

SEQUENCE LISTING

SEQ ID NO: 1

HUP promoter from Chlorella (subsequence of GenBank accession number X55349)

GATCAGACGGGCCTGACCTGCGAGATAATCAAGTGCTCGTAGGCAACCAACTCA
GCAGCTGCTTGGTGTGGGTCTGCAGGATAGTGTGCAGGGCCCCAAGGACAGC
AGGGGAACCTTACACCTTGTCCCCGACCCAGTTTATGGAGTGCATTGCCTCAAGA
GCCTAGCCGGAGCGCTAGGCTACATACTTGCCGCACCGGTATGAGGGGATATAG
TACTCGCACTGCGCTGTCTAGTGAGATGGGCAGTGCTGCCCATAAACAACCTGGCT
GCTCAGCCATTTGTTGGCGGACCATTTCTGGGGGGGCCAGCAATGCCTGACTTTTCG
GGTAGGGTGAACAACTGAACAAAGACTACCAAAACAGAATTTCTTCCTCCTTGGA
GGTAAGCGCAGGCCGGCCCGCCTGCGCCACATGGCGCTCCGAACACCTCCATA
GCTGTAAGGGCGCAAACATGGCCGGACTGTTGTCAGCACTCTTTCATGGCCATAC
AAGGTCATGTCGAGATTAGTGCTGAGTAAGACACTATCACCCCATGTTTCGATTGA
AGCCGTGACTTCATGCCAACCTGCCCCTGGGCGTAGCAGACGTATGCCATCATGA
CCACTAGCCGACATGCGCTGTCTTTTGCCACCAAAACAACCTGGTACACCGCTCGA
AGTCGTGCCGCACACCTCCGGGAGTGAGTCCGGCGACTCCTCCCCGGCGGGCCG
CGGCCCTACCTGGGTAGGGTCGCCATACGCCACGACCAAAACGACGCAGGAGGG
GATTGGGGTAGGGAATCCCAACCAGCCTAACCAAGACGGCACCTATAATAATAG
GTGGGGGGGACTAACAGCCCTATATCGCAAGCTTTGGGTGCCTATCTTGAGAAGCA
CGAGTTGGAGTGGCTGTGTACGGTCGACCCTAAGGTGGGTGTGCCGCAGCCTGA
AACAAAGCGTCTAGCAGCTGCTTCTATAATGTGTCAGCCGTTGTGTTTCAGTTAT
ATTGTATGCTATTGTTTGTTCGTGCTAGGGTGGCGCAGGCCACCTACTGTGGCG
GGCCATTGGTTGGTGCTTGAATTGCCTCACCATCTAAGGTCTGAACGCTCACTCA
AACGCCTTTGTACAACTGCAGAACTTTCCTTGGCGCTGCAACTACAGTGTGCAAA
CCAGCACATAGCACTCCCTTACATCACCCAGCAGTACAACA

SEQ ID NO: 2

Chlorella ellipsoidea nitrate reductase promoter from AY307383

CGCTGCGCACCCAGGGCCGCCAGCTCGCTGATGTCGCTCCAAATGCGGTCCCCGA
TTTTTTGTTCTTCATCTTCTCCACCTTGGTGGCCTTCTTGGCCAGGGCCCTTCAGCTG
CATGCGCACAGACCGTTGAGCTCCTGATCAGCATCCTCAGGAGGCCCTTTGACAA
GCAAGCCCCTGTGCAAGCCCATTACAGGGGTACCAGTGGTGCTGAGGTAGATGG
GTTTGAAAAGGATTGCTCGGTTCGATTGCTGCTCATGGAATTGGCATGTGCATGCA
TGTTTACAATATGCCACCAGGCTTTGGAGCAAGAGAGCATGAATGCCTTCAGGC
AGGTTGAAAGTTCCTGGGGGTGAAGAGGCAGGGCCGAGGATTGGAGGAGGAAA
GCATCAAGTCGTCGCTCATGCTCATGTTTTTCAGTCAGAGTTTGCCAAGCTCACAG
GAGCAGAGACAAGACTGGCTGCTCAGGTGTTGCATCGTGTGTGTGGTGGGGGGG
GGGGGGTTAATACGGTACGAAATGCACTTGGAATTCCCACCTCATGCCAGCGGA
CCCACATGCTTGAATTCGAGGCCTGTGGGGTGAGAAATGCTCACTCTGCCCTCGT
TGCTGAGGTACTTCAGGCCGCTGAGCTCAAAGTCGATGCCCTGCTCGTCTATCAG
GGCCTGCACCTCTGGGCTGACCGGCTCAGCCTCCTTCGCGGGCATGGAGTAGGCG
CCGGCAGCGTTTCATGTCCGGGCCCAGGGCAGCGGTGGTGCCATAAATGTCGGTG
ATGGTGGGGAGGGGGGGCCGTCGCCACACCATTGCCGTTGCTGGCTGACGCATGC
ACATGTGGCCTGGCTGGCACCGGCAGCACTGGTCTCCAGCCAGCCAGCAAGTGG
CTGTTCAAGGAAAGCGGCCATGTTGTTGGTCCCTGCGCATGTAATCCCCAGATCA
AAGGAGGGAACAGCTTGGATTTGATGTAGTGCCCAACCGGACTGAATGTGCGAT

GGCAGGTCCCTTTGAGTCTCCCGAATTACTAGCAGGGCACTGTGACCTAACGCAG
CATGCCAACCGCAAAAAAATGATTGACAGAAAATGAAGCGGTGTGTCAATATTT
GCTGTATTTATTCGTTTTAATCAGCAACCAAGTTCGAAACGCAACTATCGTGGTG
ATCAAGTGAACCTCATCAGACTTACCTCGTTCCGGCAAGGAAACGGAGGCACCAA
ATTCCAATTTGATATTATCGCTTGCCAAGCTAGAGCTGATCTTTGGGAAACCAAC
TGCCAGACAGTGGACTGTGATGGAGTGCCCCGAGTGGTGGAGCCTCTTCGATTTCG
GTTAGTCATTACTAACGTGAACCCTCAGTGAAGGGACCATCAGACCAGAAAAGAC
CAGATCTCCTCCTCGACACCGAGAGAGTGTTCGCGGCAGTAGGACGACAAG

SEQ ID NO: 3

Yeast sucrose invertase

MTNETSDRPLVHFTPNKGMNDPNGLWYDEKDAKWHL YFQYNPNDTVWGTPLF
WGHATSDDL TNWEDQPIAIAPKRND SGAFSGSMVVDYNN TSGFFNDTIDPRQRCVAI
WTYNTPESEEQYISYSLDGGYTFTEYQKNPVLAANSTQFRDPKVFWEPSQKWIMT
AAKSQDYKIEIYSSDDLKSWKLESAFANEGFLGYQYECPLIEVPTEQDPSKSYWVM
FISINPGAPAGGSFNQYFVGSFNGTHFEAFDNQSRVVD FGDYALQTFNTDPTYGS
ALGIAWASNWEYSAFVPTNPWRSSMSLVKFSLNTEYQANPETELINLKAEPILNISN
AGPWSRFATNTTLTKANSYNVDLSNSTGTLEFELVYA VNTTQTISKSVFADLSLWFK
GLEDPEEYLRMGFEVSASSFFLDRGNSKV KFKENPYFTNRMSVNNQPFKSENDLSY
YKVYGLLDQNILELYFNDGDVVSTNTYFMTTGNALGSVNMTTGVDNLFYIDKFQVR
EVK

SEQ ID NO: 4

Yeast secretion signal

MLLQAFLFLLAGFAAKISAS

SEQ ID NO: 5

Higher plants secretion signal

MANKSLLLLLLLLGSLASG

SEQ ID NO: 6

Consensus eukaryotic secretion signal

MARLPLAALG

SEQ ID NO: 7

Combination higher plant/eukaryotic secretion signal

MANKLLLLLLLLLLPLAASG

SEQ ID NO: 8

S. cerevisiae sucrose invertase NP_012104

GAATTCCCCAACATGGTGGAGCACGACACTCTCGTCTACTCCAAGAATATCAAAG
ATACAGTCTCAGAAGACCAAAGGGCTATTGAGACTTTTCAACAAAGGGTAATAT
CGGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTCATCAAAAGGAC
AGTAGAAAAGGAAGGTGGCACCTACAAATGCCATCATTGCGATAAAGGAAAGGC
TATCGTTCAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACG
AGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGAT

TGATGTGAACATGGTGGAGCACGACACTCTCGTCTACTCCAAGAATATCAAAGAT
ACAGTCTCAGAAGACCAAAGGGCTATTGAGACTTTTCAACAAAGGGTAATATCG
GGAAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTCATCAAAAAGGACAG
TAGAAAAGGAAGGTGGCACCTACAAATGCCATCATTGCGATAAAGGAAAGGCTA
TCGTTCAAGATGCCTCTGCCGACAGTGGTCCCAAGATGGACCCCCACCCACGAG
GAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCCTTCAAAGCAAGTGGATTG
ATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAA
GACCCTTCCTCTATATAAGGAAGTTCATTTTCATTTGGAGAGGACACGCTGAAATC
ACCAGTCTCTCTCTACAAATCTATCTCTGGCGCGCCATATCAATGCTTCTTCAGGC
CTTTCTTTTTCTTCTTGCTGGTTTTGCTGCCAAGATCAGCGCCTCTATGACGAACG
AAACCTCGGATAGACCACTTGTGCACCTTACACCAAACAAGGGCTGGATGAATG
ACCCCAATGGACTGTGGTACGACGAAAAAGATGCCAAGTGGCATCTGTACTTTC
AATACAACCCGAACGATACTGTCTGGGGGACGCCATTGTTTTGGGGGCCACGCCAC
GTCCGACGACCTGACCAATTGGGAGGACCAACCAATAGCTATCGCTCCGAAGAG
GAACGACTCCGGAGCATTCTCGGGTTCATGGTGGTTGACTACAACAATACTTCC
GGCTTTTTCAACGATACCATTGACCCGAGACAACGCTGCGTGGCCATATGGACTT
ACAACACACCGGAGTCCGAGGAGCAGTACATCTCGTATAGCCTGGACGGTGGAT
ACACTTTTACAGAGTATCAGAAGAACCCTGTGCTTGCTGCAAATTCGACTCAGTT
CCGAGATCCGAAGGTCTTTTGGTACGAGCCCTCGCAGAAGTGGATCATGACAGC
GGCAAAGTCACAGGACTACAAGATCGAAATTTACTCGTCTGACGACCTTAAATCC
TGGAAGCTCGAATCCGCGTTCGCAAACGAGGGCTTTCTCGGCTACCAATACGAAT
GCCCAGGCCTGATAGAGGTCCCAACAGAGCAAGATCCCAGCAAGTCTTACTGGG
TGATGTTTTATTTCCATTAATCCAGGAGCACCGGCAGGAGGTTCTTTTAATCAGTA
CTTCGTGCGGAAGCTTTAACGGAACCTCATTTTCGAGGCATTTGATAACCAATCAAGA
GTAGTTGATTTTGGAAAGGACTACTATGCCCTGCAGACTTTCTTCAATACTGACC
CGACCTATGGGAGCGCTCTTGGCATTGCGTGGGCTTCTAACTGGGAGTATTCCGC
ATTCGTTCTTACAAACCTTGGAGGTCTTCCATGTGCTCGTCTGAGGAAATTCTCTC
TCAACACTGAGTACCAGGCCAACCCGGAAACCGAACTCATAAACCTGAAAGCCG
AACCGATCCTGAACATTAGCAACGCTGGCCCCCTGGAGCCGGTTTGCAACCAACA
CCACGTTGACGAAAGCCAACAGCTACAACGTCGATCTTTCGAATAGCACC GGTA
CACTTGAATTTGAACTGGTGTATGCCGTCAATACCACCCAAACGATCTCGAAGTC
GGTGTTCGCGGACCTCTCCCTCTGGTTTTAAAGGCCTGGAAGACCCCGAGGAGTAC
CTCAGAATGGGTTTCGAGGTTTCTGCGTCCTCCTTCTTCTTCTTGATCGCGGGAACAG
CAAAGTAAAATTTGTTAAGGAGAACCCATATTTTACCAACAGGATGAGCGTTAA
CAACCAACCATTCAAGAGCGAAAACGACCTGTGCTACTACAAAGTGTATGGTTT
GCTTGATCAAAATATCCTGGAACCTCTACTTCAACGATGGTGTATGTCGTGTCCACC
AACACATACTTCATGACAACCGGGAACGCACTGGGCTCCGTGAACATGACGACG
GGTGTGGATAACCTGTTCTACATCGACAAATTCAGGTGAGGGAAGTCAAGTGA
GATCTGTGATCGACAAGCTCGAGTTTCTCCATAATAATGTGTGAGTAGTTCCCA
GATAAGGGAATTAGGGTTCCTATAGGGTTTCGCTCATGTGTTGAGCATATAAGAA
ACCTTAGTATGTATTTGTATTTGTAAAATACTTCTATCAATAAAAATTTCTAATTC
CTAAAACCAAAAATCCAGTACTAAAATCCAGATCCCCCGAATTAA

SEQ ID NO: 9

TGTTGAAGAATGAGCCGCGAC

SEQ ID NO: 10

CAGTGAGCTATTACGCACTC

SEQ ID NO: 11UTEX 329 *Prototheca kruegani*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGAAAATACT
CTGGAGCCATAGCGAAAGCAAGTTTAGTAAGCTTAGGTCATTCTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTAGTAAAATAACATGGAG
GCCCCAACCGACTAATGTTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAGTACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGAATAAGGATGTGGGTATGTCAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 12UTEX 1440 *Prototheca wickerhamii*

TGTTGAAGAATGAGCCGGCGACTTAAAAATAAATGGCAGGCTAAGAGATTTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCTTAATAGGGCGTCAATTTAACAAAACCTT
TAAATAAATTATAAAGTCATTTATTTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC
TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACCAGTTAAGGCCCAAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 13UTEX 1442 *Prototheca stagnora*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAAAATGGCATGGTTAAAGATATTTCT
CTGAAGCCATAGCGAAAGCAAGTTTTACAAGCTATAGTCATTTTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTCAAATAACATGGAG
GCCCCAACCGACTAATGGTGA AAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGAATAAGGACGTGAGTATGTCAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 14UTEX 288 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGATAATTCT
CTGGAGCCATAGCGAAAGCAAGTTTACAAGCTAAAGTCACCCTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG

GCCCGAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGACTAAGGATGTGGGTATGTTAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 15

UTEX 1439, UTEX 1441, UTEX 1435, UTEX 1437 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGACTTAAAATAAAATGGCAGGCTAAGAGAATTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCTTAATAGGGCGCTAATTTAACAAAACAT
TAAATAAAAATCTAAAGTCATTTATTTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC
TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACCAGTTAAGGCCCCAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 16

UTEX 1533 *Prototheca wickerhamii*

TGTTGAAGAATGAGCCGTCGACTTAAAATAAAATGGCAGGCTAAGAGAATTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCTTAATAGGGCGCTAATTTAACAAAACAT
TAAATAAAAATCTAAAGTCATTTATTTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC
TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACCAGTTAAGGCCCCAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 17

UTEX 1434 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCGTGGTTAAAGAAAATTCT
CTGGAACCATAGCGAAAGCAAGTTTAACAAGCTTAAGTCACTTTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG
GCCCGAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTCCGAAAG
TTGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGA
CCTTGGGGGATAAGCTCCTTGGTCGAAAGGGAAACAGCCCAGATCACCAGTTAA

GGCCCCAAAATGAAAATGATAGTGACTAAGGATGTGAGTATGTCAAAACCTCCA
GCAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 18

UTEX 1438 *Prototheca zopfii*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGAAAATTCT
CTGGAGCCATAGCGAAAGCAAGTTTAACAAGCTTAAGTCACTTTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG
GCCCGAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTCGTGGGCTTCGAAAG
TTGTACCTCAAAGTGGCAAACCTCTGAATACTCTAATTTAGATATCTACTAGTGAGA
CCTTGGGGGATAAGCTCCTTGGTCAAAAAGGGAAACAGCCCAGATCACCAGTTAA
GGCCCCAAAATGAAAATGATAGTGACTAAGGATGTGAGTATGTCAAAACCTCCA
GCAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 19

UTEX 1436 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGACTTAGAAAAGGTGGCATGGTTAAGGAAATATTC
CGAAGCCGTAGCAAAAAGCGAGTCTGAATAGGGCGATAAAATATATTAATATTTA
GAATCTAGTCATTTTTCTAGACCCGAACCCGGGTGATCTAACCATGACCAGGAT
GAAGCTTGGGTGATACCAAGTGAAGGTCCGAACCGACCGATGTTGAAAAATCGG
CGGATGAGTTGTGGTTAGCGGTGAAATACCAGTCGAACCCGGAGCTAGCTGGTT
CTCCCCGAAATGCGTTGAGGCGCAGCAGTACATCTAGTCTATCTAGGGGTAAAGC
ACTGTTTTCGGTGCGGGCTGTGAGAACGGTACCAAATCGTGGCAAACCTCTGAATAC
TAGAAATGACGATGTAGTAGTGAGACTGTGGGGGATAAGCTCCATTGTCAAGAG
GGAAACAGCCCAGACCACCAGCTAAGGCCCCAAAATGGTAATGTAGTGACAAAG
GAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGCCATCCTTTA
AAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 20

Chicorium intybus invertase: Genbank Accession No. Y11124

MSNSSNASESLFPATSEQPYRTAFHFQPPQNWMDPNPGPMCYNGVYHLFYQYNPFG
PLWNLRLMYWAHSVSHDLINWIHLDLAFAPTEPFDINGCLSGSATVLPGNKPIMLYTG
IDTENRQVQNLAVPKDLSDPYLREWVKHTGNPIISLPEEIQPDDFRDPTTTWLEEDGT
WRLLVGSQKDKTGIAFLYHSGDFVNWTKSDSPLHKVSGTGMWECVDFFPVWVDST
NGVDTSIUNPSNRVKHVLKLGIDHGKDCYLIGKYSADKENYVPEDELTLSTLRDLY
GMY YASKSFFDPVKNRRIMTAWVNESDSEADVIARGWSGVQSFPRSLWLDKNQKQ
LLQWPIEEIEMLHQNEVSFHNKKLDGGSSLEVLGITASQADV KISFKLANLEEAEELD
PSWVDPQLICSENDASKKGKFGPFGLLALASSDLREQTAIFFRVFRKNGRYVVL MCS
DQSRSSMKNGIEKRTYGAFVDIDPQQDEISLRTLIDHSIVESFGGRGKTCITTRVYPTL
AIGEQARLFAFNHGTESVEISELSAWSMKKAQMKVEEP

SEQ ID NO: 21

Schizosaccharomyces pombe Invertase: Genbank Accession No. AB011433

MFLKYILASGICLVSLLSSTNAAPRHLVVKRYPIYNASNITEVSNSTTVPPPPFVNTT
APNGTCLGNYNEYLPSSGYYNATDRPKIHFTPSSGFMNDPNGLVYTGGVYHMFQYS
PKTLTAGEVHWGHTVSKDLIHWENYPIAIYPDEHENGVLSPFSGSAVVDVHNSSGL
FSNDTIPEERIVLIYTDHWTGVAERQAIAIYTTDGGYTFKKYSGNPVLDINSLQFRDPK
VIWDFDANRWVMIVAMSQNYGIAFYSSYDLIHWTELSVFSTSGYLGLQYECPGMAR
VPVEGTDEYKWVLFISINPGAPLGGSVVQYFVGDWNGTNFVPPDDGQTRFVDLGKDF
YASALYHSSSANADVIGVGWASNWQYTNQAPTQVFRSAMTVARKFTLRDVPQNPM
TNLTSLIQTPLNVSLLRDETLFTAPVINSSSSLSGSPITLPSNTAFEFNVTLNINYTEGCT
TGYCLGRIIISDDPYRLQSSISVDVDFAASTLVINRAKAQMGWFNSLFTSPFANDIYTY
GNVTLYGIVDNGLLELYVNNGEKTYTNDFFFLQGATPGQISFAAFQGVSPNNVTVTP
LKTWNC

SEQ ID NO: 22

Picha anomala beta-fructofuranosidase (invertase): Genbank Accession No. X80640

MIQLSPLLLLPLFSVFENSIADASTEYLRPQIHLTPDQGWMMNDPNGMFYDRKDKLWHV
YFQHNPKKSIWATPVTWGHSTSKDLLTWDYHGNALPENDDGIFSGSVVVDNRN
TSGFFNDSTDPEQRIVAIYTNNAQLQTQEIAYSLDKGYSFIKYDQNPVINVNSSQQRD
PKVLWHDESNQWIMVVAKTQEFKVQIYGSPDLKKWDLKSNFTSNGYLGQYECPG
LFLKPIENPLNDTVTSKWVLLAINPGSPLGGSINEYFIGDFDGTTFHPDDGATRFMDI
GKDFYAFQSFNTEPEDGALGLAWASNWQYANTVPTENWRSSMSLVARNYTLKYVD
VNPENYGLTLIQKPVYDTKETRLNETLKTLETINEYEVNDLKLDKSSFVATDFNTERN
ATGVFEFDLKFQTDLMGYSNMTTQFGLYIHSQTVKGSQETLQLVFDLSTTWYID
RTTQHSFQRNSPVFTERISTYVEKIDTTDQGNVYTLYGVDNRNILELYFNDGSIAMTN
TFFFREGKIPTSFEVVCDSSEKSFITIDELSVRELARK

SEQ ID NO: 23

Debaryomyces occidentalis Invertase: Genbank Accession No. X17604

MVQVLSVLVIPLLTFFGYVASSSIDLSVDTSEYNRPLIHFTPEKGWMNDPNGLFYDK
TAKLWHLFYQYNPNATAWQQPLYWGHATSNDLVHWDEHEIAIGPEHDNEGIFSGSI
VVDHNNTSGFFNSSIDPNQRIVAIYTNNIPDLQTQDIAFSLDGGYTFTKYENNPVIDVS
SNQFRDPKVFVHERFKSMDHGCSEIARVKIQIFGSANLKNWVLNSNFSSGYGNQY
GMSRLIEVPIENSDDKSKWVMFLAINPGSPLGGSINQYFVGDFDGFQFVPDDSQTRFVD
IGKDFYAFQTFSEVEHGVGLAWASNWQYADQVPTNPWRSSSTSLARNYTLRYVIQM
LKL TANIDKSVLPDSINVVDKLLKKNVKLTKKKPIKTNFKGSTGLFDFNITFKVLNLN
VSPGKTHFDILINSQELNSSVDSIKIGFDSSQSLFYIDRHIPNVEFPRKQFFTDKLAAYL
EPLDYDQDLRVFSLYGIVDKNIIELYFNDGTVAMTNTFFMGEGKYPHDIQIVTDTEEP
LFELESVIHRELNK

SEQ ID NO: 24

Oryza sativa Invertase: Genbank Accession No. AF019113

MATSRLTPAYDLKNAAAA VYTPLPEQPHSAEVEIRDRKPFKIISA ILSL LLL LALILVA
VNYQAPPSHSSGDNSQPAAVMPPSRGVSQGVSEKA FRGASGAGNGVSFAWSNLMLS
WQRTSYHFQPVKNWMNDPNGLYYKGWYHLFYQYNPDSAVWGNITWGHAVSTD
LINWLHLPFAMVPDQWYDVNGVWGTGSATILPDGRIVMLYTGD TDDYVQDQNLAF
ANLSDPLLVDWVKYPNNPVIYPPPGIGVKDFRDPTTAGTAGMQNGQRLVTIGSKVG
KTGISLVYETTNTFTFKLLYGVLHAVPGTGMWECVDLYPVSTTG ENGLDTSVNLG
VKHVLKTS LDDDKHDYYALGTYPVKNKWTPDNPDL DVGIGLR LDY GKYYAARTF

YDQNKQRRILWGWIGETDLEAVDLMKGWASLQAIPRTIVFDKKTGTNVLQRPEEEV
ESWSSGDPITQRRIFEPGSVVPIHVSGATQLDITASFEVDETLLETTSESHDAGYDCSN
SGGAGTRGSLGPFGLLVVADEKLSELTPVYLYVAKGGDGKAKAHLCA YQTRSSMAS
GVEKEVYGSAPVLDGENYSARILIDHSIVESFAQAGRTCVRSDYPTKDIYGAARCF
FFNNATEASVRASLKAWQMKS FIRYPPIPDQKS

SEQ ID NO: 25

Allium cepa Invertase: Genbank Accession No. AJ006067

MSSDDLESPPSSYLPIPPSDEFHDQPPPLRSWLRLLSIPLALMFLLFLATFLSNLESPPSD
SGLVSDPVTFDVNPVAVRRGKDAGVSDKTS GVDSGFVLDPVAVDANSVVVHRGKD
AGVSDKTS GVDSGLLKDSPLGPYPWTNQMLSWQRTGFHFQPVKNWMNDPNGPLY
KGWYHFFYQYNPEGAVWGNIAWGHAVSRDLVHWTHLPLAMVPDQWYDINGVWT
GSATILPDGQIVMLYT GATNESVQVQNLAVPADQSDTLLLRWKKSEANPILVPPP
DKDFRDPTTAWYEPSDDTW RIVIGSKDSSSHSGIAIVYSTKDFINYKLIPGILHA
VERVG MWECVDFYPVATADSSHANHGLDPSARSPAVKHVLKASMDDDRHDYYAIG
TYDP AQNTWVPDDASVDVGIGLRYDWGKPYASKTFYDHAKKRILWSWIGETDSE
TADIA KGWASLQGVPRTVLLDVKTGSLNLTWPVVEIESLRTRPRDFSGITVDAGST
FKLDVG GAAQLDIEAEFKISSELEAVKEADVSYNCS SSGAAERGVLGPFGLLVLAN
QDLTE QTATYFYVSRGMDGGLNTHFCQDEKRSSKASDIVKRIVGHSVPVLDGESF
ALRILVD HSIVESFAQGGRASATSRVYPTEAIYNNARVVFVNNATGAKVTAQSLK
VWHMSTAI NEIYDPATSV M

SEQ ID NO: 26

Beta vulgaris subsp. vulgaris Invertase: Genbank Accession No. AJ278531

LFYQYNPNPVGVIWGPVWGHSTSKDLVNWVPQPLTMEPEMAANINGSWSGSATILPG
NKPAILFTGLDPKYEQVQVLA YPKDTS DPNLKEWFLAPQNPVMFPTPQNQINAT
SFR DPTTAWRLPDGVWRLLIGSKRGQRGLSLLFRSRDFVHWVQAKHPLYSDKLSGM
WE CPDFFPVYANGDQMGVDTSIIGSHVKHVLKNSLDITKHDIYTIGDYNIKKDAY
TPDIG YMNDSSLRYDYGKYYASKTFFDDAKKERILLGWANESSSVEDDIKKGWSG
IHTIPRK IWLDKLGKQLIQWPIANIEKLRQKPVNIYRKVLKGGSQIEVSGITAAQAD
VEISFKIKD LKNVEKFDASWTSPQLLCSKKGASVKGGLGPFGLLTLASXGLEEYTA
VFFRIFKAYD NKFVVLMCSDQSRSSLNPTNDKTTYGTFVDVNP IREGLSLRVLIDH
SVVESFGAKGK NVITARVYPTLAINEKAHLYVFNRGTSNVEITGLTAWSMKKANIA

SEQ ID NO: 27

Bifidobacterium breve UCC2003 beta-fructofuranosidase (invertase): Genbank Accession No. AAT28190

MTDFTPETPVLTPIRDHAAELAKAEAGVAEMAAKRNNRWYPKYHIASNGGWINDPN
GLCFYKGRWHVIFYQLHPYGTQWGPMHWGHVSSTDMLNWKREPIMFAPSLEQEKD
GVFSGSAVIDDNGDLRFYYTGHRWANGHDNTGGDWQVQMTALPDNDELTSATKQ
GMIIDCPTDKVDHHRDPKVWKTGDTWYMTFGVSSSEDKRGQMWL FSSKDMVRWE
YERVL FQHPDPDVFMLECPDFFPIKDKDGNEKWVIGFSAMGSKPSGFMNRNVNNA
G YMIGTWE PGGEFKPETEFRLWDCGHNY YAPQSFNVDGRQIVY GWMSPFVQPI
MED DGWCGQLTLPREITLDDDGDVVTAPVAEMEGLREDTL DHGSITLDM DGEQVIAD
DA EAVEIEMTIDLAAS'TADRAGLKI HATEDGAYTYVAYDDQIGRVV VDRQAMANGDH
GYRAAPLTD AELASGKLDLRV FVDRGSVEVYVNGGHQVLSSYSYASEGPRAIKLVA
EFGNLKVESLKLHHMKSIGLE

SEQ ID NO: 28

Saccharomyces cerevisiae Invertase: Genbank Accession No. NP_012104

MLLQAFLFLLAGFAAKISASMTNETSDRPLVHFTPNKGWMNDPNGLWYDEKDAKW
HLYFQYNPNDTVWGTPLFWGHATSDDLTNWEDQPIAIAPKRNDSGAFSGSMVVDY
NNTSGFFNDTIDPRQRCVAIWTYNTPESSEEQYISYSLDGGYTFTEYQKNPVLAANSTQ
FRDPKVFWEPSQKWIMTAAKSQDYKIEIYSSDDLKSWKLESAFANEGFLGYQYECF
GLIEVPTQDPSKSYWVMFISINPGAPAGGSFNQYFVGSFNGTHFEAFDNQSRVVDG
KDYALQTFNTDPTYGSALGIAWASNWEYSFVPTNPWRSSMSLVRKFSLNTEYQ
ANPETELINLKAEPILNISNAGPWSRFATNTTLTKANSYNVDLSNSTGTLEFELVYAV
NTTQTISKSVFADLSLWFKGLEDPEEYLRMGFEVSASSFFLDRGNSKVFKVKNPYFT
NRMSVNNQPFKSENDLSYKVGGLLDQNILELYFNDGDVVSTNTYFMTTGNALGSV
NMTTGVDNLFYIDKFQVREVK

SEQ ID NO: 29

Zymomonas mobilis Invertase A: Genbank Accession No. AY171597

MESPSYKNLIKAEDAQKKAGKRLLSSEWYPGFHVTPLTGWMNDPNGLIFFKGEYHL
FYQYYFAPVWGPMPHWGHAKSRDLVHWETLPVALAPGDLFDRDGCFSGCAVDNN
GVLTLIYTGHIWLSNDSPDAIREVQCMATSIDGIHFQKEGIVLEKAPMPQVAHFRDPR
VWKENDHWFMVVGYRTDDEKHQGGHVALYRSENLDWIFVKTLLGDNSQLPLGK
RAFMWECPDFSLGNRSVLMFSPQGLKASGYKNRNLFNQNGYILGKWQAPQFTPETS
FQELDYGHDFYAAQRFEAKDGRQILAWFDMWENQKPSQRDGWAGCMTLPRKLDL
IDNKIVMTVPVREMEILRQSEKIESVVTLSDAEHPFTMDSPLEIELIFDLEKSSAYQAG
LALRCNGKGQETLLYIDRSQNRILDRNRSGQNVKGIRSCPLPNTSKVRLHIFLDRSSIE
IFVGDDQTQGLYSISRIFFDKDSLKGRLFAIEGYAVFDSFKRWTLQDANLAAFSSDA
C

SEQ ID NO: 30

Cinnamomum camphora FATB1 (Genbank Q39473) amino acid sequence with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

MAIKTNRQPVEKPPFTIGTLRKAIPAHCERSALRGRAPDWSMLFAVITTIFSAAEKQ
WTNLEWKPKPNPPQLDDHFGPHGLVFRRTFAIRS YEVGPDRTSIVAVMNLHQEAA
LNHAKSVGILGDGFGTTLEMSKRDLIWVVKRTHVAVERYPAWGD TVEVEC WVGAS
GNNGRRHDFLVRDCKTGEILTRCTSLVMMNTRTRRLSKIPEEVRGEIGPAFIDNVAV
KDEEIKKPQKLNDSTADYIQGGLTPRWNDLDINQHVNNIKYVDWILETVPDSIFESH
ISSFTIEYRRECTMDSVLQSLTTVSGGSSEAGLVCEHLLQLEGGSEVLRAKTEWRPKL
TDSFRGISVIPAESSV

SEQ ID NO: 31

Relevant codon optimized expression construct of *Cinnamomum camphora* FATB1 cDNA with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCTGCCTGCCGCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCA TGCCGGGCAATGCGCATGGTGGAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA

TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTTCAGCAAAGGTAGCAGTTT
 GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
 GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
 AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTTAGCTACAACGGCAGGAGTC
 ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
 AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
 GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
 CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATAACCAAACCCTCT
 GGCTGCTTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
 GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCACACTGT
 CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
 GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
 AGAGGCGGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
 GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTACATTTCATT
 TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
 AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
 TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCGTGTT
 GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
 CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
 TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
 TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
 CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
 CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
 ACGAGGCCCGCCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
 GCTGGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
 CGGCCAGGACCTGCTGTCCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
 GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCCCTCGACC
 ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
 TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
 TCGCCCGCCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
 GCGACGCCTGCCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
 CGACTGCGGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
 CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGCT
 GTACGGCATCGCCGCCCGGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
 GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
 GCTGGTCTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
 CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
 TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTTC
 CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTTGCTATCC
 CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
 CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
 CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
 GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
 CACAATAACCACCTGACGAATGCGCTTGGTTCTTTCGTCCATTAGCGAAGCGTCCG
 GTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
 GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCCTTTCTTGCGCTATGAC
 ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
 ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
 ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGAC
 ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCCTTCTAC

ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
 GTTTCAGTCACAACCCGCAAACACTAGTATGGCTATCAAGACGAACAGGCAGCC
 TGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCA
 CTGTTTCGAGCGCTCGGCGCTTCGTGGGCGCGCCCCCGACTGGTCCATGCTGTTT
 GCCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGACCAACCTGGAGT
 GGAAGCCCAAGCCCAACCCCCCCCAGCTGCTGGACGACCACTTCGGCCCCCACG
 GCCTGGTGTTCGCGCGCACCTTCGCCATCCGCAGCTACGAGGTGGGCCCCGACCG
 CTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCCGCCCTGAACCA
 CGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCCTGGAGATGTCC
 AAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCCGTGGAGCGCTAC
 CCCGCCTGGGGCGACACCGTGGAGGTGGAGTGTGGGTGGGCGCCTCCGGCAAC
 AACGGCCGCCGCCACGACTTCCTGGTGGCGGACTGCAAGACCGGCGAGATCCTG
 ACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCCGCCGCCTGAGCA
 AGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCCGCCTTCATCGACAACGTGG
 CCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGACTCCACCGCCC
 ACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACATCAACCAGC
 ACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACCGTGCCCCGACAGCA
 TCTTCGAGAGCCACCACATCTCCTCCTTCACCATCGAGTACCGCCGCGAGTGCAC
 CATGGACAGCGTGTGTCAGTCCCTGACCACCGTGAGCGGCGGCTCCTCCGAGGC
 CGGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCG
 CGCCAAGACCGAGTGGCGCCCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGT
 GATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCACGACGGCGACTACAA
 GGACCACGACATCGACTACAAGGACGACGACGACAAGTGATGACTCGAGGCAGC
 AGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTTCGTGTGATGGACTGTTG
 CCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCC
 TCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCT
 ATTTGCGAATAACACCCCCAGCATCCCCTCCCTCGTTTCATATCGCTTGCATCCC
 AACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCT
 CACTGCCCTTCGCACAGCCTTGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCA
 ACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACA
 CAAATGGAAAGCTT

SEQ ID NO: 32

Cuphea hookeriana FATB2 (Genbank AAC49269) amino acid sequence with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

MAIKTNRQPVEKPPFTIGTLRKAIPAHCFFERSALRGRAQLPDWSRLLTAITTVFVKSK
 RPDMDHRKSKRPDMLVDSFGLESTVQDGLVFRQSFIRSIEIGTDRITASIELTMNHLQ
 ETSLNHCKSTGILLDGFGRITLEMCKRDLIWVVIKMQIKVNRYPAWGDTVEINTRFSR
 LGKIGMGRDWLISDCNTGEILVRATSAYAMMNQKTRRLSKLPYEVHQEIVPLFVDSP
 VIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVSINVKYIGWILESMPTVEVLET
 QELCSLALEYRRECGRDSVLESVTAMDPSKVGVRISQYQHLLRLEDGTAIVNGATEW
 RPKNAGANGAISTGKTSNGNSVS

SEQ ID NO: 33

Relevant codon optimized expression construct of *Cuphea hookeriana* FATB2 cDNA with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAAC
 CGACACGTGCTTGGGCGCCTGCCGCCTGCCTGCCGCATGCTTGTGCTGGTGAGGC

TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAGCTTATGTG
 TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
 TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
 TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTTCAGCAAAGGTAGCAGTTT
 GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
 GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
 AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTC
 ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
 AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
 GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
 CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATAACCAAACCTCT
 GGCTGCTTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
 GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCACACTGT
 CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
 GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
 AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
 GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTTCAT
 TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
 AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
 TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCGTGTT
 GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
 CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
 TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
 TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
 CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
 CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
 ACGAGGCCCGCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
 GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
 CGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
 GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCTTCGACC
 ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCGATGGAGGCCGGCCTGG
 TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
 TCGCCCGCCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
 GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
 CGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
 CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGCT
 GTACGGCATCGCCGCCCGGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
 GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
 GCTGGTCTGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
 CCTGCGGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
 TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCC
 CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCC
 CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTC
 CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
 CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
 GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
 CACAATAACCACCTGACGAATGCGCTTGGTTCTTTCGTCCATTAGCGAAGCGTCCG
 GTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
 GATGGTCGAAACGTTTACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
 ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA

ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
 ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGAC
 ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
 ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
 GTTTCAGTCACAACCCGCAAACACTAGTATGGCTATCAAGACGAACAGGCAGCC
 TGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCA
 CTGTTTCGAGCGCTCGGCGCTTCGTGGGCGCGCCCAGCTGCCCGACTGGAGCCGC
 CTGCTGACCGCCATCACCAACCGTGTTCGTGAAGTCCAAGCGCCCCGACATGCACG
 ACCGCAAGTCCAAGCGCCCCGACATGCTGGTGGACAGCTTCGGCCTGGAGTCCA
 CCGTGCAGGACGGCCTGGTGTTCGCGCAGTCCTTCTCCATCCGCTCCTACGAGAT
 CGGCACCGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAGAC
 CTCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTGGACGGCTTCGGCCGCAACC
 CTGGAGATGTGCAAGCGCGACCTGATCTGGGTGGTGGATCAAGATGCAGATCAAG
 GTGAACCGCTACCCCGCCTGGGGCGACACCGTGGAGATCAACACCCGCTTCAGC
 CGCCTGGGCAAGATCGGCATGGGCCGCGACTGGCTGATCTCCGACTGCAACACC
 GGCGAGATCCTGGTGCGCGCCACCAGCGCCTACGCCATGATGAACCAGAAGACC
 CGCCGCCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTGCCCCCTGTTCCG
 TGGACAGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGA
 AGACCGGCGACAGCATCCAGAAGGGCCTGACCCCCGGCTGGAACGACCTGGACG
 TGAACCAGCACGTGTCCAACGTGAAGTACATCGGCTGGATCCTGGAGAGCATGC
 CCACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCCCTGGCCCTGGAGTACCGCC
 GCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGGACCCCAGCA
 AGGTGGGCGTGCGCTCCAGTACCAGCACCTGCTGCGCCTGGAGGACGGCACCG
 CCATCGTGAACGGCGCCACCGAGTGGCGCCCCAAGAACGCCGGCGCCAACGGCG
 CCATCTCCACCGGCAAGACCAGCAACGGCAACTCCGTGTCCATGGACTACAAGG
 ACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGAC
 AAGTGACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGG
 TCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGC
 CGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCAG
 TTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTT
 TCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGC
 GCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTG
 TATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAA
 GTAGTGGGATGGGAACACAAATGGAAAGCTT

SEQ ID NO: 34

Umbellularia californica FATB1 (Genbank Q41635) amino acid sequence with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

MAIKTNRPVEKPPFTIGTLRKAIPAHCFERSALRGRAPDWSMLFAVITTIIFSAAEKQ
 WTNLEWKPKPKLPQLDDHFGLHGLVFRRTFAIRSYEVGPDRSTSILAVMNMHMQEA
 TLNHAHSVGLGDGFGTTLEMSKRDLMWVVRRTHVAVERYPTWGDVEVECWIGA
 SGNNGMRRDFLVRDCKTGEILTRCTSLSVLMNTRTRRLSTIPDEVARGEIGPAFIDNVA
 VKDDEIKKLQKLNSTADYIQGGLTPRWNDLDVNQHVNNLKYVAWVFETVPDSIFE
 SHHISFTLEYRRECTRDSVLRSLTTVSGGSSEAGLVCDHLLQLEGGSEVLRA RTEWR
 PKLTDSFRGISVIPAEPRV

SEQ ID NO: 35

Relevant codon optimized expression construct of *Umbellularia californica* FATB1 cDNA with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

GGTACCCGCTGCAACGCAAGGGCAGCCACAGCCGCTCCACCCGCCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCCTGCCGCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCCTCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTTCAGCAAAGGTAGCAGTTT
GCCAGAATGATCGGTTTCACTGTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
AGTGTGTGCGAGCCAGCATCGCCGCTGGCTGTTTCAGCTACAACGGCAGGAGTC
ATCCAACGTAAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCTCT
GGCTGCTTGTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCACACTGT
CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTTCATT
TGCATGCCITGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
TTCCATAGTGCAATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCGTGT
GCAAAACGCGCCAGCTCAGCAACCTGTCCCCTGGGTCCCCCGTGCCGATGAAAT
CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
TCCACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCCGCCGCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCTTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCTGCACACCCTGGACCCCGCCACCTGCCCTTCGACC
ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGCTGCT
GTACGGCATCGCCGCCCGGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
CACAATAACCACCTGACGAATGCGCTTGGTTCTTTCGTCCATTAGCGAAGCGTCCG

GTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
 GATGGTCGAAACGTTTACAGCCTAGGGATATCGAATTCCTTTCTTGCCTATGAC
 ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
 ACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
 ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCGGATTGCAAAGAC
 ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
 ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
 GTTTCAGTCACAACCCGCAAACACTAGTATGGCTATCAAGACGAACAGGCAGCC
 TGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCA
 CTGTTTCGAGCGCTCGGCGCTTCGTGGGCGCGCCCCGACTGGTCCATGCTGTTT
 GCCGTGATCACCACCATCTTCAGCGCCGCGGAGAAGCAGTGGACCAACCTGGAG
 TGGAAGCCCAAGCCCAAGCTGCCCCAGCTGCTGGACGACCACTTCGGCCTGCAC
 GGCCTGGTGTTCGCGCGCACCTTCGCCATCCGCTCCTACGAGGTGGGCCCCGACC
 GCAGCACCTCCATCCTGGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACC
 ACGCCAAGAGCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCCTGGAGATGT
 CCAAGCGCGACCTGATGTGGGTGGTGCGCCGACCCACGTGGCCGTGGAGCGCT
 ACCCCACCTGGGGCGACACCCTGGAGGTGGAGTGCTGGATCGGCGCCAGCGGCA
 ACAACGGCATGCGCCGCGACTTCCTGGTGCGCGACTGCAAGACCGGCGAGATCC
 TGACCCGCTGCACCTCCCTGAGCGTGCTGATGAACACCCGCACCCGCCGCTGAG
 CACCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCGCCTTCATCGACAACGT
 GGCCGTGAAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGC
 CGACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACGTGAACCA
 GCAGTGAACAACCTGAAGTACGTGGCCTGGGTGTTTCGAGACCGTGCCCGACAG
 CATCTTCGAGTCCACACATCAGCTCCTTCACCCTGGAGTACCGCCGCGAGTGC
 ACCCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCAGCTCCGAG
 GCCGGCCTGGTGTGCGACCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTG
 CGCGCCCGCACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCGCGGCATCAGC
 GTGATCCCCGCCGAGCCCCGCGTGATGGACTACAAGGACGACGACGACAAGTGATGACTCGAGGC
 AAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGATGACTCGAGGC
 AGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACT
 GTTGCCGCCACACTTGCTGCCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAAC
 AGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTG
 TGCTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCA
 TCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCC
 TGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTAC
 TGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGG
 AACACAAATGGAAAGCTT

SEQ ID NO: 36

Cuphea palustris C8 preferring thioesterase (Genbank AAC49179) amino acid sequence with
Prototheca moriformis delta 12 fatty acid desaturase transit peptide.

MAIKTNRQPVEKPPFTIGTLRKAIPAHCERSALRGRAPANGSAVTLKSGSLNTQEDT
 LSSSPPPRAFFNQLPDWSMLLTATTVFVAPEKRWTMFDRKSKRPNMLMDSFGLERV
 VQDGLVFRQSFSIRSIEICADRTASIEVTVMNHVQETSLNQCKSIGLLDDGFGRSPENC
 KRDLIWVVTRMKIMVNRYPWGDITIEVSTWLSQSGKIGMGRDWLISDCNTGEILVR
 ATSVYAMMNQKTRRFKLPHEVRQEFAPHFLDSPPAIEDNDGKLQKFDVKTGDSIRK
 GLTPGWYDLVDNQHVSNVKYIGWILESMPTVEVLETQELCSLTLEYRRECGRDSVLES
 VTSMDPSKVGDRFQYRHLLRLEDGADIMKGRTEWRPKNAGTNGAISTGKT

SEQ ID NO: 37

Relevant codon optimized expression construct of *Cuphea palustris* C8 preferring thioesterase cDNA with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCAACCCGCCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCCTGCCTGCCGCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTCAGCAAAGGTAGCAGTTT
GCCAGAATGATCGGTTTCACTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
GCAAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTCAGCTACAACGGCAGGAGTC
ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCTCT
GGCTGCTTGTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCACACTGT
CACCAGGCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
AGAGGCGGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
GGGTGCGGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCATT
TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCTGTT
GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGCCTGTCCGCCAGGGC
CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCGCCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCTTCGACC
ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCGCTGAAGGCCCGCATGCCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCGCCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTTGGTGCT
GTACGGCATCGCCGCCCGGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCGCATCCCCCTTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGTCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA

CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
GCGCAGAGGAACGCTGAAGGTCTCGCTCTGTGCGACCTCAGCGCGGCATACAC
CACAATAACCACTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG
GTTACACACAGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCTTTCTTGGCGCTATGAC
ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCGGATTGCAAAGAC
ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
GTTTCAGTCACAACCCGCAAACACTAGTATGGCTATCAAGACGAACAGGCAGCC
TGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCA
CTGTTTCGAGCGCTCGGCGCTTCGTGGGCGCGCCCCCGCGAACGGCAGCGCGGTG
ACCCTGAAGTCGGGCTCCCTGAACACCCAGGAGGACACGCTGAGCTCGTCCCCC
CCCCCCCCGCGCTTCTTCAACCAGCTGCCCCGACTGGAGCATGCTGCTGACCGCGA
TCACCACGGTCTTCGTGGCGCCCCGAGAAGCGCTGGACCATGTTGACCGCAAGTC
GAAGCGCCCCAACATGCTGATGGACTCCTTCGGCCTGGAGCGCGTGGTCCAGGA
CGGCCTGGTGTTCGCCAGAGCTTCTCGATCCGCTCCTACGAGATCTGCGCGGAC
CGCACCGCGAGCATCGAGACGGTGATGAACCACGTCCAGGAGACCTCGCTGAAC
CAGTGCAAGTCCATCGGCCTGCTGGACGACGGCTTCGGCCGCGAGCCCCGAGATG
TGCAAGCGCGACCTGATCTGGGTGGTCACCCGCATGAAGATCATGGTGAACCGC
TACCCACGTGGGGCGACACCATCGAGGTCTCGACGTGGCTGTCCCAGAGCGGC
AAGATCGGCATGGGCCGCGACTGGCTGATCTCGGACTGCAACACCGGCGAGATC
CTGGTGC GCGCGACGTCCGTCTACGCGATGATGAACCAGAAGACCCGCCGCTTC
AGCAAGCTGCCCCACGAGGTGCGCCAGGAGTTCGCGCCCCACTTCTGGACTCGC
CCCCCGCGATCGAGGACAACGACGGCAAGCTGCAGAAGTTCGACGTCAAGACGG
GCGACTCCATCCGCAAGGGCCTGACCCCCGGCTGGTACGACCTGGACGTGAACC
AGCACGTGAGCAACGTCAAGTACATCGGCTGGATCCTGGAGTCGATGCCACCG
AGGTCTTGGAGACGCAGGAGCTGTGCTCCCTGACCCTGGAGTACCGCCGCGAGT
GCGGCCGCGACTCGGTGCTGGAGAGCGTCAACAGCATGGACCCCTCGAAGGTGG
GCGACCGCTTCCAGTACCGCCACCTGCTGCGCCTGGAGGACGGCGCGGACATCA
TGAAGGGCCGCACCGAGTGGCGCCCCAAGAACGCGGGCACGAACGGCGCGATCT
CCACCGGCAAGACGTGACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACT
CTGGACGCTGGTTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGT
GAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACG
CGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCC
CCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTTG
CTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCTCGCACAGCCTTGTTTT
GGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCT
GATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTC

SEQ ID NO: 38

Cuphea lanceolata C10 preferring thioesterase (Genbank CAB60830) amino acid sequence
with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

MAIKTNRPVEKPPFTIGTLRKAIPAHCERSALRGRAPANGSAVNLSGSLNTQEDT
SSSPPPRAFLNQLPDWSMLLTAITTVFVA AEKQWTMLDRKSKRPDMLVDSVGLKSIV
RDGLVSRQSFLIRSYEIGADRTASIELTMNHLQETSINHCKSLGLLNDGFGRTPGMCK
NDLIWVLTKMQIMVNRYPWTWGDVEINTWFSQSGKIGMASDWLISDCNTGEILIRAT
SVWAMMNQKTRRFSRLPYEVRQELTPHFVDSPHVIEDNDQKLHKFDVKTGDSIRKG

LTPRWNDLDVNQHVSINVKYIGWILESMPIEVLETQELCSLTVEYRRECGMDSVLESV
TAVDPSENGGRSQYKHLLRLEDGTDIVKSRTTEWRPKNAGTNGAISTSTAKTSNGNSA
SDDDDKLG

SEQ ID NO: 39

Relevant codon optimized coding region of *Cuphea lanceolata* C10 preferring thioesterase with *Prototheca moriformis* delta 12 fatty acid desaturase transit peptide.

ACTAGTATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTCACG
ATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTC
GTGGGCGCGCCCCCGCGAACGGCAGCGCGGTGAACCTGAAGTCGGGGCTCCCTGA
ACACCCAGGAGGACACGAGCTCGTCCCCCCCCCCCCCGCGCGTTCCTGAACCAGCT
GCCCCACTGGAGCATGCTGCTGACCGCGATCACCACCGTCTTCGTGGCGGCGGA
GAAGCAGTGGACGATGCTGGACCGCAAGTCGAAGCGCCCCGACATGCTGGTGGA
CTCCGTTCGGCCTGAAGAGCATCGTGC GCGACGGCCTGGTCTCGCGCCAGTCCTTC
CTGATCCGCAGCTACGAGATCGGCGCGGACCGCACCGCGTCGATCGAGACCCTG
ATGAACCACCTGCAGGAGACGTCCATCAACCACTGCAAGAGCCTGGGCCTGCTG
AACGACGGCTTCGGCCGCACCCCCGGCATGTGCAAGAACGACCTGATCTGGGTG
CTGACCAAGATGCAGATCATGGTCAACCGCTACCCACGTGGGGCGACACCGTC
GAGATCAACACGTGGTTCTCGCAGTCCGGCAAGATCGGCATGGCGAGCGACTGG
CTGATCTCGGACTGCAACACCGGCGAGATCCTGATCCGCGCGACCTCCGTGTGGG
CGATGATGAACCAGAAGACGCGCCGCTTCAGCCGCCTGCCCTACGAGGTCCGCC
AGGAGCTGACCCCCCACTTCGTGGACTCGCCCCACGTCATCGAGGACAACGACC
AGAAGCTGCACAAGTTCGACGTGAAGACCGGCGACTCCATCCGCAAGGGCCTGA
CGCCCCGCTGGAACGACCTGGACGTCAACCAGCACGTGTCGAACGTGAAGTACA
TCGGCTGGATCCTGGAGTCCATGCCATCGAGGTCCTGGAGACCCAGGAGCTGTG
CTCGCTGACCGTGGAGTACCGCCGCGAGTGC GGCATGGACTCCGTGCTGGAGTC
GGTCACGGCGGTGGACCCCAGCGAGAACGGCGGCGCGCAGCCAGTACAAGCACCT
GCTGCGCTGGAGGACGGCACCGACATCGTCAAGTCGCGCACCGAGTGGCGCCC
CAAGAACGCGGGCACGAACGGCGCGATCTCCACCAGCACCGCGAAGACGTCGAA
CGGCAACTCCGCGAGCGATGACGATGACAAGCTGGGATGACTCGAG

SEQ ID NO: 40

Iris germanica C14 preferring thioesterase (Genbank AAG43858.1) amino acid sequence with *Chlorella protothecoides* stearoyl ACP desaturase chloroplast transit peptide.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAAQAATRVNGSKVGLKTD
NKLEDAPFIPSSAPRTFYNQLPDWSVLLAAITIFLAAEKQWTLIDWKRGGPDMLSDA
FGLPKHENGLLYRQKFSIRS YEIGADQTAS IETLMNHLQETALNHVKCAGLLGNGFG
STPEMSKMNLIWVVTKMQVLVEHYPSWGDVIEVDTWAAASGKNGMRDWHVRD
WQTGQTIMRASSNWVMMNQNTRRLSKFPEEVRAEIEPYFMERAPVIDDDNRKLPKL
DDDTADHVRNGLTPRWSDLDVNQHVKNVKYIGWILESAPISILESHELASMTLEYRR
ECGRDSVLQSLTSVSNNCTDGSEELPIECQHLLRNEGGSEIVKGRTEWRPKKCGPFGA
GRP

SEQ ID NO: 41

Relevant codon optimized coding region of *Iris germanica* C14 preferring thioesterase with *Chlorella protothecoides* stearoyl ACP desaturase transit peptide.

ACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACC
 TGGCTCGCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGAGGCCCTCCCCGTGCG
 CGGGCGCGCCGCCAGGCGGGCACCCGCGTGAACGGCAGCAAGGTGGGCTGAA
 GACCGACACCAACAAGCTGGAGGACGCGCCCTTCATCCCCTCGTCCGCCCCCGC
 ACCTTCTACAACCAGCTGCCCCACTGGAGCGTCCTGCTGGCGGCCATCACCACCA
 TCTTCCTGGCGGCCGAGAAGCAGTGGACCCTGATCGACTGGAAGCGCGGCGGCC
 CCGACATGCTGTCGGACGCGTTCGGCCTGCCCAAGATCATCGAGAACGGCCTGCT
 GTACCGCCAGAAGTTCTCCATCCGCAGCTACGAGATCGGCGCCGACCAGACCGC
 CTCGATCGAGACCCTGATGAACCACCTGCAGGAGACCGCGCTGAACCACGTCAA
 GTGCGCCGGCCTGCTGGGCAACGGCTTCGGCTCCACCCCCGAGATGAGCAAGAT
 GAACCTGATCTGGGTGGTCACCAAGATGCAGGTGCTGGTCGAGCACTACCCCTCG
 TGGGGCGACGTGATCGAGGTGGACACCTGGGCGGCCGCGTCCGGCAAGAACGGC
 ATGCGCCGCGACTGGCACGTCCGCGACTGGCAGACCGGCCAGACCATCATGCGC
 GCCAGCTCGAACTGGGTGATGATGAACCAGAACACCCGCCGCTGTCCAAGTTC
 CCCGAGGAGGTCCGCGCCGAGATCGAGCCCTACTTCATGGAGCGCGCCCCCGTG
 ATCGACGACGACAACCGCAAGCTGCCCAAGCTGGACGACGACACCGCGGACCAC
 GTGCGCAACGGCCTGACCCCCCGCTGGAGCGACCTGGACGTGAACCAGCACGTC
 AAGAACGTGAAGTACATCGGCTGGATCCTGGAGTCGGCCCCCATCTCCATCCTGG
 AGAGCCACGAGCTGGCCTCGATGACCCTGGAGTACCGCCGCGAGTGCGGCGCGC
 ACTCCGTCCTGCAGAGCCTGACCTCGGTGTCCAACAACCTGCACCGACGGCAGCG
 AGGAGCTGCCCATCGAGTGCCAGCACCTGCTGCGCAACGAGGGCGGCTCGGAGA
 TCGTCAAGGGCCGCACCGAGTGGCGCCCCAAGAAGTGCGGCCCTTCGGCGCCG
 GCCGCCCTGACTCGAG

SEQ ID NO: 42

Myristica fragrans fatty acyl thioesterase (Genbank AAB717291.1) amino acid sequence
 with *Prototheca moriformis* delta 12 fatty acid desaturase chloroplast transit peptide.

MAIKTNRPVEKPPFTIGTLRKAIPAHCFERSALRGRAANAHTVPKINGNKAGLLTPM
 ESTKDEDIVAAPTVPKRTFINQLPDWSMLLAITTIFLAAEKQWTNLDWKPRRPDM
 LVDFFDPFSLGRFVQDGLIFRQNFIRSIEIGADRTASIEITLMNHLQETALNHVRCIGLL
 DDGFGSTPEMTRRDLIWVVTMRQVLVDRYPSWGDVIEVDSWVTPSGKNGMKREW
 LRDKTGEILTRATSVWVMMNKRTRRLSKIPEEVRVEIEPYFVEHGVLDSDSRKLPK
 LNDNTANYIRRGLAPRWSLDLVNQHVN NVKYIGWILESVPSSLLESHELYGMTLEY
 RKECGKDGLLQSLTAVASDYGGGSLEAGVECDHLLRLEDGSEIMRGKTEWRPKRAA
 NTTYFGSVDDIPPANNA

SEQ ID NO: 43

Relevant codon optimized coding region of *Mysistica fragrans* fatty acyl thioesterase with
Prototheca moriformis delta 12 fatty acid desaturase chloroplast transit peptide.

ACTAGTATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTCACG
 ATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTC
 GTGGGCGCGCCGCCAACGCCCACACCGTGCCCAAGATCAACGGCAACAAGGCCG
 GCCTGCTGACCCCCATGGAGAGCACCAAGGACGAGGACATCGTCGCGGCCCCCA
 CCGTGGCGCCCAAGCGCACCTTCATCAACCAGCTGCCCCACTGGTCGATGCTGCT
 GGCCGCGATCACCACCATCTTCCTGGCGGCCGAGAAGCAGTGGACCAACCTGGA
 CTGGAAGCCCCGCCGCCCGACATGCTGGTCGACTTCGACCCCTTCTCCCTGGGC
 CGCTTCGTGCAGGACGGCCTGATCTTCCGCCAGAACTTCAGCATCCGCTCGTACG
 AGATCGGCGCGGACCGCACCGCCTCCATCGAGACCCTGATGAACCACCTGCAGG

AGACCGCGCTGAACCACGTCCGCTGCATCGGCCTGCTGGACGACGGCTTCGGCA
 GCACCCCGAGATGACCCGCCGCGACCTGATCTGGGTGGTCACCCGCATGCAGG
 TCCTGGTGGACCGCTACCCCTCGTGGGGCGACGTGATCGAGGTCGACTCCTGGGT
 GACCCCGAGCGGCAAGAACGGCATGAAGCGCGAGTGGTTCCTGCGCGACTGCAA
 GACCGGCGAGATCCTGACCCGCGCCACCTCGGTCTGGGTGATGATGAACAAGCG
 CACCCGCCGCTGTCCAAGATCCCGAGGAGGTCCGCGTGGAGATCGAGCCCTA
 CTTTCGTCGAGCACGGCGTGCTGGACGAGGACTCGCGCAAGCTGCCCAAGCTGAA
 CGACAACACCGCCAACTACATCCGCCGCGGCCTGGCGCCCCGCTGGTCCGACCTG
 GACGTCAACCAGCACGTGAACAACGTCAAGTACATCGGCTGGATCCTGGAGAGC
 GTGCCCAGCAGCCTGCTGGAGTCGCACGAGCTGTACGGCATGACCCTGGAGTAC
 CGCAAGGAGTGCGGCAAGGACGGCCTGCTGCAGTCCCTGACCGCCGTCGCCAGC
 GACTACGGCGGGCGGCTCGCTGGAGGCCGGCGTGGAGTGCGACCACCTGCTGCGC
 CTGGAGGACGGCTCCGAGATCATGCGCGGCAAGACCGAGTGGCGCCCCAAGCGC
 GCCGCGAACACCACCTACTTCGGCAGCGTCGACGACATCCCCCCCCGCCAACAAC
 GCGTGA CT CGAG

SEQ ID NO: 44

Cuphea palustris C14 preferring thioesterase (Genbank AAC49180) amino acid sequence with *Chlorella protothecoides* stearyl ACP desaturase transit peptide.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASMLLSAVTTVFGVAEKQW
 PMLDRKSKRPDMLVEPLGVDRIVYDGVSFQSFIRSIEIGADRTASIELTMNMFQET
 SLNHCKIIGLLNDGFGRTPEMCKRDLIWVVTMKEIVNRYPTWGDTHIEVNTWVSASG
 KHGMGRDWLISDCHTGEILIRATSVWAMMNQKTRRLSKIPYEVQRQIEPQFVDSAPV
 IVDDRKFHKLDLKTGDSICNGLTPRWTDLDVNQHVNNVKYIGWILQSVPTVEVFETQE
 LCGLTLEYRRECGRDSVLESVTAMDPSKEGDRSLYQHLLRLEDGADIVKGRTEWRP
 KNAGAKGAILTGKTSNGNSIS

SEQ ID NO: 45

Relevant codon optimized coding region of *Cuphea palustris* C14 preferring thioesterase with *Chlorella protothecoides* stearyl ACP desaturase transit peptide.

ACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGGCGACC
 TGGCTCGCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGAGGCCCTCCCCGTGCG
 CGGGCGCGCCAGCATGCTGCTGTCGGCGGTGACCACGGTCTTCGGCGTGGCCGA
 GAAGCAGTGGCCCATGCTGGACCGCAAGTCCAAGCGCCCCGACATGCTGGTCTGA
 GCCCCTGGGCGTGGACCGCATCGTCTACGACGGCGTGAGCTTCGCCAGTCGTTT
 TCCATCCGCAGCTACGAGATCGGCGCCGACCGCACCGCCTCGATCGAGACGCTG
 ATGAACATGTTCCAGGAGACCTCCCTGAACCACTGCAAGATCATCGGCCTGCTGA
 ACGACGGCTTCGGCCGCACGCCCCGAGATGTGCAAGCGCGACCTGATCTGGGTCTG
 TGACCAAGATGCAGATCGAGGTGAACCGCTACCCACAGTGGGGCGACACCATCG
 AGGTCAACACGTGGGTGAGCGCCTCGGGCAAGCACGGCATGGGCCGCGACTGGC
 TGATCTCCGACTGCCACACCGGCGAGATCCTGATCCGCGCGACGAGCGTCTGGGC
 GATGATGAACCAGAAGACCCGCCGCTGTCGAAGATCCCCTACGAGGTGCGCCA
 GGAGATCGAGCCCCAGTTCGTCGACTCCGCCCCCGTGATCGTGGACGACCGCAA
 GTTCCACAAGCTGGACCTGAAGACGGGCGACAGCATCTGCAACGGCCTGACCCC
 CCGCTGGACGGACCTGGACGTGAACCAGCACGTCAACAACGTGAAGTACATCGG
 CTGGATCCTGCAGTCGGTCCCCACCGAGGTGTTTCGAGACGCGAGGAGCTGTGCGG
 CCTGACCTGGAGTACCGCCGCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCGT
 CACGGCCATGGACCCCTCGAAGGAGGGCGACCGCTCCCTGTACCAGCACCTGCT

GCGCCTGGAGGACGGCGCGGACATCGTGAAGGGCCGCACCGAGTGGCGCCCCAA
GAACGCCGGCGCCAAGGGCGCCATCCTGACGGGCAAGACCAGCAACGGCAACTC
GATCTCCTGACTCGAG

SEQ ID NO: 46

Ulmus americana broad specificity thioesterase (Genbank AAB71731) amino acid sequence with *Chlorella protothecoides* stearyl ACP desaturase transit peptide.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAQLPDWSMLLAATTFLAA
EKQWMMLDWKPKRPDMLVDPFGLGRFVQDGLVFRNNSIRSIEIGADRTASIELM
NHLQETALNHVKSIVGLLEDGLGSTREMSLRNLIWVVTKMQVAVDRIPTWGDVQV
SSWATAIGKNGMRREWIVTDFRTGETLLRATSVWVMMNKLTRISKIPEEVWHEIGP
SFIDAPPLPTVEDDGRKLTRFDESSADFIRKGLTPRWSLDINQHVNNVKYIGWLLES
APPEIHESHEIASLTLEYRRECGRDSVLNSATKVSDDSSQLGKSAVECNHLVRLQNGGE
IVKGRTVWRPKRPLYNDGAVVDVPAKTS

SEQ ID NO: 47

Relevant codon optimized coding region of *Ulmus americana* broad specificity thioesterase with *Chlorella protothecoides* stearyl ACP desaturase transit peptide.

ACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACC
TGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGAGGCCCTCCCCGTGCG
CGGGCGCGCCAGCTGCCCGACTGGAGCATGCTGCTGGCCGCGATCACCACCCT
GTTCTTGGCGGCCGAGAAGCAGTGGATGATGCTGGACTGGAAGCCCAAGCGCCC
CGACATGCTGGTGGACCCCTTCGGCCTGGGCGCTTCGTGCAGGACGGCCTGGTG
TTCCGCAACAACCTTCAGCATCCGCAGCTACGAGATCGGCGCGGACCGCACCCGCC
AGCATCGAGACCCTGATGAACCACCTGCAGGAGACCGCCCTGAACCACGTGAAG
AGCGTGGGCGCTGCTGGAGGACGGCCTGGGCAGCACCCGCGAGATGAGCCTGCGC
AACCTGATCTGGGTGGTGACCAAGATGCAGGTGGCGGTGGACCGCTACCCACC
TGGGGCGACGAGGTGCAGGTGAGCAGCTGGGCGACCGCCATCGGCAAGAACGG
CATGCGCCGCGAGTGGATCGTGACCGACTTCCGCACCGGCGAGACCCTGCTGCG
CGCCACCAGCGTGTGGGTGATGATGAACAAGCTGACCCGCCGCGATCAGCAAGAT
CCCCGAGGAGGTGTGGCACGAGATCGGCCCCAGCTTCATCGACGCGCCCCCCCCT
GCCACCGTGGAGGACGACGGCCGCAAGCTGACCCGCTTCGACGAGAGCAGCGC
CGACTTCATCCGCAAGGGCCTGACCCCCCGCTGGAGCGACCTGGACATCAACCA
GCACGTGAACAACGTGAAGTACATCGGCTGGCTGCTGGAGAGCGCGCCCCCGA
GATCCACGAGAGCCACGAGATCGCCAGCCTGACCCTGGAGTACCGCCGCGAGTG
CGGCCGCGACAGCGTGCTGAACAGCGCCACCAAGGTGAGCGACAGCAGCCAGCT
GGGCAAGAGCGCCGTGGAGTGCAACCACCTGGTGCGCCTGCAGAACGGCGGCGCA
GATCGTGAAGGGCCGCACCGTGTGGCGCCCCAAGCGCCCCCTGTACAACGACGG
CGCCGTGGTGGACGTGCCCGCCAAGACCAGCTGACTCGAG

SEQ ID NO: 48

Codon optimized *Prototheca moriformis* (UTEX 1435) delta 12 fatty acid desaturase transit peptide cDNA sequence.

ACTAGTATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTCACG
ATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTC
GTGGGCGCGCC

SEQ ID NO: 49

Codon optimized *Chlorella protothecoides* (UTEX 250) stearoyl ACP desaturase transit peptide cDNA sequence.

ACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGCGACC
TGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCG
CGGGCGCGCC

SEQ ID NO: 50

Revelant homologous recombination expression construct of codon optimized coding region of *Ulmus americana* broad specificity thioesterase.

GCTCTTCGGCCCGCCGCACTCCTGCTCGAGCGCGCCCGACTCGCGCTCCGCCTGC
GCCCCGCGCTGCGCCGCCAGCGCCTTGGCCTTTTCGCCGCGCTCGTGCGCGTTCG
TGATGTCCATCACCAGGTCCATGAGGTCTGCCTTGCGCCGGCTGAGCCACTGCTT
CGTCCGGGCGGCCAAGAGGAGCATGAGGGAGGACTCCTGGTCCAGGGTCTCTGAC
GTGGTTCGCGGCTCTGGGAGCGGGCCAGCATCATCTGGCTCTGCCGCACCGAGGC
CGCCTCCAACCTGGTCTCCAGCAGCCGAGTCGCCGCCGACCCTGGCAGAGGAA
GACAGGTGAGGGGGGTATGAATTGTACAGAACAACCACGAGCCTTGTCTAGGCA
GAATCCCTACCAGTCATGGCTTTACCTGGATGACGGCCTGCGAACAGCTGTCCAG
CGACCCTCGCTGCCGCCGCTTCTCCCGCACGCTTCTTTCCAGCACCGTGATGGCG
CGAGCCAGCGCCGCACGCTGGCGCTGCGCTTCGCCGATCTGAGGACAGTCGGGG
AACTCTGATCAGTCTAAACCCCTTGC GCGTTAGTGTTGCCATCCTTTGCAGACC
GGTGAGAGCCGACTTGTGTGCGCCACCCCCACACCACCTCCTCCCAGACCAAT
TCTGTACCTTTTTTGGCGAAGGCATCGGCCTCGGCCTGCAGAGAGGACAGCAGTG
CCCAGCCGCTGGGGGTGGCGGATGCACGCTCAGGTACCCTTCTTTCGCTATGA
CACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGC
AACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCC
GATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGA
CATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTA
CACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCTCTT
CGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCCTTCTGTTC
CTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCC
GACCGCCCCCTGGTGCACCTTACCCCCAACAAGGGCTGGATGAACGACCCCAAC
GGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAAC
CCGAACGACACCGTCTGGGGGACGCCCTTGTTCGTTGGGGCCACGCCACGTCCGAC
GACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGAC
TCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCT
TCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACA
CCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTT
CACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGA
CCCGAAGGTCTTCTGGTACGAGCCCTCCAGAAGTGGATCATGACCGCGGCCAA
GTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAA
GCTGGAGTCCGCGTTCGCCAACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCC
GGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCCTACTGGGTGATG
TTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCG
TCGGCAGCTTCAACGGCACCCACTTCGAGGCCCTTCGACAACCAGTCCCGCGTGGT
GGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACC
TACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCTTCG
TGCCCAACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAAC

ACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCG
ATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACG
TTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACAGCACCGGCACCCTG
GAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGT
TCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCG
CATGGGCTTCGAGGTGTCCGCGTCCTCCTTCTTCTGACCGCGGGAACAGCAAG
GTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAAC
CAGCCCTTCAAGAGCGAGAACGACCTGTCTACTACAAGGTGTACGGCTTGCTGG
ACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTTCGTGTCCACCAACA
CCTACTTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGG
TGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATT
GGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTTCGTGTGATGG
ACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCA
AACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGC
TTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTT
GCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGC
TCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGG
TACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGCGAT
GGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGCGCGCAGAGGAACGCT
GAAGGTCTCGCCTCTGTTCGCACCTCAGCGCGGCATACACCACAATAACCACCTGA
CGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCGGTTCACACACGTGCCA
CGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCTGATGGTCGAAACGTTT
ACAGCCTAGGGATATCGAATTCTTTCTTGCGCTATGACACTTCCAGCAAAAAGGT
AGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTC
GACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCG
AGCGCTGTTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACC
AAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCT
TGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCTCTTCGTTTCAGTCACAACCC
GCAAACACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCG
GCGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCC
CCGTGCGCGGGCGCGCCCAGCTGCCCCGACTGGAGCATGCTGCTGGCCGCGATCA
CCACCCTGTTCTTGGCGGCCGAGAAGCAGTGATGATGCTGGACTGGAAGCCCA
AGCGCCCCGACATGCTGGTGGACCCCTTCGGCCTGGGCGGCTTCGTGCAGGACGG
CCTGGTGTTCGCAACAACCTTCAGCATCCGCAGCTACGAGATCGGCGCGGACCGC
ACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAGACCGCCCTGAACCAC
GTGAAGAGCGTGGGCCTGCTGGAGGACGGCCTGGGCAGCACCCGCGAGATGAGC
CTGCGCAACCTGATCTGGGTGGTGACCAAGATGCAGGTGGCGGTGGACCGCTAC
CCCACCTGGGGCGACGAGGTGCAGGTGAGCAGCTGGGCGACCGCCATCGGCAAG
AACGGCATGCGCCGCGAGTGGATCGTGACCGACTTCCGCACCGGCGAGACCCTG
CTGCGCGCCACCAGCGTGTGGGTGATGATGAACAAGCTGACCCGCCGATCAGC
AAGATCCCCGAGGAGGTGTGGCACGAGATCGGCCCCAGCTTCATCGACGCGCCC
CCCCTGCCACCGTGGAGGACGACGGCCGCAAGCTGACCCGCTTCGACGAGAGC
AGCGCCGACTTCATCCGCAAGGGCCTGACCCCCGCTGGAGCGACCTGGACATC
AACCAGCACGTGAACAACGTGAAGTACATCGGCTGGCTGCTGGAGAGCGCGCCC
CCCGAGATCCACGAGAGCCACGAGATCGCCAGCCTGACCCTGGAGTACCGCCGC
GAGTGCGGCGCGACAGCGTGTGAACAGCGCCACCAAGGTGAGCGACAGCAGC
CAGCTGGGCAAGAGCGCCGTGGAGTGAACACCTGGTGCGCCTGCAGAACGGC
GGCGAGATCGTGAAGGGCCGACCGTGTGGCGCCCCAAGCGCCCCCTGTACAAC
GACGGCGCCGTGGTGGACGTGCCCCGCAAGACCAGCGATGACGATGACAAGCTG
GGATGACTCGAGTTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCGACAC

ACTCTGGACGCTGGTTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACC
TGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGT
ACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGC
ATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGT
CCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTG
GTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAA
TGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTGTAGAGCT
CCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGATAACC
TCCAAAGCCGCTCTAATTGTGGAGGGGGTTTGAATTTAAAAGCTTGGAATGTTGG
TTCGTGCGTCTGGAACAAGCCCAGACTTGTTGCTCACTGGGAAAAGGACCATCAG
CTCCAAAAAACTTGCCGCTCAAACCGCGTACCTCTGCTTTCGCGCAATCTGCCCT
GTTGAAATCGCCACCACATTCATATTGTGACGCTTGAGCAGTCTGTAATTGCCTC
AGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCGCGGGC
GAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAGTTCATA
ACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAGTGGC
CACCCCCCGGCCCTGGTGCTTGCAGGAGGGCAGGTCAACCGGCATGGGGCTACCG
AAATCCCCGACCGGATCCCACCACCCCCGCGATGGGAAGAATCTCTCCCCGGGA
TGTGGGCCCACCACCAGCACAACTGTGGCCCAGGCGAGCGTCAAACCATAACC
ACACAAATATCCTTGGCATCGGCCCTGAATTCTTCTGCCGCTCTGCTACCCGGT
GCTTCTGTCCGAAGCAGGGGTTGCTAGGGATCGCTCCGAGTCCGCAAACCTTGT
CGCGTGGCGGGGCTTGTTTCGAGCTTGTTTCGAGCTTGAAGAGCCTCTAGAGTCGAC
CTGCAGGCATGCAAGCTTGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT
TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAG
GCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGC
CCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACG
CGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 51

Revelant homologous recombination expression construct of codon optimized coding region of *Cinnamomum camphora* C14 preferring thioesterase.

GAATTCGCCCTCCCGTGATCACACAGGTGCCTTGCGAGCGTGATCACACTATTTT
GGGGGTCTTACAGTACTGAAATGGTGAGAAGTCGTAAGTCAAGGATGAAC
AATGAAAATGGTGCTGTGGTGGCTTCTCAAAGGTCAAGAATCAGTCGCTCGCGTC
AGGAAATCGCGGCGTCAACCAGCGTGGGCGCGGTCAAGTGGCCCCGCACTGGTCA
CCATAGCCTCTCCTGCCACAGTAGCGATCCCCTGGGCGTTCACTCTCAGCAGCGG
CTGTACTGCCTCCAGATTTTCTTCTTCTGGACCTGCGGGCGTGAGAGGATGAGC
AGGGTGGGCCAAGGGCTCAATCCTGAACGGCCCTCATTCGGTTTCCAATCCCACA
ACACATACCCACAGCAGGTGAGACCACGCATTCCACCATGCGCACCAATAACGT
GTCTTACCTGATTGGGTGTGGCAGGCTCCGTGGACAGGAGTGCCTCGTCCCCCG
CCCAGACCCGCTCCCCCGTCACGGCGGCGTCCGGGACCCGCAGCGGCTCCACCG
CGGTGTGATCCGCGTTGGCGGCGCAGAGCAGCATCCCAGCCGATTTGACCCCGC
GCATGCTCCGAGGCTTGAGGTTGGCCAGCACCACCACCCGCCGGCCGACAAGGT
CCTCCAGGGTCACGTGCCGGACCAGGCCACTCACGATGGTGCGAGGGCCCCCT
CCTCGCCGAGGTCGATCTGCTCGACGTACAGACTGCGACATGCGTGGCGAGTGGT
CATCAGAAGGAAGCAGGTGTGAGAAGGGGCACGTGGTTGGTATTGAGAGTAGC
CAAAGCTTTGTGCCAATCAGAAAGTCAACGCAGCTGCCTGCCTGGCTCGCGTACA
ATTCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACG
GCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTC
GGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCC

AGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCATATTCAAACACC
TAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAA
GGGGGCGCCTCTTCCCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTG
CTGCAGGCCTTCCCTGTTCCCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCA
TGACGAACGAGACGTCCGACCGCCCCCTGGTGCACTTCACCCCCAACAAAGGGCT
GGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACC
TGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCCTGGGG
CCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGC
CCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAAC
AACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCA
TCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGG
ACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACT
CCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGAT
CATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGA
CCTGAAGTCCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGGCTTCTCCTGGCTAC
CAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAG
TCCTACTGGGTGATGTTTCATCTCCATCAACCCCGCGCCCCGGCCGGCGGCTCCT
TCAACCAGTACTTCGTGCGCAGCTTCAACGGCACCCACTTCGAGGGCCTTCGACAA
CCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTC
AACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGG
GAGTACTCCGCCTTCGTGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGC
GCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCA
ACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGT
TCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTGACCTGTCCA
ACAGCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGA
CGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGA
CCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCCTG
GACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAAC
CGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGAGAACGACCTGTCTACTAC
AAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGC
GACGTGCTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCG
TGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGTGC
GCGAGGTCAAGTGATTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCGAC
ACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGA
CCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGT
GTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCA
GCATCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTG
TCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTT
GGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCA
ATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTC
CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCTCTTCGTTTCAGTCACAACCCGCAAACACTAGTATGACGTTCCG
GGTCGCCCTCCCGGCCATGGGCGCGGTGTCTCCCTTCCCCGGCCAGGGTCGCG
GTGCGCGCCAGTCCGGCGAGTCAGGTTTTGGAGAGCGGGCGCGCCCCCGACTGG
TCCATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGA
CCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCAGCTGCTGGACGACCACT

TCGGCCCCACGGCCTGGTGTTCGCGCCGACCTTCGCCATCCGCAGCTACGAGGT
GGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGC
CGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCAC
CCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGC
CGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTGGAGTGCTGGGTGGG
CGCCTCCGGCAACAACGGCCGCCGCCACGACTTCCTGGTGCGCGACTGCAAGAC
CGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACC
CGCCGCCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCCGCCTTC
ATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAA
CGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCT
GGACATCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGAC
CGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTACCATCGAGTAC
CGCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGC
GGCTCCTCCGAGGCCGGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGC
AGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTC
CGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCAC
GACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTG
ACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTG
TGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTT
TTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCT
AGCTCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATAT
CGCTTGCAATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCT
CCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCT
CCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGT
GGGATGGGAACACAAATGGAAAGCTGGTACCCGTACCCATCAGCATCCGGGTGA
ATCTTGGCCTCCAAGATATGGCCAATCCTCACATCCAGCTTGGCAAAATCGACTA
GACTGTCTGCAAGTGGGAATGTGGAGCACAAGGTTGCTTGTAGCGATCGACAGA
CTGGTGGGGTACATTGACAGGTGGGCAGCGCCGCATCCATCGTGCCTGACGCGA
GCGCCGCCGGTTGCTCGCCCGTGCTGCGCTCAAAGAGCGGCAGAGAAATCGGG
AACCGAAAACGTACATTGCCTGATGTTGTTACATGCTGGACTAGACTTTCTTGG
CGTGGGTCTGCTCCTCGCCAGGTGCGCGACGCCTCGGGGCTGGGTGCGAGGGAG
CCGTGCGGCCACGCATTTGACAAGACCCAAAGCTCGCATCTCAGACGGTCAACC
GTTCTGATTATACATTCAACATATGGTACATACGCAAAAAGCATGCCAACGATGA
CATAGGCGAATTC

SEQ ID NO: 52

Relevant expression construct for codon optimized coding region of *Cuphea hookeriana* C10 preferring thioesterase with *Chlorella protothecoides* stearyl ACP desaturase transit peptide.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCACCCGCCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCCTGCCTGCCGCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTGAGCAAAAGGTAGCAGTTT
GCCAGAATGATCGGTTGAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
GCAAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTC
ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT

GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
 CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCCTCT
 GGCTGCTTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
 GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCACACTGT
 CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
 GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
 AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
 GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTTACATTTCATT
 TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCCGAATGG
 AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
 TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTTCGTGT
 GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
 CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
 TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
 TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
 CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
 CGCCCCGTGCTGTTCTGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
 ACGAGGCCCGCCGCCTGTCCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
 GCTGGACGTGGTGACCGAGGCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
 CGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
 GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCTTCGACC
 ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
 TGGACCAGGACGACCTGGACGAGGAGCACCAAGGGCCTGGCCCCCGCCGAGCTGT
 TCGCCCGCCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
 GCGACGCCTGCCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
 CGACTGCGGCCGCTGGGCGTGCCGACCGCTACCAGGACATCGCCCTGGCCAC
 CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCGGACCGCTTCTGGTGCT
 GTACGGCATCGCCGCCCCCGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
 GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
 GCTGGTCTGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
 CCTTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
 TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCC
 CTCGTTTTCATATCGCTTGCATCCCAAACCGCAACTTATCTACGCTGTCTGTATCC
 CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTC
 CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
 CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
 GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
 CACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG
 GTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
 GATGGTCGAAACGTTACACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
 ACTTCCAGCAAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
 ACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
 ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGAC
 ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
 ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTTCTCTCTTC
 GTTTCAGTCACAACCCGCAAACACTAGTATGGCCACCGCATCCACTTCTCGGCG
 TTCAATGCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGGGGCCCCGGCGCC
 CAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCAGCTGCCCCACTGGAGCCGCC
 TGCTGACCGCCATCACCAACCGTGTTCTGTGAAGTCCAAGCGCCCCGACATGCACGA
 CCGCAAGTCCAAGCGCCCCGACATGCTGGTGGACAGCTTCGGCCTGGAGTCCAC

CGTGCAGGACGGCCTGGTGTTCGCGCAGTCCTTCTCCATCCGCTCCTACGAGATC
GGCACCCGACCGCACCAGCATCGAGACCCTGATGAACCACCTGCAGGAGACC
TCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTGGACGGCTTCGGCCGCACCC
TGGAGATGTGCAAGCGCGACCTGATCTGGGTGGTGATCAAGATGCAGATCAAGG
TGAACCGCTACCCCGCCTGGGGCGACACCGTGGAGATCAACACCCGCTTCAGCC
GCCTGGGCAAGATCGGCATGGGCCGCGACTGGCTGATCTCCGACTGCAACACCG
GCGAGATCCTGGTGCAGCGCCACCAGCGCCTACGCCATGATGAACCAGAAGACCC
GCCGCCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTGCCCTGTTCGT
GGACAGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGAA
GACCGGCGACAGCATCCAGAAGGGCCTGACCCCGGCTGGAACGACCTGGACGT
GAACCAGCACGTGTCCAACGTGAAGTACATCGGCTGGATCCTGGAGAGCATGCC
CACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCCCTGGCCCTGGAGTACCGCCG
CGAGTGCAGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGGACCCAGCAA
GGTGGGCGTGCGCTCCAGTACCAGCACCTGCTGCGCCTGGAGGACGGCACCCGC
CATCGTGAACGGCGCCACCGAGTGGCGCCCCAAGAACGCCGGCGCCAACGGCGC
CATCTCCACCGGCAAGACCAGCAACGGCAACTCCGTGTCCATGGACTACAAGGA
CCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACA
AGTGAAGTGCAGGAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGT
CGTGTGATGGACTGTTGCCGCCACACTTGCTGCCCTTGACCTGTGAATATCCCTGC
CGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAG
TTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCGAGCATCCCTTCCCTCGTT
TCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGC
GCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGTTTGGGCTCCGCCTG
TATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAA
GTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTC

SEQ ID NO: 53

Relevant expression construct for codon optimized coding region of *Umbellularia californica* C12 preferring thioesterase with *Chlorella protothecoides* stearyl ACP desaturase transit peptide.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCACCCGCGCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCCTGCCGCGATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCATCGCGGCAGCCAGAGGGA
TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTGAGCAAAGGTAGCAGTTT
GCCAGAATGATCGGTTTCACTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTC
ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCCTCT
GGCTGCTTGTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCACACTGT
CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
GTCCCITGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTTCATT

TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCCGCAATGG
AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
TTCCATAGTGCAATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCGTGTT
GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGCCTGTCCGCCAGGGC
CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCCGCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCCCTCGACC
ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCGCCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGCT
GTACGGCATCGCCGCCCGGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTCCGACCTCAGCGCGGCATACAC
CACAATAACCACCTGACGAATGCGCTTGGTTCTTTCGTCCATTAGCGAAGCGTCCG
GTTACACACAGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCCTTTCTTGCGCTATGAC
ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGAC
ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
GTTTCAGTCACAACCCGCAAACACTAGTATGGCCACCGCATCCACTTTCTCGGCG
TTCAATGCCCCGTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCC
CAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCCCCGACTGGTCCATGCTGTTCCG
CGTGATCACCAACATCTTCAGCGCCGCCGAGAAGCAGTGGACCAACCTGGAGTG
GAAGCCCAAGCCCAAGCTGCCCCAGCTGCTGGACGACCACTTCGGCCTGCACGG
CCTGGTGTTCGCGCGACCTTCGCCATCCGCTCCTACGAGGTGGGCCCCGACCGC
AGCACCTCCATCTGGCCGTGATGAACCACATGCAGGAGGCCACCTGAACCAC
GCCAAGAGCGTGGGCATCTGGGCGACGGCTTCGGCACCACTTGGAGATGTCC
AAGCGCGACCTGATGTGGGTGGTGCGCCGACCCACGTGGCCGTGGAGCGCTAC
CCCACCTGGGGCGACACCGTGGAGGTGGAGTGCTGGATCGGCGCCAGCGGCAAC
AACGGCATGCGCCGCGACTTCCTGGTGCGCGACTGCAAGACCGGCGAGATCCTG
ACCCGCTGCACCTCCCTGAGCGTGCTGATGAACACCCGCACCCGCGCCTGAGCA
CCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCCGCCTTCATCGACAACGTGG

CCGTGAAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCG
 ACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACGTGAACCAGC
 ACGTGAACAACCTGAAGTACGTGGCCTGGGTGTTTCGAGACCGTGCCCGACAGCA
 TCTTCGAGTCCCACCACATCAGCTCCTTCACCCTGGAGTACCGCCGCGAGTGCAC
 CCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCAGCTCCGAGGC
 CGGCCTGGTGTGCGACCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCG
 CGCCCGCACCAGGTGGCGCCCCAAGCTGACCGACTCCTTCCGCGGCATCAGCGTG
 ATCCCCGCGGAGCCCCGCGTGATGGACTACAAGGACCACGACGGCGACTACAAG
 GACCACGACATCGACTACAAGGACGACGACGACAAGTGACTCGAGGCAGCAGC
 AGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCG
 CCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCA
 GTGTGTTTGATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATT
 TGCGAATACCAACCCCGAGCATCCCCCTCCCTCGTTTCATATCGCTTGATCCCAAC
 CGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCAC
 TGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCGCTGTATTCCTGCTACTGCAACC
 TGTAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGCGATGGGAACACAA
 ATGGAAAGCTTGAGCTC

SEQ ID NO: 54

Relevant expression construct for codon optimized coding region of *Ulmus americana* broad specificity thioesterase with *Chlorella protothecoides* stearyl ACP desaturase transit peptide.

GGTACCCGCTTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCGCTGAAC
 CGACACGTGCTTGGGCGCCTGCCGCTGCTGCCGCATGCTTGTGCTGGTGAGGC
 TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAGCTTATGTG
 TGTGCTGGGCTTGATGCCGGGCAATGCGCATGGTGGAAGAGGGCGGCAGCAC
 TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
 TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTGAGCAAAGGTAGCAGTTT
 GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAAGAGAAGGGGTCAAGT
 GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
 AGTGTGTGCGAGCCAGCATCGCCGCTGGCTGTTTCAGCTACAACGGCAGGAGTC
 ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
 AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
 GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
 CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCCTCT
 GGCTGCTTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
 GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCACACTGT
 CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
 GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
 AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
 GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTTCATT
 TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGACCCCGCAATGG
 AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
 TTCCATAGTGCAATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCGTGTT
 GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
 CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGGCCCTCCTTTC
 TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
 TCCACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
 CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC

CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCCGCCCGCCTGTCCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCCTGCACACCCCTGGACCCCGCCACCTGCCCCCTTCGACC
ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCGATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCGCTGAAGGCCCGCATGCCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCGCCTGGGCGTGCCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGGCGAGTGGGCGGACCGCTTCTGGTGCT
GTACGGCATCGCCGCCCGGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCGAGCATCCCCCTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGTCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTTCGCACCTCAGCGCGGCATACAC
CACAATAACCACCTGACGAATGCGCTTGTTCTTTCGTCCATTAGCGAAGCGTCCG
GTTACACACAGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
GATGGTTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCGGATTGCAAAGAC
ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
GTTTCAGTCACAACCCGCAAAACACTAGTATGGCCACCGCATCCACTTCTCGGCG
TTCAATGCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCC
CAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCAGCTGCCCGACTGGAGCATGC
TGCTGGCCGCGATCACCACCCTGTTCTTGGCGGCCGAGAAGCAGTGGATGATGCT
GGACTGGAAGCCCAAGCGCCCCGACATGCTGGTGGACCCCTTCGGCCTGGGCGG
CTTCGTGCAGGACGGCCTGGTGTTCGCAACAACCTCAGCATCCGCAGCTACGAG
ATCGGCGCGGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAG
ACCGCCCTGAACCACGTGAAGAGCGTGGGCCTGCTGGAGGACGGCCTGGGCAGC
ACCCGCGAGATGAGCCTGCGCAACCTGATCTGGGTGGTGACCAAGATGCAGGTG
GCGGTGGACCGCTACCCACCTGGGGCGACGAGGTGCAGGTGAGCAGCTGGGCG
ACCGCCATCGGCAAGAACGGCATGCGCCGCGAGTGGATCGTGACCGACTTCCGC
ACCGGCGAGACCCTGCTGCGCGCCACCAGCGTGTGGGTGATGATGAACAAGCTG
ACCGCCGCGATCAGCAAGATCCCCGAGGAGGTGTGGCACGAGATCGGCCCCAGC
TTCATCGACGCGCCCCCTGCCACCCTGGAGGACGACGGCCGCAAGCTGACC
CGCTTCGACGAGAGCAGCGCCGACTTCATCCGCAAGGGCCTGACCCCCGCTGG
AGCGACCTGGACATCAACCAGCACGTGAACAACGTGAAGTACATCGGCTGGCTG
CTGGAGAGCGCGCCCCCGAGATCCACGAGAGCCACGAGATCGCCAGCCTGACC
CTGGAGTACCGCCGCGAGTGCGGCCGCGACAGCGTGCTGAACAGCGCCACCAAG
GTGAGCGACAGCAGCCAGCTGGGCAAGAGCGCCGTGGAGTGCAACCACCTGGTG
CGCCTGCAGAACGGCGGCGAGATCGTGAAGGGCCGCACCGTGTGGCGCCCCAAG
CGCCCCCTGTACAACGACGGCGCCGTGGTGGACGTGCCCGCCAAGACCAGCGAT

GACGATGACAAGCTGGGATGACTCGAGGCAGCAGCAGCTCGGATAGTATCGACA
CACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGAC
CTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTG
TACGCGCTTTTGGCAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGC
ATCCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGT
CCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTG
GTTTGGGCTCCGCCTGTATTCTCTTGGTACTGCAACCTGTAAACCAGCACTGCAA
TGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTC

SEQ ID NO: 55

Prototheca moriformis FatB/A promoter/5'UTR

CCTGTCGATCGAAGAGAAGGAGACATGTGTACATTATTGGTGTGAGGGCGCTGA
ATCGGCCATTTTTTAAATGATCACGCTCATGCCAATAGACGCGGCACATAACGA
CGTTCAAACCCCCGCCAAAGCCGCGGACAACCCCATCCCTCCACACCCCCCACAC
AAAGAACCCGCCACCGCTTACCTTGCCACGAGGTAGGCCTTTCGTTGCGCAAAA
CCGGCCTCGGTGATGAATGCATGCCCGTTCCTGACGAGCGCTGCCCGGGCCAACA
CGCTCTTTTGTGCGTCTCCTCAGGCTTGGGGGCCTCCTTGGGCTTGGGTGCCGCC
ATGATCTGCGCGCATCAGAGAAACGTTGCTGGTAAAAAGGAGCGCCCCGGCTGCG
CAATATATATATAGGCATGCCAACACAGCCCAACCTCACTCGGGAGCCCGTCCCA
CCACCCCCAAGTCGCGTGCCTTGACGGCATACTGCTGCAGAAGCTTCATGAGAAT
GATGCCGAACAAGAGGGGACAGAGGACCCAATCCCGGACATCCTTGTCGATAAT
GATCTCGTGAGTCCCCATCGTCCGCCCGACGCTCCGGGGAGCCCGCCGATGCTCA
AGACGAGAGGGCCCTCGACCAGGAGGGGCTGGCCCCGGGCGGGCACTGGCGTCG
AAGGTGCGCCCGTTCGTTGCGCTGCAGTCTATGCCACAAAACAAGTCTTCTGACG
GGGTGCGTTTGTCTCCCGTGCGGGCAGGCAACAGAGGTATTCACCCTGGTCATGGG
GAGATCGGCGATCGAGCTGGGATAAGAGATACTTCTGGCAAGCAATGACAACTT
GTCAGGACCGGACCGTGCCATATATTTCTCACCTAGCGCCGCAAAACCTAACAAT
TTGGGAGTCACTGTGCCACTGAGTTCGACTGGTAGCTGAATGGAGTCGCTGCTCC
ACTAAACGAATTGTCAGCACCGCCAGCCGGCCGAGGACCCGAGTCATAGCGAGG
GTAGTAGCGCGCC

SEQ ID NO: 56

Prototheca moriformis NRAMP metal transporter promoter/5'UTR

ACTAATTGCAATCGTGCAGTAATCATCGATATGGTCACAAGTAGATCCCCTACTG
ACACCCTCTCGTACATGTAGGCAATGTCATCGGCGCCGTCCTGCTGACCGATGCC
GACGTAGCAGAGCAGACCCGGGCGGATCTGGGATACGAGCCGGCCCTCCACCTG
CGCTCGAGGTGGAATCAAGTAAATAACCAATACACTTTTTCGACACCACACAGAG
TTGCACGGACGGTGGCGTACCTCTACGCTCGCGCTCTTCACGCGCTGGACGACCG
CACGCATGAGCCCGGGTGGCTTGGTCTGGGCTGCAAAAATGCACAACAAACAAG
TATCAGACGCTCATGGATGCACACGCGCTCCCAAGCACGCTCAGACTAAATATTA
CAGTAGCTCGTATCTGATAAGATATCGAGACATACCGCTCAACTACCCGCAAAC
TGCGCCCCCGCCAGGTGATGCGCACAGGGCCCCACCATGCGATCCATCGCATCGCT
CCTCGAGGGCGCTATCACGTGGCCGGAGAGCGTTACAGCGTACGCCACTGTATC
TGGGCGGTATGCGGTCCGTCAACATGGAGACAGATACCCGCACCACCACCTTGC
AAGCTCTTCCATATTGGAAGTAGAAAATTGTAAATTGTATCATCGCACGAGGGGCC
AACTTGCCGTCGGCGAGCTGGGCGACGAACACCACCTGGACGTTGTGAGACTC
GCTCGTGCCGTGCGCCGGGCGGCTGGGTATCCAGACCGTCGCC

SEQ ID NO: 57

Prototheca moriformis FLAP Flagellar-associated protein promoter/5'UTR

CAACGACAACCAGCAGGCAACTCGGTCAGCGACCCAACACGCGAGTCAAATTGT
TGCGTGTCTTGCCTTGTCTATTTACTGTGATAGCAAGACTGTCGGTCAGTCAATA
CCGCGGTGCGCACGTCGGGGTGCCAAGCCTAGCAGAGCACGGGACGGCTGGTGC
TGTGCGCCAGCTCAGCTCGCTTCGCGACCAATTGTAGGACCGGCAAAGTCAACAA
AACATGCCAGCGGTGCGATTCAATTGGTCATGAGCTCTACAAAATTGTTTTGTGC
GTGCGCGCAGGTATCCAACGGCGCGGCAGAGAAAGTTTGACAGCTCTCGATTTCAT
CTCGGAAAAATGGGGAGAATTTATGACACACAAGTGCGCAGGCGGCCAGGCGG
CCAGCATATTCTGGCGTGACCTGGGCCGCCACAAAATGCTTGGATGCACTCTAA
AATAATTATATTTGCCATGAACAAGGGAAGAGTTACCGCACCCAGCCCTAGACTT
GGGCGCCCCGAGCAAGGTTACGTCAAGCCACCTTCGCCCATCGCCCAACTCCGTAT
TCCCCGACAGCCGCACGTGGCCCTCGCCGGAATGAACCCTGAATCGGCATCACG
CCACGCGTTCGCCAATCGTTCCGCTCTCTGGCTTCATCGGCCTGCGCCTTCACGTC
GTGGTCACGACAGTGCATTCATACTTCCATTTGCACCTCGGCACACACTTTTACG
CATCGCCTACCTTGCTGCGGCAGTCTAGGGTCACTTTGCAGCCATGGGACAGTG
CTACACCACCGTCGGTGCGCAAAGCTATTTCAAGTGAACCGTGGGCGGAAAAAA
GGAATGTACACTGTCTCAACCGACTCCTACAATTGTTTACCATGCAGATCAGAGC
TCGACGGCCATCATCGAGCAGGTGTGGGGCCTTGGTGGCGCGGCGCGGGGCCCC
AGGGCGTCGCAGGCATTGATGGCACTCTGAGACTTTCGCACGCGCATGAGGGAC
CCCATCAAGAGAAGAGTGTGTCTTTATGTCCCCATTTCATGATGATGTATCTTGTG
ATTGTGCGCAGTTTGGCAAGTTTAACCGGATCGCCGCTCCAGGTGTGGCGTGGCGG
ATTTTTCTAGGGGTGCTTGAGCAGTCG

SEQ ID NO: 58

Prototheca moriformis SulRed Sulfite reductase promoter/5'UTR

GGCCCAGGGCCCTGCGGATGGCCCCACACCAGATCTAGCCTCTCTTATGCCATGCC
CGCCTCGCTGCCCCTCGTATCCCCCGCCGATCCGCGCGTAGGGGACCGCGGCCT
GACCCACGCCACGAAAGAGCTTTGCTCCTCAATTTCTCGCCAACAGAACCGTATC
AAACGCTCAACGCCTATCCCGAACAATCCGTATTACACCAAATCGAGTATACCG
GACTGGTTTGCCTAGTCTTGAAGGAAATGATCCCGTCCATGCTCGGAAGGGGGA
GCGGGCGGAGGATCCTACTCATCTCTGAAATGGGATTGGTCCGAAGATGGGTTG
GGCAAGCACGTGCCAAACCCAGCGAGTTGCTGACGAGCAGGCTCATCCAATCC
CCCGGCGAATCCTCCCTCACGCCCCGCATGCATACAAGTCCCTCCCACACGCCCC
CTCCCATCCATTTTCGCCTGGTCCGAACGCGAGCGGCGTCGAGGCGGACCACTTG
CTCCGCAGCGCCGTCTGGGTCTCCACCCACAGCGGCTTTGCTGCCAGAGGCACC
CCCCTTGCCCCACCTCCTCTTGACGCC

SEQ ID NO: 59

Prototheca moriformis SugT Sugar transporter promoter/5'UTR

CCAGGCAGGCGGTAGGGTTGCCGATTGCTTGAGCGAATTGGAAGATATAATTTTT
TGTGGTGTCCCTGGACGCTGTTTGTGGCGCTCCTTTTTGGAGAAGATTGCGTGGG
GGAGCTTTCCATGTACCACGCTTCCTTCTGAAAGGATTCTGGCCGAGTCTGATG
AGCCCAAAGAAAAACCTGCCCTTTCAGTGCTGGCACTCTGAAAACGTCAACAGA
TGATTATACATGTCACAAAAGGCAGCCGATTAGGAACGGGAGCTCTGGCCGTTT
GTTTGGCTGCCTGGGCTGATTGAAGTGATCCACCCTGTTTGAATGAAGGCGGTGC
AGTCGAATTATCGACCGGAGCTGTTCGGGAAGGCGTCCGGGGCAGAGTGAGGTGC

TGCGGCCTGGTTGTCGTTCAAAAAGACCCCGGTAGCCCAACAATCACGAACGAA
AGGAATATAATTGCTTGCATACTATACATTTCAGTTTCTATGTGGCGGGTAGACAA
GTCTCATGGGCTTCTAAAGGCTGTCCCTTGAAGGCTACTTATAAAAACTTGCTGC
GCCATGGCACGGATCGCGCTTGCAGGCTGCAACCCTGCGCGCAAGGTCAAAT
ACACAGCAAAAAGATACTAACAGAATTTCTAAAAACATTTAAATATTTGTTTCGAC
CAGCCAATTGTGGTCGTAGGCACGCAAAAGACTTTGTTTTGCGCCACCGAGCAT
CCACGCTGGCAGTCAAGCCAGTCCGATGTGCATTGCGTGGCAGCATCGAGGAGC
ATCAAAAACCTCGTGCACGCTTTTCTGTCAATCATCATCAACCACTCCACCATGT
ATACCCGATGCATCGCGGTGCGCAGCGCGCCACGCGTCCCAGACCCGCCAAAA
ACCCAGCAGCGGCGAAAGCAAATCTTCACTTGCCCGAAACCCCGAGCAGCGGCA
TTCACACGTGGGCGAAAACCCCACTTGCCCTAACAGGCGTATGTCTGCTGTCACG
ATGCTGACAACGGTATTATAGATATACACTGATTAATGTTTGAGTGTGTGCGAG
TCGCGAATCAGGAATGAATTGCTAGTAGGCACCTCCGACCGGGCGGGGGCCGAGG
GACCA

SEQ ID NO: 60

Prototheca moriformis Amt03-Ammonium transporter promoter/5'UTR

GGCCGACAGGACGCGCGTCAAAGGTGCTGGGCGTGTATGCCCTGGTCGGCAGGT
CGTTGCTGTTGCTGCGCTCGTGGTTCCGCAACCCTGATTTTGCGCTCTTATTCTGG
CGTGGCAAGCGCTGACGCCCGGAGCCGGGCGGCGGCGATGCGGTGTCTCACG
GCTGCCGAGCTCCAAGGGAGGCAAGAGCGCCGGATCAGCTGAAGGGCTTTACA
CGCAAGGTACAGCCGCTCCTGCAAGGCTGCGTGGTGGACTTGAACCTGTAGGTCC
TCTGCTGAAGTTCCCTCCACTACCTCACCAGGCCAGCAGACCAAAGCACAGGCTT
TTCAGGTCCGTGTCATCCACTCTAAAACACTCGACTACGACCTACTGATGGCCCT
AGATTCTTCATCAACAATGCCTGAGACACTTGCTCAGAATTGAACTCCCTGAAG
GGACCACCAGAGGCCCTGAGTTGTTCTTCCCCCGTGGCGAGCTGCCAGCCAGG
CTGTACCTGTGATCGAGGCTGGCGGGAAAATAGGCTTCGTGTGCTCAGGTCATGG
GAGGTGCAGGACAGCTCATGAAACGCCAACAATCGCACAAATTCATGTCAAGCTA
ATCAGCTATTTCTCTTACGAGCTGTAATTGTCCCAAAATTCTGGTCTACCGGGG
GTGATCCTTCGTGTACGGGCCCTTCCCTCAACCCTAGGTATGCGCGCATGCGGTC
GCCGCGCAACTCGCGCGAGGGCCGAGGGTTTGGGACGGGCGCTCCCGAAATGCA
GTTGCACCCGGATGCGCGGCGCCTTTCTTGCGATAATTTATGCAATGGACTGCTC
TGCAAATTTCTGGGTCTGTGCGCAACCCTAGGATCAGCGGCGTAGGATTTTCGTAA
TCATTCGTCTCTGATGGGGAGCTACCGACTACCCTAATATCAGCCCGGCTGCCTGA
CGCCAGCGTCCACTTTTGCGTACACATTCCATTTCGTGCCCAAGACATTTCAATTGTG
GTGCGAAGCGTCCCCAGTTACGCTACCTGTTTCCCGACCTCCTTACTGTTCTGTC
GACAGAGCGGGCCACAGGCCGGTTCGCAGCC

SEQ ID NO: 61

Prototheca moriformis Amt02-Ammonium transporter promoter/5'UTR

TCACCAGCGGACAAAGCACCGGTGTATCAGGTCCGTGTCATCCACTCTAAAGAG
CTCGACTACGACCTACTGATGGCCCTAGATTCTTCATCAAAAACGCCTGAGACAC
TTGCCAGGATTGAAACTCCCTGAAGGGACCACCAGGGGCCCTGAGTTGTTCTT
CCCCCGTGGCGAGCTGCCAGCCAGGCTGTACCTGTGATCGGGGCTGGCGGGAA
AACAGGCTTCGTGTGCTCAGGTTATGGGAGGTGCAGGACAGCTCATTAACGCC
AACAAATCGCACAAATTCATGGCAAGCTAATCAGTTATTTCCCATTAACGAGCTATA
ATTGTCCCAAAATTCTGGTCTACCGGGGGTGATCCTTCGTGTACGGGGCCCTTCCCT
CAACCCTAGGTATGCGCACATGCGGTGCGCGCAACGCGCGCGAGGGCCGAGG

GTTTGGGACGGGCCGTCCCGAAATGCAGTTGCACCCGGATGCGTGGCACCTTTTT
TGCGATAATTTATGCAATGGACTGCTCTGCAAAATTCTGGCTCTGTCGCCAACCC
TAGGATCAGCGGTGTAGGATTTCTGTAATCATTCGTCTGATGGGGAGCTACCGAC
TGCCCTAGTATCAGCCCGACTGCCTGACGCCAGCGTCCACTTTTGTGCACACATT
CCATTCGTGCCCAAGACATTTATTGTGGTGCGAAGCGTCCCCAGTTACGCTCAC
CTGATCCCCAACCTCCTTATTGTTCTGTCGACAGAGTGGGCCCAGAGGCCGGTCG
CAGCC

SEQ ID NO: 62

Prototheca moriformis Aat-01Amino Acid Transporter promoter/5'UTR

CGAAGGGGTCTGCATCGATTCTGCGCGGTCTGGAGGCCAGCGTGACTGCTCGCGA
AAATGCTCTGCCGTGTCGGGCTCTGGCTGGGGCGGCCAGAGATCTCACCGTGCCA
CACGCAACTGCCGCACTCTGTGCCCGCCACCTGGCGCGCACATGCGACCTCTTCC
CCGTCATACCCTCTCCTCATGTGATCTTTCCACACGAGTGACGCAGGTGCGCGGA
GTGGAGGGAATCAGGACGTTTTCAAGGTACCTGCTCGAGCCGTACCAACAGCTG
CCGCCCGGCAAGGAAGAGATCGAGGCGAGAGATTGCCCGGCTGGAGGCCCGGATA
ACGGAGCTCAAGAGCAAGCTGTCCGAGTGAGACCGCCCGAGGTGCACGTGTCGAC
TCGCTATGACATGTACTCGACACAACATGAGGAATTCATCGAATTTGTAGGAAGC
GGGCATTGGTACGGGAGTGGGAAAGCGAAAAAACCTCCCTCCGGCAGTGCCATC
TGCCGGAGTCGAACGTTGATAGGGTTCTCGTGACAGGGTGTGACCTCTCAGCCTT
GCATCAATTAACGCTATAGACATTATCAGTAACCGTGAATCCCGCATTGGATGC
CACCCGCGCGACCATTTGGGGACCTGCATTACAGATCTAGGTGAGATGACAGCGA
GGCAACTTCGGCCCGCGGCCCGAGCTTGCGGCGCACCAATATTGGTCACGGGAAG
CCACACACCGACCATAAATGAATACTTGTAAGCTATGTCAACCGATCAATGGCGT
CGAAAGTGTGCCACGAGGATCCATCTGGCGGGGCGGCGTGGCGCACAAGCGCAG
TCGCAATTTCTCGGACCCATCTGACCTAGGCCAGCGCCGCGGGAGAAATCCCCG
GCGGGTCTCCACGCAGTAACCCTAATGAGTATCGAGCGCCGACCATTTACACCA
TCGCCCCCGAAATCCTTCCGACATTATTATTATCTTTTAGATCTTGGAACAGACTC
TGCCAACC

SEQ ID NO: 63

Prototheca moriformis Aat02-Amino acid transporter promoter/5'UTR

AGAGAGCGGAGGTGGGGTTGTGAGGTGGGGTTGCTGACCAGGAGCTCGCGTCGC
CGAGCGGACTCGCACACGGTCCAGTTACCCCCCTCCGCCCAAACGCAAGCCT
CCCATCTTGATGCCTTTCCGGCCACCTATACTATTTCTTAGTTTCGCTGTAACATCC
AGACCGTCCTGAATAATAACAATGCCCTGTGTCAAGTGCATTCTAAAAAAATTC
TGTCCCAACCAACAATCCCACCTGAAATACCACCAGCCCTGCCCAGTACACTCTT
CCAATACCATCTCCCTACCTCCACGCGCAAGCGACCCCATGCGCGACCAAGGCTC
GAAAGTGATTTATGACTTGAGACGAGCGAGTGCGCGCGCGGTGACTGCCCTTTTC
ATCACGTGCCGTACGTGCGCGACCGCTAGGGCTTTGCACGGCAACGCACGGCTTC
GCCAACCCGACCAAGCAGGACCTCGACTACTCTACCGCGAATTCGCCTCAAGAA
GTGCCCAAATGTGCCATACACCATTCCTTACAGCACTGTTCAAACCTTGATGCCAA
TTTTGACATTCGGGTTGCTCGTTGGCTGCGCCACATCGGCCGTGAGTGCAGCAG
GCGGGATCGGACACGGAGGACGCGGCGTACGCCCCGAACGCAGCCCGTAATCTC
TACATCAACACGACGTGTTGCGTAATCCCGCCCGGCTGCGCATCGTGCCAACCCA
TTCCGCGATGGATGGTTCGGAATAATGGTGTGCCAACTGCCCTGAGGGAGGCTCTCG
CGAAACGGGCACGTCCCTGAAACCGAAACTGTGGCCCTTGTCGTGGGCCACGCAA
GCACGTGGACCCTAAACACCAAGAAAATCAGTAAACAAGGTTGACATCCTCTAC

GGGCGAATTGTTTGCCCAACCCTTCATCGCACACTGCCATTATAATGCATCTAGC
TCGGCGACAAGTTTAGAAAAAGGCAGGCTGCATTGTTCCATTTGCGCCGTGGCGGCG
TGGGTGCCCATTTTACGAGGTTTGGGCTCCCGGGCAGCGACCGAGCCAGGTCTGA
GTCCCTCTCGCCCGTCGACAATGTTGCGAACCCCAAGCGGCTAACAACAACCTT
GATGGTACCTGTACACTGCCAATTCCTTCTTCCCGGCCGAGGTTTACACGTGAT
GGCCATGGCTTCGCATTCAGGCCGACTTCCCATTCGACTTTCCAGAGGGTCCGC
GGACGCTGGGGGTTGGCTGCCTGAGGCCACCCCTTTGTTCCCGCGTCCCGACAA
ACACAATTGCGTTACATAAGGGGGAGCCGCCCCCGTTTCAGAGTGCAGAAATCTTT
CACTATATTTTCCAGTCGTCAGCGAAATCAAGT

SEQ ID NO: 64

Prototheca moriformis Aat03-Amino acid transporter promoter/5'UTR

GATGGTGGGGTGTCTGCCTTGGGCTGGGTGATGGAGGCTGGTGGTGCGCGGGTTT
CCTGATGCATTCTATCTACGCAGTGTTCATGGTGTCCATTCACACACCAGTACAC
CCTTACACTAAGGATCCATCCCTCCTTCCCTCTTCAGGACTACATGGACCCACG
AGCTACCGACCGGGCTTTCTCAAAAACGTCAAGGTTCATGTTTGACATGCGGGACG
TGGTGGACGACGTGCAAGGTGCGTCCGGAGTGCAGCGCAAATGAGCAAGTCGGGC
AATGTGTGCGGGTGGGCACCGGGGCTGGAGATCCGCGATCCCGAGAAAAACGCC
GTACCAACCCCGCGCTATTCCCTCGATTGCGCGCAGATGTGGTGACCGACACGG
GGGACAACCTGGCGGACATGGGGCGCCGGACCTGGAAGCACGCCAAGTCGCACA
CGGGGAGGCTCGTGCAGTCCCCCCCATCGTACCTCAAGGGTCTCTTTGGTCCGCA
TCCAAAGTACGCTGGTGGCATGGCATGCCCGAAATGAACATCATGTGTGATCTCC
GATTGCCAATGGCCACCTCCACGGACCACCTTGCAGGCGGAAGCGCAATCCAGG
GCCCAGCCTGACGAGGACGGAGACTCCTCGTCCAGCGCGGGGTCCCCGACCCG
ACGCAGCAGCCGACCCCTGCTAACCCGGCAACGATCGGACCAGCAACCTTGCTG
TAGTTCCGATCCGTGATGACGGGCATTGCCGCCGCTCGATCCGCTTTGATGACTG
TCTATTATTTGCGCGGAGCCCCCTCGGAACCCTACCCCGCTCTTGCAAGCCCCCTT
CATCGGAGATCCTCGTGCGCCCGCCATGACCCCACTGGATTGCCCAACATCCTTC
TTTATCGTGTAATAATGTGATTTCCTCGGCTGCAATCGACTGGCCTTCGCTTCTGGCC
CCAAGAGGGCTCGAACGTGCGGCAGCGAGGGCGCTGACACACCCAAGCCCTAGG
GCTTTCAACGTGCGCTGCCAGGCCGGATAGGGGGATCGCCTCCTTTCCACCACCC
ACCTACGAGGGATTTCGAGTCGGCTTCCAGCTCAGCTATTTCGGCCGCGCCCCGGC
CCTGCAGACGTCTCCAGTTTCCGAACAGGTCGCTCTCAGAACACCTGCCGCGGC
TGCGATACGGCAGGCTCTCAAAGCGTCGAC

SEQ ID NO: 65

Prototheca moriformis Aat04-Amino acid transporter promoter/5'UTR

CGCGTGGAGCGGTGCGTGCGGATGCCGCGCGCCTGCCAAGGCCTTTTGTATGCCT
GGCCTGGGAAGTTTCTGACTGAAGCATCTTCAAGATGCTCTCTCACGACCAGCG
ACACCAACACCGTCACTTTTGGCCCTCCTGCCGAGGTGCCACTTTCTACTTTGA
CGTCTTCTCCAGGCGGTACATTGCGGGACTGAGCGCCAATTCGGCCAAGAACAG
CGCTGTCGACTTGAGGAGGCAGGGGTCCGTCGACTCTGCCGAGTGACACGCCTTC
GACCCGACTGTACTACGGCCTGCTGAAGAGTGGGTCTCGCCGGCCGGCGTGACC
GGCCCTGTGCCACAATCGACCATCTATTGCTCCTTGTCATCTGGCGCCGTCAAT
TGCCCGCGACTTGACGGCAACTGGCTCGATCGAGTCGTATTGAAAAAGCACGTTT
TGTCCTACAGGGCCGCGGTCCGTTACCAACGTGGTTCTCGTTAGGTTTTTCGTCCG
GCGGTGGTGCAGCAACTGTCCGATGCCATCCCGGCAAACCCAGCAAGGTCGCC
AGTCTGGTTCTGACGCAATAGAGTGCGTTTTGGGCCAGTCTAAAAATTCGTCTGG

CATGACGTGGCTCCACATCGTACCCGGAGCCTGCCTTGGTAATGTGAGGCACCGG
TGCCAACTCCATTATGGCAGGCATCGAGCGCGCAGGTGAGTACATGACCTTCCGT
GAATTGGGAAGGCGAGCTTGTGTAACGCCTGCGATCGTGCCAGTGAGGCATCGT
AAACTCAAAATATTTTGTAGAAAGTGTCTGATGCCTGGTGAGGCTGCGTAGGGCA
AGGGCAAGCCCTTGGCAGATGGGTAATGGGTCCGGACCTCACAAACAGCAACCCC
GCGTCCCCCTTAGGGCCCCTGAGGCTCGATGGCAGGGCCAGCGAGCCCGCGGCC
AAAGGGCGCCATCCCACGGTCGCCCAACGACTCCACGGGTCTTATACCTCATCTT
GAATGGCACTAAAACTATAGAATATCGGGCACTGGTGGGCGTCTGGGGTACAG
CTGGCCGAGCGCAGTGGCAAACCCTAGGTCCCGCCTCAAGGGCGATTCCCGGGT
CAATGACACGCAAGCAAGATCACATGGCGCGGTCCGCCTCGCGGCTCCACACCC
AGGCCCTAGTTTCGCAACCCATAAATATCGCCCCGATACCATCATAAGCCAGCAA
ATAATTTTTTATCAGAGTTCCAAACCTCCTCAGCTGTGGGAAACCAGCCCCTCT
GAACG

SEQ ID NO: 66

Prototheca moriformis Aat05-Amino acid transporter promoter/5'UTR

CCGAGCAGTTCATGGCCAAGTACAAGGACTAGAGACCGGAGGTCGGTAGGCTGA
ATGGAGCTGGCGTTCGTGCGCGACGTGCACGCGATGCGATACTACGACCCCA
CAAACGCATGCCTCCCATCTTGATGCCTTCCGGCCATTATATACTATTTCTCATTT
CGCTGTAACATCTTGAATAATAGAATTGCCCTGTGTCAAGTGGATTCCAAGAAAT
ATTCTGTCCCAACAAAACAACCCAAACCTGAAAACAACCTCAAATACCACCAGCC
CTGCCACCTGCCCAGTACACTTTTCCAATACCATCTCCCTACCTTCACGCGCAAG
CGGCACCCATGCGCGACCAGGCTCGAAAGGATTTACGACTCAGGACGAGCGAG
TGGCGGCGCGACCGCCTGCCTGTTCTGTCACGTGCCGTACGTTCGGCGACCGCTAGA
GCTTTGCCTGGCAACCCCCGGCTTCGTCAACCCGGCCAGCCAGGATCTCGACCAC
TCTACCGCGAAATCGCCTCAAGAAGTCGCCAAAAGTGCCGTACACCATGCTTCGC
AGCGCTGTTCAAACCTTGATGCCAATCTTGACAATCAGGTTGCTCGTTGGCTGCGT
CCACATCGGCCGTGATTGCAGCAGGCGGGGATCGGACACGGAGGACGCGGCGTC
ACGCCGCGAACGCAGCCCGTAACCTCTACATCAACGCGATATGTTGCGTAATCCCG
CCCGGCTGCGCATTGTGACAACCCATTCTGCGATGGATGGTCGGAAAATGGTGTGC
CAACTGCCCTGAGGGACTCTCTCGCGAAACGGGCACGTCCCTGTATCCGAAACTG
TGGCATGGCCTTGTGCGACCACGCAAGCACGTGGACCCTAACACCACGAAAATAA
GTAAAAAAGGTTGACATCCTCTACGAGCGAATTGTTTGCTCGACCCTTCATCGCA
CACTGTCAATTATAATGCATCTAGCTCGGCGACAAGTTTAAAAAAGGCAGGCTGCA
TTATTCCATTTTGCCGTGGCGGCATGGGTGCCCATTTTATGAGGTTTGGGCTCTTG
GGCAGCGACCGAGCCAGGTTGAGTCCCTCTCGCCCGTCGACAACGTTCCAAAGC
CCATAAGTGGCTAATAAACAACCTTGATGGTACCTGTACACTGCCAGTTCTTCTT
CCCCGGCCGAGGTTTACACGTGATGGCCATGGCTTCGCGTTTCAGGCTGACTTCC
CATTCGACTTTCCAGAGGGTCCGCGGACGCCGGGGGTTGGCTGCGTGAGGCCC
ACCCCTTGTTCCCCGCGTCCCGACAAACACAATTGCGTTACATAAGGGGGAAGCC
GCCCCCGTTTCAAGTGCAAACATCTTTCATTATATTTTTCAGTCGTCAGCGAAAT
CAAGTATGTCGCTGACAGGCATGAAGGCC

SEQ ID NO: 67

Relevant portions of the reporter construct for testing putative promoter/5'UTR of *Prototheca moriformis* Aat01.

GCTCTTCGGCCGCGCCCACTCCTGCTCGAGCGCGCCCGACTCGCGCTCCGCCTGC
GCCCGCGCGTGCGCCGCCAGCGCCTTGGCCTTTTCGCCGCGCTCGTGCGCGTCGC

TGATGTCCATCACCAGGTCCATGAGGTCTGCCTTGCGCCGGCTGAGCCACTGCTT
 CGTCCGGGCGGCCAAGAGGAGCATGAGGGAGGACTCCTGGTCCAGGGTCTGAC
 GTGGTCGCGGCTCTGGGAGCGGGCCAGCATCATCTGGCTCTGCCGCACCGAGGC
 CGCCTCCAACCTGGTCTCCAGCAGCCGCAAGTCGCCGCCGACCCTGGCAGAGGAA
 GACAGGTGAGGGGGGTATGAATTGTACAGAACAACCACGAGCCTTGTCTAGGCA
 GAATCCCTACCAGTCATGGCTTTACCTGGATGACGGCCTGCGAACAGCTGTCCAG
 CGACCCTCGCTGCCGCCGCTTCTCCCGCACGCTTCTTTCCAGCACCGTGATGGCG
 CGAGCCAGCGCCGCACGCTGGCGCTGCGCTTCGCCGATCTGAGGACAGTCGGGG
 AACTCTGATCAGTCTAAACCCCTTGCGCGTTAGTGTTGCCATCCTTTGCAGACC
 GGTGAGAGCCGACTTGTGTGCGCCACCCCCACACCACCTCCTCCCAGACCAAT
 TCTGTCACCTTTTTGGCGAAGGCATCGGCCTCGGCCTGCAGAGAGGACAGCAGTG
 CCCAGCCGCTGGGGGTGGCGGATGCACGCTCAGGTACCCTTTCTTGCGCTATGA
 CACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGC
 AACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCC
 GATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCCGATTGCAAAGA
 CATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTA
 CACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCTCTT
 CGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTT
 CTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCC
 GACCGCCCCCTGGTGCACCTCACCCCCAACAGGGCTGGATGAACGACCCCAAC
 GGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAAC
 CCGAACGACACCGTCTGGGGGACGCCCTTGTCTGGGGCCACGCCACGTCCGAC
 GACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGAC
 TCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCT
 TCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACA
 CCCCAGAGTCCGAGGAGCAGTACATCTCTACAGCCTGGACGGCGGCTACACCTT
 CACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGA
 CCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGATCATGACCGCGGCCAA
 GTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAA
 GCTGGAGTCCGCGTTTCGCCAACGAGGGCTTCTTCGGCTACCAGTACGAGTGCCCC
 GGCCTGATCGAGGTCCCCACCGAGCAGGACCCCAGCAAGTCCTACTGGGTGATG
 TTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCG
 TCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGT
 GGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACC
 TACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACTGGGAGTACTCCGCCTTCG
 TGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAAC
 ACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCG
 ATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACG
 TTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACAGCACCGGCACCCCTG
 GAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGT
 TCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCG
 CATGGGCTTCGAGGTGTCCGCGTCCTCCTTCTTCTTGGACCGCGGGAACAGCAAG
 GTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAAC
 CAGCCCTTCAAGAGCGAGAACGACCTGTCTACTACAAGGTGTACGGCTTGCTGG
 ACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACA
 CCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGG
 TGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATT
 GGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCTGTGATGG
 ACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCA
 AACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGC

TTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTT
 GCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGC
 TCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGG
 TACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGAT
 GGGAAACACAAATGGAGGATCCCGCGTCTCGAACAGAGCGCGCAGAGGAACGCT
 GAAGGTCTCGCCTCTGTCGCACCTCAGCGCGGCATACACCACAATAACCACCTGA
 CGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCGGTTCACACACGTGCCA
 CGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCTGATGGTCGAAACGTTT
 ACAGCCTAGGGATATCGAATTTCATCGAATTTGTAGGAAGCGGGCATTTGGTACGG
 GAGTGGGAAAGCGAAAAAACCTCCCTCCGGCAGTGCCATCTGCCGGAGTCGAAC
 GTTGATAGGGTTCTCGTGACAGGGTGTGACCTCTCAGCCTTGCATCAATTAAACG
 CTATAGACATTATCAGTAACCGTGAATCCCGCATTGGATGCCACCCGCGCGACCA
 TTGGGGACCTGCATTACAGATCTAGGTGAGATGACAGCGAGGCAACTTCGGCCC
 GCGGCCCAGCTTGCGGCGCACCAATATTGGTCACGGGAAGCCACACACCGACCA
 TAAATGAATACTTGTAAGCTATGTCAACCGATCAATGGCGTCGAAAGTGTGCCAC
 GAGGATCCATCTGGCGGGGCGGCGTGGCGCACAAGCGCAGTCGCAATTTCTCGG
 ACCCATCTGACCTAGGCCCAGCGCCGCGGGAGAAATCCCCGGCGGGTCTCTCCAC
 GCAGTAACCTAATGAGTATCGAGCGCCGACCATTTACACCATCGCCCCCGAA
 TCCTTCCGACATTATTATTATCTTTTAGATCTTGGAAACAGACTCTGCCAACCTA
 GTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGCGACCTGCG
 TCGCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGAGGCCCTCCCCGTGCGCGG
 GCGCGCCCCGACTGGTCCATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCC
 GCGGAGAAGCAGTGGACCAACCTGGAGTGGAAAGCCCAAGCCCAACCCCCCAG
 CTGCTGGACGACCACTTCGGCCCCACGGCCTGGTGTTCGCGCGCACCTTCGCCA
 TCCGCAGCTACGAGGTGGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGA
 ACCACCTGCAGGAGGCCGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCG
 ACGGCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGA
 AGCGCACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGG
 TGGAGTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCCTGG
 TGCGCGACTGCAAGACCGGCGGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGAT
 GATGAACACCCGCACCCGCCGCCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGA
 GATCGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAA
 GCCCCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCC
 CCGCTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGA
 CTGGATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCC
 TTCACCATCGAGTACCGCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGA
 CCACCGTGAGCGGCGGCTCCTCCGAGGGCCGGCCTGGTGTGCGAGCACCTGCTGC
 AGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGC
 TGACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGA
 CTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGA
 CGACGACAAGTGACTCGAGTTAATTAACTCGAGGCAGCAGCAGCTCGGATAGTA
 TCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGC
 CTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCT
 TGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACC
 CCCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTA
 CGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACA
 GCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCA
 CTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTGT
 AGAGCTCCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCG
 ATAACCTCCAAAGCCGCTCTAATTGTGGAGGGGGTTCGAATTTAAAAGCTTGAA

TGTTGGTTCGTGCGTCTGGAACAAGCCCAGACTTGTTGCTCACTGGGAAAAGGAC
 CATCAGCTCCAAAAAACTTGCCGCTCAAACCGCGTACCTCTGCTTTCGCGCAATC
 TGCCCTGTTGAAATCGCCACCACATTCATATTGTGACGCTTGAGCAGTCTGTAAT
 TGCCCTCAGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTGC
 CGGGCGAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAG
 TTCATAACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTG
 AGTGGCCACCCCCCGGCCCTGGTGCTTGCGGAGGGCAGGTCAACCGGCATGGGG
 CTACCGAAATCCCCGACCGGATCCCACCACCCCGCGATGGGAAGAATCTCTCCC
 CGGGATGTGGGGCCACCACCAGCACAACTGCTGGCCCAGGCGAGCGTCAAACC
 ATACCACACAAATATCCTTGGCATCGGCCCTGAATTCCTTCTGCCGCTCTGCTACC
 CGGTGCTTCTGTCCGAAGCAGGGGTGCTAGGGATCGCTCCGAGTCCGCAAACCC
 TTGTCGCGTGGCGGGGCTTGTTTCGAGCTTGTTTCGAGCTTGAAGAGCCTCTAGAGT
 CGACCTGCAGGCATGCAAGCTTGCGCTAATCATGGTCATAGCTGTTTCCTGTGTG
 AAATTGTTATCCGCTCACAATTCACACAAACATACGAGCCGGAAGCATAAAGTGT
 AAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCAC
 TGCCCGCTTTCAGTCGGGAAACCTGTCTGCGCAGCTGCATTAATGAATCGGCCA
 ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 68

C. camphora thioesterase forward primer

TACCCCGCCTGGGGCGACAC

SEQ ID NO: 69

C. camphora thioesterase reverse primer

CTTGCTCAGGCGGCGGGTGC

SEQ ID NO: 70

cd189 forward primer

CCGGATCTCGGCCAGGGCTA

SEQ ID NO: 71

cd189 reverse primer

TCGATGTCGTGCACCGTCGC

SEQ ID NO: 72

5' donor DNA sequence of *Prototheca moriformis* delta 12 FAD knockout homologous recombination targeting construct

GCTCTTCGGGTTTGCTCACCCGCGAGGTCGACGCCCAGCATGGCTATCAAGACGA
 ACAGGCAGCCTGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCA
 TCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTCGTAGCAGCATGTACCTGGCCTT
 TGACATCGCGGTCATGTCCCTGCTCTACGTCGCGTCGACGTACATCGACCCTGCG
 CCGGTGCCTACGTGGGTCAAGTATGGCGTCATGTGGCCGCTCTACTGGTTCTTCC
 AGGTGTGTGTGAGGGTTGTGGTTGCCCGTATCGAGGTCCTGGTGGCGCGCATGGG
 GGAGAAGGCGCCTGTCCCGCTGACCCCCCGGCTACCCTCCCGGCACCTTCCAGG
 GCGCCTTCGGCACGGGTGTCTGGGTGTGCGCGCACGAGTGCGGCCACCAGGCCTT

TTCCTCCAGCCAGGCCATCAACGACGGCGTGGGCCTGGTGTTCACAGCCTGCTG
 CTGGTGGCCTACTACTCCTGGAAGCACTCGCACCGCCGCCACCACTCCAACACGG
 GGTGCCTGGACAAGGACGAGGTGTTTGTGCCGCCGCACCGCGCAGTGGCGCACG
 AGGGCCTGGAGTGGGAGGAGTGGCTGCCCATCCGCATGGGCAAGGTGCTGGTCA
 CCTGACCCTGGGCTGGCCGCTGTACCTCATGTTCAACGTCGCCCTCGCGGCCGTA
 CCCGCGCTTCGCCAACCACCTTTGACCCGTGGTCGCCCATCTTCAGCAAGCGCGAG
 GTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGAC
 GGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTT
 CGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGC
 CAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACA
 CCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCT
 AAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCC

SEQ ID NO: 73

3' donor DNA sequence of *Prototheca moriformis* delta 12 FAD knockout homologous recombination targeting construct

CAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTG
 ATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTT
 ATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAG
 CTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATC
 GCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTC
 CTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTC
 CTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTG
 GGATGGGAACACAAATGGAGCATCGAGGTGGTCATCTCCGACCTCGCGTTGGTG
 GCGGTGCTCAGCGGGCTCAGCGTGCTGGGCGGCACCATGGGCTGGGCTGGCTG
 GTCAAGACCTACGTGGTGCCCTACATGATCGTGAACATGTGGCTGGTGCTCATCA
 CGCTGCTCCAGCACACGCACCCGGCCCTGCCGCACTACTTCGAGAAGGACTGGG
 ACTGGCTACGCGGCGCCATGGCCACCGTCGACCGCTCCATGGGCCCCGCCCTTCAT
 GGACAGCATCCTGCACCACATCTCCGACACCCACGTGCTGCACCACCTCTTCAGC
 ACCATCCCGCACTACCACGCCGAGGAGGCCTCCGCCGCCATCCGGCCCCATCCTGG
 GCAAGTACTACCAATCCGACAGCCGCTGGGTCCGCCCGCGCCCTGTGGGAGGACT
 GGCGCGACTGCCGCTACGTCGTCCCCGACGCGCCCGAGGACGACTCCGCGCTCTG
 GTTCCACAAGTGAGCGCGCCTGCGCGAGGACGCAGAACAAACGCTGCCGCCGTGT
 CTTTTGCACGCGCGACTCCGGCGCTTCGCTGGTGGCACCCCCATAAAGAAACCCT
 CAATTCTGTTTGTGGAAGACACGGTGTACCCCCACCCACCCACCTGCACCTCTAT
 TATTGGTATTATTGACGCGGGAGTGGGCGTTGTACCCTACAACGTAGCTTCTCTA
 GTTTTCAGCTGGCTCCACCATTTGTAAGAGCCTCTAGAGTCGACCTGCAGGCAT
 GCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCT
 CACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGC
 CTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCAG
 TCGGGAAACCTGTCTGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGA
 GGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 74

Prototheca moriformis delta 12 FAD knockout homologous recombination targeting construct

GCTCTTCGGGTTTGCTCACCCGCGAGGTGACGCCCAGCATGGCTATCAAGACGA
 ACAGGCAGCCTGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCA

TCCCCGCGCACTGTTTTGAGCGCTCGGCGCTTCGTAGCAGCATGTACCTGGCCTT
 TGACATCGCGGTCATGTCCCTGCTCTACGTCGCGTCGACGTACATCGACCCTGCG
 CCGGTGCCTACGTGGGTCAAGTATGGCGTCATGTGGCCGCTCTACTGGTTCTTCC
 AGGTGTGTGTGAGGGTTGTGGTTGCCCGTATCGAGGTCCTGGTGGCGCGCATGGG
 GGAGAAGGGCGCCTGTCCCGCTGACCCCCCGGCTACCCTCCCGGCACCTTCCAGG
 GCGCCTTCGGCACGGGTGTCTGGGTGTGCGCGCACGAGTGCGGCCACCAGGCCTT
 TTCCTCCAGCCAGGCCATCAACGACGGCGTGGGCCTGGTGTTCACAGCCTGCTG
 CTGGTGCCCTACTACTCCTGGAAGCACTCGCACCCGCCGCCACCCTCCAACACGG
 GGTGCCTGGACAAGGACGAGGTGTTTGTGCCGCCGCACCGCGCAGTGGCGCACG
 AGGGCCTGGAGTGGGAGGAGTGGCTGCCCATCCGCATGGGCAAGGTGCTGGTCA
 CCCTGACCCTGGGCTGGCCGCTGTACCTCATGTTCAACGTCCGCTCGCGGCCGTA
 CCCGCGCTTCGCCAACCCTTTGACCCGCTGGTTCGCCCATCTTCAGCAAGCGCGAG
 GTACCCTTTTCTTGCCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGAC
 GGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTT
 CGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTAAATAGC
 CAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACA
 CCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCT
 AAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATG
 CTGCTGCAGGCCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCT
 CCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCACTTCACCCCCAACAAGG
 GCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGC
 ACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCGT
 GGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCAT
 CGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTAC
 AACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGG
 CCATCTGGACCTACAACACCCCGAGTCCGAGGAGCAGTACATCTCCTACAGCCT
 GGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAA
 CTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGG
 ATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGAC
 GACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTTCGCCAACGAGGGGCTTCTCGGCT
 ACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCA
 AGTCCTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTC
 CTTCAACCAGTACTTTCGTGCGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGAC
 AACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCT
 TCAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAAC
 TGGAGTACTCCGCCTTCGTGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGT
 GCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGAT
 CAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCG
 GTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTC
 CAACAGCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCA
 GACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAG
 GACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCC
 TGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCA
 ACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGACCTGTCCTACT
 ACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACG
 GCGACGTGCTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTC
 CGTGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGT
 GCGCGAGGTCAAGTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTC
 TGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTG
 AATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGC

GCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCC
 CTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGC
 TATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTG
 GGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTG
 ATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAGCATCGAGGTGGTCATCT
 CCGACCTCGCGTTGGTGGCGGTGCTCAGCGGGCTCAGCGTGCTGGGCCGCACCAT
 GGGCTGGGCCTGGCTGGTCAAGACCTACGTGGTGCCCTACATGATCGTGAACATG
 TGGCTGGTGCTCATCACGCTGCTCCAGCACACGCACCCGGCCCTGCCGCACTACT
 TCGAGAAGGACTGGGACTGGCTACGCGGCGCCATGGCCACCGTCGACCGCTCCA
 TGGGCCCCGCCCTTCATGGACAGCATCCTGCACCACATCTCCGACACCCACGTGCT
 GCACCACCTCTTCAGCACCATCCCGCACTACCACGCCGAGGAGGCCTCCGCCGCC
 ATCCGGCCCATCCTGGGCAAGTACTACCAATCCGACAGCCGCTGGGTGCGCCGC
 GCCCTGTGGGAGGACTGGCGCGACTGCCGCTACGTCGTCCCCGACGCGCCCCGAG
 GACGACTCCGCGCTCTGGTTCCACAAGTGAGCGCGCCTGCGCGAGGACGCAGAA
 CAACGCTGCCGCCGTGTCTTTTGACGCGCGACTCCGGCGCTTCGCTGGTGGCAC
 CCCCATAAAGAAACCTCAATTCTGTTTGTGGAAGACACGGTGTACCCCCACCCA
 CCCACCTGCACCTCTATTATTGGTATTATTGACGCGGGAGTGGGCGTTGTACCCT
 ACAACGTAGCTTCTCTAGTTTTGAGCTGGCTCCCACCATTGTAAAGAGCCTCTAG
 AGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
 GTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAA
 GTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGC
 TCACTGCCCCGCTTTCAGTCGGGAAACCTGTCTGTGCCAGCTGCATTAATGAATCG
 GCCAACGCGCGGGGAGAGGCGGTTTGCCTATTGGGCGCTCTTCC

SEQ ID NO: 75

5' donor DNA sequence of *Prototheca moriformis* SAD2A knockout homologous
 recombination targeting construct

GCTCTTCCGCCTGGAGCTGGTGCAGAGCATGGGTGAGTTTGCGGAGGAGAGGGT
 GCTCCCCGTGCTGCACCCCGTGGACAAGCTGTGGCAGCCGCAGGACTTCCTGCCC
 GACCCCGAGTCGCCCCGACTTCGAGGACCAGGTGGCGGAGCTGCGCGCGCGCGCC
 AAGGACCTGCCCCGACGAGTACTTTGTGGTGTGCTGGTGGGCGACATGATCACGGAG
 GAGGCGCTGCCGACCTACATGGCCATGCTCAACACCTTGGACGGTGTGCGCGAC
 GACACGGGCGCGGCTGACCACCCGTGGGCGCGCTGGACGCGGCAGTGGGTGGCC
 GAGGAGAACCGGCACGGCGACCTGCTGAACAAGTACTGTTGGCTGACGGGGCGC
 GTCAACATGCGGGCCGTGGAGGTGACCATCAACAACCTGATCAAGAGCGGCATG
 AACCCGCAGACGGACAACAACCCTTACTTGGGCTTCGTCTACACCTCCTTCCAGG
 AGCGCGCCACCAAGTAGGTACC

SEQ ID NO: 76

3' donor DNA sequence of *Prototheca moriformis* SAD2A knockout homologous
 recombination targeting construct

CAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTG
 ATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTT
 ATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAG
 CTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATC
 GCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTC
 CTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTC
 CTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTG

GGATGGGAACACAAATGGAAGGATCGTAGAGCTCCAGCCACGGCAACACCGCGC
 GCCTGGCGGCCGAGCACGGCGACAAGGGCCTGAGCAAGATCTGCGGGCTGATCG
 CCAGCGACGAGGGCCGGCACGAGATCGCCTACACGCGCATCGTGGACGAGTTCT
 TCCGCCTCGACCCCGAGGGCGCCGTCGCCGCCTACGCCAACATGATGCGCAAGC
 AGATCACCATGCCCGCGCACCTCATGGACGACATGGGCCACGGCGAGGCCAACCC
 CGGGCCGCAACCTCTTCGCCGACTTCTCCGCCGTCGCCGAGAAGATCGACGTCTA
 CGACGCCGAGGACTACTGCCGCATCCTGGAGCACCTCAACGCGCGCTGGAAGGT
 GGACGAGCGCCAGGTCAGCGGCCAGGCCGCCGCGGACCAGGAGTACGTTCTGGG
 CCTGCCCCAGCGCTTCCGGAAACTCGCCGAGAAGACCGCCGCCAAGCGCAAGCG
 CGTCGCGCGCAGGCCCCGTCGCCTTCTCCTGGAGAGAAGAGCCTCTAGAGTCGACC
 TGCAGGCATGCAAGCTTGCGCTAATCATGGTCATAGCTGTTTCTGTGTGAAATT
 GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAG
 CCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCC
 CGCTTTCCAGTCGGGAAACCTGTCTGTGCCAGCTGCATTAATGAATCGGCCAACGC
 GCGGGGAGAGGCGGTTTTCGTATTGGGCGCTCTTCC

SEQ ID NO: 77

Prototheca moriformis SAD2A knockout homologous recombination targeting construct

GCTCTTCCGCCTGGAGCTGGTGCAGAGCATGGGTTCAGTTTGCGGAGGAGAGGGT
 GCTCCCCGTGCTGCACCCCGTGGACAAGCTGTGGCAGCCGCAGGACTTCCTGCCC
 GACCCCGAGTCGCCCGACTTCGAGGACCAGGTGGCGGAGCTGCGCGCGCGCGCC
 AAGGACCTGCCCAGCAGTACTTTGTGGTGTGGTGGGCGACATGATCACGGAG
 GAGGCGCTGCCGACCTACATGGCCATGCTCAACACCTTGGACGGTGTGCGCGAC
 GACACGGGCGCGGCTGACCACCCGTGGGCGCGCTGGACGCGGCAGTGGGTGGCC
 GAGGAGAACCGGCACGGCGACCTGCTGAACAAGTACTGTTGGCTGACGGGGCGC
 GTCAACATGCGGGCCGTGGAGGTGACCATCAACAACCTGATCAAGAGCGGCATG
 AACCCGACAGCGGACAACAACCTTACTTGGGCTTCGTCTACACCTCCTTCCAGG
 AGCGCGCCACCAAGTAGGTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGT
 AGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTC
 GACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCG
 AGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACC
 AAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCT
 TGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCC
 GCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGC
 CGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCA
 CTTACCCCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGA
 GAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTG
 GGGGACGCCCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGA
 GGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGC
 TCCATGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACC
 CGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGC
 AGTACATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGA
 ACCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTA
 CGAGCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGAT
 CGAGATCTACTCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTCGCC
 AACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCA
 CCGAGCAGGACCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAACCCCGG
 CGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTTCGGCAGCTTCAACGGCACC
 CACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACT

ACGCCCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCAT
 CGCGTGGGCCTCCAACTGGGAGTACTCCGCCTTCGTGCCCACCAACCCCTGGCGC
 TCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCC
 CGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACG
 CCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCT
 ACAACGTCGACCTGTCCAACAGCACCGGCACCCTGGAGTTTCGAGCTGGTGTACG
 CCGTCAACACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTG
 GTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTC
 CGCGTCCCTCCTTCTTCCTGGACCGCGGGAACAGCAAGGTGAAGTTTCGTGAAGGA
 GAACCCCTACTTCACCAACCGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGA
 GAACGACCTGTCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGA
 GCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACC
 GGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTAC
 ATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATTGGCAGCAGCAGCTCGG
 ATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACT
 TGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTT
 TGATCTTGTTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTTGCTATTTGCGAAT
 ACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACT
 TATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCT
 CGCACAGCCTTGTTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAA
 CCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAA
 GGATCGTAGAGCTCCAGCCACGGCAACACCGCGCGCCTGGCGGCCGAGCACGGC
 GACAAGGGCCTGAGCAAGATCTGCGGGCTGATCGCCAGCGACGAGGGCCGGCAC
 GAGATCGCCTACACGCGCATCGTGGACGAGTTCTTCCGCCTCGACCCCGAGGGCG
 CCGTCGCCGCCTACGCCAACATGATGCGCAAGCAGATCACCATGCCCGCGCACCT
 CATGGACGACATGGGCCACGGCGAGGCCAACCCGGGCGCAACCTCTTCGCCGA
 CTTCTCCGCCGTGCGCGAGAAGATCGACGTCTACGACGCCGAGGACTACTGCCCGC
 ATCCTGGAGCACCTCAACGCGCGCTGGAAGGTGGACGAGCGCCAGGTCAGCGGC
 CAGGCCCGCGCGGACCAGGAGTACGTTCTGGGCCTGCCCCAGCGCTTCCGAAA
 CTCGCCGAGAAGACCGCCGCCAAGCGCAAGCGCGTCGCGCGCAGGCCCGTCGCC
 TTCTCCTGGAGAGAAGAGCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGT
 AATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCACAC
 AACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGC
 TAACTCACATTAAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGT
 CGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGTA
 TTGGGCGCTCTTCC

SEQ ID NO: 78

5' donor DNA sequence of *Prototheca moriformis* SAD2B knockout homologous recombination targeting construct

GCTCTTCCCGCCTGGAGCTGGTGCAGAGCATGGGGCAGTTTTCGGAGGAGAGGG
 TGCTCCCCGTGCTGCACCCCGTGGACAAGCTGTGGCAGCCGAGGACTTCTGCC
 CGACCCCGAGTCGCCCCGACTTCGAGGACCAGGTGGCGGAGCTGCGCGCGCGCGC
 CAAGGACCTGCCCCGACGAGTACTTTGTGGTGTGTTGGGCGACATGATCACGGA
 GGAGGCGCTGCCGACCTACATGGCCATGCTCAACACCTTGGACGGTGTGCGCGA
 CGACACGGGCGCGGCTGACCACCCGTGGGCGCGCTGGACGCGGCAGTGGGTGGC
 CGAGGAGAACCAGGACGGCGACCTGCTGAACAAGTACTGTTGGCTGACGGGGCG
 CGTCAACATGCGGGCCGTGGAGGTGACCATCAACAACCTGATCAAGAGCGGCAT

GAACCCGCAGACGGACAACAACCCTTACTTGGGCTTCGTCTACACCTCCTTCCAG
GAGCGCGCCACCAAGTAGGTACC

SEQ ID NO: 79

3' donor DNA sequence of *Prototheca moriformis* SAD2B knockout homologous recombination targeting construct

CAGCCACGGCAACACCGCGCGCCTTGCGGCCGAGCACGGCGACAAGAACCTGAG
CAAGATCTGCGGGCTGATCGCCAGCGACGAGGGCCGGCACGAGATCGCCTACAC
GCGCATCGTGGACGAGTTCTTCCGCCTCGACCCCGAGGGCGCCGTCGCCGCCTAC
GCCAACATGATGCGCAAGCAGATCACCATGCCCCGCGCACCTCATGGACGACATG
GGCCACGGCGAGGCCAACCCGGGCCGCAACCTCTTCGCCGACTTCTCCGCGGTTCG
CCGAGAAGATCGACGTCTACGACGCCGAGGACTACTGCCGCATCCTGGAGCACC
TCAACGCGCGCTGGAAGGTGGACGAGCGCCAGGTCAGCGGCCAGGCCGCCGCGG
ACCAGGAGTACGTCTTGGGCCTGCCCCAGCGCTTCCGGAAACTCGCCGAGAAGA
CCGCCGCCAAGCGCAAGCGCGTCGCGCGCAGGCCCGTCGCCCTTCTCCTGGAGAA
GAGCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGCGCTAATCATGGTCATAGC
TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGG
AAGCATAAAGTGTAAGCCTGGGGTGCTAATGAGTGAGCTAACTCACATTAAT
TGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCAT
TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 80

Prototheca moriformis SAD2B knockout homologous recombination targeting construct

GCTCTTCCCGCCTGGAGCTGGTGCAGAGCATGGGGCAGTTTGCGGAGGAGAGGG
TGCTCCCCGTGCTGCACCCCGTGGACAAGCTGTGGCAGCCGCAGGACTTCCTGCC
CGACCCCGAGTCGCCCGACTTCGAGGACCAGGTGGCGGAGCTGCGCGCGCGCGC
CAAGGACCTGCCCCGACGAGTACTTTGTGGTGCTGGTGGGCGACATGATCACGGA
GGAGGCGCTGCCGACCTACATGGCCATGCTCAACACCTTGGACGGTGTGCGCGA
CGACACGGGCGCGGCTGACCACCCGTGGGCGCGCTGGACGCGGCAGTGGGTGGC
CGAGGAGAACC GGACGCGACCTGCTGAACAAGTACTGTTGGCTGACGGGGCG
CGTCAACATGCGGGCCGTGGAGGTGACCATCAACAACCTGATCAAGAGCGGCAT
GAACCCGCAGACGGACAACAACCCTTACTTGGGCTTCGTCTACACCTCCTTCCAG
GAGCGCGCCACCAAGTAGGTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAGG
TAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTT
CGACCCCCCGAAGCTTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGC
GAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTAC
CAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGC
TTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCC
GCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGC
CGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCA
CTTCACCCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGA
GAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTG
GGGACGCCCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGA
GGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGC
TCCATGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACC
CGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGC
AGTACATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGA
ACCCCGTGCTGGCCGCCAACTCCACCCAGTTCGCGACCCGAAGGTCTTCTGGTA

CGAGCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGAT
CGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTTCGCC
AACGAGGGCTTCCTCGGCTACCAGTACGAGTCCCCCGGCCTGATCGAGGTCCCCA
CCGAGCAGGACCCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAACCCCGG
CGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTCGGCAGCTTCAACGGCACC
CACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACT
ACGCCCTGCAGACCTTCTTCAACACCCGACCCGACCTACGGGAGCGCCCTGGGCAT
CGCGTGGGCCTCCAACTGGGAGTACTCCGCCTTCGTGCCACCAACCCCTGGCGC
TCCTCCATGTCCCTCGTGCGCAAGTTCCTCCCTCAACACCCGAGTACCAGGCCAACCC
CGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACG
CCGGCCCCCTGGAGCCGGTTCGCCACCAACACCAACGTTGACGAAGGCCAACAGCT
ACAACGTCGACCTGTCCAACAGCACCGGCACCCCTGGAGTTCGAGCTGGTGTACG
CCGTCAACACCAACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTG
GTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTC
CGCGTCTCTCTTCTTCTGACCGCGGGGAACAGCAAGGTGAAGTTCGTGAAGGA
GAACCCCTACTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGA
GAACGACCTGTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGA
GCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACC
GGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTAC
ATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATTGGCAGCAGCAGCTCGG
ATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACT
TGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTT
TGATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAAT
ACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACT
TATCTACGCTGTCTGTCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCT
CGCACAGCCTTGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAA
CCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAC
AGCCACGGCAACACCGCGCGCCTTGCGGCCGAGCACGGCGACAAGAACCTGAGC
AAGATCTGCGGGCTGATCGCCAGCGACGAGGGCCGGCACGAGATCGCCTACACG
CGCATCGTGGACGAGTTCTTCCGCCTCGACCCCGAGGGCGCCGTCCGCCCTACG
CCAACATGATGCGCAAGCAGATCACCATGCCCGCGCACCTCATGGACGACATGG
GCCACGGCGAGGCCAACCCGGGCGCAACCTCTTCGCCGACTTCTCCGCGGTCCG
CGAGAAGATCGACGTCTACGACGCCGAGGACTACTGCCGCATCCTGGAGCACCT
CAACGCGCGCTGGAAGGTGGACGAGCGCCAGGTGAGCGGCCAGGCCGCCGCGG
ACCAGGAGTACGTCTTGGGCCTGCCCCAGCGCTTCCGGAACTCGCCGAGAAGA
CCGCCGCCAAGCGCAAGCGCGTCGCGCGCAGGCCCGTCGCCTTCTCCTGGAGAA
GAGCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGC
TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGG
AAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAAT
TGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCAT
TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 81

P.moriformis Aat02 promoter/5'UTR::*C.protothecoides* stearoyl ACP desaturase

PTS::*Cinnamomum camphora* C14 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

AGAGAGCGGAGGTGGGGTTGTGAGGTGGGGTTGCTGACCAGGAGCTCGCGTCGC
CGAGCGGCACTCGCACACGGTCCAGTTACCCCCCCTCCGCCCAAACGCAAGCCT
CCCATCTTGATGCCTTTCCGGCCACCTATACTATTTCTTAGTTCCGTGTAACATCC
AGACCGTCCTGAATAATAACAATGCCCTGTGTCAAGTGCATTCTAAAAAATTC

TGTCCCAACCAACAATCCCACCTGAAATACCACCAGCCCTGCCAGTACACTCTT
CCAATACCATCTCCCTACCTCCACGCGCAAGCGACCCCATGCGCGACCAGGCTC
GAAAGTGATTTATGACTTGAGACGAGCGAGTGGCGGCGCGGTGCGACTGCCTTTTC
ATCACGTGCCGTACGTCGGCGACCGCTAGGGCTTTGCACGGCAACGCACGGCTTC
GCCAACCCGACCAGCCAGGACCTCGACTACTCTACCGCGAATTCGCCTCAAGAA
GTCGCCAAATGTGCCATACACCATTCCTTACAGCACTGTTCAAACITGATGCCAA
TTTTGACATTCGGGTTGCTCGTTGGCTGCGCCACATCGGCCGTGAGTGCAGCAG
GCGGGATCGGACACGGAGGACGCGGGCTCACGCCCCGAACGCAGCCCGTAACTC
TACATCAACACGACGTGTTGCGTAATCCCGCCCGGCTGCGCATCGTGCCAAACCA
TTCGCGATGGATGGTTCGGAAAATGGTGTGCCAACTGCCCTGAGGGAGGCTCTCG
CGAAACGGGCACGTCCCTGAAACCGAAACTGTGGCCTTGTCGTGCGCCACGCAA
GCACGTGGACCCTAAACACCAAGAAAATCAGTAAACAAGGTTGACATCCTCTAC
GGGCGAATTGTTTGCCCAACCCTTCATCGCACACTGCCATTATAATGCATCTAGC
TCGGCGACAAGTTTAGAAAAGGCAGGCTGCATTGTTCCATTTGCGCGTGCGCGGC
TGGGTGCCCATTTTACGAGGTTTGGGCTCCCGGGCAGCGACCGAGCCAGGTGCA
GTCCCTCTCGCCCGTCGACAATGTTGCGAACCCCAACAAGCGGCTAACAACAACCTT
GATGGTACCTGTACACTGCCAATTCCTTCTTCCCCGGCCGAGGTTTACACGTGAT
GGCCATGGCTTCGCATTCAGGCCGACTTCCCATTCGACTTTCCAGAGGGTCCGC
GGACGCTGGGGGTTGGCTGCCTGAGGCCACCCTTTGTTCCCCGCGTCCCGACAA
ACACAATTGCGTTACATAAGGGGGAGCCGCCCGCTTCAGAGTGCAGAAATCTTT
CACTATATTTTCCAGTCGTCAGCGAAATCAAGTACTAGTATGGCCACCGCATCCA
CTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGG
GCCCCGGCGCCAGCGAGGCCCTCCCGTGC GCGGGCGCGCCCCCGACTGGTC
CATGCTGTTGCGCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGACC
AACCTGGAGTGGAAGCCCAAGCCCAACCCCCCAGCTGCTGGACGACCACTTC
GGCCCCACGGCCTGGTGTTCGCGCGCACCTTCGCCATCCGCAGCTACGAGGTGG
GCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCCG
CCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCAACCT
GGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCCGT
GGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTGGAGTGCTGGGTGGGCGC
CTCCGGCAACAACGGCCGCCGCCACGACTTCCTGGTGC GCGACTGCAAGACCGG
CGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCCG
CGCCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCGCCTTCATC
GACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGA
CTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGA
CATCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACCGT
GCCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTACCATCGAGTACCGC
CGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCGGC
TCCTCCGAGGCCGCGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGC
GAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCCGC
GGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCACGAC
GGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGACTC
GAGTTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCTTCC
CTCGTTTCATATCGCTTGCAATCCCAACCGCAACTTATCTACGCTGTCTGTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTT

REPLACEMENT SHEET

SEQ ID NO: 82

5' 6S region genomic donor sequence

GCTCTTCGCGCCGCCACTCCTGCTCGAGCGCGCCCGCGCGTGCGCCGCCAGCGC
CTTGCCCTTTTCGCGCGCTCGTGCGCGTCTGATGTCCATCACCAGGTCCATG
AGGTCTGCCTTTCGCGCGGCTGAGCCACTGCTTCGTCCGGGCGGCCAAGAGGAGC
ATGAGGGAGGACTCCTGGTCCAGGGTCTGACGTGGTCGCGGCTCTGGGAGCGG
GCCAGCATCATCTGGCTCTGCCGCACCGAGGCGCCTCCAACCTGGTCTCCAGCA
GCCGCAGTCGCGCCGACCCCTGGCAGAGGAAGACAGGTGAGGGGGGTATGAATT
GTACAGAACAACCACGAGCCTTGTCTAGGCAGAATCCCTACCAGTCATGGCTTTA
CCTGGATGACGGCCTGCGAACAGCTGTCCAGCGACCCCTCGCTGCCGCCGCTTCTC
CCGCACGCTTCTTTCCAGCACCGTGATGGCGCGAGCCAGCGCCGCACGCTGGCGC
TGCGCTTCGCCGATCTGAGGACAGTCGGGGAACTCTGATCAGTCTAAACCCCTT
GCGCGTTAGTGTTGCCATCCTTTGCAGACCGGTGAGAGCCGACTTGTGTGCGCC
ACCCCCACACCACCTCCTCCCAGACCAATTCTGTACCTTTTGGCGAAGGCAT
CGGCCTCGGCCTGCAGAGAGGACAGCAGTGCCAGCCGCTGGGGGTGGCGGAT
GCACGCTCAGGTACC

SEQ ID NO: 83

Relevant expression construct for *Cinnamomum camphora* thioesterase

(ftub::neo::nitred::ftub::C.camphora TE::nitred)

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCTCTTCTCTTCGTTTCAGTCACAACCCGCAAACTCTAGAATATCAATGATC
GAGCAGGACGGCCTCCACGCCGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTT
GGCTACGACTGGGCCCAGCAGACCATCGGCTGCTCCGACGCCGCGCTGTTCCGCC
TGTCGCCCCAGGGCCGCCCGTGTGTTCTGTAAGACCGACCTGTCCGGCGCCCT
GAACGAGCTGCAGGACGAGGCCGCCCGCTGTCTGGCTGGCCACCACCGGCGT
GCCCTGCGCCGCGTGTGAGCTGGTGACCGAGGCCGGCCGCGACTGGCTGCT
GCTGGGCGAGGTGCCCGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAG
AAGGTGTCCATCATGGCCGACGCCATGCGCCGCTGCACACCCTGGACCCCGCCA
CCTGCCCTTCGACCACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCAT
GGAGGCCGGCCTGTTGGACAGGACGACCTGGACGAGGAGCACCAGGGCCTGG
CCCCCGCGAGCTGTTGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCTGCTGCCCAACATCATGGTGGAGAACGGCC
GCTTCTCCGGCTTCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGA
CATCGCCCTGGCCACCCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGGCCGA
CCGCTTCTGTTGCTGTACGGCATCGCCGCCCCGACTCCAGCGCATCGCCTTC
TACCGCCTGCTGGACGAGTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTAT
CGACACACTCTGGACGCTGGTCTGTGATGGACTGTTGCCGCCACACTTGCTGCC
TTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTT
GTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCC
CCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTAC
GCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAG
CCTTGGTTTGGGCTCCGCCTGTATTCTCTGTTACTGCAACCTGTAAACCAGCACT

REPLACEMENT SHEET

GCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCCG
GTCTCGAACAGAGCGCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTCCGACCTC
AGCGCGGCATACACCACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCAT
TAGCGAAGCGTCCGGTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAA
TGATCGGTGGAGCTGATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTT
TCTTGGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCC
GGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTG
CATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCC
GATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCA
CTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCG
CCTCTTCTCTTCGTTTCAGTCACAACCCGCAAACACTAGTATGGCCACCGCATCC
ACTTTCTCGGCGTTCATGCCCCGCTGCGGCGACCTGCGTCTGCTCGGCGGGCTCCG
GGCCCCGGCGCCAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCCCCGACTGGT
CCATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGAC
CAACCTGGAGTGGAAAGCCCAAGCCCAACCCCCCAGCTGCTGGACGACCACTT
CGGCCCCCAGGGCTGGTGTTCGCCCGCACCTTCGCCATCCGCAGCTACGAGGTG
GGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCC
GCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCACT
CTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCC
GTGGAGCGCTACCCCGCTGGGGCGACACCGTGGAGGTGGAGTGTCTGGGTGGGC
GCCTCCGGCAACAACGGCCGCCGCCACGACTTCCTGGTGCAGCTGCAAGACC
GGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCC
GCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCCGCCTTCA
TCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAAC
GATCCACCCGCGACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTG
GACATCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACC
GTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTACCATCGAGTACC
GCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCG
GCTCCTCCGAGGCCGCGCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCA
GCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCC
GCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCACG
ACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGA
CTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGT
GATGGACTGTTGCCGCCACACTTGTGCTTACCTGTGAATATCCCTGCCGCTTT
TATCAAACAGCCTCAGTGTGTTGATCTTGTGTGTACGCGCTTTTGGCAGTTGCTA
GCTGCTTGTGCTATTTGCGAATACCACCCCGAGCATCCCCTTCCCTCGTTTCATAT
CGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCT
CCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTCCGCTGTATTCT
CCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGT
GGGATGGGAACACAAATGGAAAGCTT

SEQ ID NO: 84

3'6S region genomic donor sequence

GAGCTCCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGA
TAACCTCCAAAGCCGCTCTAATTGTGGAGGGGGTTTGAATTTAAAAGCTTGGAAAT
GTTGGTTTCGTGCGTCTGGAACAAGCCCAGACTTGTGCTCACTGGGAAAAGGACC
ATCAGCTCCAAAAAAGTTGCCGCTCAAACCGCGTACCTCTGCTTTTCGCGCAATCT
GCCCTGTTGAAATCGCCACCACATTCATATTGTGACGCTTGAGCAGTCTGTAATT
GCCTCAGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTGCG

GGGCGAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAGTT
CATAACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAG
TGGCCACCCCCCGGCCCTGGTGCTTGCGGAGGGCAGGTCAACCGGCATGGGGCT
ACCGAAATCCCCGACCGGATCCCACCACCCCGCGATGGGAAGAATCTCTCCCC
GGGATGTGGGCCCACCACCAGCACAAACCTGCTGGCCCAGGCGAGCGTCAAACCA
TACCACACAAATATCCTTGGCATCGGCCCTGAATTCTTCTGCCGCTCTGCTACCC
GGTGCTTCTGTCCGAAGCAGGGGTTGCTAGGGATCGCTCCGAGTCCGCAAACCTT
TGTCGCGTGGCGGGGCTTGTTTCGAGCTTGAAGAGC

SEQ ID NO: 85

Amt03 forward primer

5'-GGAGGAATTCGGCCGACAGGACGCGCGTCA-3'

SEQ ID NO: 86

Amt03 reverse primer

5'-GGAGACTAGTGGCTGCGACCGGCCTGTG-3'

SEQ ID NO: 87

Amt02 forward primer

5'-GGAGGAATTCTCACCAGCGGACAAAGCACCG-3'

SEQ ID NO: 88

Amt02 reverse primer

5'-GGAGACTAGTGGCTGCGACCGGCCTCTGG-3'

SEQ ID NO: 89

amt03 promoter/UTR sequence

GGCCGACAGGACGCGCGTCAAAGGTGCTGGTCGTGTATGCCCTGGCCGGCAGGT
CGTTGCTGCTGCTGGTTAGTGATTCCGCAACCCTGATTTTGGCGTCTTATTTTGGC
GTGGCAAACGCTGGCGCCCCGCGAGCCGGGCGGCGGCGATGCGGTGCCCCACGG
CTGCCGGAATCCAAGGGAGGCAAGAGCGCCCGGGTCAGTTGAAGGGCTTTACGC
GCAAGGTACAGCCGCTCCTGCAAGGCTGCGTGGTGGAATTGGACGTGCAGGTCC
TGCTGAAGTTCTCCACCGCCTCACCAGCGGACAAAGCACCGGTGTATCAGGTCC
GTGTCATCCACTCTAAAGAGCTCGACTACGACCTACTGATGGCCCTAGATTCTTC
ATCAAAAACGCCTGAGACACTTGCCAGGATTGAAACTCCCTGAAGGGACCACC
AGGGGCCCTGAGTTGTTCTTCCCCCGTGGCGAGCTGCCAGCCAGGCTGTACCT
GTGATCGAGGCTGGCGGGAAAATAGGCTTCGTGTGCTCAGGTCATGGGAGGTGC
AGGACAGCTCATGAAACGCCAACAATCGCACAAATTCATGTCAAGCTAATCAGCT
ATTTCTCTTCACGAGCTGTAATTGTCCCAAAATTCTGGTCTACCGGGGGTGATCC
TTCGTGTACGGGCCCTTCCCTCAACCCTAGGTATGCGCGCATGCGGTGCGCCGCGC
AACTCGCGCGAGGGCCGAGGGTTTGGGACGGGCCGTCCCGAAATGCAGTTGCAC
CCGGATGCGTGGCACCTTTTTTGGCGATAATTTATGCAATGGACTGCTCTGCAAAA
TTCTGGCTCTGTGCGCAACCCTAGGATCAGCGGCGTAGGATTTCTGAATCATTCG
TCCTGATGGGGAGCTACCGACTACCCTAATATCAGCCCGACTGCCTGACGCCAGC
GTCCACTTTTGTGCACACATTCCATTTCGTGCCCAAGACATTTTCATTGTGGTGCGAA

GCGTCCCCAGTTACGCTCACCTGTTTCCCGACCTCCTTACTGTTCTGTGCGACAGAG
CGGGCCCACAGGCCGGTCGCAGCC

SEQ ID NO: 90

relevant sequence of the amt promoter/UTR::*C. camphora* thioesterase expression construct

GCTCTTCGGCCGCCGCGCCACTCCTGCTCGAGCGCGCCCGACTCGCGCTCCGCCTGC
GCCCCGCGCTGCGCCGCCAGCGCCTTGGCCCTTTTCGCCGCGCTCGTGCGCGTGC
TGATGTCCATCACCAGGTCCATGAGGTCTGCCTTGCGCCGGCTGAGCCACTGCTT
CGTCCGGGCGGCCAAGAGGAGCATGAGGGAGGACTCCTGGTCCAGGGTCTTGAC
GTGGTTCGCGGCTCTGGGAGCGGGCCAGCATCATCTGGCTCTGCCGCACCGAGGC
CGCCTCCAACCTGGTCTCCAGCAGCCGCAAGTCGCCGCCGACCCTGGCAGAGGAA
GACAGGTGAGGGGGGTATGAATTGTACAGAACAACCACGAGCCTTGTCTAGGCA
GAATCCCTACCAGTCATGGCTTTACCTGGATGACGGCCTGCGAACAGCTGTCCAG
CGACCCTCGCTGCCGCCGCTTCTCCCGCACGCTTCTTTCCAGCACCGTGATGGCG
CGAGCCAGCGCCGCACGCTGGCGCTGCGCTTCGCCGATCTGAGGACAGTCGGGG
AACTCTGATCAGTCTAAACCCCTTGC GCGTTAGTGTTGCCATCCTTTGCAGACC
GGTGAGAGCCGACTTGTGTGCGCCACCCCCACACCACCTCCTCCCAGACCAAT
TCTGTCACCTTTTTTGGCGAAGGCATCGGCCTCGGCCTGCAGAGAGGACAGCAGTG
CCCAGCCGCTGGGGGTTGGCGGATGCACGCTCAGGTACCCTTTCTTGCCTATGA
CACTTCCAGCAAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGCGCGTGCATGC
AACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCC
GATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGA
CATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTA
CACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTT
CGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTT
CTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCC
GACCGCCCCCTGGTGCACCTTACCCCCAACAAAGGGCTGGATGAACGACCCCAAC
GGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAAC
CCGAACGACACCGTCTGGGGGACGCCCTTGTTCCTGGGGCCACGCCACGTCCGAC
GACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGAC
TCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCT
TCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACA
CCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTT
CACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGA
CCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGATCATGACCGCGGCCAA
GTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAA
GCTGGAGTCCGCGTTCGCCAACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCC
GGCCTGATCGAGGTCCCCACCGAGCAGGACCCCAGCAAGTCCTACTGGGTGATG
TTCATCTCCATCAACCCCGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCG
TCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGT
GGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACC
TACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCTTCG
TGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAAC
ACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCG
ATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACG
TTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACAGCACCGGCACCCTG
GAGTTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGT
TCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCG
CATGGGCTTCGAGGTGTCCGCGTCCCTCCTTCTTCTTGGACCGCGGGAACAGCAAG

GTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAAC
CAGCCCTTCAAGAGCGAGAACGACCTGTCTACTACAAGGTGTACGGCTTGCTGG
ACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTTCGTGTCCACCAACA
CCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGG
TGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATT
GGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGG
ACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCA
AACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGC
TTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTT
GCATCCCAACCGCAACTTATCTACGCTGTCTGTATCCCTCAGCGCTGCTCCTGC
TCCTGCTCACTGCCCCCTCGCACAGCCTTGCTTTGGGCTCCGCCTGTATTCTCCTGG
TACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGAT
GGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGCGCGCAGAGGAACGCT
GAAGGTCTCGCCTCTGTGCGCACCTCAGCGCGGCATACACCACAATAACCACCTGA
CGAATGCGCTTGCTTCTTCGTCCATTAGCGAAGCGTCCGGTTCACACACGTGCCA
CGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCTGATGGTCGAAACGTTT
ACAGCCTAGGGATATCGAATTGCGCCGACAGGACGCGCGTCAAAGGTGCTGGTC
GTGTATGCCCTGGCCGGCAGGTGCTTGTGCTGCTGGTTAGTGATTCCGCAACCC
TGATTTTGGCGTCTTATTTTGGCGTGGCAAACGCTGGCGCCCGCGAGCCGGGCCG
GCGGCGATGCGGTGCCCCACGGCTGCCGGAATCCAAGGGAGGCAAGAGCGCCCCG
GGTCAGTTGAAGGGCTTTACGCGCAAGGTACAGCCGCTCCTGCAAGGCTGCGTG
GTGGAATTGGACGTGCAGGTCTGTGTAAGTTCCTCCACCGCCTCACCAGCGGAC
AAAGCACCGGTGTATCAGGTCCGTGTTCATCCACTCTAAAGAGCTCGACTACGACC
TACTGATGGCCCTAGATTCTTCATCAAAAACGCCTGAGACACTTGCCAGGATTG
AAACTCCCTGAAGGGACCACCAGGGGCCCTGAGTTGTTCTTCCCCCGTGGCGA
GCTGCCAGCCAGGCTGTACCTGTGATCGAGGCTGGCGGGAAAATAGGCTTCGTG
TGCTCAGGTTCATGGGAGGTGCAGGACAGCTCATGAAACGCCAACAAATCGCACA
TTCATGTCAAGCTAATCAGCTATTTCTCTTCACGAGCTGTAATTGTCCAAAATT
CTGGTCTACCGGGGGTGATCCTTCGTGTACGGGCCCTTCCCTCAACCCTAGGTAT
GCGCGCATGCGGTGCGCGCGCAACTCGCGCGAGGGCCGAGGGTTTGGGACGGGC
CGTCCCGAAATGCAGTTGCACCCGGATGCGTGGCACCTTTTTTGGGATAATTTAT
GCAATGGACTGCTCTGCAAAAATTCTGGCTCTGTGCGCAACCCTAGGATCAGCGGC
GTAGGATTTTCGTAATCATTCGTCTGATGGGGAGCTACCGACTACCCTAATATCA
GCCCCACTGCCTGACGCCAGCGTCCACTTTTGTGCACACATTCCATTCTGTGCCA
AGACATTTTATTGTGGTGCGAAGCGTCCCCAGTTACGCTCACCTGTTTCCCGACC
TCCTTACTGTTCTGTGACAGAGCGGGCCACAGGCCGGTCGCAGCCACTAGTAT
GGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCGC
TCGGCGGGCTCCGGGCCCGGCGCCAGCGAGGCCCTCCCCGTGCGCGGGCGC
GCCCCGACTGGTCCATGCTGTTGCGCGTGATCACCACCATCTTCTCCGCCGCCG
AGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCCAGCTGC
TGGACGACCATTTCGGCCCCCACGGCCTGGTGTTCCGCGCACCTTCGCCATCCG
CAGCTACGAGGTGGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCA
CCTGCAGGAGGCCGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGG
CTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCG
CACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTGGA
GTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCTTGGTGCG
CGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGAT
GAACACCCGCACCCGCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGAT
CGGCCCCGCTTCATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCC
CCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCG

CTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGACTG
GATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTC
ACCATCGAGTACCGCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGACC
ACCGTGAGCGGCGGCTCCTCCGAGGCGGCGCTGGTGTGCGAGCACCTGCTGCAG
CTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTG
ACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACT
ACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGAC
GACGACAAGTGACTCGAGTTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTAT
CGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCC
TTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTT
GTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCC
CCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCTATCCCAACCGCAACTTATCTAC
GCTGTCCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAG
CCTTGGTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACT
GCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTGTAG
AGCTCCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGAT
AACCTCCAAAGCCGCTCTAATTGTGGAGGGGGTTTGAATTTAAAAGCTTGGAATG
TTGGTTTCGTGCGTCTGGAACAAGCCCAGACTTGTGTGCTCACTGGGAAAAGGACCA
TCAGCTCCAAAAAACTTGCCGCTCAAACCGCGTACCTCTGCTTTTCGCGCAATCTG
CCCTGTTGAAATCGCCACCACATTCATATTGTGACGCTTGAGCAGTCTGTAATTG
CCTCAGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCGCG
GGCGAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAGTTC
ATAACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAGT
GGCCACCCCCCGGCCCTGGTGCTTGCGGAGGGCAGGTCAACCGGCATGGGGCTA
CCGAAATCCCCGACCGGATCCACACCCCCCGCGATGGGAAGAATCTCTCCCCG
GGATGTGGGCCCACCACCAGCACAACCTGCTGGCCCAGGCGAGCGTCAAACCAT
ACCACACAAATATCCTTGGCATCGGCCCTGAATTCCTTCTGCCGCTCTGCTACCC
GGTGCTTCTGTCCGAAGCAGGGGTTGCTAGGGATCGCTCCGAGTCCGCAAACCTT
TGTCGCGTGCGGGGCTTGTTCGAGCTTGTTCGAGCTTGAAGAGCCTCTAGAGTC
GACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCTGTGTGA
AATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA
AAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACT
GCCCCTTTCAGTCGGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAA
CGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 91

Codon-optimized *C. camphora* thioesterase sequence

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTC
GCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGAGGCCCCTCCCCGTGCGCGGGC
GCGCCCCCGACTGGTCCATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCCGC
CGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCAGCT
GCTGGACGACCCTTCGGCCCCCACGGCCTGGTGTTCGCCCGCACCTTCGCCATC
CGCAGCTACGAGGTGGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAAC
CACCTGCAGGAGGCGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGAC
GGCTTCGGCACCAACCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAG
CGCACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTG
GAGTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCTGGTG
CGCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATG
ATGAACACCCGCACCCGCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAG

ATCGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAG
CCCCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCC
CGCTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGAC
TGGATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCT
TCACCATCGAGTACCGCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGAC
CACCGTGAGCGGCGGCTCCTCCGAGGCGCGCCTGGTGTGCGAGCACCTGCTGCA
GCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCT
GACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGAC
TACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGAC
GACGACAAGTGA

