

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 960 692**

51 Int. Cl.:

C12N 9/10 (2006.01)

A61K 48/00 (2006.01)

C12N 15/86 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.12.2019** **PCT/US2019/064786**

87 Fecha y número de publicación internacional: **11.06.2020** **WO20118115**

96 Fecha de presentación y número de la solicitud europea: **05.12.2019** **E 19893199 (0)**

97 Fecha y número de publicación de la concesión europea: **18.10.2023** **EP 3891274**

54 Título: **Composiciones y métodos para el tratamiento de la deficiencia de ornitina transcarbamilasa**

30 Prioridad:

06.12.2018 US 201862776302 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

06.03.2024

73 Titular/es:

ARCTURUS THERAPEUTICS, INC. (100.0%)
10628 Science Center Drive Suite 250
San Diego, CA 92121, US

72 Inventor/es:

PEREZ-GARCIA, CARLOS, G.;
TACHIKAWA, KIYOSHI;
MATSUDA, DAIKI y
CHIVUKULA, PADMANABH

74 Agente/Representante:

PONS ARIÑO, Ángel

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 960 692 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Composiciones y métodos para el tratamiento de la deficiencia de ornitina transcarbamilasa

5 Antecedentes

La ornitina transcarbamilasa (OTC) es una enzima mitocondrial presente en los mamíferos que desempeña un papel esencial en la desintoxicación del organismo del amoníaco tóxico. El ARNm de OTC tiene un péptido de señalización mitocondrial (PSM) que es fundamental para redirigir el prepeptido naciente del citosol a las mitocondrias. La proteína OTC existe como precursora en el citosol, la presencia del PSM redirige el propéptido a las mitocondrias, donde experimenta la escisión del péptido de señalización y el suministro de la proteína funcional a la matriz mitocondrial. La OTC es una de las seis enzimas que desempeñan un papel en la descomposición de proteínas y la eliminación de amoníaco del organismo, un proceso conocido como ciclo de la urea, un proceso metabólico que se produce en los hepatocitos. La OTC es responsable de convertir el carbamilo fosfato y la ornitina en citrulina.

La deficiencia de la enzima OTC produce una acumulación excesiva de nitrógeno, en forma de amoníaco (hiperamoniaquemia), en la sangre. El exceso de amoníaco, que es una neurotoxina, viaja al sistema nervioso central a través de la sangre, lo produce los síntomas y hallazgos físicos asociados a la deficiencia de OTC. Estos síntomas pueden incluir vómitos, rechazo a comer, letargo progresivo o coma. Si no se trata, un episodio de hiperamoniaquemia puede avanzar a coma y a complicaciones potencialmente mortales.

La gravedad y la edad de aparición de la deficiencia de OTC varían de una persona a otra, incluso dentro de la misma familia. Una forma grave del trastorno afecta a algunos lactantes, normalmente hombres, poco después del nacimiento (período neonatal). Una forma más leve del trastorno afecta a algunos niños más adelante en la infancia. Tanto hombres como mujeres pueden desarrollar síntomas de deficiencia de OTC durante la infancia. Actualmente, el tratamiento de la deficiencia de OTC tiene como objetivo prevenir la formación excesiva de amoníaco o eliminar el exceso de amoníaco durante un episodio de hiperamoniaquemia. La terapia prolongada para la deficiencia de OTC combina restricciones dietéticas y la estimulación de métodos alternativos para convertir y excretar nitrógeno del organismo (terapia de rutas alternativas).

Las restricciones dietéticas en personas con deficiencia de OTC tienen como objetivo limitar la cantidad de ingesta de proteínas para impedir el desarrollo de un exceso de amoníaco. Sin embargo, un lactante afectado debe ingerir suficientes proteínas para garantizar un crecimiento adecuado. Los lactantes con deficiencia de OTC reciben una dieta baja en proteínas, dieta hipercalórica complementada con aminoácidos esenciales.

Además de las restricciones dietéticas, las personas con deficiencia de OTC se tratan con medicamentos que estimulan la eliminación de nitrógeno del organismo. Estos medicamentos proporcionan un método alternativo al ciclo de la urea para convertir y eliminar los desechos de nitrógeno. Estos medicamentos son desagradables para muchos pacientes y a menudo se administran a través de una sonda que se coloca en el estómago a través de la pared abdominal (sonda de gastrostomía) o una sonda estrecha que llega al estómago a través de la nariz (sonda nasogástrica).

En los casos en los que no haya mejoría o en los casos en los que se desarrolle un coma hiperamoniaquémico, puede ser necesaria la eliminación de desechos filtrando la sangre del individuo afectado a través de una máquina (hemodiálisis). La hemodiálisis también se utiliza para tratar a lactantes, niños y adultos a quienes se les diagnostica por primera vez deficiencia de OTC durante un coma hiperamoniaquémico.

En algunos casos, el trasplante de hígado, puede ser una opción de tratamiento adecuada. El trasplante de hígado puede curar la hiperamoniaquemia en la deficiencia de OTC. Sin embargo, esta cirugía es arriesgada y puede provocar complicaciones posoperatorias. Además, después del trasplante de hígado, los pacientes deberán seguir un régimen de medicación durante toda su vida para la inmunosupresión.

Siguen siendo necesarios nuevos enfoques y terapias para el tratamiento de la deficiencia de la enzima OTC. Se necesitan estrategias que superen los inconvenientes y limitaciones asociados, por ejemplo, a la terapia génica. Una mala estabilidad, conseguir suficiente OTC en las mitocondrias y el suministro eficaz a las células diana, siguen siendo retos.

Sumario

La presente divulgación proporciona una proteína OTC humana modificada de SEQ ID NO: 4 que tiene propiedades mejoradas para el tratamiento de la deficiencia de OTC en un paciente. La invención se define por un polinucleótido que codifica una proteína ornitina transcarbamilasa (OTC) que comprende una secuencia de aminoácidos de SEQ ID NO: 4 y que tiene actividad enzimática OTC, en donde el polinucleótido es un ARNm que comprende una secuencia de SEQ ID NO: 221, en donde T está sustituido por U. Además, la SEQ ID NO: 4 se ha modificado a partir de OTC de tipo silvestre para eliminar uno o más sitios de ubiquitinación previstos, lo que da como resultado una proteína que es menos susceptible a la ubiquitinación y degradación por las ubiquitina ligasas. Sin embargo, la proteína OTC

modificada de SEQ ID NO: 4 conserva la actividad enzimática catalítica de la OTC humana de tipo silvestre. La eliminación de sitios de ubiquitinación previstos comprende preferentemente reemplazar restos del extremo N que se ha observado que favorecen la ubiquitinación, tales como asparagina, arginina, leucina, lisina o fenilalanina por restos del extremo N que se ha observado que son estabilizadores contra la ubiquitinación, tales como alanina, glicina, metionina, serina, treonina, valina y prolina. La estabilización de la proteína OTC modificada de SEQ ID NO: 4 de esta manera es particularmente ventajosa para preservar la estabilidad de la proteína OTC modificada durante su transporte desde el citosol a las mitocondrias donde ejerce su actividad enzimática.

Preferentemente, la construcción de ARNm heterólogo comprende una caperuza en 5', una UTR en 5', una UTR en 3', un marco abierto de lectura que codifica una proteína modificada de SEQ ID NO: 4 y una cola de poli A en 3'. Preferentemente, la UTR en 5' derivada de un gen expresado por *Arabidopsis thaliana*. Preferentemente, la UTR en 5' derivada de un gen expresado por *Arabidopsis thaliana* se encuentra en la Tabla 2.

Las construcciones de ARNm descritas en el presente documento proporcionan una expresión de alta eficacia de las proteínas OTC descritas en el mismo. La expresión puede ser *in vitro*, *ex vivo* o *in vivo*.

La presente divulgación también proporciona composiciones farmacéuticas que comprenden las secuencias de ARNm descritas en el presente documento y métodos para tratar la deficiencia de ornitina transcarbamilasa (OTC) mediante la administración de las composiciones farmacéuticas que comprenden las secuencias de ARNm descritas en el presente documento a un paciente que lo necesite, en donde la proteína OTC comprende una secuencia de aminoácidos de SEQ ID NO: 4 y que tiene actividad enzimática OTC, en donde el polinucleótido es un ARNm que comprende una secuencia de SEQ ID NO: 221, en donde T está sustituido por U.

Breve descripción de los dibujos

Las figuras 1A-B muestran diagramas de dispersión que ilustran la expresión de la proteína ornitina transcarbamilasa (OTC) en líneas celulares de hepatocitos Hepa1,6 (ratón) y Hep3B (humano) a las 24 horas (FIG. 1A) y 48 horas (FIG. 1B) utilizando ensayos Western en células (WEC).

Las figuras 2A-B muestran diagramas de dispersión que ilustran la correlación de los compuestos de estabilidad de proteínas seleccionados en células Hepa1,6 (FIG. 2A) y células Hep3B (FIG. 2B) a las 24 h en la Ronda 1.

Las figuras 3A-B muestran diagramas de dispersión que ilustran la correlación de los compuestos de estabilidad de proteínas seleccionados en hepatocitos primarios humanos a las 24 h y 48 h en la Ronda 2 (compuestos recientemente optimizados basados en la Ronda 1) como se muestra en las figuras 3A y 3B.

Las figuras 4A-B muestran diagramas de dispersión que ilustran la correlación de los compuestos de estabilidad de proteínas seleccionados en hepatocitos primarios humanos a las 24 h y 48 h en la Ronda 3 (compuestos recientemente optimizados basados en las rondas 1 y 2) como se muestra en las figuras 4A y 4B.

La figura 5 es un gráfico que ilustra los niveles de expresión de la proteína OTC en hepatocitos primarios humanos transfectados con los ARNm de OTC iniciales. 1799.1 es un ARNm que tiene la secuencia de SEQ ID NO: 175 en donde el 100 % de las uridinas en SEQ ID NO: 175 son N¹-metilpseudouridina (N1MPU).

Las figuras 6A-B muestran gráficos de barras que representan la evolución temporal de niveles de expresión de OTC en ratones *spf/ash* (*sparse-fur/abnormal skin and hair*, piel y pelo escaso/anómalo) a los que se les administraron dosis de 10 mg/kg con epítopos de ARNm específicos de OTC humana (figura 6A) o epítopos de ARNm específicos de OTC de ratón (figura 6B).

La figura 7 es un gráfico de barras que representa los niveles de expresión de OTC en ratones *spf/ash* a los que se les administraron 3 mg/kg de ARNm de OTC usando dos químicas diferentes en las que el 100 % de las uridinas son N¹-metilpseudouridina (N1MPU) y el 100 % de las uridinas son 5-metoxiuridina (5MeOU).

La figura 8 es un gráfico que representa los niveles de expresión de OTC en ratones Balb/c que recibieron dosis de ARNm de OTC a tres dosis diferentes y usando dos químicas diferentes (N1MPU y 5MeOU).

La figura 9 es un gráfico que representa los niveles de orotato en orina medidos en ratones *spf/ash* que recibieron dosis de ARNm 1799.7 de OTC a tres dosis diferentes: 0,3 mg/kg, 1 mg/kg y 3 mg/kg. 1799.1 es un ARNm que tiene la secuencia de SEQ ID NO: 175 en donde el 100 % de las uridinas en SEQ ID NO: 175 son 5MeOU.

La figura 10 es un diagrama de dispersión que compara los niveles de expresión de OTC humana y orotato urinario a las 96 h en ratones *spf/ash* que recibieron dosis de los ARNm de OTC a 1 mg/kg y 3 mg/kg usando dos químicas diferentes (N1MPU y 5MeOU).

La figura 11 es una transferencia Western que ilustra los niveles de expresión de proteínas de los ARNm de OTC en ratones *spf/ash* que recibieron dosis de 1 mg/kg y 3 mg/kg con los ARNm de OTC iniciales.

La figura 12 es una transferencia Western que ilustra los niveles de expresión de proteínas en fracciones mitocondriales frente a citosólicas de ratones *spf/ash* tratados con los ARNm de OTC iniciales.

- 5 La figura 13 es un diagrama que ilustra la evolución temporal de la expresión de los niveles de orotato urinario en ratones *spf/ash* tratados con ARNm de OTC inicial.

La figura 14 es un diagrama que ilustra la supervivencia de ratones con deficiencia de OTC (*spf/ash*) con una dieta rica en proteínas durante el tratamiento con tres dosis diferentes de ARNm de OTC 1799.7.

- 10 La figura 15 es un diagrama que ilustra los niveles de expresión de OTCh en ratones macho C57BL/6 que recibieron dosis de los ARNm de OTC (2262) que tenían diferentes modificaciones.

Descripción detallada

- 15 Definiciones:

La expresión "ornitina transcarbamilasa", como se usa indistintamente en el presente documento con "OTC" u "OTCh" u "OTC_HUMANA", generalmente se refiere a la proteína humana asociada a UniProtKB-P00480. En el presente documento, la secuencia de aminoácidos de la proteína OTC humana de tipo silvestre se representa como SEQ ID NO: 3.

25 La expresión "ácido nucleico", en su sentido más amplio, incluye cualquier compuesto y/o sustancia que comprenda un polímero de nucleótidos. Estos polímeros suelen denominarse polinucleótidos. Los ácidos nucleicos o polinucleótidos ilustrativos descritos en el presente documento incluyen, pero sin limitación, ácidos ribonucleicos (ARN), ácidos desoxirribonucleicos (ADN), ácidos treosanucleicos (ATN), ácidos glicolnucleicos (AGN), ácidos peptidonucleicos (APN), ácidos nucleicos bloqueados (ANB, incluyendo ANB que tiene una configuración β -D-ribo, α -ANB que tiene una configuración α -L-ribo (un diastereómero de ANB), 2'-amino-ANB que tiene una funcionalización 2'-amino y 2'-amino- α -ANB que tiene una funcionalización 2'-amino) o híbridos de los mismos.

30 Como se utiliza en el presente documento, el término "polinucleótido" se usa generalmente para referirse a un ácido nucleico (p. ej., ADN o ARN). Cuando se hace referencia específicamente al ARN, tal como al ARNm, puede utilizarse el término polirribonucleótido. Los términos polinucleótido, polirribonucleótido, ácido nucleico, ácido ribonucleico, ADN, ARN, ARNm y similares, incluyen moléculas que pueden estar compuestas de restos estándar o no modificados; restos no estándar o modificados (p. ej., análogos); y mezclas de restos estándar y no estándar (p. ej., análogos). En determinadas realizaciones, un polinucleótido o un polirribonucleótido es un polinucleótido o un polirribonucleótido modificado. En el contexto de la presente divulgación, para cada secuencia de ARN (polirribonucleótido) enumerada en este documento, se contempla la secuencia de ADN (polidesoxirribonucleótido o polinucleótido) correspondiente y viceversa. "Polinucleótido" puede usarse indistintamente con el "oligómero". Las secuencias de polinucleótidos que se muestran en el presente documento son de izquierda a derecha, de 5' a 3', a menos que se indique lo contrario.

40 Como se utiliza en el presente documento, la expresión "ARN mensajero" (ARNm) se refiere a cualquier polinucleótido que codifica una proteína o polipéptido de interés y que es capaz de traducirse para producir la proteína o polipéptido de interés codificado *in vitro*, *in vivo*, *in situ* o *ex vivo*.

45 Como se utiliza en el presente documento, el término "traducción" es el proceso en el que los ribosomas crean polipéptidos. En la traducción, el ARN mensajero (ARNm) se descodifica mediante los ARN de transferencia (ARNt) en un complejo de ribosomas para producir una cadena de aminoácidos específica o polipéptido. La región codificante de una secuencia de polinucleótidos (ADN o ARN), también conocida como secuencia codificante o CDS, es capaz de convertirse en una proteína o en un fragmento de la misma mediante el proceso de traducción.

50 Como se utiliza en el presente documento, la expresión "optimizada con codones" significa una secuencia codificante natural (o variante diseñada expresamente de una natural) que se ha rediseñado eligiendo diferentes codones sin alterar la secuencia de aminoácidos de la proteína codificada. La secuencia optimizada con codones puede aumentar los niveles de expresión de las proteínas (Gustafsson *et al.*, *Codon bias and heterologous protein expression*. 2004, Trends Biotechnol 22: 346-53) de las proteínas codificadas, entre otras ventajas. Se ha observado que, variables tales como alto índice de adaptación de codones (IAC), método de LowU, estructuras secundarias de ARNm, secuencias reguladoras en cis, contenido de GC y muchas otras variables similares, se correlacionan de alguna manera con los niveles de expresión de proteínas (Villalobos *et al.*, *Gene Designer: a synthetic biology tool for constructing artificial DNA segments*. 2006, BMC Bioinformatics 7:285). El método de alto IAC (índice de adaptación de codones) elige un codón sinónimo utilizado con mayor frecuencia para toda una secuencia codificante de proteínas. El codón más utilizado para cada aminoácido se deduce de 74.218 genes codificantes de proteínas del genoma humano. El método Low U se dirige únicamente a codones que contienen U que pueden reemplazarse por un codón sinónimo con menos residuos de U. Si hay algunas opciones para el reemplazo, se seleccionará el codón utilizado con más frecuencia. Los codones restantes de la secuencia no se modifican mediante el método de Low U. Este método se puede utilizar junto con los ARNm divulgados para diseñar secuencias codificantes que se van a sintetizar con, por ejemplo, 5-

metoxiuridina o N¹-metilpseudouridina.

Como se utiliza en el presente documento, "modificado" se refiere a un cambio en el estado o estructura de una molécula desvelada en el presente documento. La molécula se puede cambiar de muchas maneras, incluidas químicamente, estructural o funcionalmente. Preferentemente, un polinucleótido o polipéptido de la divulgación se modifica en comparación con la forma nativa del polinucleótido o polipéptido o en comparación con una secuencia de polipéptido o secuencia de polinucleótidos de referencia. Por ejemplo, el ARNm desvelado en el presente documento puede modificarse mediante optimización con codones o mediante la inserción de nucleósidos o nucleótidos no naturales. Los polipéptidos pueden modificarse, por ejemplo, mediante delecciones o sustituciones de aminoácidos específicas del sitio para alterar las propiedades del polipéptido.

Como se utiliza en el presente documento, el término "homología" se refiere a la relación global entre moléculas poliméricas, p. ej., entre moléculas de ácido nucleico (p. ej., moléculas de ADN y/o moléculas de ARN) y/o entre moléculas polipeptídicas. En algunas realizaciones, se considera que las moléculas poliméricas son "homólogas" entre sí si sus secuencias son al menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 99 % idénticas o similares. El término "homólogo" se refiere necesariamente a una comparación entre al menos dos secuencias (secuencias de polinucleótidos o de polipéptidos). De conformidad con la presente divulgación, se considera que dos secuencias de polinucleótidos son homólogas si los polipéptidos que codifican son al menos aproximadamente 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o incluso 99 % durante al menos un tramo de al menos aproximadamente 20 aminoácidos. En algunas realizaciones, las secuencias de polinucleótidos homólogas se caracterizan por la capacidad de codificar un tramo de al menos 4-5 aminoácidos especificados de forma única. Para secuencias de polinucleótidos de menos de 60 nucleótidos de longitud, la homología está determinada por la capacidad de codificar un tramo de al menos 4-5 aminoácidos especificados de forma única. De conformidad con la presente divulgación, se considera que dos secuencias de proteínas son homólogas si las proteínas son al menos aproximadamente 50 %, 60 %, 70 %, 80 % o 90 % idénticas durante al menos un tramo de al menos aproximadamente 20 aminoácidos.

Como se utiliza en el presente documento, el término "identidad" se refiere a la relación global entre moléculas poliméricas, p. ej., entre moléculas oligonucleotídicas (p. ej., moléculas de ADN y/o moléculas de ARN) y/o entre moléculas polipeptídicas. El cálculo del porcentaje de identidad de dos secuencias de polinucleótidos, por ejemplo, se puede realizar alineando las dos secuencias con fines de comparación óptima. En determinadas realizaciones, la longitud de una secuencia de referencia alineada con fines comparativos es de al menos 30 %, al menos 40 %, al menos 50 %, al menos 60 %, al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 % o 100 % de la longitud de la secuencia de referencia. A continuación, se comparan los nucleótidos en las posiciones de nucleótidos correspondientes. Cuando una posición en la primera secuencia está ocupada por el mismo nucleótido que la posición correspondiente en la segunda secuencia, entonces las moléculas son idénticas en esa posición. El porcentaje de identidad entre las dos secuencias depende del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos y la longitud de cada hueco, que es necesario introducir para una alineación óptima de las dos secuencias. La comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias pueden realizarse usando un algoritmo matemático. Los métodos comúnmente empleados para determinar el porcentaje de identidad entre secuencias incluyen, pero no se limitan a los divulgados en Carillo, H. y Lipman, D., *SIAM J Applied Math.*, 48:1073 (1988); Techniques for determining identity are codified in publicly available computer programs. Los programas informáticos ilustrativos para determinar la homología entre dos secuencias incluyen, pero sin limitación, el paquete de programas GCG, Devereux, J., *et al.*, *Nucleic Acids Research*, 12(1), 387 (1984), BLASTP, BLASTN y FASTA Altschul, S. F. *et al.*, *J. Molec. Biol.*, 215, 403 (1990).

Una "cantidad eficaz" de la secuencia de ARNm que codifica una proteína de marco abierto de lectura (ORF, *open reading frame*) o una composición correspondiente de la misma es generalmente esa cantidad de ARNm que proporciona una producción eficaz de proteína de ORF en una célula. Preferentemente, la producción de proteínas utilizando una composición de ARNm descrita en el presente documento es más eficaz que una composición que contiene un ARNm de tipo silvestre correspondiente que codifica una proteína de ORF. Se puede demostrar una mayor eficacia mediante una mayor transfección celular (es decir, el porcentaje de células transfectadas con el ácido nucleico), mayor traducción de proteínas del ácido nucleico, disminución de la degradación del ácido nucleico (como se ha demostrado, p. ej., por una mayor duración de la traducción de proteínas a partir de un ácido nucleico modificado) o una respuesta inmunitaria innata reducida de la célula hospedadora. Cuando se hace referencia a una proteína de ORF descrita en el presente documento, una cantidad eficaz es aquella cantidad de proteína de ORF que supera una deficiencia de proteína de ORF en una célula.

Como se utiliza en el presente documento, la expresión "*in vitro*" se refiere a acontecimientos que se producen en un entorno artificial, p. ej., en un tubo de ensayo o recipiente de reacción, en cultivo celular, en una placa de Petri, etc., en lugar de dentro de un organismo (p. ej., animal, planta o microbio).

Como se utiliza en el presente documento, la expresión "*in vivo*" se refiere a acontecimientos que se producen dentro de un organismo (p. ej., animal, planta o microbio, o célula o tejido del mismo).

Como se utiliza en el presente documento, el término "aislada" se refiere a una sustancia o entidad que se ha separado

de al menos algunos de los componentes con los que estaba asociada (ya sea en la naturaleza o en un entorno experimental). Las sustancias aisladas pueden tener distintos niveles de pureza con respecto a las sustancias de las que se han asociado. Las sustancias y/o entidades aisladas pueden separarse de al menos aproximadamente el 10 %, aproximadamente el 20 %, aproximadamente el 30 %, aproximadamente el 40 %, aproximadamente el 50 %, aproximadamente el 60 %, aproximadamente el 70 %, aproximadamente el 80 %, aproximadamente el 90 % o más de los otros componentes con los que estaban asociados inicialmente. En algunas realizaciones, los agentes aislados tienen una pureza de más de aproximadamente el 80 %, aproximadamente el 85 %, aproximadamente el 90 %, aproximadamente el 91 %, aproximadamente el 92 %, aproximadamente el 93 %, aproximadamente el 94 %, aproximadamente el 95 %, aproximadamente el 96 %, aproximadamente el 97 %, aproximadamente el 98 %, aproximadamente el 99 % o una pureza de más de aproximadamente el 99 %. Como se utiliza en el presente documento, una sustancia es "pura" si carece sustancialmente de otros componentes. Sustancialmente aislado: por "sustancialmente aislado" se entiende que el compuesto está sustancialmente separado del entorno en el que se formó o se detectó. La separación parcial puede incluir, por ejemplo, una composición enriquecida con el compuesto descrito en el presente documento. La separación sustancial puede incluir composiciones que contengan al menos aproximadamente el 50 %, al menos aproximadamente el 60 %, al menos aproximadamente el 70 %, al menos aproximadamente el 80 %, al menos aproximadamente el 90 %, al menos aproximadamente el 95 %, al menos aproximadamente el 97 % o al menos aproximadamente el 99 % en peso del compuesto descrito en el presente documento, o una sal del mismo. Los métodos para aislar compuestos y sus sales son rutinarios en la técnica.

Como se utiliza en el presente documento, el término "sujeto" o "paciente" se refiere a cualquier organismo al que se le pueda administrar una composición de acuerdo con la presente divulgación, p. ej., con fines experimentales, diagnósticos, profilácticos y/o terapéuticos. Los sujetos típicos incluyen animales (p. ej., mamíferos tales como ratones, ratas, conejos, primates no humanos y seres humanos) y/o plantas. Preferentemente "paciente" se refiere a un sujeto humano que puede buscar o necesitar tratamiento, que requiere tratamiento, que está recibiendo tratamiento, que recibirá tratamiento, o un sujeto que está al cuidado de un profesional capacitado para una enfermedad o afección particular.

La expresión "farmacéuticamente aceptable" se emplea en el presente documento para referirse a aquellos compuestos, materiales, composiciones y/o formas farmacéuticas que son, dentro del alcance del buen criterio médico, adecuados para su uso en contacto con los tejidos de seres humanos y animales sin excesiva toxicidad, irritación, respuesta alérgica u otro problema o complicación, acorde con una relación beneficio/riesgo razonable.

Como se utiliza en el presente documento, la expresión "prevenir" se refiere a retrasar parcial o completamente la aparición de una infección, enfermedad, trastorno y/o afección; retrasar parcial o completamente la aparición de uno o más síntomas, características o manifestaciones clínicas de una particular infección, enfermedad, trastorno y/o afección; retrasar parcial o completamente la aparición de uno o más síntomas, características o manifestaciones de una particular infección, enfermedad, trastorno y/o afección; retrasar parcial o completamente el avance de una infección, una particular enfermedad, trastorno y/o afección; y/o disminuir el riesgo de desarrollar una patología asociada a la infección, la enfermedad, al trastorno y/o a la afección.

Como se utiliza en el presente documento, el término "sustancialmente" se refiere a la condición cualitativa de mostrar la extensión o el grado total o casi total de una característica o propiedad de interés. Un experto en la materia biológica comprenderá que los fenómenos biológicos y químicos rara vez, o nunca, llegan a completarse y/o transcurrir hasta completarse o consiguen o impiden un resultado absoluto. Por lo tanto, en el presente documento, el término "sustancialmente" se usa para captar la posible falta de exhaustividad intrínseca a muchos fenómenos biológicos y químicos.

Como se utiliza en el presente documento, la expresión "cantidad terapéuticamente eficaz" significa una cantidad de un agente que se va a suministrar (p. ej., ácido nucleico, proteína o péptido, fármaco, agente terapéutico, agente de diagnóstico, agente profiláctico, etc.) que es suficiente, cuando se administra a un sujeto que padece o es susceptible a una infección, enfermedad, trastorno y/o afección, para tratar, mejorar los síntomas de, diagnosticar, prevenir y/o retrasar la aparición de la infección, enfermedad, trastorno y/o afección.

Como se utiliza en el presente documento, una "dosis diaria total" es una cantidad dada o prescrita para un periodo de 24 horas. Puede administrarse como una sola dosis unitaria.

Como se utiliza en el presente documento, el término "tratar" se refiere a aliviar parcial o completamente, mejorar, aumentar, mitigar, retrasar el inicio de, inhibir el avance de, reducir la gravedad de y/o reducir la incidencia de uno o más síntomas o características de deficiencia de OTC. El tratamiento se puede administrar a un sujeto que no muestre signos de deficiencia de OTC y/o que solo muestre signos tempranos de deficiencia de OTC con el fin de disminuir el riesgo de desarrollar una patología asociada a la enfermedad, trastorno y/o afección.

Como se utiliza en el presente documento, los términos "transfectar" o "transfección" significan la introducción intracelular de un ácido nucleico en una célula, o preferentemente en una célula diana. El ácido nucleico introducido puede mantenerse estable o transitoriamente en la célula diana. El término "eficacia de transfección" se refiere a la cantidad relativa de ácido nucleico que capta la célula diana que está sujeta a transfección. En la práctica, la eficacia

de transfección se estima mediante la cantidad de un producto de ácido nucleico indicador expresado por las células diana después de la transfección. Se prefieren composiciones con altas eficacias de transfección y en particular aquellas composiciones que minimizan los efectos adversos mediados por la transfección de células y tejidos no diana.

- 5 Como se utiliza en el presente documento, la expresión "célula diana" se refiere a una célula o a un tejido al que se va a dirigir una composición de la divulgación o que es su diana. En algunas realizaciones, las células diana tienen deficiencia de una proteína o enzima de interés. Por ejemplo, donde se desea suministrar un ácido nucleico a un hepatocito, el hepatocito representa la célula diana. En algunas realizaciones, los ácidos nucleicos y las composiciones de la presente divulgación transfectan las células diana de forma discriminatoria (es decir, no transfectan células no diana). Las composiciones y métodos de la presente divulgación se pueden preparar para dirigirse preferentemente a una variedad de células diana, que incluyen, pero sin limitación, hepatocitos, células epiteliales, células hematopoyéticas, células epiteliales, células endoteliales, células pulmonares, células óseas, células madre, células mesenquimatosas, células neuronales (p. ej., en las meninges, astrocitos, neuronas motoras, células de los ganglios de la raíz dorsal y de las neuronas motoras del asta anterior), células fotorreceptoras (por ejemplo, bastones y conos),
 10 células epiteliales pigmentadas de la retina, células secretoras, células cardíacas, adipocitos, células del músculo liso vascular, cardiomiocitos, células musculoesqueléticas, células beta, células hipofisarias, células del revestimiento sinovial, células ováricas, células testiculares, fibroblastos, linfocitos B, linfocitos T, reticulocitos, leucocitos, granulocitos y células tumorales.
- 20 Después de la transfección de una o más células diana mediante las composiciones y ácidos nucleicos descritos en el presente documento, se puede estimular preferentemente la expresión de la proteína codificada por dicho ácido nucleico y aumentar la capacidad de dichas células diana para expresar la proteína de interés. Por ejemplo, la transfección de una célula diana con un ARNm de OTC permitirá la expresión del producto proteico de OTC después de la traducción del ácido nucleico. Los ácidos nucleicos de las composiciones y/o métodos proporcionados en el
 25 presente documento codifican preferentemente un producto (p. ej., una proteína, enzima, polipéptido, péptido, ARN funcional y/o molécula antisentido), y preferentemente codifican un producto cuya producción se desea realizar *in vivo*.

30 Como se utiliza en el presente documento, "una actividad enzimática de proteína de OTC" se refiere a la actividad enzimática que cataliza la reacción entre el carbamoil fosfato y la ornitina para formar citrulina como parte del ciclo de la urea en los mamíferos.

35 Como se utiliza en el presente documento, el término "alrededor de" o "aproximadamente" aplicado a uno o más valores de interés, se refiere a un valor que es similar a un valor de referencia indicado. En determinadas realizaciones, el término "alrededor de" o "aproximadamente" se refiere a un intervalo de valores que se encuentra en el 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 %, 1 % o menos en cualquier dirección (mayor o menor que) del valor de referencia indicado, a menos que se indique de otro modo o resulte evidente de otro modo a partir del contexto (excepto cuando tal número exceda el 100 % de un posible valor).

Secuencias de polinucleótidos

40 La presente divulgación proporciona métodos y composiciones mejorados para el tratamiento de la deficiencia de ornitina transcarbamilasa (OTC) usando, por ejemplo, terapia de ARNm. La presente divulgación proporciona métodos para tratar la deficiencia de ornitina transcarbamilasa (OTC), que comprenden administrar a un sujeto que necesita tratamiento una composición que comprende una secuencia de ARNm descrita en el presente documento que codifica
 45 una proteína ornitina transcarbamilasa (OTC) humana, formas modificadas de proteína OTC humana o fragmentos activos de proteína OTC en una dosis eficaz y un intervalo de administración tal que al menos un síntoma o característica de la deficiencia de OTC se reduzca en intensidad, gravedad o frecuencia, o se retrase su aparición. La presente divulgación también proporciona proteínas OTC modificadas codificadas por las secuencias de ARNm en las que las proteínas OTC modificadas tienen propiedades mejoradas tales como estabilidad y resistencia mejoradas a la degradación de proteínas y una semivida aumentada en comparación con las proteínas OTC humanas de tipo silvestre.

50 Preferentemente, la administración de una composición de ARNm descrita en el presente documento da como resultado una mayor expresión o actividad de la proteína OTC del sujeto en comparación con un nivel de control.
 55 Preferentemente, el nivel de control es un nivel de actividad o expresión de la proteína OTC en suero basal en el sujeto antes del tratamiento y/o el nivel de control es indicativo del nivel promedio de expresión o actividad de la proteína OTC en suero en pacientes con OTC sin tratamiento.

60 Preferentemente, la administración de un ARNm descrito en la presente composición da como resultado un nivel reducido de ácido orótico urinario en el sujeto en comparación con un nivel de ácido orótico de control. Preferentemente, el nivel de ácido orótico de control es un nivel de ácido orótico urinario basal en el sujeto antes del tratamiento y/o el nivel de ácido orótico de control es un nivel de referencia indicativo del nivel promedio de ácido orótico urinario en pacientes con OTC sin tratamiento.

65 Preferentemente, las proteínas OTC codificadas por el ARNm descrito en el presente documento se producen a partir de una construcción de ARNm heterólogo que comprende un marco abierto de lectura (ORF) también denominado en

el presente documento "secuencia codificante" (CDS, *coding sequence*) que codifica una proteína OTC. Preferentemente, la secuencia codificante está optimizada con codones. Preferentemente, la secuencia codificante se optimiza para que tenga un mínimo teórico posible de uridina para codificar una proteína OTC. Preferentemente, las construcciones de ARNm descritas en el presente documento comprenden una o más de las siguientes características:

- 5 una caperuza en 5'; una UTR en 5', una secuencia potenciadora de UTR en 5', una secuencia de Kozak o una secuencia de Kozak parcial, una UTR en 3', un marco abierto de lectura que codifica una proteína OTC y una cola de poli A. Preferentemente, las construcciones de ARNm descritas en el presente documento pueden proporcionar expresión de alta eficacia de una proteína OTC. La expresión puede ser *in vitro*, *ex vivo* o *in vivo*.
- 10 Preferentemente, una proteína OTC humana codificada por un ARNm descrito en el presente documento, un polinucleótido que codifica una proteína ornitina transcarbamilasa (OTC) que comprende una secuencia de aminoácidos de SEQ ID NO: 4 y que tiene actividad enzimática OTC, en donde el polinucleótido es un ARNm que comprende una secuencia de SEQ ID NO: 221, en donde T está sustituido por U.
- 15 Preferentemente, un ARNm descrito en el presente documento comprende además una secuencia de región no traducida (UTR, *untranslated region*) en 5'. Como se entiende en la técnica, la UTR en 5' y/o 3' puede afectar a la estabilidad de un ARNm o la eficacia de su traducción. La UTR en 5' puede derivar de una molécula de ARNm conocida en la técnica que es relativamente estable (p. ej., histona, tubulina, globina, gliceraldehído 1-fosfato deshidrogenasa (GAPDH), actina o enzimas del ciclo del ácido cítrico) para aumentar la estabilidad del oligómero traducible. En otras realizaciones, una secuencia UTR en 5' puede incluir una secuencia parcial de un gen inmediato-temprano 1 (IE1, *immediate-early 1*) de citomegalovirus (CMV).
- 20

Preferentemente, la UTR en 5' comprende una secuencia seleccionada de las UTR en 5' de IL-6 humana, alanina aminotransferasa 1, apolipoproteína E humana, cadena alfa de fibrinógeno humano, transtirretina humana, haptoglobina humana, alfa-1-antiquimotripsina humana, antitrombina humana, alfa-1-antitripsina humana, albúmina humana, beta globina humana, complemento C3 humano, complemento C5 humano, SynK (proteína del canal de potasio de las membranas tilacoides derivada del género de cianobacterias, *Synechocystis sp.*), beta globina de ratón, albúmina de ratón y un virus del grabado del tabaco, o fragmentos de cualquiera de los anteriores. Preferentemente, la UTR en 5' deriva de un virus del grabado del tabaco (TEV, *tobacco etch virus*). Preferentemente, un ARNm descrito en el presente documento comprende una secuencia UTR en 5' que deriva de un gen expresado por *Arabidopsis thaliana*. Preferentemente, la secuencia UTR en 5' de un gen expresado por *Arabidopsis thaliana* es AT1G58420. Las secuencias UTR en 5' preferidas comprenden las SEQ ID NO: 5-10, 125-127 y 230-250: como se muestra en la Tabla 2.

35

Tabla 2
Secuencias de UTR en 5'

Nombre	Secuencia	SEQ ID NO:
EV	UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCA UUCUACUUCUAUUGCAGCAAUUUUAAAUCAUUUCUUUUAAAAGCA AAAGCAAUUUUCUGAAAAUUUUCACCAUUUACGAACGAUAG	SEQ ID NO: 5
AT1G58420	AUUUAUUACAUCAAAACAAAAAGCCGCCA	SEQ ID NO: 6
ARC5-2	CUUAAGGGGGCGCUGCCUACGGAGGUGGCAGCCAUCUCCUUCU CGGCAUCAAGCUUACCAUGGUGCCCCAGGCCUGCUCUUGGUCC CGCUGCUGGUGUUCUUUUUUCUGCUUCGGCAAGUUCUUUUUUAUAC ACCAUCCCCGACAAGCUGGGGGCCGUGGAGCCCCAUCGACAUCCA CCACCUGUCCUGCCCCAACAAACCUCGUGGUCGAGGACGAGGGCU GCACCAACCUGAGCGGGUUCUCCUAC	SEQ ID NO: 7
VHC	UGAGUGUCGU ACAGCCUCCA GGCCCCCCCC UCCCGGGAGA GCCAUAGUGG UCUGCGGAACCGGUGAGUAC ACCGGAUUUG CCGGGAAGAC UGGGUCCUUU CUUGGAUAAA CCCACUCUAUGCCCGGCCAU UUGGGCGUGC CCCCAGCAAGA CUGCUAGCCG AGUAGUGUUG GGUUGCG	SEQ ID NO: 8
ALBÚMINA HUMANA	AAUUAUUGGUUAAAGAAGUAUAUUAGUGCUAAUUUCCCUCCG UUUGUCCUAGCUUUUCUCUUCUGUCAACCCACACGCCUUUGG CACA	SEQ ID NO: 9

ES 2 960 692 T3

(continuación)

Nombre	Secuencia	SEQ ID NO:
EMCV	CUCCCUCCCC CCCCCUAAC GUUACUGGCC GAAGCCGCUU GGAAUAAGGC CGGUGUGCGU UUGUCUAUUA GUUAUUUUCC ACCAUAUUGC CGUCUUUUGG CAAUGUGAGG GCCCGGAAAC CUGGCCCUGU CUUCUUGACG AGCAUUCCUA GGGGUCUUUC CCUCUCGCC AAAGGAAUGC AAGGUCUGUU GAAUGUCGUG AAGGAAGCAG UUCCUCUGGA AGCUUCUUGA AGACAAACAA CGUCUGUAGC GACCCUUUGC AGGCAGCGGA ACCCCCCACC UGGCGACAGG UGCCUCUGCG GCCAAAAGCC ACGUGUAUAA GAUACACCUG CAAAGGCGGC ACAACCCAG UGCCACGUUG UGAGUUGGAU AGUUGUGGAA AGAGUCAAU GGCUCUCCUC AAGCGUAUUC AACAAGGGGC UGAAGGAUGC CCAGAAGGUA CCCCAUUGUA UGGGAUCUGA UCUGGGGCCU CGGUGCACAU GCUUUACGUG UGUUUAGUCG AGGUUAAAAA ACGUCUAGGC CCCCGAACC ACGGGGACGU GGUUUUCCUU UGAAAAACAC GAUGAUAAU	SEQ ID NO: 10
AT1G67090	CACAAAGAGUAAAGAAGAACA	SEQ ID NO: 125
AT1G35720	AACACUAAAAGUAGAAGAAAA	SEQ ID NO: 126
AT5G45900	CUCAGAAAGAUAGAUCAGCC	SEQ ID NO: 127
AT5G61250	AACCAAUCGAAAGAAACCAA	SEQ ID NO: 230
AT5G46430	CUCUAAUCACCAGGAGUAAAA	SEQ ID NO: 231
AT5G47110	GAGAGAGAUUCUUAACAAAAA	SEQ ID NO: 232
AT1G03110	UGUGUAACAACAACAACA	SEQ ID NO: 233
AT3G12380	CCGCAGUAGGAAGAGAAAGCC	SEQ ID NO: 234
AT5G45910	AAAAAAAAAAGAAUUAUAAA	SEQ ID NO: 235
AT1G07260	GAGAGAAGAAAGAAGAAGACG	SEQ ID NO: 236
AT3G55500	CAAUUAAAAAUACUUACCAA	SEQ ID NO: 237
AT3G46230	GCAAACAGAGUAAGCGAAACG	SEQ ID NO: 238
AT2G36170	GCGAAGAAGACGAACGCAAAG	SEQ ID NO: 239
AT1G10660	UUAGGACUGUAUUGACUGGCC	SEQ ID NO: 240

(continuación)

Nombre	Secuencia	SEQ ID NO:
AT4G14340	AUCAUCGGAAUUCGGAAG	SEQ ID NO: 241
AT1G49310	AAAACAAAAGUUAAGCAGAC	SEQ ID NO: 242
AT4G14360	UUUAUCUCAAAUAAGAAGGCA	SEQ ID NO: 243
AT1G28520	GGUGGGGAGGUGAGAUUUCUU	SEQ ID NO: 244
AT1G20160	UGAUUAGGAAACUACAAAGCC	SEQ ID NO: 245
AT5G37370	CAUUUUUCAAUUCAUAAAAC	SEQ ID NO: 246
AT4G11320	UUACUUUUAAGCCCAACAAAA	SEQ ID NO: 247
AT5G40850	GGCGUGUGUGUGUGUUGUUGA	SEQ ID NO: 248
AT1G06150	GUGGUGAAGGGGAAGGUUAG	SEQ ID NO: 249
AT2G26080	UUGUUUUUUUUUGGUUUGGUU	SEQ ID NO: 250

Preferentemente, la secuencia UTR en 5' comprende la SEQ ID NO: 6 (AT1G58420).

- 5 Preferentemente, un ARNm descrito en el presente documento comprende una secuencia potenciadora de la traducción. Las secuencias potenciadoras de la traducción mejoran la eficacia de la traducción de un ARNm descrito en el presente documento y, por lo tanto, proporcionan una mayor producción de la proteína codificada por el ARNm. La región potenciadora de la traducción puede estar ubicada en la UTR 5' o 3' de una secuencia de ARNm. Como ejemplos de regiones potenciadoras de la traducción se incluyen regiones potenciadoras de origen natural de la UTR en 5' del TEV y la UTR en 3' de beta-globina de *Xenopus*. Las secuencias potenciadoras de UTR en 5' preferidas incluyen, pero sin limitación, las derivadas de los ARNm que codifican proteínas de choque térmico (HSP, *heat shock proteins*) humanas, incluidas HSP70-P2, HSP70-M1 HSP72-M2, HSP17.9 y HSP70-P1. Las secuencias potenciadoras de la traducción preferidas utilizadas de acuerdo con las realizaciones de la presente divulgación están representadas por las SEQ ID NOS 11-15 como se muestra en la Tabla 3.

15

Tabla 3
Potenciadores de UTR en 5'

Nombre	Secuencia	SEQ ID NO:
HSP70-P2	GUCAGCUUUCAAACUCUUUGUUUCUUGUUUGUUGAUUGAGAAUA	SEQ ID NO: 11
HSP70-M1	CUCUCGCCUGAGAAAAAAAUCCACGAACCAAUUUCUCAGCAACCAG CAGCACG	SEQ ID NO: 12
HSP72-M2	ACCUGUGAGGGUUCGAAGGAAGUAGCAGUGUUUUUUGUUCUAGA GGAAGAG	SEQ ID NO: 13
HSP17.9	ACACAGAAACAUUCGCAAAAACAAAUCCAGUAUCAAUUUCUUCU CUUUUUUUCAUAAUUCGCAAGAC	SEQ ID NO: 14
HSP70-P1	CAGAAAAAUUUGCUACAUUGUUUCACAAACUCAAUAUUAUUCAU UUAUUU	SEQ ID NO: 15

- 20 Preferentemente, un ARNm descrito en el presente documento comprende una secuencia de Kozak. Como se entiende en la técnica, una secuencia de Kozak es una secuencia de consenso corta centrada alrededor del sitio de inicio de la traducción de los ARNm eucariotas que permite el inicio eficiente de la traducción del ARNm. La maquinaria de traducción ribosómica reconoce el codón de inicio AUG en el contexto de la secuencia de Kozak. Una secuencia

de Kozak, puede insertarse cadena arriba de la secuencia codificante de la OTC, secuencia abajo de una UTR en 5' o insertarse cadena arriba de la secuencia codificante de la OTC y cadena abajo de una UTR en 5'. Preferentemente, un ARNm descrito en el presente documento comprende una secuencia de Kozak que tiene la secuencia de aminoácidos GCCACC (**SEQ ID NO: 23**). Preferentemente, un ARNm descrito en el presente documento comprende una secuencia de Kozak parcial "p" que tiene la secuencia de aminoácidos GCCA. (**SEQ ID NO: 24**).

Preferentemente, un ARNm descrito en el presente documento comprende una UTR en 3'. Preferentemente, la UTR en 3' comprende una secuencia seleccionada de las UTR en 3' de alanina aminotransferasa 1, apolipoproteína E humana, cadena alfa de fibrinógeno humano, haptoglobina humana, antitrombina humana, alfa globina humana, beta globina humana, complemento C3 humano, factor de crecimiento humano, hepcidina humana, MALAT-1, beta globina de ratón, albúmina de ratón y beta globina de *Xenopus*, o fragmentos de cualquiera de los anteriores. Preferentemente, las UTR en 3' derivan de la beta globina de *Xenopus*. Las secuencias de UTR en 3' preferidas incluyen las SEQ ID NO: 16-22 como se muestra en la Tabla 4.

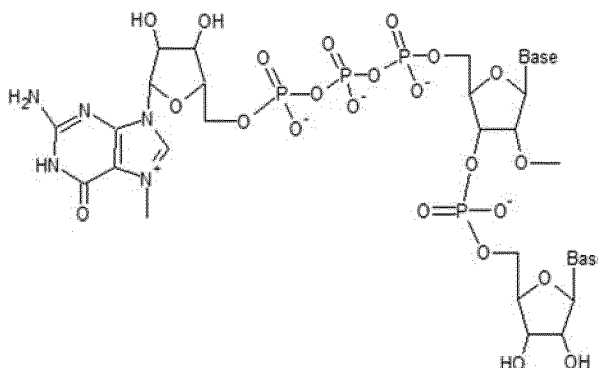
Tabla 4
Secuencias de UTR en 3'

Nombre	Secuencia	SEQ ID NO:
XBG	CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCU CAAGAACACCCGAAUGGAGUCUCUAAAGCUACAUAAUACCAA CUUACACUUAACAAAUGUUGUCCCCCAAAUGUAGCCAUUC GUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU	SEQ ID NO: 16
HAPTOGLOBINA HUMANA	UGCAAGGCUGGCCGGAAGCCCUUGCCUGAAAGCAAGAUUU CAGCCUGGAAGAGGGCAAAGUGGACGGGAGUGGACAGGAG UGGAUGCGAUAGAUGUGGUUUGAAGCUGAUGGGUGCCA GCCUGCAUUGCUGAGUCAUAUAAAGAGCUUUCUUUU GACCAU	SEQ ID NO: 17
APOLIPOPROTEÍNA E HUMANA	ACGCCGAAGCCUGCAGCCAUGCGACCCACGCCACCCCGUGC CUCCUGCCUCCGCGCAGCCUGCAGCGGGAGACCCUGUCCCC GCCCCAGCCGUCCUCCUGGGGUGGACCCUAGUUUAAUAAA GAUUCACCAAGUUUCACGCA	SEQ ID NO: 18
VHC	UAGAGCGGCAAACCCUAGCUACACUCCAUAGCUAGUUUCU UUUUUUUUUUGUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU UUUUUUUUUUUCCUUUCUUUUCCUUCUUUUUUUCCUCU UUUCUUGGUGGCUCCAUUCUAGCCCUAGUCACGGCUAGCU GUGAAAGGUCCGUGAGCCGCAUGACUGCAGAGAGUGCCGU AACUGGUCUCUCUGCAGAUCAUGU	SEQ ID NO: 19
ALBÚMINA DE RATÓN	ACACAUACAACCACAACCUUCUCAGGCUACCCUGAGAAAA AAAGACAUGAAGACUCAGGACUCAUCUUUUCUGUUGGUGU AAAUAACAACCCUAAGGAACACAAAUUCUUUAAACAUUU GACUUCUUGUCUCUGUGCUGCAAUUAUAAAAAUGGAAA GAAUCUAC	SEQ ID NO: 20
ALFA GLOBINA HUMANA	GCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCU CCCAACGGGCCCUCUCCCCUCCUUGCACCGGCCCUUCCUG GUCUUUGAAUAAAGUCUGAGUGGGCAGCA	SEQ ID NO: 21
EMCV	UAGUGCAGUCAC UGGCACAACG CGUUGCCCGG UAAGCCAAUC GGGUAUACAC GGUCGUCAUACUGCAGACAG GGUUCUUCUA CUUUGCAAGA UAGUCUAGAG UAGUAAAAUA AAUAGUAUAAG	SEQ ID NO: 22

Preferentemente, un ARNm descrito en el presente documento comprende una región de cola en 3', que puede servir para proteger el ARNm de la degradación por exonucleasa. La región de cola puede ser una región de poli(A) en 3'

y/o de poli(C) en 3'. Preferentemente, la región de cola es una cola de poli(A) en 3'. Como se utiliza en el presente documento, una "cola de poli(A) en 3'" es un polímero de nucleótidos de adenina secuenciales cuyo tamaño puede variar, por ejemplo: de 10 a 250 nucleótidos de adenina secuenciales; 60-125 nucleótidos de adenina secuenciales; 90-125 nucleótidos de adenina secuenciales, 95-125 nucleótidos de adenina secuenciales, 95-121 nucleótidos de adenina secuenciales, de 100 a 121 nucleótidos de adenina secuenciales, 110-121 nucleótidos de adenina secuenciales; 112-121 nucleótidos de adenina secuenciales; 114-121 nucleótidos de adenina secuenciales; y de 115 a 121 nucleótidos de adenina secuenciales. Preferentemente, una cola de poli A en 3' como se describe en el presente documento, comprende 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124 o 125 nucleótidos de adenina secuenciales. Se pueden añadir colas de poli(A) en 3' usando una variedad de métodos conocidos en la técnica, p. ej., usando poli(A) polimerasa para añadir colas al ARN sintético o transcrito *in vitro*. Otros métodos incluyen el uso de un vector de transcripción para codificar colas de poli A o el uso de una ligasa (p. ej., mediante ligamiento con férula usando una ARN ligasa T4 y/o ADN ligasa T4), en donde la poli(A) puede ligarse al extremo 3' de un ARN de sentido. Preferentemente, se utiliza una combinación de cualquiera de los métodos anteriores.

Preferentemente, un ARNm descrito en el presente documento comprende una caperuza en 5'. En la técnica se conocen extremos 5' con caperuza con varios grupos y análogos de los mismos. La caperuza en 5' puede seleccionarse de m7GpppA, m7GpppC; análogos de caperuza no metilada (p. ej., GpppG); análogo de caperuza dimetilada (p. ej., m2,7GpppG), un análogo de caperuza trimetilada (p. ej., m2,2,7GpppG), análogos de caperuza simétrica dimetilada (p. ej., m7Gpppm7G) o análogos de caperuza anti inversa (p. ej., ARCA; m7, 2'OmGpppG, m72'dGpppG, m7,3'OmGpppG, m7,3'dGpppG y sus derivados de tetrafosfato) (véase, p. ej., Jemielity, J. *et al.*, RNA 9: 1108-1122 (2003)). La caperuza en 5' puede ser una caperuza ARCA (3'-OMe-m7G(5')pppG). La caperuza en 5' puede ser un mCAP (m7G(5')ppp(5')G, N⁷-Metil-Guanosin-5'-Trifosfato-5'-Guanosine). La caperuza en 5' puede ser resistente a la hidrólisis. Una caperuza en 5' preferida se denomina en el presente documento "caperuza m7GpppGm", también denominada en el presente documento "Cap1" y tiene la siguiente estructura central:



Preferentemente, un ARNm descrito en el presente documento comprende uno o más nucleótidos modificados químicamente. Como ejemplos de monómeros de ácido nucleico se incluyen nucleótidos no naturales, modificados y modificados químicamente, incluyendo cualquiera de dichos nucleótidos conocidos en la técnica. Se ha demostrado que las secuencias de ARNm que comprenden nucleótidos modificados químicamente mejoran la expresión del ARNm, sus tasas de expresión, semivida y/o concentraciones de proteínas expresadas. Las secuencias de ARNm que comprenden nucleótidos modificados químicamente también han sido útiles para optimizar la localización de proteínas, impidiendo así que se produzcan respuestas biológicas nocivas tales como la respuesta inmunitaria y/o las rutas de degradación.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen 5-hidroxicitidinas, 5-alcilcitidinas, 5-hidroxicilcitidinas, 5-carboxicitidinas, 5-formilcitidinas, 5-alcoxicitidinas, 5-alcinilcitidinas, 5-halocitidinas, 2-tiocitidinas, N⁴-alcilcitidinas, N⁴-aminocitidinas, N⁴-acetilcitidinas y N⁴,N⁴-dialcicilcitidinas.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen 5-hidroxicitidina, 5-metilcitidina, 5-hidroximetilcitidina, 5-carboxicitidina, 5-formilcitidina, 5-metoxicitidina, 5-propinilcitidina, 5-bromocitidina, 5-yodocitidina, 2-tiocitidina; N⁴-metilcitidina, N⁴-aminocitidina, N⁴-acetilcitidina y N⁴,N⁴-dimetilcitidina.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen 5-hidroxiuridinas, 5-alquiluridinas, 5-hidroxiálquiluridinas, 5-carboxiuridinas, 5-carboxialquilesteruridinas, 5-formiluridinas, 5-alcoxiuridinas, 5-alquiluridinas, 5-halouridinas, 2-tiouridinas y 6-alquiluridinas.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen 5-hidroxiuridina, 5-metiluridina, 5-hidroximetiluridina, 5-carboxiuridina, 5-carboximetilesteruridina, 5-formiluridina, 5-metoxiuridina (también denominada en el presente documento "5MeOU"), 5-propiniluridina, 5-bromouridina, 5-fluorouridina, 5-yodouridina, 2-tiouridina y 6-metiluridina.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen 5-metoxycarbonilmetil-2-tiouridina, 5-metoxiaminometil-2-tiouridina, 5-carbamoilmetiluridina, 5-carbamoilmetil-2'-O-metiluridina, 1-metil-3-(3-amino-3-carboxipropil)pseudouridina, 5-metilaminometil-2-selenouridina, 5-carboximetiluridina, 5-metildihidrouridina, 5-taurinometiluridina, 5-taurinometil-2-tiouridina, 5-(isopentenilaminometil)uridina, 2'-O-metilpseudouridina, 2-tio-2'-O-metiluridina y 3,2'-O-dimetiluridina.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen N⁶-metiladenosina, 2-aminoadenosina, 3-metiladenosina, 8-azaadenosina, 7-deaza adenosina, 8-oxoadenosina, 8-bromoadenosina, 2-metiltio-N⁶-metiladenosina, N⁶-isopenteniladenosina, 2-metiltio-N⁶-isopenteniladenosina, N⁶-(cis-hidroxiisopentenil)adenosina, 2-metiltio-N⁶-(cis-hidroxiisopentenil)adenosina, N⁶-glicinilcarbamoiladenosina, N⁶-treonilcarbamoil-adenosina, N⁶-metil-N⁶-treonilcarbamoil-adenosina, 2-metiltio-N⁶-treonilcarbamoil-adenosina, N⁶,N⁶-dimetiladenosina, N⁶-hidroxinorvalilcarbamoiladenosina, 2-metiltio-N⁶-hidroxinorvalilcarbamoil-adenosina, N⁶-acetiladenosina, 7-metil-adenosina, 2-metiltioadenosina, 2-metoxi-adenosina, alfa-tio-adenosina, 2'-O-metil-adenosina, N⁶,2'-O-dimetil-adenosina, N⁶,N⁶,2'-O-trimetil-adenosina, 1,2'-O-dimetil-adenosina, 2'-O-ribosiladenosina, 2-amino-N⁶-metilpurina, 1-tio-adenosina, 2'-F-ara-adenosina, 2'-F-adenosina, 2'-OH-ara-adenosina y N⁶-(19-amino-pentaoxonadecil)-adenosina.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen N¹-alquilguanosinas, N²-alquilguanosinas, tienoguanosinas, 7-desazaguanosinas, 8-oxoguanosinas, 8-bromoguanosinas, O⁶-alquilguanosinas, xantosinas, inosinas y N¹-alquilinosinas.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen N¹-metilguanosina, N²-metilguanosina, tienoguanosina, 7-desazaguanosina, 8-oxoguanosina, 8-bromoguanosina, O⁶-metilguanosina, xantosa, inosina y N¹-metilinosina.

Como ejemplos de nucleótidos modificados o químicamente modificados se incluyen pseudouridinas. Como ejemplos de pseudouridinas se incluyen N¹-alquimpseudouridinas, N¹-cicloalquimpseudouridinas, N¹-hidroxipseudouridinas, N¹-hidroxialquimpseudouridinas, N¹-fenilpseudouridinas, N¹-fenilalquimpseudouridinas, N¹-aminoalquimpseudouridinas, N³-alquimpseudouridinas, N⁶-alquimpseudouridinas, N⁶-alcoxipseudouridinas, N⁶-hidroxipseudouridinas, N⁶-hidroxialquimpseudouridinas, N⁶-morfolinopseudouridinas, N⁶-fenilpseudouridinas y N⁶-halopseudouridinas. Como ejemplos de pseudouridinas se incluyen N¹-alquil-N⁶-alquimpseudouridinas, N¹-alquil-N⁶-alcoxipseudouridinas, N¹-alquil-N⁶-hidroxipseudouridinas, N¹-alquil-N⁶-hidroxialquimpseudouridinas, N¹-alquil-N⁶-morfolinopseudouridinas, N¹-alquil-N⁶-fenilpseudouridinas y N¹-alquil-N⁶-halopseudouridinas. En estos ejemplos, los sustituyentes alquilo, cicloalquilo y fenilo pueden estar sin sustituir o además sustituidos con sustituyentes alquilo, halo, haloalquilo, amino o nitró.

Como ejemplos de pseudouridinas se incluyen N¹-metilpseudouridina (también denominada en el presente documento "N1MPU"), N¹-etilpseudouridina, N¹-propilpseudouridina, N¹-ciclopropilpseudouridina, N¹-fenilpseudouridina, N¹-aminometilpseudouridina, N³-metilpseudouridina, N¹-hidroxipseudouridina y N¹-hidroximetilpseudouridina.

Como ejemplos de monómeros de ácido nucleico se incluyen nucleótidos modificados y químicamente modificados, incluyendo cualquiera de dichos nucleótidos conocidos en la técnica.

Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen cualquiera de dichos nucleótidos conocidos en la técnica, por ejemplo, ribonucleótidos de 2'-O-metilo, nucleótidos de 2'-O-metil purina, ribonucleótidos de 2'-desoxi-2'-fluoro, nucleótidos de 2'-desoxi-2'-fluoro pirimidina, 2'-desoxi ribonucleótidos, nucleótidos de 2'-desoxi purina, nucleótidos de base universal, 5-C-metil-nucleótidos y restos de monómeros desoxiabásicos invertidos.

Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen nucleótidos estabilizados en el extremo 3', nucleótidos de 3'-glicerilo, nucleótidos abásicos invertidos en 3' y timidina invertida en 3'.

Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen nucleótidos de ácido nucleico bloqueados (ANB), nucleótidos de 2'-O,4'-C-metilen-(D-ribofuranósilo), nucleótidos de 2'-metoxietoxi (MOE), nucleótidos de 2'-metil-tio-etilo, 2'-desoxi-2'-fluoro y 2'-O-metilo. En una realización ilustrativa, el monómero modificado es un nucleótido de ácido nucleico bloqueado (ANB).

Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen ADN modificados con 2'-O-metoxietilo y 2'-O-etilo restringidos en 2',4, (cMOE) y (cEt), respectivamente.

Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen nucleótidos de 2'-amino, nucleótidos de 2'-O-amino, nucleótidos de 2'-C-alilo y nucleótidos de 2'-O-alilo.

Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen nucleótidos de

N⁶-metiladenosina.

Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen monómeros de nucleótidos con bases modificadas 5-(3-amino)propiluridina, 5-(2-mercapto)etiluridina, 5-bromouridina; 8-bromoguanosina o 7-deaza adenosina.

Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen nucleótidos sustituidos con 2'-O-aminopropilo.

- 10 Como ejemplos de monómeros de nucleótidos modificados y químicamente modificados se incluyen el reemplazo del grupo 2'-OH de un nucleótido por un 2'-R, un 2'-OR, un 2'-halógeno, un 2'-SR o un 2'-amino, donde R puede ser H, alquilo, alqueno o alquino.

- 15 Los ejemplos de modificaciones de bases descritas anteriormente se pueden combinar con modificaciones adicionales de la estructura de nucleósidos o nucleótidos, incluyendo modificaciones de azúcar y modificaciones de enlaces. En la naturaleza pueden encontrarse algunos monómeros de nucleótidos modificados o químicamente modificados.

Las modificaciones de nucleótidos preferidas incluyen N¹-metilpseudouridina y 5-metoxiuridina.

- 20 En la Tabla 5 se proporcionan construcciones de secuencias de ARNm preferidas.

Tabla 5: Construcciones de ARNm ilustrativas

N.º de construcción de ARNm	Caperuza	UTR en 5'	Kozak*	Proteína codificada OTC	UTR en 3'	Cola de Poli A en 3'	Construcción de ARNm SEQ ID NO:
563	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	26
564	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	27
565	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	28
566	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	29
567	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	30
568	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	31
569	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	32
570	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	33
571	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	34
572	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	35
573	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	36
574	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	37
575	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	38
708	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	39
709	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	40
710	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	41
711	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	42
712	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	43
713	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	44
714	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	45
715	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	46
716	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	47
717	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	48
718	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	49
719	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	50
720	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	51
721	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	52
722	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	53
723	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	54
724	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	55
725	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	56
726	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	57
727	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	58
728	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	59
729	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	60
1787	Cap1	ARC5-2	N.º	SEQ ID NO: 3	haptoglobina humana	Sí	61
1788	Cap1	AT1G58420	P	SEQ ID NO: 3	ApoE humana	Sí	62

(continuación)

N.º de construcción de ARNm	Caperuza	UTR en 5'	Kozak*	Proteína codificada OTC	UTR en 3'	Cola de Poli A en 3'	Construcción de ARNm SEQ ID NO:
1789	Cap1	ARC5-2	N.º	SEQ ID NO: 3	haptoglobina humana	Sí	63
1790	Cap1	AT1G58420	P	SEQ ID NO: 3	ApoE humana	Sí	64
1791	Cap1	ARC5-2	N.º	SEQ ID NO: 3	haptoglobina humana	Sí	65
1792	Cap1	VHC5'	P	SEQ ID NO: 3	VHC3'	Sí	66
1793	Cap1	AT1G58420	P	SEQ ID NO: 3	a-glob humana	Sí	67
1794	Cap1	AT1G58420	P	SEQ ID NO: 3	a-glob humana	Sí	68
1795	Cap1	AT1G58420	P	SEQ ID NO: 3	a-glob humana	Sí	69
1796	Cap1	alb humana	N.º	SEQ ID NO: 3	alb murina	Sí	70
1797	Cap1	alb humana	N.º	SEQ ID NO: 3	alb murina	Sí	71
1798	Cap1	alb humana	N.º	SEQ ID NO: 3	alb murina	Sí	72
1799	Cap1	AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	73
1800	Cap1	alb humana	N.º	SEQ ID NO: 3	alb murina	Sí	74
1801	Cap1	AT1G58420	P	SEQ ID NO: 3	ApoE humana	Sí	75
1802	Cap1	ARC5-2	N.º	SEQ ID NO: 3	haptoglobina humana	Sí	76
1803	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	77
1804	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	78
1805	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	79
1806	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	80
1808	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	81
1809	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	82
1816	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	83
1822	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	84
1823	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	85
1840	Cap1	EMCV	N.º	SEQ ID NO: 3	EMCV	Sí	86
1841	Cap1	EMCV	N.º	SEQ ID NO: 3	EMCV	Sí	87
1842	Cap1	EMCV	N.º	SEQ ID NO: 3	EMCV	Sí	88
1843	Cap1	HSP70-P2-TEV	Sí	SEQ ID NO: 3	XBG	Sí	89
1844	Cap1	HSP70-M1-TEV	Sí	SEQ ID NO: 3	XBG	Sí	90
1845	Cap1	HSP70-M2-TEV	Sí	SEQ ID NO: 3	XBG	Sí	91
1846	Cap1	HSP17.9-TEV	Sí	SEQ ID NO: 3	XBG	Sí	92
1847	Cap1	HSP70-P1-TEV	Sí	SEQ ID NO: 3	XBG	Sí	93
1882	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	94
1883	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	95
1884	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	96
1885	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	97
1886	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	98
1887	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	99
1888	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	100
1889	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	101
1890	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	102
1891	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	103
1898	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	104
1899	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	105
1900	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	106

(continuación)

N.º de construcción de ARNm	Caperuza	UTR en 5'	Kozak*	Proteína codificada OTC	UTR en 3'	Cola de Poli A en 3'	Construcción de ARNm SEQ ID NO:
1903	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	107
1904	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	108
1905	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	109
1906	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	110
1907	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	111
1908	Cap1	TEV	P	SEQ ID NO: 3	XBG	Sí	112
1915	Cap1	HSP70-P1-AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	113
1916	Cap1	HSP70-P1-AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	114
1917	Cap1	HSP70-P1-AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	115
1918	Cap1	HSP70-P1-AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	116
1919	Cap1	AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	117
1920	Cap1	AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	118
1921	Cap1	AT1G58420	Sí	SEQ ID NO: 4**	a-glob humana	Sí	119
1925	Cap1	TEV	Sí	SEQ ID NO: 3	XBG	Sí	120
1926	Cap1	AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	121
1927	Cap1	AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	122
1928	Cap1	AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	123
1929	Cap1	HSP70-P1-AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	124
2016	Cap1	AT1G58420	Sí	SEQ ID NO: 3	a-glob humana	Sí	253
2260	Cap1	AT1G58420	Sí	SEQ ID NO: 4**	a-glob humana	Sí	251
2262	Cap1	AT1G58420	Sí	SEQ ID NO: 4**	a-glob humana	Sí	252

*Secuencia de Kozak definida como GCCACC (SEQ ID NO: 23). Kozak parcial (P) definida como GCCA (SEQ ID NO: 24).

** La construcción codifica la proteína OTC humana modificada de SEQ ID NO: 4.

5

Las secuencias de ARNm preferidas incluyen todas las secuencias de ARNm enumeradas en la Tabla 5. Las secuencias de ARNm preferidas incluyen todas las secuencias de ARNm enumeradas en donde, del 0 % al 100 %, preferentemente del 1 % al 100 %, preferentemente del 25 % al 100 %, preferentemente del 50 % al 100 % y preferentemente del 75 % al 100 % de los nucleótidos de uracilo de las secuencias de ARNm están modificados. Preferentemente, del 1 % al 100 % de los nucleótidos de uracilo son N¹-metilpseudouridina o 5-metoxiuridina. Preferentemente el 100 % de los nucleótidos de uracilo son N¹-metilpseudouridina. Preferentemente el 100 % de los nucleótidos de uracilo son 5-metoxiuridina.

