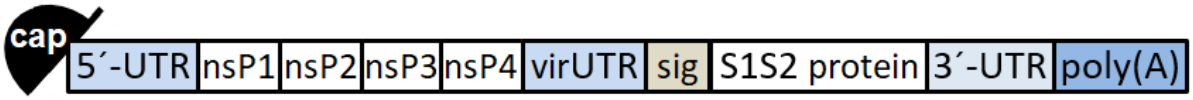
11891

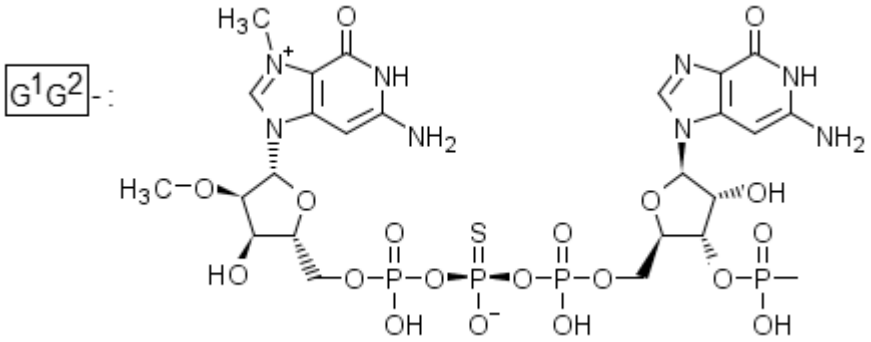
Description

Self-replicating messenger RNA encoding the full-length SARS-CoV-2 spike glycoprotein.

Schematic



UTR = Untranslated region; nsP1, nsP2, nsP3 and nsP4 = Venezuelan equine encephalitis virus (VEEV) RNA-dependent RNA polymerase; virUTR = 5´-UTR derived from VEEV; sig = extended signal sequence of the S glycoprotein. S1S2 protein = S glycoprotein sequence containing mutations K986P and V987P.



5‘- capping structure

The 5‘- capping structure is **= $m^7G^+m^{2'}$-5'-p-(*R*)p^s-p-5'-G**

Table of features

Element	Description	Position
cap	cap: $m^7G^+m^{2'}$-5'-p-(*R*)p^s-p-5'-G	1-2
5’-UTR	5´-untranslated region derived from the Venezuelan equine encephalitis virus (VEEV)	3-46

nsP1	Venezuelan equine encephalitis virus (VEEV) RNA-dependent RNA polymerase (RNA replicase) with a subgenomic promotor starting in nsP4. – In conjunction with the virUTR, these sequences introduce a self-amplifying function.	47-1651
nsP2		1652-4033
nsP3		4034-5704
nsP4		5705-7528
virUTR	VEEV untranslated region encoding 5 nucleotides of the subgenomic promotor as well as replication and translation supporting sequence elements.	7529-7569
sig	S glycoprotein signal peptide (extended leader sequence), which guides translocation of the nascent polypeptide chain into the endoplasmic reticulum.	7570 - 7617
S1S2 protein	Codon-optimized sequences encoding full-length SARS-CoV-2 S glycoprotein; two mutations, K986P and V987, ensure that the S glycoprotein remains in an antigenically optimal pre-fusion conformation.	7618 - 11394
3’-UTR	3´-untranslated region comprising sequence elements derived from the amino-terminal enhancer of split (AES) mRNA and the mitochondrial encoded 12S ribosomal RNA, to confer RNA stability and highest total protein expression, followed by a further Venezuelan equine encephalitis virus (VEEV) derived sequence of the 26S mRNA.	11395-11808
poly(A)	A 110-nucleotide poly(A)-tail consisting of a stretch of 30 adenosine residues, followed by a 10-nucleotide linker sequence and another 70 adenosine residues.	11809-11918

