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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 March 2023 (30.03.2023)

WIPO | PCT

(10) International Publication Number
WO 2023/049707 A1

(51) International Patent Classification:
A61K 38/46 (2006.01) *C12N 5/10* (2006.01)
A61K 48/00 (2006.01) *C12N 15/55* (2006.01)
C07K 14/195 (2006.01)

(21) International Application Number:
PCT/US2022/076723

(22) International Filing Date:
20 September 2022 (20.09.2022)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/246,689 21 September 2021 (21.09.2021) US
63/352,530 15 June 2022 (15.06.2022) US

(71) Applicant: **THE JOHNS HOPKINS UNIVERSITY** [US/US]; 3400 North Charles Street, Baltimore, Maryland 21218 (US).

(72) Inventor; and
(71) Applicant: **COLLER, Jeffery** [US/US]; 618 Blandwood Road, Frederick, Maryland 21701 (US).

(74) Agent: **YOON, Sohee Kim** et al.; Fish & Richardson P.C., P.O. Box 1022, Minneapolis, Minnesota 55440-1022 (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, ME, MG, MK, MN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) **Title:** MRNA REGULON THERAPY FOR THE TREATMENT OF HAPLOINSUFFICIENCY DISORDERS

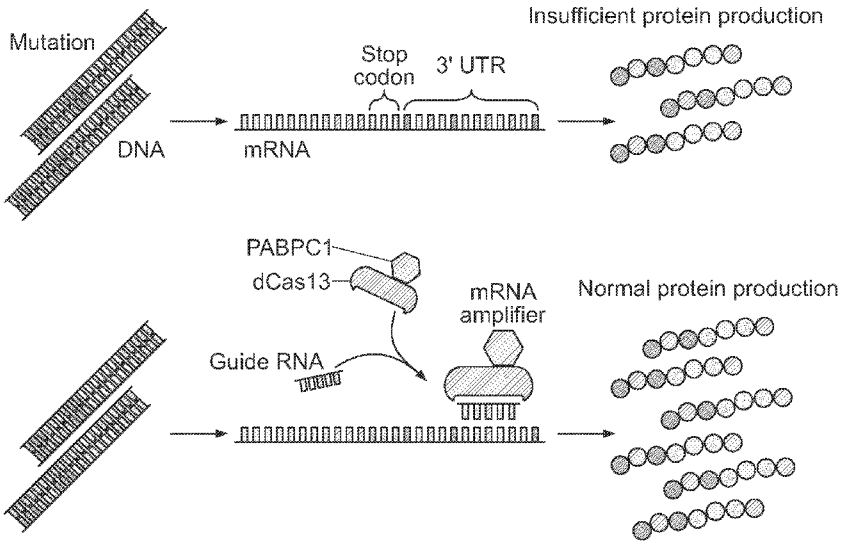


FIG. 15

(57) **Abstract:** Described herein are compositions and methods for treatment of haploinsufficiency disorders by mRNA regulation. For example, provided herein are fusion proteins that include an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA(s).

WO 2023/049707 A1

[Continued on next page]

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))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

MRNA REGULON THERAPY FOR THE TREATMENT OF HAPLOINSUFFICIENCY DISORDERS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Patent Application No. 63/246,689, filed on September 21, 2021, and U.S. Provisional Patent Application No. 63/352,530, filed on June 15, 2022, which are incorporated herein by reference in their entireties.

SEQUENCE LISTING

This application contains a Sequence Listing that has been submitted electronically as an XML file named "44807-0408WO1.XML." The XML file, created on September 20, 2022, is 468,769 bytes in size. The material in the XML file is hereby incorporated by reference in its entirety.

TECHNICAL FIELD

Described herein are compositions and methods for treatment of haploinsufficiency disorders by mRNA regulation.

BACKGROUND

Haploinsufficiency occurs when one gene allele is inactivated and the amount of gene product expressed from the remaining active allele is insufficient for proper gene function. A number of disorders are associated with, or are caused by haploinsufficiency. An example of a haploinsufficiency disorder is Dravet Syndrome.

Dravet Syndrome is a rare and catastrophic form of intractable epilepsy that begins in infancy. Initially, patients experience prolonged seizures. In their second year, additional types of seizure begin to occur, which typically coincide with a developmental decline, possibly due to repeated cerebral hypoxia. This leads to poor development of language and motor skills. Mutations in *SCNIA* (encoding the voltage-gated sodium channel a subunit), *SCNIB* (encoding the voltage-gated sodium channel ~I subunit), *SCN2A*, *SCN3A*, *SCN9A*, *GABRG2* (encoding the y-aminobutyric acid receptor y2 subunit), *GABRD* (encoding the y-aminobutyric acid receptor *11* subunit) and/or *PCDH19* genes have been linked to Dravet Syndrome.

SCN1A encodes the neuronal voltage-gated sodium channel Nav1.1 that is expressed prominently in inhibitory GABAergic neurons. Loss-of-function (LOF) mutations in *SCN1A*

including missense and premature termination codons (PTC) are the most frequently discovered cause of Dravet Syndrome [1, 2].

LOF premature termination codons (PTCs) arise from single nucleotide mutations that convert a canonical triplet nucleotide codon into one of three stop codons (TAG, TGA, or TAA). PTCs can be more deleterious than missense mutations because they result in the complete loss of protein expression [7]. There are ten amino acid codons vulnerable to PTC conversion by a single nucleotide substitution (arginine, tryptophan, tyrosine, cysteine, glutamic acid, lysine, glutamine, serine, leucine, and glycine). PTC mutations in *SCN1A* are a prevalent class of mutations associated with Dravet Syndrome (Figure 1) with more than 180 known PTC mutations.

SCN1A LOF mutations lead to neuronal network hyperactivity as a consequence of impaired inhibitory GABAergic neurotransmission. While frequent unprovoked seizures are among the presenting feature of Dravet Syndrome, patients also suffer from a range of comorbidities affecting the areas of cognition, locomotion, speech, and behavior [3]. Dravet Syndrome patients often have disrupted sleep and circadian rhythms, neurodevelopmental delay and intellectual disabilities, oculomotor deficits, and psychomotor regression. Sudden unexpected death associated with epilepsy (SUDEP) is also prevalent in this population [4-6]. Due to the severity of these comorbidities, effective treatments that can address both seizures and the range of comorbidities associated with Dravet Syndrome are urgently needed. Dravet Syndrome has available animal models and has been amenable to clinical trials with successful outcomes.

There are limited therapeutic options for repair of mutations in *SCN1A*. Small molecules, such as aminoglycosides [9], dipeptides [10], and oxadiazoles [11] promote "read-through" but this approach results in the encoding of a near-cognate amino acid [12], effectively generating a missense mutation in place of the PTC. Sodium channels, such as *SCN1A*, are poorly tolerant of missense mutations, which can cause gain-of-function and loss-of-function neurological disorders [13, 14]. Furthermore, aminoglycosides are ototoxic and nephrotoxic [15], and the first-in-class oxadiazole (Ataluren) displayed unexpectedly low efficacy in patient populations (ACT DMD Phase 3 clinical trial, NCT01826487; ACT CF, NCT02139306). Furthermore, most previous therapeutic initiatives for Dravet Syndrome were aimed only at seizure reduction. True disease modifying therapies for developmental and epileptic encephalopathies (DEE) are lacking.

Recent and ongoing advances in CRISPR/Cas9-mediated genome editing provides a potentially permanent solution for monogenic diseases. However, LOF missense mutations occur throughout the *SCN1A* coding region, making gene-editing approaches highly challenging [8]. Furthermore, aspects of this technology impart hurdles for its rapid use as a therapeutic [16,

17] including cell type specific delivery, the efficiency of homologous recombination, and the frequency of on-target repair errors and off-target editing. Moreover, at >6 kb per transcript, the sodium channel coding region of *SCN1A* exceeds the capacity of AAV thus limiting the utility of the rapid pace of advancing AAV therapies for tissue specific gene replacement. Thus, there are limited clinical options for *SCN1A* LOF associated with Dravet Syndrome, and the therapeutic pipeline remains thin. The compositions and methods described herein address this and other deficiencies in the art.

SUMMARY

The compositions and methods described herein have a broad range of applications, providing potential disease modifying therapies for a range of haploinsufficiency disorders. The key advantage of this mRNA regulon approach is that it is 100% agnostic to mutation type and location by stabilizing the mRNA of the wild-type (WT) allele. The potential to reengineer these approaches for other indications is high and facile.

Thus, provided herein are fusion proteins comprising: an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA(s).

In some embodiments, the RNA effector protein is a Cas effector protein selected from the group consisting of Cas9, Cas12, Cas13, and Cas14. In some embodiments, the RNA effector protein is Cas13b. In some embodiments, the Cas effector protein is a catalytically inactive Cas protein.

In some embodiments, the regulon moiety is PABPC1 or NAT10.

In some embodiments, the fusion proteins further comprise a linker and/or a spacer.

In some embodiments, the fusion proteins further comprise a nuclear export signal and/or an epitope tag.

In some embodiments, the RNA effector protein is N terminal to the regulon moiety. In some embodiments, the RNA effector protein is C terminal to the regulon moiety.

In some embodiments, the fusion proteins comprise or consist of SEQ ID NO: 48 or SEQ ID NO: 49 or a polypeptide having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to SEQ ID NO: 48 or SEQ ID NO: 49.

Also provided herein are polynucleotides encoding any one of the fusion proteins described herein. Also provided herein are vectors comprising any one of the polynucleotides described herein. Also provided herein are cells comprising any one of the vectors described herein.

Also provided herein are systems comprising: a fusion protein comprising: an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA(s); and a gRNA that forms a complex with the RNA effector protein and comprises a complementarity region that hybridizes with the mRNA(s) of the active allele.

In some embodiments, the RNA effector protein is a Cas effector protein selected from the group consisting of Cas9, Cas12, Cas13, and Cas14. In some embodiments, the Cas effector protein is Cas13b. In some embodiments, the Cas effector protein is a catalytically inactive Cas effector protein.

In some embodiments, the regulon moiety is PABPC1 or NAT10.

In some embodiments, the fusion protein further comprises a linker and/or a spacer. In some embodiments, the fusion protein further comprises a nuclear export signal and/or an epitope tag.

In some embodiments, the RNA effector protein is N terminal to the regulon moiety. In some embodiments, the RNA effector protein is C terminal to the regulon moiety.

In some embodiments, the fusion protein comprises or consists of SEQ ID NO: 48 or SEQ ID NO: 49 or a polypeptide having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to SEQ ID NO: 48 or SEQ ID NO: 49.

In some embodiments, the gRNA targets an mRNA encoding MeCP2, SCN1A, SYNGAP1, SHANK3, CHD2, or PTEN. In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16 , SEQ ID NO: 17, SEQ ID NO: 18, and combinations thereof. In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 17, and combinations thereof.

In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding an amino acid selected from the group consisting of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, and combinations thereof.

In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID

NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, and combinations thereof.

In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding an amino acid selected from the group consisting of SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, and combinations thereof.

In some embodiments, the gRNA is selected from the group consisting of SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, SEQ ID NO: 60, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67, SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 70, SEQ ID NO: 71, and SEQ ID NO: 72.

Also provided herein are one or more polynucleotide(s) encoding any one of the systems described herein. Also provided herein are one or more vector(s) comprising any one of the polynucleotide(s) described herein. Also provided herein are cells comprising any one of the vector(s) described herein.

Also provided herein are complexes comprising: a fusion protein comprising: an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA(s), bound to a gRNA comprising a complementarity region that hybridizes with the mRNA(s) of the active allele.

In some embodiments, the RNA effector protein is dCas13b and the regulon moiety is PABP1 or NAT10.

In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16 , SEQ ID NO: 17, SEQ ID NO: 18, and combinations thereof. In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 17, and combinations thereof.

In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding an amino acid selected from the group consisting of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, and combinations thereof.

In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, and combinations thereof.

In some embodiments, the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, and combinations thereof.

Also provided herein are complexes comprising: a fusion protein comprising: an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA(s), bound to a gRNA and the mRNA.

In some embodiments, the RNA effector protein is dCas13b and the regulon moiety is PABP1 or NAT10.

In some embodiments, the mRNA is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16 , SEQ ID NO: 17, SEQ ID NO: 18, and combinations thereof. In some embodiments, the mRNA encodes an amino acid selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 17, and combinations thereof.

In some embodiments, the mRNA is selected from the group consisting of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, and combinations thereof.

In some embodiments, the mRNA is selected from the group consisting of SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, and combinations thereof.

In some embodiments, the mRNA encodes an amino acid selected from the group consisting of is selected from the group consisting of SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, and combinations thereof.

Also provided herein are pharmaceutical compositions comprising: any one of the fusion proteins or systems described herein. In some embodiments, the pharmaceutical compositions further comprise a pharmaceutically acceptable carrier.

Also provided herein are pharmaceutical compositions comprising: one or more nucleic acids encoding any one of the fusion proteins or systems described herein.

Also provided herein are viral vectors comprising one or more nucleic acids encoding any one of the fusion proteins or systems described herein. In some embodiments, the viral vector is an adeno-associated viral vector.

Also provided herein are nanoparticles or liposomes comprising any one of the fusion proteins or systems described herein or one or more nucleic acids encoding any one of the fusion proteins or systems described herein.

Also provided herein are methods of stimulating or stabilizing mRNA(s), the method comprising: contacting the mRNA(s) with any one of the fusion proteins or systems described herein. In some embodiments, the method is carried out, in vitro, in vivo, or ex vivo.

Also provided herein are methods of treating or preventing a haploinsufficiency disorder in a subject, the method comprising: administering to the subject a fusion protein or a nucleic acid encoding a fusion protein comprising: an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA(s); and a gRNA or a nucleic acid encoding a gRNA designed to form a complex with the RNA effector protein and comprising a complementarity region designed to hybridize with the mRNA of the active allele.

In some embodiments, the RNA effector protein is a Cas effector protein selected from the group consisting of a Cas9, Cas12, Cas13, and Cas14. In some embodiments, the Cas effector protein is Cas13b. In some embodiments, the Cas effector protein is a catalytically inactive Cas protein.

In some embodiments, the regulon moiety is PABPC1 or NAT10.

In some embodiments, the fusion protein further comprises a linker and/or a spacer. In some embodiments, the fusion protein further comprises a nuclear export signal and/or an epitope tag.

In some embodiments, the RNA effector protein is N terminal to the regulon moiety. In some embodiments, the RNA effector protein is C terminal to the regulon moiety.

In some embodiments, the fusion protein comprises or consists of SEQ ID NO: 48 or SEQ ID NO: 49 or a polypeptide having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to SEQ ID NO: 48 or SEQ ID NO: 49.

In some embodiments, the haploinsufficiency disorder is selected from the group consisting from 5qsyndrome,Adams-Oliver syndrome 1, Adams-Oliver syndrome 3, Adams-Oliver syndrome 5,Adams-Oliver syndrome 6, Alagille syndrome 1, Autoimmune lymphoproliferative syndrome type IA, Autoimmune lymphoproliferative syndrome type V,

Autosomal dominant deafness-2A,Brain malformations with or without urinary tract defects (BRMUTD), Carney complex type 1,CHARGE syndrome, Cleidocranial dysplasia, Currarino syndrome, Denys-Drash syndrome/Frasier syndrome, Developmental delay, intellectual disability, obesity, and dysmorphic features(DIDOD), DiGeorge syndrome (TBXI-associated), Dravet syndrome, Duane-radial raysyndrome, Ehlers-Danlos syndrome (classic-like), Ehlers-Danlos syndrome (vascular type),Feingold syndrome 1, Frontotemporal lobar degeneration with TDP43 inclusions (FTLD-TDP),GRN-related, GLUT I deficiency syndrome, Greig cephalopolysyndactyly syndrome, Hereditary hemorrhagic telangiectasia type 1, Holoprosencephaly 3, Holoprosencephaly 4,Holoprosencephaly 5, Holt-Oram syndrome, Hypoparathyroidism, sensorineural deafness, andrenal disease (HDR), Kleefstra syndrome 1, Klippel-Trenaunay syndrome (AAGF-related), Leri-Weill dyschondrosteosis, Marfan syndrome, Mental retardation and distinctive facial features with or without cardiac defects (MRFACD), Mental retardation, autosomal dominant 1, Mental retardation, autosomal dominant 19, Mental retardation, autosomal dominant 29, Nail-patella syndrome (NPS), Phelan-McDermid syndrome, Pitt-Hopkins syndrome, Primary pulmonary hypertension 1, Rett syndrome (congenital variant), Smith-Magenis syndrome (RAII associated), Sotos syndrome 1, Sotos syndrome 2, Stickler syndrome type I, Supravalvular aorticstenosis, SYNGAPI-related intellectual disability, Treacher Collins syndrome, Trichorhinophalangeal syndrome type I, Ulnar-mammary syndrome, van der Woude syndrome1, Waardenburg syndrome type 1, W aardenburg syndrome type 2A, and Waardenburg syndrometype 4C.

In some embodiments, the haploinsufficiency disorder is a CNS haploinsufficiency disorder. In some embodiments, the CNS haploinsufficiency disorder is selected from the group consisting of episodic ataxia, familial hemiplegia migraine, CDKL5 deficiency disorder, CHD2 myoclonic encephalopathy, familial focal epilepsy with variable loci, FOXG1 syndrome, benign familial neonatal seizures, Rett syndrome, Dravat syndrome, SCN2A-epileptic encephalopathy, SCN2A-developmental encephalopathy, SCN8A-epileptic encephalopathy, SC8A familial infantile epilepsy, early infantile epileptic encephalopathy, myoclonic-atonic epilepsy, early infantile epileptic encephalopathy, SYNGAP1-related intellectual disability, tuberous sclerosis, Lennox-Gastaut Syndrome, FoxG1 syndrome, KCNQ2-related epileptic encephalopathy, PCDH19-related epilepsy, SLC6A1-related myoclonic-astatic epilepsy, STXBP1-related epileptic encephalopathy, SYNGAP1 syndrome, and combinations thereof. In some embodiments, the haploinsufficiency disorder is Dravet Syndrome or Rett syndrome.

In some embodiments, the subject has a haploinsufficiency in a gene selected from the group consisting of AGGFI, ARHGAP31, BMPR2, CHD7, COL2Al, COL3Al, CTLA4, CTNNBI, DLL4, EHMTI, ELN,ENG, FAS, FBNI, FOXGI, GATA3, GLI3, GRN, IRF6, JAGI,

KCNQ4, LMXIB, MBD5,MED13L, MITF, MNXI, MYCN, NFIA, NFIX, NOTCH!, NSDI, PAX3, PHIP, PRKARIA,RAil, RBPJ, RPS14, RUNX2, SALL4, SCNIA, SETBPI, SHANK3, SHH, SHOX,SLC2Al/GLUT1, SOXI0, SYNGAPI, TBXI, TBX3, TBX5, TCF4, TCOFI, TGIFI, TNXB,TRPSI, WTI, ZIC2, and combinations thereof. In some embodiments, the subject has a haploinsufficiency in a gene selected from the group consisting of SCN1A, SCN2A, SCN8A, SCN12A5, SPTAN1, CDKL5, CHD2, FOXG1, KCNQ2, PCDH19, SLC6A1, STXBP1, SYNGAP1, CACNA1A, DEPDC5, MECP2, TSC1, TSC2, and combinations thereof.

In some embodiments, the subject has mutation selected from the list in Table 4 and combinations thereof. In some embodiments, the subject has a mutation selected from the list in Table 6 and combinations thereof. In some embodiments, the subject is a mammal. In some embodiments, the subject is a human.

In some embodiments, the fusion protein and gRNA are administered as part of a pharmaceutical composition. In some embodiments, administering comprises administering a viral vector comprising nucleic acid sequence(s) encoding the fusion protein and gRNA to the subject. In some embodiments, administering comprises administering a nanoparticle or liposome comprising the fusion protein and gRNA or nucleic acid sequence(s) encoding the fusion protein and gRNA to the subject.

Throughout this application, various embodiments may be 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 disclosure. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges 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 subranges 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.

As used in the specification and claims, the singular forms “a”, “an” and “the” include plural references unless the context clearly dictates otherwise. For example, the term “a sample” includes a plurality of samples, including mixtures thereof.

The terms “determining,” “measuring,” “evaluating,” “assessing,” “assaying,” and “analyzing” are often used interchangeably herein to refer to forms of measurement. The terms include determining if an element is present or not (for example, detection). These terms can include quantitative, qualitative or quantitative and qualitative determinations. Assessing can be relative or absolute. “Detecting the presence of” can include determining the amount of

something present in addition to determining whether it is present or absent depending on the context.

As used herein, the term "about" a number refers to that number plus or minus 10% of that number. The term "about" a range refers to that range minus 10% of its lowest value and plus 10% of its greatest value.

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 to which this invention belongs. Methods and materials are described herein for use in the present invention; other, suitable methods and materials known in the art can also be used. The materials, methods, and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, sequences, database entries, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

Other features and advantages of the invention will be apparent from the following detailed description and figures, and from the claims.

DESCRIPTION OF DRAWINGS

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

FIG. 1 is a schematic showing SCN1A topology and premature-termination codons (PTC). Each of the 182 SCN1A PTCs is show by color.

FIG. 2 is a schematic representing the tethered function assay. The fused functional protein to the tethered protein bind to the mRNA of interest to enhance or stabilize the mRNA.

FIG. 3A is a schematic showing the experimental approach applied for this assay. Briefly, PABPC1 was fused to dCas13b tethering protein and guided to 3'UTR of the Luciferase gene via designed guide RNAs(gRNA) to enhance its translation.

FIG. 3B shows a firefly luciferase assessment depicting about 2-folds increase in PAB-tethered luciferase targeting over non-targeting in HEK-293 cells.

FIG. 4 shows a map of a plasmid pJC1211 encoding a Cas13b-PABPC1 fusion protein.

FIGS. 5A–5AB show the sequence and features of the plasmid pJC1211 (SEQ ID NO: 50). The nucleic acid sequence is shown in both 5'->3' and the reverse complement (3'->5').

FIG. 6 is a map of plasmid pJC1212 encoding a luciferase targeting Cas13b gRNA.

FIGS. 7A–7G show the sequence and features of plasmid pJC1212 (SEQ ID NO: 52). The nucleic acid sequence is shown in both 5'->3' and the reverse complement (3'->5').

FIG. 8 is a map of plasmid pJC1213 encoding a luciferase targeting Cas13b gRNA.

FIGS. 9A–9F show the sequence and features of plasmid pJC1213 (SEQ ID NO: 53).

FIG. 10 is a map of Addgene plasmid 103854 (empty vector as compared to pJC1212 and pJC1213).

FIGS. 11A–11H show the sequence and features of Addgene plasmid 103854 (empty vector as compared to pJC1212 and pJC1213) (SEQ ID NO: 54).

FIGS. 12A-12G shows the fusion protein approach enhances mRNA expression in multiple cell types. FIG. 12A is an exemplary schematic representing the tethered mRNA amplifier approach. Briefly, PABPC1 is fused to dCas13b. This fusion is recruited to specific mRNAs via a guide RNA (gRNA) targeted to the 3'UTR. In HEK293 cells, the Tethered mRNA Amplifier stimulates gRNA-dependent luciferase reporter (FIG. 12B) and endogenous MeCP2 mRNA (FIG. 12C) expression, using either a luciferase activity assay or western blot, respectively. Red bars in the corresponding schematics represent the approximate positions of two gRNAs used in each experiment. The effect of the Tethered mRNA Amplifier requires PABPC1 (FIG. 12D). The Tethered mRNA Amplifier also enhances MeCP2 mRNA levels (FIG. 12E). A similar stimulatory effect on MeCP2 can also be seen in SH-SY5Y and HepG2 cells (FIG. 12F). FIG. 12G shows that the position of the gRNA along the MeCP2 3'UTR alters the stimulatory effect. (p-values: *<0.05, ** <0.005, *** <0.0005).

FIGS. 13A-13D show the fusion protein enhances the expression of haploinsufficiency disorder-associated transcripts; SYNGAP1 (FIG. 13A), SHANK3 (FIG. 13B), PTEN (FIG. 13C), and CHD2 (FIG. 13D) mRNAs were targeted by the Tethered mRNA Amplifier in a gRNA-dependent manner. The result of protein and mRNA analysis in SH-SY5Y cells are shown. All the protein assays were performed with four distinct biological replicates and at least two biological repeats for the RNA analysis. (p-values: *<0.05, ** <0.005, *** <0.0005).

FIGS. 14A-14F show a minimal fusion protein functions in cells. FIG. 14A is an exemplary schematic of PABPC1 and its functional motifs. RRM1-4 are RNA Recognition Motifs. MLLE is the Mademoiselle domain. FIG. 14B shows AlphaFold predicted model of dCas13b fusion with the MLLE domain of PABPC1 (amino acids 545-636). FIG. 14C shows a predicted model of full-length PABPC1 – all residues except amino acids 545-636 have been hidden post-prediction. FIG. 14D shows a predicted model of full-length dCas13b alone. FIG. 14E shows PyMol alignment of FIGS. 14B-14D. FIG. 14F shows Western blots comparing tethering of full-length PABPC1 and MLLE domain alone targeting MeCP2 transcripts in HEK293 cells.

FIG. 15 is an exemplary schematic showing haploinsufficiency disorders result when one copy of a gene is mutated while the other is normal. This mutation ultimately reduces protein expression by one half, causing a disease state. The mRNA Amplifier targets the messenger RNA from the normal gene theoretically doubling protein expression to normal amounts.

DETAILED DESCRIPTION

The compositions and methods described herein have a broad range of applications, providing potential disease modifying therapies for a range of haploinsufficiency disorders. Haploinsufficiency occurs when one gene allele is inactivated and the amount of gene product expressed from the remaining active allele is insufficient for proper gene function. A number of disorders are associated with, or are caused by haploinsufficiency. An example of a haploinsufficiency disorder is Dravet Syndrome.

In the treatment of haploinsufficiencies, targeting messenger RNA metabolism offers a novel therapeutic window. Though changes in gene expression are commonly considered to reflect programmed transcriptional variability. Less commonly considered is that extensive regulation of messenger RNA expression also occurs during translation. For example, in the early oocyte, large changes in protein expression occur via coordinated activation/deactivation of maternally derived mRNA. Likewise, translationally quiescent transcripts are activated upon synaptic stimulation in neurons. A strong discordance between mRNA levels and protein levels is also observed in somatic cells, highlighting the prevalence and physiological significance of a post-transcriptional regulon.

Transcript stability and translation rate are a key feature defining the post-transcriptional regulon. All transcripts degrade and are translated at unique rates. mRNA stability and translation are often proportional to gene function; housekeeping genes typically produce long lived mRNAs that persist in cells for over 24 hours, whereas transiently expressed genes generate short-lived transcripts that survive on the order of minutes. In part, the cell achieves this vast repertoire of transcript stabilities and expression though sequence and/or structural elements that recruit specific positive or negative acting factors. The compositions and methods described herein, among other things, utilize this mRNA / post transcriptional regulon as a novel disease modifying treatment for haploinsufficiency disorders such as Dravet Syndrome.

MRNA REGULATION THERAPEUTICS

Described herein are novel molecular therapeutic strategies using posttranscriptional regulation of mRNA, e.g., SCN1A mRNA, to overcome haploinsufficiency in genetic brain disorders. The theory of the technology is based on what is termed “A Tethered Function

Approach" or a "Tethered mRNA Amplifier Approach". In brief, this technology tethers a specific factor to the 3' untranslated region (3'UTR) of a target mRNA, thereby changing its stability and/or translational rate. The approach is commonly used to determine the function of RNA binding proteins [18, 19].

Proteins and protein complexes that regulate mRNA metabolism possess two activities. They bind an mRNA specifically, and then elicit some function, that is, regulate mRNA splicing, transport, localization, translation, or stability. These two activities can often reside in different proteins in a complex, or in different regions of a single polypeptide. In the majority of instances involving mRNAs, specific RNA binding activity and function are distinct. In these cases, the separation of the two activities from one another provides a powerful inroad for a therapeutic modality. In essence, a functional activity (stripped of its native RNA binding activity) can be tethered to a new mRNA via a unique RNA binding activity.

In the tethered function assay (e.g., as shown in FIG. 2), to determine the effects of a protein X on mRNA metabolism, a chimeric protein is expressed *in vivo* in which protein X is continuous with a tethering polypeptide (see FIG. 2). The tethering protein is an RNA-binding protein that recognizes an RNA tag sequence with high specificity and affinity. The effect of the fusion protein on mRNA metabolism is determined by co-expressing the chimera with an mRNA reporter (such as lacZ or luciferase) into which a tag RNA sequence has been embedded. The fusion protein's effects on mRNA metabolism are assayed by conventional means [e.g., Western blot, Northern blot, reverse transcriptase polymerase chain reaction (RT-PCR), etc.]. The assay, however, until now, has only been utilized in model organisms / cell culture to evaluate the contribution of RNA binding proteins to the post-transcriptional regulation of mRNA. Described herein, is a novel therapeutic modality, based on tethering, for disease modification, e.g. for modification of Dravet Syndrome.

Tethered function assays were developed as a tool to dissect the function of unknown RNA binding protein in the posttranscriptional control of mRNA. The contrived system took advantage of exogenous RNA binding activities, often of bacteriophage origin (MS2 Coat Protein and Lamda N-peptide being the most common). The obvious limitation to this approach was that a corresponding RNA binding element had to be cloned and expressed in a reporter mRNA. The novel Protein/RNA binding was utilized to "tether" an activity to an mRNA and test its function. Critically, having the ability to tether a functional activity to ANY mRNA of interest would be powerful in changing that mRNAs expression pattern. And this approach could be leveraged in the treatment of haploinsufficiency.

CRISPR-based research tools established a system capable of selective recognition of RNA molecules. This novel RNA targeting system has enabled the detecting and manipulation of

specific RNA using different approaches, such as RNA-knock-down, site-specific RNA editing, RNA localization, and destruction of toxic RNAs that lead to human neurodegenerative disorders [20, 21].

In some cases, the mRNA regulon therapy described herein utilize the CRISPR-based RNA binding approach, providing a precise and advanced technology to target genetic disorders at the transcript level. The RNA targeting CRISPR-Cas13 has RNA strand specificity and binds with high affinity; enabling efficient and precise delivery to relevant utilize the mRNA-specific targeting capacity of Cas13b to tether known factors that enhance mRNA expression to ameliorate haploinsufficiency (FIG. 15).

HAPLOINSUFFICIENCY DISORDERS AND GENES

In some cases, the haploinsufficiency disorder is selected from 5qsyndrome,Adams-Oliver syndrome 1, Adams-Oliver syndrome 3, Adams-Oliver syndrome 5,Adams-Oliver syndrome 6, Alagille syndrome 1, Autoimmune lymphoproliferative syndrome type IA, Autoimmune lymphoproliferative syndrome type V, Autosomal dominant deafness-2A,Brain malformations with or without urinary tract defects (BRMUTD), Carney complex type 1,CHARGE syndrome, Cleidocranial dysplasia, Currarino syndrome, Denys-Drash syndrome/Frasier syndrome, Developmental delay, intellectual disability, obesity, and dysmorphic features(DIDOD), DiGeorge syndrome (TBXI-associated), Dravet syndrome, Duane-radial raysyndrome, Ehlers-Danlos syndrome (classic-like), Ehlers-Danlos syndrome (vascular type),Feingold syndrome 1, Frontotemporal lobar degeneration with TDP43 inclusions (FTLD-TDP),GRN-related, GLUT I deficiency syndrome, Greig cephalopolysyndactyly syndrome, Hereditary hemorrhagic telangiectasia type 1, Holoprosencephaly 3, Holoprosencephaly 4,Holoprosencephaly 5, Holt-Oram syndrome, Hypoparathyroidism, sensorineural deafness, andrenal disease (HDR), Kleefstra syndrome 1, Klippel-Trenaunay syndrome (AAGF-related), Leri-Weill dyschondrosteosis, Marfan syndrome, Mental retardation and distinctive facial features with or without cardiac defects (MRFACD), Mental retardation, autosomal dominant 1, Mental retardation, autosomal dominant 19, Mental retardation, autosomal dominant 29, Nail-patella syndrome (NPS), Phelan-McDermid syndrome, Pitt-Hopkins syndrome, Primary pulmonary hypertension 1, Rett syndrome (congenital variant), Smith-Magenis syndrome (RAII associated), Sotos syndrome 1, Sotos syndrome 2, Stickler syndrome type I, Supravalvular aorticstenosis, SYNGAPI-related intellectual disability, Treacher Collins syndrome, Trichorhinophalangeal syndrome type I, Ulnar-mammary syndrome, van der Woude syndrome1, Waardenburg syndrome type 1, W aardenburg syndrome type 2A, and

Waardenburg syndrometype 4C. In some cases, the haploinsufficient disorder is Dravet syndrome.

In some cases, the haploinsufficient gene is selected from *AGGFI, ARHGAP31, BMPR2, CHD7, COL2Al, COL3Al, CTLA4, CTNNBI, DLL4, EHMTI, ELN,ENG, FAS, FBNI, FOXGI, GATA3, GLI3, GRN, IRF6, JAGI, KCNQ4, LMXIB, MBD5,MED13L, MITF, MNXI, MYCN, NFIA, NFIX, NOTCH!, NSDI, PAX3, PHIP, PRKARIA,RAil, RBPJ, RPS14, RUNX2, SALL4, SCNIA, SETBPI, SHANK3, SHH, SHOX,SLC2Al/GLUT1, SOXI0, SYNGAPI, TBXI, TBX3, TBX5, TCF4, TCOFI, TGIFI, TNXB,TRPSI, WTI, ZIC2,* and combinations thereof. In some cases embodiments, the haploinsufficient gene is *SCNIA*.

In some cases, the combination of haploinsufficiency disorder and haploinsufficient gene combination is a combination shown in Table 1.

Table 1. Haploinsufficiency disorders and genes.

Haploinsufficiency Disorder	Haploinsufficient Gene
5q-syndrome	RPS14
Adams-Oliver syndrome I	ARHGAP31
Adams-Oliver syndrome 3	RBPJ
Adams-Oliver syndrome 5	NOTCH1
Adams-Oliver syndrome 6	DLL4
Alagille syndrome I	JAG1
Autoimmune lymphoproliferative syndrome type IA	FAS
Autoimmune lymphoproliferative syndrome type V	CTLA4
Autosomal dominant deafness-2A	KCNQ4
Brain malformations with or without urinary tract defects (BRMUTD)	NFIA
Carney complex type 1	PRKARlA
CHARGE syndrome	CHD7
Cleidocranial dysplasia	RUNX2
Currarino syndrome	MNXl
Denys-Drash syndrome/ Frasier syndrome	WTl
Developmental delay, intellectual disability, obesity, and dysmorphic features (DIDOD)	PHIP
DiGeorge syndrome (TBXl-associated)	TBXl
Dravet syndrome	SCNlA
Duane-radial ray syndrome	SALL4
Ehlers-Danlos syndrome (classic-like)	TNXB
Ehlers Danlos syndrome (vascular type)	COL3Al
Feingold syndrome 1	MYCN
Frontotemporal lobar degeneration with TDP43 inclusions (FTLD-TDP), GRN-related	GRN
GLUT1 deficiency syndrome	SLC2Al
Greig cephalopolysyndacty1y syndrome	GLI3
Hereditary hemorrhagic telangiectasia type 1	ENG
Holoprosencephaly 3	SHH
Holoprosencephaly 4	TGIFl

Haploinsufficiency Disorder	Haploinsufficient Gene
Holoprosencephaly 5	ZIC2
Holt-Oram syndrome	TBX5
Hypoparathyroidism, sensorineural deafness, and renal disease (HDR)	GATA3
Kleefstra syndrome 1	EHMTl
Klippel-Trenaunay syndrome (AAGF-related)	AGGFl
Leri-Weill dyschondrosteosis	SHOX
Marfan syndrome	FBNl
Mental retardation and distinctive facial features with or without cardiac defects (MRFACD)	MED13L
Mental retardation, autosomal dominant 1	MBD5
Mental retardation, autosomal dominant 19	CTNNBl
Mental retardation, autosomal dominant 29	SETBPl
Nail-patella syndrome (NPS)	LMXlB
Phelan-McDermid syndrome	SHANK3
Pitt-Hopkins syndrome	TCF4
Primary pulmonary hypertension 1	BMPR2
Rett syndrome (congenital variant)	FOXGl
Smith-Magenis syndrome (RAIl-associated)	RAI1
Sotos syndrome 1	NSDl
Sotos syndrome 2	NFIX
Stickler syndrome type I	COL2Al
Supravalvular aortic stenosis	ELN
SYNGAPl-related intellectual disability	SYNGAPl
Treacher Collins syndrome	TCOFl
Trichorhinophalangeal syndrome type I	TRPSl
Ulnar-mammary syndrome	TBX3
van der Woude syndrome 1	IRF6
Waardenburg syndrome type 1	PAX3
Waardenburg syndrome type 2A	MITF
Waardenburg syndrome type 4C	SOXlO

CNS Haploinsufficiency Disorders

In some cases, the haploinsufficiency disorder is a CNS haploinsufficiency disorder. In some cases, the haploinsufficiency disorder is selected from the group consisting of episodic ataxia, familial hemiplegia migraine, CDKL5 deficiency disorder, CHD2 myoclonic encephalopathy, familial focal epilepsy with variable loci, FOXG1 syndrome, benign familial neonatal seizures, Rett syndrome, Dravat syndrome, SCN2A-epileptic encephalopathy, SCN2A-developmental encephalopathy, SCN8A-epileptic encephalopathy, SC8A familial infantile epilepsy, early infantile epileptic encephalopathy, myoclonic-atonic epilepsy, early infantile epileptic encephalopathy, SYNGAP1-related intellectual disability, tuberous sclerosis, Lennox-Gastaut Syndrome, FoxG1 syndrome, KCNQ2-related epileptic encephalopathy, PCDH19-related epilepsy, SLC6A1-related myoclonic-astatic epilepsy, STXBP1-related epileptic encephalopathy, SYNGAP1 syndrome, and combinations thereof.

In some cases, the haploinsufficiency gene is selected from the group consisting of *SCN1A, SCN2A, SCN8A, SCN12A5, SPTAN1, CDKL5, CHD2, FOXG1, KCNQ2, PCDH19, SLC6A1, STXBP1, SYNGAP1, CACNA1A, DEPDC5, MECP2, TSC1, TSC2,* and combinations thereof.

In some cases, the combination of haploinsufficiency disorder and haploinsufficient gene combination is a combination shown in Table 2

Table 2. CNS haploinsufficiency disorders and genes

Haploinsufficiency Disorder	Haploinsufficient Gene
Episodic ataxia, Familial hemiplegia migraine	CACNA1A
CDKL5 deficiency disorder	CDKL5
CHD2 myoclonic encephalopathy	CHD2
Lennox-Gastaut Syndrome	CHD2
Familial focal epilepsy with variable loci	DEPDC5
FOXG1 syndrome	FOXG1
Benign familial neonatal seizures	KCNQ2
Rett syndrome	MECP2
Dravet syndrome	SCN1A
SCN2A-epileptic encephalopathy	SCN2A
SCN2A-developmental encephalopathy	SCN2A
SCN8A-epileptic encephalopathy	SCN8A
SCN8A familial infantile epilepsy	SCN8A
Early infantile epileptic encephalopathy	SLC12A5
Myoclonic-atonic epiplepsy	SLC6A1
Early infantile epileptic encephalopathy	SPTAN1
SYNGAP1-related intellectual disability	SYNGAP1
Tuberous sclerosis	TSC1
Tuberous sclerosis	TSC2
KCNQ2-related epileptic encephalopathy	KCNQ2
SLC6A1-related myoclonic-astatic epilepsy	SLC6A1
STXBP1-related epileptic encephalopathy	STXBP1
SYNGAP1 syndrome	SYNGAP1

SCN1A

SCN1A (NCBI Gene ID: 6323; RefSeq NG_011906.1) encodes a sodium voltage-gated channel alpha subunit 1. Its transcripts and protein isoforms shown in Table 3.

FIG. 1 is a schematic showing SCN1A topology and premature-termination codons (PTC).

Table 3. SCN1A transcripts and isoforms

NCBI transcript accession	Transcript #	SEQ ID NO:	NCBI protein accession	Isoform #	SEQ ID NO:
NR_148667.2	16	SEQ ID NO: 1			

NCBI transcript accession	Transcript #	SEQ ID NO:	NCBI protein accession	Isoform #	SEQ ID NO:
XR_001738884.1	X2	SEQ ID NO: 2			
XR_001738883.1	X1	SEQ ID NO: 3			
NM_001353949.2	6	SEQ ID NO: 4	NP_001340878.1	2	SEQ ID NO: 21
NM_001353958.2	13	SEQ ID NO: 5	NP_001340887.1	3	SEQ ID NO: 20
NM_001353950.2	7	SEQ ID NO: 6	NP_001340879.1	2	SEQ ID NO: 21
NM_001202435.3	4	SEQ ID NO: 7	NP_001189364.1	1	SEQ ID NO: 19
NM_001353955.2	11	SEQ ID NO: 8	NP_001340884.1	4	SEQ ID NO: 22
NM_001353957.2	12	SEQ ID NO: 9	NP_001340886.1	3	SEQ ID NO: 20
NM_001353951.2	8	SEQ ID NO: 10	NP_001340880.1	2	SEQ ID NO: 21
NM_001353948.2	5	SEQ ID NO: 11	NP_001340877.1	1	SEQ ID NO: 19
NM_001353960.2	14	SEQ ID NO: 12	NP_001340889.1	5	SEQ ID NO: 23
NM_001353954.2	10	SEQ ID NO: 13	NP_001340883.1	4	SEQ ID NO: 22
NM_001165964.3	3	SEQ ID NO: 14	NP_001159436.1	3	SEQ ID NO: 20
NM_001353952.2	9	SEQ ID NO: 15	NP_001340881.1	2	SEQ ID NO: 21
NM_001353961.2	15	SEQ ID NO: 16	NP_001340890.1	6	SEQ ID NO: 24
NM_006920.6	2	SEQ ID NO: 17	NP_008851.3	2	SEQ ID NO: 21
NM_001165963.4	1	SEQ ID NO: 18	NP_001159435.1	1	SEQ ID NO: 19

Mutations in SCN1A that are associated with Dravat Syndrome are shown in Table 4.

Table 4. Mutations in SCN1A associated with Dravat Syndrome

Position(s) in SEQ ID NO: 19	Description
17	Missing in DRVT.
58	G → V in DRVT.
61	L → F in DRVT.
63	F → L in DRVT. Corresponds to variant dbSNP:rs121917907
68	I → T in DRVT; borderline phenotype. Corresponds to variant dbSNP:rs758871507
78	E → D in DRVT. Corresponds to variant dbSNP:rs121917933

Position(s) in SEQ ID NO: 19	Description
79	D → H in DRVT; borderline phenotype. Corresponds to variant dbSNP:rs121917982
79	D → N in DRVT. Corresponds to variant dbSNP:rs121917982
84	Y → C in DRVT. 5 Publications Corresponds to variant dbSNP:rs121917964
90	F → S in DRVT and ICEGTC. Corresponds to variant dbSNP:rs121918733
91	I → T in DRVT. Corresponds to variant dbSNP:rs121918734
98	A → P in DRVT.
101	R → Q in DRVT and ICEGTC. 8 Publications Corresponds to variant dbSNP:rs121917918
101	R → W in DRVT. Corresponds to variant dbSNP:rs121917965.
103	S → G in DRVT. Corresponds to variant dbSNP:rs121918743
105	T → I in DRVT. Corresponds to variant dbSNP:rs796053089
108	L → R in DRVT.
112	T → I in DRVT. Corresponds to variant dbSNP:rs121918745
113	P → T in DRVT. Corresponds to variant dbSNP:rs794726711
118	R → S in DRVT. Corresponds to variant dbSNP:rs121917959
124	I → N in DRVT. Corresponds to variant dbSNP:rs121918761
127	H → D in DRVT; borderline phenotype. Corresponds to variant dbSNP:rs148442069
162	T → P in DRVT. Corresponds to variant dbSNP:rs121917934
171	I → K in DRVT. Corresponds to variant dbSNP:rs121918766
171	I → R in DRVT.
175	A → T in DRVT. Corresponds to variant dbSNP:rs121918767
175	A → V in DRVT.
177	G → E in DRVT; results in a non-functional channel. Corresponds to variant dbSNP:rs121918770
179	C → R in DRVT.
190	W → R in DRVT. Corresponds to variant dbSNP:rs121918773
191	N → K in DRVT.
191	N → Y in DRVT. Corresponds to variant dbSNP:rs121918762
194	D → G in DRVT.
194	D → N in DRVT. Corresponds to variant dbSNP:rs121917935
199	T → R in DRVT; borderline phenotype with spike wave activity. Corresponds to variant dbSNP:rs121917983
217	T → K in DRVT. Corresponds to variant dbSNP:rs121917936
223	A → E in DRVT.
226	T → M in DRVT; borderline phenotype; also found in a patient with cryptogenic generalized epilepsy. Corresponds to variant dbSNP:rs121917984
226	T → R in DRVT. Corresponds to variant dbSNP:rs121917984
227	I → S in DRVT; borderline phenotype with spike wave activity in some patients; results in a non-functional channel. Corresponds to variant dbSNP:rs121917937
227	I → T in DRVT.
228	S → P in DRVT. Corresponds to variant dbSNP:rs1057519530
232	G → S in DRVT.

Position(s) in SEQ ID NO: 19	Description
233	L → R in DRVT.
239	A → T in DRVT; borderline phenotype with spike wave activity in some patients. Corresponds to variant dbSNP:rs121917985
239	A → V in DRVT. Corresponds to variant dbSNP:rs121917909
243	S → Y in DRVT. Corresponds to variant dbSNP:rs794726755
252	I → N in DRVT. Corresponds to variant dbSNP:rs121918780
259	S → R in DRVT. Corresponds to variant dbSNP:rs121918735
265	G → W in DRVT. Corresponds to variant dbSNP:rs121918749
277	C → R in DRVT.
280	W → C in DRVT.
280	W → R in DRVT. Corresponds to variant dbSNP:rs121917938
281	P → A in DRVT.
281	P → L in DRVT. Corresponds to variant dbSNP:rs796052964
281	P → S in DRVT.
289	E → V in DRVT.
297	T → I in DRVT. Corresponds to variant dbSNP:rs121918771
322	R → I in DRVT. Corresponds to variant dbSNP:rs121917928
340	S → F in DRVT.
342	A → V in DRVT. Corresponds to variant dbSNP:rs794726797
343	G → D in DRVT. Corresponds to variant dbSNP:rs121918753
345	C → R in DRVT. Corresponds to variant dbSNP:rs794726782
351	C → W in DRVT.
355	G → D in DRVT.
356	R → G in DRVT. Corresponds to variant dbSNP:rs121917920
357	N → I in DRVT.
358	P → T in DRVT. Corresponds to variant dbSNP:rs121917923
359	N → S in DRVT and ICEGTC. Corresponds to variant dbSNP:rs794726713
363	T → P in DRVT. Corresponds to variant dbSNP:rs1131691465
363	T → R in DRVT.
366	D → E in DRVT. Corresponds to variant dbSNP:rs121917958
378	L → Q in DRVT.
379	M → R in DRVT.
383	F → L in DRVT. Corresponds to variant dbSNP:rs121917939
384	W → R in DRVT; borderline phenotype. Corresponds to variant dbSNP:rs1057523858
393	R → C in DRVT; also in a patient with myoclonic astatic epilepsy. Corresponds to variant dbSNP:rs121917929
393	R → H in DRVT and ICEGTC; results in a non-functional channel. Corresponds to variant dbSNP:rs121917927
393	R → S in DRVT. Corresponds to variant dbSNP:rs121917929
400	M → V in DRVT.
400	Missing in DRVT.
403	F → L in DRVT. Corresponds to variant dbSNP:rs121917966
403	F → V in DRVT.
406	V → F in DRVT. Corresponds to variant dbSNP:rs121918768
409	L → W in DRVT.
413	Y → N in DRVT. Corresponds to variant dbSNP:rs121917967

Position(s) in SEQ ID NO: 19	Description
426	Y → C in DRVT. Corresponds to variant dbSNP:rs796052973
426	Y → N in DRVT; results in decreased peak current densities; causes a negative shift in the half-maximal steady-state inactivation and delayed recovery from fast inactivation. Corresponds to variant dbSNP:rs121917940
450 – 2009	Missing in DRVT.
525	S → F in DRVT.
626	S → G in DRVT; also found in a patient with cryptogenic generalized epilepsy. Corresponds to variant dbSNP:rs121917990
674	D → G in DRVT.
762	N → D in DRVT.
783	L → P in DRVT. Corresponds to variant dbSNP:rs121917968
785	M → T in DRVT. Corresponds to variant dbSNP:rs796053095
812	T → I in DRVT; borderline phenotype.
812	T → R in DRVT. Corresponds to variant dbSNP:rs121917941
842	L → R in DRVT.
843	S → R in DRVT.
846	E → K in DRVT. Corresponds to variant dbSNP:rs121917942
854 – 855	Missing in DRVT.
859	R → C in GEFS+2 and DRVT; causes a positive shift in the voltage dependence of channel activation, slower recovery from slow inactivation and lower levels of current compared with the wild-type channel. Corresponds to variant dbSNP:rs121918784
862	R → Q in DRVT. Corresponds to variant dbSNP:rs121918785
865	R → G in DRVT; results in impaired channel fast inactivation and significantly increased persistent current.
875	T → K in DRVT. Corresponds to variant dbSNP:rs121918623
875	T → M in GEFS+2 and DRVT; borderline phenotype. Corresponds to variant dbSNP:rs121918623
876	L → I in DRVT.
890	L → P in DRVT. Corresponds to variant dbSNP:rs1553541473
896	V → F in DRVT; borderline phenotype.
896	V → L in DRVT.
902	F → C in DRVT. Corresponds to variant dbSNP:rs121918787
927	C → F in DRVT. Corresponds to variant dbSNP:rs794726811
931	R → C in DRVT. Corresponds to variant dbSNP:rs121918788
932	W → C in DRVT.
933	H → P in DRVT.
934	M → I in DRVT. Corresponds to variant dbSNP:rs121918774
939	H → P in DRVT; unknown pathological significance.
939	H → Q in DRVT; results in a non-functional channel. Corresponds to variant dbSNP:rs121918795
939	H → Y in DRVT. Corresponds to variant dbSNP:rs121918736
940	S → F in DRVT. Corresponds to variant dbSNP:rs1057521080
942	L → P in DRVT. Corresponds to variant dbSNP:rs121917943
943	I → N in DRVT.
944	V → A in DRVT and ICEGTC. Corresponds to variant dbSNP:rs121917969

Position(s) in SEQ ID NO: 19	Description
944	V → E in DRVT.
945	F → L in DRVT. Corresponds to variant dbSNP:rs121917970
946	R → C in DRVT; loss-of-function mutation resulting in complete absence of sodium current. Corresponds to variant dbSNP:rs121918775
946	R → H in DRVT and GEFS+2; GEFS+2 phenotype consists of partial epilepsy with antecedent febrile seizures and seizure aggravation by antiepileptic drugs; loss-of-function mutation resulting in complete absence of sodium current. Corresponds to variant dbSNP:rs121917971
946	R → S in DRVT. Corresponds to variant dbSNP:rs121918775
949	C → S in DRVT.
949	C → Y in DRVT.
950	G → E in DRVT. Corresponds to variant dbSNP:rs121917972
950	G → R in DRVT.
952	W → G in DRVT. Corresponds to variant dbSNP:rs121918737
954	E → K in DRVT. Corresponds to variant dbSNP:rs121918786
956	M → K in DRVT.
957	W → L in DRVT. Corresponds to variant dbSNP:rs121917917
959	C → R in DRVT; results in a non-functional channel. Corresponds to variant dbSNP:rs121918796
960	M → V in DRVT. Corresponds to variant dbSNP:rs121918750
973	M → K in DRVT.
976	M → I in DRVT and GEFS+2.
979	G → V in DRVT.
985	N → I in DRVT. Corresponds to variant dbSNP:rs121918747
986	L → F in DRVT; complete loss of function. Corresponds to variant dbSNP:rs121918625
986	L → P in DRVT.
987	F → L in DRVT.
993	S → R in DRVT; borderline phenotype.
998	D → G in DRVT. Corresponds to variant dbSNP:rs1484321812
999 – 1000	NL → LIS in DRVT.
1068	E → K in DRVT.
1207	L → P in DRVT. Corresponds to variant dbSNP:rs121917963
1208	R → K in DRVT. Corresponds to variant dbSNP:rs1559149025
1210	T → K in DRVT. Corresponds to variant dbSNP:rs121918738
1221	E → K in DRVT.
1230	L → F in DRVT and GEFS+2.
1231	S → R in DRVT. Corresponds to variant dbSNP:rs121918746
1231	S → T in DRVT. Corresponds to variant dbSNP:rs121918800
1233	G → R in DRVT. Corresponds to variant dbSNP:rs121917911
1238	E → D in DRVT. Corresponds to variant dbSNP:rs121917973
1239	D → G in DRVT.
1239	D → Y in DRVT.
1245	R → Q in DRVT. Corresponds to variant dbSNP:rs121917912
1255	A → D in DRVT.
1260	T → P in DRVT. Corresponds to variant dbSNP:rs121918739
1263	F → L in DRVT. Corresponds to variant dbSNP:rs121918752
1265	L → P in DRVT. Corresponds to variant dbSNP:rs121918794

Position(s) in SEQ ID NO: 19	Description
1266	E → A in DRVT.
1275	G → V in DRVT.
1284 – 2009	Missing in DRVT. Add BLAST
1284	W → S in DRVT.
1287	L → P in DRVT. Corresponds to variant dbSNP:rs121918740
1288	D → N in DRVT.
1289	Missing in DRVT; results in a non-functional channel. 3 Publications
1316	R → G in DRVT.
1316	R → S in DRVT.
1320	A → V in DRVT; borderline phenotype.
1326	A → P in DRVT. Corresponds to variant dbSNP:rs121918803
1328	S → P in ICEGTC and DRVT.
1335	V → M in DRVT. Corresponds to variant dbSNP:rs121917960
1339	A → V in DRVT. Corresponds to variant dbSNP:rs794726789
1344	I → M in DRVT.
1345	P → L in DRVT.
1350	V → G in DRVT.
1355	L → P in DRVT. Corresponds to variant dbSNP:rs121918776
1358	W → R in DRVT.
1358	W → S in DRVT. Corresponds to variant dbSNP:rs121917961
1367	N → K in DRVT. Corresponds to variant dbSNP:rs121918760
1370	A → P in DRVT; borderline phenotype.
1378	N → H in DRVT.
1378	N → T in DRVT. Corresponds to variant dbSNP:rs1131691775
1385	F → V in DRVT.
1390	V → M in DRVT; some patients have a borderline DRVT phenotype. Corresponds to variant dbSNP:rs121917986
1391	N → S in DRVT. Corresponds to variant dbSNP:rs1553525062
1393	H → P in DRVT; borderline phenotype.
1394	T → I in DRVT.
1396	C → G in DRVT; some patients have a borderline DRVT phenotype. Corresponds to variant dbSNP:rs121917987
1396	C → Y in DRVT.
1414	N → Y in DRVT. Corresponds to variant dbSNP:rs121917925
1416	D → G in DRVT.
1417	N → S in DRVT.
1418	V → G in DRVT.
1422	Y → C in DRVT. Corresponds to variant dbSNP:rs121917913
1423	L → F in DRVT.
1426	L → R in DRVT. Corresponds to variant dbSNP:rs121917944
1427	Q → P in DRVT.
1429	Missing in DRVT.
1431	F → I in DRVT.
1433	G → E in DRVT. Corresponds to variant dbSNP:rs121918741
1433	G → R in DRVT. Corresponds to variant dbSNP:rs121917908
1433	G → V in DRVT.
1434	W → R in DRVT. Corresponds to variant dbSNP:rs121918789
1437	I → M in DRVT.

Position(s) in SEQ ID NO: 19	Description
1441	A → P in DRVT. Corresponds to variant dbSNP:rs121917974
1450	Q → K in DRVT. Corresponds to variant dbSNP:rs121918806
1450	Q → R in DRVT. Corresponds to variant dbSNP:rs121918790
1451	P → L in DRVT. Corresponds to variant dbSNP:rs121917945
1451	P → S in DRVT.
1453	Y → C in DRVT; borderline phenotype.
1454	E → K in DRVT. Corresponds to variant dbSNP:rs1553522472
1461	L → I in DRVT. Corresponds to variant dbSNP:rs121918772
1462	Y → C in DRVT. Corresponds to variant dbSNP:rs121917962
1462	Y → H in DRVT and ICEGTC; borderline DRVT phenotype. Corresponds to variant dbSNP:rs1559119345
1463	F → S in DRVT. Corresponds to variant dbSNP:rs121917946
1470	G → W in DRVT. Corresponds to variant dbSNP:rs121917924
1472	F → S in DRVT; borderline phenotype.
1473	Missing in DRVT.
1475	L → S in DRVT. Corresponds to variant dbSNP:rs121917947
1476	N → K in DRVT; borderline phenotype.
1483	Missing in DRVT.
1484	D → G in DRVT.
1485	N → Y in DRVT.
1503	E → K in DRVT; borderline phenotype. 3 Publications
1503	Missing in DRVT; borderline phenotype.
1514	L → S in DRVT. Corresponds to variant dbSNP:rs121918764
1538	V → I in DRVT. Corresponds to variant dbSNP:rs780360360
1544	D → A in DRVT.
1544	D → G in DRVT.
1545	I → V in DRVT. Corresponds to variant dbSNP:rs121917975
1555	M → R in DRVT.
1559	Missing in DRVT.
1561	E → K in DRVT.
1579	V → E in DRVT.
1586	G → E in DRVT. Corresponds to variant dbSNP:rs121918742
1588	C → R in DRVT. Corresponds to variant dbSNP:rs121917919
1592	L → H in DRVT; borderline phenotype.
1592	L → P in DRVT.
1596	R → C in DRVT; also found in a patient with cryptogenic focal epilepsy. Corresponds to variant dbSNP:rs121917993
1596	R → L in DRVT.
1605	N → I in DRVT. Corresponds to variant dbSNP:rs1057519533
1605	N → S in DRVT; borderline phenotype.
1608	D → G in DRVT.
1608	D → Y in DRVT. Corresponds to variant dbSNP:rs121917915
1612	V → I in DRVT. Corresponds to variant dbSNP:rs121918808
1630	V → L in DRVT; borderline phenotype in some patients. Corresponds to variant dbSNP:rs121917914
1630	V → M in DRVT. Corresponds to variant dbSNP:rs121917914
1637	V → E in DRVT; also found in a child with febrile status epilepticus who developed liver failure. Corresponds to variant dbSNP:rs121918810

Position(s) in SEQ ID NO: 19	Description
1638	I → N in DRVT. Corresponds to variant dbSNP:rs1057521079
1638	I → T in DRVT; also found in a patient with an unclassified form of epilepsy.
1639	R → G in DRVT.
1642	R → S in DRVT. Corresponds to variant dbSNP:rs1131691581
1645	R → Q in DRVT. Corresponds to variant dbSNP:rs121917976
1648	R → C in DRVT. Corresponds to variant dbSNP:rs121918791
1648	R → H in GEFS+2 and DRVT. Corresponds to variant dbSNP:rs121918622
1653	A → E in DRVT; borderline phenotype.
1658	T → M in DRVT. Corresponds to variant dbSNP:rs121917922
1658	T → R in DRVT. Corresponds to variant dbSNP:rs121917922
1660	L → P in DRVT. Corresponds to variant dbSNP:rs1131691675
1661	F → S in DRVT. Corresponds to variant dbSNP:rs121918797
1662	A → V in DRVT; borderline phenotype. Corresponds to variant dbSNP:rs794726839
1664	M → K in DRVT. Corresponds to variant dbSNP:rs121918765
1667	L → P in DRVT. Corresponds to variant dbSNP:rs1131691774
1668	P → A in DRVT. Corresponds to variant dbSNP:rs121917948
1668	P → L in DRVT; borderline phenotype.
1672	N → I in DRVT.
1673	I → T in DRVT.
1674	G → R in DRVT. Corresponds to variant dbSNP:rs121918792
1675	L → R in DRVT.
1677	L → F in DRVT.
1683	I → T in DRVT; borderline phenotype.
1684	Y → D in DRVT.
1685	A → D in DRVT. Corresponds to variant dbSNP:rs121918744
1688	G → W in DRVT.
1692	F → S in DRVT. Corresponds to variant dbSNP:rs121918778
1694	Y → C in DRVT. Corresponds to variant dbSNP:rs121918777
1707	F → V in DRVT. Corresponds to variant dbSNP:rs121917977
1713	S → N in DRVT. Corresponds to variant dbSNP:rs121918816
1714	M → K in DRVT.
1714	M → R in DRVT. Corresponds to variant dbSNP:rs121917949
1716	C → R in DRVT. Corresponds to variant dbSNP:rs121917926
1721	T → R in DRVT. Corresponds to variant dbSNP:rs121917978
1725	G → C in DRVT.
1726	W → R in DRVT. Corresponds to variant dbSNP:rs121917979
1727	D → G in DRVT.
1741	C → R in DRVT.
1749	G → E in DRVT. Corresponds to variant dbSNP:rs121918798
1756	C → G in DRVT. Corresponds to variant dbSNP:rs121918809
1762	G → E in DRVT. Corresponds to variant dbSNP:rs121917950
1763	I → N in DRVT.
1766	Missing in DRVT.
1770	I → F in DRVT.
1770	I → N in DRVT.

Position(s) in SEQ ID NO: 19	Description
1770	I → T in DRVT; borderline phenotype.
1771	I → F in DRVT; borderline phenotype; also found in a patient with focal epilepsy.
1771	I → N in DRVT.
1773	S → F in DRVT. Corresponds to variant dbSNP:rs121917951
1780	M → T in DRVT. Corresponds to variant dbSNP:rs121917952
1781	Y → C in DRVT and ICEGTC. Corresponds to variant dbSNP:rs121918779
1781	Y → H in DRVT.
1782	I → M in DRVT. Corresponds to variant dbSNP:rs121918763
1782	I → S in DRVT.
1783	A → T in DRVT. Corresponds to variant dbSNP:rs121917980
1783	A → V in DRVT. Corresponds to variant dbSNP:rs121917921
1787	E → K in DRVT. Corresponds to variant dbSNP:rs121917916
1788	N → K in DRVT; unknown pathological significance. Corresponds to variant dbSNP:rs1381184010
1792	A → T in DRVT.
1807 – 1810	Missing in DRVT.
1808	F → I in DRVT.
1812 – 1815	WEKF → C in DRVT.
1812	W → G in DRVT. Corresponds to variant dbSNP:rs121918751
1812	W → S in DRVT.
1813 – 1815	Missing in DRVT.
1831	F → S in DRVT. Corresponds to variant dbSNP:rs121918748
1832	A → P in DRVT.
1835	L → F in DRVT.
1852	M → K in DRVT.
1855	P → L in DRVT. Corresponds to variant dbSNP:rs1057517958
1880	G → E in DRVT. Corresponds to variant dbSNP:rs201905405
1881	E → D in DRVT. Corresponds to variant dbSNP:rs121918804
1909	T → I in DRVT; functional channel displaying decreased peak current densities but increased persistent current. Corresponds to variant dbSNP:rs121918793
1909	Missing in DRVT. Corresponds to variant dbSNP:rs1553519902
1922	I → T in DRVT. Corresponds to variant dbSNP:rs121917981
1927	R → IIQ in DRVT.

MECP2

MECP2 (NCBI Gene ID: 4204; RefSeq NG_007107.3) encodes a methyl-CpG binding protein 2. Its transcripts and protein isoforms are shown in Table 5.

Table 5. MECP2 transcripts and isoforms

NCBI transcript accession	Transcript #	SEQ ID NO:	NCBI protein accession	Isoform #	SEQ ID NO:
NM_001386139.1	10	SEQ ID NO: 25	NP_001373068.1	4	SEQ ID NO: 39

NCBI transcript accession	Transcript #	SEQ ID NO:	NCBI protein accession	Isoform #	SEQ ID NO:
NM_001110792.2	2	SEQ ID NO: 26	NP_001104262.1	2	SEQ ID NO: 37
NM_001369393.2	6	SEQ ID NO: 27	NP_001356322.1	3	SEQ ID NO: 38
NM_001386138.1	9	SEQ ID NO: 28	NP_001373067.1	4	SEQ ID NO: 39
NM_004992.4	1	SEQ ID NO: 29	NP_004983.1	1	SEQ ID NO: 40
NM_001386137.1	8	SEQ ID NO: 30	NP_001373066.1	4	SEQ ID NO: 39
NM_001369392.2	5	SEQ ID NO: 31	NP_001356321.1	3	SEQ ID NO: 38
NM_001369391.2	4	SEQ ID NO: 32	NP_001356320.1	3	SEQ ID NO: 38
NM_001316337.2	3	SEQ ID NO: 33	NP_001303266.1	3	SEQ ID NO: 38
NM_001369394.2	7	SEQ ID NO: 34	NP_001356323.1	3	SEQ ID NO: 38
XM_024452383.1	X1	SEQ ID NO: 35	XP_024308151.1	X1	SEQ ID NO: 41
XM_011531166.2	X3	SEQ ID NO: 36	XP_011529468.1	X1	

Mutations in MECP2 associated with Rett disease are shown in Table 6.

Table 6. Mutations in MECP2 associated with Rett disease.

Position(s) in SEQ ID NO: 40	Description
10	E → Q in RTT. Corresponds to variant dbSNP:rs61754421 .
97	D → E in RTT. Corresponds to variant dbSNP:rs61754449 .
97	D → Y in RTT. Corresponds to variant dbSNP:rs61754448 .
100	L → R in RTT. Corresponds to variant dbSNP:rs61754451 .
100	L → V in RTT. Corresponds to variant dbSNP:rs28935168 .
101	P → H in RTT. Corresponds to variant dbSNP:rs61754453 .
101	P → L in RTT. Corresponds to variant dbSNP:rs61754453 .
101	P → R in RTT; also in a patient with Angelman syndrome and some typical RTT features. Corresponds to variant dbSNP:rs61754453 .
101	P → S in RTT. Corresponds to variant dbSNP:rs61754452 .
101	P → T in RTT.
106	R → Q in RTT. Corresponds to variant dbSNP:rs61754457 .
106	R → W in RTT. 1 Corresponds to variant dbSNP:rs28934907 .
111	R → G in RTT. Corresponds to variant dbSNP:rs61754459 .
120	Y → D in RTT. Corresponds to variant dbSNP:rs267608454 .
124	L → F in RTT. Corresponds to variant dbSNP:rs61755763 .
128	Q → P in RTT. Corresponds to variant dbSNP:rs61748383 .

133	R → C in RTT; impairs interaction with ATRX and abolishes ATRX recruitment to heterochromatin. Corresponds to variant dbSNP:rs28934904 .
133	R → H in RTT. Corresponds to variant dbSNP:rs61748389 .
134	S → C in RTT. Corresponds to variant dbSNP:rs61748390 .
135	K → E in RTT. Corresponds to variant dbSNP:rs61748391 .
152	P → R in RTT. Corresponds to variant dbSNP:rs61748404 .
155	F → I in RTT. Corresponds to variant dbSNP:rs61748406 .
155	F → S in RTT. Corresponds to variant dbSNP:rs28934905 .
156	D → G in RTT. Corresponds to variant dbSNP:rs61748407 .
158	T → A in RTT. Corresponds to variant dbSNP:rs61748411 .
158	T → M in RTT. Corresponds to variant dbSNP:rs28934906 .
161	G → V in RTT. Corresponds to variant dbSNP:rs61748417 .
210	K → I in RTT. Corresponds to variant dbSNP:rs61749730 .
225	P → R in RTT. Corresponds to variant dbSNP:rs61749715 .
270 – 486	Missing in RTT. Add BLAST
302	P → A in RTT. Corresponds to variant dbSNP:rs61751373 .
302	P → H in RTT. Corresponds to variant dbSNP:rs61749723 .
302	P → L in RTT. Corresponds to variant dbSNP:rs61749723 .
302	P → R in RTT. Corresponds to variant dbSNP:rs61749723 .
305	K → R in RTT; abolishes interaction with TBL1X. Corresponds to variant dbSNP:rs61751441 .
306	R → C in RTT; abolishes interaction with TBL1X and TBL1XR1. Corresponds to variant dbSNP:rs28935468 .
306	R → H in RTT. Corresponds to variant dbSNP:rs61751443 .
322	P → A in RTT. Corresponds to variant dbSNP:rs61751449 .
322	P → L in RTT. Corresponds to variant dbSNP:rs61751450 .
344	R → W in RTT. Corresponds to variant dbSNP:rs61752361 .
388	P → S in RTT; unknown pathological significance. Corresponds to variant dbSNP:rs61753000 .

FUSION PROTEINS

Described herein are non-naturally occurring or engineered fusion proteins comprising an RNA effector protein and a regulation moiety, as well as variants and mutants thereof.

In some cases, the RNA effector protein amino acid sequence is N-terminal to the regulation moiety amino acid sequence. In some cases, the RNA effector protein amino acid sequence is C-terminal to the regulation protein amino acid sequence. In some cases, the RNA effector protein amino acid sequence is inserted within the regulation protein amino acid sequence. In some cases, the regulation protein amino acid sequence is inserted within the RNA effector protein amino acid sequence.

In some cases, the fusion protein comprises a linker and/or a spacer between the RNA effector protein and the regulation moiety.

In some cases, the fusion protein further comprises a nuclear export signal. In some cases, the nuclear export signal is from the HIV Rev protein (LPPLERLTL, SEQ ID NO: 51). In

some cases, the nuclear export signal is between the RNA effector protein amino acid sequence and the regulation moiety amino acid sequence.

In some cases, the RNA effector protein is dCas13b and the regulon moiety is PABPC1. Thus in some cases, the fusion protein comprises or consists of SEQ ID NO: 48. In some cases, the fusion protein comprises or consists of SEQ ID NO: 49.

In some cases, the RNA effector protein is dCas13b and the regulon moiety is NAT10.

In some cases, the RNA effector protein and/or regulation moiety is a functionally active portion of an effector protein and/or regulation moiety. Therefore, for example, if the RNA effector protein is C terminal to the regulation moiety, one or more N terminal amino acids of the regulation moiety may be substituted or deleted (e.g., the N terminal methionine). Likewise, if the regulation moiety is C terminal to the RNA effector protein, one or more N terminal amino acids of the RNA effector moiety may be substituted or deleted (e.g., the N terminal methionine).

Also described herein are polynucleotide sequence(s) encoding the fusion protein(s) described herein, vectors comprising the polynucleotide sequence(s), and cells comprising and/or expressing the vector(s).

In some cases, the nucleic acid sequence(s) are codon optimized.

In some cases, the vector(s) comprise a promoter sequence that drives the expression of the fusion protein(s) and/or gRNA(s).

RNA Effector Proteins

The fusion proteins described herein comprise RNA effector proteins. In some cases, the RNA effector protein is a Cas effector protein or variant or mutant thereof.

In some cases, the RNA effector protein is a Cas9 effector protein (e.g., SEQ ID NO: 46) or a mutant or variant thereof. In some cases, the RNA effector protein is a catalytically inactive Cas9 effector protein, e.g., a Cas9 effector protein with eliminated cleavage activity (e.g., dCas9; e.g., SEQ ID NO: 46 with mutations D10A and H840A).

In some cases, the RNA effector protein is a Cas12 effector protein, e.g., Cas12a (Cpf1), e.g., LbCas12a or a mutant or variant thereof, e.g., a mutant or variant with eliminated cleavage activity. In some cases, the RNA effector protein is a Cas12b effector protein, e.g., AapCas12b or AacCas12b, e.g., a mutant or variant with eliminated cleavage activity.

In some cases, the RNA effector protein is Cas14 effector protein or a mutant or variant thereof. *See, e.g.,* Harrington et al., "Programmed DNA Destruction by Miniature CRISPR-Cas13 Enzymes," *Science* 362(6146):839–42 (2018); *see also* Karvelis et al., "PAM Recognition by Miniature CRISPR-Cas12f Nucleases Triggers Programmable Double-Stranded DNA Target Cleavage," *Nucleic Acids Res* 48(9):5016–23 (2020).

Type VI CRISPR-Cas systems contain the programmable single-effector RNA-guided RNases Cas13. *See, e.g.,* Cox et al., "RNA Editing with CRISPR-Cas13," *Science* 358(6366):1019–27 (2017). The Cas13 family contains at least four known subtypes, including Cas13a (formerly C2c2), Cas13b, Cas13c, and Cas13d.

In some cases, the RNA effector protein is a Cas13 effector protein, e.g., Cas13a, Cas13b, Cas13c, or Cas13d. In some cases, the RNA effector protein is a Cas13b effector protein. *See, e.g.,* Smargon et al. (2017),"Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28," *Molecular Cell* 65, 618-630 (Feb. 16, 2017); *see also* Smargon et al., "RNA-Targeting CRISPR Systems from Metagenomic Discovery to Transcriptome Engineering," *Nat Cell Biol* 22(2):143–50 (2020).

In some cases, the RNA effector protein is a catalytically inactive RNA-effector protein, e.g., a Cas13 effector protein with eliminated cleavage activity (dCas13), e.g., dCas13b (SEQ ID NO: 47)).

In some cases, the RNA effector protein has at least 80%, e.g., at least 85%, 90%, 95%, 98%, or 100% sequence identity compared to SEQ ID NO: 47.

In some cases, the Cas13 effector protein is a Cas13bt. *See, e.g.,* Kannan et al., "Compact RNA Editors with Small Cas13 Proteins," *Nature Biotechnology* 18:499–560 (2021). In some cases, the Cas13 effector protein is a catalytically inactive Cas13bt effector protein (dCas13bt). In some cases, the Cas13 effector protein is a Cas13bt with mutations corresponding to H133A and H1058 of the dCas13bt.

In some cases, the Cas13 effector protein is a Cas13 effector protein described in WO2018170333A1 (e.g., in Table 1A of WO2018170333A1). In some cases, the Cas13 is a catalytically inactive mutant of a Cas13 described in WO2018170333A1 (e.g., a mutant of a Cas13 shown in Table 1A of WO2018170333A1).

Regulon Moieties

The fusion proteins described herein comprise regulation moieties. In some cases, the regulation moiety that stimulates and/or stabilizes an mRNA, e.g., an mRNA of an active allele of a gene associated with a haplotype insufficiency disorder. In some cases, the regulation moiety stimulates activity of the mRNA. In some cases, the regulation moiety stabilizes the mRNA.

In some cases, the regulation moiety can be a translational stimulator protein. In some cases, the regulation moiety can be a wild-type protein. In some cases, the regulation moiety can be a truncated variant of a wild-type protein.

PABPC1

In some cases, the regulation moiety is PABPC1. Polyadenylate-binding protein 1 PABPC1 is a highly conserved RNA binding protein in eukaryotes. This protein has four N-terminal RNA recognition motif (RRM) domains, which bind poly(A) RNA with a nanomolar affinity [23, 24]. The RRMs are followed by a proline-rich linker and a C-terminal MLLE domain. The MLLE domain recognizes a peptide motif called poly(A)-interacting motif 2 (PAM2), which is found in a number of PABPC partner proteins that regulate mRNA metabolism (stability and translation). The presence of PABPC1 on mRNA is known to stimulate their activity, enhancing translation and mRNA stability [19].

Thus, in some cases, the regulation moiety comprises or consists of SEQ ID NO: 42 or SEQ ID NO: 43.

In some cases, the regulation moiety comprises or consists of polypeptide sequence having at least 80%, e.g., at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to SEQ ID NO: 42 or SEQ ID NO: 43.

In some cases, the regulation moiety comprises a wild-type PABPC1 protein. In some cases, the regulation moiety comprises a truncated variant of the wild-type PABPC1 protein. In some embodiments, the regulation moiety comprises the RRM domains and the MLLE domains. In some case, the regulation moiety comprises the MLLE domain.

NAT10

In some cases, the regulation moiety is NAT10. RNA cytidine acetyltransferase NAT10 is a highly conserved enzyme that catalyzes the conversion of cytidine to N4-acetylcytindine (ac4C) [25]. Generation of the "epitranscriptome" through post- transcriptional ribonucleoside modification embeds a layer of regulatory complexity into RNA structure and function. N4-acetylcytidine mRNA modification has been established to have robust stimulatory effect on mRNA stability and translation in human cells [25].

Thus, in some cases, the regulation moiety comprises or consists of SEQ ID NO: 44 or SEQ ID NO: 45. In some cases, the regulation moiety comprises or consists of a polypeptide sequence having at least 80%, e.g., at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to SEQ ID NO: 44 or SEQ ID NO: 45.

PHARMACEUTICAL COMPOSITIONS AND METHODS OF TREATMENT

Pharmaceutical Compositions

Also described herein are pharmaceutical compositions comprising a fusion protein described herein, pharmaceutical compositions comprising a nucleic acid encoding the fusion

proteins described herein, and pharmaceutical compositions comprising a vector comprising a nucleic acid encoding a fusion protein described herein.

In some cases, the pharmaceutical composition further comprises a gRNA, e.g., as described herein. In some cases, the pharmaceutical composition further comprises a nucleic acid encoding a gRNA, e.g., as described herein. In some cases, the pharmaceutical composition further comprises a vector comprising a nucleic acid encoding a gRNA, e.g., as described herein.

In some cases, the pharmaceutical composition is formulated for gene therapy, e.g., as described herein.

Fusion Proteins

In some cases, the pharmaceutical compositions described herein comprise a fusion protein, e.g., a fusion protein described herein. In some cases, the pharmaceutical compositions described herein comprise a nucleic acid encoding a fusion protein, e.g., a fusion protein described herein. In some cases, the pharmaceutical composition comprises the nucleic acid encoding a fusion protein, e.g., a fusion protein described herein.

gRNAs

In some cases, e.g., when the fusion protein comprises a Cas effector protein, e.g., as described herein, the pharmaceutical composition further comprises a guide RNA (gRNA) comprising a complementarity region designed to complex with the Cas protein and hybridize to a nucleic acid, e.g., an mRNA of an active allele of a gene associated with a haploinsufficiency disorder, e.g., the 3' UTR of an mRNA of an active allele of a gene associated with a haploinsufficiency disorder.

In some cases, the gRNA is designed to hybridize to the 3' UTR of an mRNA. In some cases, the gRNA can hybridize to a 3' end of the 3' UTR transcript. In some cases, the gRNA can hybridize to a position within the 3' UTR transcript.

In some cases, the gRNA comprises a CRISPR RNA (crRNA). In some cases, the gRNA comprises a *trans*-activating CRISPR RNA (tracrRNA). In some cases, the gRNA is a single guide RNA (sgRNA). In some cases, the gRNA does not comprise a tracrRNA.

In some cases, the gRNA is designed to complex with Cas13b and comprises a complementary region designed to hybridize to an mRNA of an active allele of a gene associated with a haploinsufficiency disorder, e.g., the 3' UTR of an mRNA of an active allele of a gene associated with a haploinsufficiency disorder.

In some cases, the Cas13b is dCas13b (SEQ ID NO: 47).

In some cases, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16 , SEQ ID NO: 17, SEQ ID NO: 18, and combinations thereof. In some cases, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 17, and combinations thereof.

In some cases, the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding an amino acid selected from the group consisting of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, and combinations thereof. In some cases, the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding SEQ ID NO: 21.

In some cases, the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, and combinations thereof. In some cases, the gRNA comprises a complementarity region designed to hybridize to SEQ ID NO: 26.

In some cases, the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, and combinations thereof. In some cases, the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding SEQ ID NO: 37.

In some cases, the gRNA targets an mRNA encoding Mecp2. In some cases, the gRNA targets an mRNA encoding SynGAP. In some cases, the gRNA targets an mRNA encoding SHANK3. In some cases, the gRNA targets an mRNA encoding PTEN. In some cases, the gRNA targets an mRNA encoding CHD2. In some cases, the gRNA comprises or consists of SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 58, SEQ ID NO: 59, SEQ ID NO: 60, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67, SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 70, SEQ ID NO: 71, or SEQ ID NO: 72.

A list of gRNAs to target the 3' UTR of genes of interest is shown in Table 7.

Table 7. List of guide RNAs to target the 3'UTR of the genes of interest.

Name	Target Gene	Sequence	SEQ ID NO:
oJC5003	pJC889-gRNA1	CACCGCACTATAGAATAGGGCCCTCTAG	55

oJC5005	pJC889-gRNA2	CACCGGAGCTCTAGCATTTAGGTGACAC	56
oJC5051	Mecp2-G1	CACCGAAACACTTAGAGTTTCGGAGCTT	57
oJC5053	Mecp2-G2	CACCGTGCCATTCAAGAAGAACTCCGAG	58
oJC5055	Mecp2-G3	CACCGATTACACCATAGTAAGGAAGCCAGG	59
oJC5057	Mecp2-G4	CACCGAAATTCACTGAAAGAGGCCCCCA	60
oJC5059	Mecp2-G5	CACCGACAGAAGATTGTCTGCACCCGCA	61
oJC5061	Mecp2-G6	CACCGGCTAAAAATGTATATGCCCAAAG	62
oJC5063	Mecp2-G7	CACCGCCAAAAGACAGGTTCAGAAGCCA	63
oJC5241	SynGAP-G1	CACCGCACCGAAGTTGAAAGTTTGGAGGTGCCA	64
oJC5243	SynGAP-G2	CACCGCACCGACCCCAGATAATTACAACAGCCA	65
oJC5556	SHANK3-G1	CACCGCACACACATAAAGGGCAGGACAG	66
oJC5558	SHANK3-G2	CACCGAAACCCAAATTGAACGGAACCAA	67
oJC5560	PTEN-G1	CACCGAAGACAACTGATAAGCCAACAGC	68
oJC5562	PTEN-G2	CACCGAGCCAGTAAATGATAGGCTACAG	69
oJC5564	CHD2-G1	CACCGGAATAAAACTGAAGGCCCCAGCA	70
oJC5566	CHD2-G2	CACCGCAGTAGAAACTTGTTCACCACAG	71
oJC5568	CHD2-G3	CACCGTACAAGAATCAAAAGTACAGCAG	72

Also provided herein are polynucleotide sequences encoding the gRNA(s) described herein, vectors comprising the polynucleotide sequence(s) encoding the gRNA(s) described herein, and cells comprising the vector(s) encoding the gRNA(s) described herein.

In some cases, the polynucleotide encoding the fusion protein and the polynucleotide encoding the gRNA are on the same vector. In some cases, the polynucleotide encoding the fusion protein and the polynucleotide encoding the gRNA are on different vectors.

Gene Therapy

The nucleic acids described herein, e.g., nucleic acids encoding fusion protein(s) and/or gRNA(s) described herein, can be incorporated into a gene construct to be used as a part of a gene therapy protocol. Thus, also provided herein are targeted expression vectors for *in vivo* transfection and expression of a polynucleotide that encode fusion protein(s) and/or gRNA(s) described herein. Expression constructs of such components can be administered in any effective carrier, e.g., any formulation or composition capable of effectively delivering the component gene to cells in vivo. Approaches include insertion of the gene in viral vectors, including recombinant retroviruses, adenovirus, adeno-associated virus, lentivirus, and herpes simplex virus-1, or recombinant bacterial or eukaryotic plasmids. Viral vectors transfect cells directly; plasmid DNA can be delivered naked or with the help of, for example, cationic liposomes (lipofectamine) or derivatized (e.g., antibody conjugated), polylysine conjugates, gramacidin S,

artificial viral envelopes or other such intracellular carriers, as well as direct injection of the gene construct or $CaPO_4$ precipitation carried out in vivo.

A preferred approach for in vivo introduction of nucleic acid into a cell is by use of a viral vector containing nucleic acid, e.g., a cDNA. Infection of cells with a viral vector has the advantage that a large proportion of the targeted cells can receive the nucleic acid. Additionally, molecules encoded within the viral vector, e.g., by a cDNA contained in the viral vector, are expressed efficiently in cells that have taken up viral vector nucleic acid.

Retrovirus vectors and adeno-associated virus vectors can be used as a recombinant gene delivery system for the transfer of exogenous genes *in vivo*, particularly into humans. These vectors provide efficient delivery of genes into cells, and the transferred nucleic acids are stably integrated into the chromosomal DNA of the host. The development of specialized cell lines (termed "packaging cells") which produce only replication-defective retroviruses has increased the utility of retroviruses for gene therapy, and defective retroviruses are characterized for use in gene transfer for gene therapy purposes (for a review see Miller, *Blood* 76:271 (1990)). A replication defective retrovirus can be packaged into virions, which can be used to infect a target cell through the use of a helper virus by standard techniques. Protocols for producing recombinant retroviruses and for infecting cells in vitro or in vivo with such viruses can be found in Ausubel, et al., eds., Current Protocols in Molecular Biology, Greene Publishing Associates, (1989), Sections 9.10-9.14, and other standard laboratory manuals. Examples of suitable retroviruses include pLJ, pZIP, pWE and pEM which are known to those skilled in the art. Examples of suitable packaging virus lines for preparing both ecotropic and amphotropic retroviral systems include ΨCrip, ΨCre, Ψ2 and ΨAm. Retroviruses have been used to introduce a variety of genes into many different cell types, including epithelial cells, in vitro and/or in vivo (see for example Eglitis, et al. (1985) Science 230:1395-1398; Danos and Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:6460-6464; Wilson et al. (1988) Proc. Natl. Acad. Sci. USA 85:3014-3018; Armentano et al. (1990) Proc. Natl. Acad. Sci. USA 87:6141-6145; Huber et al. (1991) Proc. Natl. Acad. Sci. USA 88:8039-8043; Ferry et al. (1991) Proc. Natl. Acad. Sci. USA 88:8377-8381; Chowdhury et al. (1991) Science 254:1802-1805; van Beusechem et al. (1992) Proc. Natl. Acad. Sci. USA 89:7640-7644; Kay et al. (1992) Human Gene Therapy 3:641-647; Dai et al. (1992) Proc. Natl. Acad. Sci. USA 89:10892-10895; Hwu et al. (1993) J. Immunol. 150:4104-4115; U.S. Patent No. 4,868,116; U.S. Patent No. 4,980,286; PCT Application WO 89/07136; PCT Application WO 89/02468; PCT Application WO 89/05345; and PCT Application WO 92/07573).

Another viral gene delivery system useful in the present methods utilizes adenovirus-derived vectors. The genome of an adenovirus can be manipulated, such that it encodes and

expresses a gene product of interest but is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. See, for example, Berkner et al., BioTechniques 6:616 (1988); Rosenfeld et al., Science 252:431-434 (1991); and Rosenfeld et al., Cell 68:143-155 (1992). Suitable adenoviral vectors derived from the adenovirus strain Ad type 5 dl324 or other strains of adenovirus (e.g., Ad2, Ad3, or Ad7 etc.) are known to those skilled in the art. Recombinant adenoviruses can be advantageous in certain circumstances, in that they are not capable of infecting non-dividing cells and can be used to infect a wide variety of cell types, including epithelial cells (Rosenfeld et al., (1992) supra). Furthermore, the virus particle is relatively stable and amenable to purification and concentration, and as above, can be modified so as to affect the spectrum of infectivity. Additionally, introduced adenoviral DNA (and foreign DNA contained therein) is not integrated into the genome of a host cell but remains episomal, thereby avoiding potential problems that can occur as a result of insertional mutagenesis in situ, where introduced DNA becomes integrated into the host genome (e.g., retroviral DNA). Moreover, the carrying capacity of the adenoviral genome for foreign DNA is large (up to 8 kilobases) relative to other gene delivery vectors (Berkner et al., supra; Haj-Ahmand and Graham, J. Virol. 57:267 (1986).

Yet another viral vector system useful for delivery of nucleic acids is the adeno-associated virus (AAV). Adeno-associated virus is a naturally occurring defective virus that requires another virus, such as an adenovirus or a herpes virus, as a helper virus for efficient replication and a productive life cycle. (For a review see Muzyczka et al., Curr. Topics in Micro. and Immunol.158:97-129 (1992). It is also one of the few viruses that may integrate its DNA into non-dividing cells, and exhibits a high frequency of stable integration (see for example Flotte et al., Am. J. Respir. Cell. Mol. Biol. 7:349-356 (1992); Samulski et al., J. Virol. 63:3822-3828 (1989); and McLaughlin et al., J. Virol. 62:1963-1973 (1989). Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate. Space for exogenous DNA is limited to about 4.5 kb. An AAV vector such as that described in Tratschin et al., Mol. Cell. Biol. 5:3251-3260 (1985) can be used to introduce DNA into cells. A variety of nucleic acids have been introduced into different cell types using AAV vectors (see for example Hermonat et al., Proc. Natl. Acad. Sci. USA 81:6466-6470 (1984); Tratschin et al., Mol. Cell. Biol. 4:2072-2081 (1985); Wondisford et al., Mol. Endocrinol. 2:32-39 (1988); Tratschin et al., J. Virol. 51:611-619 (1984); and Flotte et al., J. Biol. Chem. 268:3781-3790 (1993).

In addition to viral transfer methods, such as those illustrated above, non-viral methods can also be employed to cause expression of a nucleic acid compound described herein, e.g., nucleic acid compound(s) encoding the fusion protein(s) and/or gRNA(s) described herein. Typically non-viral methods of gene transfer rely on the normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In some embodiments, non-

viral gene delivery systems can rely on endocytic pathways for the uptake of the subject gene by the targeted cell. Exemplary gene delivery systems of this type include liposomal derived systems, poly-lysine conjugates, lipid nanoparticles and artificial viral envelopes. Other embodiments include plasmid injection systems such as are described in Meuli et al., J. Invest. Dermatol. 116(1):131-135 (2001); Cohen et al., Gene Ther. 7(22):1896-905 (2000); or Tam et al., Gene Ther. 7(21):1867-74 (2000).

In some embodiments, nucleic acid compound described herein, e.g., nucleic acid compound(s) encoding the fusion protein(s) and/or gRNA(s), is entrapped in liposomes bearing positive charges on their surface (e.g., lipofectins), which can be tagged with antibodies against cell surface antigens of the target tissue (Mizuno et al., No Shinkei Geka 20:547-551 (1992); PCT publication WO91/06309; Japanese patent application 1047381; and European patent publication EP-A-43075).

In clinical settings, the gene delivery systems for the therapeutic gene can be introduced into a subject by any of a number of methods, each of which is familiar in the art. For instance, a pharmaceutical preparation of the gene delivery system can be introduced systemically, e.g., by intravenous injection, and specific transduction of the protein in the target cells will occur predominantly from specificity of transfection, provided by the gene delivery vehicle, cell-type or tissue-type expression due to the transcriptional regulatory sequences controlling expression of the receptor gene, or a combination thereof. In other embodiments, initial delivery of the recombinant gene is more limited, with introduction into the subject being quite localized. For example, the gene delivery vehicle can be introduced by catheter (see U.S. Patent 5,328,470) or by stereotactic injection (e.g., Chen et al., PNAS USA 91: 3054-3057 (1994)).

The pharmaceutical preparation of the gene therapy construct can consist essentially of the gene delivery system in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is embedded. Alternatively, where the complete gene delivery system can be produced intact from recombinant cells, e.g., retroviral vectors, the pharmaceutical preparation can comprise one or more cells, which produce the gene delivery system.

Methods of Treatment

The methods described herein include methods for the treatment of disorders associated with haploinsufficiency, e.g., as described herein. In Generally, the methods include administering a therapeutically effective amount of a pharmaceutical composition as described herein, to a subject who is in need of, or who has been determined to be in need of, such treatment, e.g., but a gene therapy method described herein.

In some cases, the methods of treatment provided herein may be used to treat a subject (e.g., human, monkey, dog, cat, mouse) who has been diagnosed with or is suspected of having a haploinsufficiency disorder, e.g., as described herein. In some embodiments, the subject is a mammal. In some embodiments, the subject is a human.

In some cases, the subject may be a human who exhibits one or more symptoms associated with a haploinsufficiency disorder, e.g., as described herein. Any of the methods of treatment provided herein may be used to treat haploinsufficiency disorders at various stages.

In some cases, the disorder is Dravet Syndrome. In some cases, the subject has a mutation selected from the list in Table 4.

In some cases, the disorder is Rett Syndrome. In some cases, the subject has a mutation selected from the list in Table 6.

As used in this context, to "treat" means to ameliorate at least one symptom of the disorder associated with a haploinsufficiency disorder. Often, a haploinsufficiency disorder results in the amount of gene product expressed from mRNA(s) of the active allele being insufficient for proper gene function; thus, a treatment can result in an increase in the amount of gene product expressed from mRNA(s) of an active allele as compared to, e.g., an untreated age-matched subject.

Dosage

An "effective amount" is an amount sufficient to effect beneficial or desired results. For example, a therapeutic amount is one that achieves the desired therapeutic effect. This amount can be the same or different from a prophylactically effective amount, which is an amount necessary to prevent onset of disease or disease symptoms. An effective amount can be administered in one or more administrations, applications or dosages. A therapeutically effective amount of a therapeutic compound (i.e., an effective dosage) depends on the therapeutic compounds selected. The compositions can be administered one from one or more times per day to one or more times per week; including once every other day. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of the therapeutic compounds described herein can include a single treatment or a series of treatments.

Dosage, toxicity and therapeutic efficacy of the therapeutic compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose

therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit high therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

VARIANTS

In some cases, the fusion protein(s) described herein, or the nucleic acid sequences encoding them, have at least 80%, e.g., at least 85%, 90%, 95%, 98%, or 100% identity to the amino acid sequence of a sequence provided herein, e.g., has differences at up to 1%, 2%, 5%, 10%, 15%, or 20% of the residues of the sequence provided herein replaced, e.g., with conservative mutations, e.g., including or in addition to the mutations described herein. In preferred embodiments, the variant retains desired activity of the parent, e.g., the binding activity of the RNA-effector protein and the regulation activity of the regulon moiety.

To determine the percent identity of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). The length of a reference sequence aligned for comparison purposes is at least 80% of the length of the reference sequence, and in some embodiments is at least 90% or 100%. The nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein nucleic acid “identity” is equivalent to

nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

Percent identity between a subject polypeptide or nucleic acid sequence (i.e. a query) and a second polypeptide or nucleic acid sequence (i.e. target) is determined in various ways that are within the skill in the art, for instance, using publicly available computer software such as Smith Waterman Alignment (Smith, T. F. and M. S. Waterman (1981) J Mol Biol 147:195-7); "BestFit" (Smith and Waterman, Advances in Applied Mathematics, 482-489 (1981)) as incorporated into GeneMatcher Plus™, Schwarz and Dayhof (1979) Atlas of Protein Sequence and Structure, Dayhof, M.O., Ed, pp 353-358; BLAST program (Basic Local Alignment Search Tool; (Altschul, S. F., W. Gish, et al. (1990) J Mol Biol 215: 403-10), BLAST-2, BLAST-P, BLAST-N, BLAST-X, WU-BLAST-2, ALIGN, ALIGN-2, CLUSTAL, or Megalign (DNASTAR) software. In addition, those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the length of the sequences being compared. In general, for target proteins or nucleic acids, the length of comparison can be any length, up to and including full length of the target (e.g., 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 100%). For the purposes of the present disclosure, percent identity is relative to the full length of the query sequence.

For purposes of the present disclosure, the comparison of sequences and determination of percent identity between two sequences can be accomplished using a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

EXAMPLES

The invention is further described in the following examples, which do not limit the scope of the invention described in the claims.

Example 1: PAB tethered targeting

FIG. 3A is a schematic showing the experimental approach applied for this assay. Briefly, PABPC1 was fused to dCas13b tethering protein and guided to 3'UTR of the Luciferase gene via designed guide RNAs (gRNA) to enhance its translation.

HEK-293 cells were co-transfected with pJC1211 expressing dCas13b plasmid (pJC1211; SEQ ID NO: 50; FIG. 4; FIGS. 5A–5AB) encoding the PABPC1 tethered dCas13b (SEQ ID NO: 49) ; a plasmid expressing the luc reporter (pJC 889; Addgene plasmid # 18964); and one of either pJC1212 (expressing gRNA1 (FIG. 6; FIGS. 7A–7G; SEQ ID NO: 52) pJC1313 (expressing gRNA1 (FIG. 8; FIGS. 9A–9F; SEQ ID NO: 53) or the empty vector Addgene plasmid 103854 (pC0043-PspCas13b crRNA backbone Addgene plasmid # 103854; FIG. 10; FIGS. 11A–11H; SEQ ID NO: 54).

Cells were lysed 48 hours after transfection for luciferase detection, and the Promega Luciferase Assay System was applied according to the manufacturer's protocol (Promega E1500). The results show roughly 2-3 fold stimulation over control using two guide RNAs targeting 3'UTR of the luciferase construct (FIG. 3B).

Example 2: *In vitro* and *in vivo* Posttranscriptional Regulation of mRNA with PAB1-dCas13b and NAT10-dCas13b Fusion Proteins

Various mRNA regulon therapies will be used as disease modifying therapies in the treatment of monogenic neurological disorders with haploinsufficiency such as Dravet Syndrome and Rett Syndrome. The proteins PABPC1 and NAT10 will each be tethered to an RNA effector protein, e.g., dCas13b and expressed *in vitro* and *in vivo* with a gRNA targeting an mRNA of an active allele of a gene associated with Dravet Syndrome (e.g., SCN1A) or a gRNA targeting an mRNA of an active allele of a gene associated with Rett Syndrome (e.g., MECP2).

Example 3: Enhancing mRNA Expression *In vitro*

Cas13b in vitro optimization

A Cas13b/PABPC1 construct was generated by cloning the PCR amplified human PABPC1 in pC0054-CMV-dPspCas13b-longlinker-ADAR2DD (Addgene 103870) (pJC1206) (Table 8). The PC1-MS2V5-PABPC1 plasmid (Addgene#65807) was used as a template for PABPC1 amplification (primers are listed in Table 9). Briefly, pJC1206 was site mutated at nucleotide 5606 to make a unique BamHI site; this construct is hereafter referred to as pJC1210. The ADAR2DD sequence was then removed from using BamHI+NotI and replaced with PCR amplified PABPC1. The generated construct is pJC1211, which includes the full human PABPC1 sequence (Table 8). pJC1246 was created by digesting pJC1211 with BamH1/Not1 and inserting PCR product amplified from oJC5001/oJC5240. pJC1280 was created by cutting pJC1211 with BamH1/Not1 and re-ligating after blunting the ends.

Multiple sgRNAs targeting the 3'UTR of the genes of interest were designed using the 'nygenome' online tool for the prediction dCas13b guide (Cas13design (nygenome.org) (Table

7). These guides were individually cloned into PC0043-Cas3b-crRNA backbone (Addgene# #103854)(pJC1208) at BbsI sites. The reaction mix, including forward and reverse oligos in 1XNEB buffer 3.1 was incubated for 5 and 10 minutes at 95 and 100 degrees, then cooled down in the room temperature for 2 hours. Prepared oligos ligated into pJC1208 using T4 DNA ligase (NEB) following the GreenGate protocol.

The RNA amplifier technology was tested in HEK293, HepG2, and SH-SY5Y (ATCC CRL 2266) cell lines. HEK293 and HepG2 cell lines were grown in Dulbecco's Modified Eagle's Medium (DMEM) with 10% FBS, and SH-SY5Y cell lines were grown in Eagle's Minimum Essential Medium (EMEM) with 10% FBS following ATCC guidelines.

Cells were transiently co-transfected with individual sgRNAs along with the dCas13b-PABPC1 fusion plasmid for 48 hours using Opti-MEM Reduced Serum Medium (Gibco) and GeneXPlus Transfection Reagent (ATCC ACS-4004). RNA was isolated using the direct-zol RNA Kit (Zymo Research) following the manufacturer's protocol. cDNA was synthesized using SuperScript III First-Strand Synthesis System (Invitrogen), and qPCR was performed in Applied Biosystems Real-Time PCR instrument using PowerUp TM SYBR Green master mix (Applied Biosystems) and designed primers (Table 9). The data were analyzed using the CT value compared to a no sgRNA transfection and normalized to ACTB as a housekeeping gene.

Protein was isolated using RIPA buffer, and the western blotting was performed using Mini-protean TGX 4-15% gels (BIO-RAD). The following antibodies were used for immunoblotting according to the manufacturer's suggested concentrations; anti-GAPDH (6C5) (Santa Cruz Biotechnology), anti-MeCP2 (D4F3) (Cell Signaling Technology), anti-CHD2 (cat#4170) (Cell Signaling Technology), anti-PTEN (cat#9552) (Cell Signaling Technology), anti-SynGAP (cat# 3200) (Cell Signaling Technology), and anti-pan-Shank, clone N23B/49 (Cat# MABN24) (Millipore).

Luciferase assay

HEK293 cells were transfected with different ratios of pJC889 (Luciferase-pcDNA3 Addgene #18964), pJc1211, and individually two distinct guide RNAs targeting 3'UTR of firefly luciferase transcripts. The optimal ratio (0.2:1:1.5) of plasmid was chosen for this experiment. For luminescence detection, cells were lysed in 100µL 1X Passive Lysis Buffer (Promega). The lysate was mixed with ONE-Glo EX Reagent (Promega) following the manufacture protocol, and Luminescence was measured using a Lumat LB9507 Luminometer (Berthold Technologies).

