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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 October 2024 (31.10.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/227074 A1

(51) International Patent Classification:
A61K 31/41 (2006.01) *A61K 48/00* (2006.01)
C12N 15/861 (2006.01) *C12N 15/113* (2010.01)
A61K 31/52 (2006.01) *C12N 15/63* (2006.01)

(21) International Application Number:
PCT/US2024/026643

(22) International Filing Date:
26 April 2024 (26.04.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/499,181 28 April 2023 (28.04.2023) US

(71) Applicant: **KRIYA THERAPEUTICS, INC.** [US/US]; 4105 Hopson Rd, Morrisville, North Carolina 27560 (US).

(72) Inventors: **DIETRICH, Nicholas**; c/o Kriya Therapeutics, Inc., 4105 Hopson Rd, Morrisville, North Carolina 27560 (US). **RABINOWITZ, Joseph**; c/o Kriya Therapeutics, Inc., 4105 Hopson Rd, Morrisville, North Carolina 27560 (US). **ISRAEL, Jennifer**; c/o Kriya Therapeutics, Inc., 4105 Hopson Rd, Morrisville, North Carolina 27560 (US).

(74) Agent: **NANNENGA-COMBS, Bonnie W.** et al.; STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C., 1101 K Street NW, 10th Floor, Washington, District of Columbia 20005 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2024/227074 A1

(54) Title: RIBOSWITCHES FOR REGULATING GENE EXPRESSION AND THERAPEUTIC METHODS OF USING THE SAME

(57) Abstract: The present disclosure provides an expression cassette comprising from 5' to 3': (i) promoter, (ii) an upstream exon, (iii) an upstream intron, (iv) an alternatively spliced exon comprising a translation initiation sequence, (v) a downstream intron, (vi) a downstream exon, and (vii) a heterologous nucleic acid sequence. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon, each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

RIBOSWITCHES FOR REGULATING GENE EXPRESSION AND THERAPEUTIC METHODS OF USING THE SAME

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims the priority benefit of U.S. Provisional Application No. 63/499,181, filed April 28, 2023, the contents of which are hereby incorporated by reference in their entirety.

REFERENCE TO SEQUENCE LISTING SUBMITTED ELECTRONICALLY

[0002] The content of the electronically submitted sequence listing (Name: 4525_105PC01_Seqlisting_ST26.xml; Size: 178,656 bytes; and Date of Creation: April 25, 2024), filed with the application, is incorporated herein by reference in its entirety.

FIELD OF DISCLOSURE

[0003] The present disclosure pertains generally to the field of molecular biology and medicine (e.g., gene therapy). More particularly it relates to compositions and methods for regulating gene expression.

BACKGROUND

[0004] Gene therapy represents a robust tool for treating human disease. While gene therapy technologies have developed in recent decades, there is a need to further control gene expression, for example, to regulate symptoms and reduce protein toxicity.

[0005] A riboswitch is a regulatory segment of a messenger RNA (mRNA) molecule that binds a small molecule, which can result in a change in production of the proteins encoded by the mRNA. Most known riboswitches occur in bacteria, but functional riboswitches (e.g., the TPP riboswitch) have been discovered in archaea, plants and certain fungi. Riboswitches can operate using different mechanisms. For example, certain riboswitch structures can affect the splicing of pre-mRNA. Bacteria make extensive use of riboswitches to sense metabolites and control gene expression, and typically do so by modulating premature transcription termination or translation initiation. The most widespread riboswitch class known in bacteria responds to the coenzyme thiamine pyrophosphate (TPP), which is a derivative of vitamin B_1. *See* Cheah et al., Nature volume 447, 497–500 (2007). Methods for controlling gene expression using drug inducible alternative splicing switches have been investigated (see, e.g., Int. Publ. No. WO2021163556A1).

To use a riboswitch in a therapeutic setting, the small molecule must be both effective at regulating the gene of interest and non-toxic to the patient.

BRIEF SUMMARY

[0006] Certain aspects of the disclosure are directed to an expression cassette comprising from 5' to 3': (i) promoter, (ii) an upstream exon, (iii) an upstream intron, (iv) an alternatively spliced exon comprising a translation initiation sequence, (v) a downstream intron, (vi) a downstream exon, and (vii) a heterologous nucleic acid sequence.

[0007] In some aspects, the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0008] In some aspects, the upstream exon and the downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0009] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the a downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0010] In some aspects, the upstream exon and the downstream exon each comprise a portion of a gene, which is different from the gene for the portion of the upstream intron, the alternatively spliced exon, and the downstream intron.

[0011] In some aspects, the upstream exon and the downstream exon comprise synthetic exons.

[0012] In some aspects, the inclusion of the alternatively spliced exon is regulatable (e.g., regulated) by a splicing modulator.

[0013] In some aspects, the splicing modulator is kinetin or homocarbonyltopsentin.

[0014] In some aspects, the gene is TRPV3, STK31, or USP25; and the splicing modulator is Kinetin.

[0015] In some aspects, the gene is GTF2B, LYPD3, IK, or TEX14; and the splicing modulator is Homocarbonyltopsentin.

[0016] In some aspects, the splicing modulator promotes inclusion of the alternatively spliced exon thereby driving expression of the heterologous nucleic acid sequence.

[0017] In some aspects, the promoter is selected from the group consisting of a CBA promoter, a smCBA promoter, a CMV promoter, an EF-1a (Elongation Factor 1a) promoter, a RSV (Rous Sarcoma Virus) promoter, an Ubiquitin (UbC) promoter, a CAG promoter, a SV40 (simian vacuolating virus 40) promoter, a PGK (phosphoglycerate kinase) promoter, a human beta actin promoter, a H1 (human polymerase III RNA) promoter, a human U6 small nuclear promoter, tetracycline responsive element, and any combination thereof.

[0018] In some aspects, the expression cassette further comprises an enhancer.

[0019] In some aspects, the enhancer is selected from the group consisting of a CMV enhancer, a SV40 enhancer, and any combination thereof.

[0020] In some aspects, the expression cassette further comprises a cleavage site between the downstream exon and the heterologous nucleic acid sequence.

[0021] In some aspects, the cleavage site is selected from the group consisting of a furin cleavage site, a foot-and-mouth disease virus 18 2A peptide (F2A), an equine rhinitis A virus 2A peptide (E2A), a porcine teschovirus-1 2A peptide (P2A), a thosea asigna virus 2A peptide (T2A), and any combination thereof.

[0022] In some aspects, the expression cassette further comprises a leader sequence between the downstream exon and the heterologous nucleic acid sequence.

[0023] In some aspects, the leader sequence is selected from the group consisting of an opticin (e.g., a human opticin (hOPT)) leader sequence, an interleukin-2 (e.g., human interleukin-2 (hIL-2)) leader sequence, an interleukin-12 (e.g., human interleukin-12 (hIL-12)) leader sequence, an interleukin-6 (e.g., a human or mouse interleukin-6 (hIL-6 or mIL-6)) leader sequence, an insulin (e.g., human insulin) leader sequence, a serum albumin (e.g., human serum albumin (HSA)) leader sequence, and any combination thereof.

[0024] In some aspects, the combined length of the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon is between about 300 basepairs to about 2100 basepairs.

[0025] In some aspects, the length of the heterologous nucleic acid sequence is less than about 4100 basepairs, less than about 4000 basepairs, or less than about 3900 basepairs.

[0026] In some aspects, the heterologous nucleic acid sequence encodes a protein. In some aspects, the protein is a therapeutic protein.

[0027] In some aspects, the upstream exon comprises a sequence selected from the group consisting of SEQ ID NO: 8, 13, 18, 23, 28, 33, and 38.

[0028] In some aspects, the alternatively spliced exon comprises a sequence selected from the group consisting of SEQ ID NO: 10, 15, 20, 25, 30, 35, and 40.

[0029] In some aspects, the downstream exon comprises a sequence selected from the group consisting of SEQ ID NO: 12, 17, 22, 27, 32, 37, and 42.

[0030] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NO: 8-42.

[0031] In some aspects, the expression cassette comprises a sequence selected from the group consisting of one or more of SEQ ID NO: 1-7 and 43-66.

[0032] In some aspects, the expression cassette further comprises an untranslated region (UTR), a microRNA binding site, a polyA (pA) sequence, an intron sequence, or any combination thereof.

[0033] In some aspects, the expression cassette comprises a poly(A) (pA) sequence.

[0034] In some aspects, the pA sequence is a synthetic pA sequence, a bovine growth hormone (bGH) pA sequence, or a human growth hormone (hGH) pA sequence.

[0035] In some aspects, provided herein is a delivery vector comprising any of the expression cassettes disclosed herein.

[0036] In some aspects, the expression cassette is flanked by inverted terminal repeat (ITR) sequences.

[0037] In some aspects, the delivery vector is selected from the group consisting of a viral vector, a plasmid, a lipid, a protein particle, a bacterial vector, a lysosome, a virus-like particle, a polymeric particle, an exosome, and a vault particle.

[0038] In some aspects, the delivery vector is a viral vector.

[0039] In some aspects, the viral vector is a lentiviral vector, an adenoviral vector, an adeno-associated viral (AAV) vector, or a retroviral vector.

[0040] In some aspects, the viral vector is an adeno-associated viral (AAV) vector.

[0041] In some aspects, the delivery vector is a recombinant AAV (rAAV) vector comprising an AAV serotype selected from the group consisting of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAV9, AAV10, AAVrh10, AAV11, and AAV12.

[0042] In some aspects, provided herein is a cell comprising any of the expression cassettes disclosed herein or any of the delivery vectors disclosed herein.

[0043] In some aspects, provided herein is a pharmaceutical composition comprising any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, or any of the cells disclosed herein.

[0044] In some aspects, the pharmaceutical composition further comprises a pharmaceutically acceptable excipient.

[0045] Certain aspects of the disclosure are directed to a kit comprising i) any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the pharmaceutical compositions disclosed herein; and (ii) optionally instructions for use.

[0046] In some aspects, the kit further comprises (iii) a splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0047] In some aspects, the splicing modulator molecule is kinetin.

[0048] In some aspects, the splicing modulator molecule is homocarbonyltopsentin.

[0049] Certain aspects of the disclosure are directed to a method of regulating expression of a protein in a mammalian cell comprising administering to the cell any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, or any of the pharmaceutical compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0050] Certain aspects of the disclosure are directed to a method of expressing a protein in a cell or subject comprising administering to the cell or subject any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the pharmaceutical compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0051] Certain aspects of the disclosure are directed to a method of treating a disease in a mammal comprising administering to a subject any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the

pharmaceutical compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0052] In some aspects, the administration of the expression cassette, delivery vector, or cell is systemic or tissue specific. In some aspects, the administration is intravenous, intramuscular, intrahepatic, intraventricular, intraparenchymal, intrathecal, subcutaneous, oral, intraocular, periocular, or intratumoral.

[0053] In some aspects, the administration of the molecule that promotes the inclusion of the alternatively spliced exon is systemic.

[0054] In some aspects, the molecule that promotes the inclusion of the alternatively spliced exon is administered at least 1 day after administration of the expression cassette.

[0055] In some aspects, the method further comprises administering a second dose of the molecule that promotes the inclusion of the alternatively spliced exon.

[0056] Certain aspects of the disclosure are directed to a method of regulating transgene expression post-transcriptionally, comprising administering to a subject (a) a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) a heterologous nucleic acid sequence; and (b) a splicing modulator.

[0057] In some aspects, the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0058] In some aspects, the upstream exon and the downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0059] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the a downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31

(STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0060] In some aspects, the upstream exon and the downstream exon each comprise a portion of a gene, which is different from the gene for the portion of the upstream intron, the alternatively spliced exon, and the downstream intron. In some aspects, the upstream exon and the downstream exon comprise synthetic exons.

[0061] In some aspects, the splicing modulator is Kinetin or Homocarbonyltopsentin.

[0062] Certain aspects of the disclosure are directed to a method of producing an adeno-associated viral (AAV) vector comprising: (a) introducing a first plasmid comprising an adenovirus helper gene and a second plasmid comprising a heterologous nucleic acid sequence into a host cell; and

(b) culturing the host cell in the presence of a splicing modulator to produce AAV particles.

[0063] In some aspects, the host cell comprises an AAV capsid (Cap) gene and a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV replication (Rep) gene.

[0064] In some aspects, the host cell comprises an AAV replication (Rep) gene and a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV capsid (Cap) gene.

[0065] In some aspects, the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0066] In some aspects, the upstream exon and the downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0067] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the a downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0068] In some aspects, the upstream exon and the downstream exon each comprise a portion of a gene, which is different from the gene for the portion of the upstream intron, the alternatively spliced exon, and the downstream intron. In some aspects, the upstream exon and the downstream exon comprise synthetic exons.

[0069] In some aspects, the method further comprises (c) harvesting the AAV particles from the cultured host cell.

[0070] In some aspects, the method further comprises (d) purifying the harvested AAV particles.

[0071] In some aspects, the host cell is a HEK293 cell.

[0072] In some aspects, the splicing modulator is Kinetin or Homocarbonyltopsentin.

[0073] In some aspects, the adenovirus helper gene is selected from the group consisting of an Early 4 (E4) gene, an Early 2A (E2A) gene, and a viral associated (VA) gene.

[0074] In some aspects, the nucleic acid construct is stably integrated into the host cell genome.

[0075] In some aspects, the nucleic acid construct comprises a sequence selected from the group consisting of SEQ ID NO: 59-62.

[0076] Certain aspects of the disclosure are directed to a method of producing an AAV particle comprising contacting a host cell comprising rep genes, cap genes, and adenoviral helper genes with a topsentin alkaloid or derivative thereof.

[0077] In some aspects, the method: i) reduces viral response genes, ii) increases immediate early genes, iii) increases a NF-κB signaling pathway, iv) increases a TNF signaling pathway, v) increases a IL-17 signaling pathway, vi) inhibits apoptosis, vii) inhibits viral defense and/or vii) increases histone levels.

[0078] In some aspects, the apoptosis is inhibited by altering a signaling pathway.

[0079] In some aspects, the signaling pathway comprises a NF-κB signaling pathway, a TNF signaling pathway, IL-17 signaling pathway, a JAK/STAT signaling pathway, a P38 MAPK signaling pathway, PI3K-Akt signaling pathway, or any combination thereof.

[0080] In some aspects, the increased histone levels result in increased metabolic activity and/or increased host cell replication.

[0081] In some aspects, the host cell comprises an AAV capsid (Cap) gene, an AAV replication (Rep) gene, an AAV helper gene, or any combination thereof.

[0082] In some aspects, the host cell is a HEK293 cell.

[0083] In some aspects, the topsentin alkaloid or derivative thereof is selected from the group consisting of Homocarbonyltopsentin, Topsentin, Bromotopsentin, Bromodeoxytopsentin, Deoxytopsentin, and any combination thereof. In some aspects, the topsentin alkaloid or derivative thereof is Homocarbonyltopsentin.

[0084] In some aspects, the host cell is contacted with about 5 μM to about 50 μM (e.g., about 5 μM, about 10 μM, about 20 μM, or about 40 μM) of the topsentin alkaloid or derivative thereof (e.g., Homocarbonyltopsentin).

[0085] In some aspects, the host cell is contacted with the topsentin alkaloid or derivative thereof (e.g., Homocarbonyltopsentin) for at least 24 hours, at least 48 hours, or at least 72 hours (e.g., 24 to 90 hours).

[0086] In some aspects, the method further comprises harvesting the AAV particle.

[0087] In some aspects, the method further comprises purifying the harvested AAV particle.

BRIEF DESCRIPTION OF FIGURES

[0088] **FIGs. 1A-1B** show an exemplary schematic of a riboswitch, both with (FIG. 1B) and without (FIG. 1A) a cleavage site between the downstream exon and the heterologous nucleic acid sequence (GFP).

[0089] **FIG. 2** shows the effect of homocarbonyltopsentin (Drug H) and kinetin (Drug K) on alternatively spliced exons in HEK293T cells as measured by mRNA-seq.

[0090] **FIGs. 3A-3C** show the effect of kinetin as measured by RNA-seq on alternatively spliced exons in three genes: TRPV3 (FIG. 3A), STK31 (FIG. 3B), and USP25 (FIG. 3C).

[0091] **FIGs. 4A-4D** show the effect of homocarbonyltopsentin as measured by RNA-seq on alternatively spliced exons in four genes: IK (FIG. 4A), TEX14 (FIG. 4B), GTF2B (FIG. 4C), and LYPD3 (FIG. 4D).

[0092] **FIGs. 5A-5C** show immunofluorescent staining in cells transfected with USP25 Full (FIG. 5A), STK31 Full (FIG. 5B) or TRPV3 mini (FIG. 5C) after treatment with DMSO, Low Drug K, or High Drug K.

[0093] **FIGs. 6A-6D** show immunofluorescent staining in cells transfected with LYPD3 Full (FIG. 6A), TEX14 Full (FIG. 6B), GTF2B Mini (FIG. 6C), or IK Mini (FIG. 6D) after treatment with DMSO, Low Drug H, or High Drug H.

[0094] **FIGs. 7A-7E** show immunofluorescent staining (FIGs. 7A-7D) or qPCR measurements (FIG. 7E) of Drug K or Drug H driven splicing. GTF2B Minimal Reporter (FIG. 7A – Drug H), IK Minimal Reporter (FIG. 7B – Drug H), TRPV3 Minimal Reporter (FIG. 7C), and USP25 Minimal Reporter (FIG. 7D) were analyzed. Fold change comparing Drug H or Drug K treated samples with vehicle was analyzed by qPCR in FIG. 7E for cells transfected with Gene G (GTF2B Minimal Reporter), Gene I (IK Minimal Reporter), Gene T (TRPV3 Minimal Reporter), and Gene U (USP25 Minimal Reporter).

[0095] **FIGs. 8A-8H** show immunofluorescent staining in triplicate for DMSO, 10 μM Drug H, or 40 μM Drug H for cells transfected with GTF2B Full Reporter (FIGs. 8A-8B), GTF2B Mini Reporter (FIGs. 8C-8D), IK Mini Reporter (FIGs. 8E-8F), or LYPD3 Mini Reporter (FIGs. 8G-8H).

[0096] **FIG. 9** shows an RNA PCR splicing assay in cells transfected with LYPD3 Mini Reporter, IK Mini Reporter, and GTF2B Full Reporter after increasing Drug H treatment.

[0097] **FIGs. 10A-10C** show immunofluorescent staining 24 hours after treatment with DMSO, Low Drug H (10 μM), or High Drug H (40 μM) for cells transfected with Full Gene L (LYPD3) (FIG. 10A), Minimum Gene L (version 1 – 779 bp) (FIG. 10B), or Minimum Exon/Intron Gene L (version 2 – 414 bp) (FIG. 10C).

[0098] **FIGs. 11A-11G** show immunofluorescent staining 48 hours after treatment with DMSO, Low Drug K (50 μM), High Drug K (200 μM), Low Drug H (10 μM), or High Drug H (40 μM) for cells transfected with constitutively expressed GFP (FIG. 11A), Full Gene L (LYPD3) (FIG. 11B), Minimum Gene L (version 1 – 779 BP) (FIG. 11C), Minimum Exon/Intron Gene L (version 2 – 414 bp) (FIG. 11D), Full Gene T (TRPV3) (FIG. 11E), Minimum Gene T (version 1 – 906 bp) (FIG. 11F), or Minimum Exon/Intron Gene T (version 2 – 410 bp) (FIG. 11G).

[0099] **FIGs. 12A-12C** show immunofluorescent staining and mRNA fold change after treatment with vehicle, Low Drug H (10 μM), or High Drug H (40 μM) for cells transfected with constitutively expressed GFP (FIG. 12A) or a KT Splice – Drug H – GFP construct (FIG. 12B). FIG. 12C shows mRNA fold change of GFP expression relative to Actin and GAPDH.

[0100] **FIGs. 13A-13D** show immunofluorescent staining and mRNA fold change after treatment with vehicle, Low Drug H (10 μM), or High Drug H (40 μM) for cells transfected with constitutively expressed iRFP (FIG. 13A) a KT Splice – Drug H – iRFP construct (FIG. 13B), or a KT Splice – Constitutively Expressed Exon (FIG. 13C), which does not have a splicing requirement (a control for similar expression to constitutively expressed iRFP). FIG. 13D shows mRNA fold change of iRFP expression relative to Actin and GAPDH.

[0101] **FIG. 14** shows immunofluorescent staining of cells after transfection with inducible green fluorescent protein (GFP) construct "V1A" or "V2A" and treatment with vehicle or with small molecule "Drug 1" (Drug H).

[0102] **FIG. 15** shows ITR titer after triple transfection of "WT Rep" or inducible Rep78/52 constructs ("V2A" or "V2B") after treatment with Drug K ("Drug 2") or Drug H ("Drug 1").

[0103] **FIG. 16** shows fold change of induction of capsid titer after triple transfection of inducible Rep78/52 after treatment with Drug K ("Drug 2") or after treatment with Drug H ("Drug 1").

[0104] **FIGs. 17A-17B** show a western blot of AAV Rep78 levels (FIG. 17A) or VP1, VP2, and VP3 levels (FIG. 17B) after transfection with Drug K ("Drug 2") or Drug H ("Drug 1").

[0105] **FIG. 18** shows an exemplary scheme for the inducible Rep system in comparison to the Triple Transfection system.

[0106] **FIGs. 19A-19B** show ITR titer after triple transfection of "WT Rep" after treatment with Drug K (FIG. 19A) or Drug H (FIG. 19B).

[0107] **FIGs. 20A-20B** show fold change of induction of capsid titer after triple transfection of WT Rep. FIG. 20A shows capsid titer after treatment with Drug K. FIG. 20B shows capsid titer after treatment with Drug H.

[0108] **FIGs. 21A-21B** show a western blot of AAV VP1, VP2, and VP3 levels after treatment with Drug K (FIG. 21A) or Drug H (FIG. 21B).

[0109] **FIG. 22A** shows the expression of Cap mRNA expression after treatment with Drug H (5 μM, 10 μM, 20 μM, and 40 μM) after 0 hours, 24 hours, and 48 hours.

[0110] **FIG. 22B** shows the expression of Rep mRNA expression after treatment with Drug H (5 μM, 10 μM, 20 μM, and 40 μM) at 0 hour, 24 hours, and 48 hours.

[0111] **FIGs. 23A-23H** show changes in the transcriptome after treatment with Drug H at 24 hours (FIGs. 23A-23D) and 48 hours (FIGs. 23E-23H). Concentrations of Drug H analyzed include 5 μM (FIG. 23A and FIG. 23E), 10 μM (FIG. 23B and FIG. 23F), 20 μM (FIG. 23C and FIG. 23G), and 40 μM (FIG. 23D and FIG. 23H).

[0112] **FIG. 24** shows a pathway analysis of differentially expressed genes driven by Drug H treatment, comparing triple transfection with Drug H relative to triple transfection without Drug H.

[0113] **FIG. 25A** shows a heat map of a gene set involved in the inflammatory response after treatment with Drug H for 24 hours or 48 hours.

[0114] **FIG. 25B** shows a heat map of a gene set involved in the immediate early response after treatment with Drug H for 24 hours or 48 hours.

[0115] **FIG. 25C** shows a heat map of a gene set involved in the Defense to Virus pathway after treatment with Drug H for 24 hours or 48 hours.

[0116] **FIGs. 25D-25G** show fold change of genes involved in the Defense to Virus pathway after treatment with Drug H for 24 hours or 48 hours. Genes analyzed include BST2 (FIG. 25D), EXOSC5 (FIG. 25E), CGAS (FIG. 25F), and GBP1 (FIG. 25G).

[0117] **FIG. 26A** shows a heat map of a gene set involved in nucleosome assembly genes after treatment with Drug H for 24 hours or 48 hours.

[0118] **FIGs. 26B-26E** show fold change of genes involved in nucleosome assembly after treatment with Drug H for 24 hours or 48 hours. Genes analyzed include H1-3 (FIG. 26B), H2BC8 (FIG. 26C), H3C1 (FIG. 26D), and H4C3 (FIG. 26E).

[0119] **FIG. 27A** shows a heat map of a gene set involved in the TNF signaling pathway after treatment with Drug H for 24 hours or 48 hours.

[0120] **FIGs. 27B-27C** show normalized gene counts of TNFAIP3 (FIG. 27B) and FOS (FIG. 27C) after treatment with Drug H at 0 hours, 24 hours, and 48 hours.

[0121] **FIG. 28A** shows expression of genes within the JAK/STAT signaling pathway 48 hours after treatment with Drug H.

[0122] **FIG. 28B** shows expression of genes within the IL-17 and NF-κB signaling pathways 48 hours after treatment with Drug H.

[0123] **FIG. 28C** shows expression of genes in parallel and downstream pathways to IL-17 signaling 48 hours after treatment with Drug H (e.g. P38 MAPK signaling and PI3K-Akt signaling pathways).

[0124] **FIG. 28D** shows expression of genes in pathways that converge on apoptosis 48 hours after treatment with Drug H.

[0125] **FIG. 28E** shows expression of genes within the TNF signaling pathway after treatment with Drug H.

[0126] **FIG. 29** shows an exemplary scheme for the Mode of Action for Drug H on host cells during AAV production.

[0127] **FIGs. 30A-30C** shows the epifluorescence of non-transduced cells (FIG. 30A) and cells transduced with AAV2 packaged Positive Control Construct carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter (FIG. 30B and 30C) in HEK293 cells 24 hours after treatment with the amount indicated of DMSO, Drug H (Homocarb), or Drug K (Kinetin). Multiplicity of infection of each AAV construct was either 1e5 or 5e5 vector genomes per cell as indicated in figure title.

[0128] **FIGs. 31A-31D** shows the epifluorescence of cells transduced with AAV2 packaged TRPV3 construct #3 (FIG. 31A and 31B), or cells transduced with AAV2 packaged LYPD3 construct #3 (FIG. 31C and 31D), each carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter, in HEK293 cells 24 hours after treatment with the amount indicated of DMSO, Drug H (Homocarb), or Drug K (Kinetin). Multiplicity of infection of each AAV construct was either 1e5 or 5e5 vector genomes per cell as indicated in figure title.

[0129] **FIGs. 32A-32C** shows the epifluorescence of non-transduced cells (FIG. 32A), cells transduced with AAV2 packaged Positive Control Construct (FIG. 32B and 32C) carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter in Neuro2A cells 24 hours after treatment with the amount indicated of DMSO, Drug H (Homocarb), or Drug K (Kinetin). Multiplicity of infection of each AAV construct was either 1e5 or 5e5 vector genomes per cell as indicated in figure title.

[0130] **FIGs. 33A-33D** shows the epifluorescence of cells transduced with AAV2 packaged TRPV3 construct #3 (FIG. 33A and 33B), or cells transduced with AAV2 packaged LYPD3 construct #3 (FIG. 33C and 33D), each carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter, in Neuro2A cells 24 hours after treatment with the amount indicated of DMSO, Drug H (Homocarb), or Drug K (Kinetin). Multiplicity of infection of each AAV construct was either 1e5 or 5e5 vector genomes per cell as indicated in figure title.

[0131] **FIGs. 34A-34C** shows the epifluorescence of non-transduced cells, cell transduced with AAV2 packaged Positive Control Construct (FIG. 34A), cells transduced with AAV2 packaged TRPV3 construct #3 (FIG. 34B and 34C), each carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter, in Neuro2A cells 24 hours after treatment with the amount indicated of DMSO or Drug K (Kinetin).

Multiplicity of infection of each AAV construct was either 1e5 or 5e5 vector genomes per cell as indicated in figure title.

[0132] **FIGs. 35A-35B** shows the epifluorescence of cells transduced with AAV2 packaged LYPD3 construct #3 (FIG. 35A and 35B) carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter in Neuro2A cells 24 hours after treatment with the amount indicated of DMSO or Drug H (Homocarb). Multiplicity of infection of each AAV construct was either 1e5 or 5e5 vector genomes per cell as indicated in figure title.

[0133] **FIGs. 36A-36B** shows the epifluorescence of non-transduced cells (FIG. 36A), cell transduced with AAV2 packaged Positive Control Construct (FIG. 36B) carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter in Neuro2A cells 24 hours after treatment with the amount indicated of DMSO, Drug H (Homocarb), or Drug K (Kinetin). Multiplicity of infection of each AAV construct was either 1e5 or 5e5 vector genomes per cell as indicated in figure title.

[0134] **FIGs. 37A-37D** show the epifluorescence cells transduced with AAV2 packaged TRPV3 construct #3 (FIG. 37A and 37B), or cells transduced with AAV2 packaged LYPD3 construct #3 (FIG. 37C and 37D) carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter in Neuro2A cells 24 hours after treatment with the amount indicated of DMSO, Drug H (Homocarb), or Drug K (Kinetin). Multiplicity of infection of each AAV construct was either 1e5 or 5e5 vector genomes per cell as indicated in figure title.

[0135] **FIGs. 38A-38B** shows the quantification of the DNA of each vector in mouse livers. Each dot represents the vector genomes per nanogram of total DNA quantified by a standard curve of input DNA. Positive control represents AAV2 packaged Positive Control Construct, TRPV3 represents AAV2 packaged TRPV3 Construct #3, and LYPD3 represents AAV2 packaged LYPD3 Construct #3, each carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter, at 1e11 vector genomes per animal (FIG. 38A) and 1e12 vector genomes per animal (FIG. 38B).

[0136] **FIGs. 39A-39F** show the effect of homocarbonyltopsentin (Drug H) and kinetin (Drug K) measured by qPCR on the alternatively spliced exon of either AAV2 packaged TRPV3 Construct #3 or AAV2 packaged LYPD3 Construct #3, carrying an expression construct including enhanced GFP (eGFP) operably linked aRous Sarcoma Virus (RSV) promoter, in mouse tissues. The relative expression of the alternatively spliced exon was compared to the expression of AAV2

packaged Positive Control Construct for TRPV3 Construct #3 of RNA from mouse liver (FIG 39A-39B) or for LYPD3 Construct #3 of RNA from mouse liver (FIG 39C-39D) or mouse heart (FIG 39E-39F). Stars indicate p-value of less than 0.05 as determined by one-way ANOVA.

DETAILED DESCRIPTION OF THE DISCLOSURE

[0137] In some aspects, provided herein is an expression cassette comprising from 5' to 3': (i) promoter, (ii) an upstream exon, (iii) an upstream intron, (iv) an alternatively spliced exon comprising a translation initiation sequence, (v) a downstream intron, (vi) a downstream exon, and (vii) a heterologous nucleic acid sequence. In some aspects, the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene (reference or endogenous gene) selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0138] In some aspects, the upstream exon and the downstream exon, each comprise a portion of a gene (reference or endogenous gene) selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0139] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the a downstream exon each comprise a portion of a gene (reference or endogenous gene) selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0140] In some aspects, the upstream exon and the downstream exon each comprise a portion of a gene (reference or endogenous gene), which is different from the gene for the portion of the upstream intron, the alternatively spliced exon, and the downstream intron.

[0141] In some aspects, the upstream exon and the downstream exon comprise synthetic exons.

[0142] In some aspects, the portion of the gene comprises fragments or shortened regions of a reference or endogenous gene (e.g., upstream exon of the gene, upstream intron of the gene, an alternatively spliced exon comprising a translation initiation sequence of the gene, downstream intron of the gene, and/or downstream exon of the gene) or regions missing sections of the endogenous gene. In some aspects, the reference or endogenous gene is selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0143] In some aspects, the portion of the gene comprises a truncated region of the gene (e.g., upstream exon of the gene, upstream intron of the gene, an alternatively spliced exon comprising a translation initiation sequence of the gene, downstream intron of the gene, and/or downstream exon of the gene).

[0144] In some aspects, the portion of the gene (e.g., a truncated portion) comprises two or more portions of a region of the gene (e.g., upstream intron but missing an intervening sequence contained in the endogenous upstream intron or downstream intron but missing an intervening sequence contained in the endogenous downstream intron).

[0145] In some aspects, the portion of the gene comprises 15 to 400 basepairs. In some aspects, the portion of the gene comprises less than 400 basepairs, less than 350 basepairs, or less than 300 basepairs. In some aspects, the portion of the gene comprises at least 15 basepairs, at least 18 basepairs, at least 21 basepairs, at least 24 basepairs, at least 27 basepairs, at least 30 basepairs, at least 45 basepairs, at least 60 basepairs, at least 75 basepairs, or at least 90 basepairs.

[0146] In some aspects, the portion of the gene is 15 to 400 basepairs. In some aspects, the portion of the gene is less than 400 basepairs, less than 350 basepairs, or less than 300 basepairs. In some aspects, the portion of the gene is at least 15 basepairs, at least 18 basepairs, at least 21 basepairs, at least 24 basepairs, at least 27 basepairs, at least 30 basepairs, at least 45 basepairs, at least 60 basepairs, at least 75 basepairs, or at least 90 basepairs.

[0147] In some aspects, the expression cassette comprises (i) a promoter; (ii) a truncated upstream exon comprising the 3' end of the endogenous upstream exon; (iii) a truncated upstream intron comprising the 5' and 3' end of the endogenous upstream intron but missing an intervening

sequence contained in the endogenous upstream intron; (iv) an alternatively spliced exon comprising a translation initiation sequence, (v) a truncated downstream intron comprising the 5' and 3' end of the endogenous downstream intron but missing an intervening sequence contained in the endogenous downstream intron, (vi) a truncated downstream exon comprising the 5' end of the endogenous downstream exon, and (vii) a heterologous nucleic acid sequence.

[0148] In some aspects, provided herein is a delivery vector comprising any of the expression cassettes disclosed herein.

[0149] In some aspects, provided herein is a cell comprising any of the expression cassettes disclosed herein or any of the delivery vectors disclosed herein.

[0150] In some aspects, provided herein is a pharmaceutical composition comprising any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, or any of the cells disclosed herein.

[0151] In some aspects, provided herein is a kit comprising i) any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the pharmaceutical compositions disclosed herein; and (ii) optionally instructions for use.

[0152] In some aspects, provided herein is a method of regulating expression of a protein in a mammalian cell comprising administering to the cell any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, or any of the pharmaceutical compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0153] In some aspects, provided herein is a method of expressing a protein in a subject comprising administering to the subject any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the pharmaceutical compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0154] In some aspects, provided herein is a method of treating a disease in a mammal comprising administering to a subject any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the pharmaceutical compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0155] In some aspects, provided herein is a method of regulating transgene expression post-transcriptionally, comprising administering to a subject (a) a nucleic acid construct

comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) a heterologous nucleic acid sequence; and (b) a splicing modulator. In some aspect, the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene. In some aspects, the splicing modulator is Kinetin or Homocarbonyltopsentin.

[0156] In some aspects, provided herein is a method of producing an adeno-associated viral (AAV) vector comprising: (a) introducing a first plasmid comprising an adenovirus helper gene and a second plasmid comprising a heterologous nucleic acid sequence into a host cell; and (b) culturing the host cell in the presence of a splicing modulator to produce AAV particles. In some aspects, the host cell comprises an AAV capsid (Cap) gene and a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV replication (Rep) gene. In some aspects, the host cell comprises an AAV replication (Rep) gene and a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV capsid (Cap) gene. In some aspects, the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0157] In some aspects, provided herein is a method of producing an AAV particle comprising contacting a host cell comprising rep genes, cap genes, and adenoviral helper genes with a topsentin alkaloid or derivative thereof (e.g., Homocarbonyltopsentin).

I. Definitions

[0158] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. In case of conflict, the present application including the definitions will control. Unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. All publications, patents and other references mentioned herein are incorporated by reference in their entireties for all purposes as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

[0159] Although methods and materials similar or equivalent to those described herein can be used in practice or testing of the present disclosure, suitable methods and materials are described below. The materials, methods and examples are illustrative only and are not intended to be limiting. Other features and advantages of the disclosure will be apparent from the detailed description and from the claims.

[0160] In order that the present disclosure can be more readily understood, certain terms are first defined. Additional definitions are set forth throughout the detailed disclosure.

[0161] It is to be noted that the term "a" or "an" entity refers to one or more of that entity; for example, "a nucleic acid sequence," is understood to represent one or more nucleic acid sequences, unless stated otherwise. As such, the terms "a" (or "an"), "one or more," and "at least one" can be used interchangeably herein.

[0162] Furthermore, "and/or", where used herein, is to be taken as specific disclosure of each of the two specified features or components with or without the other. Thus, the term "and/or" as used in a phrase such as "A and/or B" herein is intended to include "A and B," "A or B," "A" (alone), and "B" (alone). Likewise, the term "and/or" as used in a phrase such as "A, B, and/or C" is intended to encompass each of the following aspects: A, B, and C; A, B, or C; A or C; A or B; B or C; A and C; A and B; B and C; A (alone); B (alone); and C (alone).

[0163] It is understood that wherever aspects are described herein with the language "comprising," otherwise analogous aspects described in terms of "consisting of" and/or "consisting essentially of" are also provided.

[0164] The term "about" is used herein to mean approximately, roughly, around, or in the regions of. When the term "about" is used in conjunction with a numerical range, it modifies that range by extending the boundaries above and below the numerical values set forth. In general, the term "about" can modify a numerical value above and below the stated value by a variance of, e.g., 10 percent, up or down (higher or lower).

[0165] The term "at least" prior to a number or series of numbers is understood to include the number adjacent to the term "at least," and all subsequent numbers or integers that could logically be included, as clear from context. For example, the number of nucleotides in a nucleic acid molecule must be an integer. For example, "at least 18 nucleotides of a 21- nucleotide nucleic acid molecule" means that 18, 19, 20, or 21 nucleotides have the indicated property. When at least is present before a series of numbers or a range, it is understood that "at least" can modify each of the numbers in the series or range. "At least" is also not limited to integers (e.g., "at least 5%" includes 5.0%, 5.1%, 5.18% without consideration of the number of significant figures).

[0166] As used herein, "no more than" or "less than" is understood as the value adjacent to the phrase and logical lower values or integers, as logical from context, to zero. When "no more than" is present before a series of numbers or a range, it is understood that "no more than" can modify each of the numbers in the series or range.

[0167] Throughout this disclosure, various aspects of this invention are presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible sub-ranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed sub-ranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6, etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range. Numeric ranges recited are inclusive of the numbers defining the range and include each integer within the defined range.

[0168] Units, prefixes, and symbols are denoted in their Système International de Unites (SI) accepted form. Numeric ranges are inclusive of the numbers defining the range. Where a range of values is recited, it is to be understood that each intervening integer value, and each fraction thereof, between the recited upper and lower limits of that range is also specifically disclosed, along with each subrange between such values. The upper and lower limits of any range can independently be included in or excluded from the range, and each range where either, neither or both limits are included is also encompassed within the disclosure. Thus, ranges recited herein are understood to be shorthand for all of the values within the range, inclusive of the recited endpoints. For example, a range of 1 to 10 is understood to include any number, combination of numbers, or sub-range from the group consisting of 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10.

[0169] Where a value is explicitly recited, it is to be understood that values which are about the same quantity or amount as the recited value are also within the scope of the disclosure. Where a combination is disclosed, each subcombination of the elements of that combination is also specifically disclosed and is within the scope of the disclosure. Conversely, where different elements or groups of elements are individually disclosed, combinations thereof are also disclosed. Where any element of a disclosure is disclosed as having a plurality of alternatives, examples of that disclosure in which each alternative is excluded singly or in any combination with the other alternatives are also hereby disclosed; more than one element of a disclosure can have such exclusions, and all combinations of elements having such exclusions are hereby disclosed.

[0170] The term "pharmaceutical composition," as used herein, represents a composition comprising a compound or molecule described herein, e.g., a vector disclosed herein, formulated with a pharmaceutically acceptable excipient, and can be manufactured or sold with the approval of a governmental regulatory agency as part of a therapeutic regimen for the treatment of disease in a mammal.

[0171] A "pharmaceutically acceptable excipient," as used herein, refers to any ingredient other than the compounds described herein (for example, a vehicle capable of suspending or dissolving the active compound) and having the properties of being substantially nontoxic and non-inflammatory in a patient.

[0172] The terms "operatively linked," "operatively inserted," "operatively positioned," "under control" or "under transcriptional control" means that the promoter is in the correct location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the gene. The term "operably linked" means that a DNA sequence and a regulatory sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins) are bound to the regulatory sequence(s).

[0173] "Nucleic acid," "polynucleotide," and "oligonucleotide," are used interchangeably in the present application. These terms refer only to the primary structure of the molecule. Thus, these terms include double- and single-stranded DNA, as well as double- and single-stranded RNA. The terms "nucleic acid," "polynucleotide," and "oligonucleotide," as used herein, are defined as it is generally understood by the skilled person as a molecule comprising two or more covalently linked nucleosides. Such covalently bound nucleosides can also be referred to as nucleic acid molecules or oligomers. Polynucleotides can be made recombinantly, enzymatically, or synthetically, e.g., by solid-phase chemical synthesis followed by purification. When referring to

a sequence of the polynucleotide or nucleic acid, reference is made to the sequence or order of nucleobase moieties, or modifications thereof, of the covalently linked nucleotides or nucleosides.

[0174] The term "mRNA" as used herein, refers to a single stranded RNA that encodes the amino acid sequence of one or more polypeptide chains.

[0175] A "coding sequence" or a sequence "encoding" a particular molecule (e.g., a therapeutic protein or peptide) is a nucleic acid that is transcribed (in the case of DNA) or translated (in the case of mRNA) into polypeptide, *in vitro* or *in vivo*, when operably linked to an appropriate regulatory sequence, such as a promoter. The boundaries of the coding sequence are determined by a start codon at the 5′ (amino) terminus and a translation stop codon at the 3′ (carboxy) terminus. A coding sequence can include, but is not limited to, cDNA from prokaryotic or eukaryotic mRNA, genomic DNA sequences from prokaryotic or eukaryotic DNA, and synthetic DNA sequences. A transcription termination sequence will usually be located 3′ to the coding sequence.

[0176] As used herein, the term "polypeptide" is intended to encompass a singular "polypeptide" as well as plural "polypeptides," and comprises any chain or chains of two or more amino acids. Thus, as used herein, a "peptide," a "peptide subunit," a "protein," an "amino acid chain," an "amino acid sequence," or any other term used to refer to a chain or chains of two or more amino acids, are included in the definition of a "polypeptide," even though each of these terms can have a more specific meaning. The term "polypeptide" can be used instead of, or interchangeably with any of these terms. The term further includes polypeptides which have undergone post-translational or post-synthesis modifications, for example, conjugation of a palmitoyl group, glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, or modification by non-naturally occurring amino acids. The term "peptide," as used herein encompasses full length peptides and fragments, variants or derivatives thereof. A "peptide" as disclosed herein, can be part of a fusion polypeptide comprising additional components such as, e.g., an Fc domain or an albumin domain, to increase half-life. A peptide as described herein can also be derivatized in a number of different ways. A peptide described herein can comprise modifications including e.g., conjugation of a palmitoyl group.

[0177] "Percent (%) sequence identity" or "Percent (%) identity" or "sequence identity" with respect to a reference polynucleotide or polypeptide sequence is defined as the percentage of nucleic acids or amino acids in a candidate sequence that are identical to the nucleic acids or amino acids in the reference polynucleotide or polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for

purposes of determining percent nucleic acid or amino acid sequence identity can be achieved in various ways that are within the capabilities of one of skill in the art, for example, using publicly available computer software such as BLAST, BLAST-2, or Megalign software. Those skilled in the art can determine appropriate parameters for aligning sequences, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For example, percent sequence identity values can be generated using the sequence comparison computer program BLAST.

[0178] As used herein, the term "promoter" refers to a DNA sequence recognized by the machinery of the cell, or introduced synthetic machinery, that can initiate the transcription of a nucleic acid (e.g., a gene of interest). The term "promoter" can also encompass those nucleic acid elements sufficient for promoter-dependent gene expression controllable for cell-type specific, tissue-specific or inducible by external signals or agents; such elements can be located in the 5′ or 3′ regions of the native gene. In some aspects, the promoter is a constitutive promoter, a cell-type specific promoter, or an inducible promoter.

[0179] As used herein, the term "enhancer" is a cis-acting element that stimulates or inhibits transcription of adjacent genes. An enhancer that inhibits transcription is also referred to as a "silencer." Enhancers can function (e.g., can be associated with a coding sequence) in either orientation, over distances of up to several kilobase pairs (kb) from the coding sequence and from a position downstream of a transcribed region.

[0180] The term "expression vector", "expression construct", or "expression cassette" means any type of genetic construct containing a nucleic acid in which part or all of the nucleic acid encoding sequence is capable of being transcribed.

[0181] As used herein, the term "delivery vector" or "vector" refers to any vehicle for the cloning of and/or transfer of a nucleic acid into a host cell, such as a plasmid, phage, transposon, cosmid, chromosome, artificial chromosome, virus, virion, etc. A vector can be a replicon to which another nucleic acid segment can be attached so as to bring about the replication of the attached segment. A "replicon" refers to any genetic element (e.g., plasmid, phage, cosmid, chromosome, virus) that functions as an autonomous unit of replication *in vivo*, i.e., capable of replication under its own control. The term "delivery vector" or "vector" includes both viral and nonviral vehicles for introducing the nucleic acid into a cell *in vitro*, *ex vivo* or *in vivo*. A large number of vectors are known and used in the art including, for example, plasmids, modified eukaryotic viruses, or modified bacterial viruses. In some aspects, insertion of a polynucleotide into a suitable vector can be accomplished by ligating the appropriate polynucleotide fragments into a chosen vector that has

complementary cohesive termini. Vectors can be engineered to encode selectable markers or reporters that provide for the selection or identification of cells that have incorporated the vector. Expression of selectable markers or reporters allows identification and/or selection of host cells that incorporate and express other coding regions contained on the vector. Examples of selectable marker genes known and used in the art include: genes providing resistance to ampicillin, streptomycin, gentamycin, kanamycin, hygromycin, bialaphos herbicide, sulfonamide, and the like; and genes that are used as phenotypic markers, i.e., anthocyanin regulatory genes, isopentanyl transferase gene, and the like. Examples of reporters known and used in the art include: luciferase (Luc), green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), β-galactosidase (LacZ), β-glucuronidase (Gus), and the like. Selectable markers can also be considered to be reporters. In some aspects, the delivery vector is selected from the group consisting of a viral vector (e.g., an AAV vector), a plasmid, a lipid, a protein particle, a bacterial vector, and a lysosome.

[0182] Some aspects of the disclosure are directed to biological vectors, which can include viruses, particularly attenuated and/or replication-deficient viruses.

[0183] A "viral vector" refers to a sequence that comprises one or more polynucleotide regions encoding or comprising a molecule of interest, e.g., a protein, a peptide, and an oligonucleotide or a plurality thereof. Viral vectors can be used to deliver genetic materials into cells. Viral vectors can be modified for specific applications. In some aspects, the delivery vector of the disclosure is a viral vector selected from the group consisting of an adeno-associated viral (AAV) vector, an adenoviral vector, a lentiviral vector, or a retroviral vector.

[0184] The term "adeno-associated virus vector" or "AAV vector" or "adeno-associated viral vector" as used herein refers to any vector that comprises or derives from components of an adeno-associated vector and is suitable to infect mammalian cells, preferably human cells. The term AAV vector typically designates an AAV-type viral particle or virion comprising a payload. The AAV vector can be derived from various serotypes, including combinations of serotypes (i.e., "pseudotyped" AAV) or from various genomes (e.g., single stranded or self-complementary). In addition, the AAV vector can be replication defective and/or targeted. As used herein, the term "adeno-associated virus" (AAV), includes but is not limited to, AAV type 1, AAV type 2, AAV type 3 (including types 3A and 3B), AAV type 4, AAV type 5, AAV type 6, AAV type 7, AAV type 8, AAV type 9, AAV type 10, AAV type 11, AAV type 12, AAV type 13, AAVrh8, AAVrh10, AAVrh.74, snake AAV, avian AAV, bovine AAV, canine AAV, equine AAV, ovine AAV, goat AAV, shrimp AAV, those AAV serotypes and clades disclosed by Gao et al. (J. Virol. 78:6381 (2004)) and Moris et al. (Virol. 33:375 (2004)), and any other AAV now known or later discovered.

See, e.g., FIELDS et al. VIROLOGY, volume 2, chapter 69 (4th ed., Lippincott-Raven Publishers). In some aspects, an "AAV vector" includes a derivative of a known AAV vector. In some aspects, an "AAV vector" includes a modified or an artificial AAV vector. In some aspects, an AAV vector includes a hybrid vector (e.g., AAV-DJ, AAV-PHP.B, AAV2-ESGHGYF, AAVM41, AAV-LK03, AAV2-BR1, AAV587MTP, AAV-Anc80L65, AAV2-7m8, AAV2HBKO, AAV2YF, AAV6-RGD or AAV6.2). In some aspects, the terms "AAV genome" and "AAV vector" can be used interchangeably. In some aspects, the AAV vector is modified relative to the wild-type AAV serotype sequence. In some aspects, the modified AAV vector is a modified AAV6, e.g., an AAV6 vector comprising the RGD peptide (an AAV6-RGD vector) or an AAV6 vector comprising mutations of surface exposed tyrosine residues as described, for example, in Sayroo et al. Gene Ther. 2016 Jan; 23(1):18-25. In some aspects, the AAV6-RGD vector further comprises modified amino acids corresponding to Y705, Y731, T492, and K531, (e.g., Y705, Y731F, T492V, and K531E; also referred to as AAV-RGD-Y705-731F+T492V+K531E).

[0185] As used herein, a "recombinant AAV particle", "recombinant AAV vector", "rAAV particle", or "rAAV vector" refers to an AAV virus that comprises a capsid protein and a vector genome (or an AAV genome) comprising at least one heterologous polynucleotide encoding a protein of interest and at least one inverted terminal repeat (ITR) region.

[0186] As used herein, the term "*in vitro*" refers to events that occur in an artificial environment, e.g., in a test tube or reaction vessel, in cell culture, in a Petri dish, etc., rather than within an organism (e.g., animal, plant, or microbe).

[0187] As used herein, the term "*in vivo*" refers to events that occur within an organism (e.g., animal, plant, or microbe or cell or tissue thereof).

[0188] As used herein, "transcription" refers to the process of copying a segment of DNA into RNA.

[0189] As used herein, "messenger RNA" or "mRNA" refer to a form of mRNA molecule that only includes the exons that code for the amino acids that make up a protein. Thus, any introns and/or any exons that are not required for the particular protein have been removed.

[0190] As used herein, "mature messenger RNA" or "mature mRNA" refer to a final form of RNA molecule that is transcribed from a single strand of DNA and carries the genetic information needed to make a protein.

[0191] As used herein the term "intron" refers to a non-coding nucleotide portion of a gene sequence that is transcribed into a precursor mRNA (e.g., pre-mRNA) and subsequently removed

by splicing during formation of the mature messenger RNA. Thus, introns are not present post-splicing in the final mature mRNA.

[0192] As used herein the term "exon" refers to a nucleotide portion of a gene or DNA sequence that is transcribed and can form part of a mature messenger RNA derived from the gene or DNA sequence, e.g., after the introns have been removed by RNA splicing.

[0193] As used herein, the terms "splicing" or "splice" refer to the process of removing introns and joining exons from pre-mRNA to generate mRNA. The terms "splice site" and "splice junction" are used interchangeably to refer to a region where a splicing event takes place, for example, at an exon-intron junction in a pre-mRNA molecule. For example, a slice donor site may be present at the 5′ end of an intron and a splice acceptor site may be present at the 3′ end of an intron.

[0194] As used herein, the term "alternative splicing" refers to a process where exons from the same gene are joined in different combinations leading to different but related mRNA. In some aspects, mRNAs can be translated to produce different proteins, e.g., with distinct structures and/or functions from a single gene.

[0195] As used herein, the terms "regulating" or "regulate" refer to controlling the activity of a molecule or biological function, such as enhancing or diminishing the activity or function (e.g., transgene or protein expression).

[0196] As used herein, the term "splicing modulator" refers to a molecule that is capable of enhancing or diminishing splicing. In some aspects, splicing can be regulated by the presence or absence of the splicing modulator, resulting in the presence or absence of particular mRNA variants (e.g., those that contain alternatively spliced exons).

[0197] As used herein, the term "stably integrated" refers to a process wherein a heterologous or foreign gene (e.g., an AAV Rep gene) becomes part of a host cell genome and is copied during replication such that descendants of the host cell also express the heterologous or foreign gene.

[0198] As used herein, the term "a portion of a gene" or "gene portion" refers to a nucleic acid sequence derived from a reference gene (e.g., an endogenous gene), wherein the portion includes a region or fragment of the reference gene. In some aspects, the gene portion can comprise a 5' splice site, a 3' splice site, an exon or fragment thereof, an intron or fragment thereof, or any combination thereof. In some aspects, the portion of the gene can comprise two or more fragments from a region of the gene (e.g., upstream exon of the gene, upstream intron of the gene, an

alternatively spliced exon comprising a translation initiation sequence of the gene, downstream intron of the gene, and/or downstream exon of the gene).

[0199] As used herein, the term "gene" refers to a nucleic acid region, also referred to as a transcribed region, which expresses a polynucleotide, such as an RNA. The transcribed polynucleotide can have a sequence encoding a polypeptide, such as a functional protein, which can be translated into the encoded polypeptide when placed under the control of an appropriate regulatory region. In some aspects, the protein coding sequences a genes can be broken into smaller portions or pieces of coding sequences (exons) separated by non-coding sequences (introns). When such genes are transcribed, those exons and introns can be included in the initial messenger RNA products.

[0200] As used herein, the term "derived from" refers to a component that is isolated from or made using a specified molecule or organism, or information (e.g., amino acid or nucleic acid sequence) from the specified molecule or organism. In some aspects, the derived species can be obtained by, for example, naturally occurring mutagenesis, artificial directed mutagenesis or artificial random mutagenesis.

[0201] As used herein, the term "Drug H" refers to homocarbonyltopsentin.

[0202] As used herein, the term "Drug K" refers to kinetin.

[0203] As used herein, the term "Gene G" refers to the General Transcription Factor IIB gene (i.e., GTF2B).

[0204] As used herein, the term "Gene L" refers to the LY6/PLAUR Domain Containing 3 gene (i.e., LYPD3).

[0205] As used herein, the term "Gene I" refers to the IK cytokine gene (i.e., IK).

[0206] As used herein, the term "Gene T" refers to the Transient Receptor Potential Cation Channel Subfamily V Member 3 gene (i.e., TRPV3).

[0207] As used herein, the term "Gene U" refers to the Ubiquitin Specific Peptidase 25 gene (i.e., USP25).

[0208] As used herein, the term "Gene S" refers to the Serine/Threonine Kinase 31 gene (i.e., STK31).

II. Alternatively Spliced Exons

[0209] Certain aspects of the disclosure are directed to an expression cassette comprising from 5' to 3': (i) promoter, (ii) an upstream exon, (iii) an upstream intron, (iv) an alternatively spliced exon comprising a translation initiation sequence, (v) a downstream intron, (vi) a

downstream exon, and (vii) a heterologous nucleic acid sequence. In some aspects, the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0210] In some aspects, the upstream intron and the downstream intron can each comprise a portion of a gene, which is different from the gene for the portion of the alternatively spliced exon. In some aspects, the upstream intron and the downstream intron can comprise synthetic introns.

[0211] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, and the downstream exon are not in the protein coding sequence of the target gene.

[0212] In some aspects, the expression cassette comprises from 5' to 3': (i) promoter, (ii) an upstream exon, (iii) an upstream intron, (iv) an alternatively spliced exon comprising an expression modulator, (v) a downstream intron, (vi) a downstream exon, and (vii) a heterologous nucleic acid sequence. In some aspects, the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0213] In some aspects, the expression modulator is a translation initiation sequence.

[0214] In some aspects, the upstream exon and the downstream exon, each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0215] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the a downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor

Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0216] In some aspects, the upstream exon and the downstream exon each comprise a portion of a gene, which is different from the gene for the portion of the upstream intron, the alternatively spliced exon, and the downstream intron.

[0217] In some aspects, the upstream exon and the downstream exon comprise synthetic exons.

[0218] In some aspects, the portion of the gene comprises fragments or shortened regions of a reference or endogenous gene (e.g., upstream exon of the gene, upstream intron of the gene, an alternatively spliced exon comprising a translation initiation sequence of the gene, downstream intron of the gene, and/or downstream exon of the gene) or regions missing sections of the endogenous gene. In some aspects, the reference or endogenous gene is selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0219] In some aspects, the portion of the gene comprises a truncated region of the gene (e.g., upstream exon of the gene, upstream intron of the gene, an alternatively spliced exon comprising a translation initiation sequence of the gene, downstream intron of the gene, and/or downstream exon of the gene).

[0220] In some aspects, the portion of the gene (e.g., a truncated portion) comprises two or more portions of a region of the gene (e.g., upstream intron but missing an intervening sequence contained in the endogenous upstream intron or downstream intron but missing an intervening sequence contained in the endogenous downstream intron).

[0221] In some aspects, the portion of the gene comprises 15 to 400 basepairs. In some aspects, the portion of the gene comprises less than 400 basepairs, less than 350 basepairs, or less than 300 basepairs. In some aspects, the portion of the gene comprises at least 15 basepairs, at least 18 basepairs, at least 21 basepairs, at least 24 basepairs, at least 27 basepairs, at least 30 basepairs, at least 45 basepairs, at least 60 basepairs, at least 75 basepairs, or at least 90 basepairs.

[0222] In some aspects, the portion of the gene is 15 to 400 basepairs. In some aspects, the portion of the gene is less than 400 basepairs, less than 350 basepairs, or less than 300 basepairs. In some aspects, the portion of the gene is at least 15 basepairs, at least 18 basepairs, at least 21 basepairs, at least 24 basepairs, at least 27 basepairs, at least 30 basepairs, at least 45 basepairs, at least 60 basepairs, at least 75 basepairs, or at least 90 basepairs.

[0223] In some aspects, the expression cassette comprises

(i) a promoter;

(ii) a truncated upstream exon comprising the 3' end of the endogenous upstream exon;

(iii) a truncated upstream intron comprising the 5' and 3' end of the endogenous upstream intron but missing an intervening sequence contained in the endogenous upstream intron;

(iv) an alternatively spliced exon comprising a translation initiation sequence,

(v) a truncated downstream intron comprising the 5' and 3' end of the endogenous downstream intron but missing an intervening sequence contained in the endogenous downstream intron,

(vi) a truncated downstream exon comprising the 5' end of the endogenous downstream exon, and

(vii) a heterologous nucleic acid sequence.

[0224] In some aspects, the alternatively spliced exon is regulatable (e.g., regulated) by a splicing modulator.

[0225] In some aspects, the splicing modulator is Kinetin or Homocarbonyltopsentin.

[0226] In some aspects, the gene is TRPV3, STK31, or USP25; and the splicing modulator is Kinetin.

[0227] In some aspects, the gene is GTF2B, LYPD3, IK, or TEX14; and the splicing modulator is Homocarbonyltopsentin.

[0228] In some aspects, the splicing modulator promotes inclusion of the alternatively spliced exon thereby driving expression of the heterologous nucleic acid sequence.

[0229] In some aspects, the promoter is selected from the group consisting of a CBA promoter, a smCBA promoter, a CMV promoter, an EF-1a (Elongation Factor 1a) promoter, a RSV (Rous Sarcoma Virus) promoter, an Ubiquitin (UbC) promoter, a CAG promoter, a SV40 (simian vacuolating virus 40) promoter, a PGK (phosphoglycerate kinase) promoter, a human beta actin promoter, a H1 (human polymerase III RNA) promoter, a human U6 small nuclear promoter, tetracycline responsive element, and any combination thereof.

[0230] In some aspects, the expression cassette further comprises an enhancer.

[0231] In some aspects, the enhancer is selected from the group consisting of a CMV enhancer, a SV40 enhancer, and any combination thereof.

[0232] In some aspects, the expression cassette further comprises a cleavage site between the downstream exon and the heterologous nucleic acid sequence.

[0233] In some aspects, the cleavage site is selected from the group consisting of a furin cleavage site, a foot-and-mouth disease virus 18 2A peptide (F2A), an equine rhinitis A virus 2A peptide (E2A), a porcine teschovirus-1 2A peptide (P2A), a thosea asigna virus 2A peptide (T2A), and any combination thereof.

[0234] In some aspects, the expression cassette further comprises a leader sequence between the downstream exon and the heterologous nucleic acid sequence. In some aspects, the leader sequence is selected from the group consisting of a human opticin (hOPT) leader sequence, an interleukin-2 (IL-2) leader sequence, an interleukin-12 (IL-12) leader sequence, a human or mouse interleukin-6 (hIL-6 or mIL-6) leader sequence, a human insulin leader sequence, a human serum albumin (HSA) leader sequence, and any combination thereof.

[0235] In some aspects, the combined length of the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon is between about 300 basepairs to about 2100 basepairs.

[0236] In some aspects, the combined length of the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon is between about 300 basepairs to about 2300 basepairs, about 400 basepairs to about 2300 basepairs, about 500 basepairs to about 2300 basepairs, about 600 basepairs to about 2300 basepairs, about 700 basepairs to about 2300 basepairs, about 800 basepairs to about 2300 basepairs, about 900 basepairs to about 2300 basepairs, about 1000 basepairs to about 2300 basepairs, about 1100 basepairs to about 2300 basepairs, about 1200 basepairs to about 2300 basepairs, about 1300 basepairs to about 2300 basepairs, about 1400 basepairs to about 2300 basepairs, about 1500 basepairs to about 2300 basepairs, about 1600 basepairs to about 2300 basepairs, about 1700 basepairs to about 2300 basepairs, about 1800 basepairs to about 2300 basepairs, about 1900 basepairs to about 2300 basepairs, about 2000 basepairs to about 2300 basepairs, about 2100 basepairs to about 2300 basepairs, about 2200 basepairs to about 2300 basepairs, about 1500 basepairs to about 2200 basepairs, about 1500 basepairs to about 2100 basepairs, about 1500 basepairs to about 2000 basepairs, about 1500 basepairs to about 1900 basepairs, about 1500 basepairs to about 1800 basepairs, about 1500 basepairs to about 1700 basepairs, about 1500 basepairs to about 1600 basepairs, about 1800 basepairs to about 2200 basepairs, about 1700

basepairs to about 2100 basepairs, about 300 basepairs to about 2200 basepairs, about 300 basepairs to about 2100 basepairs, about 300 basepairs to about 2000 basepairs, about 300 basepairs to about 1900 basepairs, about 300 basepairs to about 1800 basepairs, about 300 basepairs to about 1700 basepairs, about 300 basepairs to about 1600 basepairs, or about 1700 basepairs to about 2200 basepairs.

[0237] In some aspects, the combined length of the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon is about 300 basepairs, about 400 basepairs, about 500 basepairs, about 600 basepairs, about 700 basepairs, about 800 basepairs, about 900 basepairs, about 1000 basepairs, about 1100 basepairs, about 1200 basepairs, about 1300 basepairs, about 1400 basepairs, about 1500 basepairs, about 1600 basepairs, about 1700 basepairs, about 1800 basepairs, about 1900 basepairs, about 2000 basepairs, about 2100 basepairs, about 2200 basepairs, or about 2300 basepairs.

[0238] In some aspects, the combined length of the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon is between 300 basepairs to 2300 basepairs, 400 basepairs to 2300 basepairs, 500 basepairs to 2300 basepairs, 600 basepairs to 2300 basepairs, 700 basepairs to 2300 basepairs, 800 basepairs to 2300 basepairs, 900 basepairs to 2300 basepairs, 1000 basepairs to 2300 basepairs, 1100 basepairs to 2300 basepairs, 1200 basepairs to 2300 basepairs, 1300 basepairs to 2300 basepairs, 1400 basepairs to 2300 basepairs, 1500 basepairs to 2300 basepairs, 1600 basepairs to 2300 basepairs, 1700 basepairs to 2300 basepairs, 1800 basepairs to 2300 basepairs, 1900 basepairs to 2300 basepairs, 2000 basepairs to 2300 basepairs, 2100 basepairs to 2300 basepairs, 2200 basepairs to 2300 basepairs, 1500 basepairs to 2200 basepairs, 1500 basepairs to 2100 basepairs, 1500 basepairs to 2000 basepairs, 1500 basepairs to 1900 basepairs, 1500 basepairs to 1800 basepairs, 1500 basepairs to 1700 basepairs, 1500 basepairs to 1600 basepairs, 1800 basepairs to 2200 basepairs, 1700 basepairs to 2100 basepairs, 300 basepairs to 2200 basepairs, 300 basepairs to 2100 basepairs, 300 basepairs to 2000 basepairs, 300 basepairs to 1900 basepairs, 300 basepairs to 1800 basepairs, 300 basepairs to 1700 basepairs, 300 basepairs to 1600 basepairs, or 1700 basepairs to 2200 basepairs.

[0239] In some aspects, the combined length of the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon is 300 basepairs, 400 basepairs, 500 basepairs, 600 basepairs, 700 basepairs, 800 basepairs, 900 basepairs, 1000 basepairs, 1100 basepairs, 1200 basepairs, 1300 basepairs, 1400 basepairs, 1500 basepairs, 1600 basepairs, 1700 basepairs, 1800 basepairs, 1900 basepairs, 2000 basepairs, 2100 basepairs, 2200 basepairs, or 2300 basepairs.

[0240] In some aspects, the length of the heterologous nucleic acid sequence is less than about 4100 basepairs, less than about 4000 basepairs, less than about 3900 basepairs, less than about 3800 basepairs, less than about 3600 basepairs, less than about 3400 basepairs, less than about 3200 basepairs, less than about 3000 basepairs, less than about 2800 basepairs, less than about 2600 basepairs, less than about 2400 basepairs, or less than about 2200 basepairs.

[0241] In some aspects, the length of the heterologous nucleic acid sequence is less than about 2600 basepairs, less than about 2400 basepairs, less than about 2200 basepairs, less than about 2000 basepairs, less than about 1800 basepairs, less than about 1600 basepairs, less than about 1400 basepairs, less than about 1200 basepairs, less than about 1000 basepairs, less than about 800 basepairs, less than about 600 basepairs, less than about 400 basepairs, or less than about 200 basepairs.

[0242] In some aspects, the length of the heterologous nucleic acid sequence is between about 200 basepairs to about 4100 basepairs, between about 400 basepairs to about 4100 basepairs, between about 600 basepairs to about 4100 basepairs, between about 800 basepairs to about 4100 basepairs, between about 1000 basepairs to about 4100 basepairs, between about 1200 basepairs to about 4100 basepairs, between about 1400 basepairs to about 4100 basepairs, between about 1600 basepairs to about 4100 basepairs, between about 1800 basepairs to about 4100 basepairs, between about 2000 basepairs to about 4100 basepairs, between about 2200 basepairs to about 4100 basepairs, between about 2400 basepairs to about 4100 basepairs, between about 2600 basepairs to about 4100 basepairs, between about 2800 basepairs to about 4100 basepairs, between about 3000 basepairs to about 4100 basepairs, between about 3200 basepairs to about 4100 basepairs, between about 3400 basepairs to about 4100 basepairs, between about 3600 basepairs to about 4100 basepairs, between about 3800 basepairs to about 4100 basepairs, between about 4000 basepairs to about 4100 basepairs, about 200 basepairs to about 2400 basepairs, between about 400 basepairs to about 2400 basepairs, between about 600 basepairs to about 2400 basepairs, between about 800 basepairs to about 2400 basepairs, between about 1000 basepairs to about 2400 basepairs, between about 1200 basepairs to about 2400 basepairs, between about 1400 basepairs to about 2400 basepairs, between about 1600 basepairs to about 2400 basepairs, between about 1800 basepairs to about 2400 basepairs, between about 2000 basepairs to about 2400 basepairs, between about 2200 basepairs to about 2400 basepairs, between about 200 basepairs to about 2200 basepairs, between about 200 basepairs to about 2000 basepairs, between about 200 basepairs to about 1800 basepairs, between about 200 basepairs to about 1600 basepairs, between about 200 basepairs to about 1400 basepairs, between about 200 basepairs to about 1200 basepairs, between

about 200 basepairs to about 1000 basepairs, between about 200 basepairs to about 800 basepairs, between about 200 basepairs to about 600 basepairs, between about 200 basepairs to about 400 basepairs, between about 1000 basepairs to about 2000 basepairs, between about 1200 basepairs to about 2000 basepairs, between about 1400 basepairs to about 2000 basepairs, or between about 1600 basepairs to about 2000 basepairs.

[0243] In some aspects, the length of the heterologous nucleic acid sequence is less than 4100 basepairs, less than 4000 basepairs, less about 3900 basepairs, less than 3800 basepairs, less than 3600 basepairs, less than 3400 basepairs, less than 3200 basepairs, less than 3000 basepairs, less than 2800 basepairs, less than 2600 basepairs, less than 2400 basepairs, or less than 2200 basepairs. In some aspects, the length of the payload sequence is less than 2600 basepairs, less than 2400 basepairs, less than 2200 basepairs, less than 2000 basepairs, less than 1800 basepairs, less than 1600 basepairs, less than 1400 basepairs, less than 1200 basepairs, less than 1000 basepairs, less than 800 basepairs, less than 600 basepairs, less than 400 basepairs, or less than 200 basepairs.

[0244] In some aspects, the length of the heterologous nucleic acid sequence is between 200 basepairs to 4100 basepairs, between 400 basepairs to 4100 basepairs, between 600 basepairs to 4100 basepairs, between 800 basepairs to 4100 basepairs, between 1000 basepairs to 4100 basepairs, between 1200 basepairs to 4100 basepairs, between 1400 basepairs to 4100 basepairs, between 1600 basepairs to 4100 basepairs, between 1800 basepairs to 4100 basepairs, between 2000 basepairs to 4100 basepairs, between 2200 basepairs to 4100 basepairs, between 2400 basepairs to 4100 basepairs, between 2600 basepairs to 4100 basepairs, between 2800 basepairs to 4100 basepairs, between 3000 basepairs to 4100 basepairs, between 3200 basepairs to 4100 basepairs, between 3400 basepairs to 4100 basepairs, between 3600 basepairs to 4100 basepairs, between 3800 basepairs to 4100 basepairs, between 4000 basepairs to 4100 basepairs, 200 basepairs to 2400 basepairs, between 400 basepairs to 2400 basepairs, between 600 basepairs to 2400 basepairs, between 800 basepairs to 2400 basepairs, between 1000 basepairs to 2400 basepairs, between 1200 basepairs to 2400 basepairs, between 1400 basepairs to 2400 basepairs, between 1600 basepairs to 2400 basepairs, between 1800 basepairs to 2400 basepairs, between 2000 basepairs to 2400 basepairs, between 2200 basepairs to 2400 basepairs, between 200 basepairs to 2200 basepairs, between 200 basepairs to 2000 basepairs, between 200 basepairs to 1800 basepairs, between 200 basepairs to 1600 basepairs, between 200 basepairs to 1400 basepairs, between 200 basepairs to 1200 basepairs, between 200 basepairs to 1000 basepairs, between 200 basepairs to 800 basepairs, between 200 basepairs to 600 basepairs, between 200 basepairs to 400

basepairs, between 1000 basepairs to 2000 basepairs, between 1200 basepairs to 2000 basepairs, between 1400 basepairs to 2000 basepairs, or between 1600 basepairs to 2000 basepairs.

[0245] In some aspects, the heterologous nucleic acid sequence encodes a protein. In some aspects, the protein is a therapeutic protein.

[0246] In some aspects, the upstream exon comprises a sequence selected from the group consisting of SEQ ID NOs: 8, 13, 18, 23, 28, 33, and 38.

[0247] In some aspects, the upstream exon comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NOs: 8, 13, 18, 23, 28, 33, and 38.

[0248] In some aspects, the upstream exon comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 8, 13, 18, 23, 28, 33, and 38.

[0249] In some aspects, the alternatively spliced exon comprises a sequence selected from the group consisting of SEQ ID NO: 10, 15, 20, 25, 30, 35, and 40.

[0250] In some aspects, the alternatively spliced exon comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 10, 15, 20, 25, 30, 35, and 40.

[0251] In some aspects, the alternatively spliced exon comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 10, 15, 20, 25, 30, 35, and 40.

[0252] In some aspects, the downstream exon comprises a sequence selected from the group consisting of SEQ ID NO: 12, 17, 22, 27, 32, 37, and 42.

[0253] In some aspects, the downstream exon comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 12, 17, 22, 27, 32, 37, and 42.

[0254] In some aspects, the downstream exon comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at

least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 12, 17, 22, 27, 32, 37, and 42.

[0255] In some aspects, the upstream intron comprises a sequence selected from the group consisting of SEQ ID NOs: 9, 14, 19, 24, 29, 34, and 39.

[0256] In some aspects, the upstream intron comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NOs: 9, 14, 19, 24, 29, 34, and 39.

[0257] In some aspects, the upstream intron comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 9, 14, 19, 24, 29, 34, and 39.

[0258] In some aspects, the downstream intron comprises a sequence selected from the group consisting of SEQ ID NO: 11, 16, 21, 26, 31, 36, and 41.

[0259] In some aspects, the downstream intron comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 11, 16, 21, 26, 31, 36, and 41.

[0260] In some aspects, the downstream intron comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 11, 16, 21, 26, 31, 36, and 41.

[0261] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 1-7 or 43-66, or fragments thereof.

[0262] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 1-7 or 43-66, or fragments thereof.

[0263] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 1-7 or 43-66, or fragments thereof.

[0264] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 8-42.

[0265] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 8-42.

[0266] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 8-42.

[0267] In some aspects, the expression cassette comprises the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the downstream exon, wherein each comprises a portion of a gene, wherein the gene is a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene.

[0268] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 1, 8-12, and 43-45.

[0269] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 1, 8-12, and 43-45.

[0270] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 1, 8-12, and 43-45.

[0271] In some aspects, the expression cassette comprises a sequence corresponding to SEQ ID NO: 45.

[0272] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to SEQ ID NO: 45.

[0273] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to SEQ ID NO: 45.

[0274] In some aspects, the expression cassette comprises the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the downstream exon, wherein each comprises a portion of a gene, wherein the gene is a Serine/Threonine Kinase 31 (STK31) gene.

[0275] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 2, 13-17, 46 and 47.

[0276] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 2, 13-17, 46 and 47.

[0277] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 2, 13-17, 46 and 47.

[0278] In some aspects, the expression cassette comprises the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the downstream exon, wherein each comprises a portion of a gene, wherein the gene is an Ubiquitin Specific Peptidase 25 (USP25) gene.

[0279] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 3, 18-22, 48 and 49.

[0280] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 3, 18-22, 48 and 49.

[0281] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 3, 18-22, 48 and 49.

[0282] In some aspects, the expression cassette comprises the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the downstream exon, wherein each comprises a portion of a gene, wherein the gene is a General Transcription Factor IIB (GTF2B) gene.

[0283] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 4, 23-27, 50 and 51.

[0284] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 4, 23-27, 50 and 51.

[0285] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 4, 23-27, 50 and 51.

[0286] In some aspects, the expression cassette comprises the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the downstream exon, wherein each comprises a portion of a gene, wherein the gene is a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0287] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 5, 28-32, and 52-54.

[0288] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 5, 28-32, and 52-54.

[0289] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 5, 28-32, and 52-54.

[0290] In some aspects, the expression cassette comprises a sequence corresponding to SEQ ID NO: 54.

[0291] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to SEQ ID NO: 54.

[0292] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to SEQ ID NO: 54.

[0293] In some aspects, the expression cassette comprises the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the downstream exon, wherein each comprises a portion of a gene, wherein the gene is a IK gene.

[0294] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 6, 33-37, 55 and 56.

[0295] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 6, 33-37, 55 and 56.

[0296] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 6, 33-37, 55 and 56.

[0297] In some aspects, the expression cassette comprises the upstream exon, the upstream intron, the alternatively spliced exon comprising a translation initiation sequence, the downstream intron, the downstream exon, wherein each comprises a portion of a gene, wherein the gene is a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene.

[0298] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NOs: 7, 38-42, 57 and 58.

[0299] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 7, 38-42, 57 and 58.

[0300] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 7, 38-42, 57 and 58.

[0301] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NO: 43 to 66.

[0302] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about

94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 43 to 66.

[0303] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 43 to 66.

[0304] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NO: 59-62.

[0305] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 59-62.

[0306] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 59-62.

[0307] In some aspects, the expression cassette comprises a sequence selected from the group consisting of SEQ ID NO: 63-66.

[0308] In some aspects, the expression cassette comprises a sequence having at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity to any one of SEQ ID NO: 63 to 66.

[0309] In some aspects, the expression cassette comprises a sequence having at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to any one of SEQ ID NO: 63 to 66.

[0310] In some aspects, the expression cassette comprises one or more of any of the sequences in Table 1.

Table 1 – Genomic Sequences

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
TRPV3 (SEQ ID NO: 1)	ttttatccaagc tgtgcttgtgat actgtctgtctt cttgtacttgttt gcctacaaag	gtgcgttgagagctgcgccggcgg gagggccctcggaggatggccatt tttcagatggggaaactgagatctg gagaggaagaggggtttaccaag gggcaggtcactgctctttctcccaa	gtcattttgcatga tgttctgaagttctt gtttgtatatatcgt gtttttgcttggatt tggagtag	gtaattgatttataaataatagta actcccacatttataaacactttt atacagatgtcaaaatggcaa agtgcatggcggccttggcaa caattagaccaactaccagtg	Ccttggcctcgctgatcg agaagtgtcccaaagaca acaaggactgcagctcct acggcagcttcagcgac gcagtgctggaactcttca

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
	agtacctcgcc tgcctcgtgct ggccatggcc ctgggctggg cgaacatgctc tactatacgcg gggtttccagt ccatgggcat gtacagcgtc atgatccagaa g (SEQ ID NO: 8)	atgccatgcactccctttgctttcgg ccaggctttaagtgtcaagaccttaa atcccccaggcttgtggggtgctgg ggaggggtacagtttgtggaggaa tgttctcaccctgggcagtggcgt gctggtaaacatttaacgactgttct gtgtgtgtgtgtgtgtatgtgtgtgtg tgtaggtggaatatgtatattccatat atatacacaaaagacttatcatacatt ttactatataaacgatgtgtatcacac aatttgcaaataatgaaaatatacca tgcagtttattgtaaattccacatagc cagttgagtctcagaataatttgttgc ttttttttttttttttggagaccaagtctt gctctgctgcccaggccagagtaa agtggtgcgatctcggctccctgca acctctgcctcctgggttcaagcagt tctcctgcctcagcctccagagtag ctgggatcacaggcacgcaccacc atgttggctaatttttgtatttttagtag agacagggtttcagcatgttggcca ggattgtcttgaactcctgacctcag gtgatccacctgcctcggcctccca aagtgctgggattacaggcatgaac caccgcgccaggcctatttgttgctt tttattgaattctagtttttgttgtcaact tttggttgcaattgacaaaagagtgc tatcagacatgaatgttggttgagag ttttctttataagcagtgattgaagtg aaacaattaagatttatgttgaacttt acttgtttatcaatgatgtgagcaact tctttgctgaattggatactagtttgc aaataccagaataatattctgttgata tttttgtgctatacataatgtaaaggct gcagacagaacacattttaaaattta aactggattaatagtttctccatcact gtcttaaatcatttagatctagatatc aacaaaacaatgagtcaatcaatct ctgatttgtggcatttaccaattttcat gatgtaaatacagtagtccccccttа ttcttgggggagatgttctaagaccc cccatgaatgcctgaaactgcagat agtaccaaacattatataatatgctat gttttttcctatacatatacatctatgc aaataatgaaaatatactatacaata cagtttattgtaaattccacttagcca gttgagtcttagaataatttttgctttttt ttttttttttttgtttgtgactgagtctcg ctctgttgcccaggctggagtgcaa tagtgcaatctcggctcactgcaac ctctgcctcccgggctcaagcaatt ctcccgccccagcctcctgagtagc tgggactacaggcatgcaccatgc ccagctaatttttgtatttttagtagag atggggtttcgccatgttggccagg ctggtctccaactcttgacctcaggt gatccgtccgcctcggcctcccaaa	(SEQ ID NO: 10)	gggcacttgccatgcttggcat tttccaagcattacctttatatat atacatatatattttattatacttt aagttctagggtacatgtgcac aacatgcaggtttgttacatatg tatacatgtgccatgttggtatg ctgcacccattaactcgtcattt acattaggtatatctcctaatca agcattacctttttacttctctca acaacttgatcaagtgggaat catcaccccgatcataaggta actgaggcttagagactccag gattcaaaccaggccagcca ctcataatccaacttcaccatgt ggtatatacctctggctttccca ttcattgcttacgtggccctcac gacctaagataggttctgcaa gtgttttcagccccattttaaca gatgcggaaactaaagtccac agaaactcaactgatggtcttt gcaaagtaggagcaaactaat accctgaggtgatttcctttgg atttcacatcacgttactctcac agtttgctttagatacaactgtg agagcctctctggtctgctgg gaccaggaaggagccccctc ttaaaatctctccatttacaaca gagggaacagcacgatctgg ttgtatttctaacctctgcataga gagatatcattttccccccacgt accagttgctttttttttctttcttc agtgaccattgacatttgaaatt aagtagctgtttccatctgtctg gtacctttttctgggtcacaattt gcttatttatttgctccttcacct gcttgtttattcatccatctattta ctctctcatttaatttgctcatcc acctgtctatccgtccatccatc catccatccatccatccatcca tccatccaatcattcatccatcg atccatccattgatctatccatc tatccatccatccatccatccat ccatccatctaatcatccatcc aatcatccatccatccatccat ccatccatccatccatccagtc atccatccattcatccatttaatc atccatccatcgatcaatcggt ccatccatccagctatccatcc atccatccatccatctaatcatt catccatcaatccatccatcca tccatccatccatccatccatc catccatctatccatccatcca atcatccatccatccatcgatc aatcagtccatccatccatcta atcatccatctaatcatccatcc atccatccatccatccatccag tcatccatccattcatccatttaa	(SEQ ID NO: 12)

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
		gtgctgggggtgggggggggaa tcaacaaattattctaagactcaact ggctatgtggaatttataaaatttattt tatcaattaggcatagtgagattaat ggcaataactaacagaacaatgata acaatatactgtaataaaagtatgtg aatgttgtctctctcagaatattgtatt gtgtgtaatatttctgggtcatagttg accacaggcaaccaaaactgcaga aagtgaaatcgaagagaagggcg gactactgttctcctctcagggccaa tttccagctaccaacattatgttactg gacatagagttagggagaaacatg cagatccacatcgctctagaattcct gccatatggctatgatagagacaag taacctcaaaagcacacacaattgt aaatatagtgaaataactggacggg gatgagttttgaacatttattgtctttgt tttatttaattaaattaattaattaa ttttttgagacagagtctggctctgcc ccccaggctggagtgcagtggtgc gatctaggctcactgcaacccctc ctcctgggttcaagcaattctcccgc ctcagcctctggagtagctgggatc acaggtgtgtgccaccacacccag ctaatttttgtatttttagtagagacgg ggtttcgccatgttggccaagctgg cctccaactcctgacctcaggcgat ccacccacctcagcctctcaaagtg ttgggattacaggcgtgagtcactg caccctggcctgtctttgttttaaatat aacttattttatttaattgtaagtttatat ggtttaatttttttggcaatacaaataat aatttctttacccccagctttattgag gtacagctgacatatataattgtataa atttaagatgtacaatgtgaggatttg atataggtatatattgtgaaatgatttc caatttccaattaacacacccatcac ctcacatagttgccatttctttgtgtgt ggtgagaacatttatgatctattctctt agcaaatttctagtacacaacacgg tattgttcactatagtcatcatgctgtc cattagatccccagaacttattcatct tattcattgatttagttctgaataatgg ctgtgtttaacaaccagctcacaag cttcatgagaatttaacgatcagctct tgcgagctcttctgagctggctcca gcatgccactgcccccagcaagtg tgaggattccagtgtgagaccctat ggtgaaaggtacagggtgaatcac agggagaagagaacctggggctg gagcagaaacagaaggcttcctgg aggaggcaggatccagtgttgagc cttggagcatcaggaggatttggct gggggaagggagagtggcccaa gtgaagagacagttagggtggggg gtgggagagggtgggaagcagga		tcatccatccatccatcgatca attggtccatccatccatctatc catccatccatccatccatcta atcacccatccatccatccatc catccaatcatctatccatccat ccatctaatcatctatccatcca tccatccatccatccatccaat catccatccatccatccatcca tccctccatctaatcatccatcc ctccatccatccctccatatcc atccatccatccatccatccat ccatccaaccattcagcagtg aacctctactgtggtgtgttgg aaagaagggaagtgtgagcc tggtcttgctcacagggaactc aaaacagagcattaggagga gactcaggatcttacagccttg agggaggtgacccaggctag aggaagtcaggagctcaaag actagctggggagccaggga cacacacatggattataattag tggctctgcaaggctgcaagt gatcaagggcagactgcggg gtaaaggaaggagagaagac tctgggctttgttaggtggaag ggctttctgcaggaggccagg ccagagaagccagcctgatg aggctttgagaggacctgaaa ctggagatgggatttaatggc cagagagaaggaggcagca ggcttgtcatgtaggcaaatac acgcaagcttgtagagtaatta ccaggcagagagctgagga gagcaccctgttatgttcccttа g (SEQ ID NO: 11)	

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
		agatcagggtgccgtgaatatgtgc ttgcaccttacacttgaattatgtttatt tccttatag (SEQ ID NO: 9)			
STK31 (SEQ ID NO: 2)	aactctggtgg tctccttacaat gagcttggaa cgagatcttctt gatgctgagc ccatgaagga acttagcagca agcgtcctttg gtacgttctga ggttaatgggc agataattctgt taaag (SEQ ID NO: 13)	gtaagtctaacttcttcttacttgtgg gagattcagcagcataaaggatagt ggtgctgttacatttcaagaaagaaa aaacctcactgttagttttctaaagtg tataagttgactggcatgataattgtg cttgtagtttggttgccagaactaatt ttcacttttaaagtttccctcaattgaa caatgagaacacttggacacaggg cggggaacatcacccgtgggggc ctgtcatggggtgggggcaggg ggagggatagcattaggagaaata cctaatgtaaatgatgagttaatggg tgcagcaaaccaacatggcacatgt atacatatgtaacaaacctgcacgtt gtgcacatgtaccctagaacttaaa gtataataataaaaataagtatataa aagtcattaaataaagctttgtctcac aagaaaaaaaaataaagtttccctc atttccctaaatagaatatgtccttaat aggatatactttccttttttttgggtga acttattatagtcatccaatatctttta aaattcttgattttcttacatttctcctt gaaatctcagctgtgttatcagtaca atcttctactctttaatttcttcttctcat tgatctgggaaactgttaatcatgat aataagataattaatttggccttagtt gctacacattctagaatgtgtagcag gctacaggtagaattgcactttccttt tccaaattatagcattcaaagaagat agaaaagtacttataaaattaaaaaa taaaattagatttgatttaacttaaaat aatgaattatgtttaaattggaatgat gtataattataatgagattttcatcttg gaattacgagtgcctctttctttggta caag (SEQ ID NO: 14)	ggctattctgtgg atgttgacacaga agccaaggtgatt gagagagcagc cacctaccataga gcttggagagaa gctgaaggagac tcagggttactgc cattgatattcctg tttttatgtaag (SEQ ID NO: 15)	gtaaagttcttatgctaatctctt gcagaaaccattttatcttgca gaaactattttatttaaaataaat tttagctccaggttttagagca gcagttcattcccctaacataa aagaaaaattatcccatagaat caatgtatgctaaaaatatgttg aagaagtttttacttggagtcac attctctgttttgcttcttag (SEQ ID NO: 16)	tctgatcctatggcttatct gatggtcccatactaccct agggcaaacctgaatgct gttcaagccaacatgcctt taaattcagaa (SEQ ID NO: 17)
USP25 (SEQ ID NO: 3)	gtgtcacaaca taatgaaagtt gctcaagcca aactggaaat gataaaacctg aagaagtaaa cttggaggaat atgag (SEQ ID NO: 18)	gtaatgtgctttcttagaggttaaact acagatttagggattctgtaattcact gatattcctttcgtttcagcccaactg aagttgaatttgttatgaatgatgagt aaaaaaattaagtttccaagtcaatg ctcaagtaattaataatttttatatgtga ttttaggattaacatgatatgatagca tgtagaccaaaaagtttaaaatcaaa aaaggaagatttttggttttttatttca atactaattctcaagtacttttttaaatg aacttttctagcttaaaaacagaata acaaatgcattgttaattttgaaatca acttagaaattgctgttacataatacc aacaagtaatatttatcaagacccta ctatatgccaggcatttgtctaagcc	tgcaggctgtatg aaaattccagcaa gatgttttatggag cttgataagccaa ctgaaaaatgtta cggaaatgcaaa gatcccagtccac ctcccatctctga caagcaa (SEQ ID NO: 20)	gttagtgaacagtataaagact ttttaataaatggttctgggaaa taaataaaatttgatttcttcttta ttccatatgcaaaaattcatttc aattgtgaaagactcctgtgta aagctgtttagaagtcaacata ggtagaagcccaaagacattc atttatccataagaggcactca ctgaccagatatttttgaccag gatatgtaatctagagccaga gaaattcaaaatttgttttctctt gaccctgaggagcaagtcaa gtcatcttttagctaataatgact ccaaatctgtgccctccacatt gattgagagctcttcccttgtcc	gagtggcatcaggattata ggaaattcagggaaacaa ctatgtatctcataattggg ctagaaaattttcaaagag aaa (SEQ ID NO: 22)

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
		ctttacatatattaactgatttaattca gacagtaacactgtaagataggtag tgttcttattcccatttgacacacgag aaaataaggcatagagcagtgaag taacctgcctaaggttttcacagcca caagtttacagcttggctaatagaac ctacattcttaactgttgcactacact gtgcaactgtaccctgtgggctgtat tttttccttcactttgattcatcagatat attttggggaagttaattcattttagtc cctatatgccatatatgtttgtcatga gacagtgttgtgtgtgagagagcac ttacataggtacttgtacatgtgcttg tgggatttttttgtgcacaaaattatttt ttgtcaactttggctttaaaatattcttt attttaattgatgaattcatacaccta gtatttcttcacatttttaaatttaatctt atgcattgtctttttttttaatttataagtc tagtaaatttttaagttgttttcttgggt gcctttgcttaatatttagttccatttac aaacttaattatggtactttagattcc ctaaataattagtgtaattcactacac aaacagggaagaggaaagagaga gaaaacactcattattgatgaaaacc ttagctagtttggcgtataaggtaact accttaattcactaaagagtgtctcta aaaacaaaaataggaaataataaaa ctgcagcaaatatatagaattgtgag cactgaaagcttaaaatttttgttgtc aaaagcaaggcaatgaaaataagc gaaaagtaaaatactggaaaagaa gaaacaaagacatcattatttgtagtt ttttttaccttgaaattcaaataaaatc tacaactgatgtattagaactagtaa gacagctttgcagcatgcaaacttgt gtcagacatcaattttattccctttgc cagtaaaaaaagtatgagaaagtat aaaagtgagaccatttacaatagtat ctgaaactatggagtatatagaaata aatttaacaaaaggtgcgtgtgactt ctgtggggaaaattataaaacttatt gggagacattaacaaagacctaaat aaatgggcatacatatggtatttgga aaacttgatattattaatatatagatat cacttttcccccaaatgctttataaatt cag (SEQ ID NO: 19)		ccaagattcgtgttcatggaac ttgatttttaattgtgacttgattc tgcaatgcttccttcagtttgctt caacagattggactcattatgtt ttttttcataaaatcttctcatatt ctaactgcagggctgttaaatc atacaagtatttaaatgacttgtt tttagattctgcatattcctccta tctttgtcttgctctaagttcagg gttcagtaaaccatagcctgca gaccagcttaccacctgtttttg taaataaacttttattggaacat ggccatttgtatacatattgtct gtggatactttcacactgcagtt acagagttgagtaactgaaga agatgaggcttgcaaagtcta aaatattatctggccttttacag aaaaagtttgcctactcacaac ttctaaggtgattctaatgacaa caaggaatgacagtttgaata ccttccccagaaaagctaactt catacttgcaggacaaatttta cctttatttgttatgaagatctac taatagcatgaaaattatgtctt gaaaattaaaattaagactgga tgtggtggctcatgcttataatc ccagcacttttggaaatcaag gcacagggagttgtttgaggc caggaatttgagaccggcctg gccaacatagcaagaccctgt ctatacaaaacagattaagtaa ttaactggaactgggagcatg cacctgtaattcaagctacttg ggagacagaggtaggaggat tgctggagcccagaaggtca agtctgcagttagccgtgatcg tgccactgcactgcagactgg gcaactgagtgagaccctgtc tccaaaaaaaaaaaaaaaagt ttatattcaaggagaatatcgtt ggctaataactatgtttcttatttt tgtttataatgtttatataaaagc cttaacacttttgttatcagtatt cactttccaatttatcctacattg aatttcagttatttttcttgagtca attattataaagtaagctaaga aaatcttgactttggtatcacctt agaattatttagtcctattacatg ccctatagttgttttgggggtg gggaggggcattgaaatgtg aactgcagttattcttttaccata tctcatgtaaataaactattcgc ctacacacttaaagtcaaattta ccacatatttgcgaagttaaaa ggttaaataaatctaacttttttt agtttcttaaaattatgtattaag tatgtgtccaagctttgattcat	

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
				actcagcgtaatatttcattttct tccatcagttgatgtagatataa taatagttcaataatatgatttta tactgatgaatttttgagatgac ctgattgattatcctgcctttata tttctgttatctgctagtaaaatc tttattgtttaaaacatttcatatc catgcctaataatttttagattta atgttgtattaataaatagtcttg tgtgccagaaagtaaagtcttg ttttagttctgatgtatcttcatttt gaactctttatttctattaattca gatagtagcgaaacagtgcat tacttttgtacagtttcatagta ctttttaaatcttactctgaacatt tttaatattttctttcacctacag (SEQ ID NO: 21)	
GTF2B (SEQ ID NO: 4)	gatcgaacaa ataatttattca agcaagtatat gaacagaaga gcctgaaggg aagagctaat gatgctatagc ttctgcttgtct ctatattgcctg tagacaagaa ggggttcctag gacatttaaag (SEQ ID NO: 23)	gtaagtttttagctttgttttttgtgttt gtttgtttgtttgtttgagataggtct cactttgtcgcccaggccagagtgc agtggcgcaatcttgactcactgca acctctgcctcccaggttcaagctat actcctgcctcagcctcctgagtag ctgggattacaggcatgtgccacca cacctggttaacttttgtatttttagta gagacagggtttcgccatgttggcc aggctggtctcaaacttctgacctca ggtgatctgcccactttggcctccca aagtgctagaattaacaggcgtgag gcactgtgcccatcctaaaataacat tttctacatcatattaaagaaacaga ccaatgtcaggcagctttgacactta gaagctaagtcatctttattgtgacta cagtaatccacttgagcatcatttgc caatgcctacttatttcattccttgaca gcaaattcatagcatgttcttttaggc tgttttcccatgtagactcattttagtc cacgtaagcccactatgttcttttcta gtgagatcttttgatggactgttgcttt gtgggctttcagaaggcctataatc gtaaaccaaagtaggccgtttatac atttggtttacatggaaacagtttcca tttctagcactggtctctgtagatgag ggagtcaagctttagaaagtatcag tgaaggccaggtatggtgtctcaca cctgtaatcccagaactttgggagg ccgaggtggttgggtcatgaggcc aggagttcaagaccagcctggcca atatggtgaaaccccgtctctactaa aaatacaaaaattagccaggcgtg gtggtgcgcgcctgtagtcccagct actcaggaggctgaggcagaaga atcgcttgaaaccgggaggcagag gttgcagtgagccgagatcacgcc actacactccagcctgggcgacag agtgagacttcatctcaaaaaaaaa	tttgcagttacatg tgaatatgtgataa tttgttcagtgttgc tcactgacttcca ctaccaaagctca gtttagctgaaga ctatacttaaccct ttattcatcca (SEQ ID NO: 25)	gtgagtatttatttttttatgttttct tacgtaaaagtgaagaggtga tggcctttatcatcatctttttgt tgtttttataactaaaacaacttttt gttaaatttcttctgaaagtgtgt ggtggttgcatctgggaatag cagggtcagatattttatctcat tgagctacagcccataatcgc agggatttcatgcacacatga cattgtgttgcctaagaggcat cctttagcttgagtattttctttca attccttggtaacactgcttaca acagattaatacataaacatca aagaatttttcatttaaagaaatt tagttgtatgaattaatccttag agcagttaatattttgaaagatt aagacctaaatttagccaggc acgatggctcacaccggtaat ctcagcactttgtgaggccga ggcaggtagatcacgaggtc aggagttcaagaccagcctg gccaagatggtgaaaccccat ctctactaaaaatacaaaaaat tagccgggtgtggtggcagg cgcctgtaatcccaactactcg ggtggcaggagaatcgcttga acccaggggcagaggttgc agtgagccaagatggtgtcac tgcactccagtctggacgaca gagtgaaatttcatctcaaaaa acaaaaaaactctaaattgaaa gagtgaagaacattaggcca ggcgcagtggctcatgcctgt aatctcagcactttgggaggc cgaggtaggaggatcacctg aggtcaggagttctagacagg cctggccaacatggtgaaacc ccatctctactaaaaatacaaa agttagccaggtgtggtggtg	aaatatgtgccgtatcacg aatttctaagaaagaaatt ggtcggtgttttaaacttatt ttgaaagcgctagaaacc agtgtggatttgattacaa ctggggacttcatgtccag gttctgttccaacctttgtct tcctaaacaagtacagatg gcagctacacatatagcc cgtaaagctgtggaactg gacttggttcctgggagg agccccatctctgtggca gcggcagctatttacatgg cctcacaggcatcagctg aaaagaggacccaaaaa g (SEQ ID NO: 27)

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
		aaaaaaaaagttatcagttaagttttt tctcttgaggtgtattgtttgcactttta cacttcagttaagaactggtatttttc atttag (SEQ ID NO: 24)		cacgcctgtagtcccagctact cgggaggctgaggcaggag aatcacttgagcctaggaggt ggaggttgcagtgagctgag atcatgccactacccttagtct gggcgacagagtgagacttc atctcaaaaaaaaaggtgtgat gaacattaaagttaactgagaa atttgcagttaactctgtggga agtatcactgagatgattaggg tttcatctttaaggtatgattgaa tgttttacaattagaatgtatatt atgtttataattagaaatggtag attttaagatgaactatatggct taagtgaattgatttatttaatttt ttttcttttag (SEQ ID NO: 26)	
LYPD3 (SEQ ID NO: 5)	atggaccccg ccaggaaagc aggtgcccag gccatgatctg gactgcaggc tggctgctgct gctgctgcttc gcggag (SEQ ID NO: 28)	gtgggtctgctgcccggagaggtg gaggggtagagtgggggctctta gggaggaggctcctctggccccgt acagagaaggctgtgggaaggatc cccacataagcagagggcactgtg ggcaaagttcccccatcctaagcaa gggaggcagagggaagggtcctc caccatgggccgtaggaagggttc cccaccacggacaaggggagctg tgggcaaaattttcctatcctgaagc tgcagaaagaagtcccctgttccta ggccgggcgcggtggttcatgcct gtaatctcagcactttgggaggccg aggcaggcggatcatgaggtcaaa cgatcgagaccagcctggccaaca tggtgaaaccctgtctctactaaaaa tacaaaaattagctgggcgcggtag cgcgagcctgtagtcccagctactc gggaggctgaggcaggagaatcg tttgaatccaggaggggaggttgca gtgagccgagattgcgccactgca ctccagcctggcgacagagcgag actccgtctcaaaaaaaaaaaaaag aaaaaaagaaaaaaagaaaaaagt gttccctgtcctgggcagagaagct gtgggtagtgttctctcatttgagca gaaggggctgtgggaagggatc cctcactgagcacagtctactgtgg gaaaagttctcccacacactgtggg aaaagttctctcattctgagcagaaa accgtgaggaggtatccttcctacc ccgggatgaggggcttcatggcgg ggatcgcctacagtgggcatacag ggctgtcggaaggatgtgcccaca gttggcatagtgggtgtgaaggaaa agggtccagtgattcccgtgcttcc agatttgtggaaagcttttccctactc ctgccctggaaaagtgtgtgcaata agactgtctaatcccgggacagag	gagcgcaggcc ctggagtgctaca gctgcgtgcaga aagcagatgacg gatgctccccga acaagatgaaga cagtgaagtgcg cgccgggcgtgg acgtctgcaccg aggccgtgggg gcggtggagacc a (SEQ ID NO: 30)	gtgagtggggacccccgggc cacgcacaccactgcaggct cctccggcccagccccgcct ccttatagtcctttctggatctg gccacgcccctttaacgcggc cccacccctccctttcagcctg gccgggaccctacgacacag ccttgcaggccaggccttctct attaacaattacaagcgcgag gagcagagagttgagaactg ggtggggcggacggggaga gctcgggtttggaacgaggtc acgtcctatgcacttttctgctc caccaatgagacctccccatc cctgcctgtcaaccaggctgt ctcctccgactcagctctttga ggttccgcccccgttccgtga ccccgctcctccctgcttcggt cgttctttccgctctgagcggc tctttgggggcgtccatccgcc aattcctgcctgtcggtcctgt ggtgttggcctatttgaggccc cgcccctcccagaccctaccc aactctgggccacgccccatc tgagaccacgcccacagcta atcccgcccttgtccctctccct ag (SEQ ID NO: 31)	tccacggacaattctcgct ggcagtgcggggttgcg gttcgggactccccggca agaatgaccgcggcctg gatcttcacgggcttctgg cgttcatccagctgcagc aatgcgctcaggatcgct gcaacgccaagctcaac ctcacctcgcgggcgctc gacccggcag (SEQ ID NO: 32)

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
		ggcatgggaaggactatttctccttg ggggtggacgctttagcaaggagg gctcaggtctattcggctccgctccc ccag (SEQ ID NO: 29)			
IK (SEQ ID NO: 6)	gtgagccgttc tccaacccttt ggcccccgat ggccacgatg tggatgatcct cactccttcca cca (SEQ ID NO: 33)	gtgagtattttgtgctaggaccatgg aaatgagttgggcaatgaatagaaa agtggtggtgatggtgaaagaattc tgaatgtcttgtcatcaaggatggta aaaaattataccgctggccctttat ctctgtgagccttaactggatcagg ccaagtgcataagggtattagaagg gaaattcctgcaggaaaggtgttaa gttttttttttttttatacgtgttgtgtaa ttacacattatataggaattacttcttc caatagcaaattaaaacataggaaa taagggtaaaatcatctttttatttttttt gagactgagtttcgctcttgttgctca g (SEQ ID NO: 34)	gctggaatgcagt ggtgcgatctcg gctcaccacaac ctccgcctcccg ggttcaagcaatt ctcctgcctcagc cttccgagtagct gggattacaggc at (SEQ ID NO: 35)	gttccactacgcccggctaatt ttgtatttttagtagagatgggg tttctccatgttgatcaggctgg tcttgaaatcccgacttcaggt gatccacccgcctctgcctcc caaagtggtgggagtacagg cgtgagccaccgtgcctggc ccccgataaaatcatctttgata acatccccaattccagtccattt ctttttttttttctttttctttttctt tctttttttttttttttgagacggag tctcactgttacccaggctgga gtgcagtgatgcaatcttggct cactgcaacctctgcctcctga gttcaagcggttctcctgcctc agcctcccaagtagctgggac tacaggcatgtgccaccatgc acgactgatttttttttttgtatttttt agtagagacagggtttcactgt gttcgtcaggctggtcttgaac tcctggcctgaagtgatccac ccgcctcagcctcctaaagtg ctggcattacaggcgtgagcc accgtgcccagcctcagtcctt ttcttagaagtaaccacagtta ccttttttggcatatcttttcctcc cgttgcccaggctggagttca atggcattatttctgctcacagc aacctttgcctcctgggctcaa gcaattcttctgcctcagtttcc ccagtagctgggattacaggc atgccccaccatgcttggctaa tttttgtattagtagagacaggg tttcgctatgttggccaggctg gtcttgaactcctgacctcaaa tgatctgcccgtctcggcggc ccaaagtgctgggattacagg cgtgagccactgcgcctggc cctcctactctttttaaatgcatt tatacacatacatgcatatctga ggaaaacacaacattgctttgt gatttaaaatatttgagttaaat gatattatgctgtatatatcatttt tgcaacttgcttctttttaaaccc tgctcagtgtttctgagattatct ttcagagcagtgacaagcatt ccctccttggctaaactttagtc aggttccccaaagagtcccatt tttcagcaggactcatccgtgg cctgctgagcacagttttaaca aagaaatcaaccaccgctcca	atcaaaactcaccaatga agacttcaggaaacttctc atgacccccagggctgca cctacctctgcaccacctt ctaagtcacgtcaccatga (SEQ ID NO: 37)

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
				ccccc atacctaaccaagttg cttttagtaattttctataaattct cacttgtccctgttgtattcgga attaagttcagtctctctactcta ttacaatggtcttgactcttgtta cagtagtcttcaataaagtcttc cttggggagaagaggaacttg attagatatcaagttggattgg atatcaagagtgggaaaacttt ctgtgaactgacttaggattcct ttttaaaactgaggtcagcagg ccaaggaagagccctagtca gaagagagctttaaaggagtc tgactttaatttgatcaagagag gagtctttgttagtgtaccaatt aaaataatctcttgtaccattag tgctgtctttccctctttgggaa acaccatcatagtttagtcagt ctccagctaataaacgcttagt ctgtttcagactgattcttttcttt ttttttttttttgagacagagtct tgctctttttgcccaggctgga gtgtaatacctagatctcggct cactgcaaactccgcctccca ggttcaagcgattttcctgcctc agcctcctgagtagctaggatt acaggcgcccaccacaacgc ccagctaatttttttgtatttttag tagagatgggatatcaccatgt tggccatgctggtcttgaactc ctgacctcgggtgatccgcct gcgtcagcctcccaaagtgcg gcgattataggtgtgagccact gcacccggccagactgattat tttcaaggctataactgaattaa atttacaggccttcttatatacat atgtcagtaatggaccaggct aaatataaataaatggagttgc tggctcataggatatatatactt aaattttaacaaatactgcaaa attgcccttcaaattggttatact ggtttacattctcatcagcattat atgagagagtcccccttcccaa attctcactcacacttgatttatc aaactttctggtagttaacactc tgatgcattaaaagtggcatgg catatctggtcgggcgtggtg gcttacacctataatcccagca ctttgggaggctgaggtgggc agatcacctgaggtcagaagt ttgggaccagcctggccaaca cctctaataaaaatacaaaaat tagctgggtgtgttggcgggc acctgtattcccagctacttgtg aggctgaggcaggagaatta cttgaacccaggaggcagag gttgcagtgggctgagatcat	

Gene	Upstream Exon	Upstream Intron	Alternatively Spliced Exon	Downstream Intron	Downstream Exon
				gccactgcactccagccggg gcaacagagggagactctgt cttaaaaaaaaaaaaaaaag ctgggtgcggtggctcatccc tataatcccagcactttgggag gccgaggtgggtggatcaca aggtcaggagttcaagacca gcctggccaatatgggtgaca gtgagactccgtctcaaaaaa aaaaaaacagtgtcatatcttt gttttaatttgtatttccctgattt ctgttaatttaatcatgatttttac ttgttccttttcatttcttattgag tcctaatatttaaaatctttgcctt ttcttctag (SEQ ID NO: 36)	
TEX14 (SEQ ID NO: 7)	tggaggtgaa agaaatggca aagaaagcag ctactggaca gctcacagtac ctccttggcat cctcagagta gtctgactttag agagcgagg ctgaaaatga gcccgacgcc ctgctgcagc cccccattagg agcccagaaa acacggattg gcagcgagtt attgagtatcat agggaaaatg atgagcccag aggaaatggc aagtttgacaa gacgggcaac aatgactgtga cagtgaccag catggcagac agcccaggct tggaagcttca cca (SEQ ID NO: 38)	gtaagttcacttgcgggaatcagga taggtgaactcaatcaagtatcagtc gggcgaggtggctcagcctgtaatt ccagcactttgggaggtgaggcga gcggatcacctgagcccaggaattt gggaccagcctggtcaacatgatg aaacctggtctctacaaaaaaaaaa aaaaaaaaaaaaaaaaacattagctg ggcgtggtggcacacgcctctagtt ccaggtacccaggaggctgaggtg ggaggatcacttggacctgagagg tggaggttgcagtgagctgtgatca tgaaactgcactccagcttgggtga cagtgagaccctgtctcaaaataaa aataaaagtaaaaaaaaagttatcag ctctcaaaaaagtaaactttccagca tatttttttggtttggtcctctctccttt cctacttttctggaaaccttgaattcc aggggaagacagtatagccctatat ctcattgcatttcccacaggcaaaca gatccagacactgatcagtttttatat tactgaaacag (SEQ ID NO: 39)	gtcccttgtccaa aggaagtaaaat aatcaaatactga ccctggttttggc gaggaatcgatc cttctttacctggg gatagag (SEQ ID NO: 40)	gtaggccatcactgcctactg cccatgtctctgcattgtattcta agctgaaattaaattctcccag aatttatttagaatgtattctaag ctggcattaaatttttccagaga gtctatgctggaggaacctag agacacatgacatttattgtcat cggtttatagatttacaggttaa aaaaaatgaaggctgtatatttt cagttgataatttgcttataatca gggcatggatattgtttcacag ttgtttataagttaaagaaagat acatctaaacttgtcatttgtctt aggcagaaaatctattggcttg ggtattatttgtatttagaatctg tttcttgtcaagcatcactctcg aatgggtgggcagatgatcca gggctgcattttggggtcagg ctaaactaacagtctctactctc tggactgaacctaacgtatgg gacatggccactcttggcaga atatgtatacatctcttctttcat gccacctttatcctcttccatgtt gtctgaatatcactgaatgactt tgaccctgactgtgttacaatc attag (SEQ ID NO: 41)	gtatcaggcacccatctcc cagacaaaaggagcaac cagagcatagtgaagcct tccaagcaagttctgacac attggtggctgtagagaa atcttacagt (SEQ ID NO: 42)

[0311] In some aspects, the expression cassette further comprises an untranslated region (UTR), a microRNA binding site, a polyA (pA) sequence, an intron sequence, or any combination thereof.

[0312] In some aspects, the expression cassette comprises a poly(A) (pA) sequence.

[0313] In some aspects, the pA sequence is a synthetic pA sequence, a bovine growth hormone (bGH) pA sequence, or a human growth hormone (hGH) pA sequence.

[0314] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon, are derived from TRPV3. TRPV3 is also known as Transient Receptor Potential Cation Channel Subfamily V Member 3 (see ncbi.nlm.nih.gov/gene/162514). In some aspects, the alternatively spliced exon is derived from a sequence between exon 13 and exon 15 of TRPV3. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon has been modified relative to WT TRPV3. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon have been shortened relative to WT TRPV3.

[0315] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon, are derived from STK31. STK31 is also known as Serine/Threonine Kinase 31 (see ncbi.nlm.nih.gov/gene/56164). In some aspects, the alternatively spliced exon is derived from a sequence between exon 18 and exon 20 of STK31. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon has been modified relative to WT STK31. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon have been shortened relative to WT STK31.

[0316] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon, are derived from USP25. USP25 is also known as Ubiquitin Specific Peptidase 25 (see ncbi.nlm.nih.gov/gene/29761). In some aspects, the alternatively spliced exon is derived from a sequence between exon 22 and exon 23 of USP25. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon has been modified relative to WT USP25. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon have been shortened relative to WT USP25.

[0317] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon, are derived from IK. IK is also known as IK cytokine (see ncbi.nlm.nih.gov/gene/3550). In some aspects, the alternatively spliced exon is derived from a sequence between exon 2 and exon 3 of IK. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon has been modified relative to WT IK. In some aspects, the upstream exon, the upstream intron,

the alternatively spliced exon, the downstream intron, and/or the downstream exon have been shortened relative to WT IK.

[0318] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon, are derived from TEX14. TEX14 is also known as Testis Expressed 14, Intracellular Bridge Forming Factor (see ncbi.nlm.nih.gov/gene/56155). In some aspects, the alternatively spliced exon is derived from a sequence between exon 18 and exon 19 of TEX14. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon has been modified relative to WT TEX14. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon have been shortened relative to WT TEX14.

[0319] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon, are derived from GTF2B. GTF2B is also known as General Transcription Factor IIB (see ncbi.nlm.nih.gov/gene/2959). In some aspects, the alternatively spliced exon is derived from a sequence between exon 5 and exon 6 of GTF2B. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon has been modified relative to WT GTF2B. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon have been shortened relative to WT GTF2B.

[0320] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon, are derived from LYPD3. LYPD3 is also known as LY6/PLAUR Domain Containing 3 (see ncbi.nlm.nih.gov/gene/27076). In some aspects, the alternatively spliced exon is derived from a sequence between exon 1 and exon 3 of LYPD3. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon has been modified relative to WT LYPD3. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and/or the downstream exon have been shortened relative to WT LYPD3.

III. Splicing Modifiers or Modulators

[0321] Certain aspects of the disclosure are directed to expression constructs comprising a gene of interest which is regulatable (e.g., regulated) by a molecule (e.g., a splicing modifier or modulator). In some aspects, the molecule promotes the inclusion of an alternatively spliced exon. In some aspects, the molecule is kinetin. In some aspects, the molecule is homocarbonyltopsentin.

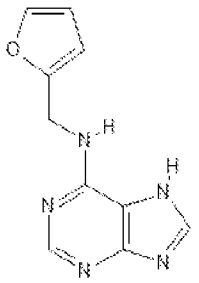
[0322] Certain aspect of the disclosure are directed to increasing AAV production by contacting or culturing a host cell with the molecule (e.g., an indole alkaloid). In some aspects, the molecule is a topsentin alkaloid or derivative thereof (e.g., Homocarbonyltopsentin, Topsentin, Bromotopsentin, Bromodeoxytopsentin, Deoxytopsentin, and any combination thereof). In some aspects, the topsentin alkaloid or derivative thereof is homocarbonyltopsentin. In some aspects, the molecule is homocarbonyltopsentin.

[0323] *Kinetin*

[0324] Kinetin or Drug K is a member of the class of 6-aminopurines that is adenine carrying a (furan-2-ylmethyl) substituent at the exocyclic amino group (PubChem CID 3830). Kinetin is a member of furans and a member of 6-aminopurines. In some aspects, kinetin has a role as a geroprotector and a cytokinin.

[0325] In some aspects, kinetin has been shown to increase inclusion of exon 20 in IKBKAP (see Hims MM, et al., J Mol Med (Berl). 85(2):149-61 (Feb. 2007)).

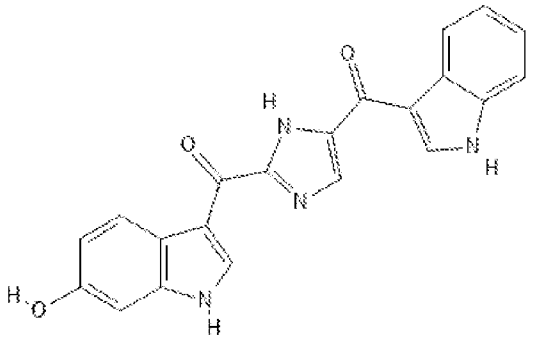
[0326] In some aspects, kinetin is represented by the structure:



[0327] *Homocarbonyltopsentin*

[0328] In some aspects, homocarbonyltopsentin, or Drug H (see PubChem CID 11326167), binds to TSL2 and converts from it from a pentaloop to triloop conformation. TSL2 is a RNA structure that regulates exon skipping, promoting skipping in the pentaloop formation (See Garcia-Lopez, A., et al., Nat Commun 9, 2032 (2018)), thus homocarbonyltopsentin leads to exon inclusion.

[0329] In some aspects, homocarbonyltopsentin is represented by the structure:



[0330] Indole alkaloids are a class of alkaloids containing a structural moiety of indole. In some aspects, the idole alkaloid is a topsentin alkaloid or derivative thereof (e.g., Homocarbonyltopsentin, Topsentin, Bromotopsentin, Bromodeoxytopsentin, Deoxytopsentin, and any combination thereof). In some aspects, the topsentin alkaloid or derivative thereof is homocarbonyltopsentin.

IV. Delivery Vectors

[0331] In some aspects, provided herein is a delivery vector comprising any of the expression cassettes disclosed herein. In some aspects, the expression cassette is flanked by inverted terminal repeat (ITR) sequences.

[0332] In some aspects, the delivery vector is selected from the group consisting of a viral vector, a plasmid, a lipid, a protein particle, a bacterial vector, a lysosome, a virus-like particle, a polymeric particle, an exosome, and a vault particle.

[0333] In some aspects, the vector is a viral vector.

[0334] In some aspects, the vector is a lentiviral vector, an adenoviral vector, an adeno-associated viral (AAV) vector, or a retroviral vector.

[0335] In some aspects, the vector is an adeno-associated viral (AAV) vector.

[0336] In some aspects, the vector is a lentiviral vector.

[0337] In some aspects, provided herein is a cell comprising any of the expression cassettes disclosed herein or any of the delivery vectors disclosed herein.

[0338] In some aspects, provided herein is a recombinant viral vector (e.g., rAAV) comprising any of the expression cassettes disclosed herein or any of the delivery vectors disclosed herein.

IV.A. Non-Viral Vectors

[0339] In some aspects, the expression cassettes can be administered using a non-viral vector. "Non-viral vector," as used herein is meant to include naked DNA, chemical formulations containing naked DNA (e.g., a formulation of DNA and cationic compounds (e.g., dextran

sulfate)), and naked DNA mixed with an adjuvant such as a viral particle (i.e., the DNA of interest is not contained within the viral particle, but the transforming formulation is composed of both naked DNA and viral particles (e.g., AAV particles) (see e.g., Curiel et al., Am. J. Respir. Cell Mol. Biol. 6:247-52 (1992)). Thus the "non-viral vector" can include vectors composed of DNA plus viral particles where the viral particles do not contain the DNA of interest within the viral genome.

[0340] In some aspects, the non-viral vector is a bacterial vector. See e.g., Baban *et al.*, Bioeng Bugs., 1(6):385–394 (2010).

[0341] In some aspects, the expression cassettes can be complexed with polycationic substances such as poly-L-lysine or DEAC-dextran, targeting ligands, and/or DNA binding proteins (e.g., histones). DNA- or RNA-liposome complex formulations comprise a mixture of lipids which bind to genetic material (DNA or RNA) and facilitate delivery of the nucleic acid into the cell. Liposomes which can be used in accordance with the disclosure include DOPE (dioleyl phosphatidyl ethanol amine), CUDMEDA (N-(5-cholestrum-3-β-ol 3-urethanyl)-N′, N′-dimethylethylene diamine).

[0342] Lipids which can be used in accordance with the disclosure include, but are not limited to, DOPE (Dioleoyl phosphatidylethanolamine), cholesterol, and CUDMEDA (N-(5-cholestrum-3-ol 3 urethanyl)-N′, N′-dimethylethylenediamine). As an example, DNA can be administered in a solution containing one of the following cationic liposome formulations: Lipofectin™ (LTI/BRL), Transfast™ (Promega Corp), Tfx50™ (Promega Corp), Tfx10™ (Promega Corp), or Tfx20™ (Promega Corp). The concentration of the liposome solutions range from about 2.5% to 15% volume:volume, preferably about 6% to 12% volume:volume. Further exemplary methods and compositions for formulation of nucleic acids (e.g., DNA, including DNA or RNA not contained within a viral particle) for delivery according to the method of the disclosure are described in U.S. Pat. Nos. 5,892,071; 5,744,625; 5,925,623; 5,527,928; 5,824,812; 5,869,715.