Sequence / Séquence / Secuencia

```
GGAUGGGCGG CGCAUGAGAG AAGCCCAGAC CAAUUACCUA CCCAAAAUGG    50
AGAAAGUUCA CGUUGACAUC GAGGAAGACA GCCCAUUCCU CAGAGCUUUG   100
CAGCGGAGCU UCCCGCAGUU UGAGGUAGAA GCCAAGCAGG UCACUGAUAA   150
UGACCAUGCU AAUGCCAGAG CGUUUUCGCA UCUGGCUUCA AAACUGAUCG   200
AAACGGAGGU GGACCCAUCC GACACGAUCC UUGACAUUGG AAGUGCGCCC   250
GCCCGCAGAA UGUAUUCUAA GCACAAGUAU CAUUGUAUCU GUCCGAUGAG   300
AUGUGCGGAA GAUCCGGACA GAUUGUAUAA GUAUGCAACU AAGCUGAAGA   350
AAAACUGUAA GGAAAUAACU GAUAAGGAAU UGGACAAGAA AAUGAAGGAG   400
CUCGCCGCCG UCAUGAGCGA CCCUGACCUG GAAACUGAGA CUAUGUGCCU   450
CCACGACGAC GAGUCGUGUC GCUACGAAGG GCAAGUCGCU GUUUACCAGG   500
```

```
AUGUAUACGC GGUUGACGGA CCGACAAGUC UCUAUCACCA AGCCAAUAAG   550
GGAGUUAGAG UCGCCUACUG GAUAGGCUUU GACACCACCC CUUUUAUGUU   600
UAAGAACUUG GCUGGAGCAU AUCCAUCAUA CUCUACCAAC UGGGCCGACG   650
AAACCGUGUU AACGGCUCGU AACAUAGGCC UAUGCAGCUC UGACGUUAUG   700
GAGCGGUCAC GUAGAGGGAU GUCCAUUCUU AGAAAGAAGU AUUUGAAACC   750
AUCCAACAAU GUUCUAUUCU CUGUUGGCUC GACCAUCUAC CACGAAAAGA   800
GGGACUUACU GAGGAGCUGG CACCUGCCGU CUGUAUUUCA CUUACGUGGC   850
AAGCAAAAUU ACACAUGUCG GUGUGAGACU AUAGUUAGUU GCGACGGGUA   900
CGUCGUUAAA AGAAUAGCUA UCAGUCCAGG CCUGUAUGGG AAGCCUUCAG   950
GCUAUGCUGC UACGAUGCAC CGCGAGGGAU UCUUGUGCUG CAAAGUGACA  1000
GACACAUUGA ACGGGGAGAG GGUCUCUUUU CCCGUGUGCA CGUAUGUGCC  1050
AGCUACAUUG UGUGACCAAA UGACUGGCAU ACUGGCAACA GAUGUCAGUG  1100
CGGACGACGC GCAAAAACUG CUGGUUGGGC UCAACCAGCG UAUAGUCGUC  1150
AACGGUCGCA CCCAGAGAAA CACCAAUACC AUGAAAAAUU ACCUUUUGCC  1200
CGUAGUGGCC CAGGCAUUUG CUAGGUGGGC AAAGGAAUAU AAGGAAGAUC  1250
AAGAAGAUGA AAGGCCACUA GGACUACGAG AUAGACAGUU AGUCAUGGGG  1300
UGUUGUUGGG CUUUUAGAAG GCACAAGAUA ACAUCUAUUU AUAAGCGCCC  1350
GGAUACCCAA ACCAUCAUCA AAGUGAACAG CGAUUUCCAC UCAUUCGUGC  1400
UGCCCAGGAU AGGCAGUAAC ACAUUGGAGA UCGGGCUGAG AACAAGAAUC  1450
AGGAAAAUGU UAGAGGAGCA CAAGGAGCCG UCACCUCUCA UUACCGCCGA  1500
GGACGUACAA GAAGCUAAGU GCGCAGCCGA UGAGGCUAAG GAGGUGCGUG  1550
AAGCCGAGGA GUUGCGCGCA GCUCUACCAC CUUUGGCAGC UGAUGUUGAG  1600
