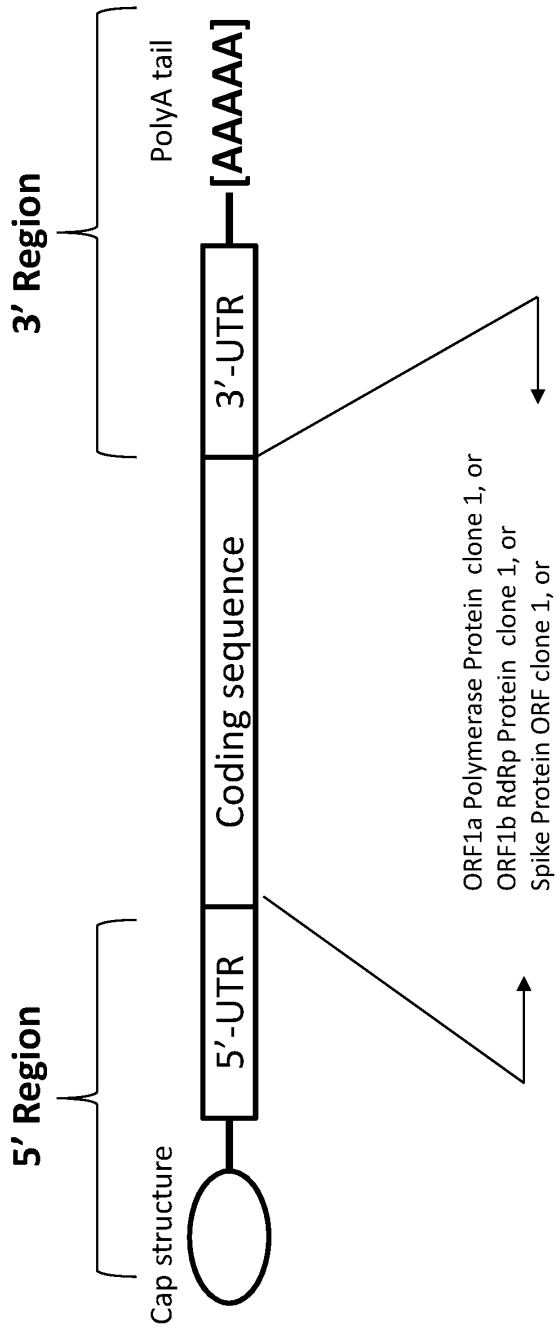


Figure 1. Schematic representation of optimized mRNA vaccine expression structure

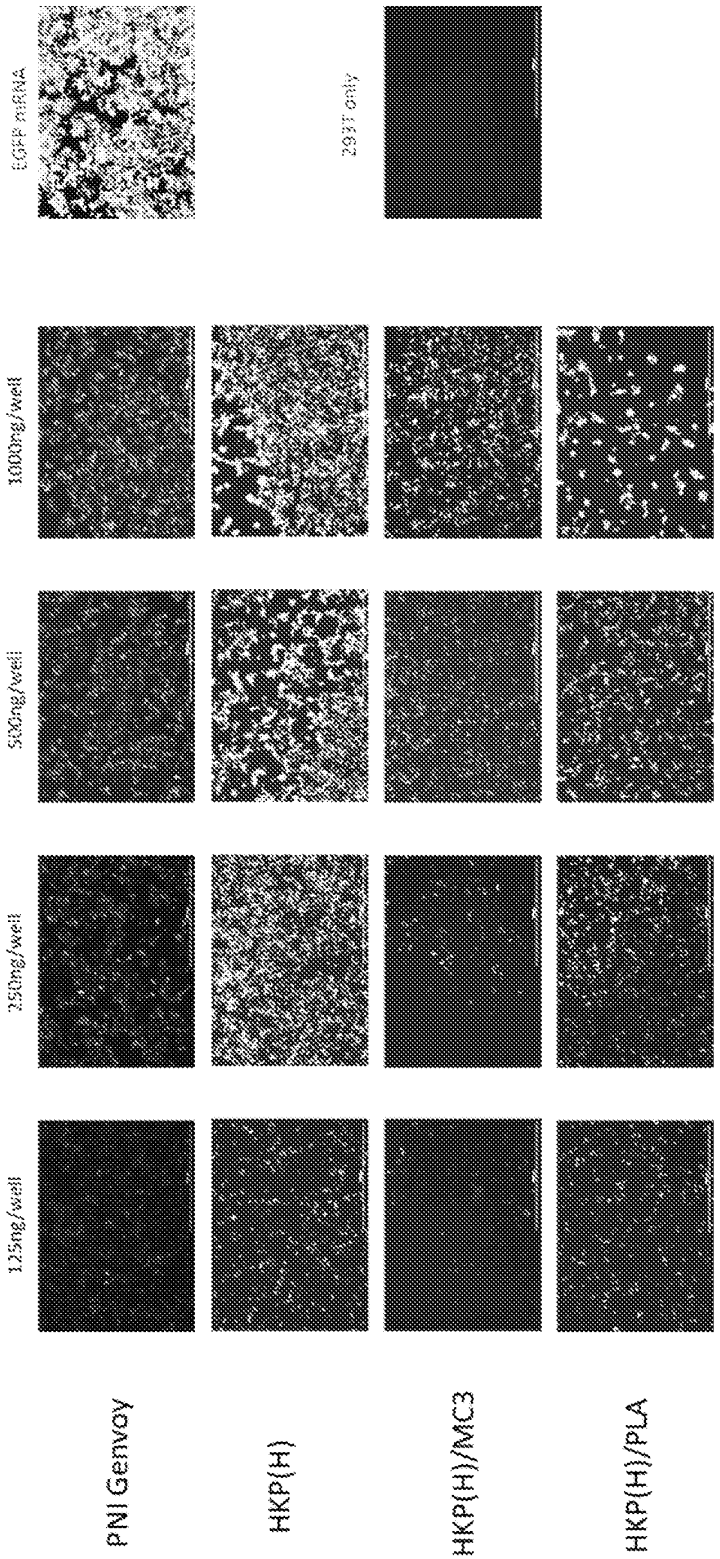


Figure 2. Immunofluorescence analysis result of Spike Protein expression mRNA transient transfection. Transfection condition : ThermoFisher Scientific Lipofectamine MessengerMAX Transfection Reagent protocol(Pin. No. MAN0010803 Rev D.0)