SEQ ID NO: 92

Codon-optimized *U. californica* thioesterase sequence

GGCGCGCCCCCGACTGGTCCATGCTGTTTCGCCGTGATCACCACCATCTTCAGCGC
CGCCGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCA
GCTGCTGGACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCCCGCACCTTCGCC
ATCCGCTCCTACGAGGTGGGCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGA
ACCACATGCAGGAGGCCACCCTGAACCACGCCAAGAGCGTGGGCATCCTGGGCG
ACGGCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATGTGGGTGGTGC
GCCGCACCACCGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTGGAGG
TGGAGTGCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGG
TGCGCGACTGCAAGACCGGCGGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCT
GATGAACACCCGCACCCGCGCCTGAGCACCATCCCCGACGAGGTGCGCGGCGA
GATCGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGACGAGATCAAGAA
GCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCC
CCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGC
CTGGGTGTTTCGAGACCGTGGCCGACAGCATCTTCGAGTCCCACCACATCAGCTCC
TTCACCCTGGAGTACCGCCGCGAGTGCACCCGCGACTCCGTGCTGCGCAGCCTGA
CCACCGTGAGCGGCGGCAGCTCCGAGGCGCGCCTGGTGTGCGACCACCTGCTGC
AGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCGCACCGAGTGGCGCCCCAAGC
TGACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGCGTGATGGA
CTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGA
CGACGACAAGTGAAGTCTGAG

SEQ ID NO: 93

Codon-optimized *U. americana* thioesterase sequence

GGCGCGCCCAGCTGCCCGACTGGAGCATGCTGCTGGCCGCGATCACCACCCTGTT
CCTGGCGGCCGAGAAGCAGTGGATGATGCTGGACTGGAAGCCCAAGCGCCCCGA
CATGCTGGTGGACCCCTTCGGCCTGGGCCGCTTCGTGCAGGACGGCCTGGTGTTC
CGCAACAACCTTCAGCATCCGCAGCTACGAGATCGGCGCGGACCGCACCGCCAGC
ATCGAGACCCTGATGAACCACCTGCAGGAGACCGCCCTGAACCACGTGAAGAGC
GTGGGCCTGCTGGAGGACGGCCTGGGCAGCACCCGCGAGATGAGCCTGCGCAAC
CTGATCTGGGTGGTGACCAAGATGCAGGTGGCGGTGGACCGCTACCCACCTGG
GGCGACGAGGTGCAGGTGAGCAGCTGGGCGACCGCCATCGGCAAGAACGGCAT
GCGCCGCGAGTGGATCGTGACCGACTTCCGCACCGGCGAGACCCTGCTGCGCGC
CACCAGCGTGTGGGTGATGATGAACAAGCTGACCCGCCGCATCAGCAAGATCCC
CGAGGAGGTGTGGCACGAGATCGGCCCCAGCTTCATCGACGCGCCCCCCCCTGCC
CACCGTGGAGGACGACGGCCGCAAGCTGACCCGCTTCGACGAGAGCAGCGCCGA

CTTCATCCGCAAGGGCCTGACCCCCCGCTGGAGCGACCTGGACATCAACCAGCA
CGTGAACAACGTGAAGTACATCGGCTGGCTGCTGGAGAGCGCGCCCCCGAGAT
CCACGAGAGCCACGAGATCGCCAGCCTGACCCTGGAGTACCGCCGCGAGTGCGG
CCGCGACAGCGTGTGAACAGCGCCACCAAGGTGAGCGACAGCAGCCAGCTGGG
CAAGAGCGCCGTGGAGTGCAACCACCTGGTGC GCCTGCAGAACGGCGGCGAGAT
CGTGAAGGGCCGCACCGTGTGGCGCCCCAAGCGCCCCCTGTACAACGACGGCGC
CGTGGTGGACGTGCCCCGCCAAGACCAGCGATGACGATGACAAGCTGGGATGACT
CGAG

SEQ ID NO: 94

Codon-optimized *C. hookeriana* thioesterase sequence

ACTAGTATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTCACG
ATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTC
GTGGGCGCGCCAGCTGCCCGACTGGAGCCGCCTGCTGACCGCCATCACCACCGT
GTTCTGTGAAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGA
CATGCTGGTGGACAGCTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTC
CGCCAGTCCCTTCTCCATCCGCTCCTACGAGATCGGCACCGACCGCACCGCCAGCA
TCGAGACCCTGATGAACCACCTGCAGGAGACCTCCCTGAACCACTGCAAGAGCA
CCGGCATCCTGCTGGACGGCTTCGGGCCGACCCCTGGAGATGTGCAAGCGCGACCT
GATCTGGGTGGTGATTAAGATGCAGATCAAGGTGAACCGCTACCCCGCCTGGGG
CGACACCGTGGAGATCAACACCCGCTTCAGCCGCCTGGGCAAGATCGGCATGGG
CCGCGACTGGCTGATCTCCGACTGCAACACCGGCGAGATCCTGGTGC GCGCCACC
AGCGCCTACGCCATGATGAACCAGAAGACCCGCCGCCTGTCCAAGCTGCCCTAC
GAGGTGCACCAGGAGATCGTGCCCTGTTCGTGGACAGCCCCGTGATCGAGGAC
TCCGACCTGAAGGTGCACAAGTTCAAGGTGAAGACCGGCGACAGCATCCAGAAG
GGCCTGACCCCCGGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGTG
AAGTACATCGGCTGGATCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAG
GAGCTGTGCTCCCTGGCCCTGGAGTACCGCCGCGAGTGCGGGCCGCGACTCCGTGC
TGGAGAGCGTGACCGCCATGGACCCAGCAAGGTGGGCGTGCGCTCCAGTACC
AGCACCTGCTGCGCCTGGAGGACGGCACCGCCATCGTGAACGGCGCCACCGAGT
GGCGCCCCAAGAACGCCGCGCCAACGGCGCCATCTCCACCGGCAAGACCAGCA
ACGGCAACTCCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACC
ACGACATCGACTACAAGGACGACGACGACAAGTGACTCGAG

SEQ ID NO: 95

U. californica thioesterase forward primer

CTGGGCGACGGCTTCGGCAC

SEQ ID NO: 96

U. californica thioesterase reverse primer

AAGTCGCGGCGCATGCCGTT

SEQ ID NO: 97

U. americana thioesterase forward primer

CCCAGCTGCTCACCTGCACC

SEQ ID NO: 98

U. americana thioesterase reverse primer

CACCCAAGGCCAACGGCAGCGCCGTG

SEQ ID NO: 99

C. hookeriana thioesterase forward primer

TACCCCGCCTGGGGCGACAC

SEQ ID NO: 100

C. hookeriana thioesterase reverse primer

AGCTTGGACAGGCGGCGGGT

SEQ ID NO: 101

FAD primer 1

5'-TCACTTCATGCCGGCGGTCC-3'

SEQ ID NO: 102

FAD primer 2

5'- GCGCTCCTGCTTGGCTCGAA-3'

SEQ ID NO: 103

pSZ1124 (FAD2B) 5' genomic targeting sequence

GCTCTTCGAGACGTGGTCTGAATCCTCCAGGCGGGTTTCCCCGAGAAAGAAAGG
GTGCCGATTTCAAAGCAGAGCCATGTGCCGGGCCCTGTGGCCTGTGTTGGCGCCT
ATGTAGTCACCCCCCTCACCCAATTGTCGCCAGTTTGCGCAATCCATAAACTCA
AAACTGCAGCTTCTGAGCTGCGCTGTTCAAGAACACCTCTGGGGTTTGCTCACCC
GCGAGGTCGACGCCAGCATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAG
CCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGC
GCTCGGCGCTTCGTAGCAGCATGTACCTGGCCTTTGACATCGCGGTCATGTCCCT
GCTCTACGTCGCGTCGACGTACATCGACCCTGCGCCGGTGCCTACGTGGGTCAAG
TATGGCGTCATGTGGCCGCTCTACTGGTTCTTCCAGGTGTGTGTGAGGGTTGTGG
TTGCCCGTATCGAGGTCCTGGTGGCGCGCATGGGGGAGAAGGCGCCTGTCCCGCT
GACCCCCCGGCTACCCTCCCGGCACCTTCCAGGGCGCCTTCGGCACGGGTGTCT
GGGTGTGCGCGCACGAGTGCGGCCACCAGGCCTTTTCCTCCAGCCAGGCCATCAA
CGACGGCGTGCGCCTGGTGTTCACAGCCTGCTGCTGGTGGCCTACTACTCCTGG
AAGCACTCGCACCGGGTACC

SEQ ID NO: 104

pSZ1124 (FAD2B) 3' genomic targeting sequence

CCGCCACCACTCCAACACGGGGTGCCTGGACAAGGACGAGGTGTTTGTGCCGCC
GCACCGCGCAGTGGCGCACGAGGGCCTGGAGTGGGAGGAGTGGCTGCCATCCG
CATGGGCAAGGTGCTGGTCACCCTGACCCTGGGCTGGCCGCTGTACCTCATGTTT

AACGTCGCCTCGCGGCCGTACCCGCGCTTCGCCAACCACCTTTGACCCGTGGTCGC
 CCATCTTCAGCAAGCGCGAGCGCATCGAGGTGGTCATCTCCGACCTGGCGCTGGT
 GGCGGTGCTCAGCGGGCTCAGCGTGCTGGGCCGCACCATGGGCTGGGCTGGCT
 GGTCAAGACCTACGTGGTGCCCTACCTGATCGTGAACATGTGGCTCGTGCTCATC
 ACGCTGCTCCAGCACACGCACCCGGCGCTGCCGCACTACTTCGAGAAGGACTGG
 GACTGGCTGCGCGGCCCATGGCCACCGTGGACCGCTCCATGGGCCCCGCCCTTCA
 TGGACAACATCCTGCACCACATCTCCGACACCCACGTGCTGCACCACCTCTTCAG
 CACCATCCCGCACTACCACGCCGAGGAGGCCTCCGCCGCCATCAGGCCCATCCTG
 GGCAAGTACTACCAGTCCGACAGCCGCTGGGTGCGCCGCGCCCTGTGGGAGGAC
 TGGCGCGACTGCCGCTACGTGCTCCCGACGCGCCCGAGGACGACTCCGCGCTCT
 GGTTCACAAGTGAGTGAGTGAGAAGAGC

SEQ ID NO: 105

S. cerevisiae suc2 cassette

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
 CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGG
 CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
 CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
 TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
 GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTG
 CAGGCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGA
 CGAACGAGACGTCCGACCGCCCCCTGGTGCACCTTACCCCCAACAAGGGCTGGA
 TGAACGACCCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGT
 ACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCGTTCTGGGGCCA
 CGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCC
 GAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAA
 CACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATC
 TGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGAC
 GCGCGCTACACCTTACCCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCC
 ACCCAGTTCGCGGACCCGAAGGTCCTTCTGGTACGAGCCCTCCCAGAAGTGGATCA
 TGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCT
 GAAGTCCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGGCTTCCTCGGCTACCA
 GTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCCAGCAAGTC
 TACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTC
 AACCAGTACTTCGTCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACC
 AGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAA
 CACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGA
 GTACTCCGCCTTCGTGCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCA
 AGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACC
 TGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCTG
 CCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACA
 GCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGA
 TCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCC
 CGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCCTCCTTCTTCCTGGAC
 CGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGC
 ATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGACCTGTCTACTACAAG
 GTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGAC
 GTCGTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGA
 ACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCCAGGTGCGCG

AGGTCAAGTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGA

SEQ ID NO: 106

pSZ1125 (FAD2C) 5' genomic targeting sequence

GCTCTTCGAGGGGCTGGTCTGAATCCTTCAGGCGGGTGTACCCGAGAAAGAAA
GGGTGCCGATTTCAAAGCAGACCCATGTGCCGGGCCCTGTGGCCTGTGTTGGCGC
CTATGTAGTACCCCCCTCACCAATTGTGCCAGTTTGCGCACTCCATAAACTC
AAAACAGCAGCTTCTGAGCTGCGCTGTTCAAGAACACCTCTGGGGTTTGCTCACC
CGCGAGGTCGACGCCAGCATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAA
GCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAG
CGCTCGGCGCTTCGTAGCAGCATGTACCTGGCCTTTGACATCGCGGTATGTCCC
TGCTCTACGTCGCGTCGACGTACATCGACCCTGCACCGGTGCCTACGTGGGTCAA
GTACGGCATCATGTGGCCGCTCTACTGGTTCTTCCAGGTGTGTTTGAGGGTTTTGG
TTGCCCGTATTGAGGTCTGGTGGCGCGCATGGAGGAGAAGGCGCCTGTCCCGCT
GACCCCCCGGCTACCCTCCCGGCACCTTCCAGGGCGCCTTCGGCACGGGTGTCT
GGGTGTGCGCGCACGAGTGCGGCCACCAGGCCTTTTCCTCCAGCCAGGCCATCAA
CGACGGCGTGGGCCTGGTGTTCACAGCCTGCTGCTGGTGCCCTACTACTCCTGG
AAGCACTCGCACCGGGTACC

SEQ ID NO: 107

pSZ1125 (FAD2C) 3' genomic targeting sequence

CCGCCACCACTCCAACACGGGGTGCCTGGACAAGGACGAGGTGTTTGTGCCGCC
GCACCGCGCAGTGGCGCACGAGGGCCTGGAGTGGGAGGAGTGGCTGCCCATCCG
CATGGGCAAGGTGCTGGTCACCTGACCCTGGGCTGGCCGCTGTACCTCATGTTC
AACGTCGCCTCGCGGCCGTACCCGCGCTTCGCCAACCACCTTTGACCCGTGGTCGC
CCATCTTCAGCAAGCGCGAGCGCATCGAGGTGGTCATCTCCGACCTGGCGCTGGT
GGCGGTGCTCAGCGGGCTCAGCGTGCTGGGCGGCACCATGGGCTGGGCCTGGCT
GGTCAAGACCTACGTGGTGCCCTACCTGATCGTGAACATGTGGCTCGTGCTCATC
ACGCTGCTCCAGCACACGCACCCGGCGCTGCCGCACTACTTCGAGAAGGACTGG
GACTGGCTGCGCGGCGCCATGGCCACCGTGGACCGCTCCATGGGCCCCGCCCTTCA
TGGACAACATCCTGCACCACATCTCCGACACCCACGTGCTGCACCACCTCTTCAG
CACCATCCCGCACTACCACGCCGAGGAGGCCTCCGCCGCCATCAGGCCCATCCTG
GGCAAGTACTACAGTCCGACAGCCGCTGGGTGCGCCGCGCCCTGTGGGAGGAC
TGGCGCGACTGCCGCTACGTCGTCCCGGACGCGCCCCGAGGACGACTCCGCGCTCT
GGTTCCACAAGTGAGTGAGTGAGAAGAGC

SEQ ID NO: 108

S. carlbergensis MEL1 amino acid sequence

MFAFYFLTACISLKGVFVSPSYNGLGLTPQMGWDNWNTFACDVSEQLLLDTADRI
SDLGLKDMGYKYIILDDCWSSGRDSDGFLVADEQKFPNGMGHVAHDLHNNSFLFG

MYSSAGEYTCAGYPGSLGREEEDAQFFANNRVDYLKYDNCYNKGQFGTPEISYHRY
KAMSDALNKTGRPIFYSLCNWGQDLTFYWGSIGANSWRMSGDVTAEFTRPDSRCP
DGDEYDCKYAGFHCSIMNILNKAAPMGQNAGVGGWNDLDNLEVGVGNLTDDEEK
AHFSMWAMVKSPLIIGANVNNLKASSYSIYSQASVIINQDSNGIPATRVWRYVSD
TDEYGQGEIQMWSGPLDNGDQVVALLNGGSVSRPMNTTLEEIFFDSNLGSKKLTST
WDIYDLWANRVDNSTASAILGRNKTATGILYNATEQSYKDGLSKNDTRLFGQKIGSL
SPNAILNTTTPAHGIAFYRLRPSS

SEQ ID NO: 109

S. carlbergensis MEL1 native signal peptide amino acid sequence

MFAFYFLTACISLKGVFG

SEQ ID NO: 110

Relevant sequence of *S. carlbergensis* MEL1 expression cassette

GCGGCCGCGTGGACGAGGGCTACAACCCCGCCTACGGCGCGCGCCCGCTGCGCC
GCGCCATCATGCGCCTGCTGGAGGACGCGCTGGCCGAGCGCATGCTCGCCGGCG
ACGTCAAGGAGGGCGACTCGGTCATCATGGACGTCGATGGCGATGGCGCCATCA
GCGTCCTCAACGGAGACCGCACCCACACCACCATCGACTCCTCCCCGGCGG
GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCGGGGGCCCTTGGGGCGCTGCTGGAGCGTGGTGAGAGCGCGGC
GGACGTGCCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTGCG
GCCCCGATCCCCCACTGCCCCCATCATCTTGCCTGTTGTCGTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTCTTTCTTGCCTATGACAC
TTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAAC
ACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGAT
GCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACAT
TATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACAC
AGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCTCTTCGT
TTCAGTCACAACCCGCAAACACTAGTATGTTTCGCGTTCTACTTCTGACGGCCTG
CATCTCCCTGAAGGGCGTGTTCGGCGTCTCCCCCTCCTACAACGGCCTGGGCCTG
ACGCCCCAGATGGGCTGGGACAACCTGGAACACGTTTCGCCTGCGACGTCTCCGAG
CAGCTGCTGCTGGACACGGCCGACCGCATCTCCGACCTGGGCCTGAAGGACATG
GGCTACAAGTACATCATCCTGGACGACTGCTGGTCCTCCGGCCGCGACTCCGACG
GCTTCCTGGTCGCCGACGAGCAGAAGTTCCCAACGGCATGGGCCACGTGCGCG
ACCACCTGCACAACAACCTCCTTCTGTTTCGGCATGTACTCCTCCGCGGGCGAGTA
CACGTGCGCCGGCTACCCCGGCTCCCTGGGCGCGAGGAGGAGGACGCCCAGTT
CTTCGCGAACAACCGCGTGGACTACCTGAAGTACGACAACCTGCTACAACAAGGG
CCAGTTTCGGCACGCCCCGAGATCTCCTACCACCGCTACAAGGCCATGTCCGACGCC
CTGAACAAGACGGGCGGCCCATCTTCTACTCCCTGTGCAACTGGGGCCAGGACC
TGACCTTCTACTGGGGCTCCGGCATCGCGAAGCTCCTGGCGCATGTCCGGCGACGT
CACGGCGGAGTTCACGCGCCCCGACTCCCGCTGCCCTGCGACGGCGACGAGTA
CGACTGCAAGTACGCCGGCTTCCACTGCTCCATCATGAACATCCTGAACAAGGCC
GCCCCCATGGGCCAGAACGCGGGCGTCCGGCGCTGGAACGACCTGGACAACCTG
GAGGTTCGGCGTTCGGCAACCTGACGGACGACGAGGAGAAGGCGCACTTCTCCATG
TGGGCCATGGTGAAGTCCCCCTGATCATCGGCGCGAAGCTGAACAACCTGAAG
GCCTCCTCCTACTCCATCTACTCCCAGGCGTCCGTCATCGCCATCAACCAGGACT
CCAACGGCATCCCCGCCACGCGCGTCTGGCGCTACTACGTGTCCGACACGGACG

AGTACGGCCAGGGCGAGATCCAGATGTGGTCCGGCCCCCTGGACAACGGCGACC
 AGGTCGTGGCGCTGCTGAACGGCGGCTCCGTGTCCCGCCCCATGAACACGACCCCT
 GGAGGAGATCTTCTTCGACTCCAACCTGGGCTCCAAGAAGCTGACCTCCACCTGG
 GACATCTACGACCTGTGGGCGAACC GCGTCGACAACCTCCACGGCGTCCGCCATCC
 TGGGCGCGAACAAGACCGCCACCGGCATCCTGTACAACGCCACCGAGCAGTCCT
 ACAAGGACGGCCTGTCCAAGAACGACACCCGCTGTTCGGCCAGAAGATCGGCT
 CCCTGTCCCCAACGCGATCCTGAACACGACCGTCCCCGCCACGGCATCGCGTT
 CTACCGCTGCGCCCCCTCCTCCTGACAATTGAAGCAGCAGCAGCTCGGATAGTAT
 CGACACACTCTGGACGCTGGTTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCC
 TTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTT
 GTGTGTACGCGCTTTTGGCAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCC
 CCAGCATCCCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTAC
 GCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAG
 CCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACT
 GCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAG
 CTCAGAATAGTATCGGGTGATGCGAAGTCAGAACCGAGGCAGGGCCTGTCGCCTG
 AGGTGGCAACGATGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATCCCG
 TGATCTCCCCACCGACACCTATCCCCGCCCATCCCGGCCACCCCTTTCAGTCCCC
 TCAGCATGCATTGTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGGTTTCAGGC
 CACGGCGCACCGAGTCCTCGCCCTTCGCGAGAGTGATCACCCCTCCCCGGAACCAGC
 CACGCTCGCTGCTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGGACA
 ACTCAGGCAACCACGCGCCTCACAAGCATGGCCGCCGTGGCATCCAACCCGCTC
 GTGACGGTGGGTGCGCAAGTGCCAGGGGCCCTCGTCGTCACGGCGTGATCCTCG
 AGGGATGCGATCCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGATGA
 GCCCCGTGTCTAGA

SEQ ID NO: 111

Codon-optimized *S. carbergensis* MEL1 sequence

ATGTTTCGCGTTCTACTTCCTGACGGCCTGCATCTCCCTGAAGGGCGTGTTTCGGCGT
 CTCCCCCTCCTACAACGGCCTGGGCCTGACGCCCCAGATGGGCTGGGACAACCTGG
 AACACGTTTCGCTGCGACGTCTCCGAGCAGCTGCTGCTGGACACGGCCGACCGC
 ATCTCCGACCTGGGCCTGAAGGACATGGGCTACAAGTACATCATCTGGACGACT
 GCTGGTCTCTCCGGCCGCGACTCCGACGGCTTCTTGGTCGCCGACGAGCAGAAGTT
 CCCCACGGCATGGGCCACGTGCGCGACCACCTGCACAACAACCTCCTTCCTGTTT
 GGCATGTACTCCTCCGCGGGCGAGTACACGTGCGCCGGCTACCCCGGCTCCCTGG
 GCCGCGAGGAGGAGGACGCCAGTTCTTCGCGAACAACCGCGTGGAACCTGA
 AGTACGACAACCTGCTACAACAAGGGCCAGTTTCGGCACGCCCCGAGATCTCCTACC
 ACCGCTACAAGGCCATGTCCGACGCCCTGAACAAGACGGGGCCGCCCATCTTCTA
 CTCCCTGTGCAACTGGGGCCAGGACCTGACCTTCTACTGGGGCTCCGGCATCGCG
 AACTCCTGGCGCATGTCCGGCGACGTACGGCGGAGTTACGCGCCCCGACTCCC
 GCTGCCCCCTGCGACGGCGACGAGTACGACTGCAAGTACGCCGGCTTCCACTGCTC
 CATCATGAACATCCTGAACAAGGCCGCCCCCATGGGCCAGAACGCGGGCGTCGG
 CGGCTGGAACGACCTGGACAACCTGGAGGTGCGCGTCGGCAACCTGACGGACGA
 CGAGGAGAAGGCGCACTTCTCCATGTGGGCCATGGTGAAGTCCCCCTGATCATC
 GCGCGAACGTGAACAACCTGAAGGCCTCCTCCTACTCCATCTACTCCAGGCGT
 CCGTCATCGCCATCAACCAGGACTCCAACGGCATCCCCGCCACGCGCGTCTGGCG
 CTACTACGTGTCCGACACGGACGAGTACGGCCAGGGCGAGATCCAGATGTGGTC
 CGGCCCCCTGGACAACGGCGACCAAGGTGCTGGCGCTGCTGAACGGCGGGCTCCGT
 GTCCCGCCCCATGAACACGACCCCTGGAGGAGATCTTCTTCGACTCCAACCTGGGC

TCCAAGAAGCTGACCTCCACCTGGGACATCTACGACCTGTGGGCGAACC GCGTC
GACAACTCCACGGCGTCCGCCATCCTGGGCGCAACAAGACCGCCACCGGCATC
CTGTACAACGCCACCGAGCAGTCCTACAAGGACGGCCTGTCCAAGAACGACACC
CGCCTGTTTCGGCCAGAAGATCGGCTCCCTGTCCCCAACGCGATCCTGAACACGA
CCGTCCCCGCCACGGCATCGCGTTCTACCGCCTGCGCCCCCTCCTCCTGA

SEQ ID NO: 112

5' Clp homologous recombination targeting sequence

GCGGCCGCGTGGACGAGGGCTACAACCCCGCCTACGGCGCGCGCCCGCTGCGCC
GCGCCATCATGCGCCTGCTGGAGGACGCGCTGGCCGAGCGCATGCTCGCCGGCG
ACGTCAAGGAGGGCGACTCGGTTCATCATGGACGTCGATGGCGATGGCGCCATCA
GCGTCCTCAACGGAGACCGCACCCACACCACCACCATCGACTCCTCCCCGGCGG
GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCGGGGGCCTTGGGGCGCTGCTGGAGCGTGGTGAGAGCGCGGG
GGACGTGCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTTCG
GCCCGCATCCCCCACTGCCCCCATCATCTTGCCTGTTGTCGTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTC

SEQ ID NO: 113

3' Clp homologous recombination targeting sequence

GAGCTCAGAATAGTATCGGGTGATGCGAAGTCAGAACCAGGCAGGGCCTGTCGC
CTGAGGTGGCAACGATGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATC
CCGTGATCTCCCCACCGACACCTATCCCCGCCCATCCCGGCCACCCTTTTCAGTC
CCCTCAGCATGCATTGTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGGTTCA
GGCCACGGCGCACCGAGTCCTCGCCCTTCGCAGAGTGATCACCCCTCCCCGGAACC
AGCCACGCTCGCTGCTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGG
ACAACCTCAGGCAACCACGCGCCTCACAAGCATGGCCGCCGTGGCATCCAACCCG
CTCGTGACGGTGGGTGCGCAAGTGCCAGGGGCCCTCGTCGTACGGCGTGATCCT
CGAGGGATGCGATCCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGAT
GAGCCCCTGTCTAGA

SEQ ID NO: 114

Chlamydomonas reinhardtii TUB2 promoter/5' UTR

CTTTCTTGCCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAAC

SEQ ID NO: 115

Chlorella vulgaris nitrate reductase 3'UTR

GCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTTCGTGTGATGGA
CTGTGCGCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAA

ACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGGCGAGTTGCTAGCTGCT
TGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTTG
CATCCCAACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCT
CCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCTGTATTCTCCTGGT
ACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATG
GGAACACAAATGGAAAGCTT

SEQ ID NO: 116

A. niger AlgC amino acid sequence

MIGSSHAVVALGLFTLYGHSAAAPAIGASNSQTIVTNGTSFALNGDNVSYRFHVNSS
TGDILSDHFGGVVSGTIPSPVEPAVNGWVGMPGRIRREFPDQGRGDFRIPAVRIRESA
GYTVSDLQYVSHEVIEGKYALPGLPATFGDAQDATTLLVVHLYDNYSSVAADLSYSIF
PKYDAIVRSVNVNTNQGPGNITIEALASISIDFPYEDLDMVSLRGDWAREANVQRSKV
QYGVQGFSGSTGYSSHLHNPFLAIVDPATTESQGEAWGFNLVYTGSFSAQVEKGSQG
FTRALLGFNPDQLSWNLGPGETLTSPECVAVYSDKGLGVSVRKFHRLYRNHLMKSK
FATSDRPVLLNSWEGVYFDYNQSSIETLAESAALGVHLFVMDDGWFGDKYPRVSD
NAGLGDWMPNPARFPDGLTPVVQDITNLTVNGTESTKLRFGIWVEPEMVNPNSTLY
HEHPEWALHAGPYPRTERRNQLVLNLALPAVQDFIIDFMTNLLQDTGISYVKWDNN
RGIHETPSPSTDHQYMLGLYRVFDTLTTRFPDVLWEGCASGGGRFDAGMLQYVPQI
WTSNDTDAIDRITIQFGTSLAYPPSAMGAHLSAVPNAQTGRTPFTFRAHVAMMGGS
FGLELDPATVEGDEIVPELLALAEKVNPIILNGDLYRLRLPQDSQWPAALFVSQDGAQ
AVLFYFQVQPNVNHAVPWVRLQGLDPKADYTVDGDQTYSGATLMNLGLQYSFDTE
YGSKVVFLEERQ

SEQ ID NO: 117

A. niger AlgC native signal peptide amino acid sequence

MIGSSHAVVALGLFTLYGHSAAAPAIGA

SEQ ID NO: 118

Codon-optimized *A. niger AlgC* coding sequence

ATGATCGGCTCCTCCCACGCGGTCGTGCGCCCTGGGCCTGTTACCCCTGTACGGCC
ACTCCGCGCGCCGCGCCCGCCATCGGCGCCTCCAACTCCCAGACCATCGTCACGAA
CGGCACCTCCTTCGCCCTGAACGGCGACAACGTGTCCTACCGCTTCCACGTGAAC
TCCTCCACGGGCGACCTGATCTCCGACCACTTCGGCGGCGTGGTGTCCGGCACCA
TCCCCCTCCCCCGTGGAGCCCGCGGTCAACGGCTGGGTGGGCATGCCCGGCCGCAT
CCGCGCGAGTTCCCCGACCAAGGGCCGCGGCGACTTCCGCATCCCCGCGGTGCGC
ATCCGCGAGTCCGCGGCTACACCGTCTCCGACCTGCAGTACGTGTCCACGAGG
TGATCGAGGGCAAGTACGCGCTGCCCGGCCTGCCCGCCACGTTCCGGCGACGCCC
AGGACGCCACCACCCTGGTGGTGCACCTGTACGACAACACTACTCCTCCGTCGCGGC
CGACCTGTCCTACTCCATCTTCCCCAAGTACGACGCGATCGTCCGCTCCGTGAAC
GTGACCAACCAGGGCCCCGGCAACATCACCATCGAGGCGCTGGCCTCCATCTCC
ATCGACTTCCCCTACGAGGACCTGGACATGGTGTCCCTGCGCGGCGACTGGGCCC
GCGAGGCGAACGTGCAGCGCTCCAAGGTCCAGTACGGCGTGCAGGGCTTCGGCT
CCTCCACCGGCTACTCCTCCCACCTGCACAACCCCTTCTGGCGATCGTCGACCC
CGCGACCACCGAGTCCCAGGGCGAGGCTGGGGCTTCAACCTGGTCTACACCGG
CTCCTTCTCCGCCCAGGTTCGAGAAGGGCTCCCAGGGCTTACGCGCGCCCTGCTG
GGCTTCAACCCCGACCAAGCTGTCTTGAACCTGGGCCCCGGCGAGACGCTGACG

TCCCCCGAGTGC GTTCGCCGTCTACTCCGACAAGGGCCTGGGCTCCGTCTCCCGCA
AGTTCCACCGCCTGTACCGCAACCACCTGATGAAGTCCAAGTTGCGCACGTCCGA
CCGCCCCGTGCTGCTGAACTCCTGGGAGGGCGTCTACTTCGACTACAACCAGTCC
TCCATCGAGACGCTGGCGGAGGAGTCCGCCGCCCTGGGCGTGCACCTGTTTCGTCA
TGGACGACGGCTGGTTCGGCGACAAGTACCCCCGCGTGTCCGACAACGCCGGCC
TGGGCGACTGGATGCCCAACCCCGCCCCGCTTCCCCGACGGCCTGACGCCCGTCGT
GCAGGACATCACCAACCTGACCGTCAACGGCACCGAGTCCACCAAGCTGCGCTT
CGGCATCTGGGTGGAGCCCCGAGATGGTGAACCCCAACTCCACCCTGTACCACGA
GCACCCCGAGTGGGCGCTGCACGCGGGCCCCCTACCCCGCACCGAGCGCCGCAA
CCAGCTGGTCTGAACCTGGCCCTGCCCGCGGTCCAGGACTTCATCATCGACTTC
ATGACCAACCTGCTGCAGGACACCGGCATCTCCTACGTCAAGTGGGACAACAAC
CGCGGCATCCACGAGACGCCCTCCCCCTCCACGGACCACCAGTACATGCTGGGCC
TGTACCGCGTGTTCGACACGCTGACCACGCGCTTCCCCGACGTCTGTGGGAGGG
CTGCGCGTCCGGCGGCGGCCGGTTCGACGCCGGCATGCTGCAGTACGTCCCCCAG
ATCTGGACGTCCGACAACACGGACGCGATCGACCGCATCACCATCCAGTTCGGC
ACCTCCCTGGCCTACCCCCCTCCGCCATGGGCGCCCACCTGTCCGCGGTGCCCA
ACGCCCAGACCGGCCGCACGGTGCCCTTCACCTTCCGCGCCCCACGTGCGCATGAT
GGGCGGCTCCTTCGGCCTGGAGCTGGACCCCGCCACCGTGGAGGGCGACGAGAT
CGTGCCCGAGCTGCTGGCGCTGGCCGAGAAGGTGAACCCCATCATCCTGAACGG
CGACCTGTACCGCCTGCGCCTGCCCCAGGACTCCCAGTGGCCCGCGGCCCTGTTC
GTGTCCCAGGACGGCGCCCAGGCCGTCTTCTACTTCCAGGTGCAGCCCAACG
TCAACCACGCCGTCCCCCTGGGTCCGCTGCAGGGCCTGGACCCCAAGGCCGACTA
CACGGTGGACGGCGACCAGACGTACTCCGGCGCGACCCTGATGAACCTGGGCCT
GCAGTACTCCTTCGACACCGAGTACGGCTCCAAGGTGGTGTTCCTGGAGCGCCAG
TAA

SEQ ID NO: 119

Relevant sequence of *A. niger* *AlgC* expression cassette

GCGGCCGCGTGGACGAGGGCTACAACCCCGCCTACGGCGCGCGCCCGCTGCGCC
GCGCCATCATGCGCCTGCTGGAGGACGCGCTGGCCGAGCGCATGCTCGCCGGCG
ACGTCAAGGAGGGCGACTCGGTATCATGGACGTCGATGGCGATGGCGCCATCA
GCGTCCTCAACGGAGACCGCACCCACACCACCACCATCGACTCCTCCCCGGCGG
GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCGGGGGCCTTGGGGCGCTGCTGGAGCGTGGTGAGAGCGCGGC
GGACGTGCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTGCG
GCCCCGATCCCCCACTGCCCCCATCATCTTGCCCTGTTGTCGTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTCTTTCTTTCGCTATGACAC
TTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAAC
ACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGAT
GCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACAT
TATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACAC
AGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCTCTTCGT
TTCAGTCACAACCCGCAAACACTAGTATGATCGGCTCCTCCACGCGGTTCGTGCG
CCTGGGCCTGTTACCCCTGTACGGCCACTCCGCCGCCGCGCCCGCCATCGGCGCC
TCCAACTCCCAGACCATCGTCACGAACGGCACCTCCTTCGCCCTGAACGGCGACA
ACGTGTCTTACCGCTTCCACGTGAACCTCTCCACGGGCGACCTGATCTCCGACCA
CTTCGGCGGGCGTGGTGTCCGGCACCATCCCCTCCCCCGTGGAGCCCCGCGGTCAAC
GGCTGGGTGGGCATGCCCGGCCGCATCCGCCGCGAGTTCCCCGACCAGGGCCGC

GGCGACTTCCGCATCCCCGCGGTGCGCATCCGCGAGTCCGCCGGCTACACCGTCT
CCGACCTGCAGTACGTGTCCCACGAGGTGATCGAGGGCAAGTACGCGCTGCCCG
GCCTGCCCCGCCACGTTCCGGCGACGCCCAGGACGCCACCACCCTGGTGGTGCACCT
GTACGACAACTACTCCTCCGTCGCGGCCGACCTGTCTACTCCATCTTCCCCAAG
TACGACGCGATCGTCCGCTCCGTGAACGTGACCAACCAGGGCCCCGGCAACATC
ACCATCGAGGCGCTGGCCTCCATCTCCATCGACTTCCCCCTACGAGGACCTGGACA
TGGTGTCCCTGCGCGGGCGACTGGGGCCCCGCGAGGCGAACGTGCAGCGCTCCAAGG
TCCAGTACGGCGTGCAGGGCTTCGGGCTCCTCCACCGGCTACTCCTCCCACCTGCA
CAACCCCTTCCTGGCGATCGTCGACCCCCGCGACCACCGAGTCCCAGGGCGAGGC
CTGGGGGCTTCAACCTGGTCTACACCGGCTCCTTCTCCGCCCAGGTGCGAGAAGGGC
TCCCAGGGCTTCACGCGCGCCCTGCTGGGGCTTCAACCCCGACCAGCTGTCTTGGA
ACCTGGGGCCCCGGCGAGACGCTGACGTCCCCCGAGTGCCTGCGCGTCTACTCCGA
CAAGGGCCTGGGCTCCGTCTCCCGCAAGTTCCACCGCCTGTACCGCAACCACCTG
ATGAAGTCCAAGTTCGCCACGTCCGACCGCCCCGTGCTGCTGAACTCCTGGGAGG
GCGTCTACTTCGACTACAACCAGTCCTCCATCGAGACGCTGGCGGAGGAGTCCGC
CGCCCTGGGCGTGCACCTGTTCGTATGGACGACGGCTGGTTTCGGCGACAAGTAC
CCCCGCGTGTCCGACAACGCCGGCCTGGGCGACTGGATGCCCAACCCCGCCCCGT
TCCCCGACGGCCTGACGCCCGTCGTGCAGGACATCACCAACCTGACCGTCAACG
GCACCGAGTCCACCAAGCTGCGCTTCGGCATCTGGGTGGAGCCCCGAGATGGTGA
ACCCCAACTCCACCCTGTACCACGAGCACCCCGAGTGGGCGCTGCACGCGGGCC
CCTACCCCCCGCACCGAGCGCCGCAACCAGCTGGTCCTGAACCTGGCCCTGCCCGC
GGTCCAGGACTTCATCATCGACTTCATGACCAACCTGCTGCAGGACACCGGCATC
TCCTACGTCAAAGTGGGACAACAACCGCGGCATCCACGAGACGCCCTCCCCCTCCA
CGGACCACCAGTACATGCTGGGCCCTGTACCGCGTGTTCGACACGCTGACCACGCG
CTTCCCCGACGTCTGTGGGAGGGCTGCGCGTCCGGCGGGCGGCCGTTTCGACGCC
GGCATGCTGCAGTACGTCCCCCAGATCTGGACGTCCGACAACACGGACGCGATC
GACCGCATCACCATCCAGTTCGGCACCTCCCTGGCCTACCCCCCTCCGCCATGG
GCGCCCACCTGTCCGCGGTGCCCAACGCCCAGACCGGCCGCACGGTGCCCTTCAC
CTTCCGCGCCACGTGCGGATGATGGGCGGGCTCCTTCGGCCTGGAGCTGGACCCC
GCCACCGTGGAGGGCGACGAGATCGTGCCCGAGCTGCTGGCGCTGGCCGAGAAG
GTGAACCCCATCATCTGAACGGCGACCTGTACCGCCTGCGCCTGCCCCAGGACT
CCCAGTGGCCCCGCGGCCCTGTTCGTGTCCAGGACGGCGCCAGGCCGTCTGT
CTACTTCCAGGTGCAGCCCAACGTCAACCACGCCGTCCCCTGGGTCCGCCTGCAG
GGCCTGGACCCCAAGGCCGACTACACGGTGGACGGCGACCAGACGTACTCCGGC
GCGACCCTGATGAACCTGGGCCTGCAGTACTCCTTCGACACCGAGTACGGCTCCA
AGGTGGTGTTCCTGGAGCGCCAGTAACAATTGAAGCAGCAGCAGCTCGGATAGT
ATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTG
CCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATC
TTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCAC
CCCCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCCAACCGCAACTTATCT
ACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCAC
AGCCTTGTTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGC
ACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTT
GAGCTCAGAATAGTATCGGGTGATGCGAAGTCAGAACCAGGCAGGGCCTGTGCG
CTGAGGTGGCAACGATGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATC
CCGTGATCTCCCCACCGACACCTATCCCCGCCCATCCCGGCCACCCCTTTCAGTC
CCCTCAGCATGCATTGTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGGTTCA
GGCCACGGCGCACCGAGTCCTCGCCCTTCGCGAGTGTATCACCTCCCCGGAAACC
AGCCACGCTCGCTGCTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGG
ACAACCTCAGGCAACCACGCGCCTCAACAAGCATGGCCGCCGTGGCATCCAACCCG

CTCGTGACGGTGGGTGCGCAAGTGCCAGGGGCCTCGTCGTCACGGCGTGTCATCCT
CGAGGGATGCGATCCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGAT
GAGCCCCCTGTCTAGA

SEQ ID NO: 120

C. tetragonobola α -galactosidase amino acid sequence

MATHYSIIIGGMIIVVLLMIIGSEGGRLLEKKNRTSAEAEHYNVRRYLAENGLGQTPPM
GWNSWNHFGCDINENVRETADAMVSTGLAALGYQYINLDDCWAEINRDSEGNM
VPNAAAFPSGIKALADYVHSKGLKLGVSADAGNQTC SKRMPGSLGHEEQDAKTFAS
WGVDYLKYDNCENLGISVKERYPPMGKALLSSGRPIFFSMCEWGWEDPQIWAKSIG
NSWRTTGDIEDNWNMTSIADSNKWSYAGPGGWNDPDMLEVGNNGMTTEEYR
SHFSIWALAKAPLLVGCDIRAMDDTTHELISNAEVIAVNQDKLGVQGKKVKSTNDLE
VWAGPLSDNKVAVILWNRSSSRATVTASWSDIGLQQGTTVDARDLWEHSTQSLVSG
EISAEIDSHACKMYVLTPRS

SEQ ID NO: 121

Codon-optimized *C. tetragonobola* α -galactosidase sequence

ATGGCCACCCACTACTCCATCATCGGCGGCATGATCATCGTCGTCCTGCTGATGA
TCATCGGCTCCGAGGGCGGCCGGCTGCTGGAGAAGAAGAACCGCACCTCCGCCG
AGGCGGAGCACTACAACGTGCGCCGCTACCTGGCCGAGAACGGCCTGGGCCAGA
CCCCCCCCATGGGCTGGAACCTCTGGAACCACTTCGGCTGCGACATCAACGAGA
ACGTCGTCCGCGAGACGGCGGACGCCATGGTGTCCACCGGCCTGGCCGCCCTGG
GCTACCAGTACATCAACCTGGACGACTGCTGGGCGGAGCTGAACCGCGACTCCG
AGGGCAACATGGTGCCCAACGCCGCCGCGTTCCTCCCGCATCAAGGCGCTGG
CGGACTACGTCCACTCCAAGGGCCTGAAGCTGGGCGTCTACTCCGACGCGGGCA
ACCAGACCTGCTCCAAGCGCATGCCCCGGCTCCCTGGGCCACGAGGAGCAGGACG
CGAAGACGTTTCGCTCCTGGGGCGTCGACTACCTGAAGTACGACAACCTGCGAGA
ACCTGGGCATCTCCGTGAAGGAGCGCTACCCCCCATGGGCAAGGCGCTGCTGTG
CTCCGGCCGCCCATCTTCTTCTCCATGTGCGAGTGGGGCTGGGAGGACCCCCAG
ATCTGGGCCAAGTCCATCGGCAACTCCTGGCGCACCAACGGCGACATCGAGGAC
AACTGGAACCTCATGACCTCCATCGCCGACTCCAACGACAAGTGGGCCTCCTACG
CCGGCCCCGGCGGCTGGAACGACCCCGACATGCTGGAGGTCGGCAACGGCGGCA
TGACGACGGAGGAGTACCGCTCCCACTTCTCCATCTGGGCCCTGGCCAAGGCGCC
CCTGCTGGTGGGCTGCGACATCCGCGCGATGGACGACACGACGACGAGCTGAT
CTCCAACGCCGAGGTGATCGCGGTGAACCAGGACAAGCTGGGCGTGCAGGGCAA
GAAGGTCAAGTCCACGAACGACCTGGAGGTCTGGGCCGGCCCCCTGTCCGACAA
CAAGGTGGCGGTGATCCTGTGGAACCGCTCCTCCTCCCGCGCCACCGTCACCGCG
TCCTGGTCCGACATCGGCCTGCAGCAGGGCACCAACCGTCGACGCGCGACCTGT
GGGAGCACTCCACGCAGTCCCTGGTGTCCGGCGAGATCTCCGCCGAGATCGACTC
CCACGCCTGCAAGATGTACGTCTGACGCCCCGCTCCTAA

SEQ ID NO: 122

Relevant sequence of *C. tetragonobola* α -galactosidase expression cassette

GCGGCCGCGTGGACGAGGGCTACAACCCCGCCTACGGCGCGCGCCCGCTGCGCC
GCGCCATCATGCGCCTGCTGGAGGACGCGCTGGCCGAGCGCATGCTCGCCGGCG
ACGTCAAGGAGGGCGACTCGGTTCATCATGGACGTCGATGGCGATGGCGCCATCA
GCGTCTCAACGGAGACCGCACCCACACCACCATCGACTCCTCCCCGGCGG

GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCGGGGGCCTTGGGGCGCTGCTGGAGCGTGGTGAGAGCGCGGC
GGACGTGCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTTCGC
GCCCCATCCCCCACTGCCCCCATCATCTTGCCTGTTGTCGTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTCCCTTCTTGCCTATGACAC
TTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAAC
ACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGAT
GCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACAT
TATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACAC
AGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCCTCTTCGT
TTCAGTCAACAACCCGCAAACACTAGTATGGCCACCCACTACTCCATCATCGGCGG
CATGATCATCGTCGTCTGCTGATGATCATCGGCTCCGAGGGCGGCGCGCTGCTG
GAGAAGAAGAACCGCACCTCCGCCGAGGCGGAGCACTACAACGTGCGCCGCTAC
CTGGCCGAGAACGGCCTGGGCCAGACCCCCCATGGGCTGGAACCTCTGGAAC
CACTTCGGCTGCGACATCAACGAGAACGTTCGTCGCGAGACGGCGGACGCCATG
GTGTCCACCGGCCTGGCCGCCCTGGGCTACCAGTACATCAACCTGGACGACTGCT
GGGCGGAGCTGAACCGCGACTCCGAGGGCAACATGGTGCCCAACGCCGCCGCGT
TCCCTCCGGCATCAAGGCGCTGGCGGACTACGTCCACTCCAAGGGCCTGAAGCT
GGGCGTCTACTCCGACGCGGGCAACCAGACCTGCTCCAAGCGCATGCCCGGCTC
CCTGGGCCACGAGGAGCAGGACGCGAAGACGTTTCGCCTCCTGGGGCGTCGACTA
CCTGAAGTACGACAACCTGCGAGAACCTGGGCATCTCCGTGAAGGAGCGCTACCC
CCCCATGGGCAAGGCGCTGCTGTCCTCCGGCCGCCCATCTTCTTCTCCATGTGC
GAGTGGGGCTGGGAGGACCCCCAGATCTGGGCCAAGTCCATCGGCAACTCCTGG
CGCACCACCGGCGACATCGAGGACAACCTGGAACCTCATGACCTCCATCGCCGAC
TCCAACGACAAGTGGGCCTCCTACGCCGGCCCCGGCGGCTGGAACGACCCCGAC
ATGCTGGAGGTCGGCAACGGCGGCATGACGACGGAGGAGTACCGCTCCCACTTC
TCCATCTGGGCCCTGGCCAAGGCGCCCCCTGCTGGTGCGGCTGCGACATCCGCGCGA
TGGACGACACGACGACGACGAGCTGATCTCCAACGCCGAGGTGATCGCGGTGAACC
AGGACAAGCTGGGCGTGCAGGGCAAGAAGGTCAAGTCCACGAACGACCTGGAG
GTCTGGGCCGCGCCCCCTGTCCGACAACAAGGTGGCGGTGATCCTGTGGAACCGCT
CCTCCTCCCGCGCCACCGTCAACCGCTCCTGGTCCGACATCGGCCTGCAGCAGGG
CACCACCGTCGACGCGCGCGACCTGTGGGAGCACTCCACGCAGTCCCTGGTGTCC
GGCGAGATCTCCGCCGAGATCGACTCCACGCCTGCAAGATGTACGTCTGACGC
CCCGCTCCTAACAATTGAAGCAGCAGCAGCTCGGATAGTATCGACACACTCTGG
ACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAAT
ATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCT
TTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTT
CCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTAT
CCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGTTTGGGC
TCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATG
CACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTCAGAATAGTA
TCGGGTGATGCGAAGTCAGAACCAGGCAGGGCCTGTCGCCTGAGGTGGCAACGA
TGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATCCCGTGATCTCCCCAC
CGACACCTATCCCCGCCCATCCCGGCCACCCCTTTCAGTCCCTCAGCATGCATT
GTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGAGGTCAGGCCACGGCGCACC
GAGTCTTCGCCCTTCGCGAGGTGATCACCTCCCCGGAACCAGCCACGCTCGCTG
CTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGGACAACCTCAGGCAAC
CACGCGCCTCAACAAGCATGGCCGCCGTGGCATCCAACCCGCTCGTGACGGTGGG
TGCGCAAGTGCCAGGGGCCTCGTCGTCACGGCGTGATCCTCGAGGGATGCGAT

CCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGATGAGCCCCTGTCTA
GA

SEQ ID NO: 123

5' primer *C. vulgaris* 3'UTR:downstream Clp sequence

5'-ACTGCAATGCTGATGCACGGGA-3'

SEQ ID NO: 124

3' primer *C. vulgaris* 3'UTR:downstream Clp sequence

5'-TCCAGGTCCTTTTCGCACT-3'

SEQ ID NO: 125

Coccomyxa C-169 THIC amino acid sequence

MTTNLAKLPLGGLSSRSSIAGAPLRVNSHSRERHLGAKTTAIAAPERLDYLDNAEEA
RLQQTDAFAELKALSSRQSVNRPQKGELSFQSPTFQDCFPGSEKCYREVEHDGKTL
KVPFRRVHLQDDNGHFDLYDTSGPQGVNPREGLPKIRSSWVEPREARGDKVQTQQY
YAKQGIITEEMAFCAARERMDPEFIRSEVARGRAIIPANKRHLELEPTVVGRNFLVKV
NANIGNSAVSSSIEEEVEKLQWSTIWGADTAMDLSGNNIHETREWVMRNSPVPVGT
VPIYQCLEKAGGIVENITWELFRETLEQAEQGVDFYFTIHAGVLLRYIPLTANRVTGIV
SRGGSIHAKLCLLDHTENFAYMHWDEILDICAQYDITLSIGDGLRPGCIADANDAAQF
AELKTQGELTRRAWAKDVQVMNEGPGHVPLHKIPENMEKQLDWCSEAPFYTLGPL
ATDIAPAYDHITS AIGAATIGALGTALLCYVTPKEHLGLPDRDDVKAGVIA YKIAAHA
ADLAKGHPYAQEWDNALSKARFEFRWYDQFHLSLDPVTARLFHDATLPQEPAKTA
HFCSMCGPKFCMQITQDLREYAKNHQMEEDEAIQTGMAEMSEQFKSIGAEVYLDE
AVLEHA

SEQ ID NO: 126

Coccomyxa C-169 THIC native transit peptide amino acid sequence

MTTNLAKLPLGGLSSRSSIAGAPLRVNSHSRERHLGAKTT

SEQ ID NO: 127

Relevant sequence of *Coccomyxa* C-169 THIC expression cassette

GCTCTTCGCCGCCGCGCACTCCTGCTCGAGCGCGCCCGCGCGTGCGCCGCCAGCGC
CTTGCCCTTTTCGCCGCGCTCGTGCGCGTCGCTGATGTCCATCACCAGGTCCATG
AGGTCTGCCTTGCGCCGGCTGAGCCACTGCTTCGTCCGGGCGGCCAAGAGGAGC
ATGAGGGAGGACTCCTGGTCCAGGGTCCTGACGTGGTCGCGGCTCTGGGAGCGG
GCCAGCATCATCTGGCTCTGCCGCACCGAGGCCCGCCTCCA ACTGGTCTCCAGCA
GCCGCACTCGCCGCCGACCCTGGCAGAGGAAGACAGGTGAGGGGGGTATGAATT
GTACAGAACAAACCAGAGCCTTGTCTAGGCAGAATCCCTACCAGTCATGGCTTTA
CCTGGATGACGGCCTGCGAACAGCTGTCCAGCGACCCTCGCTGCCGCCGCTTCTC
CCGCACGCTTCTTTCCAGCACCGTGATGGCGCGAGCCAGCGCCGCACGCTGGCGC
TGCGCTTCGCCGATCTGAGGACAGTCGGGGAACTCTGATCAGTCTAAACCCCTT
GCGCGTTAGTGTGTCATCCTTTGCAGACCGGTGAGAGCCGACTTGTGTGCGCC
ACCCCCACACCACCTCCTCCAGACCAATTCTGTACCTTTTGGCGAAGGCAT
CGGCCTCGGCCTGCAGAGAGGACAGCAGTGCCCAGCCGCTGGGGGTTGGCGGAT

GCACGCTCAGGTACCCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGG
GCTGCGAGACGGCTTCCCGCGCTGCATGCAACACCGATGATGCTTCGACCCCCC
GAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGT
TTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCAT
ATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCG
CACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGG
CGCGCCATGCTGCTGCAGGCCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAAGA
TCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCCTTCACCCC
CAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGG
CAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCC
CTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCC
CATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTG
GTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGC
GCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCT
CCTACAGCCTGGACGGCGGCTACACCTTCAACGAGTACCAGAAGAACCCCGTGC
TGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTC
CCAGAAGTGGATCATGACCGCGGCCAAGTCCAGGACTACAAGATCGAGATCTA
CTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTCCGCCAACGAGGG
CTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAG
GACCCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGG
CCGGCGGCTCCTTCAACCAAGTACTTCGTGCGCAGCTTCAACGGCACCCACTTCGA
GGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTG
CAGACCTTCTTCAACACCCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGG
GCCTCCAACCTGGGAGTACTCCGCCTTCGTGCCCACCAACCCCTGGCGCTCCTCCA
TGTCCCTCGTGCAGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGAC
GGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCC
CTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGT
CGACCTGTCCAACAGCACCCGGCACCCCTGGAGTTCGAGCTGGTGTACGCCGTCAAC
ACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGG
GCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCCCTC
CTTCTTCCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTA
CTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGACCT
GTCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTC
AACGACGGCGACGTGCTGTCCACCAACACCTACTTCATGACCACCGGGAACGCC
CTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAG
TTCCAGGTGCGCGAGGTCAAGTGACAATTGGCAGCAGCAGCTCGGATAGTATCG
ACACACTCTGGACGCTGGTCTGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTT
GACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGT
GTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCC
AGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGC
TGTCCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCC
TTGGTTTGGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGC
AATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGT
CTCGAACAGAGCGCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAG
CGCGGCATACACCACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTA
GCGAAGCGTCCGGTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATG
ATCGGTGGAGCTGATGGTTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTC
TTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCG
GCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGC
ATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCC

GATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCA
 CTACCACFTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCG
 CCTCTTCCTCTTCGTTTTAGTCACAACCCGCAAACACTAGTATGACCACCAACCT
 GGCTAAGCTGCCTCTCGGCGGCCTCTCTTCTCGGTCGTCCATCGCCGGGGCTCCTC
 TGC GCGTGAACAGCCACAGCCGCGAGCGCCACCTGGGCGCCAAGACCACCGCGA
 TCGCCGCTCCCGAGCGGCTGGACTACCTCGACAACGCGGAGGAGGCGCGGCTGC
 AGCAGACGGACGCCTTCGCCGAGCTGAAGGCGCTGTCGAGCCGCCAGTCCGTGA
 ACCGGCCCCAAAAGGGCGAGCTGTCTGTTCCGGCAGTCGCCGACCTTTCAGGACT
 GCTTCCCTGGCTCGGAGAAGTGCTACCGCGAGGTGGAGCACGATGGCAAGACGC
 TCAAGGTGCCTTTTTCGCCGGGTGCACCTGCAGGACGACAACGGCCATTTTCGACCT
 GTACGACACCTCTGGCCCCCAGGGCGTGAACCCCCGCGAGGGTCTGCCCAAGAT
 CCGCTCCTCCTGGGTGGAGCCGCGCGAGGCGCGGGGGCGACAAGGTGCAAACGCA
 GCAGTACTACGCCAAGCAAGGCATCATCACCGAGGAGATGGCCTTCTGCGCTGC
 GCGCGAGCGCATGGACCCCGAGTTCATCCGGTCTGAGGTGGCCCCGGGGTCGGGC
 CATCATCCCTGCGAACAAGCGCCATCTGGAGCTGGAGCCGACCGTCGTGGGCGG
 CAACTTTCTGGTCAAGGTGAACGCCAACATCGGCAACTCTGCGGTGAGCTCGAGC
 ATCGAGGAGGAGGTGGAGAAGCTGCAGTGGAGCACCATCTGGGGTGCCGACACG
 GCCATGGACCTGAGCACCGGCAACAACATCCACGAGACGCGCGAGTGGGTGATG
 CGCAACAGCCCGGTCCCTGTGGGTACGGTCCCTATCTACCAGTGCCTGGAGAAGG
 CGGGCGGCATCGTGGAGAACATCACCTGGGAGCTGTTCCGGGAGACGCTGATCG
 AGCAGGCCGAGCAGGGCGTGCAGTACTTTACCATCCACGCGGGGGTCTGCTCC
 GCTACATCCCCCTGACGGCGAACC GCGTCAACGGCATCGTCAGCCGCGGCGGCTC
 TATCCATGCCAAGCTCTGCCTGCTGGACCACACCGAGAACTTCGCCTACATGCAC
 TGGGACGAGATCCTGGACATCTGCGCCAGTACGACATCACCTGTGATCGGCG
 ACGGCCTGCGCCCCGGCTGCATCGCGGACGCCAACGACGCTGCCCAGTTTGCTGA
 GCTGAAGACCCAGGGCGAGCTGACCCGCGCGCCTGGGCGAAGGACGTGCAGGT
 GATGAACGAGGGTCCCGGCCATGTGCCGCTGCACAAGATCCCCGAGAACATGGA
 GAAGCAGCTGGACTGGTGCAGCGAGGCGCCCTTCTACACGCTCGGCCCCCTGGC
 GACGGACATCGCGCCTGCGTACGACCACATCACCTCCGCGATCGGCGCCGCCAC
 GATCGGCGCGCTGGGGACCGCGCTGCTGTGTTACGTGACCCCGAAGGAGCACCT
 GGGCCTCCCCGATCGCGACGACGTGAAGGCCGGGGTGATCGCCTACAAGATCGC
 GGCGCACGCGGCTGACCTGGCGAAGGGTCACCCCTACGCTCAGGAGTGGGATAA
 CGCCCTCTCTAAGGCGCGCTTCGAGTTCGGGTGGTACGACCAGTTCACCTGTGCG
 CTGGACCCCGTCAACCGCCCGCCTGTTCCATGACGCGACCCCTGCCGACGAGCCCG
 CCAAGACCGCCCACTTCTGCTCCATGTGCGGCCCAAGTTCTGCTCCATGCAGAT
 CACCCAGGACCTCCGCGAGTACGCCAAGAACCATCAAATGGAGGAGGACGAGGC
 GATCCAAACCGGCATGGCCGAGATGTCCGAGCAGTTCAAGAGCATCGGCGCGGA
 GGTCTACCTCGACGAGGCGGTGCTGGAGCACGCGTAAGCAGCAGCAGCTCGGAT
 AGTATCGACACACTCTGGACGCTGGTCTGTGATGGACTGTTGCCGCCACACTTG
 CTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTG
 ATCTTGTTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATAC
 CACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTA
 TCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCG
 CACAGCCTTGGTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAAACC
 AGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAA
 GCTTGAGCTCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCT
 CGATAACCTCCAAAGCCGCTCTAATTGTGGAGGGGGTTCGAATTTAAAAGCTTGG
 AATGTTGGTTCTGTGCTGTGGAACAAGCCAGACTTGTTGCTCACTGGGAAAAGG
 ACCATCAGCTCCAAAAAAGCTTGCCGCTCAAACCGCGTACCTCTGCTTTCGCGCAA
 TCTGCCCTGTTGAAATCGCCACCACATTCATATTGTGACGCTTGAGCAGTCTGTA

ATTGCCTCAGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGT
CGCGGGCGAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAAT
AGTTCATAACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTC
TGAGTGGCCACCCCCCGGCCCTGGTGCTTGCGGAGGGCAGGTCAACCGGCATGG
GGCTACCGAAATCCCCGACCGGATCCCACCACCCCCGCGATGGGAAGAATCTCT
CCCCGGGATGTGGGCCCACCACCAGCACAACTGCTGGCCCAGGCGAGCGTCAA
ACCATAACCACACAAATATCCTTGGCATCGGCCCTGAATTCTTCTGCCGCTCTGCT
ACCCGGTGCTTCTGTCCGAAGCAGGGGTGCTAGGGATCGCTCCGAGTCCGCAAA
CCCTTGTCGCGTGCGGGGCTTGTTTCGAGCTGAAGAGC

SEQ ID NO: 128

Codon-optimized *Coccomyxa* C-169 THIC sequence

ATGACCACCAACCTGGCTAAGCTGCCTCTCGGCGGCCCTCTCTTCTCGGTCTGCCA
TCGCCGGGGCTCCTCTGCGCGTGAACAGCCACAGCCGCGAGCGCCACCTGGGCG
CCAAGACCACCGCGATCGCCGCTCCCGAGCGGCTGGACTACCTCGACAACGCGG
AGGAGGCGCGGCTGCAGCAGACGGACGCCTTCGCCGAGCTGAAGGCGCTGTCGA
GCCGCCAGTCCGTGAACCGGCCCAAAAGGGCGAGCTGTCGTTCCGGCAGTTCG
CGACCTTTCAGGACTGCTTCCCTGGCTCGGAGAAGTGCTACCGCGAGGTGGAGCA
CGATGGCAAGACGCTCAAGGTGCCTTTTCGCCGGGTGCACCTGCAGGACGACAA
CGGCCATTTTCGACCTGTACGACACCTCTGGCCCCCAGGGCGTGAACCCCCGCGAG
GGTCTGCCCAAGATCCGCTCCTCTGGGTGGAGCCGCGCGAGGCGCGGGGCGAC
AAGGTGCAAACGCAGCAGTACTACGCCAAGCAAGGCATCATCACCGAGGAGATG
GCCTTCTGCGCTGCGCGCGAGCGCATGGACCCCGAGTTCATCCGGTCTGAGGTGG
CCCGGGGTTCGGGCCATCATCCCTGCGAACAAGCGCCATCTGGAGCTGGAGCCGA
CCGTCGTGGGCCGCAACTTTCTGGTCAAGGTGAACGCCAACATCGGCAACTCTGC
GGTCAGCTCGAGCATCGAGGAGGAGGTGGAGAAGCTGCAGTGGAGCACCATCTG
GGGTGCCGACACGGCCATGGACCTGAGCACCGGCAACAACATCCACGAGACGCG
CGAGTGGGTGATGCGCAACAGCCCCGGTCCCTGTGGGTACGGTCCCTATCTACCAG
TGCCTGGAGAAGGCGGGCGGCATCGTGGAGAACATCACCTGGGAGCTGTTCCGG
GAGACGCTGATCGAGCAGGCCGAGCAGGGCGTCTGACTACTTTACCATCCACGCG
GGGGTCCCTGCTCCGCTACATCCCCCTGACGGCGAACC CGGTACCCGGCATCGTCA
GCCGCGGCGGCTCTATCCATGCCAAGCTCTGCCTGCTGGACCACACCGAGAACTT
CGCCTACATGCACTGGGACGAGATCCTGGACATCTGCGCCCAGTACGACATCACC
CTGTCGATCGGCGACGGCCTGCGCCCCGGCTGCATCGCGGACGCCAACGACGCT
GCCCAGTTTGCTGAGCTGAAGACCCAGGGCGAGCTGACCCGCCGCGCCTGGGCG
AAGGACGTGCAGGTGATGAACGAGGGTCCCGGCCATGTGCCGCTGCACAAGATC
CCCGAGAACATGGAGAAGCAGCTGGACTGGTGCAGCGAGGCGCCCTTCTACACG
CTCGGCCCCCTGGCGACGGACATCGCGCCTGCGTACGACCACATCACCTCCGCGA
TCGGCGCCGCCACGATCGGCGCGCTGGGGACCGCGCTGCTGTGTTACGTGACCCC
GAAGGAGCACCTGGGCCTCCCCGATCGCGACGACGTGAAGGCCGGGGTGATCGC
CTACAAGATCGCGGCGCACGCGGCTGACCTGGCGAAGGGTCACCCCTACGCTCA
GGAGTGGGATAACGCCCTCTCTAAGGCGCGCTTCGAGTTCGGGTGGTACGACCA
GTTCCACCTGTCGCTGGACCCCGTCAACCGCCCGCCTGTTCCATGACGCGACCCTG
CCGAGGAGCCCGCCAAGACCGCCCACTTCTGCTCCATGTGCGGCCCAAGTTCT
GCTCCATGCAGATCACCCAGGACCTCCGCGAGTACGCCAAGAACCATCAAATGG
AGGAGGACGAGGCGATCCAAACCGGCATGGCCGAGATGTCCGAGCAGTTCAAGA
GCATCGGCGCGGAGGTCTACCTCGACGAGGCGGTGCTGGAGCACGCGTAA

SEQ ID NO: 129

Codon-optimized *S. cerevisiae* *suc2* sequence

GGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAA
GATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCACCTTCACC
CCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGAC
GCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACG
CCCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAG
CCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGG
TGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCA
GCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACAT
CTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCGT
GCTGGCCGCCAACTCCACCCAGTTCGCGGACCCGAAGGTCTTCTGGTACGAGCCC
TCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATC
TACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTTCGCCAACGAGG
GCTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCA
GGACCCAGCAAGTCTTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCG
GCCGGCGGCTCCTTCAACCAGTACTTCGTGCGCAGCTTCAACGGCACCCACTTCG
AGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCT
GCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTG
GGCCTCCAACCTGGGAGTACTCCGCCTTCGTGCCCCACCAACCCCTGGCGCTCCTCC
ATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGA
CGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCC
CCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACG
TCGACCTGTCCAACAGCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTCAA
CACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAG
GGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCT
CCTTCTTCTGACCGCGGGAACAGCAAGGTGAAGTTTCGTGAAGGAGAACCCCT
ACTTCACCAACCGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGAGAACGACC
TGTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTT
CAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGAACGC
CCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAA
GTTCCAGGTGCGCGAGGTCAAGTGACAATTG

SEQ ID NO: 130

Relevant sequence of *Coccomyxa* C-169 THIC expression construct with *C. protothecoides* transit peptide

GCGGCCGCGTGGACGAGGGCTACAACCCCGCCTACGGCGCGCGCCCGCTGCGCC
GCGCCATCATGCGCCTGCTGGAGGACGCGCTGGCCGAGCGCATGCTCGCCGGCG
ACGTCAAGGAGGGCGACTCGGTCATCATGGACGTCGATGGCGATGGCGCCATCA
GCGTCCTCAACGGAGACCGCACCCACACCACCATCGACTCCTCCCCGGCGG
GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCCGGGGCCTTGGGGCGCTGCTGGAGCGTGGTGAGAGCGCGGC
GGACGTGCCTTTTCTTCTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTTCG
GCCCGCATCCCCCACTGCCCCCATCATCTTGCTGTTGTCGTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTCTTTCTTTCGCTATGACAC
TTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAAC
ACCGATGATGCTTCGACCCCCGAAGTCTTTCGGGGCTGCATGGGCGCTCCGAT
GCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACAT

TATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACAC
AGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGT
TTCAGTCACAACCCGCAAACACTAGTATGGCCACCGCATCCACTTTCTCGGCGTT
CAATGCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCC
AGCGAGGCCCCCTCCCCGTGCGCGGGGCGCGCCGCGATCGCCGCTCCCGAGCGGCT
GGACTACCTCGACAACGCGGAGGAGGCGCGGCTGCAGCAGACGGACGCCTTCGC
CGAGCTGAAGGCGCTGTGAGCCGCCAGTCCGTGAACCGGCCCAAAAGGGCGA
GCTGTTCGTTCCGGCAGTCGCCGACCTTTCAGGACTGCTTCCCTGGCTCGGAGAAG
TGCTACCGCGAGGTGGAGCACGATGGCAAGACGCTCAAGGTGCCTTTTCGCCGG
GTGCACCTGCAGGACGACAACGGCCATTTTCGACCTGTACGACACCTCTGGCCCCC
AGGGCGTGAACCCCCGCGAGGGTCTGCCCAAGATCCGCTCCTCCTGGGTGGAGC
CGCGCGAGGCGCGGGGGCGACAAGGTGCAAACGCAGCAGTACTACGCCAAGCAA
GGCATCATCACCGAGGAGATGGCCTTCTGCGCTGCGCGCGAGCGCATGGACCCC
GAGTTCATCCGGTCTGAGGTGGCCCCGGGGTTCGGGCCATCATCCCTGCGAACAAG
CGCCATCTGGAGCTGGAGCCGACCGTCGTGGGGCCGCAACTTTCTGGTCAAGGTGA
ACGCCAACATCGGCAACTCTGCGGTGAGCTCGAGCATCGAGGAGGAGGTGGAGA
AGCTGCAGTGGAGCACCATCTGGGGTGCCGACACGGCCATGGACCTGAGCACCG
GCAACAACATCCACGAGACGCGCGAGTGGGTGATGCGCAACAGCCCGGTCCCTG
TGGGTACGGTCCCTATCTACCAGTGCCTGGAGAAGGCGGGCGGCATCGTGGAGA
ACATCACCTGGGAGCTGTTCCGGGAGACGCTGATCGAGCAGGCCGAGCAGGGCG
TCGACTACTTTACCATCCACGCGGGGGTCCCTGCTCCGCTACATCCCCCTGACGGC
GAACCGCGTCACCGGCATCGTCAGCCGCGGCGGCTCTATCCATGCCAAGCTCTGC
CTGCTGGACCACACCGAGAACTTCGCCTACATGCACTGGGACGAGATCCTGGAC
ATCTGCGCCCAGTACGACATCACCTGTGATCGGCGACGGCCTGCGCCCCGGCT
GCATCGCGGACGCCAACGACGCTGCCAGTTTGCTGAGCTGAAGACCCAGGGCG
AGCTGACCCGCGCGCCTGGGCGAAGGACGTCAGGTGATGAACGAGGGTCCCG
GCCATGTGCCGCTGCACAAGATCCCCGAGAACATGGAGAAGCAGCTGGACTGGT
GCAGCGAGGCGCCCTTCTACACGCTCGGCCCCCTGGCGACGGACATCGCGCCTGC
GTACGACCACATCACCTCCGCGATCGGCGCCGCCACGATCGGCGCGCTGGGGAC
CGCGCTGCTGTGTTACGTGACCCCGAAGGAGCACCTGGGCCTCCCCGATCGCGAC
GACGTGAAGGCCGGGGTGATCGCCTACAAGATCGCGGCGCACGCGGCTGACCTG
GCGAAGGGTCACCCCTACGCTCAGGAGTGGGATAACGCCCTCTCTAAGGCGCGC
TTCGAGTTCCGGTGGTACGACCAGTTCACCTGTGCTGGACCCCGTCACCGCCC
GCCTGTTCCATGACGCGACCTGCCGCGAGGAGCCCGCCAAGACCGCCCACTTCTG
CTCCATGTGCGGCCCAAGTTCTGCTCCATGCAGATCACCCAGGACCTCCGCGAG
TACGCCAAGAACCATCAAATGGAGGAGGACGAGGCGATCCAAACCGGCATGGCC
GAGATGTCCGAGCAGTTCAAGAGCATCGGCGCGGAGGTCTACCTCGACGAGGCG
GTGCTGGAGCACGCGTAACAATTGAAGCAGCAGCAGCTCGGATAGTATCGACAC
ACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACC
TGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGT
ACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGC
ATCCCCCTCCCTCGTTTCATATCGCTTGCAATCCCAACCGCAACTTATCTACGCTGT
CCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTG
GTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAA
TGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTCA
GAATAGTATCGGGTGATGCGAAGTCAGAACCAGGCAGGGCCTGTGCGCTGAGGT
GGCAACGATGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATCCCGTGAT
CTCCCCACCGACACCTATCCCCGCCCATCCCGGCCACCCCTTTCAGTCCCCTCAG
CATGCATTGTGCACCGCGACAAAGCATGTCTGCTCGTGCAGTGGTTCAGGCCACG
CGCACCGAGTCCTCGCCCTTCGCAGAGTGATCACCTCCCCGGAACCAGCCACG

CTCGCTGCTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGGACAACCTCA
GGCAACCACGCGCCTCACAAGCATGGCCGCCGTGGCATCCAACCCGCTCGTGAC
GGTGGGTGCGCAAGTGCCAGGGGCCCTCGTCGTACGGCGTGATCCTCGAGGGA
TGCGATCCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGATGAGCCCC
TGTCTAGA

SEQ ID NO: 131

Codon-optimized *Coccomyxa* C-169 THIC sequence with *C. protothecoides* transit peptide

ACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGCGACC
TGCGTTCGCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGAGGCCCTCCCCGTGCG
CGGGCGCGCCGCGATCGCCGCTCCCGAGCGGCTGGACTACCTCGACAACGCGGA
GGAGGCGCGGCTGCAGCAGACGGACGCCCTTCGCCGAGCTGAAGGCGCTGTCGAG
CCGCCAGTCCGTGAACCGGCCCAAAAGGGCGAGCTGTCGTTCCGGCAGTCGCC
GACCTTTCAGGACTGCTTCCCTGGCTCGGAGAAGTGCTACCGCGAGGTGGAGCAC
GATGGCAAGACGCTCAAGGTGCCTTTTCGCCGGGTGCACCTGCAGGACGACAAC
GGCCATTTTCGACCTGTACGACACCTCTGGCCCCCAGGGCGTGAACCCCCGCGAGG
GTCTGCCCAAGATCCGCTCCTCCTGGGTGGAGCCGCGCGAGGCGCGGGGCGACA
AGGTGCAAACGCGAGCAGTACTACGCCAAGCAAGGCATCATCACCGAGGAGATGG
CCTTCTGCGCTGCGCGCGAGCGCATGGACCCCGAGTTCATCCGGTCTGAGGTGGC
CCGGGGTTCGGGCCATCATCCCTGCGAACAAGCGCCATCTGGAGCTGGAGCCGAC
CGTCGTGGGCGCAACTTTCTGGTCAAGGTGAACGCCAACATCGGCAACTCTGCG
GTCAGCTCGAGCATCGAGGAGGAGGTGGAGAAGCTGCAGTGGAGCACCATCTGG
GGTGCCGACACGGCCATGGACCTGAGCACCGGCAACAACATCCACGAGACGCGC
GAGTGGGTGATGCGCAACAGCCCGGTCCCTGTGGGTACGGTCCCTATCTACCACT
GCCTGGAGAAGGCGGGCGGCATCGTGGAGAACATCACCTGGGAGCTGTTCCGGG
AGACGCTGATCGAGCAGGCCGAGCAGGGCGTCGACTACTTTACCATCCACGCGG
GGTCCCTGCTCCGCTACATCCCCCTGACGGCGAACC CGCTCACCGGCATCGTCAG
CCGCGGCGGCTCTATCCATGCCAAGCTCTGCCTGCTGGACCACACCGAGAACTTC
GCCTACATGCACTGGGACGAGATCCTGGACATCTGCGCCCAGTACGACATCACCC
TGTCGATCGGCGACGGCCTGCGCCCCGGCTGCATCGCGGACGCCAACGACGCTG
CCCAGTTTGCTGAGCTGAAGACCCAGGGCGAGCTGACCCGCCGCGCCTGGGCGA
AGGACGTGCAGGTGATGAACGAGGGTCCCGGCCATGTGCCGCTGCACAAGATCC
CCGAGAACATGGAGAAGCAGCTGGACTGGTGCAGCGAGGCGCCCTTCTACACGC
TCGGCCCCCTGGCGACGGACATCGCGCCTGCGTACGACCACATCACCTCCGCGAT
CGGCGCCGCCACGATCGGCGCGCTGGGGACCGCGCTGCTGTGTTACGTGACCCC
GAAGGAGCACCTGGGCCTCCCCGATCGCGACGACGTGAAGGCCGGGGTGATCGC
CTACAAGATCGCGGCGCACGCGGCTGACCTGGCGAAGGGTCACCCCTACGCTCA
GGAGTGGGATAACGCCCTCTCTAAGGCGCGCTTCGAGTTCGGGTGGTACGACCA
GTTCCACCTGTCGCTGGACCCCGTCACCGCCCGCCTGTTCCATGACGCGACCCCTG
CCGCAGGAGCCCCGCCAAGACCGCCCACTTCTGCTCCATGTGCGGCCCAAGTTCT
GCTCCATGCAGATCACCCAGGACCTCCGCGAGTACGCCAAGAACCATCAAATGG
AGGAGGACGAGGCGATCCAAACCGGCATGGCCGAGATGTCCGAGCAGTTCAAGA
GCATCGGCGCGGAGGTCTACCTCGACGAGGCGGTGCTGGAGCACGCGTAACAAT
TG

SEQ ID NO: 132

Chlorella protothecoides actin promoter/5'UTR

ACTAGAGAGTTTAGGTCCAGCGTCCGTGGGGGGGGACGGGCTGGGAGCTTGGGC
 CGGGAAGGGCAAGACGATGCAGTCCCTCTGGGGAGTCACAGCCGACTGTGTGTG
 TTGCACTGTGCGGCCCCGAGCACTCACACGCAAAATGCCTGGCCGACAGGCAGG
 CCCTGTCCAGTGCAACATCCACGGTCCCTCTCATCAGGCTCACCTTGCTCATTGAC
 ATAACGGAATGCGTACCGCTCTTTTCAGATCTGTCCATCCAGAGAGGGGAGCAGG
 CTCCCCACCGACGCTGTCAAACCTTGCTTCCTGCCCAACCGAAAAACATTATTGTTT
 GAGGGGGGGGGGGGGGGGGGGCAGATTGCATGGCGGGGATATCTCGTGAGGAACAT
 CACTGGGACACTGTGGAACACAGTGAGTGCAGTATGCAGAGCATGTATGCTAGG
 GGTCAGCGCAGGAAGGGGGCCCTTTCCAGTCTCCCATGCCACTGCACCGTATCCA
 CGACTCACCAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGTGGACACCAGTGTG
 TAGCCTCTGGACTCCAGGTATGCGTGCACCGCAAAGGCCAGCCGATCGTGCCGAT
 TCCTGGGGTGGAGGATATGAGTCAGCCAACTTGGGGCTCAGAGTGCACACTGGG
 GCACGATACGAAACAACATCTACACCGTGTCTCCATGCTGACACACCACAGCTT
 CGCTCCACCTGAATGTGGGCGCATGGGCCCCGAATCACAGCCAATGTCGCTGCTGC
 CATAATGTGATCCAGACCCTCTCCGCCAGATGCCGAGCGGATCGTGCGCGCTGA
 ATAGATTCTGTTTCGATCACTGTTTGGGTCTTTTCTTTTCGTCTCGGATGCGCG
 TCTCGAAACAGGCTGCGTCCGGGCTTTCCGATCCCTTTTGCTCCCTCCGTCACCATC
 CTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACCAGGGTTGGAGGGTATTA
 CCGCGTCAGGCCATTCCAGCCCGGATTCAATTCAAAGTCTGGGCCACCACCCTC
 CGCCGCTCTGTCTGATCACTCCACATTCTGTCATACACTACGTTCAAGTCCTGATC
 CAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAATCTCAAGGACCCACTCCA
 GCACAGCTGCTGGTTGACCCCGCCCTCGCAA

SEQ ID NO: 133

Chlorella protothecoides EF1A 3'UTR

ACGGAGCGTCTGTGCGGGAGGGAGTGTGCCGAGCGGGGAGTCCCGGTCTGTGCGA
 GGCCCGGCAGCTGACGCTGGCGAGCCGTACGCCCCGAGGGTCCCCCTCCCCTGC
 ACCCTCTTCCCCCTTCCCTCTGACGGCCGCGCCTGTTCTTGCATGTTACGCGAC

SEQ ID NO: 134

Relevant sequence of *A. thaliana* THIC expression cassette

GCTCTTCGCCGCCGCCACTCCTGCTCGAGCGCGCCCGCGCGTGGCGCCGCCAGCGC
 CTTGGCCTTTTCGCCGCGCTCGTGCAGCTCGCTGATGTCCATCACCAGGTCCATG
 AGGCTCTGCCTTGCGCCGGCTGAGCCACTGCTTCGTCCGGGCGGCCAAGAGGAGC
 ATGAGGGAGGACTCCTGGTCCAGGGTCTGACGTGGTTCGCGGCTCTGGGAGCGG
 GCCAGCATCATCTGGCTCTGCCGCACCGAGGCCCGCCTCCAACCTGGTCTCTCAGCA
 GCCGAGTCCGCCCGGACCCTGGCAGAGGAAGACAGGTGAGGGGGGTATGAATT
 GTACAGAACAACCACGAGCCTTGTCTAGGCAGAATCCCTACCAGTCATGGCTTTA
 CCTGGATGACGGCCTGCGAACAGCTGTCCAGCGACCCTCGCTGCCGCCGCTTCTC
 CCGCACGCTTCTTTCCAGCACCGTGATGGCGCGAGCCAGCGCCGCACGCTGGCGC
 TGCGCTTCGCCGATCTGAGGACAGTCGGGGAACTCTGATCAGTCTAAACCCCTT
 GCGCGTTAGTGTTGCCATCCTTTGCAGACCGGTGAGAGCCGACTTGTGTGCGCC
 ACCCCCCACACCACCTCCTCCAGACCAATTCTGTACCTTTTTGGCGAAGGCAT
 CGGCCTCGGCCTGCAGAGAGGACAGCAGTGCCAGCCGCTGGGGGTGGCGGAT
 GCACGCTCAGGTACCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGG
 GCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCC
 GAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGT
 TAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCAT

ATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCG
CACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACCTC
TAGAATATCAATGATCGAGCAGGACGGCCTCCACGCCGGCTCCCCCGCCGCCTG
GGTGGAGCGCCTGTTCCGGCTACGACTGGGCCCAGCAGACCATCGGCTGCTCCGA
CGCCGCCGTGTTCCGCCTGTCCGCCCAGGGCCGCCCGCTGCTGTTTCGTGAAGACC
GACCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCCTGTCCTGG
CTGGCCACCACCGGCGTGCCCTGCGCCGCCGTGCTGGACGTGGTGACCGAGGCC
GGCCGCGACTGGCTGCTGCTGGGCGAGGTGCCCGGCCAGGACCTGCTGTCCTCCC
ACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCGACGCCATGCGCCGCCTGC
ACACCCTGGACCCCGCCACCTGCCCTTTCGACCACCAGGCCAAGCACCCGCATCGA
GCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGGACCAGGACGACCTGGACGA
GGAGCACCAGGGCCTGGCCCCCGCCGAGCTGTTCCGCCGCCTGAAGGCCCGCAT
GCCCCGACGGCGAGGACCTGGTGGTGACCCACGGCGACGCCTGCCTGCCAACAT
CATGGTGGAGAACGGCCGCTTCTCCGGCTTCATCGACTGCGGCCGCCTGGGCGTG
GCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACATCGCCGAGGAGCTG
GGCGGCGAGTGGGCCGACCGCTTCTGGTGCTGTACGGCATCGCCGCCCGGACT
CCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGACAATTGGCAGC
AGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTG
CCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCC
TCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCT
ATTTGCGAATACCACCCCGAGCATCCCTTCCCTCGTTTCATATCGCTTGCATCCC
AACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCT
CACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCA
ACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACA
CAAATGGAGGATCCACTAGAGAGTTTAGGTCCAGCGTCCGTGGGGGGGGACGGG
CTGGGAGCTTGGGGCCGGGAAGGGCAAGACGATGCAGTCCCTCTGGGGAGTCACA
GCCGACTGTGTGTGTTGCACTGTGCGGCCCGCAGCACTCACACGCAAAATGCCTG
GCCGACAGGCAGGCCCTGTCCAGTGCAACATCCACGGTCCCTCTCATCAGGCTCA
CCTTGCTCATTGACATAACGGAATGCGTACCGCTCTTTCAGATCTGTCCATCCAG
AGAGGGGAGCAGGCTCCCCACCGACGCTGTCAAACCTTGCTTCCTGCCCAACCGA
AAACATTATTGTTTGAGGGGGGGGGGGGGGGGGGGCAGATTGCATGGCGGGATATC
TCGTGAGGAACATCACTGGGACACTGTGGAACACAGTGAGTGCAGTATGCAGAG
CATGTATGCTAGGGGTCAGCGCAGGAAGGGGGCCTTTCCAGTCTCCCATGCCAC
TGCACCGTATCCACGACTCACCAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGT
GGACACCAGTGTGTAGCCTCTGGACTCCAGGTATGCGTGACCCGCAAAGGCCAG
CCGATCGTGCCGATTCTGGGGTGGAGGATATGAGTCAGCCAACCTTGGGGGCTCA
GAGTGCACACTGGGGCACGATACGAAACAACATCTACACCGTGTCTCTCCATGCT
GACACACCACAGCTTCGCTCCACCTGAATGTGGGCGCATGGGCCCGAATCACAG
CCAATGTGCTGCTGCCATAATGTGATCCAGACCCTCTCCGCCAGATGCCGAGC
GGATCGTGGGCGCTGAATAGATTCTGTTCGATCACTGTTTGGGTCTTTCTTTT
TCGTCTCGGATGCGCGTCTCGAAACAGGCTGCGTCGGGCTTTTCGGATCCCTTTTG
CTCCCTCCGTCACCATCTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACC
AGGGTTGGAGGGTATTACCGCGTCAGGCCATTCCCAGCCCGGATTCAATTCAAAG
TCTGGGCCACCACCCTCCGCCGCTCTGTCTGATCACTCCACATTCTGTGCATACACT
ACGTTCAAAGTCTGATCCAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAAT
CTCAAGGACCCACTCCAGCACAGCTGCTGGTTGACCCCGCCCTCGCAAACTAGTA
TGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTTCG
CTCGGCGGGGCTCCGGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCGCGGGCG
CGCCGTCCAGGCCGCGGCCACCCGCTTCAAGAAGGAGACGACGACCACCCGCGC
CACGCTGACGTTTCGACCCCCCACGACCAACTCCGAGCGCGCCAAGCAGCGCAA