GAGCCCACUC UGGAAGCCGA UGUCGACUUG AUGUUACAAG AGGCUGGGGC  1650
CGGCUCAGUG GAGACACCUC GUGGCUUGAU AAAGGUUACC AGCUACGCUG  1700
GCGAGGACAA GAUCGGCUCU UACGCUGUGC UUUCUCCGCA GGCUGUACUC  1750
AAGAGUGAAA AAUUAUCUUG CAUCCACCCU CUCGCUGAAC AAGUCAUAGU  1800
GAUAACACAC UCUGGCCGAA AAGGGCGUUA UGCCGUGGAA CCAUACCAUG  1850
GUAAAGUAGU GGUGCCAGAG GGACAUGCAA UACCCGUCCA GGACUUUCAA  1900
GCUCUGAGUG AAAGUGCCAC CAUUGUGUAC AACGAACGUG AGUUCGUAAA  1950
CAGGUACCUG CACCAUAUUG CCACACAUGG AGGAGCGCUG AACACUGAUG  2000
AAGAAUAUUA CAAAACUGUC AAGCCCAGCG AGCACGACGG CGAAUACCUG  2050
UACGACAUCG ACAGGAAACA GUGCGUCAAG AAAGAGCUAG UCACUGGGCU  2100
AGGGCUCACA GGCGAGCUGG UCGAUCCUCC CUUCCAUGAA UUCGCCUACG  2150
AGAGUCUGAG AACACGACCA GCCGCUCCUU ACCAAGUACC AACCAUAGGG  2200
GUGUAUGGCG UGCCAGGAUC AGGCAAGUCU GGCAUCAUUA AAAGCGCAGU  2250
CACCAAAAAA GAUCUAGUGG UGAGCGCCAA GAAAGAAAAC UGUGCAGAAA  2300
UUAUAAGGGA CGUCAAGAAA AUGAAAGGGC UGGACGUCAA UGCCAGAACU  2350
GUGGACUCAG UGCUCUUGAA UGGAUGCAAA CACCCCGUAG AGACCCUGUA  2400
UAUUGACGAG GCUUUUGCUU GUCAUGCAGG UACUCUCAGA GCGCUCAUAG  2450
CCAUUAUAAG ACCUAAAAAG GCAGUGCUCU GCGGAGAUCC CAAACAGUGC  2500
GGUUUUUUUA ACAUGAUGUG CCUGAAAGUG CAUUUUAACC ACGAGAUUUG  2550
CACACAAGUC UUCCACAAAA GCAUCUCUCG CCGUUGCACU AAAUCUGUGA  2600
CUUCGGUCGU CUCAACCUUG UUUUACGACA AAAAAAUGAG AACGACGAAU  2650
CCGAAAGAGA CUAAGAUUGU GAUUGACACU ACCGGCAGUA CCAAACCUAA  2700
GCAGGACGAU CUCAUUCUCA CUUGUUUCAG AGGGUGGGUG AAGCAGUUGC  2750
AAAUAGAUUA CAAAGGCAAC GAAAUAAUGA CGGCAGCUGC CUCUCAAGGG  2800
CUGACCCGUA AAGGUGUGUA UGCCGUUCGG UACAAGGUGA AUGAAAAUCC  2850
UCUGUACGCA CCCACCUCAG AACAUGUGAA CGUCCUACUG ACCCGCACGG  2900
AGGACCGCAU CGUGUGGAAA ACACUAGCCG GCGACCCAUG GAUAAAAACA  2950
CUGACUGCCA AGUACCCUGG GAAUUUCACU GCCACGAUAG AGGAGUGGCA  3000
AGCAGAGCAU GAUGCCAUCA UGAGGCACAU CUUGGAGAGA CCGGACCCUA  3050
CCGACGUCUU CCAGAAUAAG GCAAACGUGU GUUGGGCCAA GGCUUUAGUG  3100
```

```
CCGGUGCUGA AGACCGCUGG CAUAGACAUG ACCACUGAAC AAUGGAACAC  3150
UGUGGAUUAU UUUGAAACGG ACAAAGCUCA CUCAGCAGAG AUAGUAUUGA  3200
ACCAACUAUG CGUGAGGUUC UUUGGACUCG AUCUGGACUC CGGUCUAUUU  3250
UCUGCACCCA CUGUUCCGUU AUCCAUUAGG AAUAAUCACU GGGAUAACUC  3300
CCCGUCGCCU AACAUGUACG GGCUGAAUAA AGAAGUGGUC CGUCAGCUCU  3350
