein, a luciferase protein, a beta-galactosidase protein, a glutathione S-transferase protein, a chloramphenicol acetyltransferase protein, and any combination of two or

60 more of the foregoing. In still another example, the reporter protein includes a green fluorescent protein. In yet another example, the reporter protein includes SEQ ID NO: 37. In a further example, the nucleotide sequence encoding a reporter protein includes SEQ ID NO: 38.

65 Another example includes a nucleotide sequence having at least 90% homology with SEQ ID NO: 39. An example includes SEQ ID NO: 39. Another example includes a

nucleotide sequence having at least 90% homology with SEQ ID NO: 40. An example includes SEQ ID NO: 40.

In another aspect, provided is a viral vector including any of the foregoing examples of a polynucleotide that include a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein. In an example, the viral vector includes an adenoviral associated vector.

In another aspect, provided is a cell transfected with any of the foregoing examples of a polynucleotide that include a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein. In an example, the cell was contacted with any of the foregoing examples of a viral vector that include any of the foregoing examples of a polynucleotide that include a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein.

In still another aspect, provided is an organism transfected with any of the foregoing examples of a polynucleotide that include a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein. In an example, the organism was contacted with any of the foregoing examples of a viral vector.

In another aspect, provided is a method. In an example, the method includes transfecting a cell with any of the foregoing examples of a polynucleotide that include a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein. In another example, transfecting includes contacting the cell with any of the foregoing examples of a viral vector that include any of the foregoing examples of a polynucleotide that include a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein. In still another example includes transfecting an organism with any of the foregoing examples of a polynucleotide that include a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein. In yet another example, transfecting includes contacting the organism with any of the foregoing examples of a viral vector that include any of the foregoing examples of a polynucleotide that include a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein.

In another example, the organism is a mammal. In still another example, the organism is a human.

In an example of the method, the vestigial like protein 4 does not include an HF to AA substitution in a TDU domain of the vestigial like protein 4 and the transfecting includes increasing a ratio of a volume of brown adipose tissue to a volume of white adipose tissue in the organism. In still another example, the vestigial like protein 4 does not include an HF to AA substitution in a TDU domain of the vestigial like protein 4 and the transfecting includes increasing a volume of brown adipose tissue in the organism, decreasing the volume of white adipose tissue in the organism, or both. In yet another example, the vestigial like protein 4 does not comprise an HF to AA substitution in a TDU domain and the transfecting includes reducing a ratio of a volume of adipose tissue to a volume of non-adipose tissue in the organism.

In a further example, the vestigial like protein 4 does not include an HF to AA substitution in a TDU domain, and the organism is obese or is at risk of developing obesity. In still a further example, the vestigial like protein 4 does not include an HF to AA substitution in a TDU domain, and the transfecting includes preventing obesity in the organism. In yet another example, the vestigial like protein 4 does not include an HF to AA substitution in a TDU domain, and the transfecting includes treating obesity in the organism. In another example, the vestigial like protein 4 does not include an HF to AA substitution in a TDU domain, and the transfecting includes reducing obesity in the organism.

In an example of the method, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, wherein the transfecting includes reducing a volume of adipose tissue of the organism. In another example, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, and the transfecting includes reducing a volume of brown adipose tissue of the organism. In still another example, the vestigial like protein 4 includes an HF to AA substitution in a TDU domain, and the organism is obese or is at risk of developing obesity. In yet another example, the vestigial like protein 4 includes an HF to AA substitution in a TDU domain, and the transfecting includes preventing obesity in the organism. In a further example, the transfecting includes treating obesity in the organism. In still a further example, the vestigial like protein 4 includes an HF to AA substitution in a TDU domain, and the transfecting includes reducing obesity in the organism.

In an example of the method, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, and the transfecting includes reducing fatty acid synthesis in the organism. In another example, the vestigial like protein 4 includes an HF to AA substitution in a TDU domain, and the organism has hepatic steatosis or is at risk for developing hepatic steatosis. In still another example, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, and the transfecting includes preventing hepatic steatosis in the organism. In yet another example, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, and the transfecting includes treating hepatic steatosis in the organism. In a further example, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, and the transfecting includes reducing hepatic steatosis in the organism.

In an example of the method, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, and the organism has diabetes or is at risk of developing diabetes. In another example, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, wherein the transfecting includes preventing diabetes in the organism. In still another example, the vestigial like protein 4 includes an HF to AA substitution in each of two TDU domains, wherein the transfecting includes treating diabetes in the organism.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features, aspects, and advantages of the present disclosure will become better understood when the following detailed description is read with reference to the accompanying drawings, wherein:

FIG. 1 shows a schematic view of constructs in accordance with aspects of the present disclosure.

FIG. 2 shows timing of vector administration and expression measurement in accordance with aspects of the present disclosure.

FIG. 3 shows representative bioluminescence images of a dorsal view of mice administered a viral vector in accordance with aspects of the present disclosure.

FIG. 4 shows representative bioluminescence images of a ventral view of mice administered a viral vector in accordance with aspects of the present disclosure.

FIG. 5 shows bioluminescence signal origin following viral vector administration in accordance with aspects of the present disclosure.

FIG. 6 shows immunofluorescence staining images of interscapular BAT following viral vector administration in accordance with aspects of the present disclosure.

FIG. 7 shows immunofluorescence staining images of interscapular BAT and WAT following viral vector administration in accordance with aspects of the present disclosure.

FIG. 8 shows a schematic view of a construct in accordance with aspects of the present disclosure.

FIG. 9 shows timing of vector administration and BAT and WAT measurement in accordance with aspects of the present disclosure.

FIG. 10 shows relative BAT and WAT expression following treatment with control viral vector in accordance with aspects of the present disclosure.

FIG. 11 shows relative BAT and WAT expression following treatment with a viral vector including a UCP-1 cis-regulatory element and Vgll4 coding sequence in accordance with aspects of the present disclosure.

FIG. 12 shows an effect of treatment with a viral vector including a UCP-1 cis-regulatory element and Vgll4 coding sequence on BAT volume in accordance with aspects of the present disclosure.

FIG. 13 shows an effect of treatment with a viral vector including a UCP-1 cis-regulatory element and Vgll4 coding sequence on WAT volume in accordance with aspects of the present disclosure.

FIG. 14 shows an effect of treatment with a viral vector including a UCP-1 cis-regulatory element and Vgll4 coding sequence on the ratio of WAT volume to BAT volume in accordance with aspects of the present disclosure.

FIG. 15 shows bioluminescent expression of a reporter protein, luciferase, in 6-week old mice administered viral vectors including cis-regulatory elements in accordance with the present disclosure.

FIG. 16 shows quantification of luciferase activity in brown adipose tissue (BAT).

FIG. 17 shows quantification of luciferase activity in liver.

FIG. 18 shows quantification of viral vector genome copies in BAT and liver as assessed by real-time PCR.

FIG. 19 is an illustration of the Hippo/YAP signaling cascade and Vgll4's inhibitory role in YAP/TEAD interactions.

FIG. 20 shows the mutations made to two TONDU (TDU_1 and TDU_2) domains of Vgll4 isoforms. Vgll4 isoforms A through F include two TDU domains, TDU_1 and TDU_2, with the indicated sequences (SEQ ID NO: 41 and SEQ ID NO: 42). The Vgll4-HF4A mutants of Vgll4 isoforms disclosed herein include two alanine substitutions to four amino acids, two in TDU_1 (a histidine to alanine and a phenylalanine to alanine) and two in TDU_2 (a histidine to alanine and a phenylalanine to alanine). Vgll4-HF4A have the dual-substituted TDU domains SEQ ID NO: 43 and SEQ NO ID: 44 (instead of SEQ ID NO: 41 and SEQ ID NO: 42, respectively).

FIG. 21 is an illustration showing that whereas Vgll4 disrupts a YAP/TEAD1 complex, Vgll4-HF4A mutants do not.

FIG. 22 Shows a schematic illustration of generation of brown adipocyte specific Tead1 knockout mice.

FIG. 23 is a western blot showing confirmation of depleted expression of TEAD1 in brown adipose tissue of the conditional TEAD1 knockout mice, with GADPH as a control.

FIG. 24 shows whole mount view of interscapular brown adipose tissue (iBAT) collected from 1-month old male mice. Bar=5 mm.

FIG. 25 shows the ration between iBAT and body weight ratio (* P<0.05).

FIG. 26 shows qRT-PCR measurement of various mRNA transcript levels in BAT, normalized to 36B4.

FIG. 27 is a genomic view of TEAD1 binding site in the Fgf21 promoter region.

FIG. 28 is an illustration of experimental design for injecting adult mice fed a high-fat diet with a viral vector in accordance with the present disclosure. GTT is glucose tolerance test and ITT is insulin tolerance test.

FIG. 29 shows bioluminescence imaging mice 8 weeks after infusion with viral vectors in accordance with the present disclosure.

FIG. 30 and FIG. 31 show total body weight measurements and accumulated body weight gain measurements of transfected mice, respectively.

FIGS. 32A, 32B, AND 32C show effects for transfection with a viral vector carrying a BCE-Vgll4 polynucleotide transcript on a ratio of total adipose tissue to tibia length, a ratio of total lean mass to tibia length, and a ratio of fat mass to lean mass. In 32A and 32B, AAV.Luciferase control is on the right and AAV.Vgll4 is on the right.

FIG. 33A shows an illustration of an experimental design demonstrating that pre-treatment of mice with AAV.BCE.VGLL4 mitigates body weight gain. AAV.BCE.Vgll4 was subcutaneously injected into the interscapular region of 8-weeks-old C57/BL6 mice. After 13 weeks high fat diet treatment, mice were tested for glucose tolerance (GTT). AAV.BCE.GFP was used as control.

FIGS. 33B and 33C show total body weight and accumulated body weight gain, respectively. AAV.GFP is the upper tracing and AAV.Vgll4.GFP is the lower.

FIG. 33D shows fgf21 mRNA levels in BAT following transfection with a construct driving Vgll4 with a Ucp1 cis regulatory element in accordance with aspects of the present disclosure.

FIGS. 34A and 34B show immunofluorescence images of interscapular brown adiposities (MCT1 was used to label the cell borders) and Vgll4 expression, respectively, of 8-week-old C57/BL6 mice 10 days after AAV.BCE.Vgll4 or AAV.BCE.GFP was subcutaneously injected into the interscapular region.

FIG. 34C AND 34D show real-time PCR measurements of various mRNA transcripts in brown adipose tissue. Control (AAV.BCE.GFP) is on the right and AAV.BCE.Vgll4-HF4A is on the right.

FIG. 35A shows an illustration of an experimental design demonstrating that pre-treatment with an AAV.BCE.VGLL4-HF4A mitigates body weight gain. AAV.BCE.Vgll4-HF4A was subcutaneously injected into the interscapular region of 8-weeks-old C57/BL6 mice. After 12 weeks high fat diet treatment, mice were tested for glucose (GTT) and insulin tolerance (ITT). AAV.BCE.GFP was used as control.

FIGS. 35B and 35C show total body weight and accumulated body weight gain, respectively. FIG. 35D shows serum glucose levels following glucose challenge in a glucose tolerance test. Upper traces are control and lower traces are AAV.BCE.Vgll4-HF4A.

FIGS. 36A, 36B, 36C, 36D, AND 36E are graphs showing a ratio of brown adipose tissue weight to body weight, Vgll4 mRNA expression in BAT, Ucp1 mRNA expression in BAT, mitochondrial gene mRNA levels in BAT, and fatty acid synthesis gene mRNA expression in BAT. Student t test, *, P<0.05; **, P<0.01. Control (AAV.BCE.GFP) is on the right and AAV.BCE.Vgll4-HF4A is on the left.

FIG. 37A shows that pre-treatment with AAV.BCE.Vgll4-HF4A reduces ratio of liver weight to body weight. FIG. 37B

shows liver Vgll4 mRNA expression level. FIGS. 37C and 37D are photomicrographs showing HE staining and oil red staining, respectively, of liver sections (bar=100 µm). FIG. 37E shows mRNA levels of expression of fatty acid synthesis genes in liver by quantitative real-time PCR. A, B, E, Student t test, *, P<0.05; ** P<0.01. Control (AAV.BCE.GFP) is on the right and AAV.BCE.Vgll4HFA is on the right.

DETAILED DESCRIPTION

This disclosure relates to a construct including a cis-regulatory element upstream of a coding sequence for a vestigial like 4 peptide (Vgll4). In an example, a cis-regulatory element may promote expression of a Vgll4 peptide in BAT cells. In an example, a cis-regulatory element may also promote expression of a Vgll4 peptide in liver cells. It may further specifically or enrichingly drive expression in BAT cells relative to expression driven in many or most other cells, when cells are transfected with the construct. In an example, it may further specifically or enrichingly drive expression in BAT and liver cells relative to expression driven in many or most other cells, when cells are transfected with the construct. Also disclosed is a viral vector including the construct, wherein the viral vector enables, permits, or promotes transfection of cells with the construct. In some examples, the Vgll4 peptide may include amino acid substitutions. For example, Vgll4 peptides include two TONDU (or TDU) domains, referred to herein as TDU_1 and TDU_2. Each TDU domain includes an HF dipeptide sequence. In an example, one or both HF TDU dipeptides may include amino acid substitutions, replacing HF with a dipeptide of aliphatic amino acids, such as AA.

In an example, also disclosed is causing an increase in BAT, a decrease in WAT, an increase in a ratio of BAT volume to WAT volume, or any two of the foregoing, by contacting an organism with the construct, such as by transfecting cells of an organism with the construct. In an example, also disclosed is reducing a volume of adipose tissue of the organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct. In an example, also disclosed is reducing a mass ratio BAT to body weight of an organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct. In an example, also disclosed is reducing a liver volume, liver weight, intrahepatic fat content, or any combination of two or more of the foregoing, of an organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct.

In an example, also disclosed is reducing or minimizing blood glucose levels or a rise in glucose levels in an organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct. In an example, also disclosed is increasing expression of mitochondrial genes, such as mitochondrial genes involved in mitochondrial respiration, in an organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct. In an example, also disclosed is decreasing expression of genes that promote lipogenesis, in an organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct.

In an example, a viral vector including the construct is used to transfect cells of an organism with the construct. A viral vector may be an adeno-associated viral vector or

another viral vector known to be able to transfect cells. The organism may be a mammal, such as a rodent or human or any other mammal.

Vgll4 is a transcription co-factor known to interact with cellular signaling molecules and transcription factors to influence cell survival and cell function. Vgll4 is particularly known for promoting cellular death by inhibiting YAP-TEAD1 complex. Several isoforms of Vgll4 have been identified, arising from splice variants to the Vgll4 gene. 10 These include Vgll4A, Vgll4B, Vgll4C Vgll4D, Vgll4E, and Vgll4F. Amino acid sequences of these Vgll4 proteins (referred to collectively here as Vgll4), and examples of polynucleotides encoding them encoding them, are given in Table II. Vgll4 has been linked with an anticancer effect in 15 several types of cancer, where lower levels of Vgll4 correlate or correspond with or cause increased tumor cell survival and higher levels of Vgll4 correlate or correspond with or cause an anti-tumor effect including decreased metastatic processes and decreased tumor cell survival or proliferation. 20 See Deng, Vgll4 is a transcriptional cofactor acting as a novel tumor suppressor via interacting with TEADs, Am J Cancer Res (2018), 8(6):932-943. In this respect, Vgll4 differs from other member of the vestigial like (Vgll) family (Vgll1, Vgll2, and Vgll3) of transcription co-factors, which 25 are not known to have tumor-suppressive functions. Vgll family members other than Vgll4 are not generally understood to share functional commonalities with Vgll4.

In view of the well-established role of increased Vgll4 expression in promoting cellular death processes or inhibiting cell survival, an increase in BAT volume as disclosed herein surprisingly results from Vgll4 expression influenced by a BAT-cell specific cis-regulatory element. In another example, and without being limited to any particular mechanism of action, where a Vgll4 protein includes HF to AA substitutions in both TDU domains, a reduced BAT volume, reduced intrahepatic fat accumulation, or both, may result from Vgll4 activity that does not include Vgll4 integration with a TEAD protein. In another example, and without being limited to any particular mechanism of action, where a Vgll4 35 protein includes HF to AA substitutions in both TDU domains, a reduced BAT volume, reduced intrahepatic fat accumulation, or both, may result from increased expression of mitochondrial genes involved in mitochondrial respiration, decreased expression of genes involved in lipogenesis, or both.

For driving expression under control of a cis-regulatory element, cis-regulatory elements of uncoupling protein 1 (Ucp1) may be placed adjacent to a coding sequence for a Vgll4. By cis-regulatory element, what is meant is a nucleotide sequence that regulates the transcription of neighboring gene or coding sequences. Conventionally, Ucp1 is considered to be expressed specifically in BAT cells. Thus, such a cis-regulatory element may drive expression mostly, or predominantly, or in some cases exclusively, in BAT cells. 50 Surprisingly, however, as disclosed herein, such a cis-regulatory element may drive expression in liver cells in addition to expression in BAT cells. In an example, such a cis-regulatory element may drive expression only in BAT and liver cells.

A cis-regulatory element may include a promotor, an enhancer, or both. In some cases, a sequence for a cis-regulatory element may be located within fewer than 10 nucleotides from a transcription start site, fewer than 20 nucleotides from a transcription start site, fewer than 30 nucleotides from a transcription start site, fewer than 40 nucleotides from a transcription start site, fewer than 50 nucleotides from a transcription start site, fewer than 60

nucleotides from a transcription start site, fewer than 70 nucleotides from a transcription start site, fewer than 80 nucleotides from a transcription start site, fewer than 90 nucleotides from a transcription start site, fewer than 100 nucleotides from a transcription start site, fewer than 125 nucleotides from a transcription start site, fewer than 150 nucleotides from a transcription start site, fewer than 175 nucleotides from a transcription start site, fewer than 200 nucleotides from a transcription start site, fewer than 225 nucleotides from a transcription start site, fewer than 250 nucleotides from a transcription start site, fewer than 275 nucleotides from a transcription start site, fewer than 300 nucleotides from a transcription start site, fewer than 325 nucleotides from a transcription start site, fewer than 35 nucleotides from a transcription start site, fewer than 375 nucleotides from a transcription start site, fewer than 400 nucleotides from a transcription start site, fewer than 425 nucleotides from a transcription start site, fewer than 450 nucleotides from a transcription start site, fewer than 475 nucleotides from a transcription start site, fewer than 500 nucleotides from a transcription start site, or between 500 and 1,000 nucleotides from a transcription start site

A promoter is a nucleotide sequence to which RNA polymerizing enzymes bind for initiation of transcription of a downstream gene sequence. Many genes that show tissue- or cell-type specific expression including a promotor upstream of the DNA sequence that codes for the RNA that is particularly active in cells where the gene is expressed. A

5 promoter may be more active in some cells than other, such as being active only in specific cell- or tissue-types, or highly active in certain cell- or tissue-types relative to others. Promoters include a sequence where transcription is initiated. Eukaryotic promoters may and typically do include features such as a TATA box, a transcription factor IIB recognition site, and a core promotor sequence (or an initiator). Transcription factors bind and RNA polymerase bind to a promoter for transcription initiation.

10 Also included in a cis-regulatory element may be one or more enhancer sequence. An enhancer is part of a cis-regulatory element that enhances transcription initiated in or by the promotor. An enhancer may serve to promote an initiation of transcription at a promoter, for example, such as through binding of additional transcription factors to the enhancer that facilitate or enhance recruitment of other factors and transcriptional machinery to the promotor. As with promotors, many genes have enhancers that are involved in cell- or tissue-specific or cell- or tissue-enhanced expression.

15 Ucp1 is a mitochondrial protein expressed specifically in BAT cells. The Ucp1 gene includes a cis regulatory element in which enhancer and promotor elements have been identified and characterized. Such cis-regulatory elements are responsible for promoting expression of neighboring gene sequences in BAT cells and not other tissue or cell types. Sequences that may be included in a cis regulatory element in accordance with the present disclosure as based on cis regulatory elements of Ucp1 genes are shown in Table I.