[0343] In some aspects, protein particles can be used in accordance with the disclosure for polymer-based gene delivery. See e.g., Putnam et al., PNAS 98 (3): 1200-1205 (2001).

[0344] In some aspects, the expression cassettes can be administered as a chemical formulation of DNA or RNA coupled to a carrier molecule (e.g., an antibody or a receptor ligand) which facilitates delivery to host cells for the purpose of altering the biological properties of the host cells. The term "chemical formulations" refers to modifications of nucleic acids to allow coupling of the nucleic acid compounds to a carrier molecule such as a protein or lipid, or derivative thereof.

IV.B Viral Vectors

[0345] In some aspects, the expression cassettes can be administered using a viral vector (e.g., a recombinant viral vector). In general, recombinant viral vectors used in accordance with the disclosure can comprise a viral particle derived from a naturally-occurring virus which has been genetically altered to render the virus replication-defective and to express a recombinant gene of interest in accordance with the disclosure. Once the virus delivers its genetic material to a cell, it does not generate additional infectious virus but does introduce exogenous recombinant genes into the cell, preferably into the genome of the cell.

[0346] In some aspects, the viral vector is a retrovirus, an adenovirus, an adeno-associated virus (AAV), a herpes simplex virus (HSV), a cytomegalovirus (CMV), a vaccinia or a poliovirus vectors. In some aspects, retroviral vectors are less preferred since retroviruses require replicating cells and secretory glands are composed of mostly slowly replicating and/or terminally differentiated cells. In some aspects, adenovirus and AAV are preferred viral vectors since this virus efficiently infects slowly replicating and/or terminally differentiated cells. In some aspects, the delivery vector (e.g., viral vector) is selected from the group consisting of an adeno-associated viral (AAV) vector, an adenoviral vector, a lentiviral vector, or a retroviral vector.

[0347] Where a replication-deficient virus is used as the viral vector, the production of infective virus particles containing either DNA or RNA corresponding to the DNA of interest can be produced by introducing the viral construct into a recombinant cell line which provides the missing components essential for viral replication. In some aspects, transformation of the recombinant cell line with the recombinant viral vector will not result in production of replication-competent viruses, e.g., by homologous recombination of the viral sequences of the recombinant cell line into the introduced viral vector. Methods for production of replication-deficient viral particles containing a nucleic acid of interest are well known in the art and are described in, e.g., Rosenfeld et al., Science 252:431-434 (1991) and Rosenfeld et al., Cell 68:143-155 (1992) (adenovirus); U.S. Patent No. 5,139,941 (adeno- associated virus); U.S. Patent No. 4,861,719 (retrovirus); and U.S. Patent No. 5,356,806 (vaccinia virus).

IV.C Adeno-Associated Virus

[0348] AAV, a parvovirus belonging to the genus Dependovirus, has several attractive features not found in other viruses. For example, AAV can infect a wide range of host cells, including non-dividing cells. Furthermore, AAV can infect cells from different species. Importantly, AAV has not been associated with any human or animal disease, and does not appear to alter the physiological properties of the host cell upon integration. Finally, AAV is stable at a

wide range of physical and chemical conditions, which lends itself to production, storage, and transportation requirements.

[0349] The AAV genome, a linear, single-stranded DNA molecule containing approximately 4700 nucleotides (the AAV-2 genome consists of 4681 nucleotides), generally comprises an internal non-repeating segment flanked on each end by inverted terminal repeats (ITRs). The ITRs are approximately 145 nucleotides in length (AAV-1 has ITRs of 143 nucleotides) and have multiple functions, including serving as origins of replication, and as packaging signals for the viral genome.

[0350] The internal non-repeated portion of the genome includes two large open reading frames (ORFs), known as the AAV replication (rep) and capsid (cap) regions. These ORFs encode replication and capsid gene products, respectively: replication and capsid gene products (i.e., proteins) allow for the replication, assembly, and packaging of a complete AAV virion. More specifically, a family of at least four viral proteins are expressed from the AAV rep region: Rep 78, Rep 68, Rep 52, and Rep 40, all of which are named for their apparent molecular weights. The AAV cap region encodes at least three proteins: VP1, VP2, and VP3.

[0351] AAV is a helper-dependent virus, requiring co-infection with a helper virus (e.g., adenovirus, herpesvirus, or vaccinia virus) in order to form functionally complete AAV virions. In the absence of co-infection with a helper virus, AAV establishes a latent state in which the viral genome inserts into a host cell chromosome or exists in an episomal form, but infectious virions are not produced. Subsequent infection by a helper virus "rescues" the integrated genome, allowing it to be replicated and packaged into viral capsids, thereby reconstituting the infectious virion. While AAV can infect cells from different species, the helper virus must be of the same species as the host cell. Thus, for example, human AAV will replicate in canine cells that have been co-infected with a canine adenovirus.

[0352] In some aspects, to produce recombinant AAV (rAAV) virions containing the HNA, a suitable host cell line is transfected with an AAV vector containing the HNA, but lacking rep and cap. The host cell is then infected with wild-type (wt) AAV and a suitable helper virus to form rAAV virions. Alternatively, wt AAV genes (known as helper function genes, comprising rep and cap) and helper virus function genes (known as accessory function genes) can be provided in one or more plasmids, thereby eliminating the need for wt AAV and helper virus in the production of rAAV virions. The helper and accessory function gene products are expressed in the host cell where they act in trans on the rAAV vector containing the heterologous gene. The heterologous gene is then replicated and packaged as though it were a wt AAV genome, forming a recombinant AAV

virion. When a patient's cells are transduced with the resulting rAAV virion, the HNA enters and is expressed in the patient's cells. Because the patient's cells lack the rep and cap genes, as well as the accessory function genes, the rAAV virion cannot further replicate and package its genomes. Moreover, without a source of rep and cap genes, wt AAV virions cannot be formed in the patient's cells. See e.g., U.S. Appl. Publ. No. 2003/0147853.

[0353] In some aspects, AAV vectors of the present disclosure can comprise or be derived from any natural or recombinant AAV serotype. According to the present disclosure, the AAV serotype can be, but is not limited to, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAV9, AAV10, AAVrh10, AAV11, and AAV12. In some aspects, the AAV serotype is AAV2. In some aspects, an AAV vector includes a hybrid vector (e.g., AAV-DJ, AAV-PHP.B, AAV2-ESGHGYF, AAVM41, AAV-LK03, AAV2-BR1, AAV587MTP, AAV-Anc80L65, AAV2-7m8, AAV2HBKO, AAV2YF, AAV6-RGD or AAV6.2). In some aspects, the AAV vector is modified relative to the wild-type AAV serotype sequence. In some aspects, the modified AAV vector is a modified AAV6, e.g., an AAV6 vector comprising the RGD peptide (an AAV6-RGD vector) or an AAV6 vector comprising mutations of surface exposed tyrosine residues as described, for example, in Sayroo et al. Gene Ther. 2016 Jan; 23(1):18-25. In some aspects, the AAV6-RGD vector further comprises modified amino acids corresponding to Y705, Y731, T492, and K531, (e.g., Y705, Y731F, T492V, and K531E; also referred to as AAV-RGD-Y705-731F+T492V+K531E).

[0354] The AAV vectors of the present disclosure comprise a viral genome with at least one ITR region and a heterologous payload (e.g., an expression cassette comprising an alternatively spliced exon). In some aspects, the AAV vector comprises two ITRs. These two ITRs flank the payload region at the 5' and 3' ends. The ITRs function as origins of replication comprising recognition sites for replication. ITRs comprise sequence regions which can be complementary and symmetrically arranged. ITRs incorporated into AAV vectors of the disclosure can be comprised of naturally occurring polynucleotide sequences or recombinantly derived polynucleotide sequences.

[0355] The ITRs can be derived from the same serotype as the capsid, selected from any of the serotypes listed herein, or a derivative thereof. The ITR can be of a different serotype from the capsid. In some aspects, the AAV vector has more than one ITR. In a non-limiting example, the AAV vector has a viral genome comprising two ITRs. In some aspects, the ITRs are of the same serotype as one another. In some aspects, the ITRs are of different serotypes. Non-limiting

examples include zero, one or both of the ITRs having the same serotype as the capsid. In some aspects both ITRs of the AAV vector are AAV2 ITRs.

[0356] Independently, each ITR can be about 75 to about 175 nucleotides in length. An ITR can be about 100-105 nucleotides in length, about 106-110 nucleotides in length, about 111-115 nucleotides in length, about 116-120 nucleotides in length, about 121-125 nucleotides in length, about 126-130 nucleotides in length, about 131-135 nucleotides in length, about 136-140 nucleotides in length, about 141-145 nucleotides in length or about 146-150 nucleotides in length. In some aspects, the ITRs are about 140-142 nucleotides in length. Non-limiting examples of ITR length are about 102, about 140, about 141, about 142, about 145 nucleotides in length, and those having at least 95% identity thereto.

[0357] In some aspects, the AAV vector comprises a nucleic acid sequence disclosed herein which can be located within about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30 or more than about 30 nucleotides downstream or upstream from the 5' or 3' end of an ITR (e.g., Flip or Flop ITR) in the vector.

[0358] In some aspects, the AAV vector comprises one or more filler sequences (also referred to as "stuffer sequences"). In some aspects, the AAV vector comprises one or more filler sequences in order to have the length of the AAV vector be the optimal size for packaging.

[0359] In some aspects, the AAV vector is a single stranded (ss) AAV vector.

[0360] In some aspects, the AAV vector is a self-complementary (sc) AAV vector.

[0361] In some aspects, the AAV vector can comprise one or more filler sequences between one of more regions of the AAV vector. In some aspects, the filler region can be located before a region such as, but not limited to, a payload region, an ITR, a promoter region, an intron region, an enhancer region, and/or a polyadenylation signal sequence region. In some aspects, the filler region can be located after a region such as, but not limited to, a payload region, an ITR, a promoter region, an intron region, an enhancer region, and/or a polyadenylation signal sequence region. In some aspects, the filler region can be located before and after a region such as, but not limited to, a payload region, an ITR, a promoter region, an intron region, an enhancer region, and/or a polyadenylation signal sequence region.

[0362] In some aspects, an AAV vector can comprise two filler sequences. The two filler sequences can be located between two regions as described herein.

[0363] The present disclosure provides also methods for the generation of AAV particles, by viral genome replication in a viral replication cell comprising contacting the viral replication cell with an AAV polynucleotide or AAV genome (e.g., an AAV vector of the present disclosure).

[0364] In some aspects, an AAV particle is produced by a method comprising the steps of:

(1) co-transfecting competent bacterial cells with a bacmid vector and either a viral construct vector and/or AAV payload construct vector,

(2) isolating the resultant viral construct expression vector and AAV payload construct expression vector and separately transfecting viral replication cells,

(3) isolating and purifying resultant payload and viral construct particles comprising viral construct expression vector or AAV payload construct expression vector,

(4) co-infecting a viral replication cell with both the AAV payload and viral construct particles comprising viral construct expression vector or AAV payload construct expression vector, and

(5) harvesting and purifying the viral particle comprising a parvoviral genome.

[0365] In one aspect, the present disclosure provides a method for producing an AAV particle comprising the steps of (1) simultaneously co-transfecting mammalian cells, such as, but not limited to HEK293 cells, with a synthetic promoter operably linked to a nucleic acid, a construct expressing rep and cap genes and a helper construct, and (2) harvesting and purifying the AAV particle comprising a viral genome.

[0366] In one aspect, the present disclosure provides a method for producing an AAV particle comprising the steps of (1) contacting mammalian cells, such as, but not limited to HEK293 cells, comprising i) a synthetic promoter operably linked to a nucleic acid, ii) rep gene(s), iii) cap gene(s) and iv) a helper construct comprising AAV helper genes, with a topsentin alkaloid or derivative thereof, and (2) harvesting and purifying the AAV particle comprising a viral genome.

[0367] In some aspects, the cells can be contacted with a topsentin alkaloid or derivative thereof. In some aspects, the cells can be cultured in the presence of a topsentin alkaloid or derivative thereof. In some aspects, the topsentin alkaloid or derivative thereof is homocarbonyltopsentin.

[0368] To address limitations of the Triple Transfection system for AAV production (e.g., issues of scale due to the number of cells required, the efficiency of transfection, and the cost of raw materials), efforts have been made to generate producer cell lines that contain one or more of the vector production genes, such as *Rep* or *Cap*. A significant barrier to these efforts was the cytotoxic and cytostatic effects of constitutive expression of Rep (*See* Yang, Q., et al., J Virol.

68(8):4847-56 (Aug. 1994); Schmidt, M., et al., J Virol. 74(20):9441-50 (Oct. 2000)). To overcome this issue, a system was developed in which *Rep* RNA was alternatively spliced to generate a message that only contains a start codon upon treatment with a small molecule. The treatment of the small molecule allowed for both the timing and level of expression to result in packaging of viral vectors.

[0369] In some aspects, provided herein is a method of producing an adeno-associated viral (AAV) vector comprising: (a) introducing a first plasmid comprising an adenovirus helper gene and a second plasmid comprising a heterologous nucleic acid sequence into a host cell; and (b) culturing the host cell in the presence of a splicing modulator to produce AAV particles.

[0370] In some aspects, the host cell comprises an AAV capsid (Cap) gene and a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV replication (Rep) gene.

[0371] In some aspects, the host cell comprises an AAV replication (Rep) gene and a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV capsid (Cap) gene.

[0372] In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0373] In some aspects, the method further comprises (c) harvesting the AAV particles from the cultured host cell.

[0374] In some aspects, the method further comprises (d) purifying the harvested AAV particles.

[0375] In some aspects, the host cell is a HEK293 cell.

[0376] In some aspects, the splicing modulator is Kinetin or Homocarbonyltopsentin.

[0377] In some aspects, the adenovirus helper gene is selected from the group consisting of an Early 4 (E4) gene, an Early 2A (E2A) gene, and a viral associated (VA) gene.

[0378] In some aspects, the nucleic acid construct is stably integrated into the host cell genome.

[0379] In some aspects, the nucleic acid construct comprises a sequence selected from the group consisting of SEQ ID NO: 59-62.

[0380] In some aspects, the AAV particles can be produced in a viral replication cell that comprises an insect cell. Growing conditions for insect cells in culture, and production of heterologous products in insect cells in culture are well-known in the art, see, e.g., U.S. Patent No. 6,204,059.

[0381] The viral replication cell can be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells. Viral replication cells can comprise mammalian cells such as A549, WEH1, 3T3, 10T1/2, BHK, MDCK, COS 1, COS 7, BSC 1, BSC 40, BMT 10, VERO, W138, HeLa, HEK293, Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals. Viral replication cells comprise cells derived from mammalian species including, but not limited to, human, monkey, mouse, rat, rabbit, and hamster or cell type, including but not limited to fibroblast, hepatocyte, tumor cell, cell line transformed cell, etc.

[0382] Viral production disclosed herein describes processes and methods for producing AAV particles that contact a target cell to deliver a payload.

[0383] In some aspects, the AAV particles can be produced in a viral replication cell that comprises a mammalian cell. Viral replication cells commonly used for production of recombinant AAV particles include, but are not limited to 293 cells, COS cells, HeLa cells, and KB cells.

[0384] In some aspects, AAV particles are produced in mammalian cells wherein all three VP proteins are expressed at a stoichiometry approaching 1:1:10 (VP1:VP2:VP3). The regulatory mechanisms that allow this controlled level of expression include the production of two mRNAs, one for VP1, and the other for VP2 and VP3, produced by differential splicing.

[0385] In some aspects, AAV particles are produced in mammalian cells using a triple transfection method wherein a payload construct, parvoviral Rep and parvoviral Cap and a helper construct are comprised within three different constructs. The triple transfection method of the three components of AAV particle production can be utilized to produce small lots of virus for assays including transduction efficiency, target tissue (tropism) evaluation, and stability.

[0386] In some aspects, the viral construct vector and the AAV payload construct vector can be each incorporated by a transposon donor/acceptor system into a bacmid, also known as a baculovirus plasmid, by standard molecular biology techniques known and performed by a person

skilled in the art. Transfection of separate viral replication cell populations produces two baculoviruses, one that comprises the viral construct expression vector, and another that comprises the AAV payload construct expression vector. The two baculoviruses can be used to infect a single viral replication cell population for production of AAV particles.

[0387] Baculovirus expression vectors for producing viral particles in insect cells, including but not limited to Spodoptera frugiperda (Sf9) cells, provide high titers of viral particle product. Recombinant baculovirus encoding the viral construct expression vector and AAV payload construct expression vector initiates a productive infection of viral replicating cells. Infectious baculovirus particles released from the primary infection secondarily infect additional cells in the culture, exponentially infecting the entire cell culture population in a number of infection cycles that is a function of the initial multiplicity of infection (MOI), see, e.g., Urabe, M. et al., J Virol. 2006 Feb; 80 (4): 1874-85, the contents of which are herein incorporated by reference in their entirety.

[0388] Production of AAV particles with baculovirus in an insect cell system can address known baculovirus genetic and physical instability. Baculovirus-infected viral producing cells are harvested into aliquots that can be cryopreserved in liquid nitrogen; the aliquots retain viability and infectivity for infection of large-scale viral producing cell culture (Wasilko DJ et al., Protein Expr Purif. 2009 Jun; 65(2): 122-32).

[0389] In some aspects, stable viral replication cells permissive for baculovirus infection are engineered with at least one stable integrated copy of any of the elements necessary for AAV replication and viral particle production including, but not limited to, the entire AAV genome, Rep and Cap genes, Rep genes, Cap genes, each Rep protein as a separate transcription cassette, each VP protein as a separate transcription cassette, the AAP (assembly activation protein), or at least one of the baculovirus helper genes with native or non-native promoters.

[0390] In some aspects, AAV particle production can be modified to increase the scale of production. Transfection of replication cells in large-scale culture formats can be carried out according to any methods known in the art.

[0391] In some aspects, cell culture bioreactors can be used for large scale viral production.

[0392] In some cases, bioreactors comprise stirred tank reactors.

[0393] Cells of the disclosure, including, but not limited to viral production cells, can be subjected to cell lysis according to any methods known in the art. Cell lysis can be carried out to obtain one or more agents (e.g. viral particles) present within any cells of the disclosure.

[0394] Cell lysis methods can be chemical or mechanical. Chemical cell lysis typically comprises contacting one or more cells with one or more lysis agent. Mechanical lysis typically comprises subjecting one or more cells to one or more lysis condition and/or one or more lysis force. In some aspects, chemical lysis can be used to lyse cells. As used herein, the term "lysis agent" refers to any agent that can aid in the disruption of a cell. In some cases, lysis agents are introduced in solutions, termed lysis solutions or lysis buffers. As used herein, the term "lysis solution" refers to a solution (typically aqueous) comprising one or more lysis agent. In addition to lysis agents, lysis solutions can include one or more buffering agents, solubilizing agents, surfactants, preservatives, cryoprotectants, enzymes, enzyme inhibitors and/or chelators.

[0395] Concentrations of salts can be increased or decreased to obtain an effective concentration for rupture of cell membranes. Lysis agents comprising detergents can include ionic detergents or non-ionic detergents. Detergents can function to break apart or dissolve cell structures including, but not limited to cell membranes, cell walls, lipids, carbohydrates, lipoproteins and glycoproteins.

[0396] In some aspects, mechanical cell lysis is carried out. Mechanical cell lysis methods can include the use of one or more lysis condition and/or one or more lysis force. As used herein, the term "lysis condition" refers to a state or circumstance that promotes cellular disruption. Lysis conditions can comprise certain temperatures, pressures, osmotic purity, salinity and the like. In some aspects, lysis conditions comprise increased or decreased temperatures. In some aspects, lysis conditions comprise changes in temperature to promote cellular disruption. Cell lysis carried out according to such aspects can include freeze-thaw lysis.

[0397] As used herein, the term "lysis force" refers to a physical activity used to disrupt a cell. Lysis forces can include, but are not limited to mechanical forces, sonic forces, gravitational forces, optical forces, electrical forces and the like. Cell lysis carried out by mechanical force is referred to herein as "mechanical lysis." Mechanical forces that can be used according to mechanical lysis can include high shear fluid forces.

[0398] In some aspects, a method for harvesting AAV particles without lysis can be used for efficient and scalable AAV particle production. In a non-limiting example, AAV particles can be produced by culturing an AAV particle lacking a heparin binding site, thereby allowing the AAV particle to pass into the supernatant, in a cell culture, collecting supernatant from the culture; and isolating the AAV particle from the supernatant, as described in U.S. Patent Application 20090275107.

[0399] Cell lysates comprising viral particles can be subjected to clarification. Clarification refers to initial steps taken in purification of viral particles from cell lysates. Clarification serves to prepare lysates for further purification by removing larger, insoluble debris. Clarification steps can include, but are not limited to centrifugation and filtration.

[0400] In some aspects, AAV particles can be purified from clarified cell lysates by one or more methods of chromatography. Chromatography refers to any number of methods known in the art for separating out one or more elements from a mixture. Such methods can include, but are not limited to ion exchange chromatography (e.g. cation exchange chromatography and anion exchange chromatography), immunoaffinity chromatography and size-exclusion chromatography.

V. Host Cells

[0401] Certain aspects of the disclosure are directed to a cell comprising any of the expression cassettes disclosed herein or any of the delivery vectors disclosed herein.

[0402] In some aspects, the cell is a host cell.

[0403] In some aspects, the host cell can be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and mammalian cells. Host cells can comprise mammalian cells such as A549, WEH1, 3T3, 10T1/2, BHK, MDCK, COS 1, COS 7, BSC 1, BSC 40, BMT 10, VERO. W138, HeLa, HEK293, Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals. Host cells comprise cells derived from mammalian species including, but not limited to, human, monkey, mouse, rat, rabbit, and hamster or cell type, including but not limited to fibroblast, hepatocyte, tumor cell, cell line transformed cell, etc.

[0404] In some aspects, the host cell is a HEK293 cell. In some aspects, the host cell is derived from a HEK293 cell (e.g., a HEK293T cell).

[0405] The expression cassettes of the disclosure can be transfected into the host cells by any methods know in the art. For example, the expression cassettes of the disclosure can be transfected into the viral replication cell by electroporation, cationic lipid transfection, or calcium phosphate transfection.

[0406] In some aspects, the host cell comprises a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV replication (Rep) gene. In some aspects, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon each comprise a portion of a gene selected from

the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.

[0407] In some aspects, the nucleic acid construct is stably integrated into the host cell genome.

[0408] In some aspects, the host cell comprises a nucleic acid construct corresponding to any of SEQ ID NOs: 59-62.

[0409] In some aspects, provided herein is a method of increasing production of an adeno-associated virus (AAV) packaging gene in a host cell, the method comprising culturing the host cell in the presence of a small molecule modulator (e.g., an indole alkaloid). In some aspects, the AAV packaging gene is an AAV capsid (Cap) gene. In some aspects, the AAV packaging gene is an AAV replication (Rep) gene.

[0410] In some aspects, the host cell is a HEK293 cell.

[0411] In some aspects, the small molecule modulator is a topsentin alkaloid or derivative thereof.

[0412] In some aspects, the small molecule modulator is Homocarbonyltopsentin.

VI. Pharmaceutical Constructs

[0413] Certain aspects of the disclosure are directed to a pharmaceutical composition comprising any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the viral particles (e.g., rAAV) disclosed herein.

[0414] In some aspects, the pharmaceutical further comprises a pharmaceutically acceptable excipient.

[0415] Pharmaceutically acceptable excipients or carriers are determined in part by the particular composition being administered, as well as by the particular method used to administer the composition.

[0416] Accordingly, there is a wide variety of suitable formulations of pharmaceutical compositions comprising a delivery vector of the present disclosure (e.g., an AAV vector) or a plurality thereof (see, e.g., Remington's Pharmaceutical Sciences, Mack Publishing Co., Easton, Pa. 18th ed. (1990)). The pharmaceutical compositions are generally formulated sterile and in full compliance with all Good Manufacturing Practice (GMP) regulations of the U.S. Food and Drug

Administration. In some aspects, the pharmaceutical composition comprises more than one AAV vector of the present disclosure.

[0417] Acceptable carriers, excipients, or stabilizers are nontoxic to recipients (e.g., animals or humans) at the dosages and concentrations employed.

[0418] Examples of carriers or diluents include, but are not limited to, water, saline, Ringer's solutions, dextrose solution, and 5% human serum albumin. Except insofar as any conventional media or compound is incompatible with the delivery vectors disclosed herein (e.g., AAV vectors or AAV capsids), use thereof in the compositions is contemplated. In some aspects, a pharmaceutical composition is formulated to be compatible with its intended route of administration. The delivery vectors disclosed herein (e.g., AAV vectors or AAV capsids) can be administered by intravenous, parenteral, topical, intravenous, oral, subcutaneous, intra-arterial, intradermal, transdermal, rectal, intracranial, intraperitoneal, intranasal, intratumoral, intramuscular route or as inhalants.

[0419] In some aspects, the pharmaceutical composition is administered intratumorally, e.g. by direct injection. In some aspects, the pharmaceutical composition is administered intravenously.

VII. Kits and Articles of Manufacture

[0420] In some aspects, provided herein is a kit comprising i) any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the pharmaceutical compositions disclosed herein; and (ii) optionally instructions for use.

[0421] In some aspects, the kit further comprises (iii) a splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0422] In some aspects, the splicing modulator molecule is kinetin.

[0423] In some aspects, the splicing modulator molecule is homocarbonyltopsentin.

[0424] The disclosure also provides kits and products of manufacture comprising, for example,

(i) an expression cassette of the present disclosure
(ii) a delivery vector of the present disclosure;
(iii) a viral particle (e.g., rAAV) of the present disclosure;
(iv) a cell of the present disclosure;
(v) instructions for use; or

(vi) any combination thereof.

[0425] In some aspects, the kit or product of manufacture comprises, e.g., an expression cassette of the present disclosure or delivery vector of the present disclosure, in at least one container, and one or more containers with reagents sufficient for transfection of one or more cells or expression in one or more cells.

[0426] In some aspects, the disclosure provides kits and products of manufacture comprising, for example, one or more cells genetically modified to express one or more expression cassettes of the present disclosure, or one or more delivery vectors of the present disclosure, or a pharmaceutical composition comprising the one or more cells, and optionally instructions for use.

[0427] One skilled in the art will readily recognize that the expression cassettes of the present disclosure, delivery vector of the present disclosure, cell genetically modified to express an expression cassette of the present disclosure, composition (e.g., a pharmaceutical composition comprising n expression cassette, delivery vector, or cell disclosed herein), or combinations thereof can be readily incorporated into one of the established kit formats which are well known in the art.

[0428] Having described the present invention, the same will be explained in greater detail in the following examples, which are included herein for illustration purposes only, and which are not intended to be limiting to the invention.

VIII. Methods of Use

[0429] In some aspects, provided herein is a method of regulating expression of a protein of interest in a mammalian cell comprising administering to the cell any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, or any of the pharmaceutical compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0430] In some aspects, provided herein is a method of providing a protein to a subject comprising administering to the subject any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the pharmaceutical compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0431] In some aspects, provided herein is a method of treating a disease in a mammal comprising administering to a subject any of the expression cassettes disclosed herein, any of the delivery vectors disclosed herein, any of the cells disclosed herein, or any of the pharmaceutical

compositions disclosed herein, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

[0432] In some aspects, provided herein is a method of regulating transgene expression post-transcriptionally, comprising (a) a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) a heterologous nucleic acid sequence; and (b) a splicing modulator. In some aspect, the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene. In some aspects, the splicing modulator is Kinetin or Homocarbonyltopsentin.

[0433] In some aspects, the administration of the expression cassette, delivery vector, or cell is systemic or tissue specific. In some aspects, the administration is intravenous, intramuscular, subcutaneous, oral, intraocular, or intratumoral.

[0434] In some aspects, the administration of the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon is systemic.

[0435] In some aspects, the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon is kinetin.

[0436] In some aspects, the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon is homocarbonyltopsentin.

[0437] In some aspects, the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon is administered at least 1 day after administration of the expression cassette.

[0438] In some aspects, provided herein is a method of producing an AAV particle comprising contacting a host cell comprising rep genes, cap genes, and adenoviral helper genes with a topsentin alkaloid or derivative thereof (e.g., Homocarbonyltopsentin, Topsentin, Bromotopsentin, Bromodeoxytopsentin, Deoxytopsentin, and any combination thereof).

[0439] In some aspects, the topsentin alkaloid or derivative thereof (e.g., Homocarbonyltopsentin): i) reduces one or more viral response genes, ii) increases one or more immediate early genes, iii) increases a NF-κB signaling pathway, iv) increases a TNF signaling

pathway, v) increases a IL-17 signaling pathway, vi) inhibits apoptosis, vii) inhibits viral defense and/or vii) increases histone levels.

[0440] In some aspects, the apoptosis is inhibited by altering a signaling pathway.

[0441] In some aspects, the signaling pathway is the NF-κB signaling pathway.

[0442] In some aspects, the signaling pathway is the TNF signaling pathway.

[0443] In some aspects, the signaling pathway is the IL-17 signaling pathway.

[0444] In some aspects, the signaling pathway is the JAK/STAT signaling pathway.

[0445] In some aspects, the signaling pathway is the P38 MAPK signaling pathway.

[0446] In some aspects, signaling pathway is the PI3K-Akt signaling pathway.

[0447] In some aspects, the increased histone levels result in increased metabolic activity.

[0448] In some aspects, the increased histone levels result in increased host cell replication.

[0449] In some aspects, the host cell comprises an AAV capsid (Cap) gene.

[0450] In some aspects, the host cell comprises an AAV replication (Rep) gene.

[0451] In some aspects, the host cell is a HEK293 cell.

[0452] In some aspects, the topsentin alkaloid or derivative thereof is Homocarbonyltopsentin.

[0453] In some aspects, the host cell is contacted with about 5 μM to about 50 μM (e.g., about 5 μM, about 6 μM, about 7 μM, about 8 μM, about 9 μM, about 10 μM, about 11 μM, about 12 μM, about 13 μM, about 14 μM, about 15 μM, about 16 μM, about 17 μM, about 18 μM, about 19 μM, about 20 μM, about 21 μM, about 22 μM, about 23 μM, about 24 μM, about 25 μM, about 26 μM, about 27 μM, about 28 μM, about 29 μM, about 30 μM, about 31 μM, about 32 μM, about 33 μM, about 34 μM, about 35 μM, about 36 μM, about 37 μM, about 38 μM, about 39 μM, or about 40 μM) of Homocarbonyltopsentin.

[0454] In some aspects, the host cell is contacted with between about 5 μM and about 50 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 5 μM and about 40 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 10 μM and about 40 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 15 μM and about 40 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 20 μM and about 40 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 25 μM and about 40 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 30 μM and about 40 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 35 μM and about 40 μM of Homocarbonyltopsentin. In some

aspects, the host cell is contacted with between about 5 μM and about 35 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 5 μM and about 30 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 5 μM and about 25 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 5 μM and about 20 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 5 μM and about 15 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 5 μM and about 10 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 10 μM and about 30 μM of Homocarbonyltopsentin. In some aspects, the host cell is contacted with between about 20 μM and about 30 μM of Homocarbonyltopsentin.

[0455] In some aspects, the host cell is contacted with the topsentin alkaloid or derivative (e.g., Homocarbonyltopsentin) for about 24 to 90 hours. In some aspects, the host cell is contacted with the topsentin alkaloid or derivative (e.g., Homocarbonyltopsentin) for at least about 24 hours, at least about 30 hours, at least about 36 hours, at least about 42 hours, at least about 48 hours, at least about 54 hours, at least about 60 hours, at least about 66 hours, or at least about 72 hours.

[0456] In some aspects, the host cell is contacted with the topsentin for about 24 hours.

[0457] In some aspects, the host cell is contacted with the topsentin for about 48 hours.

[0458] In some aspects, the host cell is contacted with the topsentin for about 72 hours.

[0459] In some aspects, the host cell is contacted with the topsentin for between about 24 hours and about 72 hours.

[0460] In some aspects, the method further comprises harvesting the AAV particle.

[0461] In some aspects, the method further comprises purifying the harvested AAV particle.

EXAMPLES

Example 1: Identification of inducible alternatively spliced exons

[0462] Regions of the genome were identified that undergo alternative splicing events upon small molecule treatment, for the purpose of controlling transgene protein levels. To identify these regions, HEK293T cells were treated with two previously reported splicing modulators, Kinetin (Drug K) and Homocarbonyltopsentin (Drug H) at three different concentrations (low, medium, and high) and isolated total RNA from cells in each of those conditions. Novel splice sites close in proximity were identified in treated compared to control samples. The splice sites were used to identify exons that were alternatively spliced in the presence of homocarbonyltopsentin (Drug H)

or kinetin (Drug K) (**FIG. 2**). Treatment with either of these drugs was assessed for increase the presence of either unused annotated exons or unannotated/novel exons in coding genes. Using RNA sequencing, the presence of alternatively spliced exons was identified in many different coding genes resulting from treatment of either Drug K or Drug H. A list of candidate sequences was generated based on the size of the novel exon as well as the depth of coverage at the novel or unused but annotated exon sequence based on visualization on a genome browser. These results supported that both Drug K and Drug H can drive alternative splicing events across the human genome and that these splicing events could be utilized to control expression of transgenes.

[0463] To generate reporter constructs that express a transgenic protein only upon treatment with Drug K or Drug H, 7 total candidate genes were selected. The regions of alternative splicing in genes TRPV3, STK31, and USP25 were selected as Drug-K responsive (**FIGs. 3A-3C**) and the alternative splicing regions from genes GTF2B, LYPD3, IK, and TEX14 were selected as Drug-H responsive (**FIGs. 4A-4D**). These genes were selected for several reasons including that they had novel exon sizes of less than 150 base pairs, at least 5 splice junction exon counts in the RNA-sequencing data, as well as low standard of deviation in novel exon splice junction counts, suggesting a conserved mechanism across biological replicates. The overall design of these constructs was to generate transcripts that would be transcribed with any promoter of interest, but only upon treatment with Drug K or Drug H would translation occur, due to the introduction of a start codon in the novel exon inserted by alternative splicing.

[0464] To design the splicing reporters, the entire region of the genome including the upstream exon, the upstream intron, the novel exon, the downstream intron, and the downstream exon were placed upstream of the coding sequence of the green fluorescent protein (GFP) (**FIG. 1A-1B**). Codons were identified and substituted in the novel exon with a consensus Kozak sequence and start codon so that they would be in frame with the downstream exon and GFP. To minimize background expression, any ATG start codons were removed that could drive GFP expression without drug treatment. To reduce the size of the constructs for eventual packaging into an AAV vector, minimal versions of these constructs were also generated that only contained 200 base pair introns, with only the first 100 bases at the 5' splice site and the last 100 bases at the 3' splice site of each intron maintained. Each of these reporters was placed downstream of a Rous sarcoma virus promoter for strong expression and cloned into a mammalian expression vector that included the hygromycin selection marker.

Example 2: Small molecule splicing modulators can induce transgene expression via alternative splicing

[0465] *Materials and Methods*

[0466] **Cell Culture, transfection, and drug treatments:** HEK293T Human embryonic kidney cells (ATCC) were cultured in DMEM media containing 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin at 37 degrees Celsius with 5% carbon dioxide. For plasmid transfections, cells were cultured in 48-well plates and transfected at 70-80% confluence using Lipofectamine LTX reagent according to manufacturer's recommendation. For all experiments, 24 hours after transfection cells were treated with indicated amounts of DMSO, kinetin (MedChemExpress), or homocarbonyltopsentin (MedChemExpress). 24 hours after treatment, cells were washed with Dulbecco's phosphate buffered saline (DPBS) and imaged for GFP expression.

[0467] **RNA-sequencing:** 200,000 HEK293T cells were plated into individual wells of a 24-well plate and cultured for 24 hours. For kinetin treatment, after 24 hours of culture cells were treated with 50 μM, 100 μM, or 200 μM kinetin in DMSO, or DMSO alone as a vehicle control. For homocarbonyltopsentin treatment, after 24 hours of culture, cells were treated with 20 μM, 40 μM, or 60 μM homocarbonyltopsentin in DMSO or DMSO alone as a vehicle control. After 24 hours of drug treatment, cells were washed with DPBS, and RNA was isolated using the Qiagen RNeasy kit with on-column DNase digestion according to manufacturer's instructions. RNA-sequencing libraries were constructed using the Illumina Stranded mRNA prep using 1000 nanograms of total RNA as input. 3 replicates of vehicle-treated or each concentration were sequenced on an Illumina NextSeq 2000 using a P3 300-cycle kit. The resulting fastq files were trimmed using Trim Galore! and then aligned to GRCh38 human genome using STAR. The splice junction output files were quantified using an R script (Monteys et al., Nature 596(7871):291-295 (2021); see also github.com/DavidsonLabCHOP/Xon) to identify novel splice junctions unique to treatment with Kinetin or Homocarbonyltopsentin. Novel splice isoforms were visualized using IGV. The top candidates from either kinetin or homocarbonyltopsentin treatment were selected by identifying two novel splice events within 500 bases of one another to select and design splicing-reporter constructs, followed by manual inspection of the novel splice site RNA-sequence coverage comparing drug treated samples to control samples.

[0468] **Plasmids:** For each gene-splicing reporter, the entire genomic sequence including the Exons upstream and downstream of the novel splice site were obtained from GRCh38. To generate the full gene reporter, a consensus Kozak sequence (ccaccatgg) was inserted into the novel

splice sequence so that the coding sequence would be in protein coding frame of an eGFP coding sequence at the 3' end of the downstream exon sequence. In addition, all in frame ATG codons were mutated to minimize expression in vehicle conditions. Splicing reporter sequences were generated for the following genes: GTF2B, LYPD3, IK, TEX14, TRPV3, USP25, and STK31. To generate minimal versions of each reporter construct, the intron sequences were reduced to only 200 bases each, the 100 bases at the 5' and 3' ends of each intron, to include only the essential splicing sequences. The full and minimal sequences were placed downstream of the rous sarcoma virus promoter and then cloned into pcDNA3.1-Hygro(+) between the NruI and ApaI sites. Plasmids were transfected into HEK293T cells according to the method described above. Plasmid constructs were also generated to replace the eGFP coding sequence at the 3' end of the downstream exon sequence with a Rep78/52 coding sequence (construct V2A and V2B – see FIGs. 15, 16 and 17A-17B).

[0469] **Splicing Assays:** Total RNA was extracted using the RNeasy kit (Qiagen) with on-column DNase digestion according to the manufacturer's protocol. To examine differential splicing, 1 µg of total RNA was converted to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Thermo) according to manufacturer's instructions. For qPCR, splicing determined by either qPCR using a forward primer directed to the novel exon and a reverse primer targeted to eGFP. Real-time PCR was performed using Fast SYBR Green Master Mix (Thermo) according to manufacturer's recommendations on a QuantStudio 5 Real-Time PCR system (Applied Biosystems). The percentage of fold change of novel exon splicing was determined by the delta-delta Ct method, comparing the average Ct values of 3 biological replicates of the novel exon in drug treated conditions compared to vehicle treated conditions relative to the geometric mean of two control genes, ACTB and GAPDH. For gel analysis of splicing, cDNAs from cells treated with vehicle or drug were amplified using KAPA HiFi HotStart ReadyMix and a primer pair targeted towards the upstream constitutive exon of each gene and eGFP. The PCR product was run on a 1% agarose gel and stained with Sybr Gold to examine the size of the bands from the splicing products.

[0470] **Fluorescent Microscopy:** 48 hours after transfection and 24 hours after drug treatment, cells were washed with DPBS and imaged using and EVOS M5000 (Thermo) using identical settings for each well. 3 biological replicates were imaged for each condition.

[0471] *Drug H and Drug K Treatment*

[0472] To examine the effects of Drug K or Drug H on the designed splicing reporters, the reporters were transfected into HEK293T cells. The transfected cells were then treated with the

corresponding small molecule and monitored for the expression of GFP after treatment. An increase in the expression of GFP was observed by fluorescent microscopy for 6 out of the 7 designed constructs (FIGs. **5A-5C, FIGs. 6A-6D**), supporting that Drug K and Drug H can operate as modulators of transgene expression via alternative splicing.

[0473] The experiments were repeated with 3 biological replicates of the reporters USP25-Minimal, GTF2B-Minimal, IK-Minimal, and TRPV3-Minimal (**FIGs. 7A-7D**) and again observed induction for all these constructs, however the GFP expression seen in Drug K treated USP25-Minimal cells was significantly lower than the other reporters.

[0474] To quantify the levels of expression, quantitative real-time polymerase chain reactions (q-RT-PCR) was performed on the RNA isolated from these transfected cells, targeted toward the novel exon (**FIG. 7E**). There was a significant increase in the amount of novel exon RNA detected when comparing either Drug K or Drug H compared to vehicle treated cells for GTF2B-Minimal, IK-Minimal, and TRPV3-Minimal reporters.

[0475] *Concentration Dependent*

[0476] The expression of the GFP transgene was concentration dependent, allowing for control of dosing to modulate the levels of the targeted transgene. Cells were transfected with GTF2B-Full, GTF2B-Minimal, IK-Minimal, or LYPD3-Minimal and treated with increasing amounts of Drug H (**FIGs. 8A-8H**). Cells transfected with each of these constructs demonstrated a Drug H concentration-dependent increase in GFP expression. These results demonstrated that Drug H can modulate the protein expression level of alternative-splicing driven transgenes in a dose responsive manner.

[0477] To visualize the effect of these small molecules on the splicing reporter constructs, an additional assay was designed to examine the size of the transgenic RNAs produced in cells transfected with the reporter constructs. PCR was used to amplify the upstream exon region to GFP in treated and untreated cells, expecting that there would be a shift in the size and amount of the amplicon as Drug H or Drug K was added, due to the inclusion of the novel exon. This effect was observed for cells transfected with the LYPD3-Minimal, IK-Minimal, GTF2B-Full or GTF2B-Minimal reporters (**FIG. 9, FIGs. 10A-10C,** and **FIGs. 11A-11F**). High Expression was established for Minimum Gene L (LYPD3) (FIG. 10B), Minimum Exon/Intron Gene L (LYPD3) (FIG. 10C) and Minimum Exon/Intron Gene L (LYPD3) (FIG. 11D). Further, increases in GFP and iRFP expression driven by KT Splice Drug H sensitive constructs (e.g., SEQ ID NOs: 63-66) were shown (**FIGs. 12A-12C** and **FIGs. 13A-13D**). Drug H was also able to drive high expression

in two different minimal LYPD3 constructs driving eGFP expression (V1A and V2A – see FIG. 14).

[0478] In summary, the results demonstrate that two small molecule splice modulators were able to drive alternative splicing of specific gene transcripts. These small molecules-transcript pairs were utilized to control expression of transgenes post-transcriptionally in a dose dependent manner. Promoter-independent control of a transgene that could be broadly targeted to various cell types for broad therapeutic usage has been demonstrated.

Example 3 – Treatment of cells with small molecule modulators results in improved transcriptome for AAV production

[0479] To make sure Drug H and Drug K did not negatively impact packaging, cells were treated in the presence of Drug H or Drug K without the inducible exon constructs. However, it was surprisingly found that they had an impact on the transcriptome of cells during AAV production.

[0480] Cell Culture

[0481] VPC2.0 cells (Thermo Fisher) were cultured in Viral Production Medium (Gibco) containing 4 mM GlutaMAX (Gibco) at 37 degrees Celsius with 8% carbon dioxide and shaken at 120 rpm. Triple plasmids were transfected at 0.5 pg/cell in equimolar ratios with PEIpro (Polyplus Transfection) as the DNA transporter. Cells were treated with drug additives or a DMSO vehicle control at 4- and 24-hours post-transfection. Kinetin concentrations were 50 µM, 100 µM, 200 µM and 400 µM. Homocarbonyltopsentin concentrations were 5 µM, 10 µM, 20 µM and 40 µM.

[0482] DNA Isolation and Titer

[0483] 48-hours after transfection cells were lysed using 50 mM Tris HCl, 0.05% Tween-20, pH 8.5 and freeze thawed 3 times. To digest host cell DNA, 10 mM $MgCl_2$, 500 mM NaCl and Benzonase were added to harvested cell culture and incubated at 37 °C for 1-hour. The crude lysate was then analyzed for capsid titer by GYROS and genome titer by ddPCR.

[0484] RNA Sequencing

[0485] RNA was isolated using the Qiagen RNeasy kit with on-column DNase digestion according to manufacturer's instructions. RNA-sequencing libraries generated, and 3 replicates of each concentration were sequenced. The resulting sequencing files were aligned to the human genome and a normalized count table was generated for downstream analysis. The splice junction output files were quantified to identify novel splice junctions.

[0486] Results

[0487] Genome and capsid titers and protein levels were determined 48 hours after triple transfection as shown in Table 2 and Table 3. GFPTest is a plasmid containing eGFP driven by a CMV promoter with a SV40 polyA, flanked by AAV2 ITRs. CMVAAV1_Cap is a plasmid containing the AAV1 CAP gene driven by a CMV promoter. AAV2 Rep_Helper is a plasmid containing the AAV2 Rep and adenovirus helper genes.

Table 2 – Drug K Transfection

Condition	Plasmid	Sample Number	GOI	Capsid	% Viability at Harvest	VCD (10^6 cells/mL) at Harvest	ddPCR - ITR - Corrected	ddPCR - GFP - Corrected	vp/mL Corrected
Negative	CMVAAV1_Cap	1	GFP	AAV1	87.9%	4.35	3.40E+11	2.80E+11	2.4E+11
Negative	AAV2 Rep_Helper	2	GFP	AAV1	87.7%	4.68	4.00E+11	3.00E+11	2.8E+11
Negative	GFPTest	3	GFP	AAV1	87.9%	4.11	3.20E+11	2.40E+11	2.2E+11
Vehicle Control	CMVAAV1_Cap	4	GFP	AAV1	90.0%	4.72	2.40E+11	1.88E+11	5.8E+11
Vehicle Control	AAV2 Rep_Helper	5	GFP	AAV1	87.7%	4.73	1.98E+11	1.40E+11	5.4E+11
Vehicle Control	GFPTest	6	GFP	AAV1	88.3%	4.68	2.40E+11	1.76E+11	5.2E+11
50 µM	CMVAAV1_Cap	7	GFP	AAV1	81.6%	3.30	1.08E+10	1.20E+10	1.3E+11
50 µM	AAV2 Rep_Helper	8	GFP	AAV1	82.7%	4.52	2.40E+11	2.00E+11	3.8E+11
50 µM	GFPTest	9	GFP	AAV1	84.0%	4.38	2.40E+11	1.74E+11	3.4E+11
100 µM	CMVAAV1_Cap	10	GFP	AAV1	82.6%	4.30	8.40E+10	6.20E+10	2.4E+11
100 µM	AAV2 Rep_Helper	11	GFP	AAV1	79.6%	4.32	6.00E+10	4.00E+10	2.2E+11
100 µM	GFPTest	12	GFP	AAV1	82.4%	4.48	7.20E+10	5.20E+10	2.0E+11
200 µM	CMVAAV1_Cap	13	GFP	AAV1	75.5%	3.14	6.60E+09	4.40E+09	5.8E+10
200 µM	AAV2 Rep_Helper	14	GFP	AAV1	73.9%	3.02	7.00E+09	5.00E+09	6.6E+10
200 µM	GFPTest	15	GFP	AAV1	70.4%	2.70	6.00E+09	4.20E+09	3.8E+10
400 µM	CMVAAV1_Cap	16	GFP	AAV1	27.0%	0.89	3.20E+09	3.40E+09	-
400 µM	AAV2 Rep_Helper	17	GFP	AAV1	48.8%	1.55	2.60E+09	2.40E+09	-
400 µM	GFPTest	18	GFP	AAV1	56.4%	1.92	2.40E+09	2.60E+09	-

Table 3 – Drug H Transfection

Condition	Plasmid	Sample Number	GOI	Capsid	% Viability at Harvest	VCD (10^6cells /mL) at Harvest	ddPCR-ITR - corrected	ddPCR - GFP - Corrected	vp/mL Corrected
Negative	CMVAAV1_Cap	1	GFP	AAV1	87.0%	5.20	4.20E+11	2.80E+11	3.0E+11
Negative	AAV2 Rep_Helper	2	GFP	AAV1	86.9%	5.29	4.20E+11	2.80E+11	3.2E+11
Negative	GFPTest	3	GFP	AAV1	87.8%	5.46	4.40E+11	3.00E+11	3.4E+11
Vehicle Control	CMVAAV1_Cap	4	GFP	AAV1	88.4%	5.51	4.40E+11	3.00E+11	4.6E+11
Vehicle Control	AAV2 Rep_Helper	5	GFP	AAV1	87.2%	5.33	4.80E+11	3.00E+11	4.4E+11
Vehicle Control	GFPTest	6	GFP	AAV1	87.0%	5.17	4.00E+11	2.60E+11	4.4E+11
5 µM	CMVAAV1_Cap	7	GFP	AAV1	75.2%	4.00	2.40E+11	2.00E+11	7.0E+11
5 µM	AAV2 Rep_Helper	8	GFP	AAV1	75.2%	3.84	2.60E+11	2.20E+11	7.2E+11
5 µM	GFPTest	9	GFP	AAV1	74.8%	4.16	2.80E+11	2.40E+11	7.0E+11
10 µM	CMVAAV1_Cap	10	GFP	AAV1	68.5%	3.58	1.86E+11	1.74E+11	1.2E+12
10 µM	AAV2 Rep_Helper	11	GFP	AAV1	66.7%	3.39	1.66E+11	1.56E+11	1.2E+12
10 µM	GFPTest	12	GFP	AAV1	66.4%	3.39	1.66E+11	1.54E+11	1.2E+12
20 µM	CMVAAV1_Cap	13	GFP	AAV1	61.6%	3.10	1.20E+11	1.14E+11	1.6E+12
20 µM	AAV2 Rep_Helper	14	GFP	AAV1	60.2%	3.09	1.28E+11	1.24E+11	1.7E+12
20 µM	GFPTest	15	GFP	AAV1	54.3%	2.70	1.08E+11	1.06E+11	1.8E+12
40 µM	CMVAAV1_Cap	16	GFP	AAV1	54.7%	2.60	7.40E+10	6.80E+10	1.9E+12
40 µM	AAV2 Rep_Helper	17	GFP	AAV1	54.5%	2.54	8.00E+10	7.00E+10	1.9E+12
40 µM	GFPTest	18	GFP	AAV1	55.1%	2.79	8.80E+10	8.00E+10	2.0E+12

[0488] In this system, Drug H and Drug K's effect on AAV production was observed by detecting genome titer (FIGs. 19A-19B) and capsid titer (FIGs. 20A-20B). Effect on genome titer efficiency in normal cells showed that the effect was independent of activating the inducible system. Results (FIG. 20B) showed that Drug H increased capsid titer independent of activating the inducible system as measured by GYROS quantification. Additionally, VP1, VP2, and VP3 levels (FIG. 21A-21B) were observed for both systems. These results showed that Drug H increased capsid proteins (i.e., VP1, VP2, and VP3) independent of activating the inducible system.

[0489] Further analysis revealed that treatment with Drug H led to altered expressed of AAV packaging genes such as Cap (FIG. 22A) and Rep (FIG. 22B). Additionally, the treatment resulted in changes in the host cell transcriptome (FIGs. 23A-23H). Pathway analysis of the differentially expressed genes revealed enrichment of apoptosis, viral, and signaling pathways (FIG. 24). Additionally, inflammatory response, defense response to virus pathway, and immediate early genes were dampened as compared to previously identified genes during AAV production after treatment with increasing concentrations of Drug H (FIGs. 25A-25G – from left to right: vehicle control, 5 μM, 10 μM, 20 μM, 40 μM at 24 hours (gray) and 48 hours (green); *see also* Chung, CH., et al., Mol Ther Methods Clin Dev. 28:272-283 (2023)). Furthermore, after treatment with Drug H, there was significant alteration of JAK/STAT signaling (FIG. 28A), IL-17 and NF-kB signaling pathways (FIG. 28B), P38 MAPK and PI3K-Akt signaling pathways (FIG. 28C), apoptosis pathways (FIG. 28D), and TNF signaling pathway (e.g., TNFAIP3 and FOS) (FIGs. 28E and 27A-27C). Additionally, the response of nucleosome assembly genes previously identified during AAV product was heightened after Drug H treatment (FIGs. 26A-26E).

[0490] These results showed that many of the enriched pathways identified by Drug H treatment converged on apoptosis (e.g. FIG. 28D). Taken together, the data suggested that treatment of cells with Drug H during AAV production led to increased epigenetic response (e.g., enhancement of histone levels that increased metabolic activity to improve replication), reduced viral defense that allowed survival of producer cells, and reduced apoptosis through alteration of signaling pathways (e.g., NF-κB, TNF, MAPK, IL-7), which, without being bound by theory, may have allowed for more availability of resources for viral production. An exemplary scheme for a possible mode of action for Drug H on host cells during AAV production is shown in FIG. 29.

Example 4 – Inducible Rep System for AAV production (Prophetic)

[0491] The inducible sequences identified in Example 1 were inserted upstream of Rep and used in a triple transfection test for the packaging of AAV-GFP. HEK293 cells were triple transfected with a helper plasmid that also expresses Cap, the GFP gene of interest plasmid, and a plasmid with either wild-type Rep or a version of Rep downstream of the identified splice sequence (e.g., SEQ ID NOs: 59-62).

[0492] In this system, the ability of induced Rep to package AAV-GFP was observed by detecting genome titer (FIG. 15) and capsid titer (FIG. 16). Additionally, Rep protein (FIG. 17A) and VP1, VP2, and VP3 levels (FIG. 17B) were observed for both systems. The genome and capsid titer, along with the Rep and Cap protein levels, show that the inducible system was capable of packaging.

[0493] These results show that the tested inducible systems with Drug H or Drug K can be used to induce the expression of transgenes in a post-transcriptional manner. The system can include cell lines capable of expressing various genes necessary for AAV packaging, including Rep or Cap, which allow for increased vector production by increasing the number of cells actively generating AAV particles over transient triple transfection methods (*See* FIG. 18). The inducible system allows for finely tuned temporal control of expression to produce AAV vectors.

Example 5 – Expression of vectorized reporter constructs *in vitro*

[0494] **Small Molecule Drugs Tested**

[0495] For in vitro dosing, Kinetin (MedChemExpress) ("Drug K") or Homocarbonyltopsentin (MedChemExpress) ("Drug H") were dissolved in 100% DMSO and this material was directly added to cells.

[0496] **AAV Vector Cloning and Production**

[0497] The sequences of TRPV3 construct #3 (SEQ ID NO: 45) and LYPD3 construct #3 (SEQ ID NO: 54) were each placed directly downstream of the Rous Sarcoma Virus (RSV) promoter and upstream of enhanced GFP (eGFP). The RSV promoter, eGFP, and either TRPV3 construct #3 or LYPD3 construct #3 sequence were cloned between two AAV inverted terminal repeat sequences. Additionally, a vector was generated that contained just the RSV, eGFP and ITR sequences (Positive Control Construct). Recombinant AAV2 serotype vectors were produced with these constructs, purified, and their vector titers were determined by digital droplet polymerase chain reaction (ddPCR).

[0498] **Cell Culture and in vitro AAV Vector Transduction**

[0499] HEK293T human embryonic kidney cells, HepG2 human liver cells, U87 MG human glioblastoma cells, and Neuro2a mouse neuroblastoma cells (ATCC) were cultured in DMEM media containing 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin at 37 degrees Celsius with 5% carbon dioxide. Cells were plated into individual wells of a 48-well plate and cultured for 24-hours. After 24-hours of culture, AAV vectors containing either TRPV3 construct #3 and LYPD3 construct #3 were added to each well at a multiplicity of infection of 1e5 or 5e5 for viral transduction. 24-hours post transduction, cells were treated with DMSO, Kinetin, or Homocarbonyltopsentin, grown for 24-hours, washed with Dulbecco's phosphate buffered saline (DPBS) and imaged for eGFP expression.

[0500] **Fluorescent Microscopy**

[0501] 24 hours after drug treatment, cells were washed with DPBS and imaged using and EVOS M5000 (ThermoFisher) using identical settings for each well. 2 biological replicates were imaged for each condition.

[0502] ***Methods and Results***

[0503] To examine the effectiveness of Drug K or Drug H on AAV packaged splicing reporters, the plasmids containing TRPV3 Construct #3 or LYPD3 Construct #3 were cloned into a plasmid containing two inverted terminal repeat sequences (ITRs), the Rous Sarcoma Virus promoter (RSV), and the enhanced green fluorescent protein (eGFP) for rAAV vector production. rAAV2 vectors were generated, purified, and transduced into HEK293, Neuro2A, HepG2, or U87 MG cells at a multiplicity of infection of 1e5 vector genomes or 5e5 vector genomes per cell.

[0504] The test groups included cells transduced with no vector (Negative Control), AAV2 Positive Control Construct, AAV2 TRPV3 Construct #3, or AAV2 LYPD3 Construct #3. 24-hours after transduction, cells that received no vector were treated with dimethyl sulfoxide (vehicle), 300 micromolar Drug K, or 40 micromolar Drug H. Cells that were transduced with AAV2 Positive Control Construct were treated with dimethyl sulfoxide (vehicle), 150 micromolar Kinetin, 300 micromolar Kinetin, 20 micromolar Drug H, or 40 micromolar Drug H. Cells that were transduced with AAV2 TRPV3 Construct #3 were treated with dimethyl sulfoxide (vehicle), 150 micromolar Kinetin or 300 micromolar Kinetin. Cells that were transfected with LYPD3 Construct #3 were treated with dimethyl sulfoxide (vehicle), 20 micromolar Drug H, or 40 micromolar Drug H. 24-hours after treatment with the indicated small molecule, cells were monitored for GFP expression.

[0505] An increase in the expression of GFP was observed by fluorescent microscopy for all cell types (e.g., HEK293 cells and Neuro2A cells) after small molecule treatment in the cells

that were transduced with TRPV3 Construct #3 or LYPD3 Construct #3 (**FIGs. 30A-30C** (controls), **31A-31C** (TRPV3 Construct #3 and LYPD3 Construct #3 at 1e5 or 5e5), **32A-32C** (controls), **33A-33D** (TRPV3 Construct #3 and LYPD3 Construct #3 at 1e5 or 5e5), **34A-34C** (TRPV3 Construct #3 at 1e5 or 5e5), **35A-35B** (LYPD3 Construct #3 at 1e5 or 5e5), **36A-36B** (controls), and **37A-37D** (TRPV3 Construct #3 and LYPD3 Construct #3 at 1e5 or 5e5)), supporting that Drug K and Drug H can operate as modulators of alternative splicing of AAV vector splicing reporters in many cell lines.

Example 6 – Expression of vectorized reporter construct *in vivo*

[0506] Small Molecule Drugs Tested

[0507] For *in vivo* dosing, Kinetin ("Drug K") was first dissolved to 150 milligrams per milliliter in 100% DMSO and then vortexed and incubated at 80 degrees Celsius for 10 minutes with shaking. After shaking a final dosing solution was prepared of 60% PEG300, 5% Tween-80, 10% DMSO, and 25% phosphate-buffered saline. For in vivo dosing, Homocarbonyltopsentin ("Drug H") was dissolved in 100% DMSO to make a 10 milligram per milliliter solution. The solution was vortexed and incubated at 50 degrees Celsius for 5 minutes with shaking. To make the final dosing solution, PEG300 was added at 30% of the desired final volume, followed by Tween-80 at 5% of the final volume, and then phosphate-buffered saline at 55% of the final volume. Each material was vortexed after addition. Both Kinetin and Homocarbonyltopsentin final dosing solutions were prepared the day of oral gavage.

[0508] AAV Vector Cloning and Production

[0509] The sequences of TRPV3 construct #3 and LYPD3 construct #3 (same sequences as Example 5) were each placed directly downstream of the Rous Sarcoma Virus (RSV) promoter and upstream of enhanced GFP (eGFP). The RSV promoter, eGFP, and either TRPV3 construct #3 or LYPD3 construct #3 sequence were cloned between two AAV inverted terminal repeat sequences. Additionally, a vector was generated that contained just the RSV, eGFP and ITR sequences (Positive Control Construct). Recombinant AAV2 vectors were produced with these constructs, purified, and their vector titers were determined by digital droplet polymerase chain reaction (ddPCR).

[0510] *in vivo* Animal Study

[0511] Eight- to twelve-week-old female C57BL/6 mice at Charles River Laboratories were selected and randomized based on body weight for treatment with rAAV vector and the small molecules. TRPV3 construct #3 vector, LYPD3 construct #3 vector, or Positive Control vector was

injected intravenously into a mouse at a dose of either 1e11 vector genomes or 1e12 vector genomes. Animals were housed and monitored for body weight changes for 4 weeks. 4 weeks post-injection, either Kinetin or Homocarbonyltopsentin were given by oral gavage. Kinetin was introduced at a concentration of either 70 or 140 milligrams per kilogram of body weight. Homocarbonyltopsentin was introduced at either 10 or 20 milligrams per kilograms of body weight. Twenty-four hours after small molecule introduction, animals were euthanized, and tissues were collected.

[0512] Extraction for RNA and DNA Analyses

[0513] Whole hearts and livers were extracted, stored in RNAlater at 4 degrees Celsius for twenty-four hours, and then decanted and stored at -80 degrees Celsius.

[0514] Extraction for Fluorescent Microscopy

[0515] Whole hearts and livers were extracted and then immersed in 4% paraformaldehyde overnight. After fixation, the samples were placed in 15% sucrose in phosphate-buffered saline for two hours, then transferred to 30% sucrose in phosphate-buffered saline for two hours, followed by freezing in Tissue Tec-OCT and stored at -80 degrees Celsius before sectioning. Each OCT block was sectioned once at 8-microns thick and mounted on Super-frost slides with VECTASHIELD Hardset Antifade Mounting Media with DAPI. Slides were then stored at -80 degrees Celsius until imaging.

[0516] RNA and DNA Co-extraction

[0517] From each mouse liver or heart, 7.5 to 10 milligrams of tissue were isolated with a razor blade and placed within a Precellys tube along with 6 Precellys 2.8 ceramic zirconium oxide beads and 300 microliters of DNA/RNA Shield (Zymo Research). The tubes were loaded into the Precellys 24 homogenizer and run two times with the settings 6500 rpm, 6 cycles, 50 second cycles, 15 second pause. The homogenates were then centrifuged at 10,000 x g for 5 minutes at room temperature. 15 microliters of 20 milligram per milliliter Proteinase K was added and the tubes were vortexed and then incubated for 1 hour at 30 degrees Celsius with shaking at 1200 rotations per minute. 200 microliters of sample were then transferred to a ThermoFisher KingFisher 96 deep-well plate containing 500 microliters of Quick-DNA/RNA Magbead DNA/RNA Lysis buffer (Zymo Research). The Quick-DNA/RNA Magbead kit (Zymo Research) was used to extract the DNA and RNA according to the manufacturer's protocol with the use of the KingFisher Apex automated sample purification system. The concentration of DNA and RNA was then determined by the Invitrogen Qubit system.

[0518] Vector DNA Quantification

[0519] Total extracted DNA was diluted to a concentration of 5 nanograms per microliter. Quantitative polymerase chain reaction (qPCR) was performed using the PrimeTime Gene Expression Master Mix (IDT) according to the manufacturer's protocol using a QuantStudio 5 Real-Time PCR system (Applied Biosystems). For quantification, a standard curve was generated using plasmid DNA of the construct used to produce the vector and vector expression was normalized to nanograms of input plasmid. Taqman probes (IDT) were used to detect vector DNA.

[0520] Vector RNA Quantification

[0521] 500 nanograms of total RNA was used as input to the High-Capacity cDNA Reverse Transcription Kit (Thermo) according to manufacturer's instructions. Total cDNA was diluted 1:10 and then used for quantitative reverse-transcription real-time polymerase reaction (qRT-PCR) using the PrimeTime Gene Expression Master Mix (IDT) according to the manufacturer's protocol using a QuantStudio 5 Real-Time PCR system (Applied Biosystems). Small molecule induced alternative splicing of either TRPV3 construct #3 or LYPD3 construct #3 was determined using two custom Taqman assays targeting the alternative exon sequence compared to the constitutive eGFP sequence. The fold change of induction was calculated using the delta-delta Ct method for the average Ct values from each animal in the no drug treatment group compared to the drug treatment group. Taqman probes (IDT) were used to quantify the RNA.

[0522] Protein Extraction

[0523] Mouse liver lysate supernatant samples were prepared by homogenizing liver tissue samples in complete lysis buffer (Mesoscale Discovery Tris lysis buffer, ThermoScientific Halt protease and phosphatase inhibitor cocktail, ThermoScientific EDTA). Homogenization was performed for 3 rounds of 6,000 rpm using a Precellys 24 homogenizer (Bertin Technologies). The samples were then centrifuged at 14,000 rpm for 10 minutes at 4°C, and the supernatant was collected for GFP quantification analysis. Homogenized mouse liver lysate supernatant samples were analyzed for total protein concentration by DC assay (BioRad).

[0524] Detection of Protein Expression

[0525] The concentration of GFP protein present in homogenized mouse liver lysate supernatant was quantified using a three-step (capture-analyte-detection) immunoassay specific to the protein. All analyses were performed on Gyrolab xPlore system (Gyros Protein Technologies) with Bioaffy 1000 CDs (Gyros Protein Technologies, method 1000-3W-001-A). The capture reagent consisted of biotinylated anti-GFP monoclonal antibody (Rockland Immunochemicals, 100 µg/mL), and the detection reagent consisted of a noncompeting Alexa Fluor 647 tagged anti-

GFP polyclonal antibody (Novus Biologicals, 25 nM). A standard curve (0.12 – 500 ng/mL) was generated using a recombinant GFP protein (Novus Biologicals) diluted in lysate buffer solution consisting of 1:1 negative control mouse liver lysate supernatant and Rexxip AN (Gyros Protein Technologies). Mouse liver lysate supernatant samples were diluted 1:1 in Rexxip AN. The concentration of GFP protein in mouse liver lysate supernatant was determined by comparison of sample signal at 5% PMT to the standard curve signal fit to a five-parameter logistic curve using the Gyrolab Evaluator software (Gyros Protein Technologies).

[0526] ***Methods and Results***

[0527] To examine the effect of alternative splicing driven by Drug K or Drug H in vivo, rAAV2 packaged splicing reporters containing Positive Control Construct, TRPV3 Construct #3 or LYPD3 Construct #3 were injected intravenously into C57BL/6 mice at a concentration of either 1e11 or 1e12 vector genomes per animal. 4-weeks post injection, Drug K, Drug H, or vehicle were introduced to these animals by oral gavage at a dose of either 70 milligrams per kilogram or 140 milligrams per kilogram for Drug K or 10 milligrams per kilogram or 20 milligrams per kilogram for Drug H. Twenty-four hours after oral gavage, animals were euthanized, and the livers and hearts of the treated animals were extracted. From those tissues, RNA, DNA, and protein were extracted and analyzed for vector RNA or DNA expression, and GFP protein expression.

[0528] The total level of vector DNA within each animal in a treated group was not significantly different between animals that received Positive Control Construct, TRPV3 Construct #3 or LYPD3 Construct #3 (**FIGs. 38A-38B**).

[0529] To examine the effect of alternative splicing driven by Drug K or Drug H, the RNA expression of the alternatively spliced exon was compared to the RNA expression of GFP in each drug condition. For animals that received TRPV3 Construct #3 at either vector genome dose, there was not a significant increase in the expression of the alternative exon upon treatment with 70 mg/kg or 140 mg/kg Drug K compared to the expression of the alternative exon upon treatment with vehicle in the liver (**FIGs. 39A-39B**). For animals that received LYPD3 Construct #3 at both vector dose, there was a significant increase in the in the expression of the alternative exon upon treatment with 10 mg/kg or 20 mg/kg Drug H compared to the expression of the alternative exon upon treatment with vehicle in the liver and heart (**FIGs. 39C-39F**). These results demonstrated that Drug H can be used to modulate the alternative splicing of a target transcript in animals, but Drug K was unable to modulate the splicing of the target transcript.

* * *

[0530] The foregoing description of the specific embodiments will so fully reveal the general nature of the invention that others can, by applying knowledge within the skill of the art, readily modify and/or adapt for various applications such specific embodiments, without undue experimentation, without departing from the general concept of the present invention. Therefore, such adaptations and modifications are intended to be within the meaning and range of equivalents of the disclosed embodiments, based on the teaching and guidance presented herein. It is to be understood that the phraseology or terminology herein is for the purpose of description and not of limitation, such that the terminology or phraseology of the present specification is to be interpreted by the skilled artisan in light of the teachings and guidance.

[0531] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described herein.

[0532] All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety.

Table 4 – Sequences

SEQ ID NO:	Description	Sequence
1	TRPV3 genomic	ttttatccaagctgtgcttgtgatactgtctgtcttcttgtacttgtttgcctacaaagagtacctcgcctgcctcgtgctggccatg gccctgggctgggcgaacatgctctactatacgcggggtttccagtccatgggcatgtacagcgtcatgatccagaaggtg cgttgagagctgcgccggcgggagggccctcggaggatggccatttttcagatggggaaactgagatctggagaggaag agggtttaccaaggggcaggtcactgctctttctcccaaatgccatgcactccctttgctttcggccaggctttaagtgtcaa gaccttaaatcccccaggcttgtggggtgctggggaggggtacagtttgtggaggaatgttctcacccctgggcagtggcgt gctggtaaacatttaacgactgttctgtgtgtgtgtgtgtatgtgtgtgtgtaggtggaatatgtatattccatatatatacac aaaagacttatcatacattttactatataaacgatgtgtatcacacaatttgcaaataatgaaaatataccatgcagtttattgtaaa ttccacatagccagttgagtctcagaataatttgttgctttttttttttttttttggagaccaagtcttgctctgctgcccaggccaga gtaaagtggtgcgatctcggctccctgcaacctctgcctcctgggttcaagcagttctcctgcctcagcctccagagtagctg ggatcacaggcacgcaccaccatgttggctaatttttgtatttttagtagagacagggtttcagcatgttggccaggattgtcttg aactcctgacctcaggtgatccacctgcctcggcctcccaaagtgctgggattacaggcatgaaccaccgcgccaggccta tttgttgctttttattgaattctagtttttgttgtcaacttttggttgcaattgacaaaagagtgctatcagacatgaatgttggttgag agttttcttttataagcagtgattgaagtgaaacaattaagatttatgttgaacttttacttgtttatcaatgatgtgagcaacttctttgc tgaattggatactagtttgcaaataccagaataatattctgttgatatttttgtgctatacataatgtaaaggctgcagacagaaca cattttaaaatttaaactggattaatagtttctccatcactgtcttaaatcatttagatctagatatcaacaaaacaatgagtcaatca atctctgatttgtggcatttaccaattttcatgatgtaaatacagtagtccccccttattcttgggggagatgttctaagacccccc atgaatgcctgaaactgcagatagtaccaaacattatataatatgctatgtttttttcctatacatatacatctatgcaaataatgaaa atatactatacaatacagtttattgtaaattccacttagccagttgagtcttagaataatttttgcttttttttttttttttttgtttgtgactg agtctcgctctgttgcccaggctggagtgcaatagtgcaatctcggctcactgcaacctctgcctcccgggctcaagcaattc tcccgccccagcctcctgagtagctgggactacaggcatgcaccatgcccagctaatttttgtatttttagtagagatgggttt cgccatgttggccaggctggtctccaactcttgacctcaggtgatccgtccgcctcggcctcccaaagtgctgggggtgg ggggggaatcaacaaattattctaagactcaactggctatgtggaatttataaaatttattttatcaattaggcatagtgagatta atggcaataactaacagaacaatgataacaatatactgtaataaaagtatgtgaatgttgtctctctcagaatattgtattgtgtgt aatatttctgggtcatagttgaccacaggcaaccaaaactgcagaaagtgaaatcgaagagaagggcggactactgttctcc tctcagggccaatttccagctaccaacattatgttactggacatagagttagggagaaacatgcagatccacatcgctctaga attcctgccatatggctatgatagagacaagtaacctcaaaagcacacacaattgtaaatatagtgaaataactggacgggga tgagttttgaacatttattgtctttgttttatttaattaaattaattaattaattatttttgagacagagtctggctctgccccccaggc tggagtgcagtggtgcgatctaggctcactgcaacccccttcctcctgggttcaagcaattctcccgcctcagcctctggagta gctgggatcacaggtgtgtgccaccacacccagctaatttttgtatttttagtagagacggggtttcgccatgttggccaagct ggcctccaactcctgacctcaggcgatccacccacctcagcctctcaaagtgttgggattacaggcgtgagtcactgcaccc tggcctgtctttgttttaaatataacttattttatttaattgtaagtttatatggtttaatttttggcaatacaaataataatttctttaccc ccagctttattgaggtacagctgacatatataattgtataaatttaagatgtacaatgtgaggatttgatataggtatatattgtgaa atgatttccaatttccaattaacacacccatcacctcacatagttgccatttctttgtgtgtggtgagaacatttatgatctattctctt agcaaatttctagtacacaacacggtattgttcactatagtcatcatgctgtccattagatccccagaacttattcatcttattcatt gatttagttctgaataatggctgtgtttaacaaccagctcacaagcttcatgagaatttaacgatcagctcttgcgagctcttctg agctggctccagcatgccactgcccccagcaagtgtgaggattccagtgtgagaccctatggtgaaaggtacagggtgaat cacagggagaagagaacctggggctggagcagaaacagaaggcttcctggaggaggcaggatccagtgttgagccttg gagcatcaggaggatttggctgggggaagggagagtggcccaagtgaagagacagttagggtgggggtgggagagg gtgggaagcaggaagatcagggtgccgtgaatatgtgcttgcaccttacacttgaattatgtttatttccttataggtcatttttgc atgatgttctgaagttcttgtttgtatatatcgtgtttttgcttggatttggagtaggtaattgatttataaataatagtaactcccacat ttataaacactttatacagatgtcaaaatggcaaagtgcatggcggccttggcaacaattagaccaactaccagtggggcact tgccatgcttggcattttccaagcattacctttatatatatacatatatattttattatactttaagttctagggtacatgtgcacaaca tgcaggtttgttacatatgtatacatgtgccatgttggtatgctgcacccattaactcgtcatttacattaggtatatctcctaatca agcattacctttttacttctctcaacaacttgatcaagtgggaatcatcaccccgatcataaggtaactgaggcttagagactcc aggattcaaaccaggccagccactcataatccaacttcaccatgtggtatatacctctggctttcccattcattgcttacgtggc cctcacgacctaagataggttctgcaagtgttttcagccccattttaacagatgcggaaactaaagtccacagaaactcaactg atggtctttgcaaagtaggagcaaactaatacctgaggtgatttccttggatttcacatcacgttactctcacagtttgctttag atacaactgtgagagcctctctggtctgctgggaccaggaaggagccccctcttaaaatctctccatttacaacagagggaa cagcacgatctggttgtatttctaacctctgcatagagagatatcattttccccacgtaccagttgctttttttttcttttcttcagtga ccattgacatttgaaattaagtagctgtttccatctgtctggtacctttttctgggtcacaatttgcttatttatttgctccttcacctgct tgtttattcatccatctatttactctctcatttaatttgctcatccacctgtctatccgtccatccatccatccatccatccatccatcca tccatccaatcattcatccatcgatccatccattgatctatccatctatccatccatccatccatccatccatccatctaatcatcca tccaatcatccatccatccatccatccatccatccatccagtcatccatccattcatccatttaatcatccatccatcgatca atcggtccatccatccagctatccatccatccatccatccatctaatcattcatccatcaatccatccatccatccatccatc catccatccatccatctatccatccatccaatcatccatccatccatcgatcaatcagtccatccatccatctaatcatccatctaa

		tcatccatccatccatccatccatccatccagtcatccatccattcatccatttaatcatccatccatccatcgatcaattggtccat ccatccatctatccatccatccatccatccatctaatcacccatccatccatccatccatccaatcatctatccatccatccatcta atcatctatccatccatccatccatccatccatccaatcatccatccatccatccatccatccctccatctaatcatccatccctcc atccatccctccatatccatccatccatccatccatccatccatccaaccattcagcagtgaacctctactgtggtgtgttggaa agaagggaagtgtgagcctggtcttgctcacagggaactcaaaacagagcattaggaggagactcaggatcttacagcctt gagggaggtgacccaggctagaggaagtcaggagctcaaagactagctggggagccagggacacacacatggattata attagtggctctgcaaggctgcaagtgatcaagggcagactgcggggtaaaggaaggagagaagactctgggctttgttag gtggaagggctttctgcaggaggccaggccagagaagccagcctgatgaggctttgagaggacctgaaactggagatgg gatttaatggccagagagaaggaggcagcaggcttgtcatgtaggcaaatacacgcaagcttgtagagtaattaccaggca gagagctgaggagagcaccctgttatgttcccttagccttggcctcgctgatcgagaagtgtcccaaagacaacaaggact gcagctcctacggcagcttcagcgacgcagtgctggaactcttcaagctcaccataggcctgggtgacctgaacatccagc agaactccaagtatcccattctctttctgttcctgctcatcacctatgtcatcctcacctttgttctcctcctcaacatgctcattgct ctgatgggcgagactgtggagaacgtctccaaggagagcgaacgcatctggcgcctgc
2	STK31 genomic	aactctggtggtctccttacaatgagcttggaacgagatcttcttgatgctgagcccatgaaggaacttagcagcaagcgtcct ttggtacgttctgaggttaatgggcagataattctgttaaaggtaagtctaacttcttcttacttgtgggagattcagcagcataaa ggatagtggtgctgttacatttcaagaaagaaaaaacctcactgttagttttctaaagtgtataagttgactggcatgataattgt gcttgtagtttggttgccagaactaattttcacttttaaagtttccctcaattgaacaatgagaacacttggacacagggcgggg aacatcacccgtgggggcctgtcatggggtgggggcaggggagggatagcattaggagaaatacctaatgtaaatgat gagttaatgggtgcagcaaaccaacatggcacatgtatacatatgtaacaaacctgcacgttgtgcacatgtaccctagaact taaagtataataataaaaataagtatatataaagtcattaaataaagctttgtctcacaagaaaaaaaaataaagtttccctcatttc cctaaatagaatatgtccttaataggatatacttccttttttttgggtgaacttattatagtcatccaatatctttttaaaattcttgatt ttcttacatttctccttgaaatctcagctgtgttatcagtacaatcttctactctttaatttcttcttctcattgatctgggaaactgttaa tcatgataataagataattaatttggccttagttgctacacattctagaatgtgtagcaggctacaggtagaattgcactttcctttt ccaaattatagcattcaaagaagatagaaaagtacttataaaattaaaaaataaaattagatttgatttaacttaaaataatgaatt atgtttaaattggaatgatgtataattataatgagattttcatcttggaattacgagtgcctctttctttggtacaagggctattctgt ggatgttgacacagaagccaaggtgattgagagagcagccacctaccatagagcttggagagaagctgaaggagactca gggttactgccattgatattcctgtttttatgtaaggtaaagttcttatgctaatctcttgcagaaaccatttttatcttgcagaaactat tttatttaaaataaatttttagctccaggttttagagcagcagttcattcccctaacataaaagaaaaattatcccatagaatcaatg tatgctaaaaatatgttgaagaagtttttacttggagtcacattctctgttttgcttcttagtctgatcctatggcttatctgatggtcc catactaccctagggcaaacctgaatgctgttcaagccaacatgcctttaaattcagaa
3	USP25 genomic	gtgtcacaacataatgaaagttgctcaagccaaactggaaatgataaaacctgaagaagtaaacttggaggaatatgaggta atgtgctttcttagaggttaaactacagatttagggattctgtaattcactgatattccttcgtttcagcccaactgaagttgaattt gttatgaatgatgagtaaaaaaattaagtttccaagtcaatgctcaagtaattaataatttttatatgtgattttaggattaacatgata tgatagcatgtagaccaaaaagtttaaaatcaaaaaaggaagatttttggttttttatttcaatactaattctcaagtacttttttaaat gaacttttctagcttaaaaacagaataacaaatgcattgttaatttttgaaatcaacttagaaattgctgttacataataccaacaag taatatttatcaagaccctactatatgccaggcatttgtctaagccctttacatatattaactgatttaattcagacagtaacactgt aagataggtagtgttcttattcccatttgacacacgagaaaataaggcatagagcagtgaagtaacctgcctaaggttttcaca gccacaagtttacagcttggctaatagaacctacattcttaactgttgcactacactgtgcaactgtaccctgtgggctgtatttttt tccttcactttgattcatcagatatattttggggaagttaattcattttagtccctatatgccatatatgtttgtcatgagacagtgttgt gtgtgagagagcacttacataggtacttgtacatgtgcttgtgggatttttttgtgcacaaaattatttttttgtcaactttggctttaa aatattctttattttaattgatgaattcatacacctagtatttcttcacatttttaaatttaatcttatgcattgtctttttttttaatttataagt ctagtaaattttaagttgttttcttgggtgcctttgcttaatatttagttccatttacaaacttaattatggtactttagattccctaaata attagtgtaattcactacacaaacagggaagaggaaagagagagaaaacactcattattgatgaaaaccttagctagtttggc gtataaggtaactaccttaattcactaaagagtgtctctaaaaacaaaaatagggaaataataaaactgcagcaaatatatagaa ttgtgagcactgaaagcttaaaatttttgttgtcaaaagcaaggcaatgaaaataagcgaaaagtaaaatactggaaaagaag aaacaaagacatcattatttgtagttttttttaccttgaaattcaaataaaatctacaactgatgtattagaactagtaagacagcttt gcagcatgcaaacttgtgtcagacatcaattttattccctttgccagtaaaaaaagtatgagaaagtataaaagtgagaccatt tacaatagtatctgaaactatggagtatatagaaataaatttaacaaaaggtgcgtgtgacttctgtggggaaaattataaaactt attgggagacattaacaaagacctaaataaatgggcatacatatggtatttggaaaacttgatattattaatatatagatatcactt ttccccaaatgctttataaattcagtgcaggctgtatgaaaattccagcaagatgttttatggagcttgataagccaactgaaa aatgttacggaaatgcaaagatcccagtccacctcccatctctgacaagcaagttagtgaacagtataaagacttttttaataaat ggttctgggaaataaataaaatttgatttcttctttattccatatgcaaaaattcatttcaattgtgaaagactcctgtgtaaagctgt ttagaagtcaacataggtagaagcccaaagacattcatttatccataagaggcactcactgaccagatatttttgaccaggata tgtaatctagagccagagaaattcaaaatttgttttctcttgaccctgaggagcaagtcaagtcatctttagctaataatgactcc aaatctgtgcctccacattgattgagagctcttcccttgtccccaagattcgtgttcatggaacttgatttttaattgtgacttgatt ctgcaatgcttccttcagtttgcttcaacagattggactcattatgttttttttcataaaatcttctcatattctaactgcagggctgtta aatcatacaagtatttaaatgacttgtttttagattctgcatattcctcctatctttgtcttgctctaagttcagggttcagtaaaccat agcctgcagaccagcttaccacctgtttttgtaaataaacttttattggaacatggccatttgtatacatattgtctgtggatacttt cacactgcagttacagagttgagtaactgaagaagatgaggcttgcaaagtctaaaatattatctggccttttacagaaaaagt

		ttgcctactcacaacttctaaggtgattctaatgacaacaaggaatgacagtttgaataccttccccagaaaagctaacttcata cttgcaggacaaattttacctttatttgttatgaagatctactaatagcatgaaaattatgtcttgaaaattaaaattaagactggat gtggtggctcatgcttataatcccagcacttttggaaatcaaggcacagggagttgtttgaggccaggaatttgagaccggcc tggccaacatagcaagaccctgtctatacaaaacagattaagtaattaactggaactgggagcatgcacctgtaattcaagct acttgggagacagaggtaggaggattgctggagcccagaaggtcaagtctgcagttagccgtgatcgtgccactgcactg cagactgggcaactgagtgagaccctgtctccaaaaaaaaaaaaaaaaagtttatattcaaggagaatatcgttggctaataac tatgtttcttatttttgtttataatgtttatataaaagccttaacacttttgttatcagtattcactttccaatttatcctacattgaatttcag ttatttttcttgagtcaattattataaagtaagctaagaaaatcttgactttggtatcaccttagaattatttagtcctattacatgccct atagttgttttgggggtggggaggggcattgaaatgtgaactgcagttattcttttaccatatctcatgtaaataaactattcgcct acacacttaaagtcaaatttaccacatatttgcgaagttaaaaggttaaataaatctaacttttttttagtttcttaaaattatgtattaa gtatgtgtccaagctttgattcatactcagcgtaatatttcattttcttccatcagttgatgtagatataataatagttcaataatatga ttttatactgatgaatttttgagatgacctgattgattatcctgccttttatatttctgttatctgctagtaaaatctttattgtttaaaacat ttcatatccatgcctaataatttttagatttaatgttgtattaataaatagtcttgtgtgccagaaagtaaagtcttgttttagttctgat gtatcttcattttgaactctttatttctattaattcagatagtagcgaaacagtgcattacttttgtacagtttcatagtacttttttaaat cttactctgaacatttttaatatttttctttcacctacaggagtggcatcaggattataggaaattcagggaaacaactatgtatctc ataattgggctagaaaattttcaaagagaaag
4	GTF2B genomic	gatcgaacaaataatatttattcaagcaagtatatgaacagaagagcctgaagggaagagctaatgatgctatagcttctgcttgt ctctatattgcctgtagacaagaaggggttcctaggacatttaaaggtaagtttttagctttgttttttgtgtttgtttgtttgtttgttt gagataggtctcactttgtcgcccaggccagagtgcagtggcgcaatcttgactcactgcaacctctgcctcccaggttca agctatactcctgcctcagcctcctgagtagctgggattacaggcatgtgccaccacacctggttaacttttgtatttttagtaga gacagggtttcgccatgttggccaggctggtctcaaacttctgacctcaggtgatctgcccactttggcctcccaaagtgcta gaattaacaggcgtgaggcactgtgcccatcctaaaataacattttctacatcatattaaagaaacagaccaatgtcaggcag ctttgacacttagaagctaagtcatctttattgtgactacagtaatccacttgagcatcatttgccaatgcctacttatttcattcctt gacagcaaattcatagcatgttcttttaggctgttttcccatgtagactcattttagtccacgtaagcccactatgttcttttctagtg agatcttttgatggactgttgctttgtgggctttcagaaggcctataatcgtaaaccaaagtaggccgtttatacatttggtttaca tggaaacagtttccatttctagcactggtctctgtagatgagggagtcaagctttagaaagtatcagtgaaggccaggtatggt gtctcacacctgtaatcccagaactttgggaggccgaggtggttgggtcatgaggccaggagttcaagaccagcctggcca atatggtgaaaccccgtctctactaaaaatacaaaaattagccaggcgtggtggtgcgcgcctgtagtcccagctactcagg aggctgaggcagaagaatcgcttgaaaccgggaggcagaggttgcagtgagccgagatcacgccactacactccagcct gggcgacagagtgagacttcatctcaaaaaaaaaaaaaaaaaaagttatcagttaagttttttctcttgaggtgtattgtttgcact tttacacttcagttaagaactggtatttttcatttagtttgcagttacatgtgaatatgtgataatttgttcagtgttgctcactgacttc cactaccaaagctcagtttagctgaagactatacttaacccttattcatccagtgagtatttatttttatgttttcttacgtaaaagt gaagaggtgatggcctttatcatcatctttttgttgttttataactaaaacaacttttgttaaatttcttctgaaagtgtgtggtggtt gcatctgggaatagcagggtcagatattttatctcattgagctacagcccataatcgcagggatttcatgcacacatgacattgt gttgcctaagaggcatcctttagcttgagtattttctttcaattccttggtaacactgcttacaacagattaatacataaacatcaaa gaatttttcatttaaagaaatttagttgtatgaattaatccttagagcagttaatattttgaaagattaagacctaaatttagccaggc acgatggctcacaccggtaatctcagcactttgtgaggccgaggcaggtagatcacgaggtcaggagttcaagaccagcct ggccaagatggtgaaaccccatctctactaaaaatacaaaaaattagccgggtgtggtggcaggcgcctgtaatcccaacta ctcgggtggcaggagaatcgcttgaacccaggggcagaggttgcagtgagccaagatggtgtcactgcactccagtctg gacgacagagtgaaatttcatctcaaaaaacaaaaaaactctaaattgaaagagtgaagaacattaggccaggcgcagtgg ctcatgcctgtaatctcagcactttgggaggccgaggtaggaggatcacctgaggtcaggagttctagacaggcctggcca acatggtgaaaccccatctctactaaaaatacaaaagttagccaggtgtggtggtgcacgcctgtagtcccagctactcggg aggctgaggcaggagaatcacttgagcctaggaggtggaggttgcagtgagctgagatcatgccactacccttagtctggg cgacagagtgagacttcatctcaaaaaaaaaggtgtgatgaacattaaagttaactgagaaatttgcagttaactctgtggga agtatcactgagatgattagggtttcatctttaaggtatgattgaatgttttacaattagaatgtatattatgtttataattagaaatgg tagattttaagatgaactatatggcttaagtgaattgatttatttaatttttttttctttttagaaatatgtgccgtatcacgaatttctaaga aagaaattggtcggtgttttaaacttattttgaaagcgctagaaaccagtgtggatttgattacaactggggacttcatgtccag gttctgttccaaccttttgtcttcctaaacaagtacagatggcagctacacatatagcccgtaaagctgtggaactggacttggtt cctgggaggagccccatctctgtggcagcggcagctatttacatggcctcacaggcatcagctgaaaagaggacccaaaa ag
5	LYPD3 genomic	atggaccccgccaggaaagcaggtgcccaggccatgatctggactgcaggctggctgctgctgctgctgcttcgcggagg tgggtctgctgcccggagaggtggaggggtagagtgggggctcttagggaggaggctcctctggccccgtacagagaa ggctgtgggaaggatccccacataagcagagggcactgtgggcaaagttcccccatcctaagcaagggaggcagaggg aagggtcctccaccatgggccgtaggaagggttccccaccacggacaaggggagctgtgggcaaaattttcctatcctgaa gctgcagaaagaagtcccctgttcctaggccgggcgcggtggttcatgcctgtaatctcagcactttgggaggccgaggca ggcggatcatgaggtcaaacgatcgagaccagcctggccaacatggtgaaaccctgtctctactaaaaatacaaaaattag ctgggcgcggtagcgcgagcctgtagtcccagctactcgggaggctgaggcaggagaatcgtttgaatccaggagggga ggttgcagtgagccgagattgcgccactgcactccagcctggcgacagagcgagactccgtctcaaaaaaaaaaaaaag aaaaaaagaaaaaaagaaaaaagtgttccctgtcctgggcagagaagctgtgggtagtgttctctcatttgagcagaaggg

		ggctgtgggaagggatccctcactgagcacagtctactgtgggaaaagttctcccacacactgtgggaaaagttctctcattc tgagcagaaaaccgtgaggaggtatccttcctaccccgggatgaggggcttcatggcggggatcgcctacagtgggcata cagggctgtcggaaggatgtgcccacagttggcatagtgggtgtgaaggaaaagggtccagtgattcccgtgcttccagat ttgtggaaagcttttccctactcctgccctggaaaagtgtgtgcaataagactgtctaatcccgggacagagggcatgggaa ggactatttctccttgggggtggacgctttagcaaggagggctcaggtctattcggctccgctccccaggagcgcaggcc ctggagtgctacagctgcgtgcagaaagcagatgacggatgctcccgaacaagatgaagacagtgaagtgcgcgccgg gcgtggacgtctgcaccgaggccgtgggggcggtggagaccagtgagtggggacccccgggccacgcacaccactgc aggctcctccggcccagccccgcctccttatagtcctttctggatctggccacgcccctttaacgcggccccaccctcccttt cagcctggccgggaccctacgacacagccttgcaggccaggccttctctattaacaattacaagcgcgaggagcagagag ttgagaactgggtggggcggacggggagagctcgggtttggaacgaggtcacgtcctatgcacttttctgctccaccaatg agacctccccatccctgcctgtcaaccaggctgtctcctccgactcagctctttgaggttccgcccccgttccgtgaccccgc tcctccctgcttcggtcgttctttccgctctgagcggctctttgggggcgtccatccgccaattcctgcctgtcggtcctgtggt gttggcctatttgaggccccgcccctcccagaccctacccaactctgggccacgccccatctgagaccacgcccacagcta atcccgcccttgtccctctccctagtccacggacaattctcgctggcagtgcggggttgcggttcgggactccccggcaaga atgaccgcggcctggatcttcacgggcttctggcgttcatccagctgcagcaatgcgctcaggatcgctgcaacgccaagc tcaacctcacctcgcgggcgctcgacccggcag
6	IK genomic	gtgagccgttctccaaccctttggcccccgatggccacgatgtggatgatcctcactccttccaccagtgagtattttgtgcta ggaccatggaaatgagttgggcaatgaatagaaaagtggtggtgatggtgaaagaattctgaatgtcttgtcatcaaggatg gtaaaaaattatacccgctggccctttatctctgtgagccttaactggatcaggccaagtgcataagggtattagaagggaaat tcctgcaggaaaggtgttaagttttttttttttttttatacgtgttgtgtaattacacattatataggaattacttcttccaatagcaaatta aaacataggaaataagggtaaaatcatcttttatttttttgagactgagtttcgctcttgttgctcaggctggaatgcagtggtg cgatctcggctcaccacaacctccgcctcccgggttcaagcaattctcctgcctcagccttccgagtagctgggattacagg catgttccactacgcccggctaattttgtatttttagtagagatggggtttctccatgttgatcaggctggtcttgaaatcccgact tcaggtgatccacccgcctctgcctcccaaagtggtgggagtacaggcgtgagccaccgtgcctggcccccgataaaatc atctttgataacatccccaattccagtccatttctttttttttttctttttctttttctttctttttttttttttgagacggagtctcactgtt acccaggctggagtgcagtgatgcaatcttggctcactgcaacctctgcctcctgagttcaagcggttctcctgcctcagcct cccaagtagctgggactacaggcatgtgccaccatgcacgactgatttttttttttgtatttttagtagagacagggtttcactgtg ttcgtcaggctggtcttgaactcctggcctgaagtgatccacccgcctcagcctcctaaagtgctggcattacaggcgtgagc caccgtgcccagcctcagtccttttcttagaagtaaccacagttacctttttggcatatctttcctcccgttgcccaggctgga gttcaatggcattatttctgctcacagcaacctttgcctcctgggctcaagcaattcttctgcctcagtttccccagtagctggga ttacaggcatgccccaccatgcttggctaatttttgtattagtagagacagggtttcgctatgttggccaggctggtcttgaactc ctgacctcaaatgatctgcccgtctcggcggcccaaagtgctgggattacaggcgtgagccactgcgcctggccctcctact ctttttaaatgcatttatacacatacatgcatatctgaggaaaacacaacattgctttgtgatttaaaatatttgagttaaatgatatt atgctgtatatatcatttttgcaacttgcttctttttaaaccctgctcagtgtttctgagattatctttcagagcagtgacaagcattcc ctccttggctaaactttagtcaggttccccaaagagtcccatttttcagcaggactcatccgtggcctgctgagcacagttttaa caaagaaatcaaccaccgctccacccccatacctaaccaagttgcttttagtaatttctataaattctcacttgtccctgttgtatt cggaattaagttcagtctctctactctattacaatggtcttgactcttgttacagtagtcttcaataaagtcttccttggggagaag aggaacttgattagatatcaagttggattggatatcaagagtgggaaaactttctgtgaactgacttaggattccttttttaaaact gaggtcagcaggccaaggaagagccctagtcagaagagagctttaaaggagtctgactttaatttgatcaagagaggagtc tttgttagtgtaccaattaaaataatctcttgtaccattagtgctgtctttccctctttgggaaacaccatcatagtttagtcagtctc agctaataaacgcttagtctgtttcagactgattcttttctttttttattttttgagacagagtcttgctcttttgcccaggctgga gtgtaatacctagatctcggctcactgcaaactccgcctcccaggttcaagcgattttcctgcctcagcctcctgagtagctag gattacaggcgcccaccacaacgcccagctaatttttttgtatttttagtagagatgggatatcaccatgttggccatgctggtct tgaactcctgacctcgggtgatccgcctgcgtcagcctcccaaagtgcggcgattataggtgtgagccactgcacccggcc agactgattattttcaaggctataactgaattaaatttacaggccttcttatatacatatgtcagtaatggaccaggctaaatataa ataaatggagttgctggctcataggatatatatacttaaatttttaacaaatactgcaaaattgcccttcaaattggttatactggttt acattctcatcagcattatatgagagagtccccttcccaaattctcactcacacttgatttatcaaactttctggtagttaacactct gatgcattaaaagtggcatggcatatctggtcgggcgtggtggcttacacctataatcccagcactttgggaggctgaggtg ggcagatcacctgaggtcagaagtttgggaccagcctggccaacacctctaataaaaatacaaaaattagctgggtgtgttg gcgggcacctgtattcccagctacttgtgaggctgaggcaggagaattacttgaacccaggaggcagaggttgcagtggg ctgagatcatgccactgcactccagccggggcaacagagggagactctgtcttaaaaaaaaaaaaaaaaaagctgggtgcg gtggctcatccctataatcccagcactttgggaggccgaggtgggtggatcacaaggtcaggagttcaagaccagcctggc caatatgggtgacagtgagactccgtctcaaaaaaaaaaaaacagtgtcatatctttgttttaatttgtatttccctgatttctgtta atttaatcatgatttttacttgttccttttttcatttcttattgagtcctaatatttaaaatctttgccttttcttctagatcaaaactcaccaa tgaagacttcaggaaacttctcatgacccccagggctgcacctacctctgcaccacctttctaagtcacgtcaccatga
7	TEX 14 genomic	tggaggtgaaagaaatggcaaagaaagcagctactggacagctcacagtacctccttggcatcctcagagtagtctgacttt agagagcgaggctgaaaatgagcccgacgccctgctgcagccccccattaggagcccagaaaacacggattggcagcg agttattgagtatcatagggaaaatgatgagcccagaggaaatggcaagtttgacaagacgggcaacaatgactgtgacag tgaccagcatggcagacagcccaggcttggaagcttcaccagtaagttcacttgcgggaatcaggataggtgaactcaatc

		aagtatcagtcgggcgaggtggctcagcctgtaattccagcactttgggaggtgaggcgagcggatcacctgagcccagg aatttgggaccagcctggtcaacatgatgaaacctggtctctacaaaaaaaaaaaaaaaaaaaaaaaaaacattagctggg cgtggtggcacacgcctctagttccaggtacccaggaggctgaggtgggaggatcacttggacctgagaggtggaggttg cagtgagctgtgatcatgaaactgcactccagcttgggtgacagtgagaccctgtctcaaaataaaaataaaagtaaaaaaa agttatcagctctcaaaaaagtaaactttccagcatatttttttggtttggtcctctctccttcctactttctggaaaccttgaattc caggggaagacagtatagccctatatctcattgcatttcccacaggcaaacagatccagacactgatcagtttttatattactga aacaggtccctttgtccaaaggaagtaaaataatcaaatactgaccctggttttggcgaggaatcgatccttctttacctgggg atagaggtaggccatcactgcctactgcccatgtctctgcattgtattctaagctgaaattaaattctcccagaatttatttagaat gtattctaagctggcattaaatttttccagagagtctatgctggaggaacctagagacacatgacatttattgtcatcggtttata gatttacaggttaaaaaaaatgaaggctgtatattttcagttgataatttgcttataatcagggcatggatattgtttcacagttgttt ataagttaaagaaagatacatctaaacttgtcatttgtcttaggcagaaaatctattggcttgggtattatttgtatttagaatctgtt tcttgtcaagcatcactctcgaatgggtgggcagatgatccagggctgcattttggggtcaggctaaactaacagtctctactc tctggactgaacctaacgtatgggacatggccactcttggcagaatatgtatacatctcttctttcatgccacctttatcctcttcc atgttgtctgaatatcactgaatgactttgaccctgactgtgttacaatcattaggtatcaggcacccatctcccagacaaaagg agcaaccagagcatagtgaagccttccaagcaagttctgacacattggtggctgtagagaaatcttacagt
8	TRPV3 upstream exon	ttttatccaagctgtgcttgtgatactgtctgtcttcttgtacttgtttgcctacaaagagtacctcgcctgcctcgtgctggccatg gccctgggctgggcgaacatgctctactatacgcggggtttccagtccatgggcatgtacagcgtcatgatccagaag
9	TRPV3 upstream intron	gtgcgttgagagctgcgccggcgggagggccctcggaggatggccatttttcagatggggaaactgagatctggagagg aagaggggtttaccaaggggcaggtcactgctctttctcccaaatgccatgcactccctttgctttcggccaggctttaagtgt caagaccttaaatcccccaggcttgtggggtgctggggagggtacagtttgtggaggaatgttctcacccctgggcagtg gcgtgctggtaaacatttaacgactgttctgtgtgtgtgtgtgtgtatgtgtgtgtgtaggtggaatatgtatattccatatatat acacaaaagacttatcatacattttactatataaacgatgtgtatcacacaatttgcaaataatgaaaatataccatgcagtttatt gtaaattccacatagccagttgagtctcagaataatttgttgctttttttttttttttttggagaccaagtcttgctctgctgcccaggc cagagtaaagtggtgcgatctcggctccctgcaacctctgcctcctgggttcaagcagttctcctgcctcagcctccagagta gctgggatcacaggcacgcaccaccatgttggctaatttttgtatttttagtagagacagggtttcagcatgttggccaggattg tcttgaactcctgacctcaggtgatccacctgcctcggcctcccaaagtgctgggattacaggcatgaaccaccgcgccag gcctatttgttgctttttattgaattctagtttttgttgtcaacttttggttgcaattgacaaaagagtgctatcagacatgaatgttggt tgagagttttctttataagcagtgattgaagtgaaacaattaagatttatgttgaacttttacttgtttatcaatgatgtgagcaacttc tttgctgaattggatactagtttgcaaataccagaataatattctgttgatatttttgtgctatacataatgtaaaggctgcagacag aacacattttaaaatttaaactggattaatagtttctccatcactgtcttaaatcatttagatctagatatcaacaaaacaatgagtc aatcaatctctgatttgtggcatttaccaattttcatgatgtaaatacagtagtccccccttattcttgggggagatgttctaagacc ccccatgaatgcctgaaactgcagatagtaccaaacattatataatatgctatgtttttttcctatacatatacatctatgcaaataat gaaaatatactatacaatacagtttattgtaaattccacttagccagttgagtcttagaataatttttgctttttttttttttttttgtttgtg actgagtctcgctctgttgcccaggctggagtgcaatagtgcaatctcggctcactgcaacctctgcctcccgggctcaagc aattctcccgccccagcctcctgagtagctgggactacaggcatgcaccatgcccagctaatttttgtatttttagtagagatgg ggtttcgccatgttggccaggctggtctccaactcttgacctcaggtgatccgtccgcctcggcctcccaaagtgctgggag gtgggggggggaatcaacaaattattctaagactcaactggctatgtggaatttataaaatttattttatcaattaggcatagtga gattaatggcaataactaacagaacaatgataacaatatactgtaataaaagtatgtgaatgttgtctctctcagaatattgtattg tgtgtaatatttctgggtcatagttgaccacaggcaaccaaaactgcagaaagtgaaatcgaagagaagggcggactactgt tctcctctcagggccaatttccagctaccaacattatgttactggacatagagttagggagaaacatgcagatccacatcgctc tagaattcctgccatatggctatgatagagacaagtaacctcaaaagcacacacaattgtaaatatagtgaaataactggacg gggatgagttttgaacatttattgtctttgttttatttaattaaattaattaattaattaattttttgagacagagtctggctctgccccc caggctggagtgcagtggtgcgatctaggctcactgcaacccctcctcctgggttcaagcaattctcccgcctcagcctctg gagtagctgggatcacaggtgtgtgccaccacacccagctaatttttgtatttttagtagagacggggtttcgccatgttggcc aagctggcctccaactcctgacctcaggcgatccacccacctcagcctctcaaagtgttgggattacaggcgtgagtcactg caccctggcctgtctttgttttaaatataacttatttttatttaattgtaagtttatatggtttaatttttggcaatacaaataataatttctt taccccccagctttattgaggtacagctgacatatataattgtataaatttaagatgtacaatgtgaggatttgatataggtatatatt gtgaaatgatttccaatttccaattaacacacccatcacctcacatagttgccatttctttgtgtgtggtgagaacatttatgatcta ttctcttagcaaatttctagtacacaacacggtattgttcactatagtcatcatgctgtccattagatccccagaacttattcatctta ttcattgatttagttctgaataatggctgtgtttaacaaccagctcacaagcttcatgagaatttaacgatcagctcttgcgagctc ttctgagctggctccagcatgccactgcccccagcaagtgtgaggattccagtgtgagaccctatggtgaaaggtacagggt gaatcacagggagaagagaacctggggctggagcagaaacagaaggcttcctggaggaggcaggatccagtgttgagc cttggagcatcaggaggatttggctgggggaagggagagtggcccaagtgaagagacagttagggtgggggtgggag agggtgggaagcaggaagatcagggtgccgtgaatatgtgcttgcaccttacacttgaattatgtttatttccttatag
10	TRPV3 alternatively spliced exon	gtcattttgcatgatgttctgaagttcttgtttgtatatatcgtgtttttgcttggatttggagtag

11	TRPV3 downstream intron	gtaattgatttataaataatagtaactcccacatttataaacactttatacagatgtcaaaatggcaaagtgcatggcggccttgg caacaattagaccaactaccagtggggcacttgccatgcttggcattttccaagcattacctttatatatatacatatatattttatt atactttaagttctagggtacatgtgcacaacatgcaggtttgttacatatgtatacatgtgccatgttggtatgctgcacccatta actcgtcatttacattaggtatatctcctaatcaagcattacctttttacttctctcaacaacttgatcaagtgggaatcatcacccc gatcataaggtaactgaggcttagagactccaggattcaaaccaggccagccactcataatccaacttcaccatgtggtatat acctctggctttcccattcattgcttacgtggccctcacgacctaagataggttctgcaagtgttttcagccccatttttaacagat gcggaaactaaagtccacagaaactcaactgatggtctttgcaaagtaggagcaaactaatacctgaggtgatttccttttgg atttcacatcacgttactctcacagtttgctttagatacaactgtgagagcctctctggtctgctgggaccaggaaggagcccc ctcttaaaatctctccatttacaacagagggaacagcacgatctggttgtatttctaacctctgcatagagagatatcattttccc ccacgtaccagttgctttttttttcttttcttcagtgaccattgacatttgaaattaagtagctgtttccatctgtctggtacctttttctgg gtcacaatttgcttatttatttgctccttcacctgcttgtttattcatccatctatttactctctcatttaatttgctcatccacctgtctatc cgtccatccatccatccatccatccatccatccatccatccaatcattcatccatcgatccatccattgatctatccatctatccat ccatccatccatccatccatccatctaatcatccatccaatcatccatccatccatccatccatccatccatccatccagtcatcc atccattcatccatttaatcatccatccatcgatcaatcggtccatccatccagctatccatccatccatccatctaatcattc atccatcaatccatccatccatccatccatccatccatccatctatccatccatccaatcatccatccatccatcgat caatcagtccatccatccatctaatcatccatctaatcatccatccatccatccatccatccagtcatccatccattcatcca tttaatcatccatccatccatcgatcaattggtccatccatccatctatccatccatccatccatccatctaatcacccatccatcc atccatccatccaatcatctatccatccatccatctaatcatctatccatccatccatccatccatccatccaatcatccatccatc catccatccatccctccatctaatcatccatccctccatccatccctccatatccatccatccatccatccatccatccatccaac cattcagcagtgaacctctactgtggtgtgttggaaagaagggaagtgtgagcctggtcttgctcacagggaactcaaaaca gagcattaggaggagactcaggatcttacagccttgagggaggtgacccaggctagaggaagtcaggagctcaaagact agctggggagccagggacacacacatggattataattagtggctctgcaaggctgcaagtgatcaagggcagactgcggg gtaaaggaaggagagaagactctgggctttgttaggtggaagggctttctgcaggaggccaggccagagaagccagcct gatgaggctttgagaggacctgaaactggagatgggatttaatggccagagagaaggaggcagcaggcttgtcatgtagg caaatacacgcaagcttgtagagtaattaccaggcagagagctgaggagagcaccctgttatgttcccttag
12	TRPV3 downstream exon	ccttggcctcgctgatcgagaagtgtcccaaagacaacaaggactgcagctcctacggcagcttcagcgacgcagtgctg gaactcttcaagctcaccataggcctgggtgacctgaacatccagcagaactccaagtatcccattctctttctgttcctgctca tcacctatgtcatcctcacctttgttctcctcctcaacatgctcattgctctgatgggcgagactgtggagaacgtctccaagga gagcgaacgcatctggcgcctgc
13	STK31 upstream exon	aactctggtggtctccttacaatgagcttggaacgagatcttcttgatgctgagcccatgaaggaacttagcagcaagcgtcct ttggtacgttctgaggttaatgggcagataattctgttaaag
14	STK31 upstream intron	gtaagtctaacttcttcttacttgtgggagattcagcagcataaaggatagtggtgctgttacatttcaagaaagaaaaaacctc actgttagttttctaaagtgtataagttgactggcatgataattgtgcttgtagtttggttgccagaactaatttttcacttttaaagttt ccctcaattgaacaatgagaacacttggacacagggcggggaacatcacccgtgggggcctgtcatggggtgggggca gggggagggatagcattaggagaaatacctaatgtaaatgatgagttaatgggtgcagcaaaccaacatggcacatgtata catatgtaacaaacctgcacgttgtgcacatgtaccctagaacttaaagtataataataaaaataagtatataaaagtcattaaat aaagctttgtctcacaagaaaaaaaaataaagtttccctcatttccctaaatagaatatgtccttaataggatatactttccttttttt tgggtgaacttattatagtcatccaatatcttttaaaattcttgattttcttacatttctccttgaaatctcagctgtgttatcagtacaa tcttctactctttaatttcttcttctcattgatctgggaaactgttaatcatgataataagataattaatttggccttagttgctacacatt ctagaatgtgtagcaggctacaggtagaattgcactttccttttccaaattatagcattcaaagaagatagaaaagtacttataa aattaaaaaataaaattagatttgatttaacttaaaataatgaattatgtttaaattggaatgatgtataattataatgagattttcatct tggaattacgagtgcctctttctttggtacaag
15	STK31 alternatively spliced exon	ggctattctgtggatgttgacacagaagccaaggtgattgagagagcagccacctaccatagagcttggagagaagctgaa ggagactcagggttactgccattgatattcctgtttttatgtaag
16	STK31 downstream intron	gtaaagttcttatgctaatctcttgcagaaaccatttttatcttgcagaaactattttatttaaaataaattttagctccaggttttagag cagcagttcattccccttaacataaaagaaaaattatcccatagaatcaatgtatgctaaaaatatgttgaagaagtttttacttgg agtcacattctctgttttgcttcttag
17	STK31 downstream exon	tctgatcctatggcttatctgatggtcccatactaccctagggcaaacctgaatgctgttcaagccaacatgcctttaaattcag aa
18	USP25 upstream exon	gtgtcacaacataatgaaagttgctcaagccaaactggaaatgataaaacctgaagaagtaaacttggaggaatatgag
19	USP25 upstream intron	gtaatgtgctttcttagaggttaaactacagatttagggattctgtaattcactgatattcctttcgtttcagcccaactgaagttga atttgttatgaatgatgagtaaaaaaattaagtttccaagtcaatgctcaagtaattaataatttttatatgtgattttaggattaacat gatatgatagcatgtagaccaaaaagtttaaaatcaaaaaaggaagatttttggttttttatttcaatactaattctcaagtacttttt aaatgaacttttctagcttaaaaacagaataacaaatgcattgttaattttgaaatcaacttagaaattgctgttacataataccaa caagtaatatttatcaagaccctactatatgccaggcatttgtctaagccctttacatatattaactgatttaattcagacagtaaca

		ctgtaagataggtagtgttcttattcccatttgacacacgagaaaataaggcatagagcagtgaagtaacctgcctaaggttttc acagccacaagtttacagcttggctaatagaacctacattcttaactgttgcactacactgtgcaactgtaccctgtgggctgta ttttttccttcactttgattcatcagatatattttggggaagttaattcattttagtccctatatgccatatatgtttgtcatgagacagt gttgtgtgtgagagagcacttacataggtacttgtacatgtgcttgtgggatttttttgtgcacaaaattatttttgtcaactttggc tttaaaatattctttattttaattgatgaattcatacacctagtatttcttcacatttttaaatttaatcttatgcattgtctttttttaatttat aagtctagtaaattttaagttgttttcttgggtgcctttgcttaatatttagttccatttacaaacttaattatggtactttagattcccta aataattagtgtaattcactacacaaacagggaagaggaaagagagagaaaacactcattattgatgaaaaccttagctagtt tggcgtataaggtaactaccttaattcactaaagagtgtctctaaaaacaaaaatagggaaataataaaactgcagcaaatatat agaattgtgagcactgaaagcttaaaatttttgttgtcaaaagcaaggcaatgaaaataagcgaaaagtaaaatactggaaaa gaagaaacaaagacatcattatttgtagttttttttaccttgaaattcaaataaaatctacaactgatgtattagaactagtaagac agctttgcagcatgcaaacttgtgtcagacatcaattttattccctttgccagtaaaaaaagtatgagaaagtataaaagtgag accatttacaatagtatctgaaactatggagtatatagaaataaatttaacaaaaggtgcgtgtgacttctgtggggaaaattata aaacttattgggagacattaacaaagacctaaataaatgggcatacatatggtatttggaaaacttgatattattaatatatagat atcacttttccccccaaatgctttataaattcag
20	USP25 alternatively spliced exon	tgcaggctgtatgaaaattccagcaagatgttttatggagcttgataagccaactgaaaaatgttacggaaatgcaaagatcc agtccacctcccatctctgacaagcaa
21	USP25 downstream intron	gttagtgaacagtataaagacttttaataaatggttctgggaaataaataaaatttgatttcttctttattccatatgcaaaaattcat ttcaattgtgaaagactcctgtgtaaagctgtttagaagtcaacataggtagaagcccaaagacattcatttatccataagagg cactcactgaccagatatttttgaccaggatatgtaatctagagccagagaaattcaaaatttgttttctcttgaccctgaggagc aagtcaagtcatcttttagctaataatgactccaaatctgtgccctccacattgattgagagctcttcccttgtccccaagattcgt gttcatggaacttgatttttaattgtgacttgattctgcaatgcttccttcagtttgcttcaacagattggactcattatgttttttttcat aaaatcttctcatattctaactgcagggctgttaaatcatacaagtatttaaatgacttgtttttagattctgcatattcctcctatcttt gtcttgctctaagttcagggttcagtaaaccatagcctgcagaccagcttaccacctgtttttgtaaataaacttttattggaacat ggccatttgtatacatattgtctgtggatactttcacactgcagttacagagttgagtaactgaagaagatgaggcttgcaaagt ctaaaatattatctggccttttacagaaaaagtttgcctactcacaacttctaaggtgattctaatgacaacaaggaatgacagtt tgaataccttccccagaaaagctaacttcatacttgcaggacaaattttacctttatttgttatgaagatctactaatagcatgaaa attatgtcttgaaaattaaaattaagactggatgtggtggctcatgcttataatcccagcacttttggaaatcaaggcacaggga gttgtttgaggccaggaatttgagaccggcctggccaacatagcaagaccctgtctatacaaaacagattaagtaattaactg gaactgggagcatgcacctgtaattcaagctacttgggagacagaggtaggaggattgctggagcccagaaggtcaagtct gcagttagccgtgatcgtgccactgcactgcagactgggcaactgagtgagaccctgtctccaaaaaaaaaaaaaaaagttt atattcaaggagaatatcgttggctaataactatgtttcttatttttgtttataatgtttatataaaagccttaacacttttgttatcagta ttcactttccaatttatcctacattgaatttcagttatttttcttgagtcaattattataaagtaagctaagaaaatcttgactttggtatc accttagaattatttagtcctattacatgccctatagttgttttgggggtggggagggcattgaaatgtgaactgcagttattctt ttaccatatctcatgtaaataaactattcgcctacacacttaaagtcaaatttaccacatatttgcgaagttaaaaggttaaataaa tctaacttttttagtttcttaaaattatgtattaagtatgtgtccaagctttgattcatactcagcgtaatatttcattttcttccatcagtt gatgtagatataataatagttcaataatatgattttatactgatgaatttttgagatgacctgattgattatcctgccttttatatttctgtt atctgctagtaaaatctttattgtttaaaacatttcatatccatgcctaataatttttagatttaatgttgtattaataaaatagtcttgtgt gccagaaagtaaagtcttgttttagttctgatgtatcttcattttgaactctttatttctattaattcagatagtagcgaaacagtgcat actttttgtacagtttcatagtacttttaaatcttactctgaacatttttaatatttttctttcacctacag
22	USP25 downstream exon	gagtggcatcaggattataggaaattcagggaaacaactatgtatctcataattgggctagaaaattttcaaagagaaa
23	GTF2B upstream exon	gatcgaacaaataatttattcaagcaagtatatgaacagaagagcctgaagggaagagctaatgatgctatagcttctgcttgt ctctatattgcctgtagacaagaaggggttcctaggacatttaaag
24	GTF2B upstream intron	gtaagtttttagctttgttttttgtgtttgtttgtttgtttgtttgagatagggtctcactttgtcgcccaggccagagtgcagtggcg caatcttgactcactgcaacctctgcctcccaggttcaagctatactcctgcctcagcctcctgagtagctgggattacaggca tgtgccaccacacctggttaacttttgtatttttagtagagacagggtttcgccatgttggccaggctggtctcaaacttctgacc tcaggtgatctgcccactttggcctcccaaagtgctagaattaacaggcgtgaggcactgtgcccatcctaaaataacattttc tacatcatattaaagaaacagaccaatgtcaggcagctttgacacttagaagctaagtcatctttattgtgactacagtaatcca cttgagcatcatttgccaatgcctacttatttcattccttgacagcaaattcatagcatgttcttttaggctgttttcccatgtagact cattttagtccacgtaagcccactatgttcttttctagtgagatcttttgatggactgttgctttgtgggctttcagaaggcctataat cgtaaaccaaagtaggccgtttatacatttggtttacatggaaacagtttccatttctagcactggtctctgtagatgagggagtc aagctttagaaagtatcagtgaaggccaggtatggtgtctcacacctgtaatcccagaacttgggaggccgaggtggttgg gtcatgaggccaggagttcaagaccagcctggccaatatggtgaaaccccgtctctactaaaaatacaaaaattagccagg cgtggtggtgcgcgcctgtagtcccagctactcaggaggctgaggcagaagaatcgcttgaaaccgggaggcagaggtt gcagtgagccgagatcacgccactacactccagcctgggcgacagagtgagacttcatctcaaaaaaaaaaaaaaaaaag ttatcagttaagtttttctcttgaggtgtattgtttgcacttttacacttcagttaagaactggtatttttcatttag