GCACACCATCGACCCCTCCTCCCCGACTTCCAGCCCATCCCCTCCTTCGAGGAG
TGCTTCCCCAAGTCCACGAAGGAGCACAAGGAGGTGGTGCACGAGGAGTCCGGC
CACGTCCTGAAGGTGCCCTTCCGCCGCGTGCACCTGTCCGGCGGCGAGCCCGCCT
TCGACAACTACGACACGTCCGGCCCCCAGAACGTCAACGCCACATCGGCCTGG
CGAAGCTGCGCAAGGAGTGGATCGACCGCCGCGAGAAGCTGGGCACGCCCCGCT
ACACGCAGATGTACTACGCGAAGCAGGGCATCATCACGGAGGAGATGCTGTACT
GCGCGACGCGCGAGAAGCTGGACCCCGAGTTCTGTCGCTCCGAGGTGCGCGGG
GCCGCGCCATCATCCCCTCCAACAAGAAGCACCTGGAGCTGGAGCCCATGATCG
TGGGCCGCAAGTTCTTGGTGAAGGTGAACGCGAACATCGGCAACTCCGCCGTGG
CCTCCTCCATCGAGGAGGAGGTCTACAAGGTGCAGTGGGGCCACCATGTGGGGCG
CCGACACCATCATGGACCTGTCCACGGGCCGCCACATCCACGAGACGCGCGAGT
GGATCCTGCGCAACTCCGCGGTCCCCGTGGGCACCGTCCCCATCTACCAGGCGCT
GGAGAAGGTGGACGGCATCGCGGAGAACCTGAACTGGGAGGTGTTCCGCGAGAC
GCTGATCGAGCAGGCCGAGCAGGGCGTGGACTACTTCACGATCCACGCGGGCGT
GCTGCTGCGCTACATCCCCCTGACCGCCAAGCGCCTGACGGGCATCGTGTCCCGC
GGCGGCTCCATCCACGCGAAGTGGTGCCTGGCCTACCACAAGGAGAACTTCGCC
TACGAGCACTGGGACGACATCCTGGACATCTGCAACCAGTACGACGTGCGCCCTGT
CCATCGGCGACGGCCTGCGCCCCGGCTCCATCTACGACGCCAACGACACGGCCC
AGTTTCGCCGAGCTGCTGACCCAGGGCGAGCTGACGCGCCGCGCGTGGGAGAAGG
ACGTGCAGGTGATGAACGAGGGCCCCGGCCACGTGCCCATGCACAAGATCCCCG
AGAACATGCAGAAGCAGCTGGAGTGGTGCACGAGGCGCCCTTCTACACCCTGG
GCCCCCTGACGACCGACATCGCGCCCGGCTACGACCACATCACCTCCGCCATCGG
CGCGGCCAACATCGGCGCCCTGGGCACCGCCCTGCTGTGCTACGTGACGCCCAA
GGAGCACCTGGGCCTGCCCAACCGCGACGACGTGAAGGCGGGCGTCAATCGCCTA
CAAGATCGCCGCCACGCGGCCGACCTGGCCAAGCAGCACCCCCACGCCCAGGC
GTGGGACGACGCGCTGTCCAAGGCGCGCTTCGAGTTCCGCTGGATGGACCAGTTC
GCGCTGTCCCTGGACCCCATGACGGCGATGTCTTCCACGACGAGACGCTGCCCG
CGGACGGCGCGAAGGTGCGCCACTTCTGCTCCATGTGCGGCCCCAAGTTCTGCTC
CATGAAGATCACGGAGGACATCCGCAAGTACGCCGAGGAGAACGGCTACGGCTC
CGCCGAGGAGGCCATCCGCCAGGGCATGGACGCCATGTCCGAGGAGTTCAACAT
CGCCAAGAAGACGATCTCCGGCGAGCAGCACGGCGAGGTCCGGCGGCGAGATCTA
CCTGCCCCGAGTCCTACGTCAAGGCCGCGCAGAAGTGATACGTACTCGAGACGGA
GCGTCGTGCGGGAGGGAGTGTGCCGAGCGGGGAGTCCCGGTCTGTGCGAGGCC
GGCAGCTGACGCTGGCGAGCCGTACGCCCCGAGGGTCCCCCTCCCCTGCACCCTC
TTCCCCTTCCCTCTGACGGCCGCGCCTGTTCTTGCATGTTTCAGCGACGAGCTCTTG
TTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGATAACCTCCAA
AGCCGCTCTAATTGTGGAGGGGGTTTGAATTTAAAAGCTTGGAATGTTGGTTCGT
GCGTCTGGAACAAGCCCAGACTTGTTGCTCACTGGGAAAAGGACCATCAGCTCC
AAAAAACTTGCCGCTCAAACCGCGTACCTCTGCTTTCGCGCAATCTGCCCTGTTG
AAATCGCCACCACATTTCATATTGTGACGCTTGAGCAGTCTGTAATTGCCTCAGAA
TGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCGCGGGCGAGG
ACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAGTTCATAACAG
TGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAGTGGCCACC
CCCCGGCCCTGGTGTGCTTGCGGAGGGCAGGTCAACCGGCATGGGGCTACCGAAAT
CCCCGACCGGATCCACACACCCCCGCGATGGGAAGAATCTCTCCCCGGGATGTG
GGCCACCAACAGCACAACCTGCTGGCCAGGCGAGCGTCAAACCATACACAC
AAATATCCTTGGCATCGGCCCTGAATTCCTTCTGCCGCTCTGCTACCCGGTGCTTC
TGTCCGAAGCAGGGGTTGCTAGGGATCGCTCCGAGTCCGCAAACCCCTGTGCGGT
GGCGGGGCTTGTTTCGAGCTTGAAGAGC

SEQ ID NO: 135

Codon-optimized *A. thaliana* THIC with *C. protothecoides* transit peptide

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGCGACCTGCGTC
GCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGAGGCCCCCTCCCCGTGCGCGGGC
GCGCCGTCCAGGCCGCGGCCACCCGCTTCAAGAAGGAGACGACGACCACCCGCG
CCACGCTGACGTTTCGACCCCCACGACCAACTCCGAGCGCGCCAAGCAGCGCA
AGCACACCATCGACCCCTCCTCCCCGACTTCCAGCCCATCCCCTCCTTCGAGGA
GTGCTTCCCCAAGTCCACGAAGGAGCACAAGGAGGTGGTGCACGAGGAGTCCGG
CCACGTCTGAAGGTGCCCTTCCGCGCGTGCACCTGTCCGGCGGCGAGCCCCGC
TTCGACAACCTACGACACGTCCGGCCCCCAGAACGTCAACGCCCACATCGGCCTG
GCGAAGCTGCGCAAGGAGTGGATCGACCGCCGCGAGAAGCTGGGCACGCCCCCG
TACACGCAGATGTACTACGCGAAGCAGGGCATCATCACGGAGGAGATGCTGTAC
TGCGCGACGCGCGAGAAGCTGGACCCCGAGTTTCGTCCGCTCCGAGGTGCGCGG
GGCCGCGCCATCATCCCCTCCAACAAGAAGCACCTGGAGCTGGAGCCCATGATC
GTGGGCGCGAAGTTCCTGGTGAAGGTGAACGCGAACATCGGCAACTCCGCGGTG
GCCTCCTCCATCGAGGAGGAGGTCTACAAGGTGCAGTGGGCCACCATGTGGGGC
GCCGACACCATCATGGACCTGTCCACGGGCGGCCACATCCACGAGACGCGCGAG
TGGATCCTGCGCAACTCCGCGGTCCCCGTGGGCACCGTCCCCATCTACCAGGCGC
TGGAGAAGGTGGACGGCATCGCGGAGAACCTGAACTGGGAGGTGTTCCGCGAGA
CGCTGATCGAGCAGGCCGAGCAGGGCGTGGACTACTTCACGATCCACGCGGGCG
TGCTGCTGCGCTACATCCCCCTGACCGCCAAGCGCCTGACGGGCATCGTGTCCCG
CGGCGGCTCCATCCACGCGAAGTGGTGCCTGGCCTACCACAAGGAGAACTTCGC
CTACGAGCACTGGGACGACATCCTGGACATCTGCAACCAGTACGACGTGCGCCT
GTCCATCGGCGACGGCCTGCGCCCCGGCTCCATCTACGACGCCAACGACACGGC
CCAGTTCGCGGAGCTGCTGACCCAGGGCGAGCTGACGCGCCGCGCGTGGGAGAA
GGACGTGCAGGTGATGAACGAGGGCCCCGGCCACGTGCCCATGCACAAGATCCC
CGAGAACATGCAGAAGCAGCTGGAGTGGTGAACGAGGCGCCCTTCTACACCCT
GGGCCCCCTGACGACCGACATCGCGCCCCGGCTACGACCACATCACCTCCGCCATC
GGCGCGGCCAACATCGGCGCCCTGGGCACCGCCCTGCTGTGCTACGTGACGCCC
AAGGAGCACCTGGGCCTGCCCAACCGCGACGACGTGAAGGCGGGCGTTCATCGCC
TACAAGATCGCCGCCACGCGGCCGACCTGGCCAAGCAGCACCCCCACGCCAG
GCGTGGGACGACGCGCTGTCCAAGGCGCGCTTCGAGTTCCGCTGGATGGACCAG
TTCGCGCTGTCCCTGGACCCCATGACGGCGATGTCTTCCACGACGAGACGCTGC
CCGCGGACGGCGCGAAGGTGCGCCACTTCTGCTCCATGTGCGGCCCCAAGTTCTG
CTCCATGAAGATCACGGAGGACATCCGCAAGTACGCCGAGGAGAACGGCTACGG
CTCCGCGGAGGAGGCCATCCGCCAGGGCATGGACGCCATGTCCGAGGAGTTCAA
CATCGCCAAGAAGACGATCTCCGGCGAGCAGCACGGCGAGGTGCGCGGCGAGAT
CTACCTGCCCCGAGTCTTACGTCAAGGCCGCGCAGAAGTGA

SEQ ID NO: 136

A. thaliana amino acid sequence (with native transit peptide sequence)

MAASVHCTLMSVVCNNKNHSARPKLPNSSLLPGFDVVVQAAATRFKKETTTTRATL
TFDPPTTNSERAKQRKHTIDPSSPDFQPIPSFEFCFKSTKEHKEVVHEESGHVLKVPF
RRVHLSGGEPAFDNYDTS GPQNVNAHIGLAKLRKEWIDRREKLGTPTRYTQMYYAKQ
GIITEEMLYCATREKLDPEFVRSEVARGRAIIPSNKKHLELEPMIVGRKFLVKVNAIG
NSAVASSIEEEVYKVQWATMWGADTIMDLSTGRHIIHETREWILRNSAVPVGTVPIYQ
ALEKVDGIAENLNWEVFRETLIEQAEQGVDFYFTIHAGVLLRYIPLTAKRLTGIVSRGG
SIHAKWCLAYHKENFAYEHWDDILDICNQYDVALSIGDGLRPGSIYDANDTAQFAEL

LTQGELTRRAWEKDVQVMNEGPGHVPMHKIPENMQKQLEWCNEAPFYTLGPLTTD
 IAPGYDHITS AIGAANIGALGTALLCYVTPKEHLGLPNRDDVKAGVIAKYIAAHAAD
 LAKQHPHAQAWDDALSKARFEFRWMDQFALS LDPM TAMS FHD ETL PADGAKVAH
 FCSMCGPKFCSMKITEDIRKYAEENGYGSAEEAIRQGM DAMSEEFNIAKKTISGEQH
 GEVGGEIYLPESYVKAQK

SEQ ID NO: 137

Relevant sequence of *Synechocystis* sp. *thiC* expression cassette

GCTCTTCGCCGCCGCGCCACTCCTGCTCGAGCGCGCCCGCGCGTGCGCCGCCAGCGC
 CTTGGCCTTTTCGCCGCGCTCGTGCGGTCGCTGATGTCCATCACCAGGTCCATG
 AGGTCTGCCTTGCGCCGGCTGAGCCACTGCTTCGTCCGGGCGGCCAAGAGGAGC
 ATGAGGGAGGACTCCTGGTCCAGGGTCCTGACGTGGTTCGCGGCTCTGGGAGCGG
 GCCAGCATCATCTGGCTCTGCCGCACCGAGGCCGCCTCCA ACTGGTCTCCAGCA
 GCCGCAGTCGCCGCCGACCCTGGCAGAGGAAGACAGGTGAGGGGGGTATGAATT
 GTACAGAACAAACCACGAGCCTTGCTAGGCAGAATCCCTACCAGTCATGGCTTTA
 CCTGGATGACGGCCTGCGAACAGCTGTCCAGCGACCCTCGCTGCCGCCGCTTCTC
 CCGCACGCTTCTTCCAGCACCGTGATGGCGCGAGCCAGCGCCGCACGCTGGCGC
 TGCGCTTCGCCGATCTGAGGACAGTCGGGGA ACTCTGATCAGTCTAAACCCCTT
 GCGCGTTAGTGTTGCCATCCTTTGCAGACCGGTGAGAGCCGACTTGTTGTGCGCC
 ACCCCCCACACCACCTCCTCCCAGACCAATTCTGTACCTTTTGGCGAAGGCAT
 CGGCCTCGGCCTGCAGAGAGGACAGCAGTGCCCAGCCGCTGGGGGTTGGCGGAT
 GCACGCTCAGGTACCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGG
 GCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCC
 GAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGT
 TTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCAT
 ATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCG
 CACTCCGCTAAGGGGGCGCCTCTTCTCTTCGTTTCAGTCACAACCCGCAAACTC
 TAGAATATCAATGATCGAGCAGGACGGCCTCCACGCCGGCTCCCCCGCCGCTG
 GGTGGAGCGCCTGTTCCGGCTACGACTGGGCCCAGCAGACCATCGGCTGCTCCGA
 CGCCGCCGTGTTCCGCCTGTCCGCCAGGGCCGCCCGCTGCTGTTGCTGAAGACC
 GACCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCCTGTCTGG
 CTGGCCACCACCGGCGTGCCCTGCGCCGCCGTGCTGGACGTGGTGACCGAGGCC
 GGCCGCGACTGGCTGCTGCTGGGCGAGGTGCCCGGCCAGGACCTGCTGTCTCTCC
 ACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCGACGCCATGCGCCGCCTGC
 ACACCCTGGACCCCGCCACCTGCCCTTCGACCACAGGCCAAGCACCGCATCGA
 GCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGGACCAGGACGACCTGGACGA
 GGAGCACCAGGGCCTGGCCCCCGCCGAGCTGTTGCCCCGCTGAAGGCCCGCAT
 GCCCGACGGCGAGGACCTGGTGGTGACCCACGGCGACGCCTGCCTGCCCAACAT
 CATGGTGGAGAACGGCCGCTTCTCCGGCTTCATCGACTGCGGCCGCCTGGGCGTG
 GCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACATCGCCGAGGAGCTG
 GGCGGCGAGTGGGCCGACCGCTTCTGGTGCTGTACGGCATCGCCGCCCGGACT
 CCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGACAATTGGCAGC
 AGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTG
 CCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCC
 TCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCT
 ATTTGCGAATAACACCCCGCATCCCTTCCCTCGTTTCATATCGCTTGCATCCC
 AACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCT
 CACTGCCCTTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCA
 ACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACA

CAAATGGAGGATCCACTAGAGAGTTTAGGTCCAGCGTCCGTGGGGGGGGGACGGG
 CTGGGAGCTTGGGCGGGGAAGGGCAAGACGATGCAGTCCCTCTGGGGAGTCACA
 GCCGACTGTGTGTGTTGCACTGTGCGGCCCGCAGCACTCACACGCAAAATGCCTG
 GCCGACAGGCAGGCCCTGTCCAGTGCAACATCCACGGTCCCTCTCATCAGGCTCA
 CCTTGCTCATTGACATAACGGAATGCGTACCGCTCTTTCAGATCTGTCCATCCAG
 AGAGGGGAGCAGGCTCCCCACCGACGCTGTCAAACCTTGCTTCCTGCCCCAACCGA
 AAACATTATTGTTTGAGGGGGGGGGGGGGGGGGGGCAGATTGCATGGCGGGGATATC
 TCGTGAGGAACATCACTGGGACACTGTGGAACACAGTGAGTGAGTATGCAGAG
 CATGTATGCTAGGGGTCAGCGCAGGAAGGGGGCCTTTCCCAGTCTCCCATGCCAC
 TGCACCGTATCCACGACTCACCAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGT
 GGACACCAGTGTGTAGCCTCTGGACTCCAGGTATGCGTGCACCGCAAAGGCCAG
 CCGATCGTGCCGATTCCCTGGGGTGGAGGATATGAGTCAGCCAACTTGGGGCTCA
 GAGTGCACACTGGGGCACGATACGAAACAACATCTACACCGTGTCTTCCATGCT
 GACACACCACAGCTTCGCTCCACCTGAATGTGGGCGCATGGGCCCCGAATCACAG
 CCAATGTCGCTGCTGCCATAATGTGATCCAGACCCTCTCCGCCCAGATGCCGAGC
 GGATCGTGGGCGCTGAATAGATTCCCTGTTTCGATCACTGTTTGGGTCCCTTTCTTT
 TCGTCTCGGATGCGCGTCTCGAAACAGGCTGCGTCCGGGCTTTTCGGATCCCTTTTG
 CTCCCTCCGTCACCATCCTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACC
 AGGGTTGGAGGGTATTACCGCGTCAGGCCATTCCCAGCCCGGATTCAATTCAAAG
 TCTGGGCCACCACCCTCCGCGGCTCTGTCTGATCACTCCACATTCTGTGCATACACT
 ACGTTCAAGTCCTGATCCAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAAT
 CTCAAGGACCCACTCCAGCACAGCTGCTGGTTGACCCCGCCCTCGCAAACCTAGTA
 TGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCG
 CTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCGCGGGCG
 CGCCATGCGCACCGCGTGGGTGCGCAAGCGCCAGGGCCAGACGAACGTCTCCCA
 GATGCACTACGCGCGCAAGGGCGTGATCACCGAGGAGATGGACTACGTGGCGAA
 GCGCGAGAACCTGCCCCGTGGAGCTGATCAAGGACGAGGTGGCGCGGGGCCGAT
 GATCATCCCCGCCAACATCAACCACACGAACCTGGAGCCCATGGCCATCGGCAT
 CGCCTCCAAGTGCAAGGTCAACGCGAACATCGGCGCGTCCCCCAACTCCTCCAAC
 ATCGACGAGGAGGTGGAGAAGCTGCTGCTGTCCGTGAAGTACGGCGCGGACACG
 GTGATGGACCTGTCCACCGGCGGCGGCGACCTGGACGTGATCCGCACCGCCATC
 ATCAACGCCTCCCCCGTGCCCATCGGCACGGTGCCCATCTACCAGGCCCTGGAGT
 CCGTGACGCGCTCCATCGAGAACCTGACGCCCCGACGACTTCTTGACATCATCGA
 GAAGCACGCCCAGCAGGGCGTGGACTACATGACGATCCACGCGGGGCTGTGAT
 CGAGTACCTGCCCCCTGGTGAAGTCCCGCATCACGGGCATCGTGTCCCGCGGCGGC
 GGCATCATCGCCAAGTGGATGCTGCACCACCACAAGCAGAACCCCCCTGTACACG
 CACTTCGACGAGATCATCGAGATCTTCAAGAAGTACGACGTGTCCTTCTCCCTGG
 GCGACTCCCTGCGCCCCGGCTGCACCCACGACGCCTCCGACGACGCCAGCTGTC
 CGAGCTGAAGACGCTGGGCCAGCTGACGCGCCGCGCCTGGGAGCACGACGTCCA
 GGTCATGGTGGAGGGCCCCGGCCACGTCCCCATCGACCAGATCGAGTTCAACGT
 CAAGAAGCAGATGGAGGAGTGCTCCGAGGCCCCCTTCTACGTGCTGGGCCCCCT
 GGTCACCGACATCGCGCCCGGCTACGACCACATCACCTCCGCGATCGGCGCCGCC
 ATCGCGGGCTGGCACGGCACCGCGATGCTGTGCTACGTGACGCCCAAGGAGCAC
 CTGGGCCTGCCCAACGCGGAGGACGTGCGCAACGGCCTGATCGCCTACAAGATC
 GCCGCGCACGCCGCGGACATCGCCCGCCACCGCCCCGGCGCGCGGACCGCGAC
 GACGAGCTGTCCAAGGCGCGCTACAACTTCGACTGGAACCGCCAGTTCGAGCTG
 TCCCTGGACCCCGAGCGCGCGAAGGAGTACCACGACGAGACGCTGCCCGCGGAC
 ATCTACAAGACCGCGGAGTTCTGCTCCATGTGCGGGCCCCAAGTTCTGCCCATGC
 AGACCAAGGTGGACGCCGAGATGCTGGAGGAGCTGGAGGTCTTCCCTGGCCAAGG
 ACAAGGAGATGGTGTCCAGCGCTGATACGTACTCGAGACGGAGCGTCTGTGCGG

GAGGGAGTGTGCCGAGCGGGGAGTCCCGGTCTGTGCGAGGCCCGGCAGCTGACG
CTGGCGAGCCGTACGCCCCGAGGGTCCCCCTCCCCTGCACCCTCTTCCCCCTTCCCT
CTGACGGCCGCGCCTGTTCTTGCATGTTTCAGCGACGAGCTCTTGTTCAGAAAG
GAGTTGCTCCTTGAGCCTTTTCATTCTCAGCCTCGATAACCTCCAAAGCCGCTCTAA
TTGTGGAGGGGGTTCGAATTTAAAAGCTTGGAATGTTGGTTCGTGCGTCTGGAAC
AAGCCCAGACTTGTGCTCACTGGGAAAAGGACCATCAGCTCCAAAAAACTTGC
CGCTCAAACCGCGTACCTCTGCTTTCGCGCAATCTGCCCTGTTGAAATCGCCACC
ACATTCATATTGTGACGCTTGAGCAGTCTGTAATTGCCTCAGAATGTGGAATCAT
CTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCGCGGGCGAGGACACCCGCCAC
TCGTACAGCAGACCATTATGCTACCTCACAATAGTTCATAACAGTGACCATATTT
CTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAGTGGCCACCCCCCGGCCCTG
GTGCTTGCGGAGGGCAGGTCAACCGGCATGGGGCTACCGAAATCCCCGACCGGA
TCCCACCACCCCCGCGATGGGAAGAATCTCTCCCCGGGATGTGGGCCCCACCACCA
GCACAACCTGCTGGCCCAGGCGAGCGTCAAACCATAACCACACAAATATCCTTGG
CATCGGCCCTGAATTCCCTTCTGCCGCTCTGCTACCCGGTGCTTCTGTCCGAAGCAG
GGGTTGCTAGGGATCGCTCCGAGTCCGCAAACCCTTGTCGCGTGCGGGGCTTGT
TCGAGCTTGAAGAGCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATC
ATGGTCATAGCTGTTTCCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAAC
ATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCTAATGAGTGAGCTAA
CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTGCT
GCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG
GGCGCTCTTC

SEQ ID NO: 138

Codon-optimized *Synechosystis sp. thiC* sequence with *C. protothecoides* transit peptide

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGGCAGCTGCGTC
GCTCGGGCGGGCTCCGGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCGCGGGC
GCGCCATGCGCACCGCGTGGGTGCGCAAGCGCCAGGGCCAGACGAACGTCTCCC
AGATGCAC'TACGCGCGCAAGGGCGTGATCACCGAGGAGATGGACTACGTGGCGA
AGCGCGAGAACCTGCCCGTGGAGCTGATCAAGGACGAGGTGGCGCGGGGCCGCA
TGATCATCCCCGCCAACATCAACCACACGAACCTGGAGCCCATGGCCATCGGCAT
CGCCTCCAAGTGCAAGGTCAACGCGAACATCGGCGCGTCCCCCAACTCTCCAAC
ATCGACGAGGAGGTGGAGAAGCTGCTGCTGTCCGTGAAGTACGGCGCGGACACG
GTGATGGACCTGTCCACCGGCGGCGGCGACCTGGACGTGATCCGCACCGCCATC
ATCAACGCCTCCCCCGTGCCCATCGGCACGGTGCCCATCTACCAGGCCCTGGAGT
CCGTGCACGGCTCCATCGAGAACCTGACGCCCCGACGACTTCTTGCACATCATCGA
GAAGCACGCCCAGCAGGGCGTGGACTACATGACGATCCACGCGGGCCTGCTGAT
CGAGTACCTGCCCTGGTGAAGTCCCGCATCACGGGCATCGTGTCCCGCGGGCGGC
GGCATCATCGCCAAGTGGATGCTGCACCACCACAAGCAGAACCCCTGTACACG
CACTTCGACGAGATCATCGAGATCTTCAAGAAGTACGACGTGTCTTCTCCCTGG
GCGACTCCCTGCGCCCCGGCTGCACCCACGACGCCTCCGACGACGCCAGCTGTC
CGAGCTGAAGACGCTGGGCCAGCTGACGCGCCGCGCCTGGGAGCACGACGTCCA
GGTCATGGTGGAGGGCCCCGGCCACGTCCCCATCGACCAGATCGAGTTCAACGT
CAAGAAGCAGATGGAGGAGTGCTCCGAGGCCCTTCTACGTGCTGGGCCCCCT
GGTCACCGACATCGCGCCCCGGCTACGACCACATCACCTCCGCGATCGGCGCCGCC
ATCGCGGGCTGGCACGGCACCGCGATGCTGTGCTACGTGACGCCAAGGAGCAC
CTGGGCCTGCCCAACGCGGAGGACGTGCGCAACGGCCTGATCGCCTACAAGATC
GCCGCGCACGCCGCGGACATCGCCCGCCACCGCCCCGGCGCGCGGACCGCGAC
GACGAGCTGTCCAAGGCGCGCTACAACCTTCGACTGGAACCGCCAGTTCGAGCTG

TCCCTGGACCCCGAGCGCGCGAAGGAGTACCACGACGAGACGCTGCCCCGCGGAC
ATCTACAAGACCGCGGAGTTCTGCTCCATGTGCGGCCCAAGTTCTGCCCCATGC
AGACCAAGGTGGACGCCGAGATGCTGGAGGAGCTGGAGGTCTTCCTGGCCAAGG
ACAAGGAGATGGTGTCCCAGCGCTGA

SEQ ID NO: 139

Synechosystis sp. thiC amino acid sequence (with native transit peptide sequence)

MRTAWVAKRQGQTNVSQMHYARKGVITEEMDYVAKRENLPVELIKDEVARGRMII
PANINHTNLEPMAIGIASKCKVNANIGASPNSSNIDEEVEKLLLSVKYGADTVMDLST
GGGDL DVIRTAIINASPVPIGTVPYQALESVHGSINLTPDDFLHIEKHAQQGV DYM
TIHAGLLIEYLPLVKSRTGIVSRGGGHIKWM LHHHKQNPLYTHFDEIIEIFKKYDV SF
SLGDSL RPGCTHDASDDAQLSELKTLGQLTRRAWEHDVQVMVEGPGHVPIDQIEFN
VKKQMEECSEAPFYVLGPLVTDIAPGYDHITS AIGAAIAGWHGTAMLCYVTPKEHLG
LPNAEDVRNGLIAYKIAAHAADIARHRPGARDRDELSKARYNFDWNRQFELS LDP
ERAKEYHDETLPADIYKTAEFCSMCGPKFCPMQTKVDAEMLEEEVFLAKDKEMVS
QR

SEQ ID NO: 140

Codon-optimized neoR gene

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAAC

SEQ ID NO: 141

5' THIC *Coccomyxa* C-169 confirmation primer

5'-GGGTGATCGCCTACAAGA-3'

SEQ ID NO: 142

3' THIC confirmation primer

5'-ACGTCGCGACCCATGCTTCC-3'

SEQ ID NO: 143

5' THIC *A. thaliana* confirmation primer

5'- GCGTCATCGCCTACAAGA-3'

SEQ ID NO: 144

5' thiC *Synechosystis sp.* confirmation primer

5'-CGATGCTGTGCTACGTGA-3'

SEQ ID NO: 145

C callophylla TE; UTEX 1450 SAD transit peptide underlined

[illegible]

EVHVDVTHSPQKIEIFKSLDNWAEENILVHLKPVBEKCWQPDFLPDPASDGFDEQVRE
 LREKAKEIPDDYFVVLVGDMITTEBALPTYQTMNLNDGVRDETGAAPTWSAIIWRAWTAE
 ENRHGDLNKKYLYLSGRVDMRQIEKTIQYILGSGMDPRTENSPYLGFTYSFQERATFIS
 HGNARQAKBHGDIKLAQICGTVAADEKRHEHTAYTKVIEKLFEPDGTVALAFADMRRKK
 ISMPAHLMYDGRDNLFDHFSAVAQRGLVYTAQDYADILBEFLVGRWKVADKLTGLSAEQK
 AQDVVCRTPRIRRLFEBAQGRAKKAPTMFSSWIFDRQVVKLMDYKKDHDGDKYKDHIDYKD
 DDDK

SEQ ID NO: 152

[illegible]

SEQ ID NO: 151
Chlamydomonas RBCS2 promoter

[illegible]

ctcatgttaccgatgctattcggagaacggcaactaagctgccgggttgaacacggatgatctcgcggagggtagcatgttgattg
taacgatgacagagcgttgctgctgtgatcaccgcgg

SEQ ID NO: 153

CAACCACGTCTTCAAAGCAA

SEQ ID NO: 154

AGCAATCGCGCATATGAAAT

SEQ ID NO: 155

Endogenous *Chlorella protothecoides* actin promoter and 5'UTR

GAATTCGAGTTTAGGTCCAGCGTCCGTGGGGGGGGACGGGCTGGGAGCTTGGGC
CGGGAAGGGCAAGACGATGCAGTCCCTCTGGGGAGTCACAGCCGACTGTGTGTG
TTGCACTGTGCGGCCCGCAGCACTCACACGCAAAATGCCTGGCCGACAGGCAGG
CCCTGTCCAGTGCAACATCCACGGTCCCTCTCATCAGGCTCACCTTGCTCATTGAC
ATAACGGAATGCGTACCGCTCTTTCAGATCTGTCCATCCAGAGAGGGGAGCAGG
CTCCCCACCGACGCTGTCAAACCTTGCTTCCTGCCCAACCGAAAACATTATTGTTT
GAGGGGGGGGGGGGGGGGGGGCAGATTGCATGGCGGGATATCTCGTGAGGAACAT
CACTGGGACACTGTGGAACACAGTGAGTGCAGTATGCAGAGCATGTATGCTAGG
GGTCAGCGCAGGAAGGGGGCCTTTCCAGTCTCCCATGCCACTGCACCGTATCCA
CGACTCACCAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGTGGACACCAGTGTG
TAGCCTCTGGACTCCAGGTATGCGTGCACCGCAAAGGCCAGCCGATCGTGCCGAT
TCCTGGGGTGGAGGATATGAGTCAGCCAACTTGGGGCTCAGAGTGCACACTGGG
GCACGATACGAAACAACATCTACACCGTGTCTCCATGCTGACACACCACAGCTT
CGCTCCACCTGAATGTGGGCGCATGGGGCCGAATCACAGCCAATGTTCGCTGCTGC
CATAATGTGATCCAGACCCTCTCCGCCAGATGCCGAGCGGATCGTGGGCGCTGA
ATAGATTCCCTGTTTCGATCACTGTTTGGGTCCTTTCCCTTTTCGTCTCGGATGCGCG
TCTCGAAACAGGCTGCGTCGGGCTTTCGGATCCCTTTTGCTCCCTCCGTCACCATC
CTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACCAGGGTTGGAGGGTATTA
CCGCGTCAGGCCATTCCCAGCCCGGATTCAATTCAAAGTCTGGGCCACCACCCTC
CGCCGCTCTGTCTGATCACTCCACATTCGTGCATACACTACGTTCAAGTCTGATC
CAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAATCTCAAGGACCCACTCCA
GCACAGCTGCTGGTTGACCCCGCCCTCGCAACTCCCTACCATGTCGCTGtaggtcca
gggatcttggccatgcacacaggaccccgttgtgggggtccccgggtgcattgctgtcgtgtgcaggcgccgggtgtggggcctgggc
cccggggagctcaactcctccccatatgcctgcgctccctccaccaccgcgacctggcccccttgcagAGGAAGGCGA
AGTCAGCGCCATCGTGTGCGATAATGGATCCGG

SEQ ID NO: 156

Endogenous *Chlorella protothecoides* EF1a promoter and 5'UTR

GAATTCGCCCTTGAGTTTAGGTCCAGCGTCCGTGGGGGGGGCGTGAGACTCCCCC
CTGACCTTCGTATGGCAGGGACTCCTACTTGCCAAGTAATCAGTTGACAATGCCA
CTTCAATGCTCGTTGTGGTACACTGACGCGGGTCTAACATACTGGGAAGCATGAA
TTGCCGACATGGACTCAGTTGGAGACAGTAACAGCTCTTTGTGTTCTATCTTCAG
GAACACATTTGGCAGCGCACCCATACAGTGGCGCACACGCAGCTGTACCTGATG
TGGCTCTATTCCCACATGTTTCAACTTGATCCAAAAGTCACTCAGACTCTCAGCA
GCTAGACTTGATCGCATCTTTGGCCATGAAGATGCTTGCGCAACTCTAGGAATGG

GACGAGAAAAGAGCCTGCTCTGATCGGATATTTCCATTCTCTGGATGGGACTGAG
 ATGATTCTGAAGAAATGCTGCTCGACTTATTTGGAAGAACAGCACCTGACGCATG
 CTTTGAGGCTGCTGTGGCTGGGATGTGCTGTATTTGTGTCAGCATTGAGCATCTACG
 GGTAGATGGCCATAACACGCGCTGCCTATCATGCGGTGGGTTGTGTGGAAAAC
 GTACAATGGACAGAAATCAATCCCATTCGAGCCTAGCGTGCAGCCATGCGCTC
 CCTCTGTAGCCCCGCTCCAAGACAAAGCCAGCCAATGCCAGAACCACATAGAG
 AGGGTATCTTCCTAATGACCTCGCCCATCATTTCTCCAAATTAACATAATGCCT
 TGATTGTGGAGTTGGCTTTGGCTTGCAGCTGCTCGCGCTGGCACTTTTGTAGGCA
 GCACAGGGTATGCCAGCGCCGAACCTTTGTGCCCTTGAGCAGGCCACAAGGGCAC
 AAGACTACACCATGCAGCTGGTATACTTGGAAGTATGATACCATTCTTACCAAGCAA
 GGCACAGCACAGCCTGCACCGACTCACTTTGCTTGAGCGGGGCACAGCGCCGCG
 ACTGATCCTGCGAGCTGTGGGGAGTTCCGACTGTTCTGGACCTCGGTCTCTGAAA
 GATGTGTACGATGGGATCAAGTCATTCAAGTATGCTCTTCACATGAGCAATCGGG
 GGAGACACGGTGGCCCTAAAGGTGTTTCATCTGATTCAAGTGTAGTGGGGGGGTG
 CTGTTTGTCCCGGGGCGCCCCCGCTCCCCGACCCCGGAGAAGGGCCCCAGAGG
 ACTCGGCCGCCACAGAGGAATAACCGGGCGTGGCTCGGCCCTGCGCCTCCCTCT
 TTCAATATTTACCTGGTGTTCAGTGCACGGACACGTAAAGAACTAGATACAATG
 GCCGAGGGAAAGACGgtgagagcttgccgttggtggaccgggcagcatcagaaactctctccccgccgccttgaaa
 ctcactgtaactccctctcttccccctcgcagCATCTGTCTATCGTTATCgtgagtgaagggaactgccatgtgtcgggt
 cgttgaccacggctcggtcgggcgtgctgcccgcgtcgcgaacgttcctgcaaagccgcgcagccgtcccttttctgccgcg
 cccacccccctcgtcccccttcaatcacaccgcagTGC GGACATGTCGATTCCGGCAAGTCCACC

SEQ ID NO: 157

Endogenous *Chlorella protothecoides* beta-tubulin promoter (isoform A)

GAATTCCCTGCAGGAAGAAGGCCGGCAGCAGCTGGTACTTGTCTTCACCTCCTT
 GATCGGCTGGGTGAGCTTGGCCGGGTGCGAGTCGTCGATGCCGGCATCGCCAG
 CACGCTGTGCGGGGAGCCGGCATCGACAACCTTGGCACTGCTCACCTTGGTCACC
 GGCATGGGGTTCATGGCGCTGCAGACCAGCGGCCTGTCAGCATGCTGCAGGCATC
 TGTGTTTTGTAGTAGATACTTTCTGATGCATCACACACGTTTGGAAGGTCCCCA
 AGCCCCCTTCAACAGTCTCGACATATGACACTCGCGCCCTCTTCCTCGTCCCGTGG
 CCTGATGAGGGTACGCAGGTACCGCAGCTGCGCCCCGTCCCGCCAGTTGCCCTGG
 CCCC GCCGGGCCCAATCTGTTTCATTGCCGCTCCCTGGCAGCCGTGAACCTCACAC
 TACCGCTCTCTGTGACCTTCAGCACAGCAGGAATCGCCATTTACCGGCGGTGCT
 TGCTGCGGAGCCTCAGCTGATCTCGCCTGCGAGACCCACAGTTTGAATTTGCGG
 TCCCCACACAACCTCTGACGCC

SEQ ID NO: 158

Endogenous *Chlorella protothecoides* beta-tubulin promoter (isoform B)

GAATTCCCTCAGGAAGAAGGCCGGCAGCAGCTGGTACTTGTCTTCACCTCCTTG
 ATCGGCTGGGTGAGCTTCGCAGGATCGCAGTCGTCGATGCCGGCATCGCCAGC
 ACGCTGTGCGGGGAGCCGGCATCNACAACCTTGGCACTGCTCCCTTGGTCACCG
 GCATGGGGTTCATGGCGCTGCAGCCAGCGGCCTGTCAGCATGCTGCAGGCATCT
 GTGTATTGTAGTAGGTACTTCCTGATGCATCAACACACGTTTGGAAGCTCCCCAA
 GCCCCCTTCAACAGTCTCGACGTATGACACTCGCGCCCTCTTCCTCGCCCCGTGGC
 CTGATGAGGGTACGCAGGTACCACAGCTGCGCCCCGTCCCGCCAGTTGCCCTGGC
 CCGGCCGGGCCCAATCTGTTTCATTGCCGCTCCCTGGTAGCCGTGAACCTCACATTA
 CCGCTCTCTGTGACCTTCAGCACAGCAGGAATCGCCATTTACCGGCGGTGCTTG

CTGCGGAGCCTCAGCTGATCTCGCCTGCGAGACCCACAGTTTGAATTTGCGGTC
CCCACACAACCTCTGACGCC

SEQ ID NO: 159

Acyl ACP desaturase

atcaaagcatagattcacattgttgcatcagagcaatcatcgcgaggacgaacatcgctcaccaagcacgtactgggcatcc
ggaggcctccgcaaattcctgcaacaggactcgctgatcagttcgcccaaggctctacgacgctccctatcggcgctagacttcaacac
atatctactgtcacagcctcggcATGCATCAGGCCTCAGTCTCCACCATGAAGACCATCCAGTC
TCGGCACGCCGGTCCCATCGGACATGTGCACTCGGGTCGCCGATCGGCGGGGGCG
CGCGGGATCCCGCATGGCGACCCCGTGGCCGCAGCTACCGTCGCAGCCCCCTCGC
TCGGCCCTCAACCTCTCCCCACCATCATTGACAGGAGGTGCTCCACTCCGCCA
GCGCCCAGCAACTAGACTGCGTGGCCTCCCTGGCGCCCGTCTTCGAGTCCCAGAT
CCTCCCCCTCCTGACGCCCGTGGACGAGATGTGGCAGCCCACCGACTTCCTCCCC
GCCTCGAACTCGGAGGCATTCTTCGACCAGATCGGCGACCTGCGGGCGCGATCG
GCGGCCATCCCCGACGACCTGCTGGTCTGCCTGGTGGGGGACATGATCACGGAG
GAGGCCCTGCCACCTACATGGCCATGCTGAACACCCTGGACGTCGTGCGCGATG
AGACAGGGCACAGCCAGCACCCCTACGCCAAGTGGACCAGGGCTTGATCGCGG
AGGAGAACCGCCATGGCGACCTGCTGAACAAGTACATGTGGCTGACGGGGCGGG
TGGGACATGCTGGCGGTGGAGCGCACCATCCAGCCATGCTGGCGGTGGAGCGCA
CCATCCAGCGCCTCATCTCATCGGGCATGGACCCGGGCACGGAGAACCACCCCT
ACCACGCCCTTTGTGTTACACAGCTTCCAGGAGCGCGCCACCAAGCTGAGCCACGG
CTCCACCGCCCGCCTGGCGGTGCGCGCCGGGGACGAGGCCCTGGCCAAGATCTG
CGGGACCATTTGCGCGGGACGAGTCGCGCCACGAGGCGGCGTACACGCGGACCAT
GGATGCCATCTTCCAGCGCGACCCAGCGGGGCCATGGTGGCGTTTGCGCACATG
ATGATGCGCAAGATCACCATGCCCCGCCACCTCATGGACGACGGCCAGCACGGC
GCGCGCAACGGGGGGGCGCAACTTGTTCGACGACTTTGCGGCAGTGGCGGAGCG
GGCAGGGGTGTACACCGCCGGCGACTACATCGGCATCCTGCGCCACCTCATCCG
GCGCTGGGACGTGGAGGG

SEQ ID NO: 160

Acyl ACP desaturase

MHQASVSTMKTIQSRHAGPIGHVQSGRRSAGRAGSRMATPVAAATVAAPRSALNLS
PTIIRQEVLSASAQQLDCVASLAPVFESQILPLLTPVDEMWPQPTDFLPASNSEAFFDQ
IGDLRARSAAIPDDLVLVCLVGDMMITEALPTYMAMLNLTLDVVRDETGHSHQHPYAKW
TRAWIAEENRHGDLLNKYMWLTGRVGHAGGGAHHPAMLAVERTIQRLISSGMDPG
TENHPYHAFVFTSFQERATKLSHGSTARLAVAAGDEALAKICGTIARDESHEAAYT
RTMDAIFQRDPSGAMVAFAHMMMRKITMPAHLMDDGQHGARNNGAQLVRRRLCGS
GGAGRGVHRRRLHRHPAPHPALGRGG

SEQ ID NO: 161

Geranyl geranyl diphosphate synthase

attatacatcgcatcgctcaggtttcacgatctgcatgctatctatgggactgtgactccgccggccaggtgtggtgcgcgagaalc
ctcccgcctcctgctctcatttccctgacgggagtcgccgctgagcaccggcgcgatcATGGGCGTCGGCACACT
CCAAACCCCATATACATGTGGTCTGTGATTACGCATAGCGCACGGTATGTCCCG
CGACGCGCGGCTCGAAGCCGTGGCCATCCGACGCGCTGCACGGCCGAGGTGAGG
GCACGCCCCCTCCGCCAATGGCGCGCAGCCCATGACCGCCTTCGACTTCCGGCAGT
ACATGCAGCAGCGCGCCGCGCTGGTGGACGCAGCGCTGGACCTGGCAGTGCCGC

TGCAGTACCCCGAGAAGATCAACGAGGCCATGCGGTACAGCCTGCTGGCCGGGG
GCAAGCGCGTGCGCCCCGCGCTCTGCCTCGCTGCGAGCTCGTGGGCGGCCC
CCTGGAGGGCGGCCATGCCCCGCGCCTGCGCCATGGAGATGATCCACACCATGAG
CCTCATCCACGACGACCTCCCCGCCATGGACAACGACGACTTCCGGCGCGGCCA
GCCCCGCAACCACAAGGCCTATGGCGAGGAGATTGCGATCCTGGCGGGCGACGC
GCTGCTGTCGCTGAGCTTTGAGCACATCGCGCGCGAGACGCGAGGCGTGGACCC
GGTGC GCGTCTTGCCCGCCATCTCGGAGTGGCGCGCGGTGGGCAGCCGCGGGCT
GGTGGCGGGGCGAGGTGGTGGACCTGGGTTTCGAGGGGCGGCGGCGTGGGGCTGGC
CCCGCTGCGCTACATCCACGAGCACAAAACCGCGGCGCTGCTGGAGGCGGCGGT
GGTGTCCGGCGCGCTGCTGGGCGGCGCGGAGGAGGCGGACCTGGAGCGCCTGCG
CACCTACAACCGCGCCATCGGCCTCGCTTTCCAGGTGGTGGGGGACATCCTGGAC
ATCCCGGGGACAGCGAGGAGCTGGGCAAGACCGCGGGCAAGGACCTGAGCTCC
CCCAAACCCCCCTACCCGTCCCTGGTGGGGCTGGCCAGGTCCAAAAAAATTGCG
GACGAACTGATTGAGGACGCGAAAACCCAACTACCCAGTACGAGCCGGCCCCGA
GCGGCGCCCCCTCGTAACCCTGGCCGAAAACATTTGAaaccggaagaactgactgggggcccccc
ctgccccagatacggggggctcctccatccagtttgggatgggaggagcgacaaccgaccccgtaaccctgtgacgcgttgcg
ttgcatacgtacgcatgccttgaacccatccatgacctcaacaatacctggttgtgttagcttggctctgaaaaaaaaaaaaaaaaa
aaaaaaaaaaaa

SEQ ID NO: 162

Geranyl geranyl diphosphate synthase

MGVGTLPYPTCGRAFTHSARYVPRRAARSRGHPTRCTAEVRARPSANGAQPMTAF
DFRQYMQQRAALVDAALDLAVPLQYPEKINEAMRYSLLAGGKVRPALCLAACEL
VGGPLEAAPAACAMEMIHTMSLIHDDLPAAMDNDDFRRGQPANHKAYGEEIAILAG
DALLSLSFEHIARETRGVDPVRVLAASEWRAVGSRLVAGQVVDLGFEGGGVGLA
PLRYIHEHKTAALLEAAVVSGALLGGAEEADLERLRITYNRAIGLAFQVVGDILDIPGT
SEELGKTAGKDLSSPKTPYPSLVGLARSKKIADELIEDAKTQLTQYEPARAAPLVTLA
ENI

SEQ ID NO: 163

Geranyl geranyl diphosphate synthase

cagatgccATGCGCCCTCGGGCCGCGGGCCTGAGGGTCCACGCAGCGTCCTCGGTGG
CCCAGACGCACCAAGGCCGCCCCCGGCGGACAGGAGGTTCGACGACTACCAGC
CCCGCACCGCCATCCTCTTCCCCGGCCAGGGCGCGCAGAGCGTGGGCATGGCGG
GAGAGCTGGCGAAGGCCGTCCCCGCGCGCGCGCTGTTTCGACGCCGCCTCCG
ACCAGCTCGGCTATGACCTGCTCCGCGTGTGCGTTGAGGGCCCCAAGGCGCGCCT
GGACAGCACCGCCGTCAGCCAGCCCGCCATCTACGTGGCCAGCCTGGCGGCGGT
GGAGAAGCTGCGCGCGGAGGGCGGGGAGGAGGCACTGGCCGCCATCGACGTCG
CTGCCGGTCTGTCCTTGGGCGAGTACACCGCGCTGGCCTTTGCCGGCGCCTTCTC
CTTCGCCGACGGGCTGCGCCTGGTGGCCCTGCGCGGCGCCAGCATGCAGGCCGC
CGCCGACGCCGACCCCTCGGGCATGGTCTCCGTATCGGTCTGCCCTCCGACGCG
GTGGCCGCGCTGTGCGAGGCGCCAACGCGCAGGTGGCCCCCGACCAGGCCGTG
CGCATCGCCAACCTCTGCGACGGCAACTACGCCGTACGCGGTGGGCTGGAG
GGCTGCGCGGCGGTGGAGGGCCTGGCCAAGGCCCAAGGCGCGCATGACGGTG
CGCCTGGCGGTGGCGGGCGCCTTCCACACCCCCCTTCATGCAGCCGGCGGTGGAG
GCGCTGAGCGCGGGCGCTGGCGGACACGCCGCTGGTTCGCGCCGCGCATCCCCGT
GGTCAGCAACGGGACGCC

SEQ ID NO: 164

Geranyl geranyl diphosphate synthase

MRPRAAGLRVHAASSVAQTHQAAPPADRRFDDYQPRTAILFPGQGAQSVGMAGEL
AKAVPAAAALFDAASDQLGYDLLRVCVEGPKARLDSTAVSQPAIYVASLAAVEKLR
AEGGEEALAAIDVAAGLSLGEY TALAFAGAFSADGLRLVALRGASMQAAADAAPS
GMVSVIGLPSDAVAALCEAANAQVAPDQAVRIANYLCDGNYAVSSGGLEGCAAVEG
LAKAHKARMTVRLAVAGAFHTPFMQPAVEALSAGAGGHAAGRAAHPRGQQRDA

SEQ ID NO: 165

Glyceraldehyde 3-phosphate dehydrogenase cDNA sequence

TGTCCATCTCCCCCACCCTCCATCCAACCATCGTCGACGGCATGCAGGCGCTGT
GTTCTCACCCCGCGTCCCTCACGGCGCGTGCGGTACCCCATGGGCGGGCCAGCCC
AGCACAGCGGGTGTCCAGCGCCGGCCCGGCCTACACCGGCCTGTCCCGGCACAC
CCTGGGCTGCCCCAGCACCCCCACCCTCCAGTCCCGCGCCGCGGTCCAGACCCGC
GGCTCCTCCTCCGGCTCCACCACGCGCATGACCACCACCGCCCAGCGCAAGATCA
AGGTGGCCATCAACGGGTTCGGCCGCATCGGCCGCCAGTTCCTGCGCTGCGTGGA
GGGGCGGAGGACTCGCTGCTGGAGATCGTGGCCGTGAACGACTCCGGCGGGCGT
GAAGCAGGCCAGCCACCTGCTCAAGTACGACTCCACCATGGGCACCTTCAACGC
CGACATCAAGATCTCGGGCGAGGGCACCTTCTCCGTCAACGGCCGCGACATCCG
CGTCGCTCTCTCCCGCGACCCCTGGCCCTGCCCTGGGGCGAGCTGGGCGTGGAC
CTGGTGATCGAGGGGACGGGAGTGTGTTGTGGACCGCAAGGGTGCCAGCAAGCAC
CTGCAGGCGGGGGCCAAGAAGGTCATCATCACCGCGCCGGCCAAGGGCTCCGAC
GTGCCACCTACGTGATGGGCGTGAACGCGGACCAGTACTCCAACCTCCGACGAC
ATCATCTCCAACGCCTCCTGCACCACCAACTGCCTGGCGCCCTTTGTCAAGGTGC
TCAACGACCGCTTCGGCATCGTGA

SEQ ID NO: 166

Glyceraldehyde 3-phosphate dehydrogenase

MQALCSHPASLTARAVPHGRASPAQRVSSAGPAYTGLSRHTLGCPSTPTLQSRAAVQ
TRGSSSGSTTRMTTTAQRKIKVAINGFGRIGRQFLRCVEGREDSLLEIVAVNDSGGVK
QASHLLKYDSTMGTNADIKISGEGTFSVNGRDIRVVSSRDPLALPWGELGVDLVIEG
TGVFVDRKGASKHLQAGAKKVIITAPAKGSDVPTYVMGVNADQYSNSDDIISNASCT
TNCLAPFVKVLNDRFGIV

SEQ ID NO: 167

Oxygen Evolving Complex protein (OEE33) cDNA sequence

gatgtgagaatagtagcttgctgccctgtcgccatgcagagcgtgtgcgcgcagtcgggttcatgcaagggggccttaccacagtc
tgcggacccccgatgcagcaggagccagctcgtctccgggctgatggcaaggccggagccttcataagaccgtaagagcgg
tgctgccgtctggctgcctccctcctctgtctgggggtgcgggcgactgaccttgatgagctgcagggcctgacctacctgag
gtgaagggtctggcatcgccaacacctgccccacctgtctggcggctcctccaacatcaaggacctgaagagcgggacctactcc
gtcaacaagatgtgcctggagcccacgtcctcaaggtcaaggaggaggcacagtcaagaacggcgaggccgacttgtgccac
caagctcgtcacgcgtctgacctacacctggacgagatctctggccagatgaagatcgacggcagcggcggtggagtcaagg
aggaggtatggcatgactatgctgcagtcaccgtgcagctccgggcggggagcgcgtgccctcctcttaccatcaaggagcttg
acgccaaggggaactgccgacggctcaaggcgagttaccgtgccctcctaccgtgggtcctcctctggacccaagggccgc
ggcgctccaccggctacgacaacgccgtggccctgcccgcgcggcgattccgaggagttggagaaggagaacaacaagtc
accaaggctctgaagggggaggccatcttccatcgccaagtggaacggcgagggaggtggcgggcatctttgagt

SEQ ID NO: 168

Oxygen Evolving Complex protein (OEE33)

MQSVCAQSVSCKGAFTQSLRTPRCRSRLVCRADGKAGAFIKTVKSGAAALAASLL
LSGGAGALTFDELQGLTYLQVKGSGIANTCPTLSGGSSNIKDLKSGTYSVNKMCLEP
TSFKVKEEAQFKNGEADFVPTKLVRTLYTLDEISGQMKIDGSGGVEFKEEDGIDYA
AVTVQLPGGERVFPFLFTIKELDAKGTADGFKGEFTVPSYRGSSFLDPKGRGASTGYD
NAVALPAAGDSEELEKENNKSTKALKGEAIFSLAKVDAGTGEVAGIFE

SEQ ID NO: 169

Clp protease cDNA sequence

ataatcggaaccagctgcacgcaccatcagtcgaggcagcatgcagaccgtcgcagccagctatggcgattggcgccctccggct
ccagcgtgacccggggctcgaccagcagcaagcagcacttcaccacctcactccctttccggctcaggcgccgaatcatgtgga
tcgggcggggcaggcggggtctgggagccccagaccctgcagcaggccgtgggcaaggccgtgcgccgtcgcggggccgc
accaccagcgccgtgcgcgtgacccgcatgatgtttgagcgggtcaccgagaaggccatcaagggtgcatgctcgcgcaggagga
ggctcggcgtctgggccacaactcgtggggacggagcaaatcctgctgggggttgattggggagtcacaggcacgcgcgccaagg
tcctcaagtcgatgggcgtcacgctgaaagatgcgcgtgtggaggtcgagaagatcatcgccggggggagcggccttggccgtg
gagatcccttcacccccgcgccaagcgtgtgctggagctgtccctggaggaggctcggcagctcggccacaactacattggcac
ggagcacatcctgctgggcctgctgcgcgagggtgagggcggtggcctcccgcgtgctggagaccctgggcgcgacccccagaa
gatccgcactcaggtgtacgcatggtgggtgagtcgcaggagcccgtgggcaccacgggtggcgagggtccaccggctccaa
caagatgcccacctggaggagtacggcaccaacctgaccgcccaggccg

SEQ ID NO: 170

Clp protease

MQTVAAASYGVLAPSGSSVTRGSTSSKQHFTTLTPFSGFRRLNHVDRAQAGSGSPQT
LQQAVGKAVRRSRGRITSAVRVTRMMFERFTEKAIKVVMLAQEEARRLGHNFGVT
EQILLGLIGESTGIAAKVLKSMGVTLKDARVEVEKIIGRGSGFVAVEIPFTPRAKRVLE
LSLEEARQLGHNHYIGTEHILLGLLREGEQVAVSRVLETLGADPQKIRTQVVRMVGESQ
EPVGTTVGGGSTGSNKMPITLEEYGTNLTAQA

SEQ ID NO: 171

TCCGGTGTGTTGTAAGTCCA

SEQ ID NO: 172

TTGTGGAATGTCATATCAA

SEQ ID NO: 173

AACGCCTTTGTACAACTGCA

SEQ ID NO: 174

CTGACCCGACCTATGGGAGCGCTCTTGGC

SEQ ID NO: 175

CTTGACTTCCCTCACCTGGAATTTGTCTG

SEQ ID NO: 176

GTGGCCATATGGACTTACAA

SEQ ID NO: 177

CAAGGGCTGGATGAATGACCCCAATGGACTGTGGTACGACG

SEQ ID NO: 178

CACCCGTCGTCATGTTACGGAGCCCAGTGCG

SEQ ID NO: 179

KE858 Homologous recombination construct SZ725

GCCCTTTGTCATCGTTGGCATGCTTTTTGCGTATGTACCATATGTTGAATGTATAA
TACGAACGGTTGACCGTCTGAGATGCGAGCTTTGGGTCTTGTCAAATGCGTGGCC
GCACGGCTCCCTCGCACCCAGCCCCGAGGCGTCGCGCACCTGGCGAGGAGCAGA
CCCACGCCAAGAAAGTCTAGTCCAGCATGTAACAACATCAGGCAATGTGACGTT
TTCGGTTCCCGATTTCTCTGCCGCTCTTTGACGGCAGGCACGGGCGAGCAACCGG
CGGCGCTCGCGTCAGGCACGATGGATGCGGCGCTGCCCACCTGTCAATGTACCCC
ACCAGTCTGTGATCGCTACAAGCAACCTTGTGCTCCACATTCCCAGTTGCAGAC
AGTCTAGTCGATTTTGCCAAGCTGGATGTGAGGATTGGCCATATCTTGGAGGCCA
AGATTACCCGGATGCTGATGGGTACGTACGCGAGCCAGGCAGGCAGCTGCGTT
GACTTTCTGATTGGCACAAAGCTTTGGCTACTCTCAATACCAACCACGTGCCCT
TCTGCACACCTGCTTCCTTCTGATGACCACTCGCCACGCATGTCGCAGTCTGTACG
TCGAGCAGATCGACCTCGGCGAGGAGGGGGGCCCTCGCACCATCGTGAGTGGCC
TGGTCCGGCACGTGACCCTGGAGGACCTTGTGCGCCGGCGGGTGGTGGTGCTGG
CCAACCTCAAGCCTCGGAGCATGCGCGGGGTCAAATCGGCTGGGATGCTGCTCT
GCGCCGCCAACGCGGATCACACCGCGGTGGAGCCGCTGCGGGTCCCGGACGCCG
CCGTGACGGGGGAGCGGGTCTGGGCGGGGGACGAGGCACTCCTGTCCACGGAGC
CTGCCACACCCAATCAGGTAAGGACACGTTATTGGTGCGCATGGTGCGAATGCGT
GGTCTGACCTGCTGTGGGTATGTGTTGTGGGATTGGAAACCGAATGAGGGCCGTT
CAGGATTGAGCCCTTGGCCCCACCCTGCTCATCCTCTCACGCCCCGAGGTCCAGA
AGAAGAAAATCTGGGAGGCAGTACAGCCGCTGCTGAGAGTGAACGCCCAGGGG
ATCGCTACTGTGGCAGGAGAGGCTATGGTGACCAGTGCGGGGCCACTGACCGCG
CCCACGCTGGTTGACGCCGCGATTTCTGACGCGAGCGACTGATTCTTGACCTTT
GAGAAGCCACCACAGCACCATTTTCAATTGTTTCATCCTTGATTTTCAGTACGACTTCT
CACCATTTTCAGTACTGTAGGACCCCCAAAATAGTGTGATCACGCTCGCAAGGCAC
CTGTGTGATCACGGGGGAAGGGCGAATTCTTTCTTGCGCTATGACACTTCCAGCA
AAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGA
TGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCC
AGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGA
GCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACT
CGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCTCTTCGTTTCAGTCAC
AACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCTTCTGTTCTGCTGGCCGGC
TTGCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCGGACCGCCCCCTG

GTGCACTTCACCCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTAC
 GACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACC
 GTCTGGGGGACGCCCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACT
 GGGAGGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCT
 CCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCAT
 CGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGA
 GGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCA
 GAAGAACCCCGTGCTGGCCGCCAATCCACCCAGTTCCGCGACCCGAAGGTCTTC
 TGGTACGAGCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTAC
 AAGATCGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGT
 TCGCCAACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGT
 CCCCACCGAGCAGGACCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAAC
 CCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTCGGCAGCTTCAACG
 GCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGG
 ACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCT
 GGGCATCGCGTGGGCCTCCAATGGGAGTACTCCGCCTTCGTGCCACCAACCCC
 TGGCGCTCCTCCATGTCCCTCGTGCAGAGTTCTCCCTCAACACCGAGTACCAGG
 CCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCA
 GCAACGCCCGGCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCA
 ACAGCTACAACGTGACCTGTCCAACAGCACCCGGCACCCCTGGAGTTCGAGCTGG
 TGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGTTGCGGGACCTCTC
 CCTCTGGTTCAAGGGCCTGGAGGACCCGAGGAGTACCTCCGCATGGGCTTCGA
 GGTGTCCGCGTCTCTCTTCTTCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTG
 AAGGAGAACCCCTACTTCAACCAACCGCATGAGCGTGAACAACCGACCCCTTCAAG
 AGCGAGAACGACCTGTCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATC
 CTGGAGCTGTACTTCAACGACGGCGACGTGCTGTCCACCAACACCTACTTCATGA
 CCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGT
 TCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGATTAATTAACCTCGAGGCA
 GCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTTCGTGTGATGGACTGT
 TGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAG
 CCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTG
 CTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCATC
 CCAACCGCAACTTATCTACGCTGTCTGTCTATCCCTCAGCGCTGCTCCTGCTCCTG
 CTCCTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTG
 CAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAA
 CACAAATGGAAAGCTT

SEQ ID NO: 180

Homologous recombination construct SZ726

CCCGTGATCACACAGGTGCCCTTGCGAGCGTGATCACACTATTTTGGGGGTCCTAC
 AGTACTGAAATGGTGAGAAGTCGTAAGTCAAGGATGAACAATGAAAATGG
 TGCTGTGGTGGCTTCTCAAAGGTCAAGAATCAGTCGCTCGCGTCAGGAAATCGCG
 GCGTCAACCAGCGTGGGCGCGGTCAAGTGGCCCCGCACTGGTCACCATAGCCTCTC
 CTGCCACAGTAGCGATCCCCCTGGGCGTTCACTCTCAGCAGCGGCTGTACTGCCTC
 CCAGATTTTCTTCTTCTGGACCTGCGGGCGTGAGAGGATGAGCAGGGTGGGGCCA
 AGGGCTCAATCCTGAACGGCCCTCATTCGGTTTCCAATCCACAACACATACCCA
 CAGCAGGTGAGACCACGCATTTCGCACCATGCGCACCAAATAACGTGTCCTTACCT
 GATTGGGTGTGGCAGGCTCCGTGGACAGGAGTGCCTCGTCCCCCGCCAGACCC
 GCTCCCCCGTCACGGCGGCGTCCGGGACCCGCAGCGGCTCCACCGCGGTGTGATC

CGCGTTGGCGGCGCAGAGCAGCATCCCAGCCGATTTGACCCCGCGCATGCTCCG
 AGGCTTGAGGTTGGCCAGCACCACCACCCGCCGCGACAAGGTCCTCCAGGGT
 CACGTGCCGGACCAGGCCACTCACGATGGTGCAGGGGCCCCCTCCTCGCCGAG
 GTCGATCTGCTCGACGTACAGACTGCGACATGCGTGGCGAGTGGTCATCAGAAG
 GAAGCAGGTGTGCAGAAGGGGACGCTGGTTGGTATTGAGAGTAGCCAAAGCTTT
 GTGCCAATCAGAAAGTCAACGCAGCTGCCTGCCTGGCTCGCGTACAATTCCTTTC
 TTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCG
 GCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGC
 ATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCC
 GATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCA
 CTACCACCTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCG
 CCTCTTCTCTTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAG
 GCCTTCTCTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGA
 ACGAGACGTCCGACCGCCCCCTGGTGCACCTTCACCCCAACAAGGGCTGGATGA
 ACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACT
 TCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTCTGGGGCCACGG
 CACGTCCGACGACCTGACCAACTGGGAGGACCAGCCATCGCCATCGCCCCGAA
 GCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACC
 TCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGA
 CCTACAACACCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCG
 GCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCA
 GTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGATCATGACC
 GCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAG
 TCCTGGAAGCTGGAGTCCGCGTTTCGCCAACGAGGGCTTCCTCGGCTACCAGTACG
 AGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCCTACT
 GGGTGATGTTTCATCTCCATCAACCCCGCGCCCCGGCCGGCGGCTCCTTCAACCA
 GTACTTCGTCCGACGCTTCAACGGCACCCACTTCGAGGCCCTTCGACAACCAGTCC
 CGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCG
 ACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACT
 CCGCCTTCGTGCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTC
 TCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAG
 GCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACC
 AACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACAGCACC
 GGCACCTTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCA
 AGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGA
 GTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCTTCTGGACCGCGGG
 AACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGC
 GTGAACAACCAGCCCTTCAAGAGCGAGAACGACCTGTCCTACTACAAGGTGTAC
 GGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTCGTGT
 CCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGAC
 GACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCCAGGTGCGCGAGGTCAA
 GTGATTAATTAACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGA
 CGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATA
 TCCCTGCCGCTTTATCAAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTT
 TTGCGAGTTGCTAGCTGCTTGTGCTAATTTGCGAATACCACCCCCAGCATCCCCTTC
 CCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATC
 CCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCTCGCACAGCCTTGGTTTGGGCT
 CCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGC
 ACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTCGGTACCCGTA
 CCCATCAGCATCCGGGTGAATCTTGGCCTCCAAGATATGGCCAATCCTCACATCC

AGCTTGGCAAAATCGACTAGACTGTCTGCAAGTGGGAATGTGGAGCACAAAGGTT
GCTTGTAGCGATCGACAGACTGGTGGGGTACATTGACAGGTGGGCAGCGCCGCA
TCCATCGTGCCTGACGCGAGCGCCGCGGTTGCTCGCCCGTGCCTGCCGTCAAAG
AGCGGCAGAGAAATCGGGAACCGAAAACGTACATTGCCTGATGTTGTTACATG
CTGGACTAGACTTTCTTGGCGTGGGTCTGCTCCTCGCCAGGTGCGCGACGCCTCG
GGGCTGGGTGCGAGGGAGCCGTGCGGCCACGCATTTGACAAGACCCAAAGCTCG
CATCTCAGACGGTCAACCGTTCGTATTATACATTCAACATATGGTACATACGCAA
AAAGCATG

SEQ ID NO: 181

Homologous recombination: Targeting cassette for disruption of *Prototheca moriformis*
stearoyl ACP desaturase coding region (β -tubulin driven *suc2* cassette)

TTTGGCCCCGCTTTCCAGCTCCGGATCTGCTGGCGTCCGCCGCGAGACGTGACAT
CGCACGTGCGCCGGGAGCGCCAGCTTGATCACTTGGCAGGGGGCCGTGCTCTACA
AATACCAGGCCCCGCGGCGGTTCAGTTCGCACATCCAATACCTGCCGAGCCATCTT
GCCTACACTTTTTATCGACTCCTCTACTCTGTTTCGCGAGAGCGCTCGGTCCAGGCT
TGGAATTCGCCGAATTCAGCTCGATCAGTCGCTTCTGCAACTGATCTCGGCGGTT
CGCAGACTGCCTTTTTCTCAGCTTGTCAGGTAGCGAGTTGTTGTTTTATATTTATTC
GATTTTCATCTGTGTTGCATGTCTTGTTTCGTGCTGTGCGTTCTTTCTGGGCCGCGCT
GTCGGGTCGCATGGGCTAGCTGTACTCATGTTAGTCATGCCGGTCCGACCTTGTT
CGAGGAAGGCCCCACACTGAGCGTGCCCTCTTTCTACACCCCTTGTGCAGAAATT
AGATAGAAAGCAGAATTCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGG
GCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGAC
CCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGC
GCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAA
GCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGT
GATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCA
AACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTTCCTGCTGGCCGGCTTCGCCGC
CAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCACCTC
ACCCCCAACAAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAG
GACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGG
ACGCCCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGAC
CAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCA
TGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCG
CCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTA
CATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCC
GTGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGC
CCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGA
TCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTCCGCAACGA
GGGCTTCTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAG
CAGGACCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCC
CGGCCGGCGGCTCCTTCAACCAGTACTTCGTGCGCAGCTTCAACGGCACCCACTT
CGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGC
CCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCG
TGGGCCTCCAACCTGGGAGTACTCCGCCTTCGTGCCACCAACCCCTGGCGCTCCT
CCATGTCCCTCGTGCAGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGA
GACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGG
CCCTTGAGCCGGTTTCGCCACCAACACCGTTCGACGAAGGCCAACAGCTACAA
CGTCGACCTGTCCAACAGCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTC

AACACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCA
 AGGGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGT
 CCTCCTTCTTCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACC
 CCTACTTCACCAACCGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGAGAACG
 ACCTGTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTA
 CTTCAACGACGGCGACGTGCTGTCCACCAACACCTACTTCATGACCACCGGGAAC
 GCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGAC
 AAGTTCAGGTGCGCGAGGTCAAGTGATTAATTAACCTCGAGGCAGCAGCAGCTC
 GGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACA
 CTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTG
 TTTGATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGA
 ATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAA
 CTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCC
 CTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAA
 ACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGA
 CCGACACGCCCCCGGCCAGGTCCAGTTCTCCTGGGTCTTCCAGAGGGCCCGTCGC
 CATGTAAAGTGGCAGAGATTGGCGCCTGATTTCGATTTGGATCCAAGGATCTCCAA
 TCGGTGATGGGGACTGAGTGCCCAACTACCACCTTGCACATCGTCCCTCGCACT
 ATTTATTCCCACCTTCTGCTCGCCCTGCCGGGCGATTGCGGGCGTTTCTGCCCTTG
 ACGTATCAATTTGCCCCCTGCTGGCGCGAGGATTCTTCATTCTAATAAGAACTCA
 CTCCCGCCAGCTCTGTACTTTTCTGCGGGGCCCTGCATGGCTTGTTCCTCAATGC
 TTGCTCGATCGACGGCGCCCATTTGCCACGGCGCTGCCGCATCCATGTGAAGAAA
 CACGGAAGAGTGCGAAGACTGGAAGTGAATTAAGAGTATAAGAAGAGGTACCA
 AGGGATTCTCAGGTGCTCTTAGGAACGGCTTTTCTCCTTCGCCAAGAGAACTGCTA
 CTGCTCGTGTGCGCACGGTGGTCAAGCCGCCCATCTGCGATCCACCAGGCCCCAT
 CCGCGGACTCGCGATCAGCCTGCTGGATCCGGACTGCCGACCTGACCCTCGCAT
 CCACCATTACAACCCTCCAATTGGACACCACTCCCACGTCCTAAAGTTCACCATG
 CAAGCTGATCGATCGCATTCGCCGATGCACTCGCCTGCCACAGAGGTGTGCGCTT
 CGGACTAGCGTGCAGGCGCCCCGAGGCCACCAGCATGCACCGATGGAAGCGGGC
 ACGGCCGCTGCTCCAGGTGCTGGCTCGCTCAGACCCATAGCAACCTCCGCTGCG
 TCCCTAAATGTACACAGAGCGTCTTTGATGGGTACGGATGGGAGAGAATCTGAT
 TGGGCATTGCTGGTGCAGTGCAGGAAGATGGCAAGTGCACAGTCAGTCATGCTG
 TACAAACTGGTGCCTCGTAGTATTGACTCGTATAGTGCATAGTATCATGCATGGT
 CGTTACTTGCAA

SEQ ID NO: 182

Homologous recombination: Targeting cassette for disruption of *Prototheca moriformis* stearoyl ACP desaturase coding region (*suc2* cassette alone)

TTTGGCCCCGCTTTCCAGCTCCGGATCTGCTGGCGTCCGCCGCGAGACGTGA
 CATCGCACGTGCGCCGGGAGCGCCAGCTTGATCACTTGGCAGGGGGCCGTGCT
 CTACAAATACCAGGCCCGCGGCGGTTCAGTTCGCACATCCAATACCTGCCGA
 GCCATCTTGCCTACACTTTTTATCGACTCCTCTACTCTGTTTCGCGAGAGCGCTC
 GGTCCAGGCTTGGAATTCGCCGAATTCAGCTCGATCAGTCGCTTCTGCAACT
 GATCTCGGCCGTTTCGCAGACTGCCTTTTCTCAGCTTGTGAGGTAGCGAGTTGT
 TGTTTTATATTTATTCGATTTTCATCTGTGTTGCATGTCTTGTTCGTGCTGTGCGT
 TCTTTCTGGGCCGCGCTGTGCGGTGCGATGGGCTAGCTGTACTCATGTTAGTC
 ATGCCGGTCCGACCTTGTTTCGAGGAAGGCCCCACACTGAGCGTGCCCTCTTT

CTACACCCCTTGTGCAGAAATTAGATAGAAAGCAATGCTGCTGCAGGCCTTC
 CTGTTCTTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACG
 AGACGTCCGACCGCCCCCTGGTGCACCTCACCCCCAACAAAGGGCTGGATGA
 ACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGT
 ACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCTGGGG
 CCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCAT
 CGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGAC
 TACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCT
 GCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCT
 CCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCG
 TGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGA
 GCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGAT
 CGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTC
 GCCAACGAGGGCTTCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAG
 GTCCCCACCGAGCAGGACCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCA
 TCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTCCGCAG
 CTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGAC
 TTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACCT
 ACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCTT
 CGTGCCCAACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCC
 TCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGG
 CCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCA
 CCAACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACA
 GCACCGGCACCCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGA
 CGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAG
 GACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCTCTTCTT
 CCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTT
 CACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGACCT
 GTCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTAC
 TTCAACGACGGCGACGTCTGTCCACCAACACCTACTTCATGACCACCGGGA
 ACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACAT
 CGACAAGTTCCAGGTGCGCGAGGTCAAGTGACCGACACGCCCCCGGCCAG
 GTCCAGTTCTCCTGGGTCTTCCAGAGGCCCGTCCGCATGTAAAGTGGCAGAG
 ATTGGCGCCTGATTCGATTTGGATCCAAGGATCTCCAATCGGTGATGGGGAC
 TGAGTGCCCAACTACCACCCCTTGCACTATCGTCCCTCGCACTATTTATTCCCAC
 CTTCTGCTCGCCCTGCCGGGCGATTGCGGGCGTTTCTGCCCTTGACGTATCAA
 TTTCGCCCCCTGCTGGCGCGAGGATTCTTCATTCTAATAAGAACTCACTCCCGC
 CAGCTCTGTACTTTTCTGCGGGGCCCCCTGCATGGCTTGTTCCCAATGCTTGCT
 CGATCGACGGCGCCCATTTGCCACGGCGCTGCCGCATCCATGTGAAGAAAC
 ACGGAAGAGTGCGAAGACTGGAAGTGAATTAAGAGTATAAGAAGAGGTAC
 CAAGGGATTCTCAGGTGCTCTTAGGAACGGCTTTTCTTTCGCCAAGAGAAAC
 TGCTACTGCTCGTGTGCCACGGTGGTCAAGCCGCCCCATCTGCGATCCACC
 AGGCCCATCCGCGGACTCGCGATCAGCCTGCTGGATCCGGACTGCCGACCTG

ACCGCTCGCATCCACCATTACAACCCTCCAATTGGACACCACTCCCACGTCC
TAAAGTTCACCATGCAAGCTGATCGATCGCATTCGCCGATGCACTCGCCTGC
CACAGAGGTGTGCGCTTCGGACTAGCGTGCAGGCGCCCCGAGGCCACCAGC
ATGCACCGATGGAAGCGGGCACGGCCGCTGCTCCAGGTCGCTGGCTCGCTC
AGACCCATAGCAACCTCCGCTGCGTCCCTAAATGTCACACAGAGCGTCTTTG
ATGGGTACGGATGGGAGAGAATCTGATTGGGCATTGCTGGTGCAGTGCAGG
AAGATGGCAAGTGCACAGTCAGTCATGCTGTACAAACTGGTGCCTCGTAGTA
TTGACTCGTATAGTGCATAGTATCATGCATGGTCGTTACTTGCAA

SEQ ID NO: 183

Cuphea wrightii acyl-ACP thioesterase Genbank ACCESSION U56103

VERSION U56103.1 GI:1336005

MVAAAASSAFFSVPTPGTSPKPGKFGNWPSSLSVPFKPDNGGFH
VKANASAHPKANGSAVNLKSGSLETPPRSFINQLPDL SMLLSKITTVFGAAEKQWKRP
GMLVEPFGVDRI FQDGVFFRQSF SIRS YEIGVDRTAS IETLMNIFQETSLNHCKSIGL
LNDGFGRTPEMCKRD LIWVVT KIQVEVNRYPTW GDTIEVNTWVSESGKNGMGRDWLIS
DCRTGEILIRATSVWAMMNQNT RRLSKFPYEV RQEIAPHFVDSAPVIEDDRKLHKL DV
KTGDSIRDGLTPRWNDLDVNQH VNNVKYIGWILKSVP IEFVETQELCGVTLEYRRECG
RDSVLESVT TMDPAKEGDR CVYQHLLRLEDGADITIGRTEWRPKNAGANGA ISSGKTS
NGNSVS

SEQ ID NO: 184

Cuphea wrightii acyl-ACP thioesterase Genbank ACCESSION U56104

VERSION U56104.1 GI:1336007

MVAAAASSAFFPVPAPRPTPKPGKFGNWPSSLSQPFKPKSNPN
GRFQVKANVSPHPKANGSAVSLKSGSLNTLEDPPSSPPPTFLNQLPDWSRLRTAITT
VFVAAEKQFTRLD RKS KRPDMLVDWFGSETIVQDGLVFRERFSIRS YEIGADRTASIE
TLMNHLQD TSLNHCKSVGLLNDGFGRTPEMCTRDLIWVLTKMQIVVNRYPTW GDTVEI
NSWFSQSGKIGMGREWLI SD CNTGEILVRATSAWAMMNQKTRRF SKLPCEVRQEIAPH
FVDAPPVIEDNDRKLHKFDVKTGDSICKGLTPGW NDFDVNQHVS NVKYIGWILESMPT
EVLETQELCSLTLEYRRECGRESVVESVTSMNPSKV GDRSQYQHLLRLEDGADIMKGR
TEWRPKNAGTNRAIST

SEQ ID NO: 185

Cuphea wrightii beta-ketoacyl-ACP synthase II Genbank ACCESSION No. U67317

VERSION U67317.1

MAAAASMVASPFCTWLVASCMSTSF DNDPRSPSVKRFP RRKRVL
SQRGSTFQCLVASCIDPCDQYRSSASLSFLGDNGFASLFGSKPFMSNRGHRRLRRASH
SGEAMAV ALQPAQEAGTKKKPVIKQRRVVVTGMGVVTPLGHEPDVFYNNLLDGVSGIS
EIETF DCTQFP TRIAGEIKSFSTDGWVAPKLSKRMDKFMLYLLTAGKKALADGGITDE
VMKELDKRKCGVLIGSGMGGMKVFND AIEALRVSYKKMNPF CVPFATNMGSAMLAMD
LGWMGPNYSTACATSNFCILNAANH IIRGEADMMLCGGSDAVIIP IGLGGFVACRA
LSQRNSDPTKASRPWDSNRDGFVMGEGAGVLLLEELHAKKRGATIYAEFLGGSTCD
AYHMT EPHPEGAGVILCIEKALAQAGVSKEDVNYINAHATSTSAGDIKEYQALARCFG
QNSLRVNSTKSMIGHLLGAAGGVEAVTVVQAIRTGWIHPNLNLEDPDKAVDAKLLVG
PKKERLNVKVGLSNSFGFGHNSSILFAPCNV

SEQ ID NO: 186

Cuphea hookeriana 3-ketoacyl-ACP synthase (Kas4)

Genbank ACCESSION AF060519 VERSION AF060519.2 GI:30995500

MATASCMVASPFCTWLVAACMPTSSDNDPRSLSHKRLRLSRRRR
TLSSHCSLRGSTFQCLDPCNQQRFLGDNGFASLFGSKPLRSNRGHLRLGRTSHSGEVM
AVAMQPAQEVSTNKKPATKQRRVVVTGMGVVTPLGHDPDVYYNNLLDGISGISEIENF
DCSQFPTRIAGEIKSFSTDGWVAPKFSEMRDKFMYMLTAGKKALADGGITEDAMKEL
NKRKCGVLIGSGLGGMKVFSDSIEALRTSYKKISPFVFPFSTTNMGSAILAMD LGWMG
PNYSISTACATSNFCILNAANHIIKGEADMMMLCGGSDAAVLPVGLGGFVACRALSQRN
NDPTKASRPWDSNRDGFVMGEGAGVLLLEELHAKKRGATIYAEFLGGSFCTDAYHMT
EPHPEGAGVILCIEKALAQSGVSREDVNYINAHATSTPAGDIKEYQALAHCFGQNSSEL
RVNSTKSMIGHLLGGAGGVEAVAVVQAI RTGWIHPNINLEDPDEGVDAKLLVGPKEK
LKVKVGLSNSFGFGGHNSSILFAPCN

SEQ ID NO: 187

oleoyl-acyl carrier protein thioesterase [Carthamus tinctorius] (safflower).

Genbank accession no. AAA33019 VERSION AAA33019.1 GI:404026

MLSKGAPAAP AVAAMYNASA KDTTFALTHS RSIGSVSIRR RYNVFLCNSS SSSRKVSPLL
AVATGEQPSG VASLREADKE KSLGNRLRLG SLTEDGLSYK EKFVIRCYEV GINKTATITET
IANLLQEVGG NHAQGVGFST DGFATTTTMR KLHLIWTAR MHIEIYRPA WSDVIEIETW
VQGEKVGTR RDWILKDYAN GEVIGRATSK WVMNEDTRR LQKVSDDVRE EYLVFCPRTL
RLAFPEENNN SMKKIPKLED PAEYSRLGLV PRRSDLDMNK HVNNVTYIGW ALESIPPEII
DTHLQAITL DYRRECQRDD IVDSLTSREP LGNAAGVKFK EINGSVSPKK DEQDLSRFMH
LLRSAGSGLE INRCRTEWRK KPAKR

SEQ ID NO: 188

Arabidopsis thaliana (thale cress) Genbank accession no: NP_566464 VERSION NP_566464.1

MTKEVCSNIG LWLLLTLLIG NYVVNLEASH HVYKRLTQST NTKSPSVNQP YRTGFHFQPP
KNWMNDPNPG MIYKGIYHLF YQWNPKGAVW GNIVWAHSTS TDLINWDPHP PAIFPSAPFD
INGCWSGSAT ILPNGKPVIL YTGIDPKNQ QVNIAEPKNL SDPYLREWKK SPLNPLMAPD
AVNGINASSF RDPTTAWLGQ DKKWRVIIGS KIHRRGLAIT YTSKDFLKWE KSPEPLHYDD
GSGMWECPDF FVTRFGSNG VETSSFGEPN EILKHVLKIS LDDTKHDYYT IGTYDRVKDK
FVPDNGFKMD GTAPRYDYGK YYASKTFFDS AKNRRILWG W TNESSVEDD VEKGWSGIQT
IPRKIWLD RS GKQLIQWPVR EVERLRTKQV KNLRNKVLKS GSRLEVYGV T AAQADVEVLF
KVRDLEKADV IEPSTWDPQL ICSKMNVS VK SGLGPFGLMV LASKNLEEY T SVYFRIFKAR
QNSNKYVVL M CSDQSRSSLK EDNDKTTYGA FVDINPHQPL SLRALIDHSV VESFGGKGRA
CITSRVYPKL AIGKSSHLFA FNYGYQSVDV LNLNAWSMNS AQIS

SEQ ID NO: 189Codon optimized coding region of *Cocos nucifera* FatB3-B

1 ACTAGTATGG TGGCCTCCGT GGCCGCCTCC TCCTCCTTCT TCCCCGTGCC CTCCTCCTCC
61 TCCTCCGCCT CCGCCAAGGC CTCCCGCGGC ATCCCCGACG GCCTGGACGT GCGCGGCATC
121 GTGGCCAAGC CCGCCTCCTC CTCCGCGTGG ATGCAGGCCA AGGCCTCCGC CCGCGCCATC
181 CCCAAGATCG ACGACACCAA GGTGGGCCTG CGCACCACG TGGAGGAGGA CGCCGCCTCC
241 ACCGCCCGCC GCACCTCCTA CGGGCGCGCC AACCAGCTGC CCGACTGGTC CATGCTGCTG
301 GCCGCCATCC GCACCATCTT CTCCGCCGCC GAGAAGCAGT GGACCCTGCT GGACTCCAAG
361 AAGCGCGGCG CCGACGCCGT GGCCGACGCC TCCGGCGTGG GCAAGATGGT GAAGAACGGC
421 CTGGTGATAC GCCAGAACTT CTCCATCCGC TCCTACGAGA TCGGCGTGGA CAAGCGCGCC
481 TCCGTGGAGG CCCTGATGAA CCACTTCCAG GAGACCTCCC TGAACCACTG CAAGTGCATC
541 GGCCTGATGC ACGGCGGCTT CGGCTGCACC CCCGAGATGA CCCGCCGCAA CCTGATCTGG
601 GTGGTGGCCA AGATGCTGGT GCACGTGGAG CGCTACCCCT GGTGGGGCGA CGTGGTGCAG
661 ATCAACACCT GGATCTCCTC CTCCGGCAAG AACGGCATGG GCCGCGACTG GCACGTGCAC

```

721 GACTGCCAGA CCGGCCTGCC CATCATGCGC GGCACCTCCG TGTGGGTGAT GATGGACAAG
781 CACACCCGCC GCCTGTCCAA GCTGCCCCGAG GAGGTGCGCG CCGAGATCAC CCCCTTCTTC
841 TCCGAGCGCG ACGCCGTGCT GGACGACAAC GGCCGCAAGC TGCCCAAGTT CGACGACGAC
901 TCCGCCGCCG ACGTGCGCCG CGGCCTGACC CCCCCTGGC ACGACTTCGA CGTGAACCAG
961 CACGTGAACA ACGTGAAGTA CGTGGGCTGG ATTCTGGAGT CCGTGCCCCG GTGGATGCTG
1021 GACGGCTACG AGGTGGCCAC CATGTCCCTG GAGTACCGCC GCGAGTGCCG CATGGACTCC
1081 GTGGTGCAGT CCCTGACCGC CGTGTCTCTC GACCACGCCG ACGGCTCCCC CATCGTGTGC
1141 CAGCACCTGC TGCGCCTGGA GGACGGCACC GAGATCGTGC GCGGCCAGAC CGAGTGGCGC
1201 CCCAAGCAGC AGGCCCGCGA CCTGGGCAAC ATGGGCCTGC ACCCCACCGA GTCCAAGTGA
1261 TTAATTAACG CGAG

```

SEQ ID NO: 190

Aspergillus awamori, preproglucoamylase G1, Genbank accession no. AAB59296

VERSION AAB59296.1

```

MSFRSLLALS GLVCTGLANV ISKRATLDSW LSNEATVART AILNNIGADG AWVSGADSGI
VVASPSTDNP DYFYTWTRDS GLVLKTLVDL FRNGDTSLLS TIENYISAQA IVQGISNPSPG
DLSSGAGLGE PKFNVDETAY TGSWGRPQRD GPALRATAMI GFGQWLLDNG YTSTATDIVW
PLVRNDLSYV AQYWNQTGYD LWEEVNGSSF FTIAVQHRAL VEGSAFATAV GSSCSWCDSQ
APEILCYLQS FWTGSFILAN FDSSRSGKDA NTLLGSIHTF DPEAACDDST FQPCSPRALA
NHKEVVDSEF SIYTLNDGLS DSEAVAVGRY PEDTYYNGNP WFLCTLAAAE QLYDALYQWD
KQGSLEVTDV SLDFFKALYS DAATGTYSST SSTYSSIVDA VKTFADGFVS IVETHAASNG
SMSEQYDKSD GEQLSARDLT WSYAALLTAN NRRNSVVPAS WGETSASSVP GTCAATSAIG
TYSSVTVTISW PSIVATGGTT TTATPTGSGS VTSTSKTTAT ASKTSTSTSS TSCTTPTAVA
VTFDLTATTT YGENIYLVGS ISQLGDWETS DGIALSADKY TSSDPLWYVT VTLPAGESFE
YKFIRIESDD SVEWESDPNR EYTVPQACGT STATVDTWR

```

SEQ ID NO: 191

Rhizopus oryzae (*Rhizopus delemar*) Glucoamylase 1;

1,4- α -D-glucan glucohydrolase, Genbank accession no. P07683

VERSION P07683.2

```

MQLFNLPLKV SFFLVLSYFS LLVSAASIPS SASVQLDSYN YDGSTFSGKI YVKNIAYSCK
VTVIYADGSD NWNNGNTIA ASYSAPISGS NYEYWTFAS INGIEFYIK YEVSGKTYD
NNNSANYQVS TSKPTTTTAT ATTTTAPSTS TTTTPSRSEP ATFTGNSTI SSWIKKQEGI
SRFAMLRNIN PPGSATGFIA ASLSTAGPDY YYAWTRDAAL TSNVIVYEYN TTLSGNKITL
NVLKDYVTFV VKTQSTSTVC NCLGEPKFN DASGYTGAW RPQNDGPAER ATTFILFADS
YLTQTKDASY VTGTLKPAIF KDLDYVVNVW SNGCFDLWEE VNGVHFYTLM VMRKGLLLGA
DFAKRNGDST RASTYSSTAS TIANKISSFW VSSNNWQVS QSVTGGVSKK GLDVSTLLAA
NLGSVDDGFF TPGSEKILAT AVAVEDSFAS LYPINKNLPS YLGNSIGRYP EDTYNGNGNS
QGNSWFLAVT GYAELYRAI KEWIGNGGVT VSSISLPFFK KFDSSATSGK KYTVGTSDFN
NLAQNIALAA DRFLSTVQLH AHNNGSLAE FDRTTGLSTG ARDLTWSHAS LITASYAKAG
APAA

```

SEQ ID NO: 192

Chlamydomonas reinhardtii hydroxymethylpyrimidine phosphate synthase, Genbank accession no.

