

JUSTIA

Composizioni e metodi del vaccino contro il coronavirus

30 giugno 2023 - Arcturus Therapeutics, Inc.

Nel presente documento sono fornite molecole di acido nucleico che codificano proteine di replicazione virale e proteine antigeniche del coronavirus o frammenti delle stesse. Nel presente documento sono inoltre fornite composizioni che includono molecole di acido nucleico che codificano proteine di replicazione virale e antigeniche, e lipidi. Le molecole di acido nucleico fornite nel presente documento sono utili per indurre risposte immunitarie.

Ultimi brevetti di Arcturus Therapeutics, Inc.:

- Metodo per realizzare nanoparticelle di RNA incapsulate nei lipidi
- Metodo di liofilizzazione delle nanoparticelle lipidiche
- Particelle lipidiche per la somministrazione di acidi nucleici

- RNA di trasferimento sintetico con loop anticodone esteso
- Acidi nucleici e metodi di trattamento della fibrosi cistica

Vai a: [Descrizione](#) · [Richieste](#) · [Riferimenti citati](#) · [Cronologia brevetti](#) · [Cronologia brevetti](#)

Descrizione

RIFERIMENTI INCROCIATI A DOMANDE CORRELATE

La presente domanda rivendica il beneficio della domanda provvisoria statunitense n. 62/987.191, depositata il 9 marzo 2020, e della domanda provvisoria statunitense n. 63/073.900, depositata il 2 settembre 2020.

RIFERIMENTO A UN ELENCO DI SEQUENZE

La domanda immediata contiene un Sequence Listing che è stato inviato elettronicamente in formato ASCII ed è qui incorporato per riferimento nella sua interezza. Tale copia ASCII, creata l'8 marzo 2021, è denominata 049386-538001US_SequenceListing_ST25.txt e ha una dimensione di 481.150 byte.

CAMPO TECNICO

La presente invenzione riguarda in generale l'induzione di risposte immunitarie contro agenti infettivi e antigeni tumorali e più specificatamente l'autotrascrizione e la replicazione dell'RNA per l'espressione dell'antigene.

SFONDO

Le malattie infettive e il cancro rappresentano un peso significativo per la salute in tutto il mondo. Secondo l'Organizzazione mondiale della sanità (OMS), l'infezione delle basse vie respiratorie è stata la malattia infettiva più mortale al mondo nel 2016, causando circa 3 milioni di decessi. L'impatto delle malattie infettive è illustrato dalla pandemia di coronavirus disease 2019 (COVID-19) causata dalla sindrome respiratoria acuta grave-coronavirus-2 (SARS-CoV-2). SARS-CoV-2 è un nuovo coronavirus identificato per la prima volta a dicembre 2019 a Wuhan, in Cina, che ha causato oltre 20 milioni di infezioni confermate con oltre 700.000 decessi in tutto il mondo ad agosto 2020. Le attuali misure di controllo per frenare la rapida diffusione mondiale di SARS-CoV-2, come i lockdown nazionali, la chiusura di luoghi di lavoro e scuole e la riduzione dei viaggi internazionali, minacciano di provocare una recessione economica globale in una misura mai vista dalla Grande Depressione.

Il cancro è la seconda causa di morte a livello globale, con circa 9,6 milioni di decessi nel mondo nel 2018. Il cancro è un vasto gruppo di malattie che possono colpire quasi tutti gli organi o tessuti del corpo. Il peso del cancro continua a crescere a livello globale, esercitando pressioni fisiche, emotive e finanziarie su pazienti e operatori sanitari.

Gli acidi ribonucleici autoreplicanti (RNA), ad esempio derivati da repliconi virali, sono utili per l'espressione di proteine, come le proteine eterologhe, per una varietà di scopi, come l'espressione di proteine terapeutiche e l'espressione di antigeni per vaccini. Una proprietà desiderabile di tali repliconi è la capacità di espressione sostenuta della proteina.

Sono disponibili pochi trattamenti per le infezioni causate da virus e organismi eucariotici e la resistenza agli antibiotici per il trattamento delle infezioni batteriche è in aumento. Inoltre, sono necessarie risposte rapide, tra cui lo sviluppo rapido di vaccini, per controllare efficacemente le malattie infettive emergenti e le pandemie. Inoltre, molti trattamenti contro il cancro includono interventi chirurgici e chemioterapie costosi e dolorosi che spesso non hanno successo o prolungano solo modestamente la vita nonostante gravi effetti collaterali. Pertanto, esiste la necessità di prevenzione e/o trattamento delle malattie infettive e del cancro.

RIEPILOGO

In un aspetto, la presente divulgazione fornisce una molecola di acido nucleico comprendente (i) un primo polinucleotide che codifica una o più proteine di replicazione virale, in cui il primo polinucleotide è ottimizzato per codoni rispetto a un polinucleotide di tipo selvatico che codifica una o più proteine di replicazione virale; e (ii) un secondo polinucleotide comprendente un primo transgene che codifica una prima proteina antigenica o un frammento della stessa, in cui la prima proteina antigenica è una proteina del coronavirus.

In alcune realizzazioni, una o più proteine di replicazione virale sono proteine alfavirus o proteine rubivirus.

In alcune forme di realizzazione, le proteine alfavirus provengono dal virus dell'encefalite equina venezuelana (VEEV), dal virus dell'encefalite equina orientale (EEEV), dal virus delle Everglades (EVEV), dal virus Mucambo (MUCV), dal virus della foresta Semliki (SFV), dal virus Pixuna (PIXV), dal virus Middleburg (MIDV), dal virus Chikungunya (CHIKV), dal virus O'Nyong-Nyong (ONNV), dal virus Ross River (RRV), dal virus della foresta Barmah (BFV), Virus Getah (GETV), Virus Sagiyama (SAGV), Virus Bebaru (BEBV), Virus Mayaro (MAYV), Virus Una (UNAV), Virus Sindbis (SINV), Virus Aura (AURAV), Virus Whataroa (WHAV), Virus Babanki (BABV), Virus Kyzylgach (KYZV), Virus dell'encefalite equina occidentale (WEEV), Virus Highland J (HJV), Fort Morgan Virus (FMV), Ndumu Virus (NDUV), Salmonid Alphavirus (SAV), Buggy Creek Virus (BCRV) o qualsiasi combinazione di essi.

In alcune realizzazioni, il primo polinucleotide codifica una poliproteina comprendente una proteina alfavirus nsP1, una proteina alfavirus nsP2, una proteina alfavirus nsP3, una proteina alfavirus nsP4 o una qualsiasi combinazione di esse.

In alcune realizzazioni, il primo polinucleotide codifica una poliproteina comprendente una proteina alfavirus nsP1, una proteina alfavirus nsP2, una proteina alfavirus nsP3, o una qualsiasi combinazione di esse, e una proteina alfavirus nsP4.

In alcune realizzazioni, la molecola di acido nucleico comprende inoltre una prima regione intergenica tra una sequenza che codifica la poliproteina comprendente una proteina alfavirus

nsP1, una proteina alfavirus nsP2, una proteina alfavirus nsP3 o una qualsiasi combinazione di esse, e una sequenza che codifica una proteina alfavirus nsP4.

In alcune realizzazioni, la prima regione intergenica comprende una sequenza alfavirus.

In alcune realizzazioni, il primo polinucleotide comprende una sequenza avente almeno l'80% di identità con una sequenza di SEQ ID NO:72.

In alcune realizzazioni, la molecola di acido nucleico comprende inoltre una regione 5' non tradotta (UTR).

In alcune realizzazioni, la 5' UTR comprende una 5' UTR virale, una 5' UTR non virale o una combinazione di sequenze 5' UTR virali e non virali.

In alcune realizzazioni, l'UTR 5' comprende un'UTR 5' dell'alfavirus.

In alcune forme di realizzazione, l'alfavirus 5' UTR comprende il virus dell'encefalite equina venezuelana (VEEV), il virus dell'encefalite equina orientale (EEEV), il virus delle Everglades (EVEV), il virus Mucambo (MUCV), il virus della foresta Semliki (SFV), il virus Pixuna (PIXV), il virus Middleburg (MIDV), il virus Chikungunya (CHIKV), il virus O'Nyong-Nyong (ONNV), il virus Ross River (RRV), Virus Barmah Forest (BFV), Virus Getah (GETV), Virus Sagiyama (SAGV), Virus Bebaru (BEBV), Virus Mayaro (MAYV), Virus Una (UNAV), Virus Sindbis (SINV), Virus Aura (AURAV), Virus Whataroa (WHAV), Virus Babanki (BABV), Virus Kyzylgach (KYZV), Virus dell'encefalite equina occidentale (WEEV), Highland J Sequenza 5' UTR del virus Fort Morgan (FMV), virus Ndumu (NDUV), virus Alphavirus dei salmonidi (SAV) o virus Buggy Creek (BCRV).

In alcune realizzazioni, l'UTR 5' comprende una sequenza SEQ ID NO:73, SEQ ID NO:74 o SEQ ID NO:75.

In alcune realizzazioni, la molecola di acido nucleico comprende inoltre una regione 3' non tradotta (UTR).

In alcune realizzazioni, l'UTR 3' comprende un UTR 3' virale, un UTR 3' non virale o una combinazione di sequenze UTR 3' virali e non virali. In alcune realizzazioni, l'UTR 3' comprende un UTR 3' di alphavirus.

In alcune forme di realizzazione, l'alfavirus 3' UTR comprende il virus dell'encefalite equina venezuelana (VEEV), il virus dell'encefalite equina orientale (EEEV), il virus delle Everglades (EVEV), il virus Mucambo (MUCV), il virus della foresta Semliki (SFV), il virus Pixuna (PIXV), il virus Middleburg (MIDV), il virus Chikungunya (CHIKV), il virus O'Nyong-Nyong (ONNV), il virus Ross River (RRV), Virus Barmah Forest (BFV), Virus Getah (GETV), Virus Sagiyama (SAGV), Virus Bebaru (BEBV), Virus Mayaro (MAYV), Virus Una (UNAV), Virus Sindbis (SINV), Virus Aura (AURAV), Virus Whataroa (WHAV), Virus Babanki (BABV), Virus Kyzylgach (KYZV), Virus dell'encefalite equina occidentale (WEEV), Highland J Sequenza 3' UTR del virus Fort Morgan (FMV), virus Ndumu (NDUV), virus Alphavirus dei salmonidi (SAV) o virus Buggy Creek (BCRV).

In alcune realizzazioni, la 3' UTR comprende una sequenza poli-A.

In alcune realizzazioni, l'UTR 3' comprende una sequenza SEQ ID NO:76.

In alcune realizzazioni, la proteina antigenica è una proteina SARS-CoV-2.

In alcune realizzazioni, la proteina antigenica è una glicoproteina spike del SARS-CoV-2.

In alcune realizzazioni, la glicoproteina spike del SARS-CoV-2 è una glicoproteina spike del SARS-CoV-2 di tipo selvaggio con una sequenza aminoacidica SEQ ID NO: 123.

In alcune realizzazioni, il secondo polinucleotide comprende una sequenza avente almeno l'85% di identità con una sequenza SEQ ID NO:121 o SEQ ID NO:122.

In alcune realizzazioni, il secondo polinucleotide comprende almeno due transgeni.

In alcune realizzazioni, un secondo transgene codifica una seconda proteina antigenica o un frammento della stessa oppure una proteina immunomodulatrice.

In alcune realizzazioni, il secondo polinucleotide comprende inoltre una sequenza che codifica un peptide 2A, un sito di ingresso ribosomiale interno (IRES) o una combinazione di questi, situato tra i transgeni.

In alcune realizzazioni, la proteina immunomodulatrice è una citochina, una chemiochina o un'interleuchina.

In alcune realizzazioni, il secondo transgene codifica una seconda proteina del coronavirus.

In alcune realizzazioni, il primo polinucleotide si trova 5' del secondo polinucleotide.

In alcune realizzazioni, la molecola di acido nucleico comprende inoltre una seconda regione intergenica situata tra il primo polinucleotide e il secondo polinucleotide.

In alcune realizzazioni, la seconda regione intergenica comprende una sequenza avente almeno l'85% di identità con una sequenza di SEQ ID NO:77.

In alcune realizzazioni, la molecola di acido nucleico è

- (a) una molecola di DNA; o
- (b) una molecola di RNA, in cui T è sostituita con U.

In alcune realizzazioni, la molecola di DNA comprende inoltre un promotore.

In alcune realizzazioni, il promotore si trova al 5' della 5'UTR.

In alcune realizzazioni, il promotore è un promotore T7, un promotore T3 o un promotore SP6.

In alcune realizzazioni, la molecola di RNA è una molecola di RNA autoreplicante.

In alcune realizzazioni, la molecola di RNA comprende inoltre un cappuccio 5'.

In alcune realizzazioni, il cappuccio 5' ha una struttura Cap 1, una struttura Cap 1 (A), una struttura Cap 2, una struttura Cap 0 o una qualsiasi combinazione di queste.

In un altro aspetto, la divulgazione fornisce una molecola di acido nucleico comprendente

- (a) una sequenza di SEQ ID NO:124;
- (b) una sequenza di SEQ ID NO:124, in cui T è sostituito con U;
- (c) una sequenza di SEQ ID NO:125; o
- (d) una sequenza di SEQ ID NO:125, in cui T è sostituito con U.

In alcune realizzazioni, la molecola di acido nucleico è una molecola di RNA.

In alcune realizzazioni, la molecola di acido nucleico comprende inoltre un cappuccio 5' avente una struttura Cap 1.

In un ulteriore aspetto, la divulgazione fornisce una molecola di acido nucleico comprendente:

- (i) un primo polinucleotide comprendente una sequenza avente almeno l'80% di identità con una sequenza di SEQ ID NO:72; e
- (ii) un secondo polinucleotide comprendente un primo transgene che codifica una prima proteina antigenica o un frammento della stessa, in cui la prima proteina antigenica è una proteina del coronavirus.

In alcune realizzazioni, la molecola di acido nucleico comprende inoltre una regione 5' non tradotta (UTR).

In alcune realizzazioni, la 5' UTR comprende una 5' UTR virale, una 5' UTR non virale o una combinazione di sequenze 5' UTR virali e non virali.

In alcune realizzazioni, l'UTR 5' comprende un'UTR 5' dell'alfavirus.

In alcune forme di realizzazione, l'alfavirus 5' UTR comprende il virus dell'encefalite equina venezuelana (VEEV), il virus dell'encefalite equina orientale (EEEV), il virus delle Everglades (EVEV), il virus Mucambo (MUCV), il virus della foresta Semliki (SFV), il virus Pixuna (PIXV), il virus Middleburg (MIDV), il virus Chikungunya (CHIKV), il virus O'Nyong-Nyong (ONNV), il virus Ross River (RRV), Virus Barmah Forest (BFV), Virus Getah (GETV), Virus Sagiyama (SAGV), Virus Bebaru (BEBV), Virus Mayaro (MAYV), Virus Una (UNAV), Virus Sindbis (SINV), Virus Aura (AURAV), Virus Whataroa (WHAV), Virus Babanki (BABV), Virus Kyzylgach (KYZV), Virus dell'encefalite equina occidentale (WEEV), Highland J Sequenza 5' UTR del virus Fort Morgan (FMV), virus Ndumu (NDUV), virus Alphavirus dei salmonidi (SAV) o virus Buggy Creek (BCRV).

In alcune realizzazioni, l'UTR 5' comprende una sequenza SEQ ID NO:73, SEQ ID NO:74 o SEQ ID NO:75.

In alcune realizzazioni, la molecola di acido nucleico comprende inoltre una regione 3' non tradotta (UTR).

In alcune realizzazioni, la 3' UTR comprende una 3' UTR virale, una 3' UTR non virale o una combinazione di sequenze 3' UTR virali e non virali.

In alcune realizzazioni, l'UTR 3' comprende un'UTR 3' dell'alfavirus.

In alcune forme di realizzazione, l'alfavirus 3' UTR comprende il virus dell'encefalite equina venezuelana (VEEV), il virus dell'encefalite equina orientale (EEEV), il virus delle Everglades (EVEV), il virus Mucambo (MUCV), il virus della foresta Semliki (SFV), il virus Pixuna (PIXV), il virus Middleburg (MIDV), il virus Chikungunya (CHIKV), il virus O'Nyong-Nyong (ONNV), il virus Ross River (RRV), Virus Barmah Forest (BFV), Virus Getah (GETV), Virus Sagiyama (SAGV), Virus Bebaru (BEBV), Virus Mayaro (MAYV), Virus Una (UNAV), Virus Sindbis (SINV), Virus Aura (AURAV), Virus Whataroa (WHAIV), Virus Babanki (BABV), Virus Kyzylgach (KYZV), Virus dell'encefalite equina occidentale (WEEV), Highland J Sequenza 3' UTR del virus Fort Morgan (FMV), virus Ndumu (NDUV), virus Alphavirus dei salmonidi (SAV) o virus Buggy Creek (BCRV).

In alcune realizzazioni, la 3' UTR comprende una sequenza poli-A.

In alcune realizzazioni, l'UTR 3' comprende una sequenza SEQ ID NO:76.

In alcune realizzazioni, la proteina antigenica è una proteina SARS-CoV-2.

In alcune realizzazioni, la proteina antigenica è una glicoproteina spike del SARS-CoV-2.

In alcune realizzazioni, la glicoproteina spike del SARS-CoV-2 è una glicoproteina spike del SARS-CoV-2 di tipo selvaggio con una sequenza aminoacidica SEQ ID NO: 123.

In alcune realizzazioni, il secondo polinucleotide comprende una sequenza avente almeno l'85% di identità con una sequenza SEQ ID NO:121 o SEQ ID NO:122.

In alcune realizzazioni, il secondo polinucleotide comprende almeno due transgeni.

In alcune realizzazioni, un secondo transgene codifica una seconda proteina antigenica o un frammento della stessa oppure una proteina immunomodulatrice.

In alcune realizzazioni, il secondo polinucleotide comprende inoltre una sequenza che codifica un peptide 2A, un sito di ingresso ribosomiale interno (IRES) o una combinazione di questi, situato tra i transgeni.

In alcune realizzazioni, la proteina immunomodulatrice è una citochina, una chemiochina o un'interleuchina.

In alcune realizzazioni, il secondo transgene codifica una seconda proteina del coronavirus.

In alcune realizzazioni, il primo polinucleotide si trova 5' del secondo polinucleotide.

In alcune realizzazioni, la molecola di acido nucleico comprende inoltre una seconda regione intergenica situata tra il primo polinucleotide e il secondo polinucleotide.

In alcune realizzazioni, la seconda regione intergenica comprende una sequenza avente almeno l'85% di identità con una sequenza di SEQ ID NO:77.

In alcune realizzazioni, la molecola di acido nucleico è

- (a) una molecola di DNA; o
- (b) una molecola di RNA, in cui T è sostituita con U.

In alcune realizzazioni, la molecola di DNA comprende inoltre un promotore.

In alcune realizzazioni, il promotore si trova al 5' della 5'UTR.

In alcune realizzazioni, il promotore è un promotore T7, un promotore T3 o un promotore SP6.

In alcune realizzazioni, la molecola di RNA è una molecola di RNA autoreplicante.

In alcune realizzazioni, la molecola di RNA comprende inoltre un cappuccio 5'.

In alcune realizzazioni, il cappuccio 5' ha una struttura Cap 1, una struttura Cap 1 (^{m6} A), una struttura Cap 2, una struttura Cap 0 o una qualsiasi combinazione di queste.

In un altro aspetto ancora, la divulgazione fornisce una composizione comprendente una qualsiasi delle molecole di acido nucleico fornite nel presente documento. In alcune forme di realizzazione, la composizione comprende inoltre un lipide.

In alcune realizzazioni, il lipide comprende un lipide cationico ionizzabile.

In alcune realizzazioni, il lipide cationico ionizzabile ha una struttura di

o un sale farmaceuticamente accettabile dello stesso.

In un ulteriore aspetto, la presente invenzione riguarda una composizione comprendente una qualsiasi delle molecole di acido nucleico descritte nel presente documento e una formulazione lipidica.

In alcune realizzazioni, la formulazione lipidica comprende un lipide cationico ionizzabile.

In alcune realizzazioni, il lipide cationico ionizzabile ha una struttura di

o un sale farmaceuticamente accettabile dello stesso.

In alcune realizzazioni, la formulazione lipidica è selezionata tra un lipoplex, un liposoma, una nanoparticella lipidica, un vettore a base di polimero, un esosoma, un corpo lamellare, una micella e un'emulsione.

In alcune realizzazioni, la formulazione lipidica è un liposoma selezionato tra un liposoma cationico, un nanoliposoma, un proteoliposoma, un liposoma unilamellare, un liposoma multilamellare, un nanoliposoma contenente ceramide e un liposoma multivescicolare.

In alcune realizzazioni, la formulazione lipidica è una nanoparticella lipidica.

In alcune realizzazioni, la nanoparticella lipidica ha una dimensione inferiore a circa 200 nm. In alcune realizzazioni, la nanoparticella lipidica ha una dimensione inferiore a circa 150 nm. In alcune realizzazioni, la nanoparticella lipidica ha una dimensione inferiore a circa 100 nm. In alcune realizzazioni, la nanoparticella lipidica ha una dimensione compresa tra circa 55 nm e circa 90 nm.

In alcune realizzazioni, la formulazione lipidica comprende uno o più lipidi cationici.

In alcune realizzazioni, uno o più lipidi cationici sono selezionati tra 5-carbossispermilglicinadiottadecilamide (DOGS), 2,3-dioleyloxi-N-[2(spermina-carbossamido)etil]-N,N-dimetil-1-propanamminio (DOSPA), 1,2-Dioleoyl-3-Dimetilammonio-Propano (DODAP), 1,2-Dioleoyl-3-Trimetilammonio-Propano (DOTAP), 1,2-distearoyloxi-N,N-dimetil-3-amminopropano (DSDMA), 1,2-dioleyloxi-N,N-dimetil-3-amminopropano (DODMA), 1,2-dilinoleyloxi-N,N-dimetil-3-amminopropano (DLinDMA), 1,2-dilinolenilossi-N,N-dimetil-3-amminopropano (DLenDMA), cloruro di N-dioleil-N,N-dimetilammonio (DODAC), bromuro di N,N-distearil-N,N-dimetilammonio (DDAB), bromuro di N-(1,2-dimiristilossi-prop-3-il)-N,N-dimetil-N-idrossietil ammonio (DMRIE), 3-dimetilammino-2-(colest-5-en-3-beta-ossi-butan-4-ossi)-1-(cis,cis-9,12-oc-tadecadienossi)propano (CLinDMA), 2-[5'-(colest-5-en-3-beta-ossi)-3'-oxapentossi]-3-dimetil 1-1-(cis,cis-9',1-2'-ottadecadienossi)propano (CpLinDMA), N,N-dimetil-3,4-dioleylossibenzilammina (DMOBA), 1,2-N,N'-dioleylcarbamil-3-dimetilamminopropano (DOcarbDAP), 2,3-dilinoleoilossi-N,N-dimetilpropilammina (DLinDAP), 1,2-N,N'-dilinoleylocarbamil-3-dimetilamminopropano (DLincarbDAP), 1,2-dilinoleoilcarbamil-3-dimetilamminopropano (DLinCDAP), 2,2-dilinoley-4-dimetilamminometil-[1,3]-diossolano (DLin-K-DMA) e 2,2-dilinoley-4-dimetilamminoetil-[1,3]-diossolano o (DLin-K-XTC2-DMA).

In alcune realizzazioni, la formulazione lipidica comprende un lipide cationico ionizzabile.

In alcune realizzazioni, il lipide cationico ionizzabile ha una struttura di Formula I.

- o un sale o solvato farmaceuticamente accettabile dello stesso, in cui R^5 e R^6 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile $C_{15}-C_{31}$ lineare o ramificato, un alchenile C_2-C_{31} o un alchinile C_2-C_{31} e un colesterile; L^1 e L^2 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile C_1-C_{20} lineare e un alchenile C_2-C_{20} ; X è $-C(O)O-$, per cui si forma $-C(O)O-R$ o $-OC(O)-R$; X è $-C(O)O-$ per cui si forma $-C(O)O-R$ o $-OC(O)-R$ per cui si forma $-OC(O)-R$; X è S o Q; L è alchile assente o inferiore; R^1 è alchile C_1-C_6 lineare o ramificato; e R^2 e R^3 sono ciascuno selezionati indipendentemente dal gruppo costituito da un idrogeno e un alchile C_1-C_6 lineare o ramificato.

In alcune realizzazioni, il lipide cationico ionizzabile è selezionato tra

In alcune realizzazioni, il lipide cationico ionizzabile è ATX-126:

In alcune realizzazioni, la formulazione lipidica incapsula la molecola di acido nucleico.

In alcune realizzazioni, la formulazione lipidica è complessata alla molecola di acido nucleico.

In alcune realizzazioni, la formulazione lipidica comprende inoltre un lipide helper. In alcune realizzazioni, il lipide helper è un fosfolipide. In alcune realizzazioni, il lipide helper è selezionato tra dioleoilfosfatidil etanolamina (DOPE), dimiristoilfosfatidilcolina (DMPC), distearoilfosfatidilcolina (DSPC), dimiristoilfosfatidilglicerolo (DMPG), dipalmitoilfosfatidilcolina (DPPC) e fosfatidilcolina (PC). In specifiche realizzazioni, il lipide helper è distearoilfosfatidilcolina (DSPC).

In alcune realizzazioni, la formulazione lipidica comprende inoltre colesterolo.

In alcune realizzazioni, la formulazione lipidica comprende inoltre un coniugato di polietilenglicole (PEG)-lipide. In alcune realizzazioni, il coniugato PEG-lipide è PEG-DMG. In alcune realizzazioni, il PEG-DMG è PEG2000-DMG.

In alcune realizzazioni, la porzione lipidica della formulazione lipidica comprende circa il 40% molare a circa il 60% molare del lipide cationico ionizzabile, circa il 4% molare a circa il 16% molare di DSPC, circa il 30% molare a circa il 47% molare di colesterolo e circa lo 0,5% molare a circa il 3% molare di PEG2000-DMG.

In alcune realizzazioni, la porzione lipidica della formulazione lipidica comprende circa il 42% molare a circa il 58% molare del lipide cationico ionizzabile, circa il 6% molare a circa il 14% molare di DSPC, circa il 32% molare a circa il 44% molare di colesterolo e circa l'1% molare a circa il 2% molare di PEG2000-DMG.

In alcune realizzazioni, la porzione lipidica della formulazione lipidica comprende circa il 45% molare a circa il 55% molare del lipide cationico ionizzabile, circa l'8% molare a circa il 12% molare di DSPC, circa il 35% molare a circa il 42% molare di colesterolo e circa l'1,25% molare a circa l'1,75% molare di PEG2000-DMG.

In alcune realizzazioni, la composizione ha un rapporto di peso molecolare totale lipide:acido nucleico di circa 50:1 a circa 10:1. In alcune realizzazioni, la composizione ha un rapporto di peso molecolare totale lipide:acido nucleico di circa 44:1 a circa 24:1. In alcune realizzazioni, la composizione ha un rapporto di peso molecolare totale lipide:acido nucleico di circa 40:1 a circa 28:1. In alcune realizzazioni, la composizione ha un rapporto di peso molecolare totale lipide:acido nucleico di circa 38:1 a circa 30:1. In alcune realizzazioni, la composizione ha un rapporto di peso molecolare totale lipide:acido nucleico di circa 37:1 a circa 33:1. In alcune realizzazioni, la composizione comprende un tampone HEPES o TRIS a un pH di circa 7,0 a circa 8,5.

In alcune realizzazioni, il tampone HEPES o TRIS ha una concentrazione compresa tra circa 7 mg/mL e circa 15 mg/mL.

In alcune realizzazioni, la composizione comprende inoltre circa 2,0 mg/mL a circa 4,0 mg/mL di NaCl.

In alcune realizzazioni, la composizione comprende inoltre uno o più crioprotettori.

In alcune realizzazioni, uno o più crioprotettori sono selezionati tra saccarosio, glicerolo o una combinazione di saccarosio e glicerolo.

In alcune realizzazioni, la composizione comprende una combinazione di saccarosio a una concentrazione di circa 70 mg/mL a circa 110 mg/mL di saccarosio e glicerolo a una concentrazione di circa 50 mg/mL a circa 70 mg/mL.

In alcune realizzazioni, la composizione è una composizione liofilizzata.

In alcune realizzazioni, la composizione liofilizzata comprende uno o più lioprotettori.

In alcune realizzazioni, la composizione liofilizzata comprende un polossamero, sorbato di potassio, saccarosio o una qualsiasi combinazione di questi.

In alcune realizzazioni, il polossamero è il polossamero 188.

In alcune realizzazioni, la composizione liofilizzata comprende circa lo 0,01-1,0% p/p della molecola di acido nucleico.

In alcune realizzazioni, la composizione liofilizzata comprende circa l'1,0-5,0% p/p di lipidi.

In alcune realizzazioni, la composizione liofilizzata comprende circa lo 0,5-2,5% p/p di tampone TRIS.

In alcune realizzazioni, la composizione liofilizzata comprende circa dallo 0,75 al 2,75% p/p di NaCl.

In alcune realizzazioni, la composizione liofilizzata comprende circa l'85-95% p/p di uno zucchero. In alcune realizzazioni, lo zucchero è saccarosio.

In alcune realizzazioni, la composizione liofilizzata comprende circa lo 0,01-1,0% p/p di un polossamero. In alcune realizzazioni, il polossamero è il polossamero 188.

In alcune realizzazioni, la composizione liofilizzata comprende circa l'1,0-5,0% p/p di sorbato di potassio.

In alcune realizzazioni, la molecola di acido nucleico comprende

- (a) una sequenza di SEQ ID NO: 124;
- (b) una sequenza di SEQ ID NO: 124, in cui T è sostituito con U;
- (c) una sequenza di SEQ ID NO: 125; o
- (d) una sequenza di SEQ ID NO: 125, in cui T è sostituito con U.

In un altro aspetto ancora, la divulgazione fornisce una composizione di nanoparticelle lipidiche comprendente

- a. una formulazione lipidica comprendente
 - i. circa il 45% molare a circa il 55% molare di un lipide cationico ionizzabile avente la struttura di ATX-126:

- - - ii. circa l'8% molare a circa il 12% molare di DSPC;
 - iii. circa il 35% molare a circa il 42% molare di colesterolo; e
 - iv. circa 1,25 moli % a circa 1,75 moli % PEG2000-DMG; e
 - b. una molecola di acido nucleico avente almeno l'85% di identità di sequenza con SEQ ID NO:125; in cui la formulazione lipidica incapsula la molecola di acido nucleico e la nanoparticella lipidica ha una dimensione compresa tra circa 60 e circa 90 nm.

In un altro aspetto ancora, la divulgazione fornisce un metodo per somministrare una qualsiasi delle composizioni descritte nel presente documento a un soggetto che ne abbia bisogno, in cui la composizione viene somministrata per via intramuscolare, sottocutanea, intradermica, transdermica, intranasale, orale, sublinguale, endovenosa, intraperitoneale, topica, per aerosol o per via polmonare. In forme di realizzazione specifiche, la composizione viene somministrata per via intramuscolare.

In un ulteriore aspetto, la presente invenzione fornisce un metodo per somministrare una qualsiasi delle composizioni descritte nel presente documento a un soggetto che ne abbia bisogno, in cui la composizione è liofilizzata e ricostituita prima della somministrazione.

In un ulteriore aspetto, la presente invenzione fornisce un metodo per migliorare il COVID-19, che comprende la somministrazione di una qualsiasi delle composizioni descritte nel presente documento a un soggetto che ne abbia bisogno.

In alcune realizzazioni, la composizione viene somministrata una volta. In alcune realizzazioni, la composizione viene somministrata due volte.

In un ulteriore aspetto, la presente invenzione fornisce un metodo per somministrare una dose di richiamo a un soggetto vaccinato, comprendente la somministrazione di una qualsiasi delle composizioni qui descritte a un soggetto precedentemente vaccinato contro il coronavirus.

In alcune realizzazioni, la composizione viene somministrata a un dosaggio compreso tra circa 0,01 µg e circa 1.000 µg di acido nucleico.

In alcune realizzazioni, la composizione viene somministrata a un dosaggio di circa 1, 2, 5, 7,5 o 10 µg di acido nucleico.

In un ulteriore aspetto, la presente invenzione riguarda un metodo per indurre una risposta immunitaria in un soggetto, che comprende la somministrazione al soggetto di una quantità efficace di una qualsiasi delle molecole di acido nucleico qui descritte.

In alcune realizzazioni, la molecola di acido nucleico può essere somministrata per via intramuscolare, sottocutanea, intradermica, transdermica, intranasale, orale, sublinguale, endovenosa, intraperitoneale, topica, tramite aerosol o per via polmonare.

In un ulteriore aspetto, la presente invenzione fornisce un metodo per indurre una risposta immunitaria in un soggetto, comprendente la somministrazione al soggetto di una quantità efficace di una qualsiasi delle composizioni qui descritte.

In alcune realizzazioni, la composizione può essere somministrata per via intramuscolare, sottocutanea, intradermica, transdermica, intranasale, orale, sublinguale, endovenosa, intraperitoneale, topica, mediante aerosol o per via polmonare.

In alcune realizzazioni, le molecole di acido nucleico descritte nel presente documento possono essere utilizzate per indurre una risposta immunitaria alla prima proteina antigenica o a un suo frammento.

In alcune realizzazioni, le molecole di acido nucleico descritte nel presente documento possono essere utilizzate nella produzione di un medicinale per indurre una risposta immunitaria alla prima proteina antigenica o a un suo frammento.

BREVE DESCRIZIONE DEI DISEGNI

FIGURE 1 A- 1 Dmostra progettazione ed espressione di un vaccino SARS-CoV-2 in piattaforme mRNA e RNA autoreplicante (STARR™). (1 A) Diagramma schematico dei costrutti del vaccino SARS-CoV-2 autoreplicante STARR™ RNA e mRNA. Il costrutto STARR™ codifica per le quattro proteine non strutturali, ns1-ns4, del virus dell'encefalite equina venezuelana (VEEV) e della proteina spike (S) a lunghezza intera SARS-CoV-2. Il costrutto mRNA codifica per la proteina spike S a lunghezza intera SARS-CoV-2. (1 B) Caratteristiche fisiche ed efficienza di intrappolamento dell'RNA dell'LNP nei vaccini mRNA e STARR™ (RNA autoreplicante corrispondente a SEQ ID NO: 125; qui denominato "RNA STARR™ SARS-CoV-2"). (1 C) Rilevamento tramite Western blot della proteina S di SARS-CoV-2 dopo la trasfezione di cellule HEK293 con i costrutti STARR™ RNA e mRNA. (1 D) Confronto in vivo dell'espressione proteica dopo somministrazione intramuscolare (IM) di LNP contenente STARR™ RNA o mRNA esprimente luciferasi. I topi Balb/c (n=3/gruppo) sono stati iniettati IM con 0,2 µg, 2,0 µg e 10,0 µg di STARR™ RNA o mRNA in formulazione lipidica. L'espressione della luciferasi è stata misurata tramite bioluminescenza in vivo nei giorni 1, 3 e 7 dopo la somministrazione IM. Dominio S 1=S1, dominio S 2=S2, dominio transmembrana=TM, dominio citoplasmatico=CP.

FIGURE 2 A- 2 Imostrano punteggi clinici, pesi dei topi e analisi trascrittomiche dei geni immunitari dopo la vaccinazione con candidati vaccini STARR™ RNA o mRNA SARS-CoV-2. (2 A) I topi C57BL/6 sono stati immunizzati con PBS, mRNA o STARR™ SARS-CoV-2 RNA (dosi 0,2 µg, 2 µg o 10 µg), peso e punteggi clinici valutati ogni giorno, dissanguati il giorno 1 dopo l'immunizzazione, sacrificati 7 giorni dopo la vaccinazione e linfonodi raccolti. L'espressione genica dei geni infiammatori e dei geni immunitari è stata misurata nel sangue intero (al giorno 1) e nei linfonodi (al giorno 7), rispettivamente. (2 B) Espressione di IFN e geni della risposta infiammatoria nel sangue intero presentati come mappa di calore dei punteggi z. (2 C) Pesi dei linfonodi a 7 giorni dalla vaccinazione. Analisi delle componenti principali (PCA) dell'espressione genica immunitaria dopo vaccinazione con mRNA o RNA STARR™ SARS-CoV-2 a dosi (2 D) 0,2 µg, (2 E) 2 µg e (2 F) 10 µg. Grafici a vulcano del fold change dell'RNA STARR™ SARS-CoV-2 rispetto all'mRNA (asse x) e valore P Log 10 dell'RNA STARR™ SARS-CoV-2 rispetto all'mRNA (asse y) per dosi (2 G) 0,2 µg, (2 H) 2 µg e (2 I) 10 µg.

FIGURE 3 A- 3 Jmostrano risposte immunitarie cellulari in seguito alla vaccinazione con SARS-CoV-2 STARR™ RNA e mRNA. I topi C57BL/6 (n=5 per gruppo) sono stati immunizzati con 0,2 µg, 2 µg o 10 µg di STARR™ RNA o mRNA tramite IM, sacrificati al giorno 7 dopo la vaccinazione e le milze sono state analizzate per le risposte delle cellule T cellulari mediante citometria a flusso ed ELISPOT. (3 A- 3 B) Le cellule effettrici T CD8+ e C) CD4+ sono state valutate negli animali vaccinati utilizzando la colorazione di superficie per i marcatori delle cellule T e la citometria a flusso. (3 D- 3 E) Le cellule T IFN γ +CD8+ e (3 F) Il rapporto delle cellule T IFN γ +IL4+CD4+ nelle milze dei topi immunizzati è stato valutato in seguito a stimolazione ex vivo con PMA/ionomicina (IO) e colorazione intracellulare. (3 G- 3 I) Le risposte specifiche della proteina S del SARS-CoV-2 ai peptidi della proteina S raggruppati sono state valutate utilizzando test ELISPOT IFN γ dopo la vaccinazione con mRNA (3 H) o RNA STARR™ (3 I). Uno schema dei domini della proteina S è mostrato in (3 J).

FIGURE 4 A- 4 Gmostrano risposte umorali in più ceppi di topi dopo immunizzazione con mRNA e candidati vaccini STARR™. (4 A) I topi BALB/c e C57BL/6J sono stati immunizzati tramite IM con 0,2 µg, 2 µg o 10 µg di RNA o mRNA STARR™ (n=5/gruppo). Il prelievo di sangue è stato condotto al basale e nei giorni 10, 19, 30, 40, 50 e 60 dopo la vaccinazione per BALB/c e nei giorni 10, 20 e 30 per C57BL/6J. (4 B- 4 C) IgM e (4 D- 4 E) IgG contro la proteina S SARS-CoV-2 nel tempo, valutata utilizzando la proteina S intera derivata da cellule di insetto in un immuno-test Luminex (misurato come MFI). I titoli finali di IgG per la proteina S intera derivata dai mammiferi, S1, S2 e le proteine del dominio di legame del recettore (RBD) al 30° giorno dopo la vaccinazione sono stati valutati in (4 F) BALB/c e (4 G) C57BL/6J.

FIGURE 5 A- 5 Ddimostrano che l'RNA STARR™ SARS-CoV-2 provoca risposte immunitarie Th1 distorte. Sottoclassi di IgG specifiche per la proteina spike SARS-CoV-2 e rapporto di IgG2a/c/IgG1 a 30 giorni dalla vaccinazione con RNA STARR™ e mRNA in (5 A) topi BALB/c e (5 B) topi C57BL/6J. Citochina Th2 e distorsione Th1/Th2 nelle cellule T CD4 al giorno 7 dopo la vaccinazione in topi C57BL/6J misurata da ICS come (5 C) percentuale di cellule T CD4 IL4+ e (5 D) rapporto di cellule T CD4+ IFN γ +/IL4+.

FIGURE 6 A- 6 Edimostrano che l'RNA STARR™ SARS-CoV-2 suscita una risposta umorale di qualità superiore rispetto alla piattaforma mRNA. (6 A) L'avidità delle IgG specifiche per la proteina S SARS-CoV-2 al giorno 30 post-immunizzazione è stata misurata utilizzando lavaggi con urea 8M. (6 B) Anticorpo neutralizzante (titoli PRNT50) al giorno 30 post-vaccinazione contro un virus SARS-CoV-2 vivo clinicamente isolato misurato sia in BALB/c che in C57BL/6J. Le linee tratteggiate rappresentano l'intervallo di diluizione del siero (vale a dire da 1:20 a 1:320) testato da PRNT. (6 C) PRNT50 e (6 D) PRNT70 della neutralizzazione di SARS-CoV-2 al giorno 60 post-vaccinazione e sieri convalescenti da pazienti COVID-19. (6 E) Analisi di correlazione dei titoli endpoint IgG specifici di Spike contro la neutralizzazione di SARS-CoV-2 (PRNT50). PRNT: test di neutralizzazione della riduzione della placca.

FIGURE 7 A- 7 Emostrano punteggi clinici, peso corporeo e risposte immunitarie a STARR™ SARS-CoV-2 RNA e mRNA dopo il boost al giorno 30 post-prime in C57BL/6J. (7 A) Punteggi clinici e (7 B) percentuale del peso corporeo iniziale dopo vaccinazioni boost. (7 C) Risposte IgG anti-Spike dopo il boost con mRNA e STARR™ SARS-CoV-2 RNA. La linea tratteggiata grigia segna il punto di

saturazione del test sperimentale. Risposte delle cellule effettrici T CD8+ IFN γ + (variazione di volte rispetto a PBS) in animali sottoposti a primer o primer e booster con (7 D) mRNA o (7 E) candidati vaccini STARR™ SARS-CoV-2 RNA.

FIGURE 8 A- 8 B mostra i dati trascrittomici del sangue intero a 1 giorno dalla vaccinazione primaria che mostrano i conteggi delle nanostringhe per 50 ng di RNA di geni infiammatori selezionati (8 A) IFN e (8 B).

FIGURE 9 A- 9 B mostra l'analisi di correlazione della neutralizzazione di SARS-CoV-2 vivo contro le IgG leganti e le sottoclassi di IgG nei ceppi di topi BALB/c e C57BL/6J. (9 A) Analisi di correlazione di Spearman della neutralizzazione di SARS-CoV-2 (PRNT50) contro le IgG totali specifiche di diversi antigeni SARS-CoV-2, tra cui le proteine ricombinanti S, S1 e RBD. (9 B) Analisi di correlazione di Spearman della neutralizzazione di SARS-CoV-2 (PRNT50) contro le sottoclassi di IgG specifiche per SARS-CoV-2 S (IgG1 e IgG2a o IgG2c).

FIGURA 10 mostra le curve di sopravvivenza di Kaplan-Meier per topi non vaccinati (PBS) e topi vaccinati con STARR™ SARS-CoV-2 RNA dopo la somministrazione di una dose letale di virus SARS-CoV-2. Linea superiore: STARR™ SARS-CoV-2 RNA (2 μ g, 10 μ g); linea di caduta: PBS.

FIGURA 11 dimostra che la vaccinazione STARR™ SARS-CoV-2 RNA protegge dall'infezione polmonare e cerebrale da SARS-CoV-2. Livelli di RNA virale nei polmoni (FIGURA 11, a sinistra) e nei cervelli (FIGURA 11, a destra) di topi non vaccinati (PBS) e topi vaccinati con la dose indicata di RNA STARR™ SARS-CoV-2.

FIGURA 12 mostra i titoli virali nei polmoni dei topi non vaccinati (PBS) e dei topi vaccinati con la dose indicata di RNA STARR™ SARS-CoV-2 dopo l'esposizione al SARS-CoV-2.

FIGURA 13 mostra un confronto dell'immunogenicità dose-dipendente dell'RNA tra la glicoproteina SARS CoV-2 G614 e D614 espresse dall'RNA autoreplicante.

FIGURA 14 mostra uno schema che illustra un aspetto della tecnologia STARR™ e della somministrazione mediata dai lipidi.

FIGURE 15 A- 15 C mostra la durata dell'espressione del gene reporter della luciferasi per l'RNA autoreplicante (replicone) (STARR™), come (15 A) STARR™ FLuc, (15 B) STARR™ FLuc IRES-E3L e (15 C) STARR™ FLuc IRES E3L (3' UTR corto) rispetto all'mRNA.

FIGURA 16 A - 16 D mostra i risultati del test Luminex per la glicoproteina spike anti-SARS-CoV-2 IgG in due studi preclinici. I topi BALB/c sono stati vaccinati con dosi crescenti di RNA di RNA autoreplicante (SEQ ID NO:125) formulato come nanoparticelle lipidiche liofilizzate (LYO-LNP) e nanoparticelle lipidiche liquide (congelate) (Liquid-LNP). (16 A) Primo studio 0,2 μ g, (16 B) Primo studio 2 μ g, (16 C) Secondo studio 0,2 μ g e (16 D) Secondo studio 2 μ g. Il sangue è stato raccolto ed elaborato in siero in vari momenti dopo la vaccinazione e valutato per la glicoproteina spike anti-SARS-CoV-2 IgG. L'ANOVA bidirezionale, il post-test di confronto multiplo di Tukey ha confrontato LYO-LNP con Liquid-LNP dove * $p < 0,0332$, ** $p < 0,0021$, *** $p < 0,0002$, **** $p < 0,0001$.

FIGURE 17 A- 17 B mostra l'analisi dell'area sotto la curva (AUC) per la glicoproteina Spike anti-SARS-Cov-2 IgG (dati combinati del primo e del secondo studio). I risultati del test IgG sono stati combinati da due studi per valutare l'RNA autoreplicante (SEQ ID NO:125) formulato come nanoparticelle lipidiche liofilizzate (LYO-LNP) e nanoparticelle lipidiche liquide (congelate) (Liquid-LNP) a (**17 A**) 0,2 µg e (**17 B**) 2 µg. N=10/gruppo. I risultati del primo giorno di studio 19 e 31 sono stati combinati con i risultati del secondo giorno di studio 20 e 30, rispettivamente, ed è stata eseguita un'analisi dell'area sotto la curva (AUC). ANOVA unidirezionale, il post-test di confronto multiplo di Sidak ha confrontato LYO-LNP con Liquid-LNP e non ha prodotto differenze statistiche.

FIGURE 18 A- 18 D mostra la caratterizzazione della tecnologia STARR™ con espressione del transgene della luciferasi di lucciola. (**18 A**) L'espressione della luciferasi di lucciola (FLuc) da STARR™ FLuc, SINV FLuc e mRNA FLuc è stata monitorata fino al giorno 28 da In Vivo Imaging System (IVIS). La media del flusso totale (p/s) da 6 siti di iniezione in un gruppo di topi è stata tracciata a ogni punto temporale con un errore standard della media, SEM. (**18 B**) L'immagine IVIS di tre topi (6 siti di iniezione) per gruppo al giorno 14 è mostrata per ogni gruppo a cui è stato somministrato l'articolo di prova etichettato sotto l'immagine. (**18 C**) L'espressione della luciferasi da topi a cui è stato iniettato per via intramuscolare STARR™ FLuc è stata monitorata da IVIS fino a 63 giorni dopo la somministrazione. (**18 D**) L'effetto della precedente somministrazione dello scheletro del replicone è stato esaminato per STARR™ (pannello superiore) e SINV (pannello inferiore). Il replicone che codifica FLuc è stato iniettato IM 7 giorni dopo la dose di replicone con struttura omologa con un gene/sequenza irrilevante (etichettato STARR™ irr o SINV irr) al giorno 0. Come riferimento, un gruppo di topi con somministrazione di PBS al giorno 0 è stato incluso in ciascuno dei gruppi STARR™ e SINV.

FIGURA 19 mostra che STARR™ suscita una risposta IFN-gamma specifica per l'antigene. È stato utilizzato lo spot assorbente immunitario enzimatico ELISpot per contare il numero di splenociti che sono stati specificamente stimolati da un peptide antigenico della stessa sequenza di amminoacidi codificata in TA STARR™. Né nessun peptide (solo cellula) né peptide irrilevante (Bgal) non hanno suscitato IFN-gamma significativo dagli splenociti di topi vaccinati con STARR™ FLuc o TA STARR™. La stimolazione con il peptide AH1-A5 ha portato al rilevamento di cellule produttrici di IFN-gamma specificamente dai topi vaccinati con TASTARR™. La concanavalina A (ConA) è stata utilizzata come controllo positivo della produzione di IFN-gamma.

FIGURE 20 A- 20 F illustrano il tasso di crescita tumorale ridotto mediante vaccinazione con TA STARR™ in un modello murino singeneico CT26. Le cellule di carcinoma coloretale murino CT26 (5×10^5) sono state impiantate sottocutaneamente in topi BALB/c femmine di 10 settimane (n=8 per gruppo). Nei giorni 1 e 8, i topi sono stati vaccinati con STARR™ FLuc, un controllo negativo, o TA STARR™, che codifica l'epitopo AH1A5. La crescita tumorale è stata monitorata nei topi vaccinati con (**20 A**) STARR™ FLuc senza trattamento con inibitore del checkpoint; (**20 B**) STARR™ FLuc con un trattamento combinato anti-PD1/PDL1; (**20 C**) STARR™ FLuc con un trattamento combinato anti-CTLA4; (**20 D**) vaccino STARR™ senza trattamento con inibitore del checkpoint; (**20 E**) vaccino STARR™ con un trattamento combinato di anti-PD1 e anti-PDL1; e (**20 F**) vaccino STARR™ con un trattamento combinato di anti-CTLA4. Le curve di crescita tumorale individuali di un gruppo di topi a cui sono stati somministrati STARR™ FLuc e TA STARR™ sono mostrate rispettivamente nei pannelli superiore e inferiore.

FIGURA 21 illustra una protezione prolungata mediante trattamento combinato del vaccino TA STARR™ con inibitori dei checkpoint. I topi trattati con TA STARR™ combinato con anti-PD1/PDL1 o anti-CTLA4 sono risultati resistenti alla crescita del tumore in seguito alla sfida con CT26 dal giorno 25 al giorno 42. I topi naïve sono stati utilizzati come controllo per la crescita del tumore CT26.

FIGURE 22 A- 22 C mostrano i risultati della colorazione con tetramero AH1 delle cellule T CD8+ sotto forma di (22 A) un grafico e (22 B e 22 C) grafici. Gli splenociti del gruppo di topi con trattamento combinato di TA STARR™ e anti-PD1/PDL1 al giorno 42 sono stati colorati con tetramero AH1 (H-2Ld). La colorazione era specifica per le cellule T CD8+ del gruppo di topi con trattamento con TA STARR™ e la popolazione rappresentava il 9-17% delle cellule T CD8+ totali degli splenociti.

FIGURA 23 mostra i titoli HAI ottenuti per l'RNA autoreplicante (STARR™) e i costrutti mRNA che codificano l'emoagglutinina del virus dell'influenza A/California/07/2009 (H1N1).

FIGURE 24 A- 24 B mostra i livelli di replicazione dell'RNA (FIGURA 24 UN) e livelli di espressione del gene reporter della luciferasi (FIGURA 24B) per gli RNA autoreplicanti (repliconi) indicati rispetto all'mRNA.

DESCRIZIONE DETTAGLIATA

La presente divulgazione riguarda RNA autoreplicanti e acidi nucleici che li codificano per l'espressione di transgeni quali proteine antigeniche e antigeni tumorali, ad esempio. Sono inoltre forniti nel presente documento metodi di somministrazione (ad esempio, a un ospite, come un soggetto mammifero) di RNA autoreplicanti, mediante i quali l'RNA autoreplicante viene tradotto in vivo e la sequenza eterologa di codifica proteica viene espressa e, ad esempio, può suscitare una risposta immunitaria alla sequenza eterologa di codifica proteica nel ricevente o fornire un effetto terapeutico, laddove la sequenza eterologa di codifica proteica è una proteina terapeutica. Gli RNA autoreplicanti forniti nel presente documento sono utili come vaccini che possono essere generati rapidamente e che possono essere efficaci a dosi basse e/o singole. La presente divulgazione riguarda inoltre metodi di induzione di una risposta immunitaria utilizzando RNA autoreplicanti forniti nel presente documento.

In alcune realizzazioni, può essere suscitata una risposta immunitaria contro il Coronavirus: immunogeni che includono, ma non sono limitati a, quelli derivati da un coronavirus SARS, bronchite infettiva aviaria (IBV), virus dell'epatite del topo (MHV) e virus della gastroenterite trasmissibile suina (TGEV). L'immunogeno del coronavirus può essere un polipeptide spike.

Gli RNA autoreplicanti sono descritti, ad esempio, nel documento US 2018/0036398, il cui contenuto è incorporato integralmente mediante riferimento.

Definizioni

Come utilizzato nel presente documento, il termine "frammento", quando si riferisce a una proteina o a un acido nucleico, ad esempio, significa qualsiasi sequenza più corta della proteina o dell'acido nucleico a lunghezza intera. Di conseguenza, qualsiasi sequenza di un acido nucleico o di

una proteina diversa dalla sequenza di acido nucleico o proteina a lunghezza intera può essere un frammento. In alcuni aspetti, un frammento proteico include un epitopo. In altri aspetti, un frammento proteico è un epitopo.

Come utilizzato nel presente documento, il termine "acido nucleico" si riferisce a qualsiasi molecola di acido desossiribonucleico (DNA), molecola di acido ribonucleico (RNA) o analoghi di acido nucleico. Una molecola di DNA o RNA può essere a doppio filamento o a singolo filamento e può essere di qualsiasi dimensione. Esempi di acidi nucleici includono, ma non sono limitati a, DNA cromosomico, DNA plasmidico, cDNA, DNA libero da cellule (cfDNA), DNA mitocondriale, DNA cloroplasto, DNA virale, mRNA, tRNA, rRNA, RNA lungo non codificante, siRNA, micro RNA (miRNA o miR), hnRNA e RNA virale. Esempi di analoghi nucleici includono acido peptidico nucleico, acido morfolino- e bloccato nucleico, acido glicole nucleico e acido treoso nucleico. Come utilizzato nel presente documento, il termine "molecola di acido nucleico" intende includere frammenti di molecole di acido nucleico nonché qualsiasi molecola di acido nucleico a lunghezza intera o non frammentata, ad esempio. Come utilizzato nel presente documento, i termini "acido nucleico" e "molecola di acido nucleico" possono essere utilizzati in modo intercambiabile, a meno che il contesto non indichi chiaramente il contrario.

Come utilizzato nel presente documento, il termine "proteina" si riferisce a qualsiasi catena polimerica di amminoacidi. I termini "peptide" e "polipeptide" possono essere utilizzati in modo intercambiabile con il termine proteina, a meno che il contesto non indichi chiaramente il contrario, e possono anche riferirsi a una catena polimerica di amminoacidi. Il termine "proteina" comprende proteine native o artificiali, frammenti proteici e analoghi polipeptidici di una sequenza proteica. Una proteina può essere monomerică o polimerica. Il termine "proteina" comprende frammenti e varianti (inclusi frammenti di varianti) delle stesse, a meno che non sia diversamente contraddetto dal contesto.

In generale, "identità di sequenza" o "omologia di sequenza", che possono essere utilizzate in modo intercambiabile, si riferiscono a una corrispondenza esatta nucleotide-nucleotide o amminoacido-amminoacido di due polinucleotidi o sequenze polipeptidiche, rispettivamente. In genere, le tecniche per determinare l'identità di sequenza includono la determinazione della sequenza nucleotidica di un polinucleotide e/o la determinazione della sequenza amminoacidica da essa codificata o della sequenza amminoacidica di un polipeptide, e il confronto di queste sequenze con una seconda sequenza nucleotidica o amminoacidica. Come utilizzato nel presente documento, il termine "percentuale (%) di identità di sequenza" o "percentuale (%) di identità", che include anche "omologia", si riferisce alla percentuale di residui amminoacidici o nucleotidi in una sequenza che sono identici ai residui amminoacidici o nucleotidi in una sequenza di riferimento dopo aver allineato le sequenze e introdotto lacune, se necessario, per ottenere la massima percentuale di identità di sequenza, e non considerando alcuna sostituzione conservativa come parte dell'identità di sequenza. Pertanto, due o più sequenze (polinucleotidi o amminoacidi) possono essere confrontate determinando la loro "identità percentuale", detta anche "omologia percentuale". La percentuale di identità rispetto a una sequenza di riferimento (ad esempio, sequenze di acidi nucleici o amminoacidi), che può essere una sequenza all'interno di una molecola più lunga (ad esempio, polinucleotidi o polipeptidi), può essere calcolata come il numero di corrispondenze esatte tra due sequenze allineate in modo ottimale diviso per la lunghezza della sequenza di

riferimento e moltiplicato per 100. La percentuale di identità può anche essere determinata, ad esempio, confrontando le informazioni sulla sequenza utilizzando il programma per computer avanzato BLAST, inclusa la versione 2.2.9, disponibile presso i National Institutes of Health. Il programma BLAST si basa sul metodo di allineamento di Karlin e Altschul, Proc. Natl. Acad. Sci. USA 87:2264-2268 (1990) e come discusso in Altschul et al., J. Mol. Biol. 215:403-410 (1990); Karlin e Altschul, Proc. Natl. Acad. sci. USA 90:5873-5877 (1993); e Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997). In breve, il programma BLAST definisce l'identità come il numero di simboli allineati identici (ad esempio, nucleotidi o amminoacidi), diviso per il numero totale di simboli nella più breve delle due sequenze. Il programma può essere utilizzato per determinare la percentuale di identità sull'intera lunghezza delle sequenze confrontate. Sono forniti parametri predefiniti per ottimizzare le ricerche con sequenze di query brevi, ad esempio, con il programma blastp. Il programma consente anche l'utilizzo di un filtro SEG per mascherare segmenti delle sequenze di query come determinato dal programma SEG di Wootton e Federhen, Computers and Chemistry 17: 149-163 (1993). Gli intervalli di gradi desiderati di identità di sequenza sono approssimativamente dall'80% al 100% e valori interi intermedi. Le identità percentuali tra una sequenza di riferimento e una sequenza dichiarata possono essere almeno dell'80%, almeno dell'85%, almeno del 90%, almeno del 95%, almeno del 98%, almeno del 99%, almeno del 99,5% o almeno del 99,9%. In generale, una corrispondenza esatta indica il 100% di identità per tutta la lunghezza della sequenza di riferimento. Altri programmi e metodi per confrontare sequenze e/o valutare l'identità di sequenza includono l'algoritmo Needleman-Wunsch (vedere, ad esempio, l'allineatore EMBOSS Needle disponibile su ebi.ac.uk/Tools/psa/emboss_needle/, facoltativamente con impostazioni predefinite), l'algoritmo Smith-Waterman (vedere, ad esempio, l'allineatore EMBOSS Water disponibile su ebi.ac.uk/Tools/psa/emboss_water/, facoltativamente con impostazioni predefinite), il metodo di ricerca di similarità di Pearson e Lipman, 1988, Proc. Natl. Acad. Sci. USA 85, 2444, o programmi per computer che utilizzano questi algoritmi (GAP, BESTFIT, FASTA, BLAST P, BLAST N e TFASTA in Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, Wis.). In alcuni aspetti, il riferimento all'identità percentuale di sequenza si riferisce all'identità di sequenza misurata utilizzando BLAST (Basic Local Alignment Search Tool). In altri aspetti, ClustalW viene utilizzato per l'allineamento di sequenze multiple. L'allineamento ottimale può essere valutato utilizzando qualsiasi parametro idoneo di un algoritmo scelto, inclusi i parametri predefiniti.

Come utilizzato nel presente documento, il termine "farmaco" o "medicamento" indica una formulazione o composizione farmaceutica come descritto nel presente documento.

Come qui utilizzato, le forme singolari "a", "an" e "the" includono riferimenti al plurale, a meno che il contesto non indichi chiaramente diversamente. Pertanto, ad esempio, i riferimenti a "the method" includono uno o più metodi e/o fasi del tipo descritto nel presente documento che diventeranno evidenti a quelle persone esperte nel settore dopo aver letto questa informativa e così via.

"Circa" come utilizzato nel presente documento quando si fa riferimento a un valore misurabile come una quantità, una durata temporale e simili, è inteso a comprendere variazioni di +20%, o ±10%, o ±5%, o anche ±1% rispetto al valore specificato, in quanto tali variazioni sono appropriate per i metodi divulgati o per eseguire i metodi divulgati.

Il termine "espressione" si riferisce al processo mediante il quale una sequenza di acidi nucleici o un polinucleotide viene trascritto da uno stampo di DNA (ad esempio in mRNA o altro trascritto di RNA) e/o al processo mediante il quale un mRNA trascritto o altro RNA viene successivamente tradotto in peptidi, polipeptidi o proteine. I trascritti e i polipeptidi codificati possono essere collettivamente definiti "prodotto genico".

Come qui utilizzati, i termini "RNA autoreplicante", "RNA autoreplicante e autoreplicante", "RNA autoamplificante (saRNA)" e "replicone" possono essere utilizzati in modo intercambiabile, a meno che il contesto non indichi chiaramente il contrario. In genere, il termine "replicone" o "replicone virale" si riferisce a un RNA subgenomico autoreplicante derivato da un genoma virale che include geni virali che codificano proteine non strutturali importanti per la replicazione virale e che non ha geni virali che codificano proteine strutturali. Un RNA autoreplicante può codificare ulteriori RNA subgenomici che non sono in grado di autoreplicarsi.

Come qui utilizzato, "collegato operativamente", "collegamento operabile", "collegato operativamente" o equivalenti grammaticali degli stessi si riferiscono alla giustapposizione di elementi genetici, ad esempio un promotore, un potenziatore, una sequenza di poliadenilazione, ecc., in cui gli elementi sono in una relazione che consente loro di operare nel modo previsto. Ad esempio, un elemento regolatore, che può comprendere sequenze promotore e/o potenziatore, è collegato operativamente a una regione codificante se l'elemento regolatore aiuta ad avviare la trascrizione della sequenza codificante. Possono esserci residui intermedi tra l'elemento regolatore e la regione codificante fintanto che questa relazione funzionale viene mantenuta.

Molecole di acido nucleico

In alcune forme di realizzazione, nel presente documento sono fornite molecole di acido nucleico comprendenti: (i) un primo polinucleotide che codifica una o più proteine di replicazione virale, in cui il primo polinucleotide è ottimizzato per codoni rispetto a un polinucleotide di tipo selvatico che codifica una o più proteine di replicazione virale; e (ii) un secondo polinucleotide comprendente un primo transgene che codifica una prima proteina antigenica o un frammento della stessa, in cui la prima proteina antigenica è una proteina del coronavirus.

Una molecola di RNA può codificare un singolo immunogeno polipeptidico o più polipeptidi. Più immunogeni possono essere presentati come un singolo immunogeno polipeptidico (polipeptide di fusione) o come polipeptidi separati. Se gli immunogeni sono espressi come polipeptidi separati da un replicone, uno o più di questi possono essere forniti con un IRES a monte o un elemento promotore virale aggiuntivo. In alternativa, più immunogeni possono essere espressi da una poliproteina che codifica singoli immunogeni fusi a una breve proteasi autocatalitica (ad esempio la proteina 2A del virus della febbre aftosa) o come inteine.

In alcune forme di realizzazione, sono inoltre fornite molecole di acido nucleico comprendenti: (i) un primo polinucleotide comprendente una sequenza avente almeno l'80% di identità con una sequenza di SEQ ID NO:72; e (ii) un secondo polinucleotide comprendente un primo transgene che codifica una prima proteina antigenica o un frammento della stessa.

Ottimizzazione del codone

In alcune realizzazioni, i primi polinucleotidi di molecole di acido nucleico qui forniti che codificano una o più proteine di replicazione virale includono sequenze codon-ottimizzate. Come usato qui, il termine "codon-ottimizzato" significa che un polinucleotide, una sequenza di acido nucleico o una sequenza codificante è stato riprogettato rispetto a un polinucleotide di tipo selvatico o di riferimento, una sequenza di acido nucleico o una sequenza codificante scegliendo codoni diversi senza alterare la sequenza amminoacidica della proteina codificata. Di conseguenza, l'ottimizzazione codon-ottimizzata si riferisce generalmente alla sostituzione di codoni con codoni sinonimi per ottimizzare l'espressione di una proteina mantenendo invariata la sequenza amminoacidica della proteina tradotta. L'ottimizzazione codon-ottimizzata di una sequenza può aumentare i livelli di espressione proteica (Gustafsson et al., Codon bias and heterologous protein expression. 2004, Trends Biotechnol 22: 346-53) delle proteine codificate, ad esempio, e fornire altri vantaggi. Variabili quali la preferenza nell'uso dei codoni misurata dall'indice di adattamento dei codoni (CAI), ad esempio la presenza o la frequenza di U e di altri nucleotidi, le strutture secondarie dell'mRNA, le sequenze cis-regolatrici, il contenuto di GC e altre variabili possono essere correlate ai livelli di espressione proteica (Villalobos et al., Gene Designer: uno strumento di biologia sintetica per la costruzione di segmenti di DNA artificiali. 2006, BMC Bioinformatics 7:285).

Qualsiasi metodo di ottimizzazione dei codoni può essere utilizzato per ottimizzare i polinucleotidi e le molecole di acido nucleico qui forniti, e qualsiasi variabile può essere modificata dall'ottimizzazione dei codoni. Di conseguenza, è possibile utilizzare qualsiasi combinazione di metodi di ottimizzazione dei codoni. I metodi esemplari includono il metodo dell'indice di adattamento dei codoni (CAI) elevato, il metodo Low U e altri. Il metodo CAI sceglie un codone sinonimo più frequentemente utilizzato per un'intera sequenza di codifica proteica. Ad esempio, il codone più frequentemente utilizzato per ciascun amminoacido può essere dedotto da 74.218 geni codificanti proteine da un genoma umano. Il metodo Low U ha come target i codoni contenenti U che possono essere sostituiti con un codone sinonimo con meno gruppi U, generalmente senza modificare altri codoni. Se vi è più di una scelta per la sostituzione, è possibile selezionare il codone più frequentemente utilizzato. Qualsiasi polinucleotide, sequenza di acido nucleico o sequenza di codoni qui fornita può essere ottimizzata per codoni.

In alcune realizzazioni, la sequenza nucleotidica di qualsiasi regione dei modelli di RNA o DNA descritti nel presente documento può essere ottimizzata per codoni. Preferibilmente, il modello di cDNA primario può includere la riduzione dell'occorrenza o della frequenza di comparsa di determinati nucleotidi nel filamento modello. Ad esempio, l'occorrenza di un nucleotide in un modello può essere ridotta a un livello inferiore al 25% di detti nucleotidi nel modello. In ulteriori esempi, l'occorrenza di un nucleotide in un modello può essere ridotta a un livello inferiore al 20% di detti nucleotidi nel modello. In alcuni esempi, l'occorrenza di un nucleotide in un modello può essere ridotta a un livello inferiore al 16% di detti nucleotidi nel modello. Preferibilmente, l'occorrenza di un nucleotide in un modello può essere ridotta a un livello inferiore al 15% e preferibilmente può essere ridotta a un livello inferiore al 12% di detti nucleotidi nel modello.

In alcune realizzazioni, il nucleotide ridotto è l'uridina. Ad esempio, la presente divulgazione fornisce acidi nucleici con contenuto di uracile alterato in cui almeno un codone nella sequenza di tipo selvatico è stato sostituito con un codone alternativo per generare una sequenza alterata di uracile. Le sequenze di uracile alterate possono avere almeno una delle seguenti proprietà:

- (i) un aumento o una diminuzione del contenuto globale di uracile (vale a dire, la percentuale di uracile del contenuto totale di nucleotidi nell'acido nucleico di una sezione dell'acido nucleico, ad esempio, il frame di lettura aperto);
- (ii) un aumento o una diminuzione del contenuto locale di uracile (vale a dire, i cambiamenti nel contenuto di uracile sono limitati a sottosequenze specifiche);
- (iii) un cambiamento nella distribuzione dell'uracile senza un cambiamento nel contenuto globale di uracile;
- (iv) un cambiamento nel clustering dell'uracile (ad esempio, numero di cluster, posizione dei cluster o distanza tra i cluster); o
- (v) combinazioni di questi.

In alcune realizzazioni, la percentuale di nucleobasi uraciliche nella sequenza di acidi nucleici è ridotta rispetto alla percentuale di nucleobasi uraciliche nella sequenza di acidi nucleici di tipo selvatico. Ad esempio, il 30% di nucleobasi può essere uracile nella sequenza di tipo selvatico, ma le nucleobasi che sono uracile sono preferibilmente inferiori al 15%, preferibilmente inferiori al 12% e preferibilmente inferiori al 10% delle nucleobasi nelle sequenze di acidi nucleici della divulgazione. La percentuale di contenuto di uracile può essere determinata dividendo il numero di uracile in una sequenza per il numero totale di nucleotidi e moltiplicando per 100.

In alcune realizzazioni, la percentuale di basi azotate uraciliche in una sottosequenza della sequenza di acidi nucleici è ridotta rispetto alla percentuale di basi azotate uraciliche nella corrispondente sottosequenza della sequenza di tipo selvatico. Ad esempio, la sequenza di tipo selvatico può avere una regione terminale 5' (ad esempio, 30 codoni) con un contenuto locale di uracile del 30%, e il contenuto di uracile in quella stessa regione potrebbe essere ridotto preferibilmente al 15% o inferiore, preferibilmente al 12% o inferiore e preferibilmente al 10% o inferiore nelle sequenze di acidi nucleici della divulgazione. Queste sottosequenze possono anche essere parte delle sequenze di tipo selvatico delle sequenze UTR eterologhe 5' e 3' della presente divulgazione.

In alcune realizzazioni, i codoni nella sequenza di acidi nucleici della divulgazione riducono o modificano, ad esempio, il numero, la dimensione, la posizione o la distribuzione dei cluster di uracile che potrebbero avere effetti deleteri sulla traduzione proteica. Sebbene un contenuto di uracile inferiore sia desiderabile sotto certi aspetti, il contenuto di uracile, e in particolare il contenuto di uracile locale, di alcune sottosequenze della sequenza di tipo selvatico può essere maggiore della sequenza di tipo selvatico e mantenere comunque caratteristiche benefiche (ad esempio, espressione aumentata).

In alcune realizzazioni, la sequenza modificata con uracile induce una risposta del recettore Toll-Like (TLR) inferiore rispetto alla sequenza di tipo selvatico. Diversi TLR riconoscono e rispondono agli acidi nucleici. È stato dimostrato che l'RNA a doppio filamento (ds)RNA, un frequente costituente virale, attiva TLR3. L'RNA a singolo filamento (ss)RNA attiva TLR7. Gli oligonucleotidi di RNA, ad esempio RNA con legami internucleotidici fosforotioati, sono ligandi del TLR8 umano. Il DNA contenente motivi CpG non metilati, caratteristici del DNA batterico e virale, attiva TLR9.

Come utilizzato nel presente documento, il termine "risposta TLR" è definito come il riconoscimento di RNA a singolo filamento da parte di un recettore TLR7 e preferibilmente comprende la degradazione dell'RNA e/o le risposte fisiologiche causate dal riconoscimento dell'RNA a singolo filamento da parte del recettore. I metodi per determinare e quantificare il legame di un RNA a un TLR7 sono noti nell'arte. Analogamente, i metodi per determinare se un RNA ha innescato una risposta fisiologica mediata da TLR7 (ad esempio, secrezione di citochine) sono ben noti nell'arte. In alcune realizzazioni, una risposta TLR può essere mediata da TLR3, TLR8 o TLR9 invece che da TLR7. La soppressione della risposta mediata da TLR7 può essere ottenuta tramite modifica dei nucleosidi. L'RNA subisce oltre un centinaio di diverse modifiche dei nucleosidi in natura. L'rRNA umano, ad esempio, ha dieci volte più pseudouracile (P) e 25 volte più nucleosidi 2'-O-metilati rispetto all'rRNA batterico. L'RNA batterico non contiene modifiche nucleosidiche, mentre gli RNA dei mammiferi hanno nucleosidi modificati come la 5-metilcitidina (m5C), la N6-metiladenosina (m6A), l'inosina e molti nucleosidi 2'-O-metilati oltre alla N7-metilguanosina (m7G).

In alcune realizzazioni, il contenuto di uracile dei polinucleotidi qui divulgati è inferiore a circa il 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%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2% o 1% delle basi azotate totali nella sequenza nella sequenza di riferimento. In alcune realizzazioni, il contenuto di uracile dei polinucleotidi qui divulgati è compreso tra circa il 5% e circa il 25%. In alcune realizzazioni, il contenuto di uracile dei polinucleotidi qui divulgati è compreso tra circa il 15% e circa il 25%.

In alcune realizzazioni, i primi polinucleotidi di molecole di acido nucleico qui forniti comprendono una sequenza avente almeno l'80%, almeno l'85%, almeno il 90%, almeno il 91%, almeno il 92%, almeno il 93%, almeno il 94%, almeno il 95%, almeno il 96%, almeno il 97%, almeno il 97,5%, almeno il 98%, almeno il 98,5%, almeno il 99%, almeno il 99,5%, almeno il 99,6%, almeno il 99,7%, almeno il 99,8%, almeno il 99,9% e qualsiasi numero o intervallo intermedio, di identità a una sequenza di SEQ ID NO:72. In alcune realizzazioni, i primi polinucleotidi di molecole di acido nucleico qui forniti comprendono una sequenza di SEQ ID NO:72.

In alcuni aspetti, i primi polinucleotidi e i secondi polinucleotidi delle molecole di acido nucleico qui forniti sono inclusi nella stessa (vale a dire, una singola) o in molecole di acido nucleico separate. In genere, i primi polinucleotidi e i secondi polinucleotidi delle molecole di acido nucleico qui forniti sono inclusi in una singola molecola di acido nucleico. In un aspetto, il primo polinucleotide si trova 5' del secondo polinucleotide. In un aspetto, i primi polinucleotidi e i secondi polinucleotidi delle molecole di acido nucleico qui forniti sono inclusi in molecole di acido nucleico separate. In un altro aspetto ancora, i primi polinucleotidi e i secondi polinucleotidi sono inclusi in due molecole di acido nucleico separate.

In alcuni aspetti, i primi polinucleotidi e i secondi polinucleotidi sono inclusi nella stessa (vale a dire, una singola) molecola di acido nucleico. I primi polinucleotidi e i secondi polinucleotidi delle molecole di acido nucleico qui forniti possono essere contigui, vale a dire, adiacenti l'uno all'altro senza nucleotidi nel mezzo. In un aspetto, una regione intergenica è situata tra il primo polinucleotide e il secondo polinucleotide. In un altro aspetto, la regione intergenica situata tra il

primo polinucleotide e il secondo polinucleotide è una seconda regione intergenica, con una prima regione intergenica inclusa nel primo polinucleotide come descritto di seguito. Come utilizzati nel presente documento, i termini "regione intergenica" e "sequenza intergenica" possono essere utilizzati in modo intercambiabile, a meno che il contesto non indichi chiaramente diversamente.

Una regione intergenica situata tra il primo e il secondo polinucleotide può avere qualsiasi lunghezza e qualsiasi sequenza nucleotidica. Ad esempio, la regione intergenica tra il primo polinucleotide e il secondo polinucleotide può includere circa un nucleotide, circa due nucleotidi, circa tre nucleotidi, circa quattro nucleotidi, circa cinque nucleotidi, circa sei nucleotidi, circa sette nucleotidi, circa otto nucleotidi, circa nove nucleotidi, circa dieci nucleotidi, circa 11 nucleotidi, circa 12 nucleotidi, circa 13 nucleotidi, circa 14 nucleotidi, circa 15 nucleotidi, circa 16 nucleotidi, circa 17 nucleotidi, circa 18 nucleotidi, circa 19 nucleotidi, circa 20 nucleotidi, circa 21 nucleotidi, circa 22 nucleotidi, circa 23 nucleotidi, circa 24 nucleotidi, circa 25 nucleotidi, circa 26 nucleotidi, circa 27 nucleotidi, circa 28 nucleotidi, circa 29 nucleotidi, circa 30 nucleotidi, circa 31 nucleotidi, circa 32 nucleotidi, circa 33 nucleotidi, circa 34 nucleotidi, circa 35 nucleotidi, circa 36 nucleotidi, circa 37 nucleotidi, circa 38 nucleotidi, circa 39 nucleotidi, circa 40 nucleotidi, circa 41 nucleotidi, circa 42 nucleotidi, circa 43 nucleotidi, circa 44 nucleotidi, circa 45 nucleotidi, circa 46 nucleotidi, circa 47 nucleotidi, circa 48 nucleotidi, circa 49 nucleotidi, circa 50 nucleotidi, circa 60 nucleotidi, circa 70 nucleotidi, circa 80 nucleotidi, circa 90 nucleotidi, circa 100 nucleotidi, circa 125 nucleotidi, circa 150 nucleotidi, circa 175 nucleotidi, circa 200 nucleotidi, circa 250 nucleotidi, circa 300 nucleotidi, circa 350 nucleotidi, circa 400 nucleotidi, circa 450 nucleotidi, circa 500 nucleotidi, circa 600 nucleotidi, circa 700 nucleotidi, circa 800 nucleotidi, circa 1.000 nucleotidi, circa 1.500 nucleotidi, circa 2.000 nucleotidi, circa 2.500 nucleotidi, circa 3.000 nucleotidi, circa 3.500 nucleotidi, circa 4.000 nucleotidi, circa 4.500 nucleotidi, circa 5.000 nucleotidi, circa 6.000 nucleotidi, circa 7.000 nucleotidi, circa 8.000 nucleotidi, circa 9.000 nucleotidi, circa 10.000 nucleotidi e qualsiasi numero o intervallo intermedio. In un aspetto, la regione intergenica tra il primo e il secondo polinucleotide comprende circa 10-100 nucleotidi, circa 10-200 nucleotidi, circa 10-300 nucleotidi, circa 10-400 nucleotidi o circa 10-500 nucleotidi. In un altro aspetto, la regione intergenica tra il primo e il secondo polinucleotide comprende circa 1-10 nucleotidi, circa 1-20 nucleotidi, circa 1-30 nucleotidi, circa 1-40 nucleotidi o circa 1-50 nucleotidi. In un altro aspetto ancora, la regione comprende circa 44 nucleotidi. In un aspetto, la regione intergenica tra il primo e il secondo polinucleotide delle molecole di acido nucleico qui fornite è una seconda regione intergenica.

In un aspetto, la regione intergenica tra il primo e il secondo polinucleotide include una sequenza virale. La regione intergenica tra il primo e il secondo polinucleotide può includere una sequenza di qualsiasi virus, come ad esempio gli alfavirus e i rubivirus. In un aspetto, la regione intergenica tra il primo polinucleotide e il secondo polinucleotide comprende una sequenza di alfavirus, come una sequenza di virus dell'encefalite equina venezuelana (VEEV), virus dell'encefalite equina orientale (EEEV), virus delle Everglades (EVEV), virus Mucambo (MUCV), virus della foresta di Semliki (SFV), virus Pixuna (PIXV), virus Middleburg (MIDV), virus Chikungunya (CHIKV), virus O'Nyong-Nyong (ONNV), virus Ross River (RRV), virus della foresta di Barmah (BFV), virus Getah (GETV), virus Sagiyama (SAGV), virus Bebaru (BEBV), virus Mayaro (MAYV), virus Una (UNAV), virus Sindbis (SINV), virus Aura (AURAV), virus Whataroa (WHAV), virus Babanki (BABV), virus

Kyzylagach (KYZV), Western Equine Encephalitis Virus (WEEV), Highland J Virus (HJV), Fort Morgan Virus (FMV), Ndumu Virus (NDUV), Salmonid Alphavirus (SAV), Buggy Creek Virus (BCRV) o qualsiasi combinazione di questi. In un altro aspetto, la regione intergenica tra il primo e il secondo polinucleotide comprende una sequenza del Venezuelan Equine Encephalitis Virus (VEEV). In un altro aspetto ancora, la regione intergenica tra il primo e il secondo polinucleotide comprende una sequenza avente almeno l'85%, almeno il 90%, almeno il 91%, almeno il 92%, almeno il 93%, almeno il 94%, almeno il 95%, almeno il 96%, almeno il 97%, almeno il 97,5%, almeno il 98%, almeno il 98,5%, almeno il 99%, almeno il 99,5%, almeno il 99,6%, almeno il 99,7%, almeno il 99,8%, almeno il 99,9% e qualsiasi numero o intervallo intermedio, identità con SEQ ID NO:77. In un ulteriore aspetto, la regione intergenica tra il primo e il secondo polinucleotide comprende una sequenza di SEQ ID NO:77. In un ulteriore aspetto, la regione intergenica tra il primo e il secondo polinucleotide è una seconda regione intergenica comprendente una sequenza avente almeno l'85% di identità con SEQ ID NO:77.

Nucleotidi naturali e modificati

Un RNA autoreplicante della divulgazione può comprendere uno o più nucleotidi modificati chimicamente. Esempi di monomeri di acido nucleico includono nucleotidi non naturali, modificati e modificati chimicamente, inclusi tutti i nucleotidi noti nell'arte. I nucleotidi possono essere modificati artificialmente nella porzione di base o nella porzione di zucchero. In natura, la maggior parte dei polinucleotidi comprende nucleotidi che sono nucleotidi "non modificati" o "naturali", che includono le basi puriniche adenina (A) e guanina (G) e le basi pirimidiniche timina (T), citosina (C) e uracile (U). Queste basi sono in genere fissate a un ribosio o deossi ribosio nella posizione 1'. È stato dimostrato che l'uso di polinucleotidi di RNA comprendenti nucleotidi modificati chimicamente migliora l'espressione dell'RNA, i tassi di espressione, l'emivita e/o le concentrazioni proteiche espresse. Anche i polinucleotidi di RNA comprendenti nucleotidi modificati chimicamente si sono rivelati utili nell'ottimizzazione della localizzazione delle proteine, evitando così risposte biologiche dannose, come risposte immunitarie e/o percorsi di degradazione.

Esempi di nucleotidi modificati o chimicamente modificati includono 5-idrossicitidine, 5-alchilcitidine, 5-idrossialchilcitidine, 5-carbossicitidine, 5-formilcitidine, 5-alcossicitidine, 5-alchinilcitidine, 5-alocitidine, 2-tiocitidine, N4-alchilcitidine, N4-amminocitidine, N4-acetilcitidine e N4,N4-dialchilcitidine.

Esempi di nucleotidi modificati o chimicamente modificati includono 5-idrossimetilcitidina, 5-metilcitidina, 5-idrossimetilcitidina, 5-carbossimetilcitidina, 5-formilmetilcitidina, 5-metossimetilcitidina, 5-propinilmetilcitidina, 5-bromocitidina, 5-iodocitidina, 2-tiocitidina; N4-metilcitidina, N4-amminocitidina, N4-acetilcitidina e N4,N4-dimetilcitidina.

Esempi di nucleotidi modificati o chimicamente modificati includono 5-idrossiuridine, 5-alchiluridine, 5-idrossialchiluridine, 5-carbossiuridine, 5-carbossialchilesteruridine, 5-formiluridine, 5-alcossiuridine, 5-alchiniluridine, 5-alouridine, 2-tiouridine e 6-alchiluridine.

Esempi di nucleotidi modificati o chimicamente modificati includono 5-idrossiuridina, 5-metiluridina, 5-idrossimetiluridina, 5-carbossiuridina, 5-carbossimetilesteruridina, 5-

formiluridina, 5-metossiuridina (qui indicata anche come "5MeOU"), 5-propiniluridina, 5-bromouridina, 5-fluorouridina, 5-iodouridina, 2-tiouridina e 6-metiluridina.

Esempi di nucleotidi modificati o modificati chimicamente includono 5-metossicarbonilmetil-2-tiouridina, 5-metilamminometil-2-tiouridina, 5-carbamoilmetiluridina, 5-carbamoilmetil-2'-O-metiluridina, 1-metil-3-(3-ammino-3-carbossipropi)pseudouridina, 5-metilamminometil-2-selenouridina, 5-carbossimetiluridina, 5-metildiidrouridina, 5-taurinometiluridina, 5-taurinometil-2-tiouridina, 5-(isopentenilaminometil)uridina, 2'-O-metilpseudouridina, 2-tio-2'-O-metiluridina e 3,2'-O-dimetiluridina.

Esempi di nucleotidi modificati o modificati chimicamente includono N6-metiladenosina, 2-amminoadenosina, 3-metiladenosina, 8-azaadenosina, 7-dezaadenosina, 8-ossoadenosina, 8-bromoadenosina, 2-metiltio-N6-metiladenosina, N6-isopenteniladenosina, 2-metiltio-N6-isopenteniladenosina, N6-(cis-idrossiisopentenil)adenosina, 2-metiltio-N6-(cis-idrossiisopentenil)adenosina, N6-glicinilcarbamoiladenosina, N6-treonilcarbamoil-adenosina, N6-metil-N6-treonilcarbamoil-adenosina, 2-metiltio-N6-treonilcarbamoil-adenosina, N6,N6-dimetiladenosina, N6-idrossinorvalilcarbamoiladenosina, 2-metiltio-N6-idrossinorvalilcarbamoil-adenosina, N6-acetil-adenosina, 7-metil-adenina, 2-metiltio-adenina, 2-metossi-adenina, 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-ammino-N6-metil-purina, 1-tio-adenosina, 2'-F-ara-adenosina, 2'-F-adenosina, 2'-OH-ara-adenosina e N6-(19-ammino-pentaoxanonadecil)-adenosina.

Esempi di nucleotidi modificati o chimicamente modificati includono N1-alchilguanosine, N2-alchilguanosine, tienoguanosine, 7-dezaguanosine, 8-ossoguanosine, 8-bromoguanosine, O6-alchilguanosine, xantosine, inosine e N1-alchilinosine.

Esempi di nucleotidi modificati o chimicamente modificati includono N1-metilguanosina, N2-metilguanosina, tienoguanosina, 7-dezaguanosina, 8-ossoguanosina, 8-bromoguanosina, O6-metilguanosina, xantosina, inosina e N1-metilinosina.

Esempi di nucleotidi modificati o chimicamente modificati includono le pseudouridine. Esempi di pseudouridine includono N1-alchilpseudouridine, N1-cicloalchilpseudouridine, N1-idrossipseudouridine, N1-idrossialchilpseudouridine, N1-fenilpseudouridine, N1-fenilalchilpseudouridine, N1-amminoalchilpseudouridine, N3-alchilpseudouridine, N6-alchilpseudouridine, N6-alcossipseudouridine, N6-idrossipseudouridine, N6-idrossialchilpseudouridine, N6-morfolinopseudouridine, N6-fenilpseudouridine e N6-alopseudouridine. Esempi di pseudouridine includono N1-alchil-N6-alchilpseudouridine, N1-alchil-N6-alcossipseudouridine, N1-alchil-N6-idrossipseudouridine, N1-alchil-N6-idrossialchilpseudouridine, N1-alchil-N6-morfolinopseudouridine, N1-alchil-N6-fenilpseudouridine e N1-alchil-N6-alopseudouridine. In questi esempi, i sostituenti alchilici, cicloalchilici e fenilici possono essere non sostituiti o ulteriormente sostituiti con sostituenti alchilici, alogeni, aloalchilici, amminici o nitro.

Esempi di pseudouridine includono N1-metilpseudouridina (qui indicata anche come "N1MPU"), N1-etilpseudouridina, N1-propilpseudouridina, N1-ciclopropilpseudouridina, N1-

fenilpseudouridina, N1-amminometilpseudouridina, N3-metilpseudouridina, N1-idrossipseudouridina e N1-idrossimetilpseudouridina.

Esempi di monomeri di acidi nucleici includono nucleotidi modificati e chimicamente modificati, compresi tutti i nucleotidi di questo tipo noti nel settore.

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono qualsiasi nucleotide noto nell'arte, ad esempio, 2'-O-metil ribonucleotidi, 2'-O-metil nucleotidi purinici, 2'-deossi-2'-fluoro ribonucleotidi, 2'-deossi-2'-fluoro nucleotidi pirimidinici, 2'-deossi ribonucleotidi, 2'-deossi nucleotidi purinici, nucleotidi di base universali, 5-C-metil-nucleotidi e residui monomerici desossibasici invertiti.

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono nucleotidi stabilizzati all'estremità 3', nucleotidi gliceril 3', nucleotidi abasici invertiti 3' e timidina invertita 3'.

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono nucleotidi di acido nucleico bloccati (LNA), nucleotidi 2'-O,4'-C-metilen-(D-ribofuranosil), nucleotidi 2'-metossietossi (MOE), nucleotidi 2'-metil-tio-etile, nucleotidi 2'-deossi-2'-fluoro e nucleotidi 2'-O-metil. In una forma di realizzazione esemplificativa, il monomero modificato è un nucleotide di acido nucleico bloccati (LNA).

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono DNA modificati con 2'-O-metossietile (cMOE) e 2'-O-etile (cEt) vincolati a 2',4'.

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono i nucleotidi 2'-ammino, i nucleotidi 2'-O-ammino, i nucleotidi 2'-C-allilici e i nucleotidi 2'-O-allilici.

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono i nucleotidi N6-metiladenosina.

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono monomeri nucleotidici con basi modificate 5-(3-ammino)propiluridina, 5-(2-mercapto)etiluridina, 5-bromuridina; 8-bromoguanosina o 7-deazaadenosina.

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono i nucleotidi sostituiti con 2'-O-amminopropile.

Esempi di monomeri nucleotidici modificati e chimicamente modificati includono la sostituzione del gruppo 2'-OH di un nucleotide con un 2'-R, un 2'-OR, un 2'-alogeno, un 2'-SR o un 2'-ammino, dove R può essere H, alchile, alchenile o alchinile.

L'esempio di modifiche di base sopra descritto può essere combinato con modifiche aggiuntive della struttura dei nucleosidi o dei nucleotidi, tra cui modifiche dello zucchero e modifiche di legame. Certi monomeri nucleotidici modificati o chimicamente modificati possono essere trovati in natura.

Le modifiche nucleotidiche preferite includono N1-metilpseudouridina e 5-metossiuridina.

Proteine di replicazione virale e polinucleotidi che le codificano

In alcune forme di realizzazione, nel presente documento sono fornite molecole di acido nucleico comprendenti un primo polinucleotide che codifica una o più proteine di replicazione virale. Come utilizzato nel presente documento, il termine "proteina di replicazione" o "proteina di replicazione virale" si riferisce a qualsiasi proteina o a qualsiasi subunità proteica di un complesso proteico che funziona nella replicazione di un genoma virale. In generale, le proteine di replicazione virale sono proteine non strutturali. Le proteine di replicazione virale codificate da molecole di acido nucleico fornite nel presente documento possono funzionare nella replicazione di qualsiasi genoma virale. Il genoma virale può essere un genoma di RNA a singolo filamento a senso positivo, un genoma di RNA a singolo filamento a senso negativo, un genoma di RNA a doppio filamento, un genoma di DNA a singolo filamento a senso positivo, un genoma di DNA a singolo filamento a senso negativo o un genoma di DNA a doppio filamento. I genomi virali possono includere una singola molecola di acido nucleico o più di una molecola di acido nucleico. Le molecole di acido nucleico fornite nel presente documento possono codificare una o più proteine di replicazione virale da qualsiasi virus o famiglia di virus, inclusi virus animali e virus vegetali, ad esempio. Le proteine di replicazione virale codificate dai primi polinucleotidi inclusi nelle molecole di acido nucleico fornite nel presente documento possono essere espresse da RNA autoreplicante.

Le prime sequenze polinucleotidiche di molecole di acido nucleico fornite nel presente documento possono codificare una o più proteine di replicazione del togavirus. In alcuni aspetti, una o più proteine di replicazione virale codificate dai primi polinucleotidi di molecole di acido nucleico fornite nel presente documento sono proteine di alfavirus. In alcune forme di realizzazione, una o più proteine di replicazione virale codificate dai primi polinucleotidi di molecole di acido nucleico fornite nel presente documento sono proteine di rubivirus. Le prime sequenze polinucleotidiche di molecole di acido nucleico fornite nel presente documento possono codificare qualsiasi proteina di replicazione di alfavirus e qualsiasi proteina di replicazione di rubivirus. Proteine di replicazione esemplari di alfavirus includono proteine del virus dell'encefalite equina venezuelana (VEEV), del virus dell'encefalite equina orientale (EEEV), del virus delle Everglades (EVEV), del virus Mucambo (MUCV), del virus della foresta Semliki (SFV), del virus Pixuna (PIXV), del virus Middleburg (MIDV), del virus Chikungunya (CHIKV), del virus O'Nyong-Nyong (ONNV), del virus Ross River (RRV), Virus Barmah Forest (BFV), Virus Getah (GETV), Virus Sagiyama (SAGV), Virus Bebaru (BEBV), Virus Mayaro (MAYV), Virus Una (UNAV), Virus Sindbis (SINV), Virus Aura (AURAV), Virus Whataroa (WHAV), Virus Babanki (BABV), Virus Kyzylgach (KYZV), Virus dell'encefalite equina occidentale (WEEV), Highland J Virus (HJV), Fort Morgan Virus (FMV), Ndumu Virus (NDUV), Salmonid Alphavirus (SAV), Buggy Creek Virus (BCRV) e qualsiasi combinazione di questi. Le proteine di replicazione del rubivirus esemplificative includono proteine del virus della rosolia.

Le proteine di replicazione virale codificate dai primi polinucleotidi delle molecole di acido nucleico qui fornite possono essere espresse come una o più poliproteine o come proteine separate o singole. In genere, le poliproteine sono proteine precursori che vengono scisse per generare proteine singole o separate. Di conseguenza, le proteine derivate da una poliproteina precursore possono essere espresse da un singolo open reading frame (ORF). Come utilizzato nel presente documento, il termine "ORF" si riferisce a una sequenza nucleotidica che inizia con un codone di inizio, generalmente ATG, e che termina con un codone di stop, come TAA, TAG o TGA, ad esempio.

Si comprenderà che T è presente nel DNA, mentre U è presente nell'RNA. Di conseguenza, un codone di inizio di ATG nel DNA corrisponde ad AUG nell'RNA e i codoni di stop TAA, TAG e TGA nel DNA corrispondono a UAA, UAG e UGA nell'RNA. Sarà inoltre apprezzato che per qualsiasi sequenza fornita nella presente divulgazione, T è presente nel DNA, mentre U è presente nell'RNA. Di conseguenza, per qualsiasi sequenza fornita nel presente documento, T presente nel DNA è sostituito con U per una molecola di RNA, e U presente nell'RNA è sostituito con T per una molecola di DNA.

La proteasi che scinde una poliproteina può essere una proteasi virale o una proteasi cellulare. In alcuni aspetti, il primo polinucleotide di molecole di acido nucleico qui fornito codifica una poliproteina comprendente una proteina alphavirus nsP1, una proteina alphavirus nsP2, una proteina alphavirus nsP3, una proteina alphavirus nsP4 o una qualsiasi combinazione delle stesse. In altri aspetti, il primo polinucleotide di molecole di acido nucleico qui fornito codifica una poliproteina comprendente una proteina alphavirus nsP1, una proteina alphavirus nsP2, una proteina alphavirus nsP3 o una qualsiasi combinazione delle stesse e una proteina alphavirus nsP4. In alcuni aspetti, la poliproteina è una poliproteina VEEV. In altri aspetti, le proteine alphavirus nsP1, nsP2, nsP3 e nsP4 sono proteine VEEV.