AlphaFold method

Structural predictions of fusion proteins and native Cas13b and PABP were generated using AlphaFold v2.0 as pulled from the github repository located at the URL github.com/deepmind/alphafold from commit '1d43aaff941c84dc56311076b58795797e49107b'(ref15). Both native and customized fusion FASTAs were processed according to the AlphaFold documentation using the provided Docker script with the following parameters: '--max_template_date=2020-05-14 --preset=reduced_dbs'. Relaxed predicted structures with the highest pLDDT scores were used for interpretation of the corresponding input FASTAs.

Quantifications and Statistical Analysis

All data are shown in figures as bar charts were quantified as mean ± standard error. Results were considered significant at $p < 0.05$ as noted throughout figure legends. All experiments, unless otherwise indicated, have been run based on the common practice at least for three biological repeats.

Table 8: List of backbone plasmids and generated construct

Plasmid number	Description
pJC1204	PC1-MS2V5-PABPC1 plasmid Addgene#65807
pJC1206	pC0054-CMV-dPspCas13b-longlinker-ADAR2DD Addgene 103870
pJC1208	PC0043-Cas3b-crRNA backbone Addgene# #103854
pJC1210	pJC1204 mutated at nucleotide 5606
pJC1211	pJC1210 digested to remove ADAR2DD and replace with human PABPC1
pJC1246	dCas13b-MLLE
pJC1280	dCas13b
pJC889	Luciferase-pcDNA3 Addgene #18964

Table 9: List of the primers used

	Primer	**Sequence**	**SEQ ID NO:**
oJC4996	Site mutation oligoA	TCAAGCTGCCCAAGGACCGCATCCACAGCGAGAAGTCCAAC	73
oJC4996	Site mutation oligoB	GTTGGACTTCTCGCTGTGGATGCGGTCCTTGGGCAGCTTGA	74
oJC5001	PABPC1 F	AATGGGATCCTGGCAAGCCCATC	75

oJC5001	PABPC1 R	CTAAAGGGAAGCGGCCGCTTAAA	76
oJC5065	MeCP2-R	CAGTTCCTGGAGCTTTGGGAGATT	77
oJC5065	MeCP2-F	GTGTATTTGATCAATCCCCAG GGAA	78
oJC5067	Actb-F	CACCATTGGCAATGAGCGGTTC	79
oJC5068	Actb-R	AGGTCTTTGCGGATGTCCACGT	80
oJC5240	MLL-F	GGGATATCGTTCGCAATCCTCAG	81
oJC5245	SynGAP-F	TCTAGGCGCTATAGAGCAAACAGG	82
oJC5246	SynGAP-R	CAGAGACCCTCAGCTTCCAG	83
oJC5546	SHANK3-F	AGGATCACACCCGCCGAGATTA	84
oJC5547	SHANK3-R	CTACAGACTTGGTCCGTGGAATC	85
oJC5548	PTEN-F	TGAGTTCCCTCAGCCGTTACCT	86
oJC5549	PTEN-R	GAGGTTTCCTCTGGTCCTGGTA	87
oJC5550	CHD2-F	CGAAAACAGGCACTGGACCACT	88
oJC5551	CHD2-R	GATGACGACTGTGTCCGCTGAA	89

mRNA expression was increased by tethering a known translational stimulator, PABPC1, to the 3'UTR of a target mRNA. Tethering was achieved by fusing PABPC1 to the RNA binding protein dCas13b and co-expressing a guide RNA (gRNA). The gRNA is critical in that it has anti-sense homology to specific mRNAs and a short hairpin required for dCas13b binding (FIG. 12A). It was demonstrated that this gRNA-targeted tethering enhances both reporter and endogenous mRNAs in a gRNA-dependent manner. First, using HEK293 cells the Tethered mRNA Amplifier was co-transfected alongside a luciferase reporter construct. An approximate 1.5 to 2-fold increase in reporter protein amount was seen when gRNAs directed against the 3'UTR of the luciferase reporter are present. No stimulation occurred when the dCas13b-PABPC1 fusion is expressed alone (FIG. 12B; Ctrl). Continuing in HEK293 cells, the translation of an endogenous mRNA, MeCP2 was stimulated. An approximate 1.5-fold stimulation of translation was seen using two distinct gRNAs directed against the endogenous MeCP2 transcript's 3'UTR (FIG. 12C). A mild increase of 15% in MeCP2 transcript steady-state levels was observed when in the presence of the mRNA Amplifier (FIG. 12D). These data suggest that the stimulatory role of the Tethered mRNA Amplifier is through both mRNA stability and mRNA translation - known roles for PABPC1 in regulating mRNA metabolism.

Example 4: Enhancing mRNA Expression in Multiple Cell Types

It was further shown that the Tethered mRNA Amplifier enhances mRNA expression in multiple cell types; a stimulatory effect on MeCP2 protein expression is seen in SH-SY5Y (a neuronal cell line) and HepG2 (a liver cell line; FIG. 12E). Finally, the effect of the Tethered

mRNA amplifier was shown to be tunable by moving the gRNA to distinct positions within the 3'UTR. In the case of MeCP2 the strongest stimulatory effect was observed as the gRNA was moved closer to the 3' end of the transcript (FIG. 12F).

Example 5: Enhancing mRNA Expression on Different mRNA Transcripts

Next, the Tethered mRNA Amplifier was tested on other transcripts associated with haploinsufficiency disorders. Using SH-SY5Y cells, a model for neurodegenerative disorders, a gRNA-dependent translational stimulation of SYNGAP115, SHANK316, CHD217, and PTEN18 mRNAs was observed (FIG. 13A-13D). The loss of function of one allele for each of these genes is associated with autism spectrum disorders. In all cases, the stimulatory effect seen was between 1.2 to 2.0-fold for protein expression with an approximately 15-20% increase in mRNA levels. These data demonstrate that the Tethered mRNA Amplifier can be used as a gene therapy candidate for haploinsufficiency and is portable across multiple transcripts of clinical relevance.

Example 6: Minimizing the Size of the Tethered mRNA Amplifier Fusion Protein

Finally, it is known that gene therapy vectors such as AAV have payload size limitations of approximately 4.5kb. To minimize the Tethered mRNA Amplifier (5.2kb), specific truncations of PABPC1 were made and their efficacy on MeCP2 expression was tested. PABPC1 contains four RNA-recognition motifs (RRM1-4) at its N-terminus followed by a linker and a Mademoiselle (MLLE) domain at the C-terminus (FIG. 14A). The RRM domains bind to poly(A) tails while the MLLE domain is known to regulate its stimulatory role in translation. Since PABPC1 was artificially and specifically tethered to mRNAs independent of PABPC1 poly(A)-binding capacity, it was reasoned the RRMs would be dispensable for the Tethered mRNA Amplifier's function. Therefore, just the MLLE domain was fused to dCas13b (3.2KB). As a first test, the putative folding pattern of this new fusion was analyzed with an in silico approach using AlphaFold v2.0. As seen in FIGs. 14B-14E, removal of the RRM domains had minimal impact on the folding of either the MLLE domain or dCas13b itself. In HEK293 cells it was observed that this minimal construct also stimulated MeCP2 mRNA expression in a gRNA-dependent fashion to a similar extent as the full-length construct (FIG. 14F). Thus, the Tethered mRNA Amplifier approach was shown to be adaptable for use in current clinically efficacious gene therapy vectors.

REFERENCES

1. Escayg, A., et al., *Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+ 2.* Nature genetics, 2000. 24(4): p. 343-345.

2. Claes, L., et al., *De novo mutations in the sodium-channel gene SCN1A cause severe myoclonic epilepsy of infancy.* The American Journal of Human Genetics, 2001. 68(6): p. 1327-1332.

3. Dravet, C., *The core Dravet syndrome phenotype.* Epilepsia, 2011. 52(s2): p. 3-9.

4. Genton, P., R. Velizarova, and C. Dravet, *Dravet syndrome: The long-term outcome.* Epilepsia, 2011. 52(s2): p. 44-49.

5. Dravet, C. and H. Oguni, *Chapter 65 - Dravet syndrome (severe myoclonic epilepsy in infancy)*, in *Handbook of Clinical Neurology*, O. Dulac, M. Lassonde, and H.B. Sarnat, Editors. 2013, Elsevier. p. 627-633.

6. Cooper, M.S., et al., *Mortality in Dravet syndrome.* Epilepsy Research, 2016. 128: p. 43-47.

7. Higurashi, N., V. Broccoli, and S. Hirose, *Genetics and gene therapy in Dravet syndrome.* Epilepsy & Behavior, 2021: p. 108043.

8. Lueck, J.D., et al., *Engineered transfer RNAs for suppression of premature termination codons.* Nature Communications, 2019. 10(1): p. 822.

9. Howard, M., R.A. Frizzell, and D.M. Bedwell, *Aminoglycoside antibiotics restore CFTR function by overcoming premature stop mutations.* Nat Med, 1996. 2(4): p. 467-9.

10. Arakawa, M., et al., *Negamycin restores dystrophin expression in skeletal and cardiac muscles of mdx mice.* J Biochem, 2003. 134(5): p. 751-8.

11. Welch, E.M., et al., *PTC124 targets genetic disorders caused by nonsense mutations.* Nature, 2007. 447(7140): p. 87-91.

12. Roy, B., et al., *Ataluren stimulates ribosomal selection of near-cognate tRNAs to promote nonsense suppression.* Proc Natl Acad Sci U S A, 2016. 113(44): p. 12508-12513.

13. Berecki, G., et al., *SCN1A gain of function in early infantile encephalopathy.* Ann Neurol, 2019. 85(4): p. 514-525.

14. Dhifallah, S., et al., *Gain of Function for the SCN1A/hNav1.1-L1670W Mutation Responsible for Familial Hemiplegic Migraine.* Front Mol Neurosci, 2018. 11: p. 232.

15. Kotecha, B. and G.P. Richardson, *Ototoxicity in vitro: effects of neomycin, gentamicin, dihydrostreptomycin, amikacin, spectinomycin, neamine, spermine and poly-L-lysine.* Hear Res, 1994. 73(2): p. 173-84.

16. Dai, W.J., et al., *CRISPR-Cas9 for in vivo Gene Therapy: Promise and Hurdles.* Mol Ther Nucleic Acids, 2016. 5: p. e349.

17. Peng, R., G. Lin, and J. Li, *Potential pitfalls of CRISPR/Cas9-mediated genome editing.* FEBS J, 2016. 283(7): p. 1218-31.

18. Coller, J. and M. Wickens, *Chapter Fourteen - Tethered Function Assays: An Adaptable Approach to Study RNA Regulatory Proteins*, in *Methods in Enzymology*, J. Lorsch, Editor. 2007, Academic Press. p. 299-321.

19. Coller, J.M., N.K. Gray, and M.P. Wickens, *mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation.* Genes & development, 1998. 12(20): p. 3226-3235.

20. Cox, D.B., et al., *RNA editing with CRISPR-Cas13.* Science, 2017. 358(6366): p. 1019-1027.

21. Gootenberg, J.S., et al., *Nucleic acid detection with CRISPR-Cas13a/C2c2.* Science, 2017. 356(6336): p. 438-442.

22. Terns, M.P., *CRISPR-Based Technologies: Impact of RNA-Targeting Systems.* Molecular Cell, 2018. 72(3): p. 404-412.

23. Nietfeld, W., H. Mentzel, and T. Pieler, *The Xenopus laevis poly(A) binding protein is composed of multiple functionally independent RNA binding domains.* The EMBO Journal, 1990. 9(11): p. 3699-3705.

24. Deardorff, J.A. and A.B. Sachs, *Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast Poly(A)-binding protein11Edited by K. Nagai.* Journal of Molecular Biology, 1997. 269(1): p. 67-81.

25. Arango, D., et al., *Acetylation of Cytidine in mRNA Promotes Translation Efficiency.* Cell, 2018. 175(7): p. 1872-1886.e24.

26. Yang, N., et al., *Generation of pure GABAergic neurons by transcription factor programming.* Nat Methods, 2017. 14(6): p. 621-628.

SEQUENCES

SEQ ID NO: 1 >NR_148667.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=16]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGAACAGCATAAGAATTATTTCTGAGTGGAGGTGAGGCTTGTCCAAATGTCTTTGCTATCATGG
ATTTCCTGACTCCTACCTGTTTGAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAAGATGG
AGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGAAAGAC
GCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAATAGTG
ACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGAGGACC
TGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGCCACCT
CTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATTCAGCA
TGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAATGTAG
AATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGATTTTA
CTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGACCTGG
GCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAAAACCA
TTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGTATTTG
CTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTCCTTGG
AGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTTTGACT
GGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAATAGCT
CTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAGCTTTG
ATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACTGACAT
TACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAATTTGA
TCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGCCGAAT
TTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGAACATT
CCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGCTAAGG
AAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATTCCAAA
AATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGAAAAGA
GGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAACAAGCC
TTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTTTGAGG
ATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCAGACCA
GTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGTGGTTT
CCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAACCACTGAAACTGAAA
TGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGCAATGA
GTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTATAAAT
TTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGTGATGG
ACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCCAATGA
CGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCTGAAAA
TTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAGCCTGG
TAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTTGGCAA
AATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCTCGTCT
TGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTGCAAGA
TCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGTGCTGT
GTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCATGATGG
TCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAACCTTG
CAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTATGTGA
AAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACTTGATG
ATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAAAGATG
TAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACATGTCAT
TCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACACGGAAG
ACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAGCACTG
TGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTTCACTG
AAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCTGAGAA
GGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGGTGCTC
TGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTTCACTT
ACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTGTTGGC
TGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGCCATCA
AATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGATAATCTTGCTC
CAACTTGGATGGGGTGGAGCGCTGGTTCCTCCCCTGAGCCCTTTATTATGGGTGGTTGTGAATGCCCTTTTAGGAGCA

ATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGTTT
GCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGACGTGAATAATCATACTGAT
TGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGATTTGGG
TATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAAATGTG
GAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTC
ACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTT
ATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATACCTCGA
CCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACATAAGCATCATGATTCTCATC
TGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTACCATTTTGTCACGCATCAAT
CTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCATTATTATTTTACCATTGGA
TGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTC
GTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGG
ATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCATG
TTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCAACTTT
GAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCATT
CTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCA
TCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCATC
CTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGGTT
TGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCG
CCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGATCCAC
TGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATACAGATG
GAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTACTTTAAAACGAAAACAAGAG
GAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACG
TACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAATTGACAGAATAAATGAAAAC
TCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAATT
GTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATTGGGTG
ACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACTGACTG
TTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAA
AGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGATAAGATGCACAATGGCTAG
TCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTAGT
AATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCTTTTTG
TTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAA
CAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCATGC
TGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAGGGGTA
ATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTGGG
TAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGTAT
TAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTACTCTTT
ATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTAATGAA
ATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCACAGTAT
TTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTACCATTT
ATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCT
CATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATTAAGCA
ATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATATGGTATATAGCCAGACTGT
ACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTATAAGCAAATATAAATACTGT
AAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTATCTTCAGGTTGATATCACAA
TCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGACTGGGT
AGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAGGTTCA
CGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTGCACTA
ACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTA
ATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTATA
TAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGA
TAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAACCAACATTTAAAATGTGCA
AAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTGTTAACAGAATATTCTAATGATATGGA
GATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACACACACACACACACACACACACACACAC
ACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGA
ATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAG
AAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGAAACCCTTAAATCATCATTTAAAAATTTA
GAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCC
AGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCA
CCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACT
GGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTCAG
GTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTA
ACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATGCCTCATTTCAGCATGCCATCTGTCTCCT

GCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGT
CTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCACA
GATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTAAATTAATCTATCTAAATTTCTGAGGA
ACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCTCAACCAGAAAAGGGACAGTAATTAAA
AGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACATCACAATCCTAAAGCAACACAACAATTAA
TTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAGAC
TTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAAATACAAAGAGAAGAACTGGTAAAACATAATCA
TATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGA
TGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCA
AAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAATACATACTGATGTAAGTAAATAAGGAAAA
GTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAAATAACGGAGTTAGAAAAATCACTATTTG
GCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAATGAGTATTAAAACTCATGGGTGAAAGTTTGATGAGG
AATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTGGA
GAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGACCAACTGATACTATGTGCCTCCTGATAAGATAT
ACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCAGAATCTAACTACAAGGAAAATCAAATTG
AGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAATT
CCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGAATATCAAATTAAAAAATAACAGATAAATAATA
TTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTAGTATTATAATAGTGTTAATTTTCCTAAT
TTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTTAG
GGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAATGAGTTACATTTATATATATATGTAT
GTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTGCT
GTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGTATGCCCTGGACTTTATAATTTCTACCTT
CAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTAAT
CTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAATAATCACTGTTTAATTAAAATAAGAATATTTT
AGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAACAGATCTGATTTGGATAACAGAAGGTGGC
ACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATTCCAAACACAGCTCCACAGAGCAGCACCC
TTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATAATGCCATAAATATTTATAAAGCTGAACA
TATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTTCTAAAGACTAATACCATTTAATGTTTAA
ATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAAAT
TGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAATCCAAAAGTCTAGTAACCATTGGGACAGA
CAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCCAGACACTCGAACTATATAAACTTGTCGC
ACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCACTGGAATTTCTTATTATATAAGAAAATAA
ATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTATTCTTCCTGGCACATGGAAATATCCCACTGAGG
TTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGAGGAAAATACTCTAGACAATCAAGTATAG
ATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATTTGCAATATCAATTGATTTTACTCAACTG
CTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGATAACTCAGCCCCTCACCTCAGCAACAAGCAAG
AGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATATCTAAGTCCCTTTTCCCTTCCCTTTTCC
TTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTT
TTCCTTCCTTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCT
CTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAGAGATCAATTTTTCCCATTCATTTCTACA
TTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTATTCTTTCCCTCTTGACTGATCCTTTTTGT
CTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTTGCTCAACCCTGATGCATCTTTATTCACT
GTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCTATACTCACTACCTTCACTCTTGGTTAAA
ATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTGAATAATCTTGTTGATATAAACACTTTGG
GTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGGGAAGGGCTGGAAAATACAAAAGGTAGAC
ATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACTCTGTCTCTGGTTATCATAAATACATCTCCTAT
ATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGTTCACTGTGATGGGTTTTTTTGTTTTCTC
CTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 2 >XR_001738884.1 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=X2]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGTGTGGACATA
AGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAA
GATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGA
AAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAA
TAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGA
GGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGC
CACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATT

CAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAA
TGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGA
TTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGA
CCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAA
AACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGT
ATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTC
CTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTT
TGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAA
TAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAG
CTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACT
GACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAA
TTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGC
CGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGA
ACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGC
TAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATT
CCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGA
AAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAAC
AAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTT
TGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCA
GACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGT
GGTTTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAACCACTGAAAC
TGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGC
AATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTA
TAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGT
GATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCC
AATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCT
GAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAG
CCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTT
GGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCT
CGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTG
CAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGT
GCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCAT
GATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAA
CCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTA
TGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACT
TGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAA
AGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACAT
GTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACAC
GGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAG
CACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTT
CACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCT
GAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGG
TGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTT
CACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTG
TTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGC
CATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGATAATCT
TGCTCCAACTTGGATGGGGTGGAGCGCTGGTTCCTCCCCTGAGCCCTTTATTATGGGTGGTTGTGAATGCCCTTTTAG
GAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATT
TGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGACGTGAATAATCATA
CTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGAT
TTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAA
ATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCT
TCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACA
TCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATAC
CTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACATAAGCATCATGATTC
TCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTACCATTTTGTCACGCA
TCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCATTATTATTTTACCA
TTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGT
ATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAA
AGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAG
TCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCA
ACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCAC
CCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGA

ACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGG
TCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATG
AGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTG
AACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGA
TCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATAC
AGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTACTTTAAAACGAAAAC
AAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCT
TTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAATTGACAGAATAAATG
AAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGC
CAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATT
GGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACT
GACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTG
TGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGATAAGATGCACAATG
GCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGT
GTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCT
TTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTC
TATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGG
CATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAG
GGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTT
GTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGT
TGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTAC
TCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTA
ATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCAC
AGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTAC
CATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAG
AATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATT
AAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATATGGTATATAGCCAG
ACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTATAAGCAAATATAAAT
ACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTATCTTCAGGTTGATAT
CACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGAC
TGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAG
GTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTG
CACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTG
GAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAAC
TTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTAC
TTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAACCAACATTTAAAAT
GTGCAAAACTAATAAAGATTACATTTTTTATTTTA

SEQ ID NO: 3 >XR_001738883.1 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=X1]

AGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTGTGTGT
TTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGAATCTG
AACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAATTTTC
CTTTTAATCAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGTGTGGACATAAGTTT
TCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAAGATGG
AGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGAAAGAC
GCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAATAGTG
ACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGAGGACC
TGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGCCACCT
CTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATTCAGCA
TGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAATGTAG
AATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGATTTTA
CTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGACCTGG
GCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAAAACCA
TTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGTATTTG
CTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTCCTTGG
AGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTTTGACT
GGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAATAGCT
CTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAGCTTTG
ATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACTGACAT
TACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAATTTGA

TCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGCCGAAT
TTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGAACATT
CCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGCTAAGG
AAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATTCCAAA
AATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGAAAAGA
GGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAACAAGCC
TTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTTTGAGG
ATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCAGACCA
GTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGTGGTTT
CCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGTGATAATAGATAAGCCAGCTA
CTGATGACAATGGAACAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTC
TAGAAGATCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAAT
CCAGGCAGAAATGCCCACCCTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAA
AAGTGAAACATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATA
CTCTTTTCATGGCCATGGAGCACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCA
CTGGGATCTTTACAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATA
TCTTTGACGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTT
CATTTCGATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATT
CCGTGGGGGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCT
TTGGTAAAAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCT
TCCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTG
GTCAAGCCATGTGCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCT
TGCTTCTGAGCTCATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTG
TGGATAGGATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAAC
AAAAGATTTTAGATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAG
AAATTGGGAAAGATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAA
AATACATTATTGATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAG
GAGAATCTGACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGA
ATGAAAGCAGTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTG
AAGAAACTCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAG
AAGGCAGAGGAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCA
TTGTTTTCATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGA
CGATGTTGGAATATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCT
ATCAAACATATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAA
ATGCCTTGGGTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTAT
CTCGATTTGAAGGGATGAGGGATAATCTTGCTCCAACTTGGATGGGGTGGAGCGCTGGTTCCTCCCCTGAGCCCTTTA
TTATGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGG
CTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGG
TTTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAAT
GTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGAT
ATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTAC
TTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAG
CAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTA
GGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGA
CAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGT
GAATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTC
ATCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGT
ATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGC
CGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTG
TTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAG
AGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACC
TCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCT
GGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTC
CTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTG
AGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAA
AAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATG
GATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAG
AGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAG
CCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTT
TTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAA
GAAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGT
CCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGG
AAATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAG

GACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGT
GACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCT
GACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTT
GGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATT
TTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGT
TTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAG
CTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAA
TTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAA
TCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTG
GAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACC
GCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCT
CTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGT
TTGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTT
TACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCT
GCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGT
TCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTT
AAGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGT
GTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATAT
GAAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAA
TTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGAT
TATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAA
ATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCT
GACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTT
TTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAAC
CATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATG
CTAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTA

SEQ ID NO: 4 >NM_001353949.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=6]

AAGTGTAGGAGACACACTGCTGGCCTGTGGAAACTCATGGAACTGTTCCTCCAGATTAACACTTCAGGGGTTATGGAA
GCTGGAGGAAGCTGAGCTTTTACTACATCTTTTGGGGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGG
ATGACAAGATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGG
CTATTGAAAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAA
AGCCAAATAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGC
CCCTGGAGGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGT
TCAGTGCCACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATT
CATTATTCAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGA
CAAAGAATGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTT
TAGAAGATTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGT
TTGTGGACCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAG
GCCTGAAAACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTC
TGAGCGTATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCA
ATGCTTCCTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCT
TTGAGTTTGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTAT
GTGGAAATAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCT
ACACAAGCTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTT
ATCAACTGACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACC
TAATAAATTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGA
AAGAGGCCGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTG
CCTCAGAACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCA
AGAGTGCTAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGG
ATGAATTCCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGA
CATATGAAAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATA
GCAGAACAAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACA
GCACCTTTGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACC
TGAGTCAGACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCA
ATGGTGTGGTTTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAACCA
CTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGC
AACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCT
GTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACC
TGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGC

ACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAA
TGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGA
CGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTT
TCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATT
TAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATT
GTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGT
TCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTG
TCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTG
CAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAG
TAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTA
AACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACT
ATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTG
ATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATT
TAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAG
AAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAG
CTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGT
GGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTA
GTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACA
AGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATG
CCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAAC
TTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGG
TGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTT
TCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACA
TCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAG
TAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGT
ATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTA
TTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAA
AGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGA
AAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTT
TTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATG
TGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTC
TACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTC
TTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCC
TACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACA
TCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAG
TTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTG
GCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCT
CAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTG
TGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGG
ATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTAT
CTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGC
CCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAG
AGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCA
CTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGC
GAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACA
TGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTT
CCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAA
TGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCT
TTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCT
TGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTG
CTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTG
TTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGG
AATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTT
GAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACA
AAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAA
ATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCC
CCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCAT
AAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAA
GACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCA
AGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCA
AACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAG
GTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTAT
CTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAG

ATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGA
TAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGA
ATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGG
AAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTA
AATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCA
CAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTAT
TTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCG
GTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTC
CACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGC
TCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATA
ATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTG
TTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACACAC
ACACACACACACACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATGTCACTG
GAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGG
ATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGAAA
CCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTC
TGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTC
CCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCTCC
AATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATACA
ACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCT
TTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATGCC
TCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAA
AGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAACCA
CAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTA
AATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCT
CAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACATCA
CAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTC
AGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAAATACAA
AGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGCTA
AAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACTGG
CCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAATAC
ATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAAAT
AACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAATGAGTATTAA
AACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAAT
TTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGACCAACTG
ATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCAGA
ATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTCAT
GAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGAATATCA
AATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTAGT
ATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATTCT
CTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAA
TGAGTTACATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAG
GGGAGGATTTTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGTAT
GCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTAAC
TACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAATAATCA
CTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAACAG
ATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATTCC
AAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATAAT
GCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTTCT
AAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTTTA
ATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAATCC
AAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCCAG
ACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCACTG
GAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTATTCTTCC
TGGCACATGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGAGG
AAAATACTCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATTTG
CAATATCAATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGATAACTC
AGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATATC
TAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTT
TTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCCCT
TCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAGAG
ATCAATTTTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTATTC
TTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTTGC

TCAACCCTGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCTAT
ACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTGAA
TAATCTTGTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGGGA
AGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACTCTGTCT
CTGGTTATCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGTTC
ACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 5 >NM_001353958.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=13]

AAGTGTAGGAGACACACTGCTGGCCTGTGGAAACTCATGGAACTGTTCCTCCAGATTAACACTTCAGGGGTTATGGAA
GCTGGAGGAAGCTGAGCTTTTACTACATCTTTTGGGGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGT
GCAGGATTTTGTGTGGACATAAGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAA
TTAAAATGTGCAGGATGACAAGATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAG
AGAATCTCTTGCGGCTATTGAAAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGA
CGAAAATGGCCCAAAGCCAAATAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGA
GATGGTGTCAGAGCCCCTGGAGGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAA
GGCCATCTTCCGGTTCAGTGCCACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAA
GATTTTGGTACATTCATTATTCAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAA
CCCTCCTGATTGGACAAAGAATGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGC
AAGGGGATTCTGTTTAGAAGATTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGC
GTACGTCACAGAGTTTGTGGACCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGAT
TTCAGTCATTCCAGGCCTGAAAACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCT
GACTGTGTTCTGTCTGAGCGTATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACA
ATGGCCTCCCACCAATGCTTCCTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTAT
AAATGAAACTGTCTTTGAGTTTGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTT
AGATGCACTACTATGTGGAAATAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAA
TCCCAATTATGGCTACACAAGCTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTT
CTGGGAAAATCTTTATCAACTGACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTT
GGGCTCATTCTACCTAATAAATTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGA
AGAAGCAGAACAGAAAGAGGCCGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGC
AGCAACGGCAACTGCCTCAGAACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTC
TAAGTTGAGTTCCAAGAGTGCTAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGA
AGAGAAAGATGAGGATGAATTCCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGA
AGGGAACCGATTGACATATGAAAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTC
ACCAAGGCGAAATAGCAGAACAAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGC
AGATGATGAGCACAGCACCTTTGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAG
ACGCAACAGCAACCTGAGTCAGACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAG
CACTGTGGATTGCAATGGTGTGGTTTCCTTGGGAACAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTT
CCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATAC
AGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGA
CTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCAC
CATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCCAATGACGGACCATTTCAATAATGTGCTTAC
AGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTA
TTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGA
AGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCT
AATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGC
CGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACG
CTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTG
GGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGT
CCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAAT
GAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCA
ACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTG
TATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGG
AACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGT
GACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGA
AGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGA
ACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTG
TTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACA
TAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGA
TCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCT
AAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTC

ATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCT
GAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCAT
CATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATT
CTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACT
AATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTT
GCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCC
TAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCT
GTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGA
ACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAA
ATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACAT
GGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCAT
TGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTT
TGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTAC
CCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCT
GCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGC
CATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGG
CAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAA
GCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAAT
TTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTT
CAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTT
TGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCT
GCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATAT
CTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATT
CATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGC
TGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAA
CAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGA
AAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACA
TGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTT
ACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAA
ATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGA
CCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAG
GGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCA
CTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGT
GTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCA
CCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAG
ATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATA
AGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAA
ATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAA
CCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTG
CTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTAT
TATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCA
AATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTT
GGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGA
AACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAG
CACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTT
TAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACT
TTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTT
ACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATT
TCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATC
CAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAA
GGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTA
GATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTAC
ATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATT
CGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAA
GATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTGTTAACAGAATATTCTAATGATATGGAGATTTTTTACA
TTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACACACACACACACACACACACACACACACACACACAGAGGCA
TACCCACGTTGACAACAAAACCTAGGGTAGATATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAG
CAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAG
CCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGAAACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATG
TGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGAC
TTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAA
GCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCC
TCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCC

TCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGAC
ACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGAC
ACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGC
TAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGA
AAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCT
TCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCC
ATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACATCACAATCCTAAAGCAACACAACAATTAATTCTGATATAT
TAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTA
CCCATCCACGCCATTACTCACTGTTACCTTTCCAAATACAAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCC
ATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGT
CAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGAT
AGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAATACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTG
TTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAAATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCT
AATGACAATTGATTCATACAAGAATCATCAATGAGTATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATT
TATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCA
GACAATACCTTCAAAGATATCATCAATTATGAGACCAACTGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGC
CACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCAGAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTA
TAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCA
AAGAACTATGACAACTAAACACAATTCTGGATGGAATATCAAATTAAAAAATAACAGATAAATAATATTATTGGGAAA
GTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGAC
TAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCA
ACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAATGAGTTACATTTATATATATATGTATGTATGTAAACA
GACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTGCTGTGTGTTACTG
AAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGAT
CTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAA
GTCACAGAAATGATTACATTCCTTGTTCTTCATAATAATCACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCT
GCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAACAGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTT
TAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATTCCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGG
TTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATAATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTA
AAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTTCTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGG
AAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGG
TGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAATCCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTG
CATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCCAGACACTCGAACTATATAAACTTGTCGCACTAAAGACAG
CAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCACTGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTA
AACCACCATGAATTGGGTACTGCTGATAGTACTATTCTTCCTGGCACATGGAAATATCCCACTGAGGTTGTCAATCAA
TATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGAGGAAAATACTCTAGACAATCAAGTATAGATGGGAAAATT
GCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATTTGCAATATCAATTGATTTTACTCAACTGCTTGAGTCCTT
TATCATTCCGTTAGAACACTCCTCATATTTGCAGATAACTCAGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATT
ATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATATCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCT
TCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTT
TTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCT
TCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAGAGATCAATTTTTCCCATTCATTTCTACATTCTCAGATAC
CTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTATTCTTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCA
TTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTTGCTCAACCCTGATGCATCTTTATTCACTGTGTCATTATC
TCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCTATACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAA
TATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTGAATAATCTTGTTGATATAAACACTTTGGGTTTTCAAAAT
GTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGGGAAGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGA
GAACACCACCCCTATGGAGCCAAGCTATACATACTCTGTCTCTGGTTATCATAAATACATCTCCTATATTCAATGCTT
GCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGTTCACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATG
CAACTTTTATTTAT

SEQ ID NO: 6>NM_001353950.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=7]

AAGTGTAGGAGACACACTGCTGGCCTGTGGAAACTCATGGAACTGTTCCTCCAGATTAACACTTCAGGGGTTATGGAA
GCTGGAGGAAGCTGAGCTTTTACTACATCTTTTGGGGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGT
GCAGGATTTTGTGTGGACATAAGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAA
TTAAAATGTGCAGGATGACAAGATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAG
AGAATCTCTTGCGGCTATTGAAAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGA
CGAAAATGGCCCAAAGCCAAATAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGA
GATGGTGTCAGAGCCCCTGGAGGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAA
GGCCATCTTCCGGTTCAGTGCCACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAA

GATTTTGGTACATTCATTATTCAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAA
CCCTCCTGATTGGACAAAGAATGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGC
AAGGGGATTCTGTTTAGAAGATTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGC
GTACGTCACAGAGTTTGTGGACCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGAT
TTCAGTCATTCCAGGCCTGAAAACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCT
GACTGTGTTCTGTCTGAGCGTATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACA
ATGGCCTCCCACCAATGCTTCCTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTAT
AAATGAAACTGTCTTTGAGTTTGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTT
AGATGCACTACTATGTGGAAATAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAA
TCCCAATTATGGCTACACAAGCTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTT
CTGGGAAAATCTTTATCAACTGACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTT
GGGCTCATTCTACCTAATAAATTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGA
AGAAGCAGAACAGAAAGAGGCCGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGC
AGCAACGGCAACTGCCTCAGAACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTC
TAAGTTGAGTTCCAAGAGTGCTAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGA
AGAGAAAGATGAGGATGAATTCCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGA
AGGGAACCGATTGACATATGAAAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTC
ACCAAGGCGAAATAGCAGAACAAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGC
AGATGATGAGCACAGCACCTTTGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAG
ACGCAACAGCAACCTGAGTCAGACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAG
CACTGTGGATTGCAATGGTGTGGTTTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCC
AGAGGGAACAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGA
TCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCA
GAAATGCCCACCCTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAA
ACATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTT
CATGGCCATGGAGCACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGAT
CTTTACAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGA
CGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCG
ATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGG
GGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAA
AAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTC
CTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGC
CATGTGCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCT
GAGCTCATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAG
GATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGAT
TTTAGATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGG
GAAAGATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACAT
TATTGATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATC
TGACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAG
CAGTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAAC
TCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAG
AGGAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTT
CATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTT
GGAATATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAAC
ATATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTT
GGGTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATT
TGAAGGGATGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTAT
ATTCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGG
TGACAGGTTTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATG
GAAAAATGTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATG
GATGGATATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTA
TCTTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTT
CAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAA
AAAATTAGGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGT
AACCAGACAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGA
CCAGAGTGAATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACT
GAAACTCATCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCAT
TGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAG
GATTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCC
TGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTA
TGTTAAGAGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAAT
TACAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGT
TAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCAT

ATCCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGA
GCCTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGA
ATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCAT
TGCCATGGATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCT
AGGAGAGAGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTC
CTATCAGCCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACG
CCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCT
TATAAAAGAAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGC
AGCTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGC
CAAAGGGAAATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTA
TCAACAGGACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATA
AGACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACT
ACCAGCTGACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTG
TGATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTG
TCACATTTTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAA
ACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCC
CTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAA
AAACAAATTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTT
TGCGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAA
ATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACC
CTCCACCGCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATC
TTCGGCTCTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGG
TGCACGTTTGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTG
CAAGCTTTACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTA
CCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGG
AAATTGTTCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTT
ACATTTTAAGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTA
AATATGTGTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACA
TAAATATGAAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATG
AAAGAAATTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAAT
TTTAGATTATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGG
GTTTAAAATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTT
TTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTC
ATAATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGA
AGCAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTAC
TATTATGCTAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGC
CCAGTCACTTTTTGTTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGC
GCGCGCACACACACACACACACACACACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGG
TAGATATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGA
ACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGA
GCTATGAAAAGAAACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCC
AGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTC
CTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGG
AATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGA
CTGCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCT
GCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAA
TTCTTCTATATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAG
TCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGT
AACTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTA
AAATTAACAGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAG
AGGGAATATAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCC
TCTTCATAACATCACAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCC
AACTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTAC
CTTTCCAAATACAAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGA
TATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAG
GGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAA
GAATCAGCAAATACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACA
ATTGTAGAAGAAATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCA
TCAATGAGTATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCT
ATTAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAAT
TATGAGACCAACTGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAA
ATTTTAAAACCAGAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAA
AATGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTC

TGGATGGAATATCAAATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTT
TATTAGTTATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAA
AATGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATC
AGTTCAAAGAAAAATGAGTTACATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGT
GTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACT
ATCATATTTTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATAT
ACTGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTT
CTTCATAATAATCACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTAT
TTGATGCTAAACAGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTA
TAAATAGTAATTCCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGA
AAATTAAGAATAATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTT
TTAGATTATTTTCTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATG
GGTCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACT
GGTGCATGAAATCCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGT
ATCACAAAGCCCAGACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCT
TCATGTGTGCACTGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGAT
AGTACTATTCTTCCTGGCACATGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAA
GGCACTAGAGGAGGAAAATACTCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCAC
AAGTGGATAATTTGCAATATCAATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATA
TTTGCAGATAACTCAGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATC
CAGTGAAAAATATCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTC
CTTTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCT
TTCCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTT
ATATCTTTTCAGAGATCAATTTTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTC
TTCTATCTTTATTCTTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTA
AAAAATAACTTTGCTCAACCCTGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAA
AGCCAGGGCTCTATACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAA
AATTTGACTGTGAATAATCTTGTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGT
AACATTATGAGGGAAGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTA
TACATACTCTGTCTCTGGTTATCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAG
CAATTCAATTGTTCACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 7 >NM_001202435.3 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=4]

AAGTGTAGGAGACACACTGCTGGCCTGTGGAAACTCATGGAACTGTTCCTCCAGATTAACACTTCAGGGGTTATGGAA
GCTGGAGGAAGCTGAGCTTTTACTACATCTTTTGGGGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGT
GCAGGATTTTGTGTGGACATAAGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAA
TTAAAATGTGCAGGATGACAAGATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAG
AGAATCTCTTGCGGCTATTGAAAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGA
CGAAAATGGCCCAAAGCCAAATAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGA
GATGGTGTCAGAGCCCCTGGAGGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAA
GGCCATCTTCCGGTTCAGTGCCACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAA
GATTTTGGTACATTCATTATTCAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAA
CCCTCCTGATTGGACAAAGAATGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGC
AAGGGGATTCTGTTTAGAAGATTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGC
GTACGTCACAGAGTTTGTGGACCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGAT
TTCAGTCATTCCAGGCCTGAAAACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCT
GACTGTGTTCTGTCTGAGCGTATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACA
ATGGCCTCCCACCAATGCTTCCTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTAT
AAATGAAACTGTCTTTGAGTTTGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTT
AGATGCACTACTATGTGGAAATAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAA
TCCCAATTATGGCTACACAAGCTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTT
CTGGGAAAATCTTTATCAACTGACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTT
GGGCTCATTCTACCTAATAAATTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGA
AGAAGCAGAACAGAAAGAGGCCGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGC
AGCAACGGCAACTGCCTCAGAACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTC
TAAGTTGAGTTCCAAGAGTGCTAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGA
AGAGAAAGATGAGGATGAATTCCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGA
AGGGAACCGATTGACATATGAAAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTC
ACCAAGGCGAAATAGCAGAACAAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGC
AGATGATGAGCACAGCACCTTTGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAG

ACGCAACAGCAACCTGAGTCAGACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAG
CACTGTGGATTGCAATGGTGTGGTTTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCC
AGAGGTGATAATAGATAAGCCAGCTACTGATGACAATGGAACAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAG
TTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAAC
AAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTATAAATTTTCCAACATATTCTTAAT
CTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGC
CATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCCAATGACGGACCATTTCAATAATGT
GCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTA
CTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAA
TGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAA
TATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCAT
TTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACT
CCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGAC
CATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCT
AGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAA
TGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATT
TATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGA
CAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGG
TATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCT
TACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGA
TCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGT
AGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATT
CAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGT
TGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATA
TATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAAT
GCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGA
TGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAG
AGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCC
ATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGG
CAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGACGTGAATAATCATACTGATTGCCT
AAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCT
CTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACT
CCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTT
GAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGAC
AGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGG
AAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCT
TAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTACCATTTTGTCACGCATCAATCTGGT
GTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCATTATTATTTTACCATTGGATGGAA
TATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTC
CCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCG
CACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCAT
CTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGAC
CTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAA
CAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGT
TGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGA
GAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGA
GAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCT
CAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCT
TGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGA
GCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGT
ATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAA
TAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAATTGACAGAATAAATGAAAACTCTAT
TACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGA
AAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATTGGGTGACAAA
TTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTT
ACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGG
AGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGA
CTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTG
TATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAAT
CCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGT
ATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCAC
TTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGG
GTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGC

CAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAA
AAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAA
TTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCT
ATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCACAGTATTTATT
GCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTACCATTTATAGA
TAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTT
ATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGT
TTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATATGGTATATAGCCAGACTGTACAGA
CATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAA
GTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTATCTTCAGGTTGATATCACAATCACT
TTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTG
TAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATT
CATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGA
ATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCT
GCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTC
AATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAAC
TGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAACCAACATTTAAAATGTGCAAAACT
AATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTGTTAACAGAATATTCTAATGATATGGAGATTT
TTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACACACACACACACACACACACACACACACACACACA
GAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTAC
AAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTA
TGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGAAACCCTTAAATCATCATTTAAAAATTTAGAATT
GCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAG
GGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGT
GGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACT
CATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAAT
CACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATT
GCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTG
GCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAA
TCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCACAGATAT
CATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAG
AAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCT
TTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACATCACAATCCTAAAGCAACACAACAATTAATTCTG
ATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTT
ATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAAATACAAAGAGAAGAACTGGTAAAACATAATCATATAA
ATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGG
TCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATA
AATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAATACATACTGATGTAAGTAAATAAGGAAAAGTACA
AATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAAATAACGGAGTTAGAAAAATCACTATTTGGCAAT
CACCCTAATGACAATTGATTCATACAAGAATCATCAATGAGTATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAG
GGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAAT
ACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGACCAACTGATACTATGTGCCTCCTGATAAGATATACTGA
AAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCAGAATCTAACTACAAGGAAAATCAAATTGAGGAC
ACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATG
AGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGAATATCAAATTAAAAAATAACAGATAAATAATATTATT
GGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGT
TAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTG
AATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAATGAGTTACATTTATATATATATGTATGTATG
TAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTGCTGTGTG
TTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAA
TTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCC
TGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAATAATCACTGTTTAATTAAAATAAGAATATTTTAGAAA
AGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAACAGATCTGATTTGGATAACAGAAGGTGGCACTTT
GCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATTCCAAACACAGCTCCACAGAGCAGCACCCTTATG
ACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATAATGCCATAAATATTTATAAAGCTGAACATATCC
AAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTTCTAAAGACTAATACCATTTAATGTTTAAATGTT
CTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCA
ACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAATCCAAAAGTCTAGTAACCATTGGGACAGACAACT
CTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCCAGACACTCGAACTATATAAACTTGTCGCACTAA
AGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCACTGGAATTTCTTATTATATAAGAAAATAAATGTG
TTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTATTCTTCCTGGCACATGGAAATATCCCACTGAGGTTGTC
AATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGAGGAAAATACTCTAGACAATCAAGTATAGATGGG
AAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATTTGCAATATCAATTGATTTTACTCAACTGCTTGA

GTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGATAACTCAGCCCCTCACCTCAGCAACAAGCAAGAGGAC
ATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATATCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCC
TTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCT
TCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCC
TCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAGAGATCAATTTTTCCCATTCATTTCTACATTCTC
AGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTATTCTTTCCCTCTTGACTGATCCTTTTTGTCTGCC
TTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTTGCTCAACCCTGATGCATCTTTATTCACTGTGTC
ATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCTATACTCACTACCTTCACTCTTGGTTAAAATTAA
AAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTGAATAATCTTGTTGATATAAACACTTTGGGTTTT
CAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGGGAAGGGCTGGAAAATACAAAAGGTAGACATAGG
AACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACTCTGTCTCTGGTTATCATAAATACATCTCCTATATTCA
ATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGTTCACTGTGATGGGTTTTTTTGTTTTCTCCTCAT
GGAATGCAACTTTTATTTAT

SEQ ID NO: 8 >NM_001353955.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=11]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAAGATGGAGCAAACAG
TGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGAAAGACGCATTGCAG
AAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAATAGTGACTTGGAAG
CTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGAGGACCTGGACCCCT
ACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGCCACCTCTGCCCTGT
ACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATTCAGCATGCTAATTA
TGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAATGTAGAATACACCT
TCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGATTTTACTTTCCTTC
GGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGACCTGGGCAATGTCT
CGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAAAACCATTGTGGGAG
CCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGTATTTGCTCTAATTG
GGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTCCTTGGAGGAACATA
GTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTTTGACTGGAAGTCAT
ATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAATAGCTCTGATGCAG
GCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAGCTTTGATACCTTCA
GTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACTGACATTACGTGCTG
CTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAATTTGATCCTGGCTG
TGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGCCGAATTTCAGCAGA
TGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGGCAGCAACGGCAACTGCCTCAGAACATTCCAGAGAGCCCA
GTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGCTAAGGAAAGAAGAAATC
GGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATTCCAAAAATCTGAATCTG
AGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGAAAAGAGGTACTCCTCCC
CACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAACAAGCCTTTTCAGCTTTA
GAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTTTGAGGATAACGAGAGCC
GTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCAGACCAGTAGGTCATCCC
GGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGTGGTTTCCTTGGTTGGTG
GACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAACCACTGAAACTGAAATGAGAAAGAGAA
GGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCA
TTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTATAAATTTTCCAACATAT
TCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTG
ACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCCAATGACGGACCATTTCA
ATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCTGAAAATTATTGCCATGG
ATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGAC
TCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAA
CGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCG
TCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATT
GTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGA
TAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCATGATGGTCATGGTGATTG
GAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAACCTTGCAGCCACTGATG
ATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATAT
ATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACTTGATGATCTAAACAACA
AGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAAAGATGTAAATGGAACTA
CAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACATGTCATTCATAAACAACC
CCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTG

AATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCG
CACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTAC
AAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCC
GAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAG
ATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTTCACTTACATTTTCATTC
TGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAA
TTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGA
CACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGTGGTTGTGAATGCCCTTTTAGGAG
CAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGT
TTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGACGTGAATAATCATACTG
ATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGATTTG
GGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAAATG
TGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCT
TCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCT
TTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATACCTC
GACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACATAAGCATCATGATTCTCA
TCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTACCATTTTGTCACGCATCA
ATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCATTATTATTTTACCATTG
GATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATT
TCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGG
GGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCA
TGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCAACT
TTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCA
TTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACC
CATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCA
TCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGG
TTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAAC
CGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGATCC
ACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATACAGA
TGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTACTTTAAAACGAAAACAAG
AGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTA
CGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAATTGACAGAATAAATGAAA
ACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAA
TTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATTGGG
TGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACTGAC
TGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGT
AAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGATAAGATGCACAATGGCT
AGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTA
GTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCTTTT
TGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTAT
AACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCAT
GCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAGGGG
TAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTG
GGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGT
ATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTACTCT
TTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTAATG
AAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCACAGT
ATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTACCAT
TTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAAT
CTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATTAAG
CAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATATGGTATATAGCCAGACT
GTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTATAAGCAAATATAAATACT
GTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTATCTTCAGGTTGATATCAC
AATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGACTGG
GTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAGGTT
CACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTGCAC
TAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAG
TAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTA
TATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTA
GATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAACCAACATTTAAAATGTG
CAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTGTTAACAGAATATTCTAATGATATG
GAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACACACACACACACACACACACACACAC
ACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATGTCACTGGAGGTAGGGGGTAATGACCTCCCA

GAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAG
AGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGAAACCCTTAAATCATCATTTAAAAATT
TAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTC
CCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAG
CACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCA
CTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTC
AGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCC
TAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATGCCTCATTTCAGCATGCCATCTGTCTC
CTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGG
GTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCA
CAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTAAATTAATCTATCTAAATTTCTGAG
GAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCTCAACCAGAAAAGGGACAGTAATTA
AAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACATCACAATCCTAAAGCAACACAACAATT
AATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAG
ACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAAATACAAAGAGAAGAACTGGTAAAACATAAT
CATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAAT
GATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCAT
CAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAATACATACTGATGTAAGTAAATAAGGAA
AAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAAATAACGGAGTTAGAAAAATCACTATT
TGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAATGAGTATTAAAACTCATGGGTGAAAGTTTGATGA
GGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTG
GAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGACCAACTGATACTATGTGCCTCCTGATAAGAT
ATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCAGAATCTAACTACAAGGAAAATCAAAT
TGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAA
TTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGAATATCAAATTAAAAAATAACAGATAAATAA
TATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTAGTATTATAATAGTGTTAATTTTCCTA
ATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTT
AGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAATGAGTTACATTTATATATATATGT
ATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTG
CTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGTATGCCCTGGACTTTATAATTTCTACC
TTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTA
ATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAATAATCACTGTTTAATTAAAATAAGAATATT
TTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAACAGATCTGATTTGGATAACAGAAGGTG
GCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATTCCAAACACAGCTCCACAGAGCAGCAC
CCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATAATGCCATAAATATTTATAAAGCTGAA
CATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTTCTAAAGACTAATACCATTTAATGTTT
AAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAA
ATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAATCCAAAAGTCTAGTAACCATTGGGACA
GACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCCAGACACTCGAACTATATAAACTTGTC
GCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCACTGGAATTTCTTATTATATAAGAAAAT
AAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTATTCTTCCTGGCACATGGAAATATCCCACTGA
GGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGAGGAAAATACTCTAGACAATCAAGTAT
AGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATTTGCAATATCAATTGATTTTACTCAAC
TGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGATAACTCAGCCCCTCACCTCAGCAACAAGCA
AGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATATCTAAGTCCCTTTTCCCTTCCCTTTT
CCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCT
TTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCT
CTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAGAGATCAATTTTTCCCATTCATTTCTA
CATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTATTCTTTCCCTCTTGACTGATCCTTTTT
GTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTTGCTCAACCCTGATGCATCTTTATTCA
CTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCTATACTCACTACCTTCACTCTTGGTTA
AAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTGAATAATCTTGTTGATATAAACACTTT
GGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGGGAAGGGCTGGAAAATACAAAAGGTAG
ACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACTCTGTCTCTGGTTATCATAAATACATCTCCT
ATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGTTCACTGTGATGGGTTTTTTTGTTTTC
TCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 9 >NM_001353957.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=12]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA

ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAAGATGGAGCAAACAG
TGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGAAAGACGCATTGCAG
AAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAATAGTGACTTGGAAG
CTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGAGGACCTGGACCCCT
ACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGCCACCTCTGCCCTGT
ACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATTCAGCATGCTAATTA
TGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAATGTAGAATACACCT
TCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGATTTTACTTTCCTTC
GGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGACCTGGGCAATGTCT
CGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAAAACCATTGTGGGAG
CCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGTATTTGCTCTAATTG
GGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTCCTTGGAGGAACATA
GTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTTTGACTGGAAGTCAT
ATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAATAGCTCTGATGCAG
GCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAGCTTTGATACCTTCA
GTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACTGACATTACGTGCTG
CTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAATTTGATCCTGGCTG
TGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGCCGAATTTCAGCAGA
TGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGAACATTCCAGAGAGC
CCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGCTAAGGAAAGAAGAA
ATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATTCCAAAAATCTGAAT
CTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGAAAAGAGGTACTCCT
CCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAACAAGCCTTTTCAGCT
TTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTTTGAGGATAACGAGA
GCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCAGACCAGTAGGTCAT
CCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGTGGTTTCCTTGGGAA
CAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCC
AAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCC
CACCCTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTG
TCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCA
TGGAGCACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAG
CAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTA
TTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGC
GAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGG
GAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACA
AAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGA
TTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCC
TTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCAT
TTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACA
AAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATG
AAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATC
TTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATG
AAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTG
AAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCT
CATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAAC
CAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAAC
AATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTC
TCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATG
CTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCA
CCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACT
CAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGA
TGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGC
TAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGT
TTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATG
TGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATA
TAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACT
TTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGC
AGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAG
GATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGAC
AAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTG
AATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCA
TCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTA

TGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCC
GAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGT
TTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGA
GGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCT
CTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTG
GAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCC
TGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGA
GTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAA
AATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGG
ATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGA
GTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGC
CAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTT
TAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAG
AAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTC
CACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGA
AATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGG
ACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTG
ACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTG
ACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTG
GGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTT
TTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTT
TCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGC
TACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAAT
TCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAAT
CCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGG
AGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCG
CCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTC
TCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTT
TGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTT
ACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTG
CTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTT
CAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTA
AGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTG
TTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATG
AAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAAT
TTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATT
ATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAA
TTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTG
ACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTT
TTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACC
ATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGC
TAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCAC
TTTTTGTTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCAC
ACACACACACACACACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATG
TCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGT
CAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAA
AAGAAACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCAC
TCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTT
GAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTC
TTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCT
CATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAG
AGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTA
TATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGG
AAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGG
GAACCACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAAC
AGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATA
TAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATA
ACATCACAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGG
TAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAA
ATACAAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCT
GTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCC
CACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGC
AAATACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGA

AGAAATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAATGAG
TATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTA
GAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGAC
CAACTGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAA
ACCAGAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAA
TGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGA
ATATCAAATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTT
ATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCT
TATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAA
GAAAAATGAGTTACATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTT
AGAAAGGGGAGGATTTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATT
TTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTT
TTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAA
TAATCACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCT
AAACAGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGT
AATTCCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAG
AATAATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTA
TTTTCTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTG
GTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATG
AAATCCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAA
GCCCAGACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGT
GCACTGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTAT
TCTTCCTGGCACATGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAG
AGGAGGAAAATACTCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGAT
AATTTGCAATATCAATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGA
TAACTCAGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAA
AATATCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCC
TTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTT
CTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTT
TCAGAGATCAATTTTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCT
TTATTCTTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAA
CTTTGCTCAACCCTGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGG
CTCTATACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGAC
TGTGAATAATCTTGTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTAT
GAGGGAAGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACT
CTGTCTCTGGTTATCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAA
TTGTTCACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 10 >NM_001353951.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=8]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAAGATGGAGCAAACAG
TGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGAAAGACGCATTGCAG
AAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAATAGTGACTTGGAAG
CTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGAGGACCTGGACCCCT
ACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGCCACCTCTGCCCTGT
ACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATTCAGCATGCTAATTA
TGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAATGTAGAATACACCT
TCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGATTTTACTTTCCTTC
GGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGACCTGGGCAATGTCT
CGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAAAACCATTGTGGGAG
CCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGTATTTGCTCTAATTG
GGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTCCTTGGAGGAACATA
GTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTTTGACTGGAAGTCAT
ATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAATAGCTCTGATGCAG
GCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAGCTTTGATACCTTCA
GTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACTGACATTACGTGCTG
CTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAATTTGATCCTGGCTG
TGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGCCGAATTTCAGCAGA

TGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGAACATTCCAGAGAGC
CCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGCTAAGGAAAGAAGAA
ATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATTCCAAAAATCTGAAT
CTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGAAAAGAGGTACTCCT
CCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAACAAGCCTTTTCAGCT
TTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTTTGAGGATAACGAGA
GCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCAGACCAGTAGGTCAT
CCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGTGGTTTCCTTGGTTG
GTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAACCACTGAAACTGAAATGAGAAAGA
GAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCA
GCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTATAAATTTTCCAACA
TATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGTGATGGACCCATTTG
TTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCCAATGACGGACCATT
TCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCTGAAAATTATTGCCA
TGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTG
GACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGC
CAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCA
TCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTG
ATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGT
GGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCATGATGGTCATGGTGA
TTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAACCTTGCAGCCACTG
ATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAA
TATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACTTGATGATCTAAACA
ACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAAAGATGTAAATGGAA
CTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACATGTCATTCATAAACA
ACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACACGGAAGACTTTAGTA
GTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAGCACTGTGGACATCG
GCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTG
TACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCTGAGAAGGACGTGTT
TCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTG
AAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTTCACTTACATTTTCA
TTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTGTTGGCTGGACTTCT
TAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGCCATCAAATCTCTCA
GGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGTGGTTGTGAATGCCCTTTTAG
GAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATT
TGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGACGTGAATAATCATA
CTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGAT
TTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAA
ATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCT
TCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACA
TCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATAC
CTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACATAAGCATCATGATTC
TCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTACCATTTTGTCACGCA
TCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCATTATTATTTTACCA
TTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGT
ATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAA
AGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAG
TCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCA
ACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCAC
CCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGA
ACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGG
TCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATG
AGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTG
AACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGA
TCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATAC
AGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTACTTTAAAACGAAAAC
AAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCT
TTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAATTGACAGAATAAATG
AAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGC
CAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATT
GGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACT
GACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTG
TGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGATAAGATGCACAATG

GCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGT
GTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCT
TTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTC
TATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGG
CATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAG
GGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTT
GTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGT
TGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTAC
TCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTA
ATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCAC
AGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTAC
CATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAG
AATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATT
AAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATATGGTATATAGCCAG
ACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTATAAGCAAATATAAAT
ACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTATCTTCAGGTTGATAT
CACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGAC
TGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAG
GTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTG
CACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTG
GAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAAC
TTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTAC
TTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAACCAACATTTAAAAT
GTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTGTTAACAGAATATTCTAATGAT
ATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACACACACACACACACACACACACACA
CACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATGTCACTGGAGGTAGGGGGTAATGACCTC
CCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGG
AAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGAAACCCTTAAATCATCATTTAAAA
ATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGT
GTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAG
AAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTG
GCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATACAACTGAAGCTACTTTCTCTGGG
GTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTC
TCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATGCCTCATTTCAGCATGCCATCTGT
CTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTT
GGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAACCACAGTGAAGTCCAGTGCAGAAT
TCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTAAATTAATCTATCTAAATTTCT
GAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCTCAACCAGAAAAGGGACAGTAA
TTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACATCACAATCCTAAAGCAACACAACA
ATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTCAGAATTATAGGATAAACCCTG
CAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAAATACAAAGAGAAGAACTGGTAAAACAT
AATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAG
AATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAG
CATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAATACATACTGATGTAAGTAAATAAG
GAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAAATAACGGAGTTAGAAAAATCACT
ATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAATGAGTATTAAAACTCATGGGTGAAAGTTTGA
TGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAATTTAAACAGAAGAAAGTATACT
TTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGACCAACTGATACTATGTGCCTCCTGATAA
GATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCAGAATCTAACTACAAGGAAAATCA
AATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTCATGAAAGACAAAGAAAGGCTAAA
GAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGAATATCAAATTAAAAAATAACAGATAAA
TAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTAGTATTATAATAGTGTTAATTTTC
CTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATTCTCTGAAAATGCATGCTAAAGTA
TTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAATGAGTTACATTTATATATATA
TGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTT
TTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGTATGCCCTGGACTTTATAATTTCT
ACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTAACTACTATTCTCATTTCCTAGCA
GTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAATAATCACTGTTTAATTAAAATAAGAAT
ATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAACAGATCTGATTTGGATAACAGAAG
GTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATTCCAAACACAGCTCCACAGAGCAG
CACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATAATGCCATAAATATTTATAAAGCT
GAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTTCTAAAGACTAATACCATTTAATG
TTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTTTAATATTTTATTTAGCCAAGTGG

AAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAATCCAAAAGTCTAGTAACCATTGGG
ACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCCAGACACTCGAACTATATAAACTT
GTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCACTGGAATTTCTTATTATATAAGAA
AATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTATTCTTCCTGGCACATGGAAATATCCCAC
TGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGAGGAAAATACTCTAGACAATCAAG
TATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATTTGCAATATCAATTGATTTTACTC
AACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGATAACTCAGCCCCTCACCTCAGCAACAA
GCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATATCTAAGTCCCTTTTCCCTTCCCT
TTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTT
CCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTT
CCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAGAGATCAATTTTTCCCATTCATTT
CTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTATTCTTTCCCTCTTGACTGATCCTT
TTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTTGCTCAACCCTGATGCATCTTTAT
TCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCTATACTCACTACCTTCACTCTTGG
TTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTGAATAATCTTGTTGATATAAACAC
TTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGGGAAGGGCTGGAAAATACAAAAGG
TAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACTCTGTCTCTGGTTATCATAAATACATCT
CCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGTTCACTGTGATGGGTTTTTTTGTT
TTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 11 >NM_001353948.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=5]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAAGATGGAGCAAACAG
TGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGAAAGACGCATTGCAG
AAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAATAGTGACTTGGAAG
CTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGAGGACCTGGACCCCT
ACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGCCACCTCTGCCCTGT
ACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATTCAGCATGCTAATTA
TGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAATGTAGAATACACCT
TCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGATTTTACTTTCCTTC
GGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGACCTGGGCAATGTCT
CGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAAAACCATTGTGGGAG
CCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGTATTTGCTCTAATTG
GGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTCCTTGGAGGAACATA
GTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTTTGACTGGAAGTCAT
ATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAATAGCTCTGATGCAG
GCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAGCTTTGATACCTTCA
GTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACTGACATTACGTGCTG
CTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAATTTGATCCTGGCTG
TGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGCCGAATTTCAGCAGA
TGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGAACATTCCAGAGAGC
CCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGCTAAGGAAAGAAGAA
ATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATTCCAAAAATCTGAAT
CTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGAAAAGAGGTACTCCT
CCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAACAAGCCTTTTCAGCT
TTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTTTGAGGATAACGAGA
GCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCAGACCAGTAGGTCAT
CCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGTGGTTTCCTTGGTTG
GTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGTGATAATAGATAAGCCAGCTACTGATGACA
ATGGAACAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATC
CTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGA
AATGCCCACCCTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAAC
ATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCA
TGGCCATGGAGCACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCT
TTACAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACG
GTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGAT
TGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGG
CTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAA

GCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCT
TCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCA
TGTGCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGA
GCTCATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGA
TGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTT
TAGATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGA
AAGATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTA
TTGATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTG
ACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCA
GTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTC
TTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAG
GAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCA
TGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGG
AATATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACAT
ATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGG
GTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTG
AAGGGATGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATAT
TCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTG
ACAGGTTTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGA
AAAATGTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGA
TGGATATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATC
TTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCA
ACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAA
AATTAGGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAA
CCAGACAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACC
AGAGTGAATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGA
AACTCATCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTG
TAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGA
TTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTG
CGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATG
TTAAGAGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTA
CAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTA
ACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATAT
CCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGC
CTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAAT
TTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTG
CCATGGATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAG
GAGAGAGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCT
ATCAGCCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCC
ACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTA
TAAAAGAAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAG
CTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCA
AAGGGAAATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATC
AACAGGACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAG
ACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTAC
CAGCTGACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTG
ATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTC
ACATTTTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAAC
GAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCT
CCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAA
ACAAATTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTG
CGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAAT
GTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCT
CCACCGCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTT
CGGCTCTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTG
CACGTTTGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCA
AGCTTTACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACC
ATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAA
ATTGTTCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTAC
ATTTTAAGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAA
TATGTGTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATA
AATATGAAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAA
AGAAATTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTT

TAGATTATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGT
TTAAAATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTT
TTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCAT
AATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAG
CAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTA
TTATGCTAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCC
AGTCACTTTTTGTTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGC
GCGCACACACACACACACACACACACACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTA
GATATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAAC
TGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGC
TATGAAAAGAAACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAG
AGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCT
ACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAA
TATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACT
GCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGC
ATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATT
CTTCTATATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTC
AGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAA
CTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAA
ATTAACAGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAG
GGAATATAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTC
TTCATAACATCACAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAA
CTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCT
TTCCAAATACAAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATA
TTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGG
GCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGA
ATCAGCAAATACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAAT
TGTAGAAGAAATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATC
AATGAGTATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTAT
TAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTA
TGAGACCAACTGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAAT
TTTAAAACCAGAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAA
TGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTG
GATGGAATATCAAATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTA
TTAGTTATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAA
TGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAG
TTCAAAGAAAAATGAGTTACATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGT
GTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTAT
CATATTTTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATAC
TGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCT
TCATAATAATCACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTT
GATGCTAAACAGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATA
AATAGTAATTCCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAA
ATTAAGAATAATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTT
AGATTATTTTCTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGG
TCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGG
TGCATGAAATCCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTAT
CACAAAGCCCAGACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTC
ATGTGTGCACTGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAG
TACTATTCTTCCTGGCACATGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGG
CACTAGAGGAGGAAAATACTCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAA
GTGGATAATTTGCAATATCAATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATT
TGCAGATAACTCAGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCA
GTGAAAAATATCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCT
TTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTT
CCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTAT
ATCTTTTCAGAGATCAATTTTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTT
CTATCTTTATTCTTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAA
AAATAACTTTGCTCAACCCTGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAG
CCAGGGCTCTATACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAA
TTTGACTGTGAATAATCTTGTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAA
CATTATGAGGGAAGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATA

CATACTCTGTCTCTGGTTATCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCA
ATTCAATTGTTCACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 12 >NM_001353960.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=14]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGTGTGGACATA
AGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAA
GATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGA
AAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAA
TAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGA
GGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGC
CACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATT
CAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAA
TGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGA
TTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGA
CCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAA
AACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGT
ATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTC
CTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTT
TGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAA
TAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAG
CTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACT
GACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAA
TTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGC
CGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGGCAGCAACGGCAACTGCCTCAGAACA
TTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGCTAA
GGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATTCCA
AAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGAAAA
GAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAACAAG
CCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTTTGA
GGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCAGAC
CAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGTGGT
TTCCTTGGGAACAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGA
AGATCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAG
GCAGAAATGCCCACCCTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGT
GAAACATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCT
TTTCATGGCCATGGAGCACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGG
GATCTTTACAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTT
TGACGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATT
TCGATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGT
GGGGGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGG
TAAAAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCA
CTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCA
AGCCATGTGCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCT
TCTGAGCTCATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGA
TAGGATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAA
GATTTTAGATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAAT
TGGGAAAGATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATA
CATTATTGATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGA
ATCTGACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGA
AAGCAGTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGA
AACTCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGG
CAGAGGAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGT
TTTCATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGAT
GTTGGAATATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCA
AACATATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGC
CTTGGGTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCG
ATTTGAAGGGATGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCT

TATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAAC
TGGTGACAGGTTTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCG
ATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGG
ATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACAT
GTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAA
TTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAAT
GAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTT
CGTAACCAGACAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGA
TGACCAGAGTGAATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGT
ACTGAAACTCATCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTC
CATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGC
TAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCT
TCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGC
CTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCA
AATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAA
AGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCAT
CATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGC
AGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCAT
GGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCT
CATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGT
TCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGT
CTCCTATCAGCCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAG
ACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCT
TCTTATAAAAGAAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCAC
TGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAA
AGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTT
TTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACA
ATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTC
ACTACCAGCTGACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAAC
CTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACAT
TTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTG
TAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATG
GCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAA
GAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTG
TTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATT
CAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTC
ACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCC
ATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTG
TGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGA
GTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATA
TTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCA
AGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATT
TTTACATTTTAAGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTA
TTAAATATGTGTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGT
ACATAAATATGAAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTC
ATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCT
AATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAA
AGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTT
CTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGC
TTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTAC
AGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGC
TACTATTATGCTAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTT
TGCCCAGTCACTTTTTGTTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAG
CGCGCGCGCACACACACACACACACACACACACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTA
GGGTAGATATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGA
AGAACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCT
TGAGCTATGAAAAGAAACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCC
TCCAGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCT
CTCCTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACAT
AGGAATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATC
AGACTGCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCT
GCTGCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATG
AAATTCTTCTATATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGA

AAGTCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAAC
AGTAACTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATA
GTAAAATTAACAGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAG
AAGAGGGAATATAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTT
GCCTCTTCATAACATCACAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATT
TCCAACTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGT
TACCTTTCCAAATACAAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCA
TGATATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAG
AAGGGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAA
TAAGAATCAGCAAATACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTA
ACAATTGTAGAAGAAATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAA
TCATCAATGAGTATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTC
TCTATTAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATC
AATTATGAGACCAACTGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTC
AAAATTTTAAAACCAGAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTT
AAAAATGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAA
TTCTGGATGGAATATCAAATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACT
GTTTATTAGTTATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATA
GAAAATGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGA
ATCAGTTCAAAGAAAAATGAGTTACATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTG
TGTGTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCAC
ACTATCATATTTTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATA
TATACTGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTT
GTTCTTCATAATAATCACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAG
TATTTGATGCTAAACAGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGG
GTATAAATAGTAATTCCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCC
AGAAAATTAAGAATAATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTAT
CTTTTAGATTATTTTCTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATA
ATGGGTCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTT
ACTGGTGCATGAAATCCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATC
TGTATCACAAAGCCCAGACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATAT
TCTTCATGTGTGCACTGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCT
GATAGTACTATTCTTCCTGGCACATGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGC
AAAGGCACTAGAGGAGGAAAATACTCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAAT
CACAAGTGGATAATTTGCAATATCAATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTC
ATATTTGCAGATAACTCAGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCT
ATCCAGTGAAAAATATCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTC
TTCCTTTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTA
CCTTTCCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACT
TTTATATCTTTTCAGAGATCAATTTTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTT
CTCTTCTATCTTTATTCTTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATC
TTAAAAAATAACTTTGCTCAACCCTGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTC
TAAAGCCAGGGCTCTATACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGG
GAAAATTTGACTGTGAATAATCTTGTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCC
AGTAACATTATGAGGGAAGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAG
CTATACATACTCTGTCTCTGGTTATCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACG
GAGCAATTCAATTGTTCACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 13 >NM_001353954.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=10]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGTGTGGACATA
AGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAA
GATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGA
AAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAA
TAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGA
GGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGC
CACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATT
CAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAA

TGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGA
TTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGA
CCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAA
AACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGT
ATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTC
CTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTT
TGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAA
TAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAG
CTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACT
GACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAA
TTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGC
CGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGGCAGCAACGGCAACTGCCTCAGAACA
TTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGCTAA
GGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATTCCA
AAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGAAAA
GAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAACAAG
CCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTTTGA
GGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCAGAC
CAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGTGGT
TTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAACCACTGAAACTGA
AATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGCAAT
GAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTATAA
ATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGTGAT
GGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCCAAT
GACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCTGAA
AATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAGCCT
GGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTTGGC
AAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCTCGT
CTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTGCAA
GATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGTGCT
GTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCATGAT
GGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAACCT
TGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTATGT
GAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACTTGA
TGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAAAGA
TGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACATGTC
ATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACACGGA
AGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAGCAC
TGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTTCAC
TGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCTGAG
AAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGGTGC
TCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTTCAC
TTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTGTTG
GCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGCCAT
CAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGTGGTTGTGAA
TGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCATCAT
GGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGACGT
GAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTTTGA
TAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGCAGT
TGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCATCAT
CTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTTTGG
AGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACCGCA
AAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACATAAG
CATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTACCAT
TTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCATTA
TTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGAGCT
GATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCTGAT
CAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCT
ACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGATCGA
TGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGATGG
ATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAAGGG
AGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACAT
GTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTTTGA

GATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTTTGC
AGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGTGAG
TGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGATGC
TCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTACTTT
AAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGTAAA
ACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAATTGA
CAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGACCG
GGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAATAAA
TAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAGGTC
AATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGCAAA
CTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGATAA
GATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACATGA
AACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGTTAG
TGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAATAGG
CTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGTGGT
TTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCACCAT
ATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGACCAA
GTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTTATG
TTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAATTG
ACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGACAC
TGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAATAAT
GACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAAAAT
GTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTTCTC
TCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCATTTA
TTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATATGG
TATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTATAAG
CAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTATCTT
CAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGACAT
GAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTTAAG
GAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATTTCT
GGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATGTTT
TCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAAACA
GAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTTTTC
CTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAACCA
ACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTGTTAACAGAAT
ATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACACACACACACACAC
ACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATGTCACTGGAGGTAGGGG
GTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGGATAGCTAGTG
ATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGAAACCCTTAAATC
ATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTCTGCTTCTGCA
TCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTCCCAATTGGTT
TGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCTCTCCAATTCCTTCT
TTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATACAACTGAAGCTA
CTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCTTTTGTATTCC
TTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATGCCTCATTTCAGC
ATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAAAGATGTGCCT
TGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAACCACAGTGAAGTC
CAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTAAATTAATCTA
TCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCTCAACCAGAAA
AGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACATCACAATCCTAAA
GCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTCAGAATTATAG
GATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAAATACAAAGAGAAGAAC
TGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGCTAAAAAGTAATT
ATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACTGGCCAAATCTGG
GGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAATACATACTGATGT
AAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAAATAACGGAGTTA
GAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAATGAGTATTAAAACTCATGGG
TGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAATTTAAACAGAA
GAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGACCAACTGATACTATGTG
CCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCAGAATCTAACTAC
AAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTCATGAAAGACAAA
GAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGAATATCAAATTAAAAAA
TAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTAGTATTATAATAG

TGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATTCTCTGAAAATGC
ATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAATGAGTTACAT
TTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAGGGGAGGATTT
TTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGTATGCCCTGGACT
TTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTAACTACTATTCTC
ATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAATAATCACTGTTTAATT
AAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAACAGATCTGATTTG
GATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATTCCAAACACAGCT
CCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATAATGCCATAAATA
TTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTTCTAAAGACTAAT
ACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTTTAATATTTTATT
TAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAATCCAAAAGTCTAG
TAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCCAGACACTCGAAC
TATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCACTGGAATTTCTTA
TTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTATTCTTCCTGGCACATGG
AAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGAGGAAAATACTCT
AGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATTTGCAATATCAAT
TGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGATAACTCAGCCCCTCAC
CTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATATCTAAGTCCCTT
TTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTTCCT
TCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCCCTTCCTCCCTCT
CTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAGAGATCAATTTTT
CCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTATTCTTTCCCTCTT
GACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTTGCTCAACCCTGA
TGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCTATACTCACTACC
TTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTGAATAATCTTGTT
GATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGGGAAGGGCTGGAA
AATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACTCTGTCTCTGGTTATCA
TAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGTTCACTGTGATGG
GTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 14 >NM_001165964.3 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=3]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGTGTGGACATA
AGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAA
GATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGA
AAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAA
TAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGA
GGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGC
CACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATT
CAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAA
TGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGA
TTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGA
CCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAA
AACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGT
ATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTC
CTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTT
TGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAA
TAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAG
CTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACT
GACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAA
TTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGC
CGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGA
ACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGC
TAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATT
CCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGA
AAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAAC
AAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTT

TGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCA
GACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGT
GGTTTCCTTGGGAACAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCT
AGAAGATCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATC
CAGGCAGAAATGCCCACCCTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAA
AGTGAAACATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATAC
TCTTTTCATGGCCATGGAGCACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCAC
TGGGATCTTTACAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATAT
CTTTGACGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTC
ATTTCGATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTC
CGTGGGGGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTT
TGGTAAAAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTT
CCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGG
TCAAGCCATGTGCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTT
GCTTCTGAGCTCATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGT
GGATAGGATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACA
AAAGATTTTAGATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGA
AATTGGGAAAGATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAA
ATACATTATTGATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGG
AGAATCTGACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAA
TGAAAGCAGTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGA
AGAAACTCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGA
AGGCAGAGGAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCAT
TGTTTTCATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGAC
GATGTTGGAATATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTA
TCAAACATATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAA
TGCCTTGGGTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATC
TCGATTTGAAGGGATGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTG
TCTTATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCAC
AACTGGTGACAGGTTTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGC
TCGATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAA
AGGATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTA
CATGTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGA
TAATTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGC
AATGAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGA
CTTCGTAACCAGACAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAAC
AGATGACCAGAGTGAATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTG
TGTACTGAAACTCATCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCT
CTCCATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCT
TGCTAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTC
CCTTCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTT
TGCCTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATT
CCAAATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAA
TAAAGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACAT
CATCATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAG
TGCAGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTT
CATGGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCA
GCTCATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCG
GGTTCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAA
GGTCTCCTATCAGCCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTA
CAGACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAA
TCTTCTTATAAAAGAAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTC
CACTGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGA
AAAAGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTA
TTTTTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCT
ACAATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTT
CTCACTACCAGCTGACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCA
AACCTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCA
CATTTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTT
TTGTAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCAC
ATGGCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCA
CAAGAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATC
TTGTTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTC

ATTCAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCC
CTCACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTA
GCCATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACC
CTGTGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGT
AGAGTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTT
ATATTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAG
TCAAGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTT
ATTTTTACATTTTAAGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAAC
CTATTAAATATGTGTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTA
TGTACATAAATATGAAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTT
TTCATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTT
GCTAATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGG
AAAAGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTT
GTTCTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTG
GGCTTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCT
TACAGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAA
TGCTACTATTATGCTAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATT
GTTTGCCCAGTCACTTTTTGTTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGG
GAGCGCGCGCGCACACACACACACACACACACACACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAAC
CTAGGGTAGATATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGA
GGAAGAACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACC
CCTTGAGCTATGAAAAGAAACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGC
TCCTCCAGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGG
GCTCTCCTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGA
CATAGGAATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCA
ATCAGACTGCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCT
GCTGCTGCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTC
ATGAAATTCTTCTATATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCA
TGAAAGTCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAA
AACAGTAACTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAG
ATAGTAAAATTAACAGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGC
CAGAAGAGGGAATATAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGAC
TTTGCCTCTTCATAACATCACAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCC
ATTTCCAACTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCAC
TGTTACCTTTCCAAATACAAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTG
GCATGATATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAAC
TAGAAGGGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATT
GAATAAGAATCAGCAAATACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAA
TTAACAATTGTAGAAGAAATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAA
GAATCATCAATGAGTATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTC
TTCTCTATTAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATC
ATCAATTATGAGACCAACTGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACA
GTCAAAATTTTAAAACCAGAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTC
CTTAAAAATGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACA
CAATTCTGGATGGAATATCAAATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGG
ACTGTTTATTAGTTATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAA
ATAGAAAATGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCT
TGAATCAGTTCAAAGAAAAATGAGTTACATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGAT
GTGTGTGTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTT
CACACTATCATATTTTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCA
ATATATACTGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTC
CTTGTTCTTCATAATAATCACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCC
CAGTATTTGATGCTAAACAGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTG
TGGGTATAAATAGTAATTCCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAA
GCCAGAAAATTAAGAATAATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTG
TATCTTTTAGATTATTTTCTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTG
ATAATGGGTCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTAT
TTTACTGGTGCATGAAATCCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTAT
ATCTGTATCACAAAGCCCAGACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCA
TATTCTTCATGTGTGCACTGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACT
GCTGATAGTACTATTCTTCCTGGCACATGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGC
AGCAAAGGCACTAGAGGAGGAAAATACTCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAA
AATCACAAGTGGATAATTTGCAATATCAATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTC

CTCATATTTGCAGATAACTCAGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTT
CCTATCCAGTGAAAAATATCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATT
CTCTTCCTTTTTCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTC
CTACCTTTCCCTTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCAT
ACTTTTATATCTTTTCAGAGATCAATTTTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTA
CTTCTCTTCTATCTTTATTCTTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTA
ATCTTAAAAAATAACTTTGCTCAACCCTGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGT
TTCTAAAGCCAGGGCTCTATACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTA
TGGGAAAATTTGACTGTGAATAATCTTGTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTT
CCCAGTAACATTATGAGGGAAGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCC
AAGCTATACATACTCTGTCTCTGGTTATCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCC
ACGGAGCAATTCAATTGTTCACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 15 >NM_001353952.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=9]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGTGTGGACATA
AGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAA
GATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGA
AAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAA
TAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGA
GGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGC
CACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATT
CAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAA
TGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGA
TTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGA
CCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAA
AACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGT
ATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTC
CTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTT
TGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAA
TAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAG
CTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACT
GACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAA
TTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGC
CGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGA
ACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGC
TAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATT
CCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGA
AAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAAC
AAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTT
TGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCA
GACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGT
GGTTTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAACCACTGAAAC
TGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGC
AATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTA
TAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGT
GATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCC
AATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCT
GAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAG
CCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTT
GGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCT
CGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTG
CAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGT
GCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCAT
GATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAA
CCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTA
TGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACT
TGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAA

AGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACAT
GTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACAC
GGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAG
CACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTT
CACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCT
GAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGG
TGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTT
CACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTG
TTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGC
CATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGTGGTTGT
GAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCAT
CATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGA
CGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTT
TGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGC
AGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCAT
CATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTT
TGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACC
GCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACAT
AAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTAC
CATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCA
TTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGA
GCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCT
GATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCT
CCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGAT
CGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGA
TGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAA
GGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAA
CATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTT
TGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTT
TGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGT
GAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGA
TGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTAC
TTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGT
AAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAAT
TGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGA
CCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAAT
AAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAG
GTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGC
AAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGA
TAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACA
TGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGT
TAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAAT
AGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGT
GGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCAC
CATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGAC
CAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTT
ATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAA
TTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGA
CACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAAT
AATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAA
AATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTT
CTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCAT
TTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATA
TGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTAT
AAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTAT
CTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGA
CATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTT
AAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATT
TCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATG
TTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAA
ACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTT
TTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAA
CCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTGTTAACAG

AATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACACACACACACA
CACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATGTCACTGGAGGTAG
GGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGGATAGCTA
GTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGAAACCCTTAA
ATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTCTGCTTCT
GCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTCCCAATTG
GTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCTCCAATTCCT
TCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATACAACTGAAG
CTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCTTTTGTAT
TCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATGCCTCATTTC
AGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAAAGATGTG
CCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAACCACAGTGAA
GTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTAAATTAAT
CTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCTCAACCAG
AAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACATCACAATCCT
AAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTCAGAATTA
TAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAAATACAAAGAGAAG
AACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGCTAAAAAGTA
ATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACTGGCCAAATC
TGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAATACATACTGA
TGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAAATAACGGAG
TTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAATGAGTATTAAAACTCAT
GGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAATTTAAACA
GAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGACCAACTGATACTAT
GTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCAGAATCTAAC
TACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTCATGAAAGAC
AAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGAATATCAAATTAAA
AAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTAGTATTATAA
TAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATTCTCTGAAAA
TGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAATGAGTTA
CATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAGGGGAGGA
TTTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGTATGCCCTGG
ACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTAACTACTATT
CTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAATAATCACTGTTTA
ATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAACAGATCTGAT
TTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATTCCAAACACA
GCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATAATGCCATAA
ATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTTCTAAAGACT
AATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTTTAATATTTT
ATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAATCCAAAAGTC
TAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCCAGACACTCG
AACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCACTGGAATTTC
TTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTATTCTTCCTGGCACA
TGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGAGGAAAATAC
TCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATTTGCAATATC
AATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGATAACTCAGCCCCT
CACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATATCTAAGTCC
CTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTT
CCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCCCTTCCTCCC
TCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAGAGATCAATT
TTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTATTCTTTCCCT
CTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTTGCTCAACCC
TGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCTATACTCACT
ACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTGAATAATCTT
GTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGGGAAGGGCTG
GAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACTCTGTCTCTGGTTA
TCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGTTCACTGTGA
TGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 16 >NM_001353961.2 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=15]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGTGTGGACATA
AGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCAGGATGACAA
GATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGCGGCTATTGA
AAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCCAAAGCCAAA
TAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGAGCCCCTGGA
GGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCGGTTCAGTGC
CACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACATTCATTATT
CAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTGGACAAAGAA
TGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTGTTTAGAAGA
TTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGAGTTTGTGGA
CCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCCAGGCCTGAA
AACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTGTCTGAGCGT
ATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCACCAATGCTTC
CTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGTCTTTGAGTT
TGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACTATGTGGAAA
TAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGGCTACACAAG
CTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCTTTATCAACT
GACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTACCTAATAAA
TTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACAGAAAGAGGC
CGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAACTGCCTCAGA
ACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTCCAAGAGTGC
TAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGAGGATGAATT
CCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATTGACATATGA
AAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAATAGCAGAAC
AAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCACAGCACCTT
TGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAACCTGAGTCA
GACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTGCAATGGTGT
GGTTTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAACCACTGAAAC
TGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGC
AATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTA
TAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGT
GATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCC
AATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTAAGCATATTGGAAGGTTTTCACTGGGATCTTTA
CAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTT
TTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGC
TGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTC
TGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCT
ACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCC
TGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGT
GCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCT
CATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGC
ACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAG
ATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAG
ATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTG
ATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACT
TTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTA
GCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTG
AACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAA
AACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGA
TTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAAT
ATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATT
TCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTT
ACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAG
GGATGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCT
GGCTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACA
GGTTTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAA
ATGTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGG

ATATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTT
ACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACC
AGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAAT
TAGGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCA
GACAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGA
GTGAATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAAC
TCATCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAG
GTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTG
GCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGT
TGTTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTA
AGAGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAA
CCTCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACC
CTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCT
TCCTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTC
TGAGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTG
AAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCA
TGGATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAG
AGAGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATC
AGCCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACC
TTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAA
AAGAAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTT
GTCCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAG
GGAAATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAAC
AGGACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACA
GTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAG
CTGACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATT
TTGGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACA
TTTTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAA
GTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCC
AGCTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACA
AATTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGA
AATCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTT
TGGAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCA
CCGCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGG
CTCTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCAC
GTTTGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGC
TTTACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATG
CTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATT
GTTCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATT
TTAAGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATAT
GTGTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAAT
ATGAAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGA
AATTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAG
ATTATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTA
AAATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTT
CTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAAT
TTTTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAA
ACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTA
TGCTAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGT
CACTTTTTGTTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCG
CACACACACACACACACACACACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGAT
ATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGG
TGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTAT
GAAAAGAAACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGC
CACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACC
TTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATAT
TTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCC
CCTCATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATT
GAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTT
CTATATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGC
AGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTG
GGGGAACCACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATT
AACAGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGA

ATATAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTC
ATAACATCACAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTA
AGGTAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTC
CAAATACAAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTT
TCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCT
TCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATC
AGCAAATACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGT
AGAAGAAATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAAT
GAGTATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAA
GTAGAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGA
GACCAACTGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTT
AAAACCAGAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGT
CAATGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGAT
GGAATATCAAATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTA
GTTATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGT
TCTTATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTC
AAAGAAAAATGAGTTACATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTC
TTTAGAAAGGGGAGGATTTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCAT
ATTTTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGG
TTTTTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCA
TAATAATCACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGAT
GCTAAACAGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAAT
AGTAATTCCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATT
AAGAATAATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGA
TTATTTTCTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCA
TTGGTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGC
ATGAAATCCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCAC
AAAGCCCAGACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATG
TGTGCACTGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTAC
TATTCTTCCTGGCACATGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCAC
TAGAGGAGGAAAATACTCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTG
GATAATTTGCAATATCAATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGC
AGATAACTCAGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTG
AAAAATATCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTT
TCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCC
TTTCTCCCTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATC
TTTTCAGAGATCAATTTTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTA
TCTTTATTCTTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAA
TAACTTTGCTCAACCCTGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCA
GGGCTCTATACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTT
GACTGTGAATAATCTTGTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACAT
TATGAGGGAAGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACAT
ACTCTGTCTCTGGTTATCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATT
CAATTGTTCACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 17>NM_006920.6 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=2]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGAACAGCATAAGAATTATTTCTGAGTGGAGGTGAGGCTTGTCCAAATGTCTTTGCTATCATGG
ATTTCCTGACTCCTACCTGTTTGAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGT
GTGGACATAAGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCA
GGATGACAAGATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGC
GGCTATTGAAAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCC
AAAGCCAAATAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGA
GCCCCTGGAGGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCG
GTTCAGTGCCACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACA
TTCATTATTCAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTG
GACAAAGAATGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTG
TTTAGAAGATTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGA

GTTTGTGGACCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCC
AGGCCTGAAAACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTG
TCTGAGCGTATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCAC
CAATGCTTCCTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGT
CTTTGAGTTTGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACT
ATGTGGAAATAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGG
CTACACAAGCTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCT
TTATCAACTGACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTA
CCTAATAAATTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACA
GAAAGAGGCCGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAAC
TGCCTCAGAACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTC
CAAGAGTGCTAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGA
GGATGAATTCCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATT
GACATATGAAAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAA
TAGCAGAACAAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCA
CAGCACCTTTGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAA
CCTGAGTCAGACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTG
CAATGGTGTGGTTTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGGAACAAC
CACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGTTTCCATGGACTTTCTAGAAGATCCTTCCCAAAG
GCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGAAGAACTTGAAGAATCCAGGCAGAAATGCCCACC
CTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTCTCCATATTGGTTAAAAGTGAAACATGTTGTCAA
CCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTGTATTGTCTTAAATACTCTTTTCATGGCCATGGA
GCACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGGAAACTTGGTTTTCACTGGGATCTTTACAGCAGA
AATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCAAGAAGGCTGGAATATCTTTGACGGTTTTATTGT
GACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATTATCTGTTCTCCGTTCATTTCGATTGCTGCGAGT
TTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAAGATCATCGGCAATTCCGTGGGGGCTCTGGGAAA
TTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGTCGGCATGCAGCTCTTTGGTAAAAGCTACAAAGA
TTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCACATGAATGACTTCTTCCACTCCTTCCTGATTGT
GTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTGTATGGAGGTTGCTGGTCAAGCCATGTGCCTTAC
TGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAATCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAG
TGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAATCTCCAAATTGCTGTGGATAGGATGCACAAAGG
AGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTCCTTCATTAGGAAACAAAAGATTTTAGATGAAAT
TAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTCCAATCATACAGCAGAAATTGGGAAAGATCTTGA
CTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGGCAGCAGTGTTGAAAAATACATTATTGATGAAAG
TGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGTACCAATTGCTGTAGGAGAATCTGACTTTGAAAA
TTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAGCAAAGAGAAACTGAATGAAAGCAGTAGCTCATC
AGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCCCGTAGTGGAACCTGAAGAAACTCTTGAACCAGA
AGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCAAATCAATGTGGAAGAAGGCAGAGGAAAACAATG
GTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTGGTTTGAGACCTTCATTGTTTTCATGATTCTCCT
TAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCGAAAGACGATTAAGACGATGTTGGAATATGCTGA
CAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATGGGTGGCATATGGCTATCAAACATATTTCACCAA
TGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGTCAGTTTAACAGCAAATGCCTTGGGTTACTCAGA
ACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACCTCTAAGAGCCTTATCTCGATTTGAAGGGATGAG
GGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAATGTGCTTCTGGTTTGTCTTATATTCTGGCTAAT
TTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCACTGTATTAACACCACAACTGGTGACAGGTTTGA
CATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGAAAGAAATGAGACTGCTCGATGGAAAAATGTGAA
AGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCAAGTTGCCACATTCAAAGGATGGATGGATATAAT
GTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTATGAAGAAAGTCTGTACATGTATCTTTACTTTGT
TATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTATTGGTGTCATCATAGATAATTTCAACCAGCAGAA
AAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAAGAAATACTATAATGCAATGAAAAAATTAGGATC
GAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCAAGGAATGGTCTTTGACTTCGTAACCAGACAAGT
TTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCACAATGATGGTGGAAACAGATGACCAGAGTGAATA
TGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCTATTTACTGGAGAGTGTGTACTGAAACTCATCTC
TCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTTTGTGGTTGTCATTCTCTCCATTGTAGGTATGTT
TCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTTCCGAGTGATCCGTCTTGCTAGGATTGGCCGAAT
CCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTTTGCTTTGATGATGTCCCTTCCTGCGTTGTTTAA
CATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTTTGGGATGTCCAACTTTGCCTATGTTAAGAGGGA
AGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAGCATGATCTGCCTATTCCAAATTACAACCTCTGC
TGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACCCGACTGTGACCCTAATAAAGTTAACCCTGGAAG
CTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTTTTTTGTCAGTTACATCATCATATCCTTCCTGGT
TGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGTTGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGA
GGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCCCGATGCAACTCAGTTCATGGAATTTGAAAAATT
ATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACAACCAAACAAACTCCAGCTCATTGCCATGGATTT

GCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATTTGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGG
AGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGCTTCCAATCCTTCCAAGGTCTCCTATCAGCCAAT
CACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCATTATTCAGCGTGCTTACAGACGCCACCTTTTAAA
GCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAATCAAAGGTGGGGCTAATCTTCTTATAAAAGAAGA
CATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAACTGATCTGACCATGTCCACTGCAGCTTGTCCACC
TTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCAAGAAGGCAAAGATGAAAAAGCCAAAGGGAAATA
AATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCCTGTGAAGGTGATGTATTTTTATCAACAGGACTC
CTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCCTTAAGGTCAGTGCCTACAATAAGACAGTGACCC
CTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGACAGGAGGTTACTGTTCTCACTACCAGCTGACAC
TGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCAGTTTCAAGGGGTGCAAACCTGTGATTTTGGGGT
TGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTTGCATTTCAACTGCCACATTTGTCACATTTTTAT
GGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTATTATATGTGACTATTTTTGTAAACGAAGTTTCTG
TTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGGGGGTATGGCAACCACATGGCCCTCCCAGCTACA
CAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATGCATGAGAAAAAGTCACAAGAAAAACAAATTCTT
AAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGAGGTGCTTTGTTGATCTTGTTTTGCGAAATCCAG
CCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTAGCAGGTGCAAACTTCATTCAAATGTTTGGAGTC
ATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAATAGTGAATATTGCCCCTCACCCTCCACCGCCAG
AAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTTCCTGCACTTTGTTTAGCCATCTTCGGCTCTCAG
CAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTAAATAGTCATTTTACCCTGTGGTGCACGTTTGAG
CAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATGTACCACAAGAAATGTAGAGTGCAAGCTTTACAC
AGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGCTATCAATGCATGTTTATATTACCATGCTGCTGT
ATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCATATGTCAGTGGTAAAGTCAAGGAAATTGTTCAAC
AGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTTAACAGCTTTTTGGTTATTTTTACATTTTAAGTG
GATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAAAACACACTGCTTAACCTATTAAATATGTGTTTA
GAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTTTATTTTTCAGCATTATGTACATAAATATGAAGA
GGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTCTGTCCATAGTACTTTTTCATGAAAGAAATTTGC
TAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCTTTTTAAATTACATTTGCTAATTTTAGATTATTT
CACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATTATGCTTTGCAATTGGAAAAGGGTTTAAAATTTT
ATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGTGCTTATGTTATTTTTGTTCTTTTTTTCTGACTT
CGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGTTCTAAATAGAATGTGGGCTTCATAATTTTTTTT
TCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGGACATTTTGTGTTTCTTACAGAAGCAAACCATAG
GCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAACTGCATGCTGGAAAATGCTACTATTATGCTAAA
TAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTACATTTTTTATTTTATTGTTTGCCCAGTCACTTTT
TGTTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAAATTGGGGGAGAAGGGGAGCGCGCGCGCACACAC
ACACACACACACACACACACACACACACACACACACAGAGGCATACCCACGTTGACAACAAAACCTAGGGTAGATATGTCAC
TGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTGTGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGA
GGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAGTGATGGTGGATGACCCCTTGAGCTATGAAAAGA
AACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTAGGATACTGTGTTTGCTCCTCCAGAGCCACTCTC
TCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTACACATTGCAAAAATGGGCTCTCCTACCTTTGAGC
TCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCAGAGAGAGAAGATTGACATAGGAATATTTCTTCT
CCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCGAAAAGTCATACTCCAATCAGACTGCCCCTCATA
CAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCTTGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTG
CTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTTGTAAATGTCCCCTTCATGAAATTCTTCTATATG
CCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGGTGATTCAACAGCTCATGAAAGTCAGCAGGAAGC
AAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTTTTGCAGATAAAGAAAAACAGTAACTGGGGGAAC
CACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTTACTCGGGTGGTCCAGATAGTAAAATTAACAGTC
TAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTTGTCATCAGGTGAAGCCAGAAGAGGGAATATAGC
CTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTTGTACAATGGACTGACTTTGCCTCTTCATAACAT
CACAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAGCTGAAAAAAATTCCCATTTCCAACTAAGGTAGG
TCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATCCACGCCATTACTCACTGTTACCTTTCCAAATAC
AAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTCATTTTGAAATATTTGGCATGATATTTTCTGTGC
TAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAAGACACAGGAACCAACTAGAAGGGGCTTCCCACT
GGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAAAGATTATAATTCATTGAATAAGAATCAGCAAAT
ACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTTGCAGTTGAATGTTAATTAACAATTGTAGAAGAA
ATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGACAATTGATTCATACAAGAATCATCAATGAGTATT
AAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGCATCTTAAAGTATCTCTTCTCTATTAAGTAGAAA
ATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAATACCTTCAAAGATATCATCAATTATGAGACCAAC
TGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATTACTTCTTGGTACACAGTCAAAATTTTAAAACCA
GAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAATAAGTGGACTGAACTCCTTAAAAATGTCAATGTC
ATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAACTATGACAACTAAACACAATTCTGGATGGAATAT
CAAATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAATAAATTTGAATATGGACTGTTTATTAGTTATTA
GTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGTGCCTGTTCCATGAAAATAGAAAATGTTCTTATT
CTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATGTCTGCAGCTCATTCTTGAATCAGTTCAAAGAAA

AATGAGTTACATTTATATATATATGTATGTATGTAAACAGACATAGATAAAAGTATAGATGTGTGTGTGTCTTTAGAA
AGGGGAGGATTTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGCCTATGTCTGCGTGTTCACACTATCATATTTTGT
ATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTGTTGGTAATTAATTCAATATATACTGGTTTTTTA
ACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACAGAAATGATTACATTCCTTGTTCTTCATAATAAT
CACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCATAGTGGTTAAGACCCCAGTATTTGATGCTAAAC
AGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCTGGGGACCAGACACTGTGGGTATAAATAGTAATT
CCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCATATGTCTATAGTTAAGCCAGAAAATTAAGAATA
ATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACCTTTATCCTGAAATTGTATCTTTTAGATTATTTT
CTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGATGGTGAGAATACGTGATAATGGGTCATTGGTTT
TAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGGTCCTCCCATTTGTATTTTACTGGTGCATGAAAT
CCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAGTTTGTATGTTTGTATATCTGTATCACAAAGCCC
AGACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATATGTCTGGTAATTGCATATTCTTCATGTGTGCAC
TGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCACCATGAATTGGGTACTGCTGATAGTACTATTCTT
CCTGGCACATGGAAATATCCCACTGAGGTTGTCAATCAATATTGATTTAATCAGTGTGGCAGCAAAGGCACTAGAGGA
GGAAAATACTCTAGACAATCAAGTATAGATGGGAAAATTGCATCACCAAGAAATTCTTAAAATCACAAGTGGATAATT
TGCAATATCAATTGATTTTACTCAACTGCTTGAGTCCTTTATCATTCCGTTAGAACACTCCTCATATTTGCAGATAAC
TCAGCCCCTCACCTCAGCAACAAGCAAGAGGACATCATTATAAACTTCCTCAATTTCTTTCCTATCCAGTGAAAAATA
TCTAAGTCCCTTTTCCCTTCCCTTTTCCTTCCCTTCCCTTCCTTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCC
TTTTTCTTTCCTTCCATTCTCTTCCTTTTTCCTTCCTTTTTCTTACTTTACCTCCCCCTCCTACCTTTCCCTTTCTCC
CTTCCTCCCTCTCTCCCTTCCTTCCTCTCTCCCTCTCCTTCCTTCCTTCCTTCAATCCATACTTTTATATCTTTTCAG
AGATCAATTTTTCCCATTCATTTCTACATTCTCAGATACCTTGCTCCATGACTTGTTCTACTTCTCTTCTATCTTTAT
TCTTTCCCTCTTGACTGATCCTTTTTGTCTGCCTTTGCATTTTCAAAATCAGTTTTTCTAATCTTAAAAAATAACTTT
GCTCAACCCTGATGCATCTTTATTCACTGTGTCATTATCTCATTATTTTTAAAGGTTAGTTTCTAAAGCCAGGGCTCT
ATACTCACTACCTTCACTCTTGGTTAAAATTAAAAGAAATATAGGTCCCAGTGATATTTATGGGAAAATTTGACTGTG
AATAATCTTGTTGATATAAACACTTTGGGTTTTCAAAATGTTATGCTAGCTGTTAACTTTCCCAGTAACATTATGAGG
GAAGGGCTGGAAAATACAAAAGGTAGACATAGGAACCGAGAACACCACCCCTATGGAGCCAAGCTATACATACTCTGT
CTCTGGTTATCATAAATACATCTCCTATATTCAATGCTTGCTTAGAAAGGCATGATGTCCACGGAGCAATTCAATTGT
TCACTGTGATGGGTTTTTTTGTTTTCTCCTCATGGAATGCAACTTTTATTTAT