TABLE I

Cis Regulatory Element Sequences				
	SEQ ID NO	Identity	Sequence	
1	Mouse Ucp1 enhancer		GCATGCCAATTATAGTGGCGTCACTAACAGTACTGATACTTTA CATGCTAAAGTTAAAGTGTGTCTATATTAAATTGTAAGATTGGTG AAGAGAGGTGTTATCAGATGGAAGCTGCACATTCTGGATTAATG TGGTTAAATGTTCTCTCTGTGATTACTGTCTTATTCTCTCTTT TAAATATTGTCATTGGACATCTGTATAGCTACGCCCCTGTTGTTCTCTCTT ACGTCTCTGGAGACAGATAAGAAGTTACGACGGGAGGAGCAG ATGGAGGCAAAGCGCTGTGATGCTTTGTGGTTGAGTCACACA TTGTTCAAGTGTGAAATGAGTGAAGCAAATGGTGACCGGG TGCCCTGAAATGTTCTACATCTTAAGAGAGAAACACGGACA CTAGGTAAGTGAAGCTTGCTGTCACTCTTACAGCGTCAAGAG GGTCAGTCACCTTGACCACACTGAACTAGTCGTACCTTCACT CTTCTGCCCAGAAAGCAGAAATCAGACTCTCTGGGATATCAGC CTCACCCCTACTGTTCTCTCATTATGAGGCAAACCTTCTTCACT TCCCAGGGCTCTGGGGCAGCAAGTCACCCCTTCTCAGACT CTAG	
2	Mouse Ucp1 promoter		TCTCGGAGGAGATCAGATCGCGCTTATTCAAGGGAACCCAGCCCC GCTCTCGCCCTGGCCAAGGCTGTGAAGAGTGACAAAAGGCAC CACGCTCGGGGACCGGGTGAACGCCCTCTGTGTGTCTCTGG CATATAAGGAACCTGGTGCAAATCAGAGGTGATGTGGCCAGGG CTTGGGAGTGACCGGGCTGGGAGGGTGGCAGACCAAGGCA CGCCCTGCCCAGGCAACTAGCAGCTTTGGAGACCTGGGCG GCTCAGCCACTTCCCCAGTCCCTCTCGGCAGGGCTATATA GATCTCCCAGGTCAAGGGCGAG	
3	Mouse Ucp1 enhancer- promoter		GCATGCCAATTATAGTGGCGTCACTAACAGTACTGATACTTTA CATGCTAAAGTTAAAGTGTGTCTATATTAAATTGTAAGATTGGTG AAGAGAGGTGTTATCAGATGGAAGCTGCACATTCTGGATTAATG TGGTTAAATGTTCTCTCTGTGATTACTGTCTTATTCTCTCTTT TAAATATTGTCATTGGACATCTGTATAGCTACGCCCCTGTTGTTCTCTCTT ACGTCTCTGGAGACAGATAAGAAGTTACGACGGGAGGAGCAG ATGGAGGCAAAGCGCTGTGATGCTTTGTGGTTGAGTCACACA TTGTTCAAGTGTGAAATGAGTGAAGCAAATGGTGACCGGG TGCCCTGAAATGTTCTACATCTTAAGAGAGAAACACGGACA CTAGGTAAGTGAAGCTTGCTGTCACTCTTACAGCGTCAAGAG GGTCAGTCACCTTGACCACACTGAACTAGTCGTACCTTCACT CTTCTGCCCAGAAAGCAGAAATCAGACTCTGGGATATCAGC	

TABLE I -continued

Cis Regulatory Element Sequences			
SEQ ID NO	Identity	Sequence	
		CTCACCCCTACTGCTCTCTCATTATGAGGCAAACCTTCTTCACT TCCCAGGGCTCTGGGGCAGCAAGGTCACCCCTTCTCAGACT CTAGTCTCGGAGGAGATCAGATCGGCCATTTCAGGAACCCAGC CCTGCTCTGCCCTTGCTCAAGGTGTTGAAGAGTGCACAAAG GCACCACGCTGCGGGGACGCGGGTGAAGCCCTCTGTGTGTCCTC TGGGATAATCAGGAACCTGGTGCAAAATCAGAGGTGATGTTGCC AGGGCTTGGGAGTGACGCCGGCTGGGAGGCTTGCGCACCCAA GGCACGCCCTCCAAGTCCCCACTACAGCTCTTGAGACCTGG GCCGGCTCAGCCACTTCCCCAGTCCCTCCGGCAAGGGCTA TAGATCTCCAGGTCAAGGGCGAG	
4	Rat adipose-specific UCP1 enhancer	GACGTACAGTGGGTCACTGACCTTGATCACACTGCACAGTCT TCACCTTCCACGCTCTGCCAGAGCATGAATCAGGCTCTGG GGATACGGCCTCACCCCTACTGAGGCAAACCTTCTCCACTTCTC AGAGGCTCTGAGGGCAGCAAGGTCAGCCCTTCTTGGAATCTAG AACACTCCCTGCTTGAGCTGACATCACAGGGCAGCAGATGCA GCAGGGAAGGGCTGGGACTGGGACGTTCAAGGAAAGC TGTGGAACTTTTCAGCAACATCTCAGAAATCAGATGCACTTATT	
5	Rat basal UCP1 promoter	GAAATCAGATGCACTTATTCAAAGGGGCCAGGCCCTGCTCTGCG CCCTGGTGGAGGCTCTCATGTGAAGAGTGCACAAAGGCCACAT GTTGTGATAACGGGGCGAAGCCCCTCCGGTGTGCTCCAGGCAT CATCAGGAACACTAGTGCACAAAGCAGAGGTGCTGGCCAGGGCTTTG GGAGTGCAGCGGTCTGGAGGCTTGCGGCCAGGGCACCC CTGCCGATTCCACTAGCGCTTGGGGACCTGGGCCCTCT GCCCTCTCCAGCAATCGGGTATAAGCTTCCAAGTCAGGG CGCAGAAGTGGGGGATCGGGCTTAAAGAGCGAGAGGAAGG GACGCTCACCTTGAGCTCCTCACAAATAGCCCTGGCTGCC ACAGAAGTTCAAGTTGAGGTTGG	
6	Rat UCP1 enhancer with rat basal UCP1 promoter	GACGTACAGTGGGTCACTGACCTTGATCACACTGCACAGTCT TCACCTTCCACGCTCTGCCAGAGCATGAATCAGGCTCTGG GGATACGGCCTCACCCCTACTGAGGCAAACCTTCTCCACTTCTC AGAGGCTCTGAGGGCAGCAAGGTCAGCCCTTCTTGGAATCTAG AACACTCCCTGCTTGAGCTGACATCACAGGGCAGCAGATGCA GCAGGGAAGGGCTGGGACTGGGACGTTCACTTACAAGAAAAGC TGTGGAACTTTTCAGCAACATCTCAGAAATCAGATGCACTTATT CAAAGGAGCCAGGCCCTGCTCTGCCCTGGTGGAGGCTCTCAT GTGAAGAGTGCACAAAGGCACCATGTTGTGGATACGGGGGAAG CCCCTCCCGTGTCTCCAGGCATCATCAGGAACATGCTCCAAA GCAGAGGTGCTGCCAGGGCTTGGAGGTGACGCCGTGG GGCTTGTGCCAGGGCACGCCCTGCCGATTCCACTAGCAGG CTTGGGGGACCTGGCCGGCTTGCCCTCTCCAGCAATCGGG CTATAAGCTTCCAAGTCAGGGCAGAAGTGCAGGGGATCC GGGCTTAAAGAGCGAGAGGAAGGGACGCTCACCTTGAGCTCCT CCACAAATAGCCCTGGGGCTGCCACAGAAGTTCAAGGTTGAGA GTTCGG	
7	Human Ucp1 enhancer	TGATCAAGTGCATTGTTAATGTTACATTCTACATTTCAAAAAGGAA AGGAGAATTGTTACATTCAAGAACTGCTGCCACTCTTGTACG TCATAAAGGGTCAGTTGCCCTTGCTCATACTGACCTATTCTTAC TCTCTGCTCTCTTGTGCCAGAAAGGTGAGAAATCTGACCCCTTG GGGATACCCACCTCTCCCTACTGCTCTCCAACTGAGGCAA CTTTCTCTACTTCCAGAGGCTGTCAAGAAGTGGTGAAGCCAGCC TGCTCTGGAACTCCAGAAACTACTTCAAGAATCTGAACTCTGTG ACCTCTCAGGGTCCC	
8	Human Ucp1 promoter	ACCGCCGCGGTGCGCCCTCCCTCCGACGTGCGGTGTGCGGGGCG AGACAACCAAGCGGCCGCCAGGGTTTCGGGGAGGGAAAGCAGG CTCCCGAGGGCACCGAGCGAGAATGGGAATGGGAGGGACCCGGT GCTCCCGAACAGCCCCCGCAGGTCCACGCCGGTCTCTGA GACCTCGCGGCCAGGCCAGGGAGCGCCAGCTATATAAGTCC CAGCGGAGAGACGGAACGCAGAGGGCTGTGGCGAGGGT GGTAGGGGGACGCGGGGACT	
9	Human Ucp1 enhancer-promoter	TGATCAAGTGCATTGTTAATGTTACATTCTACATTTCAAAAAGGAA AGGAGAATTGTTACATTCAAGAACTGCTGCCACTCTTGTACG TCATAAAGGGTCAGTTGCCCTTGCTCATACTGACCTATTCTTAC TCTCTGCTCTCTTGTGCCAGAAAGGTGAGAAATCTGACCCCTTG GGGATACCCACCTCTCCCTACTGCTCTCCAACTGAGGCAA CTTTCTCTACTTCCAGAGGCTGTCAAGAAGTGGTGAAGCCAGCC TGCTCTGGAACTCCAGAAACTACTTCAAGAATCTGAACTCTGTG ACCTCTCAGGGTCCCACCGCCGGTGCAGGCCCTCCCTCCGACGTG CGGTGCGGGGGACGCGGGGACT	

TABLE I -continued

Cis Regulatory Element Sequences			
SEQ	ID	NO	Identity Sequence
			GGGGAGCGAAGCAGGGCTCCGAGGCACCGAGCGAGAATGGGA ATGGGAGGGACCCGGTGTCCCCGGACAGCCCCCGGCAGGGTCCC ACGGCCGGGTCTTCTGAGACCTCGGCGGGCCAGCCCCGGGAGCGG CCCAGCTATAATAAGTCCCAGCGGAAGACCGGAACGCAGAGGGTC CTGCTGGCGCGAGGGTGGTAGGAGGGGACGCGGGACT

Examples of Ucp1 promoters include those of SEQ ID NO: 2, SEQ ID NO: 5, and SEQ ID NO: 8 (from mouse, rat, and human Ucp1 genes, respectively). Examples of Ucp1 enhancers include those of SEQ ID NO: 1, SEQ ID NO: 4, and SEQ ID NO: 7 (from mouse, rat, and human Ucp1 genes, respectively). In an example, presence of such a Ucp1 enhancer or Ucp1 promoter, or both, or of other Ucp1 enhancer or promoter elements, or both, in the cis-regulatory element of a gene may drive transcription and expression of such gene only in BAT, or only at high levels in BAT, or only at detectable levels in BAT, or at substantially higher levels in BAT compared to other cell types. In another example, presence of such a Ucp1 enhancer or Ucp1 promoter, or both, or of other Ucp1 enhancer or promoter elements, or both, in the cis-regulatory element of a gene may also drive transcription and expression of such gene in liver cells. In some examples, a cis-regulatory element may include multiple Ucp1 enhancer elements, such as more than one of SEQ ID NO: 2, SEQ ID NO: 5, and SEQ ID NO: 8, or combination or combinations thereof.

A Ucp1 cis-regulatory element may include a sequence of SEQ ID NO: 3, SEQ ID NO: 6, or SEQ ID NO: 9. Or, it may include a Ucp1 promoter without a Ucp1 enhancer. A Ucp1 cis-regulatory element may also include combinations of a Ucp1 enhancer and a Ucp1 promoter other than the afore-

mentioned examples, such as any one or more of enhancer SEQ ID NO: 1, SEQ ID NO: 4, and SEQ ID NO: 7, together with any one of promoter SEQ ID NO: 2, SEQ ID NO: 5, and SEQ ID NO: 8. All possible combinations and permutations of the foregoing are explicitly contemplated herein and explicitly included as examples of the present disclosure.

A cis-regulatory element including a rat Ucp1 enhancer of SEQ ID NO: 4 and a rat Ucp1 promoter of SEQ ID NO: 5 has previously been shown to drive expression of a neighboring gene in a BAT-specific manner. US Patent Application Publication No. 2016/0319303A1. As disclosed herein, a cis-regulatory element including a mouse Ucp1 enhancer of SEQ ID NO: 1, a mouse Ucp1 promoter of SEQ ID NO: 2, or both (as in SEQ ID NO: 3) may also drive expression of a neighboring gene in a BAT-specific manner. Surprisingly, as further disclosed herein, a cis-regulatory element including a Ucp1 enhancer and a Ucp1 promoter may also induce expression in liver cells. Some examples may have a Ucp1 enhancer sequence, Ucp1 promoter sequence, or Ucp1 cis-regulatory element, with 90% or more, or 95% or more, or 97.5% or more sequence homology with any of the corresponding, foregoing examples.

Amino acid sequences, and non-limiting examples of nucleotide sequences encoding such Vgl14 peptide sequences, are shown in Table II.

TABLE II

Vgl14 sequences			
SEQ	ID	Identity	Sequence
10	Vgl14A	nucleotide	ATGCTATTATGAAGATGGACCTGTTGAACTATCAGTACTTGGAC AAGATGAACAACAAATATCGGCATTCTGTGCTACGAAGGGAAAGC TGCTCTCAGGGGAGAACCGAATACAGACCCCTGCCGGTGGCCTC TGCCTCTCAGCAGTCAACCGCACCGCCCTCCCCAATCAGCCCCAG CAAGAGGAAGTTCACTGGAGCCAGGTGACGAGGACCTAGACT GTGACAAACGACCACTGCTCCAAAATGAGTCGCATCTCAACCCCC ATCTGAACAAGAGCTGCCAATGGAGACTGCCAGAGACCCCCGG GAGGGAGGCCAGCCCCATCGAGCGCTGTGGCCCCCACCAT GAGCCTGCACGGCAGCCACCTGTACACCTCCCTCCCCAGCCTTG CCTGGAGCAGCCCTCGCACTGACCAAGAACAGCCTGGACGCCA GCAGGCCAGCCGGCTCTGCCAACACTGACCCGGGGAGCGGG CAGCAGAACCGGCCCTCCGTGATCACCTGTGCTCGCTGGCGCC CGCAACTGCAACCTCTCGCACTGCCCATCGGCACAGCGGCTGT GCCCGCCGGGGCTGCGCAGCTCCGGAGGGCACCGAGCGCTGC CACCACTGTGACCCCTGGTGGAGGAGCATTCCCGCAGGAGCCT GGGAAGAATTACAAGGAGCCCGAGCCGCACCCAACTCCGTGT CCATCACGGGCTCCGTGGACGACCACTTGCAAAGCTCTGGGTG ACACGTGGCTCCAGATCAAAGCGCCAAGGACGGAGCATCCAGC AGCCCTGAGTCCGCCTCTGCAGGGGCCAGCCGCCAGCCCCCTCT GCCACATGGTCAGCCACAGTCACTCCCCCTCTGTGGTCTCC
11	Vgl14A	amino acid	MLFMKMDLLNYQYLDKMNNNNIGILCYEGEAALRGEPRIQLPVASA LSSHRTGPPIPSPSKRKFMSMEPGDEDLDCDNHDVSKMSRIFNPFLNKT ANGDCRRDPRERSRSPIERAVAPTMISLHGSFLYTSPLSLGLEQPLALT KNSLDASRPAGLSPTLTPGERQQNRPSPVITCASAGRCNLSHCPIAH

TABLE II -continued

Vgll4 sequences		
SEQ ID NO	Identity Sequence	
	SGCAAPGPASYRRPPSAATTCDPVVEEHFRRSLGKNYKEPEPAPNSV SITGSVDDHFAKALGDTWLQIKAAKDGASSSPESASRRGQPASPSAH MVSHSHSPSVS	
12 Vgll4B nucleotide	ATGGAGACGCCATTGGATTTGTCCAGGGCAGCATCTCTGGTG CATCTGTGACGAAAAACCGCGAAGCTGCTCTCAGGGGAGAAC CAGATACAGACCCCTGCCCTGCGCTAGCAGTCACCG CACGGCCCTCCCCAATCAGCCCCACAAAGAGGAACCTCAGCAT GGAGCAGGTGACGAGGACTAGACTGTGACAACGACCGAC CCAAAATGAGTCGCATCTCAACCCCATCTGAACAAGACTGCCA ATGGAGACTGCCAGAGACCCCCGGAGCGGAGCCGAGCCC ATCGAGCAGCTGTGGCCACATGAGCTGCACGGCAGCC ATCGAGCAGCTGTGGCCACAGCGCTGTGGCCGCCCC CTGTACACCTCCCCCAGCTTGGCTGGAGCAGCCCTCGCA CTGACCAAAGACGCTGGACGCCAGCAGGAGCCGAGC GCCCAACTGACCCCCGGGGAGCGGAGCAGAACCGCCCTCCG TGATCACCTGTGCTCGCTGGCGCCGCAACTGCAACCTCTCGC ACTGCCCCATCGGCCACAGCGCTGTGGCCGCCCC GCTACCGGAGGCCACCGAGCTGCCACCCACTGTGACCCCTGG TGGAGGAGCATTTCCGAGGAGCTGGCAAGAATTACAAGGAG CCCGAGCCGGCACCAACTCCGTGTCATCACGGGCTCGTGGAC GACCACTTGCCAAGCTCTGGGTGACACGTGGCTCCAGATCAA GCGGCAAGGACGGAGCATCCAGCAGCCCTGCCCCACATGGTCAGCCACAGT CACTCCCCCTGTGGTCTCC	
13 Vgll4B peptide	METPLDVL SRAASL VHADDE KREA ALRGE PRI QTL PVAS ALSS HRTG PPFISPSKRKFMSME PGEDD LDCNDH VS KMSR I FNP HLN K TANG DCR RDP RRS RSP IER AVAP TMSL HGSH LY TSL PSL QPL A LTK NSL DA SRPAGL SPTL TP GERO QQ NRP SVI TC ASAG ARN C NL SHCP IA HSG CA AP GPAS YRRPP SA ATT CD PV VE EH FRR SLG KN YKE PE PAP N S VIT GSVD DH FAK ALG DT WL QI KA AK DG A SS SP ES AS RR GQP AS PS AH MV SH SH SP SVS	
14 Vgll4C nucleotide	ATGATTAAGTGAGGAACAAGACTGCCAATGGAGACTGCCGAG AGACCCCGGGAGCGCGAGCCAGCCCACATCGAGCGCTGTGG CCCCCACCATGAGCTGACGCCAGCCACCTGTACACCTCCCTCC CCAGCTTGGCTTGGAGCAGCCCTCGCACTGACCAAGAACAGCC TGGACGCCAGCAGCAGCAACCGGCCCTCGTGATCACCTGTGCCCTGG GGGAGCCAGCAGCAGCAACCGGCCCTCGTGATCACCTGTGCCCTGG CTGGCGCCGCAACTGCAACCTCTCGCACTGCCCATCGCGCAC GCGGCTGTGCCGCCGGGCTGCCAGCTACGGAGGCCACCG AGGGCTGCCACACCTGTGACCCCGTGTGGAGGAGCATTTCGC AGGAGCCTGGCAAGAATTACAAGGAGCCCGAGCCGGACCCCA CTCCGTGTCATCACGGGCTCCGTGGAGCACCACTTGCAAAGC TCTGGGTGACACGTGGCTCCAGATCAAAGCGCCAGGACGGAG CATCCAGCAGCCCTGAGTCCGCTCTCGCAGGGGCCAGCCGCCA GCCCTCTGCCACATGGTCAGCCACAGTCACCCCCCTGTGGT CTCC	
15 Vgll4C amino acid	MIKVRNKTANGDCRRDPRERSRSP IER AVAP TMSL HGSH LY TSL PSL GLEQPLALT KNSL DASR PAGL SPTL TP GERO QQ NRP SVI TC ASAG ARN C NL SHCP IA HSG CA AP GPAS YRRPP SA ATT CD PV VE EH FRR SLG KN Y KE PE PAP N S VIT GSVD DH FAK ALG DT WL QI KA AK DG A SS SP ES AS RR GQP AS PS AH MV SH SH SP SVS	
16 Vgll4D nucleotide	ATGACAAGACTGCCAATGGAGACTGCCGAGAGACCCCCGGGA GCGGAGCCGAGCCCATCGAGGCCCTGTGACACTCCCTCCAGCCTTGGCC GCCCTGTGACCCCTCGCACTGACCAAGAACAGCTGGAGCCAGC AGGCCAGCCGCCCTCGCCACACTGACCCGGGGAGCGGCCA GCAGAACCGGCCCTCGTGATCACCTGTGCCCTGGCTGGCGCCG CAACTGCAACCTCTCGCACTGCCCATCGCGCACAGCGCTGTGC CGGCCCGGGCTGCCAGCTACGGAGGCCACCGAGCGCTGCCA CCACCTGTGACCCCGTGGTGAGGAGCATTTCGGCAGGAGCCTGG GCAAGAATTACAAGGAGCCGAGCCGACCCACTCCGTGTC ATCACGGGCTCGTGAGCACCACTTGCCAAGCTCTGGGTGAC ACGTGGCTCAGATCAAAGCGCCAGGACGGAGCATCCAGCAG CCCTGAGTCCGCTCTCGCAGGGCCAGCCGCCAGCCCTCTGC CCACATGGTCAGCCACAGTCACCCCCCTGTGGTCTCC	

TABLE II -continued

		Vgll4 sequences
SEQ ID NO	Identity Sequence	
17	Vgll4D amino acid	MNKTANGDCRRDPRERSRSPIERAVAPTMISLHGSHLYTSLPSLGLQE PLALTNSLDAASRPAGLSPTLTPGERQNRPSVITCASAGARNCNLSH CPIAHSGCAAPGPASYRPPSAATTCDPVVEEHFRSLGKNYKEPEPA PNSVSIITGSVDDHFAKALGDTWLQIKAAKDGASSSPESASRRGQPAS PSAHMVSHTSHSPSVVS
18	Vgll4E nucleotide	ATGACTGAGAATAACGCATTTGACAAAATCCCTGAGTCCTGTGCA CTCAAAAGTTGGAGACATCCAGGTCATGCCACCATGGCGAAGCTGCT CTCAGGGAGAACCCAAGAATAACAGACCCCTGCCGGTGGCCCTCTGCC CTCACGACTCACCGCACCGGCCCTCCCCAATCAGCCCCAGCAAG AGGAAGTTTCAGCATGGAGCCAGGTGACGAGGACCTAGACTGTGA CAACGACCACTCTCAGGAAATGAGTCGATCTTCACCCCCCATCT GAACAAGACTGCCAATGGAGACTGCCAGAGACCCCCGGGAGC GGAGCCGAGCCCCATCGAGCGGCCTGGCCCCCACCATGAGCC TGCACGGCAGGCCACCTGTACACCTCCCTCCCCAGCCTTGGCCTGG AGCAGCCCCCTCGACTGACCAAGAACAGCCTGGACGCCAGCAGG CCAGCGGCCCTCTGCCAACACTGACCCCCGGGGAGCGGCCAGCA GAACCGGCCCCCTCGTGTACACCTGTGCTCGCTGGCCCTGGGCCCC CTGCAACCTCTCGCAGCTGCCACATCGGCCACAGCGGCTGTGCCG GCCGGGGCTCTGCCAGCTACCGGAGGCCACCGAGCGCTGCCACCA CCTGTGACCCCGTGGTGGAGGAGCATTCGGCAGGAGCCTGGGCA AGAATTAAGGAGCCGAGCCGACCCAACTCCGTGTCCATCA CGGGCTCGTGGACGACCACTTTGCAAAAGCTCTGGGTGACACG GGCTCGAGATCAAGCGCCCAAGGAGCGGAGATCCACAGCAGCCT GAGTCGCGCTCTCGCAGGGGCCAGCCCCCAGCCCCCTCTGCCAC ATGGTCAGCCACAGTCACTCCCCCTCTGTGGTCTCC
19	Vgll4E amino acid	MTENTHFDKIPESALKSWRHPGLHHGEAALRGEPRIQTLPVASALS SHRTGPPPISPSKRKFMSMEPGDEDLDCNDHVSKMSRIFNPHLNKTA NGDCRDRPRERSRSPIERAVAPTMISLHGSHLYTSLPSLGLEQPLALT NSLDASRPAGLSPTLTPGERQNRPSVITCASAGARNCNLSHCPIAHS GCAAPGPASYRPPSAATTCDPVVEEHFRSLGKNYKEPEPAPNSVSI TGSVDDHFAKALGDTWLQIKAAKDGASSSPESASRRGQPASPAHM VSHSHSPSVVS
20	Vgll4F nucleotide	ATGGAGCCAGGTGACGAGGACCTAGACTGTGACAACGACCACGT CTCAAAATGAGTCGATCTCAACCCCCATCTGAACAAGACTGC CAATGGAGACTGCCAGAGACCCCCGGAGCGGAGCGGCCAGCC CCATCGAGCCGCTGTGGCCCCCAGCATGAGCCTGACGGCAGGCC ACCTGTACACCTCTCCCAGCCTTGCCCTGGAGCACCCCTCG CACTGACCAAGAACAGCCTGGAGCAGCAGGCCAGCCGCTC TCGCCACACTGACCCGGGGAGCGCAGCAGAACCGGCCCTC CGTGTACACTGTGCTCGGCTGGCGCCGCAACTGCAACCTCTC GCACTGCCCATCGGCACAGCGCTGTGCCCGCCGGGCTG CAGCTACCGGAGGCCACCGAGCGCTGCCACACCTGTGACCCGT GGTGGAGGAGCATTCGGCAGGAGCCTGGGCAAGAATTACAAGG AGGCCAGCCGGCACCAACTCGTGTCCATCACGGCTCCGTGG ACGACCACTTTGCAAAAGCTCTGGGTGACACGTGGCTCCAGATCA AAGGGCCAAGGGAGCGACATCCAGCAGCCCTGAGTCGCCCT CGCAGGGGCCAGCCCCCAGCCCCCTGTGCCACATGGTCAGCCAC AGTCACCTCCCCCTGTGGTCTCC
21	Vgll4F amino acid	MEPGDEDLDCNDHVSKMSRIFNPHLNKTA NGDCRDRPRERSRSPIERAVAPTMISLHGSHLYTSLPSLGLEQPLALT NSLDASRPAGLSPTLTPGERQNRPSVITCASAGARNCNLSHCPIAHS GCAAPGPASYRPPSAATTCDPVVEEHFRSLGKNYKEPEPAPNSVSI TGSVDDHFAKALGDTWLQIKAAKDGASSSPESASRRGQPASPAHM VSHSHSPSVVS

At least six isoforms (A-F) of Vgll4 have been identified, referred to herein as Vgll4A, Vgll4B, Vgll4C, Vgll4D, Vgll4E, and Vgll4F, having amino acid sequences SEQ ID NO 11, SEQ ID NO 13, SEQ ID NO 15, SEQ ID NO 17, SEQ ID NO 19, and SEQ ID NO 21, respectively. These six isoforms are collectively included in the term Vgll4 as used herein. Also included herein is any nucleotide sequence that encodes any of the foregoing Vgll4 isoforms, including, without limitation, SEQ ID NO 10, SEQ ID NO: 12, SEQ ID NO 14, SEQ ID NO 16, SEQ ID NO 18, and SEQ ID NO: 20, including one or more codon substitution to any of the foregoing nucleotide sequences that nevertheless still

⁵⁵ encodes a Vgll4 (e.g., A-F), owing to codon degeneracy. A construct as disclosed herein may include a nucleotide sequence encoding a Vgll4 peptide as disclosed herein with any cis-regulatory element as disclosed herein, including without limitation one or more Ucp1 enhancer and a Ucp1 promotor, including any variation thereof described above.