In un aspetto, i primi polinucleotidi di molecole di acido nucleico forniti nel presente documento sono privi di un codone di stop tra sequenze che codificano una proteina nsP3 e una proteina nsP4. Di conseguenza, in alcuni aspetti, i primi polinucleotidi di molecole di acido nucleico forniti nel presente documento codificano una poliproteina P1234 comprendente nsP1, nsP2, nsP3 e nsP4. I primi polinucleotidi di molecole di acido nucleico forniti nel presente documento possono anche includere un codone di stop tra sequenze che codificano una proteina nsP3 e una proteina nsP4. Di conseguenza, in alcuni aspetti, i primi polinucleotidi di molecole di acido nucleico forniti nel presente documento codificano una poliproteina P123 comprendente nsP1, nsP2 e nsP3 e una poliproteina P1234 comprendente nsP1, nsP2, nsP3 e nsP4 come risultato della lettura del codone di stop, ad esempio. In altri aspetti, i primi polinucleotidi di molecole di acido nucleico forniti nel presente documento codificano una poliproteina avente almeno l'80%, almeno l'85%, almeno il 90%, almeno il 91%, almeno il 92%, almeno il 93%, almeno il 94%, almeno il 95%, almeno il 96%, almeno il 97%, almeno il 97,5%, almeno il 98%, almeno il 98,5%, almeno il 99%, almeno il 99,5%, almeno il 99,6%, almeno il 99,7%, almeno il 99,8%, almeno il 99,9% e qualsiasi numero o intervallo intermedio, di identità a una sequenza di SEQ ID NO:79. In alcune realizzazioni, i primi polinucleotidi di molecole di acido nucleico forniti nel presente documento codificano una poliproteina avente una sequenza di SEQ ID NO:79. Ulteriori poliproteine esemplari comprendono una sequenza di SEQ ID NO:80 o SEQ ID NO:81. In un aspetto, le proteine nsP2 e nsP3 includono mutazioni. Le mutazioni esemplari includono le mutazioni G1309R e S1583G delle proteine VEEV. In un altro aspetto, le proteine nsP1, nsP2 e nsP4 sono proteine VEEV e la proteina nsP3 è una proteina nsP3 del virus chikungunya (CHIKV).

In alcuni aspetti, i primi polinucleotidi di molecole di acido nucleico qui forniti possono includere una prima regione intergenica. In alcuni aspetti, la prima regione intergenica è situata tra una sequenza che codifica una poliproteina comprendente una proteina alphavirus nsP1, una proteina alphavirus nsP2, una proteina alphavirus nsP3 o una qualsiasi combinazione di queste, e una sequenza che codifica una proteina alphavirus nsP4. Una prima regione intergenica può

comprendere qualsiasi sequenza, come qualsiasi sequenza virale o non virale. In un aspetto, la prima regione intergenica comprende una sequenza virale. In un altro aspetto, la prima regione intergenica comprende una sequenza alphavirus. In un altro aspetto ancora, l'alphavirus è VEEV. In un aspetto, le proteine nsP2 e nsP3 includono mutazioni. Le mutazioni esemplari includono le mutazioni G1309R e S1583G delle proteine VEEV. Sotto un altro aspetto, le proteine nsP1, nsP2 e nsP4 sono proteine VEEV e la proteina nsP3 è una proteina nsP3 del virus chikungunya (CHIKV).

In alcune realizzazioni, il primo polinucleotide può comprendere una sequenza avente almeno l'80% di identità con una sequenza di SEQ ID NO:72.

In alcune realizzazioni, la molecola di acido nucleico descritta nel presente documento può comprendere ulteriormente un secondo polinucleotide comprendente un primo transgene che codifica una prima proteina antigenica o un frammento della stessa, in cui la prima proteina antigenica è una proteina del coronavirus. In realizzazioni specifiche, la proteina antigenica può essere una proteina SARS-CoV-2. In realizzazioni specifiche, la proteina antigenica è una glicoproteina spike SARS-CoV-2. In realizzazioni specifiche, la glicoproteina spike SARS-CoV-2 è una glicoproteina spike SARS-CoV-2 di tipo selvaggio avente una sequenza amminoacidica di SEQ ID NO:123.

In alcune realizzazioni, il secondo polinucleotide comprende una sequenza avente almeno l'85% di identità con una sequenza SEQ ID NO:121 o SEQ ID NO:122.

Regione non tradotta 5' (5' UTR)

Le molecole di acido nucleico fornite nel presente documento possono inoltre comprendere regioni non tradotte (UTR). Le regioni non tradotte, tra cui 5' UTR e 3' UTR, ad esempio, possono influenzare la stabilità dell'RNA e/o l'efficienza della traduzione dell'RNA, come la traduzione di mRNA cellulari e virali, ad esempio. 5' UTR e 3' UTR possono anche influenzare la stabilità e la traduzione di RNA genomici virali e RNA autoreplicanti, tra cui RNA autoreplicanti o repliconi derivati da virus. Esempi di RNA genomici virali la cui stabilità e/o efficienza di traduzione può essere influenzata da 5' UTR e 3' UTR includono l'acido nucleico del genoma dei virus a RNA a senso positivo. Sia l'acido nucleico del genoma dei virus a RNA a senso positivo che gli RNA autoreplicanti, tra cui RNA autoreplicanti o repliconi derivati da virus, possono essere tradotti in seguito a infezione o introduzione in una cellula.

In alcuni aspetti, le molecole di acido nucleico fornite nel presente documento includono inoltre una regione 5' non tradotta (5' UTR). Qualsiasi sequenza 5' UTR può essere inclusa nelle molecole di acido nucleico fornite nel presente documento. In alcune forme di realizzazione, le molecole di acido nucleico fornite nel presente documento includono una 5' UTR virale. In un aspetto, le molecole di acido nucleico fornite nel presente documento includono una 5' UTR non virale. Qualsiasi 5' UTR non virale può essere inclusa nelle molecole di acido nucleico fornite nel presente documento, come le 5' UTR di trascrizioni espresse in qualsiasi cellula o organo, inclusi muscoli, pelle, tessuto sottocutaneo, fegato, milza, linfonodi, cellule presentanti l'antigene e altri. In un altro aspetto, le molecole di acido nucleico fornite nel presente documento includono una 5' UTR comprendente sequenze virali e non virali. Di conseguenza, una 5' UTR inclusa nelle molecole di acido nucleico fornite nel presente documento può comprendere una combinazione di sequenze 5'

UTR virali e non virali. In alcuni aspetti, il 5' UTR incluso nelle molecole di acido nucleico fornite nel presente documento si trova a monte o a 5' del primo polinucleotide che codifica una o più proteine di replicazione virale. In altri aspetti, il 5' UTR si trova a 5' o a monte del primo polinucleotide di molecole di acido nucleico fornite nel presente documento che codifica una o più proteine di replicazione virale, e il primo polinucleotide si trova a 5' o a monte del secondo polinucleotide di molecole di acido nucleico fornite nel presente documento.

In un aspetto, l'UTR 5' delle molecole di acido nucleico qui fornite comprende un UTR 5' dell'alfavirus. Un 5' UTR da qualsiasi alfavirus può essere incluso nelle molecole di acido nucleico fornite nel presente documento, comprese le sequenze 5' UTR dal virus dell'encefalite equina venezuelana (VEEV), dal virus dell'encefalite equina orientale (EEEV), dal virus delle Everglades (EVEV), dal virus Mucambo (MUCV), dal virus della foresta di Semliki (SFV), dal virus Pixuna (PIXV), dal virus Middleburg (MIDV), dal virus Chikungunya (CHIKV), dal virus O'Nyong-Nyong (ONNV), dal virus Ross River (RRV), dal virus della foresta di Barmah (BFV), dal virus Getah (GETV), dal virus Sagiyama (SAGV), dal virus Bebaru (BEBV), dal virus Mayaro (MAYV), dal virus Una (UNAV), dal virus Sindbis (SINV), dal virus Aura (AURAV), dal virus Whataroa (WHAV), dal virus Babanki (BABV), dal virus Kyzylgach (KYZV), virus dell'encefalite equina occidentale (WEEV), virus Highland J (HJV), virus Fort Morgan (FMV), virus Ndumu (NDUV), virus Alphavirus dei salmonidi (SAV) o virus Buggy Creek (BCRV). In un altro aspetto, la 5' UTR comprende una sequenza avente almeno l'80%, almeno l'85%, almeno il 90%, almeno il 91%, almeno il 92%, almeno il 93%, almeno il 94%, almeno il 95%, almeno il 96%, almeno il 97%, almeno il 97,5%, almeno il 98%, almeno il 98,5%, almeno il 99%, almeno il 99,5%, almeno il 99,6%, almeno il 99,7%, almeno il 99,8%, almeno il 99,9% e qualsiasi numero o intervallo intermedio, identità a una sequenza di SEQ ID NO:73, SEQ ID NO:74 o SEQ ID NO:75. In un altro aspetto ancora, la 5' UTR comprende una sequenza di SEQ ID NO:73, SEQ ID NO: 74 o SEQ ID NO:75.

In alcune realizzazioni, la 5' UTR comprende una sequenza selezionata dalle 5' UTR di IL-6 umana, alanina aminotransferasi 1, apolipoproteina E umana, catena alfa del fibrinogeno umano, transtiretina umana, aptoglobina umana, alfa-1-antichimotripsina umana, antitrombina umana, alfa-1-antitripsina umana, albumina umana, beta globina umana, complemento umano C3, complemento umano C5, SynK (proteina del canale del potassio tilacoide derivata dai cianobatteri, *Synechocystis* sp.), beta globina di topo, albumina di topo e un virus dell'incisione del tabacco, o frammenti di uno qualsiasi dei precedenti. Preferibilmente, la 5' UTR è derivata da un virus dell'incisione del tabacco (TEV). Preferibilmente, un mRNA descritto nel presente documento comprende una sequenza 5' UTR che è derivata da un gene espresso da *Arabidopsis thaliana*. Preferibilmente, la sequenza 5' UTR di un gene espresso da *Arabidopsis thaliana* è AT1G58420. Esempi di 5' UTR e 3' UTR sono descritti in PCT/US2018/035419, il cui contenuto è qui incorporato per riferimento. Le sequenze 5' UTR preferite comprendono SEQ ID NO: 5-10 e 25-45: come mostrato nella Tabella 1.

TABELLA 1 Sequenze UTR 5' Nome Sequenza Numero ID Seq: Veicolo elettrico
UCAACACAAACAUUAUACAAAACAAACGAAUCUCAAGCAAUC SEQ ID N.: 5
AAGCAUUCUACUUCUAUUGCAGCAAUUUAAAUCAUUUCU
UUUAAAGCAAAGCAAUUUUCUGAAAAUUUUCACCAUUU ACGAACGAUAG AT1G58420
AUUAUUACAUCAAAACAAAAGCCGCCA SEQ ID N.: 6 ARC5-2

CUUAAGGGGGCGCUGCCUACGGAGGUGGCAGCCAUCUCCU SEQ ID N.: 7
UCUCGGCAUCAAGCUUACCAUGGUGCCCCAGGCCUCUC
UUGGUCCCUGCUGGUGUCCCCUCUGCUUCGGCAAGU
UCCCAUCUACACCAUCCCCGACAAGCUGGGGCCGUGGAG
CCCAUCGACAUCCACCACUGUCCUGCCCCAACAACCUCG
UGGUCGAGGACGAGGGCUGCACCAACCUGAGCGGGUUCUC CUAC virus dell'epatite C UGAGUGUCGU
ACAGCCUCCA GGCCCCCCC SEQ ID NO: 8 UCCCGGAGA GCCAUAGUGG
UCUGCGGAACCGGUGAGUAC ACCGGAUUG CCGGGAAGAC UGGGUCCUUU CUUGGAUAAA
CCCACUCUAUGCCCGGCCAU UUGGGCUGC CCCCAGCAAGA CUGCUAGCG AGUAGUGUUG GGUUGCG
UMANO AAUUAUUGGUUAAAGAAGUAUAUAGUGCUAAUUUCCCU SEQ ID NO: 9 ALBUMINA
CCGUUUGUCCUAGCUUUUCUCUUCUGUCAACCCACACGC BUONANOTTE EMCV CUCCUCCCC
CCCCCUAAC GUUACUGGCC SEQ ID NO: 10 GAAGCCGUU GGAAUAAGGC CGGUGUGCGU
UUGUCUAAUUAU GUUAUUUCC ACCAUUUGC CGUCUUUUGG CAAUGUGAGG GCCCGGAAAC
CUGGCCUCUGU CUUCUUGACG AGCAUCCUA GGGGUCUUUC CCCUCUCGCC AAAGGAAUGC
AAGGUCUGUU GAAUGUCGUG AAGGAAGCAG UCCUCUGGA AGCUUCUUGA AGACAAACAA
CGUCUGUAGC GACCCUUGC AGGCACGGA ACCCCCCACC UGGCGACAGG UGCCUCUGCG
GCCAAAAGCC ACGUGUAUAA GAUACACUG CAAAGGCGGC ACAACCCAG UGCCACGUUG
UGAGUGUGGAU AGUUGUGGAA AGAGUCAAU GGCUCUCCUC AAGCGUAUUC AACAAAGGGC
UGAAGGAUGC CCAGAAGGUA CCCAUUGUA UGGGAUCUGA UCUGGGGCCU CGGUGCACAU
GCUUUACGUG UGUUUAGUCG AGGUUAAAAA ACGUCUAGGC CCCCCGAACC ACGGGGACGU
GGUUUCCUU UGAAAAACAC GAUGAUAAU AT1G67090 CACAAAGAGUAAAGAAGAAACA SEQ ID NO:
25 AT1G35720 AACACUAAAAGUAGAAGAAAA SEQ ID NO: 26 AT5G45900
CUCAGAAAAGAUAGAUCAGCC SEQ ID NO: 27 AT5G61250 AACCAUCGAAAAGAAACAAA SEQ ID
NO: 28 AT5G46430 CUCUAAUCACCAGGAGUAAAA SEQ ID NO: 29 AT5G47110
GAGAGAGAUUUACAAAAAA SEQ ID NO: 30 AT1G03110 UGUGUAACAACAACAACA SEQ ID NO:
31 AT3G12380 CCGCAGUAGGAAGAGAAAGCC SEQ ID NO: 32 AT5G45910 AAAAAAAAAAAGAA ... SEQ
ID NO: 33 AT1G07260 GAGAGAAGAAGAAGAAGAAGAGAGCG SEQ ID NO: 34 AT3G55500
CAUUUAAAAUACUUACAAA SEQ ID NO: 35 AT3G46230 GCAAACAGAGUAAGCGAAACG SEQ ID NO:
36 AT2G36170 GCGAAGAAGACGAACGCAAAG SEQ ID NO: 37 AT1G10660
UUAGGACUGAUUGACUGGCC SEQ ID NO: 38 AT4G14340 AUCAUCGGAUUUCGAAAAAG SEQ ID NO:
39 AT1G49310 AAAACAAAAGUUAAAGCAGAC SEQ ID NO: 40 AT4G14360
UUUAUCUCAAUAAGAAGGGCA SEQ ID NO: 41 AT1G28520 GGGGGGAGGUGAGAUUUUCUU SEQ ID
NO: 42 AT1G20160 UGAUUAGGAAACUACAAAGCC SEQ ID NO: 43 AT5G37370
CAUUUUCAAUUUCAUAAAAC SEQ ID NO: 44 AT4G11320 UUACUUUUUAAGCCCAACAAA SEQ ID
NO: 45 AT5G40850 GCGGUGUGUGUGUGUUGUUGA SEQ ID NO: 46 AT1G06150
GUGGUGAAGGGGAAGGUUAG SEQ ID NO: 47 AT2G26080 UUGUUUUUUUUUGGUUUGGUU SEQ ID
NO: 48

Regione non tradotta 3' (3' UTR)

In alcuni aspetti, le molecole di acido nucleico fornite nel presente documento includono inoltre una regione 3' non tradotta (3' UTR). Qualsiasi sequenza 3' UTR può essere inclusa nelle molecole di acido nucleico fornite nel presente documento. In un aspetto, le molecole di acido nucleico fornite

nel presente documento includono una 3' UTR virale. In un altro aspetto, le molecole di acido nucleico fornite nel presente documento includono una 3' UTR non virale. Qualsiasi 3' UTR non virale può essere inclusa nelle molecole di acido nucleico fornite nel presente documento, come le 3' UTR di trascrizioni espresse in qualsiasi cellula o organo, inclusi muscoli, pelle, tessuto sottocutaneo, fegato, milza, linfonodi, cellule presentanti l'antigene e altri. In alcuni aspetti, le molecole di acido nucleico fornite nel presente documento includono una 3' UTR comprendente sequenze virali e non virali. Di conseguenza, una 3' UTR inclusa nelle molecole di acido nucleico fornite nel presente documento può comprendere una combinazione di sequenze 3' UTR virali e non virali. In un aspetto, l'UTR 3' è situato 3' o a valle del secondo polinucleotide di molecole di acido nucleico fornito nel presente documento che comprende un primo transgene che codifica una prima proteina antigenica o un frammento della stessa. In un altro aspetto, l'UTR 3' è situato 3' o a valle del secondo polinucleotide di molecole di acido nucleico fornito nel presente documento che comprende un primo transgene che codifica una prima proteina antigenica o un frammento della stessa, e il secondo polinucleotide è situato 3' o a valle del primo polinucleotide di molecole di acido nucleico fornito nel presente documento.

In un aspetto, l'UTR 3' delle molecole di acido nucleico qui fornite comprende un UTR 3' dell'alfavirus. Un 3' UTR da qualsiasi alfavirus può essere incluso nelle molecole di acido nucleico fornite nel presente documento, comprese le sequenze 3' UTR dal virus dell'encefalite equina venezuelana (VEEV), dal virus dell'encefalite equina orientale (EEEV), dal virus delle Everglades (EVEV), dal virus Mucambo (MUCV), dal virus della foresta di Semliki (SFV), dal virus Pixuna (PIXV), dal virus Middleburg (MIDV), dal virus Chikungunya (CHIKV), dal virus O'Nyong-Nyong (ONNV), dal virus Ross River (RRV), dal virus della foresta di Barmah (BFV), dal virus Getah (GETV), dal virus Sagiya (SAGV), dal virus Bebaru (BEBV), dal virus Mayaro (MAYV), dal virus Una (UNAV), dal virus Sindbis (SINV), dal virus Aura (AURAV), dal virus Whataroa (WHAV), dal virus Babanki (BABV), dal virus Kyzylgach (KYZV), Western Equine Encephalitis Virus (WEEV), Highland J Virus (HJV), Fort Morgan Virus (FMV), Ndumu Virus (NDUV), Salmonid Alphavirus (SAV) o Buggy Creek Virus (BCRV). In un altro aspetto, l'UTR 3' comprende una sequenza avente almeno l'80%, almeno l'85%, almeno il 90%, almeno il 91%, almeno il 92%, almeno il 93%, almeno il 94%, almeno il 95%, almeno il 96%, almeno il 97%, almeno il 97,5%, almeno il 98%, almeno il 98,5%, almeno il 99%, almeno il 99,5%, almeno il 99,6%, almeno il 99,7%, almeno il 99,8%, almeno il 99,9% e qualsiasi numero o intervallo intermedio, identità a una sequenza di SEQ ID NO:76. In un altro aspetto ancora, l'UTR 3' comprende una sequenza poli-A. In un ulteriore aspetto, l'UTR 3' comprende una sequenza di SEQ ID NO:76.

In alcune realizzazioni, la 3' UTR comprende una sequenza selezionata dalle 3' UTR di alanina aminotransferasi 1, apolipoproteina E umana, catena alfa del fibrinogeno umano, aptoglobina umana, antitrombina umana, alfa globina umana, beta globina umana, complemento umano C3, fattore di crescita umano, epcidina umana, MALAT-1, beta globina di topo, albumina di topo e beta globina di Xenopus, o frammenti di uno qualsiasi dei precedenti. In alcune realizzazioni, la 3' UTR è derivata dalla beta globina di Xenopus. Sequenze esemplari di 3' UTR includono SEQ ID NO: 16-22 come mostrato nella Tabella 2.

TABELLA 2 Sequenze 3' UTR. Nome Sequenza Numero ID Seq: XBG
CUAGUGACUGACUAGGGAUCUGGUUACCACUAAACCAG SEQ ID NO: 16

CCUCAAGAACACCCGAAUUGGAGUCUCUAAGCUACAUA
 AUACCAACUACACUUACAAAUGUUGUCCCCCAAAA
 UGUAGCCAUUCGUAUCUGCUCCUAAUAAAAAGAAAGU UUCUUCACAU UMANO
 UGCAAGGCUGGCCGGAAGCCCUUGCCUGAAAGCAAGA SEQ ID NO: 17 APTOGLOBINA
 UUUCAGCCUGGAAGAGGGCAAAGUGGACGGGAGUGG
 ACAGGAGUGGAUGCGAUAGAUGUGGUUUGAAGCUG
 AUGGGUGCCAGCCUGCAUUGCUGAGUCAAUCAAUA AGAGCUUUCUUUUGACCCAU UMANO
 ACGCCGAAGCCUGCAGCCAUGCGGACCCACGCCACCCC SEQ ID NO: 18 APOLIPO-
 GUGCCUCCUGCCUCCGCGCAGCCUGCAGCGGGAGACC PROTEINA
 CUGUCCCCGCCAGCCGUCUCCUGGGGUGGACCCU E AGUUUAAUAAAGAUUCACCAAGUUUCACGCA
 virus dell'epatite C UAGAGCGGCAAACCCUAGCUACACUCCAUGCUAGUU SEQ ID NO: 19
 UCUUUUUUUUUUGUUUUUUUUUUUUUUUUUUUUUUUU
 UUUUUUUUUUUUUUUUCCUUUCUUUUCCUUUCUUUU
 UUCUUCUUUUCUUGGUGGCUCUUAAGCCUAGUC
 ACGGCUAGCUGUGAAAGGUCCGUGAGCCGCAUGACUG
 CAGAGAGUGCCGUAACUGGUCUCUCUGCAGAUCAUGU TOPO
 ACACAUCACAACCACAACCUUCUCAGGCUACCCUGAG SEQ ID NO: 20 ALBUMINA
 AAAAAAGACAUGAAGACUCAGGACUCAUCUUUUCUG
 UUGGUGUAAAAUCAACACCCUAAGGAACACAAAUUC
 UUUAAACAUUUGACUUCUUGUCUCUGUGCUGCAAUA AUAAAAAUGGAAAGAAUCUAC UMANO
 ALFA GCUGGAGCCUCGGUAGCCGUUCCUCCUGCCCGGGG SEQ ID NO: 21 GLOBINA
 CCUCCAACGGGCCUCCUCCUCCUUGCACCGGCC
 UUCCUGGUCUUUGAAUAAAGUCUGAGUGGGCAGCA EMCV UAGUGCAGUCAC UGGCACAACG
 CGUUGCCCGG SEQ ID NO: 22 AAAGCCAAUC GGUUAUACAC GGUCGUCAUACUGCAGACAG
 GGUUCUUCUA CUUUGCAAGA UAGUCUAGAG UAGUAAAAUA AAA ...

Codone di stop triplo

In alcune realizzazioni, l'RNA autoreplicante può comprendere una sequenza immediatamente a valle di una regione codificante (ad esempio, ORF) che crea un triplo codone di stop. Un triplo codone di stop è una sequenza di tre codoni di stop consecutivi. Il triplo codone di stop può garantire l'isolamento totale di una cassetta di espressione e può essere incorporato per migliorare l'efficienza della traduzione. In alcune realizzazioni, un RNA autoreplicante della divulgazione può comprendere una tripla combinazione di una qualsiasi delle sequenze UAG, UGA o UAA immediatamente a valle di un ORF descritto nel presente documento. La tripla combinazione può essere tre degli stessi codoni, tre codoni diversi o qualsiasi altra permutazione dei tre codoni di stop.

Potenziatori di traduzione e sequenze Kozak

Per l'inizio della traduzione, devono essere stabilite interazioni appropriate tra ribosomi e mRNA per determinare la posizione esatta della regione di inizio della traduzione. Tuttavia, i ribosomi devono anche dissociarsi dalla regione di inizio della traduzione per scivolare verso la sequenza a

valle durante la traduzione dell'mRNA. Gli enhancer di traduzione a monte delle sequenze di inizio degli mRNA migliorano le rese della biosintesi proteica. Diversi studi hanno esaminato gli effetti degli enhancer di traduzione. In alcune forme di realizzazione, un mRNA descritto nel presente documento comprende una sequenza enhancer di traduzione. Queste sequenze enhancer di traduzione migliorano l'efficienza della traduzione di un RNA autoreplicante della divulgazione e quindi forniscono una maggiore produzione della proteina codificata dall'mRNA. La regione enhancer di traduzione può essere situata nell'UTR 5' o 3' di una sequenza di mRNA. Esempi di regioni enhancer di traduzione includono regioni enhancer naturali dall'UTR 5' di TEV e dall'UTR 3' della beta-globina di *Xenopus*. Le sequenze di potenziamento 5' UTR esemplari includono, ma non sono limitate a, quelle derivate da mRNA che codificano proteine da shock termico (HSP) umane, tra cui HSP70-P2, HSP70-M1 HSP72-M2, HSP17.9 e HSP70-P1. Le sequenze di potenziamento della traduzione preferite utilizzate in conformità con le forme di realizzazione della presente informativa sono rappresentate da SEQ ID No: 11-15 come mostrato nella Tabella 3.

TABELLA 3 Potenziatori UTR 5' ID sequenza Nome Sequenza NO.: HSP70-
 GUCAGCUUCAAACUCUUUGUUUCUUGUUU ID SEQ P2 GUUGAUUGAGAAUA Numero: 11 HSP70-
 CUCUCGCCUGAGAAAAAAAUCCACGAACC ID SEQ La mia 1 AAUUCUCAGCAACCAGCACG Numero:
 12 HSP72- ACCUGUGAGGGGUUCGAAGGAAGUAGCAGUG ID SEQ La M2
 UUUUUUGUUCUAGAGGAAGAG Numero: 13 HSP17.9 ACACAGAAACAUUCGCAAAAACAAAUCCC
 ID SEQ AGUAUCAAAAUUCUUCUUCUUUUUUCAUAU Numero: 14 UUCGCAAAGAC HSP70-
 CAGAAAAUUUGCUACAUGUUUCACAAAC ID SEQ La prima UUCAAAUAUUAUUCAUU ... Numero:
 15

In alcune realizzazioni, un RNA autoreplicante della divulgazione comprende una sequenza Kozak. Come è noto nell'arte, una sequenza Kozak è una breve sequenza di consenso incentrata sul sito di inizio della traduzione degli mRNA eucariotici che consente un inizio efficiente della traduzione dell'mRNA. Vedere, ad esempio, Kozak, Marilyn (1988) *Mol. e Cell Biol*, 8:2737-2744; Kozak, Marilyn (1991) *J. Biol. Chem*, 266: 19867-19870; Kozak, Marilyn (1990) *Proc Natl. Acad. Sci. USA*, 87:8301-8305; e Kozak, Marilyn (1989) *J. Cell Biol*, 108:229-241. Garantisce che una proteina venga tradotta correttamente dal messaggio genetico, mediando l'assemblaggio del ribosoma e l'inizio della traduzione. Il meccanismo di traduzione ribosomiale riconosce il codone di inizio AUG nel contesto della sequenza Kozak. Una sequenza Kozak può essere inserita a monte della sequenza codificante per la proteina di interesse, a valle di un 5' UTR o inserita a monte della sequenza codificante per la proteina di interesse e a valle di un 5' UTR. In alcune realizzazioni, un RNA autoreplicante descritto nel presente documento comprende una sequenza Kozak avente la sequenza amminoacidica GCCACC (SEQ ID NO: 23). Preferibilmente un RNA autoreplicante descritto nel presente documento comprende una sequenza Kozak parziale "p" avente la sequenza amminoacidica GCCA (SEQ ID NO: 24).

Transgeni

I transgeni inclusi nelle molecole di acido nucleico fornite nel presente documento possono codificare una proteina antigenica o un frammento della stessa. In alcune realizzazioni, i secondi polinucleotidi di molecole di acido nucleico forniti nel presente documento comprendono un

primo transgene. Un primo transgene incluso nei secondi polinucleotidi di molecole di acido nucleico forniti nel presente documento può codificare una prima proteina antigenica o un frammento della stessa. Un transgene incluso nei secondi polinucleotidi di molecole di acido nucleico forniti nel presente documento può comprendere una sequenza che codifica la sequenza amminoacidica completa di una proteina antigenica o una sequenza che codifica qualsiasi porzione o frammento idoneo della sequenza amminoacidica completa di una proteina antigenica. In alcune realizzazioni, la proteina antigenica è una proteina di coronavirus.

In un'altra forma di realizzazione, la proteina antigenica, quando somministrata a un soggetto mammifero, provoca una risposta immunitaria a un patogeno, come un coronavirus. In alcune forme di realizzazione più particolari, la proteina antigenica è espressa sulla superficie esterna del coronavirus; mentre in altre forme di realizzazione più particolari, l'antigene può essere un antigene non di superficie, ad esempio, utile come epitopo delle cellule T. L'immunogeno può provocare una risposta immunitaria contro un coronavirus. La risposta immunitaria può comprendere una risposta anticorpale (solitamente includente IgG) e/o una risposta immunitaria mediata dalle cellule. L'immunogeno polipeptidico in genere provocherà una risposta immunitaria che riconosce il coronavirus corrispondente. L'immunogeno sarà in genere un polipeptide di superficie, ad esempio una glicoproteina dell'involucro, una glicoproteina spike, ecc.

In alcuni aspetti, la proteina virale codificata dai transgeni inclusi nelle molecole di acido nucleico fornite nel presente documento è una proteina del coronavirus. In alcune realizzazioni, la proteina antigenica è una proteina SARS-CoV-2.

In un aspetto, la proteina antigenica è una glicoproteina spike di SARS-CoV-2 o un frammento di essa. In un altro aspetto, la glicoproteina spike di SARS-CoV-2 è una glicoproteina spike di SARS-CoV-2 di tipo selvaggio. In alcuni aspetti, la glicoproteina spike di SARS-CoV-2 di tipo selvaggio ha una sequenza amminoacidica di SEQ ID NO:123. In un altro aspetto ancora, il secondo polinucleotide di molecole di acido nucleico fornito nel presente documento comprende una sequenza avente almeno l'85%, almeno il 90%, almeno il 91%, almeno il 92%, almeno il 93%, almeno il 94%, almeno il 95%, almeno il 96%, almeno il 97%, almeno il 97,5%, almeno il 98%, almeno il 98,5%, almeno il 99%, almeno il 99,5%, almeno il 99,6%, almeno il 99,7%, almeno il 99,8%, almeno il 99,9% e qualsiasi numero o intervallo intermedio, di identità con una sequenza di SEQ ID NO:121 o SEQ ID NO:122. In un altro aspetto, il secondo polinucleotide di molecole di acido nucleico fornito nel presente documento comprende una sequenza di SEQ ID NO:121 o SEQ ID NO:122. Di conseguenza, in alcuni aspetti, i primi transgeni inclusi nei secondi polinucleotidi di molecole di acido nucleico qui forniti comprendono una sequenza avente almeno l'85%, almeno il 90%, almeno il 91%, almeno il 92%, almeno il 93%, almeno il 94%, almeno il 95%, almeno il 96%, almeno il 97%, almeno il 97,5%, almeno il 98%, almeno il 98,5%, almeno il 99%, almeno il 99,5%, almeno il 99,6%, almeno il 99,7%, almeno il 99,8%, almeno il 99,9% e qualsiasi numero o intervallo intermedio, o il 100% di identità con una sequenza di SEQ ID NO:121 o SEQ ID NO:122.

In un aspetto, il secondo polinucleotide di molecole di acido nucleico fornito nel presente documento codifica una glicoproteina spike di SARS-CoV-2 di tipo selvaggio o un frammento della stessa. In alcuni aspetti, una glicoproteina spike di SARS-CoV-2 di tipo selvaggio comprende una sequenza di SEQ ID NO:123. In un altro aspetto, il secondo polinucleotide di molecole di acido

nucleico fornito nel presente documento codifica una proteina spike di SARS-CoV-2 comprendente una o più mutazioni rispetto a una sequenza di glicoproteina spike di SARS-CoV-2 di tipo selvaggio. Le mutazioni possono includere sostituzioni, delezioni, inserzioni e altro. Le mutazioni possono essere presenti in qualsiasi posizione o in qualsiasi combinazione di posizioni di una glicoproteina spike di SARS-CoV-2. Un numero qualsiasi di sostituzioni, inserzioni, delezioni o combinazioni delle stesse può essere presente in una o più posizioni di una glicoproteina spike di SARS-CoV-2. Ad esempio, le sostituzioni possono includere un cambiamento di un amminoacido di tipo selvatico in qualsiasi posizione o in qualsiasi combinazione di posizioni in qualsiasi altro amminoacido o combinazione di qualsiasi altro amminoacido. Le mutazioni esemplari includono mutazioni nelle posizioni 614, 936, 320, 477, 986, 987 o qualsiasi combinazione di esse. In un aspetto, una glicoproteina spike SARS-CoV-2 o un frammento di essa codificato da transgeni di secondi polinucleotidi inclusi nelle molecole di acido nucleico fornite nel presente documento include una mutazione D614G, una mutazione D936Y, una mutazione D936H, una mutazione V320G, una mutazione S477N, una mutazione S477I, una mutazione S477T, una mutazione K986P, una mutazione V987P o qualsiasi combinazione di esse. Ulteriori mutazioni e varianti possono essere trovate nel database informativo del National Bioinformatics Center 2019 Novel Coronavirus (2019nCoV), National Genomics Data Center, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Science all'indirizzo bigd.big.ac.cn/ncov/variation/annotation. In un altro aspetto, il secondo polinucleotide include un transgene che codifica una glicoproteina SARS-CoV-2 avente almeno l'85%, almeno il 90%, almeno il 91%, almeno il 92%, almeno il 93%, almeno il 94%, almeno il 95%, almeno il 96%, almeno il 97%, almeno il 97,5%, almeno il 98%, almeno il 98,5%, almeno il 99%, almeno il 99,5%, almeno il 99,6%, almeno il 99,7%, almeno il 99,8%, almeno il 99,9% e qualsiasi numero o intervallo intermedio, o il 100% di identità con una sequenza di SEQ ID NO:123.

In alcuni aspetti, il secondo polinucleotide di molecole di acido nucleico fornito nel presente documento comprende almeno due transgeni, come una seconda proteina coronavirus. Qualsiasi numero di transgeni può essere incluso nei secondi polinucleotidi di molecole di acido nucleico forniti nel presente documento, come uno, due, tre, quattro, cinque, sei, sette, otto, nove, dieci o più transgeni. In un aspetto, il secondo polinucleotide di molecole di acido nucleico fornito nel presente documento include un secondo transgene che codifica una seconda proteina antigenica o un frammento della stessa o una proteina immunomodulatrice. In un aspetto, il secondo polinucleotide comprende inoltre un sito di ingresso ribosomiale interno (IRES), una sequenza che codifica un peptide 2A o una combinazione di essi, situati tra i transgeni. Come utilizzato nel presente documento, il termine "peptide 2A" si riferisce a una piccola sequenza (generalmente 18-22 amminoacidi) che consente una produzione stechiometrica efficiente di prodotti proteici discreti all'interno di un singolo frame di lettura tramite un evento di salto ribosomiale all'interno della sequenza peptidica 2A. Come utilizzato nel presente documento, il termine "sito di ingresso ribosomiale interno" o "IRES" si riferisce a una sequenza nucleotidica che consente l'inizio della traduzione proteica di una sequenza di RNA messaggero (mRNA) in assenza di un codone di inizio AUG o senza utilizzare un codone di inizio AUG. Un IRES può essere trovato ovunque in una sequenza di mRNA, come all'inizio o vicino all'inizio, al centro o vicino alla fine o alla fine della sequenza di mRNA, ad esempio.

Qualsiasi numero di transgeni inclusi nei secondi polinucleotidi di molecole di acido nucleico qui forniti può essere espresso tramite qualsiasi combinazione di sequenze peptidiche 2A e IRES. Ad esempio, un secondo transgene situato a 3' di un primo transgene può essere espresso tramite una sequenza peptidica 2A o tramite una sequenza IRES. Come altro esempio, un secondo transgene situato a 3' di un primo transgene e un terzo transgene situato a 3' del secondo transgene possono essere espressi tramite sequenze peptidiche 2A situate tra il primo e il secondo transgene e il secondo e il terzo transgene, tramite una sequenza IRES situata tra il primo e il secondo transgene e il secondo e il terzo transgene, tramite una sequenza peptidica 2A situata tra il primo e il secondo transgene e una IRES situata tra il secondo e il terzo transgene, o tramite una sequenza IRES situata tra il primo e il secondo transgene e una sequenza peptidica 2A situata tra il secondo e il terzo transgene. Configurazioni e combinazioni simili di sequenze peptidiche 2A e IRES situate tra transgeni sono contemplate per qualsiasi numero di transgeni inclusi nei secondi polinucleotidi di molecole di acido nucleico qui forniti. Oltre all'espressione tramite sequenze peptidiche 2A e IRES, due o più transgeni inclusi nelle molecole di acido nucleico qui forniti possono anche essere espressi da RNA subgenomici separati.

Un secondo, terzo, quarto, quinto, sesto, settimo, ottavo, nono, decimo, ecc., transgene incluso nei secondi polinucleotidi di molecole di acido nucleico qui forniti può codificare una proteina immunomodulatrice o un frammento funzionale o una variante funzionale della stessa. Qualsiasi proteina immunomodulatrice o un frammento funzionale o una variante funzionale della stessa può essere codificata da un transgene incluso nei secondi polinucleotidi.

Come qui utilizzato, i termini "variante funzionale" o "frammento funzionale" si riferiscono a una molecola, tra cui un acido nucleico o una proteina, ad esempio, che comprende una sequenza di nucleotidi e/o amminoacidi che è alterata da uno o più nucleotidi e/o amminoacidi rispetto alle sequenze di nucleotidi e/o amminoacidi della molecola madre o di riferimento. Per una proteina, una variante funzionale è ancora in grado di funzionare in un modo che è simile alla molecola madre. In altre parole, le modifiche nella sequenza di amminoacidi e/o nucleotidi della molecola madre non influenzano o alterano significativamente le caratteristiche funzionali della molecola codificata dalla sequenza nucleotidica o contenente la sequenza di amminoacidi. La variante funzionale può avere modifiche conservative della sequenza tra cui sostituzioni, aggiunte e delezioni di nucleotidi e amminoacidi. Queste modifiche possono essere introdotte da tecniche standard note nell'arte, come la mutagenesi sito-diretta e la mutagenesi mediata da PCR casuale. Le varianti funzionali possono anche includere, ma non sono limitate a, derivati che sono sostanzialmente simili nella sequenza strutturale primaria, ma che contengono, ad esempio, modifiche in vitro o in vivo, chimiche e/o biochimiche, che non si trovano nella molecola madre. Tali modifiche includono, tra le altre, acetilazione, acilazione, ADP-ribosilazione, amidazione, legame covalente di flavina, legame covalente di una frazione eme, legame covalente di un nucleotide o derivato nucleotidico, legame covalente di un lipide o derivato lipidico, legame covalente di fosfatidilinositolo, reticolazione, ciclizzazione, formazione di legami disolfuro, demetilazione, formazione di legami crociati covalenti, formazione di cisteina, formazione di piroglutammato, formilazione, gamma-carbossilazione, glicosilazione, formazione di ancoraggio GPI, idrossilazione, iodinazione, metilazione, miristoilazione, ossidazione, pegilazione, elaborazione proteolitica, fosforilazione, prenilazione, racemizzazione, selenoilazione,

solfatazione, aggiunta di amminoacidi alle proteine mediata da RNA di trasferimento come arginilazione, ubiquitinazione e simili.

In un aspetto, un secondo transgene incluso nei secondi polinucleotidi di molecole di acido nucleico qui forniti codifica una citochina, una chemioquina o un'interleuchina. Le citochine esemplari includono interferoni, TNF- α , TGF- β , G-CSF e GM-CSF. Le chemiochine esemplari includono CCL3, CCL26 e CXCL7. Le interleuchine esemplari includono IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12, IL-15, IL-18, IL-21 e IL-23. Qualsiasi transgene o combinazione di transgeni che codificano qualsiasi citochina, chemioquina, interleuchina o combinazioni di queste, può essere incluso nei secondi polinucleotidi di molecole di acido nucleico qui forniti.

In alcune realizzazioni, il secondo transgene codifica una seconda proteina del coronavirus.

Molecole di DNA e RNA

Le molecole di acido nucleico qui fornite possono essere molecole di DNA o molecole di RNA. Si comprenderà che T presente nel DNA è sostituito con U nell'RNA e viceversa. In un aspetto, le molecole di acido nucleico qui fornite sono molecole di DNA. In un altro aspetto, le molecole di DNA qui fornite comprendono inoltre un promotore. Come utilizzato nel presente documento, il termine "promotore" si riferisce a una sequenza regolatrice che avvia la trascrizione. Un promotore può essere collegato operativamente al primo e al secondo polinucleotide delle molecole di acido nucleico qui fornite. In generale, i promotori inclusi nelle molecole di DNA qui fornite includono promotori per la trascrizione in vitro (IVT). Qualsiasi promotore adatto per la trascrizione in vitro può essere incluso nelle molecole di DNA qui fornite, come un promotore T7, un promotore T3, un promotore SP6 e altri. In un aspetto, le molecole di DNA qui fornite comprendono un promotore T7. In un altro aspetto, il promotore è situato a 5' del 5' UTR incluso nelle molecole di DNA fornite nel presente documento. In un altro aspetto ancora, il promotore è un promotore T7 situato a 5' del 5' UTR incluso nelle molecole di DNA fornite nel presente documento. In un altro aspetto ancora, il promotore si sovrappone al 5' UTR. Un promotore e un 5' UTR possono sovrapporsi di circa un nucleotide, circa due nucleotidi, circa tre nucleotidi, circa quattro nucleotidi, circa cinque nucleotidi, circa sei nucleotidi, circa sette nucleotidi, circa otto nucleotidi, circa nove nucleotidi, circa dieci nucleotidi, circa 11 nucleotidi, circa 12 nucleotidi, circa 13 nucleotidi, circa 14 nucleotidi, circa 15 nucleotidi, circa 16 nucleotidi, circa 17 nucleotidi, circa 18 nucleotidi, circa 19 nucleotidi, circa 20 nucleotidi, circa 21 nucleotidi, circa 22 nucleotidi, circa 23 nucleotidi, circa 24 nucleotidi, circa 25 nucleotidi, circa 26 nucleotidi, circa 27 nucleotidi, circa 28 nucleotidi, circa 29 nucleotidi, circa 30 nucleotidi, circa 31 nucleotidi, circa 32 nucleotidi, circa 33 nucleotidi, circa 34 nucleotidi, circa 35 nucleotidi, circa 36 nucleotidi, circa 37 nucleotidi, circa 38 nucleotidi, circa 39 nucleotidi, circa 40 nucleotidi, circa 41 nucleotidi, circa 42 nucleotidi, circa 43 nucleotidi, circa 44 nucleotidi, circa 45 nucleotidi, circa 46 nucleotidi, circa 47 nucleotidi, circa 48 nucleotidi, circa 49 nucleotidi, circa 50 nucleotidi o più nucleotidi.

In alcuni aspetti, le molecole di DNA fornite nel presente documento includono un promotore per la trascrizione in vivo. In genere, il promotore per la trascrizione in vivo è un promotore della RNA polimerasi II (RNA pol II). Qualsiasi promotore della RNA pol II può essere incluso nelle molecole di DNA fornite nel presente documento, inclusi promotori costitutivi, promotori inducibili e

promotori tessuto-specifici. I promotori costitutivi esemplari includono un promotore del citomegalovirus (CMV), un promotore EF1 α , un promotore SV40, un promotore PGK1, un promotore Ubc, un promotore della beta-actina umana, un promotore CAG e altri. Qualsiasi promotore tessuto-specifico può essere incluso nelle molecole di DNA fornite nel presente documento. In un aspetto, il promotore RNA pol II è un promotore specifico del muscolo, un promotore specifico della pelle, un promotore specifico del tessuto sottocutaneo, un promotore specifico del fegato, un promotore specifico della milza, un promotore specifico del linfonodo o un promotore con qualsiasi altra specificità tissutale. Le molecole di DNA fornite nel presente documento possono anche includere un potenziatore. Qualsiasi potenziatore che aumenti la trascrizione può essere incluso nelle molecole di DNA fornite nel presente documento.

In alcuni aspetti, le molecole di acido nucleico fornite nel presente documento sono molecole di RNA. Una molecola di RNA fornita nel presente documento può essere generata tramite trascrizione in vitro (IVT) di molecole di DNA fornite nel presente documento. In un aspetto, le molecole di RNA fornite nel presente documento sono molecole di RNA autoreplicanti. In un altro aspetto, le molecole di RNA fornite nel presente documento comprendono inoltre un cappuccio 5'. Qualsiasi cappuccio 5' può essere incluso nelle molecole di RNA fornite nel presente documento, inclusi i cappucci 5' aventi una struttura Cap 1, una struttura Cap 1 (m6A), una struttura Cap 2, una struttura Cap 0 o qualsiasi combinazione di queste. In un aspetto, le molecole di RNA fornite nel presente documento includono un cappuccio 5' avente una struttura Cap 1. In un altro aspetto ancora, le molecole di RNA fornite nel presente documento sono molecole di RNA autoreplicanti comprendenti un cappuccio 5' avente una struttura Cap 1. In un ulteriore aspetto, le molecole di RNA fornite nel presente documento comprendono un cappuccio avente una struttura Cap 1, in cui un m7G è collegato tramite un trifosfato 5'-5' all'estremità 5' dell'UTR 5'. In un ulteriore aspetto ancora, le molecole di RNA fornite nel presente documento comprendono un cappuccio avente una struttura Cap 1, in cui un m7G è collegato tramite un trifosfato 5'-5' all'estremità 5' dell'UTR 5' comprendente una sequenza di SEQ ID NO:73. È possibile utilizzare qualsiasi metodo di capping, incluso, ma non limitato all'utilizzo di un enzima Vaccinia Capping (New England Biolabs, Ipswich, Mass.) e capping co-trascrizionale o capping all'inizio o subito dopo l'inizio della trascrizione in vitro (IVT), ad esempio, includendo un agente di capping come parte di una reazione di trascrizione in vitro (IVT). (Nuc. Acids Symp. (2009) 53:129).

In alcune forme di realizzazione, sono fornite nel presente documento molecole di acido nucleico comprendenti (a) una sequenza di SEQ ID NO:10; (b) una sequenza di SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:76 e SEQ ID NO:77, in cui T è sostituito con U; (c) una sequenza di SEQ ID NO: 124; (d) una sequenza di SEQ ID NO:124, in cui T è sostituito con U; (e) una sequenza di SEQ ID NO:125; o (f) una sequenza di SEQ ID NO:125, in cui T è sostituito con U. In un aspetto, le molecole di acido nucleico fornite nel presente documento sono molecole di RNA. In un altro aspetto, le molecole di RNA fornite nel presente documento comprendono inoltre un cappuccio 5' avente una struttura Cap 1. Qualsiasi molecola di RNA fornita nel presente documento può essere una molecola di RNA autoreplicante.

Solo quegli mRNA che portano la struttura Cap sono attivi nella traduzione dipendente da Cap; la "decapitazione" degli mRNA determina una perdita quasi completa della loro attività modello per

la sintesi proteica (Nature, 255:33-37, (1975); J. Biol. Chem., vol. 253:5228-5231, (1978); e Proc. Natl. Acad. Sci. USA, 72:1189-1193, (1975)).

Un altro elemento dell'mRNA eucariotico è la presenza di residui di nucleoside 2'-O-metilico nella posizione di trascrizione 1 (Cap 1) e, in alcuni casi, nelle posizioni di trascrizione 1 e 2 (Cap 2). La 2'-O-metilazione dell'mRNA fornisce una maggiore efficacia della traduzione dell'mRNA in vivo (Proc. Natl. Acad. Sci. USA, 77:3952-3956 (1980)) e migliora ulteriormente la stabilità della nucleasi dell'mRNA con cappuccio 5'. L'mRNA con Cap 1 (e Cap 2) è un segno distintivo che consente alle cellule di riconoscere l'estremità 5' dell'mRNA autentico e, in alcuni casi, di discriminare le trascrizioni provenienti da elementi genetici infettivi (Nucleic Acid Research 43: 482-492 (2015)).

Alcuni esempi di strutture di cap 5' e metodi per preparare mRNA che le comprendono sono forniti in WO2015/051169A2, WO/2015/061491, US 2018/0273576 e US Pat. Nos. 8,093,367, 8,304,529 e 10,487,105. In alcune realizzazioni, il cap 5' è m7GpppAmpG, che è noto nell'arte. In alcune realizzazioni, il cap 5' è m7GpppG o m7GpppGm, che sono noti nell'arte. Le formule strutturali per le realizzazioni di strutture di cap 5' sono fornite di seguito.

In alcune realizzazioni, un RNA autoreplicante della presente informativa comprende un cappuccio 5' avente la struttura di Formula (Cap I).

in cui B^1 è una nucleobase naturale o modificata; R^1 e R^2 sono ciascuno indipendentemente selezionati da un alogeno, OH e OCH_3 ; ciascuna L è indipendentemente selezionata dal gruppo costituito da fosfato, fosforotioato e boranofosfato in cui ciascuna L è legata da legami diesterici; n è 0 o 1. e mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'. In alcune realizzazioni B^1 è G, m G^1 o A. In alcune realizzazioni, n è 0. In alcune realizzazioni n è 1. In alcune realizzazioni, B^1 è A o m A e R^1 è OCH_3 ; in cui G è guanina, m G è 7-metilguanina, A è adenina e m A è N⁶-metiladenina.

In alcune realizzazioni, un RNA autoreplicante della divulgazione comprende un cappuccio 5' avente la struttura di Formula (Cap II).

in cui B^1 e B^2 sono ciascuno indipendentemente una nucleobase naturale o modificata; R^1 , R^2 , e R^3 sono ciascuno indipendentemente selezionati da un alogeno, OH, e OCH_3 ; ogni L è indipendentemente selezionato dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legata da legami diesterici; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni B^1 è G, m G^1 o A. In alcune realizzazioni, n è 0. In alcune realizzazioni, n è 1. In alcune realizzazioni, B^1 è A o m A e R^1 è OCH_3 ; in cui G è guanina, m G è 7-metilguanina, A è adenina, e m A è N⁶-metiladenina.

In alcune realizzazioni, un RNA autoreplicante della divulgazione comprende un cappuccio 5' avente la struttura di Formula (Cap III).

in cui B1, B2 e B3 sono ciascuno indipendentemente una nucleobase naturale o modificata; R1, R2,

R3 e R4 sono ciascuno indipendentemente selezionati da un alogeno, OH e OCH₃; ciascuna L è indipendentemente selezionata dal gruppo costituito da fosfato, fosforotioato e boranofosfato in cui ciascuna L è legata da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni, almeno uno tra R1, R2, R3 e R4 è OH. In alcune realizzazioni B1 è G, m7G o A. In alcune realizzazioni, B1 è A o m6A e R1 è OCH₃; in cui G è guanina, m7G è 7-metilguanina, A è adenina e m6A è N6-metiladenina. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della presente invenzione comprende un analogo del cap 5' m7GpppG avente la struttura di Formula (Cap IV).

in cui, R¹, R², e R³ sono ciascuno indipendentemente selezionati da un alogeno, OH, e OCH₃, ogni L è indipendentemente selezionato dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legata da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; n è 0 o 1. In alcune realizzazioni, almeno uno di R¹, R², e R³ è OH. In alcune realizzazioni, il tappo 5' è m⁷GpppG in cui R¹, R², e R³ sono ciascuno OH, n è 1, e ogni L è un fosfato. In alcune realizzazioni, n è 1. In alcune realizzazioni, il cappuccio 5' è m7GpppGm, in cui R¹ e R² sono ciascuno OH, R³ è OCH₃, ciascuna L è un fosfato, mRNA è l'mRNA che codifica un enzima avente attività OTC legata alla sua estremità 5', e n è 1.

In alcune realizzazioni, un RNA autoreplicante della presente invenzione comprende un analogo del cap 5' m7Gpppm7G avente la struttura di Formula (Cap V).

in cui, R¹, R², e R³ sono ciascuno indipendentemente selezionato da un alogeno, OH, e OCH₃; ogni L è indipendentemente selezionato dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legata da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni, almeno uno di R¹, R², e R³ è OH. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della presente invenzione comprende un analogo m7Gpppm7GpN, 5' cap, in cui N è un nucleotide naturale o modificato, l'analogo 5' cap avente la struttura di Formula (Cap VI).

in cui B³ è una nucleobase naturale o modificata; R¹, R², R³, e R⁴ sono ciascuno selezionato indipendentemente da un alogeno, OH, e OCH₃; ogni L è selezionato indipendentemente dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legata da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 3. In alcune realizzazioni, almeno uno tra R¹, R², R³, e R⁴ è OH. In alcune realizzazioni B³ è G, m⁶G, o A. In alcune realizzazioni, B³ è A o m⁶A e R⁴ è OCH₃; in cui G è guanina, m⁶G è 7-metilguanina, A è adenina, e m⁶A è N⁶-metiladenina. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della presente invenzione comprende un analogo del cap 5' m7Gpppm7GpG avente la struttura di Formula (Cap VII).

in cui, R^1 , R^2 , R^3 , e R^4 sono ciascuno selezionato indipendentemente da un alogeno, OH, e OCH_3 ; ogni L è selezionato indipendentemente dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legata da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni, almeno uno di R^1 , R^2 , R^3 , e R^4 è OH. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della presente invenzione comprende un analogo del cap 5' m7Gpppm7Gpm7G avente la struttura di Formula (Cap VIII).

in cui, R^1 , R^2 , R^3 , e R^4 sono ciascuno selezionato indipendentemente da un alogeno, OH, e OCH_3 ; ogni L è selezionato indipendentemente dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legata da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; n è 0 o 1. In alcune realizzazioni, almeno uno tra R^1 , R^2 , R^3 , e R^4 è OH. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della divulgazione comprende un analogo del cap 5' m7GpppA avente la struttura di Formula (Cap IX).

in cui, R^1 , R^2 , e R^3 sono ciascuno indipendentemente selezionato da un alogeno, OH, e OCH_3 ; ogni L è indipendentemente selezionato dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legata da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni, almeno uno di R^1 , R^2 , e R^3 è OH. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della presente invenzione comprende un analogo del cappuccio 5' m7GpppApN, in cui N è un nucleotide naturale o modificato e il cappuccio 5' ha la struttura di Formula (Cap X).

in cui B³ è una nucleobase naturale o modificata; R^1 , R^2 , R^3 , e R^4 sono ciascuno selezionato indipendentemente da un alogeno, OH, e OCH_3 , ogni L è selezionato indipendentemente dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legata da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni, almeno uno tra R^1 , R^2 , R^3 , e R^4 è OH. In alcune realizzazioni B³ è G, m³G, A o m³A; in cui G è guanina, m³G è 7-metilguanina, A è adenina, e m³A è N⁷-metiladenina. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della presente invenzione comprende un analogo del cap 5' m7GpppAmpG avente la struttura di Formula (Cap XI).

in cui, R^1 , R^2 , e R^4 sono ciascuno selezionato indipendentemente da un alogeno, OH, e OCH_3 ; ogni L è selezionato indipendentemente dal gruppo costituito da fosfato, fosforotioato, e

boranofosfato in cui ogni L è legato da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni, almeno uno di R₁, R₂, e R₄ è OH. In alcune realizzazioni, il composto di Formula Cap XI è m¹GpppAmpG, in cui R₁, R₂, e R₃ sono ciascuno OH, n è 1, e ogni L è un legame fosfato. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della divulgazione comprende un analogo del cap 5' m¹GpppApm⁷G avente la struttura di Formula (Cap XII).

in cui, R₁, R₂, R₃, e R₄ sono ciascuno indipendentemente selezionato da un alogeno, OH, e OCH₃, ogni L è indipendentemente selezionato dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legato da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni, almeno uno di R₁, R₂, R₃, e R₄ è OH. In alcune realizzazioni, n è 1.

In alcune realizzazioni, un RNA autoreplicante della presente invenzione comprende un analogo del cap 5' m¹GpppApm⁷G avente la struttura di Formula (Cap XIII).

in cui, R₁, R₂, e R₄ sono ciascuno indipendentemente selezionato da un alogeno, OH, e OCH₃, ogni L è indipendentemente selezionato dal gruppo costituito da fosfato, fosforotioato, e boranofosfato in cui ogni L è legato da legami diesteri; mRNA rappresenta un mRNA della presente divulgazione legato alla sua estremità 5'; e n è 0 o 1. In alcune realizzazioni, almeno uno di R₁, R₂, e R₄ è OH. In alcune realizzazioni, n è 1.

Coda di poli-adenina (poli-A)

La poliadenilazione è l'aggiunta di una coda di poli(A), una catena di nucleotidi di adenina solitamente lunga circa 100-120 monomeri, a un mRNA. Negli eucarioti, la poliadenilazione è parte del processo che produce mRNA maturo per la traduzione e inizia quando termina la trascrizione di un gene. Il segmento più a 3' di un pre-mRNA appena creato viene prima scisso da un set di proteine; queste proteine quindi sintetizzano la coda di poli(A) all'estremità 3'. La coda di poli(A) è importante per l'esportazione nucleare, la traduzione e la stabilità dell'mRNA. La coda si accorcia nel tempo e, quando è sufficientemente corta, l'mRNA viene degradato enzimaticamente. Tuttavia, in alcuni tipi di cellule, gli mRNA con code di poli(A) corte vengono immagazzinati per una successiva attivazione tramite ri-poliadenilazione nel citosol.

Preferibilmente, un RNA autoreplicante della divulgazione comprende una regione di coda 3', che può servire a proteggere l'RNA dalla degradazione dell'esonucleasi. La regione di coda può essere una regione 3'poly(A) e/o 3'poly(C). Preferibilmente, la regione di coda è una coda 3'poly(A). Come utilizzato nel presente documento, una "coda 3'poly(A)" è un polimero di nucleotidi adenina sequenziali che può variare in dimensioni da, ad esempio: da 10 a 250 nucleotidi adenina sequenziali; da 60 a 125 nucleotidi adenina sequenziali, da 90 a 125 nucleotidi adenina sequenziali, da 95 a 125 nucleotidi adenina sequenziali, da 95 a 121 nucleotidi adenina sequenziali, da 100 a 121 nucleotidi adenina sequenziali, da 110 a 121 nucleotidi adenina sequenziali; 112-121 nucleotidi adenini sequenziali; 114-121 nucleotidi adenini sequenziali; oppure 115-121 nucleotidi adenini sequenziali. Preferibilmente, una coda di poli(A) 3' come descritto nel 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 nucleotidi di adenina sequenziali. Le code di poli(A) 3' possono essere aggiunte utilizzando una varietà di metodi noti nell'arte, ad esempio, utilizzando la poli(A) polimerasi per aggiungere code all'RNA sintetico o trascritto in vitro. Altri metodi includono l'uso di un vettore di trascrizione per codificare le code di poli(A) o l'uso di una ligasi (ad esempio, tramite legatura a splint usando una ligasi di RNA T4 e/o una ligasi di DNA T4), in cui il poli(A) può essere legato all'estremità 3' di un RNA senso. In alcune realizzazioni, viene utilizzata una combinazione di uno qualsiasi dei metodi sopra indicati.

Progettazione e sintesi di RNA autoreplicanti

I costrutti per sequenze di RNA autoreplicanti esemplari della presente informativa sono forniti nelle Tabelle 4-5.

TABELLA 4 Confronto di STARR™ autoreplicante RNA della divulgazione con comparativo RNA autoreplicante come descritto Sequenza Costruire Posizione Tipo Sequenza STELLARE™ 5' UTR nucleotide ATGGGCGGCGCATGAGAGAAGCCCAGACCAATTACCO (ID SEQ ACCAAAAA Numero: 49) STELLARE™ non- nucleotide ATGGAGAAAGTTCACGTTGACATCGAGGAAGACAGCC (ID SEQ strutturale CATTCTCAGAGCTTTGCAGCGGAGCTTCCCGCAGTTTT N.: 50) gene ORF

GAGGTAGAAGCCAAGCAGGTCCTGATAATGACCATG
CTAATCGCCAGAGCGTTTTTCGCATCTGGCTTCAAACTG
ATCGAAACGGAGGTGGACCCATCCGACACGATCCTTG
ACATTGGAAGTGCGCCCGCCGCGAGAATGTATTCTAA
GCACAAGTATCOATTGTATCTGTCCGATGAGATGTGCGG
AAGATCCGGACAGATTGTATAAGTATGCAACTAAGCT
GAAGAAAACTGTAAGGAAATAACTGATAAGGAATTG
GACAAGAAAATGAAGGAGCTGGCCGCCGTCATGAGCG
ACCCTGACCTGGAAACTGAGACTATGTGCCTCCACGA
CGACGAGTCGTGTCGCTACGAAGGGCAAGTCGCTGTT TACCAGGATGTACGCCGTCGACGGCCCCACCAGCC
TGTACCACCAGGCCAACAAGGGCGTGAGGGTGGCCTA
CTGGATCGGCTTCGACACCACACCCTTCATGTTCAAGA
ACCETTAGGGCCGCCCTACCCAGCTACAGCACCAACTG
GGCCGACGAGACCGTGCTGACCGCCAGGAACATCGGC
CTGTGCAGCAGCCGACGTGATGGAGAGGAGCCGGAGAG
GCATGAGCATCCTGAGGAAGAAATACCTGAAGCCCAG
CAACAACGTGCTGTTTCAGCGTGGGCAGCACCATCTAC
CACGAGAAGAGGGACCTGCTCAGGAGCTGGCACCTGC
CCAGCGTGTTCCACCTGAGGGGGCAAGCAGA ACTACAC
CTGCAGGTGCGAGACCATCGTGAGCTGCGACGGCTAC
GTGGTGAGAGGATCGCCATCAGCCCCGGCCTGTACG
GCAAGCCCAGCGGCTACGCCGCTACAATGCACAGGGA GGGCTTCCTGTGCTAGGTGACCGACACCCTGAAC
GGCGAGAGGGTGAGCTTCCCCGTGTGCACCTACGTGC
CCGCCACCCTGTGCGACCAGATGACCGGCATCCTGGC
CACCGACGTGAGCGCCGACGACGCCAGAAGCTGCTC
GTGGGCCTGAACCAGAGGATCGTGGTCAACGGCAGGA

CCCAGAGGAACACCAACACAATGAAGAACTACCTGCT
GCCCCGTGGTGGCCCAGGCTTTCGCCAGGTGGGCCAAG
GAGTACAAGGAGGACCAGGAAGACGAGAGGCCCCCTG
GGCCTGAGGGACAGGCAGCTGGTGATGGGCTGCTGCT
GGGCCTTCAGGCGGCACAAGATCACCAGCATCTACAA
GAGGCCCCGACACCCAGACCATCATCAAGGTGAACAGC
GACTTCCACAGCTTCGTGCTGCCAGGATCGGCAGCA
ACACCCTGGAGATCGGCCTGAGGACCCGGATCAGGAA
GATGCTGGAGGAACACAAGGAGCCCAGCCACTGATC
ACCGCCGAGGACGTGCAGGAGGCCAAGTGCGCTGCCG
ACGAGGCCAAGGAGGTGAGGGAGGCCGAGGAACTGA
GGGCCGCCCTGCCACCCCTGGCTGCCGACGTGGAGGA
ACCCACCCTGGAAGCCGACGTGGACCTGATGCTGAC
GAGGCCGGCGCCGGAAGCGTGAGACACCCAGGGGC
CTGATCAAGGTGACCAGCTACGACGGCGAGGACAAGA
TCGGCAGCTACGCCGTGCTGAGCCCACAGGCCGTGCT
GAAGTCGAGAAGCTGAGCTGCATCCACCACCCACTGGCC
GAGCAGGGTGATCGTGATCACCCACAGCGGCAGGAAGG
GCAGGTACGCCGTGGAGCCCTACCACGGCAAGGTGT
CGTGCCCGAGGGCCACGCCATCCCCGTGCAGGACTTC
CAGGCCCTGAGCGAGAGCGCCACCATCGTGTACAACG
AGAGGGAGTTTCGTGAACAGGTACCTGCACCATATCGC
CACCCACGGCGGAGCCCTGAACACCGACGAGGAATAC
TACAAGACCGTGAAGCCCAGCGAGCACGACGGCGAGT
ACCTGTACGACATCGACAGGAAGCAGTGCGTGAAGAA
AGAGCTGGTGACCGGCCTGGGACTGACCGGCGAGCTG
GTGGACCCACCCTTCCACGAGTTTCGCCTACGAGAGCCT
GAGGACCAGACCCGCCGCTCCCTACCAGGTGCCACC
ATCGGGCGTGTACGGCGTGCCCGGCAGCGGAAAGAGCG
GCATCATCAAGAGCGCCGTAGACCAAGAAAGACCTGGT
GGTCAGCGCCAAGAAAGAGAACTGCGCCGAGATCATC
AGGGACGTGAAGAAGATGAAAGGCCTGGACGTGAAC
GCGCGCACCGTGGACAGCGGTGCTGCTGAACGGCTGCA
AGCACCCCGTGGAGACCCTGTACATCGACGAGGCCTT
CGCTTGCCACGCCGGCACCCCTGAGGGCCCTGATCGCC
ATCATCAGGCCCAAGAAAGCCGTGCTGTGCGGCGACC
CCAAGCAGTGCGGCTTCTTCAACATGATGTGCCTGAAG
GTGCACTTCAACCACGAGATCTGCACCCAGGTGTTCCA
CAAGAGCATCAGCAGGCGGTGCACCAAGAGCGTGACC
AGCGTCGTGAGCACCCCTGTTCTACGACAAGAAAATGA
GGACCACCAACCCCAAGGAGACCAAAAATCGTGATCGA
CACCACAGGCAGCCAAGCCCAAGCAGGACGACCTG
ATCCTGACCTGCTTCAGGGGCTGGGTGAAGCAGCTGC
AGATCGACTACAAGGGCAACGAGATCATGACCGCCGC

TGCCAGCCAGGGCCTGACCAGGAAGGGCGTGTACGCC
GTGAGGTACAAGGTGAACGAGAACCCACTGTACGCTC
CCACCAGCGAGCACGTGAACGTGCTGCTGACCAGGAC
CGAGGACAGGATCGTGTGGAAGACCCTGGCCGGCGAC
CCCTGGATCAAGACCCTGACCGCCAAGTACCCCGGCA
ATTOTCACCGCCACCATCGAAGAGTGGCAGGCCGAGCA
CGACGCCATCATGAGGCACATCCTGGAGAGGCCCGAC CCCACCGACGTGTTCCAGAACGTGTGCT
GGGCCAAGGCCCTGGTGCCCGTGCTGAAGACCGCCGG
CATCGACATGACCACAGAGCGAGTGGAACACCGTGGAC
TATTOTCGAGACCGACAAGGCCACAGCGCCGAGATCG
TGCTGAACCAGCTGTGCGTGAGGTTCTTCGGCCTGGAC
CTGGACAGCGCCTGTTTCAGCGCCCCACCGTGCCACT
GAGCATCAGGAAACAACCACTGGGACAACAGCCCCAGC
CCAAACATGTACGGCCTGAACAAGGAGGTGGTCAGGC
AGCTGAGCAGGGCGGTACCCACAGCTGCCAGGGCCGT
GGCCACCGGCAGGGGTGTACGACATGAACACCGGCACC
CTGAGGAACTACGACCCAGGATCAACCTGGTCCCG
TGAACAGGCGGCTGCCCCACGCCCTGGTGCTGCACCA
CAACGAGCACCCACAGAGCGACTTCAGCTCCTTCGTG
AGCAAGCTGAAAGGCAGGACCGTGCTGGTCTGGGCG
AGAAGCTGAGCGTGCCCGCAAGATGGTGGACTGGCT
GAGCGACAGGCCCGAGGCCACCTTCCGGGCCAGGCTG
GACCTCGGCATCCCGGCGACGTGCCAAGTACGACA
TCATCTTCGTGAACGTCAGGACCCATAACAAGTACCAC
CATTACCAGCAGTGCGAGGACCACGCCATCAAGCTGA
GCATGCTGACCAAGAAGGCCTGCCTGCACCTGAACCC
CGGAGGCACCTGCGTGAGCATCGGCTACGGCTACGCC
GACAGGCCAGCGAGAGCATCATTGGCGCCATCGCCA
GGCTGTTCAAGTTCAGCAGGGTGTGCAAACCCAAGAG
CAGCCTGGAGGAAACCGAGGTGCTGTCGTGTTTCATC
GGCTACGACCGGAAGGCCAGGGACCCACAACCCCTACA
AGCTGAGCAGCACCCCTGACAAACATCTACACCGGCAG
CAGGCTGCACGAGGCCGGCTGCGCCCCAGCTACCAC
GTGGTCAGGGCGATATCGCCACCGCCACCGAGGGCG
TGATCATCAACGCTGCCAACAGCAAGGGCCAGCCCG
AGGCGGAGTGTGCGGGCCCTGTACAAGAAGTTCCCC
GAGAGCTTCGACCTGCAGCCATCGAGGTGGGCAAGG
CCAGGCTGGTGAAGGGCGCCGCTAAGCACATCATCCA
CGCCGTGGGCCCAACTTCAACAAGGTGAGCGAGGTG
GAAGGCGACAAGCAGCTGGCCGAAGCCTACGAGAGC
ATCGCCAAGATCGTGAACGACAATAACTACAAGAGCG
TGGCCATCCCCTGCTCAGCCGGCATCCTCAGCACCGGCATCAGCGGC
AACAAAGGACAGGCTGACCCAGAGCCTGAACCACCTGC TCACCGCCCTGGACACCGATGCCGACGTGCCAT
CTACTGCAGGGACAAGAAGTGGGAGATGACCCTGAAG

GAGGCCGTGGCCAGGCGGGAGGCCGTGGAAGAGATCT
GCATCAGCGACTCCAGCGTGACCGAGCCCGACGC
CGAGCTGGTGAGGGTGCACCCCAAGAGCTCCCTGGCC
GGCAGGAAGGGCTACAGCACCAGCGACGGCAAGACCT
TCAGCTACCTGGAGGGCACCAAGTTCACCAGGGCCGC
TAAGGACATCGCCGAGATCAACGCTATGTGGCCCGTG
GCCACCGAGGCCAACGAGCAGGTGTGCATGTACATCC
TGGGCGAGAGCATGTCCAGCATCAGGAGCAAGTGCC
CGTGGAGGAAAGCGAGGCCAGCACACCACCCAGCACC
CTGCCCTGCCTGTGCATCCACGCTATGACACCCGAGAG
GGTGCAGCGGCTGAAGGGCCAGCAGGCCCGAGCAGATC
ACCGTGTGCAGTCCTTCCCCTGCCCAAGTACAGGAT
CACCGGCGTGCAGAAGATCCAGTGCAGCCAGCCCATC
CTGTTCCAGCCCAAAGGTGCCCGCTACATCCACCCAG
GAAGTACCTGGTGGAGACCCACCCGTGGACGAGACA
CCGAGCCAAGCGCCGAGAACCAGAGCACCGAGGGC
ACACCCGAGCAGCCACCCCTGATCACCGAGGACGAGA
CAAGGACCCGACCCAGAGCCATCATTATCGAGGA
AGAGGAAGGAGACAGCATCAGCCTGCTGAGCGACGGC
CCCACCCACCAGGTGCTGCAGGTGGAGGCCGACATCC
ACGGCCCCACCCAGCGTGTCCAGCTCCAGCTGGAGCAT
CCCACACGCCAGCGACTTCGACGTGGACAGCCTGAGC
ATCCTGGACACCCTGGAGGGCGCCAGCGTGACCTCCG
GCGCCACCAGCGCCGAGACCAACAGCTACTTCGCCAA
GAGCATGGAGTTCCTGGCCAGGCCCGTGCCAGTCCC
AGGACCGTTCAGGAACCCACCCACCCAGCTCCCA
GGACCAGGACCCCAAGCCTGGCTCCCAGCAGGGCCTG
CAGCAGGAGGACCAGCCTGGTGAGCACCCACCCGGCGTG
AACAGGGTGATCACCAGGGAGGAACTGGAGCCCTGA CACCCAGCAGGACCCCAAGGTCCGTGAGCAGGAC
TAGTCTGTGTCCAACCCACCCGGCGTGAACAGGGTG
ATCACCAGGGAGGAATTCGAGGCCTTCGTGGCCCAGC
AACAGAGACGGTTCGACGCCGGCGCCTACATCTTCAG
CAGCGACACCGGCCAGGGACACCTGCAGCAAAAGAGC
GTGAGGCAGACCGTGTGAGCGAGGTGGTGCTGGAGA
GGACCGAGCTGGAATCAGCTACGCCCCAGGCTGGA
CCAGGAAGGAAGGAGGAACTGCTCAGGAAGAACTGCA
GCTGAACCCACCCAGCCAACAGGAGCAGGTACCAG
AGCAGGAAGGTGGAGAACATGAAGGCCATCACCGCCA
GGCGGATCCTGCAGGGCCTGGGACACTACCTGAAGGC
CGAGGGGCAAGGTGGAGTGCTACAGGACCCTGCACCCC
GTGCCACTGTACAGCTCCAGCGTGAACAGGGCCTTCTC
CAGCCCCAAGGTGGCCGTGGGAGGCCTGCAACGCTATG
CTGAAGGAGAACTTCCCCACCGTGGCCAGCTACTGCA
TCATCCCCGAGTACGACGCCTACCTGGACATGGTGGA

CGGCGCCAGCTGCCTGCGGACACCGCCAGCTTCTGCC
CCGCCAAGCTGAGGAGCTTCCCAAGAAACACAGCTA
CCTGGAGCCCACCATCAGGAGCGCCGTGCCAGCGCC
ATCCAGAACACCCTGCAGAACGTGCTGGCCGCTGCCA
CCAAGAGGAACTGCAACGTGACCCAGATGAGGGAGCT
GCCCCTGCTGGACAGCGCTGCCTTCAACGTGGAGTGCT
TCAAGAAATACGCCTGCAACAACGAGTACTGGGAGAC
CTTCAAGGAGAACCCCATCAGGCTGACCGAAGAGAAC
GTGGTGAACATACATCACCAGCTGAAGGGCCCAAGG
CCGCTGCCCTGTTTCGCTAAGACCCACAACCTGAACATG
CTGCAGGACATCCCAATGGACAGTTTCGTGATGGACC
TGAAGAGGGACGTGAAGGTGACACCCGGCACCAAGCA
CACCGAGGAGAGGCCAAGGTGCAGGTGATCCAGGCC
GCTGACCCACTGGCCACCGCCTACCTGTGCGGCATCCA
CAGGGAGCTGGTGAGGCGGCTGAACGCCGTGCTGCTG
CCCAACATCCACACCCTGTTTCGACATGAGCGCCGAGG
ACTTCGACGCCATCATCGCCGAGCACTTCCAGCCGGC
GACTGCGTGCTGGAGACCGACATCGCCAGCTTCGACA
AGAGCGAGGATGACGCTATGGCCCTGACCGCTCTGAT
GATCCTGGAGGACCTGGGCGTGGACGCCGAGCTGCTC
ACCCTGATCGAGGCTGCCTTCGGCGAGATCAGCTCCAT
CCACCTGCCCACCAAGACCAAGTTCAAGTTCGGCGCT
ATGATGAAAAGCGGAATGTTCTGACCCTGTTTCGTGA
ACACCGTGATCAACATTGTGATCGCCAGCAGGGTGCT
GCGGGAGAGGCTGACCGGCAGCCCCTGCGCTGCCTTC
ATCGGCGACACAAACATCGTGAAGGGCGGTGAAAAGCG
ACAAGCTGATGGCCGACAGGTGCGCCACCTGGCTGAA
CATGGAGGTGAAGATCATCGACGCCGTGGTGGGCGAG
AAGGCCCCCTACTTCTGCGGCGGATTCATCCTGTGCGA
CACGTGACCGGCACCGCCTGCAGGGTGGCCGACCCC
CTGAAGAGGGGCTGTTCAAGCTGGGCAAGCCACTGGCCG
CTGACGATGAGCACGACGATGACAGGCGGAGGGCCCT
GCACGAGGAAAGCACCAGGTGGAACAGGGTGGGCAT
CCTGAGCGAGCTGTGCAAGGCCGTGGAGAGCAGGTAC
GAGACCGTGGGCACCAGCATCATCGTGATGGCTATGA
CCCACTGGCCAGCTCCGTCAAGAGCTTCTCCTACCTG AGGGGGCCCCTATAACTCTCTACGGCTAA
STELLARE™ non- aminoacido MEKVHVDIEEDSPFLRALQRSFPQFEVEAKQVTDNDHAN (ID SEQ
strutturale ARAFSHLASKLIETEVDPSDTILDIGSAPARRMYSKHKYH N.: 51) gene ORF
CICPMRCAEDPDRLYKYATKLKKNCKEITDKELDKMK
ELAAVMSDPDLETETMCLHDDDESCRYEGQVAVYQDVY
AVDGPTSLYHQANKGVRVAYWIGFDTPFMFKNLGAY
PSYSTNWADETIVLTARNIGLCSSDVMERSRRGMSILRKK
YLPKPSNNVLFVSGSTIYHEKRDLRSWHLPSVFHLRGKQ
NYTCRCETIVSCDGYVVKRIAISPGLYGKPSGYAATMHR

EGFLCCKVDTLNGERSFPVCTYVPATLCDQMTGILAT
DVSADDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPV
VAQAFARWAKEYKEDQEDERPLGLRDRQLVMGCCWAF
RRHKITSIYKRPDTQTIKVNDFHSFVLPRIGSNTLEIGLR
TRIRKMLEEHKEPSPLITAEDVQEAKCAADEAKEVREAE
ELRAALPPLAADVEEPTLEADVLDLMLQEAGAGSVETPRG
LIKVTSYDGEDKIGSYAVLSPQAVLKSEKLSCHPLAEQVI
VITHSGRKGRYAVEPYHGKVVVPEGHAIPVQDFQALSES
ATIVYNREFVNRYLHHIATHGGALNTDEEYKTVKPSE
HDGEYLYDIDRKQCVKKELVTLGLTGLVDPFHEFAY
ESLRRTRPAAPYQVPTIGVYGVPGSGKSGIISAVTKKDLV
VSAKKENCAEIIRDVKKMKGLDVNARTVDSVLLNGCKH IDEE PER LA PUNTATA DI
VETROFABBRICATA DI AGTLRALIAIIRPKKAVLCGDPKQCG
FFNMMCLKVHFNHEICTQVFHKSISRRTKSVTSVVSTLF
YDKMRTTNPKETKIVIDTTGSTKPKQDDLILTCFRGWV
KQLQIDYKGNEMTAAASQGLTRKGVYAVRYKVNENPL
YAPTSEHVNVLLTRTEDRIVWKTLAGDPWIKTLTAKYPG
NFTATIEEWQAEHDAIMRHILERPDPTDVFQNKANVCWA
KALVPVLKTAGIDMTTEQWNTVDYFETDKAHSAEIVLN
QLCVRFFGLDLDLSDGLFSAPTVPVLSIRNNHWDNSPSPNMY
GLNKEVVRQLSRRYPQLPRAVATGRVYDMNTGTLRNYD
PRINLVPVNRRLPHALVLHHNEHPQSDFFSVSKLKGRTV
LVVGEKLSVPGKMVDWLSRPEATFRARLDLGPVGDVP
KYDIIFVNVRTPYKYHHYQCEDHAIKLSMLTKKACLHL
NPGGTCVSIQGYADRASESIIIGAIARLFKFSRVCKPKSSL
EETEVLFFVIGYDRKARTHNPYKLSSTLTNIYTGSRLEHA
GCAPSYHVVRGDIATATEGVIINAANSKGQPGGGVCGAL
YKFPESFDLQPIEVGKARLVKGAAKHIIHAVGPNFNKVS
EVEGDKQLAEAYESIAKIVNDNNYKSVAIPLLSTGIFSGN
KDRLTQSLNHLTALDTTDADVAIYCRDKKWEMTLKEA
VARREAVEEICISDDSSVTEPDAELVRVHPKSSLAGRKGY
STSDGKTFSYLEGTFKHQAADIAEINAMWPVATEANEQ
VCMYILGESMSSIRSKCPVEESEASTPPSTLPCLCIHAMTP
ERVQRLKASRPEQITVCSFPLPKYRITGVQKIQCSQPILFS
PKVPAYIHPRKYLVEVTPVDETPEPSAENQSTEGTPEQPPL
ITEDETTRTRTPEPIIIIEEEDSISLLSDGPTHQVLQVEADIH
GPPSVSSSSWSIPHASDFDVSLSILDTLLEGASVTSGATSA
ETNSYFAKSMEFLARPVPAPRTVFRNPPHPAPRTRTPSLA
PSRACRSTSLVSTPPGVNRVITREELEALTPSRTPSRVSR
TSLVSNPPGVNRVITREEFEAFVAQQRRFDAGAYIFSSD
TGQGHLLQKSVRQTVLSEVVLERTELEISYAPRLDQEKE
ELLRKKLQLNPTPANRSRYQSRKVENMKAITARRILQGL
GHYLKAEGKVECYRTLHPVPLYSSSVNRAFSSPKVAVEA
CNAMLKENFPTVASYCIPEYDAYLDMVDGASCCLDTAS

FCPAKLRSPKHSYLEPTIRSAVPSAIQNTLQNVLAAAT
KRNCNVTQMRELPVLDSAAFNVECFKKYACNNEYWETF
KENPIRLTEENVVNYITKLKGPKAAALFAKTHNLNMLQD
IPMDRFVMDLKRVDKVTGPKHTEERPKVQVIQAADPL
ATAYLCGIHRELVRRLNAVLLPNIHTLFDMSAEDFDAIA
EHFQPGDCVLETDIASFDKSEDDAMALTALMILEDLQVD
AELLTLIEAAFGEISSIHLPTKTKFKFGAMMKSGMFLTLF
VNTVINIVIASRVLRLRERLTGSPCAAFIGDDNIVKGVKSDK
LMADRCATWLNMEVKIIDAVVGEKAPYFCGGFILCDSVT
GTACRVADPLKRLFKLGKPLAADDEHDDRRRRALHEES
TRWNRVIGILSELCKAVESRYETVGTSHIIMAMTTLASSV KSFSYLRGAPITLYG* STELLARE™ intergenico
nucleotide CCTGAATGGACTACGACATAGTCTAGTCCGCCAAGGC (ID SEQ regione CGCCACC N.: 52)
STELLARE™ transgene nucleotide n/a (dipende dal gene di nostro interesse) ORF STELLARE™ 3'
UTR nucleotide ACTCGAGTATGTTACGTGCAAAGGTGATTGTCACCCCC (ID SEQ
CGAAAGACCATATTGTGACACACCCTCAGTATCACGC N.: 53)
CCAAACATTTACAGCCGCGGTGTCAAAAACCGCGTGG
ACGTGGTTAACATCCCTGCTGGGAGGATCAGCCGTAA
TTATTATAATTGGCTGGTGCTGGCTATTOATTGTGGCC
ATGTACGTGCTGACCAACCAGAAACATAATTGAATAC
AGCAGCAATTGGCAAGCTGCTTACATAGAACTCGCGG
CGATTGGCATGCCGCCTTAAAATTTTTATTTTTTTTTT
CTTTCTTTTCCGAATCGGATTTTGTTTTTAATATTCA
AAAAAAAAAAAAAAAAAAAAAAAAATCTTAGAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
Comparativo 5' UTR nucleotide sconosciuto Originale non- nucleotide
ATGCCCGAGAAGGTGCACGTGGACATCGAGGAGGACA (ID SEQ strutturale
GCCCTTCCCTGAGGGCCCTGCAGAGGAGCTTCCCACA N.: 54) gene ORF
GTTCGAAGTGGAGGCCAAGCAGGTGACCGACAACGAC
CACGCCAACGCCAGGGCCTTCAGCCACCTGGCCAGCA
AGCTGATCGAGACCGAGGTGGACCCAGCGACACCAT
CCTGGACATCGGCAGCGCCCCAGCCAGGAATGTAC AGCAAGCACAAGTACCACTGCATCTGCCCCATGAGGT
GCGCCGAGGACCCCGACAGGCTGTACAAGTACGCCAC
CAAAGTGAAGAAGAACTGCAAGGAGATCACCGACAA
GGAGCTGGACAAGAAAATGAAGGAGCTGGCCGCCGTG
ATGACGACCCCGACCTGGAGACCGAGACAATGTGCC
TGCACGACGACGAGAGCTGCAGGTACGAGGGCCAGGT
GGCCGTCTACCAGGACGTGTACGCCGTCGACGGCCCC
ACCAGCCTGTACCACCAGGGCCAACAAGGGCGTGAGGG
TGGCCTACTGGATCGGCTTCGACACCACACCCTTCATG
TTCAAGAACCTGGCCGGCGCCTACCCAGCTACAGCA
CCAACTGGGGCCGACGAGACCGTGCTGACCGCCAGGAA
CATCGGCCTGTGCAGCAGCGACGTGATGGAGAGGAGC
CGGAGAGGCATGAGCATCCTGAGGAAGAAATACCTGA

AGCCCAGCAACAACGTGCTGTTTCAGCGTGGGCAGCAC
CATCTACCACGAGAAGAGGGACCTGCTCAGGAGCTGG
CACCTGCCCACGTGTTCCACCTGAGGGGGCAAGCAGA
ATTOACACCTGCAGGTGCGAGACCATCGTGAGCTGCGA
CGGCTACGTGGTGAAGAGGATCGCCATCAGCCCCGGC
CTGTACGGCAAGCCCAGCGGCTACCGCCGCTACAATGC
ACAGGGAGGGCTTCCTGTGCAAGGTGACCGACAC CCTGAACGGCGAGAGGGTGAGCTTCCCCGTGTGCACC
TACGTGCCCCGCCACCCTGTGCGACCAGATGACCGGCA
TCCTGGCCACCGACGTGAGCGCCGACGACGCCAGAA
GCTGCTCGTGGCCTGAACCAGAGGATCGTGGTCAAC
GGCAGGACCAGAGGAACACCAACAATGAAGAAC
TACCTGCTGCCCCGTGGTGGCCAGGCTTTCGCCAGGTG
GGCCAAGGAGTACAAGGAGGACCAGGAAGACGAGAG
GCCCCGAGGCTGAGGGACAGGCAGCTGGTGTGGGC
TGCTGCTGGCCTTCAGGCGGCACAAGATCACCAGCA
TCTACAAGAGGCCCCGACACCCAGACCATCATCAAGGT
GAACAGCGACTTCCACAGCTTCGTGCTGCCAGGATC
GGCAGCAACACCCTGGAGATCGGCCTGAGGACCCGGA
TCAGGAAGATGCTGGAGGAACACAAGGAGCCCAGCCC
ACTGATCACCGCCGAGGACGTGCAGGAGGCCAAGTGC
GCTGCCGACGAGGCCAAGGAGGTGAGGGAGGCCGAG
GAACTGAGGGCCGCCCTGCCACCCCTGGGTGCCGACG
TGGAGGAACCCACCCTGGAAGCCGACGTGGACCTGAT
GCTGCAGGAGGCCGGCGCCGGAAGCGTGGAGACCCC
AGGGGCTGATCAAGGTGACCAGCTACGACGGCGAGG
ACAAGATCGGCAGCTACGCCGTGCTGAGCCCACAGGC CGTGCTGAAGTCCGAGAGCTGCATCCACCCA
CTGGCCGAGCAGGTGATCGTGATCACCCACAGCGGCA
GGAAGGGCAGGTACGCCGTGGAGCCCTACCACGGCAA
GGTGGTTCGTGCCCCGAGGGCCACGCCATCCCCGTGCAG
GACTTCCAGGCCCTGAGCGAGAGCCGCCACCATCGTGT
ACAACGAGAGGGAGTTCGTGAACAGGTACCTGCACCA
TATCGCCACCCACGGCGGAGCCCTGAACACCGACGAG
GAATACTACAAGACCGTGAAGCCCAGCGAGCACGACG
GCGAGTACCTGTACGACATCGACAGGAAGCAGTGCGT
GAAGAAAGAGCTGGTGACCGGCCCTGGGACTGACCGGC
GAGCTGGTGGACCCACCCTTCCACGAGTTCGCTACGA
GAGCCTGAGGACCAGACCCGCCGCTCCCTACCAGGTG
CCCACCATCGCGTGTACGGCGTGCCCCGGCAGCGGAA
AGAGCGGCATCATCAAGAGCGCCGTGACCAAGAAAGA
CCTGGTGGTACGCGCAAGAAAGAGAACTGCGCCGAG
ATCATCAGGGACGTGAAGAAGATGAAAGGCCTGGACG
TGAACGCGCGCACCGTGGACAGCGTGCTGCTGAACGG
CTGCAAGCACCCCGTGGAGACCCTGTACATCGACGAG
GCCTTCGCTTGCCACGCCGGCACCCCTGAGGGCCCTGAT

CGCCATCATCAGGCCCAAGAAAGCCGTGCTGTGCGGC
GACCCCAAGCAGTGC GGCTTCTTCAACATGATGTGCCT
GAAGGTGCACTTCAACCACGAGATCTGCACCCAGGTG
TTCCACAAGAGCATCAGCAGGCGGTGCACCAAGAGCG
TGACCAGCGTCGTGAGCACCTGTTCTACGACAAGAA
AATGAGGACCACCAACCCCAAGGAGACCAAAATCGTG
ATCGACACCACAGGCAGCACCAAGCCCAAGCAGGACG
ACCTGATCCTGACCTGCTTCAGGGGCTGGGTGAAGCA
GCTGCAGATCGACTACAAGGGCAACGAGATCATGACC
GCCGCTGCCAGCCAGGGCCTGACCAGGAAGGGCGTGT
ACGCCGTGAGGTACAAGGTGAACGAGAACCCACTGTA
CGCTCCCACCAGCGAGCACGTGAACGTGCTGCTGACC
AGGACCGAGGACAGGATCGTTGTGGAAGACCCTGGCCG
GCGACCCCTGATCAAGACCCTGACCGCCAAGTACCC
CGGCAACTTCAACGCCACCATCGAAGAGTGGCAGGCC
GAGCACGACGCCATCATGAGGCACATCCTGGAGAGGC
CCGACCCACCGACGTGTTCCAGAACAAGGCCAACGT
GTGCTGGGCCAAGGCCCTGGTGCCCGTGTGTAAGACC
GCCGGCATCGACATGACCACAGAGCAGTGGAACACCG
TGGACTACTTCGAGACCGACAAGGCCACAGCGCCGA
GATCGTGCTGAACCAGCTGTGCGTGAGGTTCTTCGGCC
TGGACCTGGACAGCGGCCTGTTACGCGCCCCACCCT
GCCACTGAGCATCAGGAACAACCACTGGGACAACAGC
CCCAGCCCAAACATGTACGGCCTGAACAAGGAGGTGG
TCAGGCAGCTGAGCAGGCGGTACCCACAGCTGCCAG
GGCCGTGGCCACCGGCAGGGGTGTACGACATGAACACC
GGCACCTGAGGAACTACGACCCAGGATCAACCTGG
TGCCCGTGAACAGGCGGCTGCCCCACGCCCTGGTGCT
GCACCACAACGAGCACCCACAGAGCGACTTCAGCTCC
TTCGTGAGCAAGCTGAAAGGCAGGACCGTGCTGGTCG
TGGCGGAGAAGCTGAGCGTGCCCGCAAGATGGTGGAA
CTGGCTGAGCGACAGGCCCGAGGCCACCTTCCGGCC
AGGCTGGACCTCGGCATCCCCGGCGACGTGCCCAAGT
ACGACATCATCTTCGTGAACGTCAGGACCCCATACAA
GTACCACCATTACCAGCAGTGCGAGGACCACGCCATC
AAGCTGAGCATGCTGACCAAGAAGGCCTGCCTGCACC
TGAACCCCGGAGGCACCTGCGTGAGCATCGGCTACGG
CTACGCGACAGGGCCAGCGAGAGCATCATTGGCGCC
ATCGCCAGGCTGTTCAAGTTCAGCAGGGTGTGCAAAC
CCAAGAGCAGCCTGGAGGAAACCGAGGTGCTGTTCTG
GTTTCATCGGCTACGACCGGAAGGCCAGGACCCACAAC
CCCTACAAGCTGAGCAGCACCTGACAAACATCTACA
CCGGGCAGCAGGCTGCACGAGGCCGGTGCGCCCCAG
CTACCACGTGGTCAGGGCGATATCGCCACCGCCACC

GAGGGCGTGATCATCAACGCTGCCAACAGCAAGGGCC
AGCCCCGAGGCGGAGTGTGCGGCGCCCTGTACAAGAA
GTTCCCCGAGAGCTTCGACCTGCAGCCCATCGAGGTG
GGCAAGGCCAGGCTGGTGAAGGGCGCCGCTAAGCACA
TCATCCACGCCGTGGGCCCAACTTCAACAAGGTGAG
CGAGGTGGAAGGCGACAAGCAGCTGGCCGAAGCCTAC
GAGAGCATCGCCAAGATCGTGAACGACAATAACTACA AGAGCGTGGCCATCCCCTGCTCAGCCGGCATCTC
AGCGGCAACAAGGACAGGCTGACCCAGAGCCTGAACC
ACCETTAZIONEGCCCTGCACCACCGCCCTGGACACCACCGATGCCGACGT
GGCCATCTACTGCAGGGGACAAGAAGTGGGAGATGACC
CTGAAGGAGGCCGTGGCCAGGCGGGAGGCCGTGGAA
GAGATCTGCATCAGGACGACTCCAGCGTGACCGAGC
CCGACGCCGAGCTGGTGAAGGTGCACCCCAAGAGCTC
CCTGGCCGGCAGGAAGGGCTACAGCACCAGCGACGGC
AAGACCTTACGCTACCTGGAGGGCACCAAGTTCACC
AGGCCGCTAAGGACATCGCCGAGATCAACGCTATGTG
GCCCCGTGGCCACCGAGGCCAACGAGCAGGTGTGCATG
TACATCCTGGGCGAGAGCATGTCCAGCATCAGGAGCA
AGTGCCCCGTGGAGGAAAGCGAGGCCAGCACACCACC
CAGCACCTGCCCTGCCTGTGCATCCACGCTATGACAC
CCGAGAGGTGCAGCGGCTGAAGGCCAGCAGGCCCGA
GCAGATCACCGTTGGCAGTCCTTCCCCTGCCCCAAGT
ACAGGATCACCGGCGTGCAGAAGATCCAGTGCAGCCA
GCCCATCCTGTTACAGCCCAAAGGTGCCCGCCTACATCC
ACCCAGGAAGTACCTGGTGGAGACCCACCCGTGGA
CGAGACACCCGAGCCAAGCGCCGAGAACCAGAGCACC
GAGGCACACCCGAGCAGCCACCCCTGATCACCGAGG
ACGAGACAAGGAACCCGGACCCAGAGCCCATCATTAT
CGAGGAAGAGGAAGAGGACAGCATCAGCCTGCTGAG
CGACGGCCCCACCCACCAGGTGCTGCAGGTGGAGGCC
GACATCCACGGCCCACCCAGCGTGTCCAGCTCCAGCT
GGAGCATCCACACGCCAGCGACTTCGACTGGACAG
CCTGAGCATCCTGGACACCCTGGAGGGCGCCAGCGTG
ACCTCCGGCGCCACCAGCGCCGAGACCAACAGCTATTO
TCGCCAAGAGCATGGAGTTCCTGGCCAGGCCCGTGCC
AGTCCCAGGACCGTGTTCAGGAACCCACCCACCCA
GCTCCCAGGACCAGGACCCCAAGCCTGGCTCCCAGCA
GGGCCTCAGCAGGACCAGCCTGGTGAGACCCACC
CGGCGTGAACAGGGTGATCACCCAGGGAGGAACTGGAG
GCCCTGACACCCAGCAGGACCCCCAGCAGGTCCGTGA
GCAGGACTAGTCTGTGTCCAACCCACCCGGGTGAA
CAGGGTGATCACCCAGGGAGGAATTCGAGGCCTTCGTG
GCCAGCAACAGAGACGGTTCGACGCCGGCGCCTACA
TCCTTCAGCAGCGACACCCGGCCAGGGACACCTGCAGCA