SEQ ID NO: 18 >NM_001165963.4 SCN1A [organism=Homo sapiens] [GeneID=6323] [transcript=1]

ACCATAGAGTGAGGCGAGGATGAAGCCGAGAGGATACTGCAGAGGTCTCTGGTGCATGTGTGTATGTGTGCGTTTGTG
TGTGTTTGTGTGTCTGTGTGTTCTGCCCCAGTGAGACTGCAGCCCTTGTAAATACTTTGACACCTTTTGCAAGAAGGA
ATCTGAACAATTGCAACTGAAGGCACATTGTTATCATCTCGTCTTTGGGTGATGCTGTTCCTCACTGCAGATGGATAA
TTTTCCTTTTAATCAGAACAGCATAAGAATTATTTCTGAGTGGAGGTGAGGCTTGTCCAAATGTCTTTGCTATCATGG
ATTTCCTGACTCCTACCTGTTTGAGGTTTGGGCAATTATGAATAAGGCTGCTGTATACATCCGTGTGCAGGATTTTGT
GTGGACATAAGTTTTCAACTCCTTTGGTTAAATCCTAAGGAATTTCATATGCAGAATAAATGGTAATTAAAATGTGCA
GGATGACAAGATGGAGCAAACAGTGCTTGTACCACCAGGACCTGACAGCTTCAACTTCTTCACCAGAGAATCTCTTGC
GGCTATTGAAAGACGCATTGCAGAAGAAAAGGCAAAGAATCCCAAACCAGACAAAAAAGATGACGACGAAAATGGCCC
AAAGCCAAATAGTGACTTGGAAGCTGGAAAGAACCTTCCATTTATTTATGGAGACATTCCTCCAGAGATGGTGTCAGA
GCCCCTGGAGGACCTGGACCCCTACTATATCAATAAGAAAACTTTTATAGTATTGAATAAAGGGAAGGCCATCTTCCG
GTTCAGTGCCACCTCTGCCCTGTACATTTTAACTCCCTTCAATCCTCTTAGGAAAATAGCTATTAAGATTTTGGTACA
TTCATTATTCAGCATGCTAATTATGTGCACTATTTTGACAAACTGTGTGTTTATGACAATGAGTAACCCTCCTGATTG
GACAAAGAATGTAGAATACACCTTCACAGGAATATATACTTTTGAATCACTTATAAAAATTATTGCAAGGGGATTCTG
TTTAGAAGATTTTACTTTCCTTCGGGATCCATGGAACTGGCTCGATTTCACTGTCATTACATTTGCGTACGTCACAGA
GTTTGTGGACCTGGGCAATGTCTCGGCATTGAGAACATTCAGAGTTCTCCGAGCATTGAAGACGATTTCAGTCATTCC
AGGCCTGAAAACCATTGTGGGAGCCCTGATCCAGTCTGTGAAGAAGCTCTCAGATGTAATGATCCTGACTGTGTTCTG
TCTGAGCGTATTTGCTCTAATTGGGCTGCAGCTGTTCATGGGCAACCTGAGGAATAAATGTATACAATGGCCTCCCAC
CAATGCTTCCTTGGAGGAACATAGTATAGAAAAGAATATAACTGTGAATTATAATGGTACACTTATAAATGAAACTGT
CTTTGAGTTTGACTGGAAGTCATATATTCAAGATTCAAGATATCATTATTTCCTGGAGGGTTTTTTAGATGCACTACT
ATGTGGAAATAGCTCTGATGCAGGCCAATGTCCAGAGGGATATATGTGTGTGAAAGCTGGTAGAAATCCCAATTATGG
CTACACAAGCTTTGATACCTTCAGTTGGGCTTTTTTGTCCTTGTTTCGACTAATGACTCAGGACTTCTGGGAAAATCT
TTATCAACTGACATTACGTGCTGCTGGGAAAACGTACATGATATTTTTTGTATTGGTCATTTTCTTGGGCTCATTCTA
CCTAATAAATTTGATCCTGGCTGTGGTGGCCATGGCCTACGAGGAACAGAATCAGGCCACCTTGGAAGAAGCAGAACA
GAAAGAGGCCGAATTTCAGCAGATGATTGAACAGCTTAAAAAGCAACAGGAGGCAGCTCAGCAGGCAGCAACGGCAAC
TGCCTCAGAACATTCCAGAGAGCCCAGTGCAGCAGGCAGGCTCTCAGACAGCTCATCTGAAGCCTCTAAGTTGAGTTC
CAAGAGTGCTAAGGAAAGAAGAAATCGGAGGAAGAAAAGAAAACAGAAAGAGCAGTCTGGTGGGGAAGAGAAAGATGA
GGATGAATTCCAAAAATCTGAATCTGAGGACAGCATCAGGAGGAAAGGTTTTCGCTTCTCCATTGAAGGGAACCGATT
GACATATGAAAAGAGGTACTCCTCCCCACACCAGTCTTTGTTGAGCATCCGTGGCTCCCTATTTTCACCAAGGCGAAA
TAGCAGAACAAGCCTTTTCAGCTTTAGAGGGCGAGCAAAGGATGTGGGATCTGAGAACGACTTCGCAGATGATGAGCA
CAGCACCTTTGAGGATAACGAGAGCCGTAGAGATTCCTTGTTTGTGCCCCGACGACACGGAGAGAGACGCAACAGCAA

CCTGAGTCAGACCAGTAGGTCATCCCGGATGCTGGCAGTGTTTCCAGCGAATGGGAAGATGCACAGCACTGTGGATTG
CAATGGTGTGGTTTCCTTGGTTGGTGGACCTTCAGTTCCTACATCGCCTGTTGGACAGCTTCTGCCAGAGGTGATAAT
AGATAAGCCAGCTACTGATGACAATGGAACAACCACTGAAACTGAAATGAGAAAGAGAAGGTCAAGTTCTTTCCACGT
TTCCATGGACTTTCTAGAAGATCCTTCCCAAAGGCAACGAGCAATGAGTATAGCCAGCATTCTAACAAATACAGTAGA
AGAACTTGAAGAATCCAGGCAGAAATGCCCACCCTGTTGGTATAAATTTTCCAACATATTCTTAATCTGGGACTGTTC
TCCATATTGGTTAAAAGTGAAACATGTTGTCAACCTGGTTGTGATGGACCCATTTGTTGACCTGGCCATCACCATCTG
TATTGTCTTAAATACTCTTTTCATGGCCATGGAGCACTATCCAATGACGGACCATTTCAATAATGTGCTTACAGTAGG
AAACTTGGTTTTCACTGGGATCTTTACAGCAGAAATGTTTCTGAAAATTATTGCCATGGATCCTTACTATTATTTCCA
AGAAGGCTGGAATATCTTTGACGGTTTTATTGTGACGCTTAGCCTGGTAGAACTTGGACTCGCCAATGTGGAAGGATT
ATCTGTTCTCCGTTCATTTCGATTGCTGCGAGTTTTCAAGTTGGCAAAATCTTGGCCAACGTTAAATATGCTAATAAA
GATCATCGGCAATTCCGTGGGGGCTCTGGGAAATTTAACCCTCGTCTTGGCCATCATCGTCTTCATTTTTGCCGTGGT
CGGCATGCAGCTCTTTGGTAAAAGCTACAAAGATTGTGTCTGCAAGATCGCCAGTGATTGTCAACTCCCACGCTGGCA
CATGAATGACTTCTTCCACTCCTTCCTGATTGTGTTCCGCGTGCTGTGTGGGGAGTGGATAGAGACCATGTGGGACTG
TATGGAGGTTGCTGGTCAAGCCATGTGCCTTACTGTCTTCATGATGGTCATGGTGATTGGAAACCTAGTGGTCCTGAA
TCTCTTTCTGGCCTTGCTTCTGAGCTCATTTAGTGCAGACAACCTTGCAGCCACTGATGATGATAATGAAATGAATAA
TCTCCAAATTGCTGTGGATAGGATGCACAAAGGAGTAGCTTATGTGAAAAGAAAAATATATGAATTTATTCAACAGTC
CTTCATTAGGAAACAAAAGATTTTAGATGAAATTAAACCACTTGATGATCTAAACAACAAGAAAGACAGTTGTATGTC
CAATCATACAGCAGAAATTGGGAAAGATCTTGACTATCTTAAAGATGTAAATGGAACTACAAGTGGTATAGGAACTGG
CAGCAGTGTTGAAAAATACATTATTGATGAAAGTGATTACATGTCATTCATAAACAACCCCAGTCTTACTGTGACTGT
ACCAATTGCTGTAGGAGAATCTGACTTTGAAAATTTAAACACGGAAGACTTTAGTAGTGAATCGGATCTGGAAGAAAG
CAAAGAGAAACTGAATGAAAGCAGTAGCTCATCAGAAGGTAGCACTGTGGACATCGGCGCACCTGTAGAAGAACAGCC
CGTAGTGGAACCTGAAGAAACTCTTGAACCAGAAGCTTGTTTCACTGAAGGCTGTGTACAAAGATTCAAGTGTTGTCA
AATCAATGTGGAAGAAGGCAGAGGAAAACAATGGTGGAACCTGAGAAGGACGTGTTTCCGAATAGTTGAACATAACTG
GTTTGAGACCTTCATTGTTTTCATGATTCTCCTTAGTAGTGGTGCTCTGGCATTTGAAGATATATATATTGATCAGCG
AAAGACGATTAAGACGATGTTGGAATATGCTGACAAGGTTTTCACTTACATTTTCATTCTGGAAATGCTTCTAAAATG
GGTGGCATATGGCTATCAAACATATTTCACCAATGCCTGGTGTTGGCTGGACTTCTTAATTGTTGATGTTTCATTGGT
CAGTTTAACAGCAAATGCCTTGGGTTACTCAGAACTTGGAGCCATCAAATCTCTCAGGACACTAAGAGCTCTGAGACC
TCTAAGAGCCTTATCTCGATTTGAAGGGATGAGGGTGGTTGTGAATGCCCTTTTAGGAGCAATTCCATCCATCATGAA
TGTGCTTCTGGTTTGTCTTATATTCTGGCTAATTTTCAGCATCATGGGCGTAAATTTGTTTGCTGGCAAATTCTACCA
CTGTATTAACACCACAACTGGTGACAGGTTTGACATCGAAGACGTGAATAATCATACTGATTGCCTAAAACTAATAGA
AAGAAATGAGACTGCTCGATGGAAAAATGTGAAAGTAAACTTTGATAATGTAGGATTTGGGTATCTCTCTTTGCTTCA
AGTTGCCACATTCAAAGGATGGATGGATATAATGTATGCAGCAGTTGATTCCAGAAATGTGGAACTCCAGCCTAAGTA
TGAAGAAAGTCTGTACATGTATCTTTACTTTGTTATTTTCATCATCTTTGGGTCCTTCTTCACCTTGAACCTGTTTAT
TGGTGTCATCATAGATAATTTCAACCAGCAGAAAAAGAAGTTTGGAGGTCAAGACATCTTTATGACAGAAGAACAGAA
GAAATACTATAATGCAATGAAAAAATTAGGATCGAAAAAACCGCAAAAGCCTATACCTCGACCAGGAAACAAATTTCA
AGGAATGGTCTTTGACTTCGTAACCAGACAAGTTTTTGACATAAGCATCATGATTCTCATCTGTCTTAACATGGTCAC
AATGATGGTGGAAACAGATGACCAGAGTGAATATGTGACTACCATTTTGTCACGCATCAATCTGGTGTTCATTGTGCT
ATTTACTGGAGAGTGTGTACTGAAACTCATCTCTCTACGCCATTATTATTTTACCATTGGATGGAATATTTTTGATTT
TGTGGTTGTCATTCTCTCCATTGTAGGTATGTTTCTTGCCGAGCTGATAGAAAAGTATTTCGTGTCCCCTACCCTGTT
CCGAGTGATCCGTCTTGCTAGGATTGGCCGAATCCTACGTCTGATCAAAGGAGCAAAGGGGATCCGCACGCTGCTCTT
TGCTTTGATGATGTCCCTTCCTGCGTTGTTTAACATCGGCCTCCTACTCTTCCTAGTCATGTTCATCTACGCCATCTT
TGGGATGTCCAACTTTGCCTATGTTAAGAGGGAAGTTGGGATCGATGACATGTTCAACTTTGAGACCTTTGGCAACAG
CATGATCTGCCTATTCCAAATTACAACCTCTGCTGGCTGGGATGGATTGCTAGCACCCATTCTCAACAGTAAGCCACC
CGACTGTGACCCTAATAAAGTTAACCCTGGAAGCTCAGTTAAGGGAGACTGTGGGAACCCATCTGTTGGAATTTTCTT
TTTTGTCAGTTACATCATCATATCCTTCCTGGTTGTGGTGAACATGTACATCGCGGTCATCCTGGAGAACTTCAGTGT
TGCTACTGAAGAAAGTGCAGAGCCTCTGAGTGAGGATGACTTTGAGATGTTCTATGAGGTTTGGGAGAAGTTTGATCC
CGATGCAACTCAGTTCATGGAATTTGAAAAATTATCTCAGTTTGCAGCTGCGCTTGAACCGCCTCTCAATCTGCCACA
ACCAAACAAACTCCAGCTCATTGCCATGGATTTGCCCATGGTGAGTGGTGACCGGATCCACTGTCTTGATATCTTATT
TGCTTTTACAAAGCGGGTTCTAGGAGAGAGTGGAGAGATGGATGCTCTACGAATACAGATGGAAGAGCGATTCATGGC
TTCCAATCCTTCCAAGGTCTCCTATCAGCCAATCACTACTACTTTAAAACGAAAACAAGAGGAAGTATCTGCTGTCAT
TATTCAGCGTGCTTACAGACGCCACCTTTTAAAGCGAACTGTAAAACAAGCTTCCTTTACGTACAATAAAAACAAAAT
CAAAGGTGGGGCTAATCTTCTTATAAAAGAAGACATGATAATTGACAGAATAAATGAAAACTCTATTACAGAAAAAAC
TGATCTGACCATGTCCACTGCAGCTTGTCCACCTTCCTATGACCGGGTGACAAAGCCAATTGTGGAAAAACATGAGCA
AGAAGGCAAAGATGAAAAAGCCAAAGGGAAATAAATGAAAATAAATAAAAATAATTGGGTGACAAATTGTTTACAGCC
TGTGAAGGTGATGTATTTTTATCAACAGGACTCCTTTAGGAGGTCAATGCCAAACTGACTGTTTTTACACAAATCTCC
TTAAGGTCAGTGCCTACAATAAGACAGTGACCCCTTGTCAGCAAACTGTGACTCTGTGTAAAGGGGAGATGACCTTGA
CAGGAGGTTACTGTTCTCACTACCAGCTGACACTGCTGAAGATAAGATGCACAATGGCTAGTCAGACTGTAGGGACCA
GTTTCAAGGGGTGCAAACCTGTGATTTTGGGGTTGTTTAACATGAAACACTTTAGTGTAGTAATTGTATCCACTGTTT
GCATTTCAACTGCCACATTTGTCACATTTTTATGGAATCTGTTAGTGGATTCATCTTTTTGTTAATCCATGTGTTTAT
TATATGTGACTATTTTTGTAAACGAAGTTTCTGTTGAGAAATAGGCTAAGGACCTCTATAACAGGTATGCCACCTGGG
GGGTATGGCAACCACATGGCCCTCCCAGCTACACAAAGTCGTGGTTTGCATGAGGGCATGCTGCACTTAGAGATCATG
CATGAGAAAAAGTCACAAGAAAAACAAATTCTTAAATTTCACCATATTTCTGGGAGGGGTAATTGGGTGATAAGTGGA
GGTGCTTTGTTGATCTTGTTTTGCGAAATCCAGCCCCTAGACCAAGTAGATTATTTGTGGGTAGGCCAGTAAATCTTA

GCAGGTGCAAACTTCATTCAAATGTTTGGAGTCATAAATGTTATGTTTCTTTTTGTTGTATTAAAAAAAAAACCTGAA
TAGTGAATATTGCCCCTCACCCTCCACCGCCAGAAGACTGAATTGACCAAAATTACTCTTTATAAATTTCTGCTTTTT
CCTGCACTTTGTTTAGCCATCTTCGGCTCTCAGCAAGGTTGACACTGTATATGTTAATGAAATGCTATTTATTATGTA
AATAGTCATTTTACCCTGTGGTGCACGTTTGAGCAAACAAATAATGACCTAAGCACAGTATTTATTGCATCAAATATG
TACCACAAGAAATGTAGAGTGCAAGCTTTACACAGGTAATAAAATGTATTCTGTACCATTTATAGATAGTTTGGATGC
TATCAATGCATGTTTATATTACCATGCTGCTGTATCTGGTTTCTCTCACTGCTCAGAATCTCATTTATGAGAAACCAT
ATGTCAGTGGTAAAGTCAAGGAAATTGTTCAACAGATCTCATTTATTTAAGTCATTAAGCAATAGTTTGCAGCACTTT
AACAGCTTTTTGGTTATTTTTACATTTTAAGTGGATAACATATGGTATATAGCCAGACTGTACAGACATGTTTAAAAA
AACACACTGCTTAACCTATTAAATATGTGTTTAGAATTTTATAAGCAAATATAAATACTGTAAAAAGTCACTTTATTT
TATTTTTCAGCATTATGTACATAAATATGAAGAGGAAATTATCTTCAGGTTGATATCACAATCACTTTTCTTACTTTC
TGTCCATAGTACTTTTTCATGAAAGAAATTTGCTAAATAAGACATGAAAACAAGACTGGGTAGTTGTAGATTTCTGCT
TTTTAAATTACATTTGCTAATTTTAGATTATTTCACAATTTTAAGGAGCAAAATAGGTTCACGATTCATATCCAAATT
ATGCTTTGCAATTGGAAAAGGGTTTAAAATTTTATTTATATTTCTGGTAGTACCTGCACTAACTGAATTGAAGGTAGT
GCTTATGTTATTTTTGTTCTTTTTTTCTGACTTCGGTTTATGTTTTCATTTCTTTGGAGTAATGCTGCTCTAGATTGT
TCTAAATAGAATGTGGGCTTCATAATTTTTTTTTCCACAAAAACAGAGTAGTCAACTTATATAGTCAATTACATCAGG
ACATTTTGTGTTTCTTACAGAAGCAAACCATAGGCTCCTCTTTTCCTTAAAACTACTTAGATAAACTGTATTCGTGAA
CTGCATGCTGGAAAATGCTACTATTATGCTAAATAATGCTAACCAACATTTAAAATGTGCAAAACTAATAAAGATTAC
ATTTTTTATTTTATTGTTTGCCCAGTCACTTTTTGTTAACAGAATATTCTAATGATATGGAGATTTTTTACATTACAA
ATTGGGGGAGAAGGGGAGCGCGCGCGCACACACACACACACACACACACACACACACACACACACACAGAGGCATACCCA
CGTTGACAACAAAACCTAGGGTAGATATGTCACTGGAGGTAGGGGGTAATGACCTCCCAGAATTACAAGCAGCAGGTG
TGTTCTCTGTTAGGAGGAAGAACTGGTGTCAGAGGATAGCTAGTGATTCTAGGAGGAAGAGAAGTATGGAAGCCAGAG
TGATGGTGGATGACCCCTTGAGCTATGAAAAGAAACCCTTAAATCATCATTTAAAAATTTAGAATTGCCATGTGTGTA
GGATACTGTGTTTGCTCCTCCAGAGCCACTCTCTCTGCTTCTGCATCATTCTGTGTGTCCCAGAAGGGTGACTTCTAC
ACATTGCAAAAATGGGCTCTCCTACCTTTGAGCTCCCAATTGGTTTGGCCAATGAGAAGCACCAGTGGGAAAGCACCA
GAGAGAGAAGATTGACATAGGAATATTTCTTCTCCAATTCCTTCTTTGCTGGGTTGGCACTGGACTCATTCCTCCCCG
AAAAGTCATACTCCAATCAGACTGCCCCTCATACAACTGAAGCTACTTTCTCTGGGGTCAGGTAATCACTCCTCCCCT
TGCTCCTTCAGGTCTGCTGCTGCATTGAGAGTGCTTTTGTATTCCTTGTAGCTTTCTCCTAACATTGCTGACACTTTT
GTAAATGTCCCCTTCATGAAATTCTTCTATATGCCTCATTTCAGCATGCCATCTGTCTCCTGCCTGGCTGACACAAGG
TGATTCAACAGCTCATGAAAGTCAGCAGGAAGCAAAGATGTGCCTTGCTTCAGCTTGGGGTCTTAATCTTGCTAACTT
TTGCAGATAAAGAAAAACAGTAACTGGGGGAACCACAGTGAAGTCCAGTGCAGAATTCACAGATATCATGGAAAGGTT
ACTCGGGTGGTCCAGATAGTAAAATTAACAGTCTAAATTAATCTATCTAAATTTCTGAGGAACGAGAAGCCTTCCCTT
GTCATCAGGTGAAGCCAGAAGAGGGAATATAGCCTCAACCAGAAAAGGGACAGTAATTAAAAGGCTTTTCCCATCCTT
GTACAATGGACTGACTTTGCCTCTTCATAACATCACAATCCTAAAGCAACACAACAATTAATTCTGATATATTAGTAG
CTGAAAAAAATTCCCATTTCCAACTAAGGTAGGTCAGAATTATAGGATAAACCCTGCAGACTTTTTATACTACCCATC
CACGCCATTACTCACTGTTACCTTTCCAAATACAAAGAGAAGAACTGGTAAAACATAATCATATAAATCTCCATATTC
ATTTTGAAATATTTGGCATGATATTTTCTGTGCTAAAAAGTAATTATTCTTCAAAGAATGATGAGGTCATGTCAGTAA
GACACAGGAACCAACTAGAAGGGGCTTCCCACTGGCCAAATCTGGGGCAAGTTGAGCATCAAAATAAATGATAGTAAA
AGATTATAATTCATTGAATAAGAATCAGCAAATACATACTGATGTAAGTAAATAAGGAAAAGTACAAATCTGTTTCTT
GCAGTTGAATGTTAATTAACAATTGTAGAAGAAATAACGGAGTTAGAAAAATCACTATTTGGCAATCACCCTAATGAC
AATTGATTCATACAAGAATCATCAATGAGTATTAAAACTCATGGGTGAAAGTTTGATGAGGAATAGGGTATTTATAGC
ATCTTAAAGTATCTCTTCTCTATTAAGTAGAAAATTTAAACAGAAGAAAGTATACTTTGGAGAAATACAGCAGACAAT
ACCTTCAAAGATATCATCAATTATGAGACCAACTGATACTATGTGCCTCCTGATAAGATATACTGAAAGGGCCACATT
ACTTCTTGGTACACAGTCAAAATTTTAAAACCAGAATCTAACTACAAGGAAAATCAAATTGAGGACACTCTATAAAAT
AAGTGGACTGAACTCCTTAAAAATGTCAATGTCATGAAAGACAAAGAAAGGCTAAAGAATTCCATGAGGTCAAAGAAC
TATGACAACTAAACACAATTCTGGATGGAATATCAAATTAAAAAATAACAGATAAATAATATTATTGGGAAAGTTGAA
TAAATTTGAATATGGACTGTTTATTAGTTATTAGTATTATAATAGTGTTAATTTTCCTAATTTTGTTAAGACTAGTGT
GCCTGTTCCATGAAAATAGAAAATGTTCTTATTCTCTGAAAATGCATGCTAAAGTATTTAGGGGTGAATGCAACAATG
TCTGCAGCTCATTCTTGAATCAGTTCAAAGAAAAATGAGTTACATTTATATATATATGTATGTATGTAAACAGACATA
GATAAAAGTATAGATGTGTGTGTGTCTTTAGAAAGGGGAGGATTTTTTTTTTTTTTTTGCTGTGTGTTACTGAAGTGC
CTATGTCTGCGTGTTCACACTATCATATTTTGTATGCCCTGGACTTTATAATTTCTACCTTCAAAATTAGATCTACTG
TTGGTAATTAATTCAATATATACTGGTTTTTTAACTACTATTCTCATTTCCTAGCAGTAATCTTCCTGAAAAGTCACA
GAAATGATTACATTCCTTGTTCTTCATAATAATCACTGTTTAATTAAAATAAGAATATTTTAGAAAAGATCTGCGGCA
TAGTGGTTAAGACCCCAGTATTTGATGCTAAACAGATCTGATTTGGATAACAGAAGGTGGCACTTTGCTGTTTAAGCT
GGGGACCAGACACTGTGGGTATAAATAGTAATTCCAAACACAGCTCCACAGAGCAGCACCCTTATGACAAGGTTTTCA
TATGTCTATAGTTAAGCCAGAAAATTAAGAATAATGCCATAAATATTTATAAAGCTGAACATATCCAAGTTAAAGACC
TTTATCCTGAAATTGTATCTTTTAGATTATTTTCTAAAGACTAATACCATTTAATGTTTAAATGTTCTTTGGAAATGA
TGGTGAGAATACGTGATAATGGGTCATTGGTTTTAATATTTTATTTAGCCAAGTGGAAAATTGGCAACCTGGTGTCGG
TCCTCCCATTTGTATTTTACTGGTGCATGAAATCCAAAAGTCTAGTAACCATTGGGACAGACAACTCTACTGCATAAG
TTTGTATGTTTGTATATCTGTATCACAAAGCCCAGACACTCGAACTATATAAACTTGTCGCACTAAAGACAGCAAATA
TGTCTGGTAATTGCATATTCTTCATGTGTGCACTGGAATTTCTTATTATATAAGAAAATAAATGTGTTTCTAAACCAC
CATGAA

SEQ ID NO: 19 >NP_001159435.1 SCN1A [organism=Homo sapiens] [GeneID=6323] [isoform=1]

MEQTVLVPPGPDSFNFFTRESLAAIERRIAEEKAKNPKPDKKDDDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLE
DLDPYYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKIAIKILVHSLFSMLIMCTILTNCVFMTMSNPPDWTKN
VEYTFTGIYTFESLIKIIARGFCLEDFTFLRDPWNWLDFTVITFAYVTEFVDLGNVSALRTFRVLRALKTISVIPGLK
TIVGALIQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCIQWPPTNASLEEHSIEKNITVNYNGTLINETVFEF
DWKSYIQDSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGRNPNYGYTSFDTFSWAFLSLFRLMTQDFWENLYQL
TLRAAGKTYMIFFVLVIFLGSFYLINLILAVVAMAYEEQNQATLEEAEQKEAEFQQMIEQLKKQQEAAQQAATATASE
HSREPSAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGGEEKDEDEFQKSESEDSIRRKGFRFSIEGNRLTYE
KRYSSPHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDVGSENDFADDEHSTFEDNESRRDSLFVPRRHGERRNSNLSQ
TSRSSRMLAVFPANGKMHSTVDCNGVVSLVGGPSVPTSPVGQLLPEVIIDKPATDDNGTTTETEMRKRRSSSFHVSMD
FLEDPSQRQRAMSIASILTNTVEELEESRQKCPPCWYKFSNIFLIWDCSPYWLKVKHVVNLVVMDPFVDLAITICIVL
NTLFMAMEHYPMTDHFNNVLTVGNLVFTGIFTAEMFLKIIAMDPYYYFQEGWNIFDGFIVTLSLVELGLANVEGLSVL
RSFRLLRVFKLAKSWPTLNMLIKIIGNSVGALGNLTLVLAIIVFIFAVVGMQLFGKSYKDCVCKIASDCQLPRWHMND
FFHSFLIVFRVLCGEWIETMWDCMEVAGQAMCLTVFMMVMVIGNLVVLNLFLALLLSSFSADNLAATDDDNEMNNLQI
AVDRMHKGVAYVKRKIYEFIQQSFIRKQKILDEIKPLDDLNNKKDSCMSNHTAEIGKDLDYLKDVNGTTSGIGTGSSV
EKYIIDESDYMSFINNPSLTVTVPIAVGESDFENLNTEDFSSESDLEESKEKLNESSSSSEGSTVDIGAPVEEQPVVE
PEETLEPEACFTEGCVQRFKCCQINVEEGRGKQWWNLRRTCFRIVEHNWFETFIVFMILLSSGALAFEDIYIDQRKTI
KTMLEYADKVFTYIFILEMLLKWVAYGYQTYFTNAWCWLDFLIVDVSLVSLTANALGYSELGAIKSLRTLRALRPLRA
LSRFEGMRVVVNALLGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNE
TARWKNVKVNFDNVGFGYLSLLQVATFKGWMDIMYAAVDSRNVELQPKYEESLYMYLYFVIFIIFGSFFTLNLFIGVI
IDNFNQQKKKFGGQDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFQGMVFDFVTRQVFDISIMILICLNMVTMMV
ETDDQSEYVTTILSRINLVFIVLFTGECVLKLISLRHYYFTIGWNIFDFVVVILSIVGMFLAELIEKYFVSPTLFRVI
RLARIGRILRLIKGAKGIRTLLFALMMSLPALFNIGLLLFLVMFIYAIFGMSNFAYVKREVGIDDMFNFETFGNSMIC
LFQITTSAGWDGLLAPILNSKPPDCDPNKVNPGSSVKGDCGNPSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATE
ESAEPLSEDDFEMFYEVWEKFDPDATQFMEFEKLSQFAAALEPPLNLPQPNKLQLIAMDLPMVSGDRIHCLDILFAFT
KRVLGESGEMDALRIQMEERFMASNPSKVSYQPITTTLKRKQEEVSAVIIQRAYRRHLLKRTVKQASFTYNKNKIKGG
ANLLIKEDMIIDRINENSITEKTDLTMSTAACPPSYDRVTKPIVEKHEQEGKDEKAKGK

SEQ ID NO: 20 >NP_001159436.1 SCN1A [organism=Homo sapiens] [GeneID=6323] [isoform=3]

MEQTVLVPPGPDSFNFFTRESLAAIERRIAEEKAKNPKPDKKDDDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLE
DLDPYYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKIAIKILVHSLFSMLIMCTILTNCVFMTMSNPPDWTKN
VEYTFTGIYTFESLIKIIARGFCLEDFTFLRDPWNWLDFTVITFAYVTEFVDLGNVSALRTFRVLRALKTISVIPGLK
TIVGALIQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCIQWPPTNASLEEHSIEKNITVNYNGTLINETVFEF
DWKSYIQDSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGRNPNYGYTSFDTFSWAFLSLFRLMTQDFWENLYQL
TLRAAGKTYMIFFVLVIFLGSFYLINLILAVVAMAYEEQNQATLEEAEQKEAEFQQMIEQLKKQQEAAQQAATATASE
HSREPSAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGGEEKDEDEFQKSESEDSIRRKGFRFSIEGNRLTYE
KRYSSPHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDVGSENDFADDEHSTFEDNESRRDSLFVPRRHGERRNSNLSQ
TSRSSRMLAVFPANGKMHSTVDCNGVVSLGTTTETEMRKRRSSSFHVSMDFLEDPSQRQRAMSIASILTNTVEELEES
RQKCPPCWYKFSNIFLIWDCSPYWLKVKHVVNLVVMDPFVDLAITICIVLNTLFMAMEHYPMTDHFNNVLTVGNLVFT
GIFTAEMFLKIIAMDPYYYFQEGWNIFDGFIVTLSLVELGLANVEGLSVLRSFRLLRVFKLAKSWPTLNMLIKIIGNS
VGALGNLTLVLAIIVFIFAVVGMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVFRVLCGEWIETMWDCMEVAG
QAMCLTVFMMVMVIGNLVVLNLFLALLLSSFSADNLAATDDDNEMNNLQIAVDRMHKGVAYVKRKIYEFIQQSFIRKQ
KILDEIKPLDDLNNKKDSCMSNHTAEIGKDLDYLKDVNGTTSGIGTGSSVEKYIIDESDYMSFINNPSLTVTVPIAVG
ESDFENLNTEDFSSESDLEESKEKLNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACFTEGCVQRFKCCQINVEE
GRGKQWWNLRRTCFRIVEHNWFETFIVFMILLSSGALAFEDIYIDQRKTIKTMLEYADKVFTYIFILEMLLKWVAYGY
QTYFTNAWCWLDFLIVDVSLVSLTANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVVNALLGAIPSIMNVLLVC
LIFWLIFSIMGVNLFAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNFDNVGFGYLSLLQVATFK
GWMDIMYAAVDSRNVELQPKYEESLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKFGGQDIFMTEEQKKYYNA
MKKLGSKKPQKPIPRPGNKFQGMVFDFVTRQVFDISIMILICLNMVTMMVETDDQSEYVTTILSRINLVFIVLFTGEC
VLKLISLRHYYFTIGWNIFDFVVVILSIVGMFLAELIEKYFVSPTLFRVIRLARIGRILRLIKGAKGIRTLLFALMMS
LPALFNIGLLLFLVMFIYAIFGMSNFAYVKREVGIDDMFNFETFGNSMICLFQITTSAGWDGLLAPILNSKPPDCDPN
KVNPGSSVKGDCGNPSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATEESAEPLSEDDFEMFYEVWEKFDPDATQF
MEFEKLSQFAAALEPPLNLPQPNKLQLIAMDLPMVSGDRIHCLDILFAFTKRVLGESGEMDALRIQMEERFMASNPSK
VSYQPITTTLKRKQEEVSAVIIQRAYRRHLLKRTVKQASFTYNKNKIKGGANLLIKEDMIIDRINENSITEKTDLTMS
TAACPPSYDRVTKPIVEKHEQEGKDEKAKGK

SEQ ID NO: 21 >NP_001340878.1 SCN1A [organism=Homo sapiens] [GeneID=6323] [isoform=2]

MEQTVLVPPGPDSFNFFTRESLAAIERRIAEEKAKNPKPDKKDDDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLE
DLDPYYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKIAIKILVHSLFSMLIMCTILTNCVFMTMSNPPDWTKN
VEYTFTGIYTFESLIKIIARGFCLEDFTFLRDPWNWLDFTVITFAYVTEFVDLGNVSALRTFRVLRALKTISVIPGLK
TIVGALIQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCIQWPPTNASLEEHSIEKNITVNYNGTLINETVFEF
DWKSYIQDSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGRNPNYGYTSFDTFSWAFLSLFRLMTQDFWENLYQL
TLRAAGKTYMIFFVLVIFLGSFYLINLILAVVAMAYEEQNQATLEEAEQKEAEFQQMIEQLKKQQEAAQQAATATASE
HSREPSAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGGEEKDEDEFQKSESEDSIRRKGFRFSIEGNRLTYE
KRYSSPHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDVGSENDFADDEHSTFEDNESRRDSLFVPRRHGERRNSNLSQ
TSRSSRMLAVFPANGKMHSTVDCNGVVSLVGGPSVPTSPVGQLLPEGTTTETEMRKRRSSSFHVSMDFLEDPSQRQRA
MSIASILTNTVEELEESRQKCPPCWYKFSNIFLIWDCSPYWLKVKHVVNLVVMDPFVDLAITICIVLNTLFMAMEHYP
MTDHFNNVLTVGNLVFTGIFTAEMFLKIIAMDPYYYFQEGWNIFDGFIVTLSLVELGLANVEGLSVLRSFRLLRVFKL
AKSWPTLNMLIKIIGNSVGALGNLTLVLAIIVFIFAVVGMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVFRV
LCGEWIETMWDCMEVAGQAMCLTVFMMVMVIGNLVVLNLFLALLLSSFSADNLAATDDDNEMNNLQIAVDRMHKGVAY
VKRKIYEFIQQSFIRKQKILDEIKPLDDLNNKKDSCMSNHTAEIGKDLDYLKDVNGTTSGIGTGSSVEKYIIDESDYM
SFINNPSLTVTVPIAVGESDFENLNTEDFSSESDLEESKEKLNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACF
TEGCVQRFKCCQINVEEGRGKQWWNLRRTCFRIVEHNWFETFIVFMILLSSGALAFEDIYIDQRKTIKTMLEYADKVF
TYIFILEMLLKWVAYGYQTYFTNAWCWLDFLIVDVSLVSLTANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVV
NALLGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNF
DNVGFGYLSLLQVATFKGWMDIMYAAVDSRNVELQPKYEESLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKF
GGQDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFQGMVFDFVTRQVFDISIMILICLNMVTMMVETDDQSEYVTT
ILSRINLVFIVLFTGECVLKLISLRHYYFTIGWNIFDFVVVILSIVGMFLAELIEKYFVSPTLFRVIRLARIGRILRL
IKGAKGIRTLLFALMMSLPALFNIGLLLFLVMFIYAIFGMSNFAYVKREVGIDDMFNFETFGNSMICLFQITTSAGWD
GLLAPILNSKPPDCDPNKVNPGSSVKGDCGNPSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATEESAEPLSEDDF
EMFYEVWEKFDPDATQFMEFEKLSQFAAALEPPLNLPQPNKLQLIAMDLPMVSGDRIHCLDILFAFTKRVLGESGEMD
ALRIQMEERFMASNPSKVSYQPITTTLKRKQEEVSAVIIQRAYRRHLLKRTVKQASFTYNKNKIKGGANLLIKEDMII
DRINENSITEKTDLTMSTAACPPSYDRVTKPIVEKHEQEGKDEKAKGK

SEQ ID NO: 22 >NP_001340883.1 SCN1A [organism=Homo sapiens] [GeneID=6323] [isoform=4]

MEQTVLVPPGPDSFNFFTRESLAAIERRIAEEKAKNPKPDKKDDDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLE
DLDPYYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKIAIKILVHSLFSMLIMCTILTNCVFMTMSNPPDWTKN
VEYTFTGIYTFESLIKIIARGFCLEDFTFLRDPWNWLDFTVITFAYVTEFVDLGNVSALRTFRVLRALKTISVIPGLK
TIVGALIQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCIQWPPTNASLEEHSIEKNITVNYNGTLINETVFEF
DWKSYIQDSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGRNPNYGYTSFDTFSWAFLSLFRLMTQDFWENLYQL
TLRAAGKTYMIFFVLVIFLGSFYLINLILAVVAMAYEEQNQATLEEAEQKEAEFQQMIEQLKKQQEAAQQAATATASEH
SREPSAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGGEEKDEDEFQKSESEDSIRRKGFRFSIEGNRLTYEK
RYSSPHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDVGSENDFADDEHSTFEDNESRRDSLFVPRRHGERRNSNLSQT
SRSSRMLAVFPANGKMHSTVDCNGVVSLVGGPSVPTSPVGQLLPEGTTTETEMRKRRSSSFHVSMDFLEDPSQRQRAM
SIASILTNTVEELEESRQKCPPCWYKFSNIFLIWDCSPYWLKVKHVVNLVVMDPFVDLAITICIVLNTLFMAMEHYPM
TDHFNNVLTVGNLVFTGIFTAEMFLKIIAMDPYYYFQEGWNIFDGFIVTLSLVELGLANVEGLSVLRSFRLLRVFKLA
KSWPTLNMLIKIIGNSVGALGNLTLVLAIIVFIFAVVGMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVFRVL
CGEWIETMWDCMEVAGQAMCLTVFMMVMVIGNLVVLNLFLALLLSSFSADNLAATDDDNEMNNLQIAVDRMHKGVAYV
KRKIYEFIQQSFIRKQKILDEIKPLDDLNNKKDSCMSNHTAEIGKDLDYLKDVNGTTSGIGTGSSVEKYIIDESDYMS
FINNPSLTVTVPIAVGESDFENLNTEDFSSESDLEESKEKLNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACFT
EGCVQRFKCCQINVEEGRGKQWWNLRRTCFRIVEHNWFETFIVFMILLSSGALAFEDIYIDQRKTIKTMLEYADKVFT
YIFILEMLLKWVAYGYQTYFTNAWCWLDFLIVDVSLVSLTANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVVN
ALLGAIPSIMNVLLVCLIFWLIFSIMGVNLFAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNFD
NVGFGYLSLLQVATFKGWMDIMYAAVDSRNVELQPKYEESLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKFG
GQDIFMTEEQKKYYNAMKKLGSKKPQKPIPRPGNKFQGMVFDFVTRQVFDISIMILICLNMVTMMVETDDQSEYVTTI
LSRINLVFIVLFTGECVLKLISLRHYYFTIGWNIFDFVVVILSIVGMFLAELIEKYFVSPTLFRVIRLARIGRILRLI
KGAKGIRTLLFALMMSLPALFNIGLLLFLVMFIYAIFGMSNFAYVKREVGIDDMFNFETFGNSMICLFQITTSAGWDG
LLAPILNSKPPDCDPNKVNPGSSVKGDCGNPSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATEESAEPLSEDDFE
MFYEVWEKFDPDATQFMEFEKLSQFAAALEPPLNLPQPNKLQLIAMDLPMVSGDRIHCLDILFAFTKRVLGESGEMDA
LRIQMEERFMASNPSKVSYQPITTTLKRKQEEVSAVIIQRAYRRHLLKRTVKQASFTYNKNKIKGGANLLIKEDMIID
RINENSITEKTDLTMSTAACPPSYDRVTKPIVEKHEQEGKDEKAKGK

SEQ ID NO: 23 >NP_001340889.1 SCN1A [organism=Homo sapiens] [GeneID=6323] [isoform=5]

MEQTVLVPPGPDSFNFFTRESLAAIERRIAEEKAKNPKPDKKDDDENGPKPNSDLEAGKNLPFIYGDIPPEMVSEPLE
DLDPYYINKKTFIVLNKGKAIFRFSATSALYILTPFNPLRKIAIKILVHSLFSMLIMCTILTNCVFMTMSNPPDWTKN
VEYTFTGIYTFESLIKIIARGFCLEDFTFLRDPWNWLDFTVITFAYVTEFVDLGNVSALRTFRVLRALKTISVIPGLK
TIVGALIQSVKKLSDVMILTVFCLSVFALIGLQLFMGNLRNKCIQWPPTNASLEEHSIEKNITVNYNGTLINETVFEF
DWKSYIQDSRYHYFLEGFLDALLCGNSSDAGQCPEGYMCVKAGRNPNYGYTSFDTFSWAFLSLFRLMTQDFWENLYQL
TLRAAGKTYMIFFVLVIFLGSFYLINLILAVVAMAYEEQNQATLEEAEQKEAEFQQMIEQLKKQQEAAQAATATASEH
SREPSAAGRLSDSSSEASKLSSKSAKERRNRRKKRKQKEQSGGEEKDEDEFQKSESEDSIRRKGFRFSIEGNRLTYEK
RYSSPHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDVGSENDFADDEHSTFEDNESRRDSLFVPRRHGERRNSNLSQT
SRSSRMLAVFPANGKMHSTVDCNGVVSLGTTTETEMRKRRSSSFHVSMDFLEDPSQRQRAMSIASILTNTVEELEESR
QKCPPCWYKFSNIFLIWDCSPYWLKVKHVVNLVVMDPFVDLAITICIVLNTLFMAMEHYPMTDHFNNVLTVGNLVFTG
IFTAEMFLKIIAMDPYYYFQEGWNIFDGFIVTLSLVELGLANVEGLSVLRSFRLLRVFKLAKSWPTLNMLIKIIGNSV
GALGNLTLVLAIIVFIFAVVGMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVFRVLCGEWIETMWDCMEVAGQ
AMCLTVFMMVMVIGNLVVLNLFLALLLSSFSADNLAATDDDNEMNNLQIAVDRMHKGVAYVKRKIYEFIQQSFIRKQK
ILDEIKPLDDLNNKKDSCMSNHTAEIGKDLDYLKDVNGTTSGIGTGSSVEKYIIDESDYMSFINNPSLTVTVPIAVGE
SDFENLNTEDFSSESDLEESKEKLNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACFTEGCVQRFKCCQINVEEG
RGKQWWNLRRTCFRIVEHNWFETFIVFMILLSSGALAFEDIYIDQRKTIKTMLEYADKVFTYIFILEMLLKWVAYGYQ
TYFTNAWCWLDFLIVDVSLVSLTANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVVNALLGAIPSIMNVLLVCL
IFWLIFSIMGVNLFAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNFDNVGFGYLSLLQVATFKG
WMDIMYAAVDSRNVELQPKYEESLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKFGGQDIFMTEEQKKYYNAM
KKLGSKKPQKPIPRPGNKFQGMVFDFVTRQVFDISIMILICLNMVTMMVETDDQSEYVTTILSRINLVFIVLFTGECV
LKLISLRHYYFTIGWNIFDFVVVILSIVGMFLAELIEKYFVSPTLFRVIRLARIGRILRLIKGAKGIRTLLFALMMSL
PALFNIGLLLFLVMFIYAIFGMSNFAYVKREVGIDDMFNFETFGNSMICLFQITTSAGWDGLLAPILNSKPPDCDPNK
VNPGSSVKGDCGNPSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATEESAEPLSEDDFEMFYEVWEKFDPDATQFM
EFEKLSQFAAALEPPLNLPQPNKLQLIAMDLPMVSGDRIHCLDILFAFTKRVLGESGEMDALRIQMEERFMASNPSKV
SYQPITTTLKRKQEEVSAVIIQRAYRRHLLKRTVKQASFTYNKNKIKGGANLLIKEDMIIDRINENSITEKTDLTMST
AACPPSYDRVTKPIVEKHEQEGKDEKAKGK

SEQ ID NO: 24 >NP_001340890.1 SCN1A [organism=Homo sapiens] [GeneID=6323] [isoform=6]

MFLKIIAMDPYYYFQEGWNIFDGFIVTLSLVELGLANVEGLSVLRSFRLLRVFKLAKSWPTLNMLIKIIGNSVGALGN
LTLVLAIIVFIFAVVGMQLFGKSYKDCVCKIASDCQLPRWHMNDFFHSFLIVFRVLCGEWIETMWDCMEVAGQAMCLT
VFMMVMVIGNLVVLNLFLALLLSSFSADNLAATDDDNEMNNLQIAVDRMHKGVAYVKRKIYEFIQQSFIRKQKILDEI
KPLDDLNNKKDSCMSNHTAEIGKDLDYLKDVNGTTSGIGTGSSVEKYIIDESDYMSFINNPSLTVTVPIAVGESDFEN
LNTEDFSSESDLEESKEKLNESSSSSEGSTVDIGAPVEEQPVVEPEETLEPEACFTEGCVQRFKCCQINVEEGRGKQW
WNLRRTCFRIVEHNWFETFIVFMILLSSGALAFEDIYIDQRKTIKTMLEYADKVFTYIFILEMLLKWVAYGYQTYFTN
AWCWLDFLIVDVSLVSLTANALGYSELGAIKSLRTLRALRPLRALSRFEGMRVVVNALLGAIPSIMNVLLVCLIFWLI
FSIMGVNLFAGKFYHCINTTTGDRFDIEDVNNHTDCLKLIERNETARWKNVKVNFDNVGFGYLSLLQVATFKGWMDIM
YAAVDSRNVELQPKYEESLYMYLYFVIFIIFGSFFTLNLFIGVIIDNFNQQKKKFGGQDIFMTEEQKKYYNAMKKLGS
KKPQKPIPRPGNKFQGMVFDFVTRQVFDISIMILICLNMVTMMVETDDQSEYVTTILSRINLVFIVLFTGECVLKLIS
LRHYYFTIGWNIFDFVVVILSIVGMFLAELIEKYFVSPTLFRVIRLARIGRILRLIKGAKGIRTLLFALMMSLPALFN
IGLLLFLVMFIYAIFGMSNFAYVKREVGIDDMFNFETFGNSMICLFQITTSAGWDGLLAPILNSKPPDCDPNKVNPGS
SVKGDCGNPSVGIFFFVSYIIISFLVVVNMYIAVILENFSVATEESAEPLSEDDFEMFYEVWEKFDPDATQFMEFEKL
SQFAAALEPPLNLPQPNKLQLIAMDLPMVSGDRIHCLDILFAFTKRVLGESGEMDALRIQMEERFMASNPSKVSYQPI
TTTLKRKQEEVSAVIIQRAYRRHLLKRTVKQASFTYNKNKIKGGANLLIKEDMIIDRINENSITEKTDLTMSTAACPP
SYDRVTKPIVEKHEQEGKDEKAKGK

SEQ ID NO: 25 >NM_001386139.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=10]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGA
CAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATC
AGCCCACCACTCTGCTGAGCCCGCAGAGGCAGGCAAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCC
GGAAGCTTCTGCCTCCCCCAAACAGCGGCGCTCCATCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTGCC
TGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAATCTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAATCC
CCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGCTCCAGGAACTGGCAGAGGCCGGGGA
CGCCCCAAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAA
AGTCCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACCACA

TCCACCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCATTCCCAAGAAA
CGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATC
CGATCTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGGTCAAGGAAGTG
GTGAAGCCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGAAA
AGCAAGGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACCACCATCACCAC
CACTCAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCGAGAGCTCCGAG
GACCCCACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCAGAGGAGGCTCA
CTGGAGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGGCCGCAGAAAAG
TACAAACACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTGTG
GACAGCCGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACGGAGCGGATTGCAAAGCAAACCAACAAGAATAAA
GGCAGCTGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTTCCCGATTAACTGAAATAAAAAATATTTTTTT
TTCTTTCAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTTCTTACCGACAA
GCACAGTCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTTCAACAAGGCTT
GCTGCAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCCCATGTGGTCGT
TAGAGACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTACAGCTCCCCCAG
CTCCCCCCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGGATTCTTTAGAG
AAGATGGATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCCAA
TCCAAAACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGGCTAGGAGGGGG
GAGTCCCTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGGCATGGTCCATT
CTCAAGGGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGGCTGGCCTGGGG
TCAGGCCGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCACTCAGAGCAGTCTGTGACTTAGTGGACAGGGGAG
GGGGCAAAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTTAGAATTATTTG
CACTATTGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACATTTA
GTCCCTCAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGATTTGTATATATACCTGCGCTTGTTTTAAAGTGGG
CTCAGCACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTCCTGATTGGTTT
CTCTTCTCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCCTCCTCCTTCCT
AGTTCATCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACTTGTCGGCTCCAGTCCCCAGAACTCTGCCTGCCC
TTTGTCCTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCTGAGTCCGACCT
GGCCTGTCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAAGCA
AAAGGCATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGAGCTGGGCCCGA
TTTGGTAGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGCTCCATGCCAGC
CGACCTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGCAGAACTCCCAA
GACCCGAGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAATGG
CAATGTTTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCACACATTTCACTGTCTGTCTGGTCCATAGCTGTG
GTGTAGGGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGATCAGTTTTCATGTGGGATCCCATCTTTTTAACC
TCTGTTCAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTGGTATATCCTTTTCTGTGTTTACAGAGATGTCTC
TTATATCTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGCAAATGAGACAGCAGTCTTATGCTTCCAGAAACA
CCCACAGGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCAGTTATTGTCCC
TGGCTTCCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATAGAAACTGTCTAGCACTTCCTTGCCAGTCTTTAG
TGATCAGGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGGAATGGTTAGAA
GTGAGGGAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACTTTCTAGCATATATGTGTCCATTTCCTTATGC
TGTAAAAGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCACCTGCTCTGCTGATG
ACCCCCCCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCCTCCAGAACTCT
TCCTCCAAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCCTGGCACTACCAAAGGACACTTATCCACGAGA
GCGCAGCATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTATACTTAGTCAA
ATGTAATGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCTGGTAAGAGGAG
TGCGTGGCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGTTGGGAACACAG
CATTTCAAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAGCCTCTTTCCCT
TCCAGTTTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTTTCTTCCTCATT
CCCTCGTCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGCCCTGGCAGGCA
GGGGCCCTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCCTTAAGCTGCAG
GATTCTCACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAATTTCAATTTTGTCTGTACTTGAACATTATGAA
GATGGGGGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCGACAGCTTTCCAGTACCCATGGGGCTAGGTCATT
AAGGCCACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTACTACCTCCTCTCCTGACAATACTGTATGTCG
TCGAGCTCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGATTGCCGGTCTTG
ACAGCTCAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCTGTATTCCCATA
TCAAAAGACACAGGGGACACCCAGAAACGCCACATCCCCCAATCCATCAGTGCCAAACTAGCCAACGGCCCCAGCTTC
TCAGCTCGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTGGGTGGCATCAT
TCCAGGCCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCTGCCCAGCTTTT
CATTGCTGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACAAACTGCCAAGTTGGGTGGAGAAAGGAGTTTCTT
TAGCTGACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGGATACGAGCGGA
GTCCCCTGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGTCCCCCGTCTAA
CCACAGCTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAGGTACCAGTTCT

TTAAGAAGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAAGTTCACACCTG
CAGATTCTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTGGGTCTGGAGGA
CCTGCCCGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGGGATTTGTTGAT
TCTCTCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAAAAGGCATAGTCTACTTGTTATAAATCG
TTAGGATACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGGGAAAAGCTGAACACTGAAGTCAGTTCTCAACAA
TTTAGAAGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTCGATGTTTTTGAATGAATACGAGCAAGCTTTTAC
AACAGTGCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGGAATCTGGAGGT
AGCCGACCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACCAGTGACCTGGA
AGGCCCAGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAACACC
CAGGGAACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCTCGGATCAGCTG
AGCCTTCTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGGAAGCAGGTGTG
AGGGCAGTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGAGCCCAGACCCT
GGAGGAAATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCCCCATTCTGTCC
TACCTTTAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCCTGCCCACTGAT
AGCCCCTCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATATCCAGTAACAC
ATAGACTGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTGCTTTTAGTTTT
TCTGTCCCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTTGAATTTTTATATATATATCTGTATATTGCACAA
TTATAAACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGACCTGTTAAAATTATACCTGTTGCTTAATTACAAT
ATTTCTGATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAAATCTGTCTGCT
GGTCACTTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCCAGGTGAAGGAG
GCTCCAGGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTCTGGGTGCAGGA
GCGCTCCCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGAGCTGGAGTCCA
CTGGGGTGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAGCCTCGCAGGCC
TCCCTGCCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGAGACGGGAGGCC
TGGCCTGTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGGGACCCCAGAATCATGTGCGTCAGTCCAAGGGGT
CCCCTCCAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCATTTGCTTTTGT
AATTCTTAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTAATAATTAGAAA
CACCAAAAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATTATAATCACCAGGGCCAAAAGCTTCCCTCCCTGC
TGTCATTGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGCTACCCGTGAGC
CCTTCGGAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCAGTGTTAGCAGT
GGGTCATGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCTGATGTTTGGCC
GTGAACAGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACTGACATGAAATCGACGCCCGAGTTAGCCTCAC
CCGGTGACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGACAGTGGAGTGC
AAAAGGCTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTGTCACTGCTTCA
ATCAATAACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGACCAAATATCACCAGGACTGTTACTCAATGTGTGC
CGAGCCCTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCTTCCCCTTCCTT
CTTTGCCCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTTTGTGTTCCAAA
CATGAGGTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTGGGTGAAAGGAA
TTTTGCTAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGTAACTATGTTTAAAGTAATTGTTCCAGAGACA
AATATTTCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGGGAGGGGGATGGTGACTGAGATGAGAGGGGAGAG
CTGAACAGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCACTCCAAGCCAGC
AAGCCGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTGGATTCTTGTTT
TGCTTCCCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTTCCTTGAGGGGA
GCCAGGAGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCAGGAGGCTGGGA
CGCTCTGCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGGTGGGCATTCTC
TCTCCAAGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCGTCACATCCCACCCCATCCTCCAGGGCTCAACAC
TGTGACATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGGCTG
TCACCCGATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCCATTCCATGGCGAAAACATCTCCTTAGAAAAGCC
ATTACCCTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGCGGAGAGTGCTG
ACTGTAGTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTC
ATGCTCCCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCATTGCCAGTAACACAGGCCACCCTGTGCAGAAGGG
AGCTGGCTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCGGGACACACTTA
GCTGGAGATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTTTTTGTGTTTTGGGACAATTACTTTAGAAAATAA
GTAGGTCGTTTTAAAAACAAAAATTATTGATTGCTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGTGTATAGCCAAA
TGACTGAAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTAAGGAAATTTGTACCATTTCAGTAAACCTGTCTG
AATGTACCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATCCCTGTAACCTATTTATTATATAAAGAGTTTGCC
TTATAAATTTACATAAAAATGTCCGTTTGTGTCTTTTGTTGTAAAAATCAAGTGATTTTTTCATAAGGTTCTTTTACT
ATTGGAAAAGATGGGCAGCACGCAGTTTTATTTTATTTTTGTAAGTTTTTTAATACATGTGAAAGCAAAGAATACTCA
GCATGCCTTTCTAAGTGACGCGTTTGCACCTTTTGTTGGGAAGTACTGTATCCTGTGCTGTTAGCATTCTCGATAAAT
CTCTCTGTGAAAGTGA

SEQ ID NO: 26 >NM_001110792.2 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=2]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGA
CAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATC
AGCCCACCACTCTGCTGAGCCCGCAGAGGCAGGCAAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCC
GGAAGCTTCTGCCTCCCCCAAACAGCGGCGCTCCATCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTGCC
TGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAATCTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAATCC
CCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCTAA
TGATTTTGACTTCACGGTAACTGGGAGAGGGAGCCCCTCCCGGCGAGAGCAGAAACCACCTAAGAAGCCCAAATCTCC
CAAAGCTCCAGGAACTGGCAGAGGCCGGGGACGCCCCAAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCAGA
GGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGG
GGGCAAGGCTGAGGGGGGTGGGGCCACCACATCCACCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGC
TGAGGCCGACCCTCAGGCCATTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGC
CAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACCCG
GGAGACGGTCAGCATCGAGGTCAAGGAAGTGGTGAAGCCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGG
ACTGAAGACCTGTAAGAGCCCTGGGCGGAAAAGCAAGGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACC
CCCCAAGAAGGAGCACCACCACCATCACCACCACTCAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCC
CCCACCTCCACCTGAGCCCGAGAGCTCCGAGGACCCCACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTG
CAAAGAGGAGAAGATGCCCAGAGGAGGCTCACTGGAGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGC
GGTTGCCACCGCCGCCACGGCCGCAGAAAAGTACAAACACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCCTC
CATGCCAAGGCCAAACAGAGAGGAGCCTGTGGACAGCCGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACGGA
GCGGATTGCAAAGCAAACCAACAAGAATAAAGGCAGCTGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTT
CCCGATTAACTGAAATAAAAAATATTTTTTTTTCTTTCAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTT
GGAGCATTGGGGATGTTTTTCTTACCGACAAGCACAGTCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGCAC
TTTTCTAAACTAGGCTCCTTCAACAAGGCTTGCTGCAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCC
TCCCGCCCAAACCTTTCCCCCATGTGGTCGTTAGAGACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGT
GCCATCAGTGCCCCGTCTACAGCTCCCCCAGCTCCCCCCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGTGT
GAGGCAGGAGAGACAGTTGGATTCTTTAGAGAAGATGGATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGTGG
TGGCTCAAGTCTGGCCCCACACCAGCCCCAATCCAAAACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCTGT
CTGCTCCAGCTCTGGCATGGCTAGGAGGGGGGAGTCCCTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAGAG
GATGGCCCAGGGTGAGGTGGCATGGTCCATTCTCAAGGGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAG
TAGGACAAGGTGCAGGCAGGCTGGCCTGGGGTCAGGCCGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCACTC
AGAGCAGTCTGTGACTTAGTGGACAGGGGAGGGGGCAAAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAA
TTCTCCTTTAGGGACAGCTTAGAATTATTTGCACTATTGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTCTG
AGAGCAAACTGGCTTGAATTGGTGACATTTAGTCCCTCAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGATTT
GTATATATACCTGCGCTTGTTTTAAAGTGGGCTCAGCACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGC
TGCAATTTTATAAGGACTTCCTGATTGGTTTCTCTTCTCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTCAC
TTCTTTCCCTTCCTCCATCCTCCTCCTTCCTAGTTCATCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACTTG
TCGGCTCCAGTCCCCAGAACTCTGCCTGCCCTTTGTCCTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTG
AGGAGGCCTTGGGCTCTGCTGAGTCCGACCTGGCCTGTCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAG
GTAGCCCCCTCTTCCCTGGTAAGAAAAAGCAAAAGGCATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTACTC
TAGAGAAGTGGACTGGAGGAGCTGGGCCCGATTTGGTAGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCA
TCGAGTGGCCCAACAGGGGCTCCATGCCAGCCGACCTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAG
GGCAGTAGCGGTACCAATGCAGAACTCCCAAGACCCGAGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTC
CCCCTTTTCCCTCGGAGTTCTTCTTGAATGGCAATGTTTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCACA
CATTTCACTGTCTGTCTGGTCCATAGCTGTGGTGTAGGGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGATC
AGTTTTCATGTGGGATCCCATCTTTTTAACCTCTGTTCAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTGGT
ATATCCTTTTCTGTGTTTACAGAGATGTCTCTTATATCTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGCAA
ATGAGACAGCAGTCTTATGCTTCCAGAAACACCCACAGGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGTGT
GTGTTGTCTTGTGTATTTCAGTTATTGTCCCTGGCTTCCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATAGA
AACTGTCTAGCACTTCCTTGCCAGTCTTTAGTGATCAGGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAA
ACCGTGTTTGTCTCTTCTGGAATGGTTAGAAGTGAGGGAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACT
TTCTAGCATATATGTGTCCATTTCCTTATGCTGTAAAAGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCTGA
TCCCTGATCCCTTCCACCTGCTCTGCTGATGACCCCCCCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGG
TCAGAAGAGAGGGTGAGTCCTCCAGAACTCTTCCTCCAAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCC
TGGCACTACCAAAGGACACTTATCCACGAGAGCGCAGCATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGGTC
AGTTGGGTTTTTATGTATTATACTTAGTCAAATGTAATGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCA
CCTGGGCGTCATCTGGTCCTGGTAAGAGGAGTGCGTGGCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCAGG
GCCGATGGGGCAGTCGTGGTTGGGAACACAGCATTTCAAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTCCC
TCAAAGAGGCAGTTGCCCAGCCTCTTTCCCTTCCAGTTTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGT
TTTCTCTCTATTTCCCCCTTTCTTCCTCATTCCCTCGTCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGG

CAGCCTTGGCGGTTTATCGCCCTGGCAGGCAGGGGCCCTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGAC
AGGAGGTTGGAGGGAAAGCCTTAAGCTGCAGGATTCTCACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAAT
TTCAATTTTGTCTGTACTTGAACATTATGAAGATGGGGGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCGAC
AGCTTTCCAGTACCCATGGGGCTAGGTCATTAAGGCCACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTA
CTACCTCCTCTCCTGACAATACTGTATGTCGTCGAGCTCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGG
CTTGTCATAGCCAGTGGGATTGCCGGTCTTGACAGCTCAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGCAC
AGCTCCCTTCTGTTGATGCTGTATTCCCATATCAAAAGACACAGGGGACACCCAGAAACGCCACATCCCCCAATCCAT
CAGTGCCAAACTAGCCAACGGCCCCAGCTTCTCAGCTCGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGG
GTGCAGACAATCTTCTGTTGGGTGGCATCATTCCAGGCCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCCCC
AAATTGTCACCTGCTTCTCTGCCCAGCTTTTCATTGCTGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACAAA
CTGCCAAGTTGGGTGGAGAAAGGAGTTTCTTTAGCTGACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTCAA
GCCCAGGAGGGAGCAGAGGGATACGAGCGGAGTCCCCTGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAGCC
TGACAGCCAGAACTCTGTGTCCCCCGTCTAACCACAGCTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGAC
TGGGCCAGGGGAGGTTACAGGTACCAGTTCTTTAAGAAGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGCCC
CAAATGGATTCCTGTTTCAAGTTCACACCTGCAGATTCTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCT
AGAGTTCCTACCATGGAGTGGGTCTGGAGGACCTGCCCGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTAC
TCTTCTCTCTGCTCTGACGGGATTTGTTGATTCTCTCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTT
AGAAAAGGCATAGTCTACTTGTTATAAATCGTTAGGATACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGGGA
AAAGCTGAACACTGAAGTCAGTTCTCAACAATTTAGAAGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTCGA
TGTTTTTGAATGAATACGAGCAAGCTTTTACAACAGTGCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCTGG
TGAGCATTACAGGCAAGGGGAATCTGGAGGTAGCCGACCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTA
TGGAAGGTGGAGCGTTCACCAGTGACCTGGAAGGCCCAGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAGCC
TGCCCCAGCGCTGACGTGTCAGGAAAACACCCAGGGAACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCC
CACTCCTGCTCTGCTCGCCTCGGATCAGCTGAGCCTTCTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTG
CCCTGTCAGGAGGCAGAAGGAAGCAGGTGTGAGGGCAGTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCT
CCGACTTGTGCACAGGCAGAGCCCAGACCCTGGAGGAAATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGGAA
ATCTCTTTGCCCCCAAACCCCCATTCTGTCCTACCTTTAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAA
AGGCCAAGAGGTTTGGCTCCTGCCCACTGATAGCCCCTCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCT
TCCTGGTGACCCTGATTATATCCAGTAACACATAGACTGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTT
TGTTTTGCTTTTTAGTTTTGCTTTTAGTTTTTCTGTCCCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTTGA
ATTTTTATATATATATCTGTATATTGCACAATTATAAACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGACCT
GTTAAAATTATACCTGTTGCTTAATTACAATATTTCTGATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAAAA
AAAAGAAAAAAAAAACGACAAATCTGTCTGCTGGTCACTTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCT
TTCAAGGGCTTTCCTGTGCCAGGTGAAGGAGGCTCCAGGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTG
TGAAAGAGGTCCCAAGGTTCTGGGTGCAGGAGCGCTCCCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTG
GTCGCCTTCCACCTCTGGGAGCTGGAGTCCACTGGGGTGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGG
GTGAGCCCATGTGGAGTCAGCCTCGCAGGCCTCCCTGCCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCG
AGCCTGGGGGCTGGAGCGGAGACGGGAGGCCTGGCCTGTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGGGA
CCCCAGAATCATGTGCGTCAGTCCAAGGGGTCCCCTCCAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCC
CCATCCTACGAGTAATTGCATTTGCTTTTGTAATTCTTAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGTTC
TTTTTGATCATCTTTTTTTAATAATTAGAAACACCAAAAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATTAT
AATCACCAGGGCCAAAAGCTTCCCTCCCTGCTGTCATTGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCATA
GGCCCTTTCAGTGGCCGGGCTACCCGTGAGCCCTTCGGAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGG
CCAGCTCCGGCGTGTGTTCAGTGTTAGCAGTGGGTCATGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAG
GCGAGGAGGCCGTTGCCGCTGATGTTTGGCCGTGAACAGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACT
GACATGAAATCGACGCCCGAGTTAGCCTCACCCGGTGACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTC
AGTGTTTCTGGGGAGCTGGACAGTGGAGTGCAAAAGGCTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCACGG
CCTCCTTTCCGTTTGATTTGTCACTGCTTCAATCAATAACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGACCA
AATATCACCAGGACTGTTACTCAATGTGTGCCGAGCCCTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAACG
TGTGCTGTGTTTGCTCCCCTTCCCCTTCCTTCTTTGCCCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGT
TTGGTTTTTATTTCTCCTTTTGTGTTCCAAACATGAGGTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTA
TATTTGTGTAATTTTTGGTGGGTGAAAGGAATTTTGCTAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGT
AACTATGTTTAAAGTAATTGTTCCAGAGACAAATATTTCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGGGA
GGGGGATGGTGACTGAGATGAGAGGGGAGAGCTGAACAGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCA
GTGGAGCCCAGGAGTCCCACTCCAAGCCAGCAAGCCGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAACC
AGGTTTTGTTCCGCCCAGTGGATTCTTGTTTTGCTTCCCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTAAG
GAAAGGCAAGATTGATGTTTCCTTGAGGGGAGCCAGGAGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATG
GGGCTGGGCCCACCCAAGCAGGAGGCTGGGACGCTCTGCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGCTC
TTCCTGGACAGGCCAGGTGGTGGGCATTCTCTCTCCAAGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCGTC
ACATCCCACCCCATCCTCCAGGGCTCAACACTGTGACATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAATG
ACCATGGAGGGGGCTTGCACTCTCTTGGCTGTCACCCGATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCCAT
TCCATGGCGAAAACATCTCCTTAGAAAAGCCATTACCCTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACAGC
CCCTCCCTCCTCTGAGAGGCGGAGAGTGCTGACTGTAGTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGG

CAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCATGCTCCCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCATTG
CCAGTAACACAGGCCACCCTGTGCAGAAGGGAGCTGGCTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACT
TGGGGCACAGGGAGAGGCCGGGACACACTTAGCTGGAGATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTTTT
TGTGTTTTGGGACAATTACTTTAGAAAATAAGTAGGTCGTTTTAAAAACAAAAATTATTGATTGCTTTTTTGTAGTGT
TCAGAAAAAAGGTTCTTTGTGTATAGCCAAATGACTGAAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTAAG
GAAATTTGTACCATTTCAGTAAACCTGTCTGAATGTACCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATCCC
TGTAACCTATTTATTATATAAAGAGTTTGCCTTATAAATTTACATAAAAATGTCCGTTTGTGTCTTTTGTTGTAAAAA
TCAAGTGATTTTTTCATAAGGTTCTTTTACTATTGGAAAAGATGGGCAGCACGCAGTTTTATTTTATTTTTGTAAGTT
TTTTAATACATGTGAAAGCAAAGAATACTCAGCATGCCTTTCTAAGTGACGCGTTTGCACCTTTTGTTGGGAAGTACT
GTATCCTGTGCTGTTAGCATTCTCGATAAATCTCTCTGTGAAAGTGA

SEQ ID NO: 27 >NM_001369393.2 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=6]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTCACTGAGCTAGCCTCTCCTTGCTAGAAGCTTACGGTCGAAGG
TCTCCAGCCATCAGAAGAAGCACGTGGAGCGCTGCGTCCGTGTTGTGGTTATTCATCCAGCATGTGTTGAGGAAGAAA
AGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAGAAAGAAG
AGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATCAGCCCACCACTCTGCTGAGCCCGCAGAGGCAGGCAAAGCAGAGA
CATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCCGGAAGCTTCTGCCTCCCCCAAACAGCGGCGCTCCATCATCCGTG
ACCGGGGACCCATGTATGATGACCCCACCCTGCCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAATCTGGCCGCT
CTGCTGGGAAGTATGATGTGTATTTGATCAATCCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATTGCGTACT
TCGAAAAGGTAGGCGACACATCCCTGGACCCTAATGATTTTGACTTCACGGTAACTGGGAGAGGGAGCCCCTCCCGGC
GAGAGCAGAAACCACCTAAGAAGCCCAAATCTCCCAAAGCTCCAGGAACTGGCAGAGGCCGGGGACGCCCCAAAGGGA
GCGGCACCACGAGACCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCTGGGAAGC
TCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACCACATCCACCCAGGTCA
TGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCATTCCCAAGAAACGGGGCCGAAAGC
CGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGG
AGACCGTACTCCCCATCAAGAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGGTCAAGGAAGTGGTGAAGCCCCTGC
TGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGAAAAGCAAGGAGAGCA
GCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACCACCATCACCACCACTCAGAGTCCC
CAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCGAGAGCTCCGAGGACCCCACCAGCC
CCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCAGAGGAGGCTCACTGGAGAGCGACG
GCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGGCCGCAGAAAAGTACAAACACCGAG
GGGAGGGAGAGCGCAAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTGTGGACAGCCGGACGC
CCGTGACCGAGAGAGTTAGCTGACTTTACACGGAGCGGATTGCAAAGCAAACCAACAAGAATAAAGGCAGCTGTTGTC
TCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTTCCCGATTAACTGAAATAAAAAATATTTTTTTTTCTTTCAGTAAA
CTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTTCTTACCGACAAGCACAGTCAGGTT
GAAGACCTAACCAGGGCCAGAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTTCAACAAGGCTTGCTGCAGATACTA
CTGACCAGACAAGCTGTTGACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCCCATGTGGTCGTTAGAGACAGAGCG
ACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTACAGCTCCCCCAGCTCCCCCCACCTC
CCCCACTCCCAACCACGTTGGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGGATTCTTTAGAGAAGATGGATATGA
CCAGTGGCTATGGCCTGTGCGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCCAATCCAAAACTGGCA
AGGACGCTTCACAGGACAGGAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGGCTAGGAGGGGGGAGTCCCTTGAAC
TACTGGGTGTAGACTGGCCTGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGGCATGGTCCATTCTCAAGGGACGTC
CTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGGCTGGCCTGGGGTCAGGCCGGGCAG
AGCACAGCGGGGTGAGAGGGATTCCTAATCACTCAGAGCAGTCTGTGACTTAGTGGACAGGGGAGGGGGCAAAGGGGG
AGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTTAGAATTATTTGCACTATTGAGTCT
TCATGTTCCCACTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACATTTAGTCCCTCAAGCCA
CCAGATGTGACAGTGTTGAGAACTACCTGGATTTGTATATATACCTGCGCTTGTTTTAAAGTGGGCTCAGCACATAGG
GTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTCCTGATTGGTTTCTCTTCTCCCCTT
CCATTTCTGCCTTTTGTTCATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCCTCCTCCTTCCTAGTTCATCCCTTC
TCTTCCAGGCAGCCGCGGTGCCCAACCACACTTGTCGGCTCCAGTCCCCAGAACTCTGCCTGCCCTTTGTCCTCCTGC
TGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCTGAGTCCGACCTGGCCTGTCTGTGA
AGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAAGCAAAAGGCATTTCCC
ACCCTGAACAACGAGCCTTTTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGAGCTGGGCCCGATTTGGTAGTTGAG
GAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGCTCCATGCCAGCCGACCTTGACCTC
ACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGCAGAACTCCCAAGACCCGAGCTGGG
ACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAATGGCAATGTTTTGCTT
TTGCTCGATGCAGACAGGGGGCCAGAACACCACACATTTCACTGTCTGTCTGGTCCATAGCTGTGGTGTAGGGGCTTA
GAGGCATGGGCTTGCTGTGGGTTTTTAATTGATCAGTTTTCATGTGGGATCCCATCTTTTTAACCTCTGTTCAGGAAG
TCCTTATCTAGCTGCATATCTTCATCATATTGGTATATCCTTTTCTGTGTTTACAGAGATGTCTCTTATATCTAAATC

TGTCCAACTGAGAAGTACCTTATCAAAGTAGCAAATGAGACAGCAGTCTTATGCTTCCAGAAACACCCACAGGCATGT
CCCATGTGAGCTGCTGCCATGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCAGTTATTGTCCCTGGCTTCCTTACT
ATGGTGTAATCATGAAGGAGTGAAACATCATAGAAACTGTCTAGCACTTCCTTGCCAGTCTTTAGTGATCAGGAACCA
TAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGGAATGGTTAGAAGTGAGGGAGTTTG
CCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACTTTCTAGCATATATGTGTCCATTTCCTTATGCTGTAAAAGCAAGT
CCTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCACCTGCTCTGCTGATGACCCCCCCAGCTT
CACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCCTCCAGAACTCTTCCTCCAAGGACA
GAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCCTGGCACTACCAAAGGACACTTATCCACGAGAGCGCAGCATCCGA
CCAGGTTGTCACTGAGAAGATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTATACTTAGTCAAATGTAATGTGGCT
TCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCTGGTAAGAGGAGTGCGTGGCCCACC
AGGCCCCCCTGTCACCCATGACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGTTGGGAACACAGCATTTCAAGCGTC
ACTTTATTTCATTCGGGCCCCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAGCCTCTTTCCCTTCCAGTTTATTCC
AGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTTTCTTCCTCATTCCCTCGTCTTTCC
CAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGCCCTGGCAGGCAGGGGCCCTGCAGC
TCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCCTTAAGCTGCAGGATTCTCACCAGC
TGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAATTTCAATTTTGTCTGTACTTGAACATTATGAAGATGGGGGCCTCT
TTCAGTGAATTTGTGAACAGCAGAATTGACCGACAGCTTTCCAGTACCCATGGGGCTAGGTCATTAAGGCCACATCCA
CAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTACTACCTCCTCTCCTGACAATACTGTATGTCGTCGAGCTCCCCCC
AGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGATTGCCGGTCTTGACAGCTCAGTGAG
CTGGAGATACTTGGTCACAGCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCTGTATTCCCATATCAAAAGACACAG
GGGACACCCAGAAACGCCACATCCCCCAATCCATCAGTGCCAAACTAGCCAACGGCCCCAGCTTCTCAGCTCGCTGGA
TGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTGGGTGGCATCATTCCAGGCCCGAAG
CATGAACAGTGCACCTGGGACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCTGCCCAGCTTTTCATTGCTGTGACA
GTGATGGCGAAAGAGGGTAATAACCAGACACAAACTGCCAAGTTGGGTGGAGAAAGGAGTTTCTTTAGCTGACAGAAT
CTCTGAATTTTAAATCACTTAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGGATACGAGCGGAGTCCCCTGCGCGG
GACCATCTGGAATTGGTTTAGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGTCCCCCGTCTAACCACAGCTCCTTT
TCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAGGTACCAGTTCTTTAAGAAGATCTT
TGGGCATATACATTTTTAGCCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAAGTTCACACCTGCAGATTCTAGGAC
CTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTGGGTCTGGAGGACCTGCCCGGTGGG
GGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGGGATTTGTTGATTCTCTCCATTTTG
GTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAAAAGGCATAGTCTACTTGTTATAAATCGTTAGGATACTGCC
TCCCCCAGGGTCTAAAATTACATATTAGAGGGGAAAAGCTGAACACTGAAGTCAGTTCTCAACAATTTAGAAGGAAAA
CCTAGAAAACATTTGGCAGAAAATTACATTTCGATGTTTTTGAATGAATACGAGCAAGCTTTTACAACAGTGCTGATC
TAAAAATACTTAGCACTTGGCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGGAATCTGGAGGTAGCCGACCTGAGG
ACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACCAGTGACCTGGAAGGCCCAGCACCA
CCCTCCTTCCCACTCTTCTCATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAACACCCAGGGAACTAGGA
AGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCTCGGATCAGCTGAGCCTTCTGAGCT
GGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGGAAGCAGGTGTGAGGGCAGTGCAAG
GAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGAGCCCAGACCCTGGAGGAAATCCTA
CCTTTGAATTCAAGAACATTTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCCCCATTCTGTCCTACCTTTAATCAG
GTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCCTGCCCACTGATAGCCCCTCTCCCC
GCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATATCCAGTAACACATAGACTGTGCGC
ATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTGCTTTTAGTTTTTCTGTCCCTTTTA
TTTAACGCACCGACTAGACACACAAAGCAGTTGAATTTTTATATATATATCTGTATATTGCACAATTATAAACTCATT
TTGCTTGTGGCTCCACACACACAAAAAAAGACCTGTTAAAATTATACCTGTTGCTTAATTACAATATTTCTGATAACC
ATAGCATAGGACAAGGGAAAATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAAATCTGTCTGCTGGTCACTTCTTCT
GTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCCAGGTGAAGGAGGCTCCAGGCAGCA
CCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTCTGGGTGCAGGAGCGCTCCCTTGAC
CTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGAGCTGGAGTCCACTGGGGTGGCCTG
ACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAGCCTCGCAGGCCTCCCTGCCAGTAG
GGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGAGACGGGAGGCCTGGCCTGTCTCGG
AACCTGTGAGCTGCACCAGGTAGAACGCCAGGGACCCCAGAATCATGTGCGTCAGTCCAAGGGGTCCCCTCCAGGAGT
AGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCATTTGCTTTTGTAATTCTTAATGAG
CAATATCTGCTAGAGAGTTTAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTAATAATTAGAAACACCAAAAAAATC
CAGAAACTTGTTCTTCCAAAGCAGAGAGCATTATAATCACCAGGGCCAAAAGCTTCCCTCCCTGCTGTCATTGCTTCT
TCTGAGGCCTGAATCCAAAAGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGCTACCCGTGAGCCCTTCGGAGGACC
AGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCAGTGTTAGCAGTGGGTCATGATGCT
CTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCTGATGTTTGGCCGTGAACAGGTGGG
TGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACTGACATGAAATCGACGCCCGAGTTAGCCTCACCCGGTGACCTCTA
GCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGACAGTGGAGTGCAAAAGGCTTGCAG
AACTTGAAGCCTGCTCCTTCCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTGTCACTGCTTCAATCAATAACAGCC
GCTCCAGAGTCAGTAGTCAATGAATATATGACCAAATATCACCAGGACTGTTACTCAATGTGTGCCGAGCCCTTGCCC
ATGCTGGGCTCCCGTGTATCTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCTTCCCCTTCCTTCTTTGCCCTTTAC

TTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTTTGTGTTCCAAACATGAGGTTCTCT
CTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTGGGTGAAAGGAATTTTGCTAAGTAA
ATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGTAACTATGTTTAAAGTAATTGTTCCAGAGACAAATATTTCTAGAC
ACTTTTTCTTTACAAACAAAAGCATTCGGAGGGAGGGGGATGGTGACTGAGATGAGAGGGGAGAGCTGAACAGATGAC
CCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCACTCCAAGCCAGCAAGCCGAATAGCT
GATGTGTTGCCACTTTCCAAGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTGGATTCTTGTTTTGCTTCCCCTCCC
CCCGAGATTATTACCACCATCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTTCCTTGAGGGGAGCCAGGAGGGGAT
GTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCAGGAGGCTGGGACGCTCTGCTGTGG
GCACAGGTCAGGCTAATGTTGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGGTGGGCATTCTCTCTCCAAGGTGTG
CCCCGTGGGCATTACTGTTTAAGACACTTCCGTCACATCCCACCCCATCCTCCAGGGCTCAACACTGTGACATCTCTA
TTCCCCACCCTCCCCTTCCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGGCTGTCACCCGATCGCC
AGCAAAACTTAGATGTGAGAAAACCCCTTCCCATTCCATGGCGAAAACATCTCCTTAGAAAAGCCATTACCCTCATTA
GGCATGGTTTTGGGCTCCCAAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGCGGAGAGTGCTGACTGTAGTGACCA
TTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCATGCTCCCCTGTG
CCAGCCCAGAGGCCGAGAGCTATGGACAGCATTGCCAGTAACACAGGCCACCCTGTGCAGAAGGGAGCTGGCTCCAGC
CTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCGGGACACACTTAGCTGGAGATGTCT
CTAAAAGCCCTGTATCGTATTCACCTTCAGTTTTTGTGTTTTGGGACAATTACTTTAGAAAATAAGTAGGTCGTTTTA
AAAACAAAAATTATTGATTGCTTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGTGTATAGCCAAATGACTGAAAGCAC
TGATATATTTAAAAACAAAAGGCAATTTATTAAGGAAATTTGTACCATTTCAGTAAACCTGTCTGAATGTACCTGTAT
ACGTTTCAAAAACACCCCCCCCCCACTGAATCCCTGTAACCTATTTATTATATAAAGAGTTTGCCTTATAAATTTACA
TAAAAATGTCCGTTTGTGTCTTTTGTTGTAAAAATCAAGTGATTTTTTCATAAGGTTCTTTTACTATTGGAAAAGATG
GGCAGCACGCAGTTTTATTTTATTTTTGTAAGTTTTTTAATACATGTGAAAGCAAAGAATACTCAGCATGCCTTTCTA
AGTGACGCGTTTGCACCTTTTGTTGGGAAGTACTGTATCCTGTGCTGTTAGCATTCTCGATAAATCTCTCTGTGAAAG
TGA

SEQ ID NO: 28 >NM_001386138.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=9]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTGCTCCATAAAAATACAGACTCACCAGTTCCTGCTTTGATGTG
ACATGTGACTCCCCAGAATACACCTTGCTTCTGTAGACCAGCTCCAACAGGATTCCATGGTAGCTGGGATGTTAGGGC
TCAGGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAG
ATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATCAGCCCACCACTCTGCTGAGCCCGCAGAGGCAG
GCAAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCCGGAAGCTTCTGCCTCCCCCAAACAGCGGCGCT
CCATCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTGCCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGA
AATCTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAATCCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGT
TGATTGCGTACTTCGAAAAGCTCCAGGAACTGGCAGAGGCCGGGGACGCCCCAAAGGGAGCGGCACCACGAGACCCAA
GGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCTGGGAAGCTCCTTGTCAAGATGCCTTT
TCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACCACATCCACCCAGGTCATGGTGATCAAACGCCCCGG
CAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCATTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGC
CGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGGAGACCGTACTCCCCATCAA
GAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGGTCAAGGAAGTGGTGAAGCCCCTGCTGGTGTCCACCCTCGGTGA
GAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGAAAAGCAAGGAGAGCAGCCCCAAGGGGCGCAGCAG
CAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACCACCATCACCACCACTCAGAGTCCCCAAAGGCCCCCGTGCCACT
GCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCGAGAGCTCCGAGGACCCCACCAGCCCCCCTGAGCCCCAGGACTT
GAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCAGAGGAGGCTCACTGGAGAGCGACGGCTGCCCCAAGGAGCCAGC
TAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGGCCGCAGAAAAGTACAAACACCGAGGGGAGGGAGAGCGCAAAGA
CATTGTTTCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTGTGGACAGCCGGACGCCCGTGACCGAGAGAGTTAG
CTGACTTTACACGGAGCGGATTGCAAAGCAAACCAACAAGAATAAAGGCAGCTGTTGTCTCTTCTCCTTATGGGTAGG
GCTCTGACAAAGCTTCCCGATTAACTGAAATAAAAAATATTTTTTTTTCTTTCAGTAAACTTAGAGTTTCGTGGCTTC
AGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTTCTTACCGACAAGCACAGTCAGGTTGAAGACCTAACCAGGGCCA
GAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTTCAACAAGGCTTGCTGCAGATACTACTGACCAGACAAGCTGTTG
ACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCCCATGTGGTCGTTAGAGACAGAGCGACAGAGCAGTTGAGAGGAC
ACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTACAGCTCCCCCAGCTCCCCCCACCTCCCCCACTCCCAACCACGTT
GGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGGATTCTTTAGAGAAGATGGATATGACCAGTGGCTATGGCCTGTG
CGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCCAATCCAAAACTGGCAAGGACGCTTCACAGGACAG
GAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGGCTAGGAGGGGGGAGTCCCTTGAACTACTGGGTGTAGACTGGCC
TGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGGCATGGTCCATTCTCAAGGGACGTCCTCCAACGGGTGGCGCTAG
AGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGGCTGGCCTGGGGTCAGGCCGGGCAGAGCACAGCGGGGTGAGAGG
GATTCCTAATCACTCAGAGCAGTCTGTGACTTAGTGGACAGGGGAGGGGGCAAAGGGGGAGGAGAAGAAAATGTTCTT
CCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTTAGAATTATTTGCACTATTGAGTCTTCATGTTCCCACTTCAAAA
CAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACATTTAGTCCCTCAAGCCACCAGATGTGACAGTGTTGA

GAACTACCTGGATTTGTATATATACCTGCGCTTGTTTTAAAGTGGGCTCAGCACATAGGGTTCCCACGAAGCTCCGAA
ACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTCCTGATTGGTTTCTCTTCTCCCCTTCCATTTCTGCCTTTTGTTC
ATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCCTCCTCCTTCCTAGTTCATCCCTTCTCTTCCAGGCAGCCGCGGT
GCCCAACCACACTTGTCGGCTCCAGTCCCCAGAACTCTGCCTGCCCTTTGTCCTCCTGCTGCCAGTACCAGCCCCACC
CTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCTGAGTCCGACCTGGCCTGTCTGTGAAGAGCAAGAGAGCAGCAAG
GTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAAGCAAAAGGCATTTCCCACCCTGAACAACGAGCCTT
TTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGAGCTGGGCCCGATTTGGTAGTTGAGGAAAGCACAGAGGCCTCCT
GTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGCTCCATGCCAGCCGACCTTGACCTCACTCAGAAGTCCAGAGTCT
AGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGCAGAACTCCCAAGACCCGAGCTGGGACCAGTACCTGGGTCCCCA
GCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAATGGCAATGTTTTGCTTTTGCTCGATGCAGACAGGG
GGCCAGAACACCACACATTTCACTGTCTGTCTGGTCCATAGCTGTGGTGTAGGGGCTTAGAGGCATGGGCTTGCTGTG
GGTTTTTAATTGATCAGTTTTCATGTGGGATCCCATCTTTTTAACCTCTGTTCAGGAAGTCCTTATCTAGCTGCATAT
CTTCATCATATTGGTATATCCTTTTCTGTGTTTACAGAGATGTCTCTTATATCTAAATCTGTCCAACTGAGAAGTACC
TTATCAAAGTAGCAAATGAGACAGCAGTCTTATGCTTCCAGAAACACCCACAGGCATGTCCCATGTGAGCTGCTGCCA
TGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCAGTTATTGTCCCTGGCTTCCTTACTATGGTGTAATCATGAAGGA
GTGAAACATCATAGAAACTGTCTAGCACTTCCTTGCCAGTCTTTAGTGATCAGGAACCATAGTTGACAGTTCCAATCA
GTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGGAATGGTTAGAAGTGAGGGAGTTTGCCCCGTTCTGTTTGTAGAG
TCTCATAGTTGGACTTTCTAGCATATATGTGTCCATTTCCTTATGCTGTAAAAGCAAGTCCTGCAACCAAACTCCCAT
CAGCCCAATCCCTGATCCCTGATCCCTTCCACCTGCTCTGCTGATGACCCCCCCAGCTTCACTTCTGACTCTTCCCCA
GGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCCTCCAGAACTCTTCCTCCAAGGACAGAAGGCTCCTGCCCCCATA
GTGGCCTCGAACTCCTGGCACTACCAAAGGACACTTATCCACGAGAGCGCAGCATCCGACCAGGTTGTCACTGAGAAG
ATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTATACTTAGTCAAATGTAATGTGGCTTCTGGAATCATTGTCCAGA
GCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCTGGTAAGAGGAGTGCGTGGCCCACCAGGCCCCCCTGTCACCCAT
GACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGTTGGGAACACAGCATTTCAAGCGTCACTTTATTTCATTCGGGCC
CCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAGCCTCTTTCCCTTCCAGTTTATTCCAGAGCTGCCAGTGGGGCCT
GAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTTTCTTCCTCATTCCCTCGTCTTTCCCAAAGGCATCACGAGTCAG
TCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGCCCTGGCAGGCAGGGGCCCTGCAGCTCTCATGCTGCCCCTGCCT
TGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCCTTAAGCTGCAGGATTCTCACCAGCTGTGTCCGGCCCAGTTTTG
GGGTGTGACCTCAATTTCAATTTTGTCTGTACTTGAACATTATGAAGATGGGGGCCTCTTTCAGTGAATTTGTGAACA
GCAGAATTGACCGACAGCTTTCCAGTACCCATGGGGCTAGGTCATTAAGGCCACATCCACAGTCTCCCCCACCCTTGT
TCCAGTTGTTAGTTACTACCTCCTCTCCTGACAATACTGTATGTCGTCGAGCTCCCCCCAGGTCTACCCCTCCCGGCC
CTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGATTGCCGGTCTTGACAGCTCAGTGAGCTGGAGATACTTGGTCACA
GCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCTGTATTCCCATATCAAAAGACACAGGGGACACCCAGAAACGCCA
CATCCCCCAATCCATCAGTGCCAAACTAGCCAACGGCCCCAGCTTCTCAGCTCGCTGGATGGCGGAAGCTGCTACTCG
TGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTGGGTGGCATCATTCCAGGCCCGAAGCATGAACAGTGCACCTGGG
ACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCTGCCCAGCTTTTCATTGCTGTGACAGTGATGGCGAAAGAGGGTA
ATAACCAGACACAAACTGCCAAGTTGGGTGGAGAAAGGAGTTTCTTTAGCTGACAGAATCTCTGAATTTTAAATCACT
TAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGGATACGAGCGGAGTCCCCTGCGCGGGACCATCTGGAATTGGTTT
AGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGTCCCCCGTCTAACCACAGCTCCTTTTCCAGAGCATTCCAGTCAG
GCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAGGTACCAGTTCTTTAAGAAGATCTTTGGGCATATACATTTTTAG
CCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAAGTTCACACCTGCAGATTCTAGGACCTGTGTCCTAGACTTCAGG
GAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTGGGTCTGGAGGACCTGCCCGGTGGGGGGGCAGAGCCCTGCTCCC
TCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGGGATTTGTTGATTCTCTCCATTTTGGTGTCTTTCTCTTTTAGAT
ATTGTATCAATCTTTAGAAAAGGCATAGTCTACTTGTTATAAATCGTTAGGATACTGCCTCCCCCAGGGTCTAAAATT
ACATATTAGAGGGGAAAAGCTGAACACTGAAGTCAGTTCTCAACAATTTAGAAGGAAAACCTAGAAAACATTTGGCAG
AAAATTACATTTCGATGTTTTTGAATGAATACGAGCAAGCTTTTACAACAGTGCTGATCTAAAAATACTTAGCACTTG
GCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGGAATCTGGAGGTAGCCGACCTGAGGACATGGCTTCTGAACCTGT
CTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACCAGTGACCTGGAAGGCCCAGCACCACCCTCCTTCCCACTCTTCT
CATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAACACCCAGGGAACTAGGAAGGCACTTCTGCCTGAGGG
GCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCTCGGATCAGCTGAGCCTTCTGAGCTGGCCTCTCACTGCCTCCCC
AAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGGAAGCAGGTGTGAGGGCAGTGCAAGGAGGGAGCACAACCCCCAG
CTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGAGCCCAGACCCTGGAGGAAATCCTACCTTTGAATTCAAGAACAT
TTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCCCCATTCTGTCCTACCTTTAATCAGGTCCTGCTCAGCAGTGAGA
GCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCCTGCCCACTGATAGCCCCTCTCCCCGCAGTGTTTGTGTGTCAAG
TGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATATCCAGTAACACATAGACTGTGCGCATAGGCCTGCTTTGTCTCC
TCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTGCTTTTAGTTTTTCTGTCCCTTTTATTTAACGCACCGACTAGAC
ACACAAAGCAGTTGAATTTTTATATATATATCTGTATATTGCACAATTATAAACTCATTTTGCTTGTGGCTCCACACA
CACAAAAAAAGACCTGTTAAAATTATACCTGTTGCTTAATTACAATATTTCTGATAACCATAGCATAGGACAAGGGAA
AATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAAATCTGTCTGCTGGTCACTTCTTCTGTCCAAGCAGATTCGTGGT
CTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCCAGGTGAAGGAGGCTCCAGGCAGCACCCAGGTTTTGCACTCTTG
TTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTCTGGGTGCAGGAGCGCTCCCTTGACCTGCTGAAGTCCGGAACGT
AGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGAGCTGGAGTCCACTGGGGTGGCCTGACTCCCCCAGTCCCCTTCC
CGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAGCCTCGCAGGCCTCCCTGCCAGTAGGGTCCGAGTGTGTTTCATC

CTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGAGACGGGAGGCCTGGCCTGTCTCGGAACCTGTGAGCTGCACCAG
GTAGAACGCCAGGGACCCCAGAATCATGTGCGTCAGTCCAAGGGGTCCCCTCCAGGAGTAGTGAAGACTCCAGAAATG
TCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCATTTGCTTTTGTAATTCTTAATGAGCAATATCTGCTAGAGAGTT
TAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTAATAATTAGAAACACCAAAAAAATCCAGAAACTTGTTCTTCCAA
AGCAGAGAGCATTATAATCACCAGGGCCAAAAGCTTCCCTCCCTGCTGTCATTGCTTCTTCTGAGGCCTGAATCCAAA
AGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGCTACCCGTGAGCCCTTCGGAGGACCAGGGCTGGGGCAGCCTCTG
GGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCAGTGTTAGCAGTGGGTCATGATGCTCTTTCCCACCCAGCCTGGG
ATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCTGATGTTTGGCCGTGAACAGGTGGGTGTCTGCGTGCGTCCACGT
GCGTGTTTTCTGACTGACATGAAATCGACGCCCGAGTTAGCCTCACCCGGTGACCTCTAGCCCTGCCCGGATGGAGCG
GGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGACAGTGGAGTGCAAAAGGCTTGCAGAACTTGAAGCCTGCTCCTT
CCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTGTCACTGCTTCAATCAATAACAGCCGCTCCAGAGTCAGTAGTCA
ATGAATATATGACCAAATATCACCAGGACTGTTACTCAATGTGTGCCGAGCCCTTGCCCATGCTGGGCTCCCGTGTAT
CTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCTTCCCCTTCCTTCTTTGCCCTTTACTTGTCTTTCTGGGGTTTTT
CTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTTTGTGTTCCAAACATGAGGTTCTCTCTACTGGTCCTCTTAACTG
TGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTGGGTGAAAGGAATTTTGCTAAGTAAATCTCTTCTGTGTTTGAAC
TGAAGTCTGTATTGTAACTATGTTTAAAGTAATTGTTCCAGAGACAAATATTTCTAGACACTTTTTCTTTACAAACAA
AAGCATTCGGAGGGAGGGGGATGGTGACTGAGATGAGAGGGGAGAGCTGAACAGATGACCCCTGCCCAGATCAGCCAG
AAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCACTCCAAGCCAGCAAGCCGAATAGCTGATGTGTTGCCACTTTCCA
AGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTGGATTCTTGTTTTGCTTCCCCTCCCCCCGAGATTATTACCACCA
TCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTTCCTTGAGGGGAGCCAGGAGGGGATGTGTGTGTGCAGAGCTGAA
GAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCAGGAGGCTGGGACGCTCTGCTGTGGGCACAGGTCAGGCTAATGT
TGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGGTGGGCATTCTCTCTCCAAGGTGTGCCCCGTGGGCATTACTGTT
TAAGACACTTCCGTCACATCCCACCCCATCCTCCAGGGCTCAACACTGTGACATCTCTATTCCCCACCCTCCCCTTCC
CAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGGCTGTCACCCGATCGCCAGCAAAACTTAGATGTGAG
AAAACCCCTTCCCATTCCATGGCGAAAACATCTCCTTAGAAAAGCCATTACCCTCATTAGGCATGGTTTTGGGCTCCC
AAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGCGGAGAGTGCTGACTGTAGTGACCATTGCATGCCGGGTGCAGCA
TCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCATGCTCCCCTGTGCCAGCCCAGAGGCCGAGAG
CTATGGACAGCATTGCCAGTAACACAGGCCACCCTGTGCAGAAGGGAGCTGGCTCCAGCCTGGAAACCTGTCTGAGGT
TGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCGGGACACACTTAGCTGGAGATGTCTCTAAAAGCCCTGTATCGTA
TTCACCTTCAGTTTTTGTGTTTTGGGACAATTACTTTAGAAAATAAGTAGGTCGTTTTAAAAACAAAAATTATTGATT
GCTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGTGTATAGCCAAATGACTGAAAGCACTGATATATTTAAAAACAAA
AGGCAATTTATTAAGGAAATTTGTACCATTTCAGTAAACCTGTCTGAATGTACCTGTATACGTTTCAAAAACACCCCC
CCCCCACTGAATCCCTGTAACCTATTTATTATATAAAGAGTTTGCCTTATAAATTTACATAAAAATGTCCGTTTGTGT
CTTTTGTTGTAAAAATCAAGTGATTTTTTCATAAGGTTCTTTTACTATTGGAAAAGATGGGCAGCACGCAGTTTTATT
TTATTTTTGTAAGTTTTTTAATACATGTGAAAGCAAAGAATACTCAGCATGCCTTTCTAAGTGACGCGTTTGCACCTT
TTGTTGGGAAGTACTGTATCCTGTGCTGTTAGCATTCTCGATAAATCTCTCTGTGAAAGTGA