25	GTF2B alternatively spliced exon	tttgcagttacatgtgaatatgtgataatttgttcagtgttgctcactgacttccactaccaaagctcagtttagctgaagactatac ttaacccttattcatcca
26	GTF2B downstream intron	gtgagtatttatttttatgttttcttacgtaaaagtgaagaggtgatggcctttatcatcatctttttgttgttttataactaaaacaa cttttgttaaatttcttctgaaagtgtgtggtggttgcatctgggaatagcagggtcagatattttatctcattgagctacagcccat aatcgcagggatttcatgcacacatgacattgtgttgcctaagaggcatcctttagcttgagtattttctttcaattccttggtaac actgcttacaacagattaatacataaacatcaaagaatttttcatttaaagaaatttagttgtatgaattaatccttagagcagttaa tattttgaaagattaagacctaaatttagccaggcacgatggctcacaccggtaatctcagcactttgtgaggccgaggcagg tagatcacgaggtcaggagttcaagaccagcctggccaagatggtgaaaccccatctctactaaaaatacaaaaaattagcc gggtgtggtggcaggcgcctgtaatcccaactactcgggtggcaggagaatcgcttgaacccaggggcagaggttgca gtgagccaagatggtgtcactgcactccagtctggacgacagagtgaaatttcatctcaaaaaacaaaaaaactctaaattga aagagtgaagaacattaggccaggcgcagtggctcatgcctgtaatctcagcactttgggaggccgaggtaggaggatca cctgaggtcaggagttctagacaggcctggccaacatggtgaaaccccatctctactaaaaatacaaaagttagccaggtgt ggtggtgcacgcctgtagtcccagctactcgggaggctgaggcaggagaatcacttgagcctaggaggtggaggttgcag tgagctgagatcatgccactaccttagtctgggcgacagagtgagacttcatctcaaaaaaaaaggtgtgatgaacattaaa gttaactgagaaatttgcagttaactctgtgggaagtatcactgagatgattagggtttcatctttaaggtatgattgaatgttttac aattagaatgtatattatgtttataattagaaatggtagattttaagatgaactatatggcttaagtgaattgatttatttaatttttttct ttttag
27	GTF2B downstream exon	aaatatgtgccgtatcacgaatttctaagaaagaaattggtcggtgttttaaacttattttgaaagcgctagaaaccagtgtggat ttgattacaactggggacttcatgtccaggttctgttccaaccttttgtcttcctaaacaagtacagatggcagctacacatatagc ccgtaaagctgtggaactggacttggttcctgggaggagccccatctctgtggcagcggcagctatttacatggcctcacag gcatcagctgaaaagaggacccaaaaag
28	LYPD3 upstream exon	atggaccccgccaggaaagcaggtgcccaggccatgatctggactgcaggctggctgctgctgctgctgcttcgcggag
29	LYPD3 upstream intron	gtgggtctgctgcccggagaggtggagggggtagagtgggggctcttagggaggaggctcctctggccccgtacagaga aggctgtgggaaggatccccacataagcagagggcactgtgggcaaagttcccccatcctaagcaagggaggcagagg gaagggtcctccaccatgggccgtaggaagggttccccaccacggacaaggggagctgtgggcaaaattttcctatcctg aagctgcagaaagaagtcccctgttcctaggccgggcgcggtggttcatgcctgtaatctcagcactttgggaggccgagg caggcggatcatgaggtcaaacgatcgagaccagcctggccaacatggtgaaaccctgtctctactaaaaatacaaaaatt agctgggcgcggtagcgcgagcctgtagtcccagctactcgggaggctgaggcaggagaatcgtttgaatccaggaggg gaggttgcagtgagccgagattgcgccactgcactccagcctggcgacagagcgagactccgtctcaaaaaaaaaaaaa agaaaaaaagaaaaaaagaaaaaagtgttccctgtcctgggcagagaagctgtgggtagtgttctctcatttgagcagaag ggggctgtgggaagggatccctcactgagcacagtctactgtgggaaaagttctcccacacactgtgggaaaagttctctca ttctgagcagaaaaccgtgaggaggtatccttcctaccccgggatgaggggcttcatggcggggatcgcctacagtgggc atacagggctgtcggaaggatgtgcccacagttggcatagtgggtgtgaaggaaaagggtccagtgattcccgtgcttcca gatttgtggaaagcttttccctactcctgccctggaaaagtgtgtgcaataagactgtctaatcccgggacagagggcatggg aaggactatttctccttgggggtggacgctttagcaaggagggctcaggtctattcggctccgctcccccag
30	LYPD3 alternatively spliced exon	gagcgcaggccctggagtgctacagctgcgtgcagaaagcagatgacggatgctccccgaacaagatgaagacagtga agtgcgcgccgggcgtggacgtctgcaccgaggccgtgggggcggtggagacca
31	LYPD3 downstream intron	gtgagtggggaccccgggccacgcacaccactgcaggctcctccggcccagccccgcctccttatagtcctttctggatc tggccacgcccctttaacgcggccccacccctccctttcagcctggccgggaccctacgacacagccttgcaggccaggc cttctctattaacaattacaagcgcgaggagcagagagttgagaactgggtggggcggacggggagagctcgggtttgga acgaggtcacgtcctatgcacttttctgctccaccaatgagacctccccatccctgcctgtcaaccaggctgtctcctccgac tcagctctttgaggttccgcccccgttccgtgaccccgctcctccctgcttcggtcgttctttccgctctgagcggctctttggg ggcgtccatccgccaattcctgcctgtcggtcctgtggtgttggcctatttgaggccccgcccctcccagaccctacccaact ctgggccacgccccatctgagaccacgcccacagctaatcccgcccttgtccctctccctag
32	LYPD3 downstream exon	tccacggacaattctcgctggcagtgcggggttgcggttcgggactccccggcaagaatgaccgcggcctggatcttcacg ggcttctggcgttcatccagctgcagcaatgcgctcaggatcgctgcaacgccaagctcaacctcacctcgcgggcgctcg acccggcag
33	IK upstream exon	gtgagccgttctccaacccttggcccccgatggccacgatgtggatgatcctcactccttccacca
34	IK upstream intron	gtgagtattttgtgctaggaccatggaaatgagttgggcaatgaatagaaaagtggtggtgatggtgaaagaattctgaatgt cttgtcatcaaggatggtaaaaaattatacccgctggcccttatctctgtgagccttaactggatcaggccaagtgcataagg gtattagaagggaaattcctgcaggaaaggtgttaagttttttttttttttatacgtgttgtgtaattacacattatataggaattactt cttccaatagcaaatttaaaacataggaaataagggtaaaatcatctttttatttttttgagactgagtttcgctcttgttgctcag
35	IK alternatively spliced exon	gctggaatgcagtggtgcgatctcggctcaccacaacctccgcctcccgggttcaagcaattctcctgcctcagccttccga gtagctgggattacaggcat

36	IK downstream intron	gttccactacgcccggctaattttgtatttttagtagagatggggtttctccatgttgatcaggctggtcttgaaatcccgacttca ggtgatccacccgcctctgcctcccaaagtggtgggagtacaggcgtgagccaccgtgcctggcccccgataaaatcatct ttgataacatccccaattccagtccatttcttttttttttctttttctttttctttctttttttttttttttgagacggagtctcactgttacc caggctggagtgcagtgatgcaatcttggctcactgcaacctctgcctcctgagttcaagcggttctcctgcctcagcctccc aagtagctgggactacaggcatgtgccaccatgcacgactgatttttttttttgtatttttagtagagacagggtttcactgtgttc gtcaggctggtcttgaactcctggcctgaagtgatccacccgcctcagcctcctaaagtgctggcattacaggcgtgagcca ccgtgcccagcctcagtccttttcttagaagtaaccacagttacctttttggcatatctttcctcccgttgcccaggctggagtt caatggcattatttctgctcacagcaacctttgcctcctgggctcaagcaattcttctgcctcagtttccccagtagctgggatta caggcatgccccaccatgcttggctaatttttgtattagtagagacagggtttcgctatgttggccaggctggtcttgaactcct gacctcaaatgatctgcccgtctcggcggcccaaagtgctgggattacaggcgtgagccactgcgcctggccctcctactct ttttttaaatgcatttatacacatacatgcatatctgaggaaaacacaacattgctttgtgatttaaaatatttgagttaaatgatattat gctgtatatatcatttttgcaacttgcttctttttaaaccctgctcagtgtttctgagattatctttcagagcagtgacaagcattccct ccttggctaaactttagtcaggttccccaaagagtcccatttttcagcaggactcatccgtggcctgctgagcacagttttaaca aagaaatcaaccaccgctccacccccatacctaaccaagttgcttttagtaattttctataaattctcacttgtccctgttgtattcg gaattaagttcagtctctctactctattacaatggtcttgactcttgttacagtagtcttcaataaagtcttccttggggagaagag gaacttgattagatatcaagttggattggatatcaagagtgggaaaactttctgtgaactgacttaggattccttttttaaaactga ggtcagcaggccaaggaagagccctagtcagaagagagctttaaaggagtctgactttaatttgatcaagagaggagtctttt gttagtgtaccaattaaaataatctcttgtaccattagtgctgtctttccctctttgggaaacaccatcatagtttagtcagtctcca gctaataaacgcttagtctgtttcagactgattcttttcttttttttattttttgagacagagtcttgctctttttgcccaggctggagt gtaatacctagatctcggctcactgcaaactccgcctcccaggttcaagcgattttcctgcctcagcctcctgagtagctagga ttacaggcgcccaccacaacgcccagctaatttttttgtatttttagtagagatgggatatcaccatgttggccatgctggtcttg aactcctgacctcgggtgatccgcctgcgtcagcctcccaaagtgcggcgattataggtgtgagccactgcacccggccag actgattattttcaaggctataactgaattaaatttacaggccttcttatatacatatgtcagtaatggaccaggctaaatataaata aatggagttgctggctcataggatatatatacttaaatttttaacaaatactgcaaaattgcccttcaaattggttatactggtttaca ttctcatcagcattatatgagagagtccccttcccaaattctcactcacacttgatttatcaaactttctggtagttaacactctgat gcattaaaagtggcatggcatatctggtcgggcgtggtggcttacacctataatcccagcactttgggaggctgaggtgggc agatcacctgaggtcagaagtttgggaccagcctggccaacacctctaataaaaatacaaaaattagctgggtgtgttggcg ggcacctgtattcccagctacttgtgaggctgaggcaggagaattacttgaacccaggaggcagaggttgcagtgggctga gatcatgccactgcactccagccggggcaacagagggagactctgtcttaaaaaaaaaaaaaaaaaaaagctgggtgcggtgg ctcatccctataatcccagcactttgggaggccgaggtgggtggatcacaaggtcaggagttcaagaccagcctggccaat atgggtgacagtgagactccgtctcaaaaaaaaaaaaacagtgtcatatctttgttttaatttgtatttccctgatttctgttaattta atcatgatttttacttgttccttttttcatttcttattgagtcctaatatttaaaatctttgccttttcttctag
37	IK downstream exon	atcaaaactcaccaatgaagacttcaggaaacttctcatgacccccagggctgcacctacctctgcaccaccttctaagtcac gtcaccatga
38	TEX14 upstream exon	tggaggtgaaagaaatggcaaagaaagcagctactggacagctcacagtacctccttggcatcctcagagtagtctgacttt agagagcgaggctgaaaatgagcccgacgccctgctgcagccccccattaggagcccagaaaacacggattggcagcg agttattgagtatcatagggaaaatgatgagcccagaggaaatggcaagtttgacaagacgggcaacaatgactgtgacag tgaccagcatggcagacagcccaggcttggaagcttcacca
39	TEX14 upstream intron	gtaagttcacttgcgggaatcaggataggtgaactcaatcaagtatcagtcgggcgaggtggctcagcctgtaattccagca ctttgggaggtgaggcgagcggatcacctgagcccaggaatttgggaccagcctggtcaacatgatgaaacctggtctcta caaaaaaaaaaaaaaaaaaaaaaaaaaacattagctgggcgtggtggcacacgcctctagttccaggtacccaggaggctg aggtgggaggatcacttggacctgagaggtggaggttgcagtgagctgtgatcatgaaactgcactccagcttgggtgaca gtgagaccctgtctcaaaataaaaataaaagtaaaaaaaaagttatcagctctcaaaaaagtaaactttccagcatatttttttgg tttggtcctctctccttttcctacttttctggaaaccttgaattccaggggaagacagtatagccctatatctcattgcatttcccaca ggcaaacagatccagacactgatcagtttttatattactgaaacag
40	TEX14 alternatively spliced exon	gtccctttgtccaaaggaagtaaaataatcaaatactgaccctggttttggcgaggaatcgatccttctttacctggggataga g
41	TEX14 downstream intron	gtaggccatcactgcctactgcccatgtctctgcattgtattctaagctgaaattaaattctcccagaatttatttagaatgtattct aagctggcattaaatttttccagagagtctatgctggaggaacctagagacacatgacatttattgtcatcggtttatagatttac aggttaaaaaaaatgaaggctgtatattttcagttgataatttgcttataatcagggcatggatattgtttcacagttgtttataagtt aaagaaagatacatctaaacttgtcatttgtcttaggcagaaaatctattggcttgggtattatttgtatttagaatctgtttcttgtc aagcatcactctcgaatgggtgggcagatgatccagggctgcattttggggtcaggctaaactaacagtctctactctggga ctgaacctaacgtatgggacatggccactcttggcagaatatgtatacatctcttctttcatgccacctttatcctcttccatgttgt ctgaatatcactgaatgactttgaccctgactgtgttacaatcattag
42	TEX14 downstream exon	gtatcaggcacccatctcccagacaaaaggagcaaccagagcatagtgaagccttccaagcaagttctgacacattggtgg ctgtagagaaatcttacagt

43	TRPV3 construct #1 (full)	ttttatccaagctgtgcttgtgatactgtctgtcttcttgtacttgtttgcctacaaagagtacctcgcctgcctcgtgctggccttg gccctgggctgggcgaacaagctctactatacgcggggtttccagtccatcggcttgtacagcgtcatcatccagaaggtgc gttgagagctgcgccggcgggaggggcctcggaggatggccatttttcagatggggaaactgagatctggagaggaaga ggggtttaccaaggggcaggtcactgctctttctcccaaatgccatgcactccctttgctttcggccaggctttaagtgtcaag accttaaatcccccaggcttgtggggtgctggggaggggtacagtttgtggaggaatgttctcacccctgggcagtggcgtg ctggtaaacatttaacgactgttctgtgtgtgtgtgtgtgtatgtgtgtgtgtaggtggaatatgtatattccatatatatacaca aaagacttatcatacattttactatataaacgatgtgtatcacacaatttgcaaataatgaaaatataccatgcagtttattgtaaat tccacatagccagttgagtctcagaataatttgttgctttttttttttttttttggagaccaagtcttgctctgctgcccaggccagag taaagtggtgcgatctcggctccctgcaacctctgcctcctgggttcaagcagttctcctgcctcagcctccagagtagctgg gatcacaggcacgcaccaccatgttggctaatttttgtatttttagtagagacagggtttcagcatgttggccaggattgtcttga actcctgacctcaggtgatccacctgcctcggcctcccaaagtgctgggattacaggcatgaaccaccgcgccaggcctatt tgttgctttttattgaattctagttttgttgtcaacttttggttgcaattgacaaaagagtgctatcagacatgaatgttggttgagag ttttctttataagcagtgattgaagtgaaacaattaagatttatgttgaacttacttgtttatcaatgatgtgagcaacttctttgctg aattggatactagtttgcaaataccagaataatattctgttgatatttttgtgctatacataatgtaaaggctgcagacagaacaca ttttaaaatttaaactggattaatagtttctccatcactgtcttaaatcatttagatctagatatcaacaaaacaatgagtcaatcaat ctctgatttgtggcatttaccaattttcatgatgtaaatacagtagtccccccttattcttgggggagatgttctaagaccccccat gaatgcctgaaactgcagatagtaccaaacattatataatatgctatgttttttcctatacatatacatctatgcaaataatgaaaat atactatacaatacagtttattgtaaattccacttagccagttgagtcttagaataatttttgctttttttttttttttgtttgtgactgag tctcgctctgttgcccaggctggagtgcaatagtgcaatctcggctcactgcaacctctgcctcccgggctcaagcaattctc ccgccccagcctcctgagtagctgggactacaggcatgcaccatgcccagctaatttttgtatttttagtagagatggggtttc gccatgttggccaggctggtctccaactcttgacctcaggtgatccgtccgcctcggcctcccaaagtgctgggggtgggg gggggaatcaacaaattattctaagactcaactggctatgtggaatttataaaatttatttttatcaattaggcatagtgagattaa tggcaataactaacagaacaatgataacaatatactgtaataaaagtatgtgaatgttgtctctctcagaatattgtattgtgtga atatttctgggtcatagttgaccacaggcaaccaaaactgcagaaagtgaaatcgaagagaagggcggactactgttctcct ctcagggccaatttccagctaccaacattatgttactggacatagagttagggagaaacatgcagatccacatcgctctagaa ttcctgccatatggctatgatagagacaagtaacctcaaaagcacacacaattgtaaatatagtgaaataactggacggggat gagttttgaacatttattgtctttgttttatttaattaaattaattaattaattaattttttgagacagagtctggctctgccccccaggct ggagtgcagtggtgcgatctaggctcactgcaacccccttcctcctgggttcaagcaattctcccgcctcagcctctggagtag ctgggatcacaggtgtgtgccaccacacccagctaatttttgtatttttagtagagacggggtttcgccatgttggccaagctg gcctccaactcctgacctcaggcgatccacccacctcagcctctcaaagtgttgggattacaggcgtgagtcactgcaccct ggcctgtctttgttttaaatataacttatttttatttaattgtaagtttatatggtttaatttttttggcaatacaaataataatttctttacccc cagctttattgaggtacagctgacatatataattgtataaatttaagatgtacaatgtgaggatttgatataggtatatattgtgaaa tgatttccaatttccaattaacacacccatcacctcacatagttgccatttctttgtgtgtggtgagaacatttatgatctattctctta gcaaatttctagtacacaacacggtattgttcactatagtcatcatgctgtccattagatccccagaacttattcatcttattcattg atttagttctgaataatggctgtgtttaacaaccagctcacaagcttcatgagaatttaacgatcagctcttgcgagctcttctga gctggctccagcatgccactgcccccagcaagtgtgaggattccagtgtgagaccctatggtgaaaggtacagggtgaatc acagggagaagagaacctggggctggagcagaaacagaaggcttcctggaggaggcaggatccagtgttgagccttgg agcatcaggaggatttggctgggggaagggagagtggcccaagtgaagagacagttagggtgggggtgggagaggg tgggaagcaggaagatcagggtgccgtgaatatgtgcttgcaccttacacttgaattatgtttatttccttataggtcatttgcca ccatgttctggagttcttgtttgtatatatcgtgtttttgcttggatttggagttggtaattgatttataaataatagtaactcccacatt tataaacactttatacagatgtcaaaatggcaaagtgcatggcggccttggcaacaattagaccaactaccagtggggcactt gccatgcttggcattttccaagcattacctttatatatatacatatatattttattatactttaagttctagggtacatgtgcacaacat gcaggtttgttacatatgtatacatgtgccatgttggtatgctgcacccattaactcgtcatttacattaggtatatctcctaatcaa gcattacctttttacttctctcaacaacttgatcaagtgggaatcatcaccccgatcataaggtaactgaggcttagagactcca ggattcaaaccaggccagccactcataatccaacttcaccatgtggtatatacctctggctttcccattcattgcttacgtggcc ctcacgacctaagataggttctgcaagtgttttcagccccatttttaacagatgcggaaactaaagtccacagaaactcaactga tggtctttgcaaagtaggagcaaactaataccctgaggtgatttccttggatttcacatcacgttactctcacagtttgctttaga tacaactgtgagagcctctctggtctgctgggaccaggaaggagccccctcttaaaatctctccatttacaacagagggaac agcacgatctggttgtatttctaacctctgcatagagagatatcattttcccccacgtaccagttgcttttttttctttttcttcagtgac cattgacatttgaaattaagtagctgtttccatctgtctggtacctttttctgggtcacaatttgcttatttatttgctccttcacctgctt gtttattcatccatctatttactctctcatttaatttgctcatccacctgtctatccgtccatccatccatccatccatccatccatccat ccatccaatcattcatccatcgatccatccattgatctatccatctatccatccatccatccatccatccatccatctaatcatccat ccaatcatccatccatccatccatccatccatccatccagtcatccatccattcatccatttaatcatccatccatcgatcaa tcggtccatccatccagctatccatccatccatccatctaatcattcatccatcaatccatccatccatccatccatcc atccatccatccatctatccatccatccaatcatccatccatcgatcaatcagtccatccatccatctaatcatccatctaat catccatccatccatccatccatccatccagtcatccatccattcatccatttaatcatccatccatccatcgatcaattggtccat ccatccatctatccatccatccatccatctaatcacccatccatccatccatccatccaatcatctatccatccatccatcta atcatctatccatccatccatccatccatccatccaatcatccatccatccatccatccctccatctaatcatccatccctcc atccatccctccatatccatccatccatccatccatccatccaaccattcagcagtgaacctctactgtggtgtgttggaa

		agaagggaagtgtgagcctggtcttgctcacagggaactcaaaacagagcattaggaggagactcaggatcttacagcctt gagggaggtgacccaggctagaggaagtcaggagctcaaagactagctggggagccagggacacacacatggattata attagtggctctgcaaggctgcaagtgatcaagggcagactgcggggtaaaggaaggagagaagactctgggctttgttag gtggaagggctttctgcaggaggccaggccagagaagccagcctgatgaggctttgagaggacctgaaactggagatgg gatttaatggccagagagaaggaggcagcaggcttgtcatgtaggcaaatacacgcaagcttgtagagtaattaccaggca gagagctgaggagagcaccctgttatgttcccttagccttggcctcgcttatcgagaagtgtcccaaagacaacaaggactg cagctcctacggcagcttcagcgacgcagtgctggaactcttcaagctcaccattggcctgggtgacctggacatccagca gaactccaagtatcccattctctttctgttcctgctcatcacctatgtcatcctcacctttgttctcctcctcaacatgctcattgctct tatgggcgagactgtggagaacgtctccaaggagagcgaacgcatctggcgcctgc
44	TRPV3 construct #2 (minimal v1) Also referred to as V1B	ttttatccaagctgtgcttgtgatactgtctgtcttcttgtacttgtttgcctacaaagagtacctcgcctgcctcgtgctggccttg gccctgggctgggcgaacaagctctactatacgcggggtttccagtccatcggcttgtacagcgtcatcatccagaaggtgc gttgagagctgcgccggcgggaggggcctcggaggatggccatttttcagatggggaaactgagatctggagaggaaga ggggtttaccaaggggcagttagggtgggggtgggagagggtgggaagcaggaagatcagggtgccgtgaatatgtg cttgcaccttacacttgaattatgtttatttccttataggtcatttgccaccatgttctggagttcttgtttgtatatatcgtgtttttgctt ggatttggagttggtaattgatttataaataatagtaactcccacatttataaacactttatacagatgtcaaaatggcaaagtgca tggcggccttggcaacaattagaccagaaggaggcagcaggcttgtcatgtaggcaaatacacgcaagcttgtagagtaat taccaggcagagagctgaggagagcaccctgttatgttcccttagccttggcctcgcttatcgagaagtgtcccaaagacaa caaggactgcagctcctacggcagcttcagcgacgcagtgctggaactcttcaagctcaccattggcctgggtgacctgga catccagcagaactccaagtatcccattctctttctgttcctgctcatcacctatgtcatcctcacctttgttctcctcctcaacatg ctcattgctcttatgggcgagactgtggagaacgtctccaaggagagcgaacgcatctggcgcctgc
45	TRPV3 construct #3 (minimal v2) Also referred to as V2B	tctactatacgcggggtttccagtccatcggcttgtacagcgtcatcatccagaaggtgcgttgagagctgcgccggcgggga ggggcctcggaggatggccattttgtgaatatgtgcttgcaccttacacttgaattatgtttatttccttataggtcatttgccacc atgttctggagttcttgtttgtatatatcgtgtttttgcttggatttggagttggtaattgatttataaataatagtaactcccacatttat aaacactttatagtaattaccaggcagagagctgaggagagcaccctgttatgttcccttagccttggcctcgcttatcgagaa gtgtcccaaagacaacaaggactgcagctcctacggcagcttcagcgacgcagtgctggaactct
46	STK31 construct #1 (full)	aactctggtggtctccttacaatcagcttggaacgagatcttcttgaagctgagcccttgaaggaacttagcagcaagcgtcct ttggtacgttctgaggttattgggcagataattctgttaaaggtaagtctaacttcttcttacttgtgggagattcagcagcataaa ggatagtggtgctgttacatttcaagaaagaaaaaacctcactgttagttttctaaagtgtataagttgactggcatgataattgt gcttgtagtttggttgccagaactaattttcacttttaaagtttccctcaattgaacaatgagaacacttggacacagggcgggg aacatcacccgtgggggcctgtcatggggtgggggcaggggagggatagcattaggagaaatacctaatgtaaatgat gagttaatgggtgcagcaaaccaacatggcacatgtatacatatgtaacaaacctgcacgttgtgcacatgtaccctagaact taaagtataataataaaaataagtatataaaagtcattaaataaagctttgtctcacaagaaaaaaaaataaagtttccctcatttc cctaaatagaatatgtccttaataggatatacttttcctttttttttgggtgaacttattatagtcatccaatatctttttaaaattcttgatt ttcttacatttctccttgaaatctcagctgtgttatcagtacaatcttctactctttaatttcttcttctcattgatctgggaaactgttaa tcatgataataagataattaatttggccttagttgctacacattctagaatgtgtagcaggctacaggtagaattgcactttccttttt ccaaattatagcattcaaagaagatagaaaagtacttataaaattaaaaaataaaattagatttgatttaacttaaaataatgaatt atgtttaaattggaatgatgtataattataatgagattttcatcttggaattacgagtgcctctttctttggtacaagggctattctgt ggttgttgacacagaagccaaggtgattgagagagcagccacctaccatagagcttggagagaagctgaaggagactcag ggttactgccattgatattcctgccaccatgaaaggtaaagttcttatgctaatctcttgcagaaaccatttttatcttgcagaaact attttatttaaaataaatttttagctccaggttttagagcagcagttcattcccctttaacataaaagaaaaattatcccatagaatcaa tgtatgctaaaaatatgttgaagaagttttttacttggagtcacattctctgttttgcttcttagtctatcctatggcttatctgatggtc ccatactaccctagggcaaacctgaatgctgttcaagccaacatgcctttaaattcagaa
47	STK31 construct #2 (minimal v1)	aactctggtggtctccttacaatcagcttggaacgagatcttcttgaagctgagcccttgaaggaacttagcagcaagcgtcct ttggtacgttctgaggttattgggcagataattctgttaaaggtaagtctaacttcttcttacttgtgggagattcagcagcataaa ggatagtggtgctgttacatttcaagaaagaaaaaacctcactgttagttttctaatttaacttaaaataatgaattatgtttaaattg gaatgatgtataattataatgagattttcatcttggaattacgagtgcctctttctttggtacaagggctattctgtggttgttgacac agaagccaaggtgattgagagagcagccacctaccatagagcttggagagaagctgaaggagactcagggttactgccat tgatattcctgccaccatgaaaggtaaagttcttatgctaatctcttgcagaaaccatttttatcttgcagaaactatttttatttaaaat aaattttagctccaggttttagagcagcagttcattcccctttaacataaaagaaaaattatcccatagaatcaatgtatgctaaaa atatgttgaagaagttttttacttggagtcacattctctgttttgcttcttagtctatcctatggcttatctgatggtcccatactaccct agggcaaacctgaatgctgttcaagccaacatgcctttaaattcagaa
48	USP25 construct #1 (full)	gtgtcacaacatattgaaagttgctcaagccaaactggaattgataaaacctgaagaagtaaacttggaggaatatcaggtaa tgtgctttcttagaggttaaactacagatttagggattctgtaattcactgatattccttcgtttcagcccaactgaagttgaatttg ttatgaatgatgagtaaaaaaattaagtttccaagtcaatgctcaagtaattaataatttttatatgtgattttaggattaacatgatat gatagcatgtagaccaaaaagtttaaaatcaaaaaaggaagatttttggttttttatttcaatactaattctcaagtacttttttaaatg aacttttctagcttaaaaacagaataacaaatgcattgttaattttgaaatcaacttagaaattgctgttacataataccaacaagt aatatttatcaagaccctactatatgccaggcatttgtctaagccctttacatatattaactgatttaattcagacagtaacactgta agataggtagtgttcttattcccatttgacacacgagaaaataaggcatagagcagtgaagtaacctgcctaaggttttcacag ccacaagtttacagcttggctaatagaacctacattcttaactgttgcactacactgtgcaactgtaccctgtgggctgtatttttt

		ccttcactttgattcatcagatatattttggggaagttaattcattttagtccctatatgccatatatgtttgtcatgagacagtgttgt gtgtgagagagcacttacataggtacttgtacatgtgcttgtgggatttttttgtgcacaaaattatttttgtcaactttggctttaa aatattctttattttaattgatgaattcatacacctagtatttcttcacatttttaaatttaatcttatgcattgtctttttttaatttataagt ctagtaaattttaagttgttttcttgggtgcctttgcttaatatttagttccatttacaaacttaattatggtactttagattccctaaata attagtgtaattcactacacaaacagggaagaggaaagagagagaaaacactcattattgatgaaaaccttagctagtttggc gtataaggtaactaccttaattcactaaagagtgtctctaaaaacaaaaataggaaataataaaactgcagcaaatatatagaa ttgtgagcactgaaagcttaaaatttttgttgtcaaaagcaaggcaatgaaaataagcgaaaagtaaaatactggaaaagaag aaacaaagacatcattatttgtagttttttttaccttgaaattcaaataaaatctacaactgatgtattagaactagtaagacagcttt gcagcatgcaaacttgtgtcagacatcaattttattccctttgccagtaaaaaaagtatgagaaagtataaaagtgagaccatt tacaatagtatctgaaactatggagtatatagaaataaatttaacaaaaggtgcgtgtgacttctgtggggaaaattataaaactt attgggagacattaacaaagacctaaataaatgggcatacatatggtatttggaaaacttgatattattaatatatagatatcactt ttcccccaaatgctttataaattcagtgcaggctgtatgaaaattccagcaagatgttttatcgagcttgataagccaactgaaa aatcttacggaattgccaccatgccagtccacctcccatctctgacaagcaagttagtgaacagtataaagacttttttaataaat ggttctgggaaataaataaaatttgatttcttctttattccatatgcaaaaattcatttcaattgtgaaagactcctgtgtaaagctgt ttagaagtcaacataggtagaagcccaaagacattcatttatccataagaggcactcactgaccagatatttttgaccaggata tgtaatctagagccagagaaattcaaaatttgttttctcttgaccctgaggagcaagtcaagtcatcttttagctaataatgactcc aaatctgtgccctccacattgattgagagctcttcccttgtccccaagattcgtgttcatggaacttgatttttaattgtgacttgatt ctgcaatgcttccttcagtttgcttcaacagattggactcattatgttttttttcataaaatcttctcatatttctaactgcagggctgtta aatcatacaagtatttaaatgacttgttttagattctgcatattcctcctatctttgtcttgctctaagttcagggttcagtaaaccat agcctgcagaccagcttaccacctgtttttgtaaataaacttttattggaacatggccatttgtatacatattgtctgtggatacttt cacactgcagttacagagttgagtaactgaagaagatgaggcttgcaaagtctaaaatattatctggccttttacagaaaaagt ttgcctactcacaacttctaaggtgattctaatgacaacaaggaatgacagtttgaatacctttccccagaaaagctaacttcata cttgcaggacaaatttttaccttttatttgttatgaagatctactaatagcatgaaaattatgtcttgaaaattaaaattaagactggat gtggtggctcatgcttataatcccagcacttttggaaatcaaggcacagggagttgtttgaggccaggaatttgagaccggcc tggccaacatagcaagaccctgtctatacaaaacagattaagtaattaactggaactgggagcatgcacctgtaattcaagct acttgggagacagaggtaggaggattgctggagcccagaaggtcaagtctgcagttagccgtgatcgtgccactgcactg cagactgggcaactgagtgagaccctgtctccaaaaaaaaaaaaaaaaagtttatattcaaggagaatatcgttggctaataac tatgtttcttatttttgtttataatgtttatataaaagccttaacacttttgttatcagtattcactttccaatttatcctacattgaatttcag ttatttttcttgagtcaattattataaagtaagctaagaaaatcttgactttggtatcaccttagaattatttagtcctattacatgccct atagttgttttgggggtggggaggggcattgaaatgtgaactgcagttattcttttaccatatctcatgtaaataaactattcgcct acacacttaaagtcaaatttaccacatatttgcgaagttaaaaggttaaataaatctaacttttttttagtttcttaaaattatgtattaa gtatgtgtccaagctttgattcatactcagcgtaatatttcattttcttccatcagttgatgtagatataataatagttcaataatatga ttttatactgatgaatttttgagatgacctgattgattatcctgccttttatatttctgttatctgctagtaaaatctttattgtttaaaacat ttcatatccatgcctaataatttttagatttaatgttgtattaataaatagtcttgtgtgccagaaagtaaagtcttgttttagttctgat gtatcttcattttgaactctttatttctattaattcagatagtagcgaaacagtgcattacttttgtacagtttcatagtacttttttaaat cttactctgaacatttttaatatttttctttcacctacaggagtggcatcaggattataggaaattcagggaaacaactatgtatctc aaaattgggccagaaaatttttcaaagagaaa
49	USP25 construct #2 (minimal v1)	gtgtcacaacatattgaaagttgctcaagccaaactggaattgataaaacctgaagaagtaaacttggaggaatatcaggtaa tgtgctttcttagaggttaaactacagatttagggattctgtaattcactgatattcctttcgtttcagcccaactgaagttgaatttg ttatgaaaacaaagacctaaataaatgggcatacatatggtatttggaaaacttgatattattaatatatagatatcacttttcccc caaatgctttataaattcagtgcaggctgtatgaaaattccagcaagatgttttatcgagcttgataagccaactgaaaaatctta cggaattgccaccatgccagtccacctcccatctctgacaagcaagttagtgaacagtataaagacttttttaataaatggttctg ggaaataaataaaatttgatttcttctttattccatatgcaaaaattcatttcaattgtgaattcagatagtagcgaaacagtgcatt acttttgtacagtttcatagtacttttttaaatcttactctgaacatttttaatatttttctttcacctacaggagtggcatcaggattata ggaaattcagggaaacaactatgtatctcaaaattgggccagaaaatttttcaaagagaaa
50	GTF2B construct #1 (full)	gatcgaacaaataatttattcaagcaagtataagaacagaagagcctgaagggaagagctattgatcctatagcttctgcttgt ctctatattgcctgtagacaagaaggggttcctaggacatttaaaggtaagtttttagctttgttttttgtgtttgtttgttttgtttt gagataggtctcactttgtcgcccaggccagagtgcagtggcgcaatcttgactcactgcaacctctgcctcccaggttca agctatactcctgcctcagcctcctgagtagctgggattacaggcatgtgccaccacacctggttaacttttgtatttttagtaga gacagggtttcgccatgttggccaggctggtctcaaacttctgacctcaggtgatctgcccactttggcctcccaaagtgcta gaattaacaggcgtgaggcactgtgcccatcctaaaataacattttctacatcatattaaagaaacagaccaatgtcaggcag ctttgacacttagaagctaagtcatctttattgtgactacagtaatccacttgagcatcatttgccaatgcctacttatttcattcctt gacagcaaattcatagcatgttcttttaggctgttttcccatgtagactcattttagtccacgtaagcccactatgttctttctagtg agatcttttgatggactgttgctttgtgggctttcagaaggcctataatcgtaaaccaaagtaggccgtttatacatttggtttaca tggaaacagtttccatttctagcactggtctctgtagatgagggagtcaagctttagaaagtatcagtgaaggccaggtatggt gtctcacacctgtaatcccagaactttgggaggccgaggtggttgggtcatgaggccaggagttcaagaccagcctggcca atatggtgaaaccccgtctctactaaaaatacaaaaattagccaggcgtggtggtgcgcgcctgtagtcccagctactcagg aggctgaggcagaagaatcgcttgaaaccgggaggcagaggttgcagtgagccgagatcacgccactacactccagcct gggcgacagagtgagacttcatctcaaaaaaaaaaaaaaaaaagttatcagttaagttttttctcttgaggtgtattgtttgcact

		tttacacttcagttaagaactggtatttttcatttagtttgcagttacttgtgaatatctgataatttgttcagtgttgctcactgacttc cactaccaaagctcagtttagctgaagacgccaccatggcctttattcatccagtgagtatttatttttatgttttcttacgtaaaa gtgaagaggtgatggcctttatcatcatctttttgttgttttataactaaaacaacttttgttaaatttcttctgaaagtgtgtggtgg ttgcatctgggaatagcagggtcagatattttatctcattgagctacagcccataatcgcagggatttcatgcacacatgacatt gtgttgcctaagaggcatcctttagcttgagtattttctttcaattccttggtaacactgcttacaacagattaatacataaacatca aagaatttttcatttaaagaaatttagttgtatgaattaatccttagagcagttaatattttgaaagattaagacctaaatttagccag gcacgatggctcacaccggtaatctcagcactttgtgaggccgaggcaggtagatcacgaggtcaggagttcaagaccag cctggccaagatggtgaaaccccatctctactaaaaatacaaaaaattagccgggtgtggtggcaggcgcctgtaatcccaa ctactcgggtggcaggagaatcgcttgaacccaggggcagaggttgcagtgagccaagatggtgtcactgcactccagt ctggacgacagagtgaaatttcatctcaaaaaacaaaaaaactctaaattgaaagagtgaagaacattaggccaggcgcag tggctcatgcctgtaatctcagcactttgggaggccgaggtaggaggatcacctgaggtcaggagttctagacaggcctgg ccaacatggtgaaaccccatctctactaaaaatacaaaagttagccaggtgtggtggtgcacgcctgtagtcccagctactc gggaggctgaggcaggagaatcacttgagcctaggaggtggaggttgcagtgagctgagatcatgccactacccttagtct gggcgacagagtgagacttcatctcaaaaaaaaaggtgtgatgaacattaaagttaactgagaaatttgcagttaactctgtg ggaagtatcactgagatgattagggtttcatctttaaggtatgattgaatgttttacaattagaatgtatattatgtttataattagaa atggtagattttaagatgaactatatggcttaagtgaattgatttatttaatttttttctttagaaatatgtgccgtatcacgaatttct aagaaagaaattggtcggtgttttaaacttatttagaaagcgcttgaaaccagtgtggattttattacaactggggacttcatgtc caggttctgttccaacctttgtcttcctaaacaagtacagatggcagctacacatattgcccgtaaagctgtggaactggacttg gttcctgggaggagccccatctctgtggcagcggcagctatttacatggcctcacaggcatcagctgaaaagaggacccaa aaag
51	GTF2B construct #2 (minimal v1)	gatcgaacaaataatttattcaagcaagtataagaacagaagagcctgaagggaagagctattgatcctatagcttctgcttgt ctctatattgcctgtagacaagaaggggttcctaggacatttaaaggtaagtttttagctttgttttttgtgtttgtttgtttgtttgttt gagatagggtctcacttttgtcgcccaggccagagtgcagtggcgcaatcttgactcaaaaaaaaaaaaaaaaaaaagttatcag ttaagttttttctcttgaggtgtattgtttgcacttttacacttcagttaagaactggtatttttcatttagtttgcagttacttgtgaatat ctgataatttgttcagtgttgctcactgacttccactaccaaagctcagtttagctgaagacgccgccaccatggcctttattcat ccagtgagtatttatttttttatgttttcttacgtaaaagtgaagaggtgatggcctttatcatcatctttttgttgttttttataactaaaa caacttttgttaacaattagaatgtatattatgtttataattagaaatggtagattttaagatgaactatatggcttaagtgaattgatt tatttaatttttttttctttagaaatatgtgccgtatcacgaatttctaagaaagaaattggtcggtgttttaaacttatttagaaagcg cttgaaaccagtgtggattttattacaactggggacttcatgtccaggttctgttccaacctttgtcttcctaaacaagtacagatg gcagctacacatattgcccgtaaagctgtggaactggacttggttcctgggaggagccccatctctgtggcagcggcagcta tttacatggcctcacaggcatcagctgaaaagaggacccaaaaag
52	LYPD3 construct #1 (full)	gaccccgccaggaaagcaggtgcccaggccttgatctggactgcaggctggctgctgctgctgctgcttcgcggaggtgg gtctgctgcccggagaggtggaggggtagagtgggggctcttagggaggaggctcctctggccccgtacagagaagg ctgtgggaaggatccccacataagcagagggcactgtgggcaaagttcccccatcctaagcaagggaggcagagggaa gggtcctccaccatgggccgtaggaagggttccccaccacggacaaggggagctgtgggcaaaattttcctatcctgaag ctgcagaaagaagtcccctgttcctaggccgggcgcggtggttcatgcctgtaatctcagcactttgggaggccgaggcag gcggatcatgaggtcaaacgatcgagaccagcctggccaacatggtgaaaccctgtctctactaaaaatacaaaaattagct gggcgcggtagcgcgagcctgtagtcccagctactcgggaggctgaggcaggagaatcgtttgaatccaggaggggag gttgcagtgagccgagattgcgccactgcactccagcctggcgacagagcgagactccgtctcaaaaaaaaaaaaaaga aaaaagaaaaaaagaaaaaagttgttccctgtcctgggcagagaagctgtgggtagtgttctctcatttgagcagaaggggg ctgtgggaagggatccctcactgagcacagtctactgtgggaaaagttctcccacacactgtgggaaaagttctctcattctg agcagaaaaccgtgaggaggtatccttcctacccccgggatgaggggcttcatggcggggatcgcctacagtgggcataca gggctgtcggaaggatgtgcccacagttggcatagtgggtgtgaaggaaaagggtccagtgattcccgtgcttccagatttg tggaaagcttttccctactcctgccctggaaaagtgtgtgcaataagactgtctaatcccgggacagagggcatgggaagga ctatttctccttgggggtggacgctttagcaaggagggctcaggtctattcggctccgctcccccaggagcgcaggccctgg agtgctacagctgcgtgcagagccaccatggcggatcctccccgaacaagatcaagacagttaagtgcgcgccgggcgt ggacgtctgcaccgaggccgtgggggcggtggagaccagtgagtggggaccccccgggccacgcacaccactgcaggc tcctccggcccagccccgcctccttatagtcctttctggatctggccacgcccctttaacgcggccccacccctccctttcagc ctggccgggaccctacgacacagccttgcaggccaggccttctctattaacaattacaagcgcgaggagcagagagttga gaactgggtggggcggacggggagagctcgggtttggaacgaggtcacgtcctatgcacttttctgctccaccaatgaga cctccccatccctgcctgtcaaccaggctgtctcctccgactcagctctttgaggttccgcccccgttccgtgaccccgctcct ccctgcttcggtcgttctttccgctctgagcggctctttgggggcgtccatccgccaattcctgcctgtcggtcctgtggtgttg gcctatttgaggccccgccccctcccagaccctacccaactctgggccacgccccatctgagaccacgcccacagctaatcc cgcccttgtccctctccctagtccacggacaattctcgctggcagtgcggggttgcggttcgggactccccggcaagaatga ccgcggcctggatcttcacgggcttctggcgttcatccagctgcagcattgcgctcaggatcgctgcaacgccaagctcaac ctcacctcgcgggcgctcgacccggcag
53	LYPD3 construct #2 (minimal v1)	gaccccgccaggaaagcaggtgcccaggccttgatctggactgcaggctggctgctgctgctgctgcttcgcggaggtgg gtctgctgcccggagaggtggaggggtagagtgggggctcttagggaggaggctcctctggccccgtacagagaagg ctgtgggaaggatccccatgtctaatcccgggacagagggcatgggaaggactatttctccttgggggtggacgctttagca

	Also referred to as V1A	aggagggctcaggtctattcggctccgctcccccaggagcgcaggccctggagtgctacagctgcgtgcagagccaccat ggcggatcctccccgaacaagatcaagacagttaagtgcgcgccgggcgtggacgtctgcaccgaggccgtgggggcg gtggagaccagtgagtggggaccccccgggccacgcacaccactgcaggctcctccggcccagccccgcctccttatagt cctttctggatctggccacgcccctttaacgctatttgaggccccgccctcccagaccctacccaactctgggccacgccc catctgagaccacgcccacagctaatcccgcccttgtccctctccctagtccacggacaattctcgctggcagtgcggggtt gcggttcgggactccccggcaagaatgaccgcggcctggatcttcacgggcttctggcgttcatccagctgcagcattgcg ctcaggatcgctgcaacgccaagctcaacctcacctcgcgggcgctcgacccggcag
54	LYPD3 construct #3 (minimal v2) Also referred to as V2A	gctgctgctgctgctgcttcgcggaggtgggtctgctgcccggagaggtggaggggggtagagtgggggctcttaggtgga cgctttagcaaggagggctcaggtctattcggctccgctccccaggagcgcaggccctggagtgctacagctgcgtgca gagccaccatggcggatcctccccgaacaagatcaagacagttaagtgcgcgccgggcgtggacgtctgcaccgaggcc gtgggggcggtggagaccagtgagtggggaccccccgggccacgcacaccactgcaggctcctccggcccccatctgag accacgcccacagctaatcccgcccttgtccctctccctagtccacggacaattctcgctggcagtgcggggttgcggttcg gggactccccggcaaga
55	IK construct #1 (full)	gtgagccgttctccaacccttggcccccgttggccacgatctggaagatcctcactccttccaccagtgagtattttgtgctag gaccatggaaatgagttgggcaatgaatagaaaagtggtggtgatggtgaaagaattctgaatgtcttgtcatcaaggatggt aaaaaattatacccgctggccctttatctctgtgagccttaactggatcaggccaagtgcataagggtattagaagggaaattc ctgcaggaaaggtgttaagtttttttttttttttatacgtgttgtgtaattacacattatataggaattacttcttccaatagcaaattaa aacataggaaataagggtaaaatcatcttttattttttttgagactgagtttcgctcttgttgctcaggctggaatgcagtggtgc gatctcggctcacgccaccatgggcctcccgggttcaagcaattctcctgcctcagccttccgagtagctgggattacaggc atgttccactacgcccggctaattttgtatttttagtagagatgggtttctccatgttgatcaggctggtcttgaaatcccgactt caggtgatccacccgcctctgcctcccaaagtggtgggagtacaggcgtgagccaccgtgcctggccccgataaaatca tctttgataacatccccaattccagtccatttctttttttttttctttttctttttctttctttttttttttttttgagacggagtctcactgtta cccaggctggagtgcagtgatgcaatcttggctcactgcaacctctgcctcctgagttcaagcggttctcctgcctcagcctc ccaagtagctgggactacaggcatgtgccaccatgcacgactgatttttttttttgtatttttagtagagacagggtttcactgtgt tcgtcaggctggtcttgaactcctggcctgaagtgatccacccgcctcagcctcctaaagtgctggcattacaggcgtgagc caccgtgcccagcctcagtccttttcttagaagtaaccacagttacctttttggcatatctttcctcccgttgcccaggctgga gttcaatggcattatttctgctcacagcaacctttgcctcctgggctcaagcaattcttctgcctcagtttccccagtagctggga ttacaggcatgccccaccatgcttggctaatttttgtattagtagagacagggtttcgctatgttggccaggctggtcttgaactc ctgacctcaaatgatctgcccgtctcggcggcccaaagtgctgggattacaggcgtgagccactgcgcctggccctcctact cttttttaaatgcatttatacacatacatgcatatctgaggaaaacacaacattgctttgtgatttaaaatatttgagttaaatgatatt atgctgtatatatcatttttgcaacttgcttcttttaaaccctgctcagtgtttctgagattatctttcagagcagtgacaagcattcc ctccttggctaaactttagtcaggttccccaaagagtcccatttttcagcaggactcatccgtggcctgctgagcacagttttaa caaagaaatcaaccaccgctccacccccatacctaaccaagttgcttttagtaattttctataaattctcacttgtccctgttgtatt cggaattaagttcagtctctctactctattacaatggtcttgactcttgttacagtagtcttcaataaagtcttccttggggagaag aggaacttgattagatatcaagttggattggatatcaagagtgggaaaactttctgtgaactgacttaggattccttttttaaaact gaggtcagcaggccaaggaagagccctagtcagaagagagctttaaaggagtctgactttaatttgatcaagagaggagtc tttgttagtgtaccaattaaaataatctcttgtaccattagtgctgtctttccctctttgggaaacaccatcatagtttagtcagtctc cagctaataaacgcttagtctgtttcagactgattcttttctttttttattttttgagacagagtcttgctcttttgcccaggctgga gtgtaatacctagatctcggctcactgcaaactccgcctcccaggttcaagcgattttcctgcctcagcctcctgagtagctag gattacaggcgcccaccacaacgcccagctaatttttttgtatttttagtagagatgggatatcaccatgttggccatgctggtct tgaactcctgacctcgggtgatccgcctgcgtcagcctcccaaagtgcggcgattataggtgtgagccactgcacccggcc agactgattattttcaaggctataactgaattaaatttacaggccttcttatatacatatgtcagtaatggaccaggctaaatataa ataaatggagttgctggctcataggatatatatacttaaatttttaacaaatactgcaaaattgcccttcaaattggttatactggttt acattctcatcagcattatatgagagagtccccttcccaaattctcactcacacttgatttatcaaactttctggtagttaacactct gatgcattaaaagtggcatggcatatctggtcgggcgtggtggcttacacctataatcccagcactttgggaggctgaggtg ggcagatcacctgaggtcagaagtttgggaccagcctggccaacacctctaataaaaatacaaaaattagctgggtgtgttg gcgggcacctgtattcccagctacttgtgaggctgaggcaggagaattacttgaacccaggaggcagaggttgcagtggg ctgagatcatgccactgcactccagccggggcaacagagggagactctgtcttaaaaaaaaaaaaaaaaaagctgggtgcg gtggctcatccctataatcccagcactttgggaggccgaggtgggtggatcacaaggtcaggagttcaagaccagcctggc caatatgggtgacagtgagactccgtctcaaaaaaaaaaaaaacagtgtcatatctttgttttaatttgtatttccctgatttctgtta atttaatcatgatttttacttgttccttttttcatttcttattgagtcctaatatttaaaatctttgccttttcttctagatcaaaactcaccatt gaagacttcaggaaacttctcagacccccagggctgcacctacctctgcaccaccttcaagtcacgtcaccatta
56	IK construct #2 (minimal v1)	gtgagccgttctccaacccttggcccccgttggccacgatctggaagatcctcactccttccaccagtgagtattttgtgctag gaccatggaaatgagttgggcaatgaatagaaaagtggtggtgatggtgaaagaattctgaatgtcttgtcatcaaggatgg ataggaattacttcttccaatagcaaattaaaacataggaaataagggtaaaatcatcttttattttttttgagactgagtttcgctc ttgttgctcaggctggaatgcagtggtgcgatctcggctcacgccaccatgggcctcccgggttcaagcaattctcctgcctc agccttccgagtagctgggattacaggcatgttccactacgcccggctaattttgtatttttagtagagatggggtttctccatgt tgatcaggctggtcttgaaatcccgacttcaggtgatccacccggtatttccctgatttctgttaatttaatcatgatttttacttgtt

		ccttttttcatttcttattgagtcctaatatttaaaatctttgccttttcttctagatcaaaactcaccattgaagacttcaggaaacttct cagacccccagggctgcacctacctctgcaccaccttcaagtcacgtcaccatta
57	TEX14 construct #1 (full)	tggaggtgaaagaaatggcaaagaaagcagctactggacagctcacagtacctccttggcatcctcagagtagtctgacttt agagagcgaggctgaaaatgagcccgacgccctgctgcagccccccattaggagcccagaaaacacggattggcagcg agttattgagtatcatagggaaaatgatgagcccagaggaaatggcaagtttgacaagacgggcaacaatgactgtgacag tgaccagcatggcagacagcccaggcttggaagcttcaccagtaagttcacttgcgggaatcaggataggtgaactcaatc aagtatcagtcgggcgaggtggctcagcctgtaattccagcactttgggaggtgaggcgagcggatcacctgagcccagg aatttgggaccagcctggtcaacatgatgaaacctggtctctacaaaaaaaaaaaaaaaaaaaaaaaaaaacattagctggg cgtggtggcacacgcctctagttccaggtacccaggaggctgaggtgggaggatcacttggacctgagaggtggaggttg cagtgagctgtgatcatgaaactgcactccagcttgggtgacagtgagaccctgtctcaaaataaaaataaaagtaaaaaaa agttatcagctctcaaaaaagtaaactttccagcatatttttttggtttggtcctctctccttcctacttttctggaaaccttgaattc caggggaagacagtatagccctatatctcattgcatttcccacaggcaaacagatccagacactgatcagtttttatattactga aacaggtccctttgtccaaaggaagtaaaataatcaaatactgaccctggttttggcgaggaatcgatccttgccaccatggg gatagaggtaggccatcactgcctactgcccatgtctctgcattgtattctaagctgaaattaaattctcccagaatttatttagaa tgtattctaagctggcattaaatttttccagagagtctatgctggaggaacctagagacacatgacatttattgtcatcggtttata gatttacaggttaaaaaaaatgaaggctgtatattttcagttgataatttgcttataatcagggcatggatattgtttcacagttgttt ataagttaaagaaagatacatctaaacttgtcatttgtcttaggcagaaaatctattggcttgggtattatttgtatttagaatctgtt tcttgtcaagcatcactctcgaatgggtgggcagatgatccagggctgcatttttggggtcaggctaaactaacagtctctactc tctggactgaacctaacgtatgggacatggccactcttggcagaatatgtatacatctcttctttcatgccacctttatcctcttcc atgttgtctgaatatcactgaatgactttgaccctgactgtgttacaatcattaggtatcaggcacccatctcccagacaaaagg agcaaccagagcagtgaagccttccaagcaagttctgacacattggtggctgtagagaaatcttacagt
58	TEX14 construct #2 (minimal v1)	tggaggtgaaagaaatggcaaagaaagcagctactggacagctcacagtacctccttggcatcctcagagtagtctgacttt agagagcgaggctgaaaatgagcccgacgccctgctgcagccccccattaggagcccagaaaacacggattggcagcg agttattgagtatcatagggaaaatgatgagcccagaggaaatggcaagtttgacaagacgggcaacaatgactgtgacag tgaccagcatggcagacagcccaggcttggaagcttcaccagtaagttcacttgcgggaatcaggataggtgaactcaatc aagtatcagtcgggcgaggtggctcagcctgtaattccagcactttgggaggtgaggcgaccttgaattccaggggaagac agtatagccctatatctcattgcatttcccacaggcaaacagatccagacactgatcagtttttatattactgaaacaggtccctttt gtccaaaggaagtaaaataatcaaatactgaccctggttttggcgaggaatcgatccttgccaccatggggatagaggtagg ccatcactgcctactgcccatgtctctgcattgtattctaagctgaaattaaattctcccagaatttatttagaatgtattctaagct ggcattagaatatgtatacatctcttctttcatgccacctttatcctcttccatgttgtctgaatatcactgaatgactttgaccctga ctgtgttacaatcattaggtatcaggcacccatctcccagacaaaaggagcaaccagagcagtgaagccttccaagcaagtt ctgacacattggtggctgtagagaaatcttacagt
59	Rep Construct #1	caattgtcgcgaaataaaatatctttatttttcattacatctgtgtgttggttttttgtgtgaatcgatagtactaacatacgctctccat caaaacaaaacgaaacaaaacaaactagcaaaatataggctgtccccagtgcaagtgcaggtgccagaacatttctctggcct aactggccggtacctgagctcgctagcctcgagcctgcaggcgacatgaagaatctgcttagggttaggcgttttgcgctgc ttcgcgttgtacgggccagatatacgcgtatctgaggggactagggtgtgtttaggcgaaaagcggggcttcggttgtacgc ggttaggagtcccctcaggatatagtagtttcgcttttgcatagggagggggattgcgccttttccaaggcagccctgggtttg cgcagggacgcggctgctctgggcgtggttccgggaaacgcagcggcgccgaccctgggtctcgcacattcttcacgtcc gttcgcagcgtcacccggatcttcgccgctacccttgtgggccccccggcgacgcttcctgctccgcccctaagtcgggaa ggttccttgcggttcgcggcgtgccggacgtgacaaacggaagccgcacgtctcactagtaccctcgcagacggacagcg ccagggagcaatggcagcgcgccgaccgcgatgggctgtggccaatagcggctgctcagcagggcgcgccgagagca gcggccgggaaggggcggtgcgggaggcggggtgtggggcggtagtgtgggccctgttcctgcccgcgcggtgttccg cattctgcaagcctccggagcgcacgtcggcagtcggctccctcgttgaccgaatcaccgacctctctccccaggggatc caccggagcttaccgaccccgccaggaaagcaggtgcccaggccttgatctggactgcaggctggctgctgctgctgctg cttcgcggaggtgggtctgctgcccggagaggtggaggggtagagtgggggctcttagggaggaggctcctctggccc cgtacagagaagaggctgtgggaaggatccccatgtctaatcccgggacagagggcatgggaaggactatttctccttgggg gtggacgctttagcaaggagggctcaggtctattcggctccgctccccccaggagcgcaggccctggagtgctacagctgc gtgcagagccaccatggcggatcctccccgaacaagatcaagacagttaagtgcgcgccgggcgtggacgtctgcaccg aggccgtgggggcggtggagaccagtgagtggggaccccccgggccacgcacaccactgcaggctcctccggcccagc cccgcctccttatagtcctttctggatctggccacgcccctttaacgctatttgaggccccgccccctcccagaccctacccaac tctgggccacgccccatctgagaccacgcccacagctaatcccgcccttgtccctctccctagtccacggacaattctcgct ggcagtgcggggttgcggttcgggactccccggcaagattgaccgcggcctggatcttcacgggcttctggcgttcatcca gctgcagcattgcgctcaggatcgctgcaacgccaagctcaacctcacctcgcgggcgctcgacccggcaggtgccggg gttttacgagattgtgattaaggtccccagcgaccttgacgagcatctgcccggcatttctgacagctttgtgaactgggtggc cgagaaggaatgggagttgccgccagattctgacatggatctgaatctgattgagcaggcacccctgaccgtggccgaga agctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccggaggccctttttcttttgtgcaatttgagaaggg agagagctacttccacatgcacgtgctcgtggaaaccaccggggtgaaatctatggttttgggacgtttcctgagtcagattc gcgaaaaactgattcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaaagaccagaaatg gcgccggaggcgggaacaaggtggtggatgagtgctacatccccaattacttgctccccaaaacccagcctgagctccag

		tgggcgtggactaatatggaacagtatttaagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgac gcacgtgtcgcagacgcaggagcagaacaaagagaatcagaatcccaattctgatgcgccggtgatcagatcaaaaactt cagccaggtacatggagctggtcgggtggctcgtggacaaggggattacctcggagaagcagtggattcaggaggacca ggcctcatacatctccttcaatgcggcctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaaagattatg agcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacatttccagcaatcggatttataaaattt tggaactaaacgggtacgatccccaatatgcggcttccgtctttctgggatgggccacgaaaaagttcggcaagaggaaca ccatctggctgtttgggcctgcaactaccggaaagaccaacatcgcggaggccatagcccacactgtgcccttctacgggt gcgtaaactggaccaatgagaactttcccttcaacgactgtgtcgacaagatggtgatctggtgggaggaggggaagatga ccgccaaggtcgtggagtcggccaaagccattctcggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggc ccagatagacccgactcccgtgatcgtcacctccaacaccaacatgtgcgccgtgattgacgggaactcaacgaccttcga acaccagcagccgttgcaagaccggatgttcaaatttgaactcacccgccgtctggatcatgactttgggaaggtcaccaag caggaagtcaaagacttttccggtgggcaaaggatcacgtggttgaggtggagcatgagttctacgtcaaaaagggtgga gccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgcgagtcagttgcgcagccatcg acgtcagacgcggaggcttcgatcaactacgcagacaggtaccaaaacaaatgttctcgtcacgtgggcatgaatctgatg ctgtttccctgcagacaatgcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgtttagagtgc tttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaaactgtgctacattcatcatatcatgggaaagg tgccagacgcttgcactgcctgcgatctggtcaatgtggatttggatgactgcatctttgaacaataaaggtaaacttgttaatc aataaaccgtttaattcgtttcagttgaacttttctctgctaactttctttctttctagtttggacttccctccaggtaaccggtcggg gcaacatgactacgtccggcgttccatttggcatgacactacgaccaacacgatctcggttgtctcggcgcactccgtacagt agggatcgtctacctccttttgagacagaaacccgcgctaccatactggaggatcatccgctgctgcccgaatgtaacacttt gacaatgcacaacgtgagttacgtgcgaggacttccctgcagtgtgggatttacgctgattcaggaatgggttgttccctggg atatggttctaacgcgggaggagcttgtaatcctgaggaagtgtatgcacgtgtgcctgtgttgtgccaacattgacatcatga cgagcatgatgatccacggttacgagtcctgggctctccactgtcattgttccagtcccggttccctgcagtgtatagccggc gggcaggttttggccagctggtttaggatggtggtggatggcgccatgtttaatcagaggtttatatggtatcgggaggtggt gaattacaacatgccaaaagaggtaatgtttatgtccagcgtgtttatgaggggtcgccacttaatctacctgcgcttgtggtat gatggccacgtgggttctgtggtccccgccatgagctttggatacagcgccttgcactgtgggattttgaacaatattgtggtg ctgtgctgcagttactgtgctgatttaagtgagatcagggtgcgctgctgtgcccggaggacaaggcgccttatgctgcggg cggtgcgaatcatcgctgaggagactaccgccatgttgtattcctgcaggacggagcggcggcggcagcagtttattcgcg cgctgctgcagcaccaccgccctatcctgatgcacgattatgactctacccccatgtaggttgtttgacccctccccgtgcct tccttgaccctggaaggtgccactcccactgtccttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcatt ctattctgggggtggggtggggcaggacagcaaggggatccctgcaggcgcaggcggtcagcatcgccatggagac caacacccttcccaccgccactcccccttcctctcagggtccctgtcccctccagtgaatcccagaaggactctggagagttc tgagcaggggggcggcactctggcctctgattggtccaaggaaggctgggggggcaggacgggaggcgaaaaccctgga atattcccgacctggcagcctcatcgagctcggtgattggctcagaagggaaaaggcgggtctccgtgacgacttataaaa gcccaggggcaagcggtccggataacggctacctgaggagctgctgcgacagtccactacctttttcgaggtgactcccgt tgtcccaaggcttcccagagcgaacctgtgcggctgcaggcaccggcgcgtcgagtttccggcgtccggaaggaccgag ctcttctcgctgatccgtgttccgtttccgcccccaatctctcgttgccgacgcagggaaccggcggcaaccatggccagtc gggaagaggagcagcgcgaaaccacccccgagcgcggacgcggtgcggcgcgacgtccccccaaccatggaggacgt gtcgtccccgtccccgtcgccgccgcctccccgggcgcccccaaaaaagcggatgaggcggcgtatcgagtccgagga cgaggaggactcatcacaagacgcgctggtgccgcgcacacccagcccgcggccatcgacctcggcggcggatttggc cattgcgcccaagaagaaaaagaagcgcccttctcccaagcccgagcgcccgccatcaccagaggtaatcgtggacagc gaggaagaaagagaagatgtggcgctacaaatggtgggtttcagcaaccctccggtgctaatcaagcatggcaaaggag gtaagcgcacagtgcggcggctgaatgaagatgacccagtggcgcgtggtatgcggacgcaagaggaagaggaagag cccagcgaagcggaaagtgaaattacggtgatgaacccgctgagtgtgccgatcgtgtctgcgtgggagaagggcatgg aggctgcgcgggcgctgatggacaagtaccacgtggataacgatctaaaggcgaacttcaaactactgcctgaccaagtg gaagctctggcagccgtatgcaagacctggctgaacgaggagcaccgcgggttgcagctgaccttcaccagcaacaaga cctttgtgacgatgatggggcgattcctgcaggcgtacctgcagtcgtttgcagaggtgacctacaagcatcacgagcccac gggctgcgcgttgtggctgcaccgctgcgctgagatcgaaggcgagcttaagtgtctacacggaagcattatgataaataa ggagcacgtgattgaaatggatgtgacgagcgaatttgggcagcgcgcgctgaaggagcagtctagcaaggccaagatc gtgaagaatcggtggggccgaaatgtggtgcagatttccaacaccgacgcaaggtgctgcgtgcacgacgcggcctgtcc ggccaatcagttttccggcaagtcttgcggcatgttcttctctgaaggcgcaaaggctcaggtggcttttaagcagatcaagg cttttatgcaggcgctgtatcctaacgcccagaccgggcacggtcaccttttgatgccactacggtgcgagtgcaactcaaa gcctgggcacgcgcccttttgggaaggcagctaccaaagttgactccgttcgccctgagcaacgcggaggacctggacg cggatctgatctccgacaagagcgtgctggccagcgtgcaccacccggcgctgatagtgttccagtgctgcaaccctgtgt atcgcaactcgcgggcgcagggcggaggccccaactgcgacttcaagatatcggcgcccgacctgctaaacgcgttggt gatggtgcgcagcctgtggagtgaaaacttcaccgagctgccgcggatggttgtgcctgagtttaagtggagcactaaaca ccagtatcgcaacgtgtccctgccagtggcgcatagcgatgcgcggcagaaccccttttgatttttaaacggcgcagacggc aagtaaataatcacccgagagtgtacaaataaaagatcttgccacaaaatagttacaggaccaagcgagcgtgagagtcca aggcgtgcgcagtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactcttccgtggtctggtggataa

		attcgcaagggtatcatggcggacgacccgggttcgaaccccggatcgggccgtccgccgtgatccatgcggttaccgcc cgcgtgtcgaacccaggtgtgcgacgtcagacaacgggggagcgctccttttggcttccttccaggcgcggcggctgctg cgctcttttttcggccgcggccgc
60	Rep Construct #2	caattgagggtggaggtcctgtattagaggtcacgtgagtgttttgcgacattttgcgacaccatgtggtcacgctgggtattta agcccgagtgagcacgcagggctctccatttgaagcgggaggtttgaacgcgcagccgccgaccccgccaggaaagc aggtgcccaggccttgatctggactgcaggctggctgctgctgctgcttcgcggaggtgggtctgctgcccggagagg tggagggggtagagtgggggctcttagggaggaggctcctctggccccgtacagagaaggctgtgggaaggatccccat gtctaatcccgggacagagggcatgggaaggactatttctccttgggggtggacgctttagcaaggagggctcaggtctatt cggctccgctcccccaggagcgcaggccctggagtgctacagctgcgtgcagagccaccatggcggatcctccccgaac aagatcaagacagttaagtgcgcgccgggcgtggacgtctgcaccgaggccgtgggggcggtggagaccagtgagtgg ggacccccgggccacgcacaccactgcaggctcctccggcccagccccgcctccttatagtcctttctggatctggccacg ccccttaacgctatttgaggccccgcccctcccagaccctacccaactctgggccacgccccatctgagaccacgcccac agctaatcccgcccttgtccctctccctagtccacggacaattctcgctggcagtgcggggttgcggttcgggactccccgg caagattgaccgcggcctggatcttcacgggcttctggcgttcatccagctgcagcattgcgctcaggatcgctgcaacgcc aagctcaacctcacctcgcgggcgctcgacccggcaggtgccgggggttttacgagattgtgattaaggtccccagcgacct tgacgagcatctgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttgccgccagattctgac atggatctgaatctgattgagcaggcacccctgaccgtggccgagaagctgcagcgcgactttctgacggaatggcgccgt gtgagtaaggccccggaggccctttttctttgtgcaatttgagaagggagagagctacttccacatgcacgtgctcgtggaaa ccaccggggtgaaatctatggttttgggacgtttcctgagtcagattcgcgaaaaactgattcagagaatttaccgcgggatc gagccgactttgccaaactggttcgcggtcacaaagaccagaaatggcgccggaggcgggaacaaggtggtggatgagt gctacatccccaattacttgctccccaaaacccagcctgagctccagtgggcgtggactaatatggaacagtatttaagcgcc tgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgcacgtgtcgcagacgcaggagcagaacaaaga gaatcagaatcccaattctgatgcgccggtgatcagatcaaaaacttcagccaggtacatggagctggtcgggtggctcgtg gacaaggggattacctcggagaagcagtggattcaggaggaccaggcctcatacatctccttcaatgcggcctccaactcg cggtcccaaatcaaggctgccttggacaatgcgggaaagattatgagcctgactaaaaccgcccccgactacctggtggg ccagcagcccgtggaggacatttccagcaatcggatttataaaattttggaactaaacgggtacgatccccaatatgcggctt ccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatctggctgtttgggcctgcaactaccggaaaga ccaacatcgcggaggccatagcccacactgtgcccttctacgggtgcgtaaactggaccaatgagaactttcccttcaacga ctgtgtcgacaagatggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggccaaagccattctcg gaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagatagacccgactcccgtgatcgtcacctccaa caccaacatgtgcgccgtgattgacgggaactcaacgaccttcgaacaccagcagccgttgcaagaccggatgttcaaattt gaactcacccgccgtctggatcatgactttgggaaggtcaccaagcaggaagtcaaagacttttccggtgggcaaaggat cacgtggttgaggtggagcatgagttctacgtcaaaaagggtggagccaagaaaagacccgcccccagtgacgcagatat aagtgagcccaaacgggtgcgcgagtcagttgcgcagccatcgacgtcagacgcggaggcttcgatcaactacgcagac aggtaccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaatgcgagagaatgaatcagaa ttcaaatatctgcttcactcacggacagaaagactgtttagagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaa aggcgtatcagaaactgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcgatctggtcaatgtg gatttggatgactgcatctttgaacaataaaggtaaacttgttaatcaataaaccgtttaattcgtttcagttgaacttttctctgcta actttctttctttctagtttggacttccctccaggtaaccggtcggggcaacatgactacgtccggcgttccatttggcatgacac tacgaccaacacgatctcggttgtctcggcgcactccgtacagtagggatcgtctacctcctttgagacagaaacccgcgct accatactggaggatcatccgctgctgcccgaatgtaacactttgacaatgcacaacgtgagttacgtgcgaggacttccctg cagtgtgggatttacgctgattcaggaatgggttgttccctgggatatggttctaacgcgggaggagcttgtaatcctgagga agtgtatgcacgtgtgcctgtgttgtgccaacattgacatcatgacgagcatgatgatccacggttacgagtcctgggctctcc actgtcattgttccagtcccggttccctgcagtgtatagccggcgggcaggttttggccagctggtttaggatggtggtggatg gcgccatgtttaatcagaggtttatatggtatcgggaggtggtgaattacaacatgccaaaagaggtaatgtttatgtccagcg tgtttatgaggggtcgccacttaatctacctgcgcttgtggtatgatggccacgtgggttctgtggtccccgccatgagctttgg atacagcgccttgcactgtgggattttgaacaatattgtggtgctgtgctgcagttactgtgctgatttaagtgagatcagggtg cgctgctgtgcccggaggacaaggcgccttatgctgcgggcggtgcgaatcatcgctgaggagactaccgccatgttgtat tcctgcaggacggagcggcggcggcagcagtttattcgcgcgctgctgcagcaccaccgccctatcctgatgcacgattat gactctacccccatgtaggttgtttgacccctccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaa taaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcaggacagcaagggg gatccctgcaggcgcaggcggtcagcatcgccatggagaccaacaccctcccaccgccactccccctcctctcagggtc cctgtcccctccagtgaatcccagaaggactctggagagttctgagcaggggcggcactctggcctctgattggtccaag gaaggctgggggcaggacgggaggcgaaaaccctggaatattcccgacctggcagcctcatcgagctcggtgattggc tcagaagggaaaaggcgggtctccgtgacgacttataaaagcccaggggcaagcggtccggataacggctacctgagga gctgctgcgacagtccactacctttttcgaggtgactcccgttgtcccaaggcttcccagagcgaacctgtgcggctgcagg caccggcgcgtcgagtttccggcgtccggaaggaccgagctcttctcgctgatccgtgttccgtttccgcccccaatctctcg ttgccgacgcagggaaccggcggcaaccatggccagtcgggaagaggagcagcgcgaaaccacccccgagcgcgga cgcggtgcggcgcgacgtccccaaccatggaggacgtgtcgtccccgtccccgtcgccgccgcctccccgggcgccc

		ccaaaaaagcggatgaggcggcgtatcgagtccgaggacgaggaggactcatcacaagacgcgctggtgccgcgcaca cccagcccgcggccatcgacctcggcggcggatttggccattgcgcccaagaagaaaaagaagcgcccttctcccaagc ccgagcgcccgccatcaccagaggtaatcgtggacagcgaggaagaaagagaagatgtggcgctacaaatggtgggttt cagcaaccctccggtgctaatcaagcatggcaaaggaggtaagcgcacagtgcggcggctgaatgaagatgacccagtg gcgcgtggtatgcggacgcaagaggaagaggaagagcccagcgaagcggaaagtgaaattacggtgatgaacccgctg agtgtgccgatcgtgtctgcgtgggagaagggcatggaggctgcgcgggcgctgatggacaagtaccacgtggataacg atctaaaggcgaacttcaaactactgcctgaccaagtggaagctctggcagccgtatgcaagacctggctgaacgaggag caccgcgggttgcagctgaccttcaccagcaacaagacctttgtgacgatgatggggcgattcctgcaggcgtacctgcag tcgtttgcagaggtgacctacaagcatcacgagcccacgggctgcgcgttgtggctgcaccgctgcgctgagatcgaagg cgagcttaagtgtctacacggaagcattatgataaataaggagcacgtgattgaaatggatgtgacgagcgaatttgggcag cgcgcgctgaaggagcagtctagcaaggccaagatcgtgaagaatcggtggggccgaaatgtggtgcagatttccaaca ccgacgcaaggtgctgcgtgcacgacgcggcctgtccggccaatcagttttccggcaagtcttgcggcatgttcttctctga aggcgcaaaggctcaggtggctttttaagcagatcaaggcttttatgcaggcgctgtatcctaacgcccagaccgggcacgg tcacctttttgatgccactacggtgcgagtgcaactcaaagcctgggcacgcgcccttttttgggaaggcagctaccaaagttg actccgttcgccctgagcaacgcggaggacctggacgcggatctgatctccgacaagagcgtgctggccagcgtgcacc acccggcgctgatagtgttccagtgctgcaaccctgtgtatcgcaactcgcgggcgcagggcggaggccccaactgcgac ttcaagatatcggcgcccgacctgctaaacgcgttggtgatggtgcgcagcctgtggagtgaaaacttcaccgagctgccg cggatggttgtgcctgagtttaagtggagcactaaacaccagtatcgcaacgtgtccctgccagtggcgcatagcgatgcgc ggcagaaccccctttgatttttaaacggcgcagacggcaagtaaataatcacccgagagtgtacaaataaaagatcttgccac aaaatagttacaggaccaagcgagcgtgagagtccaaggcgtgcgcagtcgttgacgctctagaccgtgcaaaaggaga gcctgtaagcgggcactcttccgtggtctggtggataaattcgcaagggtatcatggcggacgacccgggttcgaaccccg gatcgggccgtccgccgtgatccatgcggttaccgcccgcgtgtcgaacccaggtgtgcgacgtcagacaacgggggag cgctccttttggcttccttccaggcgcggcggctgctgcgctctttttcggccggcggccgc
61	Rep Construct #3	caattgtcgcgaaataaaatatctttattttcattacatctgtgtgttggttttttgtgtgaatcgatagtactaacatacgctctccat caaaacaaaacgaaacaaaacaaactagcaaaatagggctgtccccagtgcaagtgcaggtgccagaacatttctctggcct aactggccggtacctgagctcgctagcctcgagcctgcaggcgacatgaagaatctgcttagggttaggcgttttgcgctgc ttcgcgttgtacgggccagatatacgcgtatctgaggggactagggtgtgtttaggcgaaaagcggggcttcggttgtacgc ggttaggagtccctcaggatatagtagtttcgcttttgcatagggagggggattgcgccttttccaaggcagccctgggtttg cgcagggacgcggctgctctgggcgtggttccgggaaacgcagcggcgccgaccctgggtctcgcacattcttcacgtcc gttcgcagcgtcacccggatcttcgccgctacccttgtgggccccccggcgacgcttcctgctccgcccctaagtcgggaa ggttccttgcggttcgcggcgtgccggacgtgacaaacggaagccgcacgtctcactagtaccctcgcagacggacagcg ccagggagcaatggcagcgcgccgaccgcgatgggctgtggccaatagcggctgctcagcagggcgcgccgagagca gcggccgggaaggggcggtgcgggaggcggggtgtggggcggtagtgtgggccctgttcctgcccgcgcggtgttccg cattctgcaagcctccggagcgcacgtcggcagtcggctccctcgttgaccgaatcaccgacctctctccccaggggatc caccggagcttacctttttatccaagctgtgcttgtgatactgtctgtcttcttgtacttgtttgcctacaaagagtacctcgcctgc ctcgtgctggccttggccctgggctgggcgaacaagctctactatacgcggggtttccagtccatcggcttgtacagcgtcat catccagaaggtgcgttgagagctgcgccggcgggaggggcctcggaggatggccatttttcagatggggaaactgagat ctggagaggaagaggggtttaccaaggggcagttagggtgggggtgggagagggtgggaagcaggaagatcagggt gccgtgaatatgtgcttgcaccttacacttgaattatgtttatttccttataggtcatttgccaccatgttctggagttcttgtttgtat atatcgtgtttttgcttggatttggagttggtaattgatttataaataatagtaactcccacatttataaacacttttatacagatgtcaa aatggcaaagtgcatggcggccttggcaacaattagaccagaaggaggcagcaggcttgtcatgtaggcaaatacacgca agcttgtagagtaattaccaggcagagagctgaggagagcaccctgttatgttcccttagccttggcctcgcttatcgagaag tgtcccaaagacaacaaggactgcagctcctacggcagcttcagcgacgcagtgctggaactcttcaagctcaccattggc ctgggtgacctggacatccagcagaactccaagtatcccattctctttctgttcctgctcatcacctttgtcatcctcacctttgttc tcctcctcaacaagctcattgctcttatcggcgagactgtggagaacgtctccaaggagagcgaacgcatctggcgcctgcg tgccggggttttacgagattgtgattaaggtccccagcgaccttgacgagcatctgcccggcatttctgacagctttgtgaact gggtggccgagaaggaatgggagttgccgccagattctgacatggatctgaatctgattgagcaggcacccctgaccgtg gccgagaagctgcagcgcgactttctgacggaatggcgccgtgtgagtaaggccccggaggccctttttctttgtgcaatttg agaagggagagagctacttccacatgcacgtgctcgtggaaaccaccggggtgaaatctatggttttgggacgtttcctgag tcagattcgcgaaaaactgattcagagaatttaccgcgggatcgagccgactttgccaaactggttcgcggtcacaaagacc agaaatggcgccggaggcgggaacaaggtggtggatgagtgctacatccccaattacttgctccccaaaacccagcctga gctccagtgggcgtggactaatatggaacagtatttaagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcag catctgacgcacgtgtcgcagacgcaggagcagaacaaagagaatcagaatcccaattctgatgcgccggtgatcagatc aaaaacttcagccaggtacatggagctggtcgggtggctcgtggacaaggggattacctcggagaagcagtggattcagg aggaccaggcctcatacatctccttcaatgcggcctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaa agattatgagcctgactaaaaccgcccccgactacctggtgggccagcagcccgtggaggacatttccagcaatcggattt ataaaatttttggaactaaacgggtacgatccccaatatgcggcttccgtctttctgggatgggccacgaaaaagttcggcaag aggaacaccatctggctgtttgggcctgcaactaccggaaagaccaacatcgcggaggccatagcccacactgtgccttc tacgggtgcgtaaactggaccaatgagaactttcccttcaacgactgtgtcgacaagatggtgatctggtgggaggagggg

		aagatgaccgccaaggtcgtggagtcggccaaagccattctcggaggaagcaaggtgcgcgtggaccagaaatgcaagt cctcggcccagatagacccgactcccgtgatcgtcacctccaacaccaacatgtgcgccgtgattgacgggaactcaacga ccttcgaacaccagcagccgttgcaagaccggatgttcaaatttgaactcacccgccgtctggatcatgactttgggaaggtc accaagcaggaagtcaaagacttttccggtgggcaaaggatcacgtggttgaggtggagcatgagttctacgtcaaaaag ggtggagccaagaaaagacccgcccccagtgacgcagatataagtgagcccaaacgggtgcgcgagtcagttgcgcag ccatcgacgtcagacgcggaggcttcgatcaactacgcagacaggtaccaaaacaaatgttctcgtcacgtgggcatgaat ctgatgctgtttccctgcagacaatgcgagagaatgaatcagaattcaaatatctgcttcactcacggacagaaagactgttta gagtgctttcccgtgtcagaatctcaacccgtttctgtcgtcaaaaaggcgtatcagaaactgtgctacattcatcatatcatgg gaaaggtgccagacgcttgcactgcctgcgatctggtcaatgtggatttggatgactgcatctttgaacaataaaggtaaactt gttaatcaataaaccgtttaattcgtttcagttgaacttttctctgctaactttctttctttctagtttggacttccctccaggtaaccgg tcggggcaacatgactacgtccggcgttccatttggcatgacactacgaccaacacgatctcggttgtctcggcgcactccg tacagtagggatcgtctacctccttttgagacagaaacccgcgctaccatactggaggatcatccgctgctgcccgaatgtaa cactttgacaatgcacaacgtgagttacgtgcgaggacttccctgcagtgtgggatttacgctgattcaggaatgggttgttcc ctgggatatggttctaacgcgggaggagcttgtaatcctgaggaagtgtatgcacgtgtgcctgtgttgtgccaacattgacat catgacgagcatgatgatccacggttacgagtcctgggctctccactgtcattgttccagtcccggttccctgcagtgtatagc cggcgggcaggttttggccagctggtttaggatggtggtggatggcgccatgtttaatcagaggtttatatggtatcgggagg tggtgaattacaacatgccaaaagaggtaatgtttatgtccagcgtgtttatgaggggtcgccacttaatctacctgcgcttgtg gtatgatggccacgtgggttctgtggtccccgccatgagctttggatacagcgccttgcactgtgggattttgaacaatattgt ggtgctgtgctgcagttactgtgctgatttaagtgagatcagggtgcgctgctgtgcccggaggacaaggcgccttatgctg cgggcggtgcgaatcatcgctgaggagactaccgccatgttgtattcctgcaggacggagcggcggcggcagcagtttatt cgcgcgctgctgcagcaccaccgccctatcctgatgcacgattatgactctacccccatgtaggttgtttgacccctccccg tgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtg tcattctattctgggggggtggggtggggcaggacagcaagggggatccctgcaggcgcaggcggtcagcatcgccatgg agaccaacaccctcccaccgccactccccttcctctcagggtccctgtcccctccagtgaatcccagaaggactctggag agttctgagcaggggcggcactctggcctctgattggtccaaggaaggctgggggcaggacgggaggcgaaaaccc tggaatattcccgacctggcagcctcatcgagctcggtgattggctcagaagggaaaaggcgggtctccgtgacgacttata aaagcccaggggcaagcggtccggataacggctacctgaggagctgctgcgacagtccactacctttttcgaggtgactcc cgttgtcccaaggcttcccagagcgaacctgtgcggctgcaggcaccggcgcgtcgagtttccggcgtccggaaggacc gagctcttctcgctgatccgtgttccgtttccgcccccaatctctcgttgccgacgcagggaaccggcggcaaccatggcca gtcgggaagaggagcagcgcgaaaccacccccgagcgcggacgcggtgcggcgcgacgtcccccaaccatggagga cgtgtcgtccccgtccccgtcgccgccgcctccccgggcgcccccaaaaaagcggatgaggcggcgtatcgagtccgag gacgaggaggactcatcacaagacgcgctggtgccgcgcacacccagcccgcggccatcgacctcggcggcggatttg gccattgcgcccaagaagaaaaagaagcgcccttctcccaagcccgagcgcccgccatcaccagaggtaatcgtggaca gcgaggaagaaagagaagaagatgtggcgctacaaatggtgggtttcagcaaccctccggtgctaatcaagcatggcaaagga ggtaagcgcacagtgcggcggctgaatgaagatgacccagtggcgcgtggtatgcggacgcaagaggaagaggaaga gcccagcgaagcggaaagtgaaattacggtgatgaacccgctgagtgtgccgatcgtgtctgcgtgggagaagggcatg gaggctgcgcgggcgctgatggacaagtaccacgtggataacgatctaaaggcgaacttcaaactactgcctgaccaagt ggaagctctggcagccgtatgcaagacctggctgaacgaggagcaccgcgggttgcagctgaccttcaccagcaacaag acctttgtgacgatgatggggcgattcctgcaggcgtacctgcagtcgtttgcagaggtgacctacaagcatcacgagccca cgggctgcgcgttgtggctgcaccgctgcgctgagatcgaaggcgagcttaagtgtctacacggaagcattatgataaata aggagcacgtgattgaaatggatgtgacgagcgaatttgggcagcgcgcgctgaaggagcagtctagcaaggccaagat cgtgaagaatcggtggggccgaaatgtggtgcagatttccaacaccgacgcaaggtgctgcgtgcacgacgcggcctgtc cggccaatcagttttccggcaagtcttgcggcatgttcttctctgaaggcgcaaaggctcaggtggcttttaagcagatcaag gcttttatgcaggcgctgtatcctaacgcccagaccgggcacggtcacctttgatgccactacggtgcgagtgcaactcaa agcctgggcacgcgcccttttgggaaggcagctaccaaagttgactccgttcgcccctgagcaacgcggaggacctggac gcggatctgatctccgacaagagcgtgctggccagcgtgcaccacccggcgctgatagtgttccagtgctgcaaccctgtg tatcgcaactcgcgggcgcagggcggaggccccaactgcgacttcaagatatcggcgcccgacctgctaaacgcgttggt gatggtgcgcagcctgtggagtgaaaacttcaccgagctgccgcggatggttgtgcctgagtttaagtggagcactaaaca ccagtatcgcaacgtgtccctgccagtggcgcatagcgatgcgcggcagaaccccttttgatttttaaacggcgcagacggc aagtaaataatcacccgagagtgtacaaataaaagatcttgccacaaaatagttacaggaccaagcgagcgtgagagtcca aggcgtgcgcagtcgttgacgctctagaccgtgcaaaaggagagcctgtaagcgggcactcttccgtggtctggtggataa attcgcaagggtatcatggcggacgacccgggttcgaaccccggatcgggccgtccgccgtgatccatgcggttaccgcc cgcgtgtcgaacccaggtgtgcgacgtcagacaacgggggagcgctccttttggcttccttccaggcgcggcggctgctg cgctcttttttcggccgcggccgc
62	Rep Construct #4	caattgagggtggaggtcctgtattagaggtcacgtgagtgttttgcgacattttgcgacaccatgtggtcacgctgggtattta agcccgagtgagcacgcagggctctccattttgaagcgggaggtttgaacgcgcagccgccttttatccaagctgtgcttgt gatactgtctgtcttcttgtacttgtttgcctacaaagagtacctcgcctgcctcgtgctggccttggccctgggctgggcgaac aagctctactatacgcggggtttccagtccatcggcttgtacagcgtcatcatccagaaggtgcgttgagagctgcgccggc gggaggggcctcggaggatggccatttttcagatggggaaactgagatctggagaggaagaggggtttaccaaggggca

		gttagggtgggggtgggagagggtgggaagcaggaagatcagggtgccgtgaatatgtgcttgcaccttacacttgaatt atgtttatttccttataggtcatttgccaccatgttctggagttcttgtttgtatatatcgtgtttttgcttggatttggagttggtaattg atttataaataatagtaactcccacatttataaacactttatacagatgtcaaaatggcaaagtgcatggcggccttggcaacaa ttagaccagaaggaggcagcaggcttgtcatgtaggcaaatacacgcaagcttgtagagtaattaccaggcagagagctga ggagagcaccctgttatgttcccttagccttggcctcgcttatcgagaagtgtcccaaagacaacaaggactgcagctcctac ggcagcttcagcgacgcagtgctggaactcttcaagctcaccattggcctgggtgacctggacatccagcagaactccaag tatcccattctctttctgttcctgctcatcacctttgtcatcctcacctttgttctcctcctcaacaagctcattgctcttatcggcgag actgtggagaacgtctccaaggagagcgaacgcatctggcgcctgcgtgccgggggttttacgagattgtgattaaggtccc cagcgaccttgacgagcatctgcccggcatttctgacagctttgtgaactgggtggccgagaaggaatgggagttgccgcc agattctgacatggatctgaatctgattgagcaggcacccctgaccgtggccgagaagctgcagcgcgactttctgacgga atggcgccgtgtgagtaaggccccggaggccctttttctttgtgcaatttgagaagggagagagctacttccacatgcacgtg ctcgtggaaaccaccggggtgaaatctatggttttgggacgtttcctgagtcagattcgcgaaaaactgattcagagaatttac cgcgggatcgagccgactttgccaaactggttcgcggtcacaaagaccagaaatggcgccggaggcgggaacaaggtg gtggatgagtgctacatccccaattacttgctccccaaaacccagcctgagctccagtgggcgtggactaatatggaacagt atttaagcgcctgtttgaatctcacggagcgtaaacggttggtggcgcagcatctgacgcacgtgtcgcagacgcaggagc agaacaaagagaatcagaatcccaattctgatgcgccggtgatcagatcaaaaacttcagccaggtacatggagctggtcg ggtggctcgtggacaaggggattacctcggagaagcagtggattcaggaggaccaggcctcatacatctccttcaatgcgg cctccaactcgcggtcccaaatcaaggctgccttggacaatgcgggaaagattatgagcctgactaaaaccgcccccgact acctggtgggccagcagcccgtggaggacatttccagcaatcggatttataaaattttggaactaaacgggtacgatcccca atatgcggcttccgtctttctgggatgggccacgaaaaagttcggcaagaggaacaccatctggctgtttgggcctgcaact accggaaagaccaacatcgcggaggccatagcccacactgtgcccttctacgggtgcgtaaactggaccaatgagaacttt cccttcaacgactgtgtcgacaagatggtgatctggtgggaggaggggaagatgaccgccaaggtcgtggagtcggcca aagccattctcggaggaagcaaggtgcgcgtggaccagaaatgcaagtcctcggcccagatagacccgactcccgtgatc gtcacctccaacaccaacatgtgcgccgtgattgacgggaactcaacgaccttcgaacaccagcagccgttgcaagaccg gatgttcaaatttgaactcacccgccgtctggatcatgactttgggaaggtcaccaagcaggaagtcaaagacttttccggtg ggcaaaggatcacgtggttgaggtggagcatgagttctacgtcaaaaagggtggagccaagaaaagacccgcccccagt gacgcagatataagtgagcccaaacgggtgcgcgagtcagttgcgcagccatcgacgtcagacgcggaggcttcgatca actacgcagacaggtaccaaaacaaatgttctcgtcacgtgggcatgaatctgatgctgtttccctgcagacaatgcgagag aatgaatcagaattcaaatatctgcttcactcacggacagaaagactgtttagagtgctttcccgtgtcagaatctcaacccgtt tctgtcgtcaaaaaggcgtatcagaaactgtgctacattcatcatatcatgggaaaggtgccagacgcttgcactgcctgcga tctggtcaatgtggatttggatgactgcatctttgaacaataaaggtaaacttgttaatcaataaaccgtttaattcgtttcagttga acttttctctgctaactttctttctttctagtttggacttccctccaggtaaccggtcggggcaacatgactacgtccggcgttcca tttggcatgacactacgaccaacacgatctcggttgtctcggcgcactccgtacagtagggatcgtctacctccttttgagaca gaaacccgcgctaccatactggaggatcatccgctgctgcccgaatgtaacactttgacaatgcacaacgtgagttacgtgc gaggacttccctgcagtgtgggatttacgctgattcaggaatgggttgttccctgggatatggttctaacgcgggaggagctt gtaatcctgaggaagtgtatgcacgtgtgcctgtgttgtgccaacattgacatcatgacgagcatgatgatccacggttacga gtcctgggctctccactgtcattgttccagtcccggttccctgcagtgtatagccggcgggcaggttttggccagctggtttag gatggtggtggatggcgccatgtttaatcagaggtttatatggtatcgggaggtggtgaattacaacatgccaaaagaggtaa tgtttatgtccagcgtgtttatgaggggtcgccacttaatctacctgcgcttgtggtatgatggccacgtgggttctgtggtcccc gccatgagctttggatacagcgccttgcactgtgggattttgaacaatattgtggtgctgtgctgcagttactgtgctgatttaag tgagatcagggtgcgctgctgtgcccggaggacaaggcgccttatgctgcgggcggtgcgaatcatcgctgaggagacta ccgccatgttgtattcctgcaggacggagcggcggcggcagcagtttattcgcgcgctgctgcagcaccaccgccctatcc tgatgcacgattatgactctacccccatgtaggttgtttgacccctccccgtgccttccttgaccctggaaggtgccactccc actgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtggggtggggcag gacagcaaggggatccctgcaggcgcaggcggtcagcatcgccatggagaccaacaccctcccaccgccactcccc cttcctctcagggtccctgtcccctccagtgaatcccagaaggactctggagagttctgagcaggggcggcactctggcct ctgattggtccaaggaaggctgggggcaggacgggaggcgaaaaccctggaatattcccgacctggcagcctcatcga gctcggtgattggctcagaagggaaaaggcgggtctccgtgacgacttataaaagcccaggggcaagcggtccggataa cggctacctgaggagctgctgcgacagtccactacctttttcgaggtgactcccgttgtcccaaggcttcccagagcgaacc tgtgcggctgcaggcaccggcgcgtcgagtttccggcgtccggaaggaccgagctcttctcgctgatccgtgttccgtttcc gcccccaatctctcgttgccgacgcagggaaccggcggcaaccatggccagtcgggaagaggagcagcgcgaaacca cccccgagcgcggacgcggtgcggcgcgacgtccccaaccatggaggacgtgtcgtccccgtccccgtcgccgccgc ctccccgggcgcccccaaaaaagcggatgaggcggcgtatcgagtccgaggacgaggaggactcatcacaagacgcg ctggtgccgcgcacacccagcccgcggccatcgacctcggcggcggatttggccattgcgcccaagaagaaaaagaag cgcccttctcccaagcccgagcgcccgccatcaccagaggtaatcgtggacagcgaggaagaaagagaagatgtggcg ctacaaatggtgggtttcagcaaccctccggtgctaatcaagcatggcaaaggaggtaagcgcacagtgcggcggctgaa tgaagatgacccagtggcgcgtggtatgcggacgcaagaggaagaggaagagcccagcgaagcggaaagtgaaattac ggtgatgaacccgctgagtgtgccgatcgtgtctgcgtgggagaagggcatggaggctgcgcgggcgctgatggacaag taccacgtggataacgatctaaaggcgaacttcaaactactgcctgaccaagtggaagctctggcagccgtatgcaagacct

		ggctgaacgaggagcaccgcgggttgcagctgaccttcaccagcaacaagacctttgtgacgatgatggggcgattcctg caggcgtacctgcagtcgtttgcagaggtgacctacaagcatcacgagcccacgggctgcgcgttgtggctgcaccgctg cgctgagatcgaaggcgagcttaagtgtctacacggaagcattatgataaataaggagcacgtgattgaaatggatgtgacg agcgaatttgggcagcgcgcgctgaaggagcagtctagcaaggccaagatcgtgaagaatcggtggggccgaaatgtgg tgcagatttccaacaccgacgcaaggtgctgcgtgcacgacgcggcctgtccggccaatcagttttccggcaagtcttgcg gcatgttcttctctgaaggcgcaaaggctcaggtggcttttaagcagatcaaggcttttatgcaggcgctgtatcctaacgccc agaccgggcacggtcacctttgatgccactacggtgcgagtgcaactcaaagcctgggcacgcgccctttttgggaaggc agctaccaaagttgactccgttcgccctgagcaacgcggaggacctggacgcggatctgatctccgacaagagcgtgctg gccagcgtgcaccacccggcgctgatagtgttccagtgctgcaaccctgtgtatcgcaactcgcgggcgcagggcggag gccccaactgcgacttcaagatatcggcgcccgacctgctaaacgcgttggtgatggtgcgcagcctgtggagtgaaaact tcaccgagctgccgcggatggttgtgcctgagtttaagtggagcactaaacaccagtatcgcaacgtgtccctgccagtggc gcatagcgatgcgcggcagaaccccttgatttttaaacggcgcagacggcaagtaaataatcacccgagagtgtacaaat aaaagatcttgccacaaaatagttacaggaccaagcgagcgtgagagtccaaggcgtgcgcagtcgttgacgctctagac cgtgcaaaaggagagcctgtaagcgggcactcttccgtggtctggtggataaattcgcaagggtatcatggcggacgaccc gggttcgaaccccggatcgggccgtccgccgtgatccatgcggttaccgcccgcgtgtcgaacccaggtgtgcgacgtca gacaacgggggagcgctccttttttggcttccttccaggcgcggcggctgctgcgctctttttcggccggcggccgc
63	iRFP Construct #1	tcgcgaaataaaatatctttattttcattacatctgtgtgttggttttttgtgtgaatcgatagtactaacatacgctctccatcaaaa caaaacgaaacaaaacaaactagcaaaataggctgtccccagtgcaagtgcaggtgccagaacatttctctggcctaactg gccggtacctgagctcgctagcctcgagcctgcaggcgacaattgcatgaagaatctgcttagggttaggcgttttgcgctg cttcgcgttgtacgggccagatatacgcgtatctgaggggactagggtgtgtttaggcgaaaagcggggcttcggttgtacg cggttaggagtcccctcaggatatagtagtttcgcttttgcatagggagggggaaatgtagtcttatgcaatacacttgtagtct tgcaacatggtaacgatgagttagcaacatgccttacaaggagagaaaaagcaccgtgcatgccgattggtggaagtaagg tggtacgatcgtgccttattaggaaggcaacagacaggtctgacatggattggacgaaccactgaattgcgcattgcagaga taattgtatttaagtgcctagctcgatacaataaacgccatttgaccattcaccacattggtgtgcacctccaaggccacagatc cgtttaaacaagcttggcaatccggtactgttggtaaagccaccatggcggatcctccccgaacaagatcaagacagttaag tgcgcgccgggcgtggacgtctgcaccgaggccgtgggggcggtggagaccatccacggacaattctcgctggcagtg cggggttgcggttcgggactcccggcaagagcggaaggatccgtcgccaggcagcctgacctcttgacctgcgacgat gagccgatccatatccccggtgccatccaaccgcatggactgctgctcgccctcgccgccgacatgacgatcgttgccggc agcgacaaccttcccgaactcaccggactggcgatcggcgccctgatcggccgctctgcggccgatgtcttcgactcgga gacgcacaaccgtctgacgatcgccttggccgagcccggggcggccgtcggagcaccgatcactgtcggcttcacgatg cgaaaggacgcaggcttcatcggctcctggcatcgccatgatcagctcatcttcctcgagctcgagcctccccagcgggac gtcgccgagccgcaggcgttcttccgccgcaccaacagcgccatccgccgcctgcaggccgccgaaaccttggaaagc gcctgcgccgccgcggcgcaagaggtgcggaagattaccggcttcgatcgggtgatgatctatcgcttcgcctccgacttc agcggcgaagtgatcgcagaggatcggtgcgccgaggtcgagtcaaaactaggcctgcactatcctgcctcaaccgtgcc ggcgcaggcccgtcggctctataccatcaacccggtacggatcattcccgatatcaattatcggccggtgccggtcacccc agacctcaatccggtcaccgggcggccgattgatcttagcttcgccatcctgcgcagcgtctcgcccgtccatctggaattca tgcgcaacataggcatgcacggcacgatgtcgatctcgattttgcgcggcgagcgactgtggggattgatcgtttgccatca ccgaacgccgtactacgtcgatctcgatggccgccaagcctgcgagctagtcgcccaggttctggcctggcagatcggcg tgatggaagagtgagggccc
64	iRFP Construct #2	gacggatcgggagatctcccgatcccctatggtcgactctcagtacaatctgctctgatgccgcatagttaagccagtatctg ctccctgcttgtgtgttggaggtcgctgagtagtgcgcgagcaaaatttaagctacaacaaggcaaggcttgaccgacaatt gcatgaagaatctgcttagggttaggcgttttgcgctgcttcgcgaaataaaatatctttattttcattacatctgtgtgttggttttttt gtgtgaatcgatagtactaacatacgctctccatcaaaacaaaacgaaacaaaacaaactagcaaaataggctgtccccagt gcaagtgcaggtgccagaacatttctctggcctaactggccggtacctgagctcgctagcctcgagcctgcaggcgacaat tgcatgaagaatctgcttagggttaggcgttttgcgctgcttcgcgttgtacgggccagatatacgcgtatctgaggggacta gggtgtgtttaggcgaaaagcggggcttcggttgtacgcggttaggagtcccctcaggatatagtagtttcgcttttgcatagg gaggggaaatgtagtcttatgcaatacacttgtagtcttgcaacatggtaacgatgagttagcaacatgccttacaaggaga gaaaaagcaccgtgcatgccgattggtggaagtaaggtggtacgatcgtgccttattaggaaggcaacagacaggtctgac atggattggacgaaccactgaattgcgcattgcagagataattgtatttaagtgcctagctcgatacaataaacgccatttgac cattcaccacattggtgtgcacctccaaggccacagatccgtttaaacaagcttggcaatccggtactgttggtaaagctgct gctgctgctgcttcgcggaggtgggtctgctgcccggagaggtggaggggtagagtgggggctcttaggtggacgcttt agcaaggagggctcaggtctattcggctccgctccccccaggagcgcaggccctggagtgctacagctgcgtgcagagcc accatggcggatcctccccgaacaagatcaagacagttaagtgcgcgccgggcgtggacgtctgcaccgaggccgtggg ggcggtggagaccagtgagtggggaccccccgggccacgcacaccactgcaggctcctccggcccccatctgagaccac gcccacagctaatcccgcccttgtcccctctccctagtccacggacaattctcgctggcagtgcggggttgcggttcgggact ccccggcaagagtggcggaaggatccgtcgccaggcagcctgacctcttgacctgcgacgatgagccgatccatatcccc ggtgccatccaaccgcatggactgctgctcgccctcgccgccgacatgacgatcgttgccggcagcgacaaccttcccga actcaccggactggcgatcggcgccctgatcggccgctctgcggccgatgtcttcgactcggagacgcacaaccgtctga cgatcgccttggccgagcccggggcggccgtcggagcaccgatcactgtcggcttcacgatgcgaaaggacgcaggctt

		catcggctcctggcatcgccatgatcagctcatcttcctcgagctcgagcctccccagcgggacgtcgccgagccgcagg cgttcttccgccgcaccaacagcgccatccgccgcctgcaggccgccgaaaccttggaaagcgcctgcgccgccgcggc gcaagaggtgcggaagattaccggcttcgatcgggtgatgatctatcgcttcgcctccgacttcagcggcgaagtgatcgc agaggatcggtgcgccgaggtcgagtcaaaactaggcctgcactatcctgcctcaaccgtgccggcgcaggcccgtcgg ctctataccatcaacccggtacggatcattcccgatatcaattatcggccggtgccggtcaccccagacctcaatccggtcac cgggcggccgattgatcttagcttcgccatcctgcgcagcgtctcgcccgtccatctggaattcatgcgcaacataggcatg cacggcacgatgtcgatctcgattttgcgcggcgagcgactgtggggattgatcgtttgccatcaccgaacgccgtactacg tcgatctcgatggccgccaagcctgcaagagggtcgccgagaggctggccactcagatcggcgtgatggaagagtgagg gcccgtttaaacccgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgccttccttga ccctggaaggtgccactcccactgtccttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctg gggggtggggtggggcaggacagcaaggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctc tatggcttctgaggcggaaagaaccagctggggctctaggggtatccccacgcgccctgtagcggcgcattaagcgcgg cgggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccctagcgcccgctcctttcgctttcttcccttcctttct cgccacgttcgccggctttccccgtcaagctctaaatcggggcatcccttagggttccgatttagtgctttacggcacctcga ccccaaaaaacttgattagggtgatggttcacgtagtgggccatcgccctgatagacggtttttcgccctttgacgttggagtc cacgttctttaatagtggactcttgttccaaactggaacaacactcaaccctatctcggtctattcttttgatttataagggatttttg gggatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaattaattctgtggaatgtgtgtcagttaggg tgtggaaagtccccaggctccccaggcaggcagaagtatgcaaagcatgcatctcaattagtcagcaaccaggtgtggaa agtccccaggctccccagcaggcagaagtatgcaaagcatgcatctcaattagtcagcaaccatagtcccgccctaactc cgcccatcccgcccctaactccgcccagttccgcccattctccgccccatggctgactaatttttttttatttatgcagaggccga ggccgcctctgcctctgagctattccagaagtagtgaggaggcttttttggaggcctaggcttttgcaaaaagctcccgggag cttgtatatccattttcggatctgatcagcacgtgatgaaaaagcctgaactcaccgcgacgtctgtcgagaagtttctgatcga aaagttcgacagcgtctccgacctgatgcagctctcggagggcgaagaatctcgtgctttcagcttcgatgtaggagggcgt ggatatgtcctgcgggtaaatagctgcgccgatggtttctacaaagatcgttatgtttatcggcactttgcatcggccgcgctc ccgattccggaagtgcttgacattggggaattcagcgagagcctgacctattgcatctcccgccgtgcacagggtgtcacgt tgcaagacctgcctgaaaccgaactgcccgctgttctgcagccggtcgcggaggccatggatgcgatcgctgcggccgat cttagccagacgagcgggttcggcccattcggaccgcaaggaatcggtcaatacactacatggcgtgatttcatatgcgcg attgctgatccccatgtgtatcactggcaaactgtgatggacgacaccgtcagtgcgtccgtcgcgcaggctctcgatgagct gatgctttgggccgaggactgccccgaagtccggcacctcgtgcacgcggatttcggctccaacaatgtcctgacggacaa tggccgcataacagcggtcattgactggagcgaggcgatgttcggggattcccaatacgaggtcgccaacatcttcttctgg aggccgtggttggcttgtatggagcagcagacgcgctacttcgagcggaggcatccggagcttgcaggatcgccgcggct ccgggcgtatatgctccgcattggtcttgaccaactctatcagagcttggttgacggcaatttcgatgatgcagcttgggcgca gggtcgatgcgacgcaatcgtccgatccggagccgggactgtcgggcgtacacaaatcgcccgcagaagcgcggccgt ctggaccgatggctgtgtagaagtactcgccgatagtggaaaccgacgccccagcactcgtccgagggcaaaggaatag cacgtgctacgagatttcgattccaccgccgccttctatgaaaggttgggcttcggaatcgttttccgggacgccggctggat gatcctccagcgcggggatctcatgctggagttcttcgcccaccccaacttgtttattgcagcttataatggttacaaataaagc aatagcatcacaaatttcacaaataaagcatttttttcactgcattctagttgtggtttgtccaaactcatcaatgtatcttatcatgt ctgtataccgtcgacctctagctagagcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattcc acacaacatacgagccggaagcataaagtgtaaagcctggggtgcctaatgagtgagctaactcacattaattgcgttgcgc tcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcggtttg cgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactc aaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggc caggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaatcgacgctc aagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctctcctgtt ccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcaatgctcacgctgtaggtat ctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccg gtaactatcgtcttgagtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagag cgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgc gctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttt ttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctca gtggaacgaaaactcacgttaagggattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatga agttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgat ctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagt gctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagc gcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagtta atagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcc caacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaag taagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgt gactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggat

		aataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttacc gctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgag caaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctttttca atattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggttcc gcgcacatttccccgaaaagtgccacctgacgtc
65	iRFP Construct #3	tcgcgaaataaaatatctttatttttcattacatctgtgtgttggttttttgtgtgaatcgatagtactaacatacgctctccatcaaaa caaaacgaaacaaaacaaactagcaaaataggctgtccccagtgcaagtgcaggtgccagaacatttctctggcctaactg gccggtacctgagctcgctagcctcgagcctgcaggcgacaattgcatgaagaatctgcttagggttaggcgttttgcgctg cttcgcgttgtacgggccagatatacgcgtatctgaggggactagggtgtgtttaggcgaaaagcggggcttcggttgtacg cggttaggagtcccctcaggatatagtagtttcgcttttgcatagggaggggggaaatgtagtcttatgcaatacacttgtagtct tgcaacatggtaacgatgagttagcaacatgccttacaaggagagaaaaagcaccgtgcatgccgattggtggaagtaagg tggtacgatcgtgccttattaggaaggcaacagacaggtctgacatggattggacgaaccactgaattgcgcattgcagaga taattgtatttaagtgcctagctcgatacaataaacgccatttgaccattcaccacattggtgtgcacctccaaggccacagatc cgtttaaacaagcttggcaatccggtactgttggtaaagccaccatggcggaaggatccgtcgccaggcagcctgacctctt gacctgcgacgatgagccgatccatatccccggtgccatccaaccgcatggactgctgctcgccctcgccgccgacatga cgatcgttgccggcagcgacaaccttcccgaactcaccggactggcgatcggcgccctgatcggccgctctgcggccgat gtcttcgactcggagacgcacaaccgtctgacgatcgccttggccgagcccggggcggccgtcggagcaccgatcactgt cggcttcacgatgcgaaaggacgcaggcttcatcggctcctggcatcgccatgatcagctcatcttcctcgagctcgagcct ccccagcgggacgtcgccgagccgcaggcgttcttccgccgcaccaacagcgccatccgccgcctgcaggccgccgaa accttggaaagcgcctgcgccgccgcggcgcaagaggtgcggaagattaccggcttcgatcgggtgatgatctatcgcttc gcctccgacttcagcggcgaagtgatcgcagaggatcggtgcgccgaggtcgagtcaaaactaggcctgcactatcctgc ctcaaccgtgccggcgcaggcccgtcggctctataccatcaacccggtacggatcattcccgatatcaattatcggccggtg ccggtcaccccagacctcaatccggtcaccgggcggccgattgatcttagcttcgccatcctgcgcagcgtctcgcccgtc catctggaattcatgcgcaacataggcatgcacggcacgatgtcgatctcgattttgcgcggcgagcgactgtggggattga tcgtttgccatcaccgaacgccgtactacgtcgatctcgatggccgccaagcctgcaagagggtcgccgagaggctggcc actcagatcggcgtgatggaagagtgagggccc
66	iRFP Construct #4	tcgcgaaataaaatatctttatttttcattacatctgtgtgttggttttttgtgtgaatcgatagtactaacatacgctctccatcaaaa caaaacgaaacaaaacaaactagcaaaataggctgtccccagtgcaagtgcaggtgccagaacatttctctggcctaactg gccggtacctgagctcgctagcctcgagcctgcaggcgacaattgcatgaagaatctgcttagggttaggcgttttgcgctg cttcgcgttgtacgggccagatatacgcgtatctgaggggactagggtgtgtttaggcgaaaagcggggcttcggttgtacg cggttaggagtcccctcaggatatagtagtttcgcttttgcatagggaggggggaaatgtagtcttatgcaatacacttgtagtct tgcaacatggtaacgatgagttagcaacatgccttacaaggagagaaaaagcaccgtgcatgccgattggtggaagtaagg tggtacgatcgtgccttattaggaaggcaacagacaggtctgacatggattggacgaaccactgaattgcgcattgcagaga taattgtatttaagtgcctagctcgatacaataaacgccatttgaccattcaccacattggtgtgcacctccaaggccacagatc cgtttaaacaagcttggcaatccggtactgttggtaaatctactatacgcggggtttccagtccatcggcttgtacagcgtcatc atccagaaggtgcgttgagagctgcgccggcgggaggggcctcggaggatggccattttgtgaatatgtgcttgcaccttа cacttgaattatgtttatttccttataggtcatttgccaccatgttctggagttcttgtttgtatatatcgtgtttttgcttggatttggag ttggtaattgatttataaataatagtaactcccacatttataaacactttatagtaattaccaggcagagagctgaggagagcac cctgttatgttcccttagccttggcctcgcttatcgagaagtgtcccaaagacaacaaggactgcagctcctacggcagcttca gcgacgcagtgctggaactctgtggcggaaggatccgtcgccaggcagcctgacctcttgacctgcgacgatgagccgat ccatatccccggtgccatccaaccgcatggactgctgctcgccctcgccgccgacatgacgatcgttgccggcagcgaca accttcccgaactcaccggactggcgatcggcgccctgatcggccgctctgcggccgatgtcttcgactcggagacgcac aaccgtctgacgatcgccttggccgagcccggggcggccgtcggagcaccgatcactgtcggcttcacgatgcgaaagg acgcaggcttcatcggctcctggcatcgccatgatcagctcatcttcctcgagctcgagcctccccagcgggacgtcgccg agccgcaggcgttcttccgccgcaccaacagcgccatccgccgcctgcaggccgccgaaaccttggaaagcgcctgcgc cgccgcggcgcaagaggtgcggaagattaccggcttcgatcgggtgatgatctatcgcttcgcctccgacttcagcggcga agtgatcgcagaggatcggtgcgccgaggtcgagtcaaaactaggcctgcactatcctgcctcaaccgtgccggcgcagg cccgtcggctctataccatcaacccggtacggatcattcccgatatcaattatcggccggtgccggtcaccccagacctcaat ccggtcaccgggcggccgattgatcttagcttcgccatcctgcgcagcgtctcgcccgtccatctggaattcatgcgcaaca taggcatgcacggcacgatgtcgatctcgattttgcgcggcgagcgactgtggggattgatcgtttgccatcaccgaacgcc gtactacgtcgatctcgatggccgccaagcctgcgagctagtcgcccaggttctggcctggcagatcggcgtgatggaaga gtgagggccc

WHAT IS CLAIMED IS:

1. An expression cassette comprising from 5' to 3': (i) promoter, (ii) an upstream exon, (iii) an upstream intron, (iv) an alternatively spliced exon comprising a translation initiation sequence, (v) a downstream intron, (vi) a downstream exon, and (vii) a heterologous nucleic acid sequence; wherein the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.
2. The expression cassette of claim 1, wherein the upstream exon and the downstream exon each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene.
3. The expression cassette of claim 1, wherein the upstream exon and the downstream exon each comprise a portion of a gene, which is different from the gene for the portion of the upstream intron, the alternatively spliced exon, and the downstream intron.
4. The expression cassette of claim 1, wherein the upstream exon and the downstream exon comprise synthetic exons.
5. The expression cassette of any one of claims 1-4, wherein the alternatively spliced exon is regulatable by a splicing modulator.
6. The expression cassette of claim 5, wherein the splicing modulator is Kinetin or Homocarbonyltopsentin.
7. The expression cassette of any one of claims 1-6, wherein the gene is TRPV3, STK31, or USP25; and the splicing modulator is Kinetin.
8. The expression cassette of any one of claims 1-6, wherein the gene is GTF2B, LYPD3, IK, or TEX14; and the splicing modulator is Homocarbonyltopsentin.

9. The expression cassette of any one of claims 5-8, wherein the splicing modulator promotes inclusion of the alternatively spliced exon thereby driving expression of the heterologous nucleic acid sequence.
10. The expression cassette of any one of claims 1-9, wherein the alternatively spliced exon further comprises a kozak sequence.
11. The expression cassette of any one of claims 1-9, wherein the promoter is selected from the group consisting of a CBA promoter, a smCBA promoter, a CMV promoter, an EF-1a (Elongation Factor 1a) promoter, a RSV (Rous Sarcoma Virus) promoter, an Ubiquitin (UbC) promoter, a CAG promoter, a SV40 (simian vacuolating virus 40) promoter, a PGK (phosphoglycerate kinase) promoter, a human beta actin promoter, a H1 (human polymerase III RNA) promoter, a human U6 small nuclear promoter, tetracycline responsive element, and any combination thereof.
12. The expression cassette of any one of the previous claims, further comprising an enhancer.
13. The expression cassette of claim 12, wherein the enhancer is selected from the group consisting of a CMV enhancer, a SV40 enhancer, and any combination thereof.
14. The expression cassette of any one of claims 1-13, which further comprises a cleavage site between the downstream exon and the heterologous nucleic acid sequence.
15. The expression cassette of claim 14, wherein the cleavage site is selected from the group consisting of a furin cleavage site, a foot-and-mouth disease virus 18 2A peptide (F2A), an equine rhinitis A virus 2A peptide (E2A), a porcine teschovirus-1 2A peptide (P2A), a thosea asigna virus 2A peptide (T2A), and any combination thereof.
16. The expression cassette of any one of the claims 1-15, which further comprises a leader sequence between the downstream exon and the heterologous nucleic acid sequence.
17. The expression cassette of claim 16, wherein the leader sequence is selected from the group consisting of a human opticin (hOPT) leader sequence, an interleukin-2 (IL-2) leader sequence, an interleukin-12 (IL-12) leader sequence, an interleukin-6 (IL-6) leader sequence, a human insulin leader sequence, a human serum albumin (HSA) leader sequence, and any combination thereof.
18. The expression cassette of any one of the previous claims, wherein the combined length of the upstream exon, the upstream intron, the alternatively spliced exon, the downstream intron, and the downstream exon is between about 300 basepairs to about 2100 basepairs.

19. The expression cassette of any one of the previous claims, wherein the length of the heterologous nucleic acid sequence is less than about 4100 basepairs, less than about 4000 basepairs, or less than about 3900 basepairs.
20. The expression cassette of any one of the previous claims, wherein the heterologous nucleic acid sequence encodes a protein.
21. The expression cassette of claim 20, wherein the protein is a therapeutic protein.
22. The expression cassette of any one of the previous claims, wherein the upstream exon comprises a sequence selected from the group consisting of SEQ ID NO: 8, 13, 18, 23, 28, 33, and 38.
23. The expression cassette of any one of the previous claims, wherein the alternatively spliced exon comprises a sequence selected from the group consisting of SEQ ID NO: 10, 15, 20, 25, 30, 35, and 40.
24. The expression cassette of any one of the previous claims, wherein the downstream exon comprises a sequence selected from the group consisting of SEQ ID NO: 12, 17, 22, 27, 32, 37, and 42.
25. The expression cassette of any one of the previous claims, comprising a sequence selected from the group consisting of SEQ ID NO: 8-42.
26. The expression cassette of any one of the previous claims, comprising a sequence selected from the group consisting of SEQ ID NO: 1-8 or 43-66.
27. The expression cassette of any one of the previous claims, further comprising an untranslated region (UTR), a microRNA binding site, a polyA (pA) sequence, an intron sequence, or any combination thereof.
28. The expression cassette any one of the previous claims, wherein the expression cassette comprises a poly(A) (pA) sequence.
29. The expression cassette of claim 27 or claim 28, wherein the pA sequence is a synthetic pA sequence, a bovine growth hormone (bGH) pA sequence, or a human growth hormone (hGH) pA sequence.
30. A delivery vector comprising the expression cassette of any one of claims 1-26.
31. The delivery vector of claim 30, wherein the expression cassette is flanked by inverted terminal repeat (ITR) sequences.
32. The delivery vector of claim 31, wherein the delivery vector is selected from the group consisting of a viral vector, a plasmid, a lipid, a protein particle, a bacterial vector, a lysosome, a virus-like particle, a polymeric particle, an exosome, and a vault particle.