10

15

Las secuencias de ARNm preferidas comprenden una caperuza en 5', una UTR en 5' que deriva de un gen expresado por *Arabidopsis thaliana*, una secuencia potenciadora de traducción opcional, una secuencia de Kozak opcional o una secuencia de Kozak parcial, una secuencia codificante optimizada con codones (CDS/ORF) que codifica una proteína OTC, una UTR en 3' y una cola de poli A. Preferentemente, la CDS optimizada con codones codifica una proteína de SEQ ID NO: 3 o SEQ ID NO: 4. Preferentemente, la UTR en 5' que deriva de un gen expresado por *Arabidopsis thaliana* se selecciona de lo que se observa en la Tabla 5. Preferentemente, la UTR en 5' que deriva de un gen expresado por *Arabidopsis thaliana* se selecciona del grupo que consiste en: SEQ ID NO: 6, SEQ ID NO: 125-127 y SEQ ID NO: 230-250. Preferentemente, la secuencia UTR en 5' es AT1G58420 que tiene la secuencia de SEQ ID NO: 6. Preferentemente, el contenido de uracilo de la secuencia optimizada con codones se ha reducido con respecto a los porcentajes de contenido de uracilo de SEQ ID NO: 1. Preferentemente, se modifica del 0 % al 100 % de los nucleótidos de uracilo de las secuencias de ARNm. Preferentemente, del 0 % al 100 % de los nucleótidos de uracilo son N¹-metil pseudouridina o 5-metoxiuridina. Preferentemente el 100 % de los nucleótidos de uracilo son N¹-metilpseudouridina. Preferentemente el 100 % de los nucleótidos de uracilo son 5-metoxiuridina.

25

Las construcciones de ARNm preferidas comprenden secuencias codificantes optimizadas con codones y una UTR en 5' de un gen expresado por *Arabidopsis thaliana* y se seleccionan de: SEQ ID NO: 62, 67, 68, 69, 73, 113-119, 121-127.

5 Una construcción de ARNm preferida de la divulgación comprende la construcción de ARNm 1921 (SEQ ID NO: 119) que tiene un ORF optimizado que codifica la proteína OTC humana modificada de SEQ ID NO: 4 y que comprende una cola de Poli A en 3' de 121 nucleótidos. Otra construcción de ARNm preferida comprende la construcción 2260 (SEQ ID NO: 251) que codifica la proteína OTC humana modificada de SEQ ID NO: 4 y que comprende una cola de Poli A en 3' de 100 nucleótidos. Otra construcción de ARNm preferida comprende la construcción 2262 (SEQ ID NO: 252) que codifica la proteína OTC humana modificada de SEQ ID NO: 4 y que comprende una cola de Poli A en 3' de 100 nucleótidos.

15 Una secuencia de ARNm preferida de la divulgación incluye la construcción de ARNm 1799 (SEQ ID NO: 73) que tiene un ORF optimizado con codones que codifica la OTC humana de tipo silvestre de SEQ ID NO: 3 y que tiene una cola de Poli A en 3' de 121 nucleótidos. Otra construcción de ARNm preferida de la divulgación incluye la construcción de ARNm 2016 (SEQ ID NO: 253) que tiene un ORF optimizado con codones que codifica la OTC humana de tipo silvestre de SEQ ID NO: 3 y que comprende una cola de Poli A en 3' de 100 nucleótidos.

20 Preferentemente el 100 % de los nucleótidos de uridina de las construcciones de ARNm 1799, 2016, 1921, 2260 y 2262, son N¹-metilpseudouridina. Preferentemente el 100 % de los nucleótidos de uracilo de las construcciones de ARNm 1799, 2016, 1921, 2260 y 2262, son 5-metoxiuridina.

25 El ARNm para su uso de acuerdo con esta divulgación puede mostrar una mayor eficacia de traducción. Como se utiliza en el presente documento, eficacia de traducción se refiere a una medida de la producción de una proteína o un polipéptido mediante la traducción de un ARNm de acuerdo con la divulgación. Preferentemente, un ARNm de la divulgación puede mostrar una eficacia de traducción de al menos 2 veces, 3 veces, 5 veces o 10 veces mayor *in vivo* en comparación con el ARNm que codifica la SEQ ID NO: 3 o SEQ ID NO: 4 que no está optimizado con codones de acuerdo con la divulgación y/o que no comprende las UTR preferidas de la divulgación. Preferentemente, un ARNm de la divulgación puede proporcionar al menos un nivel de proteína o de polipéptido 2 veces, 3 veces, 5 veces o 10 veces mayor *in vivo* en comparación con el ARNm que codifica la SEQ ID NO: 3 o SEQ ID NO: 4 que no está optimizado con codones y/o que no comprende las UTR preferidas de la divulgación. Preferentemente, un ARNm de la divulgación puede proporcionar mayores niveles de un polipéptido o una proteína *in vivo* en comparación con el ARNm que codifica la SEQ ID NO: 3 o SEQ ID NO: 4 que no está optimizado con codones de acuerdo con la divulgación y/o que no comprende las UTR preferidas de la misma. Por ejemplo, el nivel de un polipéptido o una proteína puede aumentar en un 10 %, 20 %, 30 %, 40 %, 50 % o más.

40 Preferentemente, el ARNm de la divulgación puede proporcionar mayor semivida funcional en el citoplasma de células de mamífero sobre el ARNm que codifica la SEQ ID NO: 3 o la SEQ ID NO: 4 que no está optimizado con codones de acuerdo con la divulgación y/o que no comprende las UTR preferidas de la misma. Las moléculas traducibles de la invención pueden tener mayor semivida de actividad en comparación con el ARNm que codifica la SEQ ID NO: 3 o la SEQ ID NO: 4 que no está optimizado con codones de acuerdo con la divulgación y/o que no comprende las UTR preferidas de la misma.

45 Preferentemente, el ARNm de la divulgación puede reducir la respuesta inmunitaria celular en comparación con el ARNm que codifica la SEQ ID NO: 3 o la SEQ ID NO: 4 que no está optimizado con codones de acuerdo con la divulgación y/o que no comprende las UTR preferidas de la misma.

50 Preferentemente, el ARNm de la divulgación puede reducir los niveles de dosis necesarios para una terapia eficaz en comparación con el ARNm que codifica la SEQ ID NO: 3 o la SEQ ID NO: 4 que no está optimizado con codones de acuerdo con la divulgación y/o que no comprende las UTR preferidas de la misma.

Diseño y Síntesis de las secuencias de ARNm

55 El ARNm para uso de acuerdo con la divulgación puede prepararse de acuerdo con cualquier técnica disponible que incluya, pero sin limitación, síntesis química, transcripción *in vitro* (TIV) o escisión enzimática o química de un precursor más largo, etc. En la técnica se conocen métodos para sintetizar ARN.

60 En algunas realizaciones, el ARNm se produce a partir de una construcción de ADN complementario (ADNc) primario. El proceso de diseño y síntesis de las construcciones de ADNc primario descritas en el presente documento generalmente incluyen las etapas de construcción génica, producción de ARNm (ya sea con o sin modificaciones) y purificación. En el método de TIV, primero se selecciona una secuencia de polinucleótidos diana que codifica una proteína OTC para su incorporación en un vector que se amplificará para producir un molde de ADNc. Opcionalmente, la secuencia de polinucleótidos diana y/o cualquier secuencia flanqueante pueden optimizarse con codones. El molde de ADNc se utiliza después para producir ARNm a través de transcripción *in vitro* (TIV). Después de la producción, el ARNm puede someterse a procesos de purificación y limpieza. Las etapas de esto se proporcionan con más detalle a

continuación.

La etapa de construcción génica puede incluir, pero sin limitación, síntesis génica, amplificación vectorial, purificación de plásmido, linealización y limpieza de plásmidos, y síntesis y limpieza de moldes de ADNc. Una vez que para la producción se selecciona una proteína OTC humana (p. ej., la SEQ ID NO: 3 o SEQ ID NO: 4), se diseña una construcción primaria. Dentro de la construcción primaria, se puede construir una primera región de nucleósidos unidos que codifica el polipéptido de interés usando un marco abierto de lectura (ORF) de un transcrito de ácido nucleico (ADN o ARN) seleccionado. El ORF puede comprender el ORF de tipo silvestre, una isoforma, variante o fragmento del mismo. Como se utiliza en el presente documento, un "marco de lectura abierto" u "ORF" pretende referirse a una secuencia de ácido nucleico (ADN o ARN) que es capaz de codificar un polipéptido de interés. Los ORF suelen comenzar con el codón de inicio, ATG y finalizar con un codón o señal sin sentido o de terminación.

Asimismo, la secuencia de nucleótidos de cualquier región del molde de ARNm o ADN puede optimizarse con codones. En la técnica se conocen métodos de optimización con codones y pueden ser útiles para alcanzar uno o más de diversos objetivos. Estos objetivos incluyen hacer coincidir las frecuencias de codones en los organismos diana y hospedadores para garantizar un plegamiento adecuado, sesgar el contenido de GC para aumentar la estabilidad del ARNm o reducir las estructuras secundarias, minimizar los codones repetidos en tándem o los desplazamientos de bases que pueden alterar la construcción o expresión de genes, personalizar las regiones de control transcripcional y traduccional, insertar o eliminar secuencias de trans migración de proteínas, eliminar/añadir sitios de modificación postraducciona l en proteínas codificadas (p. ej., sitios de glucosilación), añadir, eliminar o mezclar dominios de proteínas, insertar o deleccionar sitios de restricción, modificar sitios de unión a ribosomas y sitios de degradación de ARNm, ajustar las tasas de traducción para permitir que los diversos dominios de la proteína se plieguen adecuadamente, o reducir o eliminar estructuras secundarias problemáticas dentro del ARNm. En la técnica se conocen herramientas, algoritmos y servicios adecuados de optimización con codones.

Preferentemente, el molde de ADNc primario puede incluir reducir la aparición o frecuencia de aparición de ciertos nucleótidos en la cadena molde. Por ejemplo, la aparición de un nucleótido en un molde podrá reducirse a un nivel inferior al 25 % de los nucleótidos del molde. En ejemplos adicionales, la aparición de un nucleótido en un molde podrá reducirse a un nivel inferior al 20 % de los nucleótidos del molde. En algunos ejemplos, la aparición de un nucleótido en un molde podrá reducirse a un nivel inferior al 16 % de los nucleótidos del molde. Preferentemente, la aparición de un nucleótido en un molde podrá reducirse a un nivel inferior al 15 %, y preferentemente podrá reducirse a un nivel inferior al 12 % de los nucleótidos del molde.

Por ejemplo, la presente divulgación proporciona ácidos nucleicos en los que con un contenido de uracilo alterado al menos un codón en la secuencia de tipo silvestre se ha reemplazado por un codón alternativo para generar una secuencia alterada por uracilo. Las secuencias de uracilo alteradas pueden tener al menos una de las siguientes propiedades:

- (i) un aumento o una disminución en el contenido de uracilo global (es decir, el porcentaje de uracilo del contenido total de nucleótidos en el ácido nucleico de una sección del ácido nucleico, p. ej., el marco de lectura abierto); o,
- (ii) un aumento o una disminución en el contenido de uracilo local (es decir, los cambios en el contenido de uracilo se limitan a subsecuencias específicas); o,
- (iii) un cambio en la distribución de uracilo sin un cambio en el contenido de uracilo global; o,
- (iv) un cambio en la agrupación de uracilo (p. ej., número de grupos, ubicación de grupos o distancia entre grupos); o,
- (v) combinaciones de los mismos.

Preferentemente, el porcentaje de nucleobases de uracilo en la secuencia de ácido nucleico se reduce con respecto al porcentaje de nucleobases de uracilo en la secuencia de ácido nucleico de tipo silvestre. Por ejemplo, el 30 % de las nucleobases pueden ser uracilo en la secuencia de tipo silvestre, pero las nucleobases que son uracilo son preferentemente inferiores al 15 %, preferentemente inferiores al 12 % y preferentemente inferiores al 10 % de las nucleobases en las secuencias de ácidos nucleicos de la divulgación. El porcentaje de contenido de uracilo puede determinarse dividiendo la cantidad de uracilo en una secuencia entre la cantidad total de nucleótidos y multiplicando por 100.

Preferentemente, el porcentaje de nucleobases de uracilo en una subsecuencia de la secuencia de ácido nucleico se reduce con respecto al porcentaje de nucleobases de uracilo en la subsecuencia correspondiente de la secuencia de tipo silvestre. Por ejemplo, la secuencia de tipo silvestre puede tener una región de extremo 5' (p. ej., 30 codones) con un contenido de uracilo local del 30 %, y el contenido de uracilo en esa misma región podría reducirse preferentemente a un 15 % o menor, preferentemente 12 % o menor y preferentemente 10 % o menor en las secuencias de ácidos nucleicos de la divulgación.

Preferentemente, los codones en la secuencia de ácido nucleico de la invención reducen o modifican, por ejemplo, el número, tamaño, ubicación o distribución de grupos de uracilo que podrían tener efectos nocivos sobre la traducción de proteínas. Aunque es deseable un menor contenido de uracilo, en determinados aspectos, el contenido de uracilo, y en particular el contenido de uracilo local, de algunas subsecuencias de la secuencia de tipo silvestre, puede ser

mayor que el de la secuencia de tipo silvestre y seguir conservando características beneficiosas (p. ej., mayor expresión).

Preferentemente, la secuencia modificada con uracilo induce una respuesta más baja del receptor tipo Toll (TLR, *Toll-Like Receptor*) en comparación con la de la secuencia de tipo silvestre. Varios TLR reconocen ácidos nucleicos y responden a ellos. Se ha demostrado que el ARN bicatenario (bc), un componente frecuente de los virus, activa el TLR3. El ARN monocatenario (mc) activa el TLR7. Los oligonucleótidos de ARN, por ejemplo, ARN con enlaces internucleotídicos de fosforotioato, son ligandos del TLR8 humano. El ADN que contiene motivos CpG no metilados, característicos del ADN bacteriano y vírico, activa el TLR9.

Como se utiliza en el presente documento, la expresión "respuesta del TLR" se define como el reconocimiento de ARN monocatenario por un receptor TLR7, y preferentemente abarca la degradación del ARN y/o las respuestas fisiológicas causadas por el reconocimiento del ARN monocatenario por el receptor. En la técnica se conocen métodos para determinar y cuantificar la unión de un ARN a un TLR7. De manera similar, en la técnica se conocen bien métodos para determinar si un ARN ha desencadenado una respuesta fisiológica mediada por TLR7 (p. ej., secreción de citocinas). Preferentemente, una respuesta del TLR puede estar mediada por TLR3, TLR8 o TLR9 en lugar de por TLR7. La supresión de la respuesta mediada por TLR7 se puede lograr mediante la modificación de nucleósidos. El ARN sufre más de cien modificaciones diferentes de nucleósidos en la naturaleza. El ARNr humano, por ejemplo, tiene diez veces más pseudouracilo ('P') y 25 veces más nucleósidos 2'-O-metilados que el ARNr bacteriano. El ARNm bacteriano no contiene modificaciones de nucleósidos, mientras que los ARNm de mamíferos tienen nucleósidos modificados, tales como 5-metilcitosina (m5C), N6-metiladenosina (m6A), inosina y muchos nucleósidos 2'-O-metilados además de N7-metilguanosina (m7G).

Preferentemente, el contenido de uracilo de los polinucleótidos desvelados en el presente documento y preferentemente el de los polinucleótidos que codifican la proteína OTC modificada de SEQ ID NO: 4 es inferior al 50 %, 49 %, 48 %, 47 %, 46 %, 45 %, 44 %, 43 %, 42 %, 41 %, 40 %, 39 %, 38 %, 37 %, 36 %, 35 %, 34 %, 33 %, 32 %, 31 %, 30 %, 29 %, 28 %, 27 %, 26 %, 25 %, 24 %, 23 %, 22 %, 21 %, 20 %, 19 %, 18 %, 17 %, 16 %, 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 90 %, 80 %, 70 %, 60 %, 5 %, 4 %, 3 %, 2 % o 1 % del total de nucleobases en la secuencia de referencia. Preferentemente, el contenido de uracilo de los polinucleótidos desvelados en el presente documento y preferentemente de los polinucleótidos que codifican la proteína OTC modificada de SEQ ID NO: 4, está entre aproximadamente el 5 % y aproximadamente el 25 %. Preferentemente, el contenido de uracilo de los polinucleótidos desvelados en el presente documento y preferentemente de los polinucleótidos que codifican la proteína OTC modificada de SEQ ID NO: 4 es de aproximadamente el 15 % y aproximadamente el 25 %.

Los moldes de ADNc pueden transcribirse para producir una secuencia de ARNm descrita en el presente documento usando un sistema de transcripción *in vitro* (TIV). El sistema normalmente comprende un tampón de transcripción, trifosfatos de nucleótidos (NTP), un inhibidor de la RNasa y una polimerasa. Los NTP pueden seleccionarse de, pero sin limitación, los descritos en el presente documento, incluidos los NTP naturales y no naturales (modificados). La polimerasa puede seleccionarse de, pero sin limitación, ARN polimerasa T7, ARN polimerasa T3 y polimerasas mutantes tales como, pero sin limitación, polimerasas capaces de incorporar ácidos nucleicos modificados.

El molde de ADNc primario o la secuencia de ARNm transcrita también pueden someterse a reacciones de adición de caperuza y/o cola. Una reacción de adición de caperuza puede realizarse mediante métodos conocidos en la técnica para añadir una caperuza en 5' al extremo 5' de la construcción primaria. Los métodos de adición de caperuza incluyen, pero sin limitación, el uso de una enzima Vaccinia Capping (New England Biolabs, Ipswich, Mass.) o tecnología CLEANCAP® (TriLink Biotechnologies). A través de métodos conocidos en la técnica puede realizarse una reacción de adición de cola de poli-A, tales como, pero sin limitación, 2' O-metiltransferasa y a través de métodos como los descritos en el presente documento. Si la construcción primaria generada a partir de ADNc no incluye una poli-T, puede resultar beneficioso realizar la reacción de adición de cola poli-A antes de retirar la construcción primaria.

Las construcciones de ADNc optimizadas con codones que codifican una proteína ornitina transcarbamilasa (OTC) son particularmente adecuadas para generar secuencias de ARNm descritas en el presente documento. Por ejemplo, dichas construcciones de ADNc pueden usarse como base para transcribir, *in vitro*, un polirribonucleótido que codifique una proteína ornitina transcarbamilasa (OTC). La Tabla 6 proporciona un listado de moldes de ORF de ADNc ilustrativos utilizados para la transcripción *in vitro* de las secuencias de ARNm enumeradas en la Tabla 5.

Tabla 6: Moldes de ADNc ilustrativos

N.º*** de construcción de ADN:	SEQ ID NO:	Proteína codificada por el molde de ADNc SEQ ID NO:
p563	128	SEQ ID NO: 3*
p564	129	SEQ ID NO: 3*
p565	130	SEQ ID NO: 3*
p566	131	SEQ ID NO: 3*
p567	132	SEQ ID NO: 3*
p568	133	SEQ ID NO: 3*

(continuación)

N.º*** de construcción de ADN:	SEQ ID NO:	Proteína codificada por el molde de ADNc SEQ ID NO:
p569	134	SEQ ID NO: 3*
p570	135	SEQ ID NO: 3*
p571	136	SEQ ID NO: 3*
p572	137	SEQ ID NO: 3*
p573	138	SEQ ID NO: 3*
p574	139	SEQ ID NO: 3*
p575	140	SEQ ID NO: 3*
p708	141	SEQ ID NO: 3*
p709	142	SEQ ID NO: 3*
p710	143	SEQ ID NO: 3*
p711	144	SEQ ID NO: 3*
p712	145	SEQ ID NO: 3*
p713	146	SEQ ID NO: 3*
p714	147	SEQ ID NO: 3*
p715	148	SEQ ID NO: 3*
p716	149	SEQ ID NO: 3*
p717	150	SEQ ID NO: 3*
p718	151	SEQ ID NO: 3*
p719	152	SEQ ID NO: 3*
p720	153	SEQ ID NO: 3*
p721	154	SEQ ID NO: 3*
p722	155	SEQ ID NO: 3*
p723	156	SEQ ID NO: 3*
p724	157	SEQ ID NO: 3*
p725	158	SEQ ID NO: 3*
p726	159	SEQ ID NO: 3*
p727	160	SEQ ID NO: 3*
p728	161	SEQ ID NO: 3*
p729	162	SEQ ID NO: 3*
p1787	163	SEQ ID NO: 3*
p1788	164	SEQ ID NO: 3*
p1789	165	SEQ ID NO: 3*
p1790	166	SEQ ID NO: 3*
p1791	167	SEQ ID NO: 3*
p1792	168	SEQ ID NO: 3*
p1793	169	SEQ ID NO: 3*
p1794	170	SEQ ID NO: 3*
p1795	171	SEQ ID NO: 3*
p1796	172	SEQ ID NO: 3*
p1797	173	SEQ ID NO: 3*
p1798	174	SEQ ID NO: 3*
p1799	175	SEQ ID NO: 3*
p1800	176	SEQ ID NO: 3*
p1801	177	SEQ ID NO: 3*
p1802	178	SEQ ID NO: 3*
p1803	179	SEQ ID NO: 3*
p1804	180	SEQ ID NO: 3*
p1805	181	SEQ ID NO: 3*
p1806	182	SEQ ID NO: 3*
p1808	183	SEQ ID NO: 3*
p1809	184	SEQ ID NO: 3*
p1816	185	SEQ ID NO: 3*
p1822	186	SEQ ID NO: 3*
p1823	187	SEQ ID NO: 3*
p1840	188	SEQ ID NO: 3*
p1841	189	SEQ ID NO: 3*
p1842	190	SEQ ID NO: 3*
p1843	191	SEQ ID NO: 3*
p1844	192	SEQ ID NO: 3*

(continuación)

N. ^o *** de construcción de ADN:	SEQ ID NO:	Proteína codificada por el molde de ADNc SEQ ID NO:
p1845	193	SEQ ID NO: 3*
p1846	194	SEQ ID NO: 3*
p1847	195	SEQ ID NO: 3*
p1882	196	SEQ ID NO: 3*
p1883	197	SEQ ID NO: 3*
p1884	198	SEQ ID NO: 3*
p1885	199	SEQ ID NO: 3*
p1886	200	SEQ ID NO: 3*
p1887	201	SEQ ID NO: 3*
p1888	202	SEQ ID NO: 3*
p1889	203	SEQ ID NO: 3*
p1890	204	SEQ ID NO: 3*
p1891	205	SEQ ID NO: 3*
p1898	206	SEQ ID NO: 3*
p1899	207	SEQ ID NO: 3*
p1900	208	SEQ ID NO: 3*
p1903	209	SEQ ID NO: 3*
p1904	210	SEQ ID NO: 3*
p1905	211	SEQ ID NO: 3*
p1906	212	SEQ ID NO: 3*
p1907	213	SEQ ID NO: 3*
p1908	214	SEQ ID NO: 3*
p1915	215	SEQ ID NO: 3*
p1916	216	SEQ ID NO: 3*
p1917	217	SEQ ID NO: 3*
p1918	218	SEQ ID NO: 3*
p1919	219	SEQ ID NO: 3*
p1920	220	SEQ ID NO: 3*
p1921	221	SEQ ID NO: 4**
p1925	222	SEQ ID NO: 3*
p1926	223	SEQ ID NO: 3*
p1927	224	SEQ ID NO: 3*
p1928	225	SEQ ID NO: 3*
p1929	226	SEQ ID NO: 3*
p2016	175	SEQ ID NO: 3*
p2260	221	SEQ ID NO: 4**
p2262	221	SEQ ID NO: 4**

*SEQ ID NO: 3 es la secuencia de aminoácidos de la OTC humana de tipo silvestre.

**SEQ ID NO: 4 es la secuencia de aminoácidos de la OTC humana modificada.

*** No se incluye la secuencia completa del plásmido.

5

Las secuencias molde de ADNc preferidas incluyen la secuencia de ADN de SEQ ID NO: 175 (p1779) que tiene una secuencia codificante optimizada que codifica la OTC humana de tipo silvestre de SEQ ID NO: 3. Las secuencias molde de ADNc preferidas también incluyen la secuencia de ADNc de SEQ ID NO: 221 (p1921), que tiene una secuencia codificante optimizada que codifica una proteína OTC modificada de SEQ ID NO: 4.

10

La presente divulgación también proporciona polinucleótidos (p. ej., ADN, ARN, ADNc, ARNm) que codifican una proteína OTC humana que pueden estar unidos operativamente a una o más secuencias de nucleótidos reguladoras en una construcción de expresión, tal como un vector o plásmido. En determinadas realizaciones, dichas construcciones son construcciones de ADN. Las secuencias de nucleótidos reguladoras generalmente serán apropiadas para una célula hospedadora usada para la expresión. Se conocen numerosos tipos de vectores de expresión apropiados y secuencias reguladoras adecuadas en la técnica para una diversidad de células hospedadoras.

15

Normalmente, dicha una o más secuencias de nucleótidos reguladoras pueden incluir, pero sin limitación, secuencias promotoras, secuencias líder o de señal, sitios de unión a ribosomas, secuencias de inicio y terminación de la transcripción, secuencias de inicio y terminación de la traducción y secuencias potenciadoras o activadoras. Las realizaciones de la presente divulgación contemplan promotores constitutivos o inducibles como se conocen en la técnica. Los promotores pueden ser promotores de origen natural o promotores híbridos que combinan elementos de más de un promotor.

20

Una construcción de expresión puede estar presente en una célula en un episoma, tal como un plásmido, o la construcción de expresión puede insertarse en un cromosoma. Preferentemente, el vector de expresión contiene un gen marcador de selección para permitir la selección de células hospedadoras transformadas. Los genes marcadores de selección son bien conocidos en la técnica y variarán con la célula hospedadora usada.

La presente divulgación también proporciona vectores de expresión que comprenden una secuencia de nucleótidos que codifica una proteína ornitina transcarbamilasa (OTC) que está preferentemente unida operativamente a al menos una secuencia reguladora. Las secuencias reguladoras están reconocidas en la técnica y se seleccionan para dirigir la expresión del polipéptido codificado.

En consecuencia, la expresión secuencia reguladora incluye promotores, potenciadores y otros elementos de control de la expresión. El diseño del vector de expresión puede depender de factores tales como la elección de la célula hospedadora a transformar y/o el tipo de proteína deseada a expresar.

La presente divulgación también proporciona una célula hospedadora transfectada con un ARNm o ADN descrito en el presente documento que codifica un polipéptido de ornitina transcarbamilasa (OTC) descrito en el presente documento. Preferentemente, el polipéptido de OTC humana tiene la secuencia de SEQ ID NO: 4. La célula hospedadora puede ser cualquier célula procariota o eucariota. Por ejemplo, un polipéptido de ornitina transcarbamilasa (OTC) puede expresarse en células bacterianas tales como *E. coli*, células de insecto (p. ej., usando un sistema de expresión de baculovirus), levadura o células de mamífero. Otras células hospedadoras adecuadas son conocidas por los expertos en la materia.

La presente divulgación también proporciona una célula hospedadora que comprende un vector que comprende un polinucleótido que codifica una secuencia de ARNm de una cualquiera de las SEQ ID NO: 26-229.

La presente divulgación también proporciona métodos para producir una proteína OTC humana de tipo silvestre de SEQ ID NO: 3 o una proteína OTC humana modificada SEQ ID NO: 4. Preferentemente, la proteína OTC es SEQ ID NO: 4 y está codificada por el ARNm de SEQ ID NO 119. Por ejemplo, una célula hospedadora transfectada con un vector de expresión que codifica una proteína OTC puede cultivarse en condiciones apropiadas para permitir que se produzca la expresión del polipéptido. El polipéptido puede secretarse y aislarse de una mezcla de células y medio que contiene los polipéptidos. Como alternativa, los polipéptidos pueden retenerse en el citoplasma o en una fracción de membrana y las células se recogen, se someten a lisis y se aísla la proteína. Un cultivo celular incluye células hospedadoras, medios y otros subproductos. Los medios adecuados para el cultivo celular son bien conocidos en la técnica.

Las proteínas OTC expresadas descritas en el presente documento pueden aislarse del medio de cultivo celular, de células hospedadoras o de ambos usando técnicas conocidas en la materia para purificar proteínas, incluyendo cromatografía de intercambio iónico, cromatografía de filtración en gel, ultrafiltración, electroforesis y purificación por inmunoafinidad con anticuerpos específicos para epítomos particulares del polipéptido OTC.

Formulaciones basadas en lípidos

Las formulaciones basadas en lípidos han sido cada vez más reconocidas como uno de los sistemas de suministro (también denominados en el presente documento vehículo o transportador) más prometedores para el ARN debido a su biocompatibilidad y a su facilidad de producción a gran escala. Los lípidos catiónicos se han estudiado extensamente como materiales sintéticos para el suministro de ARN. Después de mezclarlos entre sí, los ácidos nucleicos son condensados por lípidos catiónicos para formar complejos de lípidos y ácidos nucleicos conocidos como lipoplexos. Estos complejos lipídicos pueden proteger el material genético de la acción de las nucleasas y suministrarlo a las células interaccionando con la membrana celular cargada negativamente. Los lipoplexos pueden prepararse mezclando directamente lípidos cargados positivamente a pH fisiológico con ácidos nucleicos cargados negativamente.

Los liposomas convencionales consisten en una bicapa lipídica que puede estar compuesta por (fosfo)lípidos catiónicos, aniónicos o neutros y colesterol, que encierra un núcleo acuoso. Tanto la bicapa lipídica como el espacio acuoso pueden incorporar compuestos hidrófobos o hidrófilos, respectivamente. Las características y el comportamiento de los liposomas *in vivo* puede modificarse mediante la adición de un recubrimiento de polímero hidrófilo, p.ej., polietilenglicol (PEG), a la superficie del liposoma para conferir estabilización estérica. Por otra parte, los liposomas pueden utilizarse para el direccionamiento específico uniendo ligandos (p. ej., anticuerpos, péptidos e hidratos de carbono) a su superficie o al extremo terminal de las cadenas de PEG unidas (Front Pharmacol. 1 de diciembre de 2015; 6:286).

Los liposomas son sistemas de suministro basados en lípidos coloidales y tensioactivos compuestos por una bicapa de fosfolípidos que rodea un compartimento acuoso. Pueden presentarse como vesículas esféricas y su tamaño puede variar desde 20 nm hasta algunos micrómetros. Los liposomas catiónicos basados en lípidos son capaces de formar complejos con ácidos nucleicos cargados negativamente mediante interacciones electrostáticas, dando como

resultado complejos que ofrecen biocompatibilidad, baja toxicidad y la posibilidad de la producción a gran escala necesaria en aplicaciones clínicas *in vivo*. Los liposomas pueden fusionarse con la membrana plasmática para su absorción; una vez dentro de la célula, los liposomas se procesan a través de la vía endocítica y después el material genético se libera desde el endosoma/transportador al citoplasma. Los liposomas se han percibido durante mucho tiempo como vehículos de suministro de fármacos debido a su biocompatibilidad superior, dado que los liposomas son básicamente análogos de las membranas biológicas, y pueden prepararse a partir de fosfolípidos tanto naturales como sintéticos (Int J Nanomedicine. 2014; 9: 1833-1843).

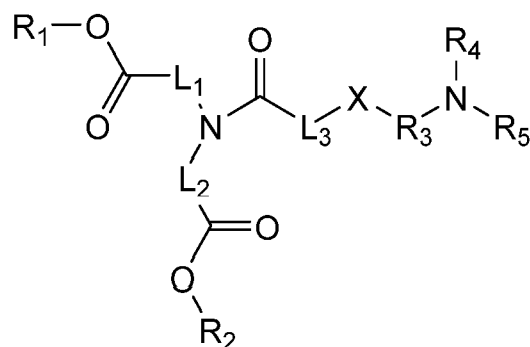
Los liposomas catiónicos han sido tradicionalmente los sistemas de suministro no vírico más utilizados de los oligonucleótidos, incluido el ADN plasmídico, los oligos antisentido y el ARNip/ARNhc (ARN de horquilla corta). Los lípidos catiónicos, tal como DOTAP, (1,2-dioleil-3-trimetilamonio-propano) y DOTMA (metilsulfato de N-[1-(2,3-dioleiloxi)propil]-N,N,N-trimetilamonio) pueden formar complejos o lipoplexos con ácidos nucleicos cargados negativamente para formar nanopartículas por interacción electrostática, proporcionando alta eficacia de transfección *in vitro*. Por otra parte, se desarrollaron nanoliposomas neutros basados en lípidos para el suministro de ARN como, por ejemplo, nanoliposomas neutros basados en 1,2-dioleil-sn-glicero-3-fosfatidilcolina (DOPC). (Adv Drug Deliv. Rev. Feb de 2014; 66: 110-116.)

Preferentemente, las construcciones de ARNm descritas en el presente documento están formuladas con lípidos. La formulación lipídica se selecciona preferentemente, pero sin limitación, de liposomas, lipoplexos, copolímeros, tales como PLGA y nanopartículas lipídicas. Preferentemente una nanopartícula lipídica (NPL) comprende:

- (a) un ácido nucleico,
(b) un lípido catiónico,
(c) un agente reductor de la agregación (tal como lípido de polietilenglicol (PEG) o lípido modificado con PEG),
(d) opcionalmente un lípido no catiónico (tal como un lípido neutro), y
(e) opcionalmente, un esteroil.

Preferentemente, la formulación de nanopartículas lipídicas consiste en (i) al menos un lípido catiónico; (ii) un lípido neutro; (iii) un estero, p. ej., colesterol; y (iv) un lípido-PEG, en una proporción molar de aproximadamente 20-60 % de lípido catiónico: 5-25 % de lípido neutro: 25-55 % de estero; 0.5-15 % de lípido-PEG.

Algunos ejemplos de lípidos y composiciones lipídicas para el suministro de una molécula activa de esta divulgación se dan en los documentos WO 2015/074085, U.S. 2018/0169268, WO 2018/119163, WO 20185/118102, U.S. 2018/0222863, WO 2016/081029, WO 2017/023817, WO 2017/117530. En determinadas realizaciones, el lípido es un compuesto de la siguiente Fórmula I:



Fórmula I

en donde

tanto R₁ como R₂ consisten en un alquilo lineal que consta de 1 a 14 carbonos, o un alqueno o alquino que consta de 2 a 14 carbonos;

tanto L₁ como L₂ consisten en un alquileno o alquenileno lineal que consta de 5 a 18 carbonos, o que forman un heterociclo con N;

X es S;

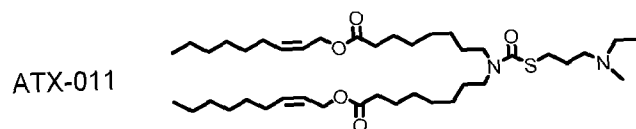
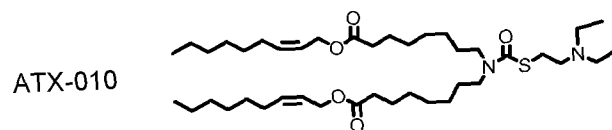
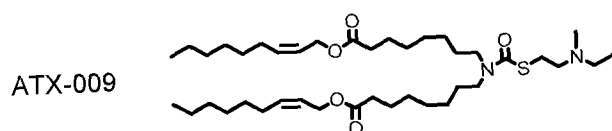
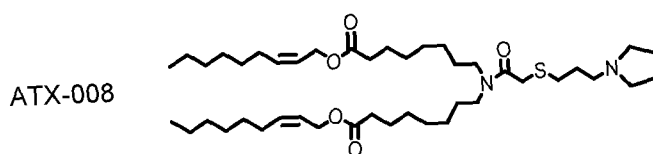
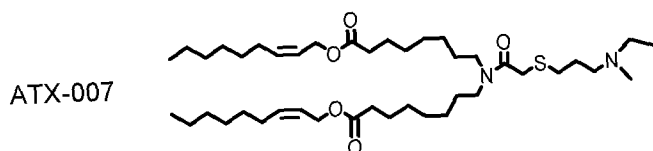
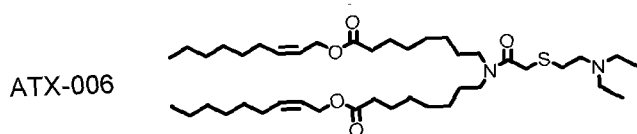
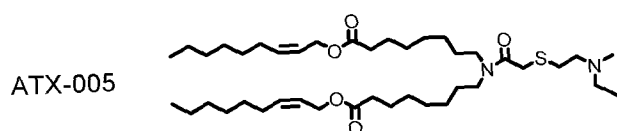
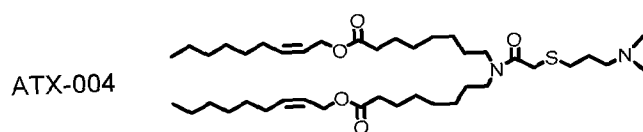
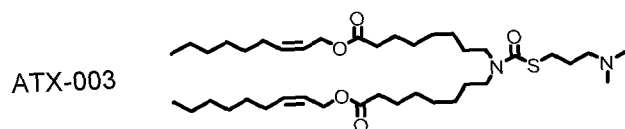
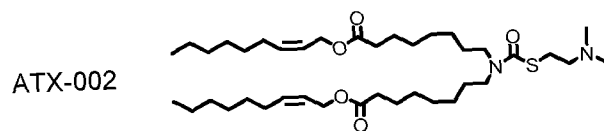
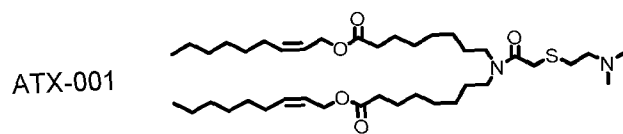
L₃ consiste en un enlace o un alquileo lineal que consta de 1 a 6 carbonos, o que forma un heterociclo con N;

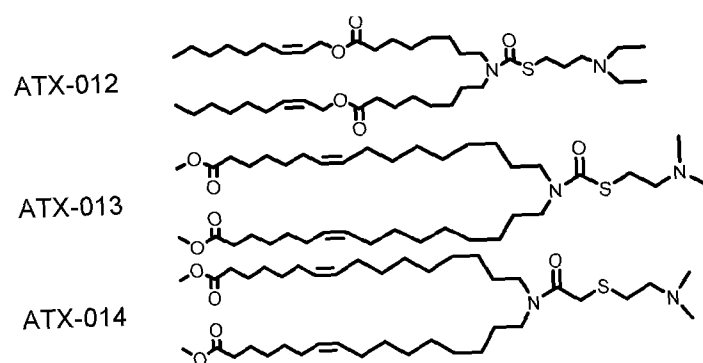
R₃ consiste en un alquileo lineal o ramificado que consta de 1 a 6 carbonos; y

R₄ y R₅ son iguales o diferentes, cada uno de los cuales consiste en un hidrógeno o un alquilo lineal o ramificado que consta de 1 a 6 carbonos:

o una de sus sales farmacéuticamente aceptables.

La formulación lipídica puede contener uno o más lípidos catiónicos ionizables seleccionados entre los siguientes (también denominados en el presente documento "lípidos ATX"):





ATX-015

ATX-016

ATX-017

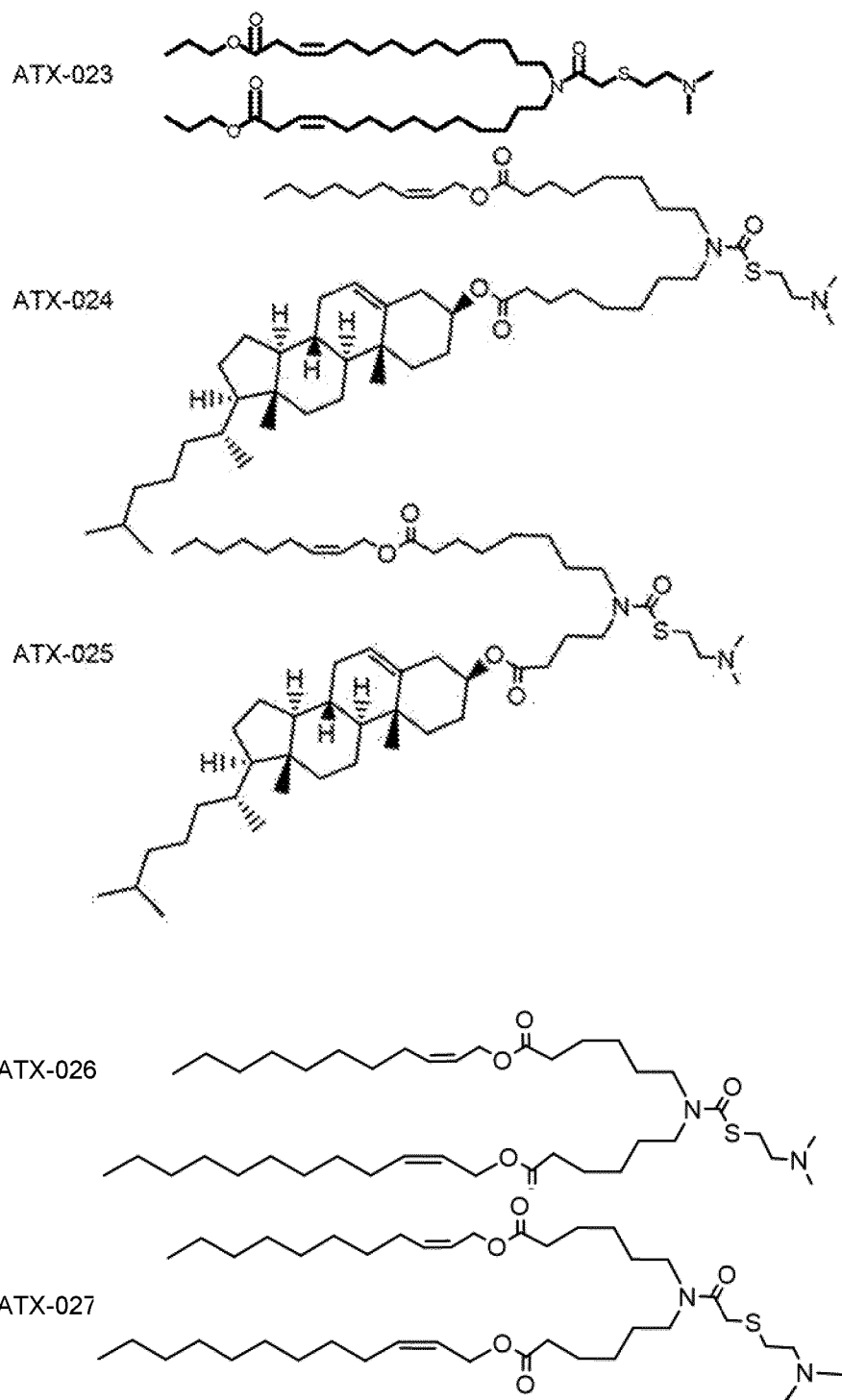
ATX-018

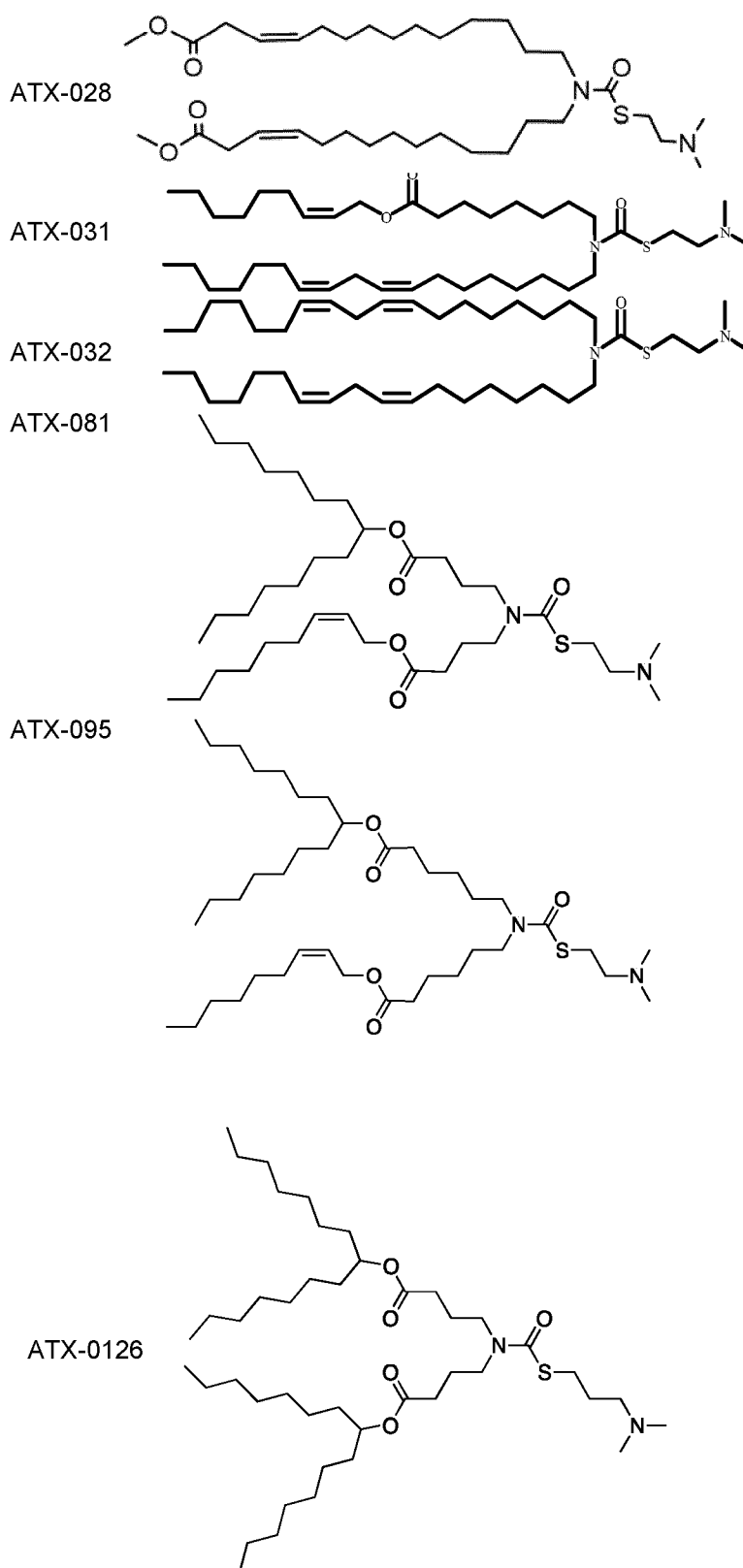
ATX-019

ATX-020

ATX-021

ATX-022





- 5 La nanopartícula lipídica incluye preferentemente un lípido catiónico adecuado para formar una nanopartícula lipídica. Preferentemente, el lípido catiónico lleva una carga neta positiva a un pH aproximadamente fisiológico. El lípido catiónico puede ser, por ejemplo, cloruro de N,N-dioleoil-N,N-dimetilamonio (DODAC), bromuro de N,N-diestearil-N,N-dimetilamonio (DDAB), cloruro de 1,2-dioleoiltrimetilamonio (DOTAP) (también conocido como cloruro de N-(2,3-dioleiloxi)propil)-N,N,N-trimetilamonio y sal de cloruro de 1,2-Dioleiloxi-3-trimetilaminopropano), cloruro de N-(1(2,3-dioleiloxi)propil)-N,N,N-trimetilamonio (DOTMA), N,N-dimetil-2,3-dioleiloxi)propilamina (DODMA), 1,2-
- 10

Dilinoleiloxi-N,N-dimetilaminopropano (DLinDMA), 1,2-Dilineniloxi-N,N-dimetilaminopropano (DLenDMA), 1,2-di-y-linoleniloxi-N,N-dimetilaminopropano (y-DLenDMA), 1,2-Dilinoileilcarbamoiloxi-3-dimetilaminopropano (DLin-C-DAP), 1,2-Dilinoileiloxi-3-(dimetilamino)acetoxipropano (DLin-DAC), 1,2-Dilinoileiloxi-3-morfolinopropano (DLin-MA), 1,2-Dilinoileil-3-dimetilaminopropano (DLinDAP), 1,2-Dilinoileil-3-dimetilaminopropano (DLin-S-DMA), 1-Linoileil-2-linoileiloxi-3-dimetilaminopropano (DLin-2-DMA), sal de cloruro de 1,2-Dilinoileiloxi-3-trimetilaminopropano (DLin-TMA.Cl), sal de cloruro de 1,2-Dilinoileil-3-trimetilaminopropano (DLin-TAP.Cl), 1,2-Dilinoileiloxi-3-(N-metilpiperazin)propano (DLin-MPZ) o 3-(N,N-Dilinoileilamino)-1,2-propanodiol (DLinAP), 3-(N,N-Dioleilamino)-1,2-propanodiol (DOAP), 1,2-Dilinoileiloxo-3-(2-N,N-dimetilamino)etoxipropano (DLin-EG-DM A), 2,2-Dilinoileil-4-dimetilaminometil-[1,3]-dioxolano (DLin-K-DMA) o análogos del mismo, (3aR,5s,6aS)-N,N-dimetil-2,2-di((9Z,12Z)-octadeca-9,12-dienil)tetrahydro-3aH-ciclopenta[d][1,3]dioxol-5-amina, (6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-il-4-(dimetilamino)butanoato (MC3), 1,1'-(2-(4-(2-(bis(2-hidroxidodecil)amino)etil)(2-hidroxidodecil)amino)etil)piperazin-1-il)etilazanedil)didodecan-2-ol (C12-200), 2,2-dilinoileil-4-(2-dimetilaminoetil)-[1,3]-dioxolano (DLin-K-C2-DMA), 2,2-dilinoileil-4-dimetilaminometil-[1,3]-dioxolano (DLin-K-DMA), (6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-il 4-(dimetilamino)butanoato (DLin-M-C3-DMA), 3-((6Z,9Z,28Z,31Z)-heptatriaconta-6,9,28,31-tetraen-19-iloxi)-N,N-dimetilpropan-1-amina (Éter MC3), 4-((6Z,9Z,28Z,31 Z)-heptatriaconta-6,9,28,31-tetraen-19-iloxi)-N,N-dimetilbutan-1-amina (Éter MC4), o cualquier combinación de cualquiera de lo anterior. Otros lípidos catiónicos incluyen, pero sin limitación, bromuro de N,N-diestearil-N,N-dimetilamonio (DDAB), 3P-(N-(N',N'-dimetilaminoetano)-carbamoil)colesterol (DC-Choi), trifluoracetato de N-(1-(2,3-dioleiloxi)propil)-N-2-(esperminacarboxamido)etil)-N,N-dimetilamonio (DOSPA), dioctadecilamidoglicil carboxispermina (DOGS), 1,2-dioleoil-sn-glicerol-3-fosfoetanolamina (DOPE), 1,2-dioleoil-3-dimetilamonio propano (DODAP), Bromuro de N-(1,2-dimiristiloxiprop-3-il)-N,N-dimetil-N-hidroxiethylamonio (DMRIE) y 2,2-Dilinoileil-4-dimetilaminoetil-[1,3]-dioxolano (XTC). Además, pueden utilizarse preparaciones comerciales de lípidos catiónicos, tales como, p. ej., LIPOFECTIN (que incluye DOTMA y DOPE, disponible en GIBCO/BRL) y Lipofectamine (que comprende DOSPA y DOPE, disponible en GIBCO/BRL).

Otros lípidos catiónicos adecuados se desvelan en las publicaciones internacionales n.º WO 09/086558, WO 09/127060, WO 10/048536, WO 10/054406, WO 10/088537, WO 10/129709 y WO 2011/153493; en las publicaciones de patente de Estados Unidos n.º 2011/0256175, 2012/0128760 y 2012/0027803; en las patentes de Estados Unidos n.º 8.158.601; y en Love *et al.*, PNAS, 107(5), 1864-69, 2010.

Otros lípidos amino adecuados, incluyen los que tienen grupos de ácidos grasos alternativos y otros grupos de dialquilamino, incluyendo aquellos, en los que los sustituyentes de alquilo son diferentes (p. ej., N-etil-N-metilamino y N-propil-N-etilamino-). En general, los lípidos amino que tienen cadenas de acilo menos saturadas se dimensionan más fácilmente, particularmente cuando los complejos deben tener un tamaño inferior a aproximadamente 0,3 micrómetros, para fines de esterilización de filtros. Pueden utilizarse lípidos amino que contengan ácidos grasos insaturados con longitudes de cadena de carbono en el intervalo de C14 a C22. También pueden utilizarse otros armazones para separar el grupo amino y la parte de ácido graso o alquilo graso del lípido amino.

Preferentemente, la NPL comprende el lípido catiónico de fórmula (III) de acuerdo con la solicitud de patente PCT/EP2017/064066.

Preferentemente, los lípidos amino o catiónicos de la divulgación tienen al menos un grupo protonable o desprotonable, de modo que el lípido esté cargado positivamente a un pH igual o inferior al pH fisiológico (p. ej., pH 7,4), y neutro a un segundo pH, preferentemente a un pH fisiológico o por encima de este. Por supuesto, se entenderá que la adición o eliminación de protones en función del pH es un proceso de equilibrio, y que la referencia a un lípido cargado o neutro se refiere a la naturaleza de la especie predominante y no requiere que todo el lípido esté presente en la forma cargada o neutra. Los lípidos que tienen más de un grupo protonable o desprotonable, o que son zwitteriónicos, no están excluidos del uso en la divulgación. En determinadas realizaciones, los lípidos protonables tienen un pKa del grupo protonable en el intervalo de aproximadamente 4 a aproximadamente 11, p. ej., un pKa de aproximadamente 5 a aproximadamente 7.

El lípido catiónico puede comprender de aproximadamente 20 % en moles a aproximadamente 70 o 75 % en moles o de aproximadamente 45 a aproximadamente 65 % en moles o aproximadamente 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, o aproximadamente 70 % en moles del lípido total presente en la partícula. En otra realización, las nanopartículas lipídicas incluyen de aproximadamente 25 % a aproximadamente 75 % en base molar de lípido catiónico, p. ej., de aproximadamente 20 a aproximadamente 70 %, de aproximadamente 35 a aproximadamente 65 %, de aproximadamente 45 a aproximadamente 65 %, aproximadamente 60 %, aproximadamente 57,5 %, aproximadamente 57,1 %, aproximadamente 50 % o aproximadamente 40 % en base molar (basado en el 100 % de moles totales de lípido en la nanopartícula lipídica). En una realización, la proporción entre lípido catiónico y ácido nucleico es de aproximadamente 3 a aproximadamente 15, tal como de aproximadamente 5 a aproximadamente 13 o de aproximadamente 7 a aproximadamente 11.

Composiciones farmacéuticas

Preferentemente, la divulgación proporciona composiciones farmacéuticas que contienen un ARNm optimizado con codones que codifica una proteína OTC humana de SEQ ID NO: 3 o SEQ ID NO: 4, preferentemente formuladas en

un sistema de suministro o transportador de lípidos y que preferentemente comprende excipientes farmacéuticamente aceptables. Las composiciones farmacéuticas desveladas en el presente documento facilitan preferentemente la expresión de ARNm *in vivo*.

5 Como vías de administración adecuadas se incluyen, por ejemplo, administración oral, rectal, vaginal, transmucosa, pulmonar, incluida intratraqueal o inhalada, o intestinal; suministro parenteral, incluyendo inyecciones intradérmicas, transdérmicas (tópicas), intramusculares, subcutáneas, intramedulares, así como intratecales, intraventriculares directas, intravenosas, intraperitoneales o intranasales.

10 Preferentemente, la administración intramuscular es en un músculo seleccionado del grupo que consiste en músculo esquelético, músculo liso y músculo cardíaco. En algunas realizaciones, la administración da como resultado el suministro de ARNm a un miocito. En algunas realizaciones, la administración da como resultado el suministro del ARNm a un hepatocito (es decir, una célula hepática). En una realización particular, la administración intramuscular da como resultado el suministro del ARNm a un miocito.

15 Preferentemente, los ARNm y sus formulaciones lipídicas pueden administrarse de manera local en lugar de sistémica, por ejemplo, mediante inyección de la composición farmacéutica directamente en un tejido específico, preferentemente en una formulación de liberación sostenida. El suministro local puede verse afectado de varias maneras, dependiendo del tejido específico al que se dirija. Por ejemplo, los aerosoles que contienen las composiciones de la presente divulgación pueden inhalarse (mediante suministro nasal, traqueal o bronquial); pueden suministrarse en forma de líquido, comprimido o cápsula para su administración en el estómago o en el intestino, pueden suministrarse en forma de supositorio mediante aplicación rectal o vaginal; o incluso pueden suministrarse al ojo usando cremas, gotas o incluso inyecciones. Las formulaciones que contienen las composiciones proporcionadas formando complejo con moléculas o ligandos terapéuticos, pueden incluso administrarse quirúrgicamente, por ejemplo junto con un polímero u otra estructura o sustancia que pueda permitir que las composiciones se difundan desde el lugar de la implantación a las células circundantes. Como alternativa, pueden aplicarse quirúrgicamente sin tener que utilizar polímeros o soportes.

20 Las composiciones farmacéuticas pueden administrarse a cualquier tejido deseado. En algunas realizaciones, el ARNm de **OTC** suministrado mediante los liposomas o las composiciones que se proporcionan, se expresa en el tejido en el que se administraron los liposomas y/o las composiciones. En algunas realizaciones, el ARNm suministrado se expresa en un tejido diferente del tejido en el que se administraron los liposomas y/o las composiciones. Como ejemplos de tejidos en los que el ARNm suministrado puede suministrarse y/ expresarse se incluyen, pero sin limitación, hígado, riñón, corazón, bazo, suero, cerebro, músculo esquelético, ganglios linfáticos, piel y/o líquido cefalorraquídeo.

25 Preferentemente, una composición farmacéutica puede contener un polinucleótido descrito en el presente documento, tal como una construcción de ADN primaria o ARNm descrito en el presente documento, dentro de un vector vírico o bacteriano.

30 Preferentemente, la construcción de ADN primaria para un ARNm descrito en el presente documento o de un ARNm descrito en el presente documento, puede formularse usando uno o más excipientes para: (1) aumentar la estabilidad; (2) aumentar la transfección celular; (3) permitir una liberación sostenida o retardada (p. ej., a partir de una formulación de depósito del polinucleótido, de la construcción primaria o del ARNm); (4) alterar la biodistribución (p. ej., dirigir el polinucleótido, la construcción primaria o el ARNm, a tipos de tejidos o de células específicos); (5) aumentar la traducción de la proteína codificada *in vivo*; y/o (6) alterar el perfil de liberación de la proteína codificada *in vivo*.

35 Además de los excipientes tradicionales, tales como todos y cada uno de los disolventes, medios de dispersión, diluyentes u otros vehículos líquidos, adyuvantes de dispersión o suspensión, agentes tensioactivos, agentes isotónicos, agentes espesantes o emulsionantes, conservantes, los excipientes de la presente divulgación pueden incluir, sin limitación, liposomas, nanopartículas lipídicas, polímeros, lipoplexos, nanopartículas con carcasa y núcleo, péptidos, proteínas, células transfectadas con una construcción de ADN primaria, o ARNm (p. ej., para su trasplante a un sujeto), hialuronidasa, miméticos de nanopartículas y combinaciones de los mismos.

40 En consecuencia, las formulaciones descritas en el presente documento pueden incluir uno o más excipientes, cada uno en una cantidad que en conjunto aumenta la estabilidad de la construcción de ADN primaria, o ARNm, aumenta la transfección celular mediante la construcción primaria o ARNm, aumenta la expresión del polinucleótido, de la construcción primaria o de la proteína codificada por ARNm, y/o altera el perfil de liberación del polinucleótido, de la construcción primaria o de las proteínas codificadas por ARNm. Asimismo, la construcción primaria y el ARNm de la presente divulgación pueden formularse usando nanopartículas de ácido nucleico autoensambladas.

45 Las formulaciones de las composiciones farmacéuticas descritas en el presente documento se pueden preparar mediante cualquier método conocido o desarrollado en el futuro en la técnica de la farmacología. En general, dichos métodos preparativos incluyen la etapa de asociar el principio activo con un excipiente y/o con uno o más principios auxiliares.

Una composición farmacéutica de acuerdo con la presente divulgación puede prepararse, envasarse y/o comercializarse a granel, como dosis unitaria individual, y/o como una pluralidad de dosis unitarias individuales. Como se utiliza en el presente documento, una "dosis unitaria" se refiere a una cantidad aislada de la composición farmacéutica que comprende una cantidad predeterminada del principio activo. La cantidad del principio activo puede ser generalmente igual a la dosificación del principio activo que se administraría a un sujeto y/o una fracción conveniente de dicha dosificación tal como, pero sin limitación, una mitad o un tercio de dicha dosificación.

Las cantidades relativas del principio activo, del excipiente farmacéuticamente aceptable y/o de cualquier ingrediente adicional en una composición farmacéutica de acuerdo con la divulgación variarán, dependiendo de la identidad, del tamaño y/o de la afección del sujeto que se esté tratando y, adicionalmente, dependiendo de la vía por la que se administre la composición. Por ejemplo, la composición puede comprender entre el 0,1 % y el 99 % (p/p) del principio activo.

Las formulaciones farmacéuticas puede comprender adicionalmente un excipiente farmacéuticamente aceptable, que, como se utiliza en el presente documento, incluye, pero sin limitación, todos y cada uno de los disolventes, medios de dispersión, diluyentes u otros vehículos líquidos, adyuvantes de dispersión o suspensión, agentes tensioactivos, agentes isotónicos, agentes espesantes o emulsionantes, conservantes, y similares, según sea adecuado para la forma farmacéutica particular deseada.

En la técnica se conocen diversos excipientes para formular composiciones farmacéuticas y técnicas para preparar la composición (véase Remington: The Science and Practice of Pharmacy, 21ª Edición, A. R. Gennaro, Lippincott, Williams & Wilkins, Baltimore, Md., 2006;) El uso de un medio excipiente convencional puede contemplarse dentro del alcance de las realizaciones de la presente divulgación, excepto en la medida en que cualquier medio excipiente convencional pueda ser incompatible con una sustancia o sus derivados, tal como mediante la producción de cualquier efecto biológico no deseable o, de otro modo, la interacción de manera perjudicial con cualquier otro componente o componentes de la composición farmacéutica.

Usos terapéuticos

Las secuencias de ARNm, las construcciones primarias de ADN que transcriben las secuencias de ARNm descritas en el presente documento y sus composiciones farmacéuticas, proporcionan numerosos métodos *in vivo* e *in vitro* y son útiles para tratar la deficiencia de OTC. El tratamiento puede comprender tratar a un paciente humano con deficiencia de OTC. De manera similar, las composiciones descritas en el presente documento, pueden utilizarse *in vitro* o *ex vivo* para estudiar la deficiencia de OTC en modelos celulares o animales. Por ejemplo, las células deficientes con respecto a la expresión de OTC pueden utilizarse para analizar la capacidad de restablecer la expresión y/o actividad de OTC, así como el período de tiempo durante el cual persiste la expresión y/o actividad. Dichas células y modelos animales también son adecuados para identificar otros factores implicados en la ruta, ya sean compañeros de unión o factores en la misma ruta bioquímica. En otras realizaciones, las composiciones descritas en el presente documento pueden utilizarse para estudiar o rastrear el suministro mitocondrial.

Los polinucleótidos descritos en el presente documento, tal como una construcción o molde de ADN o una secuencia de ARNm descrita en el presente documento, pueden suministrarse a pacientes o a células que padecen deficiencia de OTC. Preferentemente, la secuencia de ARNm comprende la SEQ ID NO: 119 que codifica una proteína OTC modificada de la SEQ ID NO: 4. Preferentemente, la secuencia de ADN comprende la SEQ ID NO: 221 que codifica una proteína OTC modificada de la SEQ ID NO: 4.

Después de la administración, la OTC se expresa en las células o en el sujeto. Preferentemente, las composiciones descritas en el presente documento se suministran a las mitocondrias. Preferentemente, las composiciones descritas en el presente documento se suministran a las células del hígado.

Preferentemente, los métodos terapéuticos descritos en el presente documento disminuyen los niveles de amoniaco en plasma y/u orina en un sujeto que lo necesita o en células en cultivo, tal como un sujeto que tiene una deficiencia de OTC. Preferentemente, los métodos terapéuticos descritos en el presente documento disminuyen los niveles de ácido orótico en plasma y/u orina en un sujeto que lo necesita o en células en cultivo. Preferentemente, los métodos terapéuticos descritos en el presente documento aumentan la citrulina en plasma y/u orina en un sujeto que lo necesita o en células en cultivo. Preferentemente, los niveles de amoniaco, ácido orótico y/o citrulina se utilizan como biomarcadores para (i) identificar sujetos que necesitan tratamiento y/o (ii) evaluar la eficacia del tratamiento usando los moldes de ARNm o ADN descritos en el presente documento.

Como ejemplos de secuencias de ARNm para su uso con estos métodos, se incluyen los que se enumeran en la Tabla 5. Preferentemente, los moldes de ADNc utilizados para transcribir las secuencias de ARNm descritas en el presente documento se enumeran en la Tabla 6. Las secuencias de ARNm preferidas para administrar a pacientes para el tratamiento de la deficiencia de OTC son la SEQ ID NO: 1799 y la SEQ ID NO: 1921. Preferentemente, las secuencias de ARNm preferidas de SEQ ID NO: 1799 y SEQ ID NO: 1921.