SEQ ID NO: 29 >NM_004992.4 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=1]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTGCTCCATAAAAATACAGACTCACCAGTTCCTGCTTTGATGTG
ACATGTGACTCCCCAGAATACACCTTGCTTCTGTAGACCAGCTCCAACAGGATTCCATGGTAGCTGGGATGTTAGGGC
TCAGGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAG
ATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATCAGCCCACCACTCTGCTGAGCCCGCAGAGGCAG
GCAAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCCGGAAGCTTCTGCCTCCCCCAAACAGCGGCGCT
CCATCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTGCCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGA
AATCTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAATCCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGT
TGATTGCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCTAATGATTTTGACTTCACGGTAACTGGGAGAGGGA
GCCCCTCCCGGCGAGAGCAGAAACCACCTAAGAAGCCCAAATCTCCCAAAGCTCCAGGAACTGGCAGAGGCCGGGGAC
GCCCCAAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAA
GTCCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACCACAT
CCACCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCATTCCCAAGAAAC
GGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCC
GATCTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGGTCAAGGAAGTGG
TGAAGCCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGAAAA
GCAAGGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACCACCATCACCACC
ACTCAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCGAGAGCTCCGAGG
ACCCCACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCAGAGGAGGCTCAC
TGGAGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGGCCGCAGAAAAGT
ACAAACACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTGTGG
ACAGCCGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACGGAGCGGATTGCAAAGCAAACCAACAAGAATAAAG

GCAGCTGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTTCCCGATTAACTGAAATAAAAAATATTTTTTTT
TCTTTCAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTTCTTACCGACAAG
CACAGTCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTTCAACAAGGCTTG
CTGCAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCCCATGTGGTCGTT
AGAGACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTACAGCTCCCCCAGC
TCCCCCCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGGATTCTTTAGAGA
AGATGGATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCCAAT
CCAAAACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGGCTAGGAGGGGGG
AGTCCCTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGGCATGGTCCATTC
TCAAGGGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGGCTGGCCTGGGGT
CAGGCCGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCACTCAGAGCAGTCTGTGACTTAGTGGACAGGGGAGG
GGGCAAAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTTAGAATTATTTGC
ACTATTGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACATTTAG
TCCCTCAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGATTTGTATATATACCTGCGCTTGTTTTAAAGTGGGC
TCAGCACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTCCTGATTGGTTTC
TCTTCTCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCCTCCTCCTTCCTA
GTTCATCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACTTGTCGGCTCCAGTCCCCAGAACTCTGCCTGCCCT
TTGTCCTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCTGAGTCCGACCTG
GCCTGTCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAAGCAA
AAGGCATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGAGCTGGGCCCGAT
TTGGTAGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGCTCCATGCCAGCC
GACCTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGCAGAACTCCCAAG
ACCCGAGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAATGGC
AATGTTTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCACACATTTCACTGTCTGTCTGGTCCATAGCTGTGG
TGTAGGGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGATCAGTTTTCATGTGGGATCCCATCTTTTTAACCT
CTGTTCAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTGGTATATCCTTTTCTGTGTTTACAGAGATGTCTCT
TATATCTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGCAAATGAGACAGCAGTCTTATGCTTCCAGAAACAC
CCACAGGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCAGTTATTGTCCCT
GGCTTCCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATAGAAACTGTCTAGCACTTCCTTGCCAGTCTTTAGT
GATCAGGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGGAATGGTTAGAAG
TGAGGGAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACTTTCTAGCATATATGTGTCCATTTCCTTATGCT
GTAAAAGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCACCTGCTCTGCTGATGA
CCCCCCCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCCTCCAGAACTCTT
CCTCCAAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCCTGGCACTACCAAAGGACACTTATCCACGAGAG
CGCAGCATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTATACTTAGTCAAA
TGTAATGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCTGGTAAGAGGAGT
GCGTGGCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGTTGGGAACACAGC
ATTTCAAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAGCCTCTTTCCCTT
CCAGTTTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTTTCTTCCTCATTC
CCTCGTCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGCCCTGGCAGGCAG
GGGCCCTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCCTTAAGCTGCAGG
ATTCTCACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAATTTCAATTTTGTCTGTACTTGAACATTATGAAG
ATGGGGGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCGACAGCTTTCCAGTACCCATGGGGCTAGGTCATTA
AGGCCACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTACTACCTCCTCTCCTGACAATACTGTATGTCGT
CGAGCTCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGATTGCCGGTCTTGA
CAGCTCAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCTGTATTCCCATAT
CAAAAGACACAGGGGACACCCAGAAACGCCACATCCCCCAATCCATCAGTGCCAAACTAGCCAACGGCCCCAGCTTCT
CAGCTCGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTGGGTGGCATCATT
CCAGGCCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCTGCCCAGCTTTTC
ATTGCTGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACAAACTGCCAAGTTGGGTGGAGAAAGGAGTTTCTTT
AGCTGACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGGATACGAGCGGAG
TCCCCTGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGTCCCCCGTCTAAC
CACAGCTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAGGTACCAGTTCTT
TAAGAAGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAAGTTCACACCTGC
AGATTCTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTGGGTCTGGAGGAC
CTGCCCGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGGGATTTGTTGATT
CTCTCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAAAAGGCATAGTCTACTTGTTATAAATCGT
TAGGATACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGGGAAAAGCTGAACACTGAAGTCAGTTCTCAACAAT
TTAGAAGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTCGATGTTTTTGAATGAATACGAGCAAGCTTTTACA
ACAGTGCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGGAATCTGGAGGTA
GCCGACCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACCAGTGACCTGGAA
GGCCCAGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAACACCC
AGGGAACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCTCGGATCAGCTGA

GCCTTCTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGGAAGCAGGTGTGA
GGGCAGTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGAGCCCAGACCCTG
GAGGAAATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCCCCATTCTGTCCT
ACCTTTAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCCTGCCCACTGATA
GCCCCTCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATATCCAGTAACACA
TAGACTGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTGCTTTTAGTTTTT
CTGTCCCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTTGAATTTTTATATATATATCTGTATATTGCACAAT
TATAAACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGACCTGTTAAAATTATACCTGTTGCTTAATTACAATA
TTTCTGATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAAATCTGTCTGCTG
GTCACTTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCCAGGTGAAGGAGG
CTCCAGGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTCTGGGTGCAGGAG
CGCTCCCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGAGCTGGAGTCCAC
TGGGGTGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAGCCTCGCAGGCCT
CCCTGCCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGAGACGGGAGGCCT
GGCCTGTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGGGACCCCAGAATCATGTGCGTCAGTCCAAGGGGTC
CCCTCCAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCATTTGCTTTTGTA
ATTCTTAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTAATAATTAGAAAC
ACCAAAAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATTATAATCACCAGGGCCAAAAGCTTCCCTCCCTGCT
GTCATTGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGCTACCCGTGAGCC
CTTCGGAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCAGTGTTAGCAGTG
GGTCATGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCTGATGTTTGGCCG
TGAACAGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACTGACATGAAATCGACGCCCGAGTTAGCCTCACC
CGGTGACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGACAGTGGAGTGCA
AAAGGCTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTGTCACTGCTTCAA
TCAATAACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGACCAAATATCACCAGGACTGTTACTCAATGTGTGCC
GAGCCCTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCTTCCCCTTCCTTC
TTTGCCCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTTTGTGTTCCAAAC
ATGAGGTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTGGGTGAAAGGAAT
TTTGCTAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGTAACTATGTTTAAAGTAATTGTTCCAGAGACAA
ATATTTCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGGGAGGGGGATGGTGACTGAGATGAGAGGGGAGAGC
TGAACAGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCACTCCAAGCCAGCA
AGCCGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTGGATTCTTGTTTT
GCTTCCCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTTCCTTGAGGGGAG
CCAGGAGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCAGGAGGCTGGGAC
GCTCTGCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGGTGGGCATTCTCT
CTCCAAGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCGTCACATCCCACCCCATCCTCCAGGGCTCAACACT
GTGACATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGGCTGT
CACCCGATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCCATTCCATGGCGAAAACATCTCCTTAGAAAAGCCA
TTACCCTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGCGGAGAGTGCTGA
CTGTAGTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCA
TGCTCCCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCATTGCCAGTAACACAGGCCACCCTGTGCAGAAGGGA
GCTGGCTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCGGGACACACTTAG
CTGGAGATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTTTTTGTGTTTTGGGACAATTACTTTAGAAAATAAG
TAGGTCGTTTTAAAAACAAAAATTATTGATTGCTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGTGTATAGCCAAAT
GACTGAAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTAAGGAAATTTGTACCATTTCAGTAAACCTGTCTGA
ATGTACCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATCCCTGTAACCTATTTATTATATAAAGAGTTTGCCT
TATAAATTTACATAAAAATGTCCGTTTGTGTCTTTTGTTGTAAAAATCAAGTGATTTTTTCATAAGGTTCTTTTACTA
TTGGAAAAGATGGGCAGCACGCAGTTTTATTTTATTTTTGTAAGTTTTTTAATACATGTGAAAGCAAAGAATACTCAG
CATGCCTTTCTAAGTGACGCGTTTGCACCTTTTGTTGGGAAGTACTGTATCCTGTGCTGTTAGCATTCTCGATAAATC
TCTCTGTGAAAGTGA

SEQ ID NO: 30 >NM_001386137.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=8]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTGCTCCATAAAAATACAGACTCACCAGTTCCTGCTTTGATGTG
ACATGTGACTCCCCAGAATACACCTTGCTTCTGTAGACCAGCTCCAACAGGATTCCATGGTAGCTGGGATGTTAGGGC
TCAGCACTGAGCTAGCCTCTCCTTGCTAGAAGCTTACGGTCGAAGGTCTCCAGCCATCAGAAGAAGCACGTGGAGCGC
TGCGTCCGTGTTGTGGTTATTCATCCAGCATGTGTTGAGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAG
GACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCA
TCAGCCCACCACTCTGCTGAGCCCGCAGAGGCAGGCAAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTG
CCGGAAGCTTCTGCCTCCCCCAAACAGCGGCGCTCCATCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTG

CCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAATCTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAAT
CCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGCTCCAGGAACTGGCAGAGGCCGGG
GACGCCCCAAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGGTCCTGGAGA
AAAGTCCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACCA
CATCCACCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCATTCCCAAGA
AACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTA
TCCGATCTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGGTCAAGGAAG
TGGTGAAGCCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGA
AAAGCAAGGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACCACCATCACC
ACCACTCAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCGAGAGCTCCG
AGGACCCCACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCAGAGGAGGCT
CACTGGAGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGGCCGCAGAAA
AGTACAAACACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTG
TGGACAGCCGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACGGAGCGGATTGCAAAGCAAACCAACAAGAATA
AAGGCAGCTGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTTCCCGATTAACTGAAATAAAAAATATTTTT
TTTTCTTTCAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTTCTTACCGAC
AAGCACAGTCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTTCAACAAGGC
TTGCTGCAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCCCATGTGGTC
GTTAGAGACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTACAGCTCCCCC
AGCTCCCCCCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGGATTCTTTAG
AGAAGATGGATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCC
AATCCAAAACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGGCTAGGAGGG
GGGAGTCCCTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGGCATGGTCCA
TTCTCAAGGGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGGCTGGCCTGG
GGTCAGGCCGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCACTCAGAGCAGTCTGTGACTTAGTGGACAGGGG
AGGGGGCAAAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTTAGAATTATT
TGCACTATTGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACATT
TAGTCCCTCAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGATTTGTATATATACCTGCGCTTGTTTTAAAGTG
GGCTCAGCACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTCCTGATTGGT
TTCTCTTCTCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCCTCCTCCTTC
CTAGTTCATCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACTTGTCGGCTCCAGTCCCCAGAACTCTGCCTGC
CCTTTGTCCTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCTGAGTCCGAC
CTGGCCTGTCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAAG
CAAAAGGCATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGAGCTGGGCCC
GATTTGGTAGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGCTCCATGCCA
GCCGACCTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGCAGAACTCCC
AAGACCCGAGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAAT
GGCAATGTTTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCACACATTTCACTGTCTGTCTGGTCCATAGCTG
TGGTGTAGGGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGATCAGTTTTCATGTGGGATCCCATCTTTTTAA
CCTCTGTTCAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTGGTATATCCTTTTCTGTGTTTACAGAGATGTC
TCTTATATCTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGCAAATGAGACAGCAGTCTTATGCTTCCAGAAA
CACCCACAGGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCAGTTATTGTC
CCTGGCTTCCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATAGAAACTGTCTAGCACTTCCTTGCCAGTCTTT
AGTGATCAGGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGGAATGGTTAG
AAGTGAGGGAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACTTTCTAGCATATATGTGTCCATTTCCTTAT
GCTGTAAAAGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCACCTGCTCTGCTGA
TGACCCCCCCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCCTCCAGAACT
CTTCCTCCAAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCCTGGCACTACCAAAGGACACTTATCCACGA
GAGCGCAGCATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTATACTTAGTC
AAATGTAATGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCTGGTAAGAGG
AGTGCGTGGCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGTTGGGAACAC
AGCATTTCAAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAGCCTCTTTCC
CTTCCAGTTTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTTTCTTCCTCA
TTCCCTCGTCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGCCCTGGCAGG
CAGGGGCCCTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCCTTAAGCTGC
AGGATTCTCACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAATTTCAATTTTGTCTGTACTTGAACATTATG
AAGATGGGGGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCGACAGCTTTCCAGTACCCATGGGGCTAGGTCA
TTAAGGCCACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTACTACCTCCTCTCCTGACAATACTGTATGT
CGTCGAGCTCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGATTGCCGGTCT
TGACAGCTCAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCTGTATTCCCA
TATCAAAAGACACAGGGGACACCCAGAAACGCCACATCCCCCAATCCATCAGTGCCAAACTAGCCAACGGCCCCAGCT
TCTCAGCTCGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTGGGTGGCATC
ATTCCAGGCCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCTGCCCAGCTT

TTCATTGCTGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACAAACTGCCAAGTTGGGTGGAGAAAGGAGTTTC
TTTAGCTGACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGGATACGAGCG
GAGTCCCCTGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGTCCCCCGTCT
AACCACAGCTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAGGTACCAGTT
CTTTAAGAAGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAAGTTCACACC
TGCAGATTCTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTGGGTCTGGAG
GACCTGCCCGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGGGATTTGTTG
ATTCTCTCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAAAAGGCATAGTCTACTTGTTATAAAT
CGTTAGGATACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGGGAAAAGCTGAACACTGAAGTCAGTTCTCAAC
AATTTAGAAGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTCGATGTTTTTGAATGAATACGAGCAAGCTTTT
ACAACAGTGCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGGAATCTGGAG
GTAGCCGACCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACCAGTGACCTG
GAAGGCCCAGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAACA
CCCAGGGAACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCTCGGATCAGC
TGAGCCTTCTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGGAAGCAGGTG
TGAGGGCAGTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGAGCCCAGACC
CTGGAGGAAATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCCCCATTCTGT
CCTACCTTTAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCCTGCCCACTG
ATAGCCCCTCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATATCCAGTAAC
ACATAGACTGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTGCTTTTAGTT
TTTCTGTCCCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTTGAATTTTTATATATATATCTGTATATTGCAC
AATTATAAACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGACCTGTTAAAATTATACCTGTTGCTTAATTACA
ATATTTCTGATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAAATCTGTCTG
CTGGTCACTTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCCAGGTGAAGG
AGGCTCCAGGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTCTGGGTGCAG
GAGCGCTCCCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGAGCTGGAGTC
CACTGGGGTGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAGCCTCGCAGG
CCTCCCTGCCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGAGACGGGAGG
CCTGGCCTGTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGGGACCCCAGAATCATGTGCGTCAGTCCAAGGG
GTCCCCTCCAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCATTTGCTTTT
GTAATTCTTAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTAATAATTAGA
AACACCAAAAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATTATAATCACCAGGGCCAAAAGCTTCCCTCCCT
GCTGTCATTGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGCTACCCGTGA
GCCCTTCGGAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCAGTGTTAGCA
GTGGGTCATGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCTGATGTTTGG
CCGTGAACAGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACTGACATGAAATCGACGCCCGAGTTAGCCTC
ACCCGGTGACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGACAGTGGAGT
GCAAAAGGCTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTGTCACTGCTT
CAATCAATAACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGACCAAATATCACCAGGACTGTTACTCAATGTGT
GCCGAGCCCTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCTTCCCCTTCC
TTCTTTGCCCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTTTGTGTTCCA
AACATGAGGTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTGGGTGAAAGG
AATTTTGCTAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGTAACTATGTTTAAAGTAATTGTTCCAGAGA
CAAATATTTCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGGGAGGGGGATGGTGACTGAGATGAGAGGGGAG
AGCTGAACAGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCACTCCAAGCCA
GCAAGCCGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTGGATTCTTGT
TTTGCTTCCCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTTCCTTGAGGG
GAGCCAGGAGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCAGGAGGCTGG
GACGCTCTGCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGGTGGGCATTC
TCTCTCCAAGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCGTCACATCCCACCCCATCCTCCAGGGCTCAAC
ACTGTGACATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGGC
TGTCACCCGATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCCATTCCATGGCGAAAACATCTCCTTAGAAAAG
CCATTACCCTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGCGGAGAGTGC
TGACTGTAGTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAG
TCATGCTCCCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCATTGCCAGTAACACAGGCCACCCTGTGCAGAAG
GGAGCTGGCTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCGGGACACACT
TAGCTGGAGATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTTTTTGTGTTTTGGGACAATTACTTTAGAAAAT
AAGTAGGTCGTTTTAAAAACAAAAATTATTGATTGCTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGTGTATAGCCA
AATGACTGAAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTAAGGAAATTTGTACCATTTCAGTAAACCTGTC
TGAATGTACCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATCCCTGTAACCTATTTATTATATAAAGAGTTTG
CCTTATAAATTTACATAAAAATGTCCGTTTGTGTCTTTTGTTGTAAAAATCAAGTGATTTTTTCATAAGGTTCTTTTA
CTATTGGAAAAGATGGGCAGCACGCAGTTTTATTTTATTTTTGTAAGTTTTTTAATACATGTGAAAGCAAAGAATACT

CAGCATGCCTTTCTAAGTGACGCGTTTGCACCTTTTGTTGGGAAGTACTGTATCCTGTGCTGTTAGCATTCTCGATAA
ATCTCTCTGTGAAAGTGA

SEQ ID NO: 31 >NM_001369392.2 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=5]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTGCTCCATAAAAATACAGACTCACCAGTTCCTGCTTTGATGTG
ACATGTGACTCCCCAGAATACACCTTGCTTCTGTAGACCAGCTCCAACAGGATTCCATGGTAGCTGGGATGTTAGGGC
TCAGCACTGAGCTAGCCTCTCCTTGCTAGAAGCTTACGGTCGAAGGTCTCCAGCCATCAGAAGAAGCACGTGGAGCGC
TGCGTCCGTGTTGTGGTTATTCATCCAGCATGTGTTGAGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAG
GACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCA
TCAGCCCACCACTCTGCTGAGCCCGCAGAGGCAGGCAAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTG
CCGGAAGCTTCTGCCTCCCCCAAACAGCGGCGCTCCATCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTG
CCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAATCTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAAT
CCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCT
AATGATTTTGACTTCACGGTAACTGGGAGAGGGAGCCCCTCCCGGCGAGAGCAGAAACCACCTAAGAAGCCCAAATCT
CCCAAAGCTCCAGGAACTGGCAGAGGCCGGGGACGCCCCAAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCA
GAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCA
GGGGGCAAGGCTGAGGGGGGTGGGGCCACCACATCCACCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAA
GCTGAGGCCGACCCTCAGGCCATTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAG
GCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACC
CGGGAGACGGTCAGCATCGAGGTCAAGGAAGTGGTGAAGCCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAA
GGACTGAAGACCTGTAAGAGCCCTGGGCGGAAAAGCAAGGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCA
CCCCCCAAGAAGGAGCACCACCACCATCACCACCACTCAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTG
CCCCCACCTCCACCTGAGCCCGAGAGCTCCGAGGACCCCACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTC
TGCAAAGAGGAGAAGATGCCCAGAGGAGGCTCACTGGAGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCC
GCGGTTGCCACCGCCGCCACGGCCGCAGAAAAGTACAAACACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCC
TCCATGCCAAGGCCAAACAGAGAGGAGCCTGTGGACAGCCGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACG
GAGCGGATTGCAAAGCAAACCAACAAGAATAAAGGCAGCTGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGC
TTCCCGATTAACTGAAATAAAAAATATTTTTTTTTCTTTCAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAG
TTGGAGCATTGGGGATGTTTTTCTTACCGACAAGCACAGTCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGC
ACTTTTCTAAACTAGGCTCCTTCAACAAGGCTTGCTGCAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCC
CCTCCCGCCCAAACCTTTCCCCCATGTGGTCGTTAGAGACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCG
GTGCCATCAGTGCCCCGTCTACAGCTCCCCCAGCTCCCCCCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGT
GTGAGGCAGGAGAGACAGTTGGATTCTTTAGAGAAGATGGATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGT
GGTGGCTCAAGTCTGGCCCCACACCAGCCCCAATCCAAAACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCT
GTCTGCTCCAGCTCTGGCATGGCTAGGAGGGGGGAGTCCCTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAG
AGGATGGCCCAGGGTGAGGTGGCATGGTCCATTCTCAAGGGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGC
AGTAGGACAAGGTGCAGGCAGGCTGGCCTGGGGTCAGGCCGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCAC
TCAGAGCAGTCTGTGACTTAGTGGACAGGGGAGGGGGCAAAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCC
AATTCTCCTTTAGGGACAGCTTAGAATTATTTGCACTATTGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTC
TGAGAGCAAACTGGCTTGAATTGGTGACATTTAGTCCCTCAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGAT
TTGTATATATACCTGCGCTTGTTTTAAAGTGGGCTCAGCACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTT
GCTGCAATTTTATAAGGACTTCCTGATTGGTTTCTCTTCTCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTC
ACTTCTTTCCCTTCCTCCATCCTCCTCCTTCCTAGTTCATCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACT
TGTCGGCTCCAGTCCCCAGAACTCTGCCTGCCCTTTGTCCTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCC
TGAGGAGGCCTTGGGCTCTGCTGAGTCCGACCTGGCCTGTCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCT
AGGTAGCCCCCTCTTCCCTGGTAAGAAAAAGCAAAAGGCATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTAC
TCTAGAGAAGTGGACTGGAGGAGCTGGGCCCGATTTGGTAGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGT
CATCGAGTGGCCCAACAGGGGCTCCATGCCAGCCGACCTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGC
AGGGCAGTAGCGGTACCAATGCAGAACTCCCAAGACCCGAGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGC
TCCCCCTTTTCCCTCGGAGTTCTTCTTGAATGGCAATGTTTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCA
CACATTTCACTGTCTGTCTGGTCCATAGCTGTGGTGTAGGGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGA
TCAGTTTTCATGTGGGATCCCATCTTTTTAACCTCTGTTCAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTG
GTATATCCTTTTCTGTGTTTACAGAGATGTCTCTTATATCTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGC
AAATGAGACAGCAGTCTTATGCTTCCAGAAACACCCACAGGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGT
GTGTGTTGTCTTGTGTATTTCAGTTATTGTCCCTGGCTTCCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATA
GAAACTGTCTAGCACTTCCTTGCCAGTCTTTAGTGATCAGGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAA
AAACCGTGTTTGTCTCTTCTGGAATGGTTAGAAGTGAGGGAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGA

CTTTCTAGCATATATGTGTCCATTTCCTTATGCTGTAAAAGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCT
GATCCCTGATCCCTTCCACCTGCTCTGCTGATGACCCCCCCCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGG
GGTCAGAAGAGAGGGTGAGTCCTCCAGAACTCTTCCTCCAAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACT
CCTGGCACTACCAAAGGACACTTATCCACGAGAGCGCAGCATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGG
TCAGTTGGGTTTTTATGTATTATACTTAGTCAAATGTAATGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGT
CACCTGGGCGTCATCTGGTCCTGGTAAGAGGAGTGCGTGGCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCA
GGGCCGATGGGGCAGTCGTGGTTGGGAACACAGCATTTCAAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTC
CCTCAAAGAGGCAGTTGCCCAGCCTCTTTCCCTTCCAGTTTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGG
GTTTTCTCTCTATTTCCCCCTTTCTTCCTCATTCCCTCGTCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCA
GGCAGCCTTGGCGGTTTATCGCCCTGGCAGGCAGGGGCCCTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTG
ACAGGAGGTTGGAGGGAAAGCCTTAAGCTGCAGGATTCTCACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCA
ATTTCAATTTTGTCTGTACTTGAACATTATGAAGATGGGGGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCG
ACAGCTTTCCAGTACCCATGGGGCTAGGTCATTAAGGCCACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGT
TACTACCTCCTCTCCTGACAATACTGTATGTCGTCGAGCTCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTG
GGCTTGTCATAGCCAGTGGGATTGCCGGTCTTGACAGCTCAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGC
ACAGCTCCCTTCTGTTGATGCTGTATTCCCATATCAAAAGACACAGGGGACACCCAGAAACGCCACATCCCCCAATCC
ATCAGTGCCAAACTAGCCAACGGCCCCAGCTTCTCAGCTCGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGC
GGGTGCAGACAATCTTCTGTTGGGTGGCATCATTCCAGGCCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCC
CCAAATTGTCACCTGCTTCTCTGCCCAGCTTTTCATTGCTGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACA
AACTGCCAAGTTGGGTGGAGAAAGGAGTTTCTTTAGCTGACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTC
AAGCCCAGGAGGGAGCAGAGGGATACGAGCGGAGTCCCCTGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAG
CCTGACAGCCAGAACTCTGTGTCCCCCGTCTAACCACAGCTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTG
ACTGGGCCAGGGGAGGTTACAGGTACCAGTTCTTTAAGAAGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGC
CCCAAATGGATTCCTGTTTCAAGTTCACACCTGCAGATTCTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTT
CTAGAGTTCCTACCATGGAGTGGGTCTGGAGGACCTGCCCGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCT
ACTCTTCTCTCTGCTCTGACGGGATTTGTTGATTCTCTCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCT
TTAGAAAAGGCATAGTCTACTTGTTATAAATCGTTAGGATACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGG
GAAAAGCTGAACACTGAAGTCAGTTCTCAACAATTTAGAAGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTC
GATGTTTTTGAATGAATACGAGCAAGCTTTTACAACAGTGCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCT
GGTGAGCATTACAGGCAAGGGGAATCTGGAGGTAGCCGACCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGG
TATGGAAGGTGGAGCGTTCACCAGTGACCTGGAAGGCCCAGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAG
CCTGCCCCAGCGCTGACGTGTCAGGAAAACACCCAGGGAACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTG
CCCACTCCTGCTCTGCTCGCCTCGGATCAGCTGAGCCTTCTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCC
TGCCCTGTCAGGAGGCAGAAGGAAGCAGGTGTGAGGGCAGTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGG
CTCCGACTTGTGCACAGGCAGAGCCCAGACCCTGGAGGAAATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGG
AAATCTCTTTGCCCCCAAACCCCCATTCTGTCCTACCTTTAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGA
AAAGGCCAAGAGGTTTGGCTCCTGCCCACTGATAGCCCCTCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTT
CTTCCTGGTGACCCTGATTATATCCAGTAACACATAGACTGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCT
TTTGTTTTGCTTTTTAGTTTTGCTTTTAGTTTTTCTGTCCCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTT
GAATTTTTATATATATATCTGTATATTGCACAATTATAAACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGAC
CTGTTAAAATTATACCTGTTGCTTAATTACAATATTTCTGATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAA
AAAAAAGAAAAAAAAACGACAAATCTGTCTGCTGGTCACTTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTT
CTTTCAAGGGCTTTCCTGTGCCAGGTGAAGGAGGCTCCAGGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCT
TGTGAAAGAGGTCCCAAGGTTCTGGGTGCAGGAGCGCTCCCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCC
TGGTCGCCTTCCACCTCTGGGAGCTGGAGTCCACTGGGGTGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCA
GGGTGAGCCCATGTGGAGTCAGCCTCGCAGGCCTCCCTGCCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGT
CGAGCCTGGGGGCTGGAGCGGAGACGGGAGGCCTGGCCTGTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGG
GACCCCAGAATCATGTGCGTCAGTCCAAGGGGTCCCCTCCAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTC
CCCCATCCTACGAGTAATTGCATTTGCTTTTGTAATTCTTAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGT
TCTTTTTGATCATCTTTTTTTAATAATTAGAAACACCAAAAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATT
ATAATCACCAGGGCCAAAAGCTTCCCTCCCTGCTGTCATTGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCA
TAGGCCCTTTCAGTGGCCGGGCTACCCGTGAGCCCTTCGGAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGG
GGCCAGCTCCGGCGTGTGTTCAGTGTTAGCAGTGGGTCATGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGG
AGGCGAGGAGGCCGTTGCCGCTGATGTTTGGCCGTGAACAGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGA
CTGACATGAAATCGACGCCCGAGTTAGCCTCACCCGGTGACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGT
TCAGTGTTTCTGGGGAGCTGGACAGTGGAGTGCAAAAGGCTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCAC
GGCCTCCTTTCCGTTTGATTTGTCACTGCTTCAATCAATAACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGAC
CAAATATCACCAGGACTGTTACTCAATGTGTGCCGAGCCCTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAA
CGTGTGCTGTGTTTGCTCCCCTTCCCCTTCCTTCTTTGCCCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTG
GTTTGGTTTTTATTTCTCCTTTTGTGTTCCAAACATGAGGTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCT
TATATTTGTGTAATTTTTGGTGGGTGAAAGGAATTTTGCTAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATT
GTAACTATGTTTAAAGTAATTGTTCCAGAGACAAATATTTCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGG
GAGGGGGATGGTGACTGAGATGAGAGGGGAGAGCTGAACAGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAG

CAGTGGAGCCCAGGAGTCCCACTCCAAGCCAGCAAGCCGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAA
CCAGGTTTTGTTCCGCCCAGTGGATTCTTGTTTTGCTTCCCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTA
AGGAAAGGCAAGATTGATGTTTCCTTGAGGGGAGCCAGGAGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAA
TGGGGCTGGGCCCACCCAAGCAGGAGGCTGGGACGCTCTGCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGC
TCTTCCTGGACAGGCCAGGTGGTGGGCATTCTCTCTCCAAGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCG
TCACATCCCACCCCATCCTCCAGGGCTCAACACTGTGACATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAA
TGACCATGGAGGGGGCTTGCACTCTCTTGGCTGTCACCCGATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCC
ATTCCATGGCGAAAACATCTCCTTAGAAAAGCCATTACCCTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACA
GCCCCTCCCTCCTCTGAGAGGCGGAGAGTGCTGACTGTAGTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTA
GGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCATGCTCCCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCAT
TGCCAGTAACACAGGCCACCCTGTGCAGAAGGGAGCTGGCTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCA
CTTGGGGCACAGGGAGAGGCCGGGACACACTTAGCTGGAGATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTT
TTTGTGTTTTGGGACAATTACTTTAGAAAATAAGTAGGTCGTTTTAAAAACAAAAATTATTGATTGCTTTTTTGTAGT
GTTCAGAAAAAAGGTTCTTTGTGTATAGCCAAATGACTGAAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTA
AGGAAATTTGTACCATTTCAGTAAACCTGTCTGAATGTACCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATC
CCTGTAACCTATTTATTATATAAAGAGTTTGCCTTATAAATTTACATAAAAATGTCCGTTTGTGTCTTTTGTTGTAAA
AATCAAGTGATTTTTTCATAAGGTTCTTTTACTATTGGAAAAGATGGGCAGCACGCAGTTTTATTTTATTTTTGTAAG
TTTTTTAATACATGTGAAAGCAAAGAATACTCAGCATGCCTTTCTAAGTGACGCGTTTGCACCTTTTGTTGGGAAGTA
CTGTATCCTGTGCTGTTAGCATTCTCGATAAATCTCTCTGTGAAAGTGA

SEQ ID NO: 32>NM_001369391.2 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=4]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTGCTCCATAAAAATACAGACTCACCAGTTCCTGCTTTGATGTG
ACATGTGACTCCCCAGAATACACCTTGCTTCTGTAGACCAGCTCCAACAGGATTCCATGGTAGCTGGGATGTTAGGGC
TCAGGACCCAGGAGGGAGAAGCAGCTGCCAGGGGAAGTCTCTTCGTAGGCGGAGGTCAGGAGTCCAAGAGGAGTGAGC
AGAGTCACAGAAGCCTCTTAAAGCCTCTTCTTCCCCCATCCCATCAACACATCTGCTGCCCAAGAACTCCTTTGTGAT
TGAATTTAACCAGAGTTGATTTCTGTTGCTTGAAACTCAGGACCTTGATGCAGTGTGCAGGGTCTGGGCAAGGAAATT
CATAAAAGAAGAGATACAGATGGCCAATGAAGATAGGAAAAGAGTTCTGCCTGCTGGTAATCAAAGAGATGCAAACGA
GAACAAAAATGATGCCTTTTCACCTACCAAATTTGTCAAGATTAAAAAGAAAGCAAAGAGCCAGCGTCAGCTGATGTT
CATACCTGCACCTGCTCGGTAGCTTGCTAATGTTCTGCCTGCTCCACACGCCAGGCCAGCCTCCACGGCGCAGCCAGG
GAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAG
AAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATCAGCCCACCACTCTGCTGAGCCCGCAGAGGCAGGCAAA
GCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCCGGAAGCTTCTGCCTCCCCCAAACAGCGGCGCTCCATC
ATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTGCCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAATCT
GGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAATCCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATT
GCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCTAATGATTTTGACTTCACGGTAACTGGGAGAGGGAGCCCC
TCCCGGCGAGAGCAGAAACCACCTAAGAAGCCCAAATCTCCCAAAGCTCCAGGAACTGGCAGAGGCCGGGGACGCCCC
AAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCT
GGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACCACATCCACC
CAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCATTCCCAAGAAACGGGGC
CGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCT
GTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGGTCAAGGAAGTGGTGAAG
CCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGAAAAGCAAG
GAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACCACCATCACCACCACTCA
GAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCGAGAGCTCCGAGGACCCC
ACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCAGAGGAGGCTCACTGGAG
AGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGGCCGCAGAAAAGTACAAA
CACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTGTGGACAGC
CGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACGGAGCGGATTGCAAAGCAAACCAACAAGAATAAAGGCAGC
TGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTTCCCGATTAACTGAAATAAAAAATATTTTTTTTTCTTT
CAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTTCTTACCGACAAGCACAG
TCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTTCAACAAGGCTTGCTGCA
GATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCCCATGTGGTCGTTAGAGA
CAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTACAGCTCCCCCAGCTCCCC
CCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGGATTCTTTAGAGAAGATG
GATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCCAATCCAAA
ACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGGCTAGGAGGGGGGAGTCC
CTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGGCATGGTCCATTCTCAAG
GGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGGCTGGCCTGGGGTCAGGC
CGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCACTCAGAGCAGTCTGTGACTTAGTGGACAGGGGAGGGGGCA

AAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTTAGAATTATTTGCACTAT
TGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACATTTAGTCCCT
CAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGATTTGTATATATACCTGCGCTTGTTTTAAAGTGGGCTCAGC
ACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTCCTGATTGGTTTCTCTTC
TCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCCTCCTCCTTCCTAGTTCA
TCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACTTGTCGGCTCCAGTCCCCAGAACTCTGCCTGCCCTTTGTC
CTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCTGAGTCCGACCTGGCCTG
TCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAAGCAAAAGGC
ATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGAGCTGGGCCCGATTTGGT
AGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGCTCCATGCCAGCCGACCT
TGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGCAGAACTCCCAAGACCCG
AGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAATGGCAATGT
TTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCACACATTTCACTGTCTGTCTGGTCCATAGCTGTGGTGTAG
GGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGATCAGTTTTCATGTGGGATCCCATCTTTTTAACCTCTGTT
CAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTGGTATATCCTTTTCTGTGTTTACAGAGATGTCTCTTATAT
CTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGCAAATGAGACAGCAGTCTTATGCTTCCAGAAACACCCACA
GGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCAGTTATTGTCCCTGGCTT
CCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATAGAAACTGTCTAGCACTTCCTTGCCAGTCTTTAGTGATCA
GGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGGAATGGTTAGAAGTGAGG
GAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACTTTCTAGCATATATGTGTCCATTTCCTTATGCTGTAAA
AGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCACCTGCTCTGCTGATGACCCCC
CCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCCTCCAGAACTCTTCCTCC
AAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCCTGGCACTACCAAAGGACACTTATCCACGAGAGCGCAG
CATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTATACTTAGTCAAATGTAA
TGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCTGGTAAGAGGAGTGCGTG
GCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGTTGGGAACACAGCATTTC
AAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAGCCTCTTTCCCTTCCAGT
TTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTTTCTTCCTCATTCCCTCG
TCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGCCCTGGCAGGCAGGGGCC
CTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCCTTAAGCTGCAGGATTCT
CACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAATTTCAATTTTGTCTGTACTTGAACATTATGAAGATGGG
GGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCGACAGCTTTCCAGTACCCATGGGGCTAGGTCATTAAGGCC
ACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTACTACCTCCTCTCCTGACAATACTGTATGTCGTCGAGC
TCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGATTGCCGGTCTTGACAGCT
CAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCTGTATTCCCATATCAAAA
GACACAGGGGACACCCAGAAACGCCACATCCCCCAATCCATCAGTGCCAAACTAGCCAACGGCCCCAGCTTCTCAGCT
CGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTGGGTGGCATCATTCCAGG
CCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCTGCCCAGCTTTTCATTGC
TGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACAAACTGCCAAGTTGGGTGGAGAAAGGAGTTTCTTTAGCTG
ACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGGATACGAGCGGAGTCCCC
TGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGTCCCCCGTCTAACCACAG
CTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAGGTACCAGTTCTTTAAGA
AGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAAGTTCACACCTGCAGATT
CTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTGGGTCTGGAGGACCTGCC
CGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGGGATTTGTTGATTCTCTC
CATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAAAAGGCATAGTCTACTTGTTATAAATCGTTAGGA
TACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGGGAAAAGCTGAACACTGAAGTCAGTTCTCAACAATTTAGA
AGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTCGATGTTTTTGAATGAATACGAGCAAGCTTTTACAACAGT
GCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGGAATCTGGAGGTAGCCGA
CCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACCAGTGACCTGGAAGGCCC
AGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAACACCCAGGGA
ACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCTCGGATCAGCTGAGCCTT
CTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGGAAGCAGGTGTGAGGGCA
GTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGAGCCCAGACCCTGGAGGA
AATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCCCCATTCTGTCCTACCTT
TAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCCTGCCCACTGATAGCCCC
TCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATATCCAGTAACACATAGAC
TGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTGCTTTTAGTTTTTCTGTC
CCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTTGAATTTTTATATATATATCTGTATATTGCACAATTATAA
ACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGACCTGTTAAAATTATACCTGTTGCTTAATTACAATATTTCT
GATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAAATCTGTCTGCTGGTCAC
TTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCCAGGTGAAGGAGGCTCCA
GGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTCTGGGTGCAGGAGCGCTC

CCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGAGCTGGAGTCCACTGGGG
TGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAGCCTCGCAGGCCTCCCTG
CCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGAGACGGGAGGCCTGGCCT
GTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGGGACCCCAGAATCATGTGCGTCAGTCCAAGGGGTCCCCTC
CAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCATTTGCTTTTGTAATTCT
TAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTTAATAATTAGAAACACCAA
AAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATTATAATCACCAGGGCCAAAAGCTTCCCTCCCTGCTGTCAT
TGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGCTACCCGTGAGCCCTTCG
GAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCAGTGTTAGCAGTGGGTCA
TGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCTGATGTTTGGCCGTGAAC
AGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACTGACATGAAATCGACGCCCGAGTTAGCCTCACCCGGTG
ACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGACAGTGGAGTGCAAAAGG
CTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTGTCACTGCTTCAATCAAT
AACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGACCAAATATCACCAGGACTGTTACTCAATGTGTGCCGAGCC
CTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCTTCCCCTTCCTTCTTTGC
CCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTTTGTGTTCCAAACATGAG
GTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTGGGTGAAAGGAATTTTGC
TAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGTAACTATGTTTAAAGTAATTGTTCCAGAGACAAATATT
TCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGGGAGGGGGATGGTGACTGAGATGAGAGGGGAGAGCTGAAC
AGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCACTCCAAGCCAGCAAGCCG
AATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTGGATTCTTGTTTTGCTTC
CCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTTCCTTGAGGGGAGCCAGG
AGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCAGGAGGCTGGGACGCTCT
GCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGGTGGGCATTCTCTCTCCA
AGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCGTCACATCCCACCCCATCCTCCAGGGCTCAACACTGTGAC
ATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGGCTGTCACCC
GATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCCATTCCATGGCGAAAACATCTCCTTAGAAAAGCCATTACC
CTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGCGGAGAGTGCTGACTGTA
GTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCATGCTC
CCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCATTGCCAGTAACACAGGCCACCCTGTGCAGAAGGGAGCTGG
CTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCGGGACACACTTAGCTGGA
GATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTTTTTGTGTTTTGGGACAATTACTTTAGAAAATAAGTAGGT
CGTTTTAAAAACAAAAATTATTGATTGCTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGTGTATAGCCAAATGACTG
AAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTAAGGAAATTTGTACCATTTCAGTAAACCTGTCTGAATGTA
CCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATCCCTGTAACCTATTTATTATATAAAGAGTTTGCCTTATAA
ATTTACATAAAAATGTCCGTTTGTGTCTTTTGTTGTAAAAATCAAGTGATTTTTTCATAAGGTTCTTTTACTATTGGA
AAAGATGGGCAGCACGCAGTTTTATTTTATTTTTGTAAGTTTTTTAATACATGTGAAAGCAAAGAATACTCAGCATGC
CTTTCTAAGTGACGCGTTTGCACCTTTTGTTGGGAAGTACTGTATCCTGTGCTGTTAGCATTCTCGATAAATCTCTCT
GTGAAAGTGA

SEQ ID NO: 33>NM_001316337.2 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=3]

GCGCGCGCTCCCTCCTCTCGGAGAGAGGGCTGTGGTAAAAGCCGTCCGGAAAATGGCCGCCGCCGCCGCCGCCGCGCC
GAGCGGAGGAGGAGGAGGAGGCGAGGAGGAGAGACTGCTCCATAAAAATACAGACTCACCAGTTCCTGCTTTGATGTG
ACATGTGACTCCCCAGAATACACCTTGCTTCTGTAGACCAGCTCCAACAGGATTCCATGGTAGCTGGGATGTTAGGGC
TCAGCTGCAAGATGGGATTCAGATCTGTTCTCAAGCCTGTCGTTCCAGGACCCAGGAGGGAGAAGCAGCTGCCAGGGG
AAGTCTCTTCGTAGGCGGAGGTCAGGAGTCCAAGAGGAGTGAGCAGAGTCACAGAAGCCTCTTAAAGCCTCTTCTTCC
CCCATCCCATCAACACGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGACAAACCCCTCAAGTTTAAAA
AGGTGAAGAAAGATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATCAGCCCACCACTCTGCTGAGC
CCGCAGAGGCAGGCAAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCCGGAAGCTTCTGCCTCCCCCA
AACAGCGGCGCTCCATCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTGCCTGAAGGCTGGACACGGAAGC
TTAAGCAAAGGAAATCTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAATCCCCAGGGAAAAGCCTTTCGCT
CTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCTAATGATTTTGACTTCACGGTAA
CTGGGAGAGGGAGCCCCTCCCGGCGAGAGCAGAAACCACCTAAGAAGCCCAAATCTCCCAAAGCTCCAGGAACTGGCA
GAGGCCGGGGACGCCCCAAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGG
TCCTGGAGAAAAGTCCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTG
GGGCCACCACATCCACCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCA
TTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGG
AGTCTTCTATCCGATCTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGG
TCAAGGAAGTGGTGAAGCCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCC
CTGGGCGGAAAAGCAAGGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACC

ACCATCACCACCACTCAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCG
AGAGCTCCGAGGACCCCACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCA
GAGGAGGCTCACTGGAGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGG
CCGCAGAAAAGTACAAACACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAG
AGGAGCCTGTGGACAGCCGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACGGAGCGGATTGCAAAGCAAACCA
ACAAGAATAAAGGCAGCTGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTTCCCGATTAACTGAAATAAAA
AATATTTTTTTTTCTTTCAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTT
CTTACCGACAAGCACAGTCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTT
CAACAAGGCTTGCTGCAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCC
CATGTGGTCGTTAGAGACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTAC
AGCTCCCCCAGCTCCCCCCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGG
ATTCTTTAGAGAAGATGGATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCAC
ACCAGCCCCAATCCAAAACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGG
CTAGGAGGGGGGAGTCCCTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGG
CATGGTCCATTCTCAAGGGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGG
CTGGCCTGGGGTCAGGCCGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCACTCAGAGCAGTCTGTGACTTAGT
GGACAGGGGAGGGGGCAAAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTT
AGAATTATTTGCACTATTGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATT
GGTGACATTTAGTCCCTCAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGATTTGTATATATACCTGCGCTTGT
TTTAAAGTGGGCTCAGCACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTC
CTGATTGGTTTCTCTTCTCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCC
TCCTCCTTCCTAGTTCATCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACTTGTCGGCTCCAGTCCCCAGAAC
TCTGCCTGCCCTTTGTCCTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCT
GAGTCCGACCTGGCCTGTCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGT
AAGAAAAAGCAAAAGGCATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGA
GCTGGGCCCGATTTGGTAGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGC
TCCATGCCAGCCGACCTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGC
AGAACTCCCAAGACCCGAGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTC
TTCTTGAATGGCAATGTTTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCACACATTTCACTGTCTGTCTGGT
CCATAGCTGTGGTGTAGGGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGATCAGTTTTCATGTGGGATCCCA
TCTTTTTAACCTCTGTTCAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTGGTATATCCTTTTCTGTGTTTAC
AGAGATGTCTCTTATATCTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGCAAATGAGACAGCAGTCTTATGC
TTCCAGAAACACCCACAGGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCA
GTTATTGTCCCTGGCTTCCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATAGAAACTGTCTAGCACTTCCTTG
CCAGTCTTTAGTGATCAGGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGG
AATGGTTAGAAGTGAGGGAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACTTTCTAGCATATATGTGTCCA
TTTCCTTATGCTGTAAAAGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCACCTG
CTCTGCTGATGACCCCCCCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCC
TCCAGAACTCTTCCTCCAAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCCTGGCACTACCAAAGGACACT
TATCCACGAGAGCGCAGCATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTA
TACTTAGTCAAATGTAATGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCT
GGTAAGAGGAGTGCGTGGCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGT
TGGGAACACAGCATTTCAAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAG
CCTCTTTCCCTTCCAGTTTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTT
TCTTCCTCATTCCCTCGTCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGC
CCTGGCAGGCAGGGGCCCTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCC
TTAAGCTGCAGGATTCTCACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAATTTCAATTTTGTCTGTACTTG
AACATTATGAAGATGGGGGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCGACAGCTTTCCAGTACCCATGGG
GCTAGGTCATTAAGGCCACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTACTACCTCCTCTCCTGACAAT
ACTGTATGTCGTCGAGCTCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGAT
TGCCGGTCTTGACAGCTCAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCT
GTATTCCCATATCAAAAGACACAGGGGACACCCAGAAACGCCACATCCCCCAATCCATCAGTGCCAAACTAGCCAACG
GCCCCAGCTTCTCAGCTCGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTG
GGTGGCATCATTCCAGGCCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCT
GCCCAGCTTTTCATTGCTGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACAAACTGCCAAGTTGGGTGGAGAA
AGGAGTTTCTTTAGCTGACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGG
ATACGAGCGGAGTCCCCTGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGT
CCCCCGTCTAACCACAGCTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAG
GTACCAGTTCTTTAAGAAGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAA
GTTCACACCTGCAGATTCTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTG
GGTCTGGAGGACCTGCCCGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGG
GATTTGTTGATTCTCTCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAAAAGGCATAGTCTACTT
GTTATAAATCGTTAGGATACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGGGAAAAGCTGAACACTGAAGTCA

GTTCTCAACAATTTAGAAGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTCGATGTTTTTGAATGAATACGAG
CAAGCTTTTACAACAGTGCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGG
AATCTGGAGGTAGCCGACCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACC
AGTGACCTGGAAGGCCCAGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTC
AGGAAAACACCCAGGGAACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCT
CGGATCAGCTGAGCCTTCTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGG
AAGCAGGTGTGAGGGCAGTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGA
GCCCAGACCCTGGAGGAAATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCC
CCATTCTGTCCTACCTTTAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCC
TGCCCACTGATAGCCCCTCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATA
TCCAGTAACACATAGACTGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTG
CTTTTAGTTTTTCTGTCCCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTTGAATTTTTATATATATATCTGT
ATATTGCACAATTATAAACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGACCTGTTAAAATTATACCTGTTGC
TTAATTACAATATTTCTGATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAA
ATCTGTCTGCTGGTCACTTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCC
AGGTGAAGGAGGCTCCAGGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTC
TGGGTGCAGGAGCGCTCCCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGA
GCTGGAGTCCACTGGGGTGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAG
CCTCGCAGGCCTCCCTGCCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGA
GACGGGAGGCCTGGCCTGTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGGGACCCCAGAATCATGTGCGTCA
GTCCAAGGGGTCCCCTCCAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCA
TTTGCTTTTGTAATTCTTAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTA
ATAATTAGAAACACCAAAAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATTATAATCACCAGGGCCAAAAGCT
TCCCTCCCTGCTGTCATTGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGC
TACCCGTGAGCCCTTCGGAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCA
GTGTTAGCAGTGGGTCATGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCT
GATGTTTGGCCGTGAACAGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACTGACATGAAATCGACGCCCGA
GTTAGCCTCACCCGGTGACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGA
CAGTGGAGTGCAAAAGGCTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTG
TCACTGCTTCAATCAATAACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGACCAAATATCACCAGGACTGTTAC
TCAATGTGTGCCGAGCCCTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCT
TCCCCTTCCTTCTTTGCCCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTT
TGTGTTCCAAACATGAGGTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTG
GGTGAAAGGAATTTTGCTAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGTAACTATGTTTAAAGTAATTG
TTCCAGAGACAAATATTTCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGGGAGGGGGATGGTGACTGAGATG
AGAGGGGAGAGCTGAACAGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCAC
TCCAAGCCAGCAAGCCGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTG
GATTCTTGTTTTGCTTCCCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTT
CCTTGAGGGGAGCCAGGAGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCA
GGAGGCTGGGACGCTCTGCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGG
TGGGCATTCTCTCTCCAAGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCGTCACATCCCACCCCATCCTCCA
GGGCTCAACACTGTGACATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCAC
TCTCTTGGCTGTCACCCGATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCCATTCCATGGCGAAAACATCTCC
TTAGAAAAGCCATTACCCTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGC
GGAGAGTGCTGACTGTAGTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCT
GAGTTGAAGTCATGCTCCCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCATTGCCAGTAACACAGGCCACCCT
GTGCAGAAGGGAGCTGGCTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCG
GGACACACTTAGCTGGAGATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTTTTTGTGTTTTGGGACAATTACT
TTAGAAAATAAGTAGGTCGTTTTAAAAACAAAAATTATTGATTGCTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGT
GTATAGCCAAATGACTGAAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTAAGGAAATTTGTACCATTTCAGT
AAACCTGTCTGAATGTACCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATCCCTGTAACCTATTTATTATATA
AAGAGTTTGCCTTATAAATTTACATAAAAATGTCCGTTTGTGTCTTTTGTTGTAAAAATCAAGTGATTTTTTCATAAG
GTTCTTTTACTATTGGAAAAGATGGGCAGCACGCAGTTTTATTTTATTTTTGTAAGTTTTTTAATACATGTGAAAGCA
AAGAATACTCAGCATGCCTTTCTAAGTGACGCGTTTGCACCTTTTGTTGGGAAGTACTGTATCCTGTGCTGTTAGCAT
TCTCGATAAATCTCTCTGTGAAAGTGA

SEQ ID NO: 34 >NM_001369394.2 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=7]

AGTAGGTGGGGATTATTATCCACAAAAGGGACTAGACATTGTGTTCTGGGTCCCACAACTCATCATAAAGAGGTGGTT
ATAGTTCCCATCAGGAGCCGTGGGTAGGGGACTGTGCGTCCAGCAGCACCCGAGGCTCTTCGGCGCCAGAGGCTCTAA
GGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATA

AGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATCAGCCCACCACTCTGCTGAGCCCGCAGAGGCAGGCA
AAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCCGGAAGCTTCTGCCTCCCCCAAACAGCGGCGCTCCA
TCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTGCCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAAT
CTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAATCCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTGA
TTGCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCTAATGATTTTGACTTCACGGTAACTGGGAGAGGGAGCC
CCTCCCGGCGAGAGCAGAAACCACCTAAGAAGCCCAAATCTCCCAAAGCTCCAGGAACTGGCAGAGGCCGGGGACGCC
CCAAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTC
CTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACCACATCCA
CCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCATTCCCAAGAAACGGG
GCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGAT
CTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGGTCAAGGAAGTGGTGA
AGCCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGAAAAGCA
AGGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACCACCATCACCACCACT
CAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCGAGAGCTCCGAGGACC
CCACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCAGAGGAGGCTCACTGG
AGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGGCCGCAGAAAAGTACA
AACACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTGTGGACA
GCCGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACGGAGCGGATTGCAAAGCAAACCAACAAGAATAAAGGCA
GCTGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTTCCCGATTAACTGAAATAAAAAATATTTTTTTTTCT
TTCAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTTCTTACCGACAAGCAC
AGTCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTTCAACAAGGCTTGCTG
CAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCCCATGTGGTCGTTAGA
GACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTACAGCTCCCCCAGCTCC
CCCCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGGATTCTTTAGAGAAGA
TGGATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCCAATCCA
AAACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGGCTAGGAGGGGGGAGT
CCCTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGGCATGGTCCATTCTCA
AGGGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGGCTGGCCTGGGGTCAG
GCCGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCACTCAGAGCAGTCTGTGACTTAGTGGACAGGGGAGGGGG
CAAAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTTAGAATTATTTGCACT
ATTGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACATTTAGTCC
CTCAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGATTTGTATATATACCTGCGCTTGTTTTAAAGTGGGCTCA
GCACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTCCTGATTGGTTTCTCT
TCTCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCCTCCTCCTTCCTAGTT
CATCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACTTGTCGGCTCCAGTCCCCAGAACTCTGCCTGCCCTTTG
TCCTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCTGAGTCCGACCTGGCC
TGTCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAAGCAAAAG
GCATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGAGCTGGGCCCGATTTG
GTAGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGCTCCATGCCAGCCGAC
CTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGCAGAACTCCCAAGACC
CGAGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAATGGCAAT
GTTTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCACACATTTCACTGTCTGTCTGGTCCATAGCTGTGGTGT
AGGGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGATCAGTTTTCATGTGGGATCCCATCTTTTTAACCTCTG
TTCAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTGGTATATCCTTTTCTGTGTTTACAGAGATGTCTCTTAT
ATCTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGCAAATGAGACAGCAGTCTTATGCTTCCAGAAACACCCA
CAGGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCAGTTATTGTCCCTGGC
TTCCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATAGAAACTGTCTAGCACTTCCTTGCCAGTCTTTAGTGAT
CAGGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGGAATGGTTAGAAGTGA
GGGAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACTTTCTAGCATATATGTGTCCATTTCCTTATGCTGTA
AAAGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCACCTGCTCTGCTGATGACCC
CCCCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCCTCCAGAACTCTTCCT
CCAAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCCTGGCACTACCAAAGGACACTTATCCACGAGAGCGC
AGCATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTATACTTAGTCAAATGT
AATGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCTGGTAAGAGGAGTGCG
TGGCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGTTGGGAACACAGCATT
TCAAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAGCCTCTTTCCCTTCCA
GTTTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTTTCTTCCTCATTCCCT
CGTCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGCCCTGGCAGGCAGGGG
CCCTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCCTTAAGCTGCAGGATT
CTCACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAATTTCAATTTTGTCTGTACTTGAACATTATGAAGATG
GGGGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCGACAGCTTTCCAGTACCCATGGGGCTAGGTCATTAAGG
CCACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTACTACCTCCTCTCCTGACAATACTGTATGTCGTCGA
GCTCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGATTGCCGGTCTTGACAG

CTCAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCTGTATTCCCATATCAA
AAGACACAGGGGACACCCAGAAACGCCACATCCCCCAATCCATCAGTGCCAAACTAGCCAACGGCCCCAGCTTCTCAG
CTCGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTGGGTGGCATCATTCCA
GGCCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCTGCCCAGCTTTTCATT
GCTGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACAAACTGCCAAGTTGGGTGGAGAAAGGAGTTTCTTTAGC
TGACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGGATACGAGCGGAGTCC
CCTGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGTCCCCCGTCTAACCAC
AGCTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAGGTACCAGTTCTTTAA
GAAGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAAGTTCACACCTGCAGA
TTCTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTGGGTCTGGAGGACCTG
CCCGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGGGATTTGTTGATTCTC
TCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAAAAGGCATAGTCTACTTGTTATAAATCGTTAG
GATACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGGGAAAAGCTGAACACTGAAGTCAGTTCTCAACAATTTA
GAAGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTCGATGTTTTTGAATGAATACGAGCAAGCTTTTACAACA
GTGCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGGAATCTGGAGGTAGCC
GACCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACCAGTGACCTGGAAGGC
CCAGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAACACCCAGG
GAACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCTCGGATCAGCTGAGCC
TTCTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGGAAGCAGGTGTGAGGG
CAGTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGAGCCCAGACCCTGGAG
GAAATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCCCCATTCTGTCCTACC
TTTAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCCTGCCCACTGATAGCC
CCTCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATATCCAGTAACACATAG
ACTGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTGCTTTTAGTTTTTCTG
TCCCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTTGAATTTTTATATATATATCTGTATATTGCACAATTAT
AAACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGACCTGTTAAAATTATACCTGTTGCTTAATTACAATATTT
CTGATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAAATCTGTCTGCTGGTC
ACTTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCCAGGTGAAGGAGGCTC
CAGGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTCTGGGTGCAGGAGCGC
TCCCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGAGCTGGAGTCCACTGG
GGTGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAGCCTCGCAGGCCTCCC
TGCCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGAGACGGGAGGCCTGGC
CTGTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGGGACCCCAGAATCATGTGCGTCAGTCCAAGGGGTCCCC
TCCAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCATTTGCTTTTGTAATT
CTTAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTAATAATTAGAAACACC
AAAAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATTATAATCACCAGGGCCAAAAGCTTCCCTCCCTGCTGTC
ATTGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGCTACCCGTGAGCCCTT
CGGAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCAGTGTTAGCAGTGGGT
CATGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCTGATGTTTGGCCGTGA
ACAGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACTGACATGAAATCGACGCCCGAGTTAGCCTCACCCGG
TGACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGACAGTGGAGTGCAAAA
GGCTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTGTCACTGCTTCAATCA
ATAACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGACCAAATATCACCAGGACTGTTACTCAATGTGTGCCGAG
CCCTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCTTCCCCTTCCTTCTTT
GCCCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTTTGTGTTCCAAACATG
AGGTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTGGGTGAAAGGAATTTT
GCTAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGTAACTATGTTTAAAGTAATTGTTCCAGAGACAAATA
TTTCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGGGAGGGGGATGGTGACTGAGATGAGAGGGGAGAGCTGA
ACAGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCACTCCAAGCCAGCAAGC
CGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTGGATTCTTGTTTTGCT
TCCCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTTCCTTGAGGGGAGCCA
GGAGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCAGGAGGCTGGGACGCT
CTGCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGGTGGGCATTCTCTCTC
CAAGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCGTCACATCCCACCCCATCCTCCAGGGCTCAACACTGTG
ACATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGGCTGTCAC
CCGATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCCATTCCATGGCGAAAACATCTCCTTAGAAAAGCCATTA
CCCTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGCGGAGAGTGCTGACTG
TAGTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCATGC
TCCCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCATTGCCAGTAACACAGGCCACCCTGTGCAGAAGGGAGCT
GGCTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCGGGACACACTTAGCTG
GAGATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTTTTTGTGTTTTGGGACAATTACTTTAGAAAATAAGTAG
GTCGTTTTAAAAACAAAAATTATTGATTGCTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGTGTATAGCCAAATGAC
TGAAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTAAGGAAATTTGTACCATTTCAGTAAACCTGTCTGAATG

TACCTGTATACGTTTCAAAAACACCCCCCCCCCCACTGAATCCCTGTAACCTATTTATTATATAAAGAGTTTGCCTTAT
AAATTTACATAAAAATGTCCGTTTGTGTCTTTTGTTGTAAAAATCAAGTGATTTTTTCATAAGGTTCTTTTACTATTG
GAAAAGATGGGCAGCACGCAGTTTTATTTTATTTTTGTAAGTTTTTTAATACATGTGAAAGCAAAGAATACTCAGCAT
GCCTTTCTAAGTGACGCGTTTGCACCTTTTGTTGGGAAGTACTGTATCCTGTGCTGTTAGCATTCTCGATAAATCTCT
CTGTGAAAGTGA

SEQ ID NO: 35 >XM_024452383.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=X1]

TAAAAATACAGACTCACCAGTTCCTGCTTTGATGTGACATGTGACTCCCCAGAATACACCTTGCTTCTGTAGACCAGC
TCCAACAGGATTCCATGGTAGCTGGGATGTTAGGGCTCAGGACCCAGGAGGGAGAAGCAGCTGCCAGGGGAAGTCTCT
TCGTAGGCGGAGGTCAGGAGTCCAAGAGGAGTGAGCAGAGTCACAGAAGCCTCTTAAAGCCTCTTCTTCCCCCATCCC
ATCAACACATCTGCTGCCCAAGAACTCCTTTGTGATTGAATTTAACCAGAGTTGATTTCTGTTGCTTGAAACTCAGGA
CCTTGATGCAGTGTGCAGGGTCTGGGCAAGTCTAGACTTGAATGCAAACAGTTCTTATTGCTAACTTACCACTAGTCA
CCACTAATGAACAAAGACTATGAACAGGAAATTCATAAAAGAAGAGATACAGATGGCCAATGAAGATAGGAAAAGAGT
TCTGCCTGCTGGTAATCAAAGAGATGCAAACGAGAACAAAAATGATGCCTTTTCACCTACCAAATTTGTCAAGATTAA
AAAGAAAGCAAAGAGCCAGCGTCAGCTGATGTTCATACCTGCACCTGCTCGGTAGCTTGCTAATGTTCTGCCTGCTCC
ACACGCCAGGCCAGCCTCCACGGCGCAGCCAGGGAAGAAAAGTCAGAAGACCAGGACCTCCAGGGCCTCAAGGACAAA
CCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAGAAAGAAGAGAAAGAGGGCAAGCATGAGCCCGTGCAGCCATCAGCC
CACCACTCTGCTGAGCCCGCAGAGGCAGGCAAAGCAGAGACATCAGAAGGGTCAGGCTCCGCCCCGGCTGTGCCGGAA
GCTTCTGCCTCCCCCAAACAGCGGCGCTCCATCATCCGTGACCGGGGACCCATGTATGATGACCCCACCCTGCCTGAA
GGCTGGACACGGAAGCTTAAGCAAAGGAAATCTGGCCGCTCTGCTGGGAAGTATGATGTGTATTTGATCAATCCCCAG
GGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGGTAGGCGACACATCCCTGGACCCTAATGAT
TTTGACTTCACGGTAACTGGGAGAGGGAGCCCCTCCCGGCGAGAGCAGAAACCACCTAAGAAGCCCAAATCTCCCAAA
GCTCCAGGAACTGGCAGAGGCCGGGGACGCCCCAAAGGGAGCGGCACCACGAGACCCAAGGCGGCCACGTCAGAGGGT
GTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCTGGGAAGCTCCTTGTCAAGATGCCTTTTCAAACTTCGCCAGGGGGC
AAGGCTGAGGGGGGTGGGGCCACCACATCCACCCAGGTCATGGTGATCAAACGCCCCGGCAGGAAGCGAAAAGCTGAG
GCCGACCCTCAGGCCATTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTGTGGTGGCAGCCGCTGCCGCCGAGGCCAAA
AAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGGAGACCGTACTCCCCATCAAGAAGCGCAAGACCCGGGAG
ACGGTCAGCATCGAGGTCAAGGAAGTGGTGAAGCCCCTGCTGGTGTCCACCCTCGGTGAGAAGAGCGGGAAAGGACTG
AAGACCTGTAAGAGCCCTGGGCGGAAAAGCAAGGAGAGCAGCCCCAAGGGGCGCAGCAGCAGCGCCTCCTCACCCCCC
AAGAAGGAGCACCACCACCATCACCACCACTCAGAGTCCCCAAAGGCCCCCGTGCCACTGCTCCCACCCCTGCCCCCA
CCTCCACCTGAGCCCGAGAGCTCCGAGGACCCCACCAGCCCCCCTGAGCCCCAGGACTTGAGCAGCAGCGTCTGCAAA
GAGGAGAAGATGCCCAGAGGAGGCTCACTGGAGAGCGACGGCTGCCCCAAGGAGCCAGCTAAGACTCAGCCCGCGGTT
GCCACCGCCGCCACGGCCGCAGAAAAGTACAAACACCGAGGGGAGGGAGAGCGCAAAGACATTGTTTCATCCTCCATG
CCAAGGCCAAACAGAGAGGAGCCTGTGGACAGCCGGACGCCCGTGACCGAGAGAGTTAGCTGACTTTACACGGAGCGG
ATTGCAAAGCAAACCAACAAGAATAAAGGCAGCTGTTGTCTCTTCTCCTTATGGGTAGGGCTCTGACAAAGCTTCCCG
ATTAACTGAAATAAAAAATATTTTTTTTTCTTTCAGTAAACTTAGAGTTTCGTGGCTTCAGGGTGGGAGTAGTTGGAG
CATTGGGGATGTTTTTCTTACCGACAAGCACAGTCAGGTTGAAGACCTAACCAGGGCCAGAAGTAGCTTTGCACTTTT
CTAAACTAGGCTCCTTCAACAAGGCTTGCTGCAGATACTACTGACCAGACAAGCTGTTGACCAGGCACCTCCCCTCCC
GCCCAAACCTTTCCCCCATGTGGTCGTTAGAGACAGAGCGACAGAGCAGTTGAGAGGACACTCCCGTTTTCGGTGCCA
TCAGTGCCCCGTCTACAGCTCCCCCAGCTCCCCCCACCTCCCCCACTCCCAACCACGTTGGGACAGGGAGGTGTGAGG
CAGGAGAGACAGTTGGATTCTTTAGAGAAGATGGATATGACCAGTGGCTATGGCCTGTGCGATCCCACCCGTGGTGGC
TCAAGTCTGGCCCCACACCAGCCCCAATCCAAAACTGGCAAGGACGCTTCACAGGACAGGAAAGTGGCACCTGTCTGC
TCCAGCTCTGGCATGGCTAGGAGGGGGGAGTCCCTTGAACTACTGGGTGTAGACTGGCCTGAACCACAGGAGAGGATG
GCCCAGGGTGAGGTGGCATGGTCCATTCTCAAGGGACGTCCTCCAACGGGTGGCGCTAGAGGCCATGGAGGCAGTAGG
ACAAGGTGCAGGCAGGCTGGCCTGGGGTCAGGCCGGGCAGAGCACAGCGGGGTGAGAGGGATTCCTAATCACTCAGAG
CAGTCTGTGACTTAGTGGACAGGGGAGGGGGCAAAGGGGGAGGAGAAGAAAATGTTCTTCCAGTTACTTTCCAATTCT
CCTTTAGGGACAGCTTAGAATTATTTGCACTATTGAGTCTTCATGTTCCCACTTCAAAACAAACAGATGCTCTGAGAG
CAAACTGGCTTGAATTGGTGACATTTAGTCCCTCAAGCCACCAGATGTGACAGTGTTGAGAACTACCTGGATTTGTAT
ATATACCTGCGCTTGTTTTAAAGTGGGCTCAGCACATAGGGTTCCCACGAAGCTCCGAAACTCTAAGTGTTTGCTGCA
ATTTTATAAGGACTTCCTGATTGGTTTCTCTTCTCCCCTTCCATTTCTGCCTTTTGTTCATTTCATCCTTTCACTTCT
TTCCCTTCCTCCATCCTCCTCCTTCCTAGTTCATCCCTTCTCTTCCAGGCAGCCGCGGTGCCCAACCACACTTGTCGG
CTCCAGTCCCCAGAACTCTGCCTGCCCTTTGTCCTCCTGCTGCCAGTACCAGCCCCACCCTGTTTTGAGCCCTGAGGA
GGCCTTGGGCTCTGCTGAGTCCGACCTGGCCTGTCTGTGAAGAGCAAGAGAGCAGCAAGGTCTTGCTCTCCTAGGTAG
CCCCCTCTTCCCTGGTAAGAAAAAGCAAAAGGCATTTCCCACCCTGAACAACGAGCCTTTTCACCCTTCTACTCTAGA
GAAGTGGACTGGAGGAGCTGGGCCCGATTTGGTAGTTGAGGAAAGCACAGAGGCCTCCTGTGGCCTGCCAGTCATCGA
GTGGCCCAACAGGGGCTCCATGCCAGCCGACCTTGACCTCACTCAGAAGTCCAGAGTCTAGCGTAGTGCAGCAGGGCA
GTAGCGGTACCAATGCAGAACTCCCAAGACCCGAGCTGGGACCAGTACCTGGGTCCCCAGCCCTTCCTCTGCTCCCCC
TTTTCCCTCGGAGTTCTTCTTGAATGGCAATGTTTTGCTTTTGCTCGATGCAGACAGGGGGCCAGAACACCACACATT
TCACTGTCTGTCTGGTCCATAGCTGTGGTGTAGGGGCTTAGAGGCATGGGCTTGCTGTGGGTTTTTAATTGATCAGTT
TTCATGTGGGATCCCATCTTTTTAACCTCTGTTCAGGAAGTCCTTATCTAGCTGCATATCTTCATCATATTGGTATAT

CCTTTTCTGTGTTTACAGAGATGTCTCTTATATCTAAATCTGTCCAACTGAGAAGTACCTTATCAAAGTAGCAAATGA
GACAGCAGTCTTATGCTTCCAGAAACACCCACAGGCATGTCCCATGTGAGCTGCTGCCATGAACTGTCAAGTGTGTGT
TGTCTTGTGTATTTCAGTTATTGTCCCTGGCTTCCTTACTATGGTGTAATCATGAAGGAGTGAAACATCATAGAAACT
GTCTAGCACTTCCTTGCCAGTCTTTAGTGATCAGGAACCATAGTTGACAGTTCCAATCAGTAGCTTAAGAAAAAACCG
TGTTTGTCTCTTCTGGAATGGTTAGAAGTGAGGGAGTTTGCCCCGTTCTGTTTGTAGAGTCTCATAGTTGGACTTTCT
AGCATATATGTGTCCATTTCCTTATGCTGTAAAAGCAAGTCCTGCAACCAAACTCCCATCAGCCCAATCCCTGATCCC
TGATCCCTTCCACCTGCTCTGCTGATGACCCCCCCCAGCTTCACTTCTGACTCTTCCCCAGGAAGGGAAGGGGGGTCAG
AAGAGAGGGTGAGTCCTCCAGAACTCTTCCTCCAAGGACAGAAGGCTCCTGCCCCCATAGTGGCCTCGAACTCCTGGC
ACTACCAAAGGACACTTATCCACGAGAGCGCAGCATCCGACCAGGTTGTCACTGAGAAGATGTTTATTTTGGTCAGTT
GGGTTTTTATGTATTATACTTAGTCAAATGTAATGTGGCTTCTGGAATCATTGTCCAGAGCTGCTTCCCCGTCACCTG
GGCGTCATCTGGTCCTGGTAAGAGGAGTGCGTGGCCCACCAGGCCCCCCTGTCACCCATGACAGTTCATTCAGGGCCG
ATGGGGCAGTCGTGGTTGGGAACACAGCATTTCAAGCGTCACTTTATTTCATTCGGGCCCCACCTGCAGCTCCCTCAA
AGAGGCAGTTGCCCAGCCTCTTTCCCTTCCAGTTTATTCCAGAGCTGCCAGTGGGGCCTGAGGCTCCTTAGGGTTTTC
TCTCTATTTCCCCCTTTCTTCCTCATTCCCTCGTCTTTCCCAAAGGCATCACGAGTCAGTCGCCTTTCAGCAGGCAGC
CTTGGCGGTTTATCGCCCTGGCAGGCAGGGGCCCTGCAGCTCTCATGCTGCCCCTGCCTTGGGGTCAGGTTGACAGGA
GGTTGGAGGGAAAGCCTTAAGCTGCAGGATTCTCACCAGCTGTGTCCGGCCCAGTTTTGGGGTGTGACCTCAATTTCA
ATTTTGTCTGTACTTGAACATTATGAAGATGGGGGCCTCTTTCAGTGAATTTGTGAACAGCAGAATTGACCGACAGCT
TTCCAGTACCCATGGGGCTAGGTCATTAAGGCCACATCCACAGTCTCCCCCACCCTTGTTCCAGTTGTTAGTTACTAC
CTCCTCTCCTGACAATACTGTATGTCGTCGAGCTCCCCCCAGGTCTACCCCTCCCGGCCCTGCCTGCTGGTGGGCTTG
TCATAGCCAGTGGGATTGCCGGTCTTGACAGCTCAGTGAGCTGGAGATACTTGGTCACAGCCAGGCGCTAGCACAGCT
CCCTTCTGTTGATGCTGTATTCCCATATCAAAAGACACAGGGGACACCCAGAAACGCCACATCCCCCAATCCATCAGT
GCCAAACTAGCCAACGGCCCCAGCTTCTCAGCTCGCTGGATGGCGGAAGCTGCTACTCGTGAGCGCCAGTGCGGGTGC
AGACAATCTTCTGTTGGGTGGCATCATTCCAGGCCCGAAGCATGAACAGTGCACCTGGGACAGGGAGCAGCCCCAAAT
TGTCACCTGCTTCTCTGCCCAGCTTTTCATTGCTGTGACAGTGATGGCGAAAGAGGGTAATAACCAGACACAAACTGC
CAAGTTGGGTGGAGAAAGGAGTTTCTTTAGCTGACAGAATCTCTGAATTTTAAATCACTTAGTAAGCGGCTCAAGCCC
AGGAGGGAGCAGAGGGATACGAGCGGAGTCCCCTGCGCGGGACCATCTGGAATTGGTTTAGCCCAAGTGGAGCCTGAC
AGCCAGAACTCTGTGTCCCCCGTCTAACCACAGCTCCTTTTCCAGAGCATTCCAGTCAGGCTCTCTGGGCTGACTGGG
CCAGGGGAGGTTACAGGTACCAGTTCTTTAAGAAGATCTTTGGGCATATACATTTTTAGCCTGTGTCATTGCCCCAAA
TGGATTCCTGTTTCAAGTTCACACCTGCAGATTCTAGGACCTGTGTCCTAGACTTCAGGGAGTCAGCTGTTTCTAGAG
TTCCTACCATGGAGTGGGTCTGGAGGACCTGCCCGGTGGGGGGGCAGAGCCCTGCTCCCTCCGGGTCTTCCTACTCTT
CTCTCTGCTCTGACGGGATTTGTTGATTCTCTCCATTTTGGTGTCTTTCTCTTTTAGATATTGTATCAATCTTTAGAA
AAGGCATAGTCTACTTGTTATAAATCGTTAGGATACTGCCTCCCCCAGGGTCTAAAATTACATATTAGAGGGGAAAAG
CTGAACACTGAAGTCAGTTCTCAACAATTTAGAAGGAAAACCTAGAAAACATTTGGCAGAAAATTACATTTCGATGTT
TTTGAATGAATACGAGCAAGCTTTTACAACAGTGCTGATCTAAAAATACTTAGCACTTGGCCTGAGATGCCTGGTGAG
CATTACAGGCAAGGGGAATCTGGAGGTAGCCGACCTGAGGACATGGCTTCTGAACCTGTCTTTTGGGAGTGGTATGGA
AGGTGGAGCGTTCACCAGTGACCTGGAAGGCCCAGCACCACCCTCCTTCCCACTCTTCTCATCTTGACAGAGCCTGCC
CCAGCGCTGACGTGTCAGGAAAACACCCAGGGAACTAGGAAGGCACTTCTGCCTGAGGGGCAGCCTGCCTTGCCCACT
CCTGCTCTGCTCGCCTCGGATCAGCTGAGCCTTCTGAGCTGGCCTCTCACTGCCTCCCCAAGGCCCCCTGCCTGCCCT
GTCAGGAGGCAGAAGGAAGCAGGTGTGAGGGCAGTGCAAGGAGGGAGCACAACCCCCAGCTCCCGCTCCGGGCTCCGA
CTTGTGCACAGGCAGAGCCCAGACCCTGGAGGAAATCCTACCTTTGAATTCAAGAACATTTGGGGAATTTGGAAATCT
CTTTGCCCCCAAACCCCCATTCTGTCCTACCTTTAATCAGGTCCTGCTCAGCAGTGAGAGCAGATGAGGTGAAAAGGC
CAAGAGGTTTGGCTCCTGCCCACTGATAGCCCCTCTCCCCGCAGTGTTTGTGTGTCAAGTGGCAAAGCTGTTCTTCCT
GGTGACCCTGATTATATCCAGTAACACATAGACTGTGCGCATAGGCCTGCTTTGTCTCCTCTATCCTGGGCTTTTGTT
TTGCTTTTTAGTTTTGCTTTTAGTTTTTCTGTCCCTTTTATTTAACGCACCGACTAGACACACAAAGCAGTTGAATTT
TTATATATATATCTGTATATTGCACAATTATAAACTCATTTTGCTTGTGGCTCCACACACACAAAAAAAGACCTGTTA
AAATTATACCTGTTGCTTAATTACAATATTTCTGATAACCATAGCATAGGACAAGGGAAAATAAAAAAAGAAAAAAAA
GAAAAAAAAACGACAAATCTGTCTGCTGGTCACTTCTTCTGTCCAAGCAGATTCGTGGTCTTTTCCTCGCTTCTTTCA
AGGGCTTTCCTGTGCCAGGTGAAGGAGGCTCCAGGCAGCACCCAGGTTTTGCACTCTTGTTTCTCCCGTGCTTGTGAA
AGAGGTCCCAAGGTTCTGGGTGCAGGAGCGCTCCCTTGACCTGCTGAAGTCCGGAACGTAGTCGGCACAGCCTGGTCG
CCTTCCACCTCTGGGAGCTGGAGTCCACTGGGGTGGCCTGACTCCCCCAGTCCCCTTCCCGTGACCTGGTCAGGGTGA
GCCCATGTGGAGTCAGCCTCGCAGGCCTCCCTGCCAGTAGGGTCCGAGTGTGTTTCATCCTTCCCACTCTGTCGAGCC
TGGGGGCTGGAGCGGAGACGGGAGGCCTGGCCTGTCTCGGAACCTGTGAGCTGCACCAGGTAGAACGCCAGGGACCCC
AGAATCATGTGCGTCAGTCCAAGGGGTCCCCTCCAGGAGTAGTGAAGACTCCAGAAATGTCCCTTTCTTCTCCCCCAT
CCTACGAGTAATTGCATTTGCTTTTGTAATTCTTAATGAGCAATATCTGCTAGAGAGTTTAGCTGTAACAGTTCTTTT
TGATCATCTTTTTTTAATAATTAGAAACACCAAAAAAATCCAGAAACTTGTTCTTCCAAAGCAGAGAGCATTATAATC
ACCAGGGCCAAAAGCTTCCCTCCCTGCTGTCATTGCTTCTTCTGAGGCCTGAATCCAAAAGAAAAACAGCCATAGGCC
CTTTCAGTGGCCGGGCTACCCGTGAGCCCTTCGGAGGACCAGGGCTGGGGCAGCCTCTGGGCCCACATCCGGGGCCAG
CTCCGGCGTGTGTTCAGTGTTAGCAGTGGGTCATGATGCTCTTTCCCACCCAGCCTGGGATAGGGGCAGAGGAGGCGA
GGAGGCCGTTGCCGCTGATGTTTGGCCGTGAACAGGTGGGTGTCTGCGTGCGTCCACGTGCGTGTTTTCTGACTGACA
TGAAATCGACGCCCGAGTTAGCCTCACCCGGTGACCTCTAGCCCTGCCCGGATGGAGCGGGGCCCACCCGGTTCAGTG
TTTCTGGGGAGCTGGACAGTGGAGTGCAAAAGGCTTGCAGAACTTGAAGCCTGCTCCTTCCCTTGCTACCACGGCCTC
CTTTCCGTTTGATTTGTCACTGCTTCAATCAATAACAGCCGCTCCAGAGTCAGTAGTCAATGAATATATGACCAAATA
TCACCAGGACTGTTACTCAATGTGTGCCGAGCCCTTGCCCATGCTGGGCTCCCGTGTATCTGGACACTGTAACGTGTG

CTGTGTTTGCTCCCCTTCCCCTTCCTTCTTTGCCCTTTACTTGTCTTTCTGGGGTTTTTCTGTTTGGGTTTGGTTTGG
TTTTTATTTCTCCTTTTGTGTTCCAAACATGAGGTTCTCTCTACTGGTCCTCTTAACTGTGGTGTTGAGGCTTATATT
TGTGTAATTTTTGGTGGGTGAAAGGAATTTTGCTAAGTAAATCTCTTCTGTGTTTGAACTGAAGTCTGTATTGTAACT
ATGTTTAAAGTAATTGTTCCAGAGACAAATATTTCTAGACACTTTTTCTTTACAAACAAAAGCATTCGGAGGGAGGGG
GATGGTGACTGAGATGAGAGGGGAGAGCTGAACAGATGACCCCTGCCCAGATCAGCCAGAAGCCACCCAAAGCAGTGG
AGCCCAGGAGTCCCACTCCAAGCCAGCAAGCCGAATAGCTGATGTGTTGCCACTTTCCAAGTCACTGCAAAACCAGGT
TTTGTTCCGCCCAGTGGATTCTTGTTTTGCTTCCCCTCCCCCCGAGATTATTACCACCATCCCGTGCTTTTAAGGAAA
GGCAAGATTGATGTTTCCTTGAGGGGAGCCAGGAGGGGATGTGTGTGTGCAGAGCTGAAGAGCTGGGGAGAATGGGGC
TGGGCCCACCCAAGCAGGAGGCTGGGACGCTCTGCTGTGGGCACAGGTCAGGCTAATGTTGGCAGATGCAGCTCTTCC
TGGACAGGCCAGGTGGTGGGCATTCTCTCTCCAAGGTGTGCCCCGTGGGCATTACTGTTTAAGACACTTCCGTCACAT
CCCACCCCATCCTCCAGGGCTCAACACTGTGACATCTCTATTCCCCACCCTCCCCTTCCCAGGGCAATAAAATGACCA
TGGAGGGGGCTTGCACTCTCTTGGCTGTCACCCGATCGCCAGCAAAACTTAGATGTGAGAAAACCCCTTCCCATTCCA
TGGCGAAAACATCTCCTTAGAAAAGCCATTACCCTCATTAGGCATGGTTTTGGGCTCCCAAAACACCTGACAGCCCCT
CCCTCCTCTGAGAGGCGGAGAGTGCTGACTGTAGTGACCATTGCATGCCGGGTGCAGCATCTGGAAGAGCTAGGCAGG
GTGTCTGCCCCCTCCTGAGTTGAAGTCATGCTCCCCTGTGCCAGCCCAGAGGCCGAGAGCTATGGACAGCATTGCCAG
TAACACAGGCCACCCTGTGCAGAAGGGAGCTGGCTCCAGCCTGGAAACCTGTCTGAGGTTGGGAGAGGTGCACTTGGG
GCACAGGGAGAGGCCGGGACACACTTAGCTGGAGATGTCTCTAAAAGCCCTGTATCGTATTCACCTTCAGTTTTTGTG
TTTTGGGACAATTACTTTAGAAAATAAGTAGGTCGTTTTAAAAACAAAAATTATTGATTGCTTTTTTTGTAGTGTTCAG
AAAAAAGGTTCTTTGTGTATAGCCAAATGACTGAAAGCACTGATATATTTAAAAACAAAAGGCAATTTATTAAGGAAA
TTTGTACCATTTCAGTAAACCTGTCTGAATGTACCTGTATACGTTTCAAAAACACCCCCCCCCCACTGAATCCCTGTA
ACCTATTTATTATATAAAGAGTTTGCCTTATAAATTTACATAAAAATGTC

SEQ ID NO: 36 >XM_011531166.2 MECP2 [organism=Homo sapiens] [GeneID=4204] [transcript=X3]

GCAGTAGTCTCACGTGGCAGACCCGGAGGTCACTCTCCTCTCTGGATGCCCTGCGGACACTCCGGTTAGCCTCAGCTG
CAAGAGCCTCCTCACCCAAGGTCACGCCATTTCCAGGACAGCACCCTGGTGACTGAGCGAGGGAAGAAAAGTCAGAAG
ACCAGGACCTCCAGGGCCTCAAGGACAAACCCCTCAAGTTTAAAAAGGTGAAGAAAGATAAGAAAGAAGAGAAAGAGG
GCAAGCATGAGCCCGTGCAGCCATCAGCCCACCACTCTGCTGAGCCCGCAGAGGCAGGCAAAGCAGAGACATCAGAAG
GGTCAGGCTCCGCCCCGGCTGTGCCGGAAGCTTCTGCCTCCCCCAAACAGCGGCGCTCCATCATCCGTGACCGGGGAC
CCATGTATGATGACCCCACCCTGCCTGAAGGCTGGACACGGAAGCTTAAGCAAAGGAAATCTGGCCGCTCTGCTGGGA
AGTATGATGTGTATTTGATCAATCCCCAGGGAAAAGCCTTTCGCTCTAAAGTGGAGTTGATTGCGTACTTCGAAAAGG
TAGGCGACACATCCCTGGACCCTAATGATTTTGACTTCACGGTAACTGGGAGAGGGAGCCCCTCCCGGCGAGAGCAGA
AACCACCTAAGAAGCCCAAATCTCCCAAAGCTCCAGGAACTGGCAGAGGCCGGGGACGCCCCAAAGGGAGCGGCACCA
CGAGACCCAAGGCGGCCACGTCAGAGGGTGTGCAGGTGAAAAGGGTCCTGGAGAAAAGTCCTGGGAAGCTCCTTGTCA
AGATGCCTTTTCAAACTTCGCCAGGGGGCAAGGCTGAGGGGGGTGGGGCCACCACATCCACCCAGGTCATGGTGATCA
AACGCCCCGGCAGGAAGCGAAAAGCTGAGGCCGACCCTCAGGCCATTCCCAAGAAACGGGGCCGAAAGCCGGGGAGTG
TGGTGGCAGCCGCTGCCGCCGAGGCCAAAAAGAAAGCCGTGAAGGAGTCTTCTATCCGATCTGTGCAGGAGACCGTAC
TCCCCATCAAGAAGCGCAAGACCCGGGAGACGGTCAGCATCGAGGTCAAGGAAGTGGTGAAGCCCCTGCTGGTGTCCA
CCCTCGGTGAGAAGAGCGGGAAAGGACTGAAGACCTGTAAGAGCCCTGGGCGGAAAAGCAAGGAGAGCAGCCCCAAGG
GGCGCAGCAGCAGCGCCTCCTCACCCCCCAAGAAGGAGCACCACCACCATCACCACCACTCAGAGTCCCCAAAGGCCC
CCGTGCCACTGCTCCCACCCCTGCCCCCACCTCCACCTGAGCCCGAGAGCTCCGAGGACCCCACCAGCCCCCCTGAGC
CCCAGGACTTGAGCAGCAGCGTCTGCAAAGAGGAGAAGATGCCCAGAGGAGGCTCACTGGAGAGCGACGGCTGCCCCA
AGGAGCCAGCTAAGACTCAGCCCGCGGTTGCCACCGCCGCCACGGCCGCAGAAAAGTACAAACACCGAGGGGAGGGAG
AGCGCAAAGACATTGTTTCATCCTCCATGCCAAGGCCAAACAGAGAGGAGCCTGTGGACAGCCGGACGCCCGTGACCG
AGAGAGTTAGCTGACTTTACACGGAGCGGATTGCAAAGCAAACCAACAAGAATAAAGGCAGCTGTTGTCTCTTCTCCT
TATGGGTAGGGCTCTGACAAAGCTTCCCGATTAACTGAAATAAAAAATATTTTTTTTTCTTTCAGTAAACTTAGAGTT
TCGTGGCTTCAGGGTGGGAGTAGTTGGAGCATTGGGGATGTTTTTCTTACCGACAAGCACAGTCAGGTTGAAGACCTA
ACCAGGGCCAGAAGTAGCTTTGCACTTTTCTAAACTAGGCTCCTTCAACAAGGCTTGCTGCAGATACTACTGACCAGA
CAAGCTGTTGACCAGGCACCTCCCCTCCCGCCCAAACCTTTCCCCCATGTGGTCGTTAGAGACAGAGCGACAGAGCAG
TTGAGAGGACACTCCCGTTTTCGGTGCCATCAGTGCCCCGTCTACAGCTCCCCCAGCTCCCCCCACCTCCCCCACTCC
CAACCACGTTGGGACAGGGAGGTGTGAGGCAGGAGAGACAGTTGGATTCTTTAGAGAAGATGGATATGACCAGTGGCT
ATGGCCTGTGCGATCCCACCCGTGGTGGCTCAAGTCTGGCCCCACACCAGCCCCAATCCAAAACTGGCAAGGACGCTT
CACAGGACAGGAAAGTGGCACCTGTCTGCTCCAGCTCTGGCATGGCTAGGAGGGGGGAGTCCCTTGAACTACTGGGTG
TAGACTGGCCTGAACCACAGGAGAGGATGGCCCAGGGTGAGGTGGCATGGTCCATTCTCAAGGGACGTCCTCCAACGG
GTGGCGCTAGAGGCCATGGAGGCAGTAGGACAAGGTGCAGGCAGGCTGGCCTGGGGTCAGGCCGGGCAGAGCACAGCG
GGGTGAGAGGGATTCCTAATCACTCAGAGCAGTCTGTGACTTAGTGGACAGGGGAGGGGGCAAAGGGGGAGGAGAAGA
AAATGTTCTTCCAGTTACTTTCCAATTCTCCTTTAGGGACAGCTTAGAATTATTTGCACTATTGAGTCTTCATGTTCC
CACTTCAAAACAAACAGATGCTCTGAGAGCAAACTGGCTTGAATTGGTGACATTTAGTCCCTCAAGCCACCAGATGTG
ACAGTGTTGAGAACTACCTGGATTTGTATATATACCTGCGCTTGTTTTAAAGTGGGCTCAGCACATAGGGTTCCCACG
AAGCTCCGAAACTCTAAGTGTTTGCTGCAATTTTATAAGGACTTCCTGATTGGTTTCTCTTCTCCCCTTCCATTTCTG
CCTTTTGTTCATTTCATCCTTTCACTTCTTTCCCTTCCTCCATCCTCCTCCTTCCTAGTTCATCCCTTCTCTTCCAGG

CAGCCGCGGTGCCCAACCACACTTGTCGGCTCCAGTCCCCAGAACTCTGCCTGCCCTTTGTCCTCCTGCTGCCAGTAC
CAGCCCCACCCTGTTTTGAGCCCTGAGGAGGCCTTGGGCTCTGCTGAGTCCGACCTGGCCTGTCTGTGAAGAGCAAGA
GAGCAGCAAGGTCTTGCTCTCCTAGGTAGCCCCCTCTTCCCTGGTAAGAAAAAGCAAAAGGCATTTCCCACCCTGAAC
AACGAGCCTTTTCACCCTTCTACTCTAGAGAAGTGGACTGGAGGAGCTGGGCCCGATTTGGTAGTTGAGGAAAGCACA
GAGGCCTCCTGTGGCCTGCCAGTCATCGAGTGGCCCAACAGGGGCTCCATGCCAGCCGACCTTGACCTCACTCAGAAG
TCCAGAGTCTAGCGTAGTGCAGCAGGGCAGTAGCGGTACCAATGCAGAACTCCCAAGACCCGAGCTGGGACCAGTACC
TGGGTCCCCAGCCCTTCCTCTGCTCCCCCTTTTCCCTCGGAGTTCTTCTTGAATGGCAATGTTTTGCTTTTGCTCGAT
GCAGACAGGGGGCCAGAACACCACACATTTCACTGTCTGTCTGGTCCATAGCTGTGGTGTAGGGGCTTAGAGGCATGG
GCTTGCTGTGGGTTTTTAATTGATCAGTTTTCATGTGGGATCCCATCTTTTTAACCTCTGTTCAGGAAGTCCTTATCT
AGCTGCATATCTTCATCATATTGGTATATCCTTTTCTGTGTTTACAGAGATGTCTCTTATATCTAAATCTGTCCAACT
GAGAAGTACCTTATCAAAGTAGCAAATGAGACAGCAGTCTTATGCTTCCAGAAACACCCACAGGCATGTCCCATGTGA
GCTGCTGCCATGAACTGTCAAGTGTGTGTTGTCTTGTGTATTTCAGTTATTGTCCCTGGCTTCCTTACTATGGTGTAA
TCATGAAGGAGTGAAACATCATAGAAACTGTCTAGCACTTCCTTGCCAGTCTTTAGTGATCAGGAACCATAGTTGACA
GTTCCAATCAGTAGCTTAAGAAAAAACCGTGTTTGTCTCTTCTGGAATGGTTAGAAGTGAGGGAGTTTGCCCCGTTCT
GTTTGTAGAGTCTCATAGTTGGACTTTCTAGCATATATGTGTCCATTTCCTTATGCTGTAAAAGCAAGTCCTGCAACC
AAACTCCCATCAGCCCAATCCCTGATCCCTGATCCCTTCCACCTGCTCTGCTGATGACCCCCCCAGCTTCACTTCTGA
CTCTTCCCCAGGAAGGGAAGGGGGGTCAGAAGAGAGGGTGAGTCCTCCAGAACTCTTCCTCCAAGGACAGAAGGCTCC
TGCCCCCATAGTGGCCTCGAACTCCTGGCACTACCAAAGGACACTTATCCACGAGAGCGCAGCATCCGACCAGGTTGT
CACTGAGAAGATGTTTATTTTGGTCAGTTGGGTTTTTATGTATTATACTTAGTCAAATGTAATGTGGCTTCTGGAATC
ATTGTCCAGAGCTGCTTCCCCGTCACCTGGGCGTCATCTGGTCCTGGTAAGAGGAGTGCGTGGCCCACCAGGCCCCCC
TGTCACCCATGACAGTTCATTCAGGGCCGATGGGGCAGTCGTGGTTGGGAACACAGCATTTCAAGCGTCACTTTATTT
CATTCGGGCCCCACCTGCAGCTCCCTCAAAGAGGCAGTTGCCCAGCCTCTTTCCCTTCCAGTTTATTCCAGAGCTGCC
AGTGGGGCCTGAGGCTCCTTAGGGTTTTCTCTCTATTTCCCCCTTTCTTCCTCATTCCCTCGTCTTTCCCAAAGGCAT
CACGAGTCAGTCGCCTTTCAGCAGGCAGCCTTGGCGGTTTATCGCCCTGGCAGGCAGGGGCCCTGCAGCTCTCATGCT
GCCCCTGCCTTGGGGTCAGGTTGACAGGAGGTTGGAGGGAAAGCCTTAAGCTGCAGGATTCTCACCAGCTGTGTCCGG
CCCAGTTTTGGGGTGTGACCTCAATTTCAATTTTGTCTGTACTTGAACATTATGAAGATGGGGGCCTCTTTCAGTGAA
TTTGTGAACAGCAGAATTGACCGACAGCTTTCCAGTACCCATGGGGCTAGGTCATTAAGGCCACATCCACAGTCTCCC
CCACCCTTGTTCCAGTTGTTAGTTACTACCTCCTCTCCTGACAATACTGTATGTCGTCGAGCTCCCCCCAGGTCTACC
CCTCCCGGCCCTGCCTGCTGGTGGGCTTGTCATAGCCAGTGGGATTGCCGGTCTTGACAGCTCAGTGAGCTGGAGATA
CTTGGTCACAGCCAGGCGCTAGCACAGCTCCCTTCTGTTGATGCTGTATTCCCATATCAAAAGACACAGGGGACACCC
AGAAACGCCACATCCCCCAATCCATCAGTGCCAAACTAGCCAACGGCCCCAGCTTCTCAGCTCGCTGGATGGCGGAAG
CTGCTACTCGTGAGCGCCAGTGCGGGTGCAGACAATCTTCTGTTGGGTGGCATCATTCCAGGCCCGAAGCATGAACAG
TGCACCTGGGACAGGGAGCAGCCCCAAATTGTCACCTGCTTCTCTGCCCAGCTTTTCATTGCTGTGACAGTGATGGCG
AAAGAGGGTAATAACCAGACACAAACTGCCAAGTTGGGTGGAGAAAGGAGTTTCTTTAGCTGACAGAATCTCTGAATT
TTAAATCACTTAGTAAGCGGCTCAAGCCCAGGAGGGAGCAGAGGGATACGAGCGGAGTCCCCTGCGCGGGACCATCTG
GAATTGGTTTAGCCCAAGTGGAGCCTGACAGCCAGAACTCTGTGTCCCCCGTCTAACCACAGCTCCTTTTCCAGAGCA
TTCCAGTCAGGCTCTCTGGGCTGACTGGGCCAGGGGAGGTTACAGGTACCAGTTCTTTAAGAAGATCTTTGGGCATAT
ACATTTTTAGCCTGTGTCATTGCCCCAAATGGATTCCTGTTTCAAGTTCACACCTGCAGATTCTAGGACCTGTGTCCT
AGACTTCAGGGAGTCAGCTGTTTCTAGAGTTCCTACCATGGAGTGGGTCTGGAGGACCTGCCCGGTGGGGGGGCAGAG
CCCTGCTCCCTCCGGGTCTTCCTACTCTTCTCTCTGCTCTGACGGGATTTGTTGATTCTCTCCATTTTGGTGTCTTTC
TCTTTTAGATATTGTATCAATCTTTAGAAAAGGCATAGTCTACTTGTTATAAATCGTTAGGATACTGCCTCCCCCAGG
GTCTAAAATTACATATTAGAGGGGAAAAGCTGAACACTGAAGTCAGTTCTCAACAATTTAGAAGGAAAACCTAGAAAA
CATTTGGCAGAAAATTACATTTCGATGTTTTTGAATGAATACGAGCAAGCTTTTACAACAGTGCTGATCTAAAAATAC
TTAGCACTTGGCCTGAGATGCCTGGTGAGCATTACAGGCAAGGGGAATCTGGAGGTAGCCGACCTGAGGACATGGCTT
CTGAACCTGTCTTTTGGGAGTGGTATGGAAGGTGGAGCGTTCACCAGTGACCTGGAAGGCCCAGCACCACCCTCCTTC
CCACTCTTCTCATCTTGACAGAGCCTGCCCCAGCGCTGACGTGTCAGGAAAACACCCAGGGAACTAGGAAGGCACTTC
TGCCTGAGGGGCAGCCTGCCTTGCCCACTCCTGCTCTGCTCGCCTCGGATCAGCTGAGCCTTCTGAGCTGGCCTCTCA
CTGCCTCCCCAAGGCCCCCTGCCTGCCCTGTCAGGAGGCAGAAGGAAGCAGGTGTGAGGGCAGTGCAAGGAGGGAGCA
CAACCCCCAGCTCCCGCTCCGGGCTCCGACTTGTGCACAGGCAGAGCCCAGACCCTGGAGGAAATCCTACCTTTGAAT
TCAAGAACATTTGGGGAATTTGGAAATCTCTTTGCCCCCAAACCCCCATTCTGTCCTACCTTTAATCAGGTCCTGCTC
AGCAGTGAGAGCAGATGAGGTGAAAAGGCCAAGAGGTTTGGCTCCTGCCCACTGATAGCCCCTCTCCCCGCAGTGTTT
GTGTGTCAAGTGGCAAAGCTGTTCTTCCTGGTGACCCTGATTATATCCAGTAACACATAGACTGTGCGCATAGGCCTG
CTTTGTCTCCTCTATCCTGGGCTTTTGTTTTGCTTTTTAGTTTTGCTTTTAGTTTTTCTGTCCCTTTTATTTAACGCA
CCGACTAGACACACAAAGCAGTTGAATTTTTATATATATATCTGTATATTGCACAATTATAAACTCATTTTGCTTGTG
GCTCCACACACACAAAAAAAGACCTGTTAAAATTATACCTGTTGCTTAATTACAATATTTCTGATAACCATAGCATAG
GACAAGGGAAAATAAAAAAAGAAAAAAAAGAAAAAAAAACGACAAATCTGTCTGCTGGTCACTTCTTCTGTCCAAGCA
GATTCGTGGTCTTTTCCTCGCTTCTTTCAAGGGCTTTCCTGTGCCAGGTGAAGGAGGCTCCAGGCAGCACCCAGGTTT
TGCACTCTTGTTTCTCCCGTGCTTGTGAAAGAGGTCCCAAGGTTCTGGGTGCAGGAGCGCTCCCTTGACCTGCTGAAG
TCCGGAACGTAGTCGGCACAGCCTGGTCGCCTTCCACCTCTGGGAGCTGGAGTCCACTGGGGTGGCCTGACTCCCCCA
GTCCCCCTTCCCGTGACCTGGTCAGGGTGAGCCCATGTGGAGTCAGCCTCGCAGGCCTCCCTGCCAGTAGGGTCCGAGT
GTGTTTCATCCTTCCCACTCTGTCGAGCCTGGGGGCTGGAGCGGAGACGGGAGGCCTGGCCTGTCTCGGAACCTGTGA
GCTGCACCAGGTAGAACGCCAGGGACCCCAGAATCATGTGCGTCAGTCCAAGGGGTCCCCTCCAGGAGTAGTGAAGAC
TCCAGAAATGTCCCTTTCTTCTCCCCCATCCTACGAGTAATTGCATTTGCTTTTGTAATTCTTAATGAGCAATATCTG