XP_001697756, VERSION XP_001697756.1

```

MQLFARSPAG VRVQQHAQRK SAKVAGRRL NVRAEAIAMP PAPAEQEKEL QRLEKGNAFE
ELKAMATSPR QSVNRPPQAE DLTFRQAPT ADCFPDSEKM YKEVKYDDEI TLRVPFRRRIH
LTTGTDFFDV DTSGPQNVDP RVGLPKLRKP WVERREANGT HARAGVTQMA LARAGIISEE
MRFAAEREGL DPEFVRSELA RGRAIIPANK CHLELEPCVI GRNFMTKVNS NFGNSAVTSS
IEEEVEKLQW STIWGADTVM DLSTGHNIFE TREWVMRNSP VPVGTVPIYE ALERADQVE
GITWELFRQV LLDQAEQGVY WTIHAGVLL RHVPLTANRI TGIVSRGGSI HAKLCLMEHK
ENFAYEHWDE ILDICAKYDI TLSIGDGLRP GCIADANDAA QFAELKTQGE LTRRAWKKNV
QVMNEGPQHV PLNKIPENMA KQLEWCSEAP FYTLGPLTTD IAPGYDHITS AIGAATIGAL

```

GTALLCYVTF KEHLGLPNRD DVKAGVIAYK IAAHAADLAK GHFAAASWDL ELSKARFEFR
WRDQFALSLD PVTAQAYHDA TLPQEPAKTA HFCSMCGPKF CSMNITQELR QMVQAEQAQE
AAAAAAAADD ADLLAAAAG MKEMSEKFKQ SGAQLYH

SEQ ID NO: 193

thiamine biosynthesis protein ThiC [Salmonella enterica subsp. enterica serovar Typhimurium str. LT2], Genbank accession no. NP_463033, VERSION NP_463033.1

MSTTTLTRRE QRAKAQHFD TLEGTAFPNS KRIYVTGSQH DIRVPMREIQ LSPTLIGGSK
DNPQFEENE VPVYDTSGPY GDPEVAINVQ QGLAKLRQPW IDARNDSEEL DDRSSAYTRE
RLADDGLDDL RFTGLLTPKR AKAGKRVTQL HYARKGIVTP EMEFIAIREN MGRERIRSEV
LRHQHFGMNF GARLPENITP EFVRDEVAAG RAIIPANINH PESEPMIIGR NFLVKVNANI
GNSAVTSSIE EEVEKLVWST RWGADTVMDL STGRYIHETR EWILRNSPVP IGTVPYIYQAL
EKVNGIAEDL TWEAFRDILL EQAEQGVDF TIHAGVLLRY VPMTAKRLTG IVSRGGSIMA
KWCLSHHKN FLFEHFREIC EICAAYDVSL SLGDGLRPGS IQDANDEAQF SELHTLGELT
KIAWEYDVQV MIEGPGHVPM HMIQRNMTEE LESCHEAPFY TLGPLTTDIA PGYDHTSGI
GAAMIGWFGC AMLCYVTPKE HLGLPNKEDV KQGLITYKIA AHAADLAKGH PGAQIRDNAM
SKARFEFRWE DQFNALDPF TARAYHDETL PQESGKVAHF CSMCGPKFCS MKISQEVVDY
AAQTIEIGM ADMSEDFRAK GGEIYLKREE A

SEQ ID NO: 194

aminoglycoside-3'-O-phosphotransferase [Cloning vector pTAP6].Genbank accession no. ADJ58594, VERSION ADJ58594.1

MIEQDGLHAG SPAAWVERLF GYDWAQQTIG CSDAAVFRLS AQGRPVLVFK TDLSGALNEL
QDEAARLSWL ATTGVPCAAY LDVVTEAGRD WLLLGEVPGQ DLLSSHLAPA EKVSIMADAM
RRLHTLDPAT CPFHDQAKHR IERARTRMEA GLVDQDDLDE EHQGLAPAEI FARLKARMPD
GEDLVVTHGD ACLPNIMVEN GRFSGFIDCG RLGVADRYQD IALATRDIAE ELGGEWADRF
LVLYGIAAPD SQRIAFYRLL DEFF

SEQ ID NO: 195

acyl ACP thioesterase [Brassica napus (rape)].Genbank accession no. CAA52070, VERSION CAA52070.1

MLKLSNVNT NLHTFSFFSD SSLFIPVNR TIAVSSSQLR KPALDPLRAV ISADQGSISP
VNSCTPADRL RAGRLMEDGY SYKEKFIVRS YEVGINKTAT VETIANLLQE VACNHVQKCG
FSTDGFATTL TMRKLHLIIV TARMHIEIYK YPAWSDVVEI ETWCQSEGRI GTRRDWILRD
SATNEVIGRA TSKWVMNQD TRRLQRTVE VRDEYLVFCP REPRLAFPEE NNSSLKKIPK
LEDPAQYSML ELKPRRADLD MNQHVNNVTY IGWVLESIPQ EIIDTHELQV ITLDYRRECQ
QDDIVDSLTT SEIPDDPISK FTGTNGSAMS SIQGHNESQF LHMLRLSENG QETNRGRTQW
RKKSSR

SEQ ID NO: 196

acyl-ACP thioesterase [Garcinia mangostana], Genbank accession no. AAB51525 VERSION AAB51525.1

MVATAATSSF FPLTSPSGDA KSGNPGKGSV SFGSMKSKSA ASSRGLQVKA NAQAPT KING
STDDAQLPAP RTFINQLPDW SMLLAITTV FLAAEKQWMM LDWKPRRPDM LIDTFGLGRI
VQDGLVFRQN FSIRSYEIGA DRTASIVTM NHLQETALNH VKTAGLLGDG FGSTPEMSKR
NLIWVVTM Q VEVDYPTWG DVVQVDTWVS ASGKNGMRRD WLLRDGNTGE TLTRASSVWV
MMNKLTRRLS KIPEEVREEI GSYFVNSDPV VEEDGRKVT LDDNTADFVR KGLTPKWN DL
DINQHVNNVK YIGWILESAP QPILETRELS AVTLEYRREC GRDSVLRSLT AVSGGGVGDL
GHAGNVECQH VLRLEDGAEI VRGRTEWRPK YINNF'SIMGQ IPTDAS

SEQ ID NO: 197

Cuphea hookeriana 8:0- and 10:0-ACP specific thioesterase (FatB2) mRNA, complete cds., Genbank accession no. U39834, VERSION U39834.1

```
MVAAAASSAF  FVPAPGASP  KPGKFGNWPS  SLSPSFKPKS  IPNGGFQVKA  NDSAHPKANG
SAVSLKSGSL  NTQEDTSSSP  PPRITFLHQLP  DWSRLLTAIT  TVFVKSKRPD  MHDRKSKRPD
MLVDSFGLES  TVQDGLVFRQ  SFSIRSYEIG  TDRTASIELT  MNHLQETSLN  HCKSTGILLD
GFGRILEMCK  RDLIWVVIKM  QIKVNRYPAW  GDTVEINTRF  SRLGKIGMGR  DWLISDCNTG
EILVRATSAY  AMMNQKTRRL  SKLPYEVHQE  IVPLFVDSVP  IEDSDLKVHK  FKVKGTGDSIQ
KGLTPGWNDL  DVNQHVSINV  YIGWILESM  TEVLETQELC  SLALEYRREC  GRDSVLESVT
AMDPSKVGVR  SQYQHLLRLE  DGTAVNGAT  EWRPKNAGAN  GAISTGKTSN  GNSVS
```

SEQ ID NO: 198

acyl-ACP thioesterase FatA [Ricinus communis] (castor bean) Genbank ACCESSION ABS30422
VERSION ABS30422.1

```
MLKVPCCNAT  DPIQSLSSQC  RFLTHFNNRP  YFTRRPSIPT  FFSSKNSSAS  LQAVVSDISS
VESAAACDsla  NRLRLGKLTE  DGFYSYKEKFI  VRSYEVGINK  TATVETIANL  LQEVGCNHAQ
SVGFSTDGFA  TTSMRKMHL  IWVTARMHIE  IYKPAWSDV  VEVETWCQSE  GRIGTRRDWI
LTDYATGQII  GRATSKWMM  NQDTRRLQKV  TDDVREEYLV  FCPRELRLAF  PEENNRSSKK
ISKLEDPAQY  SKLGLVPRRA  DLDMNQHVNN  VTYIGWVLES  IPQEIIDTHE  LQTITLDYRR
ECQHDDIVDS  LTSVEPSEN  EAVSELRGTN  GSATTTAGDE  DCRNFLHLLR  LSGDGLLEINR
GRTEWRKKSA  R
```

SEQ ID NO: 199

stearoyl-acyl-carrier protein desaturase [Ricinus communis] Genbank ACCESSION AAA74692
VERSION AAA74692.1

```
FRQITKNQKK  KVRKKTMAK  LNPFLSQTQK  LPSFALPPMA  STRSPKFYMA  STLKSGSKEV
ENLKKPFMP  REVHVQVTHS  MPPQKIEIFK  SLDNWAEENI  LVHLKPVEKC  WQPQDFLPDP
ASDGFDEQVR  ELRERAKEIP  DDYFVVLVGD  MITEEALPTY  QTMLNTLDGV  RDETASPTS
WAIWTRAWTA  EENRHGDLLN  KYLYLSGRVD  MRQIEKTIQY  LIGSGMDPRT  ENSPYLGFYI
TSFQERATFI  SHGNTARQAK  EHGDIKLAQI  CGTIAADEKR  HETAYTKIVE  KLFEIDPDGT
VLAEFADMRR  KISMPAHLMY  DGRDDNLFDH  FSAVAQRLGV  YTAQDYADIL  EFLVGRWKVD
KLTGLSAEQ  KAQDYVCRLP  PRIIRLEERA  QGRAKEAPTM  PFSWIFDRQV  KL
```

SEQ ID NO: 200

stearoyl-ACP desaturase [Olea europaea] Genbank ACCESSION AAB67840 VERSION
AAB67840.1

```
MALKLCFPPH  KMPSPFDARI  RSHRVFMAST  IHSPSMEVGK  VKKPFTPPRE  VHVQVTHSLA
PEKREIFNSL  NNWAQENILV  LLKDVDKCWQ  PSDFLPDSAS  EGFDEQVMEL  RKRCKEIPDD
YFIVLVGDM  TEEALPTYQT  MLNTLDGVRD  ETGASLTPWA  IWTRAWTAE  NRHGDLLNKY
LYLSGRVDM  QIEKTIQYLI  GSGMDPRTEN  NPYLGFYITS  FQERATFISH  GNTARLAKEH
GDLKLAQICG  IIAADEKRHE  TAYTKIVEKL  FEIDPDGTVL  ALADMMRKKV  SMPAHLMYDG
QDDNLFENFS  SVAQRLGVYT  AKDYADILEF  LVGRWDIEKL  TGLSGEGRKA  QDYVCTLPPR
IRRLEERAQS  RVKKASATPF  SWIFGREINL
```

SEQ ID NO: 201

Myristoyl-acyl carrier protein thioesterase, chloroplastic; Cuphea hookeriana, Genbank ACCESSION Q39513 VERSION Q39513.1

```
MVATAASSAF  FPLPSADTSS  RPGKLGKPKS  SLSPLKPKST  PNGGLQVKAN  ASAPPKINGS
PVGLKSGGLK  TQEDAHSAAP  PRTFINQLPD  WSMLLAAITT  VFLAAEKQWM  MLDWKPKRPD
MLVDPFGLGS  IVQDGLVFRQ  NFSIRSYEIG  ADRTASIEIV  MNHLQETALN  HVKIAGLSND
```

GFGRTPEMYK RDLIWVAKM QVMVNRYPW GDTVEVNTWV AKSGKNGMRR DWLISDCNTG
 EILTRASSVW VMNQKTRRL SKIPDEVRNE IEPHFVDSPP VIEDDDRKLK KLDEKTADSI
 RKGLTPRWND LDVNQHVNNV KYIGWILEST PPEVLETQEL CSLTLEYRRE CGRESVLESL
 TAMDPSSGGY GSQFQHLLRL EDGGEIVKGR TEWRPKNGVI NGVVPTGESS PGDYS

SEQ ID NO: 202

thioesterase FatB1-1 [*Cuphea hookeriana*], Genbank ACCESSION AAC72882 VERSION AAC72882.1

MVATAASSAF FVPSSPVTSS RPGKPGNGSS SFSPKPKFV ANGGLQVKAN ASAPPKINGS
 SVGLKSCSLK TQEDTPSAPA PRTFINQLPD WSMLLAATT VFLAAEKQWM MLDWKPKRPD
 MLVDPFGLGS IVQHGLVFRQ NFSIRSYEIG ADRTASITV MNHLQETALN HVKSAGLMND
 GFGRTPEMYK KDLIWVAKM QVMVNRYPW GDTVEVNTWV AKSGKNGMRR DWLISDCNTG
 EILTRASSVW VMNQKTRRL SKIPDEVRNE IEPHFVDSPP VIEDDDRKLK KLDDKTADSI
 RKGLTPRWND LDVNQHVNNV KYIGWILEST PPEVLETQEL CSLTLEYRRE CGRESVLESL
 TAADPSGKGF GSQFQHLLRL EDGGEIVKGR TEWRPKTAGI NGAIPSGETS PGDS

SEQ ID NO: 203

Uc FatB2 [*Umbellularia californica*], Genbank ACCESSION AAC49001 VERSION AAC49001.1

MVTSLASAF FSMKAVMLAP DGSGIKPRSS GLQVRAGKEQ NSCKMINGTK VKDTEGLKGR
 STLHGWSMPL ELITTIFSAA EKQWTLNLSK PPQLDDHLG LHGLVFRRTF AIRCSEVGPD
 RSTSIVAVMN YLQEAACNHA ESLGLLDGDF GETLEMSRRD LIWVVRTHV VVERYPAWGD
 TVEVEAWIGA AGNIGMRRHF LVRDCKTGH LARCTSVSVM MNMRTRRLSK IPQEVGEID
 PLFIEKFAVK EGEIKKLQKF NDSTADYIQG GWTPRWNDLD VNQHVNNIKY VGWIFKSVPD
 SIYENHHLSS ITLEYRRECT RGRALQSLTT VCGGSSEAGI ICEHLLQLED GSEVLRGRTD
 WRPKRTDSFE GISERFPQQE PHN

SEQ ID NO: 204

palmitoyl-ACP thioesterase [*Elaeis guineensis*], (African oil palm). Genbank ACCESSION ABD83939 VERSION ABD83939.1

MVASIAASAF FPTPSFSPTA SAKASKTIGE GSENLARGI IAKPTSSSAA MQGKVMQAV
 SKINGAKVGL KAESQKAEED AAPSSAPRTF YNQLPDWSVL LAAVTITFLA AEKQWTLDDW
 KPRRPDMLTD AFSLGKIVQD GLVFKQNFIS RSYEIGADRT ASIETLMNHL QETALNHVRS
 AGLMGDGFGA TPEMSKRNL I WVTKVRVLI EHYPWGDV EVDTWVGPA KNGMRRDWHV
 RDHRTGQTIL RATRVWVMN KNTRKLSKVP EEVRAEIGPY FVERAAIVDE DSRKLPKLDE
 DTTDYVKKGL TPRWSDLDVN QHVNNVKYIG WILESAPISI LENHELASMT LEYRRECGRD
 SVLQSLTAVA NDCTGGLPEA SIECQHLLQL ECGAEIVRGR TQWRPRRASG PTSAGSA

SEQ ID NO: 205

palmitoyl-acyl carrier protein thioesterase [*Elaeis guineensis*] (African oil palm) Genbank ACCESSION AAD42220 VERSION AAD42220.2

MVASIVAWAF FPTPSFSPTA SAKASKTIGE GSENLNVRGI IAKPTSSSAA KQGKVMQAV
 PKINGAKVGL KAESQKAEED AAPSSAPRTF YNQLPDWSVL LAAVTITFLA AEKQWTLDDW
 KPRRPDMLTG AFSLGKIVQD GLVFRQNFIS RSYEIGADRT ASIETLMNHL QETALNHVRN
 AGLLDGDFGA TPEMSKRNL I WVTKMQLVI EHYPWGDV EVDTWVGASG KNGMRRDWHV
 RDYRTGQTIL RATSIVWMD KHTRKLSKMP EEVRAEIGPY FMEHAAIVDE DSRKLPKLDD
 DTADYIKWGL TPRWSDLDVN QHVNNVKYIG WILESAPISI LENHELASMT LEYRRECGRD
 SVLQSLTAVA NDCTGGLPEA SIECQHLLQL ECGAEIVRGR TQWRPRRASG PTSAGSA

SEQ ID NO: 206

palmitoyl-ACP thioesterase [*Elaeis guineensis*]. Genbank ACCESSION AAL15645 VERSION AAL15645.1

```
FYNQLPDWSV LLAAVTTIFL AAEKQWTLDD WKPRRPDMLA DAFGLGKIVQ DGLVFKQNF'S
IRSYEIGADR TASIETLMNH LQETALNHVR SAGLMGDGFG ATPEMSKRNL IWVVTCKMRVL
IKHYPSWGDV VEVDTWVGPT GKNGMRRDWH VRDHRTGQTI LRATSVVWMM NKNTRKLSKV
PEEVRAELGP YFVERAAIVD EDSRKLPKLD EDTTDYIKKG LTPRWSDLDV NQHVNNVKYI
GWILESAPIS FLENHELASM SLEYRRECGR DSVLQSLTAV SNDLTDGLPE AGIECQHLLQ
LECGTELKVG RTEWRPKHSL ALRNMGPTPG GSA
```

SEQ ID NO: 207

FATB [*Populus tomentosa*]. Genbank ACCESSION ABC47311 VERSION ABC47311.1

```
MVATAATSSF FVPVSPPGDA KSSKVGSGSA SLGGIKSKSA SSGALQVKAN AQAPPKINGS
PVGLTASVET AKKEDVVSSP APRTFINQLP DWSMLLAAIT TMFLAAEKQW MMLDWKPKRA
DMLIDPFGIG RIVQDGLVFS QNFSIRSIEI GADRTASIEI LMNHLQETAL NHVKTAGLLG
DGFGSTPEMS KRNLIIWVTR MQILVDRYPT WGDVVHVDTW VSASGKNGMR RDWLVRDAKT
GETLTRASSL WVMNKLTRR LSKIPEDVRG EIEPYFLNSD PVVNEDSTKL PKLDDKTADY
IRKGLTPRWV DLDVNQHVNN VKYIGWILES APPPILESHE LAAITLEYRR ECGRDSVLQS
LTAVSGAGIG NLGGPGKVEC QHLLRHEDGA EIVRGRTEWR PKHANNFGMM GGQMPADESG
A
```

SEQ ID NO: 208

FATB (fatty acyl-ACP thioesterases B); acyl carrier/ acyl-[acyl-carrier-protein] hydrolase [*Arabidopsis thaliana*]. Genbank ACCESSION NP_172327 VERSION NP_172327.1

```
MVATSATSSF FVPVSSSLDP NGKGNKIGST NLAGLNSAPN SGRMKVKPNA QAPPKINGKK
VGLPGSVDIV RTDTETSSHP APRTFINQLP DWSMLLAAIT TIFLAAEKQW MMLDWKPKRS
DMLVDPFGIG RIVQDGLVFR QNFSIRSIEI GADRSASIEI VMNHLQETAL NHVKTAGLLG
DGFGSTPEMF KKNLIWVTR MQVVVDKYPT WGDVVEVDTW VSQSGKNGMR RDWLVRDCNT
GETLTRASSV WVMNKLTRR LSKIPEEVRG EIEPYFVNSD PVLAEDSRKL TKIDDKTADY
VRSGLTPRWS DLDVNQHVNN VKYIGWILES APVGIMERQK LKSMTLEYRR ECGRDSVLQS
LTAVTGCDIG NLATAGDVEC QHLLRLQDGA EVVRGRTEWS SKTPTTTWGT AP
```

SEQ ID NO: 209

acyl-(acyl carrier protein) thioesterase [*Arabidopsis thaliana*]. Genbank ACCESSION CAA85387 VERSION CAA85387.1

```
MVATSATSSF FVPVSSSLDP NGKGNKIGST NLAGLNSAPN SGRMKVKPNA QAPPKINGKR
VGLPGSVDIV RTDTETSSHP APRTFINQLP DWSMLLAAIT TIFLAAEKQW MMLDWKPKRS
DMLVDPFGIG RIVQDGLVFR QNFSIRSIEI GADRSASIEI VMNHLQETAL NHVKTAGLLG
DGFGSTPEMF KKNLIWVTR MQVVVDKYPT WGDVVEVDTW VSQSGKNGMR RDWLVRDCNT
GETLTRASSV WVMNKLTRR LSKIPEEVRG EIEPYFVNSD PVLAEDSRKL TKIDDKTADY
VRSGLTPRWS DLDVNQHVNN VKYIGWILES APVGIMERQK LKSMTLEYRR ECGRDSVLQS
LTAVTGCDIG NLATAGDVEC QHLLRLQDGA EVVRGRTEWS SKTPTTTWGT AP
```

SEQ ID NO: 210

acyl-(acyl carrier protein) thioesterase [*Arabidopsis thaliana*]. Genbank ACCESSION CAA85388 VERSION CAA85388.1

```
MVATSATSSF FVPVSSSLDP NGKGNKIGST NLAGLNSTPN SGRMKVKPNA QAPPKINGKR
VGLPGSVDIV RTDTETSSHP APRTFINQLP DWSMLLAAIT TIFLAAEKQW MMLDWKPKRS
DMLVDPFGIG RIVQDGLVFR QNFSIRSIEI GADRSASIEI VMNHLQETAL NHVKTAGLLG
DGFGSTPEMF KKNLIWVTR MQVVVDKYPT WGDVVEVDTW VSQSGKNGMR RDWLVRDCNT
```

GETLTRASSV WVMNKLTRR LSKIPPEVRG EIEPYFVNSD PVLAEDSRKL TKIDDKTADY
 VRSGLTFRWS DLDVNQHVNN VKYIGWILES APVGIMERQK LKSMTLEYRR ECGRDSVLQS
 LTAVTGCDIG NLATAGDVEC QHLLRLQDGA EVVRGRTEWS SKTPTTTWGT AP

SEQ ID NO: 211

Gossypium hirsutum Genbank ACCESSION Q9SQI3

MVATAVISAF FPGTSSPDSS DSKNKKLGSI KSKPSVSSGS LQVKANAQAP PKINGTVAST
 TPVEGSKNDD GASSPPPTF INQLPDWSML LAAITTIFLA AEKQWMLLDW KPRRPDMVID
 PFGIGKIVQD GLVFSQNFISI RSYEIGADQT ASIETLMNHL QETAINHCRS AGLLGEGFGA
 TPMECKKNLI WVVTRMQVVV DRYPTWGDVV QVDTWVSASG KNGMRRDWLV SNSETGEILT
 RATSVMNMMN KLTRRLSKIP EEVRGEIEPF FMNSDPVLAE DSQKLVLDD STAETHVCKGL
 TPKWSDLDVN QHVNNVKYIG WILESAPLPI LESHELSALT LEYRRECGRD SVLQSLTTVS
 DSNTENAVNV GEFNCQHLLR LDDGAEIVRG RTRWRPKHAK SSANMDQITA KRA

SEQ ID NO: 212

acyl-ACP thioesterase [*Cuphea lanceolata*]. Genbank ACCESSION CAA54060 VERSION
 CAA54060.1

MVATAASSAF FPLPSPDTSS RPGKLGNGSS SLSPLKPKFV ANAGLKVKAS ASAPPKINGS
 SVGLKSGSLK TQEDTPSVPP PRTFINQLPD WSMLLAIAIT VFLAAEKQWM MLDWKPKRPD
 MLVDPFGLGS IVQGGLVFRQ NFSIRSYEIG ADRTASIEIV MNHLQETALN HVKSAGLLND
 GFGRTPPEMFK RDLIWVAKM QVMVNRYPW GDTVEVNTWV AKSGKNGMRR DWLISDCNTG
 EILTRASSVW VMNQKTRKL SKIPDEVRHE IEPHFIDCAP VIEDDDRKLK KLDEKTADSI
 RKGLTPKWND LDVNQHVNNV KYIGWILEST PQEVLETQEL SSLTLEYRRE CGRESVLESL
 TAVDSSGKGF GSQFQHLLRL EDGGEIVKGR TEWRPKTAGV NGAIASGETS HGDS

SEQ ID NO: 213

acyl-acyl carrier protein thioesterase [*Cuphea calophylla* subsp. *mesostemon*]. Genbank
 ACCESSION ABB71581 VERSION ABB71581.1

MVATAASSAF FVPSPDTSS RPGKLGNGSS SLSPLKPKLM ANGGLQVKAN ASAPPKINGS
 SVGLKSGSLK TQEDTPSAPP PRTFINQLPD WSMLLAIAIT VFLAAEKQWM MLDWKPKRPD
 MLVDPFGLGR IVQDGLVFRQ NFSIRSYEIG ADRTASIEIV MNHLQETALN HVKSAGLLND
 GFGRTPPEMYK RDLIWVAKM QVMVNRYPW GDTVEVNTWV AKSGKNGMRR DWLISDCNTG
 EILTRASSVW VMNQKTRRL SKIPDEVRHE IEPHFVDSAP VIEDDDRKLK KLDEKTADSI
 RKGLTPKWND LDVNQHVNNV KYIGWILEST PPEVLETQEL CSLTLEYRRE CGRESVLESL
 TAVDPSGKGS GSQFQHLLRL EDGGEIVKGR TEWRPKTAGI NGPIASGETS PGDS

SEQ ID NO: 214

palmitoyl-acyl carrier protein thioesterase [*Gossypium hirsutum*](upland cotton). Genbank
 ACCESSION AAD01982 VERSION AAD01982.1

TAVTSAFFPV TSSPDSSDSK NKKLGSISK PSVSSGSLQV KANAQAPPKI NGTVASTTPV
 EGSKNDDGAS SPPPTFINQ LPDWSMLLAA ITTIFLAAEK QWMLLDWKPR RPDVIDPFG
 IGKIVQDGLV FSQNFISIRY EIGADQTASI ETLMNHLQET AINHCRSAGL LGEGFGATPE
 MCKKNLIWVV TRMQVVVDY PTWGDVVQVD TWVSASGKNG MRRDWLVSN ETGEILTRAT
 SVWMMNKLK RLRSKIPPEV RGEIEPFFMN SDPVLAEVSQ KLVKLDDSTA EHVCKGLTPK
 WSDLDVNQHV NNVKYIGWIL ESAPLPILES HELSALTLEY RRECGRDSVL QSLTTVSDSN
 TENAVNVGEF NCQHLLRLDD GAEIVRGTR WRPKHAKSSA NMDQITAKRA

SEQ ID NO: 215

hypothetical protein VITISV_008781 [*Vitis vinifera*]. Genbank ACCESSION CAN81819 VERSION
 CAN81819.1

MVATAATSFAF FAVASPSDDP DAKPSTKPGX GSAILRGIKS RNAPSGSLQV KANAQAPPKI
 NGTTVGYTSS AEGVKIEDDM SSPPPRTFIN QLPDWSMLLA AITTIFLAAE KQWMLLDWKP
 RRSDMLIDPF GLGKIVQDGL VFRQNFIRS YEIGADRTAS IETLMNHLQE TALNHVRTAG
 LLGDGFGSTP EMSIRNLIWV VTRMQVVVDR YPTWGDVVQV DTWVCASGKN GMRRDWIIRD
 CKTGETLTRA SSVWVMMNKQ TRRLSKIPDA VRAEIEPYFM DSAPIVDEDG RKLPLDDST
 ADYIRTGLTP RWSDLDVNQH VNNVKYIGWI LESAPLPILE SHELSSMTLE YRRECGRDSV
 LQSLTAVSGT GVGNNLDCGN VECQHLLRLE EGAEIVEGKD

SEQ ID NO: 216

palmitoyl-ACP thioesterase [Brassica juncea]. Genbank ACCESSION ABI18986 VERSION
 ABI18986.1

MVATSATSLE FPLPSSSLDP NXKTNNRVTS TNFAGLGPTP NSGGRMKVKP NAQAPPKING
 KKVGLPGSVE IETSQQQQPA PRTFINQLPD WSMLLAATT VFLAAEKQWM MLDWKPRRSD
 MIMEPFGLGR IVQDGLVFRQ NFSIRSYEIG ADRSASIETV MNHLQETALN XVKTAGLLGD
 GFGSTPEMVK KXLIWVVTM QVVVDITYPTW GDVVEVDTWV SKSGKNGMRR DWLVRDNGTG
 QILTRASSVW VMMNKLTRRL SKIPEEVRGE IEPYFVDFDP VLAEDSRKLT KLDDKTADYV
 RSGLTFRWSD LDVNQHVNNV KYIGWILESA PVGMESQKL KSMTLEYRRE CGRDSVLQSL
 TAVSGCDIGN LGTAGEVECO HLLRLQDGAE VVRGRTEWSS KPTTTTWDIT P

SEQ ID NO: 217

chloroplast stearyl/oleoyl specific acyl-acyl carrier protein thioesterase precursor [Madhuca longifolia
 var. latifolia]. Genbank ACCESSION AAX51637 VERSION AAX51637.1

INQLPDWSML LAAITTIFLA AEKQWMLLDW KPKRPDMLID PFGLGKIVQD GLVFRQNFST
 RSYEIGADRT ASIETLMNHL QETALNHVKT AGLLGDGFGV TPMECKKNLI WVVAKMQVLV
 DRYPTWGDVV QVDTWVAASG KNGMRRDWLV RDFETGDILT KASSVWMMN RETRLSKIP
 DEVRLEIGSY FVDSPPVLDE DGRKLPLKLE STADHIRTGL TPRWNLQVN QHVNNVKYIG
 WILESAPQPI LENHELASMT LEYRRECCKD SVLQSLTGVT SGGVGGLADS GHVECOHLLR
 LEGGAEIVKG RTEWRPKYAN RLGCLDQLPA GST

SEQ ID NO: 218

hypothetical protein Osl_08261 [Oryza sativa Indica Group]. Genbank ACCESSION EAY86877
 VERSION EAY86877.1

MAGSLAASAF FPGPGSSPAA SARSSKSAAV TGELPENLSV RGIVAKPNPP PAAMQVKAQA
 QTLPKVNGTK VNLKTVKPD M EETVPYSAPK TFYNQLPDWS MLLAAITTIF LAEEKQWTL
 DWKPKKPDML VDTFGFGRII QDGMVFRQNF MIRSIEIGAD RTASIEITLMN HLQETALNHV
 RTAGLLGDGF GATPEMSKRN LIWVVSQIQL LVEQYPAWGD TVQVDTWVAA AGKNGMRRDW
 HVRDYNNGRT ILRATSVWVM MHKKTRRLSK MPDEVRAEIG PYFNDRSAIT EEQSEKLAKT
 GNKVGDDATE QFIRKGLTPR WGDLDVNQHV NNVKYIGWIL ESAPISVLEK HELASMTLDY
 RKECGRDSVL QSLTTVSGEC TSIGADKQAS AIQCDHLLQL ESGADIVKAH TEWRPKRSHA
 AAENA

SEQ ID NO: 219

Os11g0659500 [Oryza sativa Japonica Group]. Genbank ACCESSION NP_001068400 VERSION
 NP_001068400.1

MAGSLAASAF FPGPGSSPAA SARSSKNAAV TGELPENLSV RGIVAKPNPP PAAMQVKAQA
 QTLPKVNGTK VNLKTVKPD M EETVPYSAPK TFYNQLPDWS MLLAAITTIF LAEEKQWTL
 DWKPKKPDML VDTFGFGRII QDGMVFRQNF MIRSIEIGAD RTASIEITLMN HLQETALNHV
 RTAGLLGDGF GATPEMSKRN LIWVVSQIQL LVEQYPAWGD TVQVDTWVAA AGKNGMRRDW
 HVRDYNNGRT ILRATSVWVM MHKKTRRLSK MPDEVRAEIG PYFNDRSAIT EEQSEKLAKT
 GNKVGDDATE QFIRKGLTPR WGDLDVNQHV NNVKYIGWIL ESAPISVLEK HELASMTLDY

RKECGRDSVL QSLTTVSGEC TSIGADKQAS AIQCDHLLQL ESGADIVKAH TEWRPKRSHA
AAENA

SEQ ID NO: 220

hypothetical protein Osl_21596 [Oryza sativa Indica Group]. Genbank ACCESSION EAY99617
VERSION EAY99617.1

MAGSLAASAF FPVPGSSPAA SARSSKNTTG ELPENLSVRG IVAKPNPSPG AMQVKAQAQA
LPKVNKTKVN LKTTSPDKED IIPYTAPKTF YNQLPDWSML LAAVTTIFLA AEKQWTLDDW
KPKKPDMLAD TFGFGRIIQD GLVFRQNFLI RSYEIGADRT ASIETLMNHL QETALNHVKT
AGLLGDGFGA TPMSKRNL I WVSQIQLLV ERYPSWGDV QVDTWVAAAG KNGMRRDWHV
RDYNSGQITL RATSVMVMN KNTRRLSKMP DEVRAEIGPY FNGRSAISEE QGEKLPKPGT
TFDGAATKQF TRKGLTPKWS DLDVNQHVNN VKYIGWILES APISILEKHE LASMTLDYRK
ECGRDSVLQS LTTVSGECDD GNTESIQC D HLLQLESGAD IVKAHTEWRP KRAQGEENMG
FFPAESA

SEQ ID NO: 221

acyl-(ACP) thioesterase type B [Cuphea lanceolata]. Genbank ACCESSION CAB60830
VERSION CAB60830.1

MVAAAATSAF FPVPAPGTSP KPGKSGNWPS SLSPTFKPKS IPNAGFQVKA NASAHPKANG
SAVNLKSGSL NTQEDTSSSP PPRAFLNQLP DWSMLLTAIT TVFVAAEKQW TMLDRKSKRP
DMLVDSVGLK SIVRDGLVSR QSFLIRSYEI GADRTAS IET LMNHLQETSI NHCKSLGLLN
DGFGRTPGMC KNDLIWVLTK MQIMVNRYPT WGD TVEINTW FSQSGKIGMA SDWLISDCNT
GEILIRATSV WAMNQKTRR FSRLPYEVQR ELTPHFVDSP HVIDEDNDQKL HKFDVKTGDS
IRKGLTPRWN DLDVNQHVSN VKYIGWILES MPEVLETQE LC SLTVEYRR ECGMDSVLES
VTAVDPSENG GRSQYKHLR LEDGTDIVKS RTEWRPKNAG TNGAISTSTA KTSNGNSAS

SEQ ID NO: 222

acyl-ACP thioesterase [Myristica fragrans]. Genbank ACCESSION AAB71730 VERSION
AAB71730.1

PDWSMLLAAI TTIFLAAEQ WTNLDWKPRR PDMLVDFDPF SLGRFVQDGL IFRQNF SIRS
YEIGADRTAS IETLMNHLQE TALNHVRCIG LLDDGFGSTP EMTRRDLI WV VTRMQVLVDR
YPSWGDVIEV DSWVTPSGKN GMKREWF LRD CKTGEILTRA TSVWVMNMNR TRRLSKIPEE
VRVEIEPYFV EHGVLDEDSR KLPKLN DNTA NYIRRLAPR WSDLDVNQHV NNVKYIGWIL
ESVPSSLLES HELYGMTLEY RKECGKDGLL QSLTAVASDY GGSLEAGVE CDHLLRLLEDG
SEIMRGKTEW RPKRAANTTY FGSVDDIPPH PIYIYIYIYI YIYIYWVGSS CSGSSTMSR
TR

SEQ ID NO: 223

Cinnamomum camphora acyl-ACP thioesterase. Genbank ACCESSION U31813 VERSION
U31813.1

MATTSLASAFCSMKAVMLARDGRGMKPRSSDLQLRAGNAQTS LK
MINGTKFSYTESLKKLPDWSMLFAVITTI FSAAEKQW TNLEWKPKPNPPQLDDHFGP
HGLVFRRTFAIRSYEVGPDRSTSIVAVMNLQEAALNHAKSVGILGDGFGTTLEMSKR
DLI WVVRKTHVAVERYPAWGD TVEVECWVGASGNNRRHDFLVRDCKTGEILTRCTSL
SVMNTRTRRLSKIPEEVRGEIGPAFIDNVAVKDEEIKKPQKLN DSTADYIQGGTLP
WNL DLDINQHVNNIKYVDWILETV PDSIFESH HISSFTIEYRRECTMDSVLQSLTTVSG
GSSEAGLVCEHLLQLEGGSEVLRAKTEWRPKLTD SFRGISVIPAESSV

SEQ ID NO: 224

acyl-ACP thioesterase [Myristica fragrans]. Genbank ACCESSION AAB71729

VERSION AAB71729.1

MVATSAASAF FVPASPSPVK PSMMLGGGGG SDNLDARGIK SRPASSGGLQ VKANAHTVPK
INGNKAGLLT PMESTKDEDI VAAPTVPKR TFINQLPDWS MLLAAITTIF LAAEKQWTNL
DWKPRRPDML VDFDPFSLGR FVQDGLIFRQ NFSIRSYEIG ADRTASIELT MNHLQETALN
HVRIGLLDD GFGSTPEMTR RDLIWVVTM QVLVDYRPSW GDVIEVDSWV TPGKNGMKR
EWFLRDCKTG EILTRATSVW VMMNKRTRRL SKIPEEVRE IEPYFVEHGV LDEDSRKLPK
LNDNTANYIR RGLAPRWSDL DVNQHVNNVK YIGWILESVP SSLLESHELY GMTLEYRKEC
GKDGLLQSLT AVASDYGGGS LEAGVECDHL LRLEDGSEIM RGKTEWRPKR AANTTYFGSV
DDIPPANNA

SEQ ID NO: 225

pSZ1420

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAGAA GAGTTGGTAG CTCTTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTGTGTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAGGAT CTTCACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCGTCGTG TAGATACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAAGT GTCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCTCTCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTTCTTCTA CTGTATGCC ATCCGTAAGA TGCTTTTCTG
961 TGA CTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCCGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCCTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAA AGTCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCAGGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGCTG TAAGCGGATG
1561 CCGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CCGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATC AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCG CCCTCTCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCG TCGTGCCTGT CGCTGATGTC
1981 CATCACCAGG TCCATGAGGT CTGCTTTCG CCGGCTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCTGGTCT CAGGGTCTTG ACGTGGTCTG GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCGGCC TCCAACCTGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCAGCA GCCTTGCTA GGCAGAATCC CTACAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCTCTGCT GCCGCCGCTT CTCCGCGACG CTTCTTTCCA
2341 GCACCGTGAT GGCAGGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGCGCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGGCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACCTTT TTGGCGAAGG CATCGGCCTC GGCCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGAGCGCT

2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCTCTT TCGTTTCAGT CACAACCCGC AAACCTCTAGA ATATCAATGC
2941 TGCTGCAGGC CTTCCTGTTC CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCCAACGG CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCCGAACGA CACCGTCTGG GGGACGCCCT TGTTCCTGGG CCACGCCACG TCCGACGACC
3181 TGACCAACTG GGAGGACCAG CCCATCGCCA TCGCCCCGAA GCGCAACGAC TCCGGCGCCT
3241 TCTCCGGCTC CATGGTGGTG GACTACAACA ACACCTCCGG CTTCTTCAAC GACACCATCG
3301 ACCCGCGCCA GCGCTGCGTG GCCAICTGGA CCTACAACAC CCCGGAGTCC GAGGAGCAGT
3361 ACATCTCCTA CAGCTGGAC GCGGGCTACA CCTTCACCGA GTACCAGAAG AACCCCGTGC
3421 TGCCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCTCCACAGA
3481 AGTGGATCAT GACCGCGGCC AAGTCCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCGCGT TCGCCAACGA GGGCTTCCTC GGCTACCAGT
3601 ACCAGTGCCC CGGCCTGATC GAGGTCCCCA CCGAGCAGGA CCCAGCGAAC TCCCTATGGG
3661 TGATGTTTAT CTCCATCAAC CCCGGCGCCC CGGCCGCGCG CTCCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CCTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCTTGGGCAT CGCGTGGGCC TCCAAC TGCG AGTACTCCGC CTTCTGTGCC ACCAACCCT
3901 GGCGCTCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCCG AGCCGATCCT GAACATCAGC AACGCCGGCC
4021 CTGGAGCCG GTTCGCCACC AACACACGCT TGACGAAGGC CAACAGTAC AACGTCGACC
4081 TGTCCAACAG CACCGGCACC CTGGAGTTCG AGCTGGTGTG CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT GCGGACCTCT CCTCTGGTT CAAGGGCCTG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCTC CTTCTTCTG GACCGCGGGA
4261 ACAGCAAGGT GAAGTTCGTG AAGGAGAACC CCTACTTCAC CAACCGCATG AGCGTGAACA
4321 ACCAGCCCTT CAAGAGCGAG AACGACCTGT CCTACTACAA GGTGTACGGC TTGCTGGACC
4381 AGAACATCCT GGAGCTGTAC TTCAACGAGC GCGACGTCGT GTCCACCAAC ACCTACTTCA
4441 TGACCACCGG GAACGCCCTG GGCTCCGTGA ACATGACGAC GGGGGTGGAG AACCTGTCTT
4501 ACATGGGACAA GTTCCAGGTG CGCGAGGTCA AGTGACAATT GGCAGCAGCA GCTCGGATAG
4561 TATCGACACA CTCTGGACGC TGGTCTGTGT ATGGACTGTT GCCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTTGTGTGT
4681 ACGCGCTTTT GCGAGTTGCT AGCTGCTGTG GCTATTTGCG AATACCACCC CCAGCATCCC
4741 CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGCTATC
4801 CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGCTCCGC
4861 CTGTATTCTC CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGAT CACGGGAAGT
4921 AGTGGGATGG GAACACAAAT GGAGGATCC GCGTCTCGAA CAGAGCGCGC AGAGAACGCG
4981 TGAAGGTCTC GCCTCTGTCT CACCTCAGCG CGGCATACAC CACAATAACC ACCTGACGAA
5041 TGCGCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGGTTCA CACACGTGCC ACGTTGGCGA
5101 GGTGGCAGGT GACAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 CGAATTCTTT TCTTGCCTTA TGACACTTCC AGCAAAAGGT AGGGCGGGCT GCGAGACGGC
5221 TTCCCGGGCG TGCAATGCAAC ACCGATGATG CTTTCACCCC CGAAGCTCC TTCGGGGCTG
5281 CATGGGCGCT CCGATGCCGC TCCAGGGCGA GCGCTGTTTA AATAGCCAGG CCCCCGATTG
5341 CAAAGACATT ATAGCGAGCT ACCAAAGCCA TATTCAAACA CCTAGATCAC TACCACTTCT
5401 ACACAGGCCA CTCGAGCTTG TGATCGCACT CCGCTAAGGG GCGCCTCTT CTTCTTCGTT
5461 TCAGTCACAA CCCGCAAACA CTAGTATGGC CACCGCATCC ACTTTCTCGG CGTTCAATGC
5521 CCGCTGCGGC GACCTGCGTC GCTCGGCGGG CTCCGGGCCC CGGCGCCCAG CGAGGCCCTT
5581 CCCCGTGC GC GGGCGCGCCA CCCAGGAGGA CGCCCACTCC GCGCGCCCCC CCGCACCTT
5641 CATCAACCAG CTGCCCAGCT GGTCCATGCT GCTGGCCGCC ATCACCACCG TGTTCTTGGC
5701 CGCCGAGAAG CAGTGGATGA TGCTGGACTG GAAGCCCAAG CGCCCCGACA TGCTGGTGGG
5761 CCCCTTCGGC CTGGGCTCCA TCGTGCAGGA CGGCCTGGTG TTCCGCCAGA ACTTCTCCAT
5821 CCGCTCTTAC GAGATCGGCG CCGACCGCAC CGCCTCCATC GAGACCGTGA TGAACACCT
5881 GCAGGAGACC GCCCTGAACC ACGTGAAGAT CGCCGGCCTG TCCAACGACG GCTTCGGCCG
5941 CACCCCGAG ATGTACAAGC GCGACCTGAT CTGGGTGGTG GCCAAGATGC AGGTGATGGT
6001 GAACCGCTAC CCCACCTGGG GCGACACCGT GGAGGTGAAC ACCTGGGTGG CCAAGTCCGG
6061 CAAGAACGGC ATGCGCCGCG ACTGGCTGAT CTCCGACTGC AACACCGGCG AGATCCTGAC
6121 CCGCGCCTCC TCCGTGTGGG TGATGATGAA CCAGAAGACC CGCCGCCTGT CCAAGATCCC
6181 CGACGAGGTG CGCAACGAGA TCGAGCCCCA CTTCTGTGAC TCCCCCCCCG TGATCGAGGA
6241 CGACGACCGC AAGCTGCCCC AGCTGGACGA GAAGACCGCC GACTCCATCC CGAAGGGCCT
6301 GACCCCCGCG TGGAACGACC TGGACGTGAA CCAGCACGTG AACAACTGTA AGTACATCGG
6361 CTGGATCCTG GAGTCCACCC CCCCCGAGGT GCTGGAGACC CAGGAGCTGT GCTCCCTGAC

```

6421 CCTGGAGTAC CGCCGCGAGT GCGGCCGCGA GTCCGTGCTG GAGTCCCTGA CCGCCATGGA
6481 CCCCTCCGGC GGGGGCTACG GCTCCAGTT CCAGCACCTG CTGCGCCTGG AGGACGGCGG
6541 CGAGATCGTG AAGGGCCGCA CCGAGTGGCG CCCCAAGAAC GGCCTGATCA ACGGCGTGGT
6601 GCCCACC GAGTCCTCCC CCGGCGACTA CTCCATGGAC TACAAGGACC ACGACGGCGA
6661 CTACAAGGAC CACGACATCG ACTACAAGGA CGACGACGAC AAGTGAATCG ATAGATCTCT
6721 TAAGGCAGCA GCAGCTCGGA TAGTATCGAC AACTCTGGA CGCTGGTCTG GTGATGGACT
6781 GTTGCCGCCA CACTTGCTGC CTTGACCTGT GAATATCCCT GCCGCTTTTA TCAAACAGCC
6841 TCAGTGTGTT TGATCTTGTT TGTACGCGCT TTGCGAGTT GCTAGCTGCT TGTGCTATTT
6901 GCGAATACCA CCCCAGCAT CCCCTTCCCT CGTTTCATAT CGCTTGCATC CCAACCGCAA
6961 CTTATCTACG CTGTCCTGCT ATCCCTCAGC GCTGCTCCTG CTCCTGCTCA CTGCCCCTCG
7021 CACAGCCTTG GTTTGGGCTC CGCCTGTATT CTCCTGGTAC TGCAACCTGT AAACCAGCAC
7081 TGCAATGCTG ATGCACGGGA AGTAGTGGGA TGGGAACACA AATGGAAGC TTAATTAAGA
7141 GCTCTTGTTT TCCAGAAGGA GTTGCTCCTT GAGCCTTTCA TTCTCAGCCT CGATAACCTC
7201 CAAAGCCGCT CTAATTGTGG AGGGGTTTCG AATTTAAAG CTTGGAATG CTGTTCGTGC
7261 GTCTGGAACA AGCCCAGACT TGTGCTCAC TGGGAAAAGG ACCATCAGCT CCAAAAACT
7321 TGCCGCTCAA ACCGCGTACC TCTGCTTTCG CGCAATCTGC CCTGTTGAAA TCGCCACCAC
7381 ATTCATATTG TGACGCTTGA GCAGTCTGTA ATTGCCTCAG AATGTGGAAT CATCTGCCCC
7441 CTGTGCGAGC CCATGCCAGG CATGTCGCGG GCGAGGACAC CCGCCACTCG TACAGCAGAC
7501 CATTATGCTA CCTCACAATA GTTCATAACA GTGACCATAT TTCTCGAAGC TCCCCAACGA
7561 GCACCTCCAT GCTCTGAGTG GCCACCCCCC GGCCCTGGTG CTTGCGGAGG GCAGGTCAAC
7621 CGGCATGGGG CTACCGAAAT CCCCAGCCGG ATCCCACCAC CCCC CGATG GGAAGAATCT
7681 CTCCCCGGGA TGTGGGCCCA CCACCAGCAC AACCTGCTGG CCCAGGCGAG CGTCAAACCA
7741 TACCACACAA ATATCCTTGG CATCGGCCCT GAATTCCTTC TGCCGCTCTG CTACCCGGTG
7801 CTTCTGTCCG AAGCAGGGGT TGCTAGGGAT CGCTCCGAGT CCGCAAACCC TTGTCGCGTG
7861 GCGGGGCTTG TTCGAGCTTG AAGAGCTCT AGAGTCGACC TGCAGGCATG CAAGCTTGGC
7921 GTAATCATGG TCATAGCTGT TTCTGTGTG AAATTGTTAT CCGCTCACAA TTCCACACAA
7981 CACACGAGCC GGAAGCATAA AGTGTAAGC CTGGGGTGCC TAATGAGTGA GCTAACTCAC
8041 ATTAATTGCG TTGCGCTCAC TGCCCGCTTT CCAGTCGGGA AACCTGTCGT GCCAGTGCA
8101 TTAATGAATC GGCCAACGCG CGGGGAGGAG CGGTTTTCGT ATTGGGCGCT CTTCGCTTC
8161 CTCGCTCACT GACTCGCTGC GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT CAGCTCACTC
8221 AAAGGCGGTA ATACGTTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC
8281 AAAAGGCCAG CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG
8341 GCTCCGCCCC CCTGACGAGC ATCACA AAAA TCGACGCTCA AGTCAGAGGT GGC GAAACCC
8401 GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT
8461 TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA CGGTGGCGCT
8521 TTCTCATAGC TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCTGG CCAAGCTGGG
8581 CTGTGTGCAC GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT
8641 TGAGTCCAAC CCGGTAAGAC ACGACTTATC GC

```

SEQ ID NO: 226
pSZ1417

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTGTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTA AAAATG AAGTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTTCG TCATCCATAG
481 TTGCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAAC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTG AAAAAAGCGG
841 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGAGTG TTACTACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGA CTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCCGG GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA

```

1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAA TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTTAT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTG AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCCG CCACTCTGTC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCGC TCGTGCCTGTC CGCTGATGTC
1981 CATCACCAGG TCCATGAGGT CTGCCTTGCG CCGGCTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCCTG ACGTGGTCGC GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCGGCC TCCAACGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGAAATC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCTCGCTT GCCGCCGCTT CTCCCGCACG CTCTTTTCCA
2341 GCACCGTGAT GGCGCGAGCC AGCGCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTCGCGCTT AGTGTGGCCA TCCTTTGAGC
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCTTCCC AGACCAATTC
2521 TGTCACCTTT TTGGCAAGG CATCGGCCTC GGCTTCGAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GGCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GGCGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCTCTT TCGTTTCAGT CACAACCCGC AAACGGCGCG CCATGCTGCT
2941 GCAGGCCCTT CTGTTCTGTC TGGCCGGCTT CGCCGCCAAG ATCAGCGCCT CCATGACGAA
3001 CGAGACGTCC GACCGCCCCC TGGTGCCTT CACCCCAAC AAGGGCTGGA TGAACGACCC
3061 CAACGGCCTG TGGTACGACG AGAAGGACGC CAAGTGGCAC CTGTACTTCC AGTACAACCC
3121 GAACGACACC GTCTGGGGGA CGCCCTTGTT CTGGGGCCAC GCCACGTCCG ACGACCTGAC
3181 CAACTGGGAG GACCAGCCCC TCGCCATCGC CCCGAAGCGC AACGACTCCG CGCCTTCTC
3241 CGGCTCCATG GTGGTGGACT ACAACAACAC CTCCGGCTTC TTCAACGACA CCATCCAGCC
3301 GCGCCAGCGC TGGTGGCCA TCTGGACCTA CAACACCCCG GAGTCCGAGG AGCAGTACAT
3361 CTCCTACAGC CTGGACGGCG GCTACACCTT CACCGAGTAC CAGAAGAACC CCGTGTGGC
3421 CGCCAACCTC ACCCAGTTCC GCGACCCGAA GGTCTTCTGG TACGAGCCCT CCCAGAAGTG
3481 GATCATGACC GCGGCCAAGT CCCAGGACTA CAAGATCGAG ATCTACTCCT CCGACGACCT
3541 GAAGTCTCTG AAGCTGGAGT CCGCGTTCGC CAACGAGGGC TTCCTCGGCT ACCAGTACGA
3601 GTGCCCGGCG CTGATCGAGG TCCCCACCGA GCAGGACCCC AGCAAGTCCT ACTGGGTGAT
3661 GTTCATCTCC ATCAACCCCG GCGCCCGGCG CGGCGGCTCC TTCAACAGT ACTTCGTGCG
3721 CAGCTTCAAC GGCACCCACT TCGAGGCCTT CGACAACCAG TCCCGCGTGG TGGACTTCGG
3781 CAAGGACTAC TACGCCCTGC AGACCTTCTT CAACACCGAC CCGACCTACG GGAGCGCCCT
3841 GGGCATCGCG TGGGCCTCCA ACTGGGAGTA CTCCGCCTTC GTGCCACCA ACCCCTGGCG
3901 CTCTCCATG TCCCTCGTGC GCAAGTTCTC CCTCAACACC GAGTACCAGG CCAACCCGGA
3961 GACGGAGCTG ATCAACCTGA AGGCCGAGCC GATCCTGAAC ATCAGCAACG CCGGCCCTG
4021 GAGCCGGTTC GCCACCAACA CCACGTTGAC GAAGGCCAAC AGCTACAACG TCGACCTGTC
4081 CAACAGCACC GGCACCCTGG AGTTCGAGC GGTGTACGCC GTCAACACCA CCCAGCAGT
4141 CTCCAAGTCC GTGTTGCGCG ACCCTCTCCT CTGGTTCAAG GGCCTGGAGG ACCCCGAGGA
4201 GTACCTCCGC ATGGGCTTCG AGGTGTCCGC GTCCTCCTTC TTCCTGGACC GCGGGAACAG
4261 CAAGGTGAAG TTCGTGAAGG AGAACCCTTA CTTACCAAC CGCATGAGCG TGAACAACCA
4321 GCCCTTCAAG AGCGAGAACG ACCTGTCTTA CTACAAGGTG TACGGCTTGC TGGACCAGAA
4381 CATCTGGAG CTGTACTTCA ACGACGCGCA CGTCGTGTCC ACCAACACCT ACTTCATGAC
4441 CACCGGGAAC GCCCTGGGCT CCGTGAACAT GACGACGGGG GTGGACAACC TGTCTACAT
4501 CGACAAGTTC CAGGTGCGCG AGGTCAAAGT ACAATTGGCA GCAGCAGTTC GGATAGTATC
4561 GACACACTCT GGACGCTGGT CGTGTGATGG ACTGTTGCCG CCACACTTGC TGCTTGACC
4621 TGTGAATATC CCTGCCGCTT TTATCAAACA GCCTCAGTGT GTTTGATCTT GTGTGTACGC
4681 GCTTTTGCGA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCCCAG CATCCCCTTC

4741 CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTATCT ACGCTGTCCT GCTATCCCTC
4801 AGCGCTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG CTCCGCTGT
4861 ATTCTCCTGG TACTGCAACC TGTAACCAG CACTGCAATG CTGATGCACG GGAAGTAGTG
4921 GGATGGGAAC ACAAATGGAG GATCCCCTGT CTCGAACAGA GCGCGCAGAG GAACGCTGAA
4981 GGTCTCGCCT CTGTCGCACC TCAGCGCGGC ATACACCACA ATAACCACCT GACGAATGCG
5041 CTTGGTTCTT CGTCCATTAG CGAAGCGTCC GGTTACACA CGTGCCACGT TGGCGAGGTG
5101 GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG GGATATCGAA
5161 TTCGGCCGAC AGGACGCGCG TCAAAGCTGC TGGTCGTGTA TGCCCTGGCC GGCAGGTCGT
5221 TGCTGCTGCT GGTTAGTGAT TCCGCAACCC TGATTTTGGC GTCTTATTTT GGCCTGGCAA
5281 ACGCTGGCGC CCGCGAGCCG GGCCGGCGGC GATGCGGTGC CCCACGGCTG CCGGAATCCA
5341 AGGGAGGCAA GAGCGCCCGG GTCAGTTGAA GGGCTTTACG CGCAAGGTAC AGCCGCTCCT
5401 GCAAGGCTGC GTGGTGGAAT TGGACGTGCA GGTCTGTGTG AAGTTCCTCC ACCGCTCAC
5461 CAGCGGACAA AGCACCAGGTG TATCAGGTCC GTGTCATCCA CTCTAAAGAA CTCGACTACG
5521 ACCTACTGAT GGCCCTAGAT TCTTCATCAA AAACGCCTGA GACACTTGCC CAGGATTGAA
5581 ACTCCCTGAA GGGACACCA GGGGCCCTGA GTTGTTCCTT CCCCCCTGG CGAGGTCCGA
5641 GCCAGGCTGT ACCTGTGATC GAGGCTGGCG GGAAAATAGG CTTCGTGTGC TCAGGTGATG
5701 GGAGGTGCAG GACAGCTCAT GAAACGCCAA CAATCGCACA ATTCATGTCA AGCTAATCAG
5761 CTATTTCCTC TTCACGAGCT GTAATTGTCC CAAAATTCTG GTCTACCGGG GGTGATCCTT
5821 CGTGACGGG CCCTTCCCTC AACCTTAGGT ATGCGCGCAT GCGGTGCGCG CGCAACTCGC
5881 GCGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGACCCCG GATGCGTGGC
5941 ACCTTTTGTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT CTGTCGCCAA
6001 CCCTAGGATC AGCGCGGTAG GATTTCGTAA TCATTCTGTC TGATGGGGAG CTACCGACTA
6061 CCCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC ATTCCATTG
6121 TGCCCAAGAC ATTTCAATTG GTTGCGAAGC GTCCCCAGTT ACGCTCACCT GTTTCCCGAC
6181 CTCCTTACTG TTCTGTGAC AGAGCGGGCC CACAGGCCGG TCGCAGCCAC TAGTATGGTG
6241 GCCACCGCCG CCTCTCCGC CTTCCTCCCG CTGCCCTCCG CCGACACCTC CTCCCGCCCC
6301 GGCAAGCTGG GCAACAAGCC CTCCTCCCTG TCCCCCTGA AGCCCAAGTC CACCCCAAC
6361 GGCGGCTGC AGGTGAAGGC CAACGCCTCC GCCCCCCCCA AGATCAACGG CTCCCCCGTG
6421 GGCCTGAAGT CCGGCGGCCT GAAGGGCGGC GCCACCCAGG AGGACGCCCA CTCCCGCCCC
6481 CCCCCCGCA CCTTCATCAA CCAGCTGCCG GACTGGTCCA TGCTGCTGGC CGCCATCACC
6541 ACCGTGTTCC TGGCCGCCGA GAAGCAGTGG ATGATGCTGG ACTGGAAGCC CAAGCGCCCC
6601 GACATGCTGG TGGACCCCTT CGGCCTGGGC TCCATCGTGC AGGACGGCCT GGTGTTCCGC
6661 CAGAACTTCT CCATCCGCTC CTACGAGATC GGCGCCGACC GCACCGCCTC CATCGAGACC
6721 GTGATGAACC ACCTGCAGGA GACCGCCCTG AACCACGTGA AGATCGCCGG CCTGTCCAAC
6781 GACGGCTTCG GCCGACCCCC CGAGATGTAC AAGCGCGACC TGATCTGGGT GGTGGCCAAG
6841 ATGCAGGTGA TGGTGAACCG CTACCCACCC TGGGGCGACA CCGTGGAGGT GAACACCTGG
6901 GTGGCCAGAT CCGGCAAGAA CGGCATGCGC CGCGACTGGC TGATCTCCGA CTGCAACACC
6961 GGCGAGATCC TGACCCGCGC CTCCTCCGTG TGGGTGATGA TGAACCAGAA GACCCGCGC
7021 CTGTCCAAGA TCCCGACGA GGTGCGCAAC GAGATCGAGC CCCACTTCGT GGACTCCCCC
7081 CCCGTGATCG AGGACGACGA CCGCAAGCTG CCAAGCTGG ACGAGAAGAC CGCCGACTCC
7141 ATCCGCAAGG GCCTGACCCC CCGCTGGAAC GACCTGGACG TGAACCAGCA CGTGAACAAC
7201 GTGAAGTACA TCGGTGGAT CCTGGAGTCC ACCCCCCCG AGGTGCTGGA GACCCAGGAG
7261 CTGTGCTCCC TGACCCTGGA GTACCGCCGC GAGTGGCGCC GCGAGTCCGT CTTGGAGTCC
7321 CTGACCGCCA TGGACCCCTC CGGCGCGCGC TACGGCTCCC AGTTCCAGCA CTGCTGCGC
7381 CTGGAGGACG GCGGCGAGAT CGTGAAGGGC CGCACCGAGT GGCGCCCCAA GAACGGCGTG
7441 ATCAACGGCG TGGTGCCAC CGGCGAGTCC TCCCCGGCG ACTACTCCAT GGACTACAAG
7501 GACCACGACG GCGACTACAA GGACCACGAC ATCGACTACA AGGACGACGA CGACAAGTGA
7561 ATCGATAGAT CTCTTAAGGC AGCAGCAGCT CGGATAGTAT CGACACACTC TGGACGCTGG
7621 TCGTGTGATG GACTGTTGCC GCCACACTTG CTGCCTTGAC CTGTGAATAT CCCTGCCGCT
7681 TTTATCAAAC AGCCTCAGTG TGTTTGAICT TGTGTGTACG CGCTTTTGGC AGTTGCTAGC
7741 TGCTTGTGCT ATTTGCGAAT ACCACCCCA GCATCCCCTT CCTCGTTTC ATATCGCTTG
7801 CATCCCAACC GCAACTTATC TACGCTGTCC TGCTATCCCT CAGCGCTGCT CCTGCTCCTG
7861 CTCACTGCCC CTCGCACAGC CTTGGTTTTG GCTCCGCTG TATTCTCCTG GTACTGCAAC
7921 CTGTAAACCA GCACTGCAAT GCTGATGCAC GGGAAAGTAG GGGATGGGAA CACAAATGGA
7981 AAGCTTAATT AAGAGCTCTT GTTTTCCAGA AGGAGTTGCT CCTTGAGCCT TTCATTCTCA
8041 GCCTCGATAA CCTCCAAAGC CGCTCTAATT GTGGAGGGGG TTCGAATTTA AAAGCTTGGA
8101 ATGTTGGTTC GTGCTCTGG AACAAGCCCA GACTTGTTC TCACTGGGAA AAGGACCATC
8161 AGCTCAAAA AACTTGCCGC TCAAACCGCG TACCTCTGCT TTCGCGCAAT CTGCCCTGTT
8221 GAAATCGCCA CCACATTAT ATTGTGACG TTGAGCAGTC TGTAATTGCC TCAGAATGTG
8281 GAATCATCTG CCCCCTGTGC GAGCCCATGC CAGGCATGTC GCGGGCGAGG ACACCCGCCA
8341 CTCGTACAGC AGACCATTAT GCTACCTCAC AATAGTTTCA AACAGTGACC ATATTTCTCG

```

8401 AAGCTCCCCA ACGAGCACCT CCATGCTCTG AGTGGCCACC CCCCAGCCCT GGTGCTTGCG
8461 GAGGGCAGGT CAACCGGCAT GGGGCIACCG AAATCCCCGA CCGGATCCCA CCACCCCGC
8521 GATGGGAAGA ATCTCTCCCC GGGATGTGGG CCCACCACCA GCACAACCTG CTGGCCAGG
8581 CGAGCGTCAA ACCATACCAC ACAAATATCC TTGGCATCGG CCCTGAATTC CTTCTGCCGC
8641 TCTGCTACCC GGTGCTTCTG TCCGAAGCAG GGGTTGCTAG GGATCGCTCC GAGTCCGCAA
8701 ACCCTTGTCG CGTGGCGGGG CTTGTTTCGAG CTTGAAGAGC CTCTAGAGTC GACCTGCAGG
8761 CATGCAAGCT TGGCGTAATC ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC
8821 ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA
8881 GTGAGCTAAC TCACATTAAT TCGTTGCGC TCACTGCCC GCTTCCAGTC GGGAAACCTG
8941 TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG
9001 CGTCTTCCG CTTCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCCGCT GCGGCGAGCG
9061 GTATCAGCTC ACTCAAAGGC GGTAAATACG TTATCCACAG AATCAGGGGA TAACGCAGGA
9121 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG
9181 GCGTTTTTCC ATAGGCTCCG CCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG
9241 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTG AACCTCCCTC
9301 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCTT TCTCCCTTCG
9361 GGAAGCGTGG CGCTTCTCA TAGCTACGCG TGTAGGTATC TCAGTTCGGT GTAGGTCGTT
9421 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCGTTCAGC CCGACCGCTG CGCCTTATCC
9481 GGTAACTATC GTCTTGAGTC CAACCGGTA AGACACGACT TATCG

```

SEQ ID NO: 227
pSZ1119

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTTTAATTG TTGCCGGGAA GCTAGATGTA GTAGTTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGTCTCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGA CTGGTGTA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAA TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACAGCG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCAGAGGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA
1741 ACTGTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCCG CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CTTGGCCTT TTCGCCGCGC TCGTGCCTCG CGCTGATGTC
1981 CATCAACAGG TCCATGAGGT CTGCCCTGCG CCGGCTGAGC CACTGCTTCG TCCGGCTGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCCTG ACGTGGTTCG GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAACGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA

```

2221 ACAACCACGA GCCTTGCTTA GGCAGAAATCC CTACCAGTCA TGGCTTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCCTCGCT GCCGCGCCTT CTCCCGCACG CTTCTTTTCCA
2341 GCACCGTGAT GGC GCGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTGCGCGCTT AGTGTTGCCA TCCTTTGCGAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACCTTT TTGGCGAAGG CATCGGCCTC GGCCCTGCAGA GAGGACAGCA GTGCCCAGCC
2581 GCTGGGGGTT GGCGGATGCA CGCTCAGGTA CCGTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GGCCTGTCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGTGTCATGG GCGCTCCGAT GCCGCTCCAG GCGGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATAITC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCCTCT TCGTTTCAGT CACAACCCGC AAACGGCGCG CCATGCTGCT
2941 GCAGGCCTTC CTGTTCTCTG TGGCCGGCTT CGCCGCCAAG ATCAGCGCCT CCATGACGAA
3001 CGAGAGCTTC GACCGCCCCC TGGTGCACTT CACCCCCAAC AAGGGCTGGA TGAACGACCC
3061 CAACGGGCTG TGGTACGAG AGAAGGACGC CAAGTGCGAC CTGTACTTCC TGATACAACC
3121 GAACGACACC GTCTGGGGGA CGCCCTTGTT CTGGGGCCAC GCCACGTCCG ACGACCTGAC
3181 CAACTGGGAG GACCAGCCCA TCGCCATCGC CCCGAAGCGC AACGACTCCG GCGCCTTCTC
3241 CGGCTCCATG GTGGTGGACT ACAACAACAC CTCGGGCTTC TTCAACGACA CCATCGACCC
3301 GCGCCAGCGC TCGTGCGCA TCTGGACCTA CAACACCCCG GAGTCCGAGG AGCAGTACAT
3361 CTCTACAGC CTGGACGGCG GCTACACCTT CACCGAGTAC CAGAAGAACC CCGTGCTGGC
3421 CGCCAACCTC ACCCAGTTCC GCGACCCGAA GGTCTTCTGG TACGAGCCCT CCCAGAAGTG
3481 GATCATGACC GCGGCCAAGT CCCAGGACTA CAAGATCGAG ATCTACTCCT CCGACGACCT
3541 GAAGTCTTGG AAGCTGGAGT CCGCGTTTCG CAACGAGGGC TTCCTCGGCT ACCAGTACGA
3601 GTGCCCCGGC CTGATCGAGG TCCCCACCGA GCAGGACCCC AGCAAGTCCT ACTGGGTGAT
3661 GTTCATCTCC ATCAACCCCG GCGCCCCGGC CGGCGGCTCC TTCAACCAGT ACTTCGTCGG
3721 CAGCTTCAAC GGCACCCACT TCGAGGCCTT CGACAACCAG TCCCGCGTGG TGGACTTCGG
3781 CAAGGACTAC TACGCCCTGC AGACCTTCTT CAACACCGAC CCGACCTACG GGAGCGCCCT
3841 GGGCATCGCG TGGGCCTCCA ACTGGGAGTA CTCCGCCTTC GTGCCACCA ACCCTTGGCG
3901 CTCTCCATG TCCCTCGTGC GCAAGTTCTC CCTCAACACC GAGTACCAGG CCAACCCGGA
3961 GACGGAGCTG ATCAACCTGA AGGCGGCGC GATCCTGAAC ATCAGCAACG CCGGCCCTTG
4021 GAGCCGGTTC GCCACCAACA CCACGTTGAC GAAGGCCAAC AGCTACAACG TCGACCTGTC
4081 CAACAGCACC GGCACCTGG AGTTGAGCT GGTGTACGCC GTCAACACCA CCCAGACGAT
4141 CTCCAAGTCC GTGTGCGCG ACCTCTCCCT CTGGTTCAAG GGCCTGGAGG ACCCCGAGGA
4201 GTACCTCCGC ATGGGCTTCG AGGTGTCCGC GTCCTCCTTC TTCCTGGACC GCGGGAACAG
4261 CAAGGTGAAG TTCGTGAAGG AGAACCCTTA CTTACCAAC CGCATGAGCG TGAACAACCA
4321 GCCCTTCAAG AGCGAGAACG ACCTGTCTTA CTACAAGGTG TACGGCTTGC TGGACAGAA
4381 CATCTGGAG CTGTACTTCA ACGACGGCGA CGTCTGTGCC ACCAACACCT ACTTCATGAC
4441 CACCGGGAAC GCCCTGGGCT CCGTGAACAT GACGACGGGG GTGGACAACC TGTTCTACAT
4501 CGACAAGTTC CAGGTGCGCG AGGTCAAGTG ACAATTGGCA GCAGCAGCTC GGATAGTATC
4561 GACACACTCT GGACGCTGGT CGTGTGATGG ACTGTTGCCG CCACACTTGC TGCCTTGACC
4621 TGTGAATATC CCTGCCGCTT TTATCAAAACA GCCTCAGTGT GTTTGATCTT GTGTGTACGC
4681 GCTTTTGCGA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCCCAG CATCCCCTTC
4741 CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTATCT ACGCTGTCTT GCTATCCCTC
4801 AGCGCTGCTC CTGCTCCTGC TCACGTGCCC TCGCACAGCC TTGGTTTGGG CTCCGCTTGT
4861 ATTCTCCTGG TACTGCAACC TGTAACCAG CACTGCAATG CTGATGCACG GGAAGTAGTG
4921 GGATGGGAAC ACAAATGGAG GATCCCGCGT CTCGAACAGA GCGCGCAGAG GAACGCTGAA
4981 GGTCTCGCCT CTGTGCGACC TCAGCGCGGC ATACACCACA ATAACCACCT GACGAATGCG
5041 CTTGGTTCTT CGTCCATTAG CGAAGCGTCC GGTTCACACA CGTGCCACGT TGGCGAGGTG
5101 GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG GGATATCGAA
5161 TTCGGCCGAC AGGACGCGCG TCAAAGGTGC TGGTCTGTGA TGCCCTGGCC GGCAGGTCTG
5221 TGCTGCTGCT GGTTAGTGAT TCCGCAACCC TGATTTTGGC GTCTTATTTT GCGGTGGCAA
5281 ACGCTGGCGC CCGCGAGCCG GGCCGGCGGC GATGCGGTGC CCCACGGCTG CCGGAATCCA
5341 AGGGAGGCAA GAGCGCCCGG GTCAGTTGAA GGGCTTTACG CGCAAGGTAC AGCCGCTCCT
5401 GCAAGGCTGC GTGGTGGAA TGGACGTGCA GGTCTCTGCT AAGTTCCTCC ACCGCTCAC
5461 CACGCGACAA AGCACCAGGT TATCAGGTCC GTGTCATCCA CTCTAAAGAG CTCGACTACG
5521 ACCTACTGAT GGCCCTAGAT TCTTCATCAA AAACGCCTGA GACACTTGCC CAGGATTGAA
5581 ACTCCCTGAA GGGACCACCA GGGGCCCTGA GTTGTTCCTT CCCCCGTGG CGAGGTGCCA
5641 GCCAGGCTGT ACCTGTGATC GAGGCTGGCG GGAATAATAG CTTCTGTGTC TCAGGTTCATG
5701 GGAGGTGCAG GACAGCTCAT GAAACGCCAA CAATCGCACA ATTCATGTCA AGCTAATCAG
5761 CTATTTCTCT TTCACGAGCT GTAATTGTCC CAAAATTCTG GTCTACCGGG GGTGATCCTT
5821 CGTGACGGG CCCTTCCCTC AACCTAGGT ATGCGCGCAT GCGGTGCGCG CGCAACTCGC

5881 GCGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGCACCCG GATGCGTGGC
5941 ACCTTTTTTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT CTGTGCCCAA
6001 CCCTAGGATC AGCGGCGTAG GATTTCGTAA TCATTCGTCC TGATGGGGAG CTACCGACTA
6061 CCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC ATTCCATTCT
6121 TGCCCAAGAC ATTTCAATTGT GGTGCGAAGC GTCCCCAGTT ACGCTACCTT GTTTCCCGAC
6181 CTCCTTACTG TTCTGTCGAC AGAGCGGGCC CACAGGCCGG TCGCAGCCAC TAGTATGGCC
6241 ACCGCATCCA CTTTCTCGGC GTTCAATGCC CGCTGCGGCG ACCTGCGTCG CTCGGCGGGC
6301 TCCGGGCCCC GCGCGCCAGC GAGGCCCTC CCCGTGCGCG GCGCGGCCA GCTGCCCGAC
6361 TGGAGCATGC TGCTGGCCGC GATCACCACC CTGTTCCCTGG CGGCCGAGAA GCAGTGGATG
6421 ATGCTGGACT GGAAGCCCAA GCGCCCCGAC ATGCTGGTGG ACCCCTTCGG CCTGGGCCGC
6481 TTCGTGCAGG ACGGCCTGGT GTTCCGCAAC AACTTCAGCA TCCGCAGCTA CGAGATCGGC
6541 GCGGACCGCA CCGCCAGCAT CGAGACCCTG ATGAACCACC TGCAGGAGAC CGCCCTGAAC
6601 CACGTGAAGA GCGTGGGCTT GCTGGAGGAC GGCTTGGGCA GCACCCGCGA GATGAGCCTG
6661 CGCAACCTGA TCTGGGTGGT GACCAAGATG CAGGTGGCGG TGGACCGCTA CCCCACCTGG
6721 GCGGAGCGAG TGCAGGTGAG CAGCTGGGCG ACCGCCATCG ACCAGAACGG GCTGCCCGAC
6781 GAGTGGATCG TGACCGACTT CCGCACCAGC GAGACCCTGC TCGCGCGCAC CAGCGTGTGG
6841 GTGATGATGA ACAAGCTGAC CCGCCGCATC AGCAAGATCC CCGAGGAGGT GTGGCACGAG
6901 ATCGGCCCCA GCTTCATCGA CGCGCCCCC CTGCCACCG TGGAGGACGA CGGCCGCAAG
6961 CTGACCCGCT TCGACGAGAG CAGCGCCGAC TTCATCCGCA AGGGCCTGAC CCCCCGCTGG
7021 AGCGACCTGG ACATCAACCA GCACGTGAAC AACGTGAAGT ACATCGGCTG GCTGCTGGAG
7081 AGCGCGCCCC CCGAGATCCA CGAGAGCCAC GAGATCGCCA GCCTGACCTT GGAGTACCGC
7141 GCGGAGTGGC GCGCGACAG GTGTCTGAAC AGCGCCACCA AGGTGAGCGA CAGCAGCCAG
7201 CTGGGCAAGA GCGCCGTGGA GTGCAACCAC CTGGTGCGCC TGCAGAACGG CGCGGAGTATC
7261 GTGAAGGGCC GCACCGTGTG GCGCCCCAAG CGCCCCCTGT ACAACGACGG CGCCGTGGTG
7321 GACGTGCCCC CCAAGACCAG CGATGACGAT GACAAGCTGG GATGACTCGA GGCAGCAGCA
7381 GCTCGGATAG TATCGACACA CTCTGGACGC TGGTCTGTGT ATGGACTGTT GCCGCCACAC
7441 TTGCTGCCTT GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA
7501 TCTTGTGTGT ACGCGCTTTT GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC
7561 CCAGCATCCC CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG
7621 TCCTGCTATC CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT
7681 TGGGCTCCGC CTGTATTCTC CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGATG
7741 CACGGGAAGT AGTGGGATGG GAACACAAAT GGAAAGCTTG AGCTCTTGTG TTCCAGAAGG
7801 AGTTGCTCCT TGAGCCTTTC ATTCTCAGCC TCGATAACCT CCAAAGCCGC TCTAATTGTG
7861 GAGGGGGTTC GAATTTAAAA GCTTGAATG TTGGTTCTGT CGTCTGGAAC AAGCCCAGAC
7921 TTGTTGCTCA CTGGGAAAAG GACCATCAGC TCCAAAAAAC TTGCCGCTCA AACCCGCTAC
7981 CTCTGCTTTT GCGCAATCTG CCCGTGTGAA ATCGCCACCA CATTCAATT GTGACGCTTG
8041 AGCAGTCTGT AATTGCTCTA GAATGTGGAA TCATCTGCCC CCTGTGCGAC CCGCTGCGAG
8101 GCATGTCGCG GCGGAGGACA CCCGCCACTC GTACAGCAGA CCATTATGCT ACCTCACAAT
8161 AGTTCATAAC AGTGACCATA TTTCTGAAG CTCCCCAACG AGCACCTCCA TGCTCTGAGT
8221 GGCCACCCCC CGGCCCTGGT GCTTGCAGAG GGCAGGTCAA CCGGCATGGG GCTACCGAAA
8281 TCCCCGACCG GATCCCACCA CCCCCGCGAT GGGAAGAATC TCTCCCCGGG ATGTGGGCCC
8341 ACCACCAGCA CAACCTGCTG GCCCAGGCGA GCGTCAAACC ATACCACACA AATATCCTTG
8401 GCATCGGCCC TGAATTCCTT CTGCCGCTCT GCTACCCGGT GCTTCTGTCC GAAGCAGGGG
8461 TTGCTAGGGA TCGCTCCGAG TCCGCAAAAC CTGTGTCGCT GCGGGGGCTT GTTCGAGCTT
8521 GAAGAGCCTC TAGAGTCGAC CTGCAGGCAT GCAAGCTTGG CGTAATCATG GTCATAGCTG
8581 TTTCTGTGTG GAAATTGTTA TCCGCTCACA ATTCCACACA ACATACGAGC CGGAAGCATA
8641 AAGTGTAAGC CCTGGGGTGC CTAATGAGTG AGCTAACTCA CATTAAATTGC GTTGCCTTCA
8701 CTGCCCCTTT TCCAGTCGGG AAACCTGTCT TGCCAGCTGC ATTAATGAAT CGGCCAACGC
8761 GCGGGGAGAG GCGGTTTGGC TATTGGGCGC TCTTCCGCTT CCTCGCTCAC TGAATCGCTG
8821 CGCTCGGTCT TTCGGCTGCG GCGAGCGGTA TCAGCTCACT CAAAGGCGGT AATACGGTTA
8881 TCCACAGAAT CAGGGGATAA CGCAGGAAAG AACATGTGAG CAAAAGGCCA GCAAAAGGCC
8941 AGGAACCGTA AAAAGGCCGC GTTGTGGCG TTTTTCATA GGCTCCGCCC CCTGACGAG
9001 CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC
9061 CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT GCCGCTTACC
9121 GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC TTTCTCATAG CTCACGCTGT
9181 AGGTATCTCA GTTCCGTGTA GGTCTTTCGC TCCAAGCTGG GCTGTGTGCA CGAACCCCCC
9241 GTTCAGCCCC ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGTAAGA
9301 CACGACTTAT CGC

SEQ ID NO: 228

```

1  CTAAATTGTA AGCGTTAATA TTTTGTAAAA ATTTCGCGTTA AATTTTTGTT AAATCAGCTC
61  ATTTTTTAAC CAATAGGCCG AAATCGGCAA AATCCCTTAT AAATCAAAAG AATAGACCGA
121  GATAGGGTTG AGTGTGTGTC CAGTTTGGAA CAAGAGTCCA CTATTAAAGA ACGTGGACTC
181  CAACGTCAAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC
241  CTAATCAAGT TTTTGGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG
301  CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GCGAACCGTG GCGAGAAAGG AAGGGAAGAA
361  AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC
421  CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCC CATTGCCCAT TCAGGCTGCG
481  CAAGTGTGGG GAAGGGCGAT CGGTGCGGGC CTCTTCGCTA TTACGCCAGC TGGCGAAAGG
541  GGGATGTGCT GCAAGGCGAT TAAGTTGGGT AACGCCAGGG TTTTCCAGT CACGACGTTG
601  TAAACGACG GCCAGTGAGC GCGCGTAATA CGACTCACTA TAGGGCGAAT TGCAGCTTTT
661  GTTCCCTTTA GTAGGGTTA ATTGCGCGCT TGGCGTAATC ATGGTCATAG CTGTTTCTCTG
721  TGTGAAATTG TTATCCGCTC ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA
781  AAGCCTGGGG TGCCTAATGA GTGAGCTAAC TCACATTAAAT TGCCTTGCGC TCACTGCCCCG
841  CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAAAT AATCGGCCAA CGCGCGGGGA
901  GAGGCGGTGC GTATTGGGCG CTCTTCGCTC CCCGTGATCA CACAGGTGCC CTGCGAGCGT
961  GATCACACTA TTTTGGGGGT CCTACAGTAC TGAAATGGTG AGAAGTCGTA CTGAAATCAA
1021  GGATGAACAA TGAATATGGT GCTGTGGTGG CTTCTCAAAG GTCAAGAATC AGTCGCTCGC
1081  GTCAGGAAAT CGCGGCGTCA ACCAGCGTGG GCGCGGTCAG TGGCCCCGCA CTGGTCACCA
1141  TAGCCTCTCC TGCCACAGTA GCGATCCCTT GGGCGTTCAC TCTCAGCAGC GGCTGTACTG
1201  CCTCCAGAT TTTCTTCTTC TGGACCTGCG GCGGTGAGAG GATGAGCAGG GTGGGGCCAA
1261  GGGCTCAATC CTGAACGGCC CTCATTGCGT TTCCAATCCC ACAACACATA CCCACAGCAG
1321  GTCAGACCAC GCATTGCGAC CATGCGCACC AATAACGTGT CTTTACCTGA TTGGGTGTGG
1381  CAGGCTCCGT GGACAGGAGT GCCTCGTCCC CCGCCCAGAC CCGCTCCCCC GTACCGCGCG
1441  CGTCCGGGAC CCGCAGCGGC TCCACCGCGG TGTGATCCGC GTTGGCGGCG CAGAGCAGCA
1501  TCCAGCCGA TTTGACCCCG CGCATGCTCC GAGGCTTGAG GTTGGCCAGC ACCACCACCC
1561  GCGGCGCGAC AAGGTCTCTC AGGGTCACGT GCCGGACCAG GCCACTCACG ATGGTGCGAG
1621  GGGCCCCCTC CTCGCCGAGG TCGATCTGCT CGACGTACAG ACTGCGACAT GCGTGGCGAG
1681  TGGTCATCAG AAGGAAGCAG GTGTGCAGAA GGGGCACGTG GTTGGTATTG AGAGTAGCCA
1741  AAGCTTTGTG CCAATCAGAA AGTCAACGCA GCTGCCTGCC TGGCTCGCTG ACGGTACCTT
1801  TTCTTGCGCT ATGACACTTC CAGCAAAAGG TAGGGCGGGC TGCGAGACGG CTTCGCCGCG
1861  CTGCATGCAA CACCGATGAT GCTTCGACCC CCCGAAGCTC CTTGCGGGCT GCATGGGCGC
1921  TCCGATGCCG CTCCAGGGCG AGCGCTGTTT AAATAGCCAG GCCCCGATT GCAAAGACAT
1981  TATAGCGAGC TACCAAAGCC ATATTCAAAC ACCTAGATCA CTACCATTTC TACACAGGCC
2041  ACTCGAGCTT GTGATCGCAC TCCGCTAAGG GGGCGCCTCT TCCTCTTCGT TTCAGTCACA
2101  ACCCGCAAAC TCTAGAATAT CAATGCTGCT GCAGGCCTTC CTGTTCTCTG TGGCCGGCTT
2161  CGCCGCCAAG ATCAGCGCCT CCATGACGAA CGAGACGTCC GACCGCCCCC TGGTGCATT
2221  CACCCCAAC AAGGGCTGGA TGAACGACCC CAACGGCCTG TGGTACGAGC AGAAGGACGC
2281  CAAGTGGCAC CTGTACTTCC AGTACAACCC GAACGACACC GTCTGGGGGA CGCCCTTGTT
2341  CTGGGGCCAC GCCACGTCCG ACGACCTGAC CAACTGGGAG GACCAGCCCA TCGCCATCGC
2401  CCCGAAGCGC AACGACTCCG GCGCCTTCTC CGGCTCCATG GTGGTGGACT ACAACAACAC
2461  CTCCGGCTTC TTCAACGACA CCATCGACCC GCGCCAGCGC TGCGTGGCCA TCTGGACCTA
2521  CAACACCCCG GAGTCCGAGG AGCAGTACAT CTCCTACAGC CTGGACGGCG GCTACACCTT
2581  CACCGAGTAC CAGAAGAACC CCGTGTGTCG CGCCAATCC ACCCAGTTCC GCGACCCGAA
2641  GGTCTTCTGG TACGAGCCCT CCCAGAAGTG GATCATGACC GCGGCCAAGT CCCAGGACTA
2701  CAAGATCGAG ATCTACTCCT CCGACGACCT GAAGTCCTGG AAGCTGGAGT CCGCGTTTCG
2761  CAACGAGGGC TTCTTCGGCT ACCAGTACGA GTGCCCCGGC CTGATCGAGG TCCCCACCGA
2821  GCAGGACCCC AGCAAGTCCT ACTGGGTGAT GTTCATCTCC ATCAACCCCG GCGCCCCGGC
2881  CGCGGGCTCC TTCAACCAGT ACTTCGTTCG CAGCTTCAAC GGCACCCACT TCGAGGCCTT
2941  CGACAACCAG TCCCGCGTGG TGGACTTCGG CAAGGACTAC TACGCCCTGC AGACCTTCTT
3001  CAACACCGAC CCGACCTACG GGAGCGCCTT GGGCATCGCG TGGGCCTCCA ACTGGGAGTA
3061  CTCCGCTTC GTGCCACCA ACCCTGCGG CTCTCCATG TCCCTCGTGC GCAAGTTCTC
3121  CCTCAACACC GAGTACCAGG CCAACCCGGA GACGGAGCTG ATCAACCTGA AGGCTGAGCC
3181  GATCCTGAAC ATCAGCAACG CCGGCCCTTG GAGCCGGTTC GCCACCAACA CCACGTTGAC
3241  GAAGGCCAAC AGCTACAACG TCGACCTGTC CAACAGCACC GGCACCCTGG AGTTCGAGCT
3301  GGTGTACGCC GTCAACACCA CCCAGACGAT CTCCAAGTCC GTGTTTCGCG ACCTCTCCCT
3361  CTGGTTCAAG GGCCTGGAGG ACCCCGAGGA GTACCTCCGC ATGGGCTTCG AGGTGTCCGC
3421  GTCTCTCTTC TTCTTGACC GCGGGAACAG CAAGGTGAAG TTCGTGAAGG AGAACCCTTA
3481  CTTACCAAC CGCATGAGCG TGAACAACCA GCCCTTCAAG AGCGAGAACG ACCTGTCTTA