AAGAGCGTGAGGCAGACCGTGCTGAGCGAGGTGGTG
CTGGAGAGGACCGAGCTGGAATCAGCTACGCCCCA
GGCTGGACCAGGAGAAGGAGGAACTGCTCAGGAAGA
AACTGCAGCTGAACCCACCCCAGCCAACAGGAGCAG
GTACCAGAGCAGGAAGGTGGAGAACATGAAGGCCATC
ACCGCCAGGCGGATCCTGCAGGGCCTGGGACTACC
TGAAGGCCGAGGGCAAGGTGGAGTGCTACAGGACCCT
GCACCCCGTGCCACTGTACAGCTCCAGCGTGAACAGG
GCCTTCTCCAGCCCCAAGGTGGCCGTGGAGGCCTGCA
ACGCTATGCTGAAGGAGAACTTCCCCACCGTGGCCAG
CTACTGCATCATCCCCGAGTACGACGCTACCTGGACA
TGGTGGACGGGCGCCAGCTGCTGCCTGGACACCGCCAG
CTTCTGCCCCGCCAAGCTGAGGAGCTTCCCCAAGAAA
CACAGCTACCTGGAGCCCACCATCAGGAGCGCCGTGC
CCAGCGCCATCCAGAACACCCTGCAGAACGTGCTGGC
CGCTGCCACCAAGAGGAACTGCAACGTGACCCAGATG
AGGGAGCTGCCCGTCTGGACAGCGCTGCCTTCAACG
TGGAGGTGCTTCAAGAAATACGCCTGCAACAACGAGTA
CTGGGAGACCTTCAAGGAGAACCCCATCAGGCTGACC
GAAGAGAACGTGGTGAACACTACATCACCAAGCTGAAGG
GCCCCAAGGCCGCTGCCCTGTTGCTAAGACCCACAA
CCTGAACATGCTGCAGGACATCCCAATGGACAGGTTC
GTGATGGACCTGAAGAGGGACGTGAAGGTGACACCCG
GCACCAAGCACACCGAGGAGAGGCCCAAGGTGCAGGT
GATCCAGGCCGCTGACCCACTGGCCACCGCCTACCTGT
GCGGCATCCACAGGGAGCTGGTGAGGGCGGCTGAACGC
CGTGCTGCTGCCAACATCCACACCCTGTTGACATGA
GCGCCGAGGACTTCGACGCCATCATCGCCGAGCACTT
CCAGCCCGGCGACTGCGTGCTGGAGACCGACATCGCC
AGCTTCGACAAGAGCGAGGATGACGCTATGGCCCTGA
CCGCTCTGATGATCCTGGAGGACCTGGGCGTGGACGC
CGAGCTGCTCACCTGATCGAGGCTGCCTTCGGCGAG
ATCAGCTCCATCCACCTGCCACCAAGACCAAGTTCAA
GTTGCGCGCTATGATGAAAAGCGGAATGTTCCCTGACC
CTGTTGCTGAACACCGTGATCAACATTGTGATCGCCAG
CAGGGTGCTGCGGGAGAGGTGCTGACCGGCAGCCCCTGC
GCTGCCTTCATCGGCGACGACAACATCGTGAAGGGCG
TGAAAAGCGACAAGCTGATGGCCGACAGGTGCGCCAC
CTGGCTGAACATGGAGGTGAAGATCATCGACGCCGTG
GTGGGCGAAGAGGGCCCCCTACTTCTGCGGCGGATTCA
TCCTGTGCGACAGCGTGACCGGCACCGCCTGCAGGT
GGCCGACCCCTGAAGAGGCTGTTCAAGCTGGGCAAG
CCACTGGCCGCTGACGATGAGCACGACGATGACAGGC
GGAGGGCCCTGCACGAGGAAAGCACCAGGTGGAACA

GGGTGGGCATCCTGAGCGAGCTGTGCAAGGCCGTGGA
GAGCAGGTACGAGACCGTGGGCACCAGCATCATCGTG
ATGGCTATGACCACACTGGCCAGCTCCGTCAAGAGCTT
CTCCTACCTGAGGGGGGCCCTATAACTCTCTACGGCT AAA Comparativo non- aminoacido
MPEKVHVDIEEDSPFLRALQRSFPQFEVEAKQVTDNDHA (ID SEQ strutturale
NARAFSHLASKLIETEVDPSDITLDIGSAPARMYSKHKY N.: 55) gene ORF
HCICPMRCAEDPDRLYKYATKLKKNCKEITDKELDKKM
KELAAVMSDPDLETETMCLHDDDESCRYEGQVAVYQDV
YAVDGPTSLYHQANKGVRVAYWIGFDTPPFMFKNLAGA
YPSYSTNWADETVLTARNIGLCSDDVMERSRRGMSILRK
KYLKPSNNVLFVSGSTIYHEKRDLLRSWHLPSVFHLRGK
QNYTCRCETIVSCDGYVVKRIAISPGLYGKPSGYAATMH
REGFLCCKVTDLTNGERVSFPVCTYVPATLCDQMTGILA
TDVSADDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLP
VVAQAFARWAKEYKEDQEDERPLGLRDRQLVMGCCWA
FRRHKITSIYKRPDTQTIKVNDFHFSVFLPRIGSNTLEIGL
RTRIRKMLEEHKEPSPLITAEDVQEAKCAADEAKEVREA
EELRAALPPLAADVEEPTLEADVLDMLQEAGAGSVETPR
GLIKVTSYDGEDKIGSYAVLSPQAVLKSEKLSCHPLAEQ
VIVITHSRKGRYAVEPYHGKVVVPEGHAIQDFQALS
ESATIVYNEREFVNRYLHHIATHGGALNTDEEYKTVKP
SEHDGEYLYDIDRKQCVKKELVTGLGLTGELVDPFHEF
AYESLRTRPAAPYQVPTIGVYGVPGSGKSGIISAVTKKD
LVVSAKKENCAEIIRDVKKMKGLDVNARTVDSVLLNGC
KHPVETLYIDEAFACHAGTLRALIAIRPKKAVLCGDPKQ
CGFFNMMCLKVHFNHEICTQVFHKSISRCTKSVTSVVS
TLFYDKKMRTTNPKETKIVIDTTGSTKPKQDDLILTCFRG QUANDO SIETE INCONTRATI
NPLYAPTSEHVNVLLTRTEDRIVWKTLAGDPWIKTLTAK
YPGNFTATIEEWQAEHDAIMRHILERPDPTDVFQNKANV
CWAKALVPVLKTAGIDMTTEQWNTVDYFETDKAHS AEI
VLNQLCVRFFGLDLSGLFSAPTVP LSIRNNHWDNSPSN
MYGLNKEVVRQLSRRYPQLPRAVATGRVYDMNTGTLR
NYDPRINLVPVNRRLPHALVLHHNEHPQSDFFSVSKLK
GRTVLSVGEKLSVPGKMVDWLSDRPEATFRARLDL GIP
GDVPKYDIIFVNVRTPYKYHHYQQCEDHAIKLSMLTKKA
CLHLNPGGTCVSI GYGADRASESII GAIARLFKFSRVCKP
KSSLEETEVLVFFIGYDRKARTHNPYKLSSTLTNIYTG SRL
HEAGCAPSYHVVRGDIATATEGVIINAANSKGP GGGVC
GALYKFPESFDLQPIEVGKARLVKGAAKHIIH AVGPNF
NKVSEVEGDKQLAEAYESIAKIVNDNNYKSVAIPLLSTGI
FSGNKDRLTQSLNHLALTALD TTDADVAIYCRDKK WEMT
LKEAVARREAVEEICISDDSSVTEPDAELVRVHPKSS LAG
RKGYSTSDGKTFSYLEGTFKHQA AKDIAEINAMWPVATE
ANEQVCMYILGESMSSIRSKCPVEESEASTPPSTLPCLCIH

AMTPERVQRLKASRPEQITVCSFPLPKYRITGVQKIQCS
 QPILFSPKVPAYIHPKYLVEVETPPVDETPEPSAENQSTEGT
 PEQPPLITEDETRRRTPPEPINIEEEEEDSISLLSDGPTHQVLQ
 VEADIHGPPSVSSSSWSIPHASDFDVSLSILDTELEGASVT
 SGATSAETNSYFAKSMEFLARPVPAPRTVFRNPPHPAPRT
 RTPSLAPSRACSRVSTPPGVNRVITREELEALTPSRTP
 SRSVSRVSLVSNPPGVNRVITREEFEAFVAQQRRFDAGA
 YIFSSDTGQGHLQKQSVRQTVLSEVVLERTELEISYAPRL
 DQEKEELLRKKLQLNPTPANRSRYQSRKVENMKAITARR
 ILQGLGHYLAEGKVECYRTLHPVPLYSSSVNRAFSSPK
 VAVEACNAMLKENFPTVASYCIPEYDAYLDMVDGASC
 CLDTASFCPAKLRSPKHSYLEPTIRSAVPSAIQNTLQNV
 LAAATKRNCNVTQMRELPVLDAAFNVECFKKYACNNE
 YWETFKENPIRLTEENVVNYITKLGPKAAALFAKTHNL
 NMLQDIPMDRFVMDLKRDKVTPGKHTTEERPQVQVIQ
 AADPLATAYLCGIHRELVRRLLNAVLLPNIHTLFDMSAED
 FDAIIAEHFQPGDCVLETDIASFDKSEDDAMALTALMILE
 DLGVDAELLTLIEAAFGISSIHLPTKTKFKFGAMMKSGM
 FLTLFVNTVINVIASRVLRERLTGSPCAAFIGDDNIVKGV
 KSDKLMADRCATWLNMEVKIIDAVVGEKAPYFCGGFIL
 CDSVTGTACRVADPLKRLFGLGKPLAADDEHDDRRRA
 LHEESTRWNRVILSELCKAVESRYETVGTSHVMAMTTL ASSVKSFSYLRGAPITLYG* Comparativo
 intergenico nucleotide sconosciuto regione Comparativo 3' UTR nucleotide sconosciuto

TABELLA 5 ORF del peptide di interesse per RNA autoreplicanti della divulgazione ORF Sequenza Identità Tipo Sequenza nCoV-2019 nucleotide

ATGTTTGTTCCTTGTTCCTTATTGCCACTAGTCTCTAGTCAGTGTGT Gene Spike
 TAATCTTACAACCAGAACTCAATTACCCCTGCATACACTAATTC (ID SEQ
 TTTCACACGTGGTGTTCCTTATTACCCTGACAAAAGTTTTTCAGATCCTCA N.: 117)
 GTTTTACATTCAACTCAGGACTTGTTCCTTACCTTTCTTTTCCCAATGT
 TACTTGGTTCATGCTATACATGTCTCTGGGACCAATGGTACTAA
 GAGGTTTGATAACCCTGTCTTACCATTAAATGATGGTGTTCATTTT
 GCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGT
 ATTOATTOTTAGATTCGAAGACCCAGTCCCTATTOTTATTGTTAATAAC
 GCTACTAATGTTGTTATTAAAGTCTGTGAATTTCAATTTTGTAATG
 ATCCATTTTTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGA
 TGAAAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTT
 TTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAAC
 AGGGTAATTTCAAAAATCTTAGGGAATTTGTGTTAAGAATATTG
 ATGTTATTTTTAAAATATATTCTAAGCACACGCCTATTAATTTAGT
 GCGTGATTCCCTCAGGGTTTTTCGGCTTTAGAACCATTGGTAGAT
 TTGCCAATAGGTATTAACATCACTAGGTTTCAAACCTTACTTGCTT
 TACATAGAAGTTATTTGACTCCTGGTGATTCTTCTCAGGTTGGAC

AGCTGGTGCTGCAGCTTATTATGTGGGTATCTTCAACCTAGGAC
TTTTTATTAATAATATAATGAAAATGGAACCATTACAGATGCTGT
AGACTGTGCACTTGACCCTCTCTCAGAAACAAAGTGTACGTTGAA
ATCCTTCACTGTAGAAAAAGGAATCTATCAAACCTCTAACTTTAG
AGTCCAACCAACAGAATCTATTGTTAGATTTCCCTAATATTACAAA
CTTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTTGCATCT
GTTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGAT
TATTCTGTCCTATATAATTCCGCATCATTTTCCACTTTTAAGTGT
ATGGAGTGTCTCTACTAAATTAATGATCTCTGCTTTACTAATGT
CTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAAT
CGCTCCAGGGCAAACCTGGAAAGATTGCTGATTATAATTATAAATT
ACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATCTAACAA
TCTTGATTCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTG
TTAGGAAGTCTAATCTCAAACCTTTTGAGAGAGATATTTCAACT
GAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTGAAGGT
TTAATTGTTACTTTCTTTACAATCATATGTTTCCAACCCACTA
ATGGTGTGGTTACCAACCATACAGAGTAGTAGTACTTTCTTTTG
AACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTA
CTAATTTGGTTAAAAACAAATGTGTCAATTTCAACTTCAATGGTTT
AACAGGCACAGGTGTTCTTACTGAGTCTAACAAAAAGTTTCTGCC
TTTCCAACAATTTGGCAGAGACATTGCTGACACTACTGATGCTGT
CCGTGATCCACAGACACTTGAGATTCCTGACATTACACCATGTT
TTTTGGTGGTGTGAGTGTATAACACCAGGAACAAATACTTCTAA
CCAGGTTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCC
TGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTAT
TCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTAATAG
GGGCTGAACATGTCAAACAACTCATATGAGTGTGACATACCCATTG
GTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCTC
GGCGGGCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTA
TGTCACTTGGTGCAGAAAATTCAGTTGCTTACTCTAATAACTCTAT
TGCCATACCCACAAAATTTATTOATTAGTGTACCACAGAAATCT
ACCAGTGTCTATGACCAAGACATCAGTAGATTGTACAATGTACAT
TTGTGGTGATTCAACTGAATGCAGCAATCTTTTGTGCAATATGG
CAGTTTTTGTACACAATTAACCGTGCTTTAACTGGAATAGCTGTT
GAACAAGACAAAAACACCAAGAAGTTTTTGCACAAGTCAAACA
AATTTACAAAAACACCACCAATTAAGATTTTGGTGGTTTTAATTTT
TCACAAATATTACCAGATCCATCAAACCAAGCAAGAGGTCATTT
ATTGAAGATCTACTTTTCAACAAAGTGACACTTGACAGATGCTGGC
TTCATCAAACAATATGGTGAATGCCTTGGTGATATTGCTGCTAGA
GACCTCATTTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCAC
CTTTGCTCACAGATGAAATGATTGCTCAATACACTTCTGCACTGTT
AGCGGGTACAATCATTCTGTTGGACCTTGGTGCAGGTGCTGC
ATTACAAATACCATTTGCTATGCAAATGGCTTATAGGTTAATGG
TATTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAAATTGAT

TGCCAACCAATTTAATAGTGCTATTGGCAAATTC AAGACTCACT
TTCTTCCACAGCAAGTGC ACTTGGAAA ACTTCAAGATGTGGTCAA
CCAAAATGCACAAGCTTTAAACACGCTTGTTAAACA ACTTAGCTC
CAATTTTGGTGCAATTTCAAGTGT TTTAAATGATATCCTTTACGT
CTTGACAAAGTTGAGGCTGAAGTGCAAATTGATAGGTTGATCACA
GGCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATT
AGAGCTGCAGAAATCAGAGCTTCTGCTAATCTTGCTGCTACTAAA
ATGTCAGAGTGTGTACTTGGACAATCAAAAAGAGTTGATTTTTGT
GGAAAGGGCTATCATCTTATGTCTTCCCTCAGTCAGCACCTCAT
GGTGTAGTCTTTCTTGCATGTGACTTATGTCCCTGCACAAGAAAAG
AACTTCACA ACTGCTCCTGCCATTTGTCATGATGGAAAAGCACAC
TTTCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTG
TAACACAAAGGAATTITTATGAACCACAAATCATTACTACAGACA
ACACATTTGTGTCTGGTAACTGTGATGTTGTAATAGGAATTGTCA
ACAACACAGTTTATGATCCTTTGCAACCTGAATTAGACTCATTCA
AGGAGGAGTTAGATAAATATTTTAAGAATCATA CATCACCAGATG
TTGATTTAGGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAACAT
TCAAAAAGAAATTGACCGCCTCAATGAGGTTGCCAAGAATTTAA
ATGAATCTCTCATCGATCTCCAAGA ACTTGGAAAGTATGAGCAGT
ATATAAAATGGCCATGGTACATTTGGCTAGGTTTATAGCTGGCT
TGATTGCCATAGTAATGGTGACAATTATGCTTTGCTGTATGACCA
GTTGCTGTAGTTGCTCAAGGGCTGTTGTTCTTGTGGATCCTGCTG
CAAATTTGATGAAGACGACTCTGAGCCAGTGCTCAAAGGAGTCA AATTACATTACACATAA nCoV-2019
aminoacido MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSV Gene Spike
LHSTQDLFLPFFSNVTWFHAIHVS GTNGTKRFDNPVLPFNDGVYFAS (ID SEQ
TEKSNIIRGWIFGTTLD SKTQSL LIVN NATNVVIKVCE FQFCNDPFLG N.: 118)
VYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFK
NLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITR
FQTLALHRSYLTPGDSSSGW TAGAAAYYVGYLQPRTFLLKYNENG
TITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNI
TNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKC
YGVSPTKLNDLCFTNVYADSFVIRGDEV RQIAPGQTGKIADYNYKLP
DDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKP FERDISTEIY
QAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQP YRVVVL SFELLH
APATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQF
GRDIADTTDAVRDPQ TLEILDITPCSF GGVS VITPGTNTSNQVAVLYQ
DVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNS
YECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAY
SNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQY
GSFCTQLNRALTGIAVEQDKNTQE VFAQVKQIYKTPPIKDFGGFNFS
QILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICA
QKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFA
MQMAYRFNGIGVTQNVLYENQKLIANQFN SAIGKIQDSL SSTASALG
KLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQI

CGCAGATACCACAGACGCCGTGCGCGACCCACAGAACCCTGGAGA
TCCTGGACATCACACCTGCTCTTTCGGCGCGTGAGCGTGATCA
CACCCGGCACCAATAACAAGCAACCAGGTGGCCGTGCTGTATCAG
GACGTGAATTGTACCGAGGTGCCCGTGCTATCCACGCCGATCAG
CTGACCCCAACATGGCGGGTGTACAGCACCGGCTCCAACGTCTTC
CAGACAAGAGCCGGATGCCTGATCGGAGCAGAGCACGTGAACAA
TTCCTATGAGTGCGACATCCAATCGGCGCCGGCATCTGTGCCTC
TTACCAGACCCAGACAAACTCTCCAGAAGAGCCCGGAGCGTGG
CCTCCCAGTCTATCATCGCCTATAACCATGTCCCTTGGGCGCCGAGA
ACAGCGTGGCCTACTCTAACAATAGCATCGCCATCCAACCAACT
TCACAATCTCTGTGACCACAGAGATCCTGCCCGTGTCCATGACCA
AGACATCTGTGGACTGCACAAATGTATATCTGTGCGATTCTACCG
AGTGCAGCAACCTGCTGCTCCAGTACGGCAGCTTTTGTACCCAGC
TGAATAGAGCCCTGACAGGCATCGCCGTGGAGCAGGATAAGAAC
ACACAGGAGGTGTTGCGCCAGGTGAAGCAAATCTACAAGACCCC
CCCTATCAAGGACTTTGGCGGCTTCAATTTTTCCAGATCCTCGCCT
GATCCATCCAAGCCTTCTAAGCGGAGCTTTATCGAGGACCTGCTG
TTCAACAAGGTGACCCTGGCCGATGCCGGCTTCATCAAGCAGTAT
GGCGATTGCTGGGCGACATCGCAGCCAGGGACCTGATCTGCGCC
CAGAAGTTTAATGGCCTGACCGTGCTGCCACCCCTGCTGACAGAT
GAGATGATCGCACAGTACACAAGCGCCCTGCTGGCCGGCACCAT
CACATCCGGATGGACCTTCGGCGCAGGAGCCGCCCTCCAGATCCC
CTTTGCCATGCAGATGGCCTATAGGTTCAACGGCATCGGCGTGAC
CCAGAATGTGCTGTACGAGAACCAGAAGCTGATCGCCAATCAGTT
TAACTCCGCCATCGGCAAGATCCAGGACAGCCTGTCCTCTACAGC
CAGCGCCCTGGGCAAGCTCCAGGATGTGGTGAATCAGAACGCCC
AGGCCCTGAATACCCTGGTGAAGCAGCTGAGCAGCAACTTCGGC
GCCATCTCTAGCGTGCTGAATGACATCCTGAGCCGGCTGGACAAG
GTGGAGGCAGGTGCAGATCGACCGGCTGATCACCGCCGGCT
CCAGAGCCTCCAGACCTATGTGACACAGCAGCTGATCAGGGCCG
CCGAGATCAGGGCCAGCGCAATCTGGCAGCAACCAAGATGTCC
GAGTGCCTGCTGGGCCAGTCTAAGAGTGGACTTTTGTGGCAAG
GGCTATCACCTGATGTCCTCCCTCAGTCTGCCCCACACGGCGTG
GTGTTTTCTGCACGTGACCTACGTGCCCCGCCAGGAGAAGAACTTC
ACCACAGCCCCTGCCATCTGCCACGATGGCAAGGCCCACTTTCCA
AGGGAGGGCGTGTTCTGTGCCAACGGCACCCACTGGTTTGTACA
CAGCGCAATTTCTACGAGCCCAGATCATCACCACAGACAACACC
TTCGTGAGCGGCAACTGTGACGTGGTCATCGGCATCGTGAACAAT
ACCGTGTATGATCCACTCCAGCCCAGCTGGACAGCTTTAAGGAG
GAGCTGGATAAGTATTTCAAGAATCACACCTCCCCTGACGTGAT
CTGGGCGACATCAGCGGCATCAATGCCTCCGTGGTGAACATCCAG
AAGGAGATCGACCGCCTGAACGAGGTGGCTAAGAATCTGAACGA
GAGCCTGATCGACCTCCAGGAGCTGGGCAAGTATGAGCAGTACA
TCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGCCTGA

TCGCCATCGTGATGGTGACCATCATGCTGTGCTGTATGACATCCT
GCTGTTCTTGCCTGAAGGGCTGCTGTAGCTGTGGCTCCTGCTGTA
AGTTTGACGAGGATGACTCTGAACCTGTGCTGAAGGGCGTGAAG CTGCATTACACCTAA nCoV-2019
aminoacido MFVFLVLLPLVSSQCVNLTRTQLPPAYTNSFTRGVYYPDKVFRSSV Gene Spike
LHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFAS (ID SEQ
TEKSNIIRGWIFGTLLDSTKQSLIVN NATNVVIVKVFCEFCNDPFLG N.: 120)
VYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFK
NLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITR
FQTLALHRSYLTPGDSSSGWTAGAAAYVGYLQPRTFLLKYNENG
TITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNI
TNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKC
YGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLP
DDFTGCVIAWNSNLDKVGGNYNLYRLFRKSNLKPFERDISTEYIY
QAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLH
APATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQF
GRDIADTTDAVRDPQTLLEILDITPCSFGGVSVITPGTNTSNQVAVLYQ
DVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNS
YECDIPIGAGICASYQTQTNPRRARSVASQSIIAYTMSLGAENSVAY
SNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECNLLLQY
GSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFS
QILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICA
QKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFA
MQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSTASALG
KLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQI
DRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRV
DFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNF TAPAICHGDK
AHFPREGVVFVSNGTHWFVTQRNFYEPQUITTDNTFVSGNCDVVIGIVN
NTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKE
IDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVM
VTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT*

Le sequenze di RNA possono includere qualsiasi combinazione delle sequenze di RNA elencate nelle Tabelle 4 e 5. In alcune realizzazioni, le sequenze di RNA della presente divulgazione includono qualsiasi combinazione delle sequenze di RNA elencate nelle Tabelle 4 e 5 in cui dallo 0% al 100%, dall'1% al 100%, dal 25% al 100%, dal 50% al 100% e dal 75% al 100% dei nucleotidi uracili delle sequenze di mRNA sono modificati. In alcune realizzazioni, dall'1% al 100% dei nucleotidi uracili sono N1-metilpseudouridina o 5-metossiuridina. In alcune realizzazioni, il 100% dei nucleotidi uracili sono N1-metilpseudouridina. In alcune realizzazioni, il 100% dei nucleotidi uracili sono 5-metossiuridina.

Un RNA autoreplicante della divulgazione può essere ottenuto con qualsiasi mezzo idoneo. I metodi per la fabbricazione di RNA autoreplicante sono noti nell'arte e sarebbero facilmente evidenti a una persona con abilità ordinaria. Un RNA autoreplicante della divulgazione può essere

preparato secondo qualsiasi tecnica disponibile, tra cui, ma non limitatamente a, sintesi chimica, trascrizione in vitro (IVT) o scissione enzimatica o chimica di un precursore più lungo, ecc.

In alcune realizzazioni, un RNA autoreplicante della divulgazione viene prodotto da un costrutto primario di DNA complementare (cDNA). I costrutti di cDNA possono essere prodotti su uno stampo di RNA mediante l'azione di una trascrittasi inversa (ad esempio, DNA polimerasi RNA-dipendente). Il processo di progettazione e sintesi dei costrutti primari di cDNA qui descritti generalmente include le fasi di costruzione genica, produzione di RNA (con o senza modifiche) e purificazione. Nel metodo IVT, una sequenza polinucleotidica target che codifica un RNA autoreplicante della divulgazione viene prima selezionata per l'incorporazione in un vettore che verrà amplificato per produrre uno stampo di cDNA. Facoltativamente, la sequenza polinucleotidica target e/o qualsiasi sequenza fiancheggiante possono essere ottimizzate per codone. Lo stampo di cDNA viene quindi utilizzato per produrre un RNA autoreplicante della divulgazione tramite trascrizione in vitro (IVT). Dopo la produzione, l'RNA autoreplicante della divulgazione può essere sottoposto a processi di purificazione e pulizia. I cui passaggi sono forniti più in dettaglio di seguito.

La fase di costruzione genica può includere, ma non è limitata a, sintesi genica, amplificazione vettoriale, purificazione del plasmide, linearizzazione e pulizia del plasmide e sintesi e pulizia del modello di cDNA. Una volta selezionata una proteina di interesse per la produzione, viene progettato un costrutto primario. All'interno del costrutto primario, una prima regione di nucleosidi collegati che codificano il polipeptide di interesse può essere costruita utilizzando un open reading frame (ORF) di un trascritto di acido nucleico (DNA o RNA) selezionato. L'ORF può comprendere l'ORF di tipo selvatico, un'isoforma, una variante o un frammento di esso. Come utilizzato nel presente documento, un "open reading frame" o "ORF" è inteso per fare riferimento a una sequenza di acido nucleico (DNA o RNA) che è in grado di codificare un polipeptide di interesse. Gli ORF spesso iniziano con il codone di inizio, ATG e terminano con un codone o segnale senza senso o di terminazione.

I modelli di cDNA possono essere trascritti per produrre un RNA autoreplicante della divulgazione utilizzando un sistema di trascrizione in vitro (IVT). Il sistema comprende in genere un tampone di trascrizione, nucleotidi trifosfati (NTP), un inibitore della RNasi e una polimerasi. Gli NTP possono essere selezionati tra, ma non sono limitati a, quelli descritti nel presente documento, inclusi NTP naturali e non naturali (modificati). La polimerasi può essere selezionata tra, ma non è limitata a, T7 RNA polimerasi, T3 RNA polimerasi e polimerasi mutanti come, ma non sono limitate a, polimerasi in grado di incorporare acidi nucleici modificati.

Il modello primario di cDNA o la sequenza di RNA trascritta possono anche subire reazioni di capping e/o tailing. Una reazione di capping può essere eseguita con metodi noti nell'arte per aggiungere un cappuccio 5' all'estremità 5' del costrutto primario. I metodi per il capping includono, ma non sono limitati a, l'uso di un enzima Vaccinia Capping (New England Biolabs, Ipswich, Mass.) o il capping all'inizio della trascrizione in vitro, ad esempio, includendo un agente di capping come parte della reazione IVT. (Nuc. Acids Symp. (2009) 53:129). Una reazione di tailing di poli(A) può essere eseguita con metodi noti nell'arte, come, ma non limitati a, 2' O-metiltransferasi e con metodi come descritto nel presente documento. Se il costrutto primario

generato da cDNA non include un poli-T, può essere utile eseguire la reazione di tailing di poli(A) prima che il costrutto primario venga pulito.

I costrutti di cDNA ottimizzati per codone che codificano le proteine non strutturali e il transgene per una proteina di RNA autoreplicante sono particolarmente adatti per generare sequenze di RNA autoreplicante descritte nel presente documento. Ad esempio, tali costrutti di cDNA possono essere utilizzati come base per trascrivere, in vitro, un poliribonucleotide che codifica una proteina di interesse come parte di un RNA autoreplicante.

La presente divulgazione fornisce anche vettori di espressione comprendenti una sequenza nucleotidica che codifica un RNA autoreplicante che è preferibilmente collegato operativamente ad almeno una sequenza regolatrice. Le sequenze regolatrici sono riconosciute dall'arte e sono selezionate per dirigere l'espressione del polipeptide codificato.

Di conseguenza, il termine sequenza regolatrice include promotori, potenziatori e altri elementi di controllo dell'espressione. La progettazione del vettore di espressione può dipendere da fattori quali la scelta della cellula ospite da trasformare e/o il tipo di proteina che si desidera esprimere.

La presente divulgazione fornisce anche polinucleotidi (ad esempio DNA, RNA, cDNA, mRNA, ecc.) diretti a un RNA autoreplicante della divulgazione che può essere collegato operativamente a una o più sequenze nucleotidiche regolatrici in un costrutto di espressione, come un vettore o un plasmide. In alcune forme di realizzazione, tali costrutti sono costrutti di DNA. Le sequenze nucleotidiche regolatrici saranno generalmente appropriate per una cellula ospite utilizzata per l'espressione. Numerosi tipi di vettori di espressione appropriati e sequenze regolatrici adatte sono noti nell'arte per una varietà di cellule ospiti.

In genere, dette una o più sequenze nucleotidiche regolatrici possono includere, ma non sono limitate a, sequenze promotore, sequenze leader o segnale, siti di legame ribosomiale, sequenze di inizio e fine trascrizionali, sequenze di inizio e fine traduzionali e sequenze potenziatrici o attivatrici. I promotori costitutivi o inducibili come noti nell'arte sono contemplati dalle forme di realizzazione della presente divulgazione. I promotori possono essere promotori naturali o promotori ibridi che combinano elementi di più di un promotore.

Un costrutto di espressione può essere presente in una cellula su un episoma, come un plasmide, oppure il costrutto di espressione può essere inserito in un cromosoma. In alcune realizzazioni, il vettore di espressione contiene un gene marcatore selezionabile per consentire la selezione di cellule ospiti trasformate. I geni marcatori selezionabili sono ben noti nell'arte e varieranno a seconda della cellula ospite utilizzata.

La presente divulgazione fornisce anche una cellula ospite transfettata con un RNA o DNA autoreplicante descritto nel presente documento. L'RNA o DNA autoreplicante può codificare qualsiasi proteina di interesse del coronavirus, ad esempio un antigene, incluso l'antigene S del virus COVID-19. La cellula ospite può essere qualsiasi cellula procariotica o eucariotica. Ad esempio, un polipeptide codificato da un RNA autoreplicante può essere espresso in cellule batteriche come *E. coli*, cellule di insetti (ad esempio, utilizzando un sistema di espressione di

baculovirus), lievito o cellule di mammiferi. Altre cellule ospiti idonee sono note agli esperti del settore.

Una cellula ospite transfettata con un vettore di espressione comprendente un RNA autoreplicante della divulgazione può essere coltivata in condizioni appropriate per consentire l'espressione dell'amplificazione dell'RNA autoreplicante e la traduzione del polipeptide. Il polipeptide può essere secreto e isolato da una miscela di cellule e terreno contenente i polipeptidi. In alternativa, i polipeptidi possono essere trattenuti nel citoplasma o in una frazione di membrana e le cellule raccolte, lisate e la proteina isolata. Una coltura cellulare include cellule ospiti, terreni e altri sottoprodotti. I terreni adatti per la coltura cellulare sono ben noti nell'arte.

Le proteine espresse descritte nel presente documento possono essere isolate dal mezzo di coltura cellulare, dalle cellule ospiti o da entrambi utilizzando tecniche note nel settore per la purificazione delle proteine, tra cui la cromatografia a scambio ionico, la cromatografia a filtrazione su gel, l'ultrafiltrazione, l'elettroforesi e la purificazione per immunoaffinità con anticorpi specifici per particolari epitopi del polipeptide.

Composizioni e composizioni farmaceutiche

In alcune forme di realizzazione, nel presente documento sono fornite composizioni comprendenti una qualsiasi delle molecole di acido nucleico qui fornite. Le composizioni qui fornite possono includere un lipide. Qualsiasi lipide può essere incluso nelle composizioni qui fornite. In un aspetto, il lipide è un lipide cationico ionizzabile. Qualsiasi lipide cationico ionizzabile può essere incluso nelle composizioni comprendenti molecole di acido nucleico qui fornite.

Le composizioni e i polinucleotidi della presente informativa possono essere utilizzati per immunizzare o vaccinare un soggetto contro un'infezione virale. In alcune realizzazioni, le composizioni e i polinucleotidi della presente informativa possono essere utilizzati per vaccinare o immunizzare un soggetto contro il virus COVID-19.

Sono inoltre fornite nel presente documento, in alcune forme di realizzazione, composizioni farmaceutiche comprendenti una qualsiasi delle molecole di acido nucleico fornite nel presente documento e una formulazione lipidica. Qualsiasi lipide può essere incluso nelle formulazioni lipidiche delle composizioni farmaceutiche fornite nel presente documento. In un aspetto, le formulazioni lipidiche delle composizioni farmaceutiche fornite nel presente documento includono un lipide cationico ionizzabile. Lipidi cationici ionizzabili esemplari di composizioni e composizioni farmaceutiche fornite nel presente documento includono quanto segue:

In un aspetto, il lipide cationico ionizzabile delle composizioni fornite nel presente documento ha una struttura di

o un sale farmaceuticamente accettabile dello stesso.

In un altro aspetto, il lipide cationico ionizzabile delle composizioni fornite nel presente documento ha una struttura di

o un sale farmaceuticamente accettabile dello stesso.

In un aspetto, il lipide cationico ionizzabile incluso nelle formulazioni lipidiche delle composizioni farmaceutiche fornite nel presente documento ha una struttura di

o un sale farmaceuticamente accettabile dello stesso.

In un altro aspetto, il lipide cationico ionizzabile incluso nelle formulazioni lipidiche delle composizioni farmaceutiche fornite nel presente documento ha una struttura di

o un sale farmaceuticamente accettabile dello stesso.

Formulazioni lipidiche/LNP

Le terapie basate sulla somministrazione intracellulare di acidi nucleici alle cellule bersaglio affrontano barriere sia extracellulari che intracellulari. Infatti, i materiali di acidi nucleici nudi non possono essere facilmente somministrati a livello sistemico a causa della loro tossicità, bassa stabilità nel siero, rapida clearance renale, ridotto assorbimento da parte delle cellule bersaglio, assorbimento da parte dei fagociti e la loro capacità di attivare la risposta immunitaria, tutte caratteristiche che ne precludono lo sviluppo clinico. Quando il materiale di acidi nucleici esogeni (ad esempio, mRNA) entra nel sistema biologico umano, viene riconosciuto dal sistema reticoloendoteliale (RES) come patogeno estraneo ed eliminato dalla circolazione sanguigna prima di avere la possibilità di incontrare le cellule bersaglio all'interno o all'esterno del sistema vascolare. È stato riportato che l'emivita dell'acido nucleico nudo nel flusso sanguigno è di circa diversi minuti (Kawabata K, Takakura Y, Hashida MPharm Res. 1995 giugno; 12(6):825-30). La modifica chimica e un metodo di somministrazione appropriato possono ridurre l'assorbimento da parte del RES e proteggere gli acidi nucleici dalla degradazione da parte delle nucleasi ubiquitarie, che aumentano la stabilità e l'efficacia delle terapie basate sugli acidi nucleici. Inoltre, gli RNA o i DNA sono polimeri idrofili anionici che non sono favorevoli all'assorbimento da parte delle cellule, che sono anche anioniche in superficie. Il successo delle terapie basate sugli acidi nucleici dipende quindi in larga parte dallo sviluppo di veicoli o vettori in grado di fornire in modo efficiente ed efficace materiale genetico alle cellule bersaglio e ottenere livelli sufficienti di espressione in vivo con una tossicità minima.

Inoltre, una volta internalizzati in una cellula bersaglio, i vettori di somministrazione di acidi nucleici sono sfidati da barriere intracellulari, tra cui l'intrappolamento dell'endosoma, la degradazione lisosomiale, lo spaccettamento degli acidi nucleici dai vettori, la traslocazione attraverso la membrana nucleare (per il DNA), il rilascio nel citoplasma (per l'RNA) e così via. Il successo della terapia basata sugli acidi nucleici dipende quindi dalla capacità del vettore di somministrare gli acidi nucleici ai siti bersaglio all'interno delle cellule per ottenere livelli sufficienti di un'attività desiderata come l'espressione di un gene.

Mentre diverse terapie geniche sono state in grado di utilizzare con successo un vettore di somministrazione virale (ad esempio, AAV), le formulazioni a base di lipidi sono state sempre più

riconosciute come uno dei sistemi di somministrazione più promettenti per RNA e altri composti di acidi nucleici grazie alla loro biocompatibilità e alla loro facilità di produzione su larga scala. Uno dei progressi più significativi nelle terapie a base di acidi nucleici a base di lipidi si è verificato nell'agosto 2018, quando Patisiran (ALN-TTR02) è stato il primo terapeutico siRNA approvato dalla Food and Drug Administration (FDA) e dalla Commissione Europea (CE). ALN-TTR02 è una formulazione siRNA basata sulla cosiddetta tecnologia di transfettazione Stable Nucleic Acid Lipid Particle (SNALP). Nonostante il successo di Patisiran, la somministrazione di terapie a base di acidi nucleici, incluso mRNA, tramite formulazioni lipidiche è ancora in fase di sviluppo.

Alcuni veicoli di somministrazione formulati con lipidi riconosciuti dall'arte per la terapia degli acidi nucleici includono, secondo varie forme di realizzazione, vettori a base di polimeri, come la polietilenimmina (PEI), nanoparticelle lipidiche e liposomi, nanoliposomi, nanoliposomi contenenti ceramide, liposomi multivescicolari, proteoliposomi, esosomi sia naturali che derivati sinteticamente, corpi lamellari naturali, sintetici e semisintetici, nanoparticelle, micelle ed emulsioni. Queste formulazioni lipidiche possono variare nella loro struttura e composizione e, come ci si può aspettare in un campo in rapida evoluzione, sono stati utilizzati diversi termini nell'arte per descrivere un singolo tipo di veicolo di somministrazione. Allo stesso tempo, i termini per le formulazioni lipidiche sono variati per quanto riguarda il loro significato previsto in tutta la letteratura scientifica e questo uso incoerente ha causato confusione per quanto riguarda il significato esatto di diversi termini per le formulazioni lipidiche. Tra le diverse potenziali formulazioni lipidiche, i liposomi, i liposomi cationici e le nanoparticelle lipidiche sono descritti in dettaglio e definiti nel presente documento ai fini della presente informativa.

Liposomi

I liposomi convenzionali sono vescicole che consistono di almeno un doppio strato e un compartimento acquoso interno. Le membrane a doppio strato dei liposomi sono tipicamente formate da molecole anfifiliche, come lipidi di origine sintetica o naturale che comprendono domini idrofili e idrofobici spazialmente separati (Lasic, Trends Biotechnol., 16: 307-321, 1998). Le membrane a doppio strato dei liposomi possono anche essere formate da polimeri e tensioattivi anfifilici (ad esempio, polimerisomi, niosomi, ecc.). In genere si presentano come vescicole sferiche e possono variare in dimensioni da 20 nm a pochi micron. Le formulazioni liposomiali possono essere preparate come dispersione colloidale o possono essere liofilizzate per ridurre i rischi di stabilità e migliorare la durata di conservazione dei farmaci a base di liposomi. I metodi di preparazione delle composizioni liposomiali sono noti nell'arte e sarebbero alla portata di un comune artigiano.

I liposomi che hanno un solo doppio strato sono detti unilamellari, mentre quelli che ne hanno più di uno sono detti multilamellari. I tipi più comuni di liposomi sono le piccole vescicole unilamellari (SUV), le grandi vescicole unilamellari (LUV) e le vescicole multilamellari (MLV). A differenza dei liposomi, i lisosomi, le micelle e le micelle invertite sono composti da monostrati di lipidi. In genere, si pensa che un liposoma abbia un singolo compartimento interno, tuttavia alcune formulazioni possono essere liposomi multivescicolari (MVL), che consistono in numerosi compartimenti acquosi interni discontinui separati da diversi doppi strati lipidici non concentrici.

I liposomi sono stati a lungo considerati veicoli per la somministrazione di farmaci a causa della loro superiore biocompatibilità, dato che i liposomi sono fondamentalmente analoghi delle membrane biologiche e possono essere preparati sia da fosfolipidi naturali che sintetici (Int J Nanomedicine. 2014; 9:1833-1843). Nel loro utilizzo come veicoli per la somministrazione di farmaci, poiché un liposoma ha un nucleo di soluzione acquosa circondato da una membrana idrofobica, i soluti idrofili disciolti nel nucleo non possono passare facilmente attraverso il doppio strato e i composti idrofobici si assoceranno al doppio strato. Pertanto, un liposoma può essere caricato con molecole idrofobiche e/o idrofile. Quando un liposoma viene utilizzato per trasportare un acido nucleico come l'RNA, l'acido nucleico sarà contenuto all'interno del compartimento liposomiale in una fase acquosa.

Liposomi cationici

I liposomi possono essere composti da lipidi cationici, anionici e/o neutri. Come importante sottoclasse di liposomi, i liposomi cationici sono liposomi realizzati in tutto o in parte da lipidi caricati positivamente, o più specificamente da un lipide che comprende sia un gruppo cationico che una porzione lipofila. Oltre alle caratteristiche generali descritte sopra per i liposomi, le frazioni caricate positivamente dei lipidi cationici utilizzate nei liposomi cationici offrono diversi vantaggi e alcune caratteristiche strutturali uniche. Ad esempio, la porzione lipofila del lipide cationico è idrofobica e quindi si allontanerà dall'interno acquoso del liposoma e si assocerà ad altre specie non polari e idrofobiche. Al contrario, la frazione cationica si assocerà a mezzi acquosi e, cosa più importante, a molecole e specie polari con cui può complessarsi nell'interno acquoso del liposoma cationico. Per queste ragioni, i liposomi cationici sono sempre più oggetto di ricerca per l'uso nella terapia genica a causa della loro predisposizione verso gli acidi nucleici caricati negativamente tramite interazioni elettrostatiche, con conseguenti complessi che offrono biocompatibilità, bassa tossicità e la possibilità di una produzione su larga scala richiesta per applicazioni cliniche in vivo. I lipidi cationici adatti all'uso nei liposomi cationici sono elencati di seguito.

Nanoparticelle lipidiche

A differenza dei liposomi e dei liposomi cationici, le nanoparticelle lipidiche (LNP) hanno una struttura che include un singolo monostrato o doppio strato di lipidi che incapsula un composto in una fase solida. Pertanto, a differenza dei liposomi, le nanoparticelle lipidiche non hanno una fase acquosa o un'altra fase liquida al loro interno, ma piuttosto i lipidi del doppio strato o guscio monostrato sono direttamente complessati al composto interno incapsulandolo in un nucleo solido. Le nanoparticelle lipidiche sono in genere vescicole sferiche con una dispersione relativamente uniforme di forma e dimensione. Mentre le fonti variano su quale dimensione qualifichi una particella lipidica come nanoparticella, c'è una certa sovrapposizione nell'accordo sul fatto che una nanoparticella lipidica può avere un diametro compreso tra 10 nm e 1000 nm. Tuttavia, più comunemente sono considerate più piccole di 120 nm o persino 100 nm.

Per i sistemi di somministrazione di acidi nucleici con nanoparticelle lipidiche, il guscio lipidico è formulato per includere un lipide cationico ionizzabile che può complessarsi e associarsi con la struttura portante caricata negativamente del nucleo dell'acido nucleico. I lipidi cationici

ionizzabili con valori pKa apparenti inferiori a circa 7 hanno il vantaggio di fornire un lipide cationico per la complessazione con la struttura portante caricata negativamente dell'acido nucleico e il caricamento nella nanoparticella lipidica a valori di pH inferiori al pKa del lipide ionizzabile dove è caricato positivamente. Quindi, a valori di pH fisiologici, la nanoparticella lipidica può adottare un esterno relativamente neutro consentendo un aumento significativo delle emivite di circolazione delle particelle dopo la somministrazione endovenosa. Nel contesto della somministrazione di acidi nucleici, le nanoparticelle lipidiche offrono molti vantaggi rispetto ad altri sistemi di somministrazione di acidi nucleici basati sui lipidi, tra cui elevata efficienza di incapsulamento dell'acido nucleico, potente trasfezione, migliore penetrazione nei tessuti per la somministrazione di farmaci e bassi livelli di citotossicità e immunogenicità.

Prima dello sviluppo di sistemi di somministrazione di nanoparticelle lipidiche per acidi nucleici, i lipidi cationici erano ampiamente studiati come materiali sintetici per la somministrazione di medicinali a base di acidi nucleici. In questi primi sforzi, dopo la miscelazione a pH fisiologico, gli acidi nucleici venivano condensati dai lipidi cationici per formare complessi lipidi-acidi nucleici noti come lipoplessi. Tuttavia, i lipoplessi si sono dimostrati instabili e caratterizzati da ampie distribuzioni dimensionali che vanno dalla scala submicronica a pochi micron. I lipoplessi, come il reagente Lipofectamine®, hanno trovato notevole utilità per la trasfezione in vitro. Tuttavia, questi lipoplessi di prima generazione non si sono dimostrati utili in vivo. Le grandi dimensioni delle particelle e la carica positiva (impartita dal lipide cationico) determinano una rapida clearance plasmatica, tossicità emolitica e di altro tipo, nonché attivazione del sistema immunitario. In alcuni aspetti, le molecole di acido nucleico qui fornite e i lipidi o le formulazioni lipidiche qui fornite formano una nanoparticella lipidica (LNP).

Sotto altri aspetti, le molecole di acido nucleico qui fornite sono incorporate in una formulazione lipidica (vale a dire un veicolo di rilascio a base di lipidi).

Nel contesto della presente divulgazione, un veicolo di somministrazione a base lipidica serve in genere a trasportare un RNA desiderato a una cellula o tessuto bersaglio. Il veicolo di somministrazione a base lipidica può essere qualsiasi veicolo di somministrazione a base lipidica idoneo noto nell'arte. In alcuni aspetti, il veicolo di somministrazione a base lipidica è un liposoma, un liposoma cationico o una nanoparticella lipidica contenente un RNA autoreplicante della divulgazione. In alcuni aspetti, il veicolo di somministrazione a base lipidica comprende una nanoparticella o un doppio strato di molecole lipidiche e un RNA autoreplicante della divulgazione. In alcuni aspetti, il doppio strato lipidico comprende inoltre un lipide neutro o un polimero. In alcuni aspetti, la formulazione lipidica comprende un mezzo liquido. In alcuni aspetti, la formulazione incapsula inoltre un acido nucleico. In alcuni aspetti, la formulazione lipidica comprende inoltre un acido nucleico e un lipide neutro o un polimero. In alcuni aspetti, la formulazione lipidica incapsula l'acido nucleico.

La descrizione fornisce formulazioni lipidiche comprendenti una o più molecole di RNA autoreplicanti incapsulate all'interno della formulazione lipidica. In alcuni aspetti, la formulazione lipidica comprende liposomi. In alcuni aspetti, la formulazione lipidica comprende liposomi cationici. In alcuni aspetti, la formulazione lipidica comprende nanoparticelle lipidiche.

In alcuni aspetti, l'RNA autoreplicante è completamente incapsulato nella porzione lipidica della formulazione lipidica, in modo tale che l'RNA nella formulazione lipidica sia resistente in soluzione acquosa alla degradazione della nucleasi. In altri aspetti, le formulazioni lipidiche descritte nel presente documento sono sostanzialmente non tossiche per animali come gli esseri umani e altri mammiferi.

Le formulazioni lipidiche della presente informativa hanno inoltre in genere un rapporto totale lipidi:RNA (rapporto massa/massa) da circa 1:1 a circa 100:1, da circa 1:1 a circa 50:1, da circa 2:1 a circa 45:1, da circa 3:1 a circa 40:1, da circa 5:1 a circa 45:1, o da circa 10:1 a circa 40:1, o da circa 15:1 a circa 40:1, o da circa 20:1 a circa 40:1; o da circa 25:1 a circa 45:1; o da circa 30:1 a circa 45:1; o da circa 32:1 a circa 42:1; o da circa 34:1 a circa 42:1. In alcuni aspetti, il rapporto lipidi:RNA totale (rapporto massa/massa) è compreso tra circa 30:1 e circa 45:1. Il rapporto può essere qualsiasi valore o sottovalore all'interno degli intervalli indicati, compresi gli endpoint.

Le formulazioni lipidiche della presente informativa hanno in genere un diametro medio da circa 30 nm a circa 150 nm, da circa 40 nm a circa 150 nm, da circa 50 nm a circa 150 nm, da circa 60 nm a circa 130 nm, da circa 70 nm a circa 110 nm, da circa 70 nm a circa 100 nm, da circa 80 nm a circa 100 nm, da circa 90 nm a circa 100 nm, da circa 70 a circa 90 nm, da circa 80 nm a circa 90 nm, da circa 70 nm a circa 80 nm, o circa 30 nm, circa 35 nm, circa 40 nm, circa 45 nm, circa 50 nm, circa 55 nm, circa 60 nm, circa 65 nm, circa 70 nm, circa 75 nm, circa 80 nm, circa 85 nm, circa 90 nm, circa 95 nm, circa 100 nm, circa 105 nm, circa 110 nm, circa 115 nm, circa 120 nm, circa 125 nm, circa 130 nm, circa 135 nm, circa 140 nm, circa 145 nm o circa 150 nm, e sono sostanzialmente non tossici. Il diametro può essere qualsiasi valore o sottovalore entro gli intervalli indicati, compresi gli endpoint. Inoltre, gli acidi nucleici, quando presenti nelle nanoparticelle lipidiche della presente informativa, sono generalmente resistenti in soluzione acquosa alla degradazione con una nucleasi.

In alcune realizzazioni, la nanoparticella lipidica ha una dimensione inferiore a circa 500 nm, inferiore a circa 400 nm, inferiore a circa 300 nm, inferiore a circa 200 nm, inferiore a circa 100 nm o inferiore a circa 50 nm. In realizzazioni specifiche, la nanoparticella lipidica ha una dimensione compresa tra circa 55 nm e circa 90 nm.

In alcuni aspetti, le formulazioni lipidiche comprendono un RNA autoreplicante, un lipide cationico (ad esempio, uno o più lipidi cationici o sali degli stessi descritti nel presente documento), un fosfolipide e un lipide coniugato che inibisce l'aggregazione delle particelle (ad esempio, uno o più coniugati PEG-lipidi). Le formulazioni lipidiche possono anche includere colesterolo. In un aspetto, il lipide cationico è un lipide cationico ionizzabile.

Nelle formulazioni di acido nucleico-lipidi, l'RNA può essere completamente incapsulato nella porzione lipidica della formulazione, proteggendo così l'acido nucleico dalla degradazione della nucleasi. In alcuni aspetti, una formulazione lipidica comprendente un RNA è completamente incapsulata nella porzione lipidica della formulazione lipidica, proteggendo così l'acido nucleico dalla degradazione della nucleasi. In alcuni aspetti, l'RNA nella formulazione lipidica non è sostanzialmente degradato dopo l'esposizione della particella a una nucleasi a 37° C. per almeno 20, 30, 45 o 60 minuti. In alcuni altri aspetti, l'RNA nella formulazione lipidica non è sostanzialmente

degradato dopo l'incubazione della formulazione nel siero a 37° C. per almeno 30, 45 o 60 minuti o almeno 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 o 36 ore. In alcuni aspetti, l'RNA è complessato con la porzione lipidica della formulazione. Uno dei vantaggi delle formulazioni della presente divulgazione è che le composizioni di acido nucleico-lipidi sono sostanzialmente non tossiche per animali come gli esseri umani e altri mammiferi.

Nel contesto degli acidi nucleici, l'incapsulamento completo può essere determinato eseguendo un test di esclusione del colorante fluorescente impermeabile alla membrana, che utilizza un colorante che ha una fluorescenza migliorata quando associato all'acido nucleico.

L'incapsulamento è determinato aggiungendo il colorante a una formulazione lipidica, misurando la fluorescenza risultante e confrontandola con la fluorescenza osservata dopo l'aggiunta di una piccola quantità di detergente non ionico. La rottura dello strato lipidico mediata dal detergente rilascia l'acido nucleico incapsulato, consentendogli di interagire con il colorante impermeabile alla membrana. L'incapsulamento dell'acido nucleico può essere calcolato come $E=(I_0-I)/I_0$, dove I_0 e I si riferisce alle intensità di fluorescenza prima e dopo l'aggiunta del detergente.

In alcuni aspetti, la presente divulgazione fornisce una composizione di acido nucleico-lipide comprendente una pluralità di liposomi di acido nucleico, liposomi di acido nucleico-cationici o nanoparticelle di acido nucleico-lipide. In alcuni aspetti, la composizione di acido nucleico-lipide comprende una pluralità di liposomi di RNA. In alcuni aspetti, la composizione di acido nucleico-lipide comprende una pluralità di liposomi di RNA-cationici. In alcuni aspetti, la composizione di acido nucleico-lipide comprende una pluralità di nanoparticelle di RNA-lipide.

In alcuni aspetti, le formulazioni lipidiche comprendono RNA che è completamente incapsulato all'interno della porzione lipidica della formulazione, in modo tale che da circa il 30% a circa il 100%, da circa il 40% a circa il 100%, da circa il 50% a circa il 100%, da circa il 60% a circa il 100%, da circa il 70% a circa il 100%, da circa l'80% a circa il 100%, da circa il 90% a circa il 100%, da circa il 30% a circa il 95%, da circa il 40% a circa il 95%, da circa il 50% a circa il 95%, da circa il 60% a circa il 95%, da circa il 70% a circa il 95%, da circa l'80% a circa il 95%, da circa l'85% a circa il 95%, da circa il 90% a circa il 95%, da circa il 30% a circa il 90%, da circa il 40% a circa il 90%, da circa il 50% a circa il 90%, da circa il 60% a circa il 90%, da circa il 70% a circa il 90%, da circa l'80% a circa il 90%, o almeno circa il 30%, circa il 35%, circa il 40%, circa il 45%, circa il 50%, circa il 55%, circa il 60%, circa il 65%, circa il 70%, circa il 75%, circa l'80%, circa l'85%, circa il 90%, circa il 91%, circa il 92%, circa il 93%, circa il 94%, circa il 95%, circa il 96%, circa il 97%, circa il 98% o circa il 99% (o qualsiasi frazione o intervallo di queste) delle particelle ha l'RNA incapsulato al suo interno. La quantità può essere qualsiasi valore o sottovalore all'interno degli intervalli recitati, compresi gli endpoint. L'RNA incluso in qualsiasi composizione RNA-lipide o formulazione RNA-lipide fornita nel presente documento può essere un RNA autoreplicante.

A seconda dell'uso previsto della formulazione lipidica, le proporzioni dei componenti possono essere variate e l'efficienza di distribuzione di una particolare formulazione può essere misurata utilizzando analisi note nel settore.

In alcuni aspetti, le molecole di acido nucleico fornite nel presente documento sono formulate in lipidi. La formulazione lipidica è preferibilmente selezionata tra, ma non limitata a, liposomi,

liposomi cationici e nanoparticelle lipidiche. In un aspetto, una formulazione lipidica è un liposoma cationico o una nanoparticella lipidica (LNP) comprendente:

- (a) un RNA della presente informativa,
- (b) un lipide cationico,
- (c) un agente di riduzione dell'aggregazione (come il lipide di glicole polietilenico (PEG) o il lipide modificato con PEG),
- (d) facoltativamente un lipide non cationico (come un lipide neutro), e
- (e) facoltativamente, uno sterolo.

In un altro aspetto, il lipide cationico è un lipide cationico ionizzabile. Qualsiasi lipide cationico ionizzabile può essere incluso nelle formulazioni lipidiche, compresi i lipidi cationici esemplari forniti nel presente documento.

Lipidi cationici

In un aspetto, la formulazione di nanoparticelle lipidiche comprende (i) almeno un lipide cationico; (ii) un lipide helper; (iii) uno sterolo (ad esempio, colesterolo); e (iv) un PEG-lipide. In un altro aspetto, il lipide cationico è un lipide cationico ionizzabile. In un altro aspetto ancora, la formulazione di nanoparticelle lipidiche comprende (i) almeno un lipide cationico; (ii) un lipide helper; (iii) uno sterolo (ad esempio, colesterolo); e (iv) un PEG-lipide, in un rapporto molare di circa 40-70% lipide cationico ionizzabile: circa 2-15% lipide helper: circa 20-45% sterolo; circa 0,5-5% PEG-lipide. In un ulteriore aspetto, il lipide cationico è un lipide cationico ionizzabile.

In un aspetto, la formulazione di nanoparticelle lipidiche è composta da (i) almeno un lipide cationico; (ii) un lipide helper; (iii) uno sterolo (ad esempio, colesterolo); e (iv) un PEG-lipide. In un altro aspetto, il lipide cationico è un lipide cationico ionizzabile. In un altro aspetto ancora, la formulazione di nanoparticelle lipidiche è composta da (i) almeno un lipide cationico; (ii) un lipide helper; (iii) uno sterolo (ad esempio, colesterolo); e (iv) un PEG-lipide, in un rapporto molare di circa 40-70% lipide cationico ionizzabile: circa 2-15% lipide helper: circa 20-45% sterolo; circa 0,5-5% PEG-lipide. In un ulteriore aspetto, il lipide cationico è un lipide cationico ionizzabile.

Nelle formulazioni lipidiche attualmente divulgate, il lipide cationico può essere, ad esempio, cloruro di N,N-dioleil-N,N-dimetilammonio (DODAC), bromuro di N,N-distearil-N,N-dimetilammonio (DDAB), cloruro di 1,2-dioleoiltrimetilamminopropano (DOTAP) (noto anche come cloruro di N-(2,3-dioleoilossi)propil)-N,N,N-trimetilammonio e sale di cloruro di 1,2-Dioleilossi-3-trimetilamminopropano), cloruro di N-(1-(2,3-dioleilossi)propil)-N,N,N-trimetilammonio (DOTMA), N,N-dimetil-2,3-dioleilossi)propilammina (DODMA), 1,2-Dilinoleilossi-N,N-dimetilamminopropano (DLinDMA), 1,2-dilinenilossi-N,N-dimetilamminopropano (DLinDMA), 1,2-di- γ -linolenilossi-N,N-dimetilamminopropano (γ -DLinDMA), 1,2-dilinoilcarbamoilossi-3-dimetilamminopropano (DLin-C-DAP), 1,2-dilinoilossi-3-(dimetilammino)acetossipropano (DLin-DAC), 1,2-dilinoilossi-3-morfolinopropano (DLin-MA), 1,2-dilinoil-3-dimetilamminopropano (DLinDAP), 1,2-dilinoiltio-3-dimetilamminopropano (DLin-S-DMA), 1-linoil-2-linoilossi-3-dimetilamminopropano (DLin-2-DMAP), Sale di cloruro di 1,2-dilinoilossi-3-trimetilamminopropano (DLin-TMA·Cl), sale di cloruro di 1,2-dilinoil-3-trimetilamminopropano (DLin-TAP·Cl), 1,2-dilinoilossi-3-(N-metilpiperazino)propano (DLin-MPZ),

o 3-(N,N-dilinoilammino)-1,2-propandiolo (DLinAP), 3-(N,N-dioleilammino)-1,2-propandiolo (DOAP), 1,2-dilinoiloloso-3-(2-N,N-dimetilammino)etossipropano (DLin-EG-DMA), 2,2-dilinoil-4-dimetilamminometil-[1,3]-diossolano (DLin-K-DMA) o analoghi degli stessi, (3aR,5s,6aS)-N,N-dimetil-2,2-di((9Z,12Z)-ottadeca-9,12-dienil)tetraidro-3aH-ciclopenta[d][1,3]diossol-5-ammina, (6Z,9Z,28Z,31Z)-eptatriaconta-6,9,28,31-tetraen-19-il-4-(dimetilammino)butanoato (MC3), 1,1'-(2-(4-(2-((2-(bis(2-idrossidodecil)ammino)etil)(2-idrossidodecil)ammino)etil)piperazin-1-il)etilazandiil)didodecan-2-olo (C12-200), 2,2-dilinoil-4-(2-dimetilamminoetil)-[1,3]-diossolano (DLin-K-C2-DMA), 2,2-dilinoil-4-dimetilamminometil-[1,3]-diossolano (DLin-K-DMA), (6Z,9Z,28Z,31Z)-eptatriaconta-6,9,28,31-tetraen-19-il-4-(dimetilammino)butanoato (DLin-M-C3-DMA), 3-((6Z,9Z,28Z,31Z)-eptatriaconta-6,9,28,31-tetraen-19-ilossi)-N,N-dimetilpropan-1-ammina (MC3 Etere), 4-((6Z,9Z,28Z,31Z)-eptatriaconta-6,9,28,31-tetraen-19-ilossi)-N,N-dimetilbutan-1-ammina (MC4 Etere), o qualsiasi combinazione di cui. Altri lipidi cationici includono, ma non sono limitati a, N,N-distearil-N,N-dimetilammonio bromuro (DDAB), 3P-(N-(N',N'-dimetilammino)etano)carbamoil)colesterolo (DC-Choi), N-(1-(2,3-dioleilossi)propil)-N-2-(sperminacarbossamido)etil)-N,N-dimetilammonio trifluoracetato (DOSPA), diottadecilamidoglicil carbossispermina (DOGS), 1,2-dioleil-sn-3-fosfoetanolammina (DOPE), 1,2-dioleil-3-dimetilammonio propano (DODAP), N-(1,2-dimiristolossiprop-3-il)-N,N-dimetil-N-idrossietil ammonio bromuro (DMRIE) e 2,2-Dilinoil-4-dimetilamminoetil-[1,3]-diossolano (XTC). Inoltre, possono essere utilizzate preparazioni commerciali di lipidi cationici, come, ad esempio, LIPOFECTIN (inclusi DOTMA e DOPE, disponibili da GIBCO/BRL), e Lipofectamine (comprendente DOSPA e DOPE, disponibile presso GIBCO/BRL).

Altri lipidi cationici adatti sono descritti nelle pubblicazioni internazionali n. WO 09/086558, WO 09/127060, WO 10/048536, WO 10/054406, WO 10/088537, WO 10/129709 e WO 2011/153493; nelle pubblicazioni di brevetti statunitensi n. 2011/0256175, 2012/0128760 e 2012/0027803; nel brevetto statunitense n. 8.158.601; e in Love et al., PNAS, 107(5), 1864-69, 2010, il cui contenuto è qui incorporato per riferimento.

Le formulazioni RNA-lipidiche della presente divulgazione possono comprendere un lipide helper, che può essere definito lipide helper neutro, lipide non cationico, lipide helper non cationico, lipide anionico, lipide helper anionico o lipide neutro. È stato scoperto che le formulazioni lipidiche, in particolare i liposomi cationici e le nanoparticelle lipidiche, hanno un aumento dell'assorbimento cellulare se i lipidi helper sono presenti nella formulazione. (Curr. Drug Metab. 2014; 15(9):882-92). Ad esempio, alcuni studi hanno indicato che i lipidi neutri e zwitterionici come 1,2-dioleil-sn-glicerolo-3-fosfatidilcolina (DOPC), Di-Oleoyl-Phosphatidyl-Ethanolamine (DOPE) e 1,2-DiStearoyl-sn-glicerolo-3-fosfocolina (DSPC), essendo più fusogeni (vale a dire, facilitando la fusione) dei lipidi cationici, possono influenzare le caratteristiche polimorfiche dei complessi lipidi-acidi nucleici, promuovendo la transizione da una fase lamellare a una esagonale e quindi inducendo la fusione e una rottura della membrana cellulare. (Nanomedicine (Lond). 2014 gennaio; 9(1):105-20). Inoltre, l'uso di lipidi helper può aiutare a ridurre qualsiasi potenziale effetto dannoso derivante dall'uso di molti lipidi cationici prevalenti come tossicità e immunogenicità.

Esempi non limitanti di lipidi non cationici adatti per le formulazioni lipidiche della presente informativa includono fosfolipidi quali lecitina, fosfatidiletanolammina, lisolecitina, lisofosfatidiletanolammina, fosfatidilserina, fosfatidilinositolo, sfingomieline, sfingomieline d'uovo (ESM), cefalina, cardiopina, acido fosfatidico, cerebrosidi, dicetilfosfato, distearoilfosfatidilcolina

(DSPC), dioleoilfosfatidilcolina (DOPC), dipalmitoilfosfatidilcolina (DPPC), dioleoilfosfatidilglicerolo (DOPG), dipalmitoilfosfatidilglicerolo (DPPG), dioleoilfosfatidiletanolammina (DOPE), palmitoiloleoil-fosfatidilcolina (POPC), palmitoiloleoil-fosfatidiletanolammina (POPE), palmitoiloleyol-fosfatidilglicerolo (POPG), dioleoilfosfatidiletanolammina 4-(N-maleimidometil)-cicloesano-1-carbossilato (DOPE-mal), dipalmitoil-fosfatidiletanolammina (DPPE), dimiristoil-fosfatidiletanolammina (DMPE), distearoil-fosfatidiletanolammina (DSPE), monometil-fosfatidiletanolammina, dimetil-fosfatidiletanolammina, dielaidoil-fosfatidiletanolammina (DEPE), stearoiloleoil-fosfatidiletanolammina (SOPE), lisofosfatidilcolina, dilinoleoilfosfatidilecolina e miscele di queste. Possono essere utilizzati anche altri fosfolipidi di diacilfosfatidilecolina e diacilfosfatidiletanolammina. I gruppi acilici in questi lipidi sono preferibilmente gruppi acilici derivati da acidi grassi con catene di carbonio da C10 a C24, ad esempio lauroile, miristoile, palmitoile, stearoile o oleoile.

Ulteriori esempi di lipidi non cationici includono steroli come il colesterolo e i suoi derivati. Come lipide ausiliario, il colesterolo aumenta la spaziatura delle cariche dello strato lipidico che si interfaccia con l'acido nucleico, facendo sì che la distribuzione delle cariche corrisponda più da vicino a quella dell'acido nucleico. (JR Soc. Interface. 2012 7 marzo; 9(68): 548-561). Esempi non limitanti di derivati del colesterolo includono analoghi polari come 5 α -colestano, 5 α -coprostanolo, colesteril-(2'-idrossi)-etil etere, colesteril-(4'-idrossi)-butil etere e 6-chetocolestano; analoghi non polari come 5 α -colestano, colestene, 5 α -colestano, 5 α -colestano e colesteril decanoato; e miscele degli stessi. Per alcuni aspetti, il derivato del colesterolo è un analogo polare come il colesteril-(4'-idrossi)-butil etere.

In alcuni aspetti, il lipide helper presente nella formulazione lipidica comprende o è costituito da una miscela di uno o più fosfolipidi e colesterolo o un suo derivato. In altri aspetti, il lipide neutro presente nella formulazione lipidica comprende o è costituito da uno o più fosfolipidi, ad esempio, una formulazione lipidica priva di colesterolo. In altri aspetti ancora, il lipide neutro presente nella formulazione lipidica comprende o è costituito da colesterolo o un suo derivato, ad esempio, una formulazione lipidica priva di fosfolipidi.

Altri esempi di lipidi helper includono lipidi non contenenti fosforo quali, ad esempio, stearylammina, dodecylammina, esadecylammina, acetil palmitato, glicerolo ricinoleato, esadecil stearato, isopropil miristato, polimeri acrilici anfoteri, trietanolammina-lauril solfato, alchil-aril solfato, ammidi di acidi grassi polietilossilati, diottadecildimetil ammonio bromuro, ceramide e sfingomielina.

Altri lipidi cationici adatti includono quelli con gruppi di acidi grassi alternativi e altri gruppi dialchilammino, compresi quelli in cui i sostituenti alchilici sono diversi (ad esempio, N-etil-N-metilammino- e N-propil-N-etilammino-). Questi lipidi fanno parte di una sottocategoria di lipidi cationici denominati amminolipidi. In alcune forme di realizzazione delle formulazioni lipidiche descritte nel presente documento, il lipide cationico è un amminolipide. In generale, gli amminolipidi con catene aciliche meno sature sono più facilmente dimensionabili, in particolare quando i complessi devono essere dimensionati al di sotto di circa 0,3 micron, ai fini della sterilizzazione del filtro. Possono essere utilizzati amminolipidi contenenti acidi grassi insaturi con lunghezze di catena di carbonio nell'intervallo da C₁₄ a C₂₂. Possono essere utilizzati anche altri

scaffold per separare il gruppo amminico e la porzione di acido grasso o alchile grasso dell'amminolipide.

In alcune realizzazioni, la formulazione lipidica comprende il lipide cationico con Formula I secondo la domanda di brevetto PCT/EP2017/064066. In questo contesto, la divulgazione di PCT/EP2017/064066 è anche incorporata nel presente documento per riferimento.

In alcune realizzazioni, i lipidi amminici o cationici della presente informativa sono ionizzabili e hanno almeno un gruppo protonabile o deprotonabile, in modo che il lipide sia caricato positivamente a un pH pari o inferiore al pH fisiologico (ad esempio, pH 7,4) e neutro a un secondo pH, preferibilmente pari o superiore al pH fisiologico. Naturalmente, si comprenderà che l'aggiunta o la rimozione di protoni in funzione del pH è un processo di equilibrio e che il riferimento a un lipide carico o neutro si riferisce alla natura della specie predominante e non richiede che tutto il lipide sia presente nella forma carica o neutra. I lipidi che hanno più di un gruppo protonabile o deprotonabile, o che sono zwitterionici, non sono esclusi dall'uso nella presente informativa. In alcune realizzazioni, i lipidi protonabili hanno un pKa del gruppo protonabile nell'intervallo da circa 4 a circa 11. In alcune realizzazioni, il lipide cationico ionizzabile ha un pKa da circa 5 a circa 7. In alcune realizzazioni, il pKa di un lipide cationico ionizzabile è da circa 6 a circa 7.

In alcune realizzazioni, la formulazione lipidica comprende un lipide cationico ionizzabile di Formula I.

- o un sale o solvato farmaceuticamente accettabile dello stesso, in cui R5 e R6 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile C1-C31 lineare o ramificato, un alchenile C2-C31 o un alchinile C2-C31 e colesterile; L5 e L6 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile C1-C20 lineare e un alchenile C2-C20; X5 è —C(O)O— , per cui si forma —C(O)O—R6 o —OC(O)— per cui si forma —OC(O)—R6 ; X6 è —C(O)O— per cui si forma —C(O)O—R5 o —OC(O)— per cui si forma —OC(O)—R5 ; X7 è S o O; L7 è alchile assente o inferiore; R4 è un alchile C1-C6 lineare o ramificato; e R7 e R8 sono ciascuno selezionati indipendentemente dal gruppo costituito da un idrogeno e un alchile C1-C6 lineare o ramificato.

In alcune realizzazioni, X7 è S.

In alcune realizzazioni, X5 è —C(O)O— , per cui si forma —C(O)O—R6 e X6 è —C(O)O— per cui si forma —C(O)O—R5 .

In alcune realizzazioni, R7 e R8 sono ciascuno selezionati indipendentemente dal gruppo costituito da metile, etile e isopropile.

In alcune realizzazioni, L5 e L6 sono ciascuno indipendentemente un alchile C1-C10. In alcune realizzazioni, L5 è un alchile C1-C3 e L6 è un alchile C1-C5. In alcune realizzazioni, L6 è un alchile C1-C2. In alcune realizzazioni, L5 e L6 sono ciascuno un alchile lineare C7. In alcune realizzazioni, L5 e L6 sono ciascuno un alchile lineare C9.

In alcune realizzazioni, R5 e R6 sono ciascuno indipendentemente un alchenile. In alcune realizzazioni, R6 è alchenile. In alcune realizzazioni, R6 è alchenile C2-C9. In alcune realizzazioni,

l'alchenile comprende un singolo doppio legame. In alcune realizzazioni, R5 e R6 sono ciascuno alchile. In alcune realizzazioni, R5 è un alchile ramificato. In alcune realizzazioni, R5 e R6 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile C9, un alchenile C9 e un alchinile C9. In alcune realizzazioni, R5 e R6 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile C11, un alchenile C11 e un alchinile C11. In alcune realizzazioni, R5 e R6 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile C7, un alchenile C7 e un alchinile C7. In alcune realizzazioni, R5 è $-\text{CH}((\text{CH}_2)_p\text{CH}_3)_2$ o $-\text{CH}((\text{CH}_2)_p\text{CH}_3)((\text{CH}_2)_{p-1}\text{CH}_3)$, in cui p è 4-8. In alcune realizzazioni, p è 5 e L5 è un alchile C1-C3. In alcune realizzazioni, p è 6 e L5 è un alchile C3. In alcune realizzazioni, p è 7. In alcune realizzazioni, p è 8 e L5 è un alchile C1-C3. In alcune realizzazioni, R5 è costituito da $-\text{CH}((\text{CH}_2)_p\text{CH}_3)((\text{CH}_2)_{p-1}\text{CH}_3)$, in cui p è 7 o 8.

In alcune realizzazioni, R4 è etilene o propilene. In alcune realizzazioni, R4 è n-propilene o isobutilene.

In alcune realizzazioni, L7 è assente, R4 è etilene, X7 è S e R7 e R8 sono ciascuno metile. In alcune realizzazioni, L7 è assente, R4 è n-propilene, X7 è S e R7 e R8 sono ciascuno metile. In alcune realizzazioni, L7 è assente, R4 è etilene, X7 è S e R7 e R8 sono ciascuno etile.

In alcune realizzazioni, X7 è S, X5 è $-\text{C}(\text{O})\text{O}-$, per cui si forma $-\text{C}(\text{O})\text{O}-\text{R}_6$, X6 è $-\text{C}(\text{O})\text{O}-$ per cui si forma $-\text{C}(\text{O})\text{O}-\text{R}_5$, L5 e L6 sono ciascuno indipendentemente un alchile lineare C3-C7, L7 è assente, R5 è $-\text{CH}((\text{CH}_2)_p\text{CH}_3)_2$ e R6 è alchenile C7-C12. In alcune ulteriori realizzazioni, p è 6 e R6 è alchenile C9.

In alcune realizzazioni, la formulazione lipidica può comprendere un lipide cationico ionizzabile selezionato dal gruppo costituito da LIPIDI #1 a LIPIDI #8:

TABELLA 6 LIPIDI # STRUTTURA 1 2 3 4 5 6 7 8

In alcune realizzazioni, la formulazione lipidica comprende un lipide cationico ionizzabile avente una struttura selezionata da

o un sale farmaceuticamente accettabile dello stesso.

In alcune forme di realizzazione preferite, la formulazione lipidica comprende un lipide cationico ionizzabile avente la struttura

o un sale farmaceuticamente accettabile dello stesso.

In alcune realizzazioni, uno o più lipidi qui menzionati possono essere espressamente esclusi.

In alcuni aspetti, il lipide ausiliario comprende da circa 2 moli % a circa 20 moli %, da circa 3 moli % a circa 18 moli %, da circa 4 moli % a circa 16 moli %, da circa 5 moli % a circa 14 moli %, da circa 6 moli % a circa 12 moli %, da circa 5 moli % a circa 10 moli %, da circa 5 moli % a circa 9 moli %, o circa 2 moli %, circa 3 moli %, circa 4 moli %, circa 5 moli %, circa 6 moli %, circa 7 moli

%, circa 8 moli %, circa 9 moli %, circa 10 moli %, circa 11 moli %, o circa 12 moli % (o qualsiasi frazione di essa o l'intervallo in essa compreso) del lipide totale presente nella formulazione lipidica.

La porzione lipidica, o il colesterolo o il derivato del colesterolo nella formulazione lipidica può comprendere fino a circa il 40% molare, circa il 45% molare, circa il 50% molare, circa il 55% molare o circa il 60% molare del lipide totale presente nella formulazione lipidica. In alcuni aspetti, il colesterolo o il derivato del colesterolo comprende circa il 15% molare a circa il 45% molare, circa il 20% molare a circa il 40% molare, circa il 25% molare a circa il 35% molare o circa il 28% molare a circa il 35% molare; o circa il 25% molare, circa il 26% molare, circa il 27% molare, circa il 28% molare, circa il 29% molare, circa il 30% molare, circa il 31% molare, circa il 32% molare, circa il 33% molare, circa il 34% molare, circa il 35% molare, circa il 36% molare o circa il 37% molare del lipide totale presente nella formulazione lipidica.

In forme di realizzazione specifiche, la porzione lipidica della formulazione lipidica è pari a circa il 35% molare fino a circa il 42% molare di colesterolo.

In alcuni aspetti, il componente fosfolipidico nella miscela può comprendere da circa 2 moli % a circa 20 moli %, da circa 3 moli % a circa 18 moli %, da circa 4 moli % a circa 16 moli %, da circa 5 moli % a circa 14 moli %, da circa 6 moli % a circa 12 moli %, da circa 5 moli % a circa 10 moli %, da circa 5 moli % a circa 9 moli %, o circa 2 moli %, circa 3 moli %, circa 4 moli %, circa 5 moli %, circa 6 moli %, circa 7 moli %, circa 8 moli %, circa 9 moli %, circa 10 moli %, circa 11 moli %, o circa 12 moli % (o qualsiasi frazione di essa o l'intervallo in essa compreso) del lipide totale presente nella formulazione lipidica.

In alcune realizzazioni, la porzione lipidica della formulazione lipidica comprende circa, ma non è necessariamente limitata a, dal 40% molare a circa il 60% molare del lipide cationico ionizzabile, da circa il 4% molare a circa il 16% molare di DSPC, da circa il 30% molare a circa il 47% molare di colesterolo e da circa lo 0,5% molare a circa il 3% molare di PEG2000-DMG.

In alcune realizzazioni, la porzione lipidica della formulazione lipidica può comprendere, ma non è necessariamente limitata a, circa il 42% molare a circa il 58% molare del lipide cationico ionizzabile, circa il 6% molare a circa il 14% molare di DSPC, circa il 32% molare a circa il 44% molare di colesterolo e circa l'1% molare a circa il 2% molare di PEG2000-DMG.

In alcune realizzazioni, la porzione lipidica della formulazione lipidica può comprendere, ma non è necessariamente limitata a, circa il 45% molare a circa il 55% molare del lipide cationico ionizzabile, circa l'8% molare a circa il 12% molare di DSPC, circa il 35% molare a circa il 42% molare di colesterolo e circa l'1,25% molare a circa l'1,75% molare di PEG2000-DMG.

La percentuale di lipidi helper presenti nella formulazione lipidica è una quantità target e la quantità effettiva di lipidi helper presente nella formulazione può variare, ad esempio, di ± 5 moli %.

Una formulazione lipidica che comprende un composto lipidico cationico o un composto lipidico cationico ionizzabile può essere su base molare circa il 30-70% di composto lipidico cationico, circa

il 25-40% di colesterolo, circa il 2-15% di lipide helper e circa lo 0,5-5% di un lipide di polietilenglicole (PEG), dove la percentuale è del lipide totale presente nella formulazione. In alcuni aspetti, la composizione è circa il 40-65% di composto lipidico cationico, circa il 25-35% di colesterolo, circa il 3-9% di lipide helper e circa lo 0,5-3% di un lipide PEG, dove la percentuale è del lipide totale presente nella formulazione.

La formulazione può essere una formulazione di particelle lipidiche, contenente ad esempio l'8-30% di composto di acido nucleico, il 5-30% di lipidi helper e lo 0-20% di colesterolo; il 4-25% di lipidi cationici, il 4-25% di lipidi helper, il 2-25% di colesterolo, il 10-35% di colesterolo-PEG e il 5% di colesterolo-ammina; oppure il 2-30% di lipidi cationici, il 2-30% di lipidi helper, l'1-15% di colesterolo, il 2-35% di colesterolo-PEG e l'1-20% di colesterolo-ammina; oppure fino al 90% di lipidi cationici e il 2-10% di lipidi helper, o addirittura il 100% di lipidi cationici.

Coniugati lipidici

Le formulazioni lipidiche descritte nel presente documento possono comprendere ulteriormente un coniugato lipidico. Il lipide coniugato è utile in quanto impedisce l'aggregazione di particelle. I lipidi coniugati adatti includono, ma non sono limitati a, coniugati lipidi-PEG, coniugati lipidi-polimero-cationico e miscele degli stessi. Inoltre, i veicoli di erogazione lipidica possono essere utilizzati per un targeting specifico mediante l'attacco di ligandi (ad esempio, anticorpi, peptidi e carboidrati) alla sua superficie o all'estremità terminale delle catene PEG attaccate (Front Pharmacol. 2015 1 dicembre; 6:286).

In alcuni aspetti, il coniugato lipidico è un PEG-lipide. L'inclusione di polietilenglicole (PEG) in una formulazione lipidica come rivestimento o legante di superficie, una tecnica denominata PEGilazione, aiuta a proteggere le nanoparticelle dal sistema immunitario e dalla loro fuga dall'assorbimento RES (Nanomedicine (Lond). 2011 giugno; 6(4):715-28). La PEGilazione è stata utilizzata per stabilizzare le formulazioni lipidiche e i loro carichi utili attraverso meccanismi fisici, chimici e biologici. I lipidi PEG simili a detergenti (ad esempio, PEG-DSPE) possono entrare nella formulazione lipidica per formare uno strato idratato e una barriera sterica sulla superficie. In base al grado di PEGilazione, lo strato superficiale può essere generalmente suddiviso in due tipi, strati a spazzola e a fungo. Per le formulazioni stabilizzate con PEG-DSPE, il PEG assumerà la conformazione a fungo a un basso grado di PEGilazione (solitamente inferiore al 5% in moli) e passerà alla conformazione a spazzola quando il contenuto di PEG-DSPE aumenta oltre un certo livello (Journal of Nanomaterials. 2011; 2011:12). La PEGilazione porta a un aumento significativo dell'emivita di circolazione delle formulazioni lipidiche (Annu. Rev. Biomed. Eng. 2011 15 agosto; 13():507-30; J. Control Release. 2010 3 agosto; 145(3):178-81).

Esempi di PEG-lipidi includono, ma non sono limitati a, PEG accoppiato a dialchilossipropili (PEG-DAA), PEG accoppiato a diacilglicerolo (PEG-DAG), metossipolietilenglicole (PEG-DMG o PEG2000-DMG), PEG accoppiato a fosfolipidi come la fosfatidiletanolammina (PEG-PE), PEG coniugato a ceramidi, PEG coniugato a colesterolo o a un suo derivato e miscele degli stessi.

Il PEG è un polimero lineare, solubile in acqua, di unità ripetute di etilene PEG con due gruppi idrossilici terminali. I PEG sono classificati in base al loro peso molecolare e includono i seguenti: monometossipolietilenglicole (MePEG-OH), monometossipolietilenglicole-succinato (MePEG-S),

monometossipolietilenglicole-succinimidil succinato (MePEG-S-NHS), monometossipolietilenglicole-ammina (MePEG-NH₂), monometossipolietilenglicole-tresilato (MePEG-TRES), monometossipolietilenglicole-imidazolil-carbonile (MePEG-IM), nonché composti contenenti un gruppo idrossilico terminale anziché un gruppo metossilico terminale (ad esempio, HO-PEG-S, HO-PEG-S—NHS, HO-PEG-NH₂).

La frazione PEG dei coniugati PEG-lipidi qui descritti può comprendere un peso molecolare medio che va da circa 550 dalton a circa 10.000 dalton. In alcuni aspetti, la frazione PEG ha un peso molecolare medio che va da circa 750 dalton a circa 5.000 dalton (ad esempio, da circa 1.000 dalton a circa 5.000 dalton, da circa 1.500 dalton a circa 3.000 dalton, da circa 750 dalton a circa 3.000 dalton, da circa 750 dalton a circa 2.000 dalton). In alcuni aspetti, la frazione PEG ha un peso molecolare medio di circa 2.000 dalton o circa 750 dalton. Il peso molecolare medio può essere qualsiasi valore o sottovalore all'interno degli intervalli indicati, inclusi gli endpoint.

In certi aspetti, il PEG può essere facoltativamente sostituito da un gruppo alchilico, alcossilico, acilico o arilico. Il PEG può essere coniugato direttamente al lipide o può essere legato al lipide tramite una frazione di collegamento. Può essere utilizzata qualsiasi frazione di collegamento adatta per accoppiare il PEG a un lipide, inclusi, ad esempio, frazioni di collegamento non contenenti esteri e frazioni di collegamento contenenti esteri. In un aspetto, la frazione di collegamento è una frazione di collegamento non contenente esteri. Esempi di gruppi di collegamento non contenenti esteri includono, ma non sono limitati a, ammido (—C(O)NH—), ammino (—NR—), carbonile (—C(O)—), carbammato (—NHC(O)O—), urea (—NHC(O)NH—), disolfuro (—S—S—), etere (—O—), succinile (—(O)CCH₂CH₂C(O)—), succinamidile (—NHC(O)CH₂CH₂C(O)NH—), etere, così come combinazioni di questi (come un collegamento contenente sia un gruppo di collegamento carbammato che un gruppo di collegamento ammidico). In un aspetto, un collegamento carbammato viene utilizzato per accoppiare il PEG al lipide.

In alcuni aspetti, una frazione linker contenente estere viene utilizzata per accoppiare il PEG al lipide. Fra le frazioni linker contenenti estere esemplificative sono inclusi, ad esempio, carbonato (—OC(O)O—), succinoile, esteri fosfatici (—O—(O)POH—O—), esteri solfonati e loro combinazioni.

Le fosfatidiletanolammine con una varietà di gruppi di catene aciliche di diverse lunghezze di catena e gradi di saturazione possono essere coniugate al PEG per formare il coniugato lipidico. Tali fosfatidiletanolammine sono disponibili in commercio o possono essere isolate o sintetizzate utilizzando tecniche convenzionali note agli esperti del settore. Sono preferite le fosfatidiletanolammine contenenti acidi grassi saturi o insaturi con lunghezze di catena di carbonio nell'intervallo da C₁₀ a C₂₀. Possono essere utilizzate anche fosfatidiletanolammine con acidi grassi mono- o di-insaturi e miscele di acidi grassi saturi e insaturi. Le fosfatidiletanolammine adatte includono, ma non sono limitate a, dimiristoil-fosfatidiletanolammina (DMPE), dipalmitoil-fosfatidiletanolammina (DPPE), dioleoil-fosfatidiletanolammina (DOPE) e distearoil-fosfatidiletanolammina (DSPE).

In alcuni aspetti, il coniugato PEG-DAA è un coniugato PEG-didecilossipropile (C₁₀), un coniugato PEG-dilaurilossipropile (C₁₂), un coniugato PEG-dimiristolossipropile (C₁₄), un coniugato PEG-dipalmitilossipropile (C₁₆) o un coniugato PEG-distearilossipropile (C₁₈). In alcuni aspetti, il PEG ha

un peso molecolare medio di circa 750 o circa 2.000 dalton. In alcuni aspetti, il gruppo idrossile terminale del PEG è sostituito con un gruppo metilico.

Oltre a quanto sopra, altri polimeri idrofili possono essere utilizzati al posto del PEG. Esempi di polimeri adatti che possono essere utilizzati al posto del PEG includono, ma non sono limitati a, polivinilpirrolidone, polimetilossazolina, polietilossazolina, poliidrossipropile, metacrilammide, polimetacrilammide e polidimetilacrilammide, acido polilattico, acido poliglicolico e cellulose derivatizzate come idrossimetilcellulosa o idrossietilcellulosa.

In alcuni aspetti, il coniugato lipidico (ad esempio, PEG-lipide) comprende da circa 0,1 moli % a circa 2 moli %, da circa 0,5 moli % a circa 2 moli %, da circa 1 moli % a circa 2 moli %, da circa 0,6 moli % a circa 1,9 moli %, da circa 0,7 moli % a circa 1,8 moli %, da circa 0,8 moli % a circa 1,7 moli %, da circa 0,9 moli % a circa 1,6 moli %, da circa 0,9 moli % a circa 1,8 moli %, da circa 1 mol % a circa 1,8 moli %, da circa 1 mol % a circa 1,7 moli %, da circa 1,2 moli % a circa 1,8 moli %, da circa 1,2 moli % a circa 1,7 moli %, da circa 1,3 moli % a circa 1,6 moli %, o da circa 1,4 moli % a circa 1,6 moli % (o qualsiasi frazione o intervallo di essa) del lipide totale presente nella formulazione lipidica. In altre realizzazioni, il coniugato lipidico (ad esempio, PEG-lipide) comprende circa 0,5%, 0,6%, 0,7%, 0,8%, 0,9%, 1,0%, 1,2%, 1,3%, 1,4%, 1,5%, 1,6%, 1,7%, 1,8%, 1,9%, 2,0%, 2,5%, 3,0%, 3,5%, 4,0%, 4,5% o 5% (o qualsiasi frazione o intervallo di essa) del lipide totale presente nella formulazione lipidica. La quantità può essere qualsiasi valore o sottovalore all'interno degli intervalli recitati, inclusi gli endpoint.

La percentuale di coniugato lipidico (ad esempio, PEG-lipide) presente nelle formulazioni lipidiche della divulgazione è una quantità target e la quantità effettiva di coniugato lipidico presente nella formulazione può variare, ad esempio, di $\pm 0,5$ moli %. Un esperto del settore comprenderà che la concentrazione del coniugato lipidico può essere variata a seconda del coniugato lipidico impiegato e della velocità alla quale la formulazione lipidica deve diventare fusogena.

In alcune realizzazioni, la formulazione lipidica per una qualsiasi delle composizioni descritte nel presente documento comprende un lipoplex, un liposoma, una nanoparticella lipidica, una particella a base di polimero, un esosoma, un corpo lamellare, una micella o un'emulsione.

Meccanismo d'azione per l'assorbimento cellulare delle formulazioni lipidiche

In alcuni aspetti, le formulazioni lipidiche per la somministrazione intracellulare di acidi nucleici, in particolare liposomi, liposomi cationici e nanoparticelle lipidiche, sono progettate per l'assorbimento cellulare penetrando nelle cellule bersaglio attraverso lo sfruttamento dei meccanismi endocitotici delle cellule bersaglio in cui il contenuto del veicolo di somministrazione lipidica viene consegnato al citosol della cellula bersaglio. (Nucleic Acid Therapeutics, 28(3):146-157, 2018). Prima dell'endocitosi, i ligandi funzionalizzati come il PEG-lipide sulla superficie del veicolo di somministrazione lipidica vengono rilasciati dalla superficie, il che innesca l'internalizzazione nella cellula bersaglio. Durante l'endocitosi, una parte della membrana plasmatica della cellula circonda il vettore e lo ingloba in una vescicola che poi si stacca dalla membrana cellulare, entra nel citosol e infine entra e si muove attraverso il percorso endolisosomiale. Per i veicoli di rilascio contenenti lipidi cationici ionizzabili, l'acidità aumentata con l'invecchiamento dell'endosoma determina un veicolo con una forte carica positiva sulla

superficie. Le interazioni tra il veicolo di rilascio e la membrana endosomiale determinano quindi un evento di fusione della membrana che porta al rilascio citosolico del carico utile. Per i carichi utili di RNA, i processi di traduzione interna della cellula tradurranno quindi l'RNA nella proteina codificata. La proteina codificata può ulteriormente subire un'elaborazione post-traduzionale, incluso il trasporto a un organello mirato o a una posizione all'interno della cellula o l'escrezione dalla cellula.

Controllando la composizione e la concentrazione del coniugato lipidico, si può controllare la velocità con cui il coniugato lipidico si scambia dalla formulazione lipidica e, a sua volta, la velocità con cui la formulazione lipidica diventa fusogena. Inoltre, altre variabili, tra cui, ad esempio, pH, temperatura o forza ionica, possono essere utilizzate per variare e/o controllare la velocità con cui la formulazione lipidica diventa fusogena. Altri metodi che possono essere utilizzati per controllare la velocità con cui la formulazione lipidica diventa fusogena diventeranno evidenti agli esperti del settore dopo aver letto questa informativa. Inoltre, controllando la composizione e la concentrazione del coniugato lipidico, si può controllare la dimensione delle particelle liposomiali o lipidiche.

Produzione di formulazioni lipidiche

Esistono molti metodi diversi per la preparazione di formulazioni lipidiche comprendenti un acido nucleico. (Curr. Drug Metabol. 2014, 15, 882-892; Chem. Phys. Lipids 2014, 177, 8-18; Int. J. Pharm. Stud. Res. 2012, 3, 14-20). Le tecniche di idratazione a film sottile, doppia emulsione, evaporazione in fase inversa, preparazione microfluidica, centrifugazione asimmetrica doppia, iniezione di etanolo, dialisi con detergente, formazione spontanea di vescicole mediante diluizione con etanolo e incapsulamento in liposomi preformati sono brevemente descritte nel presente documento.

Idratazione a film sottile

Nell'idratazione a film sottile (TFH) o nel metodo Bangham, i lipidi vengono disciolti in un solvente organico, quindi evaporati tramite l'uso di un evaporatore rotante che porta alla formazione di uno strato sottile di lipidi. Dopo l'idratazione dello strato tramite una soluzione tampone acquosa contenente il composto da caricare, si formano vescicole multilamellari (MLV), che possono essere ridotte in dimensioni per produrre vescicole unilamellari piccole o grandi (LUV e SUV) tramite estrusione attraverso membrane o tramite sonicazione della MLV di partenza.

Doppia Emulsione

Le formulazioni lipidiche possono anche essere preparate tramite la tecnica della Doppia Emulsione, che prevede la dissoluzione dei lipidi in una miscela di acqua/solvente organico. La soluzione organica, contenente goccioline d'acqua, viene miscelata con un eccesso di mezzo acquoso, portando alla formazione di una doppia emulsione acqua-in-olio-in-acqua (W/O/W). Dopo un'agitazione meccanica vigorosa, parte delle goccioline d'acqua collassa, dando origine a Grandi Vescicole Unilamellari (LUV).

Evaporazione in fase inversa

Il metodo Reverse Phase Evaporation (REV) consente inoltre di ottenere LUV caricati con acido nucleico. In questa tecnica, un sistema bifase è formato dalla dissoluzione di fosfolipidi in solventi organici e tampone acquoso. La sospensione risultante viene quindi sonicata brevemente fino a quando la miscela diventa una dispersione monofase limpida. La formulazione lipidica è ottenuta dopo l'evaporazione del solvente organico a pressione ridotta. Questa tecnica è stata utilizzata per incapsulare diverse molecole idrofile grandi e piccole, compresi gli acidi nucleici.

Preparazione microfluidica

Il metodo microfluidico, a differenza di altre tecniche di massa, offre la possibilità di controllare il processo di idratazione dei lipidi. Il metodo può essere classificato in microfluidico a flusso continuo e microfluidico basato su goccioline, in base al modo in cui il flusso viene manipolato. Nel metodo di focalizzazione idrodinamica microfluidica (MHF), che opera in modalità di flusso continuo, i lipidi vengono disciolti in alcol isopropilico che viene focalizzato idrodinamicamente in una giunzione incrociata microcanale tra due flussi tampone acquosi. La dimensione delle vescicole può essere controllata modulando le velocità di flusso, controllando così il processo di diluizione della soluzione lipidica/tampone. Il metodo può essere utilizzato per produrre formulazioni lipidiche di oligonucleotidi (ON) utilizzando un dispositivo microfluidico costituito da tre porte di ingresso e una di uscita.

Centrifugazione asimmetrica doppia

La centrifugazione asimmetrica duale (DAC) differisce dalla centrifugazione più comune in quanto utilizza una rotazione aggiuntiva attorno al proprio asse verticale. Si ottiene un'omogeneizzazione efficiente grazie ai due movimenti sovrapposti generati: il campione viene spinto verso l'esterno, come in una normale centrifuga, e poi viene spinto verso il centro della fiala grazie alla rotazione aggiuntiva. Miscelando lipidi e una soluzione di NaCl si ottiene un gel fosfolipidico vescicolare viscoso (VPC), che viene poi diluito per ottenere una dispersione di formulazione lipidica. La dimensione della formulazione lipidica può essere regolata ottimizzando la velocità DAC, la concentrazione lipidica e il tempo di omogeneizzazione.