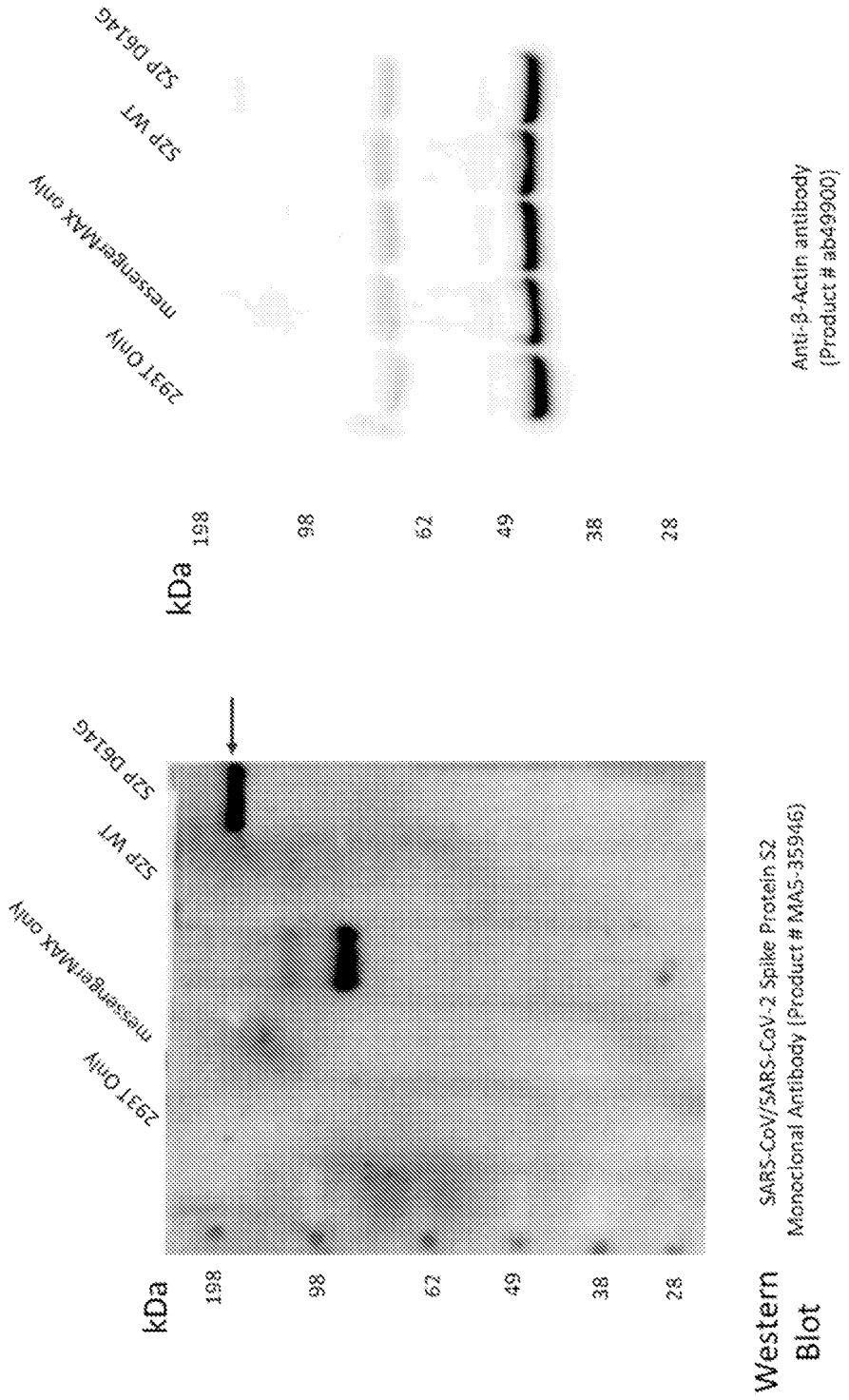


Figure 3. Western Blot result of Spike Protein expression mRNA transient transfection. Trasfection condition : ThermoFisher Scientific Lipofectamine MessengerMAX Transfection Reagent protocol(Pin. No. MAN0010803 Rev D.0). Primary antibody : SARS-CoV/SARS-CoV-2 Spike Protein S2 Monoclonal Antibody (ThermoFisher Scientific, MA5-35946). Secondary antibody : Mouse IgG (H+L) Secondary Antibody (ThermoFisher Scientific, A24512). Loading control : HRP Anti-beta Actin antibody (abcam, ab49900)

Figure 4:

Spike Protein ORF clone 1 (MN996531: viral cDNA: 21523-23040) SEQ ID NO:1

atgttctt gttaacaact 21541
aaacgaacaa tgtttgtttt tcttgtttta ttgccactag tctctagtca gtgtgttaat cttacaacca gaactcaatt
acccccctgca tacactaatt ctttcacacg tggtgtttat taccctgaca aagttttcag atcctcagtt ttacattcaa
ctcaggactt gttcttacct ttcttttcca atgttacttg gttccatgct atacatgtct ctgggaccaa tggtactaag
aggtttgata accctgtcct accatttaat gatggtgttt attttgcttc cactgagaag tctaacataa taagaggctg
gatttttggt actactttag attcgaagac ccagtcccta cttattgtta ataacgctac taatgttgtt attaaagtct
gtgaatttca attttgtaat gatccatttt tgggtgttta ttaccacaaa aacaacaaaa gttggatgga aagtgagttc
agagtttatt ctagtgcgaa taattgcact tttgaatatg tctctcagcc ttttcttatg gaccttgaag gaaaacaggg
taatttcaaa aatcttaggg aatttgtgtt taagaatatt gatggttatt ttaaaatata ttctaagcac acgcctatta
atttagtgcg tgatctccct cagggttttt cggctttaga accattggta gatttgccaa taggtattaa catcactagg
tttcaaactt tacttgcttt acatagaagt tatttgactc ctggtgattc ttcttcaggt tggacagctg gtgctgcagc
ttattatgtg ggttatcttc aacctaggac ttttctatta aaatataatg aaaatggaac cattacagat gctgtagact
gtgcacttga ccctctctca gaaacaaagt gtacgttgaa atccttcact gtagaaaaag gaatctatca aacttctaac
tttagagtcc aaccaacaga atctattgtt agatttccta atattacaaa cttgtgccct tttggtgaag tttttaacgc
caccagattt gcatctgttt atgcttggaa caggaagaga atcagcaact gtgttgctga ttattctgtc ctatataatt
ccgcatcatt ttccactttt aagtgttatg gagtgtctcc tactaaatta aatgatctct gctttactaa tgtctatgca
gattcatttg taattagagg tgatgaagtc agacaaatcg ctccagggca aactggaaag attgctgatt ataattataa
attaccagat gattttacag gctgcgttat agcttggaat tctaacaatc ttgattctaa ggttggtggt aattataatt
acctgtatag attgtttagg aagtctaatc tcaaaccttt tgagagagat atttcaactg aaatctatca ggccggtagc
acaccttgta atggtgttga aggttttaat tgttactttc ctttacaatc atatggtttc