```

3541 CTACAAGGTG TACGGCTTGC TGGACCAGAA CATCCTGGAG CTGTACTTCA ACCACGGCGA
3601 CGTCGTGTCC ACCAACACCT ACTTCAATGAC CACCGGGAAC GCCCTGGGCT CCGTGAACAT
3661 GACGACGGGG GTGGACAACC TGTCTACAT CGACAAGTTC CAGGTGCGCG AGGTCAAGTG
3721 ACAATTGGCA GCAGCAGCTC GGATAGTATC GACACACTCT GGACGCTGGT CGTGTGATGG
3781 ACTGTTGCCG CCACACTTGC TGCCTTGACC TGTGAATATC CCTGCCGCTT TTATCAAACA
3841 GCCTCAGTGT GTTTGATCTT GTGTGTACGC GCTTTTGC GA GTTGCTAGCT GCTTGTGCTA
3901 TTTGCGAATA CCACCCCCAG CATCCCCCTC CCTCGTTTCA TATCGCTTGC ATCCCAACCG
3961 CAACTTATCT ACGCTGTCTT GCTATCCCTC AGCGCTGCTC CTGCTCCTGC TCACTGCCCC
4021 TCGCACAGCC TTGGTTTGGG CTCGCTGTGT ATTCTCCTGG TACTGCAACC TGTAACCAG
4081 CACTGCAATG CTGATGCACG GGAAGTAGTG GGATGGGAAC ACAAAATGGAG GATCCCGCGT
4141 CTCGAACAGA GCGCGCAGAG GAACGCTGAA GGTCTCGCCT CTGTGCGACC TCAGCGCGGC
4201 ATACACCACA ATAACCACCT GACGAATGCG CTTGGTTCTT CGTCCATTAG CGAAGCGTCC
4261 GGTTCACACA CGTGCCACGT TGGCGAGGTG GCAGGTGACA ATGATCGGTG GAGCTGATGG
4321 TCGAAACGTT CACAGCCTAG GGATATCGAA TTCCTTTCTT GCGCTATGAC ACTTCCAGCA
4381 AAAGGTAGGG CGGGCTGCGA GACGGCTTCC CGGCGCTGCA TGCAACACCG ATGATGCTTC
4441 GACCCCCCGA AGCTCCTTCG GCGCTGCATG GCGCTCCGA TGCCGCTCCA GGGCGGCGC
4501 TGTTTAAATA GCCAGGCCCC CGATTGCAAA GACATTATAG CGAGCTACCA AAGCCATATT
4561 CAAACACCTA GATCACTACC ACTTCTACAC AGGCCACTCG AGCTTGTGAT CGCACTCCGC
4621 TAAGGGGGCG CTTCTTCCTC TTCGTTTCAG TCACAACCCG CAAACACTAG TATGGCCACC
4681 GCATCCACTT TCTCGGCGTT CAATGCCGCG TGCGGCGACC TGCGTCTGCTC GCGGGGCTCC
4741 GGGCCCCGGC GCCACGCGAG GCCCTCCCC GTGCGCGGGC GCGCCCCCGA CTGGTCCATG
4801 CTGTTGCGCG TGATCACCAC CATCTTCTCC GCGCGCGAGA AGCAGTGGAC CAACCTGGAG
4861 TGGAAGCCCA AGCCCAACCC CCCCCAGCTG CTGGACGACC ACTTCGGGCC CCACGGCCTG
4921 GTGTTCCGCC GCACCTTCGC CATCCGACGC TACGAGGTGG GCCCCGACCG CTCACACAGC
4981 ATCGTGGCCG TGATGAACCA CCTGCAGGAG GCGGCCCTGA ACCACGCCAA GTCCGTGGGC
5041 ATCCTGGGCG ACGGCTTCGG CACCACCCTG GAGATGTCCA AGCGCGACCT GATCTGGGTG
5101 GTGAAGCGCA CCCACGTGGC CGTGGAGCGC TACCCCGCCT GGGGCGACAC CGTGGAGGTG
5161 GAGTGCTGGG TGGGCGCCTC CGGCAACAAC GGCCGCGGCC ACGACTTCCT GGTGCGCGAC
5221 TGCAAGACCG GCGAGATCCT GACCCGCTGC ACCTCCCTGA GCGTGATGAT GAACACCCGC
5281 ACCCGCCGCC TGAGCAAGAT CCCCAGGAG GTGCGCGGGC AGATCGGCC CGCCTTCATC
5341 GACAACGTGG CCGTGAAGGA CGAGGAGATC AAGAAGCCCC AGAAGCTGAA AGACTCCACC
5401 GCCGACTACA TCCAGGGCGG CCTGACCCCC CGCTGGAACG ACCTGGACAT CAACCAGCAC
5461 GTGAACAACA TCAAGTACGT GGAATGATC CTGGAGACCG TGCCCGACAG CATCTTCGAG
5521 AGCCACCACA TCTCTCCTT CACCATCGAG TACCGCCGCG AGTGCACCAT GGACAGCGTG
5581 CTGCACTCCC TGACCACCGT GAGCGGCGGC TCCTCCGAGG CCGGCCTGGT GTGCGAGCAC
5641 CTGCTGCAGC TGGAGGGCGG CAGCGAGGTG CTGCGCGCCA AGACCGAGTG GCGCCCCAAG
5701 CTGACCGACT CTTCCGCGG CATCAGCGTG ATCCCCGCCG AGTCCAGCGT GATGGACTAC
5761 AAGCAACGAC ACGGCGACTA CAGGACACG GACATCGACT ACAAGACGA CGACGACAA
5821 TGAATCGAGG CAGCAGCAGC TCGGATAGTA TCGACACACT CTGGACGCTG GTCGTGTGAT
5881 GGAATGTTGC CGCCACACTT GCTGCCCTGA CCTGTGAATA TCCCTGCCGC TTTTATCAAA
5941 CAGCCTCAGT GTGTTTGATC TTGTGTGTAC GCGCTTTTGC GAGTTGCTAG CTGCTTGTGC
6001 TATTTGCGAA TACCACCCCC AGCATCCCTT TCCCTCGTTT CATATCGCTT GCATCCCAAC
6061 CGCAACTTAT CTACGCTGTC CTGCTATCCC TCAGCGCTGC TCCTGCTCCT GCTCACTGCC
6121 CCTCGCACAG CTTGGTTTG GGCTCCGCTT GTATTCTCCT GGTACTGCAA CCTGTAAACC
6181 AGCACTGCAA TGCTGATGCA CGGGAAGTAG TGGATGGGA ACACAAATGG AAAGCTTGAG
6241 CTCGTACCCA TCAGCATCCG GGTGAATCTT GGCTTCCAAG ATATGGCCAA TCCTCACATC
6301 CAGCTTGGCA AAATCGACTA GACTGTCTGC AAGTGGGAAT GTGGAGCACA AGGTTGCTTG
6361 TAGCGATCGA CAGACTGGTG GGGTACATTG ACAGGTGGGC AGCGCCGCAT CCATCGTGCC
6421 TGACGCGAGC GCGCGCGGTT GCTCGCCCGT GCCTGCGCTC AAAGAGCGGC AGAGAAATCG
6481 GGAACCGAAA ACGTCACATT GCCTGATGTT GTTACATGCT GGACTAGACT TTCTTGGCGT
6541 GGGTCTGCTC CTCGCCAGGT GCGCGACGCC TCGGGGCTGG GTGCGAGGGA GCCGTGCGGC
6601 CACGCATTTG ACAAGACCCA AAGCTCGCAT CTCAGACGGT CAACCGTTCC TATTATACAT
6661 TCAACATATG GTACATACGC AAAAAGCATG CCAACGATGA CAGCTCTTCG CCCACATGTG
6721 AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA
6781 TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
6841 CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG TCGCTCTCC
6901 TGTTCCGACC CTGCGCTTA CCGGATACCT GTCCGCTTT CTCCCTTCGG GAAGCGTGGC
6961 GCTTTCTCAT AGCTCACGCT GTAGGTATCT CAGTTCGGTG TAGGTGCTTC GCTCCAAGCT
7021 GGGCTGTGTG CACGAACCCC CCGTTACGCC CGACCGCTGC GCCTTATCCG GTAACATCG
7081 TCTTGAGTCC AACCCGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG
7141 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT GGCCTAACTA

```

7201 CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG
7261 AAAAAGAGTT GGTAGCTCTT GATCCGGCAA ACAAACCACC GCTGGTAGCG GTGGTTTTTT
7321 TGTTTGCAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT
7381 TTCTACGGGG TCTGACGCTC AGTGAACGA AAACTCACGT TAAGGGATTT TGGTCATGAG
7441 ATTATCAAAA AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAATCAAT
7501 CTAAAGTATA TATGAGTAAA CTGGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC
7561 TATCTCAGCG ATCTGTCTAT TTCGTTCAIC CATAGTTGCC TGACTIONCCG TCGTGTAGAT
7621 AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC
7681 ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG
7741 AAGTGGTCCCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGAAGCTAG
7801 AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT
7861 GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCTAAC GATCAAGGCG
7921 AGTTACATGA TCCCCATGT TGTGCAAAA AGCGGTTAGC TCCTTCGGTC CTCGATCGT
7981 TATCAGTAGT AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCTG ATGGCAATTC
8041 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
8101 ATTCTGAGAA TAGTGTATGC GCGCACCAGG TTGCTCTTGC CCGGCGTCAA TACGGGATAA
8161 TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTCGGGGCG
8221 AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC
8281 CAACTGATCT TCAGCATCTT TACTTTTAC CAGCGTTTCT GGGTGAGCAA AAACAGGAAG
8341 GCAAAATGCC GCAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT
8401 CCTTTTTCAA TATTATTGAA GCATTATCA GGGTTATTGT CTCATGAGCG GATACATATT
8461 TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCGCGC ACATTTCCCC GAAAAGTGCC
8521 AC

```

SEQ ID NO: 229
pSZ1283

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAATAATG AAGTTTAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAGCGG
841 TTAGCTCCTT CGGTCCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCGATG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTTCTTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTIONGTA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTGCCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAA TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCAACT GATCTTCAGC ATCTTTTACT TTCACGACG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAA AAAGGGAATA AGGCGGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCAGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CCGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGC TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCCG CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGCGCTT TTCGCCGCGC TCGTGCAGCT CGCTGATGTC

```

1981 CATCACCAGG TCCATGAGGT CTGCCTTGCG CCGGCTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCCTG ACGTGGTCGC GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAACCTGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCC TGGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGAAATCC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCCTCGCT GCCGCCGCTT CTCCCGCACG CTCTTTTCCA
2341 GCACCGTGAT GGC GCGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTCGCGCTT AGTGTTGCCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACCTTT TTGGCGAAGG CATCGGCCTC GGCCTGCAGA GAGGACAGCA GTGCCAGGCC
2581 GCTGGGGGTT GGC GATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGTCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACTAG ATCACTACCA CTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCTCT TCGTTTCAGT CACAACCCGC AAACGGCGCG CCATGCTGCT
2941 GCAGGCCCTT CTGTTCTGCT TGGCCGCTT CGCCGCCAAG ATCAGCGCCT CCATGACGAA
3001 CGAGACGTCC GACCGCCCCC TGGTGCATTT CACCCCAAC AAGGGCTGGA TGAACGACCC
3061 CAACGGCCTG TGGTACGACG AGAAGGACGC CAAGTGGCAC CTGTACTTCC AGTACAACCC
3121 GAACGACACC GTCTGGGGGA CGCCCTTGTT CTGGGGCCAC GCCACGTCCG ACGACCTGAC
3181 CAACTGGGAG GACCAGCCCA TCGCCATCGC CCCGAAGCGC AACGACTCCG GCGCCTTCTC
3241 CGGCTCCATG GTGGTGGACT ACAACAACAC CTCCGGCTTC TTCAACGACA CCATCGACCC
3301 GCGCCAGCGC TGGCTGGCCA TCTGGACCTA CAACACCCCG GAGTCCGAGG AGCATGACAT
3361 CTCTACAGC CTGGACGGCG GCTACACCTT CACCGAGTAC CAGAAGAACC CCGTGTGGC
3421 CGCCAACTCC ACCCAGTTCC GCGACCCGAA GGTCTTCTGG TACGAGCCCT CCCAGAAGTG
3481 GATCATGACC GCGGCCAAGT CCCAGGACTA CAAGATCGAG ATCTACTCCT CCGACGACCT
3541 GAAGTCTTGG AAGCTGGAGT CCGCGTTTCG CAACGAGGGC TTCCTCGGCT ACCAGTACGA
3601 GTGCCCCGGC CTGATCGAGG TCCCCACCGA GCAGGACCCC AGCAAGTCCT ACTGGGTGAT
3661 GTTCATCTCC ATCAACCCCG GCGCCCGCGC CGGCGGCTCC TTCAACCGT ACTTCGTCGG
3721 CAGCTTCAAC GGCACCCACT TCGAGGCTTT CGACAACCA TCCCGCTGG TGGAGTCTCG
3781 CAAGGACTAC TACGCCCTGC AGACCTTCTT CAACACCGAC CCGACCTACG GAGCGCCCT
3841 GGGCATCGCG TGGGCTCCA ACTGGGAGTA CTCCGCCTTC GTGCCACCA ACCCCTGGCG
3901 CTCTCCATG TCCCTCGTGC GCAAGTTCTC CCTCAACACC GAGTACCAGG CCAACCCGGA
3961 GACGGAGCTG ATCAACCTGA AGGCCGAGCC GATCCTGAAC ATCAGCAACG CCGGCCCTG
4021 GAGCCGGTTC GCCACCAACA CCACGTTGAC GAAGGCCAAC AGCTACAACG TCGACCTGTC
4081 CAACAGCACC GGCACCCTGG AGTTCGAGCT GGTGTACGCC GTCAACACCA CCCAGACGAT
4141 CTCCAGTCC GTGTTGCGG ACCTCTCCCT CTGGTTCAAG GGCCTGGAGG ACCCCGAGGA
4201 GTACCTCCGC ATGGGCTTCG AGGTGTCCGC GTCCTCCTTC TTCCTGGACC GCGGGAACAG
4261 CAAGGTGAAG TTCGTGAAGG AGAACCCCTA CTTACCAAC CGCATGAGCG TGAACAACCA
4321 GCCCTCAAG AGCGAGAACG ACCTGTCTTA CTACAAGGTG TACGGCTTGC TGGACCAGAA
4381 CATCTGGAG CTGTACTTCA ACGACGGCGA CGTCGTGTCC ACCAACACCT ACTTCATGAC
4441 CACCGGGAAC GCCCTGGGCT CCGTGAACAT GACGACGGGG GTGGACAACC TGTCTACAT
4501 CGACAAGTTC CAGGTGCGCG AGGTCAAGTG ACAATTGGCA GCAGCAGCTC GGATAGTATC
4561 GACACACTCT GGCAGCTGGT CGTGTGATGG ACTGTTGCCG CCACACTTGC TGCCCTGACC
4621 TGTGAATATC CCTGCCGCTT TTATCAAACA GCCTCAGTGT GTTTGATCTT GTGTGTACGC
4681 GCTTTTGC GA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCCCAG CATCCCCTTC
4741 CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTAICT ACGCTGTCTT GCTATCCCTC
4801 AGCGCTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG CTCCGCCTGT
4861 ATTCTCTGG TACTGCAACC TGTAACCAG CACTGCAATG CTGATGCACG GGAAGTAGTG
4921 GGATGGGAAC ACAAATGGAG GATCCCGCGT CTCGAACAGA GCGCGCAGAG GAACGCTGAA
4981 GGTCTCGCCT CTGTCGCACC TCAGCGCGCG ATACACCACA ATAACCACCT GACGAATGCG
5041 CTTGGTTCTT CGTCCATTAG CGAAGCGTCC GGTTCACACA CGTGCCACGT TGGCGAGGTG
5101 GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG GGATATCGAA
5161 TTCGGCCGAC AGGACGCGCG TCAAAGGTGC TGGTCGTGTA TGCCCTGGCC GGCAGGTCTG
5221 TGCTGCTGCT GGTAGTGAT TCCGCAACCC TGATTTTGGC GTCTTATTTT GCGGTGGCAA
5281 ACGCTGGCGC CGCGGAGCCG GCGCGGCGCG GATGCGGTGC CCCACGGCTG CCGGAATCCA
5341 AGGGAGGCAA GAGCGCCCGG GTCAGTTGAA GGGCTTTACG CGCAAGGTAC AGCCGCTCCT
5401 GCAAGGCTGC GTGGTGAAT TGGACGTGCA GGTCTGCTG AAGTTCTTCC ACCGCTCAC
5461 CAGCGGACAA AGCACCGGTG TATCAGGTCC GTGTATCCA CTCTAAAGCG CTCGACTACG
5521 ACCTACTGAT GGCCCTAGAT TCTTCATCAA AAACGCCTGA GACACTTGCC CAGGATTGAA
5581 ACTCCCTGAA GGGACCACCA GGGGCCCTGA GTTGTTCCTT CCCCCGTTG CGAGCTGCCA

5641 GCCAGGCTGT ACCTGTGATC GAGGCTGGCG GGAAAATAGG CTTTCGTGTGC TCAGGTCATG
5701 GGAGGTGCAG GACAGCTCAT GAAACGCCAA CAATCGCACA ATTCAATGTC AGCTAATCAG
5761 CTATTTCTCT TACAGGAGCT GTAATTGTCC CAAAATTCTG GTCTACCGGG GGTGATCCTT
5821 CGTGTACGGG CCCTTCCCTC AACCTTAGGT ATGCGCGCAT GCGGTCGCCG CGCAACTCGC
5881 GCGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGACCCCG GATGCGTGGC
5941 ACCTTTTTTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT CTGTCGCCAA
6001 CCCTAGGATC AGCGGCGTAG GATTTCGTAA TCATTCTGCC TGATGGGGAG CTACCGACTA
6061 CCCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC ATTCCATTCT
6121 TGCCCAAGAC ATTTCAATTG GGTGCGAAGC GTCCCCAGTT ACGCTCACCT GTTTCCCGAC
6181 CTCTTACTGT TTCTGTGAC AGAGCGGGCC CACAGGCCGG TCGCAGCCAC TAGTATGGTG
6241 GTGGCCGCCG CCGCCAGCAG CGCCTTCTTC CCCGTGCCCC CCCCCGCCG CACCCCCAAG
6301 CCCGGCAAGT TCGGCAACTG GCCCAGCAGC CTGAGCCAGC CTTCAAGCC CAAGAGCAAC
6361 CCCAACGGCC GCTTCCAGGT GAAGGCCAAC GTGAGCCCCC ACGGGCGCGC CCCCAGGCC
6421 AACGGCAGCG CCGTAGCCTT GAAGTCGGC AGCCTGAACA CCCTGGAGGA CCCCCCAGC
6481 AGCCCCCCCC CCCGCACCTT CCTGAACCAG CTGCCCGACT GGAGCCGCCT GCGCACCGCC
6541 ATCACCACCG TGTTCTGTGG CGCCGAGAAG CAGTTCACCC GCCTGGACCG CAAGAGCAAG
6601 CGCCCCGACA TGCTGGTGGA CTGGTTCCGG AGCGAGACCA TCGTGCAGGA CGGCCTGGTG
6661 TTCCGCGAGC GCTTCAGCAT CCGCAGCTAC GAGATCGGCG CCGACCGCAC CGCCAGCATC
6721 GAGACCCTGA TGAACCACCT GCAGGACACC AGCCTGAACC ACTGCAAGAG CGTGGGCGCT
6781 CTGAACGACG GCTTCGGCCG CACCCCGAG ATGTGCACCC GCGACCTGAT CTGGGTCTGT
6841 ACCAAGATGC AGATCGTGGT GAACCCGTAC CCCACCTGGG GCGACACCGT GGAGATCAAC
6901 AGCTGGTTCA GCCAGAGCGG CAAGATCGGC ATGGGCCGCG AGTGGCTGAT CAGCGACTGC
6961 AACACCGGCG AGATCCTGGT GCGCGCCACC AGCGCCTGGG CCATGATGAA CCAGAAGACC
7021 CGCCGCTTCA GCAAGCTGCC CTGCGAGGTG CGCCAGGAGA TCGCCCCCA CTTCGTGGAC
7081 GCCCCCCCCG TGATCGAGGA CAACGACCGC AAGCTGCACA AGTTCGACGT GAAGACCGGC
7141 GACAGCATCT GCAAGGGCCT GACCCCCGCG TGAACGACT TCGACGTGAA CCAGCACGTG
7201 AGCAACGTGA AGTACATCGG CTGGATTCTG GAGAGCATGC CCACCGAGGT GCTGGAGACC
7261 CAGGAGCTGT GCAGCCTGAC CTTGGAGTAC CGCCGCGAGT GCGGCCGCGA GAGCGTGGTG
7321 GAGAGCGTGA CCAGCATGAA CCCCAGCAAG GTGGGCGACC GCAGCCAGTA CCAGCACCTG
7381 CTGCGCCTGG AGGACGGCGC CGACATCATG AAGGGCCGCA CCGAGTGGCG CCCCAGAAG
7441 GCCGGCACCA ACCGCGCCAT CAGCACCTGA TTAATTAACT CGAGGCAGCA GCAGCTCGGA
7501 TAGTATCGAC AACTCTGGA CGCTGCTCGT GTGATGGACT GTTGCCGCCA CACTTGCTGC
7561 CTTGACCTGT GAATATCCCT GCCGCTTTTA TCAAACAGCC TCAGTGTGTT TGATCTTGTG
7621 TGTACGCGCT TTTGCGAGTT GCTAGCTGCT TGTGCTATT GCGAATACCA CCCCCAGCAT
7681 CCCCTTCCCT CGTTTCATAT CGCTTGATC CCAACCGCAA CTTATCTACG CTGTCTGCT
7741 ATCCCTCAGC GCTGCTCTG CTCCTGCTCA CTGCCCTCG CACAGCCTTG GTTTGGGCTC
7801 CGCCTGTATT CTCCTGGTAC TGCAACCTGT AAACCAGCAC TGCAATGCTG ATGCACGGGA
7861 AGTAGTGGGA TGGGAACACA AATGGAAGC TTGAGCTCTT GTTTTCCAGA AGGAGTTGCT
7921 CCTTGAGCCT TTCATTCTCA GCCTCGATAA CCTCCAAAGC CGCTCTAATT GTGGAGGGGG
7981 TTCGAATTTA AAAGCTTGGA ATGTTGGTTC GTGCGTCTGG AACAAGCCCA GACTTGTGTC
8041 TCACTGGGAA AAGGACCATC AGCTCCAAA AACTTGCCGC TCAAACCGCG TACCTCTGCT
8101 TTCGCGCAAT CTGCCCTGTT GAAATCGCCA CCACATTCAT ATTGTGACG TTGAGCAGTC
8161 TGTAAATTGCC TCAGAATGTG GAATCATCTG CCCCCTGTGC GAGCCCATGC CAGGCATGTC
8221 GCGGGCGAGG ACACCCGCCA CTCGTACAGC AGACCATTAT GCTACCTCAC AATAGTTTAT
8281 AACAGTGACC ATATTTCTCG AAGCTCCCCA ACGAGCACCT CCATGCTCTG AGTGGCCACC
8341 CCCCAGCCCT GGTGCTTGG GAGGGCAGGT CAACCGGCAT GGGGCTACCG AAATCCCCGA
8401 CCGGATCCCA CCACCCCGC GATGGGAAGA ATCTCTCCCC GGGATGTGGG CCCACCACCA
8461 GCACAACCTG CTGGCCAGG CGAGCGTCAA ACCATAACCAC ACAAATATCC TTGGCATCGG
8521 CCCTGAATTC CTTCTGCCG TCTGTACCC GGTGCTTCTG TCCGAAGCAG GGTGTGCTAG
8581 GGATCGCTCC GAGTCCGCAA ACCCTTGTG CGTGGCGGGG CTTGTTTCAG CTTGAAGAGC
8641 CTCTAGAGTC GACCTGCAGG CATGCAAGCT TGGCGTAATC ATGGTCATAG CTGTTTCTCT
8701 TGTGAAATTG TTATCCGCTC ACAATTCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA
8761 AAGCTGGGG TGCTTAATGA GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG
8821 CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCAATTAAT AATCGGCCAA CGCGCGGGGA
8881 GAGGCGGTTT GCGTATTGGG CGCTCTTCCG CTTCTCTCGT CACTGACTCG CTGCGCTCGG
8941 TCGTTCGGCT GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG
9001 AATCAGGGGA TAACGAGGA AAGAACATGT GAGCAAAAG CCAGCAAAAG GCCAGGAACC
9061 GTAAAAAGGC CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCTGAC GAGCATCACA
9121 AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAGA TACCAGGCGT
9181 TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CTTGCCGCTT ACCGGATACC
9241 TGTCCGCCTT TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC

9301 TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTACGC
 9361 CCGACCGCTG CGCCTTATCC GGTAACATATC GTCTTGAGTC CAACCCGGTA AGACACGACT
 9421 TATCGC

SEQ ID NO: 230

pSZ1410

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
 61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
 121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAA
 181 CCACCGCTGG TAGCGGTGGT TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
 241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC
 301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
 361 ATTAATAATG AAGTTTAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
 421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
 481 TTGCTGACT CCCCCTCGTG TAGATACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
 541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAAC
 601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
 661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
 721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
 781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
 841 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
 901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
 961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
 1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
 1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
 1141 GTTCGATGTA ACCCACTCGT GCACCCAAC TATCTTTCAGC ATCTTTTACT TTCACCAGCG
 1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
 1261 GGAATGTTG AATACTCATA CTCTTCCTT TCAATATTA TTGAAGCATT ATCAGGGTT
 1321 ATTGTCTCAT GAGCGGATAC ATATTTGAA GTATTTAGAA AAATAACAA ATAGGGGTT
 1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
 1441 TAACCTATAA AAATAGGCGT ATCAGGAGGC CCTTTCGCT CGCGCGTTT GGTGATGACG
 1501 GTGAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
 1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CCGGGCTGGC
 1621 TTAATATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
 1681 CGCAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT AGGCTGCGCA
 1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGTT ACGAAGGGG
 1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
 1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCCGCCTGG AGCTGGTGCA GAGCATGGGG
 1921 CAGTTTGCGG AGGAGAGGGT GCTCCCGTGT CTGCACCCG TGGACAAGCT GTGGCAGCCG
 1981 CAGGACTTCC TGCCCGACCC CGAGTCGCCC GACTTCGAGG ACCAGGTGGC GGAGCTGCGC
 2041 GCGCGCGCCA AGGACCTGCC CGACGAGTAC TTTGTGGTGC TGGTGGGCGA CATGATCAG
 2101 GAGGAGGCGC TGCCGACCTA CATGGCCATG CTCAACACCT TGGACGGTGT GCGCGACGAC
 2161 ACGGCGCGCG CTGACCACCC GTGGGCGCGC TGGACGCGGC AGTGGGTGGC CGAGGAGAAC
 2221 CGGCACGGCG ACCTGCTGAA CAAGTACTGT TGGCTGACGG GCGCGTCAA CATGCGGGCC
 2281 GTGGAGGTGA CCATCAACAA CCTGATCAAG AGCGGCATGA ACCCGCAGAC GGACAACAAC
 2341 CCTTACTTGG GCTTCGTCTA CACCTCCTTC CAGGAGCGCG CCACCAAGTA GGTACCTTTT
 2401 CTTGCGCTAT GACACTTCCA GCAAAAGGTA GGGCGGGCTG CGAGACGGCT TCCCGGCGCT
 2461 GCATGCAACA CCGATGATGC TTCGACCCCC CGAAGCTCCT TCGGGGCTGC ATGGGCGCTC
 2521 CGATGCCGCT CCAGGGCGAG CGCTGTTTTA ATAGCCAGGC CCCCATTGC AAAGACATTA
 2581 TAGCGAGCTA CCAAAGCCAT ATTCAAACAC CTAGATCACT ACCACTTCTA CACAGGCCAC
 2641 TCGAGCTTGT GATCGCACTC CGCTAAGGGG GCGCCTCTTC CTCTTCGTTT CAGTCAACAAC
 2701 CCGCAAACCTC TAGAATATCA ATGCTGCTGC AGGCCTTCCT GTTCTGTCTG GCCGGCTTCG
 2761 CCGCAAGAT CAGCGCCTCC ATGACGAACG AGACGTCCGA CCGCCCCCTG GTGCACTTCA
 2821 CCCCCAACAA GGGCTGGATG AACGACCCCA ACGGCCTGTG GTACGACGAG AAGGACGCCA
 2881 AGTGGCACCT GTACTTCCAG TACAACCCGA ACGACACCGT CTGGGGGACG CCCTTGTTC
 2941 GGGGCCACGC CACGTCCGAC GACCTGACCA ACTGGGAGGA CCAGCCCATC GCCATCGCCC
 3001 CGAAGCGCAA CGACTCCGGC GCCTTCTCCG GCTCCATGGT GGTGGACTAC AACACACCT
 3061 CCGGCTTCTT CAACGACACC ATCGACCCGC GCCAGCGCTG CGTGGCCATC TGGACCTACA
 3121 ACACCCCGGA GTCCGAGGAG CAGTACATCT CCTACAGCCT GGACGGCGGC TACACCTTCA

3181 CCGAGTACCA GAAGAACCCC GTGCTGGCCG CCAACTCCAC CCAGTTCGCG GACCCGAAGG
3241 TCTTCTGGTA CGAGCCCTCC CAGAAGTGGA TCATGACCGC GGCCAAGTCC CAGGACTACA
3301 AGATCGAGAT CTACTCCTCC GACGACCTGA AGTCCTGGAA GCTGGAGTCC GCGTTCGCCA
3361 ACGAGGGCTT CCTCGGCTAC CAGTACGAGT GCCCCGGCCT GATCGAGGTC CCCACCGAGC
3421 AGGACCCAG CAAGTCCTAC TGGGTGATGT TCATCTCCAT CAACCCCGGC GCCCCGGCCG
3481 GCGGCTCCTT CAACCACTAC TTCGTCGGCA GCTTCAACGG CACCCACTTC GAGGCCTTCG
3541 ACAACCAAGT CCGCGTGGTG GACTTCGGCA AGGACTACTA CGCCCTGCAG ACCTTCTTCA
3601 ACACCGACCC GACCTACGGG AGCGCCCTGG GCATCGCGTG GGCCTCCAAC TGGGAGTACT
3661 CCGCCTTCGT GCCACCAAC CCTGGCGCT CCTCCATGTC CCTCGTGC GC AAGTTCTCCC
3721 TCAACACCGA GTACCAGGCC AACCCGAGA CGGAGCTGAT CAACCTGAAG GCCGAGCCGA
3781 TCCTGAACAT CAGCAACGCC GGCCCTTGGA GCCGGTTCGC CACCAACACC ACCTTGACGA
3841 AGGCCAACAG CTACAACGTC GACCTGTCCA ACAGCACCGG CACCCTGGAG TTCGAGCTGG
3901 TGTACGCCGT CAACACCACC CAGACGATCT CCAAGTCCGT GTTCGCGGAC CTCTCCCTCT
3961 GGTTCAGGG CCTGGAGGAC CCCGAGGAGT ACCTCCGCAT GGGCTTCGAG GTGTCCCGCT
4021 CCTCCTTCTT CCTGGACCGC GGGAACAGCA AGGTGAAGTT CGTGAAGAG AACCCCTACT
4081 TCACCAACCG CATGAGCGTG AACAAACAGC CCTTCAAGAG CGAGAACGAC CTGTCTACT
4141 ACAAGGTGTA CGGCTTGCTG GACCAGAACA TCCTGGAGCT GTACTTCAAC GACGGCGACG
4201 TCGTGTCCAC CAACACCTAC TTCATGACCA CCGGGAACGC CCTGGGCTCC GTGAACATGA
4261 CGACGGGGGT GGACAACCTG TTCTACATCG ACAAGTTCCA GGTGCGCGAG GTCAAGTGAC
4321 AATTGGCAGC AGCAGCTCGG ATAGTATCGA CACACTCTGG ACCTGGTTCG TGTGATGGAC
4381 TGTGCGCGC ACCTTGCTG CCTTGACCTG TGAATATCCC TGCCGCTTTT ATCAAACAGC
4441 CTCAGTGTGT TTGATCTTGT GTGTACGCGC TTTTGCAGT TGCTAGCTATT TTGTGCTATT
4501 TCGGAATACC ACCCCAGCA TCCCTTCCC TCGTTTCATA TCGCTTGCAAT CCCAACCAGCA
4561 ACTTATCTAC GCTGTCTGTC TATCCCTCAG CGCTGCTCCT GCTCCTGCTC ACTGCCCCCTC
4621 GCACAGCCTT GGTTCGGGCT CCGCTGTAT TCTCCTGGTA CTGCAACCTG TAAACCAGCA
4681 CTGCAATGCT GATGCACGGG AAGTAGTGGG ATGGGAACAC AAATGGAGGA TCCCGCGTCT
4741 CGAACAGAGC GCGCAGAGGA ACGCTGAAGG TCTCGCCTCT GTGCGACCTC AGCGCGGCAT
4801 ACACCACAAT AACCACTGA CGAATGCGCT TGGTTCCTCG TCCATTAGCG AAGCGTCCGG
4861 TTCACACAG TGCCACGTTG GCGAGGTGGC AGGTGACAAT GATCGGTGGA CTGATGCTGC
4921 GAAACGTTCA CAGCCTAGGG ATATCGAATT CGGCCGACAG GACGCGCGTC AAAGGTGCTG
4981 GTCGTGTATG CCTGGCCGG CAGGTGCTTG CTGCTGCTGG TTAGTGATTC CGCAACCCTG
5041 ATTTTGGCGT CTTATTTTGG CGTGGCAAAC GCTGGCGCCC GCGAGCCGGG CCGGCGGCGA
5101 TGGCGTGCCC CACGCTGCG GGAATCCAAG GGAGGCAAGA GCGCCCGGGT CAGTTGAAGG
5161 GCTTTACGCG CAAGGTACAG CCGCTCCTGC AAGGCTGCGT GGTGGAATTG GACGTGCAGG
5221 TCCTGTGAA GTTCTCTCAC CGCTCACCA GCGGACAAA CACCGGTGTA TCAGGTCCGT
5281 GTCATCCACT CTAAAGAACT GCACTACGAC CTACTGATGG CCCTAGATT TTAACAATA
5341 ACGCCTGAGA CACTTGCCCA GGATTGAAAC TCCCTGAAGG GACCACAGG GGCCTGAGT
5401 TGTTCTTCC CCCCCTGGCG AGCTGCCAGC CAGGCTGTAC CTGTGATCGA GGCTGGCGGG
5461 AAAATAGGCT TCGTGTGCTC AGGTCATGGG AGGTGCAGGA CAGCTCATGA AACGCCAACA
5521 ATCGCACAAAT TCATGTCAAG CTAATCAGCT ATTTCTCTT CACGAGCTGT AATTGTCCCA
5581 AAATTCTGGT CTACCGGGGG TGATCCTTCG TGTACGGGCC CTTCCCTCAA CCCTAGGTAT
5641 GCGCGCATGC GGTGCGCGCG CAACTCGCGC GAGGGCCGAG GGTTCGGGAC GGGCCGTCCC
5701 GAAATGCAGT TGCACCCGGA TCGTGCGCAC CTTTTTTGCG ATAATTTATG CAATGGACTG
5761 CTCTGCAAAA TTCTGGCTCT GTCGCCAACC CTAGGATCAG CGGCGTAGGA TTTCGTAATC
5821 ATTCGTCTCTG ATGGGGAGCT ACCGACTACC CTAATATCAG CCCGACTGCC TGACGCCAGC
5881 GTCCACTTTT GTGCACACAT TCCATTCTGT CCAAGACAT TTCATTGTGG TGCGAAGCGT
5941 CCCCAGTTAC GCTCACCTGT TTCCCGACCT CTTACTGTT CTGTGACAG AGCGGGCCCA
6001 CAGGCCGGTC GCAGCCACTA GTATGTTGGT GGCCGCCGCC GCCAGCAGCG CCTTCTTCCC
6061 CGTGCCCGCC CCCC GCCCAAGCC CGGCAAGTTC GGCAACTGGC CCAGCAGCCT
6121 GAGCCAGCCC TTCAAGCCCC AGAGCAACCC CAACGGCCGC TTCCAGGTGA AGGCCAACGT
6181 GAGCCCCCAC GGGCGCGCCC CCAAGGCCAA CGGCAGCGCC GTGAGCCTGA AGTCCGGCAG
6241 CCTGAACACC CTGGAGGACC CCCCCAGCAG CCCCCCCCC CGCACCTTCC TGAACCAGCT
6301 GCCCGACTGG AGCCGCCTGC GCACCGCCAT CACCACCGTG TTCGTGGCCG CCGAGAAGCA
6361 GTTCAACCGC CTGGACCGCA AGAGCAAGCG CCCCACATG CTGGTGGACT GGTTCGGCAG
6421 CGAGACCATC GTGACGAGCG GCCTGTGTGT CCGCGAGCGC TTCAGCATCC GCAGCTACGA
6481 GATCGGCGCC GACCGCACCG CCAGCATCGA GACCCTGATG AACCACTGC AGGACACCAG
6541 CCTGAACACC TGCAAGAGCG TGGGCTGTGT GAACGACGGC TTCGGCCGCA CCCCAGAT
6601 GTGCACCCGC GACCTGATCT GGGTGTGAC CAAGATGAC ATCGTGGTGA ACCCTGATCC
6661 CACCTGGGGC GACACCGTGG AGATCAACAG CTGGTTACGC CAGAGCGGCA AGATCGGCAT
6721 GGGCCGCGAG TGGCTGATCA GCGACTGCAA CACCGGCGAG ATCCTGGTGC GCGCCACCAG
6781 CGCCTGGGCC ATGATGAACC AGAAGACCCG CCGCTTCAGC AAGCTGCCCT GCGAGGTGCG

```

6841 CCAGGAGATC GCGGGGCACT TCGTGGACGC CCGGGGCGTG ATCGAGGACA ACGACCGCAA
6901 GCTGCACAAG TTCGACGTGA AGACCGGCGA CAGCATCTGC AAGGGCCTGA CCGGGGCTG
6961 GAACGACTTC GACGTGAACC AGCAGGTGAG CAACGTGAAG TACATCGGCT GGATTCTGGA
7021 GAGCATGCCC ACCGAGGTGC TGGAGACCCA GGAGCTGTGC AGCCTGACCC TGGAGTACCG
7081 CCGCGAGTGC GGCCGCGAGA GCGTGGTGGG GAGCGTGACC AGCATGAACC CCAGCAAGGT
7141 GGGCGACCGC AGCCAGTACC AGCACCTGCT GCGCCTGGAG GACGGCGCCG ACATCATGAA
7201 GGGCGCACCC GAGTGGCGCC CCAAGAACGC CGGCACCAAC CGCGCCATCA GCACCTGATT
7261 AATTAACCTG AGGCAGCAGC AGCTCGGATA GTATCGACAC ACTCTGGACG CTGGTCTGTG
7321 GATGGACTGT TGCCGCCACA CTTGCTGCCT TGACCTGTGA ATATCCCTGC CGCTTTTATC
7381 AAACAGCCTC AGTGTGTTTG ATCTGTGTGT TACGCGCTTT TCGGAGTTGC TAGCTGCTTG
7441 TGCTATTTGC GAATACCACC CCCAGCATCC CCTTCCCTCG TTTTATATCG CTTGCATCCC
7501 AACCGBAACT TATCTACGCT GTCCTGCTAT CCTCAGCGC TGCTCCTGCT CCTGCTCACT
7561 GCGGCTCGCA CAGCCTTGGT TTGGGCTCCG CTTGTATTCT CCTGGTACTG CAACCTGTAA
7621 ACCAGCACTG CAATGCTGAT GCACGGGAAG TAGTGGGATG GGAACACAAA TGGAAAAGCTT
7681 GAGCTCCAGC CACGGCAACA CCGCGCGCCT TCGCGCCGAG CACGGCGACA AGAACCTGAG
7741 CAAGATCTGC GGGCTGATCG CCAGCGACGA GGGCCGGCAC GAGATCGCCT ACACGCGCAT
7801 CGTGGACGAG TTCTTCCGCC TCAGCCCCGA GGGCGCCGTC GCCGCTACG CCAACATGAT
7861 GCGCAAGCAG ATCACCATGC CCGCGCACCT CATGGACGAC ATGGGCCACG GCGAGGCCAA
7921 CCGGGGCGCG AACCTCTTCG CCGACTTCTC CGCGGTGCGC GAGAAGATCG ACGTCTACGA
7981 CGCCGAGGAC TACTGCCGCA TCCTGGAGCA CCTCAACGCG CGCTGGAAGG TGGACGAGCG
8041 CCAGGTCAGC GGCCAGGCCG CCGCGGACCA GGAGTACGTC CTGGGCCTGC CCCAGCGCTT
8101 CCGGAAACTC GCCGAGAAGA CCGCCGCCAA GCGCAAGCGC GTCGCGCGCA GGCCCGTCGC
8161 CTTCTCTTGG AGAAGAGCCT CTAGAGTCGA CCTGCAGGCA TGCAAGCTTG GCGTAATCAT
8221 GGTCTATAGT GTTTCCTGTG TGAATTTGTT ATCCGCTCAC AATTCCACAC AACATACGAG
8281 CCGGAAGCAT AAAGTGTAAG GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG
8341 CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA
8401 TCGGCCAACG CGCGGGGAGA GCGGTTTTCG GTATTGGGCG CTCTTCCGCT TCCTCGCTCA
8461 CTGACTCGCT GCGCTCGGTC GTTCGGCTGC GCGGAGCGGT ATCAGCTCAC TCAAAGGCGG
8521 TAATACGGTT ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC
8581 AGCAAAAGGC CAGGAACCGT AAAAAAGGCC CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC
8641 CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC
8701 TATAAAGATA CCAGGCGTTT CCCCTTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC
8761 TGCCGCTTAC CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA
8821 GCTCAGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCT CTCCAAGCTG GGCTGTGTGC
8881 ACGAACCCTC CGTTCAGCCC GACCGCTGCG CTTATCCGG TAACATCTGT CTTGAGTCCA
8941 ACCCGGTAAG ACACGACTTA TCGC

```

SEQ ID NO: 231

pSZ1413

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAA
181 CCACCGCTGG TAGCGGTGGT TTTTGTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAATAATG AAGTTTAA TCAATCTAAA GTATATATGA GTAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCTATCCATG
481 TTGCTGACT CCGCTGCTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCACTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCTAGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTCTTTCG GGGCGAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG

```

1201 TTTCTGGGTG AGCAAAAACA GGAAGGCCAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTG AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGCTG CGGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCCGCTGG AGCTGGTGCA GAGCATGGGG
1921 CAGTTTGCAG AGGAGAGGGT GCTCCCCGTG CTGCACCCCG TGGACAAGCT GTGGCAGCCG
1981 CAGGACTTCC TGCCCGACCC CGAGTCGCCC GACTTCGAGG ACCAGGTGGC GGAGCTGCGC
2041 GCGCGCGCCA AGGACCTGCC CGACGAGTAC TTTGTGGTGC TGGTGGGCGA CATGATCAGC
2101 GAGGAGGCGC TGCCGACCTA CATGGCCATG CTCAACACCT TGGACGGTGT GCGCGAGAC
2161 ACGGGCGCGG CTGACCCACC GTGGCGCGCG TGGACGCGGC AGTGGGTGGC CGAGGAGAAC
2221 CGGCACGGCG ACCTGCTGAA CAAGTACTGT TGGCTGACGG GCGCGCTCAA CATGCGGGCC
2281 GTGGAGGTGA CCATCAACAA CCTGATCAAG AGCGGCATGA ACCCGCAGAC GGACAACAAC
2341 CCTTACTTGG GCTTCGTCTA CACCTCCTTC CAGGAGCGCG CCACCAAGTA GGTACCCTTT
2401 CTTGCGCTAT GACACTTCCA GCAAAGGTA GGGCGGGCTG CGAGACGGCT TCCCGGCGCT
2461 GCATGCAACA CCGATGATGC TTCGACCCCT CGAAGCTCCT TCGGGGCTGC ATGGGCGCTC
2521 CGATGCCGCT CCAGGGCGAG CGCTGTTTAA ATAGCCAGGC CCCCGATTGC AAAGACATTA
2581 TAGCGAGCTA CCAAAGCCAT ATTCAAACAC CTAGATCACT ACCACTCTA CACAGGCCAC
2641 TCGAGCTTGT GATCGCACTC CGCTAAGGGG GCGCCTCTTC CTCTTCGTTT CAGTCACAAC
2701 CCGCAAACCT TAGAATATCA ATGCTGCTGC AGGCCTTCCT GTTCCTGCTG GCCGGCTTCG
2761 CCGCCAAGAT CAGCGCCTCC ATGACGAACG AGACGTCCGA CCGCCCCCTG GTGCACTTCA
2821 CCCCCAACAA GGGCTGGATG AACGACCCCA ACGGCCTGTG GTACGACGAG AAGGACGCCA
2881 AGTGGCACCT GTACTTCCAG TACAACCCGA ACGACACCGT CTGGGGGACG CCCTTGTTC
2941 GGGGCCACGC CACGTCCGAC GACCTGACCA ACTGGGAGGA CCAGCCCATC GCCATCGCCC
3001 CGAAGCGCAA CGACTCCGGC GCCTTCTCCG GCTCCATGGT GGTGGACTAC AACAACACCT
3061 CCGGCTTCTT CAACGACACC ATCGACCCGC GCCAGCGCTG CGTGGCCATC TGGACCTACA
3121 ACACCCCGGA GTCCGAGGAG CAGTACATCT CCTACAGCCT GGACGGCGGC TACACCTTCA
3181 CCGAGTACCA GAAGAACCCC GTGCTGGCCG CCAACTCCAC CCAGTTCCGC GACCCGAAGG
3241 TCTTCTGGTA CGAGCCCTCC CAGAAGTGGA TCATGACCGC GGCCAAGTCC CAGGACTACA
3301 AGATCGAGAT CTACTCCTCC GACGACCTGA AGTCCTGGAA GCTGGAGTCC GCGTTCGCCA
3361 ACGAGGGCTT CCTCGGCTAC CAGTACGAGT GCCCCGGCCT GATCGAGGTC CCCACCGAGC
3421 AGGACCCCGA CAAGTCCTAC TGGGTGATGT TCATCTCCAT CAACCCGGC GCGCCGGCCG
3481 GCGGCTCCTT CAACCAGTAC TTCGTCCGGA GCTTCAACGG CACCCACTTC GAGGCTTTCG
3541 ACAACAGTGC CCGCGTGGTG GACTTCGGCA AGGACTACTA CGCCCTGCAG ACCTTCTTCA
3601 ACACCGACCC GACCTACGGG AGCGCCCTGG GCATCGCGTG GGCTTCCAAC TGGGAGTACT
3661 CCGCCCTCGT GCCACCAAC CCCTGGCGCT CCTCCATGTC CCTCGTGCAG AAGTTCTCCC
3721 TCAACACCGA GTACCAGGCC AACCCGAGGA CGGAGCTGAT CAACCTGAAG GCCGAGCCGA
3781 TCCTGAACAT CAGCAACGCC GGCCCTTGGG GCCGGTTTCG CACCAACACC ACGTTGACGA
3841 AGGCCAACAG CTACAACGTC GACCTGTCCA ACAGCACCGG CACCTGGAG TTCGAGCTGG
3901 TGTACGCCGT CAACACCACC CAGACGATCT CCAAGTCCGT GTTCGCGGAC CTCTCCCTCT
3961 GGTTCAGGGG CCTGGAGGAC CCCGAGGAGT ACCTCCGCAT GGGCTTCGAG GTGTCCGCGT
4021 CCTCCTTCTT CTGGACCGC GGGAACAGCA AGGTGAAGTT CGTGAAGGAG AACCCCTACT
4081 TCACCAACCG CATGAGCGTG AACAACCAGC CCTTCAAGAG CGAGAACGAC CTGTCTTACT
4141 ACAAGGTGTA CGGCTTGCTG GACCAGAACA TCCTGGAGCT GTACTTCAAC GACGGCGACG
4201 TCGTGTCCAC CAACACCTAC TTCATGACCA CCGGGAACGC CCTGGGCTCC GTGAACATGA
4261 CGACGGGGGT GGACAACCTG TTCTACATCG ACAAGTTCCA GGTGCGCGAG GTCAAGTGAC
4321 AATTGGCAGC AGCAGCTCGG ATAGTATCGA CACACTCTGG ACGCTGGTCG TGTGATGGAC
4381 TGTGCGCGCC ACACTTGCTG CCTTGACCTG TGAATATCCC TGCCGCTTTT ATCAAACAGC
4441 CTCAGTGTGT TTGATCTTGT GTGTACGCGC TTTTGCGAGT TGCTAGCTGC TTGTGCTATT
4501 TGCGAATACC ACCCCAGCA TCCCCTTCCC TCGTTTCATA TCGCTTGCAT CCAACCGCA
4561 ACTTATCTAC GCTGTCTGCT TATCCCTCAG CGCTGCTCCT GCTCCTGCTC ACTGCCCTC
4621 GCACAGCCTT GGTTCGGGCT CCGCCTGTAT TCTCCTGGTA CTGCAACCTG TAAACCAGCA
4681 CTGCAATGCT GATGCACGGG AAGTAGTGGG ATGGGAACAC AAATGGAGGA TCCGCGCTCT
4741 CGAACAGAGC GCGCAGAGGA ACGCTGAAGG TCTCGCCTCT GTCGCACCTC AGCGCGCAT
4801 ACACCACAAT AACCACCTGA CGAATGCGCT TGGTTCTTCG TCCATTAGCG AAGCGTCCGG

4861 TTCACACACG TGCCACGTTG GCGAGGTGGC AGGTGACAAT GATCGGTGGA GCTGATGGTC
4921 GAAACGTTCA CAGCCTAGGG ATATCGAATT CGGCCGACAG GACGCGCGTC AAAGGTGCTG
4981 GTCGTGTATG CCCTGGCCGG CAGGTCGTTG CTGCTGCTGG TTAGTGATTG CGCAACCCCTG
5041 ATTTTGGCGT CTTATTTTGG CGTGGCAAAC GCTGGCGCCC GCGAGCCGGG CCGGCGGCGA
5101 TGCGGTGCCC CACGGCTGCC GGAATCCAAG GGAGGCAAGA GCGCCCGGT CAGTTGAAGG
5161 GCTTTACGCG CAAGGTACAG CCGCTCCTGC AAGGCTGCGT GGTGGAATTG GACGTGCAGG
5221 TCCTGCTGAA GTTCTCCAC CGCTCACCA GCGGACAAAG CACCGGTGTA TCAGGTCCGT
5281 GTCATCCACT CTAAAGAACT CGACTACGAC CTACTGATGG CCCTAGATTG TTCATCAAAA
5341 ACGCCTGAGA CACTTGCCCA GGATTGAAAC TCCCTGAAGG GACCACCAGG GGCCCTGAGT
5401 TGTTCTTTC CCCCCTGGCG AGCTGCCAGC CAGGCTGTAC CTGTGATCGA GGCTGGCGGG
5461 AAAATAGGCT TCGTGTGCTC AGGTGATGGG AGGTGCAGGA CAGCTCATGA AACGCCAACA
5521 ATCGCACAAAT TCATGTCAAG CTAATCAGCT ATTTCTCTTT CACGAGCTGT AATTGTCCCA
5581 AAATTCGTGT CTACCGGGGG TGATCCTTCG GTACGGGGC CTTCCCTCAA CCCTAGGTAT
5641 GCGCGCATGC GGTGCGCGCG CAACTCGCGC GAGGCGCGAG GGTGTTGGGAC GGGCCGTCCC
5701 GAAATGCAGT TGCACCCGGA TGCCTGGCAC CTTTTTTGCG ATAATTTATG CAATGGACTG
5761 CTCTGCAAAA TTCTGGCTCT GTCCCAACC CTAGGATCAG CGGCGTAGGA TTTCGTAATC
5821 ATTCGTCTCT ATGGGGAGCT ACCGACTACC CTAATATCAG CCCGACTGCC TGACGCCAGC
5881 GTCCACTTTT GTGCACACAT TCCATTCTGT CCCAAGACAT TTCATTGTGG TGCGAAGCGT
5941 CCCCAGTTAC GCTACCTGT TTTCCGACCT CCTTACTGTT CTGTCGACAG AGCGGGCCCCA
6001 CAGGCGGGT GCAGCCACTA CAAGACGAAC AGGCAGCCCTG TGGAGAACCC
6061 TCCGTTACAG ATCGGGACGC TGCCCAAGGC CATCCCGCG CACTGTTTCG AGCGCTCGGC
6121 GCTTCGTGGG CGCGCCCCCA AGGCCAACGG CAGCGCCGTG AGCCTGAAGT CCGGCAGCCT
6181 GAACACCCTG GAGGACCCCC CCAGCAGCCC CCCCCCCGC ACCTTCCTGA ACCAGCTGCC
6241 CGACTGGAGC CGCTGCGCA CCGCCATCAC CACCGTGTTC GTGGCCGCGG AGAAGCAGTT
6301 CACCCGCCTG GACCGCAAGA GCAAGCGCCC CGACATGCTG GTGGACTGGT TCGGCAGCGA
6361 GACCATCGTG CAGGACGGCC TGGTGTTCG CGAGCGCTTC AGCATCCGCA GCTACGAGAT
6421 CGGCGCCGAC CGCACCGCCA GCATCGAGAC CCTGATGAAC CACCTGCAGG ACACCAGCCT
6481 GAACCACTGC AAGAGCGTGG GCCTGCTGAA CGACGGCTTC GGCCGCAACC CCGAGATGTG
6541 CACCCGCGAC CTGATCTGGG TGCTGACCAA GATGCAGATC GTGGTGAACC GCTACCCAC
6601 CTGGGGCGAC ACCGTGGAGA TCAACAGCTG GTTCAGCCAG AGCGGCAAGA TCGGCATGGG
6661 CCGCGAGTGG CTGATCAGCG ACTGCAACAC CGGCGAGATC CTGGTGCAGG CCACCAGCGC
6721 CTGGGCCATG ATGAACCAGA AGACCCGCGG CTTACGCAAG CTGCCCTGCG AGGTGCGCCA
6781 GGAGATCGCC CCCCCTTCG TGGACGCCCC CCCCCTGATC GAGGACAACG ACCGCAAGCT
6841 GCACAAGTTC GACGTGAAGA CCGGCGACAG CATCTGCAAG GGCCTGACCC CCGCTGGGAA
6901 CGACTTCGAC GTGAACGAGC ACGTGAGCAA CGTGAAGTAC ATCGGCTGGA TTCTGGAGAG
6961 CATGCCACAC GAGGTGCTGG AGACCCAGGA GCTGTGCAGC CTGACCCTGG AGTACCGCCG
7021 CGAGTGCAGG CGCGAGAGCG TGGTGGAGAG CGTGACCAGC ATGAACCCCA GCAAGGTGGG
7081 CGACCGCAGC CAGTACCAGC ACCTGCTGCG CCTGGAGGAC GCGCGCGACA TCATGAAGGG
7141 CCGCACCGAG TGGCGCCCCA AGAACGCGCG CACCAACCGC GCCATCAGCA CCTGATTAAT
7201 TAACTCGAGG CAGCAGCAGC TCGGATAGTA TCGACACACT CTGGACGCTG GTCGTGTGAT
7261 GGACTGTTGC CGCCACACTT GCTGCCTTGA CTTGTGAATA TCCCTGCCGC TTTTATCAAA
7321 CAGCCTCAGT GTGTTTGATC TTGTGTGTAC GCGCTTTTGC GAGTTGTAGC GTGTTGTGTC
7381 TATTTGCGAA TACCACCCCC AGCATCCCC TCCCTCGTTT CATATCGCTT GCATCCCAAC
7441 CGCAACTTAT CTACGCTGTC CTGCTATCCC TCAGCGCTGC TCCTGCTCCT GCTCACTGCC
7501 CCTCGCACAG CCTTGGTTTG GGCTCCGCTT GTATTCTCCT GGTACTGCAA CCTGTAAACC
7561 AGCACTGCAA TGCTGATGCA CGGGAAGTAG TGGGATGGGA ACACAAATGG AAAGCTTGAG
7621 CTCCAGCCAC GGCAACACCG CGCGCCTTGC GGCCGAGCAC GCGGACAAGA ACCTGAGCAA
7681 GATCTGCGGG CTGATCGCCA GCGACGAGGG CCGGCACGAG ATCGCCTACA CGCGCATGCT
7741 GGACGAGTTC TTCCGCTCG ACCCCGAGGG CGCCGTCGCC GCCTACGCCA ACATGATCGT
7801 CAAGCAGATC ACCATGCCCG CGCACCTCAT GGACGACATG GGCCACGGCG AGGCCAACCC
7861 GGGCCGCAAC CTCTTCGCGG ACTTCTCCGC GGTGCGCGAG AAGATCGACG TCTACGACGC
7921 CGAGGACTAC TGCCGCATCC TGGAGCACCT CAACGCGCGC TGGGAAGGTG ACGAGCGCCA
7981 GGTGAGCGGC CAGGCCGCGG CGGACGAGGA GTACGTCTTG GGCCTGCCCG AGCGCTTCCG
8041 GAAACTCGCC GAGAAGACCG CCGCCAAGCG CAAGCGCGTC GCGCGCAGGC CCGTCGCCTT
8101 CTCTGGAGA AGAGCCTCTA GAGTCGACCT GCAGGCATGC AAGCTTGGCG TAATCATGGT
8161 CATAGCTGTT TCCTGTGTGA AATTGTTATC CGCTCACAAT TCCACACAAC ATACGAGCCG
8221 GAAGCATAAA GTGTAAAGCC TGGGGTGCTT AATGAGTGAG CTAACCTACA TTAATGCGT
8281 TGCGCTCACT GCCCGCTTTC CAGTCGGGAA ACCTGTCGTG CCAGCTGCAT TAATGAATCG
8341 GCCAACGCGC GGGGAGAGGC GGTGTCGTG TGGGCGCTC TTCCGCTTCC TCGCTCACTG
8401 ACTCGCTGCG CTCGGTCTGT CCGCTGCGGC GAGCGGTATC AGCTCACTCA AAGGCGGTAA
8461 TACGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC

```

8521 AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
8581 CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
8641 AAAGATACCA GGCCTTTCCC CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT CCGACCCTGC
8701 CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTGCGGAAG CGTGGCGCTT TCTCATAGCT
8761 CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGITCGCTC CAAGCTGGGC TGTGTGCACG
8821 AACCCCCCGT TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
8881 CGGTAAGACA CGACTTATCG C

```

SEQ ID NO: 232
pSZ1491

```

1  CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61  AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTITGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTTCT TCATCCATAG
481 TTGCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GTCCCGGTTT CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCTCTCG ATCGTTGTCA GAAGTAAAGT GGCCGCAGTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTGCCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTCTCTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCCAAT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTCTGCGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCGC CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCGC TCGTGCCTGT CGCTGATGTC
1981 CATCACCAGG TCCATGAGGT CTGCCTTGCG CCGGCTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGTGC CAGGGTCCTG ACGTGGTGC GGTCTGCGG
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAACCTGG CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGAAATC CTACCAGTCA TGGCTTTTAC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCTCGCT GCCGCCGCTT CTCCGCGACG CTCTTTTCCA
2341 GCACCGTGAT GCGCGGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGCGCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACCTTT TTGGCGAAGG CATCGGCCCT GGCCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGGACGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCCTCT TCGTTTCAGT CACAACCCGC AAACCTCTAG ATATCAATGC

```

2941 TGCTGCAGGC CTTCTGTGTC CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCCAACGG CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCCGAACGA CACCGTCTGG GGGACGCCCT TGTTCCTGGG CCACGCCACG TCCGACGACC
3181 TGACCAACTG GGAGGACCAG CCCATCGCCA TCGCCCCGAA GCGCAACGAC TCCGGCGCCT
3241 TCTCCGGCTC CATGGTGGTG GACTACAACA ACACCTCCGG CTTCTTCAAC GACACCATCG
3301 ACCCGCGCCA GCGCTGCGTG GCCATCTGGA CCTACAACAC CCCGGAGTCC GAGGAGCAGT
3361 ACATCTCCTA CAGCCTGGAC GCGGGCTACA CTTTCACCGA GTACCAGAAG AACCCCGTGC
3421 TGGCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCGTCCCGA
3481 AGTGGATCAT GACCGCGGCC AAGTCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCGCGT TCGCCAACGA GGGCTTCCTC GGCTACCAGT
3601 ACGAGTGCCC CGGCTGATC GAGGTCCCCA CCGAGCAGGA CCCCAGCAAG TCCTACTGGG
3661 TGATGTTTAT CTCCATCAAC CCCGGCGCCC CGGCCGCGGG CTCCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CTTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCCTGGGCAT CGCGTGGGCC TCCAACTGGG AGTACTCCGC CTTCTGTCCC ACCAACCCTT
3901 GGGCTCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCCG AGCCGATCCT GAACATCAGC AACGCCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTCGACC
4081 TGTCCAACAG CACCGGCACC CTGGAGTTCG AGCTGGTGTG CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT GCGGACCTCT CCCTCTGGTT CAAGGGCCTG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCCTC CTTCTTCCTG GACCGCGGGA
4261 ACAGCAAGGT GAAGTTCGTG AAGGAGAACC CCTACTTCAC CAACCGCATG AGCGTGAACA
4321 ACCAGCCCTT CAAGAGCGAG AACGACCTGT CCTACTACAA GGTGTACGGC TTGCTGGACC
4381 AGAACATCCT GGAGCTGTAC TTCAACGACG GCGACGTCGT GTCCACCAAC ACCTACTTCA
4441 TGACCACCGG GAACGCCCTG GGCTCCGTGA ACATGACGAC GGGGGTGGAC AACCTGTTCT
4501 ACATCGACAA GTTCCAGGTG CGCGAGGTCA AGTGACAATT GGCAGCAGCA GCTCGGATAG
4561 TATCGACACA CTCTGACGCG TGGTCTGTGT ATGGACTGTT GCCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTTGTGTGT
4681 ACGCGCTTTT GCGAGTTGCT AGCTGCTTGT GCTATTTCG AATACCACCC CCAGCATCCC
4741 CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCTGCTATC
4801 CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGCTCCGC
4861 CTGTATTCTC CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGATG CACGGGAAGT
4921 AGTGGGATGG GAACACAAAT GGAGGATCCC GCGTCTCGAA CAGAGCGCGC AGAGGAACGC
4981 TGAAGGTCTC GCCTCTGTCT CACCTCAGCG CGGCATACAC CACAATAACC ACCTGACGAA
5041 TGGCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGGTTCA CACACGTGCC ACGTTGGCGA
5101 GGTGGCAGGT GACAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 GCAATTTCGG CGACAGGACG CGCGTCAAAG GTGCTGGTCG TGTATGCCCT TCTGCTCAGG
5221 TCGTTGCTGC TGCTGGTTAG TGATTCCGCA ACCCTGATTT TGGCGCTTGA TTTTGGCGTG
5281 GCAAACGCTG GCGCCCGCGA GCCGGGCGCG CGGCATGCG GTGCCCCACG GCTGCCGGAA
5341 TCCAAGGGAG GCAAGAGCGC CCGGGTCAGT TGAAGGGCTT TACGCGCAAG GTACAGCCGC
5401 TCCTGCAAGG CTGCGTGGTG GAATTGGACG TGCAGGTCTT GCTGAAGTTC CTCCACCGCC
5461 TCACCAGCGG ACAAAGCACC GGTGTATCAG GTCCGTGTCA TCCACTCTAA AGAACTCGAC
5521 TACGACCTAC TGATGGCCCT AGATTCTTCA TCAAAAACGC CTGAGACACT TGCCCAGGAT
5581 TGAAACTCCC TGAAGGGACC ACCAGGGGCC CTGAGTTGTT CTTTCCCCC GTGGCGAGCT
5641 GCCAGCCAGG CTGTACCTGT GATCGAGGCT GCGGGGAAAA TAGGCTTCGT GTGCTCAGGT
5701 CATGGGAGGT GCAGGACAGC TCAIGAAACG CCAACAATCG CACAATTAT GTCAAGCTAA
5761 TCAGCTATTT CCTCTTCACG AGCTGTAAAT GTCCCAAAT TCTGGTCTAC CGGGGGTGAT
5821 CTTCTGTGTA CGGGCCCTTC CCTCAACCCCT AGGTATGCGC GCATGCGGTC GCCGCGCAAC
5881 TCGCGCGAGG GCCGAGGGTT TGGGACGGGC CGTCCCGAAA TGCAGTTGCA CCCGGATGCG
5941 TGGCACCTTT TTTGCGATAA TTTATGCAAT GGACTGCTCT GCAAAATTCT GGCTCTGTCT
6001 CCAACCCTAG GATCAGCGGC GTAGGATTTT GTAATCATTC GTCTGTATGG GGAGCTACCG
6061 ACTACCTTAA TATCAGCCCG ACTGCCTGAC GCCAGCGTCC ACTTTTGTGC ACACATTCCA
6121 TTCGTGCCCA AGACATTTCA TTGTGGTGGC AAGCGTCCCC AGTTACGCTC ACCTGTTTCC
6181 CGACCTCCTT ACTGTTCTGT CGACAGAGCG GGGCCACAGG CCGGTGCGAG CCCATATGGC
6241 TTCCGCGGCA TTCACCATGT CGGCGTGCCC CGCGATGACT GGCAGGGCCC CTGGGGCAGC
6301 TCGCTCCGGA CGGCCAGTCG CCACCCGGCT GAGGTACGTA TTCCAGTGCC TGGTGGCCAG
6361 CTGCATCGAC CCTTGCAGCC AGTACCGCAG CAGCGCCAGC CTGAGCTTCC TGGGCGACAA
6421 CGGCTTCGCC AGCCTGTTCG GCAGCAAGCC CTTATGAGC AACCGCGGCC ACCCGGCCT
6481 GCGCCGCGCC AGCCACAGCG GCGAGGCCAT GGCCTGGGCC CTGCAGCGCC CCCAGGAGG
6541 CGGCACCAAG AAGAAGCCCC TGATCAAGCA GCGCCGCGTG GTGGTGACCG GCATGGGCGT

6601 GGTGACCCCC CTGGGCCACG AGCCCGACGT GTTCTACAAC AACCTGCTGG ACGGCGTGAG
6661 CGGCATCAGC GAGATCGAGA CCTTCGACTG CACCCAGTTC CCCACCCGCA TCGCCGGCGA
6721 GATCAAGAGC TTCAGCACCG ACGGCTGGGT GGCCCCAAG CTGAGCAAGC GCATGGACAA
6781 GTTCATGCTG TACCTGCTGA CCGCCGCAAA GAAGGCCCTG GCCGACGGCG GCATCACCGA
6841 CGAGGTGATG AAGGAGCTGG ACAAGCGCAA GTGCGGCGTG CTGATCGGCA GCGGCATGGG
6901 CGGCATGAAG GTGTTCAACG ACGCCATCGA GGCCCTGCGC GTGAGCTACA AGAAGATGAA
6961 CCCCTTCTGC GTGCCCTTCG CCACCACCAA CATGGGCAGC GCCATGCTGG CCATGGACCT
7021 GGGCTGGATG GGCCCCAACT ACAGCATCAG CACCGCCTGC GCCACCAGCA ACTTCTGCAT
7081 CCTGAACGCC GCCAACCACA TCATCCGCGG CGAGGCCGAC ATGATGCTGT GCGGCGGCAG
7141 CGACGCCGTG ATCATCCCCA TCGGCCCTGG CGGCTTCGTG GCCTGCCGCG CCCTGAGCCA
7201 GCGCAACAGC GACCCACCA AGGCCAGCCG CCCCTGGGAC AGCAACCCG ACGGCTTCGT
7261 GATGGGCGAG GGCGCCGGCG TGCTGCTGCT GGAGGAGCTG GAGCACGCCA AGAAGCGCGG
7321 CGCCACCATC TACGCCGAGT TCCTGGGCGG CAGCTTCACC TCGACGCCT ACCCATGAC
7381 CGAGCCCCAC CCCGAGGGCG CGGCGTGAT CCTGTGCATC GAGAAGGCC GCCCCAGGC
7441 CGGCGTGAGC AAGGAGGACG TGAACTACAT CAACGCCAC GCCACCAGCA CCAGCGCCG
7501 CGACATCAAG GAGTACCAGG CCTGGGCCG CTGCTTCGGC CAGAACAGCG AGCTGCGCGT
7561 GAACAGCACC AAGAGCATGA TCGGCCACCT GCTGGGCGCC GCCGGCGGCG TGGAGGCCGT
7621 GACCGTGGTG CAGGCCATCC GCACCGGCTG GATTCAACCC AACCTGAACC TGGAGGACCC
7681 CGACAAGGCC GTGGACGCCA AGCTGCTGGT GGGCCCCAAG AAGGAGCGCC TGAACGTGAA
7741 GGTGGGCCTG AGCAACAGCT TCGGCTTCGG CGGCCACAAC AGCAGCATCC TGTTTCGCCCC
7801 CTGCAAAAGT TGAATGCATA CGGAGCGTCC TCGGGAGGG AGTGTGCCGA GCGGGGAGTC
7861 CCGGTCTGTG CGAGGCCCGG CAGCTGACGC TGGCGAGCCG TACGCCCGA GGGTCCCCCT
7921 CCCCTCCCCC CTCTTCCCCC TCCCTCTGAC GGCCGCGCCT GTTTTTCGAT GTTCAGCGAC
7981 CTTAAGGATC TAAGTAAGAT TCGAAGCGCT CGACCGTGCC GGACGGACTG CAGCCCCATG
8041 TCGTAGTGAC CGCCAATGTA AGTGGGCTGG CGTTTCCCTG TACGTGACTC AACGTCAGT
8101 CACGCGCACC ACCCTCTCGA CCGGCAGGAC CAGGCATCGC GAGATACAGC GCGAGCCAGA
8161 CACGGAGTGC CGAGCTATGC GCACGCTCCA ACTAGATATC GAATTCGGCC GACAGGACGC
8221 GCGTCAAAGG TGCTGGTCTG GTATGCCCTG GCCGCGAGGT CGTTGCTGCT GCTGGTTAGT
8281 GATTCGCAA CCCTGATTTT GGCGTCTTAT TTTGGCGTGG CAAACGCTGG CGCCCGCGAG
8341 CCGGGCCGGC GCGGATGCGG TGCCCCACGG CTGCCGGAAT CCAAGGGAGG CAAGAGCGCC
8401 CGGGTCAGTT GAAGGGCTTT ACGCGCAAGG TACAGCCGCT CCTGCAAGGC TCGGTGGTGG
8461 AATGGACGT GCAGGTCCTG CTGAAGTCC TCCACCGCCT CACCAGCGGA CAAAGCACCG
8521 GTGTATCAGG TCCGTGTCAT CCACTCTAAA GAACTCGACT ACGACCTACT GATGGCCCTA
8581 GATTCTTCAT CAAAAACGCC TGAGACACTT GCCCAGGATT GAAACTCCCT GAAGGGACCA
8641 CCAGGGGCCC TGAGTTGTTT CTTCCTCCCG TGGCGAGCTG CCAGCCAGG TGTACCTGTG
8701 ATCGAGGCTG GCGGGAAAAA AGGCTTCGTG TGCTCAGGTC ATGGGAGGTG CAGGACGACT
8761 CATGAAACGC CAACAATCGC ACAATTCTATG TCAAGCTAAT CAGCTATTTT CTCTTCACGA
8821 GCTGTAATTG TCCAAAAATT CTGGTCTACC GGGGGTGATC CTTCTGTGAC GGGCCCTTCC
8881 CTCAACCCTA GGTATGCGCG CATGCGGTG CCGCGCAACT CGCGCGAGGG CCGAGGGTTT
8941 GGGACGGGCC GTCCCGAAAT GCAGTTGCAC CCGGATGCGT GGCACCTTTT TTGCGATAAT
9001 TTATGCAATG GACTGCTCTG CAAAATTCTG GCTCTGTGCG CAACCCTAGG ATCAGCGGCG
9061 TAGGATTTTG TAATCATTCG TCCTGATGGG GAGCTACCGA CTACCCTAAT ATCAGCCCGA
9121 CTGCTTCGAC CCAGCGTCCA CTTTGTGACA CATATTCAT TCGTGCCCAA GACATTTTCA
9181 TGTGGTGCGA AGCGTCCCCA GTTACGCTCA CCTGTTTCCC GACCTCCTTA CTGTTCTGTC
9241 GACAGAGCGG GCCCACAGGC CGGTGCGAGC CACTAGTATG GCCACCGCAT CCACTTTCTC
9301 GGCCTTCAAT GCCCGCTGCG GCGACCTGCG TCGCTCGGCG GGCTCCGGGC CCCGGCGCCC
9361 AGCGAGGCC CTCCCGGTGC GCGGGCGGCG CCCCAGGCC AACGGCAGCG CCGTGAGCCT
9421 GAAGTCCGGC AGCCTGAACA CCCTGGAGGA CCCCCCAGC AGCCCCCCCC CCCGCACCTT
9481 CCTGAACCAG CTGCCCGACT GGAGCCGCTT GCGCACCGCC ATCACCACCG TGTTCTGTGG
9541 CGCCGAGAAG CAGTTCACCC GCCTGGACCG CAAGAGCAA AGCCCCGACA TGCTGGTGGG
9601 CTGGTTCGGC AGCGAGACCA TCGTGCAGGA CGGCCTGGTG TTCCGCGAGC GCTTCAGCAT
9661 CCGCAGCTAC GAGATCGGCG CCGACCGCAC CGCCAGCATC GAGACCCTGA TGAACCACCT
9721 GCAGGACACC AGCCTGAACC ACTGCAAGAG CGTGGGCCTG CTGAACGACG GCTTCGGCCG
9781 CACCCCGAG ATGTGACACC GCGACCTGAT CTGGGTGCTG ACCAAGATGC AGATCGTGGT
9841 GAACCGCTAC CCCACCTGGG GCGACACCGT GGAGATCAAC AGCTGGTTCA GCCAGAGCGG
9901 CAAGATCGGC ATGGGCCGCG AGTGGCTGAT CAGCGACTGC AACACCGGCG AGATCCTGGT
9961 GCGCGCCACC AGCGCCTGGG CCATGATGAA CCAGAAGACC CGCCGCTTCA GCAAGCTGCC
10021 CTGCGAGGTG CGCCAGGAGA TCGCCCCCA CTTCTGTGAC GCCCCCCCCG TGCATGAGGA
10081 CAACGACCGC AAGCTGCACA AGTTCGACGT GAAGACCGGC GACAGCATCT GCAAGGGCCT
10141 GACCCCGGCG TGGAACGACT TCGACGTGAA CCAGCACGTG AGCAACGTGA AGTACATCGG
10201 CTGGATTCTG GAGAGCATGC CCACCGAGGT GCTGGAGACC CAGGAGCTGT GCAGCCTGAC

```

10261 CCTGGAGTAC CGCCGCGAGT GCGGCCGCGA GAGCGTGGTG GAGAGCGTGA CCAGCATGAA
10321 CCCCAGCAAG GTGGGCGACC GCAGCCAGTA CCAGCACCTG CTGCGCCTGG AGGACGGCGC
10381 CGACATCATG AAGGGCCGCA CCGAGTGGCG CCCCAGAAC GCCGGCACCA ACCGCGCCAT
10441 CAGCACCTGA TTAATTAAC TCGAGGAGCA GCAGCTCGGA TAGTATCGAC AACTCTGGA
10501 CGCTGGTCGT GTGATGGACT GTTGCCGCCA CACTTGCTGC CTTGACCTGT GAATATCCCT
10561 GCCGCTTTTA TCAAACAGCC TCAGTGTGTT TGATCTTGTT TGTACGCGCT TTTGCGAGTT
10621 GCTAGCTGCT TGIGCTATTT GCGAATACCA CCCCAGCAT CCCCTTCCCT CGTTTCATAT
10681 CGCTTGCATC CCAACCGCAA CTTATCTACG CTGTCCTGCT ATCCCTCAGC GCTGCTCCTG
10741 CTCTGCTCA CTGCCCTCG CACAGCCTTG GTTTGGGCTC CGCCTGTATT CTCCTGGTAC
10801 TGCAACCTGT AAACCAGCAC TGCAATGCTG ATGCACGGGA AGTAGTGGGA TGGGAACACA
10861 AATGGAAGC TTGAGCTCTT GTTTTCAGA AGGAGTTGCT CCTTGAGCCT TTCATTCTCA
10921 GCCTCGATAA CCTCCAAAGC CGCTCTAATT GTGGAGGGGG TTCGAATTTA AAAGCTTGGA
10981 ATGTTGGTTC GTGCGTCTGG AACAAAGCCA GACTTGTTGC TCACTGGGAA AAGGACCATC
11041 AGCTCCAAAA AACTTGCCGC TCAAACCGCG TACCTCTGCT TTCGCGCAAT CTGCCCTGTT
11101 GAAATCGCCA CCACATTCAT ATTGTGACGC TTGAGCAGTC TGTAATTGCT TCAGAATTG
11161 GAATCATCTG CCCCCTGTGC GAGCCCATGC CAGGCATGTC CGCGGCGAGG ACACCCGCCA
11221 CTCGTACAGC AGACCATTAT GCTACCTCAC AATAGTTCAT AACAGTGACC ATATTTCTCG
11281 AAGCTCCCCA ACGAGCACCT CCATGCTCTG AGTGGCCACC CCCCAGCCCT GGTGCTTGCG
11341 GAGGGCAGGT CAACCGGCAT GGGGCTACCG AAATCCCCGA CCGGATCCCA CCACCCCGC
11401 GATGGGAAGA ATCTCTCCCC GGGATGTGGG CCCACCACCA GCACAACCTG CTGGCCAGG
11461 CGAGCGTCAA ACCATACCAC ACAAATATCC TTGGCATCGG CCCTGAATTC CTCTGCCGC
11521 TCTGCTACCC GGTGCTTCTG TCCGAAGCAG GGGTTGCTAG GGATCGCTCC GAGTCCGCAA
11581 AACCTTGTG CGTGCGGGG CTTGTTTCGAG CTTGAAGAGC CTCTAGAGTC GACCTGCAGG
11641 CATGCAAGCT TGGCGTAATC ATGGTTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC
11701 ACAATTCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA
11761 GTGAGCTAAC TCACATTAAT TCGGTTGCGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG
11821 TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG
11881 CGCTCTTCCG CTTCTCTGCT CACTGACTCG CTGCGCTCGG TCGTTCCGGT GCGGCGAGCG
11941 GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA TAACGAGGA
12001 AAGAACATGT GAGCAAAAGG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG
12061 GCGTTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCGAAGTCAG
12121 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC
12181 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCTT TCTCCCTTCG
12241 GGAAGCGTGG CGCTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTCGTT
12301 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTGAGC CCGACCGCTG CGCCTTATCC
12361 GGTAACATAT GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGC

```

SEQ ID NO: 233

pSZ1500

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAA
181 CCACCGCTGG TAGCGGTGGT TTTTTTGTIT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCTCTCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGTTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG

```

1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCCCGT CAGCGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTTCGTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGGGCTGG TCTGAATCCT TCAGGCGGGT
1921 GTTACCCGAG AAAGAAAGGG TGCCGATTTT AAAGCAGACC CATGTGCCGG CCCCTGTGGC
1981 CTGTGTTGGC GCCTATGTAG TCACCCCGCT TCACCCAATT GTCGCCAGTT TCGGCACTCC
2041 ATAAACTCAA AACAGCAGCT TCTGAGCTGC GCTGTTCAAG AACACCTCTG GGGTTTGCTC
2101 ACCCGCGAGG TCGACGCCCA GCATGGCTAT CAAGACGAAC AGGCAGCCTG TGGAGAAGCC
2161 TCGTTTCACG ATCGGGACGC TGCGCAAGGC CATCCCGCG CACTGTTTCG AGCGCTCGGC
2221 GCTTCGTAGC AGCATGTACC TGGCCTTTGA CATCGCGGTC ATGTCCCTGC TCTACGTCGC
2281 GTCGACGTAC ATCGACCCTG CACCGGTGCG TACGTGGGTC AAGTACGGCA TCATGTGGCC
2341 GCTCTACTGG TTCTTCCAGG TGTGTTGAG GGTTTTGGT GCCCGTATTG AGGTCTGGT
2401 GCGCGCATG GAGGAGAAGG CGCCTGTCCC GCTGACCCCC CCGGCTACCC TCCCGGCACC
2461 TTCCAGGGCG CCTTCGGCAC GGGTGCTGG GTGTGCGCGC ACGAGTGCGG CCACCAGGCC
2521 TTTTCCTCCA GCCAGGCCAT CAACGACGGC GTGGGCCTGG TGTTCCACAG CCTGCTGCTG
2581 GTGCCCTACT ACTCCTGGAA GCACTCGCAC CGGGTACCCT TTCTTGCGCT ATGACACTTC
2641 CAGCAAAAGG TAGGGCGGGC TGCGAGACGG CTTCCCGCG CTGCATGCAA CACCGATGAT
2701 GCTTCGACCC CCCGAAGCTC CTTCCGGGCT GCATGGGCGC TCCGATGCCG CTCCAGGGCG
2761 AGCGCTGTTT AAATAGCCAG GCCCCGATT GCAAAGACAT TATAGCGAGC TACCAAAGCC
2821 ATATTCAAAC ACCTAGATCA CTACCACTTC TACACAGGCC ACTCGAGCTT GTGATCGCAC
2881 TCCGCTAAGG GGGCGCCTCT TCCTCTTCGT TTCAGTCACA ACCCGCAAAC TCTAGAATAT
2941 CAATGCTGCT GCAGGCCCTT CTGTTCTGCT TGGCCGGCTT CGCCGCCAAG ATCAGCGCCT
3001 CCATGACGAA CGAGACGTCC GACCGCCCCC TGGTGCACCT CACCCCCAAC AAGGGCTGGA
3061 TGAACGACCC CAACGGCCTG TGGTACGACG AGAAGGACGC CAAGTGCCAC CTGTACTTCC
3121 AGTACAACCC GAACGACACC GTCTGGGGGA CGCCCTTGTT CTGGGGCCAC GCCACGTCCG
3181 ACGACCTGAC CAACTGGGAG GACCAGCCCA TCGCCATCGC CCCGAAGCGC AACCACTCCG
3241 GCGCCTTCTC CGGCTCCATG GTGGTGGACT ACAACAACAC CTCCGGCTTC TTCAACGACA
3301 CCATCGACCC GCGCCAGCGC TGCGTGCCCA TCTGGACCTA CAACACCCCG GAGTCCGAGG
3361 AGCAGTACAT CTCCTACAGC CTGGACGGCG GCTACACCTT CACCGAGTAC CAGAAGAACC
3421 CCGTGCTGGC CGCCAACTCC ACCCAGTTCC GCGACCCGAA GGTCTTCTGG TACGAGCCCT
3481 CCAGAAAGTG GATCATGACC GCGGCCAAGT CCCAGGACTA CAAGATCGAG ATCTACTCCT
3541 CCGACGACCT GAAGTCCTGG AAGCTGGAGT CCGCGTTTCG CAACGAGGGC TTCCTCGGCT
3601 ACCAGTACGA GTGCCCCGGC CTGATCGAGG TCCCCACCGA GCAGGACCCC AGCAAGTCCT
3661 ACTGGGTGAT GTTCATCTCC ATCAACCCCG GCGCCCGGCG CGGCGGCTCC TCAACCACT
3721 ACTTCGTCGG CAGCTTCAAC GGCACCCACT TCGAGGCCCT CGACAACCAG TCCCGCGTGG
3781 TGGACTTCGG CAAGGACTAC TACGCCCTGC AGACCTTCTT CAACACCGAC CCGACCTACG
3841 GGAGCGCCCT GGGCATCGCG TGGGCTCCA ACTGGGAGTA CTCCGCCTTC GTGCCACCA
3901 ACCCTTGGCG CTCCTCCATG TCCCTCGTGC GCAAGTTCTC CCTCAACACC GAGTACCAGG
3961 CCAACCCGGA GACGGAGCTG ATCAACCTGA AGGCCGAGCC GATCCTGAAC ATCAGCAACG
4021 CCGGCCCTG GAGCCGGTTC GCCACCAACA CCACGTTGAC GAAGGCCAAC AGCTACAACG
4081 TCGACCTGTC CAACAGCACC GGCACCTGG AGTTTCGAGT GGTGTACGCC GTCAACACCA
4141 CCCAGACGAT CTCCAAGTCC GTGTTCGCGG ACCTCTCCCT CTGGTTCAAG GGCCTGGAGG
4201 ACCCCGAGGA GTACCTCCGC ATGGGCTTCG AGGTGTCCGC GTCCTCCTTC TTCCTGGACC
4261 GCGGGAACAG CAAGGTGAAG TTCGTGAAGG AGAACCCTTA CTTACCAAC CGCATGAGCG
4321 TGAACAACCA GCCCTTCAAG AGCGAGAACG ACCTGTCTTA CTACAAGGTG TACGGCTTGC
4381 TGGACCAGAA CATCTGGAG CTGTACTTCA ACGACGGCGA CGTCGTGTCC ACCAACACCT
4441 ACTTCATGAC CACCGGGAAC GCCCTGGGCT CCGTGAACAT GACGACGGGG GTGGACAACC
4501 TGTTCATCAT CGACAAGTTC CAGGTGCGCG AGGTCAAGTG ACAATTGGCA GCAGCAGCTC
4561 GGATAGTATC GACACACTCT GGACGCTGGT CGTGTGATGG ACTGTTCCCG CCACACTTGC
4621 TGCCTTGACC TGTGAATATC CCTGCCGCTT TTATCAAACA GCCTCAGTGT GTTTGATCTT
4681 GTGTGTACGC GCTTTTGCGA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCCGAG
4741 CATCCCTTTC CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTATCT ACGCTGTCTT
4801 GCTATCCCTC AGCGCTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG

4861 CTCGCCCTGT ATTCTCCTGG TACTGCAACC TGTAACCAG CACTGCAATG CTGATGCACG
4921 GGAAGTAGTG GGATGGGAAC ACAAATGGAG GATCCCGCGT CTCGAACAGA GCGCGCAGAG
4981 GAACGCTGAA GGTCTCGCCT CTGTGCGACC TCAGCGCGGC ATACACCACA ATAACCACCT
5041 GACGAATGCG CTTGGTTCTT CGTCCATTAG CGAAGCGTCC GGTTCACACA CGTGCCACGT
5101 TGGCGAGGTG GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG
5161 GGATATCGAA TTCGGCCGAC AGGACGCGCG TCAAAGGTGC TGGTCGTGTA TGCCTGGCC
5221 GGCAGGTCGT TGCTGCTGCT GGTTAGTGAT TCCGCAACCC TGATTTTGGC GTCTTATTTT
5281 GGCGTGGCAA ACGCTGGCGC CCGCGAGCCG GGCCGGCGGC GATGCGGTGC CCCACGGCTG
5341 CCGGAATCCA AGGGAGGCAA GAGCGCCCGG GTCAGTTGAA GGGCTTTACG CGCAAGGTAC
5401 AGCCGCTCCT GCAAGGCTGC GTGGTGGAAT TGGACGTGCA GGTCTGCTG AAGTTCTCTC
5461 ACCGCCTCAC CAGCGGACAA AGCACCAGGT TATCAGGTCC GTGTCATCCA CTCTAAAGAA
5521 CTCGACTACG ACCTACTGAT GGCCCTAGAT TCTTCATCAA AAACGCCTGA GACACTTGCC
5581 CAGGATTGAA ACTCCCTGAA GGGACACCA GGGGCCCTGA GTTGTTCCTT CCCCCCGTGG
5641 CGAGCTGCCA GCCAGGCTGT ACCTGTGATC GAGGCTGGCG GGAAAATAGG CTTGCTGTGC
5701 TCAGGTGATG GGAGGTGACG GACAGTCACT GAAACGCCAA CAATCGCACA ATTCATGTCA
5761 AGCTAATCAG CTATTTCTCT TTCACGAGCT GTAATTGTCC CAAAATTTCTG TCTACCGGG
5821 GGTGATCCTT CGTGTACGGG CCTTTCCTC AACCCTAGGT ATGCGCGCAT GCGGTGCGCG
5881 CGCAACTCGC GCGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGACCCG
5941 GATGCGTGGC ACCTTTTGTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT
6001 CTGTGCGCAA CCCTAGGATC AGCGGCGTAG GATTTCTGTA TCATTGCTCC TGATGGGGAG
6061 CTACCGACTA CCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC
6121 ATTCCATTCTG TGCCCAAGAC ATTTCAATTG GGTGCGAAGC GTCCCCAGTT ACGCTCACCT
6181 GTTTCCCGAC CTCCTACTG TTCTGTGAC AGAGCGGGCC CACAGGCCG TCGCAGCCAC
6241 TAGTATGGCC ACCGCATCCA CTTTCTCGGC GTTCAATGCC CGTGTGCGCG ACCTGCGTGC
6301 CTCGGCGGGC TCCGGCCCCC GCGCGCCAGC GAGGCCCTC CCGTGTGCGG GCGCGCGCCG
6361 CACCGGCGAG CAGCCCTCCG GCGTGGCCTC CTGCGCGAG GCGGACAAGG AGAAGTCCCT
6421 GGGCAACCGC CTGCGCCTGG GCTCCCTGAC CGAGGACGGC CTGTCTTACA AGGAGAAGTT
6481 CGTGATCCGC TGCTACGAGG TGGGCATCAA CAAGACCGCC ACCATCGAGA CCATCGCCAA
6541 CCTGTGTCAG GAGGTGGGCG GCAACCACGC CCAGGGCGTG GGCTTCTCCA CCGACGGCTT
6601 CGCCACCACC ACCACCATGC GCAAGCTGCA CCTGATCTGG GTGACCGCCC GCATGCACAT
6661 CGAGCTCTAC CGCTACCCCG CCTGGTCCGA CGTGATCGAG ATCGAGACCT GGGTGCAGGG
6721 CGAGGGCAAG GTGGGCACCC GCGCGGACTG GATCCTGAAG GACTACGCCA ACGGCGAGGT
6781 GATCGGCGGC GCCACCTCCA AGTGGGTGAT GATGAACGAG GACACCGGCC GCCTGCAGAA
6841 GGTGTCCGAC GACGTGCGCG AGGAGTACCT GGTGTTCTGC CCGCGCACCC TCGCCTGGC
6901 CTTCCTCCGAG GAGAACAAACA ACTCCATGAA GAAGATCCCC AAGCTGGAGG ACCCCGCCGA
6961 GTACTCCCGC CTGGGCCTGG TGCCCCGCCG CTCCGACCTG GACATGAACA AGCAGTGAA
7021 CAACGTGACC TACATCGGCT GGGCCCTGGA GTCCATCCCC CCGAGATCA TCGACACCCA
7081 CGAGCTCTAG GCCATCACCC TGGACTCCG CCGCGAGTGC CAGCGCGACG ACATCGTGA
7141 CTCCCTGACC TCCCGCGAGC CCCTGGGCAA CGCCGCCGGC GTGAAGTTCA AGGAGATCAA
7201 CGGCTCCGTG TCCCCAAGA AGGACGAGCA GGACCTGTCC CGCTTCATGC ACCTGCTGCG
7261 CTCGCGCGGC TCCGGCCTGG AGATCAACCG CTGCGCGACC GAGTGGCGCA AGAAGCCCGC
7321 CAAGCGCATG GACTACAAGG ACCACGACGG CGACTACAAG GACCACGACA TCGACTACAA
7381 GGACGACGAC GACAAGTGAA TCGATAGATC TCTTAAGGCA GCAGCAGCTC GGATAGTATC
7441 GACACACTCT GGACGCTGGT CGTGTGATGG ACTGTTGCCG CCACACTTGC TGCCTTGACC
7501 TGTGAATATC CCTGCCGCTT TTATCAAACA GCCTCAGTGT GTTTGATCTT GTGTGTACGC
7561 GCTTTTGCGA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCCGAG CATCCCCTTC
7621 CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTATCT ACGCTGTCTT GCTATCCCTC
7681 AGCGTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG CTCCGCCTGT
7741 ATTCTCCTGG TACTGCAACC TGTAACCAG CACTGCAATG CTGATGCACG GGAAGTAGTG
7801 GGATGGGAAC ACAAATGGAA AGCTTAATTA AGAGCTCCCG CCACCACTCC AACACGGGGT
7861 GCCTGGACAA GGACGAGGTG TTTGTGCCGC CGCACCGCGC AGTGGCGCAC GAGGGCCTGG
7921 AGTGGGAGGA GTGGCTGCCC ATCCGCATGG GCAAGGTGCT GGTACCCCTG ACCCTGGGCT
7981 GGCCGCTGTA CCTCATGTTT AACGTGCGCT CGCGGCCGTA CCGCGCTTC GCCAACCACT
8041 TTGACCCGTG GTCGCCCATC TTCAGCAAGC GCGAGCGCAT CGAGGTGGTC ATCTCCGACC
8101 TGGCGCTGGT GCGGTGCTC AGCGGGCTCA GCGTGTGGG CCGCACCATG GGCTGGGCTT
8161 GGCTGGTCAA GACCTACGTG GTGCCCTACC TGATCGTGAA CATGTGGCTC GTGCTCATCA
8221 CGCTGCTCCA GCACACGCAC CCGGCGCTGC CGCACTACTT CGAGAAGGAC TGGGACTGGC
8281 TGCGGCGCGC CATGGCCACC GTGGACCGCT CCATGGGCCC GCCCTTCATG GACAACATCC
8341 TGCACCACAT CTCCGACACC CACGTGCTGC ACCACCTCTT CAGCACCATC CCGCACTACC
8401 ACGCGGAGGA GGCCTCCGCC GCAATCTGGG CCATCTGCGC CAAGTACTAC CATCCGACA
8461 GCCGCTGGGT CCGCCGCGCC CTGTGGGAGG ACTGGCGCGA CTGCCGCTAC GTCGTCCCGG