Iniezione di etanolo

Il metodo Ethanol Injection (EI) può essere utilizzato per l'incapsulamento degli acidi nucleici. Questo metodo prevede l'iniezione rapida di una soluzione etanolica, in cui i lipidi sono disciolti, in un mezzo acquoso contenente acidi nucleici da incapsulare, tramite l'uso di un ago. Le vescicole si formano spontaneamente quando i fosfolipidi sono dispersi nel mezzo.

Dialisi con detergente

Il metodo di dialisi detergente può essere utilizzato per incapsulare acidi nucleici. In breve, lipidi e plasmidi vengono solubilizzati in una soluzione detergente di forza ionica appropriata, dopo aver rimosso il detergente tramite dialisi, si forma una formulazione lipidica stabilizzata. L'acido nucleico non incapsulato viene quindi rimosso tramite cromatografia a scambio ionico e le vescicole vuote tramite centrifugazione a gradiente di densità di saccarosio. La tecnica è altamente

sensibile al contenuto di lipidi cationici e alla concentrazione salina del tampone di dialisi, e il metodo è anche difficile da scalare.

Formazione spontanea di vescicole mediante diluizione dell'etanolo

Le formulazioni lipidiche stabili possono anche essere prodotte attraverso il metodo di formazione spontanea di vescicole mediante diluizione con etanolo, in cui una diluizione graduale o goccia a goccia dell'etanolo fornisce la formazione istantanea di vescicole caricate con acido nucleico mediante l'aggiunta controllata di lipidi disciolti in etanolo a un tampone acquoso a rapida miscelazione contenente l'acido nucleico.

Incapsulamento in liposomi preformati

L'intrappolamento degli acidi nucleici può essere ottenuto anche a partire da liposomi preformati attraverso due metodi diversi: (1) una semplice miscelazione di liposomi cationici con acidi nucleici che dà complessi elettrostatici chiamati "lipoplex", dove possono essere utilizzati con successo per trasfettare colture cellulari, ma sono caratterizzati dalla loro bassa efficienza di incapsulamento e scarse prestazioni in vivo; e (2) una destabilizzazione liposomiale, aggiungendo lentamente etanolo assoluto a una sospensione di vescicole cationiche fino a una concentrazione del 40% v/v seguita dall'aggiunta goccia a goccia di acidi nucleici ottenendo vescicole caricate; tuttavia, i due passaggi principali che caratterizzano il processo di incapsulamento sono troppo sensibili e le particelle devono essere ridimensionate.

Eccipienti

Le composizioni farmaceutiche qui divulgate possono essere formulate utilizzando uno o più eccipienti per: (1) aumentare la stabilità; (2) aumentare la trasfezione cellulare; (3) consentire un rilascio prolungato o ritardato (ad esempio, da una formulazione di deposito del polinucleotide, del costrutto primario o dell'RNA); (4) alterare la biodistribuzione (ad esempio, indirizzare il polinucleotide, il costrutto primario o l'RNA a tessuti o tipi di cellule specifici); (5) aumentare la traduzione della proteina codificata in vivo; e/o (6) alterare il profilo di rilascio della proteina codificata in vivo.

Le composizioni farmaceutiche descritte nel presente documento possono essere preparate con qualsiasi metodo noto o sviluppato in seguito nell'arte della farmacologia. In generale, tali metodi preparatori includono la fase di associazione del principio attivo (ad esempio, acido nucleico) con un eccipiente e/o uno o più altri ingredienti accessori. Una composizione farmaceutica conforme alla presente informativa può essere preparata, confezionata e/o venduta in grandi quantità, come dose unitaria singola e/o come una pluralità di dosi unitarie singole.

Le composizioni farmaceutiche possono inoltre comprendere un eccipiente farmaceuticamente accettabile che, come qui utilizzato, include, ma non è limitato a, tutti i solventi, mezzi di dispersione, diluenti o altri veicoli liquidi, ausiliari di dispersione o sospensione, agenti tensioattivi, agenti isotonici, agenti addensanti o emulsionanti, conservanti e simili, come adatto alla particolare forma di dosaggio desiderata.

Oltre agli eccipienti tradizionali quali tutti i solventi, mezzi di dispersione, diluenti o altri veicoli liquidi, ausili di dispersione o sospensione, agenti tensioattivi, agenti isotonici, agenti addensanti o emulsionanti, conservanti, gli eccipienti della presente informativa possono includere, senza limitazioni, liposomi, nanoparticelle lipidiche, polimeri, lipoplessi, nanoparticelle core-shell, peptidi, proteine, cellule transfettate con costrutto di DNA primario o RNA (ad esempio, per il trapianto in un soggetto), ialuronidasi, imitatori di nanoparticelle e combinazioni degli stessi.

Di conseguenza, le composizioni farmaceutiche descritte nel presente documento possono includere uno o più eccipienti, ciascuno in una quantità che insieme aumenta la stabilità dell'acido nucleico nella formulazione lipidica, aumenta la trasfezione cellulare da parte dell'acido nucleico, aumenta l'espressione della proteina codificata e/o altera il profilo di rilascio delle proteine codificate. Inoltre, l'RNA della presente divulgazione può essere formulato utilizzando nanoparticelle di acido nucleico autoassemblate.

Vari eccipienti per la formulazione di composizioni farmaceutiche e tecniche per la preparazione della composizione sono noti nell'arte (vedere Remington: The Science and Practice of Pharmacy, 21a edizione, AR Gennaro, Lippincott, Williams & Wilkins, Baltimora, Md., 2006; incorporato nel presente documento per riferimento nella sua interezza). L'uso di un mezzo eccipiente convenzionale può essere contemplato nell'ambito delle forme di realizzazione della presente divulgazione, eccetto nella misura in cui qualsiasi mezzo eccipiente convenzionale può essere incompatibile con una sostanza o i suoi derivati, ad esempio producendo qualsiasi effetto biologico indesiderato o interagendo in altro modo in modo deleterio con qualsiasi altro componente della composizione farmaceutica.

Le composizioni farmaceutiche di questa divulgazione possono inoltre contenere come sostanze trasportatrici farmaceuticamente accettabili, come richiesto per approssimare le condizioni fisiologiche, come agenti di regolazione del pH e tamponanti, agenti di regolazione della tonicità e agenti bagnanti, ad esempio, acetato di sodio, lattato di sodio, cloruro di sodio, cloruro di potassio, cloruro di calcio, monolaurato di sorbitano, oleato di trietanolammina e miscele degli stessi. Per le composizioni solide, possono essere utilizzati trasportatori farmaceuticamente accettabili non tossici convenzionali che includono, ad esempio, gradi farmaceutici di mannitolo, lattosio, amido, stearato di magnesio, saccarina sodica, talco, cellulosa, glucosio, saccarosio, carbonato di magnesio e simili.

In alcune forme di realizzazione della divulgazione, la formulazione RNA-lipide può essere somministrata in una formulazione a rilascio temporizzato, ad esempio in una composizione che include un polimero a rilascio lento. L'agente attivo può essere preparato con vettori che proteggeranno dal rilascio rapido, ad esempio un veicolo a rilascio controllato come un polimero, un sistema di rilascio microincapsulato o un gel bioadesivo. Il rilascio prolungato dell'RNA, in varie composizioni della divulgazione, può essere ottenuto includendo nella composizione agenti che ritardano l'assorbimento, ad esempio, idrogel di monostearato di alluminio e gelatina.

Metodi per indurre risposte immunitarie

In alcune forme di realizzazione, sono forniti qui metodi per indurre una risposta immunitaria in un soggetto. Qualsiasi tipo di risposta immunitaria può essere indotta utilizzando i metodi forniti

qui, comprese le risposte immunitarie adattative e innate. In un aspetto, le risposte immunitarie indotte utilizzando i metodi forniti qui includono una risposta anticorpale, una risposta immunitaria cellulare o sia una risposta anticorpale che una risposta immunitaria cellulare.

I metodi di induzione di una risposta immunitaria qui forniti includono la somministrazione a un soggetto di una quantità efficace di qualsiasi molecola di acido nucleico qui fornita. In un aspetto, i metodi di induzione di una risposta immunitaria includono la somministrazione a un soggetto di una quantità efficace di qualsiasi composizione comprendente una molecola di acido nucleico e un lipide qui forniti. In un altro aspetto, i metodi di induzione di una risposta immunitaria includono la somministrazione a un soggetto di una quantità efficace di qualsiasi composizione farmaceutica comprendente una molecola di acido nucleico e una formulazione lipidica qui fornita. In alcuni aspetti, le molecole di acido nucleico, le composizioni e la composizione farmaceutica qui fornite sono vaccini che possono suscitare una risposta immunitaria protettiva o terapeutica, ad esempio.

Come utilizzato nel presente documento, il termine "soggetto" si riferisce a qualsiasi individuo o paziente su cui vengono eseguiti i metodi qui divulgati. Il termine "soggetto" può essere utilizzato in modo intercambiabile con il termine "individuo" o "paziente". Il soggetto può essere un essere umano, sebbene il soggetto possa essere un animale, come sarà apprezzato da coloro che sono nel settore. Pertanto, altri animali, inclusi mammiferi come roditori (inclusi topi, ratti, criceti e porcellini d'India), gatti, cani, conigli, animali da fattoria tra cui mucche, cavalli, capre, pecore, maiali, ecc. e primati (inclusi scimmie, scimpanzé, oranghi e gorilla) sono inclusi nella definizione di soggetto. Come utilizzato nel presente documento, il termine "quantità efficace" o "quantità terapeutamente efficace" si riferisce a quella quantità di una molecola di acido nucleico, composizione o composizione farmaceutica qui descritta che è sufficiente per effettuare l'applicazione prevista, inclusa ma non limitata all'induzione di una risposta immunitaria e/o trattamento di una malattia, come definito nel presente documento. La quantità terapeutamente efficace può variare a seconda dell'applicazione prevista (ad esempio, induzione di una risposta immunitaria, trattamento, applicazione in vivo), o del soggetto o paziente e della condizione della malattia in cura, ad esempio, il peso e l'età del soggetto, la specie, la gravità della condizione della malattia, la modalità di somministrazione e simili, che possono essere facilmente determinati da una persona con normale competenza nell'arte. Il termine si applica anche a una dose che indurrà una particolare risposta in una cellula bersaglio. La dose specifica varierà a seconda della particolare molecola di acido nucleico, composizione o composizione farmaceutica scelta, del regime di dosaggio da seguire, se viene somministrata in combinazione con altri composti, tempi di somministrazione, il tessuto a cui viene somministrata e il sistema di somministrazione fisica in cui viene trasportata.

Le dosi esemplari di molecole di acido nucleico che possono essere somministrate includono circa 0,01 µg, circa 0,02 µg, circa 0,03 µg, circa 0,04 µg, circa 0,05 µg, circa 0,06 µg, circa 0,07 µg, circa 0,08 µg, circa 0,09 µg, circa 0,1 µg, circa 0,2 µg, circa 0,3 µg, circa 0,4 µg, circa 0,5 µg, circa 0,6 µg, circa 0,7 µg, circa 0,8 µg, circa 0,9 µg, circa 1,0 µg, circa 1,5 µg, circa 2,0 µg, circa 2,5 µg, circa 3,0 µg, circa 3,5 µg, circa 4,0 µg, circa 4,5 µg, circa 5,0 µg, circa 5,5 µg, circa 6,0 µg, circa 6,5 µg, circa 7,0 µg, circa 7,5 µg, circa 8,0 µg, circa 8,5 µg, circa 9,0 µg, circa 9,5 µg, circa 10 µg, circa 11 µg, circa 12 µg, circa 13 µg, circa 14 µg, circa 15 µg, circa 16 µg, circa 17 µg, circa 18 µg, circa 19 µg, circa 20 µg, circa 21 µg, circa 22 µg, circa 23 µg, circa 24 µg, circa 25 µg, circa 26 µg, circa 27 µg, circa 28 µg, circa 29 µg,

circa 30 µg, circa 35 µg, circa 40 µg, circa 45 µg, circa 50 µg, circa 55 µg, circa 60 µg, circa 65 µg, circa 70 µg, circa 75 µg, circa 80 µg, circa 85 µg, circa 90 µg, circa 95 µg, circa 100 µg, circa 125 µg, circa 150 µg, circa 175 µg, circa 200 µg, circa 250 µg, circa 300 µg, circa 350 µg, circa 400 µg, circa 450 µg, circa 500 µg, circa 600 µg, circa 700 µg, circa 800 µg, circa 900 µg, circa 1.000 µg, o più, e qualsiasi numero o intervallo intermedio. In un aspetto, le molecole di acido nucleico sono molecole di RNA. In un altro aspetto, le molecole di acido nucleico sono molecole di DNA. Le molecole di acido nucleico possono avere un dosaggio unitario comprendente circa 0,01 µg a circa 1.000 µg o più di acido nucleico in una singola dose.

In alcuni aspetti, le composizioni fornite nel presente documento che possono essere somministrate includono circa 0,01 µg, circa 0,02 µg, circa 0,03 µg, circa 0,04 µg, circa 0,05 µg, circa 0,06 µg, circa 0,07 µg, circa 0,08 µg, circa 0,09 µg, circa 0,1 µg, circa 0,2 µg, circa 0,3 µg, circa 0,4 µg, circa 0,5 µg, circa 0,6 µg, circa 0,7 µg, circa 0,8 µg, circa 0,9 µg, circa 1,0 µg, circa 1,5 µg, circa 2,0 µg, circa 2,5 µg, circa 3,0 µg, circa 3,5 µg, circa 4,0 µg, circa 4,5 µg, circa 5,0 µg, circa 5,5 µg, circa 6,0 µg, circa 6,5 µg, circa 7,0 µg, circa 7,5 µg, circa 8,0 µg, circa 8,5 µg, circa 9,0 µg, circa 9,5 µg, circa 10 µg, circa 11 µg, circa 12 µg, circa 13 µg, circa 14 µg, circa 15 µg, circa 16 µg, circa 17 µg, circa 18 µg, circa 19 µg, circa 20 µg, circa 21 µg, circa 22 µg, circa 23 µg, circa 24 µg, circa 25 µg, circa 26 µg, circa 27 µg, circa 28 µg, circa 29 µg, circa 30 µg, circa 35 µg, circa 40 µg, circa 45 µg, circa 50 µg, circa 55 µg, circa 60 µg, circa 65 µg, circa 70 µg, circa 75 µg, circa 80 µg, circa 85 µg, circa 90 µg, circa 95 µg, circa 100 µg, circa 125 µg, circa 150 µg, circa 175 µg, circa 200 µg, circa 250 µg, circa 300 µg, circa 350 µg, circa 400 µg, circa 450 µg, circa 500 µg, circa 600 µg, circa 700 µg, circa 800 µg, circa 900 µg, circa 1.000 µg, o più, e qualsiasi numero o intervallo intermedio, acido nucleico e lipidi. In altri aspetti, le composizioni farmaceutiche fornite nel presente documento che possono essere somministrate includono circa 0,01 µg, circa 0,02 µg, circa 0,03 µg, circa 0,04 µg, circa 0,05 µg, circa 0,06 µg, circa 0,07 µg, circa 0,08 µg, circa 0,09 µg, circa 0,1 µg, circa 0,2 µg, circa 0,3 µg, circa 0,4 µg, circa 0,5 µg, circa 0,6 µg, circa 0,7 µg, circa 0,8 µg, circa 0,9 µg, circa 1,0 µg, circa 1,5 µg, circa 2,0 µg, circa 2,5 µg, circa 3,0 µg, circa 3,5 µg, circa 4,0 µg, circa 4,5 µg, circa 5,0 µg, circa 5,5 µg, circa 6,0 µg, circa 6,5 µg, circa 7,0 µg, circa 7,5 µg, circa 8,0 µg, circa 8,5 µg, circa 9,0 µg, circa 9,5 µg, circa 10 µg, circa 11 µg, circa 12 µg, circa 13 µg, circa 14 µg, circa 15 µg, circa 16 µg, circa 17 µg, circa 18 µg, circa 19 µg, circa 20 µg, circa 21 µg, circa 22 µg, circa 23 µg, circa 24 µg, circa 25 µg, circa 26 µg, circa 27 µg, circa 28 µg, circa 29 µg, circa 30 µg, circa 35 µg, circa 40 µg, circa 45 µg, circa 50 µg, circa 55 µg, circa 60 µg, circa 65 µg, circa 70 µg, circa 75 µg, circa 80 µg, circa 85 µg, circa 90 µg, circa 95 µg, circa 100 µg, circa 125 µg, circa 150 µg, circa 175 µg, circa 200 µg, circa 250 µg, circa 300 µg, circa 350 µg, circa 400 µg, circa 450 µg, circa 500 µg, circa 600 µg, circa 700 µg, circa 800 µg, circa 900 µg, circa 1.000 µg, o più, e qualsiasi numero o intervallo intermedio, formulazione di acidi nucleici e lipidi.

In un aspetto, le composizioni fornite nel presente documento possono avere un dosaggio unitario comprendente circa 0,01 µg a circa 1.000 µg o più di acido nucleico e lipidi in una singola dose. In un altro aspetto, le composizioni farmaceutiche fornite nel presente documento possono avere un dosaggio unitario comprendente circa 0,01 µg a circa 1.000 µg o più di acido nucleico e formulazione lipidica in una singola dose. Un dosaggio unitario di vaccino può corrispondere al dosaggio unitario di molecole di acido nucleico, composizioni o composizioni farmaceutiche fornite nel presente documento e che può essere somministrato a un soggetto. In un aspetto, le

composizioni di vaccino della presente divulgazione hanno un dosaggio unitario comprendente circa 0,01 µg a circa 1.000 µg o più di acido nucleico e formulazione lipidica in una singola dose. In un altro aspetto, le composizioni di vaccino della presente divulgazione hanno un dosaggio unitario comprendente circa 0,01 µg a circa 50 µg di acido nucleico e formulazione lipidica in una singola dose. In un ulteriore aspetto, le composizioni vaccinali della presente invenzione hanno un dosaggio unitario comprendente circa 0,2 µg a circa 20 µg di formulazione di acidi nucleici e lipidi in una singola dose.

Una forma di dosaggio della composizione di questa divulgazione può essere solida, che può essere ricostituita in un liquido prima della somministrazione. Il solido può essere somministrato come polvere. Il solido può essere sotto forma di capsula, compressa o gel. In alcune realizzazioni, la composizione farmaceutica comprende una formulazione lipidica di acido nucleico che è stata liofilizzata. In alcune realizzazioni, la composizione liofilizzata può comprendere uno o più lioprotettori, quali, tra cui, ma non necessariamente limitati a, glucosio, trealosio, saccarosio, maltosio, lattosio, mannitolo, inositolo, idrossipropil-β-ciclodestrina e/o polietilenglicole. In alcune realizzazioni, la composizione liofilizzata comprende un poloxamer, sorbato di potassio, saccarosio o qualsiasi combinazione di questi. In forme di realizzazione specifiche, il polossamero è il polossamero 188. In alcune forme di realizzazione, le composizioni liofilizzate descritte nel presente documento possono comprendere circa lo 0,01-1,0% p/p di un polossamero. In alcune forme di realizzazione, le composizioni liofilizzate descritte nel presente documento possono comprendere circa l'1,0-5,0% p/p di sorbato di potassio. Le percentuali possono essere qualsiasi valore o sottovalore entro gli intervalli indicati, compresi gli endpoint.

In alcune realizzazioni, la composizione liofilizzata può comprendere circa lo 0,01-1,0% p/p della molecola di acido nucleico. In alcune realizzazioni, la composizione può comprendere circa l'1,0-5,0% p/p di lipidi. In alcune realizzazioni, la composizione può comprendere circa lo 0,5-2,5% p/p di tampone TRIS. In alcune realizzazioni, la composizione può comprendere circa lo 0,75-2,75% p/p di NaCl. In alcune realizzazioni, la composizione può comprendere circa l'85-95% p/p di uno zucchero. Le percentuali possono essere qualsiasi valore o sottovalore entro gli intervalli indicati, inclusi gli endpoint.

In una forma di realizzazione preferita, la forma di dosaggio delle composizioni farmaceutiche descritte nel presente documento può essere una sospensione liquida di nanoparticelle lipidiche di RNA autoreplicanti descritte nel presente documento. In alcune forme di realizzazione, la sospensione liquida è in una soluzione tamponata. In alcune forme di realizzazione, la soluzione tamponata comprende un tampone selezionato dal gruppo costituito da HEPES, MOPS, TES e TRIS. In alcune forme di realizzazione, il tampone ha un pH di circa 7,4. In alcune forme di realizzazione preferite, il tampone è HEPES. In alcune ulteriori forme di realizzazione, la soluzione tamponata comprende inoltre un crioprotettore. In alcune forme di realizzazione, il crioprotettore è selezionato da uno zucchero e glicerolo o una combinazione di uno zucchero e glicerolo. In alcune forme di realizzazione, lo zucchero è uno zucchero dimerico. In alcune forme di realizzazione, lo zucchero è saccarosio. In alcune forme di realizzazione preferite, il tampone comprende HEPES, saccarosio e glicerolo a un pH di 7,4. In alcune realizzazioni, la composizione comprende un tampone HEPES, MOPS, TES o TRIS a un pH di circa 7,0 a circa 8,5. In alcune realizzazioni, il tampone HEPES, MOPS, TES o TRIS può essere a una concentrazione che va da 7 mg/ml a circa 15

mg/ml. Il pH o la concentrazione possono essere qualsiasi valore o sottovalore entro gli intervalli recitati, inclusi gli endpoint.

In alcune realizzazioni, la sospensione viene congelata durante la conservazione e scongelata prima della somministrazione. In alcune realizzazioni, la sospensione viene congelata a una temperatura inferiore a circa 70° C. In alcune realizzazioni, la sospensione viene diluita con acqua sterile durante la somministrazione endovenosa. In alcune realizzazioni, la somministrazione endovenosa comprende la diluizione della sospensione con circa 2 volumi a circa 6 volumi di acqua sterile. In alcune realizzazioni, la sospensione comprende circa 0,1 mg a circa 3,0 mg di RNA autoreplicante/mL, circa 15 mg/mL a circa 25 mg/mL di un lipide cationico ionizzabile, circa 0,5 mg/mL a circa 2,5 mg/mL di un lipide PEG, circa 1,8 mg/mL a circa 3,5 mg/mL di un lipide helper, circa 4,5 mg/mL a circa 7,5 mg/mL di un colesterolo, circa 7 mg/mL a circa 15 mg/mL di un tampone, circa 2,0 mg/mL a circa 4,0 mg/mL di NaCl, circa 70 mg/mL a circa 110 mg/mL di saccarosio e circa 50 mg/mL a circa 70 mg/mL di glicerolo. In alcune realizzazioni, una formulazione di nanoparticelle di RNA-lipidi autoreplicanti liofilizzate può essere risospesa in un tampone come descritto nel presente documento.

In alcune realizzazioni, le composizioni della presente informativa vengono somministrate a un soggetto in modo tale che una concentrazione di RNA autoreplicante di almeno circa 0,05 mg/kg, almeno circa 0,1 mg/kg, almeno circa 0,5 mg/kg, almeno circa 1,0 mg/kg, almeno circa 2,0 mg/kg, almeno circa 3,0 mg/kg, almeno circa 4,0 mg/kg, almeno circa 5,0 mg/kg di peso corporeo venga somministrata in una singola dose o come parte di un singolo ciclo di trattamento. In alcune realizzazioni, le composizioni della presente informativa vengono somministrate a un soggetto in modo tale che una quantità totale di almeno circa 0,1 mg, almeno circa 0,5 mg, almeno circa 1,0 mg, almeno circa 2,0 mg, almeno circa 3,0 mg, almeno circa 4,0 mg, almeno circa 5,0 mg, almeno circa 6,0 mg, almeno circa 7,0 mg, almeno circa 8,0 mg, almeno circa 9,0 mg, almeno circa 10 mg, almeno circa 15 mg, almeno circa 20 mg, almeno circa 25 mg, almeno circa 30 mg, almeno circa 35 mg, almeno circa 40 mg, almeno circa 45 mg, almeno circa 50 mg, almeno circa 55 mg, almeno circa 60 mg, almeno circa 65 mg, almeno circa 70 mg, almeno circa 75 mg, almeno circa 80 mg, almeno circa 85 mg, almeno circa 90 mg, almeno circa 95 mg, almeno circa 100 mg, almeno circa 105 mg, almeno circa 110 mg, almeno circa 115 mg, almeno circa 120 mg o almeno circa 125 mg di RNA autoreplicante vengono somministrati in una o più dosi fino a una dose massima di circa 300 mg, circa 350 mg, circa 400 mg, circa 450 mg o circa 500 mg di RNA autoreplicante.

Qualsiasi via di somministrazione può essere inclusa nei metodi forniti nel presente documento. In alcuni aspetti, le molecole di acido nucleico, le composizioni e le composizioni farmaceutiche fornite nel presente documento vengono somministrate per via intramuscolare, sottocutanea, intradermica, transdermica, intranasale, orale, sublinguale, endovenosa, intraperitoneale, topica, per aerosol o per via polmonare, come per inalazione o nebulizzazione, ad esempio. In alcune forme di realizzazione, le composizioni farmaceutiche descritte vengono somministrate per via sistemica. Le vie di somministrazione idonee includono, ad esempio, la somministrazione orale, rettale, vaginale, transmucosale, polmonare, inclusa la somministrazione intratracheale o inalatoria, o intestinale; somministrazione parenterale, inclusa l'iniezione intradermica, transdermica (topica), intramuscolare, sottocutanea, intramidollare, nonché intratecale, intraventricolare diretta, endovenosa, intraperitoneale o intranasale. In particolari realizzazioni, la

somministrazione intramuscolare avviene a un muscolo selezionato dal gruppo costituito da muscolo scheletrico, muscolo liscio e muscolo cardiaco. In alcune realizzazioni, la composizione farmaceutica viene somministrata per via endovenosa.

Le composizioni farmaceutiche possono essere somministrate a qualsiasi tessuto desiderato. In alcune realizzazioni, l'RNA autoreplicante somministrato è espresso in un tessuto diverso dal tessuto in cui è stata somministrata la formulazione lipidica o la composizione farmaceutica. In realizzazioni preferite, l'RNA autoreplicante è somministrato ed espresso nel fegato.

In altri aspetti, le molecole di acido nucleico, le composizioni e le composizioni farmaceutiche qui fornite vengono somministrate per via intramuscolare.

In alcuni aspetti, il soggetto in cui viene indotta una risposta immunitaria è un soggetto sano. Come utilizzato nel presente documento, il termine "soggetto sano" si riferisce a un soggetto che non ha una condizione o una malattia, tra cui una malattia infettiva o un cancro, ad esempio, o che non ha una condizione o una malattia contro cui viene indotta una risposta immunitaria. Di conseguenza, in alcuni aspetti, una molecola di acido nucleico, una composizione o una composizione farmaceutica qui fornita viene somministrata a scopo profilattico per prevenire una malattia infettiva o un cancro, ad esempio. In altri aspetti, il soggetto in cui viene indotta una risposta immunitaria ha il cancro. Il soggetto può soffrire di qualsiasi cancro o avere qualsiasi tumore, inclusi tumori solidi e liquidi. In un aspetto, il cancro è cancro al rene, cancro renale, cancro alla vescica urinaria, cancro alla prostata, cancro all'utero, cancro al seno, cancro cervicale, cancro alle ovaie, cancro ai polmoni, cancro al fegato, cancro allo stomaco, cancro al colon, cancro al retto, cancro alla cavità orale, cancro alla faringe, cancro al pancreas, cancro alla tiroide, melanoma, cancro alla pelle, cancro alla testa e al collo, cancro al cervello, cancro ematopoietico, leucemia, linfoma, cancro alle ossa o sarcoma. Di conseguenza, una molecola di acido nucleico, una composizione o una composizione farmaceutica qui fornita può essere somministrata terapeuticamente, vale a dire, per trattare una condizione o una malattia, come il cancro, dopo l'insorgenza della condizione o della malattia.

Come qui utilizzati, i termini "trattare", "trattamento", "terapia", "terapeutico" e simili si riferiscono all'ottenimento di un effetto farmacologico e/o fisiologico desiderato, inclusi, ma non limitati a, alleviare, ritardare o rallentare la progressione, ridurre gli effetti o i sintomi, prevenire l'insorgenza, inibire, migliorare l'insorgenza di una malattia o di un disturbo, ottenere un risultato benefico o desiderato rispetto a una malattia, un disturbo o una condizione medica, come un beneficio terapeutico e/o un beneficio profilattico. "Trattamento", come qui utilizzato, include qualsiasi trattamento di una malattia in un mammifero, in particolare in un essere umano, e include: (a) prevenire che la malattia si verifichi in un soggetto, incluso un soggetto che è predisposto alla malattia o a rischio di contrarre la malattia ma non è stato ancora diagnosticato come affetto; (b) inibire la malattia, ovvero arrestarne lo sviluppo; e (c) alleviare la malattia, ovvero causare la regressione della malattia. Un beneficio terapeutico include l'eradicazione o il miglioramento del disturbo sottostante in cura. Inoltre, un beneficio terapeutico si ottiene con l'eradicazione o il miglioramento di uno o più sintomi fisiologici associati al disturbo sottostante, in modo tale che si osservi un miglioramento nel soggetto, nonostante il soggetto possa ancora essere afflitto dal disturbo sottostante. In alcuni aspetti, per un beneficio profilattico, il trattamento o le

composizioni per il trattamento, comprese le composizioni farmaceutiche, vengono somministrati a un soggetto a rischio di sviluppare una particolare malattia, o a un soggetto che segnala uno o più sintomi fisiologici di una malattia, anche se una diagnosi di questa malattia potrebbe non essere stata fatta. I metodi della presente divulgazione possono essere utilizzati con qualsiasi mammifero o altro animale. In alcuni aspetti, il trattamento determina una diminuzione o cessazione dei sintomi. Un effetto profilattico include il ritardo o l'eliminazione della comparsa di una malattia o condizione, il ritardo o l'eliminazione dell'insorgenza dei sintomi di una malattia o condizione, il rallentamento, l'arresto o l'inversione della progressione di una malattia o condizione, o qualsiasi combinazione di questi.

Le molecole di acido nucleico, le composizioni e le composizioni farmaceutiche qui fornite possono essere somministrate una o più volte. Di conseguenza, le molecole di acido nucleico, le composizioni e le composizioni farmaceutiche qui fornite possono essere somministrate una, due, tre, quattro, cinque, sei, sette, otto, nove, dieci o più volte. Il tempo tra due o più somministrazioni può essere una settimana, due settimane, tre settimane, quattro settimane, cinque settimane, sei settimane, sette settimane, otto settimane, nove settimane, settimane, dieci settimane, 11 settimane, 12 settimane, 13 settimane, 14 settimane, 15 settimane, 16 settimane, 17 settimane, 18 settimane, 19 settimane, 20 settimane, 21 settimane, 22 settimane, 23 settimane, 24 settimane, 25 settimane, 26 settimane, 27 settimane, 28 settimane, 29 settimane, 30 settimane, 31 settimane, 32 settimane, 33 settimane, 34 settimane, 35 settimane, 36 settimane, 37 settimane, 38 settimane, 39 settimane, 40 settimane, 41 settimane, 42 settimane, 43 settimane, 44 settimane, 45 settimane, 46 settimane, 47 settimane, 48 settimane, 49 settimane, 50 settimane, 51 settimane, 52 settimane o più settimane e qualsiasi numero o intervallo intermedio. In alcuni aspetti, la tempistica tra due o più somministrazioni è di un mese, due mesi, tre mesi, quattro mesi, cinque mesi, sei mesi, sette mesi, otto mesi, nove mesi, dieci mesi, 11 mesi, 12 mesi, 13 mesi, 14 mesi, 15 mesi, 16 mesi, 17 mesi, 18 mesi, 19 mesi, 20 mesi, 21 mesi, 22 mesi, 23 mesi, 24 mesi o più mesi e qualsiasi numero o intervallo intermedio. In altri aspetti, la tempistica tra due o più somministrazioni può essere di un anno, due anni, tre anni, quattro anni, cinque anni, sei anni, sette anni, otto anni, nove anni, dieci anni o più anni e qualsiasi numero o intervallo intermedio. La tempistica tra la prima e qualsiasi somministrazione successiva può essere la stessa o diversa. In un aspetto, le molecole di acido nucleico, le composizioni o le composizioni farmaceutiche qui fornite vengono somministrate una volta.

Più di una molecola di acido nucleico, composizione o composizione farmaceutica può essere somministrata nei metodi forniti nel presente documento. In un aspetto, due o più molecole di acido nucleico, composizioni o composizioni farmaceutiche fornite nel presente documento vengono somministrate simultaneamente. In un altro aspetto, due o più molecole di acido nucleico, composizioni o composizioni farmaceutiche fornite nel presente documento vengono somministrate in sequenza. Le somministrazioni simultanee e sequenziali possono includere qualsiasi numero e qualsiasi combinazione di molecole di acido nucleico, composizioni o composizioni farmaceutiche fornite nel presente documento. Più molecole di acido nucleico, composizioni o composizioni farmaceutiche che vengono somministrate insieme o in sequenza possono includere transgeni che codificano diverse proteine antigeniche o frammenti delle stesse. In questo modo, possono essere indotte risposte immunitarie contro diversi bersagli antigenici.

Due, tre, quattro, cinque, sei, sette, otto, nove, dieci o più molecole di acido nucleico, composizioni o composizioni farmaceutiche che includono transgeni che codificano diverse proteine antigeniche o frammenti delle stesse possono essere somministrate simultaneamente o in sequenza. Qualsiasi combinazione di molecole di acido nucleico, composizioni e composizioni farmaceutiche, inclusa qualsiasi combinazione di transgeni, può essere somministrata simultaneamente o sequenzialmente. In alcuni aspetti, la somministrazione è simultanea. In altri aspetti, la somministrazione è sequenziale. Il tempo tra due o più somministrazioni può essere una settimana, due settimane, tre settimane, quattro settimane, cinque settimane, sei settimane, sette settimane, otto settimane, nove settimane, settimane, dieci settimane, 11 settimane, 12 settimane, 13 settimane, 14 settimane, 15 settimane, 16 settimane, 17 settimane, 18 settimane, 19 settimane, 20 settimane, 21 settimane, 22 settimane, 23 settimane, 24 settimane, 25 settimane, 26 settimane, 27 settimane, 28 settimane, 29 settimane, 30 settimane, 31 settimane, 32 settimane, 33 settimane, 34 settimane, 35 settimane, 36 settimane, 37 settimane, 38 settimane, 39 settimane, 40 settimane, 41 settimane, 42 settimane, 43 settimane, 44 settimane, 45 settimane, 46 settimane, 47 settimane, 48 settimane, 49 settimane, 50 settimane, 51 settimane, 52 settimane o più settimane e qualsiasi numero o intervallo intermedio. In alcuni aspetti, la tempistica tra due o più somministrazioni è di un mese, due mesi, tre mesi, quattro mesi, cinque mesi, sei mesi, sette mesi, otto mesi, nove mesi, dieci mesi, 11 mesi, 12 mesi, 13 mesi, 14 mesi, 15 mesi, mesi, 16 mesi, 17 mesi, 18 mesi, 19 mesi, 20 mesi, 21 mesi, 22 mesi, 23 mesi, 24 mesi o più mesi e qualsiasi numero o intervallo intermedio. In altri aspetti, la tempistica tra due o più somministrazioni può essere di un anno, due anni, tre anni, quattro anni, cinque anni, sei anni, sette anni, otto anni, nove anni, dieci anni o più anni e qualsiasi numero o intervallo intermedio. La tempistica tra la prima e qualsiasi successiva somministrazione può essere la stessa o diversa. Le molecole di acido nucleico, le composizioni e le composizioni farmaceutiche qui fornite possono essere somministrate con qualsiasi altro vaccino o trattamento.

Dopo la somministrazione della composizione al soggetto, il prodotto proteico codificato dall'RNA autoreplicante della divulgazione (ad esempio, un antigene) è rilevabile nei tessuti bersaglio per almeno circa uno-sette giorni o più. La quantità di prodotto proteico necessaria per ottenere un effetto terapeutico varierà a seconda del titolo anticorpale necessario per generare un'immunità al COVID-19 nel paziente. Ad esempio, il prodotto proteico può essere rilevabile nei tessuti bersaglio a una concentrazione (ad esempio, una concentrazione terapeutica) di almeno circa 0,025-1,5 µg/ml (ad esempio, almeno circa 0,050 µg/ml, almeno circa 0,075 µg/ml, almeno circa 0,1 µg/ml, almeno circa 0,2 µg/ml, almeno circa 0,3 µg/ml, almeno circa 0,4 µg/ml, almeno circa 0,5 µg/ml, almeno circa 0,6 µg/ml, almeno circa 0,7 µg/ml, almeno circa 0,8 µg/ml, almeno circa 0,9 µg/ml, almeno circa 1,0 µg/ml, almeno circa 1,1 µg/ml, almeno circa 1,2 µg/ml, almeno circa 1,3 µg/ml, almeno circa 1,4 µg/ml o almeno circa 1,5 µg/ml), per almeno circa 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45 giorni o più dopo la somministrazione della composizione al soggetto.

In alcune realizzazioni, la composizione descritta nel presente documento può essere somministrata una volta. In alcune realizzazioni, la composizione descritta nel presente documento può essere somministrata due volte.

In alcune realizzazioni, la composizione può essere somministrata sotto forma di dose di richiamo a un soggetto precedentemente vaccinato contro il coronavirus.

In alcune realizzazioni, una composizione farmaceutica della presente informativa viene somministrata a un soggetto una volta al mese. In alcune realizzazioni, una composizione farmaceutica della presente informativa viene somministrata a un soggetto due volte al mese. In alcune realizzazioni, una composizione farmaceutica della presente informativa viene somministrata a un soggetto tre volte al mese. In alcune realizzazioni, una composizione farmaceutica della presente informativa viene somministrata a un soggetto quattro volte al mese.

In alternativa, le composizioni della presente informativa possono essere somministrate in modo locale piuttosto che sistemico, ad esempio, tramite iniezione della composizione farmaceutica direttamente in un tessuto mirato, preferibilmente in una formulazione a rilascio prolungato o depot. La somministrazione locale può essere influenzata in vari modi, a seconda del tessuto da trattare. Ad esempio, gli aerosol contenenti composizioni della presente informativa possono essere inalati (per somministrazione nasale, tracheale o bronchiale); le composizioni della presente informativa possono essere iniettate nel sito di lesione, manifestazione di malattia o dolore, ad esempio; le composizioni possono essere fornite in pastiglie per applicazione orale, tracheale o esofagea; possono essere fornite in forma liquida, compressa o capsula per somministrazione allo stomaco o all'intestino, possono essere fornite in forma di supposta per applicazione rettale o vaginale; o possono anche essere somministrate all'occhio mediante l'uso di creme, gocce o persino iniezione. Le formulazioni contenenti composizioni della presente divulgazione complessate con molecole terapeutiche o ligandi possono anche essere somministrate chirurgicamente, ad esempio in associazione con un polimero o altra struttura o sostanza che può consentire alle composizioni di diffondersi dal sito di impianto alle cellule circostanti. In alternativa, possono essere applicate chirurgicamente senza l'uso di polimeri o supporti.

Combinazioni

L'RNA autoreplicante, le sue formulazioni o le proteine codificate descritte nel presente documento possono essere utilizzate in combinazione con uno o più altri agenti terapeutici, profilattici, diagnostici o di imaging. Con "in combinazione con" non si intende implicare che gli agenti debbano essere somministrati contemporaneamente e/o formulati per la somministrazione insieme, sebbene questi metodi di somministrazione rientrino nell'ambito della presente informativa. Le composizioni possono essere somministrate contemporaneamente, prima o dopo una o più altre terapie o procedure mediche desiderate. In generale, ogni agente verrà somministrato a una dose e/o secondo un programma temporale determinato per quell'agente. Preferibilmente, i metodi di trattamento della presente informativa comprendono la somministrazione di composizioni farmaceutiche, profilattiche, diagnostiche o di imaging in combinazione con agenti che possono migliorare la loro biodisponibilità, ridurre e/o modificare il loro metabolismo, inibire la loro escrezione e/o modificare la loro distribuzione all'interno del corpo. Come esempio non limitativo, un RNA autoreplicante della divulgazione può essere utilizzato in combinazione con un agente farmaceutico per immunizzare o vaccinare un soggetto. In generale, ci si aspetta che gli agenti utilizzati in combinazione con l'RNA autoreplicante attualmente divulgato e le relative formulazioni siano utilizzati a livelli che non superino i livelli a cui sono utilizzati individualmente. In alcune realizzazioni, i livelli utilizzati in combinazione saranno inferiori a quelli utilizzati individualmente. In una realizzazione, le combinazioni,

ciascuna o insieme, possono essere somministrate secondo i regimi di dosaggio frazionati come sono noti nell'arte.

Intervalli

In questa informativa, vari aspetti possono essere presentati in formato intervallo. Si deve comprendere che qualsiasi descrizione in formato intervallo è solo per comodità e brevità e non intende essere limitativa. Di conseguenza, la descrizione di un intervallo deve essere considerata come avente divulgato specificamente tutti i possibili sottointervalli nonché singoli valori numerici all'interno di tale intervallo. Ad esempio, la descrizione di un intervallo come da 1 a 6 deve essere considerata come avente divulgato specificamente sottointervalli come da 1 a 3, da 1 a 4, da 1 a 5, da 2 a 4, da 2 a 6, da 3 a 6, ecc., nonché singoli numeri all'interno di tale intervallo, ad esempio 1, 2, 2,1, 2,2, 2,5, 3, 4, 4,75, 4,8, 4,85, 4,95, 5, 5,5, 5,75, 5,9, 5,00 e 6. Ciò si applica a un intervallo di qualsiasi ampiezza.

Esempio 1

Questo esempio descrive un confronto tra la progettazione e l'espressione delle piattaforme mRNA e RNA autoreplicante (STARR™).

Sia i costrutti del vaccino mRNA che STARR™ sono stati progettati per codificare la proteina S SARS-CoV-2 a lunghezza intera (1273 aa), con l'RNA autoreplicante STARR™ che codifica inoltre per i geni della replicasi del virus dell'encefalite equina venezuelana (VEEV) (FIGURA 1 UNA; Costrutto del vaccino STARR™ corrispondente a un RNA con una sequenza di SEQ ID NO:125, con U al posto di T, qui denominato "RNA STARR™ SARS-CoV-2"; mRNA corrispondente a una sequenza di SEQ ID NO: 126, con U al posto di T e comprendente un virus dell'incisione del tabacco (TEV) 5' UTR, una beta-globina di Xenopus (Xbg) 3' UTR e un open reading frame ottimizzato per codoni che codifica la glicoproteina SARS-CoV-2). Le caratteristiche di questi diversi costrutti sono state studiate per prime. I costrutti sono stati incapsulati nella stessa composizione LNP. In breve, i costrutti di RNA sono stati incapsulati in nanoparticelle lipidiche (LNP) che includevano quattro eccipienti lipidici (un lipide cationico ionizzabile, 1,2-distearoil-sn-glicerolo-3-fosfolina (DSPC), colesterolo e PEG2000-DMG) dispersi in tampone HEPES (pH 8,0) contenente cloruro di sodio e i crioprotettori saccarosio e glicerolo. Il lipide cationico ionizzabile aveva la seguente struttura:

Nonostante le differenze nelle lunghezze dell'RNA dell'mRNA e del costrutto RNA STARR™ SARS-CoV-2, il diametro LNP, l'indice di polidispersità e l'efficienza di intrappolamento dell'RNA erano simili (FIGURA 1). L'espressione in vitro del vaccino mRNA e del costrutto RNA STARR™ SARS-CoV-2 è stata confermata nel lisato cellulare 24 ore dopo la trasfezione attraverso il rilevamento positivo del Western blot della proteina S (FIGURA 1 C). È stato inoltre osservato che sia i vaccini mRNA che STARR™ esprimevano una miscela di proteina S a lunghezza intera e proteina S scissa, ovvero la S era scissa nei domini transmembrana S1 e S2 e nella membrana citoplasmatica (FIGURA 1 C). L'espressione proteica in vivo delle due piattaforme RNA nei topi BALB/c è stata confrontata utilizzando mRNA e costrutti STARR™ che esprimevano un reporter della luciferasi (FIGURA 1). Gli animali iniettati con il costrutto del vaccino mRNA hanno mostrato un'elevata espressione in vivo della luciferasi al giorno 1, sebbene i livelli di espressione siano diminuiti nel tempo. Al contrario,

l'espressione della luciferasi nei topi iniettati con STARR™ ha mostrato segnali sostenuti o addirittura aumentati, a parte quelli a cui è stata somministrata la dose di 0,2 µg, fino al giorno 7 post-inoculazione (FIGURA 1).

Questi dati dimostrano che l'espressione dell'antigene, dose per dose, è stata più prolungata con il vaccino STARR™ rispetto al vaccino mRNA.

Esempio 2

Questo esempio descrive l'espressione genica immunitaria in seguito al costrutto STARR™ e alla vaccinazione con mRNA.

I topi C57BL/6J sono stati inoculati con RNA STARR™ SARS-CoV-2 (che codifica la glicoproteina SARS-CoV-2 come descritto sopra (Esempio 1)) e vaccini mRNA a dosi di 0,2 µg, 2 µg e 10 µg o controllo PBS. Non si è verificata alcuna perdita media significativa del peso degli animali nei primi 4 giorni, ad eccezione di quelli che hanno ricevuto 10 µg di RNA STARR™ SARS-CoV-2 (FIGURA 2 UNA). Tuttavia, a parte la perdita di peso, c'erano pochi altri segni clinici come indicato dalle minime differenze nei punteggi clinici. Sia il peso che i punteggi clinici sono migliorati dopo il terzo giorno dopo la vaccinazione.

È stato precedentemente dimostrato che la risposta immunitaria innata, come la risposta all'interferone di tipo I (IFN), è associata all'immunogenicità del vaccino in seguito alla vaccinazione contro la febbre gialla, ad esempio. Inoltre, è stato dimostrato che le risposte proinfiammatorie guidate dalle specie reattive dell'ossigeno sono alla base degli eventi avversi sistemici nella vaccinazione contro la febbre gialla. Pertanto, l'espressione dei geni immunitari innati e proinfiammatori nel sangue intero di topi C57BL/6 inoculati è stata misurata con PBS, vaccino a mRNA o il costrutto RNA STARR™ SARS-CoV-2. I geni nel percorso IFN di tipo I erano i più altamente espressi negli animali inoculati con RNA STARR™ SARS-CoV-2 rispetto al vaccino a mRNA o al PBS (FIGURA 2EFICO. 8). Al contrario, i geni associati alle risposte pro-infiammatorie sono stati per lo più ridotti in abbondanza a seguito della vaccinazione STARR™ SARS-CoV-2 RNA rispetto al vaccino mRNA o PBS (FIGURA 2EFICO. 8).

Poiché le risposte immunitarie adattative si sviluppano nei centri germinali nei linfonodi drenanti, i linfonodi drenanti sono stati sezionati al 7° giorno post-inoculazione (schema dello studio inFIGURA 2 UNA). I linfonodi inguinali dei topi inoculati con STARR™ SARS-CoV-2 RNA hanno mostrato un aumento di peso dose-dipendente, a differenza di quelli dei topi inoculati con vaccino a mRNA o PBS; il peso medio dei linfonodi dei topi a cui erano stati somministrati 10 µg di STARR™ SARS-CoV-2 RNA era significativamente più alto di quelli a cui era stato somministrato il vaccino a mRNA equivalente (FIGURA 2 C). L'analisi dei componenti principali (PCA) dell'espressione genica immunitaria ha mostrato il raggruppamento delle risposte a ciascuna delle 3 dosi di RNA STARR™ SARS-CoV-2 lontano dal controllo PBS (RNA STARR™): rappresentato come sfera inferiore inFIGURA 2, la sfera più piccola inFIGURA 2 E, e sfera inferiore inFIGURA 2 F.; Controllo PBS: rappresentato come sfera superiore inFIGURA 2, sfera inferiore allungata e stretta inFIGURA 2 E, e sfera superiore inFIGURA 2 F.) indicando chiare differenze nell'espressione genica immunitaria tra i gruppi vaccinati con RNA STARR™ SARS-CoV-2 e placebo. Queste tendenze erano anche diverse da quelle

dei topi a cui era stato somministrato il vaccino mRNA, dove a tutte le dosi testate, il PCA mostrava una sovrapposizione sostanziale con il placebo (mRNA: mostrato come sfera centrale inFIGURA 2, grande sfera verticale inFIGURA 2 E e come una linea piatta con quattro punti dati lungo la parte inferiore del quadrato centrale inFIGURA 2 F.; placebo (controllo PBS): mostrato come sfera superiore inFIGURA 2, sfera inferiore allungata e stretta inFIGURA 2 E, e sfera superiore inFIGURA 2 F.).

Successivamente sono stati valutati i geni differenzialmente espressi nei linfonodi dei topi a cui è stato somministrato STARR™ SARS-CoV-2 RNA rispetto a quelli inoculati con vaccino mRNA. L'analisi del grafico a vulcano ha identificato una significativa sovraregolazione di diversi geni innati, delle cellule B e delle cellule T negli animali immunizzati con STARR™ SARS-CoV-2 RNA (FIGURA 2 G- 2 I). Alcuni dei geni espressi in modo più altamente differenziale includevano, ad esempio, GZMB (importante per l'uccisione delle cellule bersaglio da parte delle cellule immunitarie citotossiche), S100A8 e S100A9 (fattori che regolano le risposte immunitarie tramite TLR4), TNFRSF17 (noto anche come BCMA e regola l'immunità umorale), CXCR3 (recettore della chemioquina coinvolto nel traffico e nella funzione delle cellule T) e AICDA (media il cambio di classe di anticorpi e l'ipermutazione somatica nelle cellule B).

Questi risultati indicano collettivamente che le risposte immunitarie adattive nei linfonodi drenanti dei topi inoculati con RNA STARR™ SARS-CoV-2 sembravano essere significativamente diverse rispetto alle risposte immunitarie nei topi inoculati con un vaccino mRNA non replicante.

Esempio 3

Questo esempio descrive le risposte delle cellule T indotte dall'RNA del SARS-CoV-2 STARR™.

Successivamente è stata studiata la risposta immunitaria cellulare in seguito alla vaccinazione di topi C57BL/6 (n=5 per gruppo) con mRNA o il costrutto RNA STARR™ SARS-CoV-2 che codifica la glicoproteina SARS-CoV-2 descritta sopra (Esempio 1). Al giorno 7 dopo la vaccinazione, le milze sono state raccolte e valutate per le cellule T CD8 e CD4 mediante citometria a flusso. Il sottoinsieme effetore/memoria delle cellule T CD8⁺ CD44⁺CD62L⁻ è stato significativamente espanso nei topi vaccinati con RNA STARR™ SARS-CoV-2 rispetto a quelli a cui è stato somministrato il vaccino PBS o mRNA (FIGURA 3 A-B). Non vi era alcuna differenza statisticamente significativa nella proporzione di cellule T effettrici CD4⁺ di questi animali (FIGURA 3 C). Le cellule T IFN γ ⁺CD8⁺ (con dosi di 2 μ g e 10 μ g) e le cellule T IFN γ ⁺CD4⁺ (in dosi di 0,2 μ g e 10 μ g) erano proporzionalmente più elevate, come riscontrato utilizzando la colorazione intracellulare (ICS) con citometria a flusso, nell'RNA STARR™ SARS-CoV-2 rispetto agli animali vaccinati con mRNA (FIGURA 3 D- 3 F).

Le risposte cellulari specifiche per SARS-CoV-2 sono state valutate negli animali vaccinati tramite ELISPOT. Un set di peptidi 15-mer che coprono la proteina S di SARS-CoV-2 è stato suddiviso in 4 pool e testato per le risposte IFN γ ⁺ negli splenociti di animali vaccinati e non vaccinati. Le risposte cellulari specifiche per SARS-CoV-2 (visualizzate come cellule IFN γ ⁺ SFU/106) sono state rilevate tramite ELISPOT sia negli animali immunizzati con RNA STARR™ SARS-CoV-2 che con mRNA rispetto al controllo PBS (FIGURA 3 G- 3 I). Queste risposte sono state più elevate in tutte le dosi di

STARR™ SARS-CoV-2 RNA rispetto ai gruppi vaccinati con mRNA (FIGURA 3 G- 3 I). Anche la dose più alta testata (10 µg) di vaccino mRNA ha prodotto risposte ELISPOT IFN γ + notevolmente inferiori a quelle ottenute con la dose più bassa (0,2 µg) di RNA STARR™ SARS-CoV-2.

Questi risultati dimostrano che il costrutto RNA STARR™ SARS-CoV-2 ha indotto forti risposte delle cellule T.

Esempio 4

Questo esempio illustra le risposte umorali in seguito alla vaccinazione con RNA STARR™ SARS-CoV-2.

Le risposte umorali specifiche di SARS-CoV-2 in seguito alla vaccinazione sono state caratterizzate in due diversi modelli di topi, BALB/c e C57BL/6. I topi femmina (n=5 per gruppo) sono stati vaccinati al giorno 0 e sottoposti a sanguinamento ogni 10 giorni, fino al giorno 60 per BALB/c e al giorno 30 per C57BL/6 (FIGURA 4 UN). Le risposte IgM S-specifiche per SARS-CoV-2 sono state testate a una diluizione sierica di 1:2000 utilizzando un immunodosaggio Luminex interno. Tutte le dosi testate di vaccino mRNA e RNA STARR™ SARS-CoV-2 (corrispondenti a SEQ ID NO: 125, come descritto nell'esempio 1 sopra) hanno prodotto risposte IgM S-specifiche rilevabili in entrambi i modelli di topi (FIGURA 4B - 4). Confrontando mRNA con topi BALB/c vaccinati con STARR™ SARS-CoV-2 RNA, non è stata osservata alcuna differenza nelle risposte IgM; i livelli di IgM nei topi C57BL/6 erano più alti nei topi C57BL/6 vaccinati con STARR™ SARS-CoV-2 RNA al giorno 10 dopo la vaccinazione. Al contrario, i livelli di IgG S-specifiche per SARS-CoV-2 (a diluizione sierica 1:2000) erano più alti dal giorno 20 in poi negli animali inoculati con STARR™ SARS-CoV-2 RNA rispetto al vaccino mRNA (FIGURA 4 D- 4 E). Sorprendentemente, i livelli di IgG hanno continuato a mostrare un trend crescente nei topi vaccinati con STARR™ SARS-CoV-2 RNA, sia BALB/c che C57BL/6, fino al giorno 50 post-vaccinazione con una singola inoculazione in tutte le dosi. Questo trend è in contrasto con i topi che hanno ricevuto il vaccino mRNA, dove nei topi BALB/c i livelli di anticorpi si sono stabilizzati dopo il giorno 10 post-vaccinazione; livelli crescenti di IgG S-specifici sono stati osservati nei topi C57BL/6 vaccinati con mRNA, ma erano inferiori a quelli osservati nei topi che hanno ricevuto STARR™ SARS-CoV-2 RNA.

Un'ulteriore caratterizzazione della risposta IgG specifica per SARS-CoV-2 negli animali vaccinati è stata condotta al giorno 30 post-immunizzazione per valutare quali regioni della proteina S sono mirate. I titoli endpoint IgG sono stati stimati per la proteina S dell'ectodominio completo, S1, S2 e regioni del dominio di legame del recettore (RBD). Per entrambi i candidati vaccini la maggior parte delle IgG specifiche per SARS-CoV-2 ha riconosciuto S1, sebbene siano stati rilevati anche titoli endpoint IgG elevati per la proteina S2 (FIGURA 4 F- 4 G). Tuttavia, i titoli endpoint IgG indotti dall'RNA STARR™ SARS-CoV-2 erano significativamente più alti rispetto a quelli prodotti dalla vaccinazione a mRNA (FIGURA 4 F- 4 G). In particolare, le IgG che legano il dominio di legame del recettore (RBD) della proteina S, che è un sito immunodominante di anticorpi neutralizzanti, erano anche più elevate nell'RNA STARR™ SARS-CoV-2 rispetto agli animali vaccinati con mRNA. Inoltre, a dosi inferiori, il vaccino mRNA ma non l'RNA STARR™ SARS-CoV-2 ha fatto fatica a suscitare titoli elevati di IgG specifici per SARS-CoV-2 nel ceppo di topi C57BL/6 più dominante Th1 (FIGURA 4 G.). Nel complesso, una singola dose di RNA STARR™ SARS-CoV-2 ha indotto differenze significative

nell'espressione genica immunitaria e risposte immunitarie cellulari superiori nei linfonodi drenanti rispetto al vaccino mRNA e, di conseguenza, risposte immunitarie umorali.

Questi dati dimostrano che la vaccinazione STARR™ SARS-CoV-2 RNA ha indotto risposte umorali più elevate rispetto alla vaccinazione mRNA.

Esempio 5

Questo esempio illustra il rischio ridotto di potenziamento immunitario con la vaccinazione STARR™ SARS-CoV-2 RNA.

Una considerazione sulla sicurezza del vaccino contro il coronavirus è il rischio di potenziamento immunitario delle malattie respiratorie mediato dal vaccino. In effetti, lo sviluppo del vaccino contro SARS-CoV e MERS-CoV ha evidenziato l'importanza delle risposte Th1 distorte per evitare il potenziamento immunitario indotto dal vaccino. Pertanto, è stato studiato l'equilibrio Th1/Th2 suscitato sia dalla vaccinazione con mRNA che con RNA STARR™ SARS-CoV-2 (costruito di RNA autoreplicante come descritto nell'esempio 1 sopra). Il destino della sottoclasse IgG delle plasmacellule è influenzato dalle cellule T helper (Th). Al giorno 30 post-vaccinazione, sia mRNA che RNA STARR™ SARS-CoV-2, ad eccezione della dose da 0,2 µg nei topi C56BL/6J, hanno indotto quantità comparabili di IgG1 S-specifica per SARS-CoV-2, una sottoclasse di IgG associata a Th2 nei topi (FIGURA 5A - 5B). Al contrario, le sottoclassi di IgG associate a Th1, IgG2a in BALB/c e IgG2c in C56BL/6J, erano significativamente maggiori negli animali vaccinati con RNA STARR™ SARS-CoV-2. I rapporti di IgG2a/IgG1 specifici per la proteina S (BALBc) e IgG2c/IgG1 (C57BL/6) erano maggiori di 1 negli animali vaccinati con RNA STARR™ SARS-CoV-2 (FIGURA 5A - 5B). Ad eccezione della dose da 0,2 µg, questi rapporti erano tutti significativamente maggiori con STARR™ SARS-CoV-2 RNA rispetto agli animali vaccinati con mRNA.

L'ICS è stato utilizzato per studiare la produzione di IFN γ (citochina Th1) e IL4 (citochina Th2) da parte delle cellule T CD4⁺ nelle milze dei topi C56BL/6J vaccinati al giorno 7. Come mostrato sopra (Esempio 3), rispetto alla vaccinazione con mRNA, i livelli di IFN γ erano significativamente maggiori negli animali vaccinati con STARR™ SARS-CoV-2 RNA (FIGURA 3 F). L'espressione di IL4 nelle cellule T CD4 era leggermente più alta nell'mRNA rispetto all'RNA STARR™ SARS-CoV-2 a dosi di 0,2 µg e 2 µg (FIGURA 5 C). Confrontando i livelli di IFN γ e IL4 nei singoli topi, i rapporti di IFN γ /IL4 nelle cellule T CD4⁺ sia per i topi vaccinati con RNA che con mRNA STARR™ SARS-CoV-2 erano superiori a 1 (FIGURA 5D). Il rapporto IFN γ /IL4 nelle cellule T CD4⁺ nei topi a cui sono state somministrate dosi di 0,2 µg e 2 µg è stato significativamente maggiore con la vaccinazione STARR™ SARS-CoV-2 RNA rispetto alla vaccinazione mRNA (FIGURA 5 F). Tuttavia, senza essere limitati dalla teoria, i rapporti elevati in queste dosi sembravano essere dovuti alla ridotta espressione di IL4 a livelli inferiori allo sfondo (ad esempio, topi di controllo PBS), piuttosto che alla ridotta attività di IFN γ e quindi Th1.

Presi nel loro insieme, questi dati mostrano che l'RNA STARR™ SARS-CoV-2 ha prodotto risposte immunitarie adattative distorte di tipo Th1 anziché Th2.

Esempio 6

Questo esempio illustra la qualità delle risposte immunitarie umorali indotte dall'RNA del virus SARS-CoV-2 STARR™.

Successivamente sono state valutate la forza di legame (avidità) e la capacità neutralizzante della risposta anticorpale suscitata dall'RNA autoreplicante STARR™ SARS-CoV-2 (costruito come descritto nell'Esempio 1 sopra) e dai costrutti del vaccino mRNA. L'avidità delle IgG sieriche è stata misurata al giorno 30 post-vaccinazione utilizzando un immunodosaggio Luminex modificato con lavaggi di urea 8M. L'RNA STARR™ SARS-CoV-2 ha suscitato un'avidità maggiore delle IgG specifiche per la proteina S rispetto all'mRNA in entrambi i modelli di topi a tutte le dosi testate (FIGURA 6 UN). Queste differenze sono state osservate, ad eccezione di 0,2 µg in BALB/c, in tutte le dosi (FIGURA 6 UN), indicando che l'RNA STARR™ SARS-CoV-2 ha suscitato anticorpi di qualità migliore rispetto all'mRNA convenzionale.

La neutralizzazione del SARS-CoV-2 vivo da parte del siero di animali vaccinati è stata valutata utilizzando il test di neutralizzazione della riduzione della placca (PRNT). Al giorno 30, BALB/c vaccinato con STARR™ SARS-CoV-2 RNA ha mostrato un chiaro aumento dose-dipendente nei titoli PRNT50; 4 su 5 (80%) dei topi nel gruppo STARR™ SARS-CoV-2 RNA da 10 µg hanno mostrato titoli PRNT50 superiori al limite superiore di 320 (FIGURA 6). Simili tendenze dose-dipendenti nei titoli PRNT50 sono state riscontrate anche nei topi C57BL/6, sebbene in questi animali i titoli PRNT50 di diversi animali abbiano superato il limite superiore di 320 anche con una bassa dose di vaccinazione di 0,2 µg (FIGURA 6). Al contrario, i titoli PRNT50 negli animali inoculati con il vaccino mRNA erano, ad eccezione di un topo C57BL/6J che aveva ricevuto una dose di 10 µg, tutti <20 (FIGURA 6). Inaspettatamente e sorprendentemente, i titoli PRNT50 e PRNT70 dei topi BALBc vaccinati con RNA STARR™ SARS-CoV-2 hanno continuato ad aumentare tra il giorno 30 e il giorno 60 dopo una singola dose di vaccinazione (FIGURA 6 C- 6 D). Questi titoli erano anche comparabili ai titoli PRNT70 nei sieri di pazienti convalescenti affetti da COVID-19 (FIGURA 6 D).

I titoli di IgG della proteina S sono risultati positivamente correlati anche con i titoli di PRNT50 in entrambi i modelli di topi (FIGURA 6 E). Correlazioni positive simili sono state osservate anche con IgG contro S1 e RBD (FIGURA 9). Al contrario, non è stata trovata alcuna correlazione tra i titoli di IgG e PRNT50 nei topi vaccinati con mRNA (FIGURA 6 E). Nel complesso, senza essere limitate dalla teoria, queste analisi della risposta anticorpale indicano che i titoli PRNT50 più elevati a seguito della vaccinazione STARR™ SARS-CoV-2 RNA non sono solo attribuibili alla quantità di IgG prodotta, ma anche alla superiorità della qualità degli anticorpi anti-SARS-CoV-2.

In sintesi, l'RNA STARR™ SARS-CoV-2 ha indotto una risposta immunitaria umorale qualitativamente superiore rispetto all'mRNA convenzionale.

Esempio 7

Questo esempio illustra l'effetto di una seconda dose di RNA STARR™ SARS-CoV-2.

Successivamente è stato esplorato un possibile beneficio aggiuntivo di una seconda dose di RNA STARR™ SARS-CoV-2 (costruito di RNA autoreplicante come descritto nell'Esempio 1 sopra) per le risposte immunitarie cellulari e umorali alla proteina S di SARS-CoV-2. I punteggi clinici dopo la

seconda dose erano più alti rispetto alla prima dose (FIGURA 7 UN). Come la prima dose, i topi che hanno ricevuto 2 µg e 10 µg di RNA STARR™ SARS-CoV-2 hanno sperimentato una perdita di peso (FIGURA 7B). La risposta IgG a una seconda dose di RNA STARR™ SARS-CoV-2 ha prodotto un apprezzabile aumento dei livelli di IgG specifiche per la proteina S, ma solo con 0,2 µg e 2 µg di RNA STARR™ SARS-CoV-2 (FIGURA 7 C). Senza essere limitati dalla teoria, una probabile ragione per la mancanza di aumento nei livelli di IgG specifiche per la proteina anti-S per la dose da 10 µg è che la quantità di fluorescenza è vicina al punto di saturazione del rilevatore e i sieri non sono stati ulteriormente diluiti per osservare e aumentare. Tuttavia, in un successivo studio sui topi Balb/c, i sieri dei topi vaccinati con una dose di RNA da 5 µg somministrata unilateralmente in un volume di iniezione di 0,05 mL hanno prodotto un aumento significativo nei titoli di anticorpi neutralizzanti come saggiato utilizzando un formato di saggio di microneutralizzazione a 96 pozzetti. I topi sono stati dissanguati ogni 14 giorni e una seconda vaccinazione da 5 µg è stata somministrata il giorno 28. 4 topi sono stati iniettati con un RNA replicone VEEV che esprimeva luciferasi come controllo negativo e 6 topi sono stati vaccinati con RNA STARR™ SARS-CoV-2. I risultati sono mostrati nella Tabella 8 di seguito.

TABELLA 7 Titoli di microneutralizzazione (MN50) nei topi Balb/c

Topo	Titolo di microneutralizzazione (MN50) NO	Trattamento	Settimana 0	Settimana 2	Settimana 4	Settimana 6	Settimana 8
1	Luciferasi	<10	<10	<10	<10	<10	<10
2	Luciferasi	<10	<10	<10	<10	<10	<10
3	Luciferasi	<10	<10	<10	<10	<10	<10
4	Luciferasi	<10	<10	<10	<10	<10	<10
5	STARR™ SARS-	<10	1.280	5.120	327.680	81.920	RNA del CoV-2
6	STARR™ SARS-	<10	640	20.480	327.680	327.680	RNA del CoV-2
7	STARR™ SARS-	<10	1.280	2.560	163.840	163.840	RNA del CoV-2
8	STARR™ SARS-	<10	1.280	10.240	327.680	163.840	RNA del CoV-2
9	STARR™ SARS-	<10	640	40.960	327.680	327.680	RNA del CoV-2
10	STARR™ SARS-	<10	1.280	10.240	327.680	327.680	RNA del CoV-2
	Media	1.016	10.240	29,1930	206.426	Geometrico	Significare

I titoli di neutralizzazione sono aumentati di 10 volte tra il giorno 14 e il giorno 28 dopo la vaccinazione. Dopo il boost del giorno 28, i titoli di neutralizzazione sono aumentati di altre 20 volte 14 giorni dopo il boost.

Per determinare se vi fosse un beneficio aggiunto nei conteggi delle cellule T IFNγ+CD8+ da una seconda dose di vaccinazione, le risposte delle cellule T CD8 IFNγ nei topi a cui era stato somministrato solo un prime sono state confrontate con le risposte dei topi a cui era stato somministrato un prime e un boost. Il fold change nelle cellule T IFNγ+CD8+ nei topi di controllo vaccinati su PBS è stato calcolato per i topi a cui era stato somministrato solo un prime o un prime e un boost. Il fold change delle cellule T IFNγ+CD8+ è stato simile in seguito al prime e al prime+boost per dosi da 2 µg e 10 µg di RNA STARR™ SARS-CoV-2 (FIGURA 7 D- 7 E); la dose da 0,2 µg ha mostrato un cambiamento di piega più elevato delle cellule T IFNγ+CD8+ tra prime (al giorno 7) e prime+boost (giorno 50). La vaccinazione con 0,2 µg di mRNA ha anche mostrato un aumento delle cellule T IFNγ+CD8+ rispetto al controllo PBS dopo due dosi di vaccinazione. Senza essere limitati dalla teoria, questi risultati suggeriscono che una seconda dose da 10 µg di RNA STARR™ SARS-CoV-2 non ha prodotto un'immunità cellulare superiore rispetto alla vaccinazione a dose singola. Pertanto, non vi è stato alcun beneficio apparente da una seconda dose da 10 µg di RNA STARR™ SARS-CoV-2.

Nel complesso, questi dati suggeriscono che 10 gSTARR™ SARS-CoV-2 RNA offre l'opportunità di una vaccinazione a dose singola per proteggere dal COVID-19.

Esempio 8

Questo esempio illustra la protezione contro la sfida virale SARS-CoV-2 nei topi a seguito della vaccinazione con RNA autoreplicante STARR™ SARS-CoV-2.

È stato condotto uno studio di sfida virale su topi con topi transgenici ACE2 umani. I topi sono stati immunizzati con dosi di RNA da 2 µg e 10 µg di RNA STARR™ SARS-CoV-2 (costruito di RNA come descritto nell'esempio 1 sopra) o iniettati con PBS. C'erano tre diverse coorti con 5 topi in ciascun gruppo di trattamento. Le coorti 1 e 3 hanno ricevuto un carico di sfida virale SARS-CoV-2 letale di 5×10^5 TCID₅₀. La coorte 1 è stata monitorata per la sopravvivenza e la coorte 3 è stata soppressa 5 giorni dopo la sfida. I polmoni sono stati analizzati per la carica virale e processati per l'istopatologia. Le coorti 2 hanno ricevuto un carico virale subletale di 5×10^4 TCID₅₀. La coorte 2 è stata soppressa 5 giorni dopo la sfida virale e i polmoni sono stati analizzati per il virus infettivo e processati per l'istopatologia. Tutti i topi sono stati inoculati per via intratracheale 30 giorni dopo la vaccinazione con una singola dose di RNA STARR™ SARS-CoV-2.

Tutti i topi iniettati con PBS nella coorte 1 erano morti entro il 7° giorno, mentre tutti i topi vaccinati non mostravano segni di infezione 15 giorni dopo la sfida virale (FIGURA 10). Per la coorte 2 che ha ricevuto una carica virale subletale, sono state misurate da 10 a 3.300 copie di virus tramite RT-PCR nei polmoni con una media di 1.200 copie, mentre nessuna copia di RNA virale è stata misurata nei topi vaccinati con ARTC-021 a dosi di 2 µg e 10 µg di RNA (LOD era 0,1 copie; FIGURA 11, a sinistra). Sono state osservate anche copie di RNA virale nel cervello che andavano da 20 a 80 nel gruppo di trattamento con PBS, mentre nessuna copia di RNA virale è stata misurata nel cervello dei topi vaccinati con dosi di RNA da 2,0 µg o 10,0 µg (FIGURA 11, a destra). Infine, i polmoni sono stati attentamente elaborati e analizzati per i titoli di placca polmonare. I titoli medi di placca per il gruppo iniettato con PBS erano 8×10^3 /mL di omogenato polmonare, mentre non sono state rilevate placche per i topi vaccinati con 2,0 µg o 10,0 µg o STARR™ SARS-CoV-2 RNA (FIGURA 12). I tessuti polmonari e cerebrali della coorte 3 sono in fase di analisi per il numero di copie virali e il virus infettivo. L'istopatologia dei polmoni per le coorti 2 e 3 è in corso.

Questi risultati dimostrano che la vaccinazione con RNA autoreplicante STARR™ SARS-CoV-2 ha protetto i topi da un'infezione letale da SARS-CoV-2 e li ha protetti dall'infezione polmonare e cerebrale in seguito alla somministrazione di una dose subletale di SARS-CoV-2.

Esempio 9

La pandemia di COVID-19 è causata dall'infezione con il virus SARS-CoV-2. Una delle principali mutazioni rilevate finora nella proteina spike dell'involucro virale SARS-CoV-2, che è responsabile dell'attaccamento del virus all'ospite ed è anche il bersaglio principale per gli anticorpi dell'ospite, è una mutazione di un aspartato (D) in posizione 614 trovato frequentemente nei ceppi cinesi in una glicina (G). I trascritti del replicone VEEV che esprimono le versioni D614 e G614 della glicoproteina spike SARS-CoV-2 sono stati formulati con la stessa identica formulazione lipidica

degli studi descritti negli esempi 1-8. I topi Balb/c sono stati vaccinati con una singola somministrazione di RNA di 0,2 µg, 2,0 µg e 10,0 µg di RNA. C'erano 5 topi per dose. I topi sono stati dissanguati nei giorni 14, 28 e 42 dopo la vaccinazione. I sieri sono stati diluiti 1/2000 e incubati con sfere Luminex derivate con la glicoproteina spike SARS-CoV-2 contenente la sequenza amminoacidica D614. Un anticorpo secondario di topo derivato con un fluoroforo è stato utilizzato per analizzare l'anticorpo legato alle sfere e l'intensità di fluorescenza media aggiustata (MFI) è stata misurata in funzione della dose di RNA, mostrata in FIGURA 13I risultati hanno mostrato che l'MFI è aumentato in funzione della dose di RNA con un MFI leggermente più alto osservato per il siero di topi immunizzati con la glicoproteina spike G614. Questo leggero aumento è attribuito a una percentuale inferiore di RNA a lunghezza intera con la sequenza di amminoacidi D614. Una conclusione importante è che il siero di topi immunizzati con il costrutto di RNA della glicoproteina spike G614 è stato in grado di legarsi alla glicoproteina spike con la sequenza di amminoacidi D614, indicativo di reattività crociata.

Questi risultati dimostrano che l'immunizzazione con una glicoproteina spike G614 espressa da RNA autoreplicante determina la produzione di anticorpi in grado di legarsi a una glicoproteina spike D614.

Discussione degli esempi 1-9

La pandemia di COVID-19 ha reso necessario uno sviluppo rapido di vaccini, poiché il distanziamento fisico per prevenire la trasmissione di SARS-CoV-2 non è una soluzione sostenibile a lungo termine. Diversi candidati vaccini contro il COVID-19 sono ora in fase di sperimentazione clinica e altri stanno entrando in sperimentazioni first-in-human. Tuttavia, la maggior parte dei candidati vaccini in fase di sviluppo richiede due dosi per un'immunità adattativa sufficiente. Un vaccino a dose singola che generi sia immunità cellulare che umorale, senza aumentare il rischio di potenziamento immunitario mediato dal vaccino, rimane un'esigenza insoddisfatta. Senza essere limitato dalla teoria, l'implementazione di un vaccino a dose singola consentirebbe un maggiore livello di conformità e consentirebbe la distribuzione di una produzione finita di vaccini a persone più suscettibili a livello globale.

Tra i vaccini autorizzati, i vaccini vivi attenuati possono offrire una protezione duratura contro le malattie virali. I vaccini vivi infettano e si replicano nei siti di inoculazione e alcuni persino nei linfonodi drenanti. La replicazione consente l'espressione endogena di antigeni virali che consentono alla presentazione dell'antigene di stimolare le cellule T CD8+ citotossiche. Gli antigeni espressi verrebbero anche assorbiti dalle cellule presentanti l'antigene per innescare l'aiuto delle cellule T CD4+ che guida la maturazione dell'affinità nelle cellule B. Gli studi sul vaccino vivo attenuato contro la febbre gialla hanno dimostrato che un periodo più lungo di stimolazione della risposta immunitaria adattativa si traduce in risposte immunitarie adattative superiori. Senza essere limitati dalla teoria, la simulazione dei processi di vaccinazione viva potrebbe offrire una possibilità di immunità duratura contro il COVID-19.

In una crisi come quella del COVID-19, una piattaforma di vaccini con acido nucleico offre opportunità per uno sviluppo accelerato. Negli studi descritti nel presente documento, è stato condotto un confronto affiancato dell'immunogenicità suscitata da due candidati vaccini SARS-CoV-

2, un costrutto mRNA non replicativo e STARR™ SARS-CoV-2 RNA. Rispetto a un vaccino mRNA, STARR™ SARS-CoV-2 RNA ha prodotto un'espressione proteica più elevata e più lunga in vivo e ha sovraregolato l'espressione genica di diversi geni di risposta innata, delle cellule B e delle cellule T nel sangue e nei linfonodi drenanti. Queste proprietà si sono tradotte in risposte delle cellule T CD8+ significativamente maggiori, risposte ELISPOT IFN γ + e risposte distorte IgG e Th1 specifiche per SARS-CoV-2. È interessante notare che, nonostante la dose più alta testata di mRNA che suscita anticorpi specifici per la proteina S comparabili alla dose più bassa testata di RNA STARR™ SARS-CoV-2, le IgG suscitate da mRNA non hanno mostrato un'avidità o un'attività di neutralizzazione simili a quelle della vaccinazione con RNA STARR™ SARS-CoV-2. Questi risultati evidenziano quindi i vantaggi immunologici dell'RNA autoreplicante rispetto alle piattaforme mRNA. Inoltre, gli studi di sfida sui topi con il virus SARS-CoV-2 hanno mostrato che la vaccinazione con una singola dose alta (10 μ g) o una singola dose bassa (2 μ g) di RNA autoreplicante STARR™ SARS-CoV-2 ha protetto i topi da un'infezione letale da SARS-CoV-2 e da infezioni polmonari e cerebrali in seguito alla sfida con una dose subletale di SARS-CoV-2.

La misura in cui i vaccini STARR™ riproducono le caratteristiche dei vaccini vivi deve ancora essere definita sperimentalmente. Senza essere limitati dalla teoria, la qualità superiore delle risposte immunitarie suscitate dall'RNA STARR™ SARS-CoV-2 rispetto al vaccino mRNA potrebbe essere attribuibile a molteplici fattori, tutti associati alla vaccinazione viva. Ad esempio, un'espressione più elevata e prolungata di immunogeni produce una migliore immunità, probabilmente attraverso un migliore coinvolgimento delle cellule T follicolari helper e portando così a bersagli anticorpali più diversificati e a più risposte anticorpali neutralizzanti. La replicazione dell'RNA STARR™ SARS-CoV-2 si tradurrebbe nella formazione di un modello a filamento negativo per la produzione di più mRNA a filamento positivo e mRNA subgenomico che esprime il transgene S. L'interazione tra i filamenti negativi e positivi formerebbe RNA a doppio filamento (dsRNA), che interagirebbe con i recettori TLR3 e RIG-I-like per stimolare le risposte all'interferone, che hanno dimostrato di essere correlate a risposte immunitarie adattive superiori. La produzione di IFN γ può quindi stimolare lo sviluppo di cellule T CD8+ citotossiche. È importante notare che la proteina S contiene epitopi di cellule T CD8+ umane. Senza essere limitati dalla teoria, lo sviluppo della memoria delle cellule T potrebbe essere importante per l'immunità a lungo termine, come suggerito da recenti scoperte sulle risposte delle cellule T al SARS-CoV-2 e ad altre infezioni da coronavirus.

Non è chiaro se il VEEV nsP1-4 che forma il complesso di replicazione contenga proprietà immunogeniche, sebbene sia stato dimostrato che le mutazioni nelle proteine nsP influenzano l'induzione dell'IFN di tipo I. È stato anche dimostrato che i repliconi VEEV adiuvano le risposte immunitarie nei siti della mucosa, illustrando ulteriormente i vantaggi dell'utilizzo della piattaforma STARR™ per sviluppare un vaccino COVID-19. Senza essere limitati dalla teoria, non sembra esserci una risposta immunitaria alle proteine non strutturali del replicone, come indicato da un aumento della produzione di IgG specifiche per l'antigene dopo una seconda somministrazione di replicone agli animali. In presenza di una risposta immunitaria alle proteine non strutturali, potrebbe essersi verificato un aumento limitato o nullo nella produzione di IgG specifiche per l'antigene dopo una seconda somministrazione. L'RNA è incapsulato in nanoparticelle lipidiche (LNP), che insieme possono formare potenti adiuvanti che portano a

robuste risposte immunitarie. Inoltre, utilizzando la sequenza genetica di un antigene, tra cui un antigene virale come la proteina spike del SARS-CoV-2, ad esempio, i vaccini STARR™ possono essere rapidamente generati e prodotti utilizzando tecniche acellulari e rapidamente scalabili.

In conclusione, un vaccino STARR™ come quello esemplificato da STARR™ SARS-CoV-2 RNA offre un approccio per simulare molte delle proprietà del vaccino vivo e offre la possibilità di una vaccinazione a dose singola contro il COVID-19.

SEQUENZE SEQ ID NO: 72

```
ATGGAGAAAGTTCACGTTGACATCGAGGAAGACAGCCATTCCTCAGAGCTTTG
CAGCGGAGCTTCCCGCAGTTTGAGGTAGAAGCCAAGCAGGTCACTGATAATGAC
CATGCTAATGCCAGAGCGTTTTTCGCATCTGGCTTCAAACACTGATCGAAACGGAGG
TGGACCCATCCGACACGATCCTTGACATTGGAAGTGCGCCCGCCCGCAGAATGT
ATTCTAAGCACAAAGTATCATTGTATCTGTCCGATGAGATGTGCGGAAGATCCGGA
CAGATTGTATAAGTATGCAACTAAGCTGAAGAAAACTGTAAGGAAATAACTGA
TAAGGAATTGGACAAGAAAATGAAGGAGCTGGCCGCCGTCATGAGCGACCCTGA
CCTGGAAACTGAGACTATGTGCCTCCACGACGACGAGTCGTGTCGCTACGAAGG
GCAAGTCGCTGTTTACCAGGATGTATACGCCGTCGACGGCCCCACCAGCCTGTAC
CACCAGGGCCAACAAGGGCGTGAGGGTGGCCTACTGGATCGGCTTCGACACCACA
CCCTTCATGTTCAAGAACCTGGCCGGCGCCTACCCAGCTACAGCACCAACTGGG
CCGACGAGACCGTGCTGACCGCCAGGAACATCGGCCTGTGCAGCAGCGACGTGA
TGGAGAGGGAGCCGGAGAGGCATGAGCATCCTGAGGAAGAAATACCTGAAGCCC
AGCAACAACGTGCTGTTTACGCGTGGGCAGCACCATCTACCACGAGAAGAGGGAC
CTGCTCAGGAGCTGGCACCTGCCACGTGTTCCACCTGAGGGGGCAAGCAGAAC
TACACCTGCAGGTGCGAGACCATCGTGAGCTGCGACGGCTACGTGGTGAAGAGG
ATCGCCATCAGCCCCGGCCTGTACGGCAAGCCCAGCGGCTACGCCGCTACAATG
CACAGGGAGGGCTTCTGTGCTGCAAGGTGACCGACACCCTGAACGGCGAGAGG
GTGAGCTTCCCCGTGTGCACCTACGTGCCCGCCACCCTGTGCGACCAGATGACCG
GCATCCTGGCCACCGACGTGAGCGCCGACGACGCCAGAAGCTGCTCGTGGGCC
TGAACCAGAGGATCGTGGTCAACGGCAGGACCCAGAGGAACACCAACACAATG
AAGAACTACCTGCTGCCCGTGGTGGCCAGGCTTTCGCCAGGTGGGCCAAGGAG
TACAAGGAGGACCAGGAAGACGAGAGGCCCTGGGCCTGAGGGACAGGCAGCT
GGTGATGGGCTGCTGCTGGGCCTTACAGGCGGCAAGATCACCAGCATCTACAA
GAGGCCCCGACACCCAGACCATCATCAAGGTGAACAGCGACTTCCACAGCTTCGT
GCTGCCAGGATCGGCAGCAACACCCTGGAGATCGGCCTGAGGACCCGGATCAG
GAAGATGCTGGAGGAACACAAGGAGCCCAGCCCCTGATCACCGCCGAGGACGT
GCAGGAGGCCAAGTGCGCTGCCGACGAGGCCAAGGAGGTGAGGGAGGCCGAGG
AACTGAGGGCCGCCCTGCCACCCCTGGCTGCCGACGTGGAGGAACCCACCCTGG
AAGCCGACGTGGACCTGATGCTGCAGGAGGCCGGCGCCGGAAGCGTGGAGACA
CCCAGGGCCTGATCAAGGTGACCAGCTACGACGGCGAGGACAAGATCGGCAGC
TACGCCGTGCTGAGCCCACAGGCCGTGCTGAAGTCCGAGAAGCTGAGCTGCATC
CACCCACTGGCCGAGCAGGTGATCGTGATCACCCAGCGGCAGGAAGGGCAGG
TACGCCGTGGAGCCCTACCACGGCAAGGTGGTCGTGCCCCGAGGGCCACGCCATC
CCCGTGCAGGACTTCCAGGCCCTGAGCGAGAGCGCCACCATCGTGTACAACGAG
```

AGGGAGTTCGTGAACAGGTACCTGCACCATATCGCCACCCACGGCGGAGCCCTG
AACACCGACGAGGAATACTACAAGACCGTGAAGCCCAGCGAGCACGACGGCGA
GTACCTGTACGACATCGACAGGAAGCAGTGCCTGAAGAAAGAGCTGGTGACCGG
CCTGGGACTGACCGCGAGCTGGTGGACCCACCCTTCCACGAGTTCGCCTACGA
GAGCCTGAGGACCAGACCCGCCGCTCCCTACCAGGTGCCACCATCGGCGTGTA
CGGCGTGCCCGCAGCGGAAAGAGCGGCATCATCAAGAGCGCCGTGACCAAGA
AAGACCTGGTGGTCAGCGCCAAGAAAGAGAACTGCGCCGAGATCATCAGGGAC
GTGAAGAAGATGAAAGGCCTGGACGTGAACGCGCGCACCGTGGACAGCGTGCTG
CTGAACGGCTGCAAGCACCCGTGGAGACCCTGTACATCGACGAGGCCTTCGCTT
GCCACGCCGGCACCCCTGAGGGCCCTGATCGCCATCATCAGGCCAAGAAAGCCG
TGCTGTGCGGGACCCCAAGCAGTGCGGCTTCTTCAACATGATGTGCCTGAAGGT
GCATCTAACCACGAGATCTGCACCCAGGTGTTCCACAAGAGCATCAGCAGGCG
GTGCACCAAAGAGCGTGACCAGCGTCGTGAGCACCCCTGTTCTACGACAAGAAAAT
GAGGACCACCAACCCCAAGGAGACCAAAATCGTGATCGACACCACAGGCAGCA
CCAAGCCCAAGCAGGACGACCTGATCCTGACCTGCTTACGGGGCTGGGTGAAGC
AGCTGCAGATCGACTACAAGGGCAACGAGATCATGACCGCCGCTGCCAGCCAGG
GCCTGACCAGGAAGGGCGTGTACGCCGTGAGGTACAAGGTGAACGAGAACCAC
TGTACGTCCCACCAGCGAGCACGTGAACGTGCTGCTGACCAGGACCGAGGACA
GGATCGTGTGGAAGACCCCTGGCCGGCGACCCCTGATCAAGACCCTGACCGCCA
AGTACCCCGGCAACTTACCCGCCACCATCGAAGAGTGGCAGGCCGAGCACGACG
CCATCATGAGGCACATCCTGGAGAGGCCCGACCCACCGACGTGTTCCAGAACA
AGGCCAACGTGTGCTGGGCCAAGGCCCTGGTGCCCGTCTGAAGACCGCCGGCA
TCGACATGACCACAGAGCAGTGGAACACCGTGGACTACTTCGAGACCGACAAGG
CCCACAGCGCCGAGATCGTGCTGAACCAGCTGTGCGTGAGGTTCTTCGGCCTGGA
CCTGGACAGCGGCCTGTTACGCGCCCCCACCCTGCCACTGAGCATCAGGAACAA
CCACTGGGACAACAGCCCCAGCCAAACATGTACGGCCTGAACAAGGAGGTGGT
CAGGCAGCTGAGCAGGCGGTACCCACACGCTGCCAGGGCCGTGGCCACCGGCAG
GGTGTACGACATGAACACCGGCACCCTGAGGAACTACGACCCCAAGGATCAACCT
GGTGCCCGTGAACAGGCGGCTGCCCCACGCCCTGGTGTGCACCACAACGAGCA
CCCACAGAGCGACTTCAGCTCCTTCGTGAGCAAGCTGAAAGGCAGGACCGTGCT
GGTCGTGGGCGAGAAGCTGAGCGTGCCCGCAAGATGGTGGACTGGCTGAGCGA
CAGGCCCGAGGCCACCTTCCGGGCCAGGCTGGACCTCGGCATCCCCGGCGACGT
GCCAAGTACGACATCATCTTCGTGAACGTCAGGACCCATAACAAGTACCACCAT
TACCAGCAGTGCAGGACCACGCCATCAAGCTGAGCATGCTGACCAAGAAGGCC
TGCCTGCACCTGAACCCCGGAGGCACCTGCGTGAGCATCGGCTACGGCTACGCC
GACAGGGCCAGCGAGAGCATCATTGGCGCCATCGCCAGGCTGTTCAAGTTCAGC
AGGGTGTGCAAACCAAGAGCAGCCTGGAGGAAACCGAGGTGCTGTTTCGTGTTT
ATCGGCTACGACCGGAAGGCCAGGACCCAGGACCCACAACCCCTACAAGCTGAGCAGCACC
CTGACAAACATCTACACCGGCAGCAGGCTGCACGAGGCCGGCTGCGCCCCCAGC
TACCACGTGGTCAGGGGGATATCGCCACCGCCACCGAGGGCGTGATCATCAAC
GCTGCCAACAGCAAGGGCCAGCCCGGAGGCGGAGTGTGCGGCGCCCTGTACAAG
AAGTTCCCCGAGAGCTTCGACCTGCAGCCATCGAGGTGGGCAAGGCCAGGCTG
GTGAAGGGCGCCGCTAAGCACATCACACGCCGTGGGCCCAACTTCAACAAG
GTGAGCGAGGTGGAAGGCGACAAGCAGCTGGCCGAAGCCTACGAGAGCATCGC

CAAGATCGTGACGACAATAACTACAAGAGCGTGGCCATCCCACTGCTCAGCAGCAC
CGGCATCTTCAGCGGCAACAAGGACAGGCTGACCCAGAGCCTGAACCACCTGCT
CACCGCCCTGGACACCACCGATGCCGACGTGCCATCTACTGCAGGGACAAGAA
GTGGGAGATGACCCTGAAGGAGGCCGTGGCCAGGCGGGAGGCCGTGGAAGAGA
TCTGCATCAGCGACGACTCCAGCGTGACCGAGCCCGACGCCGAGCTGGTGAGGG
TGCACCCCAAGAGCTCCCTGGCCGGCAGGAAGGGCTACAGCACCAGCGACGGCA
AGACCTTCAGCTACCTGGAGGGCACCAAGTTCACCAGGCCGCTAAGGACATCG
CCGAGATCAACGCTATGTGGCCCCGTGGCCACCGAGGCCAACGAGCAGGTGTGCA
TGTACATCCTGGGCGAGAGCATGTCCAGCATCAGGAGCAAGTGCCCCGTGGAGG
AAAGCGAGGCCAGCACACCACCCAGCACCCCTGCCCTGCCTGTGCATCCACGCTA
TGACACCCGAGAGGGTGCAGCGGCTGAAGGCCAGCAGGCCCGAGCAGATCACC
GTGTGCAGCTCCTTCCCACTGCCCAAGTACAGGATCACCGGCGTGCAGAAGATCC
AGTGCAGCCAGCCCATCCTGTTAGCCCAAAGGTGCCCGCTACATCCACCCAG
GAAGTACCTGGTGGAGACCCACCCGTGGACGAGACACCCGAGCCAAGCGCCGA
GAACCAGAGCACCGAGGGCACACCCGAGCAGCCACCCCTGATCACCGAGGACG
AGACAAGGACCCGGACCCAGAGCCATCATTATCGAGGAAGAGGAAGAGGAC
AGCATCAGCCTGCTGAGCGACGGCCCCACCCACCAGGTGCTGCAGGTGGAGGCC
GACATCCACGGCCCACCCAGCGTGTCCAGCTCCAGCTGGAGCATCCACAGCC
AGCGACTTCGACGTGGACAGCCTTGAGCATCCTGGACACCCTGGAGGGCGCCAGC
GTGACCTCCGGCGCCACCAGCGCCGACCAACAGCTACTTCGCCAAGAGCATG
GAGTTCCTGGCCAGGCCCGTGCCAGCTCCCAGGACCGTGTTTCAGGAACCCACCCC
ACCCAGCTCCCAGGACCAGGACCCCAAGCCTGGCTCCCAGCAGGGCCTGCAGCA
GGACCAGCCTGGTGAGCACCCACCCGGCGTGAACAGGGTGATCACAGGG
AACTGGAGGCCCTGACACCCAGCAGGACCCCCAGCAGGTCCGTGAGCAGGACTA
GTCTGGTGTCCAACCCACCCGGCGTGAACAGGGTGATCACAGGGAGGAATTCG
AGGCCTTCGTGGCCCAGCAACAGAGACGGTTCGACGCCGGCGCCTACATCTTCA
GCAGCGACACCCGGCCAGGGACACCTGCAGCAAAAGAGCGTGAGGCAGACCGTG
CTGAGCGAGGTGGTGTGGAGAGGACCGAGCTGGAAATCAGCTACGCCCCCAGG
CTGGACCAGGAAGGAAGGAACTGCTCAGGAAGAACTGCAGCTGAACCCAC
CCCAGCCAACAGGAGCAGGTACCAGAGCAGGAAGGTGGAGAACATGAAGGCCA
TCACCGCCAGGCGGATCCTGCAGGGCCTGGGACACTACCTGAAGGCCGAGGGCA
AGGTGGAGTGCTACAGGACCCTGCACCCCGTGCCACTGTACAGCTCCAGCGTGA
ACAGGGCCTTCTCCAGCCCCAAGGTGGCCGTGGAGGCCTGCAACGCTATGCTGA
AGGAGAACTTCCCACCGTGGCCAGCTACTGCATCATCCCCGAGTACGACGCCTA
CCTGGACATGGTGGACGGCGCCAGCTGCTGCCTGGACACCGCCAGTTTGTCCCC
GCCAAGCTGAGGAGCTTCCCCAAGAAACACAGCTACCTGGAGCCCACCATCAGG
AGCGCCGTGCCAGCGCCATCCAGAACACCCTGCAGAACGTGCTGGCCGCTGCC
ACCAAGAGGAACTGCAACGTGACCCAGATGAGGGAGCTGCCCGTGGTGGACAGC
GCTGCCTTCAACGTGGAGTGCTTCAAGAAATACGCCTGCAACAACGAGTACTGG
GAGACCTTCAAGGAGAACCCCATCAGGCTGACCGAAGAGAACGTGGTGAACACTAC
ATCACCAGCTGAAGGGCCCCAAGGCCGTGCCCTGTTTCGCTAAGACCCACAAC
CTGAACATGCCTGCAGGACATCCCAATGGACAGGTTTCGTGATGGACCTGAAGAGG
GACGTGAAGGTGAACACCCGGCACCAAGCACACCGAGGAGAGGCCCAAGGTGCA
GGTGATCCAGGCCGCTGACCCACTGGCCACCGCCTACCTGTGCGGCATCCACAG