CTAGAGAGTTTAGCTGTAACAGTTCTTTTTGATCATCTTTTTTTAATAATTAGAAACACCAAAAAAATCCAGAAACTT
GTTCTTCCAAAGCAGAGAGCATTATAATCACCAGGGCCAAAAGCTTCCCTCCCTGCTGTCATTGCTTCTTCTGAGGCC
TGAATCCAAAAGAAAAACAGCCATAGGCCCTTTCAGTGGCCGGGCTACCCGTGAGCCCTTCGGAGGACCAGGGCTGGG
GCAGCCTCTGGGCCCACATCCGGGGCCAGCTCCGGCGTGTGTTCAGTGTTAGCAGTGGGTCATGATGCTCTTTCCCAC
CCAGCCTGGGATAGGGGCAGAGGAGGCGAGGAGGCCGTTGCCGCTGATGTTTGGCCGTGAACAGGTGGGTGTCTGCGT
GCGTCCACGTGCGTGTTTTCTGACTGACATGAAATCGACGCCCGAGTTAGCCTCACCCGGTGACCTCTAGCCCTGCCC
GGATGGAGCGGGGCCCACCCGGTTCAGTGTTTCTGGGGAGCTGGACAGTGGAGTGCAAAAGGCTTGCAGAACTTGAAG
CCTGCTCCTTCCCTTGCTACCACGGCCTCCTTTCCGTTTGATTTGTCACTGCTTCAATCAATAACAGCCGCTCCAGAG
TCAGTAGTCAATGAATATATGACCAAATATCACCAGGACTGTTACTCAATGTGTGCCGAGCCCTTGCCCATGCTGGGC
TCCCGTGTATCTGGACACTGTAACGTGTGCTGTGTTTGCTCCCCTTCCCCTTCCTTCTTTGCCCTTTACTTGTCTTTC
TGGGGTTTTTCTGTTTGGGTTTGGTTTGGTTTTTATTTCTCCTTTTGTGTTCCAAACATGAGGTTCTCTCTACTGGTC
CTCTTAACTGTGGTGTTGAGGCTTATATTTGTGTAATTTTTGGTGGGTGAAAGGAATTTTGCTAAGTAAATCTCTTCT
GTGTTTGAACTGAAGTCTGTATTGTAACTATGTTTAAAGTAATTGTTCCAGAGACAAATATTTCTAGACACTTTTTCT
TTACAAACAAAAGCATTCGGAGGGAGGGGGATGGTGACTGAGATGAGAGGGGAGAGCTGAACAGATGACCCCTGCCCA
GATCAGCCAGAAGCCACCCAAAGCAGTGGAGCCCAGGAGTCCCACTCCAAGCCAGCAAGCCGAATAGCTGATGTGTTG
CCACTTTCCAAGTCACTGCAAAACCAGGTTTTGTTCCGCCCAGTGGATTCTTGTTTTGCTTCCCCTCCCCCCGAGATT
ATTACCACCATCCCGTGCTTTTAAGGAAAGGCAAGATTGATGTTTCCTTGAGGGGAGCCAGGAGGGGATGTGTGTGTG
CAGAGCTGAAGAGCTGGGGAGAATGGGGCTGGGCCCACCCAAGCAGGAGGCTGGGACGCTCTGCTGTGGGCACAGGTC
AGGCTAATGTTGGCAGATGCAGCTCTTCCTGGACAGGCCAGGTGGTGGGCATTCTCTCTCCAAGGTGTGCCCCGTGGG
CATTACTGTTTAAGACACTTCCGTCACATCCCACCCCATCCTCCAGGGCTCAACACTGTGACATCTCTATTCCCCACC
CTCCCCTTCCCAGGGCAATAAAATGACCATGGAGGGGGCTTGCACTCTCTTGGCTGTCACCCGATCGCCAGCAAAACT
TAGATGTGAGAAAACCCCTTCCCATTCCATGGCGAAAACATCTCCTTAGAAAAGCCATTACCCTCATTAGGCATGGTT
TTGGGCTCCCAAAACACCTGACAGCCCCTCCCTCCTCTGAGAGGCGGAGAGTGCTGACTGTAGTGACCATTGCATGCC
GGGTGCAGCATCTGGAAGAGCTAGGCAGGGTGTCTGCCCCCTCCTGAGTTGAAGTCATGCTCCCCTGTGCCAGCCCAG
AGGCCGAGAGCTATGGACAGCATTGCCAGTAACACAGGCCACCCTGTGCAGAAGGGAGCTGGCTCCAGCCTGGAAACC
TGTCTGAGGTTGGGAGAGGTGCACTTGGGGCACAGGGAGAGGCCGGGACACACTTAGCTGGAGATGTCTCTAAAAGCC
CTGTATCGTATTCACCTTCAGTTTTTGTGTTTTGGGACAATTACTTTAGAAAATAAGTAGGTCGTTTTAAAAACAAAA
ATTATTGATTGCTTTTTTGTAGTGTTCAGAAAAAAGGTTCTTTGTGTATAGCCAAATGACTGAAAGCACTGATATATT
TAAAAACAAAAGGCAATTTATTAAGGAAATTTGTACCATTTCAGTAAACCTGTCTGAATGTACCTGTATACGTTTCAA
AAACACCCCCCCCCCACTGAATCCCTGTAACCTATTTATTATATAAAGAGTTTGCCTTATAAATTTACATAAAAATGT
C

SEQ ID NO: 37 >NP_001104262.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [isoform=2]

MAAAAAAAPSGGGGGGEEERLEEKSEDQDLQGLKDKPLKFKKVKKDKKEEKEGKHEPVQPSAHHSAEPAEAGKAETSE
GSGSAPAVPEASASPKQRRSIIRDRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELIAYFEK
VGDTSLDPNDFDFTVTGRGSPSRREQKPPKKPKSPKAPGTGRGRGRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLV
KMPFQTSPGGKAEGGGATTSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETV
LPIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHSESPKA
PVPLLPPLPPPPPEPESSEDPTSPPEPQDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQPAVATAATAAEKYKHRGEG
ERKDIVSSSMPRPNREEPVDSRTPVTERVS

SEQ ID NO: 38 >NP_001303266.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [isoform=3]

MYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTSLDPNDFDFTVTGRGSPSRREQK
PPKKPKSPKAPGTGRGRGRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPFQTSPGGKAEGGGATTSTQVMVIK
RPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETVLPIKKRKTRETVSIEVKEVVKPLLVST
LGEKSGKGLKTCKSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHSESPKAPVPLLPPLPPPPPEPESSEDPTSPPEP
QDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQPAVATAATAAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTE

SEQ ID NO: 39 >NP_001373066.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [isoform=4]

MPFQTSPGGKAEGGGATTSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETVL
PIKKRKTRETVSIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHSESPKAP
VPLLPPLPPPPPEPESSEDPTSPPEPQDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQPAVATAATAAEKYKHRGEGE
RKDIVSSSMPRPNREEPVDSRTPVTERVS

SEQ ID NO: 40 >NP_004983.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [isoform=1]

MVAGMLGLREEKSEDQDLQGLKDKPLKFKKVKKDKKEEKEGKHEPVQPSAHHSAEPAEAGKAETSEGSGSAPAVPEAS
ASPKQRRSIIRDRGPMYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTSLDPNDFD
FTVTGRGSPSRREQKPPKKPKSPKAPGTGRGRGRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPFQTSPGGKA
EGGGATTSTQVMVIKRPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETVLPIKKRKTRETV
SIEVKEVVKPLLVSTLGEKSGKGLKTCKSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHSESPKAPVPLLPPLPPPP
PEPESSEDPTSPPEPQDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQPAVATAATAAEKYKHRGEGERKDIVSSSMPR
PNREEPVDSRTPVTERVS

SEQ ID NO: 41 >XP_011529468.1 MECP2 [organism=Homo sapiens] [GeneID=4204] [isoform=X1]

MYDDPTLPEGWTRKLKQRKSGRSAGKYDVYLINPQGKAFRSKVELIAYFEKVGDTSLDPNDFDFTVTGRGSPSRREQK
PPKKPKSPKAPGTGRGRGRPKGSGTTRPKAATSEGVQVKRVLEKSPGKLLVKMPFQTSPGGKAEGGGATTSTQVMVIK
RPGRKRKAEADPQAIPKKRGRKPGSVVAAAAAEAKKKAVKESSIRSVQETVLPIKKRKTRETVSIEVKEVVKPLLVST
LGEKSGKGLKTCKSPGRKSKESSPKGRSSSASSPPKKEHHHHHHHSESPKAPVPLLPPLPPPPPEPESSEDPTSPPEP
QDLSSSVCKEEKMPRGGSLESDGCPKEPAKTQPAVATAATAAEKYKHRGEGERKDIVSSSMPRPNREEPVDSRTPVTE
RVS

SEQ ID NO: 42 >sp|P11940|PABP1_HUMAN Polyadenylate-binding protein 1 OS=Homo sapiens OX=9606 GN=PABPC1 PE=1 SV=2

MNPSAPSYPMASLYVGDLHPDVTEAMLYEKFSPAGPILSIRVCRDMITRRSLGYAYVNFQQPADAERALDTMNFDVIK
GKPVRIMWSQRDPSLRKSGVGNIFIKNLDKSIDNKALYDTFSAFGNILSCKVVCDENGSKGYGFVHFETQEAAERAIE
KMNGMLLNDRKVFVGRFKSRKEREAELGARAKEFTNVYIKNFGEDMDDERLKDLFGKFGPALSVKVMTDESGKSKGFG
FVSFERHEDAQKAVDEMNGKELNGKQIYVGRAQKKVERQTELKRKFEQMKQDRITRYQGVNLYVKNLDDGIDDERLRK
EFSPFGTITSAKVMMEGGRSKGFGFVCFSSPEEATKAVTEMNGRIVATKPLYVALAQRKEERQAHLTNQYMQRMASVR
AVPNPVINPYQPAPPSGYFMAAIPQTQNRAAYYPPSQIAQLRPSPRWTAQGARPHPFQNMPGAIRPAAPRPPFSTMRP
ASSQVPRVMSTQRVANTSTQTMGPRPAAAAAAATPAVRTVPQYKYAAGVRNPQQHLNAQPQVTMQQPAVHVQGQEPLT
ASMLASAPPQEQKQMLGERLFPLIQAMHPTLAGKITGMLLEIDNSELLHMLESPESLRSKVDEAVAVLQAHQAKEAAQ
KAVNSATGVPTV

SEQ ID NO: 43 >sp|P11940-2|PABP1_HUMAN Isoform 2 of Polyadenylate-binding protein 1 OS=Homo sapiens OX=9606 GN=PABPC1

MNPSAPSYPMASLYVGDLHPDVTEAMLYEKFSPAGPILSIRVCRDMITRRSLGYAYVNFQQPADAERALDTMNFDVIK
GKPVRIMWSQRDPSLRKSGVGNIFIKNLDKSIDNKALYDTFSAFGNILSCKVVCDENGSKGYGFVHFETQEAAERAIE
KMNGMLLNDRKVFVGRFKSRKEREAELGARAKEFTNVYIKNFGEDMDDERLKDLFGKFGPALSVKVMTDESGKSKGFG
FVSFERHEDAQKAVDEMNGKELNGKQIYVGRAQKKVERQTELKRKFEQMKQDRITRYQGVNLYVKNLDDGIDDERLRK
EFSPFGTITSAKVMMEGGRSKGFGFVCFSSPEEATKAVTEMNGRIVATKPLYVALAQRKEERQAHLTNQYMQRMASVR
AVPNPVINPYQPAPPSGYFMAAIPQTQNRAAYYPPSQIAQLRPSPRWTAQGARPHPAVHVQGQEPLTASMLASAPPQE
QKQMLGERLFPLIQAMHPTLAGKITGMLLEIDNSELLHMLESPESLRSKVDEAVAVLQAHQAKEAAQKAVNSATGVPT
V

SEQ ID NO: 44 >sp|Q9H0A0|NAT10_HUMAN RNA cytidine acetyltransferase OS=Homo sapiens OX=9606 GN=NAT10 PE=1 SV=2

MHRKKVDNRIRILIENGVAERQRSLFVVVGDRGKDQVVILHHMLSKATVKARPSVLWCYKKELGFSSHRKKRMRQLQK
KIKNGTLNIKQDDPFELFIAATNIRYCYYNETHKILGNTFGMCVLQDFEALTPNLLARTVETVEGGGLVVILLRTMNS
LKQLYTVTMDVHSRYRTEAHQDVVGRFNERFILSLASCKKCLVIDDQLNILPISSHVATMEALPPQTPDESLGPSDLE
LRELKESLQDTQPVGVLVDCCKTLDQAKAVLKFIEGISEKTLRSTVALTAARGRGKSAALGLAIAGAVAFGYSNIFVT
SPSPDNLHTLFEFVFKGFDALQYQEHLDYEIIQSLNPEFNKAVIRVNVFREHRQTIQYIHPADAVKLGQAELVVIDEA
AAIPLPLVKSLLGPYLVFMASTINGYEGTGRSLSLKLIQQLRQQSAQSQVSTTAENKTTTTARLASARTLYEVSLQES
IRYAPGDAVEKWLNDLLCLDCLNITRIVSGCPLPEACELYYVNRDTLFCYHKASEVFLQRLMALYVASHYKNSPNDLQ
MLSDAPAHHLFCLLPPVPPTQNALPEVLAVIQVCLEGEISRQSILNSLSRGKKASGDLIPWTVSEQFQDPDFGGLSGG
RVVRIAVHPDYQGMGYGSRALQLLQMYYEGRFPCLEEKVLETPQEIHTVSSEAVSLLEEVITPRKDLPPLLLKLNERP

AERLDYLGVSYGLTPRLLKFWKRAGFVPVYLRQTPNDLTGEHSCIMLKTLTDEDEADQGGWLAAFWKDFRRRFLALLS
YQFSTFSPSLALNIIQNRNMGKPAQPALSREELEALFLPYDLKRLEMYSRNMVDYHLIMDMIPAISRIYFLNQLGDLA
LSAAQSALLLGIGLQHKSVDQLEKEIELPSGQLMGLFNRIIRKVVKLFNEVQEKAIEEQMVAAKDVVMEPTMKTLSDD
LDEAAKEFQEKHKKEVGKLKSMDLSEYIIRGDDEEWNEVLNKAGPNASIISLKSDKKRKLEAKQEPKQSKKLKNRETK
NKKDMKLKRKK

SEQ ID NO: 45 >sp|Q9H0A0-2|NAT10_HUMAN Isoform 2 of RNA cytidine acetyltransferase OS=Homo sapiens OX=9606 GN=NAT10

MRQLQKKIKNGTLNIKQDDPFELFIAATNIRYCYYNETHKILGNTFGMCVLQDFEALTPNLLARTVETVEGGGLVVIL
LRTMNSLKQLYTVTMDVHSRYRTEAHQDVVGRFNERFILSLASCKKCLVIDDQLNILPISSHVATMEALPPQTPDESL
GPSDLELRELKESLQDTQPVGVLVDCCKTLDQAKAVLKFIEGISEKTLRSTVALTAARGRGKSAALGLAIAGAVAFGY
SNIFVTSPSPDNLHTLFEFVFKGFDALQYQEHLDYEIIQSLNPEFNKAVIRVNVFREHRQTIQYIHPADAVKLGQAEL
VVIDEAAAIPLPLVKSLLGPYLVFMASTINGYEGTGRSLSLKLIQQLRQQSAQSQVSTTAENKTTTTARLASARTLYE
VSLQESIRYAPGDAVEKWLNDLLCLDCLNITRIVSGCPLPEACELYYVNRDTLFCYHKASEVFLQRLMALYVASHYKN
SPNDLQMLSDAPAHHLFCLLPPVPPTQNALPEVLAVIQVCLEGEISRQSILNSLSRGKKASGDLIPWTVSEQFQDPDF
GGLSGGRVVRIAVHPDYQGMGYGSRALQLLQMYYEGRFPCLEEKVLETPQEIHTVSSEAVSLLEEVITPRKDLPPLLL
KLNERPAERLDYLGVSYGLTPRLLKFWKRAGFVPVYLRQTPNDLTGEHSCIMLKTLTDEDEADQGGWLAAFWKDFRRR
FLALLSYQFSTFSPSLALNIIQNRNMGKPAQPALSREELEALFLPYDLKRLEMYSRNMVDYHLIMDMIPAISRIYFLN
QLGDLALSAAQSALLLGIGLQHKSVDQLEKEIELPSGQLMGLFNRIIRKVVKLFNEVQEKAIEEQMVAAKDVVMEPTM
KTLSDDLDEAAKEFQEKHKKEVGKLKSMDLSEYIIRGDDEEWNEVLNKAGPNASIISLKSDKKRKLEAKQEPKQSKKL
KNRETKNKKDMKLKRKK

SEQ ID NO: 46 spCas9

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNR
ICYLQEIFSNE**M**AKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAH**M**IKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI
TKAPLSAS**M**IKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK**M**DGTEELL
VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW**M**TRK
SEETITPWNFEEVVDKGASAQSFIER**M**TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG**M**RKPAFLSGEQKK
AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT
LFEDRE**M**IEERLKTYAHLFDDKV**M**KQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF**M**QLIHDDSL
TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV**M**GRHKPENIVIE**M**ARENQTTQKGQKNSRER
MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD**M**YVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT
RSDKNRGKSDNVPSEEVVKK**M**KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD
SR**M**NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK
VYDVRK**M**IAKSEQEIGKATAKYFFYSNI**M**NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLS**M**PQV
NIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITI**M**E
RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKR**M**LASAGELQKGNELALPSKYVNFLYLASHYEKLKGS
PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY
FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

SEQ ID NO: 47 dCas13

MNIPALVENQKKYFGTYSVMAMLNAQTVLDHIQKVADIEGEQNENNENLWFHPVMSHLYNAKNGYDKQPEKTMFIIER
LQSYFPFLKIMAENQREYSNGKYKQNRVEVNSNDIFEVLKRAFGVLKMYRDLTNAYKTYEEKLNDGCEFLTSTEQPLS
GMINNYYTVALRNMNERYGYKTEDLAFIQDKRFKFVKDAYGKKKSQVNTGFFLSLQDYNGDTQKKLHLSGVGIALLIC
LFLDKQYINIFLSRLPIFSSYNAQSEERRIIIRSFGINSIKLPKDRIHSEKSNKSVAMDMLNEVKRCPDELFTTLSAE
KQSRFRIISDDHNEVLMKRSSDRFVPLLLQYIDYGKLFDHIRFHVNMGKLRYLLKADKTCIDGQTRVRVIEQPLNGFG
RLEEAETMRKQENGTFGNSGIRIRDFENMKRDDANPANYPYIVDTYTHYILENNKVEMFINDKEDSAPLLPVIEDDRY
VVKTIPSCRMSTLEIPAMAFHMFLFGSKKTEKLIVDVHNRYKRLFQAMQKEEVTAENIASFGIAESDLPQKILDLISG
NAHGKDVDAFIRLTVDDMLTDTERRIKRFKDDRKSIRSADNKMGKRGFKQISTGKLADFLAKDIVLFQPSVNDGENKI
TGLNYRIMQSAIAVYDSGDDYEAKQQFKLMFEKARLIGKGTTEPHPFLYKVFARSIPANAVEFYERYLIERKFYLTGL
SNEIKKGNRVDVPFIRRDQNKWKTPAMKTLGRIYSEDLPVELPRQMFDNEIKSHLKSLPQMEGIDFNNANVTYLIAEY
MKRVLDDDFQTFYQWNRNYRYMDMLKGEYDRKGSLQHCFTSVEEREGLWKERASRTERYRKQASNKIRSNRQMRNASS
EEIETILDKRLSNSRNEYQKSEKVIRRYRVQDALLFLLAKKTLTELADFDGERFKLKEIMPDAEKGILSEIMPMSFTF
EKGGKKYTITSEGMKLKNYGDFFVLASDKRIGNLLELVGSDIVSKEDIMEEFNKYDQCRPEISSIVFNLEKWAFDTYP

ELSARVDREEKVDFKSILKILLNNKNINKEQSDILRKIRNAFDANNYPDKGVVEIKALPEIAMSIKKAFGEYAIMKGS
LQ

SEQ ID NO: 48 PABC1-dCas13 Fusion Protein

MNIPALVENQKKYFGTYSVMAMLNAQTVLDHIQKVADIEGEQNENNENLWFHPVMSHLYNAKNGYDKQPEKTMFIIER
LQSYFPFLKIMAENQREYSNGKYKQNRVEVNSNDIFEVLKRAFGVLKMYRDLTNAYKTYEEKLNDGCEFLTSTEQPLS
GMINNYYTVALRNMNERYGYKTEDLAFIQDKRFKFVKDAYGKKKSQVNTGFFLSLQDYNGDTQKKLHLSGVGIALLIC
LFLDKQYINIFLSRLPIFSSYNAQSEERRIIIRSFGINSIKLPKDRIHSEKSNKSVAMDMLNEVKRCPDELFTTLSAE
KQSRFRIISDDHNEVLMKRSSDRFVPLLLQYIDYGKLFDHIRFHVNMGKLRYLLKADKTCIDGQTRVRVIEQPLNGFG
RLEEAETMRKQENGTFGNSGIRIRDFENMKRDDANPANYPYIVDTYTHYILENNKVEMFINDKEDSAPLLPVIEDDRY
VVKTIPSCRMSTLEIPAMAFHMFLFGSKKTEKLIVDVHNRYKRLFQAMQKEEVTAENIASFGIAESDLPQKILDLISG
NAHGKDVDAFIRLTVDDMLTDTERRIKRFKDDRKSIRSADNKMGKRGFKQISTGKLADFLAKDIVLFQPSVNDGENKI
TGLNYRIMQSAIAVYDSGDDYEAKQQFKLMFEKARLIGKGTTEPHPFLYKVFARSIPANAVEFYERYLIERKFYLTGL
SNEIKKGNRVDVPFIRRDQNKWKTPAMKTLGRIYSEDLPVELPRQMFDNEIKSHLKSLPQMEGIDFNNANVTYLIAEY
MKRVLDDDFQTFYQWNRNYRYMDMLKGEYDRKGSLQHCFTSVEEREGLWKERASRTERYRKQASNKIRSNRQMRNASS
EEIETILDKRLSNSRNEYQKSEKVIRRYRVQDALLFLLAKKTLTELADFDGERFKLKEIMPDAEKGILSEIMPMSFTF
EKGGKKYTITSEGMKLKNYGDFFVLASDKRIGNLLELVGSDIVSKEDIMEEFNKYDQCRPEISSIVFNLEKWAFDTYP
ELSARVDREEKVDFKSILKILLNNKNINKEQSDILRKIRNAFDANNYPDKGVVEIKALPEIAMSIKKAFGEYAIMKGS
LQMNPSAPSYPMASLYVGDLHPDVTEAMLYEKFSPAGPILSIRVCRDMITRRSLGYAYVNFQQPADAERALDTMNFDV
IKGKPVRIMWSQRDPSLRKSGVGNIFIKNLDKSIDNKALYDTFSAFGNILSCKVVCDENGSKGYGFVHFETQEAAERA
IEKMNGMLLNDRKVFVGRFKSRKEREAELGARAKEFTNVYIKNFGEDMDDERLKDLFGKFGPALSVKVMTDESGKSKG
FGFVSFERHEDAQKAVDEMNGKELNGKQIYVGRAQKKVERQTELKRKFEQMKQDRITRYQGVNLYVKNLDDGIDDERL
RKEFSPFGTITSAKVMMEGGRSKGFGFVCFSSPEEATKAVTEMNGRIVATKPLYVALAQRKEERQAHLTNQYMQRMAS
VRAVPNPVINPYQPAPPSGYFMAAIPQTQNRAAYYPPSQIAQLRPSPRWTAQGARPHPFQNMPGAIRPAAPRPPFSTM
RPASSQVPRVMSTQRVANTSTQTMGPRPAAAAAAATPAVRTVPQYKYAAGVRNPQQHLNAQPQVTMQQPAVHVQGQEP
LTASMLASAPPQEQKQMLGERLFPLIQAMHPTLAGKITGMLLEIDNSELLHMLESPESLRSKVDEAVAVLQAHQAKEA
AQKAVNSATGVPTV

SEQ ID NO: 49 PABC1-dCas13 Fusion Protein

MNIPALVENQKKYFGTYSVMAMLNAQTVLDHIQKVADIEGEQNENNENLWFHPVMSHLYNAKNGYDKQPEKTMFIIER
LQSYFPFLKIMAENQREYSNGKYKQNRVEVNSNDIFEVLKRAFGVLKMYRDLTNAYKTYEEKLNDGCEFLTSTEQPLS
GMINNYYTVALRNMNERYGYKTEDLAFIQDKRFKFVKDAYGKKKSQVNTGFFLSLQDYNGDTQKKLHLSGVGIALLIC
LFLDKQYINIFLSRLPIFSSYNAQSEERRIIIRSFGINSIKLPKDRIHSEKSNKSVAMDMLNEVKRCPDELFTTLSAE
KQSRFRIISDDHNEVLMKRSSDRFVPLLLQYIDYGKLFDHIRFHVNMGKLRYLLKADKTCIDGQTRVRVIEQPLNGFG
RLEEAETMRKQENGTFGNSGIRIRDFENMKRDDANPANYPYIVDTYTHYILENNKVEMFINDKEDSAPLLPVIEDDRY
VVKTIPSCRMSTLEIPAMAFHMFLFGSKKTEKLIVDVHNRYKRLFQAMQKEEVTAENIASFGIAESDLPQKILDLISG
NAHGKDVDAFIRLTVDDMLTDTERRIKRFKDDRKSIRSADNKMGKRGFKQISTGKLADFLAKDIVLFQPSVNDGENKI
TGLNYRIMQSAIAVYDSGDDYEAKQQFKLMFEKARLIGKGTTEPHPFLYKVFARSIPANAVEFYERYLIERKFYLTGL
SNEIKKGNRVDVPFIRRDQNKWKTPAMKTLGRIYSEDLPVELPRQMFDNEIKSHLKSLPQMEGIDFNNANVTYLIAEY
MKRVLDDDFQTFYQWNRNYRYMDMLKGEYDRKGSLQHCFTSVEEREGLWKERASRTERYRKQASNKIRSNRQMRNASS
EEIETILDKRLSNSRNEYQKSEKVIRRYRVQDALLFLLAKKTLTELADFDGERFKLKEIMPDAEKGILSEIMPMSFTF
EKGGKKYTITSEGMKLKNYGDFFVLASDKRIGNLLELVGSDIVSKEDIMEEFNKYDQCRPEISSIVFNLEKWAFDTYP
ELSARVDREEKVDFKSILKILLNNKNINKEQSDILRKIRNAFDANNYPDKGVVEIKALPEIAMSIKKAFGEYAIMKGS
LQLENPSAPSYPMASLYVGDLHPDVTEAMLYEKFSPAGPILSIRVCRDMITRRSLGYAYVNFQQPADAERALDTMNFD
VIKGKPVRIMWSQRDPSLRKSGVGNIFIKNLDKSIDNKALYDTFSAFGNILSCKVVCDENGSKGYGFVHFETQEAAER
AIEKMNGMLLNDRKVFVGRFKSRKEREAELGARAKEFTNVYIKNFGEDMDDERLKDLFGKFGPALSVKVMTDESGKSK
GFGFVSFERHEDAQKAVDEMNGKELNGKQIYVGRAQKKVERQTELKRKFEQMKQDRITRYQGVNLYVKNLDDGIDDER
LRKEFSPFGTITSAKVMMEGGRSKGFGFVCFSSPEEATKAVTEMNGRIVATKPLYVALAQRKEERQAHLTNQYMQRMA
SVRAVPNPVINPYQPAPPSGYFMAAIPQTQNRAAYYPPSQIAQLRPSPRWTAQGARPHPFQNMPGAIRPAAPRPPFST
MRPASSQVPRVMSTQRVANTSTQTMGPRPAAAAAAATPAVRTVPQYKYAAGVRNPQQHLNAQPQVTMQQPAVHVQGQE
PLTASMLASAPPQEQKQMLGERLFPLIQAMHPTLAGKITGMLLEIDNSELLHMLESPESLRSKVDEAVAVLQAHQAKE
AAQKAVNSATGVPTV*

SEQ ID NO: 50 >pJC1211 (dCAS13-PABP) (10,647 bp)

CTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGAT
TTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGGAATGTG
TGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCA
ACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATA
GTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATT

TTTTTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGC
CTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTT
CGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCA
CAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACC
GACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGC
GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTG
TCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCT
ACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAG
GATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGC
GAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATC
GACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGC
GGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTT
CTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGAGATT
TCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATGATCCTCCAGC
GCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATA
GCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTT
ATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTT
ATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAAC
TCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCC
AACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTC
GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGA
ACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCC
CCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTT
CGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCT
GTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGAC
ACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT
TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGA
TTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACT
CACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTA
AATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGA
TCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAA
GGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAA
GTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTA
TGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCT
CCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATT
CTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTA
TGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCA
TCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTC
GTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCG
CAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATC
AGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTC
CCCGAAAAGTGCCACCTGACGTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTC
TGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTA
AGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGA
TGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
TAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCC
ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTT
ACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAA
ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT
CGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG
TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTC
CGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCTGGCTAACTAGAGA
ACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTAAG
CTTGCCACCATGAACATCCCCGCTCTGGTGGAAAACCAGAAGAAGTACTTTGGCACCTACAGCGTGATGGCCATGCTG
AACGCTCAGACCGTGCTGGACCACATCCAGAAGGTGGCCGATATTGAGGGCGAGCAGAACGAGAACAACGAGAATCTG
TGGTTTCACCCCGTGATGAGCCACCTGTACAACGCCAAGAACGGCTACGACAAGCAGCCCGAGAAAACCATGTTCATC
ATCGAGCGGCTGCAGAGCTACTTCCCATTCCTGAAGATCATGGCCGAGAACCAGAGAGAGTACAGCAACGGCAAGTAC
AAGCAGAACCGCGTGGAAGTGAACAGCAACGACATCTTCGAGGTGCTGAAGCGCGCCTTCGGCGTGCTGAAGATGTAC
AGGGACCTGACCAACGCATACAAGACCTACGAGGAAAAGCTGAACGACGGCTGCGAGTTCCTGACCAGCACAGAGCAA
CCTCTGAGCGGCATGATCAACAACTACTACACAGTGGCCCTGCGGAACATGAACGAGAGATACGGCTACAAGACAGAG

GACCTGGCCTTCATCCAGGACAAGCGGTTCAAGTTCGTGAAGGACGCCTACGGCAAGAAAAAGTCCCAAGTGAATACC
GGATTCTTCCTGAGCCTGCAGGACTACAACGGCGACACACAGAAGAAGCTGCACCTGAGCGGAGTGGGAATCGCCCTG
CTGATCTGCCTGTTCCTGGACAAGCAGTACATCAACATCTTTCTGAGCAGGCTGCCCATCTTCTCCAGCTACAATGCC
CAGAGCGAGGAACGGCGGATCATCATCAGATCCTTCGGCATCAACAGCATCAAGCTGCCCAAGGACCGGATCCACAGC
GAGAAGTCCAACAAGAGCGTGGCCATGGATATGCTCAACGAAGTGAAGCGGTGCCCCGACGAGCTGTTCACAACACTG
TCTGCCGAGAAGCAGTCCCGGTTCAGAATCATCAGCGACGACCACAATGAAGTGCTGATGAAGCGGAGCAGCGACAGA
TTCGTGCCTCTGCTGCTGCAGTATATCGATTACGGCAAGCTGTTCGACCACATCAGGTTCCACGTGAACATGGGCAAG
CTGAGATACCTGCTGAAGGCCGACAAGACCTGCATCGACGGCCAGACCAGAGTCAGAGTGATCGAGCAGCCCCTGAAC
GGCTTCGGCAGACTGGAAGAGGCCGAGACAATGCGGAAGCAAGAGAACGGCACCTTCGGCAACAGCGGCATCCGGATC
AGAGACTTCGAGAACATGAAGCGGGACGACGCCAATCCTGCCAACTATCCCTACATCGTGGACACCTACACACACTAC
ATCCTGGAAAACAACAAGGTCGAGATGTTTATCAACGACAAAGAGGACAGCGCCCCACTGCTGCCCGTGATCGAGGAT
GATAGATACGTGGTCAAGACAATCCCCAGCTGCCGGATGAGCACCCTGGAAATTCCAGCCATGGCCTTCCACATGTTT
CTGTTCGGCAGCAAGAAAACCGAGAAGCTGATCGTGGACGTGCACAACCGGTACAAGAGACTGTTCCAGGCCATGCAG
AAAGAAGAAGTGACCGCCGAGAATATCGCCAGCTTCGGAATCGCCGAGAGCGACCTGCCTCAGAAGATCCTGGATCTG
ATCAGCGGCAATGCCCACGGCAAGGATGTGGACGCCTTCATCAGACTGACCGTGGACGACATGCTGACCGACACCGAG
CGGAGAATCAAGAGATTCAAGGACGACCGGAAGTCCATTCGGAGCGCCGACAACAAGATGGGAAAGAGAGGCTTCAAG
CAGATCTCCACAGGCAAGCTGGCCGACTTCCTGGCCAAGGACATCGTGCTGTTTCAGCCCAGCGTGAACGATGGCGAG
AACAAGATCACCGGCCTGAACTACCGGATCATGCAGAGCGCCATTGCCGTGTACGATAGCGGCGACGATTACGAGGCC
AAGCAGCAGTTCAAGCTGATGTTCGAGAAGGCCCGGCTGATCGGCAAGGGCACAACAGAGCCTCATCCATTTCTGTAC
AAGGTGTTCGCCCGCAGCATCCCCGCCAATGCCGTCGAGTTCTACGAGCGCTACCTGATCGAGCGGAAGTTCTACCTG
ACCGGCCTGTCCAACGAGATCAAGAAAGGCAACAGAGTGGATGTGCCCTTCATCCGGCGGGACCAGAACAAGTGGAAA
ACACCCGCCATGAAGACCCTGGGCAGAATCTACAGCGAGGATCTGCCCGTGGAACTGCCCAGACAGATGTTCGACAAT
GAGATCAAGTCCCACCTGAAGTCCCTGCCACAGATGGAAGGCATCGACTTCAACAATGCCAACGTGACCTATCTGATC
GCCGAGTACATGAAGAGAGTGCTGGACGACGACTTCCAGACCTTCTACCAGTGGAACCGCAACTACCGGTACATGGAC
ATGCTTAAGGGCGAGTACGACAGAAAGGGCTCCCTGCAGCACTGCTTCACCAGCGTGGAAGAGAGAGAAGGCCTCTGG
AAAGAGCGGGCCTCCAGAACAGAGCGGTACAGAAAGCAGGCCAGCAACAAGATCCGCAGCAACCGGCAGATGAGAAAC
GCCAGCAGCGAAGAGATCGAGACAATCCTGGATAAGCGGCTGAGCAACAGCCGGAACGAGTACCAGAAAAGCGAGAAA
GTGATCCGGCGCTACAGAGTGCAGGATGCCCTGCTGTTTCTGCTGGCCAAAAAGACCCTGACCGAACTGGCCGATTTC
GACGGCGAGAGGTTCAAACTGAAAGAAATCATGCCCGACGCCGAGAAGGGAATCCTGAGCGAGATCATGCCCATGAGC
TTCACCTTCGAGAAAGGCGGCAAGAAGTACACCATCACCAGCGAGGGCATGAAGCTGAAGAACTACGGCGACTTCTTT
GTGCTGGCTAGCGACAAGAGGATCGGCAACCTGCTGGAACTCGTGGGCAGCGACATCGTGTCCAAAGAGGATATCATG
GAAGAGTTCAACAAATACGACCAGTGCAGGCCCGAGATCAGCTCCATCGTGTTCAACCTGGAAAAGTGGGCCTTCGAC
ACATACCCCGAGCTGTCTGCCAGAGTGGACCGGGAAGAGAAGGTGGACTTCAAGAGCATCCTGAAAATCCTGCTGAAC
AACAAGAACATCAACAAAGAGCAGAGCGACATCCTGCGGAAGATCCGGAACGCCTTCGATGCAAACAATTACCCCGAC
AAAGGCGTGGTGGAAATCAAGGCCCTGCCTGAGATCGCCATGAGCATCAAGAAGGCCTTTGGGGAGTACGCCATCATG
AAGGGAAGCCTGCAGCTGCCTCCACTTGAAAGACTGACACTGGGATCCGGCAAGCCCATCCCTAACCCACTGCTGGGC
CTGGACAGCACCCTCGAgAACCCCAGTGCCCCCAGCTACCCGATGGCCTCGCTCTACGTGGGGGACCTCCACCCCGAC
GTGACCGAGGCGATGCTCTACGAGAAGTTCAGCCCGGCCGGGCCCATCCTCTCCATCCGGGTCTGCAGGGACATGATC
ACCCGCCGCTCCTTGGGCTACGCGTATGTGAACTTCCAGCAGCCGGCGGACGCGGAGCGTGCTTTGGACACCATGAAT
TTTGATGTTATAAAGGGCAAGCCAGTACGCATCATGTGGTCTCAGCGTGATCCATCACTTCGCAAAAGTGGAGTAGGC
AACATATTCATTAAAAATCTGGACAAATCCATTGATAATAAAGCACTGTATGATACATTTTCTGCTTTTGGTAACATC
CTTTCATGTAAGGTGGTTTGTGATGAAAATGGTTCCAAGGGCTATGGATTTGTACACTTTGAGACGCAGGAAGCAGCT
GAAAGAGCTATTGAAAAAATGAATGGAATGCTCCTAAATGATCGCAAAGTATTTGTTGGACGATTTAAGTCTCGTAAA
GAACGAGAAGCTGAACTTGGAGCTAGGGCAAAAGAATTCACCAATGTTTACATCAAGAATTTTGGAGAAGACATGGAT
GATGAGCGCCTTAAGGATCTCTTTGGCAAGTTTGGGCCTGCCTTAAGTGTGAAAGTAATGACTGATGAAAGTGGAAAA
TCCAAAGGATTTGGATTTGTAAGCTTTGAAAGGCATGAAGATGCACAGAAAGCTGTGGATGAGATGAACGGAAAGGAG
CTCAATGGAAAACAAATTTATGTTGGTCGAGCTCAGAAAAAGGTGGAACGGCAGACGGAACTTAAGCGCAAATTTGAA
CAGATGAAACAAGATAGGATCACCAGATACCAGGGTGTTAATCTTTATGTGAAAAATCTTGATGATGGTATTGATGAT
GAACGTCTCCGGAAAGAGTTTTCTCCATTTGGTACAATCACTAGTGCAAAGGTTATGATGGAAGGTGGTCGCAGCAAA
GGGTTTGGTTTTGTATGTTTCTCCTCCCCAGAAGAAGCCACTAAAGCAGTTACAGAAATGAACGGTAGAATTGTGGCC
ACAAAGCCATTGTATGTAGCTTTAGCTCAGCGCAAAGAAGAGCGCCAGGCTCACCTCACTAACCAGTATATGCAGAGA
ATGGCAAGTGTACGAGCTGTTCCCAACCCTGTAATCAACCCCTACCAGCCAGCACCTCCTTCAGGTTACTTCATGGCA
GCTATCCCACAGACTCAGAACCGTGCTGCATACTATCCTCCTAGCCAAATTGCTCAACTAAGACCAAGTCCTCGCTGG
ACTGCTCAGGGTGCCAGACCTCATCCATTCCAAAATATGCCCGGTGCTATCCGCCCAGCTGCTCCTAGACCACCATTT
AGTACTATGAGACCAGCTTCTTCACAGGTTCCACGAGTCATGTCAACACAGCGTGTTGCTAACACATCAACACAGACA
ATGGGTCCACGTCCTGCAGCTGCAGCCGCTGCAGCTACTCCTGCTGTCCGCACCGTTCCACAGTATAAATATGCTGCA
GGAGTTCGCAATCCTCAGCAACATCTTAATGCACAGCCACAAGTTACAATGCAACAGCCTGCTGTTCATGTACAAGGT
CAGGAACCTTTGACTGCTTCCATGTTGGCATCTGCCCCTCCTCAAGAGCAAAAGCAAATGTTGGGTGAACGGCTGTTT
CCTCTTATTCAAGCCATGCACCCTACTCTTGCTGGTAAAATCACTGGCATGTTGTTGGAGATTGATAATTCAGAACTT
CTTCACATGCTGGAGTCTCCAGAGTCACTCCGTTCTAAGGTTGATGAAGCTGTAGCTGTACTACAAGCCCACCAAGCT
AAAGAGGCTGCCCAGAAAGCAGTTAACAGTGCCACCGGTGTTCCAACTGTTTAAGCGGCCGCTCGAGCCTAGAGGGCC
CGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTC
CTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTG

TCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCCCTGTAGCGG
CGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTT
CGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTT
CCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTG
ATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTT

SEQ ID NO: 51 Nuclear export signal

LPPLERLTL

SEQ ID NO: 52 >pJC1212 (2963 bp)

AGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG
GTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCG
TGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGG
AGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGA
GCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCT
GGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTT
TGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCC
AATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGC
GGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGC
TCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACGCAGGAAACAGCTATGACCATGATTACGCCAGAGGGCC
TATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAA
ACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTT
TTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGAC
GAAACACCgcactatagaatagggccctctagGTTGTGGAAGGTCCAGTTTTGAGGGGCTATTACAACTTTTTTGGTA
CCGAGCTCGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTT
GCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGC
CTGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACT
CTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGG
GCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCG
TCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTT
TCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT
ATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACAT
TTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTA
AAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGT
TTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGAC
GCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAG
CATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTA
CTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGAT
CGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACG
TTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAA
GTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGG
TCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAG
GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAA
GTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGAT
AATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCT
TCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTG
CCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTA
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACC

SEQ ID NO: 53 >pJC1213 (2962 bp)

AGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCG
GTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCG
TGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGG
AGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGA
GCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCT
GGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTT
TGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCC
AATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGC
GGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGC

TCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACGCAGGAAACAGCTATGACCATGATTACGCCAGAGGGCC
TATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAA
ACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTT
TTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGAC
GAAACACCgagctctagcatttaggtgacacGTTGTGGAAGGTCCAGTTTTGAGGGGCTATTACAACTTTTTTGGTAC
CGAGCTCGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTG
CAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCC
TGAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATATGGTGCACTC
TCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGG
CTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGT
CATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTT
CTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATA
TGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATT
TCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA
AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTT
TTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACG
CCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGC
ATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTAC
TTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATC
GTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGT
TGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAG
TTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGT
CTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGG
CAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAG
TTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATA
ATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT
CTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGC
CGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAG
TGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACC

SEQ ID NO: 54>addgene-plasmid-103854-sequence-254736 (2962 bp)

agtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcg
gtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcg
tgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacagg
agagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttga
gcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctttttacggttcct
ggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctt
tgagtgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgccc
aatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagc
gggcagtgagcgcaacgcaattaatgtgagttagctcactcattaggcaccccaggctttacactttatgcttccggc
tcgtatgttgtgtggaattgtgagcggataacaatttcacgcaggaaacagctatgaccatgattacgccagagggcc
tatttcccatgattccttcatatttgcatatacgatacaaggctgttagagagataattggaattaatttgactgtaa
acacaaagatattagtacaaaatacgtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgtt
ttaaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaaggac
gaaacaccatgtcttcctgggacgaagacaagttgtggaaggtccagttttgaggggctattacaacttttttggtac
cgagctcgaattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaacttaatcgccttg
cagcacatcccccttcgccagctggcgtaatagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcc
tgaatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatatggtgcactc
tcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgccctgacggg
cttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggttttcaccgt
catcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgataataatggttt
cttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgtttatttttctaaatacattcaaata
tgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaacatt
tccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaa
aagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagtt
ttcgccccgaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacg
ccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcaccagtcacagaaaagc
atcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttac
ttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatc
gttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatggcaacaacgt
tgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggaggcggataaag

ttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatctggagccggtgagcgtgggt
ctcgcggtatcattgcagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcagg
caactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaactgtcagaccaag
tttactcatatatactttagattgatttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttgata
atctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatctt
cttgagatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctag
tgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttacc

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A fusion protein comprising:

an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and

a regulon moiety that stimulates and/or stabilizes the mRNA(s).

2. The fusion protein of claim 1, wherein the RNA effector protein is a Cas effector protein selected from the group consisting of Cas9, Cas12, Cas13, and Cas14.

3. The fusion protein of claim 2, wherein the RNA effector protein is Cas13b.

4. The fusion protein of claim 2 or claim 3, wherein the Cas effector protein is a catalytically inactive Cas protein.

5. The fusion protein of any one of the preceding claims, wherein the regulon moiety is PABPC1 or NAT10.

6. The fusion protein of any one of the preceding claims, further comprising a linker and/or a spacer.

7. The fusion protein of any one of the preceding claims, further comprising a nuclear export signal and/or an epitope tag.

8. The fusion protein of any one of the preceding claims, wherein the RNA effector protein is N terminal to the regulon moiety.

9. The fusion protein of any one of the preceding claims, wherein the RNA effector protein is C terminal to the regulon moiety.

10. The fusion protein of claim 1 comprising or consisting of SEQ ID NO: 48 or SEQ ID NO: 49 or a polypeptide having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to SEQ ID NO: 48 or SEQ ID NO: 49.

11. A polynucleotide encoding the fusion protein of any one of the preceding claims.

12. A vector comprising the polynucleotide of claim 11.

13. A cell comprising the vector of claim 12.

14. A system comprising:

a fusion protein comprising:

an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and

a regulon moiety that stimulates and/or stabilizes the mRNA(s); and

a gRNA that forms a complex with the RNA effector protein and comprises a complementarity region that hybridizes with the mRNA(s) of the active allele.

15. The system of claim 13, wherein the RNA effector protein is a Cas effector protein selected from the group consisting of Cas9, Cas12, Cas13, and Cas14.

16. The system of claim 14, wherein the Cas effector protein is Cas13b.

17. The system of claim 14 or claim 15, wherein the Cas effector protein is a catalytically inactive Cas effector protein.

18. The system of any one of the preceding claims, wherein the regulon moiety is PABPC1 or NAT10.

19. The system of any one of the preceding claims, wherein the fusion protein further comprises a linker and/or a spacer.

20. The system of any one of the preceding claims, wherein the fusion protein further comprises a nuclear export signal and/or an epitope tag.

21. The system of any one of the preceding claims, wherein the RNA effector protein is N terminal to the regulon moiety.

22. The system of any one of the preceding claims, wherein the RNA effector protein is C terminal to the regulon moiety.

23. The system of any one of the preceding claims, wherein the fusion protein comprises or consists of SEQ ID NO: 48 or SEQ ID NO: 49 or a polypeptide having at least

80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to SEQ ID NO: 48 or SEQ ID NO: 49.

24. The system of any one of the preceding claims, wherein the gRNA targets an mRNA encoding MeCP2, SCN1A, SYNGAP1, SHANK3, CHD2, or PTEN.

25. The system of any one of the preceding claims, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16 , SEQ ID NO: 17, SEQ ID NO: 18, and combinations thereof.

26. The system of any one of the preceding claims, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 17, and combinations thereof.

27. The system of any one of the preceding claims, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding an amino acid selected from the group consisting of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, and combinations thereof.

28. The system of any one of the preceding claims, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, and combinations thereof.

29. The system of any one of the preceding claims, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding an amino acid selected from the group consisting of SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, and combinations thereof.

30. The system of any one of the preceding claims, wherein the gRNA is selected from the group consisting of SEQ ID NO: 55, SEQ ID NO: 56, SEQ ID NO: 57, SEQ ID NO: 58,

SEQ ID NO: 59, SEQ ID NO: 60, SEQ ID NO: 61, SEQ ID NO: 62, SEQ ID NO: 63, SEQ ID NO: 64, SEQ ID NO: 65, SEQ ID NO: 66, SEQ ID NO: 67, SEQ ID NO: 68, SEQ ID NO: 69, SEQ ID NO: 70, SEQ ID NO: 71, and SEQ ID NO: 72.

31. One or more polynucleotide(s) encoding the system of any one of the preceding claims.

32. One or more vector(s) comprising the polynucleotide(s) of claim 31.

33. A cell comprising the vector(s) of claim 32.

34. A complex comprising:

a fusion protein comprising:

an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and

a regulon moiety that stimulates and/or stabilizes the mRNA(s),

bound to a gRNA comprising a complementarity region that hybridizes with the mRNA(s) of the active allele.

35. The complex of claim 34, wherein the RNA effector protein is dCas13b and the regulon moiety is PABP1 or NAT10.

36. The complex of claim 34 or 35, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16 , SEQ ID NO: 17, SEQ ID NO: 18, and combinations thereof.

37. The complex of claim 34 or 35, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 17, and combinations thereof.

38. The complex of claim 34 or 35, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding an amino acid selected from the group

consisting of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, and combinations thereof.

39. The complex of claim 34 or 35, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA selected from the group consisting of SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, and combinations thereof.

40. The complex of claim 34 or 35, wherein the gRNA comprises a complementarity region designed to hybridize to an mRNA encoding SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, and combinations thereof.

41. A complex comprising:

a fusion protein comprising:

an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and

a regulon moiety that stimulates and/or stabilizes the mRNA(s),

bound to a gRNA and the mRNA.

42. The complex of claim 41, wherein the RNA effector protein is dCas13b and the regulon moiety is PABP1 or NAT10.

43. The complex of claim 41 or 42, wherein the mRNA is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16 , SEQ ID NO: 17, SEQ ID NO: 18, and combinations thereof.

44. The complex of claim 41 or 42, wherein the mRNA encodes an amino acid selected from the group consisting of SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 10, SEQ ID NO: 15, SEQ ID NO: 17, and combinations thereof.

45. The complex of claim 41 or 42, wherein the mRNA is selected from the group consisting of SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, and combinations thereof.

46. The complex of claim 41 or 42, wherein the mRNA is selected from the group consisting of SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 33, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, and combinations thereof.

47. The complex of claim 41 or 42, wherein the mRNA encodes an amino acid selected from the group consisting of is selected from the group consisting of SEQ ID NO: 37, SEQ ID NO: 38, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, and combinations thereof.

48. A pharmaceutical composition comprising:
the fusion protein or system of any one of the preceding claims.

49. The pharmaceutical composition of claim 48, further comprising a pharmaceutically acceptable carrier.

50. A pharmaceutical composition comprising:
one or more nucleic acids encoding the fusion protein or system of any one of the preceding claims.

51. A viral vector comprising one or more nucleic acids encoding the fusion protein or system of any one of the preceding claims.

52. The viral vector of claim 51, wherein the viral vector is an adeno-associated viral vector.

53. A nanoparticle or liposome comprising the fusion protein or system of any one of the preceding claims or one or more nucleic acids encoding the fusion protein or system of any one of the preceding claims.

54. A method of stimulating or stabilizing mRNA(s), the method comprising:
contacting the mRNA(s) with the fusion protein or system of any one of the preceding claims.

55. The method of claim 54, wherein the method is carried out, *in vitro*, *in vivo*, or *ex vivo*.

56. A method of treating or preventing a haploinsufficiency disorder in a subject, the method comprising:

administering to the subject a fusion protein or a nucleic acid encoding a fusion protein comprising:

an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and

a regulon moiety that stimulates and/or stabilizes the mRNA(s); and

a gRNA or a nucleic acid encoding a gRNA designed to form a complex with the RNA effector protein and comprising a complementarity region designed to hybridize with the mRNA of the active allele.

57. The method of claim 56, wherein the RNA effector protein is a Cas effector protein selected from the group consisting of a Cas9, Cas12, Cas13, and Cas14.

58. The method of claim 57, wherein the Cas effector protein is Cas13b.

59. The method of claim 57 or 58, wherein the Cas effector protein is a catalytically inactive Cas protein.

60. The method of any one of the preceding claims, wherein the regulon moiety is PABPC1 or NAT10.

61. The method of any one of the preceding claims, wherein the fusion protein further comprises a linker and/or a spacer.

62. The method of any one of the preceding claims, wherein the fusion protein further comprises a nuclear export signal and/or an epitope tag.

63. The method of any one of the preceding claims, wherein the RNA effector protein is N terminal to the regulon moiety.

64. The method of any one of the preceding claims, wherein the RNA effector protein is C terminal to the regulon moiety.

65. The method of any one of the preceding claims, wherein the fusion protein comprises or consists of SEQ ID NO: 48 or SEQ ID NO: 49 or a polypeptide having at least

80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to SEQ ID NO: 48 or SEQ ID NO: 49.

66. The method of any one of the preceding claims, wherein the haploinsufficiency disorder is selected from the group consisting from 5qsyndrome,Adams-Oliver syndrome 1, Adams-Oliver syndrome 3, Adams-Oliver syndrome 5,Adams-Oliver syndrome 6, Alagille syndrome 1, Autoimmune lymphoproliferative syndrome type IA, Autoimmune lymphoproliferative syndrome type V, Autosomal dominant deafness-2A,Brain malformations with or without urinary tract defects (BRMUTD), Carney complex type 1,CHARGE syndrome, Cleidocranial dysplasia, Currarino syndrome, Denys-Drash syndrome/Frasier syndrome, Developmental delay, intellectual disability, obesity, and dysmorphic features(DIDOD), DiGeorge syndrome (TBXI-associated), Dravet syndrome, Duane-radial raysyndrome, Ehlers-Danlos syndrome (classic-like), Ehlers-Danlos syndrome (vascular type),Feingold syndrome 1, Frontotemporal lobar degeneration with TDP43 inclusions (FTLD-TDP),GRN-related, GLUT I deficiency syndrome, Greig cephalopolysyndactyly syndrome, Hereditary hemorrhagic telangiectasia type 1, Holoprosencephaly 3, Holoprosencephaly 4,Holoprosencephaly 5, Holt-Oram syndrome, Hypoparathyroidism, sensorineural deafness, andrenal disease (HDR), Kleefstra syndrome 1, Klippel-Trenaunay syndrome (AAGF-related), Leri-Weill dyschondrosteosis, Marfan syndrome, Mental retardation and distinctive facial features with or without cardiac defects (MRFACD), Mental retardation, autosomal dominant 1, Mental retardation, autosomal dominant 19, Mental retardation, autosomal dominant 29, Nail-patella syndrome (NPS), Phelan-McDermid syndrome, Pitt-Hopkins syndrome, Primary pulmonary hypertension 1, Rett syndrome (congenital variant), Smith-Magenis syndrome (RAII associated), Sotos syndrome 1, Sotos syndrome 2, Stickler syndrome type I, Supravalvular aorticstenosis, SYNGAPI-related intellectual disability, Treacher Collins syndrome, Trichorhinophalangeal syndrome type I, Ulnar-mammary syndrome, van der Woude syndrome1, Waardenburg syndrome type 1, W aardenburg syndrome type 2A, and Waardenburg syndrometype 4C.

67. The method of any one of the preceding claims, wherein the haploinsufficiency disorder is a CNS haploinsufficiency disorder.

68. The method of claim 67, wherein the CNS haploinsufficiency disorder is selected from the group consisting of episodic ataxia, familial hemiplegia migraine, CDKL5 deficiency disorder, CHD2 myoclonic encephalopathy, familial focal epilepsy with variable loci, FOXG1 syndrome, benign familial neonatal seizures, Rett syndrome, Dravat syndrome, SCN2A-epileptic

encephalopathy, SCN2A-developmental encephalopathy, SCN8A-epileptic encephalopathy, SC8A familial infantile epilepsy, early infantile epileptic encephalopathy, myoclonic-atonic epilepsy, early infantile epileptic encephalopathy, SYNGAP1-related intellectual disability, tuberous sclerosis, Lennox-Gastaut Syndrome, FoxG1 syndrome, KCNQ2-related epileptic encephalopathy, PCDH19-related epilepsy, SLC6A1-related myoclonic-astatic epilepsy, STXBP1-related epileptic encephalopathy, SYNGAP1 syndrome, and combinations thereof.

69. The method of any one of the preceding claims, wherein the haploinsufficiency disorder is Dravet Syndrome or Rett syndrome.

70. The method of any one of the preceding claims, wherein the subject has a haploinsufficiency in a gene selected from the group consisting of *AGGFI, ARHGAP31, BMPR2, CHD7, COL2Al, COL3Al, CTLA4, CTNNBI, DLL4, EHMTI, ELN,ENG, FAS, FBNI, FOXGI, GATA3, GLI3, GRN, IRF6, JAGI, KCNQ4, LMXIB, MBD5,MED13L, MITF, MNXI, MYCN, NFIA, NFIX, NOTCH!, NSDI, PAX3, PHIP, PRKARIA,RAil, RBPJ, RPS14, RUNX2, SALL4, SCNIA, SETBPI, SHANK3, SHH, SHOX,SLC2Al/GLUT1, SOXI0, SYNGAPI, TBXI, TBX3, TBX5, TCF4, TCOFI, TGIFI, TNXB,TRPSI, WTI, ZIC2,* and combinations thereof.

71. The method of any one of the preceding claims, wherein the subject has a haploinsufficiency in a gene selected from the group consisting of *SCN1A, SCN2A, SCN8A, SCN12A5, SPTAN1, CDKL5, CHD2, FOXG1, KCNQ2, PCDH19, SLC6A1, STXBP1, SYNGAP1, CACNA1A, DEPDC5, MECP2, TSC1, TSC2,* and combinations thereof.

72. The method of any one of the preceding claims, wherein the subject has mutation selected from the list in Table 4 and combinations thereof.

73. The method of any one of the preceding claims, wherein the subject has a mutation selected from the list in Table 6 and combinations thereof.

74. The method of any one of the preceding claims, wherein the subject is a mammal.

75. The method of claim 74, wherein the subject is a human.

76. The method of any one of the preceding claims, wherein the fusion protein and gRNA are administered as part of a pharmaceutical composition.

77. The method of any one of the preceding claims, wherein administering comprises administering a viral vector comprising nucleic acid sequence(s) encoding the fusion protein and gRNA to the subject.

78. The method of any one of the preceding claims, wherein administering comprises administering a nanoparticle or liposome comprising the fusion protein and gRNA or nucleic acid sequence(s) encoding the fusion protein and gRNA to the subject.

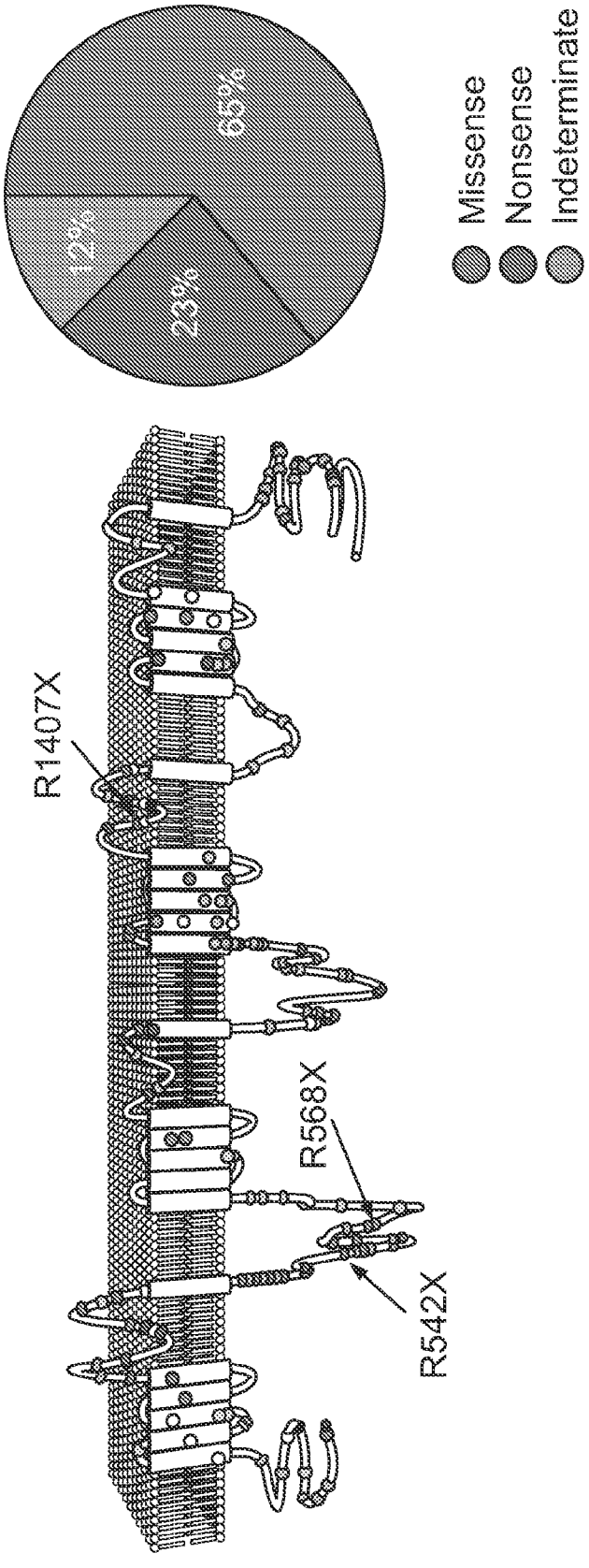

65%
23%
12%
Missense
Nonsense
Indeterminate
R1407X
R568X
R542X
Arg
Cys
Gln
Glu
Gly
Leu
Lys
Ser
Trp
Tyr

FIG. 1

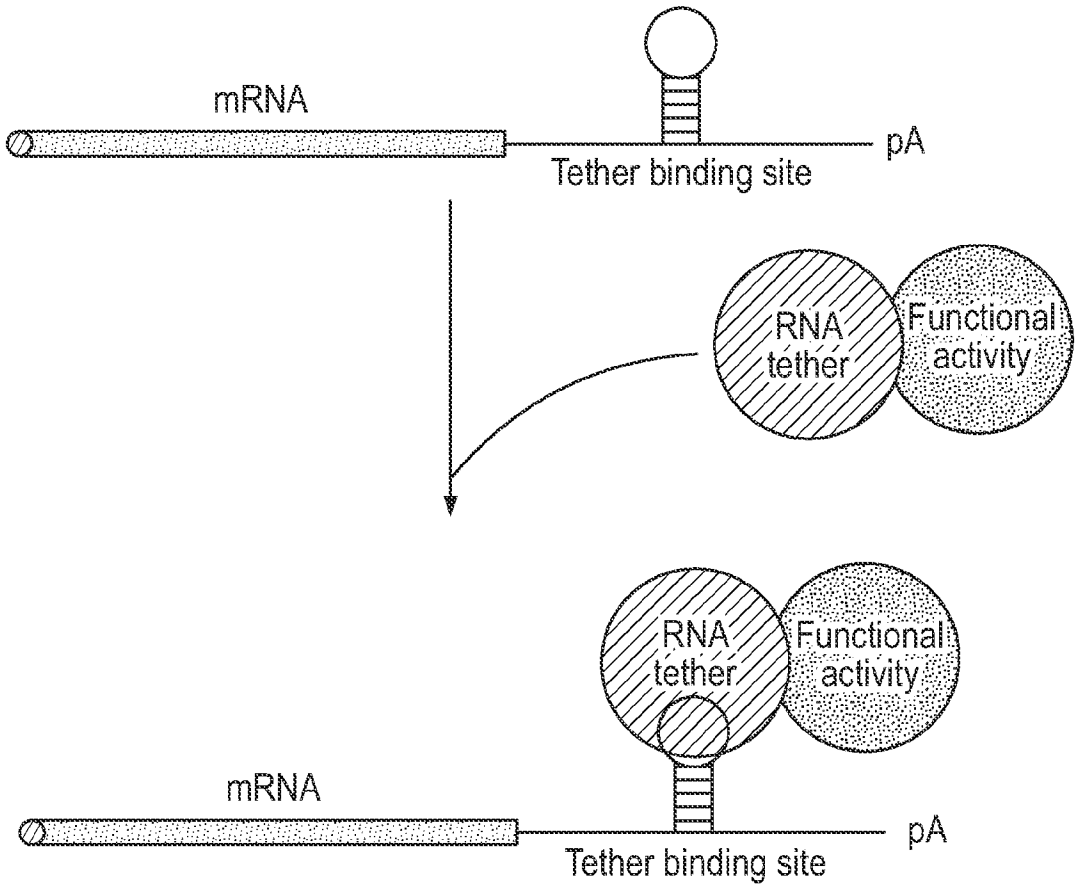

mRNA
pA
Tether binding site
RNA tether
Functional activity
RNA tether
Functional activity
mRNA
pA
Tether binding site

FIG. 2

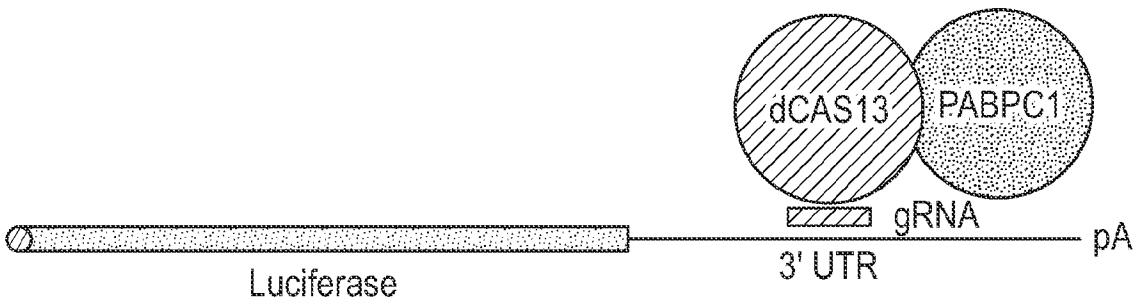

dCAS13
PABPC1
gRNA
pA
3' UTR
Luciferase

FIG. 3A

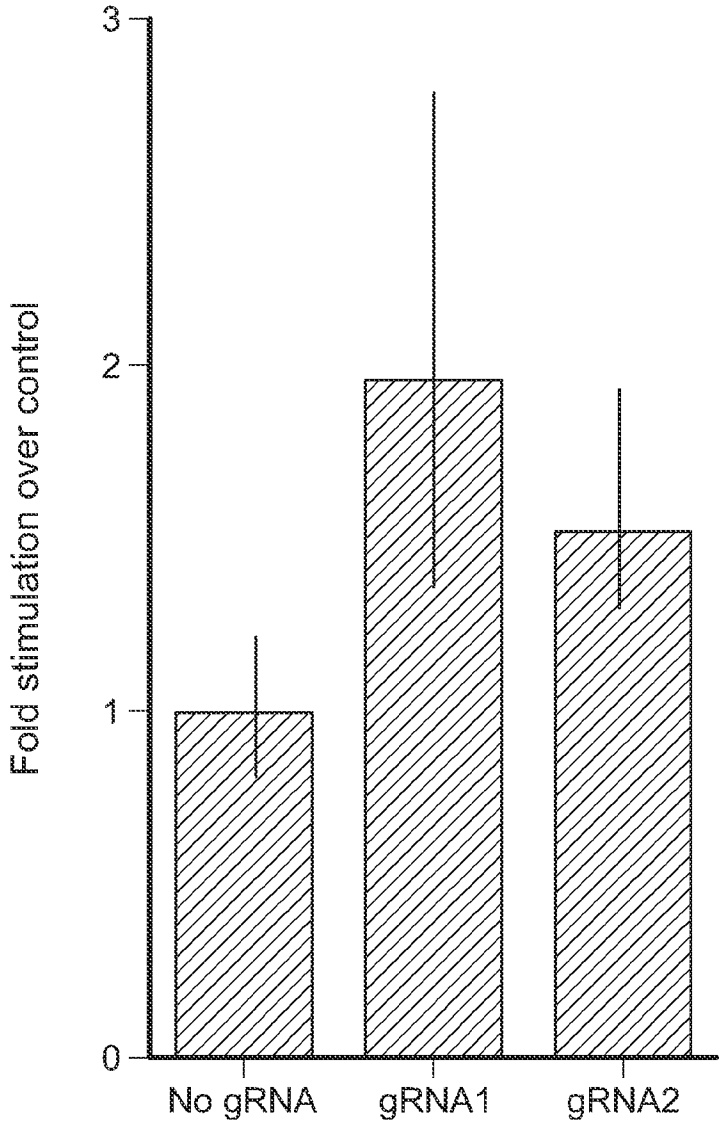


FIG. 3B

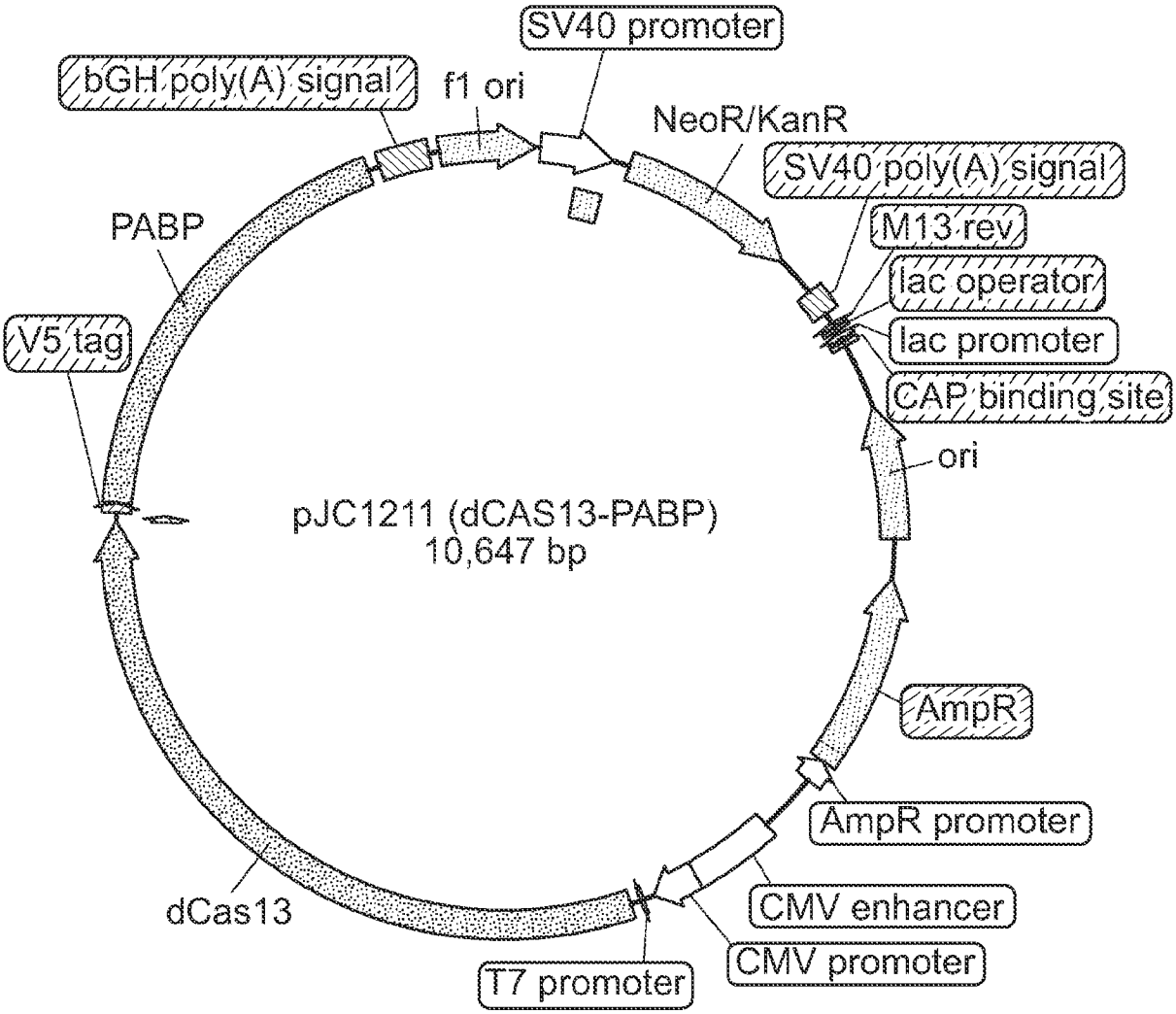

SV40 promoter
bGH poly(A) signal
f1 ori
NeoR/KanR
SV40 poly(A) signal
M13 rev
PABP
lac operator
V5 tag
lac promoter
CAP binding site
ori
pJC1211 (dCAS13-PABP)
10,647 bp
AmpR
AmpR promoter
CMV enhancer
dCas13
CMV promoter
T7 promoter

FIG. 4

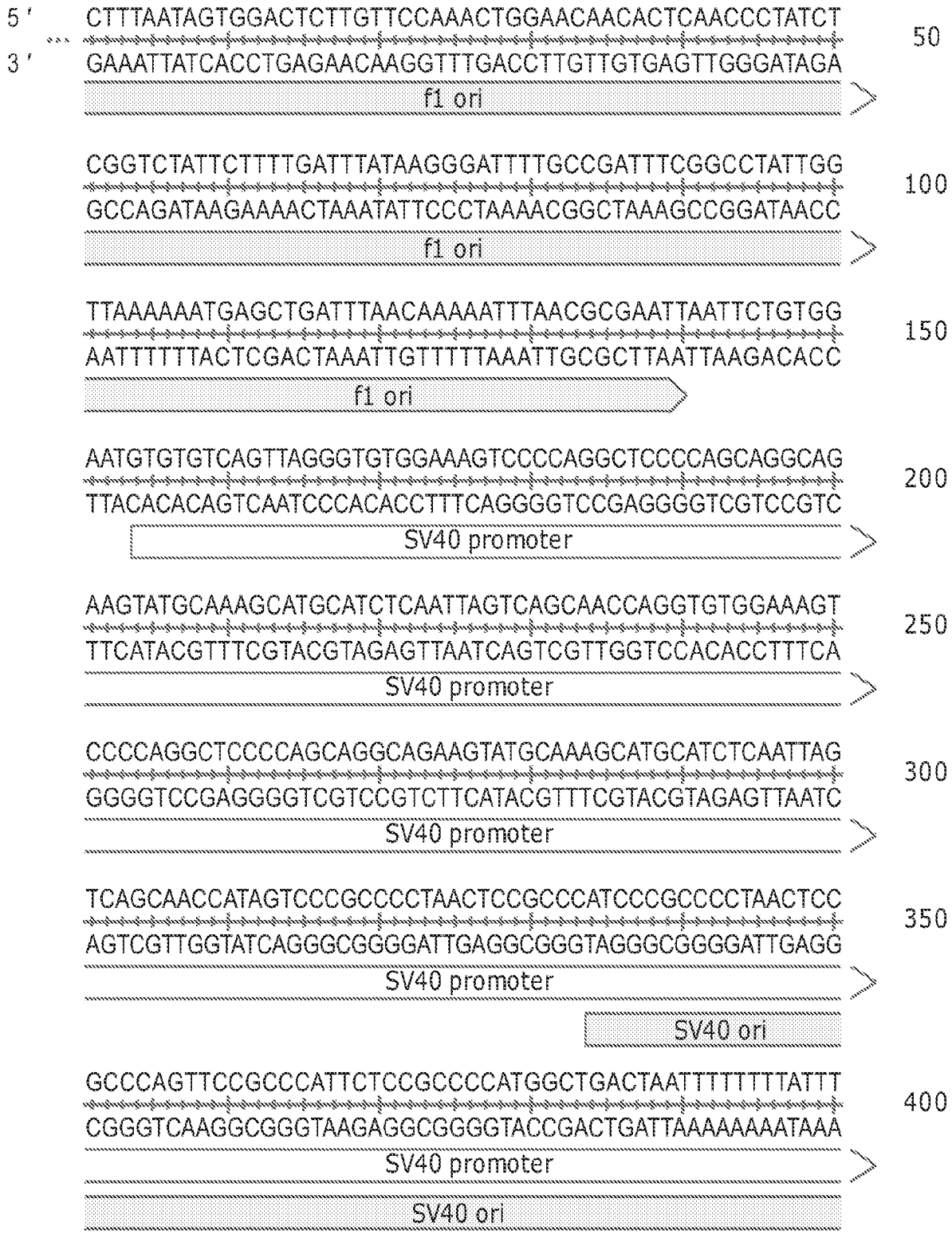

5′
3′
CTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCT
GAAATTATCACCTGAGAACAAGGTTTGACCTTGTTGTGAGTTGGGATAGA
50
f1 ori
CGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGG
GCCAGATAAGAAAACTAAATATTCCCTAAAACGGCTAAAGCCGGATAACC
100
f1 ori
TTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTGG
AATTTTTTACTCGACTAAATTGTTTTTAAATTGCGCTTAATTAAGACACC
150
f1 ori
AATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAG
TTACACACAGTCAATCCCACACCTTTCAGGGGTCCGAGGGGTCGTCCGTC
200
SV40 promoter
AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGT
TTCATACGTTTCGTACGTAGAGTTAATCAGTCGTTGGTCCACACCTTTCA
250
SV40 promoter
CCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAG
GGGGTCCGAGGGGTCGTCCGTCTTCATACGTTTCGTACGTAGAGTTAATC
300
SV40 promoter
TCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCC
AGTCGTTGGTATCAGGGCGGGGATTGAGGCGGGTAGGGCGGGGATTGAGG
350
SV40 promoter
SV40 ori
GCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTT
CGGGTCAAGGCGGGTAAGAGGCGGGGTACCGACTGATTAAAAAAAATAAA
400
SV40 promoter
SV40 ori

FIG. 5A

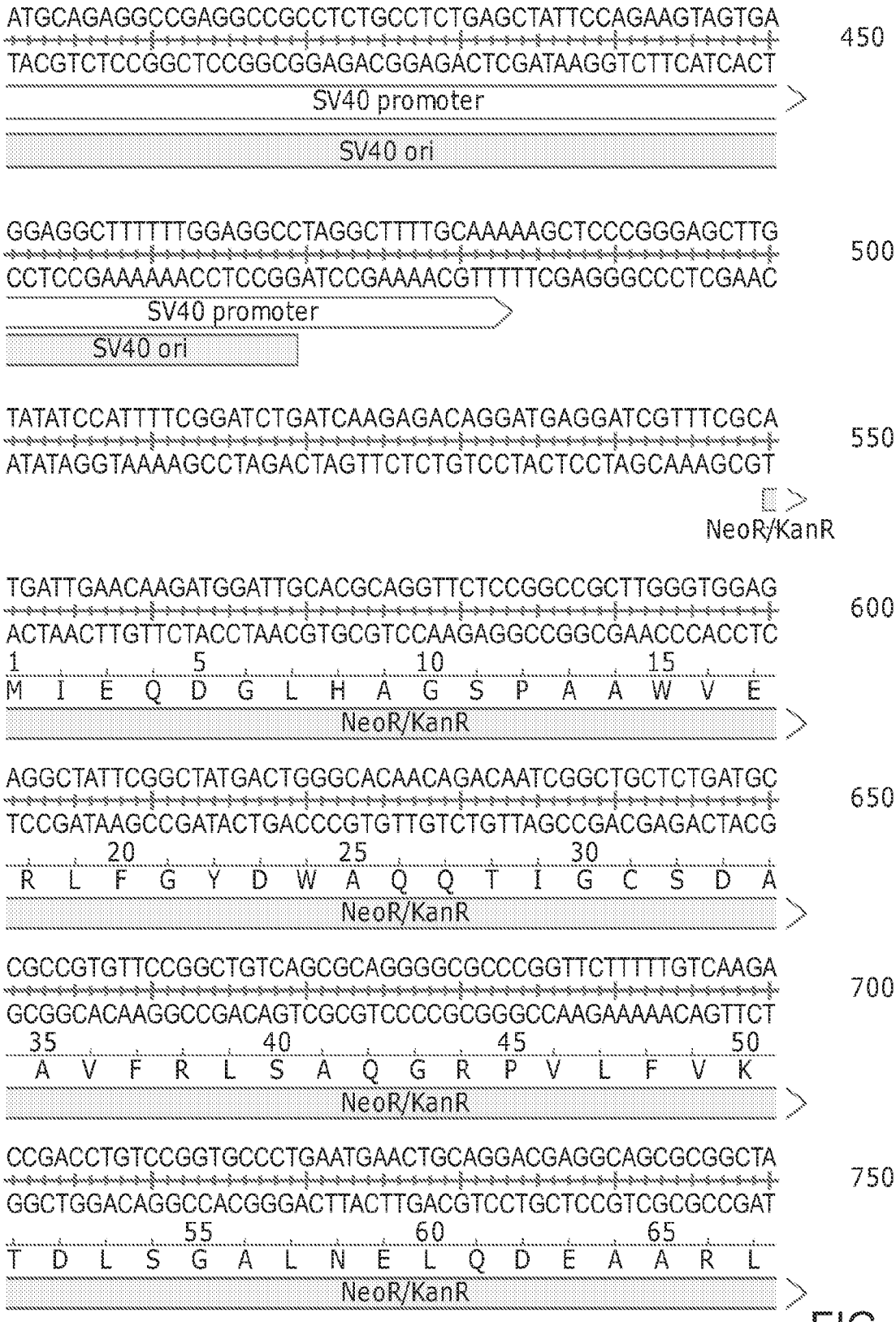

ATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGA
450
TACGTCTCCGGCTCCGGCGGAGACGGAGACTCGATAAGGTCTTCATCACT
SV40 promoter
SV40 ori
GGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTCCCGGGAGCTTG
500
CCTCCGAAAAAACCTCCGGATCCGAAAACGTTTTTCGAGGGCCCTCGAAC
SV40 promoter
SV40 ori
TATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCA
550
ATATAGGTAAAAGCCTAGACTAGTTCTCTGTCCTACTCCTAGCAAAGCGT
NeoR/KanR
TGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAG
600
ACTAACTTGTTCTACCTAACGTGCGTCCAAGAGGCCGGCGAACCCACCTC
1 5 10 15
M I E Q D G L H A G S P A A W V E
NeoR/KanR
AGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGC
650
TCCGATAAGCCGATACTGACCCGTGTTGTCTGTTAGCCGACGAGACTACG
20 25 30
R L F G Y D W A Q Q T I G C S D A
NeoR/KanR
CGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGA
700
GCGGCACAAGGCCGACAGTCGCGTCCCCGCGGGCCAAGAAAAACAGTTCT
35 40 45 50
A V F R L S A Q G R P V L F V K
NeoR/KanR
CCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTA
750
GGCTGGACAGGCCACGGGACTTACTTGACGTCCTGCTCCGTCGCGCCGAT
55 60 65
T D L S G A L N E L Q D E A A R L
NeoR/KanR

FIG. 5B

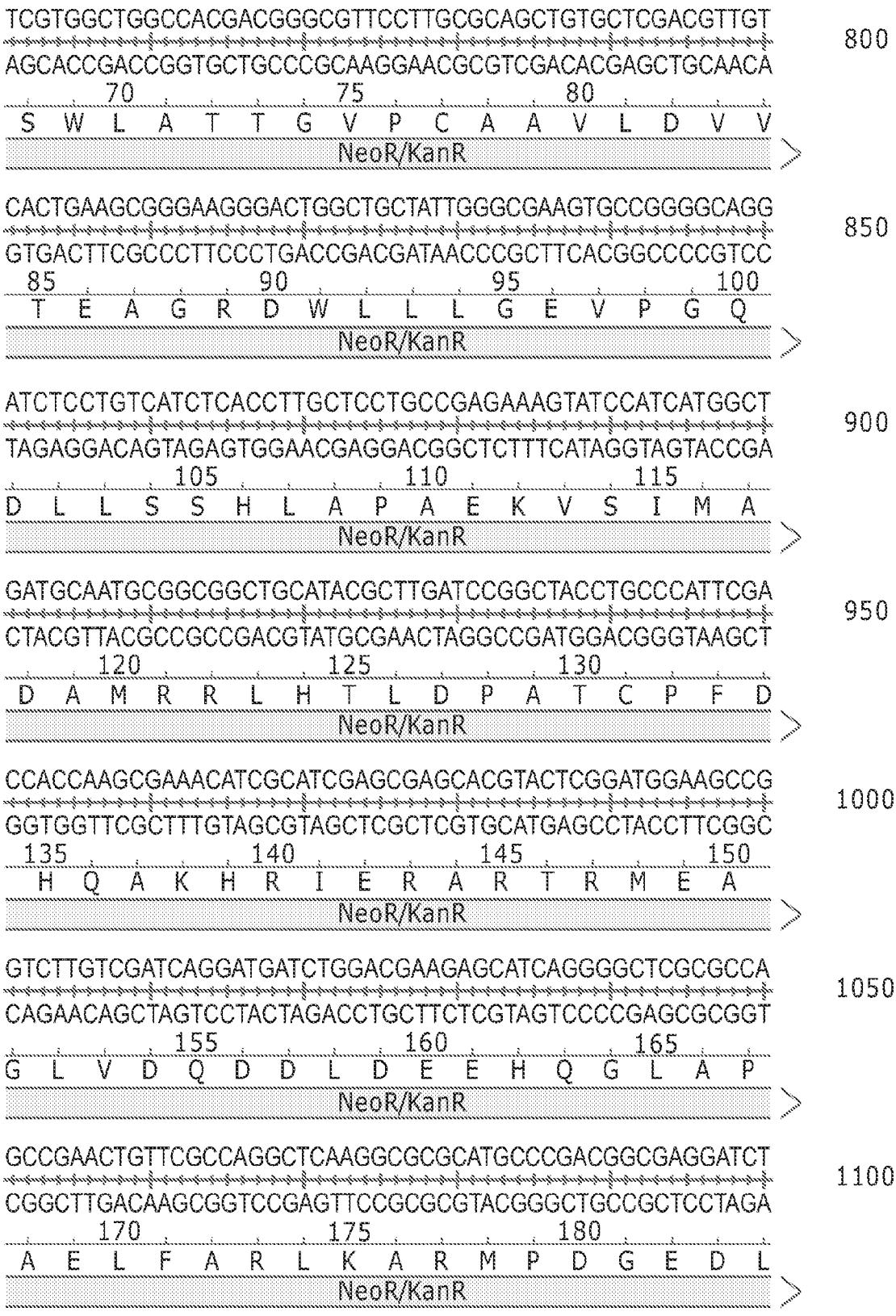

TCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGT
800
AGCACCGACCGGTGCTGCCCGCAAGGAACGCGTCGACACGAGCTGCAACA
70 75 80
S W L A T T G V P C A A V L D V V
NeoR/KanR
CACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGG
850
GTGACTTCGCCCTTCCCTGACCGACGATAACCCGCTTCACGGCCCCGTCC
85 90 95 100
T E A G R D W L L L G E V P G Q
NeoR/KanR
ATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCT
900
TAGAGGACAGTAGAGTGGAACGAGGACGGCTCTTTCATAGGTAGTACCGA
105 110 115
D L L S S H L A P A E K V S I M A
NeoR/KanR
GATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGA
950
CTACGTTACGCCGCCGACGTATGCGAACTAGGCCGATGGACGGGTAAGCT
120 125 130
D A M R R L H T L D P A T C P F D
NeoR/KanR
CCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCG
1000
GGTGGTTCGCTTTGTAGCGTAGCTCGCTCGTGCATGAGCCTACCTTCGGC
135 140 145 150
H Q A K H R I E R A R T R M E A
NeoR/KanR
GTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCA
1050
CAGAACAGCTAGTCCTACTAGACCTGCTTCTCGTAGTCCCCGAGCGCGGT
155 160 165
G L V D Q D D L D E E H Q G L A P
NeoR/KanR
GCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCT
1100
CGGCTTGACAAGCGGTCCGAGTTCCGCGCGTACGGGCTGCCGCTCCTAGA
170 175 180
A E L F A R L K A R M P D G E D L
NeoR/KanR

FIG. 5C

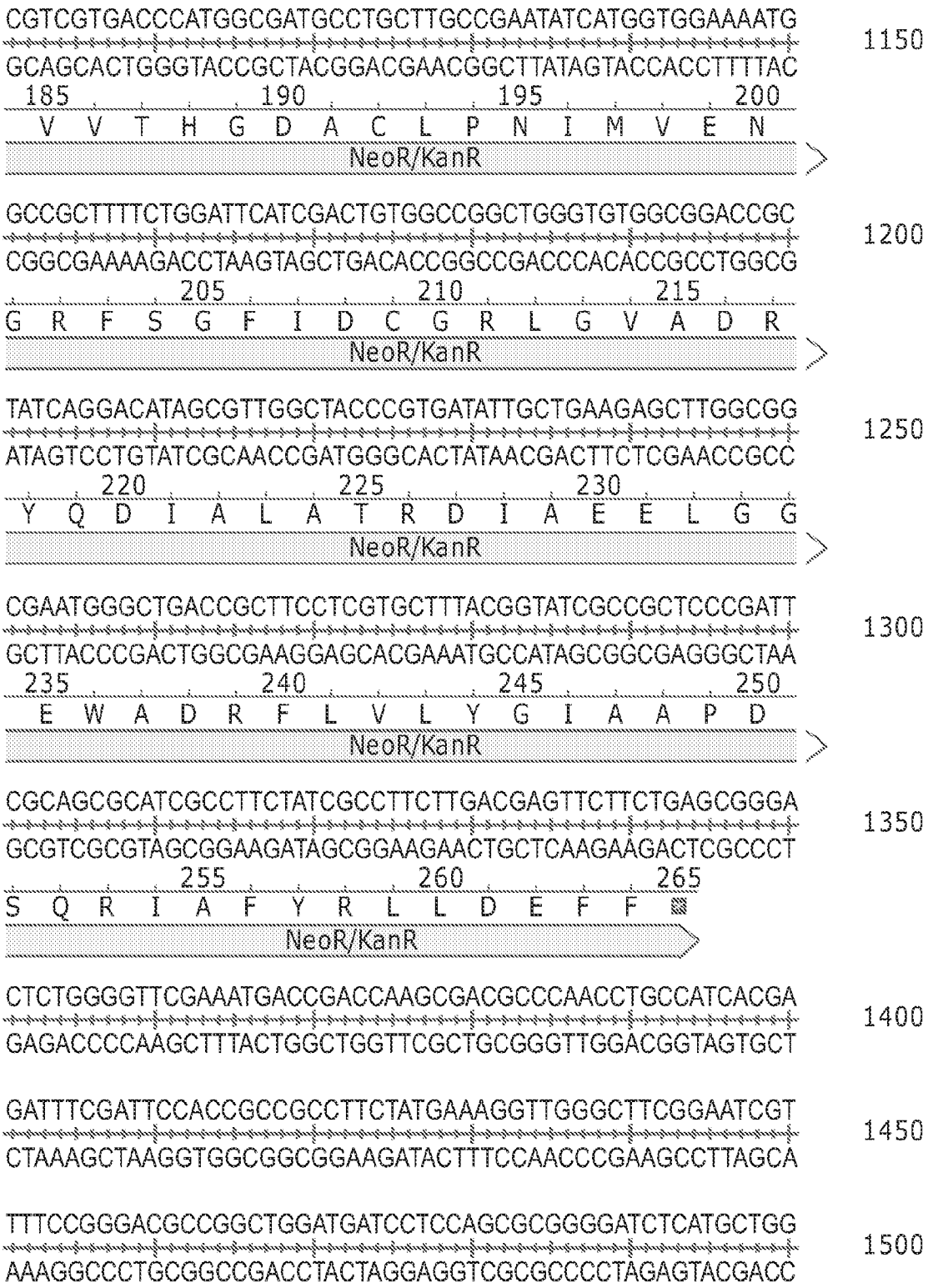

CGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATG
1150
GCAGCACTGGGTACCGCTACGGACGAACGGCTTATAGTACCACCTTTTAC
185 190 195 200
V V T H G D A C L P N I M V E N
NeoR/KanR
GCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGC
1200
CGGCGAAAAGACCTAAGTAGCTGACACCGGCCGACCCACACCGCCTGGCG
205 210 215
G R F S G F I D C G R L G V A D R
NeoR/KanR
TATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGG
1250
ATAGTCCTGTATCGCAACCGATGGGCACTATAACGACTTCTCGAACCGCC
220 225 230
Y Q D I A L A T R D I A E E L G G
NeoR/KanR
CGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATT
1300
GCTTACCCGACTGGCGAAGGAGCACGAAATGCCATAGCGGCGAGGGCTAA
235 240 245 250
E W A D R F L V L Y G I A A P D
NeoR/KanR
CGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGA
1350
GCGTCGCGTAGCGGAAGATAGCGGAAGAACTGCTCAAGAAGACTCGCCCT
255 260 265
S Q R I A F Y R L L D E F F ■
NeoR/KanR
CTCTGGGGTTCGAAATGACCGACCAAGCGACGCCCAACCTGCCATCACGA
1400
GAGACCCCAAGCTTTACTGGCTGGTTCGCTGCGGGTTGGACGGTAGTGCT
GATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATCGT
1450
CTAAAGCTAAGGTGGCGGCGGAAGATACTTTCCAACCCGAAGCCTTAGCA
TTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGG
1500
AAAGGCCCTGCGGCCGACCTACTAGGAGGTCGCGCCCCTAGAGTACGACC

FIG. 5D

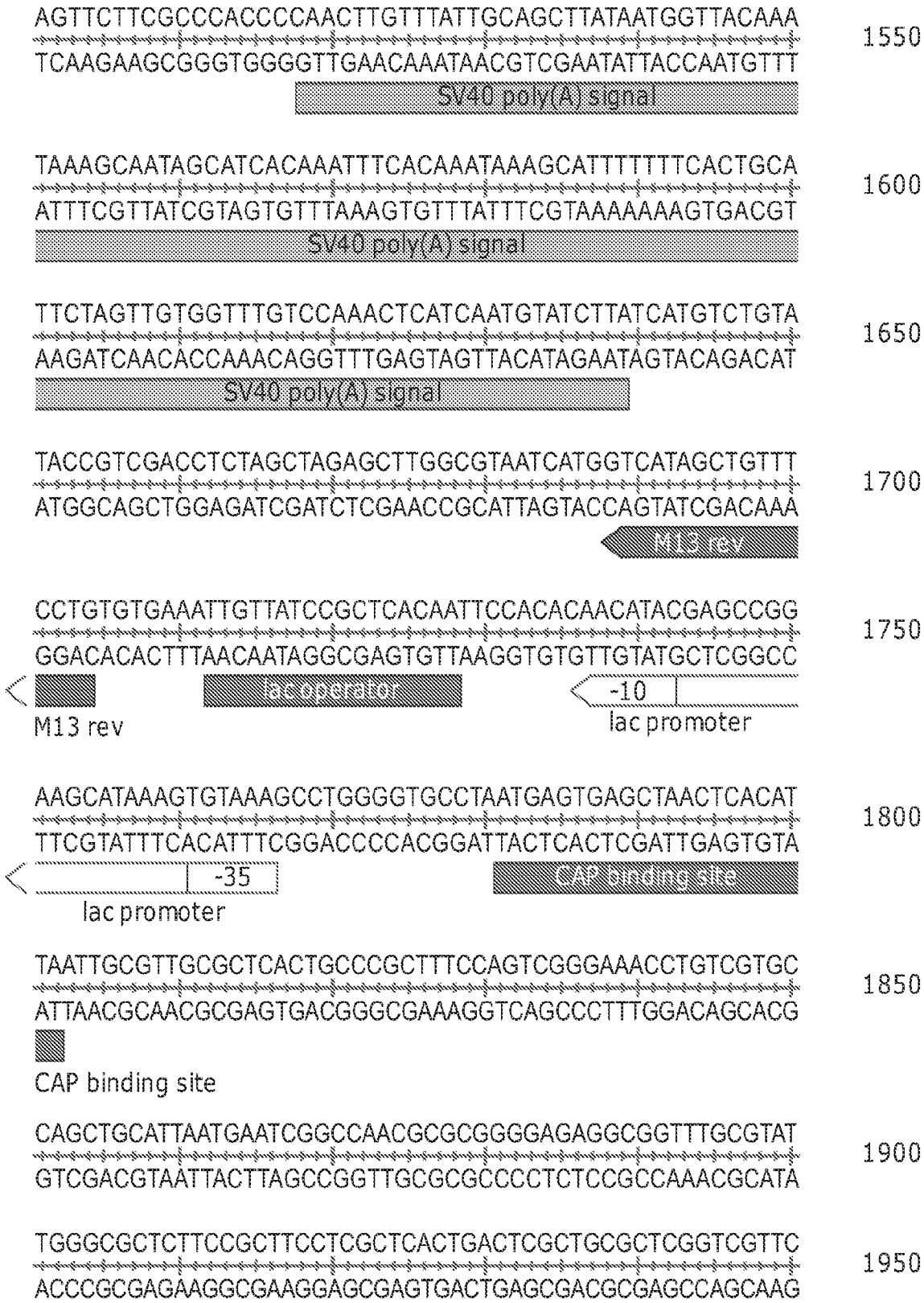

AGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAA
TCAAGAAGCGGGTGGGGGTTGAACAAATAACGTCGAATATTACCAATGTTT
SV40 poly(A) signal
1550
TAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCA
ATTTCGTTATCGTAGTGTTTAAAGTGTTTATTTCGTAAAAAAAGTGACGT
SV40 poly(A) signal
1600
TTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTA
AAGATCAACACCAAACAGGTTTGAGTAGTTACATAGAATAGTACAGACAT
SV40 poly(A) signal
1650
TACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTT
ATGGCAGCTGGAGATCGATCTCGAACCGCATTAGTACCAGTATCGACAAA
M13 rev
1700
CCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGG
GGACACACTTTAACAATAGGCGAGTGTTAAGGTGTGTTGTATGCTCGGCC
lac operator
-10
M13 rev
lac promoter
1750
AAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT
TTCGTATTTCACATTTCGGACCCCACGGATTACTCACTCGATTGAGTGTA
-35
CAP binding site
lac promoter
1800
TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGC
ATTAACGCAACGCGAGTGACGGGCGAAAGGTCAGCCCTTTGGACAGCACG
CAP binding site
1850
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTAT
GTCGACGTAATTACTTAGCCGGTTGCGCGCCCCTCTCCGCCAAACGCATA
1900
TGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTC
ACCCGCGAGAAGGCGAAGGAGCGAGTGACTGAGCGACGCGAGCCAGCAAG
1950