```

8521 ACGCGCCCGA GGACGACTCC GCGCTCTGGT TCCACAAGTG AGTGAGTGAG AAGAGCCTCT
8581 AGAGTCGACC TGCAGGCATG CAAGCTTGCG GTAATCATGG TCATAGCTGT TTCCTGTGTG
8641 AAATTGTTAT CCGCTCACAA TTCCACACAA CATAAGAGCC GGAAGCATAA AGTGTAAGC
8701 CTGGGGTGCC TAATGAGTGA GCTAACTCAC ATTAATTGCG TTGCGCTCAC TGCCCGCTTT
8761 CCAGTCGGGA AACCTGTCGT GCCAGCTGCA TTAATGAATC GGCCAACGCG CGGGGAGAGG
8821 CGGTTTTCGT ATTGGGCGCT CTTCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCTG
8881 TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC
8941 AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA
9001 AAAGGCCGCG TTGCTGGCGT TTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA
9061 TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC
9121 CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC
9181 CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA GGTATCTCAG
9241 TTCGGTGTAG GTCGTTGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCCG
9301 CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC
9361 GC

```

SEQ ID NO: 234
pSZ1501_FADc-inv

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGCCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGAAAAAA GAGTTGGTAG CTCCTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTTGTITT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTTTAA TCAATCTAAA GTATATATGA GTAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCGTCGTG TAGATACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAAC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCCGC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGGTTATGCG AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA ATCCCTTCTG
961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTGCCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATGTGCTCAT GAGCGGATAC ATATTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCAGGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGGGCTGG TCTGAATCCT TCAGGCGGGT
1921 GTTACCCGAG AAAGAAAGGG TGCCGATTTC AAAGCAGACC CATGTGCCGG GCCCTGTGGC
1981 CTGTGTTGGC GCCTATGTAG TCACCCCCC TCACCAATT GTCGCCAGTT TGCGCACTCC
2041 ATAAACTCAA AACAGCAGCT TCTGAGCTGC GCTGTTCAAG AACACCTCTG GGGTTTGCTC
2101 ACCCGCGAGG TCGACGCCCA GCATGGCTAT CAAGACGAAC AGGCAGCCTG TGGAGAAGCC
2161 TCCGTTACAG ATCGGGACGC TGCGCAAGGC CATCCCCGCG CACTGTTTCG AGCGCTCGGC
2221 GCTTCGTAGC AGCATGTACC TGGCCTTTGA CATCGCGGTC ATGTCCCTGC TCTACGTGCG
2281 GTCGACGTAC ATCGACCCGT CACCGGTGCC TACGTGGGTC AAGTACGGCA TCATGTGGCC
2341 GCTCTACTGG TTCTTCCAGG TGTGTTTGAG GGTTTTGGTT GCCCGTATTG AGGTCTTGGT
2401 GCGCGCATG GAGGAGAAGG CGCCTGTCCC GCTGACCCCC CCGGCTACCC TCCCGGCACC

```

2461 TTCCAGGGCG CTTTCGGCAC GGGTGTCTGG GTGTGCGCGC ACGAGTGC GG CCACCAGGGC
2521 TTTTCTCTCA GCCAGGCCAT CAACGACGGC GTGGGCTGG TGTTCACAG CCTGCTGCTG
2581 GTGCCCTACT ACTCCTGGAA GCACTCGCAC CGGGTACCCT TTCTTGCGCT ATGACACTTC
2641 CAGCAAAAGG TAGGGCGGGC TGGGAGACGG CTTCCCGGGC CTGCATGCAA CACCGATGAT
2701 GCTTCGACCC CCCGAAGCTC CTTGGGGGCT GCATGGGCGC TCCGATGCG CTCCAGGGCG
2761 AGCGCTGTTT AAATAGCCAG GCCCCGATT GCAAAGACAT TATAGCGAGC TACCAAAGCC
2821 ATATTCAAAC ACCTAGATCA CTACCACTTC TACACAGGCC ACTCGAGCTT GTGATCGCAC
2881 TCCGCTAAGG GGGCGCCTCT TCCTCTTCGT TTCAGTACA ACCCGCAAAC TCTAGAATAT
2941 CAATGCTGCT GCAGGCCTTC CTGTTCTGTC TGGCCGGCTT CGCCGCCAAG ATCAGCGCCT
3001 CCATGACGAA CGAGACGTCC GACCGCCCCC TGGTGCACCT CACCCCCAAC AAGGGCTGGA
3061 TGAACGACCC CAACGGCCTG TGGTACGACG AGAAGGACGC CAAGTGGCAC CTGTACTTCC
3121 AGTACAACCC GAACGACACC GTCTGGGGGA CGCCCTTGT CTGGGGCCAC GCCACGTCCG
3181 ACGACCTGAC CAACTGGGAG GACCAGCCCA TCGCCATCGC CCCGAAGCGC AACGACTCCG
3241 GCGCCTTCTC CGGCTCCATG GTGGTGGACT ACAACAACAC CTCCGGCTTC TTCAACGACA
3301 CCATCGACCC GCGCCAGCGC TCGTGGCCA TCTGGACCTA CAACACCCTG GAGTCCGAGG
3361 AGCAGTACAT CTCCTACAGC CTGGACGGCG GCTACACCTT CACCGAGTAC CAGCAAGAAC
3421 CCGTGCTGGC CGCCAACTCC ACCCAGTTCC GCGACCCGAA GGTCTTCTGG TACGAGCCCT
3481 CCCAGAAGTG GATCATGACC GCGGCCAAGT CCCAGGACTA CAAGATCGAG ATCTACTCCT
3541 CCGACGACCT GAAGTCTTGG AAGCTGGAGT CCGCGTTCGC CAACGAGGGC TTCCTCGGCT
3601 ACCAGTACGA GTGCCCCGGC CTGATCGAGG TCCCCACCGA GCAGGACCCC AGCAAGTCCT
3661 ACTGGGTGAT GTTCATCTCC ATCAACCCCG GCGCCCCGGC CGGCGGCTCC TTCAACCAGT
3721 ACTTCGTCTG CAGCTTCAAC GGCACCCACT TCGAGGCCTT CGACAACCAG TCCCGCGTGG
3781 TGGACTTCGG CAAGGACTAC TACGCCCTGC AGACCTTCTT CAACACCAGC CCGACCTACG
3841 GGAGCGCCCT GGGCATCGCG TGGGCCCTCA ACTGGGAGTA CTCCGCCTTC GTGCCACCA
3901 ACCCCTGGCG CTCCTCCATG TCCCTCGTGC GCAAGTTCTC CCTCAACACC GAGTACCAGG
3961 CCAACCCGGA GACGGAGCTG ATCAACCTGA AGGCCGAGCC GATCCTGAAC ATCAGCAACG
4021 CCGGCCCTTG GAGCCGGTTC GCCACCAACA CCACGTTGAC GAAGGCCAAC AGCTACAACG
4081 TCGACCTGTC CAACAGCACC GGCACCCCTG AGTTCGAGCT GGTGTACGCC GTCAACACCA
4141 CCCAGACGAT CTCCAAGTCC GTGTTGCGCG ACCTCTCCCT CTGGTTCAAG GGCCTGGAGG
4201 ACCCGAGGA GTACCTCCGC ATGGGCTTCG AGGTGTCCGC GTCTCTCTT TTCCTGGACC
4261 GCGGGAACAG CAAGGTGAAG TTCGTGAAGG AGAACCCTA CTTACCAAAC CCACATGAGC
4321 TGAACAACCA GCCCTTCAAG AGCGAGAACG ACCTGTCTTA CTACAAGGTG TACGGCTTGC
4381 TGGACCAGAA CATCTGGAG CTGTACTTCA ACGACGGCGA CGTCGTGTCC ACCAACACCT
4441 ACTTCATGAC CACCGGGAAC GCCCTGGGCT CCGTGAACAT GACGACGGGG GTGGACAACC
4501 GTTCTACAT CGACAAGTTC CAGGTGCGCG AGGTCAAGTG ACAATTGGCA GCAGCAGCTC
4561 GGATAGTATC GACACACTCT GGACGCTGGT CGTGTGATGG ACTGTTGCCG CCACACTTGC
4621 TGCCTTGACC TGTGAATATC CCTGCGCTT TTATCAAACA GCCTCAGTGT GTTTGATCTT
4681 GTGTGTACGC GCTTTTGCGA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCGAG
4741 CATCCCTTTC CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACCTATCT ACCTGTCTCT
4801 GCTATCCCTC AGCGCTGCTC CTGCTCTGTC TCACTGCCCC TCGCACAGCC TTGGTTTGGG
4861 CTCCGCTGT ATTCTCTGG TACTGCAACC TGTAACACAG CACTGCAATG CTGATGCACG
4921 GGAAGTAGTG GGATGGGAAC ACAATGGAG GATCCCGCGT CTCGAACAGA GCGCGCAGAG
4981 GAACGCTGAA GTTCTCGCCT CTGTCGCACC TCAGCGCGGC ATACACCACA ATAACCACCT
5041 GACGAATGCG CTTGGTTCTT CGTCCATTAG CGAAGCGTCC GGTTCACACA CGTGCCACGT
5101 TGGCGAGGTG GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG
5161 GGATATCGAA TTCCTTTCTT GCGCTATGAC ACTTCCAGCA AAAGGTAGGG CGGGCTGCGA
5221 GACGGCTTCC CGGCGCTGCA TGCAACACCG ATGATGCTTC GACCCCCCGA AGCTCCTTCG
5281 GGGCTGCATG GGGCTCCGA TGCCGCTCCA GGGCGAGCGC TGTTTAAATA GCCAGGCCCC
5341 CGATTGCAAA GACATTATAG CGAGCTACCA AAGCCATATT CAAACACCTA GATCACTACC
5401 ACTTCTACAC AGGCCACTCG AGCTTGATGAT CGCACTCCGC TAAGGGGGCG CCTCTTCTCT
5461 TTCGTTTCAG TCACAACCCG CAAACACTAG TATGGCCACC GCATCCACTT TCTCGGCGTT
5521 CAATGCCCGC TGCGGCGACC TCGTTCGCTC GGGCGGCTCC GGGCCCCGGC GCCAGCGAG
5581 GCCCTTCCCC GTGCGCGGGC GCGCCGCCAC CGGCGAGCAG CCCTCCGGCG TGGCCTCCCT
5641 GCGCGAGGCC GACAAGGAGA AGTCCCTGGG CAACCGCCTG CGCCTGGGCT CCCTGACCGA
5701 GGACGGCCTG TCCTACAAGG AGAAGTTCGT GATCCGCTGC TACGAGGTGG GCATCAACAA
5761 GACCGCCACC ATCAGACCA TCGCCAACCT GCTGCAGGAG GTGGGCGGCA ACCACGCCCA
5821 GGGCGTGGG TTCTCCACCG ACGGCTTCGC CACCACCACC ACCATGCGCA AGCTGCACCT
5881 GATCTGGGTG ACCGCCCGCA TGCACATCGA GATCTACCGC TACCCCGCCT GGTCCGACGT
5941 GATCGAGATC GAGACCTGGG TGCAGGGCGA GGGCAAGGTG GGCACCCGCC GCGACTGGAT
6001 CCTGAAGGAC TACGCCAACG CCGAGGTGAT CGGCCCGGCC ACCTCCAAGT GGGGTGATGAT
6061 GAACGAGGAC ACCCGCCGCC TGCAGAAGGT GTCCGACGAC GTGCGCGAGG AGTACCTGGT

```

6121 GTTCTGCCCC CGCACCCCTGC GCCTGGCCTT CCCCAGAGGAG AACAACAACCT CCATGAAGAA
6181 GATCCCCAAG CTGGAGGACC CCGCCGAGTA CTCCCGCCTG GGCCTGGTGC CCCGCCGCTC
6241 CGACCTGGAC ATGAACAAGC ACGTGAACAA CGTGACCTAC ATCGGCTGGG CCCTGGAGTC
6301 CATCCCCCCC GAGATCATCG ACACCACGA GCTGCAGGCC ATCACCCCTGG ACTACCGCCG
6361 CGAGTGCCAG CGCGACGACA TCGTGGACTC CCTGACCTCC CGCGAGCCCC TGGGCAACGC
6421 CGCCGGCGTG AAGTTCAAGG AGATCAACGG CTCCGTGTCC CCCAAGAAGG ACGAGCAGGA
6481 CCTGTCCCGC TTCATGCACC TGCTGCGCTC CGCCGGCTCC GGCCTGGAGA TCAACCGCTG
6541 CCGCACCGAG TGGCGCAAGA AGCCCCCAA GCGCATGGAC TACAAGGACC ACGACGGCGA
6601 CTACAAGGAC CACGACATCG ACTACAAGGA CGACGACGAC AAGTGAATCG ATAGATCTCT
6661 TAAGGCAGCA GCAGCTCGGA TAGTATCGAC AACTCTGGA CGCTGGTCGT GTGATGGACT
6721 GTTCCCGCCA CACTTGCTGC CTTGACCTGT GAATATCCCT GCCGCTTTTA TCAAACAGCC
6781 TCAGTGTGTT TGATCTTGTT TGTACGCGCT TTTGCGAGTT GCTAGCTGCT TGTGCTATTT
6841 GCGAATACCA CCCCAGCAT CCCCTTCCCT CGTTTCATAT CGCTTGCTAT CCAACCGCAA
6901 CTTATCTACG CTGTCTGCT CTCCCTCAGC GCTGCTCCTG CTCCTGCTCA CTGCCCTCG
6961 CACAGCCTTG GTTTGGGCTC CGCCTGTATT CTCCTGGTAC TGCAACCTGT AAACCGCAC
7021 TGCAATGCTG ATGCACGGGA AGTAGTGGGA TGGGAACACA AATGAAAGC TTAATTAAGA
7081 GCTCCCGCCA CCACTCCAAC ACGGGGTGCC TGGACAAGGA CGAGGTGTTT GTGCCGCCGC
7141 ACCGCGCAGT GCGCGACGAG GGCCTGGAGT GGGAGGAGTG GCTGCCCATC CGCATGGGCA
7201 AGGTGCTGGT CACCTGACC CTGGGCTGGC CGCTGTACCT CATGTTCAAC GTCGCTCGC
7261 GGCCGTACCC GCGCTTCGCC AACCACCTTG ACCCGTGGTC GCCCATCTTC AGCAAGCGCG
7321 AGCGCATCGA GGTGGTCATC TCCGACCTGG CGCTGGTGGC GGTGCTCAGC GGGCTCAGCG
7381 TGCTGGGCGC CACCATGGGC TGGGCTGGC TGGTCAAGAC CTACGTGGTG CCTACCTGA
7441 TCGTGAACAT GTGGCTCGTG CTCATCACGC TGCTCCAGCA CACGCACCCG GCGCTGCCGC
7501 ACTACTTCGA GAAGGACTGG GACTGGCTGC GCGGCGCCAT GGCCACCGTG GACCGCTCCA
7561 TGGGCCCGCC CTTATGGAC AACATCCTGC ACCACATCTC CGACACCCAC GTGCTGCACC
7621 ACCTCTTCAG CACCATCCCG CACTACCACG CCGAGGAGGC CTCCGCCGCC ATCAGGCCCA
7681 TCCTGGGCAA GTACTACCA TCCGACAGCC GCTGGGTCCG CCGCGCCCTG TGGGAGGACT
7741 GGCGCGACTG CCGCTACGTC GTCCCGAGC CGCCGAGGA CGACTCCGCG CTCTGGTTCC
7801 ACAAGTGAGT GAGTGAGAAG AGCCTTAGA GTCGACCTGC AGGCATGCAA GCTTGCGTGA
7861 ATCATGGTCA TAGCTGTTTC CTGTGTGAAA TTGTTATCCG CTCACAATTC CACACAACAT
7921 ACGAGCCGGA AGCATAAAGT GTAAAGCCTG GGGTGCCTAA TGAGTGAGCT AACTCACATT
7981 AATTGCGTTG CGCTCACTGC CCGCTTTCCA GTCGGGAAAC CTGTCGTGCC AGCTGCATTA
8041 ATGAATCGGC CAACGCGCGG GGAGAGCGCG TTTGCGTATT GGGCGCTCTT CCGCTTCCTC
8101 GCTCACTGAC TCGCTGCGCT CGGTCGTTCC GCTGCGGCGA GCGGTATCAG CTCACTCAAA
8161 GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA TGTGAGCAAA
8221 AGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG CTGGCGTTTT TCCATGGCT
8281 CCGCCCCCTT GACGAGCATC ACAAAAATCG ACGCTCAAGT CAGAGGTGGC GAAACCCGAC
8341 AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC
8401 GACCTGCCG CTTACGGAT ACCTGTCCGC CTTTCTCCCT TCGGGAAGCG TGGCGCTTTC
8461 TCATAGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA AGCTGGGCTG
8521 TGTGCACGAA CCCCCGTTT AGCCGACCG CTGCGCCTTA TCCGGTAAC ATCGTCTTGA
8581 GTCCAACCCG GTAAGACACG ACTTATCGC

```

SEQ ID NO: 235
pSZ1570

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC
301 CAGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCA GACCCAGCT CACCGGCTCC AGATTTATCA GCAATAAAC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG

```

841 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCCA CCGAGTTGCT
1021 CTGCCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAG ACGTCTCTCG GGGCGAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCTTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCAATTAT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CCGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGGGG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCAATGCTC TTCGCCGCCG CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGCCCTT TTCGCCCGCG TCGTGCCTGT CGCTGATGTC
1981 CATCACCAGG TCCATGAGGT CTGCCTTGCG CCGGCTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCTTG ACGTGGTCTG GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAACCTGT CCTCCAGCAG
2161 CCGCAGTCCG CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGAAATC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCTCTGCT GCCGCCGCTT CTCCCGCACG CTTCTTTCCA
2341 GCACCGTGAT GCGCGAGGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTCGCGGTT AGTGTGCGCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACCTTT TTGGCGAAGG CATCGGCCTC GGCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GGCGGATGCA GCCTCAGGTA CCCTTTCTTG CGTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCGAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCTCTT TCGTTTCAGT CACAACCCGC AAACGGCGCG CCATGCTGCT
2941 GCAGGCCTTC CTGTTCTCTG TGGCCGGCTT CGCCGCCAAG ATCAGCGCCT CCATGACGAA
3001 CGAGACGTCC GACCGCCCCC TGGTGCATCT CACCCCAAC AAGGGCTGGA TGAACGACCC
3061 CAACGGCCTG TGGTACGAG AGAAGGACGC CAAGTGGCAC CTGTACTTCC AGTACAACCC
3121 GAACGACACC GTCTGGGGGA CGCCCTTGTT CTGGGGCCAC GCCACGTCCG ACGACCTGAC
3181 CAACTGGGAG GACCAGCCCA TCGCCATCGC CCCGAAGCGC AACGACTCCG GCGCCTTCTC
3241 CGGCTCCATG GTGGTGGACT ACAACAACAC CTCCGGCTTC TTCAACGACA CCATCGACCC
3301 GCGCCAGCGC TGGTGGCCA TCTGGACCTA CAACACCCCG GAGTCCGAGG AGCAGTACAT
3361 CTCTACAGC CTGGACGGCG GCTACACCTT CACCGAGTAC CAGAAGAACC CCGTGTGGC
3421 CGCCAACCTC ACCCAGTTCC GCGACCCGAA GGTCTTCTGG TACGAGCCCT CCCAGAAGTG
3481 GATCATGACC GCGGCCAAGT CCCAGGACTA CAAGATCGAG ATCTACTCCT CCGACGACCT
3541 GAAGTCTTGG AAGCTGGAGT CCGCGTTCGC CAACGAGGGC TTCCTCGGCT ACCAGTACGA
3601 GTGCCCCGGC CTGATCGAGG TCCCCACCGA GCAGGACCCC AGCAAGTCCT ACTGGGTGAT
3661 GTTCATCTCC ATCAACCCCG GCGCCCCGGC CGGCGGCTCC TTCAACAGT ACTTCGTGCG
3721 CAGCTTCAAC GGCACCCACT TCGAGGCCTT CGACAACCAG TCCCGCGTGG TGGACTTCGG
3781 CAAGGACTAC TACGCCCTGC AGACCTTCTT CAACACCGAC CCGACCTACG GGAGCGCCCT
3841 GGGCATCGCG TGGGCCTCCA ACTGGGAGTA CTCCGCCTTC GTGCCACCA ACCCTTGGCG
3901 CTCTCCATG TCCCTCGTGC GCAAGTTCTC CCTCAACACC GAGTACCAGG CCAACCCGGA
3961 GACGGAGCTG ATCAACCTGA AGGCCGAGCC GATCCTGAAC ATCAGCAACG CCGGCCCTG
4021 GAGCCGGTTC GCCACCAACA CCACGTTGAC GAAGGCCAAC AGCTACAACG TCGACCTGTC
4081 CAACAGCACC GGCACCTTGG AGTTCGAGCT GGTGTACGCC GTCAACACCA CCCAGACGAT
4141 CTCCAAGTCC GTGTTCGCGG ACCTCTCCCT CTGGTTCAAG GGCCTGGAGG ACCCCGAGGA
4201 GTACCTCCGC ATGGGCTTCG AGGTGTCCGC GTCCTCCTTC TTCCTGGACC GCGGGAACAG
4261 CAAGGTGAAG TTCGTGAAGG AGAACCCCTA CTTACCAAC CGCATGAGCG TGAACAACCA
4321 GCCCTTCAAG AGCGAGAACG ACCTGTCTTA CTACAAGGTG TACGGCTTGC TGGACCAGAA
4381 CATCTGGAG CTGTACTTCA ACGACGCGCA CGTCGTGTCC ACCAACACCT ATTCATGAC
4441 CACCGGGAAC GCCCTGGGCT CCGTGAACAT GACGACGGGG GTGGACAACC TGTTCTACAT

4501 CGACAAGTTC CAGGTGCGCG AGGTCAAGTG ACAATTGGCA GCAGCAGCTC GGATAGTATC
4561 GACACACTCT GGACGCTGGT CGTGTGATGG ACTGTTGCCG CCACACTTGC TGCCCTGACC
4621 TGTGAATATC CCTGCCGCTT TTATCAAACA GCCTCAGTGT GTTTGATCTT GTGTGTACGC
4681 GCTTTTGCGA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCCCAG CATCCCCTTC
4741 CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTAICT ACGCTGTCTT GCTATCCCTC
4801 AGCGCTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG CTCCGCCTGT
4861 ATTCTCTTGG TACTGCAACC TGTAACCAG CACTGCAATG CTGATGCACG GGAAGTAGTG
4921 GGATGGGAAC ACAAATGGAG GATCCCGCGT CTCGAACAGA GCGCGCAGAG GAACGCTGAA
4981 GGTCCTGCCT CTGTGCGACC TCAGCGCGGC ATACACCACA ATAACCACCT GACGAATGCG
5041 CTTGGTTCTT CGTCCATTAG CGAAGCGTCC GGTTACACAC CGTGCCACGT TGGCGAGGTG
5101 GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG GGATATCGAA
5161 TTCGCCCGAC AGGACGCGCG TCAAAGGTGC TGGTCGTGTA TGCCCTGGCC GGCAGGTGCT
5221 TGCTGCTGCT GGTTAGTGAT TCCGCAACCC TGATTTTGGC GTCTTATTTT GGCCTGGCAA
5281 ACGCTGGCGC CCGCGAGCCG GGCCGCGCGC GATGCGGTGC CCCACGGCTG CCGGAATCCA
5341 AGGGAGGCAA GAGCGCCCGG GTCAGTTGAA GGGCTTTACG CGCAAGGTAC AGCCGCTCCT
5401 GCAAGGCTGC GTGGTGGAAT TGGACGTGCA GTCTCTGCTG AAGTTCTCTC ACCGCTCAC
5461 CAGCGGACAA AGCACCGGTG TATCAGGTCC GTGTCAATCA CTCTAAAGAG CTGAGCTACG
5521 ACCTACTGAT GGCCCTAGAT TCTTCATCAA AAACGCCTGA GACACTTGCC CAGGATTGAA
5581 ACTCCCTGAA GGGACCACCA GGGGCCCTGA GTTGTTCCTT CCCCCCGTGG CGAGCTGCCA
5641 GCCAGGCTGT ACCTGTGATC GAGGCTGGCG GGAAATAGG CTTCTGTGTC TCAGGTCATG
5701 GGAGGTGCAG GACAGCTCAT GAAACGCCAA CAATCGCACA ATTCATGTCA AGCTAATCAG
5761 CTATTTCTCT TTACAGAGCT GTAATTGTCC CAAAATTCGT GTCTACCGGG GGTGATCCTT
5821 CGTGACGGG CCTTCCCTC AACCCTAGGT ATGCGCGCAT GCGGTGCGCC CGCAACTCGC
5881 CGGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGCACCCG GATGCGTGGC
5941 ACCTTTTTTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT CTGTGCGCAA
6001 CCCTAGGATC AGCGGCGTAG GATTTCTGTA TCATTCTGTC TGATGGGGAG CTACCGACTA
6061 CCCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC ATTCCATTCT
6121 TGCCCAAGAC ATTTCAATTG GTTGCGAAGC GTCCCCAGTT ACGCTCACCT GTTTCCCGAC
6181 CTCCTTACTG TTCTGTGAC AGAGCGGGCC CACAGGCCCG TCGCAGCCAC TAGTATGGCT
6241 TCCGCGGCAT TCACCATGTC GCGGTGCCCC GCGATGACTG GCAGGGGCCC TGGGGACGCT
6301 CGCTGCGGAC GGCCAGTCGC CACCCGCTG AGGGGGCGCG CCCCCAAGG CAACGCTCTC
6361 GCCGTGTCCC TGAAGTCCGG CTCCCTGGAC ACCCAGGAGG ACACCTCCTC CTCTCCTCTC
6421 CCCCCCGCA CCTTCATCAA CCAGCTGCCC GACTGGTCCA TGCTGCTGTC CGCCATCACC
6481 ACCGTGTTCT TGGCCGCGGA GAAGCAGTGG ACCATGCTGG ACCGCAAGTC CAAGCGCCCC
6541 GACACCCTGA TGGACCCTT CGGCGTGGAC CGCGTGGTGC AGGACGGCGT GGTGTTCCGC
6601 CAGTCCCTCT CCATCCGCTC CTACGAGATC GGCGCCGACC GCACCGCCTC CATCGAGACC
6661 CTGATGAACA TCTTCCAGGA GACCTCCCTG AACCCTGCA AGTCCATCGG CCTGTGAAC
6721 GACGCGCTCG GCCGACCCC CAGAGATGTC AAGCGCGACC TGATCTGGGT GGTGACCAAG
6781 ATGCACATCG AGGTGAACCG CTACCCACC TGGGGCGACA CCATCGAGGT GAACACCTGG
6841 GTGTCGAGT CCGGCAAGAC CGGCATGGGC CGCGACTGGC TGATCTCCGA CTGCCACACC
6901 GGCAGATCC TGATCCGCGC CACCTCCGTG TCGGCCATGA TGAACCAGAC CACCCGCGGC
6961 TTCTCCAAGT TCCCTACGA GGTGCGCCAG GAGCTGGCCC CCCACTTCTG GACTCCGCC
7021 CCCGTGATCG AGGACTACCA GAAGCTGCAC AAGCTGGACG TGAAGACCGG CGACTCCATC
7081 TGCAACGGCC TGACCCCCCG CTGGAACGAC CTGGACGTGA ACCAGCACGT GAACAACGTG
7141 AAGTACATCG GCTGGATTCT GGAGTCCGTG CCAAGGAGG TGTTCGAGAC CCAGGAGCTG
7201 TGCGGCCTGA CCCTGGAGTA CCGCCGCGAG TCGGGCCGCG ACTCCGTGCT GAAGTCCGTG
7261 ACCGCCATGG ACCCTCCAA GGAGGGCGAC CGCTCCCTGT ACCAGCACCT GCTGCGCCTG
7321 GAGGACGGCA CCGACATCGC CAAGGGCCGC ACCAAGTGGC GCCCCAAGAA CGCCGGCACC
7381 AACGGCGCCA TCTCCACCGG CAAGACCTCC AACGGCAACT CCATCTCCTG ATTAATTAAC
7441 TCGAGGCAGC AGCAGCTCGG ATAGTATCGA CACACTCTGG ACGCTGGTCT TGTGATGGAC
7501 TGTGCGCGCC AACTTGCTG CCTTGACCTG TGAATATCCC TGCCGCTTTT ATCAAACAGC
7561 CTCAGTGTGT TTGATCTGT GTGTACGCGC TTTTGCAGT TGCTAGCTGC TTGTGCTATT
7621 TCGGAATACC ACCCCAGCA TCCCTTCCC TCGTTTCATA TCGCTTGCAT CCAACCGCA
7681 ACTTATCTAC GCTGTCCTGC TATCCCTCAG CGCTGCTCCT GCTCCTGCTC ACTGCCCTC
7741 GCACAGCCTT GGTTTGGGCT CCGCCTGTAT TCTCCTGGTA CTGCAACCTG TAAACCAGCA
7801 CTGCAATGCT GATGCACGGG AAGTAGTGGG ATGGGAACAC AAATGGAAAG CTGAGCTCT
7861 TGTTTCCAG AAGGAGTTGC TCCTTGAGCC TTTCATTCTC AGCCTCGATA ACCTCCAAAG
7921 CCGCTCTAAT TGTGGAGGGG GTTCGAATT AAAAGCTTGG AATGTTGGTT CGTGCGTCTG
7981 GAACAAGCCC AGACTTGTG CTCACTGGGA AAAGGACCAT CAGCTCCAAA AACTTGCCG
8041 CTCAAACCGC GTACCTCTG TTTGCGCAA TCTGCCCTGT TGAAATCGCC ACCACATTCA
8101 TATTGTGACG CTTGAGCAGT CTGTAATTGC CTCAGAATGT GGAATCATCT GCCCCCTGTG

```

8161 CGAGCCCATG CCAGGCATGT CGCGGGCGAG GACACCCGCC ACTCGTACAG CAGACCATTA
8221 TGCTACCTCA CAATAGTTCA TAACAGTGAC CATATTTCTC GAAGCTCCCC AACGAGCACC
8281 TCCATGCTCT GAGTGGCCAC CCCCCGCCCC TGGTGCTTGC GGAGGGCAGG TCAACCGGCA
8341 TGGGGCTACC GAAATCCCCG ACCGGATCCC ACCACCCCGG CGATGGGAAG AATCTCTCCC
8401 CGGGATGTGG GCCCACCACC AGCACAACTT GCTGGCCAG GCGAGCGTCA AACCATACCA
8461 CACAAATATC CTTGGCATCG GCCCTGAATT CCTTCTGCCG CTCTGCTACC CGGTGCTTCT
8521 GTCCGAAGCA GGGGTTGCTA GGGATCGCTC CGAGTCCGCA AACCCTTGTC GCGTGGCGGG
8581 GCTTGTTTGA GCTTGAAGAG CCTCTAGAGT CGACCTGCAG GCATGCAAGC TTGGCGTAAT
8641 CATGGTCATA GCTGTTTCCT GTGTGAAATT GTTATCCGCT CACAATTCCA CACAACATAC
8701 GAGCCGGAAG CATAAAGTGT AAAGCCTGGG GTGCCTAATG AGTGAGCTAA CTCACATTAA
8761 TTGCGTTGCG CTCACTGCCC GCTTTCAGT CGGGAAACCT GTCGTGCCAG CTGCATTAAT
8821 GAATCGGCCA ACGCGCGGGG AGAGGCGGTT TGCGTATTGG GCGCTCTTCC GCTTCTCGC
8881 TCACTGACTC GCTGCGCTCG GTGCTTCGGC TGCGGCGAGC GGTATCAGCT CACTCAAAGG
8941 CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG TGAGCAAAAG
9001 GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GGCCTTTTTT CATAGGCTCC
9061 GCCCCCTGA CGAGCATCAC AAAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCAGACAG
9121 GACTATAAAG ATACCAGGCG TTTCCCCCTG GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA
9181 CCCTGCCGCT TACCGGATAC CTGTCCGCCT TTCTCCCTTC GGAAGCGTG GCGCTTCTC
9241 ATAGCTCACG CTGTAGGTAT CTCAGTTGGG TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG
9301 TGCACGAACC CCCCCTTCAG CCCGACCCTG GCGCCTTATC CGGTAACATAT CGTCTTGAGT
9361 CCAACCCGGT AAGACACGAC TTATCGC

```

SEQ ID NO: 236

pSZ941

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAA
181 CCACCGCTGG TAGCGGTGGT TTTTTTGTTC GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAAC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTACTACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGCGCA CCGAGTTGCT
1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAA TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCCAAC TATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCCTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT CAGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCCGG CCACTCTGCT TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCCGC TCGTGCGCGT CGGTGATGTC
1981 CATCACCAGG TCCATGAGGT CTGCCTTGCG CCGGCTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCTTG ACGTGGTCTG GGCTCTGGGA

```

2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAAC TGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGAAATCC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCCTCGCT GCCGCCGCTT CTCCCGCACG CTTCTTTCCA
2341 GCACCGTGAT GGC GCGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTCGCGCTT AGTGTGCCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTACCTTTT TTGGCGAAGG CATCGGCCTC GGCCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGTTT GCGGATGCA CGCTCAGGTA CCTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCGAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGTGCGATGG GCGCTCCGAT GCCGCTCCAG GCGGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCTCT TCGTTTCAGT CACAACCCGC AAACCTCTAGA ATATCAATGA
2941 TCGAGCAGGA CGGCCTCCAC GCCGGCTCCC CCGCCGCTTG GGTGGAGCGC CTGTTCCGCC
3001 ACAGCTGGGC CCAGCAGACC ATCGGCTGCT CCGACGCCGC CGTGTTCGC CTGTCCGCC
3061 AGGGCCGCC CGTGCTGTTT GTGAAGACCG ACCTGTCCGG CGCCCTGAAC GAGCTGCAGG
3121 ACAGGGCCGC CCGCTGTCC TGGCTGGCCA CCACCGCGCT GCCCTGCGCC GCCGTGCTGG
3181 ACGTGGTGAC CGAGGCCGCG CCGGACTGGC TGCTGCTGGG CGAGGTGCCC GGCCAGGACC
3241 TGCTGTCTC CCACCTGGCC CCGCGCGAGA AGGTGTCCAT CATGGCCGAC GCCATGCGCC
3301 GCCTGCACAC CCTGGACCCC GCCACCTGCC CTTTCGACCA CCAGGCCAAG CACCGCATCG
3361 AGCGCGCCCG CACCCGCATG GAGGCCGGCC TGGTGGACCA GGACGACCTG GACGAGGAGC
3421 ACCAGGGCCT GGGCCCCGCC GAGCTGTTCC CCGCCTGAA GGGCCGCAIG CCCGACGGCG
3481 AGGACCTGGT GGTGACCCAC GCGGACGCC CTCTGCCCA CATCATGGTG GAGAACGGCC
3541 GCTTCTCCGG CTTTATCGAC TCGCGCCGCC TGGGCGTGGC CGACCGCTAC CAGGACATCG
3601 CCTGGCCAC CCGGACATC GCCGAGGAGC TGGGCGGCGA GTGGGCGGAC CGCTTCTTGG
3661 TGCTGTACGG CATCGCCGCC CCGACTCCC AGCGCATCGC CTTCTACCGC CTGCTGGACG
3721 AGTTCTTCTG ACAATTGGCA GCAGCAGCTC GGATAGTATC GACACACTCT GGACGCTGGT
3781 CGTGTGATGG ACTGTTGCCG CCACACTTGC TGCCTTGACC TGTGAATATC CCTGCCGCTT
3841 TTATCAAACA GCCTCAGTGT GTTTGATCTT GTGTGTACGC GCTTTTGCGA GTTGTAGCT
3901 GCTTGTGCTA TTTGCGAATA CCACCCCAAG CATCCCTTC CTTGTTTCA TATCGCTTGC
3961 ATCCCAACCG CAACTTATCT ACGCTGTCTT GCTATCCCTC AGCGCTGCTC CTGCTCCTGC
4021 TCACTGCCCC TCGCACAGCC TTGGTTTGGG CTCCGCTGT ATTCTCCTGG TACTGCAACC
4081 TGTAACACAG CACTGCAATG CTGATGCACG GGAAGTAGTG GGATGGGAAC ACAATGGAG
4141 GATCCCGCGT CTCGAACAGA GCGCGCAGAG GAACGCTGAA GGTCTCGCCT CTGTGCGACC
4201 TCAGCGCGGC ATACACCACA ATAACCACCT GACGAATGCG CTTGTTTCTT CGTCCATTAG
4261 CGAAGCGTCC GGTTCACACA CGTGCCACGT TGGCGAGGTG GCAGGTGACA ATGATCGGTG
4321 GAGCTGATGG TCGAAACGTT CACAGCCTAG GGATATCGAA TTCCTTTCTT GCGCTATGAC
4381 ACTTCCAGCA AAAGGTAGGG CGGGCTGCGA GACGGCTTCC CCGCGCTGCA TGCAACACCG
4441 ATGATGCTTC GACCCCCCGA AGCTCCTTCG GGGCTGCGAT GCGGCTCCGA TGCCGCTCCA
4501 GGGCGAGCGC TGTTTAAATA GCCAGGCCCC CGATTGCAAA GACATTATAG CGAGCTACCA
4561 AAGCCATATT CAAACACCTA GATCACTACC ACTTCTACAC AGGCCACTCG AGCTTGTGAT
4621 CGCACTCCGC TAAGGGGGCG CCTCTTCTC TTCGTTTCAG TCACAACCCG CAAACACTAG
4681 TATGGCCACC GCATCCACTT TCTCGGCGTT CAATGCCCGC TCGGCGGACC TGCGTCGCTC
4741 GGGGGGCTCC GGGCCCCGCG GCCAGCGAG GCCCTCCCC GTGCGCGGGC GCGCCCCCGA
4801 CTGGTCCATG CTGTTGCGCG TGATCACCAC CATCTTCTCC GCGCGCGAGA AGCAGTGAC
4861 CAACCTGGAG TGAAGCCCA AGCCCAACCC CCCCCAGCTG CTGGACGACC ACTTCGGCCC
4921 CCACGGCCTG GTGTTCCGCC GCACCTTCGC CATCCGAGC TACGAGGTGG GCGCCGACCG
4981 CTCCACCAGC ATCGTGGCCG TGATGAACCA CCTGCAGGAG GCGGCCCTGA ACCACGCCAA
5041 GTCCGTGGGC ATCTGGGGCG ACGGCTTCGG CACCACCTG GAGATGTCCA AGCGCGACCT
5101 GATCTGGGTG GTGAAGCGCA CCCACGTGGC CGTGGAGCGC TACCCCGCTT GGGGCGACAC
5161 GTGGAGGTG GAGTGCTGGG TGGGCGCCTC CGGCAACAAC GCGCCCGCTT ACCTTCTT
5221 GGTGCGGAC TGCAAGACCG GCGAGATCCT GACCCGCTGC ACCTCCTTGA GCGTGTATGAT
5281 GAACACCCGC ACCCGCCGCC TGAGCAAGAT CCCCAGGAG GTGCGCGGCG AGATCGGGCC
5341 CGCCTTCATC GACAACGTGG CCGTGAAGGA CGAGGAGATC AAGAAGCCCC AGAAGCTGAA
5401 CGACTCCACC GCCGACTACA TCCAGGGCGG CCTGACCCCC CGCTGGAACG ACCTGGACAT
5461 CAACACGAC GTGAACAACA TCAAGTACGT GGAAGGATC CTGGAGACCG TGCCCGACAG
5521 CATCTTCGAG AGCCACCACA TCTCTCCTT CACCATCGAG TACCGCCGCG AGTGACCAT
5581 GGACAGCGTG CTGCAGTCCC TGACCACCGT GAGCGGCGGC TCCTCCGAG CCGGCTGGT
5641 GTGCGGACAC CTGCTGCAGC TGGAGGCGG CAGCGAGGTG CTGCGCGCCA AGACCGAGTG
5701 GCGCCCCAAG CTGACCGACT CCTTCCGCGG CATCAGCGTG ATCCCGCGCG AGTCCAGCGT

5761 GATGGACTAC AAGGACCACG ACGGCGACTA CAAGGACCAC GACATCGACT ACAAGGACGA
5821 CGACGACAAG TGA CTGAGG CAGCAGCAGC TCGGATAGTA TCGACACACT CTGGACGCTG
5881 GTCGTGTGAT GGA CTGTTG CCGCACA CTGTCCTTGA CCTGTGAATA TCCCTGCCGC
5941 TTTTATCAAA CAGCCTCAGT GTGTTTGATC TTGTGTGTAC GCGCTTTTGC GAGTTGCTAG
6001 CTGCTTGTGC TATTTGCGAA TACCACCCCC AGCATCCCCC TCCCTCGTTT CATATCGCTT
6061 GCATCCCAAC CGCAACTTAT CTACGCTGTC CTGCTATCCC TCAGCGCTGC TCCTGCTCCT
6121 GCTCACTGCC CCTCGCACAG CCTTGGTTTG GGCTCCGCCT GTATTCTCCT GGTACTGCAA
6181 CCTGTAAACC AGCACTGCAA TGCTGATGCA CGGGAAGTAG TGGGATGGGA ACACAAATGG
6241 AAAGCTTGAG CTCTTGTTTT CCAGAAGGAG TTGCTCCTTG AGCCTTTTCAT TCTCAGCCTC
6301 GATAACCTCC AAAGCCGCTC TAATTGTGGA GGGGGTTTGA ATTTAAAGC TTGGAATGTT
6361 GGTTCTGTGC TCTGGAACAA GCCCAGACTT GTTGCTCCTT GGGAAAAGGA CCATCAGCTC
6421 CAAAAAACTT GCCGCTCAAA CCGCGTACCT CTGCTTTTCGC GCAATCTGCC CTGTTGAAAT
6481 CGCCACCACA TTCATATTGT GACGCTTGAG CAGTCTGTAA TTGCCTCAGA ATGTGGAATC
6541 ATCTGCCCCC TGTGCGAGCC CATGCCAGGC ATGTGCGGGG CGAGGACACC CGCCACTCGT
6601 ACAGCAGACC ATTATGCTAC CTCACAATAG TTCATAACAG TGACCATATT TCTCGAAGCT
6661 CCCCACGAG CACCTCCATG CTCTGAGTGG CCACCCCCCG GCCCTGCTGC TTGCGGAGGG
6721 CAGGTCAACC GGCATGGGGC TACCGAAATC CCCGACCGGA TCCCACCACC CCCGCGATGG
6781 GAAGAATCTC TCCCCGGGAT GTGGGCCCCAC CACCAGCACA ACCTGCTGGC CCAGGCGAGC
6841 GTCAAACCAT ACCACACAAA TATCCTTGGC ATCGGCCCTG AATTCTCTCT GCCGCTCTGC
6901 TACCCGGTGC TTCTGTCCGA AGCAGGGGTT GCTAGGGATC GCTCCGAGTC CGCAAACCCT
6961 TGTGCGGTGG CGGGGCTTGT TCGAGCTTGA AGAGCCTCTA GAGTCGACCT CGAGGCTATC
7021 AAGCTTGGCG TAATCATGGT CATAGCTGTT TCCTGTGTGA AATTGTTATC TCTCACAAT
7081 TCCACACAAC ATACGAGCCG GAAGCATAAA GTGTAAAGCC TGGGGTGCCT AATGAGTGAG
7141 CTAACCTACA TTAATTGCGT TCGCTCACT GCGCTTCTT CAGTCGGGAA ACCTGCTGTG
7201 CCAGCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTCGCTA TTGGGCGCTC
7261 TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCTGT CGGCTGCGGC GAGCGGTATC
7321 AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA
7381 CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGTGGCGTT
7441 TTTCTATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG
7501 GCGAAACCCG ACAGGACTAT AAAGATACCA GCGGTTTCCC CCTGGAAGCT CCCTCGTGCG
7561 CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTGCGGAAG
7621 CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC
7681 CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCCT TATCCGGTAA
7741 CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG C

SEQ ID NO: 237
pSZ1439

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTGTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCTTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGTCTCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTGGTGA GACTCAACC AAGTCATTCT GAGAATAGTG TATGCGCGCA CCGAGTTGCT
1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAA TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCCAACCT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCTTTT TTCAATATTA TTGAAGCATT TATCAGGGTT

1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGCTTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA
1741 ACTGTTGGGA AGGCGGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCCG CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCCG TCGTGCCTGT CGCTGATGTC
1981 CATCACCAGG TCCATGAGGT CTGCCTTGCG CCGGCTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCCTG ACGTGGTCGC GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGAC CGAGGCCGCC TCCAACGTGT TCCCTCAGAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGAAATCC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCTCGCT GCCGCCGCTT CTCCCGCACG CTTCTTTCCA
2341 GCACCGTGAT GGCAGGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTCGCCGTT AGTGTTGCCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTACCTTTT TTGGCGAAGG CATCGGCCTC GGCCTGCAGA GAGGACAGCA GTGCCCAGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCTCTT TCGTTTCAGT CACAACCCGC AAACCTCTAGA ATATCAATGC
2941 TGCTGCAGGC CTTCTGTTC CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAAGCAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCCAACGG CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCCGAACGA CACCGTCTGG GGGACGCCCT TGTTCCTGGG CCACGCCACG TCCGACGACC
3181 TGACCAACTG GGAGGACCAG CCCATCGCCA TCGCCCCGAA GCGCAACGAC TCCGGCGCCT
3241 TCTCCGGCTC CATGGTGGTG GACTACAACA ACACCTCCGG CTTCTTCAAC GACACCATCG
3301 ACCCGCGCCA GCGCTGCGTG GCCATCTGGA CCTACAACAC CCCGAGTCC GAGGAGCAGT
3361 ACATCTCCTA CAGCCTGGAG GCGGCTTACA CTTTCACCGA GTACCAGAAG AACCCCGTGC
3421 TGGCCGCCAA CTCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCAGAG
3481 AGTGGATCAT GACCGCGGCC AAGTCCGAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCGCGT TCGCCAACGA GGGCTTCCTC GGCTACCAGT
3601 ACGAGTGCCC CGGCCTGATC GAGGTCCCA CCGAGCAGGA CCCCAGCAAG TCCTACTGGG
3661 TGATGTTTAT CTCCATCAAC CCCGGCGCCC CCGCCGGCGG CTCCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CTTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCCTGGGCAT CGCGTGGGCC TCCAACCTGG AGTACTCCGC CTTCTGTCCC ACCAACCCCT
3901 GCGGCTCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCCG AGCCGATCCT GAACATCAGC AACGCCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTCGACC
4081 TGTCCAACAG CACCGGCACC CTGGAGTTCG AGCTGGTGTA CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT GCGGACCTCT CCCTCTGGTT CAAGGGCTTG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCTCT CTTCTTCTCT GACCGCGGGA
4261 ACAGCAAGGT GAAGTTCGTG AAGGAGAACC CCTACTTCAC CAACCGCATG AGCGTGAACA
4321 ACCAGCCCTT CAAGAGCGAG AACGACCTGT CTAATACAA GGTGTACGGC TTGCTGGACC
4381 AGAACATCCT GGAGCTGTAC TTCAACGACG GCGACGTCGT GTCCACCAAC ACCTACTTCA
4441 TGACCACCGG GAACGCCCTG GGCTCCGTGA ACATGACGAC GGGGGTGGAC AACCTGTTCT
4501 ACATCGACAA GTTCCAGGTG CGCGAGGTCA AGTGACAATT GGCAGCAGCA GCTCGGATAG
4561 TATCGACACA CTCTGACGCG TGGTCTGTGT ATGGACTGTT GCCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTTGTGTGT
4681 ACGCGCTTTT GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC CCAGCATCCC
4741 CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCACTT ATCTACGCTG TCCTGCTATC
4801 CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGCTCCGC
4861 CTGTATTCTC CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGATG CACGGGAAGT
4921 AGTGGGATGG GAACACAAAT GGAGGATCCC GCGTCTCGAA CAGAGCGCGC AGAGGAACGC

4981 TGAAGGTCTC GCCTCTGTCTG CACCTCAGCG CGGCATACAC CACAATAACC ACCTGACGAA
5041 TGCGCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGGTTCA CACACGTGCC ACGTTGGCGA
5101 GGTGGCAGGT GACAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 CGAATTCCCTT TCTTGGCTTA TGACACTTCC AGCAAAAGGT AGGGCGGGCT GCGAGACGGC
5221 TTCCCGGCGC TGCATGCAAC ACCGATGATG CTTCGACCCC CCGAAGCTCC TTCGGGGCTG
5281 CATGGGCGCT CCGATGCCGC TCCAGGGCGA GCGCTGTTTA AATAGCCAGG CCCCCGATTG
5341 CAAAGACATT ATAGCGAGCT ACCAAAGCCA TATTCAAACA CCTAGATCAC TACCACTTCT
5401 ACACAGGCCA CTCGAGCTTG TGATCGCACT CCGCTAAGGG GGCGCCTCTT CCTCTTCGTT
5461 TCAGTCACAA CCCGCAAACA CTAGTATGGC CACCGCATCC ACTTTCTCGG CGTTCAATGC
5521 CCGCTGCGGC GACCTGCGTC GCTCGGCGGG CTCCGGGCCC CGGCGCCCAG CGAGGCCCTT
5581 CCCCCTGCGC GGGCGCGCCG CCTCGGCCAA GGCCTCCAAG ACCATCGGCG AGGGCTCCGA
5641 GAACCTGAAC GTGCGCGGCA TCATCGCCAA GCCCACCTCC TCCTCCGCGG CCAAGCAGGG
5701 CAAGGTGATG GCCCAGGCCG TGCCCAAGAT CAACGGCGCC AAGGTGGGCC TGAAGGCCGA
5761 GTCCCAAGG GCCGAGGAGG ACGCCGCCCC CTCCTCCGCC CCCCACCTT TCTACAACCA
5821 GCTGCCCGAC TGGTCCGTGC TGCTGGCCGC CGTGACCACC ATCTTCTTGG CCGCTCGAGA
5881 GCAGTGGACC CTGCTGGACT GGAAGCCCCG CCGCCCCGAC ATGCTGACCG GCGCCTTCTC
5941 CCTGGGCAAG ATCGTGCAGG ACGGCCTGGT GTTCCGCCAG AACTTCTCCA TCCGCTCCTA
6001 CGAGATCGGC GCCGACCGCA CCGCTCCAT CGAGACCCTG ATGAACCACC TGCAGGAGAC
6061 CGCCCTGAAC CACGTGCGCA ACGCCGGCCT GCTGGGCGAC GGCTTCGCGG CCACCCCCGA
6121 GATGTCCAAG CGCAACCTGA TCTGGGTGGT GACCAAGATG CAGGTGCTGA TCGAGACTA
6181 CCCCTCTCTG GCGGACGTGG TGGAGGTGGA CACCTGGGTG GCGCCTCCG GCAAGAACGG
6241 CATGCGCCG GACTGGCACG TCGCGGACTA CCGCACCGGC CAGACCATCC TGCAGCGCAC
6301 CTCCATCTGG GTGATGATGG ACAAGCACAC CCGCAAGCTG TCCAAGATGC CCGAGGAGGT
6361 GCGCGCCGAG ATCGGCCCTT ACTTCATGGA GCACGCCGCC ATCGTGGACG AGGACTCCCG
6421 CAAGCTGCCC AAGCTGGACG ACGACACCGC CGACTACATC AAGTGGGGCC TGACCCCCCG
6481 CTGGTCCGAC CTGGACGTGA ACCAGCACGT GAACAACGTG AAGTACATCG GCTGGATCCT
6541 GGAGTCCGCC CCCATCTCCA TCCTGGAGAA CCACGAGCTG GCCTCCATGA CCCTGGAGTA
6601 CCGCCGCGAG TCGGCGCCG GACTCCGTGT GCAGTCCCTG ACCGCGCTGG CCAACGACTG
6661 CACCGCGCGC CTGCCGAGG CCTCCATCGA GTGCCAGCAC CTGCTTCGAG TGGAGTCGG
6721 CGCCGAGATC GTGCGCGGCC GCACCCAGTG GCGCCCCCGC CGCGCCTCCG GCGCCACCTC
6781 CGCCGGCTCC GCCATGGACT ACAAGGACCA CGACGGCGAC TACAAGGACC ACGACATCGA
6841 CTACAAGGAC GACGACGACA AGTGAATCGA TAGATCTCTT AAGGCAGCAG CAGCTCGGAT
6901 AGTATCGACA CACTCTGGAC GCTGGTCTGT TGATGGACTG TTGCCGCCAC ACTTGCTGCC
6961 TTGACCTGTG AATATCCCTG CCGCTTTTAT CAAACAGCCT CAGTGTGTTT GATCTTGTGT
7021 GTACGCGCTT TTGCGAGTTG CTAGCTGCTT GTGCTATTTG CGAATACCAC CCCCAGCATC
7081 CCCTTCCCTC GTTTCATATC GCTTGCATCC CAACCGCAAC TTATCTACGC TGTCTTGCTA
7141 TCCCTCAGCG CTGCTCCTGC TCCTGTCTAC TGCCCCCTCG ACAGCCTTGG TTGGGGCTCC
7201 GCCTGTATTG TCCTGGTACT GCAACCTGTA AACCAGCACT GCAATGCTGA TGCACGGGAA
7261 GTAGTGGGAT GGGAACACAA ATGGAAAGCT TAATTAAGAG CTCTTGTTTT CCAGAAGGAG
7321 TTGCTCCTTG AGCCTTTCAT TCTCAGCCTC GATAACCTCC AAAGCCGCTC TAATTGTGGA
7381 GGGGGTTTGA ATTTAAAAGC TTGGAAATGT GGTTCGTGCG TCTGGAACAA GCCCAGACTT
7441 GTTGCTCACT GGGAAAAGGA CCATCAGCTC CAAAAAACTT GCCGCTCAAA CCGCGTACCT
7501 CTGCTTTTCG GCAATCTGCC CTGTTGAAAT CGCCACCACA TTCATATTGT GACGCTTGAG
7561 CAGTCTGTAA TTGCCTCAGA ATGTGGAATC ATCTGCCCCC TGTGCGAGCC CATGCCAGGC
7621 ATGTGCGGGG CGAGGACACC CGCCACTCGT ACAGCAGACC ATTATGCTAC CTCACAATAG
7681 TTCATAACAG TGACCATATT TCTCGAAGCT CCCCACGAG CACCTCCATG CTCTGAGTGG
7741 CCACCCCCCG GCCCTGGTGC TTGCGGAGGG CAGGTCAACC GGCATGGGGC TACCGAAATC
7801 CCCGACCGGA TCCCACCACC CCCGCGATGG GAAGAATCTC TCCC CGGGAT GTGGGCCAC
7861 CACCAGCACA ACCTGCTGGC CCAGGCGAGC GTCAAACCAT ACCACACAAA TATCCTTGGC
7921 ATCGGCCCTG AATTCCTTCT GCCGCTCTGC TACCCGGTGC TTCTGTCCGA AGCAGGGGTT
7981 GCTAGGGATC GCTCCGAGTC GCAAAACCTT TGTGCGGTGG CGGGGCTTGT TCGAGCTTGA
8041 AGAGCCTCTA GAGTCGACCT CGAGGCATGC AAGCTTGGCG TAATCATGGT CATAGCTGTT
8101 TCCTGTGTGA AATTGTTATC CGCTCACAAT TCCACACAAC ATACGAGCCG GAAGCATAAA
8161 GIGTAAAGCC TGGGTGCTCT AATGAGTGAG CTAACCTACA TTAATTGCGT TGCCTCACT
8221 GCCGCTTTTC CAGTCGGGAA ACCTGTCTGT CCAGCTGCAT TAATGAATCG GCCAACGCGC
8281 GGGGAGAGGC GGTTCGCTGA TTGGGCGCTC TTCCGCTTCC TCGCTCACTG ACTCGCTGCG
8341 CTCGGTCTGT CGGCTGCGGC GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC
8401 CACAGAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAG AAAAGGCCAG
8461 GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCGCGCCCC CTGACGAGCA
8521 TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA
8581 GCGGTTTCCC CCTGGAAGCT CCCTCGTGGC CTCTCTGTGT CCGACCCTGC CGCTTACCGG

```

8641 ATACCTGTCC GCCTTTCTCC CTTCGGGAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG
8701 GTATCTCAGT TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT
8761 TCAGCCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA
8821 CGACTTATCG C

```

SEQ ID NO: 238
pSZ1449

```

1  CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61  AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAGAA GAGTTGGTAG CTCTTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTGTGTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCTTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA CTAGAGTAA GTAGTTGCGC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAAGT GGCCGCAGTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTGGTGA GACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTGCCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAG ACCTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCAGGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGCTG TAAGCGGATG
1561 CCGGGATGAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGCTGGC
1621 TTAATATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGCCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCCG CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCGC TCGTGCGCGT CGCTGATGTC
1981 CATCACCAGG TCCATGAGGT CTGCCTTGCG CCGGTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCTTG ACGTGGTCCG GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC CTGCGCCGAC CGAGGCCGCC TCCAATGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGATCC CTACCAGTCA TGCTTTTACC TGGATGACGG
2281 CCGCGAACA GCTGTCCAGC GACCTCGCT GCGCCGCTT CTCCGCGACG CTCTTTTCCA
2341 GCACCGTGAT GCGCGGAGCC AGCGCCGAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGCCA TCCTTTGTCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGGCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTACACCTT TTGGCGAAGG CATCGGCCTC GGCTTGCGA GAGGACAGCA GTGCCAGGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCAGCAA
2641 AAGGTAGGGC GGGTTCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTTCTCTCT TCGTTTCAGT CACAACCCGC AAACCTAGA ATATCAATGC
2941 TGCTGACAGG CTTCCTGTTT CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCCAACGG CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA

```

3121 ACCCGAACGA CACCGTCTGG GGGACGCCCT TGTTCCTGGG CCACGCCACG TCCGACGACC
3181 TGACCAACTG GGAGGACCAG CCCATCGCCA TCGCCCCGAA GCGCAACGAC TCCGGCGCCT
3241 TCTCCGGCTC CATGGTGGTG GACTACAACA ACACCTCCGG CTTCTTCAAC GACACCATCG
3301 ACCCGCGCCA GCGCTGCGTG GCCATCTGGA CCTACAACAC CCCGGAGTCC GAGGAGCAGT
3361 ACATCTCCTA CAGCCTGGAC GGGCGCTACA CCTTCACCGA GTACCAGAAG AACCCCGTGC
3421 TGGCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCCAGA
3481 AGTGGATCAT GACCGCGGCC AAGTCCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACGTGAAGTC CTGGAAGCTG GAGTCCGCGT TCGCCAACGA GGGCTTCCCT GAGTACCAGT
3601 ACGAGTGCCC CGGCCTGATC GAGGTCCCCA CCGAGCAGGA CCCCAGCAAG TCCTACTGGG
3661 TGATGTTTAT CTCCATCAAC CCCGGCGCCC CGGCCGGCGG CTCCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CCTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCCTGGGCAT CGCGTGGGCC TCCAACTGGG AGTACTCCGC CTTCTGTCCC ACCAACCCCT
3901 GCGCTCTCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGCAGGA GCTGATCAAC CTGAAGGCCG AGCCGATCCT GAACATCAGC AACGCCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTGACCC
4081 TGTCCAACAG CACCGGCACC CTGGAGTTCG AGCTGGTGTA CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT GCGGACCTCT CCCTCTGGTT CAAGGGCCTG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCCCT CTTCTTCCCT GACCGCGGGA
4261 ACAGCAAGGT GAAGTTCGTG AAGGAGAACC CCTACTTCAC CAACCGCATG AGCGTGAACA
4321 ACCAGCCCTT CAAGAGCGAG AACGACCTGT CTTACTACAA GGTGTACGGC TTGCTGGACC
4381 AGAATACCTT GGAGCTGTAC TTCAACGACG GCGACGTCTG GTCCACCAAC ACCTACTTCA
4441 TGACCACCGG GAACGCCCTG GGCTCCGTGA ACATGACGAC GGGGGTGGAC AACCTGTTCT
4501 ACATCGACAA GTTCCAGGTG CGCGAGGTCA AGTGACAATT GGCAGCAGCA GCTCGGATAG
4561 TATCGACACA CTCTGGACGC TGGTCTGTGT ATGGACTGTT GCCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTGTGTGT
4681 ACGCGCTTTT GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC CCAGCATCCC
4741 CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGCTATC
4801 CCTCAGCGCT GCTCCTGCTC CTGCTCAGTG CCCCTCGCAC AGCCTTGATT TGGGCTCCGC
4861 CTGTATTCTC CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGTCTGAC CACGGGAAGT
4921 AGTGGGATGG GAACACAAAT GGAGGATCCC GCGTCTCGAA CAGAGCGCGC AGAGGAACGC
4981 TGAAGGTCTC GCCTCTGTCT CACCTCAGCG CGGCATACAC CACAATAACC ACCTGACGAA
5041 TCGCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGGTTCA CACACGTGCC ACGTTGGCGA
5101 GGTGGCAGGT GACAAATGATC GGTGGAGCTG ATGGTCAAAA CGTTTACAGC CTAGGGATAT
5161 CGAATTCGGC CGACAGGACG CGCGTCAAAAG GTGCTGGTCTG TGTATGCCCT GGCCGGCAGG
5221 TCGTTGCTGC TGCTGGTTAG TGATTCCGCA ACCCTGATT TGGCGTCTTA TTTTGGCTG
5281 GCAAAACGTG GCGCCCGCGA GCCGGGCGCG CGGCGATGCG GTGCCCCAGT GCTGCCGAA
5341 TCCAAGGGAG GCAAGAGCGC CCGGGTCAGT TGAAGGGCTT TACGCGCAAG GTACAGCCGC
5401 TCCTGCAAGG CTGCGTGGTG GAATTGGACG TGCAGGTCCT GCTGAAGTTC CTCCACCGCC
5461 TACCAGCGG ACAAAGCACC GGTGTATCAG GTCCGTGTCA TCCACTCTAA AGAACTCGAC
5521 TACGACCTAC TGATGGCCCT AGATTCTTCA TCAAAAACGC CTGAGACACT TGCCAGGAT
5581 TGAAACTCCC TGAAGGGACC ACCAGGGGCC CTGAGTTGTT CCTTCCCCC GTGGCGAGCT
5641 GCCAGCCAGG CTGTACCTGT GATCGAGGCT GCGGGGAAAA TAGGCTTCGT GTGCTCAGGT
5701 CATGGGAGGT GCAGGACAGC TCATGAAACG CCAACAATCG CACAATTCAT GTCAAGCTAA
5761 TCAGCTATTT CCTCTTCACG AGCTGTAATT GTCCCAAAT TCTGGTCTAC CGGGGGTGAT
5821 CCTTCGTGTA CGGGCCCTTC CCTCAACCCT AGGTATGCGC GCATGCGGTC GCCGCGCAAC
5881 TCGCGCGAGG GCCGAGGGTT TGGGACGGGC CGTCCCGAAA TGCAGTTGCA CCCGGATGCG
5941 TGGCACCTTT TTTGCGATAA TTTATGCAAT GGACTGCTCT GCAAAATTCT GGCTCTGTCT
6001 CCAACCCTAG GATCAGCGGC GTAGGATTTT GTAATCATTC GTCCTGATGG GGAGCTACCG
6061 ACTACCCTAA TATCAGCCCG ACTGCCTGAC GCCAGCGTCC ACTTTTGTGC ACACATTCCA
6121 TTCGTGCCCC AGACATTCA TTGTGGTGCG AAGCGTCCCC AGTTACGCTC ACCTGTTTCC
6181 CGACCTCCTT ACTGTTCTGT CGACAGAGCG GGGCCACAGG CCGGTGCGAG CCACTAGTAT
6241 GGCCACCGCA TCCACTTTCT CGGCGTTCAA TGCCCGCTGC GCGACCTGC GTCGCTCGGC
6301 GGGCTCCGGG CCCC GGCGCC CAGCGAGGCC CCTCCCCGTG CGCGGGCGCG CCTCCCAGCT
6361 GCGCAAGCCC GCCCTGGACC CCTGCGCGC CGTGATCTCC GCCGACCAGG GCTCCATCTC
6421 CCCC GTGAAC TCCTGCACCC CCGCCGACCG CCTGCGCGCC GGCCGCCTGA TGGAGGACGG
6481 CTACTCTTAC AAGGAGAAGT TCATCGTGCG CTCCTACGAG GTGGGCATCA ACAAGACCGC
6541 ACCGTGGAG ACCATCGCCA ACCTGCTGCA GGAGGTGGCC TGCAACCACG TGCAGAAGTG
6601 CCGCTTCTCC ACCGACGGCT TCGCCACCAC CCTGACCATG CGCAAGCTGC ACCTGATCTG
6661 GGTGACCGCC CGCATGCACA TCGAGATCTA CAAGTACCCC GCCTGGTCCG ACGTGGTGGA
6721 GATCGAGACC TGGTGCCAGT CCGAGGGCCG CATCGGCACC CGCCGCGACT GGATCTGCG

```

6781 CGACTCCGCC ACCAACGAGG TGATCGGCCG CGCCACCTCC AAGTGGGTGA TGATGAACCA
6841 GGACACCCGC CGCCTGCAGC GCGTGACCGA CGAGGTGCGC GACGAGTACC TGGTGTCTTG
6901 CCCCCGCGAG CCCCCGCTGG CCTTCCCCGA GGAGAACAAC TCCTCCCTGA AGAAGATCCC
6961 CAAGCTGGAG GACCCCGCCC AGTACTCCAT GCTGGAGCTG AAGCCCCGCC GCGCCGACCT
7021 GGACATGAAC CAGCACGTGA ACAACGTGAC CTACATCGGC TGGGTGCTGG AGTCCATCCC
7081 CCAGGAGATC ATCGACACCC ACGAGCTGCA GGTGATCACC CTGGACTACC GCCGCGAGTG
7141 CCAGCAGGAC GACATCGTGG ACTCCCTGAC CACCTCCGAG ATCCCCGACG ACCCCATCTC
7201 CCAAGTTCACC GGCACCAACG GCTCCGCCAT GTCCCTCCATC CAGGGCCACA ACGAGTCCCA
7261 GTTCCTGCAC ATGTGTGCGC TGTCGAGAAA CGGCCAGGAG ATCAACCGCG GCCGCACCCA
7321 GTGGCGCAAG AAGTCTTCCC GCATGGACTA CAAGGACCAC GACGGCGACT ACAAGGACCA
7381 CGACATCGAC TACAAGGACG ACGACGACAA GTGAATCGAT AGATCTCTTA AGGCAGCAGC
7441 AGCTCGGATA GTATCGACAC ACTCTGACG CTGGTCTGTG TATGGACTGT TGCCGCCACA
7501 CTTGCTGCCT TGACCTGTGA ATATCCCTGC CGCTTTTATC AAACAGCCTC AGTGTGTTTG
7561 ATCTTGTGTG TACGCGCTTT TGCGAGTTG TAGCTGCTTG TGCTATTTGC GAATACCACC
7621 CCCAGCTACC CCTTCCCTCG TTTTCATATC CTTCATCCC AACCGCACT TATCTAGCCT
7681 GTCTGTCTAT CCTCAGCGC TGCTCCTGCT CTTGCTCACT GCCCTCGCA CAGCCTTGGT
7741 TTGGGCTCCG CCTGTATTCT CTTGGTACTG CAACCTGTAA ACCAGCACTG CAATGCTGAT
7801 GCACGGGAAG TAGTGGGATG GGAACACAAA TGGAAAGCTT AATTAAGAGC TCTTGTTTTC
7861 CAGAGGAGT TGCTCCTTGA GCCTTTTCAAT CTCAGCCTCG ATAACCTCCA AAGCCGCTCT
7921 AATTGTGGAG GGGGTTCGAA TTTAAAAGCT TGAATGTTG GTTCGTGCGT CTGGAACAAG
7981 CCCAGACTTG TTGCTCACTG GGAAGGAGC CATCAGCTCC AAAAACTTG CCGCTCAAAC
8041 CGCGTACCTC TGCTTTGCGC CAATCTGCCC TGTTGAAATC GCCACCACAT TCATATTGTG
8101 ACGCTTGAGC AGTCTGTAAT TGCCTCAGAA TGTGGAATCA TCTGCCCCCT GTGCGAGCCC
8161 ATGCCAGGCA TGTGCGGGC GAGGACACCC GCCACTCGTA CAGCAGACCA TTATGCTACC
8221 TCACAATAGT TCATAACAGT GACCATAATT CTCGAAGCTC CCCAACGAGC ACCTCCATGC
8281 TCTGAGTGGC CACCCCCCGG CCCTGGTGCT TGCGGAGGGC AGGTCAACCG GCATGGGGCT
8341 ACCGAAATCC CCGACCGGAT CCCACCACCC CCGCGATGGG AAGAATCTCT CCCCGGGATG
8401 TGGGCCACCC ACCAGCACAA CCTGCTGGCC CAGGCGAGCG TCAAACCATA CCACACAAAT
8461 ATCCTTGGCA TCGGCCCTGA ATTCTTCTG CCGCTCTGCT ACCCGGTGCT TCTGTCCGAA
8521 GCGAGGGTTG CTAGGGATCG CTCGAGTCC GCAAACCCTT GTCGCGTGGC GGGGCTTGTT
8581 CGAGCTTGAA GAGCCTCTAG AGTCGACCTG CAGGCATGCA AGCTTGGCGT AATCATGGTC
8641 ATAGCTGTTT CCTGTGTGAA ATTGTTATCC GCTCACAATT CCACACAACA TACGAGCCGG
8701 AAGCATAAAG TGTAAGCCT GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT
8761 GCGCTCACTG CCGCTTTTCC AGTCGGGAAA CCTGTCTGTC CAGCTGCATT AATGAATCGG
8821 CCAACGCGCG GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCCT CGCTCACTGA
8881 CTCGCTGCGC TCGGTCGTTT GGTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT
8941 ACGGTTATCC ACAGAAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCAGCA
9001 AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCC
9061 TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA
9121 AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGC GC TCTCCTGTTT CGACCCTGCC
9181 GCTTACCGGA TACCTGTCCG CCTTCTCTCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC
9241 ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA
9301 ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC
9361 GGTAAGACAC GACTTATCGC

```

SEQ ID NO: 239
pSZ1451

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCIGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTGTGTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCGGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG

```

721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGIATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCTATGCC ATCCGTAAGA TGCTTTTCTG
961 TGA CTGGTGA G TACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCCGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCCTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCCAAGT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATGTCTCAT GAGCGGATAC ATATTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTC
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAGCCCGT CAGGGCGGTG TGGCGGGTGT CCGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT CAGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCCG CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCCG TCGTGCCTG CGCTGATGTC
1981 CATCACAGG TCCATGAGGT CTGCTTGGC CCGGCTGAGC CACTGCTTCG TCCGGCGCGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCTG ACGTGGTTCG GGTCTGTTGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAACGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACACCCACGA GCCTTGCTTA GGCAGATCC CTACCACTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCTCTGCT GCCGCCGCTT CTCCCGCACG CTTCTTTCCA
2341 GCACCGTGAT GCGCGAGACC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGTCCA TCCTTTGACG
2461 ACCGGGGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACCTTT TTGGCGAAGG CATCGGCCTC GGCCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCTCT TCGTTTCACT CACAACCCGC AAACCTTAGA AGATCAATGC
2941 TGCTGCAGGC CTTCTGTTC CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCAACGG CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCGAACGA CACCGTCTGG GGGACGCCCT TGTCTTGGGG CCACGCCACG TCCGACGACC
3181 TGACCAACTG GGAGGACCAG CCCATCGCCA TCGCCCCGAA GCGCAACGAC TCCGGCGCCT
3241 TCTCCGGCTC CATGGTGGTG GACTACAACA ACACCTCCGG CTTCTTCAAC GACACCATCG
3301 ACCCGCGCCA GCGCTGCGTG GCCATCTGGA CCTACAACAC CCGGAGTCC CCGGAGTCC
3361 ACATCTCCTA CAGCCTGGAC GCGGGTACA CCTTCACCGA GTACCAAGG AACCCCGTGC
3421 TGGCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCAGA
3481 AGTGATCAT GACCGCGGCC AAGTCCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCCGT TCGCCAACGA GGGCTTCTCT GGCTACCACT
3601 ACGAGTGCCC CGGCCTGATC GAGGTCCCCA CCGAGCAGGA CCCCAGCAAG TCCTACTGGG
3661 TGA TGTTTAT CTCCATCAAC CCCGGCGCCC CCGCCGGCGG CTCTTCAAC CAGTACTTCG
3721 TCGGAGCCTT CAACGGCACC CACTTCGAGG CTTTCGACAA CCAGTCCCGC GTGGTGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCTGGGGCAT CGCGTGGGCC TCCAACGGG AGTACTCCGC CTTCTGTGCC ACCAACCCTT
3901 GCGCTCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCGG AGCCGATCCT GAACATCAGC AACGCCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTGACCC
4081 TGTTCAACAG CACCGGCACC CTGGAGTTCG AGCTGGTGTG CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT CCGGACCTCT CCCTCTGGTT CAAGGGCTCG GAGGACCCCG
4201 AAGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCCCT CTTCTTCTCT GACCGCGGGA
4261 ACAGCAAGGT GAAGTTCGTG AAGGAGAACC CCTACTTCAC CAACCGCATG AGCGTGAACA
4321 ACCAGCCCTT CAAGAGCGAG AACGACCTGT CCTACTACAA GGTGTACGGC TTGCTGGACC

4381 AGAACATCCT GGAGCTGTAC TTCAACGACG GCGACGTCGT GTCCACCAAC ACCTACTTCA
4441 TGACCACCGG GAACGCCCTG GGCTCCGTGA ACATGACGAC GGGGGTGGAC AACCTGTTCT
4501 ACATCGACAA GTTCCAGGTG CGCGAGGTCA AGTGACAATT GGCAGCAGCA GCTCGGATAG
4561 TATCGACACA CTCTGGACGC TGGTCGTGTG ATGGACTGTT GCCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTTGTGTGT
4681 ACGCGCTTTT GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC CCAGCATCCC
4741 CTTCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGCTATC
4801 CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGCTCCGC
4861 CTGTATTCTC CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGATG CACGGAAGT
4921 AGTGGGATGG GAACACAAAT GGAGGATCCC GCGTCTCGAA CAGAGCGCGC AGAGGAACGC
4981 TGAAGGTCTC GCCTCTGTCG CACCTCAGCG CGGCATACAC CACAATAACC ACCTGACGAA
5041 TGCGCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGGTTCA CACACGTGCC ACGTTGGCGA
5101 GGTGGCAGGT GACAAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 CGAATTCGGC CGACAGGACG CGCGTCAAAG GTGCTGGTCG TGTATGCCCT GCGCGGACAG
5221 TCGTTGCTGC TGCTGGTTAG TGATTCCAA ACCCTGATT TTTGCGCTG TTTTGGCTG
5281 GCAAACGCTG GCGCCCGCGA GCCGGGCGCG CGGCGATGCG GTGCCCCACG GCTGCCGGA
5341 TCCAAGGGAG GCAAGAGCGC CCGGGTCAGT TGAAGGGCTT TACGCGCAAG GTACAGCCGC
5401 TCCTGCAAGG CTGCGTGGTG GAATTGGACG TGCAGGTCCT GCTGAAGTTC CTCCACCGCC
5461 TCACCAGCGG ACAAAGCACC GGTGTATCAG GTCCGTGTCA TCCACTCTAA AGAACTCGAC
5521 TACGACCTAC TGATGGCCCT AGATTCTTCA TCAAAAACGC CTGAGACACT TGCCAGGAT
5581 TGAAACTCCC TGAAGGGACC ACCAGGGGCC CTGAGTTGTT CCTTCCCCC GTGGCGAGCT
5641 GCCAGCCAGG CTGTACCTGT GATCGAGGCT GGCGGAAAA TAGGCTTCGT TAGCTCAGGT
5701 CATGGGAGGT GCAGGACAGC TCATGAAACG CCAACAATCG CACAATTCAT GTCAAGCTAA
5761 TCAGCTATTT CCTCTTCACG AGCTGTAATT GTCCCAAAT TCTGGTCTAC CGGGGGTGAT
5821 CCTTCGTGTA CGGGCCCTTC CCTCAACCCT AGGTATGCGC GCATGCGGTC GCCGCGCAAC
5881 TCGCGCGAGG GCCGAGGGTT TGGGACGGGC CGTCCCGAAA TGCAGTTGCA CCCGGATGCG
5941 TGGCACCTTT TTTGCGATAA TTTATGCAAT GGACTGCTCT GCAAAATTCT GGCTCTGTCT
6001 CCAACCCTAG GATCAGCGGC GTAGGATTTT GTAATCATT CTCCTGATGG GGAGCTACCC
6061 ACTACCTTAA TATCAGCCCG ACTGCCTGAC GCCAGGTCC ACTTTTGTG ACACATTCCA
6121 TTCGTGCCCC AGACATTTCA TTGTGTTGCG AAGCGTCCCC AGTTACGCTC ACCTGTTTCC
6181 CGACCTCCTT ACTGTTCTGT CGACAGAGCG GGCCACAGG CCGGTGCGAG CCACTAGTAT
6241 GGCCACCGCA TCCACTTTCT CGGCGTTCAA TGCCCGCTGC GGCGACCTGC GTCGCTCGGC
6301 GGGCTCCGGG CCCCAGCGCC CAGCGAGGCC CCTCCCCGTG CGCGGGCGCG CCGCCACCGG
6361 CGAGCAGCCC TCCGGCGTGG CCTCCCTGCG CGAGGCCGAC AAGGAGAAGT CCCTGGGCAA
6421 CCGCTGCGC CTGGGCTCCC TGACCAGGGA CGGCCTGTCC TACAAGGAGA AGTTCGTGAT
6481 CCGCTGCTAC GAGGTGGGCA TCAACAAGAC CGCCACCATC GAGACCATCG CCAACCTGCT
6541 GCAGGAGGTG GGCGGCAACC ACGCCGAGGG CGTGGGCTTC TCCACCGACG GCTTCGCTAC
6601 CACCACCACC ATGCGCAAGC TGCACCTGAT CTGGGTGACC GCCCGCATGC ACATCGAGAT
6661 CTACCCTAC CCGCCTGGT CCGACGTGAT CGAGATCGAG ACCTGGGTGC AGGGCGAGGG
6721 CAAGTGGGG ACCCGCCGCG ACTGGATCCT GAAGGACTAC GCCAACGGCG AGGTGATCGG
6781 CCGCGCCACC TCCAAGTGGG TGATGATGAA CGAGGACACC CGCCGCCTGC AGAAGGTGTC
6841 CGACGACGTG CGCGAGGAGT ACCTGGTGTG CTGCCCCCGC ACCCTGCGCC TGGCCTTCCC
6901 CGAGGAGAAC ACAACTCCA TGAAGAAGAT CCCAAGCTG GAGGACCCG CCGAGTACTC
6961 CCGCTGGGC CTGGTGCCCC GCCGCTCCGA CCTGGACATG AACAAGCAG TGAACAACGT
7021 GACCTACATC GGCTGGGCCC TGGAGTCCAT CCCCCCGAG ATCATCGACA CCCACGAGCT
7081 GCAGGCCATC ACCCTGGACT ACCGCGCGCA GTGCCAGCGC GACGACATCG TGGACTCCCT
7141 GACCTCCCGC GAGCCCTGG GCAACGCCGC CGGCGTGAAG TTCAAGGAGA TCAACGGCTC
7201 CGTGTCCCC AAGAAGGACG AGCAGGACCT GTCCCGCTTC ATGCACCTGC TGCCTCCGC
7261 CGGCTCCGGC CTGGAGATCA ACCGTGCGG CACCGAGTGG CGCAAGAAGC CCGCAAGCG
7321 CATGGACTAC AAGGACCAG ACGGCGACTA CAAGGACCAC GACATCGACT ACAAGGACGA
7381 CGACGACAAG TGAATCGATA GATCTCTTAA GGCAGCAGCA GCTCGGATAG TATCGACACA
7441 CTCTGGACGC TGGTCGTGTG ATGGACTGTT GCCGCCACAC TTGCTGCCCT GACCTGTGAA
7501 TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTTGTGTGT ACGCGCTTTT
7561 GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC CCAGCATCCC CTTCCTCGT
7621 TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGCTATC CCTCAGCGCT
7681 GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGCTCCGC CTGTATTCTC
7741 CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGATG CACGGAAGT AGTGGGATGG
7801 GAACACAAAT GGAAAGCTTA ATTAAGAGCT CTTGTTTTCC AGAAGGAGTT GCTCCTTGAG
7861 CCTTTCATT TCAGCCTCGA TAACCTCAA AGCCGCTCTA ATTGTGGAG GGGTTCGAAT
7921 TTAAGGCTT GGAATGTTGG TTCGTGCGTC TGGAACAAGC CCAGACTTGT TGCTCACTGG
7981 GAAAAGGACC ATCAGCTCCA AAAAATTGCG CGCTCAAACC GCGTACCTCT GCTTTCGCGC

```

8041 AATCTGCCCT GTTGAAATCG CCACCACATT CATATTGTGA CGCTTGAGCA GTCTGTAATT
8101 GCCTCAGAAT GTGGAATCAT CTGCCCCCTG TGCGAGCCCA TGCCAGGCAT GTCGCGGGCG
8161 AGGACACCCG CCACTCGTAC AGCAGACCAT TATGCTACCT CACAATAGTT CATAACAGTG
8221 ACCATATTTT TCGAAGCTCC CCAACGAGCA CCTCCATGCT CTGAGTGGCC ACCCCCCGGC
8281 CCTGGTGCTT GCGGAGGGCA GGTCAACCGG CATGGGGCTA CCGAAATCCC CGACCGGATC
8341 CCACCACCCC CGCGATGGGA AGAATCTCTC CCCGGGATGT GGGCCCACCA CCAGCACAAAC
8401 CTGCTGGCCC AGGCGAGCGT CAAACCATAC CACACAAATA TCCTTGGCAT CGGCCCTGAA
8461 TTCCTTCTGC CGCTCTGCTA CCCGGTGCTT CTGTCCGAAG CAGGGGTTCG TAGGGATCGC
8521 TCCGAGTCCG CAAACCCTTG TCGCGTGGCG GGGCTTGTTT GAGCTTGAAG AGCCTCTAGA
8581 GTCGACCTGC AGGCATGCAA GCTTGGCGTA ATCATGGTCA TAGCTGTTTC CTGTGTGAAA
8641 TTGTTATCCG CTCACAATTC CACACAACAT ACGAGCCGGA AGCATAAAGT GTAAAGCCTG
8701 GGGTGCCATA TGAGTGAGCT AACTCACATT AATTGCGTTG CGCTCACTGC CCGCTTTCCA
8761 GTCGGGAAAC CTGTGCTGCC AGCTGCATTA ATGAATCGGC CAACGCGCGG GGAGAGGCGG
8821 TTTGCGTATT GGGCGCTCTT CCGCTTCCTC GCTCACTGAC TCGCTGCGCT CGGTCTGTTG
8881 GCTGCGGCGA GCGGTATCAG CTCAGTCAAA GGCGGTAAATA CGGTTATCCA CAGAATCAGG
8941 GGATAACGCA GGAAGAACA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA
9001 GGCCGCGTTG CTGGCGTTTT TCCATAGGCT CCGCCCCCTT GACGAGCATC ACAAAAATCG
9061 ACGCTCAAGT CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC
9121 TGAAGCTCC CTGCTGCGCT CTCCTGTTCG GACCCTGCCG CTTACCGGAT ACCTGTCCGC
9181 CTTTCTCCCT TCGGGAAGCG TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC
9241 GGTGTAGGTC GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCGTTT AGCCCGACCG
9301 CTGCGCCTTA TCCGTAACCT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGC

```

SEQ ID NO: 240
pSZ1452

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTAAA TCAATCTAAA GTATATATGA GTAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCCTGACT CCGCTGCTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GTCCTGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCCTCCG ATCGTTTGCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TCTTTTCTG
961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAG ACGTTCTTCG GGGCGAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC ATCTTTTACT TTCACCAGCG
1201 TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT CAGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCTT CTTGCTATT ACGCCAGTCA GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TCCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCGC CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCGC TCGTGCGCGT CGCTGATGTC
1981 CATACCAGG TCCATGAGGT CTGCCTTGCG CCGGCTGAGC CACTGCTTCG TCCGGGCGGC