GGAGCTGGTGAGGCGGCTGAACGCCGTGCTGCTGCCAACATCCACACCCTGTT
GACATGAGCGCCGAGGACTTCGACGCCATCATCGCCGAGCACTTCCAGCCCGGC
GACTGCGTGCTGGAGACCGACATCGCCAGCTTCGACAAGAGCGAGGATGACGCT
ATGGCCCTGACCGCTCTGATGATCCTGGAGGACCTGGGCGTGGACGCCGAGCTG
CTCACCTGATCGAGGCTGCCTTCGGCGAGATCAGCTCCATCCACCTGCCACCA
AGACCAAGTTCAAGTTCGGCGCTATGATGAAAAGCGGAATGTTCTGACCCTGTT
CGTGAACACCGTGATCAACATTGTGATCGCCAGCAGGCTGCTGCGGGAGAGGCT
GACCGGCAGCCCCTGCGCTGCCTTCATCGGCGACGACAACATCGTGAAGGGCGT
GAAAAGCGACAAGCTGATGGCCGACAGGTGCGCCACCTGGCTGAACATGGAGGT
GAAGATCATCGACGCCGTGGTGGGCGAGAAGGCCCCCTACTTCTGCGGCGGATT
CATCCTGTGCGACAGCGTGACCGGCACCGCTGCAGGCTGGCCGACCCCCTGAA
GAGGCTGTTCAAGCTGGGCAAGCCACTGGCCGCTGACGATGAGCAGCAGATGA
CAGGCGGAGGGCCCTGCACGAGGAAAGCACCAGGTGGAACAGGGTGGGCATCC
TGAGCGAGCTGTGCAAGGCCGTGGAGAGCAGGTACGAGACCGTGGGCACCAGC
ATCATCGTGATGGCTATGACCACACTGGCCAGCTCCGTCAAGAGCTTCTCCTACC
TGAGGGGGCCCCCTATAACTCTCTACGGCTAA SEQ ID NO: 73

ATGGGCGGCGCATGAGAGAAGCCCAGACCAATTACCTACCCAAA SEQ ID NO: 74

GATGGGCGGCGCATGAGAGAAGCCCAGACCAATTACCTACCCAAA SEQ ID NO: 75

GATAGGCGGCGCATGAGAGAAGCCCAGACCAATTACCTACCCAAA SEQ ID NO: 76

ACTCGAGTATGTTACGTGCAAAGGTGATTGTCACCCCCGAAAGACCATATTGTG

ACACACCCTCAGTATCACGCCAAACATTTACAGCCGCGGTGTCAAAAACCGCG

TGGACGTGGTTAACATCCCTGCTGGGAGGATCAGCCGTAATTATTATAATTGGCT

TGGTGCTGGCTACTATTGTGGCCATGTACGTGCTGACCAACCAGAAACATAATTG

AATACAGCAGCAATTGGCAAGCTGCTTACATAGAACTCGCGGCGATTGGCATGC

CGCCTTAAAATTTTTATTTTTTTTTCTTTTTCTTTCCGAATCGGATTTTGTTTTA

ATATTTCAAAAAAAAAAAAAAAAAAAAAAAAAAATCTAGAAAAAAAAAAAAAAAAAA

AA

AA SEQ ID NO: 121

ATGTTTGTTTTTCTTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACA

ACCAGAACTCAATTACCCCTGCATACACTAATTCTTTCACACGTGGTGTTTATTA

CCCTGACAAAGTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTTCTTAC

CTTTCTTTTCCAATGTTACTTGGTTCCATGCTATACATGTCTCTGGGACCAATGGT

ATTOAAGAGGTTTGATAACCCTGTCCTACCATTTAATGATGGTGTTTATTTTGCTTC

CACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTAGATTTCG

AAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTTGTTATTAAAGTCTG

TGAATTTCAATTTTGTAATGATCCATTTTTGGGTGTTTATTACCACAAAAACAACA

AAAGTTGGATGGAAGTGAGTTCAGAGTTTATTCTAGTGCGAATAATTGCACTTT

TGAATATGTCTCTCAGCCTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTT

AAAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTTATTTTAAAATATATT

CTAAGCACACGCCTATTAATTTAGTGCGGTGATCTCCCTCAGGGTTTTTCGGCTTTA

GAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTTCAAACCTTAC

TTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTTCAGGTTGGACAGCT

GGTGCTGCAGCTTATTATGTGGTTATCTTCAACCTAGGACTTTTCTATTAATAATA

TAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACTTGACCCTCTCTCA

GAAACAAAGTGTACGTTGAAATCCTTCACTGTAGAAAAAGGAATCTATCAAAC
TCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTCTAATATTACAA
ATTOTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCAGATTTGCATCTGTTTATGCT
TGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCCTATATAAATT
CCGCATCATTTTCCACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAATGAT
CTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCA
GACAAATCGCTCCAGGGCAAACCTGGAAAGATTGCTGATTATAATTATAAATTAC
CAGATGATTTTACAGGCTGCGTTATAGCTTGGAAATTCTAACAATCTTGATTCTAA
GGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAA
CCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTA
ATGGTGTGAAGTTTTAATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCC
ACTAATGGTGTGGTTACCAACCATACAGAGTAGTAGTACTTTCTTTTTGAACTTCT
ACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTTAAAAAC
AAATGTGTCAATTTCAACTTCAATGGTTTAAACAGGCACAGGTGTTCTTACTGAGT
CTAACAAAAAGTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGACACTAC
TGATGCTGTCCGTGATCCACAGACACTTGAGATTCTTGACATTACACCATGTTCTT
TTGGTGGTGTGAGTGTATAACACCAGGAACAAATACTTCTAACCAGGTTGCTGT
TCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAA
CTTACTCCTACTTGGCGTGTATTATTCTACAGGTTCTAATGTTTTTCAAACACGTGC
AGGCTGTTTAAATAGGGGCTGAACATGTCAACAACCTCATATGAGTGTGACATACCC
ATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCTCGGCGGG
CACGTAGTGTAGCTAGTCAATCCATCATTGCCCTACTATGCTCACTTGGGTGCAGA
AAATTCAGTTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTTTACTATTA
GTGTTACCACAGAAATTCTACCAGTGTCTATGACCAAGACATCAGTAGATTGTAC
AATGTACATTTGTGGTGATTCAACTGAATGCAGCAATCTTTTGTGCAATATGGC
AGTTTTTGTACACAATTAACCGTGCTTTAACTGGAATAGCTGTTGAACAAGACA
AAAACACCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAACACCACCAA
TTAAAGATTTTGGTGTTTAATTTTTTACAAATATTACCAGATCCATCAAACCA
AGCAAGAGGTCATTTATTGAAGATCTACTTTTCAACAAAGTGACACTTGCAGATG
CTGGCTTCATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCTAGAGACCT
CATTTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCACCTTTGCTCACAGATG
AAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACTTCTGGTTG
GACCTTGGTGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAATGGCTTAT
AGGTTTAAATGGTATTGGAGTTACACAGAATGTTCTCTATGAGAACCAAAAATTGA
TTGCCAACCAATTTAATAGTGTATTGGCAAATTTCAAGACTCACTTTCTTCCAC
AGCAAGTGCACCTTGAAAACCTCAAGATGTGGTCAACCAAAATGCACAAGCTTT
AAACACGCTTGTTAAACAACCTTAGCTCCAATTTTGGTGCAATTTCAAGTGTTTTA
AATGATATCCTTTACGTCCTTGACAAAGTTGAGGCTGAAGTGCAAATTGATAGGT
TGATCACAGGCAGACTTCAAAGTTTGCAGACATATGTGACTCAACAATTAATTAG
AGCTGCAGAAATCAGAGCTTCTGCTAATCTTGCTGCTACTAAAATGTCAGAGTGT
GTACTTGGACAATCAAAAAGAGTTGATTTTTGTGGAAAGGGCTATCATCTTATGT
CCTTCCCTCAGTCAGCACCTCATGGTGTAGTCTTCTTGCATGTGACTTATGTCCCT
GCACAAGAAAAGAACTTCACAACCTGCTCCTGCCATTTGTCATGATGAAAAGCA
CACTTTCCTCGTGAAGGTGTCTTTGTTTCAAATGGCACACACTGGTTTGTAAACAC

AAAGGAATTTTTATGAACCACAAATCATTATTOACAGACAACACATTTGTGTCTGG
TAACTGTGATGTTGTAATAGGAATTGTCAACAACACAGTTTATGATCCTTTGCAA
CCTGAATTAGACTCATTCAAGGAGGAGTTAGATAAATATTTTAAAGAATCATAACAT
CACCAGATGTTGATTTAGGTGACATCTCTGGCATTAAATGCTTCAGTTGTAAACAT
TCAAAAAGAAATTGACCGCCTCAATGAGGTTGCCAAGAATTTAAATGAATCTCTC
ATCGATCTCCAAGAACTTGGAAAGTATGAGCAGTATATAAAATGGCCATGGTAC
ATTGGCTAGGTTTTATAGCTGGCTTGATTGCCATAGTAATGGTGACAATTATGCT
TTGCTGTATGACCAGTTGCTGTAGTTGTCTCAAGGGCTGTTGTTCTTGTGGATCCT
GCTGCAAATTTGATGAAGACGACTCTGAGCCAGTGCTCAAAGGAGTCAAATTAC ATTACCA CACAA SEQ ID
NO: 122 ATGTTCTGCTTCTCCTGGTCTGCTGCCTCTGGTCTCCTCACAGTGCGTCAATCTGAC
AACTCGGACTCAGCTGCCACCTGCTTATACTAATAGCTTCACCAGAGGCGTGTAC
TATCCTGACAAGGTGTTTAGAAGCTCCGTGCTGCACTCTACACAGGATCTGTTTC
TGCCATTCTTTAGCAACGTGACCTGGTTCACGCCATCCACGTGAGCGGCACCAA
TGGCACAAAGCGGTTTCGACAATCCCGTGTGCCTTTTAAACGATGGCGTGTACTTC
GCCTTACCAGAGAAGTCCAACATCATCAGAGGCTGGATCTTTGGCACCACACTGG
ATTOCCAAGACACAAGTCTCTGCTGATCGTGAACAATGCCACCAACGTGGTCATCAA
GGTGTGCGAGTTCCAGTTTTGTAATGATCCCTTCTGGGCGTGTACTATCACAAG
ACAATAAGAGCTGGATGGAGTCCGAGTTTAGAGTGTATTCTAGCGCCAACAAC
TGCACATTTGAGTACGTGAGCCAGCCTTTCCTGATGGACCTGGAGGGCAAGCAG
GGCAATTTCAAGAACCTGAGGGAGTTCGTGTTTAAAGAATATCGACGGCTACTTCA
AAATCTACTCTAAGCACACCCCCATCAACCTGGGTGCGCGACCTGCCTCAGGGCTT
CAGCGCCCTGGAGCCCCTGGTGGATCTGCCTATCGGCATCAACATCACCCGTTT
CAGACACTGCTGGCCCTGCACAGAAGCTACCTGACACCCGGGACTCCTCTAGC
GGATGGACCGCCGGCGCTGCCGCCTACTATGTGGGCTACCTCCAGCCCCGGACCT
TCCTGCTGAAGTACAACGAGAATGGCACCATCACAGACGCAGTGGATTGCGCCC
TGGACCCCTGAGCGAGACAAAGTGTAACCTGAAGTCTTTACCGTGGAGAAGG
GCATCTATCAGACATCAGACATCCAATTTAGGGTGCAGCCAACCGAGTCTATCGTGCCTT
TCCTAATATCACAAACCTGTGCCATTTGGCGAGGTGTTCAACGCAACCCGCTT
GCCAGCGTGTACGCCTGGAATAGGAACGGATCAGCAACTGCGTGGCCGACTAT
AGCGTGTGTACAACCTCCGCCTCTTTCAGCACCTTTAAGTGCTATGGCGTGTCCC
CCACAAAGCTGAATGACCTGTGCTTTACCAACGTCTACGCCGATTCTTTCGTGAT
CAGGGGCGACGAGGTGCGCCAGATCGCCCCGGCCAGACAGGCAAGATCGCAG
ATTOCAATTATAAGCTGCCAGACGATTTACCGGCTGCGTGATCGCCTTGAACAG
CAACAATCTGGATTCCAAAGTGGGCGGCAACTACAATTATCTGTACCGGCTGTTT
AGAAAGAGCAATCTGAAGCCCTTCGAGAGGGACATCTCTACAGAAATCTACCAG
GCCGGCAGCACCCCTTGAATGGCGTGGAGGGCTTTAACTGTTATTTCCCACTCC
AGTCTACGGCTTCCAGCCCACAAACGGCGTGGGCTATCAGCCTTACCGCGTGGT
GGTGTGAGCTTTGAGCTGCTGCACGCCCCAGCAACAGTGTGCGGCCCCAAAGAA
GTCCACCAATCTGGTGAAGAACAAGTGCAGTGAACCTTCAACTTCAACGGCCTGAC
CGGCACAGGCGTGTGACCGAGTCCAACAAGAAGTTCCTGCCATTTACAGCAGTTC
GGCAGGGACATCGCAGATACCACAGACGCCGTGCGCGACCCACAGAACCCTGGAG
ATCCTGGACATCACACCCTGCTCTTTCGGCGGCGTGAGCGTGATCACACCCGGCA
CCAATACAAGCAACCAGGTGGCCGTGCTGTATCAGGACGTGAATTGTACCGAGG
TGCCCGTGGCTATCCACGCCGATCAGCTGACCCCAACATGGCGGGTGTACAGCA

CCGGCTCCAACGTCTTCCAGACAAGAGCCGGATGCCTGATCGGAGCAGACCG
TGAACAATTCTATGAGTGCACATCCCAATCCGGCGCCGGCATCTGTGCCTCTTA
CCAGACCCAGACAAACTCTCCAGACGGGCCGGAGCGTGGCCTCCCAGTCTAT
CATCGCTATAACCATGTCCCTGGCGCCGAGAACAGCGTGGCCTACTCTAACAAT
AGCATCGCCATCCCAACCAACTTCACAATCTCTGTGACCACAGAGATCCTGCCCC
TGTCATGACCAAGACATCTGTGGACTGCACAATGTATATCTGTGCGATTCTAC
CGAGTGCAGCAACCTGCTGCTCCAGTACGGCAGCTTTTGTACCCAGCTGAATAGA
GCCCTGACAGGCATCGCCCGTGGAGCAGGATAAGAACACACAGGAGGTGTTCCGC
CAGGTGAGCAAATCTACAAGACCCCCCTATCAAGGGACTTTGGCGGCTTCAATT
TTCCCAGATCCTGCCTGATCCATCCAAGCCTTCTAAGCGGAGCTTTATCGAGGA
CCTGCTGTTCAACAAGTTCGCCGATGCCGGCTTCATCAAGCAGTATGGC
GATTGCCTGGGCGACATCGCAGCCAGGGACCTGATCTGCGCCAGAAGTTTAAAT
GGCCTGACCGTGTGCCACCCCTGCTGACAGATGAGATGATCGCACAGTACACA
AGCGCCCTGCTGGCCGGCACCATCATCATCCGGATGGACCTTCGGCGCAGGAGCC
GCCCTCCAGATCCCTTTGCCATGCAGATGGCCTATAGGTTCAACGGCATCGGCG
TGACCCAGAATGTGCTGTACGAGAACCAGAAGCTGATCGCCAATCAGTTTAACT
CCGCCATCGGCAAGATCCAGGACAGCCTGTCTCTACAGCCAGCGCCCTGGGCA
AGCTCCAGGATGTGGTGAATCAGAACGCCCCAGGCCCTGAATACCCTGGTGAAGC
AGCTGAGCAGCAACTTCGGCGCCATCTCTAGCGTGTGAATGACATCCTGAGCCG
GCTGGACAAGGTGGAGGCAGAGGTGCAGATCGACCGGCTGATCACCGGCCGGCT
CCAGAGCCTCCAGACCTATGTGACACAGCAGCTGATCAGGGCCGCCGAGATCAG
GGCCAGCGCCAATCTGGCAGCAACCAAGATGTCCGAGTGCCTGCTGGCCAGTC
TAAGAGAGGTGGACTTTTGTGGCAAGGGCTATCACCTGATGTCCTTCCCTCAGTCT
GCCCCACAGGGCGGTGGTGTCTTCTGCACGTGACCTACGTGCCCGCCAGGAGAAG
AACTTCACCACAGCCCCTGCCATCTGCCACGATGGCAAGGCCCACTTTCCAAGGG
AGGGCGTGTTCGTGTCCAACGGCACCCACTGGTTTGTGACACAGCGCAATTTCTA
CGAGCCCCAGATCATCACCACAGACAACACCTTCGTGAGCGGCAACTGTGACGT
GGTCATCGGCATCGTGAACAATACCGTGTATGATCCACTCCAGCCCAGCTGGAC
AGCTTTAAGGAGGAGCTGGATAAGTATTTCAAGAATCACACCTCCCCTGACGTG
GATCTGGGCGACATCAGCGGCATCAATGCCTCCGTGGTGAACATCCAGAAGGAG
ATCGACCGCCTGAACGAGGTGGCTAAGAATCTGAACGAGAGCCTGATCGACCTC
CAGGAGCTGGGCAAGTATGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTG
GGCTTCATCGCCGGCCTGATCGCCATCGTGTGATGGTGACCATCATGCTGTGCTGTA
TGACATCCTGCTGTTCTTGCCCTGAAGGGCTGCTGTAGCTGTGGCTCCTGCTGTAA
GTTTGACGAGGATGACTCTGAACCTGTGCTGAAGGGCGGTGAAGCTGCATTACAC Associazione CTAAA
SEQ ID NO: 123 MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVVYYPDKVFRSSVLHSTQDLFLPF
FSNVTWFHAIHVSNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQS
LLIVNNATNVVIKVEFQFCNDPFLGVVYHKNNKSWMESEFRVYSSANNCTFEYVS
QPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLP
IGINITRFQTLALHRSYLPDSSSGWTAGAAAYVGYLQPRTFLLKYNENGTITDA
VDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRF
ASVYAWNKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNLDLCTNVYADSFVIRG
DEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRFRKSN
LKPFERDISTEIYQAGSTPCNGVEGENCYFPLQSYGFQPTNGVGYQPYRVVVLSEFLL

HAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTD
AVRDPQTLLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPT
WRVYSTGSNVFQTRAGCLIGAEHVNNNSYECDIPIGAGICASYQTQTNSPRRARSVAS
QSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTEC
SNLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILP
DPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLT
DEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLI
ANQFNSAIGKIQDLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDIL
SRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSK
RVDFCGKGYHLSFPQSAHPGVVFLHVTVYVPAQEKNFTTAPAICHHDGKAHFPREGV
FVSNGTHWFVTQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEEL
DKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYI
KWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVLKG VKLHYT SEQ ID NO:
77 CCTGAATGACTACGACATAGTCTAGTCCGCCAAGGCCCGCCACC SEQ ID NO: 78
ATGGGCGGCGCATGAGAGAAGCCAGACCAATTACCTACCCAAAATGGAGAAA
GTTACGTTGACATCGAGGAAGACAGCCCATTCCTCAGAGCTTTGCAGCGGGAGCT
TCCCGCAGTTTGAGGTAGAAAGCCAAGCAGGTCCTGATAATGACCATGCTAATG
CCAGAGCGTTTTTCGCATCTGGCTTCAAAAAGTATCGAAACGGAGGTGGACCCATC
CGACACGATCCTTGACATTGGAAGTGCAGCCCGCCGAGAATGTATTCTAAGCAC
AAGTATCATTGTATTCTGTCCGATGAGATGTGCGGAAGATCCGGACAGATTGTATA
AGTATGCAACTAAGCTGAAGAAAACTGTAAGGAAATAACTGATAAGGAATTGG
ACAAGAAAATGAAGGAGCTGGCCGCCGTCATGAGCGACCCTGACCTGGAAACTG
AGACTATGTGCCTCCACGACGACGAGTCGTGTGCTACGAAGGGCAAGTCGCTG
TTTACCAGGATGTATACGCCGTCGACGGCCCCACCAGCCTGTACCACCAGGGCCAA
CAAGGGCGTGAGGGTGGCCTACTGGATCGGCTTCGACACCACACCCTTCATGTTC
AAGAACCTGGCCGGCGCTACCCAGCTACAGCACCAACTGGGCCGACGAGACC
GTGCTGACCGCCAGGAACATCGGCCTGTGCAGCAGCGACGTGATGGAGAGGAGC
CGGAGAGGCATGAGCATCCTGAGGAAGAAATACCTGAAGCCAGCAACAACGT
GCTGTTACAGCGTGGCAGCACCATCTACCACGAGAAGAGGGACCTGCTCAGGAG
CTGGCACCTGCCACGTGTTCCACCTGAGGGGGCAAGCAGAACTACACCTGCAG
GTGCGACCATCGTGAGCTGCGACGGCTACGTGGTGAAGAGGATCGCCATCAG
CCCCGGCCTGTACGGCAAGCCAGCGGCTACGCCGCTACAATGCACAGGGAGGG
CTTCTGTGCTGCAAGGTGACCGACACCCTGAACGGCGAGAGGGTGAGCTTCCC
CGTGTGCACCTACGTGCCCCGCCACCCTGTGCGACCAGATGACCGGCATCCTGGCC
ACCGACGTGAGCGCCGACGACGCCAGAAAGCTGCTCGTGGGCCTGAACCAGAGG
ATCGTGGTCAACGGCAGGACCCAGAGGAACACCAACACAATGAAGAACTACCTG
CTGCCCCGTGGTGGCCAGGCTTTCGCCAGGTGGGCCAAGGAGTACAAGGAGGAC
CAGGAAGACGAGAGGCCCTGGGCTGAGGGACAGGCAGCTGGTGATGGGGCTG
CTGCTGGGCCTTCAGGCGGCACAAGATCACCAGCATCTACAAGAGGCCCGACAC
CCAGACCATCATCAAGGTGAACAGCGACTTCCACAGCTTCGTGCTGCCAGGATC
GGCAGCAACACCCTGGAGATCGGCCTGAGGACCCGGATCAAGGAAGATGCTGGAG
GAACACAAGGAGCCAGCCACTGATCACCGCCGAGGACGTGCAGGAGGCCAA
GTGCGCTGCCGACGAGGCCAAGGAGGTGAGGGAGGCCGAGGAACTGAGGGCCG
CCCTGCCACCCTGGCTGCCGACGTGGAGGAACCCACCCTGGAAGCCGACGTGG

ACCTGATGCTGACAGGAGGCCGGCGCCGGAAGCGTGGAGACACCCAGGGCCTGA
TCAAGGTGACCAGCTACGACGCGGAGGACAAGATCGGCAGCTACGCCGTGCTGA
GCCCACAGGCCGTGCTGAAGTCCGAGAAGCTGAGCTGCATCCACCCACTGGCCG
AGCAGGTGATCGTGATCCCACAGCGGCAGGAAGGGCAGGTACGCCGTGGAGC
CCTACCACGGCAAGGTGGTCTGTCGTCGCCGAGGGCCACGCCATCCCCGTGCAGGACT
TCCAGGCCCTGAGCGAGAGCGCCACCATCGTGTACAACGAGAGGGAGTTCGTGA
ACAGGTACCTGCACCATATCGCCACCCACGGCGGAGCCCTGAACACCGACGAGG
AATACTACAAGACCGTGAAGCCCAGCGAGCACGACGGCGAGTACCTGTACGACA
TCGACAGGAAGCAGTGCCTGAAGAAAGAGCTGGTGACCGGCCTGGGACTGACCG
GCGAGCTGGTGGACCCACCCCTCCACGAGTTCGCCTACGAGAGCCTGAGGACCA
GACCCGCGCTCCCTACCAGGTGCCACCATCGCGTGTACGGCGTGCCCGGCA
GCGGAAAGAGCGGCATCATCAAGAGCGCCGTGACCAAGAAAGACCTGGTGGTC
AGCGCCAAGAAAGAGAACTGCGCCGAGATCATCAGGGACGTGAAGAAGATGAA
AGGCCTGGACGTGAACGCGCGCACCGTGGACAGCGTGCTGCTGAACGGCTGCAA
GCACCCCGTGGAGACCCTGTACATCGACGAGCCTTCGCTTGCCACGCCGGCACC
CTGAGGGCCCTGATCGCCATCATCAGGCCAAGAAAGCCGTGCTGTGCGGGCAG
CCCAAGCAGTGCGGCTTCTTCAACATGATGTGCCTGAAGGTGCACTTCAACCAG
AGATCTGCACCCAGGTGTTCCACAAGAGCATCAGCAGGCGGTGCACCAAGAGCG
TGACCAGCGTCGTGAGCACCTGTTCTACGACAAGAAAATGAGGACCACCAACC
CCAAGGAGACCAAAATCGTGATCGACACCACAGGCAGCACCAAGCCCAAGCAG
GACGACCTGATCCTGACCTGCTTACGGGGCTGGGTGAAGCAGCTGCAGATCGAC
TACAAGGGCAACGAGATCATGACCGCCGCTGCCAGCCAGGGCCTGACCAGGAAG
GGCGTGTACGCCGTGAGGTACAAGGTGAACGAGAACCCACTGTACGCTCCCACC
AGCGAGCACGTGAACGTGCTGCTGACCAGGACCGAGGACAGGATCGTGTGGAAG
ACCCTGGCCGGCGACCCCTGATCAAGACCCTGACCGCCAAGTACCCCGGCAAC
TTCACCGCCACCATCGAAGAGTGGCAGGCCGAGCACGACGCCATCATGAGGCAC
ATCCTGGAGAGGCCCGACCCACCGACGTGTTCCAGAACAAGGCCAACGTGTGC
TGGGCCAAGGCCCTGGTGCCCGTGTGAAGACCGCCGGCATCGACATGACCACA
GAGCAGTGGAACACCGTGGACTACTTCGAGACCGACAAGGCCACAGCGCCGAG
ATCGTGTGAACCAGCTGTGCGTGAGGTTCTTCGGCCTGGACCTGGACAGCGGCC
TGTTACGCGCCCCACCGTGCCACTGAGCATCAGGAACAACCACTGGGACAACA
GCCCCAGCCAAACATGTACGGCCTGAACAAGGAGGTGGTCAGGCAGCTGAGCA
GGCGGTACCCACACGCTGCCAGGGCCGTGGCCACCGGCAGGGGTGTACGACATGA
ACACCGGCACCCTGAGGAACTACGACCCACAGGATCAACCTGGTCCCCTGAACA
GGCGGCTGCCCCACGCCCTGGTGTGCACCACAACGAGCACCCACAGAGCGACT
TCAGCTCCTTCGTGAGCAAGCTGAAAGGCAGGACCGTGCTGGTCTGGGCGAGA
AGCTGAGCGTGCCCGCAAGATGGTGGACTGGCTGAGCGACAGGCCCGAGGCCA
CCTCCGGGCCAGGCTGGACCTCGGCATCCCCGGCGACGTGCCCAAGTACGACA
TCATCTTCGTGAACGTCAGGACCCATACAAGTACCACCATTACCAGCAGTGCGA
GGACCACGCCATCAAGCTGAGCATGCTGACCAAGAAGGCCTGCCTGCACCTGAA
CCCCGGAGGCACCTGCGTGAGCATCGGCTACGGCTACGCCGACAGGGCCAGCGA
GAGCATATTGGCGCCATCGCCAGGCTGTTCAAGTTCAGCAGGGTGTGCAAACC
CAAGAGCAGCCTGGAGGAAACCGAGGTGCTGTTTCGTGTTTCATCGGCTACGACCG
GAAGGCCAGGACCCACAACCCCTACAAGCTGAGCAGCACCCCTGACAAACATCTA

CACCGGCAGCAGGCTGCACGAGGCCGGCTGCGCCCCAGCTACCACGTGGTCAG
GGGCGATATCGCCACCGCCACCGAGGGCGTGATCATCAACGCTGCCAACAGCAA
GGGCCAGCCCGGAGGGCGGAGTGTGCGGCGCCCTGTACAAGAAGTCCCCGAGAG
CTTCGACCTGCAGCCCATCGAGGTGGGCAAGGCCAGGCTGGTGAAGGGCGCCGC
TAAGCACATCATCCACCGCCGTGGGCCCAACTTCAACAAGGTGAGCGAGGTGGA
AGGCGACAAGCAGCTGGCCGAAGCCTACGAGAGCATCGCCAAGATCGTGAACG
ACAATAACTACAAGAGCGTGGCCATCCCACTGCTCAGCG
GCAACAAGGACAGGCTGACCCAGAGCCTGAACCACCTGCTCACCGCCCTGGACA
CCACCGATGCCGACGTGCCATCTACTGCAGGGACAAGAAGTGGGAGATGACCC
TGAAGGAGGCCGTGGCCAGGGCGGAGGCCGTGGAAGAGATCTGCATCAGCGAC
GACTCCAGCGTGACCGAGCCGACGCCGAGCTGGTGAAGGTGCACCCCAAGAGC
TCCCTGGCCGGCAGGAAGGGCTACAGCACCAGCGACGGCAAGACCTTCAGCTAC
CTGGAGGGCACCAAGTTCACCAGGCCGCTAAGGACATCGCCGAGATCAACGCT
ATTGGGCCGTGGCCACCGAGGCCAACGAGCAGGTGTGCATGTACATCCTGGGC
GAGAGCATGTCCAGCATCAGGAGCAAGTGCCCCGTGGAGGAAAGCGAGGCCAG
CACACCACCCAGCACCCCTGCCCTGCCTGTGCATCCACGCTATGACACCCGAGAGG
GTGCAGCGGCTGAAGGCCAGCAGGCCCGAGCAGATCACCGTTGGCAGTCCTTC
CCACTGCCCAAGTACAGGATCACCGGCGTGCAGAAGATCCAGTGCAGCCAGCCC
ATCCTGTTACGCCAAAGGTGCCCGCTACATCCACCCAGGAAGTACCTGGTGG
AGACCCACCCGTGGACGAGACACCCGAGCCAAGCGCCGAGAACCAGAGCACC
GAGGGCACACCCGAGCAGCCACCCCTGATCACCGAGGACGAGACAAGGACCCG
GACCCAGAGCCCATCATTATCGAGGAAGAGGAAGAGGACAGCATCAGCCTGCT
GAGCGACGGCCCCACCCACCAGGTGCTGCAGGTGGAGGCCGACATCCACGGCCC
ACCCAGCGTGTCCAGCTCCAGCTGGAGCATCCACACGCCAGCGACTTCGACGT
GGACAGCCTGAGCATCCTGGACACCCTGGAGGGCGCCAGCGTGACCTCCGGCGC
CACCAGCGCCGAGACCAACAGTACTTCGCCAAGAGCATGGAGTTCCTGGCCAG
GCCCCGTGCCAGTCCCAGGACCGTGTTCAGGAACCCACCCACCCAGTCCCAG
GACCAGGACCCCAAGCCTGGCTCCCAGCAGGGCCTGCAGCAGGACCAGCCTGGT
GAGCACCCACCCGGCGTGAACAGGGTGATCACCGGGAGGAACTGGAGCCCT
GACACCCAGCAGGACCCCAAGCAGGTCCGTGAGCAGGACTAGTCTGGTGTCCAA
CCCACCCGGCGTGAACAGGGTGATCACCGGGAGGAATTCGAGGCCTTCGTGGC
CCAGCAACAGAGACGGTTCGACGCCGGCGCCTACATCTTCAGCAGCGACACCGG
CCAGGGACACCTGCAGCAAAAGAGCGTGAGGCAGCCGTGCTGAGCGAGGTGG
TGCTGGAGAGGACCGAGCTGGAATCAGCTACGCCCCAGGCTGGACCAGGAGA
AGGAGGAACTGCTCAGGAAGAACTGCAGCTGAACCCACCCAGCCAACAGG
AGCAGGTACCAGAGCAGGAAGGTGGAGAACATGAAGGCCATCACCGCCAGGCG
GATCCTGCAGGGCCTGGGACACTACCTGAAGGCCGAGGGCAAGGTGGAGTGCTA
CAGGACCCTGCACCCCGTGCCACTGTACAGCTCCAGCGTGACAGGGCCTTCTCC
AGCCCCAAGGTGGCCGTGGAGGCCTGCAACGCTATGCTGAAGGAGAACTTCCCC
ACCGTGGCCAGCTACTGCATCATCCCCGAGTACGACGCCTACCTGGACATGGTGG
ACGGCGCCAGCTGCTGCCTGGACACCGCCAGCTTCTGCCCCGCAAGCTGAGGA
GCTTCCCCAAGAAACACAGTACCTGGAGCCCACCATCAGGAGCGCCGTGCCCA
GGCCATCCAGAACACCCTGCAGAACGTGCTGGCCGCTGCCACCAAGAGGAACT
GCAACGTGACCCAGATGAGGGAGCTGCCCGTGGTGGACAGCGCTGCCTTCAACG

CGGAGAGGCATGAGCATCCTGAGGAAGAAATACCTGAAGCCCAGCAACAACGT
GCTGTTACAGCGTGGGCAGCACCATCTACCACGAGAAGAGGGACCTGCTCAGGAG
CTGGCACCTGCCACGTGTTCCACCTGAGGGGGCAAGCAGAACTACACCTGCAG
GTGCGACCATCGTGAGCTGCGACGGCTACGTGGTGAAGAGGATCGCCATCAG
CCCCGGCCTGTACGGCAAGCCCAGCGGCTACGCCGCTACAATGCACAGGGAGGG
CTTCTGTGCTGCAAGGTGACCGACACCCTGAACGGCGAGAGGGTGAAGTCCC
CGTGTGCACCTACGTGCCCCGCCACCCTGTGCGACCAGATGACCGGCATCCTGGCC
ACCGACGTGAGCGCCGACGACGCCCAGAAGCTGCTCGTGGGCCTGAACCAGAGG
ATCGTGGTCAACGGCAGGACCCAGAGGAACACCAACACAATGAAGAACTACCTG
CTGCCCCGTGGTGGCCAGGCTTTCGCCAGGTGGGCCAAGGAGTACAAGGAGGAC
CAGGAAGACGAGAGGCCCTGGGCCTGAGGGACAGGCAGCTGGTGTATGGGGCTG
CTGCTGGGCCTTCAGGCGGCACAAGATCACCAGCATCTACAAGAGGCCCGACAC
CCAGACCATCATCAAGGTGAACAGCGACTTCCACAGCTTCGTGCTGCCAGGATC
GGCAGCAACACCCTGGAGATCGGCCTGAGGACCCGGATCAAGGAAGATGCTGGAG
GAACACAAGGAGCCCAGCCACTGATCACCGCCGAGGACGTGCAGGAGGCCAA
GTGCGCTGCCGACGAGGCCAAGGAGGTGAGGGAGGCCGAGGAACTGAGGGCCG
CCCTGCCACCCCTGGCTGCCGACGTGGAGGAACCCACCCTGGAAGCCGACGTGG
ACCTGATGCTGACAGGAGGCCGGCGCCGGAAGCGTGGAGACACCCAGGGCCTGA
TCAAGGTGACCAGCTACGACGGCGAGGACAAGATCGGCAGCTACGCCGTGCTGA
GCCACAGGCCGTGCTGAAGTCCGAGAAGCTGAGCTGCATCCACCCACTGGCCG
AGCAGGTGATCGTGATCCCACAGCGGCAGGAAGGGCAGGTACGCCGTGGAGC
CCTACCACGGCAAGGTGGTTCGTGCTGCCCGAGGGCCACGCCATCCCCGTGCAGGACT
TCCAGGCCCTGAGCGAGAGCGCCACCATCGTGTACAACGAGAGGGAGTTCGTGA
ACAGGTACCTGCACCATATCGCCACCCACGGCGGAGCCCTGAACACCGACGAGG
AATACTACAAGACCGTGAAGCCCAGCGAGCACGACGGCGAGTACCTGTACGACA
TCGACAGGAAGCAGTGCGTGAAGAAAGAGCTGGTGACCGGCCTGGGACTGACCG
GCGAGCTGGTGGACCCACCTTCCACGAGTTCGCCTACGAGAGCCTGAGGACCA
GACCCGCCGCTCCCTACCAGGTGCCACCATCGCGTGTACGGCGTGCCCGGCA
GCGGAAAGAGCGGCATCATCAAGAGCGCCGTGACCAAGAAAGACCTGGTGGTC
AGCGCCAAGAAAGAGAAGTGCGCCGAGATCATCAGGGACGTGAAGAAGATGAA
AGGCCTGGACGTGAACGCGCGCACCGTGGACAGCGTGTGCTGAACGGCTGCAA
GCACCCCGTGGAGACCCTGTACATCGACGAGGCCTTCGCTTGCCACGCCGGCACC
CTGAGGGCCCTGATCGCCATCATCAGGCCAAGAAAGCCGTGCTGTGCGGGCAG
CCCAAGCAGTGCGGCTTCTTCAACATGATGTGCCTGAAGGTGCACTTCAACCACG
AGATCTGCACCCAGGTGTTCCACAAGAGCATCAGCAGGCGGTGCACCAAGAGCG
TGACCAGCGTCGTGAGCACCTGTTCTACGACAAGAAAATGAGGACCACCAACC
CCAAGGAGACCAAAATCGTGATCGACACCACAGGCAGCACCAAGCCCAAGCAG
GACGACCTGATCCTGACCTGCTTACGGGGCTGGGTGAAGCAGCTGCAGATCGAC
TACAAGGGCAACGAGATCATGACCGCCGCTGCCAGCCAGGGCCTGACCAGGAAG
GGCGTGTACGCCGTGAGGTACAAGGTGAACGAGAACCCACTGTACGCTCCCACC
AGCGAGCACGTGAACGTGCTGCTGACCAGGACCGAGGACAGGATCGTGTGGAAG
ACCCTGGCCGGCGACCCCTGATCAAGACCCTGACCGCCAAGTACCCCGGCAAC
TTCACCGCCACCATCGAAGAGTGGCAGGCCGAGCACGACGCCATCATGAGGCAC
ATCCTGGAGAGGCCCGACCCACCGACGTGTTCCAGAACAAGGCCAACGTGTGC

TGGGCCAAGGCCCTGGTGCCCGTGCTGAAGACCGCCGGCATCGACATGACCACA
GAGCAGTGAACACCGTGGACTACTTTCGAGACCGACAAGGCCACAGCGCCGAG
ATCGTGCTGAACCAGCTGTGCGTGAGGTTCTTCGGCCTGGACCTGGACAGCGGCC
TGTTACGCGCCCCACCGTGCCACTGAGCATCAGGAACAACCACTGGGACAACA
GCCCCAGCCAAACATGTACGGCCTGAACAAGGAGGTGGTCAGGCAGCTGAGCA
GGCGGTACCCACACGCTGCCAGGGCCGTGGCCACCGGCAGGGGTGTACGACATGA
ACACCGGCACCCTGAGGAACTACGACCCAGGATCAACCTGGTCCCCTGAACA
GGCGGCTGCCCCACGCCCTGGTGCTGCACCACAACGAGCACCCACAGAGCGACT
TCAGCTCCTTCGTGAGCAAGCTGAAAGGCAGGACCGTGCTGGTCGTGGGCGAGA
AGCTGAGCGTGCCCGGCAAGATGGTGGACTGGCTGAGCGACAGGCCCGAGGCCA
CCTTCCGGGCCAGGCTGGACCTCGGCATCCCCGGCGACGTGCCAAGTACGACA
TCATCTTCGTGAACGTCAGGACCCATACAAGTACCACCATTACCAGCAGTGCGA
GGACCACGCCATCAAGCTGAGCATGCTGACCAAGAAGGCCTGCCTGCACCTGAA
CCCCGAGGCACCTGCGTGAGCATCGGCTACGGCTACGCCGACAGGGCCAGCGA
GAGCATCATTGGCGCCATCGCCAGGCTGTTCAAGTTCAGCAGGGTGTGCAAACC
CAAGAGCAGCCTGGAGGAAACCGAGGTGCTGTTTCGTGTTTCATCGGCTACGACCG
GAAGGCCAGGACCCACAACCCCTACAAGCTGAGCAGCACCCCTGACAAACATCTA
CACCGGCAGCAGGCTGCACGAGGCCGGCTGCGCCCCAGCTACCACGTGGTCAG
GGGCGATATCGCCACCGCCACCGAGGGCGTGATCATCAACGCTGCCAACAGCAA
GGGCCAGCCCGGAGGCGGAGTGTGCGGCGCCCTGTACAAGAAGTCCCCGAGAG
CTTCGACCTGCAGCCATCGAGGTGGGCAAGGCCAGGCTGGTGAAGGGCGCCGC
TAAGCACATCATCCACCGCCGTGGGCCCAACTTCAACAAGGTGAGCGAGGTGGA
AGGCGACAAGCAGCTGGCCGAAGCCTACGAGAGCATCGCCAAGATCGTGAACG
ACAATAACTACAAGAGCGTGGCCATCCCACTGCTCAGCG
GCAACAAGGACAGGCTGACCCAGAGCCTGAACCACCTGCTACCGCCCTGGACA
CCACCGATGCCGACGTGCCATCTACTGCAGGGACAAGAAGTGGGAGATGACCC
TGAAGGAGGCCGTGGCCAGGCGGGAGGCCGTGGAAGAGATCTGCATCAGCGAC
GACTCCAGCGTGACCGAGCCGACGCCGAGCTGGTGAGGGTGCACCCCAAGAGC
TCCCTGGCCGGCAGGAAGGGCTACAGCACCAGCGACGGCAAGACCTTCAGCTAC
CTGGAGGGCACCAAGTTCACCAGGCCGCTAAGGACATCGCCGAGATCAACGCT
ATTGGGCCGTGGCCACCGAGGCCAACGAGCAGGTGTGCATGTACATCCTGGGC
GAGAGCATGTCCAGCATCAGGAGCAAGTGCCCCGTGGAGGAAAGCGAGGCCAG
CACACCACCCAGCACCTGCCCTGCCTGTGCATCCACGCTATGACACCCGAGAGG
GTGCAGCGGCTGAAGGCCAGCAGGCCCGAGCAGATCACCGTTGGCAGTCCTTC
CCACTGCCAAGTACAGGATCACCGGCGTGCAGAAGATCCAGTGCAGCCAGCCC
ATCCTGTTACGCCAAAGGTGCCCGCTACATCCACCCAGGAAGTACCTGGTGG
AGACCCACCCGTGGACGAGACACCCGAGCCAAGCGCCGAGAACCAGAGCACC
GAGGGCACACCCGAGCAGCCACCCCTGATCACCGAGGACGAGACAAGGACCCG
GACCCAGAGCCCATATTATCGAGGAAGAGGAAGAGGACAGCATCAGCCTGCT
GAGCGACGGCCCCACCCACCAGGTGCTGCAGGTGGAGGCCGACATCCACGGCCC
ACCCAGCGTGTCCAGCTCCAGCTGGAGCATCCACACGCCAGCGACTTCGACGT
GGACAGCCTGAGCATCCTGGACACCCTGGAGGGCGCCAGCGTGACCTCCGGCGC
CACCAGCGCCGAGACCAACAGCTACTTCGCCAAGAGCATGGAGTTCCTGGCCAG
GCCCGTGCCAGCTCCCAGGACCGTGTTCAGGAACCCACCCACCCAGCTCCCAG

GACCAGGACCCCAAGCCTGGCTCCCAGCAGGGCCTGCAGCAGGACCAGCCTGGT
GAGACCCCCACCCGGCGTGAACAGGGTGATCACCAGGGAGGAACTGGAGCCCT
GACACCCAGCAGGACCCCAAGCAGGTCCGTGAGCAGGACTAGTCTGGTGTCCAA
CCCACCCGGCGTGAACAGGGTGATCACCAGGGAGGAATTCGAGGCCTTCGTGGC
CCAGCAACAGAGACGGTTCGACGCCGGCCTACATCTTCAGCAGCGACACCG
CCAGGGACACCTGCAGCAAAAGAGCGTGAGGCAGCCGTGCTGAGCGAGGTGG
TGCTGGAGAGGACCGAGCTGAAAATCAGCTACGCCCCAGGCTGGACCAGGAGA
AGGAGGAACTGCTCAGGAAGAACTGCAGCTGAACCCACCCAGCCAACAGG
AGCAGGTACCAGAGCAGGAAGGTGGAGAACATGAAGGCCATCACCGCCAGGCG
GATCCTGCAGGGCCTGGGACACTACCTGAAGGCCGAGGGCAAGGTGGAGTGCTA
CAGGACCCTGCACCCCGTGCCACTGTACAGCTCCAGCGTGACAGGGCCTTCTCC
AGCCCCAAGGTGGCCGTGGAGGCCTGCAACGCTATGCTGAAGGAGAACTTCCCC
ACCGTGGCCAGCTACTGCATCATCCCCGAGTACGACGCCTACCTGGACATGGTGG
ACGGCGCCAGCTGCTGCCTGGACACCGCCAGCTTCTGCCCCGCAAGCTGAGGA
GCTTCCCCAAGAAACACAGCTACCTGGAGCCACCATCAGGAGCGCCGTGCCA
GGCCATCCAGAACACCCTGCAGAACGTGCTGGCCGCTGCCACCAAGAGGAACT
GCAACGTGACCCAGATGAGGGAGCTGCCCCGTGCTGGACAGCGCTGCCTTCAACG
TGGAGTGCTTCAAGAAATACGCCTGCAACAACGAGTACTGGGAGACCTTCAAGG
AGAACCCCATCAGGCTGACCGAAGAGAACGTGGTGAACACTACATACCAAGCTGA
AGGGCCCCAAGGCCGTGCCCTGTTTCGCTAAGACCCACAACCTGAACATGCTGC
AGGACATCCCAATGGACAGGTTTCGTGATGGACCTGAAGAGGGACGTGAAGGTGA
CACCCGGCACCAAGCACACCGAGGAGAGGCCCAAGGTGCAGGTGATCCAGGCC
GCTGACCCACTGGCCACCGCCTACCTGTGCGGCATCCACAGGGAGCTGGTGAGG
CGGCTGAACGCCGTGCTGCTGCCAACATCCACACCCTGTTTCGACATGAGCGCCG
AGGACTTCGACGCCATCATCGCCGAGCACTTCCAGCCGGCGACTGCGTGCTGG
AGACCGACATCGCCAGCTTCGACAAGAGCGAGGATGACGCTATGGCCCTGACCG
CTCTGATGATCCTGGAGGACCTGGGCGTGACGCCGAGCTGTCACCCTGATCGA
GGCTGCCTTCGGCGAGATCAGCTCCATCCACCTGCCACCAAGACCAAGTTCAAG
TTCGGCGCTATGATGAAAAGCGGAATGTTCTGACCCTGTTTCGTGAACACCGTGA
TCAACATTGTGATCGCCAGCAGGGTGCTGCGGGAGAGGCTGACCGGCAGCCCCCT
GCGCTTTCATCGGCGACGACAACATCGTGAAGGGCGTGAAAAGCGACAAGC
TGATGGCCGACAGGTGCGCCACCTGGCTGAACATGGAGGTGAAGATCATCGACG
CCGTGGTGGGCGAGAAGGCCCCCTACTTCTGCGGCGGATTCATCCTGTGCGACAG
CGTGACCGGCACCGCCTGCAGGGTGGCCGACCCCTGAAGAGGCTGTTCAAGCT
GGGCAAGCCACTGGCCGTGACGATGAGCACGACGATGACAGGCGGAGGGCCCT
GCACGAGGAAAGCACCAGGTGGAACAGGGTGGGGCATCCTGAGCGAGCTGTGCA
AGGCCGTGGAGAGCAGGTACGAGACCGTGGGCACCAGCATCATCGTGATGGCTA
TGACCACACTGGCCAGCTCCGTCAAGAGCTTCTCCTACCTGAGGGGGGCCCTAT
AACTCTTACGGCTAACCTGAATGGACTACGACATAGTCTAGTCCGCCAAGGCCG
CCACCATGTTTTGTTTTTCTTGTTTTTATTGCCACTAGTCTTAGTCAGTGTGTTAATC
TTACAACCAGAACTCAATTACCCCTGCATACACTAATTCTTTACACGTGGTGT
TTATTACCCTGACAAAGTTTTTACAGATCCTCAGTTTTTACATTCAACTCAGGACTTGT
TCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATACATGTCTCTGGGACC
AATGGTACTAAGAGGTTTTGATAACCCTGTCTACCATTTAATGATGGTGTTTATTT

TGCTTCCACTGAGAAGTCTAACATAATAAGAGGCTGGATTTTTGGTACTACTTTA
GATTCGAAGACCCAGTCCCTACTTATTGTTAATAACGCTACTAATGTGTGTTATTAA
AGTCTGTGAATTTCAATTTTGTAAATGATCCATTTTGGGTGTTTATTACCACAAAA
ACAACAAAAGTTGGATGGAAAGTGAGTTCAGAGTTTTATTCTAGTGCGAATAATT
GCACTTTTGAATATGTCTCTCAGCCTTTTCTTATGGACCTTGAAGGAAAACAGGG
TAATTTCAAAAATCTTAGGGAATTTGTGTTTAAAGAATATTGATGGTTATTTTAAA
ATATATTCTAAGCACACGCCTATTAATTTAGTGCGGTGATCTCCCTCAGGGTTTTTC
GGCTTTAGAACCATTGGTAGATTTGCCAATAGGTATTAACATCACTAGGTTTCAA
ACTTTACTTGCTTTACATAGAAGTTATTTGACTCCTGGTGATTCTTCTTCAGGTTG
GACAGCTGGTGCTGCAGCTTATTATGTGGGTTATCTTCAACCTAGGACTTTTCTAT
TAAAATATAATGAAAATGGAACCATTACAGATGCTGTAGACTGTGCACCTTGACC
CTCTCTCAGAAACAAAGTGTACGTTGAAATCCTTCACTGTAGAAAAAGGAATCTA
TCAAACCTTCTAACTTTAGAGTCCAACCAACAGAATCTATTGTTAGATTTTCTAAT
ATTACAACTTGTGCCCTTTTGGTGAAGTTTTTAAACGCCACCAGATTTGCATCTGT
TTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGTCTTA
TATAATCCGCATCATTTTCCACTTTTAAAGTGTTATGGAGTGTCTCTACTAAATT
AAATGATCTCTGCTTTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATG
AAGTCAGACAAATCGCTCCAGGGCAAACCTGGAAAGATTGCTGATTATAATTATA
AATTACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAAATTCTAACAATCTTGA
TTCTAAGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATC
TCAAACCTTTTGGAGAGAGATTTTCAACTGAAATCTATCAGGCCGGTAGCACACC
TTGTAATGGTGTGGAAGGTTTTAATTGTTACTTTTCTTTTACAATCATATGGTTTCC
AACCCACTAATGGTGTGTTGTTACCAACCATACAGAGTAGTAGTACTTTCTTTTTGA
ACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAAGTCTACTAATTTGGTT
AAAAACAAATGTGTCAATTTCAACTTCAATGGTTTAAACAGGCACAGGTGTTCTTA
CTGAGTCTAACAAAAGTTTTCTGCCTTTCCAACAATTTGGCAGAGACATTGCTGA
CACTACTGATGCTGTCCGTGATCCAGACACTTGAGATTCTTGACATTACACCA
TGTTCTTTTGGTGGTGTGTCAGTGTATAACACCAGGAACAAATACTTCTAACCAGG
TTGCTGTTCTTTATCAGGATGTTAACTGCACAGAAGTCCCTGTTGCTATTTCATGCA
GATCAACTTACTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAAC
ACGTGCAGGCTGTTTAAATAGGGGCTGAACATGTCAAACAACCTCATATGAGTGTGA
CATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCT
CGGCGGGCACGTAGTGTAGCTAGTCAATCCATCATTGCCTACACTATGTCACTTG
GTGCAGAAAATTCAGTGTGCTTACTCTAATAACTCTATTGCCATACCCACAAATTT
TATTOATTACCATGTGTTTACCACAGAAATTTCTACCAGTGTCTATGACCAAGACATCAGTA
GATTGTACAATGTACATTTGTGGGATTCAACTGAATGCAGCAATCTTTTGTGTC
AATATGGCAGTTTTTGTACACAATTAACCGTGCTTTAACTGGAATAGCTGTTGA
ACAAGACAAAACACCCAAGAAGTTTTTGCACAAGTCAAACAAATTTACAAAAC
ACCACCAATTAAGATTTTGGTGTTTTAAATTTTTTCAAAATATTACCAGATCCAT
CAAAACCAAGCAAGAGGTCATTTATTGAAGATCTACTTTTCAACAAAGTGACACT
TGCAGATGCTGGCTTCATCAAACAATATGGTGATTGCCTTGGTGATATTGCTGCT
AGAGACCTCATTTGTGCACAAAAGTTTAAACGGCCTTACTGTTTTGCCACCTTTGCT
CACAGATGAATGATTGCTCAATACACTTCTGCACTGTTAGCGGGTACAATCACT
TCTGTTGGACCTTTGGGCAGGTGCTGCATTACAAATACCATTTGCTATGCAAA

CTGGCACCTGCCACGTGTTCCACCTGAGGGGGCAAGCAGAACTACACCTGCAG
GTGCGACCATCGTGAGCTGCGACGGCTACGTGGTGAAGAGGATCGCCATCAG
CCCCGGCCTGTACGGCAAGCCCAGCGGCTACGCCGCTACAATGCACAGGGAGGG
CTTCTGTGCTGCAAGGTGACCGACACCCTGAACGGCGAGAGGGTGAGCTTCCC
CGTGTGCACCTACGTGCCCCGCCACCCTGTGCGACCAGATGACCGGCATCCTGGCC
ACCGACGTGAGCGCCGACGACGCCAGAAGCTGCTCGTGGGCCTGAACCAGAGG
ATCGTGGTCAACGGCAGGACCCAGAGGAACACCAACACAATGAAGAACTACCTG
CTGCCCCGTGGTGCCAGGCTTTCGCCAGGTGGGCCAAGGAGTACAAGGAGGAC
CAGGAAGACGAGAGGCCCTGGGCCTGAGGGACAGGCAGCTGGTGTGGGGCTG
CTGCTGGGCCTTCAGGCGGCACAAGATCACCAGCATCTACAAGAGGCCCGACAC
CCAGACCATCATCAAGGTGAACAGCGACTTCCACAGCTTCGTGCTGCCAGGATC
GGCAGCAACACCCTGGAGATCGGCCTGAGGACCCGGATCAAGGAAGATGCTGGAG
GAACACAAGGAGCCCAGCCCCTGATCACCGCCGAGGACGTGCAGGAGGCCAA
GTGCGCTGCCGACGAGGCCAAGGAGGTGAGGGAGGCCGAGGAACTGAGGGCCG
CCCTGCCACCCCTGGCTGCCGACGTGGAGGAACCCACCCTGGAAGCCGACGTGG
ACCTGATGCTGACAGGAGGCCGGCGCCGGAAGCGTGGAGACACCCAGGGCCTGA
TCAAGGTGACCAGCTACGACGGCGAGGACAAGATCGGCAGCTACGCCGTGCTGA
GCCACAGGCCGTGCTGAAGTCCGAGAAGCTGAGCTGCATCCACCCACTGGCCG
AGCAGGTGATCGTGATCCACAGCGGCAGGAAGGGCAGGTACGCCGTGGAGC
CCTACCACGGCAAGGTGGTTCGTGCTGCCCCGAGGGCCACGCCATCCCCGTGCAGGACT
TCCAGGCCCTGAGCGAGAGCGCCACCATCGTGTACAACGAGAGGGAGTTCGTGA
ACAGGTACCTGCACCATATCGCCACCCACGGCGGAGCCCTGAACACCGACGAGG
AATACTACAAGACCGTGAAGCCCAGCGAGCACGACGGCGAGTACCTGTACGACA
TCGACAGGAAGCAGTGCGTGAAGAAAGAGCTGGTGACCGGCCTGGGACTGACCG
GCGAGCTGGTGGACCCACCTTCCACGAGTTCGCCTACGAGAGCCTGAGGACCA
GACCCGCCGCTCCCTACCAGGTGCCACCATCGCGTGTACGGCGTGCCCGGCA
GCGGAAAGAGCGGCATCATCAAGAGCGCCGTGACCAAGAAAGACCTGGTGGTC
AGCGCCAAGAAAGAGAAGTGCGCCGAGATCATCAGGGACGTGAAGAAGATGAA
AGGCCTGGACGTGAACGCGCGCACCGTGGACAGCGTGCTGCTGAACGGCTGCAA
GCACCCCGTGGAGACCCTGTACATCGACGAGGCCTTCGCTTGCCACGCCGGCACC
CTGAGGGCCCTGATCGCCATCATCAGGCCCAAGAAAGCCGTGCTGTGCGGCGAC
CCCAAGCAGTGCGGCTTCTTCAACATGATGTGCCTGAAGGTGCACTTCAACCACG
AGATCTGCACCCAGGTGTTCCACAAGAGCATCAGCAGGCGGTGCACCAAGAGCG
TGACCAGCGTCGTGAGCACCTGTTCTACGACAAGAAAATGAGGACCACCAACC
CCAAGGAGACCAAAATCGTGATCGACACCACAGGCAGCACCAAGCCCAAGCAG
GACGACCTGATCCTGACCTGCTTACAGGGGCTGGGTGAAGCAGCTGCAGATCGAC
TACAAGGGCAACGAGATCATGACCGCCGCTGCCAGCCAGGGCCTGACCAGGAAG
GGCGTGTACGCCGTGAGGTACAAGGTGAACGAGAACCCACTGTACGCTCCCACC
AGCGAGCACGTGAACGTGCTGCTGACCAGGACCGAGGACAGGATCGTGTGGAAG
ACCCTGGCCGGCGACCCCTGATCAAGACCCTGACCGCCAAGTACCCCGGCAAC
TTCACCGCCACCATCGAAGAGTGGCAGGCCGAGCACGACGCCATCATGAGGCAC
ATCCTGGAGAGGCCCGACCCACCGACGTGTTCCAGAACAAGGCCAACGTGTGC
TGGGCCAAGGCCCTGGTGCCCGTGTGAAGACCGCCGGCATCGACATGACCACA
GAGCAGTGGAACACCGTGGACTACTTCGAGACCGACAAGGCCACAGCGCCGAG

ATCGTGCTGAACCAGCTGTGCGTGAGGTTCTTCGGCCTGGACCTGGACAGCGGCC
TGTTACAGCGCCCCACCGTGCCACTGAGCATCAGGAACAACCACTGGGACAACA
GCCCCAGCCAAACATGTACGGCCTGAACAAGGAGGTGGTCAGGCAGCTGAGCA
GGCGGTACCCACACGCTGCCAGGGCCGTGGCCACCGGCAGGGGTGTACGACATGA
ACACCGGCACCCTGAGGAACTACGACCCCAGGATCAACCTGGTCCCCTGAACA
GGCGGCTGCCCCACGCCCTGGTGCTGCACCACAACGAGCACCCACAGAGCGACT
TCAGCTCCTTCGTGAGCAAGCTGAAAGGCAGGACCGTGCTGGTCGTGGGCGAGA
AGCTGAGCGTGCCCGCAAGATGGTGGACTGGCTGAGCGACAGGCCCGAGGCCA
CCTTCCGGGCCAGGCTGGACCTCGGCATCCCCGGCGACGTGCCAAGTACGACA
TCATCTTCGTGAACGTCAGGACCCATACAAGTACCACCATTACCAGCAGTGCGA
GGACCACGCCATCAAGCTGAGCATGCTGACCAAGAAGGCCTGCCTGCACCTGAA
CCCCGGAGGCACCTGCGTGAGCATCGGCTACGGCTACGCCGACAGGGCCAGCGA
GAGCATCATTGGCGCCATCGCCAGGCTGTTCAAGTTCAGCAGGGTGTGCAAACC
CAAGAGCAGCCTGGAGGAAACCGAGGTGCTGTTTCGTGTTTCATCGGCTACGACCG
GAAGGCCAGGACCCACAACCCCTACAAGCTGAGCAGCACCCCTGACAAACATCTA
CACCGGCAGCAGGCTGCACGAGGCCGCTGCGCCCCAGCTACCACGTGGTCAG
GGGCGATATCGCCACCGCCACCGAGGGCGTGATCATCAACGCTGCCAACAGCAA
GGGCCAGCCCGGAGGCGGAGTGTGCGGCGCCCTGTACAAGAAGTCCCCGAGAG
CTTCGACCTGCAGCCATCGAGGTGGGCAAGGCCAGGCTGGTGAAGGGCGCCGC
TAAGCACATCATCCACCGCCGTGGGCCCAACTTCAACAAGGTGAGCGAGGTGGA
AGGCGACAAGCAGCTGGCCGAAGCCTACGAGAGCATCGCCAAGATCGTGAACG
ACAATAACTACAAGAGCGTGGCCATCCCCTGCTCAGCG
GCAACAAGGACAGGCTGACCCAGAGCCTGAACCACCTGCTCACCGCCCTGGACA
CCACCGATGCCGACGTGCCATCTACTGCAGGGACAAGAAGTGGGAGATGACCC
TGAAGGAGGCCGTGGCCAGGCGGGAGGCCGTGGAAGAGATCTGCATCAGCGAC
GACTCCAGCGTGACCGAGCCGACGCCGAGCTGGTGAGGGTGCACCCCAAGAGC
TCCCTGGCCGCGAGGAAGGGCTACAGCACCAGCGACGGCAAGACCTTCAGCTAC
CTGGAGGGCACCAAGTTCACCAGGCCGCTAAGGACATCGCCGAGATCAACGCT
ATTGGGCCGTGGCCACCGAGGCCAACGAGCAGGTGTGCATGTACATCCTGGGC
GAGAGCATGTCCAGCATCAGGAGCAAGTGCCCCGTGGAGGAAAGCGAGGCCAG
CACACCACCCAGCACCTGCCCCTGCCTGTGCATCCACGCTATGACACCCGAGAGG
GTGCAGCGGCTGAAGGCCAGCAGGCCCGAGCAGATCACCGTTGGCAGTCCTTC
CCACTGCCCAAGTACAGGATCACCGGCCGTGCAGAAGATCCAGTGCAGCCAGCCC
ATCCTGTTACGCCAAAGGTGCCCCCTACATCCACCCAGGAAGTACCTGGTGG
AGACCCACCCGTGGACGAGACACCCGAGCCAAGCGCCGAGAACCAGAGCACC
GAGGGCACACCCGAGCAGCCACCCCTGATCACCGAGGACGAGACAAGGACCCG
GACCCAGAGCCCATCATTATCGAGGAAGAGGAAGAGGACAGCATCAGCCTGCT
GAGCGACGGCCCCACCCACCAGGTGCTGCAGGTGGAGGCCGACATCCACGGCCC
ACCCAGCGTGTCCAGCTCCAGCTGGAGCATCCACACGCCAGCGACTTCGACGT
GGACAGCCTGAGCATCCTGGACACCCTGGAGGGCGCCAGCGTGACCTCCGGCGC
CACCAGCGCCGAGACCAACAGCTACTTCGCCAAGAGCATGGAGTTCCTGGCCAG
GCCCCGTGCCAGCTCCAGGACCGTGTTCAGGAACCCACCCACCCAGCTCCCAG
GACCAGGACCCCAAGCCTGGCTCCAGCAGGGCCTGCAGCAGGACCAGCCTGGT
GAGCACCCACCCGGCGTGAACAGGGTGATCACCGGGAGGAACTGGAGCCCT

GACACCCAGCAGGACCCCCAAGCAGGTCCGTGAGCAGGACTAGTCTGGTGTCCAA
CCCACCCGGCGTGAACAGGGTGATCACCAGGGAGGAATTCGAGGCCTTCGTGGC
CCAGCAACAGAGACGGTTCGACGCCGCGCCTACATCTTCAGCAGCGACACCGG
CCAGGGACACCTGCAGCAAAAGAGCGTGAGGCAGCCGTGCTGAGCGAGGTGG
TGCTGGAGAGGACCGAGCTGGAAATCAGCTACGCCCCAGGCTGGACCAGGAGA
AGGAGGAACTGCTCAGGAAGAACTGCAGCTGAACCCACCCCAGCCAACAGG
AGCAGGTACCAGAGCAGGAAGGTGGAGAACATGAAGGCCATCACCGCCAGGCG
GATCCTGCAGGGCCTGGGACACTACCTGAAGGCCGAGGGCAAGGTGGAGTGCTA
CAGGACCCTGCACCCCGTGCCACTGTACAGCTCCAGCGTGACAGGGCCTTCTCC
AGCCCCAAGGTGGCCGTGGAGGCCTGCAACGCTATGCTGAAGGAGAACTTCCCC
ACCGTGGCCAGCTACTGCATCATCCCCGAGTACGACGCTACCTGGACATGGTGG
ACGGCGCCAGCTGCTGCCTGGACACCGCCAGCTTCTGCCCCGCAAGCTGAGGA
GCTTCCCCAAGAAACACAGCTACCTGGAGCCCACCATCAGGAGCGCCGTGCCA
GGCCATCCAGAACACCCTGCAGAACGTGCTGGCCGCTGCCACCAAGAGGAACT
GCAACGTGACCAGATGAGGGAGCTGCCCCGTGCTGGACAGCGCTGCCTTCAACG
TGGAGTGCTTCAAGAAATACGCCTGCAACAACGAGTACTGGGAGACCTTCAAGG
AGAACCCCATCAGGCTGACCGAAGAGAACGTGGTGAACACTACATCACCAAGCTGA
AGGGCCCCAAGGCCGCTGCCCTGTTTCGCTAAGACCCACAACCTGAACATGCTGC
AGGACATCCCAATGGACAGGTTTCGTGATGGACCTGAAGAGGGACGTGAAGGTGA
CACCCGGCACCAAGCACACCGAGGAGAGGCCCAAGGTGCAGGTGATCCAGGCC
GCTGACCCACTGGCCACCGCCTACCTGTGCGGCATCCACAGGGAGCTGGTGAGG
CGGCTGAACGCCGTGCTGCTGCCAACATCCACACCCTGTTTCGACATGAGCGCCG
AGGACTTCGACGCCATCATCGCCGAGCACTTCCAGCCGGCGACTGCGTGCTGG
AGACCGACATCGCCAGCTTCGACAAGAGCGAGGATGACGCTATGGCCCTGACCG
CTCTGATGATCCTGGAGGACCTGGGCGTGACGCCGAGCTGTCACCCTGATCGA
GGCTGCCTTCGGCGAGATCAGTCCATCCACCTGCCACCAAGACCAAGTTCAAG
TTCGGCGCTATGATGAAAAGCGGAATGTTCTGACCCTGTTTCGTGAACACCGTGA
TCAACATTGTGATCGCCAGCAGGGTGCTGCGGGAGAGGCTGACCGGCAGCCCCCT
GCGCTTTCATCGGGCAGCACAACATCGTGAAGGGCGTGAAAAGCGACAAGC
TGATGGCCGACAGGTGCGCCACCTGGCTGAACATGGAGGTGAAGATCATCGACG
CCGTGGTGGGCGAGAAGGCCCTACTTCTGCGGCGGATTCATCCTGTGCGACAG
CGTGACCGGCACCGCCTGCAGGGTGGCCGACCCCTGAAGAGGCTGTTCAAGCT
GGGCAAGCCACTGGCCGCTGACGATGAGCACGACGATGACAGGCGGAGGGCCCT
GCACGAGGAAAGCACCAGGTGGAACAGGGTGGGGCATCCTGAGCGAGCTGTGCA
AGGCCGTGGAGAGCAGGTACGAGACCGTGGGCACCAGCATCATCGTGATGGCTA
TGACCACACTGGCCAGCTCCGTCAAGAGCTTCTCCTACCTGAGGGGGGCCCTAT
AACTCTTACGGCTAACCTGAATGGACTACGACATAGTCTAGTCCGCCAAGGCCG
CCACCATGTTTCGTCTTCTGGTCCCTGCTGCCTCTGGTCTCCTCACAGTGCGTCAAT
CTGACAACCTCGGACTCAGCTGCCACCTGCTTATACTAATAGCTTACCAGAGGGCG
TGTAATATCCTGACAAGGTGTTTAGAAGCTCCGTGCTGCACTCTACACACAGGATCT
GTTCTGCCATTCTTTAGCAACGTGACCTGGTTCCACGCCATCCACGTGAGCGGC
ACCAATGGCACAAAGCGGTTTCGACAATCCCGTGCTGCCTTTTAAACGATGGCGTGT
ATTOTCGCCTTACCGAGAAGTCCAACATCATCAGAGGCTGGATCTTTGGCACCAC
ACTGGACTCCAAGACACAAGTCTCTGCTGATCGTGAACAATGCCACCAACGTGGTC

ATCAAGGTTGTGCGAGTTCCAGTTTTGTAATGATCCCTTCTGGGCGTGTACTATC
ACAAGAACAATAAGAGCTGGATGGAGTCCGAGTTTAGAGTGTATTCTAGCGCCA
ACAAGTGCACATTTGAGTACGTGAGCCAGCCTTTCCTGATGGACCTGGAGGGCA
ACGAGGGCAATTTCAAGAACCTGAGGGAGTTTCGTGTTTAAGAATATCGACGGCT
ATTOCAAAATCTACTCTAAGCACACCCCATCAACCTGGGTGCGCGACCTGCCTCA
GGGCTTCAGCGCCCTGGAGCCCCCTGGTGGATCTGCCTATCGGCATCAACATCACC
CGTTTTAGACACTGCTGGCCCTGCACAGAAGCTACCTGACACCCGGCGACTCCT
CTAGCGGATGGACCGCCGGCGCTGCCGCTACTATGTGGGCTACCTCCAGCCCCG
GACCTTCTGCTGAAGTACAACGAGAATGGCACCATCACAGACGCAGTGGATTG
CGCCCTGGACCCCTGAGCGAGACAAAGTGTACTGAAGTCCTTTACCGTGA
GAAGGGCATCTATCATCAGACATCCAATTTAGGGTGCAGCCAACCGAGTCTATCGT
GCGCTTTCCTAATATCACAAACCTGTGCCATTTGGCGAGGTGTTCAACGCAACC
CGTTTCGCCAGCGTGTACGCCTGGAATAGGAAGCGGATCAGCAACTGCGTGGCC
GACTATAGCGTGTGTACAACCTCCGCCTTTTCAGCACCTTTAAGTGTATGGCG
TGTCACCCACAAAGTGAATGACCTGTGCTTTACCAACGTCTACGCCGATTCTTT
CGTGATCAGGGGCGACGAGGTGCGCCAGATCGCCCCGGCCAGACAGGCAAGAT
CGCAGACTACAATTATAAGCTGCCAGACGATTTACCGGCTGCGTGATCGCCTGG
AACAGCAACAATCTGGATTCCAAGTGGGCGGCAACTACAATTATCTGTACCGG
CTGTTTAGAAAGAGCAATCTGAAGCCCTTCGAGAGGGACATCTCTACAGAAATC
TACCAGGGCCGGCAGCACCCCTTGCAATGGCGTGGAGGGCTTTAACTGTTATTTCC
CACTCCAGTCTACGGCTTCCAGCCCACAAACGGCGTGGGCTATCAGCCTTACCG
CGTGGTGGTGCTGAGCTTTGAGCTGCTGCACGCCCCAGCAACAGTGTGCGGCCCC
AAGAAGTCCACCAATCTGGTGAAGAACAAGTGCAGTGAACCTCAACTCAACGGC
CTGACCGGCACAGGCGTGCTGACCGAGTCCAACAAGAAGTTCCTGCCATTTAG
CAGTTCGGCAGGGACATCGCAGATACCACAGACGCCGTGCGCGACCCACAGACC
CTGGAGATCCTGGACATCACACCCTGCTCTTTCGGCGGCGTGAGCGTGATCACAC
CCGGCACCAATACAAGCAACCAGGTGGCCGTCGCTGTATCAGGACGTGAATTGTA
CCGAGGTGCCCCGTGGCTATCCACGCCGATCAGCTGACCCCAACATGGCGGGTGT
ACAGCACCGGCTCCAACGTCTTCCAGACAAGAGCCGGATGCCTGATCGGAGCAG
AGCACGTGAACAATTCCTATGAGTGCAGACATCCAATCGGCGCCGGCATCTGTGC
CTCTTACCAGACCCAGACAACTCTCCCAGACGGGCCCCGGAGCGTGGCCTCCA
GTCTATCATCGCTATAACATGTCCCTTGGGCGCCGAGAACAGCGTGGCCTACTCT
AACAAATAGCATCGCCATCCCAACCAACTTACAATCTCTGTGACCACAGAGATCC
TGCCCCTGTCCATGACCAAGACATCTGTGGACTGCACAATGTATATCTGTGGCGA
TTCTACCGAGTGCAGCAACCTGCTGCTCCAGTACGGCAGCTTTTGTACCCAGCTG
AATAGAGCCCTGACAGGCATCGCCGTGGAGCAGGATAAGAACACACAGGAGGT
GTTCCGCCAGGTGAAGCAAATCTACAAGACCCCCCTATCAAGGGACTTTGGCGG
CTTCAATTTTCCAGATCCCTGCCTGATCCATCCAAGCCTTCTAAGCGGAGCTTTA
TCGAGGACCTGCTGTTCAACAAGTGACCCTGGCCGATGCCGGCTTCATCAAGCA
GTATGGCGATTGCCTGGGCGACATCGCAGCCAGGGACCTGATCTGCGCCAGAA
GTTTAATGGCCTGACCGTGTGCCACCCCTGCTGACAGATGAGATGATCGCACAG
TACACAAGCGCCCTGCTGGCCGGCACCATCATCATCCGGATGGACCTTCGGCGCA
GGAGCCGCCCTCCAGATCCCCTTTGCCATGCAGATGGCCTATAGGTTCAACGGCA
TCGGCGTGACCCAGAATGTGCTGTACGAGAACCAGAAGCTGATCGCCAATCAGT

TTAACTCCGCCATCGGCAAGATCCAGGACAGCTGTCCTCTACAGCCAGCGCCCT
GGGCAAGCTCCAGGATGTGGTGAATCAGAACGCCCCAGGCCCTGAATACCTGGT
GAAGCAGCCTGAGCAGCAACTTCGGCGCCATCTCTAGCGTGCTGAATGACATCCT
GAGCCGGCTGGACAAGGTGGAGGCAGAGGTGCAGATCGACCGGCTGATCACCG
GCCGGCTCCAGAGCCTCCAGACCTATGTGACACAGCAGCTGATCAGGGCCGCGC
AGATCAGGGCCAGCGCCAATCTGGCAGCAACCAAGATGTCGGAGTGCGTGCTGG
GCCAGTCTAAGAGAGTGGACTTTTTGTGGCAAGGGCTATCACCTGATGTCCTTCCC
TCAGTCTGCCCCACAGGCGTGTTTCTGCACGTGACCTACGTGCCCCGCCAG
GAGAAGAACTTACCACAGCCCCTGCCATCTGCCACGATGGCAAGGCCACTTTC
CAAGGGAGGGCGTGTTCTGTGTCACCGCACCCACTGGTTTGTGACACAGCGCA
ATTCTACGAGCCCCAGATCATCACACAGACAACACCTTCGTGAGCGGCAACTG
TGACGTGTCATCGGCATCGTGAACAATACCGTGTATGATCCACTCCAGCCGAG
CTGGACAGCTTTAAGGAGGAGCTGGATAAGTATTTCAAGAATCACACCTCCCCTG
ACGTGGATCTGGGCGACATCAGCGGCATCAATGCCTCCGTGGTGAACATCCAGA
AGGAGATCGACCGCCTGAACGAGGTGGCTAAGAATCTGAACGAGAGCCTGATCG
ACCTCCAGGAGCTGGGCAAGTATGAGCAGTACATCAAGTGGCCCTGGTACATCT
GGCTGGGCTTCATCGCCGCGCTGATCGCCATCGTGATGGTGACCATCATGCTGTG
CTGTATGACATCCTGCTGTTCTTGCCTGAAGGGCTGCTGTAGCTGTGGCTCCTGCT
GTAAGTTTGACGAGGATGACTCTGAACCTGTGCTGAAGGGCGGTGAAGCTGCATT
ACACCTAAACTCGAGTATGTTACGTGCAAAGGTGATTGTCACCCCCGAAAGAC
CATATTGTGACACACCCTCAGTATCACGCCAAACATTTACAGCCGCGGTGTCAA
AAACCGCGTGGACGTGGTTAACATCCCTGCTGGGAGGATCAGCCGTAATTATTAT
AATTGGCTTGGGTGCTGGCTACTATTGTGGCCATGTACGTGCTGACCAACCAGAAA
CATAATTGAATACAGCAGCAATTGGCAAGCTGCTTACATAGAACTCGCGGCGAT
TGGCATGCCGCTTAAAATTTTTATTTTTATTTTTTTCTTTTCTTTTCCGAATCGGATT
TTGTTTTTAATATTTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTCTAGAAAAAAAAA
AA
AA SEQ ID NO: 79
MEKVHVDIEEDSPFLRALQRSFPQFEVEAKQVTDNDHANARAFSHLASKLIETEVDP
SDTILDIGSAPARMYSKHKYHCICPMRCAEDPDRLYKYATKLKKNCKEITDKELDK
KMKELAAVMSDPDLETETMCLHDESCRYEGQVAVYQDVYAVDGPSTLYHQANK
GVRVAYWIGFDTPPFMFKNLGAYPSYSTNWADETVLTARNIGLCSSDVMERSRRG
MSILRKKYLKPSNNVLFVSGSTIYHEKRDLLRSWHLPSVFHLRGKQNYTCRCETIVSC
DGYVVKRIAISPGLYGKPSGYAATMHREGFLCCKVTDLTNGERSFPVCTYVPATLC
DQMTGILATDVSADDAQKLLVGLNQRIVVNGRTQRNTNTMKNYLLPVVAQAFARW
AKEYKEDQEDERPLGLRDRQLVMGCCWAFRRHKITSIYKRPDTQTIKVNNSDFHSFV
LPRIGSNTLEIGLRTRIRKMLEEHKEPSPLITAEDVQEAKCAADEAKEVREAEELRAA
LPPLAADVEEPTLEADVLDMLQEAGAGSVETPRGLIKVTSYDGEDKIGSYAVLSPQA
VLKSEKLSCHPLAEQVIVITHSGRKGRYAVEPYHGKVVVPEGHAIQVDFQALSESA
TIVIYNEREFVNRYLHHIATHGGALNTDEEYKTVKPSHDGEYLYDIDRKQCVKKE
VTGLGLTGELVDPPFHEFAYESLRTRPAAPYQVPTIGVYGVPGSGKSGIISAVTKKD
LVVSAKKENCAEIRDVKMKMGLDVNARTVDSVLLNGCKHPVETLYIDEAFACHAG
TLRALIAIIRPKKAVLCGDPKQCGFFNMMCLKVHFNHEICTQVFHKSISRRTKSVTS
VVSTLFYDKKMRTTNPKETKIVIDTTGSTKPKQDDLILTCFRGWVKQLQIDYKNEI

ACAGAAGCTACCTGACACCCGGCGACTCCTCTAGCGGATGGACCGCCGGCGCTG
CCGCCTACTATGTGGGCTACCTCCAGCCCCGACCTTCCTGCTGAAGTACAACGA
GAATGGCACCATCACAGACGCAGTGGATTGCGCCCTGGACCCCTGAGCGAGAC
AAAGTGTACTACTGAAGTCCTTTACCGTGGAGAAGGGCATCTATCAGACATCCAA
TTTCAGGGTGCAGCCAACCGAGTCTATCGTGCGCTTTCCTAATATCACAAACCTG
TGCCCATTTGGCGAGGTGTTCAACGCAACCCGCTTCGCCAGCGTGTACGCCTGGA
ATAGGAAGCGGATCAGCAACTGCGTGGCCGACTATAGCGTGTGTACAACCTCCG
CCTCTTTCAGCACCTTTAAGTGTATGGCGTGTCCCCACAAAGCTGAATGACCT
GTGCTTTACCAACGTCTACGCCGATTCTTTTCGTGATCAGGGGCGACGAGGTGCGC
CAGATCGCCCCGGCCAGACAGGCAAGATCGCAGACTACAATTATAAGCTGCCA
GACGATTTACCGGCTGCGTGTATCGCCTGGAACAGCAACAATCTGGATTCCAAA
GTGGGCGGCAACTACAATTATCTGTACCGGCTGTTTAGAAAGAGCAATCTGAAG
CCCTTCGAGAGGGACATCTCTACAGAAATCTACCAGGGCCGGCAGCACCCCTTGC
AATGGCGGTGGAGGGCTTTAACTGTTATTTCCCACTCCAGTCCTACGGCTTCCAGC
CCACAAACGGCGTGGGCTATCAGCCTTACCGCTGGTGGTGTGCTGAGCTTTGAGCT
GCTGCACGCCCCAGCAACAGTGTGCGGCCCAAGAAGTCCACCAATCTGGTGAA
GAACAAGTGCCTGAACCTCAACTTCAACGGCCTGACCGGCACAGGCGTGTGAC
CGAGTCCAACAAGAAGTTCCTGCCATTTACAGCAGTTCGGCAGGGACATCGCAGA
TACCACAGACGCCGTGCGCGACCCACAGACCCTGGAGATCCTGGACATCACACC
CTGCTCTTTCGGCGGCGTGTGAGCGTGTATCACACCCGGCACCAATACAAGCAACCA
GGTGGCCGTGTGTATCAGGACGTGAATTGTACCGAGGTGCCCGTGGGCTATCCAC
GCCGATCAGCTGACCCCAACATGGCGGGTGTACAGCACCGGCTCCAACGTCTTCC
AGACAAGAGCCGGATGCCTGATCGGAGCAGAGCACGTGAACAATTCCTATGAGT
GCGACATCCCAATCGGCGCCGGCATCTGTGCCTCTTACCAGACCCAGACAAACTC
TCCCAGACGGGCCCGGAGCGTGGCCTCCAGTCTATCATCGCCTATAACCATGTCC
CTGGGCGCCGAGAACAGCGTGGCCTACTCTAACAATAGCATCGCCATCCCAACC
AACTTCACAATCTCTGTGACCACAGAGATCCTGCCCGTGTCCATGACCAAGACAT
CTGTGGACTGCACAAGTATATCTGTGGGATTCTACCGAGTGCAGCAACCTGCT
GCTCCAGTACGGCAGCTTTTGTACCCAGCTGAATAGAGCCCTGACAGGCATCGCC
GTGGAGCAGGATAAGAACACACAGGAGGTGTTTCGCCAGGTGAAGCAAATCTAC
AAGACCCCCCTATCAAGGACTTTGGCGGCTTCAATTTTCCAGATCCTGCCTG
ATCCATCCAAGCAGCCTTCTAAGCGGAGCTTTATCGAGGACCTGCTGTTCAACAAGGT
GACCCTGGCCGATGCCGGCTTCATCAAGCAGTATGGCGATTGCCTGGGCGACATC
GCAGCCAGGGACCTGATCTGCGCCAGAAGTTTAAATGGCCTGACCGTGTGCCA
CCCCTGCTGACAGATGAGATGATCGCACAGTACACAAGCGCCCTGCTGGCCGGC
ACCATCACATCAGATCCGGATGGACCTTCGGCGCAGGAGCCGCCCTCCAGATCCCCTTG
CCATGCAGATGGCCTATAGGTTCAACGGCATCGGCGTGACCCAGAATGTGCTGT
ACGAGAACCAGAAGCTGATCGCCAATCAGTTTAACTCCGCCATCGGCAAGATCC
AGGACAGCCTGTCTCTACAGCCAGCGCCCTGGCAAGCTCCAGGATGTGGTGA
ATCAGAACGCCCCAGGCCCTGAATACCCTGGTGAAGCAGCTGAGCAGCAACTTCG
GGCCATCTCTAGCGTGTGAATGACATCCTGAGCCGGCTGGACAAGGTGGAGG
CAGAGGTGCAGATCGACCGGCTGATCACCGGCCGGCTCCAGAGCCTCCAGACCT
ATGTGACACAGCAGCTGATCAGGGCCGCCGAGATCAGGGCCAGCGCCAATCTGG
CAGCAACCAAGATGTCCGAGTGCCTGCTGGGCCAGTCTAAGAGAGTGGACTTTT

GTGGCAAGGGCTATCACCTGATGTCCTTCCCTCAGTCTGCCCCACACGGCGTGGT
GTTTCTGCACTGACCTACGTGCCCCCCCAGGAGAAGAACTTCACCACAGCCCCT
GCCATCTGCCACGATGGCAAGGCCACTTTCCAAGGGAGGGCGTGTTCGTGTCCA
ACGGCACCCACTGGTTTGTGACACAGCGCAATTTCTACGAGCCCCAGATCATCAC
CACAGACAACACCTTCGTGAGCGGCAACTGTGACGTGGTCATCGGCATCGTGAA
CAATACCGTGTATGATCCACTCCAGCCCCGAGCTGGACAGCTTTAAGGAGGAGCT
GGATAAGTATTTCAAGAATCACACCTCCCCTGACGTGGATCTGGGCGACATCAGC
GGCATCAATGCCTCCGTGGTGAACATCCAGAAGGAGATCGACCGCCTGAACGAG
GTGGCTAAGAATCTGAACGAGAGCCTGATCGACCTCCAGGAGCTGGGCAAGTAT
GAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGCCTGA
TCGCCATCGTGATGGTGACCATCATGCTGTGCTGTATGACATCCTGCTGTTCTTGC
CTGAAGGGCTGCTGTAGCTGTGGCTCCTGCTGTAAGTTTGACGAGGATGACTCTG
AACCTGTGCTGAAGGGCGTGAGCTGCATTACACCTAAACTCGAGCTAGTGACT
GACTAGGATCTGTTACCACTAAACCAGCCTCAAGAACACCCGAATGGAGTCTC
TAAGCTACATAATAACCACTTACACTTACAAAATGTTGTCCCCAAAATGTAGCC
ATTCGTATCTGCTCCTAATAAAAAGAAAAGTTTCTTCACATTCTAGAAAAAAAAAA
AA
AA SEQ ID NO: 82
AGGAACTTAAGTCAACACAACATATACAAAACAAACGAATCTCAAGCAATCAA
GCATTCTACTTCTATTGCAGCAATTTAAATCATTCTTTTAAAGCAAAGCAATTT
TCTGAAAATTTTACCATTTACGAACGATAGCCACC SEQ ID NO: 83
ATTOGAGCTAGTGACTGACTAGGATCTGTTACCACTAAACCAGCCTCAAGAAC
ACCCGAATGGAGTCTCTAAGCTACATAATAACCACTTACACTTACAAAATGTTGT
CCCCAAAATGTAGCCATTTCGTATCTGCTCCTAATAAAAAGAAAAGTTTCTTCACA
TTCTAGAA
AAA AAA SEQ ID N.: 1
GAGGAAATTO AAGAUGGG SEQ ID N.: 2 GGAUGGG SEQ ID N.: 3 GGAUAGG SEQ ID N.: 4 GGAGAGG
SEQ ID NO: 58 GGGAUGGG SEQ ID NO: 59 GAGAGG SEQ ID NO: 60 BUFFALO SEQ ID NO: 61
GAGAUGGG SEQ ID NO: 62 BUFFALO SEQ ID NO: 63 BAA ... SEQ ID NO: 64 GAGAGG SEQ ID NO: 65
BUFFALO (Sequenza di RNA per un costrutto con due promotori subgenomici, Luc ed E3L) SEQ ID
NO: 127 atgggcgccgatgagagaagcccagaccaattacctacccaaatggagaaagttcacgttgacatcgaggaagacagccatt
cctcagagcttgcagcggagcttcccgagttgaggtagaagccaagcaggtcactgataatgacatgctaataccagagcgtt
ttcgcacatggttcaaaactgatcgaaacggaggtggaccatccgacacgatccttgacattggaagtgcgcccggcgcagaat
gtatttaagcacaagtatcattgtatctgctccgatgatgatgtgcggaagatccggacagattgtataagtatgcaactaagctgaa
gaaaaagiregtaaggaaataagiregataaggaattggacaagaaaatgaaggagctggccgctcatgagcgaccctgacctgga
aactgagactatgtgctccacgacgacgagtcgtgctgctacgaagggaagtcgctgtttaccaggatgtatacgcCGTCGAC
GGCCCCACCAGCCTGTACCACCAGGGCCAACAAGGGCGTGAGGGTGGCCTACTGGATCGGCTTCGAC
ACCACACCCTTCATGTTCAAGAACCTGGCCGGCGCCTACCCAGCTACAGCACCAACTGGGCCGACG
AGACCGTGCTGACCGCCAGGAACATCGCCTGTGACGACGACGCTGATGGAGAGGAGCCGGAGA
GGCATGAGCATCCTGAGGAAGAAATACCTGAAGCCCAGCAACAACGTGCTGTTTACGCGTGGGGCAGC
ACCATCTACCACGAGAAGAGGGACCTGCTCAGGAGCTGGCACCTGCCACGTGTTCCACCTGAGG
GGCAAGCAGAACTACACCTGCAGGTGCGAGACCATCGTGAGCTGCGACGGCTACGTGGTGAAGAG
GATCGCCATCAGCCCCGGCCTGTACGGCAAGCCCAGCGGCTACGCCGCTACAATGCACAGGGAGGG
CTTCTGTGCTGCAAGGTGACCGACACCCTGAACGGCGAGAGGGTGAGCTTCCCCGTGTGCACCTA

CGTGCCCCACCCTGTGCGACCAGATGACCGGCATCCTGGCCACCGACGTGAGCGCCGACGACGC
CCAGAAGCTGCTCGTGGGCTGAACCAGAGGATCGTGGTCAACGGCAGGACCCAGAGGAACACCA
ACACAATGAAGAACTACCTGCTGCCCCGTGGTGGCCAGGCTTTCGCCAGGTGGGCCAAGGAGTACA
AGGAGGACCAGGAAGACGAGAGGCCCTGGGCTGAGGGACAGGCAGCTGGTGATGGGCTGCTG
CTGGGCTTCAGGCGGCACAAGATCACCAGCATCTACAAGAGGCCCGACACCCAGACCATCATCAA
GGTGAACAGCGACTTCCACAGCTTTCGTGCTGCCAGGATCGGCAGCAACACCCCTGGAGATCGGCCT
GAGGACCCGGATCAGGAAGATGCTGGAGGAACACAAGGAGCCAGCCACTGATCACCGCCGAGG
ACGTGCAGGAGGCCAAGTGCCTGCCGACGAGGCCAAGGAGGTGAGGGAGGCCGAGGAACTGAG
GGCCGCCCTGCCACCCCTGGCTGCCGACGTGGAGGAACCCACCCTGGAAGCCGACGTGGACCTGAT
GCTGCAGGAGGCCGGCGCCGAAGCGTGGAGACACCCAGGGCCCTGATCAAGGTGACCAGCTACG
ACGGCGAGGACAAAGATCGGCAGCTACGCCGTGCTGAGCCACAGGCCGTGCTGAAGTCCGAGAAG
CTGAGCTGCATCCACCCACTGGCCGAGCAGGTGATCGTGATCACCCACAGCGGCAGGAAGGGCAG
GTACGCCGTGGAGCCCTTACCACGGCAAGGTGGTCGTGCCGAGGGCCACGCCATCCCCGTGCAGGA
CTTCCAGGCCCTGAGCGAGAGCGCCACCATCGTGTACAACGAGAGGGAGTTCGTGAACAGGTACCT
GCACCATATCGCCACCCACGGCGGAGCCCTGAACACCGACGAGGAATACTACAAGACCGTGAAGCC
CAGCGAGCACGACGGCGAGTACCTGTACGACATCGACAGGAAGCAGTGCCTGAAGAAAGAGCTGG
TGACCGCCTGGGACTGACCGCGAGCTGGTGGACCCACCCTTCCACGAGTTCGCCTACGAGAGCC
TGAGGACCAGACCCGCCGCTCCCTACCAGGTGCCACCATCGGCCGTGTACGGCGTGCCCGGCAGCG
GAAAGAGCGGCATCATCAAGAGCGCCGTGACCAAGAAAAGACCTGGTGGTCAGCGCAAGAAAAGAG
AACTGCGCCGAGATCATCAGGGACGTGAAGAAGATGAAAGGCCCTGGACGTGAACGCGCGCACCGT
GGACAGGTGCTGCTGAACGGCTGCAAGCACCCCGTGGAGACCCTGTACATCGACGAGGCCTTCGC
TTGCCACGCCGGCACCCCTGAGGGCCCTGATCGCCATCATCAGGCCCAAGAAAGCCGTGCTGTGCGG
CGACCCCAAGCAGTGGGCTTCTTCAACATGATGTGCCTGAAGGTGCACTTCAACCACGAGATCTGC
ACCCAGGTGTTCCACAAGAGCATCAGCAGGCGGTGCACCAAGAGCGTGACCAGCGTCGTGAGCACC
CTGTTCTACGACAAGAAAATGAGGACCACCAACCCCAAGGAGACCAAAAATCGTGATCGACACCACA
GGCAGCACCAAGCCCAAGCAGGACGACCTGATCCTGACCTGCTTACGGGGGCTGGGTGAAGCAGCTG
CAGATCGACTACAAGGGCAACGAGATCATGACCGCCGCTGCCAGCCAGGGCCTGACCAGGAAGGG
CGTGTACGCCGTGAGGTACAAGGTGAACGAGAACCCACTGTACGCTCCCACCAGCGAGCAGTGAA
CGTGCTGCTGACCAGGACCGAGGACAGGATCGTGTGGAAGACCCTGGCCGGCGACCCCTGATCA
AGACCCTGACCGCAAGTACCCGGCAACTTACCGCCACCATCGAAGAGTGGCAGGCCGAGCACG
ACGCCATCATGAGGCACATCCTGGAGAGGCCCGACCCACCGACGTGTTCCAGAACAAGGCCAACG
TGTGTCTGGGCCAAGGCCCTGGTGGCCGTGCTGAAGACCGCCGGCATCGACATGACCACAGAGCAGT
GGAACACCGTGGACTACTTCGAGACCGACAAGGCCACAGCGCCGAGATCGTGCTGAACCAGCTGT
GCGTGAGGTTCTTCGGCCTGGACCTGGACAGCGGCCTGTTACGCGCCCCACCGTGCCACTGAGCAT
CAGGAACAACCACTGGGACAACAGCCCCAGCCAAACATGTACGGCCTGAACAAGGAGGTGGTCA
GGCAGCTGAGCAGGCGGTACCCACACGCTGCCAGGGCCGTGGCCACCGGCAGGGTGTACGACATG
AACACCGCACCCCTGAGGAACTACGACCCAGGATCAACCTGGTGGCCGTGAACAGGCGGCTGCC
CACGCCCTGGTGTGACCAACAACGAGCACCCACAGAGCGACTTCAGCTCCTTCGTGAGCAAGCTGA
AAGGGCAGGACCGTGTGGTCTGGGCGAGAAGCTGAGCGTGGCCGCAAGATGGTGGACTGGCTG
AGCGACAGGCCCCGAGGCCACCTTCCGGGCCAGGCTGGACCTCGGCATCCCCGGCGACGTGCCAAG
TACGACATCATCTTCGTGAACGTCAGGACCCATAACAAGTACCACCATTACCAGCAGTGCAGGACC
ACGCCATCAAGCTGAGCATGCTGACCAAGAAGGCCTGCCTGCACCTGAACCCCGGAGGCACCTGCG
TGAGCATCGGCTACGGCTACGCCGACAGGGCCAGCGAGAGCATATTGGCGCCATCGCCAGGCTGT
TCAAGTTCAGCAGGGTGTGCAAACCAAGAGCAGCCTGGAGGAAACCGAGGTGCTGCTGTTCA