CUCGCAGGUA CCCACAACUG CCUCGGGCAG UUGCCACUGG UAGAGUCUAU  3400
GACAUGAACA CUGGUACACU GCGCAAUUAU GAUCCGCGCA UAAACCUAGU  3450
ACCUGUAAAC AGAAGACUGC CUCAUGCUUU AGUCCUCCAC CAUAAUGAAC  3500
ACCCACAGAG UGACUUUUCU UCAUUCGUCA GCAAAUUGAA GGGCAGAACU  3550
GUCCUGGUGG UCGGGGAAAA GUUGUCCGUC CCAGGCAAAA UGGUUGACUG  3600
GUUGUCAGAC CGGCCUGAGG CUACCUUCAG AGCUCGGCUG GAUUUAGGCA  3650
UCCCAGGUGA UGUGCCCAAA UAUGACAUAA UAUUUGUUAA UGUGAGGACC  3700
CCAUAUAAAU ACCAUCACUA UCAGCAGUGU GAAGACCAUG CCAUUAAGCU  3750
AAGCAUGUUG ACCAAGAAAG CAUGUCUGCA UCUGAAUCCC GGCGGAACCU  3800
GUGUCAGCAU AGGUUAUGGU UACGCUGACA GGGCCAGCGA AAGCAUCAUU  3850
GGUGCUAUAG CGCGGCAGUU CAAGUUUUCC CGAGUAUGCA AACCGAAAUC  3900
CUCACUUGAG GAGACGGAAG UUCUGUUUGU AUUCAUUGGG UACGAUCGCA  3950
AGGCCCGUAC GCACAAUCCU UACAAGCUAU CAUCAACCUU GACCAACAUU  4000
UAUACAGGUU CCAGACUCCA CGAAGCCGGA UGUGCACCCU CAUAUCAUGU  4050
GGUGCGAGGG GAUAUUGCCA CGGCCACCGA AGGAGUGAUU AUAAAUGCUG  4100
CUAACAGCAA AGGACAACCU GGCGGAGGGG UGUGCGGAGC GCUGUAUAAG  4150
AAAUUCCCGG AAAGUUUCGA UUUACAGCCG AUCGAAGUAG GAAAAGCGCG  4200
ACUGGUCAAA GGUGCAGCUA AACAUAUCAU UCAUGCCGUA GGACCAAACU  4250
UCAACAAAGU UUCGGAGGUU GAAGGUGACA AACAGUUGGC AGAGGCUUAU  4300
GAGUCCAUCG CUAAGAUUGU CAACGAUAAC AAUUACAAGU CAGUAGCGAU  4350
UCCACUGUUG UCCACCGGCA UCUUUUCCGG GAACAAAGAU CGACUAACCC  4400
AAUCAUUGAA CCAUUUGCUG ACAGCUUUAG ACACCACUGA UGCAGAUGUA  4450
GCCAUAUACU GCAGGGACAA GAAAUGGGAA AUGACUCUCA AGGAAGCAGU  4500
GGCUAGGAGA GAAGCAGUGG AGGAGAUAUG CAUAUCCGAC GAUUCUUCAG  4550
UGACAGAACC UGAUGCAGAG CUGGUGAGGG UGCAUCCCAA GAGUUCUUUG  4600
GCUGGAAGGA AGGGCUACAG CACAAGCGAU GGCAAAACUU UCUCAUAUUU  4650
GGAAGGGACC AAGUUUCACC AGGCGGCCAA GGAUAUAGCA GAAAUUAAUG  4700
CCAUGUGGCC CGUUGCAACG GAGGCCAAUG AGCAGGUAUG CAUGUAUAUC  4750
CUCGGAGAAA GCAUGAGCAG UAUUAGGUCG AAAUGCCCCG UCGAGGAGUC  4800
GGAAGCCUCC ACACCACCUA GCACGCUGCC UUGCUUGUGC AUCCAUGCCA  4850
UGACUCCAGA AAGAGUACAG CGCCUAAAAG CCUCACGUCC AGAACAAAUU  4900
ACUGUGUGCU CAUCCUUUCC AUUGCCGAAG UAUAGAAUCA CUGGUGUGCA  4950
GAAGAUCCAA UGCUCCCAGC CUAUAUUGUU CUCACCGAAA GUGCCUGCGU  5000
AUAUUCAUCC AAGGAAGUAU CUCGUGGAAA CACCACCGGU AGACGAGACU  5050