A Vgll4 protein may be of human Vgll4 (SEQ ID NO 11, SEQ ID NO 13, SEQ ID NO 15, SEQ ID NO 17, SEQ ID NO 19, or SEQ ID NO 21), or mouse or rat Vgll4, or a Vgll4 sequence having at least 90%, at least 95%, or at least 97.5% homology with any of the foregoing examples in Table II. In an example, A Vgll4 peptide may include one or more amino

acid substitution (relative to the examples disclosed in Table II). In an example, a Vgll4 peptide may include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acid substitutions (relative to the examples disclosed in Table II). In an example, a Vgll4 peptide may have from 1 to 3 amino acid substitutions, or 2 amino acid substitutions, or 1 amino acid substitutions (relative to the examples disclosed in Table II). In an example, any of the foregoing amino acid substitutions may be outside of a TDU_1 and TDU_2 domain of the Vgll4 peptide.

In another example, a Vgll4 peptide may include two amino acid substitutions an a TDU domain, or two amino acid substitutions in each of two TDU domains. A TDU_1 domain has the amino acid sequence SEQ ID NO: 41. A TDU_2 domain has the amino acid sequence SEQ ID NO: 42. Each of SEQ ID NO 11, SEQ ID NO 13, SEQ ID NO 15, SEQ ID NO 17, SEQ ID NO 19, and SEQ ID NO 21 includes a TDU_1 domain with amino acid sequence SEQ ID NO: 41 and a TDU_2 domain with amino acid sequence SEQ ID NO: 42. In some examples, a TDU_1 domain may have an HF dipeptide amino acid sequence substituted with an AA dipeptide amino acid sequence, to yield the TDU_1 amino acid sequence SEQ ID NO: 43. In some examples, a

TDU_2 domain may have an HF dipeptide amino acid sequence substituted with an AA dipeptide amino acid sequence, to yield the TDU_2 amino acid sequence SEQ ID NO: 44. A Vgll4 peptide may include a TDU_1 domain having an amino acid sequence of SEQ ID NO: 43 instead of an amino acid sequence of SEQ ID NO: 41. A Vgll4 peptide may include a TDU_2 domain having an amino acid sequence of SEQ ID NO: 44 instead of an amino acid sequence of SEQ ID NO: 42. A Vgll peptide may include a TDU_1 domain having an amino acid sequence of SEQ ID NO: 43 instead of an amino acid sequence of SEQ ID NO: 41 and include a TDU_2 domain having an amino acid sequence of SEQ ID NO: 44 instead of an amino acid sequence of SEQ ID NO: 42. A Vgll4 peptide (e.g., Vgll4A, Vgll4B, Vgll4C, Vgll4D, Vgll4E, or Vgll4F) having TDU domains with amino acid sequences SEQ ID NO: 43 and SEQ ID NO: 44 instead of SEQ ID NO: 41 and SEQ ID NO: 42 is referred to herein as a Vgll4-HF4A peptide.

At least six isoforms (A-F) of Vgll4-HF4A are disclosed herein, referred to herein as Vgll4A-HF4A, Vgll4B-HF4A, Vgll4C-HF4A, Vgll4D-HF4A, Vgll4E-HF4A, and Vgll4F-HF4A. Amino acid sequences, and non-limiting examples of nucleotide sequences encoding such V 114-HF4A peptide sequences, are shown in Table III.

TABLE III

Vgll4-HF4A sequences				
	SEQ ID NO	Identity	Sequence	
22	Vgll4A- HF4A nucleotide		ATGCTATTATGAAAGATGGACCTGTGAACATTCAGTACTTGGAC AAGATGAACAACAAATATCGGCATTCTGTCTACGAAGGCGAAC TGCTCTCAGGGGAACCCAGAATACAGACCCCTGCCGTGGCC TGCCCTCAGCAGTCACCGCACCGCCCTCCCCAATCAGCCCC CAAGAGGAAGTTCAGCATGGAGCCAGGTGACGAGGAACCTAGACT GTGACAACGACCACGCTCCAAAATGAGTCGATCTTCACACCCCC ATCTGAACAAGAGACTGCCAATGGAGACTGCCGCAGAGACCCCC GAGCGGAGCGCAGCCCCATCGAGCGCTGTGGCCCCACCAT GAGCCTGACGGCAGGCCACTGTACACCTCCCCAGGCCCTGG CCTGGAGCAGCCCCCTCGCACTGACCAAGAACAGCTGGACGCC GCAGGGCAGCGGCCCTCGGCCACACTGACCCCCGGGGAGCG CAGCAGAACCGGCCCTCGCTGATCACCTGTGCTCGCTGGGCC CGCAACTGCAACCTCTCGCACTGCCCATCGCGCACAGCGCTGT GCCGGCAGGGGGCTGCCAGCTACCGAGGCCACCGAGCCTG CACCACTGTGACCCCTGCTGGAGGAGGCCAGCCCCAGGGAGCC TGGCAAGAATTACAAGGAGGCCAGGCCAACCAACTCCGT TCCATCACGGCTCCGTGGACGAGCAGCTGCCAACAGCTCGGT GACACGTGGCTCCAGATCAAAGGGCCAAGGGAGGCCATCCAG CAGCCCTGAGTCGCCCTCGCAGGGGCCAGGCCGCCAGCCCC TGCCCCATGGTCAGGCCACAGTCACCCCCCTGTGGCTCC	
23	Vgll4A- HF4A amino acid		MLFMKMDLLNYQYLDKMBNNNIGILCYESEAALRGEPRIQLTPVASA LSSHRTGPPPISPSKRKFMSMEPGDEDLDCNDHVSKMSRIFNPHLNK ANGDCRRDPRSRSPRIERAVAPTMISLHGSFLYTSPLSLGEQPLALT KNSLDAASRPAGLSPTLTPGERQONRPSVITCASAGARNCLSHCPIAH SGCAAPGPASYRPPSAATTCDPVBVEAARRSLGKNYKEPEPAPNSV SITGSVDDAAAKALGDTWLQIAKAADKGASSSPESASRRGQPASPAH MVSHTSHSPSVS	
24	Vgll4B- HF4A nucleotide		ATGGAGACGCCATTGGATGTTTGTCCAGGGCAGCATCTCTGGTG CATGCTGATGACGAAAAACCGCAAGCTGCTCTCAGGGAGAAC CAGAAATACAGACCCCTGCCGTGGCCCTCGCCCTCAGCAGTCACCG CACCGCCCTCCCCAATCAGCCCCAGCAAGAGGAAGTTCAAGCAT GGAGCCAGGTGACGAGGCCACTAGACTGTGACAAACGACCACGTCT CCAAAATGAGTCGATCTCAACCCCCATCTGAACAAGACTGCCA ATGGAGACTGCCAGAGACCCCCGGGAGCGGAGGCCAGCCCC ATCGAGCGCGCTGTGGCCCCCACCAGGCCCTGAGCCTGACGGCAGGCC CTGTACACCTCCCTCCAGCCTGTGGCTGGAGCAGGCCCTCGCA CTGACCAAGAACAGCTGGACGCCAGCAGGCCAGGCCCTCTC GCCCAACTGACCCGGGGAGGGCAGCAGAACCGGCCCTCG TGATCACCTGTGCTCGGCTGGCCCGCAACTGCAACCTCTCGC ACTGCCCATCGCGCACAGCGGCTGTGCCGCCGGGCTGCCA GCTACCGGAGGCCACCGAGCGCTGCCACACCTGTGACCCCCCTGG	

TABLE III-continued

Vgll14-HF4A sequences			
SEQ ID NO	Identity	Sequence	
		TGGAGGAGGCAGCCCGCAGGAGCTGGCAAGAATTACAAGGAG CCCGAGCGGCACCCAACTCCGTGTCATCACGGGCTCCGTGGAC GACCGAGCTGCCAAAGCTCTGGTGACACGTGGCTCCAGATCAA AGCGCCCAAGGGAGCATCCAGCAGCCCTGAGTCCGCTCTC GCAGGGCCAGGCCAGGCCCTGCCCACATGGTCAGCCACA GTCACTCCCCCTGTGGTCTCC	
25	Vgll14B- HF4A peptide	METPLDVLSLRAASLVHADDEKREAALRGEPRIQTLPVASALSSHRTG PPPISPSKRKFMSMEPGDEDLDCNDHVKMSRIFPNPHLNKTANGDR RDRPRSRSPIERAVAPTMSSLHGSHLYTSLPSLGLEPALTKNSLDA SRPAGLSPLTPGERQQNRPSTITCASAGARNCLSHCPIAHSGCAAP GPASYRRPSAATTCDPVVEEAARRSLGKNYKEPEPAPNSVSITGSVD DAAAKALGDTWLQIKAAKDGASSSPESASRRGQPSPSAHMVSHSH SPSVS	
26	Vgll14C- HF4A nucleotide	ATGATTAAGTGAGGAACAAGACTGCCAATGGAGACTGCCGAG AGACCCCGGGAGCGGAGCCGAGCCCACATCGAGCGCTGTGG CCCCACCATGAGCTGACGCCACCTGTACACCTCCCTCC CCAGCCTTGGCCCTGGAGCAGGCCCTCGCACTGACCAAGAACAGCC TGGACGCCAGCAGGCCAGCCGCTCTGCCACACTGACCCCG GGGAGCGGAGCAGAACCGGCCCTCGTGTACACCTGTGCCCTGG CTGGGCCGCAACTGCAACCTCTCGCACTGCCACATCGCGACA GCCGCTGTGCCGCCGGCTGCAAGCTACCGGAGGCCACCG AGGCTGCCACACCTGTGACCCCTGGAGCAGGCCAGCCCG CAGGAGCTGGCAAGAATTACAAGGAGGCCAGGCCACCCA ACTCCGTGTCATCACGGCTCCGTGGACGCGAGCTGCCAAAG CTCTGGGTGACACGTGGCTCCAGATCAAAGCGGCCAAGGACGGA GATCACAGCAGCCCTGAGTCCGCTCTGCCAGGGGCCAGGCC AGCCCCCTGCCACATGGTCAGCCACAGTCACCTCCCTCTGTG GTCTCC	
27	Vgll14C- HF4A amino acid	MIKVRNKTANGDCRRDPRERSRSPIERAVAPTMSSLHGSHLYTSLPSL GLEPALTKNSLDA SRPAGLSPLTPGERQQNRPSTITCASAGARN NLSHCPIAHSGCAAPGPASYRRPSAATTCDPVVEEAARRSLGKNY EPEPAPNSVSITGSVDDAAAKALGDTWLQIKAAKDGASSSPESASRR GQPSPSPSAHMVSHSHSPSVS	
28	Vgll14D- HF4A nucleotide	ATGAACAAGACTGCCAATGGAGACTGCCGAGAGACCCCGGGA GCGGAGCCGCAGCCCACATCGAGCCGCTGTGGCCCCCACCATGA GCCGACGGCAGCCACTGTACACCTCCCTCCGCCCTTGCC TGGAGCAGCCCTCGCACTGACCAAGAACAGCCTGGACGCCAGC AGGCCAGCGGCCCTCGTGTACCCCTGCCACACTGACCCGGGGAGCGGCA GCAGAACCCGGCCCTCGTGTACCCCTGCCACAGGGCTGTG CAACTGCAACCTCTCGCACTGCCACATCGGCCACAGGGCTGTG CGGCCCGGGCCCTGCCACAGGGAGGCCACCGAGCGCTGCC CCACCTGTGACCCCGTGGTGGAGGAGGCCAGGCCAGGAGCTG GGCAAGAATTACAAGGAGGCCAGGCCAGGCCACCAACTCCGTG CATCACGGGCTCCGTGGACGACCCAGCTGCCAAAGCTCTGGGTGA CACGTGGCTCCAGATCAAAGCGGCCAAGGAGCGGACATCCAGCA GCCCTGGAGTCCGCTCTGCCAGGGGCCAGGCCAGCCCCCTTG CCCCACATGGTCAGCCACAGTCACCTCCCTCTGTGGTCTCC	
29	Vgll14D- HF4A amino acid	MNKTANGDCRRDPRERSRSPIERAVAPTMSSLHGSHLYTSLPSLGLBQ PLALTKNSLDA SRPAGLSPLTPGERQQNRPSTITCASAGARNCLSH CPIAHSGCAAPGPASYRRPSAATTCDPVVEEAARRSLGKNYKEPEP APNSVSITGSVDDAAAKALGDTWLQIKAAKDGASSSPESASRRGQP SPSAHMVSHSHSPSVS	
30	Vgll14E- HF4A nucleotide	ATGACTGAGAATACGCATTTGACAAAATCCCTGAGTCTGTGCA CTCAAAGTTGGAGACATCCAGGTCTGCACCATGGGAAGCTGCT CTCAGGGAGAACCGAGAATACAGACCCCTGCCGGTGGCTCTGCC CTCAGCAGTCACCGCACCGGCCCTCCCCAATCAGCCCCAGCAAG AGGAAGTTGAGCATGGAGCCAGGTGACGAGGACCTAGACTGTGA CAACGACACAGTCTCAAATGAGTCGCATCTCAACCCCATCT GAACAAGACTGCCAATGGAGACTGCCAGAGAACCCGGGAGC GGAGCCGAGCCCCATCGAGCGGCTGTGGCCCCACCATGAGCC TGCACGGCAGCACCTGTACACCTCCCTCCCCAGCCTGGCTGG AGCAGCCCCCTCGCACTGACCAAGAACAGCCTGGACGCCAGCAG CCAGCGGCCCTCGCCCCACACTGACCCGGGGAGCGGAGCAGCA GAACCGGCCCCCTCGTGTACCTGTGCTGGCTGGCGCCCAA CTGCAACCTCTGCCACTGCCACATCGGCCACAGCGCTGTGCCG	

TABLE III-continued

Vgll4-HF4A sequences			
	SEQ ID NO	Identity	Sequence
			GCCCCGGCCTGCCAGCTACCGGAGGCCACCGAGCGCTGCCACCA CCTGTGACCCCGTGGTGGAGGAGGCAGCCCGCAGGAGCCTGGC AAAGATTACAAGGAGCCCGAGCCGGCACCCAACCTCCGTGTCATC ACGGGCTCCGTGGACGACGCAGCTGCCAAAGCTCTGGGTGACAC GTGGCTCCAGATAAAGCGGCAAGGACGGAGCATCCAGCAGCC CTGAGTCCGCCCTCTCGCAGGGCCAGGCCAGGCCCTCTGCC ACATGGTCAGCACAGTCACTCCCCCTCTGTGGTCTCC
31	Vgll4E- HF4A amino acid		MTENTHFDKIPESCALKSWRHPGLHGEAALRGEPRIQTLPVASALS SHRTGPPPISPSKRFSMEPGDDELDCDNDHVSKMSRIFNPHLNKTA NGDCRDRPRERSRSPIERAVENTMSLHGSHLYTSLPSLGLEQPLALT NSLDASRPAGLSPTLTGERQQNRPSVITCASAGARNCNLHSCHPIAHS GCAAPGPASYRRPPSAATTCDPVVEEARRLGKNYKEPEPAPNSVS ITGSVDDAAAKALGDTWLQIKAAKDGASSSPESASRRGQPASPSAH MVSHSHSPSVVS
32	Vgll4F- HF4A nucleotide		ATGGAGGCCAGGTGACCGAGGACCTAGACTGTGACAACGACCACGT CTCCAAAATGAGTCGCATCTCAACCCCCATCTGAACAAGACTGC CAATGGAGACTTGCGCGAGAACCCCGGGAGCGGAGCGCAGCC CCATCGAGCGCGCTGTGGCCCCACATGAGCTGACGGCAGCC ACCTGTACACCTCCCCTCCCAGCCTTGGCTGGAGCAGCCCCCTCG CACTGACCAAGAACAGCCTGGACGCCAGCAGGGCAGGCCCTC TCGCCACACTGACCCCCGGGGAGCGGGAGCAGAACCGGCCCTC CGTGTACACTGTGCTCTGGCTGGCCGCAACTGCAACCTCTC GCACTGCCCCATCGGCACAGCGCTGTGCCGCCGGGCTGC CAGCTACCGGAGGCCACCGAGGGCTGCCACCTGTGACCCCCGT GGTGGAGGAGGCAGCCGGCAGGGCTGGCAAGAAATTACAAGG AGCCCGAGCCGGCACCAACTCGTGTCCATCACGGCTCCGTGG ACGACGAGCTGACCAAGCTCTGGTGACACGTGGTCCAGATCA AAGCGGCCAAGGAGCGGAGCATCCAGCAGCCCTGAGTCCGCCCT CGCAGGGCCAGCCCGCCAGGCCCTCTGCCACATGGTCAGCCAC AGTCACCTCCCCCTGTGGTCTCC
33	Vgll4F- HF4A amino acid		MEPGDEDLDCNDHVSKMSRIFNPHLNKTANGDCRDRPRERSRSPIE RAVPTMSLHGSHLYTSLPSLGLEQPLALTNSLDASRPAGLSPTLT GERQQNRPSVITCASAGARNCNLHSCHPIAHSGCAAPGPASYRRPPSA ATTCDPVVEEAARRSLGKNYKEPEPAPNSVSITGSVDDAAAKALGDT WLQIKAAKDGASSSPESASRRGQPASPSAHMVSHSHSPSVVS

As disclosed herein, a Vgll4-HF4A peptide may have an amino acid sequence of s SEQ ID NO 23, SEQ ID NO 25, SEQ ID NO 27, SEQ ID NO 29, SEQ ID NO 31, or SEQ ID NO 33. These six isoforms are collectively included in the term Vgll4-HF4A as used herein. Also included herein is any nucleotide sequence that encodes any of the foregoing Vgll4 isoforms, including, without limitation, SEQ ID NO 22, SEQ ID NO: 24, SEQ ID NO 26, SEQ ID NO 28, SEQ ID NO 30, and SEQ ID NO: 32, including one or more codon substitution to any of the foregoing nucleotide sequences that nevertheless still encodes a Vgll4-HF4A (e.g., A-F), owing to codon degeneracy. A construct as disclosed herein may include a nucleotide sequence encoding a Vgll4-HF4A peptide as disclosed herein with any cis-regulatory element as disclosed herein, including without limitation one or more Ucp1 enhancer and a Ucp1 promotor, including any variation thereof described above.

A Vgll4-HF4A protein may be a human Vgll4, or mouse or rat Vgll4, bearing an HF to AA substitution in its TDU domains, or a Vgll4-HF4A sequence having at least 90%, at least 95%, or at least 97.5% homology with any of the foregoing examples in Table III. In an example, a Vgll4-HF4A peptide may include one or more amino acid substitution (relative to the examples disclosed in Table III)

⁴⁰ outside a TDU_1 and TDU_2 domain. In an example, a Vgll4-HF4A peptide may include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acid substitutions (relative to sequences disclosed in Table III) outside a TDU_1 and TDU_2 domain. In an example, a Vgll4-HF4A peptide may have from 1 to 3 amino acid substitutions, or 2 amino acid substitutions, or 1 amino acid substitution (relative to sequences disclosed in Table III) outside a TDU_1 and TDU_2 domain.

⁴⁵ In some examples, an intron may be included between a cis-regulatory element and a gene encoding Vgll4. In some examples, an intron may enhance or promote transcription or promote stability of an RNA transcript. Other examples do not include an intron. Various intronic sequences are known by skilled artisans to be able to be included in recombinant constructs for promoting gene expression, any of which could be included in a construct as disclosed herein. In an example, an intron of SEQ ID NO: 34 (a chimeric intron of human b-globin and immunoglobulin heavy chain genes) may be included, or a sequence having at least 90%, at least 95%, or at least 97.5% sequence homology therewith.

⁵⁰ A summary of aspects of a construct including a cis regulatory element and a Vgll4-encoding nucleotide sequence, and a cis regulatory element and a Vgll4-HF4A-encoding nucleotide sequence, as disclosed herein are shown in Tables IV and V, respectively.