TCGGCTACGACCGGAAGGCCAGGACCCACAACCCCTACAAGCTGAGCAGCACCCCTGACAAACATCT
ACACCGGCAGCAGGCTGCACGAGGCCGGCTGCGCCCCAGCTACCACGTGGTCAGGGCGATATC
GCCACCGCCACCGAGGGCGTGATCATCAACCGTGCCAACAGCAAGGGCCAGCCCGAGGGCGGAGT
GTGCGGCGCCCTGTACAAGAAGTTCCCCGAGAGCTTCGACCTGCAGCCCATCGAGGTGGGCAAGGC
CAGGCTGGTGAAGGGCGCCGCTAAGCACATCATCCACGCCGTGGGCCCCAACTTCAACAAGGTGAG
CGAGGTGGAAGGGCACAAGCAGCTGGCCGAAGCCTACGAGAGCATCGCCAAGATCGGTGAACGACA
ATAACTACAAGAGCGTGCCATCCACTGCTCAGCCACCGCATCTTCAGCGGCAACAAGGACAGGC
TGACCCAGAGCCTGAACCACCTGCTCACCGCCCTGGACACCACCGATGCCGACGTGCCATCTACTG
CAGGGACAAGAAGTGGGAGATGACCCTGAAGGAGGCCGTGGCCAGGCGGGAGGCCGTGGAAGAG
ATCTGCATCAGCGACGACTCCAGCGTGACCGAGCCCAGCGCCGAGCTGGTGAGGGTGCACCCCAAG
AGCTCCCTGGCCGGCAGGAAGGGCTACAGCACCAGCGACGGCAAGACCTTCAGCTACCTGGAGGG
CACCAAGTTCCACCAGGCCGCTAAGGACATCGCCGAGATCAACGCTATGTGGCCCCGTGGCCACCGA
GGCCAACGAGCAGGTGTGCATGTACATCCTGGGCGAGAGCATGTCCAGCATCAGGAGCAAGTGCCC
CGTGGAGGAAAGCGAGGCCAGCACACCACCCAGCACCCCTGCCCTGCCTGTGCATCCACGCTATGAC
ACCCGAGAGGGTGCAGCGGCTGAAGGCCAGCAGGCCCGAGCAGATCACCGTGGCAGTCTTTCCC
ACTGCCCAAGTACAGGATCACCGGCGTGCAGAAGATCCAGTGCAGCCAGCCCATCCTGTTACGCCCA
AAGGTGCCCCCTACATCCACCCCAGGAAGTACCTGGTGGAGACCCCACCCGTGGACGAGACACC
GAGCCAAGCGCCGAGAACCAGAGCACCAGGGCACACCCGAGCAGCCACCCCTGATCACCGAGGA
CGAGACAAGGACCCGGACCCCAGAGCCATCATTATCGAGGAAGAGGAAGGAGGACAGCATCAGCC
TGCTGACGACGGCCCCACCCACCAGGTGCTGCAGGTGGAGGCCGACATCCACGGCCCCACCCAGCG
TGTCAGCTCCAGCTGGAGCATCCACACGCCAGCGACTTCGACGTGGACAGCTGAGCATCCTGG
ACACCCTGGAGGGCGCCAGCGTGACCTCCGGCGCCACCAGCGCCGACCAACAGCTACTTCCGCA
AGAGCATGGAGTTCCTGGCCAGGGCCCGTGCCAGCTCCCAGGACCGTGTTCAGGAACCCACCCCACC
CAGCTCCCAGGACCAGGACCCCAAGCCTGGCTCCCAGCAGGGCCTGCAGCAGGACCAGCCTGGTGA
GCACCCCACCCGGCGTGAACAGGGTGTACACCAGGGAGGAACTGGAGGCCCTGACACCCAGCAGG
ACCCCCAGCAGGTCCGTGAGCAGGACTAGTCTGTGTCCAACCCACCCGGCGTGAACAGGGTGGATC
ACCAGGGAGGAATTCGAGGCCTTCGTGGCCCAGCAACAGAGACGGTTCGACGCCGGCGCCTACATC
TTCAGCAGCGACACCGGCCAGGGACACCTGCAGCAAAAAGAGCGTGAGGCAGACCGTGCTGAGCGA
GGTGGTGTGGAGAGGACCGAGCTGAAATCAGCTACGCCCCAGGCTGGACCAGGAGAAGGAG
GAACTGCTCAGGAAGAACTGCAGCTGAACCCACCCCAGCCAACAGGAGCAGGTACCAGAGCAG
GAAGGTGGAGAACATGAAGGCCATCACCGCCAGGCGGATCCTGCAGGGCCTGGGACACTACCTGA
AGGCCGAGGGCAAGGTGGAGTGCTACAGGACCCTGCACCCCGTGCCACTGTACAGCTCCAGCGTGA
ACAGGGCCTTCTCCAGCCCCAAGGTGGCCGTGGAGGCCTGCAACGCTATGCTGAAGGAGAACTTCC
CCACCGTGGCCAGCTACTGCATCATCCCCGAGTACGACGCCTACCTGGACATGGTGGACGGCGCCA
GCTGCTGCCTGGACACCGCCAGCTTCTGCCCCGCAAGCTGAGGAGCTTCCCCAAGAAACACAGCTA
CCTGGAGCCCACCATCAGGAGCGCCGTGCCAGCGCCATCCAGAACACCCTGCAGAACGTGCTGGC
CGCTGCCACCAAGAGGAACTGCAACGTGACCCAGATGAGGGAGCTGCCCGTGTGGACAGCGCTG
CCTTCAACGTGGAGTGCTTCAAGAAATACGCTGCAACAACGAGTACTGGGAGACCTTCAAGGAGA
ACCCCATCAGGCTGACCGAAGAGAACGTGGTGAACCTACATACCAAGCTGAAGGGCCCCAAGGCCG
CTGCCCTGTTTCGTAAGACCCACAACCTGAACATGCTGCAGGACATCCCAATGGACAGGTTTCGTGAT
GGACCTGAAGAGGGACGTGAAGGTGACACCCGGCACCAAGCACACCCGAGGAGGCCCAAGGTG
CAGGTGATCCAGGGCCGTGACCCACTGGCCACCGCCTACCTGTGCGGCATCCACAGGGAGCTGGTG
AGGCGGTGAACGCCGTGCTGCTGCCAACATCCACACCCTGTTTCGACATGAGCGCCGAGGACTTC
GACGCCATCATCGCCGAGCACTTCCAGCCCGGCGACTGCGTGCTGGAGACCGACATCGCCAGCTTC

GACAAGAGCGAGGATGACGCTATGGCCCTGACCGCTCTGATGATCCTGGAGGACCTGGGCGTGGA
CGCCGAGCTGCTCACCTGATCGAGGCTGCCTTCGGCGAGATCAGTCCATCCACCTGCCCACCAAG
ACCAAGTTCAAGTTCGGCGCTATGATGAAAAGCGGAATGTTCCCTGACCCTGTTCTGTAACACCGTGA
TCAACATTGTGATCGCCAGCAGGGTGTGCGGGAGAGGCTGACCGGCAGCCCCTGCGCTGCCTTCA
TCGGCGACGACAACATCGTGAAGGGCGTGAAAAGCGACAAGCTGATGGCCGACAGGTGCGCCACC
TGGCTGAACATGGAGGTGAAGATCATCGACGCCGTGGTGGGCGAGAAGGCCCCCTACTTCTGCGGC
GGATTCATCCTGTCGACAGCGTGACCGGCACCGCTGCAGGGTGGCCGACCCCCTGAAGAGGCTG
TTCAAGCTGGGCAAGCCACTGGCCGCTGACGATGAGCAGCAGATGACAGGCGGAGGGCCCTGCA
CGAGGAAAGCACCAGGTGGAACAGGGTGGGGCATCCTGAGCGAGCTGTGCAAGGCCGTGGAGAGC
AGGTACGAGACCGTGGGCACCAGCATCATCGTATGGCTATGACCACACTGGCCAGCTCCGTCAAG
AGCTTCTCTACCTGAGGGGCCCCCTATAACTCTCTACGGCTAACCTGAATGGACTACGACATAGT
CTAgccaccATGagcaagatctacatcgacgagcggagcaacgccgagatcgtgtgagggccatcaagaccatggcatcga
ggcgcccaccgcccagctgaccaggcagctgaacatggagaagcgggaggtgaacaaggccctgacgacctgacagaggag
cgctatggtgactccagcgacgacatccctccccggtggttcatgaccaccgaggccgacaagcccgacgcccagcgtatggcgg
acgtgatcatcgacgacgtgagcagggagaagtccatgaggaggaccacaagagcttcgacgacgtgatccccccaagaaga
tcatcgactggaagggcgcaaccccgtgacctgatcaacgagtactgccagatcaccaggaggactggagcttcggatcga
gagcgtgggccccagcaacagccccacttctacgctgctggacatcgacggcaggggttgcacaaggccgacggcaagagc
aagcgggacgccaagaacaacgcccgaagctggccgtggacaagctgctgggctacgtgatcatccggttcTAAactcgagcta
gtgactgactaggatctggttaccactaaaccagcctcaagaacaccggaatggagtctctaagctacataataccaacttacactt
acaaaatggttcccccaaatgtagccattcgtatctgctcctaataaaaagaaagtttctcacattctagAGCTCCGTC AAG
AGCTTCTCTACCTGAGGGGCCCCCTATAACTCTCTACGGCTAACCTGAATGGACTACGACATAGT
CTAGCCACCATGGAAGATGCCAAAAACATTAAGAAGGGCCCAGCGCCATTCTACCCACTCGAAGAC
GGGACCGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGTGGCCGGCACCATCGCC
TTTACCGACGCACATATCGAGGTGGACATTACCTACGCCGAGTACTTCGAGATGAGCGTTTCGGCTGG
CAGAAGCTATGAAGCGCTATGGGCTGAATACAAACCATCGGATCGTGGTGTGCAGCGAGAATAGCT
TGCAGTTCTTCATGCCCCTGTTGGGTGCCCTGTTTCATCGGTGTGGCTGTGGCCCCAGCTAACGACAT
CTACAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCCACCGTCGTATTTCGTGAGCAA
GAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTACCGATCATACAAAAGATCATCATCAT
GGATAGCAAGACCGACTACCAGGGCTTCCAAAGCATGTACACCTTCGTGACTTCCCATTTGCCACCC
GGCTTCAACGAGTACGACTTCGTGCCCCGAGAGCTTCGACCCGGACAAAACCATCGCCCTGATCATGA
ACAGTAGTGGCAGTACCGGATTGCCAAGGGCGTAGCCCTACCGCACCGCACCGCTTGTGTCCGATT
CAGTCATGCCCCGACCCCATCTTCGGCAACCAGATCATCCCCGACACCGCTATCCTCAGCGTGGTG
CCATTTACCACGGCTTCGGCATGTTACCACGCTGGGCTACTTGATCTGCGGCTTTCGGGTCTGTGCT
CATGTACCGCTTCGAGGAGGAGCTATTCTTGCGCAGCTTGCAAGACTATAAGATTCAATCTGCCCTG
CTGGTGGCCACACTATTTAGCTTCTTCGCTAAGAGCACTCTCATCGACAAGTACGACCTAAGCAACTT
GCACGAGATCGCCAGCGGGCGGGGCGCCGCTCAGCAAGGAGGTAGGTGAGGCGGTGGCCAAACGC
TTCCACCTACCAGGCATCCGACAGGGCTACGGCCTGACAGAAACAACCAGCGCCATTCTGATCACCC
CCGAAGGGGACGACAAGCCTGGCGCAGTAGGCAAGGTGGTGGCCCTTCTTCGAGGCTAAGGTGGTG
GACTTGGACACCGGTAAGACACTGGGTGTGAACCAGCGCGGCGAGCTGTGCGTCCGTGGCCCCATG
ATCATGAGCGGCTACGTTAAACAACCCCGAGGCTACAAACGCTCTATCGACGACAAGGACGGCTGGCTG
CACAGCGGGCGACATCGCCTACTGGGACGAGGACGAGCACTTCTTCATCGTGGACCGGCTGAAGTCC
CTGATCAAATACAAGGGCTACCAGGTAGCCCCAGCCGAAGTGGAGAGCATCCTGCTGCAACACCCC
AACATCTTCGACGCCGGGGTCGCCGGCCTGCCCCGACGACGATGCCGGCGAGCTGCCCCGCCGACGTC
GTCGTGCTGGAACACGGTAAAACCATGACCGAGAAGGAGATCGTGGACTATGTGGCCAGCCAGGTT

ACAACCGCCAAGAAGCTGCGCGGTGGTGTTCGTGGACGAGGTGCCTAAAGGACTGACCGG
CAAGTTGGACGCCCGCAAGATCCGCGAGATTCTCATTAAGGCCAAGAAGGGCGGCAAGATCGCCGT
GTA ACTCGAGTATGTTACGTGCAAAGGTGATTGTCACCCCCGAAAGACCATATTGTGACACACCCT
CAGTATCACGCCCAAACATTTACAGCCGCGGTGTCAAAAACCGCGTGGACGTGGTTAACATCCCTGC
TGGGAGGATCAGCCGTAATTATTATAATTGGCTGGTGCTGGCTATTOATTGTGGCCATGTACGTGCT
GACCAACCAGAAACATAATTGAATACAGCAGCAATTGGCAAGCTGCTTACATAGA ACTCGCGGCGA
TTGGCATGCCGCTTAAAATTTTTATTTTATTTTTCTTTCTTTTCCGAATCGGATTTTGTTTTAAATAT
TTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATctagAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

(Sequenza di RNA per STARR Fluc IRES-E3L) SEQ ID NO: 128

AUGGGCGCGCAUGAGAGAAGCCCAGACCAAUUACCUACCCAAAAUGGAGAAA
GUUCACGUUGACAUCGAGGAAGACAGCCAUUCCUCAGAGCUUUGCAGCGGAG
CUUCCCGCAGUUUGAGGUAGAAGCCAAGCAGGUCACUGAUA AUGACCAUGCUA
AUGCCAGAGCGUUUUCGCAUCUGGCUUCAAAACUGAUCGAAACGGAGGUGGA
CCAUCCGACACGAUCCUUGACAUUGGAAGUGCGCCCGCCGCAGAAUGUAUU
CUAAGCACAAGUAUCAUUGUAUCUGUCCGAUGAGAUGUGCGGAAGAUCCCGGA
CAGAUUGUAUAAGUAUGCAACUAAGCUGAAGAAAACUGUAAGGAAUAACU
GAUAAGGAAUUGGACAAGAAAUGAAGGAGCUGGCCCGCCGUC AUGAGCGACC
CUGACCUGGAAACUGAGACUAUGUGCCUCCACGACGACGAGUCGUGUCGUAC
GAAGGGCAAGUCCGUGUUUACCAGGAUGUAUACGCCGUCGACGGCCCCACCAG
CCUGUACCACCAGGGCCAACAAGGGCGUGAGGGUGGCCUACUGGAUCGGCUUCG
ACACCACACCCUUAUGUUAAGAACCUGGCCGGCGCCUACCCAGCUACAGC
ACCAACUGGGCCGACGAGACCGUGCUGACCGCCAGGAACAUCGGCCUGUGCAG
CAGCGACGUGAUGGAGAGGCCGGAGAGGCAUGAGCAUCCUGAGGAAGAAA
UACCUGAAGCCCAGCAACAACGUGCUGUUCAGCGUGGGCAGCACCAUCUACCA
CGAGAAGAGGGACCUGCUCAGGAGCUGGCACCUGCCCAGCGUGUUCACCCUGA
GGGGCAAGCAGAACUACACCGCAGGUGCGAGACCAUCGUGAGCUGCGACGGC
UACGUGGUGAAGAGGAUCGCCAUCAGCCCCGGCCUGUACGGCAAGCCCAGCGG
CUACGCCGUACAAUGCACAGGGGAGGGCUUCCUGUGCUGCAAGGUGACCGACA
CCUGAACGGCGAGAGGGUGAGCUUCCCCGUGUGCACCUACGUGCCC GCCACC
CUGUGCAGACCAGAUGACCGCAUCCUGGCCACCGACGUGAGCGCCGACGACGC
CCAGAAGCUGCUCGUGGGCCUGAACAGAGGAUCGUGGUCAACGGCAGGACCC
AGAGGAACACCAACACAAUGAAGAACUACCGCUGCCCCGUGGUGGCCAGGCU
UUCGCCAGGUGGGCCAAGGAGUACAAGGAGGACCAGGAAGACGAGAGGCCCU
GGGCCUGAGGGACAGGCAGCUGGUGAUGGGCUGCUGCUGGGCCUUCAGGCGG
ACAAGAUACACAGCAUCUACAAAGAGGCCCGACCCAGACCAUCAUAAGGUG
AACAGCGACUCCACAGCUUCGUGCUGCCAGGAUCGGCAGCAACACCCUGGA
GAUGGGCCUGAGGACCCGGAUCAGGAAGAUGCUGGAGGAACACAAGGAGCCCA
GCCACUGAUCACCGCCGAGGACGUGCAGGAGGCCAAGUGCGCUGCCGACGAG
GCCAAGGAGGUGAGGGAGGCCGAGGAACUGAGGGCCGCCUGCCACCCUGGC
UGCCGACGUGGAGGAACCCACCCUGGAAGCCGACGUGGACCUGAUGCUGCAGG
AGGCCGGCGCCGGAAGCGUGGAGACACCCAGGGGCCUGAUAAGGGUGACCAGC
UACGACGGCGAGGACAAGAUCGGCAGCUACGCCGUGCUGAGCCCACAGGCCGU
GCUGAAGUCCGAGAAGCUGAGCUGCAUCCACCCACUGGCCGAGCAGGUGAUCG

UGAUCACCCACAGCGGCAGGAAGGGCAGGUACGCCGUGGAGCCCUACCACGGC
AAGGUGGUCGUGCCCAGGGCCACGCCAUCCCCUGGCAGGACUUCAGGCCCU
GAGCGAGAGCGCCACCAUCGUGUACAACGAGAGGGAGUUCGUGAACAGGUACC
UGCACCAUAUCGCCACCCACGGCGGAGCCCUGAACACCGACGAGGAAUACUAC
AAGACCGUGAAGCCCAGCGAGCACGACGGCGAGUACCGUACGACAUCGACAG
GAAGCAGUGCGUGAAGAAAGAGCUGGUGACCGGCCUGGGACUGACCGGCGAG
CUGGUGGACCCACCCUUCACGAGUUCGCCUACGAGAGCCUGAGGACCAGACC
CGCCGCUCCCUACCAGGUGCCCACCAUCGGCGUGUACGGCGUGCCCGGCAGCG
GAAAGAGCGGCAUCAUCAAGAGCGCCGUGACCAAGAAAGACCUGGUGGUCAGC
GCCAAGAAAGAGAACUGCGCCGAGAUAUCAGGGACGUGAAGAAGAUGAAAG
GCCUGGACGUGAACGCGCGCACCGUGGACAGCGUGCUGCUGAACGGCUGCAAG
CACCCCGUGGAGACCCUGUACAUCGACGAGGCCUUCGCUUGCCACGCCGGCAC
CCUGAGGGCCCUGAUCGCCAUCAUCAGGCCAAGAAAGCCGUGCUGUGCGGGC
ACCCCAAGCAGUGCGGCUUCUUAACAUGAUGUGCCUGAAGGUGCACUUAAC
CACGAGAUGCACCCAGGUGUCCACAAGAGCAUCAGCAGGCGGUGCACCAA
GAGCGUGACCAGCGUCGUGAGCACCCUGUUCUACGACAAGAAAUGAGGACCA
CCAACCCCAAGGAGACCAAAAUCGUGAUCGACACCACAGGCAGCACCAAGCCC
AAGCAGGACGACCCUGAUCCUGACCUGCUUCAGGGCUGGGUGAAGCAGCUGCA
GAUCGACUACAAGGGCAACGAGAUAUGACCGCCGUGCGAGCCAGGGCCUGA
CCAGGAAGGGCGUGUACGCCGUGAGGUACAAGGUGAACGAGAACCCACUGUAC
GCUCCCACCAGCGAGCACGUGAACGUGCUGCUGACCAGGACCGAGGACAGAU
CGUGUGGAAGACCCUGGCCGGCGACCCCUUGGAUCAAGACCCUGACCGCCAAGU
ACCCCGGCAACUUCACCGCCACCAUCGAAGAGUGGCAGGCCGAGCACGACGCC
AUCAUGAGGCACAUCUCCUGGAGAGGCCCGACCCACCGACGUGUUCAGAACAA
GGCCAACGUGUGCUGGGCCAAGGCCCUUGGUGCCCGUGCUGAAGACCGCCGGCA
UCGACAUGACCACAGAGCAGUGGAACACCGUGGACUACUUCGAGACCGACAAG
GCCACAGCGCCGAGAUCGUGCUGAACAGCUGUGCGUGAGGUUCUUCGGCCU
GGACCUUGGACAGCGGCCCUUUCAGCGCCCCACCGUGCCACUGAGCAUCAGGA
ACAACCACUGGGACAACAGCCCCAGCCAAACAUGUACGGCCUGAACAAAGGAG
GUGGUCAGGCAGCGUGAGCAGGCGGUACCCACAGCUGCCCCAGGGCCGUGGCCAC
CGGCAGGGUGUACGACAUGAACACCGGCACCCUGAGGAACUACGACCCACAGGA
UCAACCUUGGUGCCCUGAACAGGCGGCUGCCCCACGCCCUUGGUGCUGCACCAC
AACGAGCACCCACAGAGCGACUUCAGCUCCUUCGUGAGCAAGCUGAAAGGCAG
GACCGUGCUGGUCGUGGGCGAGAAGCUGAGCGUGCCCGGCAAGAUGGUGGAC
UGGCUGAGCGACAGGCCCGAGGCCACCUUCGGGCCAGGCUGCCACCUUGGCAU
CCCCGGCGACGUGCCCAAGUACGACAUAUCUUCGUGAACGUCAGGACCCCAU
ACAAGUACCACCAUUAACCAGCAGUGCGAGGACCACGCCAUAAGCUGAGCAUG
CUGACCAAGAAGGCCUGCCUGCACCUGAACCCCGGAGGCACCUGCGUGAGCAU
CGGCUACGGCUACGCCGACAGGGCCAGCGAGCAUCAUUGGCGCCAUCGCCA
GGCUGUUAAGUUCAGCAGGGUGUGCAAACCCAAGAGCAGCCUGGAGGAAACC
GAGGUGCUGUUCGUGUUAUCGGCUACGACCGGAAGGCCAGGACCCACAACCC
CUACAAGCUGAGCAGCACCCUGACAAACUACACCGGCAGCAGGCUGCACG
AGGCCGGCUGCGCCCCAGCUACCACGUGGUCAGGGGCGAUUUCGCCACCGCC
ACCGAGGGCGUGAUAUCAACGUGCCAACAGCAAGGGCCAGCCCGGAGGGCG

AGUGUGCGGCCCCUGUACAAGAAGUUCCCCGAGAGCUUCGACCUGCAGCCCA
UCGAGGUGGGCAAGCCAGGCUGGUGAAGGGCGCCGCUAAGCACAUCAUCCAC
GCCGUGGGCCCCAACUUAACAAGGUGAGCGAGGUGGAAGGCGACAAGCAGCU
GGCCGAAGCCUACGAGAGCAUCGCCAAGAUCGUGAACGACAAUAACUACAAGA
GCGUGGCCAUCCCACUGCUCAGCACCGGCAUCUUCAGCGGCAACAAGGACAGG
CUGACCCAGAGCCUGAACCACCUGCUCACCGCCCCUGGACACCACCGAUGCCGA
CGUGGCCAUUCACUGCAGGGACAAGAAGUGGGGAAUGACCCUGAAGGAGGCC
GUGGCCAGGCGGGAGGCCGUGGAAGAGAUCUGCAUCAGCGACGACUCCAGCGU
GACCGAGCCCAGCGCCGAGCUGGUGAGGGUGCACCCCAAGAGCUCCUGGCCG
GCAGGAAGGGCUACAGCACCGAGCGACGGCAAGACCUUCAGCUACCCUGGAGGGC
ACCAAGUUCACCAGGCCGCUAAGGACAUCGCCGAGAUCAACGCUAUGUGGCC
CGUGGCCACCGAGGCCAACGAGCAGGUGCAUGUACAUCUCCUGGGCGAGAGCA
UGUCCAGCAUCAGGAGCAAGUGCCCCGUGGAGGAAAGCGAGGCCAGCACACCA
CCCAGCACCCUGCCCUGCCUGUGCAUCCACGCUAUGACACCCGAGAGGGUGCA
GCGGCUGAAGGCCAGCAGGCCCGAGCAGAUACCGUGUGCAGCUCCUCCCAC
UGCCCAAGUACAGGAUACCCGGCGUGCAGAAGAUCAGUGCAGCCAGCCAU
CUGUUCAGCCAAAGGUGCCCGCCUACAUCCACCCAGGAAGUACCGUGGGA
GACCCACCCGUGGACGAGACACCCGAGCCAAGCGCCGAGAACCAGCACCG
AGGGCACACCCGAGCAGCCACCCUGAUCACCGAGGACGAGACAAGGACCCGG
ACCCAGAGCCAUCAUUAUCGAGGAAGAGGAAGAGGACAGCAUCAGCCUGCU
GAGCGACGGCCCCACCCACCAGGUGCUGCAGGUGGAGGCCGACAUCCACGGCC
CACCCAGCGUGUCCAGCUCCAGCUGGAGCAUCCACACGCCAGCGACUUCGAC
GUGGACAGCCUGAGCAUCCUGGACACCCUGGAGGGCGCCAGCGUGACCUCCGG
CGCCACCAGCGCCGAGACCAACAGCUACUUCGCCAAGAGCAUGGAGUUCUGG
CCAGGCCCUGCCAGCUCCAGGACCGUGUUCAGGAACCCACCCACCCAGCUC
CCAGGACCAGGACCCCAAGCCUGGCUCCAGCAGGGCCUGCAGCAGGACCAGC
CUGGUGAGCACCCACCCGGCGUGAACAGGGUGAUCACCAGGGAGGAACUGGA
GGCCCUGACACCCAGCAGGACCCCCAGCAGGUCCGUGAGCAGGACUAGUCUGG
UGUCCAACCCACCCGGCGUGAACAGGGUGAUCACCAGGGAGGAAUUCGAGGCC
UUCGUGGCCAGCAACAGAGACGGUUCGACGCCGGCGCCUACAUCUUCAGCAG
CGACACCGGCCAGGGACACCUGCAGCAAAAGAGCGUGGAGGCACCGUGCUGA
GCGAGGUGGUGCUGGAGAGGACCGAGCUGGAAAUCAGCUACGCCCCAGGCUG
GACCAGGAGAAGGAGGAACUGCUCAGGAAGAAACUGCAGCUGAACCCACCC
AGCCAACAGGAGCAGGUACCAGAGCAGGAAGGUGGAGAAUCAUGAAGGCCAUC
ACCGCCAGGCGGAUCCUGCAGGGCCUGGGACACUACCUGAAGGCCGAGGGCAA
GGUGGAGUGCUACAGGACCCUGCACCCCGUGCCACUGUACAGCUCCAGCGUGA
ACAGGGCCUUCUCCAGCCCCAAGGUGGCCGUGGAGGCCUGCAACGCUAUGCUG
AAGGAGAACUCCCCACCGUGGCCAGCUACUGCAUCAUCCCCGAGUACGACGC
CUACCUGGACAUGGUGGACGGCGCCAGCUGCUGCCUGGACACCGCCAGCUUCU
GCCCCGCAAGCUGAGGAGCUUCCCCAAGAAACACAGCUACCGGAGCCCACC
AUCAGGAGCGCCGUGCCCAGCGCCAUCCAGAACACCCUGCAGAACGUGCUGGC
CGCUGCCACCAAGAGGAACUGCAACGUGACCCAGAUGAGGGAGCUGCCCUGC
UGGACAGCGCUGCCUUAACGUGGAGUGCUUCAAGAAAUCGCCUGCAACAAC
GAGUACUGGGAGACCUUCAAGGAGAACCCCAUCAGGCUGACCGAAGAGAACGU

GGUGAACUACAUCACCAAGCUGAAGGGCCCCAAGGCCGUGCCCUGUUCGCUA
AGACCCACAACCUGAACAUGCUGCAGGACAUCCAAUGGACAGGUUCGUGAUG
GACCUGAAGAGGGACGUGAAGGUGACACCCGGCACCAAGCACACCGAGGAGAG
GCCAAGGUGCAGGUGAUCCAGGCCGUGACCCACUGGCCACCGCCUACCUGU
GCGGAUCCACAGGGAGCUGGUGAGGGCGGUGAACGCCGUGCUGCUGCCCAAC
AUCCACACCCUGUUCGACAUGAGCGCCGAGGACUUCGACGCCAUCAUCGCCGA
GCACUCCAGCCCGGCACUGCGUGCUGGAGACCGACAUCGCCAGCUUCGACA
AGAGCGAGGAUGACGCUAUGGCCUGACCGCUCUGAUGAUCCUGGAGGACCUG
GGCGUGGACCGCGAGCUGCUCACCCUGAUCGAGGCUGCCUUCGGCGAGAUCAG
CUCCAUCCACCUGCCACCAAGACCAAGUUCAAGUUCGGCGCUAUGAUGAAAA
GCGGAAUGUUCUGACCCUGUUCGUGAACACCGUGAUAACAUGUGAUCGCC
AGCAGGGUGCUGCGGGAGAGGCUGACCGGCAGCCCCUGCGCUGCCUUCAUCGG
CGACGACAACAUCGUGAAGGGCGUGAAAAGCGACAAGCUGAUGGCCGACAGG
UGCGCCACCUGGCUGAACAUUGGAGGUGAAGAUCAUCGACGCCGUGGUGGGCGA
GAAGGCCCCUACUUCUGCGGGCGGAUUAUCCUGUGCGGACAGCGUGACCGGCA
CCGCCUGCAGGGUGGCCGACCCCCUGAAGAGGCUGUUAAGCUGGGCAAGCCA
CUGGCCGUGACGAUGAGCACGACGAUGACAGGCGGAGGGCCUGCACGAGGA
AAGCACAGGUGGAACAGGGUGGGCAUCCUGAGCGAGCUGUGCAAGGCCGUG
GAGAGCAGGUACGAGACCGUGGGCACCAGCAUCAUCGUGAUGGCUAUGACCAC
ACUGGCCAGCUCCGUCAAGAGCUUCUCCUACCUAGGGGGCCCCUAUAACUC
UCUACGGCUAACCUGAAUGGACUACGACAUAUCUAGUCCGCCAAGGCCGCCA
CCAUGGAAGAUGCCAAAACAUAAGAAGGGCCAGCGCAUUCUACCCACUC
GAAGACGGGACCGCCGGCAGCAGCUGCACAAAGCCAUGAAGCGCUACGCCU
GGUGCCCGGCACCAUCGCCUUUACCGACGCACAUAUCGAGGUGGACAUAACCU
ACGCCGAGUACUUCGAGAUGAGCGUUCGGCUGGCAGAAGCUAUGAAGCGCUA
UGGGCUGAAUACAAACCAUCGGAUCGUGGUGUGCAGCGAGAAUAGCUUGCAG
UUCUUAUGCCCUGUUGGGUGCCCUGUUAUCGGUGUGGCUGUGGCCCCAGC
UAACGACAUCUACAACGAGCGGAGCUGCUGAACAGCAUGGGCAUCAGCCAGC
CCACCGUCGUAUUCGUGAGCAAGAAAGGGCUGCAAAAGAUCCUCAACGUGCAA
AAGAAGCUACCGAUCAUACAAAAGAUCAUCAUAUGGAUAGCAAGACCGACU
ACCAGGGCUUCCAAAGCAUGUACACCUUCGUGACUCCCAUUGCCACCCGGC
UUAACGAGUACGACUUCGUGCCCGAGAGCUUCGACCGGGACAAAACCAUCGC
CCUGAUGAACAGUAGUGGCAGUACCGGAUUGCCCAAGGGCGUAGCCCUAC
CGCACCGCACCGCUUGUGUCCGAUUCAGUCAUGCCCGCGACCCCAUCUUCGGC
AACCAGAUCAUCCCCGACACCGCUAUCCUCAGCGUGGUGCCAUUCACCACGG
CUUCGGCAUGUUCACCACGUGGGCUACUUGAUCUGCGGCUUUCGGGUCGUGC
UCAUGUACCGCUUCGAGGAGGAGCUAUUCUUGCGCAGCUUGCAAGACUAUAA
GAUUCAAUCUGCCCUGCUGGUGCCCACACUAUUAAGCUUCUUCGCUAAGAGCA
CUCUCAUCGACAAGUACGACCUAAGCAACUUGCACGAGAUCCAGCGGGCGGG
GCGCCGUCAGCAAGGAGGUAGGUGAGGGCCGUGGCCAAACGCUUCCACCUACC
AGGCAUCCGACAGGGCUACGGCCUGACAGAAACAACCAGCGCAUUCUGAUCA
CCCCGAAGGGGACGACAAGCCUGGCGCAGUAGGCAAGGUGGGUGCCCUUCUUC
GAGGCUAAGGUGGUGGACUUGGACACCGGUAAGACACUGGGUGUGAACCAGC
GCGGCGAGCUGUGCGUCCGUGGCCCAUGAUAUGAGCGGCUACGUUAACAAC

corto)) SEQ ID NO: 129 AUGGGCGGCGCAUGAGAGAAGCCCAGACCAAUACCACCCAAAAUGGAGAAA
GUUCACGUUGACAUCGAGGAAGACAGCCCAUCCUCAGAGCUUUGCAGCGGAG
CUUCCCGCAGUUUGAGGUAGAAGCCAAGCAGGUCACUGAUAUUGACCAUGC
AUGCCAGAGCGUUUUCGCAUCUGGCUUCAAACUGAUCGAAACGGAGGUGGA
CCCAUCCGACACGAUCCUUGACAUUGGAAGUGCGCCCGCCCGCAGAAUGUAUU
CUAAGCACAAGUAUCAUUGUAUCUGUCCGAUGAGAUUGUGCGGAAGAUCCCGGA
CAGAUUGUAUAAGUAUGCAACUAAGCUGAAGAAAAACUGUAAGGAAUAACU
GAUAAGGAAUUGGACAAGAAAAUGAAGGAGCUGGCCCGCCGUAUGAGCGACC
CUGACCUGGAAACUGAGACUAUGUGCCUCCACGACGACGAGUCGUGUCGUAC
GAAGGGCAAGUCCGCGUUUACCAGGAUGUAUACGCCGUCGACGGCCCCACCAG
CCUGUACCACCAGGGCCAACAAGGGCGUGAGGGUGGCCUACUGGAUCGGCUUCG
ACACCACACCCUUCAGUUCAGAACCUGGCCGGCGCCUACCCAGCUACAGC
ACCAACUGGGCCGACGAGACCGUGCUGACCGCCAGGAACAUCGGCCUGUGCAG
CAGCGACGUGAUGGAGAGGCCGGAGAGGCAUGAGCAUCCUGAGGAAGAAA
UACCUGAAGCCCAGCAACAACGUGCUGUUCAGCGUGGGCAGCACCAUCUACCA
CGAGAAGAGGGACCUGCUCAGGAGCUGGCACCUGCCCAGCGUGUUCACCCUGA
GGGGCAAGCAGAACUACACCUGCAGGUGCGAGACCAUCGUGAGCUGCGACGGC
UACGUGGUGAAGAGGAUCGCCAUCAGCCCCGGCCUGUACGGCAAGCCCAGCGG
CUACGCCGCUACAAUGCACAGGGGAGGGCUUCCUGUGCUGCAAGGUGACCGACA
CCUGAACGGCGAGAGGGUGAGCUUCCCCGUGUGCACCUACGUGCCCCGCCACC
CUGUGCGACCAGAUGACCGGAUCCUGGCCACCGACGUGAGCGCCGACGACGC
CCAGAAGCUGCUCGUGGGCCUGAACAGAGGAUCGUGGUAACGGCAGGACCC
AGAGGAACACCAACACAAUGAAGAACUACCUGCUGCCCGUGGUGGCCAGGCU
UUCGCCAGGUGGGCCAAGGAGUACAAGGAGGACCAGGAAGACGAGAGGCCCU
GGGCCUGAGGGACAGGCAGCUGGUGAUGGGCUGCUGCUGGGCCUUCAGGCGG
ACAAGAUACCCAGCAUCUACAAAGAGGCCCGACACCCAGACCAUCAUAAGGUG
AACAGCGACUCCACAGCUUCGUGCUGCCCAGGAUCGGCAGCAACACCCUGGA
GAUGGGCCUGAGGACCCGGAUCAGGAAGAUGCUGGAGGAACACAAGGAGCCCA
GCCACUGAUCACCGCCGAGGACGUGCAGGAGGCCAAGUGCGCUGCCGACGAG
GCCAAGGAGGUGAGGGAGGCCGAGGAACUGAGGGCCCGCCUGCCACCCUUGC
UGCCGACGUGGAGGAACCCACCCUGGAAGCCGACGUGGACCUGAUGCUGCAGG
AGGCCGGCGCCGGAAGCGUGGAGACACCCAGGGGCCUGAUAAGGGUGACCAG
UACGACGGCGAGGACAAGAUCGGCAGCUACGCCGUGCUGAGCCCACAGGCCGU
GCUGAAGUCCGAGAAGCUGAGCUGCAUCCACCCACUGGCCGAGCAGGUGAUCG
UGAUCACCCACAGCGGCAGGAAGGGCAGGUACGCCGUGGAGCCCUACCACGGC
AAGGUGGUCGUGCCCAGGGCCACGCCAUCCCCGUGCAGGACUUCAGGCCCU
GAGCGAGAGCGCCACCAUCGUGUACAACGAGAGGGAGUUCGUGAACAGGUACC
UGCACCAUAUCGCCACCCACGGCGGAGCCUGAACACCCGACGAGGAAUACUAC
AAGACCGUGAAGCCCAGCGAGCACGACGGCGAGUACCGUACGACAUCGACAG
GAAGCAGUGCGUGAAGAAAGAGCUGGUGACCGGCCUGGGACUGACCGGCGAG
CUGGUGGACCCACCCUUCACGAGUUCGCCUACGAGAGCCUGAGGACCAGACC
CGCCGCUCCUACCAGGUGCCCACCAUCGGCGUGUACGGCGUGCCCCGGCAGCG
GAAAGAGCGGCAUCAUAAGAGCGCCGUGACCAAGAAAGACCUGGUGGUCAGC
GCCAAGAAAGAGAACUGCGCCGAGAUAUCAGGGACGUGAAGAAGAUAAAAG

GCCUGGACGUGAACGCGCGCACCGUGGACAGCGUGCUGCUGAACGGCUGCAAG
CACCCCGUGGAGACCCUGUACAUCGACGAGGCCUUCGCUUGCCACGCCGGCAC
CCUGAGGGCCCGUAUCGCCAUCAUCAGGCCCAAGAAAGCCGUGCUGUGCGGCG
ACCCCAAGCAGUGCGGCUUCUUAACAUGAUGUGCCUGAAGGUGCACUUAAC
CACGAGAUGCACCCAGGUGUCCACAAGAGCAUCAGCAGGCGGUGCACCAA
GAGCGUGACCAGCGUCGUGAGCACCCUGUUCUACGACAAGAAAUGAGGACCA
CCAACCCCAAGGAGACCAAAAUCGUGAUCGACACCACAGGCAGCACCAAGCCC
AAGCAGGACGACCCUGAUCCUGACCUGCUUCAGGGCUGGGUGAAGCAGCUGCA
GAUCGACUACAAGGGCAACGAGAUCAUGACCGCCGUGCGAGCCAGGGCCUGA
CCAGGAAGGGCGUGUACGCCGUGAGGUACAAGGUGAACGAGAACCCACUGUAC
GCUCCACCAGCGAGCACGUGAACGUGCUGCUGACCAGGACCGAGGACAGAU
CGUGUGGAAGACCCUGGCCGCGACCCCGGAUCAAGACCCUGACCGCCAAGU
ACCCCGGCAACUUCACCGCCACCAUCGAAGAGUGGCAGGCCGAGCACGACGCC
AUCAUGAGGCACAUCUGGAGAGGCCCGACCCACCGACGUGUUCAGAAACA
GGCCAACGUGUGCUGGGCCAAGGCCUGGUGCCCGUGCUGAAGACCGCCGGCA
UCGACAUGACCACAGAGCAGUGGAACACCGUGGACUACUUCGAGACCGACAAG
GCCACAGCGCCGAGAUUCGUGCUGAACAGCUGUGCGUGAGGUUCUUCGGCCU
GGACCUGGACAGCGGCCCGUUCAGCGCCCCACCGUGCCACUGAGCAUCAGGA
ACAACCACUGGGACAACAGCCCCAGCCAAACAUGUACGGCCUGAACAAAGGAG
GUGGUCAGGCAGCGUGAGCAGGCGGUACCCACAGCUGCCCCAGGGCCGUGGCCAC
CGGCAGGGUGUACGACAUGAACACCGGCACCCUGAGGAACUACGACCCACAGGA
UCAACCGUGGUGCCCGUGAACAGGCGGCUGCCCCACGCCUGGUGCUGCACCAC
AACGAGCACCCACAGAGCGACUUCAGCUCCUUCGUGAGCAAGCUGAAAGGCAG
GACCGUGCUGGUCGUGGGCGAGAAGCUGAGCGUGCCCGCAAGAUGGUGGAC
UGGCUGAGCGACAGGCCCGAGGCCACCUUCCGGGCCAGGCUGCCACCUCGGCAU
CCCCGGCGACGUGCCCAAGUACGACAUCUUCGUGAACGUCAGGACCCCAU
ACAAGUACCACCAUUAACAGCAGUGCGAGGACCACGCCAUCAAGCUGAGCAUG
CUGACCAAGAAGGCCUGCCUGCACCUGAACCCCGGAGGCACCGUGCUGAGCAU
CGGCUACGGCUACGCCGACAGGGCCAGCGAGCAUCAUUGGCGCAUCGCCA
GGCUGUUAAGUUCAGCAGGGUGUGCAAACCCAAGAGCAGCCUGGAGGAAACC
GAGGUGCUGUUCGUGUUAUCGGCUACGACCGGAAGGCCAGGACCCACAACCC
CUACAAGCUGAGCAGCACCCUGACAAACAUCUACACCGGCAGCAGGCUGCACG
AGGCCGGCUGCGCCCCAGCUACCACGUGGUCAGGGGCGAUUUCGCCACCGCC
ACCGAGGGCGUGAUCAUACGCUGCCAACAGCAAGGGCCAGCCGGAGGCGG
AGUGUGCGGCCCCUGUACAAGAAGUUCGCCGAGAGCUUCGACCUGCAGCCA
UCGAGGUGGGCAAGCCAGGCUGGUGAAGGGCGCCGUAAGCACAUCAUCCAC
GCCGUGGGCCCAACUUAACAAGGUGAGCGAGGUGGAAGGCGACAAGCAGCU
GGCCGAAGCCUACGAGAGCAUCGCCAAGAUCGUGAACGACAAUAACUACAAGA
GCGUGGCCAUCCACUGCUCAGCACCGGCAUCUUCAGCGCAACAAGGACAGG
CUGACCCAGAGCCUGAACACCUGCUCACCGCCCCUGGACACCACCGAUGCCGA
CGUGGCCAUUCUACUGCAGGGACAAGAAGUGGGGAAUGACCCUGAAGGAGGCC
GUGGCCAGGCGGGAGGCCGUGGAAGAGAUCUGCAUCAGCGACGACUCCAGCGU
GACCGAGCCCGACGCCGAGCUGGUGAGGGUGCACCCCAAGAGCUCCUGGCCG
GCAGGAAGGGCUACAGCACCGACGGCAAGACCUUCAGCUACCCUGGAGGGC

ACCAAGUUCACCAGGCCGCUAAGGACAUCGCCGAGAUAACGCUAUGUGGCC
CGUGGCCACCGAGGCCAACGAGCAGGUGCAUGUACAUCCUGGGCGAGAGCA
UGUCCAGCAUCAGGAGCAAGUGCCCCGUGGAGGAAAGCGAGGCCAGCACACCA
CCCAGCACCCUGCCCUGCCUGUGCAUCCACGCUAUGACACCCGAGAGGGUGCA
GCGGCUGAAGGCCAGCAGGCCCGAGCAGAUACCGUGUGCAGCUCCUCCCAC
UGCCCAAGUACAGGAUACCCGGCGUGCAGAAGAUCCAGUGCAGCCAGCCAU
CUGUUCAGCCAAAGGUGCCCGCCUACAUCCACCCCAGGAAGUACCUUGGUGGA
GACCCACCCGUGGACGAGACACCCGAGCCAAGCGCCGAGAACCAGCACCG
AGGGCACACCCGAGCAGCCACCCUGAUACCCGAGGACGAGACAAGGACCCGG
ACCCAGAGCCAUCAUUAUCGAGGAAGAGGAAGAGGACAGCAUCAGCCUGCU
GAGCGACGGCCCCACCCACCAGGUGCUGCAGGUGGAGGCCGACAUCCACGGCC
CACCCAGCGUGUCCAGCUCCAGCUGGAGCAUCCACACGCCAGCGACUUCGAC
GUGGACAGCCUGAGCAUCCUGGACACCCUGGAGGGCGCCAGCGUGACCUCCGG
CGCCACCAGCGCCGAGACCAACAGCUACUUCGCCAAGAGCAUGGAGUUCUGG
CCAGGCCCGUGCCAGCUCCAGGACCGUGUUCAGGAACCCACCCACCCAGCUC
CCAGGACCAGGACCCCAAGCCUGGCUCCAGCAGGGCCUGCAGCAGGACCAGC
CUGGUGAGCACCCACCCGGCGUGAACAGGGUGAUCACCAGGGAGGAACUGGA
GGCCCUGACACCCAGCAGGACCCCAAGGUGGAGGAGGACUAGUCUGG
UGUCCAACCCACCCGGCGUGAACAGGGUGAUCACCAGGGAGGAUUCGAGGCC
UUCGUGGCCAGCAACAGAGACGGUUCGACGCCGGCGCCUACAUCUUCAGCAG
CGACACCGGCCAGGGACACCUGCAGCAAAAGAGCGUGGAGGCACCGUGCUGA
GCGAGGUGGUGCUGGAGAGGACCGAGCUGGAAAUCAGCUACGCCCCAGGCUG
GACCAGGAGAAGGAGGAACUGCUCAGGAAGAAACUGCAGCUGAACCCACCC
AGCCAACAGGAGCAGGUACCAGAGCAGGAAGGUGGAGAACAUGAAGGCCAUC
ACGCCAGGCGGAUCCUGCAGGGCCUGGGACACUACCUGAAGGCCGAGGGCAA
GGUGGAGUGCUACAGGACCCUGCACCCCGUGCCACUGUACAGCUCCAGCGUGA
ACAGGGCCUUCUCCAGCCCCAAGGUGGCCGUGGAGGCCUGCAACGCUAUGCUG
AAGGAGAACUCCCCACCGUGGCCAGCUACUGCAUCAUCCCCGAGUACGACGC
CUACCUGGACAUGGUGGACGGCGCCAGCUGCUGCCUGGACACCGCCAGCUUCU
GCCCCGCAAGCUGAGGAGCUUCCCCAAGAAACACAGCUACCUGGAGCCCACC
AUCAGGAGCGCCUGCCCAGCGCCAUCCAGAACACCCUGCAGAACGUGCUGGC
CGCUGCCACCAAGAGGAACUGCAACGUGACCCAGAUGAGGGAGCUGCCCCGUGC
UGGACAGCGCUGCCUUAACGUGGAGUGCUUCAAGAAAACGCCUGCAACAAC
GAGUACUGGGAGACCUUCAAGGAGAACCCCAUCAGGCUGACCGAAGAGAACGU
GGUGAACUACAUCACCAAGCUGAAGGGCCCCAAGGCCGUGCCCUGUUCGCUA
AGACCCACAACCUGAACAUGCUGCAGGACAUCCCAAUGGACAGGUUCGUGAUG
GACCUGAAGAGGGACGUGAAGGUGACACCCGGCACCAAGCACACCGAGGAGAG
GCCAAGGUGCAGGUGAUCCAGGCGCUGACCCACUGGCCACCGCCUACCUGU
GCGGAUCCACAGGGAGCUGGUGAGGGCGGCUGAACGCCGUGCUGCUGCCCAAC
AUCCACACCCUGUUCGACAUGAGCGCCGAGGACUUCGACGCCAUCAUCGCCGA
GCACUUCAGCCGGCGACUGCGUGCUGGAGACCGACAUCGCCAGCUUCGACA
AGAGCGAGGAUGACGCUAUGGCCUGACCGCUCUGAUGAUCCUGGAGGACCUG
GGCGUGGACGCCGAGCUGCUCACCCUGAUCGAGGCUGCCUUCGGCGAGAUCAG
CUCCAUCCACCUGCCACCAAGACCAAGUUCAAGUUCGGCGCUAUGAUGAAAA

GCGGAAUGUCCUGACCCUGUUCGUGAACACCGUGAUCAACAUUGUGAUCGCC
AGCAGGGUGUCGCGGGAGAGGCUGACCGGCAGCCCCUGCGCUGCCUUCAUCGG
CGACGACAACAUUCGUGAAGGGCGUGAAAAGCGACAAGCUGAUGGCCGACAGG
UUGCACCACCGGUGAACAUGGAGGUGAAGAUAUCGACGCCGUGGUGGGCGA
GAAGGCCCCUACUUCUGCGGCGGAUUCAUCCUGUGCGGACAGCGUGACCGGCA
CCGCCUGCAGGGUGGCCGACCCCCUGAAGAGGCUGUUCAAGCUGGGCAAGCCA
CUGGCCCGUGACGAUGAGCACGACGAUGACAGGCGGAGGGCCUGCACGAGGA
AAGCACCAGGUGGAACAGGGUGGGCAUCCUGAGCGAGCUGUGCAAGGCCGUG
GAGAGCAGGUACGAGACCGUGGGCACCAGCAUCAUCGUGAUGGCUAUGACCAC
ACUGGCCAGCUCGUAAGAGCUUCUCCUACCUGAGGGGCCCCUAUAACUC
UCUACGGCUAACCUGAAUGGACUACGACAUAGUCUAGUCCGCCAAGGCCGCCA
CCAUGGAAGAUGCCAAAAACAUAAGAAGGGCCAGCGCCAUUCUACCCACUC
GAAGACGGGACCGCCGCGAGCAGCUGCACAAAGCCAUGAAGCGCUACGCCCU
GGUGCCCGGCACCAUCGCCUUUACCGACGCACAUAUCGAGGUGGACAUUACCU
ACGCCGAGUACUUCGAGAUGAGCGUUCGGCUGGCAGAAGCUAUGAAGCGCUA
UGGGCUGAAUACAAACCAUCGGAUCGUGGUGUGCAGCGAGAAUAGCUUGCAG
UUCUUCAUGCCCCGUGUUGGGUGCCCUGUUCAUCGGUGUGGCUGUGCCCCAGC
UAACGACAUCUACAACGAGCGCGAGCUGCUGAACAGCAUGGGCAUCAGCCAGC
CCACCGUCGUAUUCGUGAGCAAGAAAGGGCUGCAAAAGAUCUCAACGUGCAA
AAGAAGCUACCGAUCAUACAAAAGAUCAUCAUCAUGGAUAGCAAGACCGACU
ACCAGGGCUUCCAAAGCAUGUACACCUUCGUGACUCCCAUUUGCCACCCGGC
UUCAACGAGUACGACUUCGUGCCCAGAGCUUCGACCCGGGACAAAACCAUCGC
CCUGAUGAACAGUAGUGGCAGUACCGGAUUGCCCAAGGGCGUAGCCCUAC
CGCACCGCACCGCUUGUGUCCGAUUCAGUCAUGCCCAGCGACCCCAUCUUCGGC
AACCAGAUCAUCCCCGACACCGCUAUCCUCAGCGUGGUGCCAUUUCACCACGG
CUUCGGCAUGUUCACCACGCUGGGCUACUUGAUCUGCGGCUUUCGGGUCGUGC
UCAUGUACCGCUUCGAGGAGGAGCUAUUCUUGCGCAGCUUGCAAGACUAUAA
GAUUCAAUCUGCCCUGCUGGUGCCACACUAUUUAGCUUCUUCGCUAAGAGCA
CUCUCAUCGACAAGUACGACCUAAGCAACUUGCACGAGAUCCAGCGGGGG
GCGCCGUCAGCAAGGAGGUAGGUGAGGGCCGUGGCCAAACGCUUCCACCUACC
AGGCAUCCGACAGGGCUACGGCCUGACAGAAACAACCAGCGCCAUUCUGAUCA
CCCCGAAGGGGACGACAAGCCUGGCGCAGUAGGCAAGGUGGUGCCCUUCUUC
GAGGCUAAGGUGGUGGACUUGGACACCGGUAAGACACUGGGUGUGAACCAGC
GCGGCGAGCUGUGCGUCCGUGGCCCAUGAUCAUGAGCGGCUACGUUAAACAAC
CCGAGGCUACAAACGCUCUCAUCGACAAGGACGGCUGGCUGCACAGCGGGCA
CAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGUGGACCGGCUGAAGU
CCUGAUCAAAUAACAAGGGCUACCAGGUAGCCCCAGCCGAACUGGAGAGCAUC
CUGCUGCAACACCCCAACUUCGACGCCGGGGUCGCCGGCCUGCCCCGACGA
CGAUGCCGCGAGCUGCCCGCCGAGUCGUCGUGCUGGAACACGGUAAAACCA
UGACCGAGAAGGAGAUUCGUGGACUAUGUGGGCCAGCCAGGUACAACCGCCCAAG
AAGCUGCGGGUGGUGUUGUGUUCGUGGACGAGGUGCCUAAAGGACUGACG
GCAAGUUGGACGCCGCAAGAUCCGCGAGAUUCUCAUUAAGGCCAAGAAGGGC
GGCAAGAUCGCCGUGUAACUCGAGCCGAAACGAAUAGCCGAAAAACAAAAA
ACAAAAAAAACAAAAAAAACCAAAAAACAAAAACACAUUAAAACAGCCUG

UGGGUUGAUCCCACCCACAGGCCCAUGGGCGCUAGCACUCUGGUAUCACGGU
ACCUUUGUGCGCCUGUUUAUAACCCCCUCCCCAACUGUAACUUAGAAGUAAC
ACACACCGAUCAACAGUCAGCGUGGCACACCAGCCACGUUUUGAUCAAGCACU
UCUGUUACCCCGGACUGAGUAUCAAUAGACUGCUCACGCGGUUGAAGGAGAA
AGCGUUCGUUAUCCGGCCAACUACUUCGAAAAAACCUAGUAACACCGUGGAAGU
UGCAGAGUGUUUCGUCAGCACUACCCCAGUGUAGAUCAAGUCAGGAUGAGUCAC
CGCAUUCCCACGGGCGACCGUGGGGUGGCGUUGGCGGCCUGCCCAUGG
GAAACCCAUGGGACGCUCUAUACAGACAUGGUGCGAAGAGUCUAUUGAGC
UAGUUGGUAGUCCUCCGGCCCCUGAAUGCGGCUAAUCCUAAACUGCGGGAGCACA
CACCCUCAAGCCAGAGGGCAGUGUGUCGUAACGGGCAACUCUGCAGCGGAACC
GACUACUUUGGGGUGUCCGUGUUUCAUUUUAUUCUUAUACUGGCUGCUUAUGG
UGACAAUUGAGAGAUCGUUACCAUAUAGCUAUUGGAUUGGCCAUCCGGUGAC
UAAUAGAGCUAUUAUAUUAUCCCUUUGUUGGGUUUAUACCACUUAGCUOGAAA
GAGGUUAAAACAUAUACAUAUUGUUAAGUUGAAUACAGCAAAAUGAGCAA
GAUCUACAUCGACGAGCGGAGCAACGCCGAGAUCGUGUGCGAGGCCAUCAAGA
CCAUCGGCAUCCGAGGGCGCCACCGCCGCCAGCUGACCAGGCAGCUGAACAU
GAGAAGCGGGAGGUGAACAAAGGCCUGUACGACCUGCAGAGGAGCGCUAUGG
UGUACUCCAGCGACGACAUCCUCCCCGUGGUUCAUGACCACCGAGGCCGAC
AAGCCCGACGCCGACGCUAUGGCCGACGUGAUCAUCGACGAGGAGGGA
GAAGUCCAUGAGGGAGGACCACAAGAGCUUCGACGACGUGAUCCCCGCCAAGA
AGAUAUCGACUGGAAGGGCGCAACCCCGUGACCGUGAUCAACGAGUACUGC
CAGAUACCAGGAGGGACUGGAGCUUCCGGAUCGAGAGCGUGGGCCCCAGCAA
CAGGCCCCACCUUCUACGCCUGCGUGGACAUCGACGGCAGGGUGUUCGACAAG
CCGACGGCAAGAGCAAGCGGGACGCCAAGAACAACGCCGCCAAGCUGGCCGUG
GACAAGCUGCUGGGCUACGUGAUCAUCCGGGUUCUAAACAAUUGGCAAGCUGCU
UACAUAGAACUCGCGGCGAUUGGCAUGCCGCCUAAAAUUUUUAUUUUUUUU
UUUCUUUUUUUUUCCGAAUCGGAUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU
AAAAAAAAAAAAAAAAAUCUAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA
AAAAAAAAAAAAAAAAA

TABELLA 8 Numero ID SEQ Descrizione SEQ ID NO: 72 nsP1-4 ORF, codone-ottimizzato SEQ ID NO:
73 5' UTR SEQ ID NO: 74 5' UTR SEQ ID NO: 75 5' UTR SEQ ID NO: 76 3' UTR SEQ ID NO: 121
Glicoproteina spike SARS-CoV-2 (non codone) acido nucleico ottimizzato) SEQ ID NO: 122
Glicoproteina spike della SARS-CoV-2 (codone- acido nucleico ottimizzato) SEQ ID NO: 123
Glicoproteina spike del SARS-CoV-2 (proteina di tipo selvaggio) SEQ ID NO: 77 Regione intergenica
tra nsP1-4 ORF e proteina antigenica ORF SEQ ID NO: 78 Sequenza del replicone comprendente SEQ
ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 76 e SEQ ID NO: 77 SEQ ID NO: 124 Sequenza del replicone
comprendente SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 76, SEQ ID NO: 121 e SEQ ID NO: 77 SEQ ID
NO: 125 Sequenza del replicone comprendente SEQ ID NO: 72, SEQ ID NO: 73, SEQ ID NO: 76, SEQ
ID NO: 122 e SEQ ID NO: 77 (ARCT-021; noto anche come "RNA STARR™ SARS-CoV-2) SEQ ID NO: 79
Sequenza proteica nsP1-4 SEQ ID NO: 80 Sequenza proteica nsP1-4 SEQ ID NO: 81 Sequenza

proteica nsP1-4 SEQ ID NO: 126 mRNA che codifica la glicoproteina SARS-CoV-2 SEQ ID NO: 82 5' UTR (VET) SEQ ID NO: 83 3' UTR (Xbg)

Esempio 10

Questo esempio descrive la caratterizzazione della tecnologia autoreplicante (STARR™) utilizzando l'espressione del transgene luciferasi della lucciola.

Le trascrizioni in vitro sono state formulate con nanoparticelle lipidiche (LNP) a una concentrazione di 0,1 mg/ml e iniettate per via intramuscolare in entrambe le zampe di topi BALB/C femmina (n=3) a una dose di 5 ug per zampa. L'espressione della luciferasi di lucciola (FLuc) è stata misurata da IVIS Lumina LT Series III (PerkinElmer) somministrando 100 ul di 1,5 mg di Xenolight D-luciferina (PerkinElmer) in PBS tramite iniezione intraperitoneale ~10 min prima della misurazione. Sono stati ottenuti sei punti dati per gruppo di topi a ciascun punto temporale (FIGURE 18 A- 18 D).

L'espressione della luciferasi di lucciola (FLuc) è stata monitorata da STARR™ Fluc, SINV FLuc e mRNA FLuc fino al giorno 28 tramite In Vivo Imaging System (IVIS). Sono stati osservati livelli e durate migliorate dell'espressione del transgene da STARR™. L'espressione da STARR™ Fluc ha raggiunto il picco tra il giorno 3 e il giorno 7 ed è diminuita fino al giorno 22. L'espressione di Fluc da SINV FLuc ha raggiunto il picco anche il giorno 10, tuttavia, l'espressione è stata ridotta a una velocità significativamente più rapida rispetto a STARR™ FLuc. Inoltre, l'espressione al giorno 3 è stata significativamente inferiore a STARR™ FLuc. L'espressione di FLuc dal backbone mRNA convenzionale è stata più alta al giorno 1, il momento più precoce in questo studio, ed è diminuita a una velocità leggermente più rapida rispetto a quella di STARR™-Fluc (FIGURA 18 UN). FIGURA 18B mostra che a 14 giorni dalla somministrazione, l'espressione di FLuc da STARR™ FLuc era superiore rispetto agli altri gruppi di circa due ordini di grandezza. FIGURA 18 D dimostra che l'effetto del backbone STARR™ è rimasto minimo per tutto il periodo sperimentale (fino al giorno 28), mentre la precedente somministrazione del backbone del replicone SINV ha determinato una riduzione dell'espressione del transgene FLuc di circa 2 ordini di grandezza.

Successivamente è stato costruito un substrato per il vaccino contro il cancro, TA STARR™, con la struttura portante STARR™ che codifica l'epitopo AH1A5 da gp70, una glicoproteina dell'involucro del virus endogeno della leucemia murina. AH1 (SPSYVYHQF) (SEQ ID NO:110) è un antigene ristretto a H-2Ld di gp70423-431, che è espresso in cellule tumorali come la linea cellulare del cancro coloretale CT26, ma non espresso nella maggior parte dei tessuti normali. AH1-A5 è una sequenza mutata con SPSYAYHQF (SEQ ID NO:111) (la mutazione sottolineata) con affinità migliorata per il recettore delle cellule T (Slansky, et al., 2000, Immunity 13: 529-538). Il frame di lettura aperto dell'RNA subgenomico TA STARR™ contiene una cassetta con un peptide segnale dall'antigene HLA di classe I, sequenza gp70 contenente l'epitopo AH1A5, l'epitopo ovalbumina (OVA323-339) e il segnale di traffico MHC di classe I (Kreiter, et al. 2008, J Immunol 180: 309-318). Tre topi BALB/c femmine sono stati iniettati per via intramuscolare con 10 ug di trascrizioni STARR™ formulate da LNP, STARR™ FLuc o TA STARR™, il giorno 0 e il giorno 7. Il giorno 16, le milze sono state raccolte e gli splenociti sono stati isolati. Gli splenociti (2,5×10⁵ cellule) sono stati

incubati con o senza AH1A5 (SPSYAYHQF) (SEQ ID NO:111), peptide beta-gal (TPHPARIGL) (SEQ ID NO: 112) a 1 ug/ml e 1× Concanavalin A (Life Technologies). ELISpot che rileva IFN-gamma murino (ImmunoSpot) è stato eseguito secondo le istruzioni del produttore. Come si può vedere in FIGURA 19, TA STARR™ ha suscitato risposte IFN-gamma specifiche per l'antigene.

Topi BALB/c, femmine di 10 settimane, sono stati impiantati sottocutaneamente nel fianco destro con 5×10⁵ cellule di cellule CT26 in PBS. Un giorno dopo, RNA STARR™ formulato con LNP è stato iniettato per via intramuscolare nella gamba sinistra a una dose di 10 ug in 100 ul. Ai topi è stata somministrata un'altra dose di richiamo l'ottavo giorno con la stessa dose. Per un gruppo con trattamento combinato di anti-PD1 di topo (RMP1-14, BioXCell) e anti-PDL1 di topo (10F.9G2, BioXCell), l'inibitore del checkpoint combinato (100 ug ciascuno) è stato somministrato tramite iniezione intraperitoneale nel quadrante destro due volte alla settimana per due settimane a partire dal giorno 3. Per un gruppo con trattamento di anti-CTLA4 di topo (9H10, BioXCell), 200 ug dell'inibitore del checkpoint sono stati somministrati nello stesso modo ma a partire dal giorno 7. Cinque topi del gruppo con trattamento combinato di vaccino TA STARR™ e inibitori del checkpoint sono rimasti esenti da tumore al giorno 25 e sono stati ulteriormente sottoposti a impianto sottocutaneo di CT26 (5×10⁵ cellule) nel fianco destro dove il sito di impianto era leggermente al di sopra del primo sito di impianto. I topi naïve sono stati utilizzati come gruppo di controllo. La crescita del tumore è stata monitorata per altri 17 giorni (ossia fino al giorno 42 dal primo impianto di CT26) prima dell'eutanasia. FIGURE 20 A- 20 B illustra la ridotta crescita del tumore risultante dalla vaccinazione TA STARR™ e FIGURA 21 mostra una protezione prolungata risultante dal trattamento con il vaccino TA STARR™ in combinazione con inibitori dei checkpoint.

Gli splenociti del gruppo di trattamento combinato con TA STARR™ e anti-PD1/PDL1 sono stati raccolti per la colorazione del tetramero con il peptide AH1. Gli splenociti del gruppo di controllo con il tampone di formulazione LNP con lo stesso schema di dosaggio sono stati utilizzati come controllo negativo. Gli splenociti (2×10⁶ cellule) sono stati incubati con AH1 (H-2Ld)-tetramero (MBL) seguito da anticorpi marcati con fluorescenza appropriati (Alexa Fluor 488 anti-CD8a (53-6,7), Pacific Orange anti-CD4 (RM4-5) e Pacific Blue anti-topo CD3ε (145-2C11), eBioscience) e DRAQ7 (Invitrogen) seguendo le raccomandazioni del produttore e gli eventi 500K sono stati analizzati da ZE5 Cell Analyzer (Bio-Rad). I risultati sono mostrati in FIGURE 22 A- 22 C.

TABELLA 9 Sequenza nucleotidica ORF del transgene RNA Marcia Indietro # osso Transgene

Sequenza 2809 STELLARE™ Fluttuare

AUGGAAGAUGCCAAAAACAUAAGAAGGGCCAGCGCAUUCUACC (SEGUENTE

CACUCGAAGACGGGACCGCCGGCGAGCAGCUGCACAAAGCCAUGAA ID

GCGCUACGCCUGGUGCCCGGCACCAUCGCCUUUACCGACGCACAU N.: 84)

AUCGAGGUGGACAUUACCUACGCCGAGUACUUCGAGAUGAGCGUUC

GGCUGGCAGAAGCUAUGAAGCGCUAUGGGCUGAAUACAAACCAUCG

GAUCGUGGUGUGCAGCGAGAAUAGCUUGCAGUUCUUAUGCCCGUG

UUGGGUGCCUGUUAUCGGUGUGGCUGUGGCCCCAGCUAACGACA

UCUACAACGAGCGCGAGCUGCUGAACAGCAUGGGCAUCAGCCAGCC

CACCGUCGAUUCGUGAGCAAGAAAGGGCUGCAAAAGAUCUCAAC

GUGCAAAAGAAGCUACCGAUCAUACAAAAGAUAUCAUCAUGGAUA

GCAAGACCGACUACCAGGGCUUCCAAAGCAUGUACACCUUCGUGAC

UUCCCAUUUGCCACCCGGCUUCAACGAGUACGACUUCGUGCCCGAG
AGCUUCGACCGGGACAAAACCAUCGCCUGAUCGAACAGUAGUG
GCAGUACCGGAUUGCCCAAGGGCGUAGCCCUACCGCACCGCACCGC
UUGUGUCCGAUUCAGUCAUGCCCGCGACCCCAUCUUCGGCAACCAG
AUCAUCCCCGACACCGCUAUCCUCAGCGUGGUGCCAUUUCACCACG
GCUUCGGCAUGUUCACCACGCUGGGCUACUUGAUCUGCGGCUUUCG
GGUCGUGCUCAUGUACCGCUUCGAGGAGCUAUUCUUGCGCAGC
UUGCAAGACUAUAAGAUUCAUUCUGCCCUGCUGGGUCCACACUAU
UUAGCUAGCUAAGACUACGUCUUCGCUAAGACUCUCUAAGCAUCGACAAGUACGACCUAAG
CAACUUGCACGAGAUCGCCAGCGGGCGCCGCUAGCAAGGAG
GUAGGUGAGGCCGUGGCCAAACCGCUUCCACCUACCAGGCAUCCAG
AGGGCUACGGCCUGACAGAAACAACCAGCGCCAUUCUGAUCACCCC
CGAAGGGGACGACAAGCCUGGCGCAGUAGGCAAGGUGGUGCCCUUC
UUCGAGGCUAAGGUGGUGGACUUGGACACCGGUAAGACACUGGGUG
UGAACCAGCGCGGCGAGCUGUGCGUCCGUGGCCCCCAUGAUCAGAG
CGGCUACGUUAACAACCCCGAGGCUACAAACGCCUCUCAUCGACAAG
GACGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACG
AGCACUUCUUCUUCGUGGACCGGCUGAAGUCCUGAUCAAAAUACAA
GGGCUACCAGGUAGCCCCAGCCGAACUGGAGAGCAUCCUGCUGCAA
CACCCCAACAUUCGACGCCGGGGUGCGCGGCCUGCCCGACGACG
AUGCCGGCGAGCUGCCCGCCGAGUCGUCGUGCUGGAACACGGUAA
AACCAUGACCGAGAAGGAGAUCGUGGACUAUGUGGGCCAGCCAGGUU
ACAACCGCCAAGAAGCUGCGCGGUGGUGUUGUGUUCGUGGACGAGG
UGCCUAAAGGACUGACCGGCAAGUUGGACGCCCGCAAGAUCGCGA
GAUUCUCAUUAAGGCCAAGAAGGGCGGCAAGAUCCCGUGUAA 2842 Sinv. Fluttuare
AUGGAAGAUGCCAAAACAUAAGAAGGGCCAGCGCAUUCUACC (SEGUENTE replicone
CACUCGAAGACGGGACCGCCGGCGAGCAGCUGCACAAGCCAUGAA ID
GCGCUACGCCUGGUGCCCGGCACCAUCGCCUUUACCGACGCACAU N.: 85)
AUCGAGGUGGACAUUACCUACGCCGAGUACUUCGAGAUGAGCGUUC
GGCUGGCAGAAGCUAUGAAGCGCUAUGGGCUGAAUACAAACCAUCG
GAUCGUGGUGUGCAGCGAGAAUAGCUUGCAGUUCUUCUUCGCCCUG
UUGGGUGCCCUGUUCUUCGGUGUGGCUUGGGCCCCAGCUAACGACA
UCUACAACGAGCGGAGCUGCUGAACAGCAUGGGCAUCAGCCAGCC
CACCGUCGAUUCGUGAGCAAGAAAGGGCUGCAAAGAUCCUCAAC
GUGCAAAGAAGCUACCGAUCAUACAAAAGAUCAUCAUCAUGGAUA
GCAAGACCGACUACCAGGGCUUCCAAAGCAUGUACACCUUCGUGAC
UUCCCAUUUGCCACCCGGCUUCAACGAGUACGACUUCGUGCCCGAG
AGCUUCGACCGGGACAAAACCAUCGCCUGAUCGAACAGUAGUG
GCAGUACCGGAUUGCCCAAGGGCGUAGCCCUACCGCACCGCACCGC
UUGUGUCCGAUUCAGUCAUGCCCGCGACCCCAUCUUCGGCAACCAG
AUCAUCCCCGACACCGCUAUCCUCAGCGUGGUGCCAUUUCACCACG
GCUUCGGCAUGUUCACCACGCUGGGCUACUUGAUCUGCGGCUUUCG
GGUCGUGCUCAUGUACCGCUUCGAGGAGCUAUUCUUGCGCAGC
UUGCAAGACUAUAAGAUUCAUUCUGCCCUGCUGGGUCCACACUAU

UUAGCUAGCUAAGACUACGUCUUCGCUAAGACUCUCUAAGCAUCGACAAGUACGACCUAAG
CAACUUGCACGAGAUCCGAGCGGGCGGGCGCCGCUAGCAAGGAG
GUAGGUGAGGCCGUGGCCAAACCGCUUCCACCUACCAGGCAUCCAG
AGGGCUACGGCCUGACAGAAACAACCAGCGCCAUUCUGAUCACCCC
CGAAGGGGACGACAAGCCUGGCGCAGUAGGCAAGGUGGUGCCCUUC
UUCGAGGCUAAGGUGGUGGACUUGGACACCGGUAAGACACUGGGUG
UGAACCAGCGCGGCGAGCUGUGCGUCCGUGGCCCCCAUGAUC AUGAG
CGGCUACGUUAACAACCCCGAGGCUACAAACGCCUCUCAUCGACAAG
GACGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACG
AGCACUUCUUC AUCGUGGACCGGCUGAAGUCCCUGAUCAAAAUACAA
GGGCUACCAGGUAGCCCCAGCCGAACUGGAGAGCAUCCUGCUGCAA
CACCCCAACAUCUUCGACGCCGGGGUCGCCGGCCUGCCCGACGACG
AUGCCGGCGAGCUGCCCGCCGAGUCGUCGUGCUGGAAACACGGUAA
AACCAUGACCGAGAAGGAGAUCGUGGACUAUGUGGGCCAGCCAGGUU
ACAACCGCCAAGAAGCUGCGCGGUGGUGUUGUGUUCGUGGACGAGG
UGCCUAAAGGACUGACCGGCAAGUUGGACGCCCGCAAGAUCGCGA
GAUUCUCAUUAAGGCCAAGAAGGGCGGCAAGAUCCGUGUAA 1782 mRNA Fluttuare
AUGGAAGAUGCCAAAAACAUAAGAAGGGCCAGCGCAUUCUACC (SEGUENTE (TEV-
CACUCGAAGACGGGACCGCCGGCGAGCAGCUGCACAAGCCAUGAA ID (XbG)
GCGCUACGCCUGGUGCCCGGCACCAUCGCCUUUACCGACGCACAU N.: 86)
AUCGAGGUGGACAUUACCUACGCCGAGUACUUCGAGAUGAGCGUUC
GGCUGGCAGAAGCUAUGAAGCGCUAUGGGCUGAAUACAAACCAUCG
GAUCGUGGUGUGCAGCGAGAAUAGCUUGCAGUUCUUC AUGCCCGUG
UUGGGUGCCUGUUC AUGGUGUGGCUUGGCCCCAGCUAACGACA
UCUACAACGAGCGGAGCUGCUGAACAGCAUGGGCAUCAGCCAGCC
CACCGUCGAUUCGUGAGCAAGAAAGGGCUGCAAAAGAUCCUCAAC
GUGCAAAAGAAGCUACCGAUCAUACAAAAGAUCAUCAUCAUGGAUA
GCAAGACCGACUACCAGGGCUUCCAAAGCAUGUACACCUUCGUGAC
UUCCCAUUGCCACCCGGCUUCAACGAGUACGACUUCGUGCCCGAG
AGCUUCGACCGGGACAAAACCAUCGCCUGAUC AUGAACAGUAGUG
GCAGUACCGGAUUGCCCAAGGGCGUAGCCCUACCGCACCGCACCGC
UUGUGUCCGAUUCAGUCAUGCCCGCGACCCCAUCUUCGGCAACCAG
AUCAUCCCCGACACCGCUAUCCUCAGCGUGGUGCCAUUUCACCACG
GCUUCGGCAUGUUCACCACGCUUGGGCUACUUGAUCUGCGGCUUUCG
GGUCGUGCUAUGUACCGCUUCGAGGAGCUAUUCUUGCGCAGC
UUGCAAGACUAUAAGAUUCAUUCUGCCUGGUGGCCACACUAU
UUAGCUAGCUAAGACUACGUCUUCGCUAAGACUCUCUAAGCAUCGACAAGUACGACCUAAG
CAACUUGCACGAGAUCCGAGCGGGCGGGCGCCGCUAGCAAGGAG
GUAGGUGAGGCCGUGGCCAAACCGCUUCCACCUACCAGGCAUCCAG
AGGGCUACGGCCUGACAGAAACAACCAGCGCCAUUCUGAUCACCCC
CGAAGGGGACGACAAGCCUGGCGCAGUAGGCAAGGUGGUGCCCUUC
UUCGAGGCUAAGGUGGUGGACUUGGACACCGGUAAGACACUGGGUG
UGAACCAGCGCGGCGAGCUGUGCGUCCGUGGCCCCCAUGAUC AUGAG
CGGCUACGUUAACAACCCCGAGGCUACAAACGCCUCUCAUCGACAAG

GACGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACG
AGCACUUCUUCUUCGUGGACCGGCUGAAGUCCCUGAUCAAAAUACAA
GGGCUACCAGGGUAGCCCCAGCCGAACUGGAGAGCAUCCUGCUGCAA
CACCCCAACAUCUUCGACGCCGGGUGCGCCGCCUGCCCCGACGACG
AUGCCGGCGAGCUGCCCGCCGAGUCGUCGUGCUGGAACACGGUAA
AACCAUGACCGAGAAGGAGAUCGUGGACUAUGUGGGCCAGCCAGGUU
ACAACCGCCAAGAAGCUGCGCGGUGGUGUUGUGUUCGUGGACGAGG
UGCCUAAAGGACUGACCGGCAAGUUGGACGCCCGCAAGAUCGCGGA
GAUUCUCAUUAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAA 2847 STELLARE™ Il KRAS
AUGAAGUUGGGUGGUUGUGGGGGCCGGGGUGUUGGCAAAGCGCCC (SEGUENTE epitopo wt
UUACAAUUGA ID N.: 87) 2862 Sinv. Vuoto
AUGGAUCCUAGACGCUACGCCCAAUGAUCCGACCAGCAAACUCG (SEGUENTE replicone
AUGUACUCCGAGGAACUGA ID N.: 88) 3060 STELLARE™ Segnale
AUGAGAGUGACAGCCCCUAGAACCUUACUGCUUCUGCUUUGGGGAG (SEGUENTE peptide-
CUGUUGCUCUGACAGAGACAUGGGCUGGAUCUCUGAGCGAGGUGAC ID gp70 con
CGGCCAGGGCCUGUGCAUCGGCGCCGUGCCCAAGACCCACCAGGUG N.: 89) AH1A5-
CUGUGCAACACCACCAGAAGACCAGCGACGGCAGCUACUACCUGG MITD
CCGCUCCACCAGCACCACCUUGGGCCUGCAGCACCAGCCUGACCCC
UUGCAUCAGCACCACCAUCCUGAACCUGACCACCGACUACUGCGUG
CUGGUGGAGCUGUGGCCAGGGUGACCUACCACAGCCCCAGCUACG
CCUACCACCAGUUCGAGAGGAGGGCCAAGUACAAGAGGGAGCCCGU
GAGCCUGACCCUGGCCUGCUGCGGGCGGCCUGACAAUGGGCGGC
AUCGCCCGCGGUGGGCACCAGCCACCAGCCUGGUGGCCACCC
AGCAGUUCAGCAGCUGCAGGCCGCAUGCACGACGACCUGAAGGA
GGUGGAGAAGUCCAUCACCAACCUGGAGAAGUCCCUGACCAGCCUG
AGCGAGGUGGUGCUGCAGAACAGGAGGGGCCUGGACCUGCUGUCC
UGAAGGAGGGCGGCCUGUGCGCCGCCUGAAGGAGGAGUGCUGCCU
GUACGCCGACCACCCGGCCUGGUGAUCGUGGGCAUUGUCGUGGC
CUGGCCGUCUCGCCGUGGUGGUGAUUGGAGCUGUGGUCGAGCUG
UUAUGUGCAGAAGAAAGUCAUCCGGCGGAAAGGGAGGCUCUACUC
UCAGGCUGCUUCUGCUACAGUGCCUAGAGCUCUUAUGUGUUUAUCU CAGCUGUAA 3061 STELLARE™
Segnale AUGAGAGUGACAGCCCCUAGAACCUUACUGCUUCUGCUUUGGGGAG (SEGUENTE peptide-
CUGUUGCUCUGACAGAGACAUGGGCUGGAUCUACCACAGCCCCAG ID AH1A5 OVA-
CUACGCCUACCACCAGUUCGAGAGGGGGGGGAGGAGGCUCGGGGGA N.: 90) MITD
GGAGGCUCUUGAAGAUCAGCCAGGCCGUGCAGCCGCCACGCCG
AGAUAACGAGGCCGGCCGAGGUGAUCGUGGGCAUUGUCGUGG
CCUGGCCGUCUCGCCGUGGUGGUGAUUGGAGCUGUGGUCGAGCU
GUUAUGUGCAGAAGAAAGUCAUCCGGCGGAAAGGGAGGCUCUACU
CUCAGGCUGCUUCUGCUACAGUGCCUAGAGCUCUAUGUGUUUAUC UCAGCUGUAA 3076 STELLARE™
Segnale AUGAGAGUGACAGCCCCUAGAACCUUACUGCUUCUGCUUUGGGGAG (SEGUENTE peptide-
CUGUUGCUCUGACAGAGACAUGGGCUGGAUCUCUGAGCGAGGUGAC ID gp70 con
CGGCCAGGGCCUGUGCAUCGGCGCCGUGCCCAAGACCCACCAGGUG N.: 91) AH1A5-
CUGUGCAACACCACCAGAAGACCAGCGACGGCAGCUACUACCUGG MITD-BANDIERA
CCGCUCCACCAGCACCACCUUGGGCCUGCAGCACCAGCCUGACCCC

UUGCAUCAGCACCACCAUCCUGAACCUGACCACCGACUACUGCGUG
CUGGUGGAGCUGUGGCCAGGGUGACCUACCACAGCCCAGCUACG
CCUACCACCAGUUCGAGAGGAGGGCCAAGUACAAGAGGGAGCCCGU
GAGCCUGACCCUGGCCUGCUGCUGGGCGGCCUGACAAUGGGCGGC
AUCGCCGCCGGCGUGGGCACCGGCACCACCGCCCUGGUGGCCACCC
AGCAGUUCAGCAGCUGCAGGCCGCAUGCACGACGACCUGAAGGA
GGUGGAGAAGUCCAUCACCAACCUGGAGAAGUCCCUGACCAGCCUG
AGCGAGGUGGUGCUGCAGAACAGGAGGGGCCUGGACCUGCUGUCC
UGAAGGAGGGCGGCCUGUGCGCCGCCUGAAGGAGGAGUGCUGCCU
GUACGCCGACCACACCGGCCUGGUGAUCGUGGGCAUUGUCGUGGC
CUGGCCGUCUCGCGUGGUGGUGAUUGGAGCUGUGGUCGACGUG
UUAUGUGCAGAAGAAAGUCAUCCGGCGGAAAGGGAGGCUCUACUC
UCAGGCUGCUUCUGCUACAGUGCCUAGAGCUCUUAUGUGUUUAUCU
CAGCUGGGCGGGGAGGCAGCGACUACAAGGACGACGAUGACAAGU AAA 3068 STELLARE Segnale
AUGAGAGUGACAGCCCUAGAACCUUACUGCUUCUGCUUUGGGGAG (SEGUENTE peptide-
CUGUUGCUCUGACAGAGACAUGGGCUGGAUCUACCACAGCCCAG ID AH1A5 OVA-
CUACGCCUACCACCAGUUCGAGAGGGGGGGAGGAGGCUCGGGGGA N.: 92) MITD-BANDIERA
GGAGGCUCUUGAAGAUCAGCCAGGCCGUGCAGCCGCCACGCCG
AGAUCAACGAGGCCGGCCGAGGUGAUCGUGGGCAUUGUCGUGG
CCUGGCCGUCUCGCGUGGUGGUGAUUGGAGCUGUGGUCGACGCU
GUUAUGUGCAGAAGAAAGUCAUCCGGCGGAAAGGGAGGCUCUACU
CUCAGGCUGCUUCUGCUACAGUGCCUAGAGCUCUUAUGUGUUUAUC
UCAGCUGGGCGGGGAGGCAGCGACUACAAGGACGACGAUGACAAG SAU Sequenza di amminoacidi
ORF del transgene Marcia transgene # descrizione Sequenza 2809, Fluttuare
MEDAKNIKKGPAPFYPLEDGTAGEQLHKAMKRYALVPGTIAFTDAH 2842,
IEVDITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPV 1782
LGALFIGVAVAPANDIYNERELLSMGSQPTVVFVSKKGLQKILN (SEGUENTE
VQKKLPPIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFVPE ID
SFDRDKTIALIMNSSGSTGLPKGVALPHRTACVRFSHARDPIFGNQ N.: 93)
IIPDTAILSVPFHHGFGMFTTLGYLICGFRVFLMYRFEELFLRS
LQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNLHEIASGGAPLSKE
VGEAVAKRFHLPGIRQGYGLTETTSAILITPEGDDKPGAVGKVVPF
FEAKVVDLDTGKTLGVNQRGELCVRGPMIMSGYVNNPEATNALIDK
DGWLHSGDIAYWDEDEHFFIVDRLKSLIKYKGYQVAPAELESILLQ
HPNIFDAGVAGLPDDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQV
TTAKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKKGGKIAV* 2847 Il KRAS NESSUNA RICHIESTA*
(SEGUENTE epitopo wt ID N.: 94) 2862 Vuoto MDPRRYAPMIRPAKLDVLPN* (SEGUENTE ID N.: 95)
3060 Segnale MRVTAPRTLLLLWGAVALTETWAGSLSEVTGQGLCIGAVPKTHQV (SEGUENTE peptide-
LCNTTQKTSDSYLAAPTGTWACSTGLTPCISTTILNLT TDYCV ID gp70 con
LVELWPRVTYHSPSYAYHQFERRAKYKREPVSLTLALLLGGGLTMGG N.: 96) AH1A5-
IAAGVGTGTTALVATQQFQQLQAAMHDDLKEVEKSITNLEKSLTSL MITD
SEVVLQNRRLDLLFLKEGGLCAALKEECCLYADHTGLVIVGIVAG
LAVLAVVVIGAVVAAVMCRKSSGGKGGSYSQAASATVPRALMCLS QL* 3061 Segnale
MRVTAPRTLLLLWGAVALTETWAGSYHSPSYAYHQFERGGGGSGG (SEGUENTE peptide-

GGSLKISQAVHAAHAEINEAGREVIVGIVAGLAVLAVVVIGAVVAA ID AH1A5 OVA-
VMCRRKSSGGKGGSYSQAASATVPRALMCLSQL* N.: 97) MITD 3076 Segnale
MRVTAPRTL LLLLWGAVALTETWAGSLSEVTGQGLCIGAVPKTHQV (SEGUENTE peptide-
LCNTTQKTS DGSYLAAPTGTTWACSTGLTPCISTTILNLTTDYCV ID gp70 con
LVELWPRVTYHSPSYAYHQFERRAKYKREPVS LTLALLLGGGLTMGG N.: 98) AH1A5-
IAAGVGTGTTALVATQQFQQLQAAMHDDLKEVEKSITNLEKSLTSL MITD-BANDIERA
SEVVLQNRRLDLLFLKEGGLCAALKEECCLYADHTGLVIVGIVAG
LAVLAVVVIGAVVAAVMCRRKSSGGKGGSYSQAASATVPRALMCLS QLGGGSDYKDDDDK* 3068
Segnale MRVTAPRTL LLLLWGAVALTETWAGSYHSPSYAYHQFERGGGGSGG (SEGUENTE peptide-
GGSLKISQAVHAAHAEINEAGREVIVGIVAGLAVLAVVVIGAVVAA Numero: 99) AH1A5 OVA-
VMCRRKSSGGKGGSYSQAASATVPRALMCLSQLGGGSDYKDDDDK MITD-BANDIERA * sequenza
intera di RNA Marcia breve # nome Sequenza 2809 STELLARE™ 2809
AUGGGCGCGCAUGAGAGAAGCCCAGAACCAUUAACCUACCCAAAAU (SEGUENTE Fluttuare
GGAGAAAGUUCACGUUGACAUCGAGGAAGACAGCCCAUUCUCAGA ID
GCUUUGCAGCGGAGCUUCCCGCAGUUUGAGGUAGAAGCCAAGCAGG N.: 100)
UCACUGAUA AUGACCAUGCUAAUGCCAGAGCGUUUUCGCAUCUGGC
UUCAAAACUGAUCGAAACGGAGGUGGACCAUCCGACACGAUCCU
GACAUUGGAAGUGCGCCCGCCGCGAGAAUGUAUUCUAAGCACAAAGU
AUCAUUGUAUCUGUCCGAUGAGAUGUGCGGAAGAUCCGGACAGAU
GUAUAAGUAUGCAACUAAGCUGAAGAAAAACUGUAAGGAAUAACU
GAUAAGGAUUGGACAAGAAAUGAAGGAGCUGGCCCGCGUCAUGA
GCGACCCUGACCUGGAAACUGAGACUAUUGGCCUCCACGACGACGA
GUCGUGUCGCUACGAAGGGCAAGUCGUGUUUACCAGGAUGUAUAC
GCCGUCGACGGCCCCACCAGCCUGUACCACCAGGCCAACAAAGGGCG
UGAGGGUGGCCUACUGGAUCGGCUUCGACACCACACCCUUCAUGU
CAAGAACCUGGCCGGCGCCUACCCCAGCUACAGCACCAACUGGGCC
GACGAGACCGUGCUGACCGCCAGGAACAUCGGCCUGUGCAGCAGCG
ACGUGAUGGAGAGGAGCCGGAGGGCAUGAGCAUCCUGAGGAAGAA
AUACCUGAAGCCCAGCAACAACGUGCUGUUCAGCGUGGGCAGCACC
AUCUACCACGAGAAGAGGGACCUGCUCAGGAGCUGGCACCUGCCCA
GCGUGUCCACCCUGAGGGGCAAGCAGAACUACACCUGCAGGUGCGA
GACCAUCGUGAGCUGCGACGGCUACGUGGUGAAGAGGAUCGCCAUC
AGCCCCGGCCUGUACGCAAGCCCAGCGGCUACGCCGCUACAAUGC
ACAGGGAGGGCUUCCUGUGCUGCAAGGUGACCGACCCUGAACGG
CGAGAGGGUGAGCUUCCCCGUGUGCACCUCGUGCCC GCCACCCUG
UGCGACCAGAUGACCGGCAUCCUGGCCACCGACGUGAGCGCCGACG
ACGCCCAGAAGCUGCUCGUGGGCCUGAACAGAGGAUCGUGGUCAA
CGGCAGGACCCAGAGGAACACCAACACAAUGAAGAACUACCUGCUG
CCCGUGGUGGGCCAGGCUUUCGCCAGGUGGGCCAAGGAGUACAAGG
AGGACCAGGAAGACGAGAGGCCCCUGGGCCUGAGGGACAGGCAGCU
GGUGAUGGGCUGCUGCUGGGCCUUCAGGGCGCACAAGAUCACCAGC
AUCUACAAGAGGCCCGACCCAGACCAUCAUCAAGGUGAACAGCG
ACUCCACAGCUUCGUGCUGCCCAGGAUCGGCAGCAACACCCUGGA
GAUCGGCCUGAGGACCCGGAUCAGGAAGAUGCUGGAGGAACACAAG

GAGCCCAGCCCACUGAUCACCGCCGAGGACGUGCAGGAGGCCAAGU
GCGCUGCCGACGAGGCCAAGGAGGUGAGGGAGGCCGAGGAACUGAG
GGCCGCCUGCCACCCUGGCUGCCGACGUGGAACCCACCCUG
GAAGCCGACGUGGACCUGAUGCUGCAGGAGGCCGCGCCGGAAGCG
UGGAGACACCCAGGGGCCUGAUCAAGGGUGACCAGCUACGACGGCGA
GGACAAGAUCGGCAGCUACGCCGUGCUGAGCCCACAGGCCGUGCUG
AAGUCCGAGAAGCUGAGCUGCAUCCACCCACUGGCCGAGCAGGUGA
UCGUGAUCACCCACAGCGGCAGGAAGGGCAGGUACGCCGUGGAGCC
CUACCACGGCAAGGUGGUCGUGCCGAGGGCCACGCCAUCCCCGUG
CAGGACUUCAGGCCUGAGCGAGAGCGCCACCAUCGUGUACAACG
AGAGGGAGUUCGUGAACAGGUACCUGCACCAUAUCGCCACCCACGG
CGGAGCCCUGAACACCGACGAGGAAUACUACAAGACCGUGAAGCCC
AGCGAGCACGACGGCGAGUACCUUGUACGACAUCGACAGGAAGCAGU
GCGUGAAGAAAGAGCUGGUGACCGGCCUGGGACUGACCGGCGAGCU
GGUGGACCCACCCUUCACGAGUUCGCCUACGAGAGCCUGAGGACC
AGACCCGCCCGCUCCCUACCAGGUGCCACCAUCGGCGUGUACGGCG
UGCCCCGGCAGCGGAAAGAGCGGCAUCAUCAAGAGCGCCGUGACCAA
GAAAGACCUGGUGGUCAGCGCCAAGAAAGAGAACUGCGCCGAGAUC
AUCAGGGACGUGAAGAAGAUGAAAGGCCUGGACGUGAACGCGCGCA
CCGUGGACAGCGUGCUGCUGAACGGCUGCAAGCACCCCGUGGAGAC
CCUGUACAUCGACGAGGCCUUCGCUUGCCACGCCGGCACCCUGAGG
GCCCUGAUCGCCAUCAUCAGGCCCAAGAAAGCCGUGCUGUGCGGCG
ACCCAAAGCAGUGCGGCUUCUUAACAUGAUGUGCCUGAAGGUGCA
CUUCAACCAGGAGAUUCGCACCCAGGUGUUCACAAAGAGCAUCAGC
AGGCGGUGCACCAAGAGCGUGACCAGCGUCGUGAGCACCCUGUUCU
ACGACAAGAAAAUGAGGACCACCAACCCCAAGGAGACCAAAAUCGU
GAUCGACACCACAGGCAGCACCAAGCCCAAGCAGGACGACCUGAUC
CUGACCUGCUUCAGGGGCGUGGUGAAGCAGCUGCAGAUCGACUACA
AGGGCAACGAGAUCAUGACCGCCGUCGCCAGCCAGGGCCUGACCAG
GAAGGGCGUGUACGCCGUGAGGUACAAGGUGAACGAGAACCACUG
UACGCUCCCACCAGCGAGCACGUGAACGUGCUGCUGACCAGGACCG
AGGACAGGAUCGUGUGGAAGACCCUGGCCGGCGACCCUGGAUCAA
GACCCUGACCGCCAAGUACCCCGCAACUUCACCGCCACCAUCGAA
GAGUGGCAGGCCGAGCACGACGCCAUCAUGGAGGCAAUCCUGGAGA
GGCCCCGACCCACCGACGUGUUCAGAAACAAGGCCAACGUGUGCUG
GGCCAAGGCCUGGUGCCCUGCUGAAGACCGCCGGCAUCGACAUG
ACCACAGAGCAGUGGAACACCGUGGACUACUUCGAGACCGACAAGG
CCCACAGCGCCGAGAUCGUGCUGAACCCAGCUGUGCGUGAGGUUCU
CGGCCUGGACCUGGACAGCGGCCUGUUCAGCGCCCCACCGUGCCA
CUGAGCAUCAGGAACAACCACUGGGACAACAGCCCCAGCCAAACA
UGUACGGCCUGAACAAGGAGGUGGUCAGGCAGCUGAGCAGGCGGUA
CCCACAGCUGCCCAGGGCCGUGGCCACCGGCAGGGUGUACGACAUG
AACACCGGCACCCUGAGGAACUACGACCCCAAGGAUCAACCUGGUGC
CCGUGAACAGGCGGCGUCCCCACGCCUGGUGCUGCACCACAACGA