Dosificación

- Una dosis eficaz de un ARNm, una proteína o formulaciones farmacéuticas de la misma de la presente divulgación, puede ser una cantidad que sea suficiente para tratar la deficiencia de proteína ORF en una célula y/o en un paciente. Una dosis terapéuticamente eficaz puede ser una cantidad de un agente o una formulación que sea suficiente para
- 5 provocar un efecto terapéutico. Una dosis terapéuticamente eficaz puede administrarse en una o más administraciones distintas y por diferentes vías. Por lo general, una cantidad terapéuticamente eficaz es suficiente para lograr un beneficio significativo para el sujeto (p. ej., tratando, modulando, curando, impidiendo y/o mejorando la fenilcetonuria). Por ejemplo, una cantidad terapéuticamente eficaz puede ser una cantidad suficiente para lograr un efecto terapéutico y/o profiláctico deseado. Por lo general, la cantidad de un agente terapéutico administrada a un sujeto que lo necesita
- 10 dependerá de las características del sujeto. Dichas características incluyen la afección, la gravedad de la enfermedad, el estado de salud general, la edad, el género y el peso corporal del paciente. Un experto en la materia podrá determinar fácilmente las dosis apropiadas dependiendo de estos y otros factores relacionados. Adicionalmente, como opción, para identificar intervalos de dosificación óptimos, pueden emplear ensayos tanto objetivos como subjetivos.
- 15 Los métodos proporcionados en el presente documento contemplan administraciones sencillas y múltiples de una cantidad terapéuticamente eficaz de una secuencia de ARNm descrita en el presente documento. Las composiciones farmacéuticas que comprenden una secuencia de ARNm que codifica una proteína ORF descrita en el presente documento pueden administrarse a intervalos regulares, dependiendo de la naturaleza, de la gravedad y del grado de la afección del sujeto. Preferentemente, una cantidad terapéuticamente eficaz de una secuencia de ARNm de la
- 20 presente divulgación se puede administrar periódicamente a intervalos regulares (p. ej., una vez al año, una vez cada seis meses, una vez cada cuatro meses, una vez cada tres meses, una vez cada dos meses, una vez al mes), cada quince días, cada semana, cada día, dos veces al día, tres veces al día, cuatro veces al día, cinco veces al día, seis veces al día o de manera continuada.
- 25 Preferentemente, las composiciones farmacéuticas del ARNm de la presente divulgación se formulan de tal manera que sean adecuadas para la liberación prolongada del compuesto traducible que codifica una proteína modificada descrita en el presente documento contenida en el mismo. Dichas composiciones de liberación prolongada pueden administrarse convenientemente a un sujeto a intervalos de dosificación prolongados. Por ejemplo, en una realización, las composiciones farmacéuticas de la presente divulgación se administran a un sujeto dos veces al día, diariamente
- 30 o cada dos días. En algunas realizaciones, las composiciones farmacéuticas de la presente divulgación se administran a un sujeto dos veces por semana, una vez a la semana, cada 10 días, cada dos semanas, cada 28 días, cada mes, cada seis semanas, cada ocho semanas, cada dos meses, cada tres meses, cada cuatro meses, cada seis meses, cada nueve meses o una vez al año. También se contemplan en el presente documento composiciones farmacéuticas que se formulan para administración mediante depósito (p. ej., por vía subcutánea, intramuscular) para suministrar o
- 35 liberar una secuencia de ARNm que codifica una proteína OTC descrita en el presente documento durante períodos de tiempo prolongados. Preferentemente, para mejorar la estabilidad, los medios de liberación prolongada empleados se combinan con modificaciones realizadas en el compuesto traducible que codifica una proteína OTC descrita en el presente documento.
- 40 Una dosis terapéuticamente eficaz, tras la administración, puede dar lugar a niveles séricos o plasmáticos de OTC de 1 a 1000 pg/ml, o de 1 a 1000 ng/ml, o de 1 a 1000 µg/ml, o mayores. En algunas realizaciones, la administración de una dosis terapéuticamente eficaz de una composición que comprende una secuencia de ARNm descrita en el presente documento puede dar como resultado mayores niveles de proteína hepática modificada en un sujeto tratado. Preferentemente, la administración de una composición que comprende un ARNm descrito en el presente documento
- 45 da como resultado un aumento de un 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % o 95 % en los niveles de proteína hepática modificada en relación con un nivel inicial de proteína modificada en el sujeto antes del tratamiento. Preferentemente, la administración de una dosis terapéuticamente eficaz de una composición que comprende un ARNm descrito en el presente documento dará como resultado un aumento en los niveles de OTC en el hígado en relación con los niveles basales de OTC en el hígado en el sujeto antes del tratamiento. En algunas
- 50 realizaciones, el aumento de los niveles de OTC en el hígado en relación con los niveles iniciales de OTC en el hígado será de al menos un 5 %, 10 %, 20 %, 30 %, 40 %, 50 %, 100 %, 200 % o mayor.
- Preferentemente, una dosis terapéuticamente eficaz, cuando se administra regularmente, da como resultado en una mayor expresión de OTC en el hígado en comparación con los niveles iniciales antes del tratamiento. Preferentemente,
- 55 la administración de una dosis terapéuticamente eficaz de una composición que comprende una secuencia de ARNm descrita en el presente documento, da como resultado la expresión de un nivel de proteína modificada igual o superior a aproximadamente 10 ng/mg, aproximadamente 20 ng/mg, aproximadamente 50 ng/mg, aproximadamente 100 ng/mg, aproximadamente 150 ng/mg, aproximadamente 200 ng/mg, aproximadamente 250 ng/mg, aproximadamente 300 ng/mg, aproximadamente 350 ng/mg, aproximadamente 400 ng/mg, aproximadamente 450 ng/mg,
- 60 aproximadamente 500 ng/mg, aproximadamente 600 ng/mg, aproximadamente 700 ng/mg, aproximadamente 800 ng/mg, aproximadamente 900 ng/mg, aproximadamente 1000 ng/mg, aproximadamente 1200 ng/mg o aproximadamente 1500 ng/mg de la proteína total en el hígado de un sujeto tratado.
- Preferentemente, una dosis terapéuticamente eficaz, cuando se administra regularmente, da como resultado una
- 65 reducción de los niveles de ácido orótico en una muestra biológica. En algunas realizaciones, la administración de una dosis terapéuticamente eficaz de una composición que comprende un ARNm descrito en el presente documento, da

como resultado una reducción de los niveles de ácido orótico en una muestra biológica (p. ej., una muestra de orina, plasma o suero) en al menos aproximadamente un 5 %, al menos aproximadamente el 10 %, al menos aproximadamente el 15 %, al menos aproximadamente el 20 %, al menos aproximadamente el 25 %, al menos aproximadamente el 30 %, al menos aproximadamente el 35 %, al menos aproximadamente el 40 %, al menos aproximadamente el 45 %, al menos aproximadamente el 50 %, al menos aproximadamente el 55 %, al menos aproximadamente el 60 %, al menos aproximadamente el 65 %, al menos aproximadamente el 70 %, al menos aproximadamente el 75 %, al menos aproximadamente el 80 %, al menos aproximadamente el 85 %, al menos aproximadamente el 90 % o al menos aproximadamente el 95 % en comparación con los niveles iniciales de ácido orótico antes del tratamiento. Preferentemente, la muestra biológica se selecciona del plasma, suero, sangre completa, orina o líquido cefalorraquídeo. Preferentemente, la administración de una dosis terapéuticamente eficaz de una composición que comprende un ARNm descrito en el presente documento, da como resultado una reducción de los niveles de ácido orótico a aproximadamente 1000 $\mu\text{mol/l}$ o menor, aproximadamente 900 $\mu\text{mol/l}$ o menor, aproximadamente 800 $\mu\text{mol/l}$ o menor, aproximadamente 700 $\mu\text{mol/l}$ o menor, aproximadamente 600 $\mu\text{mol/l}$ o menor, aproximadamente 500 $\mu\text{mol/l}$ o menor, aproximadamente 400 $\mu\text{mol/l}$ o menor, aproximadamente 300 $\mu\text{mol/l}$ o menor, aproximadamente 200 $\mu\text{mol/l}$ o menor, aproximadamente 100 $\mu\text{mol/l}$ o menor o aproximadamente 50 $\mu\text{mol/l}$ o menor en suero o plasma. Preferentemente, una dosis terapéuticamente eficaz, cuando se administra regularmente, da como resultado una reducción de los niveles de ácido orótico a aproximadamente 600 $\mu\text{mol/l}$ o menor en suero o plasma. En otra realización ilustrativa, una dosis terapéuticamente eficaz, cuando se administra regularmente, da como resultado una reducción de los niveles de ácido orótico a aproximadamente 360 $\mu\text{mol/l}$ o menor en suero o plasma. Preferentemente, una dosis terapéuticamente eficaz, cuando se administra regularmente, da como resultado una reducción de los niveles de ácido orótico a aproximadamente 120 $\mu\text{mol/l}$ o menor en suero o plasma.

Una dosis terapéuticamente eficaz de un ARNm *in vivo* descrito en el presente documento puede ser una dosis de aproximadamente 0,001 a aproximadamente 500 mg/kg de peso corporal. Por ejemplo, la dosis terapéuticamente eficaz puede ser de aproximadamente 0,001-0,01 mg/kg de peso corporal o de 0,01-0,1 mg/kg o de 0,1-1 mg/kg o de 1-10 mg/kg o de 10-100 mg/kg. Preferentemente, un ARNm modificado con agente nucleomonomérico desbloqueado y habilitado para lípidos (LUNAR)-ARNm (véanse los documentos WO 2015/074085 y US 2018/0169268), que codifica una proteína OTC descrita en el presente documento, se proporciona en una dosis que varía de aproximadamente 0,1 a aproximadamente 10 mg/kg de peso corporal.

Combinaciones

Las construcciones primarias de ADNc, el ARNm o las proteínas OTC codificadas descritas en el presente documento, pueden utilizarse junto con uno o más agentes terapéuticos, profilácticos, de diagnóstico o de obtención de imágenes distintos. La expresión "junto con", no pretende implicar que los agentes deban administrarse al mismo tiempo y/o formularse para su suministro conjunto, aunque estos métodos de suministro están dentro del alcance de la presente divulgación. Las composiciones pueden administrarse simultáneamente con, antes de o después de, uno o más agentes terapéuticos o procedimientos médicos deseados distintos. En general, cada agente se administrará a una dosis y/o siguiendo una pauta determinada para ese agente. Preferentemente, los métodos de tratamiento de la presente divulgación abarcan el suministro de composiciones farmacéuticas, profilácticas, de diagnóstico o de obtención de imágenes, junto con agentes que pueden mejorar su biodisponibilidad, reducir y/o modificar su metabolismo, inhibir su excreción y/o modificar su distribución en el organismo. Como ejemplo no limitativo, el ARNm desvelado en el presente documento y preferentemente una secuencia de ARNm que comprende la SEQ ID NO: 119, que codifica una proteína OTC modificada de SEQ ID NO: 4 puede utilizarse junto con un agente farmacéutico para el tratamiento de la deficiencia de OTC. El agente farmacéutico incluye, pero sin limitación, uno o más de: fenilbutirato de sodio, fenilbutirato de glicerol, fenil acetato sódico, benzoato de sodio, arginina, citrulina, múltiples vitaminas, complementos de calcio o combinados con una dieta baja en proteínas/alta en calorías. En general, se espera que los agentes utilizados en combinación se utilicen a niveles que no superen los niveles a los que se utilicen individualmente. En algunas realizaciones, los niveles utilizados en combinación serán menores que los utilizados individualmente. En una realización, las combinaciones, cada una o conjuntamente, pueden administrarse según los regímenes de dosificación fraccionada conocidos en la técnica.

Ejemplos

Ejemplo 1: Material y Métodos

Protocolo de transcripción *in vitro*

Los ARNm se sintetizaron *in vitro* utilizando la transcripción de ARN dependiente de ADN mediada por la ARN polimerasa T7, donde el trifosfato de uridina (UTP) se sustituyó y no se sustituyó por UTP modificados tales como 5 metoxi UTP (5MeOU), N1-metoximetil pseudo UTP (N1-MOM), 5-hidroximetil UTP, 5-carboxi UTP y mezcla de modificaciones utilizando un molde linealizado para cada combinación de UTR. El ARNm se purificó mediante cromatografía en columna, la contaminación por ADN y ARNm bicatenario de todos los ARNm se eliminó mediante una reacción enzimática, se concentró el ARNm y se intercambió el tampón.

Preparación de ARNm encapsulado en lípidos

Se prepararon partículas de ARNm encapsuladas en lípidos mezclando lípidos (lípidos ATX: DSPC: Colesterol: PEG-DMG) en etanol con ARNm de OTC disuelto en tampón citrato. El material mezclado se diluyó instantáneamente con tampón fosfato. El etanol se eliminó mediante diálisis contra tampón fosfato usando una membrana de celulosa regenerada (MWCO, *molecular weight cutoff*, límite de peso molecular de 100 kD) o mediante filtración de flujo tangencial (FFT) usando membranas de fibra hueca de polietersulfona modificada (mPES, *modified polyethersulfone*) (MWCO de 100 kD). Una vez que se eliminó completamente el etanol, el tampón se intercambió con tampón HEPES (ácido 4-(2-hidroxietil)-1-piperazina etano sulfónico) que contenía NaCl 50 mM y sacarosa al 9 %, a un pH de 7,3. La formulación se concentró seguido de una filtración usando filtros de PES de 0,2 µm. Después, se midió la concentración de ARNm en la formulación mediante el ensayo fluorimétrico Ribogreen, después de lo cual la concentración se ajustó a una concentración final deseada diluyendo con tampón HEPES que contenía NaCl 50 mM, sacarosa al 9 %, a un pH de 7,3 que contenía glicerol. Después, la formulación final se filtró a través de un filtro de 0,2 µm y se vertió en viales de vidrio, tapados, protegidos y llevados a $-70 \pm 5^\circ\text{C}$. Las formulaciones congeladas se caracterizaron con respecto a su contenido de ARNm y porcentaje de encapsulación mediante ensayo Ribogreen, a su integridad del ARNm mediante un analizador de fragmentos, a su contenido de lípidos mediante cromatografía líquida de alta resolución (HPLC), a su tamaño de partícula mediante dispersión dinámica de luz en un instrumento Zetasizer Nano ZS de Malvern, a su pH y osmolalidad.

Western en células (WEC)

Se usaron placas de colágeno de 96 pocillos para sembrar las células a la densidad adecuada en medios de cultivo de medio Eagle modificado de Dulbecco (DMEM)/suero bovino fetal (FBS). Al llegar a la confluencia óptima, las células se transfectaron con los ARNm específicos diluidos en la mezcla de reactivos de transfección (MessengerMax y Opti-MEM). Las células se llevaron a una incubadora con CO₂ y se dejó que crecieran. En el punto temporal deseado, los medios se eliminaron y las células se fijaron en paraformaldehído (PFA) reciente al 4 % durante 20 minutos. Después de esto, se eliminó el fijador y las células se permeabilizaron en solución salina tamponada con tris con TWEEN (TBST) durante 5 minutos varias veces. Cuando finalizaron los lavados de permeabilización, las células se incubaron con un tampón de bloqueo (tampón de bloqueo ODYSSEY® (PBS) (Li-Cor, Lincoln, NE)) durante 45 min. Después se añadió el anticuerpo primario y se incubó durante 1 hora a temperatura ambiente. Posteriormente, las células se lavaron varias veces en TBST y se incubaron durante 1 h con un anticuerpo secundario diluido en tampón de bloqueo y que contenía una tinción CellTag 700. Para finalizar, las células se lavaron varias veces en TBST seguido de un último lavado en solución salina tamponada con Tris (TBS). Se tomaron imágenes de la placa utilizando el sistema de detección Licor y los datos se normalizaron al número total de células marcadas con la tinción CellTag 700.

Ejemplo 2: Exploración de las UTR en Hepa1,6 y Hep3B - Correlación a las 24h y 48h.

Se exploró una biblioteca de UTR *in vitro* usando la construcción de ARNm n.º 571 que, como CDS (secuencia codificante), comprendía la secuencia de SEQ ID NO: 34. Para transfectar los diferentes ARNm en células Hepa1,6 y Hep3B, se utilizaron ensayos Western en células como se describe en el Ejemplo 1, usando reactivos de transfección disponibles en el comercio. Los niveles de expresión de proteína OTC se midieron mediante sistemas de obtención de imágenes fluorescentes en el infrarrojo cercano. Para la detección se utilizaron anticuerpos contra OTC disponibles en el comercio. Como controles internos, se utilizaron secuencias no transfectadas y de referencia. La figura 1, panel A, es un diagrama de dispersión de los niveles de expresión de proteína OTC encontrados en células Hepa1,6 y Hep3B a las 24 horas. La figura 1, panel B, es un diagrama de dispersión de los niveles de expresión de proteína OTC encontrados en células Hepa1,6 y Hep3B a las 48 horas. El objetivo de la exploración era determinar un impacto específico de la UTR sobre los niveles de expresión de OTC en una línea celular de hígado humano (Hep3B) y otra de ratón (Hepa1,6), lo que será beneficioso para determinar cuáles son las UTR con mejor funcionamiento en ambos modelos, en particular, su traducibilidad de ratón a ser humano. Las UTR de máxima expresión se utilizaron en estudios posteriores de elaboración de perfiles.

Ejemplo 3: Ronda 1 de compuestos de estabilidad de proteínas explorados en Hepa1,6 y Hep3B a las 24 h - Correlación.

Se realizó una exploración *in vitro* de determinadas construcciones de ARNm de la Tabla 5 que se diseñaron basándose en un enfoque de estabilidad de proteínas. Las construcciones de ARNm se analizaron en dos químicas diferentes: N'-metilpseudouridina (N1MPU) y 5-metoxiuridina (5MeOU), lo que significa que el 100 % de las uridinas en cada ARNm eran solo N1MPU o solo 5MeOU (no una combinación de 5MeOU o N1MPU). Para transfectar los diferentes ARNm en células Hepa1,6 y Hep3B, se utilizaron ensayos Western en células (WEC) como se describe en el Ejemplo 1, usando reactivos de transfección disponibles en el comercio. Los niveles de expresión de proteína OTC se midieron mediante sistemas de obtención de imágenes fluorescentes en el infrarrojo cercano. Para la detección se utilizaron anticuerpos contra OTC disponibles en el comercio. Como controles internos, se utilizaron secuencias no transfectadas y de referencia. La figura 2, panel A, es un diagrama de dispersión que muestra la correlación de los niveles de expresión de proteína OTC en células Hepa1,6 a las 24 horas en función de los ARNm analizados en las químicas N1MPU y 5MeOU. La figura 2, panel B, es un diagrama de dispersión que muestra la correlación de los niveles de expresión de proteína OTC en células Hep3B a las 24 horas en función de los ARNm analizados en las químicas N1MPU y 5MeOU. Estas cifras muestran el grado de variabilidad de los niveles de expresión cuando los

ARNm de dos químicas diferentes se analizan en una línea celular de hígado de ratón y de ser humano. Se puede observar que, en este experimento, la mayoría de los compuestos se expresan mejor cuando se utiliza una química N1MPU en el ARNm.

5 Ejemplo 4: Ronda 2 de compuestos de estabilidad de proteínas analizados en hepatocitos primarios humanos a las 24 h y 48 h - Correlación.

Se realizó una exploración *in vitro* de determinadas construcciones de ARNm de la Tabla 5 que se diseñaron basándose en un enfoque de estabilidad de proteínas. Los ARNm se analizaron en dos químicas diferentes, el 100 % de las uridinas que son N1MPU se indica con el nombre de las construcciones de ARNm seguido de ".1" y el 100 % de las uridinas que son 5MeOU se indica con el nombre de las construcciones de ARNm seguido de ".7". Para transfectar los diferentes ARNm en hepatocitos primarios humanos, se utilizaron ensayos Western en células (WEC) como se describe en el Ejemplo 1, usando reactivos de transfección disponibles en el comercio. Los niveles de expresión de proteína OTC se midieron mediante sistemas de obtención de imágenes fluorescentes en el infrarrojo cercano. Para la detección se utilizaron anticuerpos contra OTC disponibles en el comercio. Como controles internos, se utilizaron secuencias no transfectadas y de referencia. (Figura 3, Paneles A y B). Los resultados indican que, a diferencia de los experimentos realizados en líneas celulares de cáncer (Hepa1,6; Hep3B; Ejemplo 3), ambas químicas se expresan de manera similar en hepatocitos primarios humanos.

20 Ejemplo 5: Ronda 3 de exploración de compuestos de estabilidad de proteínas en hepatocitos primarios humanos a las 24 h y 48 h - Correlación.

Se realizó una exploración *in vitro* de nuevos compuestos diseñados basándose en un enfoque de estabilidad de proteínas. Los ARNm se analizaron en dos químicas diferentes, N1MPU indicada con el nombre de las construcciones de ARNm seguido de ".1" y 5MeOU indicada con el nombre de las construcciones de ARNm seguido de ".7". Para transfectar los diferentes ARNm en hepatocitos primarios humanos, se utilizaron ensayos Western en células (WEC) como se describe en el Ejemplo 1, usando reactivos de transfección disponibles en el comercio. Los niveles de expresión de proteína OTC se midieron mediante sistemas de obtención de imágenes fluorescentes en el infrarrojo cercano. Para la detección se utilizaron anticuerpos contra la proteína OTC disponibles en el comercio. Como controles internos, se utilizaron secuencias no transfectadas y de referencia. (Figura 4, Paneles A y B). Los resultados indican que, a diferencia de los experimentos realizados en líneas celulares de cáncer (Hepa1,6; Hep3B; Ejemplo 3), ambas químicas se expresan de manera similar en hepatocitos primarios humanos.

35 Ejemplo 6: Niveles de expresión de proteína OTC en células primarias humanas transfectadas con construcciones de ARNm de OTC 1799.7 (química 5MeOU) que codifican la proteína OTC de SEQ ID NO: 3 y 1921.7 (química 5MeOU) que codifican la proteína OTC modificada de SEQ ID NO: 4.

Para transfectar los ARNm de OTC en hepatocitos primarios humanos, se utilizaron ensayos Western en células (WEC) como se describe en el Ejemplo 1, usando reactivos de transfección disponibles en el comercio. Los niveles de expresión de proteína OTC se midieron mediante sistemas de obtención de imágenes fluorescentes en el infrarrojo cercano durante un estudio de evolución temporal de hasta 96 h. Para la detección, se utilizaron anticuerpos contra OTC disponibles en el comercio. Como control interno, se utilizaron células no transfectadas. El gráfico muestra niveles de proteína OTC normalizados a controles no transfectados. (Figura 5). El propósito de este estudio era evaluar la semivida de la secuencia de proteína no modificada frente a la modificada (codificada por las construcciones 1799.7, 1921.7, respectivamente) en condiciones *in vitro* en hepatocitos primarios humanos transfectados. Los resultados indican que 1921,7 demostró una expresión más estable que 1799,7.

Ejemplo 7: Niveles de expresión de OTC medidos mediante espectrometría de masas con monitorización de reacciones múltiples (MRM) en ratones *spf/ash* con dosis de 10 mg/kg. Ratones *spf/ash* recibieron una inyección i.v. de PBS o de ARNm de OTC formulados con lípidos (como se describe en el Ejemplo 1) a una dosis de 10 mg/kg. Para determinar los niveles endógenos, se utilizaron ratones de TS como controles internos. Se realizó una evolución temporal (6 h, 24 h y 48 h) y los niveles de expresión se midieron mediante MRM utilizando epítomos específicos de OTC humana y de ratón. Se realizaron gráficos que representaban la cantidad de proteína (ng/mg de tejido) detectada por MRM específica para OTC humana (figura 6, panel A) o de ratón (figura 6, panel B). Para medir los niveles totales de OTC en ambas especies, se diseñaron péptidos pesados específicos de ser humano y de ratón. Este conjunto de datos muestra que se detectan niveles cuantitativos de OTC específica de ser humano derivada de la traducción de los ARNm suministrados. Esta expresión mantiene un alto nivel de estabilidad de hasta 48 h.

60 Ejemplo 8: Niveles de expresión de OTC medidos mediante transferencia Western en ratones de TS con dosis de 3 mg/kg.

Ratones *spf/ash* recibieron una inyección i.v. de solución salina tamponada con fosfato (PBS) o de ARNm de OTC formulados con lípidos a una dosis de 3 mg/kg utilizando dos químicas diferentes (N1MPU y 5MeOU). Para determinar los niveles endógenos, se utilizaron ratones de TS como controles internos. Los animales se sacrificaron 24 h después de la dosis. Los niveles de expresión de OTC se midieron mediante transferencia Western (WB, *Western Blot*) utilizando un anticuerpo específico contra OTC. En los resultados proporcionados en la figura 7, las barras representan

el porcentaje de expresión en relación con los niveles de TS (100 %). Los datos generados en esta figura muestran que se lograron niveles de OTC total de TS en un acervo de ratón para varias secuencias optimizadas con codones.

Ejemplo 9: Niveles de expresión de OTC medidos por MRM en un estudio de búsqueda de intervalo de dosis.

Ratones Balb/c recibieron una inyección i.v. de PBS o de ARNm de OTC formulados con lípidos a tres dosis diferentes: 0,3 mg/kg, 1 mg/kg y 3 mg/kg y usando dos químicas diferentes (N1MPU y 5MeOU). Los animales se sacrificaron 24 h después de la dosis y los niveles de expresión se midieron mediante MRM utilizando epítomos específicos de OTC humana y de ratón. El gráfico de la figura 8 representa el porcentaje de expresión de la OTC humana (ng) por mg de tejido hepático en ratones Balb/c. La línea de puntos horizontal representa los niveles relativos de OTC en ratones Balb/c. (Figura 8). Los niveles de expresión de la proteína OTCh para la construcción de ARNm 713 5MeOU y la construcción de ARNm 571 N1MPU, se indican con flechas en la figura 8. En esta figura, para determinar cuantitativamente los niveles de proteína OTC selectivos de ser humano y selectivos de ratón se utilizó MRM. Los datos generados en esta figura muestran, de una manera dependiente de la dosis, que los niveles de TS de la OTC humana en un acervo de ratón, se logran con las secuencias optimizadas con codones desveladas en el presente documento.

Ejemplo 10: Niveles de orotato urinario medidos en ratones spf/ash tratados con PBS.

Ratones spf/ash recibieron una inyección i.v. de PBS o de la construcción 1799.7 de ARNm de OTC formulada con lípidos (química 5MeOU) a tres dosis diferentes: 0,3 mg/kg, 1 mg/kg y 3 mg/kg. Para determinar los niveles iniciales y elevados de orotato urinario, se utilizaron ratones de TS y spf/ash, respectivamente. Se determinó una evolución temporal de spf/ash y se midieron los niveles de orotato urinario en cada punto temporal. En la figura 9 pueden observarse los resultados. El orotato urinario se normalizó frente a la creatinina, que en el gráfico se representa en el eje Y a lo largo de la evolución temporal y sirve como prueba demostrativa de restablecimiento funcional de la actividad de OTC después de la inyección. A 3 mg/kg, se observó una reducción sostenible de los niveles de orotato urinario de hasta 14 días.

Ejemplo 11: Análisis farmacocinético/farmacodinámico (FC/FD) que compara los niveles de expresión de la OTC humana y del orotato urinario a las 96 h.

Ratones Spf/ash recibieron una inyección i.v. de PBS o de determinados ARNm de OTC formulados con los lípidos de la Tabla 5 a 1 mg/kg y 3 mg/kg usando dos químicas diferentes (N1MPU y 5MeOU). Como controles internos se utilizaron ratones de TS. Los niveles de OTC específica de ser humano se midieron mediante MRM, mientras que los del orotato urinario se determinaron en cada muestra y se normalizaron a los de la creatinina. El análisis FC/FD se representa gráficamente en la figura 10. El análisis FC/FD muestra la correlación entre los niveles de expresión de proteínas y la reducción del orotato urinario de una manera específica del compuesto. La construcción 1799.7 muestra una alta correlación de FC/FD.

Ejemplo 12: Fraccionamiento de muestras *in vivo* de ratones spf/ash tratados con ARNm seleccionados.

Ratones spf/ash recibieron una inyección i.v. de PBS o de ARNm de OTC formulados con lípidos (como se describe en el Ejemplo 1) a 1 mg/kg y 3 mg/kg. Como controles internos se utilizaron ratones de TS. El fraccionamiento de muestras se realizó en las muestras de hígado, separando una fracción citosólica y otra mitocondrial. Los niveles de OTC se midieron mediante WB usando anticuerpos humanos (OTCh) específicos y de reacción cruzada (OTCrc) (Figura 11). Como control mitocondrial se utilizó ciclooxigenasa IV (CoxIV). Los niveles de expresión de proteína OTC se midieron mediante sistemas de obtención de imágenes fluorescentes en el infrarrojo cercano y se normalizaron con respecto a los de la proteína total. WB indica diferencias en los niveles de expresión de OTC dentro de las fracciones mitocondriales y citosólicas cuando se suministraron los ARNm 2016 y 2260 a ratones spf/ash. Estos resultados indican que ambos compuestos pueden dirigirse de manera eficaz a las mitocondrias.

Ejemplo 13: Gráfico de las fracciones mitocondriales frente a citosólicas de muestras de ratones spf/ash tratados con las construcciones de ARNm 2016 y 2260.

Ratones spf/ash recibieron una inyección i.v. de PBS o de ARNm de OTC formulados con lípidos a 3 mg/kg. Como controles internos se utilizaron ratones de TS. El fraccionamiento de muestras se realizó en las muestras de hígado, separando una fracción citosólica y otra mitocondrial. Los niveles de OTC se midieron mediante transferencia Western utilizando un anticuerpo humano específico. Los niveles de expresión de proteína OTC se midieron mediante sistemas de obtención de imágenes fluorescentes en el infrarrojo cercano y ambas fracciones se normalizaron con respecto a los de la proteína total (FIG. 12). La gráfica de los niveles de expresión de proteína que se muestran en la figura 11 (ejemplo 12), indican que aunque los dos compuestos, 2016 y 2260, suministran niveles similares de proteína en el citosol, es en las mitocondrias donde el compuesto 2260 suministra más OTC humana que el compuesto 2016. El compuesto 2260 incluye una secuencia peptídica de señalización mitocondrial modificada de la invención.

Ejemplo 14: Niveles de orotato urinario en ratones spf/ash tratados con la construcción de ARNm 2260.

Ratones spf/ash recibieron una inyección i.v. de PBS o de ARNm de OTC formulado con lípidos (como se describe en el Ejemplo 1) a 1 mg/kg y 3 mg/kg. Como control interno se incluyeron ratones de TS. Se midieron los niveles de orotato urinario a los 0, 1, 3, 7 y 14 días y los niveles se normalizaron a los de la creatinina (Figura 13). La lectura funcional de este ensayo muestra que los niveles de orotato urinario se reducen hasta 14 días con el compuesto 2260, de una manera dependiente de la dosis.

Ejemplo 15: Supervivencia de ratones spf/ash con una dieta alta en proteínas durante el tratamiento con la construcción 1799,7 de ARNm de OTC (química 5MeOU).

Ratones spf/ash recibieron una inyección i.v. de PBS o de ARNm de OTC formulado con lípidos (1799.7) a tres dosis: 0,3 mg/kg, 1 mg/kg y 3 mg/kg. Los ratones estuvieron con una dieta alta en proteínas desde el día 0 hasta el final del estudio. Los animales tratados recibieron una inyección por vía intravenosa (i.v.) los días 0, 7, 14, 21 y 28 (flechas). Las tasas de supervivencia se determinaron cada semana. La gráfica de la figura 14 resume todo el desarrollo cronológico del estudio y las tasas de supervivencia observadas en los diferentes grupos. Los resultados muestran que los animales tratados con los ARNm de OTC humana descritos en el presente documento tienen una mayor probabilidad de sobrevivir durante una crisis hiperamonemiaquémica, lo que sugiere el papel protector de los ARNm de OTC descritos en el presente documento para desintoxicar a los animales del amoníaco tóxico. Esta tasa de supervivencia era dependiente de la dosis, y los animales tratados con una dosis de 3 mg/kg tuvieron una tasa de supervivencia mayor que la de los animales tratados con 1 mg/kg o 0,3 mg/kg.

Ejemplo 16: Comparación de niveles de expresión de OTCh con diferentes modificaciones.

Por vía i.v., se inyectaron dosis de la construcción 2262 de ARNm de OTC libre formulada con lípidos, a ratones macho C57BL/6 de 8 semanas de vida a una dosis de 1 mg/kg. En este estudio se utilizaron diferentes químicas como se indica en el eje inferior del gráfico proporcionado en la figura 15. Los hígados de los ratones se extirparon 24 horas después de la administración i.v. y se realizó transferencia Western utilizando un anticuerpo específico contra OTCh. Los niveles se normalizaron a los de la proteína total (Figura 15). Cada construcción se formuló como una nanopartícula lipídica que comprendía un lípido ATX como se describe en el Ejemplo 1. Este conjunto de datos indica el impacto de diferentes químicas de uridina sobre los niveles de expresión de nuestros ARNm optimizados con codones.

LISTADO DE SECUENCIAS

Secuencia codificante de ARNm de OTC humana de tipo silvestre (SEQ ID NO: 1)

AUGCUGUUUAAUCUGAGGAUCCUGUUAACAAUGCAGCUUUUAGAAAUGGUC
ACAACUUC AUGGUUCGAAAUUUUCGGUGUGGACAACCACUACAAAUAAGU
GCAGCUGAAGGGCCGUGACCUUCUCACUCUAAAAAAC
UUUACCGGAGAAGAAAUUAAUAUAUGCUAUGGCUAUCAGCAGAUCUGAAAU
UUAGGAUAAAACAGAAAGGAGAGUAUUUGCCUUUAUUGCAAGGGAAGUCCUU
AGGCAUGAUUUUUGAGAAAAGAAGUACUCGAACAAGAUUGUCUACAGAAACA
GGCUUUGCACUUCUGGGAGGACAUCUUGUUUCUUAACACACAAGAUUAUUA
UUUGGGUGUGAAUGAAAGUCUCACGGACACGGCCCGUGUAUUGUCUAGCAUG
GCAGAUGCAGUAUUGGCUCGAGUGUAUAAACAUCAGAUUUGGACACCCUGG
CUAAAGAAGCAUCCAUCCCAAUUAUCAUUGGGCUGUCAGAUUUGUACCAUCCU
AUCCAGAUCUGGCUGAUUACCUCACGCUCAGGAACACUAUAGCUCUCUGAA
AGGUCUUAACCCUCAGCUGGAUCGGGGAUGGGAACAUAUCCUGCACUCCAUA
UGAUGAGCGCAGCGAAAUUCGGAAUGCACCUUCAGGCAGCUACUCCAAAGGGU
UAUGAGCCGGAUGCUAGUGUAACCAAGUUGGCAGAGCAGUAUGCCAAAGAGA
AUGGUACCAAGCUGUUGCUGACAAAUGAUCCAUGGAAGCAGCGCAUGGAGG
CAAUGUAUUAAUUAACAGACACUUGGAUAAGCAUGGGACAAGAAGAGGAGAAG
AAAAAGCGGCUCAGGCUUCCAAGGUUACCAGGUUACAAUGAAGACUGCUA
AAGUUGCUGCCUCUGACUGGACAUUUUUAACACUGCUUGCCCAGAAAGCCAGAA
GAAGUGGAUGAUGAAGUCUUUUUAUUCUCCUCGAUCACUAGUGUCCCAGAGG
CAGAAAACAGAAAGUGGACAAUCAUGGCUGUCAUGGUGUCCCUGCUGACAGA
UUACUCACCUCAGCUCCAGAAGCCUAAAUUUUGA

Secuencia codificante de ADN de OTC humana de tipo silvestre (SEQ ID NO: 2)

ATGCTGTTTAATCTGAGGATCCTGTAAACAATGCAGCTTTTAGAAATGGTCACA
 ACTTCATGGTTCGAAATTTTCGGTGTGGACAACCACTACAAAATAAAGTGCAGCT
 GAAGGGCCGTGACCTTCTCACTCTAAAAAACTTTACCGGAGAAGAAATTAAATA
 TATGCTATGGCTATCAGCAGATCTGAAATTTAGGATAAAACAGAAAGGAGAGTA
 TTTGCCTTTATTGCAAGGGAAGTCCTTAGGCATGATTTTTGAGAAAAGAAGTACT
 CGAACAAAGATTGTCTACAGAAACAGGCTTTGCACTTCTGGGAGGACATCCTTGTT
 TTCTTACCACACAAGATATTCATTTGGGTGTGAATGAAAGTCTCACGGACACGGC
 CCGTGTATTGTCTAGCATGGCAGATGCAGTATTGGCTCGAGTGTATAAACAATCA
 GATTTGGACACCCTGGCTAAAGAAGCATCCATCCCAATTATCAATGGGCTGTCAG
 ATTTGTACCATCCTATCCAGATCCTGGCTGATTACCTCACGCTCCAGGAACACTA
 TAGCTCTCTGAAAGGTCTTACCCTCAGCTGGATCGGGGATGGGAACAATATCCTG
 CACTCCATCATGATGAGCGCAGCGAAATTCGGAATGCACCTTCAGGCAGCTACTC
 CAAAGGGTTATGAGCCGGATGCTAGTGTAACCAAGTTGGCAGAGCAGTATGCCA
 AAGAGAATGGTACCAAGCTGTTGCTGACAAATGATCCATTGGAAGCAGCGCATG
 GAGGCAATGTATTAATTACAGACACTTGGATAAGCATGGGACAAGAAGAGGAGA
 AGAAAAAGCGGCTCCAGGCTTTCCAAGGTTACCAGGTTACAATGAAGACTGCTA
 AAGTTGCTGCCTCTGACTGGACATTTTTTACACTGCTTGCCCAGAAAGCCAGAAGA
 AGTGGATGATGAAGTCTTTTATTCTCCTCGATCACTAGTGTTCCAGAGGCAGAA
 AACAGAAAGTGGACAATCATGGCTGTCATGGTGTCCCTGCTGACAGATTACTCAC
 CTCAGCTCCAGAAGCCTAAATTTTGA

Secuencia de aminoácidos de OTC humana de tipo silvestre

(El péptido señal de importación mitocondrial se indica subrayado*) (SEQ ID NO: 3)

MLFNLRLILLNNAAFRNGHNFMVRFRCGOPLQNKVQLKGRDLLTLKNFTGEEIK
 YMLWLSADLKFRICKQGEYLP LLQGKSLGMIFEKRSTRTRLSTETGFALLGGHPCFL
 TTQDIHLGVNESLTD TARVLSSMADAVLARVYKQSDLDTLAKEASIPIF
 GLSDLYHPIQILADYLT LQEHYSSLKGLT LSWIGDGNNILHSIMMSAAKFGMHLQAA
 TPKGYEPDASVTKLAEQYAKENGTKLLL TNDPLEAAHGGNVLITDTWISMGQEEEEK
 KKRLQAFQGYQVTMKTAKVAASDWTFLHCLPRKPEEVDDEVFYSRSLVFPEAENR
 KWTIMAVMVSLLDYSPQLQKPKF

5

Secuencia de aminoácidos de OTC modificada

(El péptido señal de importación mitocondrial se indica subrayado*) (SEQ ID NO: 4)

MLVFNLRILLNNAAFRNGHNFMVRFRCGOPLQNRVQLKGRDLLTLKNFTGEEI
 RYMLWLSADLKFRICKQGEYLP LLQGKSLGMIFEKRSTRTRLSTETGFALLGGHPCF
 LTTQDIHLGVNESLTD TARVLSSMADAVLARVYKQSDLDTLAKEASIPIINGLSDLYH
 PIQILADYLT LQEHYSSLKGLT LSWIGDGNNILHSIMMSAAKFGMHLQAATPKGYEP
 DASVTKLAEQYAKENGTKLLL TNDPLEAAHGGNVLITDTWISMGQEEEEKKKRLQAF
 QGYQVTMKTAKVAASDWTFLHCLPRKPEEVDDEVFYSRSLVFPEAENRKWTIMA
 VMVSLLDYSPQLQKPKF

10

TEV (SEQ ID NO: 5)

TCAACACAACATATACAAAACAAACGAATCTCAAGCAATCAAGCATTCTACTTCT
 ATTGCAGCAATTTAAATCATTTCTTTTAAAGCAAAAGCAATTTTCTGAAAATTTTC
 ACCATTTACGAACGATAG

AT1G58420 (SEQ ID NO: 6)

ATTATTACATCAAAACAAAAAAGCCGCCA

ARC5-2 (SEQ ID NO: 7)

CTTAAGGGGGCGCTGCCTACGGAGGTGGCAGCCATCTCCTTCTCGGCATCAAGCT
TACCATGGTGCCCCAGGCCCTGCTCTT_gGTCCCGCTGCTGGTGTTCCTCCCTCTGCT
TCGGCAAGTTCCCCATCTACACCATCCCCGACAAGCTGGGGCCGTGGAGCCCCAT
CGACATCCACCACCTGTCCTGCCCAACAACCTCGTGGTCGAGGACGAGGGCTG
CACCAACCTGAGCGGGTTCTCCTAC

HCV (SEQ ID NO: 8)

TGAGTGTCTGT ACAGCCTCCA GGCCCCCCCC TCCCGGGAGA GCCATAGTGG
TCTGCGGAACCGGTGAGTAC ACCGGAATTG CCGGGAAGAC TGGGTCCTTT
CTTGATAAA CCCACTCTATGCCCGGCCAT TTGGGCGTGC CCCC GCAAGA
CTGCTAGCCG AGTAGTGTTG GGTTCG

ALBÚMINA HUMANA (SEQ ID NO: 9)

AATTATTGGTTAAAGAAGTATATTAGTGCTAATTTCCTCCGTTTGTCTTAGCTTT
TCTCTTCTGTCAACCCACACGCCTTTGGCACA

5

EMCV (SEQ ID NO: 10)

CTCCCTCCCC CCCCCCTAAC GTTACTGGCC GAAGCCGCTT GGAATAAGGC
CGGTGTGCGT TTGTCTATAT GTTATTTTCC ACCATATTGC CGTCTTTTGG
CAATGTGAGG GCCCGGAAAC CTGGCCCTGT CTTCTTGACG AGCATTCCTA
GGGGTCTTTC CCCTCTCGCC AAAGGAATGC AAGGTCTGTT GAATGTCGTG

AAGGAAGCAG TTCCTCTGGA AGCTTCTTGA AGACAAACAA CGTCTGTAGC
GACCCTTTGC AGGCAGCGGA ACCCCCCACC TGGCGACAGG TGCCTCTGCG
GCCAAAAGCC ACGTGTATAA GATACACCTG CAAAGGCGGC ACAACCCAG
TGCCACGTTG TGAGTTGGAT AGTTGTGGAA AGAGTCAAAT GGCTCTCCTC
AAGCGTATTC AACAAGGGGC TGAAGGATGC CCAGAAGGTA CCCCATTGTA
TGGGATCTGA TCTGGGGCCT CGGTGCACAT GCTTTACGTG TGTTTAGTCG
AGGTTAAAAA ACGTCTAGGC CCCCCGAACC ACGGGGACGT GGTTTTCTTT
TGAAAAACAC GATGATAAT

10

HSP70-P2 (SEQ ID NO: 11)

GTCAGCTTTCAAACTCTTTGTTTCTTGTTTGTGATTGAGAATA

HSP70-M1 (SEQ ID NO: 12)

CTCTCGCCTGAGAAAAAAAATCCACGAACCAATTTCTCAGCAACCAGCAGCACG

HSP72-M2 (SEQ ID NO: 13)

ACCTGTGAGGGTTCTGAAGGAAGTAGCAGTGTTTTTTGTTTCCTAGAGGAAGAG

15

HSP17.9 (SEQ ID NO: 14)

ACACAGAAACATTCGCAAAAACAAAATCCCAGTATCAAAATTCTTCTCTTTTTTT
CATATTTTCGCAAAGAC

HSP70-P1 (SEQ ID NO: 15)

CAGAAAAATTTGCTACATTGTTTCACAAACTTCAAATATTATTCATTTATTT

XBG (SEQ ID NO: 16)

CTAGTGACTGACTAGGATCTGGTTACCACTAAACCAGCCTCAAGAACACCCGAA
TGGAGTCTCTAAGCTACATAATACCAACTTACACTTACAAAATGTTGTCCCCCAA
AATGTAGCCATTTCGTATCTGCTCCTAATAAAAAGAAAGTTTCTTCACAT

5 HAPTOGLOBINA HUMANA (SEQ ID NO: 17)

TGCAAGGCTGGCCGGAAGCCCTTGCTGAAAGCAAGATTTTCAGCCTGGAAGAGG
GCAAAGTGGACGGGAGTGGACAGGAGTGGATGCGATAAGATGTGGTTTGAAGCT
GATGGGTGCCAGCCCTGCATTGCTGAGTCAATCAATAAAGAGCTTTCTTTTGACC
CAT

APOLIPOPROTEÍNA E HUMANA (SEQ ID NO: 18)

10

ACGCCGAAGCCTGCAGCCATGCGACCCCAACGCCACCCCGTGCCTCCTGCCTCCGC
GCAGCCTGCAGCGGGAGACCCTGTCCCCGCCCGCCCGTCCCTCCTGGGGTGGAC
CCTAGTTTAATAAAGATTCACCAAGTTTCACGCA

HCV (SEQ ID NO: 19)

TAGAGCGGCAAACCCTAGCTACACTCCATAGCTAGTTTCTTTTTTTTTTTGTTTTTTT
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTT
TTCTTGGTGGCTCCATCTTAGCCCTAGTCACGGCTAGCTGTGAAAGGTCCGTGAG
CCGCATGACTGCAGAGAGTGCCGTAACCTGGTCTCTCTGCAGATCATGT

15 ALBÚMINA DE RATÓN (SEQ ID NO: 20)

ACACATCACAACCACAACCTTCTCAGGCTACCCTGAGAAAAAAGACATGAAGA
CTCAGGACTCATCTTTTCTGTTGGTGTAAAATCAACACCCTAAGGAACACAAATT

TCTTTAAACATTTGACTTCTTGTCTCTGTGCTGCAATTAATAAAAAATGGAAAGA
ATCTAC

20

ALFA GLOBINA HUMANA (SEQ ID NO: 21)

GCTGGAGCCTCGGTAGCCGTTCCCTCCTGCCCGCTGGGCCTCCCAACGGGGCCCTCC
TCCCCTCCTTGACCGGCCCTTCCTGGTCTTTGAATAAAGTCTGAGTGGGCAGCA

EMCV (SEQ ID NO: 22)

TAGTGCAGTCAC TGGCACAACG CGTTGCCCGG TAAGCCAATC GGGTATACAC
GGTCGTCATACTGCAGACAG GGTCTTCTA CTTTGCAAGA TAGTCTAGAG
TAGTAAAATA AATAGTATAAG

25

(SEQ ID NO: 23)

GCCACC

(SEQ ID NO: 24)

GCCA

(SEQ ID NO: 25)
AUAAGUGAA

>mARM563 (SEQ ID NO: 26)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUUAAUCUGAGGAUCCUG
UUAAACAAUGCAGCUUUUAGAAAUGGUCACAACUUCAUGGUUCGAAAUUUUC
GGUGUGGACAACCAACUACAAAUAAGUGCAGCUGAAGGGCCGUGACCUUCUC
ACUCUAAAAAACUUUACCGGAGAAGAAAUUAAAUAUAUGCUAUGGCUAUCAG
CAGAUUCUGAAAUUUAGGAUAAAACAGAAAGGAGAGUAUUUGCCUUUAUUGCA
AGGGAAGUCCUUAGGCAUGAUUUUUGAGAAAAGAAGUACUCGAACAAGAUUG
UCUACAGAAACAGGCUUUGCACUUCUGGGAGGACAUCUUGUUUUCUUACCAC
ACAAGAUAUUCAUUUGGGUGUGAAUGAAAGUCUCACGGACACGGCCCGUGUA
UUGUCUAGCAUGGCAGAUUGCAGUAUUGGCUCGAGUGUAUAAACAAUCAGAUU
UGGACACCCUGGCUAAAGAAGCAUCCAUCCCAAUUUAUCAUUGGGCUGUCAGAU
UUGUACCAUCCUAUCCAGAUCCUGGCUGAUUACCUCACGCUCCAGGAACACUA
UAGCUCUCUGAAAGGUCUUACCCUCAGCUGGAUCGGGGAUGGGAACAAUAUCC
UGCACUCCAUCAUGAUGAGCGCAGCGAAAUUCGGAUUGCACCUCAGGCAGCU
ACUCCAAAGGGUUAUGAGCCGGAUGCUAGUGUAACCAAGUUGGCAGAGCAGU
AUGCCAAAGAGAAUGGUACCAAGCUGUUGCUGACAAAUGAUCCAUUGGAAGC
AGCGCAUGGAGGCAAUGUAUUAAUUACAGACACUUGGAUAAGCAUGGGACAA
GAAGAGGAGAAAGAAAAGCGGCUC CAGGCUUCCAAGGUUACCAGGUUACAA
UGAAGACUGCUAAAGUUGCUGCCUCUGACUGGACAUUUUUACACUGCUUGCCC
AGAAAGCCAGAAGAAGUGGAUGAUGAAGUCUUUUUAUUCUCCUCGAUCACUAG
UGUUCCCAGAGGCAGAAAACAGAAAGUGGACAAUCAUGGCUGUCAUGGUGUC
CCUGCUGACAGAUUACUCACCUCAGCUCCAGAAGCCUAAAUUUUGACUCGAGC
UAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAA
UGGAGUCUCUAAGCUACAUAUAACCAACUACACUUAACAAAUGUUGUCCCCC
AAAAGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUU
CUAG

>mARM564 (SEQ ID NO: 27)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUCUUUAAUCUGCGCAUCUUA
CUGAACAACGCCGCAUUCGGAACGGUCACAACUUCAUGGUCCGCAAUUUCCG
CUGUGGCCAGCCGCUUCAAAACAAGGUCCAGCUGAAGGGACGGGAUCUGCUGA
CACUGAAGAACUUCACCGGAGAAGAGAUCAAGUACAUGCUGUGGCUCAGCGCA
GACUUGAAGUCCGGAUCAAGCAGAAGGGAGAAUACUUGCCCCUGCUGCAAG
GAAAGUCGCUGGGAAUGAUUUUUGAGAAGCGGUCAACUCGCACCAGACUCUCC
ACCGAAACUGGUUUCGCACUGCUUGGCGGGCACCCUUGCUUCCUGACGACUCA
GGACAUCCACCUCGGCGUGAACGAAUCGCUAACCGAUACCGCCAGAGUGCUUU
CUUCCAUGGCCGACGCGGUGCUGGCCAGGGUGUACAAGCAGUCCGACCUCGAU
ACCUUGGCAAAGGAGGCUUCCAUUCCCAUCAUCAACGGCCUGAGCGACCUGUA
CCACCCAAUCCAAAUCCUGGCUGACUACCUGACCCUGCAAGAGCACUACAGCA
GCCUGAAGGGUCUGACCCUGUCAUGGAUUGGCGAUGGAAACAUAUUCUGCA
CUCCAUCAUGAUGUCCGCCGCGAAGUUCGGAUUGCAUCUGCAAGCCGCCACUC
CAAAAGGAUACGAACCGGAUGCGUCCGUGACCAAGUUGGCGGAACAGUACGCG
AAGGAGAACGGAACCAAGCUUCUGCUGACUACGACCCCCUCGAGGCUGCGCA
UGGGGGCAACGUGCUGAUUACCGACACCUGGAUCUCCAUGGGGGCAGGAGGAA
GAGAAGAAGAAGAGACUGCAGGCAUUCAGGGGUACCAGGUCACCAUGAAAA
CCGCAAAAGUGGCAGCUUCGGACUGGACUUCUCCUGCAUUGCCUGCCGAGGAAG
CCGGAGGAAGUCGACGACGAAGUGUUCUACUCGCCUCGGUCCUGGUGUUC
CGAGGCCGAAAACCGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCUUGCUGA
CUGACUAUAGCCCGCAGCUGCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUCUUCACAUCUAG

>mARM565 (SEQ ID NO: 28)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUUAACCUACGUUUUUG
CUCAACAAUGCAGCCUUUAGAAACGGACAUAACUUUAUGGUUCGAAACUUUC
GCUGCGGGCAGCCACUGCAGAACAAGGUCCAGCUGAAAGGGAGAGAUUUGCUC
ACGCUGAAGAACUUUACUGGCGAAGAAAUCAAGUAUAUGCUGUGGUUGUCCG
CGGACCUCAGUUUCGGAUUAAGCAGAAAGGGGAGUAUCUGCCACUGCUGCA
AGGAAAGAGCCUCGGCAUGAUCUUCGAGAAGCGGAGCACUCGGACCAGGCUGA
GUACCGAAACUGGCUUCGCAUUGUUGGGUGGACAUCUUGUUUUCUGACAAC
GCAGGACAUAUCUGGGCGUGAACGAGAGUCUGACGGACACAGCUCGCGUUC
UGUCCUCUAUGGCUGAUGCGGUGUUGGCCCGGGUCUAUAAGCAGUCCGAUUU
GGACACCUUGGCUAAGGAAGCUAGCAUACCGAUUAUCAUUGGGCUGUCCGACC
UGUAUCACCCUAUUCAAAUCCUGGCCGACUACCUCACACUGCAAGAACACUAU
AGCUCAUUGAAGGGACUGACCCUGAGCUGGAUAGGGGACGGAAACAACAUC
UACAUAGCAUUAUGAUGUCCGCUGCCAAGUUUGGCAUGCAUCUUCAAGCCGCC
ACGCCAAAGGGUUAUGAGCCCGACGCGUCAGUGACAAAGCUGGCCGAGCAGUA
CGCUAAGGAGAAUGGUACCAAAUACUGCUGACUAAUGAUCCACUGGAGGCU
GCACAUGGCGGCAAUGUACUGAUCACCGACACAUGGAUCUCGAUGGGGCCAGGA
GGAAGAAAAGAAGAAGAGGCUUCAGGCCUCCAAGGCUACCAGGUCACCAUG
AAACAGCUAAGGUUGCAGCAUCUGAUUGGACC UUUCUGCACUGUCUGCCAAG
GAAGCCCGAAGAGGUGGACGAUGAAGUAUUCUAUAGCCCACGGAGUUUGGUG

UCCCCUGAGGCUGAAAAUAGGAAGUGGACAAUUAUGGCCGUA AUGGUGUCCC
 UGUUAACCGACUACUCUCCGCAACUGCAGAAACCUAAGUUUUAGCUCGAGCUA
 GUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUG
 GAGUCUCUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAA
 AAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCU
 AG

>mARM566 (SEQ ID NO: 29)

UCAACACAACAUUAACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUUAAACUUAAGGAUCCUG
 CUGAACAACGCCGCUUUUCGUAACGGUCAUAACUUUAUGGUCCGGAACUUUAG
 AUGUGGCCAGCCGCUUGCAGAACAAGGUUCAGCUGAAGGGGAGGGGAUCUGCUG
 ACCUUGAAGAACUUUACCGGCGAAGAGAUCAAGUACAUGUUGUGGCUGAGCG
 CCGAUCUGAAGUUUAGGAUUAAGCAGAAGGGGGAGUAUUUGCCACUGCUGCA
 AGGAAA AUCCUUGGGGAUGAUCUUCGAGAAGCGCUCCACUAGAACCCGGCUAA
 GCACAGAAACCGGCUUCGCACUUCUGGGUGGACAUCCUGUUUUCUGACGACG
 CAGGAUUAACACCUGGGCGUGAAUGAGAGUCUGACGGACACAGCUAGGGUGU
 UGAGCAGCAUGGCCGAUGCAGUACUGGCCCGCGUUUAUAAGCAGAGCGACUUG
 GACACACUGGCCAAGGAAGCGUCAAUUCCGAUUAUCA AUGGGCUGUCAGACCU
 GUAUCAUCCCAUUCAAAUCUUGGCUGACUAUCUGACCCUGCAAGAACAUUACA
 GCUCCCGUAAGGGCCUCACGUUGUCCUGGAUUGGCGACGGAAACAACAUUCUG
 CAUUCGAUCAUGAUGAGCGCUGCUAAGUUUGGCAUGCACCUCCAAGCCGCUAC
 ACCUAAGGGGAUAUGAGCCUGAUGCCAGCGUAACCAAGCUGGCCGAACAGUACG
 CGAAGGAGAAUGGCACGAAACUGCUGUUGACAAAUGACCCACUGGAGGCAGC
 UCACGGUGGCAACGUGCUGAUCACCGACACGUGGAUAUCUAUGGGACAGGAA
 GAAGAGAAGAAGAAGCGGCUGCAGGCAUUCCAAGGGUAUCAGGUCACCAUGA
 AAACGGCCAAGGUUGCUGCAUCCGACUGGACA UUUUCUGCAUUGCUUGCCCCGC
 AAACCAGAAGAAGUAGACGACGAAGUCUUUUAU UCCCCACGGUCGCUGGUGU
 UCCCCGAGGCGGAGAAUCGAAAGUGGACGAUUAUGGCCGUGAUGGUGUCCCU
 GCUGACUGAUUACUCUCCCCAACUGCAAAAGCCUAAGUUUUAGCUCGAGCUAG
 UGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGG
 AGUCUCUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAA
 AUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUA
 G

>mARM567 (SEQ ID NO: 30)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAACCUGAGGAUCCUC
CUGAACAACGCCGCCUUUCGCAAUGGUCACAACUUUAUGGUCCGGAACUUCAG
AUGCGGCCAGCCGCUGCAGAACAAGGUCCAGCUGAAGGGACGGGAUCUGCUGA
CUCUGAAGAACUUCACCGGAGAAGAGAUCAAGUACAUGCUGUGGCUGUCGGCC
GACCUGAAGUUCAGGAUCAAGCAGAAGGGAGAAUACCUCGCCUGCUGCAAGG
AAAGUCCUGGGCAUGAUUUUCGAGAAGCGCUCGACCAGAACUCGGUUGUCCA
CCGAAACCGGGUUUGCGCUGCUGGGCGGACAUCUUGCUUCCUGACGACUCAG
GAUAUUCACCUGGGAGUGAACGAGUCGCUGACCGACACCGCCAGAGUGCUGAG
CUCGAUGGCCGACGCCGUGUUGGCACGCGUGUACAAGCAGUCCGAUCUGGAUA
CCCUGGCCAAAGAAGCUUCCAUCCCGAUCAUUAACGGGCUGAGCGACCUCUAC
CACCCCAUUCAAAUCCUGGCCGACUACCUGACUCUGCAAGAACACUACAGCUC

GCUGAAGGGGUUGACUCUGUCCUGGAUCGGCGACGGAAACAACAUCCUGCACU
CCAUCAUGAUGUCGGCCGCAAAGUUCGGCAUGCAUUUGCAAGCCGCCACCCCA
AAGGGCUACGAACCAGACGCGAGCGUCACCAAGCUGGCCGAACAGUACGCGAA
GGAAAUGGUACUAAGCUGCUGCUGACCAACGACCCAUUGGAAGCUGCCCAUG
GUGGAAACGUGCUGAUCACCGACACCUUGGAUCUCGAUGGGCCAGGAAGAGGA
GAAGAAGAAGCGGCUGCAGGCGUCCAGGGGUAUCAGGUCACCAUGAAAACA
GCCAAAGUGGCAGCGUCAGACUGGACCUUCCUCCACUGUCUGCCUCGCAAGCC
AGAGGAGGUGGACGACGAGGUGUUCUACUCCCCUCGGUCCCUUGUGUCCCCUG
AGGCUGAGAACCGGAAGUGGACCAUUAUGGCCGUGAUGGUGUCACUCCUGAC
UGAUUACUCCCCGCAACUGCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
UAAGCUACAUAUAACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAUGUAG
CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM568 (SEQ ID NO: 31)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUUAACCUGAGGAUCCUA
UGAACAUAUGCUGCUUUUCGUAAUUGGCCAUAACUUUAUGGUUCGGAACUUUA
GAUGCGGGCAGCCACUGCAGAACAAGGUCCAGUUGAAAGGCCGCGAUCUGUUG
ACAUUGAAGAACUUUACCGGCGAAGAGAUUAAGUAUAUGCUGUGGCUGUCUG
CUGACCUCAAGUUUCGAAUCAAGCAGAAGGGCGAAUAUCUCCCCCUGCUGCAA
GGAAAGUCUCUCGGCAUGAUCUUUGAGAAGCGGAGUACCCGAACACGGCUGA
GCACCGAAACGGGCUUCGCACUGCUGGGGGGGCCAUCCUGUUUUCUGACAACG
CAGGACAUCCACUUGGGGGGUUAACGAAUCAUUGACUGAUACCGCCCCGCGUACU
GUCAUCCAUGGCCGACGCUGUGCUGGCUAGGGUGUACAAGCAGUCAGAUCUGG
AUACACUGGCCAAGGAAGCUAGCAUACCAAUCAUCAUUGGACUGAGUGACCUU
UAUCACCCGAUUCAAAUAUAGCCGAUUAUCUGACCCUGCAAGAGCAUUAUCUC
CUCGCUGAAAGGCCUCACGCUGUCCUGGAUCGGCGACGGCAACAACAUUCUGC
AUAGUAUUAUGAUGUCUGCUGCCAAAUUCGGCAUGCAUCUGCAAGCUGCUAC
GCCGAAGGGUUAUGAACCCGACGCGUCAGUUAACGAAGCUCGCUGAGCAGUAUG
CAAAGGAGAAUGGCACAAAGCUGUUGCUUACCAACGAUCCCCUGGAAGCUGCU
CAUGGCGGCAAUGUGCUGAUUACUGACACCUGGAUUUCAUUGGGCCAGGAGG
AGGAGAAGAAGAAGAGGUUACAGGCUUUUCAAGGUUACCAAGUCACGAUGAA
AACCGCUAAGGUCGCAGCCAGCGACUGGACAUUCCUGCACUGUCUGCCAAGAA
AGCCGGAAGAAGUGGACGACGAGGUGUUCUAUUCCCCGCGGUCUUUGGUGUU
UCCGGAGGCCGAAAACAGGAAAUGGACCAUUAUGGCCGUGAUGGUUAUCGUUG
CUGACGGACUACAGCCCUCAGUUGCAAAAGCCCAAGUUCUAGCUCGAGCUAGU
GACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGA
GUCUCUAAGCUACAUAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAA
UGUAGCCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM569 (SEQ ID NO: 32)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUCUUUAACCUCCGCAUCCUC
CUCAACAACGCCGCCUUCGGAAUGGGCAUAACUUAUGGUCCGGAACUUCAG
AUGCGGCCAGCCCCUGCAAAACAAGGUCCAGUUGAAGGGACGGGACCUCCUUA
CGCUGAAGAACUUUACCGGAGAAGAGAUUAAGUACAUGCUGUGGUUGUCCGC
UGACCUCAAGUUCGCAUUAAGCAGAAGGGAGAAUAUCUGCCGCGUGCUGCAAG

GAAAGAGCCUGGGCAUGAUCUUCGAAAAGCGCUCCACUAGAACCCGGCUGUCC
 ACUGAGACUGGAUUCGCCUUGCUCGGUGGACACCCGUGCUUCCUGACGACCCA
 GGACAUCCACCUGGGAGUGAACGAGUCACUUAACGGAUACCGCGAGGGUGCUGU
 CCUCAAUUGGCCGACGCAGUGCUCGCGCGCGUGUACAAGCAGUCAGAUCUGGAU
 ACCCUGGGCCAAGGAAGCCAGCAUUCCCAUCAUCAACGGACUGAGCGACCUUUA
 CCACCCAAUCCAGAUCUCGCCGACUACUUAACCCUGCAAGAGCACUACAGCU
 CCCUGAAGGGGACUGACUCUGUCCUGGAUCGGGGAUGGAAACAACAUCUGCAC
 UCCAUCAUGAUGUCUGCCGCUAAGUUUGGGAUGCAUCUGCAAGCCGCAACCCC
 UAAGGGAUACGAGCCCGACGCCUCGGUGACCAAACUUGCGGAACAGUACGCCA
 AGGAAAACGGUACCAAGCUGCUGCUGACCAACGACCCUCUGGAAGCGGCCAC
 GGAGGAAAUGUGCUGAUUACCGACACCUGGAUUUCGAUGGGCCAGGAGGAGG
 AGAAGAAGAAGAGACUGCAGGCGUUCAGGGAUUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGCCGCCAGCGACUGGACCUUCCUGCACUGUCUCCUCGGAAAC
 CGGAAGAAGUGGAUGACGAGGUGUUCUACUCCCCGCGCUCGCUGGUGUCCCCG
 GAGGCUGAAAACAGGAAGUGGACAAUCAUGGCCGUGAUGGUGUCCCUUGUA
 CCGACUACUCCCCACAACUGCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUACCAACUUAACACUACAAAAUGUUGUCCCCCAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUCUAG

>mARM570 (SEQ ID NO: 33)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACU
 CUAUUGCAGCAAUUUAAAUCAUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUGCGCAUCCUC
 CUGAACAAACGCCGCCUUCGCAAUGGACACAACUUUAUGGUCCGCAACUCCG
 CUGUGGGCAGCCGCGUCGAGAACAAGGUCCAGCUCAAGGGGAGAGAUCCUGA
 CCCUGAAGAACUUCACUGGAGAGGAGAUCAAGUACAUGCUGUGGCUGUCCGCC
 GACCUGAAAUUUCGGAUUAAAGCAGAAGGGCGAAUACCUCCACUGCUGCAAGG
 AAAGUCUUUGGGCAUGAUCUUCGAAAAGAGAAGCACCCGGACCCGGUUGAGC
 ACCGAAACUGGGUUCGCGCUCCUCGGUGGACACCCGUGCUUCCUGACCACCCA
 AGAUAUUCAUCUGGGUGUCAACGAAAGCCUGACCGACACCGCCAGGGUGCUGU
 CAUCCAUGGCUGACGCAGUGCUCGCCCCGGGUGUACAAGCAGUCAGACCUGGAC
 ACCCUCGCCAAGGAAGCUUCGAUCCCUAUAUCAACGGACUUUCCGACCUGUA
 CCACCCCAUCCAAAUUCUGGGCCGACUACCUGACUCUGCAAGAACACUAUAGCU
 CGCUGAAAGGACUUACUCUGUCCUGGAUCGGGGACGGCAACAACAUUCUCCAU
 UCCAUCAUGAUGUCCGCGUGCCAAGUUCGGAUUGCACCUUCAAGCAGCGACUCC
 CAAGGGAUACGAACCUGAUGCCUCCGUGACUAAGCUGGCAGAGCAGUACGCCA
 AGGAGAACGGUACAAGCUGCUGCUCACGAACGACCCCCUGGAGGCGGCCAC
 GGCGGAAACGUGCUGAUUACCGAUACCUGGAUCUCAUUGGGCCAGGAAGAGG
 AGAAGAAGAAGCGGCUCAGGCGUUUCAAGGCUACCAGGUCACCAUGAAAACC
 GCGAAGGUCGCCGCCUCCGACUGGACUUUCUUGCACUGCCUGCCGCGGAAGCC
 CGAGGAAGUGGAUGACGAAGUGUUCUACUCGCCGAGAUCGUUGGUGUCCCCU
 GAGGCCGAAAACAGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCCUGCUGAC
 UGAUUACAGCCCACAGCUGCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACUG
 ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
 UAAGCUACAUAUACCAACUUAACACUACAAAAUGUUGUCCCCCAAAUGUAG
 CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUCUAG

5 >mARM571 (SEQ ID NO: 34)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAAUCCCUACCCGAUACCGCCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGGCGUCAAUUCCUAUUAUCAACGGCCUUAUGUGACCUCUA
CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCA
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
GAAGGGUACGAGCCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCCAC
GGGGGCAACGUGCUUAUUAUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUUAUCAGGUCACCAUGAAAAC
CGCCAAGGUCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCUCC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM572 (SEQ ID NO: 35)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAACCUGAGAAUCCUC
CUGAACAAACGCCGCCUUCGCAAUGGUCAUAACUUCAUGGUCCGCAACUUUCG
CUGCGGACAGCCUCUCCAAAACAAGGUCCAGCUCAAGGGGCGCGACCUCUCA
CACUGAAGAACUUCACUGGAGAAGAAAUCAAGUACAUGCUGUGGCUGAGCGC
CGAUCUGAAGUUCGGAUCAAGCAGAAGGGAGAGUACCUUCCUCUGCUGCAAG
GGAAGUCCUUGGGAAUGAUUUUCGAGAAGCGGUCCACCCGGACCAGGCUGAGC
ACUGAAACUGGCUUCGCCUGCUGGGAGGCCACCCUUGUUUCCUGACCACUCA
GGACAUCCACCUGGGCGUGAACGAGUCCUGACC GAUACUGCCAGAGUGCUGU
CCUCCAUGGCCGACGCCGUGCUCGCCCGGGUGUACAAGCAGUCAGACCUCGAU
ACGCUGGCCAAGGAAGCCUCCAUUC CCAUUAUCA AUGGUCUGUCGGACCUCUA
CCAUCCA AUCCAAA UCCUCGCCGACUACCUGACUCUGCAAGAACACUACAGCU
CACUCAAGGGCCUCACCCUCUCCUGGAUCGGCGACGGAAACAACAUCCUUCAC
UCGAUUAUGAUGUCGGCCGCGAAGUUCGGGAUGCACCUC CAAGCUGCCACUCC
AAAAGGCUACGAGCCGGAUGCCUCAGUGACUAAGUUGGCGGAACAGUAUGCG
AAGGAGAACGGUACCAAGCUCCUGCUGACUAACGACCCGCUGGAGGCCGCCCA
CGGGGGAAACGUGCUCAUACCGAUACUUGGAUUUCCAUGGGACAGGAGGAA
GAGAAGAAGAAGCGGUUGCAGGCAUUUCAGGGCUACCAGGUCACCAUGAAAA
CUGCCAAAGUCGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAG
CCUGAAGAAGUGGACGACGAGGUGUUCUACUCUCCCCGGUCCCU CGUGUUC C
UGAGGCCGAAAACAGGAAGUGGACCAUCAUGGCUGUGAUGGUGUCCCUCCUG

ACCGACUACAGCCCUCAGCUCCAAAAACCCAAGUUUUAGCUCGAGCUAGUGAC
UGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUC
UCUAAGCUACAUA AUACCAACUACACUUA CAAAAUGUUGUCCCCCAAA AUGU
AGCCAUUCGU AUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM573 (SEQ ID NO: 36)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAACCUGAGAAUCCUC
UUGAACAAUGCUGCUUUUCGGAUUGGCCACAACUUUAUGGUUCGGAACUUC
GUUGCGGCCAGCCUUUACAAAACAAGGUCCAGCUGAAGGGCCGGGAUUUGCUC
ACACUGAAGAACUUUACUGGGGGAGGAGAUUAAGUAUAUGCUGUGGCUGUCCG
CUGACCUGAAGUUUAGGAUCAAGCAGAAGGGCGAAUAUCUGCCGCUGCUGCA
AGGGAAAAGUCUGGGCAUGAUUUUUGAAAAGCGCUCUACCCGGACCAGACUG
UCUACGGAAACAGGCUUUGCCCUGCUGGGCGGCCACCCUGUUUUCUGACAAC
GCAGGACAUCCAUCUGGGCGUGAACGAAUCACUGACCGAUACUGCUCGGGUAC
UCAGUUCUAUGGCUGACGCAGUGCUGGCUAGGGUGUACAAGCAGAGCGACUU
GGACACACUGGCUAAGGAGGCCAGCAUCCCCAUUAUCAUUGGCCUGUCUGAUU
UGUACCAUCCCAUUCAAAUCCUGGCUGAUUAUCUGACACUACAAGAGCAUUA
UCAAGUCUGAAGGGUUUGACUCUCUCCUGGAUCGGCGACGGCAACAACAUUUU
ACAUCCAUAUGAUGAGUGCUGCUAAGUUUGGCAUGCAUUUGCAAGCUGCU
ACCCCAAAGGGCUAUGAACCUGACGCUAGCGUAACCAAGUUGGCCGAACAGUA
UGCUAAGAGAAUGGCACCAAGCUGCUCCUGACGAAUGACCCCGUGGAAGCUG
CUAUGGCGGAAACGUACUUAUAAACUGAUACAUGGAUUAGCAUGGGGCCAGGA
AGAGGAGAAGAAGAAGAGACUGCAGGCCUCCAAGGCUAUCAGGUCACCAUG
AAAACUGCCAAGGUUGCAGCUAGCGACUGGACCUUCCUGCACUGUUUGCCGAG
GAAACCCGAGGAGGUGGACGAUGAAGUCUUUUAUUCUCCCCGCUCUUGGUGU
UUCCCGAGGCUGAAAAUCGAAAGUGGACGAUAAUGGCAGUGAUGGUGUCCCU
ACUGACCGACUAUUCUCCACAACUGCAGAAGCCUAAAUUCUAGCUCGAGCUAG
UGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGG
AGUCUCUAAGCUACAUAUACCAACUUAACACUACAAAAUGUUGUCCCCCAA
AUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUA
G