TABLE IV

Vgll4 constructs			
cis-Regulatory Elements	Promoter sequences	Optional intron	Vgll4
Enhancer sequences		SEQ ID NO: 34	Amino acid sequences:
SEQ ID NO: 1	SEQ ID NO: 2	GTAAGTATCAAGG	SEQ ID NO: 11
SEQ ID NO: 4	SEQ ID NO: 5	TTACAAGACAGGT	SEQ ID NO: 13
SEQ ID NO: 7	SEQ ID NO: 8	TTAAGGGAGACCAA	SEQ ID NO: 15
		TAGAAACTGGGCT	SEQ ID NO: 17
		TGTCGAGACAGAG	SEQ ID NO: 19
		AAGACTCTTGCCTT	SEQ ID NO: 21
Enhancer-promoter sequences:		TCTGATAGGCACCT	Nucleotide sequences:
SEQ ID NO: 3		ATTGGTCTTACTGA	SEQ ID NO: 10
SEQ ID NO: 6		CATCCACTTTGCCT	SEQ ID NO: 12
SEQ ID NO: 9		TTCTCTCCACAG	SEQ ID NO: 14
			SEQ ID NO: 16
			SEQ ID NO: 18
			SEQ ID NO: 20

TABLE V

Vgll4-HF4A constructs			
cis-Regulatory Elements	Promoter sequences	Optional intron	Vgll4-HF4A
Enhancer sequences		SEQ ID NO: 34	Amino acid sequences:
SEQ ID NO: 1	SEQ ID NO: 2	GTAAGTATCAAGG	SEQ ID NO: 23
SEQ ID NO: 4	SEQ ID NO: 5	TTACAAGACAGGT	SEQ ID NO: 25
SEQ ID NO: 7	SEQ ID NO: 8	TTAAGGGAGACCAA	SEQ ID NO: 27
		TAGAAACTGGGCT	SEQ ID NO: 29
		TGTCGAGACAGAG	SEQ ID NO: 31
		AAGACTCTTGCCTT	SEQ ID NO: 33
Enhancer-promoter sequences:		TCTGATAGGCACCT	Nucleotide sequences:
SEQ ID NO: 3		ATTGGTCTTACTGA	SEQ ID NO: 22
SEQ ID NO: 6		CATCCACTTTGCCT	SEQ ID NO: 24
SEQ ID NO: 9		TTCTCTCCACAG	SEQ ID NO: 26
			SEQ ID NO: 28
			SEQ ID NO: 30
			SEQ ID NO: 32

A construct as disclosed herein may include a cis regulatory element and a nucleotide sequence encoding a Vgll4 or Vgll4-HF4A peptide. A cis regulatory element may include, for example, any one or more of enhancer sequence SEQ ID NO: 1, SEQ ID NO: 4, and SEQ ID NO: 7, and any one of promoter sequence SEQ ID NO: 2, SEQ ID NO: 5, and SEQ ID NO: 8. All permutations of the foregoing are expressly contemplated and included in the present disclosure. In an example, a cis-regulatory element may include 2, 3, or 4 enhancer sequences each independently selected from SEQ ID NO: 1, SEQ ID NO: 4, and SEQ ID NO: 7. Examples of cis regulatory elements include SEQ ID NO: 3, SEQ ID NO: 6, and SEQ ID NO: 9.

A Vgll4 peptide encoded by a Vgll4 peptide-encoding nucleotide sequence of a construct may include, for example, any of SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, and SEQ ID NO: 21, or any variation thereof as further explained above. Examples include SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, and SEQ ID NO: 20. Vgll4-HF4A peptide encoded by a Vgll4-HF4A peptide-encoding nucleotide sequence of a construct may include, for example, any of SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, and SEQ ID NO: 33, or any variation thereof as further explained above. Examples include SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, and SEQ ID NO: 32. A construct may include any cis regulatory element as disclosed herein and any Vgll4- or Vgll4-HF4A-encoding

nucleotide sequence as disclosed herein. Optionally, a construct may also include an intron between a cis regulatory element and a Vgll4- or Vgll4-HF4A-encoding nucleotide sequence. A non-limiting example of an optional intron is SEQ ID NO: 34. In other examples, nucleotides other than an intron or having an intronic nucleotide sequence other than SEQ ID NO: 34 may be included in a construct between a cis regulatory element and a Vgll4- or Vgll4-HF4A-encoding nucleotide sequence. An example of a Vgll4 construct is SEQ ID NO: 35 and an example of a Vgll4-HF4A construct is SEQ ID NO: 36.

A cell may be transfected with a construct as disclosed above by various methods, such as chemical transfection, electroporation, impalefaction, gene gun transfection, or viral vector mediated gene transfer, or any other method known to skilled persons in the relevant field. In an example, a Vgll4 gene with associated Ucp1 cis-regulatory element is packaged in a viral vector for cellular transfection. Viral vector in this case refers to a viral-like particle that contains or includes a payload gene construct or cassette capable of attaching to a cell and delivering the payload into the cell. In some examples, a viral vector may be of a type wherein a payload, once introduced into a transfected cell, integrates into the cell's genomic DNA, though such genomic integration is not an essential feature of a viral vector as disclosed herein. Viral vector may also refer to a gene sequence including a gene construct or cassette structured for inclusion in a viral-like particle. Examples of viral vectors include retroviruses, lentiviruses, adenoviruses, and adeno-

associated viruses (AAV). Several serotypes of AAV vectors are useful for cellular transfection, including any of serotypes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, and 11, or subtypes thereof. Sequences for such AAV serotypes are known and may be found in publicly accessible databases, as are methods of packaging a construct of interest in viral vector particles for cellular transfection and promotion of construct expression in transfected cells.

An AAV vector includes sequences bounding a payload construct referred to as inverted terminal repeats (ITRs). ITR sequences are involved in transcription of AAV genome, encapsulation of payload in a vector particle, genome multiplication for particle generation, and integration into host genome. A cassette, construct, transgene, payload, etc., placed between ITRs of an AAV vector may promote production of an AAV vector and/or expression of transfected gene within cells. In an example, a Ucp1 cis-regulatory element neighboring a nucleotide sequence encoding a Vgll4 or Vgll4-HF4 peptide may be placed between ITRs and used for generation of an AAV particle, wherein said particle may be contacted with cells of an organism to transfect them with such construct. An example includes an AAV9 serotype AAV containing such construct, though other serotypes may also be used.

In some cases a reporter gene may be used or included in a construct to verify expression of a construct gene included in a vector or for testing tissue- or cell-type specific expression of a gene under control of a given cis-regulatory element. Numerous reporter genes are known and have been widely used in the relevant field. A non-limiting list of examples includes a green fluorescent protein (for example, having an amino acid sequence of SEQ ID NO: 37, encoded for by a nucleotide sequence of SEQ ID NO: 38 or any other nucleotide sequence as may encode an amino acid sequence of SEQ ID NO: 37 according to principles codon degeneracy), a yellow fluorescent protein, a red fluorescent protein, a blue fluorescent protein, a luciferase protein, a beta-galactosidase protein, a glutathione S-transferase protein, a chloramphenicol acetyltransferase protein, and any combination of two or more of the foregoing. Other reporters may also be included. In other examples, no reporter is included. By detecting expression of a reporter protein, the ability of a given cis-regulatory element, or viral vector, to promote transfection and/or expression in various cell ad tissue types can be evaluated. A reporter protein sequence may occur immediately before the N-terminal or immediately after the C-terminal amino acid of a Vgll4 or Vgll4-HF4A peptide as disclosed herein, or may be separated by and one or more amino acids from the N- or C-terminal amino acid of a Vgll4 or Vgll4-HF4A peptide. A construct may include any nucleotide sequence for encoding any reporter protein. A non-limiting example of a Vgll4 construct including a sequence encoding a GFP-encoding reported protein is SEQ ID NO: 39. In another non-limiting example, SEQ ID NO: 39 may be modified to replace SEQ ID NO: 41 therein and SEQ ID NO: 42 therein with SEQ ID NO: 43 and SEQ ID NO: 44, respectively, to encode a Vgll4-HF4A and a GFP reporter protein.

A viral vector or viral-like particle, such as an AAV vector, can be injected into an organism, such as subcutaneously, intramuscularly, intravenously, intraperitoneally, or by other methods for introduction of the vector into the organism for contact with cells thereof. A vector may contact various different cell and tissue types and transfet them. However, inclusion of a cell- or tissue-specific cis-regulatory element (enhancer, promoter, or both) may restrict expression of the transfected gene to a given cell or tissue

type or types, wherein the construct is not transcribed or is otherwise dormant or at most barely or minimally expressed in other cell types. As explained above, a cis-regulatory element may include elements that are known or believed to drive expression in adipocytes, or fat cells, specifically, including in specific subtypes of fat cells, such as predominantly in BAT cells. However, it is not necessary that expression be limited absolutely to a given cell type, including only in BAT cells, even under control of a Ucp1 cis-regulatory element. For example, although Ucp1 expression is believed to be restricted to mature BAT cells, it is possible that other cell types may from time to time express a construct whose expression is influenced by a neighboring cis-regulatory element such as a Ucp1 enhancer, promoter, or both. Surprisingly, in some circumstances, as disclosed herein, a Ucp1 cis-regulatory element may include expression in liver cells in addition to BAT cells.

In an example, contacting the cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element and a coding sequence for a Vgll4 protein or variant may increase BAT volume, lower WAT volume, increase a ratio of BAT volume to WAT volume, or any combination of the foregoing, even if Vgll4 expression under control of the cis-regulatory element is not strictly limited to BAT cells containing the transgene. In another example, a construct may

In another example, transfecting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element and a coding sequence for a Vgll4 protein or variant may reduce a volume of adipose tissue of the organism. In another example, transfecting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element and a coding sequence for a Vgll4 protein or variant may reduce a mass ratio BAT to body weight of the organism.

In another example, disclosed herein is a method for prevention or treatment of obesity, by transfecting cells of an organism with any of the foregoing constructs disclosed herein including a cis-regulatory element and nucleotide sequence encoding a Vgll4 peptide or a Vgll4-HF4A peptide.

In another example, transfecting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element and a coding sequence for a Vgll4 protein or variant may reduce a liver volume, liver weight, intrahepatic fat content, or any combination of two or more of the foregoing, of the organism. An intrahepatic fat content of at least 5% of liver weight is referred to as hepatic steatosis. Obesity, or risk of developing obesity, such as genetic or life-style factors (e.g., high-calorie or high-fat diet, low exercise or caloric burn rate, sedentary lifestyle, etc.), are risk factors for developing elevated hepatic steatosis. Obesity may be defined as having a body mass index (BMI) of 30 or higher. An example of a risk factor for developing obesity may be having a BMI of from 25 to 29, which is considered being overweight. Prolonged hepatic steatosis is a risk factor for disorders including liver metabolic dysfunction, inflammation, and advanced forms of nonalcoholic fatty liver disease. Disclosed herein is a method for prevention or treatment of hepatic steatosis, by transfecting cells of an organism with any of the foregoing constructs disclosed herein including a cis-regulatory element and nucleotide sequence encoding a Vgll4 peptide or a Vgll4-HF4A peptide.

In another example, transfecting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element and a coding sequence for a Vgll4 protein or variant may reduce or minimize blood glucose levels or a rise in

29

glucose levels or duration of such rise in an organism. Obesity is a risk factor for diabetes, which includes pathological dysregulation of glucose levels, specifically pathological elevations in serum glucose levels or pathologically elevated duration of elevated serum glucose levels such as following calorie intake such as a meal. In an example, a rise in serum glucose may be measured following administration of a glucose challenge (i.e., consuming a glucose solution). Normally, a rise in serum glucose follows such a challenge, which rises then returns to baseline or near baseline. In individuals with diabetes, however, glucose may rise pathologically higher and/or for a pathologically longer duration than in individuals without diabetes. For individuals with diabetes or at risk for developing diabetes (e.g., family history, genetic or other biomarker-evinced predisposition, obesity, etc.), a treatment for preventing pathological rise in serum glucose levels, or a pathological extension of a rise in serum glucose levels, following a meal or a glucose challenge is advantageous. An example in accordance with the present disclosure includes reducing or minimizing blood glucose levels or a rise in glucose levels in an organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct. An example in accordance with the present disclosure includes preventing development of a pathologically high rise in blood glucose levels or a pathologically high duration of a rise in blood glucose levels in an organism with diabetes or at risk for developing diabetes by contacting an organism with the construct, such as by transfecting cells of the organism. The organism may be an obese person, or a person at risk of developing obesity, or a person diagnosed with diabetes, or a person at risk of developing diabetes. Accordingly, an example disclosed herein includes a method for prevention or treatment of diabetes, by transfecting cells of an organism with any of the foregoing constructs disclosed herein including a cis-regulatory element and nucleotide sequence encoding a Vgll4 peptide or a Vgll4-HF4A peptide.

In an example, also disclosed is increasing expression of mitochondrial genes, such as mitochondrial genes involved in mitochondrial respiration, in an organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct. In an example, also disclosed is decreasing expression of genes that promote lipogenesis, in an organism by contacting an organism with the construct, such as by transfecting cells of an organism with the construct. Increasing mitochondrial genes that promote mitochondrial respiration, or decreasing expression of genes involved in lipogenesis, such as in BAT or liver cells of an organism transfected with a construct as disclosed herein (e.g., including a Ucp1 cis regulatory element and a nucleotide sequence encoding a Vgll4- or Vgll4-HFA-encoding nucleotide, as disclosed herein), may include advantageously promote BAT levels, decrease lipogenesis in adipose cells, decrease hepatic steatosis, or any combination of the foregoing.

As disclosed herein, in an example, contacting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element neighboring a Vgll4 coding sequence surprisingly increases BAT volume, i.e. the volume occupied by BAT cells. In another example, contacting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element neighboring a Vgll4 coding sequence surprisingly decreases WAT volume. In another example, contacting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element neighboring a Vgll4 coding sequence surprisingly increases a ratio of

30

BAT volume to WAT volume. An increase in BAT by driving Vgll4 expression under control of a Ucp1 cis-regulatory element, known to increase expression of a neighboring gene in BAT cells, is particularly unexpected given that increased Vgll4 expression is known to promote apoptosis or otherwise have anti-tumor cell effects, unlike other members of the Vgll family. By comparison, Vgll3 levels are increased in WAT cells in obese mice, suggesting that Vgll3 may promote WAT cells, whereas over-expression of Vgll3 inhibits adipogenesis overall. U.S. Pat. No. 8,852,939.

As further disclosed herein, in an example, contacting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element neighboring a Vgll4 coding sequence or a Vgll4-HF4A sequence in some cases may promote expression of Vgll4 or Vgll4-HF4A respectively, in liver.

In another example, contacting cells of an organism with a nucleotide sequence including a Ucp1 cis-regulatory element neighboring a Vgll4-H1F4A coding sequence may surprisingly decrease adipose tissue volume. In an example, a volume of BAT is decreased. Without limiting the present disclosure to any particular mechanism of action, decreased volume of BAT following transfection with such a construct may be related to stimulation of expression of mitochondrial genes involved in mitochondrial respiration, inhibition of expression of genes involved in lipogenesis, an inhibition of lipogenesis, or any combination of two or more of the foregoing, in BAT, caused by the transfection. In another example, hepatic steatosis is decreased. Without limiting the present disclosure to any particular mechanism of action, decreased hepatic steatosis following transfection with such a construct may be related to stimulation of expression of mitochondrial genes involved in mitochondrial respiration, inhibition of expression of genes involved in lipogenesis, an inhibition of lipogenesis, or any combination of two or more of the foregoing, in liver, caused by the transfection.

EXAMPLES

The following examples are intended to illustrate particular embodiments of the present disclosure, but are by no means intended to limit the scope thereof.

Example 1: A Ucp1 Cis-Regulatory Element Drives Expression in BAT when Transfection Occurs During Early Development

FIG. 1 shows examples of polynucleotides in accordance with certain aspects of the present disclosure. The example includes a cis-regulatory element (BCE) including a mouse Ucp1 enhancer and mouse Ucp1 promoter upstream of a coding sequence of a reporter protein (either a green fluorescent protein (GFP) or firefly luciferase (Luc)). In these examples, a chimeric intron of human b-globin and immunoglobulin heavy chain genes is included between the BCE and the reporter sequence. Inclusion of an intron may increase expression of the payload gene sequence, here the reporter constructs. Also included at the 5' and 3' ends, flanking the BCE and reporter, are inverted terminal repeat (ITR) sequences. Adeno-associated viral vectors were synthesized incorporating constructs as illustrated for determining an ability of BCE to drive expression of a downstream coding sequence. In an example, adeno-associated viral vectors of serotype 9 (AAV9) were constructed carrying constructs as illustrated in FIG. 1. Constructs include a BCE having a sequence of SEQ ID NO: 3 and an intron of SEQ ID NO: 34.

FIG. 2 illustrates time course of treatment of mice with a viral vectors as illustrated in FIG. 1 for determining expression patterns of reporter proteins driven by BCE. At 5 days of age (P5), neonatal mice received dorsal sc AAV injection (of an AAV9 carrying one of the constructs shown in FIG. 1), at a dose of 1×10^{10} genome copies per gram of body weight, in phosphate-buffered saline. At 42 days after birth (P42), bioluminescence was assessed to determine expression of marker proteins. Expression was determined by an in vivo imaging system (IVISTM, Perkin Elmer) as shown in FIG. 3 (dorsal view) and 4 (ventral view), as well as by micro CT scanning to show topographical expression patterns, as shown in FIG. 5. AAV9.BCE.luci transduced subjects were first imaged for bioluminescence, and then scanned by micro CT, with AAV9.BCE.GFP serving as a control. In FIG. 5, bioluminescence signal origins were matched with tissues mapped by micro CT (pink=brown adipose tissue, blue or red=bioluminescence signal positive tissues). Expression was specifically elevated in BAT, demonstrating the BAT-specific expression driven by the BCE cis-regulatory element following transfection during the neonatal period.

As shown in FIGS. 6 and 7, interscapular adipose tissues (i.e., in the region where BAT is located) were collected at postnatal day 60 and used for immunofluorescence staining. FIG. 6 shows immunofluorescence staining images of interscapular BAT (using UCP1 imaging to identify BAT cells), and FIG. 7 shows immunofluorescence staining images of both BAT and white adipose tissue (using perilipin staining to mark adipose tissue, both BAT and WAT). Nuclear stain DAPI is also shown in FIG. 6 and GFP expression was used to stain cells with BCE-driven reporter expression in animals treated with AAV.BCE.GFP (with AAV.BCE.Luci treatment serving as control). Bar=200 μ m. BCE drove expression of reporter protein in adipocytes, and in BAT in particular.

FIG. 8 shows a schematic view of a pAAV.BCE.Vgll4-GFP construct administered via an AAV9 carrier. The construct resembles that shown in FIG. 1 except that Vgll4-GFP is the coding sequence whose expression is driven by BCE rather than merely GFP or luciferase. The sequence of the BCA-intron-Vgll4-GFP transcript is SEQ ID NO:39, and is a polynucleotide including a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein, wherein the cis-regulatory element includes an uncoupling protein 1 enhancer and an uncoupling protein 1 promoter. An intron is present between the cis-regulatory element and nucleotide sequence encoding a vestigial like 4 protein. The construct was packaged in an adeno-associated viral vector (AAV9) for transfection of cells of an organism with the construct.

FIG. 9 illustrates time course of treatment of mice with a viral vectors as illustrated in FIG. 8 for determining effects on adipose tissue volume. On postnatal day 5 mice were injected dorsally sc with AAV (AAV9.BCE.Vgll4-GFP, or AAV9.BCE.GFP, carrying the construct shown in FIG. 1, as a control), at a dose of 1×10^{10} genome copies per gram of body weight, in phosphate-buffered saline. On postnatal day 42, a micro CT scan was taken to measure intrascapular adipose tissue volume, using standard commercially available micro CT scan software (analyze 12TM). Intrascapular adipose tissue was differentiable from neighboring tissue, and BAT was differentiable from WAT, due to differences in Hounsfield units associated with differing tissue types according to standard micro CT scan techniques. An example scan is shown in FIG. 10 (control) and 11 (for a subject injected with AAV9.BCE.Vgll4-GFP). BAT and

WAT are as indicated. Increased volume of BAT and decreased volume of WAT can be seen following BCE-driven expression of Vgll4.

BAT volume was measured as graphically represented in FIG. 12. In subjects with BCE-driven Vgll4 expression, BAT volume was an average of 102.2 ± 2.8 mm³, compared to 84.2 ± 1.9 mm³ in controls (*= $p < 0.05$). WAT volume was measured as graphically represented in FIG. 13. In subjects with BCE-driven Vgll4 expression, WAT volume was an average of 27.1 ± 0.6 mm³, compared to 40.0 ± 2.8 mm³ in controls (**= $p < 0.01$). Ratio of BAT/WAT volume was measured as graphically represented in FIG. 14. In subjects with BCE-driven Vgll4 expression, BAT/WAT volume ration was an average of 3.8 ± 0.6 mm³, compared to 40.0 ± 2.8 mm³ in controls (**= $p < 0.01$). Thus, transfection of cells of an organism with a construct including a cis-regulatory element driving expression of Vgll4, wherein the cis-regulatory element includes an uncoupling protein 1 enhancer and an uncoupling protein 1 promoter, increased BAT volume, decreased WAT volume, and increased a ration of BAT volume to WAT volume.

Example 2: A Ucp1 Cis-Regulatory Element Drives Expression in BAT and Liver when Transfection Occurs During Adulthood

An AAV vector (AAV9) containing a construct with a BCE cis regulatory element (SEQ ID NO: 3) driving expression of luciferase (AAV.BCE.Luci) was administered to 6 weeks old mice. In other mice, the cis regulatory element (MiniUcp1) included a Ucp1 enhancer (SEQ ID NO: 4) but not a Ucp1 promoter (AAV.MiniUcp1.Luci). Subjects were tested for luciferase signals one week later. Results are illustrated in FIGS. 15-18. Both BAT (FIG. 16) and liver (FIG. 17) had luciferase signals, with BCE cis regulatory element driving higher liver expression than MiniUcp1. AAV.BCE.Luci drives expression in liver when transfection occurs later in development such as in adulthood. FIG. 18 shows that BAT and liver included viral vector genome copies, as assessed by real-time PCR.

AAV has very low chance of integrating into the host cell genome, existing primarily as episomes in host cells. If the host cells proliferate rapidly, daughter cells may easily lose AAV copy number. In cell cycle quiescent cells, in contrast, AAV coexist until cells die. During development, such as in neonates, hepatocytes are rapidly proliferating, whereas BAT are mostly cell cycle quiescent cells. Without being limited to any particular theory or mechanism of action, during the growth of AAV.BCE.Luci transduced pups, hepatocytes but not the brown adipocytes may have shed AAVs, which could explain why luciferase signals can only be detected in the BAT but not in the liver following transfection early in development (as in Example 1). In the AAV.BCE.Luci transduced adult mice of the present Example (Example 2), BAT and hepatocytes were not actively proliferating, and luciferase signals were detected in both tissues.