GCACCCACAGAGCGACUUCAGCUCCUUCGUGAGCAAGCUGAAAGGC
AGGACCGUGCUGGUCGUGGGCGAGAAGCUGAGCGUGCCCGGCAAGA
UGGUGGACUGGCUGAGCGACAGGCCCGAGGCCACCUUCCGGGCCAG
GCUGGACCUCCGCAUCCCCGGCGACGUGCCCAAGUACGACAUCAUC
UUCGUGAACGUCAGGACCCCAUCAAGUACCACCCAUUACCAGCAGU
GCGAGGACCACGCCAUCAAGCUGAGCAUGCUGACCAAGAAGGCCUG
CCUGCACCUGAACCCCGGAGGCACCUGCGUGAGCAUCGGCUACGGC
UACGCCGACAGGGCCAGCGAGCAUCAUUGGCGCCAUCGCCAGGC
UGUUCAAGUUCAGCAGGGUGUGCAAACCCAAGAGCAGCCUGGAGGA
AACCGAGGUGCUGUUCGUGUUAUCGGCUACGACCGGAAGGCCAGG
ACCCACAACCCCUACAAGCUGAGCAGCACCCUGACAAACAUCUACA
CCGGCAGGCUGCACGAGGCCGGCUGCGCCCCCAGCUACCACGU
GGUCAGGGGCGAUUUCGCCACCGCCACCGAGGGCGUGAUCAUCAA
GCUGCCAACAGCAAGGGCCAGCCCGGAGGGCGAGUGUGCGGGCCCC
UGUACAAGAAGUUCGGGAGAGCUUCGACCUGCAGCCAUCGAGGU
GGGCAAGGCCAGGCUGGUGAAGGGCGCCGCUAAGCACAUCAUCCAC
GCCGUGGGCCCCAACUUAACAAGGUGAGCGAGGUGGAAGGCGACA
AGCAGCUGGCCGAAGCCUACGAGAGCAUCGCCAAGAUCGUGAACGA
CAUAACUACAAGAGCGUGGCCAUCCACUGCUCAGCACCGGCAUC
UUCAGCGGCAACAAGGACAGGCUGACCCAGAGCCUGAACACCUCG
UCACCGCCCUGGACACCACCGAUGCCGACGUGGCCAUUACUGCAG
GGACAAGAAGUGGGAGAUAGACCCUGAAGGAGGCCGUGGCCAGGCCG
GAGGCCGUGGAAGAGAUUCGCAUCAGCGACGACUCCAGCGUGACCG
AGCCCCGACGCCGAGCUGGUGAGGGUGCACCCCAAGAGCUCCCUGGC
CGGCAGGAAGGGCUACAGCACCGACGGCAAGACCUUCAGCUAC
CUGGAGGGGCACCAAGUUCACCAGGCCGCUAAGGACAUCGCCGAGA
UCAACGCUAUGUGGCCCGUGGCCACCGAGGCCAACGAGCAGGUGUG
CAUGUACAUCUGGGCGAGAGCAUGUCCAGCAUCAGGAGCAAGUGC
CCCUGGAGGAAAGCGAGGCCAGCACACCACCCAGCACCCUGCCCU
GCCUGUGCAUCCACGCUAUGACACCCGAGGGUGCAGCGGCUGAA
GGCCAGCAGGCCCGAGCAGAUACCGUGUGCAGCUCCUCCCACUG
CCCAAGUACAGGAUACCGGCGUGCAGAAGAUCCAGUGCAGCCAGC
CCAUCCUGUUCAGCCAAAGGUGCCCGCCUACAUCCACCCCAGGAA
GUACCGUGGAGACCCACCCGUGGACGAGACACCCGAGCCAAGC
GCCGAGAACCAGAGCACCGAGGGCACACCCGAGCAGCCACCCUGA
UCACCGAGGACGAGACAAGGAACCCGGACCCAGAGCCAUCAUUAU
CGAGGAAGAGGAAGAGGACAGCAUCAGCCUGCUGACGACGGCCCC
ACCCACAGGUGCUGCAGGUGGAGGCCGACAUCCACGGCCCACCCA
GCGUGUCCAGCUCCAGCUGGAGCAUCCACACGCCAGCGACUUCGA
CGUGGACAGCCUGAGCAUCCUGGACACCCUGGAGGGCGCCAGCGUG
ACCUCCGCGCCACCAGCGCCGAGACCAACAGCUACUUCGCCAAGA
GCAUGGAGUUCUGGCCAGGCCCGUGCCAGCUCCAGGACCGUGUU
CAGGAACCCACCCACCCAGCUCCAGGACCAGGCCCAAGCCUG
GCUCCAGCAGGGCCUGCAGCAGGACCAGCCUGGUGAGCACCCAC

CCGGCGUGAACAGGGUGAUCACCAGGGAGGAACUGGAGGCCUGAC
ACCCAGCAGGACCCAGCAGGUCCGUGAGCAGGACUAGUCUGGUG
UCCAACCCACCCGGCGUGAACAGGGUGAUCACCAGGGAGGAAUUCG
AGGCCUUCGUGGCCAGCAACAGAGACGGUUCGACGCCGGCGCCUA
CAUCUUCAGCAGCGACACCGGCCAGGGACACCUAGCAGAAAAGAGC
GUGAGGCAGACCGUGCUGAGCGAGGUGGUGCUGGAGAGGACCGAGC
UGGAAAUCAGCUACGCCCCAGGCUGGACCAGGAGAAGGAGGAACU
GCUAGGAAGAAACUGCAGCUGAACCCACCCAGCCAACAGGAGC
AGGUACCAGAGCAGGAAGGUGGAGAACAUGAAGGCCAUCACCGCCA
GGCGGAUCCUGCAGGGCCUGGGACACUACCUGAAGGCCGAGGGCAA
GGUGGAGUGCUACAGGACCCUGCACCCCGUGCCACUGUACAGCUCC
AGCGUGAACAGGGCCUUCUCCAGCCCCAAGGUGGCCGUGGAGGCCU
GCAACGCUAUGCUGAAGGAGAACUCCCCACCGUGGCCAGCUACUG
CAUCAUCCCCGAGUACGACGCCUACCUGGACAUGGUGGACGGCGCC
AGCUGCUGCCUGGACACCGCCAGCUUCUGCCCCGCCAAGCUGAGGA
GCUUCCCCAAGAAACACAGCUACCUGGAGCCCACCAUCAGGAGCGC
CGUGCCCAGCGCAUCCAGAACACCCUGCAGAACGUGCUGGCCGCU
GCCACCAAGAGGAACUGCAACGUGACCCAGAUGAGGGAGCUUGCCCCG
UGCUGGACAGCGCUGCCUUAACGUGGAGUGCUUCAAGAAAUAACGC
CUGCAACAACGAGUACUGGGAGACCUUCAAGGAGAACCCCAUCAGG
CUGACCGAAGAGAACGUGGUGAACUACAUCACCAAGCUGAAGGGCC
CCAAGGCCGUGCCUGUUCGCUAAGACCCACAACCUGAACAUUCU
GCAGGACAUCCAAUGGACAGGUUCGUGAUGGACCUGAAGAGGGAC
GUGAAGGUGACACCCGGCACCAAGCACACCGAGGAGGCCCAAGG
UGCAGGUGAUCCAGGCCGUGACCCACUGGCCACCGCCUACCUGUG
CGGCAUCCACAGGGAGCUGGUGAGGGCGGUGAACGCCGUGCUGCUG
CCCAACAUCCACACCCUGUUCGACAUGAGCGCCGAGGACUUCGACG
CCAUCAUCGCCGAGCACUUCAGCCCGGCGACUGCGUGCUGGAGAC
CGACAUCGCCAGCUUCGACAAGAGCGAGGAUGACGCUAUGGCCUG
ACCGCUCUGAUGAUCCUGGAGGACCUGGGCGUGGACGCCGAGCUGC
UCACCCUGAUCGAGGCUGCCUUCGGCGAGAUAGCUCCAUCCACCU
GCCACCAAGACCAAGUUAAGUUCGGCGCUAUGAUGAAAAGCGGA
AUGUUCUGACCCUGUUCGUGAACACCGUGAUCAACAUUGUGAUCG
CCAGCAGGGUGCUGCGGGAGAGGCUGACCGGCAGCCCCUGCGCUGC
CUUCAUCGGCGACGACAACAUCGUGAAGGGCGUGAAAAGCGACAAG
CUGAUGGCCGACAGGUGCGCCACCUGGCUGAACAUUGGAGGUGAAGA
UCAUCGACCCGUGGUGGGCGAGAAGGCCCCUACUUCUGCGGCGG
AUUCAUCCUGUGCGACAGCGUGACCGGCACCGCCUGCAGGGUGGCC
GACCCCUGAAGAGGCUGUUAAGCUGGGCAAGCCACUGGGCCGUG
ACGAUGAGCACGACGAUGACAGGCGGAGGGCCUGCACGAGGAAAG
CACCAGGUGGAACAGGGUGGGCAUCCUGAGCGAGCUGUGCAAGGCC
GUGGAGAGCAGGUACGAGACCGUGGGCACCAGCAUCAUCGUGAUGG
CUAUGACCACACUGGCCAGCUCCGUCAAGAGCUUCUCCUACCUGAG
GGGGGCCCUAUAACUCUCAACGGCUAACCGAUGGACUACGACA

UAGUCUAGUCCGCCAAGGCCGCCACCAUGGAAGAUGCCAAAAACAU
UAAGAAGGGCCCAGCGCCAUUCUACCCACUCGAAGACGGGACCGCC
GGCGAGCAGCUGCACAAAGCCAUGAAGCGCUACCGCCUGGUGCCCG
GCACCAUCGCCUUUACCGACGCACAUAUCGAGGUGGACAUAACCUA
CGCCGAGUACUUCGAGAUGAGCGUUCGGCUGGCAGAAGCUAUGAAG
CGCUAUGGGCUGAAUACAACCAUCGGGAUCGUGGUGUGCAGCGAGA
AUAGCUUGCAGUUCUUAUGCCCGUGUUGGGUGCCCGUUAUCGCG
UGUGGCUGUGGCCCCAGCUAACGACAUCUACAACGAGCGCGAGCUG
CUGAACAGCAUGGGCAUCAGCCAGCCCACCGUCGUUAUCGUGAGCA
AGAAAGGGCUGCAAAGAUCUCAACGUGCAAAGAAGCUACCGAU
CAUACAAAAGAUCAUCAUCAUGGAUAGCAAGACCGACUACCAGGGC
UUCCAAAGCAUGUACACCUUCGUGACUUCCCAUUUGCCACCCGGCU
UCAACGAGUACGACUUCGUGCCCAGAGCUUCGACCGGGACAAAAC
CAUCGCCCCUGAUGAACAGUAGUGGCAGUACCGGAUUGCCCAAG
GGCGUAGCCUACCGCACCGCACCUCUUGUGUCCGAUUCAGUCAUG
CCGCGACCCCAUCUUCGGCAACCAGAUCAUCCCGACACCGCUAU
CCUCAGCGUGGUGCCAUUUACCCACGGCUUCGGCAUGUUCACCACG
CUGGGCUACUUGAUCUGCGGCUUUCGGGUCGUCUCAUGUACCGCU
UCGAGGAGGGAGCUAUUCUUGCGCAGCUUGCAAGACUAUAAGAUUCA
AUCUGCCCGCUGGUGCCCACACUAUUUAGCUUCUUCGCUAAGAC
ACUCUCAUCGACAAGUACGACCUAAGCAACUUGCACGAGAUCCCA
GCGGCGGGGCGCCGCUACAGCAAGGAGGGUAGGUGAGGCCGUGGCCAA
ACGCUCCACCUACCAGGAUCCAGGCAGGGCUACGGCCUGACAGAA
ACAACCAGCGCAUUCUGAUCACCCCGAAGGGGACGACAAGCCUG
GCGCAGUAGGCAAGGUGGUGCCUUCUUCGAGGCUAAGGUGGUGGA
CUUGGACACCGGUAAGACACUGGGUGUGAACCAGCGCGGCGAGCUG
UGCGUCCGUGGCCCAUGAUCUAGAGCGGCUACGUUAACAACCCCG
AGGCUACAAACGCUCUCAUCGACAAGGACGGCUGGCUGCACAGCGG
CGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUAUCGUGGAC
CGGCUGAAGUCCUGAUCAAAUACAAGGGCUACCAGGUAGCCCCAG
CCGAACUGGAGAGCAUCCUGCUGCAACACCCCAACAUCUUCGACGC
CGGGGUCGCCGGCCUGCCCAGCAGAUUGCCGGCGAGCUGCCCCGC
GCAGUCGUCGUGCUGGAACACGGUAAAACCAUGACCGAGAAGGAGA
UCGUGGACUAUGUGGCCAGCCAGGUUACAACCGCCAAGAAGCUGCG
CGGUGGUGUUGUGUUCGUGGACGAGGUGCCUAAAGGACUGACCGGC
AAGUUGGACGCCCGAAGAUCCGCGAGAUUCUCAUUAAGGCCAAGA
AGGGCGGCAAGAUCCCGUGUAACUCGAGUAUGUUAACGUGCAAAGG
UGAUUGUACCCCCCGAAAGACCAUAUUGUGACACCCCCUCAGUAU
CACGCCAAACAUUUACAGCCGCGGUGUCAAAAACCGCGUGGACGU
GGUUAACAUCUCCUGCUGGGAGGAUCAGCCGUAAUUAUUAUUAUUGG
CUUGGUGCUGGCUACUAUUGUGGCAUGUACGUGCUGACCAACCAG
AAACAUAUUGAAUACAGCAGCAAUUGGCAAGCUGCUUACAUAAGAA
CUCGCGCAUUGGCAUGCCGCCUAAAAAU ...
CUUUUCUUUCCGAAUCGGAUUUUGUUUUAAUAUUUCAAAAAAAA

AAAAAAAAAAAAAAAAAAUCUAGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2842 Sinv. 2842
AUUGACGGGCGUAGUACACACUAUUGAAUCAACAGCCGACCAAUUG (SEGUENTE Fluttuare
CACUACCAUCACAAUGGAGAAGCCAGUAGUAAACGUAGACGUAGAC ID
CCCCAGAGUCCGUUUGUCGUGCAACUGCAAAAAAGCUUCCCGCAAU N.: 101)
UUGAGGUAGUAGCACAGCAGGUCACUCCAAAUGACCAUGCUGAAUGC
CAGAGCAUUUUCGCAUCUGGCCAGUAAACUAAUCGAGCUGGAGGUU
CCUACCACAGCGACGAUCUUGGACAUAGGCAGCGCACCGGCUCGUA
GAAUGUUUUCGAGCACCAGUAUCAUUGUGUCUGCCCCAUGCGUAG
UCCAGAAGACCCGACCGCAUGAUGAAAUAUGCCAGUAAACUGGCG
GAAAAAGCGUGCAAGAUUACAAACAAGAACUUGCAUGAGAAGAUUA
AGGAUCUCCGGACCGUACUUGAUACGCCGGAUGCUGAAACACCAUC
GCUCUGCUUUCACAACGAUGUUACCUGCAACAUGCGUGCCGAAUAU
UCCGUCAUGCAGGACGUGUAUAUCAACGCUCCCCGGAACUAUCUAUC
AUCAGGCUAUGAAAGGCGUGCGGACCCUGUACUGGAUUGGCUUCGA
CACCACCCAGUUCAUGUUCUCGGCUAUGGCAGGUUCGUACCCUGCG
UACAACACCAACUGGGCCGACGAGAAAGUCCUUGAAGCGCGUAACA
UCGACUUUGCAGCACAAAGCUGAGUGAAGGUAGGACAGGAAAAUU
GUCGAUAAUGAGGAAGAAGGAGUUGAAGCCCGGGUCGCGGGUUUAU
UUCUCCGUAGGAUCGACACUUUAUCCAGAACACAGAGCCAGCUUGC
AGAGCUGGCAUCUCCAUCGGUGUCCACUUGAAUGGAAAGCAGUC
GUACACUUGCCGUGUGAUACAGUGGUGAGUUGCGAAGGCUACGUA
GUGAAGAAAUAACCAUCAGUCCCGGGAUCACGGGAGAAACCGUGG
GAUACGCGGUUACACACAAUAGCGAGGGCUUCUUGCUAUGCAAAGU
UACUGACACAGUAAAAGGAGAACGGGUAUCGUUCCCGUGUGGCACG
UACAUCCCGGCCACCAUUGCGAUCAGAUUGAUCUGGUUAUAAUGGCCA
CGGAUUAUACCCUGACGAUGCACAAAAACUUCUGGUUGGGCUCAA
CCAGCGAAUUGUCAUUAACGGUAGGACUAAACAGGAACACCAACACC
AUGCAAAAUUACCUUCUGCCGAUCAUAGCACAAGGGUUCAGCAAU
GGGCUAAGGAGCGCAAGGAUGAUCUUGAUAAACGAGAAAUGCUGGG
UACUAGAGACGCAAGCUUACGUAUGGCUGCUUGUGGGGCGUUUCG
ACUAAGAAAGUACAUCGUUUUAUCGCCACCUGGAACGCAGACCU
GCGUAAAAGUCCAGCCUCUUUAGCGCUUUUCCCAUGUCGUCCGU
AUGGACGACCUCUUGGCCAUGUCGUGAGGCAGAAAUUGAAACUG
GCAUUGCAACCAAAGAAGGAGGAAAAACUGCUGCAGGUCUCGGAGG
AAUUAGUCAUGGAGGCCAAAGGCUGCUUUUGAGGAUGCUCAGGAGGA
AGCCAGAGCGGAGAAGCUCCGAGAAGCACUCCACCAUAGUGGCA
GACAAAGGCAUCGAGGCAGCCGAGAAGUUGUCUGCGAAGUGGAGG
GGCUCAGGCGGACAUCGGAGCAGCAUUAUUGAAACCCCGCGCGG
UCACGUAAGGAUAAUACCUCAAGCAAUGACCGUAUGAUCGGACAG
UAUAUCGUUGUCUCGCCAAACUCUGUGCUGAAGAAUGCCAAACUCG
CACCAGCGCACCCGCUAGCAGAUAGGUUAAGAUCUAACACACUC
CGGAAGAUCAGGAAGGUACGCGGUCGAACCAUACGACGCUAAAGUA

CUGAUGCCAGCAGGAGGUGCCGUACCAUGGCCAGAAUCCUAGCAC
UGAGUGAGAGCGCCACGUUAGUGUACAACGAAAGAGAGUUUGUGAA
CCGCAAACUAUACCACAUUGCCAUGCAUGGCCCCGCAAGAAUACA
GAAGAGGAGCAGUACAAGGUUACAAGGCAGAGCUUGCAGAAACAG
AGUACGUGUUUGACGUGGACAAGAAGCGUUGCGUUAAGAAGGAAGA
AGCCUCAGGUCUGGUCCUCUCGGGAACUGACCAACCCUCCCCUAU
CAUGAGCUAGCUCUGGAGGGACUGAAGACCCGACCUGCGGUCCCGU
ACAAGGUCGAAACAUAAGGAGUGAUAGGCACACCGGGGUCGGGCAA
GUCAGCUAUUAUCAAGUCAACUGUCACGGCACGAGAUCUUGUACC
AGCGGAAAGAAAGAAAUAUGUCGCGAAAUAUGAGGCCGACGUCUAA
GACUGAGGGGUAUGCAGAUUACGUCGAAGACAGUAGAUUCGGUUAU
GCUCAACGGAUGCCACAAAGCCGUAGAAGUGCUGUACGUUGACGAA
GCGUUCGCGUGCCACGCAGGAGCACUACUUGCCUUGAUUGCUAUCG
UCAGCCCCGCAAGAAGGUAGUACUAUGCGGAGACCCCAUGCAAUG
CGGAUUCUUAACAUGAUGCAACUAAAGGUACAUAUCAAUCACCCU
GAAAAAGACAUAGCACCAAGACAUUCUACAAGUAUAUCUCCCGC
GUUGCACACAGCCAGUACAGCUAUUGUAUCGACACUGCAUACGA
UGGAAAGAUGAAAACCACGAACCCGUGCAAGAAGAACAUAUGAAAUC
GAUAUACAGGGGCCACAAAGCCGAAGCCAGGGGAUAUCAUCCUGA
CAUGUUUCCGCGGGUGGGUUAAGCAAUUGCAAUUCGACUAUCCCGG
ACAUGAAGUAAUGACAGCCGCGCCUCACAAGGGCUAACCAGAAAA
GGAGUGUAUGCCGUCCGGCAAAAAGUCAUAUGAAAACCCACUGUACG
CGAUCACAUCAGAGCAUGUGAACGUGUUGCUCACCCGCACUGAGGA
CAGGCUAGUGUGGAAAACCUUGCAGGGCGACCCAUGGAUUAAGCAG
CUCACUAACAUACCUAAAGGAAACUUCAGGCUACUAUAGAGGACU
GGGAAGCUGAACACAAGGGAAUAAUUGCUGCAAUAAACAGCCCCAC
UCCCCGUGCCA AUCCGUUCAGCUGCAAGACCAACGUUUGCUGGGCG
AAAGCAUUGGAACCGAUACUAGCCACGGCCGUAUCGUACUACCG
GUUGCCAGUGGAGCGAACUGUUCACAGUUUGCGGAUGACAAACC
ACAUUCGGCAUUAACGCCUAGACGUAAUUGCAUUAAGUUUUUC
GGCAUGGACUUGACAAGCGGACUGUUUUCUAAACAGAGCAUCCAC
UAACGUACCAUCCCGCCGAUUCAGCGAGGCCGGUAGCUCAUUGGGA
CAACAGCCCAGGAACCCGCAAGUAUGGGUACGAUCACGCCAUUGCC
GCCGAACUCUCCCGUAGAUUUCGGGUGUUCAGCUAGCUGGGAAGG
GCACACAACUUGAUUUGCAGACGGGGAGAACCAGAGUUAUCUCUGC
ACAGCAUAACCGUGUCCCGGGUGAACCGCAAUCUCCUCACGCCUUA
GUCCCCGAGUACAAGGAGAAGCAACCCGCCCCGUCGAAAAAUUCU
UGAACCAGUUCAAAACACCACUCAGUACUUGUGGUAUCAGAGGAAAA
AAUUGAAGCUCCCCGUAAGAGAAUCGAAUGGAUCGCCCCGAUUGGC
AUAGCCGGUGCAGAUAAAGAACUACAACCUUGGCUUUCGGGUUCCGC
CGCAGGCACGGUACGACCUGGUGUUAUCAACAUUGGAACUAAAUA
CAGAAACCACCACUUCAGCAGUGCGAAGACCAUGCGGGCGACCUUA
AAAACCCUUCGCGUUCGGCCUGAAUUGCCUUAACCCAGGAGGCA
CCCUCGUGGUGAAGUCCUAUGGCUACGCCGACCGCAACAGUGAGGA

CGUAGUCACCGCUCUUGCCAGAAAGUUUGUCAGGGUGUCUGCAGCG
AGACCAGAUUGUGUCUCAAGCAAUACAGAAAUGUACCUGAUUUUCC
GACAACUAGACAACAGCCGUACACGGCAAUUCACCCCGCACCAUCU
GAAUUGCGUGAUUUCGUCCGUGUAUGAGGGUACAAGAGAUGGAGUU
GGAGCCGCGCCGUCAUACCGCACAAAAGGGAGAAUAUUGCUGACU
GUCAAGAGGAAGCAGUUGUCAACGCAGCCAAUCCGCUGGGUAGACC
AGGCGAAGGAGUCUGCCGUGCCAUCUAUAAACGUUGGCCGACCAGU
UUUACCGAUUCAGCCACGGGAGACAGGCACCGCAAGAAUGACUGUGU
GCCUAGGAAAGAAAGUGAUCCACGCGGUCGGCCUGAUUUCCGGAA
GCACCCAGAAGCAGAAGCCUUGAAAUUGCUACAAAACGCCUACCAU
GCAGUGGCAGACUAGUAAAUGAACAUAAUCAAGUCUGUCGCCA
UUCACUGCUAUUCUACAGGCAUUUACGCAGCCGAAAAGACCGCCU
UGAAGUAUCACUUAACUGCUUGACAACCGCGCUAGACAGAACUGAC
GCGGACGUAACCAUCUAUUGCCUGGAUAAGAAGUGGAAGGAAAGAA
UCGAGCGGCACUCCAACUUAAGGAGUCUGUUAACAGAGCUGAAGGA
UGAAGAUUUGGAGAUUCGACGAUGAGUUAGUAUGGAUCCAUCCAGAC
AGUUGCUUGAAGGGAAGAAAGGGAUUCAGUACUACAAAAGGAAAAU
UGUAUUCGUACUUCGAAGGCACAAAUCCAUAAGCAGCAAAAAGA
CAUGGGCGGAGAUAAAGGUCCUGUUCCCUAAUGACCAGGAAAAGUAAU
GAACAACUGUGUGCCUACAUAUUGGGUGAGACCAUGGAAGCAAUCC
GCGAAAAGUGCCCGGUCGACCAUAACCCGUCGUCUAGCCCGCCAA
AACGUUGCCGUGCCUUGCAUGUAUGCCAUGACGCCAGAAAGGGUC
CACAGACUJAGAAGCAAUAACGUCAAAGAAGUUACAGUAUGCUCU
CCACCCCCUUCUUAAGCACAAAUAAGAAUGUUCAGAAGGUUCA
GUGCACGAAAGUAGUCCUGUUUAAUCCGCACACUCCCGCAUUCGUU
CCCGCCCGUAAGUACAUAAGAAGUGCCAGAACAGCCUACCGCUCCUC
CUGCACAGGCCGAGGAGGCCCCCGAAGUUGUAGCGACACCGUCACC
AUCUACAGCUGUAAACACCUCGCUUGAUGUCACAGACAUCUCACUG
GAUAGGAUGACAGAGCGAAGGCUCACUUUUUUCGAGCUUUAGCG
GAUCGGACAACUCUAUUACUAGUAUGGACAGUUGGUCGUCAGGACC
UAGUUCACUAGAGAUAGUAGACCGAAGGCAGGUGGUGGUGGUCGAC
GUUCAUGCCGUCCAAGAGCCUGCCCCCUAUUCCACCGCCAAGGCUAA
AGAAGAUGGCCCGCCUGGCAGCGGCAAGAAAAGAGCCACUCCACC
GGCAAGCAAUAGCUCUGAGUCCCUCCACCUCUCUUUGGUGGGGUA
UCCAUGUCCUCGGAUCAAUUUUCGACGGAGAGACGGCCCGCCAGG
CAGCGGUACAACCCUGGCAACAGGCCCCACGGAUGUGCCUAUGUC
UUUCGGAUCGUUUUCCGACGGAGAGAUUGAUGAGCUGAGCCGCAGA
GUAACUGAGUCCGAACCCGUCUGUUGGAUCAUUUGAACCGGGCG
AAGUGAACUCAAUUAUAUCGUCCCGAUCAGCCGUAUCUUUUCCUCU
ACGCAAGCAGAGACGUAGACGCAGGAGCAGGAGGACUGAAUACUGA
CUAACCGGGUAGGUGGGUACAUAUUUUCGACGGACACAGGCCUG
GGCACUUGCAAAGAAGUCCGUUCUGCAGAACCAGCUACAGAACC
GACCUUGGAGCCGCAAUGUCCUGGAAAGAAUUAUGCCCCGGUGCUC
GACACGUCGAAAGAGGAACAACUCAAACUCAGGUACCAGAUGAUGC

CCACCGAAGCCAACAAAAGUAGGUACCAGUCUCGUAAGUAGAAAA
UCAGAAAGCCAUAAACCACUGAGCGACUACUGUCAGGACUACGACUG
UAUAACUCUGCCACAGAUACGCCAGAAUGCUAUAAGAUCACCUAUC
CGAAACCAUUGUACUCCAGUAGCGUACCGGCGAACUACUCCGAUCC
ACAGUUCGCUGUAGCUGUCUGUAACAACUACUGCAUGAGAACUAU
CCGACAGUAGCAUCUUAUCAGAUUACUGACGAGUACGAUGCUUACU
UGGAUAUGGUAGACGGGACAGUCGCCUGCCUGGACACUGCAACCUU
CUGCCCCGCUAAGCUUAGAAGUACCCGAAAAACAUGAGUAUAGA
GCCCCGAAUAUCCGCAGUGCGGUUCCAUCAGCGAUGCAGAACACGC
UACAAAUGUGUCUAUUGCCGCAACUAAAAGAAAUUGCAACGUCAC
GCAGAU GCGUGAACUGCCAACACUGGACUCAGCGACAUCAAUGUC
GAAUGCUUUCGAAAUAUGCAUGUAAUGACGAGUAUUGGGAGGAGU
UCGCUCGGAAGCCAAUUAAGGAUUAACCACUGAGUUUGUCACCGCAUA
UGUAGCUAGACCUGAAAGGCCCUAAGGCCGCCGCACUAUUUGCAAAG
ACGUAAAUUUUGGUCCAUUGCAAGAAGUGCCUAUGGAUAGAUUCG
UCAUGGACAUGAAAAGAGACGUGAAAGUUACACCAGGCACGAAACA
CACAGAAGAAAGACCGAAAGUACAAGUGAUACAAGCCGCAGAACCC
CUGGGCAGUCGUUACUUAUGCGGGAUUCACCGGGAAUUAGUGCGUA
GGCUUACGGCCGUCUUGCUUCCAACAUCACACGCUUUUUGACAU
GUCGGCGGAGGAUUUUGAUGCAAUCAUAGCAGAACACUUCAAGCAA
GGCGACCCGGUACUGGAGACGGAAUAUCGCAUCAUUCGACAAAAGCC
AAGACGACGCUAUGGCGUUAACCGGUCUGAUGAUCUUGGAGGACCU
GGGUGUGGAUCAACCACUACUCGACUUGAUCGAGUGCGCCUUGGA
GAAUAUCAUCCACCAUCUACCUACGGGUACUCGUUUUAAAUUCG
GGGCGAUGAUGAAAUCCGGAUGUCCUCACACUUUUUGUCAACAC
AGUUUUGAAUGUCGUUAUCGCCAGCAGAGUACUAGAGGAGCGGCUU
AAAACGUCCAGAUGUGCAGCGUUAUUGGCGACGACAACAUCAUAC
AUGGAGUAGUAUCUGACAAAGAAAUGGCUGAGAGGUGCGCCACCUG
GCUCAACAUGGAGGUUAAGAUCAUCGACGCAGUCAUCGGGUGAGAGA
CCACCUUACUUCUGCGGCGGAUUUAUCUUGCAAGAUUCGGUACU
CCACAGCGUGCCGCGUGGCGGAUCCCCUGAAAAGGCUGUUUAAGUU
GGGUAAAACCGCUCCAGCCGACGACGAGCAAGACGAAGACAGAAGA
CGCGCUCUGCUAGAUGAAACAAAGGCGUGGUUUAGAGUAGGUAAUA
CAGGCACUUUAGCAGUGGCCGUGACGACCCGGUAUGAGGUAGACAA
UAUUACACCUGUCCUACUGGCAUUGAGAACUUUUGCCCAGAGCAAA
AGAGCAUUCCAAGCCAUCAGAGGGGAAAUAAGCAUCUCUACGGUG
GUCCUAAAUAAGUCAGCAUAGUACAUUUCAUCUGACUAAUACUACAA
CACCACCACUUGGAAGAUGCCAAAAACAUAAGAAGGGCCCAGCG
CCAUUCUACCCACUCGAAGACGGGACCGCCGGCGAGCAGCUGCACA
AAGCCAUGAAGCGCUACGCCUGGUGCCCGGCACCAUCGCCUUUAC
CGACGCACAUACGAGGUGGACAUUACCUACGCCGAGUACUUCGAG
AUGAGCGUUCGGCUGGACAGCUAUGAAGCGCUAUGGGCUGAAUA
CAAACCAUCGGAUCGUGGUGUGCAGCGAGAAUAGCUUGCAGUUCU
CAUGCCCUGUUGGGUGCCUGUUAUCGGUGUGGCUGUGGCCCA

CUACCCACUCGAAGACGGGACCGCCGGCGAGCAGCUGCACAAAGCC
AUGAAGCGCUACGCCUGGGUGCCCGGCACCAUCGCCUUUACCGACG
CACAUUCGAGGUGGACAUUACCUACGCCGAGUACUUCGAGAUGAG
CGUUCGGCUGGCAGAAGCUAUGAAGCGCUAUUGGGCUGAAUACAAAC
CAUCGGGAUCGUGGUGUGCAGCGAGAAUAGCUUGCAGUUCUUAUGC
CCGUGUUGGGUGCCUGUUAUCGGUGUGGCUGUGGGCCCAGCUAA
CGACAUUCUACAACGAGCGCGAGCUGCUGAACAGCAUGGGCAUCAGC
CAGCCCACCGUCGUUUCGUGAGCAAGAAAGGGCUGCAAAGAUAUC
UCAACGUGCAAAGAAGCUACCGAUCAUACAAAGAUAUCAUCAU
GGAUAGCAAGACCGACUACCAGGGCUUCCAAAGCAUGUACACCUUC
GUGACUUCCAUUGCCACCCGGCUUACAGAGUACGACUUCGUGC
CCGAGAGCUUCGACCGGGACAAAACCAUCGCCUGAUCAUGAACAG
UAGUGGCAGUACCGGAUUGCCCAAGGGCGUAGCCCUACCGCACCGC
ACCGCUUGUGUCCGAUUCAGUCAUGCCCGCGACCCAUUCUGGCA
ACCAGAUCAUCCCCGACACCGCUAUCCUCAGCGUGGUGCCAUUUCA
CCACGGCUUCGGCAUGUUCACCACGCUGGGCUACUIGUAUCUGCGGC
UUUCGGGUCGUGCUAUGUACCGCUUCGAGGAGGAGCUAUUCUUGC
GCAGCUUGCAAGACUAUAAGAUUCAUUCUGCCCUGCUGGUGCCCAC
ACUAAUUAGCUUCUUCGCUAAGAGCACUCUCAUCGACAAGUACGAC
CUAAGCAACUUGCACGAGAUUCGCCAGCGGGGGCGCCGUCAGCA
AGGAGGUAGGUGAGGCCGUGGCCAAACCGCUUCCACCUACCAGGCAU
CCGACAGGGCUACGGCCUGACAGAAACAACCAGCGCAUUCUGAUC
ACCCCGAAGGGGACGACAAGCCUGGGCGAGUAGGCAAGGUGGUGC
CCUUCUUCGAGGCUAAGGUGGUGGACUUGGACACCGGUAAGACACUCU
GGGUGUGAACCAGCGCGGCGAGCUGUGCGUCCGUGGCCCAUGAUC
AUGAGCGGCUACGUUAACAACCCCGAGGCUACAAACGCUCUCAUCG
ACAAGGACGGCUGGCUACACCGGCGACAUCGCCUACUGGGACGA
GGACGAGCACUUCUUCAUUCGUGGACCGGCUGAAGUCCCUGAUCAAA
UACAAGGGCUACAGGUAGCCCCAGCCGAACUGGAGAGCAUCCUGC
UGCAACACCCCAACAUCUUCGACGCCGGGGUCGCCGGCCUGCCCGA
CGACGAUGCCGGCGAGCUGCCCGCCGAGUCGUCGUGCUGGAACAC
GGUAAAACCAUGACCGAGAAGGAGAUUCGUGGACUAGUGGGCCAGCC
AGGUUACAACCGCCAAGAAGCUGCGCGGUGGUGUUGUGUUCGUGGA
CGAGGUGCCUAAAGGACUGACCGGCAAGUUGGACGCCCGCAAGAUC
CGCGAGAUUCUAUUAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGU
AACUCGAGCUAGUGACUGACUAGGAUCUGGUUACCACUAAACCAGC
CUCAAGAACACCCGAAUGGAGUCUCUAAGCUACAUAUACCAACU
ACACUACAAAAUGUUGUCCCCCAAAUGUAGCCAUUCGUAUCUGC
UCCUAAUAAAAAGAAAGUUCUUCACAUUCUAGAAAAAAAAAAAAAA
AA
AA 2847
STELLARE™ 2847 AUGGGCGGCGCAUGAGAGAAGCCCAGAACCAUUACCCAAAAU (SEGUENTE II
KRAS GGAGAAAGUUCACGUUGACAUCGAGGAAGACAGCCCAUUCUCAGA ID peso
GCUUUGCAGCGGAGCUUCCCGAGUUUGAGGUAGAAGCCAAGCAGG N.: 103)

UCACUGAUAAUGACCAUGC UAAUGCCAGAGCGUUUUCGCAUCUGGC
UUCAAAACUGAUCGAAACGGAGGUGGACCCAUCCGACACGAUCCUU
GACAUUGGAAGUGCGCCCCGCCGCAGAAUGUAUUCUAAGCACAAGU
AUCAUUGUAUCUGUCCGAUGAGAUUGUGCGGAAGAUCGGACAGAUU
GUAUAAGUAUGCAACUAAGCUGAAGAAAAACUGUAAGGAAUAACU
GAUAAGGAAUUGGACAAGAAAAUGAAGGAGCUGGCCCGCCGUCAUGA
GCGACCCUGACCUGGAAACUGAGACUAUUGGCCUCCACGACGACGA
GUCGUGUCGCUACGAAGGGCAAGUCGCGUUUUACCAGGAUGUAUAC
GCCGUCGACGGCCCCACCAGCCUGUACCACCAGGCCAACAAGGGCG
UGAGGGUGGCCUACUGGAUCGGCUUCGACACCACACCCUUCAUGUU
CAAGAACCUGGCCGGCGCCUACCCCAGCUACAGCACCAACUGGGCC
GACGAGACCGUGCUGACCGCCAGGAACAUCGGCCUGUGCAGCAGCG
ACGUGAUGGAGAGGAGCCGGAGGGCAUGAGCAUCCUGAGGAAGAA
AUACCUGAAGCCCAGCAACAACGUGCUGUUCAGCGUGGGCAGCACC
AUCUACCACGAGAAGAGGGACCUGCUCAGGAGCUGGCACCUGCCCA
GCGUGUUCACCCUGAGGGGCAAGCAGAACUACACCUGCAGGUGCGA
GACCAUCGUGAGCUGCGACGGCUACGUGGUGAAGAGGAUCGCCAUC
AGCCCCGGCCUGUACGCAAGCCCAGCGGCUACGCCGCUACAAUGC
ACAGGGAGGGCUUCCUGUGCUGCAAGGUGACCGACACCUGAACGG
CGAGAGGGUGAGCUUCCCCGUGUGCACCUACGUGCCCCGCCACCCUG
UGCGACCAGAUGACCGGCAUCCUGGCCACCGACGUGAGCGCCGACG
ACGCCCAGAAGCUGCUCGUGGGCCUGAACCCAGAGGAUCGUGGUCAA
CGGCAGGACCCAGAGGAACACCAACACAAUGAAGAACUACCUGCUG
CCCUGUGGUGGCCAGGCUUUCGCCAGGUGGGCCAAGGAGUACAAGG
AGGACCAGGAAGACGAGAGGCCCUUGGGCCUGAGGGACAGGCAGCU
GGUGAUGGGCUGCUGCUGGGCCUUCAGGCGGCACAAGAUACCCAGC
AUCUACAAGAGGCCCGACACCCAGACCAUCAUAAGGUGAACAGCG
ACUUCCACAGCUUCGUGCUGCCCAGGAUCGGCAGCAACACCCUGGA
GAUCGGCCUGAGGACCCGGAUCAGGAAGAUGCUGGAGGAACACAAG
GAGCCCAGCCCACUGAUCACCGCCGAGGACGUGCAGGAGGCCAAGU
GCGCUGCCGACGAGGCCAAGGAGGUGAGGGAGGCCGAGGAACUGAG
GGCCGCCUGCCACCCUGGCUGCCGACGUGGAACCCACCCUG
GAAGCCGACGUGGACCUGAUGCUGCAGGAGGCCGGCGCCGGAAGCG
UGGAGACACCCAGGGGCCUGAUCAAGGGUGACCAGCUACGACGGCGA
GGACAAGAUCGGCAGCUACGCCGUGCUGAGCCCACAGGCCGUGCUG
AAGUCCGAGAAGCUGAGCUGCAUCCACCCACUGGCCGAGCAGGUGA
UCGUGAUCACCCACAGCGGCAGGAAGGGCAGGUACGCCGUGGAGCC
CUACCACGGCAAGGUGGUCGUGCCGAGGGCCACGCCAUCCCCGUG
CAGGACUUCAGGCCUGAGCGAGAGCGCCACCAUCGUGUACAACG
AGAGGGAGUUCGUGAACAGGUACCUGCACCAUAUCGCCACCCACGG
CGGAGCCUGAACACCGACGAGGAAUACUACAAGACCGUGAAGCCC
AGCGAGCACGACGGCGAGUACCUUGUACGACAUCGACAGGAAGCAGU
GCGUGAAGAAAGAGCUGGUGACCGGCCUGGGACUGACCGGCGAGCU
GGUGGACCCACCCUUCACGAGUUCGCCUACGAGAGCCUGAGGACC

AGACCCGCCCGCUCCCUACCAGGUGCCCACCAUCGGCGUGUACGGCG
UGCCCCGGCAGCGGAAAGAGCGGCAUCAUCAAGAGCGCCGUGACCAA
GAAAGACCUGGUGGUCAGCGCCAAGAAAGAGAACUGCGCCGAGAUC
AUCAGGGACGUGAAGAAGAUGAAAGGCCUGGACGUGAACGCGCGCA
CCGUGGACAGCGUGCUGUAACGGCUGCAAGCACCCCGUGGAGAC
CCUGUACAUCGACGAGGCCUUCGCUUGCCACGCCGGCACCCUGAGG
GCCUGAUCGCCAUCAUCAAGGCCAAGAAAGCCGUGCUGUGCGGCG
ACCCCAAGCAGUGCGGCUUCUUAACAUGAUGUGCCUGAAGGUGCA
CUUCAACCACGAGAUUCGACCCAGGUGUCCACAAGAGCAUCAGC
AGGCGGUGCACCAAGAGCGUGACCAGCGUCGUGAGCACCCUGUUCU
ACGACAAGAAAUGAGGACCACCAACCCCAAGGAGACCAAAAUCGU
GAUCGACACCACAGGCAGCACCAAGCCCAAGCAGGACGACCUGAUC
CUGACCUGCUUCAGGGGCGUGGGUGAAGCAGCUGCAGAUUCGACUACA
AGGGCAACGAGAUCAUGACCGCCGCGUGCCAGCCAGGGCCUGACCAG
GAAGGGCGUGUACGCCGUGAGGUACAAGGUGAACGAGAACCACUG
UACGCUCCACCAGCGAGCACGUGAACGUGCUGCUGACCAGGACCG
AGGACAGGAUCGUGUGGAAGACCCUGGCCGGCGACCCUGGAUCAA
GACCCUGACCGCCAAGUACCCCGGCAACUUCACCGCCACCAUCGAA
GAGUGGCAGGCCGAGCACGACGCCAUCAUUGGAGGCAAUCCUGGAGA
GGCCCCACCCACCAGCUGUUCAGAACAAAGGCCAACGUGUGCUG
GGCCAAGGCCUGGUGCCCGUGCUGAAGACCGCCGCAUCGACAUG
ACCACAGAGCAGUGGAACACCGUGGACUACUUCGAGACCGACAAGG
CCCACAGCGCCGAGAUUCGUGCUGAACCAGCUGUGCGUGAGGUUCU
CGGCCUGGACCUGGACAGCGGCCUGUUCAGCGCCCCACCGUGCCA
CUGAGCAUCAGGAACAACCACUGGGACAACAGCCCCAGCCAAACA
UGUACGGCCUGAACAAAGGAGGUGGUCAGGCAGCUGAGCAGGCGGUA
CCCACAGCUGCCCAGGGCCGUGGCCACCGGCAGGGUGUACGACAUG
AACACCGGCACCCUGAGGAACUACGACCCCAAGAUCAACCUGGUGC
CCGUGAACAGGCGGCUGCCCCACGCCUGGUGCUGCACCACAACGA
GCACCCACAGAGCGACUUCAGCUCCUUCGUGAGCAAGCUGAAAGGC
AGGACCGUGCUGGUCGUGGGCGAGAAGCUGAGCGUGCCCGGCAAGA
UGGUGGACUGGCUGAGCGACAGGCCCGAGGCCACCUCCGGGCCAG
GCUGGACCUCGGCAUCCCCGGCGACGUGCCCAAGUACGACAUCAUC
UUCGUGAACGUCAGGACCCAUACAAGUACCACCAUUAACCAGCAGU
GCGAGGACCACGCCAUCUAAGCUGAGCAUGCUGACCAAGAAGGCCUG
CCUGCACCUGAACCCCGGAGGCCACCGUGCGUGAGCAUCGGCUACGGC
UACGCCGACAGGGCCAGCGAGCAUCAUUGGCGCCAUCGCCAGGC
UGUUCAAGUUCAGCAGGGUGUGCAAACCCAAGAGCAGCCUGGAGGA
AACCGAGGUGCUGUUCGUGUUAUCGGCUACGACCGGAAGGCCAGG
ACCCACAACCCUACAAGCUGAGCAGCACCCUGACAAACAUCUACA
CCGGCAGGCUGCACGAGGCCGGCUGCGCCCCAGCUACCACGU
GGUCAGGGGCGAUUCGCCACCGCCACCGAGGGCGUGAUCAUCAA
GCUGCCAACAGCAAGGGCCAGCCCGGAGGCGGAGUGUGCGGCGCCC
UGUACAAGAAGUUCCCCGAGAGCUUCGACCUGCAGCCCAUCGAGGU

GGGCAAGGCCAGGCUGGUGAAGGGCGCCGCUAAGCACAUCAUCCAC
GCCGUGGGCCCAACUUCAACAAGGUGAGCGAGGUGGAAGGCGACA
AGCAGCUGGCCGAAGCCUACGAGAGCAUCGCCAAGAUCGUGAACGA
CAAUAACUACAAGAGCGUGGCCAUCCACUGCUCAGCACCGGCAUC
UUCAGCGGCAACAAGGACAGGCUGACCCAGAGCCUGAACCACCUGC
UCACCGCCCUGGACACCACCGAUGCCGACGUGGCCAUUCUACUGCAG
GGACAAGAAGUGGGAGAUGACCCUGAAGGAGGCCGUGGCCAGGCGG
GAGGCCGUGGAAGAGAUCUGCAUCAGCGACGACUCCAGCGUGACCG
AGCCCCAGCCGAGCUGGUGAGGGUGCACCCCAAGAGCUCCCUGGC
CGGCAGGAAGGGCUACAGCACCGACGGCAAGACCUUCAGCUAC
CUGGAGGGGCACCAAGUCCACCAGGCCGCUAAGGACAUCGCCGAGA
UCAACGCUAUGUGGCCCGUGGCCACCGAGGCCAACGAGCAGGUGUG
CAUGUACAUCCUGGGCGAGAGCAUGUCCAGCAUCAGGAGCAAGUGC
CCCGUGGAGGAAAGCGAGGCCAGCACACCACCCAGCACCCUGCCCU
GCCUGUGCAUCCACGCUAUGACACCCGAGGGUGCAGCGGCUGAA
GGCCAGCAGGCCCGAGCAGAUACCGUGUGCAGCUCCUCCCACUG
CCCAAGUACAGGAUACCCGGCGUGCAGAAGAUCCAGUGCAGCCAGC
CCAUCCUGUUCAGCCCAAAGGUGCCCGCCUACAUCCACCCCAGGAA
GUACCUGGUGGAGACCCACCCGUGGACGAGACACCCGAGCCAAGC
GCCGAGAACCAGAGCACCGAGGGCACACCCGAGCAGCCACCCUGA
UCACCGAGGACGAGACAAGGAACCCGGACCCAGAGCCAUCAUUAU
CGAGGAAGAGGAAGAGGACAGCAUCAGCCUGCUGACGACGGCCCC
ACCCACCAGGUGCUGCAGGUGGAGGCCGACAUCCACGGCCCACCA
GCGUGUCCAGCUCCAGCUGGAGCAUCCACACGCCAGCGACUUCGA
CGUGGACAGCCUGAGCAUCCUGGACACCCUGGAGGGCGCCAGCGUG
ACCUCCGCGCCACCAGCGCCGAGACCAACAGCUACUUCGCCAAGA
GCAUGGAGUUCUGGCCAGGCCCGUGCCAGCUCCAGGACCCGUGUU
CAGGAACCCACCCACCCAGCUCCAGGACCAGGACCCCAAGCCUG
GCUCCAGCAGGGCCUGCAGCAGGACCAGCCUGGUGAGCACCCAC
CCGGCGUGAACAGGGUGAUACCCAGGGAGGAACUGGAGGCCUGAC
ACCCAGCAGGACCCCAAGCAGGUCCGUGAGCAGGACUAGUCUGGUG
UCCAACCCACCCGGCGUGAACAGGGUGAUACCCAGGGAGGAAUUCG
AGGCCUUCGUGGCCAGCAACAGAGACGGUUCGACGCCGGCGCCUA
CAUCUUCAGCAGCGACACCGGCCAGGGACACCUGCAGCAAAAGAGC
GUGAGGCAGACCGUGCUGAGCGAGGUGGUGCUGGAGAGGACCGAGC
UGGAAAUCAGCUACGCCCCAGGCUGGACCAGGAGAAGGAGGAACU
GCUCAGGAAGAAACUGCAGCUGAACCCACCCAGCCAACAGGAGC
AGGUACCAGAGCAGGAAGGUGGAGAACAUGAAGGCCAUACCCGCCA
GGCGGAUCCUGCAGGGCCUGGACACUACCUGAAGGCCGAGGGCAA
GGUGGAGUGCUACAGGACCCUGCACCCCGUGCCACUGUACAGCUCC
AGCGUGAACAGGGCCUUCUCCAGCCCCAAGGUGGCCGUGGAGGCCU
GCAACGCUAUGCUGAAGGAGAACUCCCCACCGUGGCCAGCUACUG
CAUCAUCCCCGAGUACGACGCCUACCUGGACAUGGUGGACGGCGCC
AGCUGCUGCCUGGACACCGCCAGCUUCUGCCCCGCCAAGCUGAGGA

UUGAGGUAGUAGCACAGCAGGUCACUCCAAAUGACCAUGC UAAUGC
CAGAGCAUUUUCGCAUCUGGCCAGUAAACUAAUCGAGCUGGAGGU
CCUACCACAGCGACGAUCUUGGACAUAGGCAGCGCACCGGCUCGUA
GAAUGUUUUCGAGCACCAGUAUCAUUGUGUCUGCCCCAUGCGUAG
UCCAGAAGACCCGGACCGCAUGAUGAAAUAUGCCAGUAAACUGGCG
GAAAAAGCGUGCAAGAUUACAAACAAGAACUUGCAUGAGAAGAUUA
AGGAUCUCCGGACCGUACUUGAUACGCCGGAUGCUGAAACACCAUC
GCUCUGCUUUCACAACGAUGUUAACCGCAACAUGCGUGCCGAAUAU
UCCGUCAUGCAGGACGUGUAUAUCAACGCUCCCCGAACUAUCUAUC
AUCAGGCUAUGAAAGGCGUGCGGACCCUGUACUGGAUUGGCUUCGA
CACCACCCAGUUCAUGUUCUGGCUAUGGCAGGUUCGUACCCUGCG
UACAACACCAACUGGGCCGACGAGAAAGUCCUUGAAGCGCGUAACA
UCGACUUUGCAGCACAAAGCUGAGUGAAGGUAGGACAGGAAAAUU
GUCGAUAAUGAGGAAGAAGGAGUUGAAGCCCGGGUCGCGGGUUUAU
UUCUCCGUAGGAUCGACACUUUAUCCAGAACACAGAGCCAGCUUGC
AGAGCUGGCAUCUCCAUCGGUGUCCACUUGAAUGGAAAGCAGUC
GUACACUUGCCGUGUGAUACAGUGGUGAGUUGCGAAGGCUACGUA
GUGAAGAAAAUCACCAUCAGUCCCGGGAUCACGGGAGAAACCGUGG
GAUACGCGGUACACACAAUAGCGAGGGCUUCUUGCUAUGCAAAGU
UACUGACACAGUAAAAGGAGAACGGGUAUCGUUCCCGUGUGCAGC
UACAUCCCGGCCACCAUAUGCGAUCAGAUGAUCUGGUUAUAAUGGCCA
CGGAUUAUACCCUGACGAUGCACAAAAACUUCUGGUUGGGCUCAA
CCAGCGAAUUGUCAUUAACGGUAGGACUAACAGGAACACCAACACC
AUGCAAAAUUACCUUCUGCCGAUCAUAGCACAAGGGUUCAGCAAU
GGGCUAAGGAGCGCAAGGAUGAUCUUGAUAAACGAGAAAAUGCUGGG
UACUAGAGACGCAAGCUUACGUAUGGCUGCUUGUGGGGCGUUUCG
ACUAAGAAAGUACAUCGUUUUAUCGCCACCUGGAACGCAGACCU
GCGUAAAAGUCCAGCCUCUUUAGCGCUUUUCCCAUGUCGUCCGU
AUGGACGACCUCUUUGCCCAUGUCGUGAGGCAGAAAUUGAAACUG
GCAUUGCAACCAAAGAAGGAGGAAAAACUGCUGCAGGUCUCGGAGG
AAUUAUGUCAUGGAGGCCAAAGGCUGCUUUUGAGGAUGCUCAGGAGGA
AGCCAGAGCGGAGAAGCUCCGAGAAGCACUUCACCAUUAUGGGCA
GACAAAGGCAUCGAGGCAGCCGAGAAGUUGUCUGCGAAGUGGAGG
GGCUCCAGGCGGACAUCGGAGCAGCAUUAAGUUGAAACCCCGCGCGG
UCACGUAAGGAUAAUACCUCAAGCAAUUGACCGUAUGAUCGGACAG
UAUAUCGUUGUCUCGCCAAACUCUGUGCUGAAGAAUGCCAAACUCG
CACCAGCGCACCCGCUAGCAGAUCAAGGUUAAGAUCUAACACACUC
CGGAAGAUCAGGAAGGUACGCGGUCGAACCAUACGACGCUAAAGUA
CUGAUGCCAGCAGGAGGUGCCGUACCAUGGCCAGAAUUCUAGCAC
UGAGUGAGAGCGCCACGUUAGUGUACAACGAAAGAGAGUUUGUGAA
CCGCAAACUAUACCACAUUGCCAUGCAUGGCCCCCGCCAAGAAUACA
GAAGAGGAGCAGUACAAGGUUACAAAGGCAGAGCUUGCAGAAACAG
AGUACGUGUUUGACGUGGACAAGAAGCGUUGCGUUAAGAAGGAAGA
AGCCUCAGGUCUGGUCCUCUGGGAACUGACCAACCCUCCCUAU

CAUGAGCUAGCUCUGGAGGGACUGAAGACCCGACCUGCGGUCCCGU
ACAAGGUCGAAACAUAAGGAGUGAUAGGCACACCGGGGUCGGGCAA
GUCAGCUAUUAUCAAGUCAACUGUCACGGCACGAGAUUCUUGUACC
AGCGGAAAGAAAGAAAUAUGUCGCGAAAUAUGAGGCCGACGUCUAA
GACUGAGGGGUUAUGCAGAUUACGUCGAAGACAGUAGAUUCGGUUAU
GCUCAACGGAUGCCACAAAGCCGUAGAAGUGCGUACGUUGACGAA
GCGUUCGCGUGCCACGCAGGAGCACUACUUGCCUUGAUUGCUAUCG
UCAGCCCCGCAAGAAGGUAGUACUAUGCGGAGACCCCAUGCAAUG
CGGAUUCUUAACAUGAUGCAACUAAAGGUACAUAUCAAUCACCCU
GAAAAAGACAUAGCACCAAGACAUUCUACAAGUAUAUCUCCCGG
GUUGCACACAGCCAGUACAGCUAUUGUAUCGACACUGCAUACGA
UGGAAAGAUGAAAACCGAACCCGUGCAAGAAGAACAUAUGAAAUC
GAUAUACAGGGGCCACAAAGCCGAAGCCAGGGGAUAUCAUCCUGA
CAUGUUUCCGCGGGGUGGGUUAAGCAAUUGCAAUUCGACUAUCCCGG
ACAUGAAGUAAUGACAGCCGCGGCCUACAAGGGCUAACCCAGAAAA
GGAGUGUAUGCCGUCCGGCAAAAAGUCAUAUGAAAACCCACUGUACG
CGAUCACAUCAGAGCAUGUGAACGUGUUGCUCACCCGCACUGAGGA
CAGGCUAGUGUGGAAAACCUUGCAGGGCGACCCAUGGAUUAAGCAG
CUCACUAACAUACCUAAAAGGAAACUUCAGGCUACUAUAGAGGACU
GGGAAGCUGAACACAAGGGAAUAAUUGCUGCAAUAAACAGCCCCAC
UCCCGUGCCAUCCGUUCAGCUGCAAGACCAACGUUUGCUGGGCG
AAAGCAUUGGAACCGAUACUAGCCACGGCCGUAUCGUACUUAACG
GUUGCCAGUGGAGCGAACUGUUCACAGUUUGCGGAUGACAAACC
ACAUUCGGCCAUUACGCCUAGACGUAAUUGCAUUAAGUUUUUC
GGCAUGGACUUGACAAGCGGACUGUUUUCUAAACAGAGCAUCCAC
UAACGUACCAUCCCGCCGAUUCAGCGAGGCCGGUAGCUCAUUGGGA
CAACAGCCCAGGAACCCGCAAGUAUGGGUACGAUCACGCCAUUGCC
GCCGAACUCUCCCGUAGAUUUCGGGUGUUCAGCUAGCUGGGAAGG
GCACACAACUUGAUUUGCAGACGGGAGAACCAGAGUUAUCUCUGC
ACAGCAUAACCGUGUCCCGGGUGAACCGCAAUCUCCUCACGCCUUA
GUCCCCGAGUACAAGGAGAAGCAACCCGCCCCGUCGAAAAAUUCU
UGAACAGUUCAAACACCACUCAGUACUUGUGGUUAUCAGAGGAAAA
AAUUGAAGCUCUCCCGUAAGAGAAUCGAAUGGAUCGCCCCGAUUGGC
AUAGCCGGUGCAGAUAAAGAACUACAACCUUGGCUUUCGGGUUCCGC
CGCAGGCACGGUACGACCUGGUGUUAUCAACAUAUGGAACUAAAUA
CAGAAACCACCACUUCAGCAGUGCGAAGACCAUGCGGGCGACCUUA
AAAACCCUUCGCGUUCGGCCUGAAUUGCCUUAACCCAGGAGGCA
CCCUCGUGGUGAAGUCCUAUGGCUACGCCGACCGCAACAGUGAGGA
CGUAGUCACCGCUCUUGCCAGAAAGUUUGUCAGGGUGUCUGCAGCG
AGACCAGAUUGUGUCUAAGCAAUACAGAAAUGUACCUGAUUUUCC
GACAACUAGACAACAGCCGUACACGGCAAUUCACCCCGCACCAUCU
GAAUUGCGUGAUUUCGUCUCCGUGUAUGAGGGUACAAGAGAUGGAGUU
GGAGCCGCGCCGUAUACCGCACAAAAGGGAGAAUAUUGCUGACU
GUCAAGAGGAAGCAGUUGUCAACGCAGCCAAUCCGCUGGGUAGACC

AGGCGAAGGAGUCUGCCGUGCCAUCUAUAAAACGUUGGCCGACCAGU
UUUACCGAUUCAGCCACGGGAGACAGGCACCGCAAGAAUGACUGUGU
GCCUAGGAAAGAAAAGUGAUCCACGCGGUCGGCCUGAUUUCCGGAA
GCACCCAGAAGCAGAAGCCUUGAAAUUGCUACAAAACGCCUACCAU
GCAGUGGCAGACUAGUAAAUGAACAUAAUCAAGUCUGUCGCCA
UUCACUGCUAUUCUACAGGCAUUUACGCAGCCGGAAAAGACCGCCU
UGAAGUAUCACUUAACUGCUUGACAACCGCGCUAGACAGAACUGAC
GCGGACGUAACCAUCUAUUGCCUGGAUAAGAAGUGGAAGGAAAGAA
UCGAGCGGCACUCCAACUUAAGGAGUCUGUUAACAGAGCUGAAGGA
UGAAGAUUUGGAGAUCGACGAUGAGUUAGUAUGGAUCCAUCCAGAC
AGUUGCUUGAAGGGAAGAAAGGGAUUCAGUACUACAAAAAGGAAAAU
UGUAUUCGUACUUCGAAGGCACCAAUUCUACAAGCAGCAAAAAGA
CAUGGGCGGAGAUAAAAGGUCCUGUCCCUGAAUGACCAGGAAAAGUAAU
GAACAACUGUGUGCCUACAUAUUGGGUGAGACCAUGGAAGCAAUCC
GCGAAAAGUGCCCGGUCGACCAUAACCCGUCGUCUAGCCCGCCAA
AACGUUGCCGUGCCUUGCAUGUAUGCCAUGACGCCAGAAAGGGUC
CACAGACUAGAAGCAUAACGUCAAAGAAGUUACAGUAUGCUCCU
CCACCCCCUUCUUAAGCACAAAUAAGAAUGUUCAGAAGGUUCA
GUGCACGAAAGUAGUCCUGUUUAUCCGCACACUCCCGCAUUCGUU
CCCGCCCGUAAGUACAUAAGAAGUGCCAGAACAGCCUACCGCUCCUC
CUGCACAGGCCGAGGAGGCCCCGAAGUUGUAGCGACACCGUCACC
AUCUACAGCUGUAAACACCUCGCUUGAUGUCACAGACAUUCACUG
GAUAGGAUGACAGAGCGAAGGCUCACUUUUUUCGAGCUUAGCG
GAUCGGACAACUCUAUUAUAGUAUGGACAGUUGGUCGUCAGGACC
UAGUUCACUAGAGAUAGUAGACCGAAGGCAGGUGGUGGUGGCUGAC
GUUCAUGCCGUCCAAGAGCCUGCCCCUAUUCACCAGCAAGGCUAA
AGAAGAUGGCCCGCCUGGCAGCGGCAAGAAAAGAGCCACUCCACC
GGCAAGCAAUAGCUCUGAGUCCCUCCACCUCUUAUUGGUGGGGUA
UCCAUGUCCUCGGAUCAAUUUUCGACGGAGAGACGGCCCGCCAGG
CAGCGGUACAACCCUGGCAACAGGCCCCACGGAUGUGCCUAUGUC
UUUCGGAUCGUUUUCCGACGGAGAGAUUGAUGAGCUGAGCCGCAGA
GUAACUGAGUCCGAACCCGUCUUGGAUCAUUUGAACCGGGCG
AAGUGAACUCAAUUAUAUCGUCCCAGCAGCCGUAUCUUUUCCUCU
ACGCAAGCAGAGACGUAGACGCAGGAGCAGGAGGACUGAAUACUGA
CUAACCGGGUAGGUGGGUACAUAUUUUCGACGGACACAGGCCUG
GGCACUUGCAAAGAAGUCCGUUCUGCAGAACCAGCUACAGAACC
GACCUUGGAGCCCAAUGUCCUGGAAAGAAUUAUGCCCCGGUGCUC
GACACGUCGAAAGAGGAACAACUCAAACUCAGGUACCAGAUGAUGC
CCACCGAAGCCAACAAAAGUAGGUACCAGUCUCGUAAGUAGAAAA
UCAGAAAAGCAAUAAACACUGAGCGACUACUGUCAGGACUACGACUG
UAUAACUCUGCCACAGAUCAGCCAGAAUGCUAUAAGAUCACCUAUC
CGAAACCAUUGUACUCCAGUAGCGUACCGGCGAACUACUCCGAUCC
ACAGUUCGCUAGCUGUCUGUAACAACUAUCUGCAUGAGAACUAU
CCGACAGUAGCAUCUUAUCAGAUUACUGACGAGUACGAUGCUUACU

UGGAUAUGGUAGACGGGACAGUCGCCUGCCUGGACACUGCAACCUU
CUGCCCCGCUAAGCUUAGAAGUUACCCGAAAAACAUGAGUAUAGA
GCCCCGAAUAUCCGCAGUGCGGUUCCAUCAGCGAUGCAGAACACGC
UACAAAUGUGCUCAUUGCCGCAACUAAAAGAAUUGCAACGUCAC
GCAGAU GCGUGAACUGCCAACACUGGACUCAGCGACAUCAAUGUC
GAAUGCUUUCGAAAUAUGCAUGUAAUGACGAGUAUUGGGAGGAGU
UCGCU CGGAAGCCAAUUAAGGAUUAACCACUGAGUUUGUCACCGCAUA
UGUAGCUAGACCUGAAAGGCCCUAAGGCCGCCGCACUAUUUGCAAAG
ACGUAAAUUUGGUCCAUUGCAAGAAGUGCCUAUGGAUAGAUUCG
UCAUGGACAUGAAAAGAGACGUGAAAGUUACACCAGGCACGAAACA
CACAGAAGAAAGACCGAAAGUACAAGUGAUACAAGCCGCAGAACCC
CUGGGCAGUCGUUACUUAUGCGGGAUUCACCGGGAAUUAUGGCGUA
GGCUUACGGCCGUCUUGCUUCCAACAUCACACGCUUUUUGACAU
GUCGGCGGAGGAUUUUGAUGCAAUCAUAGCAGAACACUUCAAGCAA
GGCGACCCGGUACUGGAGACGGAAUAUCGCAUCAUUCGACAAAAGCC
AAGACGACGCUAUGGCGUUAACCGGUCUGAUGAUCUUGGAGGACCU
GGGUGUGGAUCAACCACUACUCGACUUGAUCGAGUGCGCCUUGGA
GAAUAUCAUCCACCAUCUACCUACGGGUACUCGUUUUAAAUUCG
GGGCGAUGAUGAAAUCCGGAUUGUCCUCACACUUUUUGUCAACAC
AGUUUUGAAUGUCGUUAUCGCCAGCAGAGUACUAGAGGAGCGGCUU
AAAACGUCCAGAUGUGCAGCGUUCAUUGGCGACGACAACAUCAUAC
AUGGAGUAGUAUCUGACAAAAGAAAUGGCUGAGAGGUGCGCCACCUG
GCUCAACAUGGAGGUUAAGAUCAUCGACGCAGUCAUCGGGUGAGAGA
CCACCUUACUUCUGCGGCGGAUUUAUCUUGCAAGAUUCGGUACU
CCACAGCGUGCCGCGUGGCGGAUCCCCUGAAAAGGCUGUUUAAGUU
GGGUA AACCGCUCCCAGCCGACGACGAGCAAGACGAAGACAGAAGA
CGCGCUCUGCUAGAUGAAACAAAGGCGUGGUUUAGAGUAGGUUAUA
CAGGCACUUUAGCAGUGGCGGUGACGACCCGGUAUGAGGUAGACAA
UAUUACACCGUCCUACUGGCAUUGAGAACUUUUGCCCAGAGCAAA
AGAGCAUUCCAAGCCAUCAGAGGGGAAAUAAGCAUCUCUACGGUG
GUCCUAAAUAUGCAGCAUAGUACAUUUCAUCUGACUAAUACUACAA
CACCACCACCGGUGCUAGACCAUGGAUCCUAGACGCUACGCC
CAAUGAUCCGACCAGCAAAACUCGAUGUACUCCGAGGAACUGAUG
UGCAUAAUGCAUCAGGCUGGUACAUAUAGAUC CCGCUUACCGCGGG
CAAUAUAGCAACACUAAAAACUCGAUGUACUUCGAGGAAGCGCAG
UGCAUAAUGCUGCGCAGUGUUGCCACAUAACCACUAUAUUAACCAU
UUAUCUAGCGGACGCCAAAAACUCAAUGUAUUUCUGAGGAAGCGUG
GUGCAUAAUGCCACGCAGCGUCUGCAUAAAUUUUAUUUUUUUU
AUUAAUCAACAAAUUUUGUUUUUAACAUUUCAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAUCUAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA
AA
AA
AUGGGCGGCGCAUGAGAGAAGCCAGAACCAAUACCUACCCAAAAU (SEGUENTE gp70
GGAGAAAGUUCACGUUGACAUCGAGGAAGACAGCCCAUCCUCAGA ID

GCUUUGCAGCGGAGCUUCCCGCAGUUUGAGGUAGAAGCCAAGCAGG N.: 105)

UCACUGAUAUUGACCAUGCUAAUGCCAGAGCGUUUUCGCAUCUGGC
UUCAAAACUGAUCGAAACGGAGGUGGACCCAUCCGACACGAUCCUJ
GACAUUGGAAGUGCGCCCGCCCGCAGAAUGUAUUCUAAGCACAAAGU
AUCAUUGUAUCUGUCCGAUGGAGAUGUGCGGAAGAUCCGGACAGAUJ
GUAUAAGUAUGCAACUAAGCUGAAGAAAAACUGUAAGGAAUAACU
GAUAAGGAAUUGGACAAGAAAAUGAAGGAGCUGGCCGCCGUCAUGA
GCGACCCUGACCUGGAAACUGAGACUAUUGGCCUCCACGACGACGA
GUCGUGUCGCUACGAAGGGCAAGUCGCUUUUACCAGGAUGUAUAC
GCCGUCGACGGCCCCACCAGCCUGUACCACCAGGCCAACAAAGGGCG
UGAGGGUGGCCUACUGGAUCGGCUUCGACACCACACCCUUCAUGUJ
CAAGAACCUGGCCGGCGCCUACCCCAGCUACAGCACCAACUGGGCC
GACGAGACCGUGCUGACCGCCAGGAACAUCGGCCUGUGCAGCAGCG
ACGUGAUGGAGAGGAGCCGGAGGGCAUGAGCAUCCUGAGGAAGAA
AUACCUGAAGCCCAGCAACAACGUGCUGUUCAGCGUGGGCAGCACC
AUCUACCACGAGAAGAGGGACCUGCUCAGGAGCUGGCACCUGCCCA
GCGUGUCCACCCUGAGGGGCAAGCAGAACUACACCCUGCAGGUGCGA
GACCAUCGUGAGCUGCGACGGCUACGUGGUGAAGAGGAUCGCCAUC
AGCCCCGGCCUGUACGCAAGCCCAGCGGCUACGCCGCUACAAUGC
ACAGGGAGGGCUUCCUGUGCUGCAAGGUGACCGACACCCUGAACGG
CGAGAGGGUGAGCUUCCCCGUGUGCACCUACGUGCCCGCCACCCUG
UGCGACCAGAUGACCGGCAUCCUGGCCACCGACGUGAGCGCCGACG
ACGCCCAGAAGCUGCUCGUGGGCCUGAACCCAGAGGAUCGUGGUCAA
CGGCAGGACCCAGAGGAACACCAACACAAUGAAGAACUACCUGCUG
CCCGUGGUGGGCCAGGCUUUCGCCAGGUGGGCCAAGGAGUACAAGG
AGGACCAGGAAGACGAGAGGCCCCUGGGCCUGAGGGACAGGCAGCU
GGUGAUGGGCUGCUGCUGGGCCUUCAGGCGGCACAAGAUCACCAGC
AUCUACAAGAGGCCCGACACCCAGACCAUCAUCAAGGUGAACAGCG
ACUUCCACAGCUUCGUGCUGCCCAGGAUCGGCAGCAACACCCUGGA
GAUCGGCCUGAGGACCCGGAUCAGGAAGAUGCUGGAGGAACACAAG
GAGCCCAGCCCACUGAUCACCGCCGAGGACGUGCAGGAGGCCAAGU
GCGCUGCCGACGAGGCCAAGGAGGUGAGGGAGGCCGAGGAACUGAG
GGCCGCCUGCCACCCUGGCUGCCGACGUGGAACCCACCCUG
GAAGCCGACGUGGACCUGAUGCUGCAGGAGGCCGGCGCCGGAAGCG
UGGAGACACCCAGGGGCCUGAUAAGGGUGACCAGCUACGACGGCGA
GGACAAGAUCGGCAGCUACGCCGUGCUGAGCCCACAGGCCGUGCUG
AAGUCCGAGAAGCUGAGCUGCAUCCACCCACUGGCCGAGCAGGUGA
UCGUGAUCACCCACAGCGGCAGGAAGGGCAGGUACGCCGUGGAGCC
CUACCACGGCAAGGUGGUCGUGCCGAGGGCCACGCCAUCCCCGUG
CAGGACUUCAGGCCUGAGCGAGAGCGCCACCAUCGUGUACAACG
AGAGGGAGUUCGUGAACAGGUACCUGCACCAUAUCGCCACCCACGG
CGGAGCCUGAACACCCGACGAGGAAUACUACAAGACCGUGAAGCCC
AGCGAGCACGACGGCGAGUACCGUACGACAUCGACAGGAAGCAGU
GCGUGAAGAAAGAGCUGGUGACCGGCCUGGGACUGACCGGCGAGCU

GGUGGACCCACCCUCCACGAGUUCGCCUACGAGAGCCUGAGGACC
AGACCCGCCCCGUCCCUACCAGGUGCCCACCAUCGGCGUGUACGGCG
UGCCCCGCGAGCGGAAAGAGCGGCAUCAUCAAGAGCGCCGUGACCAA
GAAAGACCUGGUGGUCAGCGCCAAGAAAGAGAACUGCGCCGAGAUC
AUCAGGGACGUGAAGAAGAUCAAAGGCCUGGACGUGAACGCGCGCA
CCGUGGACAGCGUGCUGCUAACGGCUGCAAGCACCCCGUGGAGAC
CCUGUACAUCGACGAGGCCUUCGCUUGCCACGCCGGCACCCUGAGG
GCCUGAUCGCCAUCAUCAGGCCCAAGAAAGCCGUGCUGUGCGGCG
ACCCAAGCAGUGCGGCUUCUUAACAUGAUGUGCCUGAAGGUGCA
CUUCAACCACGAGAUUCGCACCCAGGUGUCCACAAGAGCAUCAGC
AGGCGGUGCACCAAGAGCGUGACCAGCGUCGUGAGCACCCUGUUCU
ACGACAAGAAAUGAGGACCACCAACCCCAAGGAGACCAAAAUCGU
GAUCGACACCACAGGCAGCACCAAGCCCAAGCAGGACGACCUGAUC
CUGACCUGCUUCAGGGGCGUGGUGAAGCAGCUGCAGAUCGACUACA
AGGGCAACGAGAUCAUGACCGCCGUGCCAGCCAGGGCCUGACCAG
GAAGGGCGUGUACGCCGUGAGGUACAAGGUGAACGAGAACCACUG
UACGCUCCACCAGCGAGCACGUGAACGUGCUGCUGACCAGGACCG
AGGACAGGAUCGUGUGGAAGACCCUGGCCGGCGACCCUGGAUCAA
GACCCUGACCGCCAAGUACCCCGCAACUUCACCGCCACCAUCGAA
GAGUGGCAGGCCGAGCACGACGCCAUCAUGGAGGCAAUCCUGGAGA
GGCCCGACCCACCGACGUGUCCAGAACAAGGCCAACGUGUGCUG
GGCCAAGGCCUGGUGCCCGUGCUGAAGACCGCCGGCAUCGACAUG
ACCACAGAGCAGUGGAACACCGUGGACUACUUCGAGACCGACAAGG
CCCACAGCGCCGAGAUUCGUGCUGAACCCAGCUGUGCGUGAGGUUCU
CGGCCUGGACCUGGACAGCGGCCUGUUCAGCGCCCCACCGUGCCA
CUGAGCAUCAGGAACAACCACUGGGACAACAGCCCCAGCCCCAAACA
UGUACGGCCUGAACAAGGAGGUGGUCAGGCAGCUGAGCAGGCGGUA
CCCACAGCUGCCCAGGGCCGUGGCCACCGGCAGGGUGUACGACAUG
AACACCGGCACCCUGAGGAACUACGACCCACAGGAUCAACCGGUGC
CCGUGAACAGGCGGCGUGCCCCACGCCUGGUGCUGCACCACAACGA
GCACCCACAGAGCGACUUCAGCUCCUUCGUGAGCAAGCUGAAAGGC
AGGACCGUGCUGGUCGUGGGCGAGAAGCUGAGCGUGCCCGGCAAGA
UGGUGGACUGGCUGAGCGACAGGCCCGAGGCCACCUUCCGGGCCAG
GCUGGACCUCCGCAUCCCCGGCGACGUGCCCAAGUACGACAUCAUC
UUCGUGAACGUCAGGACCCAUACAAGUACCACCAUUAACCAGCAGU
GCGAGGACCACGCCAUCAAGCUGAGCAUGCUGACCAAGAAGGCCUG
CCUGCACCUGAACCCCGGAGGCACCUUGCGUGAGCAUCGGCUACGGC
UACGCCGACAGGGCCAGCGAGCAUCAUUGGCGCCAUCGCCAGGC
UGUUAAGUUCAGCAGGGUGUGCAAACCAAGAGCAGCCUGGAGGA
AACCGAGGUGCUGUUCGUGUUAUCGGCUACGACCGGAAGGCCAGG
ACCCACAACCCCUACAAGCUGAGCAGCACCCUGACAAACAUCUACA
CCGGCAGGCUGCACGAGGCCGGCUGCGCCCCAGCUACCACGU
GGUCAGGGGCGAUUCGCCACCGCCACCGAGGGCGUGAUCAUCAAAC
GCUGCCAACAGCAAGGGCCAGCCCCGAGGCGGAGUGUGCGGCGCCC

UGUACAAGAAGUUCCTCCGAGAGCUUCGACCUGCAGCCAU CGAGGU
GGGCAAGGCCAGGCUGGUGAAGGGCGCCGCUAAGCACAUCAUCCAC
GCCGUGGGCCCCAACUUAACAAGGUGAGCGAGGUGGAAGGCGACA
AGCAGCUGGCCGAAGCCUACGAGAGCAUCGCCAAGAUCGUGAACGA
CAAUAACUACAAGAGCGUGGCCAUCCCACUGCUCAGCACCGGCAUC
UUCAGCGGCAACAAGGACAGGCUGACCCAGAGCCUGAACACCUCGC
UCACCGCCCUGGACACCACCGAUGCCGACGUGGCCAUUCUACUGCAG
GGACAAGAAGUGGGGAGAUGACCCUGAAGGAGGCCGUGGCCAGGCGG
GAGGCCGUGGAAGAGAUCUGCAUCAGCGACGACUCCAGCGUGACCG
AGCCCCGACGCCGAGCUGGUGAGGGUGCACCCCAAGAGCUCCCUGGC
CGGCAGGAAGGGCUACAGCACGACGACGGCAAGACCUUCAGCUAC
CUGGAGGGGCACCAAGUUCACCAGGCCGCUAAGGACAUCGCCGAGA
UCAACGCUAUGUGGCCCGUGGCCACCGAGGCCAACGAGCAGGUGUG
CAUGUACAUCUGGGCGAGAGCAUGUCCAGCAUCAGGAGCAAGUGC
CCCUGGAGGAAAGCGAGGCCAGCACACCACCCAGCACCCUGCCCU
GCCUGUGCAUCCACGCUAUGACACCCGAGGGUGCAGCGGCUGAA
GGCCAGCAGGCCCGAGCAGAUACCGUGUGCAGCUCCUCCCACUG
CCCAAGUACAGGAUACCGGCGUGCAGAAGAUCCAGUGCAGCCAGC
CCAUCCUGUUCAGCCCCAAGGUGCCCGCCUACAUCCACCCCAGGAA
GUACCUUGGUGGAGACCCACCCGUGGACGAGACACCCGAGCCAAGC
GCCGAGAACCAGAGCACCGAGGGCACACCCGAGCAGCCACCCUGA
UCACCGAGGACGAGACAAGGAACCCGGACCCCAGAGCCAUCAUUAU
CGAGGAAGAGGAAGAGGACAGCAUCAGCCUGCUGACGACGGCCCC
ACCCACCAGGUGCUGCAGGUGGAGGCCGACAUCACGGCCCCACCCA
GCGUGUCCAGCUCCAGCUGGAGCAUCCCACACGCCAGCGACUUCGA
CGUGGACAGCCUGAGCAUCCUGGACACCCUGGAGGGCGCCAGCGUG
ACCUCCGCGCCACCAGCGCCGAGACCAACAGCUACUUCGCCAAGA
GCAUGGAGUUCUGGCCAGGCCCGUGCCAGCUCCAGGACCGUGUU
CAGGAACCCACCCACCCAGCUCCAGGACCAGGACCCCAAGCCUG
GCUCCAGCAGGGCCUGCAGCAGGACCAGCCUGGUGAGCACCCAC
CCGGCGUGAACAGGGUGAUCACCAGGGAGGAACUGGAGGCCUGAC
ACCCAGCAGGACCCCAAGGUGCCGUGAGCAGGACUAGUCUGGUG
UCCAACCCACCCGGCGUGAACAGGGUGAUCACCAGGGAGGAAUUCG
AGGCCUUCGUGGCCAGCAACAGAGACGGUUCGACGCCGGCGCCUA
CAUCUUCAGCAGCGACACCGGCCAGGGACACCUAGCAGAAAAGAGC
GUGAGGCAGACCGUGCUGAGCGAGGUGGUGCUGGAGAGGACCGAGC
UGGAAAUCAGCUACGCCCCAGGCUGGACCAGGAGAAGGAGGAACU
GCUCAGGAAGAAACUGCAGCUGAACCCCAACCCAGCCAACAGGAGC
AGGUACCAGAGCAGGAAGGUGGAGAACAUGAAGGCCAUCACCGCCA
GGCGGAUCCUGCAGGGCCUGGGACACUACCUGAAGGCCGAGGGCAA
GGUGGAGUGCUACAGGACCCUGCACCCCGUGCCACUGUACAGCUCC
AGCGUGAACAGGGCCUUCUCCAGCCCCAAGGUGGCCGUGGAGGCCU
GCAACGCUAUGCUGAAGGAGAACUCCCCACCGUGGCCAGCUACUG
CAUCAUCCCCGAGUACGACGCCUACCGGACAUGGUGGACGGCGCC

AGCUGCUGCCUGGACACCGCCAGCUUCUGCCCCGCAAGCUGAGGA
GCUUCCCCAAGAAACACAGCUACCUGGAGCCCACCAUCAGGAGCGC
CGUGCCCAGCGCCAUCCAGAACACCCUGCAGAACGUGCUGGCCGCU
GCCACCAAGAGGAACUGCAACGUGACCCAGAUGAGGGAGCUUCCCCG
UGCUGGACAGCGCUGCCUUCAACGUGGAGUGCUUCAAGAAAAUACGC
CUGCAACAACGAGUACUGGGGAGACCUUCAAGGAGAACCCCAUCAGG
CUGACCGAAGAGAACGUGGUGAACUACAUCACCAAGCUGAAGGGCC
CCAAGGCCGUGCCUGUUCGCUAAGACCCACAACCUGAACAUAGCU
GCAGGACAUCCAAUGGACAGGUUCGUGAUGGACCUGAAGAGGGAC
GUGAAGGUGACACCCGGCACCAAGCACACCCGAGGAGGCCCAAGG
UGCAGGUGAUCCAGGCCGUGACCCACUGGCCACCGCCUACCUGUG
CGGCAUCCACAGGGAGCUGGUGAGGGCGGUGAACGCCGUGCUGCUG
CCCAACAUCCACACCCUGUUCGACAUGAGCGCCGAGGACUUCGACG
CCAUCAUCGCCGAGCACUUCAGCCCGGCGACUGCGUGCUGGAGAC
CGACAUCGCCAGCUUCGACAAGAGCGAGGAUGACGCUAUGGCCUG
ACCGCUCUGAUGAUCCUGGAGGACCUGGGCGUGGACGCCGAGCUGC
UCACCCUGAUCGAGGCUGCCUUCGGCGAGAUCAGCUCCAUCCACCU
GCCACCAAGACCAAGUUCAAGUUCGGCGCUAUGAUGAAAAGCGGA
AUGUUCUGACCCUGUUCGUGAACACCGUGAUAACAUGUGAUCG
CCAGCAGGGUGCUGCGGGAGAGGCUGACCGGCAGCCCCUGCGCUGC
CUUCAUCGGCGACGACAACAUCGUGAAGGGCGUGAAAAGCGACAAG
CUGAUGGCCGACAGGUGCGCCACCUGGCUGAACAUUGGAGGUGAAGA
UCAUCGACGCCGUGGUGGGCGAGAAGGCCCCUACUUCUGCGGCGG
AUUCAUCCUGUGCGACAGCGUGACCGGCACCGCCUGCAGGGUGGCC
GACCCCUGAAGAGGCUGUUCAAGCUGGGCAAGCCACUGGGCCGUG
ACGAUGAGCACGACGAUGACAGGGCGAGGGCCUGCACGAGGAAAG
CACCAGGUGGAACAGGGUGGGCAUCCUGAGCGAGCUGUGCAAGGCC
GUGGAGAGCAGGUACGAGACCGUGGGCACCAGCAUCAUCGUGAUGG
CUAUGACCACACUGGCCAGCUCCGUCAAGAGCUUCUCCUACCUGAG
GGGGGCCCUAUAACUCUCAACGGCUAACCUGAAUGGACUACGACA
UAGUCUAGUCCGCCAAGGCCGCCACCAUGAGAGUGACAGCCCUAG
AACCUUACUGCUUCUGCUUUGGGGAGCUGUUGCUCUGACAGAGACA
UGGGCUGGAUCUCUGAGCGAGGUGACCGGCCAGGGCCUGUGCAUCG
GCGCCGUGCCCAAGACCCACCAGGUGCUGUGCAACACCACCCAGAA
GACCAGCGACGGCAGCUACUACCUGGCCGCCUCCACCGGCACCACC
UGGCCUGCAGCACCGGCCUGACCCCUUGCAUCAGCACCACCAUCC
UGAACCUGACCACCGACUACUGCGUGCUGGUGGAGCUGUGGCCAG
GGUGACCUACCACAGCCCCAGCUACGCCUACCACCAGUUCGAGAGG
AGGGCCAAGUACAAGAGGGAGCCCGUGAGCCUGACCCUGGCCUGC
UGCUGGGCGGCCUGACAAUGGGCGGCAUCGCCGCCGGCGUGGGCAC
CGGCACCACCGCCUGGUGGCCACCCAGCAGUUCAGCAGCUGCAG
GCCGCCAUGCACGACGACCUGAAGGAGGUGGAGAAGUCCAUCACCA
ACCUGGAGAAGUCCUGACCAGCCUGAGCGAGGUGGUGCUGCAGAA
CAGGAGGGCCUGGACCUGCUGUCCUGAAGGAGGGCGGCCUGUGC

GCCCCCCUGAAGGAGGAGUGCUGCCUGUACGCCGACCACACCGGCC
UGGUGAUCGUGGGCAUUGUCGUGGCCUGGCCGUCCUCGCCGUGGU
GGUGAUUGGAGCUGUGGUCGAGCUGUUAUGUGCAGAAGAAAGUCA
UCCGGCGGAAAAGGAGGCUCUACUCUCAGGCUGCUUCUGCUACAG
UGCCUAGAGCUCUUAUGUGUUAUCUCAGCUGUAAACUCGAGUAUG
UUACGUGCAAAGGUGAUUGUCACCCCCGAAAGACCAUAUUGUGAC
ACACCCUCAGUAUCACGCCCAAACAUAUUACAGCCGCGGUGUCAAAA
ACCGCGUGCUGGACGUGGUUAACAUCCUGCUGGGAGGAUCAGCCGUAA
UUAUUAUAAUUGGCUUGGUGCUGGCUACUAUUGUGGGCCAUGUACGU
GCUGACCAACCAGAAACAUAUUGAAUACAGCAGCAAUUGGCAAGC
UGCUIACAUAGAACUCGCGGCGAUUGGCAUGCCGCCUAAAAUUUU
UAUUUUUAUUUUUUUCUUUUUCUUUUUCCGAAUCGGAUUUUUGUUUUUAU
AUUUCAA
AA
AA 3061 STELLARE™ 3061
AUGGGCGGCGCAUGAGAGAAGCCCAGAACCAUUACCUACCCAAA (SEGUENTE AH1A5
AUGGAGAAAGUUCACGUUGACAUCGAGGAAGACAGCCAUUCCU ID
CAGAGCUUUGCAGCGGAGCUUCCCCGAGUUUGAGGUAGAAGCCA N.: 106)
AGCAGGUCACUGAUAAUGACCAUGCUAUUGCCAGACGUUUUCG
CAUCUGGCUUCAAAACUGAUCGAAACGGAGGUGGACCCAUCCGA
CACGAUCCUUGACAUUGGAAGUGCGCCCGCCGCGAGAAUGUAU
CUAAGCACAAGUAUCAUUGUAUCUGUCCGAUGAGAUUGUGCGGAA
GAUCCGGACAGAUUGUAUAAGUAUGCAACUAAGCUGAAGAAAA
CUGUAAGGAAUAACUGAUAAAGGAAUUGGACAAGAAAAUGAAGG
AGCUGGCCCGGCUAUGAGCGACCCUGACCUGGAAACUGAGACU
AUGUGCCUCCACGACGACGAGUCGUGUCGUACGAAGGGCAAGU
CGCUGUUUACCAGGAUGUAUACGCCGUCGACGGCCCCACCAGCC
UGUACCACCAGGGCAACAAGGGCGUGAGGGUGGCCUACUGGAUC
GGCUUCGACACCACCCCUUCAUGUUAAGAACCUGGCCGGCGC
CUACCCAGCUACAGCACCAACUGGGCCGACGAGACCGUGCUGA
CCGCCAGGAACAUCGGCCUGUGCAGCAGCGACGUGAUGGAGAGG
AGCCGGAGAGGCAUGAGCAUCCUGAGGAAGAAAUACCUGAAGCC
CAGCAACAACGUGCUGUUCAGCGUGGGCAGCACCAUCUACCACG
AGAAGAGGGACCUGCUCAGGAGCUGGCACCUGCCCAGCGUGUUC
CACCUGAGGGGGCAAGCAGAACUACACCUGCAGGUGCGAGACCAU
CGUGAGCUGCGACGGCUACGUGGUGGAAGAGGAUCGCCAUCAGCC
CCGGCCUGUACGCAAGCCCAGCGGCUACGCCGCUACAAUGCAC
AGGGAGGGCUUCCUGUGCUGCAAGGUGACCGACCCUGAACGG
CGAGAGGGUGAGCUUCCCCGUGUGCACCUACGUGCCC GCCACCC
UGUGCGACCAGAUAGCCGGCAUCCUGGCCACCGACGUGAGCGCC
GACGACGCCAGAAGCUGCUCGUGGGCCUGAACCAGAGGAUCGU
GGUCAACGGCAGGACCCAGAGGAACACCAACACAAUGAAGAACU
ACCUGCUGCCCUGGUGGGCCAGGCUUUCGCCAGGUGGGCCAAG
GAGUACAAGGAGGACCAGGAAGACGAGAGGCCCCUGGGCCUGAG

GGACAGGCAGCUGGUGAUGGGCUGCUGCUGGGCCUUCAGGCGGC
ACAAGAUCACCAGCAUCUACAAGAGCCCGACACCCAGACCAUC
AUCAAGGUGAACAGCGACUCCACAGCUUCGUGCUGCCCAGGAU
CGGCAGCAACACCCUGGAGAUCGGCCUGAGGACCCGGAUCAGGA
AGAUGCUGGAGGAACACAAGGAGCCCAGCCCACUGAUCACCGCC
GAGGACGUGCAGGAGGCCAAGUGCGCUGCCGACGAGGCCAAGGA
GGUGAGGGAGGCCGAGGAACUGAGGGCCGCCUGCCACCCUGG
CUGCCGACGUGGAGGAACCCACCCUGGAAGCCGACGUGGACCUG
AGOSTOCUGCAGGAGGCCGGCGCCGGAAGCGUGGAGACACCCAGGGG
CCUGAUCAAGGUGACCAGCUACGACGGCGAGGACAAGAUCGGCA
GCUACGCCGUGCUGAGCCCACAGGCCGUGCUGAAGUCCGAGAAG
CUGAGCUGCAUCCACCCACUGGCCGAGCAGGUGAUCGUGAUCAC
CCACAGCGGCAGGAAGGGCAGGUACGCCGUGGAGCCCUACCACG
GCAAGGUGGUCGUGCCGAGGGCCACGCCAUCCCCGUGCAGGAC
UCCAGGCCUGAGCGAGAGCGCCACCAUCGUGUACAACGAGAG
GGAGUUCGUGAACAGGUACCUGCACCAUAUCGCCACCCACGGCG
GAGCCUGAACACCGACGAGGAAUACUACAAGACCGUGAAGCCC
AGCGAGCACGACGGCGAGUACCGUACGACAUCGACAGGAAGCA
GUGCGUGAAGAAAGAGCUGGUGACCGGCCUGGGACUGACCGGCG
AGCUGGGUGACCCACCCUUCACGAGUUCGCCUACGAGAGCCUG
AGGACCAGACCCGCCGCUCCUACCAGGUGCCCACCAUCGGCGU
GUACGGCGUGCCCGCAGCGGAAAGAGCGGCAUCAUCAAGAGCG
CCGUGACCAAGAAAGACCUGGUGGUCAGCGCCAAGAAAGAGAAC
UGCGCCGAGAUCAUCAGGGACGUGAAGAAGAUCAAAGGCCUGGA
CGUGAACGCGCGCACCGUGGACAGCGUGCUGCUGAACGGCUGCA
AGCACCCCGUGGAGACCCUGUACAUCGACGAGGCCUUCGCUUGC
CACGCCGGCACCCUGAGGGCCUGAUCGCCAUCAUCAGGCCCAA
GAAAGCCGUGCUGUGCGGGACCCCAAGCAGUGCGGCUUCUUCA
ACAUGAUGUGCCUGAAGGUGCACUUAACCACGAGAUCUGCACC
CAGGUGUUCACAAGAGCAUCAGCAGGGCGGUGCACCAAGAGCGU
GACCAGCGUCGUGAGCACCCUGUUCUACGACAAGAAAUGAGGA
CCACCAACCCCAAGGAGACCAAAAUCGUGAUCGACACCACAGGC
AGCACCAAGCCCAAGCAGGACGACCUGAUCUGACCUGCUUCAG
GGGCUGGGUGAAGCAGCUGCAGAUCAAGGGCAACGAGA
UCAUGACCCGCCGUGCCAGCCAGGGCCUGACCAGGAAGGGCGUG
UACGCCGUGAGGUACAAGGUGAACGAGAACCCACUGUACGCUCC
CACCAGCGAGCACGUGAACGUGCUGCUGACCAGGACCGAGGACA
GGAUCGUGUGAAGACCCUGGCCGGCGACCCUGGAUCAAGACC
CUGACCGCCAAGUACCCCGCAACUUCACCGCCACCAUCGAAGA
GUGGCAGGCCGAGCACGACGCCAUCAUGAGGCACAUCUGGAGA
GGCCCGACCCACCGACGUGUUCGAGAACAAGGCCAACGUGUGC
UGGGCCAAGGCCUGGUGCCCGUGCUGAAGACCGCCGGCAUCGA
CAUGACCACAGAGCAGUGGAACACCGUGGACUACUUCGAGACCG
ACAAGGCCACAGCGCCGAGAUCGUGCUGAACCAGCUGUGCGUG

AGGUUCUUCGGCCUGGACCUGGACAGCGGCCCUUUCAGCGCCCC
CACCGUGCCACUGAGCAUCAGGAACAACCACUGGGACAACAGCC
CCAGCCCAAACAUGUACGGCCUGAACAAAGGAGGUGGUCAGGCAG
CUGAGCAGGCGGUACCCACAGCUGCCCCAGGGCCUGGCCACCGG
CAGGGUGUACGACAUGAACACCGGCACCCUGAGGAACUACGACC
CCAGGAUCAACCUUGGUGCCCGUGAACAGGGCGCUGCCCCACGCC
CUGGUGCUGCACCACAACGAGCACCCACAGAGCGACUUCAGCUC
CUUCGUGAGCAAGCUGAAAGGCAGGACCGUGCUGGUCUGGGCG
AGAAGCUGAGCGUGCCCGCAAGAUGGUGGACUGGCUGAGCGAC
AGGCCCGAGGCCACCUUCGGGCCAGGCUGACCUCGGCAUCCC
CGGCGACGUGCCCAAGUACGACAUCUUCGUGAACGUCAGGA
CCCCAUACAAGUACCACCAUUAACCAGCAGUGCGAGGACCACGCC
AUCAAGCUGAGCAUGCUGACCAAGAAGGCCUGCCUGCACCUGAA
CCCCGGAGGCACCUGCGUGAGCAUCGGCUACGGCUACGCCGACA
GGGCCAGCGAGAGCAUCAUUGGCGCCAUCGCCAGGCUGUUAAG
UUCAGCAGGGUGUGCAAACCCAAGAGCAGCCUGGAGGAAACCGA
GGUGCUGUUCGUGUUAUCGGCUACGACCGGAAGGCCAGGACCC
ACAACCCCUACAAGCUGAGCAGCACCCUGACAAACAUCUACACC
GGCAGCAGGCUGCACGAGGCCGGCUGCGCCCCAGCUACCACGU
GGUCAGGGGGCAUAUCGCCACCGCCACCGAGGGCGUGAUCAUCA
ACGUCGCCAACAGCAAGGGCCAGCCCGAGGGCGGAGUGUGGGCGG
GCCUGUACAAGAAGUUCGCCGAGAGCUUCGACCUGCAGCCCAU
CGAGGUGGGCAAGGCCAGGCUGGUGAAGGGCGCCGUAAGCACA
UCAUCCACGCCGUGGGCCCCAACUUAACAAGGUGAGCGAGGUG
GAAGGCGACAAGCAGCCUGGCCGAAGCCUACGAGAGCAUCGCCAA
GAUCGUGAACGACAUAACUACAAGAGCGUGGCCAUCCACUGC
UCAGCACGGCAUCUUCAGCGGCAACAAGGACAGGCUGACCCAG
AGCCUGAACCACCGUCACCCGCCUGGACACCACCGAUGCCGA
CGUGGCCAUUCUACUGCAGGGACAAGAAGUGGGGAAUGACCCUGA
AGGAGGCCGUGGCCAGGCGGGAGGCCGUGGAAGAGAUCUGCAUC
AGCGACGACUCCAGCGUGACCGAGCCCGACGCCGACGUGGUGAG
GGUGCACCCCAAGAGCUCCUGGCCGGCAGGAAGGGCUACAGCA
CCAGCGACGGCAAGACCUUCAGCUACCUGGAGGGCACCAAGUUC
CACCAGGCCGUAAGGACAUCGCCGAGAUCAACGCUAUGUGGCC
CGUGGCCACCGAGGCCAACGAGCAGGUGUGCAUGUACAUCUCCUGG
GCGAGAGCAUGUCCAGCAUCAGGAGCAAGUGCCCCGUGGAGGAA
AGCGAGGCCAGCACACCACCCAGCACCCUGCCCUGCCUGUGCAU
CCACGCUAUGACACCCGAGAGGGUGCAGCGGCUGAAGGCCAGCA
GGCCCCGAGCAGAUACCGUGUGCACUCCUCCCCACUGCCCAAG
UACAGGAUACCGGCGUGCAGAAGAUCAGUGCAGCCAGCCCAU
CCUGUUCAGCCCAAAGGUGCCCGCCUACAUCACCCCCAGGAAGU
ACCGUGGAGACCCACCCGUGGACGAGACACCCGAGCCAAGC
GCCGAGAACCAGAGCACCGAGGGCACACCCGAGCAGCCACCCCU
GAUACCCGAGGACGAGACAAGGACCACCGACCCACAGCCCAUCA