33. The delivery vector of claim 31 or claim 32, wherein the delivery vector is a viral vector.
34. The delivery vector of claim 33, wherein the viral vector is a lentiviral vector, an adenoviral vector, an adeno-associated viral (AAV) vector, or a retroviral vector.
35. The delivery vector of any one of claims 32-34, wherein the viral vector is an adeno-associated viral (AAV) vector.
36. The delivery vector of any one of claims 31-35, which is a recombinant AAV (rAAV) vector comprising an AAV serotype selected from the group consisting of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAVrh8, AAV9, AAV10, AAVrh10, AAV11, and AAV12.
37. A cell comprising the expression cassette of any one of claims 1-30, or the delivery vector of any one of claims 31-36.
38. A pharmaceutical composition comprising the expression cassette of any one of claims 1-30, the delivery vector of any one of claims 31-36, or the cell of claim 37.
39. The pharmaceutical composition of claim 38, further comprising a pharmaceutically acceptable excipient.
40. A kit comprising i) the expression cassette of any one of claims 1-30, the delivery vector of any one of claims 31-36, or the pharmaceutical composition of claim 38 or claim 39; and (ii) optionally instructions for use.
41. The kit of claim 40, further comprising (iii) a splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.
42. The kit of claim 41, wherein the splicing modulator molecule is kinetin.
43. The kit of claim 42, wherein the splicing modulator molecule is homocarbonyltopsentin.
44. A method of regulating expression of a protein in a mammalian cell comprising administering to the cell the expression cassette of any one of claims 1-30, the delivery vector of any one of claims 31-36, or the pharmaceutical composition of claim 38 or claim 39, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.
45. A method of expressing a protein in a subject comprising administering to the subject the expression cassette of any one of claims 1-30, the delivery vector of any one of claims 31-36, the cell of claim 37, or the pharmaceutical composition of claim 38 or claim 39, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

46. A method of treating a disease in a mammal comprising administering to a subject the expression cassette of any one of claims 1-30, the delivery vector of any one of claims 31-36, the cell of claim 37, or the pharmaceutical composition of claim 38 or claim 39, followed by administering the splicing modulator molecule that promotes the inclusion of the alternatively spliced exon.

47. The methods of any one of claims 44-46, wherein the administration of the expression cassette, delivery vector, or cell is systemic or tissue specific.

48. The methods of any one of claims 44-47, wherein the administration of the molecule that promotes the inclusion of the alternatively spliced exon is systemic.

49. The method of any one of claims 44-48, wherein the molecule that promotes the inclusion of the alternatively spliced exon is administered at least 1 day after administration of the expression cassette.

50. The method of any one of claims 44-49, further comprising administering a second dose of the molecule that promotes the inclusion of the alternatively spliced exon.

51. A method of regulating transgene expression post-transcriptionally, comprising administering to a cell or subject (a) a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) a heterologous nucleic acid sequence; wherein the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene; and (b) a splicing modulator, wherein the splicing modulator is Kinetin or Homocarbonyltopsentin.

52. A method of producing an adeno-associated viral (AAV) vector comprising:
(a) introducing a first plasmid comprising an adenovirus helper gene and a second plasmid comprising a heterologous nucleic acid sequence into a host cell, wherein the host cell comprises an AAV capsid (Cap) gene and a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV replication (Rep) gene; wherein the upstream intron, the alternatively spliced

exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene; and

(b) culturing the host cell in the presence of a splicing modulator to produce AAV particles.

53. A method of producing an adeno-associated viral (AAV) vector comprising:

(a) introducing a first plasmid comprising an adenovirus helper gene and a second plasmid comprising a heterologous nucleic acid sequence into a host cell, wherein the host cell comprises an AAV replication (Rep) gene and a nucleic acid construct comprising (i) an upstream exon, (ii) an upstream intron, (iii) an alternatively spliced exon comprising a translation initiation sequence, (iv) a downstream intron, (v) a downstream exon, and (vi) an AAV capsid (Cap) gene; wherein the upstream intron, the alternatively spliced exon, and the downstream intron each comprise a portion of a gene selected from the group consisting of a Transient Receptor Potential Cation Channel Subfamily V Member 3 (TRPV3) gene, a Serine/Threonine Kinase 31 (STK31) gene, an Ubiquitin Specific Peptidase 25 (USP25) gene, a IK gene, a Testis Expressed 14, Intercellular Bridge Forming Factor (TEX14) gene, a General Transcription Factor IIB (GTF2B) gene, and a LY6/PLAUR Domain Containing 3 (LYPD3) gene; and

(b) culturing the host cell in the presence of a splicing modulator to produce AAV particles.

54. The method of claim 52 or 53, further comprising (c) harvesting the AAV particles from the cultured host cell.

55. The method of claim 54, further comprising (d) purifying the harvested AAV particles.

56. The method of claim 52 or 53, wherein the host cell is a HEK293 cell.

57. The method of any one of claims 52 to 56, wherein the splicing modulator is Kinetin or Homocarbonyltopsentin.

58. The method of any one of claims 52 to 57, wherein the adenovirus helper gene is selected from the group consisting of an Early 4 (E4) gene, an Early 2A (E2A) gene, and a viral associated (VA) gene.

59. The method of any one of claims 52 to 58, wherein the nucleic acid construct is stably integrated into the host cell genome.

60. The method of any one of claims 52 to 58, wherein the nucleic acid construct comprises a sequence selected from the group consisting of SEQ ID NO: 59-62.

61. A method of producing an AAV particle comprising contacting a host cell comprising rep genes, cap genes, and adenoviral helper genes with a topsentin alkaloid or derivative thereof.

62. The method of claim 61, wherein topsentin alkaloid or derivative thereof is selected from the group consisting of Homocarbonyltopsentin, Topsentin, Bromotopsentin, Bromodeoxytopsentin, Deoxytopsentin, and any combination thereof.

63. The method of claim 62, wherein the method: i) reduces one or more viral response genes, ii) increases one or more immediate early genes, iii) increases a NF-κB signaling pathway, iv) increases a TNF signaling pathway, v) increases a IL-17 signaling pathway, vi) inhibits apoptosis, vii) inhibits viral defense and/or vii) increases histone levels.

64. The method of any one of claims 61-63, wherein the apoptosis is inhibited by altering a signaling pathway selected from the group consisting of a NF-κB signaling pathway, a TNF signaling pathway, a IL-17 signaling pathway, a JAK/STAT signaling pathway, a P38 MAPK signaling pathway, a PI3K-Akt signaling pathway, or any combination thereof.

65. The method of any one of claims 61-65, wherein the increased histone levels result in increased metabolic activity.

66. The method of any one of claims 61-65, wherein the increased histone levels result in increased host cell replication.

67. The method of any one of claims 61-66, wherein the host cell comprises an AAV capsid (Cap) gene.

68. The method of any one of claims 61-67, wherein the host cell comprises an AAV replication (Rep) gene.

69. The method of any one of claims 61-68, wherein the host cell is a HEK293 cell.

70. The method of any one of claims 61 to 69, wherein the topsentin alkaloid or derivative thereof is Homocarbonyltopsentin.

71. The method of claim 70, wherein the host cell is contacted with about 5 μM to about 40 μM of Homocarbonyltopsentin.

72. The method of any one of claims 61 to 71, wherein the host cell is contacted with the topsentin alkaloid or derivative thereof for at least about 24 hours, 48 hours, or 72 hours.

73. The method of any one of claims 61 to 72, further comprising harvesting the AAV particle.

74. The method of claim 73, further comprising purifying the harvested AAV particle.

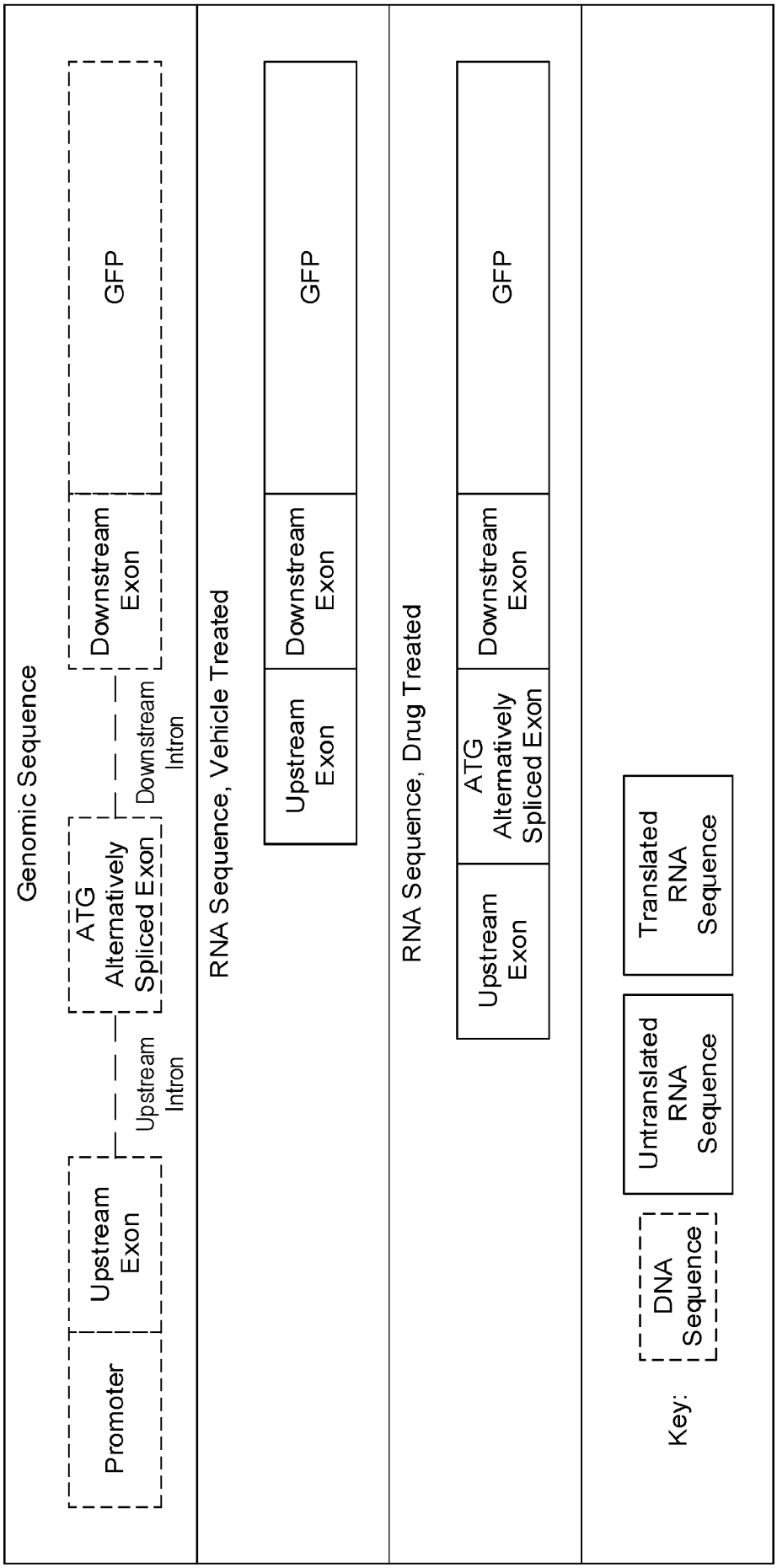

Genomic Sequence
Promoter
Upstream Exon
Upstream Intron
ATG Alternatively Spliced Exon
Downstream Intron
Downstream Exon
GFP
RNA Sequence, Vehicle Treated
Upstream Exon
Downstream Exon
GFP
RNA Sequence, Drug Treated
Upstream Exon
ATG Alternatively Spliced Exon
Downstream Exon
GFP
Key:
DNA Sequence
Untranslated RNA Sequence
Translated RNA Sequence

FIG. 1A

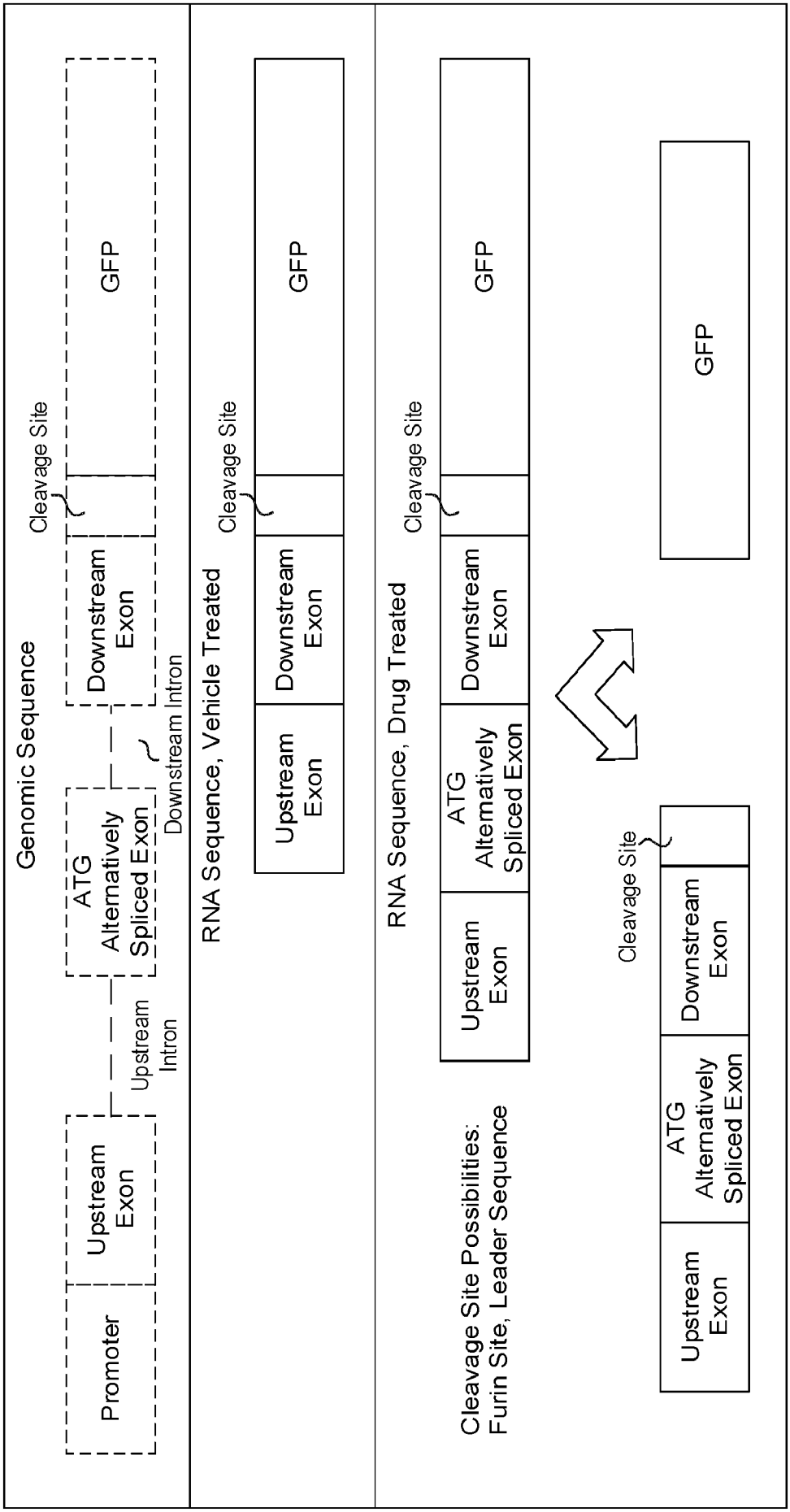

Genomic Sequence
Promoter
Upstream Exon
Upstream Intron
ATG Alternatively Spliced Exon
Downstream Intron
Cleavage Site
Downstream Exon
GFP
RNA Sequence, Vehicle Treated
Upstream Exon
Downstream Exon
Cleavage Site
GFP
RNA Sequence, Drug Treated
Upstream Exon
ATG Alternatively Spliced Exon
Downstream Exon
Cleavage Site
GFP
Cleavage Site Possibilities: Furin Site, Leader Sequence
Upstream Exon
ATG Alternatively Spliced Exon
Downstream Exon
Cleavage Site
GFP

FIG. 1B

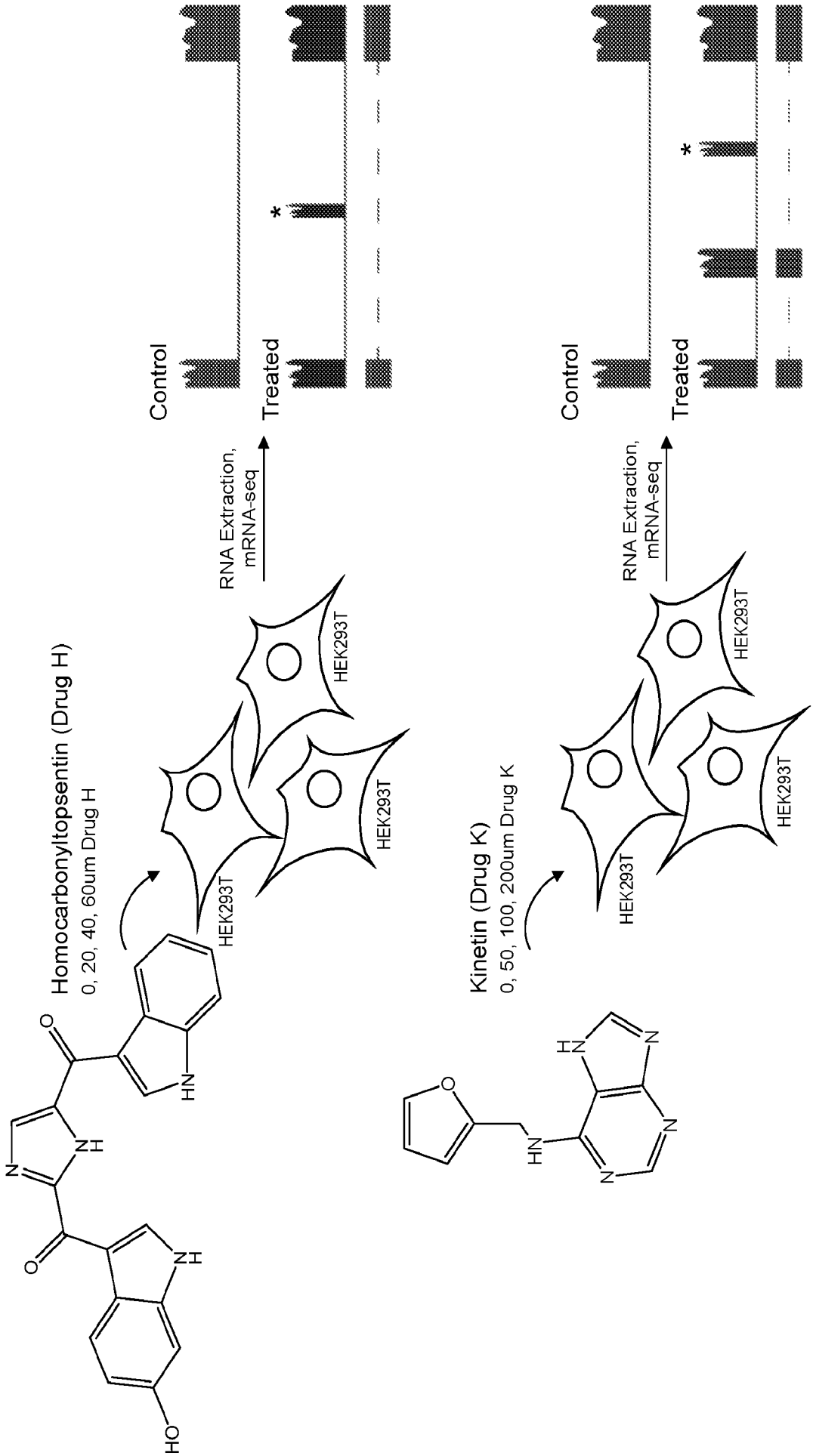

Homocarbonyltopsentin (Drug H)
0, 20, 40, 60um Drug H
HEK293T
HEK293T
HEK293T
RNA Extraction, mRNA-seq
Control
Treated
*
Kinetin (Drug K)
0, 50, 100, 200um Drug K
HEK293T
HEK293T
HEK293T
RNA Extraction, mRNA-seq
Control
Treated
*

FIG. 2

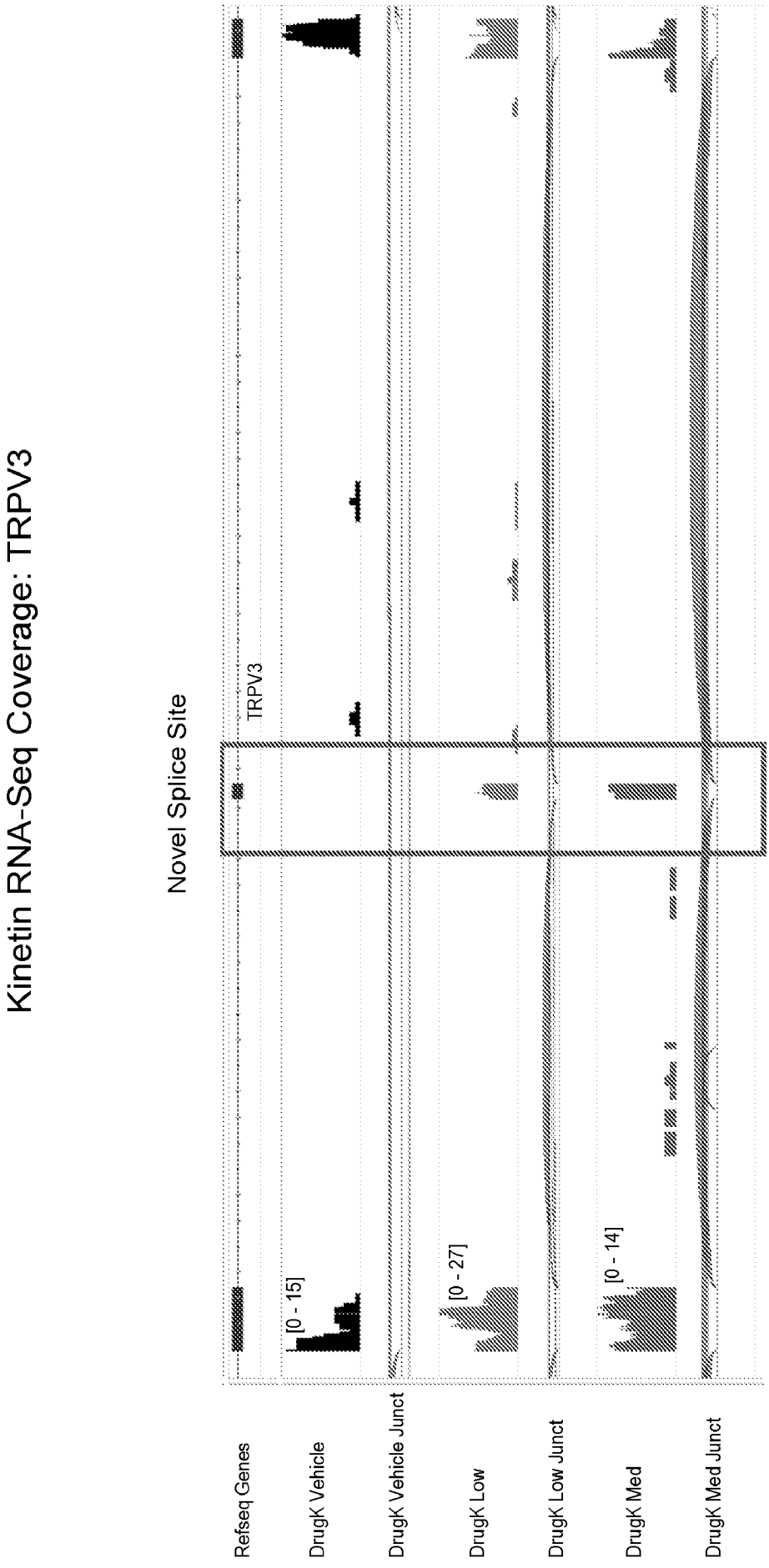

Kinetin RNA-Seq Coverage: TRPV3
Novel Splice Site
TRPV3
[0 - 15]
[0 - 27]
[0 - 14]
Refseq Genes
DrugK Vehicle
DrugK Vehicle Junct
DrugK Low
DrugK Low Junct
DrugK Med
DrugK Med Junct

FIG. 3A

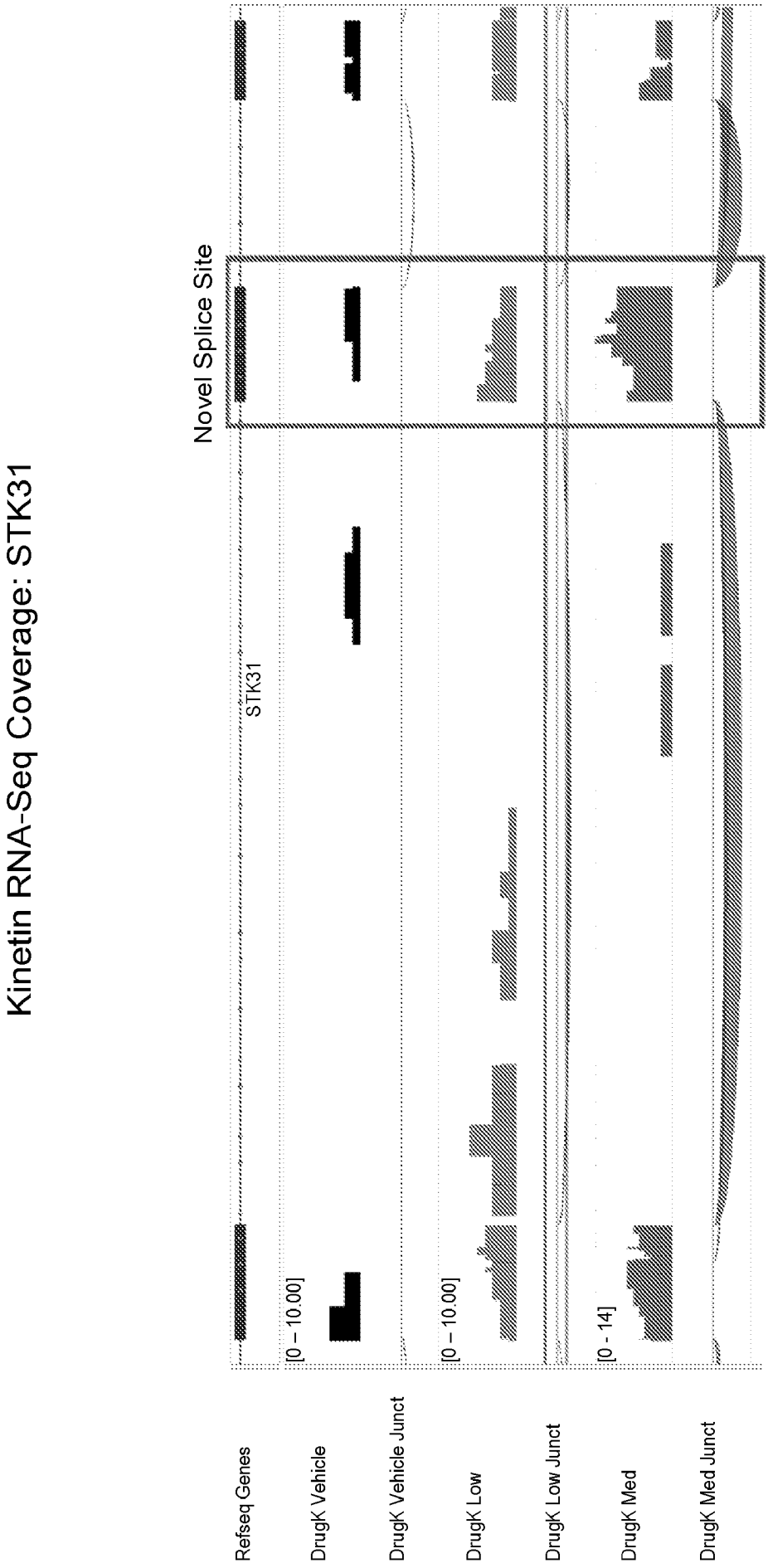

Kinetin RNA-Seq Coverage: STK31
Novel Splice Site
STK31
Refseq Genes
DrugK Vehicle
[0 – 10.00]
DrugK Vehicle Junct
DrugK Low
[0 – 10.00]
DrugK Low Junct
DrugK Med
[0 - 14]
DrugK Med Junct

FIG. 3B

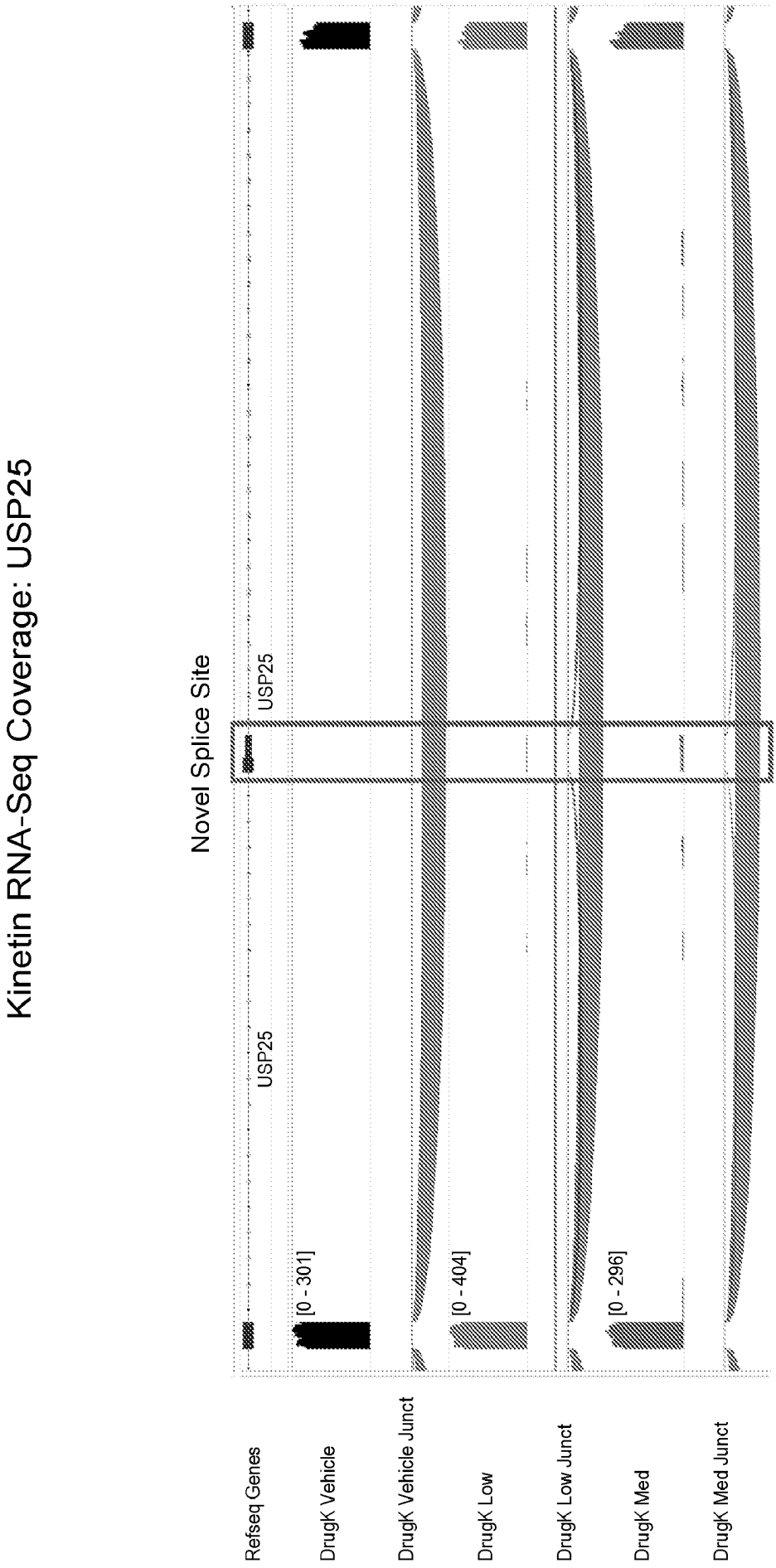

Kinetin RNA-Seq Coverage: USP25
Novel Splice Site
USP25
USP25
[0 - 301]
[0 - 404]
[0 - 296]
Refseq Genes
DrugK Vehicle
DrugK Vehicle Junct
DrugK Low
DrugK Low Junct
DrugK Med
DrugK Med Junct

FIG. 3C

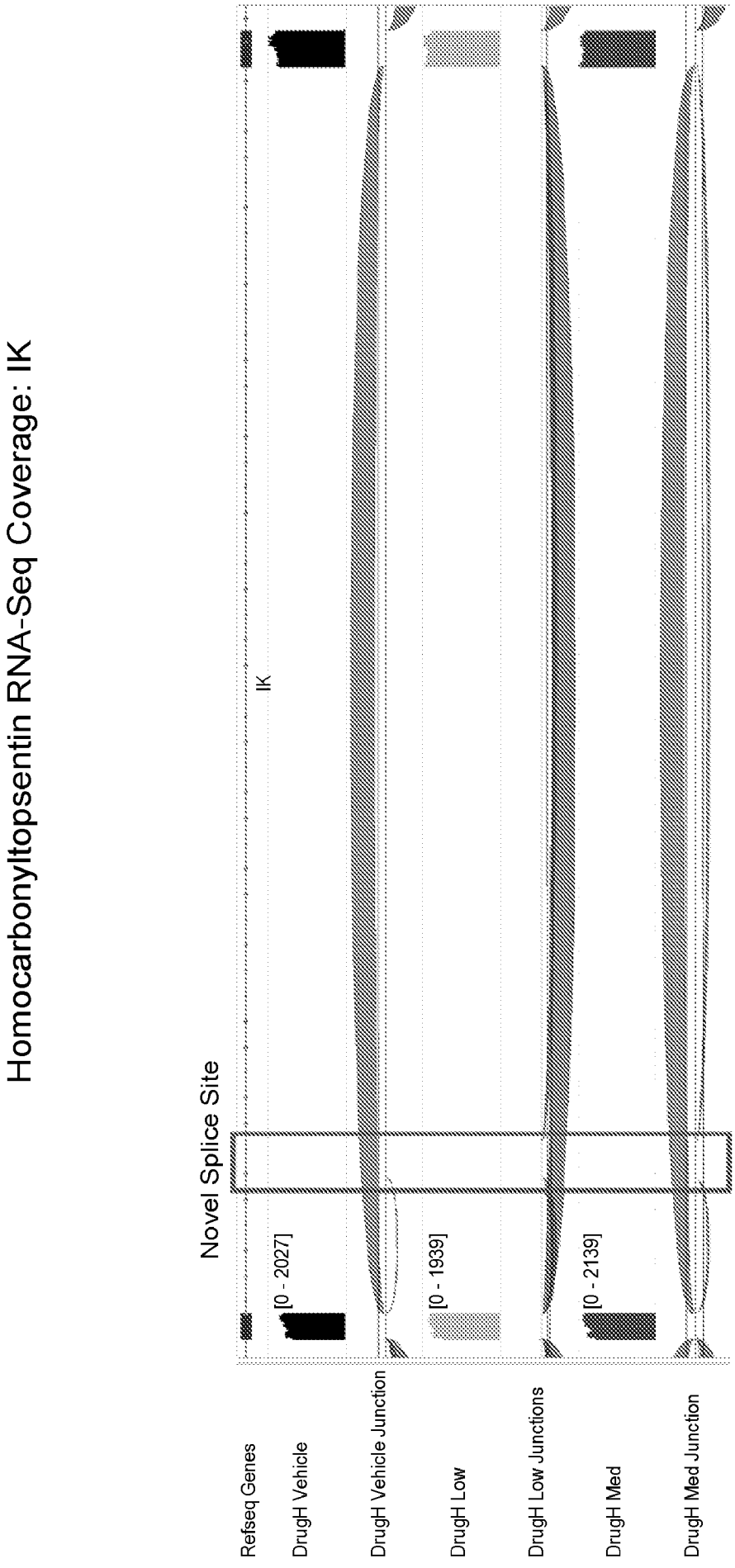

Homocarbonyltopsentin RNA-Seq Coverage: IK
Novel Splice Site
IK
[0 - 2027]
[0 - 1939]
[0 - 2139]
Refseq Genes
DrugH Vehicle
DrugH Vehicle Junction
DrugH Low
DrugH Low Junctions
DrugH Med
DrugH Med Junction

FIG. 4A

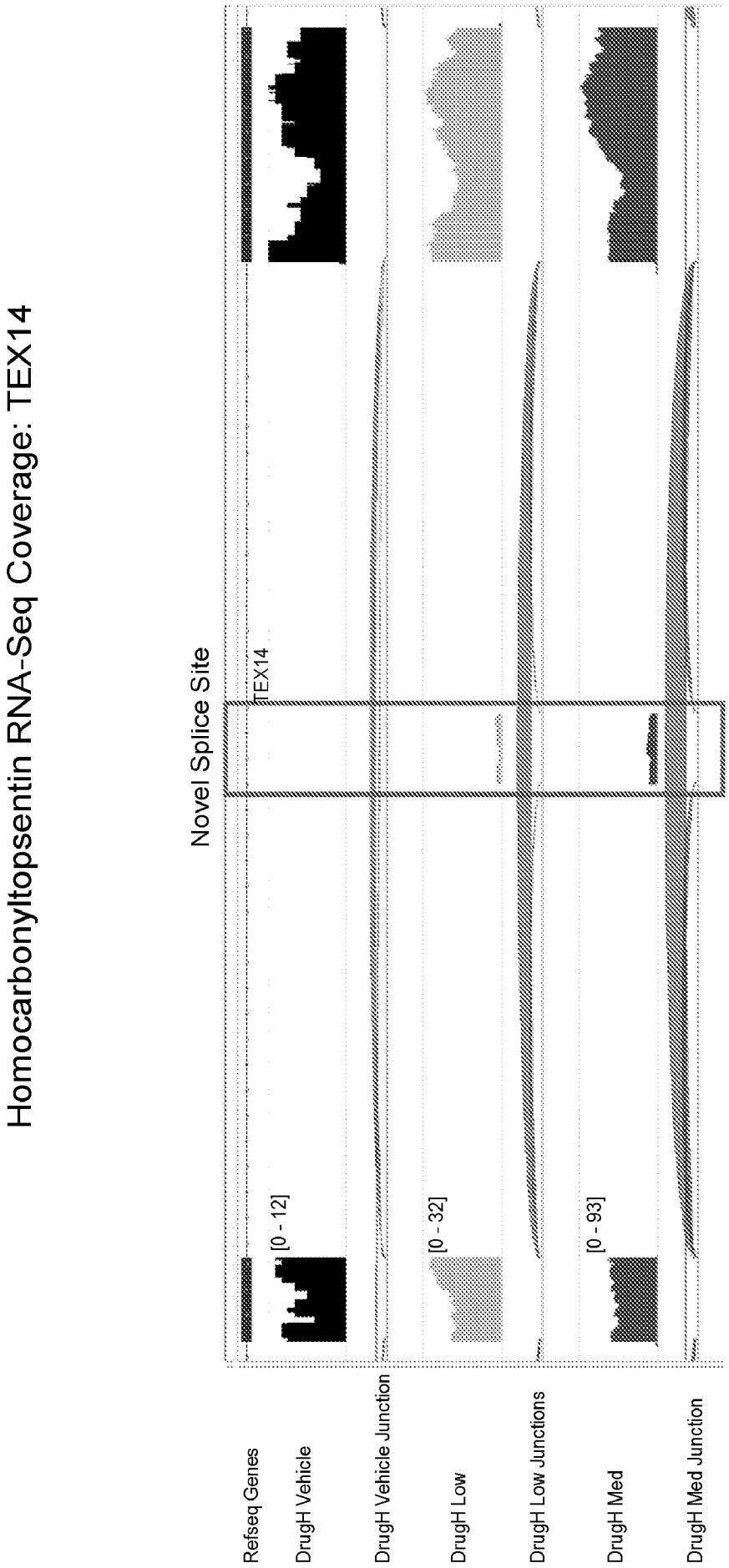

Homocarbonyltopsentin RNA-Seq Coverage: TEX14
Novel Splice Site
TEX14
[0 - 12]
[0 - 32]
[0 - 93]
Refseq Genes
DrugH Vehicle
DrugH Vehicle Junction
DrugH Low
DrugH Low Junctions
DrugH Med
DrugH Med Junction

FIG. 4B

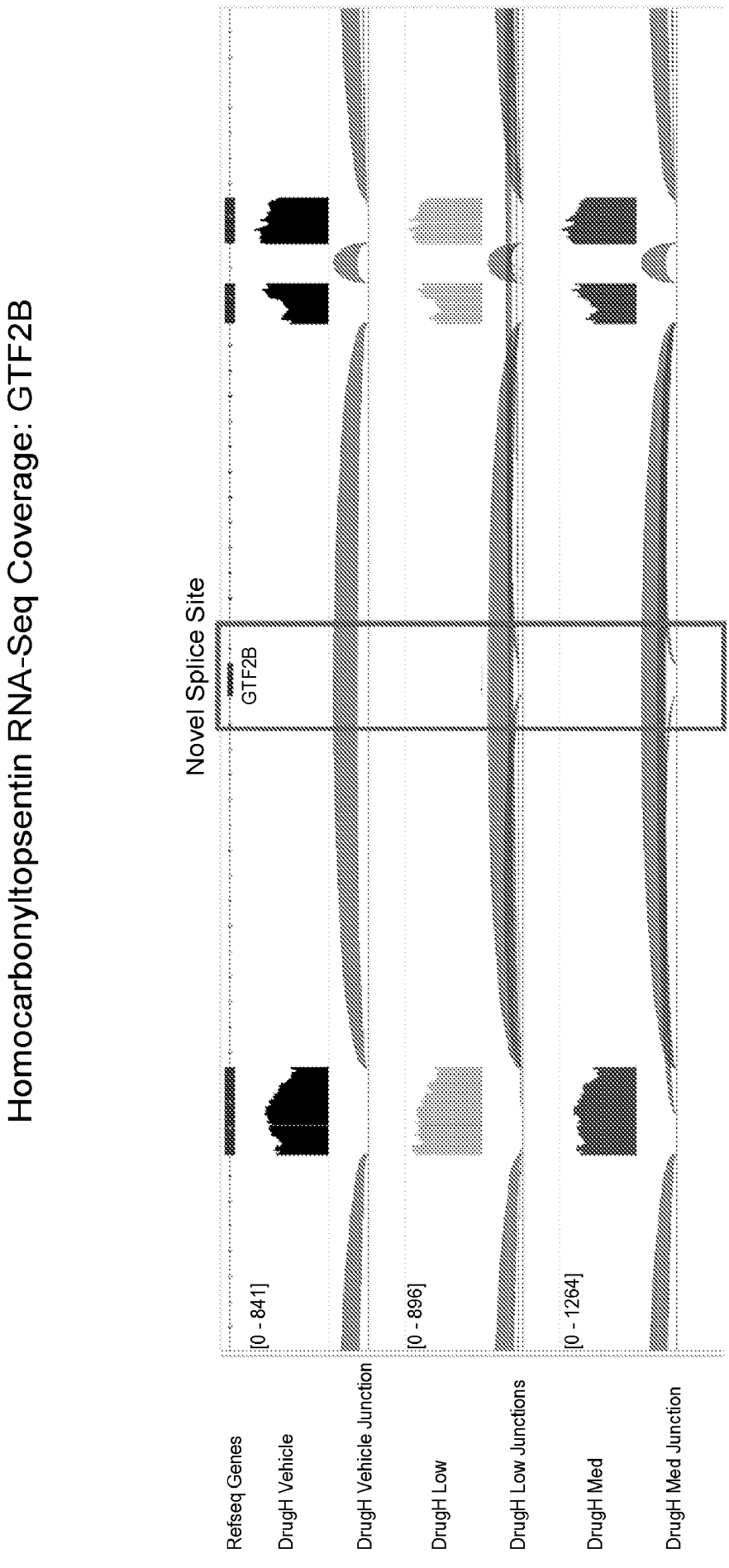

Homocarbonyltopsentin RNA-Seq Coverage: GTF2B
Novel Splice Site
GTF2B
[0 - 841]
[0 - 896]
[0 - 1264]
Refseq Genes
DrugH Vehicle
DrugH Vehicle Junction
DrugH Low
DrugH Low Junctions
DrugH Med
DrugH Med Junction

FIG. 4C

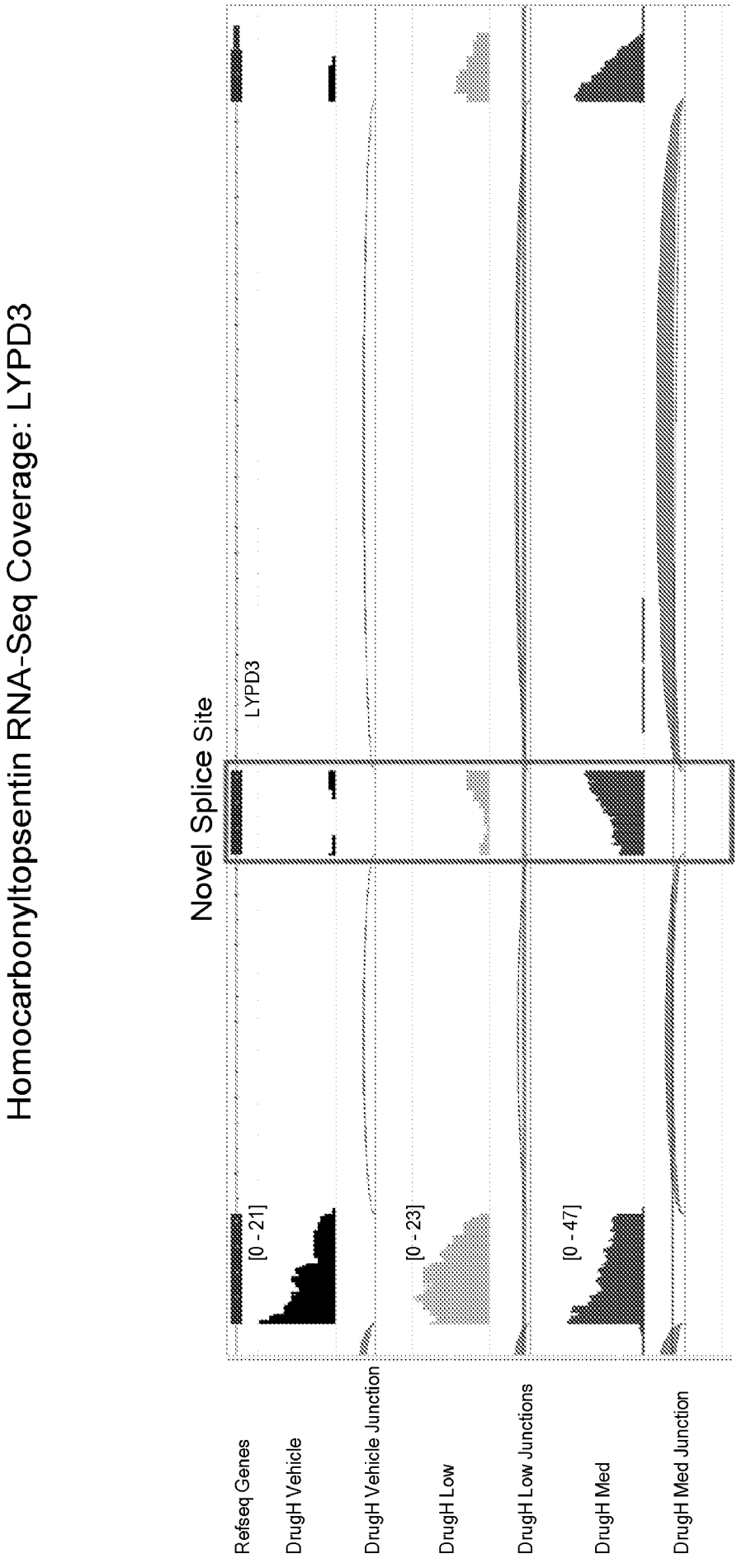

Homocarbonyltopsentin RNA-Seq Coverage: LYPD3
Novel Splice Site
LYPD3
[0 - 21]
[0 - 23]
[0 - 47]
Refseq Genes
DrugH Vehicle
DrugH Vehicle Junction
DrugH Low
DrugH Low Junctions
DrugH Med
DrugH Med Junction

FIG. 4D

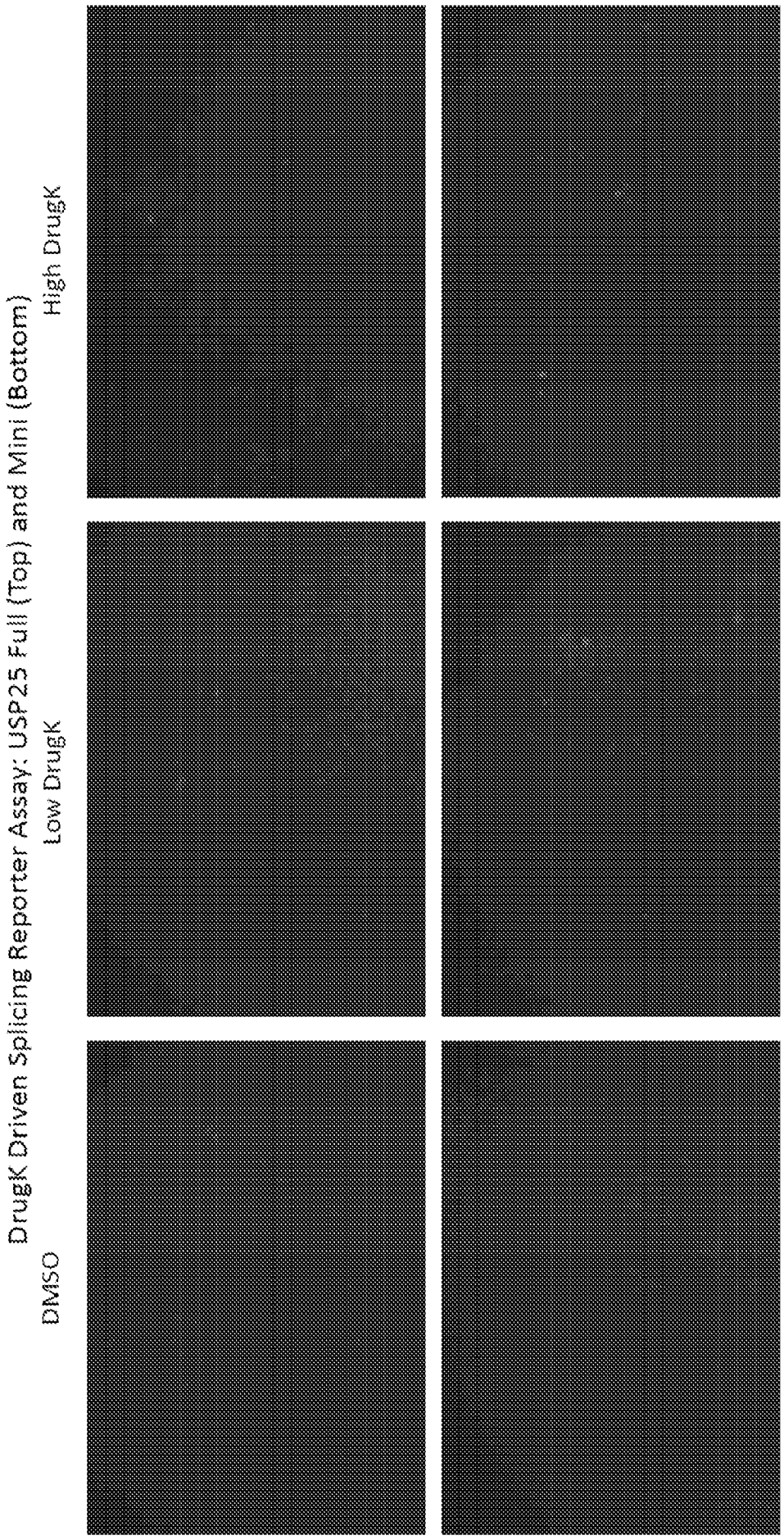

DrugK Driven Splicing Reporter Assay: USP25 Full (Top) and Mini (Bottom)
DMSO
Low DrugK
High DrugK

FIG. 5A

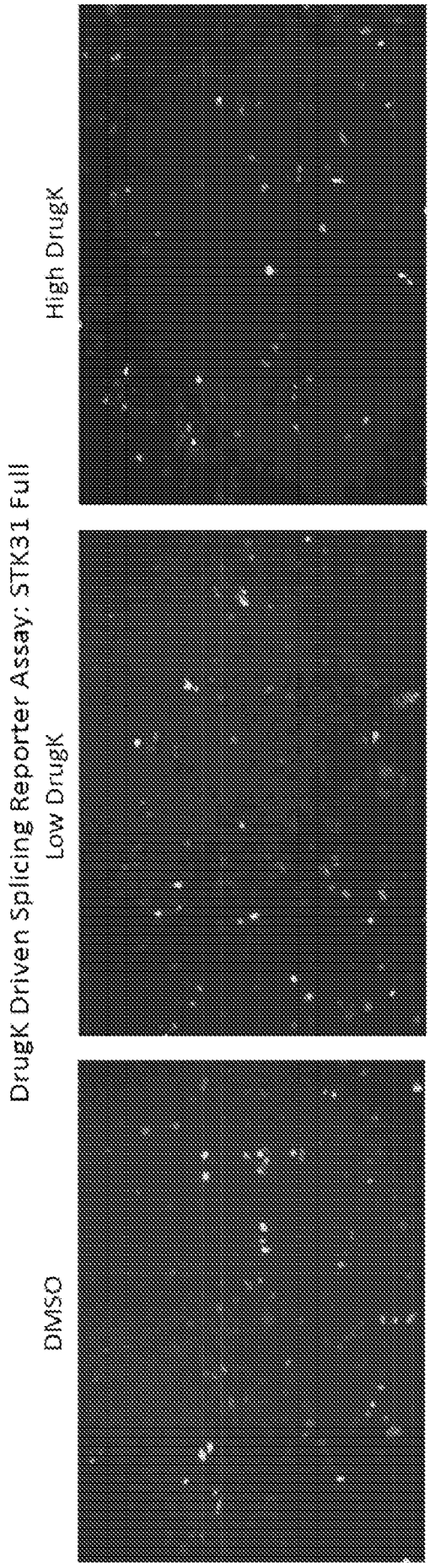

DrugK Driven Splicing Reporter Assay: STK31 Full
DMSO
Low DrugK
High DrugK

FIG. 5B

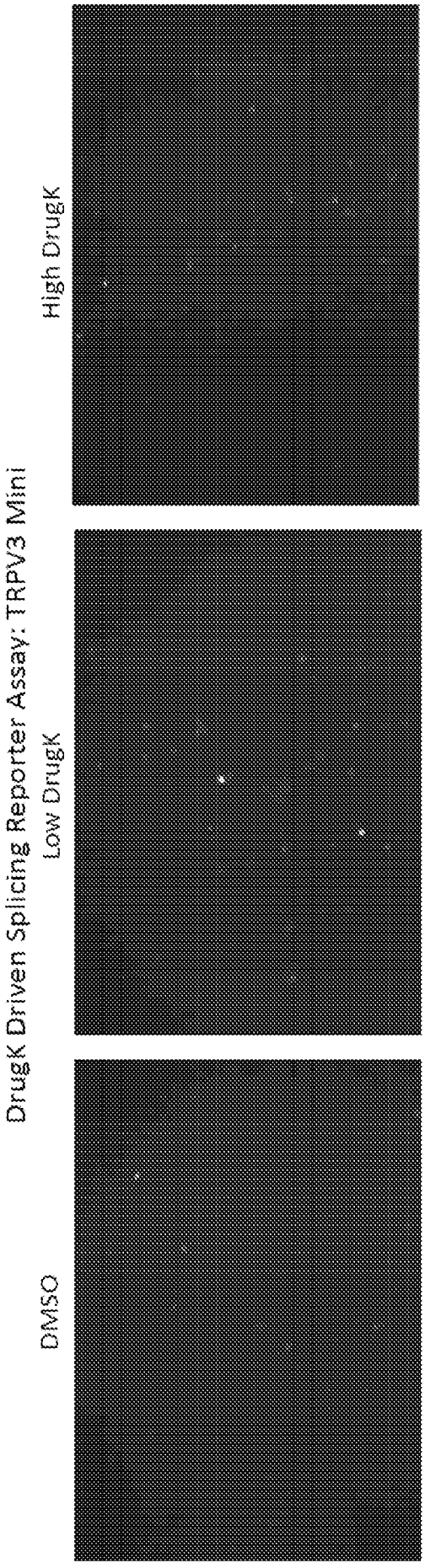

DrugK Driven Splicing Reporter Assay: TRPV3 Mini
DMSO
Low DrugK
High DrugK

FIG. 5C

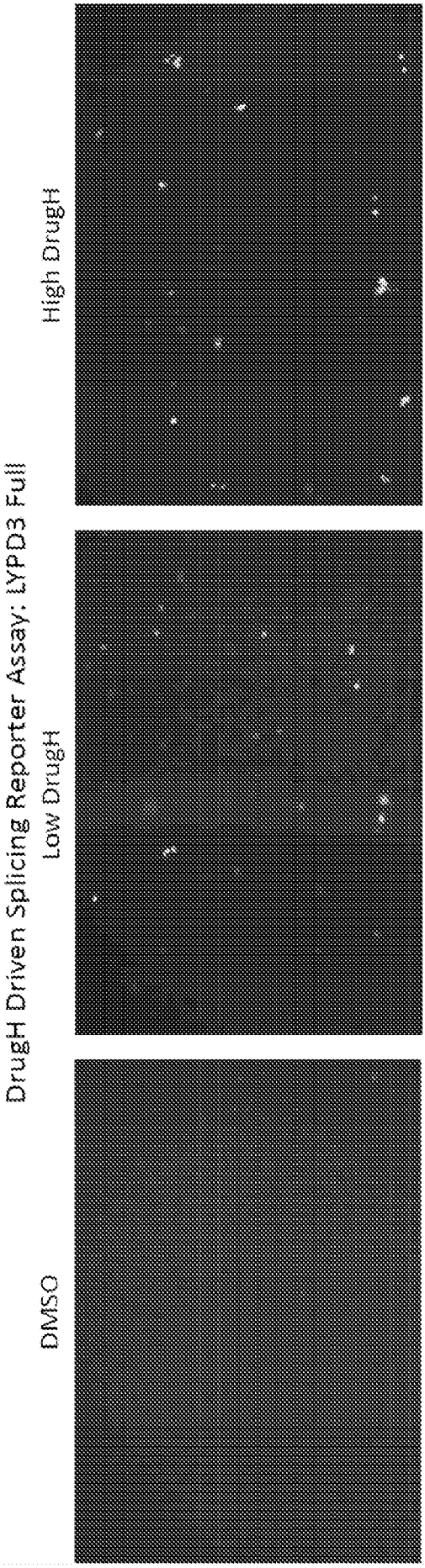

DrugH Driven Splicing Reporter Assay: LYPD3 Full
DMSO
Low DrugH
High DrugH

FIG. 6A

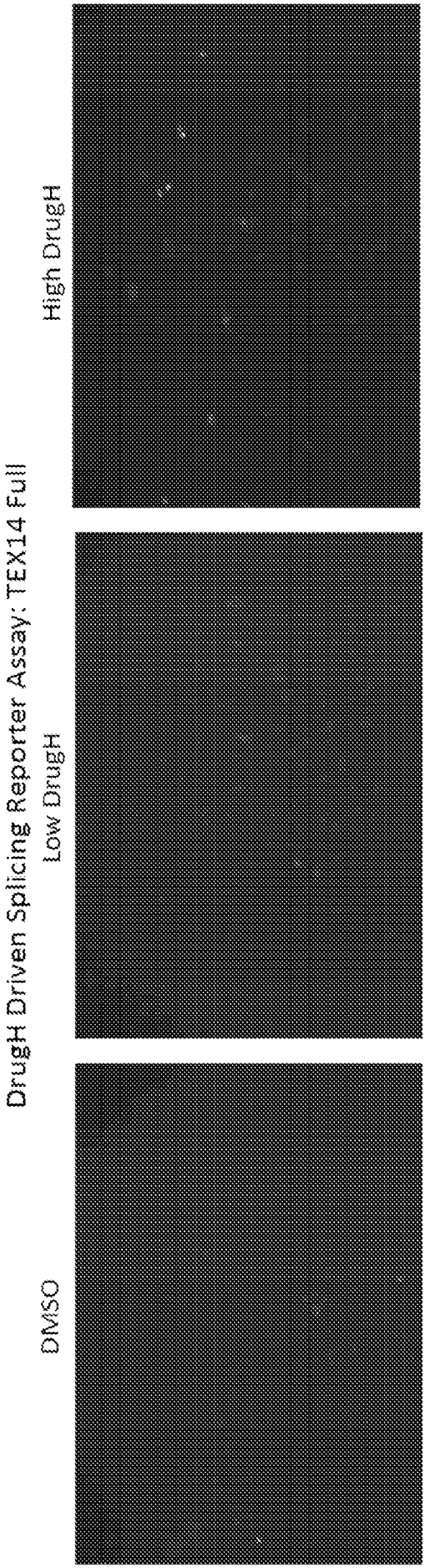

DrugH Driven Splicing Reporter Assay: TEX14 Full
DMSO
Low DrugH
High DrugH

FIG. 6B

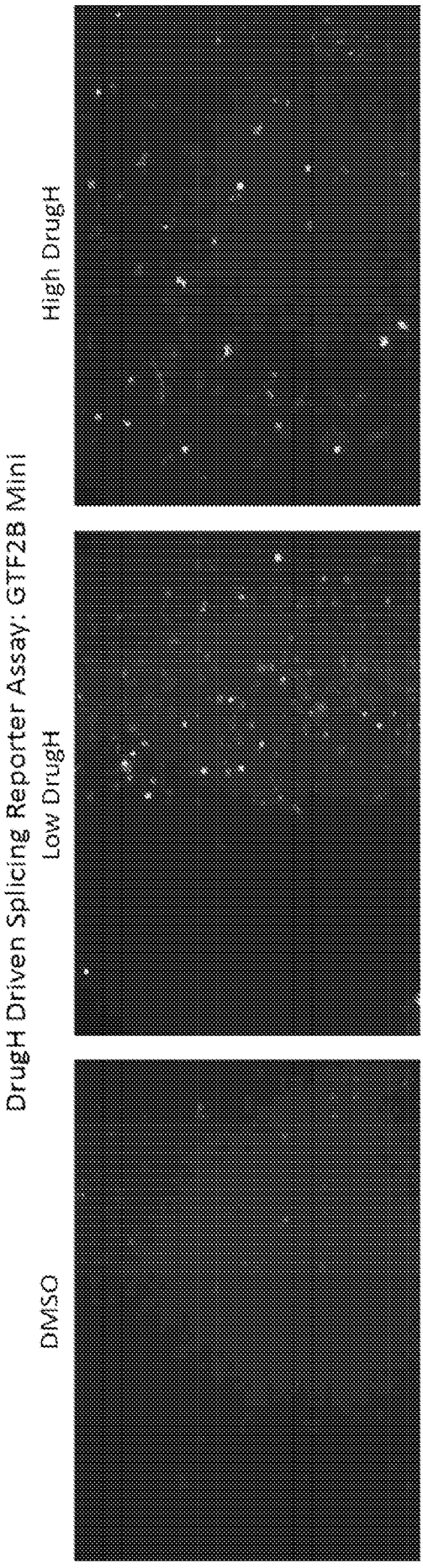

DrugH Driven Splicing Reporter Assay: GTF2B Mini
DMSO
Low DrugH
High DrugH

FIG. 6C

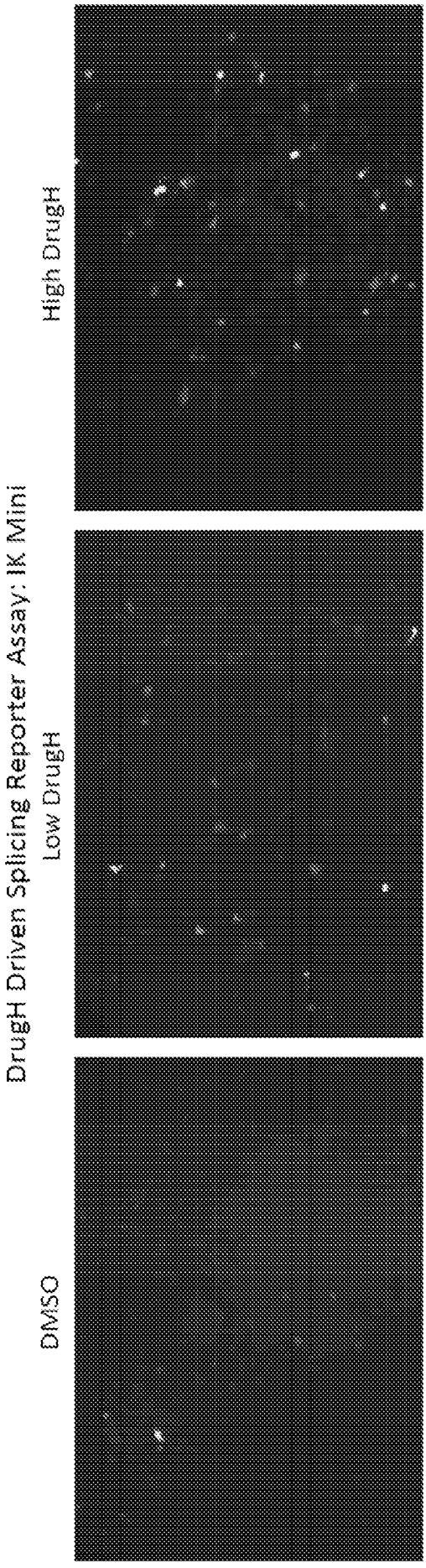

DrugH Driven Splicing Reporter Assay: IK Mini
DMSO
Low DrugH
High DrugH

FIG. 6D

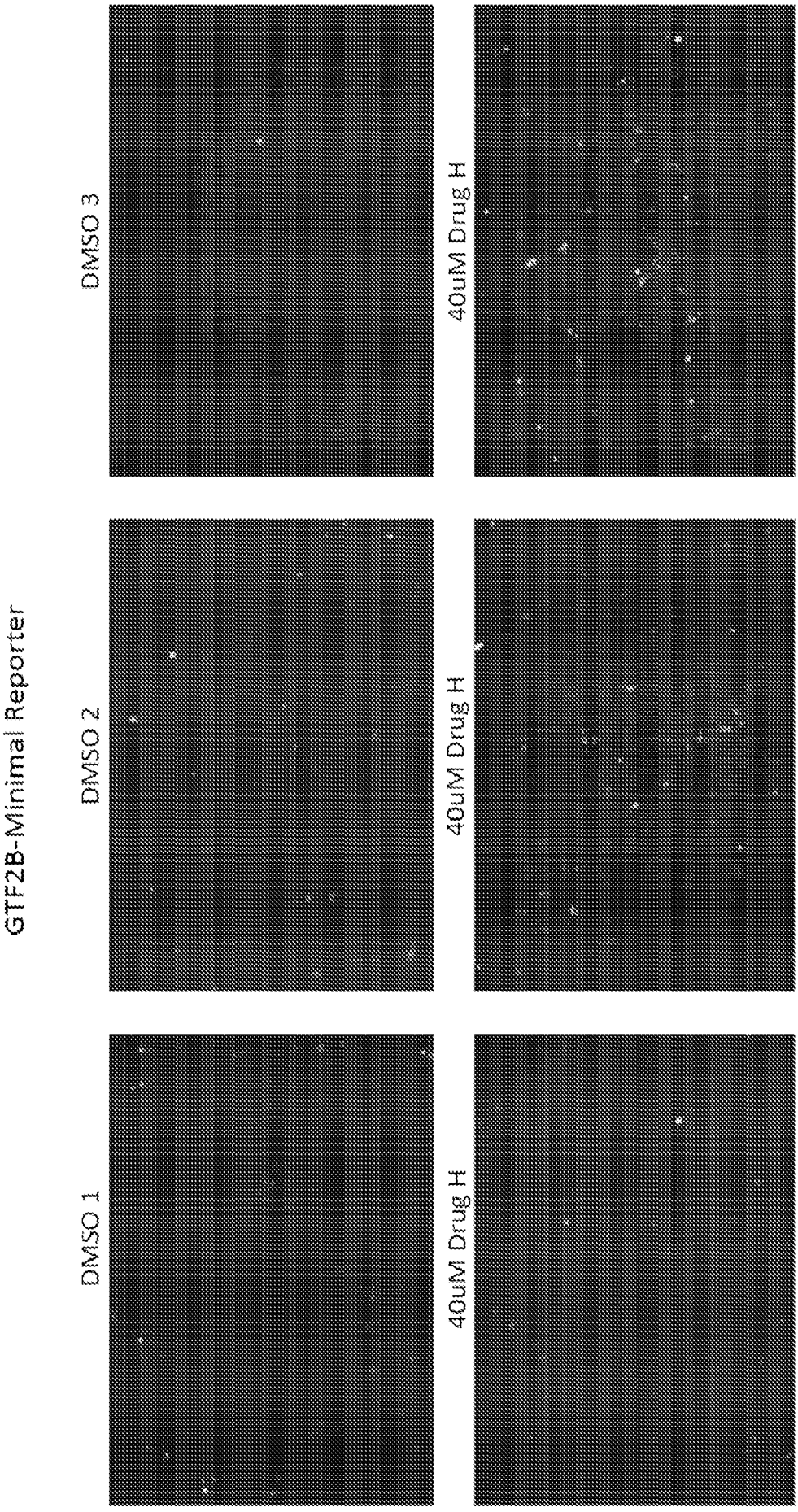

GTF2B-Minimal Reporter
DMSO 1
40uM Drug H
DMSO 2
40uM Drug H
DMSO 3
40uM Drug H

FIG. 7A

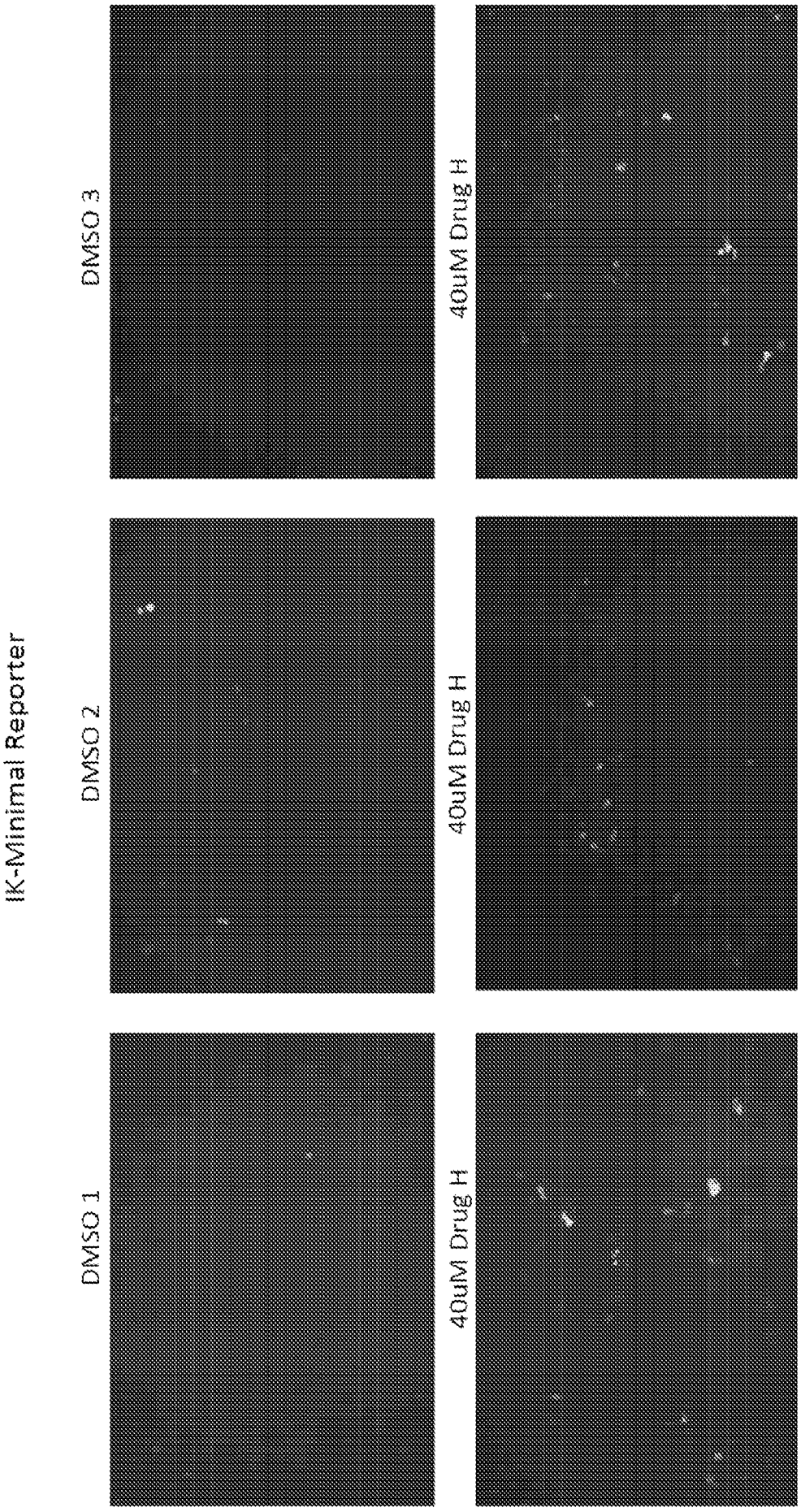

IK-Minimal Reporter
DMSO 1
40uM Drug H
DMSO 2
40uM Drug H
DMSO 3
40uM Drug H

FIG. 7B

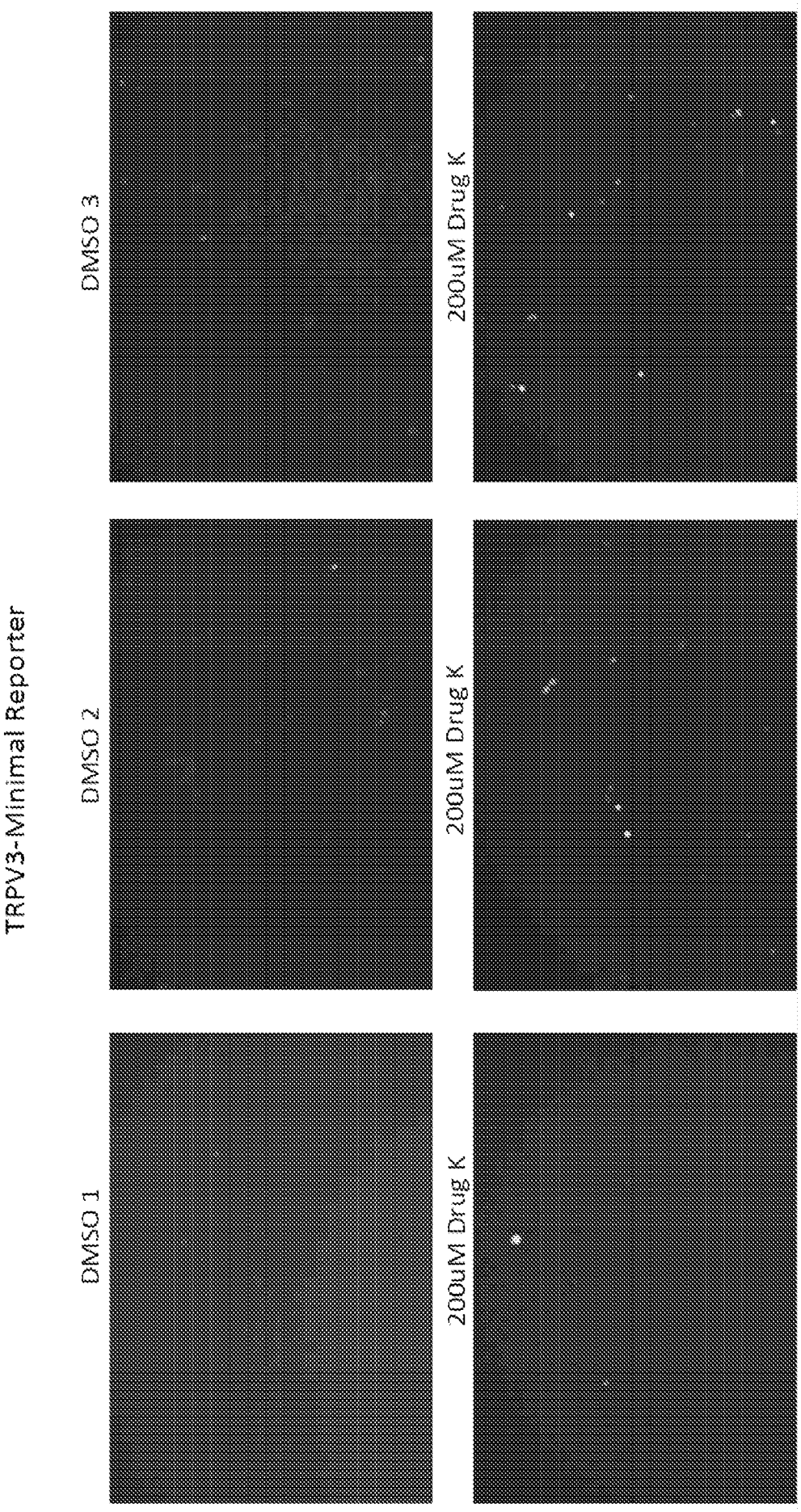

TRPV3-Minimal Reporter
DMSO 1
200uM Drug K
DMSO 2
200uM Drug K
DMSO 3
200uM Drug K

FIG. 7C

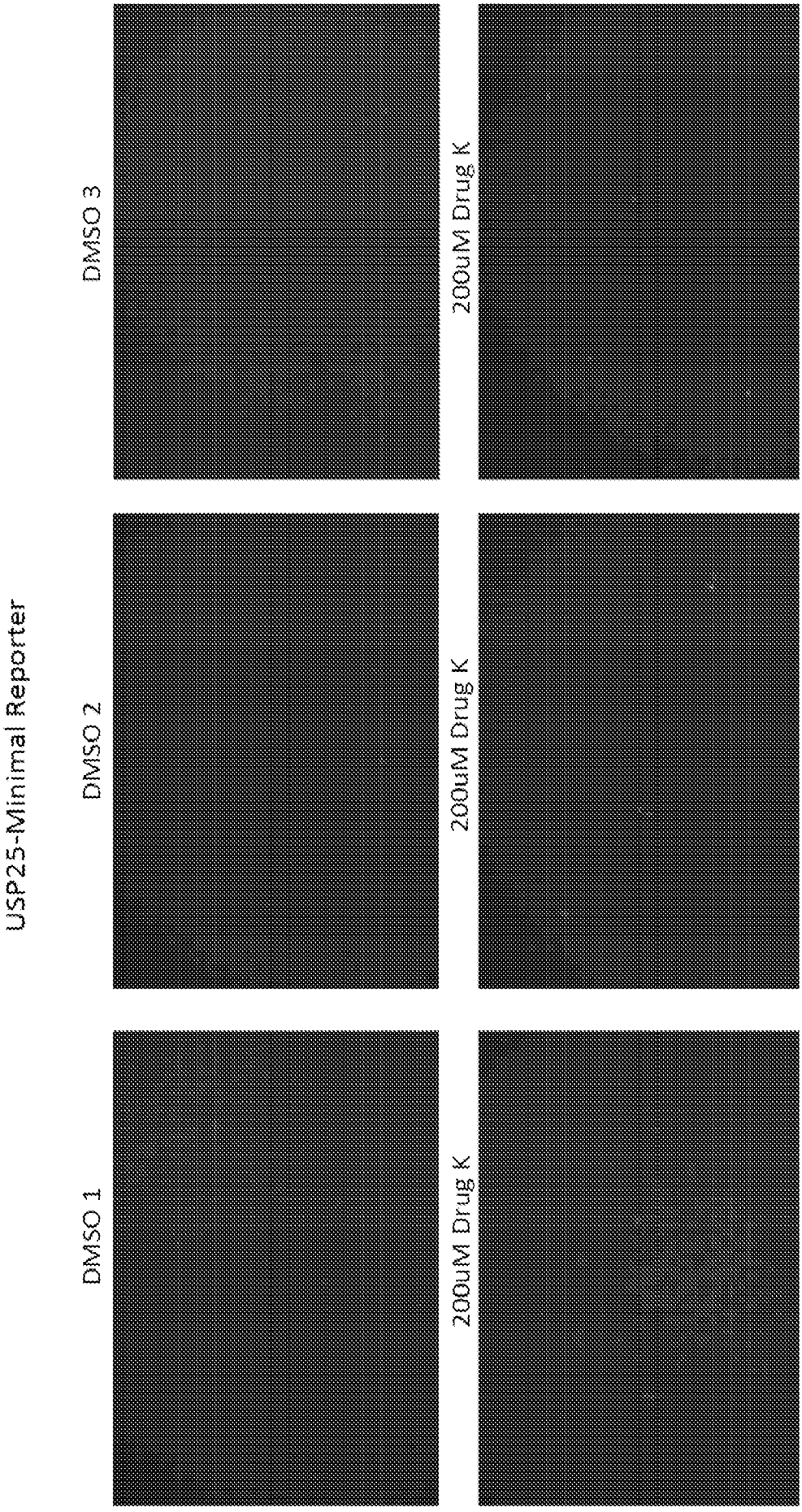

USP25-Minimal Reporter
DMSO 1
200uM Drug K
DMSO 2
200uM Drug K
DMSO 3
200uM Drug K

FIG. 7D

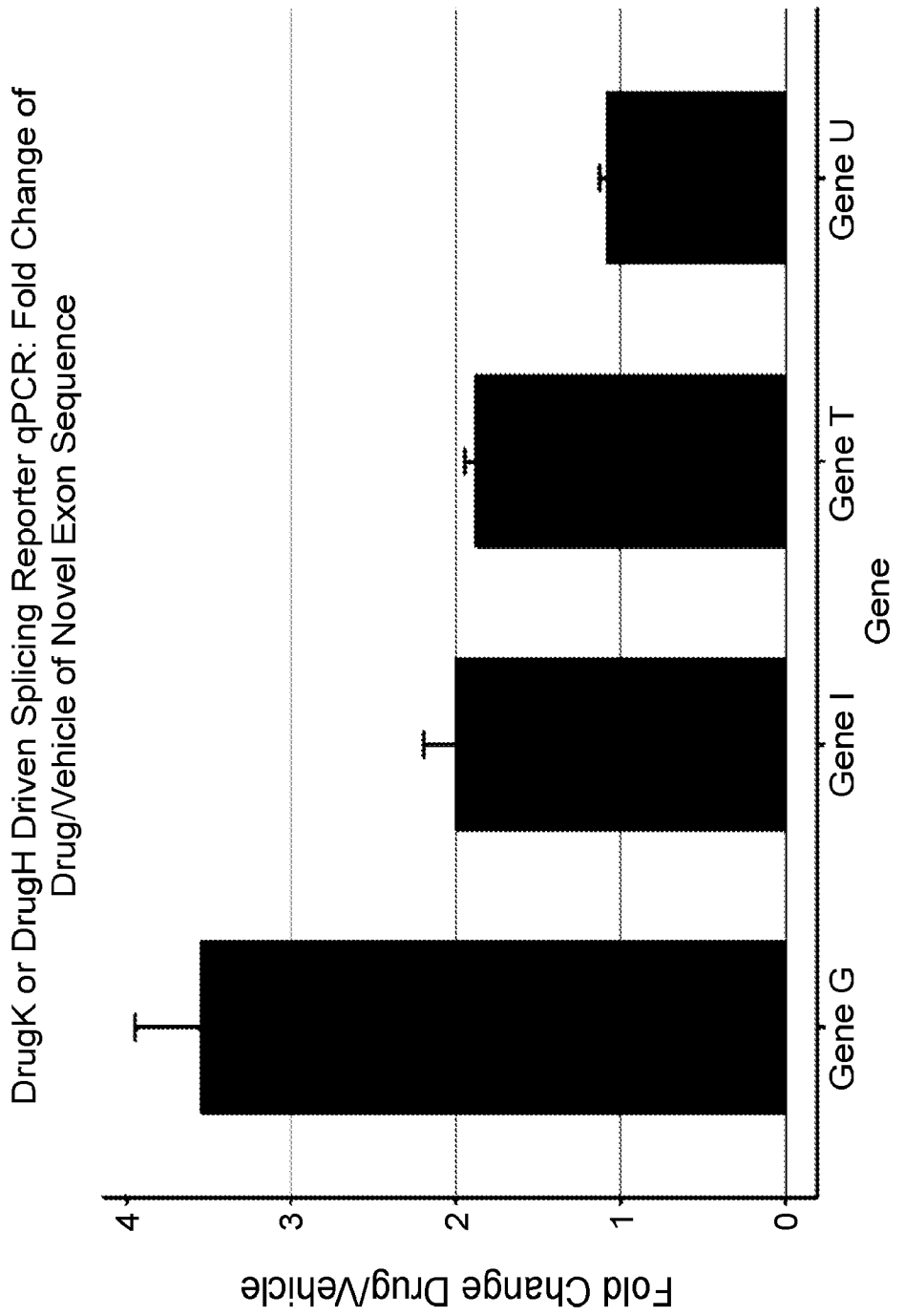

DrugK or DrugH Driven Splicing Reporter qPCR: Fold Change of Drug/Vehicle of Novel Exon Sequence
Fold Change Drug/Vehicle
0
1
2
3
4
Gene G
Gene I
Gene T
Gene U
Gene

FIG. 7E

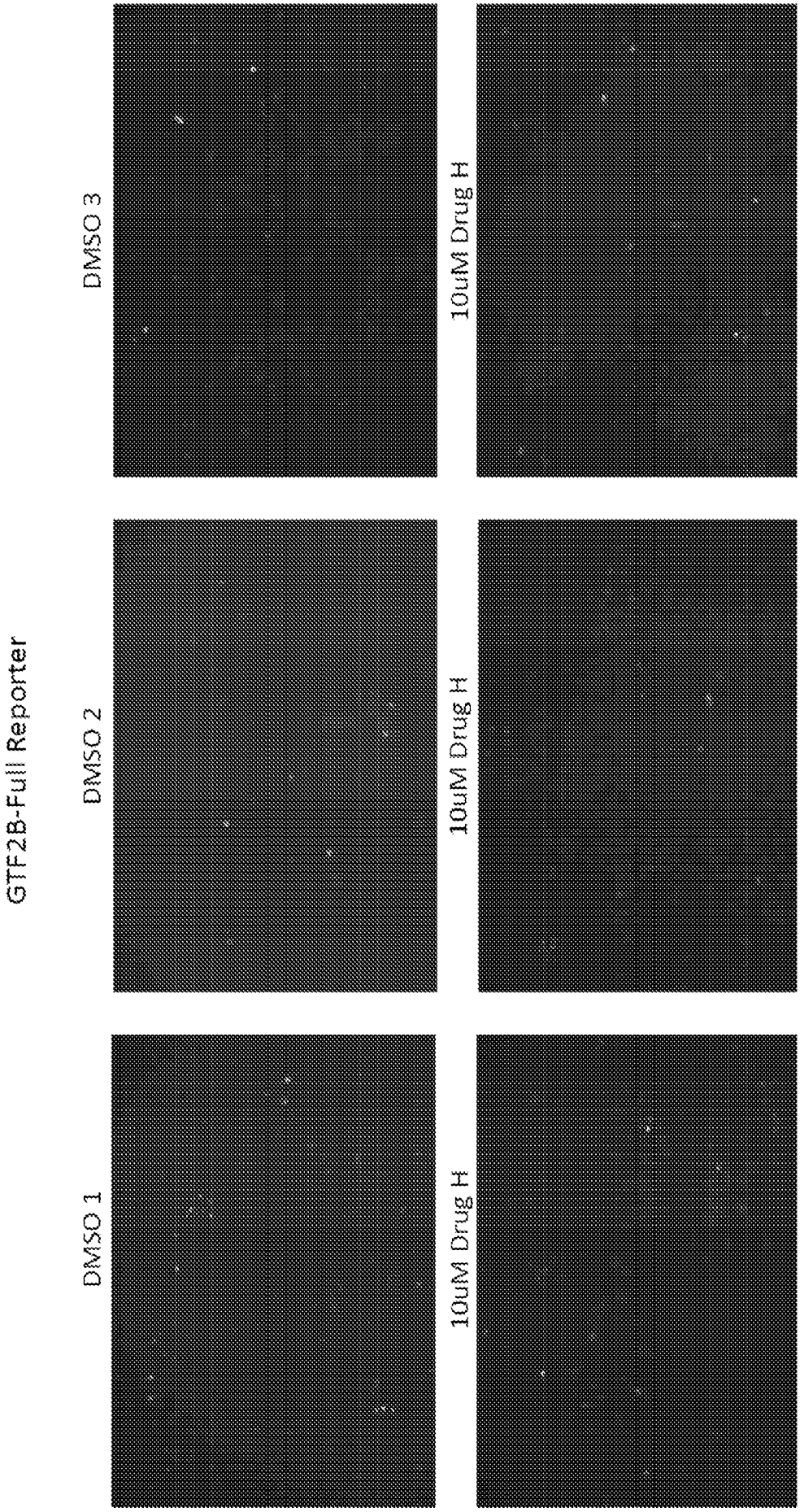

GTF2B-Full Reporter
DMSO 1
10uM Drug H
DMSO 2
10uM Drug H
DMSO 3
10uM Drug H

FIG. 8A

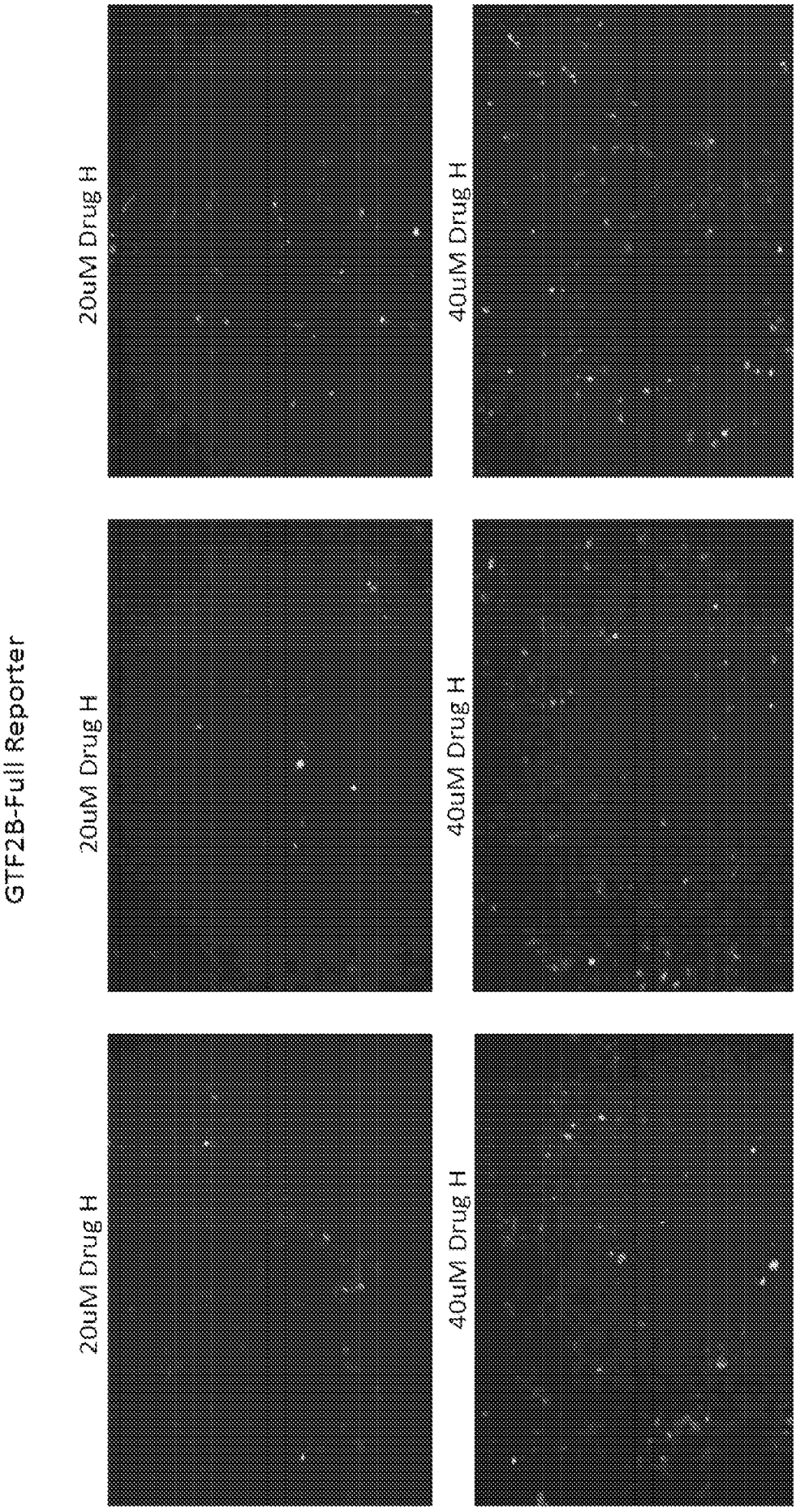

GTF2B-Full Reporter
20uM Drug H
40uM Drug H
20uM Drug H
40uM Drug H
20uM Drug H
40uM Drug H

FIG. 8B

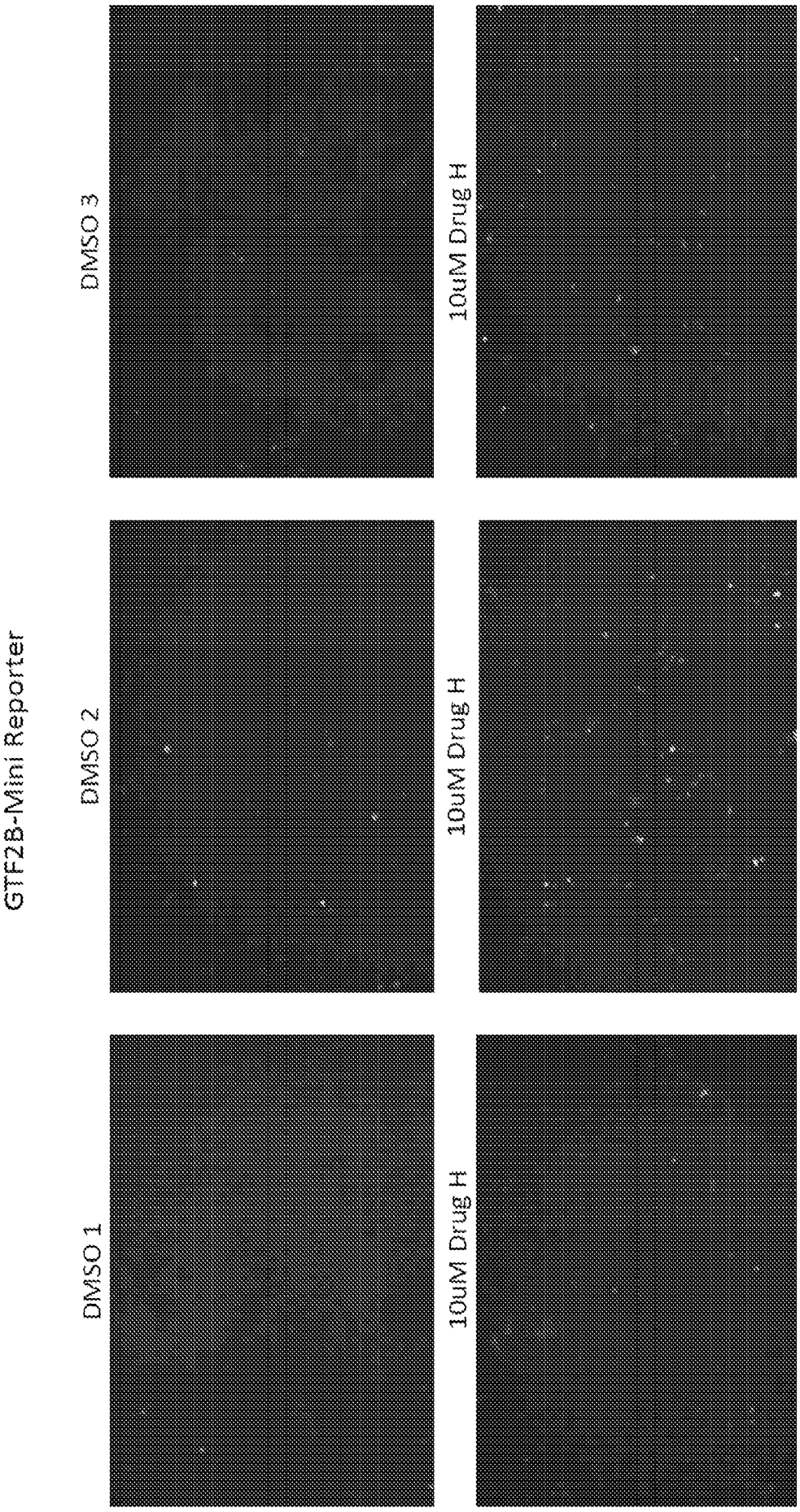

GTF2B-Mini Reporter
DMSO 1
10uM Drug H
DMSO 2
10uM Drug H
DMSO 3
10uM Drug H

FIG. 8C

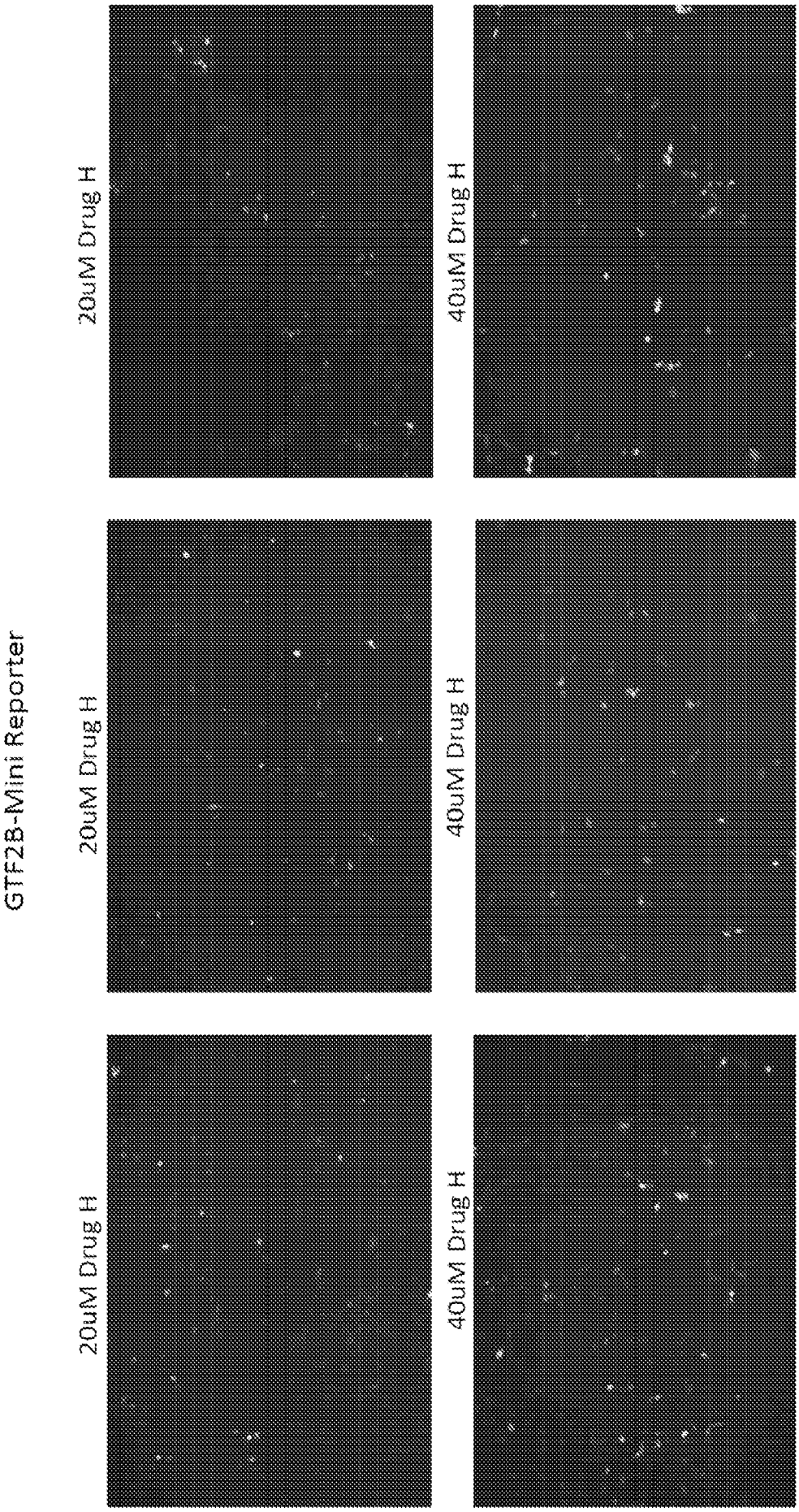

GTF2B-Mini Reporter
20uM Drug H
40uM Drug H
20uM Drug H
40uM Drug H
20uM Drug H
40uM Drug H

FIG. 8D

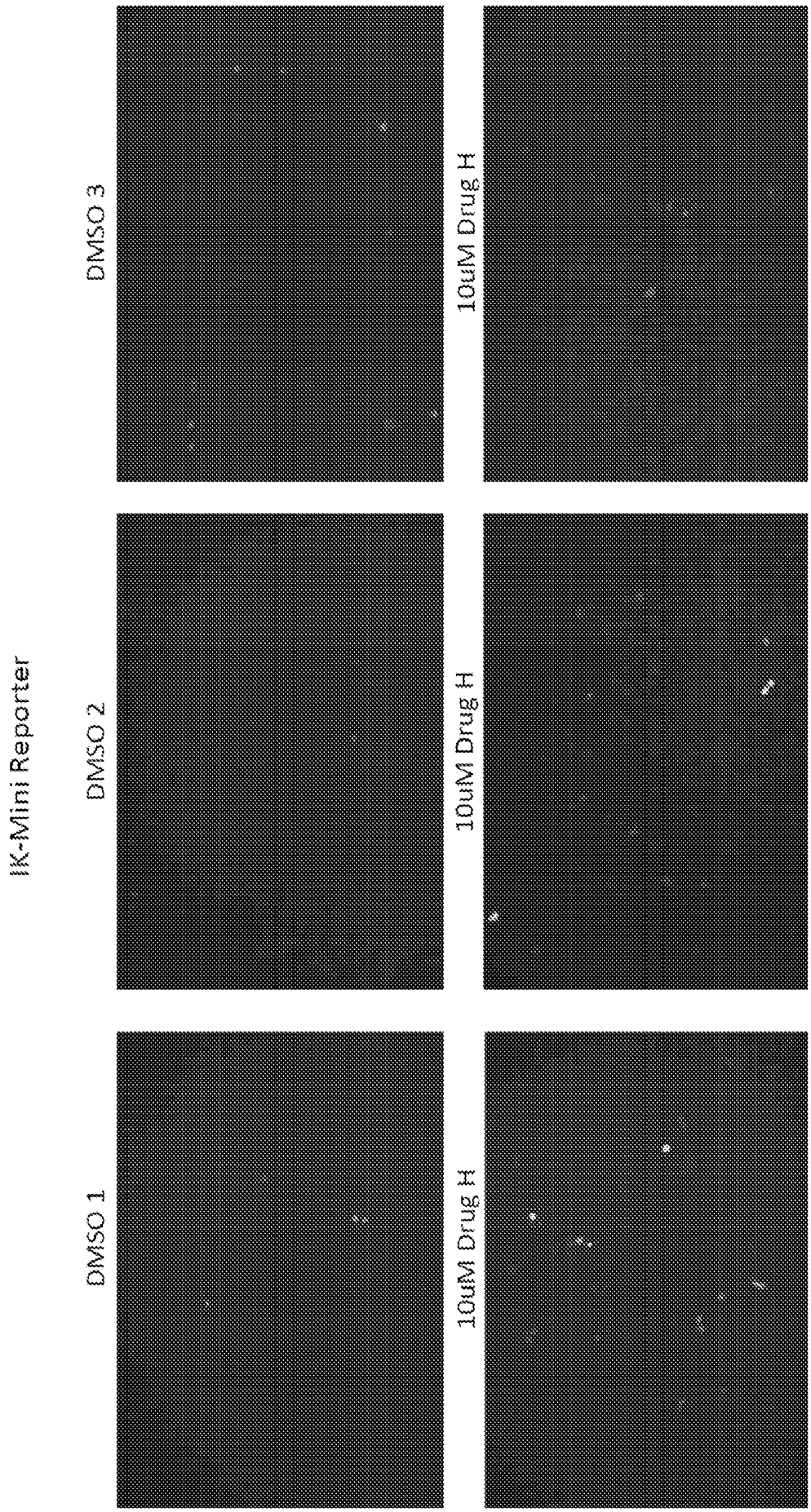

IK-Mini Reporter
DMSO 1
10uM Drug H
DMSO 2
10uM Drug H
DMSO 3
10uM Drug H

FIG. 8E

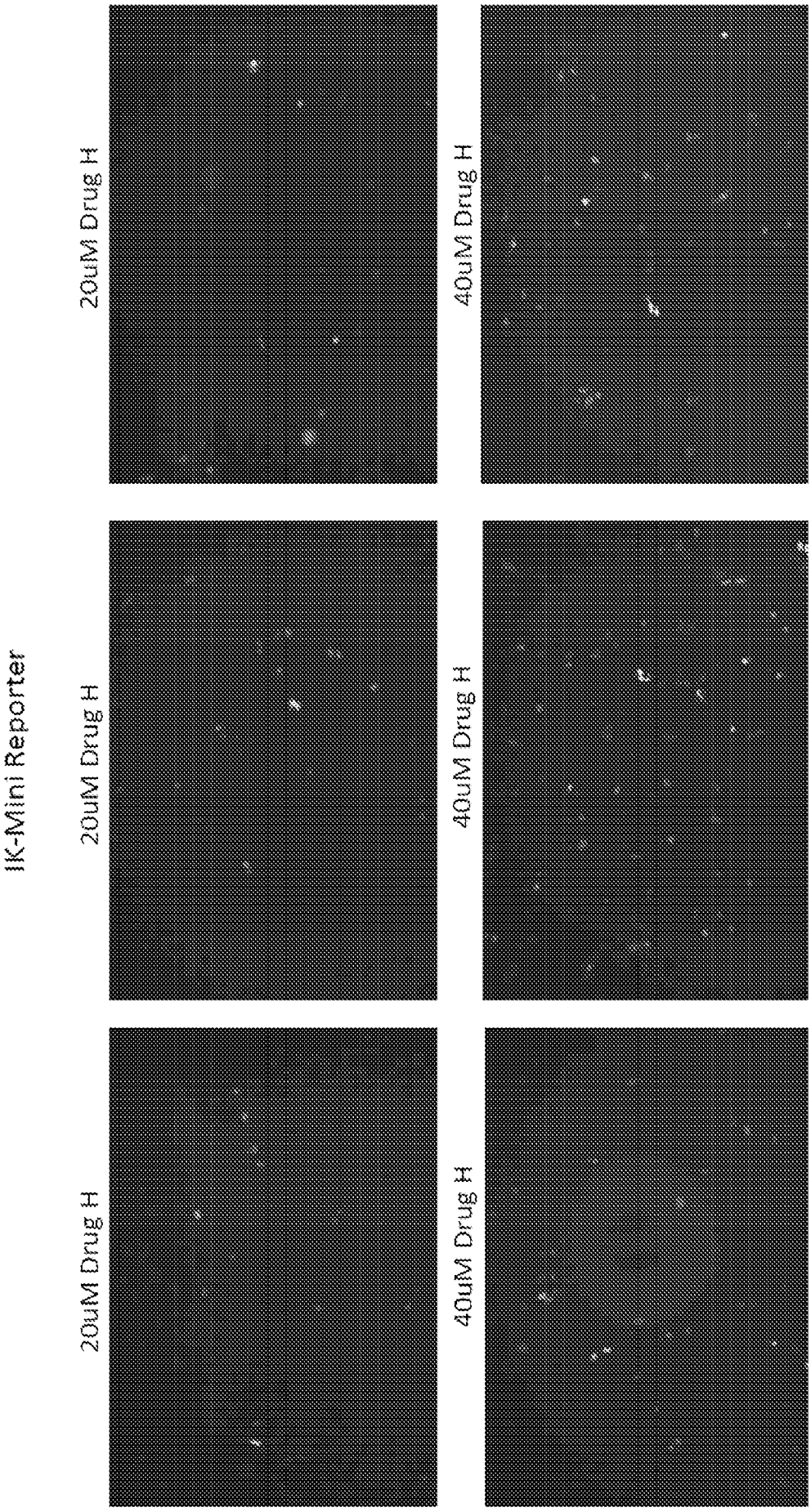

IK-Mini Reporter
20uM Drug H
40uM Drug H
20uM Drug H
40uM Drug H
20uM Drug H
40uM Drug H

FIG. 8F

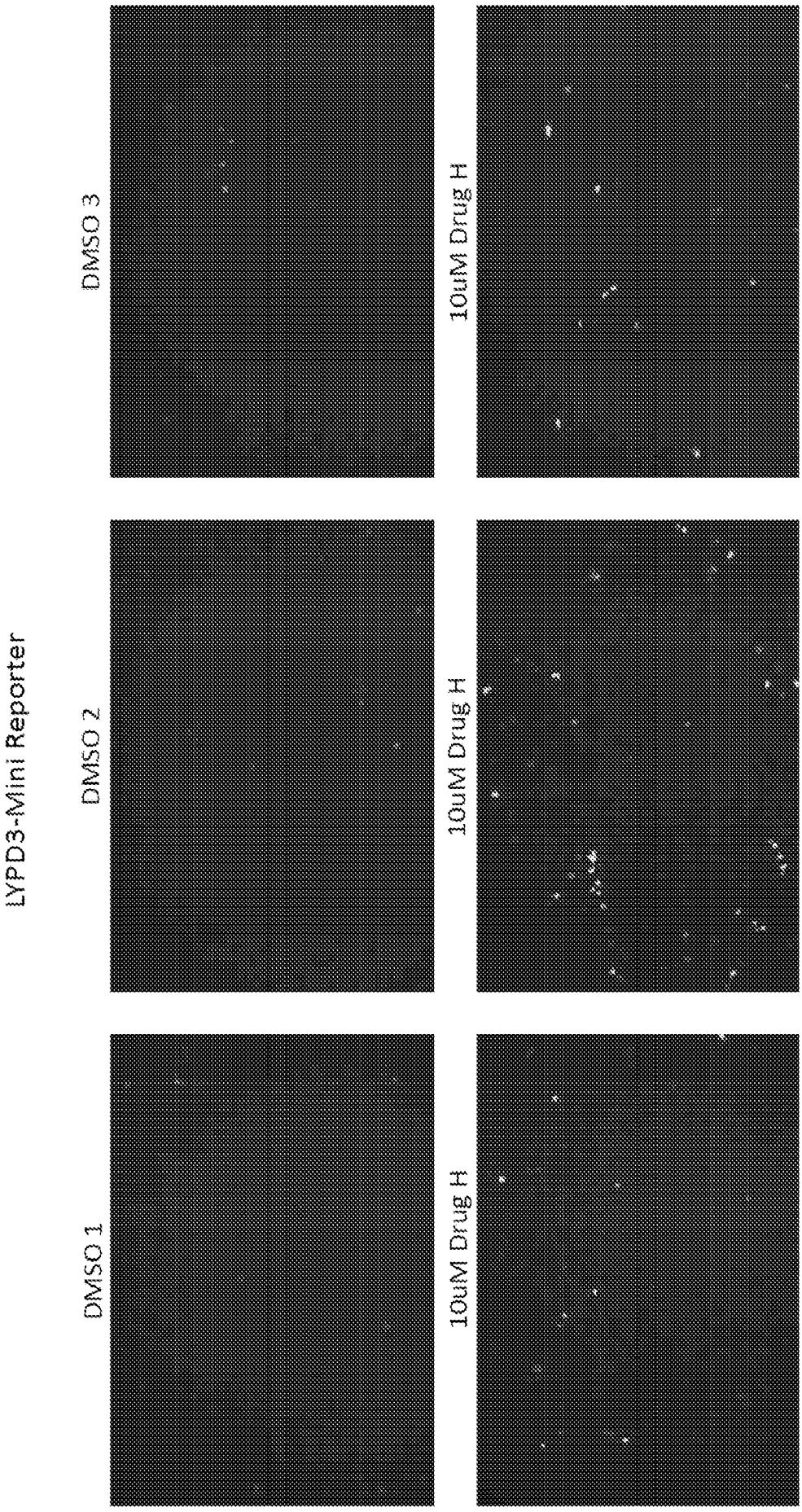

LYPD3-Mini Reporter
DMSO 1
10uM Drug H
DMSO 2
10uM Drug H
DMSO 3
10uM Drug H

FIG. 8G

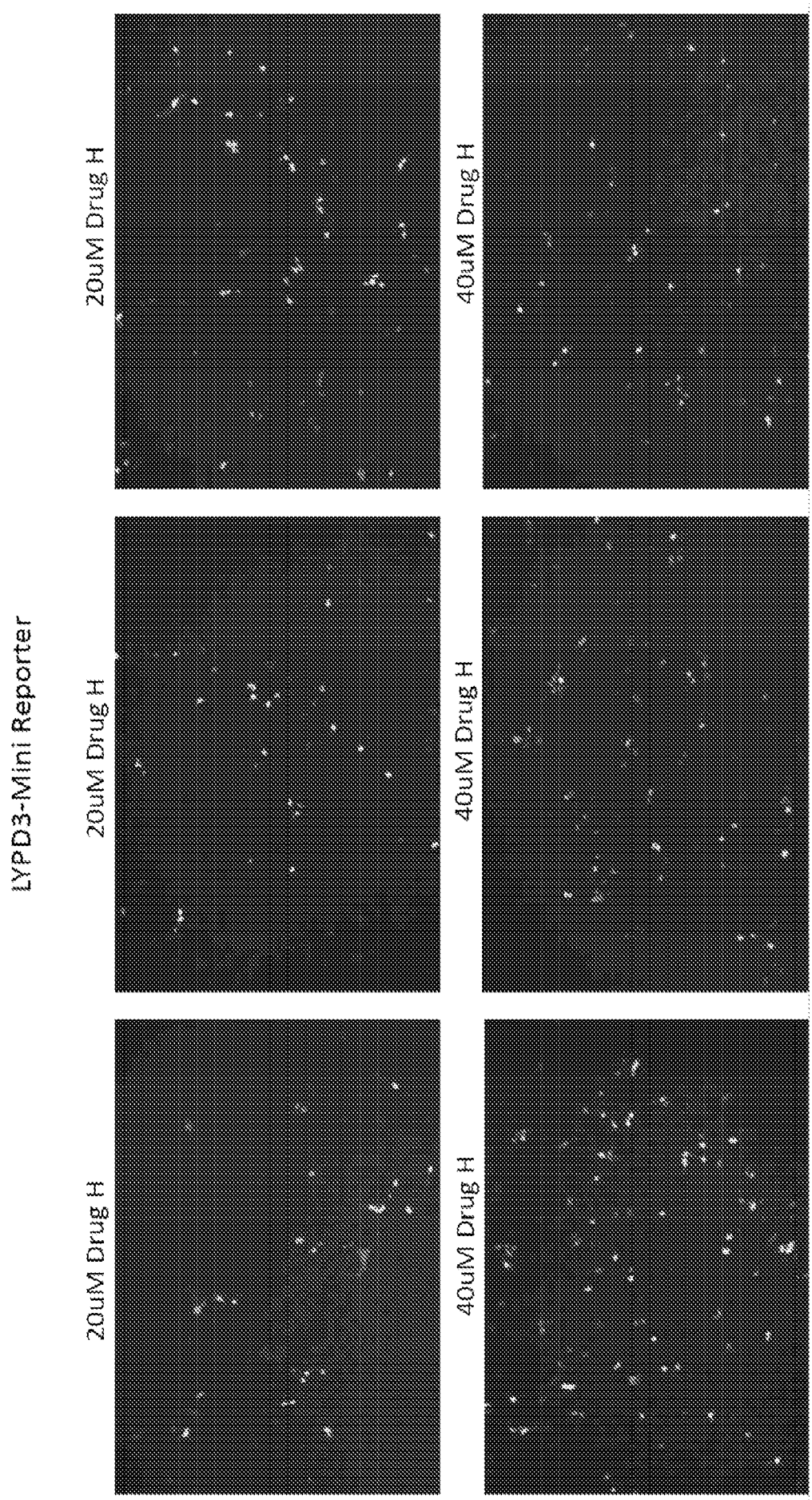

LYPD3-Mini Reporter
20uM Drug H
40uM Drug H
20uM Drug H
40uM Drug H
20uM Drug H
40uM Drug H

FIG. 8H

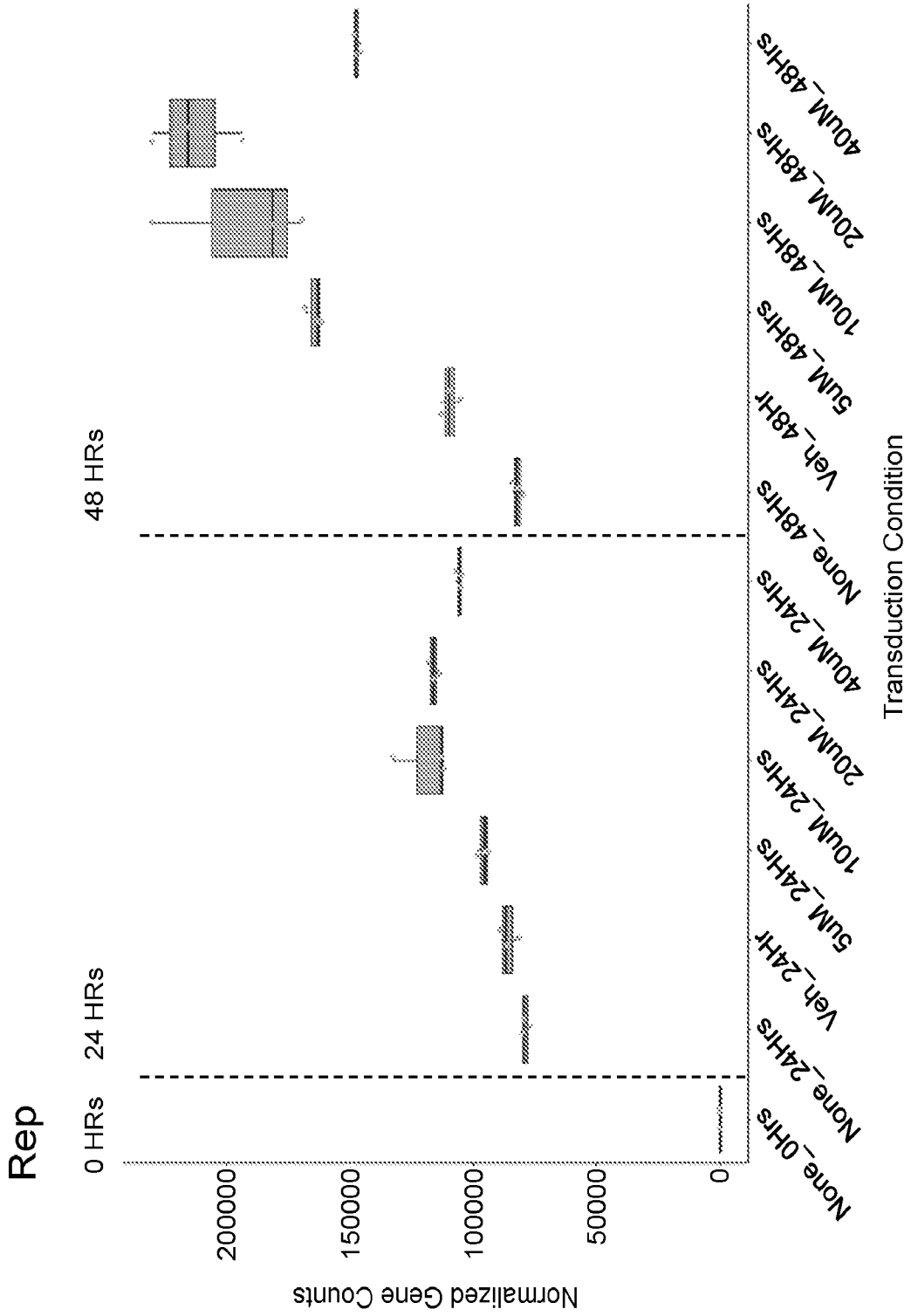

Rep
0 HRs
24 HRs
48 HRs
Normalized Gene Counts
200000
150000
100000
50000
0
None_0Hrs
None_24Hrs
Veh_24Hr
5uM_24Hrs
10uM_24Hrs
20uM_24Hrs
40uM_24Hrs
None_48Hrs
Veh_48Hr
5uM_48Hrs
10uM_48Hrs
20uM_48Hrs
40uM_48Hrs
Transduction Condition