FIG. 5E

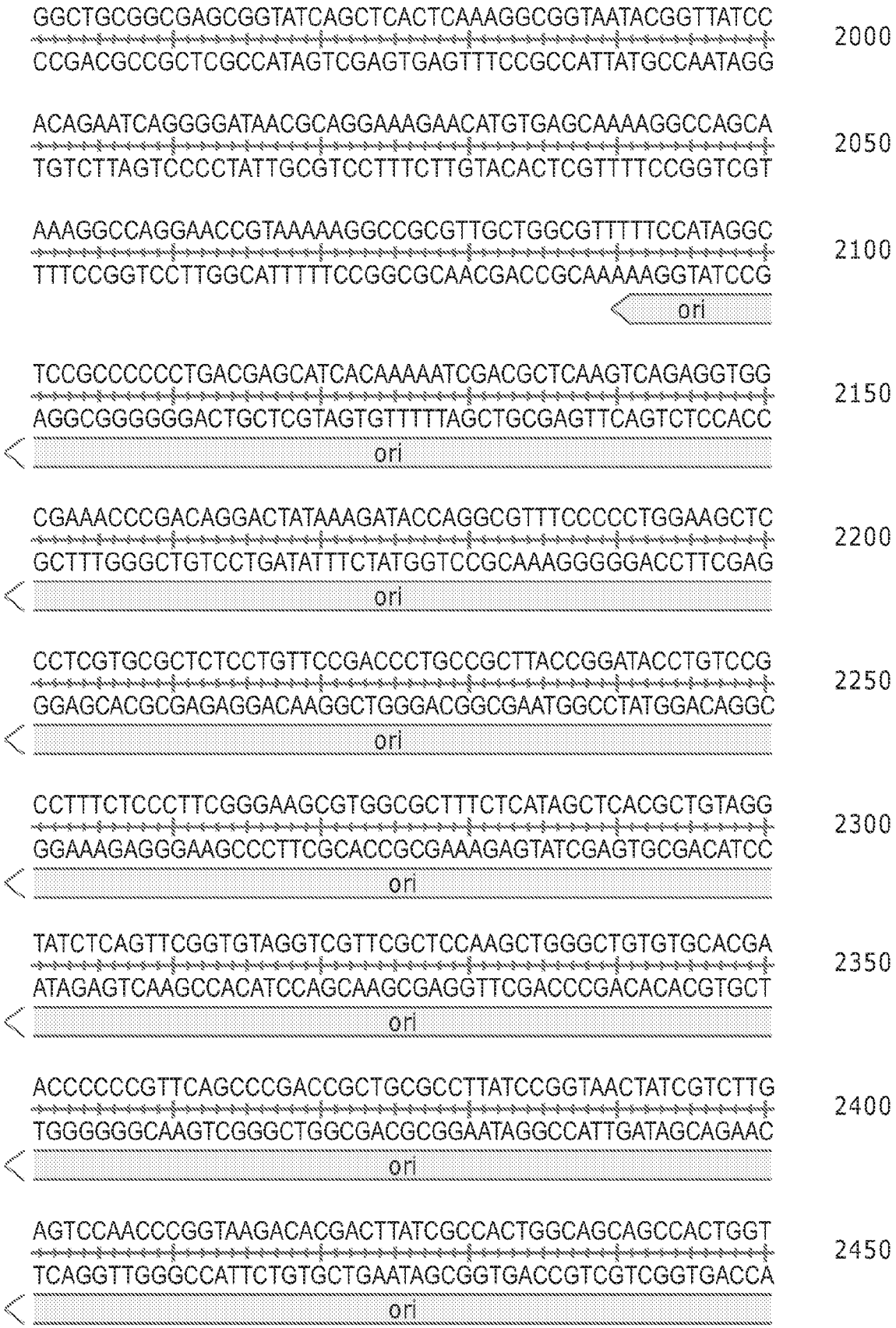

GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC
CCGACGCCGCTCGCCATAGTCGAGTGAGTTTCCGCCATTATGCCAATAGG
2000
ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCA
TGTCTTAGTCCCCTATTGCGTCCTTTCTTGTACACTCGTTTTCCGGTCGT
2050
AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGC
TTTCCGGTCCTTGGCATTTTTCCGGCGCAACGACCGCAAAAAGGTATCCG
2100
ori
TCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG
AGGCGGGGGGACTGCTCGTAGTGTTTTTAGCTGCGAGTTCAGTCTCCACC
2150
ori
CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTC
GCTTTGGGCTGTCCTGATATTTCTATGGTCCGCAAAGGGGGACCTTCGAG
2200
ori
CCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG
GGAGCACGCGAGAGGACAAGGCTGGGACGGCGAATGGCCTATGGACAGGC
2250
ori
CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG
GGAAAGAGGGAAGCCCTTCGCACCGCGAAAGAGTATCGAGTGCGACATCC
2300
ori
TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGA
ATAGAGTCAAGCCACATCCAGCAAGCGAGGTTCGACCCGACACACGTGCT
2350
ori
ACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG
TGGGGGGCAAGTCGGGCTGGCGACGCGGAATAGGCCATTGATAGCAGAAC
2400
ori
AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGT
TCAGGTTGGGCCATTCTGTGCTGAATAGCGGTGACCGTCGTCGGTGACCA
2450
ori

FIG. 5F

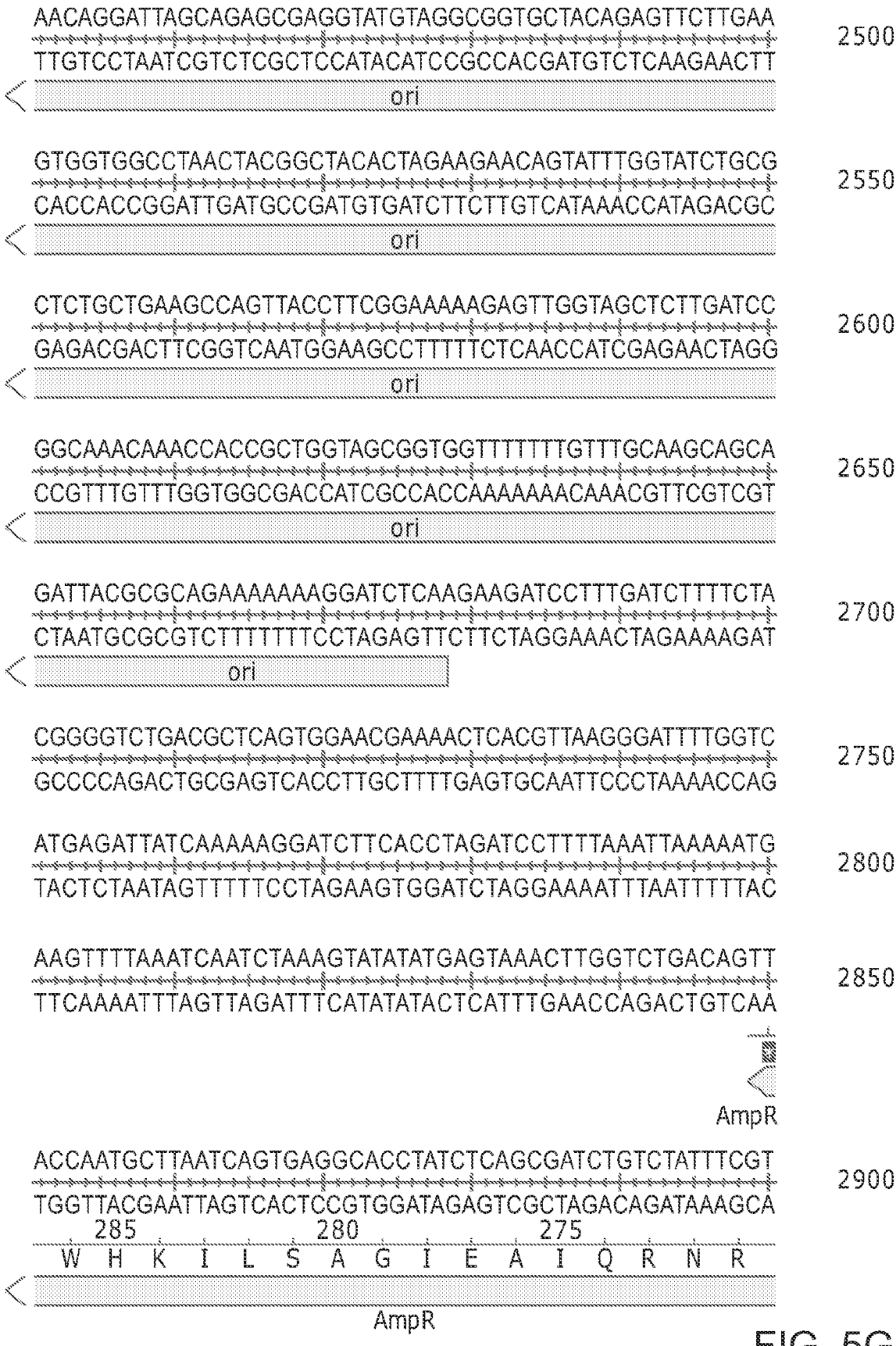

AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
TTGTCCTAATCGTCTCGCTCCATACATCCGCCACGATGTCTCAAGAACTT
2500
ori
GTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCG
CACCACCGGATTGATGCCGATGTGATCTTCTTGTCATAAACCATAGACGC
2550
ori
CTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC
GAGACGACTTCGGTCAATGGAAGCCTTTTTCTCAACCATCGAGAACTAGG
2600
ori
GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA
CCGTTTGTTTGGTGGCGACCATCGCCACCAAAAAAACAAACGTTCGTCGT
2650
ori
GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTA
CTAATGCGCGTCTTTTTTTCCTAGAGTTCTTCTAGGAAACTAGAAAAGAT
2700
ori
CGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTC
GCCCCAGACTGCGAGTCACCTTGCTTTTGAGTGCAATTCCCTAAAACCAG
2750
ATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATG
TACTCTAATAGTTTTTCCTAGAAGTGGATCTAGGAAAATTTAATTTTTAC
2800
AAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTT
TTCAAAATTTAGTTAGATTTCATATATACTCATTTGAACCAGACTGTCAA
2850
AmpR
ACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGT
TGGTTACGAATTAGTCACTCCGTGGATAGAGTCGCTAGACAGATAAAGCA
2900
285 280 275
W H K I L S A G I E A I Q R N R
AmpR

FIG. 5G

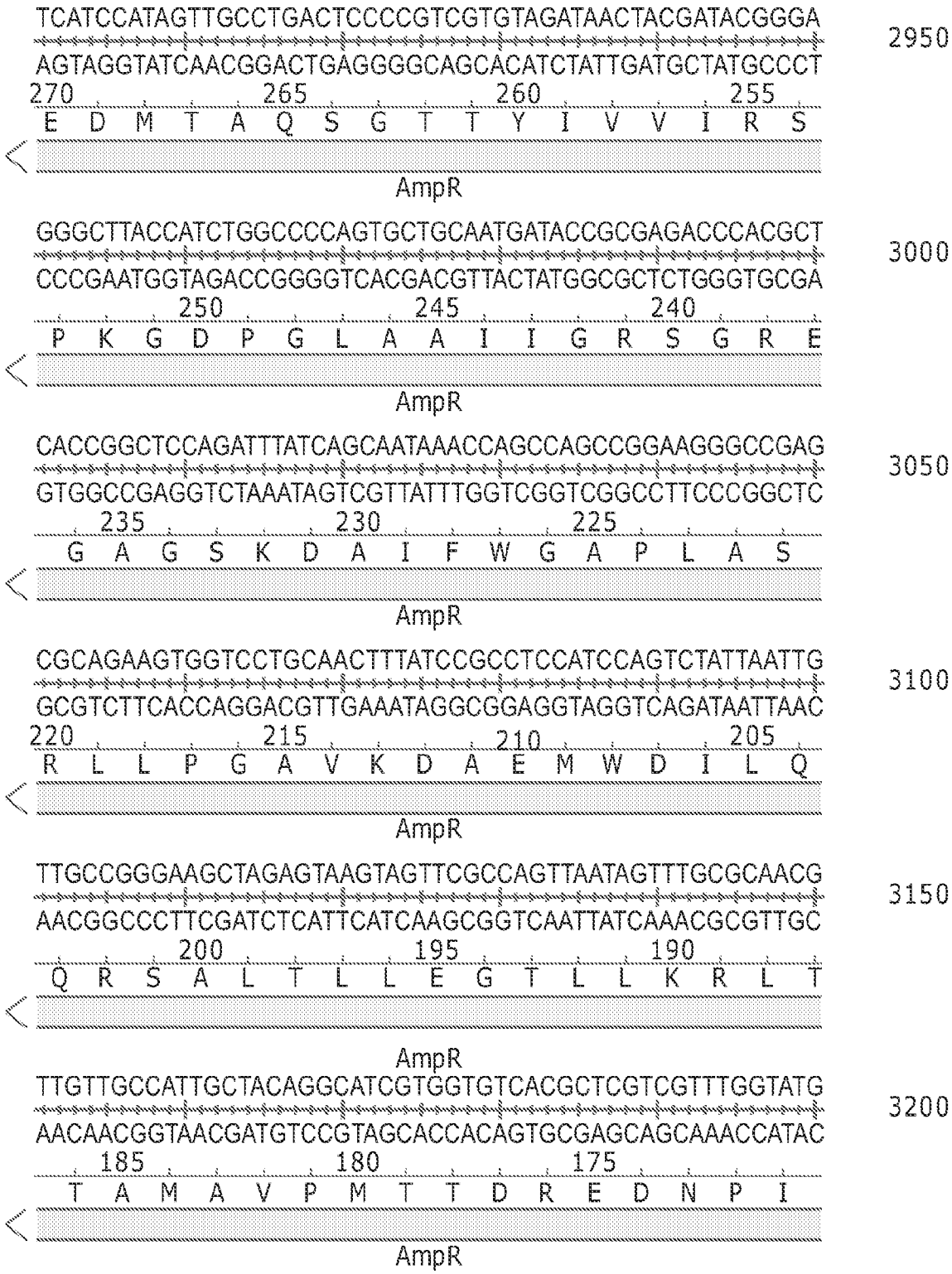

TCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGA
2950
AGTAGGTATCAACGGACTGAGGGGCAGCACATCTATTGATGCTATGCCCT
270 265 260 255
E D M T A Q S G T T Y I V V I R S
AmpR
GGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
3000
CCCGAATGGTAGACCGGGGTCACGACGTTACTATGGCGCTCTGGGTGCGA
250 245 240
P K G D P G L A A I I G R S G R E
AmpR
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAG
3050
GTGGCCGAGGTCTAAATAGTCGTTATTTGGTCGGTCGGCCTTCCCGGCTC
235 230 225
G A G S K D A I F W G A P L A S
AmpR
CGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTG
3100
GCGTCTTCACCAGGACGTTGAAATAGGCGGAGGTAGGTCAGATAATTAAC
220 215 210 205
R L L P G A V K D A E M W D I L Q
AmpR
TTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACG
3150
AACGGCCCTTCGATCTCATTCATCAAGCGGTCAATTATCAAACGCGTTGC
200 195 190
Q R S A L T L L E G T L L K R L T
AmpR
TTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG
3200
AACAACGGTAACGATGTCCGTAGCACCACAGTGCGAGCAGCAAACCATAC
185 180 175
T A M A V P M T T D R E D N P I
AmpR

FIG. 5H

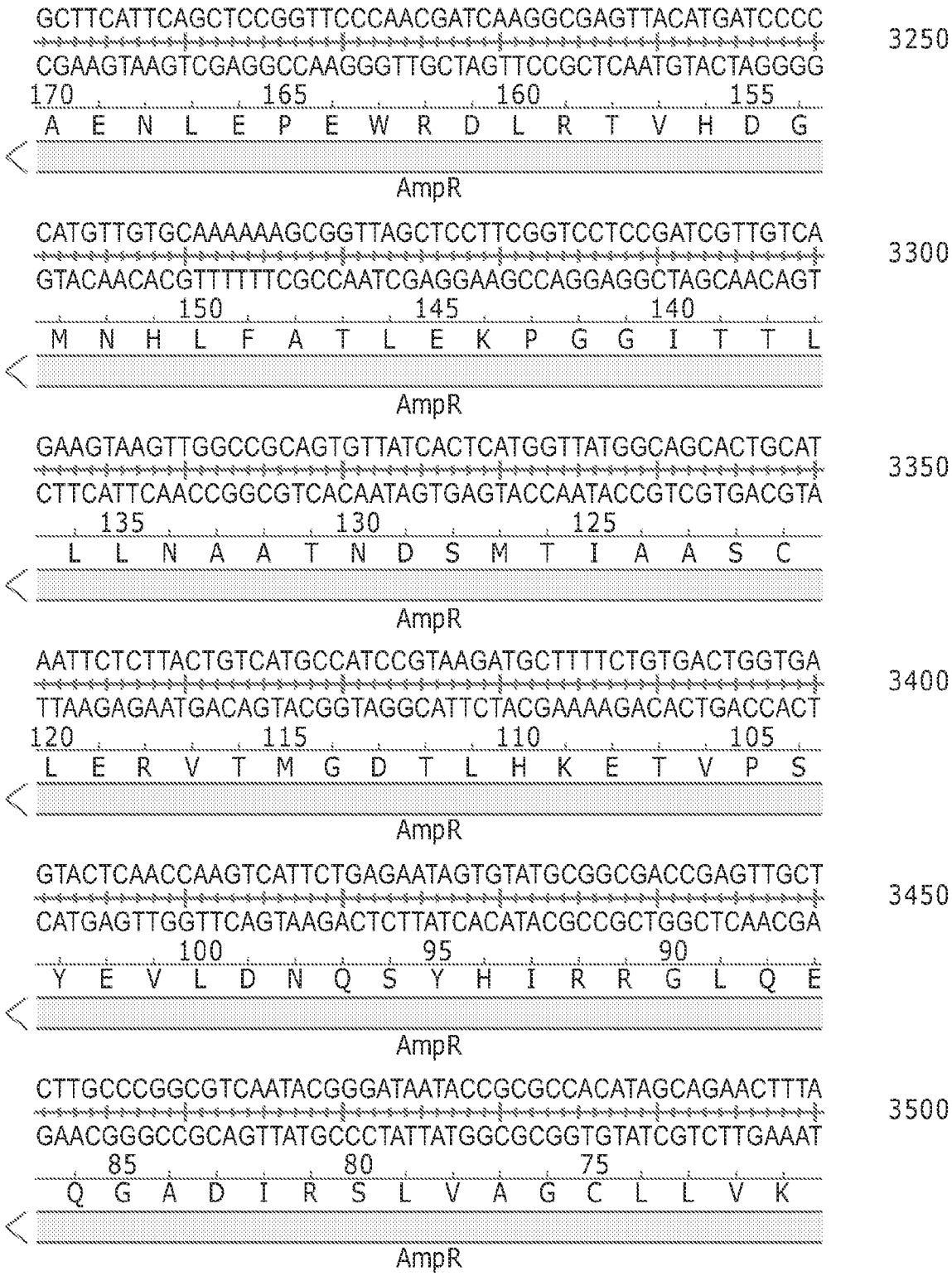

GCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCC
3250
CGAAGTAAGTCGAGGCCAAGGGTTGCTAGTTCCGCTCAATGTACTAGGGG
170 165 160 155
A E N L E P E W R D L R T V H D G
AmpR
CATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCA
3300
GTACAACACGTTTTTTCGCCAATCGAGGAAGCCAGGAGGCTAGCAACAGT
150 145 140
M N H L F A T L E K P G G I T T L
AmpR
GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
3350
CTTCATTCAACCGGCGTCACAATAGTGAGTACCAATACCGTCGTGACGTA
135 130 125
L L N A A T N D S M T I A A S C
AmpR
AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGA
3400
TTAAGAGAATGACAGTACGGTAGGCATTCTACGAAAAGACACTGACCACT
120 115 110 105
L E R V T M G D T L H K E T V P S
AmpR
GTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCT
3450
CATGAGTTGGTTCAGTAAGACTCTTATCACATACGCCGCTGGCTCAACGA
100 95 90
Y E V L D N Q S Y H I R R G L Q E
AmpR
CTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA
3500
GAACGGGCCGCAGTTATGCCCTATTATGGCGCGGTGTATCGTCTTGAAAT
85 80 75
Q G A D I R S L V A G C L L V K
AmpR

FIG. 5I

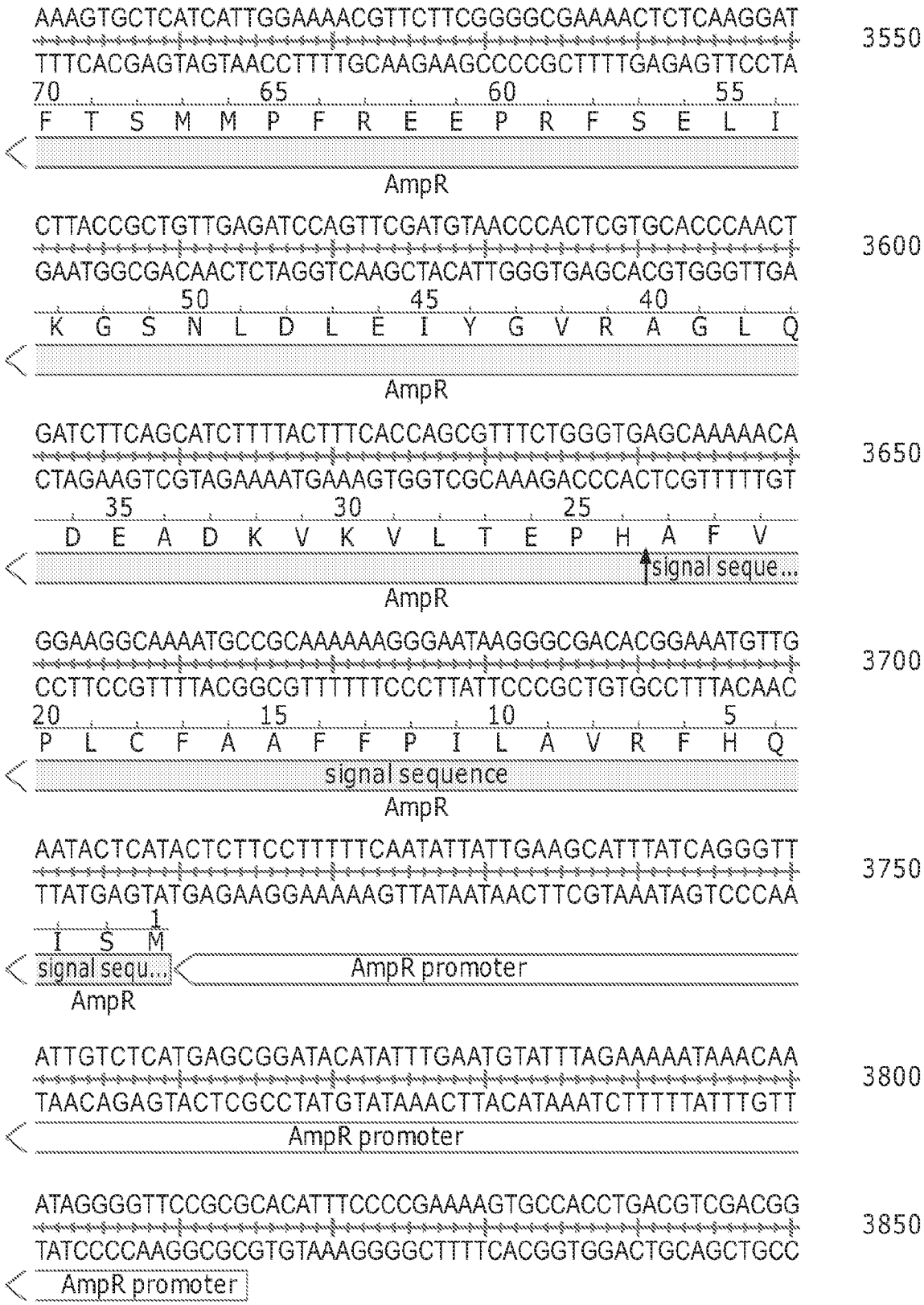

AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT
3550
TTTCACGAGTAGTAACCTTTTGCAAGAAGCCCCGCTTTTGAGAGTTCCTA
70 65 60 55
F T S M M P F R E E P R F S E L I
AmpR
CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACT
3600
GAATGGCGACAACTCTAGGTCAAGCTACATTGGGTGAGCACGTGGGTTGA
50 45 40
K G S N L D L E I Y G V R A G L Q
AmpR
GATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
3650
CTAGAAGTCGTAGAAAATGAAAGTGGTCGCAAAGACCCACTCGTTTTTGT
35 30 25
D E A D K V K V L T E P H A F V
signal seque...
AmpR
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTG
3700
CCTTCCGTTTTACGGCGTTTTTTCCCTTATTCCCGCTGTGCCTTTACAAC
20 15 10 5
P L C F A A F F P I L A V R F H Q
signal sequence
AmpR
AATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTT
3750
TTATGAGTATGAGAAGGAAAAAGTTATAATAACTTCGTAAATAGTCCCAA
1
I S M
signal sequ...
AmpR promoter
AmpR
ATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA
3800
TAACAGAGTACTCGCCTATGTATAAACTTACATAAATCTTTTTATTTGTT
AmpR promoter
ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCGACGG
3850
TATCCCCAAGGCGCGTGTAAAGGGGCTTTTCACGGTGGACTGCAGCTGCC
AmpR promoter

FIG. 5J

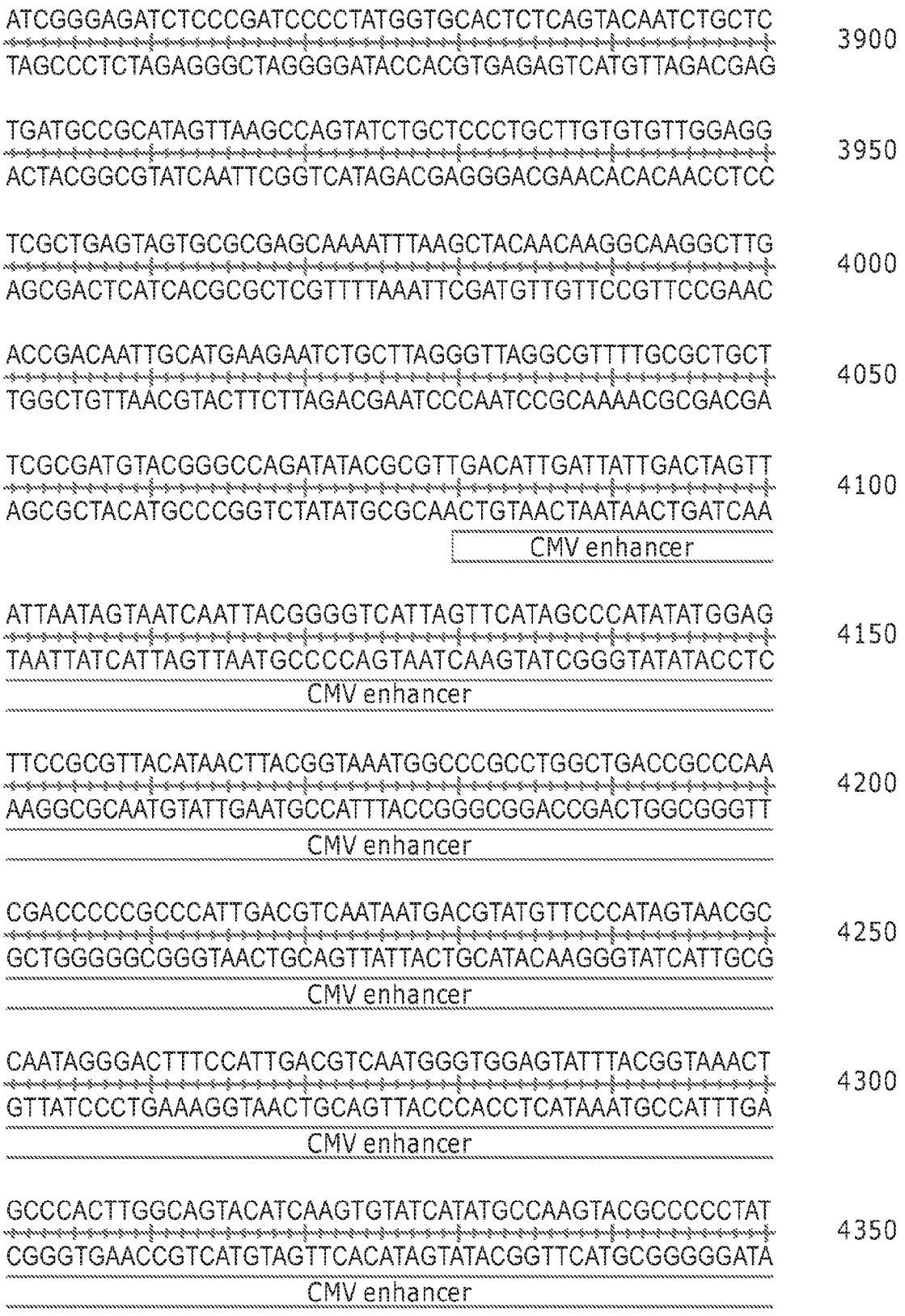

ATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGCTC
TAGCCCTCTAGAGGGCTAGGGGATACCACGTGAGAGTCATGTTAGACGAG
3900
TGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGG
ACTACGGCGTATCAATTCGGTCATAGACGAGGGACGAACACACAACCTCC
3950
TCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTG
AGCGACTCATCACGCGCTCGTTTTAAATTCGATGTTGTTCCGTTCCGAAC
4000
ACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCT
TGGCTGTTAACGTACTTCTTAGACGAATCCCAATCCGCAAAACGCGACGA
4050
TCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTT
AGCGCTACATGCCCGGTCTATATGCGCAACTGTAACTAATAACTGATCAA
CMV enhancer
4100
ATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAG
TAATTATCATTAGTTAATGCCCCAGTAATCAAGTATCGGGTATATACCTC
CMV enhancer
4150
TTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAA
AAGGCGCAATGTATTGAATGCCATTTACCGGGCGGACCGACTGGCGGGTT
CMV enhancer
4200
CGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGC
GCTGGGGGCGGGTAACTGCAGTTATTACTGCATACAAGGGTATCATTGCG
CMV enhancer
4250
CAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT
GTTATCCCTGAAAGGTAACTGCAGTTACCCACCTCATAAATGCCATTTGA
CMV enhancer
4300
GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTAT
CGGGTGAACCGTCATGTAGTTCACATAGTATACGGTTCATGCGGGGGATA
CMV enhancer
4350

FIG. 5K

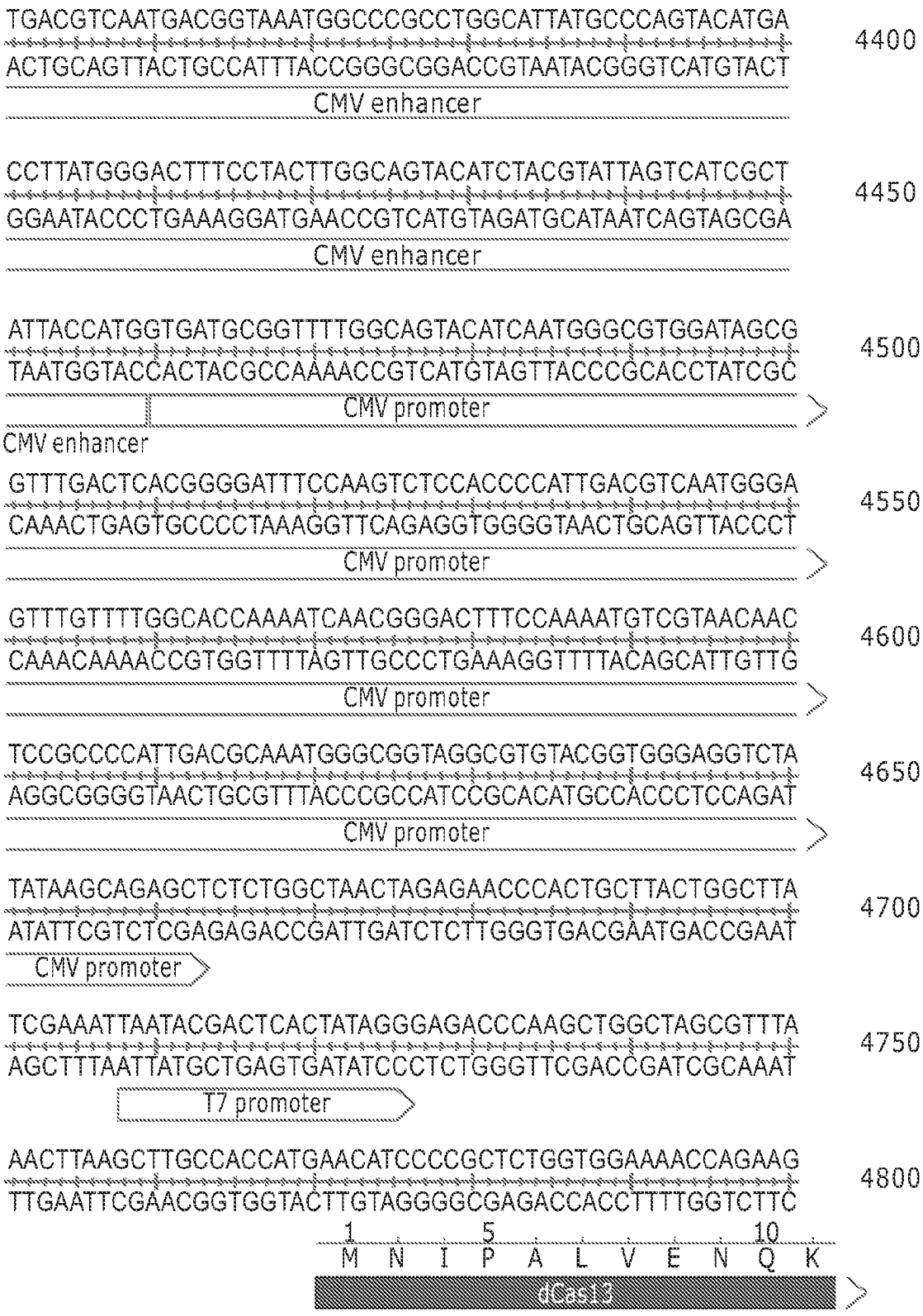

TGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA
ACTGCAGTTACTGCCATTTACCGGGCGGACCGTAATACGGGTCATGTACT
4400
CMV enhancer
CCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCT
GGAATACCCTGAAAGGATGAACCGTCATGTAGATGCATAATCAGTAGCGA
4450
CMV enhancer
ATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCG
TAATGGTACCACTACGCCAAAACCGTCATGTAGTTACCCGCACCTATCGC
4500
CMV promoter
CMV enhancer
GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGA
CAAACTGAGTGCCCCTAAAGGTTCAGAGGTGGGGTAACTGCAGTTACCCT
4550
CMV promoter
GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAAC
CAAACAAAACCGTGGTTTTAGTTGCCCTGAAAGGTTTTACAGCATTGTTG
4600
CMV promoter
TCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA
AGGCGGGGTAACTGCGTTTACCCGCCATCCGCACATGCCACCCTCCAGAT
4650
CMV promoter
TATAAGCAGAGCTCTCTGGCTAACTAGAGAACCCACTGCTTACTGGCTTA
ATATTCGTCTCGAGAGACCGATTGATCTCTTGGGTGACGAATGACCGAAT
4700
CMV promoter
TCGAAATTAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTA
AGCTTTAATTATGCTGAGTGATATCCCTCTGGGTTCGACCGATCGCAAAT
4750
T7 promoter
AACTTAAGCTTGCCACCATGAACATCCCCGCTCTGGTGGAAAACCAGAAG
TTGAATTCGAACGGTGGTACTTGTAGGGGCGAGACCACCTTTTGGTCTTC
4800
1 5 10
M N I P A L V E N Q K
dCas13

FIG. 5L

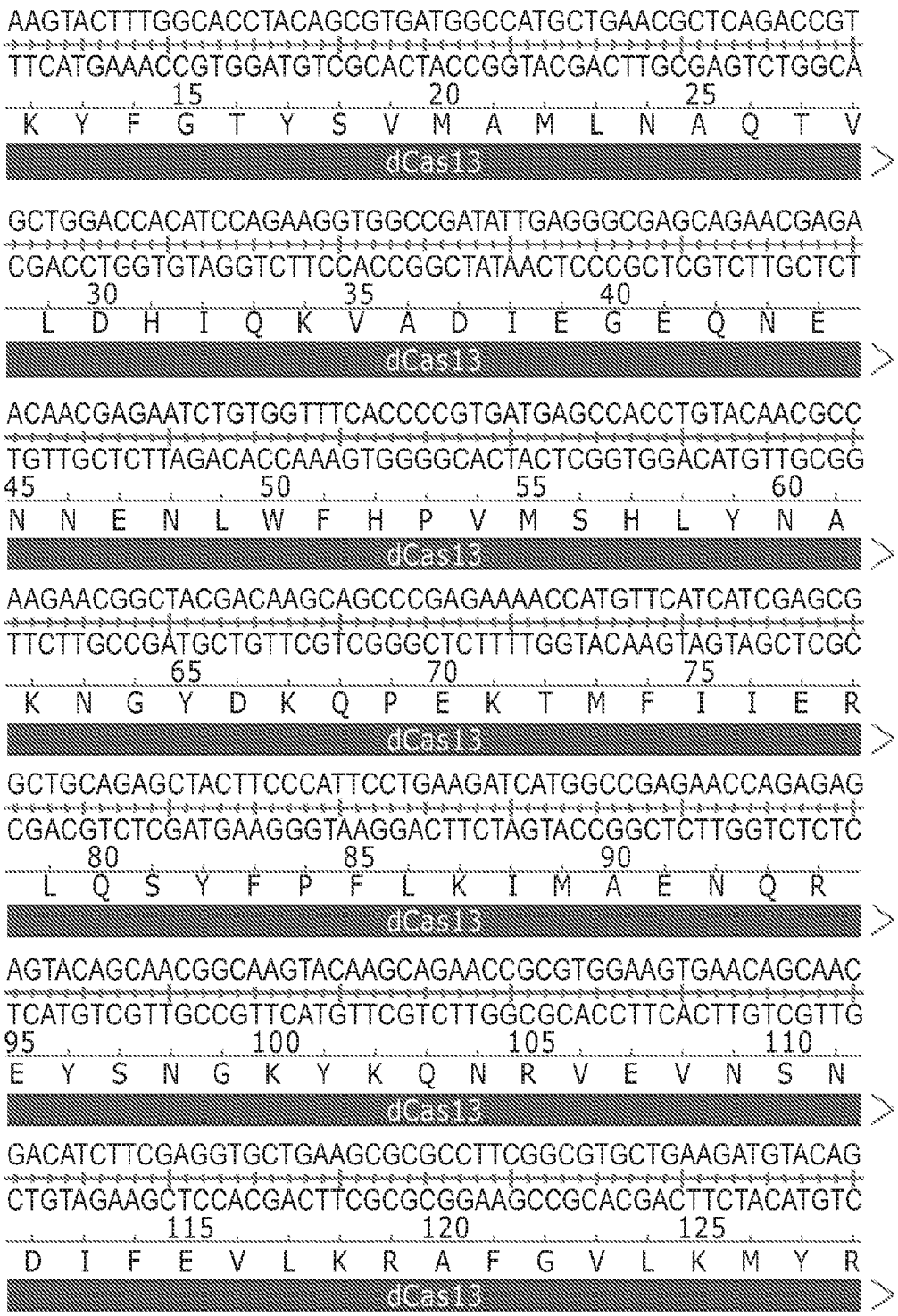

AAGTACTTTGGCACCTACAGCGTGATGGCCATGCTGAACGCTCAGACCGT
TTCATGAAACCGTGGATGTCGCACTACCGGTACGACTTGCGAGTCTGGCA
4850
15 20 25
K Y F G T Y S V M A M L N A Q T V
dCas13
GCTGGACCACATCCAGAAGGTGGCCGATATTGAGGGCGAGCAGAACGAGA
CGACCTGGTGTAGGTCTTCCACCGGCTATAACTCCCGCTCGTCTTGCTCT
4900
30 35 40
L D H I Q K V A D I E G E Q N E
dCas13
ACAACGAGAATCTGTGGTTTCACCCCGTGATGAGCCACCTGTACAACGCC
TGTTGCTCTTAGACACCAAAGTGGGGCACTACTCGGTGGACATGTTGCGG
4950
45 50 55 60
N N E N L W F H P V M S H L Y N A
dCas13
AAGAACGGCTACGACAAGCAGCCCGAGAAAACCATGTTCATCATCGAGCG
TTCTTGCCGATGCTGTTCGTCGGGCTCTTTTGGTACAAGTAGTAGCTCGC
5000
65 70 75
K N G Y D K Q P E K T M F I I E R
dCas13
GCTGCAGAGCTACTTCCCATTCCTGAAGATCATGGCCGAGAACCAGAGAG
CGACGTCTCGATGAAGGGTAAGGACTTCTAGTACCGGCTCTTGGTCTCTC
5050
80 85 90
L Q S Y F P F L K I M A E N Q R
dCas13
AGTACAGCAACGGCAAGTACAAGCAGAACCGCGTGGAAGTGAACAGCAAC
TCATGTCGTTGCCGTTCATGTTCGTCTTGGCGCACCTTCACTTGTCGTTG
5100
95 100 105 110
E Y S N G K Y K Q N R V E V N S N
dCas13
GACATCTTCGAGGTGCTGAAGCGCGCCTTCGGCGTGCTGAAGATGTACAG
CTGTAGAAGCTCCACGACTTCGCGCGGAAGCCGCACGACTTCTACATGTC
5150
115 120 125
D I F E V L K R A F G V L K M Y R
dCas13

FIG. 5M

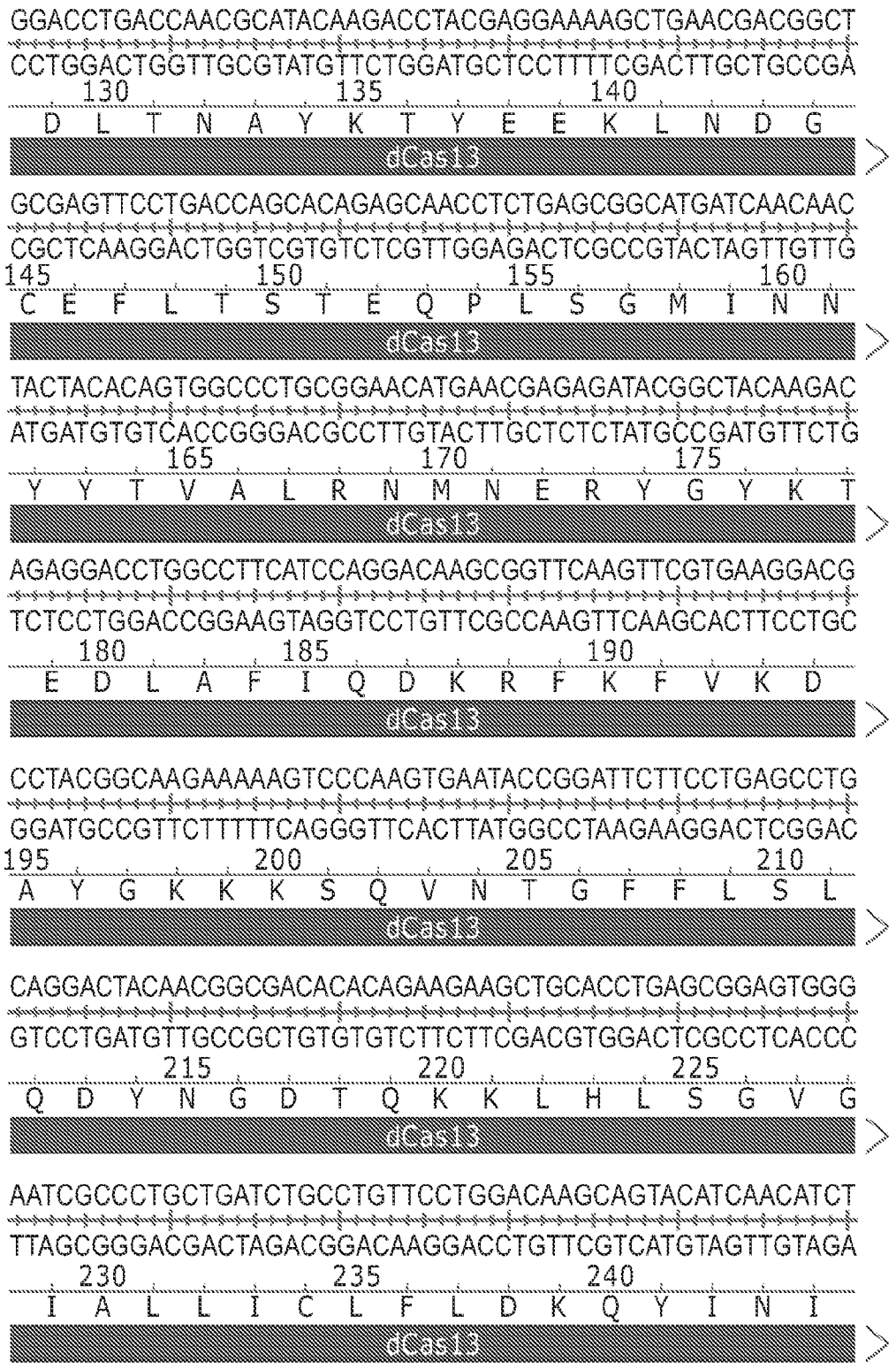

GGACCTGACCAACGCATACAAGACCTACGAGGAAAAGCTGAACGACGGCT
CCTGGACTGGTTGCGTATGTTCTGGATGCTCCTTTTCGACTTGCTGCCGA
5200
130 135 140
D L T N A Y K T Y E E K L N D G
dCas13
GCGAGTTCCTGACCAGCACAGAGCAACCTCTGAGCGGCATGATCAACAAC
CGCTCAAGGACTGGTCGTGTCTCGTTGGAGACTCGCCGTACTAGTTGTTG
5250
145 150 155 160
C E F L T S T E Q P L S G M I N N
dCas13
TACTACACAGTGGCCCTGCGGAACATGAACGAGAGATACGGCTACAAGAC
ATGATGTGTCACCGGGACGCCTTGTACTTGCTCTCTATGCCGATGTTCTG
5300
165 170 175
Y Y T V A L R N M N E R Y G Y K T
dCas13
AGAGGACCTGGCCTTCATCCAGGACAAGCGGTTCAAGTTCGTGAAGGACG
TCTCCTGGACCGGAAGTAGGTCCTGTTCGCCAAGTTCAAGCACTTCCTGC
5350
180 185 190
E D L A F I Q D K R F K F V K D
dCas13
CCTACGGCAAGAAAAAGTCCCAAGTGAATACCGGATTCTTCCTGAGCCTG
GGATGCCGTTCTTTTTCAGGGTTCACTTATGGCCTAAGAAGGACTCGGAC
5400
195 200 205 210
A Y G K K K S Q V N T G F F L S L
dCas13
CAGGACTACAACGGCGACACACAGAAGAAGCTGCACCTGAGCGGAGTGGG
GTCCTGATGTTGCCGCTGTGTGTCTTCTTCGACGTGGACTCGCCTCACCC
5450
215 220 225
Q D Y N G D T Q K K L H L S G V G
dCas13
AATCGCCCTGCTGATCTGCCTGTTCCTGGACAAGCAGTACATCAACATCT
TTAGCGGGACGACTAGACGGACAAGGACCTGTTCGTCATGTAGTTGTAGA
5500
230 235 240
I A L L I C L F L D K Q Y I N I
dCas13

FIG. 5N

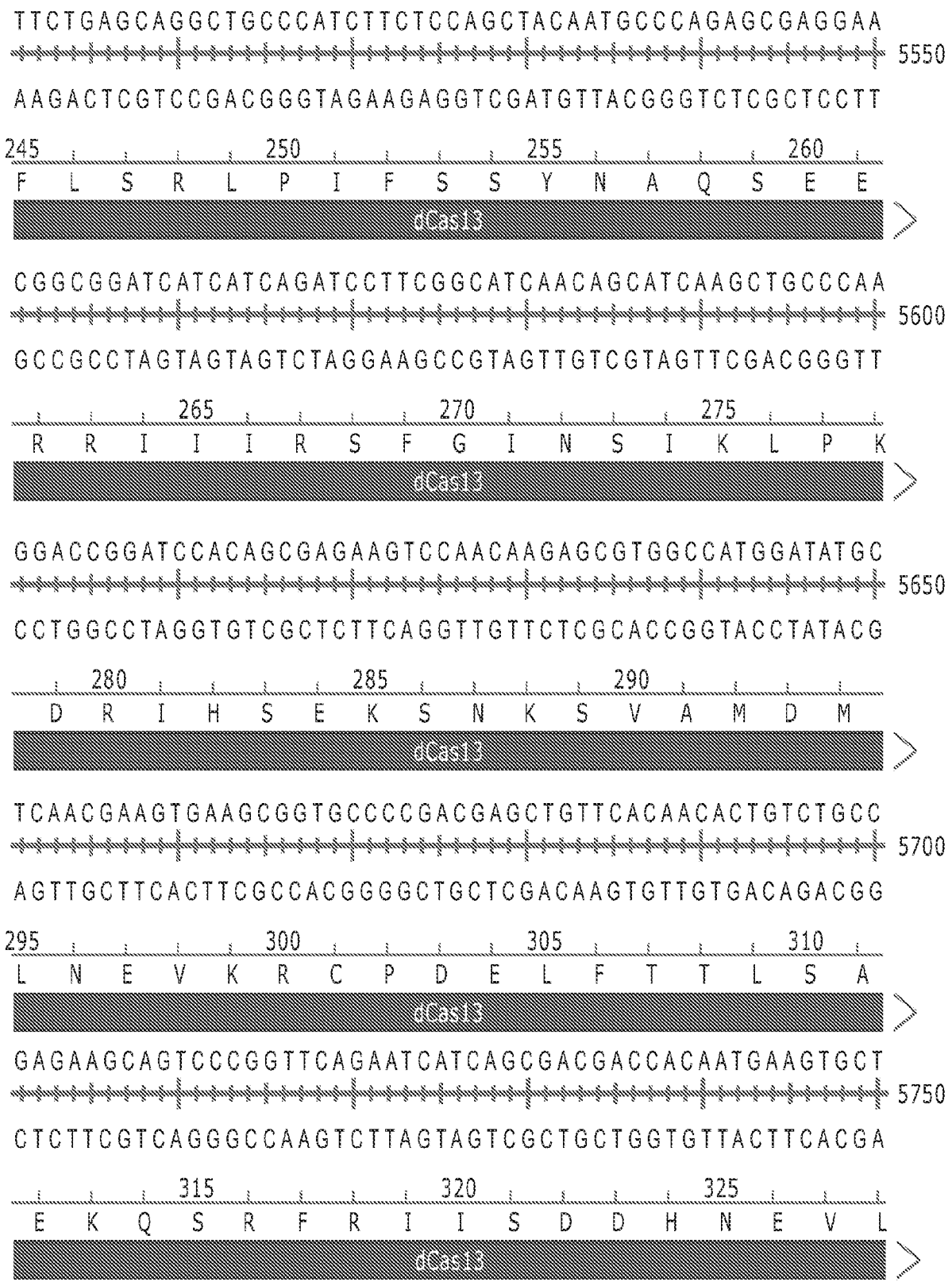

TTCTGAGCAGGCTGCCCATCTTCTCCAGCTACAATGCCCAGAGCGAGGAA
5550
AAGACTCGTCCGACGGGTAGAAGAGGTCGATGTTACGGGTCTCGCTCCTT
245 250 255 260
F L S R L P I F S S Y N A Q S E E
dCas13
CGGCGGATCATCATCAGATCCTTCGGCATCAACAGCATCAAGCTGCCCAA
5600
GCCGCCTAGTAGTAGTCTAGGAAGCCGTAGTTGTCGTAGTTCGACGGGTT
265 270 275
R R I I I R S F G I N S I K L P K
dCas13
GGACCGGATCCACAGCGAGAAGTCCAACAAGAGCGTGGCCATGGATATGC
5650
CCTGGCCTAGGTGTCGCTCTTCAGGTTGTTCTCGCACCGGTACCTATACG
280 285 290
D R I H S E K S N K S V A M D M
dCas13
TCAACGAAGTGAAGCGGTGCCCCGACGAGCTGTTCACAACACTGTCTGCC
5700
AGTTGCTTCACTTCGCCACGGGGCTGCTCGACAAGTGTTGTGACAGACGG
295 300 305 310
L N E V K R C P D E L F T T L S A
dCas13
GAGAAGCAGTCCCGGTTCAGAATCATCAGCGACGACCACAATGAAGTGCT
5750
CTCTTCGTCAGGGCCAAGTCTTAGTAGTCGCTGCTGGTGTTACTTCACGA
315 320 325
E K Q S R F R I I S D D H N E V L
dCas13

FIG. 5O

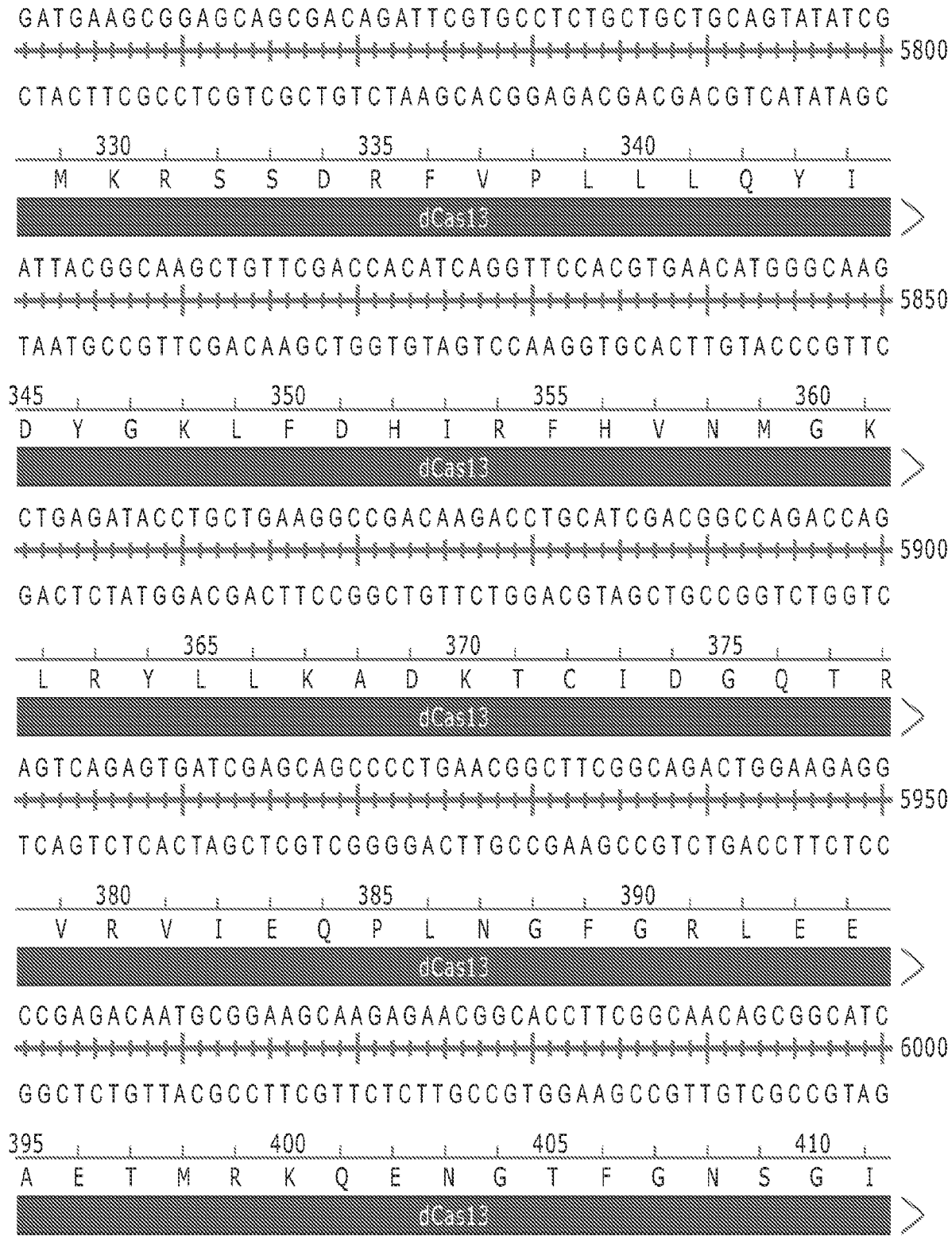

GATGAAGCGGAGCAGCGACAGATTCGTGCCTCTGCTGCTGCAGTATATCG
5800
CTACTTCGCCTCGTCGCTGTCTAAGCACGGAGACGACGACGTCATATAGC
330
335
340
M K R S S D R F V P L L L Q Y I
dCas13
ATTACGGCAAGCTGTTCGACCACATCAGGTTCCACGTGAACATGGGCAAG
5850
TAATGCCGTTCGACAAGCTGGTGTAGTCCAAGGTGCACTTGTACCCGTTC
345
350
355
360
D Y G K L F D H I R F H V N M G K
dCas13
CTGAGATACCTGCTGAAGGCCGACAAGACCTGCATCGACGGCCAGACCAG
5900
GACTCTATGGACGACTTCCGGCTGTTCTGGACGTAGCTGCCGGTCTGGTC
365
370
375
L R Y L L K A D K T C I D G Q T R
dCas13
AGTCAGAGTGATCGAGCAGCCCCTGAACGGCTTCGGCAGACTGGAAGAGG
5950
TCAGTCTCACTAGCTCGTCGGGGACTTGCCGAAGCCGTCTGACCTTCTCC
380
385
390
V R V I E Q P L N G F G R L E E
dCas13
CCGAGACAATGCGGAAGCAAGAGAACGGCACCTTCGGCAACAGCGGCATC
6000
GGCTCTGTTACGCCTTCGTTCTCTTGCCGTGGAAGCCGTTGTCGCCGTAG
395
400
405
410
A E T M R K Q E N G T F G N S G I
dCas13

FIG. 5P

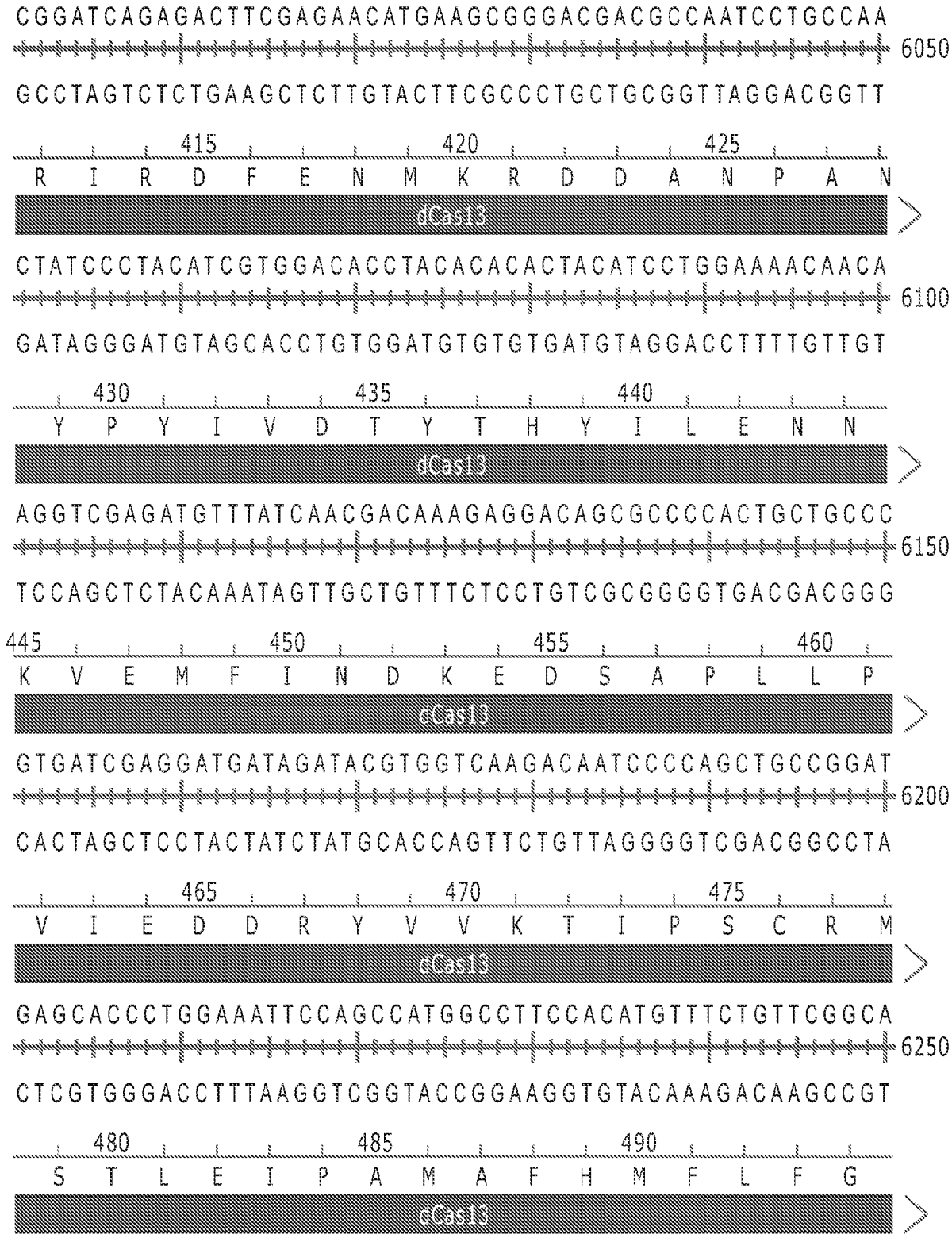

CGGATCAGAGACTTCGAGAACATGAAGCGGGACGACGCCAATCCTGCCAA
6050
GCCTAGTCTCTGAAGCTCTTGTACTTCGCCCTGCTGCGGTTAGGACGGTT
415 420 425
R I R D F E N M K R D D A N P A N
dCas13
CTATCCCTACATCGTGGACACCTACACACACTACATCCTGGAAAACAACA
6100
GATAGGGATGTAGCACCTGTGGATGTGTGTGATGTAGGACCTTTTGTTGT
430 435 440
Y P Y I V D T Y T H Y I L E N N
dCas13
AGGTCGAGATGTTTATCAACGACAAAGAGGACAGCGCCCCACTGCTGCCC
6150
TCCAGCTCTACAAATAGTTGCTGTTTCTCCTGTCGCGGGGTGACGACGGG
445 450 455 460
K V E M F I N D K E D S A P L L P
dCas13
GTGATCGAGGATGATAGATACGTGGTCAAGACAATCCCCAGCTGCCGGAT
6200
CACTAGCTCCTACTATCTATGCACCAGTTCTGTTAGGGGTCGACGGCCTA
465 470 475
V I E D D R Y V V K T I P S C R M
dCas13
GAGCACCCTGGAAATTCCAGCCATGGCCTTCCACATGTTTCTGTTCGGCA
6250
CTCGTGGGACCTTTAAGGTCGGTACCGGAAGGTGTACAAAGACAAGCCGT
480 485 490
S T L E I P A M A F H M F L F G
dCas13

FIG. 5Q

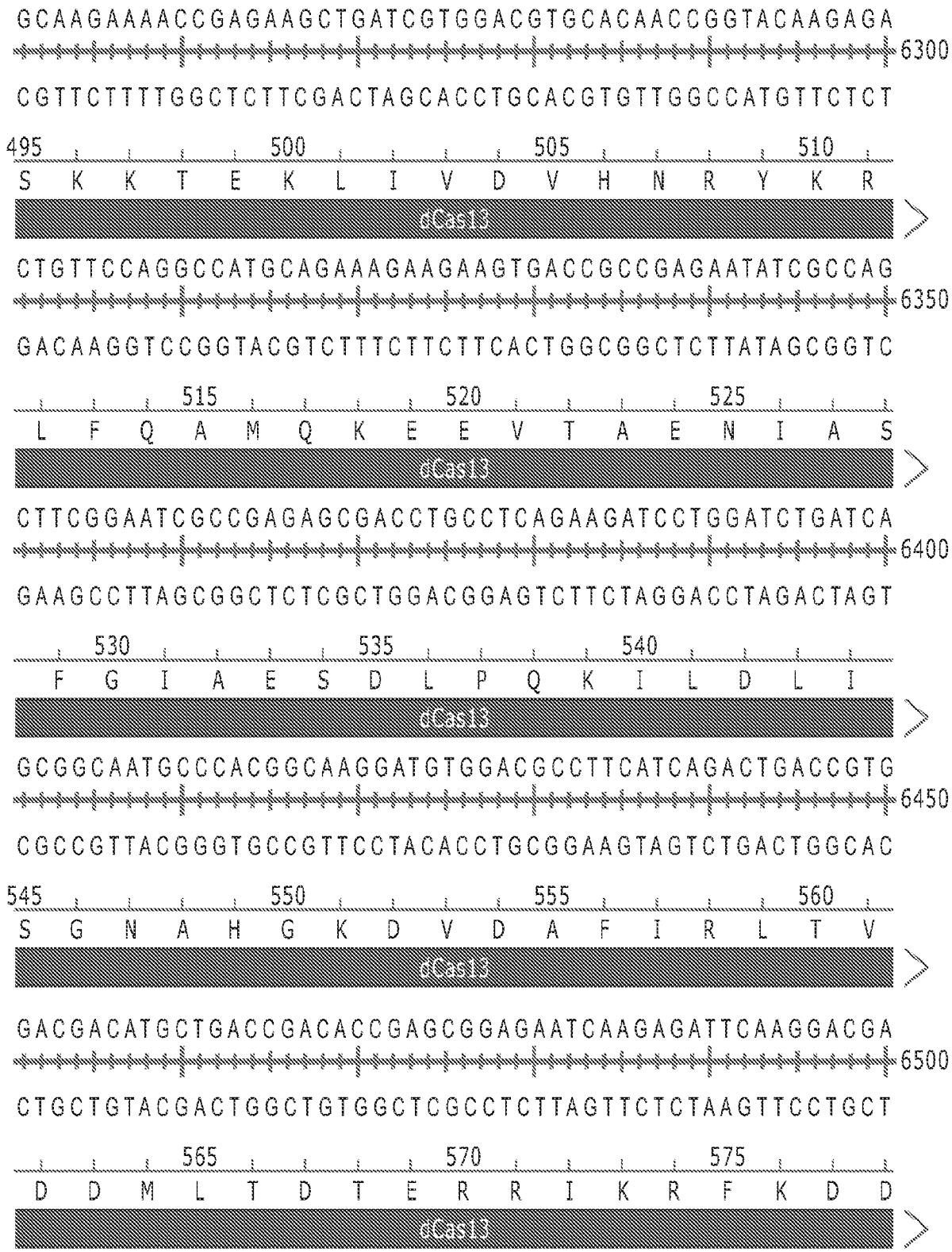

GCAAGAAAACCGAGAAGCTGATCGTGGACGTGCACAACCGGTACAAGAGA
6300
CGTTCTTTTGGCTCTTCGACTAGCACCTGCACGTGTTGGCCATGTTCTCT
495 500 505 510
S K K T E K L I V D V H N R Y K R
dCas13
CTGTTCCAGGCCATGCAGAAAGAAGAAGTGACCGCCGAGAATATCGCCAG
6350
GACAAGGTCCGGTACGTCTTTCTTCTTCACTGGCGGCTCTTATAGCGGTC
515 520 525
L F Q A M Q K E E V T A E N I A S
dCas13
CTTCGGAATCGCCGAGAGCGACCTGCCTCAGAAGATCCTGGATCTGATCA
6400
GAAGCCTTAGCGGCTCTCGCTGGACGGAGTCTTCTAGGACCTAGACTAGT
530 535 540
F G I A E S D L P Q K I L D L I
dCas13
GCGGCAATGCCCACGGCAAGGATGTGGACGCCTTCATCAGACTGACCGTG
6450
CGCCGTTACGGGTGCCGTTCCTACACCTGCGGAAGTAGTCTGACTGGCAC
545 550 555 560
S G N A H G K D V D A F I R L T V
dCas13
GACGACATGCTGACCGACACCGAGCGGAGAATCAAGAGATTCAAGGACGA
6500
CTGCTGTACGACTGGCTGTGGCTCGCCTCTTAGTTCTCTAAGTTCCTGCT
565 570 575
D D M L T D T E R R I K R F K D D
dCas13

FIG. 5R

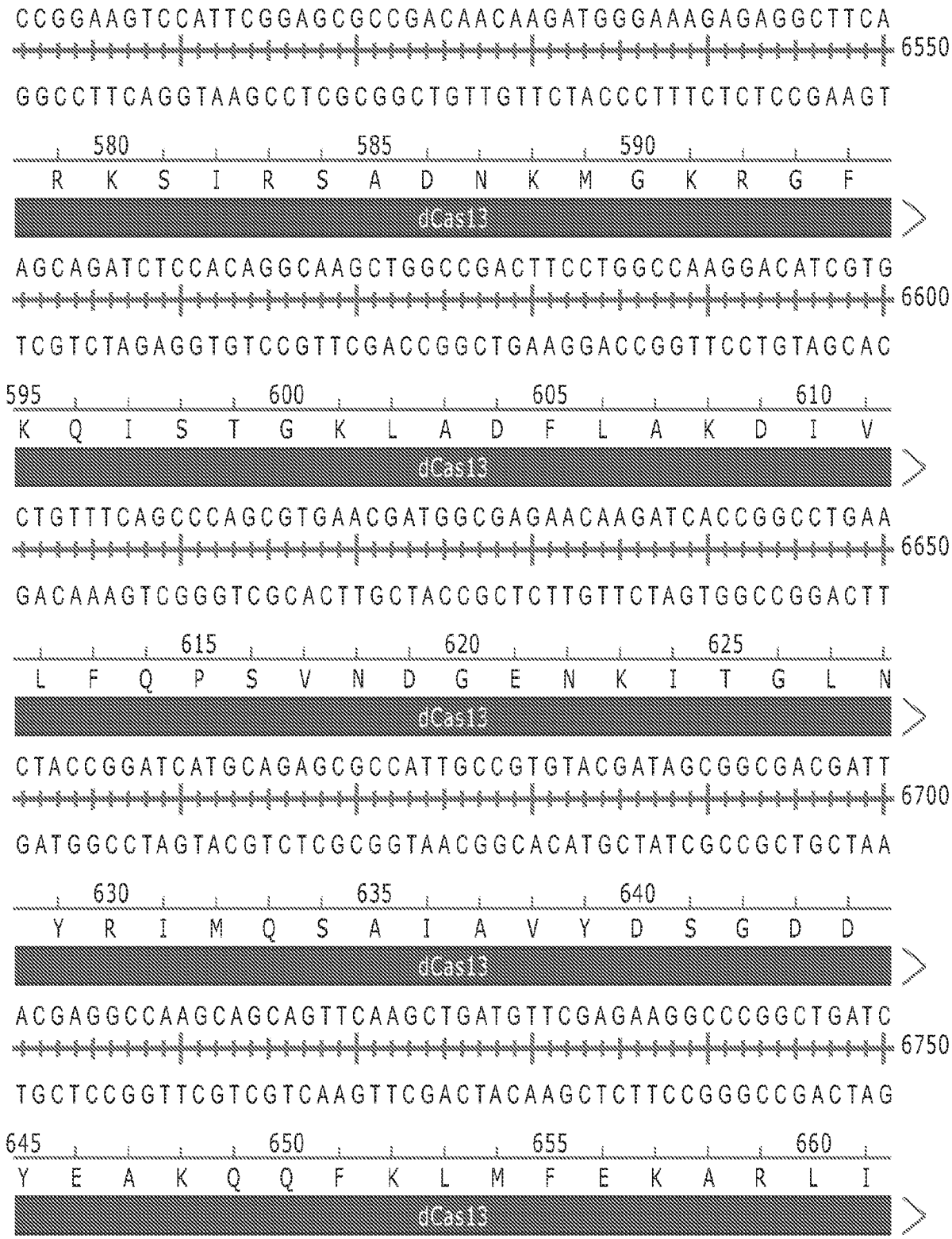

CCGGAAGTCCATTCGGAGCGCCGACAACAAGATGGGAAAGAGAGGCTTCA
6550
GGCCTTCAGGTAAGCCTCGCGGCTGTTGTTCTACCCTTTCTCTCCGAAGT
580
585
590
R K S I R S A D N K M G K R G F
dCas13
AGCAGATCTCCACAGGCAAGCTGGCCGACTTCCTGGCCAAGGACATCGTG
6600
TCGTCTAGAGGTGTCCGTTCGACCGGCTGAAGGACCGGTTCCTGTAGCAC
595
600
605
610
K Q I S T G K L A D F L A K D I V
dCas13
CTGTTTCAGCCCAGCGTGAACGATGGCGAGAACAAGATCACCGGCCTGAA
6650
GACAAAGTCGGGTCGCACTTGCTACCGCTCTTGTTCTAGTGGCCGGACTT
615
620
625
L F Q P S V N D G E N K I T G L N
dCas13
CTACCGGATCATGCAGAGCGCCATTGCCGTGTACGATAGCGGCGACGATT
6700
GATGGCCTAGTACGTCTCGCGGTAACGGCACATGCTATCGCCGCTGCTAA
630
635
640
Y R I M Q S A I A V Y D S G D D
dCas13
ACGAGGCCAAGCAGCAGTTCAAGCTGATGTTCGAGAAGGCCCGGCTGATC
6750
TGCTCCGGTTCGTCGTCAAGTTCGACTACAAGCTCTTCCGGGCCGACTAG
645
650
655
660
Y E A K Q Q F K L M F E K A R L I
dCas13

FIG. 5S

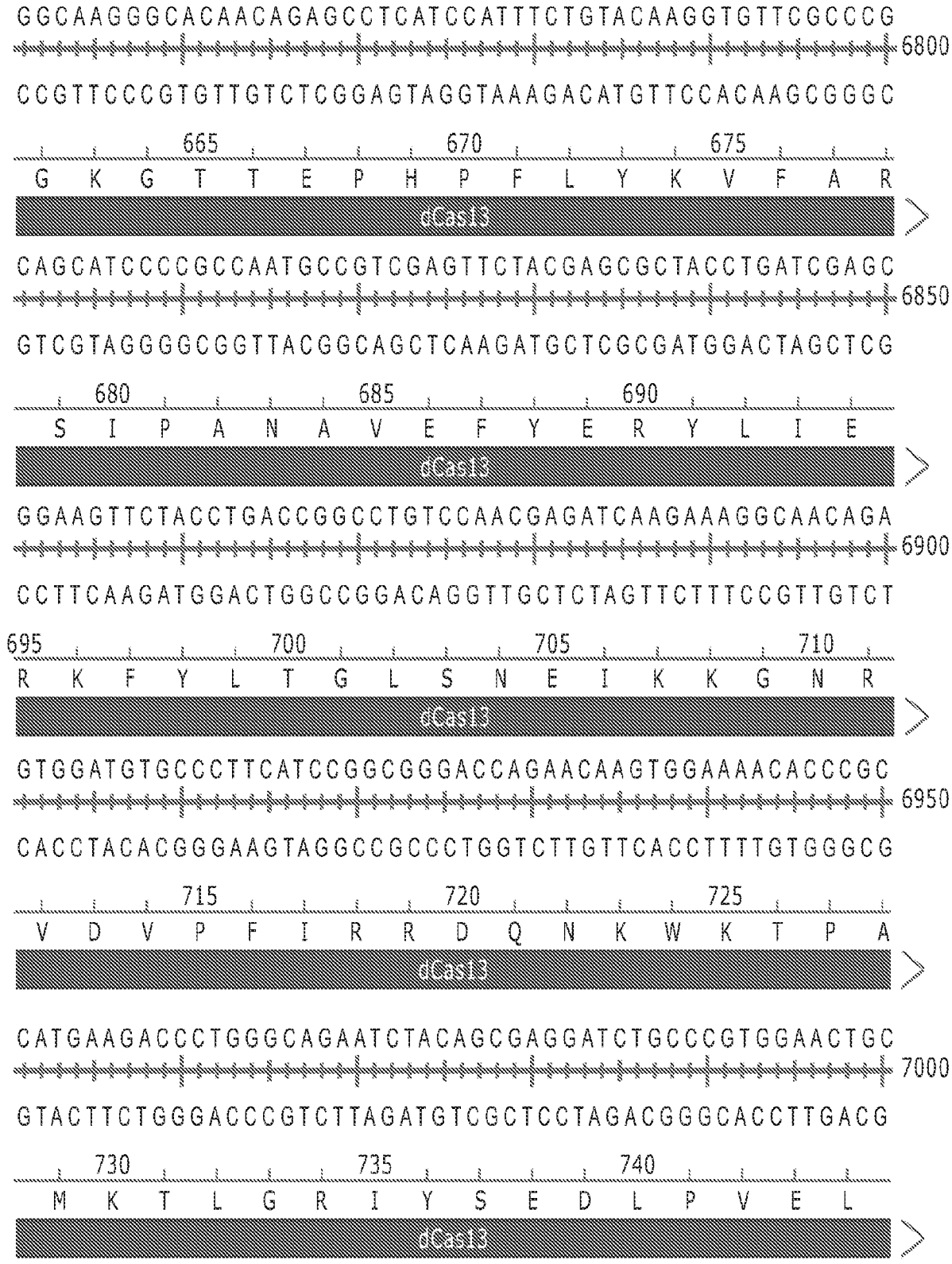

GGCAAGGGCACAACAGAGCCTCATCCATTTCTGTACAAGGTGTTCGCCCG
6800
CCGTTCCCGTGTTGTCTCGGAGTAGGTAAAGACATGTTCCACAAGCGGGC
665
670
675
G K G T T E P H P F L Y K V F A R
dCas13
CAGCATCCCCGCCAATGCCGTCGAGTTCTACGAGCGCTACCTGATCGAGC
6850
GTCGTAGGGGCGGTTACGGCAGCTCAAGATGCTCGCGATGGACTAGCTCG
680
685
690
S I P A N A V E F Y E R Y L I E
dCas13
GGAAGTTCTACCTGACCGGCCTGTCCAACGAGATCAAGAAAGGCAACAGA
6900
CCTTCAAGATGGACTGGCCGGACAGGTTGCTCTAGTTCTTTCCGTTGTCT
695
700
705
710
R K F Y L T G L S N E I K K G N R
dCas13
GTGGATGTGCCCTTCATCCGGCGGGACCAGAACAAGTGGAAAACACCCGC
6950
CACCTACACGGGAAGTAGGCCGCCCTGGTCTTGTTCACCTTTTGTGGGCG
715
720
725
V D V P F I R R D Q N K W K T P A
dCas13
CATGAAGACCCTGGGCAGAATCTACAGCGAGGATCTGCCCGTGGAACTGC
7000
GTACTTCTGGGACCCGTCTTAGATGTCGCTCCTAGACGGGCACCTTGACG
730
735
740
M K T L G R I Y S E D L P V E L
dCas13

FIG. 5T

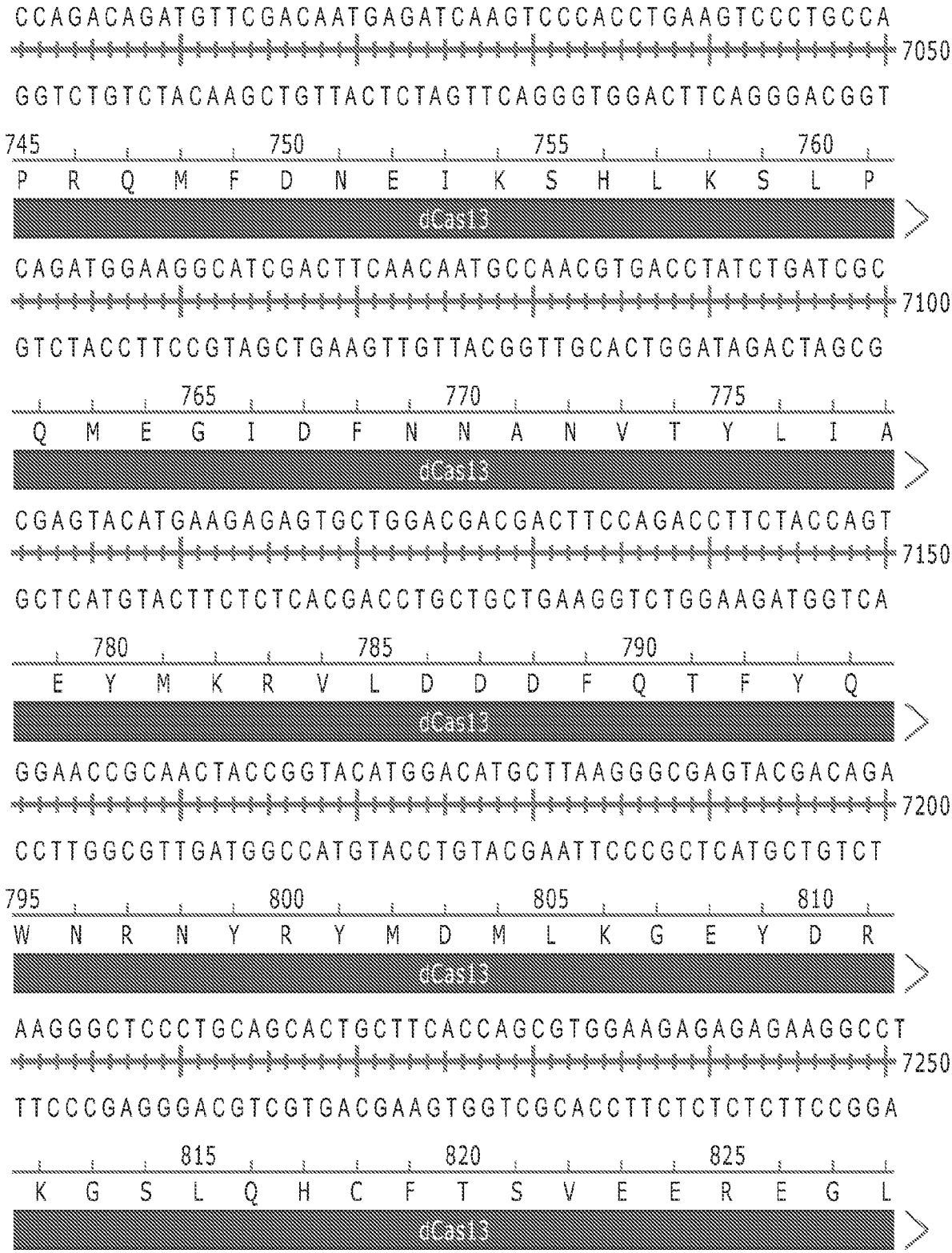

CCAGACAGATGTTCGACAATGAGATCAAGTCCCACCTGAAGTCCCTGCCA
7050
GGTCTGTCTACAAGCTGTTACTCTAGTTCAGGGTGGACTTCAGGGACGGT
745
750
755
760
P R Q M F D N E I K S H L K S L P
dCas13
CAGATGGAAGGCATCGACTTCAACAATGCCAACGTGACCTATCTGATCGC
7100
GTCTACCTTCCGTAGCTGAAGTTGTTACGGTTGCACTGGATAGACTAGCG
765
770
775
Q M E G I D F N N A N V T Y L I A
dCas13
CGAGTACATGAAGAGAGTGCTGGACGACGACTTCCAGACCTTCTACCAGT
7150
GCTCATGTACTTCTCTCACGACCTGCTGCTGAAGGTCTGGAAGATGGTCA
780
785
790
E Y M K R V L D D D F Q T F Y Q
dCas13
GGAACCGCAACTACCGGTACATGGACATGCTTAAGGGCGAGTACGACAGA
7200
CCTTGGCGTTGATGGCCATGTACCTGTACGAATTCCCGCTCATGCTGTCT
795
800
805
810
W N R N Y R Y M D M L K G E Y D R
dCas13
AAGGGCTCCCTGCAGCACTGCTTCACCAGCGTGGAAGAGAGAGAAGGCCT
7250
TTCCCGAGGGACGTCGTGACGAAGTGGTCGCACCTTCTCTCTCTTCCGGA
815
820
825
K G S L Q H C F T S V E E R E G L
dCas13

FIG. 5U

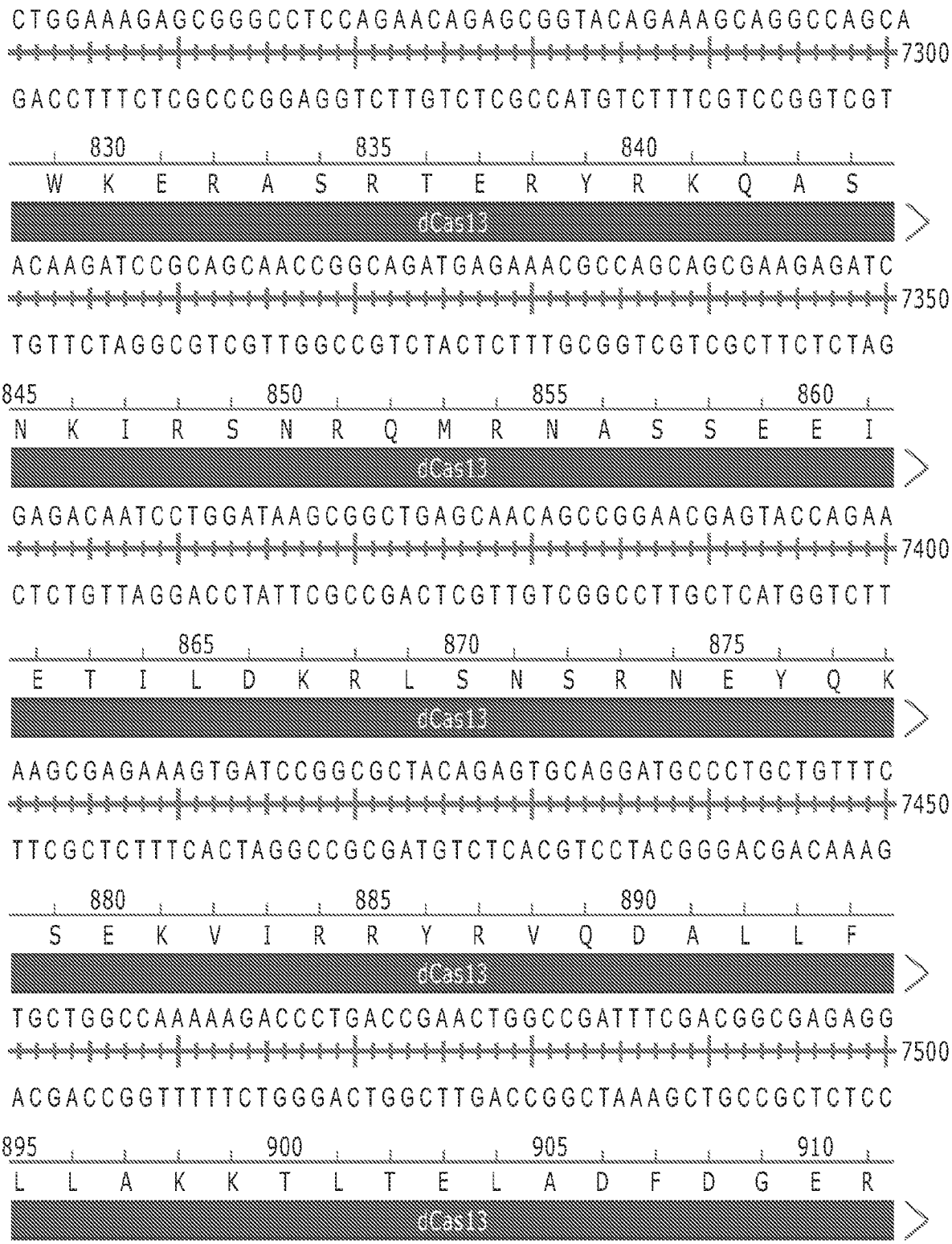

CTGGAAAGAGCGGGCCTCCAGAACAGAGCGGTACAGAAAGCAGGCCAGCA
7300
GACCTTTCTCGCCCGGAGGTCTTGTCTCGCCATGTCTTTCGTCCGGTCGT
830
835
840
W K E R A S R T E R Y R K Q A S
dCas13
ACAAGATCCGCAGCAACCGGCAGATGAGAAACGCCAGCAGCGAAGAGATC
7350
TGTTCTAGGCGTCGTTGGCCGTCTACTCTTTGCGGTCGTCGCTTCTCTAG
845
850
855
860
N K I R S N R Q M R N A S S E E I
dCas13
GAGACAATCCTGGATAAGCGGCTGAGCAACAGCCGGAACGAGTACCAGAA
7400
CTCTGTTAGGACCTATTCGCCGACTCGTTGTCGGCCTTGCTCATGGTCTT
865
870
875
E T I L D K R L S N S R N E Y Q K
dCas13
AAGCGAGAAAGTGATCCGGCGCTACAGAGTGCAGGATGCCCTGCTGTTTC
7450
TTCGCTCTTTCACTAGGCCGCGATGTCTCACGTCCTACGGGACGACAAAG
880
885
890
S E K V I R R Y R V Q D A L L F
dCas13
TGCTGGCCAAAAAGACCCTGACCGAACTGGCCGATTTCGACGGCGAGAGG
7500
ACGACCGGTTTTTCTGGGACTGGCTTGACCGGCTAAAGCTGCCGCTCTCC
895
900
905
910
L L A K K T L T E L A D F D G E R
dCas13

FIG. 5V

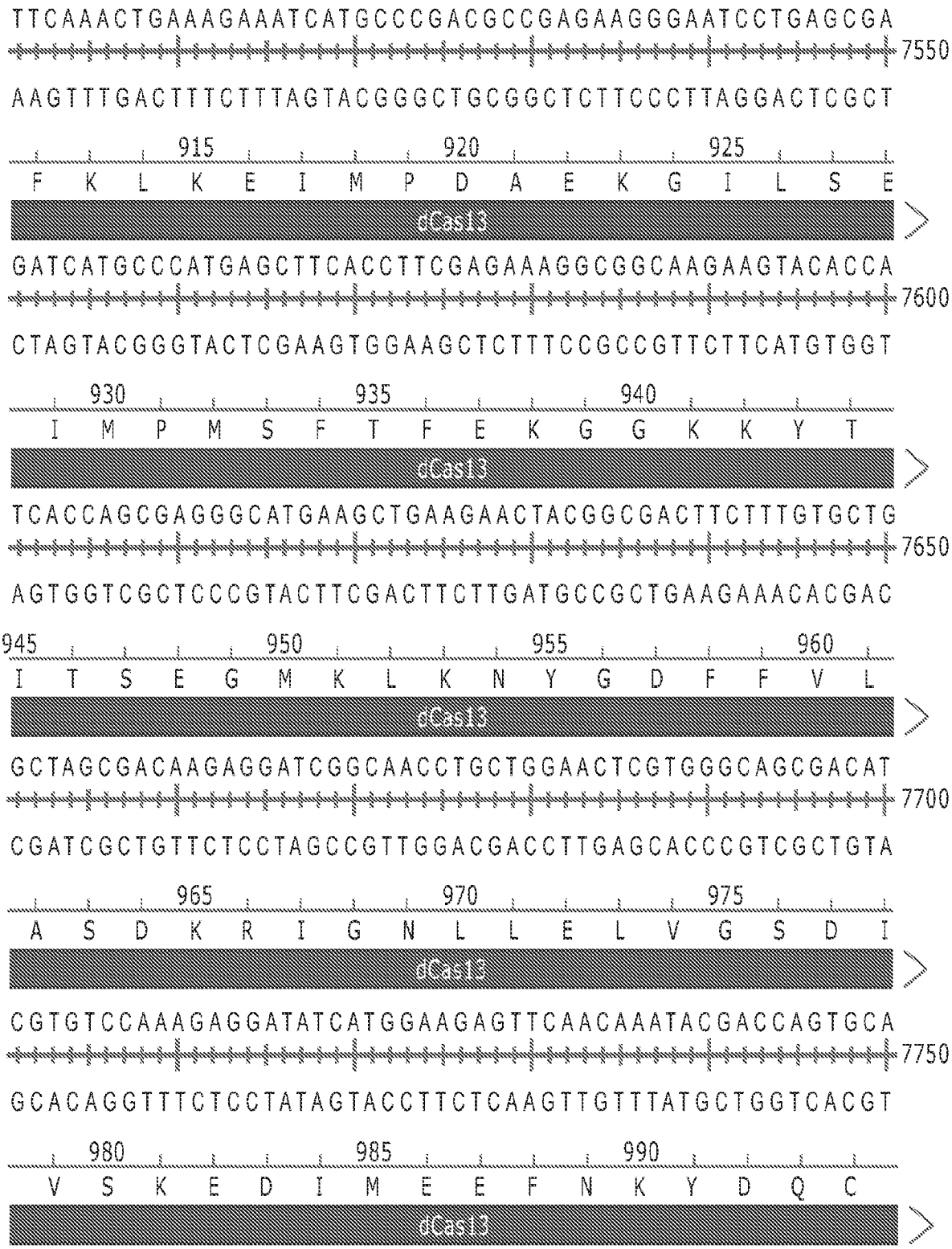

TTCAAACTGAAAGAAATCATGCCCGACGCCGAGAAGGGAATCCTGAGCGA
7550
AAGTTTGACTTTCTTTAGTACGGGCTGCGGCTCTTCCCTTAGGACTCGCT
915 920 925
F K L K E I M P D A E K G I L S E
dCas13
GATCATGCCCATGAGCTTCACCTTCGAGAAAGGCGGCAAGAAGTACACCA
7600
CTAGTACGGGTACTCGAAGTGGAAGCTCTTTCCGCCGTTCTTCATGTGGT
930 935 940
I M P M S F T F E K G G K K Y T
dCas13
TCACCAGCGAGGGCATGAAGCTGAAGAACTACGGCGACTTCTTTGTGCTG
7650
AGTGGTCGCTCCCGTACTTCGACTTCTTGATGCCGCTGAAGAAACACGAC
945 950 955 960
I T S E G M K L K N Y G D F F V L
dCas13
GCTAGCGACAAGAGGATCGGCAACCTGCTGGAACTCGTGGGCAGCGACAT
7700
CGATCGCTGTTCTCCTAGCCGTTGGACGACCTTGAGCACCCGTCGCTGTA
965 970 975
A S D K R I G N L L E L V G S D I
dCas13
CGTGTCCAAAGAGGATATCATGGAAGAGTTCAACAAATACGACCAGTGCA
7750
GCACAGGTTTCTCCTATAGTACCTTCTCAAGTTGTTTATGCTGGTCACGT
980 985 990
V S K E D I M E E F N K Y D Q C
dCas13

FIG. 5W

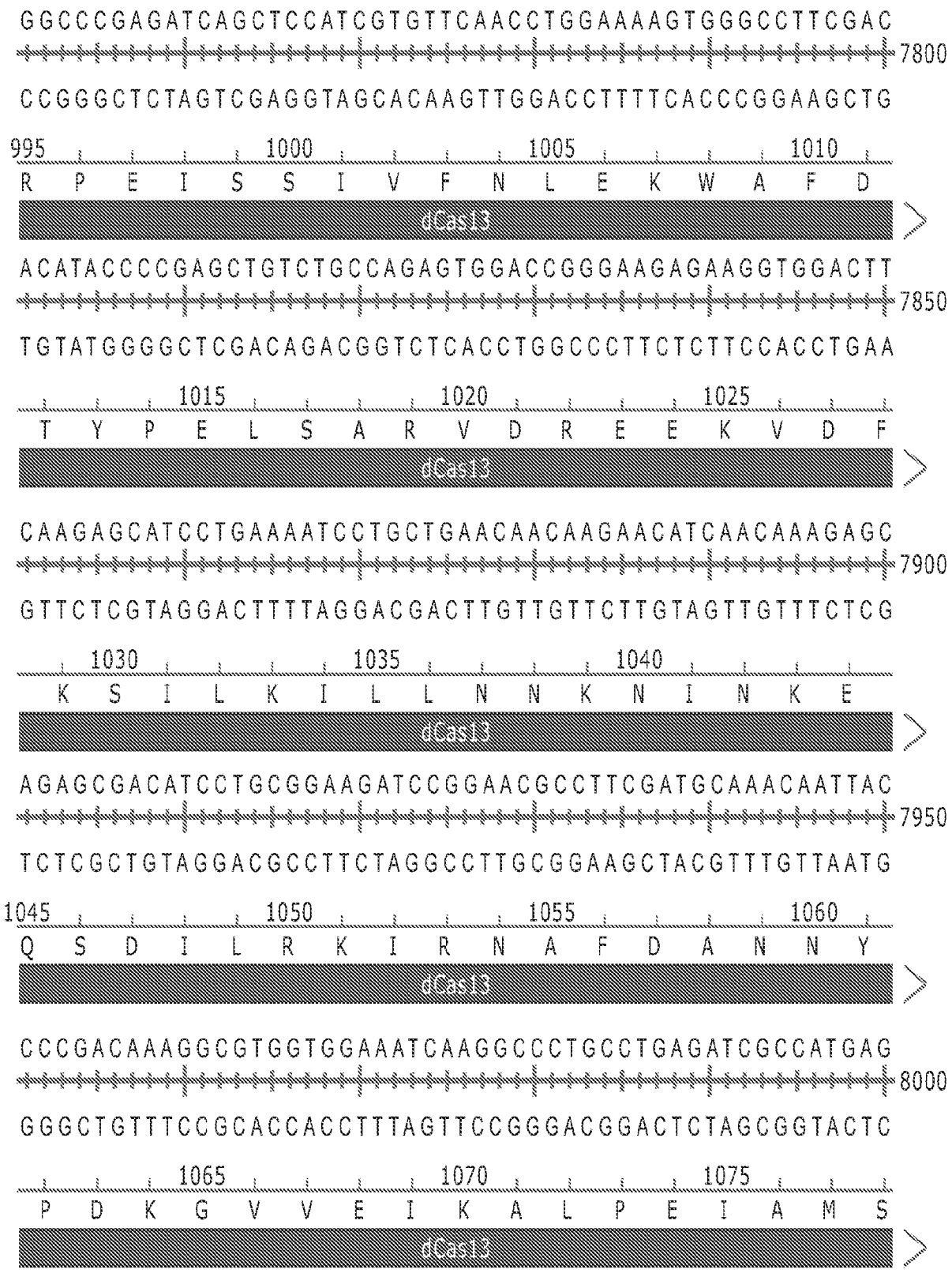

GGCCCGAGATCAGCTCCATCGTGTTCAACCTGGAAAAGTGGGCCTTCGAC
7800
CCGGGCTCTAGTCGAGGTAGCACAAGTTGGACCTTTTCACCCGGAAGCTG
995
1000
1005
1010
R P E I S S I V F N L E K W A F D
dCas13
ACATACCCCGAGCTGTCTGCCAGAGTGGACCGGGAAGAGAAGGTGGACTT
7850
TGTATGGGGCTCGACAGACGGTCTCACCTGGCCCTTCTCTTCCACCTGAA
1015
1020
1025
T Y P E L S A R V D R E E K V D F
dCas13
CAAGAGCATCCTGAAAATCCTGCTGAACAACAAGAACATCAACAAAGAGC
7900
GTTCTCGTAGGACTTTTAGGACGACTTGTTGTTCTTGTAGTTGTTTCTCG
1030
1035
1040
K S I L K I L L N N K N I N K E
dCas13
AGAGCGACATCCTGCGGAAGATCCGGAACGCCTTCGATGCAAACAATTAC
7950
TCTCGCTGTAGGACGCCTTCTAGGCCTTGCGGAAGCTACGTTTGTTAATG
1045
1050
1055
1060
Q S D I L R K I R N A F D A N N Y
dCas13
CCCGACAAAGGCGTGGTGGAAATCAAGGCCCTGCCTGAGATCGCCATGAG
8000
GGGCTGTTTCCGCACCACCTTTAGTTCCGGGACGGACTCTAGCGGTACTC
1065
1070
1075
P D K G V V E I K A L P E I A M S
dCas13

FIG. 5X

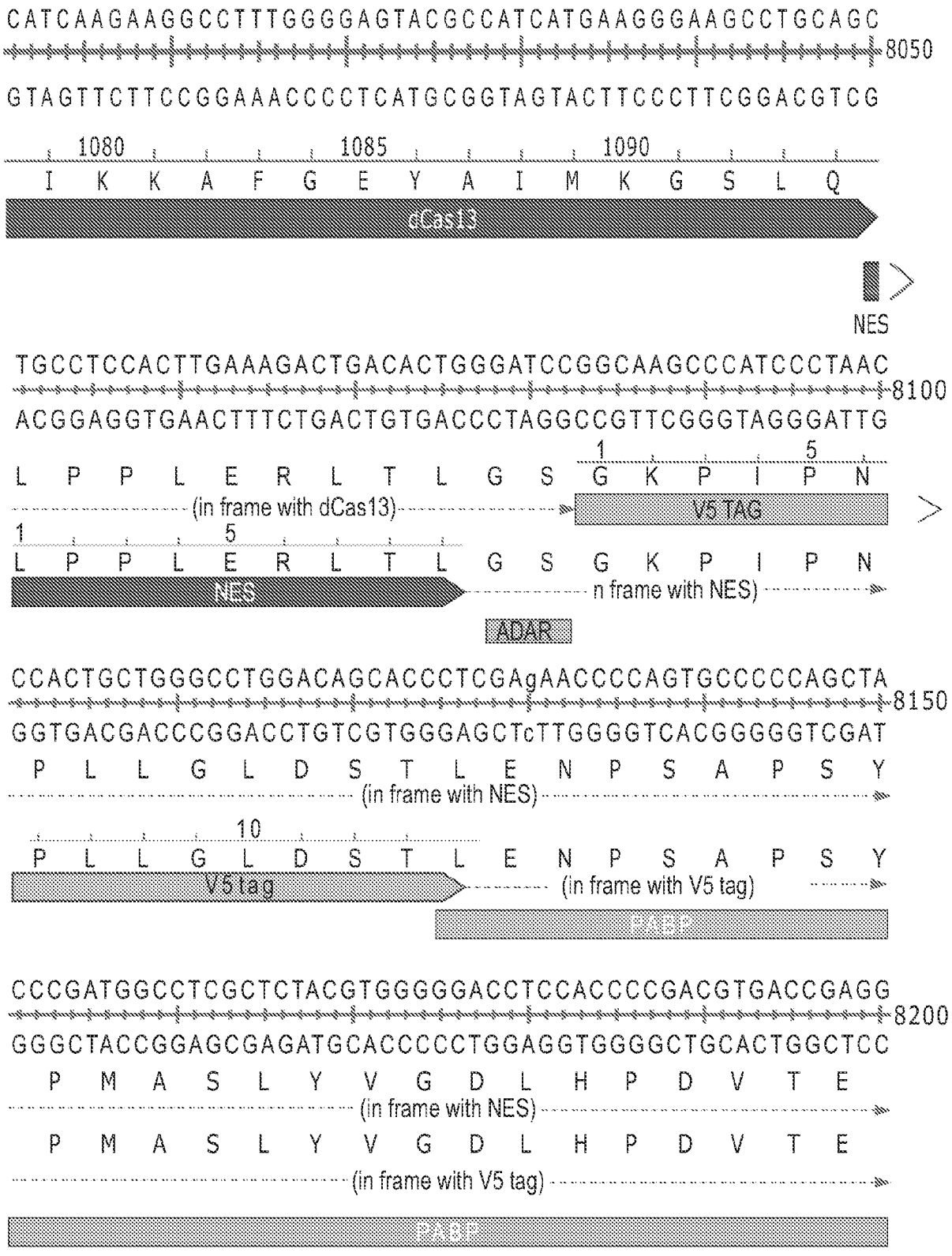

CATCAAGAAGGCCTTTGGGGAGTACGCCATCATGAAGGGAAGCCTGCAGC
8050
GTAGTTCTTCCGGAAACCCCTCATGCGGTAGTACTTCCCTTCGGACGTCG
1080
1085
1090
I K K A F G E Y A I M K G S L Q
dCas13
NES
TGCCTCCACTTGAAAGACTGACACTGGGATCCGGCAAGCCCATCCCTAAC
8100
ACGGAGGTGAACTTTCTGACTGTGACCCTAGGCCGTTCGGGTAGGGATTG
L P P L E R L T L G S G K P I P N
(in frame with dCas13)
V5 TAG
L P P L E R L T L G S G K P I P N
NES
n frame with NES)
ADAR
CCACTGCTGGGCCTGGACAGCACCCTCGAgAACCCCAGTGCCCCCAGCTA
8150
GGTGACGACCCGGACCTGTCGTGGGAGCTcTTGGGGTCACGGGGGTCGAT
P L L G L D S T L E N P S A P S Y
(in frame with NES)
P L L G L D S T L E N P S A P S Y
V5 tag
(in frame with V5 tag)
PABP
CCCGATGGCCTCGCTCTACGTGGGGGACCTCCACCCCGACGTGACCGAGG
8200
GGGCTACCGGAGCGAGATGCACCCCCTGGAGGTGGGGCTGCACTGGCTCC
P M A S L Y V G D L H P D V T E
(in frame with NES)
P M A S L Y V G D L H P D V T E
(in frame with V5 tag)
PABP