>mARM574 (SEQ ID NO: 37)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUGAGGAUCCUG
CUGAACAAACGCUGCUUUUCGCAACGGUCAUAACUUUAUGGUUCGCAAUUUUCG
UUGUGGCCAGCCGCUGCAGAACAAAGGUUCAGCUGAAGGGCAGAGAUCUGCUG
ACUCUGAAGAACUUCACUGGGGAAGAAAUCAAGUAUAUGUUUUGGCUGUCCG
CGGAUCUGAAAUUUCGAAUCAAGCAGAAGGGCGAAUAUCUCCCCUGCUGCAA
GGGAAAUCCUUGGGCAUGAUUUUUGAGAAGAGGAGCACUAGGACUAGAUUGU
CAACAGAAACAGGCUUUGCUUUGUUGGGCGGACAUCCUGCUUUCUGACGACA
CAGGAUAUCCACCUCGGCGUAAACGAGUCCCUACCGACACUGCUAGGGUACU
GAGCAGCAUGGCCGACGCUGUGCUAGCCCGGGUUUACAAGCAGUCAGACCUGG
ACACCCUUGCCAAGGAAGCUUCUAUUCCAAUUUAUCAACGGCCUGAGUGACCUG
UAUCACCCUAUUCAAAUAUCUGCCGACUAUUUGACGCUUCAAGAACAUAACAG
CAGCCUCAAGGGCUUAACCUUGAGUUGGAUAGGCGACGGCAACAUAUCCUGC
AUUCCAUAUGAUGUCUGCCGCUAAGUUUGGCAUGCAUCUACAAGCCGCAACA

CCCAAGGGCUAUGAACCCGACGCUAGCGUGACCAAGCUGGGCCGAGCAGUAUGC
 UAAGGAAAAAUGGCACAAAGCUCCUUCUUACCAACGAUCCCCUGGAGGCUGCUC
 ACGGCGGCAACGUGCUGAUUACCGAUACAUGGAUUAGCAUGGGCCAGGAGGA
 GGAGAAAAAGAAGCGGCUCAGGCUUUUAAGGCUAUCAGGUCACCAUGAAA
 ACUGCAAAGGUCGCGUCCUCCGACUGGACUUUCCUGCAUUGUCUACCCCGCAA
 GCCUGAGGAAGUGGACGAUGAGGUGUUCUACUCCCCACGGAGUCUGGUGUUC
 CGGAAGCAGAGAAUCGGAAGUGGACCAUCAUGGCUGUCAUGGUGUCGUCUU
 GACUGACUAUUCUCCCCAACUGCAAAAACCCAAGUUUUAGCUCGAGCUAGUGA
 CUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGU
 CUCUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUG
 UAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM575 (SEQ ID NO: 38)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGUUAUUCACCUUCGUAUCCUG
 CUAACAAUGCUGCUUUUCGAAUGGCCAUAAACUUUAUGGUUCGCAACUUUA
 GAUGCGGCCAGCCGUGCAGAACAAAGGUUCAGCUGAAGGGCCGGGACUUGCUG
 ACGCUGAAAAACUUUACCGGGGAAGAGAUUAAGUAUAUGCUGUGGCUAAGCG
 CUGAUCUGAAGUUUAGGAUCAAGCAGAAGGGCGAAUAUCUGCCACUGCUGCA
 AGGGAAGAGUCUUGGCAUGAUUUUUGAAAAGCGGUCUACCAGAACCCGGCUG
 UCGACCGAGACAGGUUUUGCUCUGCUGGGGGGCCAUCCUGUUUUCUGACAAC
 UCAGGACAUUCACCUGGGCGUGAAUGAGUCCUGACCGAUACUGCUAGGGUGU
 UGAGUAGCAUGGCCGACGCGUACUCGCUUGAGUGUAUAAGCAGUCUGAUCU
 GGACACUCUGGCUAAGGAAGCUUCCAUUCCUAUUAUCAACGGCUUGAGCGACC
 UGUACCACCCCAUUCAAAUCCUCGCUGAUUACUUGACUUUGCAAGAACAUAAC
 AGCAGCUUGAAGGGCUUAACACUGAGCUGGAUAGGCGACGGAAACAACAUUCU
 UGCAUUCCAUAAUGAUGUCCGCCGCUAAGUUCGGGAUGCACCUCCAAGCAGCC
 ACACCCAAGGGCUAUGAACCGGAUGCUUCCGUGACAAAACUGGCUGAGCAGUA
 UGCUAAGGAGAAUGGCACGAAACUGCUGCUCACCAACGACCCAUUGGAAGCUG
 CACAUGGUGGCAACGUACUGAUCACUGACACUUGGAUCUCAAUGGGGCCAGGAG
 GAAGAGAAGAAGAAAAGGCUGCAGGCAUUUCAGGGAUACCAAGUCACUAUGA
 AAACUGCCAAGGUCGCGUCCUCCGACUGGACAUUCCUGCAUUGUCUGCCACGG
 AAGCCUGAGGAAGUCGAUGACGAAGUGUUCUAUAGCCCACGAAGCUUGGUGU
 UUCCCGAGGCUGAGAAUAGGAAGUGGACCAUUAUGGCUGUUAUGGUGUCCCU
 GCUCACCGACUAUUCCCCUCAACUGCAAAAACCCAAGUUUUAGCUCGAGCUAG
 UGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGG
 AGUCUCUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCAAA
 AUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUA
 G

>mARM708 (SEQ ID NO: 39)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCAUUCUUUUUAUCUCCGCAUCCUCCUU
 AACAAACGCCGCGUUUAGAAACGGCCACAACUUCUUGGUCCGGAACUUCAGAUG
 UGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGACCC
 UGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCGGAC
 UUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUCUUAAGGAA
 AGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCACACAGGACCCGCCUUCUACU

GAAACUGGGUUCGCGCUGCUCGGUGGCCACCCCUGCUUCCUGACGACCCAGGA
CAUCCACCUCGGAGUGAACGAAUCCCUCACCGAUACCGCCCCGGGUGUUAUCGA
GCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGACACU
CUGGCCAAGGAGGCGUCAAUUCCCAUCAUAACGGCCUGAGCGACCUGUACCA
CCCAAUCCAAAUCCUGGCUGACUACCUGACCCUGCAAGAGCACUACAGCAGCC
UGAAGGGUCUGACCCUGUCAUGGAUUGGCGAUGGAAACAAUAUUCUGCACUC
CAUCAUGAUGUCCGCCGCGAAGUUCGGAAUGCAUCUGCAAGCCGCCACGCCAA
AAGGAUACGAACCGGAUGCGUCCGUGACGAAGUUGGCGGAACAGUACGCGAA
GGAGAACGGAACCAAGCUUCUGCUGACUAAACGACCCCCUCGAGGCUGCGCAUG
GGGGCAACGUGCUGAUUACCGACACCUGGAUCUCCAUGGGGGCAGGAGGAAGA
GAAGAAGAAGAGACUGCAGGCAUUCAGGGGUACCAGGUCACCAUGAAAACC
GCAAAGUGGCAGCUUCGGACUGGACUUUCCUGCAUUGCCUGCCGAGGAAGCC
GGAGGAAGUCGACGACGAAGUGUUCUACUCGCCUCGGUCCUGGUGUUCCCCG
AGGCCGAAAACCGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCUUGCUGACU
GACUAUAGCCCGCAGCUGCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACUGA
CUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCU
AAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCCCAAAUGUAGC
CAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM709 (SEQ ID NO: 40)

UCAACACAACAUUAACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUAUUCUUUUAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCAUGCUUUUCAACCUGAGAAUCCUCUUG
AACAAUGCUGCUUUUCGGAUUGGCCACAACUUUAUGGUUCGGAACUUCGCU
GCGGCCAGCCUUUACAAAACAAGGUCCAGCUGAAGGGCCGGGAUUUGCUCACA
CUGAAGAACUUUACUGGAGAAGAGAUCAAGUACAUGCUGUGGCUGUCGGCCG
ACCUGAAGUUCAGGAUCAAGCAGAAGGGAGAAUACCUUCCGCUGCUUCAAGG
AAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCUA
CUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCCUGCUUCCUGACGACCCAG
GACAUCCACCUCGGAGUGAACGAAUCCCUCACCGAUACCGCCCCGGGUGUUAUC
GAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUCGAUA
CCUUGGCAAAGGAGGCUUCCAUUCCCAUCAUAACGGCCUGAGCGACCUGUAC
CACCCAAUCCAAAUCCUGGCUGACUACCUGACCCUGCAAGAGCACUACAGCAG
CCUGAAGGGUCUGACCCUGUCAUGGAUUGGCGAUGGAAACAAUAUUCUGCAC
UCCAUCAUGAUGUCCGCCGCGAAGUUCGGAAUGCAUCUGCAAGCCGCCACUCC
AAAAGGAUACGAACCGGAUGCGUCCGUGACCAAGUUGGCGGAACAGUACGCG
AAGGAGAACGGAACCAAGCUUCUGCUGACUAAACGACCCCCUCGAGGCUGCGCA
UGGGGGCAACGUGCUGAUUACCGACACCUGGAUCUCCAUGGGGCAGGAGGAA
GAGAAGAAGAAGAGACUGCAGGCAUUCAGGGGUACCAGGUCACCAUGAAAA
CCGCAAAAGUGGCAGCUUCGGACUGGACUUUCCUGCAUUGCCUGCCGAGGAAG
CCGGAGGAAGUCGACGACGAAGUGUUCUACUCGCCUCGGUCCUGGUGUUC
CGAGGCCGAAAACCGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCUUGCUGA
CUGACUAUAGCCCGCAGCUGCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCCCAAAUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM710 (SEQ ID NO: 41)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA

AUUUUCACCAUUUACGAACGAUAGCCAUGCUUUUCAACCUGAGAAUCCUCUUG
AACAAUGCUGCUUUUCGGAUUGGCCACAACUUUAUGGUUCGGAACUUCGCUU
GCGGCCAGCCUUUACAAAACAAGGUCCAGCUGAAGGGCCGGGAUUUUGCUCACA
CUAAAGAACUUUACUGGAGAAGAGAUCAAGUACAUGCUAUGGCUGUCGGCCG
ACCUGAAGUUCGUAUCAAGCAGAAGGGAGAAUACCUUCCGCUGCUUCAAGGA
AAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCUAC
UGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCAGG
ACAUCCACCUCGGAGUGAACGAAUCCCUACCGAUACCGCCCGGGUGUUAUCG
AGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUCGAUAC
CUUGGCAAAGGAGGCUUCCAUUCUCAUCAACGGCCUGAGCGACCUGUACC
ACCCAAUCCAAAUCCUGGCUGACUACCUGACCCUGCAAGAGCACUACAGCAGC
CUGAAGGGUCUGACCCUGUCAUGGAUUGGCGAUGGAAACAAUUAUUCUGCACU
CCAUCAUGAUGUCCGCCGCGAAGUUCGGAAUGCAUCUGCAAGCCGCCACUCCA
AAAGGAUACGAACCGGAUGCAUCCGUGACCAAGUUGGCGGAACAGUACGCGA
AGGAGAACGGAACCAAGCUCCUGCUGACUAAACGACCCGCUCGAGGCUGCGCAU
GGGGGUAACGUGCUGAUUACGGACACCUGGAUCUCCAUGGGGCAGGAGGAAG
AGAAGAAGAAGAGACUGCAGGCAUUCAGGGGUACCAGGUCACCAUGAAAAC
CGCAAAGUGGCAGCUUCGGACUGGACUUUCCUGCAUUGCCUGCCGAGGAAGC
CGGAGGAAGUCGACGACGAAGUGUUCUACUCGCCUCGGUCCCUUGGUGUUCUCC
GAGGCCGAAAACCGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCUUGCUGAC
UGACUAUAGCCCGCAGCUGCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
UAAGCUACAUAUAAUACCAACUUACACUUACAAAUGUUGUCCCCCAAAUGUAG
CCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM711 (SEQ ID NO: 42)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCAUGCUUUUCAACCUGAGAAUCCUCUUG
AACAAUGCUGCUUUUUCGGAAUGGGCCACAACUUUAUGGUUCGGAACUUCGCUU
GCGGCCAGCCUUUACAAAACAAGGUCCAGCUGAAGGGCCGGGAUUUUGCUCACA
CUAAAGAACUUUACUGGAGAAGAGAUCAAGUACAUGCUAUGGCUGUCGGCCG
ACCUGAAGUUCGUAUCAAGCAGAAGGGAGAAUACCUUCCGCUGCUUCAAGGA
AAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCUAC
UGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCAGG
ACAUCCACCUCGGAGUGAACGAUCCCUACCGAUACCGCCCGGGUGUUAUCG
AGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUCGAUAC
CUUGGCAAAGGAGGCUUCCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUACC
ACCCAAUCCAAAUCCUGGCUGACUACCUGACCCUGCAAGAGCACUACAGCAGC
CUGAAGGGUCUGACCCUGUCAUGGAUUGGCGAUGGAAACAAUAUUCUGCACU
CCAUCAUGAUGUCCGCCGCGAAGUUCGGAUUGCAUCUGCAAGCCGCCACUCCA
AAAGGAUACGAACCGGAUGCGUCCGUGACCAAGUUGGCGGAACAGUACGCGA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCCCUCGAGGCUGCGCAU
GGGGGCAACGUGCUGAUUACCGACACCUGGAUCUCCAUGGGGCAGGAGGAAG
AGAAGAAGAAGAGACUGCAGGCAUUCAGGGGUACCAGGUCACCAUGAAAAC
CGCAAAAGUGGCAGCUUCGGACUGGACUUUCCUGCAUUGCCUGCCGAGGAAGC
CGGAGGAAGUCGACGACGAAGUGUUCUACUCGCCUCGGUCCCUGGUGUUCUCC
GAGGCCGAAAACCGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCUUGCUGAC
UGACUAUAGCCCGCAGCUGCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC

UAAGCUACAUAUACCAACUUACACUUACAAAAUGUUGUCCCCCAAAUUGUAG
CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM712 (SEQ ID NO: 43)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUAACCUGCGAAUCCUG
CUGAACAAACGCCGCUUUUCGGAACGGGCACAACUUUAUGGUGAGGAACUUUCG
CUGCGGACAGCCCCUCCAGAAUAAGGUCCAGCUGAAGGGCAGGGACCUGCUGA
CCCUGAAAAAUUUCACAGGGGAGGAAAUCAAGUAUAUGCUGUGGCUGUCAGC
UGAUCUGAAGUUCGGAUCAAGCAGAAGGGCGAAUAUCUGCCUCUGCUCACAGG
GCAAAAGCCUGGGGAUGAUCUUCGAAAAGCGCAGUACUCGGACCAGACUGUCA
ACCGAGACUGGAUUCGCUCUGCUGGGAGGACACCCUUGUUUUCUGACCACUCA
GGACAUUCACCUGGGAGUGAACGAGUCCUGACCGACACUGCUCGCGUCCUGA
GCUCUAUGGCCGACGCUGUGCUGGCUCGAGUCUACAAACAGUCCGACCUGGAU
ACCCUGGCCAAGGAAGCUUCUAUCCCAAUUAUUAACGGCCUGUCAGACCUGUA
UCACCCCAUCCAGAUUCUGGCCGAUUAACUGACCCUCCAGGAGCACUAUUCUA
GUCUGAAAGGGCUGACACUGAGUUGGAUUGGGGACGGAAACAAUAUCCUGCA
CUCUAUUUAUGAUGUCAGCCGCCAAGUUUGGAUUGCACCUCAGGCUGCAACCC
CAAAAGGCUACGAACCCGAUGCCUCAGUGACAAAGCUGGCUGAACAGUACGCC
AAAGAGAACGGCACUAAGCUGCUGCUGACCAACGACCCUCUGGAGGCCGCUCA
CGGAGGCAACGUGCUGAUCACCGAUACCUGGAUUAGUAUGGGACAGGAGGAA
GAGAAGAAGAAGCGGCUCACAGGCCUUCAGGGCUACCAGGUGACAAUGAAAAC
CGCUAAGGUCGCAGCCAGCGAUUGGACCUUUCUGCACUGCCUGCCCAGAAAGC
CCGAAGAGGUGGACGACGAGGUCUUCUACUCUCCAGAAAGCCUGGUGUUUCCC
GAAGCUGAGAAUAGGAAGUGGACAAUUAUGGCAGUGAUGGUCAGCCUGCUGA
CUGAUUAUUCACCUCAGCUCACAGAAACCAAAGUUCUGAUAACUCGAGCUAGUG
ACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAG
UCUCUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAAU
GUAGCCAUUCGUUAUCUGCUCCUAAUAAAAGAAAGUUUCUUCACAUUCUAG

>mARM713 (SEQ ID NO: 44)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUAACCUGCGCAUCCUG
CUGAACAAACGCCGCCUUCGCAACGGCCACAACUUUAUGGUGCGCAACUCCG
CUGCGGCCAGCCCCUGCAGAACAAAGGUGCAGCUGAAGGGCCGCGACCUGCUGA
CCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAGCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCAGGG
CAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGACCCGCCUGAGCA
CCGAGACAGGCCUGGCCUGCUGGGCGGCCACCCUGCUUCCUGACCACCCAG
GACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCCGCGUGCUGAG
CAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGAGCGACCUGGACA
CCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUAC
CACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACUACAGCAG
CCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGCACA
GCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCACCCCC
AAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACGCCAA
GGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCCACG
GCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGAGGAG

AAGAAGAAGCGCCUGCAGGCCUUCCAGGGCUACCAGGUGACCAUGAAGACCGC
CAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCCCG
AGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUCCCCGAG
GCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACCGA
CUACAGCCCCCAGCUGCAGAAGCCCCAAGUUCUGAUAAACUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
UAAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCCCAAAAUGUAG
CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM714 (SEQ ID NO: 45)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUAACCUGCGAAUCCUG
CUGAACAAACGCCGCUUUUCGGAACGGGCACAACUUUAUGGUGAGGAACUUUCG
CUGCGGACAGCCCCUCCAGAAUAAGGUCCAGCUGAAGGGCAGGGACCUGCUGA
CCCUGAAAAAUUUCACAGGGGAGGAAAUCAAGUAUAUGCUGUGGCUGUCAGC
UGAUCUGAAGUUCGGAUCAAGCAGAAGGGCGAAUAUCUGCCUCUGCUCCAGG
GCAAAAGCCUGGGGAUGAUCUUCGAAAAGCGCAGUACUCGGACCAGACUGUCA
ACCGAGACUGGAUUCGCUCUGCUGGGAGGACACCCUUGUUUUCUGACCACUCA
GGACAUUCACCUGGGAGUGAACGAGUCCUGACCGACACUGCUCGCGUCCUGA
GCUCUAUGGCCGACGCUGUGCUAGCUCGAGUCUACAAACAGUCCGACCUGGAU
ACCCUGGCCAAGGAAGCUUCUAUCCCAAUUAUUAACGGCCUGUCAGACCUGUA
UCACCCCAUCCAGAUUUCUGGCCGAUUACCUGACCCUCCAGGAGCACUAUUCUA
GUCUGAAAGGGCUGACACUGAGUUGGAUUGGGGACGGAAACAAUAUCCUGCA
CUCUAUUAUGAUGUCAGCCGCCAAGUUUGGAAUGCACCUCAGGCUGCAACCC
CAAAAGGCUACGAACCCGAUGCCUCAGUGACAAAGCUGGCUGAACAGUACGCC
AAAGAGAACGGCACUAAGCUGCUGCUGACCAACGACCCUCUGGAGGCCGCUCA
CGGAGGCAACGUGCUGAUCACCGAUACCUGGAUUAGUAUGGGACAGGAGGAA
GAGAAGAAGAAGCGGCUCCAGGCCUCCAGGGCUACCAGGUGACAAUGAAAAC
CGCUAAGGUCGCAGCCAGCGAUUGGACCUUUCUGCACUGCCUGCCCAGAAAGC
CCGAAGAGGUGGACGACGAGGUCUUCUACUCUCCAGAAAGCCUGGUGUUUCCC
GAAGCUGAGAAUAGGAAGUGGACAAUUAUGGCAGUGAUGGUCAGCCUGCUGA
CUGAUUAUUCACCUCAGCUCCAGAAACCAAAGUUCUGAUAAACUCGAGCUAGUG
ACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAG
UCUCUAAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCCCAAAU
GUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM715 (SEQ ID NO: 46)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUCAACCUGCGAAUCCUG
CUGAACAAUGCCGCUUUUCGGAACGGGCACAAUUUCAUGGUGAGGAACUUUC
GCUGCGGACAGCCCCUCCAGAACAAGGUCCAGCUGAAGGGCAGGGACCUGCUG
ACCCUGAAAAAUUUCACAGGGGAGGAAAUCAAGUACAUGCUGUGGCUGUCAG
CCGAUCUGAAGUUCGGAUCAAGCAGAAGGGCGAAUAUCUGCCUCUGCUCCAG
GGCAAAAGCCUGGGGAUGAUCUUCGAAAAGCGCAGUACUCGGACCAGACUGUC
AACAGAGACUGGAUUCGCACUGCUGGGAGGACACCCAUGUUUUCUGACCACAC
AGGACAUAUCUGGGAGUGAACGAGUCCCUGACCGACACAGCACGCGUCCUG
AGCUCCAUGGCUGAUGCAGUGCUGGCUCGAGUCUACAAACAGUCUGACCUGGA
UACCCUGGCCAAGGAAGCUUCUAUCCCAAUCAUUAUUGGCCUGAGUGACCUGU

AUCACCCCAUCCAGAUUCUGGCCGAUUACCUGACCCUCCAGGAGCAUUAUUCU
AGUCUGAAAGGGCUGACACUGAGCUGGAUUGGGGACGGAAACAAUAUCCUGC
ACUCCAUAUGAUGAGCGCCGCCAAGUUUGGAAUGCACCUCAGGCUGCAACC
CCAAAAGGCUACGAACCCGAUGCCUCCGUGACAAAGCUGGCAGAACAGUAUGC
CAAAGAGAACGGCACUAAGCUGCUGCUGACCAUUGACCCUCUGGAGGCCGCUC
ACGGAGGCAACGUGCUGAUCACUGAUACCUGGAUUAGUAUGGGACAGGAGGA
AGAGAAGAAGAAGCGGCUCCAGGCCUUCAGGGCUACCAGGUGACAAUGAAA
ACUGCUAAGGUCGCAGCCAGCGACUGGACCUUUCUGCAUUGCCUGCCCAGAAA
GCCUGAAGAGGUGGACGAUGAGGUCUUCUACUCACCCAGAAGCCUGGUGUUUC
CUGAAGCUGAGAAUAGGAAGUGGACAAUCAUGGCAGUGAUGGUCAGCCUGCU
GACUGAUUAUUCUCCUCCAGCUCCAGAAACCAAAGUUCUGAUAAUCUGAGCUAG
UGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGG
AGUCUCUAAGCUACAUAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAA
AUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUA
G

>mARM716 (SEQ ID NO: 47)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAACCUUCGCAUUCUC
CUCAACAACGCCGCGUUUAGAAACGGACACAACUUCAUGGUCCGCAACUUCCG
CUGCGGACAGCCGCGUGCAGAACAAGGUCCAGCUCAAGGGUCGCGGAUCUCCUGA
CGCUGAAGAACUUUACCGGCGAAGAGAUUAAGUACAUGCUGUGGCUGUCCGCC
GACCUUAAGUUCGGAUCAAGCAGAAGGGCGAAUACCUUCCCCUGCUGCAAGG
AAAGUCCCUGGGCAUGAUCUUCGAGAAGCGCAGUACCAGAACCAGACUCUCCA
CUGAAACCGGGUUCGCGCUGCUUGGCGGGCCACCCGUGUUUCCUCACUACGCAA
GACAUCCAUCUUGGCGUGAACGAGUCCCUUACCGACACCGCCAGGGUGCUGUC
AAGCAUGGCCGACGCCGUCCUUGCGCGCGUGUACAAGCAGUCAGACCUUGAUA
CUCUGGCCAAGGAAGCCUCCAUCCCUAUUAUCAACGGCCUAUCCGACCUUUAC
CACCCGAUCCAGAUCUCGCGUGACUACCUGACCCUGCAAGAACACUACAGCAG
CCUCAAGGGACUGACUCUGUCCUGGAUCGGCGACGGGAACAACAUCUGCACU
CAAUCAUGAUGAGCGCAGCCAAGUUCGGCAUGCAUCUCCAAGCCGCUACACCC
AAGGGUUAUGAACCGGACGCCUCUGUGACCAAGUUGGCAGAACAGUACGCCAA
GGAGAACGGUACUAAGCUCCUUUUAACCAACGACCCCCUCGAAGCAGCCCAUG
GCGGGAAUGUGCUCAUUACCGAUACCUGGAUUUUCGAUGGGGCCAGGAGGAGGA
GAAGAAGAAGCGGCUGCAGGCGUCCAGGGCUACCAGGUCACCAUGAAAACUG
CCAAAGUGGCCCGCCUCGGAUUGGACCUUUCUCCACUGCCUGCCUCGGAAGCCU
GAGGAGGUGGACGACGAAGUGUUCUACUCCCCACGGUCCCUCGUGUCCCCGA
GGCCGAAAAUAGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCCUCUUGACCG
AUUACAGCCCGCAGCUUCAGAAGCCUAAAUUCUAGCUCGAGCUAGUGACUGAC
UAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCUA
AGCUACAUAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAUGUAGCC
AUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM717 (SEQ ID NO: 48)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUUCGCAUCCUG
UUGAACAACGCCGCCUUCGCAUUGGUCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGACAGCCUCUCCAAAACAAGGUCCAGCUGAAGGGAAGGGACCUCUUA

CCCUCAAAAACUUUACUGGAGAGGAGAUCAAGUACAUGCUGUGGCCUUAGCGCC
 GACCUUAAGUUCCGGAUCAAGCAGAAGGGAGAGUACCUCCCGCUGCUGCAAGG
 AAAGAGUCUUGGAAUGAUCUUCGAGAAGCGGUCCACCAGAACUCGCCUCUCCA
 CUGAAACCGGAUUCGCACUCCUGGGUGGACACCCGUGCUUUCUGACCACCCAA
 GACAUCCACCUCGGAGUGAACGAGAGCCUCACGGACACCCGCGAGAGUGCUGUC
 AUCCAUGGCCGACGCCGUGCUUGCACGGGUCUACAAGCAGUCCGAUCUGGACA
 CUCUUGCCAAGGAAGCCUCCAUAUCCUAUCAUUAACGGUCUGUCGGAUCUGUAC
 CACCCGAUUCAGAUCUUGCGGACUACCUCACACUUCAAGAACACUAUUCAAG
 CCUAAAGGGUCUGACCCUGUCCUGGAUCGGAGAUGGAAACAACAUUCUCCAUAU
 CCAUCAUGAUGAGCGCUGCCAAGUUCGGAAUGCAUCUCCAAGCAGCGACUCCU
 AAGGGUUAACGAGCCGGACGCCUCAGUGACUAAGCUGGCCGAGCAGUACGCCAA
 GGAGAACGGUACCAAACUGUUGCUUACUAACGACCCGCUUGAAGCGGCCCAUG
 GAGGAAACGUGCUGAUUACCGACACCUGGAUUUUCGAUGGGACAGGAAGAGGA
 GAAGAAGAAGCGGCUCCAGGCGUCCAGGGAUACCAGGUCACCAUGAAAACGG
 CCAAAGUGGCCGCUAGCGAUUGGACCUUUCUGCACUGCCUCCCGCGCAAGCCU
 GAAGAAGUGGACGACGAAGUGUUCUACUCCCCUCGCUCUCUUGUGUUCCCGGA
 AGCCGAAAACAGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCCUCCUGACCG
 AUUACAGCCCGCAGCUGCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACUGAC
 UAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCUA
 AGCUACAUAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAAUGUAGCC
 AUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM718 (SEQ ID NO: 49)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAAUCUCCGCAUCCUC
 CUCAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
 AUGUGGCCAGCCGCUUCAGAACAAAGGUCCAGCUCAAGGGCCGGGAUCUUCUGA
 CCCUGAAGAACUUUACUGGCGAAGAAAUCAAGUACAUGCUCUGGCUCUCCGCC
 GACUUGAAGUUCCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCUGCUGCAAG
 GAAAGUCGCUCGGCAUGAUCUUUGAGAAGCGCUCAACCCGCACCAGGCUGUCC
 ACUGAAACCGGGUUCGCGCUGCUUGGUGGCCACCCUGCUUCCUGACCAACCA
 AGACAUUCACCUCGGAGUGAACGAAUCGCUCACUGAUACUGCCCGGGUGCUGU
 CGUCGAUGGCCGAUGCAGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
 ACUCUGGCCAAGGAGGCGUCCAUCCCUAUUAUCAACGGCCUUUCCGACCUCUA
 CCACCCGAUUCAGAUCUUGCCGAUUACCUCACCCUGCAAGAACACUACUCGU
 CACUGAAGGGUCUGACCUUGUCCUGGAUCGGCGACGGCAACAACAUCCUCCA
 UCCAUAUGAUGUCCGCCGCCAAAUUCGGCAUGCAUCUUAAGCCGCAACCCC
 UAAGGGUUAACGAGCCGGACGCUUCCGUGACCAAGCUCGCCGAGCAGUACGCUA
 AGGAGAACGGAACCAAGCUUCUGCUGACUAAACGACCCCCUAGAGGCAGCCAC
 GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGACAGGAAGAAG
 AGAAGAAGAAGCGGUUACAGGCGUCCAGGGCUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGCUGCCUCGGACUGGACCUUCCUGCAUUGCCUGCCUCGCAAGC
 CCGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGUCCCUUGUGUUCCCU
 GAGGCCGAGAAUAGAAAGUGGACCAUUAUGGCCGUGAUGGUGUCCCUUCUCA
 CCGACUACUCGCCGCAACUGCAGAAACCCAAGUUCUAGCUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM719 (SEQ ID NO: 50)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAAUUCUUCGCAUCCUC
CUCAACAACGCCGCCUUCGGAACGGUCACAACUUCAUGGUCCGGAACUUCG
CUGCGGCCAGCCGCUCCAAAACAAAGUGCAGCUUAAGGGCCGCGAUCUCCUGA
CCCUGAAGAACUUCACCGGAGAGGAAAUCAAGUACAUGCUGUGGCUCUCGGCG
GACCUGAAGUUUAGGAUUUAGCAGAAGGGGGAGUAUCUGCCGCGUCUCCAAG
GGAAGUCCCUUGGCAUGAUCUUCGAAAAGAGGUCCACCCGGACUCGGCUCAGC
ACCGAAACAGGUUUUGCACUUCUGGGGGGGCCACCCGUGCUUCCUGACGACCCA
GGACAUCCAUCUGGGUGUCAACGAGAGUUUGACCGACACUGCCAGAGUGCUGU
CAUCCAUGGCGGACGCGGUGCUCGCGAGAGUGUACAAGCAGUCCGAUCUUGAC
ACCCUGGCAAAAGAGGCUUCAAUCCCGAUCAUUAACGGACUCUCGGAUCUGUA
CCACCCUAUCCAAAUCUUGGCCGACUACCUGACCCUGCAAGAACACUACAGCU
CCCUGAAGGGCCUGACUCUUUCCUGGAUUGGCGAUGGAAACAACAUUCUCCA
UCUAUUAUGAUGUCCGCCGCCAAGUUCGGCAUGCACCUUCAAGCCGCCACCCC
GAAGGGCUACGAACCUGACGCCUCCGUGACUAAGCUAGCCGAACAGUACGCUA
AGGAGAACGGCACUAAGCUUCUCCUUACCAACGAUCCGCUUGGAGGCGGCCCAU
GGCGGAAAUGUGCUUAUCACCGACACCUGGAUUAGCAUGGGGCAGGAAGAAG
AGAAGAAGAAACGGCUCCAGGCAUUCAGGGCUACCAGGUCACCAUGAAAACU
GCCAAGGUCGCCGCUAGCGACUGGACCUUCCUCCACUGUCUGCCUCGCAAGCC
UGAAGAAGUGGACGACGAGGUGUUCUACUCCCCGCGCUCCCUCGUGUUUCCUG
AGGCCGAGAACAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCAUUACUUAC
GGACUACAGCCCGCAGCUGCAGAAGCCGAAGUUCUAGCUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
UAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUUGUAG
CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM720 (SEQ ID NO: 51)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUUAACUUGAGAAUCCUU
CUGAACAACGCCGCUUUCGCAACGGUCAUAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCCCUCCAAAACAAAGUGCAGCUGAAGGGCCGGGACCUUCUUA
CGCUGAAGAAUUUCACCGGCGAAGAAAUCAAGUACAUGCUCUGGCUGUCCGCC
GAUCUUAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUCCCGCUGCUGCAAGG
GAAGUCGCUGGGCAUGAUUUUUGAGAAGCGGUCAACUCGCACCCGCCUGUCCA
CUGAAACUGGAUUCGCACUGCUCGGUGGCCAUCCUGCUUCCUGACCACCCAA
GACAUCCACCUCGGCGUGAACGAGUCCCUGACUGACACCGCCCGGGUCUUAUC
CUCGAUGGCCGAUGCUGUGCUUGCGAGGGUGUACAAGCAGUCCGACCUCGACA
CACUCGCGAAGGAGGCCUCCAUCCCCAUCAUCAACGGCCUGUCCGACCUUUAC
CACCCAAUUCAGAUAUCCUCGCCGAUUACCUGACCCUGCAAGAGCACUACUCGUC
GCUCAAGGGGCUUACCCUCUCGUGGAUUGGCGACGGCAACAACAUCUUCACU
CCAUCAUGAUGUCGGCAGCGAAGUUCGGCAUGCAUCUGCAAGCCGCCACGCCU
AAGGGUUAUGAACCAGGAUGCCUCAGUGACCAAGCUCGCCGAACAGUACGCGAA
AGAGAAUGGAACCAAGCUACUUCUGACCAACGACCCCCUGGAGGCCGCUCACG
GCGGCAACGUCCUCAUUACCGAUACUUGGAUUUCGAUGGGACAGGAAGAGGA
AAAGAAGAAGAGACUGCAGGCGUCCAGGGAUACCAGGUCACCAUGAAAACU
GCCAAAGUGGCAGCCUCCGACUGGACCUUCCUUCACUGCCUGCCGAGGAAGCC
UGAAGAGGUGGACGACGAGGUGUUCUACUCCCCGCGCUCCUUGGUGUUUCCUG

AGGCCGAAAACCGGAAGUGGACUAUCAUGGCCGUGAUGGUGUCCCUCCUACC
GACUACUCGCCGCAACUGCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACUGA
CUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCU
AAGCUACAUAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAUGUAGC
CAUUCGUUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM721 (SEQ ID NO: 52)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGUUAUUAACCUUAGAAUUCUC
CUUAACAACGCCGCCUUCGGAUUGGGCAUAACUUUAUGGUCCGCAAUUUCCG
CUGUGGACAGCCUCUGCAAAACAAGGUCCAGCUCAAGGGCCGGGAUCUGCUGA
CUCUCAAGAACUUCACUGGGGAAGAAAUCAAGUACAUGCUCUGGCUGAGCGCC
GACCUC AAGUUCGCAUCAAGCAGAAGGGAGAGUACCUCCCGCUGCUCCAAGG
GAAGUCCUGGGCAUGAUCUUCGAGAAGAGAUAUCCACCCGCACCAGACUUUCCA
CUGAGACUGGCUUCGCCUUGCUGGGAGGCCACCCAUGCUUCCUGACGACCCAG
GACAUUCACCUUGGCGUGAACGAGUCCCUGACUGACACCGCAAGGGUGUUGUC
CUCGAUGGCCGACGCCGUGCUUGCCCGGGUGUACAAGCAGAGCGAUUCUUGACA
CCCUGGCUAAGGAAGCUUCCAUCUCCAUCAUCAACGGUCUGAGCGACCUGUAC
CACCCGAUUCAGAUCCUGGCGGACUACCUAACCUCUGCAAGAGCACUAUAGCUC
CCUGAAGGGCCUCACACUUUCAUGGAUCGGCGACGGCAACAACAUCUGCACU
CUAUUAUGAUGAGCGCUGCCAAAUCGGCAUGCACCUCCAAGCCGCCACGCCU
AAAGGCUACGAGCCCGACGCCUCGGUGACCAAGCUUGCGGAGCAGUACGCGAA
GGAAAACGGCACCAAGCUGCUUUCUACCAACGAUCCUCUGGAAGCGGCCCAUG
GUGGCAACGUGCUCAUUACCGACACUUGGAUCUCCAUGGGACAGGAGGAGGA
AAAGAAGAAGCGGCUCACGGCGUUUCAGGGUUUACCAGGUCACCAUGAAAACCG
CCAAGGUCGCAGCCUCCGACUGGACCUUCCUUCAUUGCCUUCGCGCAAGCCC
GAAGAAGUGGACGAUGAAGUGUUUUACUCACCUCGGUCACUCGUGUUCGCGG
AAGCAGAGAACAGGAAAUGGACCAUUAUGGCCGUGAUGGUGUCCUGCUCACC
GAUUACAGUCCGCAACUGCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACUGA
CUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCU
AAGCUACAUAUACCAACUUCACUUAUACAAAUGUUGUCCCCCAAAUUGUAGC
CAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM722 (SEQ ID NO: 53)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAAUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUC AUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUACCGCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAUGUGACCUCUA
CCAUCCGAUUCAAAUCCUGGCCGAUUAACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUCUCCA
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC

UAAGGGUUAACGAACCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
 AGGAGAACGGAACCAAGCUGCUGCUGACUAACGACCCGCUAGAAGCAGCCAC
 GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGGCCAGGAGGAAG
 AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
 CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUCCCC
 GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUUCUCA
 CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM723 (SEQ ID NO: 54)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUCUCCGCAUCCUC
 CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCUAGGUCCGGAACUUCAG
 AUGUGGCCAGCCGCUUCAAAACAAGGUCCAGCUUAAGGGCCGGGAUCUCCUCA
 CCCUUA AAAACUUCACCGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
 GACCUUAAGUUCGCAUUAAGCAGAAGGGGGAUACCUUCCGCUUCUUAAGG
 AAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCAGGCUUUCUA
 CUGAAACUGGGUUCGCGCUUCUCGGCGGUCAUCCUGCUUCCUCACGACCCAA
 GACAUCCACCUCGGAGUGAACGAAUCCCUACGGAUACUGCCCGCGUGCUUUC
 GAGCAUGGCAGACGCCGUGCUCGCCCGGGUGUACAAACAGUCCGAUCUCGACA
 CUCUCGCCAAGGAGGCGUCAAUUCCUAUUUAUCAACGGUCUUAUGUGACCUUUAC
 CACCCGAUCCAGAUCUCCUGCCGAUUAACCUACACUCCAAGAACACUACAGCUC
 CCUUAAGGGUCUUAACCUUCCUGGAUCGGCGACGGCAACAACAUUCUCCACU
 CCAUCAUGAUGUCCGCCGCAAAGUUCGGCAUGCAUCUUAAGCCGCCACCCCG
 AAGGGCUACGAGCCUGAUGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUAA
 GGAGAACGGAACCAAGCUUCUUCUACUAACGACCCACUCGAAGCAGCCCAUG
 GGGGCAACGUGCUUAUCACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAGAG
 AAGAAGAAGCGGCUCAGGCGUUCAGGGAUAUCAGGUCACCAUGAAAACCGC
 CAAGGUCGUGCCUCCGACUGGACCUUUCUCCACUGCCUCCUCGCAAACCUG
 AAGAAGUGGACGACGAGGUGUUCUACUCGCCCGGAGCCUCGUGUUCCCCGAG
 GCCGAGAAUAGAAAGUGGACCAUUAUGGCCGUGAUGGUGUCACUCCUCACCGA
 CUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACUGACU
 AGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAA
 GCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUAGCCA
 UUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM724 (SEQ ID NO: 55)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUCUCCGCAUCCUC
 CUUAACAACGCCGCGUUUAGAAACGGACAUAACUUCUAGGUCCGGAACUUCAG
 AUGUGGACAGCCGCUUCAAAACAAGGUCCAGCUGAAGGGUCCGGGAUCUUCUG
 ACCCUGAAGAACUUUACCGGAGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGC
 GGACUUGAAGUUCGCAUUAAGCAGAAGGGAGAAUACCUCCCGCUGCUUCAAG
 GAAAGAGCCUCGGAAUGAUUUUUGAGAAGCGCUCAACCAGGACCCGCCUUCU
 ACUGAAACUGGAUUCGCGCUGCUGGGUGGACACCCUGCUUCCUGACGACCCA

GGACAUCCACCUCGGAGUGAACGAAUCCCUCACUGAUACCGCCCCGGGUGUUAU
 CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
 ACUCUGGCCAAGGAGGCGUCAAUCCUAUCAUAACGGACUUAGUGACCUCUA
 CCAUCCGAUUCAAAUCCUGGCCGACUACCUCACCCUGCAAGAACACUACAGCU
 CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGAGAUGGAAACAACAUUCUCCAC
 UCCAUCAUGAUGUCCGCCGCAAAAUUCGGAAUGCAUCUUAAGCCGCCACGCC
 UAAGGGUUAACGAACCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
 AGGAGAACGGUACCAAGCUUCUCCUGACCAACGACCCACUAGAAGCAGCCCAC
 GGUGGAAACGUGCUUAUUACUGACACUUGGAUCUCCAUGGGACAGGAGGAAG
 AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
 CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCGCGGAGCCUCGUGUUCGCC
 GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
 CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUAUACCAACUUAACAUUACAAAAUGUUGUCCCCCAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM725 (SEQ ID NO: 56)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAACCUCGCCAUUCUC
 CUCAACAACGCGUGCCUCCGGAUUGGACAUAACUUAUGGUCCGGAACUUCAG
 AUGCGGACAGCCGCUUCAGAACAAGGUCCAGCUUAAGGGGAGAGAUCCCUUA
 CCCUCAAAAACUUCACUGGCGAAGAAUUAAGUACAUGCUCUGGCUUAGUGCG
 GAUCUCAAGUUCGCAUCAAGCAGAAGGGAGAAUACCUCGCCUCCUUAAGG
 AAAGAGCCUCGGCAUGAUUUUUGAGAAGAGGUCCACCAGAACUCGCCUUAAC
 CCGAGACUGGGUUCGCCUUGGCGGUCACCCUGCUUCCUCACUACCCAA
 GACAUCCACCUCGGCGUGAACGAGAGCCUUAACCGACACCGCCCGCGUGCUC
 CUCAAUGGCCGACGCGUGGUCGCCCCGGGUGUACAAGCAGUCCGACCUUGAUA
 CUCUCGCCAAGGAGGCCUCCAUCCCAAUUAUCAACGGGCUCUCUGAUCUCUAC
 CACCCUAUCCAAAUCCUCGCGGACUACCUCACCCUCCAAGAGCACUAUAGCUC
 GCUCAAGGGCCUCACCCUUCUGGAUUGGCGACGGCAACAACAUUCUUCACU
 CGAUCAUGAUGUCCGCCGCAAGUUCGGCAUGCAUCUCCAAGCCGCGACCCCC
 AAGGGCUACGAGCCUGACGCAUCCGUGACCAAGCUCGCCGAGCAGUACGCGAA
 GGAAAAUGGCACCAAGCUUCUUCUACCAACGACCCCCUUGAGGCCGCUAUG
 GCGGCAACGUGCUCAUCACUGACACUUGGAUCAGCAUGGGCCAGGAGGAGGAA
 AAGAAGAAGCGCCUUCAGGCAUCCAGGGUUAACAGGUCACCAUGAAAACCGC
 CAAAGUGGCCGCCUCCGACUGGACCUUUCUUCACUGUCUCCCGCGGAAGCCUG
 AAGAAGUGGAUGACGAAGUGUUUACUCCCCUCGGUCACUCGUGUUCGCCGAA
 GCAGAAAACAGGAAGUGGACCAUUAUGGCGGUCAUGGUGUCCCUCCUACCGA
 CUACAGCCCGCAGCUUCAGAAACCCAAGUUCUAGCUCGAGCUAGUGACUGACU
 AGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCUAA
 GCUACAUAUAUACCAACUUAACAUUACAAAAUGUUGUCCCCCAAAUGUAGCCA
 UUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM726 (SEQ ID NO: 57)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUCCGCAUCCUC

CUUAACAACGCAGCGUUUAGAAACGGUCACAACUUCAUGGUCCGGAACUUCCG
CUGUGGACAGCCGCUUCAAACAAGGUCCAGCUGAAGGGUCGGGACCUUCUGA
CCCUGAAGAACUUUACUGGAGAAGAGAUCAAGUACAUGCUUUGGCUGUCCGC
GGACUUGAAGUUCGCAUUAAGCAGAAGGGAGAAUACCUUCCGCGUGCUCCAAG
GAAAGAGCCUGGGAAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGAUUCGCGCUGCUGGGUGGUCACCCUUGCUUCCUGACGACCCA
GGACAUUCACCUCGGAGUGAACGAGUCCCUCACUGAUACCGCCAGAGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCUAGGGUGUACAAACAGUCCGAUCUGGAC
ACCCUGGCCAAGGAGGCAUCAAUUCCUAUUUAUCAACGGACUAGUGACCUCUA
CCAUCCGAUUCAAAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGAGAUGGAAACAACAUUCUCCA
UCCAUCAUGAUGUCCGCGGCCAAGUUCGGAAUGCAUCUCCAAGCCGCCACGCC
GAAAGGAUACGAGCCGGACGCUUCCGUGACUAAGCUUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCGCUAGAAGCCGCCAC
GGUGGAAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGACAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUAUCAGGUCACCAUGAAAAC
CGCCAAGGUCGCCGCCUCCGACUGGACCUUCCUUCACUGCCUGCCUCGGAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCGCGGAGCCUCGUGUCCCCU
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUCCUCAC
CGACUACAGCCCGCAGCUUCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
UAAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCCCAAAUAGUAG
CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM727 (SEQ ID NO: 58)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUCUCCGCAUUCUC
CUCAACAACGCAGCCUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAGAACAAGGUCCAGCUCAAGGGCCGGGACCUCUCA
CCCUCAAAAACUUUACCGGCGAAGAGAUCAAGUACAUGCUCUGGCUUUCGGCC
GACCUUAAGUCCGCAUCAAGCAGAAGGGGGAAUACCUUCCGCUGCUUCAAGG
AAAGUCCCUCGGCAUGAUCUUUGAAAAGCGCUCGACCAGGACCCGCCUUUCCA
CUGAAACCGGGUUCGCGCUUCUCGGUGGCCACCCUGCUUCCUCACCACCCAA
GACAUUCACCUCGGAGUGAACGAAUCCCUUACCGAUACCGCAAGAGUGCUUUC
GUCGAUGGCCGAUGCCGUGCUUGCGCGGGUGUACAAGCAGUCAGAUCUCGACA
CUCUCGCCAAGGAGGGCGUCCAUUCUUAUUAUCAACGGCCUUUCCGACCUUAC
CACCCGAUUCAGAUCCUCGCCGAUUACCUCACCCUGCAAGAGCACUACUCGUC
ACUCAAGGGUCUUACCCUCUCCUGGAUCGGCGACGGAAACAACAUCCUCCA
CGAUCAUGAUGUCCGCCGCCAAAUUCGGCAUGCACCUCCAAGCCGCGACCCCG
AAGGGUUACGAGCCCGACGCUUCCGUGACCAAGCUCGCCGAACAGUACGCUAA
GGAAAACGGCACCAAGCUCCUCCUCACUAACGACCCUCUCGAAGCAGCCCAUG
GGGGCAACGUGCUCAUUACUGACACUUGGAUCUCGAUGGGCCAGGAAGAGGA
GAAAAAGAAGCGGCUUCAGGCGUCCAGGGGAUAUCAGGUCACCAUGAAAACC
GCCAAGGUCGCUGCCUCGGACUGGACCUUCCUUCACUGCCUUCGCGCAAGCC
UGAAGAGGUGGACGAUGAGGUGUUCUACUCCCCACGGUCCCUUGUGUUCCCCG
AGGCCGAGAAUAGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCGUCCUCACU
GACUACUCCCCGCAACUUCAGAAGCCUAAGUUCUAGCUCGAGCUAGUGACUGA
CUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUCU

AAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCCCAAAUGUAGC
CAUUCGUAUCUGCUCCUAAUAAAAGAAAGUUUCUUCACAUUCUAG

>mARM728 (SEQ ID NO: 59)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUUAAUCUGAGAAUACU
UCUAAAACAACGCCGCCUUCGGAUUGGCCAUAACUUUAUGGUUCGGAAUUUCC
GCUGCGGCCAGCCGCUGCAGAACAAAGGUCCAGCUGAAGGGAAGAGACUUGCUG
ACCCUCAAGAACUUCACCGGAGAAGAAAUCAAGUAUAUGCUGUGGCUGUCCGC
CGACCUGAAAUUCGCAUCAAGCAGAAGGGCGAAUAUCUGCCGCUGUUGCAAG
GGAAGUCCUGGGGAUGAUCUUCGAGAAGAGGUCCACCAGAACACGGCUUUA
ACCGAAACCGGGUUUGCACUGCUGGGUGGACACCCUGUUUUCUGACCACUCA
AGAUAUCCACCUGGGCGUGAACGAGUCCCUUACCGACACUGCUAGGGUGUUGU
CCAGCAUGGCCGAUGCCGUCCUGGCUCGCGUGUACAAGCAGUCCGACCUGGAU
ACCCUGGCAAAGGAAGCGUCCAUUCCTCAUUAUCAACGGGCUGUCCGACCUGUA
CCAUCCGAUUCAAAUCCUGGCGGACUACCUGACUCUGCAAGAGCAUUACAGCA
GCUUGAAGGGGCUUACUCUCUCUGUGGAUCGGCGACGGGAACAACAUCCUGCAC
UCCAUCAUGAUGUCCGCCGCCAAGUUCGGGAUGCAUUUGCAAGCUGCGACCCC
GAAAGGUUACGAGCCCGAUGCUAGCGUAACUAAGCUUGCCGAACAGUACGCCA
AAGAGAAUGGUACAAAACUGCUUCUGACUAACGACCCGCUGGAAGCAGCCAC
GGCGGGAACGUGCUGAUAAACCGACACCUGGAUUUCAUUGGGGCAGGAGGAAG
AGAAGAAGAAGCGACUGCAGGCGUUCCAAGGCUAUCAGGUUACCAUGAAAAC
CGCCAAAGUGGCAGCCAGCGAUUUGGACUUUCCUGCACUGUCUGCCGCGGAAGC
CCGAGGAAGUUGAUGACGAAGUAUUCUACUCACCCCGGAGCCUCGUGUUCCCC
GAGGCCGAAAACCGGAAGUGGACUAUUUUGGCCGUGAUGGUGUCGUGUUGA
CCGACUACAGCCCGCAACUGCAGAAGCCGAAGUUUAGCUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAGAAAGUUUCUUCACAUUCUAG

>mARM729 (SEQ ID NO: 60)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUUUAACCGUGAGGAUCCUU
UUGAACAACGCCGCCUUCGCAACGGCCACAACUUUAUGGUCCGCAAUUUCCG
CUGCGGGCAGCCGCUGCAGAACAAAGGUCCAGCUGAAGGGCCGGGAUCUGCUGA
CCCUGAAGAACUUCACCGGGGAGGAAAUCAAGUACAUGCUUUGGCUCUCCGCC
GAUCUGAAGUUCAGAAUCAAGCAGAAGGGAGAGUACCUCCCGUUGCUGCAAG
GAAAGUCACUCGGAAUGAUUUUCGAAAAGAGAAGCACUAGGACCCGCCUCUCA
ACUGAAACCGGGUUCGCGCUGCUCGGGGGCCAUCCGUGUUUCCUGACUACCCA
AGACAUCCACCUGGGAGUGAACGAGUCGUGACCGACACCGCACGCGUGCUGU
CAUCCAUGGCGGACGCAGUGCUUUGCCCGGGUGUACAAGCAGUCGGACCUGGAC
ACUCUUGCCAAGGAGGCAUCAAUCCCCAUCAUUAACGGACUGUCCGAUCUCUA
CCACCCGAUUCAGAUCUGGCUGACUACCUAACCCUGCAAGAGCACUACUCAA
GCCUGAAGGGGCUGACCCUGUCGUGGAUCGGGGACGGCAACAACAUUCUGCAC
UCCAUCAUGAUGUCGGCGGCUAAGUUCGGGAUGCAUUUGCAAGCGGCAACUCC
GAAGGGUUAUGAACCCGACGCCUCCGUGACCAAGCUGGCCGAACAGUACGCCA
AGGAAAACGGAACCAAGUUGCUGCUGACUAAUGAUCCCCUGGAGGCGGCCAC
GGGGGGAACGUGCUGAUAAACCGAUACCUGGAUCUCCAUGGGGCAGGAAGAAG

AGAAGAAAAAGCGGCUGCAGGCAUUCCAGGGAUACCAGGUCACCAUGAAAACC
GCAAAAGUGGGCAGCCAGCGACUGGACUUUCCUCCAUGGCCUGCCGCGAAAGCC
GGAGGAGGUCGAUGACGAGGUGUUCUACUCCCCGCGGUCGCUGGUGUUCGCCG
AGGCGGAAAACCGGAAGUGGACCAUUAUGGCCGUGAUGGUGUCACUCCUGAC
UGACUACAGCCCCGCAACUGCAGAAGCCGAAGUUCUAGCUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
UAAGCUACAUAUAACCAACUACACUACAAAAUGUUGUCCCCCAAAUUGUAG
CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUCUAG

>mARM1787 (SEQ ID NO: 61)

CUUAAGGGGGGCGCUGCCUACGGAGGUGGCAGCCAUCUCCUUCUCGGCAUCAAG
CUUACCAUGGUGCCCCAGGCCUGCUCUUGGUCCCCGUGCUGGUGUUCUUUU
CUGCUUCGGCAAGUUCUUUACUACACCAUCCCCGACAAGCUGGGGGCCGUGGA
GCCCCAUCGACAUCACCACCUGUCCUGCCCCAACACCUCGUGGUCGAGGAC
GAGGGCUGCACCAACCUGAGCGGGUUCUCCUACAUGCUUUUCAAUCUCCGCAU
CCUCCUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCUUGGUCCGGAACU
UCAGAUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGGCCGGGAUCUU
CUGACCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUC
CGCGGACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCUGCUUC
AAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUU
UCUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGAC
CCAGGACAUCACCUCGGAGUGAACGAAUCCCUCACCGAUACCGCCCCGGGUGU
UAUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUG
GACACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAGUGACCU
CUACCAUCCGAUUCAGAUCUGGCCGAUUACCUCACCCUGCAAGAACACUACA
GCUCCCUUGAAGGGUCUGACAUCUUGUCCUGGAUCGGCGACGGCAACAACAUUCUC
CAUUCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCAC
GCCGAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACG
CUAAGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCC
CACGGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAAGA
AGAGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUUAUCAGGUCACCAUGAAA
ACCGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAA
GCCUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUC
CCGAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUC
ACCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGAUAAGUGAAUGCAA
GGCUGGCCGGAAGCCCUUGCCUGAAAGCAAGAUUUCAGCCUGGAAGAGGGCAA
AGUGGACGGGAGUGGACAGGAGUGGAUGCGAUAAAGAUGUGGUUUGAAGCUGA
UGGGUGCCAGCCCUGCAUUGCUGAGUCAAUCAAUAAAGAGCUUUCUUUUGACC
CAUUCUAGAUCUAG

>mARM1788 (SEQ ID NO: 62)

AGGAAACUUAAGAUUAUUACAUCAAAAACAAAAAGCCGCCAAUGCUGUUAAC
CUGCGCAUCCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCUAGGU
GCGCAACUUCGCGUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCC
GCGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUG
UGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCC
CCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCA
CCCGCCUGAGCACCGAGACAGGCCUGGCCUGCUGGGCGGCCACCCUGCUUC
CUGACCACCCAGGACAUCACCUGGGCGUGAACGAGAGCCUGACCGACACCGC
CCGCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGA

GCGACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUG
 AGCGACCUGUACCAACCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGA
 GCACUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACA
 ACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAG
 GCCGCCACCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGA
 GCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGG
 AGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGC
 CAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUCCAGGGCUACCAGGUGAC
 CAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUCCUGCACUGCCUGC
 CCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUG
 GUGUUCCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAG
 CCUGCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAACGCCGA
 AGCCUGCAGCCAUGCGACCCACGCCACCCCGUGCCUCCUGCCUCCGCGCAGCC
 UGCAGCGGGAGACCCUGUCCCCGCCCCAGCCGUCCUCCUGGGGUGGACCCUAG
 UUUAUAUAAAGAUUCACCAAGUUUCACGCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

>mARM1789 (SEQ ID NO: 63)

CUUAAGGGGGGCGCUGCCUACGGAGGUGGCAGCCAUCUCCUUCUCGGCAUCAAG
 CUUACCAUGGUGCCCCAGGCCUGCUCUUGGUCCCCGUGCUGGUGUCCCCCU
 CUGCUUCGGCAAGUUCCTCAUCUACACCAUCCCCGACAAGCUGGGGGCCGUGGA
 GCCCCAUCGACAUCCACCACCUUGUCCUGCCCCAACCAACCUCUGGUGCAGGAC
 GAGGGCUGCACCAACCUGAGCGGGUUCUCCUACAUGCUGUUAACCUGCGCAU
 CCUGCUGAACAAACGCCGCCUUCGCAACGGCCACAACUUAUGGUGCGCAACU
 UCCGCUGCGGCCAGCCCCUGCAGAACAAAGGUGCAGCUGAAGGGCCGCGACCUG
 CUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAG
 CGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGC
 AGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUG
 AGCACCGAGACAGGCCUGGGCCUGCUGGGCGGCCACCCUGCUUCCUGACCAC
 CCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCGUGC
 UGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGAGCGACCUG
 GACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCU
 GUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACUACA
 GCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCUG
 CACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCAC
 CCCCAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGGCCGAGCAGUACG
 CCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCCGCC
 CACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGA
 GGAGAAGAAGAAGCGCCUGCAGGCCUCCAGGGCUACCAGGUGACCAUGAAGA
 CCGCCAAGGUGGCCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAG
 CCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUCCCC
 CGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGA
 CCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAAUAAGUGAUGCAAGG
 CUGGCCGGAAGCCCUUGCCUGAAAGCAAGAUUUCAGCCUGGAAGAGGGCAAAG
 UGGACGGGAGUGGACAGGAGUGGAUGCGAUAAAGAUUGGUGUUGAAGCUGAUG
 GGUGCCAGCCUGCAUUGCUGAGUCAUAUAUAAAGAGCUUUCUUUUGACCCA
 UUCUAGAUCUAG

>mARM1790 (SEQ ID NO: 64)

AGGAAACUUAAGAUUAUUACAUCAAAACAAAAAGCCGCCAAUGCUUUUCAAU
 CUCCGCAUCCUCCUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGU
 CCGGAACUUCAGAUUGGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCC
 GGGAUUCUUCUGACCCUGAAGAACUUUACUGGGCGAAGAGAUCAAGUACAUGCU
 CUGGCUUCUCCGCGGACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUC
 CGCUGCUUCAAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCACCCAGG
 ACCCGCCUUUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGGCCACCCUGCUU
 CCUGACGACCCAGGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUACCG
 CCCGGGUGUUUUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAG
 UCCGAUCUGGACACUCUGGCCAAGGAGGCGUCAAUUCCUAUUUAUCAACGGCCU
 UAGUGACCUCUACCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAG
 AACACUACAGCUCUCCUGAAGGGUCUGACAUAUUGUCCUGGAUCGGCGACGGCAAC
 AACAUUCUCCAUAUCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUA
 AGCCGCCACGCCGAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCG
 AGCAGUACGCUAAGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUA
 GAAGCAGCCCACGGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGG
 CCAGGAAGAAGAGAAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUUAUCAGGUC
 ACCAUGAAAACCGCCAAGGUUCGUGCCUCCGACUGGACCUUCCUGCACUGCCU
 GCCUCGCAAGCCUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCC
 UCGUGUCCCCGAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGU
 GUCACUGCUCACCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGACGC
 CGAAGCCUGCAGCCAUGCGACCCACGCCACCCCGUGCCUCCUGCCUCCGCGCA
 GCCUGCAGCGGGAGACCCUGUCCCCGCCCCAGCCGUCCUCCUGGGGUGGACCC
 UAGUUUAAUAAAGAUUCACCAAGUUUCACGCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

>mARM1791 (SEQ ID NO: 65)

CUUAAGGGGGCGCUGCCUACGGAGGUGGCAGCCAUCUCCUUCUCGGCAUCAAG
 CUUACCAUGGUGCCCCAGGCCUGCUCUUGGUCCCGCUGCUGGUGUUCUUUU
 CUGCUUCGGCAAGUUCUCCCAUCUACACCAUCCCGACAAGCUGGGGCGGUGGA
 GCCCCAUCGACAUCCACCACCUGUCCUGCCCCAACAACCUCGUGGUCGAGGAC
 GAGGGCUGCACCAACCUGAGCGGGUUCUCCUACAUGCUUUUCAACCUGAGAAU
 CCUCUUGAACAUAUGCUGCUUUUCGGAAUGGCCACAACUUUAUGGUUCGGAACU
 UCCGUUGCGGCCAGCCUUUACAAAACAAGGUCCAGCUGAAGGGCCGGGAUUUG
 CUCACACUAAAGAACUUUACUGGAGAAGAGAUCAAGUACAUGCUAUGGCUGU
 CGGCCGACCUGAAGUUCGUAUCAAGCAGAAGGGAGAAUACCUUCCGCUGCUU
 CAAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCU
 UUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGA
 CCCAGGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUACCGCCCGGGUG
 UUAUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCU
 CGAUACCUUGGCAAAGGAGGCUUCCAUAUCCCAUCAUAACGGCCUGAGCGACC
 UGUACCACCCAAUCCAAAUCCUGGCUGACUACCUGACCCUGCAAGAGCACUAC
 AGCAGCCUGAAGGGUCUGACCCUGUCAUGGAUUGGCGAUGGAAACAAUAUUC
 UGCACUCCAUCAUGAUGUCCGCCGCGAAGUUCGGAAUGCAUCUGCAAGCCGCC
 ACUCCAAAAGGAUACGAACCGGAUGCAUCCGUGACCAAGUUGGCGGAACAGUA
 CGCGAAGGAGAACGGAACCAAGCUCCUGCUGACUAACGACCCGCUCGAGGCUG
 CGAUGGGGGUAACGUGCUGAUUACGGACACCUGGAUCUCCAUGGGGGCAGGA
 GGAAGAGAAGAAGAAGAGACUGCAGGCAUUCAGGGGUACCAGGUCACCAUG
 AAAACCGCAAAAGUGGCAGCUUCGGACUGGACUUUCCUGCAUUGCCUGCCGAG

GAAGCCGGAGGAAGUCGACGACGAAGUGUUCUACUCGCCUCGGUCCCUGGUGU
 UCCCCGAGGCCGAAAACCGGAAGUGGACCAUCAUGGCCGUGAUGGUGUCCUUG
 CUGACUGACUAUAGCCCCGACGUCGAGAAGCCUAAGUUCUAGAUAAAGUGAUGC
 AAGGCUGGCCGGAAGCCCUUGCCUGAAAGCAAGAUUUCAGCCUGGAAGAGGGC
 AAAGUGGACGGGAGUGGACAGGAGUGGAUGCGAUAAAGAUGUGGUUUGAAGCU
 GAUGGGUGCCAGCCCUGCAUUGCUGAGUCAAUCAAUAAAGAGCUUUCUUUUG
 ACCCAUUCUAGAUCUAG

>mARM1792 (SEQ ID NO: 66)

UGAGUGUCGUACAGCCUCCAGGCCCCCCCCCCUCCCGGGAGAGCCAUAGUGGUCU
 GCGGAACCGGUGAGUACACCGGAUUGCCGGGAAGACUGGGUCCUUUCUUGG
 AUAAACCCACUCUAUGCCCCGGCCAUUUGGGCGUGCCCCCGCAAGACUGCUAGC
 CGAGUAGUGUUGGGUUGCGAUGCUGUUAACCUGCGCAUCCUGCUGAACAACG
 CCGCCUUCGCAACGGGCCACAACUUCAUGGUGCGCAACUUCGCGUGCGGGCCAG
 CCCUGCAGAACAAAGGUGCAGCUGAAGGGCCGCGACCUGCUGACCCUGAAGAA
 CUUCACCGGCGAGGAGAUAAGUACAUGCUGUGGCUGAGCGCCGACCUGAAGU
 UCCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCAGGGCAAGAGCCUG
 GGCAUGAUCUUCGAGAAGCGCAGCACCCGCAACCCGCCUGAGCACCGAGACAGG
 CCUGGGCCCUGCUGGGCGGGCCACCCUGCUUCCUGACCACCCAGGACAUCCACCU
 GGGCGUGAACGAGAGCCUGACCGACACCGCCCCGUGCUGAGCAGCAUGGCCG
 ACGCCGUGCUGGCCCGCGUGUACAAGCAGAGCGACCUGGACACCCUGGCCAAG
 GAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUACCACCCCAUCCA
 GAUCCUGGCCGACUACCUGACCCUGCAGGAGCACUACAGCAGCCUGAAGGGCC
 UGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGCACAGCAUCAUGAUG
 AGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCACCCCCAAGGGGCUACGA
 GCCCCGACGCCAGCGUGACCAAGCUGGGCCGAGCAGUACGCCAAGGAGAACGGCA
 CCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCCACGGCGGCAACGUG
 CUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGAGGAGAAGAAGAAGC
 GCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGACCGCCAAGGUGGCC
 GCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCCCGAGGAGGUGGA
 CGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUCCCCCGAGGCCGAGAACC
 GCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACCGACUACAGCCCC
 CAGCUGCAGAAGCCCCAAGUUCUGAAUAAGUGAUAGAGCGGCAAAACCCUAGCUA
 CACUCCAUAAGCUAGUUUCUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
 UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
 UUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
 GGCUCCAUCUAGCCCUAGUCACGGCUAGCUGUGAAAGGUCCGUGAGCCGCAU
 GACUGCAGAGAGUGCCGUAACUGGGCCUCUCUGCAGAUCAUGUUCUAG

>mARM1793 (SEQ ID NO: 67)

AGGAAACUUAAGAUUAUUAACAUCAAAACAAAAAGCCGCCAAUGCUUUUCAAU
 CUCCGCAUCCUCCUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGU
 CCGGAACUUCAGAUGUGGCCAGCCGCUUCAAAACAAGGUCCAGCUGAAGGGCC
 GGAUCUUCUGACCCUGAAGAACUUUACUGGCGAAGAGAUAAGUACAUGC
 CUGGCUCUCCGCGGACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUC
 CGCUGCUUCAAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGG
 ACCCGCCUUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUU
 CCUGACGACCCAGGACAUCCACCUCGGAGUGAACGAAUCCCUCACCGAUACCG
 CCCGGGUGUUUUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAG
 UCCGAUCUGGACACUCUGGCCAAGGAGGCGUCAAUUCCUAUUUAUCAACGGCCU
 UAGUGACCUCUACCAUCCGAUUCAGAUCUGGCCGAUUAACCUCACCCUGCAAG

AACACUACAGCUCCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAAC
 AACAUUCUCCAUCCAUAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUA
 AGCCGCCACGCCGAAGGGUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCG
 AGCAGUACGCUAAGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUA
 GAAGCAGCCCACGGGGGCAACGUGCUUAUUAUGACACCUGGAUCUCCAUGGG
 CCAGGAAGAAGAGAAAAAGAAGCGGCUGCAGGCGUCCAGGGAUUAUCAGGUC
 ACCAUGAAAACCGCCAAGGUCGCGUGCCUCCGACUGGACCUUCCUGCACUGCCU
 GCCUCGCAAGCCUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCC
 UCGUGUUCGCCGAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGU
 GUCACUGCUCACCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGGCUG
 GAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGCCCUCCUCC
 CCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAGCAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAA

>mARM1794 (SEQ ID NO: 68)

AGGAAACUUAAGAUUAUUAUCAAAAACAAAAAGCCGCCAAUGCUGUUAAC
 CUGCGCAUCCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUAUGGU
 GCGCAACUUCGCGUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCC
 GCGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUG
 UGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCC
 CCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCA
 CCCGCCUGAGCACCGAGACAGGCCUGGCCUGCUGGGCGGCCACCCUGCUUC
 CUGACCACCCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGC
 CCGCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGA
 GCGACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUG
 AGCGACCUGUACCACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGA
 GCACUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACA
 ACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAG
 GCCGCCACCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGA
 GCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGG
 AGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGC
 CAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGAC
 CAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGC
 CCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUG
 GUGUUCGCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAG
 CCUGCUGACCGACUACAGCCCCAGCUGCAGAAGCCCAAGUUCUGAGCUGGAG
 CCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGCCCUCCUCCCU
 CCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAGCAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAA

>mARM1795 (SEQ ID NO: 69)

AGGAAACUUAAGAUUAUUAACAUCAAAACAAAAAGCCGCCAAUGCUUUUCAAC
CUGAGAAUCCUCUUGAACAAUGCUGCUUUUCGGAUUGGCCACAACUUAUGG
UUCGGAACUUCCGUUGCGGCCAGCCUUUACAAAACAAGGUCCAGCUGAAGGGC
CGGGAUUUGCUCACACUAAAGAACUUUACUGGAGAAGAGAUCAAGUACAUGC
UAUGGCUGUCGGCCGACCUGAAGUUCGUAUCAAGCAGAAGGGAGAAUACCU

UCCGCUGCUUCAAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUC AACCA
GGACCCGCCUUUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUCG
UUCUGACGACCCAGGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUAC
CGCCCGGGUGUUAUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAAC
AGUCCGAUCUCGAUACCUUGGCAAAGGAGGCUUCCAUUCCCAUCAUCAACGGC
CUGAGCGACCUGUACCACCCAAUCCAAAUCCUGGCUGACUACCUGACCCUGCA
AGAGCACUACAGCAGCCUGAAGGGUCUGACCCUGUCAUGGAUUUGGCGAUGGA
AACAAUAUUCUGCACUCCAUCAUGAUGUCCGCCGCGAAGUUCGGAAUGCAUCU
GCAAGCCGCCACUCCAAAAGGAUACGAACCGGAUGCAUCCGUGACCAAGUUGG
CGGAACAGUACGCGAAGGAGAACGGAACCAAGCUCCUGCUGACU AACGACCCG
CUCGAGGCUGCGCAUGGGGGU AACGUGCUGAUUACGGACACCUGGAUCUCCA
GGGGCAGGAGGAAGAGAAGAAGAAGAGACUGCAGGCAU UCCAGGGGUACCAG
GUCACCAUGAAAACCGCAAAAGUGGCAGCUUCGGACUGGACUUUCCUGCAUUG
CCUGCCGAGGAAGCCGGAGGAAGUCGACGACGAAGUGUUCUACUCGCCUCGGU
CCCUGGUGU UCCCCGAGGCCGAAAACCGGAAGUGGACCAUCAUGGCCGUGAUG
GUGUCCUUGCUGACUGACUAUAGCCCGCAGCUGCAGAAGCCUAAGUUCUAGGC
UGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGGCCUCC
UCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAG
CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

>mARM1796 (SEQ ID NO: 70)

AGGAAACUUAAGAAUUAUUGGUUAAAGAAGUAUAUUAGUGCUAAUUUCCCUC
CGUUUGUCCUAGCUUUUCUCUUCUGUCAACCCACACGCCUUUGGCACAAUGC
UUUUCAAUCUCCGCAUCCUCCUUAACAACGCCGCGUUUAGAAACGGCCACAAC
UUCAUGGUCCGGAACUUCAGAUGUGGCCAGCCGCUUCAAACAAGGUCCAGCU
GAAGGGCCGGGAUCUUCUGACCCUGAAGAACUUUACUGGCGAAGAGAUAAG
UACAUGCUCUGGCUCUCCGCGGACUUGAAGUCCGCAUUAAGCAGAAGGGGGA
AUACCUUCCGCUGCUUCAAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCU
CAACCAGGACCCGCCUUUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCAC
CCCUGCUUCCUGACGACCCAGGACAUCCACCUCGGAGUGAACGAAUCCCUCAC
CGAUACCGCCCCGGGUGUUAUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGU
ACAAACAGUCCGAUCUGGACACUCUGGCCAAGGAGGCGUCAAUUCCUAUUUUC
AACGGCCUAGUGACCUCUACCAUCCGAUUCAGAUCUGGCCGAUUACCUCAC
CCUGCAAGAACACUACAGCUCUCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCG
ACGGCAACAACAUUCUCCAUCCAUCGAUGUCCGCCGCAAAAUUCGGCAUG
CAUCUUAAGCCGCCACGCCGAAGGGUACGAGCCCGACGCUUCCGUGACUAA
GCUCGCCGAGCAGUACGCUAAGGAGAACGGAACCAAGCUUCUGCUGACUACG
ACCCACUAGAAGCAGCCCACGGGGGCAACGUGCUUAUUACUGACACCUGGAUC
UCCAUGGGCCAGGAAGAAGAGAAAAAGAAGCGGCUGCAGGCGUCCAGGGAU
AUCAGGUCACCAUGAAAACCGCCAAGGUCGCGUGCCUCCGACUGGACCUCCUG
CACUGCCUGCCUCGCAAGCCUGAAGAAGUGGACGACGAGGUGUUCUACUCGCC
ACGGAGCCUCGUGUCCCCGAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCG
UGAUGGUGUCACUGCUCACCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUC
UAGCUCGAGACACAUCACAACCACAACCUUCUCAGGCUACCCUGAGAAAAAAA
GACAUGAAGACUCAGGACUCAUCUUUUCUGUUGGUGUAAAAUCAACACCCUA
AGGAACACAAAUUUCUUUAAACAUUUGACUUCUUGUCUCUGUGCUGCAAUUA
AUAAAAAUGGAAAGAAUCUAUCUAG

>mARM1797 (SEQ ID NO: 71)

AGGAAACUUAAGAAUUAUUGGUUAAAGAAGUAUAUUAGUGCUAAUUUCCCUC
CGUUUGUCCUAGCUUUUCUCUUCUGUCAACCCACACGCCUUUGGCACAAUGC
UGUUCAACCUGCGCAUCCUGCUGAACAACGCCGCCUUCCGCAACGGCCACAAC
UUCAUGGUGCGCAACUUCGCGUGCGGCCAGCCCCUGCAGAACAAAGGUGCAGCU
GAAGGGCCCGGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUAAGU
ACAUGCUGUGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAG
UACCUGCCCCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAG
CACCCGCACCCGCCUGAGCACCGAGACAGGCCUGGGCCUGCUGGGCGGGCCACCC
CUGCUUCCUGACCACCCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCG
ACACCGCCCCGCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGGCCCCGCGUGUAC
AAGCAGAGCGACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAA
CGGCCUGAGCGACCUGUACCACCCCAUCCAGAUCUCCUGGCCGACUACCUGACCC
UGCAGGAGCACUACAGCAGCCUGAAGGGGCCUGACCCUGAGCUGGAUCGGCGAC
GGCAACAACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCA
CCUGCAGGCCGCCACCCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGC
UGGCCGAGCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGAC
CCCCUGGAGGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAG
CAUGGGCCAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACC
AGGUGACCAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCAC
UGCCUGCCCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCG
CAGCCUGGUGUUCCTCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGA
UGGUGAGCCUGCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGA
CUCGAGACACAUCACAACCACAACCUUCUCAGGCUACCCUGAGAAAAAAAGAC
AUGAAGACUCAGGACUCAUCUUUUCUGUUGGUGUAAAAUCAACACCCUAAGG
AACACAAAUUUCUUUAAACAUUUGACUUCUUGUCUCUGUGCUGCAAUUAUA
AAAAAUGGAAAGAAUCUAUCUAG

>mARM1798 (SEQ ID NO: 72)

AGGAAACUUAAGAAUUAUUGGUUAAAGAAGUAUAUUAGUGCUAAUUUCCCUC
CGUUUGUCCUAGCUUUUCUCUUCUGUCAACCCACACGCCUUUGGCACAAUGC
UUUUCAACCUGAGAAUCCUCUUGAACAACUGCUGCUUUUCGGAUUGGCCACAAC
UUUAUGGUUCGGAACUUCGUGCGGCCAGCCUUUACAAAACAAGGUCCAGCU
GAAGGGCCGGGAUUUGCUCACACUAAAGAACUUUACUGGAGAAGAGAUAAG
UACAUGCUAUGGCUGUCGGCCGACCUGAAGUUCGUAUCAAGCAGAAGGGAG
AAUACCUUCCGUGCUUCAAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGC
UCAACCAGGACCCGCCUUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCA
CCCCUGCUUCCUGACGACCCAGGACAUCCACCUCGGAGUGAACGAAUCCCUCA
CCGAUACCGCCCCGGGUGUUAUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUG
UACAAACAGUCCGAUCUCGAUACCUUGGCAAAGGAGGCUUCCAUUCCCAUCAU
CAACGGCCUGAGCGACCUGUACCACCCAAUCCAAAUCCUGGCUGACUACCUGA
CCCUGCAAGAGCACUACAGCAGCCUGAAGGGGUCUGACCCUGUCAUGGAUUGGC
GAUGGAAACAAUAUUCUGCACUCCAUCAUGAUGUCCGCCGCGAAGUUCGGAU
GCAUCUGCAAGCCGCCACUCCAAAAGGAUACGAACCGGAUGCAUCCGUGACCA
AGUUGGCGGAACAGUACGCGAAGGAGAACGGAACCAAGCUCCUGCUGACUAC
GACCCGCUCGAGGCUGCGCAUGGGGGUAAACGUGCUGAUUACGGACACCUGGAU
CUCCAUGGGGCAGGAGGAAGAGAAGAAGAAGAGACUGCAGGCAUUCAGGGG
UACCAGGUCACCAUGAAAACCGCAAAAGUGGCAGCUUCGGACUGGACUUCCU
GCAUUGCCUGCCGAGGAAGCCGGAGGAAGUCGACGACGAAGUGUUCUACUCGC
CUCGGUCCCUGGUGUUCCTCCGAGGCCGAAAACCGGAAGUGGACCAUCAUGGCC

GUGAUGGUGUCCUUGCUGACUGACUAUAGCCCGCAGCUGCAGAAGCCUAAGUU
CUAGCUCGAGACACAUCACAACCACAACCUUCUCAGGCUACCCUGAGAAAAAA
AGACAUAGAAGACUCAGGACUCAUCUUUUCUGUUGGUGUAAAAUCAACACCCU
AAGGAACACAAAUUUCUUUAAACAUUUGACUUCUUGUCUCUGUGCUGCAAUU
AAUAAAAAAUGGAAAGAAUCUAUCUAG

>mARM1799 (SEQ ID NO: 73)

AUUUAUUACAUCAAAACAAAAAGCCGCCACCAUGCUGUUCAACCUGCGCAUCCU
GCUGAACAAACGCCGCCUUCCGCAACGGCCACAACUUCAUGGUGCGCAACUUC
GCUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCCGCGACCUGCUG
ACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAGCGC
CGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCAGG
GCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCGACCCGCCUGAGC
ACCGAGACAGGCUUCGCCUGCUGGGCGGGCCACCCUGCUUCCUGACCACCCA
GGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCGUGCUGA
GCAGCAUGGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGAGCGACCUGGAC
ACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGGCCUGAGCGACCUGUA
CCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACUACAGCA
GCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGCAC
AGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCACCCC
CAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGGCCGAGCAGUACGCCA
AGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCCAC
GGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGAGGA
GAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGACCG
CCAAGGUGGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCC
GAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUCCCCGA
GGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACCG
ACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAAUGAGCU
GGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGGCCCUCCU
CCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAUC
UAG

>mARM1800 (SEQ ID NO: 74)

AGGAAACUUAAGAAUUAUUGGUUAAAGAAGUAUAUUAGUGCUAAUUUCCCUC
CGUUUGUCCUAGCUUUUCUCUUCUGUCAACCCACACGCCUUUGGCACAAUGC
UGUUCAACCUGCGCAUCCUGCUGAACAACGCCGCCUUCCGCAACGGCCACAAC
UUCAUGGUGCGCAACUUCGCGUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCU
GAAGGGCCGCGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGU
ACAUGCUGUGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAG
UACCUGCCCCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAG
CACCCGCACCCGCCUGAGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCC
CUGCUUCCUGACCACCCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCG
ACACCGCCCGCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUAC
AAGCAGAGCGACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCA
CGGCCUGAGCGACCUGUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCC
UGCAGGAGCACUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGAC
GGCAACAACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCA
CCUGCAGGCCGCCACCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGC
UGGCCGAGCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGAC
CCCCUGGAGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAG

CAUGGGCCAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCCAGGGCUACC
AGGUGACCAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCAC
UGCCUGCCCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCG
CAGCCUGGUGUUCCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGA
UGGUGAGCCUGCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGA
CUCGAGACACAUCACAACCACAACCUUCUCAGGCUACCCUGAGAAAAAAGAC
AUGAAGACUCAGGACUCAUCUUUUCUGUUGGUGUAAAAUCAACACCCUAAGG
AACACAAAUUUCUUUAAACAUUUGACUUCUUGUCUCUGUGCUGCAAUUAUA
AAAAAUGGAAAGAAUCUAUCUAG

>mARM1801 (SEQ ID NO: 75)

AGGAAACUUAAGAUUAUUACAUCAAAACAAAAAGCCGCCAAUGCUGUUAAC
CUGCGCAUCCUGCUGAACAACGCCGCCUUCCGCAACGGCCACAACUUAUGGU
GCGCAACUUCCGCUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCC
GCGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUAAGUACAUGCUG
UGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCC
CCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCA
CCCGCCUGAGCACCGAGACAGGCUUCGCCUGCUGGGCGGCCACCCUGCUUC
CUGACCACCCAGGACAUCACCCUGGGCGUGAACGAGAGCCUGACCGACACCGC
CCGCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGA
GCGACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUG
AGCGACCUGUACCACCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGA
GCACUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACA
ACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAG
GCCGCCACCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGGCCGA
GCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGG
AGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGC
CAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUCCAGGGCUACCAGGUGAC
CAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGC
CCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUG
GUGUUCGGGAGGCGGAGAACCAGCAAGUGGACCAUCAUGGCCGUGAUGGUGAG
CCUGCUGACCGACUACAGCCCCAGCUGCAGAAGCCCAAGUUCUGAACGCCGA
AGCCUGCAGCCAUGCGACCCCCAGCCACCCCGUGCCUCCUGCCUCCGCGCAGCC
UGCAGCGGGAGACCCUGUCCCCGCCCCAGCCGUCCUCCUGGGGUGGACCCUAG
UUUAAUAAAGAUUCACCAAGUUCACGCAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

>mARM1802 (SEQ ID NO: 76)

CUUAAGGGGGCGCUGCCUACGGAGGUGGCAGCCAUCUCCUUCUCGGCAUCAAG
CUUACCAUGGUGCCCCAGGCCUGCUCUUGGUCCCGCUGCUGGUGUUCUUCCCU
CUGCUUCGGCAAGUUCUCCCAUCUACACCAUCCCGACAAGCUGGGGGCCGUGGA
GCCCCAUCGACAUCACCAACCUGUCCUGCCCCAACAACCUCGUGGUCGAGGAC
GAGGGCUGCACCAACCUGAGCGGGUUCUCCUACAUGCUGUUAACCUGCGCAU
CCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUAUGGUGCGCAACU
UCCGCUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCCGCGACCUG
CUGACCCUGAAGAACUUCACCGGCGAGGAGAUAAGUACAUGCUGUGGCUGAG
CGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGC
AGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUG
AGCACCGAGACAGGCUUCGCCUGCUGGGCGGCCACCCUGCUUCCUGACCAC

CCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCCGCGUGC
 UGAGCAGCAUGGCCGACGCCGUGCUGGCCCCGCGUGUACAAGCAGAGCGACCUG
 GACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCU
 GUACCACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGAGCACUACA
 GCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUG
 CACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCAC
 CCCCAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACG
 CCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCCGCC
 CACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGA
 GGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGA
 CCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAG
 CCGGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUCCCC
 CGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGA
 CCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAAUAAGUGAAUGCAAG
 GCUGGCCGGAAGCCCUUGCCUGAAAGCAAGAUUUCAGCCUGGAAGAGGGCAA
 GUGGACGGGAGUGGACAGGAGUGGAUGCGAUAAGAUGUGGUUUGAAGCUGAU
 GGGUGCCAGCCUGCAUUGCUGAGUCAAUCAAUAAAGAGCUUUCUUUGACCC
 AUUCUAGAUCUAG

>mARM1803 (SEQ ID NO: 77)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGGGCGUCUUAACCUGCGGAUC
 CUGCUGAACAAACGCCGCCUUCGGAACGGCCACAACUUAUGGUCCGCAACUU
 CAGAUGC GGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCCGGGACCUGC
 UGACCCUGAAGAACUUCACCGGCCAAGAGAUCAAGUACAUGCUGUGGCUGAGC
 GCCGACCUGAAGUUCGGAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCA
 AGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGGAGCACCCGGACCCGGCUGA
 GCACCGAGACAGGCUUUGCCUGCUGGGAGGCCACCCUGCUUUCUGACCACC
 CAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCAGAGUGCU
 GAGCAGCAUGGCCGACGCCGUGCUGGCCCGGGUGUACAAGCAGAGCGACCUGG
 ACACCCUGGCCAAAGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUG
 UACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAACACUACAG
 CUCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGC
 ACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCAUCUGCAGGCCGCCACC
 CCAAGGGCUACGAGCCUGAUGCCAGCGUGACCAAGCUGGCCGAGCAGUACGC
 CAAAGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAAGCCGCC
 ACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAAGAG
 GAAAAGAAGAAGCGGCUGCAGGCCUUCAGGGCUACCAGGUCACAAUGAAGAC
 CGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGC
 CCGAAGAGGUGGACGACGAGGUGUUCUACAGCCCCCGGUCCCUGGUGUCCCC
 GAGGCCGAGAACCGGAAGUGGACCAUUAUGGCCGUGAUGGUGUCCCUGCUGAC
 CGACUACUCCCCCAGCUGCAGAAGCCCAAGUUCUAGAUAAGUGAACUCGAGC
 UAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAA
 UGGAGUCUCUAAGCUACAUAUAACCAACUUAACAUUACAAAUGUUGUCCCCC
 AAAAUGUAGCCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUU
 CUAG

>mARM1804 (SEQ ID NO: 78)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGGGCGUCUUAACCUGCGGAUC
CUGCUGAACAACGCCGCCUUCGGGAACGGCCACAACUUAUGGUCCGCAACUU
CAGAUGCGGGCCAGCCCCUGCAGAACAGGGUGCAGCUGAAGGGCCGGGACCUGC
UGACCCUGAAGAACUUCACCGGGCGAAGAGAUCAGGUACAUGCUGUGGCUGAGC
GCCGACCUGAAGUUCGGGAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCA
AGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGGAGCACCCGGACCCGGCUGA
GCACCGAGACAGGCUUUGCCCUGCUGGGAGGGCCACCCUGCUUUCUGACCACC
CAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCAGAGUGCU
GAGCAGCAUGGCCGACGCCGUGCUGGCCCGGGUGUACAAGCAGAGCGACCUGG
ACACCCUGGCCAAAGAGGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUG
UACCACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGAACACUACAG
CUCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGC
ACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCAUCUGCAGGCCGCCACC
CCCAAGGGCUACGAGCCUGAUGCCAGCGUGACCAAGCUGGCCGAGCAGUACGC
CAAAGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAAGCCGCC
ACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAAGAG
GAAAAGAAGAAGCGGCUGCAGGCCUUCAGGGCUACCAGGUCACAAUGAAGAC
CGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGC
CCGAAGAGGUGGACGACGAGGUGUUCUACAGCCCCCGGUCCUGGUGUUCUCC
GAGGCCGAGAACCAGGAAGUGGACCAUUAUGGCCGUGAUGGUGUCCUGCUGAC
CGACUACUCCCCCAGCUGCAGAAGCCCAAGUUCUAGAUAAGUGAACUCGAGC
UAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAA
UGGAGUCUCUAAGCUACAUAUAACCAACUUAACACUUAACAAAUGUUGUCCCC
AAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUI
CUAG

>mARM1805 (SEQ ID NO: 79)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGGUCUUAACCUGCGGAUC
CUGCUGAACAACGCCGCCUUCGGGAACGGCCACAACUUAUGGUCCGCAACUU
CAGAUGCGGGCCAGCCCCUGCAGAACAGGGUGCAGCUGAAGGGCCGGGACCUGC
UGACCCUGAAGAACUUCACCGGGCGAAGAGAUCAGGUACAUGCUGUGGCUGAGC
GCCGACCUGAAGUUCGGGAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCA
AGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGGAGCACCCGGACCCGGCUGA
GCACCGAGACAGGCUUUGCCCUGCUGGGAGGGCCACCCUGCUUUCUGACCACC
CAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCAGAGUGCU
GAGCAGCAUGGCCGACGCCGUGCUGGCCCGGGUGUACAAGCAGAGCGACCUGG
ACACCCUGGCCAAAGAGGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUG
UACCACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGAACACUACAG
CUCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGC
ACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCAUCUGCAGGCCGCCACC
CCCAAGGGCUACGAGCCUGAUGCCAGCGUGACCAAGCUGGCCGAGCAGUACGC
CAAAGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAAGCCGCC
ACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAAGAG
GAAAAGAAGAAGCGGCUGCAGGCCUUCAGGGCUACCAGGUCACAAUGAAGAC
CGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGC
CCGAAGAGGUGGACGACGAGGUGUUCUACAGCCCCCGGUCCUGGUGUUCUCC

GAGGCCGAGAACCGGAAGUGGACCAUUAUGGCCGUGAUGGUGUCCCUGCUGAC
CGACUACUCCCCCAGCUGCAGAAGCCCAAGUUCUAGAUAAAGUGAACUCGAGC
UAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAA
UGGAGUCUCUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCC
AAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUU
CUAG

>mARM1806 (SEQ ID NO: 80)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUAACCUGAGGAUCCUG
CUGAACAAACGCAGCUUUCAGGAACGGCCACAACUUCAUGGUGAGGAACUUCCG
GUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCAGGGACCUGCUGA
CCCUGAAGAACUUCACCGGAGAGGAGAUCAAGUACAUGCUGUGGCUGAGCGCA
GACCUGAAGUUCAGGAUCAAGCAGAAGGGAGAGUACCUGCCCCUGCUGCAGGG
GAAGUCCCUGGGCAUGAUCUUCGAGAAGAGGAGUACCAGGACCAGGCUGAGC
ACCGAAACCGGCUUCGCCCUGCUGGGAGGACACCCCUGCUUCCUGACCACCCA
GGACAUCCACCUGGGCGUGAACGAGAGUCUGACCGACACCGCCAGGGUGCUGU
CUAGCAUGGCCGACGCCGUGCUGGCCAGGGUGUACAAGCAGUCAGACCUGGAC
ACCCUGGCUAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUA
CCACCCCAUCCAGAUCCUGGCUGACUACCUGACCCUGCAGGAGCACUACAGCU
CUCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGGAACAACAUCCUGCAC
AGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCUACCCC
CAAGGGUUACGAGCCCGACGCCAGCGUGACCAAGCUGGCAGAGCAGUACGCCA
AGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCCGCCAC
GGAGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGACAGGAGGAGG
AGAAGAAGAAGCGGCUGCAGGCUUUCAGGGUUACCAGGUGACCAUGAAGAC
CGCCAAGGUGGCUGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCAGGAAGC
CCGAGGAGGUGGACGACGAGGUGUUCUACUCUCCCAGGAGCCUGGUGUCCCC
GAGGCCGAGAACAGGAAGUGGACCAUCAUGGCUGUGAUGGUGUCCCUGCUGA
CCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAAUAAGUGAACUCGAG
CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGA
AUGGAGUCUCUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCC
CAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU
UCUAG

>mARM1808 (SEQ ID NO: 81)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUUCUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUCAACCUGAGGAUCCUG CUGAACAACGCAGCUUUCAGGAACGGCCACAACUUCAUGGUGAGGAACUUCGG GUGCGGCCAGCCCCUGCAGAACAAAGGUGCAGCUGAAGGGCAGGGACCUGCUGA CCCUGAAGAACUUCACCGGAGAGGAGAUCAAGUACAUGCUGUGGCUGAGCGCA GACCUGAAGUUCAGGAUCAAGCAGAAGGGAGAGUACCUGCCCCUGCUGCAGGG GAAGUCCCUGGGCAUGAUCUUCGAGAAGAGGAGUACCAGGACCAGGCUGAGC ACCGAAACCGGCUUCGCCUGCUGGGAGGACACCCCUGCUUCCUGACGACCCA GGACAUCCACCUCGGAGUGAACGAAUCCCUCACCGAUACCGCCCCGGGUGUUAU CAAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC ACUCUGGCUAAGGAGGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUA CCACCCCAUCCAGAUCUGGCUGACUACCUGACCCUGCAGGAGCACUACAGCU

CUCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGGAACAACAUCCUGCAC UCCAUCAUGAUGUCCGCCGCGAAGUUCGGAAUGCAUCUGCAAGCCGCCACGCC AAAAGGAUACGAACCGGAUGCGCCCGUGACAAAGUUGGCGGAACAGUACGCU AAGGAGAACGGAACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCCGCCA CGGAGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGACAGGAGGAGG AGAAGAAGAAGCGGCUGCAGGCUUUCAGGGUUACCAGGUGACCAUGAAGAC CGCCAAGGUGGCUGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCAGGAAGC CCGAGGAGGUGGACGACGAGGUGUUCUACUCUCCCAGGAGCCUGGUGUUC CCC GAGGCCGAGAACAGGAAGUGGACCAUCAUGGCUGUGAUGGUGUCCCUGCUGA CCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAAUAAGUGAACUCGAG CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGA AUGGAGUCUCUAAGCUACAUAUAACCAACUUAACACUUAACAAAUGUUGUCCCC CAAAUGUAGCCAUUCGUUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU UCUAG

>mARM1809 (SEQ ID NO: 82)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUCAACCUGCGCAUCCUG
CUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGCGCAACUUCGG
CUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGGCCGCGACCUGCUGA
CCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAGCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCAGGG
CAAGAGCCUGGGCAUGAUUCUUCGAGAAGCGCAGCACCCGCAACCCGCCUGAGCA
CCGAGACAGGCCUUCGCCUGCUGGGGCGGCCACCCUGCUUCCUGACCACCCAG
GACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCGUGCUGAG
CAGCAUGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGAGCGACCUGGACA
CCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGGCCUGAGCGACCUGUAC
CACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGAGCACUACAGCAG
CCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCUGCACA
GCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCACCCCC
AAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGGCCGAGCAGUACGCCAA
GGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCCACG
GCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGGCCAGGAGGAGGAG
AAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGACCGC
CAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCCG
AGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUCGCCGAG
GCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACCGA
CUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAAUAAGUGAACUCGAGCUAG
UGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGG
AGUCUCUAAGCUACAUAUAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAA
AUGUAGCCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUA
GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

>mARM1816 (SEQ ID NO: 83)

AGGAAACUUAAGUCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUC
AAGCAUUCUACUUCUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAG
CAAUUUUCUGAAAAUUUUCACCAUUUACGAACGAUAGCCAUGCUGUUCAACCU
GCGCAUCCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGC

GCAACUUCCGCUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCCGC
 GACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUG
 GCUGAGCGCCGACCUGAAGUUCCGCAUCAAGCAGAAGGGCGAGUACCUGCCCC
 UGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACC
 CGCCUGAGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCCUGCUUCCU
 GACCACCCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCC
 GCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGAGC
 GACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAG
 CGACCUGUACCACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGAGC
 ACUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAAC
 AUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGC
 CGCCACCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGC
 AGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAG
 GCCGCCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCA
 UGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCC
 CGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGU
 GUUCCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCC
 UGCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGACUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUAACCAACUUACACUUACAAAAUGUUGUCCCCCAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1822 (SEQ ID NO: 84)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAACUUGAGAAUCCUG
CUGAACAAACGCCGCCUUUCGCAACGGUCACAAUUUUAUGGUCAGAAACUUCAG
AUGCGGACAGCCCCUCCAAAACAAGGUCCAGCUGAAGGGCCGCGAUCUCCUCA
CCCUGAAGAACUUCACGGGGGAGGAGAUCAAGUACAUGCUGUGGCUCUCCGCU
GACCUGAAGUUCAGGAUCAAGCAGAAGGGAGAAUAUCUGCCGCUGCUGCAAG
GGAAGUCCCUGGGGAUGAUUUUCGAGAAGCGGAGCACCCGGACUCGGCUCUCC
ACUGAAACUGGUUUUCGCCCUUCUGGGCGGUACCCCUUGCUUCCUGACCACUCA
AGACAUUCACCUCGGAGUGAACGAGUCCUUGACUGACACCGCCCGGGUGCUGU
CGAGCAUGGCAGACGCCGUGCUAGCCCGCGUGUACAAGCAGUCAGACCUCGAU
ACCCUGGCCAAGGAGGCUUCGAUCCCGAUCAUCAACGGGUUGUCCGACCUGUA
CCACCCGAUUCAGAUUUCUGCCGACUACCUCACCCUGCAAGAGCAUUAACAGCU
CCCUGAAGGGGCUUACCCUGUCCUGGAUUGGCGACGGAAACAACAUCCUGCAC
UCCAUUAUGAUGUCGGCGGCCAAGUUCGGCAUGCACCUCCAAGCCGCGACCCC
UAAGGGUUAACGAACCAGACGCGUCAGUGACUAAGCUGGCCGAACAGUACGCAA
AGGAAAAUGGCACGAAGCUGCUCCUGACCAACGAUCCGUUGGAAGCCGCCAU
GGCGGAAAUGUGCUCAUCACCGACACCUUGGAUCUCGAUGGGACAGGAGGAAG
AGAAGAAGAAGCGGCUGCAGGCGUUCAGGGCUACCAGGUCACCAUGAAAACU
GCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUUCCGCGCAAGCC
UGAGGAGGUGGACGAUGAAGUGUUCUACUCUCCACGGUCCCUUGGUGUCCCCG
AGGCGGAGAACCGCAAAUGGACCAUCAUGGCUGUGAUGGUCAGCCUGCUGACC
GAUUACAGCCCUAGUUGCAAAAGCCGAAGUUUUGAUAAACUCGAGCUAGUGA
CUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGU
CUCUAAGCUACAUAUAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAUG
UAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1823 (SEQ ID NO: 85)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGUUCAACCUCCGCAUCCUC
CUCAACAACGCCGCAUUCAGAAACGGGCACAACUUCAUGGUCAGAAACUUCG
CUGCGGGCAACCCCUACAAAACAAGGUCCAGCUCAAGGGGCGGGACCUCUGA
CCCUGAAGAACUUCACCGGCGAAGAGAUCAAGUACAUGCUGUGGCUCUCCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGAGAGUACCUCCCGCUGCUGCAAGG
GAAGUCGCUGGGGAUGAUUCGAGAAGCGGUCAACCAGAACCCGGCUGUCAA
CCGAAACCGGGUUCGCACUGCUGGGGGGACACCCGUGCUUCCUGACCACCCAA
GACAUCCACCUGGGAGUGAACGAAUCGCUGACCGACACCGCCCCGCGUGCUGAG
CUCAAUGGCGGACGCCGUGCUGGCCCCGCGUGUACAAGCAGUCCGACCUGGACA
CCCUGGCCAAGGAAGCGUCCAUCCCGAUCAUCAACGGACUGUCCGACCUGUAC
CACCCGAUCCAGAUCUGGCAGACUACCUGACCCUGCAAGAACACUACAGCUC
CCUGAAGGGCCUGACCCUGUCAUGGAUCGGGGACGGGAACAACAUCCUGCACU
CCAUAUAUGAUGUCAGCCGCCAAGUUCGGAAUGCACCUCCAAGCCGCAACCCCG
AAGGGCUACGAACCGGACGCAUCAGUGACCAAACUGGCCGAGCAGUACGCCAA
GGAAAACGGCACCAAGCUCCUGCUGACCAACGACCCGCUGGAGGCCGCACACG
GGGGGAACGUGCUGAUCACCGACACCUGGAUCUCCAUGGGACAGGAGGAGGA
AAAGAAGAAGCGGCUGCAGGCGUCCAGGGGUACCAGGUCACCAUGAAAACCG
CGAAGGUCGCGGCAUCAGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGCCG
GAAGAGGUGGACGACGAGGUGUUCUACUCGCCGCGCUCGCUGGUGUCCCCGA
GGCGGAGAACAGGAAGUGGACCAUCAUGGCGGUGAUGGUCAGCCUCCUGACCG
ACUACUCGCCGCAGCUGCAGAAGCCGAAGUUCUGAUAACUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUAACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAUUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUCUUCACAUUCUAGAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

>mARM1840 (SEQ ID NO: 86)

CUCCCUCCCCCCCCCUAACGUUACUGGCCGAAGCCGCUUGGAAUAAGGCCGG
UGUGCGUUUGUCUAUAUGUUAUUUCCACCAUAUUGCCGUCUUUUGGCAAUG
UGAGGGCCCCGGAACCUUGGCCUGUCUUCUUGACGAGCAUUCUAGGGGUCUU
UCCCCUCUCGCCAAAGGAUGCAAGGUCUGUUGAAUGUCGUGAAGGAAGCAG
UUCCUCUGGAAGCUUCUUGAAGACAAACAACGUCUGUAGCGACCCUUGCAGG
CAGCGGAACCCCCCACCUGGCGACAGGUGCCUCUGCGGCCAAAAGCCACGUGU
AUAAGAUACACCUGCAAAGGCGGCACAACCCAGUGCCACGUUGUGAGUUGGA
UAGUUGUGGAAAGAGUCAAAUGGCUCUCCUCAAGCGUAUUAACAAGGGGCU
GAAGGAUGCCCAGAAGGUACCCCAUUGUAUGGGAUCUGAUCUGGGGCCUCGG
UGCACAUGCUUACGUGUGUUUAGUCGAGGUUAAAAACGUCUAGGCCCCCCG
AACCACGGGGACGUGGUUUUCCUUUGAAAAACACGAUGAUAAUAUGCUUUUC
AAUCUCCGCAUCCUCCUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCU
GGUCCGGAACUUCAGAUUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGG
GCCGGGAUCUUCUGACCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAU
GCUCUGGCUCUCCGCGGACUUGAAGUCCGCAUUAAGCAGAAGGGGGAAUACC
UCCGCGUCUUAAGGAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACC
AGGACCCGCCUUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUG
CUCCUGACGACCCAGGACAUCCACCUCGGAGUGAACGAAUCCUACCCGAUA

CCGCCCCGGGUGUUAUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAA
 CAGUCCGAUCUGGACACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGG
 CCUUAGUGACCUCUACCAUCCGAUUCAGAUCUGGCCGAUUACCUCACCCUGC
 AAGAACACUACAGCUCCCUGAAGGGGUCUGACAUUGUCCUGGAUCGGCGACGGC
 AACAAACAUUCUCCAUUCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCU
 UCAAGCCGCCACGCCGAAGGGUACGAGCCCGACGCUUCCGUGACUAAGCUCG
 CCGAGCAGUACGCUAAGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCA
 CUAGAAGCAGCCCACGGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCA
 GGGCCAGGAAGAAGAGAAAAAGAAGCGGCUGCAGGCGUCCAGGGGAUAUCAG
 GUCACCAUGAAAACCGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUG
 CCUGCCUCGCAAGCCUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGA
 GCCUCGUGUCCCCGAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUG
 GUGUCACUGCUCACCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUGAAU
 AAGUAGAUAGUGCAGUCACUGGCACAACGCGUUGCCCGGUAAGCCAAUCGGGU
 AUACACGGUCGUCAUACUGCAGACAGGGUUCUUCUACUUUGCAAGAUAGUCU
 AGAGUAGUAAAAUAAAUAGUAUAAGUCUAG

>mARM1841 (SEQ ID NO: 87)

CUCCCUCCCCCCCCCUAACGUUACUGGCCGAAGCCGCUUGGAAUAAGGCCGG
 UGUGCGUUUGUCUAUAUGUUAUUUCCACCAUAUUGCCGUCUUUUGGCAAUG
 UGAGGGCCCGGAAACCUGGCCUGUCUUCUUGACGAGCAUCCUAGGGGUCUU
 UCCCCUCUCGCCAAAGGAAUGCAAGGUCUGUUGAAUGUCGUGAAGGAAGCAG
 UUCCUCUGGAAGCUUCUUGAAGACAAACAACGUCUGUAGCGACCCUUUGCAGG
 CAGCGGAACCCCCACCUGGCCGACAGGUGCCUCUGCGGCCAAAAGCCACGUGU
 AUAAGAUACACCUGCAAAGGCCGGCACAACCCCAGUGCCACGUUGUGAGUUGGA
 UAGUUGUGGAAAGAGUCAAAUGGCUCUCCUCAAGCGUAUUCAACAAGGGGCU
 GAAGGAUGCCCAGAAGGUACCCCAUUGUAUGGGAUCUGAUCUGGGGGCCUCGG
 UGCACAUGCUUUACGUGUGUUUAGUCGAGGUUAAAAACGUCUAGGCCCCCCG
 AACCACGGGGACGUGGUUUUCCUUUGAAAAACACGAUGAUAAUAUGCUUUUC
 AACCUGAGAAUCCUCUUGAACA AUGCUGCUUUUCGGA AUGGCCACAACUUUAU
 GGUUCGGAACUCCGUUGCGGCCAGCCUUUACAAAACAAGGUCCAGCUGAAGG
 GCCGGGAUUUGCUCACACUAAAGAACUUUACUGGAGAAGAGAUCAAGUACAU
 GCUAUGGCUGUCGGCCGACCUGAAGUCCGUUAUCAAGCAGAAGGGAGAAUACC
 UUCCGCUGCUUCAAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACC
 AGGACCCGCCUUUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUG
 CUUCCUGACGACCCAGGACAUCACCUCGGAGUGAACGAUCCCUACCGAUA
 CCGCCCGGGUGUUAUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAA
 CAGUCCGAUCUCGAUACCUUGGCAAAGGAGGCUUCCA UUCCAUCAUCAACGG
 CCUGAGCGACCUGUACCACCCAAUCCAAAUCCUGGCUGACUACCUGACCCUGC
 AAGAGCACUACAGCAGCCUGAAGGGUCUGACCCUGUCAUGGAUUGGCGAUGG
 AAACAAUAUUCUGCACUCCAUCAUGAUGUCCGCCGCGAAGUUCGGA AUGCAUC
 UGCAAGCCGCCACUCCAAAAGGAUACGAACCGGAUGCAUCCGUGACCAAGUUG
 GCGGAACAGUACGCGAAGGAGAACGGAACCAAGCUCCUGCUGACUAACGACCC
 GCUCGAGGCUGCGCAUGGGGGUAACGUGCUGAUUACGGACACCUUGGAUCUCCA
 UGGGGCAGGAGGAAGAGAAGAAGAAGAGACUGCAGGCAUUC CAGGGGUACCA
 GGUCACCAUGAAAACCGCAAAAGUGGCAGCUUCGGACUGGACUUUCCUGCAUU
 GCCUGCCGAGGAAGCCGGAGGAAGUCGACGACGAAGUGUUCUACUCGCCUCGG
 UCCCUGGUGUUCCCCGAGGCCGAAAACCGGAAGUGGACCAUCAUGGCCGUGAU
 GGUGUCCUUGCUGACUGACUAUAGCCCCGACGUCGAGAAGCCUAAGUUCUGAA
 UAAGUAGAUAGUGCAGUCACUGGCACAACGCGUUGCCCGGUAAGCCAAUCGGG

UAUACACGGUCGUCAUACUGCAGACAGGGUUCUUCUACUUUGCAAGAUAGUC
 UAGAGUAGUAAAAUAAAUAGUAUAAGUCUAG

>mARM1842 (SEQ ID NO: 88)

CUCCCCUCCCCCCCCCUAACGUUACUGGCCGAAGCCGCUUGGAAUAAGGCCGG
UGUGCGUUUGUCUAUAUGUUAUUUCCACCAUAUUGCCGUCUUUUGGCAAUG
UGAGGGCCCCGGAACCUGGCCUGUCUUCUUGACGAGCAUUCCUAGGGGUCUU
UCCCCUCUCGCCAAAGGAAUGCAAGGUCUGUUGAAUGUCGUGAAGGAAGCAG
UUCCUCUGGAAGCUUCUUGAAGACAAACAACGUCUGUAGCGACCCUUUGCAGG
CAGCGGAACCCCCACCUGGCGACAGGUGCCUCUGCGGCCAAAAGCCACGUGU
AUAAGAUACACCUGCAAAGGCGGCACAACCCCAGUGCCACGUUGUGAGUUGGA
UAGUUGUGGAAAGAGUCAAAUGGCUCUCCUCAAGCGUAUUAACAAGGGGCU
GAAGGAUGCCCAGAAGGUACCCCAUUGUAUGGGAUCUGAUCUGGGGGCCUCGG
UGCACAUGCUUUACGUGUGUUUAGUCGAGGUUAAAAACGUCUAGGCCCCCCG
AACCACGGGGACGUGGUUUUCCUUUGAAAAACACGAUGAUAAUAGCUGUUC
AACCUGCGCAUCCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAU
GGUGCGCAACUUCGCGUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGG
GCCGCGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUG
CUGUGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCU
GCCCCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCC
GCACCCGCCUGAGCACCGAGACAGGCCUGGCCUGCUGGGCGGCCACCCUGC
UUCUGACCAACCAGGACAUCACCUGGGCGUGAACGAGAGCCUGACCGACAC
CGCCCGCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGC
AGAGCGACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGC
CUGAGCGACCUGUACCAACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCA
GGAGCACUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCA
ACAACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUG
CAGGCCGCCACCCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGC
CGAGCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCC
UGGAGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUG
GGCCAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGU
GACCAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCC
UGCCCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGC
CUGGUGUUCCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGU
GAGCCUGCUGACCGACUACAGCCCCAGCUGCAGAAGCCCAAGUUCUGAAUAA
GUAGAUAGUGCAGUCACUGGCACAACGCGUUGCCCGGUAAGCCAAUCGGGUAU
ACACGGUCGUAUACUGCAGACAGGGUUCUUCUACUUUGCAAGAUAGUCUAG
AGUAGUAAAAUAAAUAGUAUAAGUCUAG

>mARM1843 (SEQ ID NO: 89)

UCAACACAACAUUAUACAAAACAAACGAUUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAUCCUCACCGAUACCGCCCGGGUGUUUAU

CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
 ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUUAUCAACGGCCUAGUGACCUCUA
 CCAUCCGAUUCAGAUCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
 CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCA
 UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
 GAAGGGUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
 AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
 GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGGCCAGGAAGAAG
 AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGGAUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
 CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUCCCC
 GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUC
 CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUCUAG

>mARM1844 (SEQ ID NO: 90)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUCCGCAUCCUC
 CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
 AUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
 CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
 GACUUGAAGUCCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCUUCUUAAG
 GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
 ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
 GGACAUCCACCUCGGAGUGAACGAUCCUCACCGAUACCGCCCGGGUGUUAU
 CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
 ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUUAUCAACGGCCUAGUGACCUCUA
 CCAUCCGAUUCAGAUCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
 CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCA
 UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
 GAAGGGUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
 AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
 GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGGCCAGGAAGAAG
 AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGGAUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
 CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUCCCC
 GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUC
 CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUCUAG

>mARM1845 (SEQ ID NO: 91)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAAUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG

AUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUC AACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAAUCCCUACCCGAUACCGCCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAGUGACCUCUA
CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACA UUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCAU
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
AGAAAAAGAAAGCGGCUGCAGGCGUUC CAGGGAU AUCAGGUCACCAUGAAAAC
CGCCAAGGUCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUC CCC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUA AUACCAACUACACUUA CAAAAUGUUGUCCCCCAAAAUGUA
GCCAUUCGU AUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1846 (SEQ ID NO: 92)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUACCGCCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAGUGACCUCUA
CCAUCCGAUUCAGAUCUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCAU
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUCAAGCCGCCACGCC
GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGGAUAUCAGGUCACCAUGAAAAC
CGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCUCC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUACCAACUACACUUACAAAAUGUUGUCCCCCAAAUUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1847 (SEQ ID NO: 93)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUUUCAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUACCGCCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAUGUGACCUCUA
CCAUCCGAUUCAGAUCUCCUGGCCGAUUAACCUACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUCUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCAU
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCCAC
GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGGAUAUCAGGUCACCAUGAAAAC
CGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCGCC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUACCAACUUAACAUUACAAAUGUUGUCCCCCAAAUUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUCUAG

>mARM1882 (SEQ ID NO: 94)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGGUGUUAACCUCGCAUCCUC
CUCAACAACGCCGCAUUCAGAAACGGGCACAACUUCAUGGUCAGAAACUUCG
CUGCGGGCAACCCCUACAAAACAAGGUCCAGCUCUAGGGGGCGGGACCUCCUGA
CCCUGAAGAACUUCACCGGCGAAGAGAUCAAGUACAUGCUGUGGCUCUCCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGAGAGUACCUCCCGCUGCUGCAAGG
GAAGUCGCUGGGGGAUGAUCUUCGAGAAGCGGUCAACCAGAACCCGGCUGUCA
CCGAAACCGGGUUCGCACUGCUGGGGGGACACCCGUGCUUCCUGACCACCCAA
GACAUCCACCUGGGAGUGAACGAAUCGCUGACCGACACCGCCCGCGUGCUGAG
CUCAAUGGCGGACGCCGUGCUGGCCCGCGUGUACAAGCAGUCCGACCUGGACA
CCCUGGCCAAGGAAGCGUCCAUCCCGAUCAUCAACGGACUGUCCGACCUGUAC
CACCCGAUCCAGAUCUCCUGGCAGACUACCUGACCCUGCAAGAACACUACAGCUC
CCUGAAGGGCCUGACCCUGUCAUGGAUCGGGGACGGGAACAACAUCCUGCACU
CCAUAUAUGAUGUCAGCCGCCAAGUUCGGAAUGCACCUCCAAGCCGCAACCCCG
AAGGGCUACGAACCGGACGCAUCAGUGACCAAACUGGCCGAGCAGUACGCCAA
GGAAAACGGCACCAAGCUCCUGCUGACCAACGACCCGCGUGGAGGCCGACACG
GGGGGAACGUGCUGAUCACCGACACCUGGAUCUCCAUGGGACAGGAGGAGGA
AAAGAAGAAGCGGCUGCAGGCGUUCAGGGGUUACCAGGUCACCAUGAAAACCG
CGAAGGUCGCGGCAUCAGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGCCG

GAAGAGGUGGACGACGAGGUGUUCUACUCGCCGCGCUCGCUUGGUGUUCCTCGA
GGCGGAGAACAGGAAGUGGACCAUCAUGGCGGUGAUGGUCAGCCUCCUGACCG
ACUACUCGCCGACGUGCAGAAAGCCGAAGUUCUGAUAACUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUAACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1883 (SEQ ID NO: 95)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGGUGUUAACCUCCGCAUCCUC
CUCAACAACGCCGCAUUCAGAAACGGGCACAACUUAUGGUCAGAAACUUCG
CUGCGGGCAACCCCUACAAAACCGGGUCCAGCUCAAGGGGCGGGACCUCUGA
CCCUGAAGAACUUCACCGGCGAAGAGAUCAAGUACAUGCUGUGGCUCUCCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGAGAGUACCUCGCGUCUGCAAGG
GAAGUCGCUUGGGGAUGAUCUUCGAGAAGCGGUAACCAAGAACCCGGCUGUCA
CCGAAACCGGGUUCGCACUGCUGGGGGGACACCCGUGCUUCCUGACCACCCAA
GACAUCCACCUGGGAGUGAACGAUUCGCUAGCCGACACCGCCCGCGUGCUGAG
CUCAAUGGCGGACGCCGUGCUGGCCCCGCGUGUACAAGCAGUCCGACCUGGACA
CCCUGGCCAAGGAAGCGUCCAUCCCGAUCAUCAACGGACUGUCCGACCUGUAC
CACCCGAUCCAGAUCUGGCAGACUACCUGACCCUGCAAGAACACUACAGCUC
CCUGAAGGGCCUGACCCUGUCAUGGAUCGGGGACGGGAACAACAUCUGCACU
CCAUAUAUGAUGUCAGCCGCCAAGUUCGGAAUGCACCUCCAAGCCGCAACCCCG
AAGGGCUACGAACCGGACGCAUCAGUGACCAAACUGGCCGAGCAGUACGCCAA
GGAAAACGGCACCAAGCUCCUGCUGACCAACGACCCGCUUGGAGGCCGACACG
GGGGGAACGUGCUGAUCACCGACACCUUGGAUCUCCAUGGGACAGGAGGAGGA
AAAGAAGAAGCGGCUGCAGGCGUCCAGGGGUACCAGGUCACCAUGAAAACCG
CGAAGGUCGCGGCAUCAGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGCCG
GAAGAGGUGGACGACGAGGUGUUCUACUCGCCGCGCUCGCUUGGUGUUCCTCGA
GGCGGAGAACAGGAAGUGGACCAUCAUGGCGGUGAUGGUCAGCCUCCUGACCG
ACUACUCGCCGACGUGCAGAAAGCCGAAGUUCUGAUAACUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUAACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1884 (SEQ ID NO: 96)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGGUGUUAACCUCCGCAUCCUC
CUCAACAACGCCGCAUUCAGAAACGGGCACAACUUAUGGUCAGAAACUUCG
CUGCGGGCAACCCCUACAAAACCGGGUCCAGCUCAAGGGGCGGGACCUCUGA
CCCUGAAGAACUUCACCGGCGAAGAGAUCCGGUACAUGCUGUGGCUCUCCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGAGAGUACCUCGCGUCUGCAAGG
GAAGUCGCUUGGGGAUGAUCUUCGAGAAGCGGUAACCAAGAACCCGGCUGUCA
CCGAAACCGGGUUCGCACUGCUGGGGGGACACCCGUGCUUCCUGACCACCCAA
GACAUCCACCUGGGAGUGAACGAUUCGCUAGCCGACACCGCCCGCGUGCUGAG
CUCAAUGGCGGACGCCGUGCUGGCCCCGCGUGUACAAGCAGUCCGACCUGGACA
CCCUGGCCAAGGAAGCGUCCAUCCCGAUCAUCAACGGACUGUCCGACCUGUAC
CACCCGAUCCAGAUCUGGCAGACUACCUGACCCUGCAAGAACACUACAGCUC
CCUGAAGGGCCUGACCCUGUCAUGGAUCGGGGACGGGAACAACAUCUGCACU

CCAUAAUGAUGUCAGCCGCCAAGUUCGGAAUGCACCUCCAAGCCGCAACCCCG
 AAGGGCUACGAACCGGACGCAUCAGUGACCAAACUGGCCGAGCAGUACGCCAA
 GGAAAACGGCACCAAGCUCCUGCUGACCAACGACCCGCUGGAGGCCGCACACG
 GGGGGAACGUGCUGAUCACCGACACCUGGAUCUCCAUGGGACAGGAGGAGGA
 AAAGAAGAAGCGGCUGCAGGCGUUCAGGGGUACCAGGUCACCAUGAAAACCG
 CGAAGGUCGCGGCAUCAGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGCCG
 GAAGAGGUGGACGACGAGGUGUUCUACUCGCCGCGCUCGCUGGUGUUCCCCGA
 GGCGGAGAACAGGAAGUGGACCAUCAUGGCGGUGAUGGUCAGCCUCCUGACCG
 ACUACUCGCCGCAGCUGCAGAAGCCGAAGUUCUGAUAAACUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1885 (SEQ ID NO: 97)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGGUCAACCUCGCAUCCUC
 CUCAACAACGCCGCAUUCAGAAACGGGCACAACUUAUGGUCAGAAACUCCG
 CUGCGGGCAACCCCUACAAAACAAGGUCCAGCUCAAGGGGCGGGACCUCUGA
 CCCUGAAGAACUUCACCGGCGAAGAGAUCAAGUACAUGCUGUGGCUCUCCGCC
 GACCUGAAGUUCGCAUCAAGCAGAAGGGAGAGUACCUCGCGUGCUGCAAGG
 GAAGUCGCUGGGGAUGAUCUUCGAGAAGCGGUCAACCAGAACCCGGCUGUCA
 CCGAAACCGGGUUCGCACUGCUGGGGGGACACCCGUGCUUCCUGACCACCCAA
 GACAUCCACCUGGGAGUGAACGAUUCGCUGACCGACACCGCCCGCGUGCUGAG
 CUCAAUGGCGGACGCCGUGCUGGCCCGCGUGUACAAGCAGUCCGACCUGGACA
 CCCUGGCCAAGGAAGCGUCCAUCCCGAUCAUCAACGGACUGUCCGACCUGUAC
 CACCCGAUCCAGAUCUGGCGAGACUACCUGACCCUGCAAGAACACUACAGCUC
 CCUGAAGGGCCUGACCCUGUCAUGGAUCGGGGACGGGAACAACAUCCUGCACU
 CCAUAAUGAUGUCAGCCGCCAAGUUCGGAAUGCACCUCCAAGCCGCAACCCCG
 AAGGGCUACGAACCGGACGCAUCAGUGACCAAACUGGCCGAGCAGUACGCCAA
 GGAAAACGGCACCAAGCUCCUGCUGACCAACGACCCGCUGGAGGCCGCACACG
 GGGGGAACGUGCUGAUCACCGACACCUGGAUCUCCAUGGGACAGGAGGAGGA
 AAAGAAGAAGCGGCUGCAGGCGUUCAGGGGUACCAGGUCACCAUGAAAACCG
 CGAAGGUCGCGGCAUCAGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGCCG
 GAAGAGGUGGACGACGAGGUGUUCUACUCGCCGCGCUCGCUGGUGUUCCCCGA
 GGCGGAGAACAGGAAGUGGACCAUCAUGGCGGUGAUGGUCAGCCUCCUGACCG
 ACUACUCGCCGCAGCUGCAGAAGCCGAAGUUCUGAUAAACUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1886 (SEQ ID NO: 98)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGGUCAACCUCCGCAUCCUC
CUCAACAACGCCGCAUUCAGAAACGGGCACAACUUCAUGGUCAGAAACUUCCG
CUGCGGGCAACCCCUACAAAACCGGGUCCAGCUCAAGGGGGCGGGACCUCUGA
CCCUGAAGAACUUCACCGGCGAAGAGAUAAGUACAUGCUGUGGGCUCUCCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGAGAGUACCUCCCGCUGCUGCAAGG
GAAGUCGCUGGGGAUGAUCUUCGAGAAGCGGUCAACCAGAACCCGGCUGUCAA

CCGAAACCGGGUUCGCACUGCUGGGGGGACACCCGUGCUUCCUGACCACCCAA
GACAUCCACCUGGGAGUGAACGAAUCGCUGACCGACACCGCCCCGUGCUGAG
CUCAAUGGCGGACGCCGUGCUGGGCCCGCGUGUACAAGCAGUCCGACCUGGACA
CCCUGGCCAAGGAAGCGUCCAUCCCGAUCAUCAACGGACUGUCCGACCUGUAC
CACCCGAUCCAGAUCUGGCAGACUACCUGACCCUGCAAGAACACUACAGCUC
CCUGAAGGGCCUGACCCUGUCAUGGAUCGGGGACGGGAACAACAUCCUGCACU
CCAUAUAUGAUGUCAGCCGCCAAGUUCGGAAUGCACCUCCAAGCCGCAACCCCG
AAGGGCUACGAACCGGACGCAUCAGUGACCAAACUGGGCCGAGCAGUACGCCAA
GGAAAACGGCACCAAGCUCCUGCUGACCAACGACCCGCUGGAGGCCGCACACG
GGGGGAACGUGCUGAUCACCGACACCUGGAUCUCCAUGGGACAGGAGGAGGA
AAAGAAGAAGCGGCUGCAGGCGUUCAGGGGUACCAGGUCACCAUGAAAACCG
CGAAGGUCGCGGCAUCAGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGCCG
GAAGAGGUGGACGACGAGGUGUUCUACUCGCCGCGCUCGCUGGUGUCCCCGA
GGCGGAGAACAGGAAGUGGACCAUCAUGGCGGUGAUGGUCAGCCUCCUGACCG
ACUACUCGCCGCAGCUGCAGAAGCCGAAGUUCUGAUAACUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAAUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAGAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

>mARM1887 (SEQ ID NO: 99)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGGUCAACCUCCGCAUCCUC
CUCAACAACGCCGCAUUCAGAAACGGGCACAACUUCAUGGUCAGAAACUUCCG
CUGCGGGCAACCCCUACAAAACCGGGUCCAGCUCAAGGGGGCGGGACCUCUGA
CCCUGAAGAACUUCACCGGCGAAGAGAUCGCGUACAUGCUGUGGCUCUCCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGAGAGUACCUCGCGCUGCUGCAAGG
GAAGUCGCUGGGGAUGAUCUUCGAGAAGCGGUCAACCAGAACCCGGCUGUCA
CCGAAACCGGGUUCGCACUGCUGGGGGGACACCCGUGCUUCCUGACCACCCAA
GACAUCCACCUGGGAGUGAACGAAUCGCUGACCGACACCGCCCGCGUGCUGAG
CUCAAUGGCGGACGCCGUGCUGGCCCGCGUGUACAAGCAGUCCGACCUGGACA
CCCUGGCCAAGGAAGCGUCCAUCCCGAUCAUCAACGGACUGUCCGACCUGUAC
CACCCGAUCCAGAUCCUGGCAGACUACCUGACCCUGCAAGAACACUACAGCUC
CCUGAAGGGCCUGACCCUGUCAUGGAUCGGGGACGGGAACAACAUCCUGCACU
CCAUAUAUGAUGUCAGCCGCCAAGUUCGGAAUGCACCUCCAAGCCGCAACCCCG
AAGGGCUACGAACCGGACGCAUCAGUGACCAAACUGGCCGAGCAGUACGCCAA
GGAAAACGGCACCAAGCUCCUGCUGACCAACGACCCGCGUGGAGGCCGCACACG
GGGGGAACGUGCUGAUCACCGACACCUGGAUCUCCAUGGGACAGGAGGAGGA
AAAGAAGAAGCGGCUGCAGGCGUUCAGGGGUACCAGGUCACCAUGAAAACCG
CGAAGGUCGCGGCAUCAGACUGGACCUUCCUGCACUGCCUGCCCCGGAAGCCG
GAAGAGGUGGACGACGAGGUGUUCUACUCGCCGCGCUCGCUGGUGUUCGCCGA
GGCGGAGAACAGGAAGUGGACCAUCAUGGCGGUGAUGGUCAGCCUCCUGACCG
ACUACUCGCCGCAGCUGCAGAAGCCGAAGUUCUGAUAACUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUAACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
GCCAUUCGUUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1888 (SEQ ID NO: 100)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGGUCAACCUGCGCAUCCUG
CUGAACAAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGCGCAACUUCG
CUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGGCCGCGACCUGCUGA
CCCUGAAGAACUUCACCGGCGAGGAGAUAAGUACAUGCUGUGGCUGAGCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGGCGAGUACCUGCCCCUGCUGCAGGG
CAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUGAGCA
CCGAGACAGGCUUCGCCUGCUGGGCGGGCCACCCUGCUUCCUGACCACCCAG
GACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCCGCGUGCUGAG
CAGCAUGGCCGACGCCGUGCUGGGCCGCGUGUACAAGCAGAGCGACCUGGACA
CCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUAC
CACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGAGCACUACAGCAG
CCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGCACA
GCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCGGCCACCCCC
AAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACGCCAA
GGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCCACG
GCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGAGGAG
AAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGACCGC
CAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCCG
AGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUCGCCGAG
GCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACCGA
CUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAUAAACUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
UAAGCUACAUAUACCAACUUAACUUAACAAAUGUUGUCCCCCAAAUUGUAG
CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

>mARM1889 (SEQ ID NO: 101)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUGGUCAACCUGCGCAUCCUG
CUGAACAAACGCCGCCUUCGCAACGGGCCACAACUUCAUGGUGCGCAACUUCG
CUGCGGCCAGCCCCUGCAGAACC GGGUGCAGCUGAAGGGCCGCGACCUGCUGA
CCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAGCGCC
GACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCAGGG
CAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUGAGCA
CCGAGACAGGCUUCGCCUGCUGGGCGGCCACCCUGCUUCCUGACCACCCAG
GACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCGUGCUGAG
CAGCAUGGCCGACGCCGUGCUGGGCCGCGUGUACAAGCAGAGCGACCUGGACA
CCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUAC
CACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGAGCACUACAGCAG
CCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCUGCACA
GCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCACCCCC
AAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACGCCAA
GGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCCGCCACG
GCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGAGGAG
AAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGACCGC
CAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCCG

AGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUCCCCGAG
GCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACCGA
CUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAUAACUCGAGCUAGUGACUG
ACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCUC
UAAGCUACAUAUAACCAACUACACUACAAAAUGUUGUCCCCCAAAUAGUAG
CCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1890 (SEQ ID NO: 102)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUGUCAAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAAACAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAUCCUCACCGAUACCGCCCGGGUGUUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAGUGACCUCUA
CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCAU
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
GAAGGGUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUUAUCAGGUCACCAUGAAAAC
CGCCAAGGUCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCUCC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
CUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUGUA
GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1891 (SEQ ID NO: 103)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGCUUGUCAAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAAACCGGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAUCCUCACCGAUACCGCCCGGGUGUUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAGUGACCUCUA
CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCAU

UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUCAAGCCGCCACGCC
 GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
 AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
 GGGGGCAACGUGCUUAUUACUGACACCUUGGAUCUCCAUGGGCCAGGAAGAAG
 AGAAAAAGAAGCGGCUGCAGGCGUCCAGGGGAUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
 CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCGCC
 GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
 CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGACU
 GACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUCU
 CUAAGCUACAUAUAACCAACUACACUUAACAAAUGUUGUCCCCCAAAAUGUA
 GCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUCUAG

>mARM1898 (SEQ ID NO: 104)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAGCAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGGGCCUUGUCAAUUCUCCGCAUC
 CUCCUUAACAACGCCGCGUUUAGAAACGGCCACAACUUAUGGUCCGGAACUU
 CAGAUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCCGGGAUCUUC
 UGACCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCC
 GCGGACUUGAAGUUCGCAUUUAGCAGAAGGGGGAUACCUUCCGCUGCUUCA
 AGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUU
 CUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACC
 CAGGACAUCCACCUCGGAGUGAACGAAUCCCUCACCGAUACCGCCCCGGGUGUU
 AUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGG
 ACACUCUGGCCAAGGAGGCGUCAAUUCCUAUUUAUCAACGGCCUUAGUGACCUC
 UACCAUCCGAUUCAGAUCUGGCCGAUUACCUCACCCUGCAAGAACACUACAG
 CUCCUGAAGGGUCUGACAUUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCC
 AUUCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUCAAGCCGCCACG
 CCGAAGGGUUAACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGC
 UAAGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCC
 ACGGGGGCAACGUGCUUAUUACUGACACCUUGGAUCUCCAUGGGCCAGGAAGAA
 GAGAAAAAGAAGCGGCUGCAGGCGUCCAGGGGAUAUCAGGUCACCAUGAAAA
 CCGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAG
 CCUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCGCC
 CGAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUC
 ACCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGAC
 UGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUC
 UCUAAGCUACAUAUAACCAACUACACUUAACAAAUGUUGUCCCCCAAAAUGU
 AGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUCUAG

>mARM1899 (SEQ ID NO: 105)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGGGCCUUGUCAAUUCUCCGCAUC
CUCCUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUU
CAGAUGUGGCCAGCCGCUUCAAAACCGGGUCCAGCUGAAGGGCCGGGAUCUUC
UGACCCUGAAGAACUUUACUGGCCGAAGAGAUCAAGUACAUGCUCUGGCUCUCC
GCGGACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCUGCUUCA
AGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUU

CUACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCCUGCUUCCUGACGACC
CAGGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUACCGCCCCGGGUGUU
AUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGG
ACACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAUGUGACCUC
UACCAUCCGAUUCAGAUCUGGCCGAUUACCUCACCCUGCAAGAACACUACAG
CUCCCUGAAGGGUCUGACAUGUCCUGGAUCGGCGACGGCAACAACAUUCUCC
AUUCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACG
CCGAAGGGUUAACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGC
UAAGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCC
ACGGGGGCAACGUGCUUAUUAACUGACACCUGGAUCUCCAUGGGCCAGGAAGAA
GAGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUUAUCAGGUCACCAUGAAAA
CCGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAG
CCUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCCC
CGAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUC
ACCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUGAC
UGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAGUC
UCUAAGCUACAUAAUACCAACUACACUUAACAAAUGUUGUCCCCCAAAUUGU
AGCCAUUCGUUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAA

>mARM1900 (SEQ ID NO: 106)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCACCAUGGGCGGACUUGUCAAUUCUCCGC
AUCCUCCUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAA
CUUCAGAUUGUGGCCAGCCGCUUCAAACAAGGUCCAGCUGAAGGGCCGGGAUC
UUCUGACCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUC
UCCGCGGACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCUGCU
UCAAGGAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCC
UUUCUACUGAAACUGGGUUCGCGCUGCUCGGUGGGCCACCCUGCUUCCUGACG
ACCCAGGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUACCGCCCCGGGU
GUUAUCGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUC
UGGACACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAUGUGAC
CUCUACCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUA
CAGCUCCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUC
UCCAUUCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCC
ACGCCGAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUA
CGCUAAGGAGAACGGAACCAAGCUUCUGCUGACUAAACGACCCACUAGAAGCAG
CCCACGGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAA
GAAGAGAAAAAGAAGCGGCUGCAGGCGUCCAGGGAUUAUCAGGUCACCAUGA
AAACCGCCAAGGUUCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGC
AAGCCUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUU
CCCCGAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGC
UCACCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGCUCGAGCUAGUG
ACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGAAUGGAG
UCUCUAAGCUACAUAUAUACCAACUUAACACUUAACAAAUGUUGUCCCCCAAAU
GUAGCCAUUCGUUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAUUCUAG

>mARM1903 (SEQ ID NO: 107)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCAUGGCCCCUUUCAAUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAGGCAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAUCCCUACCGAUACCGCCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGGCGUCAAUUCCUAUUAUCAACGGCCUUAGUGACCUCUA
CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCA
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
GAAGGGUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
GGGGGCAACGUGCUUAUACUGACACCUUGGAUCUCCAUGGGGCCAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUUAUCAGGUCACCAUGAAAAC
CGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCUCC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGAUAAGUGAACUCGAG
CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGA
AUGGAGUCUCUAAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCC
CAAAAUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU
UCUAG

>mARM1904 (SEQ ID NO: 108)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCAUGGCCUUUUCAAUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAGGCCGGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAUCCUCACCGAUACCGCCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAGUGACCUCUA
CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCA
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUCAAGCCGCCACGCC
GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUCCAGGGAUUUCAGGUCACCAUGAAAAC
CGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC

CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGAUAAAGUGAACUCGAG
CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGA
AUGGAGUCUCUAAGCUACAUAUACCAACUUAACACUUAACAAAUGUUGUCCCC
CAAAAUGUAGCCAUUCGUUUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU
UCUAG

>mARM1905 (SEQ ID NO: 109)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCAUGGCCCUUUUCAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAGGCCGGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAUCCUCACCGAUACCGCCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUUAUCAACGGCCUUAGUGACCUCUA
CCAUCCGAUUCAGAUCCUGGCCGAUUAACCUACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCA
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGGAUAUCAGGUCACCAUGAAAAC
CGCCAAGGUCGCGUCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCUCC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGAUAAAGUGAACUCGAG
CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGA
AUGGAGUCUCUAAGCUACAUAUAACCAACUUAACACUUAACAAAUGUUGUCCCC
CAAAAUAGUAGCCAUUCGUUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU
UCUAG