Example 3. The Hippo-YAP Pathway

The Hippo-YAP signaling pathway is well known for controlling organ growth. In mammals, the Hippo kinase cascade includes MST1/2, LATS1/2, and the scaffold protein Salvador (Sav). Activation of these kinases results in phosphorylation and inactivation of YAP and WWTR1 (more commonly known as TAZ), orthologous transcriptional coactivators that are terminal effectors of this pathway.

33

YAP/TAZ interact with TEAD family transcription factors to regulate downstream target genes expression. Vestigial like 4 (VGLL4) is another co-transcriptional factor that serves as a suppressor of a YAP-TEAD complex. Mechanistically, VGLL4 directly binds to TEAD through its two TONDU (TDU) domains, and the binding of VGLL4 or YAP to TEAD is mutually exclusive (FIG. 19).

Each VGLL4 TDU domain has two essential amino acid residues (HF) mediating VGLL4-TEAD interaction (SEQ ID NO: 41 and SEQ ID NO: 42, respectively). Replacing the HFs in the TDU domains with four alanine residues (FIG. 20) minimizes the interaction between VGLL4 and TEAD. Unlike VGLL\$, VGLL-HF4A does not suppress a YAP-TEAD complex (FIG. 21). YAP/TAZ may promote BAT thermogenesis, raising the possibility of manipulating this pathway to reduce obesity.

Example 3. TEAD1 Directly Regulates the Expression of Fgf21

Double heterozygous YAP and TAZ knockout mice have previously been shown to have much smaller BAT than their littermate controls at four weeks after birth. In the Hippo-YAP pathway, YAP/TAZ interacts with TEAD proteins to regulate downstream targets expression. Thus, TEAD1 may regulated the postnatal growth of BAT. As disclosed herein, Ucp1::Cre transgenic mice were crossed with Tead1 flox allele to specifically delete Tead1 in the BAT (FIG. 22). TEAD1 depletion in BAT of Tead1 cKO mice was confirmed by western blot (FIG. 23). Compared with controls, the Tead1 cKO mice had smaller interscapular BAT deposits (FIGS. 24 and 25). Knocking out TEAD1 in the brown adipocytes significantly decreased the expression of Fgf21 in BAT (FIG. 26). Chromatin immunoprecipitation sequencing data demonstrated that TEAD1 directly binds to the promoter region of Fgf21 (FIG. 27). Fibroblast growth factor 21 (FGF21) is an important myokine that regulates glucose-lipid metabolism.

Example 4. Activation of VGLL4 Reduces Adiposity

BAT plays important roles in non-shivering thermogenesis and energy homeostasis. As disclosed herein, AAV-mediated overexpression of VGLL4 increased BAT volume. To demonstrate whether activation of VGLL4 in an obesity model (mice fed on a high fat diet) would reduce body weight, AAV.BCE.VGLL4 (including SEQ ID NO: 39) into high fat diet induced obesity mice, and their body weight monitored for 7 weeks (FIG. 28). Controls received AAV with luciferase controlled by the cis regulatory element (SEQ ID NO: 3). 8 weeks after AAV delivery, luciferase signals were easily detected in the AAV.BCE.luci transduced mice (FIG. 29). Body weight and body weight gain values were not distinguishable between control and VGLL4 treated mice (FIGS. 30 and 31). However, using micro CT, the volume of adipose and non-adipose tissue (lean mass) were measured. 4 weeks after AAV transduction, although the body weight gain was similar between control and VGLL4 mice, the control but not the VGLL4 mice showed a significantly increase in adipose tissue mass. Meanwhile, VGLL4 but not the control mice had a significant increase of lean mass. Consequently, 4 weeks after AAV infusion, the VGLL4 mice had a lower fat/lean test ratio than the control mice (FIGS. 32A-C). Western blot showed that exogenous VGLL4 was expressed in BAT of AAV.BCE.VGLL4 transduced mice (not shown).

34

Example 5. AAV.BCE.VGLL4 Mitigates Body Weight Gain

To demonstrate whether activation of VGLL4 in normal mice would prevent or mitigate the progression of obesity, AAV.BCE.VGLL4 (including SEQ ID NO: 39) was administered to 8-week-old mice. Beginning 1 week after injection, mice were fed a high-fat diet (FIG. 33A). At the end of 13 weeks high fat diet treatment, the body weight of VGLL4 mice was lower than the control mice, though the difference did not reach statistical significance (FIG. 33B). Starting from week 5, however, VGLL4 mice had a significantly lower accumulated body weight gain than controls (FIG. 33C). As disclosed herein (Example 2), AAV.BCE.VGLL4 targets both BAT and liver when administered later in development such as in adulthood. Fgf21 is a target of TEAD1 (FIGS. 26 and 27), and is mainly produced by liver and adipose tissue. Expression of VGLL4 and Fgf21 was therefore measure in liver. VGLL4 was overexpressed in liver, and that Fgf21 was significantly decreased in liver of VGLL4 mice (FIG. 33D).

Example 6. AAV.BCE.VGLL4HF4A Increases BAT Mitochondrial Genes Expression

As disclosed herein, VGLL4 expression driven by a Ucp1 cis regulatory element mitigated body weight gain, it also suppressed the expression of Fgf21, which is important for glucose metabolism. VGLL4 may therefore have multiple roles, interacting with TEAD1 to decrease Fgf21 expression, while also interacting with other unknown factors to improve energy expenditure. Without being limited to any particular theory or mechanism of action, this possibility may indicate why VGLL4 may mitigate body weight gain without improving glucose metabolism. An AAV.BCE.VGLL4HF4A vector was created, which expresses a mutated VGLL4 that does not interact with TEAD (H1F4A mutations), including SEQ ID NO: 36, and also including a GFP reporter protein (SEQ ID NO: 40). 8-week-old normal mice received subcutaneous injection of AAV.BCE.VGLL4HF4A, resulting in transfection of BAT (FIG. 34A and B). qRT-PCR results showed that over-expression of VGLL4HF4A did not affect the expression of Fgf21 but reduced the expression of Cidea and Fasn (FIG. 34C), which are two genes involved in lipogenesis, in BAT. Additionally, VGLL4HF4A increased the expression of Cox2 (Cytochrome C Oxidase Subunit II, encoded by MT-CO2) and Cox6a2 (Cytochrome C Oxidase Subunit 6A2) in BAT (FIG. 34D). These data indicate that VGLL4HF4A may increase mitochondrial respiration activity without affecting Fgf21 expression, and also reduce lipogenesis.

Example 7. AAV.BCE.VGLL4HF4A Mitigates Body Weight Gain and Reduces Serum Glucose Level

AAV.BCE.VGLL4HF4A (including SEQ ID NO: 36) was injected sc to the inter-scapular region of 8 week old C57/BL6 mice, at a dosage of 2×10^9 GC/gram body weight. AAV.BCE.GFP was used as control. One week after virus injection, 12 weeks of feeding with high fat diet (HFD) began (FIG. 35A). During HFD treatment, body weight gain rate of AAV.BCE.VGLL4HF4A mice (VGLL4HF4A) was slower than that of the AAV.BCE.GFP mice, and the difference reached to significance at 9 weeks after HFD treatment (FIG. 35B). 11 weeks after high fat diet treatment, the body weight of AAV.BCE.VGLL4HF4A mice started to become

35

significantly lower than that of the AAV.BCE.GFP mice (FIG. 35C). A glucose tolerance test (GTT) showed that the starving serum glucose level and glucose peak level following glucose challenge was significantly lower in the AAV.BCE.VGLL4HF4A mice (FIG. 35D).

Example 8. AAV.BCE.VGLL4HF4A Reduces BAT Weight

Compared to that of the GFP control mice, the mass of BAT was significantly lower in the VGLL4HF4A mice (FIG. 36A). qRT-PCR confirmed that VGLL4 was overexpressed in the BAT of VGLL4HF4A mice (FIG. 36B). The expression of Ucp1 was not affected by VGLL4HF4A (FIG. 36C). Expression levels of three more genes that regulate mitochondria respiration activity were also measured in BAT: Cox2, Cox6a, Ndufsa8. Cox2 was significantly upregulated in the VGLL4HF4A BAT (FIG. 36D). VGLL4HF4A also suppressed expression of Acc1, a gene involved in fatty acid synthesis (FIG. 36E). VGLL4HF4A may therefore preserve BAT function by both increasing the mitochondria respiration activity and attenuating fatty acid synthesis.

Example 9. AAV.BCE.VGLL4HF4A Reduces Liver Weight and Fatty Acid Synthesis

A BCE cis regulatory element including a Ucp1 enhancer and Ucp1 promoter as disclosed herein drives gene expression in liver when adult subjects are transfected, as disclosed herein. Interestingly, liver weight of VGLL4HF4A mice was

36

significantly lower than that of the GFP control mice (FIG. 37A). qRT-PCR confirmed that VGLL4 was robustly over-expressed in liver of VGLL4HF4A mice (FIG. 37B). Histology determined by haematoxylin and eosin (H&E) staining showed moderate lipid droplets accumulation and microsteatosis (FIG. 37C). Oil red staining confirmed that VGLL4HF4A liver had much less lipid droplets accumulation than the GFP control liver (FIG. 37D). The expression of fatty acid synthesis genes Acc1 and Fasn in liver were significantly reduced by VGLL4HF4A (FIG. 37E). VGLL4HF4A may prevent HFD induced liver pathologies such as liver metabolic dysfunction, inflammation, or non-alcoholic fatty liver disease

It should be appreciated that all combinations of the foregoing concepts and additional concepts discussed in greater detail herein (provided such concepts are not mutually inconsistent) are contemplated as being part of the inventive subject matter disclosed herein. In particular, all combinations of claimed subject matter appearing at the end of this disclosure are contemplated as being part of the inventive subject matter disclosed herein and may be used to achieve the benefits and advantages described herein.

Although examples have been depicted and described in detail herein, it will be apparent to those skilled in the relevant art that various modifications, additions, substitutions, and the like can be made without departing from the spirit of the present disclosure and these are therefore considered to be within the scope of the present disclosure as defined in the claims that follow.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 44

<210> SEQ ID NO 1
<211> LENGTH: 638
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 1

gcatgc	aat ttat	agt g	ccat	ttt a	acat	gc ta	agg t	ttaaa	60	
gtgtgt	ctata	attt	gtt a	ggtg	a aag	ttt g	aggat	g aag	ctgcac	120
at	ttt c	tggat	aat	gttat	c t	tct	cgt	tga	ttact	180
ttt	aaaat	at	gtc	at	ttt	gtt	gt	tct	tttct	240
ttt	aaaaat	at	at	at	ttt	gtt	at	ttt	cttgg	300
ttt	at	at	at	at	ttt	gtt	at	ttt	gtt	360
ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	420
ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	480
ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	540
ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	600
ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	ttt	638

<210> SEQ ID NO 2
<211> LENGTH: 335
<212> TYPE: DNA
<213> ORGANISM: Mus musculus

<400> SEQUENCE: 2

tctcg	ggagg	aa	gtc	atcg	cg	ctt	atttca	agg	aaacc	cc	ctgtct	gc	ccctgg	60
-------	-------	----	-----	------	----	-----	--------	-----	-------	----	--------	----	--------	----

-continued

ccaaggctgt	tgaagagtga	caaaaaggcac	cacgctcggg	ggacgcgggt	gaagccccctc	120
tgtgtgtccct	ctgggcataa	tcaaggaactg	gtgccaaatac	agaggtgatg	tggccagggc	180
tttggggagtq	acgcgcggct	gggaggcttg	cgcacccaaag	gcacgcggct	gcacaagtccc	240
actagcagct	ctttggagac	ctggggccggc	tcaagccactt	cccccaagtcc	ctcctccggc	300
aaggggctat	ataqatctcc	caggtaaaaa	ccggaa			335

```
<210> SEQ ID NO 3
<211> LENGTH: 973
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Laboratory-synthesized sequence.
```

<400> SEQUENCE: 3

gcatgccaat ttatagtgcc gtcactaaca gtactgatac tttacatgc taagttaaa
gtgtgtgtca tattaattgt aagattggtg aagagaggta ttatcagatg gaaggctgcac 120
atttctggat taatgtggtt aaatgtatct tctccctgtgta ttactgtctt tattttctt 180
tttaaaatat tgtcatttgg acatctatct gtatagctac gccctgacac gtcctctgg 240
agacagataa aaagttacga cgggaggaggc agatggaggc aaagcgctgt gatgttttg 300
tggtttgagt gcacacatTT gttcagtgtat tctgtgaaat gagtgagcaa atggtgaccg 360
ggtgcccgtg aaatgggttt ctacatcttta agagaagaac acggacacta ggtaagtggaa 420
gcttgctgtc actcctctac agcgtcacag agggtcagtc acccttgacc acactgaact 480
agtcgtcacc ttccactct tcctgccaga agagcagaaa tcagactctc tggggatatc 540
agcctcaccc ctactgctct ctccattatg aggcaaaactt tctttcactt cccagaggct 600
ctggggggcaag caaggtaaac ccttcctca gactctagtc tcggaggaga tcagatcgcg 660
cttattcaag ggaaccagcc cctgtctgc gccctggtcc aaggctgtt aagagtgaca 720
aaaggccacca cgctgegggg acgcgggtga agccccctgc tgggtctctt gggcataatc 780
aggaacttgtt gccaaatcag aggtgtatgtt gcccagggtt tgggagtgac ggcggctgg 840
ggggcttgcg cacccaaaggc acggccctgc caagtccccac tagcagetct ttggagacct 900
gggcccggctc agccacttcc cccagtcctt cctccggcaa ggggtatata agatctccca 960
ggtcaggggc cag 973

<210> SEQ ID NO 4
<211> LENGTH: 295
<212> TYPE: DNA
<213> ORGANISM: *Rattus norvegicus*

<400> SEQUENCE: 4

gacgtcacag tgggtcagtc acccttgatc acactgcacc agtcttcacc tttccacgct	60
tcctgccaga gcatgaatca ggctctctgg ggataccggc ctcaccccta ctgaggcaaa	120
ctttctccca cttctcagag gctctgaggg cagcaaggctc agccctttct ttggaatcta	180
gaaccactcc ctgtctttag ctgacatcac agggcaggca gatgcagcag ggaaggcct	240
qqdqactqqqa cqttcatctt acaaaqaqtc tqtqqaactt ttcaqcaaca tctca	295

<210> SEQ ID NO 5
<211> LENGTH: 427
<212> TYPE: DNA
<213> ORGANISM: *Rattus norvegicus*

-continued

<400> SEQUENCE: 5

gaaaatcagat	cgcaacttatt	caaaggagcc	aggcccgtct	ctgcgcctg	gtggaggctc	60
ctcatgtgaa	gagtacaaa	aggcaccatg	tttgtggatac	ggggcgaagc	ccctccggtg	120
tgtcctccag	gcatcatcag	gaactagtgc	caaagcagag	gtgctggcca	gggccttggg	180
agtgacgcgc	gtctgggagg	cttgtgcgcgc	cagggcacgc	ccctgccat	tcccactagc	240
aggcttggg	ggacctgggc	cggtctgcgc	cctcctccag	caatcggtct	ataaaagctct	300
tccaagtcag	ggcgcagaag	tgccgggcga	tccgggctta	aagagcgaga	ggaaggacg	360
ctcacctttg	agtcctcca	caaatacgcc	tggtggtcgc	cacagaagtt	cgaagtttag	420
agttcg						427

<210> SEQ ID NO 6

<211> LENGTH: 722

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 6

gacgtcacag	tgggtcagtc	acccttgcata	acactgcacc	agtcttcacc	tttccacgct	60
tcctgccaga	gcatgaatca	ggctctctgg	ggataccggc	ctcaccctta	ctgaggcaaa	120
ctttctccca	cttctcagag	gctctgaggg	cagcaaggtc	agccctttct	ttggaatcta	180
gaaccactcc	ctgtcttgcag	ctgacatcac	agggcaggca	gatgcagcag	ggaaggcct	240
gggactggga	cgttcatcct	acaagaaago	tgtggaaact	ttcagcaaca	tctcagaaat	300
cagatcgcac	ttattcaaag	gagccaggcc	ctgctctgcg	ccctggtgga	ggctcctcat	360
gtgaagagt	acaaaaggca	ccatgttgc	gatacggggc	gaagccccctc	cggtgtgtcc	420
tccaggcatc	atcaggaact	agtgcacaaag	cagaggtgc	ggccagggtct	ttgggagtga	480
cgcgcgtctg	ggaggcttgc	gcccgcagg	cacgccccctg	ccgattccca	ctagcaggct	540
ttgggggacc	tggccggct	ctgccccctc	tccagcaatc	gggctataaa	gctttccaa	600
gtcaggcgc	agaagtgcgc	ggcgtatccgg	gcttaaagag	cgagaggaag	ggacgctcac	660
ctttgagetc	ctccacaaat	agccctggtg	gctgccacag	aagttcgaag	ttgagagttc	720
gg						722

<210> SEQ ID NO 7

<211> LENGTH: 334

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

tgatcaagtg	catttgttaa	tgtgttctac	atttcaaaa	aggaaaggag	aatttgttac	60
attcagaact	tgctgcact	cctttgtctac	gtcataaaagg	gtcagttgcc	cttgcata	120
ctgacattt	ctttacctct	ctgcttctc	tttgtgccag	aagagtagaa	atctgaccc	180
ttggggatac	caccctctcc	cctactgctc	tctccaaacct	gaggcaaact	ttctcctact	240
tcccagagcc	tgtcagaagt	ggtgaagcc	gcctgctcct	tggaatccag	aactactttc	300
agaatcttga	acttctgtga	cctctcagg	tccc			334

<210> SEQ ID NO 8

<211> LENGTH: 289

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 8

accggcgccgg	tgcgcctcc	ctccgacgtg	cggtgtcgcg	ggcgccagaca	accagcgcc	60
ggcccaggc	tttcggggag	cgaagcagg	ctcccgagge	accgagcgag	aatggaaatg	120
ggagggaccc	ggtgcetccc	gacacgcccc	cggcagggtcc	cacgccccgg	tcttctgaga	180
cctcgcgccg	cccagcccg	gagcggccca	gctatataag	tcccagcgga	agaccggAAC	240
gcagagggtc	ctgctggcgc	gagggtgggt	aggaggggac	gcggggact		289

<210> SEQ ID NO 9

<211> LENGTH: 623

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

tgatcaagt	catttgtta	tgtgttctac	atttcaaaa	aggaaaggag	aatttgttac	60
attcagaact	tgctgcact	cctttgttac	gtcataaaagg	gtcagttgc	cttgcata	120
ctgacctatt	ctttacctct	ctgcttctt	tttgcacag	aagagtagaa	atctgaccct	180
ttggggatac	caccctctcc	cctactgtct	tctccaaacct	gaggcaact	ttctctact	240
tcocagagcc	tgtcagaagt	ggtgaagcca	gcctgtct	tggaaatccag	aactacttcc	300
agaatcttga	acttctgtga	cctctcagg	tcccacccgc	gggggtgcgc	ctccctccga	360
cgtgcggtgt	gcggggcgca	gacaaccgc	ggccggccca	gggtttcgg	ggagegaagc	420
agggctcccg	aggcaccgc	cgagaatggg	aatggggagg	accgggtgt	cccgacacgc	480
ccccccggcag	gtccccacgc	cgggtcttct	gagacctcgc	ggggcccagc	ccgggagcgg	540
cccgactata	taagtcccg	cggaaagaccc	gaacgcagag	ggtcctgtc	gcgcgagggt	600
gggttaggagg	ggacgcgggg	act				623

<210> SEQ ID NO 10

<211> LENGTH: 888

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10

atgctatTTA	tgaagatgg	cctgttgaac	tatcagtact	tggacaagat	gaacaacaat	60
atcggcatTC	tgtgtcacGA	aggegaagCT	gctctcaggG	gagaacccAG	aatacagacc	120
ctgccggTGG	cctctgcCT	cagcagtac	cgcacccggC	ctcccccaat	cagccccagC	180
aagaggaagt	tcaGcatgg	gccaggtgac	gaggacctag	actgtgacAA	cgaccacgtc	240
tccaaaatga	gtcgcatctt	caaccccat	ctgaacaaga	ctgccaatgg	agactgcgc	300
agagaccccc	gggagcggag	ccgcacggcc	atcgagcgc	ctgtggccccc	caccatgagc	360
ctgcacggca	gcacacgttA	cacccctctc	cccagccttG	gcctggagca	gcccctcgca	420
ctgaccaaga	acagcctgg	cgccagcagg	ccagccggcc	tctcgcccac	actgaccccg	480
ggggagcggc	agcagaaccg	gccctccgt	atcacctgt	cctcggttg	cgcccgcaac	540
tgcacacTCT	cgcaactgccc	catcgccac	agggcgtgt	cgcgcacccgg	gcctgcacgc	600
taccggaggc	caccgagcgc	tgccaccacc	tgtgaccccg	tggtgagga	gcatttccgc	660
aggaggctgg	gcaagaatta	caaggagccc	gagccggcac	ccaactccgt	gtccatcacg	720
ggctccgtgg	acgaccactt	tgccaaagct	ctgggtgaca	cgtggctcca	gatcaaagcg	780
gccaaggacg	gagcatccag	cagccctgag	tccgcctctc	gcagggggcc	gcggccacg	840

-continued

ccctctgccc acatggtcag ccacagtcac tccccctctg tggtctcc 888

<210> SEQ ID NO 11
<211> LENGTH: 296
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

Met	Leu	Phe	Met	Lys	Met	Asp	Leu	Leu	Asn	Tyr	Gln	Tyr	Leu	Asp	Lys
1					5				10			15			

Met	Asn	Asn	Asn	Ile	Gly	Ile	Leu	Cys	Tyr	Glu	Gly	Glu	Ala	Ala	Leu
				20				25			30				

Arg	Gly	Glu	Pro	Arg	Ile	Gln	Thr	Leu	Pro	Val	Ala	Ser	Ala	Leu	Ser
				35				40			45				

Ser	His	Arg	Thr	Gly	Pro	Pro	Ile	Ser	Pro	Ser	Lys	Arg	Lys	Phe
				50				55			60			

Ser	Met	Glu	Pro	Gly	Asp	Glu	Asp	Leu	Asp	Cys	Asp	Asn	Asp	His	Val
65				70				75			80				

Ser	Lys	Met	Ser	Arg	Ile	Phe	Asn	Pro	His	Leu	Asn	Lys	Thr	Ala	Asn
				85				90			95				

Gly	Asp	Cys	Arg	Arg	Asp	Pro	Arg	Glu	Arg	Ser	Arg	Ser	Pro	Ile	Glu
				100				105			110				

Arg	Ala	Val	Ala	Pro	Thr	Met	Ser	Leu	His	Gly	Ser	His	Leu	Tyr	Thr
				115				120			125				

Ser	Leu	Pro	Ser	Leu	Gly	Leu	Glu	Gln	Pro	Leu	Ala	Leu	Thr	Lys	Asn
				130				135			140				