FIG. 5Y

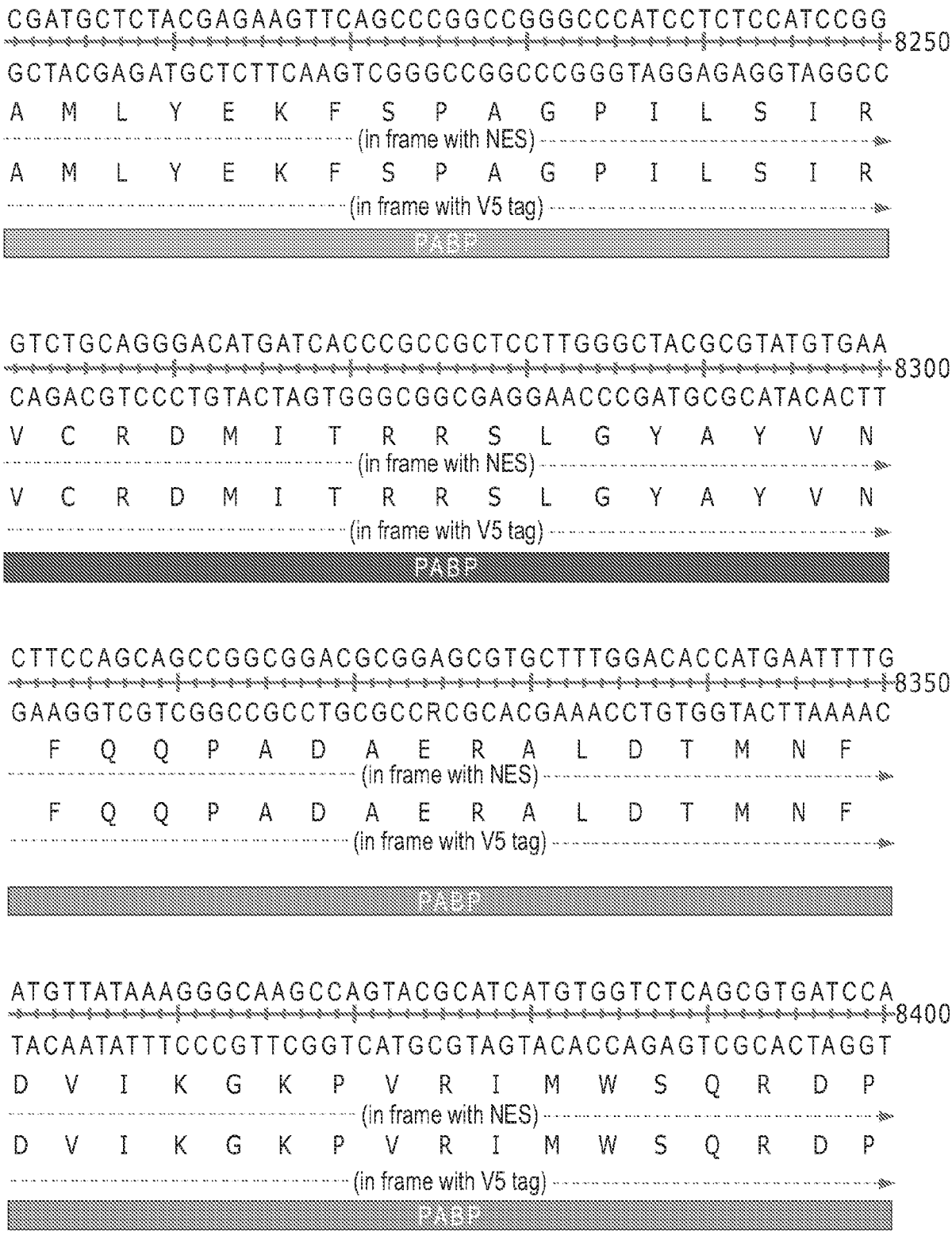

CGATGCTCTACGAGAAGTTCAGCCCGGCCGGGCCCATCCTCTCCATCCGG
8250
GCTACGAGATGCTCTTCAAGTCGGGCCGGCCCGGGTAGGAGAGGTAGGCC
A M L Y E K F S P A G P I L S I R
(in frame with NES)
A M L Y E K F S P A G P I L S I R
(in frame with V5 tag)
PABP
GTCTGCAGGGACATGATCACCCGCCGCTCCTTGGGCTACGCGTATGTGAA
8300
CAGACGTCCCTGTACTAGTGGGCGGCGAGGAACCCGATGCGCATACACTT
V C R D M I T R R S L G Y A Y V N
(in frame with NES)
V C R D M I T R R S L G Y A Y V N
(in frame with V5 tag)
PABP
CTTCCAGCAGCCGGCGGACGCGGAGCGTGCTTTGGACACCATGAATTTTG
8350
GAAGGTCGTCGGCCGCCTGCGCCRCGCACGAAACCTGTGGTACTTAAAAC
F Q Q P A D A E R A L D T M N F
(in frame with NES)
F Q Q P A D A E R A L D T M N F
(in frame with V5 tag)
PABP
ATGTTATAAAGGGCAAGCCAGTACGCATCATGTGGTCTCAGCGTGATCCA
8400
TACAATATTTCCCGTTCGGTCATGCGTAGTACACCAGAGTCGCACTAGGT
D V I K G K P V R I M W S Q R D P
(in frame with NES)
D V I K G K P V R I M W S Q R D P
(in frame with V5 tag)
PABP

FIG. 5Z

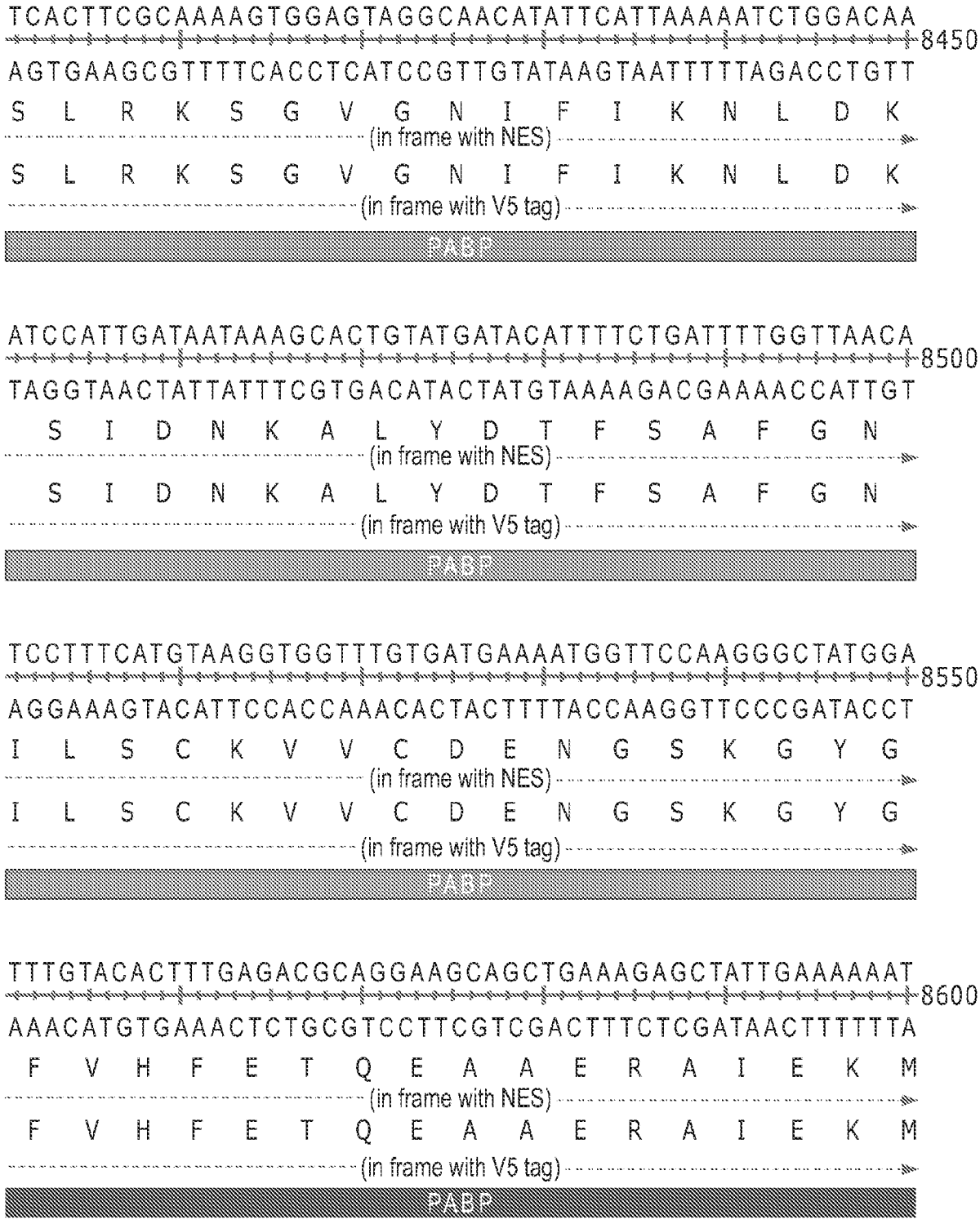

TCACTTCGCAAAAGTGGAGTAGGCAACATATTCATTAAAAATCTGGACAA
8450
AGTGAAGCGTTTTCACCTCATCCGTTGTATAAGTAATTTTTAGACCTGTT
S L R K S G V G N I F I K N L D K
(in frame with NES)
S L R K S G V G N I F I K N L D K
(in frame with V5 tag)
PABP
ATCCATTGATAATAAAGCACTGTATGATACATTTTCTGATTTTGGTTAACA
8500
TAGGTAACTATTATTTCGTGACATACTATGTAAAAGACGAAAACCATTGT
S I D N K A L Y D T F S A F G N
(in frame with NES)
S I D N K A L Y D T F S A F G N
(in frame with V5 tag)
PABP
TCCTTTCATGTAAGGTGGTTTGTGATGAAAATGGTTCCAAGGGCTATGGA
8550
AGGAAAGTACATTCCACCAAACACTACTTTTACCAAGGTTCCCGATACCT
I L S C K V V C D E N G S K G Y G
(in frame with NES)
I L S C K V V C D E N G S K G Y G
(in frame with V5 tag)
PABP
TTTGTACACTTTGAGACGCAGGAAGCAGCTGAAAGAGCTATTGAAAAAAT
8600
AAACATGTGAAACTCTGCGTCCTTCGTCGACTTTCTCGATAACTTTTTTA
F V H F E T Q E A A E R A I E K M
(in frame with NES)
F V H F E T Q E A A E R A I E K M
(in frame with V5 tag)
PABP

FIG. 5AA

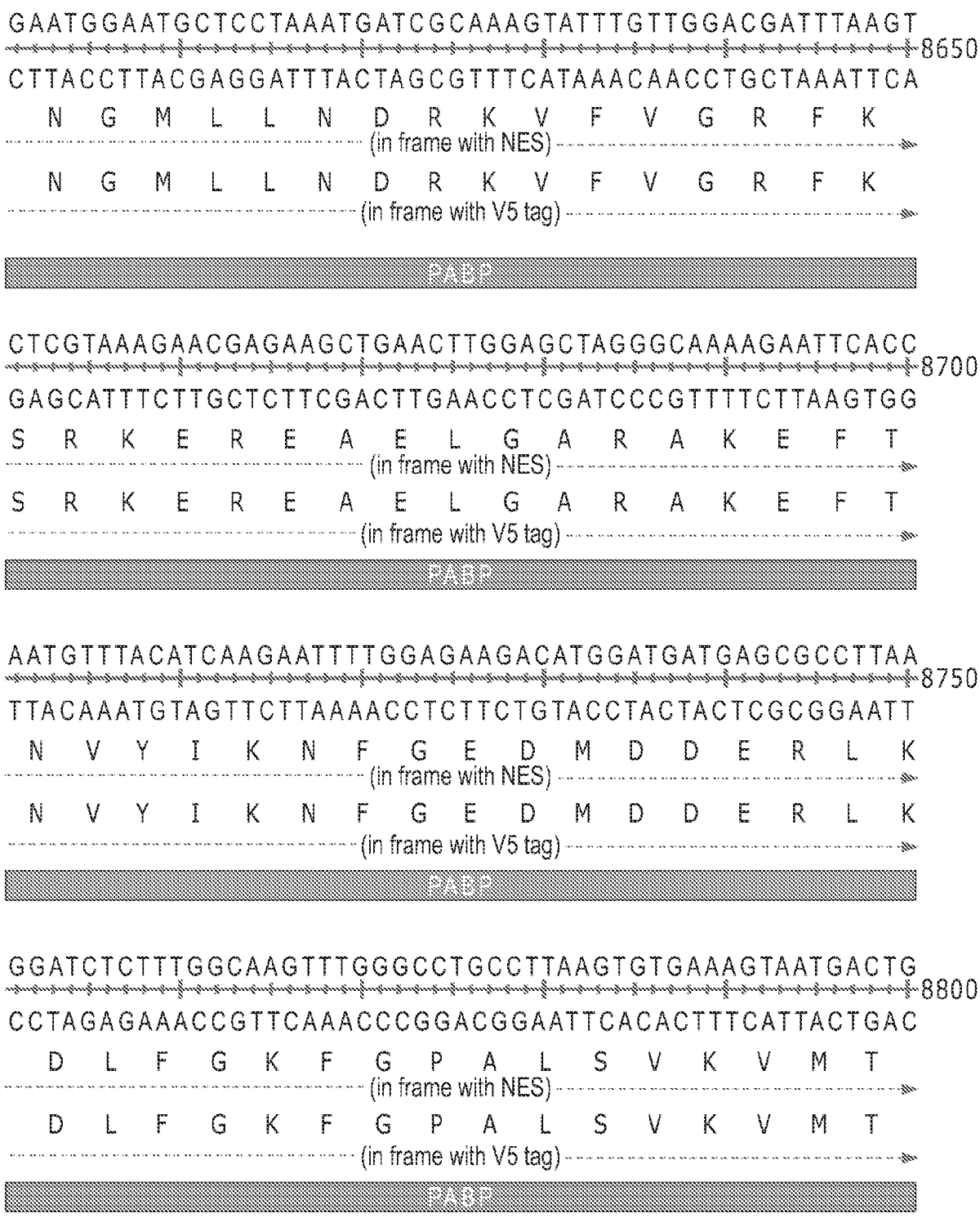

GAATGGAATGCTCCTAAATGATCGCAAAGTATTTGTTGGACGATTTAAGT
8650
CTTACCTTACGAGGATTTACTAGCGTTTCATAAACAACCTGCTAAATTCA
N G M L L N D R K V F V G R F K
(in frame with NES)
N G M L L N D R K V F V G R F K
(in frame with V5 tag)
PABP
CTCGTAAAGAACGAGAAGCTGAACTTGGAGCTAGGGCAAAAGAATTCACC
8700
GAGCATTTCTTGCTCTTCGACTTGAACCTCGATCCCGTTTTCTTAAGTGG
S R K E R E A E L G A R A K E F T
(in frame with NES)
S R K E R E A E L G A R A K E F T
(in frame with V5 tag)
PABP
AATGTTTACATCAAGAATTTTGGAGAAGACATGGATGATGAGCGCCTTAA
8750
TTACAAATGTAGTTCTTAAAACCTCTTCTGTACCTACTACTCGCGGAATT
N V Y I K N F G E D M D D E R L K
(in frame with NES)
N V Y I K N F G E D M D D E R L K
(in frame with V5 tag)
PABP
GGATCTCTTTGGCAAGTTTGGGCCTGCCTTAAGTGTGAAAGTAATGACTG
8800
CCTAGAGAAACCGTTCAAACCCGGACGGAATTCACACTTTCATTACTGAC
D L F G K F G P A L S V K V M T
(in frame with NES)
D L F G K F G P A L S V K V M T
(in frame with V5 tag)
PABP

FIG. 5AB

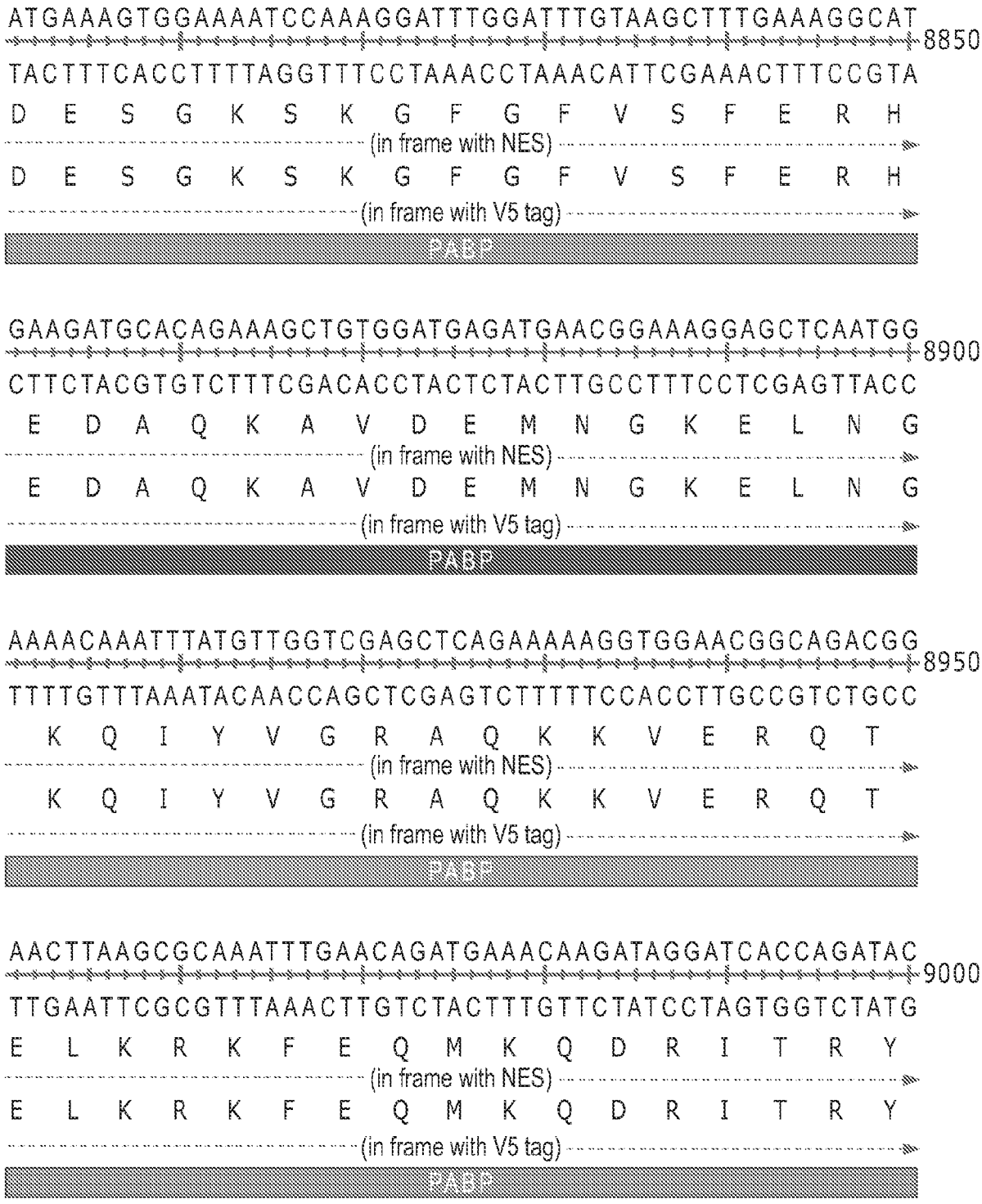

ATGAAAGTGGAAAATCCAAAGGATTTGGATTTGTAAGCTTTGAAAGGCAT
8850
TACTTTCACCTTTTAGGTTTCCTAAACCTAAACATTCGAAACTTTCCGTA
D E S G K S K G F G F V S F E R H
(in frame with NES)
D E S G K S K G F G F V S F E R H
(in frame with V5 tag)
PABP
GAAGATGCACAGAAAGCTGTGGATGAGATGAACGGAAAGGAGCTCAATGG
8900
CTTCTACGTGTCTTTCGACACCTACTCTACTTGCCTTTCCTCGAGTTACC
E D A Q K A V D E M N G K E L N G
(in frame with NES)
E D A Q K A V D E M N G K E L N G
(in frame with V5 tag)
PABP
AAAACAAATTTATGTTGGTCGAGCTCAGAAAAAGGTGGAACGGCAGACGG
8950
TTTTGTTTAAATACAACCAGCTCGAGTCTTTTTCCACCTTGCCGTCTGCC
K Q I Y V G R A Q K K V E R Q T
(in frame with NES)
K Q I Y V G R A Q K K V E R Q T
(in frame with V5 tag)
PABP
AACTTAAGCGCAAATTTGAACAGATGAAACAAGATAGGATCACCAGATAC
9000
TTGAATTCGCGTTTAAACTTGTCTACTTTGTTCTATCCTAGTGGTCTATG
E L K R K F E Q M K Q D R I T R Y
(in frame with NES)
E L K R K F E Q M K Q D R I T R Y
(in frame with V5 tag)
PABP

FIG. 5AC

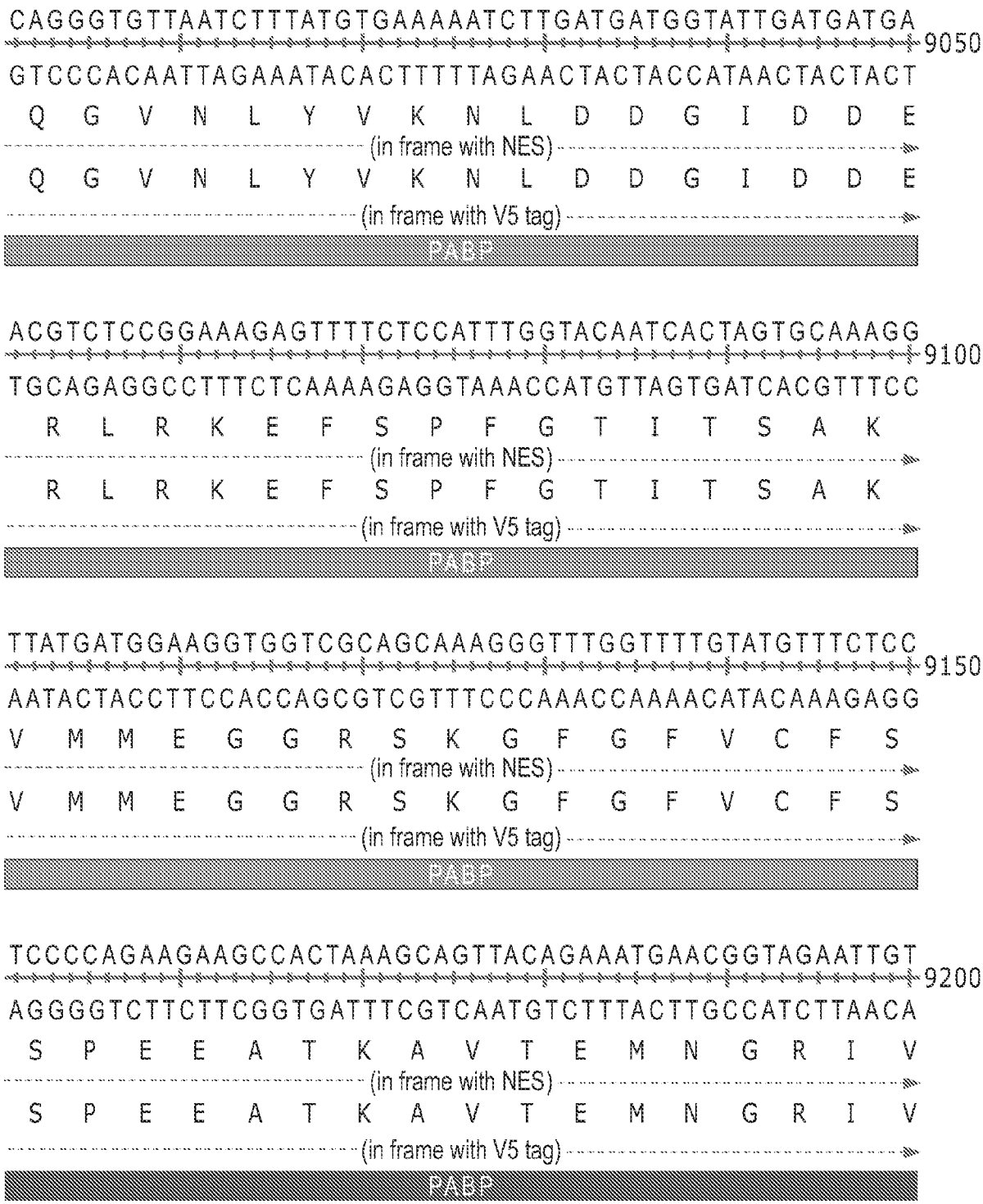

CAGGGTGTTAATCTTTATGTGAAAAATCTTGATGATGGTATTGATGATGA
9050
GTCCCACAATTAGAAATACACTTTTTAGAACTACTACCATAACTACTACT
Q G V N L Y V K N L D D G I D D E
(in frame with NES)
Q G V N L Y V K N L D D G I D D E
(in frame with V5 tag)
PABP
ACGTCTCCGGAAAGAGTTTTCTCCATTTGGTACAATCACTAGTGCAAAGG
9100
TGCAGAGGCCTTTCTCAAAAGAGGTAAACCATGTTAGTGATCACGTTTCC
R L R K E F S P F G T I T S A K
(in frame with NES)
R L R K E F S P F G T I T S A K
(in frame with V5 tag)
PABP
TTATGATGGAAGGTGGTCGCAGCAAAGGGTTTGGTTTTGTATGTTTCTCC
9150
AATACTACCTTCCACCAGCGTCGTTTCCCAAACCAAAACATACAAAGAGG
V M M E G G R S K G F G F V C F S
(in frame with NES)
V M M E G G R S K G F G F V C F S
(in frame with V5 tag)
PABP
TCCCCAGAAGAAGCCACTAAAGCAGTTACAGAAATGAACGGTAGAATTGT
9200
AGGGGTCTTCTTCGGTGATTTCGTCAATGTCTTTACTTGCCATCTTAACA
S P E E A T K A V T E M N G R I V
(in frame with NES)
S P E E A T K A V T E M N G R I V
(in frame with V5 tag)
PABP

FIG. 5AD

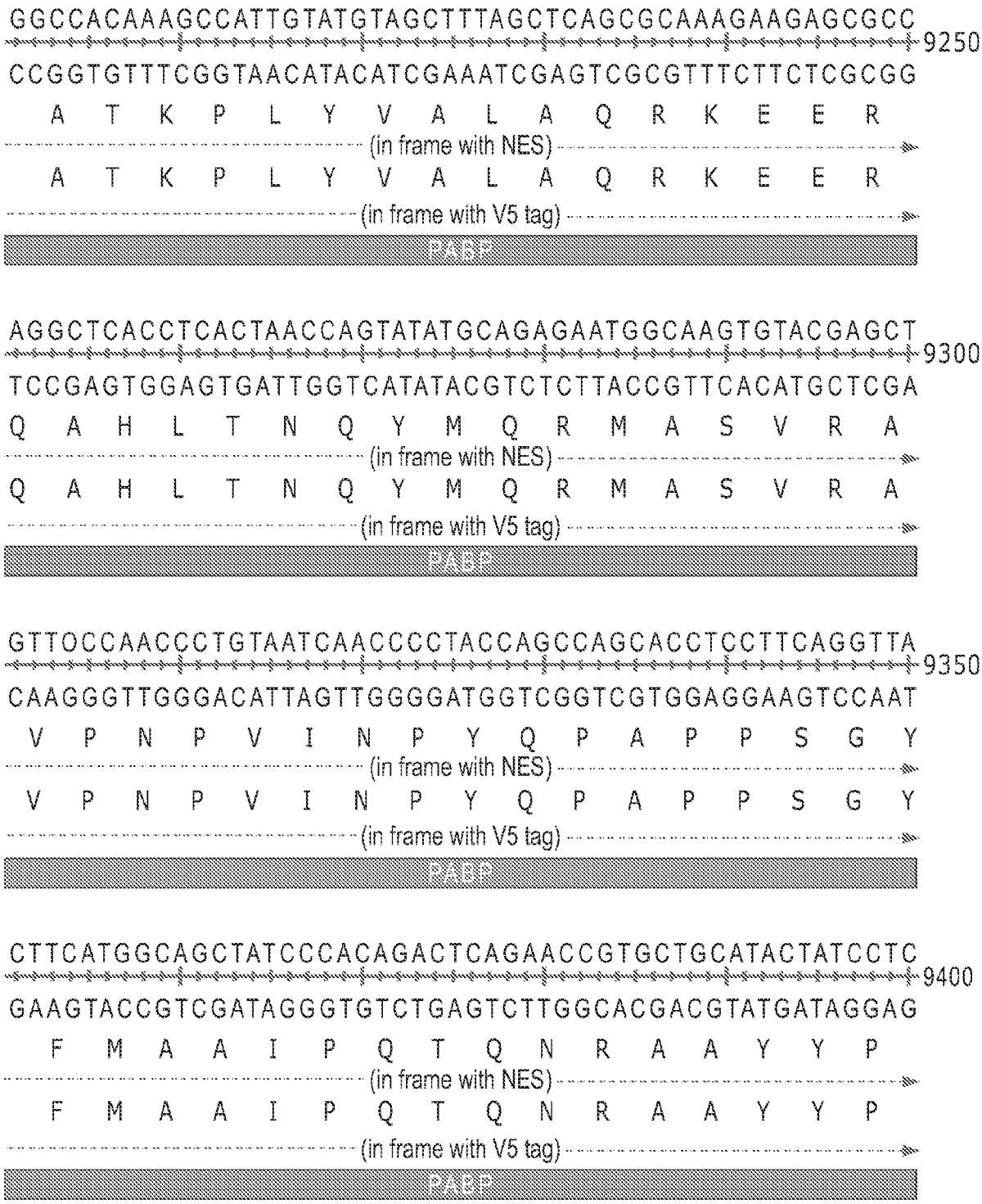

GGCCACAAAGCCATTGTATGTAGCTTTAGCTCAGCGCAAAGAAGAGCGCC
9250
CCGGTGTTTCGGTAACATACATCGAAATCGAGTCGCGTTTCTTCTCGCGG
A T K P L Y V A L A Q R K E E R
(in frame with NES)
A T K P L Y V A L A Q R K E E R
(in frame with V5 tag)
PABP
AGGCTCACCTCACTAACCAGTATATGCAGAGAATGGCAAGTGTACGAGCT
9300
TCCGAGTGGAGTGATTGGTCATATACGTCTCTTACCGTTCACATGCTCGA
Q A H L T N Q Y M Q R M A S V R A
(in frame with NES)
Q A H L T N Q Y M Q R M A S V R A
(in frame with V5 tag)
PABP
GTTCCCAACCCTGTAATCAACCCCTACCAGCCAGCACCTCCTTCAGGTTA
9350
CAAGGGTTGGGACATTAGTTGGGGATGGTCGGTCGTGGAGGAAGTCCAAT
V P N P V I N P Y Q P A P P S G Y
(in frame with NES)
V P N P V I N P Y Q P A P P S G Y
(in frame with V5 tag)
PABP
CTTCATGGCAGCTATCCCACAGACTCAGAACCGTGCTGCATACTATCCTC
9400
GAAGTACCGTCGATAGGGTGTCTGAGTCTTGGCACGACGTATGATAGGAG
F M A A I P Q T Q N R A A Y Y P
(in frame with NES)
F M A A I P Q T Q N R A A Y Y P
(in frame with V5 tag)
PABP

FIG. 5AE

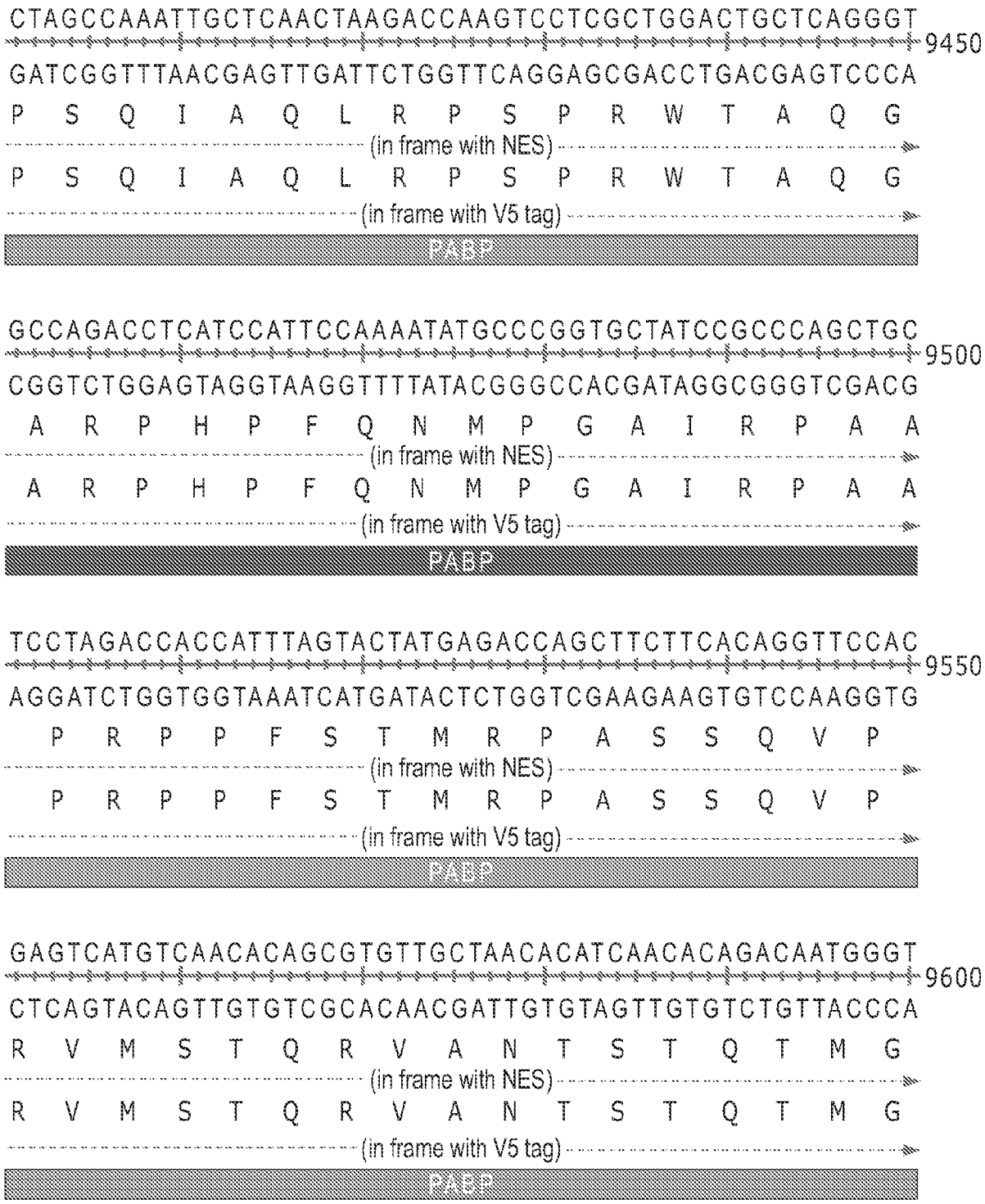

CTAGCCAAATTGCTCAACTAAGACCAAGTCCTCGCTGGACTGCTCAGGGT
9450
GATCGGTTTAACGAGTTGATTCTGGTTCAGGAGCGACCTGACGAGTCCCA
P S Q I A Q L R P S P R W T A Q G
(in frame with NES)
P S Q I A Q L R P S P R W T A Q G
(in frame with V5 tag)
PABP
GCCAGACCTCATCCATTCCAAAATATGCCCGGTGCTATCCGCCCAGCTGC
9500
CGGTCTGGAGTAGGTAAGGTTTTATACGGGCCACGATAGGCGGGTCGACG
A R P H P F Q N M P G A I R P A A
(in frame with NES)
A R P H P F Q N M P G A I R P A A
(in frame with V5 tag)
PABP
TCCTAGACCACCATTTAGTACTATGAGACCAGCTTCTTCACAGGTTCCAC
9550
AGGATCTGGTGGTAAATCATGATACTCTGGTCGAAGAAGTGTCCAAGGTG
P R P P F S T M R P A S S Q V P
(in frame with NES)
P R P P F S T M R P A S S Q V P
(in frame with V5 tag)
PABP
GAGTCATGTCAACACAGCGTGTTGCTAACACATCAACACAGACAATGGGT
9600
CTCAGTACAGTTGTGTCGCACAACGATTGTGTAGTTGTGTCTGTTACCCA
R V M S T Q R V A N T S T Q T M G
(in frame with NES)
R V M S T Q R V A N T S T Q T M G
(in frame with V5 tag)
PABP

FIG. 5AF

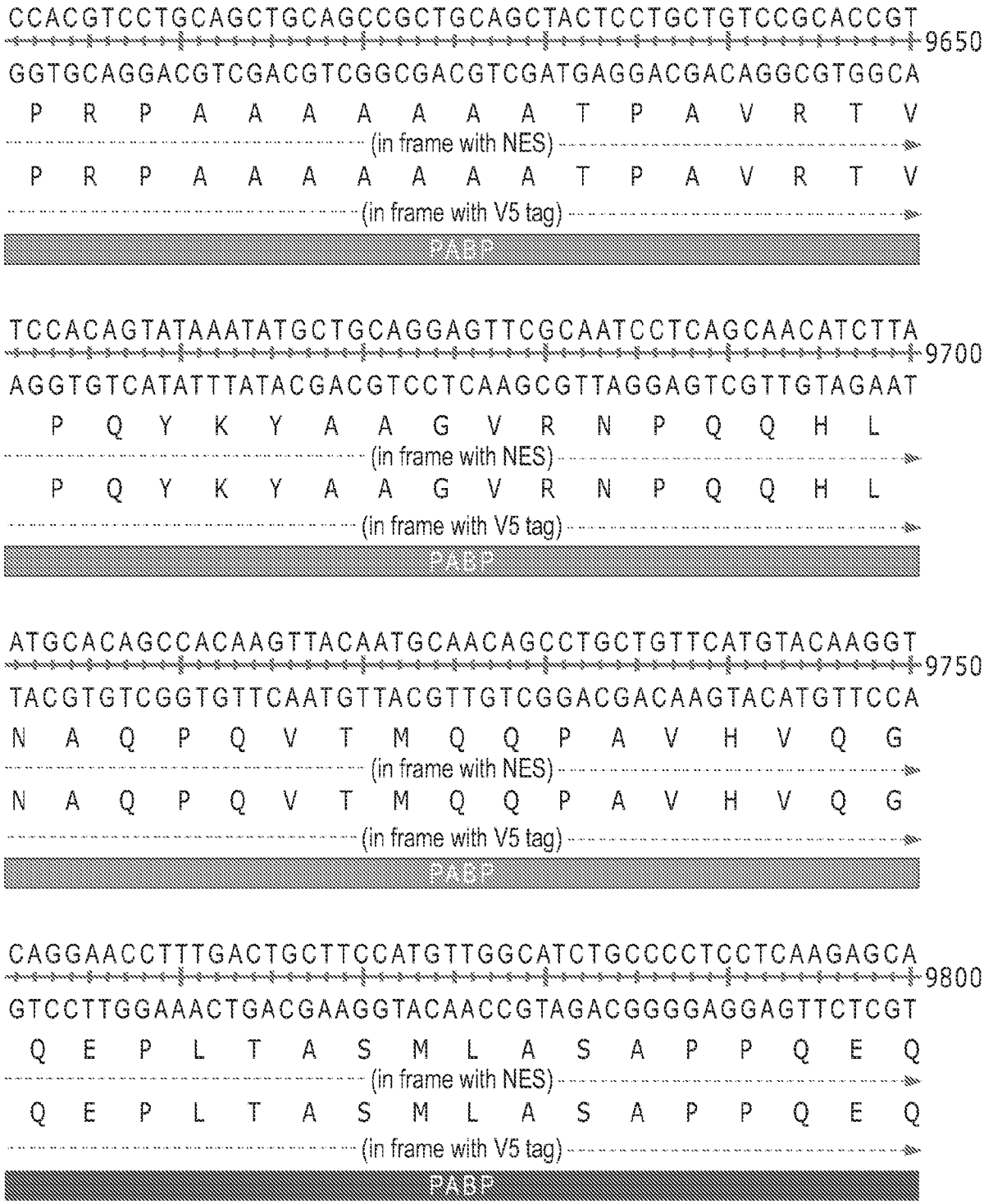

CCACGTCCTGCAGCTGCAGCCGCTGCAGCTACTCCTGCTGTCCGCACCGT
9650
GGTGCAGGACGTCGACGTCGGCGACGTCGATGAGGACGACAGGCGTGGCA
P R P A A A A A A A T P A V R T V
(in frame with NES)
P R P A A A A A A A T P A V R T V
(in frame with V5 tag)
PABP
TCCACAGTATAAATATGCTGCAGGAGTTCGCAATCCTCAGCAACATCTTA
9700
AGGTGTCATATTTATACGACGTCCTCAAGCGTTAGGAGTCGTTGTAGAAT
P Q Y K Y A A G V R N P Q Q H L
(in frame with NES)
P Q Y K Y A A G V R N P Q Q H L
(in frame with V5 tag)
PABP
ATGCACAGCCACAAGTTACAATGCAACAGCCTGCTGTTCATGTACAAGGT
9750
TACGTGTCGGTGTTCAATGTTACGTTGTCGGACGACAAGTACATGTTCCA
N A Q P Q V T M Q Q P A V H V Q G
(in frame with NES)
N A Q P Q V T M Q Q P A V H V Q G
(in frame with V5 tag)
PABP
CAGGAACCTTTGACTGCTTCCATGTTGGCATCTGCCCCTCCTCAAGAGCA
9800
GTCCTTGGAAACTGACGAAGGTACAACCGTAGACGGGGAGGAGTTCTCGT
Q E P L T A S M L A S A P P Q E Q
(in frame with NES)
Q E P L T A S M L A S A P P Q E Q
(in frame with V5 tag)
PABP

FIG. 5AG

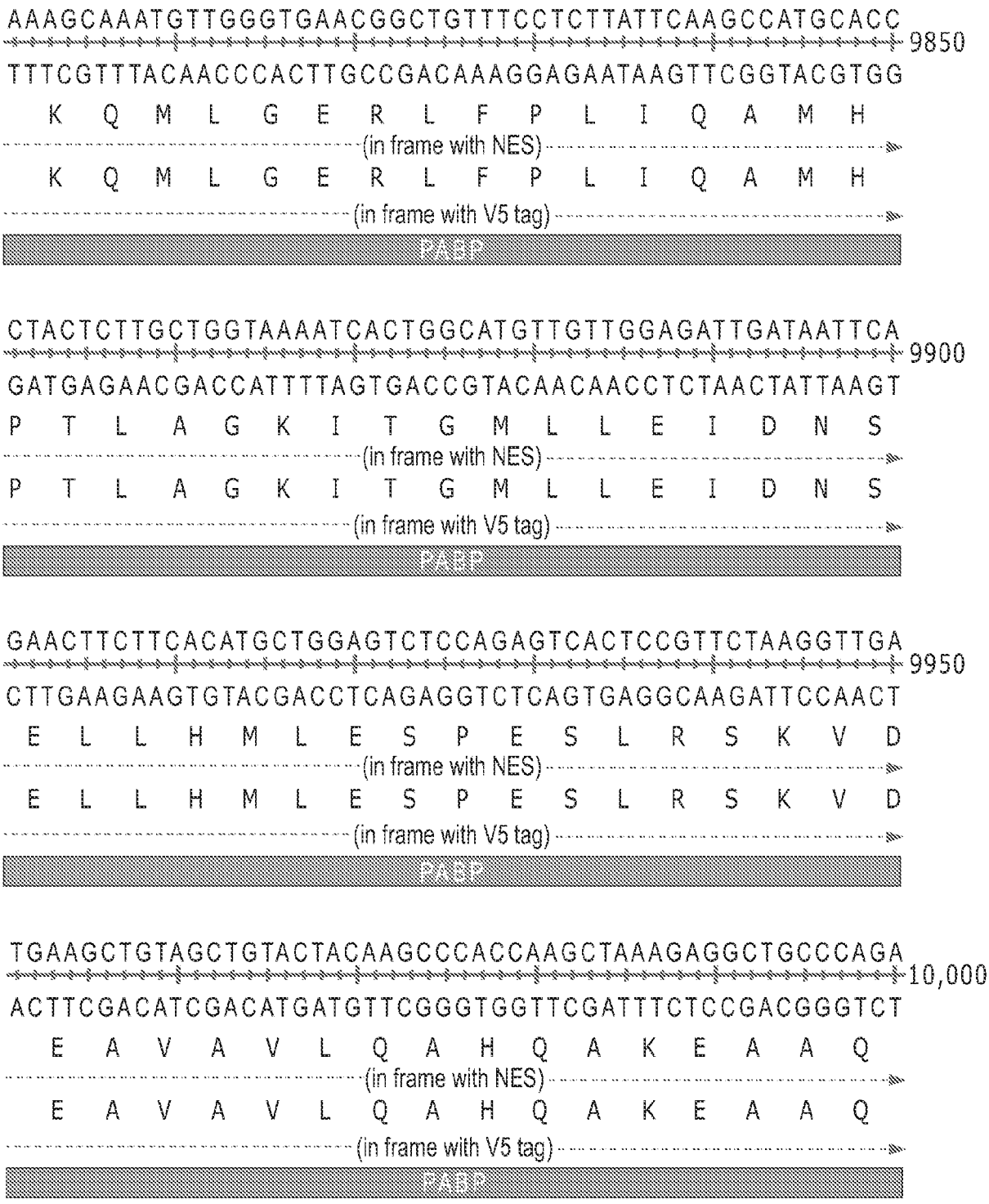

AAAGCAAATGTTGGGTGAACGGCTGTTTCCTCTTATTCAAGCCATGCACC
9850
TTTCGTTTACAACCCACTTGCCGACAAAGGAGAATAAGTTCGGTACGTGG
K Q M L G E R L F P L I Q A M H
(in frame with NES)
K Q M L G E R L F P L I Q A M H
(in frame with V5 tag)
PABP
CTACTCTTGCTGGTAAAATCACTGGCATGTTGTTGGAGATTGATAATTCA
9900
GATGAGAACGACCATTTTAGTGACCGTACAACAACCTCTAACTATTAAGT
P T L A G K I T G M L L E I D N S
(in frame with NES)
P T L A G K I T G M L L E I D N S
(in frame with V5 tag)
PABP
GAACTTCTTCACATGCTGGAGTCTCCAGAGTCACTCCGTTCTAAGGTTGA
9950
CTTGAAGAAGTGTACGACCTCAGAGGTCTCAGTGAGGCAAGATTCCAACT
E L L H M L E S P E S L R S K V D
(in frame with NES)
E L L H M L E S P E S L R S K V D
(in frame with V5 tag)
PABP
TGAAGCTGTAGCTGTACTACAAGCCCACCAAGCTAAAGAGGCTGCCCAGA
10,000
ACTTCGACATCGACATGATGTTCGGGTGGTTCGATTTCTCCGACGGGTCT
E A V A V L Q A H Q A K E A A Q
(in frame with NES)
E A V A V L Q A H Q A K E A A Q
(in frame with V5 tag)
PABP

FIG. 5AH

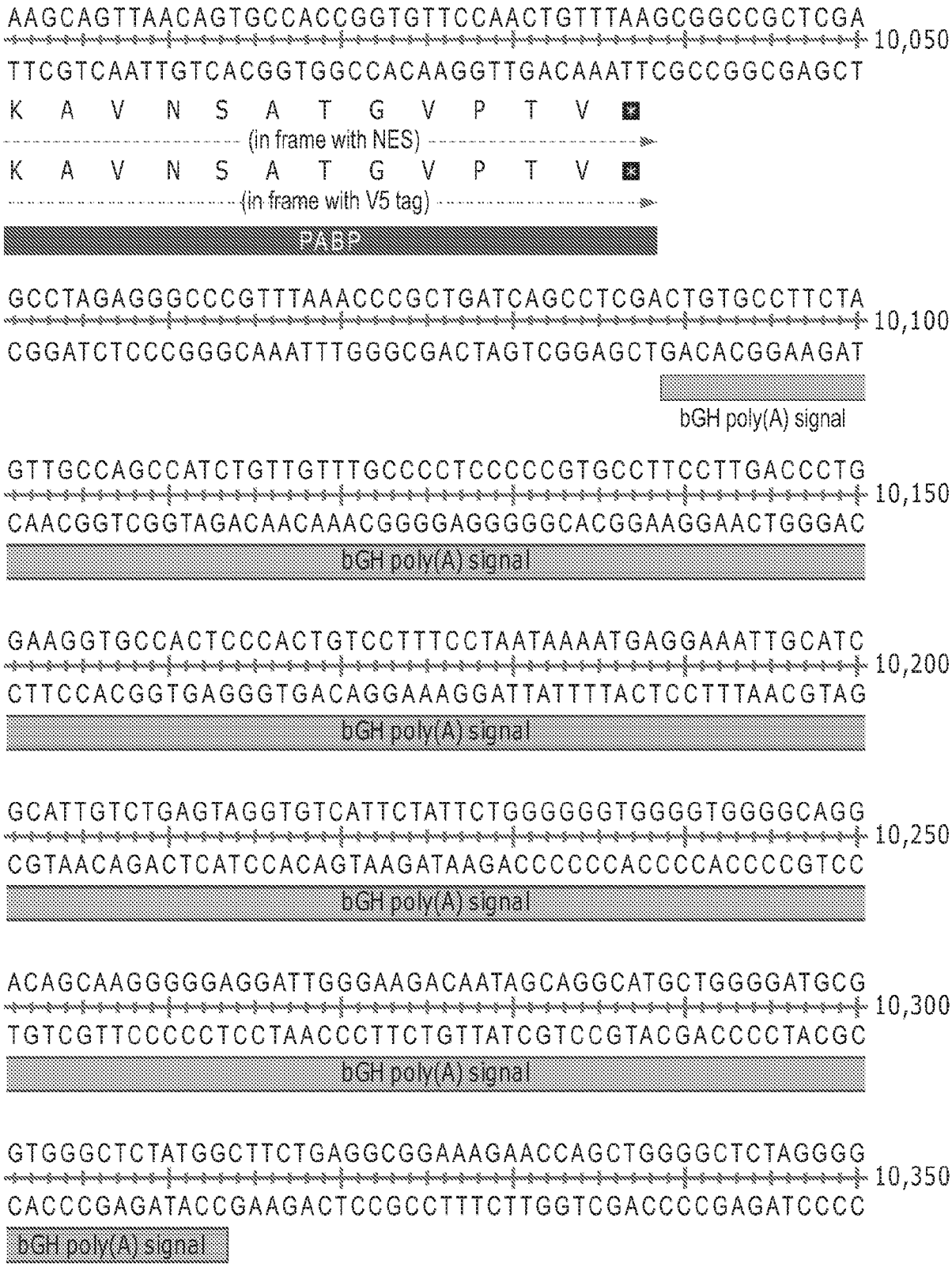

AAGCAGTTAACAGTGCCACCGGTGTTCCAACTGTTTAAGCGGCCGCTCGA
10,050
TTCGTCAATTGTCACGGTGGCCACAAGGTTGACAAATTCGCCGGCGAGCT
K A V N S A T G V P T V
(in frame with NES)
K A V N S A T G V P T V
(in frame with V5 tag)
PABP
GCCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTA
10,100
CGGATCTCCCGGGCAAATTTGGGCGACTAGTCGGAGCTGACACGGAAGAT
bGH poly(A) signal
GTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTG
10,150
CAACGGTCGGTAGACAACAAACGGGGAGGGGGCACGGAAGGAACTGGGAC
bGH poly(A) signal
GAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATC
10,200
CTTCCACGGTGAGGGTGACAGGAAAGGATTATTTTACTCCTTTAACGTAG
bGH poly(A) signal
GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGG
10,250
CGTAACAGACTCATCCACAGTAAGATAAGACCCCCCACCCCACCCCGTCC
bGH poly(A) signal
ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCG
10,300
TGTCGTTCCCCCTCCTAACCCTTCTGTTATCGTCCGTACGACCCCTACGC
bGH poly(A) signal
GTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGG
10,350
CACCCGAGATACCGAAGACTCCGCCTTTCTTGGTCGACCCCGAGATCCCC
bGH poly(A) signal

FIG. 5AI

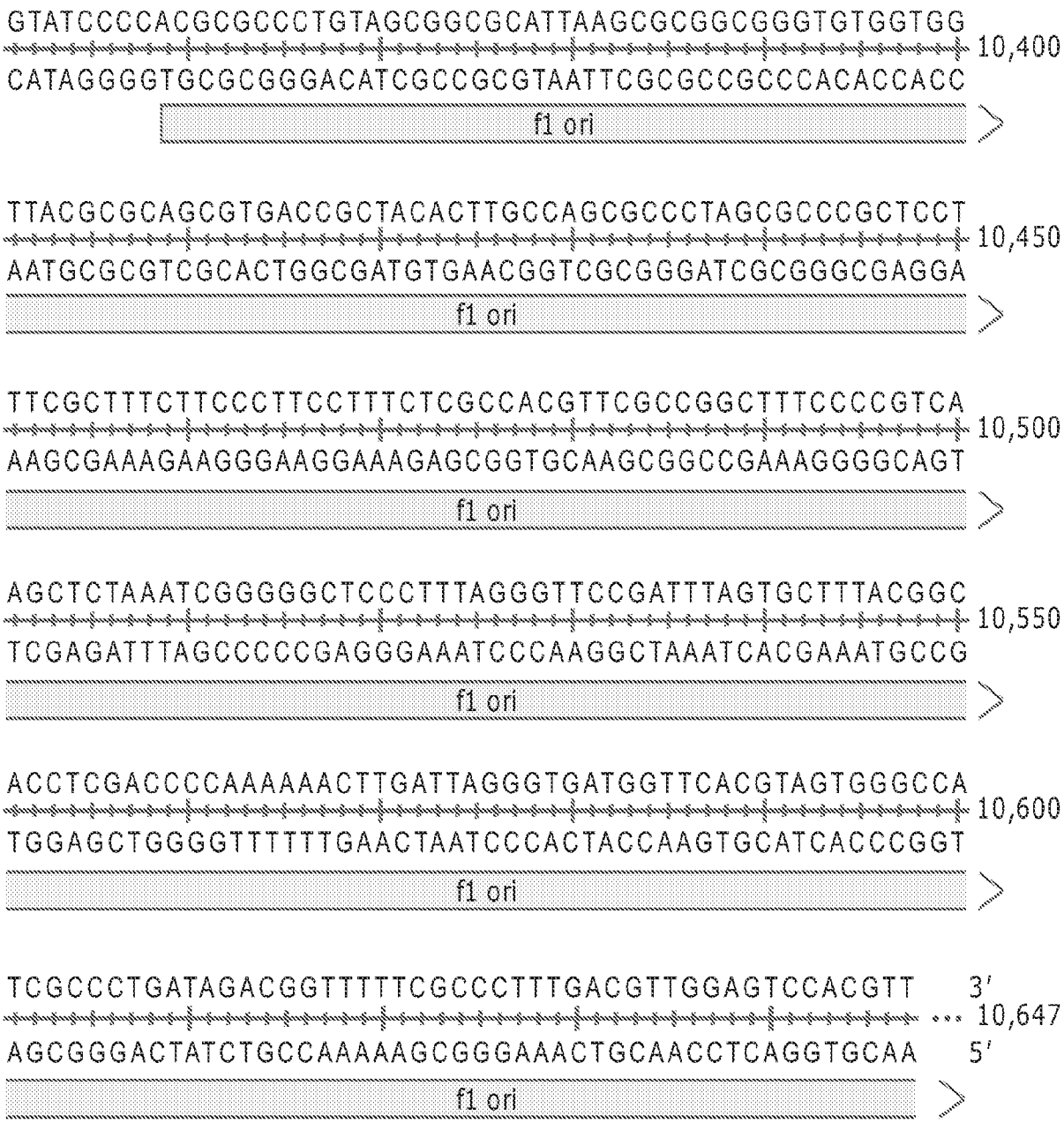

GTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG
10,400
CATAGGGGTGCGCGGGACATCGCCGCGTAATTCGCGCCGCCCACACCACC
f1 ori
TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCT
10,450
AATGCGCGTCGCACTGGCGATGTGAACGGTCGCGGGATCGCGGGCGAGGA
f1 ori
TTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCA
10,500
AAGCGAAAGAAGGGAAGGAAAGAGCGGTGCAAGCGGCCGAAAGGGGCAGT
f1 ori
AGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGC
10,550
TCGAGATTTAGCCCCCGAGGGAAATCCCAAGGCTAAATCACGAAATGCCG
f1 ori
ACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCA
10,600
TGGAGCTGGGGTTTTTTGAACTAATCCCACTACCAAGTGCATCACCCGGT
f1 ori
TCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTT 3′
... 10,647
AGCGGGACTATCTGCCAAAAAGCGGGAAACTGCAACCTCAGGTGCAA 5′
f1 ori

FIG. 5AJ

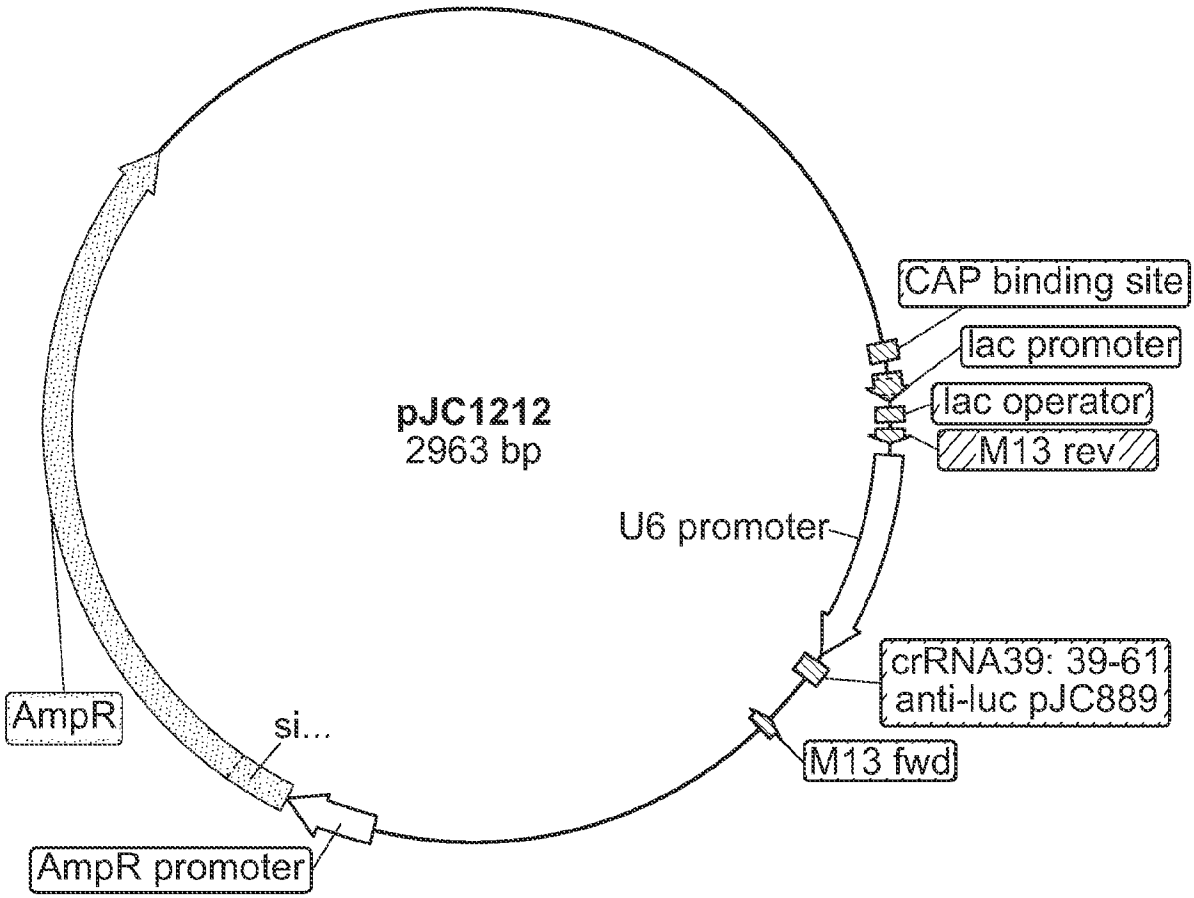

pJC1212
2963 bp
CAP binding site
lac promoter
lac operator
M13 rev
U6 promoter
crRNA39: 39-61
anti-luc pJC889
M13 fwd
AmpR
si...
AmpR promoter

FIG. 6

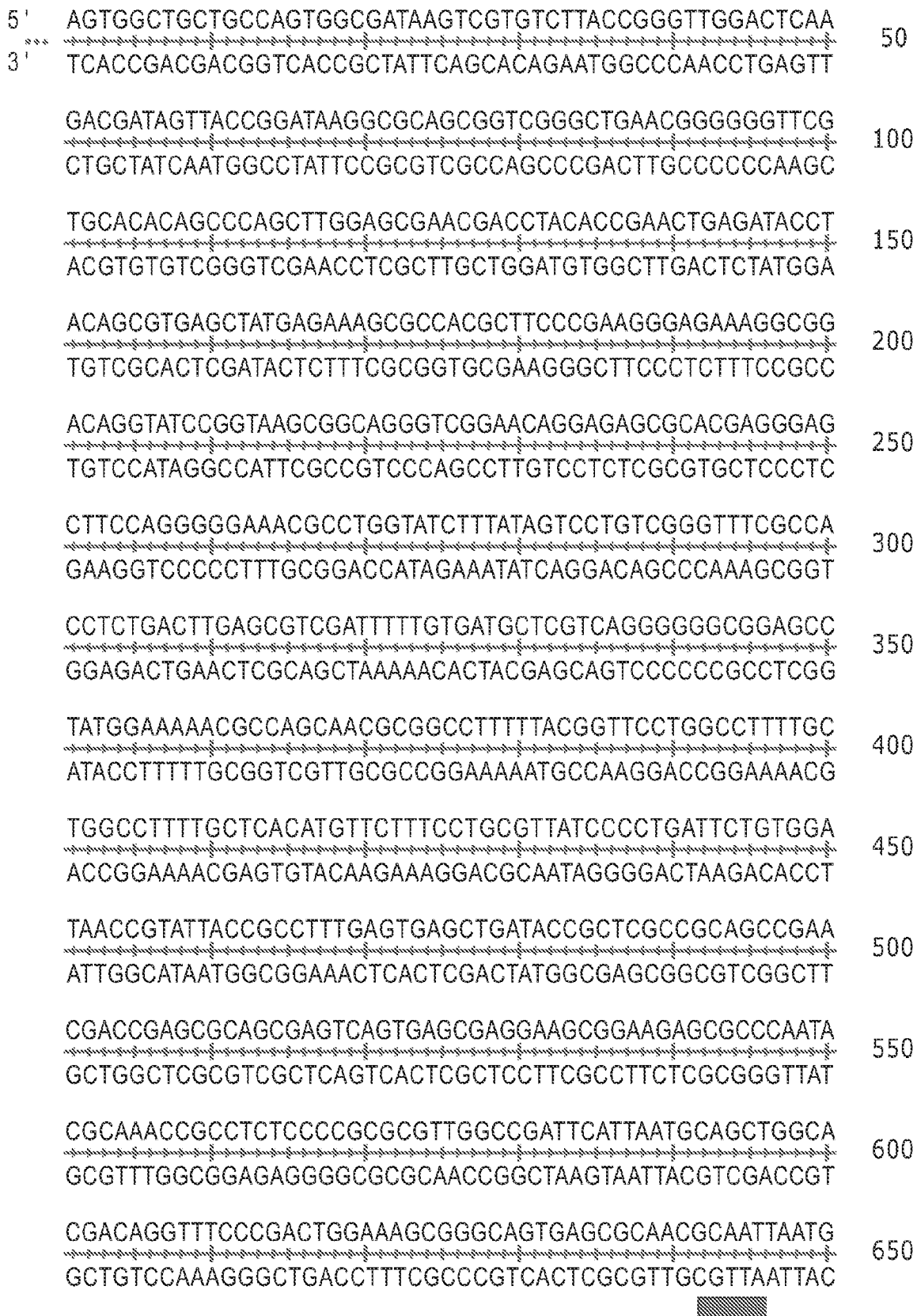

5'
3'
AGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAA
TCACCGACGACGGTCACCGCTATTCAGCACAGAATGGCCCAACCTGAGTT
50
GACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCG
CTGCTATCAATGGCCTATTCCGCGTCGCCAGCCCGACTTGCCCCCCAAGC
100
TGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACGTGTGTCGGGTCGAACCTCGCTTGCTGGATGTGGCTTGACTCTATGGA
150
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGG
TGTCGCACTCGATACTCTTTCGCGGTGCGAAGGGCTTCCCTCTTTCCGCC
200
ACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAG
TGTCCATAGGCCATTCGCCGTCCCAGCCTTGTCCTCTCGCGTGCTCCCTC
250
CTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCA
GAAGGTCCCCCTTTGCGGACCATAGAAATATCAGGACAGCCCAAAGCGGT
300
CCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCC
GGAGACTGAACTCGCAGCTAAAAACACTACGAGCAGTCCCCCCGCCTCGG
350
TATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGC
ATACCTTTTTGCGGTCGTTGCGCCGGAAAAATGCCAAGGACCGGAAAACG
400
TGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGA
ACCGGAAAACGAGTGTACAAGAAAGGACGCAATAGGGGACTAAGACACCT
450
TAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAA
ATTGGCATAATGGCGGAAACTCACTCGACTATGGCGAGCGGCGTCGGCTT
500
CGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATA
GCTGGCTCGCGTCGCTCAGTCACTCGCTCCTTCGCCTTCTCGCGGGTTAT
550
CGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
GCGTTTGGCGGAGAGGGGCGCGCAACCGGCTAAGTAATTACGTCGACCGT
600
CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
GCTGTCCAAAGGGCTGACCTTTCGCCCGTCACTCGCGTTGCGTTAATTAC
650
CAP binding site

FIG. 7A

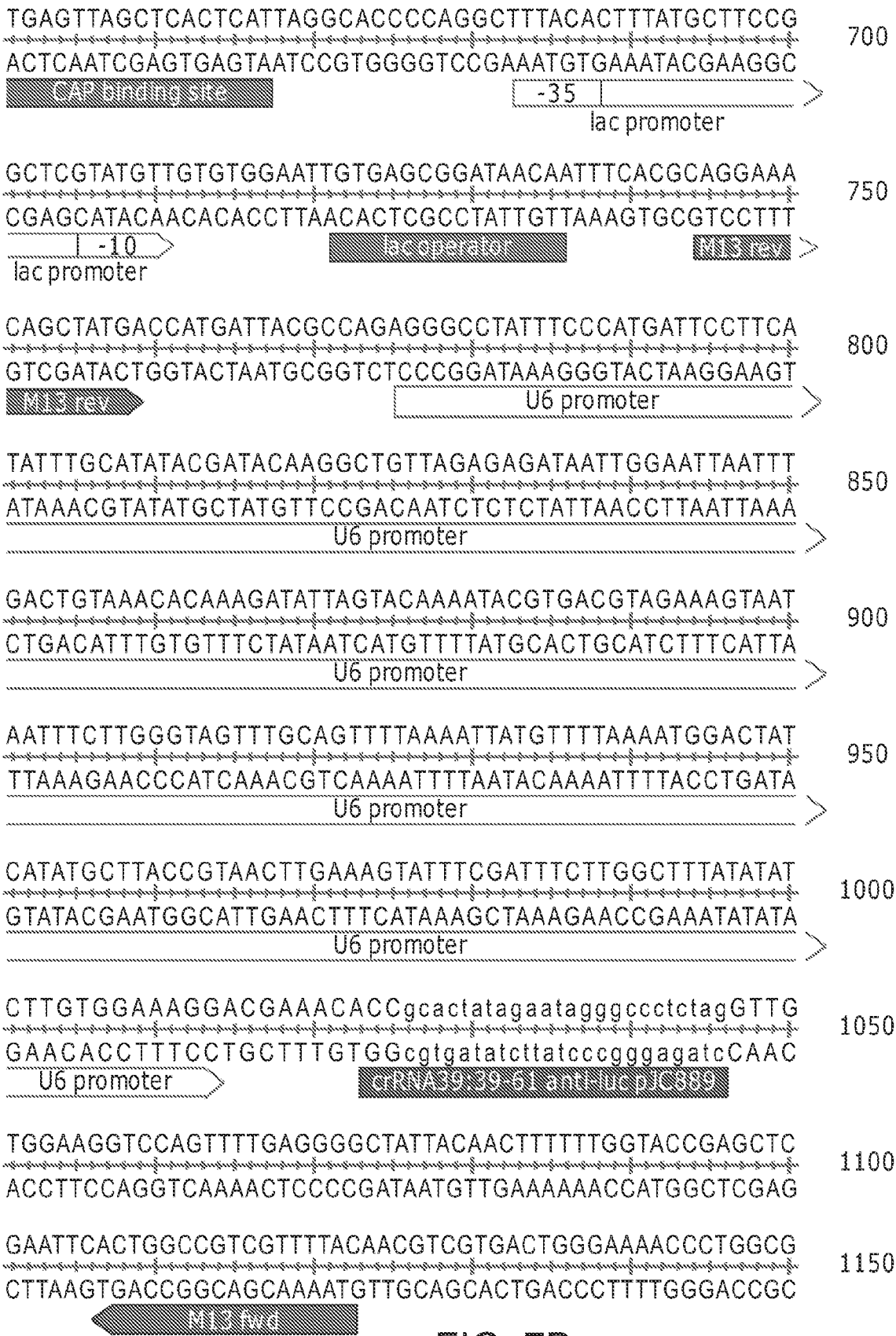

TGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCG
ACTCAATCGAGTGAGTAATCCGTGGGGTCCGAAATGTGAAATACGAAGGC
700
CAP binding site
-35
lac promoter
GCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACGCAGGAAA
CGAGCATACAACACACCTTAACACTCGCCTATTGTTAAAGTGCGTCCTTT
750
-10
lac promoter
lac operator
M13 rev
CAGCTATGACCATGATTACGCCAGAGGGCCTATTTCCCATGATTCCTTCA
GTCGATACTGGTACTAATGCGGTCTCCCGGATAAAGGGTACTAAGGAAGT
800
M13 rev
U6 promoter
TATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTT
ATAAACGTATATGCTATGTTCCGACAATCTCTCTATTAACCTTAATTAAA
850
U6 promoter
GACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAAT
CTGACATTTGTGTTTCTATAATCATGTTTTATGCACTGCATCTTTCATTA
900
U6 promoter
AATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTAT
TTAAAGAACCCATCAAACGTCAAAATTTTAATACAAAATTTTACCTGATA
950
U6 promoter
CATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATAT
GTATACGAATGGCATTGAACTTTCATAAAGCTAAAGAACCGAAATATATA
1000
U6 promoter
CTTGTGGAAAGGACGAAACACCgcactatagaatagggccctctagGTTG
GAACACCTTTCCTGCTTTGTGGcgtgatatcttatcccgggagatcCAAC
1050
U6 promoter
crRNA 39-39-61 anti-UC pJC389
TGGAAGGTCCAGTTTTGAGGGGCTATTACAACTTTTTTGGTACCGAGCTC
ACCTTCCAGGTCAAAACTCCCCGATAATGTTGAAAAAACCATGGCTCGAG
1100
GAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCG
CTTAAGTGACCGGCAGCAAAATGTTGCAGCACTGACCCTTTTGGGACCGC
1150
M13 fwd

FIG. 7B

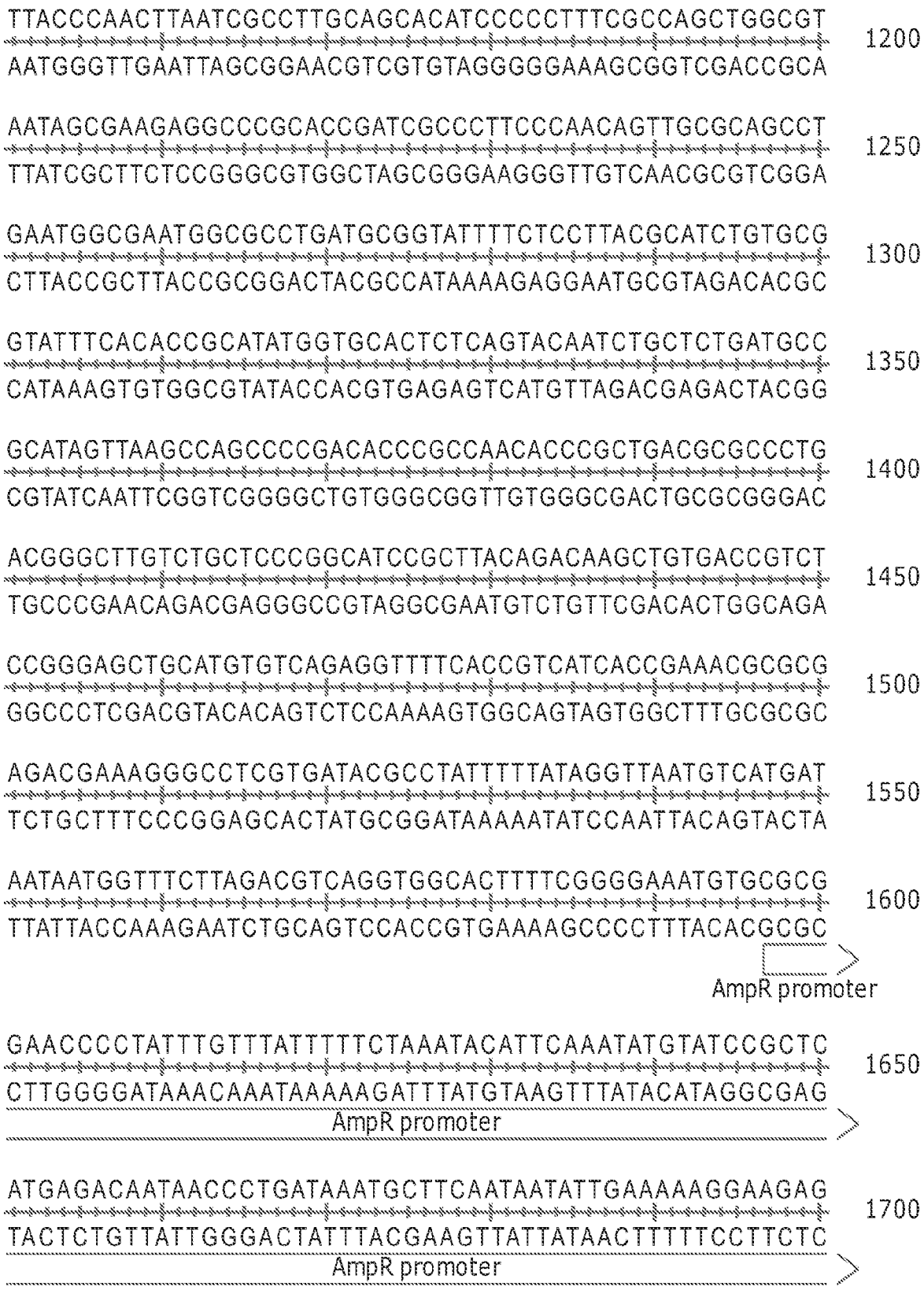

TTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGT
AATGGGTTGAATTAGCGGAACGTCGTGTAGGGGGAAAGCGGTCGACCGCA
1200
AATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCT
TTATCGCTTCTCCGGGCGTGGCTAGCGGGAAGGGTTGTCAACGCGTCGGA
1250
GAATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCG
CTTACCGCTTACCGCGGACTACGCCATAAAAGAGGAATGCGTAGACACGC
1300
GTATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCC
CATAAAGTGTGGCGTATACCACGTGAGAGTCATGTTAGACGAGACTACGG
1350
GCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTG
CGTATCAATTCGGTCGGGGCTGTGGGCGGTTGTGGGCGACTGCGCGGGAC
1400
ACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCT
TGCCCGAACAGACGAGGGCCGTAGGCGAATGTCTGTTCGACACTGGCAGA
1450
CCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCG
GGCCCTCGACGTACACAGTCTCCAAAAGTGGCAGTAGTGGCTTTGCGCGC
1500
AGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGAT
TCTGCTTTCCCGGAGCACTATGCGGATAAAAATATCCAATTACAGTACTA
1550
AATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCG
TTATTACCAAAGAATCTGCAGTCCACCGTGAAAAGCCCCTTTACACGCGC
1600
AmpR promoter
GAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC
CTTGGGGATAAACAAATAAAAAGATTTATGTAAGTTTATACATAGGCGAG
1650
AmpR promoter
ATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAG
TACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTC
1700
AmpR promoter

FIG. 7C

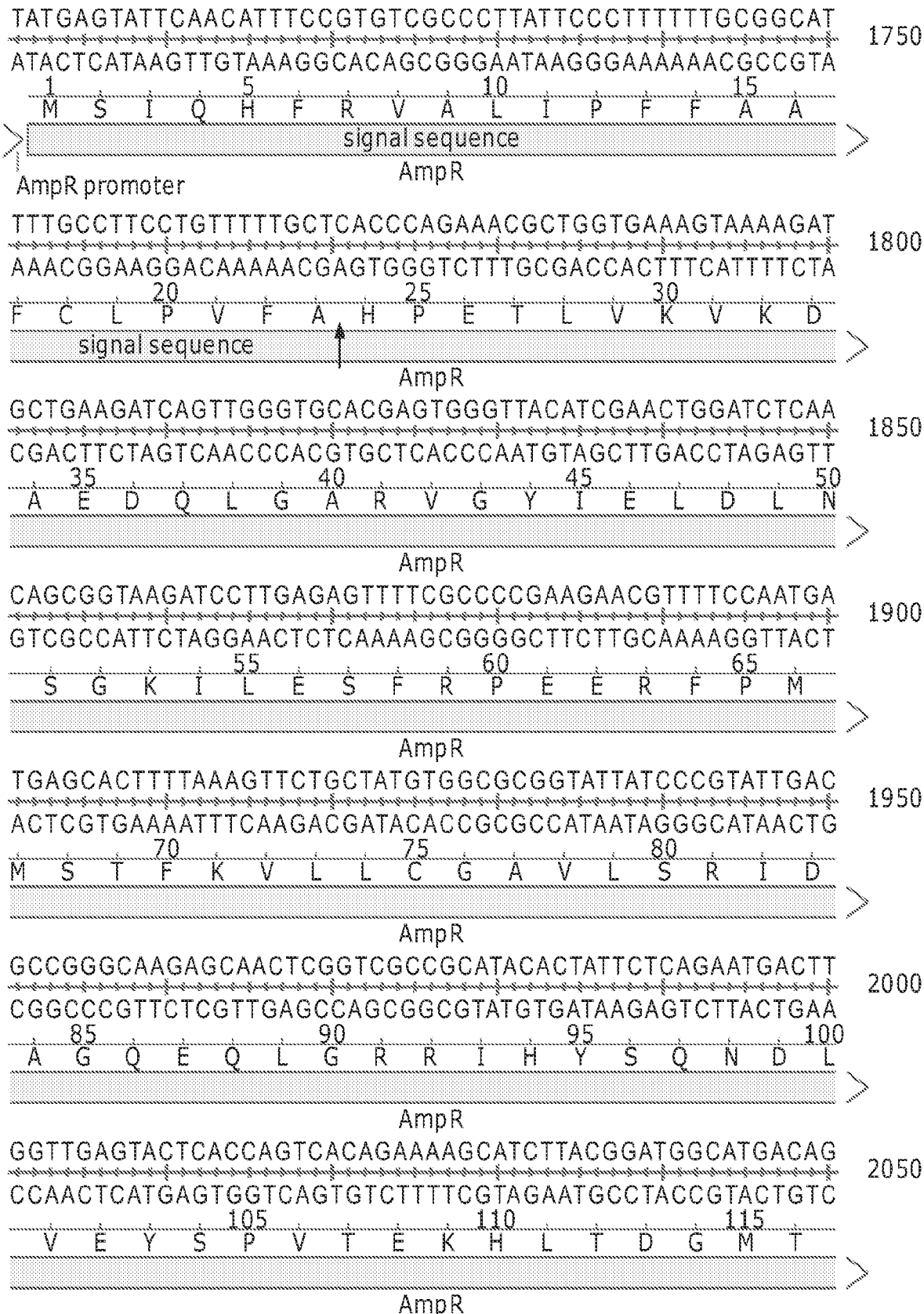

TATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCAT
1750
ATACTCATAAGTTGTAAAGGCACAGCGGGAATAAGGGAAAAAACGCCGTA
1 5 10 15
M S I Q H F R V A L I P F F A A
signal sequence
AmpR
AmpR promoter
TTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGAT
1800
AAACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTCATTTTCTA
20 25 30
F C L P V F A H P E T L V K V K D
signal sequence
AmpR
GCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAA
1850
CGACTTCTAGTCAACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTT
35 40 45 50
A E D Q L G A R V G Y I E L D L N
AmpR
CAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGA
1900
GTCGCCATTCTAGGAACTCTCAAAAGCGGGGCTTCTTGCAAAAGGTTACT
55 60 65
S G K I L E S F R P E E R F P M
AmpR
TGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGAC
1950
ACTCGTGAAAATTTCAAGACGATACACCGCGCCATAATAGGGCATAACTG
70 75 80
M S T F K V L L C G A V L S R I D
AmpR
GCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTT
2000
CGGCCCGTTCTCGTTGAGCCAGCGGCGTATGTGATAAGAGTCTTACTGAA
85 90 95 100
A G Q E Q L G R R I H Y S Q N D L
AmpR
GGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAG
2050
CCAACTCATGAGTGGTCAGTGTCTTTTCGTAGAATGCCTACCGTACTGTC
105 110 115
V E Y S P V T E K H L T D G M T
AmpR

FIG. 7D

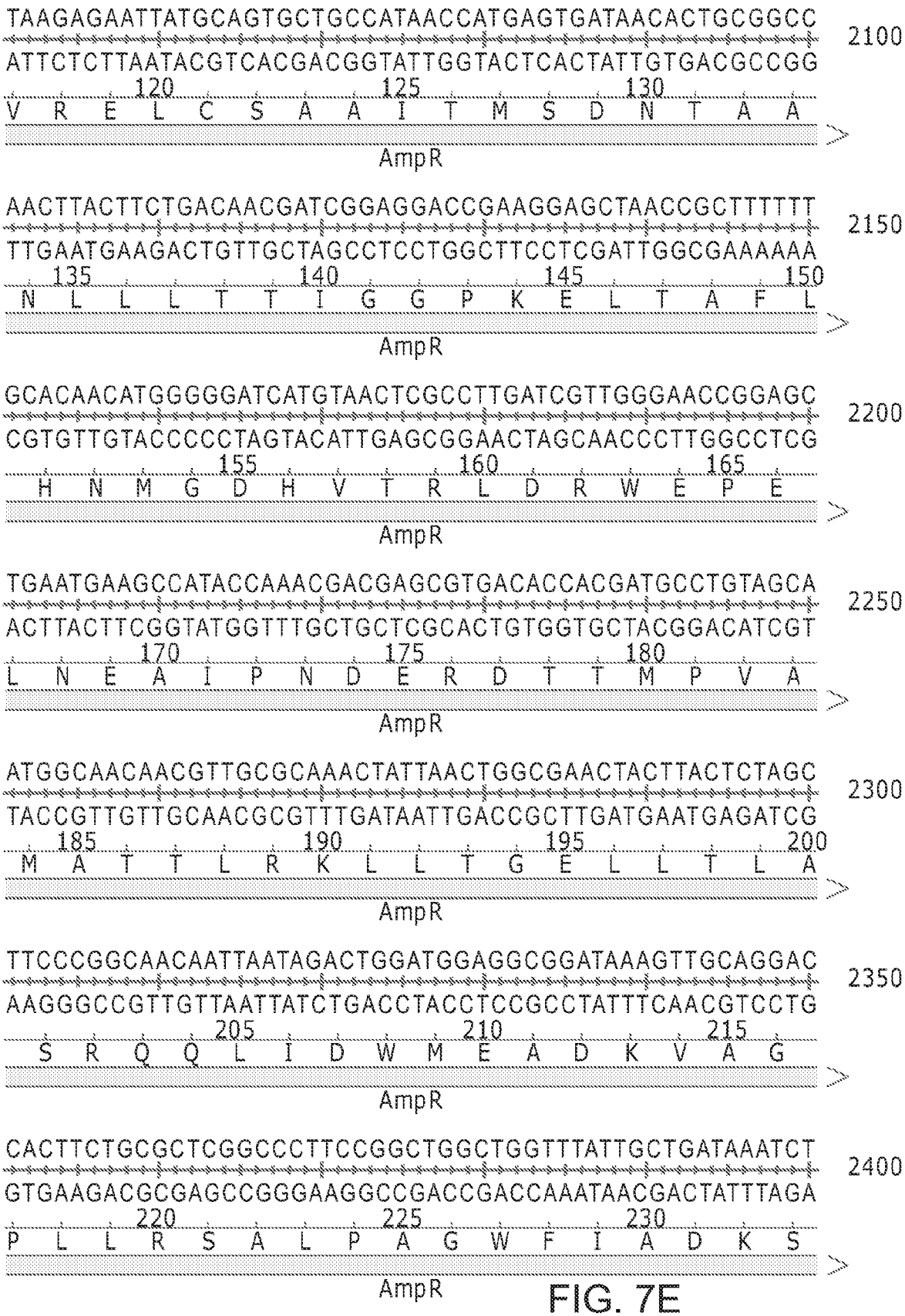

TAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCC
2100
ATTCTCTTAATACGTCACGACGGTATTGGTACTCACTATTGTGACGCCGG
120 125 130
V R E L C S A A I T M S D N T A A
AmpR
AACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTT
2150
TTGAATGAAGACTGTTGCTAGCCTCCTGGCTTCCTCGATTGGCGAAAAAA
135 140 145 150
N L L L T T I G G P K E L T A F L
AmpR
GCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGC
2200
CGTGTTGTACCCCCTAGTACATTGAGCGGAACTAGCAACCCTTGGCCTCG
155 160 165
H N M G D H V T R L D R W E P E
AmpR
TGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCA
2250
ACTTACTTCGGTATGGTTTGCTGCTCGCACTGTGGTGCTACGGACATCGT
170 175 180
L N E A I P N D E R D T T M P V A
AmpR
ATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGC
2300
TACCGTTGTTGCAACGCGTTTGATAATTGACCGCTTGATGAATGAGATCG
185 190 195 200
M A T T L R K L L T G E L L T L A
AmpR
TTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGAC
2350
AAGGGCCGTTGTTAATTATCTGACCTACCTCCGCCTATTTCAACGTCCTG
205 210 215
S R Q Q L I D W M E A D K V A G
AmpR
CACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCT
2400
GTGAAGACGCGAGCCGGGAAGGCCGACCGACCAAATAACGACTATTTAGA
220 225 230
P L L R S A L P A G W F I A D K S
AmpR

FIG. 7E

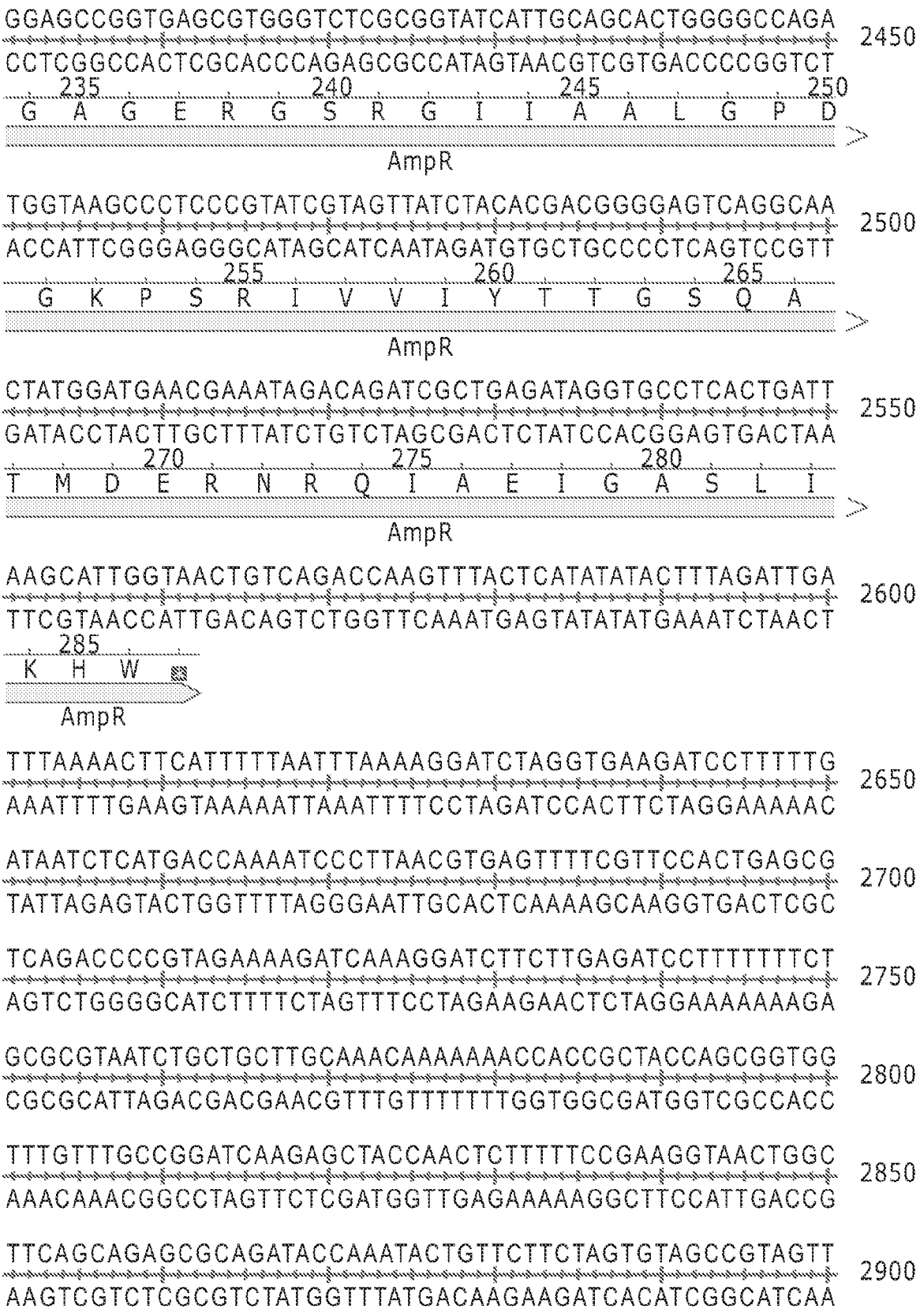

GGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGA
2450
CCTCGGCCACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCT
235 240 245 250
G A G E R G S R G I I A A L G P D
AmpR
TGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAA
2500
ACCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTT
255 260 265
G K P S R I V V I Y T T G S Q A
AmpR
CTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATT
2550
GATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAA
270 275 280
T M D E R N R Q I A E I G A S L I
AmpR
AAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGA
2600
TTCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACT
285
K H W
AmpR
TTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTG
2650
AAATTTTGAAGTAAAAATTAAATTTTCCTAGATCCACTTCTAGGAAAAAC
ATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCG
2700
TATTAGAGTACTGGTTTTAGGGAATTGCACTCAAAAGCAAGGTGACTCGC
TCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCT
2750
AGTCTGGGGCATCTTTTCTAGTTTCCTAGAAGAACTCTAGGAAAAAAAGA
GCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGG
2800
CGCGCATTAGACGACGAACGTTTGTTTTTTTGGTGGCGATGGTCGCCACC
TTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGC
2850
AAACAAACGGCCTAGTTCTCGATGGTTGAGAAAAAGGCTTCCATTGACCG
TTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTT
2900
AAGTCGTCTCGCGTCTATGGTTTATGACAAGAAGATCACATCGGCATCAA

FIG. 7F

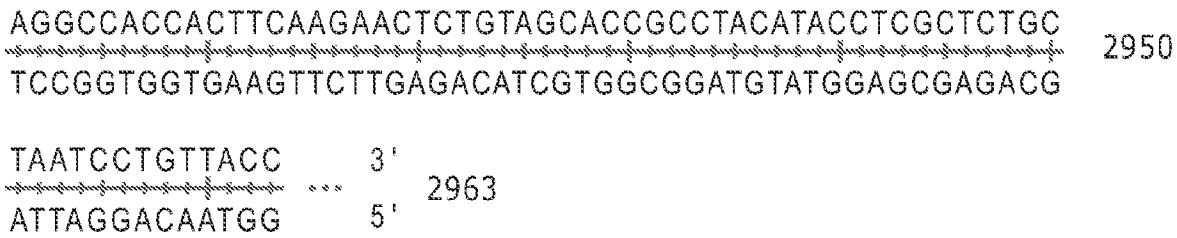

AGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGC
2950
TCCGGTGGTGAAGTTCTTGAGACATCGTGGCGGATGTATGGAGCGAGACG
TAATCCTGTTACC
3'
2963
ATTAGGACAATGG
5'

FIG. 7G

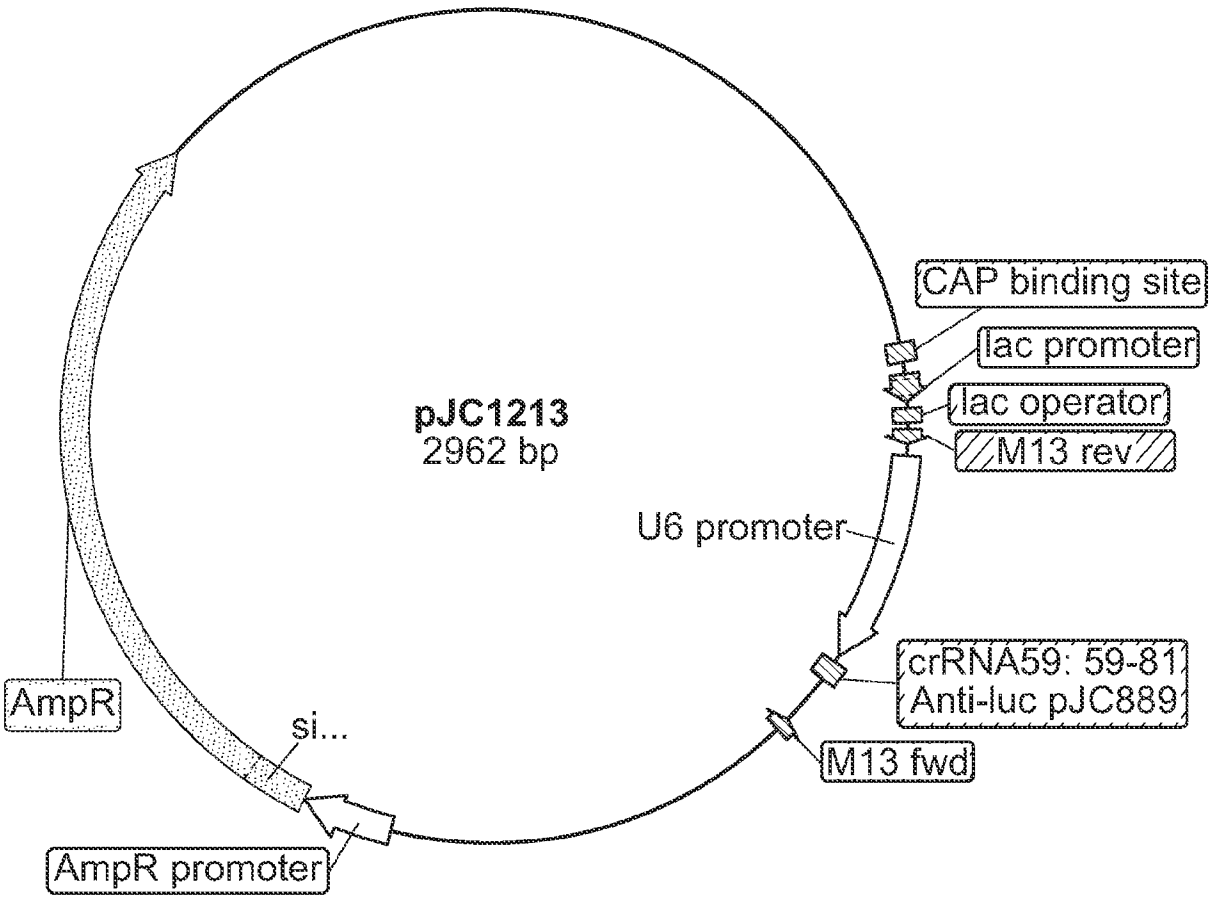

pJC1213
2962 bp
CAP binding site
lac promoter
lac operator
M13 rev
U6 promoter
crRNA59: 59-81
Anti-luc pJC889
M13 fwd
si...
AmpR
AmpR promoter

FIG. 8


5' AGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAA
3' TCACCGACGACGGTCACCGCTATTCAGCACAGAATGGCCCAACCTGAGTT
50
GACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCG
CTGCTATCAATGGCCTATTCCGCGTCGCCAGCCCGACTTGCCCCCCAAGC
100
TGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACGTGTGTCGGGTCGAACCTCGCTTGCTGGATGTGGCTTGACTCTATGGA
150
ACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGG
TGTCGCACTCGATACTCTTTCGCGGTGCGAAGGGCTTCCCTCTTTCCGCC
200
ACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAG
TGTCCATAGGCCATTCGCCGTCCCAGCCTTGTCCTCTCGCGTGCTCCCTC
250
CTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCA
GAAGGTCCCCCTTTGCGGACCATAGAAATATCAGGACAGCCCAAAGCGGT
300
CCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCC
GGAGACTGAACTCGCAGCTAAAAACACTACGAGCAGTCCCCCCGCCTCGG
350
TATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGC
ATACCTTTTTGCGGTCGTTGCGCCGGAAAAATGCCAAGGACCGGAAAACG
400
TGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGA
ACCGGAAAACGAGTGTACAAGAAAGGACGCAATAGGGGACTAAGACACCT
450
TAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAA
ATTGGCATAATGGCGGAAACTCACTCGACTATGGCGAGCGGCGTCGGCTT
500
CGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATA
GCTGGCTCGCGTCGCTCAGTCACTCGCTCCTTCGCCTTCTCGCGGGTTAT
550
CGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCA
GCGTTTGGCGGAGAGGGGCGCGCAACCGGCTAAGTAATTACGTCGACCGT
600
CGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATG
GCTGTCCAAAGGGCTGACCTTTCGCCCGTCACTCGCGTTGCGTTAATTAC
650
CAP binding site