CCGGAGCCAU CGGCAGAGAA CCAAUCCACA GAGGGGACAC CUGAACAACC  5100
ACCACUUAUA ACCGAGGAUG AGACCAGGAC UAGAACGCCU GAGCCGAUCA  5150
UCAUCGAAGA AGAAGAAGAA GAUAGCAUAA GUUUGCUGUC AGAUGGCCCG  5200
ACCCACCAGG UGCUGCAAGU CGAGGCAGAC AUUCACGGGC CGCCCUCUGU  5250
AUCUAGCUCA UCCUGGUCCA UUCCUCAUGC AUCCGACUUU GAUGUGGACA  5300
GUUUAUCCAU ACUUGACACC CUGGAGGGAG CUAGCGUGAC CAGCGGGGCA  5350
ACGUCAGCCG AGACUAACUC UUACUUCGCA AAGAGUAUGG AGUUUCUGGC  5400
GCGACCGGUG CCUGCGCCUC GAACAGUAUU CAGGAACCCU CCACAUCCCG  5450
CUCCGCGCAC AAGAACACCG UCACUUGCAC CCAGCAGGGC CUGCUCCAGA  5500
ACCAGCCUAG UUUCCACCCC GCCAGGCGUG AAUAGGGUGA UCACUAGAGA  5550
GGAGCUCGAA GCGCUUACCC CGUCACGCAC UCCUAGCAGG UCGGUCUCCA  5600
GAACCAGCCU GGUCUCCAAC CCGCCAGGCG UAAAUAGGGU GAUUACAAGA  5650
GAGGAGUUUG AGGCGUUCGU AGCACAACAA CAAUGACGGU UUGAUGCGGG  5700
```

UGCAUACAUC UUUUCCUCCG ACACCGGUCA AGGGCAUUUA CAACAAAAAU 5750
CAGUAAGGCA AACGGUGCUA UCCGAAGUGG UGUUGGAGAG GACCGAAUUG 5800
GAGAUUUCGU AUGCCCCGCG CCUCGACCAA GAAAAAGAAG AAUUACUACG 5850
CAAGAAAUUA CAGUUAAAUC CCACACCUGC UAACAGAAGC AGAUACCAGU 5900
CCAGGAAGGU GGAGAACAUG AAAGCCAUAA CAGCUAGACG UAUUCUGCAA 5950
GGCCUAGGGC AUUAUUUGAA GGCAGAAGGA AAAGUGGAGU GCUACCGAAC 6000
CCUGCAUCCU GUUCCUUUGU AUUCAUCUAG UGUGAACCGU GCCUUUUCAA 6050
GCCCCAAGGU CGCAGUGGAA GCCUGUAACG CCAUGUUGAA AGAGAACUUU 6100
CCGACUGUGG CUUCUUACUG UAUUAUUCCA GAGUACGAUG CCUAUUUGGA 6150
CAUGGUUGAC GGAGCUUCAU GCUGCUUAGA CACUGCCAGU UUUUGCCCUG 6200
CAAAGCUGCG CAGCUUUCCA AAGAAACACU CCUAUUUGGA ACCCACAAUA 6250
CGAUCGGCAG UGCCUUCAGC GAUCCAGAAC ACGCUCCAGA ACGUCCUGGC 6300
AGCUGCCACA AAAAGAAAUU GCAAUGUCAC GCAAAUGAGA GAAUUGCCCG 6350
UAUUGGAUUC GGCGGCCUUU AAUGUGGAAU GCUUCAAGAA AUAUGCGUGU 6400
AAUAAUGAAU AUUGGGAAAC GUUUAAAGAA AACCCCAUCA GGCUUACUGA 6450
AGAAAACGUG GUAAAUUACA UUACCAAAUU AAAAGGACCA AAAGCUGCUG 6500
CUCUUUUUGC GAAGACACAU AAUUUGAAUA UGUUGCAGGA CAUACCAAUG 6550
GACAGGUUUG UAAUGGACUU AAAGAGAGAC GUGAAAGUGA CUCCAGGAAC 6600
AAAACAUACU GAAGAACGGC CCAAGGUACA GGUGAUCCAG GCUGCCGAUC 6650
CGCUAGCAAC AGCGUAUCUG UGCGGAAUCC ACCGAGAGCU GGUUAGGAGA 6700
UUAAAUGCGG UCCUGCUUCC GAACAUUCAU ACACUGUUUG AUAUGUCGGC 6750