Ser	Leu	Asp	Ala	Ser	Arg	Pro	Ala	Gly	Leu	Ser	Pro	Thr	Leu	Thr	Pro
145				150				155			160				

Gly	Glu	Arg	Gln	Gln	Asn	Arg	Pro	Ser	Val	Ile	Thr	Cys	Ala	Ser	Ala
				165				170			175				

Gly	Ala	Arg	Asn	Cys	Asn	Leu	Ser	His	Cys	Pro	Ile	Ala	His	Ser	Gly
				180				185			190				

Cys	Ala	Ala	Pro	Gly	Pro	Ala	Ser	Tyr	Arg	Arg	Pro	Pro	Ser	Ala	Ala
				195				200			205				

Thr	Thr	Cys	Asp	Pro	Val	Val	Glu	Glu	His	Phe	Arg	Arg	Ser	Leu	Gly
				210				215			220				

Lys	Asn	Tyr	Lys	Glu	Pro	Glu	Pro	Ala	Pro	Asn	Ser	Val	Ser	Ile	Thr
225				230				235			240				

Gly	Ser	Val	Asp	Asp	His	Phe	Ala	Lys	Ala	Leu	Gly	Asp	Thr	Trp	Leu
				245				250			255				

Gln	Ile	Lys	Ala	Ala	Lys	Asp	Gly	Ala	Ser	Ser	Ser	Pro	Glu	Ser	Ala
				260				265			270				

Ser	Arg	Arg	Gly	Gln	Pro	Ala	Ser	Pro	Ser	Ala	His	Met	Val	Ser	His
				275				280			285				

Ser	His	Ser	Pro	Ser	Val	Val	Ser								
				290				295							

<210> SEQ ID NO 12

<211> LENGTH: 870

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

atggagacgc	cattggatgt	tttgtccagg	gcagcatctc	tggtgcatgc	tgtatgcacaa	60
------------	------------	------------	------------	------------	-------------	----

aaacgcgaag	ctgctctcag	gggagaaccc	agaatacaga	ccctgccgtt	ggccctctgcc	120
------------	------------	------------	------------	------------	-------------	-----

US 12,384,823 B2

45**46**

-continued

ctcagcagtc accgcacccgg ccctccccca atcagccccca gcaagaggaa gttcagcatg	180
gagccagggtg acgaggacct agactgtgac aacgaccacg tctccaaaat gagtcgcatac	240
ttcaacccccc atctgaacaa gactgccaat ggagactgcc gcagagaccc cggggagcgg	300
agccgcagcc ccatcgagcg cgctgtggcc cccaccatga gcctgcacgg cagccacctg	360
tacacccccc tccccagcct tggctggag cagccctcg cactgaccaa gaacagctg	420
gacgccagca ggccagccgg cctctcgccc acactgaccc cggggagcgg gcagcagaac	480
cggccctccg tgatcacctg tgctcggtt ggcccccggca actgcaacct ctgcactgc	540
cccatcgcc acagcggctg tgccgcgccc gggcctgcca gctacggag gccaccgagc	600
gctgccacca cctgtgaccc cgtggtaggag gacattcc gcaggagcct gggcaagaat	660
tacaaggagc ccgagccggc acccaactcc gtgtccatca cgggctccgt ggacgaccac	720
tttgccaaag ctctgggtga cacgtggctc cagatcaaag cggccaagga cggagcatcc	780
agcagccctg agtccgcctc tcgcaggggc cagcccgcca gcccctctgc ccacatggc	840
agccacagtc actccccctc tgtggtctcc	870

<210> SEQ ID NO 13

<211> LENGTH: 290

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

Met Glu Thr Pro Leu Asp Val Leu Ser Arg Ala Ala Ser Leu Val His			
1	5	10	15

Ala Asp Asp Glu Lys Arg Glu Ala Ala Leu Arg Gly Glu Pro Arg Ile			
20	25	30	

Gln Thr Leu Pro Val Ala Ser Ala Leu Ser Ser His Arg Thr Gly Pro			
35	40	45	

Pro Pro Ile Ser Pro Ser Lys Arg Lys Phe Ser Met Glu Pro Gly Asp			
50	55	60	

Glu Asp Leu Asp Cys Asp Asn Asp His Val Ser Lys Met Ser Arg Ile			
65	70	75	80

Phe Asn Pro His Leu Asn Lys Thr Ala Asn Gly Asp Cys Arg Arg Asp			
85	90	95	

Pro Arg Glu Arg Ser Arg Ser Pro Ile Glu Arg Ala Val Ala Pro Thr			
100	105	110	

Met Ser Leu His Gly Ser His Leu Tyr Thr Ser Leu Pro Ser Leu Gly			
115	120	125	

Leu Glu Gln Pro Leu Ala Leu Thr Lys Asn Ser Leu Asp Ala Ser Arg			
130	135	140	

Pro Ala Gly Leu Ser Pro Thr Leu Thr Pro Gly Glu Arg Gln Gln Asn			
145	150	155	160

Arg Pro Ser Val Ile Thr Cys Ala Ser Ala Gly Ala Arg Asn Cys Asn			
165	170	175	

Leu Ser His Cys Pro Ile Ala His Ser Gly Cys Ala Ala Pro Gly Pro			
180	185	190	

Ala Ser Tyr Arg Arg Pro Pro Ser Ala Ala Thr Thr Cys Asp Pro Val			
195	200	205	

Val Glu Glu His Phe Arg Arg Ser Leu Gly Lys Asn Tyr Lys Glu Pro			
210	215	220	

Glu Pro Ala Pro Asn Ser Val Ser Ile Thr Gly Ser Val Asp Asp His			
225	230	235	240

US 12,384,823 B2

47**48**

-continued

Phe Ala Lys Ala Leu Gly Asp Thr Trp	Leu Gln Ile Lys Ala Ala Lys	
245	250	255

Asp Gly Ala Ser Ser Ser Pro Glu Ser Ala Ser Arg Arg	Gly Gln Pro	
260	265	270

Ala Ser Pro Ser Ala His Met Val Ser His Ser His Ser Pro Ser Val		
275	280	285

Val Ser
290

<210> SEQ ID NO 14

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

atgattaaag tgaggaacaa gactgccaat ggagactgcc gcagagaccc ccgggagcgg	60
agccgcagcc ccatcgagcg cgctgtggcc cccaccatga gcctgcacgg cagccacctg	120
tacacacctcc tccccagcct tggcctggag cagcccctcg cactgaccaa gaacagcctg	180
gacgccagca ggccagccgg cctetcgccc acactgaccc cgggggagcg gcagcagaac	240
cggccctccg tgatcacctg tgcctcggt ggcgcccgcg actgcaacct ctgcactgc	300
cccatcgccg acagcggctg tgccgcgccc gggcctgcca gctaccggag gccaccgagc	360
gctgccacca cctgtgaccc cgtggtggag gaggcattcc gcaggagcct gggcaagaat	420
tacaaggagc ccgagccggc acccaactcc gtgtccatca cgggctccgt ggacgacac	480
tttgccaaag ctctgggtga cacgtggctc cagatcaaag cggccaagga cggagcatcc	540
agcagccctg agtccgcctc tcgcaggggc cagccgcgc gcccctctgc ccacatggc	600
agccacagtc actccccctc tgtggtctcc	630

<210> SEQ ID NO 15

<211> LENGTH: 210

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 15

Met Ile Lys Val Arg Asn Lys Thr Ala Asn Gly Asp Cys Arg Arg Asp			
1	5	10	15

Pro Arg Glu Arg Ser Arg Ser Pro Ile Glu Arg Ala Val Ala Pro Thr		
20	25	30

Met Ser Leu His Gly Ser His Leu Tyr Thr Ser Leu Pro Ser Leu Gly		
35	40	45

Leu Glu Gln Pro Leu Ala Leu Thr Lys Asn Ser Leu Asp Ala Ser Arg		
50	55	60

Pro Ala Gly Leu Ser Pro Thr Leu Thr Pro Gly Glu Arg Gln Gln Asn			
65	70	75	80

Arg Pro Ser Val Ile Thr Cys Ala Ser Ala Gly Ala Arg Asn Cys Asn		
85	90	95

Leu Ser His Cys Pro Ile Ala His Ser Gly Cys Ala Ala Pro Gly Pro		
100	105	110

Ala Ser Tyr Arg Arg Pro Pro Ser Ala Ala Thr Thr Cys Asp Pro Val		
115	120	125

Val Glu Glu His Phe Arg Arg Ser Leu Gly Lys Asn Tyr Lys Glu Pro		
130	135	140

Glu Pro Ala Pro Asn Ser Val Ser Ile Thr Gly Ser Val Asp Asp His			
145	150	155	160

-continued

Phe Ala Lys Ala Leu Gly Asp Thr Trp Leu Gln Ile Lys Ala Ala Lys
 165 170 175
 Asp Gly Ala Ser Ser Ser Pro Glu Ser Ala Ser Arg Arg Gly Gln Pro
 180 185 190
 Ala Ser Pro Ser Ala His Met Val Ser His Ser His Ser Pro Ser Val
 195 200 205
 Val Ser
 210

<210> SEQ ID NO 16
<211> LENGTH: 618
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

atgaacaaga	ctgccaatgg	agactgccgc	agagacccccc	gggagcggag	ccgcagcccc	60
atcgagcgcg	ctgtggcccc	caccatgago	ctgcacggca	gccacctgta	caccccttc	120
cccagccttg	gcctggagca	gcccotcgca	ctgaccaaga	acagcctgga	cgccagcagg	180
ccagccggcc	tctcgcccac	actgaccccg	ggggagcggc	agcagaacccg	gccctccgt	240
atcaccttgt	cctcggtctgg	cgcggcgaac	tgcaacctct	cgcactgccc	cacgcgcac	300
agcggctgt	ccgcgccccgg	gcctgcccgc	taccggaggc	caccgagcgc	tgccaccacc	360
tgtgaccccg	tggtggagga	gcatttccgc	aggagcctgg	gcaagaattta	caaggagccc	420
gagccggcac	ccaactccgt	gtccatcaeg	ggctccgtgg	acgaccactt	tgccaaagct	480
ctgggtgaca	cgtggctcca	gatcaaageg	gccaaggacg	gagcatccag	cagccctgag	540
tccgcctctc	gcagggggca	gccccccgc	ccctctgccc	acatggtcag	ccacagtac	600
tccccctctg	tggtctcc					618

<210> SEQ ID NO 17
<211> LENGTH: 206
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

Met Asn Lys Thr Ala Asn Gly Asp Cys Arg Arg Asp Pro Arg Glu Arg			
1	5	10	15
Ser Arg Ser Pro Ile Glu Arg Ala Val Ala Pro Thr Met Ser Leu His			
20	25	30	
Gly Ser His Leu Tyr Thr Ser Leu Pro Ser Leu Gly Leu Glu Gln Pro			
35	40	45	
Leu Ala Leu Thr Lys Asn Ser Leu Asp Ala Ser Arg Pro Ala Gly Leu			
50	55	60	
Ser Pro Thr Leu Thr Pro Gly Glu Arg Gln Gln Asn Arg Pro Ser Val			
65	70	75	80
Ile Thr Cys Ala Ser Ala Gly Ala Arg Asn Cys Asn Leu Ser His Cys			
85	90	95	
Pro Ile Ala His Ser Gly Cys Ala Ala Pro Gly Pro Ala Ser Tyr Arg			
100	105	110	
Arg Pro Pro Ser Ala Ala Thr Thr Cys Asp Pro Val Val Glu Glu His			
115	120	125	
Phe Arg Arg Ser Leu Gly Lys Asn Tyr Lys Glu Pro Glu Pro Ala Pro			
130	135	140	
Asn Ser Val Ser Ile Thr Gly Ser Val Asp Asp His Phe Ala Lys Ala			

US 12,384,823 B2

51**52**

-continued

145	150	155	160
Leu Gly Asp Thr Trp	Leu Gln Ile Lys Ala Ala Lys Asp Gly Ala Ser		
165	170	175	
Ser Ser Pro Glu Ser Ala Ser Arg Arg Gly Gln Pro Ala Ser Pro Ser			
180	185	190	
Ala His Met Val Ser His Ser His Ser Pro Ser Val Val Ser			
195	200	205	

<210> SEQ ID NO 18

<211> LENGTH: 885

<212> TYPE: DNA

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 18

atgactgaga atacgcattt tgacaaaatc cctgagtcgt gtgcactcaa aagttggaga	60
catccaggta tgcaccatgg cgaagctgct ctcaggggag aacccagaat acagaccctg	120
ccggtgtggct ctgccctcag cagtaccgc accggccctc ccccaatcag ccccgcaag	180
aggaagttca gcatggagcc aggtgacgag gaccttagact gtgacaacga ccacgtctcc	240
aaaaatgatgc gcatcttcaa cccccatctg aacaagactg ccaatggaga ctggccgaga	300
gaccggccgg agcggagccg cagccccato gagcgcgttg tggcccccac catgagctg	360
cacggcagcc acctgtacac ctccctcccc agecttggcc tggagcagcc cctcgactg	420
accaagaaca gcctggacgc cagcaggcca gcggccctc cgccccacact gacccgggg	480
gagcggcagc agaaccggcc ctccgtgatc acctgtgcgtt cggctggcgc ccgcaactgc	540
aacctctcgc actgccccat cgccacago ggetgtgccc cgccggggcc tgccagctac	600
cggaggccac cgagcgctgc caccacctgt gacccctgg tggaggagca ttccgcagg	660
agcctgggca agaattacaa ggagcccgag ccggcaccca actccgtgtc catcacggc	720
tccgtggacg accactttgc caaagctctg ggtgacacgt ggctccagat caaagcggcc	780
aaggacggag catccagcag ccctgagtcg gcctctcgca ggggccagcc cgccagcccc	840
tctgcccaca tggtcagcca cagtcaactcc ccctctgtgg tctcc	885

<210> SEQ ID NO 19

<211> LENGTH: 295

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 19

Met Thr Glu Asn Thr His Phe Asp Lys Ile Pro Glu Ser Cys Ala Leu			
1	5	10	15
Lys Ser Trp Arg His Pro Gly Leu His His Gly Glu Ala Ala Leu Arg			
20	25	30	
Gly Glu Pro Arg Ile Gln Thr Leu Pro Val Ala Ser Ala Leu Ser Ser			
35	40	45	
His Arg Thr Gly Pro Pro Pro Ile Ser Pro Ser Lys Arg Lys Phe Ser			
50	55	60	
Met Glu Pro Gly Asp Glu Asp Leu Asp Cys Asp Asn Asp His Val Ser			
65	70	75	80
Lys Met Ser Arg Ile Phe Asn Pro His Leu Asn Lys Thr Ala Asn Gly			
85	90	95	
Asp Cys Arg Arg Asp Pro Arg Glu Arg Ser Arg Ser Pro Ile Glu Arg			
100	105	110	
Ala Val Ala Pro Thr Met Ser Leu His Gly Ser His Leu Tyr Thr Ser			

US 12,384,823 B2

53**54**

-continued

115	120	125
Leu Pro Ser Leu Gly Leu Glu Gln Pro Leu Ala Leu Thr Lys Asn Ser		
130	135	140
Leu Asp Ala Ser Arg Pro Ala Gly Leu Ser Pro Thr Leu Thr Pro Gly		
145	150	155 160
Glu Arg Gln Gln Asn Arg Pro Ser Val Ile Thr Cys Ala Ser Ala Gly		
165	170	175
Ala Arg Asn Cys Asn Leu Ser His Cys Pro Ile Ala His Ser Gly Cys		
180	185	190
Ala Ala Pro Gly Pro Ala Ser Tyr Arg Arg Pro Pro Ser Ala Ala Thr		
195	200	205
Thr Cys Asp Pro Val Val Glu Glu His Phe Arg Arg Ser Leu Gly Lys		
210	215	220
Asn Tyr Lys Glu Pro Glu Pro Ala Pro Asn Ser Val Ser Ile Thr Gly		
225	230	235 240
Ser Val Asp Asp His Phe Ala Lys Ala Leu Gly Asp Thr Trp Leu Gln		
245	250	255
Ile Lys Ala Ala Lys Asp Gly Ala Ser Ser Ser Pro Glu Ser Ala Ser		
260	265	270
Arg Arg Gly Gln Pro Ala Ser Pro Ser Ala His Met Val Ser His Ser		
275	280	285
His Ser Pro Ser Val Val Ser		
290	295	

<210> SEQ ID NO 20
<211> LENGTH: 693
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

```

atggagccag gtgacgagga cctagactgt gacaacgacc acgtctccaa aatgagtgc 60
atcttcaacc cccatctgaa caagactgcc aatggagact gccgcagaga cccccggag 120
cggagccgca gccccatcgca gcgcgcgtgtg gcccccacca tgagcctgca cggcagccac 180
ctgtacacct ccctccccag cttggcctg gagcagcccc tcgcactgac caagaacagc 240
ctggacgcca gcaggccgc cggcctctcg cccacactga cccccgggga gcccgcagcag 300
aacccgcct ccgtgatcac ctgtgcctcg gctggcgecc gcaactgcaa cctctcgac 360
tgccccatcg cgcacacgccc ctgtgcccgcc cccgggcctg ccagctacccg gaggccaccg 420
agcgctgcca ccacctgtga ccccggtgtg gaggagcatt tccgcaggag cctggcaag 480
aattacaagg agcccgagcc ggcacccaaac tccgtgtcca tcacgggctc cgtggacgac 540
cactttgcca aagctctggg tgacacgtgg ctccagatca aagcggccaa ggacggagca 600
tccagcagcc ctgagtcgc ctcctcgccagg ggccagcccc ccagccccctc tgcccacatg 660
gtcagccaca gtcactcccc ctctgtggtc tcc 693

```

<210> SEQ ID NO 21
<211> LENGTH: 231
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21

Met Glu Pro Gly Asp Glu Asp Leu Asp Cys Asp Asn Asp His Val Ser	5	10 15
---	---	-------

Lys Met Ser Arg Ile Phe Asn Pro His Leu Asn Lys Thr Ala Asn Gly

US 12,384,823 B2

55**56**

-continued

20 25 30

Asp Cys Arg Arg Asp Pro Arg Glu Arg Ser Arg Ser Pro Ile Glu Arg		
35	40	45

Ala Val Ala Pro Thr Met Ser Leu His Gly Ser His Leu Tyr Thr Ser		
50	55	60

Leu Pro Ser Leu Gly Leu Glu Gln Pro Leu Ala Leu Thr Lys Asn Ser		
65	70	75

Leu Asp Ala Ser Arg Pro Ala Gly Leu Ser Pro Thr Leu Thr Pro Gly		
85	90	95

Glu Arg Gln Gln Asn Arg Pro Ser Val Ile Thr Cys Ala Ser Ala Gly		
100	105	110

Ala Arg Asn Cys Asn Leu Ser His Cys Pro Ile Ala His Ser Gly Cys		
115	120	125

Ala Ala Pro Gly Pro Ala Ser Tyr Arg Arg Pro Pro Ser Ala Ala Thr		
130	135	140

Thr Cys Asp Pro Val Val Glu Glu His Phe Arg Arg Ser Leu Gly Lys		
145	150	155

Asn Tyr Lys Glu Pro Glu Pro Ala Pro Asn Ser Val Ser Ile Thr Gly		
165	170	175

Ser Val Asp Asp His Phe Ala Lys Ala Leu Gly Asp Thr Trp Leu Gln		
180	185	190

Ile Lys Ala Ala Lys Asp Gly Ala Ser Ser Ser Pro Glu Ser Ala Ser		
195	200	205

Arg Arg Gly Gln Pro Ala Ser Pro Ser Ala His Met Val Ser His Ser		
210	215	220

His Ser Pro Ser Val Val Ser		
225	230	

<210> SEQ ID NO 22

<211> LENGTH: 888

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 22

atgctattta tgaagatggaa cctgttgaac tatcagtact tggacaagat gaacaacaat	60
atcggcattc tgtgtcacga aggccaagct gctctcaggg gagaaccagg aatacagacc	120
ctgcgggtgg cctctgcctt cagcagtcac cgccacggcc ctcccccaat cagccccagc	180
aagaggaagt tcagcatgga gccaggtgac gaggacctag actgtgacaa cgaccacgtc	240
tccaaaatga gtgcgtatctt caaccccat ctgaacaaga ctgccaatgg agactgcccgc	300
agagacccccc gggagcggag ccgcgcgcgc atcgagcgcg ctgtggccccc caccatgagc	360
ctgcacggca gccacctgta cacccctctt cccagccttg gcctggagca gcccctcgca	420
ctgaccaaga acagcctgga cgccagcagg ccagccggcc tctcgcccac actgaccccg	480
ggggagcggc agcagaaccg gccctccgtg atcacctgtg cctcggtgg cgcccgcaac	540
tgcacacccctt cgcactgccc catcgccac agcggtgtg cgcgcggccgg gcctggccagc	600
taccggaggc caccgagcgc tgccaccacc tggaccccg tggtgagga ggcagccgc	660
aggaggctgg gcaagaatta caaggagccc gagccggcac ccaactccgt gtccatcacg	720
ggctccgtgg acgacgcagc tgccaaagct ctgggtgaca cgtggctcca gatcaaagcg	780
gccaaggacg gagcatccag cagccctgag tccgcctctc gcagggggcca gcccggccagc	840

US 12,384,823 B2

57

-continued

ccctctgcc acatggtcag ccacagtcac tccccctctg tggctc

888

<210> SEQ ID NO 23
<211> LENGTH: 296
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 23