>mARM1906 (SEQ ID NO: 110)

UCAACACAACAUUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAGCAAAGCAAUUUUCUGAAA
AUUUUCACCAUUUACGAACGAUAGCCAUGGCCCUUGUCAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAGGCAGGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCAACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAUCCUCACCGAUACCGCCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUUAUCAACGGCCUUAGUGACCUCUA

CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
 CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCA
 UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
 GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
 AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCCAC
 GGGGGCAACGUGCUUAUUAACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
 AGAAAAAGAAGCGGCUGCAGGCGUCCAGGGGAUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
 CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCGCC
 GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
 CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAAGUGAAUAGACUCGAG
 CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGA
 AUGGAGUCUCUAAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCC
 CAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU
 UCUAG

>mARM1907 (SEQ ID NO: 111)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACU
 CUAUUGCAGCAAUUUAAAUCAUUCUUAUAAAGCAAAAGCAAUUUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCAUGGCCCUUUUCAUUCGCCAUCCUC
 CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCUUGGUCCGGAACUUCAG
 AUGUGGCCAGCCGCUUCAAGUCAAGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
 CCCUGAAGAACUUUACUGGCCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
 GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUCUUAAG
 GAAAGAGCCUCGGCAUGAUCUUGAGAAGCGCUCUACCAAGGACCCGCCUUCU
 ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUUGCUUCCUGACGACCA
 GGACAUCCACCUCGGAGUGAACGAAUCCCUACCCGAUACCGCCCGGGUGUUAU
 CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
 ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUAGUGACCUCUA
 CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
 CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCA
 UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
 GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
 AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCCAC
 GGGGGCAACGUGCUUAUUAACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
 AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGGAUAUCAGGUCACCAUGAAAAC
 CGCCAAGGUCGCUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
 CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCGCC
 GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
 CCGACUACAGCCCGCAGCUUCAGAAGCCCAAGUUCUAGAUAAAGUGAACUCGAG
 CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGA
 AUGGAGUCUCUAAGCUACAUAUACCAACUACACUACAAAAUGUUGUCCCC
 CAAAAUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU
 UCUAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAA

>mARM1908 (SEQ ID NO: 112)

UCAACACAACAUAUACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
CUAUUGCAGCAAUUUAAAUCAUUUCUUUUAAAAGCAAAAGCAAUUUUCUGAAA

AUUUUCACCAUUUACGAACGAUAGCCAUGGCCCUUUUCAAUUCUCCGCAUCCUC
CUUAACAACGCCGCGUUUAGAAACGGCCACAACUUCAUGGUCCGGAACUUCAG
AUGUGGCCAGCCGCUUCAAGUCAGGGUCCAGCUGAAGGGCCGGGAUCUUCUGA
CCCUGAAGAACUUUACUGGCGAAGAGAUCAAGUACAUGCUCUGGCUCUCCGCG
GACUUGAAGUUCGCAUUAAGCAGAAGGGGGAAUACCUUCCGCGUGCUUCAAG
GAAAGAGCCUCGGCAUGAUCUUUGAGAAGCGCUCACACCAGGACCCGCCUUUCU
ACUGAAACUGGGUUCGCGCUGCUCGGUGGCCACCCUGCUUCCUGACGACCCA
GGACAUCCACCUCGGAGUGAACGAAUCCUCACCGAUACCGCCCGGGUGUUAU
CGAGCAUGGCAGAUGCCGUGCUGGCCAGGGUGUACAAACAGUCCGAUCUGGAC
ACUCUGGCCAAGGAGGCGUCAAUUCCUAUUAUCAACGGCCUUAGUGACCUCUA
CCAUCCGAUUCAGAUCCUGGCCGAUUACCUCACCCUGCAAGAACACUACAGCU
CCCUGAAGGGUCUGACAUUGUCCUGGAUCGGCGACGGCAACAACAUUCUCCAU
UCCAUCAUGAUGUCCGCCGCAAAAUUCGGCAUGCAUCUUAAGCCGCCACGCC
GAAGGGUUACGAGCCCGACGCUUCCGUGACUAAGCUCGCCGAGCAGUACGCUA
AGGAGAACGGAACCAAGCUUCUGCUGACUAACGACCCACUAGAAGCAGCCAC
GGGGGCAACGUGCUUAUUACUGACACCUGGAUCUCCAUGGGCCAGGAAGAAG
AGAAAAAGAAGCGGCUGCAGGCGUUCAGGGAUUACAGGUCACCAUGAAAAC
CGCCAAGGUCGUGCCUCCGACUGGACCUUCCUGCACUGCCUGCCUCGCAAGC
CUGAAGAAGUGGACGACGAGGUGUUCUACUCGCCACGGAGCCUCGUGUUCUCC
GAGGCCGAGAAUAGAAAGUGGACCAUCAUGGCCGUGAUGGUGUCACUGCUCA
CCGACUACAGCCCGCAGCUUCAGAAAGCCCAAGUUCUAGAUAAAGUGAACUCGAG
CUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACACCCGA
AUGGAGUCUCUAAGCUACAUAUACCAACUACACUUAACAAAUGUUGUCCCC
CAAAUUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUCACAU
UCUAG

>mARM1915 (SEQ ID NO: 113)

GGCAGAAAAAUUUGCUACAUUGUUUCACAAACUUCAAAUAUUUAUUCAUUUUAU
 UUAGAUCUAUUUAUUACAUCAAAACAAAAAGCCGCCACCAUGCUGUUCAACCUG
 CGCAUCCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGCG
 CAACUUCGCGUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCCGCG
 ACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUAAGUACAUGCUGUGG
 CUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCU
 GCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCC
 GCCUGAGCACCGAGACAGGCUUCGCCUGCUGGGCGGCCACCCUGCUUCCUG
 ACCACCCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCG
 CGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGAGCG
 ACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGC
 GACCUGUACCACCCCAUCCAGAUCUUGGCCGACUACCUGACCCUGCAGGAGCA
 CUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACA
 UCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCC
 GCCACCCCCAAGGGCUACGAGCCCCGACGCCAGCGUGACCAAGCUGGCCGAGCA
 GUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGG
 CCGCCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAG
 GAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAU
 GAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCC
 GCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUG
 UCCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCU
 GCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGU
 AAUGAGCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGGCCUCCCAACGG

GCCCUCCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAG
 UGGGCAUCUAG

>mARM1916 (SEQ ID NO: 114)

GGCAGAAAAAUUUGCUACAUUGUUUCACAAACUUCAAAUAUUAUUCAUUUUAU
UUAGAUCUAUUUAUACAUCAAAACAAAAAGCCGCCACCAUGGGAGUAUUCAA
CCUGCGCAUCCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGG
UGCGCAACUUCGCGUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGC
CGCGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCU
GUGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGC
CCCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGC
ACCCGCCUGAGCACCGAGACAGGCUUCGCCUUGCUGGGCGGCCACCCUUGC
CCUGACCACCCAGGACAUCACCUGGGCGUGAACGAGAGCCUGACCGACACCG
CCCGCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAG
AGCGACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCU
GAGCGACCUGUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGG
AGCACUACAGCAGCCUGAAGGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAAC
AACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCA
GGCCGCCACCCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCG
AGCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUG
GAGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGG
CCAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGA
CCAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUG
CCCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCU
GGUGUUCCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGA
GCCUGCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUC
UAGUAAUGAGCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCCGCUGGGCCUCCCA
ACGGGCCCUCCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUC
UGAGUGGGCAUCUAG

>mARM1917 (SEQ ID NO: 115)

GGCAGAAAAAUUUGCUACAUUGUUUCACAAACUUCAAAUAUUAUUCAUUUUAU
UUAGAUCUAUUUAUACAUCAAAACAAAAAGCCGCCACCAUGGGAGUAUUCAA
CCUGCGCAUCCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGG
UGCGCAACUUCGCGUGCGGCCAGCCCCUGCAGAACCAGGUGCAGCUGAAGGGC
CGCGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCCGGUACAUGCU
GUGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGC
CCCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGC
ACCCGCCUGAGCACCGAGACAGGCUUCGCCUUGCUGGGCGGCCACCCUUGC
CCUGACCACCCAGGACAUCACCUGGGCGUGAACGAGAGCCUGACCGACACCG
CCCGCGUGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAG
AGCGACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCU
GAGCGACCUGUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGG
AGCACUACAGCAGCCUGAAGGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAAC
AACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCA
GGCCGCCACCCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCG
AGCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUG
GAGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGG
CCAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGA
CCAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUG

CCCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCU
GGUGUUCCCCGAGGCCGAGAACC GCAAGUGGACCAUCAUGGCCGUGAUGGUGA
GCCUGCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUC
UAGUAAUGAGCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGGCCUCCCA
ACGGGCCCUCCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUC
UGAGUGGGCAUCUAG

>mARM1918 (SEQ ID NO: 116)

GGCAGAAAAAUUUGCUACAUAUUGUUUCACAAACUUCAAAUAUUAUUCAUUUAU
UUAGAUCUAUUUAUACAUCAAAACAAAAAGCCGCCACCAUGCUGGUAUUAAC
CUGCGCAUCCUGCUGAACAACGCCGCCUUCCGCAACGGCCACAACUUAUGGU
GCGCAACUUCGCGUGCGGGCCAGCCCCUGCAGAACC GGGUGCAGCUGAAGGGCC
GCGACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCCGGUACAUGCUG
UGGCUGAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCC
CCUGCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCA
CCCGCCUGAGCACCGAGACAGGCUUCGCCCUGCUGGGCGGGCCACCCUGCUUC
CUGACCACCCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGC
CCGCGUGCUGAGCAGCAUGGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGA
GCGACCUGGACACCCUGGCCAAGGAGGGCCAGCAUCCCCAUCAUACACGGCCUG
AGCGACCUGUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGA
GCACUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACA
ACAUCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAG
GCCGCCACCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGA
GCAGUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGG
AGGCCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGC
CAGGAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGAC
CAUGAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGC
CCCGCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUG
GUGUUCCCCGAGGCCGAGAACC GCAAGUGGACCAUCAUGGCCGUGAUGGUGAG
CCUGCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCU
AGUAAUGAGCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGGCCUCCCA
CGGGCCCUCCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCU
GAGUGGGCAUCUAG

>mARM1919 (SEQ ID NO: 117)

AUUUUUACAUCAAAAACAAAAAGCCGCCACCAUGGGAGUAUUCAACCUGCGCAU
 CCUGCUGAACAAACGCCGCCUUCCGCAACGGCCACAACUUCAUGGUGCGCAACU
 UCCGCUGCGGCCAGCCCCUGCAGAACAAAGGUGCAGCUGAAGGGCCGCGACCUG
 CUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAG
 CGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGC
 AGGGCAAGAGCCUGGGCAUGAUUCUUCGAGAAGCGCAGCACCCGCACCCGCCUG
 AGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCUUGCUUCCUGACCAC
 CCAGGACAUCACCCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCCGCGUGC
 UGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGAGCGACCUG
 GACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCU
 GUACCACCCCAUCCAGAUCUGGCCGACUACCUGACCCUGCAGGAGCACUACA
 GCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCUG
 CACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCAC
 CCCCAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACG
 CCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCCGCC

CACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGA
 GGAGAAGAAGAAGCGCCUGCAGGCCUCCAGGGCUACCAGGUGACCAUGAAGA
 CCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAG
 CCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUC
 CGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGA
 CCGACUACAGCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAAUGA
 GCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGCCCU
 CCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUGAAUAAAGUCUGAGUGGGC
 AUCUAG

>mARM1920 (SEQ ID NO: 118)

AUUUAUUACAUCAAAACAAAAAGCCGCCACCAUGGGAGUAUUCAACCUGCGCAU
 CCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGCGCAACU
 UCCGCUGCGGCCAGCCCCUGCAGAACCGGGUGCAGCUGAAGGGCCGCGACCUG
 CUGACCCUGAAGAACUUCACCGGCGAGGAGAUCGGUACAUGCUGUGGCUGAG
 CGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGC
 AGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUG
 AGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCUGCUUCCUGACCAC
 CCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCGUGC
 UGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGAGCGACCUG
 GACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCU
 GUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACUACA
 GCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCUG
 CACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCAC
 CCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACG
 CCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCCGCC
 CACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGA
 GGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGA
 CCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAG
 CCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUCCCC
 CGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGA
 CCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAAUGA
 GCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGCCCCU
 CCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGC
 AUCUAG

>mARM1921 (SEQ ID NO: 119)

AUUUAUUACAUCAAAACAAAAAGCCGCCACCAUGCUGGUUAUUCAACCUGCGCAU
 CCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGCGCAACU
 UCCGCUGCGGCCAGCCCCUGCAGAACCGGGUGCAGCUGAAGGGCCGCGACCUG
 CUGACCCUGAAGAACUUCACCGGCGAGGAGAUCGGUACAUGCUGUGGCUGAG
 CGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGC
 AGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUG
 AGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCUGCUUCCUGACCAC
 CCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCGUGC
 UGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGAGCGACCUG
 GACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCU
 GUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACUACA
 GCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCUG
 CACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCAC

CCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGGCCGAGCAGUACG
 CCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCC
 CACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGGCCAGGAGGA
 GGAGAAGAAGAAGCGCCUGCAGGGCCUCCAGGGCUACCAGGUGACCAUGAAGA
 CCGCCAAGGUGGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAG
 CCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUCCC
 CGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGA
 CCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAAUGA
 GCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCCGCUGGGGCCUCCCAACGGGGCCCU
 CCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGC
 AUCUAG

>mARM1925 (SEQ ID NO: 120)

UCAACACAACAUUAACAAAACAAACGAAUCUCAAGCAAUCAAGCAUUCUACUU
 CUAUUGCAGCAAUUUAAAUCAUUUCUUUAAAAGCAAUUAUUCUGAAA
 AUUUUCACCAUUUACGAACGAUAGCCACCAUGUUGUUAACUUGAGGAUCUU
 GUUGAACAACGCCGCCUUCAGGAACGGACACAACUUCAUUGGUAAGGAACUUA
 GGUGCGGACAGCCCUUGCAGAACAAAGUACAGUUGAAAGGAAGGGACUUGUU
 GACAUUGAAAAACUUCACAGGAGAAGAAAUCAAUACAUGUUGUGGUUGUCG
 GCCGACUUGAAAUUCAGGAUCAAAACAGAAAGGAGAAUACUUGCCCUUGUUGC
 AGGGAAAAUCGUUGGGAAUGAUCUUCGAAAAAAGGUCGACAAGGACAAGGUU
 GUCGACAGAAACAGGAUUCGCCUUGUUGGGAGGACACCCCUGCUUCUUGACAA
 CACAGGACAUCCACUUGGGAGUAAACGAAUCGUUGACAGACACAGCCAGGGUA
 UUGUCGUCGAUGGCCGACGCCGUAUUGGCCAGGGUAUACAAACAGUCGGACUU
 GGACACAUUGGCCAAAGAAGCCUCGAUCCCCAUCAUCAACGGAUUGUCGGACU
 UGUACCACCCCAUCCAGAUUCUUGGCCGACUACUUGACAUUGCAGGAACACUAC
 UCGUCGUUGAAAGGAUUGACAUUGUCGUGGAUCGGAGACGGAAACAACAUUC
 UGCACUCGAUCAUGAUGUCGGCCGCCAAAUUCGGAAUGCACUUGCAGGCCGCC
 ACACCCAAAGGAUACGAACCCGACGCCUCGGUAACAAAAUUGGCCGAACAGUA
 CGCCAAAGAAAAACGGAACAAAAUUGUUGUUGACAAACGACCCCUUGGAAGCCG
 CCCACGGAGGAAACGUUUUGAUCACAGACACAUGGAUCUCGAUGGGACAGGA
 AGAAGAAAAAAGGUUGCAGGCCUCCAGGGAUACCAGGUAACAAUG
 AAAACAGCCAAAGUAGCCGCCUCGGACUGGACAUCUUGCACUGCUUGCCCAG
 GAAACCCGAAGAAGUAGACGACGAAGUAUUCUACUCGCCCAGGUCGUUGGUA
 UUCCCCGAAGCCGAAAACAGGAAAUGGACAAUCAUGGCCGUAUUGGUAUCGU
 UGUUGACAGACUACUCGCCCCAGUUGCAGAAACCCAAAUUCUGAAUAGUGAAC
 UCGAGCUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGCCUCAAGAACA
 CCCGAAUGGAGUCUCUAAGCUACAUAUACCAACUACACUACAAAAUGUUG
 UCCCCCAAAUUGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGUUUCUUC
 ACAUUCUAG

>mARM1926 (SEQ ID NO: 121)

AUUUUUACAUCAAAACAAAAAGCCGCCACCAUGCUGUUAACCUUGCGCAUCCU
 GCUGAACAAACGCCGCCUCCGCAACGGCCACAACUUCAUUGGUGCGCAACUUC
 GCUGCGGCCAGCCCCUGCAGGGCAAGGUGCAGCUGAAGGGCCGCGACCUGCUG
 ACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAGCGC
 CGACCUGAAGUUCGCAUCAAGCAGAAAGGGCGAGUACCUGCCCCUGCUGCAGG
 GCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUGAGC
 ACCGAGACAGGCUUCGCCUGCUGGGCGGCCACCCUGCUUCCUGACCACCCA
 GGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCGUGCUGA

GCAGCAUGGCCGACGCCGUGCUGGCCCCGCGUGUACAAGCAGAGCGACCUGGAC
 ACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAAACGGCCUGAGCGACCUGUA
 CCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACUACAGCA
 GCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGCAC
 AGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCACCCC
 CAAGGGCUACGAGCCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACGCCA
 AGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCCAC
 GGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUAGGGCCAGGAGGAGGA
 GAAGAAGAAGCGCCUGCAGGCCUCCAGGGCUACCAGGUGACCAUGAAGACCG
 CCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCC
 GAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUCCCCGA
 GGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACCG
 ACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAAUGAGCU
 GGAGCCUCGGUAGCCGUUCCUCCUGCCCCGCUGGGCCUCCCAACGGGGCCUCCU
 CCCCUCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAUC
 UAG

>mARM1927 (SEQ ID NO: 122)

AUUAAUACAUCAAAACAAAAAGCCGCCACCAUGCUGUUAACCUGCGCAUCCU
 GCUGAACAACGCCGCCUUCCGCAACGGCCACAACUUAUGGUGCGCAACUCC
 GCUGCGGCCAGCCCCUGCAGGGCCGGGUGCAGCUGAAGGGCCGCGACCUGCUG
 ACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAGCGC
 CGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCAGG
 GCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCAACCCGCCUGAGC
 ACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCUUGCUUCCUGACCACCCA
 GGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCCGCGUGCUGA
 GCAGCAUGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGAGCGACCUGGAC
 ACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUA
 CCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACUACAGCA
 GCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGCAC
 AGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCACCCC
 CAAGGGCUACGAGCCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACGCCA
 AGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCCAC
 GGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUAGGGCCAGGAGGAGGA
 GAAGAAGAAGCGCCUGCAGGCCUCCAGGGCUACCAGGUGACCAUGAAGACCG
 CCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCC
 GAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUCCCCGA
 GGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACCG
 ACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAAUGAGCU
 GGAGCCUCGGUAGCCGUUCCUCCUGCCCCGCUGGGCCUCCCAACGGGGCCUCCU
 CCCCUCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAUC
 UAG

>mARM1928 (SEQ ID NO: 123)

AUUAUUACAUCAAAACAAAAAGCCGCCACCAUGCUGUUCAACCUGCGCAUCCU
GCUGAACAAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGCGCAACUUC
GCUGCGGCCAGCCCCUGCAGGGCCGGGUGCAGCUGAAGGGCCGCGACCUGCUG
ACCCUGAAGAACUUCACCGGCCGAGGAGAUCGCGUACAUGCUGUGGCUGAGCGC
CGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGCAGG
GCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUGAGC

ACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCUUGC UUCCUGACCACCCA
GGACAUCCACCUGGGCGUGAACGAGAGCCUGACCAGACCCGCCCGCGUGCUGA
GCAGCAUGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGAGCGACCUGGAC
ACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCUGUA
CCACCCCAUCCAGAUCCUGGCCGACUACCUGACCUGCAGGAGCACUACAGCA
GCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUGCAC
AGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGGCCGCCACCC
CAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACGCCA
AGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGGCCGCCAC
GGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGAGGAGGA
GAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGACCG
CCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCCC
GAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUCGCCGA
GGCCGAGAACCGCAAGUGGACCAUCAUGGCCCGUGAUGGUGAGCCUGCUGACCG
ACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAAUGAGCU
GGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGCCCUCCU
CCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAUC
UAG

>mARM1929 (SEQ ID NO: 124)

GGCAGAAAAAUUUGCUACAUGUUUCACAAACUUCAAAUAUUAUUCAUUUUAU
 UUAGAUCUAUUUAUACAUCAAAACAAAAAGCCGCCACCAUGCUGUUCAACCG
 CGCAUCCUGCUGAACAACGCCGCCUUCCGCAACGGCCACAACUUCAUGGUGCG
 CAACUUCCGCUGCGGCCAGCCCCUGCAGGGCAAGGUGCAGCUGAAGGGCCGCG
 ACCUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGG
 CUGAGCGCCGACCUGAAGUUCCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCU
 GCUGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCC
 GCCUGAGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCUGCUUCCUG
 ACCACCCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCG
 CGUGCUGAGCAGCAUGGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGAGCG
 ACCUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGC
 GACCUGUACCACCCCAUCCAGAUCUCCUGGCCGACUACCUGACCCUGCAGGAGCA
 CUACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACA
 UCCUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCC
 GCCACCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCA
 GUACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGG
 CCGCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAG
 GAGGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAU
 GAAGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCC
 GCAAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUG
 UUCCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCU
 GCUGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGU
 AAUGAGCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGG
 GCCUCCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAG
 UGGGCAUCUAG

AT1G67090> (SEQ ID NO: 125)

CACAAAGAGUAAAGAAGAACA

AT1G35720> (SEQ ID NO: 126)

AACACUAAAAGUAGAAGAAAA

AT5G45900> (SEQ ID NO: 127)

CUCAGAAAGAUAGAUCAGCC

>pARM563 (SEQ ID NO: 128)

ATGCTGTTTAATCTGAGGATCCTGTAAACAATGCAGCTTTTAGAAATGGTCACA
 ACTTCATGGTTCGAAATTTTCGGTGTGGACAACCACTACAAAATAAAGTGCAGCT
 GAAGGGCCGTGACCTTCTCACTCTAAAAAACTTTACCGGAGAAGAAATTAATA
 TATGCTATGGCTATCAGCAGATCTGAAATTTAGGATAAAACAGAAAGGAGAGTA
 TTTGCCTTTATTGCAAGGGAAGTCCTTAGGCATGATTTTTGAGAAAAGAAGTACT
 CGAACAAGATTGTCTACAGAAACAGGCTTTGCACTTCTGGGAGGACATCCTTGTT
 TTCTTACCACACAAGATATTCATTTGGGTGTGAATGAAAGTCTCACGGACACGGC
 CCGTGTATTGTCTAGCATGGCAGATGCAGTATTGGCTCGAGTGTATAAACAATCA
 GATTTGGACACCCTGGCTAAAGAAGCATCCATCCCAATTATCAATGGGCTGTCAG
 ATTTGTACCATCCTATCCAGATCCTGGCTGATTACCTCACGCTCCAGGAACACTA
 TAGCTCTCTGAAAGGTCTTACCCTCAGCTGGATCGGGGATGGGAACAATATCCTG
 CACTCCATCATGATGAGCGCAGCGAAATTCGGAATGCACCTTCAGGCAGCTACTC
 CAAAGGGTTATGAGCCGGATGCTAGTGTAACCAAGTTGGCAGAGCAGTATGCCA
 AAGAGAATGGTACCAAGCTGTTGCTGACAAATGATCCATTGGAAGCAGCGCATG
 GAGGCAATGTATTAATTACAGACACTTGGATAAGCATGGGACAAGAAGAGGAGA
 AGAAAAAGCGGCTCCAGGCTTTCCAAGGTTACCAGGTTACAATGAAGACTGCTA
 AAGTTGCTGCCTCTGACTGGACATTTTTACTGCTTGCCAGAAAGCCAGAAGA
 AGTGGATGATGAAGTCTTTTATTCTCCTCGATCACTAGTGTTCCAGAGGCAGAA
 AACAGAAAGTGGACAATCATGGCTGTCATGGTGTCCCTGCTGACAGATTACTCAC
 CTCAGCTCCAGAAGCCTAAATTTTGA

>pARM564 (SEQ ID NO: 129)

ATGCTCTTTAATCTGCGCATCTTACTGAACAACGCCGCATTCCGGAACGGTCACA
 ACTTCATGGTCCGCAATTTCCGCTGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGACGGGATCTGCTGACACTGAAGAACTTCACCGGAGAAGAGATCAAGTA
 CATGCTGTGGCTCAGCGCAGACTTGAAGTTCGGGATCAAGCAGAAGGGAGAATA
 CTTGCCCCTGCTGCAAGGAAAGTCGCTGGGAATGATTTTTGAGAAGCGGTCAACT
 CGCACCAGACTCTCCACCGAAACTGGTTTCGCACTGCTTGGCGGGCACCCTTGCT
 TCCTGACGACTCAGGACATCCACCTCGGCGTGAACGAATCGCTAACCGATAACCG
 CCAGAGTGCTTTCTTCCATGGCCGACGCGGTGCTGGCCAGGGTGTACAAGCAGTC
 CGACCTCGATACCTTGGCAAAGGAGGCTTCCATTCCCATCATCAACGGCCTGAGC
 GACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCCTGCAAGAGCACT
 ACAGCAGCCTGAAGGGTCTGACCCTGTCATGGATTGGCGATGGAAACAATATTC
 TGCCTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCCAC
 TCCAAAAGGATACGAACCGGATGCGTCCGTGACCAAGTTGGCGGAACAGTACGC
 GAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCCTCGAGGCTGCGCA
 TGGGGGCAACGTGCTGATTACCGACACCTGGATCTCCATGGGGCAGGAGGAAGA
 GAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAACCG
 CAAAAGTGGCAGCTTCGGACTGGACTTTCTGCAATTGCCTGCCGAGGAAGCCGG
 AGGAAGTCGACGACGAAGTGTTCTACTCGCCTCGGTCCCTGGTGTTCCCCGAGGC
 CGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCCCTTGCTGACTGACTAT
 AGCCCGCAGCTGCAGAAGCCTAAGTTCTAG

>pARM565 (SEQ ID NO: 130)

ATGCTGTTTAACTACGTATTTTGCTCAACAATGCAGCCTTTAGAAACGGACATA
 ACTTTATGGTTCGAACTTTCGCTGCGGGCAGCCACTGCAGAACAAGGTCCAGCT
 GAAAGGGAGAGATTTGCTCACGCTGAAGAACTTTACTGGCGAAGAAATCAAGTA
 TATGCTGTGGTTGTCCGCGGACCTCAAGTTTCGGATTAAGCAGAAAGGGGAGTAT
 CTGCCACTGCTGCAAGGAAAGAGCCTCGGCATGATCTTCGAGAAGCGGAGCACT
 CGGACCAGGCTGAGTACCGAAACTGGCTTCGCATTGTTGGGTGGACATCCATGTT
 TTCTGACAACGCAGGACATTCATCTGGGCGTGAACGAGAGTCTGACGGACACAG
 CTCGCGTTCTGTCTCTATGGCTGATGCGGTGTTGGCCCCGGGTCTATAAGCAGTC
 CGATTTGGACACCTTGGCTAAGGAAGCTAGCATACCGATTATCAATGGGCTGTCC
 GACCTGTATCACCTATTCAAATCCTGGCCGACTACCTCACACTGCAAGAACACT
 ATAGCTCATTGAAGGGACTGACCCTGAGCTGGATAGGGGACGGAAACAACATCC
 TACATAGCATTATGATGTCCGCTGCCAAGTTTGGCATGCATCTTCAAGCCGCCAC
 GCCAAAGGGTTATGAGCCCGACGCGTCAGTGACAAAGCTGGCCGAGCAGTACGC
 TAAGGAGAATGGTACCAAATTACTGCTGACTAATGATCCACTGGAGGCTGCACA
 TGGCGGCAATGTACTGATCACCGACACATGGATCTCGATGGGCCAGGAGGAAGA
 AAAGAAGAAGAGGCTTCAGGCCTTCCAAGGCTACCAGGTCACCATGAAAACAGC
 TAAGGTTGCAGCATCTGATTGGACCTTTCTGCACTGTCTGCCAAGGAAGCCCGAA
 GAGGTGGACGATGAAGTATTCTATAGCCACGAGTTTGGTGTTCCCTGAGGCTG
 AAAATAGGAAGTGGACAATTATGGCCGTAATGGTGTCCCTGTTAACCGACTACTC
 TCCGCAACTGCAGAAACCTAAGTTTTAG

>pARM566 (SEQ ID NO: 131)

ATGCTGTTTAACTTAAGGATCCTGCTGAACAACGCCGCTTTTCGTAACGGTCATA
 ACTTTATGGTCCGGAACCTTAGATGTGGCCAGCCGCTGCAGAACAAGGTTTCAGCT
 GAAGGGGAGGGATCTGCTGACCTTGAAGAACTTTACCGGCGAAGAGATCAAGTA
 CATGTTGTGGCTGAGCGCCGATCTGAAGTTTAGGATTAAGCAGAAGGGGGAGTA
 TTTGCCACTGCTGCAAGGAAAATCCTTGGGGATGATCTTCGAGAAGCGCTCCACT
 AGAACCCGGCTAAGCACAGAAACCGGCTTCGCACTTCTGGGTGGACATCCCTGTT
 TTCTGACGACGCAGGATATACACCTGGGCGTGAATGAGAGTCTGACGGACACAG
 CTAGGGTGTTGAGCAGCATGGCCGATGCAGTACTGGCCCGCGTTTATAAGCAGA
 GCGACTTGGACACACTGGCCAAGGAAGCGTCAATTCCGATTATCAATGGGCTGT
 CAGACCTGTATCATCCCATTCAAATCTTGGCTGACTATCTGACCCTGCAAGAACA
 TTACAGCTCCCTGAAGGGCCTCACGTTGTCCTGGATTGGCGACGGAAACAACATT
 CTGCATTCGATCATGATGAGCGCTGCTAAGTTTGGCATGCACCTCCAAGCCGCTA
 CACCTAAGGGATATGAGCCTGATGCCAGCGTAACCAAGCTGGCCGAACAGTACG
 CGAAGGAGAATGGCACGAAACTGCTGTTGACAAATGACCCACTGGAGGCAGCTC
 ACGGTGGCAACGTGCTGATCACCGACACGTGGATATCTATGGGACAGGAAGAAG
 AGAAGAAGAAGCGGCTGCAGGCATTCCAAGGGTATCAGGTCACCATGAAAACG
 GCCAAGGTTGCTGCATCCGACTGGACATTTCTGCATTGCTTGCCCCGCAAACCAG
 AAGAAGTAGACGACGAAGTCTTTTATTCCCCACGGTCGCTGGTGTTCCCCGAGGC
 GGAGAATCGAAAGTGGACGATTATGGCCGTGATGGTGTCCCTGCTGACTGATTA
 CTCTCCCCAACTGCAAAAGCCTAAGTTTTAG

>pARM567 (SEQ ID NO: 132)

ATGCTTTTCAACCTGAGGATCCTCCTGAACAACGCCGCCTTTCGCAATGGTCACA
 ACTTTATGGTCCGGAACCTTCAGATGCGGCCAGCCGCTGCAGAACAAGGTCCAGC
 TGAAGGGACGGGATCTGCTGACTCTGAAGAACTTCACCGGAGAAGAGATCAAGT
 ACATGCTGTGGCTGTCGGCCGACCTGAAGTTCAGGATCAAGCAGAAGGGAGAAT
 ACCTCCCGCTGCTGCAAGGAAAGTCCCTGGGCATGATTTTCGAGAAGCGCTCGAC

CAGAACTCGGTTGTCCACCGAAACCGGGTTTGCGCTGCTGGGCGGACATCCTTGC
 TTCCTGACGACTCAGGATATTACCTGGGAGTGAACGAGTCGCTGACCGACACC
 GCCAGAGTGCTGAGCTCGATGGCCGACGCCGTGTTGGCACGCGTGTACAAGCAG
 TCCGATCTGGATACCCTGGCCAAAGAAGCTTCCATCCCGATCATTAAACGGGCTGA
 GCGACCTCTACCACCCCATTCAAATCCTGGCCGACTACCTGACTCTGCAAGAACA
 CTACAGCTCGCTGAAGGGGTTGACTCTGTCCTGGATCGGCGACGGAAACAACAT
 CCTGCACTCCATCATGATGTCGGCCGCAAAGTTCGGCATGCATTTGCAAGCCGCC
 ACCCCAAAGGGCTACGAACCAGACGCGAGCGTCACCAAGCTGGCCGAACAGTAC
 GCGAAGGAAAATGGTACTAAGCTGCTGCTGACCAACGACCCATTGGAAGCTGCC
 CATGGTGGAACGTGCTGATCACCGACACCTGGATCTCGATGGGCCAGGAAGAG
 GAGAAGAAGAAGCGGCTGCAGGCGTTCCAGGGGTATCAGGTCACCATGAAAAC
 AGCCAAAGTGGCAGCGTCAGACTGGACCTTCCTCCACTGTCTGCCTCGCAAGCCA
 GAGGAGGTGGACGACGAGGTGTTCTACTCCCCTCGGTCCCTCGTGTTCCCTGAGG
 CTGAGAACCGGAAGTGGACCATTATGGCCGTGATGGTGTCACCTCCTGACTGATTA
 CTCCCCGCAACTGCAGAAGCCCAAGTTCTAG

>pARM568 (SEQ ID NO: 133)

ATGCTGTTTAACTGAGGATCCTATTGAACAATGCTGCTTTTCGTAATGGCCATA
 ACTTTATGGTTCGGAACCTTTAGATGCGGGCAGCCACTGCAGAACAAGGTCCAGTT
 GAAAGGCCGCGATCTGTTGACATTGAAGAACTTTACCGGCGAAGAGATTAAGTA
 TATGCTGTGGCTGTCTGCTGACCTCAAGTTTTCGAATCAAGCAGAAGGGCGAATAT
 CTCCCCCTGCTGCAAGGAAAGTCTCTCGGCATGATCTTTGAGAAGCGGAGTACCC
 GAACACGGCTGAGCACCGAAACGGGCTTCGCACTGCTGGGGGGGCCATCCCTGTT
 TTCTGACAACGCAGGACATCCACTTGGGGGTAAACGAATCATTGACTGATACCGC
 CCGCGTACTGTCATCCATGGCCGACGCTGTGCTGGCTAGGGTGTACAAGCAGTCA
 GATCTGGATACTGGCCAAGGAAGCTAGCATACCAATCATCAATGGACTGAGT
 GACCTTTATACCCGATTCAAATACTAGCCGATTATCTGACCCTGCAAGAGCATT
 ACTCCTCGCTGAAAGGCCCTACGCTGTCCTGGATCGGCGACGGCAACAACATTCT
 GCATAGTATTATGATGTCTGCTGCCAAATTCGGCATGCATCTGCAAGCTGCTACG
 CCGAAGGGTTATGAACCCGACGCGTCAGTTACGAAGCTCGCTGAGCAGTATGCA
 AAGGAGAATGGCACAAAGCTGTTGCTTACCAACGATCCCCTGGAAGCTGCTCAT
 GCGGCAATGTGCTGATTACTGACACCTGGATTTCAATGGGCCAGGAGGAGGAG
 AAGAAGAAGAGGTTACAGGCTTTTCAAGGTTACCAAGTCACGATGAAAACCGCT
 AAGGTCGCAGCCAGCGACTGGACATTCTGCACTGTCTGCCAAGAAAGCCGGAA
 GAAGTGGACGACGAGGTGTTCTATTCCCGCGGTCTTTGGTGTTCGCGAGGCCG
 AAAACAGGAAATGGACCATTATGGCCGTGATGGTATCGTTGCTGACGGACTACA
 GCCCTCAGTTGCAAAAGCCCAAGTTCTAG

>pARM569 (SEQ ID NO: 134)

ATGCTCTTTAACTCCGCATCCTCCTCAACAACGCCGCCTTCCGGAATGGGCATA
 ACTTCATGGTCCGGAACCTTCAGATGCGGCCAGCCCTGCAAAACAAGGTCCAGTT
 GAAGGGACGGGACCTCCTTACGCTGAAGAACTTTACCGGAGAAGAGATTAAGTA
 CATGCTGTGGTTGTCCGCTGACCTCAAGTTCCGCATTAAGCAGAAGGGAGAATAT
 CTGCCGCTGCTGCAAGGAAAGAGCCTGGGCATGATCTTCGAAAAGCGCTCCACT
 AGAACCCGGCTGTCGACTGAGACTGGATTTCGCCTTGCTCGGTGGACACCCGTGCT
 TCCTGACGACCCAGGACATCCACCTGGGAGTGAACGAGTCACTTACGGATACCG
 CGAGGGTGCTGTCTCAATGGCCGACGCAGTGCTCGCGCGCGTGTACAAGCAGT
 CAGATCTGGATACCTGGCCAAGGAAGCCAGCATTCCCATCATCAACGGACTGA
 GCGACCTTTACCACCCAATCCAGATCCTCGCCGACTACTTAACCCTGCAAGAGCA
 CTACAGCTCCCTGAAGGGACTGACTCTGTCTGGATCGGGGATGGAAACAACAT
 CCTGCACTCCATCATGATGTCTGCCGCTAAGTTTGGGATGCATCTGCAAGCCGCA

ACCCCTAAGGGATACGAGCCCGACGCCTCGGTGACCAAACCTTGCGGAACAGTAC
 GCCAAGGAAAACGGTACCAAGCTGCTGCTGACCAACGACCCTCTGGAAGCGGCC
 CACGGAGGAAATGTGCTGATTACCGACACCTGGATTTTCGATGGGCCAGGAGGAG
 GAGAAGAAGAAGAGACTGCAGGCGTTCCAGGGATATCAGGTCAACCATGAAAAC
 CGCCAAGGTCGCCGCCAGCGACTGGACCTTCCTGCACTGTCTCCCTCGGAAACCG
 GAAGAAGTGGATGACGAGGTGTTCTACTCCCCGCGCTCGCTGGTGTTCGCGGAG
 GCTGAAAACAGGAAGTGGACAATCATGGCCGTGATGGTGTCCCTGTTGACCGAC
 TACTCCCCACAACCTGCAGAAGCCCAAGTTCTAG

>pARM570 (SEQ ID NO: 135)

ATGCTTTTCAATCTGCGCATCCTCCTGAACAACGCCGCCTTCCGCAATGGACACA
 ACTTTATGGTCCGCAACTTCCGCTGTGGGCAGCCGCTGCAGAACAAGGTCCAGCT
 CAAGGGGAGAGATCTCCTGACCCTGAAGAACTTCACTGGAGAGGAGATCAAGTA
 CATGCTGTGGCTGTCCGCCGACCTGAAATTTTCGGATTAAGCAGAAGGGCGAATA
 CCTCCCCTGCTGCAAGGAAAGTCTTTGGGCATGATCTTCGAAAAGAGAAGCAC
 CCGGACCCGGTTGAGCACCGAACTGGGTTCGCGCTCCTCGGTGGACACCCGTG
 CTTCTGACCACCCAAGATATTCATCTGGGTGTCAACGAAAGCCTGACCGACACC
 GCCAGGGTGTGTATCCATGGCTGACGCAGTGCTCGCCCGGGTGTACAAGCAG
 TCAGACCTGGACACCCCTCGCCAAGGAAGCTTCGATCCCTATCATCAACGGACTTT
 CCGACCTGTACCACCCCATCCAAATTCTGGCCGACTACCTGACTCTGCAAGAACA
 CTATAGCTCGCTGAAAGGACTTACTCTGTCTTGGATCGGGGACGGCAACAACATT
 CTCCATTCCATCATGATGTCCGCTGCCAAGTTCGGAATGCACCTTCAAGCAGCGA
 CTCCCAAGGGATACGAACCTGATGCCTCCGTGACTAAGCTGGCAGAGCAGTACG
 CCAAGGAGAACGGTACAAAGCTGCTGCTCACGAACGACCCCCCTGGAGGCGGCC
 ACGGCGGAACGTGCTGATTACCGATACCTGGATCTCAATGGGCCAGGAAGAGG
 AGAAGAAGAAGCGGCTCCAGGCGTTTCAAGGCTACCAGGTCACCATGAAAACCG
 CGAAGGTCGCCGCCTCCGACTGGACTTTCTTGCACTGCCTGCCGCGGAAGCCCGA
 GGAAGTGGATGACGAAGTGTTCTACTCGCCGAGATCGTTGGTGTTCCTGAGGCC
 GAAAACAGGAAGTGGACCATCATGGCCGTGATGGTGTCCCTGCTGACTGATTAC
 AGCCACAGCTGCAGAAGCCTAAGTTCTAG

>pARM571 (SEQ ID NO: 136)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTCAGATGTGGCCAGCCGCTTCAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCGCGATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTGAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGAAAATTTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTCGCTGCCTCCGACTGGACCTTCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCTCGAGGCC

GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM572 (SEQ ID NO: 137)

ATGCTTTTCAACCTGAGAATCCTCCTGAACAACGCCGCTTCCGCAATGGTCATA
 ACTTCATGGTCCGCAACTTTCGCTGCGGACAGCCTCTCCAAAACAAGGTCCAGCT
 CAAGGGGCGCGACCTCCTCACACTGAAGAACTTCACTGGAGAAGAAATCAAGTA
 CATGCTGTGGCTGAGCGCCGATCTGAAGTTCGGGATCAAGCAGAAGGGAGAGTA
 CCTTCCTCTGCTGCAAGGGAAGTCCTTGGGAATGATTTTCGAGAAGCGGTCCACC
 CGGACCAGGCTGAGCACTGAAACTGGCTTCGCCCTGCTGGGAGGCCACCCTTGTT
 TCCTGACCACTCAGGACATCCACCTGGGCGTGAACGAGTCCCTGACCGATACTGC
 CAGAGTGCTGTCTCCATGGCCGACGCCGTGCTCGCCCGGGTGTACAAGCAGTCA
 GACCTCGATACGCTGGCCAAGGAAGCCTCCATTCCCATTATCAATGGTCTGTCTGG
 ACCTCTACCATCCAATCCAAATCCTCGCCGACTACCTGACTCTGCAAGAACAATA
 CAGCTCACTCAAGGGCCTCACCTCTCCTGGATCGGCGACGGAAACAACATCCTT
 CACTCGATTATGATGTCGGCCGCGAAGTTCGGGATGCACCTCCAAGCTGCCACTC
 CAAAAGGCTACGAGCCGGATGCCTCAGTGACTAAGTTGGCGGAACAGTATGCGA
 AGGAGAACGGTACCAAGCTCCTGCTGACTAACGACCCGCTGGAGGGCCGCCACG
 GGGGAAACGTGCTCATCACCGATACTTGGATTTCCATGGGACAGGAGGAAGAGA
 AGAAGAAGCGGTTGCAGGCATTTCAAGGGCTACCAGGTCACCATGAAAACCTGCCA
 AAGTCGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCAAGCCTGAAG
 AAGTGGACGACGAGGTGTTCTACTCTCCCCGGTCCCTCGTGTTCCCTGAGGCCGA
 AAACAGGAAGTGGACCATCATGGCTGTGATGGTGTCCCTCCTGACCGACTACAG
 CCCTCAGCTCCAAAAACCCAAGTTTTAG

>pARM573 (SEQ ID NO: 138)

ATGCTTTTCAACCTGAGAATCCTCTTGAACAATGCTGCTTTTCGGAATGGCCACA
 ACTTTATGGTTCGGAACCTTCCGTTGCGGCCAGCCTTTACAAAACAAGGTCCAGCT
 GAAGGGGCCGGGATTTGCTCACACTGAAGAACTTTACTGGGGAGGAGATTAAGTA
 TATGCTGTGGCTGTCCGCTGACCTGAAGTTTAGGATCAAGCAGAAGGGCGAATA
 TCTGCCGCTGCTGCAAGGGAAAAGTCTGGGCATGATTTTTTGAAAAGCGCTCTACC
 CGGACCAGACTGTCTACGGAAACAGGCTTTGCCCTGCTGGGCGGCCACCCCTGTT
 TTCTGACAACGCAGGACATCCATCTGGGCGTGAACGAATCACTGACCGATACTG
 CTCGGGTACTCAGTTCTATGGCTGACGCAGTGCTGGCTAGGGTGTACAAGCAGA
 GCGACTTGGACACACTGGCTAAGGAGGCCAGCATCCCCATTATCAATGGCCTGTC
 TGATTTGTACCATCCCATTCAAATCCTGGCTGATTATCTGACACTACAAGAGCAT
 TACTCAAGTCTGAAGGGTTTGACTCTCTCCTGGATCGGCGACGGCAACAACATTT
 TACATTCCATTATGATGAGTGCTGCTAAGTTTGGCATGCATTTGCAAGCTGCTAC
 CCCAAAGGGCTATGAACCTGACGCTAGCGTAACCAAGTTGGCCGAACAGTATGC
 TAAAGAGAATGGCACCAAGCTGCTCCTGACGAATGACCCCCTGGAAGCTGCTCA
 TGGCGGAAACGTACTTATAACTGATACATGGATTAGCATGGGCCAGGAAGAGGA
 GAAGAAGAAGAGACTGCAGGCCTTCCAAGGCTATCAGGTCACCATGAAAACCTGC
 CAAGGTTGCAGCTAGCGACTGGACCTTCCTGCACTGTTTGCCGAGGAAACCCGA
 GGAGGTGGACGATGAAGTCTTTTATTCTCCCCGCTCCTTGGTGTTCCTGAGGCT
 GAAAATCGAAAGTGGACGATAATGGCAGTGATGGTGTCCCTACTGACCGACTAT
 TCTCCACAACCTGCAGAAGCCTAAATTCTAG

>pARM574 (SEQ ID NO: 139)

ATGCTTTTCAATCTGAGGATCCTGCTGAACAACGCTGCTTTTCGCAACGGTCATA
 ACTTTATGGTTCGCAATTTTCGTTGTGGCCAGCCGCTGCAGAACAAAGGTTACAGCT

GAAGGGCAGAGATCTGCTGACTCTGAAGAACTTCACTGGGGAAGAAATCAAGTA
TATGTTATGGCTGTCCGCGGATCTGAAATTTTGAATCAAGCAGAAGGGCGAATAT
CTTCCCCTGCTGCAAGGGAAATCCTTGGGCATGATTTTTGAGAAGAGGAGCACTA
GGACTAGATTGTCAACAGAAACAGGCTTTGCTTTGTTGGGCGGACATCCCTGCTT
TCTGACGACACAGGATATCCACCTCGGCGTAAACGAGTCCCTCACCGACACTGCT
AGGGTACTGAGCAGCATGGCCGACGCTGTGCTAGCCCGGGTTTACAAGCAGTCA
GACCTGGACACCCTTGCCAAGGAAGCTTCTATTCCAATTATCAACGGCCTGAGTG
ACCTGTATCACCTATTCAAATACTCGCCGACTATTTGACGCTTCAAGAACATTA
CAGCAGCCTCAAGGGCTTAACCTTGAGTTGGATAGGCGACGGCAACAATATCCT
GCATTCCATTATGATGTCTGCCGCTAAGTTTGGCATGCATCTACAAGCCGCAACA
CCCAAGGGCTATGAACCCGACGCTAGCGTGACCAAGCTGGCCGAGCAGTATGCT
AAGGAAAATGGCACAAAGCTCCTTCTTACCAACGATCCCCTGGAGGCTGCTCAC
GGCGGCAACGTGCTGATTACCGATACATGGATTAGCATGGGCCAGGAGGAGGAG
AAAAAGAAGCGGCTCCAGGCTTTTCAAGGCTATCAGGTCACCATGAAAACCTGCA
AAGGTCGCTGCCTCCGACTGGACTTTCCTGCATTGTCTACCCCGCAAGCCTGAGG
AAGTGGACGATGAGGTGTTCTACTCCCCACGGAGTCTGGTGTTCCCGGAAGCAG
AGAATCGGAAGTGGACCATCATGGCTGTCATGGTGTCGCTCTTGACTGACTATTC
TCCCCAACTGCAAAAACCCAAGTTTTAG

>pARM575 (SEQ ID NO: 140)

ATGCTTTTCAATCTGAGGATCCTGCTGAACAACGCTGCTTTTCGCAACGGTCATA
ACTTTATGGTTCGCAATTTTCGTTGTGGCCAGCCGCTGCAGAACAAGGTTTCAGCT
GAAGGGCAGAGATCTGCTGACTCTGAAGAACTTCACTGGGGAAGAAATCAAGTA
TATGTTATGGCTGTCCGCGGATCTGAAATTTTGAATCAAGCAGAAGGGCGAATAT
CTTCCCCTGCTGCAAGGGAAATCCTTGGGCATGATTTTTGAGAAGAGGAGCACTA
GGACTAGATTGTCAACAGAAACAGGCTTTGCTTTGTTGGGCGGACATCCCTGCTT
TCTGACGACACAGGATATCCACCTCGGCGTAAACGAGTCCCTCACCGACACTGCT
AGGGTACTGAGCAGCATGGCCGACGCTGTGCTAGCCCGGGTTTACAAGCAGTCA
GACCTGGACACCCTTGCCAAGGAAGCTTCTATTCCAATTATCAACGGCCTGAGTG
ACCTGTATCACCTATTCAAATACTCGCCGACTATTTGACGCTTCAAGAACATTA
CAGCAGCCTCAAGGGCTTAACCTTGAGTTGGATAGGCGACGGCAACAATATCCT
GCATTCCATTATGATGTCTGCCGCTAAGTTTGGCATGCATCTACAAGCCGCAACA
CCCAAGGGCTATGAACCCGACGCTAGCGTGACCAAGCTGGCCGAGCAGTATGCT
AAGGAAAATGGCACAAAGCTCCTTCTTACCAACGATCCCCTGGAGGCTGCTCAC
GGCGGCAACGTGCTGATTACCGATACATGGATTAGCATGGGCCAGGAGGAGGAG
AAAAAGAAGCGGCTCCAGGCTTTTCAAGGCTATCAGGTCACCATGAAAACCTGCA
AAGGTCGCTGCCTCCGACTGGACTTTCCTGCATTGTCTACCCCGCAAGCCTGAGG
AAGTGGACGATGAGGTGTTCTACTCCCCACGGAGTCTGGTGTTCCCGGAAGCAG
AGAATCGGAAGTGGACCATCATGGCTGTCATGGTGTCGCTCTTGACTGACTATTC
TCCCCAACTGCAAAAACCCAAGTTTTAG

>pARM708 (SEQ ID NO: 141)

ATGCTTTTAAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCTTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT

CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCCATCATCAACGGCCTGA
 GCGACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCCTGCAAGAGCA
 CTACAGCAGCCTGAAGGGTCTGACCCTGTCATGGATTGGCGATGGAAACAATAT
 TCTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCC
 ACGCCAAAAGGATACGAACCGGATGCGTCCGTGACGAAGTTGGCGGAACAGTAC
 GCGAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCCCCTCGAGGCTGCG
 CATGGGGGCAACGTGCTGATTACCGACACCTGGATCTCCATGGGGCAGGAGGAA
 GAGAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAAC
 CGCAAAAGTGGCAGCTTCGGACTGGACTTTCTGTCATTGCCTGCCGAGGAAGCC
 GGAGGAAGTCGACGACGAAGTGTTCTACTCGCCTCGGTCCCTGGTGTTCCTCGAG
 GCCGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCTTGCTGACTGAC
 TATAGCCCGCAGCTGCAGAAGCCTAAGTTCTAG

>pARM709 (SEQ ID NO: 142)

ATGCTTTTCAACCTGAGAATCCTCTTGAACAATGCTGCTTTTCGGAATGGCCACA
 ACTTTATGGTTCGGAACCTTCCGTTGCGGCCAGCCTTTACAAAACAAGGTCCAGCT
 GAAGGGCCGGGATTTGCTCACACTGAAGAACTTTACTGGAGAAGAGATCAAGTA
 CATGCTGTGGCTGTGCGGCCGACCTGAAGTTCAGGATCAAGCAGAAGGGAGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCTTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTCGATACCTTGGCAAAGGAGGCTTCCATTCCCATCATCAACGGCCTGAG
 CGACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCCTGCAAGAGCAC
 TACAGCAGCCTGAAGGGTCTGACCCTGTCATGGATTGGCGATGGAAACAATATT
 CTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCCA
 CTCCAAAAGGATACGAACCGGATGCGTCCGTGACCAAGTTGGCGGAACAGTACG
 CGAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCCCCTCGAGGCTGCGC
 ATGGGGGCAACGTGCTGATTACCGACACCTGGATCTCCATGGGGCAGGAGGAAAG
 AGAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAACC
 GCAAAAGTGGCAGCTTCGGACTGGACTTTCTGTCATTGCCTGCCGAGGAAGCCG
 GAGGAAGTCGACGACGAAGTGTTCTACTCGCCTCGGTCCCTGGTGTTCCTCGAGG
 CCGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCTTGCTGACTGACT
 ATAGCCCGCAGCTGCAGAAGCCTAAGTTCTAG

>pARM710 (SEQ ID NO: 143)

ATGCTTTTCAACCTGAGAATCCTCTTGAACAATGCTGCTTTTCGGAATGGCCACA
 ACTTTATGGTTCGGAACCTCCGTTGCGGCCAGCCTTTACAAAACAAGGTCCAGCT
 GAAGGGCCGGGATTTGCTCACACTAAAGAACTTTACTGGAGAAGAGATCAAGTA
 CATGCTATGGCTGTCGGCCGACCTGAAGTTCCGTATCAAGCAGAAGGGAGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTCGATACCTTGGCAAAGGAGGCTTCCATTCCCATCATCAACGGCCTGAG
 CGACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCCTGCAAGAGCAC
 TACAGCAGCCTGAAGGGTCTGACCCTGTATGGATTGGCGATGGAAACAATATT
 CTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCCA
 CTCCAAAAGGATACGAACCGGATGCATCCGTGACCAAGTTGGCGGAACAGTACG
 CGAAGGAGAACGGAACCAAGCTCCTGCTGACTAACGACCCGCTCGAGGCTGCGC

ATGGGGGTAACGTGCTGATTACGGACACCTGGATCTCCATGGGGCAGGAGGAAG
 AGAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAACC
 GCAAAAGTGGCAGCTTCGGACTGGACTTTCCTGCATTGCCTGCCGAGGAAGCCG
 GAGGAAGTCGACGACGAAGTGTTCTACTCGCCTCGGTCCCTGGTGTTCCTCCGAGG
 CCGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCTTGCTGACTGACT
 ATAGCCCGCAGCTGCAGAAGCCTAAGTTCTAG

>pARM711 (SEQ ID NO: 144)

ATGCTTTTCAACCTGAGAATCCTCTTGAACAATGCTGCTTTTCGGAATGGCCACA
 ACTTTATGGTTCGGAACCTCCGTTGCGGCCAGCCTTTACAAAACAAGGTCCAGCT
 GAAGGGCCGGGATTTGCTCACACTAAAGAACTTTACTGGAGAAGAGATCAAGTA
 CATGCTATGGCTGTCGGCCGACCTGAAGTTCCGTATCAAGCAGAAGGGAGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTCGATACCTTGGCAAAGGAGGCTTCCATTCCCATCATCAACGGCCTGAG
 CGACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCCTGCAAGAGCAC
 TACAGCAGCCTGAAGGGTCTGACCCTGTATGGATTGGCGATGGAAACAATATT
 CTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCCA
 CTCCAAAAGGATACGAACCGGATGCGTCCGTGACCAAGTTGGCGGAACAGTACG
 CGAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCCTCGAGGCTGCGC
 ATGGGGGCAACGTGCTGATTACCGACACCTGGATCTCCATGGGGCAGGAGGAAG
 AGAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAACC
 GCAAAAGTGGCAGCTTCGGACTGGACTTTCCTGCATTGCCTGCCGAGGAAGCCG
 GAGGAAGTCGACGACGAAGTGTTCTACTCGCCTCGGTCCCTGGTGTTCCTCCGAGG
 CCGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCTTGCTGACTGACT
 ATAGCCCGCAGCTGCAGAAGCCTAAGTTCTAG

>pARM712 (SEQ ID NO: 145)

ATGCTGTTCAACCTGCGAATCCTGCTGAACAACGCCGCTTTTCGGAACGGGCACA
 ACTTTATGGTGAGGAACTTTTCGCTGCGGACAGCCCCTCCAGAATAAGGTCCAGCT
 GAAGGGCAGGGACCTGCTGACCCTGAAAAATTTACAGGGGAGGAAATCAAGTA
 TATGCTGTGGCTGTCAGCTGATCTGAAGTTCCGGATCAAGCAGAAGGGCGAATA
 TCTGCCTCTGCTCCAGGGCAAAGCCTGGGGATGATCTTCGAAAAGCGCAGTACT
 CGGACCAGACTGTCAACCGAGACTGGATTTCGCTCTGCTGGGAGGACACCCTTGTT
 TTCTGACCACTCAGGACATTCACCTGGGAGTGAACGAGTCCCTGACCGACACTGC
 TCGCGTCCTGAGCTCTATGGCCGACGCTGTGCTGGCTCGAGTCTACAAACAGTCC
 GACCTGGATACCCTGGCCAAGGAAGCTTCTATCCCAATTATTAACGGCCTGTCAG
 ACCTGTATCACCCCATCCAGATTCTGGCCGATTACCTGACCCTCCAGGAGCACTA
 TTCTAGTCTGAAAGGGCTGACACTGAGTTGGATTGGGGACGGAAACAATATCCT
 GCACTCTATTATGATGTCAGCCGCCAAGTTTGAATGCACCTCCAGGCTGCAACC
 CCAAAAGGCTACGAACCCGATGCCTCAGTGACAAAGCTGGCTGAACAGTACGCC
 AAAGAGAACGGCACTAAGCTGCTGCTGACCAACGACCCTCTGGAGGCCGCTCAC
 GGAGGCAACGTGCTGATCACCGATACCTGGATTAGTATGGGACAGGAGGAAGAG
 AAGAAGAAGCGGCTCCAGGCCTTCCAGGGCTACCAGGTGACAATGAAAACCGCT
 AAGGTGCGAGCCAGCGATTGGACCTTTCTGCACTGCCTGCCCAGAAAGCCCGAA
 GAGGTGGACGACGAGGTCTTCTACTCTCCCAGAAGCCTGGTGTTTCCCGAAGCTG
 AGAATAGGAAGTGGACAATTATGGCAGTGATGGTCAGCCTGCTGACTGATTATC
 ACCTCAGCTCCAGAAACCAAAGTTCTGA

>pARM713 (SEQ ID NO: 146)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCCTGGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTCC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM714 (SEQ ID NO: 147)

ATGCTGTTCAACCTGCGAATCCTGCTGAACAACGCCGCTTTTCGGAACGGGCACA
 ACTTTATGGTGAGGAACCTTCGCTGCGGACAGCCCCTCCAGAATAAGGTCCAGCT
 GAAGGGCAGGGACCTGCTGACCCTGAAAAATTTACAGGGGAGGAAATCAAGTA
 TATGCTGTGGCTGTCAGCTGATCTGAAGTTCCGGATCAAGCAGAAGGGCGAATA
 TCTGCCTCTGCTCCAGGGCAAAAGCCTGGGGATGATCTTCGAAAAGCGCAGTACT
 CGGACCAGACTGTCAACCGAGACTGGATTTCGCTCTGCTGGGAGGACACCCTTGTT
 TTCTGACCACTCAGGACATTCACCTGGGAGTGAACGAGTCCCTGACCGACACTGC
 TCGCGTCCTGAGCTCTATGGCCGACGCTGTGCTAGCTCGAGTCTACAAACAGTCC
 GACCTGGATAACCTGGCCAAGGAAGCTTCTATCCCAATTATTAACGGCCTGTGAG
 ACCTGTATCACCCCATCCAGATTCTGGCCGATTACCTGACCCTCCAGGAGCACTA
 TTCTAGTCTGAAAGGGCTGACACTGAGTTGGATTGGGGACGGAAACAATATCCT
 GCACTCTATTATGATGTCAGCCGCCAAGTTTGGAATGCACCTCCAGGCTGCAACC
 CAAAAGGCTACGAACCCGATGCCTCAGTGACAAAGCTGGCTGAACAGTACGCC
 AAAGAGAACGGCACTAAGCTGCTGCTGACCAACGACCCTCTGGAGGCCGCTCAC
 GGAGGCAACGTGCTGATCACCGATACCTGGATTAGTATGGGACAGGAGGAAGAG
 AAGAAGAAGCGGCTCCAGGCCTTCCAGGGCTACCAGGTGACAATGAAAACCGCT
 AAGGTCGCAGCCAGCGATTGGACCTTTCTGCACTGCCTGCCAGAAAGCCCGAA
 GAGGTGGACGACGAGGTCTTCTACTCTCCAGAAAGCCTGGTGTTCCTCGAAGCTG
 AGAATAGGAAGTGGACAATTATGGCAGTGATGGTCAGCCTGCTGACTGATTATTC
 ACCTCAGCTCCAGAAACCAAAGTTCTGA

>pARM715 (SEQ ID NO: 148)

ATGCTGTTCAACCTGCGAATCCTGCTGAACAATGCCGCTTTTCGGAACGGGCACA
 ATTTTCATGGTGAGGAACCTTCGCTGCGGACAGCCCCTCCAGAACAAGGTCCAGCT
 GAAGGGCAGGGACCTGCTGACCCTGAAAAATTTACAGGGGAGGAAATCAAGTA
 CATGCTGTGGCTGTCAGCCGATCTGAAGTTCCGGATCAAGCAGAAGGGCGAATA

TCTGCCTCTGCTCCAGGGCAAAAGCCTGGGGATGATCTTCGAAAAGCGCAGTACT
 CGGACCAGACTGTCAACAGAGACTGGATTTCGCACTGCTGGGAGGACACCCATGT
 TTTCTGACCACACAGGACATTCATCTGGGAGTGAACGAGTCCCTGACCGACACA
 GCACGCGTCCTGAGCTCCATGGCTGATGCAGTGCTGGCTCGAGTCTACAAACAGT
 CTGACCTGGATAACCTGGCCAAGGAAGCTTCTATCCCAATCATTAAATGGCCTGAG
 TGACCTGTATCACCCCATCCAGATTCTGGCCGATTACCTGACCCTCCAGGAGCAT
 TATTCTAGTCTGAAAGGGCTGACACTGAGCTGGATTGGGGACGGAAACAATATC
 CTGCACTCCATTATGATGAGCGCCGCCAAGTTTGGAATGCACCTCCAGGCTGCAA
 CCCCCAAAAGGCTACGAACCCGATGCCTCCGTGACAAAGCTGGCAGAACAGTATG
 CCAAAGAGAACGGCACTAAGCTGCTGCTGACCAATGACCCTCTGGAGGCCGCTC
 ACGGAGGCAACGTGCTGATCACTGATACCTGGATTAGTATGGGACAGGAGGAAG
 AGAAGAAGAAGCGGCTCCAGGCCTTCCAGGGCTACCAGGTGACAATGAAAAGTGA
 CTAAGGTCGCAGCCAGCGACTGGACCTTTCTGCATTGCCTGCCAGAAAGCCTGA
 AGAGGTGGACGATGAGGTCTTCTACTACCCAGAAAGCCTGGTGTTCCTGAAGCT
 GAGAATAGGAAGTGGACAATCATGGCAGTGATGGTCAGCCTGCTGACTGATTAT
 TCCCCTCAGCTCCAGAAACCAAAGTTCTGA

>pARM716 (SEQ ID NO: 149)

ATGCTTTTCAACCTTCGCATTCTCCTCAACAACGCCGCGTTTAGAAACGGACACA
 ACTTCATGGTCCGCAACTTCCGCTGCGGACAGCCGCTGCAGAACAAGGTCCAGCT
 CAAGGGTCGGGATCTCCTGACGCTGAAGAAGTTTACCGGCGAAGAGATTAAGTA
 CATGCTGTGGCTGTCCGCCGACCTTAAGTTCCGGATCAAGCAGAAGGGCGAATA
 CCTTCCCCTGCTGCAAGGAAAGTCCCTGGGCATGATCTTCGAGAAGCGCAGTACC
 AGAACCAGACTCTCCACTGAAACCGGGTTTCGCGCTGCTTGGCGGGCCACCCGTGTT
 TCCTCACTACGCAAGACATCCATCTTGGCGTGAACGAGTCCCTTACCGACACCGC
 CAGGGTGCTGTCAAGCATGGCCGACGCCGTCTTGC CGCGCTGTACAAGCAGTC
 AGACCTTGATACTCTGGCCAAGGAAGCCTCCATCCCTATTATCAACGGCCTATCC
 GACCTTTACCACCCGATCCAGATCCTCGCTGACTACCTGACCCTGCAAGAACACT
 ACAGCAGCCTCAAGGGACTGACTCTGTCTTGGATCGGCGACGGGAACAACATCC
 TGCACTCAATCATGATGAGCGCAGCCAAGTTCGGCATGCATCTCCAAGCCGCTAC
 ACCCAAGGGTTATGAACCGGACGCCTCTGTGACCAAGTTGGCAGAACAGTACGC
 CAAGGAGAACGGTACTAAGCTCCTTTTAACCAACGACCCCTCGAAGCAGCCCA
 TGGCGGGAATGTGCTCATTACCGATACTTGGATTTTCGATGGGCCAGGAGGAGGA
 GAAGAAGAAGCGGCTGCAGGCGTTCCAGGGCTACCAGGTCACCATGAAAACCTGC
 CAAAGTGGCCGCCTCGGATTGGACCTTTCTCCACTGCCTGCCTCGGAAGCCTGAG
 GAGGTGGACGACGAAGTGTTCTACTCCCCACGGTCCCTCGTGTTCCCCGAGGCCG
 AAAATAGGAAGTGGACCATCATGGCCGTGATGGTGTCCCTCTTGACCGATTACA
 GCCCCGACGCTTCAGAAGCCTAAATTCTAG

>pARM717 (SEQ ID NO: 150)

ATGCTTTTCAATCTTCGCATCCTGTTGAACAACGCCGCCTTCCGCAATGGTCACA
 ACTTCATGGTCCGGAAGTTTCAAGATGTGGACAGCCTCTCCAAAACAAGGTCCAGCT
 GAAGGGAAGGGACCTCTTAACCTCAAAAAGTTTACTGGAGAGGAGATCAAGTA
 CATGCTGTGGCTTAGCGCCGACCTTAAGTTCCGGATCAAGCAGAAGGGAGAGTA
 CCTCCCGCTGCTGCAAGGAAAGAGTCTTGGGAATGATCTTCGAGAAGCGGTCCAC
 CAGAACTCGCCTCTCCACTGAAACCGGATTCGCACTCCTGGGTGGACACCCGTGC
 TTTCTGACCACCCAAGACATCCACCTCGGAGTGAACGAGAGCCTCACGGACACC
 GCGAGAGTGCTGTCATCCATGGCCGACGCCGTGCTTGCACGGGTCTACAAGCAG
 TCCGATCTGGACACTCTTGCCAAGGAAGCCTCCATTCCTATCATTAACGGTCTGT
 CGGATCTGTACCACCCGATTCAGATCCTTGCGGACTACCTCACACTTCAAGAACA

CTATTCAAGCCTAAAGGGTCTGACCCTGTCCTGGATCGGAGATGGAAACAACATT
 CTCCATTCCATCATGATGAGCGCTGCCAAGTTCGGAATGCATCTCCAAGCAGCGA
 CTCCTAAGGGTTACGAGCCGGACGCCTCAGTGACTAAGCTGGCCGAGCAGTACG
 CCAAGGAGAACGGTACCAAAGTGTGCTTACTAACGACCCGCTTGAAGCGGCC
 ATGGAGGAAACGTGCTGATTACCGACACCTGGATTTTCGATGGGACAGGAAGAGG
 AGAAGAAGAAGCGGCTCCAGGCGTTCCAGGGATACCAGGTCACCATGAAAACG
 GCCAAAGTGGCCGCTAGCGATTGGACCTTTCTGCACTGCCTCCCGCGCAAGCCTG
 AAGAAGTGGACGACGAAGTGTTCTACTCCCTCGCTCTCTTGTGTTCCCGGAAGC
 CGAAAACAGGAAGTGGACCATCATGGCCGTGATGGTGTCCCTCCTGACCGATTAC
 CAGCCCCGACGCTGCAGAAGCCTAAGTTCTAG