UGAAGACUUU GACGCUAUUA UAGCCGAGCA CUUCCAGCCU GGGGAUUGUG 6800
UUCUGGAAAC UGACAUCGCG UCGUUUGAUA AAAGUGAGGA CGACGCCAUG 6850
GCUCUGACCG CGUUAAUGAU UCUGGAAGAC UUAGGUGUGG ACGCAGAGCU 6900
GUUGACGCUG AUUGAGGCGG CUUUCGGCGA AAUUUCAUCA AUACAUUUGC 6950
CCACUAAAAC UAAAUUUAAA UUCGGAGCCA UGAUGAAAUC UGGAAUGUUC 7000
CUCACACUGU UUGUGAACAC AGUCAUUAAC AUUGUAAUCG CAAGCAGAGU 7050
GUUGAGAGAA CGGCUAACCG GAUCACCAUG UGCAGCAUUC AUUGGAGAUG 7100
ACAAUAUCGU GAAAGGAGUC AAAUCGGACA AAUUAAUGGC AGACAGGUGC 7150
GCCACCUGGU UGAAUAUGGA AGUCAAGAUU AUAGAUGCUG UGGUGGGCGA 7200
GAAAGCGCCU UAUUUCUGUG GAGGGUUUAU UUUGUGUGAC UCCGUGACCG 7250
GCACAGCGUG CCGUGUGGCA GACCCCCUAA AAAGGCUGUU UAAGCUAGGC 7300
AAACCUCUGG CAGCAGACGA UGAACAUGAU GAUGACAGGA GAAGGGCAUU 7350
GCAUGAGGAG UCAACACGCU GGAACCGAGU GGGUAUUCUU UCAGAGCUGU 7400
GCAAGGCAGU AGAAUCAAGG UAUGAAACCG UAGGAACUUC CAUCAUAGUU 7450
AUGGCCAUGA CUACUCUAGC UAGCAGUGUU AAAUCAUUCA GCUACCUGAG 7500
AGGGGCCCCU AUAACUCUCU ACGGCUAA*CC UGAAUGGACU ACGACAUAGU* 7550
*CUAGUCCGCC AAGACUAGU***A UGUUCGUGUU CCUGGUGCUG CUGCCUCUGG** 7600
UGUCCAGCCA GUGUGUGAAC CUGACCACCA GAACACAGCU GCCUCCAGCC 7650
UACACCAACA GCUUUACCAG AGGCGUGUAC UACCCCGACA AGGUGUUCAG 7700
AUCCAGCGUG CUGCACUCUA CCCAGGACCU GUUCCUGCCU UUCUUCAGCA 7750
ACGUGACCUG GUUCCACGCC AUCCACGUGU CCGGCACCAA UGGCACCAAG 7800
AGAUUCGACA ACCCCGUGCU GCCCUUCAAC GACGGGGUGU ACUUUGCCAG 7850
CACCGAGAAG UCCAACAUCA UCAGAGGCUG GAUCUUCGGC ACCACACUGG 7900
ACAGCAAGAC CCAGAGCCUG CUGAUCGUGA ACAACGCCAC CAACGUGGUC 7950
AUCAAAGUGU GCGAGUUCCA GUUCUGCAAC GACCCCUUCC UGGGCGUCUA 8000
CUACCACAAG AACAACAAGA GCUGGAUGGA AAGCGAGUUC CGGGUGUACA 8050
GCAGCGCCAA CAACUGCACC UUCGAGUACG UGUCCCAGCC UUUCCUGAUG 8100
GACCUGGAAG GCAAGCAGGG CAACUUCAAG AACCUGCGCG AGUUCGUGUU 8150
UAAGAACAUC GACGGCUACU UCAAGAUCUA CAGCAAGCAC ACCCCUAUCA 8200
ACCUCGUGCG GGAUCUGCCU CAGGGCUUCU CUGCUCUGGA ACCCCUGGUG 8250
GAUCUGCCCA UCGGCAUCAA CAUCACCCGG UUUCAGACAC UGCUGGCCCU 8300

GCACAGAAGC UACCUGACAC CUGGCGAUAG CAGCAGCGGA UGGACAGCUG 8350
GUGCCGCCGC UUACUAUGUG GGCUACCUGC AGCCUAGAAC CUUCCUGCUG 8400
AAGUACAACG AGAACGGCAC CAUCACCGAC GCCGUGGAUU GUGCUCUGGA 8450
UCCUCUGAGC GAGACAAAGU GCACCCUGAA GUCCUUCACC GUGGAAAAGG 8500