FIG. 22B

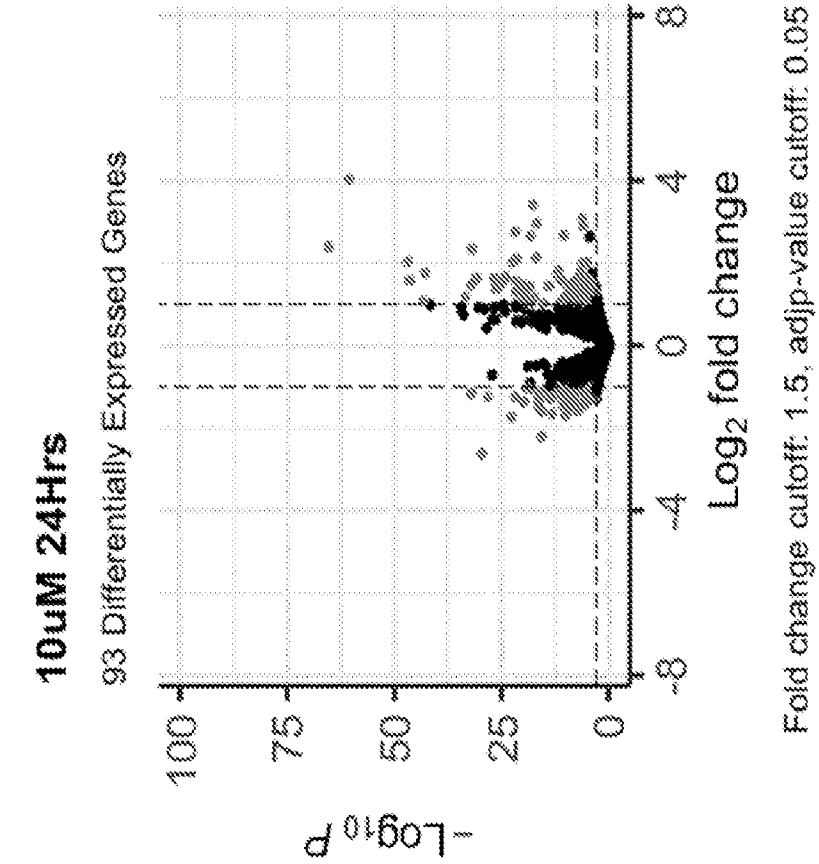


FIG. 23B

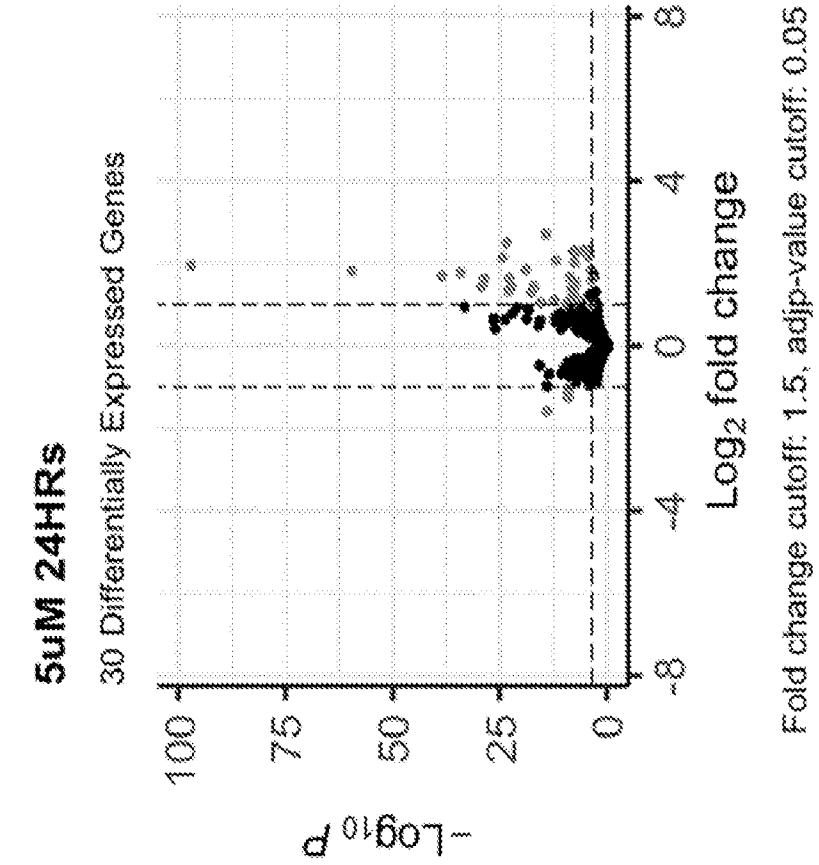


FIG. 23A

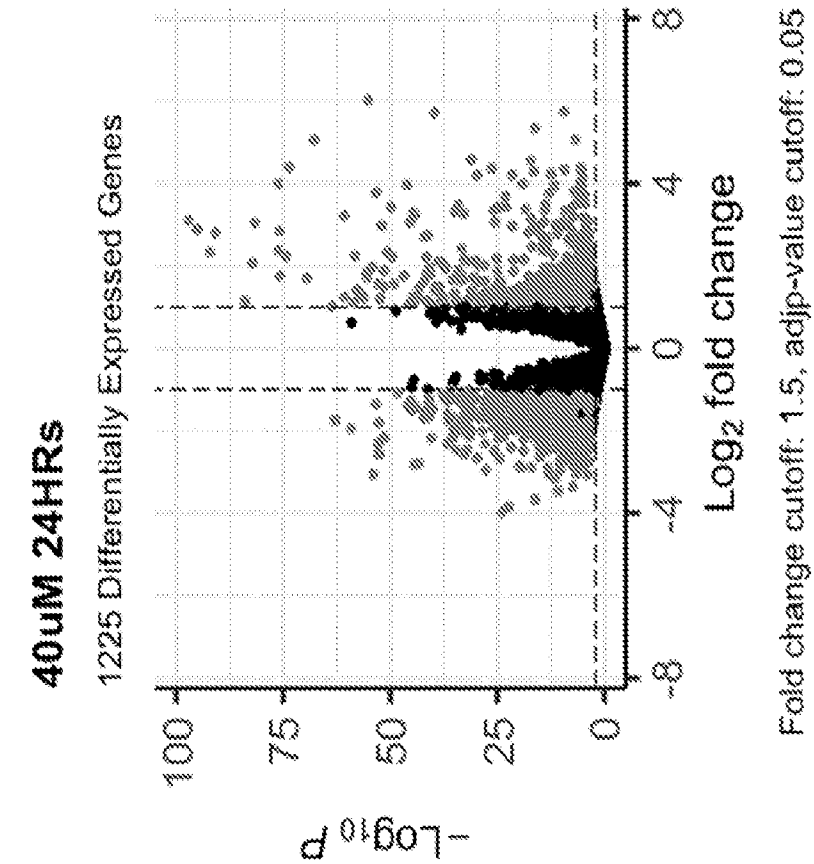


FIG. 23D

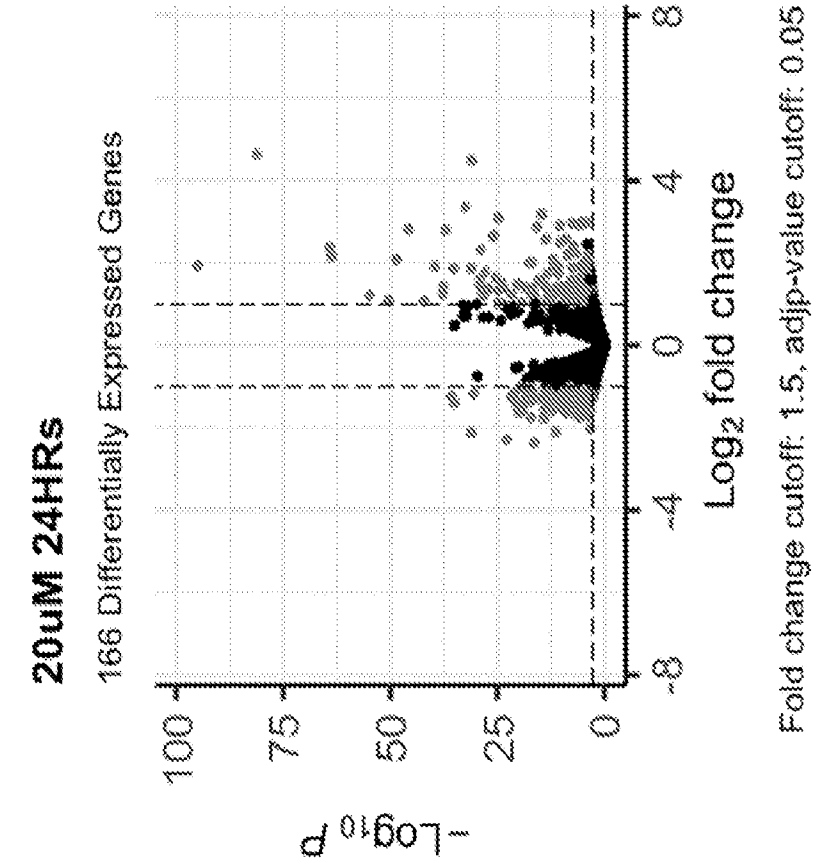


FIG. 23C

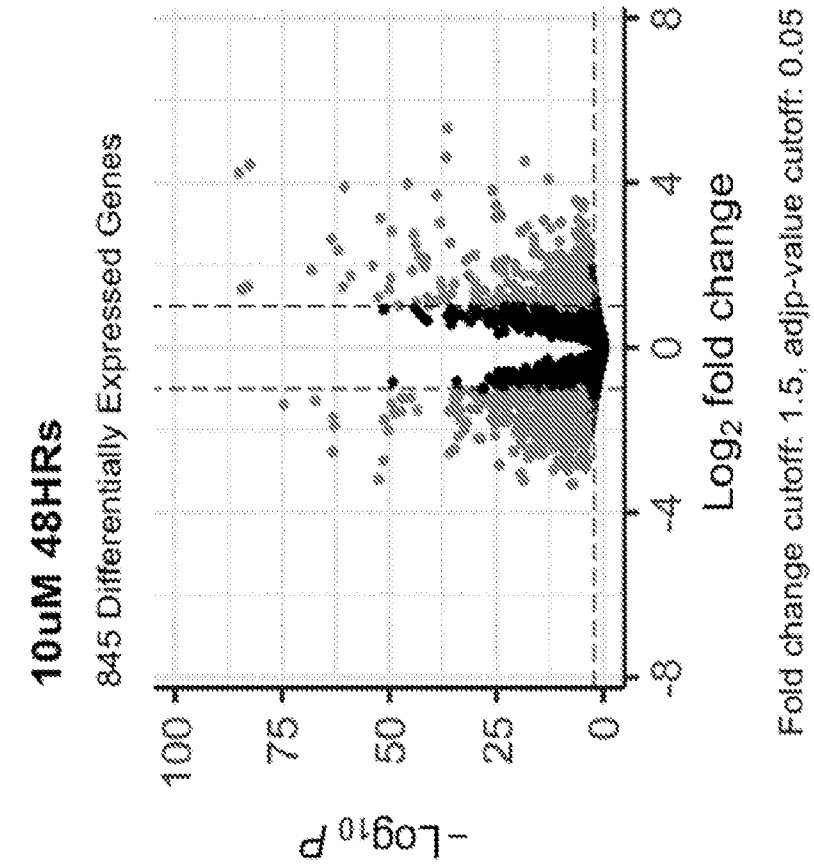


FIG. 23F

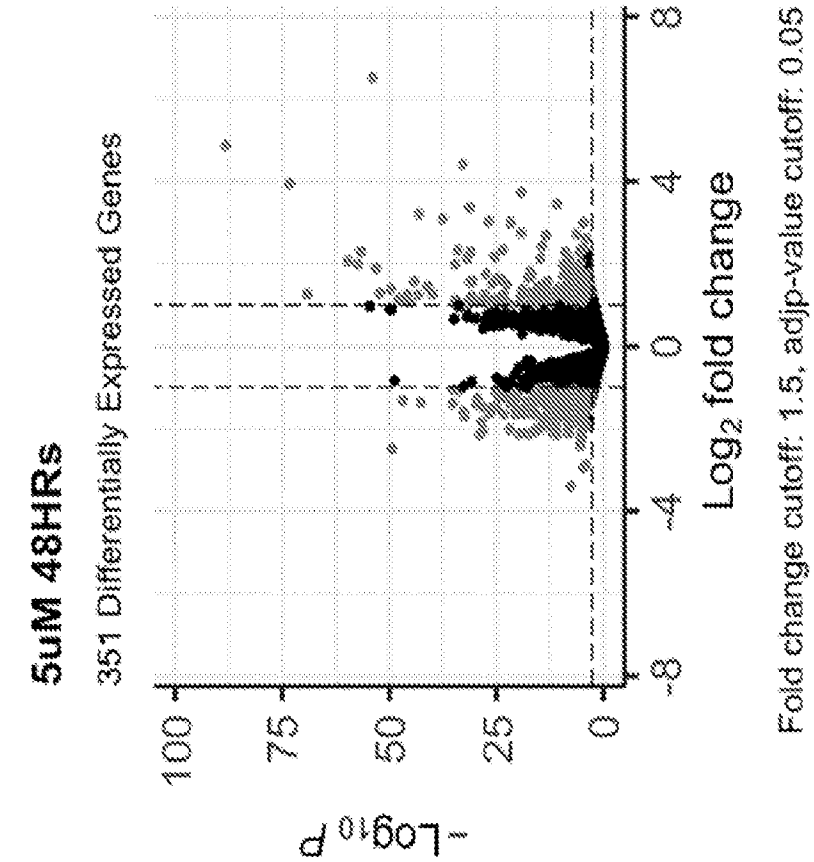


FIG. 23E

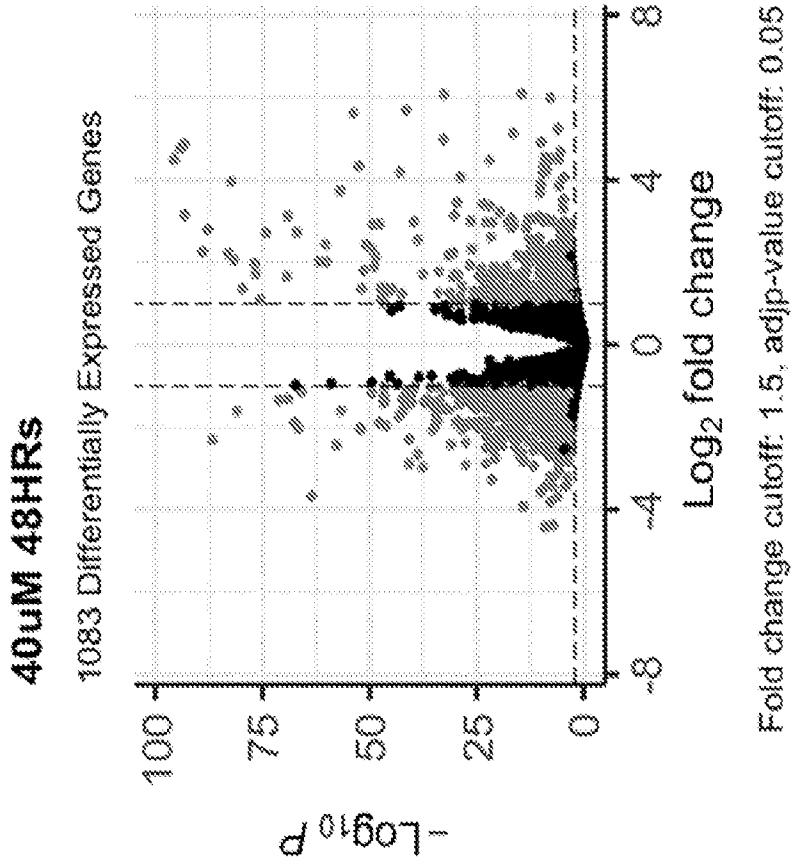


FIG. 23H

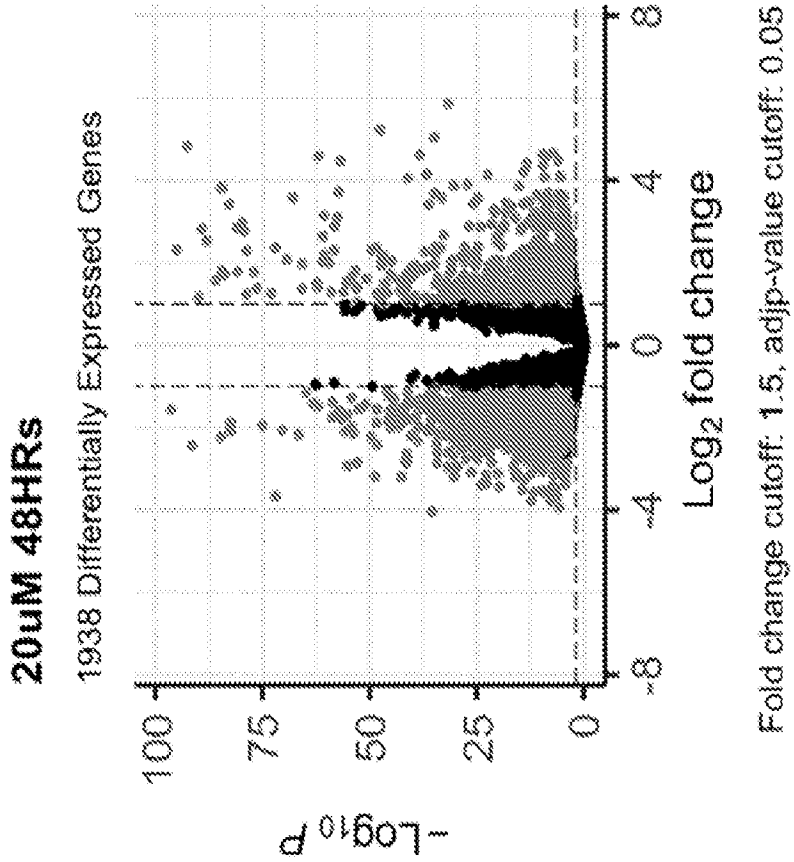


FIG. 23G

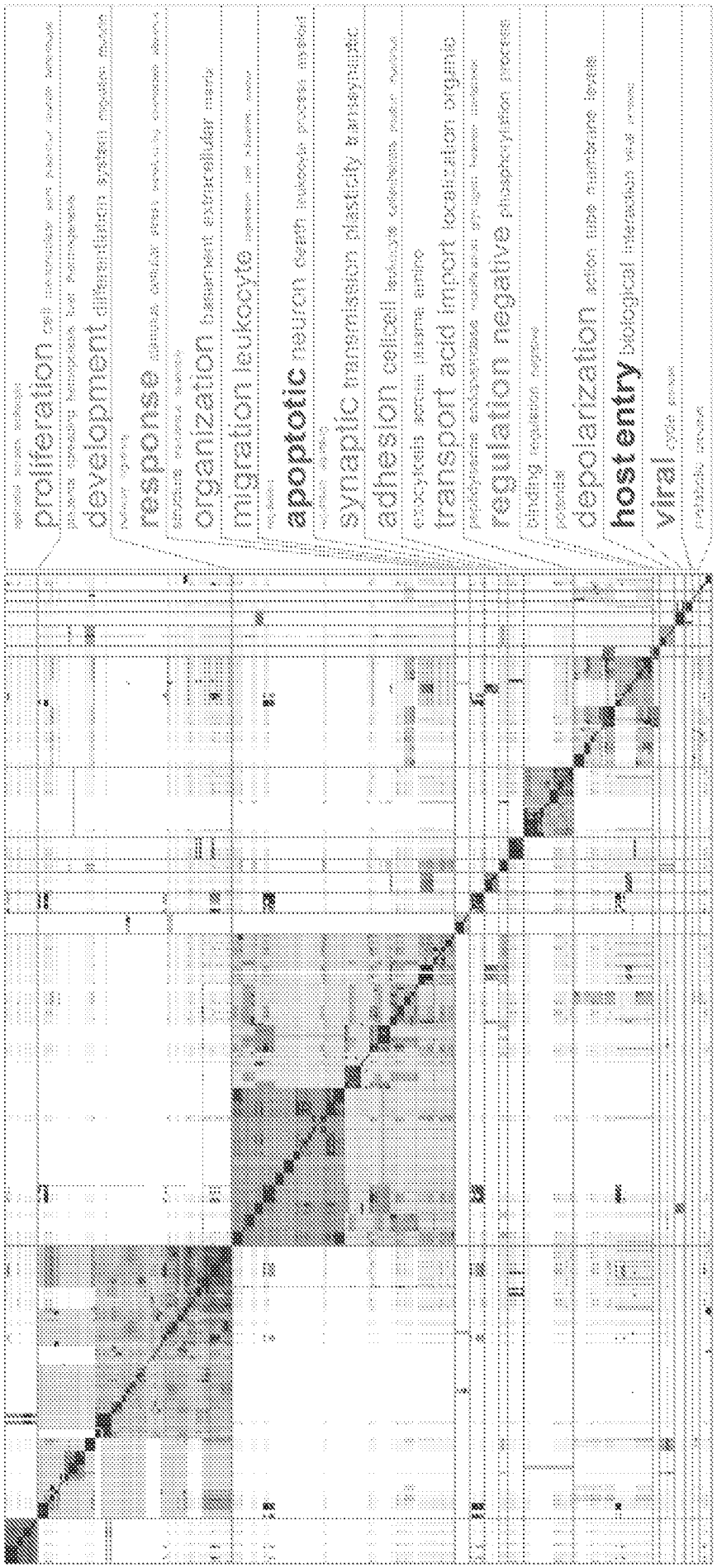

proliferation
development
response
organization
migration leukocyte
apoptotic
synaptic
adhesion cellular
transport acid import
regulation negative
depolarization
host entry
viral
Similarity
0.8
0.6
0.4
0.2
0

FIG. 24

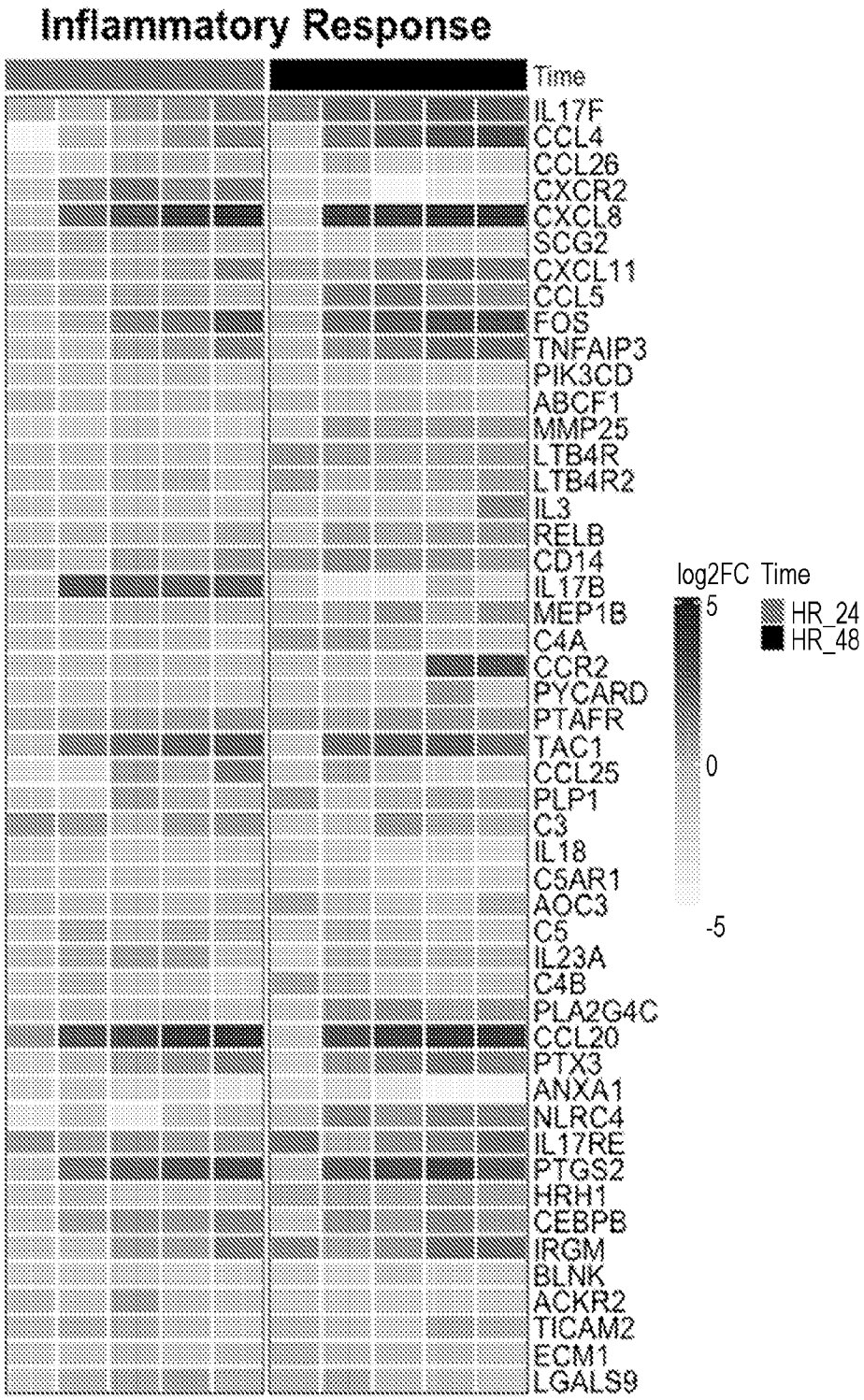

Inflammatory Response
Time
IL17F
CCL4
CCL26
CXCR2
CXCL8
SCG2
CXCL11
CCL5
FOS
TNFAIP3
PIK3CD
ABCF1
MMP25
LTB4R
LTB4R2
IL3
RELB
CD14
IL17B
MEP1B
C4A
CCR2
PYCARD
PTAFR
TAC1
CCL25
PLP1
C3
IL18
C5AR1
AOC3
C5
IL23A
C4B
PLA2G4C
CCL20
PTX3
ANXA1
NLRC4
IL17RE
PTGS2
HRH1
CEBPB
IRGM
BLNK
ACKR2
TICAM2
ECM1
LGALS9
log2FC
5
0
-5
Time
HR_24
HR_48

FIG. 25A

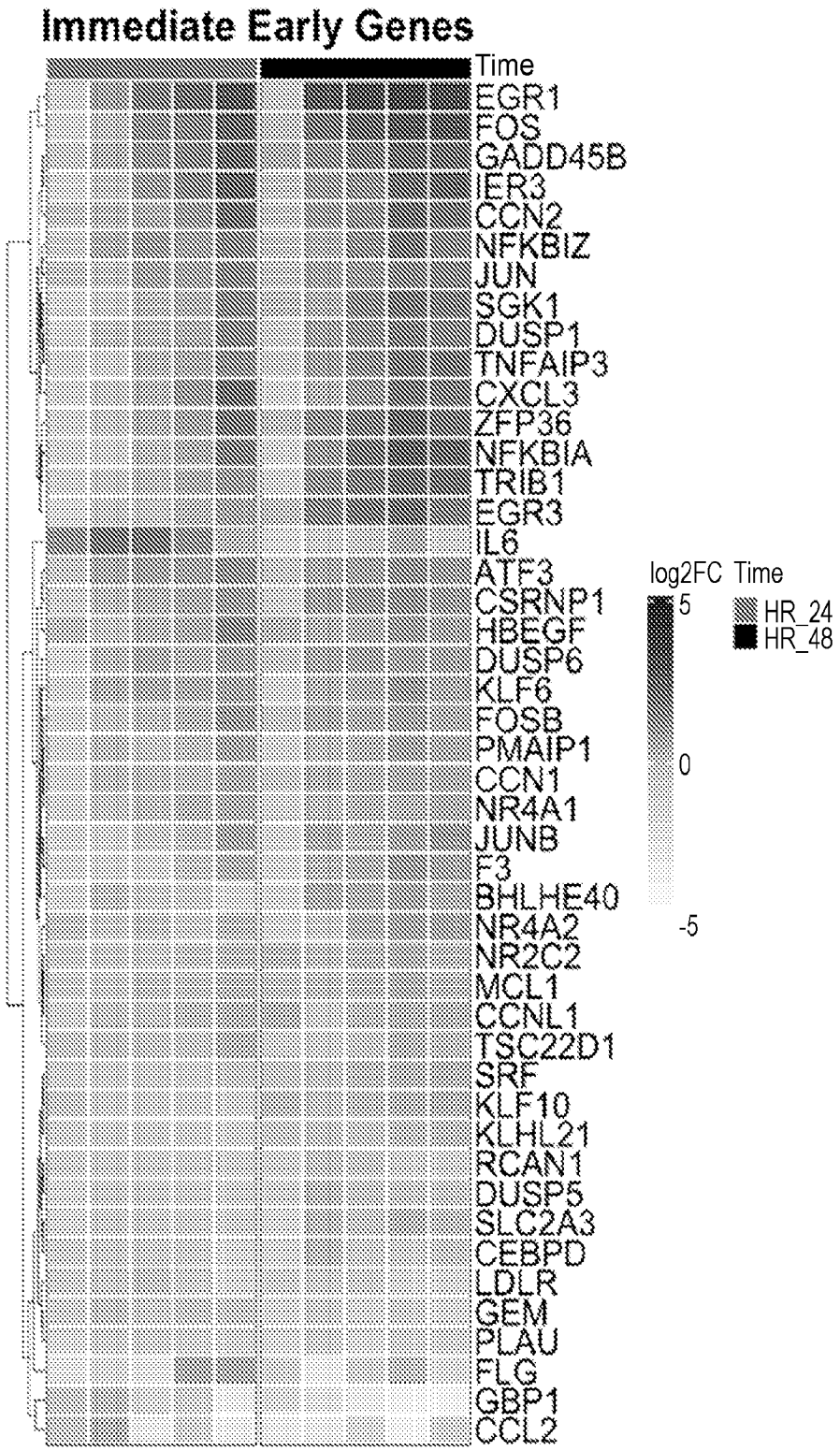

Immediate Early Genes
Time
EGR1
FOS
GADD45B
IER3
CCN2
NFKBIZ
JUN
SGK1
DUSP1
TNFAIP3
CXCL3
ZFP36
NFKBIA
TRIB1
EGR3
IL6
ATF3
CSRNP1
HBEGF
DUSP6
KLF6
FOSB
PMAIP1
CCN1
NR4A1
JUNB
F3
BHLHE40
NR4A2
NR2C2
MCL1
CCNL1
TSC22D1
SRF
KLF10
KLHL21
RCAN1
DUSP5
SLC2A3
CEBPD
LDLR
GEM
PLAU
FLG
GBP1
CCL2
log2FC
5
0
-5
Time
HR_24
HR_48

FIG. 25B

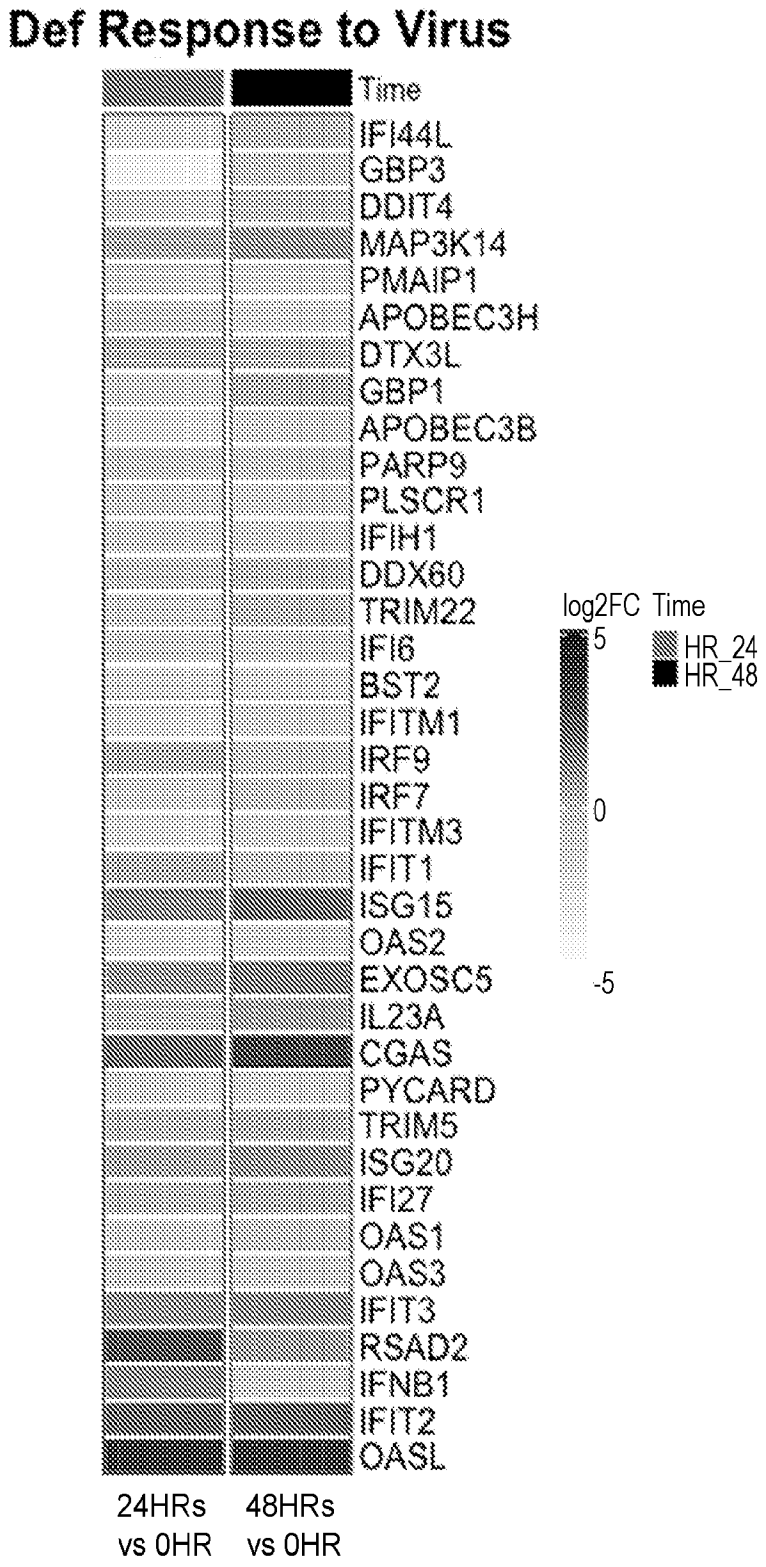

Def Response to Virus
Time
IFI44L
GBP3
DDIT4
MAP3K14
PMAIP1
APOBEC3H
DTX3L
GBP1
APOBEC3B
PARP9
PLSCR1
IFIH1
DDX60
TRIM22
IFI6
BST2
IFITM1
IRF9
IRF7
IFITM3
IFIT1
ISG15
OAS2
EXOSC5
IL23A
CGAS
PYCARD
TRIM5
ISG20
IFI27
OAS1
OAS3
IFIT3
RSAD2
IFNB1
IFIT2
OASL
log2FC Time
5
0
-5
HR_24
HR_48
24HRs vs 0HR
48HRs vs 0HR

FIG. 25C

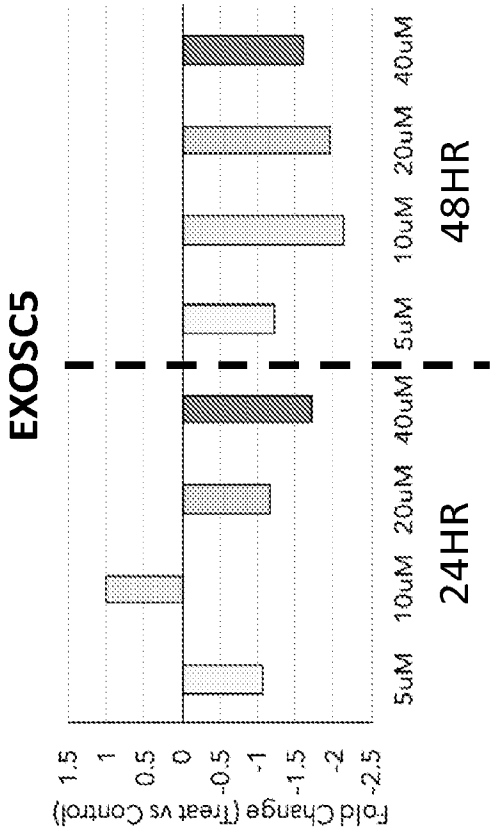


FIG. 25E

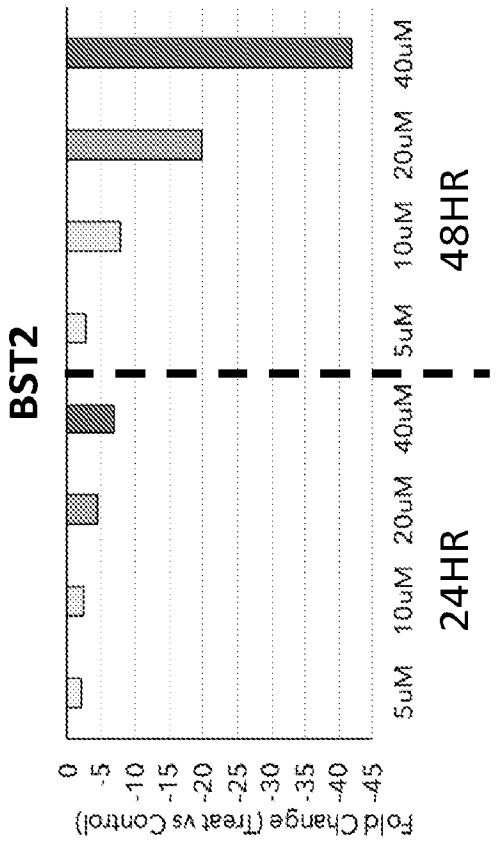


FIG. 25D

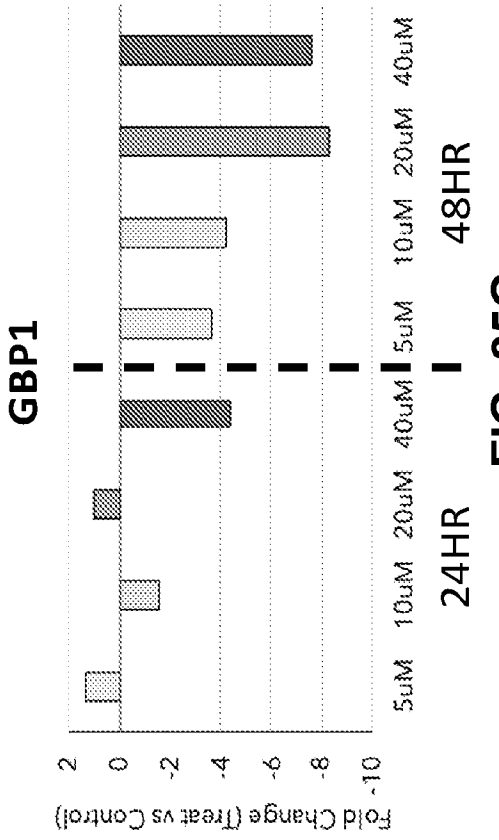


FIG. 25G

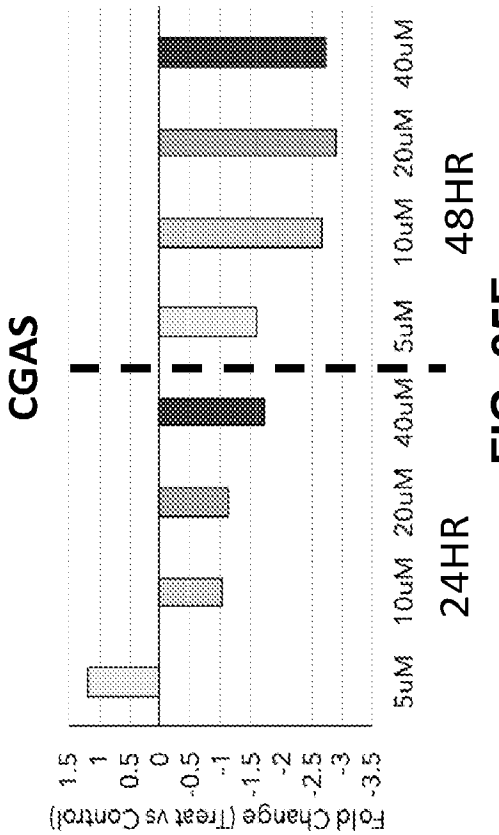


FIG. 25F

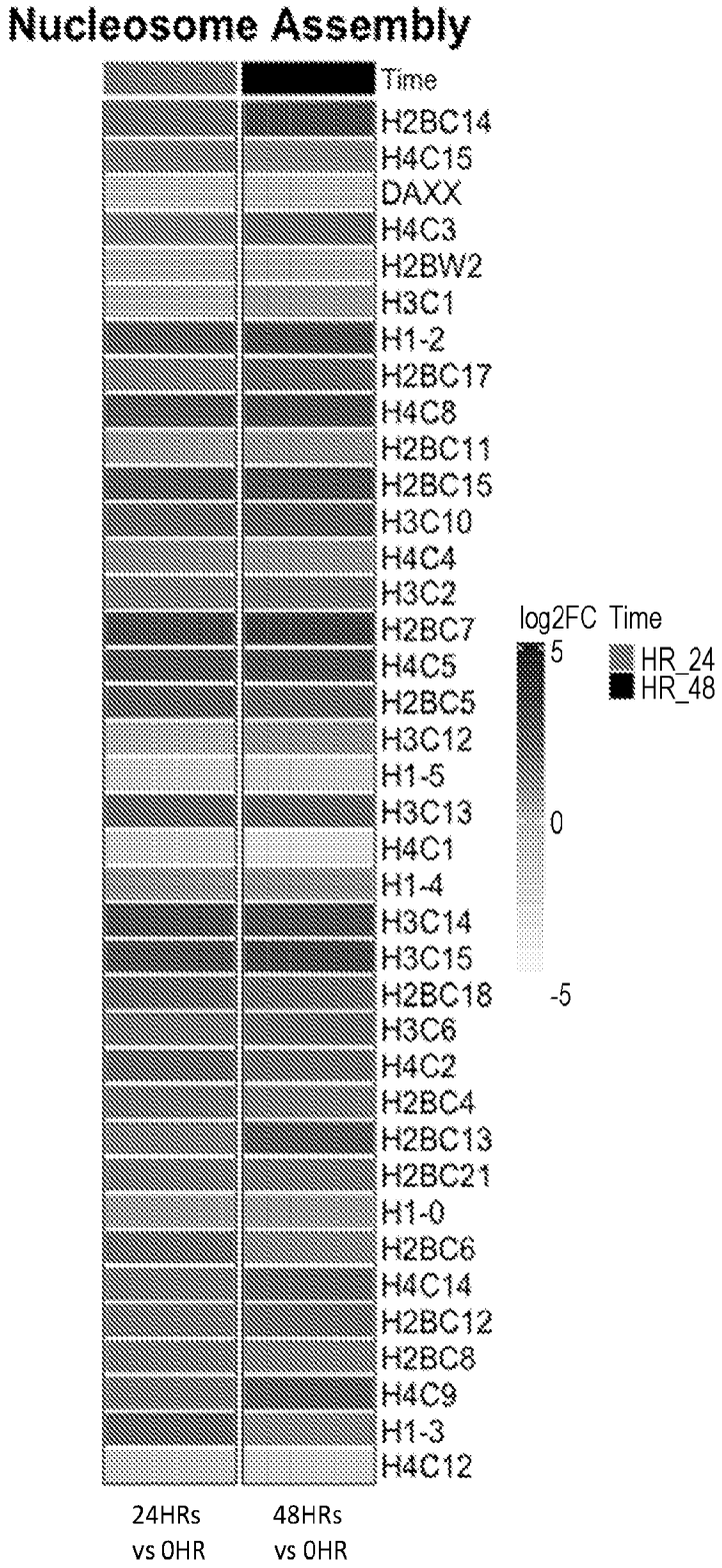

Nucleosome Assembly
Time
H2BC14
H4C15
DAXX
H4C3
H2BW2
H3C1
H1-2
H2BC17
H4C8
H2BC11
H2BC15
H3C10
H4C4
H3C2
H2BC7
H4C5
H2BC5
H3C12
H1-5
H3C13
H4C1
H1-4
H3C14
H3C15
H2BC18
H3C6
H4C2
H2BC4
H2BC13
H2BC21
H1-0
H2BC6
H4C14
H2BC12
H2BC8
H4C9
H1-3
H4C12
log2FC
5
0
-5
Time
HR_24
HR_48
24HRs vs 0HR
48HRs vs 0HR

FIG. 26A

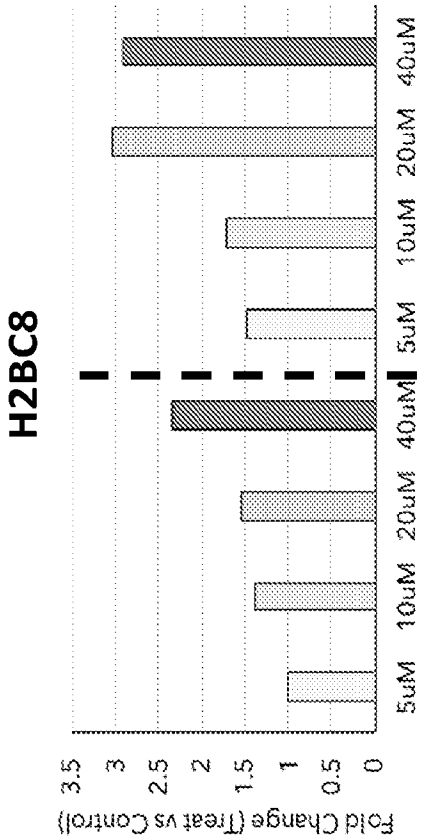


FIG. 26B

FIG. 26C

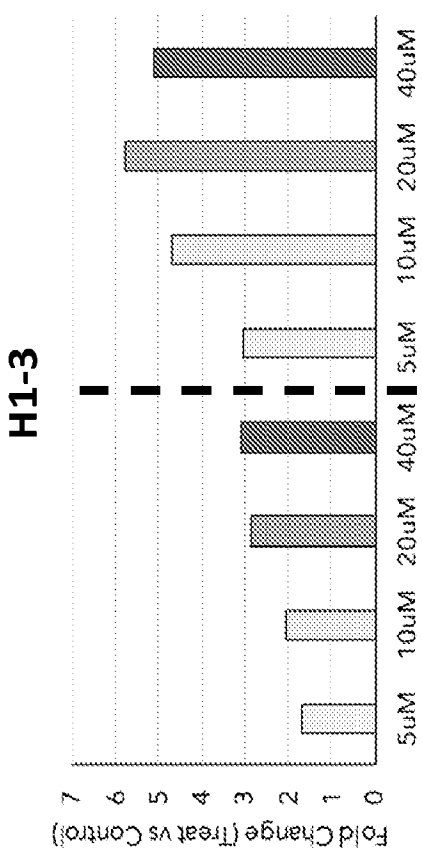


FIG. 26D

FIG. 26E

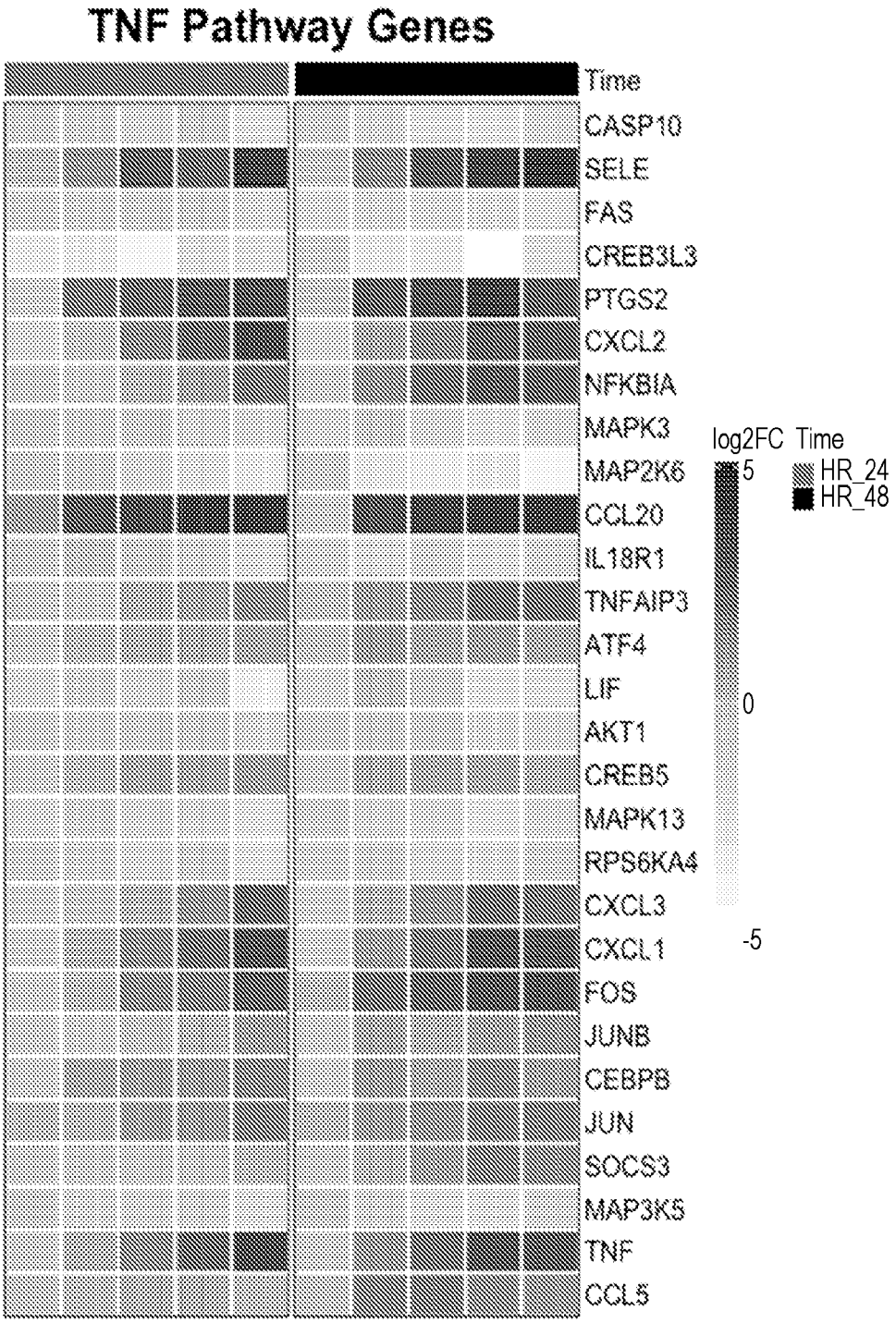

TNF Pathway Genes
Time
CASP10
SELE
FAS
CREB3L3
PTGS2
CXCL2
NFKBIA
MAPK3
MAP2K6
CCL20
IL18R1
TNFAIP3
ATF4
LIF
AKT1
CREB5
MAPK13
RPS6KA4
CXCL3
CXCL1
FOS
JUNB
CEBPB
JUN
SOCS3
MAP3K5
TNF
CCL5
log2FC Time
5
0
-5
HR_24
HR_48

FIG. 27A

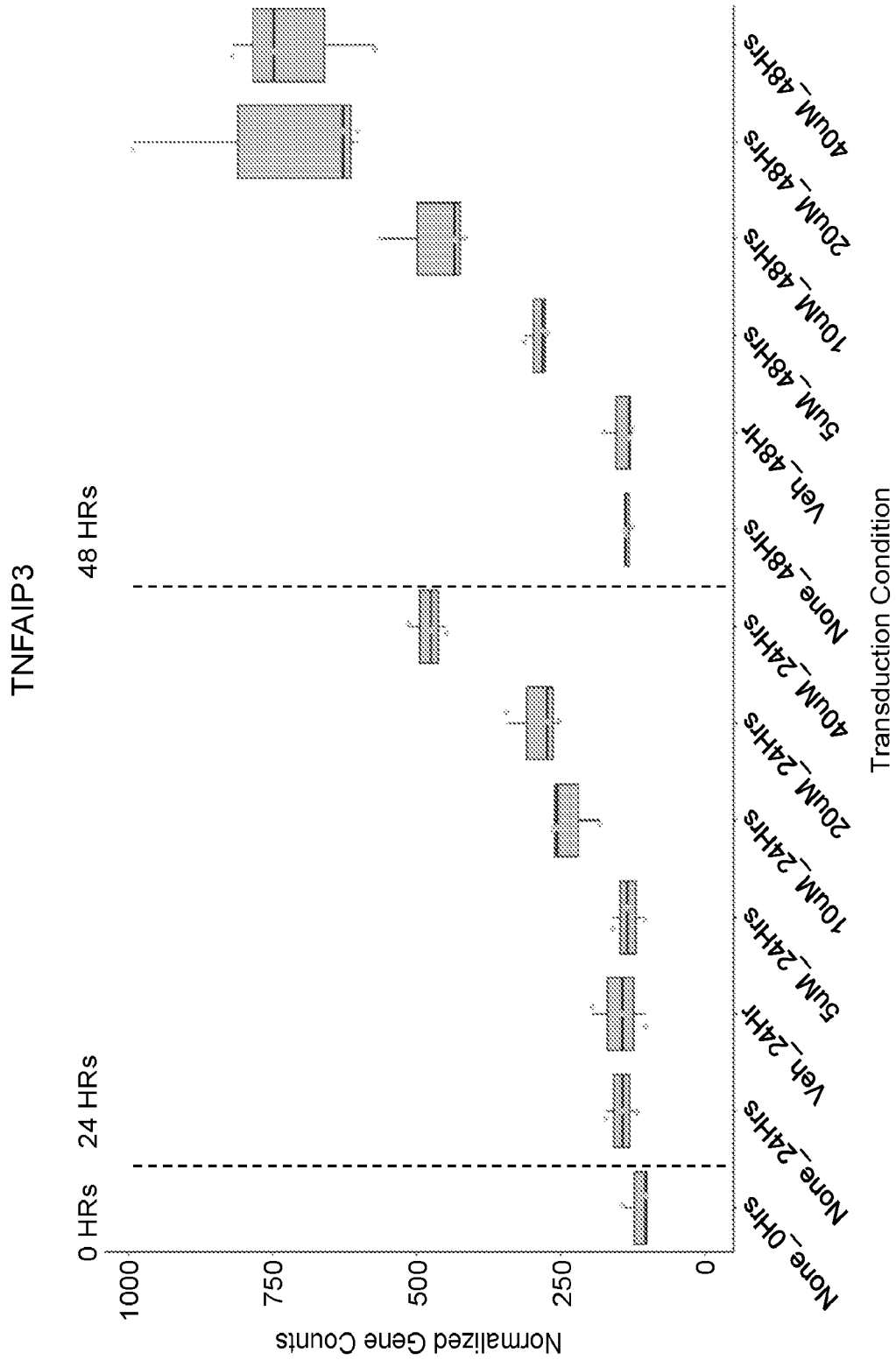

TNFAIP3
0 HRs
24 HRs
48 HRs
Normalized Gene Counts
1000
750
500
250
0
None_0Hrs
None_24Hrs
Veh_24Hr
5uM_24Hrs
10uM_24Hrs
20uM_24Hrs
40uM_24Hrs
None_48Hrs
Veh_48Hr
5uM_48Hrs
10uM_48Hrs
20uM_48Hrs
40uM_48Hrs
Transduction Condition

FIG. 27B

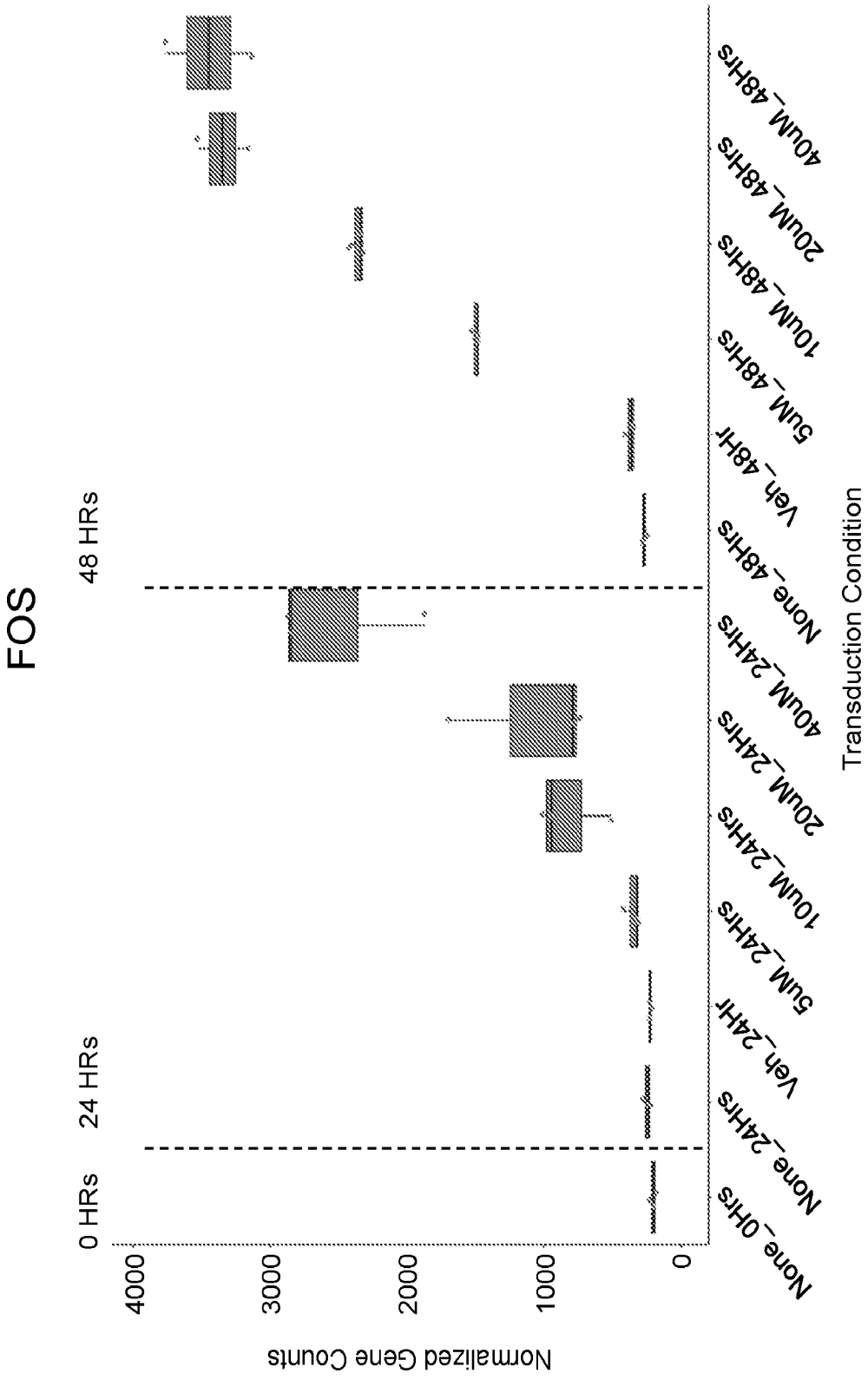

FOS
0 HRs
24 HRs
48 HRs
Normalized Gene Counts
4000
3000
2000
1000
0
None_0Hrs
None_24Hrs
Veh_24Hr
5uM_24Hrs
10uM_24Hrs
20uM_24Hrs
40uM_24Hrs
None_48Hrs
Veh_48Hr
5uM_48Hrs
10uM_48Hrs
20uM_48Hrs
40uM_48Hrs
Transduction Condition

FIG. 27C

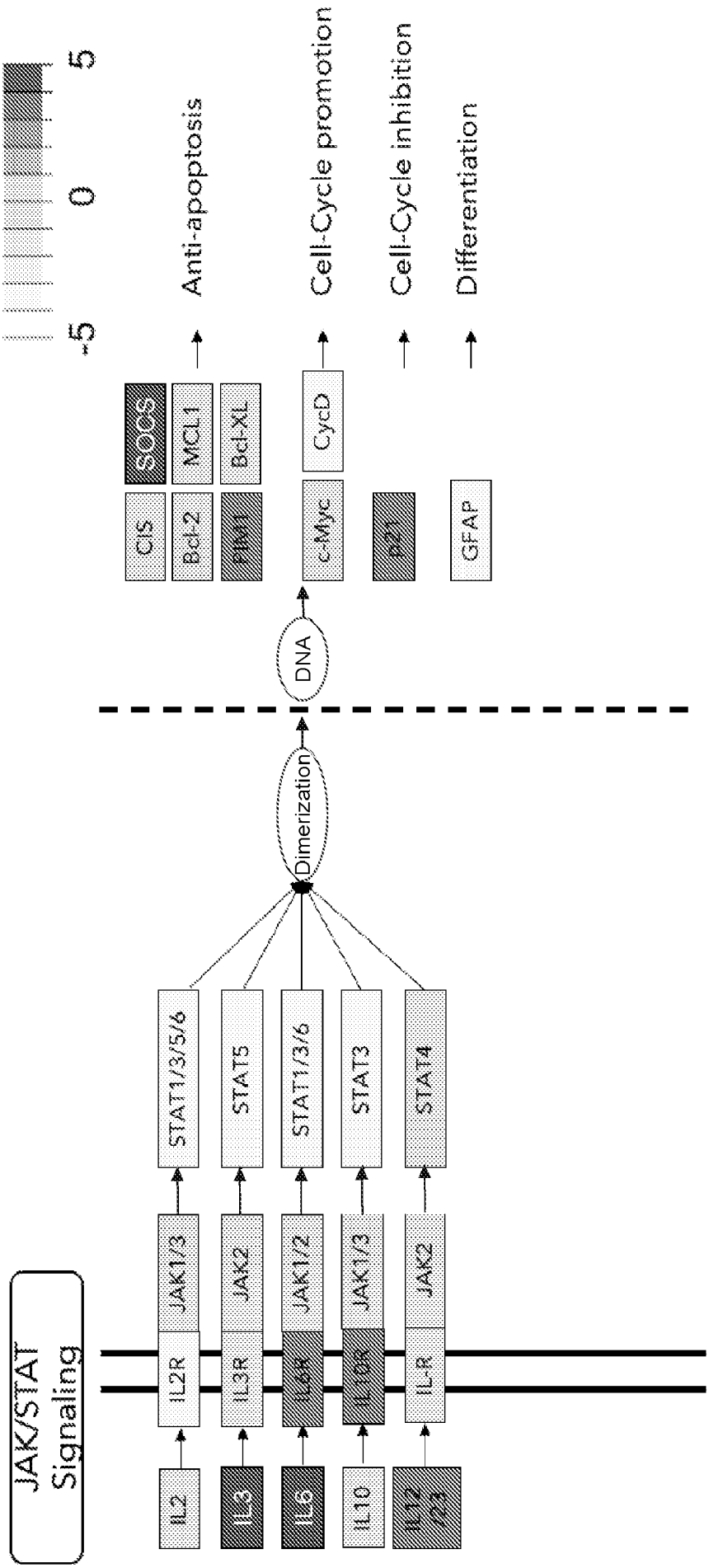

JAK/STAT Signaling
IL2
IL3
IL6
IL10
IL12 /23
IL2R
IL3R
IL6R
IL10R
IL-R
JAK1/3
JAK2
JAK1/2
JAK1/3
JAK2
STAT1/3/5/6
STAT5
STAT1/3/6
STAT3
STAT4
Dimerization
DNA
CIS
Bcl-2
Pim1
SOCS
MCL1
Bcl-XL
c-Myc
CycD
p21
GFAP
Anti-apoptosis
Cell-Cycle promotion
Cell-Cycle inhibition
Differentiation
-5
0
5

FIG. 28A

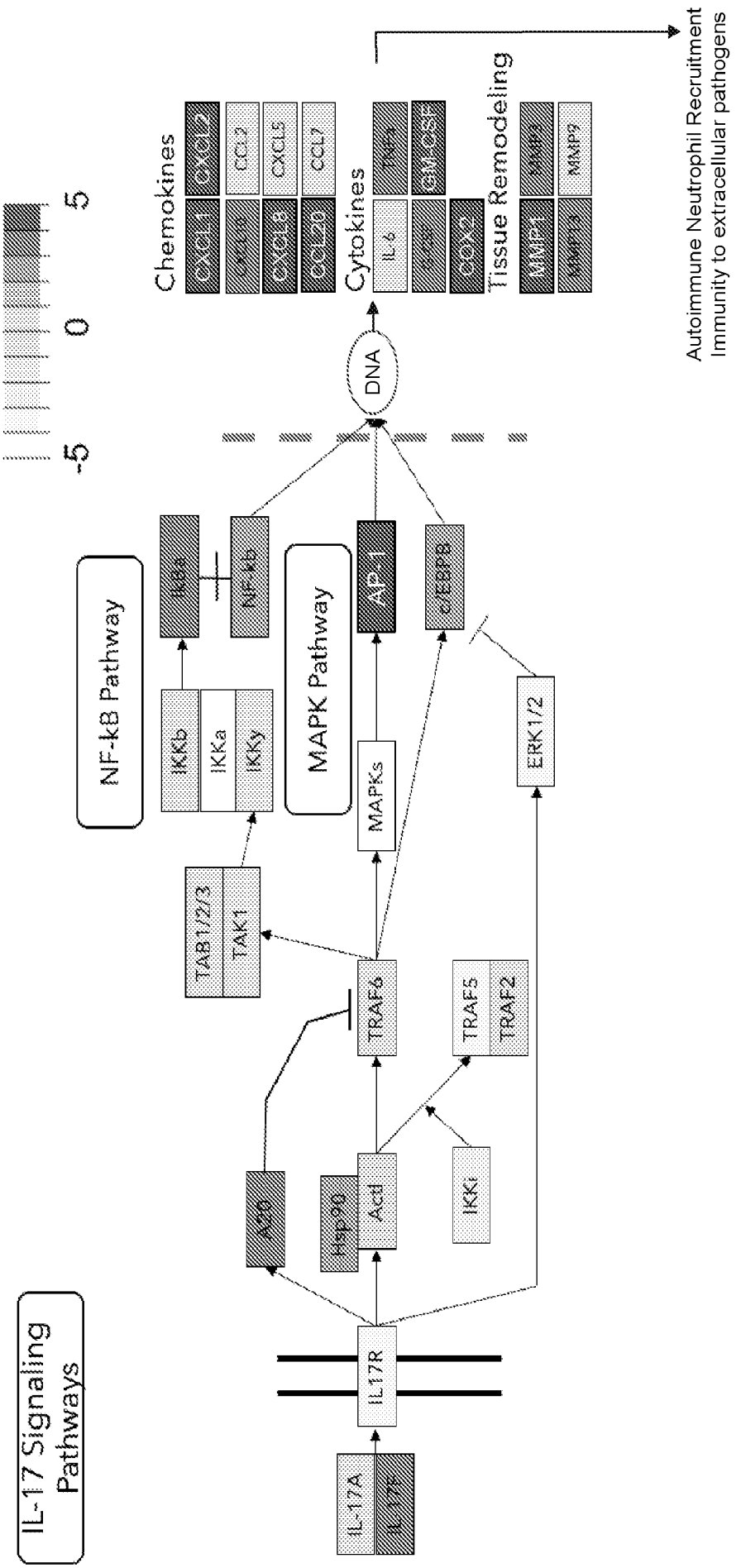

IL-17 Signaling Pathways
IL-17A
IL-17F
IL17R
Act1
Hsp90
A20
IKKi
TRAF6
TRAF5
TRAF2
TAB1/2/3
TAK1
IKKb
IKKa
IKKγ
NF-kB Pathway
IκBα
NF-kB
MAPKs
MAPK Pathway
AP-1
C/EBPβ
ERK1/2
DNA
Chemokines
CXCL1
CXCL2
CCL3
CXCL5
CXCL8
CCL20
CCL7
Cytokines
IL-6
TNFα
GM-CSF
GROX
Tissue Remodeling
MMP1
MMP3
MMP7
MMP9
-5
0
5
Autoimmune Neutrophil Recruitment
Immunity to extracellular pathogens

FIG. 28B

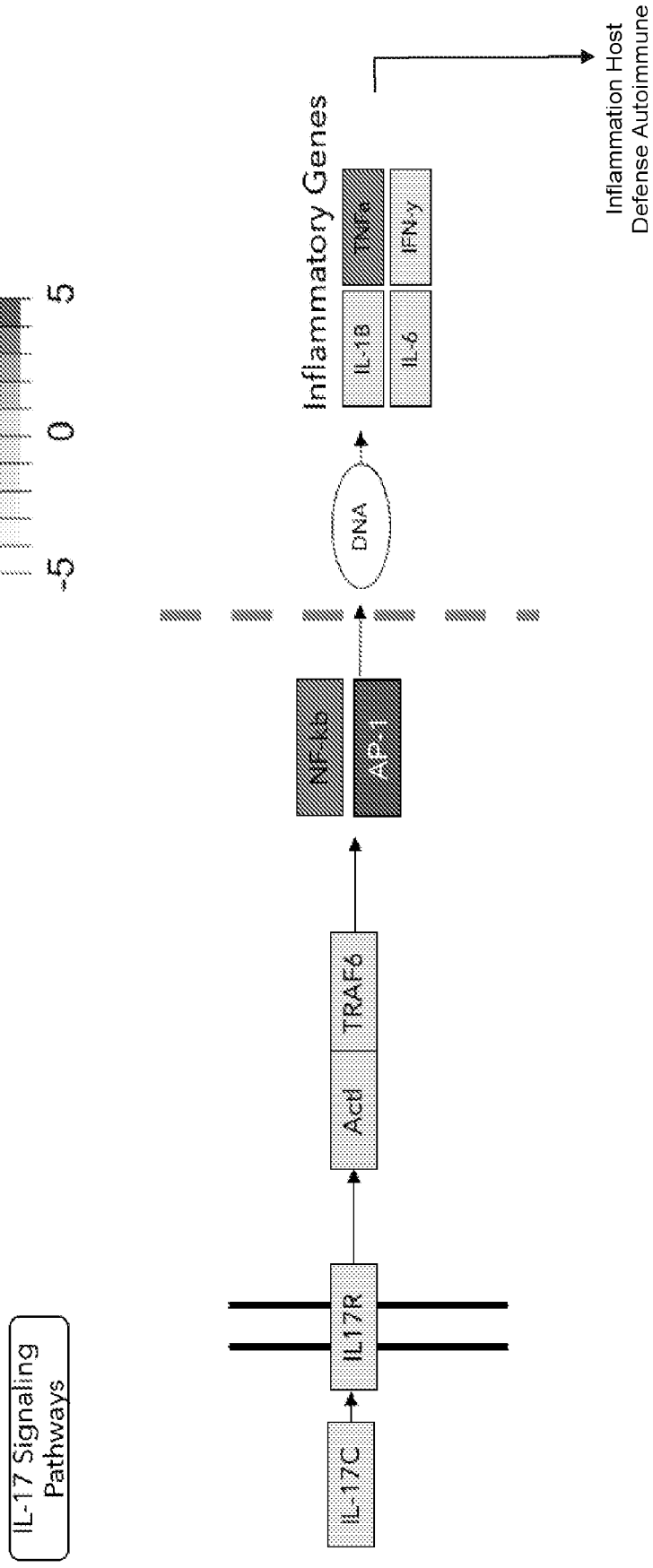

IL-17 Signaling Pathways
IL17C
IL17R
Act1
TRAF6
NF-κB
AP-1
DNA
Inflammatory Genes
IL-18
IL-6
TNFα
IFN-γ
Inflammation Host Defense Autoimmune
5
0
-5

FIG. 28B (cont.)

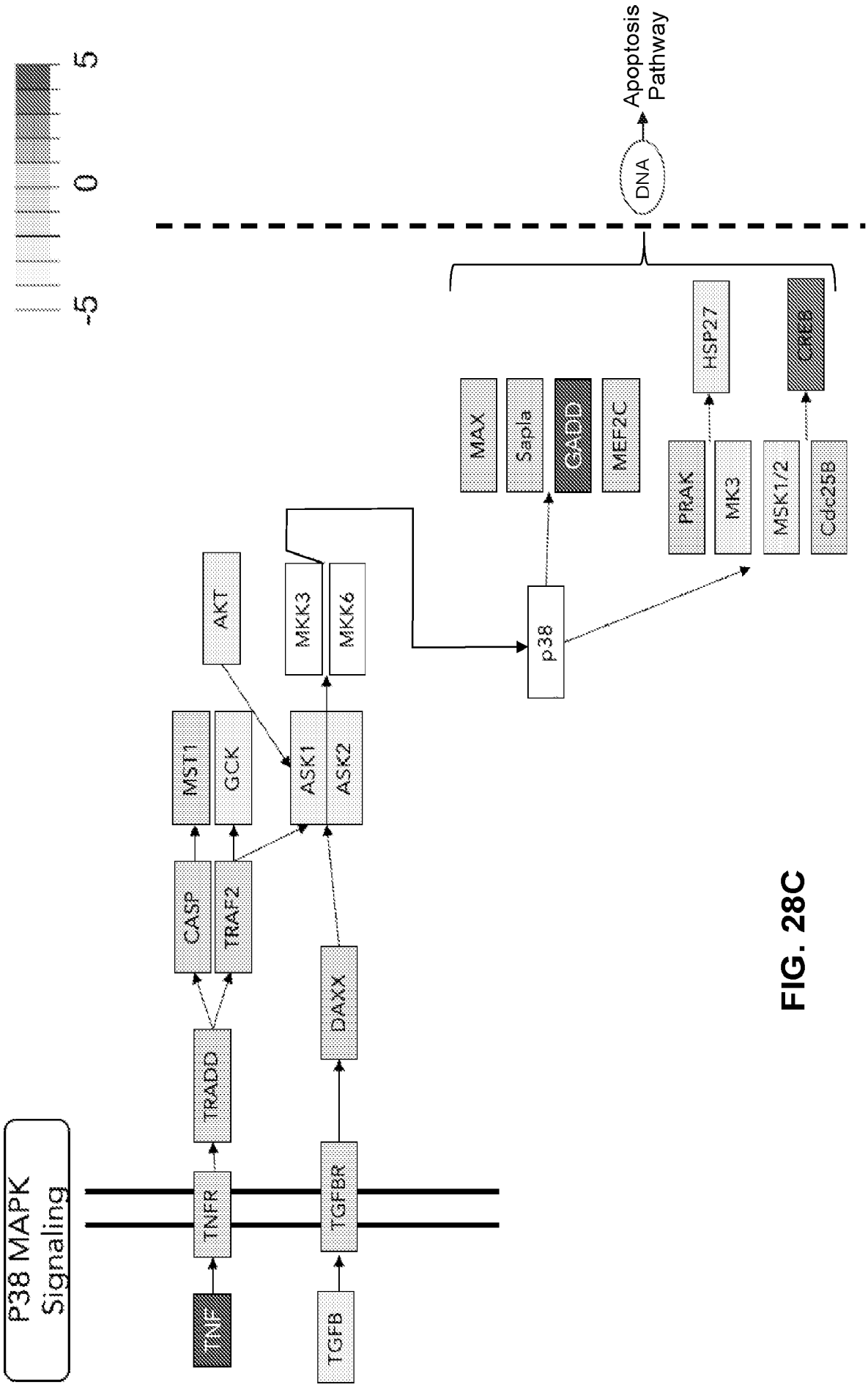

P38 MAPK Signaling
TNF
TGFB
TNFR
TGFBR
TRADD
DAXX
CASP
TRAF2
MST1
GCK
ASK1
ASK2
AKT
MKK3
MKK6
p38
MAX
Sapla
GADD
MEF2C
PRAK
MK3
MSK1/2
Cdc25B
HSP27
CREB
DNA
Apoptosis Pathway
5
0
-5

FIG. 28C

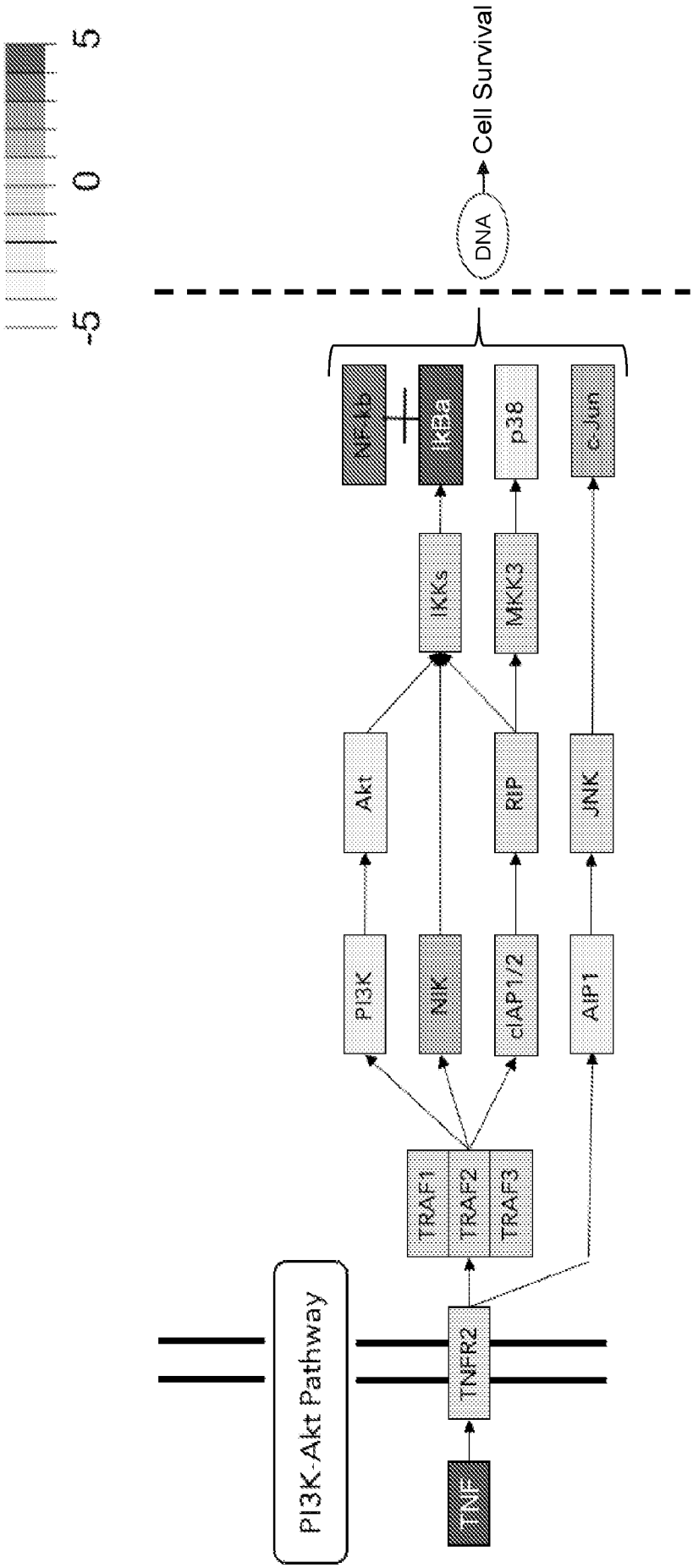

5
0
-5
Cell Survival
DNA
NF-kb
IKBa
p38
c-Jun
IKKs
MKK3
Akt
RIP
JNK
PI3K
NIK
cIAP1/2
AIP1
TRAF1
TRAF2
TRAF3
TNFR2
TNF
PI3K-Akt Pathway

FIG. 28C (cont.)

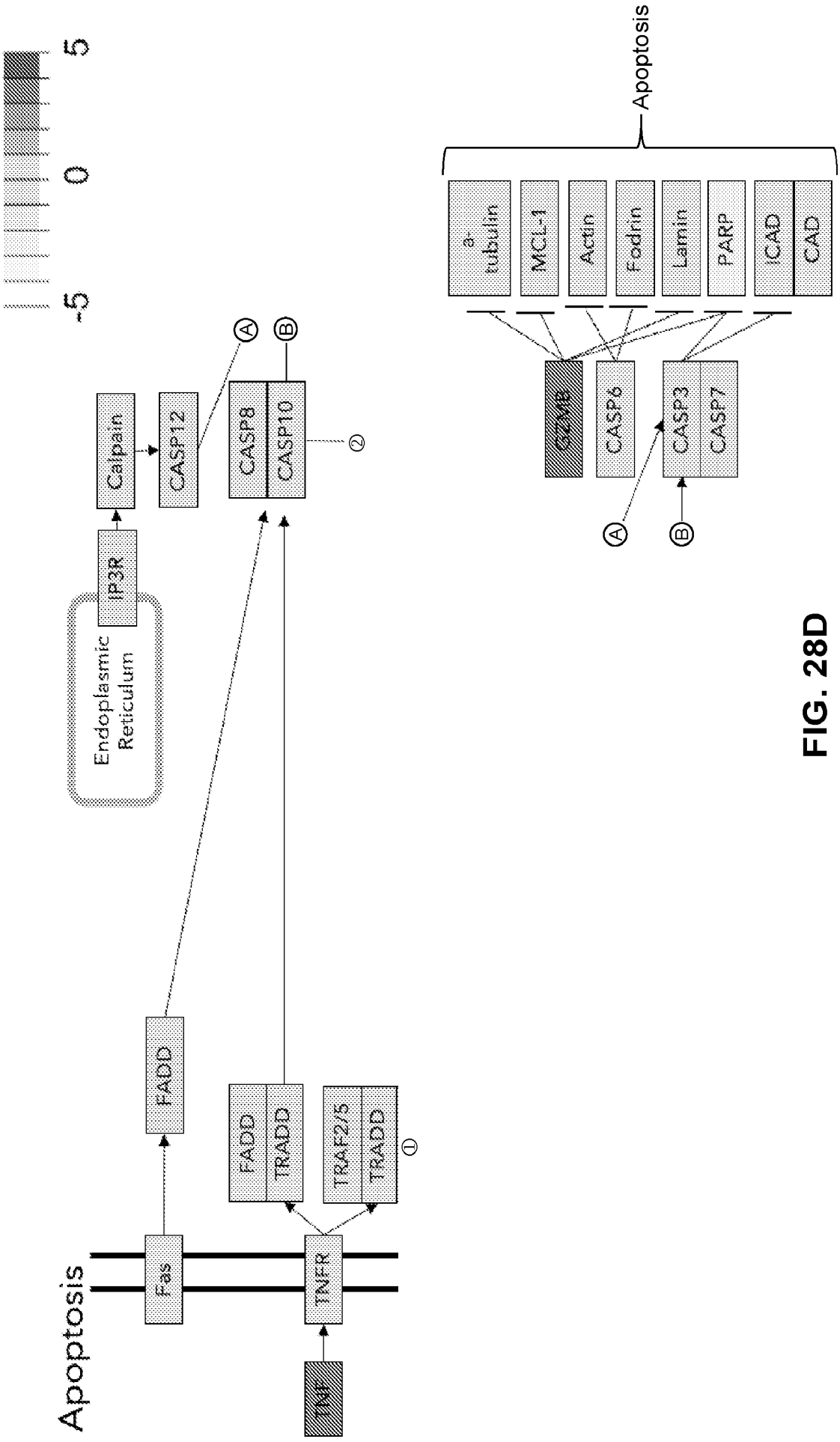

5
0
-5
Apoptosis
a-tubulin
MCL-1
Actin
Fodrin
Lamin
PARP
ICAD
CAD
GZMB
CASP6
CASP3
CASP7
A
B
Calpain
CASP12
CASP8
CASP10
2
IP3R
Endoplasmic Reticulum
FADD
FADD
TRADD
TRAF2/5
TRADD
1
Fas
TNFR
TNF
Apoptosis

FIG. 28D

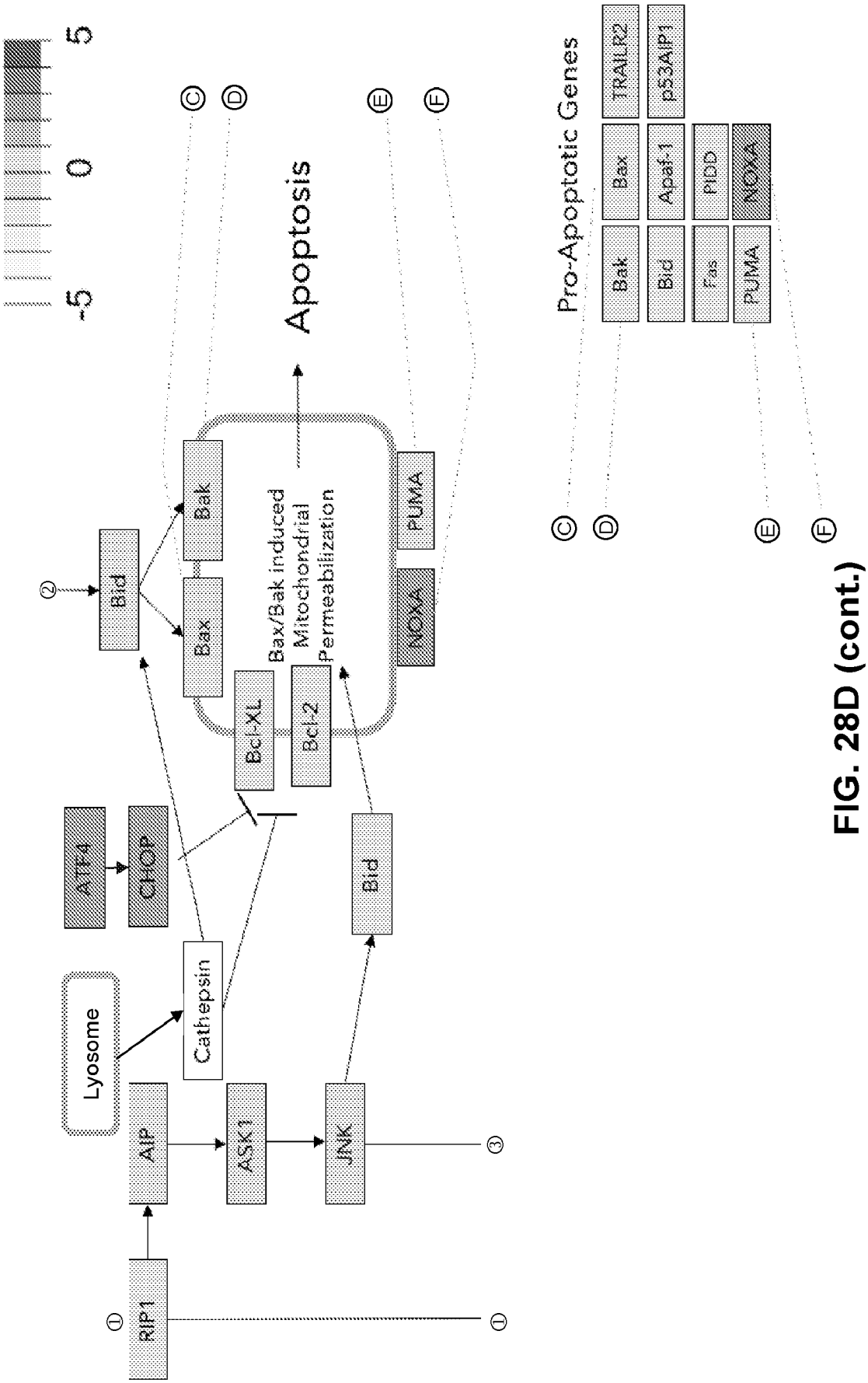

5
0
-5
Apoptosis
Pro-Apoptotic Genes
TRAILR2
p53AIP1
Bax
Apaf-1
Bak
Bid
PIDD
Fas
NOXA
PUMA
Bak
Bax
Bid
PUMA
NOXA
Bax/Bak induced
Mitochondrial
Permeabilization
Bcl-XL
Bcl-2
ATF4
CHOP
Lysosome
Cathepsin
AIP
ASK1
JNK
RIP1
①
②
③
Ⓒ
Ⓓ
Ⓔ
Ⓕ

FIG. 28D (cont.)

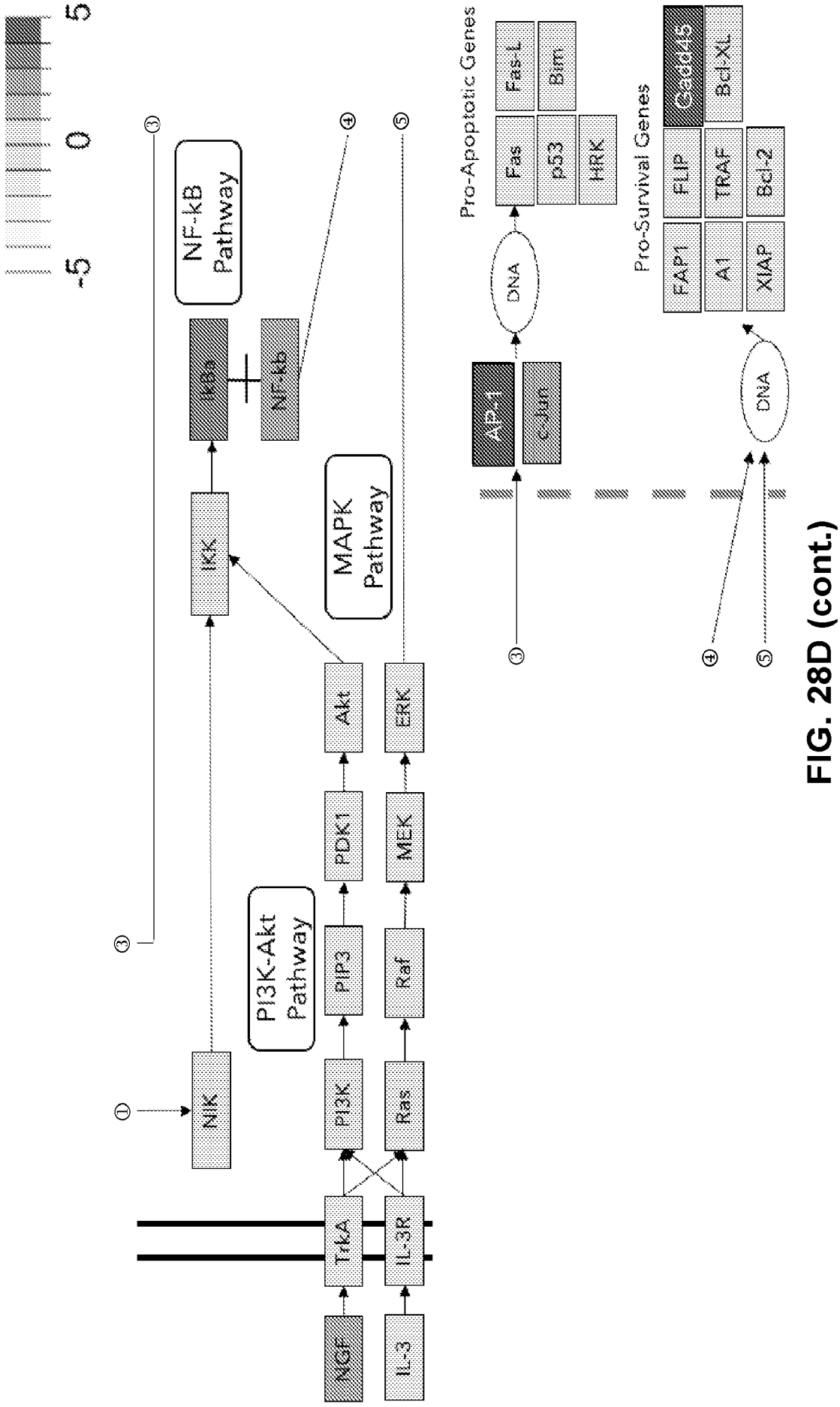

5
0
-5
NF-kB Pathway
MAPK Pathway
PI3K-Akt Pathway
NIK
IKK
IκBα
NF-kb
PI3K
Ras
PIP3
Raf
PDK1
MEK
Akt
ERK
TrkA
IL-3R
NGF
IL-3
Pro-Apoptotic Genes
Fas
p53
HRK
Fas-L
Bim
DNA
AP-1
c-Jun
Pro-Survival Genes
FAP1
A1
XIAP
FLIP
TRAF
Bcl-2
Caspase8
Bcl-XL
DNA

FIG. 28D (cont.)

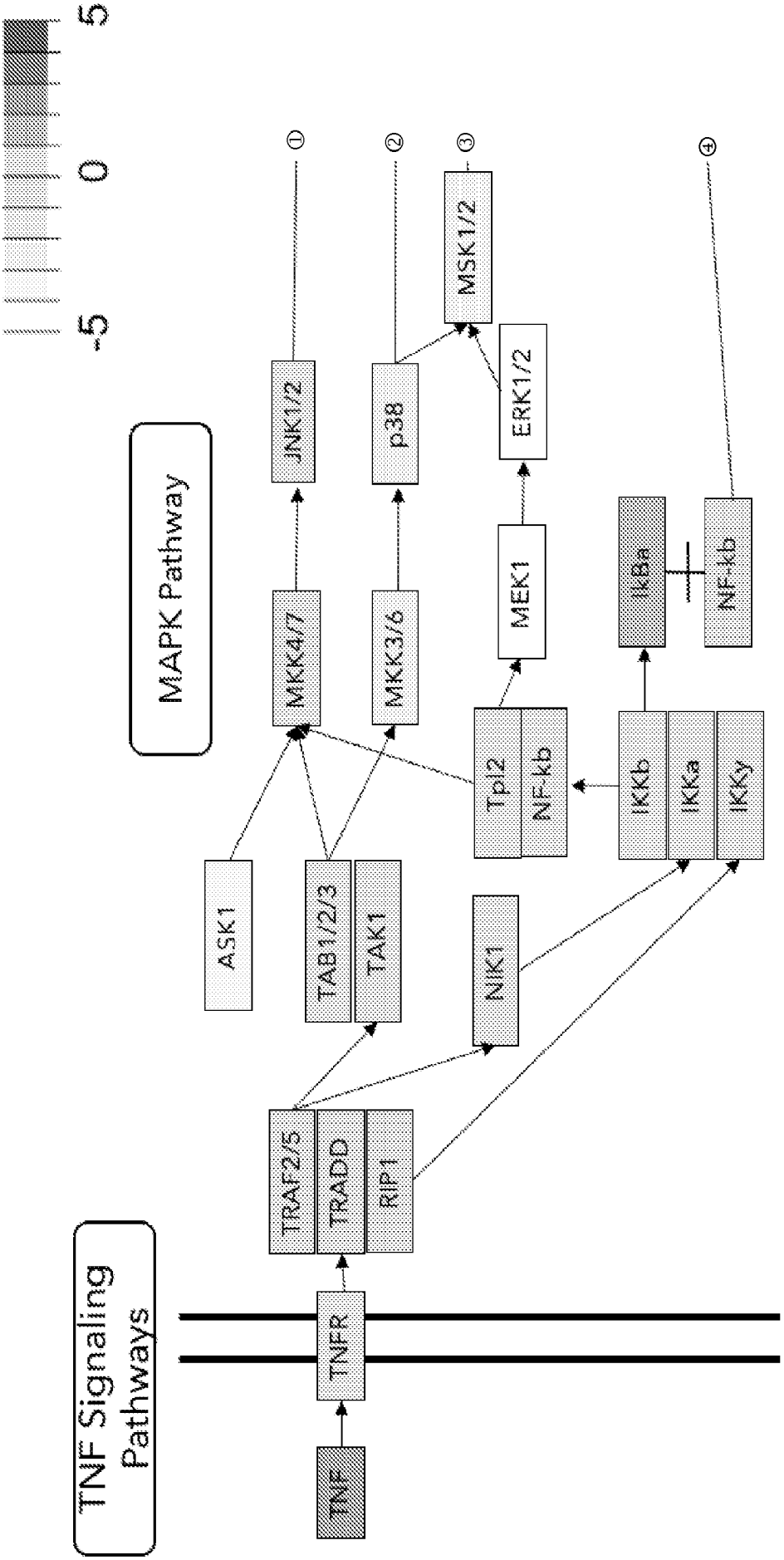

5
0
-5
TNF Signaling Pathways
TNF
TNFR
TRAF2/5
TRADD
RIP1
ASK1
TAB1/2/3
TAK1
NIK1
Tpl2
NF-kb
IKKb
IKKa
IKKy
IkBa
NF-kb
MAPK Pathway
MKK4/7
MKK3/6
MEK1
JNK1/2
p38
ERK1/2
MSK1/2
①
②
③
④

FIG. 28E

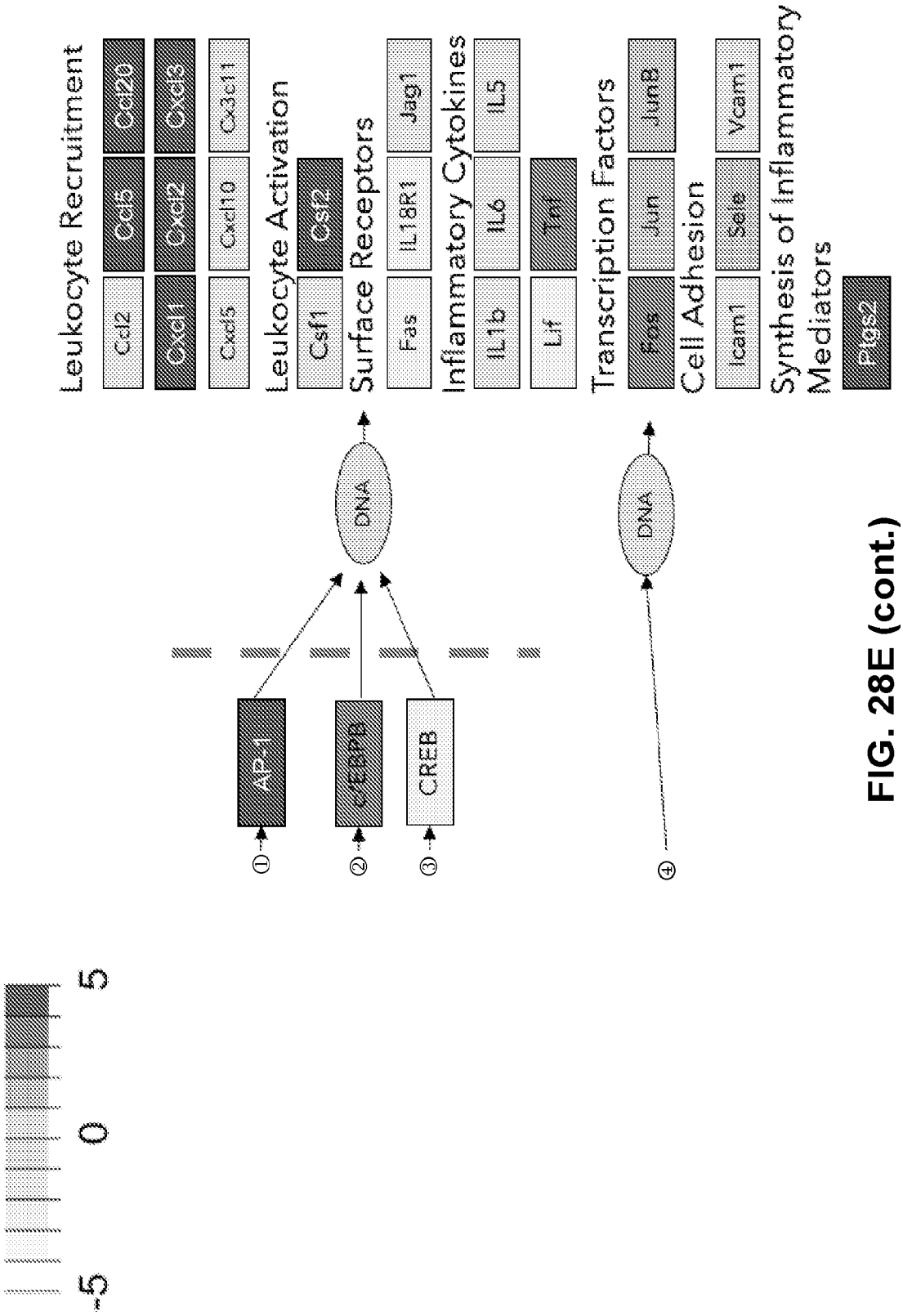

Leukocyte Recruitment
Ccl2
Ccl5
Ccl20
Cxcl1
Cxcl2
Cxcl3
Cxcl5
Cxcl10
Cx3cl1
Leukocyte Activation
Csf1
Csf2
Surface Receptors
Fas
IL18R1
Jag1
Inflammatory Cytokines
IL1b
IL6
IL5
Lif
Tnf
Transcription Factors
Fos
Jun
JunB
Cell Adhesion
Icam1
Sele
Vcam1
Synthesis of Inflammatory Mediators
Ptgs2
DNA
DNA
AP-1
C/EBPB
CREB
①
②
③
④
5
0
-5

FIG. 28E (cont.)

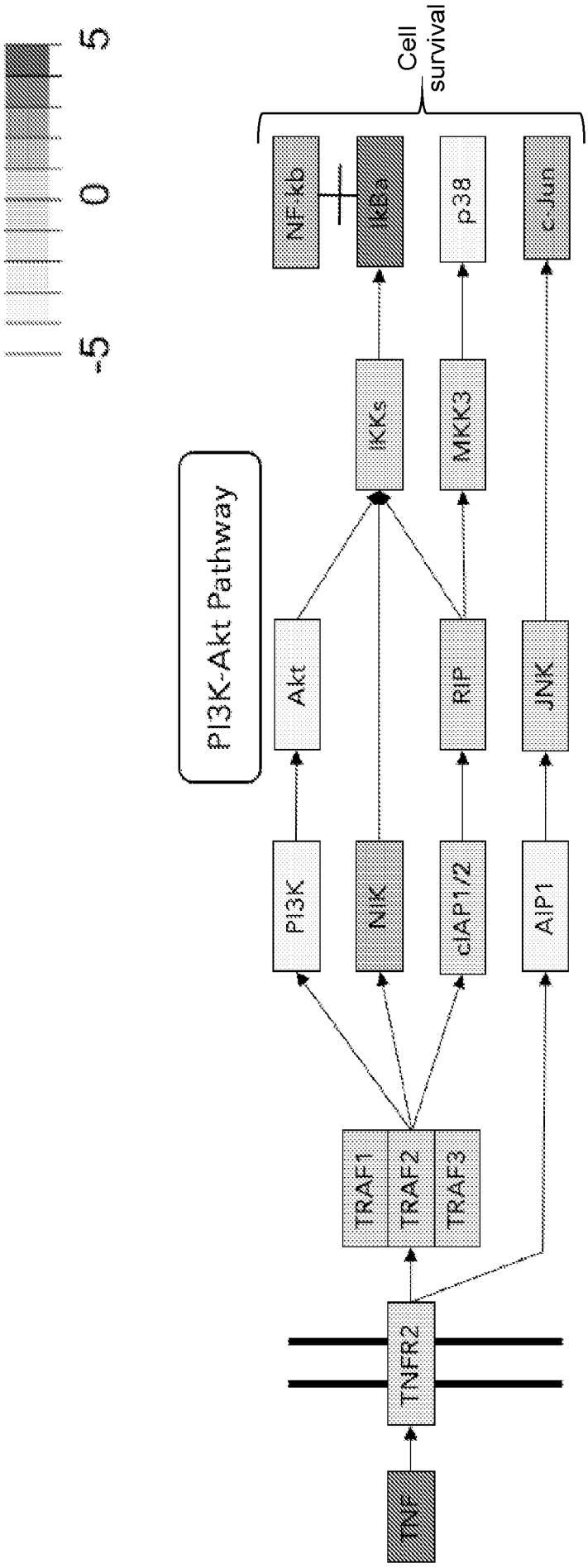

5
0
-5
Cell survival
NF-kb
IκBα
p38
c-Jun
IKKs
MKK3
PI3K-Akt Pathway
Akt
RIP
JNK
PI3K
NIK
cIAP1/2
A1P1
TRAF1
TRAF2
TRAF3
TNFR2
TNF

FIG. 28E (cont.)