>pARM718 (SEQ ID NO: 151)

ATGCTTTTCAATCTCCGCATCCTCCTCAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAGAACAAGGTCCAGCT
 CAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAAATCAAGTA
 CATGCTCTGGCTCTCCGCCGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATAC
 CTTCCGCTGCTGCAAGGAAAGTCGCTCGGCATGATCTTTGAGAAGCGCTCAACCC
 GCACCAGGCTGTCCACTGAAACCGGGTTCGCGCTGCTTGGTGGCCACCCCTGCTT
 CCTGACCACCCAAGACATTCACCTCGGAGTGAACGAATCGCTCACTGATACTGCC
 CGGGTGCTGTCTGTCGATGGCCGATGCAGTGCTGGCCAGGGTGTACAAACAGTCC
 GATCTGGACACTCTGGCCAAGGAGGCGTCCATCCCTATTATCAACGGCCTTTCCG
 ACCTCTACCACCCGATTCAGATCCTTGCCGATTACCTCACCCCTGCAAGAACTA
 CTCGTCACTGAAGGGTCTGACCTTGTCCTGGATCGGCGACGGCAACAACATCCTC
 CATTCCATTATGATGTCCGCCGCCAAATTCGGCATGCATCTTCAAGCCGCAACCC
 CTAAGGGTTACGAGCCGGACGCTTCCGTGACCAAGCTCGCCGAGCAGTACGCTA
 AGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCCTAGAGGCAGCCACG
 GGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGACAGGAAGAAGAGA
 AGAAGAAGCGGTTACAGGCGTTCCAGGGCTATCAGGTCACCATGAAAACCGCCA
 AGGTGCTGCCTCGGACTGGACCTTCTGCATTGCCTGCCTCGCAAGCCCGAAGA
 AGTGGACGACGAGGTGTTCTACTCGCCACGGTCCCTTGTGTTCCCTGAGGCCGAG
 AATAGAAAGTGGACCATTATGGCCGTGATGGTGTCCCTTCTCACCGACTACTCGC
 CGCAACTGCAGAAACCCAAGTTCTAG

>pARM719 (SEQ ID NO: 152)

ATGCTTTTCAATCTTCGCATCCTCCTCAACAACGCCGCTTCCGGAACGGTCACA
 ACTTCATGGTCCGGAACCTTCGCTGCGGCCAGCCGCTCCAAAACAAAGTGCAGCT
 TAAGGGCCGCGATCTCCTGACCCTGAAGAACTTCACCGGAGAGGAAATCAAGTA
 CATGCTGTGGCTCTCGGCGGACCTGAAGTTTAGGATTAAGCAGAAGGGGGAGTA
 TCTGCCGCTGCTCCAAGGGAAGTCCCTTGGCATGATCTTCGAAAAGAGGTCCACC
 CGGACTCGGCTCAGCACCGAAACAGGTTTTTGCATTCTGGGGGGGCCACCCGTGCT
 TCCTGACGACCCAGGACATCCATCTGGGTGTCAACGAGAGTTTGACCGACACTGC
 CAGAGTGCTGTATCCATGGCGGACGCGGTGCTCGCGAGAGTGTACAAGCAGTC
 CGATCTTGACACCCTGGCAAAAAGAGGCTTCAATCCCGATCATTAACGGACTCTCG
 GATCTGTACCACCCTATCCAAATCTTGGCCGACTACCTGACCCTGCAAGAACT
 ACAGCTCCCTGAAGGGCCTGACTCTTTCCTGGATTGGCGATGGAAACAACATTCT
 CCATTCTATTATGATGTCCGCCGCCAAGTTCGGCATGCACCTTCAAGCCGCCACC
 CCGAAGGGCTACGAACCTGACGCCTCCGTGACTAAGCTAGCCGAACAGTACGCT
 AAGGAGAACGGCACTAAGCTTCTCCTTACCAACGATCCGCTGGAGGCGGCCCAT
 GGCGGAAATGTGCTTATCACCGACACCTGGATTAGCATGGGGCAGGAAGAAGAG
 AAGAAGAAACGGCTCCAGGCATTCCAGGGCTACCAGGTCACCATGAAAACCTGCC

AAGGTCGCCGCTAGCGACTGGACCTTCCCTCCACTGTCTGCCTCGCAAGCCTGAAG
 AAGTGGACGACGAGGTGTTCTACTCCCCGCGCTCCCTCGTGTTTCTGAGGCCGA
 GAACAGAAAGTGGACCATTATGGCCGTGATGGTGTCACTTACGGACTACAG
 CCCGACGCTGCAGAAAGCCGAAGTTCTAG

>pARM720 (SEQ ID NO: 153)

ATGCTTTTAACTTGAGAATCCTTCTGAACAACGCCGCTTTCCGCAACGGTCATA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCCCCTCCAAAACAAAGTGCAGCT
 GAAGGGCCGGGACCTTCTTACGCTGAAGAATTTACCCGGCGAAGAAATCAAGTA
 CATGCTCTGGCTGTCCGCCGATCTTAAGTTCCGCATTAAGCAGAAGGGGGAATAC
 CTCCCGCTGCTGCAAGGGAAGTCGCTGGGCATGATTTTTGAGAAGCGGTCAACTC
 GCACCCGCCTGTCCACTGAAACTGGATTTCGCACTGCTCGGTGGCCATCCCTGCTT
 CCTGACCACCCAAGACATCCACCTCGGCGTGAACGAGTCCCTGACTGACACCGC
 CCGGGTCTTATCCTCGATGGCCGATGCTGTGCTTGCGAGGGTGTACAAGCAGTCC
 GACCTCGACACACTCGCGAAGGAGGCCTCCATCCCCATCATCAACGGCCTGTCCG
 ACCTTTACCACCCAATTCAGATCCTCGCCGATTACCTGACCCTGCAAGAGCACTA
 CTCGTCGCTCAAGGGGCTTACCCTCTCGTGGATTGGCGACGGCAACAACATCCTT
 CACTCCATCATGATGTGCGCAGCGAAGTTCGGCATGCATCTGCAAGCCGCCACGC
 CTAAGGGTTATGAACCGGATGCCTCAGTGACCAAGCTCGCCGAACAGTACGCGA
 AAGAGAATGGAACCAAGCTACTTCTGACCAACGACCCCCCTGGAGGCCGCTCACG
 GCGGCAACGTCCTCATTACCGATACTTGGATTTTCGATGGGACAGGAAGAGGAAA
 AGAAGAAGAGACTGCAGGCGTTCCAGGGATAACCAGGTCACCATGAAAAGTCCA
 AAGTGGCAGCCTCCGACTGGACCTTCCCTTCACTGCCTGCCGAGGAAGCCTGAAG
 AGGTGGACGACGAGGTGTTCTACTCCCCGCGCTCCTTGGTGTTTCTGAGGCCGA
 AAACCGGAAGTGGACTATCATGGCCGTGATGGTGTCCCTCCTCACCAGCTACTCG
 CCGCAACTGCAGAAGCCTAAGTTCTAG

>pARM721 (SEQ ID NO: 154)

ATGTTATTCAACCTTAGAATTCTCCTTAACAACGCCGCCTTCCGGAATGGGCATA
 ACTTTATGGTCCGCAATTTCCGCTGTGGACAGCCTCTGCAAAACAAGGTCCAGCT
 CAAGGGCCGGGATCTGCTGACTCTCAAGAACTTCACTGGGGAAGAAATCAAGTA
 CATGCTCTGGCTGAGCGCCGACCTCAAGTTCCGCATCAAGCAGAAGGGAGAGTA
 CCTCCCGCTGCTCCAAGGGAAGTCCCTGGGCATGATCTTCGAGAAGAGATCCACC
 CGCACCAGACTTTCCACTGAGACTGGCTTCGCCTTGCTGGGAGGCCACCCATGCT
 TCCTGACGACCCAGGACATTCACCTTGGCGTGAACGAGTCCCTGACTGACACCGC
 AAGGGTGTGTGCTCGATGGCCGACGCCGTGCTTGCCCGGGTGTACAAGCAGAG
 CGATCTTGACACCCTGGCTAAGGAAGCTTCCATTCCCATCATCAACGGTCTGAGC
 GACCTGTACCACCCGATTCAGATCCTGGCGGACTACCTAACCCTGCAAGAGCACT
 ATAGCTCCCTGAAGGGCCTCACACTTTTCATGGATCGGCGACGGCAACAACATCCT
 GCACTCTATTATGATGAGCGCTGCCAAATTCGGCATGCACCTCCAAGCCGCCACG
 CCTAAAGGCTACGAGCCCGACGCCTCGGTGACCAAGCTTGCGGAGCAGTACGCG
 AAGGAAAACGGCACCAAGCTGCTTCTACCAACGATCCTCTGGAAGCGGCCCAT
 GGTGGCAACGTGCTCATTACCGACACTTGGATCTCCATGGGACAGGAGGAGGAA
 AAGAAGAAGCGGCTCCAGGCGTTTCAGGGTTACCAGGTCACCATGAAAACCGCC
 AAGGTCGACGCTCCGACTGGACCTTCCCTTCAATTGCCTTCCGCGCAAGCCCGAAG
 AAGTGGACGATGAAGTGTTTTACTACCTCGGTCACTCGTGTTCCCGGAAGCAGA
 GAACAGGAAATGGACCATTATGGCCGTGATGGTGTCCCTGCTCACCATTACAGT
 CCGCAACTGCAGAAGCCCAAGTTCTAG

>pARM722 (SEQ ID NO: 155)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCTTTCTACTGAAACTGGGTTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCTTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAAATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCTTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCTAAGGGTTACGAACCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTGCTGCTGACTAACGACCCGCTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAGGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTGCTGCCTCCGACTGGACCTTCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTTCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM723 (SEQ ID NO: 156)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 TAAGGGCCGGGATCTCCTCACCTTAAAAACTTCACCGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACCTTAAGTTCCGCATTAAGCAGAAGGGGGAATAC
 CTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACCA
 GGACCAGGCTTTCTACTGAAACTGGGTTTCGCGCTTCTCGGCGGTATCCCTGCTT
 CCTCACGACCCAAGACATCCACCTCGGAGTGAACGAATCCCTCACGGATACTGC
 CCGCGTGCTTTCGAGCATGGCAGACGCCGTGCTCGCCCGGGTGTACAAACAGTCC
 GATCTCGACACTCTCGCCAAGGAGGCGTCAATTCTTATTATCAACGGTCTTAGTG
 ACCTTTACCACCCGATCCAGATCCTCGCCGATTACCTCACACTCCAAGAACTACTA
 CAGCTCCCTTAAGGGTCTTACCCTCTCTGGATCGGCGACGGCAACAACATTCTC
 CACTCCATCATGATGTCCGCCGCAAGTTCCGGCATGCATCTTCAAGCCGCCACCC
 CGAAGGGCTACGAGCCTGATGCTTCCGTGACTAAGCTCGCCGAGCAGTACGCTA
 AGGAGAACGGAACCAAGCTTCTTCTACTAAGCAGCCCACTCGAAGCAGCCCATG
 GGGGCAACGTGCTTATCACTGACACCTGGATCTCCATGGGCCAGGAAGAAGAGA
 AGAAGAAGCGGCTCCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGCCA
 AGGTGCTGCTGCCCTCCGACTGGACCTTCTTCCACTGCCTCCCTCGCAAACCTGAAGA
 AGTGGACGACGAGGTGTTCTACTCGCCCCGGAGCCTCGTGTTCCCCGAGGCCGA
 GAATAGAAAGTGGACCATTATGGCCGTGATGGTGTCACTCCTCACCGACTACAG
 CCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM724 (SEQ ID NO: 157)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGACATA
 ACTTCATGGTCCGGAACCTTCAGATGTGGACAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGTCCGGGATCTTCTGACCCTGAAGAACTTTACCGGAGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGAGAATA
 CCTCCCGCTGCTTCAAGGAAAGAGCCTCGGAATGATTTTTGAGAAGCGCTCAACC

AGGACCCGCCTTTCTACTGAAACTGGATTTCGCGCTGCTGGGTGGACACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACTGATACCGC
 CCGGGTGTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGTC
 CGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATCATCAACGGACTTAGT
 GACCTCTACCATCCGATTCAAATCCTGGCCGACTACCTCACCTGCAAGAACACT
 ACAGCTCCCTGAAGGGTCTGACATTGTCTGGATCGGAGATGGAAACAACATTCT
 CCACTCCATCATGATGTCCGCCGCAAAATTCGGAATGCATCTTCAAGCCGCCACG
 CCTAAGGGTTACGAACCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGCT
 AAGGAGAACGGTACCAAGCTTCTCCTGACCAACGACCCACTAGAAGCAGCCCAC
 GGTGGAAACGTGCTTATTACTGACACTTGGATCTCCATGGGACAGGAGGAAGAG
 AAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGCC
 AAGGTCGCTGCCTCCGACTGGACCTTCTGCACTGCCTGCCTCGCAAGCCTGAAG
 AAGTGGACGACGAGGTGTTCTACTCGCCGCGGAGCCTCGTGTTCCCCGAGGCCG
 AGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGACTACA
 GCGCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM725 (SEQ ID NO: 158)

ATGCTTTTCAACCTCCGCATTCTCCTCAACAACGCTGCCTTCCGGAATGGACATA
 ACTTCATGGTCCGGAACCTTCAGATGCGGACAGCCGCTTCAGAACAAAGGTCCAGC
 TTAAGGGGAGAGATCTCCTTACCCTCAAAAACCTTCACTGGCGAAGAAATCAAGT
 ACATGCTCTGGCTTAGTGCGGATCTCAAGTTCCGCATCAAGCAGAAGGGAGAAT
 ACCTCCCGCTCCTTCAAGGAAAGAGCCTCGGCATGATTTTTGAGAAGAGGTCCAC
 CAGAACTCGCCTTTCAACCGAGACTGGGTTTCGCCCTGCTTGGCGGTACCCCTGC
 TTCCTCACTACCCAAGACATCCACCTCGGCGTGAACGAGAGCCTTACCGACACCG
 CCCGCGTGCTCTCCTCAATGGCCGACGCTGTGCTCGCCCGGGTGTACAAGCAGTC
 CGACCTTGATACTCTCGCCAAGGAGGCCTCCATCCCAATTATCAACGGGCTCTCT
 GATCTCTACCACCCTATCCAAATCCTCGCGGACTACCTCACCTCCAAGAGCACT
 ATAGCTCGCTCAAGGGCCTCACCTTTCTGATTGGCGACGGCAACAACATTCT
 TCACTCGATCATGATGTCCGCCGCCAAGTTCGGCATGCATCTCCAAGCCGCGACC
 CCAAGGGCTACGAGCCTGACGCATCCGTGACCAAGCTCGCCGAGCAGTACGCG
 AAGGAAAATGGCACCAAGCTTCTTCTACCAACGACCCCCCTTGAGGCCGCTCATG
 GCGGCAACGTGCTCATCACTGACACTTGGATCAGCATGGGCCAGGAGGAGGAAA
 AGAAGAAGCGCCTTCAGGCATTCCAGGGTTACCAGGTACCATGAAAACCGCCA
 AAGTGGCCGCCTCCGACTGGACCTTCTTCACTGTCTCCCGCGGAAGCCTGAAGA
 AGTGGATGACGAAGTGTCTTACTCCCTCGGTCACTCGTGTTCCCGGAAGCAGAA
 AACAGGAAGTGGACCATATGGCGGTGATGGTGTCCCTCCTCACCGACTACAGCC
 CGCAGCTTCAGAAACCCAAGTTCTAG

>pARM726 (SEQ ID NO: 159)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCAGCGTTTAGAAACGGTCACA
ACTTCATGGTCCGGAACCTCCGCTGTGGACAGCCGCTTCAAAACAAGGTCCAGCT
GAAGGGTCGGGACCTTCTGACCCTGAAGAACTTTACTGGAGAAGAGATCAAGTA
CATGCTTTGGCTGTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGAGAATA
CCTTCCGCTGCTCCAAGGAAAGAGCCTGGGAATGATCTTTGAGAAGCGCTCAAC
CAGGACCCGCCTTTCTACTGAAACTGGATTTCGCGCTGCTGGGTGGTCACCCTTGC
TTCCTGACGACCCAGGACATTACCTCGGAGTGAACGAGTCCCTCACTGATACCG
CCAGAGTGTTATCGAGCATGGCAGATGCCGTGCTGGCTAGGGTGTACAAACAGT
CCGATCTGGACACCCTGGCCAAGGAGGCATCAATTCCTATTATCAACGGACTTAG
TGACCTCTACCATCCGATTCAAATCCTGGCCGATTACCTCACCTGCAAGAACAC
TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGAGATGGAAACAACATTC

TCCATTCCATCATGATGTCCGCGGCCAAGTTCGGAATGCATCTCCAAGCCGCCAC
GCCGAAAGGATACGAGCCGGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCGCTAGAAGCCGCCCA
CGGTGGAAACGTGCTTATTACTGACACCTGGATCTCCATGGGACAGGAAGAAGA
GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
CAAGGTCGCCGCCTCCGACTGGACCTTCCTTCACTGCCTGCCTCGGAAGCCTGAA
GAAGTGGACGACGAGGTGTTCTACTCGCCGCGGAGCCTCGTGTTCCCTGAGGCC
GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTCCTCACCGACTAC
AGCCCGCAGCTTCAGAAGCCTAAGTTCTAG

>pARM727 (SEQ ID NO: 160)

ATGCTTTTCAATCTCCGCATTCTCCTCAACAACGCAGCCTTTAGAAACGGCCACA
ACTTCATGGTCCGGAACCTCAGATGTGGCCAGCCGCTTCAAGAACAAGGTCCAGCT
CAAGGGCCGGGACCTCCTCACCTCAAAAACCTTACC GGCGAAGAGATCAAGTA
CATGCTCTGGCTTTCGGCCGACCTTAAGTTCCGCATCAAGCAGAAGGGGGAATAC
CTTCCGCTGCTTCAAGGAAAGTCCCTCGGCATGATCTTTGAAAAGCGCTCGACCA
GGACCCGCCTTTCCACTGAAACCGGGTTTCGCGCTTCTCGGTGGCCACCCCTGCTT
CCTCACCACCCAAGACATTACCTCGGAGTGAACGAATCCCTTACCGATACCGCA
AGAGTGCTTTCGTTCGATGGCCGATGCCGTGCTTGCGCGGGTGTACAAGCAGTCAG
ATCTCGACACTCTCGCCAAGGAGGCGTCCATTCTATTATCAACGGCCTTTCCGA
CCTTTACCACCCGATTACAGATCCTCGCCGATTACCTCACCTGCAAGAGCACTAC
TCGTCACTCAAGGGTCTTACCCTCTCCTGGATCGGCGACGGAAACAACATCCTCC
ATTGATCATGATGTCCGCCGCCAAATTCGGCATGCACCTCCAAGCCGCGACCCC
GAAGGGTTACGAGCCCGACGCTTCCGTGACCAAGCTCGCCGAACAGTACGCTAA
GGAAAACGGCACCAAGCTCCTCCTACTAACGACCCTCTCGAAGCAGCCCATGG
GGCAACGTGCTCATTACTGACACTTGGATCTCGATGGGCCAGGAAGAGGAGAA
AAAGAAGCGGCTTCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGCCAA
GGTCGCTGCCTCGGACTGGACCTTCCTTCACTGCCTTCCGCGCAAGCCTGAAGAG
GTGGACGATGAGGTGTTCTACTCCCCACGGTCCCTTGTTGTTCCCCGAGGCCGAGA
ATAGGAAGTGGACCATCATGGCCGTGATGGTGTGCTCCTCACTGACTACTCCCC
GCAACTTCAGAAGCCTAAGTTCTAG

>pARM728 (SEQ ID NO: 161)

ATGCTGTTTAATCTGAGAATACTTCTAAACAACGCCGCCTTCCGGAATGGCCATA
 ACTTTATGGTTCGGAATTTCCGCTGCGGCCAGCCGCTGCAGAACAAGGTCCAGCT
 GAAGGGAAGAGACTTGCTGACCCTCAAGAAGTTCACCGGAGAAGAAATCAAGTA
 TATGCTGTGGCTGTCCGCCGACCTGAAATTCCGCATCAAGCAGAAGGGCGAATA
 TCTGCCGCTGTTGCAAGGGAAGTCCCTGGGGATGATCTTCGAGAAGAGGTCCAC
 CAGAACACGGCTTTCAACCGAAACCGGGTTTGCCTGCTGGGTGGACACCCCTGT
 TTTCTGACCACTCAAGATATCCACCTGGGCGTGAACGAGTCCCTTACCGACACTG
 CTAGGGTGTGTCCAGCATGGCCGATGCCGTCTGGCTCGCGTGTACAAGCAGTC
 CGACCTGGATACCCTGGCAAAGGAAGCGTCCATTCCCATTATCAACGGGCTGTCC
 GACCTGTACCATCCGATTCAAATCCTGGCGGACTACCTGACTCTGCAAGAGCATT
 ACAGCAGCTTGAAGGGGCTTACTCTCTCGTGGATCGGCGACGGGAACAACATCC
 TGCCTCCATCATGATGTCCGCCGCCAAGTTCGGGATGCATTTGCAAGCTGCGAC
 CCCGAAAGGTTACGAGCCCGATGCTAGCGTAACTAAGCTTGCCGAACAGTACGC
 CAAAGAGAATGGTACAAAAGTCTTCTGACTAACGACCCGCTGGAAGCAGCCCA
 CGGCGGGAACGTGCTGATAACCGACACCTGGATTTCAATGGGGCAGGAGGAAGA
 GAAGAAGAAGCGACTGCAGGCGTTCCAAGGCTATCAGGTTACCATGAAAACCGC
 CAAAGTGGCAGCCAGCGATTGGACTTTCCTGCACTGTCTGCCGCGGAAGCCCGA

GGAAGTTGATGACGAAGTATTCTACTCACCCCGGAGCCTCGTGTTCCCCGAGGCC
 GAAAACCGGAAGTGGACTATTATGGCCGTGATGGTGTGCTGTTGACCGACTAC
 AGCCCGCAACTGCAGAAGCCGAAGTTT TAG

>pARM729 (SEQ ID NO: 162)

ATGCTTTTCAACCTGAGGATCCTTTTGAACAACGCCGCCTTTCGCAACGGCCACA
 ACTTTATGGTCCGCAATTTCCGCTGCGGGCAGCCGCTGCAGAACAAGGTCCAGCT
 GAAGGGCCGGGATCTGCTGACCCTGAAGAAGTTCACCGGGGAGGAAATCAAGTA
 CATGCTTTGGCTCTCCGCCGATCTGAAGTTCAGAATCAAGCAGAAGGGAGAGTA
 CCTCCCGTTGCTGCAAGGAAAGTCACTCGGAATGATTTTCGAAAAGAGAAGCAC
 TAGGACCCGCCTCTCAACTGAAACCGGGTTTCGCGCTGCTCGGGGGCCATCCGTGT
 TTCCTGACTACCCAAGACATCCACCTGGGAGTGAACGAGTCGCTGACCGACACC
 GCACGCGTGCTGTATCCATGGCGGACGCAGTGCTTGCCCGGGTGTACAAGCAG
 TCGGACCTGGACACTCTTGCCAAGGAGGCATCAATCCCCATCATTAACGGACTGT
 CCGATCTCTACCACCCGATTGAGATCCTGGCTGACTACCTAACCTGCAAGAGCA
 CTACTCAAGCCTGAAGGGGCTGACCCTGTCGTGGATCGGGGACGGCAACAACAT
 TCTGCACTCCATCATGATGTGCGCGGCTAAGTTCGGGATGCATTTGCAAGCGGCA
 ACTCCGAAGGGTTATGAACCCGACGCCTCCGTGACCAAGCTGGCCGAACAGTAC
 GCCAAGGAAAACGGAACCAAGTTGCTGCTGACTAATGATCCCCTGGAGGCGGCC
 CACGGGGGGAACGTGCTGATAACCGATACCTGGATCTCCATGGGGCAGGAAGAA
 GAGAAGAAAAAGCGGCTGCAGGCATTCCAGGGATACCAGGTCACCATGAAAAC
 CGCAAAAGTGGCAGCCAGCGACTGGACTTTCCTCCATTGCCTGCCGCGAAAGCC
 GGAGGAGGTCGATGACGAGGTGTTCTACTCCCCGCGGTGCTGGTGTTCCTGGA
 GGCGGAAAACCGGAAGTGGACCATATGGCCGTGATGGTGTCACTCCTGACTGA
 CTACAGCCCGCAACTGCAGAAGCCGAAGTTCTAG

>pARM1787 (SEQ ID NO: 163)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1788 (SEQ ID NO: 164)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCGCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCCTGGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGA

>pARM1789 (SEQ ID NO: 165)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTACCCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCCTGGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGA

>pARM1790 (SEQ ID NO: 166)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC

AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACCTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1791 (SEQ ID NO: 167)

ATGCTTTTCAACCTGAGAATCCTCTTGAACAATGCTGCTTTTCGGAATGGCCACA
 ACTTTATGGTTCGGAACCTTCCGTTGCGGCCAGCCTTTACAAAACAAGGTCCAGCT
 GAAGGGCCGGGATTTGCTCACACTAAAGAACTTTACTGGAGAAGAGATCAAGTA
 CATGCTATGGCTGTCGGCCGACCTGAAGTTCGATATCAAGCAGAAGGGAGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTCGATACCTTGGCAAAGGAGGCTTCCATTCCCATCATCAACGGCCTGAG
 CGACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCTGCAAGAGCAC
 TACAGCAGCCTGAAGGGTCTGACCCTGTCATGGATTGGCGATGGAAACAATATT
 CTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCCA
 CTCCAAAAGGATACGAACCGGATGCATCCGTGACCAAGTTGGCGGAACAGTACG
 CGAAGGAGAACGGAACCAAGCTCCTGCTGACTAACGACCCGCTCGAGGCTGCGC
 ATGGGGGTAAACGTGCTGATTACGGACACCTGGATCTCCATGGGGCAGGAGGAAG
 AGAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAACC
 GCAAAAGTGGCAGCTTCGGACTGGACTTTCTGCAATTGCCTGCCGAGGAAGCCG
 GAGGAAGTCGACGACGAAGTGTCTACTCGCCTCGGTCCCTGGTGTTCCTCCGAGG
 CCGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCTTGCTGACTGACT
 ATAGCCCGCAGCTGCAGAAGCCTAAGTTCTAGATAAGTGA

>pARM1792 (SEQ ID NO: 168)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCGATCAAGCAGAAGGGCGAGTA
 CCTGCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCCTGGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTAACAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA

CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGA

>pARM1793 (SEQ ID NO: 169)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTGAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1794 (SEQ ID NO: 170)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCCTGGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA

AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAAGCCTGCTGA
 CCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGA

>pARM1795 (SEQ ID NO: 171)

ATGCTTTTCAACCTGAGAATCCTCTTGAACAATGCTGCTTTTCGGAATGGCCACA
 ACTTTATGGTTCGGAACCTCCGTTGCGGCCAGCCTTTACAAAACAAGGTCCAGCT
 GAAGGGCCGGGATTTGCTCACACTAAAGAACTTTACTGGAGAAGAGATCAAGTA
 CATGCTATGGCTGTCGGCCGACCTGAAGTTCGCTATCAAGCAGAAGGGAGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTCGATACCTTGGCAAAGGAGGCTTCCATTCCCATCATCAACGGCCTGAG
 CGACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCCTGCAAGAGCAC
 TACAGCAGCCTGAAGGGTCTGACCCTGTCATGGATTGGCGATGGAAACAATATT
 CTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCCA
 CTCCAAAAGGATACGAACCGGATGCATCCGTGACCAAGTTGGCGGAACAGTACG
 CGAAGGAGAACGGAACCAAGCTCCTGCTGACTAACGACCCGCTCGAGGCTGCGC
 ATGGGGGTAAACGTGCTGATTACGGACACCTGGATCTCCATGGGGCAGGAGGAAG
 AGAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAACC
 GCAAAAGTGGCAGCTTCGGACTGGACTTTCTGCAATTGCCTGCCGAGGAAGCCG
 GAGGAAGTCGACGACGAAGTGTCTACTCGCCTCGGTCCCTGGTGTTCCTCCGAGG
 CCGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCTTGCTGACTGACT
 ATAGCCCGCAGCTGCAGAAGCCTAAGTTCTAGATAAGTGA

>pARM1796 (SEQ ID NO: 172)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTCAGATGTGGCCAGCCGCTTCAAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCGCGATTAAGCAGAAGGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCTCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1797 (SEQ ID NO: 173)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCCTGGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGA

>pARM1798 (SEQ ID NO: 174)

ATGCTTTTCAACCTGAGAATCCTCTTGAACAATGCTGCTTTTCGGAATGGCCACA
 ACTTTATGGTTTCGGAACCTCCGTTGCGGCCAGCCTTTACAAAACAAGGTCCAGCT
 GAAGGGCCGGGATTTGCTCACACTAAAGAACTTTACTGGAGAAGAGATCAAGTA
 CATGCTATGGCTGTGCGCCGACCTGAAGTTCCGTATCAAGCAGAAGGGAGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTCGATACCTTGGCAAAGGAGGCTTCCATTCCCATCATCAACGGCCTGAG
 CGACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCCTGCAAGAGCAC
 TACAGCAGCCTGAAGGGTCTGACCCTGTCATGGATTGGCGATGGAAACAATATT
 CTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCCA
 CTCCAAAAGGATACGAACCGGATGCATCCGTGACCAAGTTGGCGGAACAGTACG
 CGAAGGAGAACGGAACCAAGCTCCTGCTGACTAACGACCCGCTCGAGGCTGCGC
 ATGGGGGTAAACGTGCTGATTACGGACACCTGGATCTCCATGGGGCAGGAGGAAG
 AGAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAACC
 GCAAAAGTGGCAGCTTCGGACTGGACTTTCCTGCATTGCCTGCCGAGGAAGCCG
 GAGGAAGTCGACGACGAAGTGTTCCTACTCGCCTCGGTCCCTGGTGTTCCTCGAGG
 CCGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCTTGTGCTGACTGACT
 ATAGCCCGCAGCTGCAGAAGCCTAAGTTCTAGATAAGTGA

>pARM1799 (SEQ ID NO: 175)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACAAGGTGCAGCT
GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC

CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
CTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCCGGCATGCACCTGCAGGC
CGCCACCCCCAAGGGCTACGAGCCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGAGGC
CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTCC
CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1800 (SEQ ID NO: 176)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACAAGGTGCAGCT
GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
CTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCCGGCATGCACCTGCAGGC
CGCCACCCCCAAGGGCTACGAGCCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGAGGC
CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTCC
CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1801 (SEQ ID NO: 177)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA

CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTCC
 CCGAGGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1802 (SEQ ID NO: 178)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTCC
 CCGAGGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGAA

>pARM1803 (SEQ ID NO: 179)

ATGGGCGTCTTCAACCTGCGGATCCTGCTGAACAACGCCGCCTTCCGGAACGGCC
 ACAACTTCATGGTCCGCAACTTCAGATGCGGCCAGCCCCTGCAGAACAAGGTGC
 AGCTGAAGGGCCGGGACCTGCTGACCCTGAAGAACTTCACCGGCGAAGAGATCA
 AGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGGATCAAGCAGAAGGGCG
 AGTACCTGCCCCTGCTGCAAGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGGA
 GCACCCGGACCCGGCTGAGCACCGAGACAGGCTTTGCCCTGCTGGGAGGCCACC
 CCTGCTTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCAGAGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGGGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAAGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAACACTACAGCTCCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCATCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCTGATGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAAGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCTGGA
 AGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAAGAGGAAAAGAAGAAGCGGCTGCAGGCCTTCCAGGGCTACCAGGTCACAA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC

GGAAGCCCGAAGAGGTGGACGACGAGGTGTTCTACAGCCCCCGGTCCCTGGTGT
 TCCCCGAGGCCGAGAACCGGAAGTGGACCATTATGGCCGTGATGGTGTCCCTGC
 TGACCGACTACTCCCCCAGCTGCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1804 (SEQ ID NO: 180)

ATGGGCGTCTTCAACCTGCGGATCCTGCTGAACAACGCCGCCTTCCGGAACGGCC
 ACAACTTCATGGTCCGCAACTTCAGATGCGGCCAGCCCCTGCAGAACAGGGTGC
 AGCTGAAGGGCCGGGACCTGCTGACCCTGAAGAACTTCACCGGCGAAGAGATCA
 GGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGGATCAAGCAGAAGGGCG
 AGTACCTGCCCCTGCTGCAAGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGGA
 GCACCCGGACCCGGCTGAGCACCGAGACAGGCTTTGCCCTGCTGGGAGGCCACC
 CCTGCTTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCAGAGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGGGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAAGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAACACTACAGCTCCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCATCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCTGATGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAAGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCTGGA
 AGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAAGAGGAAAAGAAGAAGCGGCTGCAGGCCTTCCAGGGCTACCAGGTCACAA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GGAAGCCCGAAGAGGTGGACGACGAGGTGTTCTACAGCCCCCGGTCCCTGGTGT
 TCCCCGAGGCCGAGAACCGGAAGTGGACCATTATGGCCGTGATGGTGTCCCTGC
 TGACCGACTACTCCCCCAGCTGCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1805 (SEQ ID NO: 181)

ATGCTGGTCTTCAACCTGCGGATCCTGCTGAACAACGCCGCTTCCGGAACGGCC
 ACAACTTCATGGTCCGCAACTTCAGATGCGGCCAGCCCCTGCAGAACAGGGTGC
 AGCTGAAGGGCCGGGACCTGCTGACCCTGAAGAACTTCACCGGCGAAGAGATCA
 GGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGGATCAAGCAGAAGGGCG
 AGTACCTGCCCCCTGCTGCAAGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGGA
 GCACCCGGACCCGGCTGAGCACCAGACAGGCTTTGCCCTGCTGGGAGGCCACC
 CCTGCTTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCAGAGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGGGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAAGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAACACTACAGCTCCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCATCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCTGATGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAAGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGA
 AGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAAGAGGAAAAGAAGAAGCGGCTGCAGGCCTTCCAGGGCTACCAGGTCACAA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GGAAGCCCGAAGAGGTGGACGACGAGGTGTTCTACAGCCCCCGGTCCCTGGTGT
 TCCCCGAGGCCGAGAACCGGAAGTGGACCATTATGGCCGTGATGGTGTCCCTGC
 TGACCGACTACTCCCCCAGCTGCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1806 (SEQ ID NO: 182)

ATGCTGTTCAACCTGAGGATCCTGCTGAACAACGCAGCTTTCAGGAACGGCCAC
 AACTTCATGGTGAGGAACTTCCGGTGCGGCCAGCCCCTGCAGAACAAAGGTGCAG
 CTGAAGGGCAGGGACCTGCTGACCCTGAAGAACTTCACCGGAGAGGAGATCAAG
 TACATGCTGTGGCTGAGCGCAGACCTGAAGTTCAGGATCAAGCAGAAGGGAGAG
 TACCTGCCCCCTGCTGCAGGGGAAGTCCCTGGGCATGATCTTCGAGAAGAGGAGT
 ACCAGGACCAGGCTGAGCACCAGAAACCGGCTTCGCCCTGCTGGGAGGACACCCC
 TGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGTCTGACCGAC
 ACCGCCAGGGTGCTGTCTAGCATGGCCGACGCCGTGCTGGCCAGGGTGTACAAG
 CAGTCAGACCTGGACACCCTGGCTAAGGAGGCCAGCATCCCCATCATCAACGGC
 CTGAGCGACCTGTACCACCCCATCCAGATCCTGGCTGACTACCTGACCCTGCAGG
 AGCACTACAGCTCTCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGGAACA
 ACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGG
 CCGCTACCCCCAAGGGTTACGAGCCCGACGCCAGCGTGACCAAGCTGGCAGAGC
 AGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGAGG
 CCGCCACGGAGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGACAGG
 AGGAGGAGAAGAAGAAGCGGCTGCAGGCCTTTCAGGGTTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCTGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCAGGA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACTCTCCAGGAGCCTGGTGTTC
 CCGAGGCCGAGAACAGGAAGTGGACCATCATGGCTGTGATGGTGTCCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGAA

>pARM1808 (SEQ ID NO: 183)

ATGCTGTTCAACCTGAGGATCCTGCTGAACAACGCAGCTTTCAGGAACGGCCAC
AACTTCATGGTGAGGAACTTCCGGTGCGGCCAGCCCCTGCAGAACAAAGGTGCAG
CTGAAGGGCAGGGACCTGCTGACCCTGAAGAACTTCACCGGAGAGGAGATCAAG
TACATGCTGTGGCTGAGCGCAGACCTGAAGTTCAGGATCAAGCAGAAGGGAGAG
TACCTGCCCCCTGCTGCAGGGGAAGTCCCTGGGCATGATCTTCGAGAAGAGGAGT
ACCAGGACCAGGCTGAGCACCGAAACCGGCTTCGCCCTGCTGGGAGGACACCCC
TGCTTCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATA
CCGCCCCGGGTGTTATCAAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAAC
AGTCCGATCTGGACACTCTGGCTAAGGAGGCCAGCATCCCCATCATCAACGGCCT
GAGCGACCTGTACCACCCCATCCAGATCCTGGCTGACTACCTGACCCTGCAGGAG
CACTACAGCTCTCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGGAACAAC
ATCCTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCG
CCACGCCAAAAGGATACGAACCGGATGCGCCCCGTGACAAAGTTGGCGGAACAGT
ACGCTAAGGAGAACGGAACCAAGCTGCTGCTGACCAACGACCCCCCTGGAGGCCG
CCCACGGAGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGACAGGAGG
AGGAGAAGAAGAAGCGGCTGCAGGCTTTCAGGGTTACCAGGTGACCATGAAGA
CCGCCAAGGTGGCTGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCAGGAAGC
CCGAGGAGGTGGACGACGAGGTGTTCTACTCTCCAGGAGCCTGGTGTTCCTCCG
AGGCCGAGAACAGGAAGTGGACCATCATGGCTGTGATGGTGTCCCTGCTGACCG
ACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGAA

>pARM1809 (SEQ ID NO: 184)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAAGGTGCAGCT
GAAGGGCCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC

CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGGCCACCCCTG
CTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGAGGC
CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGAA

>pARM1816 (SEQ ID NO: 185)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACC GCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTGAA

>pARM1822 (SEQ ID NO: 186)

ATGCTTTTCAACTTGAGAATCCTGCTGAACAACGCCGCCTTTCGCAACGGTCACA
 ATTTTATGGTCAGAACTTCAGATGCGGACAGCCCCTCCAAAACAAGGTCCAGCT
 GAAGGGCCGCGATCTCCTCACCTGAAGAACTTCACGGGGGAGGAGATCAAGTA
 CATGCTGTGGCTCTCCGCTGACCTGAAGTTCAGGATCAAGCAGAAGGGAGAATA
 TCTGCCGCTGCTGCAAGGGAAGTCCCTGGGGATGATTTTCGAGAAGCGGAGCAC
 CCGGACTCGGCTCTCCACTGAACTGGTTTCGCCCTTCTGGGCGGTACCCCTGC
 TTCCTGACCACTCAAGACATTCACCTCGGAGTGAACGAGTCCTTGACTGACACCG
 CCCGGGTGCTGTGCGAGCATGGCAGACGCCGTGCTAGCCCGCGTGTACAAGCAGT
 CAGACCTCGATACCCTGGCCAAGGAGGCTTCGATCCCGATCATCAACGGGTTGTC
 CGACCTGTACCACCCGATTCAGATTCTCGCCGACTACCTCACCTGCAAGAGCAT
 TACAGCTCCCTGAAGGGGCTTACCCTGTCTGGATTGGCGACGGAAACAACATCC

TGCACTCCATTATGATGTCGGCGGCCAAGTTCGGCATGCACCTCCAAGCCGCGAC
 CCCTAAGGGTTACGAACCAGACGCGTCAGTGACTAAGCTGGCCGAACAGTACGC
 AAAGGAAAATGGCACGAAGCTGCTCCTGACCAACGATCCGTTGGAAGCCGCCCA
 TGCGGAAATGTGCTCATCACCGACACCTGGATCTCGATGGGACAGGAGGAAGA
 GAAGAAGAAGCGGCTGCAGGCGTTCCAGGGCTACCAGGTCACCATGAAAAGTGC
 CAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTTCCGCGCAAGCCTGAG
 GAGGTGGACGATGAAGTGTCTACTCTCCACGGTCCCTGGTGTTCCTCCGAGGCGG
 AGAACCGCAAATGGACCATCATGGCTGTGATGGTCAGCCTGCTGACCGATTACA
 GCCCTCAGTTGCAAAAGCCGAAGTTTGA

>pARM1823 (SEQ ID NO: 187)

ATGCTGTTCAACCTCCGCATCCTCCTCAACAACGCCGCATTCAGAAACGGGCACA
 ACTTCATGGTCAGAAACTTCCGCTGCGGGCAACCCCTACAAAACAAGGTCCAGC
 TCAAGGGGCGGGACCTCCTGACCCTGAAGAACTTCACCGGCGAAGAGATCAAGT
 ACATGCTGTGGCTCTCCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGAGAGT
 ACCTCCCGCTGCTGCAAGGGAAGTCGCTGGGGATGATCTTCGAGAAGCGGTCAA
 CCAGAACCCGGCTGTCAACCGAAACCGGGTTCGCACTGCTGGGGGGACACCCGT
 GCTTCCTGACCACCAAGACATCCACCTGGGAGTGAACGAATCGCTGACCGACA
 CCGCCCGCGTGCTGAGCTCAATGGCGGACGCCGTGCTGGCCCGCGTGTACAAGC
 AGTCCGACCTGGACACCCTGGCCAAGGAAGCGTCCATCCCGATCATCAACGGAC
 TGTCCGACCTGTACCACCCGATCCAGATCCTGGCAGACTACCTGACCCTGCAAGA
 ACACTACAGCTCCCTGAAGGGCCTGACCCTGTCATGGATCGGGGACGGGAACAA
 CATCCTGCACTCCATAATGATGTCAGCCGCCAAGTTCGGAATGCACCTCCAAGCC
 GCAACCCCGAAGGGCTACGAACCGGACGCATCAGTGACCAAACCTGGCCGAGCAG
 TACGCCAAGGAAAACGGCACCAAGCTCCTGCTGACCAACGACCCGCTGGAGGCC
 GCACACGGGGGGAACGTGCTGATCACCGACACCTGGATCTCCATGGGACAGGAG
 GAGGAAAAGAAGAAGCGGCTGCAGGCGTTCCAGGGGTACCAGGTCACCATGAA
 AACCGCGAAGGTTCGCGGCATCAGACTGGACCTTCCTGCACTGCCTGCCCCGGAA
 GCCGGAAGAGGTGGACGACGAGGTGTTCTACTCGCCGCGCTCGCTGGTGTTCCTCC
 GAGGCGGAGAACAGGAAGTGGACCATCATGGCGGTGATGGTCAGCCTCCTGACC
 GACTACTCGCCGCAGCTGCAGAAGCCGAAGTTCTGA

>pARM1840 (SEQ ID NO: 188)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCCTGAA

GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCTCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTGAATAAGTAGA

>pARM1841 (SEQ ID NO: 189)

ATGCTTTTCAACCTGAGAATCCTCTTGAACAATGCTGCTTTTCGGAATGGCCACA
 ACTTTATGGTTCGGAACCTTCCGTTGCGGCCAGCCTTTACAAAACAAGGTCCAGCT
 GAAGGGCCGGGATTTGCTCACACTAAAGAACTTTACTGGAGAAGAGATCAAGTA
 CATGCTATGGCTGTCGGCCGACCTGAAGTTCGGTATCAAGCAGAAGGGAGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTCGATACTTGGCAAAGGAGGCTTCCATTCCCATCATCAACGGCCTGAG
 CGACCTGTACCACCCAATCCAAATCCTGGCTGACTACCTGACCCTGCAAGAGCAC
 TACAGCAGCCTGAAGGGTCTGACCCTGTCATGGATTGGCGATGGAAACAATATT
 CTGCACTCCATCATGATGTCCGCCGCGAAGTTCGGAATGCATCTGCAAGCCGCCA
 CTCCAAAAGGATACGAACCGGATGCATCCGTGACCAAGTTGGCGGAACAGTACG
 CGAAGGAGAACGGAACCAAGCTCCTGCTGACTAACGACCCGCTCGAGGCTGCGC
 ATGGGGGTAAACGTGCTGATTACGGACACCTGGATCTCCATGGGGCAGGAGGAAG
 AGAAGAAGAAGAGACTGCAGGCATTCCAGGGGTACCAGGTCACCATGAAAACC
 GCAAAAGTGGCAGCTTCGGACTGGACTTTCTGCAATTGCCTGCCGAGGAAGCCG
 GAGGAAGTCGACGACGAAGTGTTCTACTCGCCTCGGTCCCTGGTGTTCCTCCGAGG
 CCGAAAACCGGAAGTGGACCATCATGGCCGTGATGGTGTCTTGCTGACTGACT
 ATAGCCCGCAGCTGCAGAAGCCTAAGTTCTGAATAAGTAGA

>pARM1842 (SEQ ID NO: 190)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCTGCAGAACAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTACC GGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCCTGGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGAATAAGTAGA

>pARM1843 (SEQ ID NO: 191)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCATAACG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCCTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1844 (SEQ ID NO: 192)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCATAACG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCCTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1845 (SEQ ID NO: 193)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC

AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACG
CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
CAAGGTGCTGCCTCCGACTGGACCTTCCCTGCACTGCCTGCCTCGCAAGCCTGAA
GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCCTGCTCACCGACTAC
AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1846 (SEQ ID NO: 194)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAACG
CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
CAAGGTGCTGCCTCCGACTGGACCTTCCCTGCACTGCCTGCCTCGCAAGCCTGAA
GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCCTGCTCACCGACTAC
AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1847 (SEQ ID NO: 195)

ATGCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCTTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCTGGATCGGCGACGGCAACAACATTC

TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTGCTGCCTCCGACTGGACCTTCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCCTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1882 (SEQ ID NO: 196)

ATGGTGTTCAACCTCCGCATCCTCCTCAACAACGCCGCGATTTCAGAAACGGGCACA
 ACTTCATGGTCAGAACTTCCGCTGCGGGCAACCCCTACAAAACAAGGTCCAGC
 TCAAGGGGCGGGACCTCCTGACCCTGAAGAACTTCACCGGCGAAGAGATCAAGT
 ACATGCTGTGGCTCTCCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGAGAGT
 ACCTCCCGCTGCTGCAAGGGAAGTCGCTGGGGATGATCTTCGAGAAGCGGTCAA
 CCAGAACCCGGCTGTCAACCGAAACCGGGTTTCGCACTGCTGGGGGGACACCCGT
 GCTTCCTGACCACCCAAGACATCCACCTGGGAGTGAACGAATCGCTGACCGACA
 CCGCCCGCGTGCTGAGCTCAATGGCGGACGCCGTGCTGGCCCGCGTGTACAAGC
 AGTCCGACCTGGACACCCTGGCCAAGGAAGCGTCCATCCCGATCATCAACGGAC
 TGTCCGACCTGTACCACCCGATCCAGATCCTGGCAGACTACCTGACCCTGCAAGA
 AACTACAGCTCCCTGAAGGGCCTGACCCTGTCATGGATCGGGGACGGGAACAA
 CATCCTGCACTCCATAATGATGTCAGCCGCCAAGTTCGGAATGCACCTCCAAGCC
 GCAACCCCGAAGGGCTACGAACCGGACGCATCAGTGACCAAACCTGGCCGAGCAG
 TACGCCAAGGAAAACGGCACCAAGCTCCTGCTGACCAACGACCCGCTGGAGGCC
 GCACACGGGGGGAACGTGCTGATCACCGACACCTGGATCTCCATGGGACAGGAG
 GAGGAAAAGAAGAAGCGGCTGCAGGCGTTCCAGGGGTACCAGGTCAACATGAA
 AACCGCGAAGGTGCGGGCATCAGACTGGACCTTCTGCACTGCCTGCCCCGGAA
 GCCGGAAGAGGTGGACGACGAGGTGTTCTACTCGCCGCGCTCGCTGGTGTTCCCC
 GAGGCGGAGAACAGGAAGTGGACCATCATGGCGGTGATGGTCAGCCTCCTGACC
 GACTACTCGCCGACGCTGCAGAAGCCGAAGTTCTGA

>pARM1883 (SEQ ID NO: 197)

ATGGTGTTC AACCTCCGCATCCTCCTCAACAACGCCGCATTCAGAAACGGGCACA
 ACTTCATGGTCAGAACTTCCGCTGCGGGCAACCCCTACAAAACCGGGTCCAGCT
 CAAGGGGCGGGACCTCCTGACCCTGAAGAACTTCACCGGCGAAGAGATCAAGTA
 CATGCTGTGGCTCTCCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGAGAGTA
 CCTCCCGCTGCTGCAAGGGAAGTCGCTGGGGATGATCTTCGAGAAGCGGTCAAC
 CAGAACCCGGCTGTCAACCGAAACCGGGTTCGCACTGCTGGGGGGACACCCGTG
 CTTCTGACCACCCAAGACATCCACCTGGGAGTGAACGAATCGCTGACCGACAC
 CGCCCGCTGCTGAGCTCAATGGCGGACGCCGTGCTGGCCCGCTGTACAAGCA
 GTCCGACCTGGACACCCTGGCCAAGGAAGCGTCCATCCCGATCATCAACGGACT
 GTCCGACCTGTACCACCCGATCCAGATCCTGGCAGACTACCTGACCCTGCAAGAA
 CACTACAGCTCCCTGAAGGGCCTGACCCTGTCATGGATCGGGGACGGGAACAAC
 ATCCTGCACTCCATAATGATGTCAGCCGCCAAGTTCGGAATGCACCTCCAAGCCG
 CAACCCCGAAGGGCTACGAACCGGACGCATCAGTGACCAAACCTGGCCGAGCAGT
 ACGCCAAGGAAAACGGCACCAAGCTCCTGCTGACCAACGACCCGCTGGAGGCCG
 CACACGGGGGGAACGTGCTGATCACCGACACCTGGATCTCCATGGGACAGGAGG
 AGGAAAAGAAGAAGCGGCTGCAGGCGTTCCAGGGGTACCAGGTCACCATGAAA
 ACCGCGAAGGTCGCGGCATCAGACTGGACCTTCTGCACTGCCTGCCCCGGAAG

CCGGAAGAGGTGGACGACGAGGTGTTCTACTCGCCGCGCTCGCTGGTGTTCCTCCG
 AGGCGGAGAACAGGAAGTGGACCATCATGGCGGTGATGGTCAGCCTCCTGACCG
 ACTACTCGCCGCAGCTGCAGAAGCCGAAGTTCTGA

>pARM1884 (SEQ ID NO: 198)

ATGGTGTTC AACCTCCGCATCCTCCTCAACAACGCCGCATTCAGAAACGGGCACA
 ACTTCATGGTCAGAACTTCCGCTGCGGGCAACCCCTACAAAACCGGGTCCAGCT
 CAAGGGGCGGGACCTCCTGACCCTGAAGAACTTCACCGGCGAAGAGATCCGGTA
 CATGCTGTGGCTCTCCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGAGAGTA
 CCTCCCGCTGCTGCAAGGGAAGTCGCTGGGGATGATCTTCGAGAAGCGGTCAAC
 CAGAACCCGGCTGTCAACCGAAACCGGGTTCGCACTGCTGGGGGGACACCCGTG
 CTTCTGACCACCCAAGACATCCACCTGGGAGTGAACGAATCGCTGACCGACAC
 CGCCCGCTGCTGAGCTCAATGGCGGACGCCGTGCTGGCCCGCTGTACAAGCA
 GTCCGACCTGGACACCCTGGCCAAGGAAGCGTCCATCCCGATCATCAACGGACT
 GTCCGACCTGTACCACCCGATCCAGATCCTGGCAGACTACCTGACCCTGCAAGAA
 CACTACAGCTCCCTGAAGGGCCTGACCCTGTCATGGATCGGGGACGGGAACAAC
 ATCCTGCACTCCATAATGATGTCAGCCGCCAAGTTCGGAATGCACCTCCAAGCCG
 CAACCCCGAAGGGCTACGAACCGGACGCATCAGTGACCAAACCTGGCCGAGCAGT
 ACGCCAAGGAAAACGGCACCAAGCTCCTGCTGACCAACGACCCGCTGGAGGCCG
 CACACGGGGGGAACGTGCTGATCACCGACACCTGGATCTCCATGGGACAGGAGG
 AGGAAAAGAAGAAGCGGCTGCAGGCGTTCCAGGGGTACCAGGTCACCATGAAA
 ACCGCGAAGGTCGCGGCATCAGACTGGACCTTCTGCACTGCCTGCCCCGGAAG
 CCGGAAGAGGTGGACGACGAGGTGTTCTACTCGCCGCGCTCGCTGGTGTTCCTCCG
 AGGCGGAGAACAGGAAGTGGACCATCATGGCGGTGATGGTCAGCCTCCTGACCG
 ACTACTCGCCGCAGCTGCAGAAGCCGAAGTTCTGA

>pARM1885 (SEQ ID NO: 199)

ATGCTGGTCAACCTCCGCATCCTCCTCAACAACGCCGCATTCAGAAACGGGCACA
 ACTTCATGGTCAGAAACTTCCGCTGCGGGCAACCCCTACAAAACAAGGTCCAGC
 TCAAGGGGCGGGACCTCCTGACCCTGAAGAACTTCACCGGCGAAGAGATCAAGT
 ACATGCTGTGGCTCTCCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGAGAGT
 ACCTCCCGCTGCTGCAAGGGAAGTCGCTGGGGATGATCTTCGAGAAGCGGTCAA
 CCAGAACCCGGCTGTCAACCGAAACCGGGTTTCGCACTGCTGGGGGGACACCCGT
 GCTTCCTGACCACCCAAGACATCCACCTGGGAGTGAACGAATCGCTGACCGACA
 CCGCCCGCGTGCTGAGCTCAATGGCGGACGCCGTGCTGGCCCGCGTGTAACAGC
 AGTCCGACCTGGACACCCTGGCCAAGGAAGCGTCCATCCCGATCATCAACGGAC
 TGTCCGACCTGTACCACCCGATCCAGATCCTGGCAGACTACCTGACCCTGCAAGA
 AACTACAGCTCCCTGAAGGGCCTGACCCTGTCATGGATCGGGGACGGGAACAA
 CATCCTGCACTCCATAATGATGTCAGCCGCCAAGTTCGGAATGCACCTCCAAGCC
 GCAACCCCGAAGGGCTACGAACCGGACGCATCAGTGACCAAACCTGGCCGAGCAG
 TACGCCAAGGAAAACGGCACCAAGCTCCTGCTGACCAACGACCCGCTGGAGGCC
 GCACACGGGGGGGAACGTGCTGATCACCGACACCTGGATCTCCATGGGACAGGAG
 GAGGAAAAGAAGAAGCGGCTGCAGGCGTTCCAGGGGTACCAGGTCACCATGAA
 AACCGCGAAGGTCGCGGCATCAGACTGGACCTTCCTGCACTGCCTGCCCCGGAA
 GCCGGAAGAGGTGGACGACGAGGTGTTCTACTCGCCGCGCTCGCTGGTGTTCCTCC
 GAGGCGGAGAACAGGAAGTGGACCATCATGGCGGTGATGGTCAGCCTCCTGACC
 GACTACTCGCCGCAGCTGCAGAAGCCGAAGTTCTGA

>pARM1886 (SEQ ID NO: 200)

ATGCTGGTCAACCTCCGCATCCTCCTCAACAACGCCGCATTCAGAAACGGGCACA
 ACTTCATGGTCAGAAACTTCCGCTGCGGGCAACCCCTACAAAACCGGGTCCAGCT
 CAAGGGGCGGGACCTCCTGACCCTGAAGAACTTCACCGGCGAAGAGATCAAGTA
 CATGCTGTGGCTCTCCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGAGAGTA
 CCTCCCGCTGCTGCAAGGGAAGTCGCTGGGGATGATCTTCGAGAAGCGGTCAAC
 CAGAACCCGGCTGTCAACCGAAACCGGGTTTCGCACTGCTGGGGGGACACCCGTG
 CTTCTGACCACCCAAGACATCCACCTGGGAGTGAACGAATCGCTGACCGACAC
 CGCCCGCGTGCTGAGCTCAATGGCGGACGCCGTGCTGGCCCGCGTGTAACAGCA
 GTCCGACCTGGACACCCTGGCCAAGGAAGCGTCCATCCCGATCATCAACGGACT
 GTCCGACCTGTACCACCCGATCCAGATCCTGGCAGACTACCTGACCCTGCAAGAA
 CACTACAGCTCCCTGAAGGGCCTGACCCTGTCATGGATCGGGGACGGGAACAAC
 ATCCTGCACTCCATAATGATGTCAGCCGCCAAGTTCGGAATGCACCTCCAAGCCG
 CAACCCCGAAGGGCTACGAACCGGACGCATCAGTGACCAAACCTGGCCGAGCAGT
 ACGCCAAGGAAAACGGCACCAAGCTCCTGCTGACCAACGACCCGCTGGAGGCCG
 CACACGGGGGGGAACGTGCTGATCACCGACACCTGGATCTCCATGGGACAGGAGG
 AGGAAAAGAAGAAGCGGCTGCAGGCGTTCCAGGGGTACCAGGTCACCATGAAA
 ACCGCGAAGGTCGCGGCATCAGACTGGACCTTCCTGCACTGCCTGCCCCGGAAG
 CCGGAAGAGGTGGACGACGAGGTGTTCTACTCGCCGCGCTCGCTGGTGTTCCTCC
 AGGCGGAGAACAGGAAGTGGACCATCATGGCGGTGATGGTCAGCCTCCTGACCG
 ACTACTCGCCGCAGCTGCAGAAGCCGAAGTTCTGA

>pARM1887 (SEQ ID NO: 201)

ATGCTGGTCAACCTCCGCATCCTCCTCAACAACGCCGCATTCAGAAACGGGCACA
 ACTTCATGGTCAGAAACTTCCGCTGCGGGCAACCCCTACAAAACCGGGTCCAGCT
 CAAGGGGCGGGACCTCCTGACCCTGAAGAACTTCACCGGCGAAGAGATCCGGTA
 CATGCTGTGGCTCTCCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGAGAGTA
 CCTCCCGCTGCTGCAAGGGAAGTCGCTGGGGATGATCTTCGAGAAGCGGTCAAC
 CAGAACCCGGCTGTCAACCGAAACCGGGTTCGCACTGCTGGGGGGACACCCGTG
 CTTCTGACCACCCAAGACATCCACCTGGGAGTGAACGAATCGCTGACCGACAC
 CGCCCGCTGCTGAGCTCAATGGCGGACGCCGTGCTGGCCCGCTGTACAAGCA
 GTCCGACCTGGACACCCTGGCCAAGGAAGCGTCCATCCCGATCATCAACGGACT
 GTCCGACCTGTACCACCCGATCCAGATCCTGGCAGACTACCTGACCCTGCAAGAA
 CACTACAGCTCCCTGAAGGGCCTGACCCTGTATGGATCGGGGACGGGAACAAC
 ATCTGCACTCCATAATGATGTGACCCGCCAAGTTCGGAATGCACCTCCAAGCCG
 CAACCCCGAAGGGCTACGAACCGGACGCATCAGTGACCAAACCTGGCCGAGCAGT
 ACGCCAAGGAAAACGGCACCAAGCTCCTGCTGACCAACGACCCGCTGGAGGCCG
 CACACGGGGGGAACGTGCTGATCACCGACACCTGGATCTCCATGGGACAGGAGG
 AGGAAAAGAAGAAGCGGCTGCAGGCGTTCCAGGGGTACCAGGTCACCATGAAA
 ACCGCGAAGGTCGCGGCATCAGACTGGACCTTCCTGCACTGCCTGCCCCGGAAG
 CCGGAAGAGGTGGACGACGAGGTGTTCTACTCGCCGCGCTCGCTGGTGTTCCCCG
 AGGCGGAGAACAGGAAGTGGACCATCATGGCGGTGATGGTCAGCCTCCTGACCG
 ACTACTCGCCGACGTGCAGAAGCCGAAGTTCTGA

>pARM1888 (SEQ ID NO: 202)

ATGCTGGTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCTGCAGAACAAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC

CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTCC
 CCGAGGCCGAGAACC GCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1889 (SEQ ID NO: 203)

ATGCTGGTCAACCTGCGCATCCTGCTGAACAACGCCGCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACCGGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1890 (SEQ ID NO: 204)

ATGCTTGTCATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCTTTTCTACTGAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC

TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTCGCTGCCTCCGACTGGACCTTCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCCTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1891 (SEQ ID NO: 205)

ATGCTTGTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCCACA
 ACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACCGGGTCCAGCT
 GAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAAGTA
 CATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGAATA
 CCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCAACC
 AGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCTGCT
 TCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATACCG
 CCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACAGT
 CCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTTAG
 TGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCCCTGCAAGAACAC
 TACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACATTC
 TCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGCCAC
 GCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTACGC
 TAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGCCCA
 CGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGAAGA
 GAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAACCGC
 CAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCCTGAA
 GAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGAGGCC
 GAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGACTAC
 AGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1898 (SEQ ID NO: 206)

ATGGGCCTTGTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCC
 ACAACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGGTCC
 AGCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCA
 AGTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGG
 AATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTC
 AACCAGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCC
 TGCTTCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATA
 CCGCCCGGGTGTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAAC
 AGTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCT
 TAGTGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCCCTGCAAGAA
 CACTACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAAC
 ATTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCG
 CCACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGT
 ACGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAG
 CCCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAG
 AAGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAA
 ACCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGC

CTGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCG
 AGGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCG
 ACTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1899 (SEQ ID NO: 207)

ATGGGCCTTGTCATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCC
 ACAACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACCGGGTCCA
 GCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAA
 GTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGA
 ATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCA
 ACCAGGACCCGCTTTTCTACTGAACTGGGTTTCGCGCTGCTCGGTGGCCACCCCT
 GCTTCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAC
 CGCCCGGGTGTATATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACA
 GTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCCTATTATCAACGGCCTT
 AGTGACCTCTACCATCCGATTTCAGATCCTGGCCGATTACCTCACCCCTGCAAGAAC
 ACTACAGCTCCCTGAAGGGTCTGACATTGTCTGATCGGCGACGGCAACAACA
 TTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGC
 CACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTA
 CGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGC
 CCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGA
 AGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAA
 CCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCC
 TGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCGA
 GGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGA
 CTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1900 (SEQ ID NO: 208)

ATGGGCGGACTTGTCATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACG
 GCCACAACCTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAAACAAGG
 TCCAGCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGA
 TCAAGTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGG
 GGAATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCG
 CTCAACCAGGACCCGCTTTTCTACTGAACTGGGTTTCGCGCTGCTCGGTGGCCAC
 CCCTGCTTCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCG
 ATACCGCCCGGGTGTATATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACA
 AACAGTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCCTATTATCAACGG
 CCTTAGTGACCTCTACCATCCGATTTCAGATCCTGGCCGATTACCTCACCCCTGCAA
 GAACACTACAGCTCCCTGAAGGGTCTGACATTGTCTGATCGGCGACGGCAAC
 AACATTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAG
 CCGCCACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGC
 AGTACGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAG
 CAGCCCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGG
 AAGAAGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATG
 AAAACCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCA
 AGCCTGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCC
 CCGAGGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCA
 CCGACTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAG

>pARM1903 (SEQ ID NO: 209)

ATGGCCCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCC
 ACAACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAGGCAAGGTCC
 AGCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCA
 AGTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGG
 AATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTC
 AACCAGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCC
 TGCTTCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCATA
 CCGCCCGGGTGTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAAC
 AGTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCT
 TAGTGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAA
 CACTACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAAC
 ATTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCG
 CCACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGT
 ACGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAG
 CCCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAG
 AAGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAA
 ACCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGC
 CTGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCG
 AGGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACC
 ACTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1904 (SEQ ID NO: 210)

ATGGCCCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCC
 ACAACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAGGCCGGGTCCA
 GCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAA
 GTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGA
 ATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCA
 ACCAGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCT
 GCTTCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCATA
 CGCCCGGGTGTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACA
 GTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTT
 AGTGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAAC
 ACTACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACA
 TTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGC
 CACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTA
 CGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGC
 CCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGA
 AGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAA
 CCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCC
 TGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGA
 GGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACC
 CTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1905 (SEQ ID NO: 211)

ATGGCCCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCC
 ACAACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAGGCCGGGTCCA
 GCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAG
 GTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGA
 ATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCA

ACCAGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCCT
GCTTCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAC
CGCCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACA
GTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTT
AGTGACCTCTACCATCCGATTACAGATCCTGGCCGATTACCTCACCTGCAAGAAC
ACTACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAACA
TTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGC
CACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTA
CGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGC
CCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGA
AGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAA
CCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCC
TGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCTCCGA
GGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGA
CTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1906 (SEQ ID NO: 212)

ATGGCCCTTGTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCC
ACAACCTTCATGGTCCGGAACCTCAGATGTGGCCAGCCGCTTCAAGGCAGGGTCC
AGCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCA
AGTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCGCGATTAAGCAGAAGGGGG
AATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTC
AACCAGGACCCGCCTTTCTACTGAAACTGGGTTCGCGCTGCTCGGTGGCCACCCC
TGCTTCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATA
CCGCCCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAAC
AGTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCT
TAGTGACCTCTACCATCCGATTACAGATCCTGGCCGATTACCTCACCTGCAAGAA
CACTACAGCTCCCTGAAGGGTCTGACATTGTCCTGGATCGGCGACGGCAACAAC
ATTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCG
CCACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGT
ACGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAG
CCCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAG
AAGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAA
ACCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGC
CTGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCTCCG
AGGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCG
ACTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAAGTGAATAGA

>pARM1907 (SEQ ID NO: 213)

ATGGCCCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCC
 ACAACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAGTCAAGGTCCA
 GCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAA
 GTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGA
 ATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCA
 ACCAGGACCCGCCTTTCTACTGAAACTGGGTTTCGCGCTGCTCGGTGGCCACCCCT
 GCTTCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAC
 CGCCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACA
 GTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTT
 AGTGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAAC
 ACTACAGCTCCCTGAAGGGTCTGACATTGTCTGGATCGGCGACGGCAACAACA

TTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGC
 CACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTA
 CGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGC
 CCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGA
 AGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAA
 CCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCC
 TGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGA
 GGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGA
 CTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1908 (SEQ ID NO: 214)

ATGGCCCTTTTCAATCTCCGCATCCTCCTTAACAACGCCGCGTTTAGAAACGGCC
 ACAACTTCATGGTCCGGAACCTTCAGATGTGGCCAGCCGCTTCAAGTCAGGGTCCA
 GCTGAAGGGCCGGGATCTTCTGACCCTGAAGAACTTTACTGGCGAAGAGATCAA
 GTACATGCTCTGGCTCTCCGCGGACTTGAAGTTCCGCATTAAGCAGAAGGGGGA
 ATACCTTCCGCTGCTTCAAGGAAAGAGCCTCGGCATGATCTTTGAGAAGCGCTCA
 ACCAGGACCCGCCTTTCTACTGAAACTGGGTTTCGCGCTGCTCGGTGGCCACCCCT
 GCTTCCTGACGACCCAGGACATCCACCTCGGAGTGAACGAATCCCTCACCGATAC
 CGCCCGGGTGTTATCGAGCATGGCAGATGCCGTGCTGGCCAGGGTGTACAAACA
 GTCCGATCTGGACACTCTGGCCAAGGAGGCGTCAATTCCTATTATCAACGGCCTT
 AGTGACCTCTACCATCCGATTCAGATCCTGGCCGATTACCTCACCTGCAAGAAC
 ACTACAGCTCCCTGAAGGGTCTGACATTGTCTGGATCGGCGACGGCAACAACA
 TTCTCCATTCCATCATGATGTCCGCCGCAAAATTCGGCATGCATCTTCAAGCCGC
 CACGCCGAAGGGTTACGAGCCCGACGCTTCCGTGACTAAGCTCGCCGAGCAGTA
 CGCTAAGGAGAACGGAACCAAGCTTCTGCTGACTAACGACCCACTAGAAGCAGC
 CCACGGGGGCAACGTGCTTATTACTGACACCTGGATCTCCATGGGCCAGGAAGA
 AGAGAAAAAGAAGCGGCTGCAGGCGTTCCAGGGATATCAGGTCACCATGAAAA
 CCGCCAAGGTCGCTGCCTCCGACTGGACCTTCCTGCACTGCCTGCCTCGCAAGCC
 TGAAGAAGTGGACGACGAGGTGTTCTACTCGCCACGGAGCCTCGTGTTCCCCGA
 GGCCGAGAATAGAAAGTGGACCATCATGGCCGTGATGGTGTCACTGCTCACCGA
 CTACAGCCCGCAGCTTCAGAAGCCCAAGTTCTAGATAAGTGAA