GCAUCUACCA GACCAGCAAC UUCCGGGUGC AGCCCACCGA AUCCAUCGUG 8550
CGGUUCCCCA AUAUCACCAA UCUGUGCCCC UUCGGCGAGG UGUUCAAUGC 8600
CACCAGAUUC GCCUCUGUGU ACGCCUGGAA CCGGAAGCGG AUCAGCAAUU 8650
GCGUGGCCGA CUACUCCGUG CUGUACAACU CCGCCAGCUU CAGCACCUUC 8700
AAGUGCUACG GCGUGUCCCC UACCAAGCUG AACGACCUGU GCUUCACAAA 8750
CGUGUACGCC GACAGCUUCG UGAUCCGGGG AGAUGAAGUG CGGCAGAUUG 8800
CCCCUGGACA GACAGGCAAG AUCGCCGACU ACAACUACAA GCUGCCCGAC 8850
GACUUCACCG GCUGUGUGAU UGCCUGGAAC AGCAACAACC UGGACUCCAA 8900
AGUCGGCGGC AACUACAAUU ACCUGUACCG GCUGUUCCGG AAGUCCAAUC 8950
UGAAGCCCUU CGAGCGGGAC AUCUCCACCG AGAUCUAUCA GGCCGGCAGC 9000
ACCCCUUGUA ACGGCGUGGA AGGCUUCAAC UGCUACUUCC CACUGCAGUC 9050
CUACGGCUUU CAGCCCACAA AUGGCGUGGG CUAUCAGCCC UACAGAGUGG 9100
UGGUGCUGAG CUUCGAACUG CUGCAUGCCC CUGCCACAGU GUGCGGCCCU 9150
AAGAAAAGCA CCAAUCUCGU GAAGAACAAA UGCGUGAACU UCAACUUCAA 9200
CGGCCUGACC GGCACCGGCG UGCUGACAGA GAGCAACAAG AAGUUCCUGC 9250
CAUUCCAGCA GUUUGGCCGG GAUAUCGCCG AUACCACAGA CGCCGUUAGA 9300
GAUCCCCAGA CACUGGAAAU CCUGGACAUC ACCCCUUGCA GCUUCGGCGG 9350
AGUGUCUGUG AUCACCCCUG GCACCAACAC CAGCAAUCAG GUGGCAGUGC 9400
UGUACCAGGA CGUGAACUGU ACCGAAGUGC CCGUGGCCAU UCACGCCGAU 9450
CAGCUGACAC CUACAUGGCG GGUGUACUCC ACCGGCAGCA AUGUGUUUCA 9500
GACCAGAGCC GGCUGUCUGA UCGGAGCCGA GCACGUGAAC AAUAGCUACG 9550
AGUGCGACAU CCCCAUCGGC GCUGGAAUCU GCGCCAGCUA CCAGACACAG 9600
ACAAACAGCC CUCGGAGAGC CAGAAGCGUG GCCAGCCAGA GCAUCAUUGC 9650
CUACACAAUG UCUCUGGGCG CCGAGAACAG CGUGGCCUAC UCCAACAACU 9700
CUAUCGCUAU CCCCACCAAC UUCACCAUCA GCGUGACCAC AGAGAUCCUG 9750
CCUGUGUCCA UGACCAAGAC CAGCGUGGAC UGCACCAUGU ACAUCUGCGG 9800
CGAUUCCACC GAGUGCUCCA ACCUGCUGCU GCAGUACGGC AGCUUCUGCA 9850
CCCAGCUGAA UAGAGCCCUG ACAGGGAUCG CCGUGGAACA GGACAAGAAC 9900
ACCCAAGAGG UGUUCGCCCA AGUGAAGCAG AUCUACAAGA CCCCUCCUAU 9950
CAAGGACUUC GGCGGCUUCA AUUUCAGCCA GAUUCUGCCC GAUCCUAGCA 10000
AGCCCAGCAA GCGGAGCUUC AUCGAGGACC UGCUGUUCAA CAAAGUGACA 10050
CUGGCCGACG CCGGCUUCAU CAAGCAGUAU GGCGAUUGUC UGGGCGACAU 10100
UGCCGCCAGG GAUCUGAUUU GCGCCCAGAA GUUUAACGGA CUGACAGUGC 10150
UGCCUCCUCU GCUGACCGAU GAGAUGAUCG CCCAGUACAC AUCUGCCCUG 10200
CUGGCCGGCA CAAUCACAAG CGGCUGGACA UUUGGAGCAG GCGCCGCUCU 10250