Met	Leu	Phe	Met	Lys	Met	Asp	Leu	Leu	Asn	Tyr	Gln	Tyr	Leu	Asp	Lys
1					5				10				15		
Met	Asn	Asn	Asn	Ile	Gly	Ile	Leu	Cys	Tyr	Glu	Gly	Glu	Ala	Ala	Leu
	20					25				30					
Arg	Gly	Glu	Pro	Arg	Ile	Gln	Thr	Leu	Pro	Val	Ala	Ser	Ala	Leu	Ser
	35						40				45				
Ser	His	Arg	Thr	Gly	Pro	Pro	Ile	Ser	Pro	Ser	Lys	Arg	Lys	Phe	
	50					55				60					
Ser	Met	Glu	Pro	Gly	Asp	Glu	Asp	Leu	Asp	Cys	Asp	Asn	Asp	His	Val
	65					70				75			80		
Ser	Lys	Met	Ser	Arg	Ile	Phe	Asn	Pro	His	Leu	Asn	Lys	Thr	Ala	Asn
	85						90				95				
Gly	Asp	Cys	Arg	Arg	Asp	Pro	Arg	Glu	Arg	Ser	Arg	Ser	Pro	Ile	Glu
	100					105				110					
Arg	Ala	Val	Ala	Pro	Thr	Met	Ser	Leu	His	Gly	Ser	His	Leu	Tyr	Thr
	115					120				125					
Ser	Leu	Pro	Ser	Leu	Gly	Leu	Glu	Gln	Pro	Leu	Ala	Leu	Thr	Lys	Asn
	130					135				140					
Ser	Leu	Asp	Ala	Ser	Arg	Pro	Ala	Gly	Leu	Ser	Pro	Thr	Leu	Thr	Pro
	145					150				155			160		
Gly	Glu	Arg	Gln	Gln	Asn	Arg	Pro	Ser	Val	Ile	Thr	Cys	Ala	Ser	Ala
	165					170				175					
Gly	Ala	Arg	Asn	Cys	Asn	Leu	Ser	His	Cys	Pro	Ile	Ala	His	Ser	Gly
	180						185				190				
Cys	Ala	Ala	Pro	Gly	Pro	Ala	Ser	Tyr	Arg	Arg	Pro	Pro	Ser	Ala	Ala
	195					200				205					
Thr	Thr	Cys	Asp	Pro	Val	Val	Glu	Ala	Ala	Arg	Arg	Ser	Leu	Gly	
	210					215				220					
Lys	Asn	Tyr	Lys	Glu	Pro	Glu	Pro	Ala	Pro	Asn	Ser	Val	Ser	Ile	Thr
	225					230				235			240		
Gly	Ser	Val	Asp	Asp	Ala	Ala	Lys	Ala	Leu	Gly	Asp	Thr	Trp	Leu	
	245					250				255					
Gln	Ile	Lys	Ala	Ala	Lys	Asp	Gly	Ala	Ser	Ser	Ser	Pro	Glu	Ser	Ala
	260					265				270					
Ser	Arg	Arg	Gly	Gln	Pro	Ala	Ser	Pro	Ser	Ala	His	Met	Val	Ser	His
	275					280				285					
Ser	His	Ser	Pro	Ser	Val	Val	Ser								
	290					295									

<210> SEQ ID NO 24
<211> LENGTH: 870
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 24

58

US 12,384,823 B2

59**60**

-continued

atggagacgc cattggatgt tttgtccagg gcagcatctc tggtgcatgc tgatgacgaa	60
aaacgcgaag ctgctctcg gggagaaccc agaatacaga ccctgccggt ggcctctgcc	120
ctcagcagtc accgcacccg ccctccccca atcagcccc acaagaggaa gttcagcatg	180
gagccaggta acgaggacct agactgtgac aacgaccacg tctccaaaat gagtgcacatc	240
ttcaacccccc atctgaacaa gactgccaat ggagactgccc gcagagaccc cccggagcgg	300
agccgcagcc ccatecgagcg cgctgtggcc cccaccatcg gcctgcacgg cagccacactg	360
tacacctccc tccccagcct tgccctggag cagccctcg cactgaccaa gaacagctg	420
gacgcccagca ggccagccgg cctctcgccc acactgaccc cggggagcgc gcagcagaac	480
cggccctccg tgatcacctg tgccctcggt ggcccccgc actgcaaccc ctgcactgc	540
cccatcgccg acagcggctg tgccgcgccc gggcctgcga gctaccggag gccaccgagc	600
gctgccacca cctgtgaccc cgtggtagag gaggcagccc gcaggagcct gggcaagaat	660
tacaaggagc ccgagccggc acccaactcc gtgtccatca cgggctccgt ggacgacgca	720
gctgccaag ctctgggtga cacgtggctc cagatcaaag cggccaagga cggagcatcc	780
agcagccctg agtccgcctc tcgcaggggc cagccgcga gcccctctgc ccacatggc	840
agccacagtc actccccctc tgtggctcc	870

<210> SEQ ID NO 25

<211> LENGTH: 290

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 25

Met Glu Thr Pro Leu Asp Val Leu Ser Arg Ala Ala Ser Leu Val His			
1	5	10	15

Ala Asp Asp Glu Lys Arg Glu Ala Ala Leu Arg Gly Glu Pro Arg Ile			
20	25	30	

Gln Thr Leu Pro Val Ala Ser Ala Leu Ser Ser His Arg Thr Gly Pro			
35	40	45	

Pro Pro Ile Ser Pro Ser Lys Arg Lys Phe Ser Met Glu Pro Gly Asp			
50	55	60	

Glu Asp Leu Asp Cys Asp Asn Asp His Val Ser Lys Met Ser Arg Ile			
65	70	75	80

Phe Asn Pro His Leu Asn Lys Thr Ala Asn Gly Asp Cys Arg Arg Asp			
85	90	95	

Pro Arg Glu Arg Ser Arg Ser Pro Ile Glu Arg Ala Val Ala Pro Thr			
100	105	110	

Met Ser Leu His Gly Ser His Leu Tyr Thr Ser Leu Pro Ser Leu Gly			
115	120	125	

Leu Glu Gln Pro Leu Ala Leu Thr Lys Asn Ser Leu Asp Ala Ser Arg			
130	135	140	

Pro Ala Gly Leu Ser Pro Thr Leu Thr Pro Gly Glu Arg Gln Gln Asn			
145	150	155	160

Arg Pro Ser Val Ile Thr Cys Ala Ser Ala Gly Ala Arg Asn Cys Asn			
165	170	175	

Leu Ser His Cys Pro Ile Ala His Ser Gly Cys Ala Ala Pro Gly Pro			
180	185	190	

Ala Ser Tyr Arg Arg Pro Pro Ser Ala Ala Thr Thr Cys Asp Pro Val			
195	200	205	

-continued

Val Glu Glu Ala Ala Arg Arg Ser Leu Gly Lys Asn Tyr Lys Glu Pro
 210 215 220

Glu Pro Ala Pro Asn Ser Val Ser Ile Thr Gly Ser Val Asp Asp Ala
 225 230 235 240

Ala Ala Lys Ala Leu Gly Asp Thr Trp Leu Gln Ile Lys Ala Ala Lys
 245 250 255

Asp Gly Ala Ser Ser Ser Pro Glu Ser Ala Ser Arg Arg Gly Gln Pro
 260 265 270

Ala Ser Pro Ser Ala His Met Val Ser His Ser His Ser Pro Ser Val
 275 280 285

Val Ser
 290

<210> SEQ ID NO 26

<211> LENGTH: 630

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 26

```

atgattaaag tgaggaacaa gactgc当地 ggagactgcc gcagagaccc ccgggagcgg      60
agccgc当地 ccatcgagcg cgctgtggcc cccaccatga gcctgc当地 cagccacctg      120
tacacctccc tccccagcct tggc当地 tggag cagccc当地 cactgaccaa gaacagcctg      180
gacgccc当地 gagccagccgg cctctcgccc acactgaccc cgggggagcg gcagcagaac      240
cggccctccg tgc当地 tggctggct ggccccc当地 actgcaaccc ctc当地 cactgc      300
cccatcgccg acagcggctg tgccgccc当地 gggcctgcca gctaccggag gccaccgagc      360
gctgccacca cctgtgaccc cgtggtggag gaggcagccc gcaggagcct gggcaagaat      420
tacaaggagc cc当地 gagccggc acccaactcc gt当地 tccatca cgggctccgt ggacgacgca      480
gctgccaag ctctgggtga cacgtggctc cagatcaaag cggccaaagga cggagcatcc      540
agcagccctg agtccgctc tc当地 cagggccca gccc当地 ctgc ccacatggc      600
agccacagtc actccccctc tgc当地 ggtctcc      630

```

<210> SEQ ID NO 27

<211> LENGTH: 210

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 27

Met Ile Lys Val Arg Asn Lys Thr Ala Asn Gly Asp Cys Arg Arg Asp
 1 5 10 15

Pro Arg Glu Arg Ser Arg Ser Pro Ile Glu Arg Ala Val Ala Pro Thr
 20 25 30

Met Ser Leu His Gly Ser His Leu Tyr Thr Ser Leu Pro Ser Leu Gly
 35 40 45

Leu Glu Gln Pro Leu Ala Leu Thr Lys Asn Ser Leu Asp Ala Ser Arg
 50 55 60

Pro Ala Gly Leu Ser Pro Thr Leu Thr Pro Gly Glu Arg Gln Gln Asn
 65 70 75 80

Arg Pro Ser Val Ile Thr Cys Ala Ser Ala Gly Ala Arg Asn Cys Asn
 85 90 95

Leu Ser His Cys Pro Ile Ala His Ser Gly Cys Ala Ala Pro Gly Pro

US 12,384,823 B2

63**64**

-continued

100 105 110

Ala Ser Tyr Arg Arg Pro Pro Ser Ala Ala Thr Thr Cys Asp Pro Val
115 120 125

Val Glu Glu Ala Ala Arg Arg Ser Leu Gly Lys Asn Tyr Lys Glu Pro
130 135 140

Glu Pro Ala Pro Asn Ser Val Ser Ile Thr Gly Ser Val Asp Asp Ala
145 150 155 160

Ala Ala Lys Ala Leu Gly Asp Thr Trp Leu Gln Ile Lys Ala Ala Lys
165 170 175

Asp Gly Ala Ser Ser Ser Pro Glu Ser Ala Ser Arg Arg Gly Gln Pro
180 185 190

Ala Ser Pro Ser Ala His Met Val Ser His Ser His Ser Pro Ser Val
195 200 205

Val Ser
210

<210> SEQ ID NO 28

<211> LENGTH: 618

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 28

atgaacaaga	ctgccaatgg	agactgccgc	agagacccccc	gggagcggag	ccgcagcccc	60
atcgagcg	ctgtggcccc	caccatgago	ctgcacggca	gccacctgta	caccccttc	120
cccagccttg	gcctggagca	gcacctcgca	ctgaccaaga	acagcctgga	cgccagcagg	180
ccagccggcc	tctcgcccac	actgaccccg	ggggagcggc	agcagaacccg	gccctccgtg	240
atcacctgtg	cctcggtctgg	cggccgcaac	tgcacccctct	cgcaactgccc	catacgcac	300
agcggctgtg	ccgcgccccgg	gcctgcccago	tacggagggc	caccgagcgc	tgccaccacc	360
tgtgaccccg	tgggtggagga	ggcagcccg	aggagcctgg	gcaagaatta	caaggagccc	420
gagccggcac	ccaactccgt	gtccatcaeg	ggctccgtt	acgacgcagc	tgccaaagct	480
ctgggtgaca	cgtggctcca	gatcaaageg	gccaaggacg	gagcatccag	cagccctgag	540
tccgcctctc	gcagggggcca	gcceggccago	ccctctgecc	acatggtcag	ccacagtac	600
tccccctctg	tggtctcc					618

<210> SEQ ID NO 29

<211> LENGTH: 206

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 29

Met Asn Lys	Thr Ala Asn Gly	Asp Cys Arg Arg	Asp Pro Arg	Glu Arg		
1	5	10	15			

Ser Arg Ser Pro Ile	Glu Arg Ala Val	Ala Pro Thr	Met Ser	Leu His		
20	25	30				

Gly Ser His	Leu Tyr Thr Ser	Leu Pro Ser	Leu Gly	Leu Glu	Gln Pro	
35	40	45				

Leu Ala	Leu Thr Lys Asn	Ser Leu Asp	Ala Ser Arg	Pro Ala Gly	Leu	
50	55	60				

Ser Pro Thr	Leu Thr Pro	Gly Glu Arg	Gln Gln Asn	Arg Pro	Ser Val	
65	70	75	80			

-continued

Ile Thr Cys Ala Ser Ala Gly Ala Arg Asn Cys Asn Leu Ser His Cys
 85 90 95

Pro Ile Ala His Ser Gly Cys Ala Ala Pro Gly Pro Ala Ser Tyr Arg
 100 105 110

Arg Pro Pro Ser Ala Ala Thr Thr Cys Asp Pro Val Val Glu Glu Ala
 115 120 125

Ala Arg Arg Ser Leu Gly Lys Asn Tyr Lys Glu Pro Glu Pro Ala Pro
 130 135 140

Asn Ser Val Ser Ile Thr Gly Ser Val Asp Asp Ala Ala Ala Lys Ala
 145 150 155 160

Leu Gly Asp Thr Trp Leu Gln Ile Lys Ala Ala Lys Asp Gly Ala Ser
 165 170 175

Ser Ser Pro Glu Ser Ala Ser Arg Arg Gly Gln Pro Ala Ser Pro Ser
 180 185 190

Ala His Met Val Ser His Ser His Ser Pro Ser Val Val Ser
 195 200 205

<210> SEQ ID NO 30
 <211> LENGTH: 885
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 30

atgactgaga atacgcattt tgacaaaatc cctgagtcgt gtgcactcaa aagttggaga 60
 catccaggta tgcaccatgg cgaagctgtc ctcaggggg aacccagaat acagaccctg 120
 ccgggtggcc ctggccctcg cagtccaccgc accggccctc ccccaatcag ccccgacaaag 180
 aggaagttca gcatggagcc aggtgacgag gaccttagact gtgacaacga ccacgtctcc 240
 aaaaatgatgc gcatcttcaa cccccatctg aacaagactg ccaatggaga ctggccgaga 300
 gaccccccggg agcggagccg cagcccatc gaggcgctg tggcccccac catgagctg 360
 cacggcagcc acctgtacac ctccctcccc agecttgccg tggagcagcc cctcgactg 420
 accaagaaca gcctggacgc cagcaggcca gcccgcctc cgcccacact gaccccgggg 480
 gagccggcagc agaaccggcc ctccgtgatc acctgtgect cggctggcgc ccgcaactgc 540
 aacctctcgc actgccccat cgccacago ggctgtgecg cgcccgccg tgccagctac 600
 cggaggccac cgagcgctgc caccacctgt gacccctgtt tggaggaggg agcccgagg 660
 agcctgggca agaattacaa ggagcccgag ccggcaccca actccgtgtc catcacggc 720
 tccgtggacg acgcagctgc caaaagctctg ggtgacacgt ggctccagat caaagcggcc 780
 aaggacggag catcccgacag ccctgagtc gcctctcgca gggccagcc cgccagcccc 840
 tctgcccaca tggtcagcca cagtcaactcc ccctctgtgg tctcc 885

<210> SEQ ID NO 31
 <211> LENGTH: 295
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 31

Met Thr Glu Asn Thr His Phe Asp Lys Ile Pro Glu Ser Cys Ala Leu
 1 5 10 15

Lys Ser Trp Arg His Pro Gly Leu His His Gly Glu Ala Ala Leu Arg

US 12,384,823 B2

67

-continued

68

20	25	30
Gly	Glu	Pro
35	40	45
His	Arg	Thr
50	55	60
Met	Glu	Pro
65	70	75
Lys	Met	Ser
85	90	95
Asp	Cys	Arg
100	105	110
Ala	Val	Ala
115	120	125
Leu	Pro	Ser
130	135	140
Leu	Asp	Ala
145	150	155
Glu	Arg	Gln
165	170	175
Ala	Arg	Asn
180	185	190
Ala	Ala	Pro
195	200	205
Thr	Cys	Asp
210	215	220
Asn	Tyr	Lys
225	230	235
Ser	Val	Asp
245	250	255
Ile	Lys	Ala
260	265	270
Arg	Arg	Gly
275	280	285
His	Ser	Pro
290	295	

<210> SEQ ID NO 32
 <211> LENGTH: 693
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 32

atggagccag	gtgacgagga	cctagactgt	gacaacgacc	acgtctccaa	aatgagtcgc	60
atcttcaacc	cccatctgaa	caagactgcc	aatggagact	gccgcagaga	ccccgggag	120
cggagccgca	gccccatcga	gcgcgctgtg	gccccacca	tgagcctgca	cgccagccac	180
ctgtacacct	ccctccccag	ccttggcctg	gaggcccc	tgcactgac	caagaacagc	240
ctggacgcca	gcaggccagc	cggcctctcg	cccacactga	ccccggggga	gcggcagcag	300
aaccggccct	ccgtgatcac	ctgtgcctcg	gctggcgccc	gcaactgcaa	cctctcgcac	360
tgc(ccatcg	cgcacagcgg	ctgtgcccgg	ccggggcctg	ccagctaccg	gaggccacccg	420
agcgctgcca	ccacctgtga	ccccgtggtg	gaggaggcag	ccgcaggag	cctggcaag	480

-continued

aattacaagg	agcccgagcc	ggcacccaac	tccgtgtcca	tcacgggctc	cgtggacgac	540
gcagctgcc	aagctctggg	tgacacgtgg	ctccagatca	aagcggccaa	ggacggagca	600
tccagcagcc	ctgagtccgc	ctctcgcagg	ggccagcccc	ccagccccctc	tgcccacatg	660
gtcagccaca	gtcactcccc	ctctgtggtc	tcc			693

<210> SEQ ID NO 33
<211> LENGTH: 231
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 33

Met	Glu	Pro	Gly	Asp	Glu	Asp	Leu	Asp	Cys	Asp	Asn	Asp	His	Val	Ser
1	5			10					15						

Lys	Met	Ser	Arg	Ile	Phe	Asn	Pro	His	Leu	Asn	Lys	Thr	Ala	Asn	Gly
	20				25				30						

Asp	Cys	Arg	Arg	Asp	Pro	Arg	Glu	Arg	Ser	Arg	Ser	Pro	Ile	Glu	Arg
	35			40			45								

Ala	Val	Ala	Pro	Thr	Met	Ser	Leu	His	Gly	Ser	His	Leu	Tyr	Thr	Ser
	50				55		60								

Leu	Pro	Ser	Leu	Gly	Leu	Glu	Gln	Pro	Leu	Ala	Leu	Thr	Lys	Asn	Ser
	65				70		75		80						

Leu	Asp	Ala	Ser	Arg	Pro	Ala	Gly	Leu	Ser	Pro	Thr	Leu	Thr	Pro	Gly
	85				90		95								

Glu	Arg	Gln	Gln	Asn	Arg	Pro	Ser	Val	Ile	Thr	Cys	Ala	Ser	Ala	Gly
	100				105		110								

Ala	Arg	Asn	Cys	Asn	Leu	Ser	His	Cys	Pro	Ile	Ala	His	Ser	Gly	Cys
	115				120		125								

Ala	Ala	Pro	Gly	Pro	Ala	Ser	Tyr	Arg	Arg	Pro	Pro	Ser	Ala	Ala	Thr
	130				135		140								

Thr	Cys	Asp	Pro	Val	Val	Glu	Glu	Ala	Ala	Arg	Arg	Ser	Leu	Gly	Lys
	145				150		155		160						

Asn	Tyr	Lys	Glu	Pro	Glu	Pro	Ala	Pro	Asn	Ser	Val	Ser	Ile	Thr	Gly
	165				170		175								

Ser	Val	Asp	Asp	Ala	Ala	Lys	Ala	Leu	Gly	Asp	Thr	Trp	Leu	Gln	
	180				185		190								

Ile	Lys	Ala	Ala	Lys	Asp	Gly	Ala	Ser	Ser	Ser	Pro	Glu	Ser	Ala	Ser
	195				200		205								

Arg	Arg	Gly	Gln	Pro	Ala	Ser	Pro	Ser	Ala	His	Met	Val	Ser	His	Ser
	210				215		220								

His	Ser	Pro	Ser	Val	Val	Ser									
	225				230										

<210> SEQ ID NO 34
<211> LENGTH: 133
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 34

gttagtatca	aggttacaag	acaggttaa	ggagaccaat	agaaaactggg	cttgcgaga	60
------------	------------	-----------	------------	-------------	-----------	----

cagagaagac	tcttgcgttt	ctgataggca	cctattggtc	ttactgacat	ccactttgcc	120
------------	------------	------------	------------	------------	------------	-----

tttctctcca	cag					133
------------	-----	--	--	--	--	-----

-continued

<210> SEQ ID NO 35
<211> LENGTH: 2096
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 35

gcatgc	aat ttatagtgcc gtcactaaca gtactgatac tttaacatgc taagttaaa	60
gtgtgt	gtcta tattaattgt aagattggtg aagagagggtt ttatcagatg gaagctgcac	120
atttctgg	at taatgtggtt aaatgttatct tctcctgtga ttactgtctt tatttctct	180
tttaaaat	at tgcatttgg acatctatct gtatagtcac gcctgacac gtcctcctgg	240
agacagata	aa agatgtacga cgggaggagc agatggaggc aaagcgctgt gatgttttgc	300
tggttt	gagt gcacacattt gttcagtat tctgtgaaat gagtgagcaa atggtgaccg	360
ggtgccctgt	aa atgggtgtt ctacatctta agagaagaac acggacacta ggtaagtgaa	420
gcttgctgtc	actcctctac agcgtcacag agggtcagtc acccttgacc acactgaact	480
agtcgtcacc	tttccactct tcctgccaga agagcagaaa tcagactctc tggggatatc	540
agcctcaccc	ctactgtct ctccattatg aggcaaaactt tctttcactt cccagaggct	600
ctgggggcag	caaggtaaac cc tttccctca gactcttagtc tcggaggaga tcagatcg	660
cttattcaag	ggaaccagcc cctgctctgc gcccctgtcc aaggctgtg aagagtgaca	720
aaaggcacca	cgctgcgggg acgcgggtga agccctctg tgcgtccctt gggcataatc	780
aggaacttgt	gc cccaaatcag aggtgatgtg gccagggtt tgggagtgac gcgccgtgg	840
gaggcttgcg	cacccaaggc acgccccctgc caagtcccac tagcagctct ttggagacct	900
ggggccggc	tc agccacttcc cccagtcct cctccggcaaa ggggctatat agatctccca	960
ggtcagggcg	cagctgcaga agttggtcgt gaggcactgg gcaggtaagt atcaaggta	1020
caagacaggt	ttaaggagac caatagaaac tggcctgtc gagacagaga agactttgc	1080
gtttctgata	ggcacctatt ggtttaactg acatccactt tgccttctc tccacaggt	1140
tccactccca	gttcaattac agctttaag gctagagtac ttaatacgc acactatagg	1200
ctagcatgt	at ttatgaag atggacctgt tgaactatca gtactggac aagatgaaca	1260
acaatatcg	cattctgtgc tacgaaggcg aagctgtctc caggggagaa cccagaatgc	1320
agaccctgcc	ggtggctctt gcccctcagca gtcaccgcac cggccctccc ccaatcagcc	1380
ccagcaagag	gaagttcagc atggagccag gtgacgaggaa cctagactgt gacaacgacc	1440
acgtctccaa	aatgagtcgc atcttcaacc cccatctgaa caagactgcc aatggagact	1500
gcccgcagaga	ccccccggag cggagccgca gccccatcga ggcgcgtgtg gccccacca	1560
tgagcctgca	cgccagccac ctgtacaccc ccttccccag cttggcctg gagcagcccc	1620
tcgcactgac	caagaacagc ctggacgcca gcaggccagc cggccctctcg cccacactga	1680
ccccggggga	gcccggccatc aaccggccctt ccgtgatcac ctgtgcctcg gctggccccc	1740
gcaactgc	aactctcgac tggccatcg cgacagccgg ctgtgcggcg cccgggctg	1800
ccagctaccg	gaggccaccg agcgtgcaca ccacctgtga cccctgtgg gaggagcatt	1860
tccgcaggag	cctggcaag aattacaagg agcccgagcc ggcacccaac tccgtgtcca	1920
tcacgggctc	cgtggacgac cactttgcca aagctctggg tgacacgtgg ctccagatca	1980
aagcggccaa	ggacggagca tccagcagcc ctgagtcgc ctctcgagg ggcagcccg	2040
ccagccccctc	tgcccacatg gtcagccaca gtcactcccc ctctgtggtc tccgc	2096