FIG. 9A

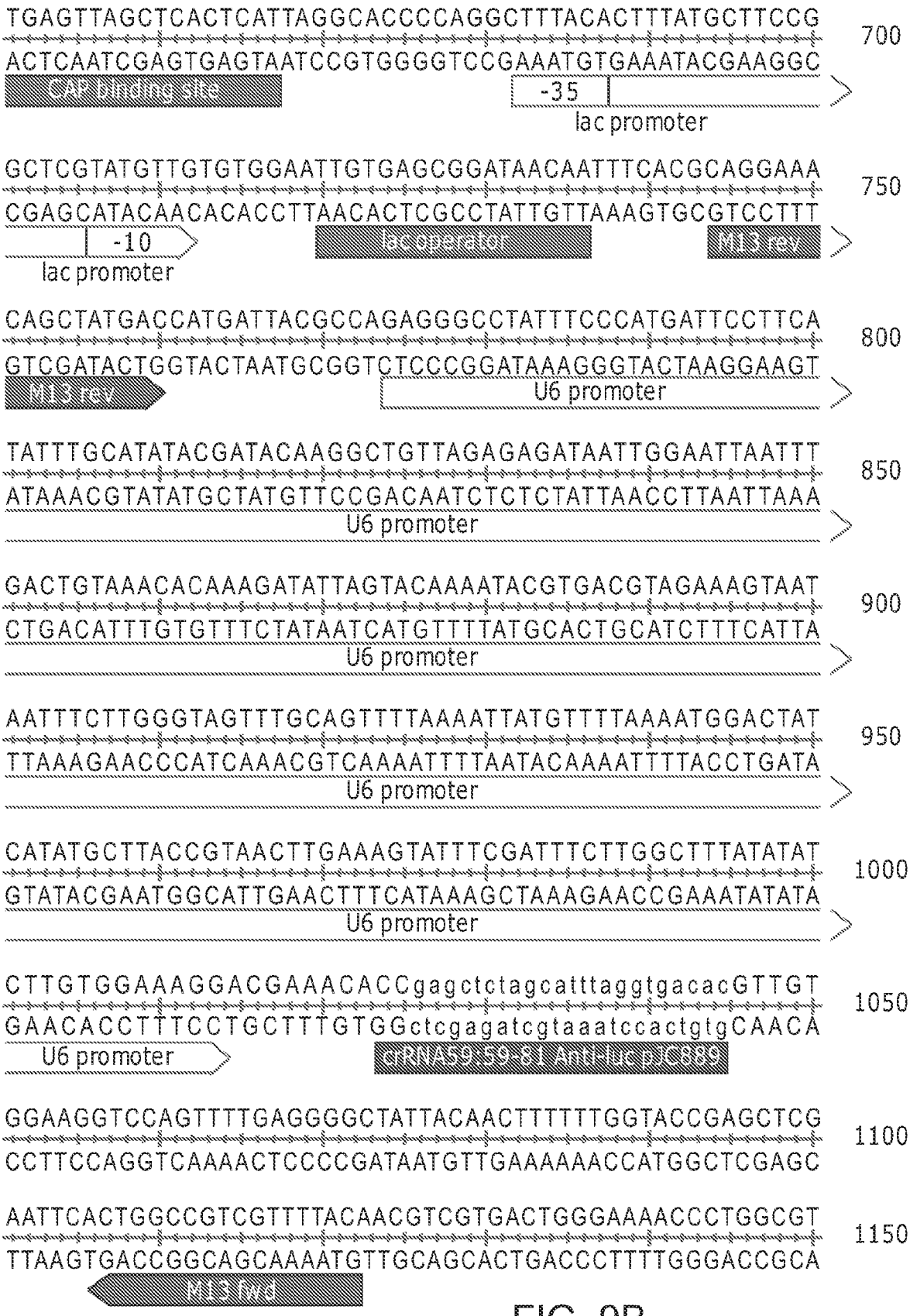

TGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCG
700
ACTCAATCGAGTGAGTAATCCGTGGGGTCCGAAATGTGAAATACGAAGGC
CAP binding site
-35
lac promoter
GCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACGCAGGAAA
750
CGAGCATACAACACACCTTAACACTCGCCTATTGTTAAAGTGCGTCCTTT
-10
lac promoter
lac operator
M13 rev
CAGCTATGACCATGATTACGCCAGAGGGCCTATTTCCCATGATTCCTTCA
800
GTCGATACTGGTACTAATGCGGTCTCCCGGATAAAGGGTACTAAGGAAGT
M13 rev
U6 promoter
TATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTGGAATTAATTT
850
ATAAACGTATATGCTATGTTCCGACAATCTCTCTATTAACCTTAATTAAA
U6 promoter
GACTGTAAACACAAAGATATTAGTACAAAATACGTGACGTAGAAAGTAAT
900
CTGACATTTGTGTTTCTATAATCATGTTTTATGCACTGCATCTTTCATTA
U6 promoter
AATTTCTTGGGTAGTTTGCAGTTTTAAAATTATGTTTTAAAATGGACTAT
950
TTAAAGAACCCATCAAACGTCAAAATTTTAATACAAAATTTTACCTGATA
U6 promoter
CATATGCTTACCGTAACTTGAAAGTATTTCGATTTCTTGGCTTTATATAT
1000
GTATACGAATGGCATTGAACTTTCATAAAGCTAAAGAACCGAAATATATA
U6 promoter
CTTGTGGAAAGGACGAAACACCgagctctagcatttaggtgacacGTTGT
1050
GAACACCTTTCCTGCTTTGTGGctcgagatcgtaaatccactgtgCAACA
U6 promoter
crRNA59.59-81 AntiFluc pJC889
GGAAGGTCCAGTTTTGAGGGGCTATTACAACTTTTTTGGTACCGAGCTCG
1100
CCTTCCAGGTCAAAACTCCCCGATAATGTTGAAAAAACCATGGCTCGAGC
AATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGT
1150
TTAAGTGACCGGCAGCAAAATGTTGCAGCACTGACCCTTTTGGGACCGCA
M13 fwd

FIG. 9B

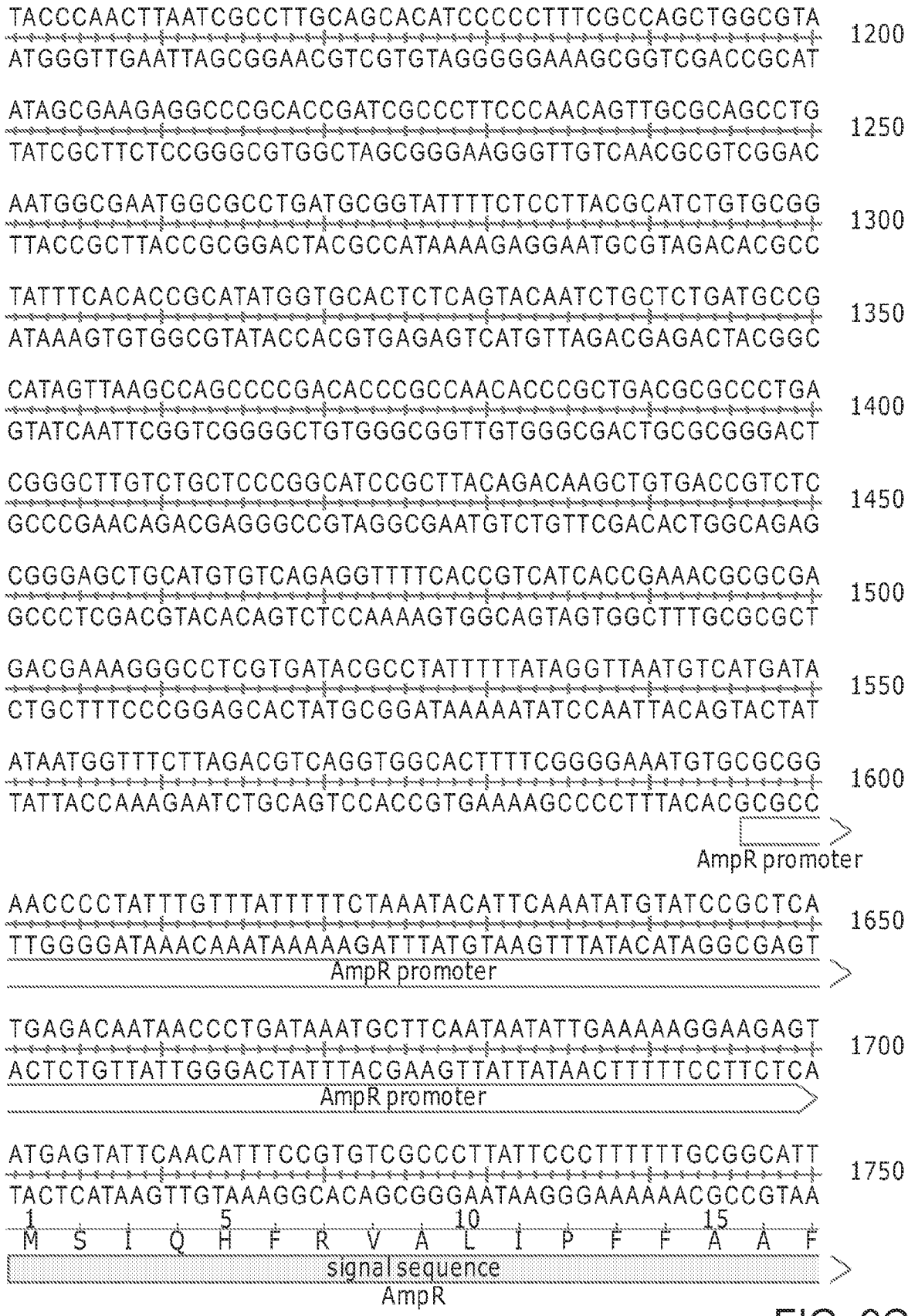

TACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTA
ATGGGTTGAATTAGCGGAACGTCGTGTAGGGGGAAAGCGGTCGACCGCAT
1200
ATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTG
TATCGCTTCTCCGGGCGTGGCTAGCGGGAAGGGTTGTCAACGCGTCGGAC
1250
AATGGCGAATGGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGG
TTACCGCTTACCGCGGACTACGCCATAAAAGAGGAATGCGTAGACACGCC
1300
TATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCG
ATAAAGTGTGGCGTATACCACGTGAGAGTCATGTTAGACGAGACTACGGC
1350
CATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGA
GTATCAATTCGGTCGGGGCTGTGGGCGGTTGTGGGCGACTGCGCGGGACT
1400
CGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTC
GCCCGAACAGACGAGGGCCGTAGGCGAATGTCTGTTCGACACTGGCAGAG
1450
CGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGA
GCCCTCGACGTACACAGTCTCCAAAAGTGGCAGTAGTGGCTTTGCGCGCT
1500
GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATA
CTGCTTTCCCGGAGCACTATGCGGATAAAAATATCCAATTACAGTACTAT
1550
ATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGG
TATTACCAAAGAATCTGCAGTCCACCGTGAAAAGCCCCTTTACACGCGCC
1600
AmpR promoter
AACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA
TTGGGGATAAACAAATAAAAAGATTTATGTAAGTTTATACATAGGCGAGT
1650
AmpR promoter
TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGT
ACTCTGTTATTGGGACTATTTACGAAGTTATTATAACTTTTTCCTTCTCA
1700
AmpR promoter
ATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATT
TACTCATAAGTTGTAAAGGCACAGCGGGAATAAGGGAAAAAACGCCGTAA
1750
1 5 10 15
M S I Q H F R V A L I P F F A A F
signal sequence
AmpR

FIG. 9C

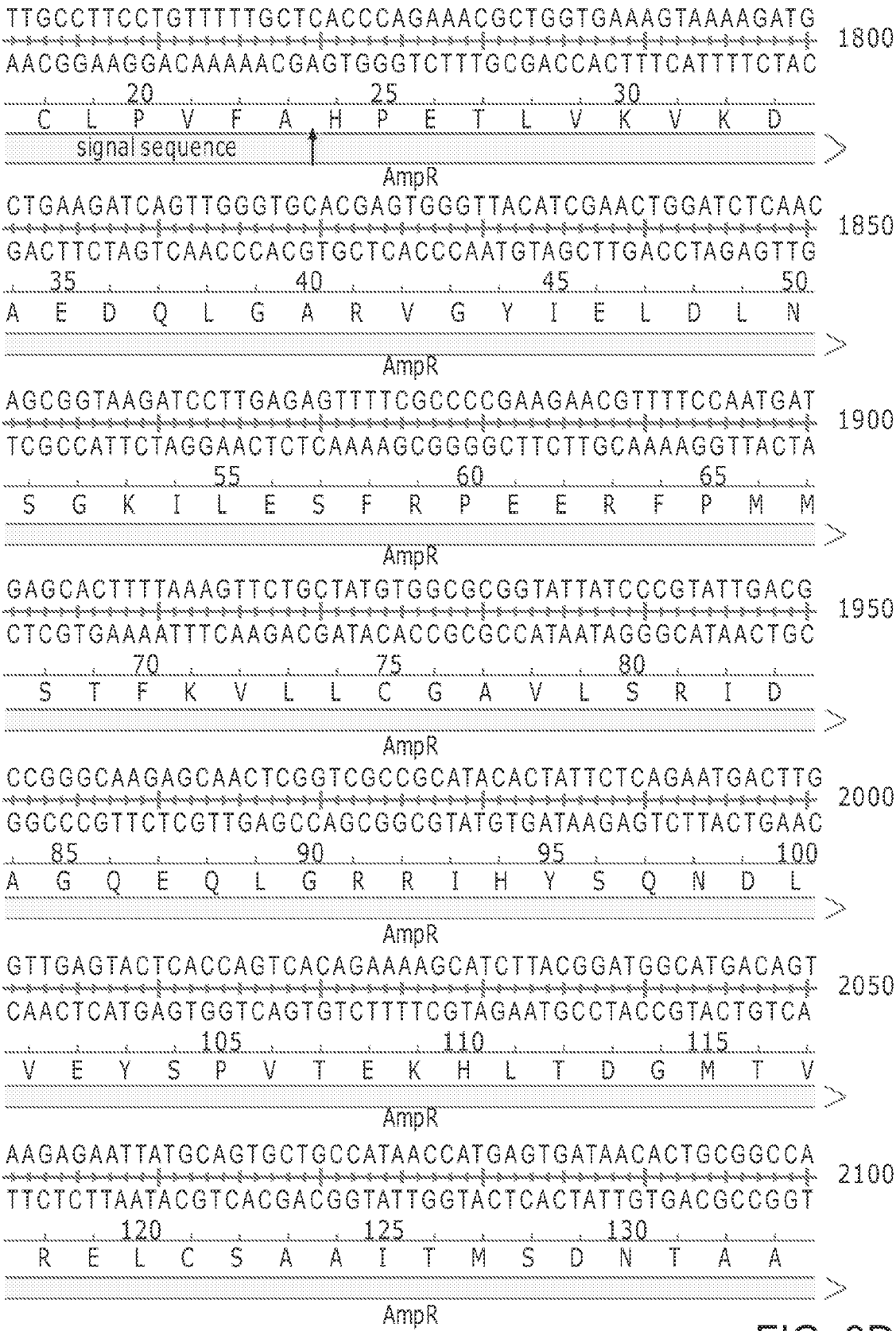

TTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATG
AACGGAAGGACAAAAACGAGTGGGTCTTTGCGACCACTTTCATTTTCTAC
1800
20 25 30
C L P V F A H P E T L V K V K D
signal sequence
AmpR
CTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAAC
GACTTCTAGTCAACCCACGTGCTCACCCAATGTAGCTTGACCTAGAGTTG
1850
35 40 45 50
A E D Q L G A R V G Y I E L D L N
AmpR
AGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGAT
TCGCCATTCTAGGAACTCTCAAAAGCGGGGCTTCTTGCAAAAGGTTACTA
1900
55 60 65
S G K I L E S F R P E E R F P M M
AmpR
GAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACG
CTCGTGAAAATTTCAAGACGATACACCGCGCCATAATAGGGCATAACTGC
1950
70 75 80
S T F K V L L C G A V L S R I D
AmpR
CCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTG
GGCCCGTTCTCGTTGAGCCAGCGGCGTATGTGATAAGAGTCTTACTGAAC
2000
85 90 95 100
A G Q E Q L G R R I H Y S Q N D L
AmpR
GTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGT
CAACTCATGAGTGGTCAGTGTCTTTTCGTAGAATGCCTACCGTACTGTCA
2050
105 110 115
V E Y S P V T E K H L T D G M T V
AmpR
AAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCA
TTCTCTTAATACGTCACGACGGTATTGGTACTCACTATTGTGACGCCGGT
2100
120 125 130
R E L C S A A I T M S D N T A A
AmpR

FIG. 9D

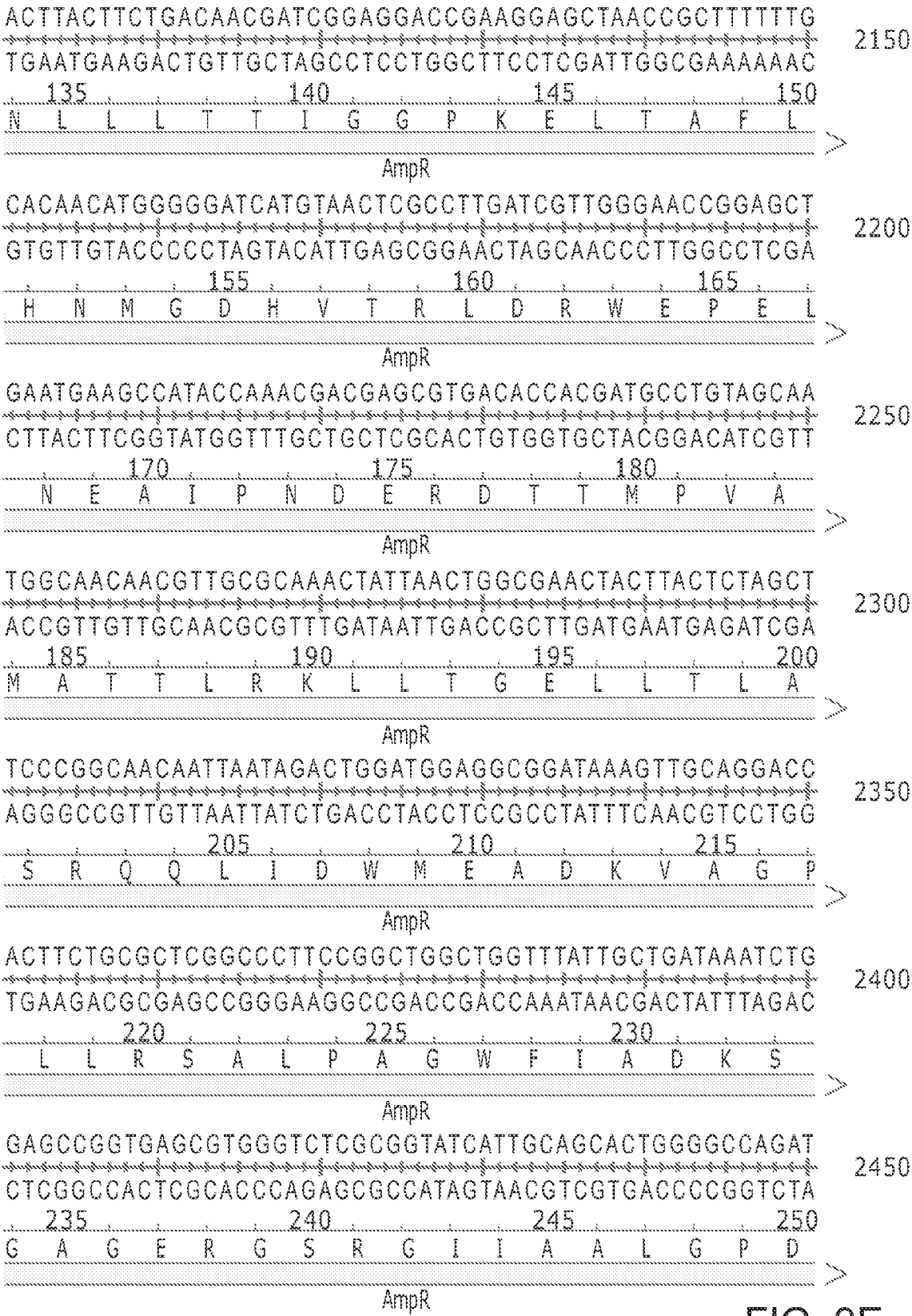

ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTG
2150
TGAATGAAGACTGTTGCTAGCCTCCTGGCTTCCTCGATTGGCGAAAAAAC
135 140 145 150
N L L L T T I G G P K E L T A F L
AmpR
CACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCT
2200
GTGTTGTACCCCCTAGTACATTGAGCGGAACTAGCAACCCTTGGCCTCGA
155 160 165
H N M G D H V T R L D R W E P E L
AmpR
GAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAA
2250
CTTACTTCGGTATGGTTTGCTGCTCGCACTGTGGTGCTACGGACATCGTT
170 175 180
N E A I P N D E R D T T M P V A
AmpR
TGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCT
2300
ACCGTTGTTGCAACGCGTTTGATAATTGACCGCTTGATGAATGAGATCGA
185 190 195 200
M A T T L R K L L T G E L L T L A
AmpR
TCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACC
2350
AGGGCCGTTGTTAATTATCTGACCTACCTCCGCCTATTTCAACGTCCTGG
205 210 215
S R Q Q L I D W M E A D K V A G P
AmpR
ACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTG
2400
TGAAGACGCGAGCCGGGAAGGCCGACCGACCAAATAACGACTATTTAGAC
220 225 230
L L R S A L P A G W F I A D K S
AmpR
GAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGAT
2450
CTCGGCCACTCGCACCCAGAGCGCCATAGTAACGTCGTGACCCCGGTCTA
235 240 245 250
G A G E R G S R G I I A A L G P D
AmpR

FIG. 9E

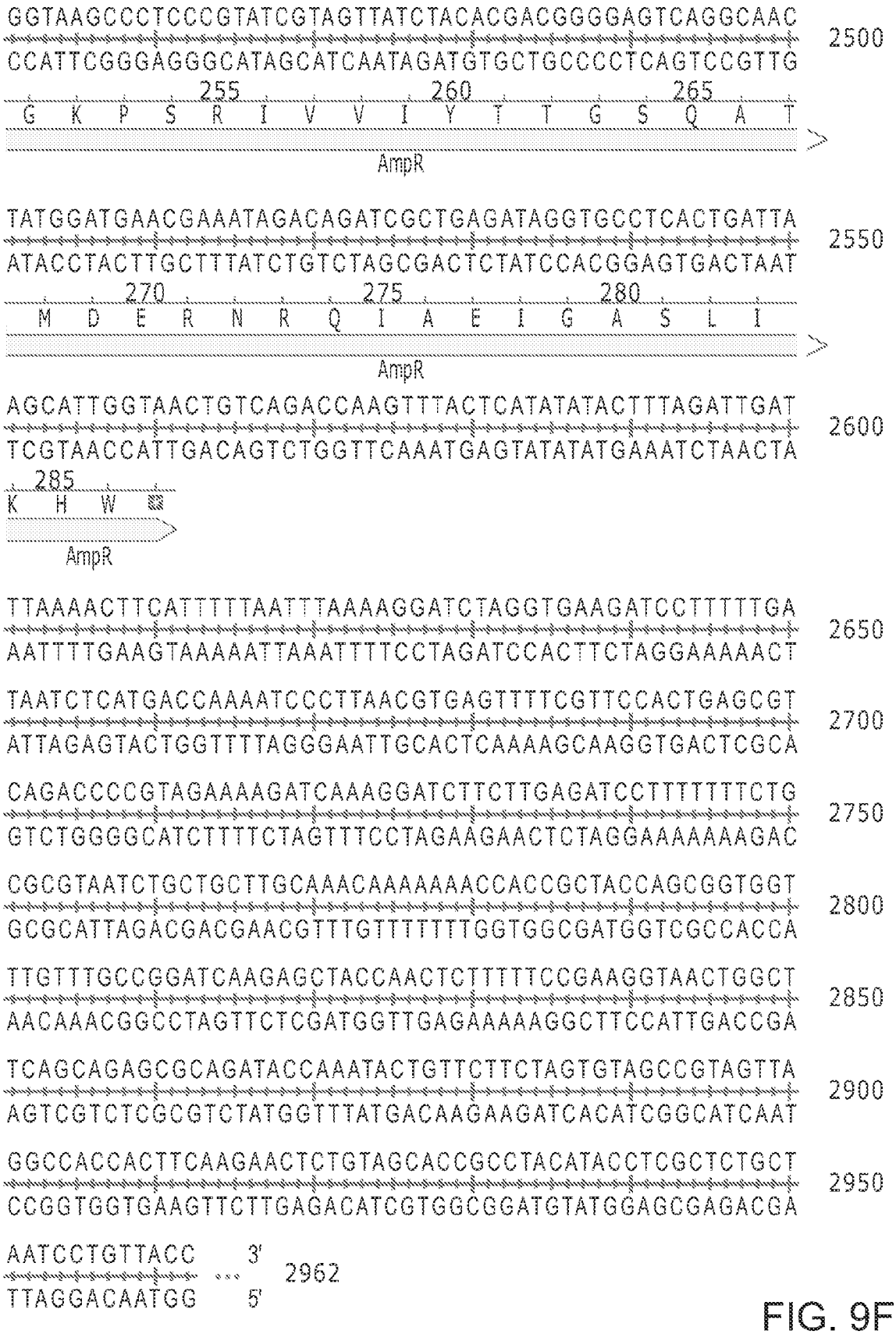

GGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC
CCATTCGGGAGGGCATAGCATCAATAGATGTGCTGCCCCTCAGTCCGTTG
2500
255 260 265
G K P S R I V V I Y T T G S Q A T
AmpR
TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTA
ATACCTACTTGCTTTATCTGTCTAGCGACTCTATCCACGGAGTGACTAAT
2550
270 275 280
M D E R N R Q I A E I G A S L I
AmpR
AGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGAT
TCGTAACCATTGACAGTCTGGTTCAAATGAGTATATATGAAATCTAACTA
2600
285
K H W
AmpR
TTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGA
AATTTTGAAGTAAAAATTAAATTTTCCTAGATCCACTTCTAGGAAAAACT
2650
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGT
ATTAGAGTACTGGTTTTAGGGAATTGCACTCAAAAGCAAGGTGACTCGCA
2700
CAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTG
GTCTGGGGCATCTTTTCTAGTTTCCTAGAAGAACTCTAGGAAAAAAAGAC
2750
CGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGT
GCGCATTAGACGACGAACGTTTGTTTTTTTGGTGGCGATGGTCGCCACCA
2800
TTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCT
AACAAACGGCCTAGTTCTCGATGGTTGAGAAAAAGGCTTCCATTGACCGA
2850
TCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTA
AGTCGTCTCGCGTCTATGGTTTATGACAAGAAGATCACATCGGCATCAAT
2900
GGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCT
CCGGTGGTGAAGTTCTTGAGACATCGTGGCGGATGTATGGAGCGAGACGA
2950
AATCCTGTTACC 3'
TTAGGACAATGG 5'
2962

FIG. 9F

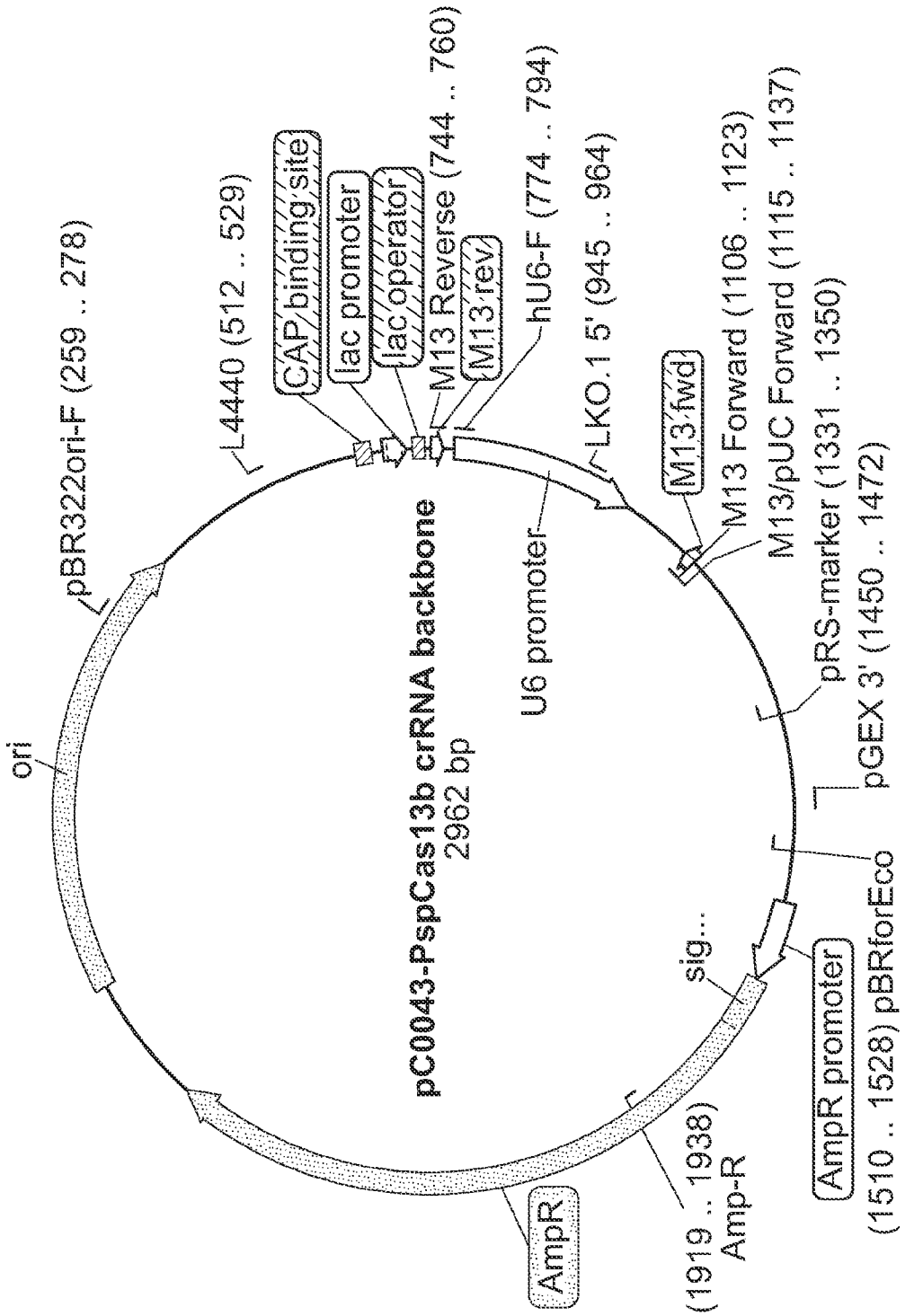

pC0043-PspCas13b crRNA backbone
2962 bp
ori
pBR322ori-F (259 .. 278)
L4440 (512 .. 529)
CAP binding site
lac promoter
lac operator
M13 Reverse (744 .. 760)
M13 rev
hU6-F (774 .. 794)
LKO.1 5' (945 .. 964)
U6 promoter
M13 fwd
M13 Forward (1106 .. 1123)
M13/pUC Forward (1115 .. 1137)
pRS-marker (1331 .. 1350)
pGEX 3' (1450 .. 1472)
pBRforEco
AmpR promoter
(1510 .. 1528)
sig...
(1919 .. 1938)
Amp-R
AmpR

FIG. 10

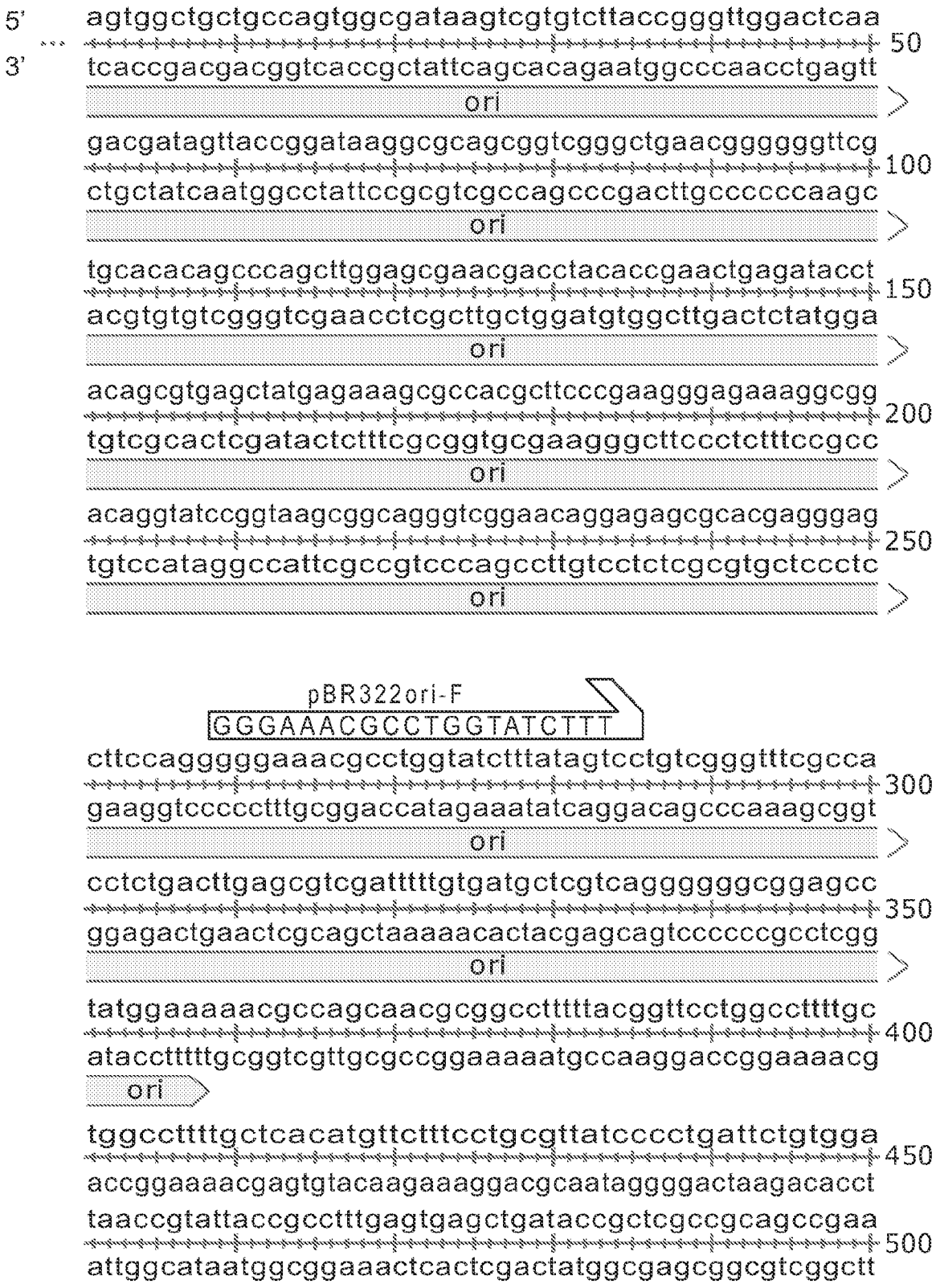

5'
3'
agtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaa
tcaccgacgacggtcaccgctattcagcacagaatggcccaacctgagtt
50
ori
gacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcg
ctgctatcaatggcctattccgcgtcgccagcccgacttgccccccaagc
100
ori
tgcacacagcccagcttggagcgaacgacctacaccgaactgagatacct
acgtgtgtcgggtcgaacctcgcttgctggatgtggcttgactctatgga
150
ori
acagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcgg
tgtcgcactcgatactctttcgcggtgcgaagggcttccctctttccgcc
200
ori
acaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggag
tgtccataggccattcgccgtcccagccttgtcctctcgcgtgctccctc
250
ori
pBR322ori-F
GGGAAACGCCTGGTATCTTT
cttccaggggggaaacgcctggtatctttatagtcctgtcgggtttcgcca
gaaggtccccctttgcggaccatagaaatatcaggacagcccaaagcggt
300
ori
cctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcc
ggagactgaactcgcagctaaaaacactacgagcagtccccccgcctcgg
350
ori
tatggaaaaacgccagcaacgcggccttttacggttcctggccttttgc
ataccttttgcggtcgttgcgccggaaaaatgccaaggaccggaaaacg
400
ori
tggccttttgctcacatgttctttcctgcgttatcccctgattctgtgga
accggaaaacgagtgtacaagaaaggacgcaataggggactaagacacct
450
taaccgtattaccgcctttgagtgagctgataccgctcgccgcagccgaa
attggcataatggcggaaactcactcgactatggcgagcggcgtcggctt
500

FIG. 11A

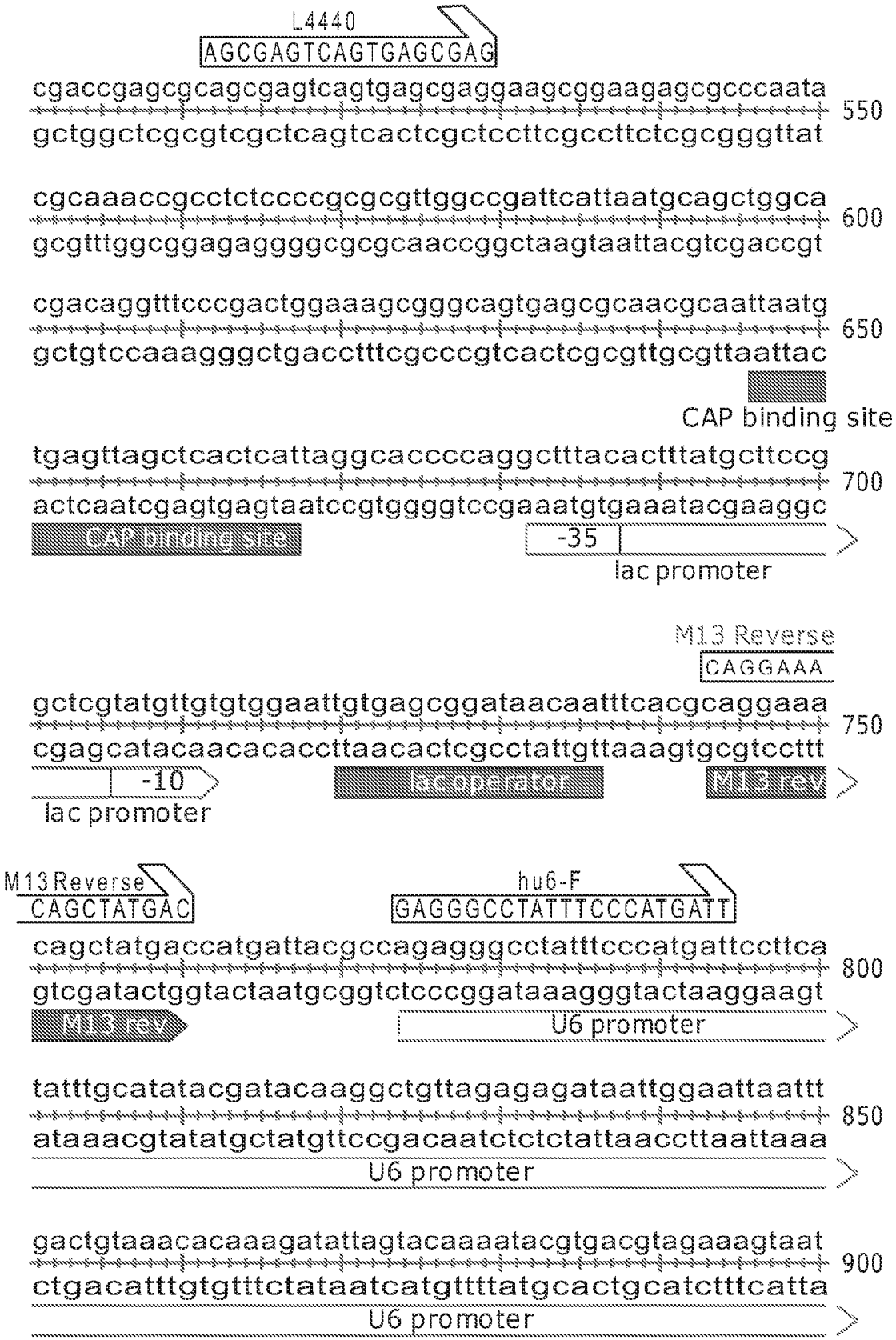

L4440
AGCGAGTCAGTGAGCGAG
cgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaata
gctggctcgcgtcgctcagtcactcgctccttcgccttctcgcgggttat
550
cgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggca
gcgtttggcggagaggggcgcgcaaccggctaagtaattacgtcgaccgt
600
cgacaggtttcccgactggaaagcgggcagtgagcgcaacgcaattaatg
gctgtccaaagggctgacctttcgcccgtcactcgcgttgcgttaattac
650
CAP binding site
tgagttagctcactcattaggcaccccaggctttacactttatgcttccg
actcaatcgagtgagtaatccgtggggtccgaaatgtgaaatacgaaggc
700
CAP binding site
-35
lac promoter
M13 Reverse
CAGGAAA
gctcgtatgttgtgtggaattgtgagcggataacaatttcacgcaggaaa
cgagcatacaacacaccttaacactcgcctattgttaaagtgcgtccttt
750
-10
lac promoter
lac operator
M13 rev
M13Reverse
CAGCTATGAC
hu6-F
GAGGGCCTATTTCCCATGATT
cagctatgaccatgattacgccagagggcctatttcccatgattccttca
gtcgatactggtactaatgcggtctcccggataaagggtactaaggaagt
800
M13 rev
U6 promoter
tatttgcatatacgatacaaggctgttagagagataattggaattaattt
ataaacgtatatgctatgttccgacaatctctctattaaccttaattaaa
850
U6 promoter
gactgtaaacacaaagatattagtacaaaatacgtgacgtagaaagtaat
ctgacatttgtgtttctataatcatgttttatgcactgcatctttcatta
900
U6 promoter

FIG. 11B

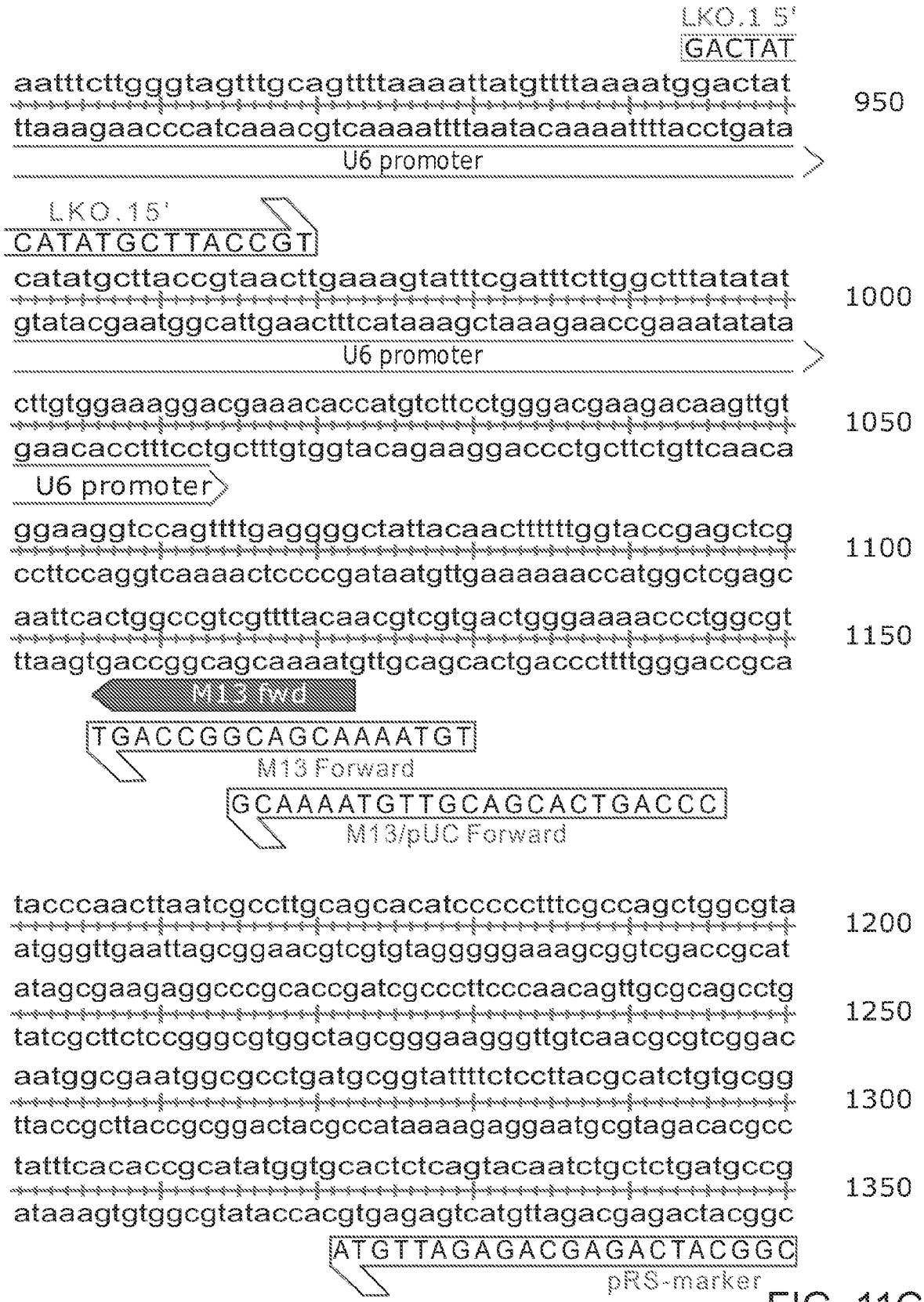

LKO.1 5'
GACTAT
aatttcttgggtagtttgcagttttaaaattatgttttaaaatggactat
ttaaagaacccatcaaacgtcaaaatttaatacaaaatttacctgata
950
U6 promoter
LKO.1 5'
CATATGCTTACCGT
catatgcttaccgtaacttgaaagtatttcgatttcttggctttatatat
gtatacgaatggcattgaactttcataaagctaaagaaccgaaatatata
1000
U6 promoter
cttgtggaaaggacgaaacaccatgtcttcctgggacgaagacaagttgt
gaacacctttcctgctttgtggtacagaaggaccctgcttctgttcaaca
1050
U6 promoter
ggaaggtccagttttgaggggctattacaacttttttggtaccgagctcg
ccttccaggtcaaaactccccgataatgttgaaaaaaccatggctcgagc
1100
aattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgt
ttaagtgaccggcagcaaaatgttgcagcactgacccttttgggaccgca
1150
M13 rvs
TGACCGGCAGCAAAATGT
M13 Forward
GCAAAATGTTGCAGCACTGACCC
M13/pUC Forward
tacccaacttaatcgccttgcagcacatccccctttcgccagctggcgta
atgggttgaattagcggaacgtcgtgtaggggggaaagcggtcgaccgcat
1200
atagcgaagaggcccgcaccgatcgcccttcccaacagttgcgcagcctg
tatcgcttctccgggcgtggctagcgggaagggttgtcaacgcgtcggac
1250
aatggcgaatggcgcctgatgcggtattttctccttacgcatctgtgcgg
ttaccgcttaccgcggactacgccataaaagaggaatgcgtagacacgcc
1300
tatttcacaccgcatatggtgcactctcagtacaatctgctctgatgccg
ataaagtgtggcgtataccacgtgagagtcatgttagacgagactacggc
1350
ATGTTAGAGACGAGACTACGGC
pRS-marker

FIG. 11C

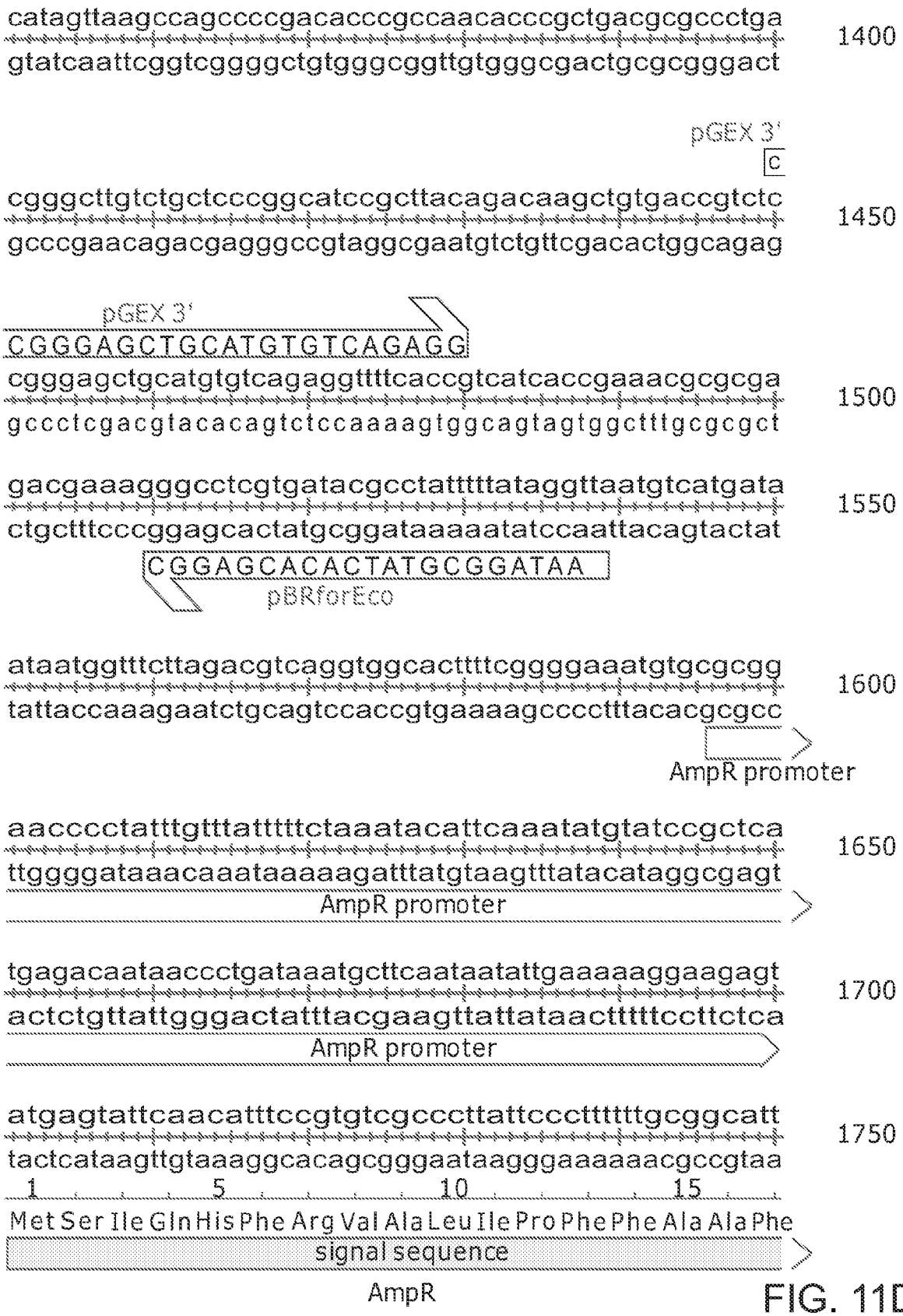

catagttaagccagccccgacacccgccaacacccgctgacgcgccctga
gtatcaattcggtcggggctgtgggcggttgtgggcgactgcgcgggact
1400
pGEX 3'
C
cgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctc
gcccgaacagacgagggccgtaggcgaatgtctgttcgacactggcagag
1450
pGEX 3'
CGGGAGCTGCATGTGTCAGAGG
cgggagctgcatgtgtcagaggttttcaccgtcatcaccgaaacgcgcga
gccctcgacgtacacagtctccaaaagtggcagtagtggctttgcgcgct
1500
gacgaaagggcctcgtgatacgcctatttttataggttaatgtcatgata
ctgctttcccggagcactatgcggataaaaatatccaattacagtactat
1550
CGGAGCACACTATGCGGATAA
pBRforEco
ataatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcgg
tattaccaaagaatctgcagtccaccgtgaaaagccccttacacgcgcc
1600
AmpR promoter
aaccccctatttgtttatttttctaaatacattcaaatatgtatccgctca
ttggggataaacaaataaaaagatttatgtaagtttatacataggcgagt
1650
AmpR promoter
tgagacaataaccctgataaatgcttcaataatattgaaaaaggaagagt
actctgttattgggactatttacgaagttattataacttttttccttctca
1700
AmpR promoter
atgagtattcaacatttccgtgtcgcccttattccctttttttgcggcatt
tactcataagttgtaaaggcacagcgggaataagggaaaaaaacgccgtaa
1750
1 5 10 15
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala Phe
signal sequence
AmpR

FIG. 11D

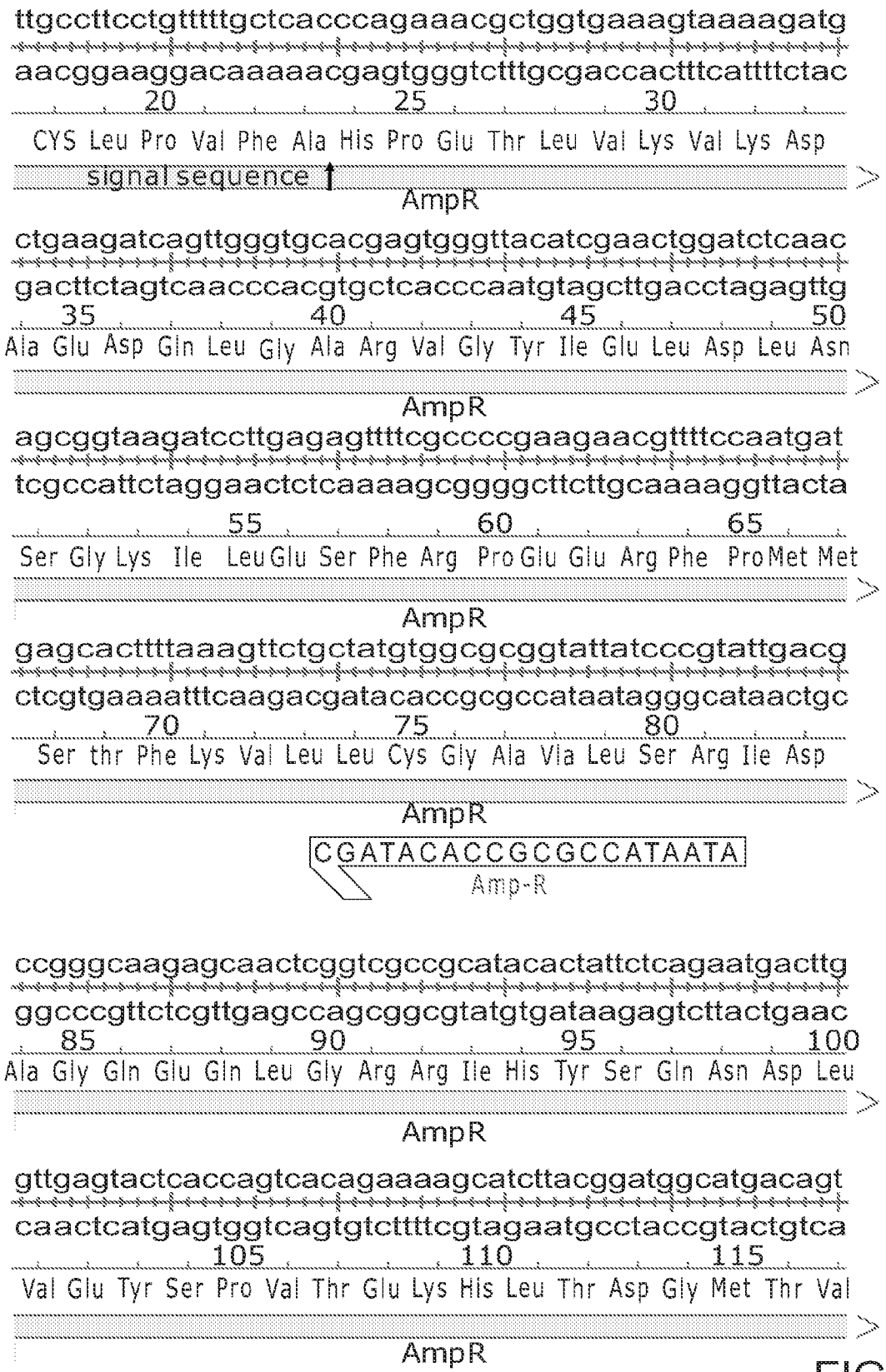

ttgccttcctgtttttgctcacccagaaacgctggtgaaagtaaaagatg
1800
aacggaaggacaaaaacgagtgggtctttgcgaccactttcattttctac
20 25 30
CYS Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys Asp
signal sequence
AmpR
ctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaac
1850
gacttctagtcaacccacgtgctcacccaatgtagcttgacctagagttg
35 40 45 50
Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp Leu Asn
AmpR
agcggtaagatccttgagagttttcgccccgaagaacgttttccaatgat
1900
tcgccattctaggaactctcaaaagcggggcttcttgcaaaaggttacta
55 60 65
Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe Pro Met Met
AmpR
gagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacg
1950
ctcgtgaaaatttcaagacgatacaccgcgccataatagggcataactgc
70 75 80
Ser thr Phe Lys Val Leu Leu Cys Gly Ala Via Leu Ser Arg Ile Asp
AmpR
CGATACACCGCGCCATAATA
Amp-R
ccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttg
2000
ggcccgttctcgttgagccagcggcgtatgtgataagagtcttactgaac
85 90 95 100
Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser Gln Asn Asp Leu
AmpR
gttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagt
2050
caactcatgagtggtcagtgtcttttcgtagaatgcctaccgtactgtca
105 110 115
Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr Asp Gly Met Thr Val
AmpR

FIG. 11E


aagagaattatgcagtgctgccataaccatgagtgataacactgcggcca
2100
ttctcttaatacgtcacgacggtattggtactcactattgtgacgccggt
120 125 130
Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala
AmpR
acttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttg
2150
tgaatgaagactgttgctagcctcctggcttcctcgattggcgaaaaaac
135 140 145 150
Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu
AmpR
cacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagct
2200
gtgttgtaccccctagtacattgagcggaactagcaacccttggcctcga
155 160 165
His Asn Met Gly Asp His Val Thr Arg Leu Asp Arg Trp Trp Glu Pro Glu Leu
AmpR
gaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaa
2250
cttacttcggtatggtttgctgctcgcactgtggtgctacggacatcgtt
170 175 180
Asn Glu Ala Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val Ala
AmpR
tggcaacaacgttgcgcaaactattaactggcgaactacttactctagct
2300
accgttgttgcaacgcgtttgataattgaccgcttgatgaatgagatcga
185 190 195 200
Met Ala Thr Thr Leu Arg Lys Leu Leu Thr Gly Glu Glu Leu Leu Thr Leu Ala
AmpR
tcccggcaacaattaatagactggatggaggcggataaagttgcaggacc
2350
agggccgttgttaattatctgacctacctccgcctatttcaacgtcctgg
205 210 215
Ser Arg Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro
AmpR

FIG. 11F

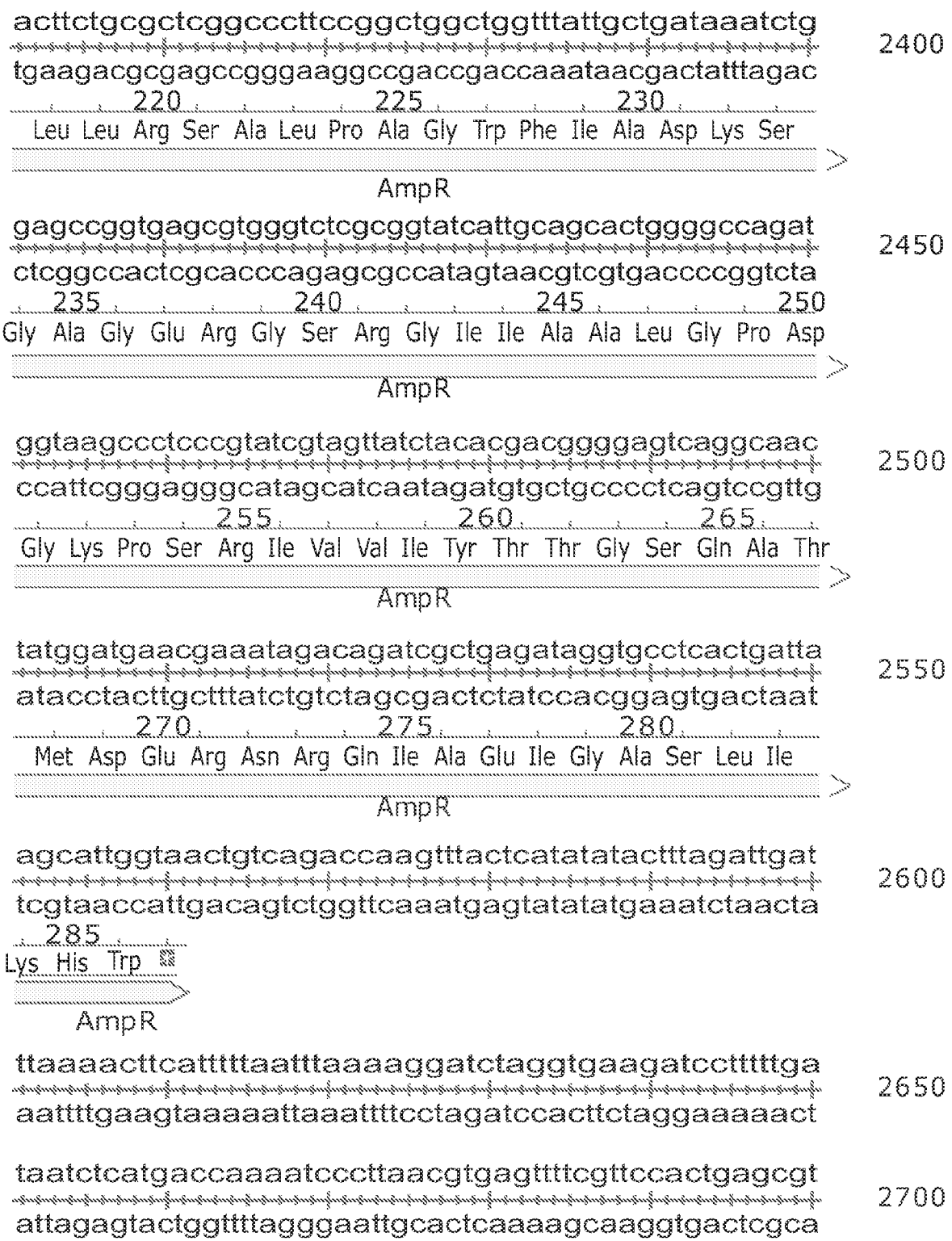

acttctgcgctcggcccttccggctggctggtttattgctgataaatctg
2400
tgaagacgcgagccgggaaggccgaccgaccaaataacgactatttagac
220 225 230
Leu Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser
AmpR
gagccggtgagcgtgggtctcgcggtatcattgcagcactggggccagat
2450
ctcggccactcgcacccagagcgccatagtaacgtcgtgaccccggtcta
235 240 245 250
Gly Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp
AmpR
ggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaac
2500
ccattcgggagggcatagcatcaatagatgtgctgcccctcagtccgttg
255 260 265
Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr
AmpR
tatggatgaacgaaatagacagatcgctgagataggtgcctcactgatta
2550
atacctacttgctttatctgtctagcgactctatccacggagtgactaat
270 275 280
Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile
AmpR
agcattggtaactgtcagaccaagtttactcatatatactttagattgat
2600
tcgtaaccattgacagtctggttcaaatgagtatatatgaaatctaacta
285
Lys His Trp
AmpR
ttaaaacttcatttttaatttaaaaggatctaggtgaagatccttttga
2650
aattttgaagtaaaaattaaatttcctagatccacttctaggaaaaact
taatctcatgaccaaaatcccttaacgtgagttttcgttccactgagcgt
2700
attagagtactggttttagggaattgcactcaaaagcaaggtgactcgca

FIG. 11G

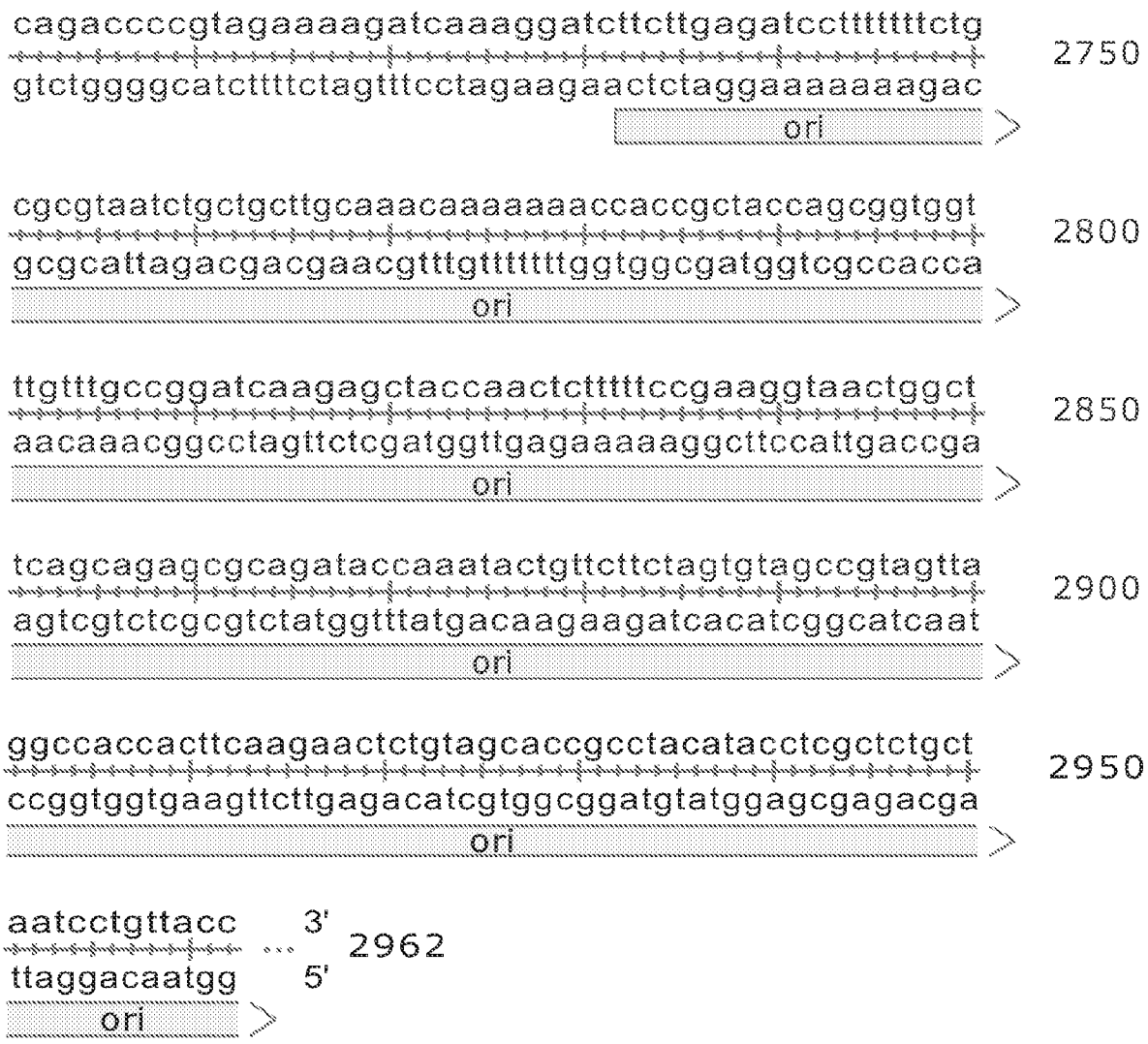

cagaccccgtagaaaagatcaaaggatcttcttgagatccttttttttctg
gtctggggcatcttttctagtttcctagaagaactctaggaaaaaaaagac
ori
2750
cgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggt
gcgcattagacgacgaacgtttgtttttttggtggcgatggtcgccacca
ori
2800
ttgtttgccggatcaagagctaccaactcttttttccgaaggtaactggct
aacaaacggcctagttctcgatggttgagaaaaaggcttccattgaccga
ori
2850
tcagcagagcgcagataccaaatactgttcttctagtgtagccgtagtta
agtcgtctcgcgtctatggtttatgacaagaagatcacatcggcatcaat
ori
2900
ggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgct
ccggtggtgaagttcttgagacatcgtggcggatgtatggagcgagacga
ori
2950
aatcctgttacc
ttaggacaatgg
ori
3'
5'
2962

FIG. 11H

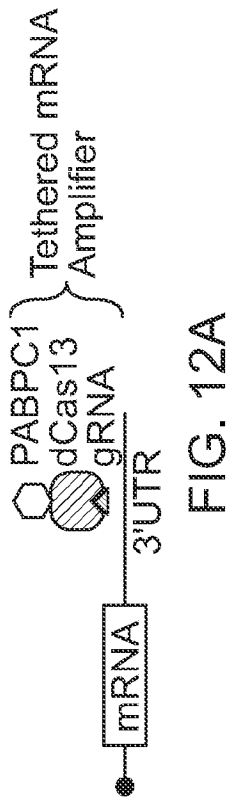


FIG. 12A

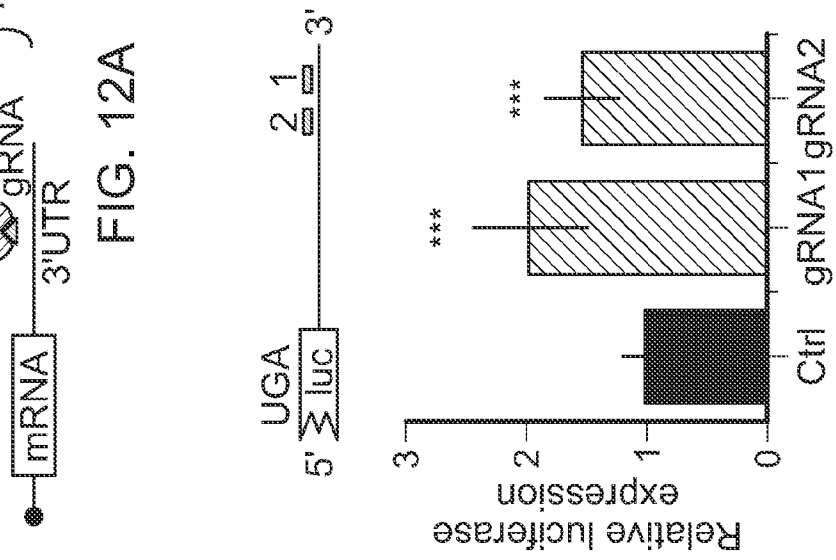
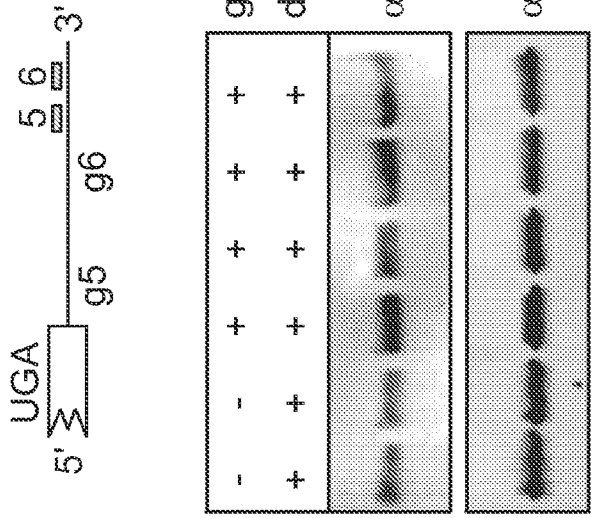


FIG. 12B



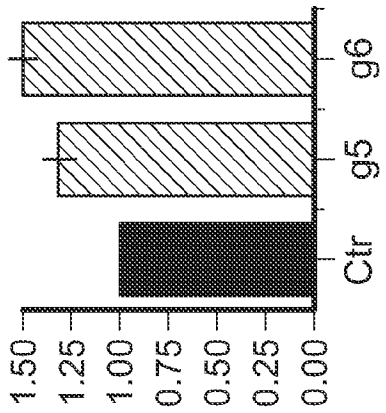


FIG. 12C

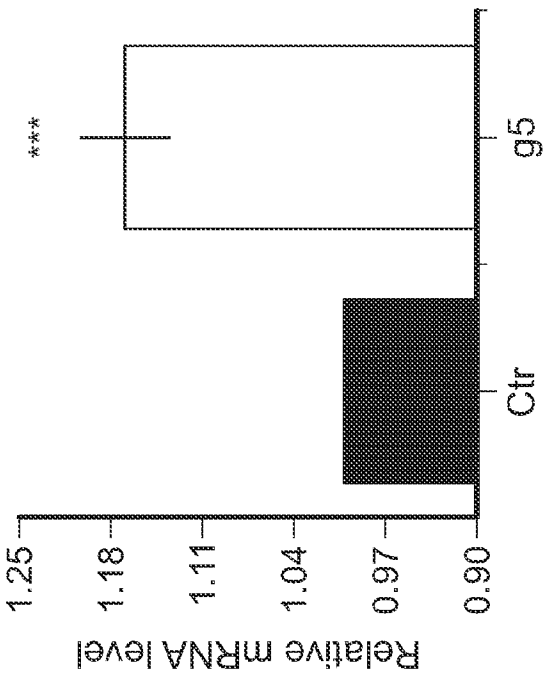


FIG. 12E

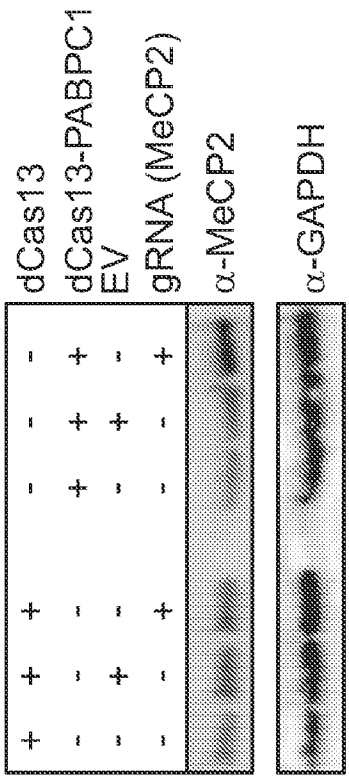


FIG. 12D

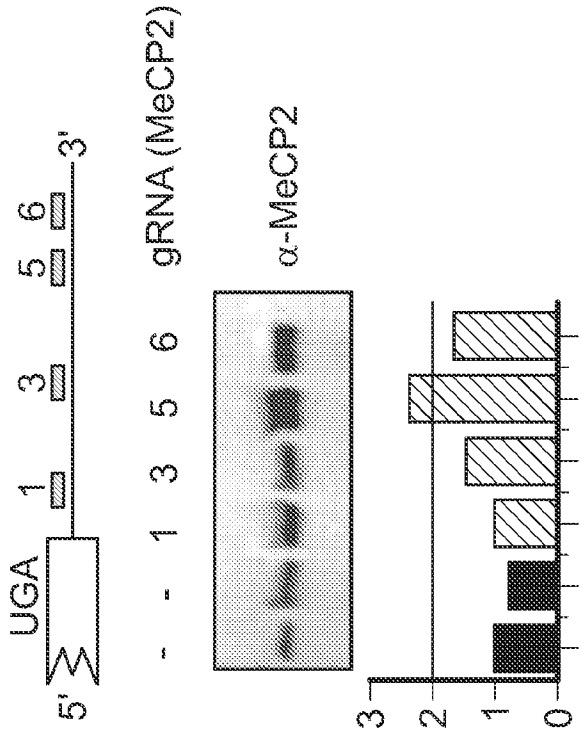


FIG. 12G

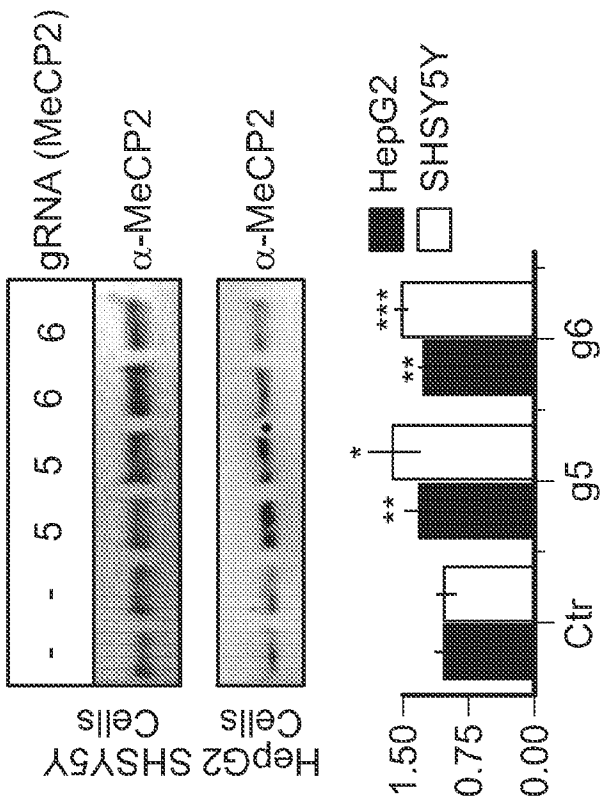


FIG. 12F

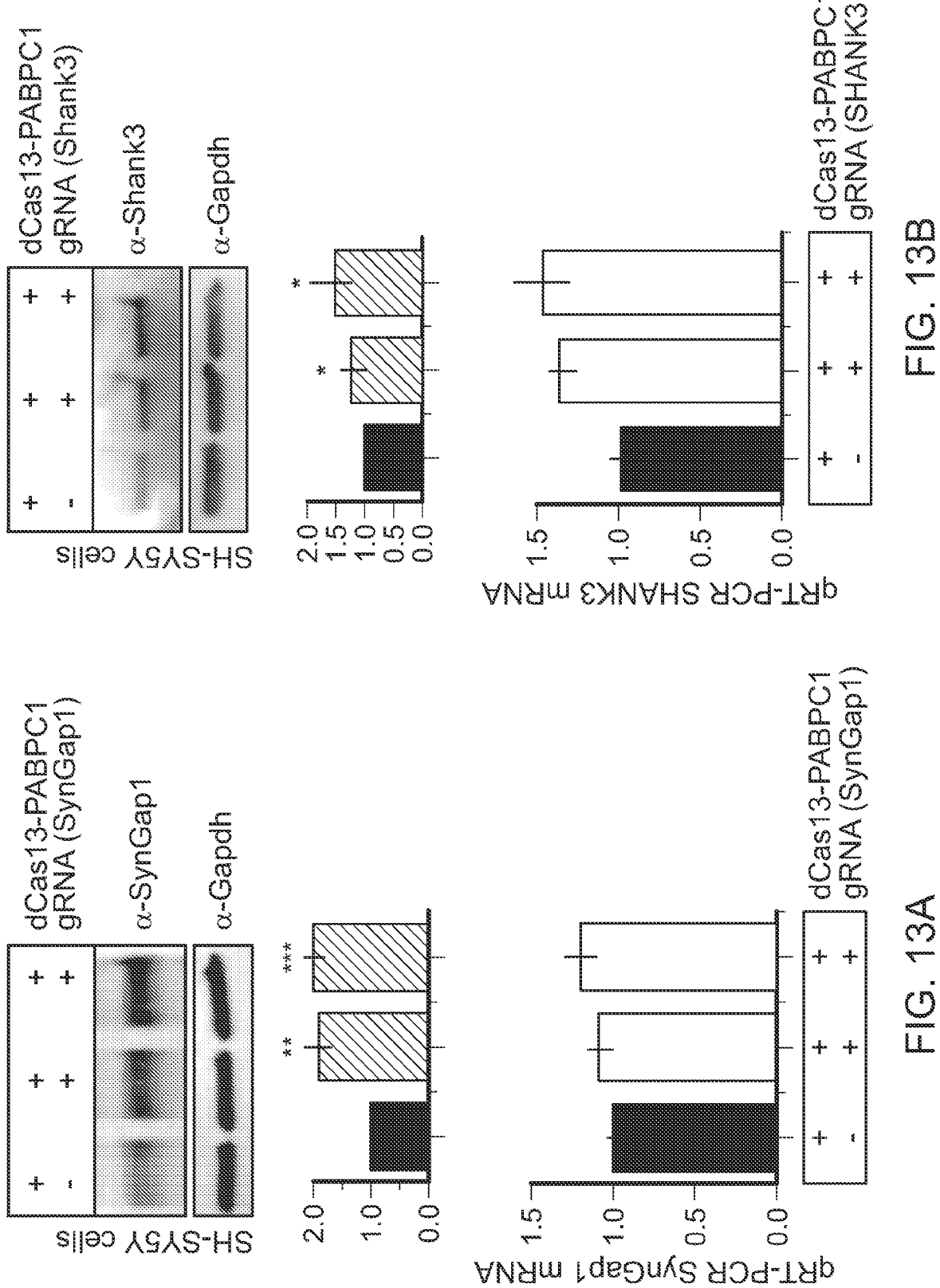


FIG. 13A

FIG. 13B

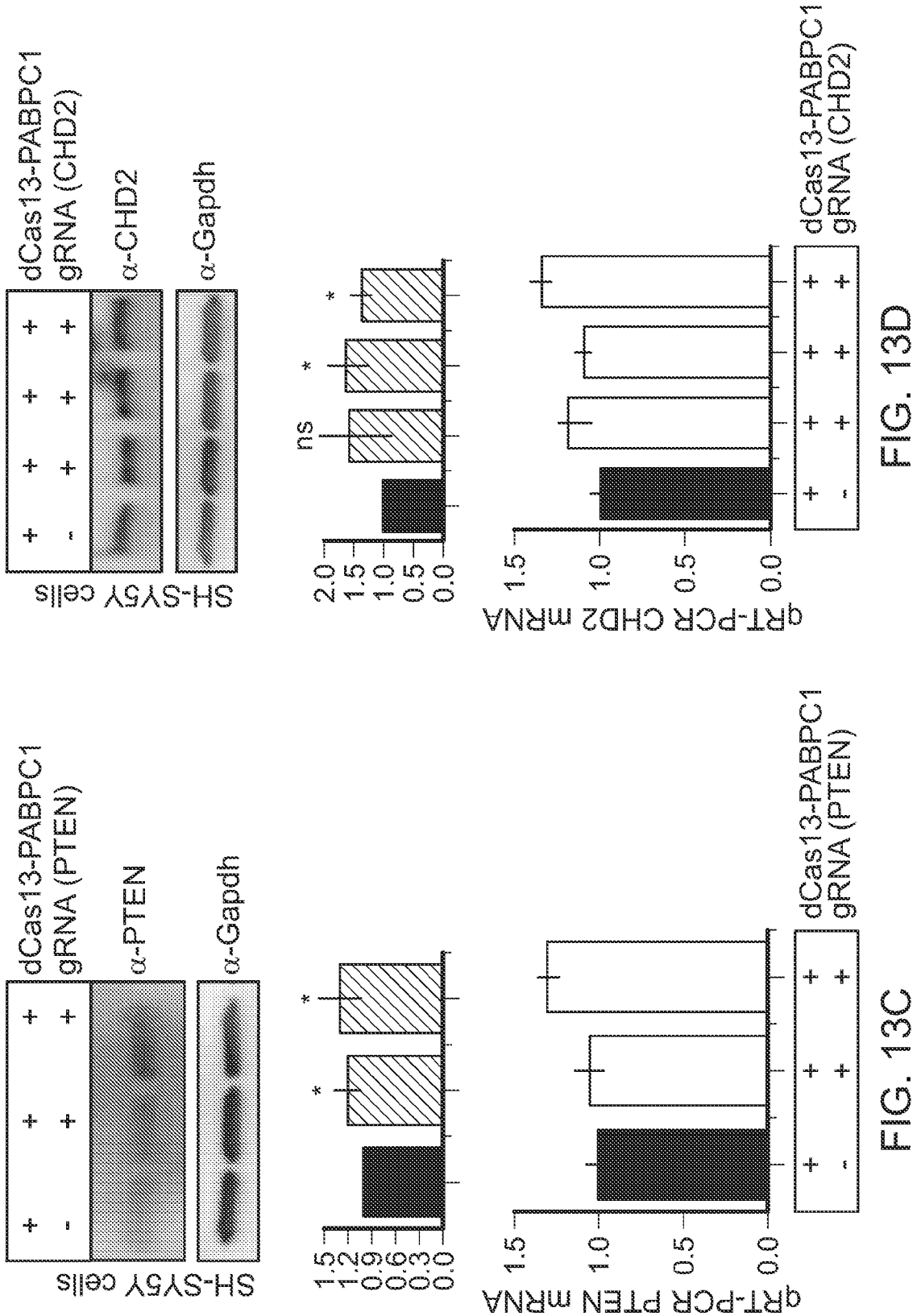


FIG. 13C

FIG. 13D

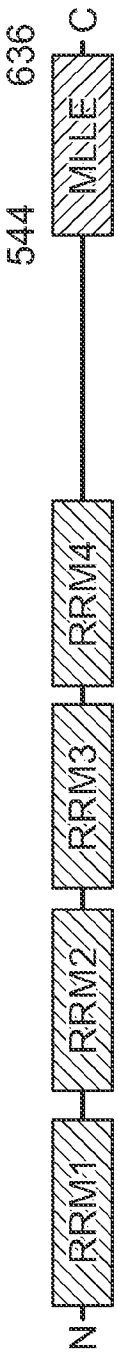


FIG. 14A

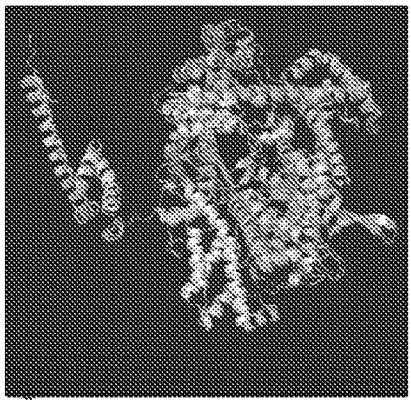


FIG. 14E

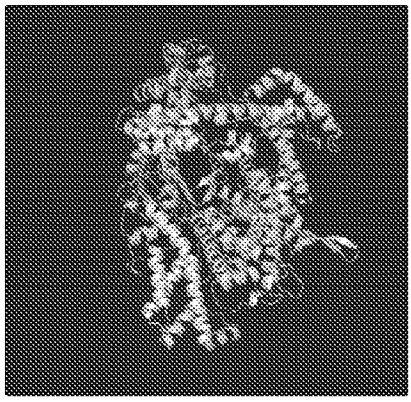


FIG. 14D

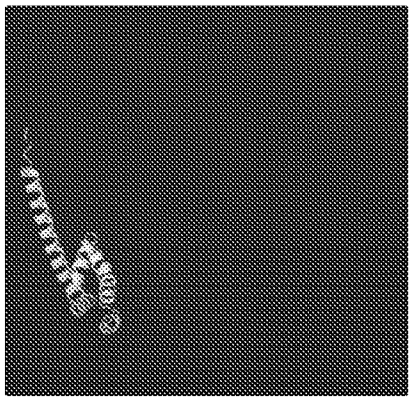


FIG. 14C

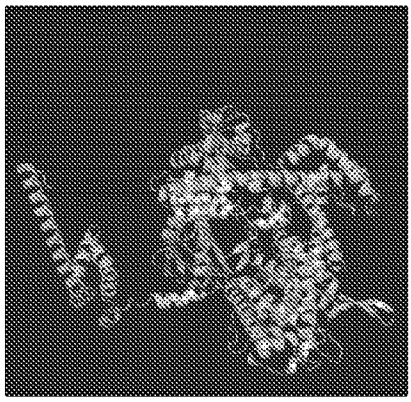


FIG. 14B

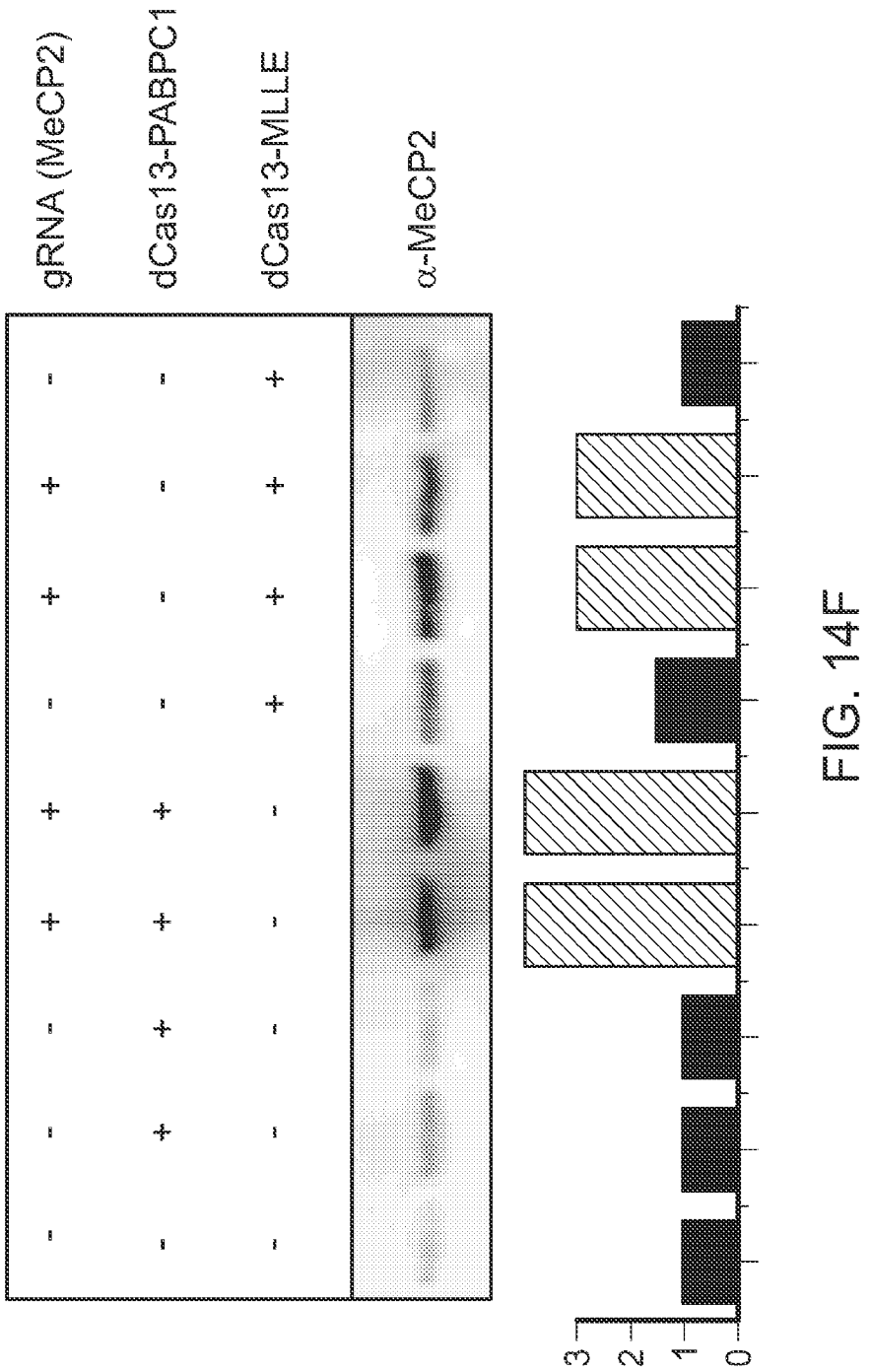


FIG. 14F

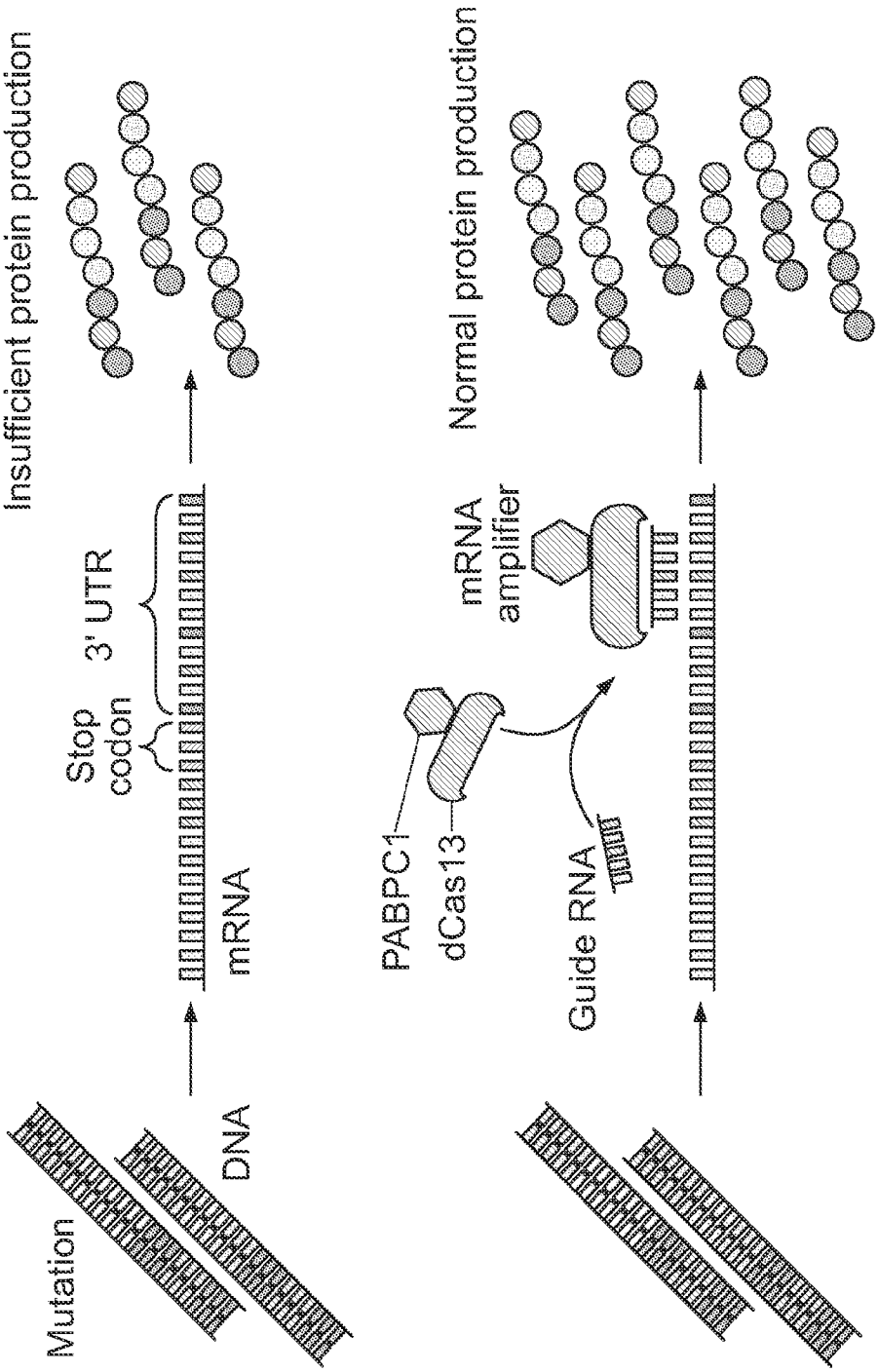

Insufficient protein production
Normal protein production
Mutation
DNA
mRNA
Stop codon
3' UTR
PABPC1
dCas13
Guide RNA
mRNA amplifier
PABPC1: Cytoplasmic polyadenylate-binding protein 1
dCas13: Deactivated Cas13
3' UTR: 3' Untranslated region

FIG. 15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/076723

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - INV. - A61K 38/46; A61K 48/00; C07K 14/195; C12N 5/10; C12N 15/55 (2023.01)
ADD.

CPC - INV. - A61K 38/465; A61K 48/0008; C07K 14/195; C12N 5/10; C12N 15/62 (2023.01)
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)
See Search History document

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

Electronic database consulted during the international search (name of database and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2015-221026 (NAGOYA CITY UNIVERSITY) 10 December 2015 (10.12.2015) entire document	1, 2, 4, 10, 14, 15, 17, 34, 36, 41, 43, 56, 57, 59
Y	US 2021/0071178 A1 (MASSACHUSETTS INSTITUTE OF TECHNOLOGY) 11 March 2021 (11.03.2021) entire document	1, 2, 4, 10, 14, 15, 17, 34, 36, 41, 43, 56, 57, 59
Y	US 2020/0024603 A1 (STOKE THERAPEUTICS INC.) 23 January 2020 (23.01.2020) entire document	1, 2, 4, 10, 14, 15, 17, 34, 36, 41, 43, 56, 57, 59
A	US 2018/0312873 A1 (10X GENOMICS INC.) 01 November 2018 (01.11.2018) entire document	1, 2, 4, 10, 14, 15, 17, 34, 36, 41, 43, 56, 57, 59
A	US 2019/0351074 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 21 November 2019 (21.11.2019) entire document	1, 2, 4, 10, 14, 15, 17, 34, 36, 41, 43, 56, 57, 59

☐ Further documents are listed in the continuation of Box C. ☐ 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
"D" document cited by the applicant in the international application
"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
05 January 2023	FEB 15 2023
Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, VA 22313-1450 Facsimile No. 571-273-8300	Authorized officer Taina Matos Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2022/076723

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. ☒ 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/US2022/076723

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

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

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

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

3. ☒ Claims Nos.: 5-9, 11-13, 18-33, 48-55, 60-78
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See extra sheet(s).

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

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

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

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

1, 2, 4, 10, 14, 15, 17, 34, 36, 41, 43, 56, 57, 59

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

Continued from Box No. III Observations where unity of invention is lacking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees need to be paid.

Group I+: claims 1-4, 10, 14-17, 34-47, and 56-59 are drawn to a fusion protein comprising: an RNA effector protein that targets mRNA (s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA a system and complex comprising the fusion protein and a gRNA that forms a complex with the RNA effector protein and comprises a complementarity region that hybridizes with the mRNA(s) of the active allele; and a method of treating or preventing a haploinsufficiency disorder in a subject, the method comprising: administering to the subject the fusion protein and gRNA.

The first invention of Group I+ is restricted to a fusion protein comprising: Cas9 fused to PABPC1, wherein the fusion protein comprises SEQ ID NO: 48; and the gRNA comprises a region designed to hybridize with SEQ ID NO: 1. It is believed that claims 1, 2, 4, 10, 14, 15, 17, 34-36, 41-43, 56, 57, and 59 read on this first named invention and thus these claims will be searched without fee to the extent that they read on the above embodiment.

Applicant is invited to elect additional fusion protein(s) or gRNA(s) for each additional embodiment to be searched in a specific combination by paying an additional fee for each election. Each additional elected embodiment requires the selection of a single definition for each fusion protein(s) or gRNA(s). An exemplary election would be a fusion protein comprising SEQ ID NO: 49. Additional embodiment(s) will be searched upon the payment of additional fees. Applicants must specify the claims that read on any additional elected inventions. Applicants must further indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined.

The inventions listed in Groups I+ do not relate to a single general inventive concept under PCT Rule 13.1, because under PCT Rule 13.2 they lack the same or corresponding special technical features for the following reasons:

The Groups I+ formulae do not share a significant structural element, requiring the selection of alternatives for fusion proteins and/or gRNAs, and accordingly these groups lack unity a priori.

Additionally, even if Groups I+ were considered to share the technical features of: a system and complex comprising: a fusion protein comprising: an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA(s), bound to a gRNA comprising a complementarity region that hybridizes with the mRNA(s) of the active allele, and the mRNA; and a method of treating or preventing a haploinsufficiency disorder in a subject, the method comprising: administering to the subject a fusion protein or a nucleic acid encoding a fusion protein comprising: an RNA effector protein that targets mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder; and a regulon moiety that stimulates and/or stabilizes the mRNA(s); and a gRNA or a nucleic acid encoding a gRNA designed to form a complex with the RNA effector protein and comprising a complementarity region designed to hybridize with the mRNA of the active allele, these shared technical features do not represent a contribution over the prior art, as disclosed by JP 2015-221026 to Nagoya City University (hereinafter, "Nagoya"), and US 2019/0351074 to The Regents of the University of California (hereinafter, "California").

Specifically, Nagoya discloses a system and complex comprising: a fusion protein comprising: an RNA effector protein that targets mRNA(s); and a regulon moiety that stimulates and/or stabilizes the mRNA(s) (a nucleic acid construct that expresses a translating-related factor fused to an RNA binding peptide, Para. [0008]; examples of said translation related factors are PABPC1, Para. [0016]); and a method comprising: administering to a subject a fusion protein or a nucleic acid encoding a fusion protein comprising: an RNA effector protein; and a regulon moiety that stimulates and/or stabilizes the mRNA(s) (a nucleic acid construct that expresses a translating-related factor fused to an RNA binding peptide, Para. [0008]; elements will be administered to the target cell; Para. [0009]; medicament of the invention is applied to a subject by oral administration or parenteral administration, Para. [0065]).

Further, California teaches mRNA(s) of an active allele of a gene associated with a haploinsufficiency disorder (Using CRISPRa targeting for either the promoter or enhancer of Sim1, we were able to rescue the obesity phenotype in a tissue-specific manner in mice that are haploinsufficient for Sim1, Para. [0156]); bound to a gRNA (dCAS9/gRNA complex, Para. [0006]); and a method of treating or preventing a haploinsufficiency disorder in a subject (Using CRISPRa targeting for either the promoter or enhancer of Sim1, we were able to rescue the obesity phenotype in a tissue-specific manner in mice that are haploinsufficient for Sim1, Para. [0156]).

The inventions listed in Groups I+ therefore lack unity under Rule 13 because they do not share a same or corresponding special technical feature.