>pARM1915 (SEQ ID NO: 215)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA

AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1916 (SEQ ID NO: 216)

ATGGGAGTATTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCC
 ACAACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACAGGTGC
 AGCTGAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCAGGAGATCA
 AGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCG
 AGTACCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCA
 GCACCCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACC
 CCTGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCA
 GGCCGCCACCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGA
 GGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCAGGGCTACCAGGTGACCA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GCAAGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGT
 TCCCCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGC
 TGACCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1917 (SEQ ID NO: 217)

ATGGGAGTATTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCC
 ACAACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACCGGGTGC
 AGCTGAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCC
 GGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCG
 AGTACCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCA
 GCACCCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACC
 CCTGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTTCGGCATGCACCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGA
 GGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GCAAGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGT
 TCCCCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGC
 TGACCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1918 (SEQ ID NO: 218)

ATGCTGGTATTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCC
 ACAACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACCGGGTGC
 AGCTGAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCC
 GGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCG
 AGTACCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCA
 GCACCCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACC
 CCTGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTTCGGCATGCACCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGA
 GGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GCAAGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGT
 TCCCCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGC
 TGACCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1919 (SEQ ID NO: 219)

ATGGGAGTATTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCC
 ACAACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACAAGGTGC
 AGCTGAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCA
 AGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCG
 AGTACCTGCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCA
 GCACCCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGGCCACC
 CCTGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAGGAGGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGA
 GGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GCAAGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGT
 TCCCCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGC
 TGACCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1920 (SEQ ID NO: 220)

ATGGGAGTATTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCC
 ACAACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGAACCAGGTGC
 AGCTGAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCC
 GGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCG
 AGTACCTGCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCA

GCACCCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGGCCACC
 CCTGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAGGAGGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGA
 GGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GCAAGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGT
 TCCCCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGC
 TGACCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1921 (SEQ ID NO: 221)

ATGCTGGTATTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCC
 ACAACTTCATGGTGCACAACTTCCGCTGCGGCCAGCCCCCTGCAGAACCGGGTGC
 AGCTGAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCC
 GGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCG
 AGTACCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCA
 GCACCCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGGCCACC
 CCTGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACA
 AGCAGAGCGACCTGGACACCTGGCCAAGGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGA
 GGCCGCCACGGCGGCAACGTGCTGATACCGACACCTGGATCAGCATGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GCAAGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGT
 TCCCCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGC
 TGACCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1925 (SEQ ID NO: 222)

ATGTTGTTCAACTTGAGGATCTTGTTGAACAACGCCGCCTTCAGGAACGGACACA
 ACTTCATGGTAAGGAACTTCAGGTGCGGACAGCCCTTGCAGAACAAAGTACAGT
 TGAAAGGAAGGGACTTGTTGACATTGAAAACTTCACAGGAGAAGAAATCAAAT
 ACATGTTGTGGTTGTCGGCCGACTTGAAATTCAGGATCAAACAGAAAGGAGAAT
 ACTTGCCCTTGTTGCAGGGAAAATCGTTGGGAATGATCTTCGAAAAAAGGTCGA
 CAAGGACAAGGTTGTCGACAGAAACAGGATTCGCCTTGTTGGGAGGACACCCCT
 GCTTCTTGACAACACAGGACATCCACTTGGGAGTAAACGAATCGTTGACAGACA
 CAGCCAGGGTATTGTCGTCGATGGCCGACGCCGTATTGGCCAGGGTATACAAAC
 AGTCGGACTTGGACACATTGGCCAAAGAAGCCTCGATCCCCATCATCAACGGAT
 TGTCGGACTTGTACCACCCCATCCAGATCTTGGCCGACTACTTGACATTGCAGGA
 ACATACTCGTCGTTGAAAGGATTGACATTGTCGTGGATCGGAGACGGAAACAA

CATCTTGCACTCGATCATGATGTCGGCCGCCAAATTCGGAATGCACTTGCAGGCC
 GCCACACCCAAAGGATACGAACCCGACGCCTCGGTAACAAAATTGGCCGAACAG
 TACGCCAAAGAAAACGGAACAAAATTGTTGTTGACAAACGACCCCTTGGAAGCC
 GCCACGGAGGAAACGTATTGATCACAGACACATGGATCTCGATGGGACAGGAA
 GAAGAAAAAAAAAAAAAGGTTGCAGGCCTTCCAGGGATACCAGGTAACAATGAA
 AACAGCCAAAGTAGCCGCCTCGGACTGGACATTCTTGCACTGCTTGCCCAGGAA
 ACCCGAAGAAGTAGACGACGAAGTATTCTACTCGCCAGGTCGTTGGTATTCCCC
 GAAGCCGAAAACAGGAAATGGACAATCATGGCCGTAATGGTATCGTTGTTGACA
 GACTACTCGCCCCAGTTGCAGAAACCCAAATTCTGAATAGTGAA

>pARM1926 (SEQ ID NO: 223)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGGGGCAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1927 (SEQ ID NO: 224)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGGGGCCGGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA

AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1928 (SEQ ID NO: 225)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGGGCCGGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCCGGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM1929 (SEQ ID NO: 226)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCCTGCAGGGCAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM2016 (SEQ ID NO: 227)

ATGCTGTTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCCACA
 ACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACAAGGTGCAGCT
 GAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCAAGTA
 CATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCGAGTA
 CCTGCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCAGCAC
 CCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACCCCTG
 CTTCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCGACAC
 CGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACAAGCA
 GAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACGGCCT
 GAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCAGGA
 GCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAACAA
 CATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCAGGC
 CGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGAGCA
 GTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGAGGC
 CGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCAGGA
 GGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCATGA
 AGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCCGCA
 AGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGTTCC
 CCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGCTGA
 CCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM2260 (SEQ ID NO: 228)

ATGCTGGTATTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCC
 ACAACTTCATGGTGCGCAACTTCCGCTGCGGCCAGCCCCTGCAGAACCAGGGTGC
 AGCTGAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCC
 GGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCG
 AGTACCTGCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCA
 GCACCCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACC
 CCTGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCTGGA
 GGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GCAAGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTG
 TCCCCGAGGCCGAGAACCGCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGC
 TGACCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

>pARM2262 (SEQ ID NO: 229)

ATGCTGGTATTCAACCTGCGCATCCTGCTGAACAACGCCGCCTTCCGCAACGGCC
 ACAACTTCATGGTGCCTCACTTCCGCTGCGGCCAGCCCCTGCAGAACCAGGGTGC
 AGCTGAAGGGCCGCGACCTGCTGACCCTGAAGAACTTCACCGGCGAGGAGATCC
 GGTACATGCTGTGGCTGAGCGCCGACCTGAAGTTCCGCATCAAGCAGAAGGGCG
 AGTACCTGCCCCCTGCTGCAGGGCAAGAGCCTGGGCATGATCTTCGAGAAGCGCA

GCACCCGCACCCGCCTGAGCACCGAGACAGGCTTCGCCCTGCTGGGCGGCCACC
 CCTGCTTCCTGACCACCCAGGACATCCACCTGGGCGTGAACGAGAGCCTGACCG
 ACACCGCCCGCGTGCTGAGCAGCATGGCCGACGCCGTGCTGGCCCGCGTGTACA
 AGCAGAGCGACCTGGACACCCTGGCCAAGGAGGCCAGCATCCCCATCATCAACG
 GCCTGAGCGACCTGTACCACCCCATCCAGATCCTGGCCGACTACCTGACCCTGCA
 GGAGCACTACAGCAGCCTGAAGGGCCTGACCCTGAGCTGGATCGGCGACGGCAA
 CAACATCCTGCACAGCATCATGATGAGCGCCGCCAAGTTCGGCATGCACCTGCA
 GGCCGCCACCCCCAAGGGCTACGAGCCCGACGCCAGCGTGACCAAGCTGGCCGA
 GCAGTACGCCAAGGAGAACGGCACCAAGCTGCTGCTGACCAACGACCCCCCTGGA
 GGCCGCCACGGCGGCAACGTGCTGATCACCGACACCTGGATCAGCATGGGCCA
 GGAGGAGGAGAAGAAGAAGCGCCTGCAGGCCTTCCAGGGCTACCAGGTGACCA
 TGAAGACCGCCAAGGTGGCCGCCAGCGACTGGACCTTCCTGCACTGCCTGCCCC
 GCAAGCCCGAGGAGGTGGACGACGAGGTGTTCTACAGCCCCCGCAGCCTGGTGT
 TCCCCGAGGCCGAGAACCGCCAAGTGGACCATCATGGCCGTGATGGTGAGCCTGC
 TGACCGACTACAGCCCCCAGCTGCAGAAGCCCAAGTTCTGA

AT5G61250	AACCAAUCGAAAGAAACCAAA	SEQ ID NO: 230
AT5G46430	CUCUAAUCACCAGGAGUAAAA	SEQ ID NO: 231
AT5G47110	GAGAGAGAUUCUUAACAAAAAA	SEQ ID NO: 232
AT1G03110	UGUGUAACAACAACAACA	SEQ ID NO: 233
AT3G12380	CCGCAGUAGGAAGAGAAAGCC	SEQ ID NO: 234
AT5G45910	AAAAAAAAAAGAAUCAUAAA	SEQ ID NO: 235
AT1G07260	GAGAGAAAGAAAGAAGACG	SEQ ID NO: 236
AT3G55500	CAAUUAAAAAUACUUACCAAA	SEQ ID NO: 237
AT3G46230	GCAAACAGAGUAAGCGAAACG	SEQ ID NO: 238
AT2G36170	GCGAAGAAGACGAACGCAAAG	SEQ ID NO: 239
AT1G10660	UUAGGACUGUAUUGACUGGCC	SEQ ID NO: 240
AT4G14340	AUCAUCGGAAUUCGGA AAAAG	SEQ ID NO: 241
AT1G49310	AAAACAAAAGUUAAGCAGAC	SEQ ID NO: 242
AT4G14360	UUUAUCUCAAAUAAGAAGGCA	SEQ ID NO: 243
AT1G28520	GGUGGGGAGGUGAGAUUUCUU	SEQ ID NO: 244
AT1G20160	UGAUUAGGAAACUACAAAGCC	SEQ ID NO: 245
AT5G37370	CAUUUUUCAAUUUCAUAAAAC	SEQ ID NO: 246
AT4G11320	UUACUUUUUAAGCCCAACAAA	SEQ ID NO: 247
AT5G40850	GGCGUGUGUGUGUGUUGUUGA	SEQ ID NO: 248
AT1G06150	GUGGUGAAGGGGAAGGUUAG	SEQ ID NO: 249
AT2G26080	UUGUUUUUUUUUGGUUUGGUU	SEQ ID NO: 250

mARM2260 (SEQ ID NO: 251)

AGGAUUAUUACAUCAAAACAAAAAGCCGCCACCAUGCUGGUUAUUCAACCUGCG
CAUCCUGCUGAACAACGCCGCCUUCCGCAACGGCCACAACUUCAUGGUGCGCA
ACUUCCGCUGCGGCCAGCCCCUGCAGAACCGGGUGCAGCUGAAGGGGCCGCGAC
CUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCCGGUACAUGCUGUGGCU
GAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGC
UGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGC
CUGAGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCCUGCUUCCUGAC
CACCCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCCGCG
UGCUGAGCAGCAUGGCCGACGCCGUGCUGGCCCGCGUGUACAAGCAGAGCGAC

CUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGA
CCUGUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACU
ACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUC
CUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGC
CACCCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGGCCGAGCAGU
ACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCC
GCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGA
GGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGA
AGACCGCCAAGGUGGCCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGC
AAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUU
CCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGC
UGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAA
UGAGCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGC
CCUCCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUG
GGCAGCAUCUAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAA

mARM2262 (SEQ ID NO: 252)

AGGAUUAUUAUCAAAAACAAAAAGCCGCCACCAUGCUGGUUAUUAACCUGCG
CAUCCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGCGCA
ACUUCGCGUGCGGCCAGCCCCUGCAGAACCAGGUGCAGCUGAAGGGCCGCGAC
CUGCUGACCCUGAAGAACUUCACCGGCGAGGAGAUCCGGUACAUGCUGUGGCU
GAGCGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCUUGC
UGCAGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGC
CUGAGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCUGCUUCCUGAC
CACCCAGGACAUCCACCUUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCG
UGCUGAGCAGCAUGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGAGCGAC
CUGGACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGA
CCUGUACCACCCCAUCCAGAUCCUGGCCGACUACCUGACCCUGCAGGAGCACU
ACAGCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUC
CUGCACAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGC
CACCCCCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGU
ACGCCAAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCC
GCCACGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGCCAGGA
GGAGGAGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGA
AGACCGCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGC
AAGCCCGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUU
CCCCGAGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGC
UGACCGACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAA
UGAGCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGCUGGGCCUCCCAACGGGC
CCUCCUCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUG
GGCAGCAUCUAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAGCU

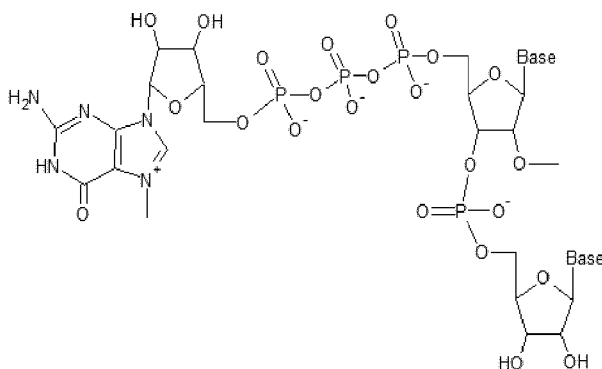
mARM2016 (SEQ ID NO: 253)

AGGAUUAUUAUCAAAAACAAAAAGCCGCCACCAUGCUGUUAACCUGCGCAU
CCUGCUGAACAACGCCGCCUUCGCAACGGCCACAACUUCAUGGUGCGCAACU
UCCGCUGCGGCCAGCCCCUGCAGAACAAGGUGCAGCUGAAGGGCCGCGACCUG

CUGACCCUGAAGAACUUCACCGGCGAGGAGAUCAAGUACAUGCUGUGGCUGAG
 CGCCGACCUGAAGUUCGCAUCAAGCAGAAGGGCGAGUACCUGCCCCUGCUGC
 AGGGCAAGAGCCUGGGCAUGAUCUUCGAGAAGCGCAGCACCCGCACCCGCCUG
 AGCACCGAGACAGGCUUCGCCCUGCUGGGCGGCCACCCUGCUUCCUGACCAC
 CCAGGACAUCCACCUGGGCGUGAACGAGAGCCUGACCGACACCGCCCGCGUGC
 UGAGCAGCAUGGCCGACGCCGUGCUGGGCCCGCGUGUACAAGCAGAGCGACCUG
 GACACCCUGGCCAAGGAGGCCAGCAUCCCCAUCAUCAACGGCCUGAGCGACCU
 GUACCACCCCAUCCAGAUCCUGGGCCGACUACCUGACCCUGCAGGAGCACUACA
 GCAGCCUGAAGGGCCUGACCCUGAGCUGGAUCGGCGACGGCAACAACAUCCUG
 CA
 CAGCAUCAUGAUGAGCGCCGCCAAGUUCGGCAUGCACCUGCAGGCCGCCACCC
 CCAAGGGCUACGAGCCCGACGCCAGCGUGACCAAGCUGGCCGAGCAGUACGCC
 AAGGAGAACGGCACCAAGCUGCUGCUGACCAACGACCCCCUGGAGGCCGCCCA
 CGGCGGCAACGUGCUGAUCACCGACACCUGGAUCAGCAUGGGGCCAGGAGGAGG
 AGAAGAAGAAGCGCCUGCAGGCCUUCAGGGCUACCAGGUGACCAUGAAGACC
 GCCAAGGUGGCCGCCAGCGACUGGACCUUCCUGCACUGCCUGCCCCGCAAGCC
 CGAGGAGGUGGACGACGAGGUGUUCUACAGCCCCCGCAGCCUGGUGUUCCTCG
 AGGCCGAGAACCGCAAGUGGACCAUCAUGGCCGUGAUGGUGAGCCUGCUGACC
 GACUACAGCCCCCAGCUGCAGAAGCCCAAGUUCUGAGGUCUCUAGUAAUGAGC
 UGGAGCCUCGGUAGCCGUUCCUCCUGCCCCGCUGGGCCUCCCAACGGGCCCUCC
 UCCCCUCCUUGCACCGGCCCUUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAG
 CAUCUAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
 AAA

REIVINDICACIONES

1. Un polinucleótido que codifica una proteína ornitina transcarbamilasa (OTC) que comprende una secuencia de aminoácidos de SEQ ID NO: 4 y que tiene actividad enzimática OTC, en donde el polinucleótido es un ARNm que comprende una secuencia de SEQ ID NO: 221, en donde T está sustituido por U.
2. El polinucleótido de la reivindicación 1, que comprende la UTR en 5' derivada de un gen expresado por *Arabidopsis thaliana*.
3. El polinucleótido de la reivindicación 1 o 2, que comprende una UTR en 5' de SEQ ID NO: 6.
4. El polinucleótido de las reivindicaciones 1-3, que comprende:
 - (a) una UTR en 3'; o
 - (b) a secuencia de Kozak; o
 - (c) una caperuza en 5'; o
 - (d) una caperuza en 5' de m7GpppGm que tiene la siguiente estructura:



5. El polinucleótido de las reivindicaciones 1-3, que comprende una UTR en 3' de SEQ ID NO: 21.
6. El polinucleótido de las reivindicaciones 1-3, que comprende una secuencia de Kozak de SEQ ID NO: 23.
7. El polinucleótido de las reivindicaciones 1-3, que comprende:
 - (i) una cola de poli A en 3'; o
 - (ii) una cola de poli A en 3' que comprende de aproximadamente 60 nucleótidos de adenina secuenciales a aproximadamente 125 nucleótidos de adenina secuenciales; o
 - (iii) una cola de poli A en 3' que comprende de aproximadamente 100 nucleótidos de adenina secuenciales a aproximadamente 125 nucleótidos de adenina secuenciales.
8. El polinucleótido de la reivindicación 1 o 2, que comprende además una secuencia de SEQ ID NO: 6 y una secuencia de SEQ ID NO: 21.
9. El polinucleótido de la reivindicación 1 o 2, que comprende una secuencia de SEQ ID NO: 251 o SEQ ID NO: 252.
10. El polinucleótido de la reivindicación 1 o 2, en donde:
 - (a) el porcentaje de nucleobases de uracilo en la región codificante del polinucleótido se reduce con respecto al porcentaje de nucleobases de uracilo en una secuencia de ácido nucleico de OTC de tipo silvestre; o
 - (b) el 1-100 % de las uridinas son análogos de uridina modificados; o
 - (c) el 100 % de los nucleótidos de uridina son 5-metoxiuridina; o
 - (d) el 100 % de los nucleótidos de uridina son N¹-metilpseudouridina.
11. El polinucleótido de la reivindicación 8 o 9, en donde:
 - (a) el 1-100 % de las uridinas son análogos de uridina modificados; o
 - (b) el 100 % de los nucleótidos de uridina son 5-metoxiuridina; o
 - (c) el 100 % de los nucleótidos de uridina son N¹-metilpseudouridina.
12. El polinucleótido de la reivindicación 1 o 2, que comprende:

- (a) una UTR en 5' que comprende una secuencia de SEQ ID NO: 6; o
- (b) una UTR en 3' que comprende una secuencia de SEQ ID NO: 21, en donde el polinucleótido comprende análogos de uridina modificados.

- 5
13. Una composición farmacéutica que comprende el polinucleótido de las reivindicaciones 8-12.
14. La composición farmacéutica de la reivindicación 13, que comprende además un transportador farmacéuticamente aceptable en donde el transportador farmacéuticamente aceptable es una formulación lipídica.
- 10
15. El polinucleótido de una cualquiera de las reivindicaciones 1 a 12 o la composición farmacéutica de la reivindicación 13 o 14, para su uso en un método para tratar la deficiencia de OTC en un paciente identificado como que padece deficiencia de OTC.

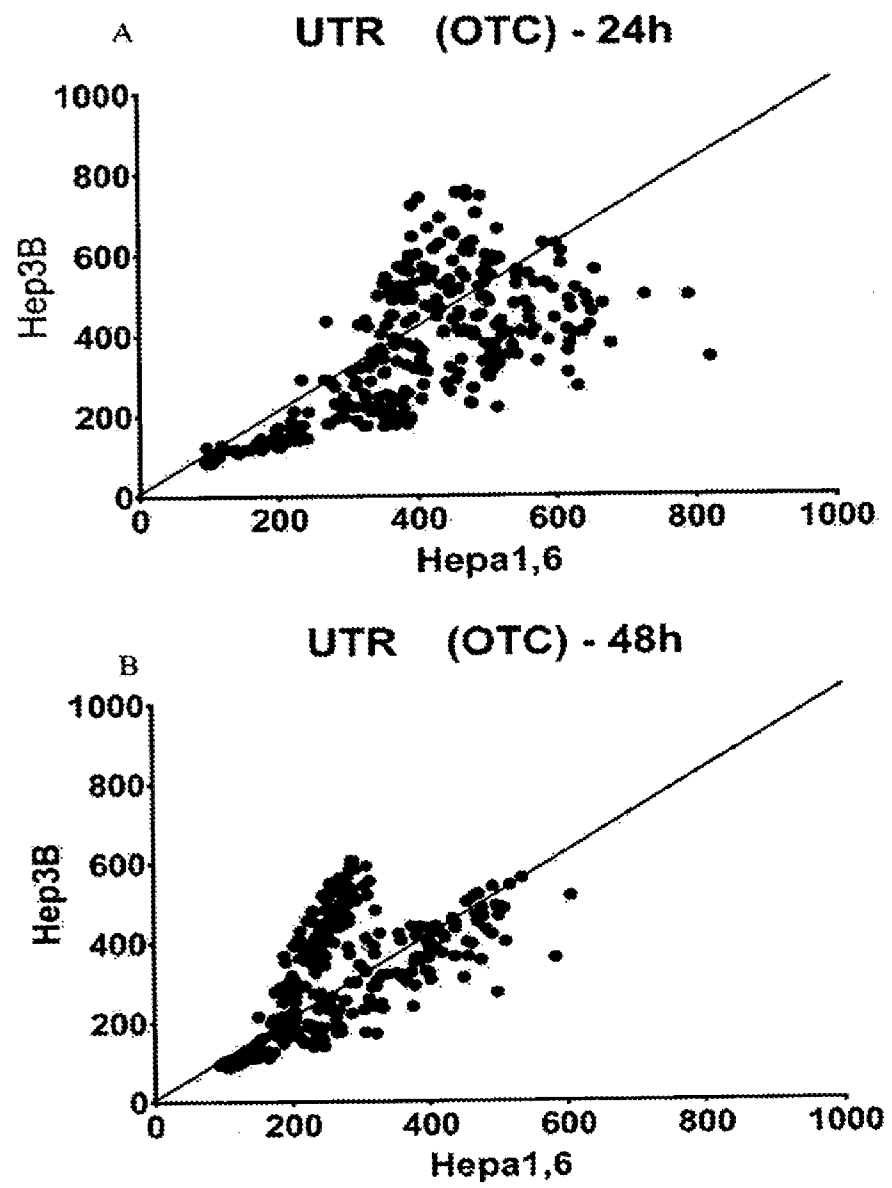


FIG. 1

FIG. 2

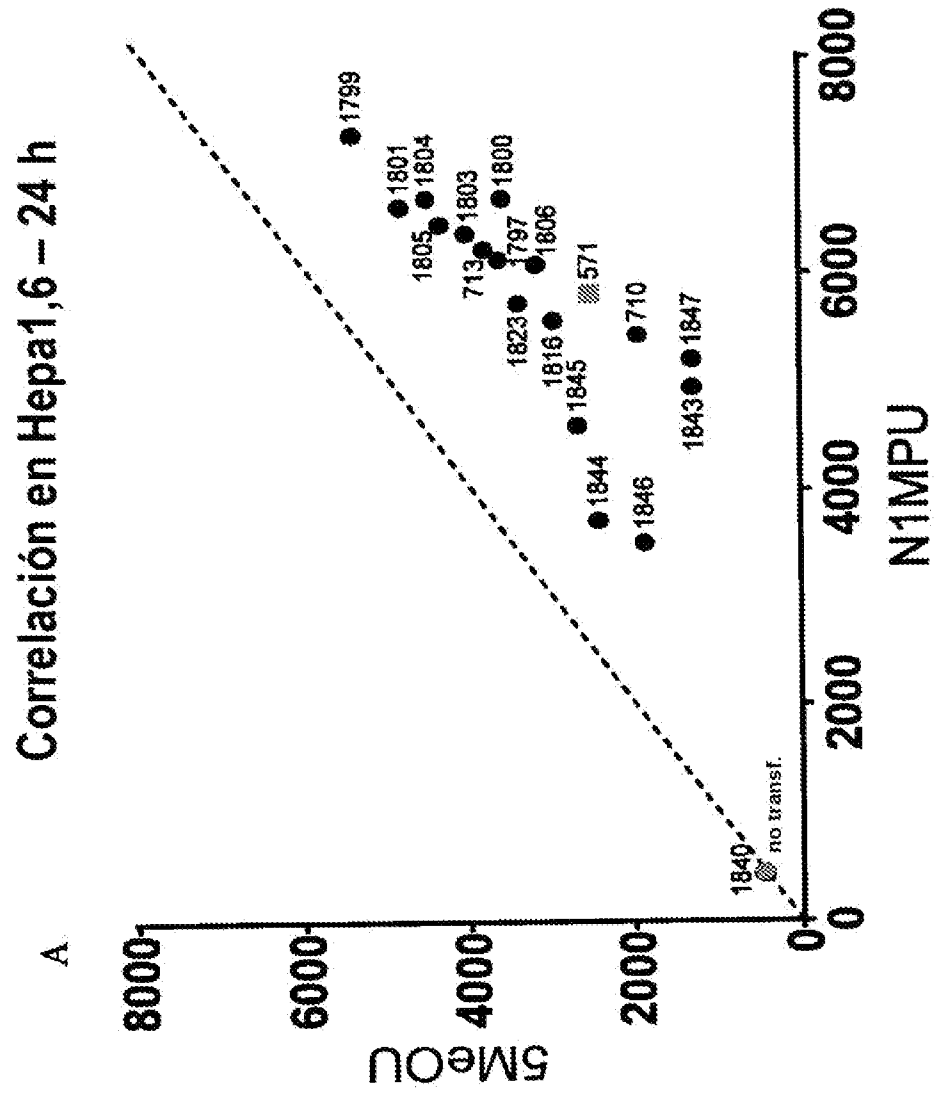


FIG. 2 Continuación

Correlación en Hep3B – 24 h

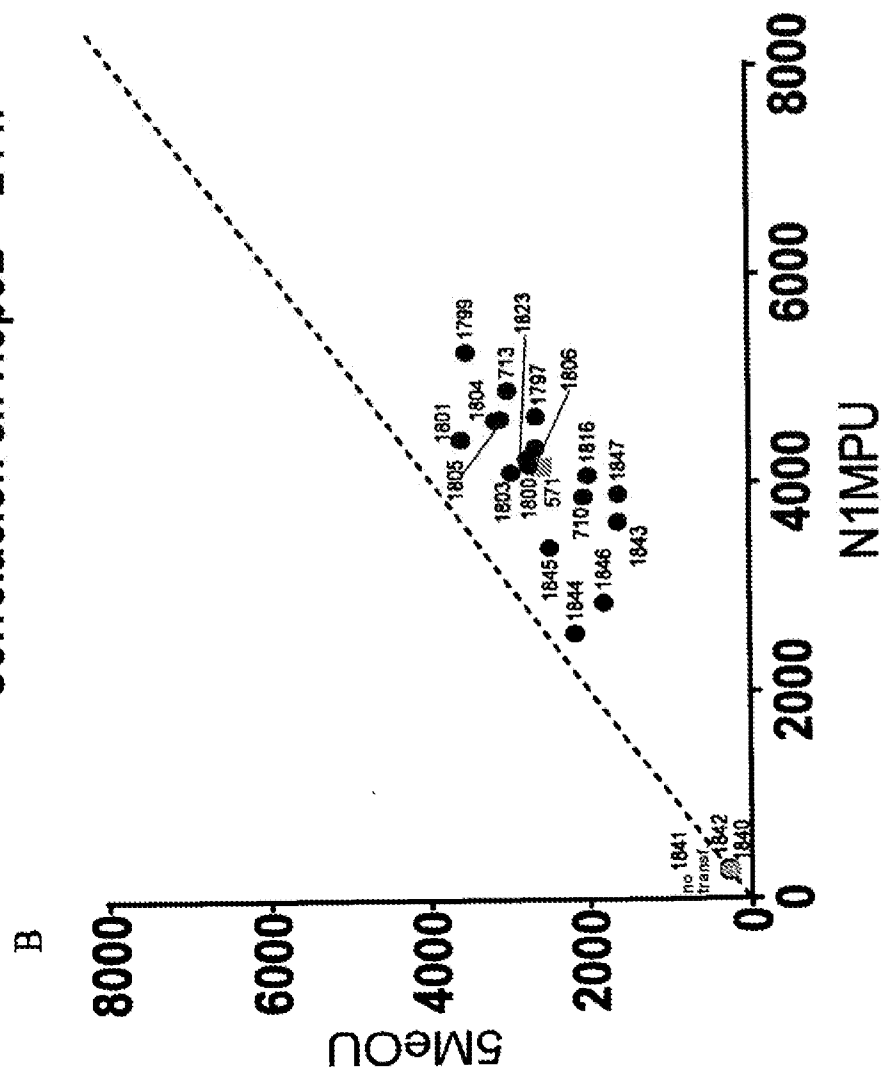


FIG. 3

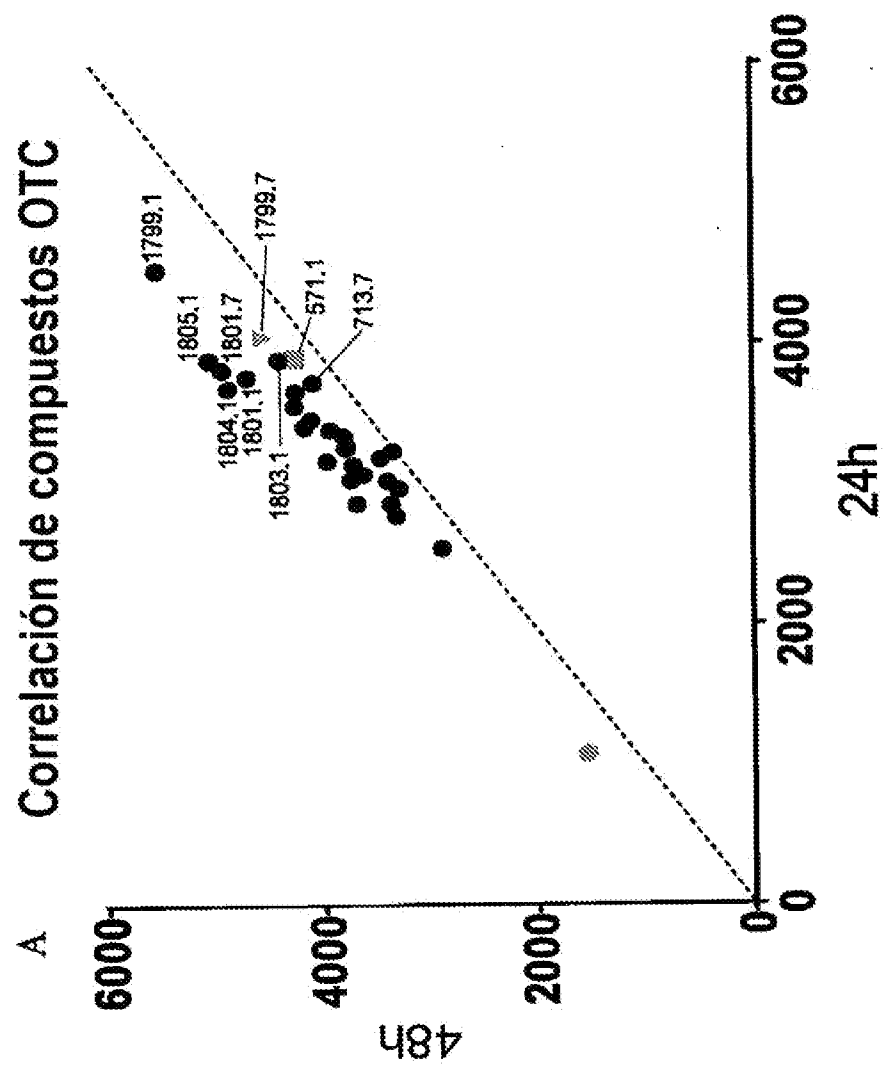


FIG. 3 Continuación

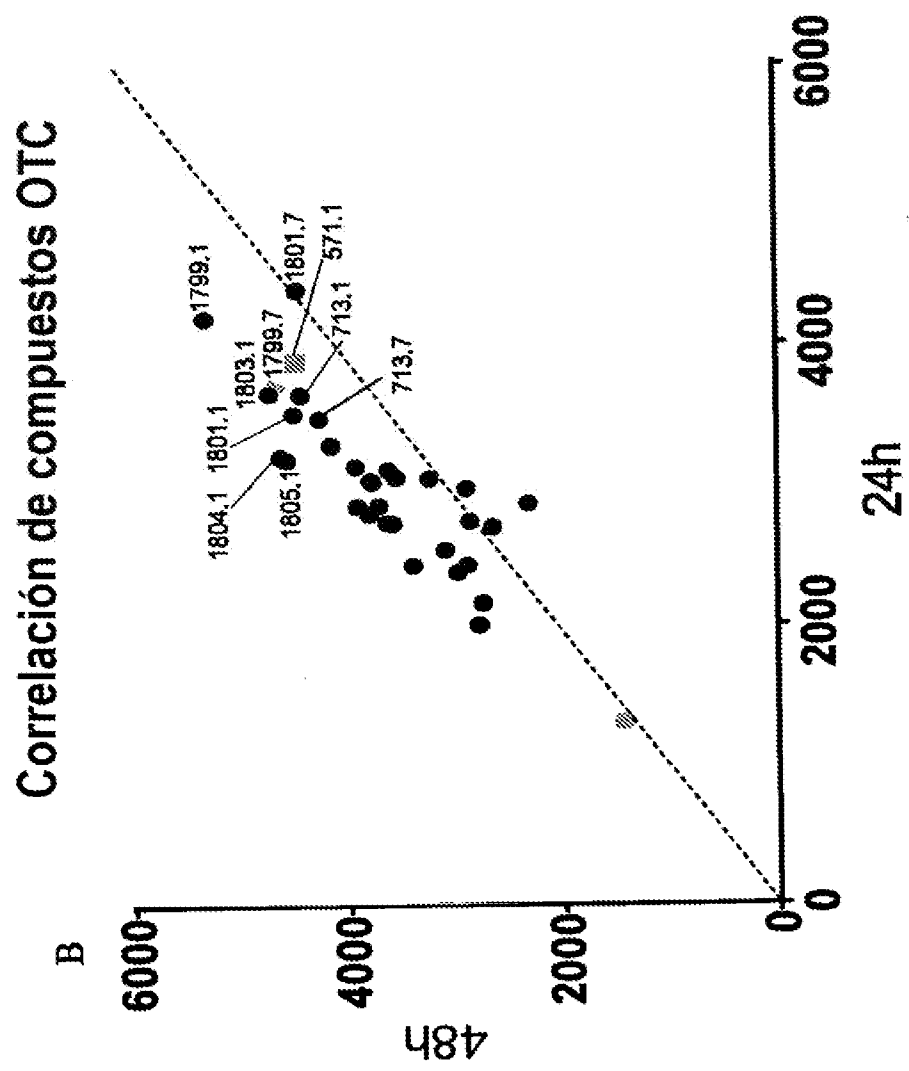


FIG. 4

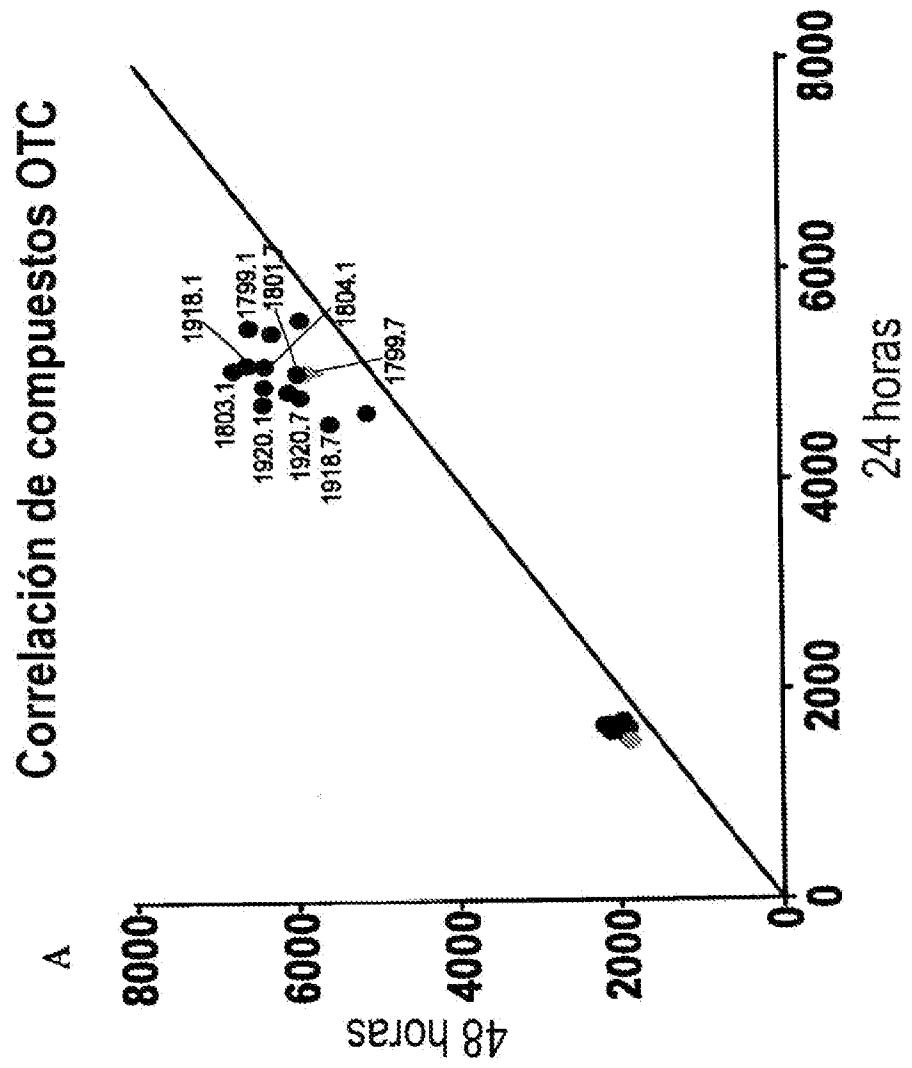
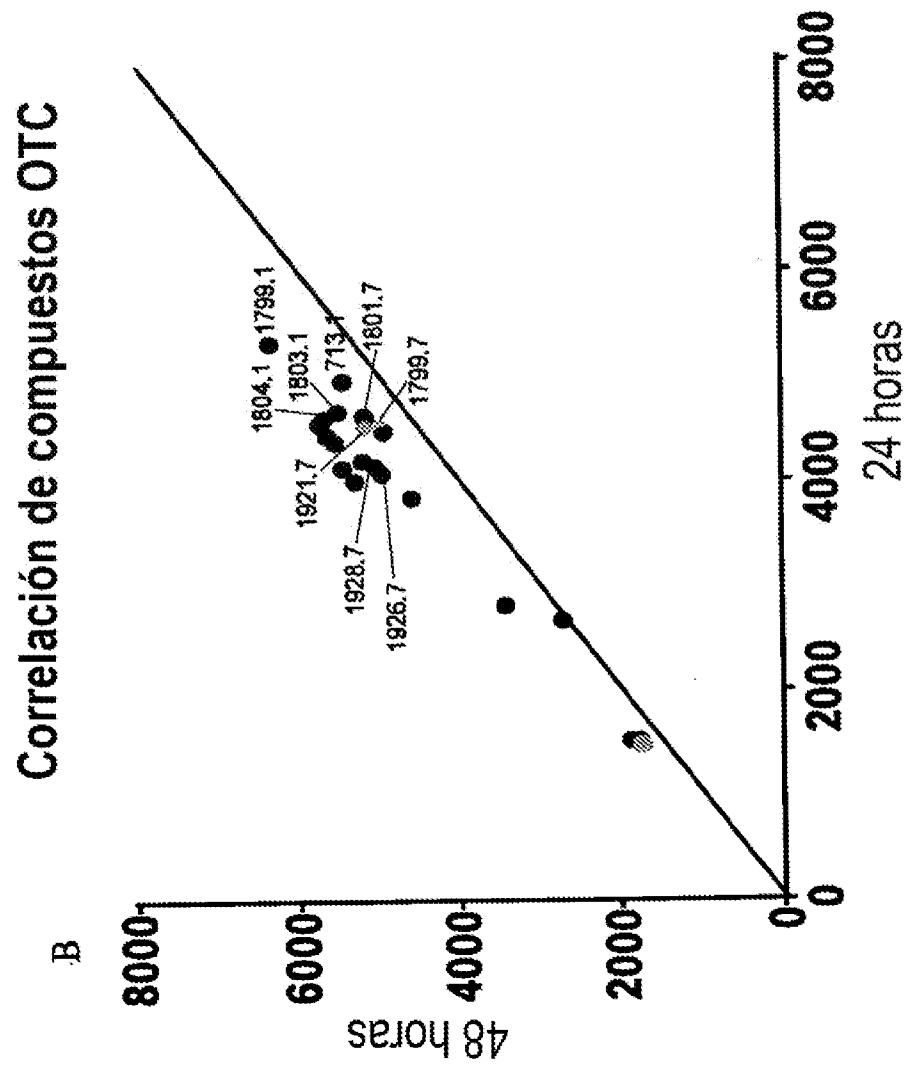


FIG. 4 Continuación



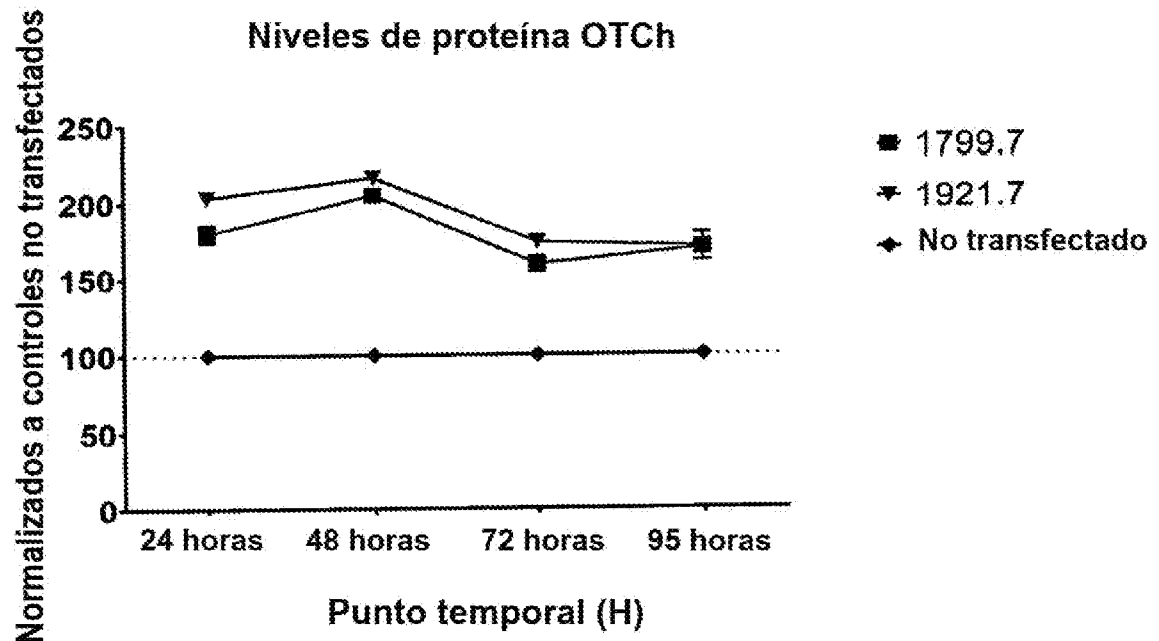


FIG. 5

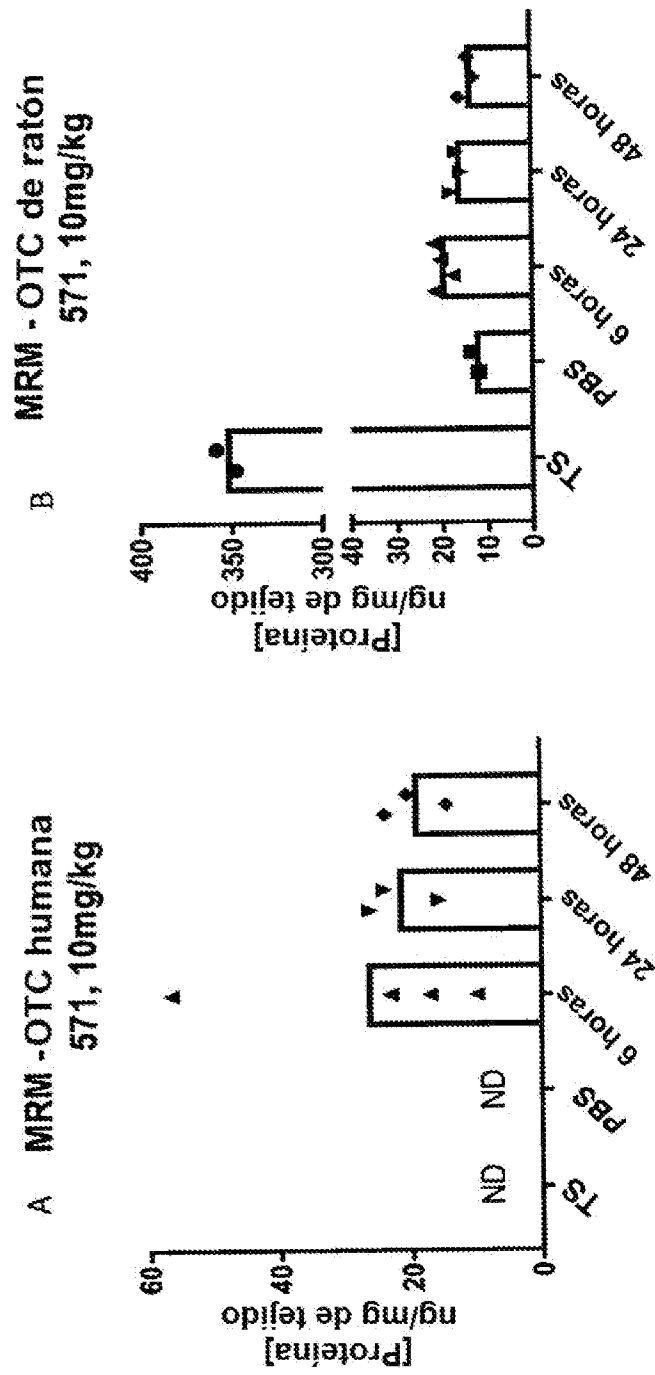


FIG. 6

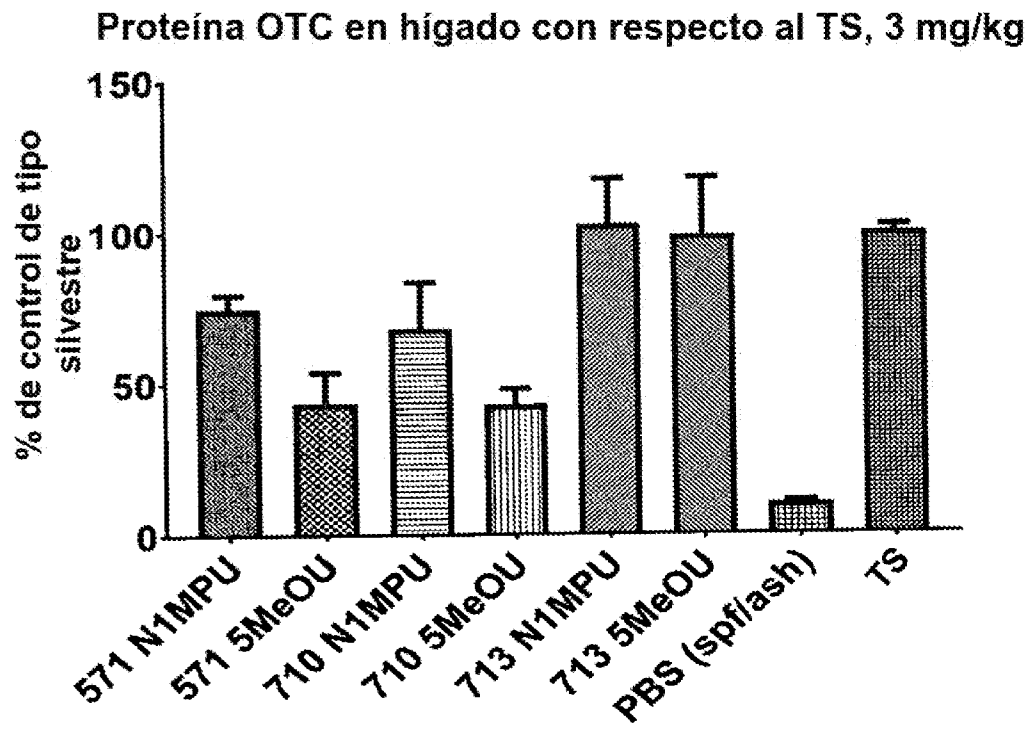


FIG. 7

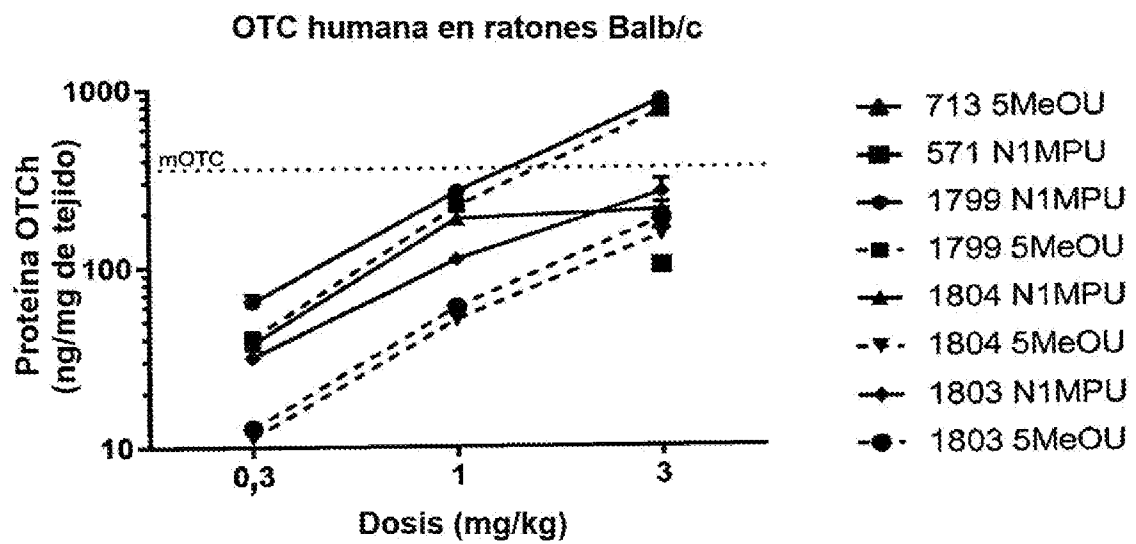


FIG. 8

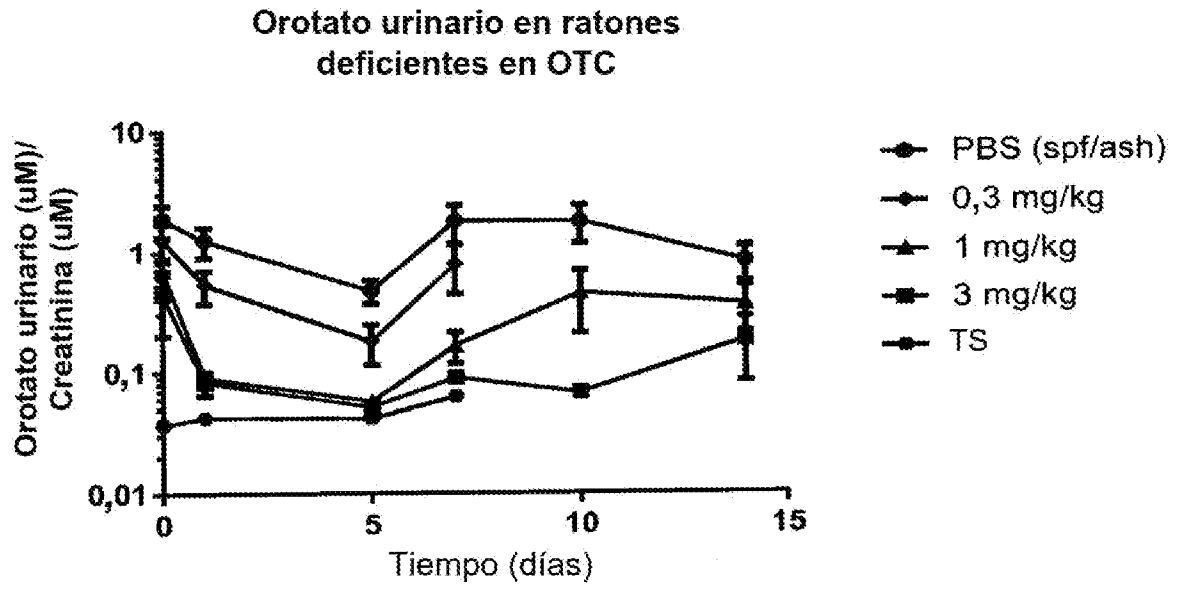


FIG. 9

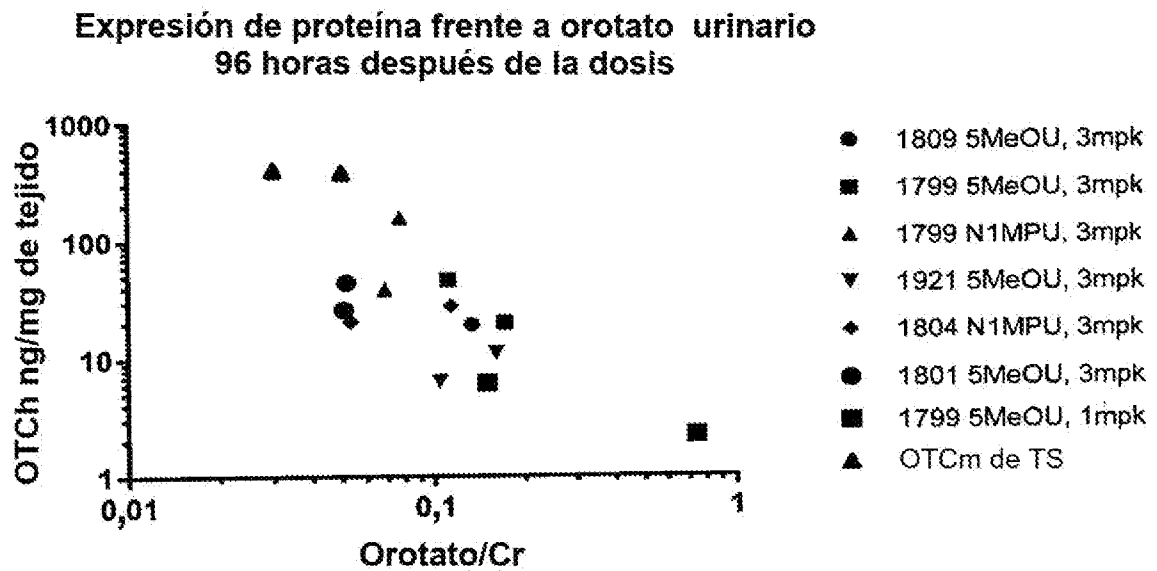


FIG. 10

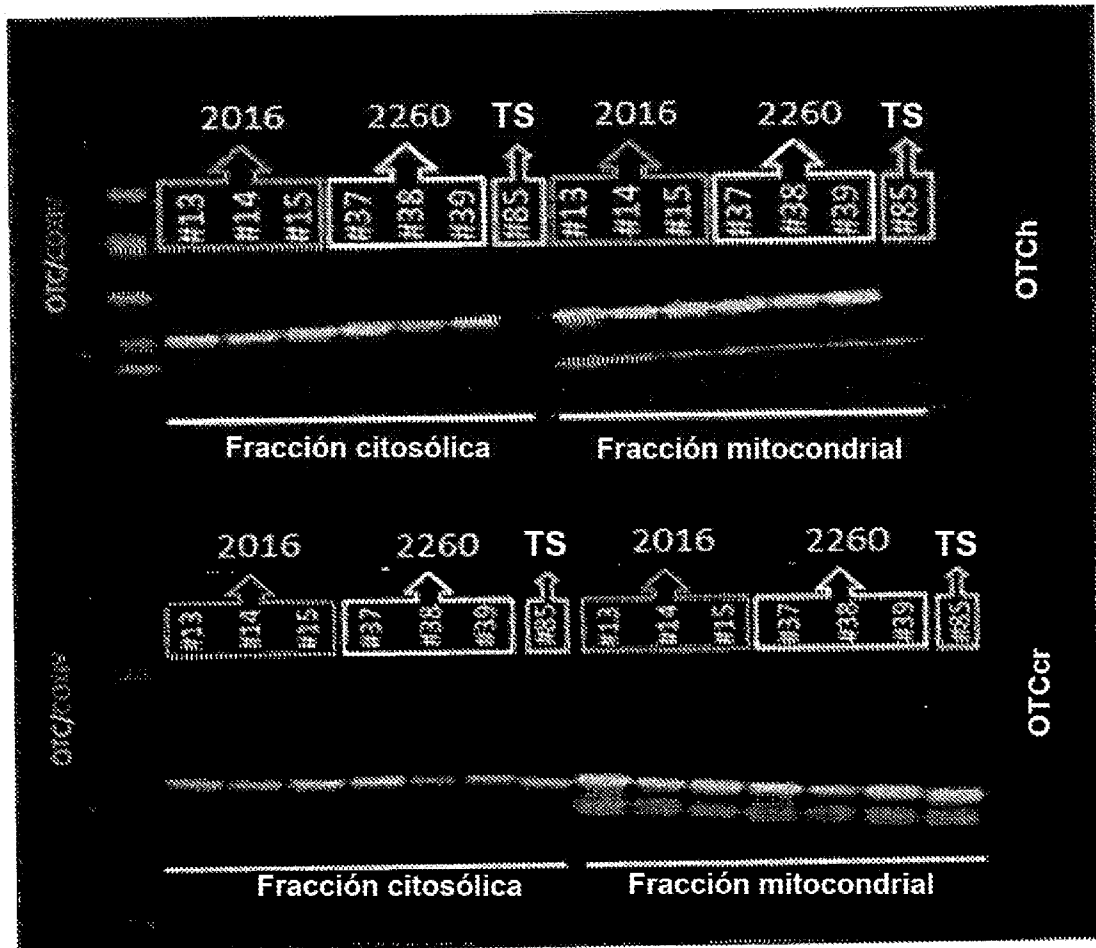


FIG. 11

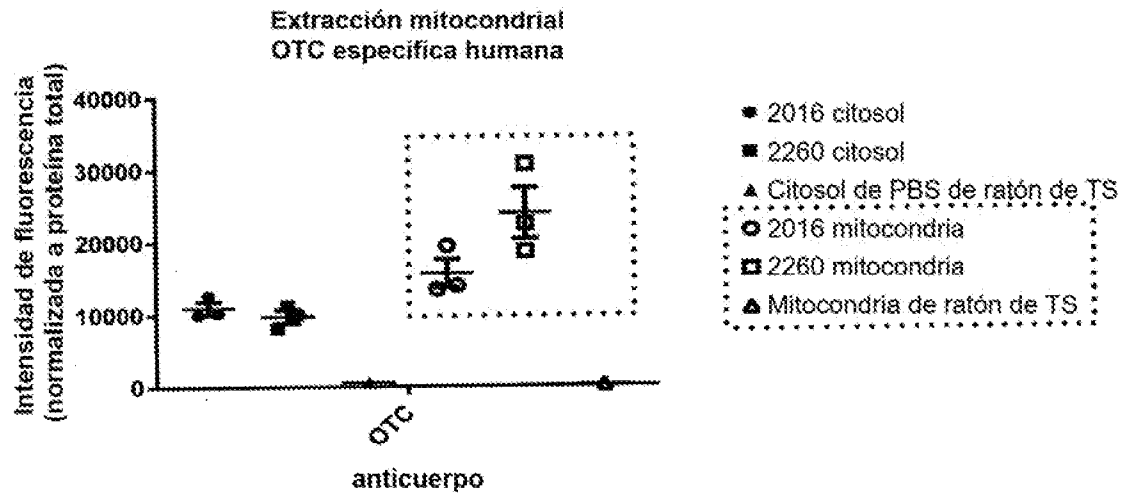


FIG. 12

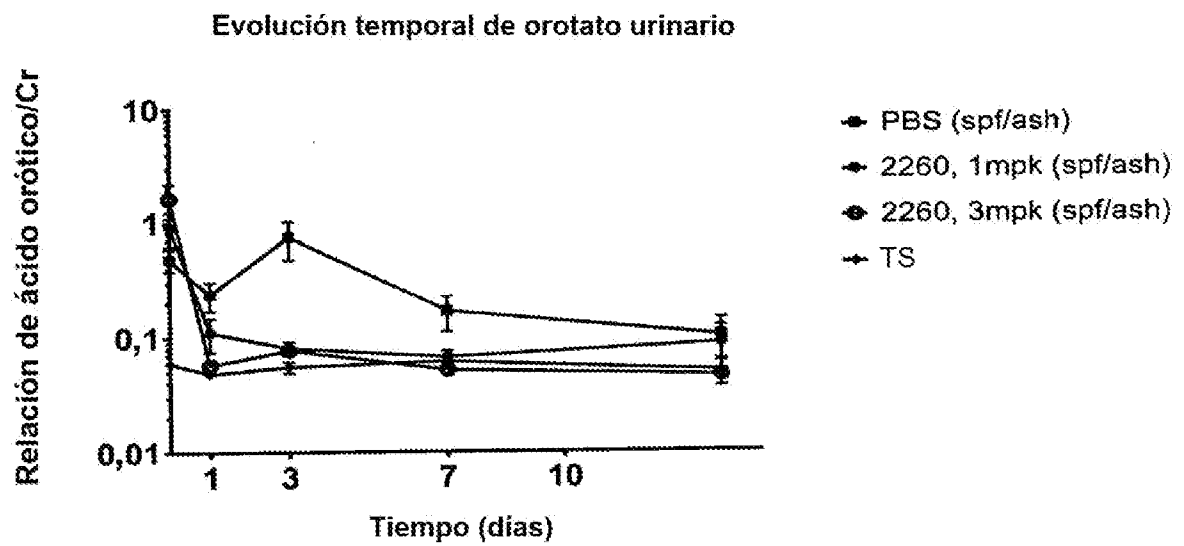


FIG. 13

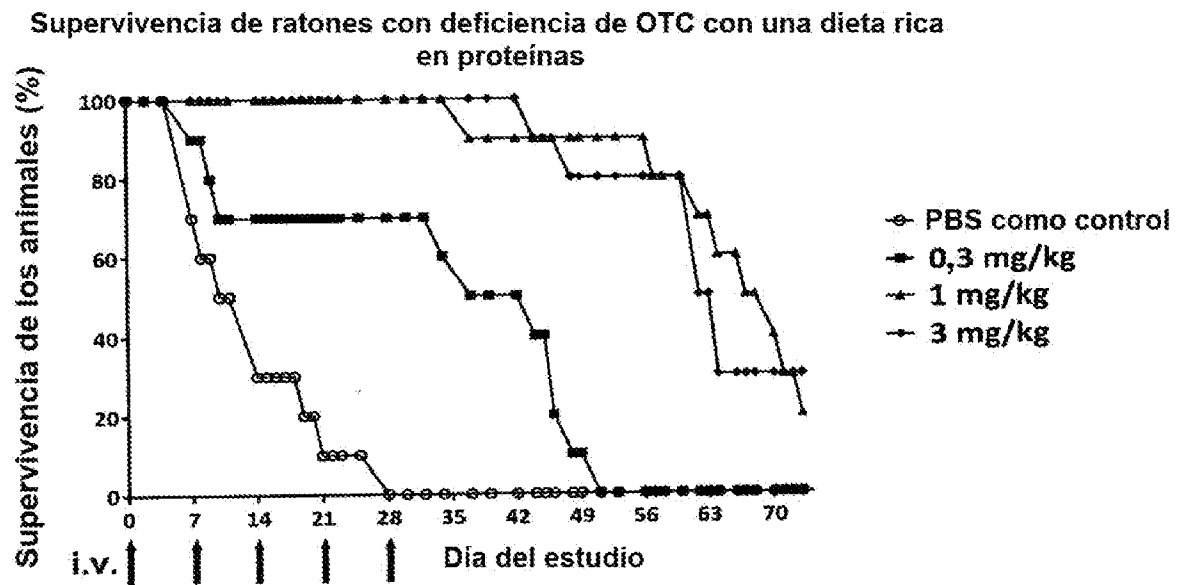


FIG. 14

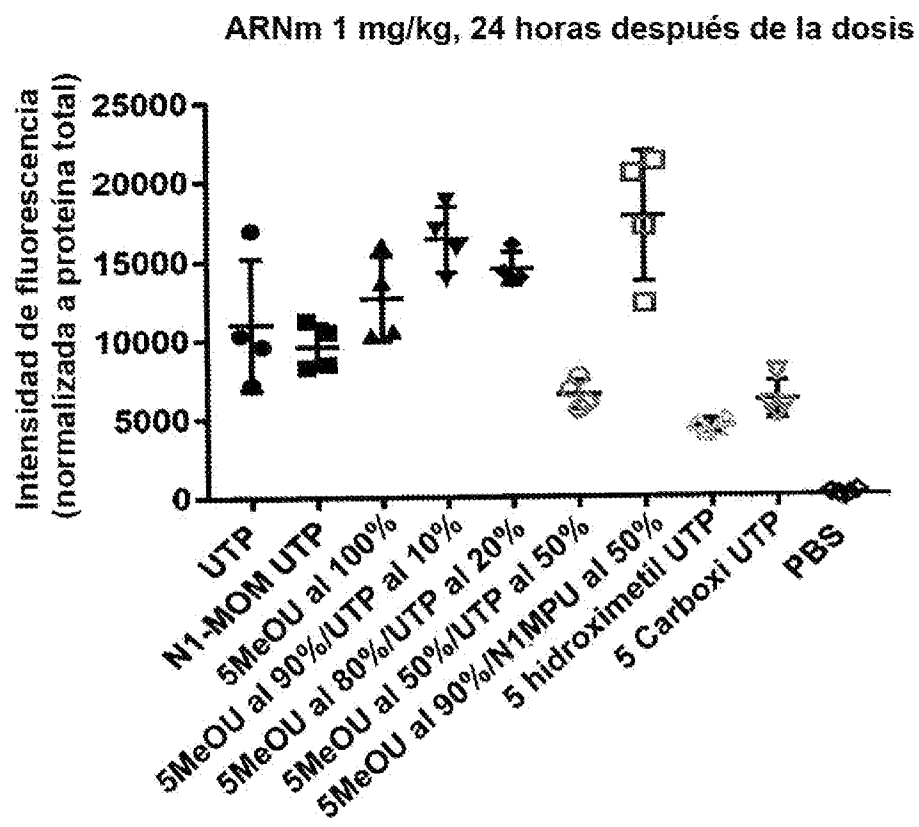


FIG. 15