UUAUCGAGGAAGAGGAAGAGGACAGCAUCAGCCUGCUGAGCGAC
GGCCCCACCCACCAGGUGCUGCAGGUGGAGGCCGACAUCCACGG
CCCACCCAGCGUGUCCAGCUCCAGCUGGAGCAUCCACACGCCA
GCGACUUCGACGUGGACACGCCUGAGCAUCCUGGACACCCUGGAG
GGCGCCAGCGUGACCUCGCGCCACCAGCGCCGAGACCAACAG
CUACUUCGCCAAGAGCAUUGGAGUUCUGGCCAGGCCCGUGCCAG
CUCCCAGGACCGUGUUCAGGAACCCACCCACCCAGCUCCCAGG
ACCAGGACCCCAAAGCCUGGCUCCAGCAGGGCCUGCAGCAGGAC
CAGCCUGGUGAGCACCCACCCGCGUGAACAGGGUGAUCACCA
GGGAGGAACUGGAGGCCUGACACCCAGCAGGACCCCCAGCAGG
UCCGUGAGCAGGACUAGUCUGGUGUCCAACCCACCCGCGUGAA
CAGGGUGAUCACCAGGGAGGAUUCGAGGCCUUCGUGGCCCAGC
AACAGAGACGGUUCGACGCCGCGCCUACAUCUUCAGCAGCGAC
ACCGGCCAGGGACACCUGCAGCAAAAGAGCGUGAGGCAGACCGU
GCUGAGCGAGGUGGUGCUGGAGAGGACCGAGCUGGAAAUCAGCU
ACGCCCCCAGGCUGGACCAGGAGAAGGAGGAACUGCUCAGGAAG
AAACUGCAGCUGAACCCACCCAGCCAACAGGAGCAGGUACCA
GAGCAGGAAGGUGGAGAACAUGAAGGCCAUCACCGCCAGGCGGA
UCCUGCAGGGCCUGGGACACUACCUGAAGGCCGAGGGCAAGGUG
GAGUGCUCAGGACCCUGCACCCCGUGCCACUGUACAGCUCCAG
CGUGAACAGGGCCCUUCUCCAGCCCCAAGGUGGCCGUGGAGGCCU
GCAACGCUAUGCUGAAGGAGAACUCCCCACCGUGGCCAGCUAC
UGCAUCAUCCCCGAGUACGACGCCUACCUGGACAUGGUGGACGG
CGCCAGCUGCUGCCUGGACACCGCCAGCUUCUGCCCCGCCAAGC
UGAGGAGCUUCCCCAAGAAACACAGCUACCUGGAGCCCACCAUC
AGGAGCGCCGUGCCCAGCGCAUCCAGAACACCCUGCAGAACGU
GCUGGCCCGCUGCCACCAAGAGGAACUGCAACGUGACCCAGAUGA
GGGAGCUGCCCGUGCUGGACAGCGCUGCCUUAACGUGGAGUGC
UUCAAGAAAUACGCCUGCAACAACGAGUACUGGGAGACCUUCA
GGAGAACCCCAUCAGGCUGACCGAAGAGAACGUGGUGAACUACA
UCACCAAGCUGAAGGGCCCCAAGGCCGUGCCUGUUCGCUAAG
ACCCACAACCUGAACAUUGCUGCAGGACAUCCAAUGGACAGGUU
CGUGAUGGACCCUGAAGAGGGACGUGAAGGUGACACCCGGCACCA
AGCACACCGAGGAGAGGCCCAAGGUGCAGGUGAUCCAGGCCGCU
GACCCACUGGCCACCGCCUACCUGUGCGGCAUCCACAGGGAGCU
GGUGAGGGCGGCUGAACGCCGUGCUGCUGCCCAACAUCCACACCC
UGUUCGACAUGAGCGCCGAGGACUUCGACGCCAUCAUCGCCGAG
CACUUCAGCCCGGCGACUGCGUGCUGGAGACCGACAUCGCCAG
CUUCGACAAGAGCGAGGAUGACGCUAUGGCCUGACCGCUCUGA
UGAUCCUGGAGGACCUGGGCGUGGACGCCGAGCUGCUCACCCUG
AUCGAGGCUGCCUUCGGCGAGAUACGCUCCAUCCACCGCCAC
CAAGACCAAGUUCAAGUUCGGCGCUAUGAUGAAAAGCGGAAUGU
UCCUGACCCUGUUCGUGAACACCGUGAUAACAUAUGUGAUCGCC
AGCAGGGUGCUGCGGGAGAGGCUGACCGGCAGCCCCUGCGCUGC

GACGAGACCGUGCUGACCGCCAGGAACAUCGGCCUGUGCAGCAGCG
ACGUGAUGGAGAGGAGCCGGAGGGCAUGAGCAUCCUGAGGAAGAA
AUACCUGAAGCCCAGCAACAACGUGCUGUUCAGCGUGGGCAGCACC
AUCUACCACGAGAAGAGGGACCUUGCUCAGGAGCUGGCACCUGCCA
GCGUGUUCACCCUGAGGGGCAAGCAGAACUACACCUGCAGGUGCGA
GACCAUCGUGAGCUGCGACGGCUACGUGGUGAAGAGGAUCGCCAUC
AGCCCCGGCCUGUACGCAAGCCCAGCGGCUACGCCGCUACAAUGC
ACAGGGAGGGCUUCCUGUGCUGCAAGGUGACCGACCCUGAACGG
CGAGAGGGUGAGCUUCCCCGUGUGCACCUACGUGCCCGCCACCCUG
UGCGACCAGAUGACCGGCAUCCUGGCCACCGACGUGAGCGCCGACG
ACGCCCAGAAGCUGCUCGUGGGCCUGAACAGAGGAUCGUGGUCAA
CGGCAGGACCCAGAGGAACACCAACACAUGAAGAACUACCUGCUG
CCCGUGGUGGGCCAGGCUUUCGCCAGGUGGGCCAAGGAGUACAAGG
AGGACCAGGAAGACGAGAGGCCCCUGGGCCUGAGGGACAGGCAGCU
GGUGAUGGGCUGCUCGUGGGCCUUCAGGCGGCACAAGAUCACCAGC
AUCUACAAGAGGGCCGACACCCAGACCAUCAUCAAGGUGAACAGCG
ACUUCCACAGCUUCGUGCUGCCCAGGAUCGGCAGCAACACCCUGGA
GAUCGGCCUGAGGACCCGGAUCAGGAAGAUGCUGGAGGAACACAAG
GAGCCCAGCCCACUGAUCACCGCCGAGGACGUGCAGGAGGCCAAGU
GCGCUGCCGACGAGGCCAAGGAGGUGAGGGAGGCCGAGGAACUGAG
GGCCGCCCUGCCACCCUGGCUGCCGACGUGGAACCCACCCUG
GAAGCCGACGUGGACCUGAUGCUGCAGGAGGCCGGCGCCGGAAGCG
UGGAGACACCCAGGGGCCUGAUCAAGGGUGACCAGCUACGACGGCGA
GGACAAGAUCGGCAGCUACGCCGUGCUGAGCCCACAGGCCGUGCUG
AAGUCCGAGAAGCUGAGCUGCAUCCACCCACUGGCCGAGCAGGUGA
UCGUGAUCACCCACAGCGGCAGGAAGGGCAGGUACGCCGUGGAGCC
CUACCACGGCAAGGUGGUCGUGCCGAGGGCCACGCCAUCCCCGUG
CAGGACUUCAGGCCCCUGAGCGAGAGCGCCACCAUCGUGUACAACG
AGAGGGAGUUCGUGAACAGGUACCUGCACCAUAUCGCCACCCACGG
CGGAGCCUGAACACCCGACGAGGAAUACUACAAGACCGUGAAGCCC
AGCGAGCACGACGGCGAGUACCUGUACGACAUCGACAGGAAGCAGU
GCGUGAAGAAAGAGCUGGUGACCGGCCUGGGACUGACCGGCGAGCU
GGUGGACCCACCCUUCACGAGUUCGCCUACGAGAGCCUGAGGACC
AGACCCGCCCUCUCCUACCAGGUGCCACCAUCGGCGUGUACGGCG
UGCCCCGGCAGCGAAAGAGCGGCAUCAUCAAGAGCGCCGUGACCAA
GAAAGACCUGGUGGUCAGCGCCAAGAAAGAGAACUGCGCCGAGAUC
AUCAGGGACGUGAAGAAGAUGAAAGGCCUGGACGUGAACGCGCGCA
CCGUGGACAGCGUGCUGUAACGGCUGCAAGCACCCCGUGGAGAC
CCUGUACAUCGACGAGGCCUUCGCUUGCCACGCCGGCACCCUGAGG
GCCUGAUCGCCAUCAUACAGGCCCAAGAAAGCCGUGCUGUGCGGCG
ACCCAAGCAGUGCGGCUUCUUAACAUGAUGUGCCUGAAGGUGCA
CUUCAACCACGAGAUUCGACCCAGGUGUUCACAAGAGCAUCAGC
AGGCGGUGCACCAAGAGCGUGACCAGCGUCGUGAGCACCCUGUUCU
ACGACAAGAAAUGAGGACCACCAACCCCAAGGAGACCAAAAUCGU

GAUCGACACCACAGGCAGCACCAAGCCCAAGCAGGACGACCUGAUC
CUGACCUGCUUCAGGGGCGUGGGUGAAGCAGCUGCAGAUCGACUACA
AGGGCAACGAGAUCAUGACCGCCGUGCCAGCCAGGGCCUGACCAG
GAAGGGCGUGUACGCCGUGAGGUACAAGGUGAACGAGAACCCACUG
UACGCUCCCACCAGCGAGCACGUGAACGUGCUGCUGACCAGGACCG
AGGACAGGAUCGUGUGGAAGACCCUGGCCGGCGACCCCUUGGAUCAA
GACCCUGACCGCCAAGUACCCCGCAACUUCACCGCCACCAUCGAA
GAGUGGCAGGCCGAGCACGACGCCAUCAUGGAGGCAAUCCUGGAGA
GGCCCCACCCACCGACGUGUCCAGAACAAGGCCAACGUGUGCUG
GGCCAAGGCCUGGUGCCCUGCUGAAGACCGCCGGAUCGACAUG
ACCACAGAGCAGUGGAACACCGUGGACUACUUCGAGACCGACAAGG
CCCACAGCGCCGAGAUUCGUGCUGAACCAGCUGUGCGUGAGGUUCU
CGGCCUGGACCUGGACAGCGGCCUGUUCAGCGCCCCACCGUGCCA
CUGAGCAUCAGGAACAACCACUGGGACAACAGCCCCAGCCCAAACA
UGUACGGCCUGAACAAGGAGGUGGUCAGGCAGCUGAGCAGGCGGUA
CCCACAGCUGCCCAGGGCCGUGGCCACCGGCAGGGUGUACGACAUG
AACACCGGCACCCUGAGGAACUACGACCCCAAGAUCAACCUGGUGC
CCGUGAACAGGCGGCUGCCCCACGCCUGGUGCUGCACCACAACGA
GCACCCACAGAGCGACUUCAGCUCCUUCGUGAGCAAGCUGAAAGGC
AGGACCGUGCUGGUCGUGGGCGAGAAGCUGAGCGUGCCCGCAAGA
UGGUGGACUGGCUGAGCGACAGGCCCGAGGCCACCUUCCGGGCCAG
GCUGGACCUCGGCAUCCCCGGCGACGUGCCCAAGUACGACAUCAUC
UUCGUGAACGUCAGGACCCAUACAAGUACCACCAUUAACCAGCAGU
GCGAGGACCACGCCAUCAAGCUGAGCAUGCUGACCAAGAAGGCCUG
CCUGCACCUGAACCCCGAGGCACCUUGCGUGAGCAUCGGCUACGGC
UACGCCGACAGGGCCAGCGAGCAUCAUUGGCGCCAUCGCCAGGC
UGUUCAAGUUCAGCAGGGUGUGCAAACCCAAGAGCAGCCUGGAGGA
AACCGAGGUGCUGUUCGUGUUAUCGGCUACGACCGGAAGGCCAGG
ACCCACAACCCUACAAGCUGAGCAGCACCCUGACAAACAUCUACA
CCGGCAGGCUGCACGAGGCCGGCUGCGCCCCCAGCUACCACGU
GGUCAGGGGGCGAUUUCGCCACCGCCACCGAGGGCGUGAUCAUCAA
GCUGCCAACAGCAAGGGCCAGCCCGGAGGCGGAGUGUGCGGGCGCC
UGUACAAGAAGUUCCCCGAGAGCUUCGACCUGCAGCCAUUCGAGGU
GGGCAAGGCCAGGCUGGUGAAGGGCGCCGUAAGCACAUCAUCCAC
GCCUGGGCCCAACUUAACAAGGUGAGCGAGGUGGAAGGCGACA
AGCAGCUGGGCCGAAGCCUACGAGAGCAUCGCCAAGAUCGUGAACGA
CAAUAACUACAAGAGCGUGGCCAUCCCACUGCUCAGCACCGGCAUC
UUCAGCGGCAACAAGGACAGGCUGACCCAGAGCCUGAACCCUGC
UCACCGCCCUGGACACCACCGAUGCCGACGUGGCCAUUCUACUGCAG
GGACAAGAAGUGGGAGAUAGCCCUGAAGGAGGCCGUGGCCAGGCGG
GAGGCCGUGGAAGAGAUCUGCAUCAGCGACGACUCCAGCGUGACCG
AGCCCAGCGCCGAGCUGGUGAGGGUGCACCCCAAGAGCUCCCUGGC
CGGCAGGAAGGGCUACAGCACCGACGGCAAGACCUUCAGCUAC
CUGGAGGGGCACCAAGUUCACCAGGCCGUAAGGACAUCGCCGAGA

UCAACGCUAUGUGGCCCGUGGCCACCGAGGCCAACGAGCAGGUGUG
CAUGUACAUCUGGGCGAGAGCAUGUCCAGCAUCAGGAGCAAGUGC
CCCUGGAGGAAAGCGAGGCCAGCACACCACCCAGCACCCUGCCCU
GCCUGUGCAUCCACGCUAUGACACCCGAGGGUGCAGCGGCUGAA
GGCCAGCAGGCCCGAGCAGAUACCGUGUGCAGCUCCUUCCCACUG
CCAAGUACAGGAUACCGGCGUGCAGAAGAUCCAGUGCAGCCAGC
CCAUCCUGUUCAGCCCAAAGGUGCCCGCCUACAUCCACCCCAGGAA
GUACCUGGUGGAGACCCACCCGUGGACGAGACACCCGAGCCAAGC
GCCGAGAACCAGAGCACCGAGGGCACACCCGAGCAGCCACCCUGA
UCACCGAGGACGAGACAAGGAACCCGGACCCCAGAGCCCAUCAUUAU
CGAGGAAGAGGAAGAGGACAGCAUCAGCCUGCUGACGACGGCCCC
ACCCACCAGGUGCUGCAGGUGGAGGCCGACAUCCACGGCCCACCCA
GCGUGUCCAGCUCCAGCUGGAGCAUCCCACACGCCAGCGACUUCGA
CGUGGACAGCCUGAGCAUCCUGGACACCCUGGAGGGCGCCAGCGUG
ACCUCCGCGCCACCAGCGCCGAGACCAACAGCUACUUCGCCAAGA
GCAUGGAGUUCUGGCCAGGCCGUGCCAGCUCCAGGACCGUGUU
CAGGAACCCACCCACCCAGCUCCAGGACCAGGACCCCAAGCCUG
GCUCCAGCAGGGCCUGCAGCAGGACCAGCCUGGUGAGCACCCAC
CCGGCGUGAACAGGGUGAUCACCAGGGAGGAACUGGAGGCCUGAC
ACCCAGCAGGACCCCAAGCAGGUCCGUGAGCAGGACUAGUCUGGUG
UCCAACCCACCCGGCGUGAACAGGGUGAUCACCAGGGAGGAAUUCG
AGGCCUUCGUGGCCAGCAACAGAGACGGUUCGACGCCGGCGCCUA
CAUCUUCAGCAGCGACACCGGCCAGGGACACCUGCAGCAAAAGAGC
GUGAGGCAGACCGUGCUGAGCGAGGUGGUGCUGGAGAGGACCGAGC
UGGAAAUCAGCUACGCCCCAGGCUGGACCAGGAGAAGGAGGAACU
GCUCAGGAAGAAACUGCAGCUGAACCCCAACCCAGCCAACAGGAGC
AGGUACCAGAGCAGGAAGGUGGAGAACAUGAAGGCCAUCACCGCCA
GGCGGAUCCUGCAGGGCCUGGGACACUACCUGAAGGCCGAGGGCAA
GGUGGAGUGCUACAGGACCCUGCACCCCGUGCCACUGUACAGCUCC
AGCGUGAACAGGGCCUUCUCCAGCCCCAAGGUGGCCGUGGAGGCCU
GCAACGCUAUGCUGAAGGAGAACUCCCCACCGUGGCCAGCUACUG
CAUCAUCCCCGAGUACGACGCCUACCUGGACAUGGUGGACGGCGCC
AGCUGCUGCCUGGACACCGCCAGCUUCUGCCCCGCCAAGCUGAGGA
GCUUCCCCAAGAAACACAGCUACCUGGAGCCCACCAUCAGGAGCGC
CGUGCCCAGCGCCAUCCAGAACACCCUGCAGAACGUGCUGGCCCGCU
GCCACCAAGAGGAACUGCAACGUGACCCAGAUGAGGGAGCUUGCCCCG
UGCUGGACAGCGCUGCCUUAACGUGGAGUGCUUCAAGAAAUAACGC
CUGCAACAACGAGUACUGGGAGACCUUCAAGGAGAACCCCAUCAGG
CUGACCGAAGAGAACGUGGUGAACUACAUCACCAAGCUGAAGGGCC
CCAAGGCCGUGCCUGUUCGCUAAGACCCACAACCUGAACAUGCU
GCAGGACAUCCAAUGGACAGGUUCGUGAUGGACCUGAAGAGGGAC
GUGAAGGUGACACCCGGCACCAAGCACACCGAGGAGGCCCAAGG
UGCAGGUGAUCCAGGCCGUGACCCACUGGCCACCGCCUACCUGUG
CGGCAUCCACAGGGAGCUGGUGAGGGCGGUGAACGCCGUGCUGCUG

CCCAACAUCCACACCCUGUUCGACAUGAGCGCCGAGGACUUCGACG
CCAUCAUCGCCGAGCACUUCAGCCC GCGACUGCGUGCUGGAGAC
CGACAUCGCCAGCUUCGACAAGAGCGAGGAUGACGCUAUGGCCUG
ACCGCUCUGAUGAUCCUGGAGGACCUGGGCGUGGACGCCGAGCUGC
UCACCCUGAUCGAGGCUGCCUUCGGCGAGAUCAUCUCCAUCCACCU
GCCCACCAAGACCAAGUUCAAGUUCGGCGCUAUGAUGAAAAGCGGA
AUGUUCUGACCCUGUUCGUGAACACCGUGAUCAACAUGUGAUCG
CCAGCAGGGUGCUGCGGGAGAGGCUGACCGGCAGCCCCUGCGCUGC
CUUCAUCGGCGACGACAACAUCGUGAAGGGCGUGAAAAGCGACAAG
CUGAUGGCCGACAGGUGCGCCACCUGGCUGAACAUUGGAGGUGAAGA
UCAUCGACGCCGUGGUGGGCGAGAAGGCCCCUACUUCUGCGGCGG
AUUCAUCCUGUGCGACAGCGUGACCGGCACCGCCUGCAGGGUGGCC
GACCCCUGAAGAGGCUGUUAAGCUGGGCAAGCCACUGGGCCGUG
ACGAUGAGCACGACGAUGACAGGCGGAGGGCCUGCACGAGGAAAG
CACCAGGUGGAACAGGGUGGGCAUCCUGAGCGAGCUGUGCAAGGCC
GUGGAGAGCAGGUACGAGACCGUGGGCACCAGCAUCAUCGUGAUGG
CUAUGACCACACUGGCCAGCUCGGUCAAGAGCUUCUCCUACCUGAG
GGGGGCCCUAUAAUCUCUAACGGCUAACCUGAAUGGACUACGACA
UAGUCUAGUCCGCCAAGGCCGCCACCAUGAGAGUGACAGCCCUAG
AACCUUACUGCUUCUGCUUUGGGGAGCUGUUGCUCUGACAGAGACA
UGGGCUGGAUCUCUGAGCGAGGUGACCGGCCAGGGCCUGUGCAUCG
GCGCCUGCCCCAAGACCCACCAGGUGCUGUGCAACACCACCCAGAA
GACCAGCGACGGCAGCUACUACCUGGCCGCCUCCACCAGGCACCACC
UGGCCUGCAGCACCGGCCUGACCCCUUGCAUCAGCACCACCAUCC
UGAACCUGACCACCGACUACUGCGUGCUGGUGGAGCUGUGGCCAG
GGUGACCUACCACAGCCCCAGCUACGCCUACCACCAGUUCGAGAGG
AGGGCCAAGUACAAGAGGGAGCCCUGAGCCUGACCCUGGCCUGC
UGCUGGGCGGCCUGACAAUGGGCGGCAUCGCCGCCGGCGUGGGCAC
CGGCACCACCGCCUGGUGGCCACCAGCAGUUCAGCAGCUGCAG
GCCGCCAUGCACGACGACCUGAAGGAGGUGGAGAAGUCCAUCACCA
ACCUGGAGAAGUCCUGACCAGCCUGAGCGAGGUGGUGCUGCAGAA
CAGGAGGGCCUGGACCUGCUGUCCUGAAGGAGGGCGGCCUGUGC
GCCC GCCUGAAGGAGGAGUGCUGCCUGUACGCCGACCACACCGGCC
UGGUGAUCGUGGGCAUUGUCGUGGCCUGGCCGUCCUGCCGUGGU
GGUGAUUGGAGCUGUGGUCGAGCUGUUAUGUGCAGAAGAAAGUCA
UCCGGCGGAAAAGGGAGGCUCUACUCUCAGGCUGCUUCUGCUACAG
UGCCUAGAGCUCUUAUGUGUUAUCUCAGCUGGGCGGCGGAGGCAG
CGACUACAAGGACGACGAUGACAAGUAAACUCGAGUAUGUUACGUG
CAAAGGUGAUUGUACCCCCGAAAGACCAUAUUGUGACACACCCU
CAGUAUCACGCCAAACAUUUACAGCCGCGGUGUAAAAACCGCGU
GGACGUGGUUAAAUCCUGCUGGGAGGAUCAGCCGUAUUUAUUUAU
AAUUGGCUUGGUGCUGGCCUACUUAUGUGGCCAUGUACGUGCUGACC
AACCAGAAACAUAUUGAAUACAGCAGCAAUUGGCAAGCUGCUUAC
AUAGAACUCGCGGCGAUUGGCAUGCCGCCUAAAAUUUUUUUUUA

CUACCACGGCAAGGUGGUCGUGCCGAGGGCCACGCCAUCCCCGUG
CAGGACUUCAGGCCUGAGCGAGAGCGCCACCAUCGUGUACAACG
AGAGGGAGUUCGUGAACAGGUACCUGCACCAUAUCGCCACCCACGG
CGGAGCCCUGAACACCGACGAGGAAUACUACAAGACCGUGAAGCCC
AGCGAGCACGACGGCGAGUACCUUGUACGACAUCGACAGGAAGCAGU
GCGUGAAGAAAAGAGCUGGUGACCGGCCUGGGACUGACCGGCGAGCU
GGUGGACCCACCCUUCACGAGUUCGCCUACGAGAGCCUGAGGACC
AGACCCGCCCGCUCCCUACCAGGUGCCCACCAUCGGCGUGUACGGCG
UGCCCCGGCAGCGGAAAAGAGCGGCAUCAUCAAGAGCGCCGUGACCAA
GAAAGACCUUGGUGGUCAGCGCCAAGAAAAGAGAACUGCGCCGAGAUC
AUCAGGGACGUGAAGAAGAUGAAAGGCCUGGACGUGAACGCGCGCA
CCGUGGACAGCGUGCUGCUGAACGGCUGCAAGCACCCCGUGGAGAC
CCUGUACAUCGACGAGGCCUUCGCUUGCCACGCCGGCACCCUGAGG
GCCUGAUCGCCAUCAUCAGGCCCAAGAAAGCCGUGCUGUGCGGCG
ACCCAAAGCAGUGCGGCUUCUUAACAUGAUGUGCCUGAAGGUGCA
CUUCAACCACGAGAUUCGACCCAGGUGUUCACAAGAGCAUCAGC
AGGCGGUGCACCAAGAGCGUGACCAGCGUCGUGAGCACCCUGUUCU
ACGACAAGAAAUGAGGACCACCAACCCCAAGGAGACCAAAAUCGU
GAUCGACACCACAGGCAGCACCAAGCCCAAGCAGGACGACCUGAUC
CUGACCUUCUUCAGGGGCGUGGUGAAGCAGCUGCAGAUCGACUACA
AGGGCAACGAGAUCAUGACCGCCGUGCCAGCCAGGGCCUGACCAG
GAAGGGCGUGUACGCCGUGAGGUACAAGGUGAACGAGAACCACUG
UACGCUCCACCAGCGAGCACGUGAACGUGCUGCUGACCAGGACCG
AGGACAGGAUCGUGUGGAAGACCCUGGCCGGCGACCCUGGAUCAA
GACCCUGACCGCCAAGUACCCCGGCAACUUCACCGCCACCAUCGAA
GAGUGGCAGGCCGAGCACGACGCCAUCAUGGAGGCAAUCCUGGAGA
GGCCCCACCCACCAGCUGUUCAGAAACAAGGCCAACGUGUGCUG
GGCCAAGGCCUGGUGCCCUGCUGAAGACCGCCGGCAUCGACAUG
ACCACAGAGCAGUGGAACACCGUGGACUACUUCGAGACCGACAAGG
CCCACAGCGCCGAGAUCGUGCUGAACCCAGCUGUGCGUGAGGUUCU
CGGCCUGGACCUGGACAGCGGCCUGUUCAGCGCCCCACCGUGCCA
CUGAGCAUCAGGAACAACCACUGGGACAACAGCCCCAGCCCAAACA
UGUACGGCCUGAACAAGGAGGUGGUCAGGCAGCUGAGCAGGCGGUA
CCCACAGCUGCCCAGGGCCGUGGCCACCGGCAGGGUGUACGACAUG
AACACCGGCACCCUGAGGAACUACGACCCCAGGAUCAACCGGUGC
CCGUGAACAGGCGGCGUCCCCACGCCUGGUGCUGCACCACAACGA
GCACCCACAGAGCGACUUCAGCUCCUUCGUGAGCAAGCUGAAAGGC
AGGACCGUGCUGGUCGUGGGCGAGAAGCUGAGCGUGCCCGGCAAGA
UGGUGGACUGGCUGAGCGACAGGCCCGAGGCCACCUCCGGGCCAG
GCUGGACCUCCGCAUCCCCGGCGACGUGCCCAAGUACGACAUCAUC
UUCGUGAACGUCAGGACCCAUACAAGUACCACCAUUACCAGCAGU
GCGAGGACCACGCCAUCAAGCUGAGCAUGCUGACCAAGAAGGCCUG
CCUGCACCUGAACCCCGGAGGCACCUUGCGUGAGCAUCGGCUACGGC
UACGCCGACAGGGCCAGCGAGCAUCAUUGGCGCCAUCGCCAGGC

UGUUCAAGUUCAGCAGGGUGUGCAAACCCAAGAGCAGCCUGGAGGA
AACCGAGGUGCUGUUCGUGUUAUCGGCUACGACCGGAAGGCCAGG
ACCCACAACCCCUACAAGCUGAGCAGCACCCUGACAAACAUCUACA
CCGGCAGGCUGCAGGAGCCGGCUGCGCCCCAGCUACCACGU
GGUCAGGGGCGAUUUCGCCACCGCCACCGAGGGCGUGAUCAUCAA
GCUGCCAACAGCAAGGGCCAGCCCGGAGGGGAGUGUGCGGGCCCC
UGUACAAGAAGUUCGAGAGCUUCGACCUGCAGCCAUCGAGGU
GGGCAAGGCCAGGCUGGUGAAGGGCGCCGCUAAGCACAUCAUCCAC
GCCGUGGGCCCCAACUUAACAAGGUGAGCGAGGUGGAAGGCGACA
AGCAGCUGGCCGAAGCCUACGAGAGCAUCGCCAAGAUCGUGAACGA
CAUAACUACAAGAGCGUGGCCAUCCACUGCUCAGCACCGGCAUC
UUCAGCGGCAACAAGGACAGGCUGACCCAGAGCCUGAACACCUCG
UCACCGCCUGGACACCACCGAUGCCGACGUGGCCAUACUGCAG
GGACAAGAAGUGGGAGAUGACCCUGAAGGAGGCCGUGGCCAGGCCG
GAGGCCGUGGAAGAGAUCUGCAUCAGCGACGACUCCAGCGUGACCG
AGCCCCGACGCCGAGCUGGUGAGGGUGCACCCCAAGAGCUCCCUGGC
CGGCAGGAAGGGCUACAGCACCGACGGCAAGACCUUCAGCUAC
CUGGAGGGGCACCAAGUUCACCAGGCCGCUAAGGACAUCGCCGAGA
UCAACGCUAUGUGGCCCGUGGCCACCGAGGCCAACGAGCAGGUGUG
CAUGUACAUCUGGGCGAGAGCAUGUCCAGCAUCAGGAGCAAGUGC
CCCUGGAGGAAAGCGAGGCCAGCACACCACCCAGCACCCUGCCCU
GCCUGUGCAUCCACGCUAUGACACCCGAGGGUGCAGCGGCUGAA
GGCCAGCAGGCCCGAGCAGAUACCGUGUGCAGCUCCUCCCACUG
CCCAAGUACAGGAUACCGGCGUGCAGAAGAUCCAGUGCAGCCAGC
CCAUCCUGUUCAGCCCAAAGGUGCCCGCCUACAUCCACCCAGGAA
GUACCGUGGAGACCCACCCGUGGACGAGACACCCGAGCCAAGC
GCCGAGAACCAGAGCACCGAGGGCACACCCGAGCAGCCACCCUGA
UCACCGAGGACGAGACAAGGAACCCGGACCCCAAGGACCAUCAUUAU
CGAGGAAGAGGAAGAGGACAGCAUCAGCCUGCUGACGACGGCCCC
ACCCACAGGUGCUGCAGGUGGAGGCCGACAUCCACGGCCCCACCA
GCGUGUCCAGCUCCAGCUGGAGCAUCCACACGCCAGCGACUUCGA
CGUGGACAGCCUGAGCAUCCUGGACACCCUGGAGGGCGCCAGCGUG
ACCUCCGCGCCACCAGCGCCGAGACCAACAGCUACUUCGCCAAGA
GCAUGGAGUUCUGGCCAGGCCCGUGCCAGCUCCAGGACCGUGUU
CAGGAACCCACCCACCCAGCUCCAGGACCAGGACCCCAAGCCUG
GCUCCAGCAGGGCCUGCAGCAGGACCAGCCUGGUGAGCACCCAC
CCGGCGUGAACAGGGUGAUCACCAGGGAGGAACUGGAGGCCUGAC
ACCCAGCAGGACCCCAAGCAGGUCGUGAGCAGGACUAGUCUGGUG
UCCAACCCACCCGGCGUGAACAGGGUGAUCACCAGGGAGGAAUUCG
AGGCCUUCGUGGCCAGCAACAGAGACGGUUCGACGCCGGCGCCUA
CAUCUUCAGCAGCGACACCGGCCAGGGACACCUAGCAGCAAAAGAGC
GUGAGGCAGACCGUGCUGAGCGAGGUGGUGCUGGAGAGGACCGAGC
UGGAAAUCAGCUACGCCCCAGGCUGGACCAGGAGAAGGAGGAACU
GCUCAGGAAGAAACUGCAGCUGAACCCACCCAGCCAACAGGAGC

AGGUACCAGAGCAGGAAGGUGGAGAACAUGAAGGCCAUCACCGCCA
GGCGGAUCCUGCAGGGCCUGGGACACUACCUGAAGGCCGAGGGCAA
GGUGGAGUGCUACAGGACCCUGCACCCCGUGCCACUGUACAGCUCC
AGCGUGAACAGGGCCUUCUCCAGCCCCAAGGUGGCCGUGGAGGCCU
GCAACGCUAUGCUGAAGGAGAACUCCCCACCGUGGCCAGCUACUG
CAUCAUCCCCGAGUACGACGCCUACCUGGACAUGGUGGACGGCGCC
AGCUGCUGCCUGGACACCGCCAGCUUCUGCCCCGCCAAGCUGAGGA
GCUUCCCCAAGAAACACAGCUACCUGGAGCCCACCAUCAGGAGCGC
CGUGCCCAGCGCCAUCCAGAACACCCUGCAGAACGUGCUGGCCGCU
GCCACCAAGAGGAACUGCAACGUGACCCAGAUGAGGGAGCUUCCCCG
UGCUGGACAGCGCUGCCUUAACGUGGAGUGCUUCAAGAAAUAACGC
CUGCAACAACGAGUACUGGGAGACCUUCAAGGAGAACCCCAUCAGG
CUGACCGAAGAGAACGUGGUGAACUACAUCACCAAGCUGAAGGGCC
CCAAGGCCGUGCCCUGUUCGCUAAGACCCACAACCUGAACAUAGCU
GCAGGACAUCCAAUGGACAGGUUCGUGAUGGACCUGAAGAGGGAC
GUGAAGGUGACACCCGGCACCAAGCACACCGAGGAGGCCCAAGG
UGCAGGUGAUCCAGGCCGUGACCCACUGGCCACCGCCUACCUGUG
CGGCAUCCACAGGGAGCUGGUGAGGGCGGCUAACGCCGUGCUGCUG
CCCAACAUCCACACCCUGUUCGACAUGAGCGCCGAGGACUUCGACG
CCAUCAUCGCCGAGCACUUCAGCCCCGGCGACUGCGUGCUGGAGAC
CGACAUCGCCAGCUUCGACAAGAGCGAGGAUGACGCUAUGGCCUG
ACCGCUCUGAUGAUCCUGGAGGACCUGGGCGUGGACGCCGAGCUGC
UCACCCUGAUCGAGGCUGCCUUCGGCGAGAUCAUCUCCAUCCACCU
GCCACCAAGACCAAGUUCAGUUCGGCGCUAUGAUGAAAAGCGGA
AUGUUCUGACCCUGUUCGUGAACACCGUGAUCAACAUUGUGAUCG
CCAGCAGGGUGCUGCGGGAGAGGCUGACCGGCAGCCCCUGCGCUGC
CUUCAUCGGCGACGACAACUUCGUGAAGGGCGUGAAAAGCGACAAG
CUGAUGGCCGACAGGUGCGCCACCUGGCUGAACAUUGGAGGUGAAGA
UCAUCGACGCCGUGGUGGGCGAGAAGGCCCCUACUUCUGCGGCGG
AUUCAUCCUGUGCGACAGCGUGACCGGCACCGCCUGCAGGGUGGCC
GACCCCUGAAGAGGCUGUUAAGCUGGGCAAGCCACUGGGCCGUG
ACGAUGAGCACGACGAUGACAGGCGGAGGGCCUGCACGAGGAAAG
CACCAGGUGGAACAGGGUGGGCAUCCUGAGCGAGCUGUGCAAGGCC
GUGGAGAGCAGGUACGAGACCGUGGGCACCAGCAUCAUCGUGAUGG
CUAUGACCACACUGGCCAGCUCCGUCAAGAGCUUCUCCUACCUGAG
GGGGGCCCCUAUAACUCUCUAACGGCUAACCUGAAUGGACUACGACA
UAGUCUAGUCCGCAAGGCCGCCACCAUGAGAGUGACAGCCCCUAG
AACCUUACUGCUUCUGCUUUGGGGAGCUGUUGCUCUGACAGAGACA
UGGGCUGGAUCUUAACACAGCCCCAGCUACGCCUACCACCAGUUCG
AGAGGGGGGGGAGGAGGCUCGGGGGAGGAGGCUCCUGAAGAUCAG
CCAGGCCGUGCACGCCGCCACGCCGAGAUCAACGAGGCCGGCCGG
GAGGUGAUCGUGGGCAUUGUCGUGGCCUGGCCGUCUCCGCGGUGG
UGGUGAUUGGAGCUGUGGUCGAGCUGUUAUGUGCAGAAGAAAGUC
AUCCGGCGGAAAGGGAGGCUCUACUCUCAGGCUGCUUCUGCUACA

GUGCCUAGAGCUCUUAUGUGUUUAUCUCAGCUGGGCGGCGGAGGCA
GCGACUACAAGGACGACGAUGACAAGUAAACUCGAGUAUGUUACGU
GCAAAGGUGAUUGUCACCCCCGAAAGACCAUAUUGUGACACACC
UCAGUAUCACGCCCAAACAUUUACAGCCGCGGUGUCAAAAACCGCG
UGGACGUGGUUAAACAUCCCUGCUGGGAGGAUCAGCCGUAUUUAUUA
UAAUUGGCUUGGUGCUGGCUACUAUUGUGGGCCAUGUACGUGCUGAC
CAACCAGAAACAUAAUUGAAUACAGCAGCAAUUGGCAAGCUGCUUA
CAUAGAACUCGCGGCGAUUGGCAUGCCGCCUAAAAUUUUUAAAAUU
AUUUUUUCUUUUUUUUUCCGAAUCGGAUUUUGUUUUUAAUAUUUCA
AA
AA
AA proteina non strutturale di SINV Marcia #
2842 Sinv. MEKPVVNVDPQSPFVVLQKSFQFEVVAQQVTPNDHANARAFS E nsP1-
HLASKLIELEVPTTATILDIGSAPARRMFSEHQYHCVCMPMRSPEDP 2862 4 anni fa
DRMMKYASKLAEKACKITNKNLHEKIKDLRTVLDTPDAETPSLCFH (SEGUENTE
NDVTCNMRAEYSVMQDVYINAPGTIYHQAMKGVRTLYWIGFDTTQF ID
MFSAMAGSYPAYNTNWADEKVLARNIGLCSTKLSEGRGKLSIMR N.: 109)
KKELKPGSRVYFVSGSTLYPEHRASLQSWHLPSVFHLNGKQSYTCR
CDTVVSCGYVVKKITISPGITGETVGYAVTHNSEGFLCKVTDTV
KGERVSFPVCTYIPATICDQMTGIMATDISPDDAQKLLVGLNQRIV
INGRTNNTNTMQNYLLPIAQGFSKWAKERKDDLDNEKMLGTRER
KLTYGCLWAFRTKKVHSFYRPPGTQTCVKVPASFSAFPMSVWTT
LPMSLRQKLKALQPKKEEKLLQVSEELVMEAKAAFEDAQEEARAE
KLREALPPLVADKGIEMAAEVVCEVEGLQADIGAALVETPRGHVRI
IPQANDRMIGQYIVVSPNSVLKNAKLAPAHPLADQVKIITHSGRSG
RYAVEPYDAKVLMPAGAVPWPEFLALSESATLVYNREFVNRKLY
HIAMHGPAKNTEEEQYKVTKAELAETEYVFDVDKKRCVKKEEASGL
VLSGELTNPPYHELALGLKTRPAVPYKQVETIGVIGTPGSGKSAII
KSTVTARDLVTSGKKENCREIEADVLRLRGMQITSKTVDSVMLNGC
HKAVEVLYVDEAFACHAGALLALIAIVRPRKKVVLGDPMQCGFFN
MMQLKVHFNHPEKDICTKTFYKYISRRCTQPVTAVSTLHYDGKMK
TTNPCKKNIEIDITGATKPKPGDIILTCFRGWVKQLQIDYPGHEVM
TAAASQGLTRKGVYAVRQKVNENPLYAITSEHVNVLLTRTEDRLVW
KTLQGDPWIKQLTNIPKGNFQATIEDWEAEHKGIIAAINSPTPRAN
PFCKTNCVWAKALEPILATAGIVLTGCQWSELFPQFADDKPHSAI
YALDVICIKFFGMDLTSGLFSKQSIPLTYHPADSARPVVAHWNSPG
TRKYGYDHAIAAELSRRFPVFLAGKGTQLDLQTRTRVISAQHNL
VPVNRNLPHALVPEYKEKQPGPVEKFLNQFKHHSVLVSEEKIEAP
RKRIEWIAPIGIAGADKNYNLAFGFPPQARYDLVFINIGTKYRNHH
FQQCEDHAATLKTLSRSALNCLNPGGTLVVKSYGYADRNSDVTVA
LARKFVRVSAARPDCVSSNTEMYLIFRQLDNSRTRQFTPHHLNCVI
SSVYEGTRDGVGAAPSYRTKRENIADCQEEAVVNAANPLGRPGEGV
CRAIYKRWPTSFTDSATETGTARMTVCLGKKVIHAVGPDFRKHPEA
EALKLLQNAYHAVADLVNEHNIKSVAIPLLSTGIYAAGKDRLEVSL

NCLTTALDRTDADVITYCLDKKKKERIDAALQLKESVTELKDEDME
IDDELVWIHPDSCCLKGRKGFSTTKGKLYSYFEGTKFHQAAKDMAEI
KVLFPNDQESNEQLCAYILGETMEAIREKCPVDHNPSSSPKTLPC
LCMYAMTPERVHRLRSNNVKEVTVCSSSTPLPKHKIKNVQKVQCTKV
VLFNPHTPAFVPARKYIEVPEQPTAPPAQAEAEPEVVATPSPSTAD
NTSLDVTDISLDMDDSSEGLFSSFSGSDNSITSMDSWSSGPSSLE
IVDRRQVVVADVHAVQEPAPIPPPRLKKMARLAAARKEPTPPASNS
SESLHLSFGGVMSLGSIFDGETARQAAVQPLATGPTDVPMSFGSF
SDGEIDELSRVTESEPVLFSGFEPGEVNSIISSRSVAVSFPLRKQR
RRRRSRTEY*LTGVGGYIFSTDTTGPGLKQKSVLQNLTEPTLER
NVLERIHAPVLDTSKEEQLKLRQMMPTANKSRYSRKVENQKAI
TTERLLSGLRLYNSATDQPECYKITYPKPLYSSVSPANYSQPFAV
AVCNLYLHENYPTVASYQITDEYDAYLDMVDGTVACLDTATFCPAK
LRSYPKKHEYRAPNIRSAVPSAMQNTLQNVLIAATKRNCNVTQMRE
LPTLDSATFNVECFRKYACNDEYWEEFARKPIRITTEFVTAYVARL
KGPKAAALFAKTYNLVPLQEVPMDFVMDMKRDVKVTPGKHTTEER
PKVQVIQAAEPLATAYICGIHRELVRRLTAVLLPNIHTLFDMSAED
FDAIIAEHFKQGDPVLETDIASFDKSQDDAMALTGLMILEDLGVDDQ
PLLDLIECAFGEISSTHLPTGTRFKFGAMMKSGMFLTLFVNTVLNV
VIASRVLEERLKTSRCAAFIGDDNIIHGVSDEKEMAERCATWLNME
VKIIDAVIGERPPYFCGGFILQDSVTSTACRVADPLKRLFGLGKPL
PADDEQDEDRRRALLDETKAWFRVGITGLAVAVTRYEVDNITPV Domande frequenti sulla
LLALRTSKRAFQAIRGEIKHLYGGPK

Esempio 11

Questo esempio descrive l'analisi dell'immunogenicità dell'emoagglutinina (HA) influenzale espressa da RNA autoreplicante o mRNA.

I costrutti di vaccino a RNA e mRNA autoreplicanti sono stati progettati per codificare la proteina emoagglutinina (HA) a lunghezza intera del virus dell'influenza A/California/07/2009 (H1N1) (SEQ ID NO:113 e 114). Come descritto sopra per l'esempio 1, il costrutto di vaccino a mRNA che codifica per HA includeva un virus dell'incisione del tabacco (TEV) 5' UTR e una beta-globina di Xenopus (Xbg) 3' UTR. Sia l'RNA autoreplicante (SEQ ID NO:56; intero RNA mARM3039) sia i costrutti del vaccino mRNA (SEQ ID NO:116; intera sequenza RNA mARM3038) sono stati incapsulati nella stessa composizione di nanoparticelle lipidiche (LNP) che includeva quattro eccipienti lipidici (un lipide cationico ionizzabile, 1,2-distearoil-sn-glicerolo-3-fosfocolina (DSPC), colesterolo e PEG2000-DMG) dispersi in tampone HEPES (pH 8,0) contenente cloruro di sodio e i crioprotettori saccarosio e glicerolo. Il rapporto N:P del lipide complessante e dell'RNA era di circa 9:1. Il lipide cationico ionizzabile aveva la seguente struttura:

Cinque topi Balb/c femmine di 8-10 settimane sono stati iniettati per via intramuscolare con 2 mg di mRNA o RNA autoreplicante che codifica HA. I topi sono stati dissanguati nei giorni 14, 28, 42 e 56, seguiti da un test di inibizione dell'emoagglutinazione (HAI) utilizzando sieri diluiti in serie. Il

reciproco della più alta diluizione del siero che ha causato l'inibizione dell'emoagglutinazione è stato considerato il titolo HAI, con un titolo di 1/40 protettivo contro l'infezione da virus influenzale e titoli quattro volte superiori al basale che indicano sierconversione.

Risultati in FIGURA 23 mostrano che titoli HAI maggiori sono stati ottenuti con RNA autoreplicante che codifica HA rispetto a mRNA che codifica HA. I titoli HAI per il costrutto di RNA autoreplicante che codifica HA erano maggiori dei titoli HAI per mRNA che codifica HA in tutti i punti temporali a partire dal giorno 28. Inoltre, titoli HAI protettivi sono stati osservati per il costrutto di RNA autoreplicante che codifica HA a partire dal giorno 28 che sono stati mantenuti per almeno 56 giorni. Al contrario, mRNA che codifica HA ha mostrato titoli HAI protettivi solo al giorno 56 che erano inferiori ai titoli HAI osservati per il costrutto di RNA autoreplicante HA. In tutti gli altri punti temporali, i titoli HAI per il costrutto di mRNA che codifica HA erano al di sotto della soglia del titolo protettivo, con un titolo HAI che era paragonabile all'iniezione con controllo PBS al giorno 28.

Questi risultati mostrano che il costrutto di RNA autoreplicante che codifica HA ha suscitato titoli di anticorpi protettivi contro HA, con titoli HAI maggiori rispetto al costrutto di mRNA che codifica HA.

Esempio 12

Questo esempio descrive la produzione di dsRNA e l'espressione della luciferasi per l'RNA autoreplicante.

Diversi sistemi di RNA autoreplicanti da diversi alfavirus sono stati testati per l'espressione in vitro utilizzando la proteina fluorescente verde (GFP) o la luciferasi di lucciola (Luc) come geni reporter. La trasfezione iniziale di cellule con quantità crescenti di RNA autoreplicante ha determinato l'espressione di geni reporter a una dose inferiore rispetto all'mRNA. Tuttavia, all'aumentare della quantità di RNA autoreplicante in ingresso, l'espressione rilevabile del gene reporter è diminuita.

L'RNA autoreplicante produce RNA a doppio filamento (dsRNA) come intermedio nel processo di amplificazione. La sovrapproduzione di dsRNA può sopprimere la traduzione. Per valutare l'effetto della produzione di dsRNA sull'espressione del transgene, sono stati misurati simultaneamente il dsRNA e l'espressione del gene reporter luciferasi. Le cellule HEK293 sono state transfettate con 2 µg di replicone A (SEQ ID NO:115; intera sequenza RNA mARM2826) o replicone B (SEQ ID NO:100, intera sequenza RNA mARM2809) RNA autoreplicante, o mRNA che esprime Luc (SEQ ID NO:102, intera sequenza mRNA mARM1782) utilizzando un reagente di transfettazione dell'RNA commerciale. Le cellule non transfettate (UTC) sono servite come controllo. Produzione di dsRNA (FIGURA 24 UN) è stata quantificata utilizzando la colorazione immunoistochimica per dsRNA, seguita dalla quantificazione della fluorescenza utilizzando uno scanner a fluorescenza 24 ore dopo la trasfezione. L'espressione della luciferasi (FIGURA 24B) è stato analizzato misurando parallelamente la bioluminescenza.

Il replicone A ha prodotto un livello di dsRNA 3 volte superiore rispetto al replicone B 24 ore dopo la trasfezione (FIGURA 24 UN). Tuttavia, il replicone B ha prodotto un livello di espressione della

luciferasi 2,4 volte superiore rispetto al replicone A. Inoltre, il livello di espressione della luciferasi dal replicone A era equivalente a quello osservato per l'mRNA. Pertanto, anche se il replicone A aveva la capacità di amplificare la quantità di RNA del replicone e di mRNA trascritto che codificava la luciferasi, la traduzione dell'mRNA amplificato è stata inibita, in linea con la sovrapproduzione di dsRNA che inibisce la traduzione. Inoltre, sono stati osservati livelli più elevati di espressione genica della luciferasi per l'RNA del replicone rispetto all'mRNA a 24, 48 e 72 ore dopo la trasfezione delle cellule HEK293 (FIGURA 15 UN). L'RNA autoreplicante con una cassetta di espressione che includeva un gene reporter della luciferasi seguito da un IRES e E3L ha anche mostrato una robusta espressione della luciferasi (FIGURE 15 B, 15 C; SEQ ID NO: 128 e 129). L'espressione della luciferasi è stata osservata anche per un RNA autoreplicante che esprimeva E3L da un primo promotore subgenomico e un gene reporter della luciferasi da un secondo promotore subgenomico situato a 3' del frame di lettura aperto E3L (non mostrato). Pertanto, non solo l'RNA del replicone ha prodotto livelli più elevati di espressione del gene della luciferasi rispetto all'mRNA, ma l'RNA del replicone ha anche mostrato una maggiore durata dell'espressione in un periodo di 72 ore.

Esempio 13

Questo esempio descrive l'immunogenicità di formulazioni di RNA autoreplicante liquido e liofilizzato. L'immunogenicità dell'RNA autoreplicante (SEQ ID NO: 125) formulato come nanoparticella lipidica liofilizzata (LYO-LNP) è stata testata su topi BALB/c in due studi preclinici separati e confrontata con la formulazione LNP liquida (congelata) (Liquid-LNP). Ogni studio ha incluso l'uso di un gruppo di dosaggio PBS come controllo negativo e di un gruppo di dosaggio Liquid (Liquid-LNP) come controllo positivo. Entrambe le formulazioni LYO-LNP e Liquid-LNP sono state dosate a 0,2 e 2 µg. C'erano n=5 animali per gruppo di dosaggio in ogni studio. Le formulazioni di prova sono state somministrate per via intramuscolare (IM) e il siero è stato raccolto in vari momenti (giorni 10, 19, 31 per il primo studio e giorni 10, 20, 30 per il secondo studio) dopo l'immunizzazione per misurare la produzione di IgG della proteina spike anti-SARS-CoV-2 utilizzando un test fluorescente a sfere Luminex.

In entrambi gli studi, le IgG della proteina spike anti-SARS-CoV-2 sono state rilevate nel siero in modo dipendente dal tempo e dalla dose per entrambe le formulazioni Liquid-LNP e LYO-LNP, mentre l'iniezione di PBS non ha provocato una risposta immunogenica (FIGURA 16A - 16D). Non è stata osservata alcuna differenza statistica nell'immunogenicità tra i gruppi di dosaggio Liquid-LNP e LYO-LNP nel primo studio, mentre LYO-LNP ha prodotto IgG statisticamente diverse e maggiori rispetto a Liquid-LNP nel secondo studio. Senza essere limitati dalla teoria, la sotto-potenza (n=5/gruppo) di questi due studi separati può aver contribuito alle differenze statistiche nei risultati di immunogenicità osservati nei due studi. Combinando i risultati di entrambi gli studi, non sono state osservate differenze statisticamente significative tra le formulazioni Liquid-LNP e LYO-LNP ai livelli di dose di 0,2 e 2 µg (FIGURA 17A , 17B). Nel complesso, i risultati di questi studi dimostrano che l'immunogenicità delle formulazioni liquide e liofilizzate era comparabile.

In sintesi, le formulazioni liquide e liofilizzate del vaccino a RNA autoreplicante (SEQ ID NO:125) hanno mostrato immunogenicità comparabile. Il vaccino può indurre risposte immunitarie efficaci

e adattative umorali (anticorpi neutralizzanti) e cellulari (CD8+) che prendono di mira la glicoproteina S del SARS-CoV-2. Il vaccino provoca anche l'induzione di livelli di anticorpi anti-glicoproteina spike (IgG) più elevati di un vaccino a mRNA convenzionale e induce anche la produzione di anticorpi IgG a una velocità maggiore rispetto a un vaccino a mRNA convenzionale. Continua a produrre livelli crescenti di IgG fino a 50 giorni dopo la vaccinazione, mentre il vaccino a mRNA convenzionale raggiunge un plateau entro il 10° giorno dopo la vaccinazione. Produce un aumento dose-dipendente dell'RNA nei linfociti T CD8+ e una risposta immunitaria bilanciata delle cellule T helper CD4+ a predominanza Th1 senza alcuna deviazione verso una risposta Th2.

Tutti i riferimenti e le citazioni ad altri documenti, quali brevetti, domande di brevetto, pubblicazioni di brevetti, riviste, libri, documenti, contenuti web, presenti nella presente informativa sono qui incorporati nella loro interezza a tutti gli effetti.

Sebbene l'invenzione sia stata descritta con riferimento agli esempi di cui sopra, si comprenderà che modifiche e varianti sono comprese nello spirito e nell'ambito dell'invenzione. Di conseguenza, l'invenzione è limitata solo dalle seguenti rivendicazioni.

Esempio 14

Liofilizzazione di materiali e metodi di formulazione di nanoparticelle di RNA-lipidi autoreplicanti in generale

I processi condotti in questo esempio sono stati condotti utilizzando composizioni di nanoparticelle lipidiche che sono state prodotte secondo processi ben noti, ad esempio quelli descritti nella domanda statunitense Ser. No. 16/823,212, il cui contenuto è incorporato per riferimento allo scopo specifico di insegnare i processi di produzione di nanoparticelle lipidiche. Le composizioni di nanoparticelle lipidiche e i prodotti liofilizzati sono stati caratterizzati per diverse proprietà. I materiali e i metodi per questi processi di caratterizzazione, nonché un metodo generale di produzione delle composizioni di nanoparticelle lipidiche che sono state utilizzate per gli esperimenti di liofilizzazione sono forniti in questo esempio.

Produzione di nanoparticelle lipidiche

Le formulazioni di nanoparticelle lipidiche utilizzate in questo esempio sono state prodotte mescolando lipidi (lipide cationico ionizzabile (ATX-126):lipide helper:colesterolo:PEG-lipide) in etanolo con RNA disciolto in tampone citrato. Il materiale miscelato è stato diluito istantaneamente con tampone fosfato. L'etanolo è stato rimosso tramite dialisi contro tampone fosfato utilizzando membrana di cellulosa rigenerata (100 kD MWCO) o tramite filtrazione a flusso tangenziale (TFF) utilizzando membrane a fibre cave di polietersulfone modificato (mPES) (100 kD MWCO). Una volta rimosso completamente l'etanolo, il tampone è stato scambiato con tampone HEPES (acido 4-(2-idrossietil)-1-piperazineetansolfonico) contenente 10-300 (ad esempio, 40-60) mM di NaCl e 5-15% di saccarosio, pH 7,3. La formulazione è stata concentrata, seguita da filtrazione a 0,2 µm utilizzando filtri PES. La concentrazione di RNA nella formulazione è stata quindi misurata mediante saggio fluorimetrico RiboGreen e la concentrazione è stata regolata a una concentrazione finale desiderata diluendo con tampone HEPES contenente 10-100 (ad esempio 40-60) mM di NaCl,

0-15% di saccarosio, pH 7,2-8,5 contenente glicerolo. Se non utilizzata immediatamente per ulteriori studi, la formulazione finale è stata quindi filtrata attraverso un filtro da 0,2 µm e riempita in fiale di vetro, tappate, tappate e poste a $-70\pm 5^\circ$ C. Le formulazioni di nanoparticelle lipidiche sono state caratterizzate per il loro pH e osmolalità. Il contenuto di lipidi e il contenuto di RNA sono stati misurati mediante cromatografia liquida ad alte prestazioni (HPLC) e l'integrità dell'mRNA è stata misurata mediante analizzatore di frammenti.

Diffusione dinamica della luce (DLS)

La dimensione media delle particelle (z) e l'indice di polidispersità (PDI) delle formulazioni di nanoparticelle lipidiche utilizzate negli esempi sono stati misurati mediante diffusione dinamica della luce su un Malvern Zetasizer Nano ZS (Regno Unito).

Saggio RiboGreen

L'efficienza di incapsulamento delle formulazioni di nanoparticelle lipidiche è stata caratterizzata utilizzando il test fluorometrico RiboGreen. RiboGreen è un colorante fluorescente proprietario (Molecular Probes/Invitrogen, una divisione di Life Technologies, ora parte di Thermo Fisher Scientific di Eugene, Oregon, Stati Uniti) che viene utilizzato nel rilevamento e nella quantificazione degli acidi nucleici, inclusi sia RNA che DNA. Nella sua forma libera, RiboGreen mostra poca fluorescenza e possiede una firma di assorbimento trascurabile. Quando legato agli acidi nucleici, il colorante emette fluorescenza con un'intensità che è di diversi ordini di grandezza maggiore rispetto alla forma non legata. La fluorescenza può quindi essere rilevata da un sensore (fluorimetro) e l'acido nucleico può essere quantificato.

Processo di liofilizzazione

Gli RNA autoreplicanti (noti anche come RNA replicone) sono in genere più grandi dell'mRNA medio e sono stati progettati dei test per determinare se le formulazioni di nanoparticelle lipidiche di RNA autoreplicante potessero essere liofilizzate con successo. La qualità delle formulazioni di nanoparticelle lipidiche liofilizzate è stata valutata analizzando le formulazioni dopo la liofilizzazione e confrontandole con la formulazione di nanoparticelle lipidiche prima della liofilizzazione e dopo un ciclo di congelamento/scongelo convenzionale (ovvero, congelate a $\sim -70^\circ$ C e poi lasciate scongelare a temperatura ambiente).

L'analisi delle formulazioni di nanoparticelle lipidiche ha incluso l'analisi della dimensione delle particelle e della polidispersità (PDI) e dell'efficienza di incapsulamento (% Encap). La dimensione delle particelle post-liofilizzazione è stata confrontata con la dimensione delle particelle pre-liofilizzazione e la differenza può essere riportata come un delta (Δ). Le varie composizioni testate sono state sottoposte a screening per verificare se fosse stata raggiunta una soglia di proprietà, tra cui un aumento minimo della dimensione delle particelle ($6 < 10$ nm), il mantenimento del PDI ($< 0,2$) e il mantenimento di un'elevata efficienza di incapsulamento ($> 85\%$).

Le formulazioni di nanoparticelle lipidiche sono state preparate come descritto sopra, con RNA autoreplicante (SEQ ID NO: 125). La formulazione di nanoparticelle lipidiche risultante è stata quindi elaborata con uno scambio di tampone per formare una sospensione di preliofilizzazione

avente una concentrazione di RNA autoreplicante da 0,05 a 2,0 mg/mL, sorbato di potassio da 0,01 a 0,05 M, Poloxamer 188 (Kolliphor®) da 0,01 a 0,10% p/v, saccarosio da 14 a 18% p/v, NaCl da 25 a 75 mM e tampone Tris a pH 8,0 da 15 a 25 mM. La formulazione di preliofilizzazione è stata quindi liofilizzata in un Millrock Revo Freeze Dryer (Modello n. RV85S4), utilizzando aliquote di 2,0 mL di sospensione e il ciclo di liofilizzazione fornito nella Tabella 10 di seguito. Le formulazioni liofilizzate di questo esempio sono state quindi applicate agli studi dell'esempio 13 sopra come "LYO-LNP".

TABELLA 10 Ciclo di liofilizzazione per nanoparticelle di RNA-lipidi autoreplicanti Formulazione
Ciclo di liofilizzazione mensola fare un passo camera temperatura durata vuoto Fare un passo
Temperatura ambiente (° C) - ± 2° C. (ore:min) (barra) Congelamento iniziale -50 4:00 atmosfera
Evacuazione -50 dalle 00:30 alle 01:45 dall'atmosfera pressione a 0,05 Essiccazione primaria -50 →
0 63:00 0,05 (rampa di discesa) Essiccazione secondaria 0 → +25 39:30 0,05 (rampa) Riempimento
con N₂ e 25 00:10-00:20 700 ± 50 tappatura Aerazione con aria 5 00:10-00:20 atmosfera

Le particelle liofilizzate preparate seguendo i metodi descritti sopra sono state ricostituite in 2 mL di acqua e caratterizzate utilizzando DLS e RiboGreen. I risultati forniti nella Tabella 11 sottostante mostrano che le composizioni liofilizzate hanno prodotto formulazioni di nanoparticelle lipidiche liofilizzate con dimensioni, polidispersità e valori delta adeguati (~5,3 nm) dopo la ricostituzione.

TABELLA 11 Caratteristiche delle nanoparticelle di RNA-lipidi autoreplicanti Pre e post LYO
Dimensione media delle particelle (nm) PDI incapsulamento (%) Pre-LYO 76.3 0,129 97 Post-LYO
81.6 0,152 93

Affermazioni

1. Una molecola di acido nucleico comprendente:

(i) un primo polinucleotide che codifica una o più proteine di replicazione virale, in cui il primo polinucleotide è ottimizzato per codoni rispetto a un polinucleotide di tipo selvaggio che codifica una o più proteine di replicazione virale, e in cui il primo polinucleotide comprende una sequenza avente almeno l'80% di identità con la sequenza di SEQ ID NO:72; e

(ii) un secondo polinucleotide comprendente un transgene che codifica una proteina antigenica o un frammento della stessa, in cui la proteina antigenica è una proteina del coronavirus codificata da una sequenza avente almeno il 90% di identità con la sequenza di SEQ ID NO: 122.

2. Molecola di acido nucleico secondo la rivendicazione 1, comprendente inoltre una regione non tradotta 5' (UTR), in cui la 5' UTR comprende una sequenza 5' UTR dell'alfavirus.

3. Molecola di acido nucleico secondo la rivendicazione 2, in cui la 5' UTR comprende una sequenza di SEQ ID NO:73, SEQ ID NO:74 o SEQ ID NO:75.

4. Molecola di acido nucleico secondo la rivendicazione 1, comprendente inoltre una regione non tradotta 3' (UTR), in cui la UTR 3' comprende una sequenza UTR 3' dell'alfavirus.

5. Molecola di acido nucleico secondo la rivendicazione 4, in cui la 3' UTR comprende una sequenza poli-A.

6. Molecola di acido nucleico secondo la rivendicazione 1, in cui il primo polinucleotide è localizzato 5' del secondo polinucleotide.

7. Molecola di acido nucleico secondo la rivendicazione 6, comprendente inoltre una regione intergenica situata tra il primo polinucleotide e il secondo polinucleotide.

8. Molecola di acido nucleico secondo la rivendicazione 7, in cui la regione intergenica comprende una sequenza avente almeno l'85% di identità con una sequenza di SEQ ID NO:77.

9. Molecola di acido nucleico secondo la rivendicazione 1, in cui la molecola di acido nucleico è

(a) una molecola di DNA; o

(b) una molecola di RNA, in cui T è sostituita con U.

10. Molecola di acido nucleico secondo la rivendicazione 9, in cui la molecola di DNA comprende inoltre un promotore situato 5' della 5' UTR, in cui il promotore è un promotore T7, un promotore T3 o un promotore SP6.

11. Molecola di acido nucleico secondo la rivendicazione 9, in cui la molecola di RNA è una molecola di RNA autoreplicante.

12. Molecola di acido nucleico secondo la rivendicazione 9, in cui la molecola di RNA comprende inoltre un cappuccio 5' avente una struttura Cap 1, una struttura Cap 1 (m6A), una struttura Cap 2 o una struttura Cap 0.

13. Una composizione comprendente la molecola di acido nucleico della rivendicazione 1 e una formulazione lipidica selezionata tra un lipoplex, un liposoma, una nanoparticella lipidica, un vettore a base di polimero, un esosoma, un corpo lamellare, una micella e un'emulsione.

14. Composizione secondo la rivendicazione 13, in cui la formulazione lipidica è una nanoparticella lipidica.

15. Composizione secondo la rivendicazione 13, in cui la formulazione lipidica comprende un lipide cationico ionizzabile.

16. Composizione della rivendicazione 15, in cui il lipide cationico ionizzabile ha una struttura di Formula I:

o un sale o solvato farmaceuticamente accettabile dello stesso, in cui R5 e R6 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile C1-C31 lineare o ramificato, un alchenile C2-C31 o un alchinile C2-C31 e colesterile; L5 e L6 sono ciascuno indipendentemente selezionati dal gruppo costituito da un alchile C1-C20 lineare e un

alchenile C2-C20; X5 è —C(O)O—, per cui si forma —C(O)O—R6 o —OC(O)— per cui si forma —OC(O)—R6; X6 è —C(O)O— per cui si forma —C(O)O—R5 o —OC(O)— per cui si forma —OC(O)—R5; X7 è S o O; L7 è alchile assente o inferiore; R4 è un alchile C1-C6 lineare o ramificato; e R7 e R8 sono ciascuno selezionati indipendentemente dal gruppo costituito da un idrogeno e un alchile C1-C6 lineare o ramificato.

17. Composizione della rivendicazione 15, in cui il lipide cationico ionizzabile ha una struttura di:
o un sale farmaceuticamente accettabile dello stesso.

18. Composizione secondo la rivendicazione 13, in cui la formulazione lipidica incapsula la molecola di acido nucleico o è complessata alla molecola di acido nucleico.

19. Composizione della rivendicazione 13, in cui la formulazione lipidica comprende

- (a) un lipide ausiliario;
- (b) un fosfolipide;
- (c) un coniugato di polietilenglicole (PEG) e lipidi; o
- (d) qualsiasi combinazione di questi.

20. Composizione secondo la rivendicazione 19, in cui il fosfolipide è selezionato tra dioleoilfosfatidil etanolamina (DOPE), dimiristoilfosfatidilcolina (DMPC), distearoilfosfatidilcolina (DSPC), dimiristoilfosfatidilglicerolo (DMPG), dipalmitoilfosfatidilcolina (DPPC) e fosfatidilcolina (PC).

21. Composizione della rivendicazione 13, in cui la porzione lipidica della formulazione lipidica comprende circa il 40% molare a circa il 60% molare del lipide cationico ionizzabile, circa il 4% molare a circa il 16% molare di DSPC, circa il 30% molare a circa il 47% molare di colesterolo e circa lo 0,5% molare a circa il 3% molare di PEG2000-DMG.

22. Composizione secondo la rivendicazione 13, in cui la composizione ha un rapporto di peso totale delle molecole di lipidi:acido nucleico di circa 50:1 a circa 10:1.

23. Composizione secondo la rivendicazione 13, in cui la composizione comprende

- (a) un tampone HEPES o TRIS a un pH compreso tra circa 7,0 e circa 8,5;
- (b) un tampone HEPES o TRIS ad una concentrazione da circa 7 mg/mL a circa 15 mg/ml;
- (c) circa 2,0 mg/mL a circa 4,0 mg/mL di NaCl;
- (d) uno o più crioprotettori;
- (e) uno o più crioprotettori selezionati tra saccarosio, glicerolo o una combinazione di saccarosio e glicerolo; o
- (f) qualsiasi combinazione di questi.

24. Composizione secondo la rivendicazione 13, in cui la composizione è una composizione liofilizzata.

25. Composizione secondo la rivendicazione 24, in cui la composizione liofilizzata comprende uno o più lioprotettori.

26. Composizione secondo la rivendicazione 24, in cui la composizione liofilizzata comprende un poloxamer, sorbato di potassio, saccarosio o una qualsiasi combinazione di questi.

27. Composizione della rivendicazione 24, in cui la composizione liofilizzata comprende

(a) circa 0,01 a circa 1,0% p/p della molecola di acido nucleico;

(b) circa 1,0 a circa 5,0% p/p di lipidi;

(c) circa 0,5 a circa 2,5% p/p di tampone TRIS;

(d) circa 0,75 a circa 2,75% p/p di NaCl;

(e) circa l'85-95% p/p di uno zucchero;

(f) circa 0,01 a circa 1,0% p/p di un poloxamer;

(g) circa 1,0 a circa 5,0% p/p di sorbato di potassio; o

(h) qualsiasi combinazione di questi.

28. Metodo per somministrare la composizione della rivendicazione 13 a un soggetto che ne ha bisogno, in cui la composizione è liofilizzata e ricostituita prima della somministrazione.

29. Metodo per prevenire o migliorare il COVID-19, comprendente la somministrazione della composizione della rivendicazione 13 a un soggetto che ne ha bisogno.

30. Metodo per somministrare una dose di richiamo a un soggetto vaccinato, comprendente la somministrazione della composizione della rivendicazione 13 a un soggetto precedentemente vaccinato contro il coronavirus.

31. Metodo secondo la rivendicazione 29, in cui la composizione viene somministrata a un dosaggio di circa 0,01 ug a circa 1.000 ug di acido nucleico.

32. Un metodo per indurre una risposta immunitaria contro un coronavirus in un soggetto comprendente:

somministrando al soggetto una quantità efficace di una molecola di acido nucleico secondo la rivendicazione 1.

33. Un metodo per indurre una risposta immunitaria contro un coronavirus in un soggetto comprendente:

somministrando al soggetto una quantità efficace di una composizione della rivendicazione
13.

Riferimenti Citati

Documenti di brevetti USA

7332322	19 febbraio 2008	Frolov e altri
7425337	16 settembre 2008	Smith e altri
7442381	28 ottobre 2008	Smith e altri
8093367	10 gennaio 2012	Kore e altri
8158601	17 aprile 2012	Chen e altri
8304529	6 novembre 2012	Kore e altri
8961995	24 febbraio 2015	Frolov e altri
9254265	9 febbraio 2016	Geall e altri
9295646	29 marzo 2016	Brito e altri
9730997	15 agosto 2017	Perri e altri
9770463	26 settembre 2017	Geall e altri
10238733	26 marzo 2019	Brito e altri
10487105	26 novembre 2019	Chivukula e altri
11135283	5 ottobre 2021	Berglund e altri
11744887	5 settembre 2023	Sullivan e altri
20090075384	19 marzo 2009	Kamrud e altri
20090155301	18 giugno 2009	Mason e altri
20110171255	14 luglio 2011	Kliver e altri
20110207223	25 agosto 2011	Tang e altri
20110256175	20 ottobre 2011	Speranza e altri
20120027803	2 febbraio 2012	Manoharan e altri
20120128760	24 maggio 2012	Manoharan e altri
20120156251	21 giugno 2012	Brito e altri
20130171241	4 luglio 2013	Gelato
20130195968	1 agosto 2013	Geall e altri
20140227346	14 agosto 2014	Geall e altri

20140242152	28 agosto 2014	Geall e altri
20150024002	22 gennaio 2015	Perri e altri
20160074500	17 marzo 2016	Pushko e altri
20160348132	1 dicembre 2016	Rayner e altri
20180036398	8 febbraio 2018	Hagen e altri
20180104359	19 aprile 2018	Camrud
20180169268	21 giugno 2018	Payne e altri
20180171340	21 giugno 2018	Kamrud e altri
20180273576	27 settembre 2018	Hogrefe e altri
20180327471	15 novembre 2018	Limphong e altri
20190091329	28 marzo 2019	Brito e altri
20190224299	25 luglio 2019	Kamrud e altri
20190321458	24 ottobre 2019	Sahin e altri
20190374650	12 dicembre 2019	Luna e altri
20200010849	9 gennaio 2020	Blair e altri
20200113830	16 aprile 2020	Geall e altri
20200113831	16 aprile 2020	Geall e altri
20200222332	16 luglio 2020	Irvine e altri
20200230058	23 luglio 2020	Geall e altri
20200230225	23 luglio 2020	Vogels e altri
20200297634	24 settembre 2020	Karmali e altri
20200330585	22 ottobre 2020	Mogler e altri
20210030859	4 febbraio 2021	Bucala e altri
20210284974	16 settembre 2021	Chivukula e altri
20210290752	23 settembre 2021	Sullivan e altri
20210290756	23 settembre 2021	Sullivan e altri
20220347298	3 novembre 2022	Sullivan e altri
20220395570	15 dicembre 2022	Rauch e altri
20220401550	22 dicembre 2022	Simon-Loriere
20230219996	13 luglio 2023	Matsuda e altri
20240115692	11 aprile 2024	Sullivan e altri

Documenti di brevetti esteri

2591114	Giugno 2016	EP
3471778	Aprile 2019	EP
3433369	Marzo 2020	EP
2729126	Dicembre 2020	EP
2008119827	Ottobre 2008	LUI
2009079185	Giugno 2009	LUI
2009086558	Luglio 2009	LUI
2009127060	Ottobre 2009	LUI
2010048536	Aprile 2010	LUI
2010054406	Maggio 2010	LUI
2010088537	Agosto 2010	LUI
2010129709	Novembre 2010	LUI
2011153493	Dicembre 2011	LUI
2012006369	Gennaio 2012	LUI
2012006380	Gennaio 2012	LUI
2012170431	Dicembre 2012	LUI
2014170493	Ottobre 2014	LUI
2015051169	Aprile 2015	LUI
2015061491	Aprile 2015	LUI
2016184822	Novembre 2016	LUI
2017083356	Maggio 2017	LUI
2017223085	Dicembre 2017	LUI
2018078053	Maggio 2018	LUI
2018208856	Novembre 2018	LUI
2018222890	Dicembre 2018	LUI
2018222926	Dicembre 2018	LUI
2019023566	Gennaio 2019	LUI
2020014654	Gennaio 2020	LUI
2020035609	Febbraio 2020	LUI
2020254535	Dicembre 2020	LUI

2020254804	Dicembre 2020	LUI
2020255055	Dicembre 2020	LUI
2021067181	Aprile 2021	LUI
2021183563	Settembre 2021	LUI
2021183564	Settembre 2021	LUI
2023010128	Febbraio 2023	LUI

Altri riferimenti

- 27 aprile 2019) “Vettore di clonazione pCMV-VEE-GFP”, sequenza completa, ID GenBank: MH891622.1, 7 pagine.
- Rapporto di ricerca internazionale e parere scritto per la domanda n. PCT/US21/21572, inviato il 20 luglio 2021, 11 pagine.
- Rapporto di ricerca internazionale e parere scritto per la domanda n. PCT/US21/21573, inviato per posta il 1° luglio 2021, 12 pagine.
- Rapporto di ricerca internazionale e parere scritto per la domanda n. PCT/US22/74337, inviato il 30 dicembre 2022, 16 pagine.
- (18 luglio 2020) “Glicoproteina di superficie [Sindrome respiratoria acuta grave Coronavirus 2]”, ID GenBank: YP_009724390, 3 pagine.
- Altschul et al. (5 ottobre 1990) “Strumento di ricerca di allineamento locale di base”, *Journal of Molecular Biology*, 215(3):403-410.
- Altschul et al. (1 settembre 1997) “Gapped BLAST e PSI-BLAST: una nuova generazione di programmi di ricerca di database proteici”, *Nucleic Acids Research*, 25(17):3389-3402.
- Baden et al. (4 febbraio 2021) “Efficacia e sicurezza del vaccino mRNA-1273 SARS-CoV-2”, *The New England Journal of Medicine*, 384(5):403-416.
- Bochicchio et al. (2014) “Liposomi come vettori di somministrazione di siRNA”, *Current Drug Metabolism*, 15(9):882-892.
- Boles et al. (2017) “Gene HA H7N9 sintetico costruito, CDS completo”, GenBank KY199425.1, National Library of Medicine, 4 pagine.
- Both et al. (1 marzo 1975) “Traduzione dipendente dalla metilazione di RNA messaggeri virali in vitro”, *Proceedings of the National Academy of Sciences*, 72(3): 1189-1193.
- Bouloy et al. (1 luglio 1980) “Sia il gruppo 7-metilico che quello 2'-O-metilico nel cappuccio dell'mRNA influenzano fortemente la sua capacità di agire come primer per la trascrizione dell'RNA del virus dell'influenza”, *Proceedings of the National Academy of Sciences*, 77(7):3952-3956.
- Chan et al. (19 settembre 2016) “Gli anticorpi cross-reattivi migliorano l'infezione da virus attenuato vivo per una maggiore immunogenicità”, *Nature Microbiology*, 1(12):16164 (10 pagine).
- Chan et al. (5 ottobre 2017) “Correlati molecolari precoci di eventi avversi a seguito della vaccinazione contro la febbre gialla”, *JCI Insight*, 2(19):e96031 (12 pagine).

- Chan et al. (agosto 2019) “Le perturbazioni metaboliche e lo stress cellulare sono alla base della suscettibilità all’infezione sintomatica da febbre gialla attenuata”, *Nature Medicine*, 25(8):1218-1224 (21 pagine).
- Chu et al. (agosto 1978) “Osservazioni paradossali sul termine 5’ dell’acido ribonucleico messaggero dell’ovalbumina”, *Journal of Biological Chemistry*, 253(15):5228-5231.
- Cirelli et al. (16 maggio 2019) “L’immunizzazione a somministrazione lenta migliora le risposte degli anticorpi neutralizzanti l’HIV e del centro germinativo tramite la modulazione dell’immunodominanza”, *Cell*, 177(5):1153-1171.e28 (57 pagine).
- Conticello et al. (22 agosto 2008) “Interazione tra l’enzima di diversificazione degli anticorpi e il fattore associato allo spliceosoma CTNNB1”, *Molecular Cell*, 31(4):474-484.
- Corbett et al. (22 ottobre 2020) “Progettazione del vaccino mRNA SARS-CoV-2 abilitata dalla preparazione del prototipo del patogeno”, *Nature*, 586(7830):567-571.
- Corbett et al. (11 giugno 2020) “Sviluppo del vaccino mRNA contro SARS-CoV-2 reso possibile dalla preparazione al prototipo di patogeno”, *bioRxiv.*, 39 pagine.
- Dabkowska et al. (7 marzo 2012) “L’effetto dei lipidi helper neutri sulla struttura dei monostrati lipidici cationici”, *Journal of the Royal Society Interface*, 9(68):548-561.
- Dua et al. (aprile-giugno 2012) “Liposome: metodi di preparazione e applicazioni”, *International Journal of Pharmaceutical Studies and Research*, 3(3):14-20.
- Dupuis et al. (1 settembre 2000) “La distribuzione dei vaccini a DNA determina la loro immunogenicità dopo l’iniezione intramuscolare nei topi”, *The Journal of Immunology*, 165(5):2850-2858.
- Ehrchen et al. (settembre 2009) “L’agonista endogeno del recettore Toll-like 4 S100A8/S100A9 (calprotectina) come amplificatore innato di infezione, autoimmunità e cancro”, *Journal of Leukocyte Biology*, 86(3):557-566.
- Enright et al. (12 dicembre 2003) “Obiettivi dei microRNA in *Drosophila*”, *Genome Biology*, 5:R1 (14 pagine).
- Geall et al. (4 settembre 2012) “Somministrazione non virale di vaccini a Rna autoamplificanti”, *Proceedings of the National Academy of Sciences*, 109(36):14604-14609.
- Groom et al. (10 marzo 2011) “CXCR3 nella funzione delle cellule T”, *Experimental Cell Research*, 317(5):620-631.
- Gustafsson et al. (luglio 2004) “Codon Bias ed espressione proteica eterologa”, *Trends in Biotechnology*, 22(7):346-353.
- Hashem et al. (8 ottobre 2019) “Un vaccino altamente immunogenico, protettivo e sicuro basato sull’adenovirus che esprime la proteina di fusione S1-cd40l del coronavirus della sindrome respiratoria mediorientale in un modello murino transgenico di dipeptidil peptidasi 4 umana”, *The Journal of Infectious Diseases*, 220(10):1558-1567.
- Hassett et al. (15 aprile 2019) “Ottimizzazione delle nanoparticelle lipidiche per la somministrazione intramuscolare di vaccini a mRNA”, *Molecular Therapy—Nucleic Acids*, 15:1-11.
- Higgins et al. (aprile 2019) “Programmazione della funzione delle plasmacellule isotipo-specifiche”, *Trends Immunology*, 40(4):345-357.
- Honda-Okubo et al. (marzo 2015) “I vaccini contro il coronavirus associati alla sindrome respiratoria acuta grave formulati con adiuvanti di delta inulina forniscono una protezione

migliorata migliorando al contempo l'immunopatologia eosinofila polmonare", *Journal of Virological Methods*, 89(6):2995-3007.

- Hsieh et al. (18 settembre 2020) "Progettazione basata sulla struttura di spike SARS-CoV-2 stabilizzati tramite prefusione", *Science*, 369(6510):1501-1505 (10 pagine).
- Huang et al. (15 agosto 2011) "Sommario in vivo di RNAi con nanoparticelle a base di lipidi", *Annual Review of Biomedical Engineering*, 13:507-530.
- Hyde et al. (3 agosto 2015) "Le estremità 5' e 3' degli RNA non codificanti dell'Alphavirus non sono non funzionali", *Virus Research*, 206:99-107 (8 pagine).
- Ishikawa et al. (27 settembre 2009) "Preparazione di mRNA eucariotico con adenosina diversamente metilata al 5'-terminale ed effetto del gruppo metilico nella traduzione", *Nucleic Acids Symposium*, 53(1):129-130.
- Jackson et al. (4 febbraio 2020) "La promessa dei vaccini a mRNA: una prospettiva biotecnologica e industriale", *NPJ Vaccines*, 5:11 (6 pagine).
- Jin et al. (5 luglio 2010) "Effetti immunomodulatori del dsRNA e il suo potenziale come adiuvante del vaccino", *Journal of Biomedicine and Biotechnology*, 2010:690438.
- Jokerst et al. (giugno 2011) "PEGilazione di nanoparticelle per imaging e terapia", *Nanomedicine (Londra)*, 6(4):715-728.
- Kalnin et al. (2021) "Immunogenicità ed efficacia del vaccino mRNA COVID-19 MRT5500 nei modelli animali preclinici", *NPJ Vaccines*, 6(61):12 pagine.
- Karlin et al. (giugno 1993) "Applicazioni e statistiche per segmenti multipli ad alto punteggio in sequenze molecolari", *Proceedings of the National Academy of Sciences*, 90(12):5873-5877.
- Karlin et al. (marzo 1990) "Metodi per valutare la significatività statistica delle caratteristiche della sequenza molecolare mediante l'uso di schemi di punteggio generali", *Proceedings of the National Academy of Sciences*, 87(6):2264-2268.
- Kasturi et al. (24 febbraio 2011) "Programmazione dell'entità e della persistenza delle risposte anticorpali con l'immunità innata", *Nature*, 470(7335):543-547 (20 pagine).
- Kawabata et al. (1995) "Il destino del DNA plasmidico dopo iniezione endovenosa nei topi: coinvolgimento dei recettori spazzini nel suo assorbimento epatico", *Pharmaceutical Research*, 12:825-830.
- Keech et al. (10 dicembre 2020) "Studio di fase 1-2 di un vaccino con nanoparticelle di proteina Spike ricombinante contro SARS-CoV-2", *The New England Journal of Medicine*, 383:2320-2332.
- Kirchdoerfer et al. (24 ottobre 2018) "I picchi stabilizzati del coronavirus sono resistenti ai cambiamenti conformazionali indotti dal riconoscimento del recettore o dalla proteolisi", *Science Reports*. 8(1):15701 (11 pagine).
- Kowalski et al. (aprile 2019) "Consegna del messaggero: progressi nelle tecnologie per la somministrazione terapeutica di mRNA", *Molecular Therapy*, 27(4):710-728.
- Kozak Marilyn. (novembre 1990) "La struttura secondaria a valle facilita il riconoscimento dei codoni iniziatori da parte dei ribosomi eucariotici", *Proceedings of the National Academy of Sciences*, 87(21):8301-8305.
- Kozak Marilyn. (luglio 1988) "La lunghezza del leader e la struttura secondaria modulano la funzione dell'mRNA in condizioni di stress.", *Biologia molecolare e cellulare*, 8(7):2737-2744.
- Kozak Marilyn. (25 ottobre 1991) "Caratteristiche strutturali negli mRNA eucariotici che modulano l'inizio della traduzione", *Journal of Biological Chemistry*, 266(30):19867-19870.

- Kozak Marilyn. (febbraio 1989) “Il modello di scansione per la traduzione: un aggiornamento”, *Journal of Cell Biology*, 108(2):229-241.
- Kreiter et al. (1 gennaio 2008) “Aumento dell’efficienza di presentazione dell’antigene mediante l’accoppiamento degli antigeni ai segnali di traffico MHC di classe I”, *Journal of Immunology*, 180(1):309-318.
- Kulasegaran-Shylini et al. (25 aprile 2009) “La mutazione specifica 5’UTR nel genoma Veev TC-83 ha un forte effetto sulla replicazione dell’RNA e sulla sintesi dell’RNA subgenomico, ma non sulla traduzione delle proteine codificate”, *Virology*, 387(1):211-221 (24 pagine).
- Kulkarni et al. (giugno 2018) “Nanoparticelle lipidiche che consentono terapie geniche: dai concetti all’utilità clinica”, *Nucleic Acid Therapeutics*, 28(3):146-157.
- Lasic Dand. (1 luglio 1998) “Nuove applicazioni dei liposomi”, *Trends in Biotechnology*, 16(7):307-321.
- Li et al. (gennaio 2011) “Biosintesi delle nanoparticelle da parte dei microrganismi e loro applicazioni”, *Journal of Nanomaterials*, ID articolo 270974, 2011:17 pagine.
- Li et al. (3 agosto 2010) “Nanoparticelle stealth: PEG ad alta densità ma rilasciabile è la chiave per il targeting dei tumori”, *Journal of Controlled Release*, 145(3):178-181.
- Lin et al. (gennaio 2014) “Nanoparticelle a base di lipidi nella somministrazione sistemica di siRNA”, *Nanomedicina*, 9(1):105-120.
- Love et al. (2 febbraio 2010) “Materiali simili ai lipidi per il silenziamento genico in vivo a basso dosaggio”, *Proceedings of the National Academy of Sciences*, 107(5):1864-1869.
- Magini et al. (15 agosto 2016) “I vaccini mRNA autoamplificanti che esprimono più antigeni influenzali conservati conferiscono protezione contro le sfide virali omologhe ed eterosubtipiche”, *PLoS one*, 11(8):e0161193 (25 pagine).
- Maruggi et al. (dicembre 2013) “I vaccini replicanti Alphavirus ingegnerizzati basati su mutanti virali attenuati noti mostrano effetti limitati sull’immunogenicità”, *Virology*, 447(1-2):254-264.
- Maruggi et al. (10 aprile 2019) “mRNA come tecnologia trasformativa per lo sviluppo di vaccini per il controllo delle malattie infettive”, *Molecular Therapy*, 27(4):757-772.
- Muthukrishnan et al. (1 maggio 1975) “La 7-metilguanosina 5’-terminale nell’RNA eucariotico è necessaria per la traduzione”, *Nature*, 255:33-37.
- Olmedillas et al. (6 maggio 2021) "Progettazione basata sulla struttura di un trimero di spike SARS-CoV-2 altamente stabile e legato covalentemente con proprietà strutturali e immunogenicità migliorate", *bioRxiv*, 51 pagine.
- Patil et al. (gennaio 2014) “Nuovi metodi per la preparazione dei liposomi”, *Chimica e fisica dei lipidi*, 177:8-18.
- Pearson et al. (2013) “Introduzione alla ricerca di similarità di sequenza (“omologia”)”, *Current Protocols in Bioinformatics Book*, 42:3.1.1-3.1.8.
- Pearson et al. (aprile 1988) “Strumenti migliorati per il confronto delle sequenze biologiche”, *Proceedings of the National Academy of Sciences*, 85:2444-2448.
- Pepini et al. (15 maggio 2017) “Induzione di una risposta antivirale mediata da IFN da parte di un vaccino a RNA autoamplificante: implicazioni per la progettazione del vaccino”, *The Journal of Immunology*, 198(10):4012-4024 (13 pagine).
- Petkov et al. (4 giugno 2018) “Il sito di immunizzazione del DNA determina il livello di espressione genica e l’entità, ma non il tipo di risposta immunitaria indotta”, *PLoS One*, 13(6):

e0197902 (22 pagine).

- Polack et al. (31 dicembre 2020) “Sicurezza ed efficacia del vaccino Covid-19 a mRNA BNT162b2”, *The New England Journal of Medicine*, 383:2603-2615.
- Querec et al. (gennaio 2009) “L’approccio della biologia dei sistemi predice l’immunogenicità del vaccino contro la febbre gialla negli esseri umani”, *National Immunology*, 10(1):116-125 (26 pagine).
- Querec et al. (2007) “Comprendere il ruolo dell’immunità innata nel meccanismo d’azione del vaccino vivo attenuato contro la febbre gialla 17D”, *Advances in Experimental Medicine and Biology*, 590:43-53.
- Ramanathan et al. (19 settembre 2016) “mRNA Capping: funzioni biologiche e applicazioni”, *Nucleic Acids Research*, 44(16):7511-7526.
- Rodriguez-Gascon et al., (10 aprile 2014) “Sviluppo di vaccini con acidi nucleici: uso di RNA autoamplificante in nanoparticelle lipidiche”, *International Journal of Nanomedicine*, 9:1833-1843.
- Sahin et al. (11 dicembre 2020) “BNT162b2 induce anticorpi neutralizzanti SARS-CoV-2 e cellule T negli esseri umani”, medRxiv, 49 pagine.
- Sahin et al. (27 maggio 2021) “Il vaccino BNT162b2 induce anticorpi neutralizzanti e cellule T polispecifiche negli esseri umani”, *Nature*, 595:572-577.
- Saltl et al. (15 dicembre 2011) “Il granzima B regola le risposte delle cellule T CD8+ antivirali”, *Journal of Immunology*, 187(12):6301-6309 (19 pagine).
- Sercombe et al. (1 dicembre 2015) “Progressi e sfide della somministrazione di farmaci assistita da liposomi”, *Frontiers in Pharmacology*, 6(286):13 pagine.
- Slansky et al. (ottobre 2000) “Immunità antitumorale antigene-specifica migliorata con ligandi peptidici alterati che stabilizzano il complesso MHC-peptide-TCR”, *Immunity*, 13(4):529-538.
- Tam et al. (4 ottobre 2016) “La disponibilità sostenuta dell’antigene durante l’inizio del centro germinativo migliora le risposte anticorpali alla vaccinazione”, *Proceedings of the National Academy of Sciences*, 113(43):e6639-e6648.
- Taverniti et al. (9 gennaio 2015) “L’eliminazione delle strutture del cappuccio generate dal decadimento dell’mRNA coinvolge il nuovo enzima scavenger di decappaggio dell’mRNA Aph1/FHIT insieme a DcpS”, *Nucleic Acids Research*, 43(1):482-492.
- Thompson et al. (7 marzo 2006) “Attività adiuvante mucosale e sistemica delle particelle repliconiche dell’alfavirus”, *Proceedings of the National Academy of Sciences*, 103(10):3722-3727.
- Villalobos et al. (6 giugno 2006) “Gene Designer: uno strumento di biologia sintetica per la costruzione di segmenti di DNA artificiali”, *BMC Bioinformatics*, 7:285 (8 pagine).
- Von Herrath et al. (giugno 2003) “Risposta immunitaria, tolleranza e dsRNA: implicazioni per i paradigmi tradizionali”, *Trends in Immunology*, 24(6):289-293 (4 pagine).
- Wootton et al. (giugno 1993) “Statistiche della complessità locale nelle sequenze di amminoacidi e nei database di sequenze”, *Computers & Chemistry*, 17(2):149-163.
- Wrapp et al. (13 marzo 2020) “Struttura crio-EM dello spike 2019-nCoV nella conformazione pre-fusione”, *Science*, 367(6483):1260-1263.
- Wu et al. (12 marzo 2020) “Un nuovo coronavirus associato a malattie respiratorie umane in Cina”, *Nature*, 579(7798):265-269.

- Yu et al. (settembre 2000) “APRIL e TALL-I e recettori BCMA e TACI: sistema per la regolazione dell’immunità umorale”, Nature Immunology, 1(3):252-256.
- Rapporto di ricerca europeo esteso per la domanda n. EP 21767978.6, inviato per posta il 30 aprile 2024, 11 pagine.
- Rapporto di ricerca europeo esteso per la domanda n. EP 21768525.4, inviato per posta il 22 marzo 2024, 10 pagine.
- (Data di recupero 28 agosto 2024) "Tris", recuperato da <https://www.sigmaaldrich.com/US/en/products/chemistry-and-biochemicals/biochemicals/biological-buffers/tris>, 7 pagine.
- Keyer et al. (2019) “Poliproteina non strutturale [vettore di clonazione pCMV-VEE-GFP]”, GenBank: QCD25069.1, 2 pagine.
- Kim et al. (2014) “Miglioramento dell’espressione proteica da parte dei repliconi dell’alfavirus mediante la progettazione di RNA subgenomici autoreplicanti”, Proceedings of the National Academy of Sciences degli Stati Uniti d’America, 111(29):10708-10713.
- Lundstrom, Kenneth (2016) “Vettori virali a RNA repliconici come vaccini”, Vaccini, 4(4):39.
- Lundstrom, Kenneth (2018) “Virus a RNA autoreplicanti per la terapia a RNA”, Molecules, 23(12):3310.
- Shustov et al. (2006) “VEEV Replicon Vector YFV-C1, sequenza completa”, GenBank: DQ322637.1, 4 pagine.

Cronologia dei brevetti

Numero di brevetto : 12220455

Tipo: Concessione

Depositato : 30 giugno 2023

Data del brevetto : 11 febbraio 2025

Numero di pubblicazione del brevetto : 20240115691

Assegnatario : Arcturus Therapeutics, Inc. (San Diego, CA)

Inventori : Sean Michael Sullivan (Escondido, CA), Daiki Matsuda (San Diego, CA), Kiyoshi Tachikawa (San Diego, CA), Padmanabh Chivukula (San Diego, CA), Priya Prakash Karmali (San Diego, CA), Jared Henry Davis (Poway, CA), Yanjie Bao (San Diego, CA), Amit Sagi (San Diego, CA)

Esaminatore principale : Rachel B Gill

Numero di domanda : 18/345,893

Classificazioni

Classificazione Internazionale : A61K 39/215 (20060101); A61K 9/51 (20060101); A61K 38/00 (20060101); A61K 39/00 (20060101); A61K 39/12 (20060101); A61K 47/10 (20170101); A61K 47/20 (20060101); A61K 47/26 (20060101); C07K 14/005 (20060101); C07K 14/18 (20060101); C12N 7/00 (20060101); C12N 15/86 (20060101);