Spike Protein ORF clone 2, (MN996531: viral cDNA: 23041-23740) SEQ ID NO:2

23041
caacccacta atggtgttgg ttaccaacca tacagagtag tagtactttc ttttgaactt ctacatgcac cagcaactgt
ttgtggacct aaaaagtcta ctaatttggt taaaaacaaa tgtgtcaatt tcaacttcaa tggtttaaca ggcacaggtg
ttcttactga gtctaacaaa aagtttctgc ctttccaaca atttggcaga gacattgctg acactactga tgctgtccgt
gatccacaga cacttgagat tcttgacatt acaccatgtt cttttggtgg tgtcagtgtt ataacaccag gaacaaatac
ttctaaccag gttgctgttc tttatcagga tgttaactgc acagaagtcc ctgttgctat tcatgcagat caacttactc
ctacttggcg tgtttattct acaggttcta atgtttttca aacacgtgca ggctgtttaa taggggctga acatgtcaac
aactcatatg agtgtgacat acccattggt gcaggtatat gcgctagtta tcagactcag actaattctc
ctcggcgggc acgtagtgta gctagtcaat ccatcattgc ctacactatg tcacttggtg cagaaaattc agttgcttac
tctaataact ctattgccat acccacaaat tttactatta gtgttaccac agaaattcta ccagtgtcta

Spike Protein ORF clone 3 (MN996531: viral cDNA: 23740-24540) SEQ ID NO:3

23740
tgaccaagac atcagtagat tgtacaatgt acatttgtgg tgattcaact gaatgcagca atcttttgtt gcaatatggc
agtttttgta cacaattaaa ccgtgcttta actggaatag ctgttgaaca agacaaaaac acccaagaag tttttgcaca
agtcaaacaa atttacaaaa caccaccaat taaagatttt ggtggtttta atttttcaca aatattacca gatccatcaa
aaccaagcaa gaggtcattt attgaagatc tacttttcaa caaagtgaca cttgcagatg ctggcttcat caaacaatat
ggtgattgcc ttggtgatat tgctgctaga gacctcattt gtgcacaaaa gtttaacggc cttactgttt tgccaccttt
gctcacagat gaaatgattg ctcaatacac ttctgcactg ttagcgggta caatcacttc tggttggacc tttggtgcag
gtgctgcatt acaaatacca tttgctatgc aaatggctta taggtttaat ggtattggag ttacacagaa tgttctctat
gagaaccaaa aattgattgc caaccaattt aatagtgcta ttggcaaaat tcaagactca ctttcttcca cagcaagtgc
acttggaaaa cttcaagatg tggtcaacca aaatgcacaa gctttaaaca cgcttgttaa acaacttagc tccaattttg
gtgcaatttc aagtgtttta aatgatatcc tttcacgtct tgacaaagtt gaggctgaag tgcaaattga taggttgatc

Spike Protein ORF clone 4 (MN996531: viral cDNA: 24540-25371) SEQ ID NO:4

24540
acaggcagac ttcaaagttt gcagacatat gtgactcaac aattaattag agctgcagaa atcagagctt
ctgctaatct tgctgctact aaaatgtcag agtgtgtact tggacaatca aaaagagttg atttttgtgg aaagggctat
catcttatgt ccttccctca gtcagcacct catggtgtag tcttcttgca tgtgacttat gtccctgcac aagaaaagaa
cttcacaact gctcctgcca tttgtcatga tggaaaagca cactttcctc gtgaaggtgt ctttgtttca aatggcacac
actggtttgt aacacaaagg aatttttatg aaccacaaat cattactaca gacaacacat ttgtgtctgg taactgtgat
gttgtaatag gaattgtcaa caacacagtt tatgatcctt tgcaacctga attagactca ttcaaggagg agttagataa
atattttaag aatcatacat caccagatgt tgatttaggt gacatctctg gcattaatgc ttcagttgta aacattcaaa
aagaaattga ccgcctcaat gaggttgcca agaatttaaa tgaatctctc atcgatctcc aagaacttgg
aaagtatgag cagtatataa aatggccatg gtacatttgg ctaggtttta tagctggctt gattgccata gtaatggtga
caattatgct ttgctgtatg accagttgct gtagttgtct caagggctgt tgttcttgtg gatcctgctg caaatttgat
gaagacgact ctgagccagt gctcaaagga gtcaaattac attacacata a

ORF1a Polymerase Protein clone 1 (MN996531: viral cDNA: 6253-7140) SEQ ID NO:5

6253
gccacgta taaaccaaat acctggtgta tacgttgtct ttggagcaca aaaccagttg aaacatcaaa ttcgtttgat
gtactgaagt cagaggacgc gcagggaatg gataatcttg cctgcgaaga tctaaaacca gtctctgaag
aagtagtgga aaatcctacc atacagaaag acgttcttga gtgtaatgtg aaaactaccg aagttgtagg
agacattata cttaaaccag caaataatag tttaaaaatt acagaagagg ttggccacac agatctaatg gctgcttatg
tagacaattc tagtcttact attaagaaac ctaatgaatt atctagagta ttaggtttga aaacccttgc tactcatggt
ttagctgctg ttaatagtgt cccttgggat actatagcta attatgctaa gccttttctt aacaaagttg ttagtacaac
tactaacata gttacacggt gtttaaaccg tgtttgtact aattatatgc cttatttctt tactttattg ctacaattgt
gtacttttac tagaagtaca aattctagaa ttaaagcatc tatgccgact actatagcaa agaatactgt taagagtgtc
ggtaaatttt gtctagaggc ttcatttaat tatttgaagt cacctaattt ttctaaactg ataaatatta taatttggtt

tttactatta agtgtttgcc taggttcttt aatctactca accgctgctt taggtgtttt aatgtctaat ttaggcatgc cttcttactg tactggttac agagaaggct atttgaactc tactaatgtc actattgcaa cctactgtac tggttctata ccttgtagtg tttgtcttag tggtttagat

ORF1b RdRp Protein clone 1 (MN996531: viral cDNA: 15253-16153) SEQ ID NO:6

15253
tagaaaac cctcacctta tgggttggga ttatcctaaa tgtgatagag ccatgcctaa catgcttaga attatggcct cacttgttct tgctcgcaaa catacaacgt gttgtagctt gtcacaccgt ttctatagat tagctaatga gtgtgctcaa gtattgagtg aaatggtcat gtgtggcggt tcactatatg ttaaaccagg tggaacctca tcaggagatg ccacaactgc ttatgctaat agtgttttta acatttgtca agctgtcacg gccaatgtta atgcactttt atctactgat ggtaacaaaa ttgccgataa gtatgtccgc aatttacaac acagacttta tgagtgtctc tatagaaata gagatgttga cacagacttt gtgaatgagt tttacgcata tttgcgtaaa catttctcaa tgatgatact ctctgacgat gctgttgtgt gtttcaatag cacttatgca tctcaaggtc tagtggctag cataaagaac tttaagtcag ttctttatta tcaaaacaat gtttttatgt ctgaagcaaa atgttggact gagactgacc ttactaaagg acctcatgaa ttttgctctc aacatacaat gctagttaaa cagggtgatg attatgtgta ccttccttac ccagatccat caagaatcct aggggccggc tgttttgtag atgatatcgt aaaaacagat ggtacactta tgattgaacg gttcgtgtct ttagctatag atgcttaccc acttactaaa catcctaatc aggagtatgc tgatgtcttt catttgtact tacaatacat aagaaagcta catgatgagt taacaggaca catgttagac atgtattctg ttatgcttac taa