GCAGAUCCCC UUUGCUAUGC AGAUGGCCUA CCGGUUCAAC GGCAUCGGAG 10300
UGACCCAGAA UGUGCUGUAC GAGAACCAGA AGCUGAUCGC CAACCAGUUC 10350
AACAGCGCCA UCGGCAAGAU CCAGGACAGC CUGAGCAGCA CAGCAAGCGC 10400
CCUGGGAAAG CUGCAGGACG UGGUCAACCA GAAUGCCCAG GCACUGAACA 10450
CCCUGGUCAA GCAGCUGUCC UCCAACUUCG GCGCCAUCAG CUCUGUGCUG 10500
AACGAUAUCC UGAGCAGACU GGACCCUCCU GAGGCCGAGG UGCAGAUCGA 10550
CAGACUGAUC ACAGGCAGAC UGCAGAGCCU CCAGACAUAC GUGACCCAGC 10600
AGCUGAUCAG AGCCGCCGAG AUUAGAGCCU CUGCCAAUCU GGCCGCCACC 10650
AAGAUGUCUG AGUGUGUGCU GGGCCAGAGC AAGAGAGUGG ACUUUUGCGG 10700
CAAGGGCUAC CACCUGAUGA GCUUCCCUCA GUCUGCCCCU CACGGCGUGG 10750
UGUUUCUGCA CGUGACAUAU GUGCCCGCUC AAGAGAAGAA UUUCACCACC 10800
GCUCCAGCCA UCUGCCACGA CGGCAAAGCC CACUUUCCUA GAGAAGGCGU 10850
GUUCGUGUCC AACGGCACCC AUUGGUUCGU GACACAGCGG AACUUCUACG 10900

AGCCCCAGAU CAUCACCACC GACAACACCU UCGUGUCUGG CAACUGCGAC 10950
GUCGUGAUCG GCAUUGUGAA CAAUACCGUG UACGACCCUC UGCAGCCCGA 11000
GCUGGACAGC UUCAAAGAGG AACUGGACAA GUACUUUAAG AACCACACAA 11050
GCCCCGACGU GGACCUGGGC GAUAUCAGCG GAAUCAAUGC CAGCGUCGUG 11100
AACAUCCAGA AAGAGAUCGA CCGGCUGAAC GAGGUGGCCA AGAAUCUGAA 11150
CGAGAGCCUG AUCGACCUGC AAGAACUGGG GAAGUACGAG CAGUACAUCA 11200
AGUGGCCCUG GUACAUCUGG CUGGGCUUUA UCGCCGGACU GAUUGCCAUC 11250
GUGAUGGUCA CAAUCAUGCU GUGUUGCAUG ACCAGCUGCU GUAGCUGCCU 11300
GAAGGGCUGU UGUAGCUGUG GCAGCUGCUG CAAGUUCGAC GAGGACGAUU 11350
CUGAGCCCGU GCUGAAGGGC GUGAAACUGC ACUACACAUG AUGACUCGAG 11400
CUGGUACUGC AUGCACGCAA UGCUAGCUGC CCCUUUCCCG UCCUGGGUAC 11450
CCCGAGUCUC CCCCGACCUC GGGUCCCAGG UAUGCUCCCA CCUCCACCUG 11500
CCCCACUCAC CACCUCUGCU AGUUCCAGAC ACCUCCCAAG CACGCAGCAA 11550
UGCAGCUCAA AACGCUUAGC CUAGCCACAC CCCCACGGGA AACAGCAGUG 11600
AUUAACCUUU AGCAAUAAAC GAAAGUUUAA CUAAGCUAUA CUAACCCCAG 11650
GGUUGGUCAA UUUCGUGCCA GCCACACCGC GGCCGCAUGA AUACAGCAGC 11700
AAUUGGCAAG CUGCUUACAU AGAACUCGCG GCGAUUGGCA UGCCGCCUUA 11750
AAAUUUUUAU UUUAUUUUUU CUUUUCUUUU CCGAAUCGGA UUUUGUUUUU 11800
AAUAUUUCAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAGC AUAUGACUAA 11850
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 11900
AAAAAAAAAA AAAAAAAA 11918