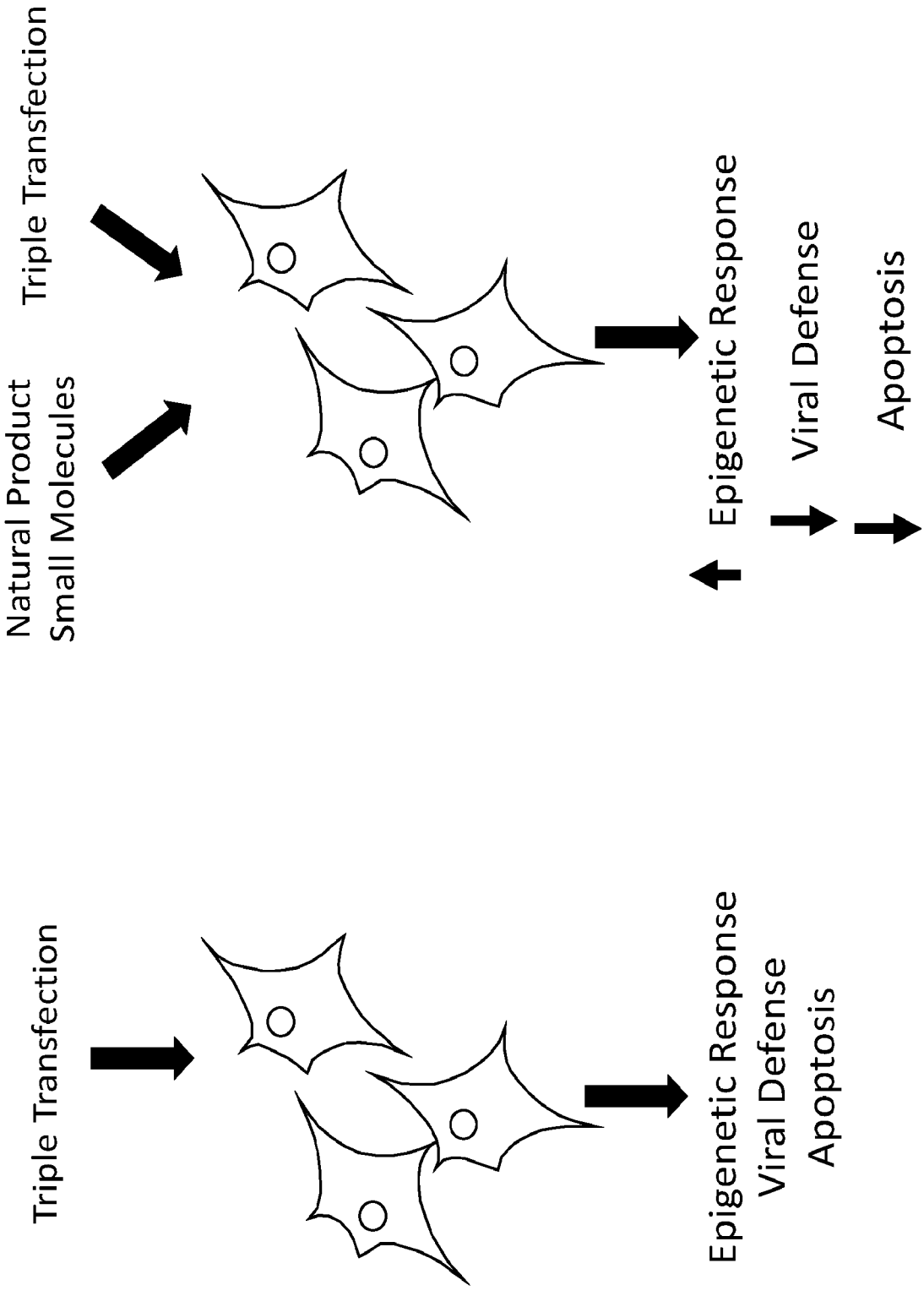

Triple Transfection
Natural Product
Small Molecules
Epigenetic Response
Viral Defense
Apoptosis
Triple Transfection
Epigenetic Response
Viral Defense
Apoptosis

FIG. 29

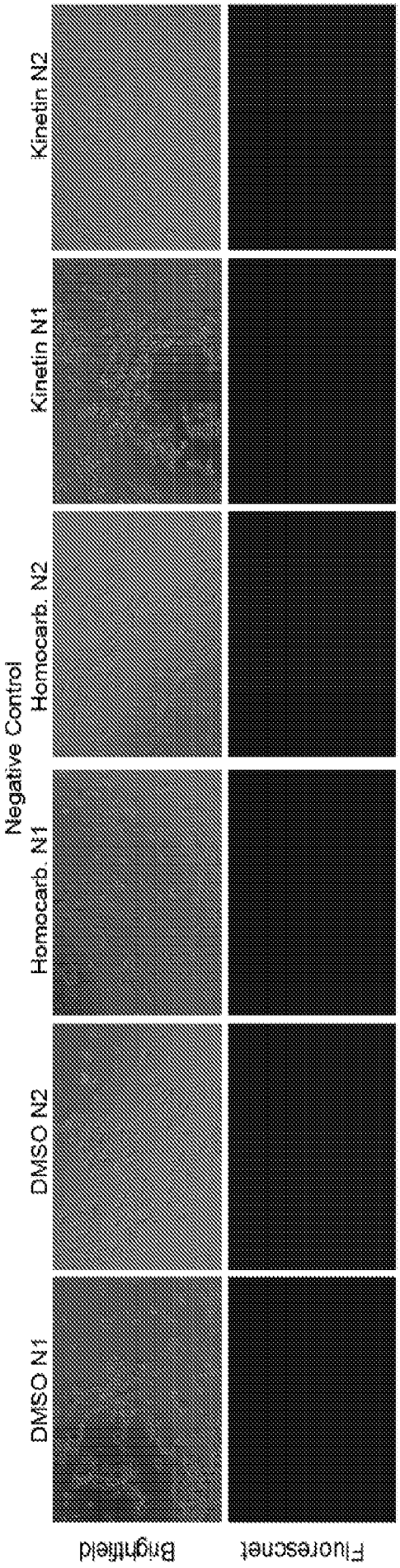


FIG. 30A

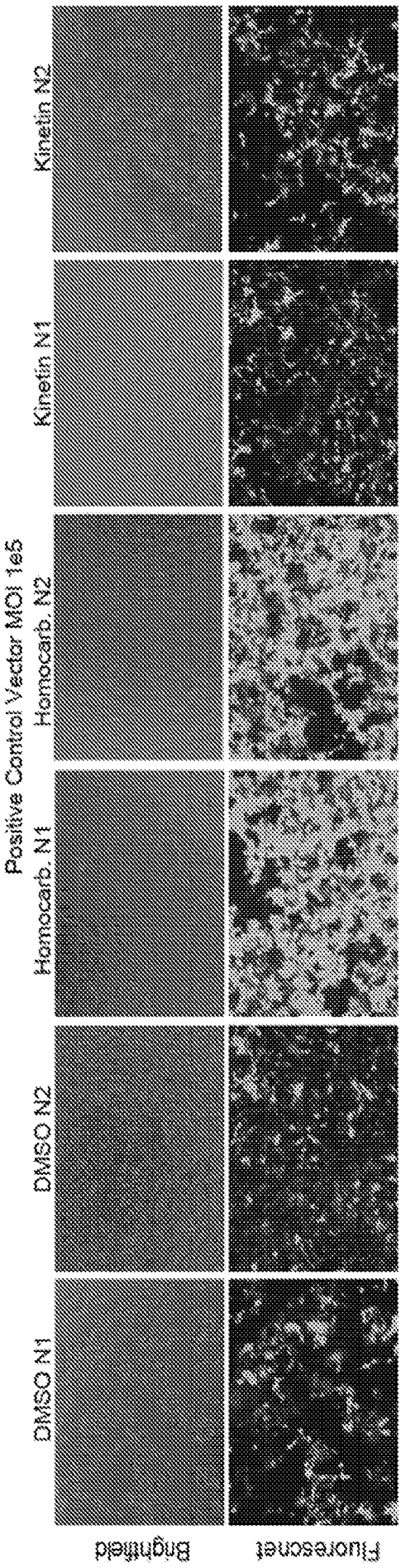


FIG. 30B

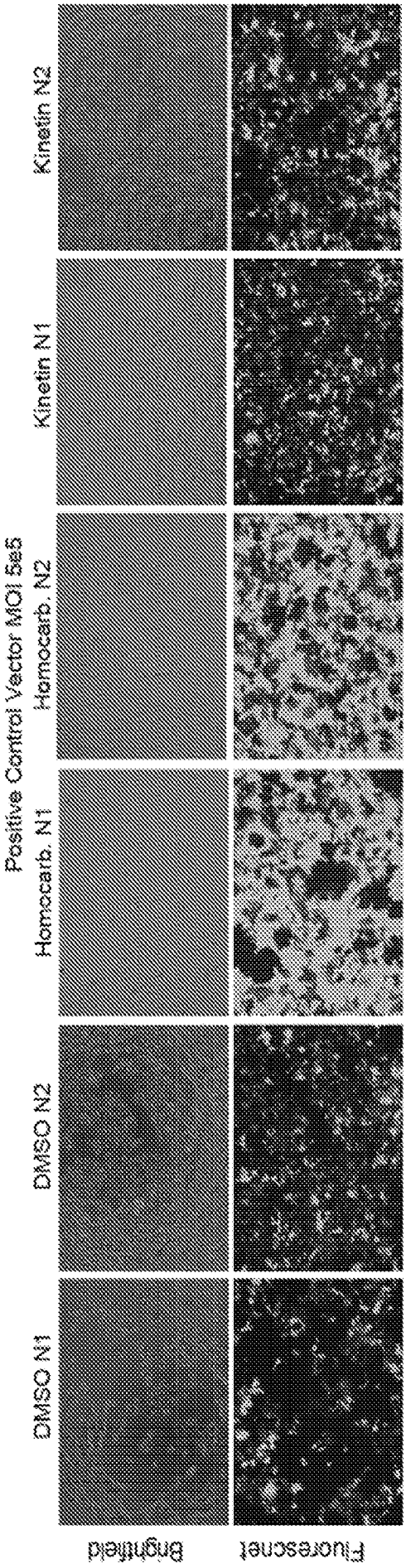

Positive Control Vector MOI 5e5
DMSO N1
DMSO N2
Homocarb. N1
Homocarb. N2
Kinetin N1
Kinetin N2
Fluorescnet
Brightfield

FIG. 30C

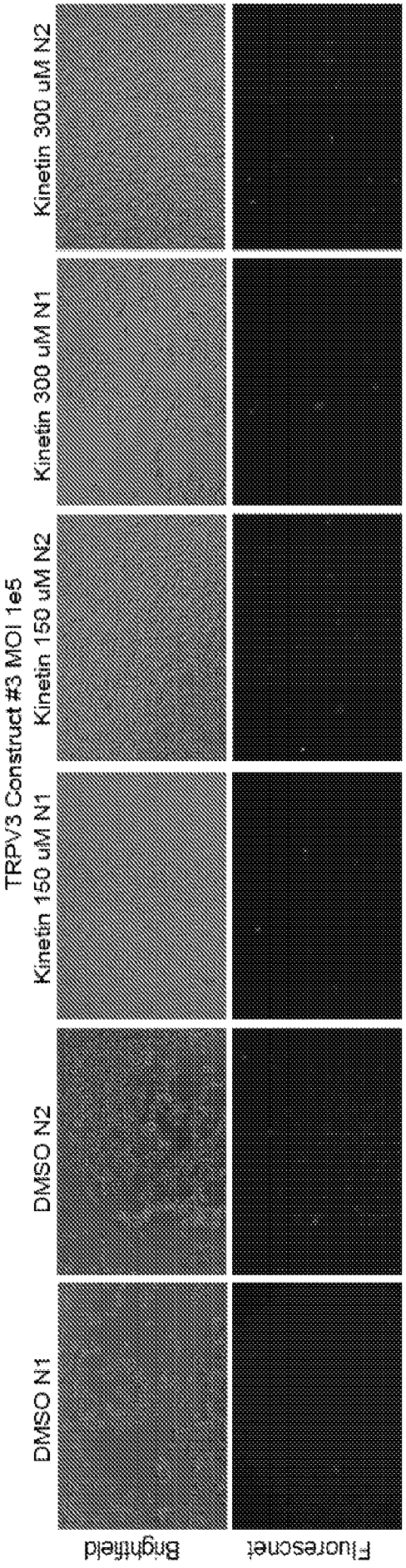


FIG. 31A

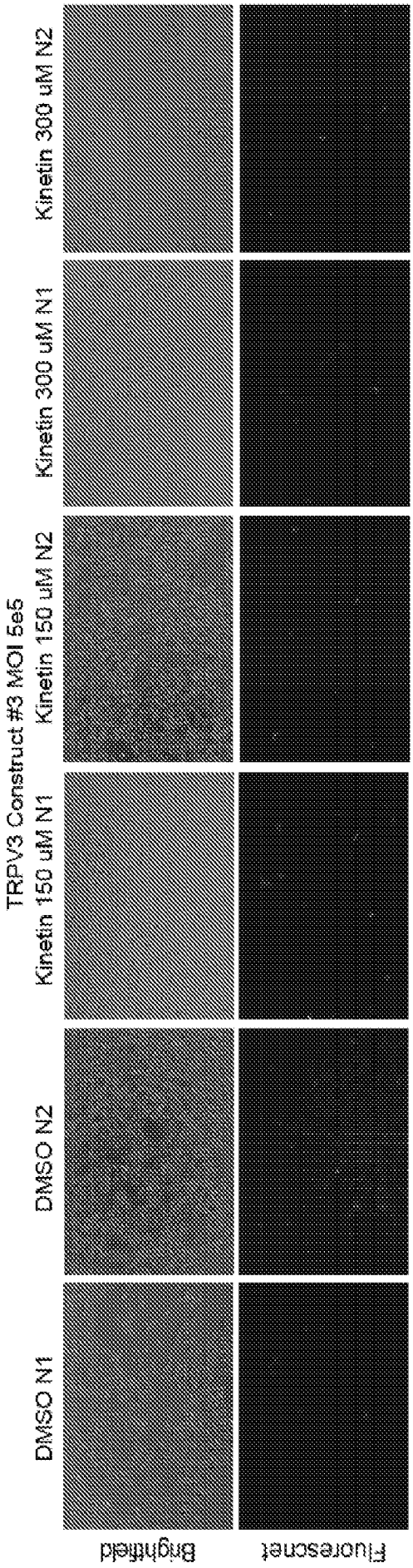


FIG. 31B

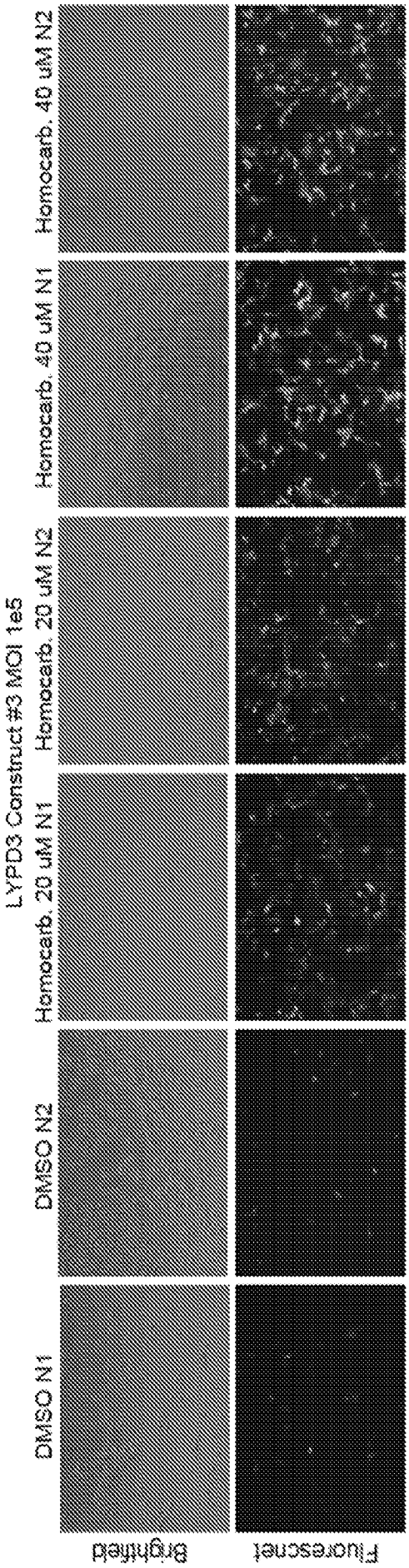


FIG. 31C

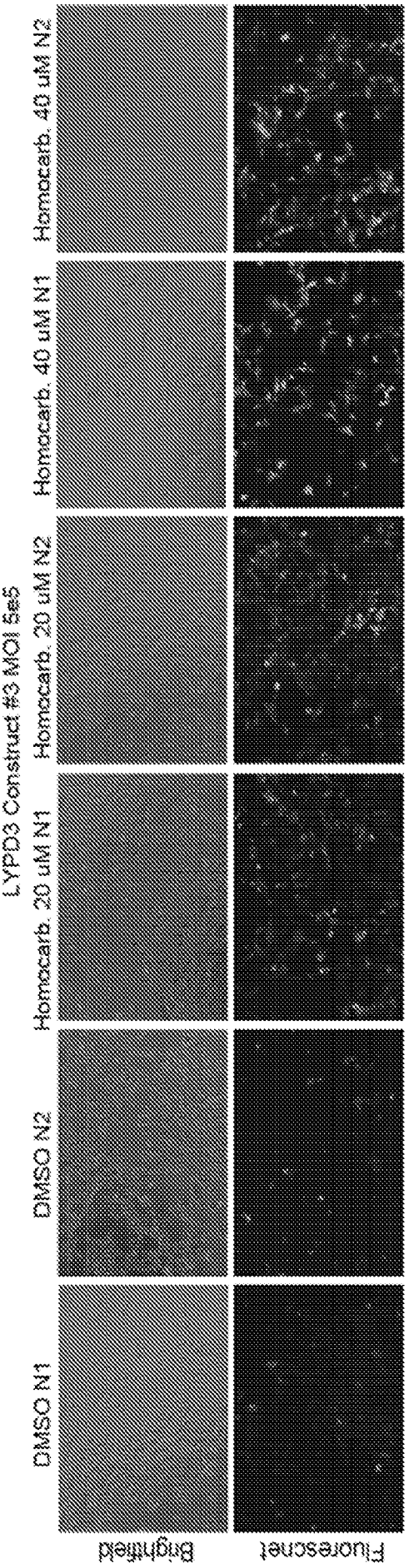


FIG. 31D

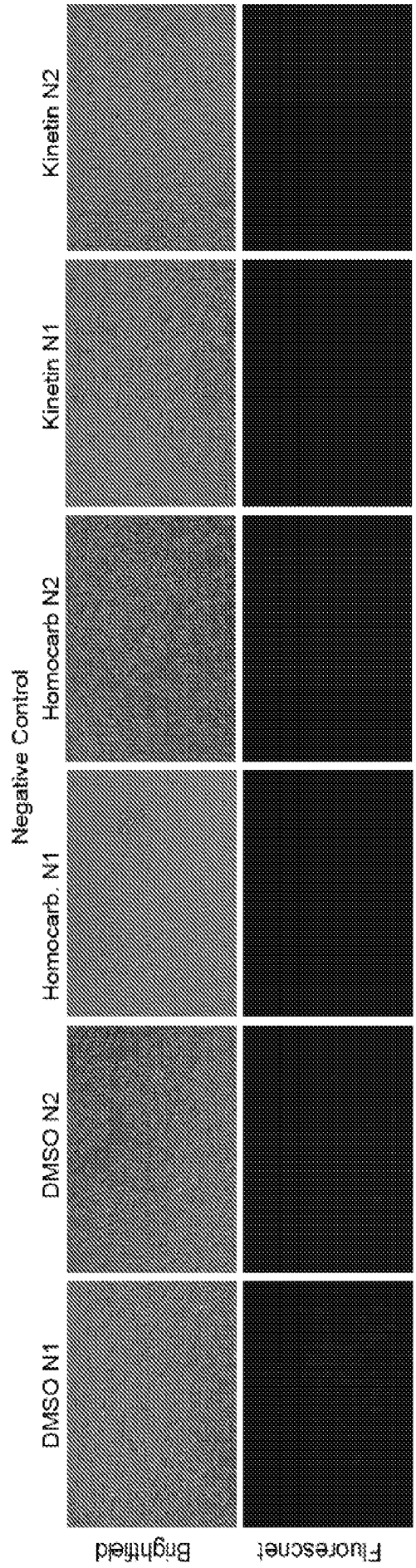


FIG. 32A

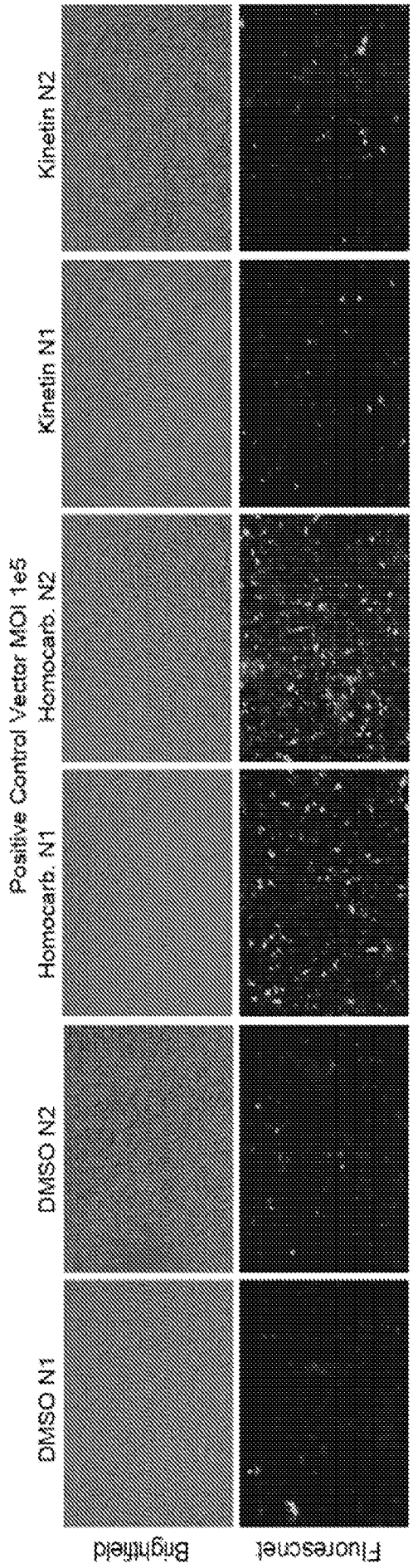


FIG. 32B

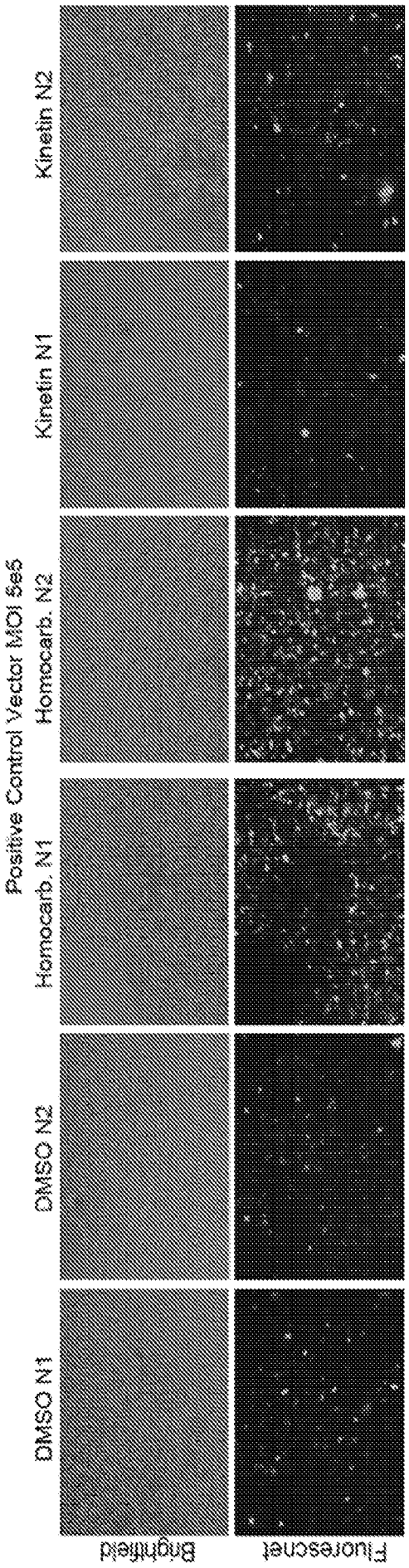

Positive Control Vector MOI 5e5
DMSO N1
DMSO N2
Homocarb. N1
Homocarb. N2
Kinetin N1
Kinetin N2
Brightfield
Fluorescnet

FIG. 32C

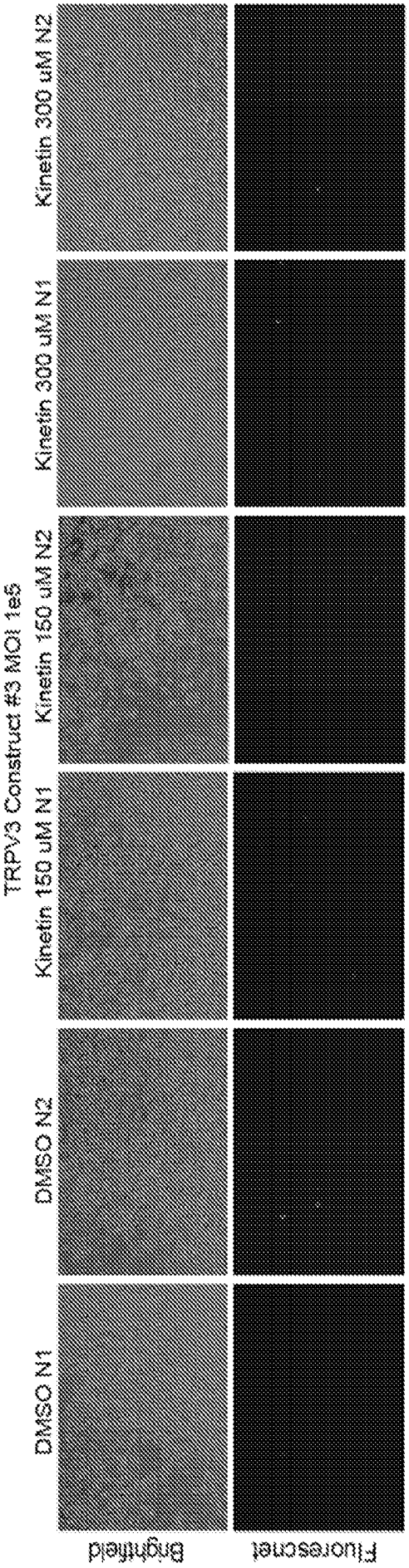


FIG. 33A

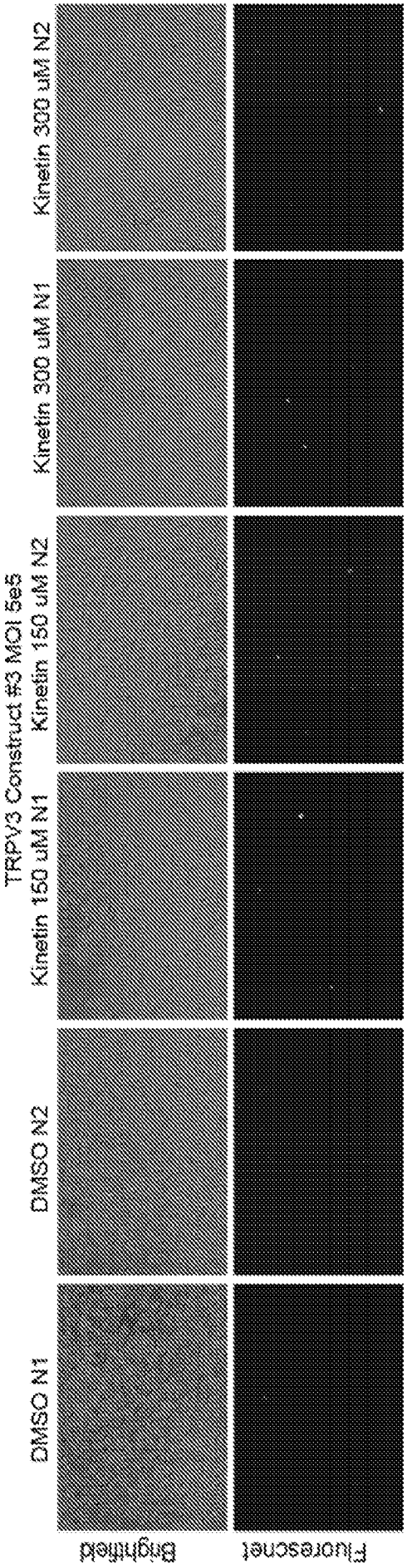


FIG. 33B

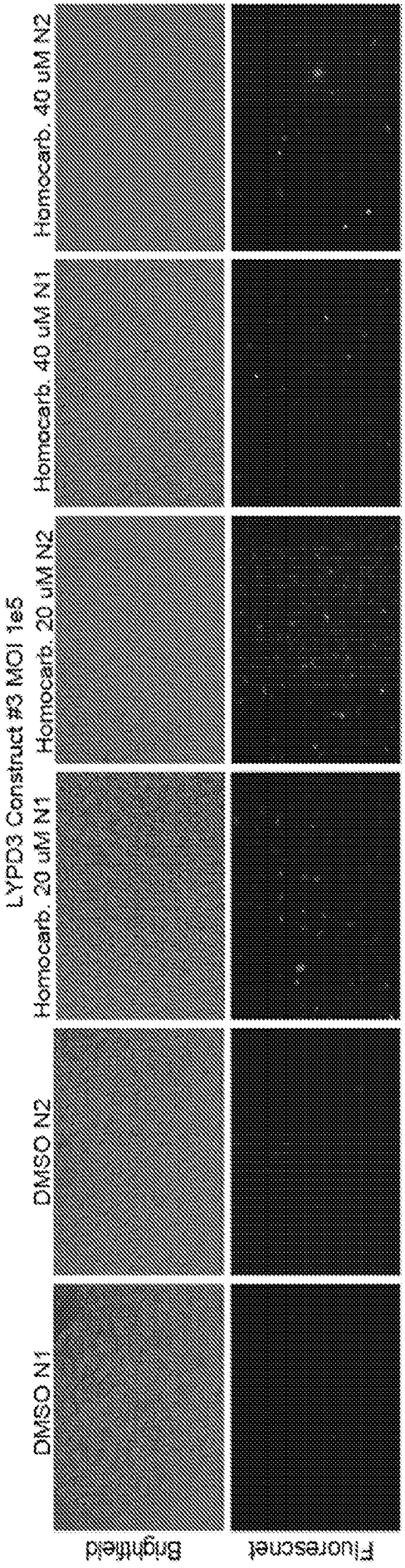


FIG. 33C

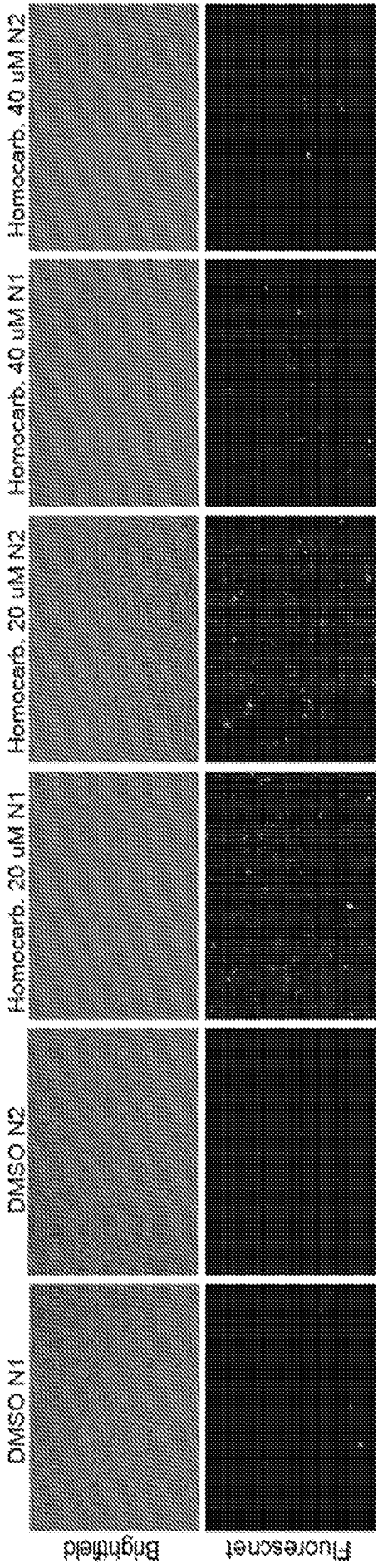


FIG. 33D

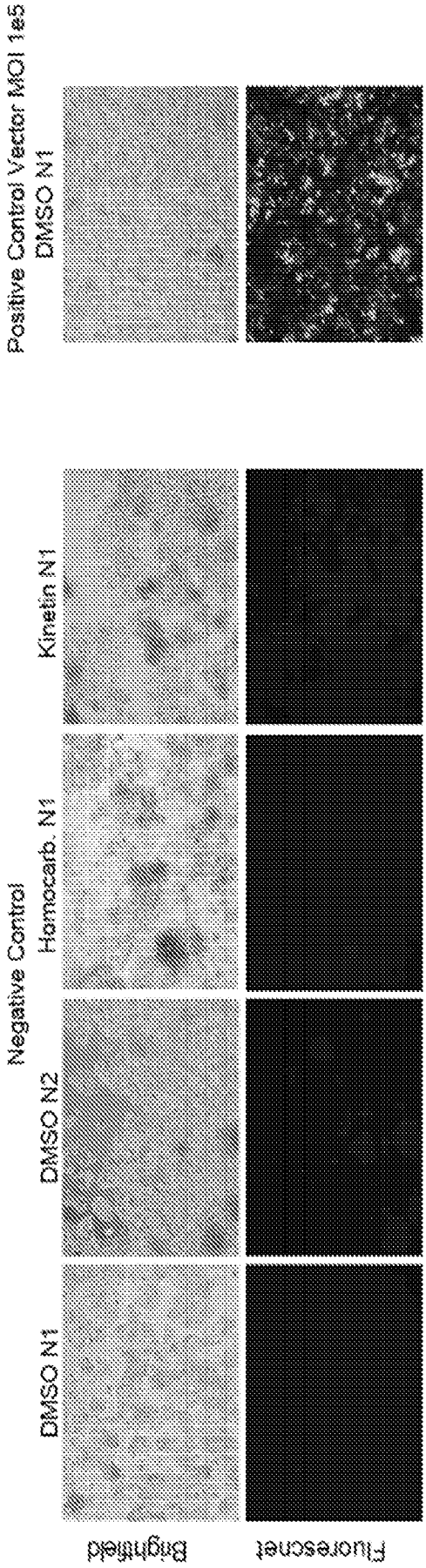


FIG. 34A

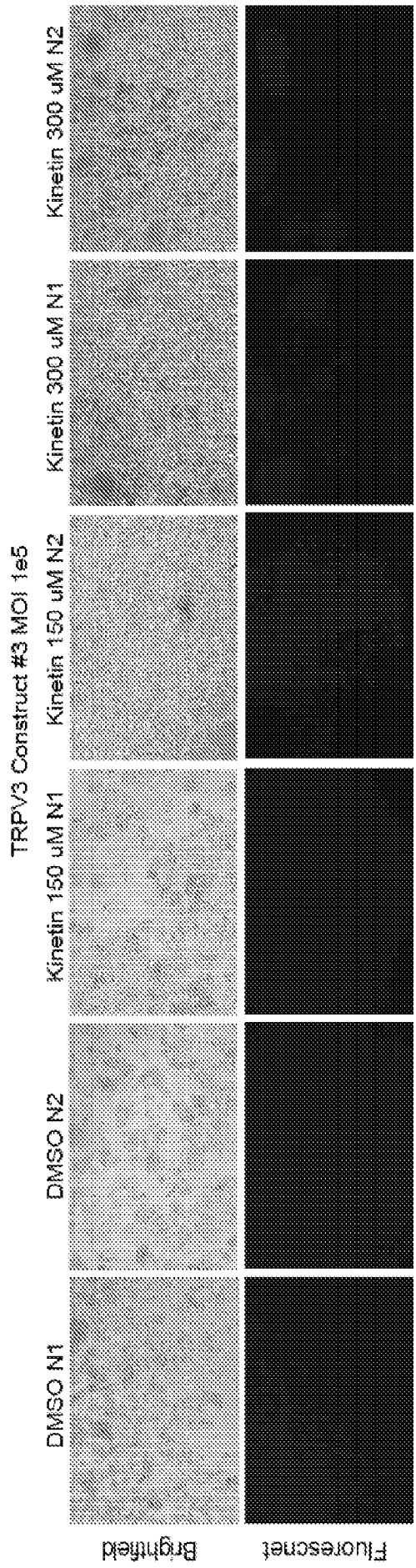


FIG. 34B

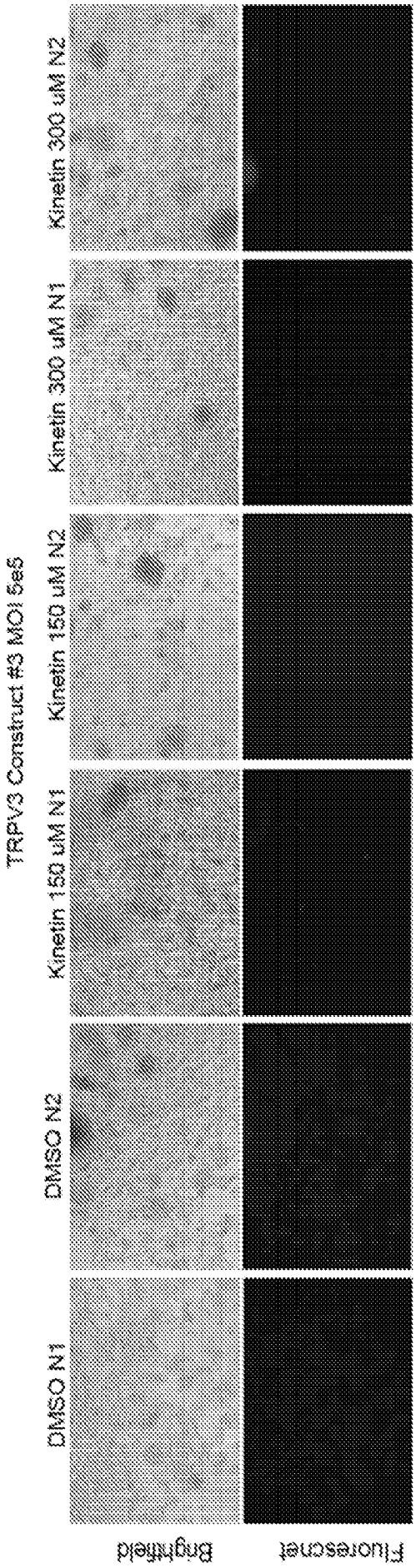

TRPV3 Construct #3 MOI 5e5
DMSO N1
DMSO N2
Kinetin 150 uM N1
Kinetin 150 uM N2
Kinetin 300 uM N1
Kinetin 300 uM N2
Brightfield
Fluorescnet

FIG. 34C

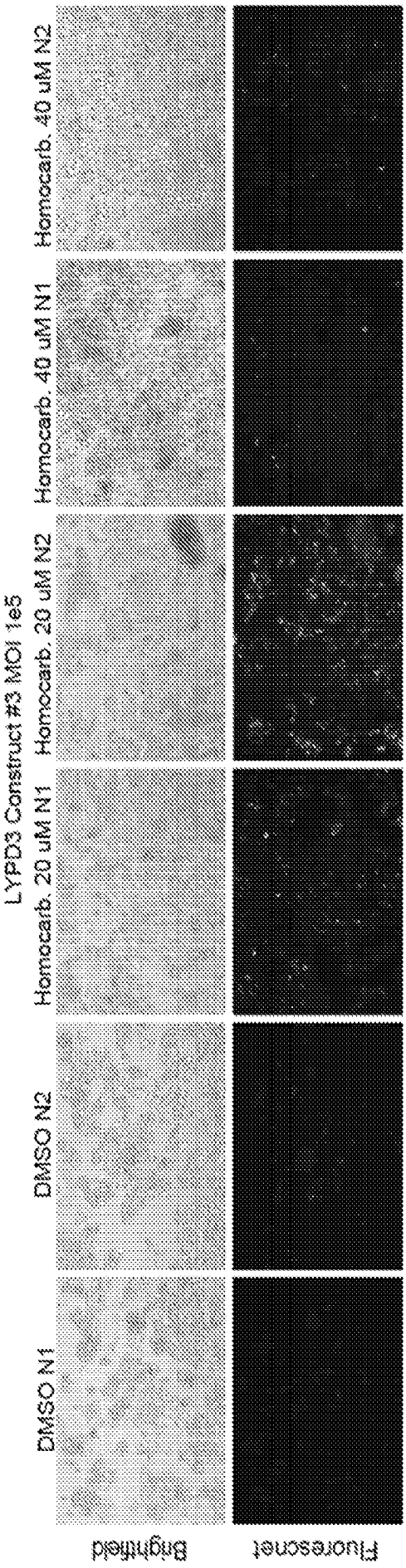


FIG. 35A

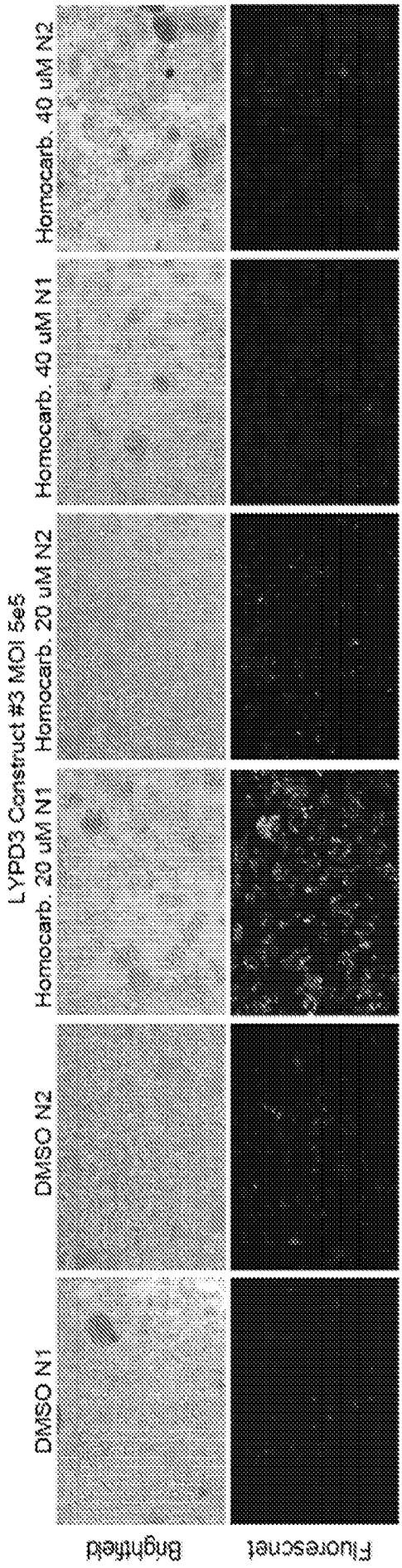


FIG. 35B

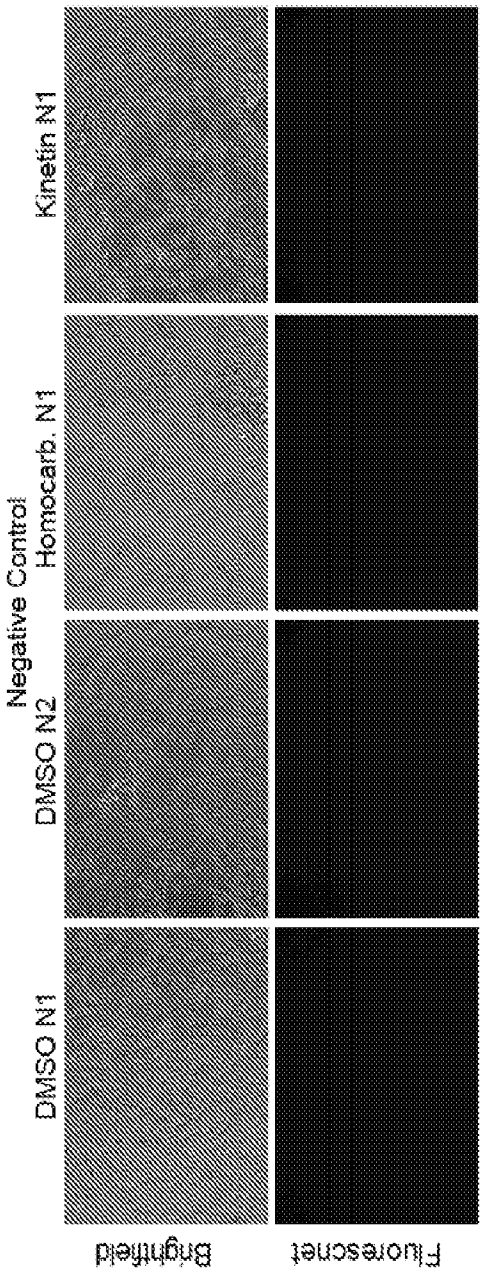


FIG. 36A

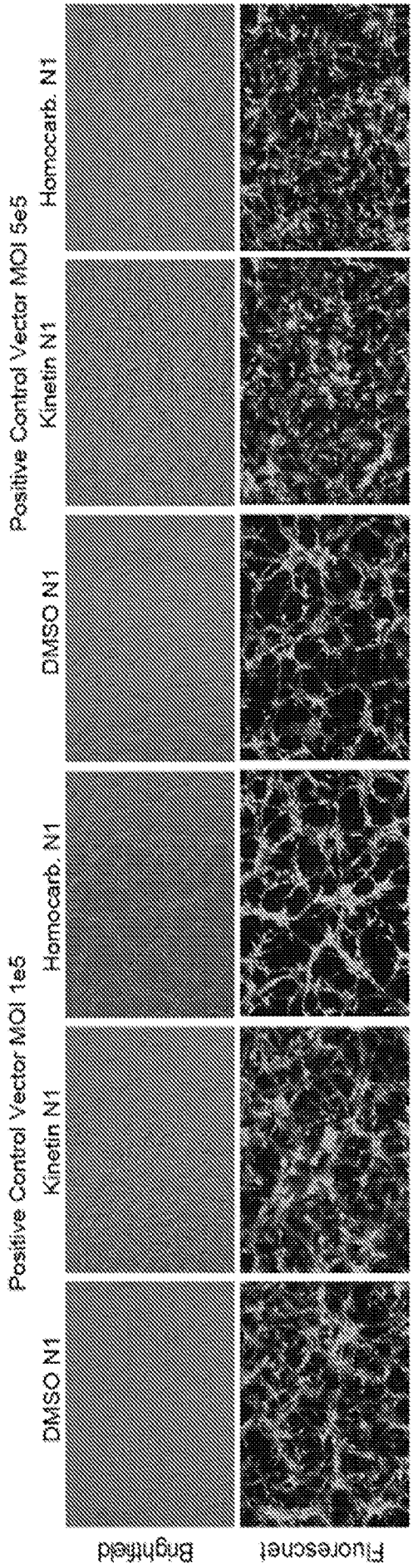


FIG. 36B

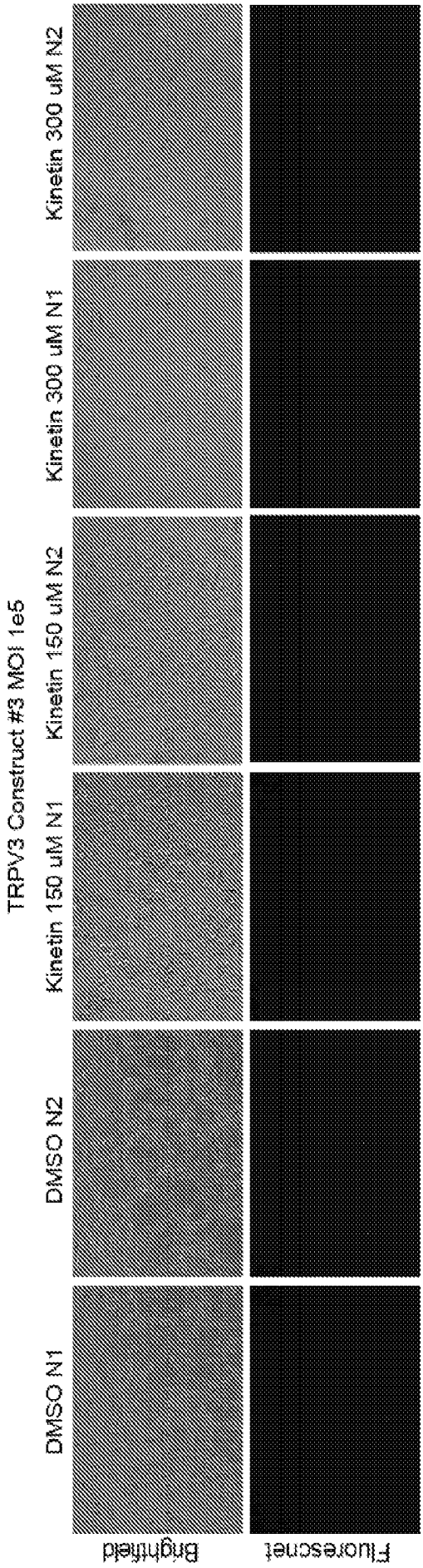


FIG. 37A

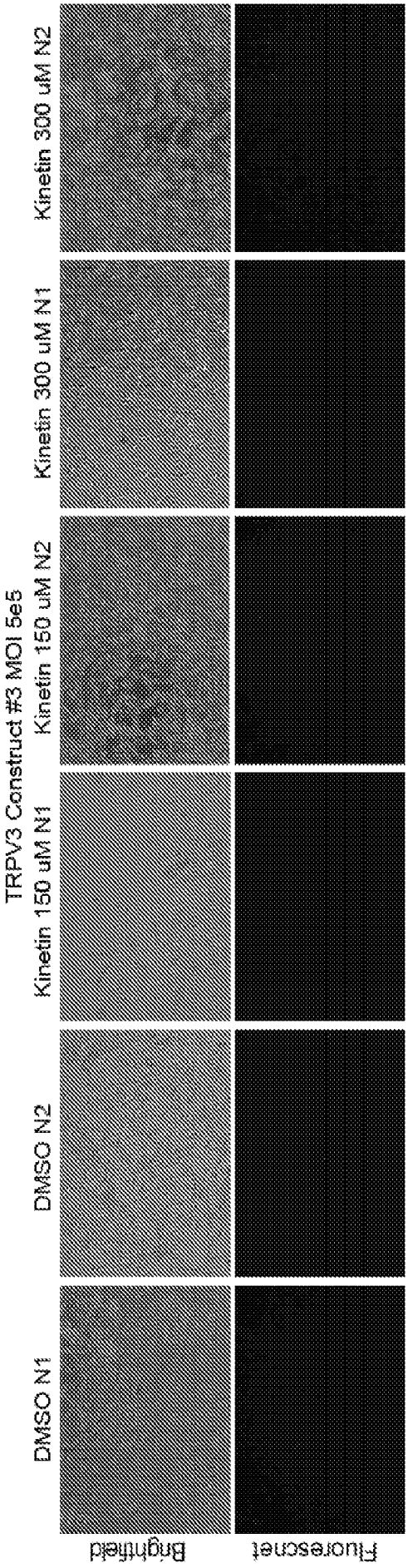


FIG. 37B

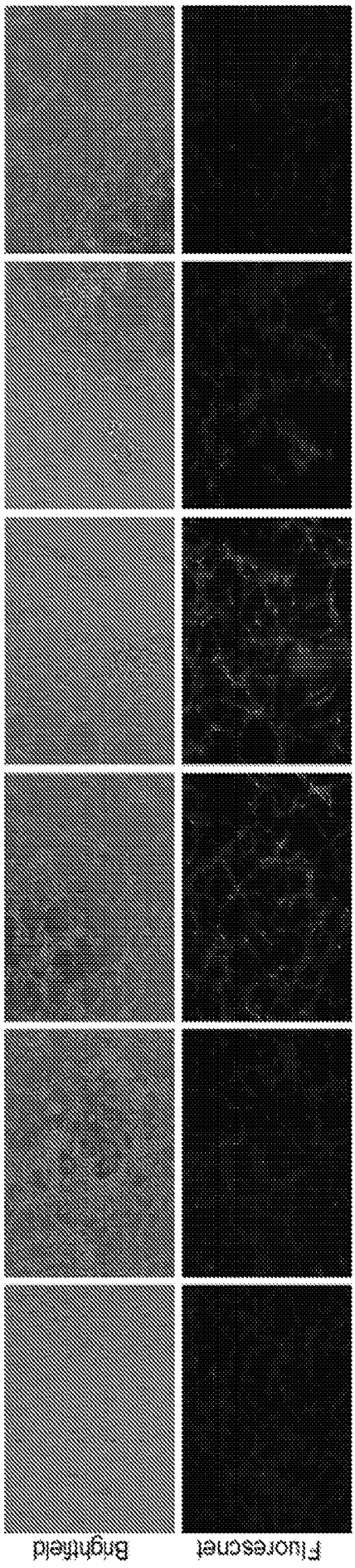


FIG. 37C

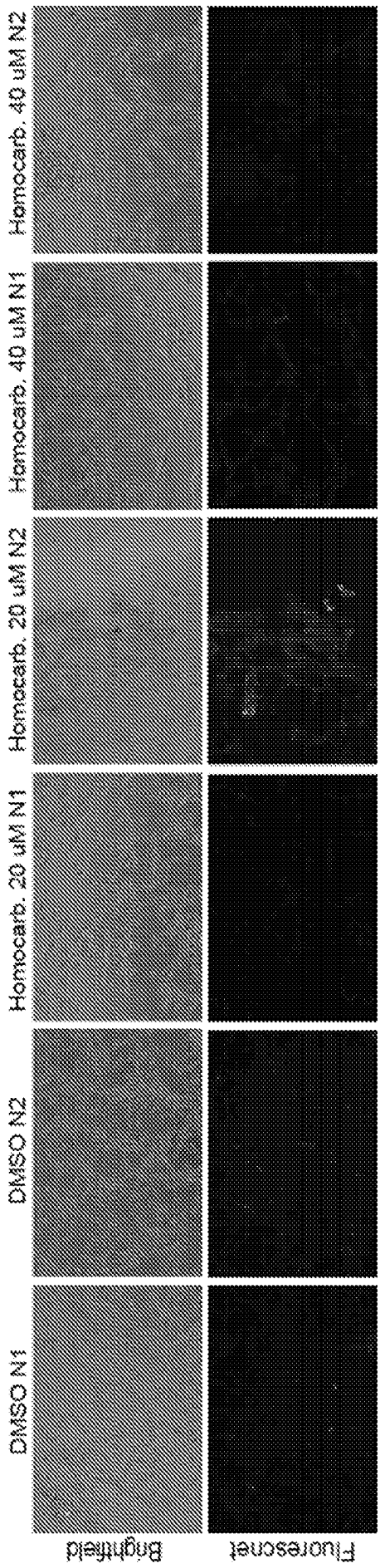


FIG. 37D

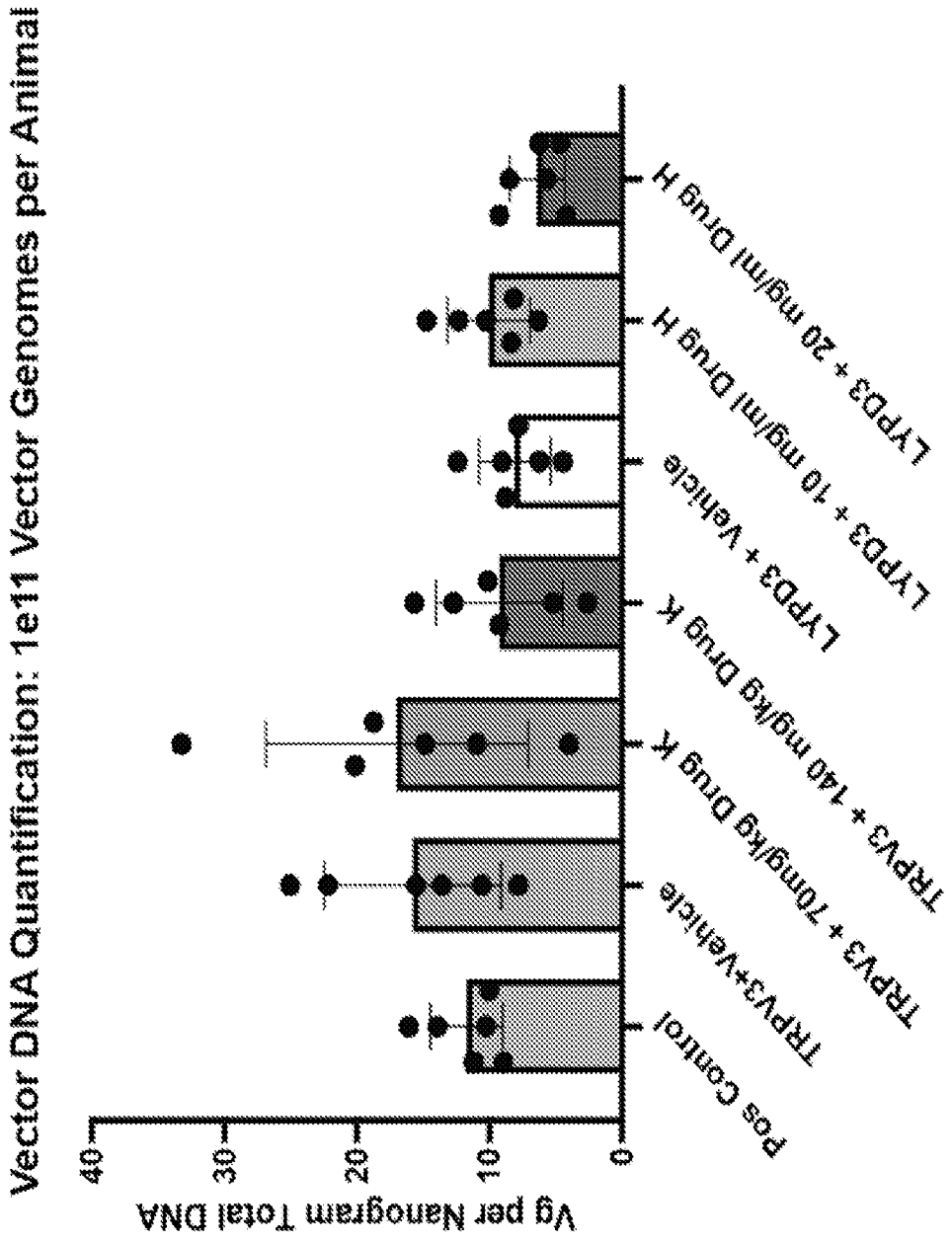

Vector DNA Quantification: 1e11 Vector Genomes per Animal
Vg per Nanogram Total DNA
0
10
20
30
40
Pos Control
TRPV3+Vehicle
TRPV3 + 70mg/kg Drug K
TRPV3 + 140 mg/kg Drug K
LYPD3 + Vehicle
LYPD3 + 10 mg/ml Drug H
LYPD3 + 20 mg/ml Drug H

FIG. 38A

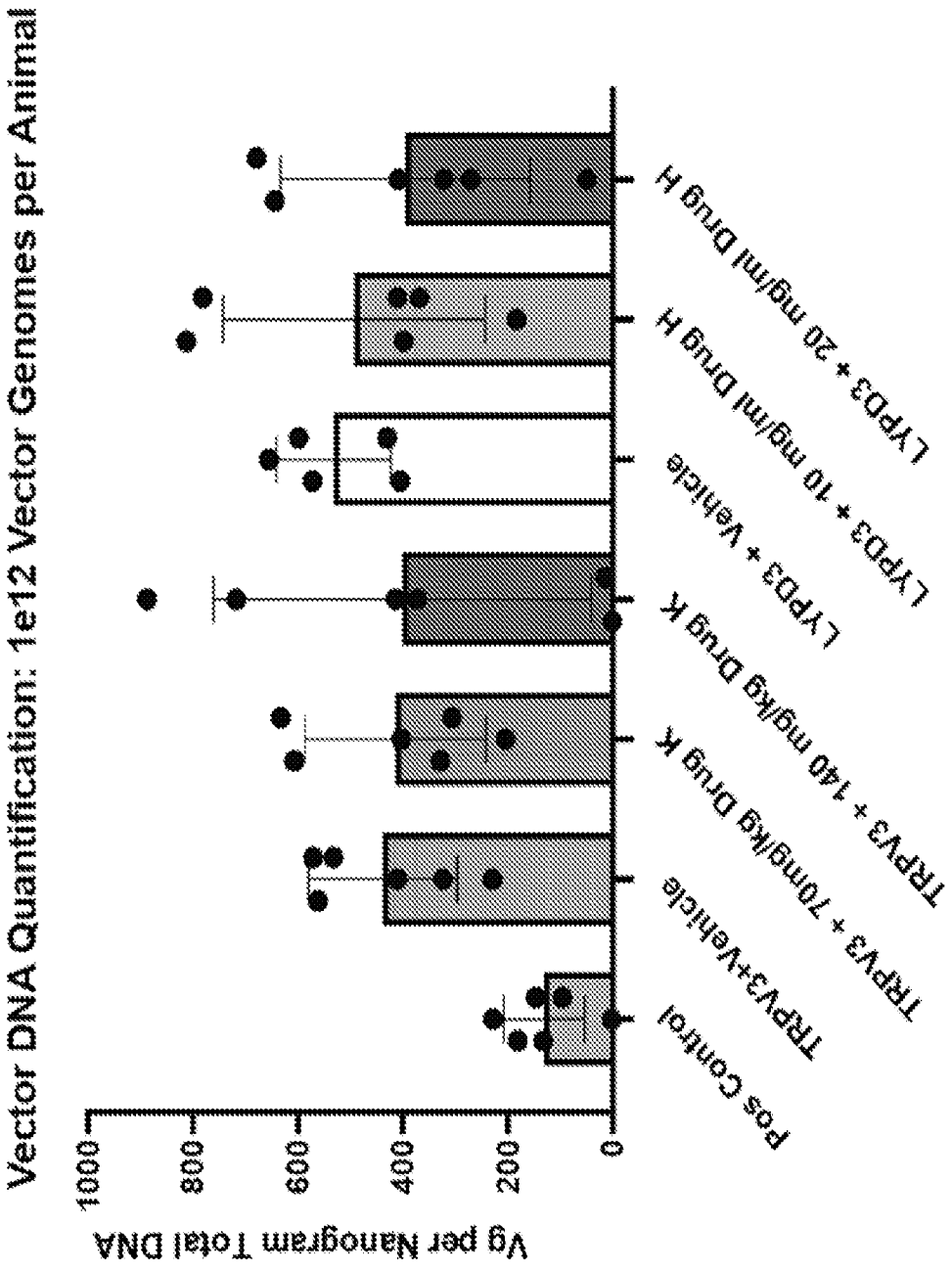

Vector DNA Quantification: 1e12 Vector Genomes per Animal
Vg per Nanogram Total DNA
1000
800
600
400
200
0
Pos Control
TRPV3+Vehicle
TRPV3 + 70mg/kg Drug K
TRPV3 + 140 mg/kg Drug K
LYPD3 + Vehicle
LYPD3 + 10 mg/ml Drug H
LYPD3 + 20 mg/ml Drug H

FIG. 38B

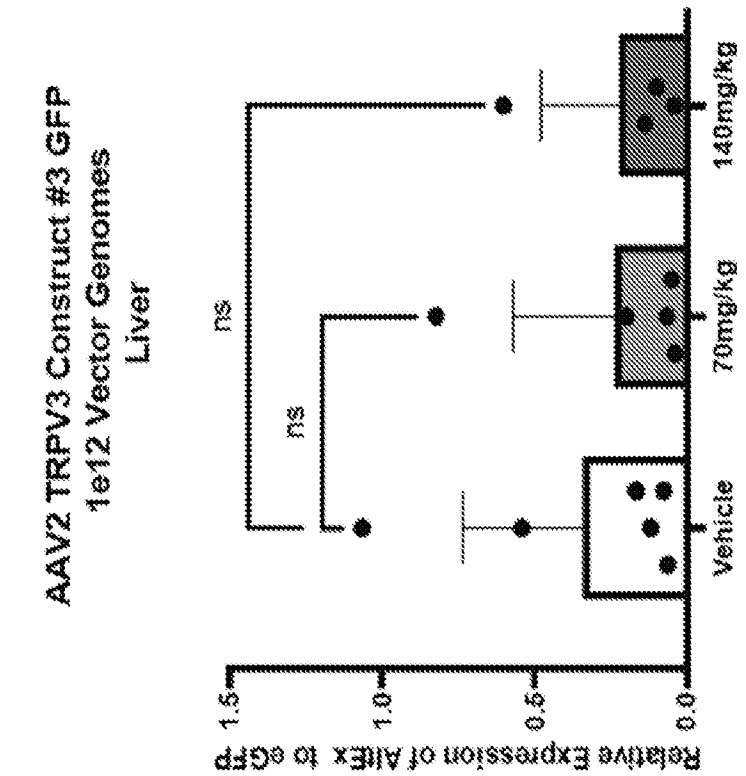


FIG. 39B

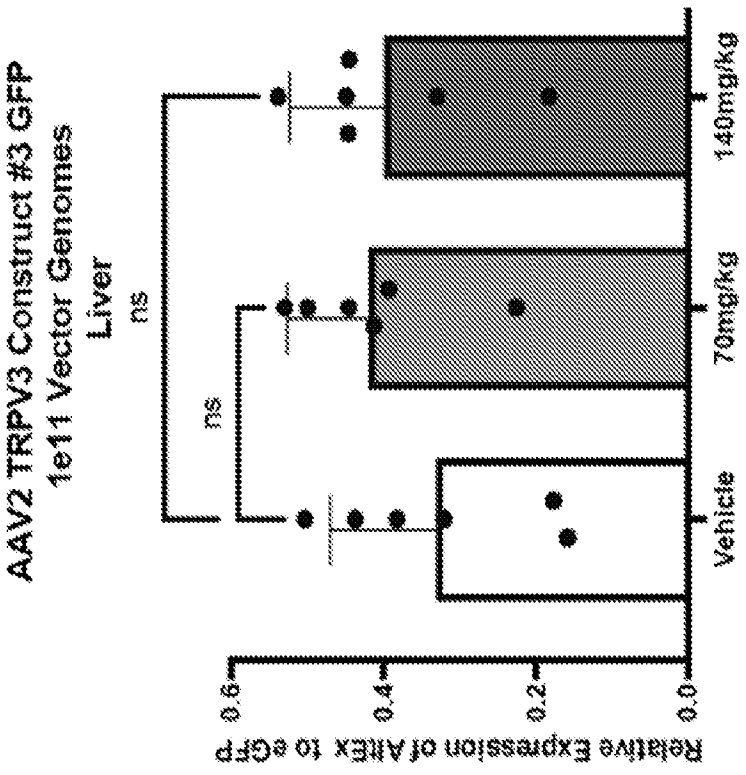


FIG. 39A

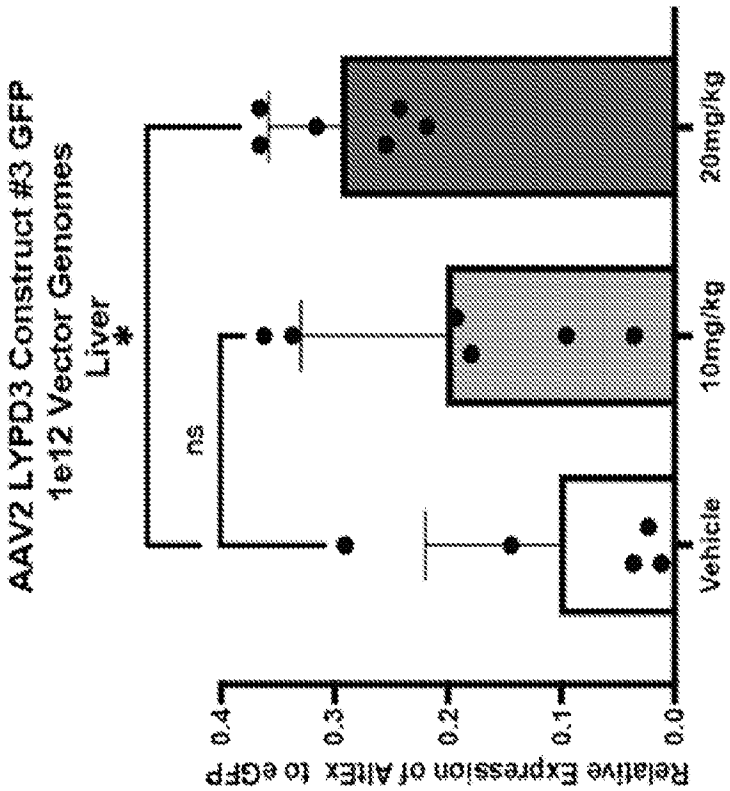


FIG. 39D

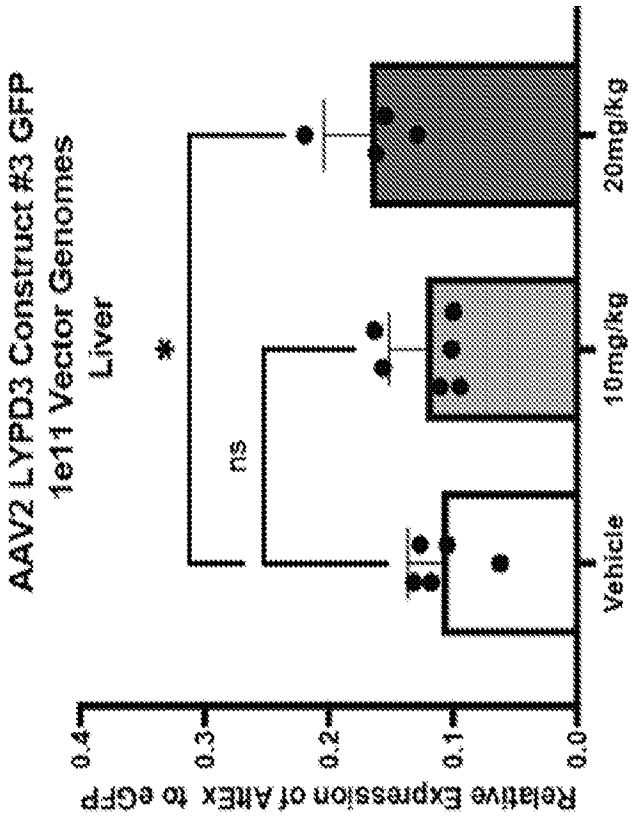


FIG. 39C

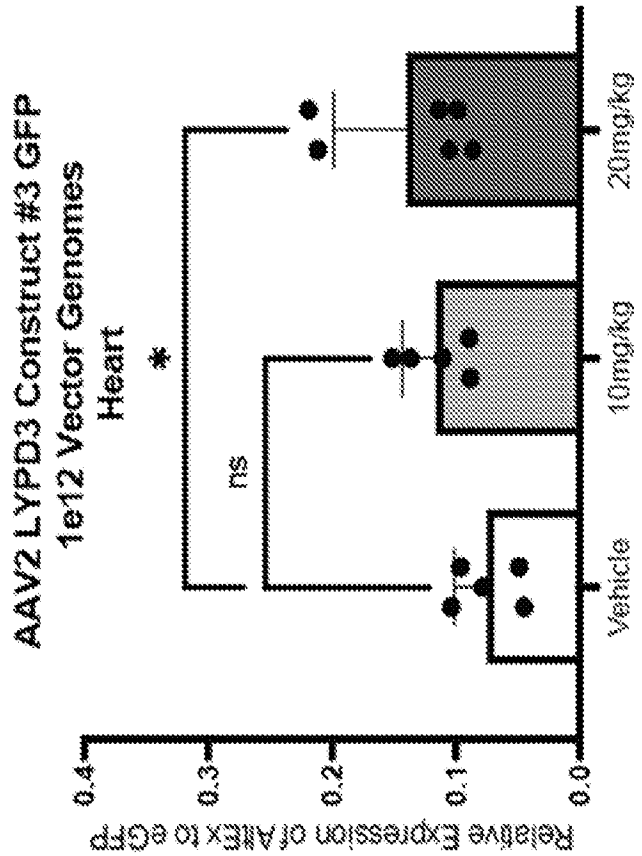


FIG. 39F

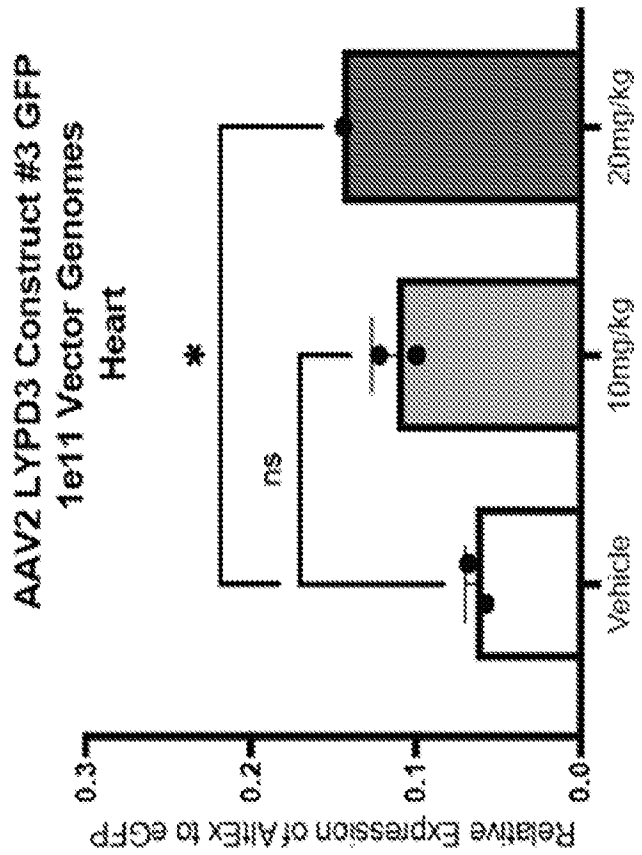


FIG. 39E

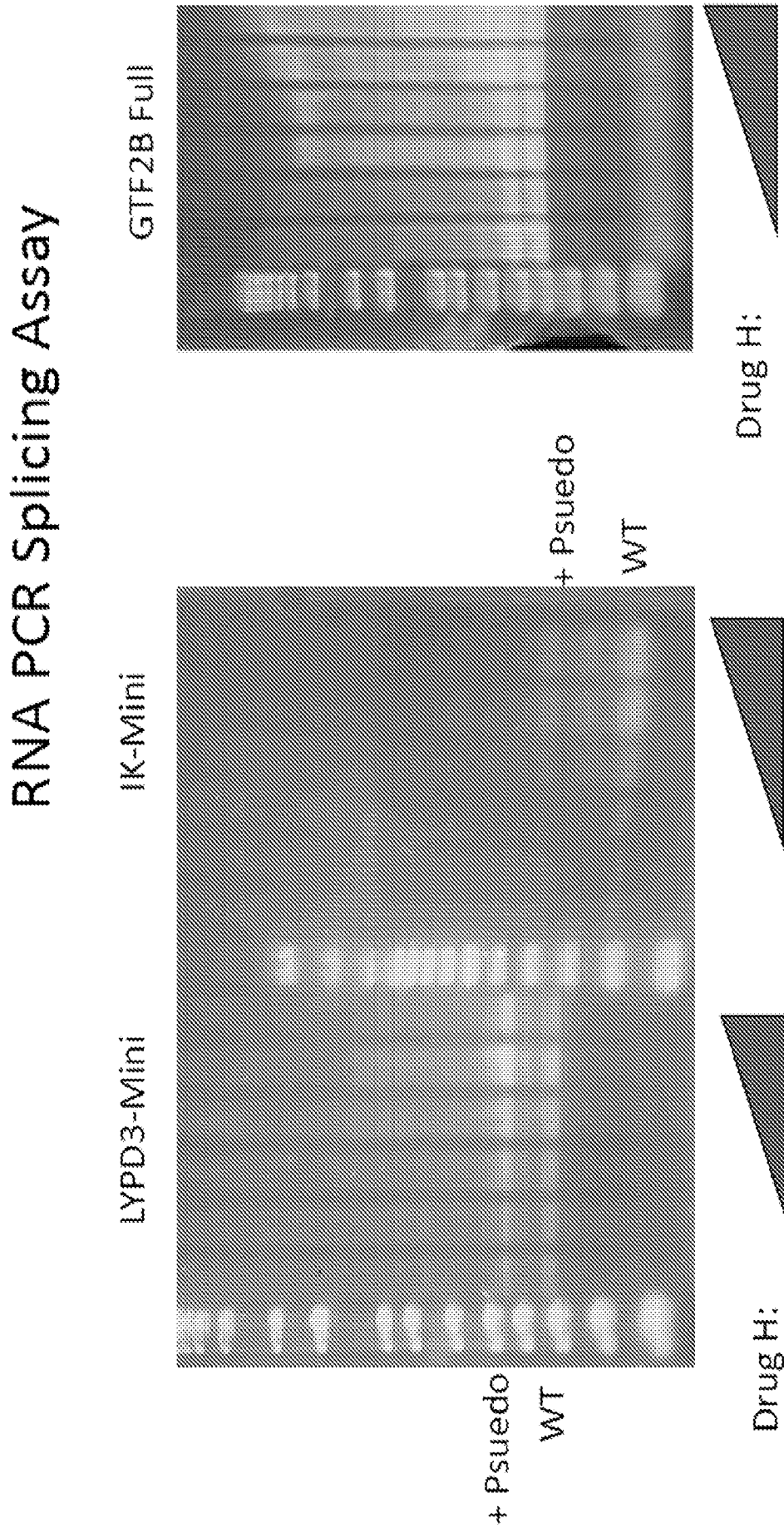

RNA PCR Splicing Assay
LYPD3-Mini
IK-Mini
GTF2B Full
+ Psuedo
WT
Drug H:
+ Psuedo
WT
Drug H:

FIG. 9

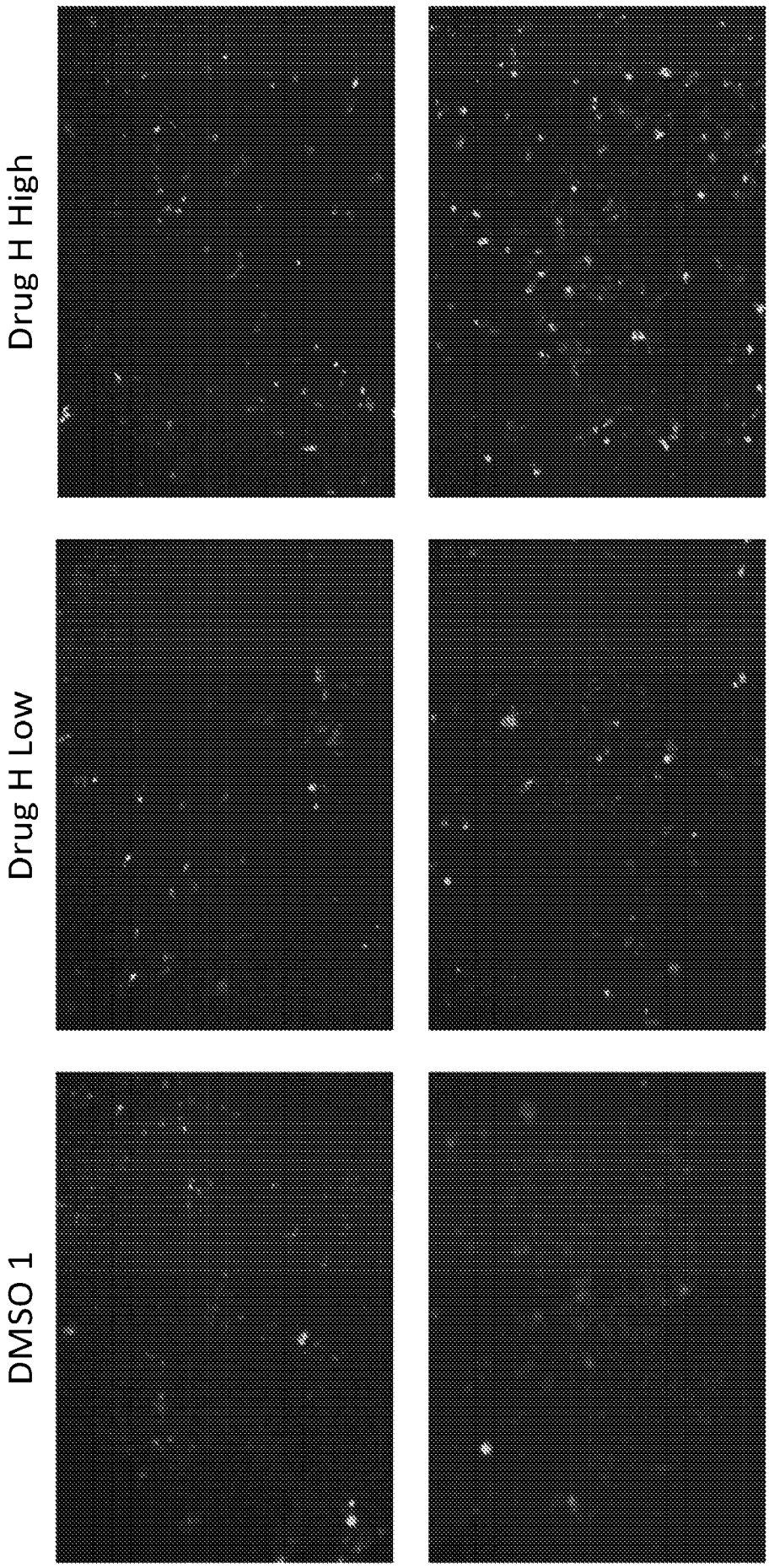

Drug H High
Drug H Low
DMSO 1
Original Drug H Full Gene L (V1) -> 1970bp

FIG. 10A

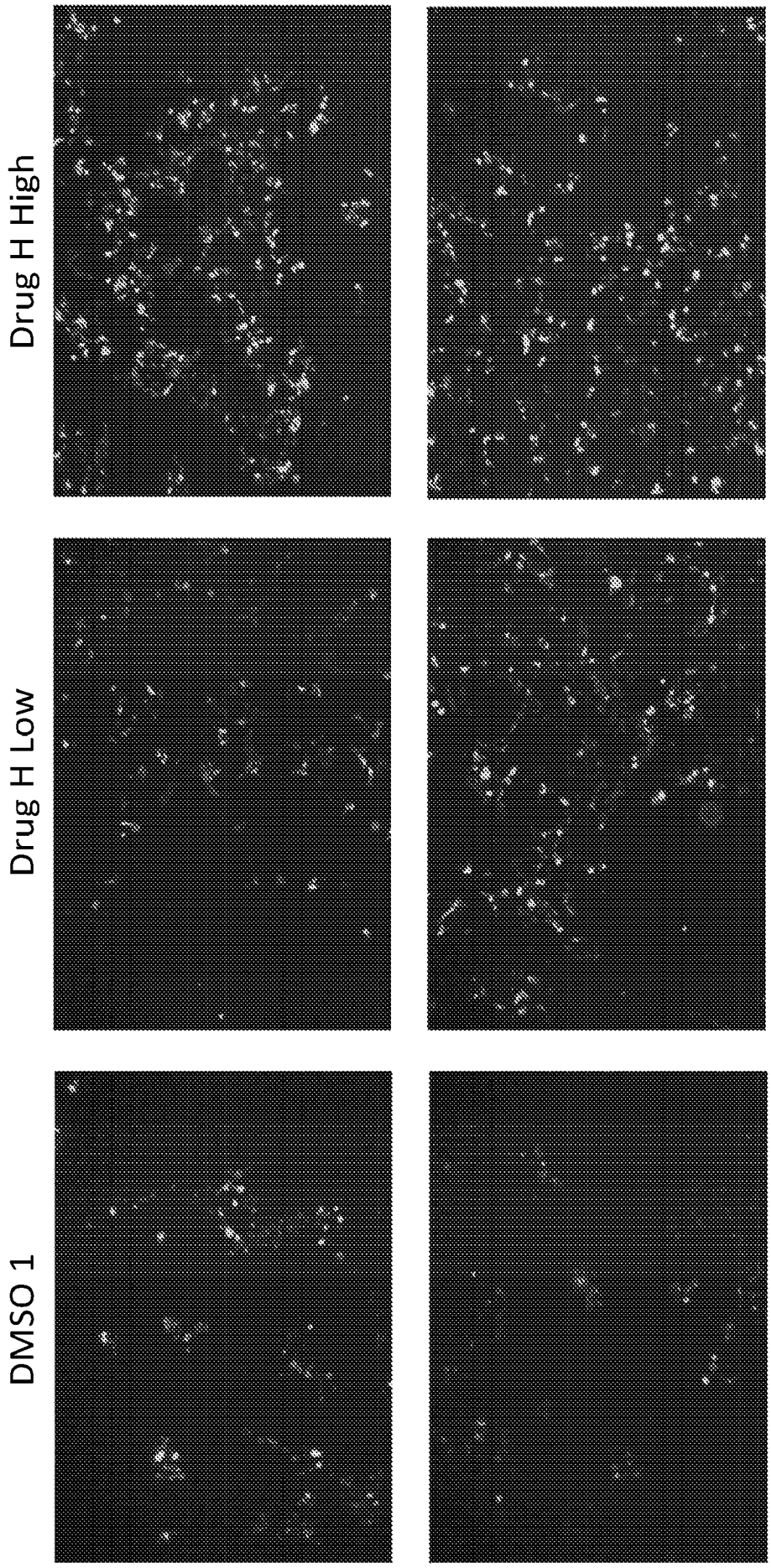


FIG. 10B

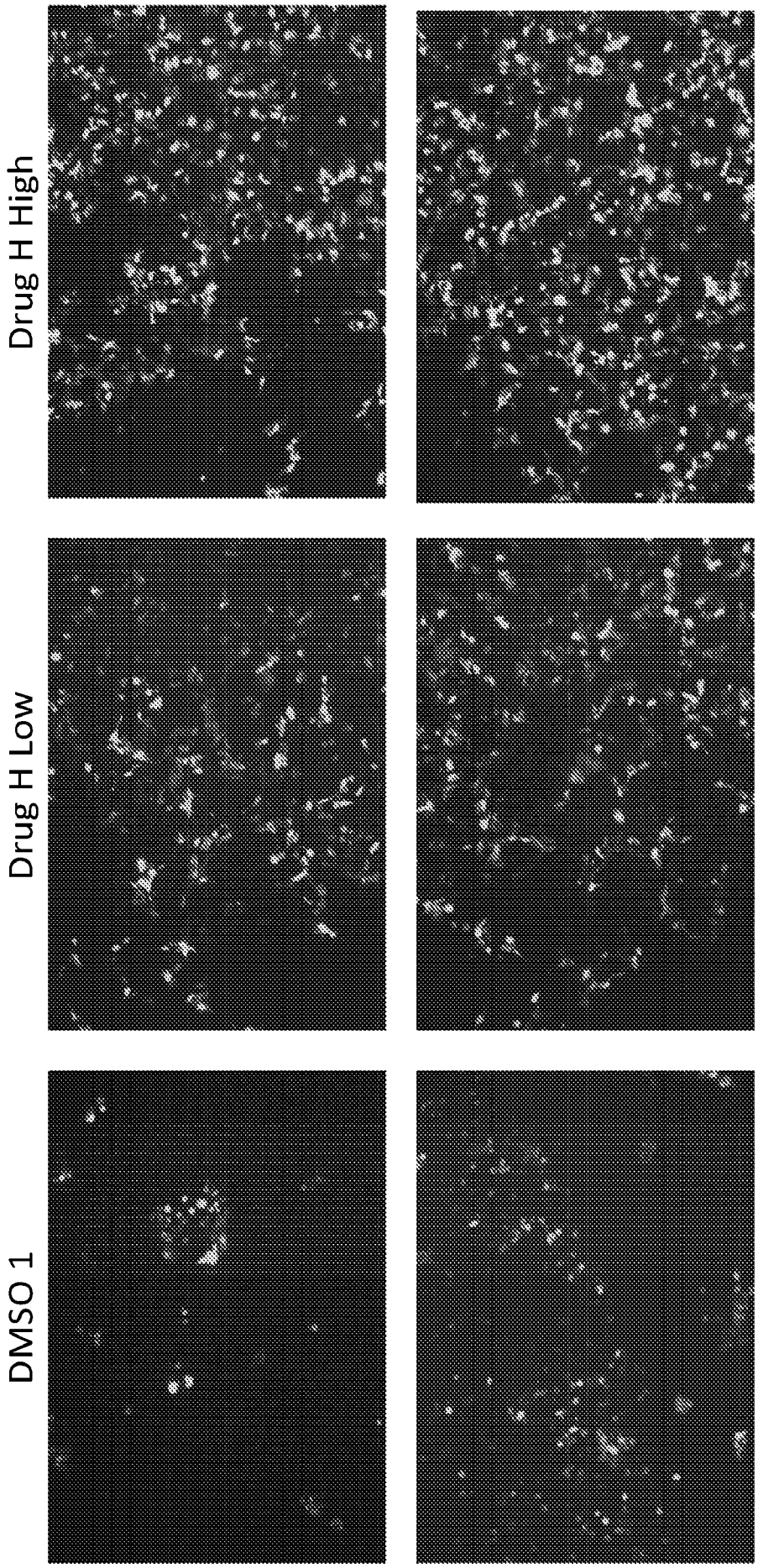

Drug H High
Drug H Low
DMSO 1
Minimum Exon/Intron Drug H Gene L (V2) -> 414bp

FIG. 10C

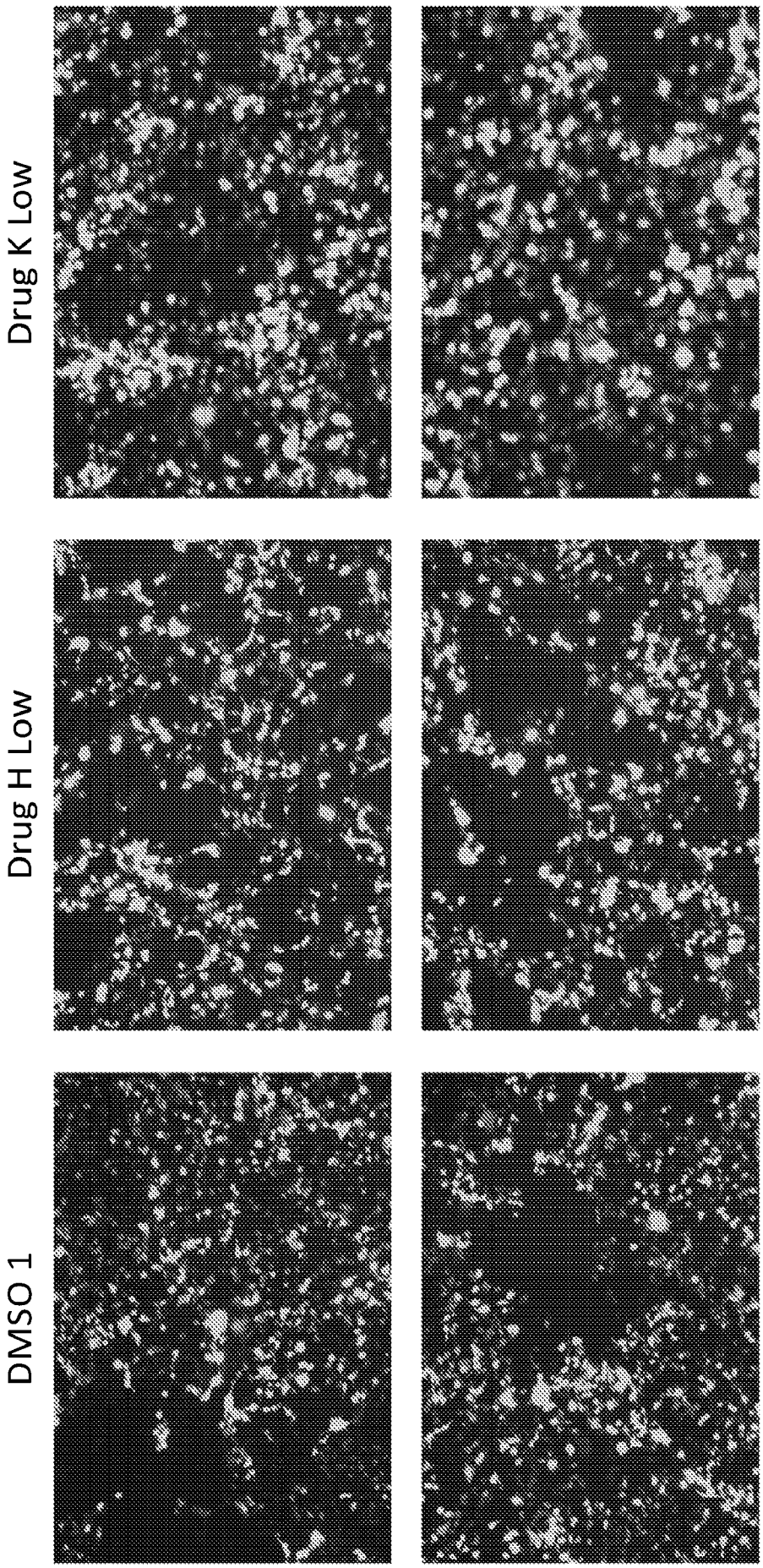


FIG. 11A

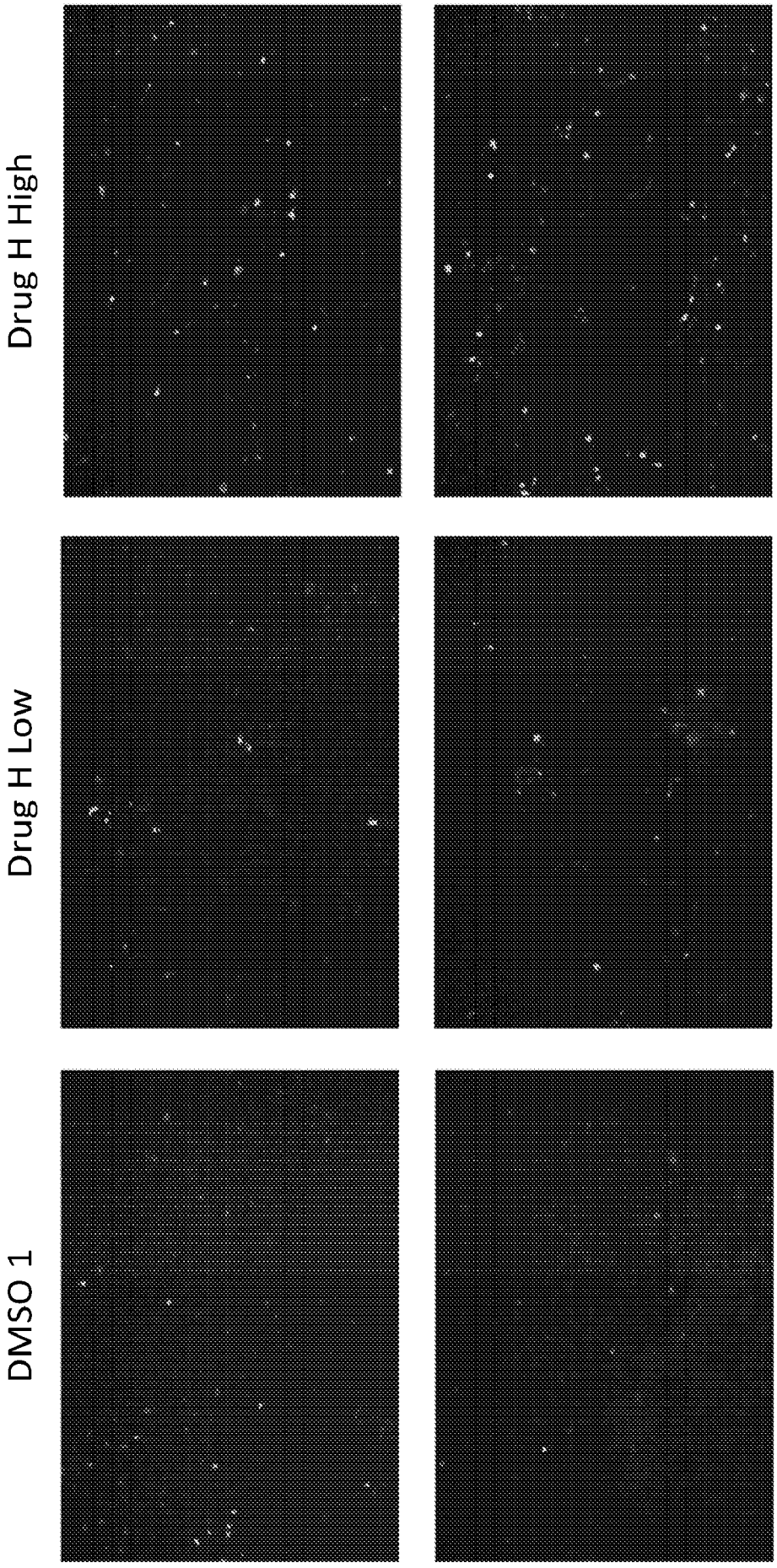

Drug H High
Drug H Low
DMSO 1
Original Drug H Full Gene L (V1) -> 1970bp

FIG. 11B

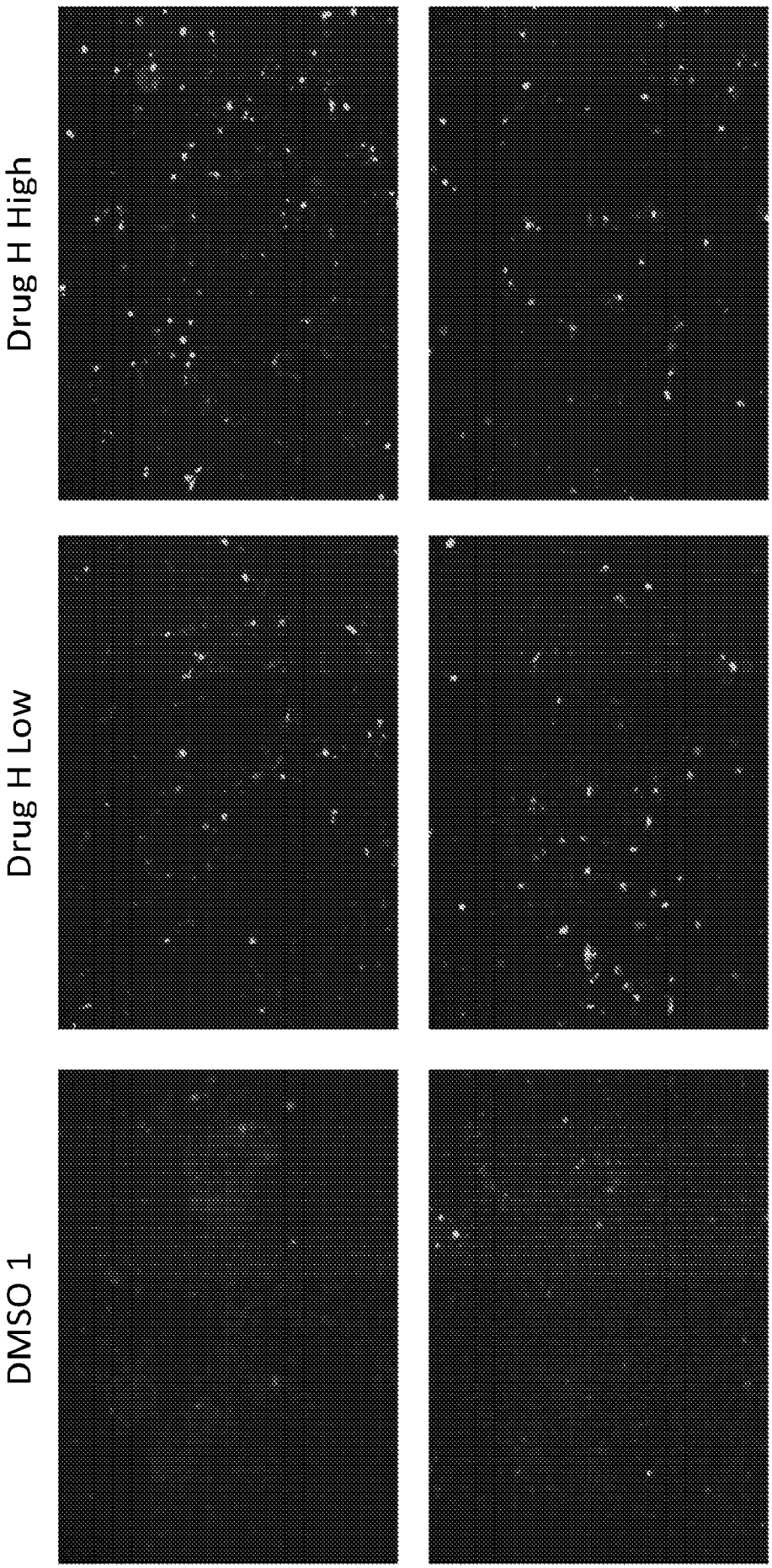


FIG. 11C

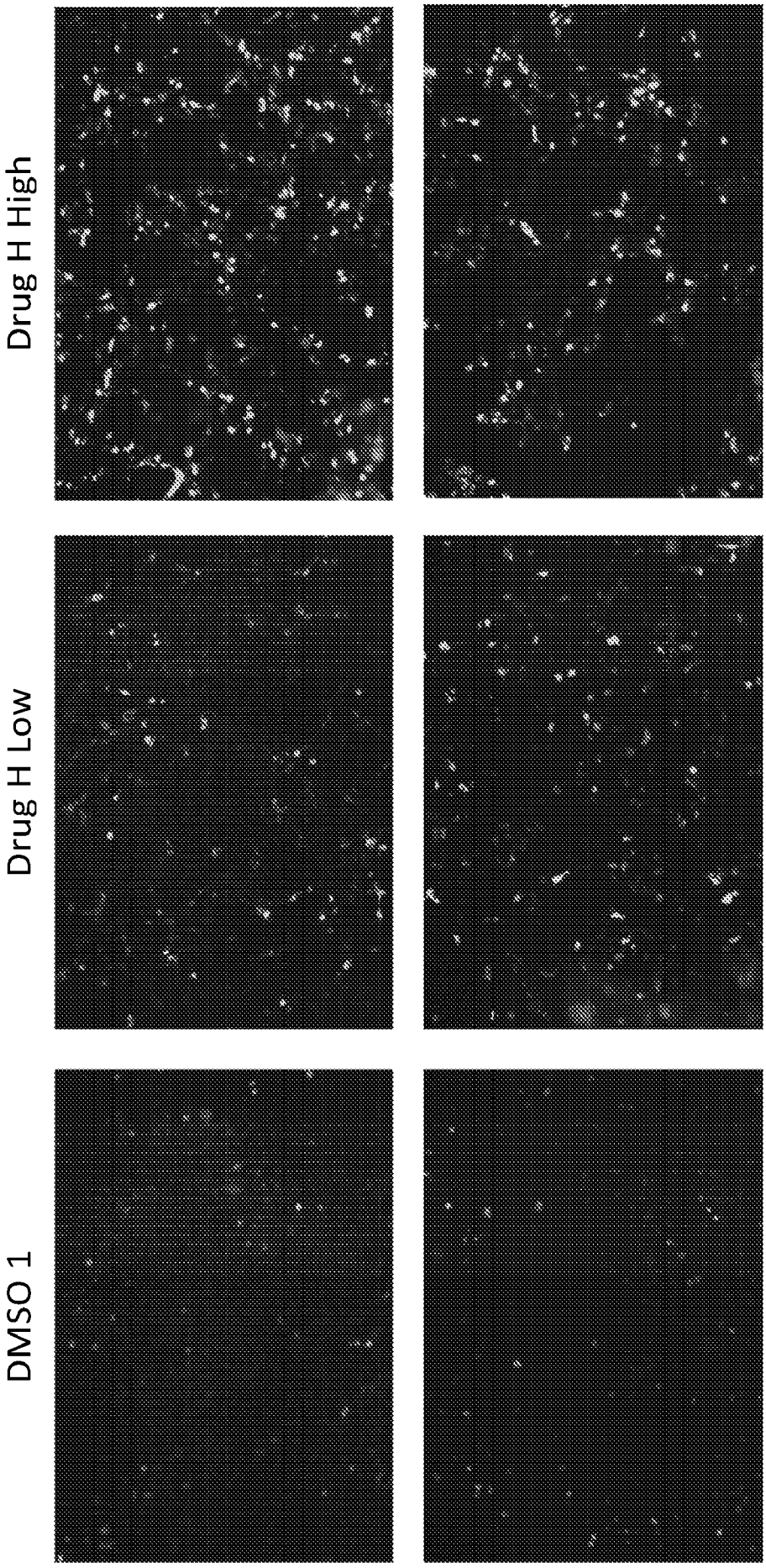

Drug H High
Drug H Low
DMSO 1
Minimum Exon/Intron Drug H Gene L (V2) -> 414bp

FIG. 11D

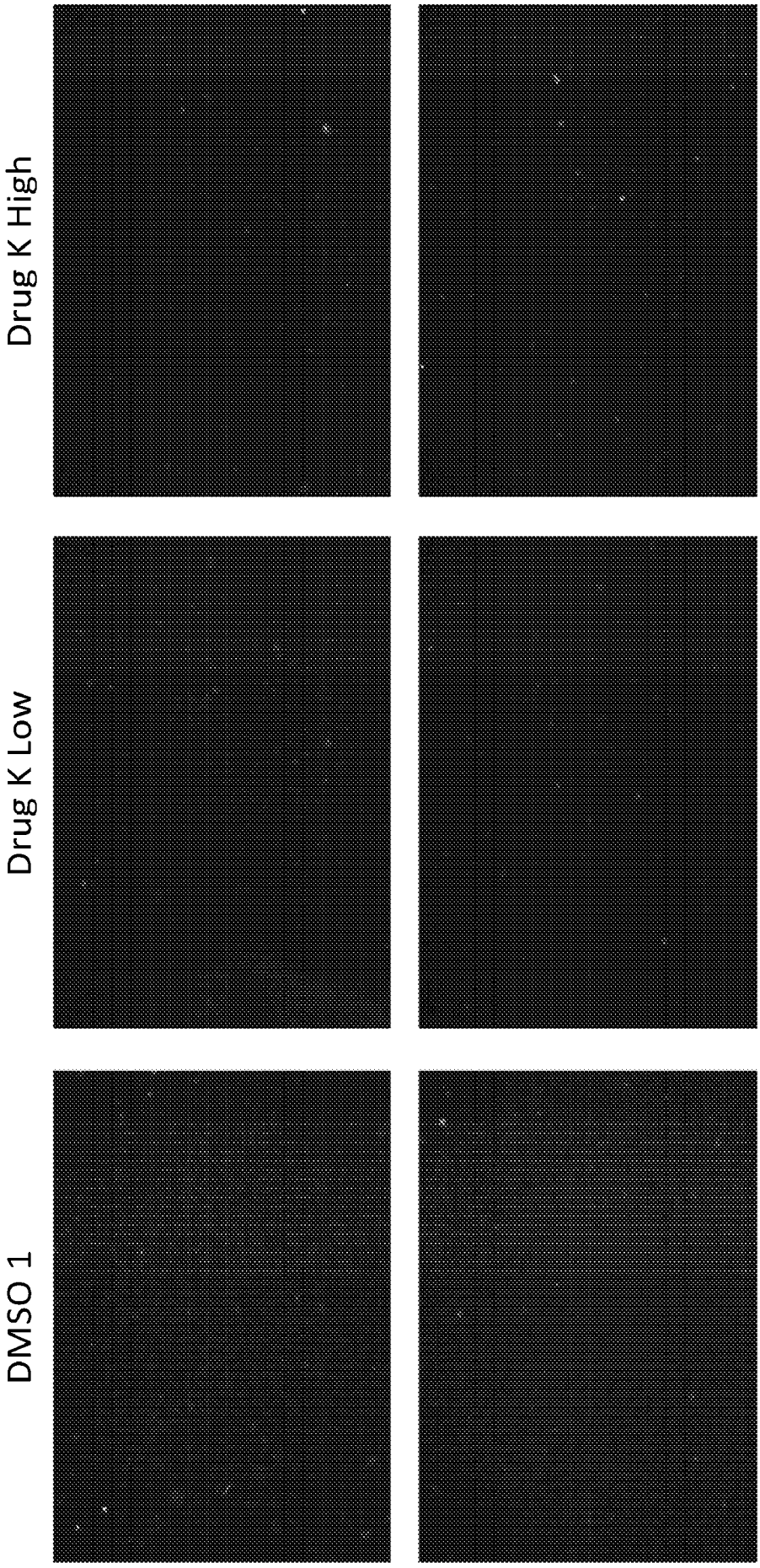

Drug K High
Drug K Low
DMSO 1
Original Drug K Full Gene T (V1) -> 5786 bp

FIG. 11E

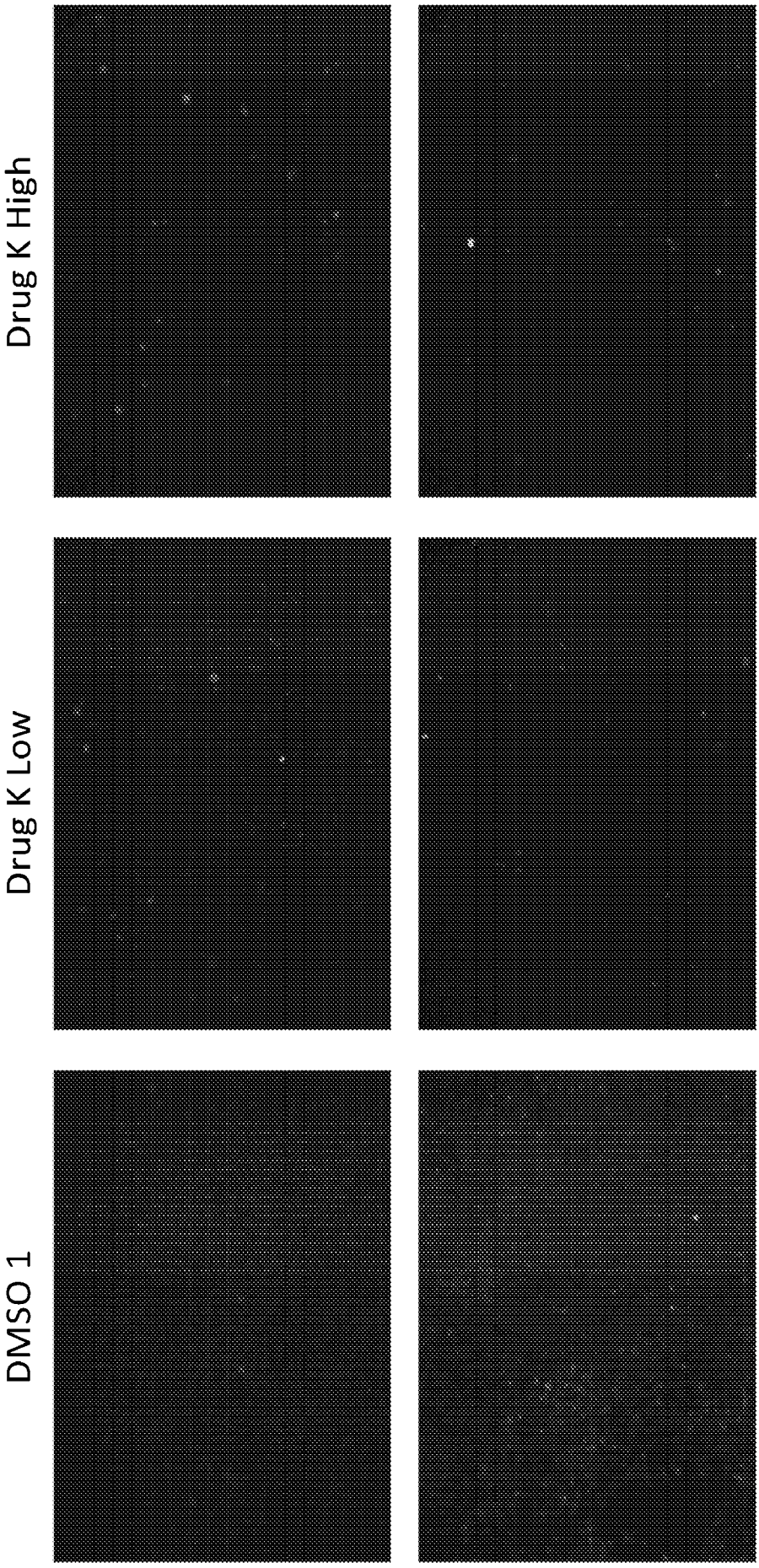

Drug K High
Drug K Low
DMSO 1
Minimum Drug K Full Gene T (V1) -> 906 bp

FIG. 11F

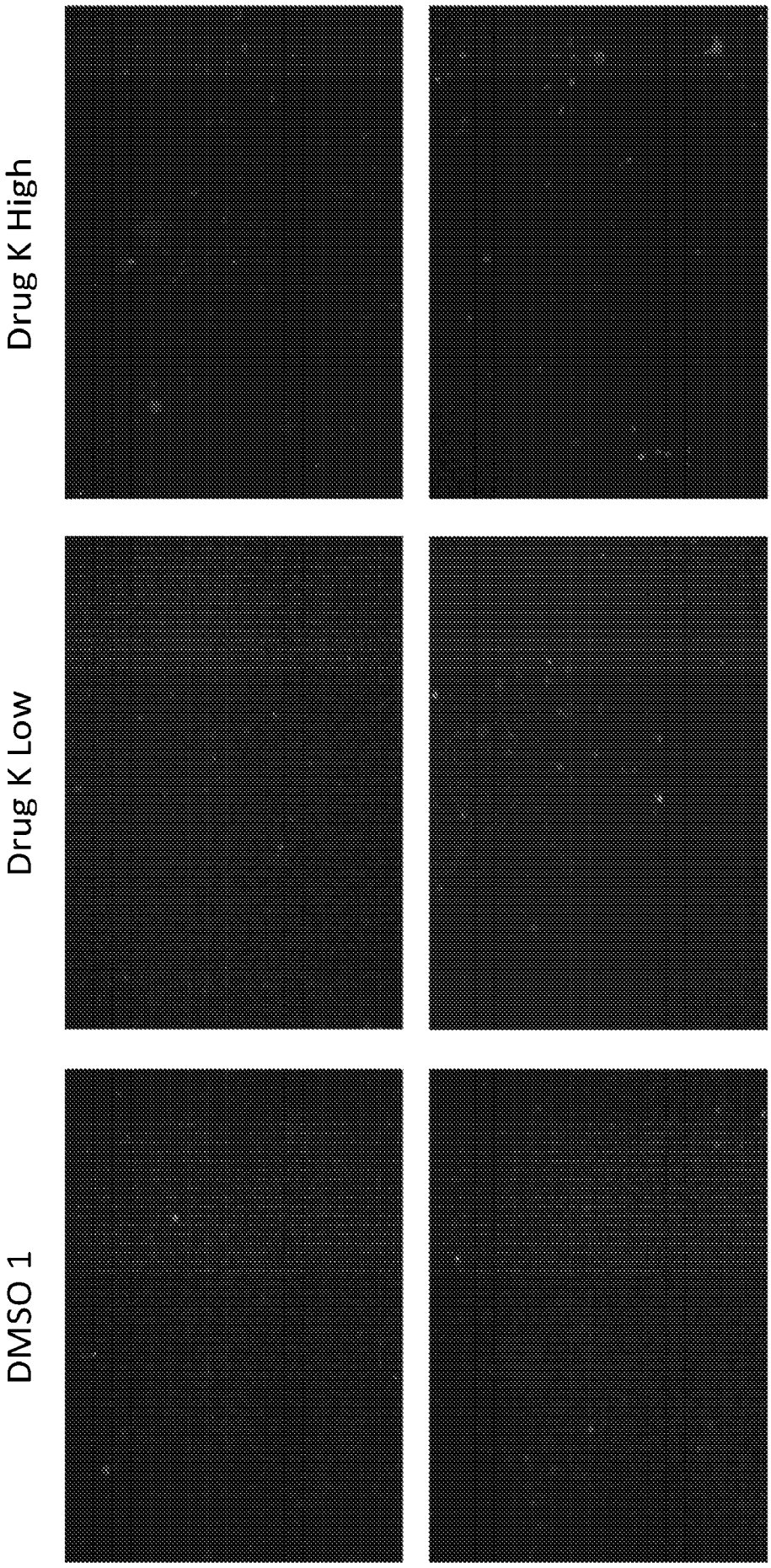

Drug K High
Drug K Low
DMSO 1
Minimum Exon/Intron Drug K Full Gene T (V2) -> 410bp

FIG. 11G

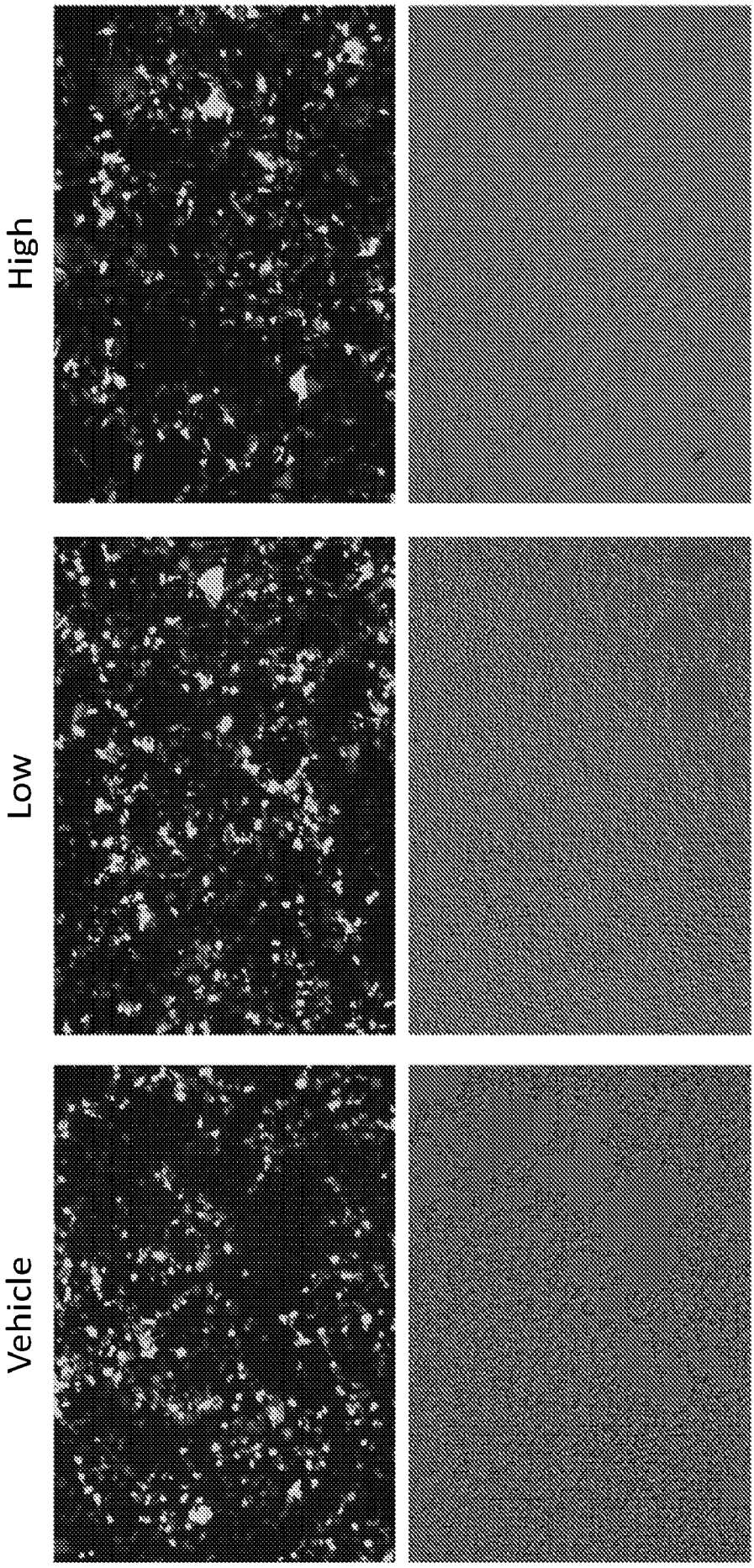


FIG. 12A

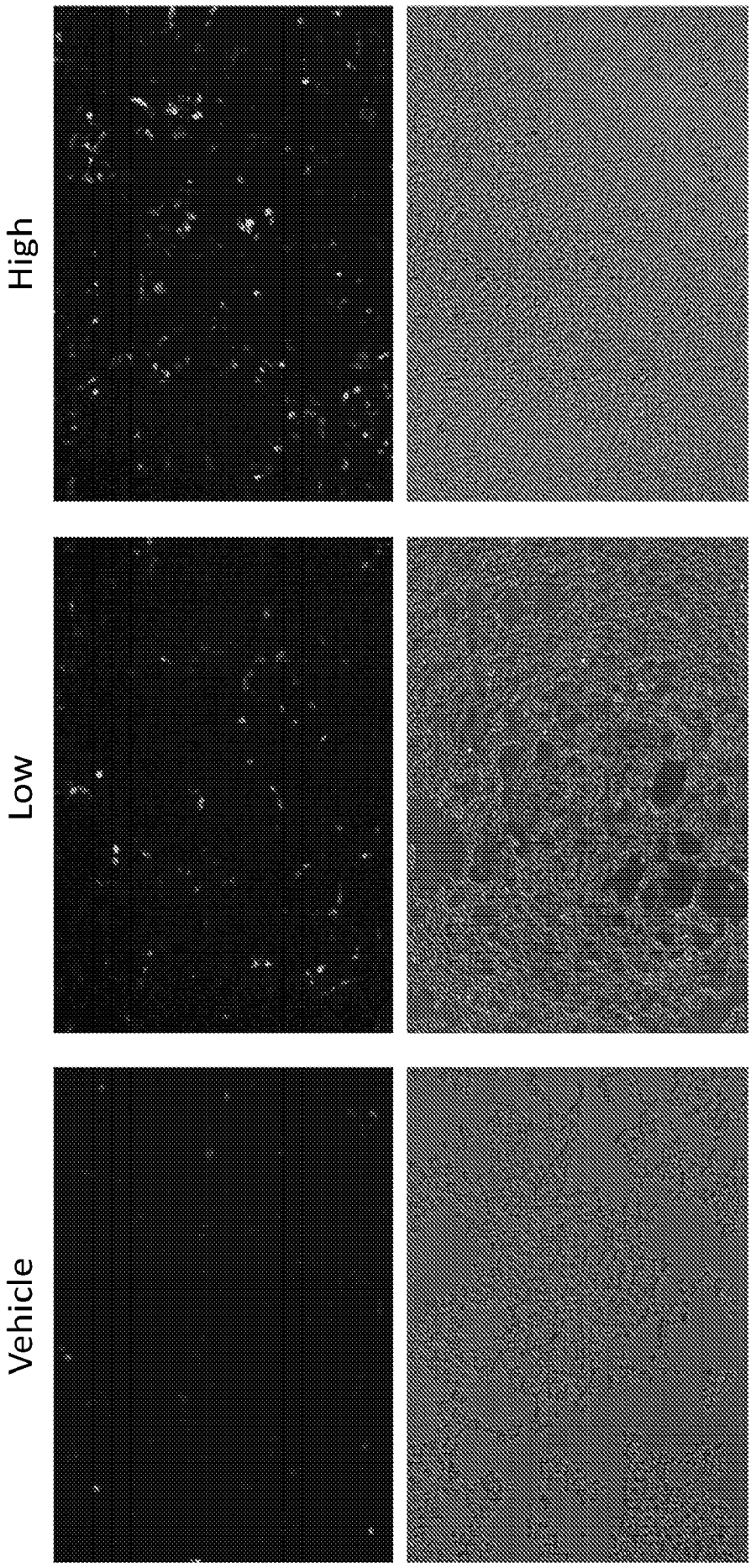


FIG. 12B

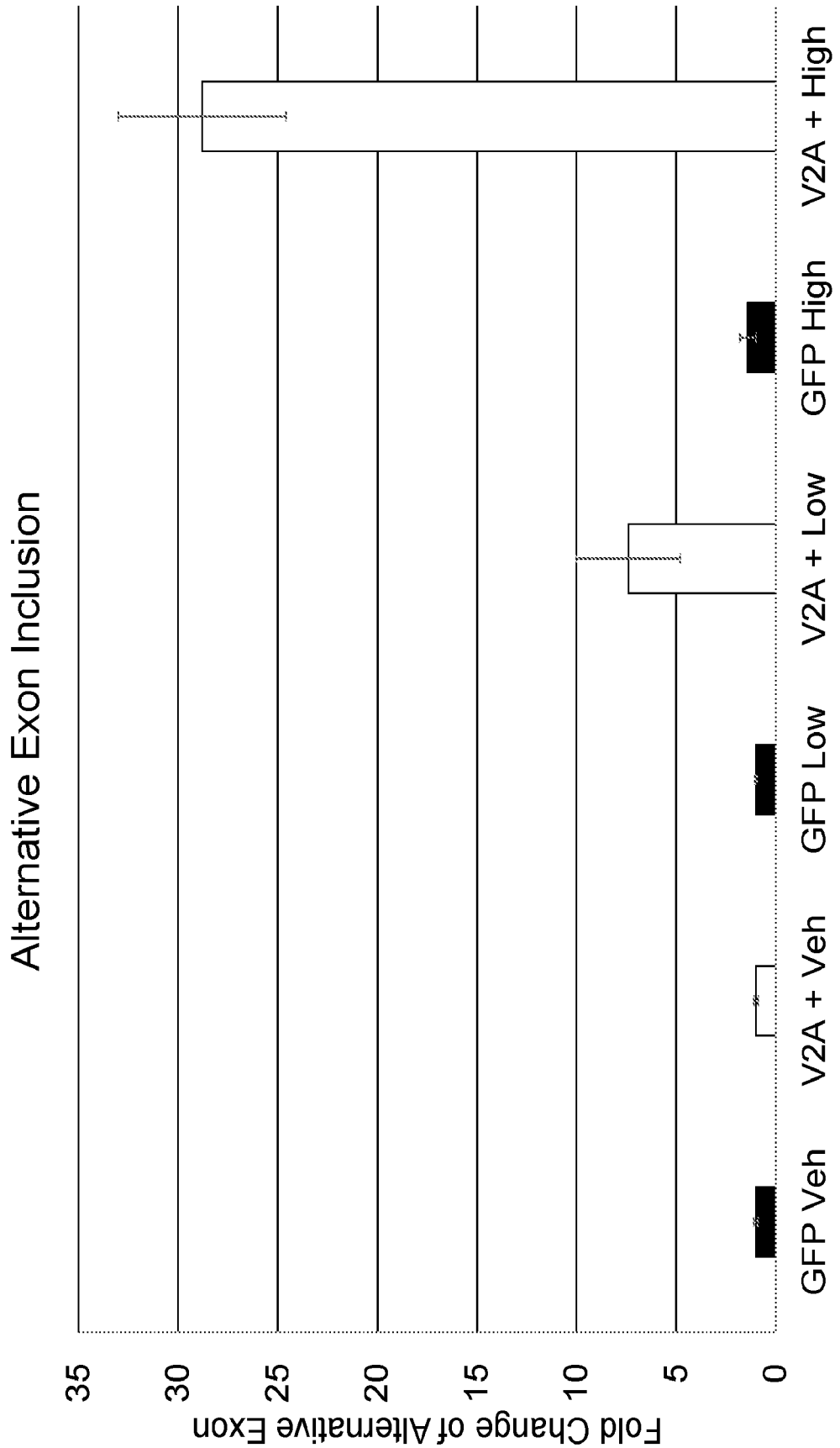

Alternative Exon Inclusion
Fold Change of Alternative Exon
35
30
25
20
15
10
5
0
GFP Veh
V2A + Veh
GFP Low
V2A + Low
GFP High
V2A + High