M Protein ORF (MN996531: viral cDNA: 26510-27178) SEQ ID NO:7

a tggcagattc 26521 caacggtact attaccgttg aagagcttaa aaagctcctt gaacaatgga acctagtaat aggtttccta ttccttacat ggatttgtct tctacaattt gcctatgcca acaggaatag gtttttgtat ataattaagt taattttcct ctggctgtta tggccagtaa ctttagcttg ttttgtgctt gctgctgttt acagaataaa ttggatcacc ggtggaattg ctatcgcaat ggcttgtctt gtaggcttga tgtggctcag ctacttcatt gcttctttca gactgtttgc gcgtacgcgt tccatgtggt cattcaatcc agaaactaac attcttctca acgtgccact ccatggcact attctgacca gaccgcttct agaaagtgaa ctcgtaatcg gagctgtgat ccttcgtgga catcttcgta ttgctggaca ccatctagga cgctgtgaca tcaaggacct gcctaaagaa atcactgttg ctacatcacg aacgctttct tattacaaat tgggagcttc gcagcgtgta gcaggtgact caggttttgc tgcatacagt cgctacagga ttggcaacta taaattaaac acagaccatt ccagtagcag tgacaatatt gctttgcttg tacagtaa

N Protein ORF (MN996531: viral cDNA: 28261-29520) SEQ ID NO:8

28261
atgtctgata atggacccca aaatcagcga aatgcacccc gcattacgtt tggtggaccc tcagattcaa ctggcagtaa ccagaatgga gaacgcagtg gggcgcgatc aaaacaacgt cggccccaag gtttacccaa taatactgcg tcttggttca ccgctctcac tcaacatggc aaggaagacc ttaaattccc tcgaggacaa ggcgttccaa ttaacaccaa tagcagtcca gatgaccaaa ttggctacta ccgaagagct accagacgaa ttcgtggtgg tgacggtaaa atgaaagatc tcagtccaag atggtatttc tactacctag gaactgggcc agaagctgga cttccctatg gtgctaacaa agacggcatc atatgggttg caactgaggg agccttgaat

acaccaaaag atcacattgg cacccgcaat cctgctaaca atgctgcaat cgtgctacaa cttcctcaag
gaacaacatt gccaaaaggc ttctacgcag aagggagcag aggcggcagt caagcctctt ctcgttcctc
atcacgtagt cgcaacagtt caagaaattc aactccaggc agcagtaggg gaacttctcc tgctagaatg
gctggcaatg gcggtgatgc tgctcttgct ttgctgctgc ttgacagatt gaaccagctt gagagcaaaa
tgtctggtaa aggccaacaa caacaaggcc aaactgtcac taagaaatct gctgctgagg cttctaagaa
gcctcggcaa aaacgtactg ccactaaagc atacaatgta acacaagctt tcggcagacg tggtccagaa
caaacccaag gaaattttgg ggaccaggaa ctaatcagac aaggaactga ttacaaacat tggccgcaaa
ttgcacaatt tgcccccagc gcttcagcgt tcttcggaat gtcgcgcatt ggcatggaag tcacaccttc
gggaacgtgg ttgacctaca caggtgccat caaattggat gacaaagatc caaatttcaa agatcaagtc
attttgctga ataagcatat tgacgcatac aaaacattcc caccaacaga gcctaaaaag gacaaaaaga
agaaggctga tgaaactcaa gccttaccgc agagacagaa gaaacagcaa actgtgactc ttcttcctgc
tgcagatttg gatgatttct ccaaacaatt gcaacaatcc atgagcagtg ctgactcaac tcaggcctaa

E Protein ORF (MN996531: viral cDNA: 26232-26459) SEQ ID NO:9

26232
atgtactca ttcgtttcgg aagagacagg tacgttaata gttaatagcg tacttctttt tcttgctttc gtggtattct
tgctagttac actagccatc cttactgcgc ttcgattgtg tgcgtactgc tgcaatattg ttaacgtgag tcttgtaaaa
ccttcttttt acgtttactc tcgtgttaaa aatctgaatt cttctagagt tcctgatctt ctggtctaa

Figure 5

D614G Mutated Spike Protein ORF SEQ ID NO:10

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttcttttcc
aatgttacttggttccatgctatacatgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatgg
tgtttattttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctact
tattgttaataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttattaccacaaaaa
caacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttat
ggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttttaaaatatattctaa
gcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaaca
tcactaggtttcaaactttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcag
cttattatgtgggttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcac
ttgaccctctctcagaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtcca
accaacagaatctattgttagatttcctaatattacaaacttgtgccctttggtgaagtttttaacgccaccagatttgcatctgttt
atgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagac
aaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaa
ttctaacaatcttgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagaga
gatatttcaactgaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatg
gtttccaacccactaatggtgttggttaccaaccatacagagtagtagtacttttcttttgaacttctacatgcaccagcaactgttt
gtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttact
gagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgctgacactactgatgctgtccgtgatccacagaca
cttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgtt
ctttatcagggtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggt
tctaatgtttttcaaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtg
caggtatatgcgctagttatcagactcagactaattctcctcggcgggcacgtagtgtagctagtcaatccatcattgcctaca
ctatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccatacccacaaattttactattagtgttaccacag
aaattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttg
ttgcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttt
ttgcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaa
aaccaagcaagaggtcatttattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtg
attgccttggtgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagat
gaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtgctgcattacaaa
taccatttgctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgcc
aaccaatttaatagtgctattggcaaaattcaagactcactttcttccacagcaagtgcacttggaaaacttcaagatgtggtca
accaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatccttt
cacgtcttgacaaagttgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgact
caacaattaattagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatca
aaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgact
tatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtc
tttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacagacaacacatttgtgtc

tggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggag
ttagataaatattttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaa
aaagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagca
gtatataaaatggccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtat
gaccagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctc
aaaggagtcaaattacattacacataa

D614G Mutated Spike Protein Amino Acid Sequence SEQ ID NO:11

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQM
AYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGV
VFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEP
QIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDL
GDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLG
FIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHY
T