-continued

<210> SEQ_ID NO 36
 <211> LENGTH: 2096
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Laboratory-synthesized sequence
 <400> SEQUENCE: 36

gcatgccaat ttatagtgcc	gtcactaaca gtactgatac	ttaacatgc taagttaaa	60
gtgtgtgcta tattaattgt	aagattggtg aagagagggt	ttatcagatg gaagctgcag	120
cagcctggat taatgtggtt	aatgttatct tccctgtga	ttactgtctt tatttcttct	180
tttaaaatat tgtcatttgg	acatctatct gtatactac	gccctgacac gtcctctgg	240
agacagataa	gaagttacga cggggaggago	agatggaggc aaagcgctgt	300
tggttttagt	gcacacattt gttcagtgtat	tctgtgaaat gagtgagcaa atggtgaccg	360
ggtgcccgtt	aatgtgtttt ctacatctta	agagaagaac acggacacta ggtaagtgaa	420
gcttgctgtc	actcctctac agcgtcacag	agggtcagtc acccttgacc acactgaact	480
agtgcgtacc	tttccactct tcctgccaga	agagcagaaaa tcagactctc tggggatatc	540
agccctcaccc	ctactgctct ctccattatg	aggccaaactt tctttcaattt cccagaggt	600
ctggggggcag	caaggtaaac ccttcctca	gactctagtc tcggaggaga tcagatcgcg	660
cttattcaag	ggaaccagcc cctgtctgc	gccctggtcc aaggctgtg aagagtgaca	720
aaaggccacca	cgctgcccc	acgggggtga agetcctctg tgggtctct gggcataatc	780
aggaacttgtt	gccaatcag aggtgatgt	gccagggctt tgggagtgac gcgcggctgg	840
gaggcttgcg	caccaaggc acgccccctgc	caagttccac tagcagctct ttggagacct	900
gggcccggctc	agccacttcc cccagtcct	cctccggcaa gggctatata agatctccca	960
ggtcagggcg	cagctgcaga agttggctgt	gaggcaactgg gcaggtaatg atcaaggta	1020
caagacaggt	ttaaggagac caatagaaac	tgggcttgc gagaacagaga agactctgc	1080
gtttctgata	ggcacctatt ggtttactg	acatcgacgc tgccttctc tccacaggt	1140
tccactccca	gttcaattac agctcttaag	gctagagtac ttaatacgac tcactatagg	1200
ctagcatgt	atttatgaag atggacctgt	tgaactatca gtacttggac aagatgaaca	1260
acaatatcg	cattctgtgc tacgaaggg	aagctgtctc cagggagaa cccagaatgc	1320
agaccctgcc	ggtggctctc	gccctcagca gtcacccgcac cggccctccc ccaatcagcc	1380
ccagcaagag	gaagttcagc atggagccag	gtgacgagga cctagactgt gacaacgacc	1440
acgtctccaa	aatgagtcgc atcttcaacc	cccatctgaa caagactgcc aatggagact	1500
gccgcagaga	ccccggggag	cgaggccgcac gccccatcga gcgcgtgtg gccccacca	1560
tgagcctgca	cggcagccac	ctgtacacct ccctccccag cttggccctg gagcagcccc	1620
tcgcactgac	caagaacagc	ctggacgcca gcaggccagc cggccctctcg cccacactga	1680
ccccggggga	gcggcagcag	aaccggccct ccgtgatcac ctgtgcctcg gctggccccc	1740
gcaactgcaa	cctctcgac	tgccocatcg cgacacggcg ctgtgcogcg cccgggctg	1800
ccagctaccg	gaggccaccg	agcgtgtcca ccacctgtga cccctgggtg gaggaggcag	1860
cccgcgaggag	cctgggcaag	aattacaagg agcccgagcc ggcacccaac tccgtgtcca	1920
tcacgggctc	cgtggacgac	gcagctgcca aagctctggg tgacacgtgg ctccagatca	1980
aagcggccaa	ggacggagca	tccagcagec ctgagtcggc ctctcgagg ggccagccg	2040

US 12,384,823 B2

75

76

-continued

ccagccccctc tgcccacatg gtcagccaca gtcactcccc ctctgtggtc tccgcc 2096

<210> SEQ ID NO 37
<211> LENGTH: 239
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 37

Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu
1				5				10					15		

Val	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly
	20				25				30						

Glu	Gly	Gly	Asp	Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	
	35				40			45							

Cys	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr
	50				55			60							

Leu	Thr	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys
	65				70			75		80					

Gln	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu
	85				90			95							

Arg	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu
	100				105			110							

Val	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly
	115				120			125							

Ile	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr
	130				135			140							

Asn	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn
	145				150			155		160					

Gly	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser
	165				170			175							

Val	Gln	Leu	Ala	Asp	His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly
	180				185			190							

Pro	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu
	195				200			205							

Ser	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Glu	Phe	
	210				215			220							

Val	Thr	Ala	Ala	Gly	Ile	Thr	Leu	Gly	Met	Asp	Glu	Leu	Tyr	Lys	
	225				230			235							

<210> SEQ ID NO 38
<211> LENGTH: 720
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 38

atggtgagca	aggggcgagga	gctgttcacc	gggggtgggc	ccatcctgg	cgagctggac	60
------------	-------------	------------	------------	-----------	------------	----

ggcgcacgtaa	acggccacaa	gttcagcggt	tccggcgagg	gcgaggggcg	tgccacccat	120
-------------	------------	------------	------------	------------	------------	-----

ggcaagctga	ccctgaaagt	catctgcacc	accggcaagc	tgcccgtgcc	ctggccacc	180
------------	------------	------------	------------	------------	-----------	-----

ctcgtgacca	ccctgaccta	cggcggtcag	tgcttcagcc	gctaccccg	ccacatgaag	240
------------	------------	------------	------------	-----------	------------	-----

cagcacgact	tcttcaagtc	cgccatgccc	gaaggctacg	tccaggagcg	caccatctc	300
------------	------------	------------	------------	------------	-----------	-----

ttcaaggacg	acggcaacta	caagacccgc	gccgagggtg	agttcgagg	cgacaccctg	360
------------	------------	------------	------------	-----------	------------	-----

-continued

gtgaaccgca tcgagctgaa gggcatcgac ttcaaggagg acggcaacat cctggggcac	420
aagctggagt acaactacaa cagccacaac gtctatatca tggccgacaa gcagaagaac	480
ggcatcaagg tgaacttcaa gatecgccac aacatcgagg acggcagcgt gcagetcgcc	540
gaccactacc agcagaacac ccccatcgga gacggccccg tgctgtgcc cgacaaccac	600
tacctgagca cccagtcgc cctgagcaaa gacccaacg agaagcgcga tcacatggc	660
ctgctggagt tcgtgaccgc cgccggata actctcgca tggacgagct gtacaagtaa	720

<210> SEQ ID NO 39
 <211> LENGTH: 2816
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Laboratory-synthesized sequence
 <400> SEQUENCE: 39

gcatgccaat ttatagtgcc gtcactaaca gtactgatac tttaacatgc taagttaaa	60
gtgtgtgcta tattaattgt aagattggtg aagagagggt ttatcagatg gaagctgcac	120
atttctggat taatgtggtt aaatgttatct tctcctgtga ttactgtctt tatttcttct	180
tttaaaatat tgcatttgg acatctatct gtatagctac gcccgtacac gtcctcctgg	240
agacagataa gaagttacga cgggaggagc agatggaggc aaagcgctgt gatgctttg	300
tggtttgggt gcacacatTTt gttcagtat tctgtgaaat gagttagacaa atggtagccg	360
ggtgccctgt aaatgggttt ctatcttta agagaagaac acggacacta ggtaagtgaa	420
gcttgctgtc actcctctac agcgtcacag agggtcagtc acccttgacc acactgaact	480
agtcgtcacc ttccactct tcctgccaga agacgacaaa tcagactctc tggggatatc	540
agcctcaccct ctactgctct ctccattatg aggccaaactt tctttactt cccagaggct	600
ctggggccag caaggtaaac ccttcctca gactctagtc tcggaggaga tcagatcg	660
cttattcaag ggaaccagcc cctgtctgc gcccgtgtcc aaggctgtt aagagtgaca	720
aaaggcacca cgctgcgggg acgcgggtga agccctctg tgggtctct gggcataatc	780
aggaacttgtt gccaaatcag aggtgatgtg gccagggttt tgggagtgac gcgccgtgg	840
gaggcttgcg caccaaggc acgcacctgc caagtcccac tagcagctct ttggagac	900
ggcccggttc agccacttcc cccagttcc cctccggcaaa ggggtatat agatctccca	960
ggtcagggcg cagctgcaga agttggctgt gaggactgg gcaggtaagt atcaaggta	1020
caagacaggtt ttaaggagac caatagaaac tggcttgatc gagacagaga agactcttc	1080
gtttctgata ggcacctatt ggttacttgc acatccactt tgcctttctc tccacagg	1140
tccactccca gttcaattac agctttaag gctagactac ttaatacgac tcactatagg	1200
ctagcatgct atttatgaag atggacactgt tgaactatca gtacttgac aagatgaaca	1260
acaatatcggtt cattctgtgc tacgaaggcg aagctgtct caggggagaa cccagaatgc	1320
agacccctgcc ggtggctctt gcccgtacca gtcaccgcac cggccctccc ccaatcagcc	1380
ccagcaagag aggttcagc atggagccag gtgacgaggc cctagactgt gacaacgacc	1440
acgtctccaa aatgagtcgc atcttcaacc cccatctgaa caagactgcc aatggagact	1500
ggccgcagaga cccccgggag cggagccgca gccccatcgaa ggcgcgtgtg gccccacca	1560
tgagcctgca cggcagccac ctgtacactt ccctcccaag cttggctgt gaggccccc	1620
tgcactgac caagaacagc ctggacgcca gcaggccagc cggccctctcg cccacactga	1680
ccccggggga gggcagcag aacggccctt ccgtgatcac ctgtgcctcg gctggccccc	1740

-continued

gcaactgcaa cctctcgac tgccccatcg cgcacagcgg ctgtgccgcg cccgggcctg	1800
ccagctaccg gaggccaccg agcgctgcca ccacctgtga ccccgtggtg gaggagcatt	1860
tccgcaggag cctggcaag aattacaagg agcccgagcc ggcacccaac tccgtgtcca	1920
tcacgggctc cgtggacgac cacttgcga aagctctggg tgacacgtgg ctccagatca	1980
aagcggccaa ggacggagca tccagcagcc ctgagtcgc ctctcgcaagg ggccagcccg	2040
ccagccccctc tgccccatcg tgcagccaca gtcaactccca ctctgtggtc tccgcacatgg	2100
tgagcaaggcgaggagctg ttacccgggg tggtgcccat cctggtcgag ctggacggcg	2160
acgttaaacccg ccacaagttc agcgtgtccg gcgaggggcga gggcgatgcc acctacggca	2220
agctgaccct gaagttcatc tgcaccacccg gcaagctgca cgtgccttgg cccaccctcg	2280
tgaccacccct gacccatcgcc gtgcagtgtc tcagccgcta ccccgaccac atgaaggcagc	2340
acgacttctt caagtccgccc atgcccgaag gctacgtcca ggagcgcacc atcttcttca	2400
aggacgacgg caactacaag accccgcggcc aggtgaagtt cgagggcgac accctggta	2460
accgcacatcgaa gctgaagggc atcgacttca aggaggacgg caacatctg gggcacaagg	2520
tggagtacaa ctacaacacgc cacaacgtct atatcatggc cgacaagcag aagaacggca	2580
tcaagggtgaa cttcaagatc cgccacaaca tgcaggacgg cagcgtgcag ctgcggacc	2640
actaccagca gaacacccccc atcggcgacg gccccgtgct gtcggccgac aaccactacc	2700
tgagcaccca gtcggccctg agcaaagacc ccaacgagaa ggcgcgtac atggccctgc	2760
tggagttcggt gaccggccccc gggatcactc tcggcatggc cgagctgtac aagtaa	2816

<210> SEQ ID NO 40
 <211> LENGTH: 2816
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Laboratory-synthesized sequence

<400> SEQUENCE: 40	
gcatgcacat ttatagtgcc gtcactaaca gtactgatac tttaacatgc taagttaaa	60
gtgtgtgtcta tattaatttgt aagattggtg aagagagggtt ttatcagatg gaagctgcag	120
cagcctggat taatgtggtt aaatgttatct tctcctgtga ttactgtctt tatttctct	180
tttaaaatat tgcattttgg acatctatct gtatagtac gcccgtacac gtcctcctgg	240
agacagataa gaagttacga cgggaggagc agatggaggc aaagcgtgt gatgttttgc	300
tggtttggat gcacacattt gttcagtgtatc tctgtgaaat gagtgagcaa atggtgaccg	360
gggtggccctgt aaatgggttt ctacatcttta agagaagaac acggacacta ggtaagtgaa	420
gcttgcgttc actcctctac agcgtcacag agggtcagtc acccttgacc acactgaact	480
agtcgtcacc tttccactct tcctgtccaga agagcagaaaa tcagactctc tggggatatc	540
agcctcaccct ctactgtctc ctccattatg aggcaaactt tctttcactt cccagaggct	600
ctggggccag caaggtcaac ctttcctca gactctagtc tcggaggaga tcagatcg	660
cttattcaag ggaaccagcc cctgtctgc gcccgtgtt aaggctgtt aagagtgaca	720
aaaggcaccatc cgtcgccggg acgcgggtga agccctctg tgcgtcttgc gggcataatc	780
aggaactggt gccaaatcag aggtgatgtg gccagggtt tggggatgtac ggcgcggctgg	840
gaggcttgcg cacccaaaggc acggccctgc caagtcccac tagcagctct ttggagaccc	900
ggccggctc agccacttcc cccagtcctt cttccggcaa gggctatata agatctccca	960

US 12,384,823 B2

81**82**

-continued

```

ggtcagggcg cagctgcaga agttggcgtg gaggcactgg gcaggtaagt atcaaggta 1020
caagacaggt ttaaggagac caatagaaac tgggcttgcc gagacagaga agactctgc 1080
gtttctgata ggcacctatt ggtttaactg acatcgccgc tgccttcctc tccacaggtg 1140
tccactccca gttcaattac agctcttaag gctagagtgatc ttaatacgac tcactatagg 1200
ctagcatgct atttatgaag atggacctgt tgaactatca gtacttggac aagatgaaca 1260
acaatatcggtt cattctgtgc tacgaaggg aagctgtct caggggagaa cccagaatgc 1320
agaccctgcc ggtggcctct gcctcagca gtcacccgac cggccctccc ccaatcagcc 1380
ccagcaagag gaagttcagc atggagccag gtgacgagga cctagactgt gacaacgacc 1440
acgtctccaa aatgagtcgc atcttcaacc cccatctgaa caagactgcc aatggagact 1500
ggcgcaagaga cccccgggag cgagccgcga gccccatcga ggcgcgtgtg gccccacca 1560
tgagcctgca cggcagccac ctgtacaccc ctctccctag ctttggcctg gagcagcccc 1620
tcgcactgac caagaacacgc ctggacgcac gcaggccagc cggccctctcg cccacactga 1680
ccccggggga gggggcggcag aaccggccct ccgtgatcac ctgtgcctcg gctggcggcc 1740
gcaactgcaa cctctcgac tggccatcg cgcacagcgg ctgtgcgcgc cccgggcctg 1800
ccagctaccg gaggccaccg agcgtgcacca ccacctgtga cccctgtgtg gaggaggcag 1860
ccgcaggag cctggcaag aattacaagg agcccgagcc ggcacccaac tccgtgtcca 1920
tcacgggctc cgtggacgcac gcaagctgcca aagctctggg tgacacgtgg ctccagatca 1980
aagcggccaa ggacggagca tccagcagcc ctgagtcgc ctctcgcagg ggcagcccg 2040
ccagccccctc tgccccatcg tgcagccaca gtcaactcccc ctctgtggc tccgcctatgg 2100
tgagcaaggg cgaggagctg ttccacgggg tgggtccccat cctggcgtgag ctggacggcg 2160
acgtaaacgg ccacaagttc agcgtgtccg gcgaggccgca gggcgtatgcc acctacggca 2220
agctgaccct gaagttcatc tgcaccaccg gcaagctgcc cgtgcctcg cccaccctcg 2280
tgaccaccct gacccatccgc gtgcgtgtct tcagccgcta ccccgaccac atgaagcagc 2340
acgacttctt caagtcgcgc atgcccgaag gctacgtcca ggagcgcacc atcttcttca 2400
aggacgacgg caactacaag accccgcgcg aggtgaattt cgagggcgcac accctggta 2460
accgcacatcg gctgaaggcc atcgacttca aggaggacgg caacatctg gggcacaaggc 2520
tggagtacaa ctacaacacgc cacaacgtct atatcatggc cgacaacgcag aagaacggca 2580
tcaaggtgaa cttcaagatc cgccacaaca tcgaggacgg cagcgtgcag ctcgcgcacc 2640
actaccagca gaacacccccc atcggcgcacg gccccgtgtc gctgcccgcac aaccactacc 2700
tgagcaccca gtccgcctg agcaaagacc ccaacgagaa ggcgcgtatcac atggcctgc 2760
tggagttcggtt gaccggccgc gggatcaatc tcggcatggc cgagctgtac aagtaa 2816

```

<210> SEQ ID NO 41

<211> LENGTH: 16

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 41

Asp	Pro	Val	Val
Glu	Glu	His	Phe
1	5	10	15

<210> SEQ ID NO 42

<211> LENGTH: 16

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 42

Thr	Gly	Ser	Val	Asp	Asp	His	Phe	Ala	Lys	Ala	Leu	Gly	Asp	Thr	Trp
1				5				10					15		

<210> SEQ ID NO 43

<211> LENGTH: 16

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Homo sapiens

<400> SEQUENCE: 43

Asp	Pro	Val	Val	Glu	Glu	Ala	Ala	Arg	Arg	Ser	Leu	Gly	Lys	Asn	Tyr
1				5				10					15		

<210> SEQ ID NO 44

<211> LENGTH: 16

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Homo sapiens

<400> SEQUENCE: 44

Thr	Gly	Ser	Val	Asp	Asp	Ala	Ala	Lys	Ala	Leu	Gly	Asp	Thr	Trp
1				5				10				15		

What is claimed is:

1. A method, comprising administering a polynucleotide to a subject, wherein the polynucleotide comprises:
 a cis-regulatory element and a nucleotide sequence encoding a vestigial like 4 protein, wherein
 the cis-regulatory element comprises an uncoupling protein 1 enhancer and an uncoupling protein 1 promoter, and
 the vestigial like 4 protein comprises a first TDU domain having a first amino acid sequence and a second TDU domain having a second amino acid sequence, wherein
 the first amino acid sequence comprises the amino acid sequence as set forth in SEQ ID NO: 41 with two amino acid substitutions, wherein the seventh amino acid of the amino acid sequence as set forth in SEQ ID NO: 41 is not H and the eighth amino acid of the amino acid sequence as set forth in SEQ ID NO: 41 is not F, and
 the second amino acid sequence comprises the amino acid sequence as set forth in SEQ ID NO: 42, wherein the seventh amino acid of the amino acid sequence as set forth in SEQ ID NO: 42 is not H and the eighth amino acid of the amino acid sequence as set forth in SEQ ID NO: 42 is not F.

2. The method of claim 1, wherein the uncoupling protein 1 enhancer has at least 90% identity with a sequence selected from SEQ ID NO: 1, SEQ ID NO 4, and SEQ ID NO: 7.

3. The method of claim 1, wherein the uncoupling protein 1 promotor has at least 90% identity with a sequence selected from SEQ ID NO: 2, SEQ ID NO 5, and SEQ ID NO: 8.

4. The method of claim 1, wherein the cis-regulatory element has at least 90% homology with a sequence selected from SEQ ID NO: 3, SEQ ID NO: 6, and SEQ ID NO: 9.

5. The method of claim 1, wherein the vestigial like 4 protein has at least 90% homology with a sequence selected from SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, and SEQ ID NO: 33.

6. The method of claim 1, wherein the sequence encoding a vestigial like 4 protein has at least 90% identity with a sequence selected from SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, and SEQ ID NO: 32.

7. The method of claim 1, wherein the vestigial like 4 protein has from 0 to 3 substitutions to a sequence selected from SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, and SEQ ID NO: 33, wherein the substitutions are not in the first TDU domain or the second TDU domain.

8. The method of claim 1, wherein the polynucleotide further comprises an intron between the cis-regulatory element and the nucleotide sequence encoding a vestigial like 4 protein.

9. The method of claim 8, wherein the intron has at least 90% homology with SEQ ID NO: 34.

10. The method of claim 1, wherein the administering comprises administering a viral vector, and the viral vector comprises the polynucleotide.

11. The method of claim 10, wherein the viral vector is comprises an adenoviral associated vector.

12. The method of claim 10, wherein the uncoupling protein 1 enhancer has at least 90% identity with a sequence selected from SEQ ID NO: 1, SEQ ID NO 4, and SEQ ID NO: 7.

13. The method of claim 10, wherein the uncoupling protein 1 promotor has at least 90% identity with a sequence selected from SEQ ID NO: 2, SEQ ID NO 5, and SEQ ID NO: 8.

14. The method of claim 10, wherein the cis-regulatory element has at least 90% homology with a sequence selected from SEQ ID NO: 3, SEQ ID NO: 6, and SEQ ID NO: 9.

15. The method of claim 10, wherein the vestigial like 4 protein has at least 90% homology with a sequence selected from SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, and SEQ ID NO: 33.

16. The method of claim **10**, wherein the sequence encoding a vestigial like 4 protein has at least 90% identity with a sequence selected from SEQ ID NO: 22, SEQ ID NO: 24, SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 30, and SEQ ID NO: 32. ⁵

17. The method of claim **10**, wherein the vestigial like 4 protein has from 0 to 3 substitutions to a sequence selected from SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, and SEQ ID NO: 33, wherein the substitutions are not in the first TDU domain or ¹⁰ the second TDU domain.

18. The method of claim **1**, wherein the subject is obese or is at risk for developing obesity.

19. The method of claim **1**, wherein the subject has hepatic steatosis or is at risk for developing hepatic steatosis. ¹⁵

20. The method of claim **1**, wherein the subject has diabetes or is at risk for developing diabetes.

* * * * *