FIG. 12C

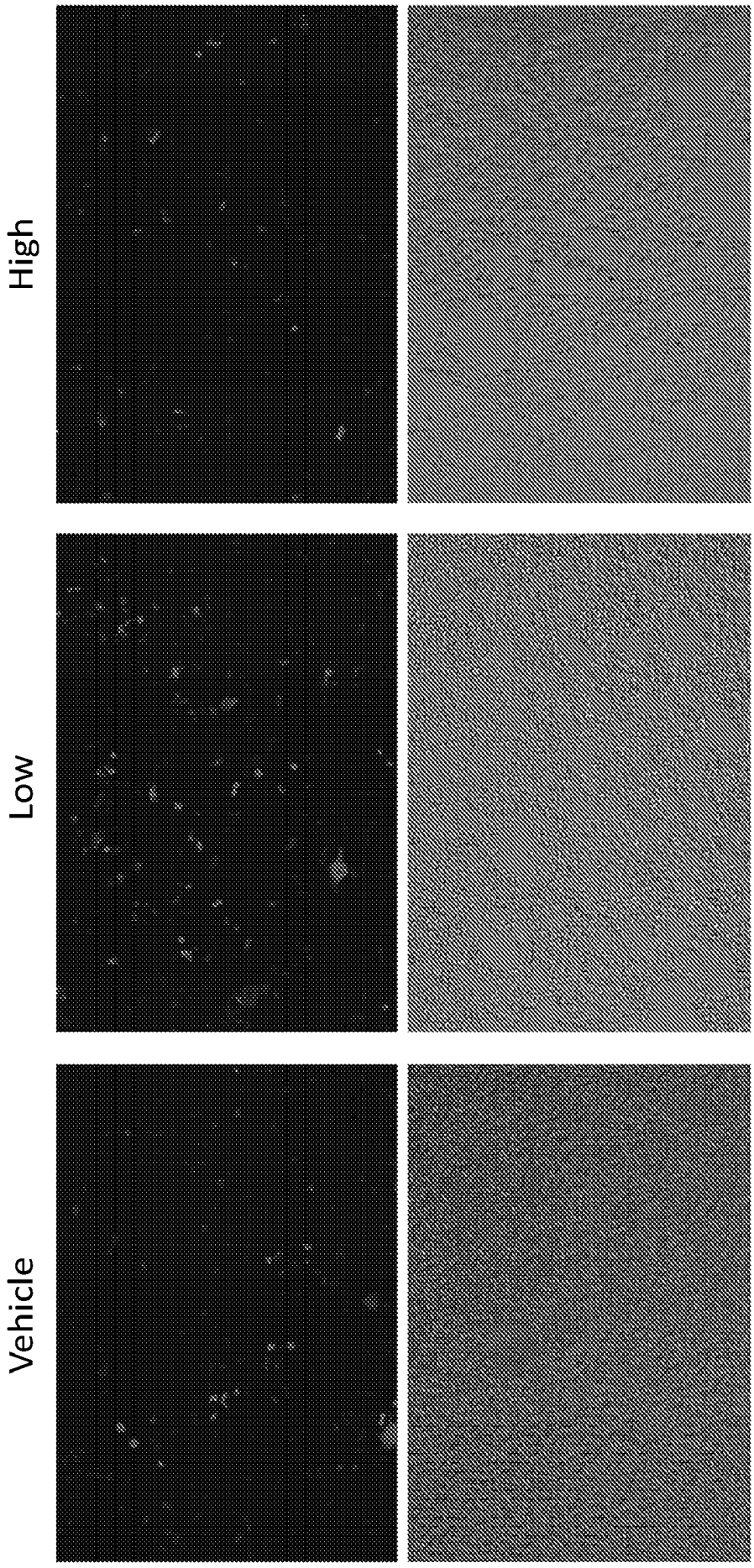


FIG. 13A

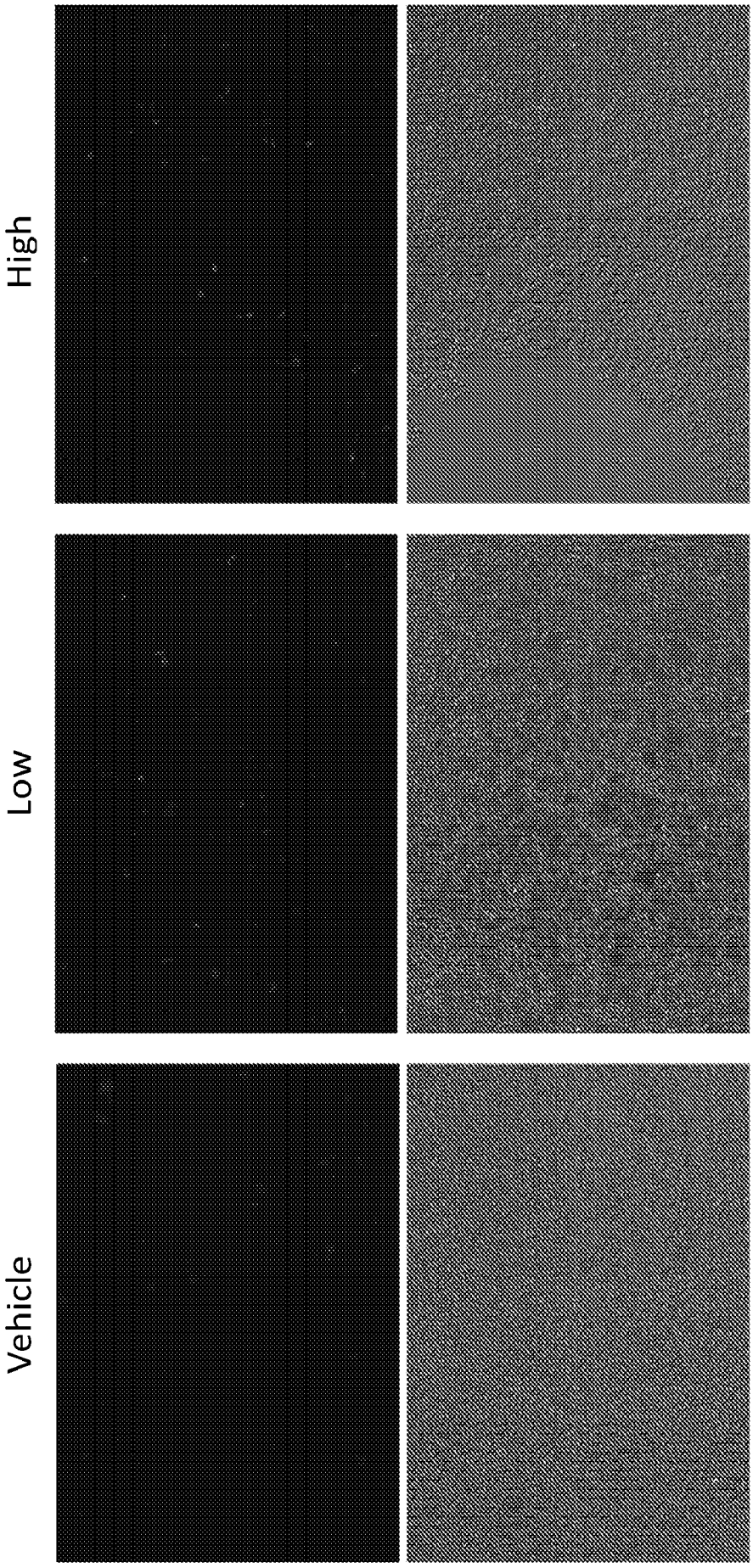


FIG. 13B

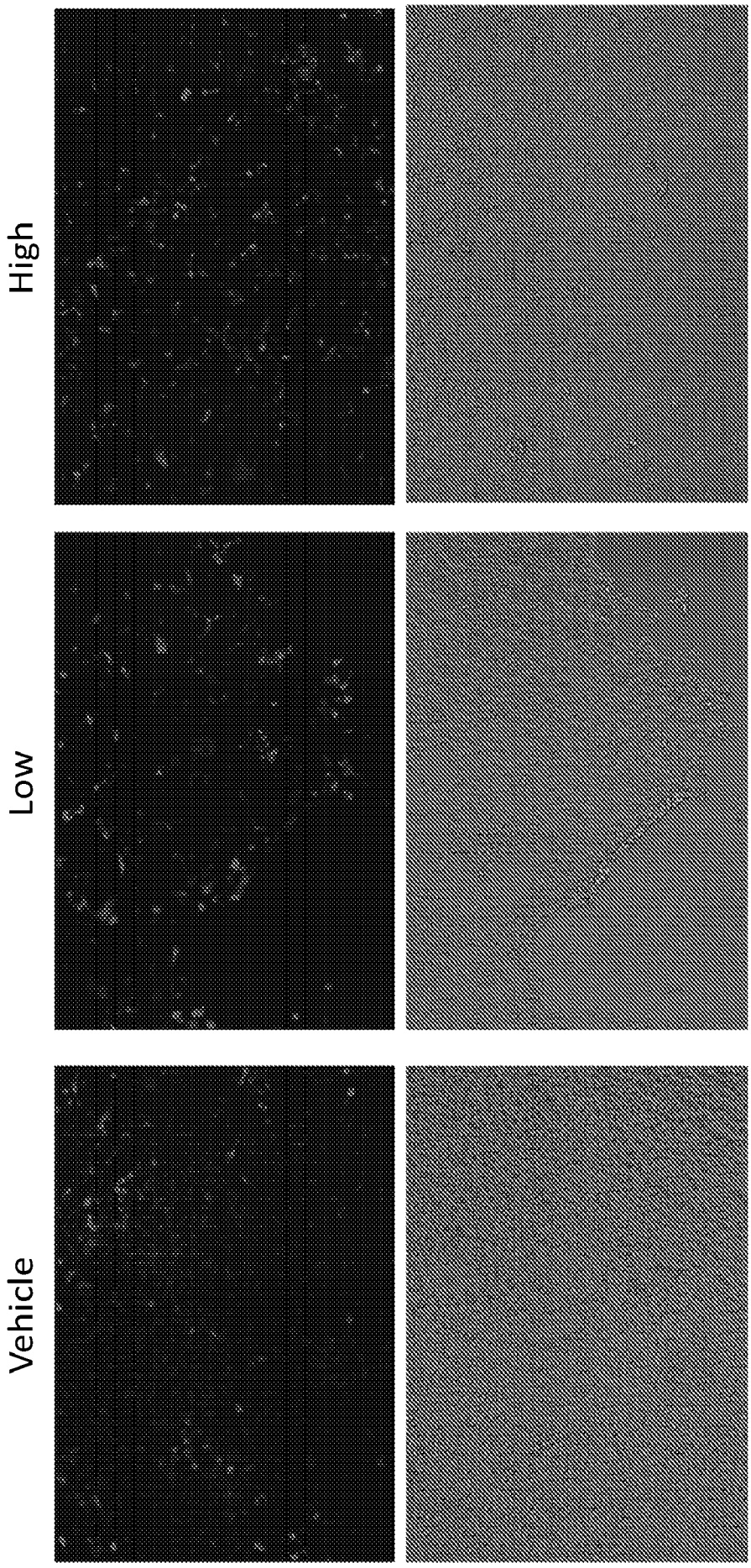

High
Low
Vehicle
iRFP KT Splice Constitutively Expressed Exons

FIG. 13C

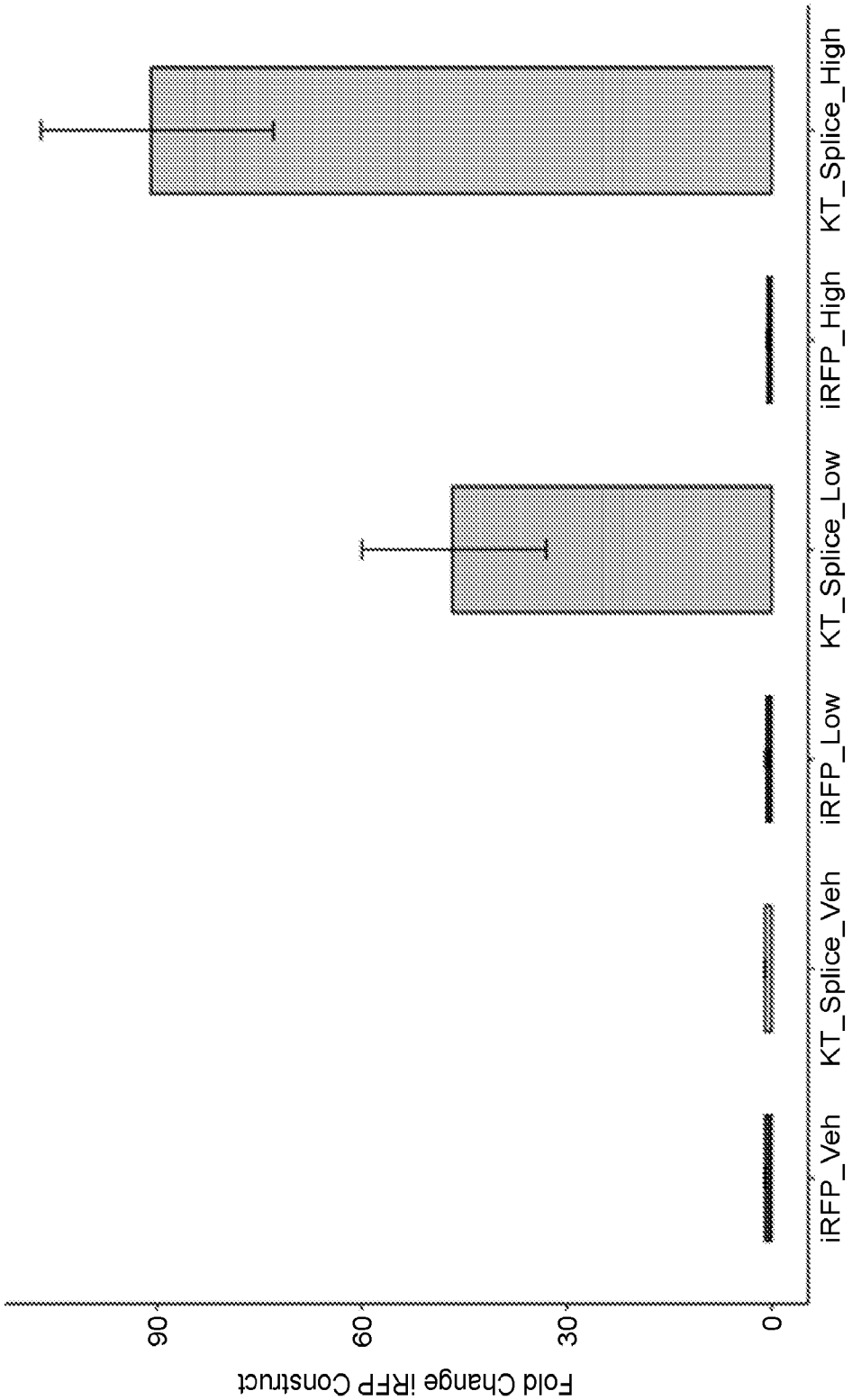

Fold Change iRFP Construct
0
30
60
90
iRFP_Veh
KT_Splice_Veh
iRFP_Low
KT_Splice_Low
iRFP_High
KT_Splice_High
Transfection

FIG. 13D

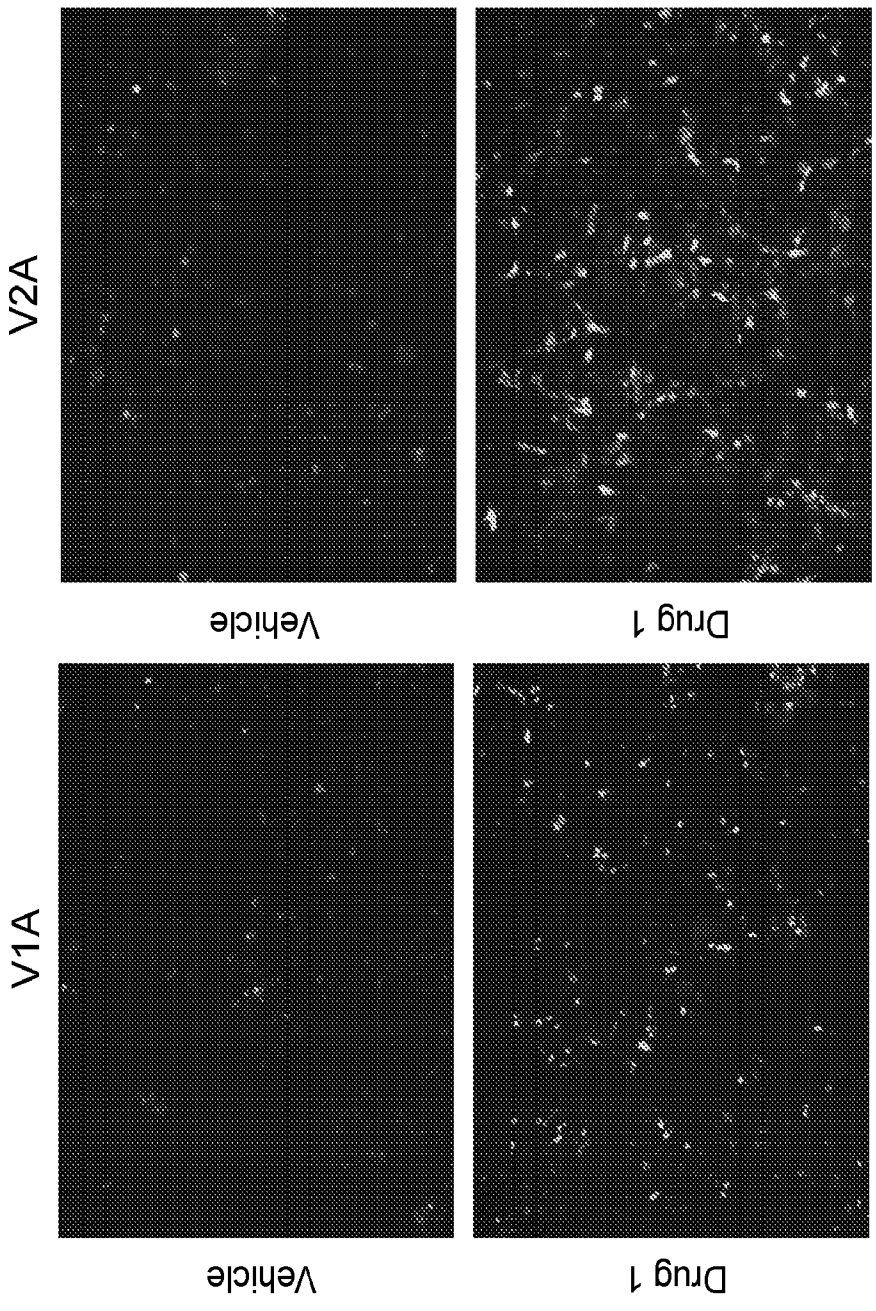


FIG. 14

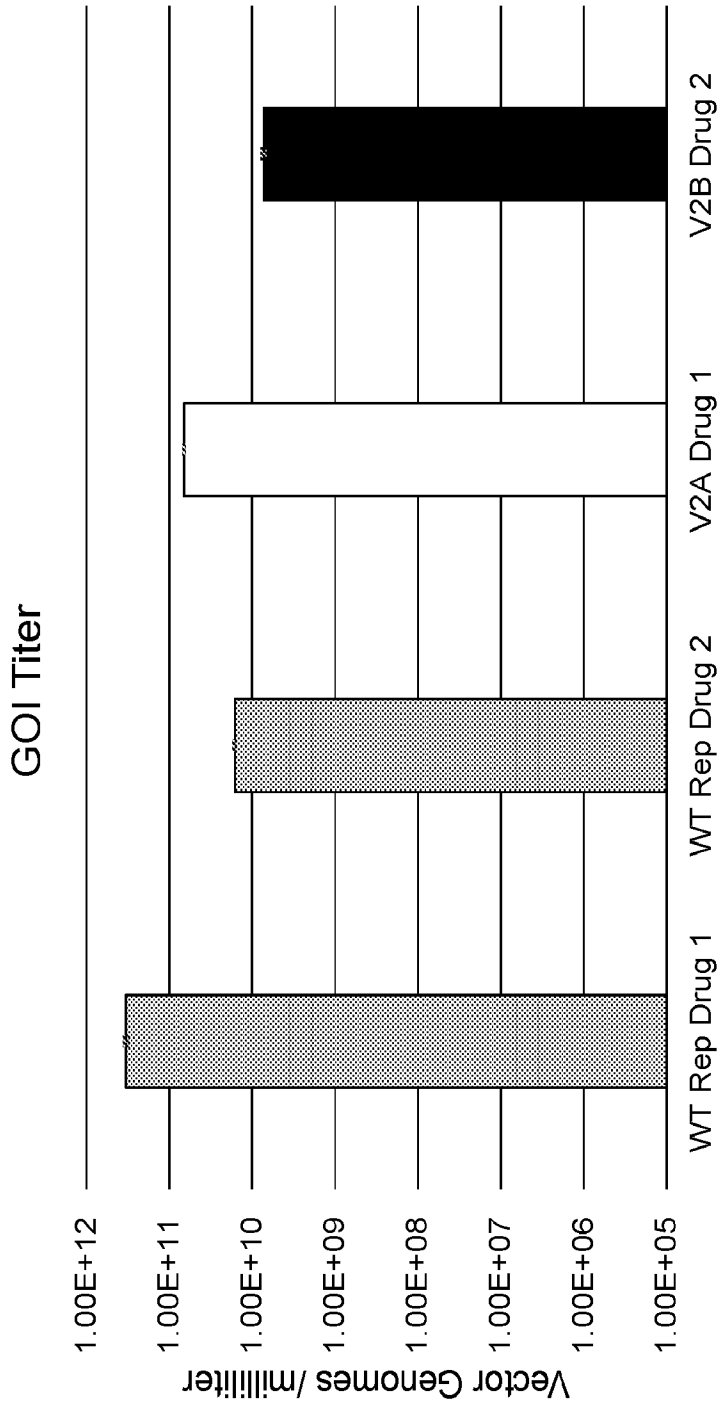

GOI Titer
Vector Genomes /milliliter
1.00E+12
1.00E+11
1.00E+10
1.00E+09
1.00E+08
1.00E+07
1.00E+06
1.00E+05
WT Rep Drug 1
WT Rep Drug 2
V2A Drug 1
V2B Drug 2

FIG. 15

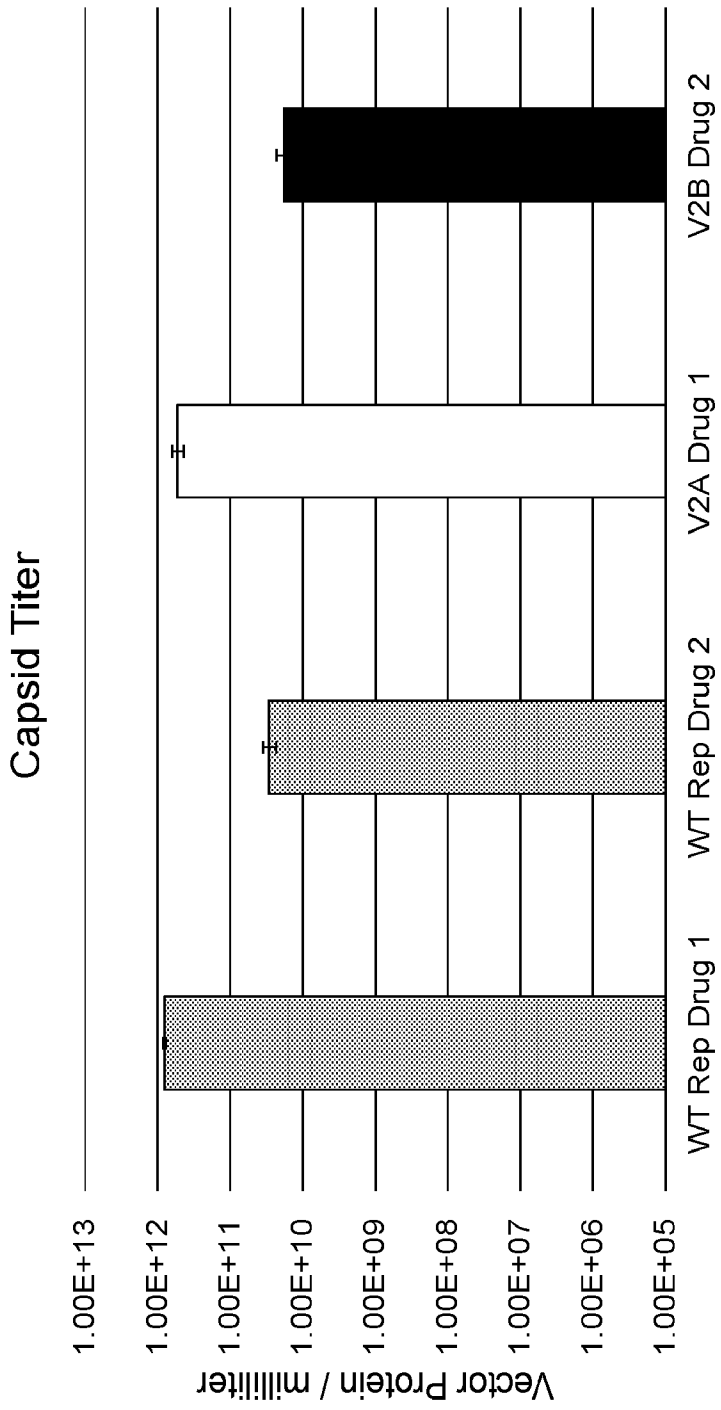
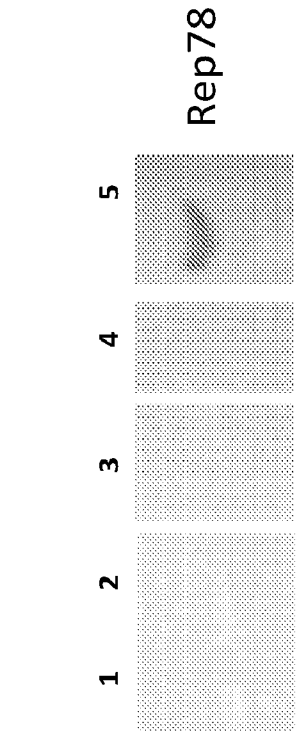
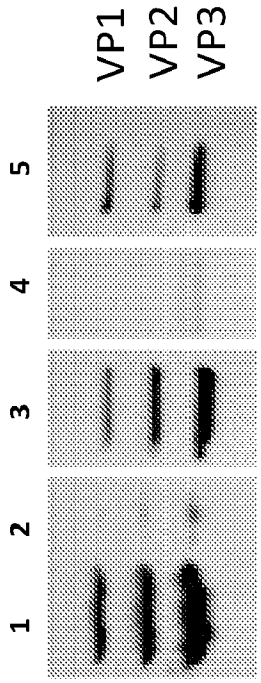

Capsid Titer
Vector Protein / milliliter
1.00E+13
1.00E+12
1.00E+11
1.00E+10
1.00E+09
1.00E+08
1.00E+07
1.00E+06
1.00E+05
WT Rep Drug 1
WT Rep Drug 2
V2A Drug 1
V2B Drug 2

FIG. 16



ID	
1	WT Rep Drug 1
2	WT Rep Drug 2
3	V2A Drug 1
4	V2B Drug 2
5	Purified Rep78

FIG. 17A



ID	
1	WT Rep Drug 1
2	WT Rep Drug 2
3	V2A Drug 1
4	V2B Drug 2
5	Purified AAV

FIG. 17B

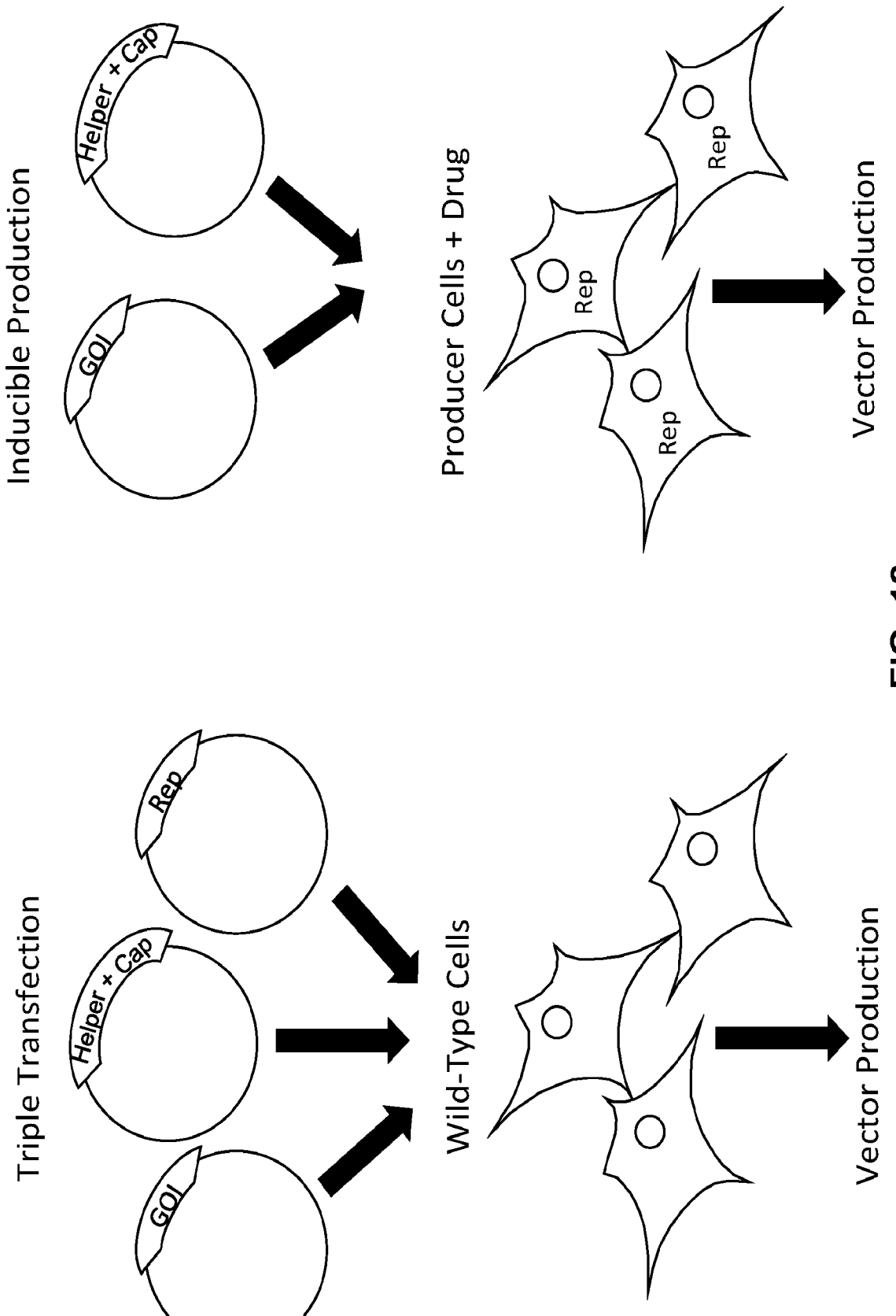
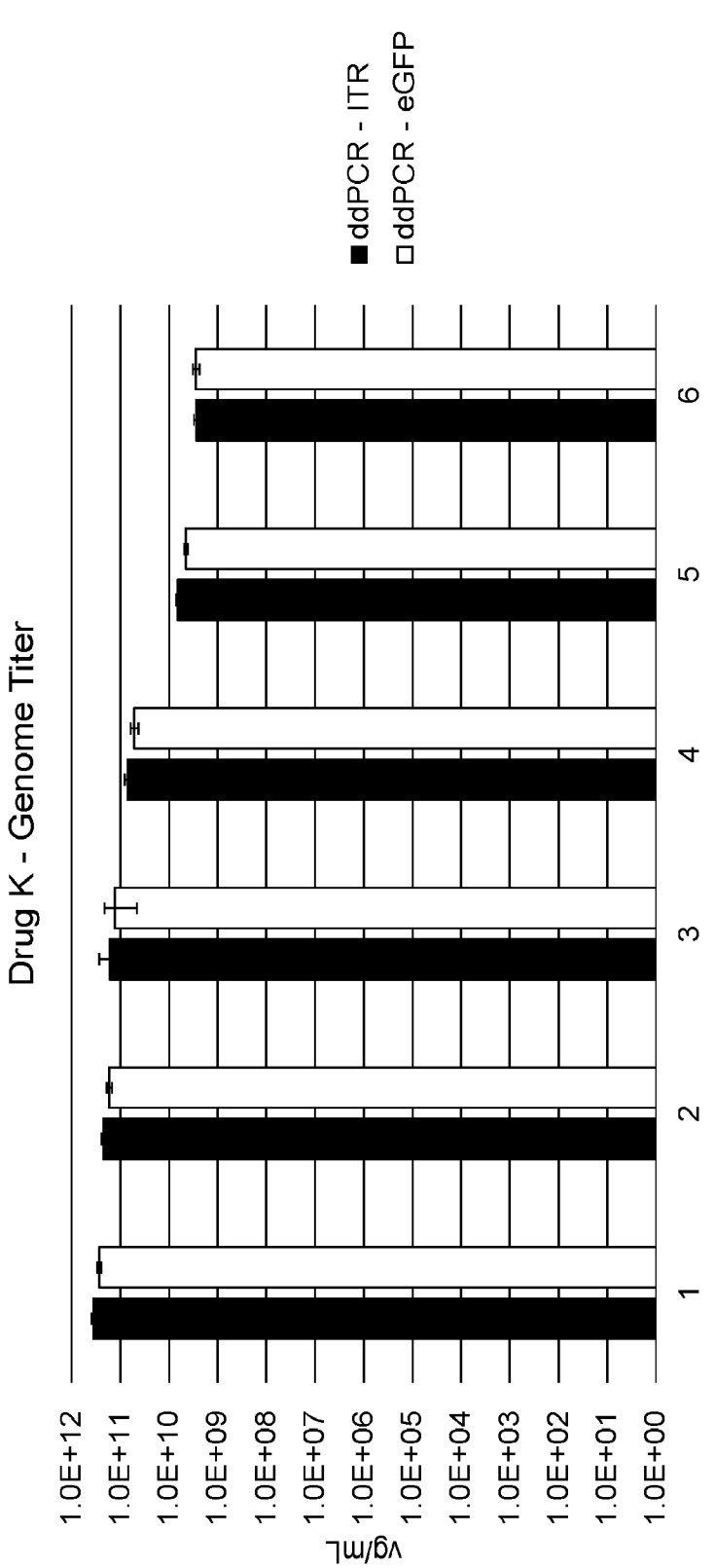
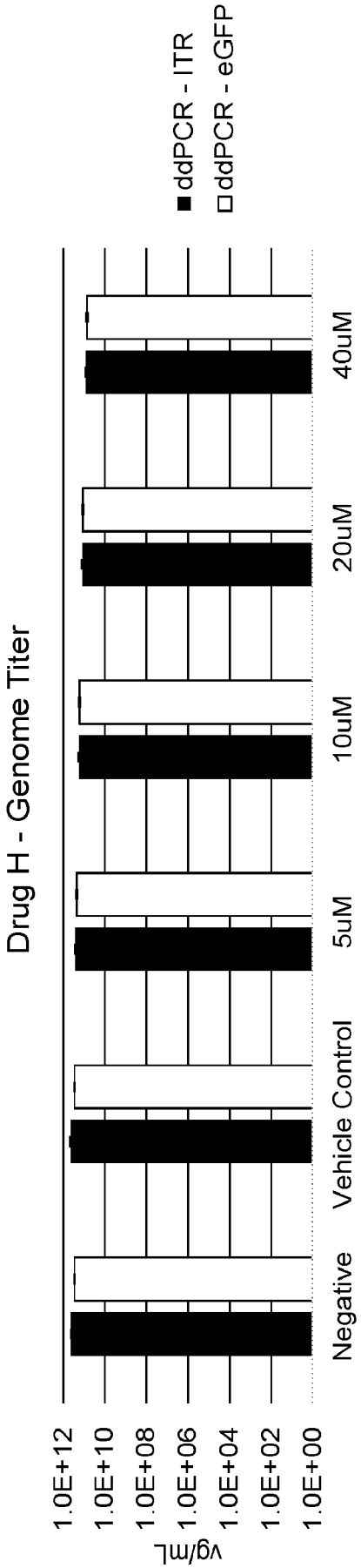

Inducible Production
Helper + Cap
GOI
Producer Cells + Drug
Rep
Rep
Rep
Vector Production
Triple Transfection
Rep
Helper + Cap
GOI
Wild-Type Cells
Vector Production

FIG. 18



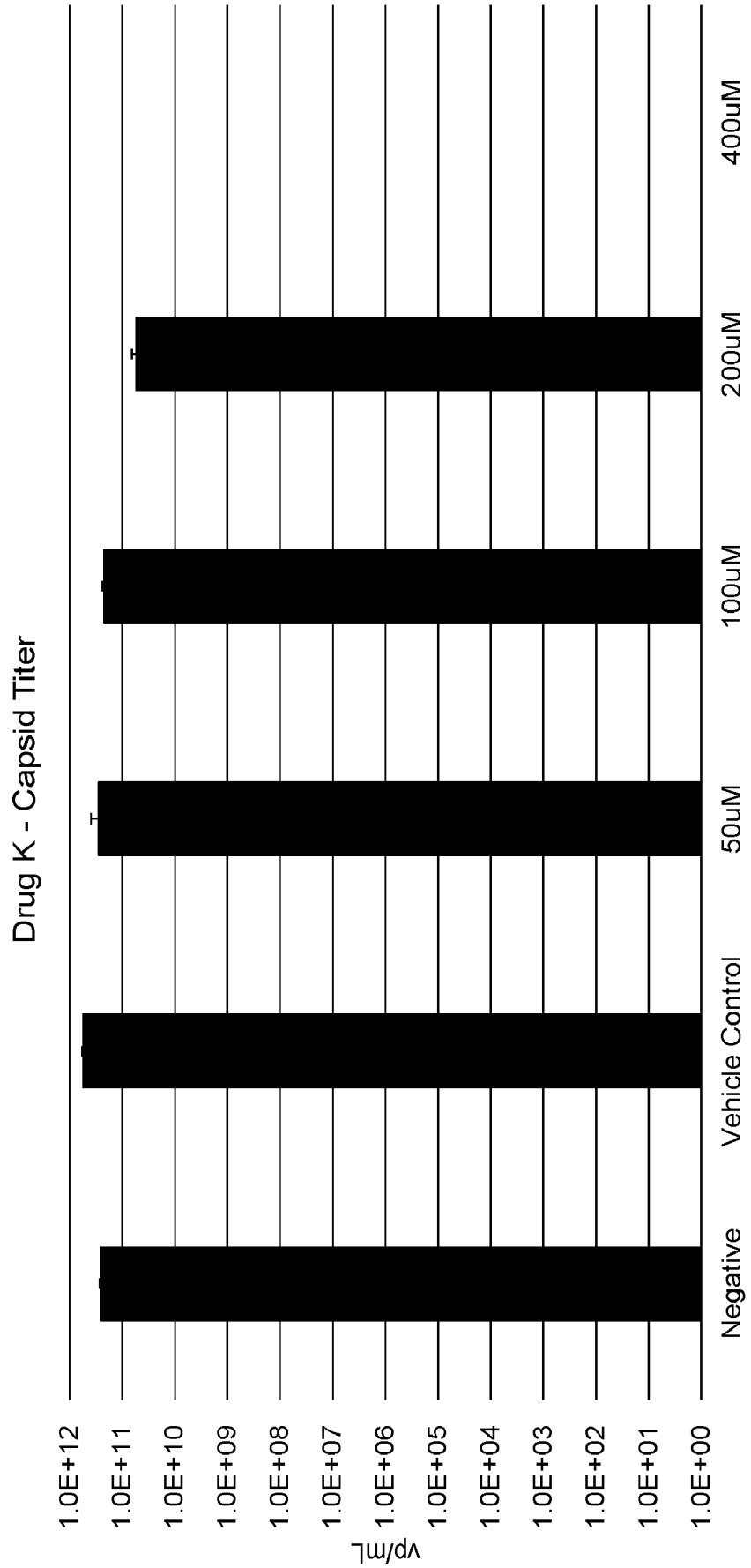
	ddPCR ITR		ddPCR - eGFP	
	vp/mL	std.dev	vp/mL	std.dev
Negative	3.5E+11	3.4E+10	2.7E+11	2.5E+10
Vehicle Control	2.3E+11	2.0E+10	1.7E+11	2.0E+10
50uM	1.6E+11	1.1E+11	1.3E+11	8.3E+10
100uM	7.2E+10	9.8E+09	5.1E+10	9.0E+09
200uM	6.5E+09	4.1E+08	4.5E+09	3.4E+08
400uM	2.7E+09	3.4E+08	2.8E+09	4.3E+08

FIG. 19A



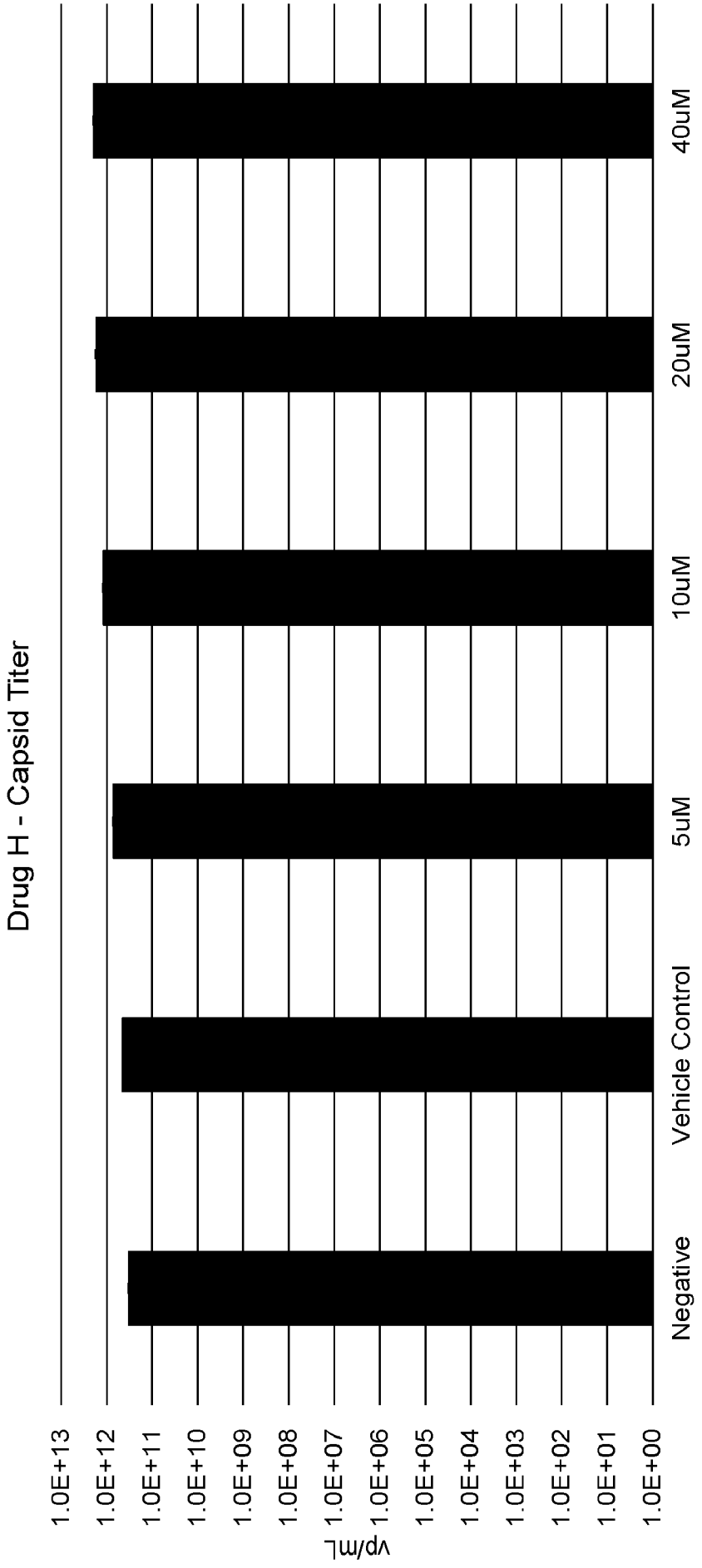
	ddPCR - ITR		ddPCR - eGFP	
	vp/mL	std.dev	vp/mL	std.dev
Negative	4.3E+11	9.4E+09	2.9E+11	9.4E+09
Vehicle Control	4.4E+11	3.3E+10	2.9E+11	1.9E+10
5uM	2.6E+11	1.6E+10	2.2E+11	1.6E+10
10uM	1.7E+11	9.4E+09	1.6E+11	9.0E+09
20uM	1.2E+11	8.2E+09	1.1E+11	7.4E+09
40uM	8.1E+10	5.7E+09	7.3E+10	5.2E+09

FIG. 19B



	Average	
	vp/mL	std.dev
Negative	2.5E+11	2.5E+10
Vehicle Control	5.5E+11	2.5E+10
50uM	2.8E+11	1.1E+11
100uM	2.2E+11	1.6E+10
200uM	5.4E+10	1.2E+10
400uM	<LOQ	<LOQ

FIG. 20A



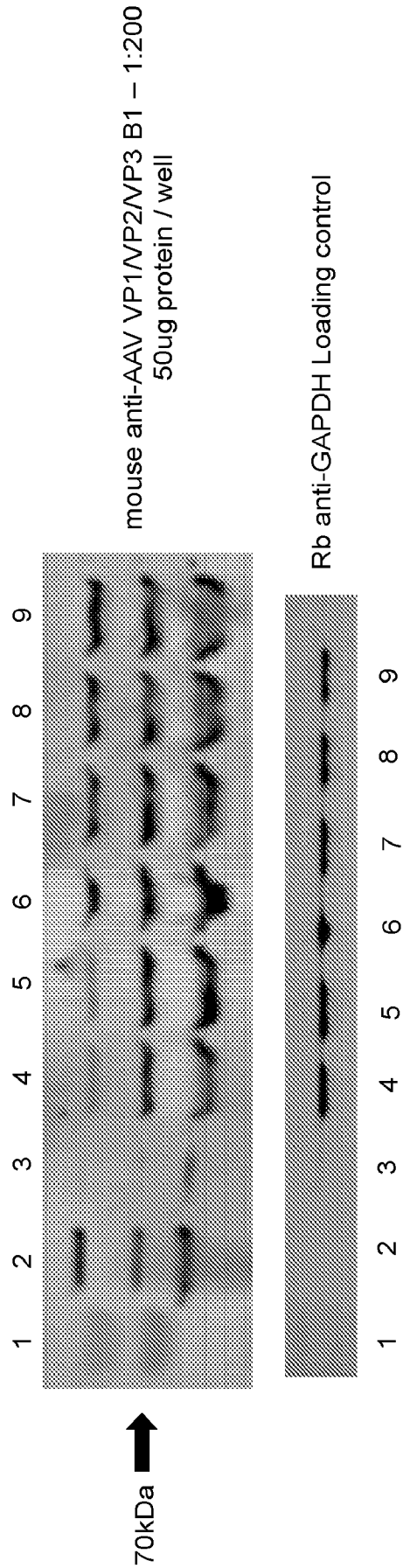
	Average	
	vp/mL	std.dev
Negative	3.2E+11	1.6E+10
Vehicle Control	4.5E+11	9.4E+09
5uM	7.1E+11	9.4E+09
10uM	1.2E+12	3.8E+10
20uM	1.7E+12	9.8E+10
40uM	1.9E+12	6.6E+10

FIG. 20B



LANE	Long ID	Capsid vp/mL
1	Marker	--
2	1.25e10vg – Purified AAV2	1.25e10/well
3	EMPTY	--
4	K-Negative	50
5	K-VC	50
6	K-50uM	50
7	K-100uM	50
8	K-200uM	50
9	K-400uM	30
10	EMPTY	--
11	K-50uM	50

FIG. 21A



LANE	Long ID	Total Protein (ug)
1	Marker	--
2	1.25e10vg – Purified AAV2	1.25e10/well
3	EMPTY	--
4	H-Negative	50
5	H-VC	50
6	H-5uM	50
7	H-10uM	50
8	H-20uM	50
9	H-40uM	50

FIG. 21B

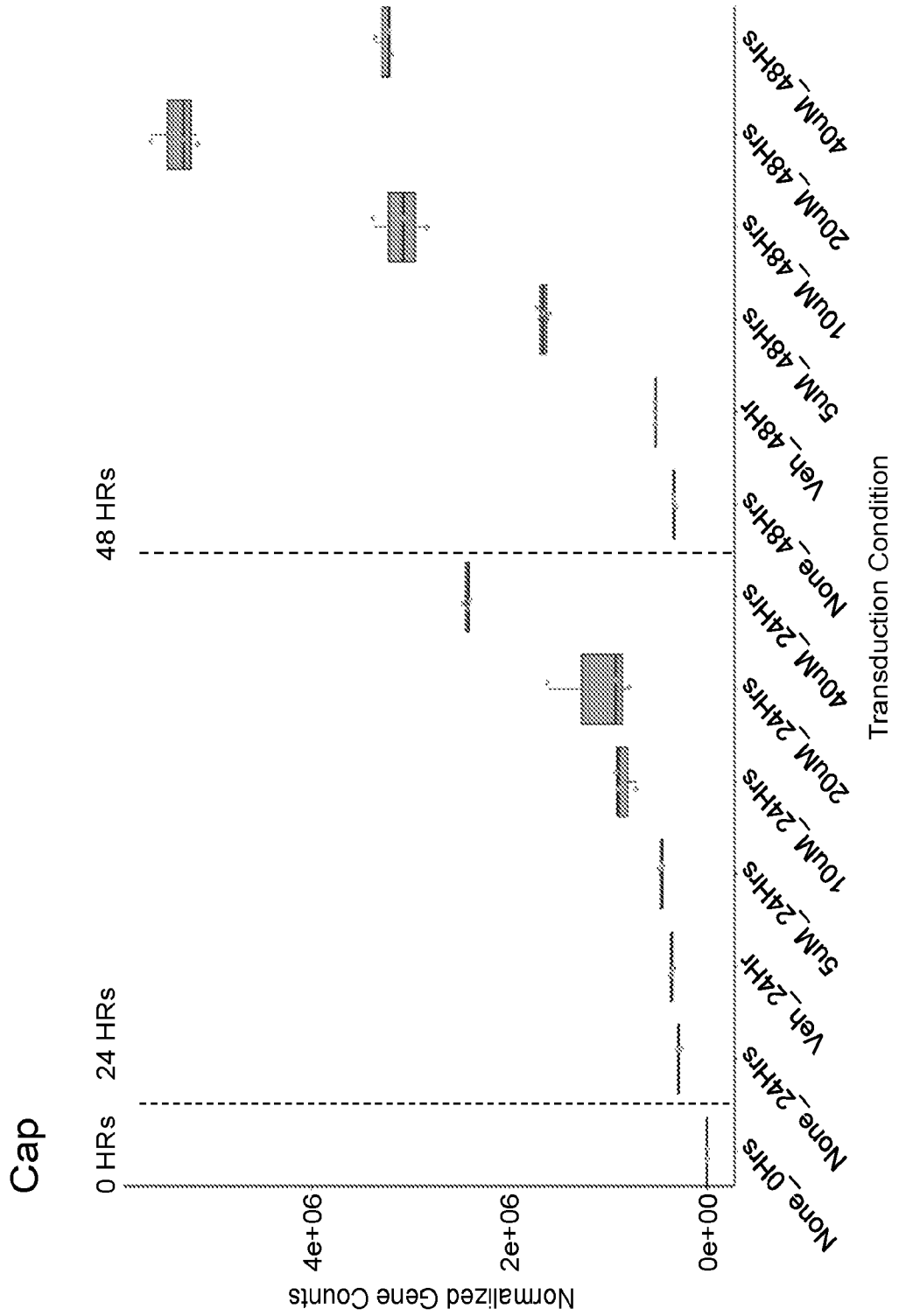

Cap
0 HRs
24 HRs
48 HRs
Normalized Gene Counts
0e+00
2e+06
4e+06
None_0Hrs
None_24Hrs
Veh_24Hr
5uM_24Hrs
10uM_24Hrs
20uM_24Hrs
40uM_24Hrs
None_48Hrs
Veh_48Hr
5uM_48Hrs
10uM_48Hrs
20uM_48Hrs
40uM_48Hrs
Transduction Condition

FIG. 22A

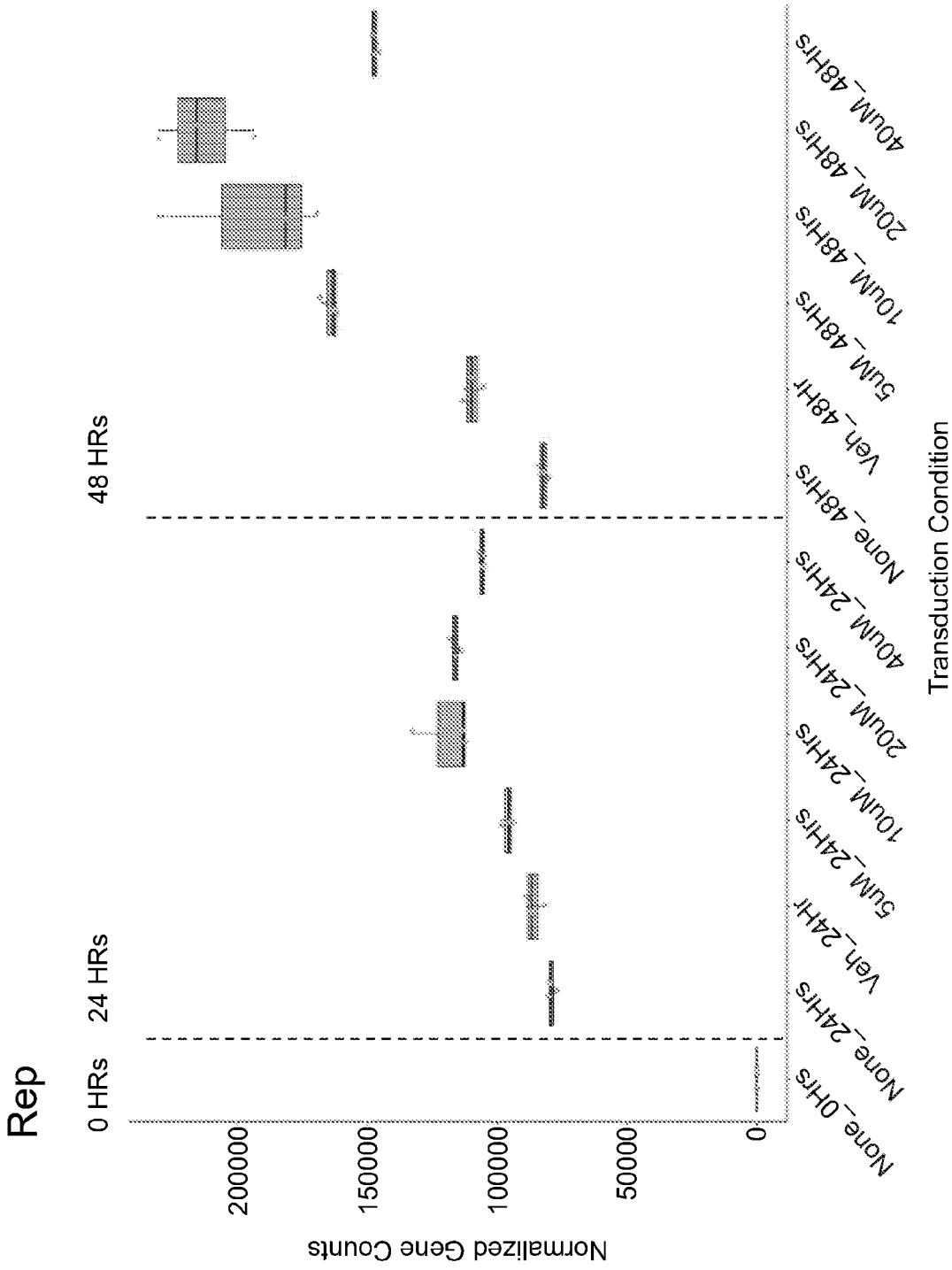

Rep
0 HRs
24 HRs
48 HRs
Normalized Gene Counts
200000
150000
100000
50000
0
None_0Hrs
None_24Hrs
Veh_24Hr
5uM_24Hrs
10uM_24Hrs
20uM_24Hrs
40uM_24Hrs
None_48Hrs
Veh_48Hr
5uM_48Hrs
10uM_48Hrs
20uM_48Hrs
40uM_48Hrs
Transduction Condition

FIG. 22B

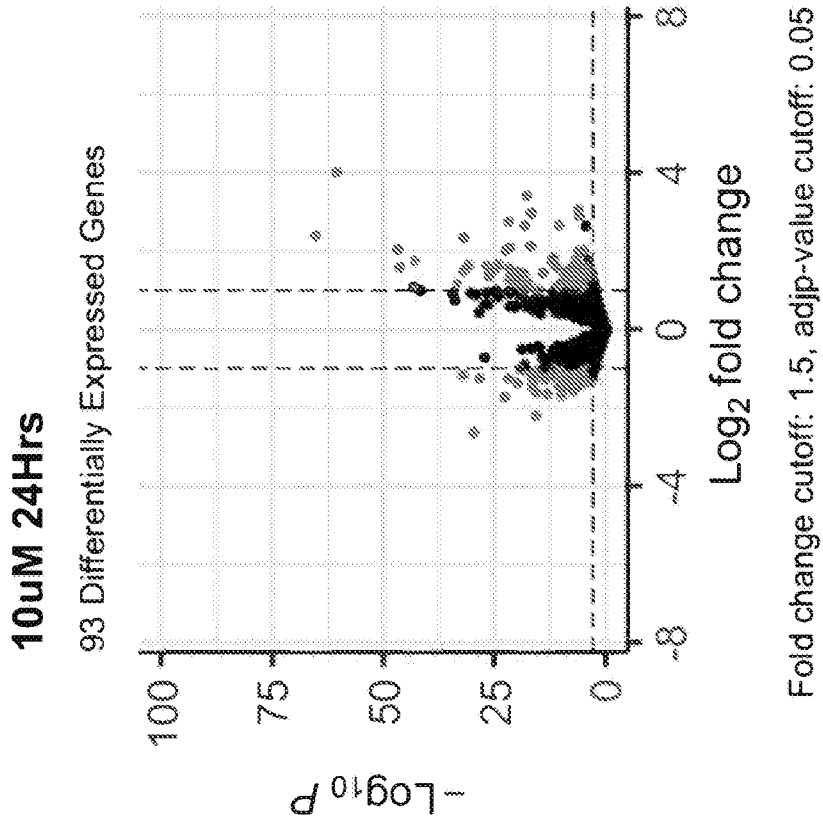


FIG. 23B

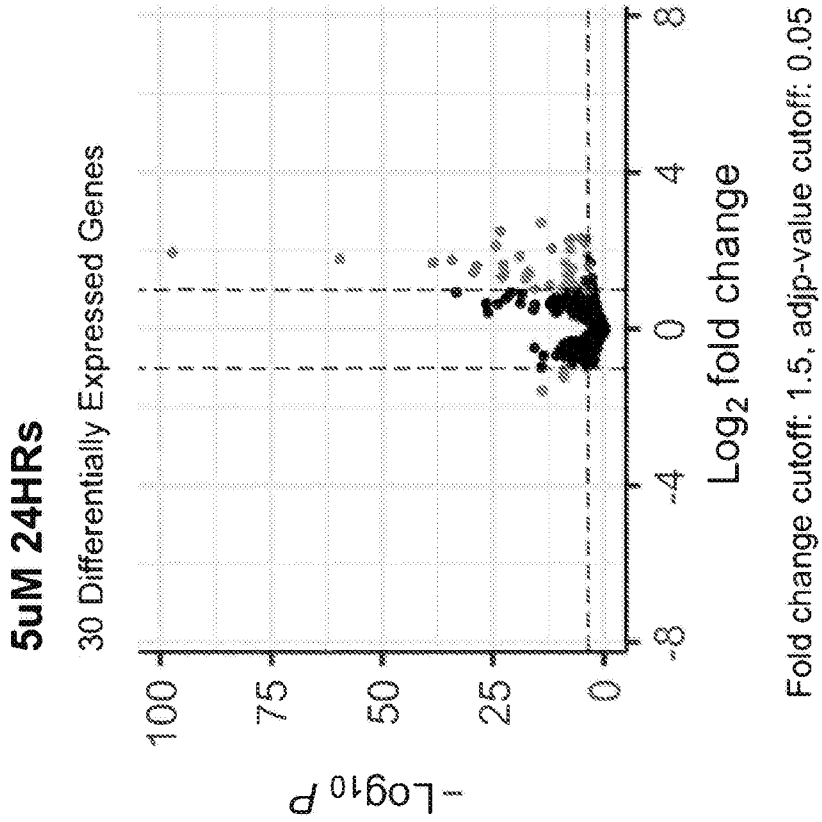


FIG. 23A

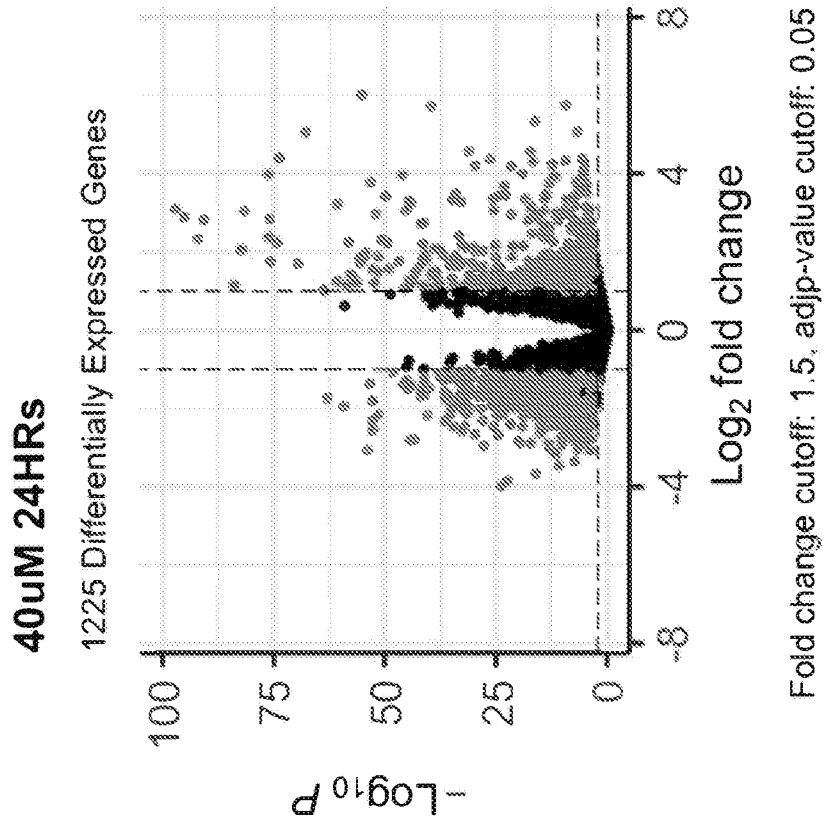


FIG. 23D

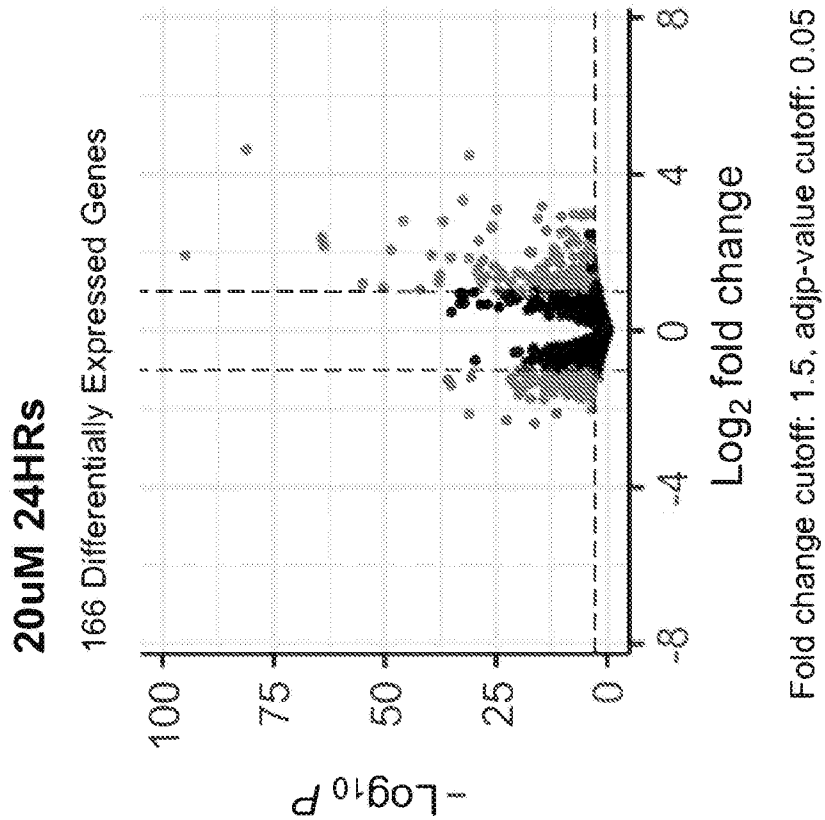


FIG. 23C

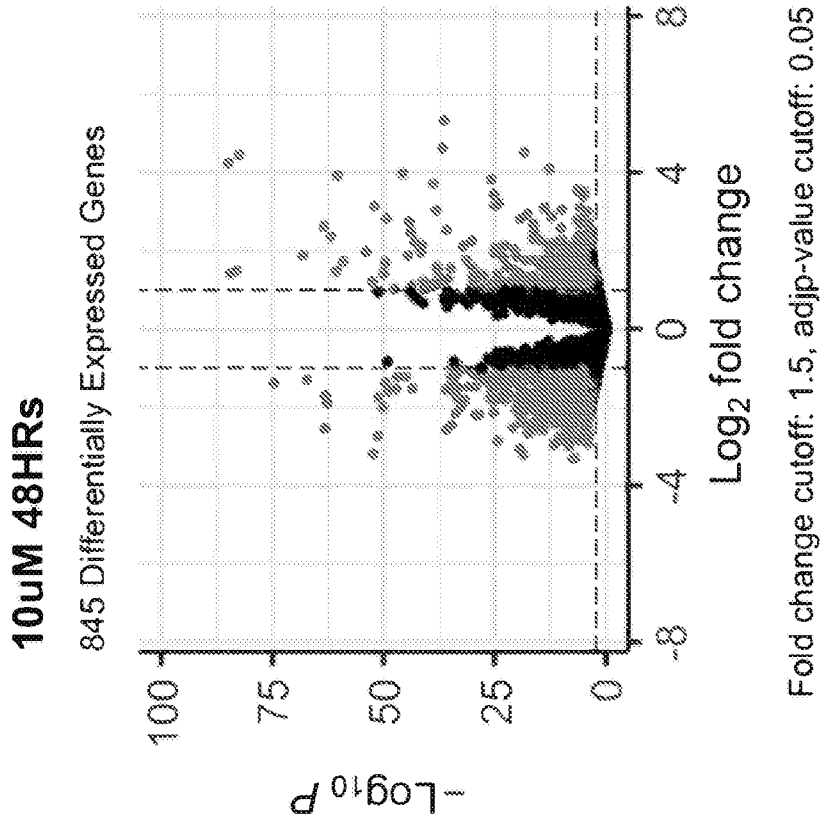


FIG. 23F

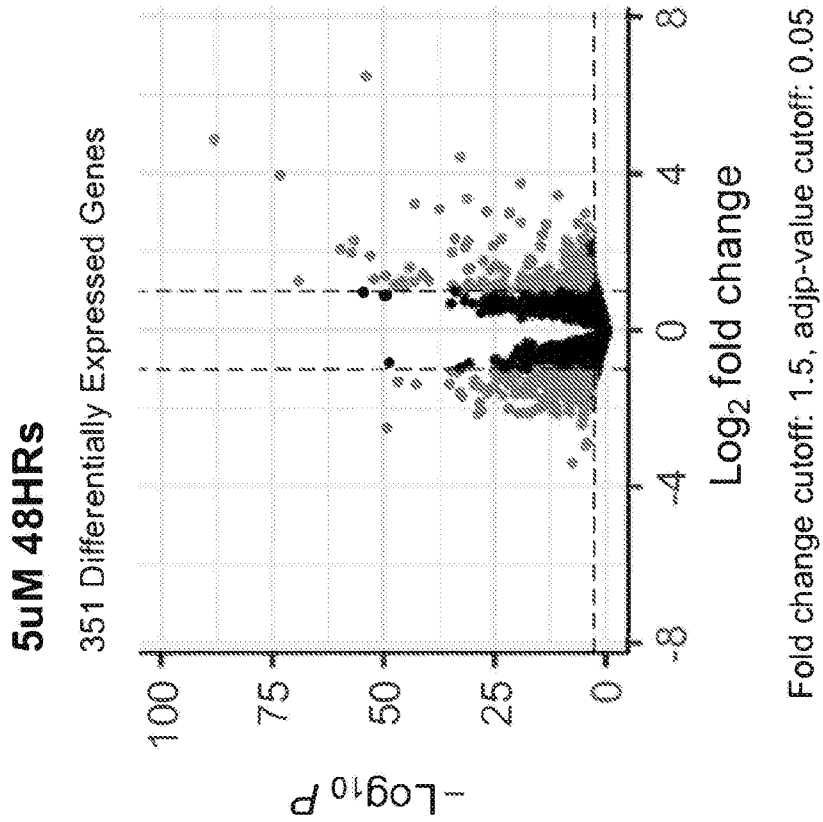


FIG. 23E

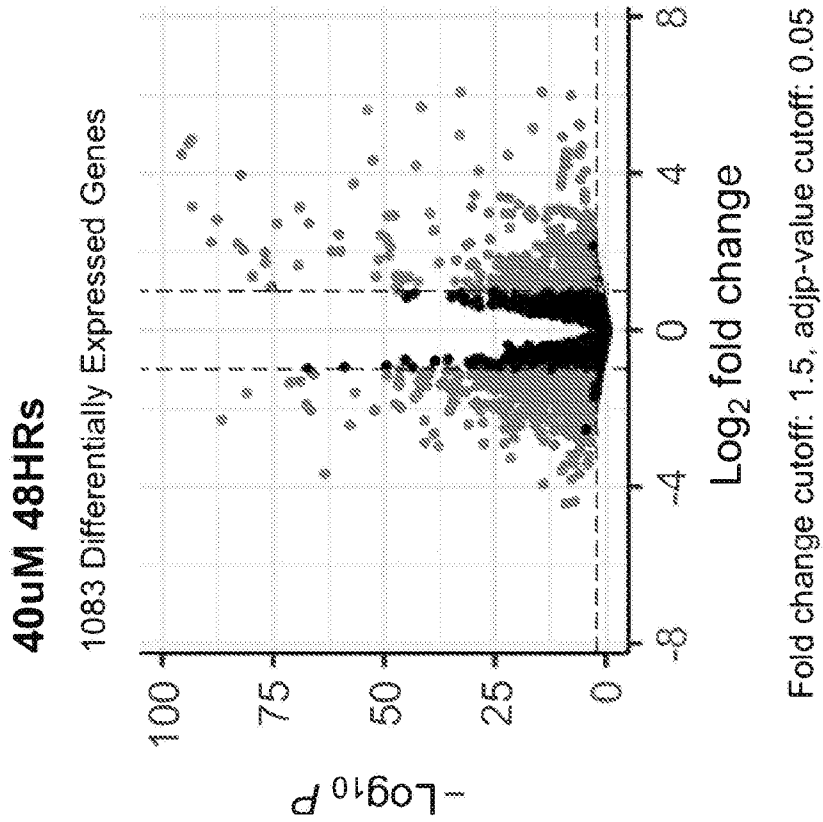


FIG. 23H

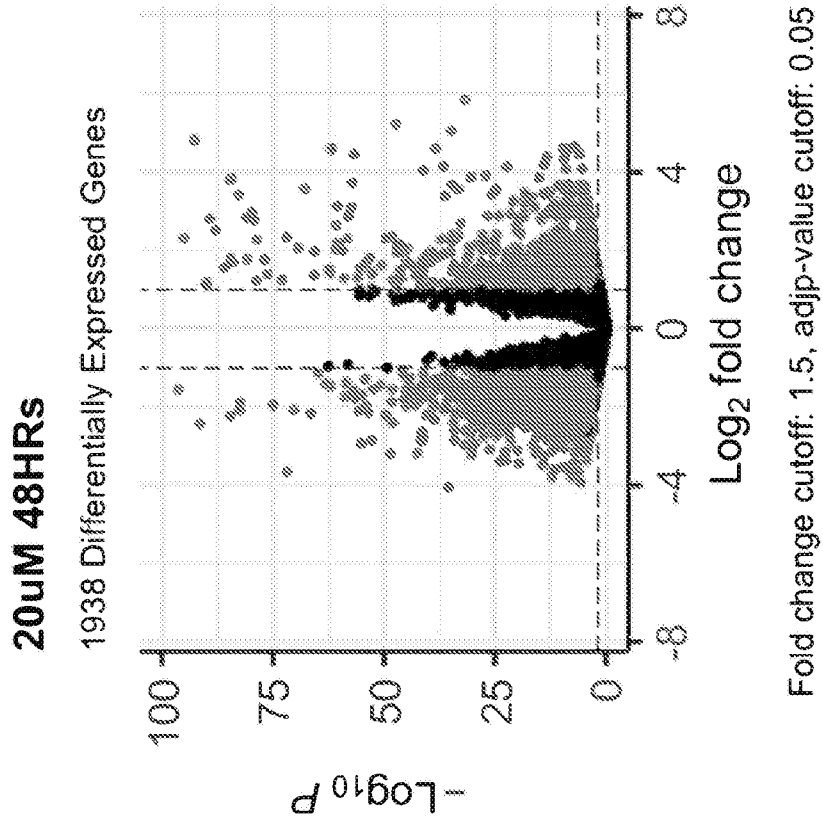


FIG. 23G

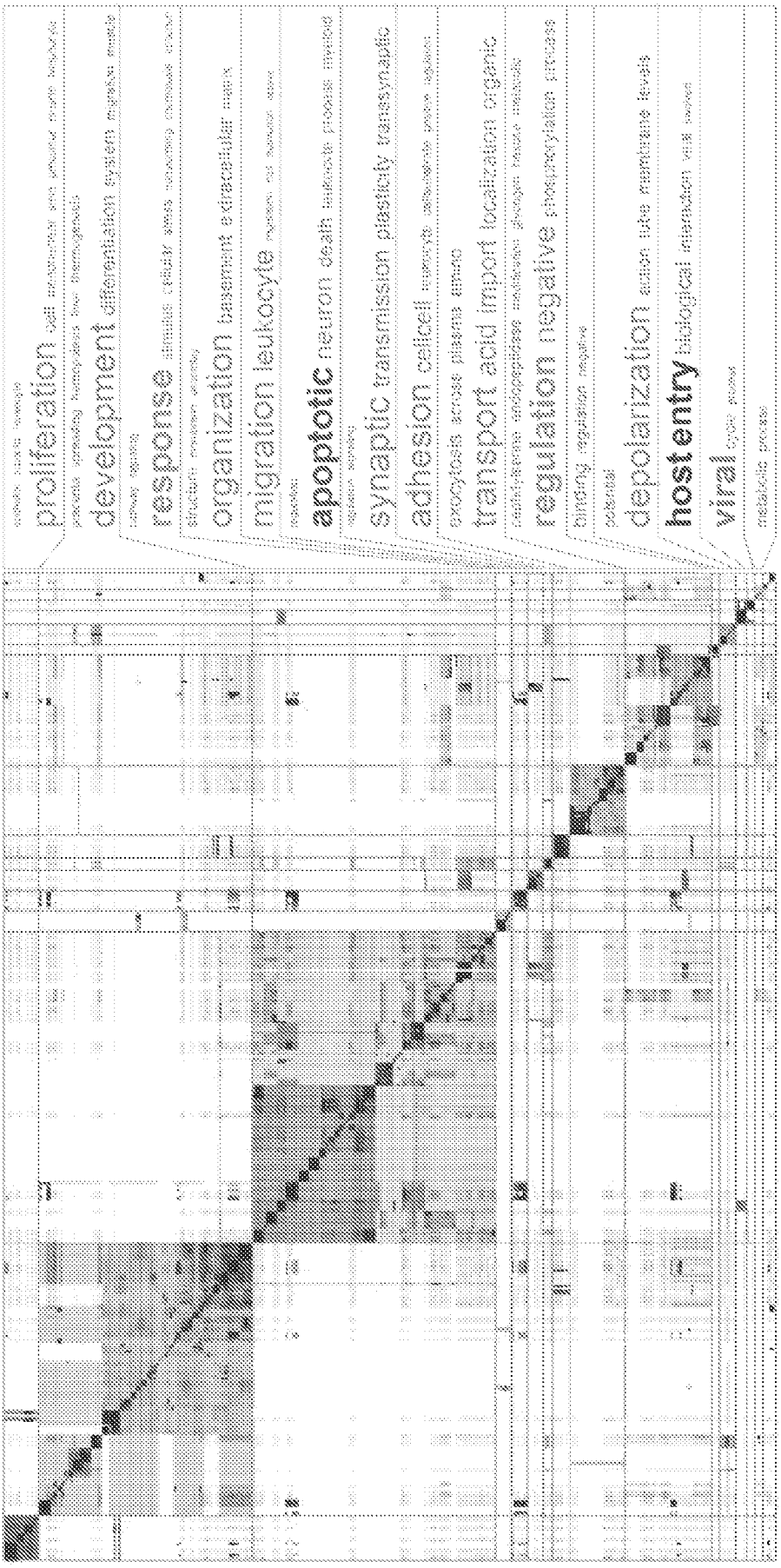

proliferation cell
development differentiation system
response
organization basement extracellular
migration leukocyte
apoptotic neuron death
synaptic transmission plasticity transsynaptic
adhesion cell
transport acid import localization organic
regulation negative
depolarization
hostentry
viral
Similarity
0.8
0.6
0.4
0.2
0

FIG. 24

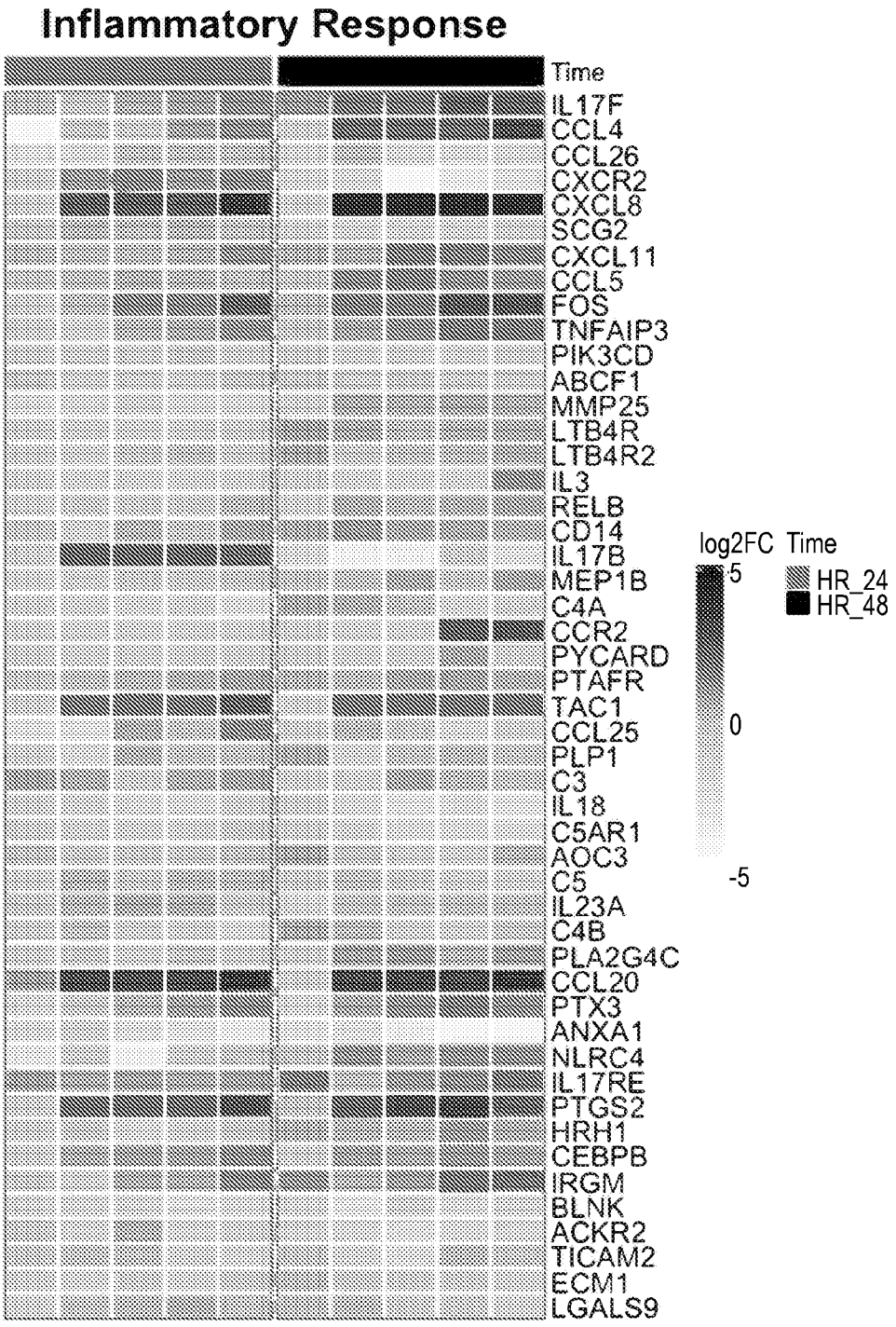

Inflammatory Response
Time
IL17F
CCL4
CCL26
CXCR2
CXCL8
SCG2
CXCL11
CCL5
FOS
TNFAIP3
PIK3CD
ABCF1
MMP25
LTB4R
LTB4R2
IL3
RELB
CD14
IL17B
MEP1B
C4A
CCR2
PYCARD
PTAFR
TAC1
CCL25
PLP1
C3
IL18
C5AR1
AOC3
C5
IL23A
C4B
PLA2G4C
CCL20
PTX3
ANXA1
NLRC4
IL17RE
PTGS2
HRH1
CEBPB
IRGM
BLNK
ACKR2
TICAM2
ECM1
LGALS9
log2FC
5
0
-5
Time
HR_24
HR_48

FIG. 25A

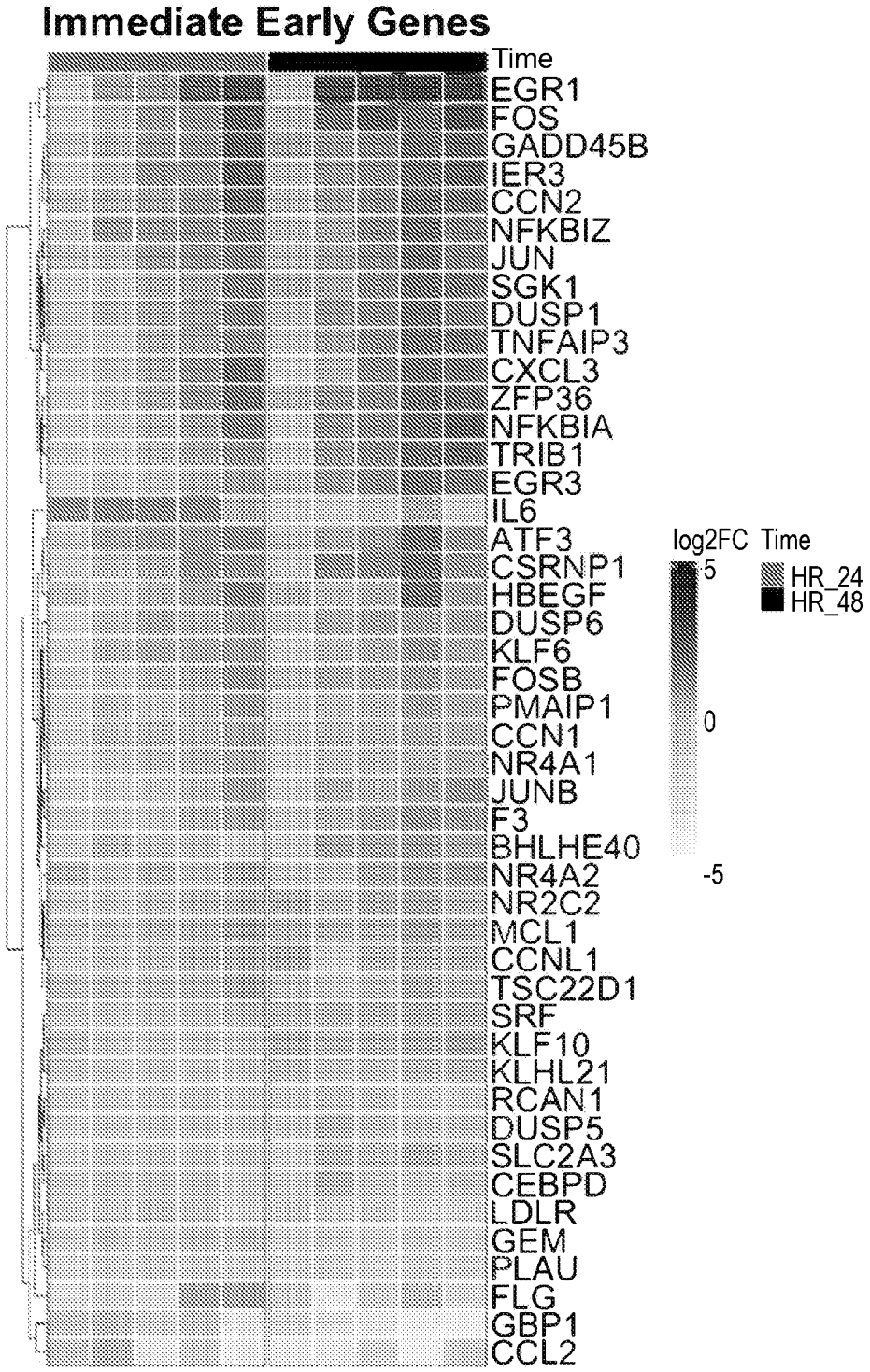

Immediate Early Genes
Time
EGR1
FOS
GADD45B
IER3
CCN2
NFKBIZ
JUN
SGK1
DUSP1
TNFAIP3
CXCL3
ZFP36
NFKBIA
TRIB1
EGR3
IL6
ATF3
CSRNP1
HBEGF
DUSP6
KLF6
FOSB
PMAIP1
CCN1
NR4A1
JUNB
F3
BHLHE40
NR4A2
NR2C2
MCL1
CCNL1
TSC22D1
SRF
KLF10
KLHL21
RCAN1
DUSP5
SLC2A3
CEBPD
LDLR
GEM
PLAU
FLG
GBP1
CCL2
log2FC
5
0
-5
Time
HR_24
HR_48

FIG. 25B

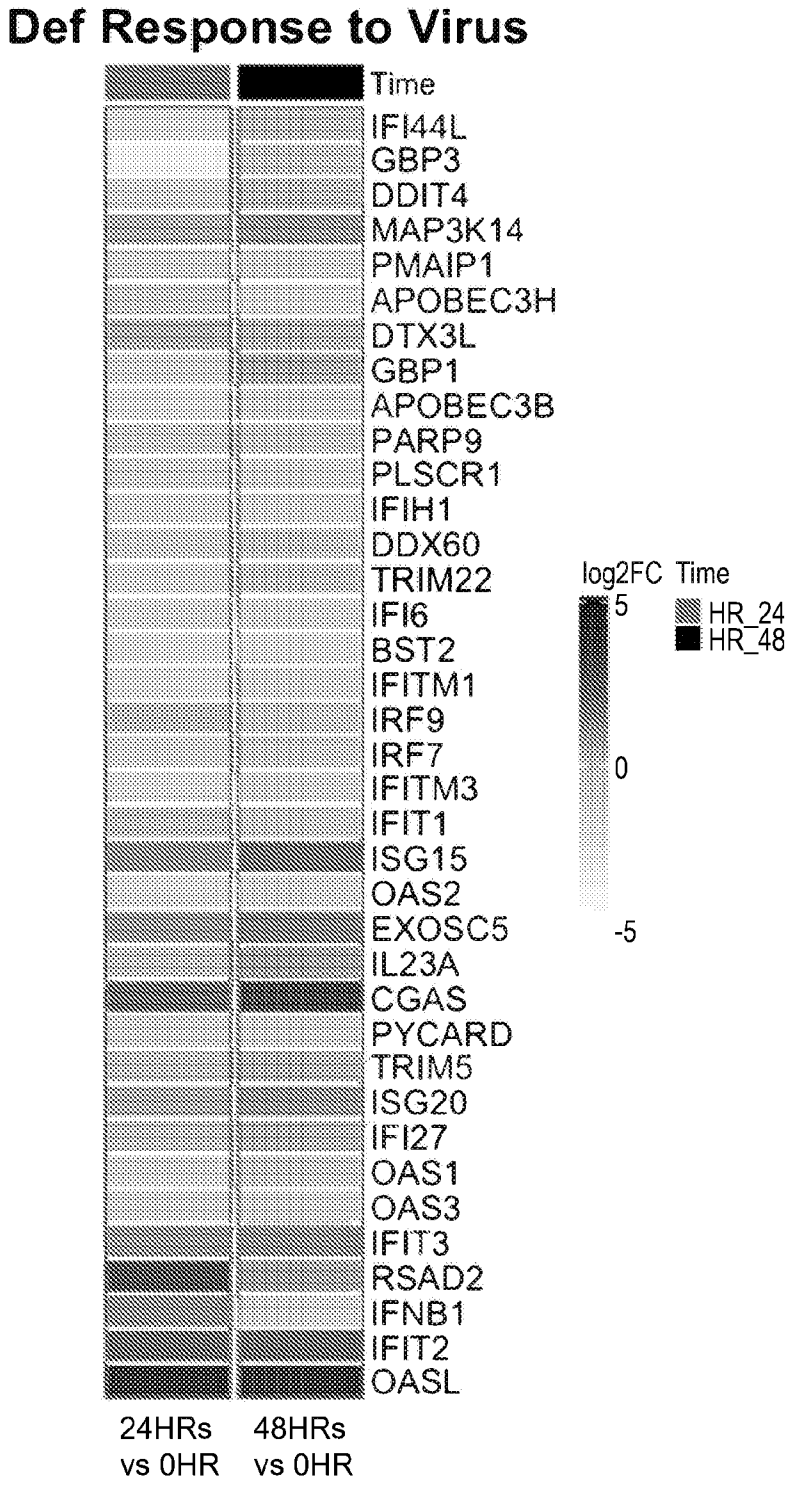

Def Response to Virus
Time
IFI44L
GBP3
DDIT4
MAP3K14
PMAIP1
APOBEC3H
DTX3L
GBP1
APOBEC3B
PARP9
PLSCR1
IFIH1
DDX60
TRIM22
IFI6
BST2
IFITM1
IRF9
IRF7
IFITM3
IFIT1
ISG15
OAS2
EXOSC5
IL23A
CGAS
PYCARD
TRIM5
ISG20
IFI27
OAS1
OAS3
IFIT3
RSAD2
IFNB1
IFIT2
OASL
log2FC
5
0
-5
Time
HR_24
HR_48
24HRs vs 0HR
48HRs vs 0HR

FIG. 25C

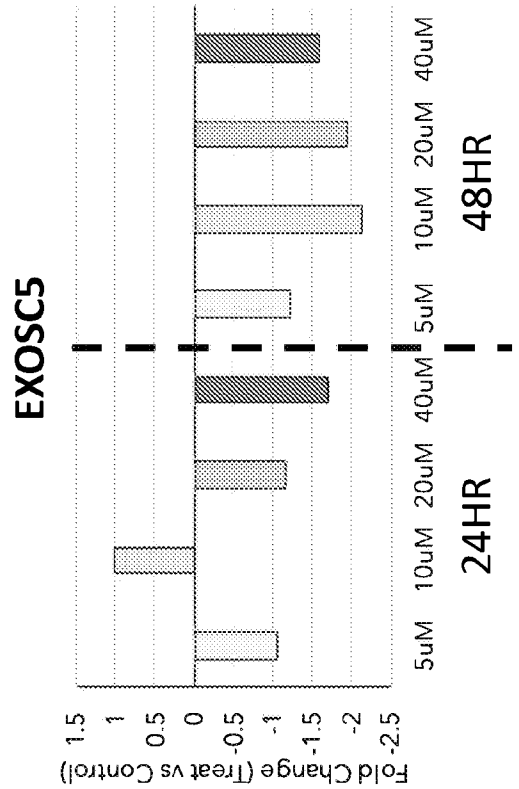


FIG. 25E

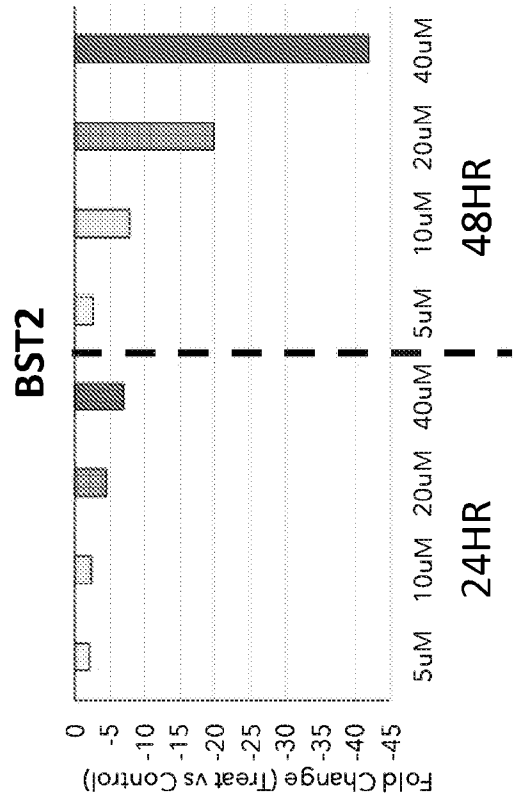


FIG. 25D

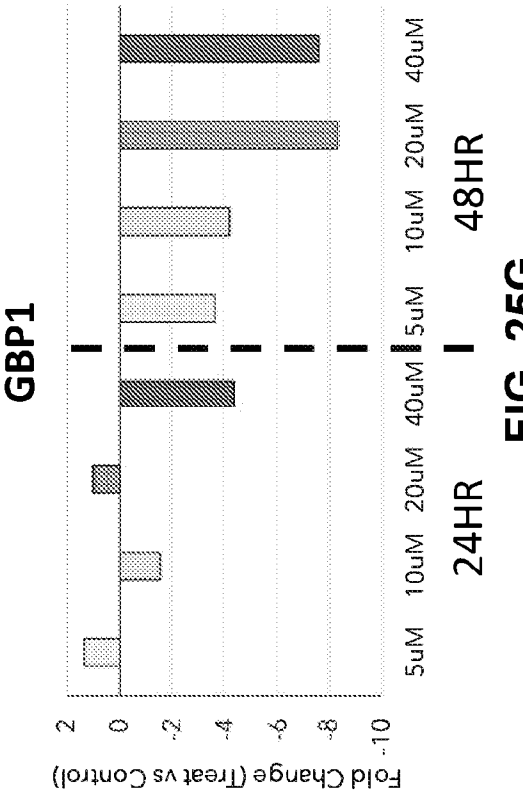


FIG. 25G

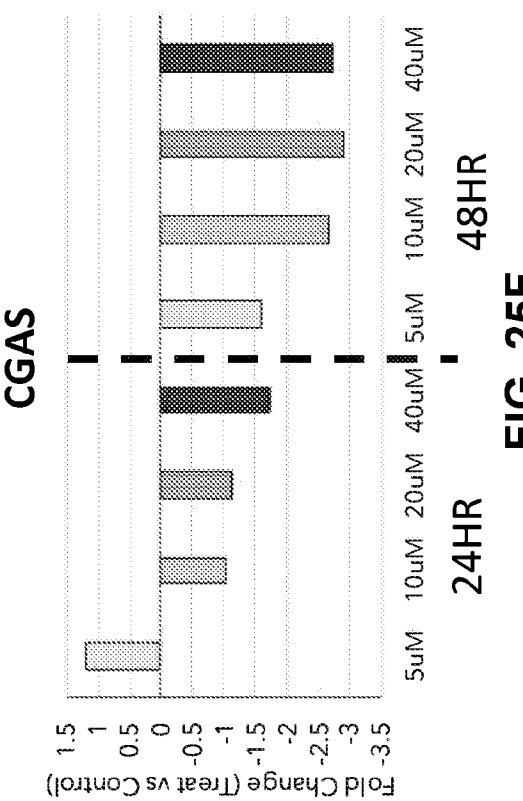


FIG. 25F

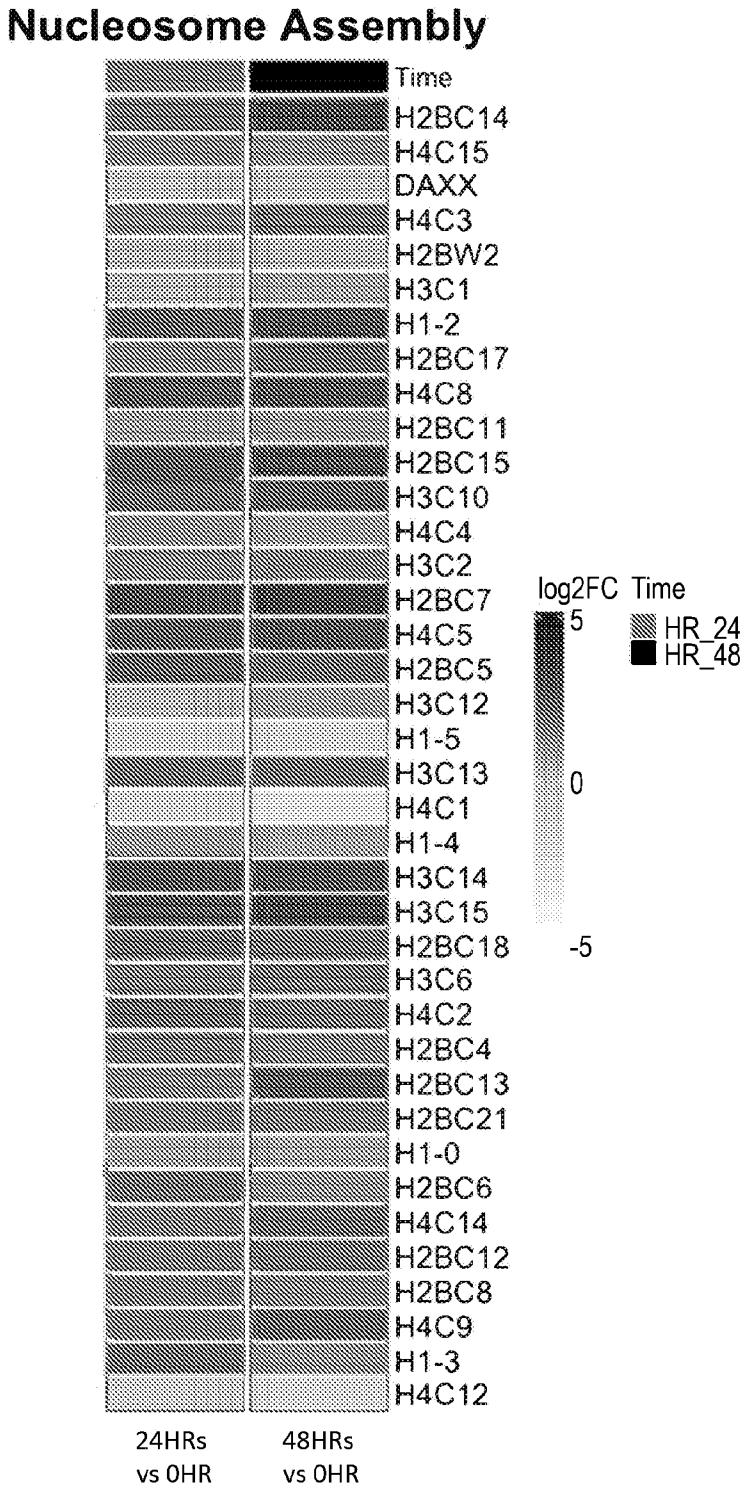

Nucleosome Assembly
Time
H2BC14
H4C15
DAXX
H4C3
H2BW2
H3C1
H1-2
H2BC17
H4C8
H2BC11
H2BC15
H3C10
H4C4
H3C2
H2BC7
H4C5
H2BC5
H3C12
H1-5
H3C13
H4C1
H1-4
H3C14
H3C15
H2BC18
H3C6
H4C2
H2BC4
H2BC13
H2BC21
H1-0
H2BC6
H4C14
H2BC12
H2BC8
H4C9
H1-3
H4C12
log2FC Time
5
0
-5
HR_24
HR_48
24HRs vs 0HR
48HRs vs 0HR

FIG. 26A

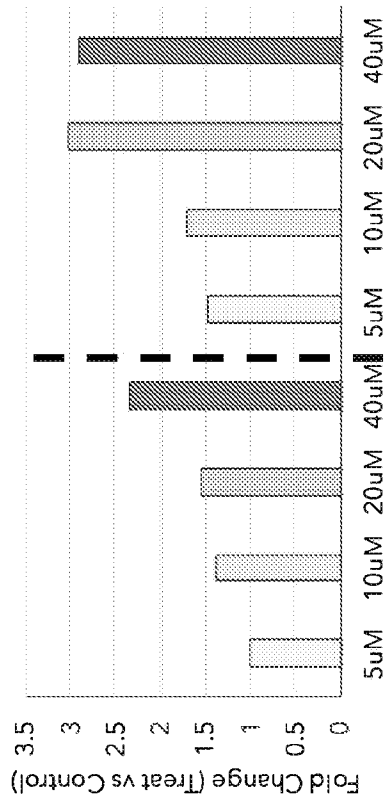


FIG. 26C

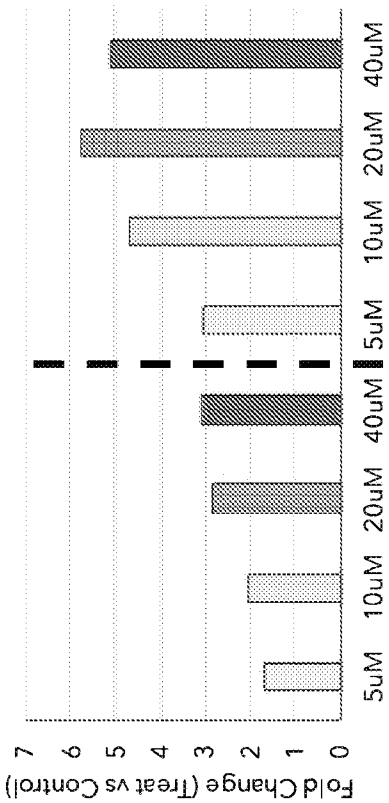


FIG. 26B

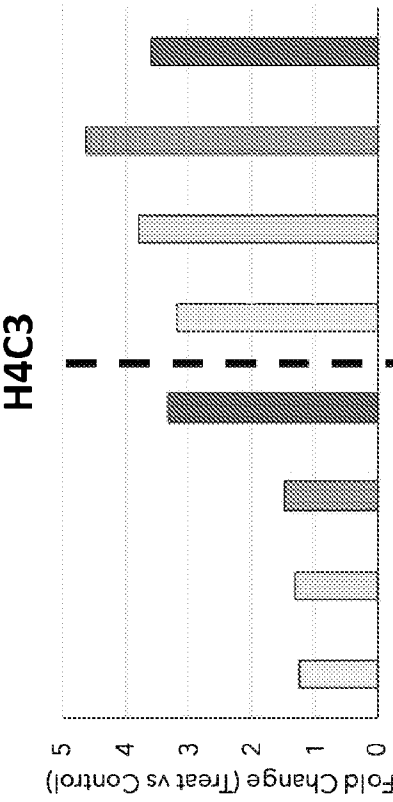


FIG. 26E

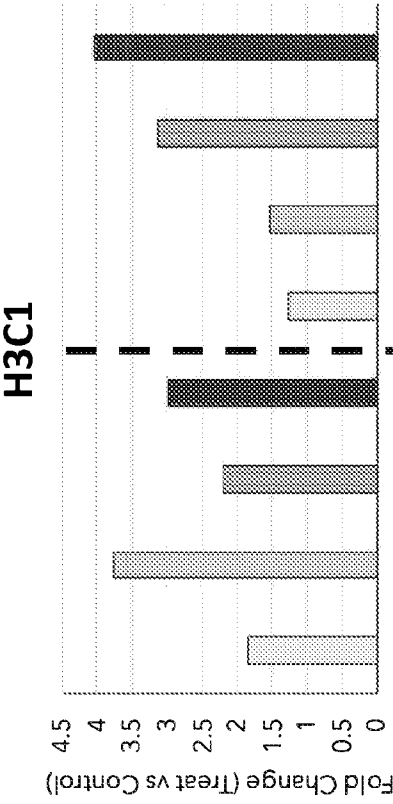


FIG. 26D

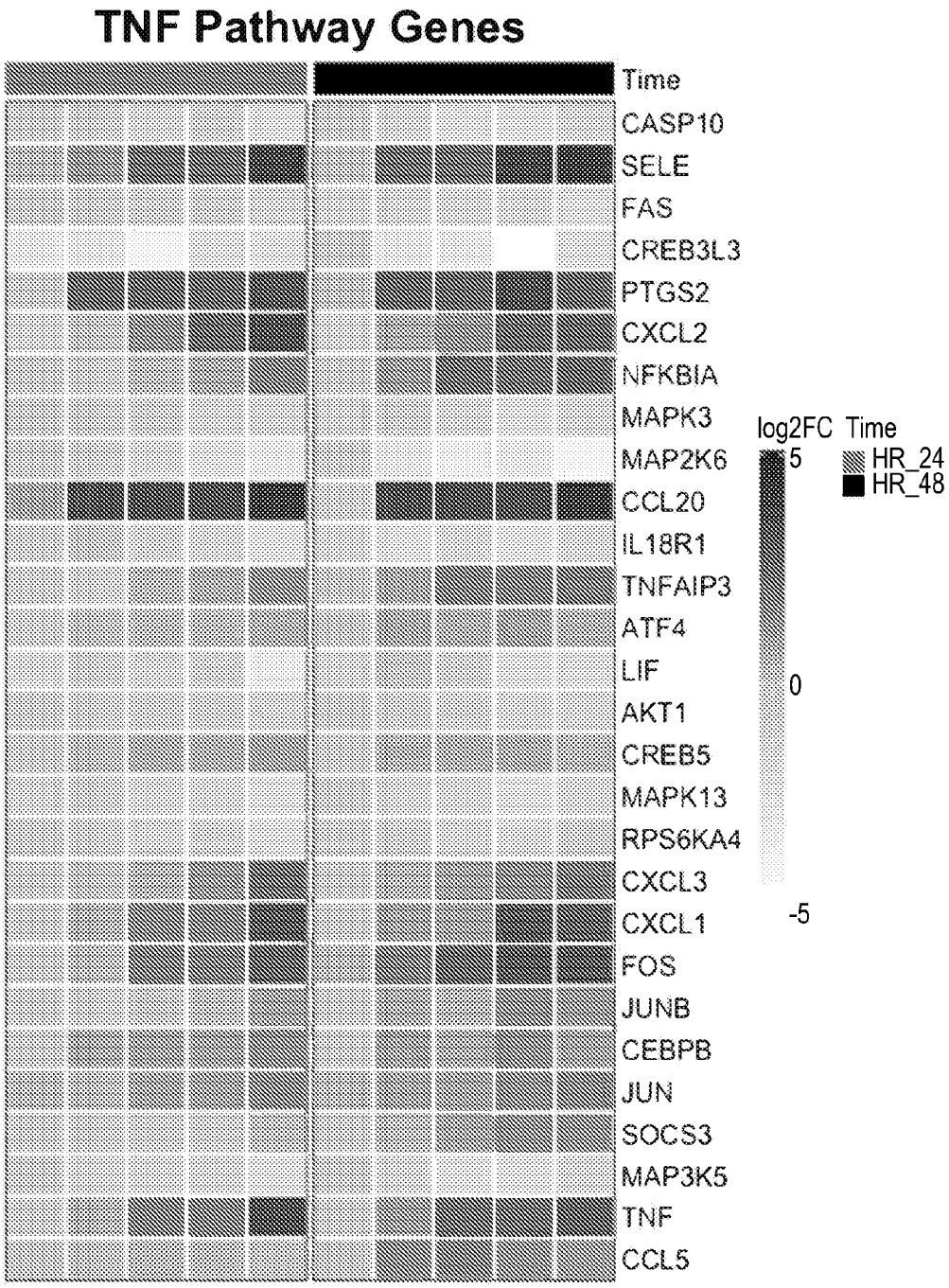

TNF Pathway Genes
Time
CASP10
SELE
FAS
CREB3L3
PTGS2
CXCL2
NFKBIA
MAPK3
MAP2K6
CCL20
IL18R1
TNFAIP3
ATF4
LIF
AKT1
CREB5
MAPK13
RPS6KA4
CXCL3
CXCL1
FOS
JUNB
CEBPB
JUN
SOCS3
MAP3K5
TNF
CCL5
log2FC
5
0
-5
Time
HR_24
HR_48

FIG. 27A

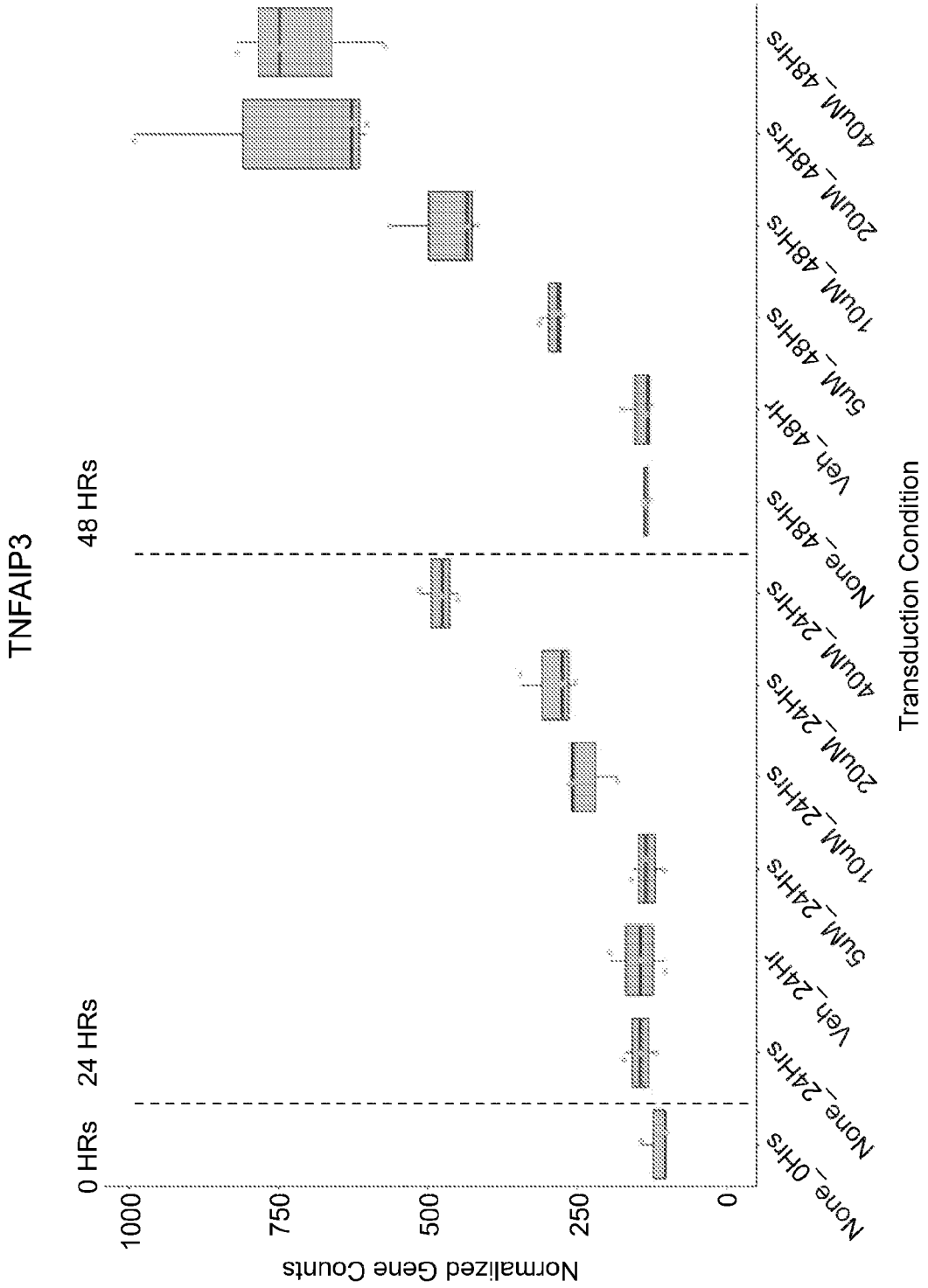

TNFAIP3
0 HRs
24 HRs
48 HRs
Normalized Gene Counts
1000
750
500
250
0
None_0Hrs
None_24Hrs
Veh_24Hr
5uM_24Hrs
10uM_24Hrs
20uM_24Hrs
40uM_24Hrs
None_48Hrs
Veh_48Hr
5uM_48Hrs
10uM_48Hrs
20uM_48Hrs
40uM_48Hrs
Transduction Condition

FIG. 27B

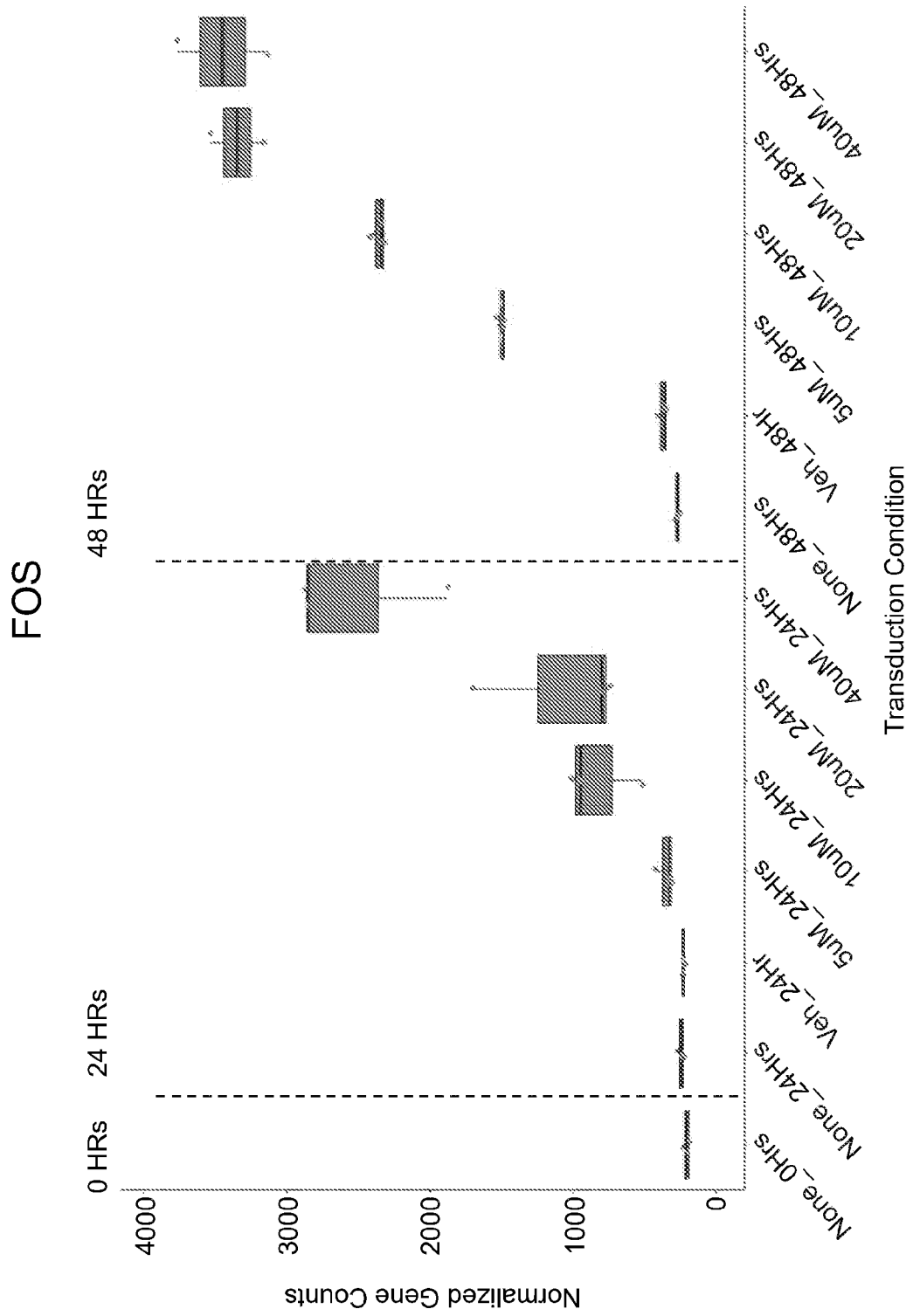

FOS
0 HRs
24 HRs
48 HRs
Normalized Gene Counts
4000
3000
2000
1000
0
None_0Hrs
None_24Hrs
Veh_24Hr
5uM_24Hrs
10uM_24Hrs
20uM_24Hrs
40uM_24Hrs
None_48Hrs
Veh_48Hr
5uM_48Hrs
10uM_48Hrs
20uM_48Hrs
40uM_48Hrs
Transduction Condition

FIG. 27C

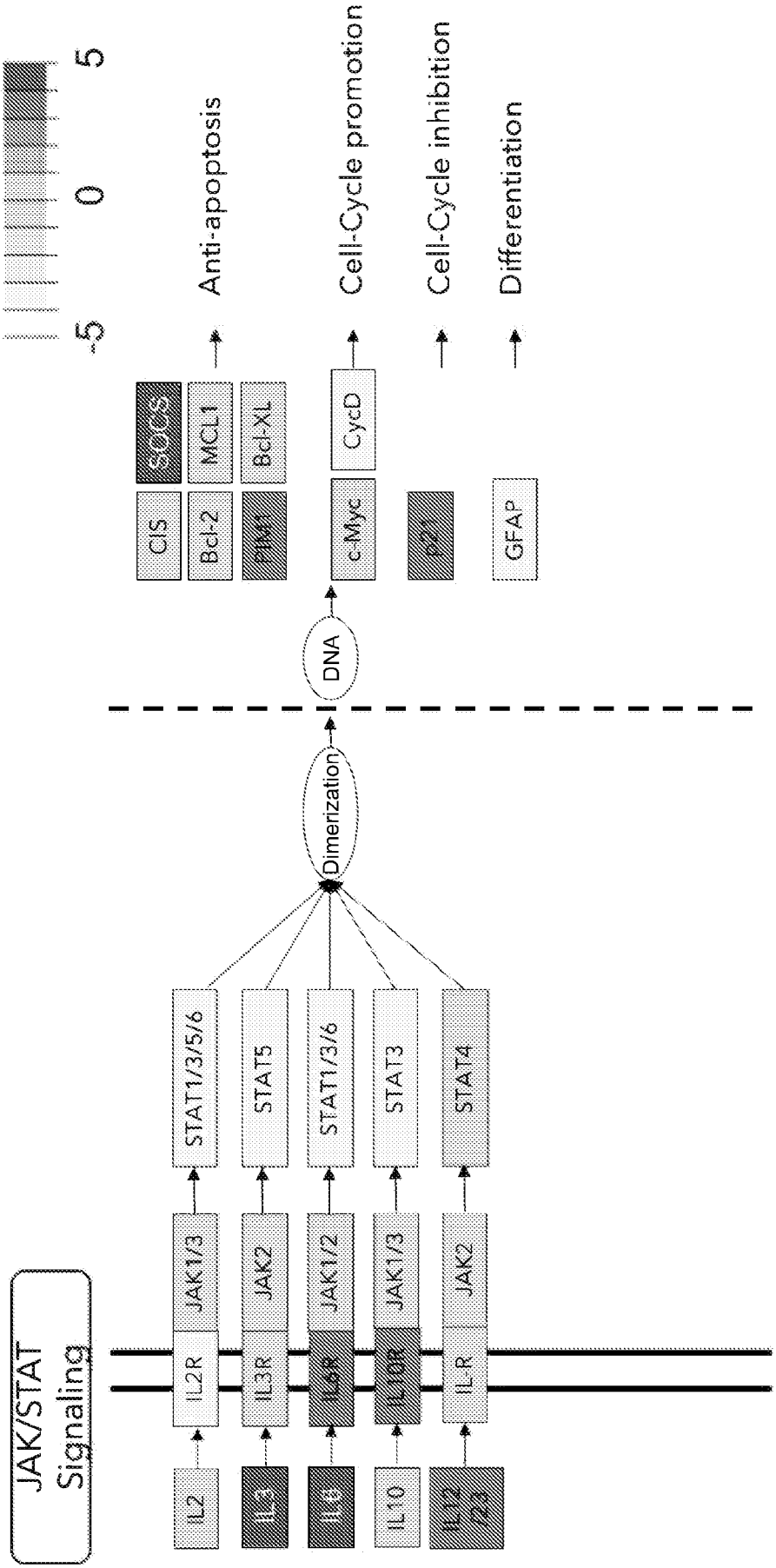

JAK/STAT Signaling
IL2
IL3
IL4
IL10
IL12 /23
IL2R
IL3R
IL4R
IL10R
IL-R
JAK1/3
JAK2
JAK1/2
JAK1/3
JAK2
STAT1/3/5/6
STAT5
STAT1/3/6
STAT3
STAT4
Dimerization
DNA
CIS
SOCS
Bcl-2
MCL1
Pim1
Bcl-XL
c-Myc
CycD
p21
GFAP
Anti-apoptosis
Cell-Cycle promotion
Cell-Cycle inhibition
Differentiation
-5
0
5

FIG. 28A

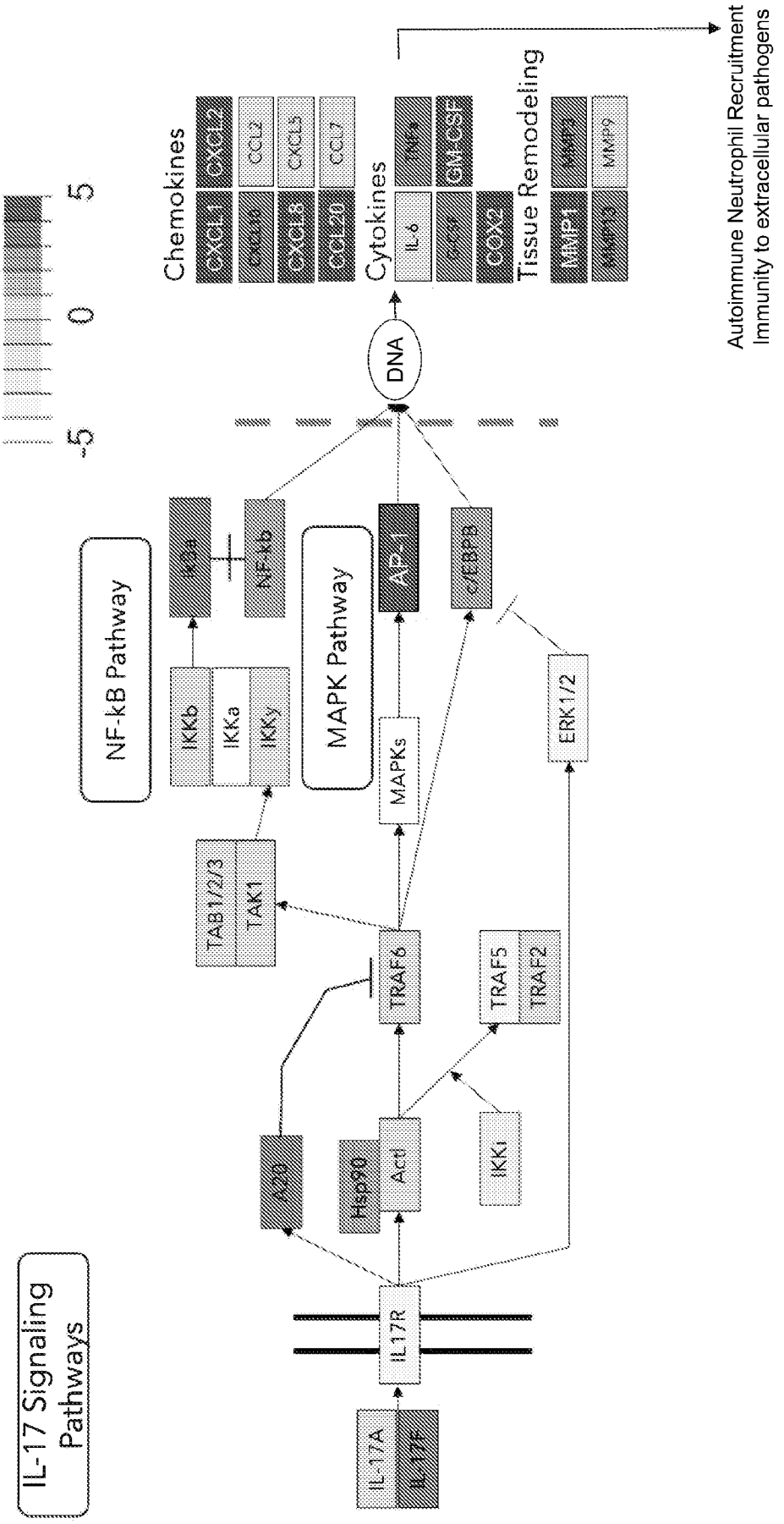

IL-17 Signaling Pathways
IL-17A
IL-17F
IL17R
Act1
Hsp90
A20
IKKi
TRAF6
TRAF5
TRAF2
TAB1/2/3
TAK1
IKKb
IKKa
IKKγ
MAPKs
ERK1/2
NF-kB Pathway
MAPK Pathway
p38
NF-kb
AP-1
cEBPβ
DNA
Chemokines
CXCL1
CXCL2
CXCL10
CCL2
CXCL5
CXCL8
CCL20
CCL7
Cytokines
IL6
TNFa
G-CSF
GM-CSF
COX2
Tissue Remodeling
MMP1
MMP3
MMP13
MMP9
-5
0
5
Autoimmune Neutrophil Recruitment
Immunity to extracellular pathogens

FIG. 28B

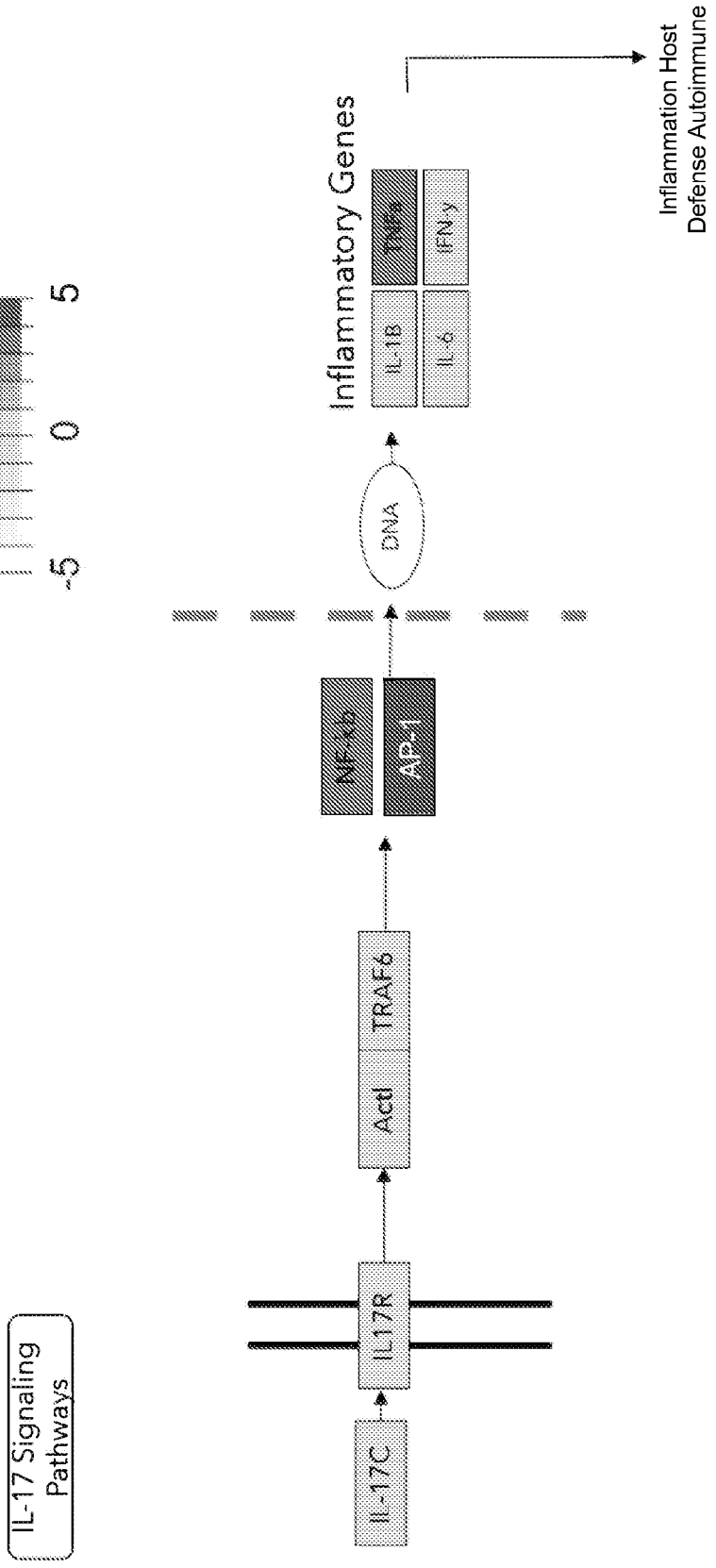

IL-17 Signaling Pathways
IL-17C
IL17R
Act1
TRAF6
NF-κB
AP-1
DNA
Inflammatory Genes
IL-1B
IL-6
TNFa
IFN-γ
Inflammation Host Defense Autoimmune
5
0
-5

FIG. 28B (cont.)