Figure 6

SARS RBD mRNA clone# 2874,
SARS2 Receptor Binding Domain (RBD)-foldon mRNA modified with Cleancap analogue and N1-Methyl-Pseudouridine-5'-Triphosphate. Codon optimized for human SEQ ID No.12

2874

ATTAATACGACTCACTATAAGGAGGCCACCATGTTCGTGTTTCTGGTGCTG
CTGCCTCTGGTGTCCAGCCAGCGGGTGCAGCCCACCGAATCCATCGTGCG
GTTCCCCAATATCACCAATCTGTGCCCCTTCGGCGAGGTGTTCAATGCCAC
CAGATTCGCCTCTGTGTACGCCTGGAACCGGAAGCGGATCAGCAATTGCG
TGGCCGACTACTCCGTGCTGTACAACTCCGCCAGCTTCAGCACCTTCAAGT
GCTACGGCGTGTCCCCTACCAAGCTGAACGACCTGTGCTTCACAAACGTG
TACGCCGACAGCTTCGTGATCCGGGGAGATGAAGTGCGGCAGATTGCCCC
TGGACAGACAGGCAAGATCGCCGACTACAACTACAAGCTGCCCGACGACT
TCACCGGCTGTGTGATTGCCTGGAACAGCAACAACCTGGACTCCAAAGTC
GGCGGCAACTACAATTACCTGTACCGGCTGTTCCGGAAGTCCAATCTGAA
GCCCTTCGAGCGGGACATCTCCACCGAGATCTATCAGGCCGGCAGCACCC
CTTGTAACGGCGTGGAAGGCTTCAACTGCTACTTCCCACTGCAGTCCTACG
GCTTTCAGCCCACAAATGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTG
CTGAGCTTCGAACTGCTGCATGCCCCTGCCACAGTGTGCGGCCCTAAGAA
AAGCACCAATCTCGTGAAGAACAAATGCGTGAACTTCGGATCCGGAGGTG
GATACATCCCGGAGGCCCCTAGGGACGGTCAAGCTTACGTGAGAAAGGA
CGGCGAATGGGTTCTGCTGTCGACCTTCTTGGGATAAGCAAGCTGCTTAC
ATAGAACTCGCGGCGATTGGCATGCCGCCTTAAAATTTTTATTTTATTTTTT
CTTTTCTTTTCCGAATCGGATTTTGTTTTTAATATTTCAAAAAAAAAAAAA
AAAAAAAAAAAACGCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAA

Protein sequence for D614G Mutated Spike Protein ORF SEQ ID NO:13

MFVFLVLLPLVSSQRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRI
SNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIA
PGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLK
PFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSF
ELLHAPATVCGPKKSTNLVKNKCVNFGSGGGYIPEAPRDGQAYVRKDGEWV
LLST- FLG

Figure 7

18 polypeptide epitopes for 2019-nCoV spike proteins

SEQ ID	start	stop	peptide
14	576	589	RDIADTTDAVRDPQ
15	416	429	VRQIAPGQTGKIAD
16	684	696	QTQTNSPRRARSV
17	258	270	LTPGDSSSGWTAG
18	1146	1157	VYDPLQPELDSF
19	814	825	ILPDPSKPSKRS
20	504	515	YGFQPTNGVGYQ
21	1166	1176	KNHTSPDVDLG
22	482	492	YQAGSTPCNGV
23	80	90	SGTNGTKRFDN
24	30	40	RTQLPPAYTNS
25	1265	1274	FDEDDSEPVL
26	797	806	IYKTPPIKDF
27	606	615	VITPGTNTSN
28	532	541	TVCGPKKSTN
29	1078	1086	PAQEKNFTT
30	781	789	VEQDKNTQE
31	448	456	NNLDSKVGG

SEQ ID NO:32 - 49

The mRNA sequences corresponding to 18 epitopes of 2019-nCoV spike protein

SEQ ID	start	stop	Nucleotide sequence
32	576	589	AGGGACATCGCCGACACCACCGACGCCGTGAGGGACCCCCAG
33	416	429	GTGAGGCAGATCGCCCCCGGCCAGACCGGCAAGATCGCCGAC
34	684	696	CAGACCCAGACCAACAGCCCCAGGAGGGCCAGGAGCGTG
35	258	270	CTGACCCCCGGCGACAGCAGCAGCGGCTGGACCGCCGGC
36	1146	1157	GTGTACGACCCCCTGCAGCCCGAGCTGGACAGCTTC
37	814	825	ATCCTGCCCGACCCCAGCAAGCCCAGCAAGAGGAGC
38	504	515	TACGGCTTCCAGCCCACCAACGGCGTGGGCTACCAG
39	1166	1176	AAGAACCACACCAGCCCCGACGTGGACCTGGGC
40	482	492	TACCAGGCCGGCAGCACCCCCTGCAACGGCGTG

41	80	90	AGCGGCACCAACGGCACCAAGAGGTTCGACAAC
42	30	40	AGGACCCAGCTGCCCCCCGCCTACACCAACAGC
43	1265	1274	TTCGACGAGGACGACAGCGAGCCCGTGCTG
44	797	806	ATCTACAAGACCCCCCCCATCAAGGACTTC
45	606	615	GTGATCACCCCCGGCACCAACACCAGCAAC
46	532	541	ACCGTGTGCGGCCCCAAGAAGAGCACCAAC
47	1078	1086	CCCGCCCAGGAGAAGAACTTCACCACC
48	781	789	GTGGAGCAGGACAAGAACACCCAGGAG
49	448	456	AACAACCTGGACAGCAAGGTGGGCGGC

SEQ ID NO: 50-55
Six unique insertions identified by aligning against other coronavirus recognized by the human immune system

SEQ ID	start	stop	peptide	nucleotide
50	153	167	YYHKNNKSWMESEFR	TACTACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCAGG
51	76	87	AIHVSGTNGTKR	GCCATCCACGTGAGCGGCACCAACGGCACCAAGAGG
52	252	266	ALHRSYLTPGDSSSG	GCCCTGCACAGGAGCTACCTGACCCCCGGCGACAGCAGCAGCGGC
53	447	457	NNLDSKVGGNY	AACAACCTGGACAGCAAGGTGGGCGGCAACTAC
54	479	495	TEIYQAGSTPCNGVEGF	ACCGAGATCTACCAGGCCGGCAGCACCCCCTGCAACGGCGTGGAGGGCTTC
55	683	692	QTQTNSPRRA	CAGACCCAGACCAACAGCCCCAGGAGGGCC

SEQ ID NO:56 - 66
The RNA sequence coded for various viral peptides serving as potential antigens for vaccination against 2019-nCoV viral infections.

SEQ ID	Start	Stop	Peptide Sequence	Nucleotide Sequence
56	606	615	VITPGTNTSN	GUGAUCACCCCCGGCACCAACACCAGCAAC
57	532	541	TVCGPKKSTN	ACCGUGUGCGGCCCCAAGAAGAGCACCAAC

58	1078	1086	PAQEKNFTT	CCCGCCCAGGAGAAGAACUUCACCACC
59	781	789	VEQDKNTQE	GUGGAGCAGGACAAGAACACCCAGGAG
60	448	456	NNLDSKVGG	AACAACCUGGACAGCAAGGUGGGCGGC
61	153	167	YYHKNNKSWMESEFR	UACUACCACAAGAACAACAAGAGCUGGAUGGAGAGCGAGUUCAGG
62	76	87	AIHVSGTNGTKR	GCCAUCCACGUGAGCGGCACCAACGGCACCAAGAGG
63	252	266	ALHRSYLTPGDSSSG	GCCCUGCACAGGAGCUACCUGACCCCCGGCGACAGCAGCAGCGGC
64	447	457	NNLDSKVGGNY	AACAACCUGGACAGCAAGGUGGGCGGCAACUAC
65	479	495	TEIYQAGSTPCNGVEGF	ACCGAGAUCUACCAGGCCGGCAGCACCCCCUGCAACGGCGUGGAGGGCUUC
66	683	692	QTQTNSPRRA	CAGACCCAGACCAACAGCCCCAGGAGGGCC

Figure 8

SEQ ID No. 67
D614G_S6P_DNA_sequence

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttcttttcc
aatgttacttggttccatgctatacatgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatgg
tgtttattttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctact
tattgttaataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttattaccacaaaaa
caacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttat
ggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttttaaaatatattctaa
gcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaaca
tcactaggtttcaaactttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcag
cttattatgtgggttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcac
ttgaccctctctcagaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtcca
accaacagaatctattgttagatttcctaatattacaaacttgtgccctttggtgaagtttttaacgccaccagatttgcatctgttt
atgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagac
aaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaa
ttctaacaatcttgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagaga
gatatttcaactgaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatg
gtttccaacccactaatggtgttggttaccaaccatacagagtagtagtacttttcttttgaacttctacatgcaccagcaactgttt
gtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttact
gagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgctgacactactgatgctgtccgtgatccacagaca
cttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgtt
ctttatcagggtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggt
tctaatgtttttcaaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtg
caggtatatgcgctagttatcagactcagactaattctccttcgcgggcaggtagtgtagctagtcaatccatcattgcctaca
ctatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccatacccacaaattttactattagtgttaccacag
aaattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttg
ttgcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttt
ttgcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaa
aaccaagcaagaggtcacctattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggt
gattgccttggtgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacaga
tgaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtcctgcattacaa
ataccatttcctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgc
caaccaatttaatagtgctattggcaaaattcaagactcactttcttccacaccaagtgcacttggaaaacttcaagatgtggtc
aaccaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatcctt
tcacgtcttgacccacctgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgac
tcaacaattaattagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatca
aaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgact
tatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtc
tttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacagacaacacatttgtgtc

tggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggag
ttagataaatattttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaa
aaagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagca
gtatataaaatggccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtat
gaccagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctc
aaaggagtcaaattacattacacataa

D614G_S6P_amino_acid_sequence SEQ ID No. 68

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPSRAGSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMA
YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNA
QALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYVTQ
QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF
LHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQII
TTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFI
AGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

Figure 9

RBD_trimer_DNA_sequence SEQ ID No. 69

Gccaccatgttcgtgtttctggtgctgctgcctctggtgtccagccagcgggtgcagcccaccgaatccatcgtgcggttcc
ccaatatcaccaatctgtgccccttcggcgaggtgttcaatgccaccagattcgcctctgtgtacgcctggaaccggaagcg
gatcagcaattgcgtggccgactactccgtgctgtacaactccgccagcttcagcaccttcaagtgctacggcgtgtcccct
accaagctgaacgacctgtgcttcacaaacgtgtacgccgacagcttcgtgatccggggagatgaagtgcggcagattgc
ccctggacagacaggcaagatcgccgactacaactacaagctgcccgacgacttcaccggctgtgtgattgcctggaaca
gcaacaacctggactccaaagtcggcggcaactacaattacctgtaccggctgttccggaagtccaatctgaagcccttcg
agcgggacatctccaccgagatctatcaggccggcagcacccttgtaacggcgtggaaggcttcaactgctacttcccac
tgcagtcctacggctttcagcccacaaatggcgtgggctatcagccctacagagtggtggtgctgagcttcgaactgctgca
tgcccctgccacagtgtgcggccctaagaaaagcaccaatctcgtgaagaacaaatgcgtgaacttcggatccggaggtg
gatacatcccggaggcccctagggacggtcaagcttacgtgagaaaggacggcgaatgggttctgctgtcgaccttcttg
ggataa

RBD_trimer_amino_acid_sequence SEQ ID No. 70

ATMFVFLVLLPLVSSQRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNR
KRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVR
QIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKS
NLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVV
LSFELLHAPATVCGPKKSTNLVKNKCVNFGSGGGYIPEAPRDGQAYVRKDGE
WVLLSTFLG

Figure 10

VUI-202012_01_S6P_DNA_sequence SEQ ID No. 71

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttcttttcc
aatgttacttggttccatgctatacatgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatgg
tgtttatttttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctact
tattgttaataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttattaccacaaaaa
caacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttat
ggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttttaaaatatattctaa
gcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaaca
tcactaggtttcaaactttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcag
cttattatgtgggttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcac
ttgaccctctctcagaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtcca
accaacagaatctattgttagatttcctaatattacaaacttgtgccctttggtgaagtttttaacgccaccagatttgcatctgttt
atgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagac
aaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaa
ttctaacaatcttgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagaga
gatatttcaactgaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatg
gtttccaacccacttatggtgttggttaccaaccatacagagtagtagtacttttcttttgaacttctacatgcaccagcaactgttt
gtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttact
gagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgatgacactactgatgctgtccgtgatccacagaca
cttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgtt
ctttatcagggtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggt
tctaatgtttttcaaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtg
caggtatatgcgctagttatcagactcagactaattctcattcgcgggcaggtagtgtagctagtcaatccatcattgcctaca
ctatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccatacccataaattttactattagtgttaccacaga
aattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttgtt
gcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttttt
gcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaaa
accaagcaagaggtcacctattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtg
attgccttggtgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagat
gaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtcctgcattacaaa
taccatttcctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgcc
aaccaatttaatagtgctattggcaaaattcaagactcactttcttccacaccaagtgcacttggaaaacttcaagatgtggtca
accaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatccttg
cacgtcttgacccacctgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgact
caacaattaattagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatca
aaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgact
tatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtc
tttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacacacaacacatttgtgtct

ggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggagt
tagataaatattttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaaa
aagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagcag
tatataaaatggccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtatg
accagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctca
aaggagtcaaattacattacacataa