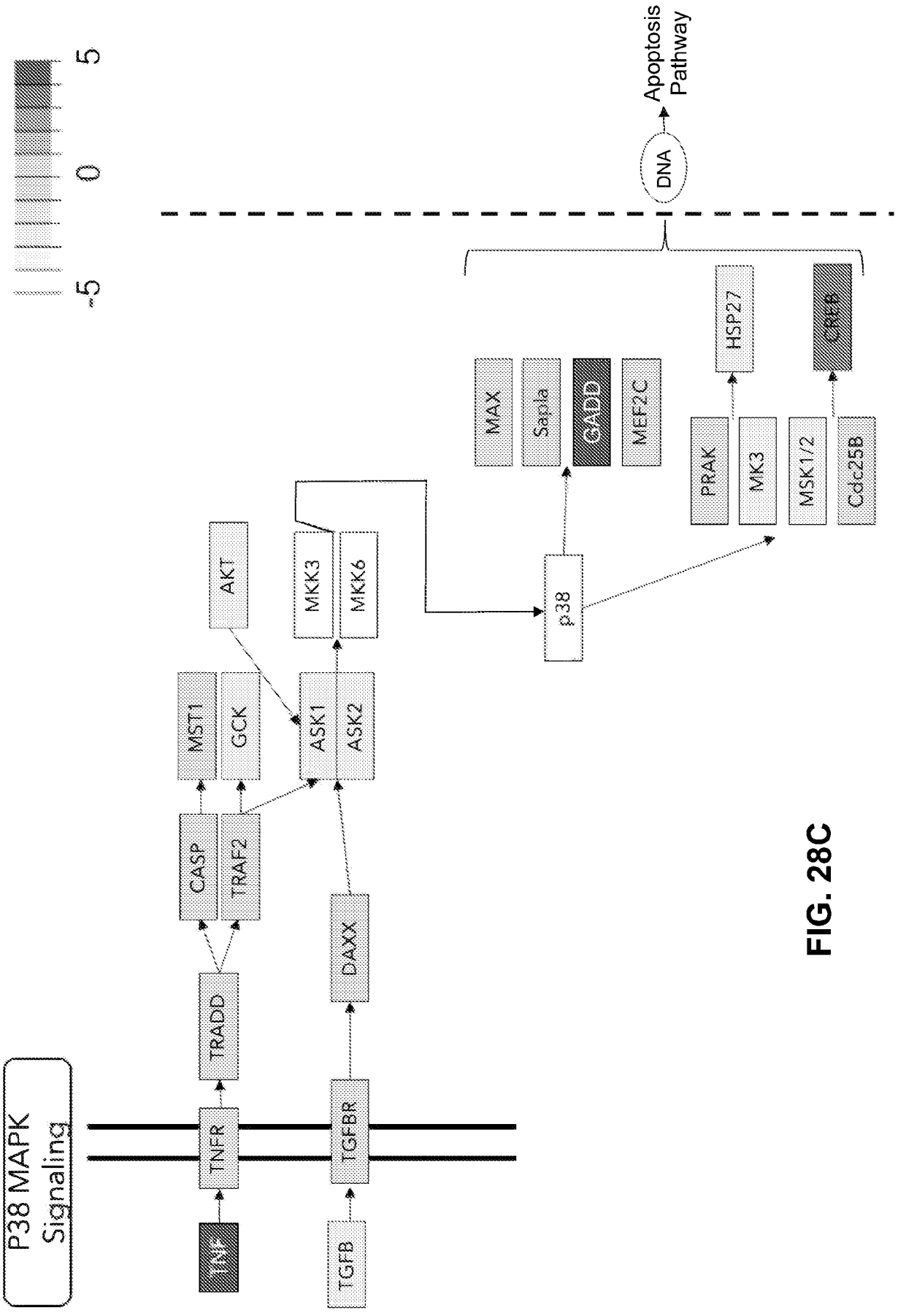

P38 MAPK Signaling
TNF
TGFB
TNFR
TGFBR
TRADD
DAXX
CASP
TRAF2
MST1
GCK
ASK1
ASK2
AKT
MKK3
MKK6
p38
MAX
Sap1a
cJUN
MEF2C
PRAK
MK3
MSK1/2
Cdc25B
HSP27
CREB
DNA
Apoptosis Pathway
5
0
-5

FIG. 28C

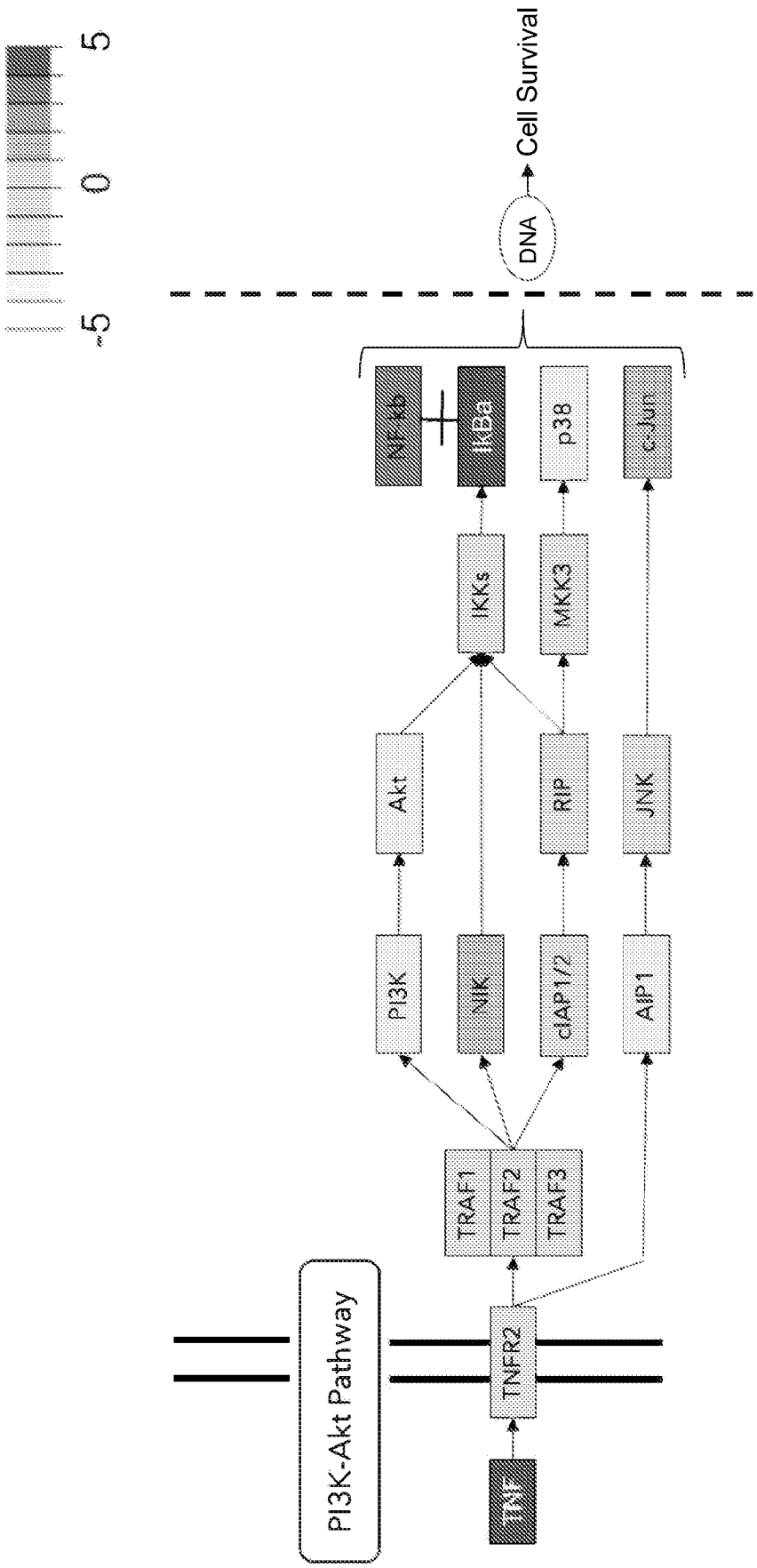

5
0
-5
Cell Survival
DNA
NFκb
IκBα
p38
c-Jun
IKKs
MKK3
Akt
RIP
JNK
PI3K
NIK
cIAP1/2
AIP1
TRAF1
TRAF2
TRAF3
TNFR2
TNF
PI3K-Akt Pathway

FIG. 28C (cont.)

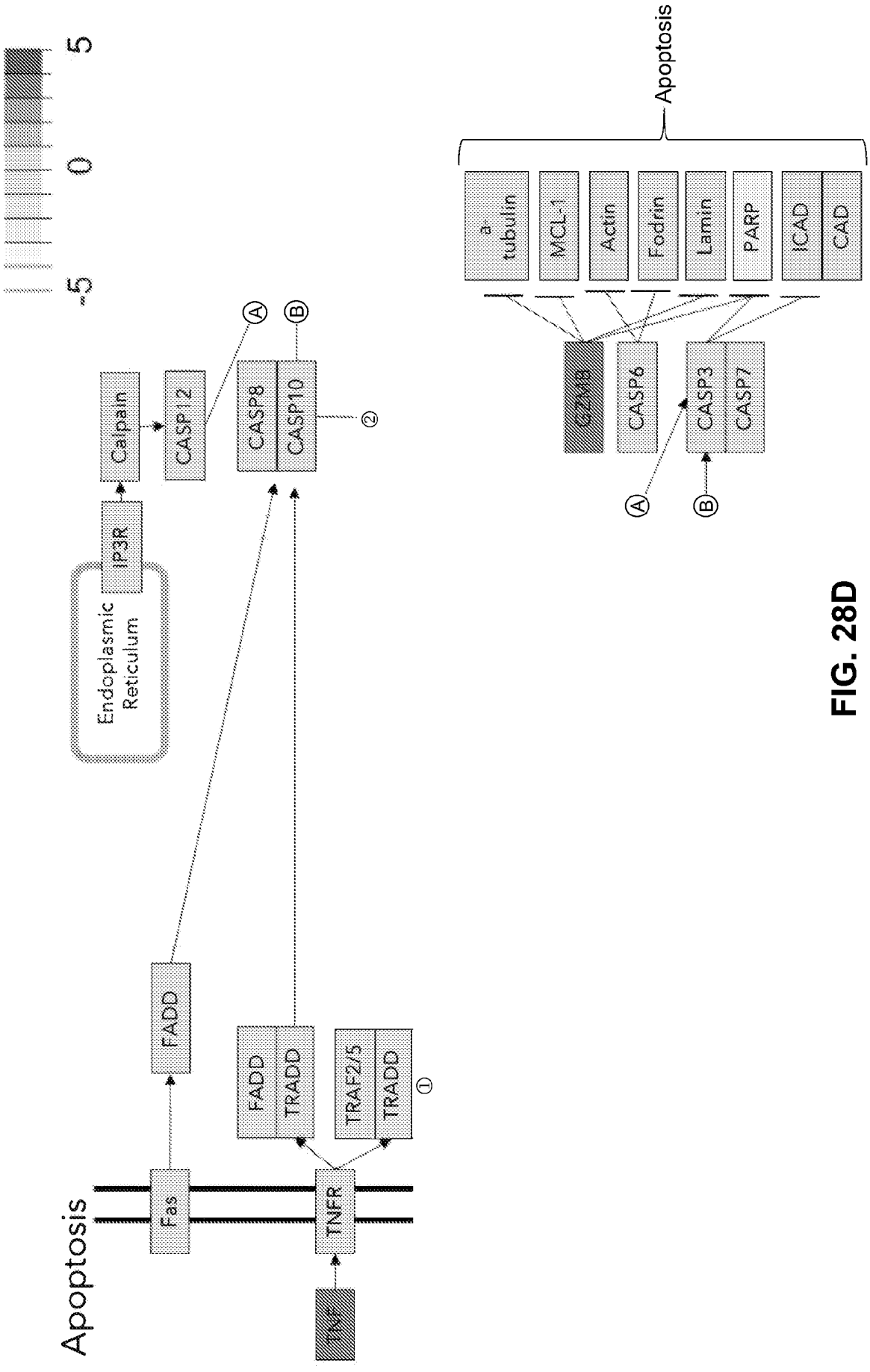

5
0
-5
Apoptosis
a-tubulin
MCL-1
Actin
Fodrin
Lamin
PARP
ICAD
CAD
GZMB
CASP6
CASP3
CASP7
A
B
Calpain
CASP12
CASP8
CASP10
IP3R
Endoplasmic Reticulum
FADD
FADD
TRADD
TRAF2/5
TRADD
Fas
TNFR
TNF
Apoptosis

FIG. 28D

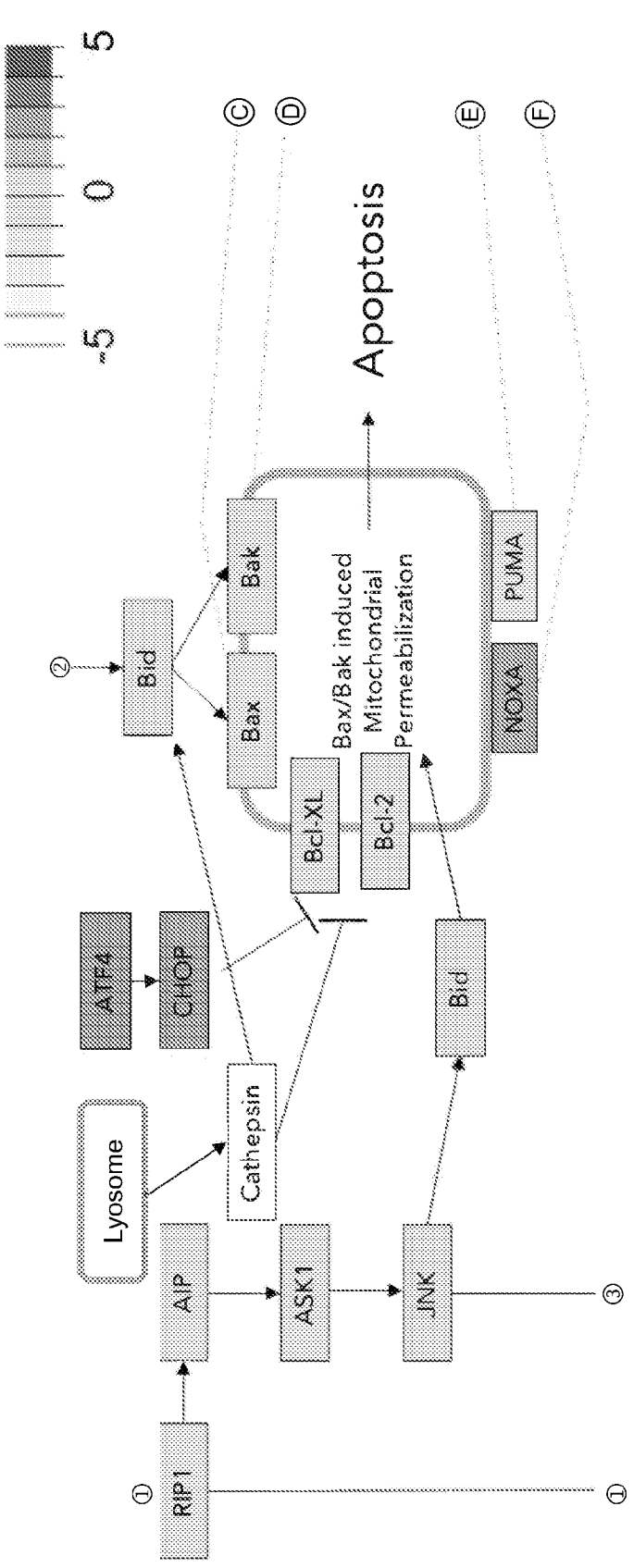

5
0
-5
Apoptosis
Pro-Apoptotic Genes
Bak
Bax
Apaf-1
PIDD
MOXA
TRAIL R2
p53AIP1
Bid
Fas
PUMA
RIP1
AIP
ASK1
JNK
Lyosome
Cathepsin
ATF4
CHOP
Bid
Bcl-XL
Bcl-2
Bax/Bak induced
Mitochondrial
Permeabilization
Bid
Bax
Bak
PUMA
MOXA
①
②
③
Ⓒ
Ⓓ
Ⓔ
Ⓕ

FIG. 28D (cont.)

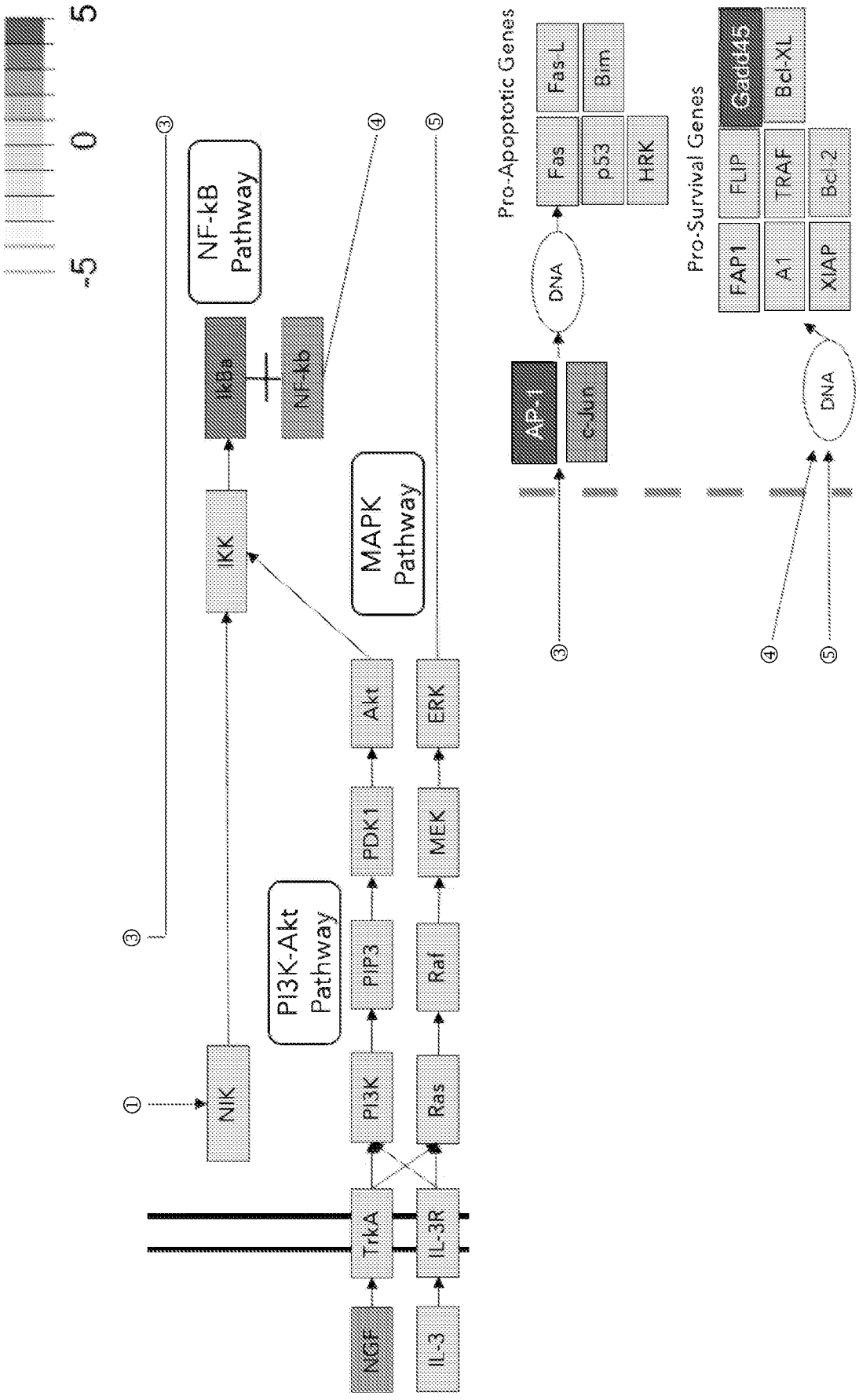

5
0
-5
NF-kB Pathway
MAPK Pathway
PI3K-Akt Pathway
NIK
IKK
IκBα
NF-kb
PI3K
Ras
PIP3
Raf
PDK1
MEK
Akt
ERK
TrkA
IL-3R
NGF
IL-3
①
③
④
⑤
Pro-Apoptotic Genes
Fas L
Bim
Fas
p53
HRK
DNA
AP-1
c-Jun
Pro-Survival Genes
Gadd45
FLIP
Bcl-XL
TRAF
FAP1
A1
Bcl-2
XIAP
DNA

FIG. 28D (cont.)

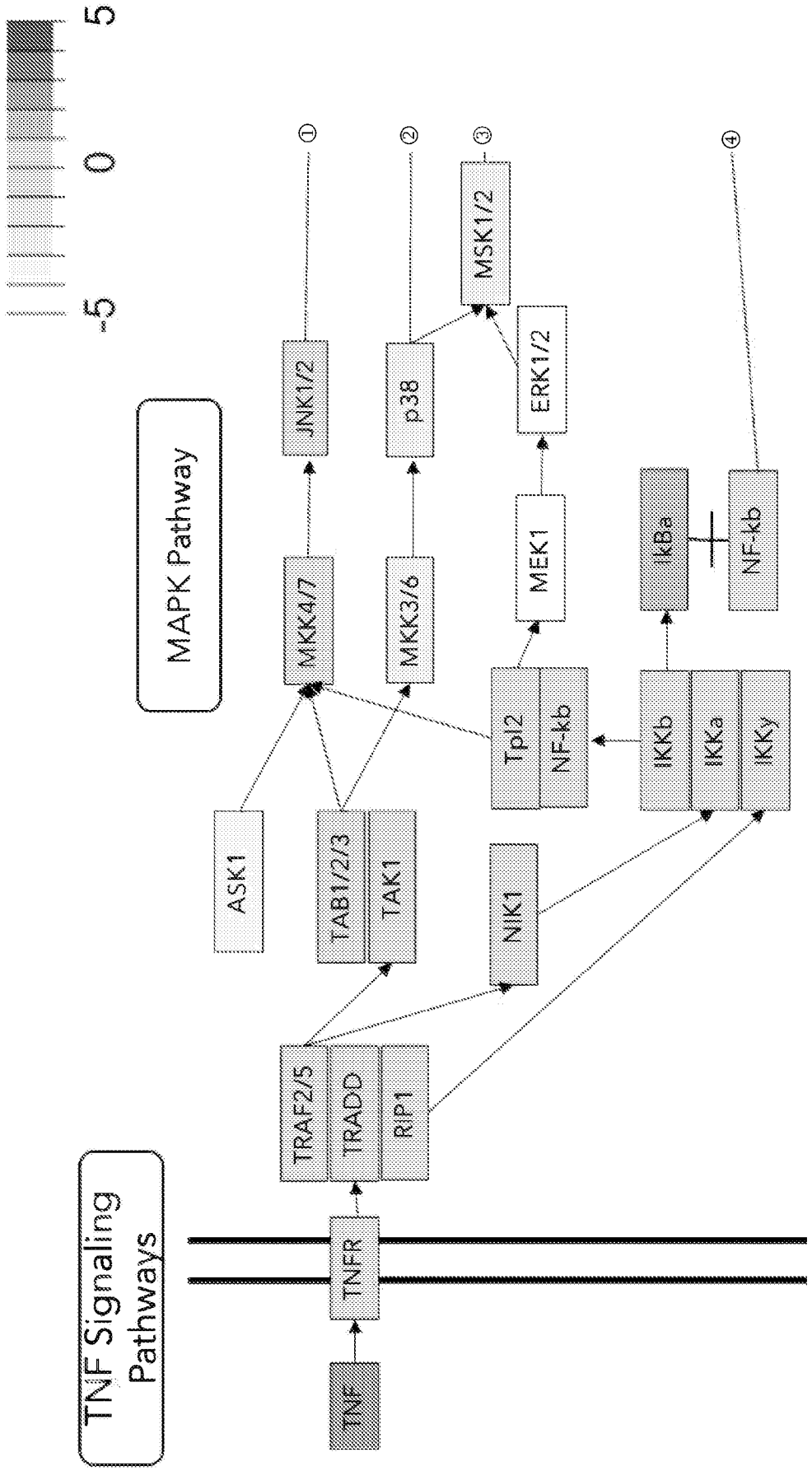

5
0
-5
TNF Signaling Pathways
TNF
TNFR
TRAF2/5
TRADD
RIP1
ASK1
TAB1/2/3
TAK1
NIK1
MAPK Pathway
MKK4/7
MKK3/6
Tpl2
NF-kb
MEK1
IKKb
IKKa
IKKy
IkBa
NF-kb
JNK1/2
p38
ERK1/2
MSK1/2
①
②
③
④

FIG. 28E

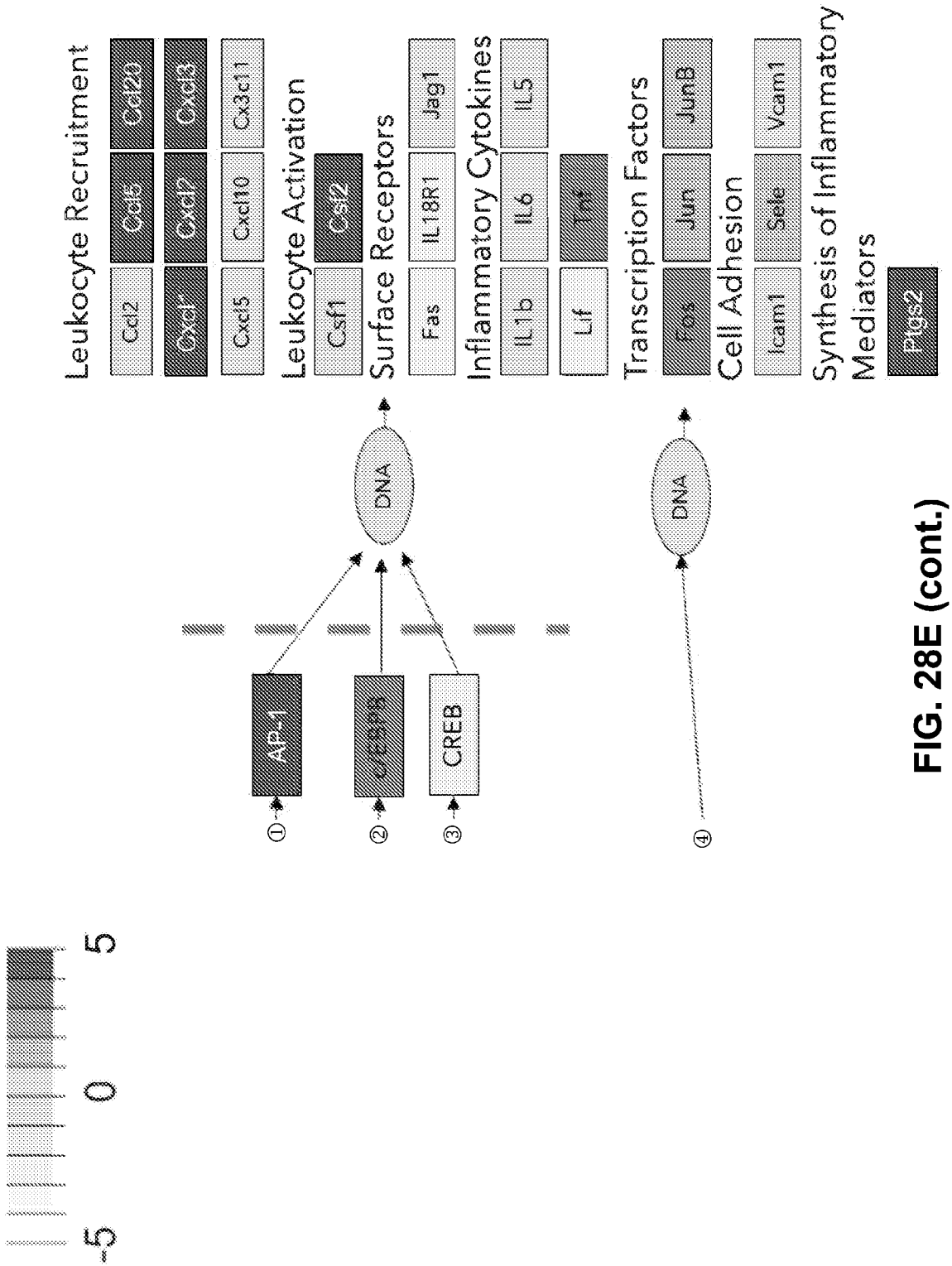

Leukocyte Recruitment
Ccl2
Ccl5
Ccl20
Cxcl1
Cxcl2
Cxcl3
Cxcl5
Cxcl10
Cx3cl1
Leukocyte Activation
Csf1
Csf2
Surface Receptors
Fas
IL18R1
Jag1
Inflammatory Cytokines
IL1b
IL6
IL5
Lif
Tnf
Transcription Factors
Fos
Jun
JunB
Cell Adhesion
Icam1
Sele
Vcam1
Synthesis of Inflammatory Mediators
P16s2
DNA
DNA
AP-1
C/EBPβ
CREB
①
②
③
④
5
0
-5

FIG. 28E (cont.)

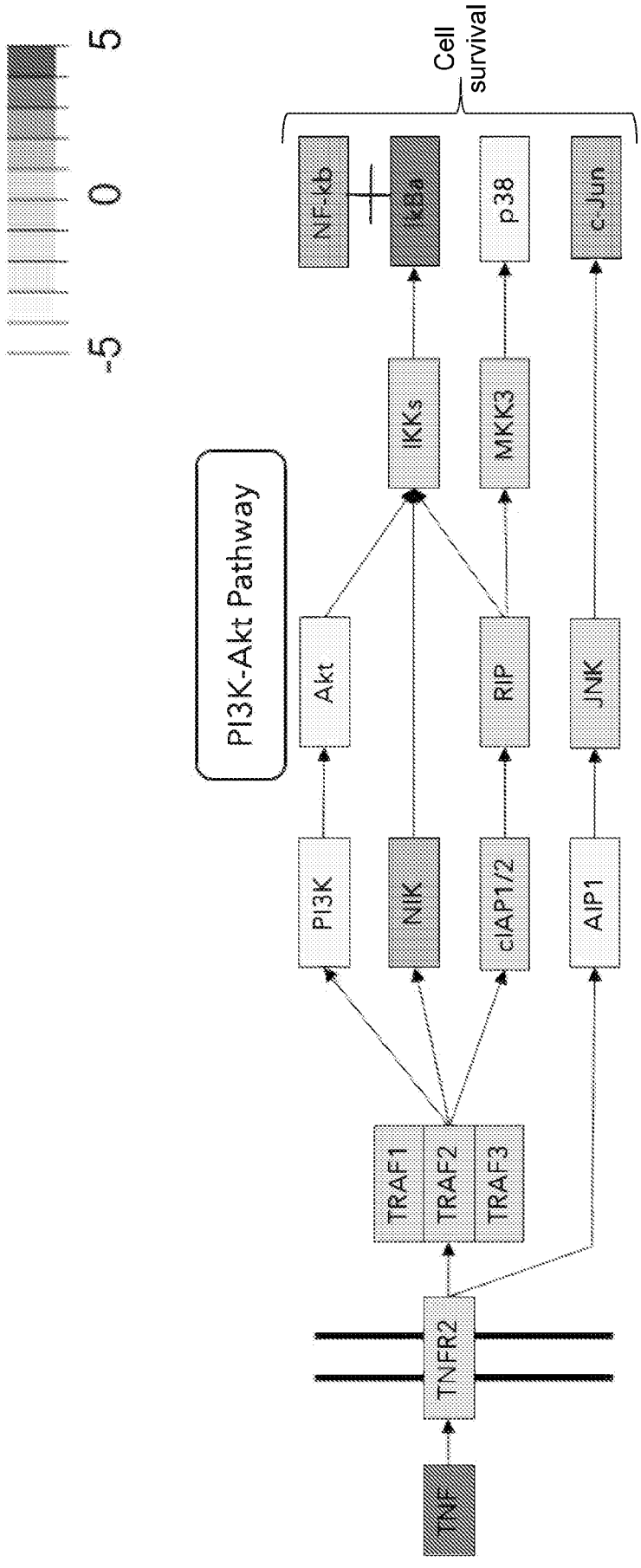

5
0
-5
Cell survival
NF-kb
IκBα
p38
c-Jun
IKKs
MKK3
PI3K-Akt Pathway
Akt
RIP
JNK
PI3K
NIK
cIAP1/2
AIP1
TRAF1
TRAF2
TRAF3
TNFR2
TNF

FIG. 28E (cont.)

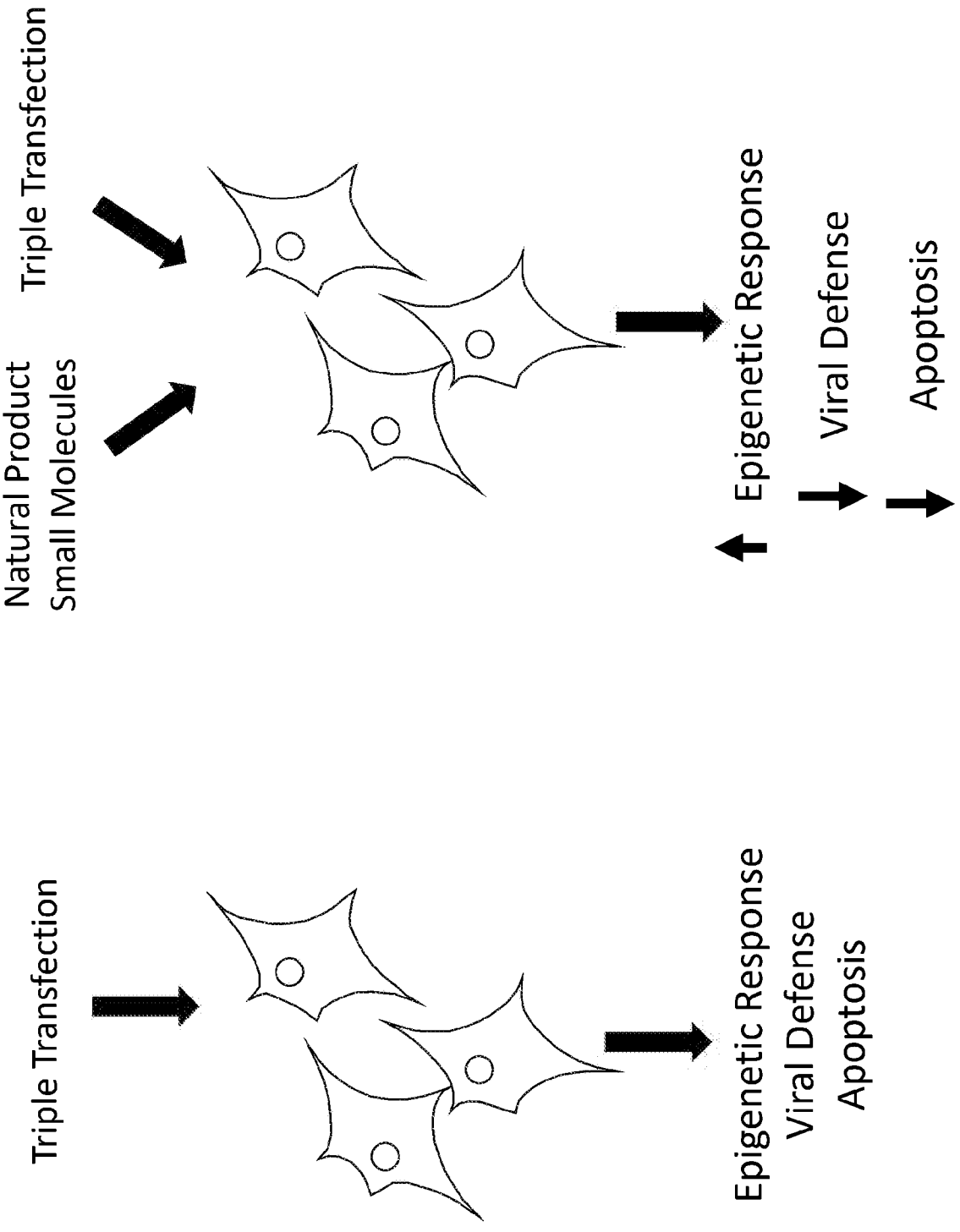

Triple Transfection
Natural Product
Small Molecules
Epigenetic Response
Viral Defense
Apoptosis
Triple Transfection
Epigenetic Response
Viral Defense
Apoptosis

FIG. 29

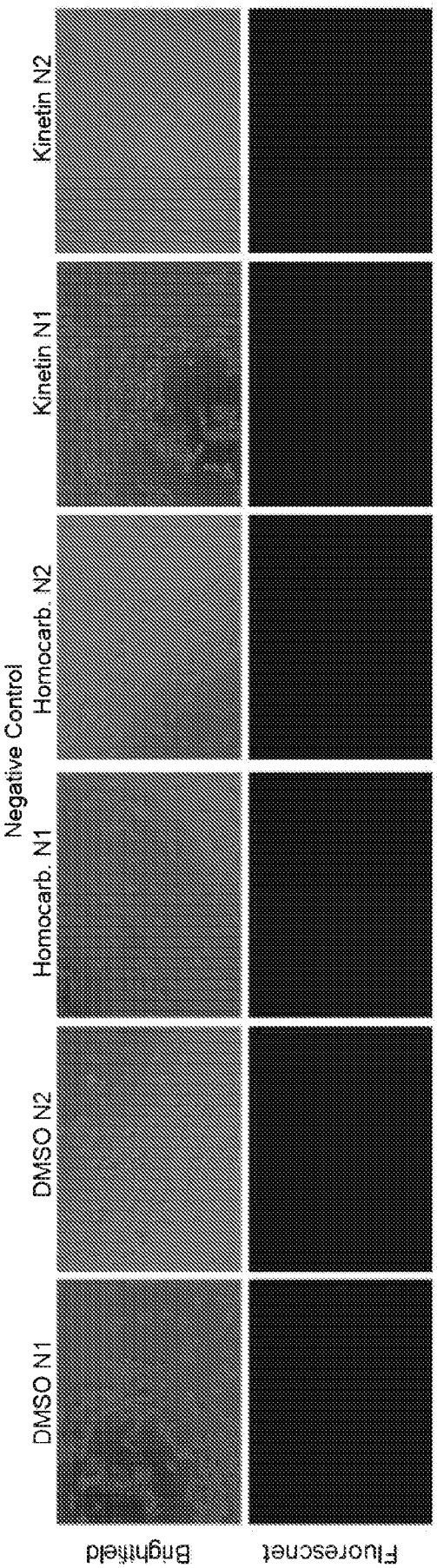


FIG. 30A

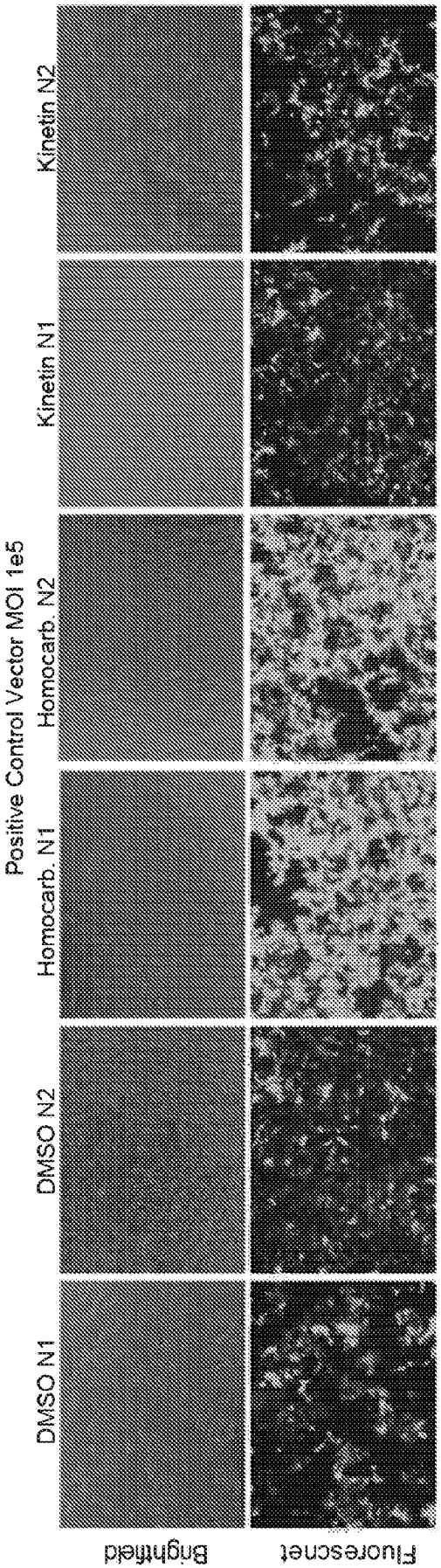


FIG. 30B

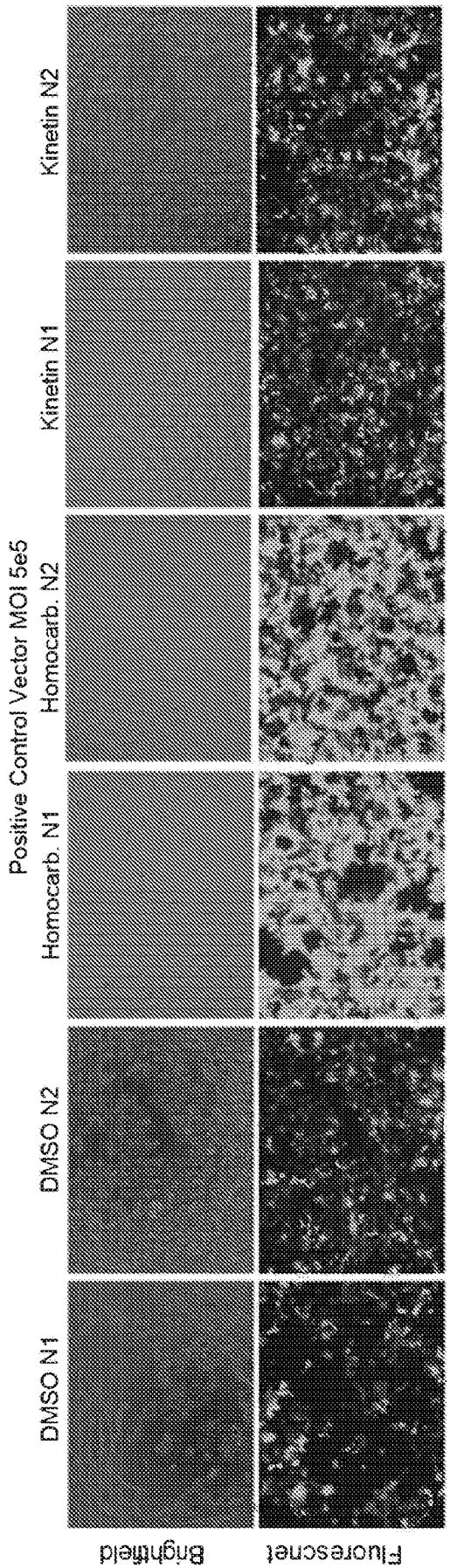

Positive Control Vector MOI 5e5
DMSO N1
DMSO N2
Homocarb. N1
Homocarb. N2
Kinetin N1
Kinetin N2
Brightfield
Fluorescnet

FIG. 30C

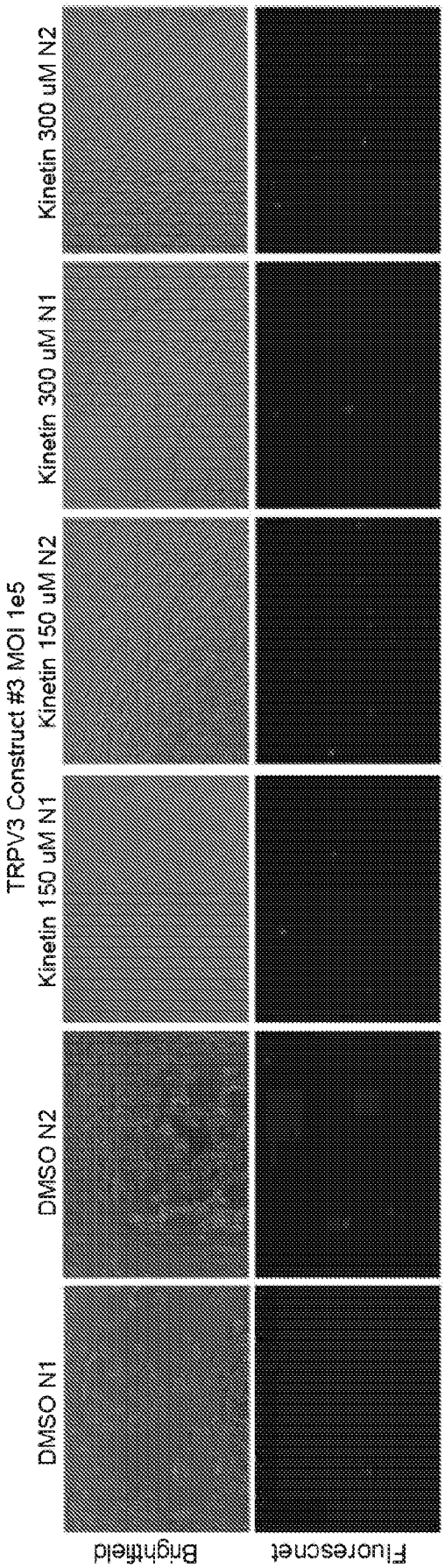


FIG. 31A

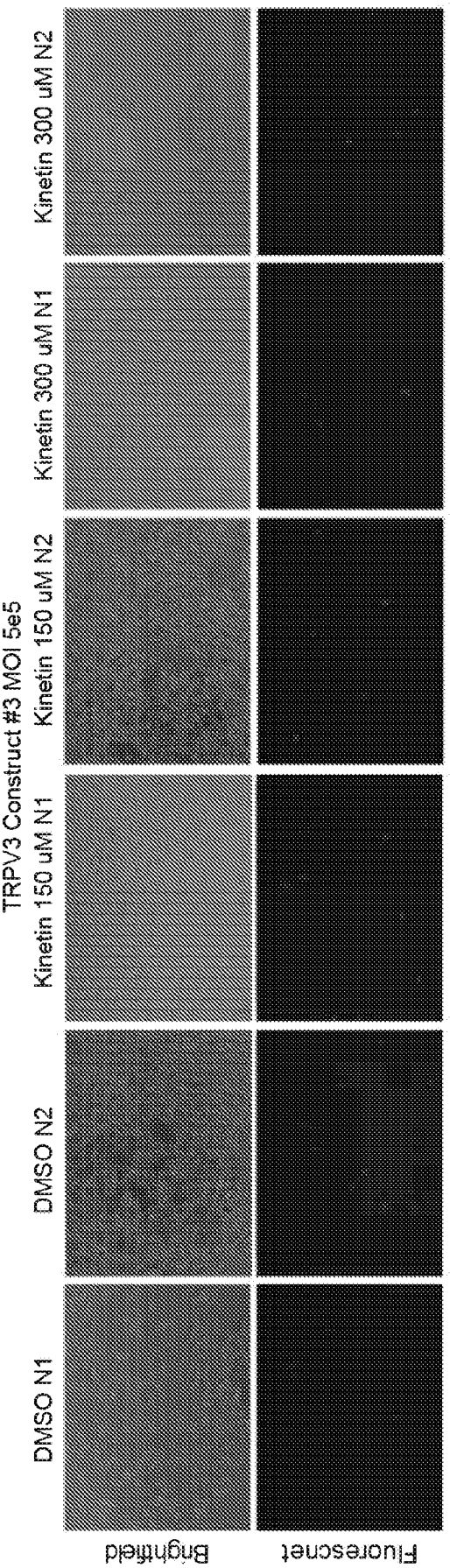


FIG. 31B

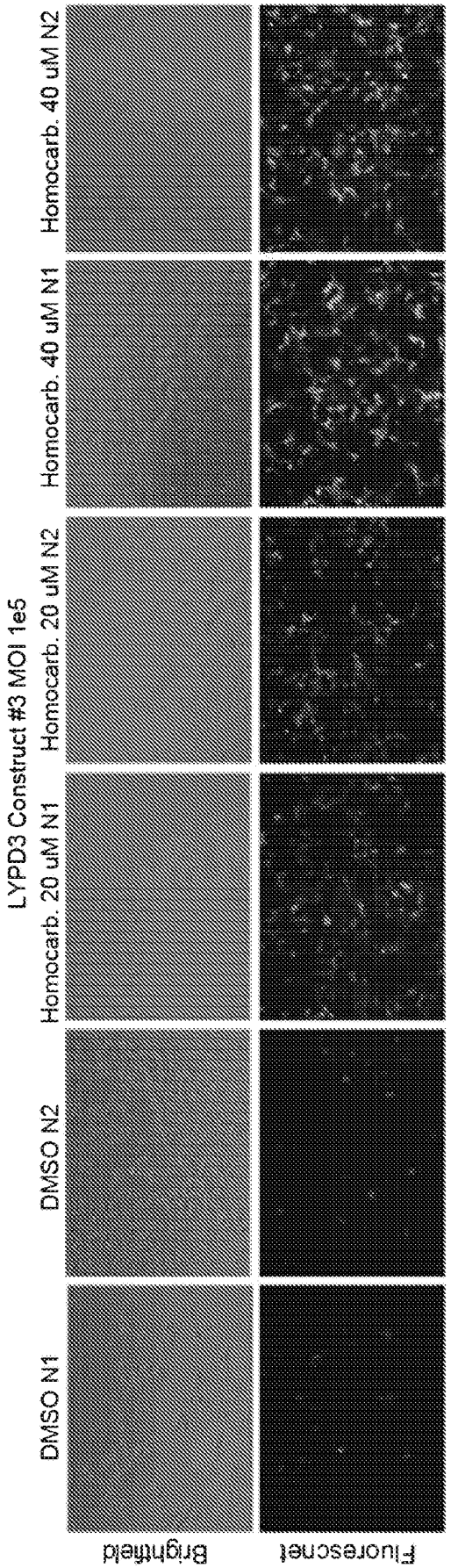


FIG. 31C

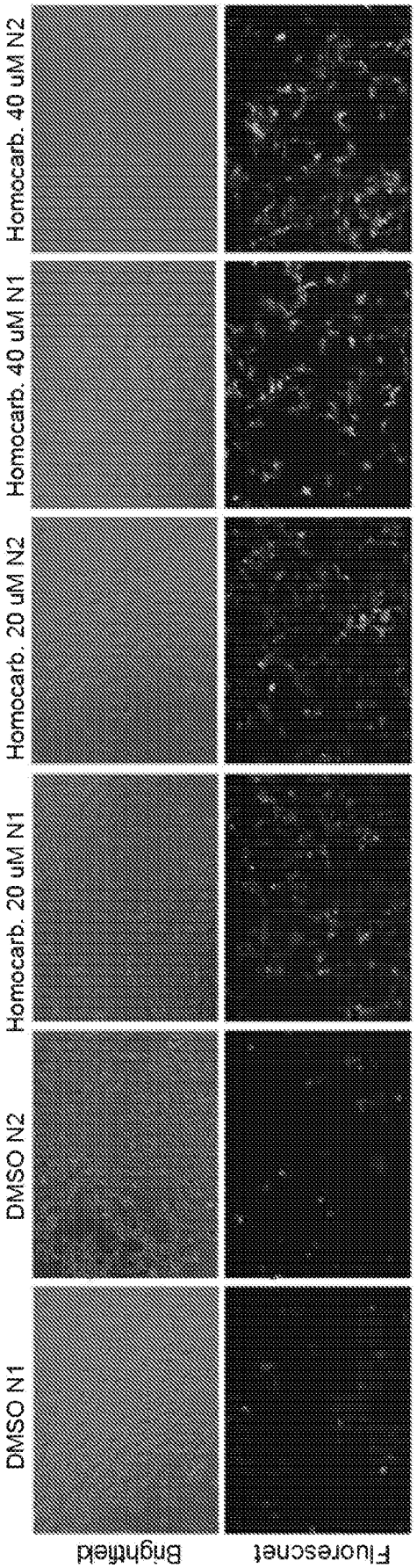


FIG. 31D

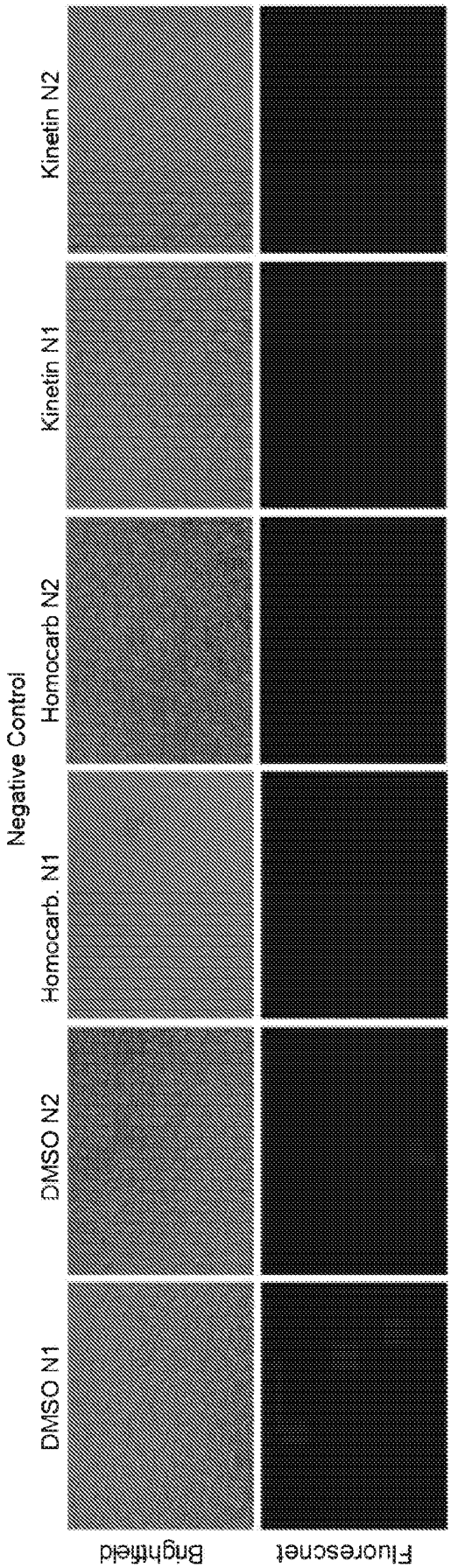


FIG. 32A

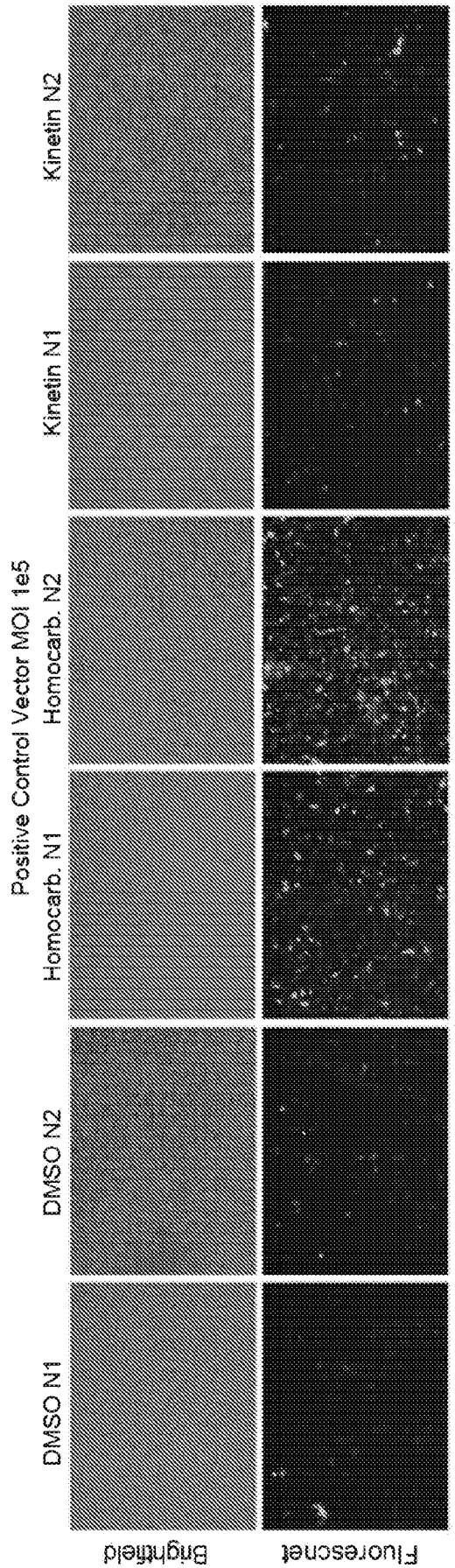


FIG. 32B

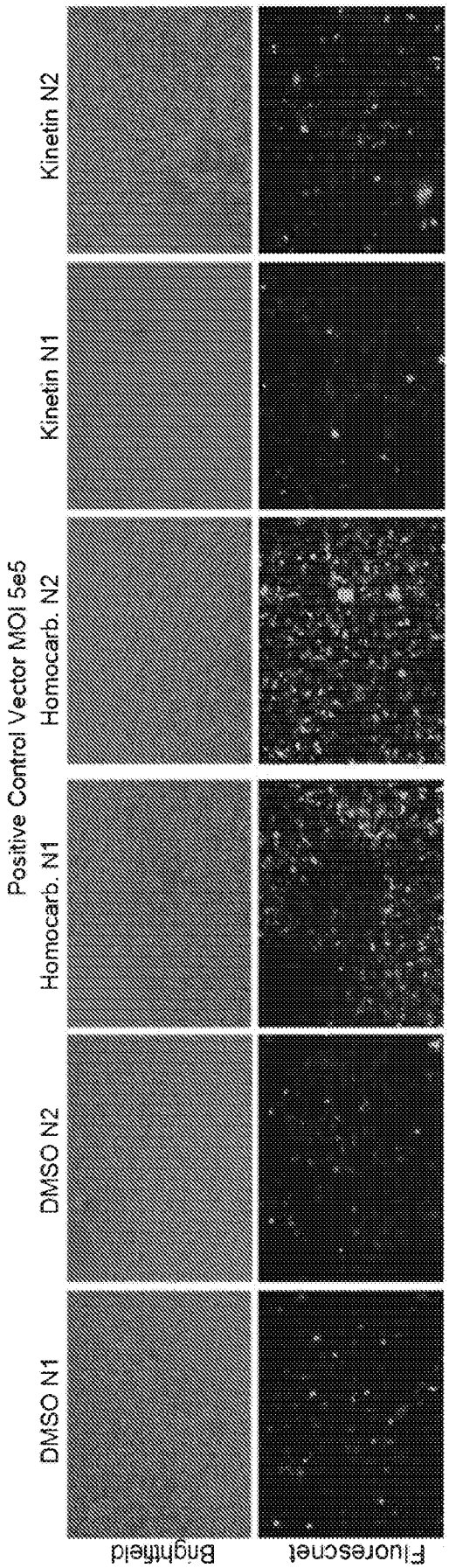

Positive Control Vector MOI 5e5
DMSO N1
DMSO N2
Homocarb. N1
Homocarb. N2
Kinetin N1
Kinetin N2
Brightfield
Fluorescnet

FIG. 32C

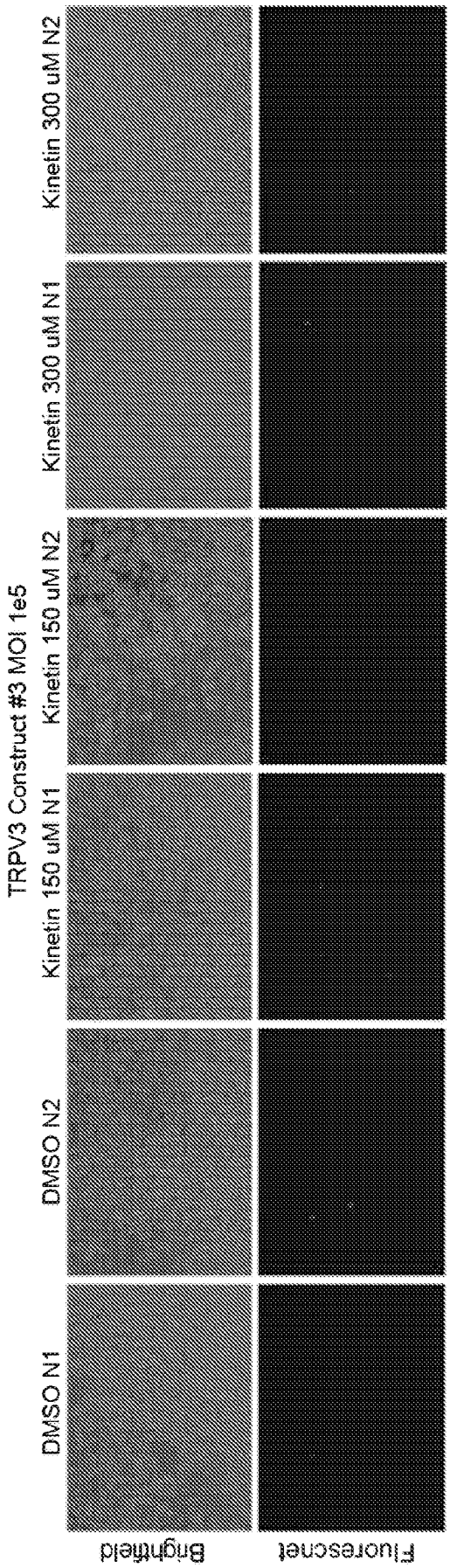


FIG. 33A

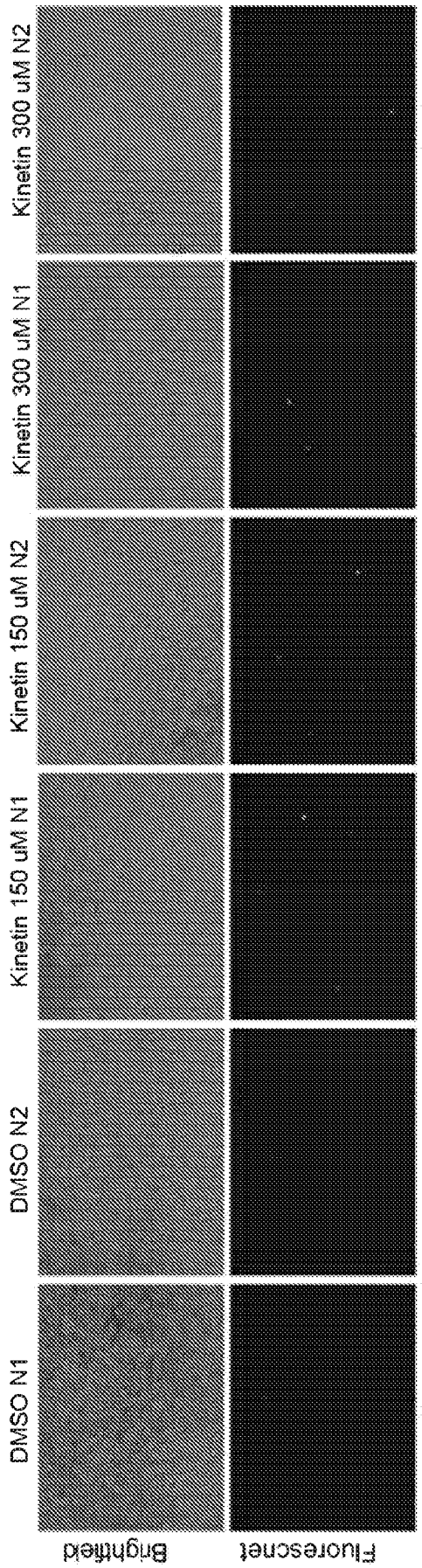


FIG. 33B

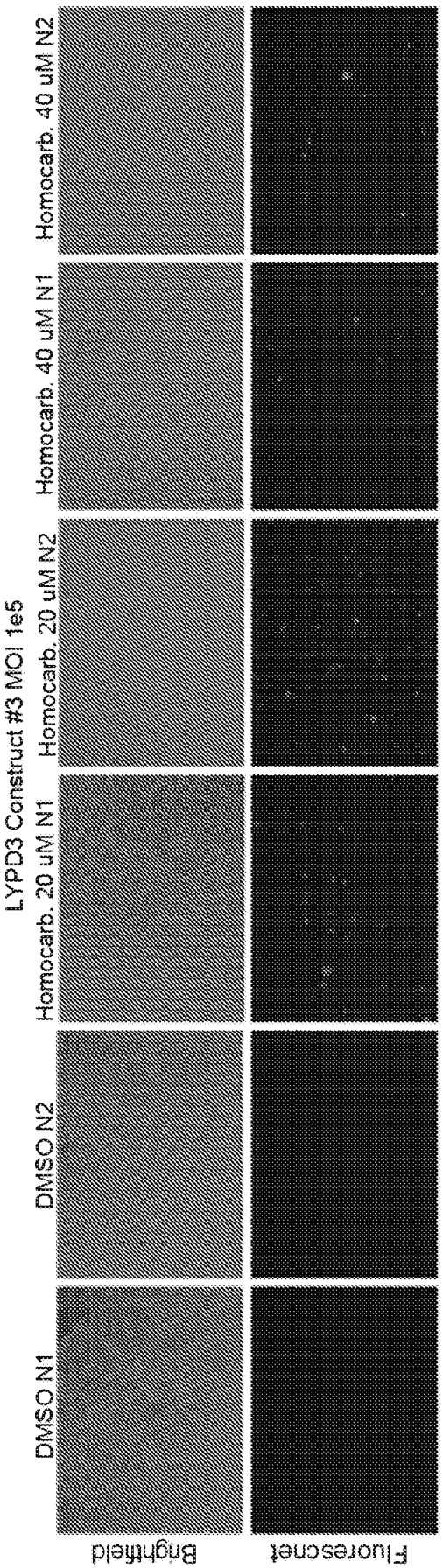


FIG. 33C

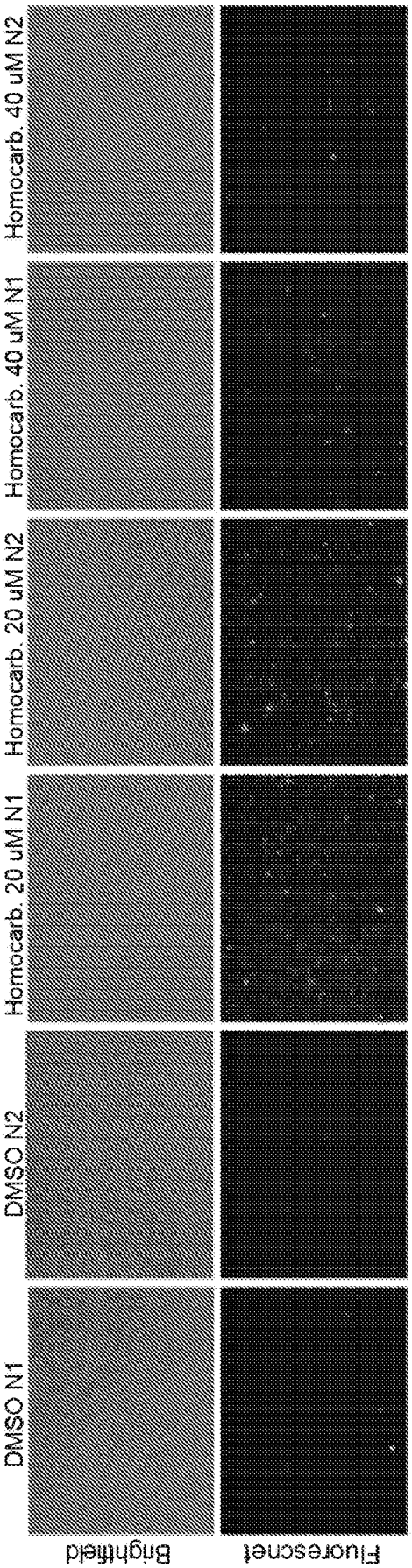


FIG. 33D

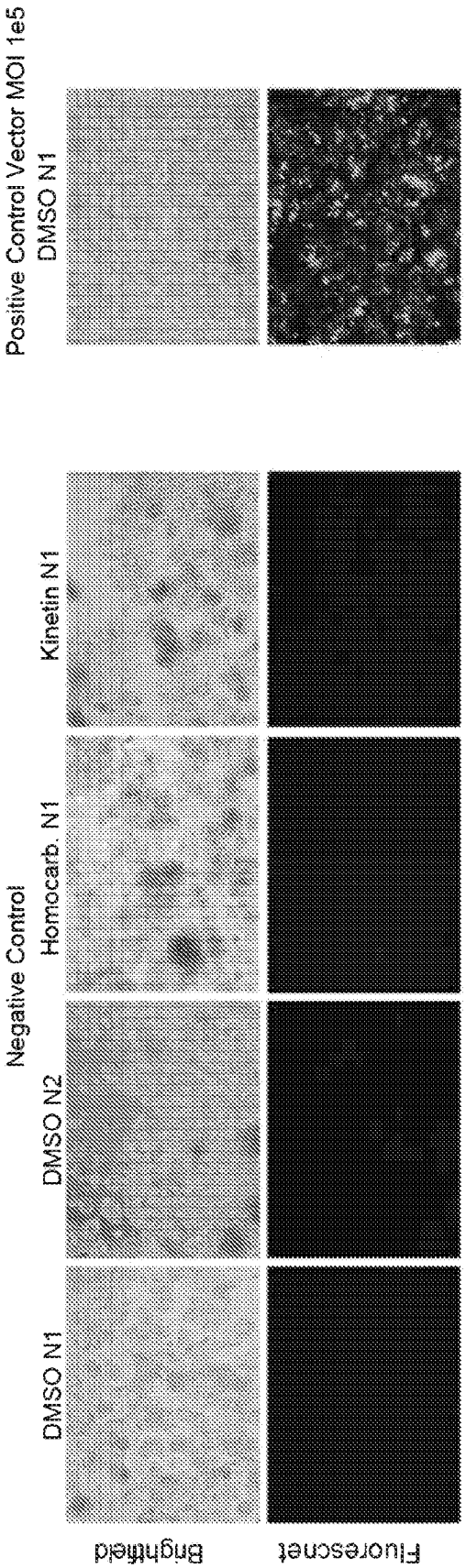


FIG. 34A

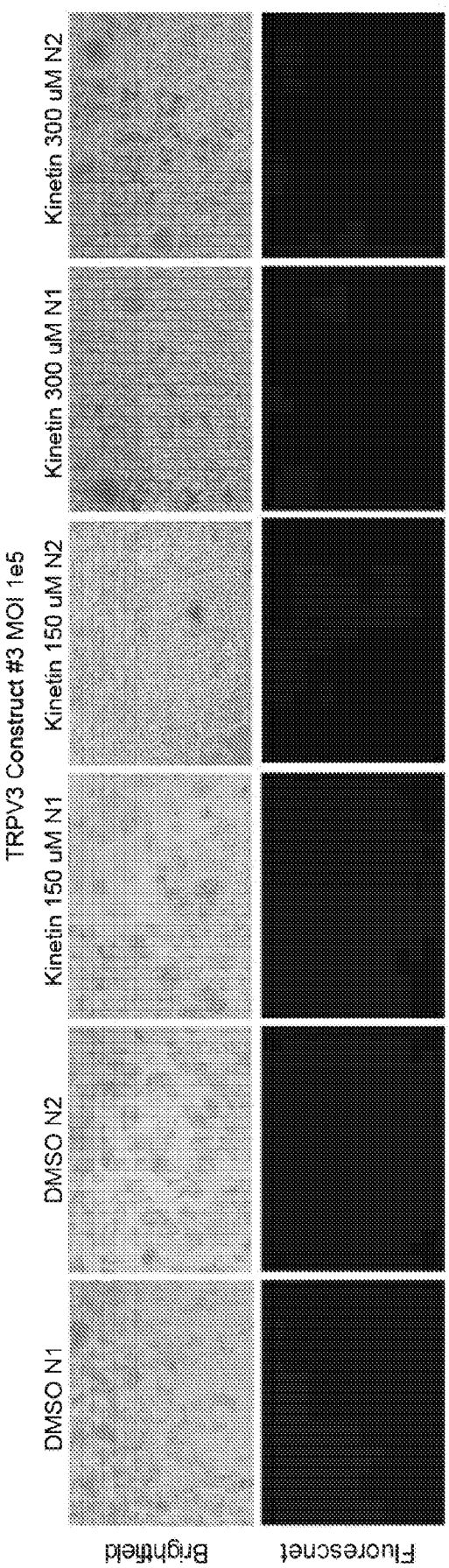


FIG. 34B

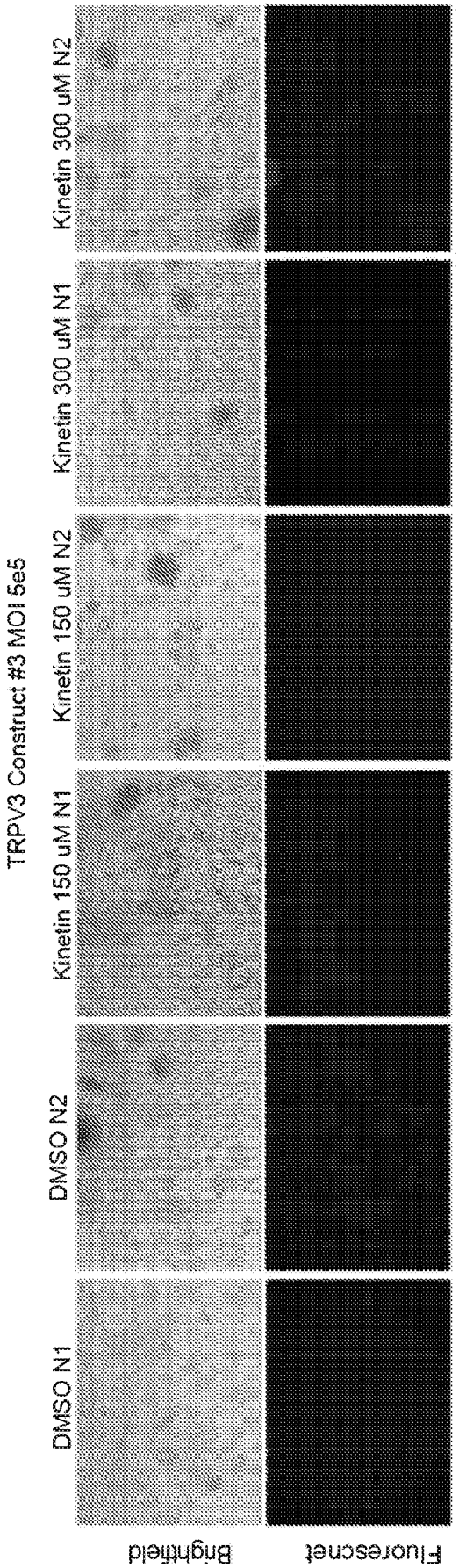

TRPV3 Construct #3 MOI 5e5
DMSO N1
DMSO N2
Kinetin 150 uM N1
Kinetin 150 uM N2
Kinetin 300 uM N1
Kinetin 300 uM N2
Brightfield
Fluorescnet

FIG. 34C

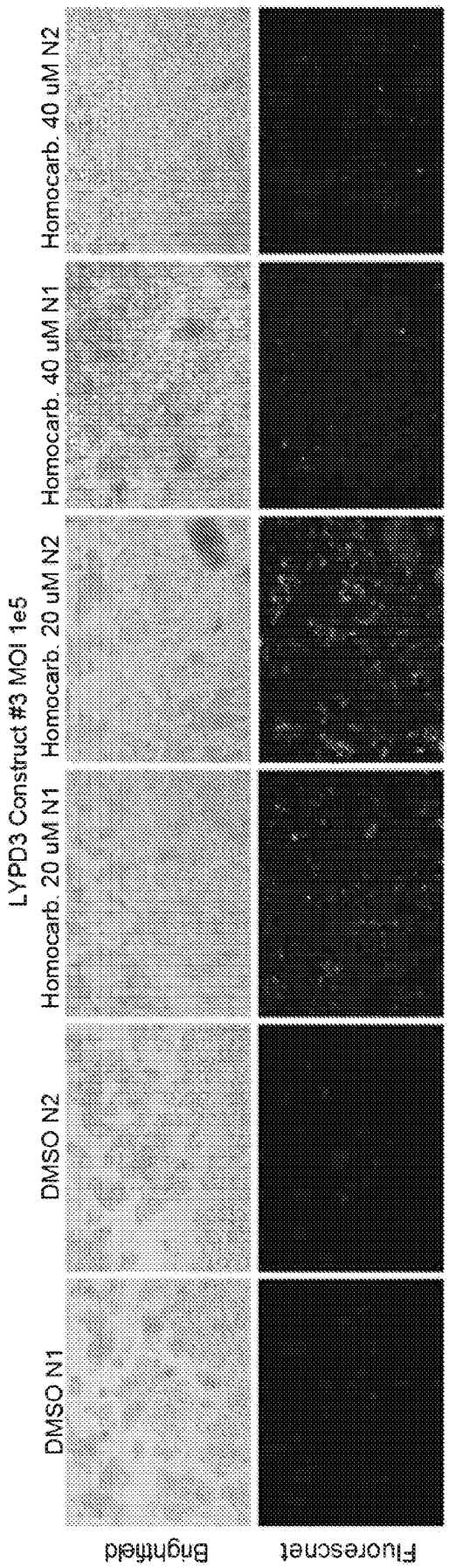


FIG. 35A

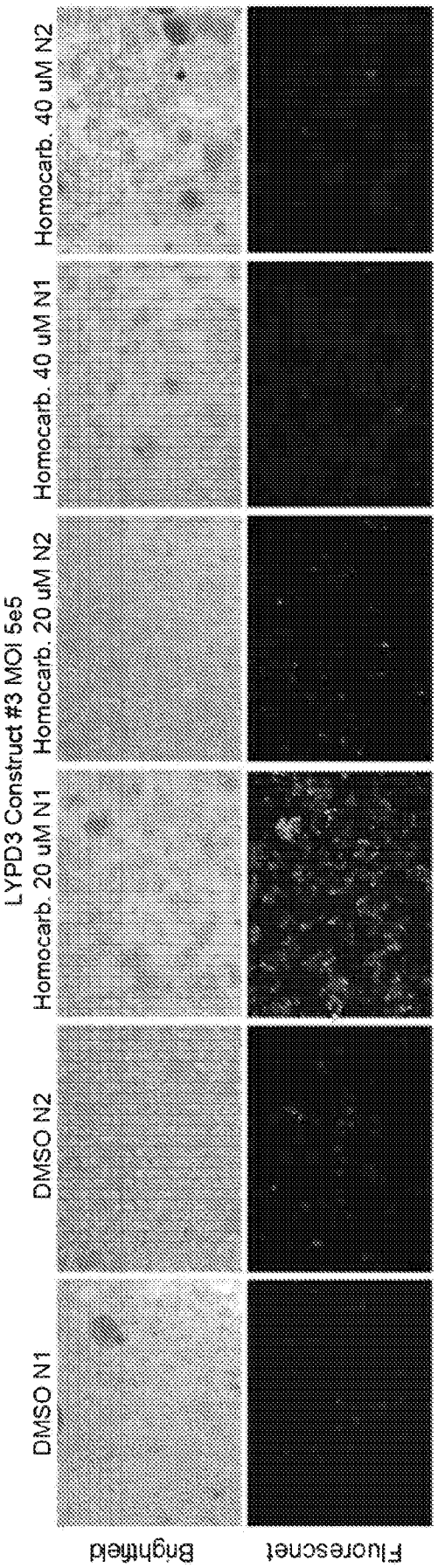


FIG. 35B

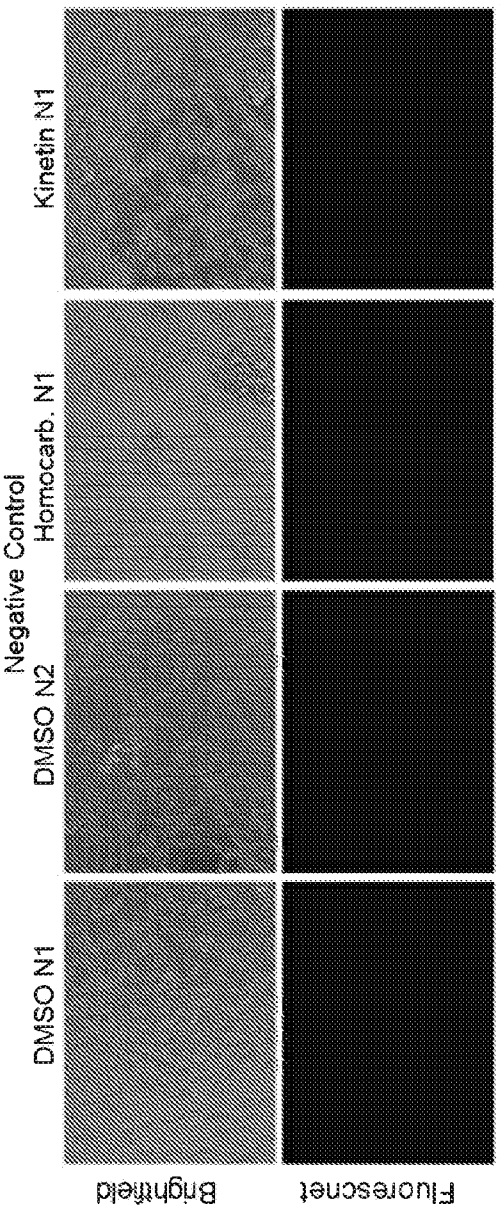


FIG. 36A

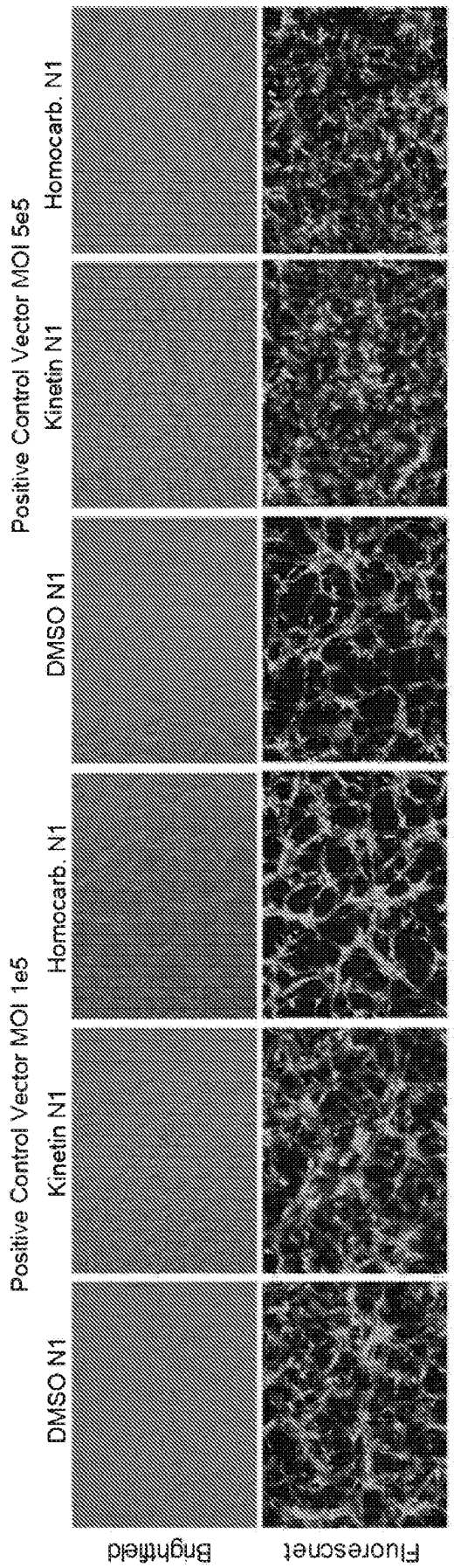


FIG. 36B

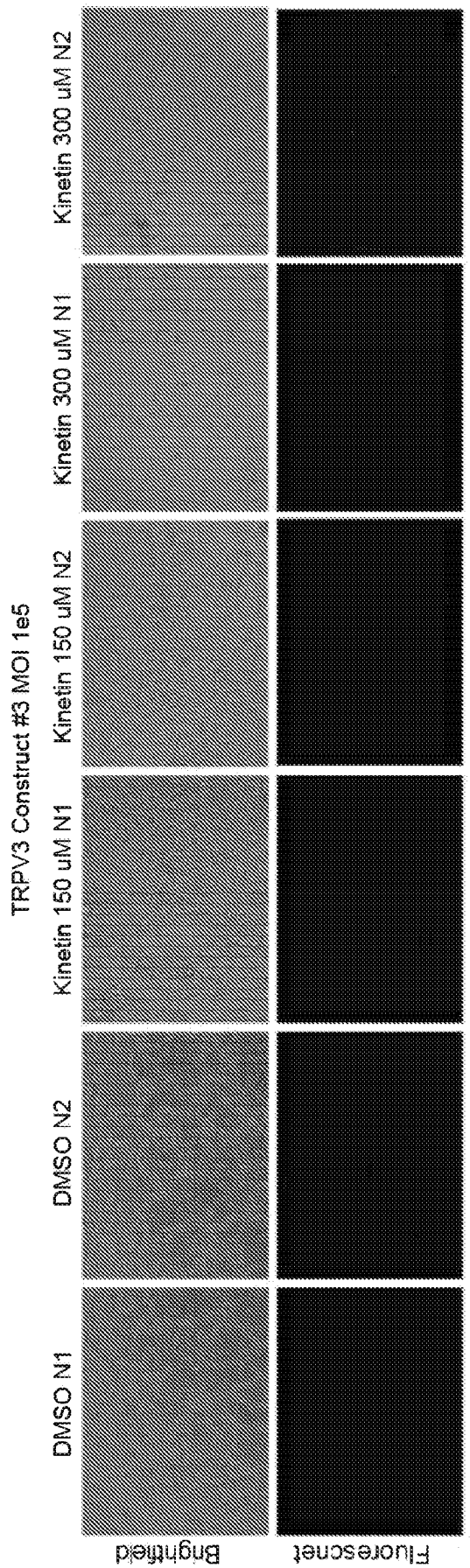


FIG. 37A

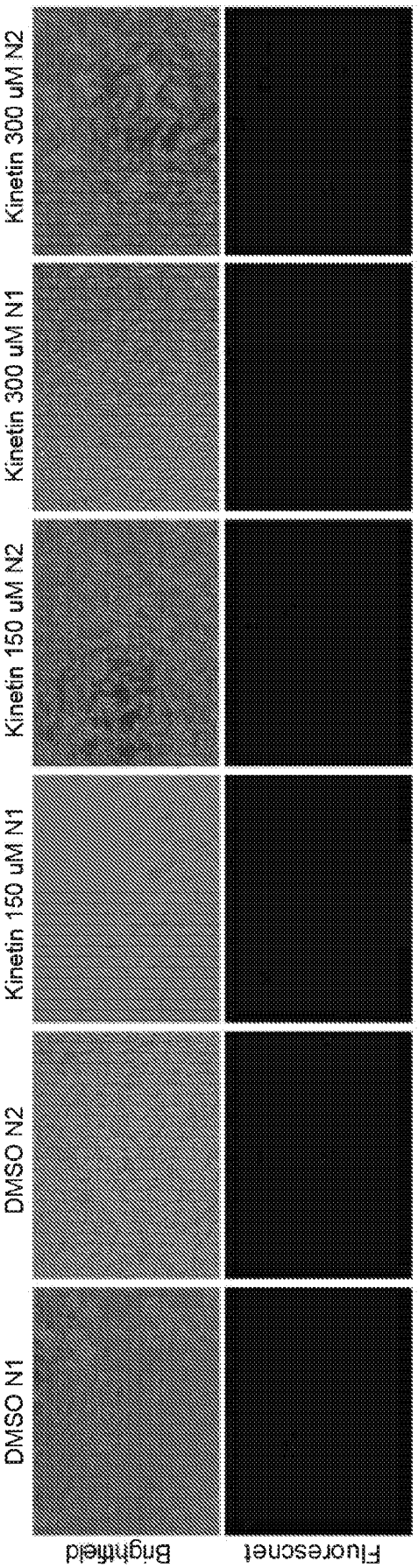


FIG. 37B

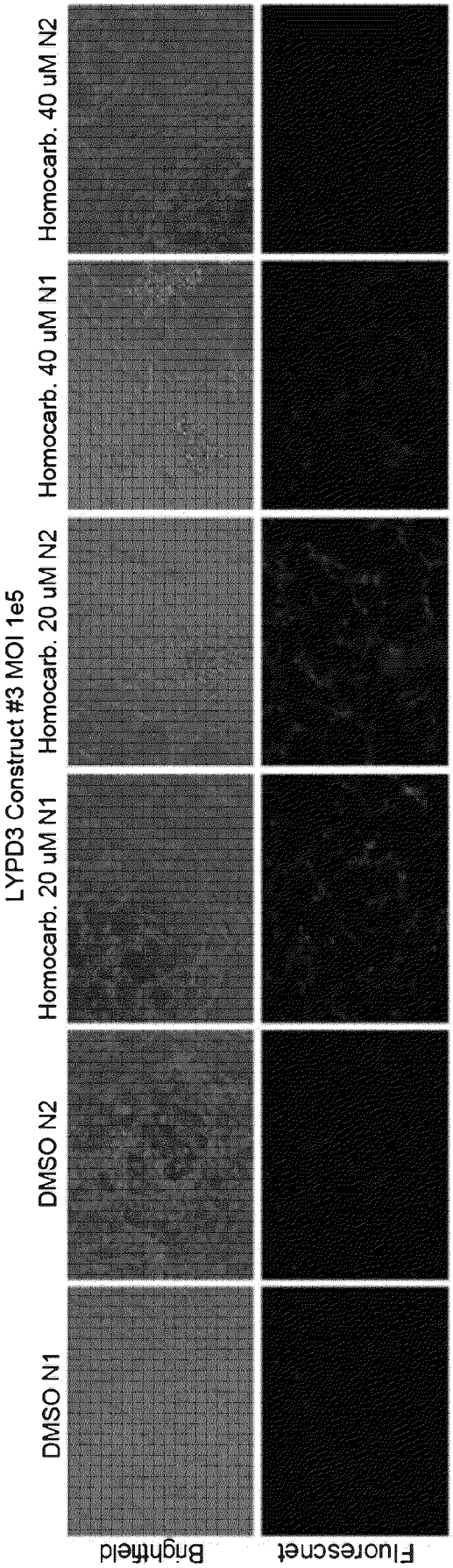


FIG. 37C

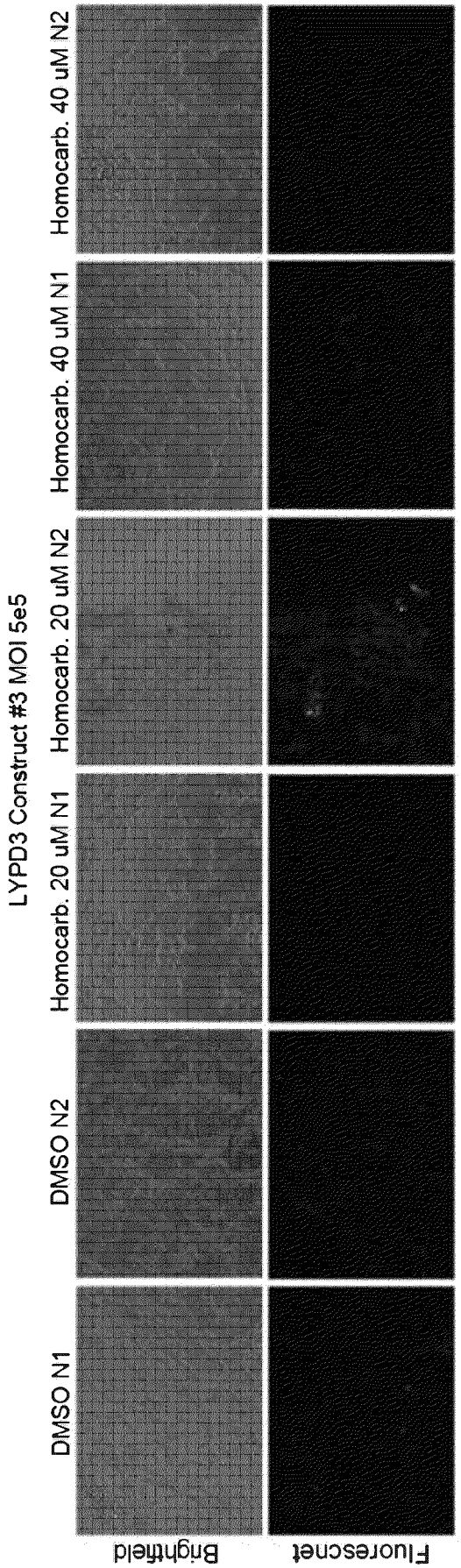


FIG. 37D

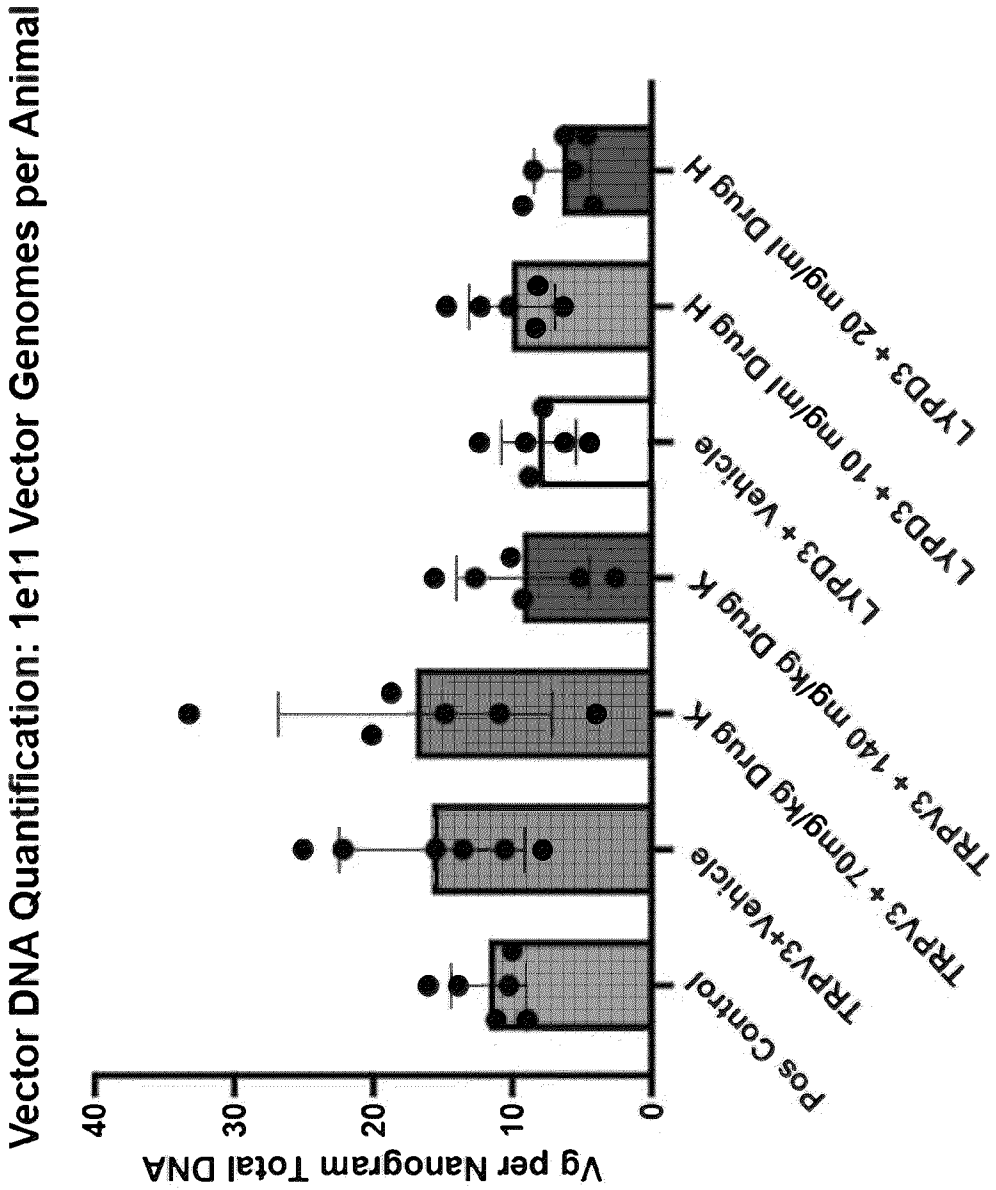

Vector DNA Quantification: 1e11 Vector Genomes per Animal
Vg per Nanogram Total DNA
0
10
20
30
40
Pos Control
TRPV3+Vehicle
TRPV3 + 70mg/kg Drug K
TRPV3 + 140 mg/kg Drug K
LYPD3 + Vehicle
LYPD3 + 10 mg/ml Drug H
LYPD3 + 20 mg/ml Drug H

FIG. 38A

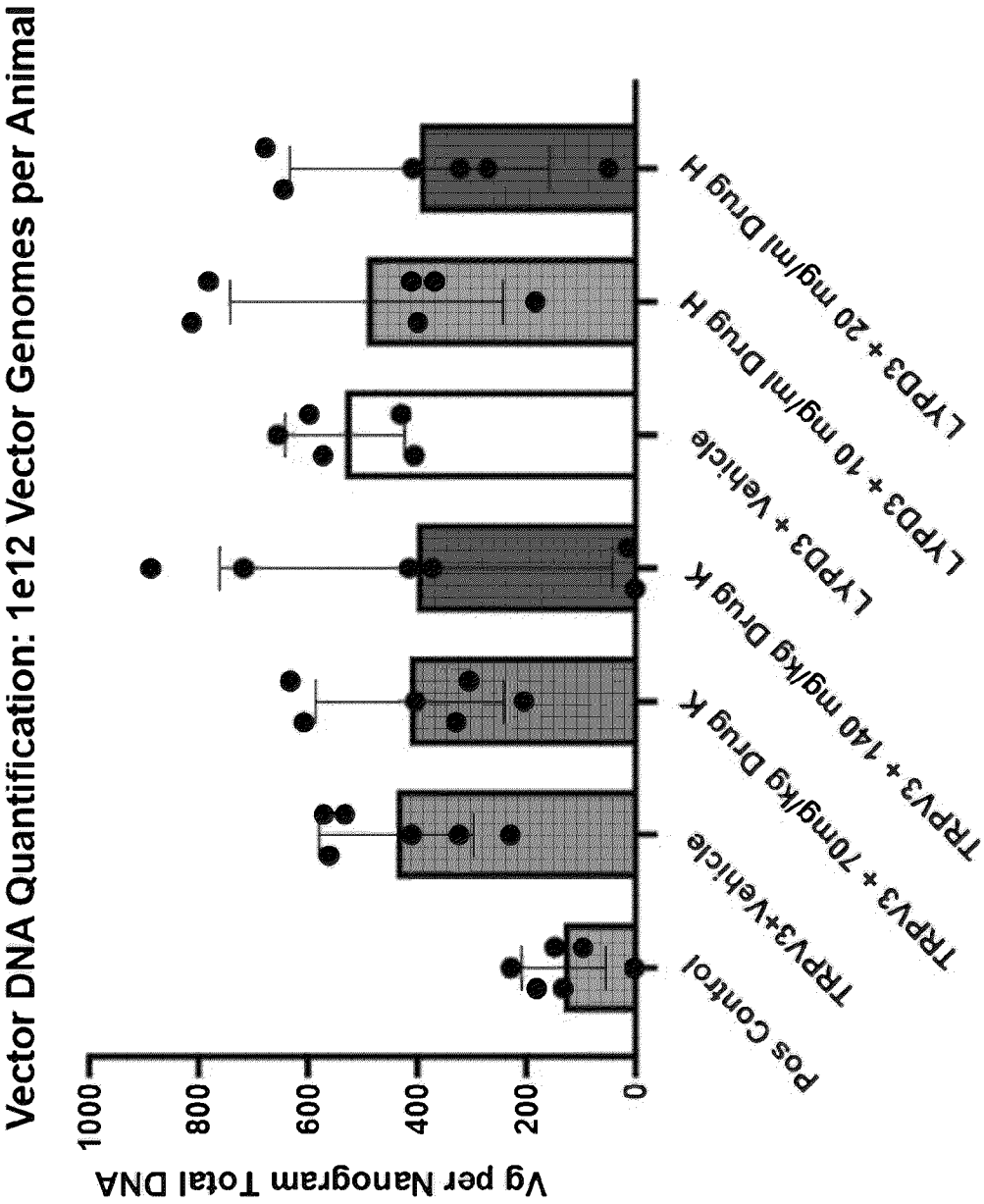

Vector DNA Quantification: 1e12 Vector Genomes per Animal
Vg per Nanogram Total DNA
0
200
400
600
800
1000
Pos Control
TRPV3+Vehicle
TRPV3 + 70mg/kg Drug K
TRPV3 + 140 mg/kg Drug K
LYPD3 + Vehicle
LYPD3 + 10 mg/ml Drug H
LYPD3 + 20 mg/ml Drug H

FIG. 38B

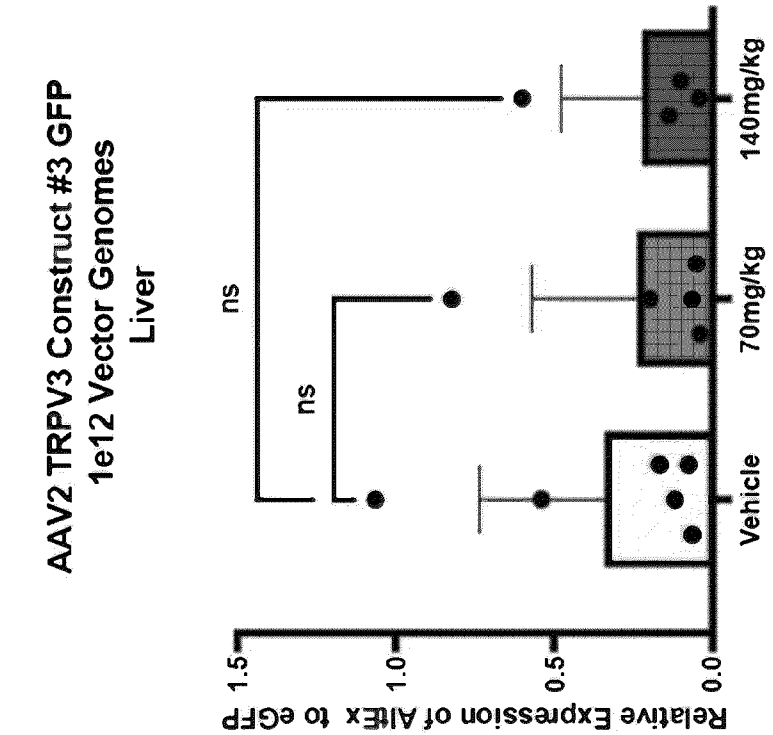


FIG. 39B

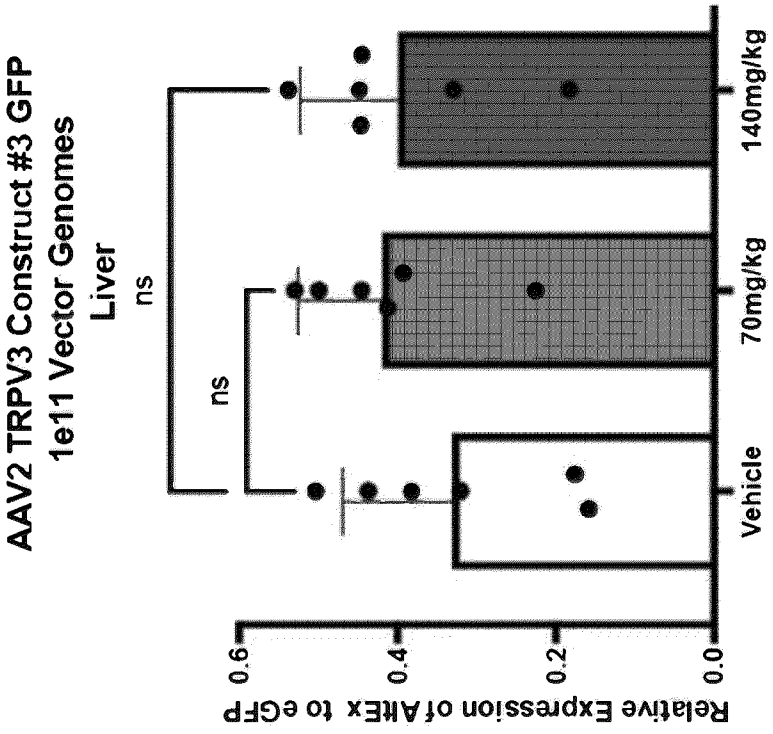


FIG. 39A

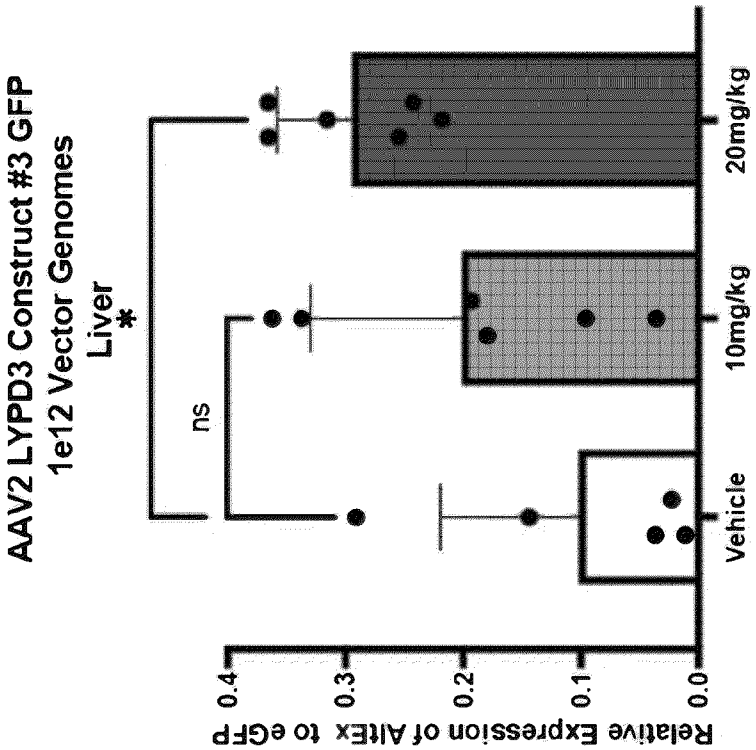


FIG. 39D

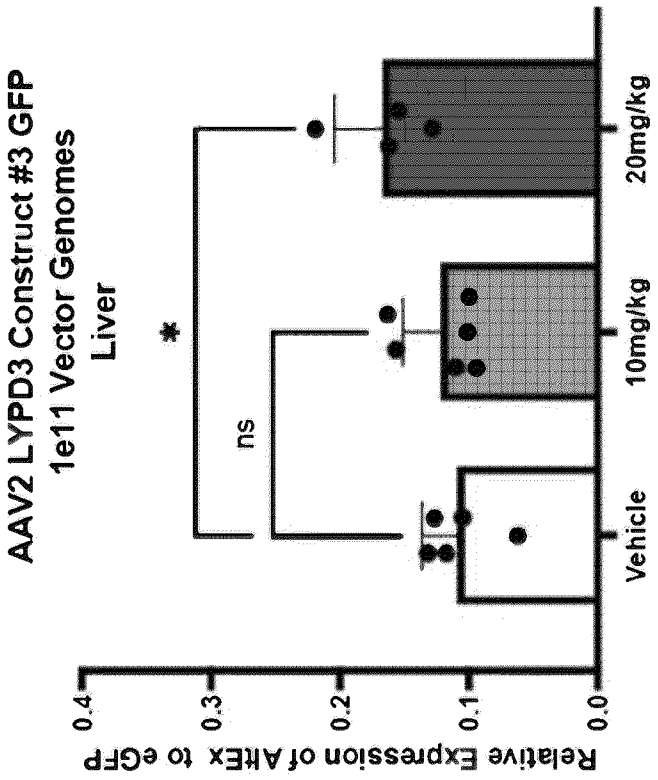


FIG. 39C

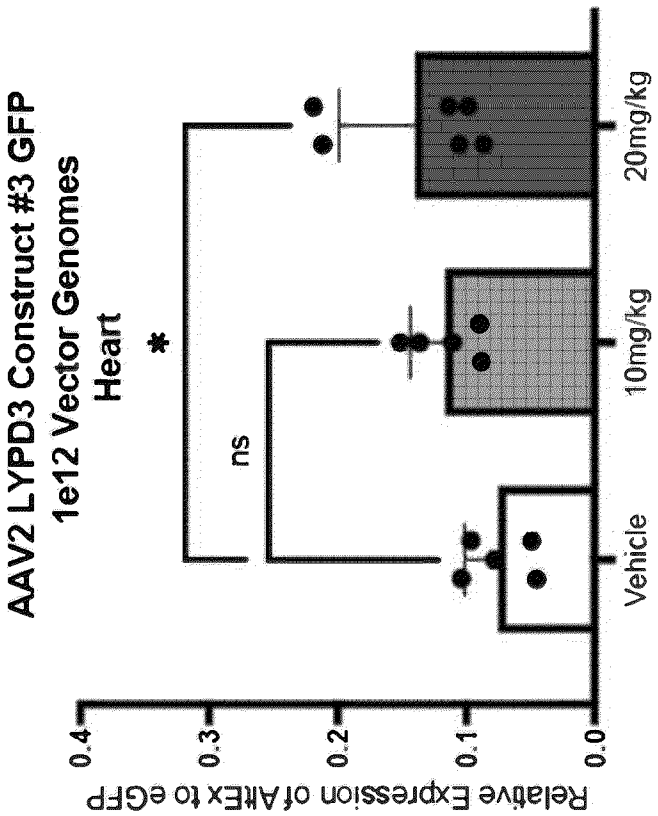


FIG. 39F

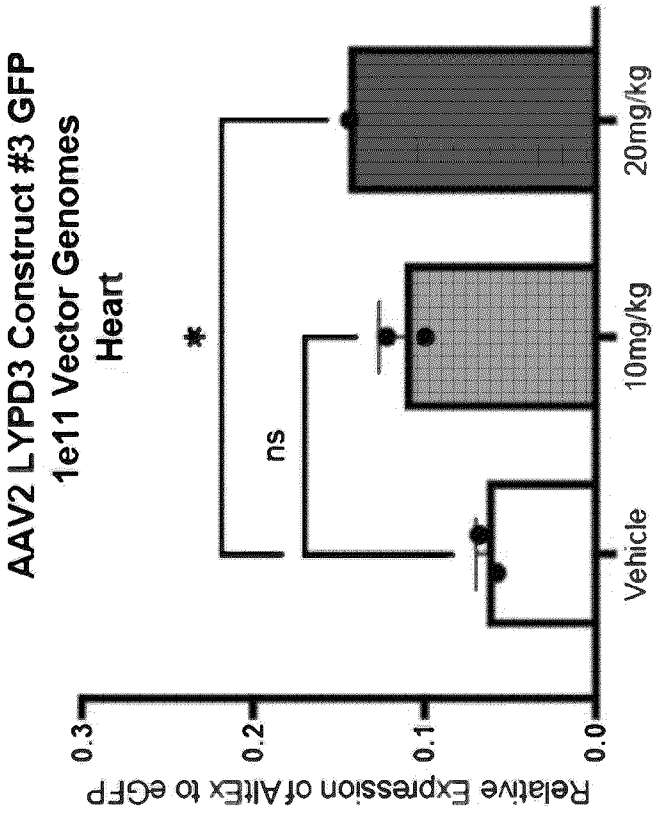


FIG. 39E

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2024/026643

A. CLASSIFICATION OF SUBJECT MATTER

INV. A61K31/41 C12N15/861 A61K31/52 A61K48/00 C12N15/113
C12N15/63
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

A61K C07K C12N

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

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

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BANNING ANTJE ET AL: "Towards Splicing Therapy for Lysosomal Storage Disorders: Methylxanthines and Luteolin Ameliorate Splicing Defects in Aspartylglucosaminuria and Classic Late Infantile Neuronal Ceroid Lipofuscinosis", CELLS, vol. 10, no. 11, 1 January 2021 (2021-01-01), page 2813, XP055953121, DOI: 10.3390/cells10112813 page 3, paragraph 3 - page 4, paragraph 2; figures 4, 11	1-10,14, 18,20, 30, 37-42, 44,49,51

-/--

[x] Further documents are listed in the continuation of Box C.

[x] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 July 2024	07/08/2024

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Salminen, Aaro

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2024/026643

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. [X] forming part of the international application as filed.

   b. [ ] furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).

      [ ] accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.

2. [ ] With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2024/026643

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2021/163556 A1 (CHILDRENS HOSPITAL PHILADELPHIA [US]) 19 August 2021 (2021-08-19) claims 1-90; figures 1, 3; examples 1, 2	1-5, 8-14, 16-21, 27-41, 43-48, 50-74
X	WO 2020/033473 A1 (CHILDRENS HOSPITAL PHILADELPHIA [US]) 13 February 2020 (2020-02-13) claims 1-93; figure 2; examples 1-2	1-5, 9-16, 18-21, 27, 30-41, 44-48,51
A	GARCIA-LOPEZ AMPARO ET AL: "Targeting RNA structure in SMN2 reverses spinal muscular atrophy molecular phenotypes", NATURE COMMUNICATIONS, vol. 9, no. 1, 23 May 2018 (2018-05-23), XP093190246, UK ISSN: 2041-1723, DOI: 10.1038/s41467-018-04110-1 the whole document	1-74

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2024/026643

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2021163556 A1	19-08-2021	AU 2021218807 A1	15-09-2022
		CA 3167836 A1	19-08-2021
		CN 115605596 A	13-01-2023
		EP 4103717 A1	21-12-2022
		IL 295542 A	01-10-2022
		JP 2023514242 A	05-04-2023
		KR 20220146501 A	01-11-2022
		US 2023340498 A1	26-10-2023
		WO 2021163556 A1	19-08-2021
WO 2020033473 A1	13-02-2020	AU 2019317573 A1	25-02-2021
		CA 3108293 A1	13-02-2020
		CN 112805009 A	14-05-2021
		EP 3833357 A1	16-06-2021
		IL 280636 A	25-03-2021
		JP 2021532812 A	02-12-2021
		KR 20210042123 A	16-04-2021
		US 2021355509 A1	18-11-2021
		WO 2020033473 A1	13-02-2020