VUI-202012_01_S6P_amino_acid_sequence SEQ ID No. 72

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTYGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIDDTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSHSRAGSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPINFTISVTTEILPVSMTKTSVDCTMYIC
GDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKD
FGGFNFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLIC
AQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMA
YRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNA
QALNTLVKQLSSNFGAISSVLNDILARLDPPEAEVQIDRLITGRLQSLQTYVTQ
QLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF
LHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQII
TTHNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLG
DISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFI
AGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

VUI-202012_del_S6P_DNA_sequence SEQ ID No. 73

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttcttttcc
aatgttacttggttccatgctatatctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatggtgtttatt
ttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctacttattgtta
ataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttaccacaaaaacaacaaaag
ttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttatggaccttgaa

ggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttattttaaaatatattctaagcacacgcct
attaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaacatcactaggtttc
aaactttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcagcttattatgtggg
ttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcacttgaccctctctc
agaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtccaaccaacagaat
ctattgttagatttcctaatattacaaacttgtgccctttggtgaagtttttaacgccaccagatttgcatctgtttatgcttggaac
aggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgttatggagtgtctc
ctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagacaaatcgctcca
gggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaattctaacaatct
tgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagagagatatttcaact
gaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatggtttccaaccc
acttatggtgttggttaccaaccatacagagtagtagtactttctttgaacttctacatgcaccagcaactgtttgtggacctaa
aaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttactgagtctaaca
aaaagtttctgcctttccaacaatttggcagagacattgatgacactactgatgctgtccgtgatccacagacacttgagattct
tgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgttctttatcagg
gtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggttctaatgttttt
caaacacgtgcaggctgtttaataggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtgcaggtatatg
cgctagttatcagactcagactaattctcattcgcgggcaggtagtgtagctagtcaatccatcattgcctacactatgtcactt
ggtgcagaaaattcagttgcttactctaataactctattgccatacccataaattttactattagtgttaccacagaaattctacca
gtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttgttgcaatatgg
cagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagtttttgcacaagt
caaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaaaaccaagca
agaggtcacctattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtgattgccttgg
tgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagatgaaatgattg
ctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtcctgcattacaaataccatttcct
atgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgccaaccaatttaa
tagtgctattggcaaaattcaagactcactttcttccacaccaagtgcacttggaaaacttcaagatgtggtcaaccaaaatgc
acaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatccttgcacgtcttga
cccacctgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgactcaacaattaa
ttagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatcaaaaagagttg
atttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgacttatgtccctgc
acaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtctttgtttcaaat
ggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacacacaacacatttgtgtctggtaactgtg
atgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggagttagataaatat
tttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaaaaagaaattga
ccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagcagtatataaaatg
gccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtatgaccagttgctg
tagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctcaaaggagtcaa
attacattacacataa

VUI-202012_del_S6P_amino_acid_sequence SEQ ID No. 74

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAISGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTT
LDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYHKNNKSWMESEFRVYSS
ANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVG
YLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQ
PTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFS
TFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPD
DFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPC
NGVEGFNCYFPLQSYGFQPTYGVGYQPYRVVVLSFELLHAPATVCGPKKSTN
LVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIDDTTDAVRDPQTLEIL
DITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWRVYST
GSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSHSRAGSVASQSI
IAYTMSLGAENSVAYSNNSIAIPINFTISVTTEILPVSMTKTSVDCTMYICGDST
ECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGF
NFSQILPDPSKPSKRSPIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQK
FNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGPALQIPFPMQMAYRFN
GIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTPSALGKLQDVVNQNAQALN
TLVKQLSSNFGAISSVLNDILARLDPPEAEVQIDRLITGRLQSLQTYVTQQLIRA
AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVT
YVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTHN
TFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGI
NASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLI
AIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

Figure 11

SEQ ID No.75 – 88

Polymer	Branch Sequence	SEQ ID No.
-	R= $KH_nKH_nKH_nKH_nK$–	75
-	R= $H_nKH_nKH_nKH_nKH_nK$–	76
-	R= $KH_nKH_nKH_nKH_nKH_n$–	77
-	R= $H_nKH_nKH_nKH_nKH_nKH_n$–	78
H2K4b	R_A= KHKHHKHHKHHKHHKHHKHK–	79
H3K4b	R_B= KHHHKHHHKHHHKHHHK–	80
H3K(+H)4b	R_C= KHHHKHHHKHHHHKHHHK–	81
H3k(+H)4b	R_D= kHHHkHHHkHHHHkHHHk–	82
H-H3K(+H)4b	R_E= HKHHHKHHHKHHHHKHHHK–	83
HH-H3K(+H)4b	R_F= HHKHHHKHHHKHHHHKHHHK–	84
H4K4b	R_G= KHHHHKHHHHKHHHHKHHHHK–	85
H3K(1+H)4b	R_H= KHHHKHHHKHHHKHHHHK–	86
H3K(3+H)4b	R_I= KHHHKHHHKHHHKHHHK–	87
H3K(1,3+H)4b	R_J= KHHHKHHHKHHHKHHHHK–	88

5

Figure 12

D614G S2P DNA Sequence SEQ ID No.89:

atgtttgtttttcttgttttattgccactagtctctagtcagtgtgttaatcttacaaccagaactcaattacccccctgcatacactaat
tctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagttttacattcaactcaggacttgttcttacctttctttttcc
aatgttacttggttccatgctatacatgtctctgggaccaatggtactaagaggtttgataaccctgtcctaccatttaatgatgg
tgtttatttttgcttccactgagaagtctaacataataagaggctggatttttggtactactttagattcgaagacccagtccctact
tattgttaataacgctactaatgttgttattaaagtctgtgaatttcaattttgtaatgatccatttttgggtgtttattaccacaaaaa
caacaaaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttttcttat
ggaccttgaaggaaaacagggtaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttatttaaaatatattctaa
gcacacgcctattaatttagtgcgtgatctccctcagggtttttcggctttagaaccattggtagatttgccaataggtattaaca
tcactaggtttcaaactttacttgctttacatagaagttatttgactcctggtgattcttcttcaggttggacagctggtgctgcag
cttattatgtgggttatcttcaacctaggacttttctattaaaatataatgaaaatggaaccattacagatgctgtagactgtgcac
ttgaccctctctcagaaacaaagtgtacgttgaaatccttcactgtagaaaaaggaatctatcaaacttctaactttagagtcca
accaacagaatctattgttagatttcctaatattacaaacttgtgcccttttggtgaagtttttaacgccaccagatttgcatctgttt
atgcttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccacttttaagtgtt
atggagtgtctcctactaaattaaatgatctctgctttactaatgtctatgcagattcatttgtaattagaggtgatgaagtcagac
aaatcgctccagggcaaactggaaagattgctgattataattataaattaccagatgattttacaggctgcgttatagcttggaa
ttctaacaatcttgattctaaggttggtggtaattataattacctgtatagattgtttaggaagtctaatctcaaacctttgagaga
gatatttcaactgaaatctatcaggccggtagcacaccttgtaatggtgttgaaggttttaattgttactttcctttacaatcatatg
gtttccaacccactaatggtgttggttaccaaccatacagagtagtagtacttttcttttgaacttctacatgcaccagcaactgttt
gtggacctaaaaagtctactaatttggttaaaaacaaatgtgtcaatttcaacttcaatggtttaacaggcacaggtgttcttact
gagtctaacaaaaagtttctgcctttccaacaatttggcagagacattgctgacactactgatgctgtccgtgatccacagaca
cttgagattcttgacattacaccatgttcttttggtggtgtcagtgttataacaccaggaacaaatacttctaaccaggttgctgtt
ctttatcagggtgttaactgcacagaagtccctgttgctattcatgcagatcaacttactcctacttggcgtgtttattctacaggt
tctaatgtttttcaaacacgtgcaggctgtttaatagggggctgaacatgtcaacaactcatatgagtgtgacatacccattggtg
caggtatatgcgctagttatcagactcagactaattctccttcgcgggcaggtagtgtagctagtcaatccatcattgcctaca
ctatgtcacttggtgcagaaaattcagttgcttactctaataactctattgccatacccacaaattttactattagtgttaccacag
aaattctaccagtgtctatgaccaagacatcagtagattgtacaatgtacatttgtggtgattcaactgaatgcagcaatcttttg
ttgcaatatggcagtttttgtacacaattaaaccgtgctttaactggaatagctgttgaacaagacaaaaacacccaagaagttt
ttgcacaagtcaaacaaatttacaaaacaccaccaattaaagattttggtggttttaatttttcacaaatattaccagatccatcaa
aaccaagcaagaggtcatttattgaagatctacttttcaacaaagtgacacttgcagatgctggcttcatcaaacaatatggtg
attgccttggtgatattgctgctagagacctcatttgtgcacaaaagtttaacggccttactgttttgccacctttgctcacagat
gaaatgattgctcaatacacttctgcactgttagcgggtacaatcacttctggttggacctttggtgcaggtgctgcattacaaa
taccatttgctatgcaaatggcttataggtttaatggtattggagttacacagaatgttctctatgagaaccaaaaattgattgcc
aaccaatttaatagtgctattggcaaaattcaagactcactttcttccacagcaagtgcacttggaaaacttcaagatgtggtca
accaaaatgcacaagctttaaacacgcttgttaaacaacttagctccaattttggtgcaatttcaagtgttttaaatgatatccttt
cacgtcttgacccacctgaggctgaagtgcaaattgataggttgatcacaggcagacttcaaagtttgcagacatatgtgact
caacaattaattagagctgcagaaatcagagcttctgctaatcttgctgctactaaaatgtcagagtgtgtacttggacaatca
aaaagagttgatttttgtggaaagggctatcatcttatgtccttccctcagtcagcacctcatggtgtagtcttcttgcatgtgact
tatgtccctgcacaagaaaagaacttcacaactgctcctgccatttgtcatgatggaaaagcacactttcctcgtgaaggtgtc
tttgtttcaaatggcacacactggtttgtaacacaaaggaatttttatgaaccacaaatcattactacagacaacacatttgtgtc
tggtaactgtgatgttgtaataggaattgtcaacaacacagtttatgatcctttgcaacctgaattagactcattcaaggaggag

ttagataaatattttaagaatcatacatcaccagatgttgatttaggtgacatctctggcattaatgcttcagttgtaaacattcaa
aaagaaattgaccgcctcaatgaggttgccaagaatttaaatgaatctctcatcgatctccaagaacttggaaagtatgagca
gtatataaaatggccatggtacatttggctaggttttatagctggcttgattgccatagtaatggtgacaattatgctttgctgtat
gaccagttgctgtagttgtctcaagggctgttgttcttgtggatcctgctgcaaatttgatgaagacgactctgagccagtgctc
aaaggagtcaaattacattacacataa

Figure 13

D614G S2P amino acid sequence SEQ ID No.90

MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQD
LFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFG
TTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRV
YSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINL
VRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAY
YVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNF
RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYK
LPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGS
TPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK
STNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQT
LEILDITPCSFGGVSVITPGTNTSNQVAVLYQGVNCTEVPVAIHADQLTPTWR
VYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPSRAGSV
ASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYI
CGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLI
CAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQM
AYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQ
NAQALNTLVKQLSSNFGAISSVLNDILSRLDPPEAEVQIDRLITGRLQSLQTYV
TQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGV
VFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEP
QIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDL
GDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLG
FIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHY
T

Figure 14

SEQ ID No.90 to 95

SEQ ID	start	stop	nucleotide
90	153	167	TACTACCACAAGAACAACAAGAGCTGGATG GAGAGCGAGTTCAGG
91	76	87	GCCATCCACGTGAGCGGCACCAACGGCACC AAGAGG
92	252	266	GCCCTGCACAGGAGCTACCTGACCCCCGGC GACAGCAGCAGCGGC
93	447	457	AACAACCTGGACAGCAAGGTGGGCGGCAAC TAC
94	479	495	ACCGAGATCTACCAGGCCGGCAGCACCCCC TGCAACGGCGTGGAGGGCTTC
95	683	692	CAGACCCAGACCAACAGCCCCAGGAGGGCC

SEQ ID No. 96 to 106

SEQ ID	Start	Stop	Nucleotide Sequence
96	606	615	GUGAUCACCCCCGGCACCAACACCAGCAA C
97	532	541	ACCGUGUGCGGCCCCAAGAAGAGCACCAA C
98	1078	1086	CCCGCCCAGGAGAAGAACUUCACCACC
99	781	789	GUGGAGCAGGACAAGAACACCCAGGAG
100	448	456	AACAACCUGGACAGCAAGGUGGGCGGC
101	153	167	UACUACCACAAGAACAACAAGAGCUGGAU GGAGAGCGAGUUCAGG
102	76	87	GCCAUCCACGUGAGCGGCACCAACGGCAC CAAGAGG
103	252	266	GCCCUGCACAGGAGCUACCUGACCCCCGG CGACAGCAGCAGCGGC
104	447	457	AACAACCUGGACAGCAAGGUGGGCGGCAA CUAC

105	479	495	ACCGAGAUCUACCAGGCCGGCAGCACCCC CUGCAACGGCGUGGAGGGCUUC
106	683	692	CAGACCCAGACCAACAGCCCCAGGAGGGC C