```

2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCCTG ACGTGGTCGC GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAAC TGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGAATCC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCTCGCT GCCGCCGCTT CTCCCGCAGC CTTCTTTCCA
2341 GCACCGTGAT GGCGCGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGCCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTACCTTTT TTGGCGAAGG CATCGGCCTC GGCCTGCAGA GAGGACAGCA GTGCCAGGCC
2581 GCTGGGGGTT GGCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGGAGCGCT
2761 GTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTCTACACA GGCCACTCGA GCTTGTGATC GCCTCCGCT
2881 AACGGGGCGC CTCTCCTCT TCGTTTCAGT CACAACCCGC AAACCTAGA ATATAGTGC
2941 TGCTGCAGGC CTTCTGTTC CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCCAACGG CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCCGAACGA CACGCTCTGG GGGACGCCCT TGTCTGCGG CCACGCCAGC TCCGACGACC
3181 TGACCAACTG GGAGGACCGA CCCATCGCCA TCGCCCCGAA GCGCAACGAC TCCGGCGCCT
3241 TCTCCGGCTC CATGGTGGTG GACTACAACA ACACCTCCGG CTTCTTCAAC GACACCATCG
3301 ACCCGCGCCA GCGCTGCGTG GCCAICTGGA CCTACAACAC CCCGGAGTCC GAGGAGCAGT
3361 ACATCTCCTA CAGCCTGGAC GCGCGCTACA CCTTCACCGA GTACCAGAAG AACCCCGTGC
3421 TGCGCCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCAGA
3481 AGTGGATCAT GACCGCGGCC AAGTCCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCGCGT TCGCCAACGA GGGCTTCCTC GGCTACCAGT
3601 ACGAGTGCCC CGGCTTGATC GAGGTCCCCA CCGAGCAGGA CCCCAGCAAG TCCTACTGGG
3661 TGATGTTTAT CTCCATCAAC CCCGGCGCCC CGGCCGGCGG CTCCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CCTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCGAGACT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCTGGGCAT CGCGTGGGCC TCCAAC TGGG AGTACTCCGC CTTCTGTCCT ACCAACCCCT
3901 GCGCTCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGCCG AGCCGATCCT GAACATCAGC AACGCCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTCGACC
4081 TGTC AACAG CACCGGCACC CTGGAGTTCG AGCTGGTGTA CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT GCGGACCTCT CCCTCTGGTT CAAGGGCTCG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGTGT CCGCGTCTC CTCTCTCCTG CACCGGAGGA
4261 ACAGCAAGGT GAAGTTCGTG AAGGAGAACC CCTACTTCAC CAACCGCATG AGCGTGAACA
4321 ACCAGCCCTT CAAGAGCGAG AACGACCTGT CCTACTACAA GGTGTACGGC TTGCTGGACC
4381 AGAACATCCT GGAGCTGTAC TTCAACGACG GCGACGTCGT GTCCACCAAC ACCTACTTCA
4441 TGACCAACGG GAACGCCCTG GGCTCCGTGA ACATGACGAC GGGGGTGGAC AACCTGTTCT
4501 ACATCGACAA GTTCCAGGTG CGCGAGGTCA AGTGACAATT GGCAGCAGCA GCTCGGATAG
4561 TATCGACACA CTCTGGACGC TGGTCGTGTG ATGGACTGTT GCCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC CTTTTATCA AACAGCCTCA GTGTGTTGA TCTGTGTGT
4681 ACGCGCTTTT GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC CCAGCATCCC
4741 CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGCTATC
4801 CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGCTCCGC
4861 CTGTATTCTC CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGATG CACGGGAAGT
4921 AGTGGGATGG GAACACAAAT GGAGGATCCC GCGTCTCGAA CAGAGCGCGC AGAGGAACGC
4981 TGAAGGTCTC GCCTCTGTCT CACCTCAGCG CGGCATACAC CACAATAACC ACCTGACGAA
5041 TGCCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGGTTCA CACAGTGCC ACCTTGGCGA
5101 GGTGGCAGGT GACAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 CGAATTCGGC CGACAGGACG CGCGTCAAAG GTGCTGGTCG TGTATGCCCT GGCCGGCAGG
5221 TCGTTGCTGC TGCTGGTTAG TGATTCCGCA ACCCTGAITT TGGCGTCTTA TTTTGGCGTG
5281 GCAAACGCTG GCGCCCGCGA GCCGGGCGCG CCGCGATGCG GTGCCCCACG GCTGCCGGAA
5341 TCCAGGGGAG GCAAGAGCGC CCGGGTCAGT TGAAGGGCTT TACGCGCAAG GTACAGCCGC
5401 TCCTGCAAGG CTGCGTGGTG GAATTGGACG TGCAGGTCCT GCTGAAGTTC CTCCACCGCC
5461 TCACCAGCGG ACAAAGCACC GGTGTATCAG GTCCGTGTCA TCCACTCTAA AGAAGCTGAC
5521 TACGACCTAC TGATGGCCCT AGATTCTTCA TCAAAAACGC CTGAGACACT TGCCCAAGAT
5581 TGAAGTCCC TGAAGGGACC ACCAGGGGCC CTGAGTTGTT CCTTCCCCC GTGGCGAGCT
5641 GCCAGCCAGG CTGTACCTGT GATCGAGGCT GGCGGGAAAA TAGGCTTCGT GTGCTCAGGT

5701 CATGGGAGGT GCAGGACAGC TCATGAAACG CCAACAATCG CACAATTTCAT GTCAAGCTAA
5761 TCAGCTATTT CCTCTTCACG AGCTGTAATT GTCCCAAAAT TCTGGTCTAC CGGGGGTGAT
5821 CCTTCGTGTA CGGGCCCTTC CCTCAACCT AGGTATGCGC GCATGCGGTC GCCGCGCAAC
5881 TCGCGCGAGG GCCGAGGGTT TGGGACGGGC CGTCCCGAAA TGCAGTTGCA CCCGGATGCG
5941 TGGCACCTTT TTTGCGATAA TTTATGCAAT GGACTGCTCT GCAAAATTCT GGCTCTGTGC
6001 CCAACCTAG GATCAGCGGC GTAGGATTTT GTAATCATTG GTCCGTATGG GGAGCTACCG
6061 ACTACCCTAA TATCAGCCCG ACTGCCTGAC GCCAGCGTCC ACTTTTGTGC ACACATTCCA
6121 TTCGTGCCCA AGACATTTCA TTGTGGTGGC AAGCGTCCCC AGTTACGCTC ACCTGTTTTCC
6181 CGACCTCCTT ACTGTTCTGT CGACAGAGCG GGCCACAGG CCGGTCGCAG CCACTAGTAT
6241 GGCCACCGCA TCCACTTTCT CGGCGTTCAA TGCCCGCTGC GCGGACCTGC GTCGCTCGGC
6301 GGGCTCCGGG CCCC GGCGCC CAGCGAGGCC CCTCCCCGTG CGCGGGCGCG CCGCCAACGC
6361 CCAGGCCCCC ACCAAGATCA ACGGCTCCAC CGACGACGCC CAGCTGCCCC CCCCCCGCAC
6421 CTTGCCTAAC CAGCTGCCCC ACTGGTCCAT GCTGCTGGCC GCCATACCA CGCTGTTCCT
6481 GGCCGCGGAG AAGCAGTGGA TGATGCTGGA CTGGAAGCCC CGCCGCCCGG ACATGTTGAT
6541 CGACACCTTC GGCTTGGGCC GCATCGTGCA GGACGGCCTG GTGTTCCGCC AGAATTCTC
6601 CATCCGCTCC TACGAGATCG GCGCCGACCG CACCGCTCC ATCGAGACCG TGATGAACCA
6661 CCTGCAGGAG ACCGCCCTGA ACCAGGTGAA GACCGCGGCG CTGCTGGGCG ACGGCTTCGG
6721 CTCCACCCCC GAGATGTCCA AGCGCAACCT GATCTGGGTG GTGACCAAGA TGCAGGTGGA
6781 GGTGGACCGC TACCCACCT GGGGCGACGT GGTGCAGGTG GACACCTGGG TGTCCGCCCT
6841 CGGCAAGAAC GGCATGCGCC GCGACTGGCT GCTGCGCGAC GGCAACACCG GCGAGACCTC
6901 GACCCGCGCC TCCCTCGTGT GGGTGATGAT GAACAAGCTG ACCCGCGGCC TGTCCAAGAT
6961 CCCCAGGAGG GTGCGCGAGG AGATCGGCTC CTACTTCGTG AACTCCGACC CCGTGGTGGA
7021 GGAGGACGGC CGCAAGGTGA CCAAGCTGGA CGACAACACC GCCGACTTCG TGCGCAAGGG
7081 CCTGACCCCC AAGTGAACG ACCTGGACAT CAACCAGCAC GTGAACAACG TGAAGTACAT
7141 CGGCTGGATC CTGGAGTCCG CCCCCAGCC CATCTGGAG ACCCGCGAGC TGTCCGCCGT
7201 GACCCTGGAG TACCGCCGCG AGTGGCGCCG CGACTCCGTG CTGCGCTCCC TGACCGCCGT
7261 GTCCGGCGGC GCGTGGGCC ACCTGGGCCA CGCCGGCAAC GTGGAGTGCC AGCACGTGCT
7321 GCGCCTGGAG GACGGCGCCG AGATCGTGCG CGCCCGCAAC GAGTGGCGCC CCAAGTACAT
7381 CAACAACCTT TCCATCATGG GCCAGATCCC CACCGACGCC TCCATGGACT ACAAGGACCA
7441 CGACGGCGAC TACAAGGACC ACGACATCGA CTACAAGGAC GACGACGACA AGTGAATCGA
7501 TAGATCTCTT AAGGCAGCAG CAGCTCGGAT AGTATCGACA CACTCTGGAC GCTGGTCGTG
7561 TGATGGACTG TTGCCGCCAC ACTTGCTGCC TTGACCTGTG AATATCCCTG CCGCTTTTAT
7621 CAAACAGCCT CAGTGTGTTT GATCTTGTGT GTACGCGCTT TTGCGAGTTG CTAGCTGCTT
7681 GTGCTATTTG CGAATACCAC CCCCAGCATC CCCTTCCCTC GTTTCATATC GCTTGCATCC
7741 CAACCGCAAC TTATCTACGC TGTCTGTCTA TCCCTCAGCG CTGCTCTCAC TCCGTCTCAC
7801 TGCCCCCTCG ACAGCCTTGG TTTGGGCTCC GCCTGTATTG TCCTGGTACT GCAACCTGTA
7861 AACCAGCACT GCAATGCTGA TGACGGGAA GTAGTGGGAT GGGAACACAA ATGGAAAGCT
7921 TAATTAAGAG CTCTTGTTTT CCAGAAGGAG TTGCTCCTTG AGCCTTTCAT TCTCAGCCTC
7981 GATAACCTCC AAAGCCGCTC TAATGTGGA GGGGGTTCGA ATTTAAAGC TTGGAATGTT
8041 GGTTCGTGCG TCTGGAACAA GCCCAGACTT GTTGCTCACT GGGAAAAGGA CCATCAGCTC
8101 CAAAAACACT GCCGCTCAAA CCGCTACCT CTGCTTTCGC GCAATCTGCC CTGTTGAAAT
8161 CGCCACCACA TTCATATTGT GACGCTTGAG CAGTCTGTAA TTGCCTCAGA ATGTGGAATC
8221 ATCTGCCCCC TGTGCGAGCC CATGCCAGGC ATGTCGCGGG CGAGGACACC CGCCACTCGT
8281 ACAGCAGACC ATTATGCTAC CTCACAATAG TTCATAACAG TGACCATAAT TCTCGAAGCT
8341 CCCCACGAG CACCTCCATG CTCTGAGTGG CCACCCCCCG GCCCTGGTGC TTGCGGAGGG
8401 CAGGTCAACC GGCATGGGGC TACCGAAATC CCCGACCGGA TCCCACCACC CCCGCGATGG
8461 GAAGAATCTC TCCCCGGGAT GTGGGCCAC CACCAGCACA ACCTGCTGGC CCAGGCGAGC
8521 GTCAACCGAT ACCACACAAA TATCCTTGGC ATCGGCCCTG AATTCCTTCT GCCGCTCTGC
8581 TACCCGGTGC TTCTGTCCGA AGCAGGGGT GCTAGGGATC GCTCCGAGTC CGCAAACCTC
8641 TGTGCGGTGG CGGGGCTTGT TCGAGCTTGA AGAGCCTCTA GAGTCGACCT GCAGGCATGC
8701 AAGCTTGGCG TAATCATGGT CATAGCTGTT TCCTGTGTGA AATTGTTATC CGCTCACAAT
8761 TCCACACAAC ATACGAGCCG GAAGCATAAA GTGTAAAGCC TGGGGTGCCT AATGAGTGAG
8821 CTAATCACA TTAATTGCGT TGCCTCACT GCCCGCTTTC CAGTCGGGAA ACCTGTCTGT
8881 CCAGCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTCGCTA TTGGGCGCTC
8941 TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCTGT CGGCTGCGGC GAGCGGTATC
9001 AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA
9061 CATGTAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCTT TGCTGGCGTT
9121 TTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG
9181 GCGAAACCCG ACAGGACTAT AAAGATACCA GCGCTTTCCC CCTGGAAGCT CCCTCGTGCG
9241 CTCTCTGTG CCGACCTGCG CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCCGGGAG
9301 CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTTCGCTC

9361 CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCCT TATCCGGTAA
 9421 CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG C

SEQ ID NO: 241
 pSZ1493

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
 61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG GACAGTATTT GGTATCTGCG
 121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAA
 181 CCACCGCTGG TAGCGGTGGT TTTTGTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
 241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACT
 301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
 361 ATTAAAAATG AAGTTTTTAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
 421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
 481 TTGCCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
 541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAAC
 601 AGCCAGCCGG AAGGGCCGAG CGCAGAAAGT GTCTTGCAAC TTTATCCGCC TCCATCCAGT
 661 CTATTAATGG TTGCCGGGAA GTAGAGTAA GTAGTTCCGC AGTTAATAGT TTGCGCAACG
 721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
 781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
 841 TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCGAGT TTATCACTCA
 901 TGGTTATGGC AGCACTGCAT AATCTCTTCT CTGTCATGCC ATCCGTAGA TGCTTTTCTG
 961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAAATAGT TATGCGCGCA CCGAGTTGCT
 1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
 1081 TCATTGGAAA ACGTCTCTCG GGGCGAAAAA TCTCAAGGAT CTTACCGCTG TTGAGATCCA
 1141 GTTCGATGTA ACCCACTCGT GCACCCAACAT GATCTTCAGC ATCTTTTACT TTCACGCGG
 1201 TTTCTGGGTG AGCAAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC
 1261 GGAAATGTTG AATACTCATA CTCTCTCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
 1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
 1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
 1441 TAACCTATAA AAATAGGCGT ATCAGCAGGC CCTTTCGCTC CGCGCGTTTC GGTGATGACG
 1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGCTG TAAGCGGATG
 1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CCGGGCTGGC
 1621 TTAATATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
 1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTC AGGCTGCGCA
 1741 ACTGTTGGGA AGGGCGATCG GTGCGGCCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
 1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
 1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCCGCTGG AGCTGGTGCA GAGCATGGGG
 1921 CAGTTTGGCG AGGAGAGGGT GCTCCCGCTG CTGCACCCCG TGGACAAGCT GTGGCAGCCG
 1981 CAGGACTTCC TGCCCGACCC CGAGTCGCC GACTTCGAGG ACCAGGTGGC CGGGCTGCGC
 2041 GCGCGCGCCA AGGACCTGCC CGACGAGTAC TTTGTGGTGC TGGTGGGCGA CATGATACG
 2101 GAGGAGGCGC TGCCGACCTA CATGGCCATG CTCAACACCT TGGACGGTGT GCGCGACGAC
 2161 ACGGGCGCGG CTGACCACCC GTGGGCGCGC TGGACGCGGC AGTGGGTGGC CGAGGAGAAC
 2221 CGGCACGGCG ACCTGCTGAA CAAGTACTGT TGGCTGACGG GCGCGCTCAA CATGCGGGCC
 2281 GTGGAGGTGA CCAATCAACA CCTGATCAAG AGCGGCATGA ACCCGCAGAC GGACAACAAC
 2341 CCTTACTTGG GCTTCGTCTA CACCTCCTTC CAGGAGCGCG CCACCAAGTA GGTACCCCTT
 2401 CTGCGCTAT GACACTTCCA GCAAAAGGTA GGGCGGGCTG CGAGACGGCT ATCGGCGCT
 2461 GCATGCAACA CCGATGATGC TTCGACCCCC CGAAGCTCCT TCGGGGCTGC TGGGCGCTC
 2521 CGATGCCGCT CCAGGGCGAG CGCTGTTTAA ATAGCCAGGC CCCCATTGCA AAAGACATTA
 2581 TAGCGAGCTA CCAAAGCCAT ATTCAAACAC CTAGATCACT ACCACTTCTA CACAGGCCAC
 2641 TCGAGCTTGT GATCGCACTC CGCTAAGGGG GCGCCTCTTC CTCTTCGTTT CAGTCACAAC
 2701 CCGCAAACCT TAGAATATCA ATGCTGCTGC AGGCCTTCCT GTTCTGCTG GCCGCTTCG
 2761 CCGCCAAGAT CAGCGCCTCC ATGACGAACG AGACGTCCGA CCGCCCCCTG GTGCACTTCA
 2821 CCCCCAACAA GGGCTGGATG AACGACCCCA ACGGCCTGTG GTACGACGAG AAGGACGCCA
 2881 AGTGCGACCT GTACTTCCAG TACAACCCGA ACGACACCGT CTGGGGGCTG CCCTTGTTCT
 2941 GGGGCCACGC CACGTCCGAC GACCTGACCA ACTGGGAGGA CCAGCCCATC GCCATCGCCC
 3001 CGAAGCGCAA CGACTCCGGC GCCTTCTCCG GCTCCATGGT GGTGGACTAC AACAACACCT
 3061 CCGGCTTCTT CAACGACACC ATCGACCCGC GCCAGCGCTG CGTGGCCATC TGGACCTACA
 3121 ACACCCCGGA GTCCGAGGAG CAGTACATCT CCTACAGCCT GGACGGCGGC TACACCTTCA
 3181 CCGAGTACCA GAAGAACCCC GTGCTGGCCG CCAACTCCAC CCAGTTCGCG GACCCGAAGG

3241 TCTTCTGGTA CGAGCCCTCC CAGAAGTGGA TCATGACCGC GGCCAAGTCC CAGGACTACA
3301 AGATCGAGAT CTACTCCTCC GACGACCTGA AGTCCTGGAA GCTGGAGTCC GCGTTTCGCCA
3361 ACGAGGGCTT CCTCGGCTAC CAGTACGAGT GCCCCGGCCT GATCGAGGTC CCCACCGAGC
3421 AGGACCCAG CAAGTCTTAC TGGGTGATGT TCATCTCCAT CAACCCCGGC GCCCCGGCCG
3481 GCGGCTCCTT CAACCAGTAC TTCGTGCGCA GCTTCAACGG CACCCACTTC GAGGCCTTCG
3541 ACAACCAGTC CCGCGTGGTG GACTTCGGCA AGGACTACTA CGCCCTGCAG ACCTTCTTCA
3601 ACACCGACCC GACCTACGGG AGCGCCCTGG GCATCGCGTG GGCCTCCAAC TGGGAGTACT
3661 CCGCCTTCGT GCCACCAAC CCTTGGCGCT CCTCCATGTC CCTCGTGC GC AAGTTCTCCC
3721 TCAACACCGA GTACCAGGCC AACCCGAGGA CGGAGCTGAT CAACCTGAAG GCCGAGCCGA
3781 TCCTGAACAT CAGCAACGCC GGCCCTGGGA GCCGGTTCGC CACCAACACC AC GTTGACGA
3841 AGGCCAACAG CTACAACGTC GACCTGTCCA ACAGCACCGG CACCCTGGAG TTCGAGCTGG
3901 GTTACGCCGT CAACACCACC CAGACGATCT CCAAGTCCGT GTTCGCGGAC CTCTCCCTCT
3961 GGTTCAGGG CCTGGAGGAC CCCGAGGAGT ACCTCCGCAT GGGCTTCGAG GTGTCCGCGT
4021 CCTCCTTCTT CCTGGACCGC GGGAACAGCA AGGTGAAGTT CGTGAAGGAG AACCCCTACT
4081 TCACCAACCG CATGAGCGTG AACAACGAGC CCTTCAAGAG CGAGAACGAC CTGTCTTACT
4141 ACAAGGTGTA CGGCTTGCTG GACCAGAACA TCCTGGAGCT GTACTTCAAG GCCGAGCCGA
4201 TCGTGTCCAC CAACACCTAC TTCATGACCA CCGGGAACGC CCTGGGCTCC GTGAACATGA
4261 CGACGGGGGT GGACAACCTG TTCTACATCG ACAAGTTCCA GGTGCGCGAG GTCAAGTGAC
4321 AATTGGCAGC AGCAGCTCGG ATAGTATCGA CACACTCTGG ACGCTGGTCG TGTGATGGAC
4381 TGTTCGCGCC ACCTTGCTG CCTTGACCTG TGAATATCCC TGCCGCTTTT ATCAAACAGC
4441 CTCAGTGTGT TTGATCTTGT GTGTACGCGC TTTTGCAGT TGCTAGCTGC TTGTGCTATT
4501 TGGGAATACC ACCCCAGCA TCCCCTTCCC TCGTTTCATA TCGCTTGCAT CCAAACCGCA
4561 ACTTATCTAC GCTGTCTGC TATCCCTCAG CGCTGCTCCT GCTCCTGCTC ACTGCCCTC
4621 GCACAGCCTT GGTTCGGCT CCGCTGTAT TCTCCTGGTA CTGCAACCTG TAAACCAGCA
4681 CTGCAATGCT GATGCACGGG AAGTAGTGGG ATGGGAACAC AAATGGAGGA TCCGCGCTCT
4741 CGAACAGAGC GCGCAGAGGA ACGCTGAAGG TCTCGCCTCT GTCGCACCTC AGCGCGGCAT
4801 ACACCACAAT AACCACCTGA CGAATGCGCT TGGTTCCTCG TCCATTAGCG AAGCGTCCCG
4861 TTCACACACG TGCCACGTTG GCGAGGTGGC AGGTGACAAT GATCGGTGGA GCTGATGGTC
4921 GAAACGTTCA CAGCCTAGGG ATATCGAATT CGGCCGACAG GACGCGCGTC AAAGGTGCTG
4981 GTCGTGTATG CCCTGGCCGG CAGGTGCTTG CTGCTGCTGG TTAGTGATTC CGCAACCCTG
5041 ATTTTGGCGT CTTATTTTGG CGTGCAAAAC GCTGGCGCCC GCGAGCCGGG CCGCGGCCGA
5101 TGCGGTGCC CACGGCTGCC GGAATCCAAG GGAGGCAAGA GCGCCCGGGT CAGTTGAAGG
5161 GCTTTACGCG CAAGGTACAG CCGCTCCTGC AAGGTGCGT GGTGGAATTG GACGTGCAGG
5221 TCCTGCTGAA GTTCTCCAC CGCTCACC CAACGACAAAG CACCGGTGTA TCAGGTCCGT
5281 GTCATCCACT CTAAAGAACT CGACTACGAC CTACTGATGG CCCTAGATT CTTATCAAAA
5341 ACGCCTGAGA CACTTGCCCA GGATTGAAAC TCCCTGAAGG GACCACCAGG GGCCCTGAGT
5401 GTTTCCTTCC CCGGTGGCG AGCTGCCAGC CAGGCTGTAC CTGTGATCGA GGCTGGCGGG
5461 AAAATAGGCT TCGTGTGCTC AGGTGATGGG AGGTGAGGA CAGCTCAGTA AAGCGGACCA
5521 ATCGCACAA TCAATGCAAG CTAATCAGCT ATTTCTCTT CACGAGCTGT AATTGTCCCA
5581 AAATTCTGGT CTACCGGGG TGATCCTTCG TGTACGGGCC CTTCCCTCAA CCCTAGGTAT
5641 GCGCGCATGC GGTGCGCGCG CAACTCGCGC GAGGGCCGAG GGTTCGGAC GGGCCGTCCC
5701 GAAATGCAGT TGACCCCGGA TCGGTGGCAC CTTTTTTCG ATAATTTATG CAATGGACTG
5761 CTCTGCAAAA TTCTGGCTCT GTCGCCAACC CTAGGATCAG CGGCGTAGGA TTTCGTAATC
5821 ATTCGTCTG ATGGGAGCT ACCGACTACC CTAATATCAG CCCGACTGCC TGACGCCAGC
5881 GTCCACTTTT GTGCACACAT TCCATTCTGT CCAAGACAT TTCATTGTGG TCGGAAGCGT
5941 CCCCAGTTAC GCTCACCTGT TTCCGACCT CTTACTGTG CTGTGACAG AGCGGGCCCA
6001 CAGGCCGGTC GCAGCCACTA GTATGGCCAC CGCATCCACT TTCTCGGCGT TCAATGCCCG
6061 CTGCGGCGAC CTGCTGCTG CCGCGGGCTC CGGGCCCCGG CGCCAGCGA GGGCCCTCCC
6121 CGTGC CGGG CGCGCCCCG ACTGTTCCAT GCTGTTCCG GTGATCACC CAATCTTCAG
6181 CGCCGCCGAG AAGCAGTGGA CCAACCTGGA GTGGAAGCCC AAGCCCAAGC TGCCCCAGCT
6241 GCTGGACGAC CACTTCGGCC TGCACGGCCT GGTGTTCCGC CGCACCTTCG CCATCCGCTC
6301 CTACGAGGTG GGCCCCGACC GCAGCACCTC CATCCTGGCC GTGATGAACC ACATGACGGA
6361 GGCCACCCTG AACCAACGCA AGAGCGTGGG CATCCTGGGC GACGGCTTCG GCACACCCT
6421 GGAGATGTCC AAGCGCGACC TGATGTGGGT GGTGCGCGC ACCCACGTGG CCGTGGAGCG
6481 CTACCCACC TGGGGCGACA CCGTGGAGGT GGAGTGCTGG ATCGGCGCCA GCGGCAACAA
6541 CGGCATGCGC CGCGACTTCC TGGTGC GCGA CTGCAAGACC GGCGAGATCC TGACCCGCTG
6601 CACCTCCCTG AGCGTGCTGA TGAACACCCG CACCCGCGC CTGAGCACC TCCCGACGA
6661 GGTGCGCGGC GAGATCGGCC CCGCTTCAT CGACAACGTG GCCGTGAAGG ACGACGAGAT
6721 CAAGAAGCTG CAGAAGCTGA ACGACTCCAC CGCGACTAC ATCCAGGGCG GCCTGACCCC
6781 CCGCTGGAAC GACCTGGACG TGAACACGCA CGTGAACAAC CTGAAGTACG TGGCTGGGT
6841 GTTCGAGACC GTGCCCGACA GCATCTTCGA GTCCACCAC ATCAGCTCCT TCACCCTGGA

```

6901 GTACCGCCGC GAGTGCACCC GCGACTCCGT GCTGCGCAGC CTGACCACCG TGAGCGGCGG
6961 CAGCTCCGAG GCCGGCCTGG TGTGCGACCA CCTGCTGCAG CTGGAGGGCG GCAGCGAGGT
7021 GCTGCGCGCC CGCACCGAGT GCGCCCCAA GCTGACCGAC TCCTTCCGCG GCATCAGCGT
7081 GATCCCCGCC GAGCCCCGCG TGATGGACTA CAAGGACCAC GACGGCGACT ACAAGGACCA
7141 CGACATCGAC TACAAGGACG ACGACGACAA GTGACTCGAG GCAGCAGCAG CTCGGATAGT
7201 ATCGACACAC TCTGGACGCT GGTCTGTGTA TGGACTGTTG CCGCCACACT TGCTGCCTTG
7261 ACCTGTGAAT ATCCCTGCCG CTTTTATCAA ACAGCCTCAG TGTGTTTGAT CTTGTGTGTA
7321 CGCGCTTTTG CGAGTTGCTA GCTGCTGTG CTATTTGCGA ATACCACCCC CAGCATCCCC
7381 TTCCCTCGTT TCATATCGCT TGCATCCCAA CCGCAACTTA TCTACGCTGT CCTGCTATCC
7441 CTCAGCGCTG CTCTGCTCC TGCTCACTGC CCCTCGCACA GCCTTGTTTT GGGCTCCGCC
7501 TGTATTCTCC TGGTACTGCA ACCTGTAAAC CAGCACTGCA ATGCTGATGC ACGGGAAGTA
7561 GTGGGATGGG AACACAAATG GAAAGCTTGA GCTCCAGCCA CGGCAACACC GCGCGCCTTG
7621 CGGCCGAGCA CGGCGACAAG AACCTGAGCA AGATCTGCGG GCTGATCGCC AGCGACGAGG
7681 GCCGGCACGA GATCGCCTAC ACGCGCATCG TGGACGAGTT CTTCCGCCTC GACCCCGAGG
7741 GCGCGTTCGC CGCCTACGCC AACATGATCG GCAAGCAGAT CACCATGCCG GCGCATCCCC
7801 TGGACGACAT GGGCCACGGC GAGGCCAACC CGGGCCGCAA CCTCTTCGCC GACTTCTCCG
7861 CGGTGCGCGA GAAGATCGAC GTCTACGACG CCGAGGACTA CTGCCGCATC CTGGAGCACC
7921 TCAACGCGCG CTGGAAGGTG GACGAGCGCC AGGTCAGCGG CCAGGCCGCC GCGGACCAGG
7981 AGTACGTCTT GGGCCTGCCC CAGCGCTTCC GGAAACTCGC CGAGAAGACC GCCGCCAAGC
8041 GCAAGCGCGT CGCGCGCAGG CCCGTCGCCCT TCTCCTGGAG AAGAGCCTCT AGAGTCGACC
8101 TGCAGGCATG CAAGCTTGGC GTAATCATG TCATAGCTGT TTCCTGTGTG AAATTGTTAT
8161 CCGTTCACAA TTCCACACAA CATAAGAGCC GGAAGCATAA AGTGTAAAGC CTGGGGTGCC
8221 TAATGAGTGA GCTAACTCAC ATTAATTGCG TTGCGCTCAC TGCCCGCTTT CCAGTCGGGA
8281 AACCTGTCTG GCCAGCTGCA TTAATGAATC GGCCAACGCG CGGGGAGAGG CGGTTTGCCT
8341 ATTGGGCGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCTG TCGGCTGCGG
8401 CGAGCGGTAT CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC
8461 GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA AAAGGCCGCG
8521 TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA
8581 AGTCAGAGGT GCGGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTTGAAGC
8641 TCCTCGTGC GCTCTCCTGT TCCGACCTTG CCGCTTACCG GATACCGTGC CGCCTTTCTC
8701 CCTTCGGGAA GCGTGCGCT TTCTCATAGC TCACGCTGTA GGTATCTCAG TTCGGTGTAG
8761 GTCGTTCTCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA CCGCTGCGCC
8821 TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GC

```

SEQ ID NO: 242
pSZ1321

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAGG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAGAA GAGTTGGTAG CTCTTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAAC
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCCCTCGTG TAGATACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGCTCCTT CGGTCTCTCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGTTATGGC AGCACTGCAT AATTCTCTTA CTGTATGCC ATCCGTAAGA TGCTTTTCTG
961 TGA CTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCCGG GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GGACCCAACT GATCTTCAGC ATCTTTTACT TTCACGACG
1201 TTTCTGGGTG AGCAAAAACT GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGCCAGAC
1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT

```

1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
1441 TAACCTATAA AAATAGGCGT ATCACGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CCGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT CAGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCCGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCCGCTGG AGCTGGTGCA GAGCATGGGG
1921 CAGTTTTCGG AGGAGAGGGT GCTCCCGTG CTGCACCCCG TGGACAAGCT GTGGCAGCCG
1981 CAGGACTTCC TGCCCGACCC CGAGTCGCCC GACTTCGAGG ACCAGGTGGC GGAGCTGCGC
2041 CGCGCGGCCA AGGACCTGCC CGACGAGTAC TTTGTGGTGC TGGTGGGCGA CATGATCAGC
2101 GAGGAGGCGC TGCCGACCTA CATGGCCATG CTCAACACCT TGGACGGTGT GCGCGACGAC
2161 ACGGGCGCGG CTGACCACCC GTGGGCGCGC TGGACGCGGC AGTGGGTGGC CGAGGAGAAC
2221 CGGACGCGG ACCTGCTCAA CAAGTACTGT TGGCTGACGG GCGCGCTCAA CATGCGGGCC
2281 GTGGAGGTGA CCATCAACAA CCTGATCAAG AGCGGCATGA ACCCGCAGAC GGACAACAAC
2341 CCTTACTTGG GCTTCGTCTA CACCTCCTTC CAGGAGCGCG CCACCAAGTA GGTACCCCTT
2401 CTGCGCTAT GACACTTCCA GCAAAGGTA GGGCGGGCTG CGAGACGGCT TCCCGGCGCT
2461 GCATGCAACA CCGATGATGC TTCGACCCCC CGAAGCTCCT TCGGGGCTGC ATGGGCGCTC
2521 CGATGCCGCT CCAGGGCGAG CGCTGTTTAA ATAGCCAGGC CCCCATTGC AAAGACATTA
2581 TAGCGAGCTA CCAAAGCCAT ATTCAAACAC CTAGATCACT ACCACTTCTA CACAGGCCAC
2641 TCGAGCTTGT GATCGCACTC CGCTAAGGG GCGCCTCTTC CTCTTCGTTT CAGTCAACAC
2701 CCGCAAATC TAGAATATCA ATGCTGCTGC AGGCCCTCCT GTTCCTGCTG GCCGGCTTCG
2761 CCGCAAAGAT CAGCGCCTCC ATGACGAACG AGACGTCCGA CCGCCCCCTG GTGCACTTCA
2821 CCCCCAACAA GGGCTGGATG AACGACCCCA ACGGCCTGTG GTACGACGAG AAGGACGCCA
2881 AGTGGCACCT GTACTTCCAG TACAACCCGA ACGACACCGT CTGGGGGACG CCCTTGTTC
2941 GGGGCCACGC CACGTCCGAC GACCTGACCA ACTGGGAGGA CCAGCCCATC GCCATCGCCC
3001 CGAAGCGCAA CGACTCCGGC GCCTTCTCCG GCTCCATGGT GGTGGACTAC AACAAACCT
3061 CCGGCTTCTT CAACGACACC ATCGACCCG GCCAGCGCTG CGTGGCCATC TACACCTACA
3121 ACACCCCGGA GTCCGAGGAG CAGTACATCT CCTACAGCCT GGACGGCGGC TACACCTTCA
3181 CCGAGTACCA GAAGAACCCC GTGCTGGCCG CCAACTCCAC CCAGTTCCGC GACCCGAAGG
3241 TCTTCTGGTA CGAGCCCTCC CAGAAGTGA TCATGACCGC GGCCAAGTCC CAGGACTACA
3301 AGATCGAGAT CTACTCCTCC GACGACCTGA AGTCCTGGAA GCTGGAGTCC GCGTTCGCCA
3361 ACGAGGGCTT CCTCGGCTAC CAGTACGAGT GCCCGGCCCT GATCGAGGTC CCCACCGAGC
3421 AGGACCCAG CAAGTCCTAC TGGGTGATGT TCATCTCCAT CAACCCCGGC GCCCGGCCG
3481 GCGGCTCCTT CAACAGTAC TTCGTGGCA GCTTCAACGG CACCCACTTC GAGGCTTCG
3541 ACACCCGATC CCGCGTGGTG GACTTCGGCA AGGACTACTA CGCCCTTCGAG GTGTCTCTCA
3601 ACACCGACCC GACCTACGGG AGCGCCCTGG GCATCGCGTG GGCCTCCAAC TGGGAGTACT
3661 CCGCCTTCGT GCCCACCAAC CCCTGGCGCT CCTCCATGTC CCTCGTGCGC AAGTTCTCCC
3721 TCAACACCGA GTACCAGGCC AACC CGGAGA CGGAGCTGAT CAACCTGAAG GCCGAGCCGA
3781 TCTGAACAT CAGCAACGCC GGCCCTTGA GCGGTTTCGC CACCAACACC ACGTTGACGA
3841 AGGCCAACAG CTACAACGTC GACCTGTCCA ACAGCACCGG CACCCTGGAG TTCGAGCTGG
3901 TGTACGCCGT CAACACCAAC CAGACGATCT CCAAGTCCGT GTTCGCGGAC CTCTCCCTCT
3961 GGTTCGAAGG CCTGGAGGAC CCGGAGGAGT ACCTCCGCAT GGGCTTCGAG GTGTCCCGCT
4021 CCTCCTTCTT CCTGGACCGC GGGAACAGCA AGGTGAAGTT CGTGAAGGAG AACCCCTACT
4081 TCACCAACCG CATGAGCGTG AACAAACCAGC CCTTCAAGAG CGAGAACGAC CTGTCTACT
4141 ACAAGGTGTA CGGCTTGCTG GACCAGAACA TCCTGGAGCT GTACTTCAAC GACGGCGACG
4201 TCGTGTCCAC CAACACCTAC TTCATGACCA CCGGGAACGC CCTGGGCTCC GTGAACATGA
4261 CGACGGGGGT GGACAACCTG TTCTACATCG ACAAGTTCCA GGTGCGCGAG GTCAAGTGAC
4321 AATTGGCAGC AGCAGCTCGG ATAGTATCGA CACACTCTGG ACGCTGGTCG TGTGATGGAC
4381 TGTGCGGCC ACACCTGCTG CTTGACCTG TGAATATCCC TGCCGCTTTT ATCAACAGC
4441 CTCAGTGTGT TTGATCTTGT GTGTACGCGC TTTTGCAGT TGCTAGCTGC TTGTGCTATT
4501 TGCGAATACC ACCCCCAGCA TCCCCTTCCC TCGTTTCATA TCGCTTGCAT CCAACCGCA
4561 ACTTATCTAC GCTGTCCTGC TATCCCTCAG CGCTGCTCCT GCTCCTGCTC ACTGCCCTC
4621 GCACAGCCTT GGTTCGGGCT CCGCCTGTAT TCTCCTGGTA CTGCAACCTG TAAACCAGCA
4681 CTGCAATGCT GATGCACGGG AAGTAGTGGG ATGGGAACAC AAATGGAGGA TCCCGCGTCT
4741 CGAACAGAGC GCGCAGAGGA ACGCTGAAGG TCTCGCCTCT GTGCGACCTC AGCGCGGCAT
4801 ACACCACAAT AACCACCTGA CGAATGCGCT TGGTTCTTCG TCCATTAGCG AAGCGTCCGG
4861 TTCACACAG TGCCACGTTG CCGAGGTGGC AGGTGACAAAT GATCGGTGGA GCTGATGGTC
4921 GAAACGTTCA CAGCCTAGGG ATATCGAATT CGGCCGACAG GACGCGCGTC AAAGGTGCTG
4981 GTCGTGTATG CCCTGGCCGG CAGGTCGTTG CTGCTGCTGG TTAGTGATT CCAACCCCTG

5041 ATTTTGGCGT CTTATTTTGG CGTGGCAAAC GCTGGCGCCC GCGAGCCGGG CCGGCGGCGA
5101 TGCGGTGCCC CACGGCTGCC GGAATCCAAG GGAGGCAAGA GCGCCCGGGT CAGTTGAAGG
5161 GCTTTACGCG CAAGGTACAG CCGCTCCTGC AAGGCTGCGT GGTGGAATTG GACGTGCAGG
5221 TCCTGCTGAA GTTCTCCAC CGCCTACCA GCGGACAAAG CACCGGTGTA TCAGGTCCGT
5281 GTCATCCACT CTAAAGAACT CGACTACGAC CTACTGATGG CCCTAGATTG TTCATCAAAA
5341 ACGCTGAGA CACTTGCCCA GGATTGAAAC TCCCTGAAGG GACCACCAGG GGCCCTGAGT
5401 TGTTCCTTCC CCCCCTGGCG AGCTGCCAGC CAGGCTGTAC CTGTGATCGA GGCTGGCGGG
5461 AAAATAGGCT TCGTGTGCTC AGGTTCATGGG AGGTGCAGGA CAGCTCATGA AACGCCAACA
5521 ATCGCACAAAT TCATGTCAAG CTAATCAGCT ATTTCTCTT CACGAGCTGT AATTGTCCCA
5581 AAATCTGGT CTACCGGGGG TGATCCTTCG TGTACGGGCC CTTCCCTCAA CCCTAGGTAT
5641 GCGCGCATGC GGTCCCGCG CAACTCGCGC GAGGGCCGAG GGTCTGGGAC GGGCCGTCCC
5701 GAAATGCAGT TGCACCCGGA TCGTGGCAC CTTTTTTGCG ATAATTTATG CAATGGACTG
5761 CTCTGCAAAA TTCTGGCTCT GTCCCAACC CTAGGATCAG CCGCGTAGGA TTTCTGAATC
5821 ATTGGTGGAG ATGGGGAGCT ACCGACTACC CTAATATCAG CCCGACTGCC CCGGCTGAGT
5881 GTCCACTTTT GTGCACACAT TCCATTCTGT CCAAGACAT TTCATTGTGG TGCGAAGCGT
5941 CCCCAGTTAC GCTCACCTGT TTCCCGACCT CTTACTGTT CTGTGACAG AGCGGGCCCA
6001 CAGGCCGGTC GCAGCCACTA GTATGGCCAC CGCATCCACT TTCTCGGCGT TCAATGCCCG
6061 CTGCGCGAC CTGCGTCGCT CGGCGGCTC CGGGCCCGG CGCCAGCGA GGCCCTCCC
6121 CGTGGCGGG CGCGCCAGC TGCCCGACTG GAGCATGCTG CTGGCCGCGA TCACCACCCT
6181 GTTCTGGCG GCCGAGAAGC AGTGGATGAT GCTGGACTGG AAGCCCAAGC GCCCGCAT
6241 GCTGGTGGAC CCCTTCGGCC TGGGCGCTT CGTGCAGGAC GGCCTGTGTT TCCGCAACAA
6301 CTTAGCATC CGCAGCTACG AGATCGGCGC GGACCGCACC GCCAGCATCG AGACCTGAT
6361 GAACCACCTG CAGGAGACCG CCCTGAACCA CGTGAAGAGT GTGGGCTGCG TGGAGGACGG
6421 CCTGGGCAGC ACCGCGAGA TGAGCTGCG CAACCTGATC TGGGTGGTGA CCAAGATGCA
6481 GGTGGCGGTG GACCGCTACC CCACCTGGGG CGACGAGGTG CAGGTGAGCA GCTGGGCGAC
6541 CGCCATCGGC AAGAACGGCA TGCGCCGCGA GTGGATCGTG ACCGACTTCC GCACCGGCGA
6601 GACCCTGCTG CGCGCCACCA GCGTGTGGGT GATGATGAAC AAGCTGACCC GCCGCATCAG
6661 CAAGATCCCC GAGGAGGTGT GGCACGAGAT CGGCCCCAGC TTCATCGACG CGCCCCCCT
6721 GCCCACCCTG GAGGACGACG CCGCAAGCT GACCCGCTT CACGAGACGA GCGCCGACTT
6781 CATCCGCAAG GGCTTGACCC CCCGCTGGAG CGACCTGGAC ATCAACCAGC ACGTGAACAA
6841 CGTGAAGTAC ATCGGCTGGC TGCTGGAGAG CGCGCCCGC GAGATCCACG AGAGCCACGA
6901 GATCGCCAGC CTGACCCTGG AGTACCGCGC CGAGTGCAGC CGCGACAGCG TGCTGAACAG
6961 CGCCACCAAG GTGAGCGACA GCAGCCAGCT GGGCAAGAGC GCCGTGGAGT GCAACCACCT
7021 GGTGCGCCTG CAGAACGGCG GCGAGATCGT GAAGGGCCGC ACCGTGTGGC GCCCAAGCG
7081 CCCCCTGTAC AACGACGGCG CCGTGGTGA CGTGCCCGC AAGACCAGT GACTCGAGGC
7141 AGCAGCAGCT CGGATAGTAT CGACACATC TGGACGCTGG TCGTGTGATG GACTGTGCT
7201 GCCCACTTGT CTGCCTTGAC CTGTGAATAT CCCTGCCGCT TTTATCAAAC AGCCTCAGTG
7261 TGTTTGATCT TGTGTGTACG CGCTTTTGGC AGTTGCTAGC TGCTTGTGCT ATTTGCGAAT
7321 ACCACCCCA GCATCCCTT CCCTCGTTTC ATATCGCTTG CATCCCAACC GCAACTTATC
7381 TACGCTGTCC TGCTATCCCT CAGCGCTGCT CCTGCTCCTG CTCACTGCCC CTCGCACAGC
7441 CTTGGTTTGG GCTCCGCTG TATTCTCCTG TACTGCAAC CTGTAAACCA GCACTGCAAT
7501 GCTGATGCAC GGAAGTAGT GGGATGGAA CACAAATGGA AAGCTTGAGC TCCAGCCAGC
7561 GCAACACCGC GCGCCTTGGC GCGGAGCAGC GCGACAAGAA CTTGAGCAAG ATCTGCGGGC
7621 TGATCGCCAG CGACGAGGGC CGGCACGAGA TCGCCTACAC GCGCATCGTG GACGAGTTCT
7681 TCCGCTCGA CCCCAGGGGC GCGTCCGCG CCTACGCCAA CATGATGCGC AAGCAGATCA
7741 CCATGCCCGC GCACCTCATG GACGACATGG GCCACGGCGA GGCCAACCCG GGCCGCAACC
7801 TCTTCGCCGA CTTCTCCGCG GTCGCCGAGA AGATCGACGT CTACGACGCC GAGGACTACT
7861 GCCGCATCCT GGAGCACCTC AACGCGCGCT GGAAGGTGGA CGAGCGCCAG GTCAGCGGCC
7921 AGGCCGCCGC GGACCAGGAG TACGTCTTGG GCCTGCCCA GCGCTTCCGG AAACCTCGCCG
7981 AGAAGACCGC CGCCAAGCGC AAGCGGTGCG CGCGCAGGCC CGTCGCCTT TCCTGGAGAA
8041 GAGCCTCTAG AGTCGACCTG CAGGCATGCA AGCTTGGCGT AATCATGGTC ATAGCTGTTT
8101 CCTGTGTGAA ATTGTTATCC GCTCACAATT CCACACAACA TACGAGCCGG AAGCATAAAG
8161 TGTAAGCCT GGGGTGCCTA ATGAGTGAGC TAACTACAT TAATTGCGTT GCGCTCACTG
8221 CCCGCTTTC AGTCGGGAAA CCTGTCGTGC CAGCTGCATT AATGAATCGG CCAACGCGCG
8281 GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCT CGCTCACTGA CTCGCTGCGC
8341 TCGGTCGTTT GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTTATCC
8401 ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG
8461 AACCCTAAAA AGGCCGCTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCT TGACGAGCAT
8521 CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG
8581 GCGTTTCCCC CTGGAAGCTC CCTCGTGCAG TCTCCTGTTT CGACCCTGCC GCTTACCGGA
8641 TACCTGTCCG CCTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG

SEQ ID NO: 243

8701 TATCTCAGTT CGGTGTAAGT CGTTGCTCC AAGCTGGGCT GTGTGACAGA ACCCCCCGTT
8761 CAGCCCCGACC GCTGCGGCCTT ATCCGGGTAC TATCGTCTTG AGTCCAAACC GGTAAGACAC
8821 GACTTATCCG

[illegible]

C. hookeriana KAS IV enzyme construct

SEQ ID NO: 245

acgacgacgaagTGA

[illegible]

coding region for *C. hookeriana* C10:0 specific thioesterase (codon-optimized)

SEQ ID NO: 244

[illegible]

cctcgtttcataatcgcttgcatcccaaccgcaacttatctacgtgtcctgctatccctcagcgctgctcctgctcctgctcactgcccct
cgacagccttggttgggctccgctgtattctcctggtagtgaacctgtaaaccagcactgcaatgctgatgcacgggaagtagt
gggatgggaacacaaatggaaagcttgagctc

SEQ ID NO: 246

coding region for *C. hookeriana* KAS IV enzyme (codon-optimized)

ATGgcgaccgccagctgcatggtggcgctgcccttctgcacctggctggcgccgctgcatgccacctccagcgacaacga
ccccgcctgctgtcccaagaagcgctgagcgccgcccgcacacctgagctgcactgctcctgcgggcagcacc
ttccagtgcctggacccctgcaaccagcagcgttctggggcgacaacggcttcgctgctgttccggctccaagccccgcgca
gcaaccgcggccacctggcgctggcgccacctgcactccggcgaggtgatggccgctgcgatgcagccccagggaggt
gagcaccacaagaagcccgcgaccaagcagcgccgctggctgtagccggcatggcgctgacccccctgggccagca
ccccgacgtgtattataacaacctgctggacggcatctcgggcatctccgagatcgagaacttcgactgcagccagttcccccac
ccgcatcgccggcgagatcaagtcgttccaccgacggctgggtcgcgcccaagttcagcgagcgcatggacaagttcatgc
tgtatatgctgaccgcccgaagaaggcgctggccgacggcgccatcacccgaggacgcgatgaaggagctgaacaagcgc
aagtgcggcgctgctgatcggtcggcgctggcgccatgaaggcttctccgacagcatcgaggccctgcgcacctcgtataa
gaagatctcccccttctgctgcccctcagcaccaccaacatgggctcgccgatcctggcgatggacctgggctggatgggccc
caactattccatcagcaccgcgtgcgcccacctgaacttctgcactcctgaacggcgcccaaccacatcatcaagggcgaggcgg
acatgatgctgtgcgggcgctccgacgccggtgctgcccgtcgccctggcggttcgtggcctgcgcgctgagccag
cgcaacaacgacccccaccaaggcctcgcgccctgggactccaaccgcgacggcttcgtcatggcgagggcgcgggcggtgc
tgtgctggaggagctggagcagccaagaagcgcgccgacccatctatgccgagttcctggcgagcttcacctgcga
cgcgatcacatgaccgagccccccccgagggcgccggtcatcctgtgcatcgagaaggcgctggcccagtcgggcgtgt
cccgcgaggacgtgaactatatcaacgcgcacgccaccagcaccgccggcgacatcaaggagtatcaggccctggcgc
actgtctcgccagaaactcggagctgcgctgaactccaccaagagcatgatcgccacctgtggcgcgccggcgcggt
ggaggcggtcgccgtggtcaggcgatccgcacggctggatccacccaacatcaacctggaggacccccgacgagggcgt
ggacgccaagctgctggtcgccccaagaaggagaagctgaaggtagaggtcgccctgtcgaactccttcggcttcggcggg
cacaacagctcgatcctgttcgccccctgcaac**TGA**

SEQ ID NO: 247

C. hookeriana KAS IV enzyme

MATASCMVASPFWLVAACMPTSSDNDPRSLSHKRLRLSRRRRTLSSHCSLRGSTFQCL
DPCNQQRFLGDNFGASLFGSKPLRSNRGHLRLGRTSHSGEVMAMVAMQPAQEVSTNKKPAT
KQRRVVVTGMGVVTPLGHPDVYYNNLLDGISGISEIENFDCSQFPTRIAGEIKSFSTDG
WVAPKFSERMDKFMLYMLTAGKKALADGGITEDAMKELNKRKCGVLIGSGLGGMKVFSDS
IEALRTSYKKISPFVFPSTTNMGSAILAMD LGWMGPNYSISTACATSNFCILNAANHII
KGEADMMLCGGSDAAVLVGLGGFVACRALSRNNDPTKASRPWDSNRDGFVMGEGAGVL
LLEELEHAKKRGATIAEFLGGSFTCDAYHMTPEHPEGAGVILCIEKALAQSGVSREDVN
YINAHATSTPAGDIKEYQALAHCFGQNSELRVNSTKSMIGHLLGGAGGVEAVAVVQAIRT
GWIHPNINLEDPDDEGVDAKLLVGPKKEKLKVKVGLSNSFGFGGHNSILFAPCN

SEQ ID NO: 248

C. hookeriana C10:0 specific thioesterase plus *C. hookeriana* KAS IV enzyme construct

ggtagccgcctgcaacgcaaggcgagccaagccgctccacccgcgctgaaccgacacgtgcttggcgccctgccgctgcct
gccgcatgctgtgctggtaggctgggagctgctgcatgctgattgaggcttggttcacgggtggaagcttatgtgtgctgg
gcttgcagtcgggcaatgcgcatggtggcaagagggcgccagcacttgcctggagctgcgcggtgcctccagggtgttcaatcg
cggcagccagagggatttcagatgatcgcgctacaggttagcagcagtgtagcagaaaggttagcagtttccagaatgatcggt

tcagctgttaatcaatgccagcaagagaaggggtcaagtgcacacgggcatgccacagcacgggcacggggagtggaatg
gcaccaccaagtgtgtgagccagcatcgccgttggtgtttcagctacaacggcaggagtcaccaacgtaacatgagctg
atcaacactgcaatcatcgggcggtgtatgcaagcatgcttggcgaagacacatgggtgtgaggtgtgctgctgctg
ctgcgcacgcccgttgagttggcagcaggctcagccatgactggatggcagctgggctgccaactgcaatgtggtggaataggatgc
aagtggagcgaataccaaacccctctggctgcttgggttgcatggcatgcacatcagcaggagcgcatgcaagggaactgg
ccccatgcacgcatgccaaacgggagcgaccgagtgccacactgtcaccaggcccgcaagctttgcagaacatgctcatgg
acgcatgtagcgtgacgtcccttgacggcgctcctctgggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcgca
gcagagcggcgagcagcggcgggggccaccttcttgcggggctcgccccagccagcgggtgatgctgatcccaacgag
ttcacattcatttgatgcctggagaagcgaggctggggcctttgggtggtgcagccgcaatggaatgcgggacggccaggcta
gcagcaaaggcgctccctactccgcatcgatgttccatagtgcattggactgcatttgggtggggcgggcggtgtttcttctg
ttgcaaaacgcgcatcagcaacctgtcccgtgggtccccgtgcccgtgaaatcgtgtgcacgcccagtcagctgattgccgg
ctgcgaagtaggcgcctcttctgctgcctctctccgtcccgtctagatctagaatatcaATGatcgagcaggacggcc
tccacgcgggtccccgcgctgggtggagcgctgttcggtacgactgggcccagcagaccatcggtgctccgacgccc
ccgtgttccgctgtccgcccaggcgcccgctgtgttctgaagaccgacctgtccggcgccctgaacgagctgcaggacg
agggcgcccgctgtcctggtgcccaccacggcgctgcccgtcgccgctgtgagctggtgaccgagggcgccgagc
tggctgctgctggcgaggtgcccggccaggacctgctgtcctcccactgcccccgccgagaaggtgtccatcatgcccagc
gcatgcccgcctgcacacctggaccccgccacctgcccctcgaccaccaggccaagcaccgcatcgagcgcccgcacc
cgcatggaggccgctggtggaccaggacgacctggacgaggagcaccaggcctggcccccgagctgttccgcccgc
ctgaaggcccgatgccgacggcgaggacctggtggtgaccacggcgacgctgctgcccacatcatggtggagAAC
ggccgcttctccgcttcatcgactgcggccgctgggctggtggcgaccgtaccaggacatcgccctggccaccgcgacatc
gcccaggagctgggcgaggtggccgaccgcttctggtgtgtacggcatcgccgccccgactcccagcgcatcgctt
ctaccgctgctggagcaggttcttcTGAcaattggcagcagcagctcgatagatgcacacactctggacgctggtgctgtgatg
gactgttccgcccacactgtgctccttgacctgtgaatatccctgcccgtttatcaaacagcctcagtggttgaatctgtgtgacg
cgctttgcgagttgctagctgctgtgtctatttgcgaataccacccccagcatcccttccctcgtttcataatcgcttgcatccaacc
gcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcactgcccctgcacagccttgggttgggtccgctg
tattctcctggtactgcaacctgtaaacagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaggatcc
cgctctcgaacagagcgcgagaggaacgctgaaggtctcgctctgtgcacctcagcgcgccataaccacaataaccacctg
acgaatgcgcttgggttcttctgctcattagcgaagcgtccggttcacacacgtgccacgttggcgaggtggcaggtgacaatgatcgg
tggagctgatggtcgaacgttcacagcctaggaggtttaggtccagcgctccgtggggggggagcgggctgggagcttggccggg
aaggggcaagacgatgcagtcctcttggggagtcacagccgactgtgtgtgtgactgtgcggcccgagcactcacagcAAAA
gcctggccgacagggcaggccctgtccagtgcacatccacggctcctctcatcaggctcaccttgcctattgacataacggaatgcg
taccgctcttccagatctgtccatccagaggggagcaggctccccaccgacgctgtcaaaactgcttcttcccacccgaaaca
ttattgttggggggggggggggggggcagattgcattggcgggatctcgtgagggaacatcactgggacactgtggaacacagt
gagtgagatgcagagcatgtatgctaggggtcagcgaggaagggggcctttccagctctccatgccactgcacgatatccag
actaccaggaccagcttctgatcggttccgctccctggacaccagtggtgtagcctctgactccaggatgctgtgcaccgcaaa
ggccagccgatcgtccgattcttgggtggaggatagagtcagccaacttggggctcagagtgacactggggcacgatacgaa
acaacatctacacgctgtcctcatgctgacacaccacagcttgcctccacctgaatgtggcgcatggggccgaatcacagccaat
gtcgtgctgcccataatgtgatccagaccctctccgcccagatgccgagcgatcgtggggcgctgaatagattctgtttcgtact
gtttgggtccttcttcttctcgtcgtcggtctcgaacaggctgcgtcggttctcggtatccctttgtcctcctcgtcaccatcc
tgcgcggggcaagttgcttgcacctgggtggtaccagggttggagggtattaccgctcaggccattccagcccggattcaatt
caaagtctggccaccacccctccgctctgtctgatcactccacattcgtgcatacactacgttcaagtctgatccaggcgtgtct
cgggacaaggtgtgctttagtttgaatctcaaggacccactccagcacagctgctggttgaccccgccctcgaacatATGgcgac
cgccagctgcatggtggtgctgccccttctgcacctgggtgtcgccgctgcatgcccacctccagcgacaacgacccccgctcg
ctgtcccaaacgctgctgagccgcccgcgcccagccctgagctcgactgctcctgagggcagcacttccagtgc
tggacccctgcaaccagcagcgttctggggcgaacggcttcgctgctgttccgctccaagccctgcgagcaacccg
gccacctgctggtggcgacactcgactccggcgaggtgatggcgctgcgatgcagcccggaggaggtgagcaca
caagaagcccgcgaccaagcagcggcggtggtgacggcatggcgctgtagccccctggggcacgaccccgacgtg

desaturase
coding region for *C. lanceolata* C10:0 specific thioesterase with UTBX 1435 A12 fatty acid
SEQ ID NO: 249

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 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 148
--

[illegible]

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. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 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. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 841. 842. 843. 844. 845. 846. 847. 848. 849. 850. 851. 852. 853. 854. 855. 856. 857. 858. 859. 860. 861. 862. 863. 864. 865. 866. 867. 868. 869. 870. 871. 872. 873. 874. 875. 876. 877. 878. 879. 880. 881. 882. 883. 884. 885. 886. 887. 888. 889. 890. 891. 892. 893. 894. 895. 896. 897. 898. 899. 900. 901. 902. 903. 904. 905. 906. 907. 908. 909. 910. 911. 912. 913. 914. 915. 916. 917. 918. 919. 920. 921. 922. 923. 924. 925. 926. 927. 928. 929. 930. 931. 932. 933. 934. 935. 936. 937. 938. 939. 940. 941. 942. 943. 944. 945. 946. 947. 948. 949. 950. 951. 952. 953. 954. 955. 956. 957. 958. 959. 960. 961. 962. 963. 964. 965. 966. 967. 968. 969. 970. 971. 972. 973. 974. 975. 976. 977. 978. 979. 980. 981. 982. 983. 984. 985. 986. 987. 988. 989. 990. 991. 992. 993. 994. 995. 996. 997. 998. 999. 1000.

SEQ ID NO: 253

[illegible]

SEQ ID NO: 254

[illegible]

SEQ ID NO: 256

[illegible]

SEQ ID NO: 255

[illegible]

caaccagtcgccggtggtggaacttcggaaggactactacgccctgcagaccttctcaacaccgacccgacctacgggagcg
 cctgggcatcgctgggctccaactgggagtagtccgcttgcgcccaccaacccctggcgctcctcatgtccctcgctgcg
 aagttctccctcaacaccgagtagtaccaggccaacccggagacggagctgatcaacctgaaggccgagccgagctctgaacatca
 gcaacgccggccctggagccggttcgccaaccaacaccacgttgacgaaggccaacagctacaacgtcgacctgtccaacag
 caccggcaccttgaggttcgagctggttacgccgtcaacaccacccagacgatctccaagtccgtgttcgaggacctctccctc
 tggttcaaggccctggaggaccccgaggagtagtcccgcatgggcttcgaggtgtccgctcctcttcttgagccgagg
 aacagcaaggtgaagtctgtgaaggagaacccctacttccaacccgcatgagcggtgaacaaccagccctcaagagcgag
 aacgacctgtcctactacaaggtgtacggcttgcggaccagaacatcctggagctgtactcaacgacggcgacgtcgtgtcc
 accaaccctacttcatgaccacggggaacgcctgggctccgtgaacatgacgacgggggtggacaacctgttcatatcga
 caagttccaggtgcgcgaggtcaag**TGA**caattggcagcagcagctcgatagtagtacacactctggacgctggtcgtgtga
 tggactgttgcgccacacttgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtggtgttgatctgtgtgtac
 gcgcttttgcgagttgtagctgcttgcctatttgcgaataccaccccgacatcccttccctcgtttcatatcgcttgcatccaacc
 gcaacttatctacgctgtcctgtatccctcagcgctgctcctgctcctgctcactgcccccgacagccttggttgggctccgctgt
 attctcctggtagtcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggagga**tccc**
 gcgtctgaacagagcgcgaggaacgctgaaggtctgccttgcgcacctcagcgcgacatacaccacaataaacacctga
 cgaatgcgcttggttcttctgctcattagcgaagcgtccggttcacacgctgccacgttggcgaggtggcaggtgacaatgatcggt
 ggagctgatggtcgaacgttcacagcctaggatata**gaattc**cttcttgcgctatgacacttcagcaadaaggtaggcgagg
 ctgcgagacggcttccggcgctgcatgcaacaccgatgatgttcgacccccgaagctcctcggggctgcatggcgctcc
 gatgccgctccagggcgagcgctgtttaaatagccaggcccccgattgcaagacattatagcgagctaccaagccatattc
 aaacacctagatcactaccacttctacacaggccactcgagcttgtgatcgactccgctaaggggcgcccttctccttctggtt
 cagtcacaacccgcaaac**actagtATG**gccaccgcatccactttctcggcggtcaatgcccgtgcggcgacctgcgtgcctc
 ggcgggctccgggccccggcgccagcgaggccccctcccgctgcgcggcgccgaggtgcacgtgcaggtgacccactc
 cctggccccgagaagcgcgagatcttcaactcctgaacaactgggcccagggagaacatcctggtgctgtgaaggacgt
 ggacaagtgcgtggcagccctccgacttctgcccgaactccgctccgagggcttcgacgagcaggtgatggagctgcgcaa
 gcgctgcaaggagatccccgacgactacttcatcgtgctggtgggcgacatgatcaccgaggaggccctgccacctacca
 gacctgtgaacacctggacggcgctgcgcgacgagaccggcgctcctgacccctgggcatctggacccgagcctgg
 accgcccagggagaaccgccaaggcgacctgtgaacaagtacctgtacctgtccggccgctggacatgaagcagatcga
 gaagaccatccagtagctgatcggtccggcatggacccccgacccgagaaacacccctacctgggcttcatctacacctct
 tccaggagcgcgccaccttcatctcccacggcaacaccgcccgcctggccaaggagcacggcgacctgaagctggcccaga
 tctgcggcatcatcgccgacgagagaagcgccacgagaccgctacaccaagatcgtggagaagctgttcgagatcgac
 cccgacggcacctgtgtggcctggccgacatgatgcgaagaaggtgtccatgcccgccacctgatgtacgacggccag
 gacgacaacctgttcgagaacttctcctcgtggcccagcgctggcggtgtacacccgaaggactacggcgacatcctgg
 agttcctggtgggcccgtgggacatcgagaagctgacccgctgtccggcgagggccgcaaggccaggactacgtgtgc
 accctgccccccgcatccgcccgtggaggagcgcccagtcggcggtgaagaaggcctccgccaccccttctcctggtatc
 ttcggccgagatcaacctgatggactacaaggaccagacggcgactacaaggaccagacatcgactacaaggacg
 acgacgacaag**TGA**atcgatagatctcttaaggcagcagcagctcgatagtagtacacactctggacgctggtcgtgtgatg
 gactgttgcgccacacttgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtggtgttgatcttgtgtgtacg
 cgcttttgcgagttgtagctgcttgtgctatttgcgaataccaccccgacatcccttccctcgtttcatatcgcttgcatcccaaccg
 caacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgcccctcgacagccttggttgggctccgctgta
 ttctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagcttaa
 ttaag**agctctt**tttccagaaggagttgctccttgagcctttcattctcagcctcgataacctccaaagccgctctaattgtggag
 ggggttgaatttaaagcttgaatgttggtctgctgctggaacaagcccagacttgtgctcactgggaaaaggacatcag
 ctcaaaaaaacttgcgctcaaacccgctaccttgccttgcgcaatctgcccgttgaaatgccaccaattcatattgtgacgc
 ttgacgactgttaattgcctcagaatgtggaatcatctgccccctgtgcgagccatgccaggcatgtcgggcgaggacccc
 gccactcgtacgacagaccattatgtacctcacaaatagttcataacagtgaacctatttctcgaagctcccaacgagcacctcc
 atgctctgagtgccacccccggccctgtgcttgcggaggcgaggtcaacccggcatggggctaccgaaatccccgacgggac
 ccaccacccccgcatgggaagaatctctcccggtatgtgggcccaccaccagcacaacctgctggccaggcgagcgtcaaac

[illegible]

SEQ ID NO: 258
C. hookeriana C16:0 thioesterase construct

[illegible]

SEQ ID NO: 257
coding region for *O. europaea* stearyl-ACP desaturase (codon-optimized)

[illegible]

[illegible]

SEQ ID NO: 260

[illegible]

SEQ ID NO: 259

[illegible]

[illegible]

SEQ ID NO: 268

SEQ ID NO: 267

SEQ ID NO: 266

[illegible]

[illegible]

atgaaccagcacgtgaacaacgtgacctacatcggtgggtgctggagtcacatccccaggagatcatcgacacccacgag
 ctgcagaccatcacctggactaccgcccgcgagtgccagcacgacgacatcgtggactccctgacctccgtggagccctccg
 agaacctggaggccgtgtccgagctgcgcggcaccaacggctccgccacccaccacgcccggcgacgaggactgccgaac
 ttctgcacctgtgcgcctgtccggcgacggcctggagatcaaccgcccgcacccgagtgggcgcaagaagtccgcccga
 tggactacaaggaccagcggcgactacaaggaccagacatcgactacaaggacgacgacgacaagTGAatcgata
 gatctcttaaggcagcagcagctcggatagtatcgacacactctggacgctggctggtgatggactgttgcgccacacttgctgcc
 ttgacctgtgaatatccctgccgcttttatcaaacagcctcagtggtttgatcttgtgtgtacgcgcttttcgagttgctagctgcttg
 tgctatttgcgaataaccacccccagcatcccttccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgctctgctatc
 cctcagcgctgctcctgctcctgctcactgcccctcgcaagccttgggttgggctccgctgtattctcctggtactgcaacctgtaa
 ccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagcttaattaagagctcttgtttccagaagg
 agttgctccttgagcctttcattctcagcctcgataacctccaaagccgctctaattgtggagggggttcgaatttaaaagcttgga
 atgttgggttcgtgcttggaacaagcccagactgttctcactgggaaaaggaccatcagctccaaaaaacttgccgctcaaac
 cgcttacctctgtttcgcgcaatctgcctgttgaaatcgccaccacattcatattgtgacgcttgagcagctgtaattgcctcag
 aatgtggaatcatctgccccctgtgcgagccatgcccagcatgtcgcggcgaggacacccgacctcgtacagcagaccattat
 gctacctacaatagttcataacagtgaccatatttctgaagctcccaacgagcacctccatgctctgagtgccacccccggc
 cctggtgcttgaggagggcaggtcaaccggcatggggctaccgaaatccccgacggatccacaccccccgatgggaagaa
 tctctcccggtggtggccaccaccagcacaacctgctggccaggcgagctaaaccataccacacaaatatccttggcat
 cggcctgaattccttctgctgctgtaaccggtgcttctgtccgaagcaggggttgctagggatcgtccgagtcgcaaacct
 tgtcgcgtggcgggcttgttcgagcttgaagagc

SEQ ID NO: 269

coding region for *Ricinus communis* ACP-thioesterase (codon-optimized)

ATGgccaccgcatccactttctcggcgttcaatgcccgtgcggcgacctgcgtcgtcggcgggctccgggccccggcgccca
 gcgagggccctccccgtgcgcggcgccgctcctacgaggtggcatcaacaagaccgccacctggagaccatcgcca
 acctgctgaggaggtgggctgcaaccacgcccagtcctggtggttctccaccgacggcttcgccaccaccacctccatgcgca
 agatgcacctgatctgggtgaccgcccgcagtcacatcgagatctacaagtacccgcctggtccgacgtggtggaggtgga
 gacctggtgccagtcggaggccgcacgcgacccgcccgcagctggatcctgaccgactacgccaccggccagatcatcgccc
 gcgccacctcaagtgggtgatgatgaaccaggacacccgcccgcctgcagaaggtgaccgacgacgtgcgcgaggagtacc
 tgggtgttctgccccgcgagctgcgcctggcctccccgaggagaacaaccgctcctccagaagatctcaagctggaggac
 cccgcccagctactcaagctggcctggtgccccgcgcgacccctggacatgaaccagcacgtgaacaacgtgacctacat
 cggctgggtgctggagtcacatccccaggagatcatcgacacccacgagctgcagaccatcacctggactacccgcccgag
 gccagcacgacgacatcgtggactccctgacctccgtggagccctccgagaacctggaggccgtgtccgagctgcgcggcacc
 aacggctccgccaccaccacccgcccgcgacgaggactgcgcaacttctgcacctgtgcgcctgtccggcgacggcctgga
 gatcaaccgcccgcacccgagtgggcgcaagaagtccgcccgcagtgactacaaggaccacgacggcgactacaaggacc
 acgacatcgactacaaggacgacgacgacaagTGA

SEQ ID NO: 270

C. camphora C14:0 thioesterase construct

gctcttcgccgccactcctgctcagcgcgcccgcgctgcgcgccagcgcttggccttttcgccgctcgtgcgcgtcgt
 gatgtccatcacagggtccatgaggtctgccttcgcccggctgagccaactgcttcggcgccgaaggaggagcatgaggag
 gactcctggtccagggtcctgacgtggtcgcggctctggagcgggcccagcatcatctggtctgcgcacccaggccgctccaa
 ctggtcctccagcagccgagtcgcgcggcaccctggcagaggaagacaggtgaggggggtatgaattgtacagaacaaccag
 agccttgtctaggcagaatccctaccagtcaggtttacctggatgacggcctgcgaacagctgtccagcgacctcgtgcgcgc
 gcttctccgcacgcttcttcagcaccgtgatggcgcgagccagcgcgcacgctggcgctgcgcttcggatctgaggacgt
 cggggaaactctgatcagcttaaaaccccttcgcgcttagtgttgcacatccttgcagaccggtgagagccgactgttgtgcgcac
 cccacacacacctcctccagaccaattctgcaccttttggcgaaggcatcgccctcgccctgcagagaggacagcagtgccc

agccgctgggggtggcgatgcacgctcaggtaccctttcttgcgtatgacacttcagcaaaaggtaggcgggctgcgagac
 ggcttccggcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctccgatgccgtcca
 gggcgagcgctgtttaaataagccaggccccgattgcaaagacattatagcgagctaccaagccatattcaaacacctagatcac
 taccacttctacacaggccactcgagcttggatgcactccgctaagggggcgcttctctcttctggttcagtcacaacccgcaaa
 ctctagaatatca**ATG**atcgagcaggacgacctccacgcccgtccccgcgcctgggtggagcgctgttcggctacgact
 gggccagcagacatcggtgctccgacgcccgtgttcgctgctccgcccagggcgccccgtgctgttcgtgaagacgg
 acctgtccggcgccctgaacgagctgcaggacgaggccgcccgtgctgctggctggccaccacggcgctgacctgcgcgccc
 gtgctggagctggtagccgagggccgcccgcgactggctgctgctggcgaggtgcccggccaggacctgctgtcctccacct
 ggccccgcgcgagaaggtgtccatcatggccgacgcatgcccgcctgcacacctggacccccgacctgccccctcgacca
 ccaggccaagcaccgcatcgagcgccccgacccgcatggaggccggcctgggtggaccaggacgacctggacgaggagc
 accagggcctggccccgcgagctgttcgcccgtgaaggccgcatgcccgacggcgaggacctggtggtgacctacgg
 cgacgctgctgcccacatcatggtggagacggccgcttctccggttcacgactgcccgcctggcggtggcgacgg
 ctaccaggacatcgccctggccacccgcgacatcgccgaggagctggcgcgagtgggcgacggcttctggtgctgtacg
 gcatcgccgccccgactcccagcgcatgccttctaccgctgctggacgagttctc**TGA**caattggcagcagcagctcggat
 agtatcgacacactctggacgctggtcgtgtagtggactgttgcgccacactgctgcttgacctgtgaatatccctgcccgtttat
 caaacagcctcagtggtttgatcttgtgtacgcgcttttgcgagttgctagctgcttgcctatttgcgaataccacccccagcatc
 ccttctcctgcttcatatcgcttgcacccaaccgcaacttatctacgctgctcctgctatccctcagcgctgctcctgctcctgctact
 gccccctgcacagccttgggttgggctccgctgtattctctggtactgcaacctgtaaacagcactgcaatgctgtagcacggga
 agtagtgggatgggaacacaaatggagga**atcc**cgctctcgaacagagcgcgagaggaacgctgaaggtctgcctctgtcgca
 cctcagcgcgccataccacaataaccacctgacgaatgcgcttgggttcttctgctcattagcgaagcgctccggttcacacagctgc
 cacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatc**gaattc**ctttcttgc
 gctatgacacttcagcaaaaggtaggcgggcgctgcgagacggcttccggcgctgcatgcaacaccgatgatgcttcgaccccc
 gaagctccttcggggctgcatggcgctccgatgcccgtccagggcgagcgctgtttaaataagccaggccccgattgcaaagaca
 ttatagcgagctaccaaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatcgactcogcta
 agggggcgcttctcttctggttcagtcacaacccgcaaac**actag****ATG**gccaccgcatccactttctcggcgttcaatgcc
 gctgcgagcactgctgcgctcggcggtcctcgggccccgcgcccagcgaggccctccccgtgcgcg**ggcgagcccc**gac
 tgggtccatgctgttcgctggtatcaccacatcttctcggcgccgagaaagcagtggaaccaactggagtggaagccaaagccc
 aacccccccagctgctggacgaccacttcggccccacggcctggtgttcgcccaccttcgcatccgagctacgaggtg
 ggccccgaccgctccaccagcatcggtggcgtgatgaaccactgcaggaggccgcccgaaccacgccaagtcctggtggca
 tcttggcgacggcttcggcaccacccctggagatgtccaagcgcgacctgatctgggtggtgaagcgacccaactggtggcgtg
 gagcgctaccccgctggggcgacaccgtggaggtggagtgctgggtggcgccctccggcaacaacggccgcccacgac
 ttctggtgcgagctgcaagacggcgagatcctgacccgctgcacctccctgagcgtgatgatgaacacccgacccgccc
 cctgagcaagatccccgaggaggtgcgcgcgagatcgccccgcttcacgacaacgtggcgtgaaggacgaggagat
 caagaagccccagaagctgaacgactccaccgcccagctacatccaggcgccgctgacccccgctggaacgacctggacatc
 aaccagcagtgacaacatcaagtacgtggactggatctggagaccgtgcccgaagcatcttcgagagccaccacatct
 cctcttcaccatcgagtaccgcccaggtgcacatggacagcgtgctgcagctccctgacacccgtgagcgggcgtcctccg
 agggccgctggtgtgcgagcacctgctgcagctggaggcgccgagcgaggtgctgcgccaagaccgagtggtgccccca
 agctgaccgactcctccgcgcatcagcgtgatccccgagctcagcgtgatggactacaaggaccacgacggcgactac
 aaggaccacgacatcgactacaaggacgacgacaag**TGActc**gaggcagcagcagctcggatagtatcgacacactc
 ggacgctggtcgtgtagtggactgttgcgccacacttgccttgacctgtgaatatccctgcccgtttatcaaacagcctcagtg
 tgtttagcttctgtgtacgcgcttttgcgagttgctagctgcttgcctatttgcgaataccacccccagcatcccttccctggttcat
 atcgcttgcacccaaccgcaacttatctacgctgctcctgctatccctcagcgctgctcctgctcctgctcactgccccctgcacagcct
 tgggttgggctcgcctgtattctcctggtactgcaacctgtaaacagcactgcaatgctgatgcaggggaagtagtgggatggga
 acacaaatggaaagct**gagctc**ttgtttccagaaggagttgctccttgagcctttcattctcagcctcgataacctccaaagccg
 ctctaattgtggaggggttgaatttaaaagcttgaatgttggctgctgctggaacaagcccagactgttgcctcactggga
 aaaggacctcagctcaaaaaacttgcgctcaaacggctacctctgcttctgcgcaatctgcccgttgaatcgccaccaca

[illegible]

C. campyhora C14:0 specific thioesterase construct

SEQ ID NO: 272

[illegible]

coding region for *C. camphora* C14:0 thioesterase (codon-optimized)

SEQ ID NO: 271

[illegible]

[illegible]

[illegible]

SEQ ID NO: 275
coding region for *U. Americana* C10:0-C16:0 specific thioesterase (codon-optimized)

[illegible]

cagcccgaactgcctgacgccagcgctccacttttgtgcacacattccattcgctgcccaagacatttcatttggtgcaagcgctccca
 gttacgctcacctgtttcccgacctccttactgttctgtgcagagcgggccacaggccggtcgagcc**actagATG**gccacc
 gcatccactttctcggttcaatgccgctgcggcgacctgcgtcgctcgggcggtctcgggccccggcgccagcgaggccc
 ctccccgtgcg**ggcgcgcccc**aaggccaacggcagcgctgagcctgaagtcggcagcctgaacacctggaggacc
 ccccgagcagcccccccccgaccttctgaaccagctgccgactggagccgctgcgcaccgccatcaccacctgttcgt
 ggccgcccagagaagcagttcaccgctggaccgcaagagcaagcgccccgacatgctggtggactggttcggcagcgagac
 catcgtgcaggacggcctggtgttcgcgagcgcttcagcatccgcagctacgagatcggcgcgaccgcaccgccagcatcg
 agacctgatgaaccacctgcaggacaccagcctgaaccactgcaagagcgtggcctgctgaacgacggcttcggccgcac
 ccccgagatgtgcacccgcgacctgatctgggtgctgaccaagatgcagatcgtggtgaaccgctacccacctggggcgaca
 ccgtggagatcaacagctggttcagccagagcggaagatcggtatggccgcgagtggtgatcagcgactgcaacaccg
 gcgagatcctggtgcgcgccaccagcgctgggcatgatgaaccagaagaccgcccgttcagcaagctgcccgcgaggt
 gcgcaggagatgcccccactcgtggacgccccccgtgatcgaggacaacgaccgcaagctgcacaagttcgacgtg
 aagaccggcgacagcatctgcaagggcctgaccccggtggaaacgacttcgacgtgaaccagcacgtgagcaacgtgaag
 tacatcggtggttctggagagcatgccaccgaggtgctggagaccaggagctgtgcagcctgacctggagtaccgcc
 gcgagtgcgccgcgagagcgtggtggagagcgtgaccagcatgaacccagcaaggtgggcgaccgcagccagtacca
 gcacctgctgcgcctggaggacggcgccgacatcatgaagggcgaccgagtggcgccccagaacgcccgcaccaaccg
 cgccatcagcacc**TGA**taattaa**ctcgaggcagcagcagctcggatagtatgcacac**ctctggacgctggtcgtgtgatggac
 tgttgcgccacacttgccttgacctgtgaatatccctgcgccttttatcaaacagcctcagtggtttgatcttctgtgtacgcg
 cttttgcgagttgtagctgcttgcctatcttgcgaatacaccacccagcatccccttccctcgtttcatatcgcttgcatccaaccgc
 aacttatctacgtgtcctgctatccctcagcgctgctcctgctcctgctcactgccctcgcaagccttggttgggctccgctgta
 ttctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagcttg
 agctctgttttccagaaggagttgctccttgagcctttcattctcagcctcgataacctccaaagccgctctaatttgaggggggttc
 gaattaaaagcttggaaatgttggctgcgtctggaacaagccagactgttgcctcactgggaaaaggaccatcagctccaaa
 aacttgccgctcaaacgcgtacctctgcttgcgcgaatctgcctgttgaatcgccaccacattcataatttgacgcttgagcagt
 ctgtaattgcctcagaatgtggaatcatctgccctgtgcgagccatgccaggcatgtcgcgggcgaggacaccgccactcgta
 cagcagaccattatgctacctcacaatagttcataacagtgaccatatttctgaagctcccaacgagcacctccatgctctgagtg
 gccacccccggcctggtgcttcggagggcaggtcaaccggcatggggctaccgaaatccccgaccggatccaccacccccgc
 gatgggaagaatcttccccgggatgtggccaccaccagcacaacctgctggcccaggcgagcgtcaaaccataccacacaaa
 tatccttggcatcgccctgaattccttctgcgctctgctaccgggtgcttctgtccgaagcaggggttgctagggatcgctccgagt
 ccgcaaaccttgcgcgtggcggggcttgcgagcttgaagagc

SEQ ID NO: 277

coding region for *C. wrightii* KASA1 (codon-optimized)

ATGgccaccgcatccactttctcggttcaatgccgctgcggcgacctgcgtcgctcgggcggtctcgggccccggcgccca
 gcgaggccctccccgtgcgct**acgtatt**ccagtgctggtggccagctgcatcgaccctgcgaccagtaccgcagcagcgcc
 agcctgagcttctgggcgacaacggcttcgccagcctgttcggcagcaagccctcatgagcaacggcgccaccgcccct
 gcgcgcgcagccacagcgcgaggccatggcctggccctgcagccgcccaggaggccggcaccagaagaagcccg
 tgatcaagcagcgcgcgctggtggtgaccggcatggcgctggtgacccccctgggccacgagcccgacgtgttacaacaa
 cctgctggacggcgtgagcggcatcagcgagatcgagaccttcgactgcaaccagttccccaccgcacgcggcgagatca
 agagcttcagcaccgacggctgggtggccccaaagctgagcaagcgcatggacaagttcatgctgtacctgtgaccgcccg
 caagaaggccctggccgacggcgccatcaccgacgaggtgatgaaggagctggacaagcgcaagtgcggcgtgctgatcg
 gcagcggcatgggcggcatgaaggtgttaacgacgccatcgaggccctgcgcgtgagctacaagaagatgaacccctctg
 cgtgcccttcgccaccaccaacatgggcagcgccatgctggccatggacctgggctggatgggccccaaactacagcatcagca
 ccgctgcgccaccagcaacttctgcatctgaacgcgccaaccacatcatccgcgcgaggccgacatgatgctgtgcggc
 ggcagcgacgcctgatcatcccatcggcctggcggttcgtggcctgcgcgcctgagccagcgcaacagcgacccac
 caaggccagccgcccctgggacagcaacccgcagcggttcgtgatgggcgagggcgccggcgtgctgctgtggaggagct

accacgcgcattcagcaccTGA

[illegible]

cctgttcgccccctgcaacgtgTGA

[illegible]

SEQ ID NO: 280
KASII-allele_2_cDNA

ATGCAGACCGCGCACCCAGCGGGCCCCGACCGAGGGGCACTGCTTCGGTGCGAGGCTGCCACGGCGTC
GAGGCGGGGCGGTGCGCCGGGCGTGGTCCCGCATCGCGCGCGGGCGGGCCGACGCGACCCCCG
CCCGCCCTCCGCGCCGCGTGGTCTGTACGGGGCAGGGCGTGGTGACCAGCCTGGGCCAGACGATCGAG
CAGTTTTTACAGCAGCCTGCTGGAGGGCGTGAGCGGCATCTCGCAGATCCAAAAGTTTTGACACCACGGG
CTACACGACGACGATCGCGGGCGAGATCAAGTCGCTGCAGCTGGACCCGTACGTGCCCAAGCGCTGGG
CCAAGCGCGTGGACGACGTCATCAAGTACGTCTACATCGCGGGCAAGCAGGCGCTGGAGAACGCGGGG
CTGCCGATCGAGGCGGCGGGGCTGGCGGGCGCGGGGCTGGACCCCGCGCTGTGCGGCGTGTCTCATCGG
CACCGCCATGGCGGGCATGACGTCCTTCGCGGGCGGGCGTGGAGGCGCTGACGCGCGGGCGGCGTGC
AGATGAACCCCTTTTGCATCCCTTCTCCATCTCCAACATGGGCGGCGCGATGCTGGCGATGGACATC
GGCTTCATGGGCCCCAACTACTCCATCTCCACGGCCTGCGCGACGGGCAACTACTGCATCCTGGGCGC
GGCGGACCACATCCGGCGCGGGCGACGCGGACGTGATGCTGGCCGGCGGGCGCGGACGCGGCCATCATCC
CCTCGGGCATCGGCGGCTTCATCGCGTGCAAGGCGCTGAGCAAGCGCAACGACGAGCCCCGAGCGCGCG
AGCCGGCCCTGGGACGCCGACCGCGACGGCTTCGTATGGGCGAGGGCGCGGGCGTGTGGTGCTGGA
GGAGCTGGAGCACGCCAAGCGCCGCGGGCGCGACCATCCTGGCCGAATTCGTGCGGCGCGCGGCCACCT
CGGACGCGCACCATGACCGAGCCGGACCCGCGAGGGCGCGGGCGTGCCTCTGCCTCGAACGCGCG
CTCGAGCGCGCGCGCTCGCGCCGAGCGCGTGGCTACGTCAACGCGCACGGCACAGCACGCCCCG
GGGCGACGTGGCCGAGTACCGCGCCATCCGCGCGTATCCCGCAGGACTCGCTGCGCATCAACTCCA
CCAAGTCCATGATCGGGCACCTGCTCGGCGGCGCGGGCGGTCGAGGCCGTGGCCGCCATCCAGGCC
CTGCGCACCCGGTGGCTCCACCCCAACCTCAACCTCGAGAACCCCGCACCCGGGGTGCACCCCGTCTG
GCTCGTGGGGCCGCGCAAGGAGCGCGCCGAAGACCTCGACGTGCTCTCTCCAACCTCCTTTGGCTTCG
CGGGGCACAACTCGTGCGTCATCTTCCAAAAGTACGACGAGTGA

SEQ ID NO: 281
KASII-allele_1

MQTAHQRPPTGEGHCFGARLPTASRRVRRAWSRIARAAAAADANPARPERRVVITGQGVVTSLGQTIE
QFYSSLLEGVSGISQIQKFDTTGYTTTTIAGEIKSLQLDPYVPKRWAKRVDDVIKYVYIAGKQALESAG
LPIEAAGLAGAGLDPALCGVLIGTAMAGMTSFAAGVEALTRGGVRKMNPF CIPFSISNMGGAMLAMDI
GFMGPNY SISTACATGNYCILGAADHIRRGDANVLAGGADAAIIPSGIGGFIACKALSKRNDEPERA
SRPWDADRDGFVMGEGAGVVLVLEELEHAKRRGATILAEVLGGAATSDAHMTEPDPQGRGVRLCLERA
LERARLAPERVGYVNAHGTSTPAGDVAEYRAIRAVIPQDSLRLNSTKSMIGHLLGGAGAVEAVAAIQAL
LRTGWLHPNLNLNENPAGVDPVVLVGP RKERAEDLDVVLNSNFGFGGHNSCVIFRKYDE

SEQ ID NO: 282
KASII-allele_2

MQTAHQRPPTGEGHCFGARLPTASRRVRRAWSRIARAAAAADATPARPPRRVVVTGQGVVT
SLGQTIEQFYSSLLEGVSGISQIQKFDTTGYTTTTIAGEIKSLQLDPYVPKRWAKRVDDVIKYVY
IAGKQALENAGLPIEAAGLAGAGLDPALCGVLIGTAMAGMTSFAAGVEALTRGGVRKMNPF
CIPFSISNMGGAMLAMDIGFMGPNY SISTACATGNYCILGAADHIRRGDADVLAGGADAAI
IPSGIGGFIACKALSKRNDEPERASRPWDADRDGFVMGEGAGVVLVLEELEHAKRRGATILAEF
VGAATSDAHMTEPDPQGRGVRLCLERALERARLAPERVGYVNAHGTSTPAGDVAEYRAI
RAVIPQDSLRLNSTKSMIGHLLGGAGAVEAVAAIQALRTGWLHPNLNLNENPAGVDPVVLV
GPRKERAEDLDVVLNSNFGFGGHNSCVIFQKYDE

SEQ ID NO: 283
pSZ1503

gctcttccgcacccggtggctccacccaacttgaacctcgagaaccccgcgctggcgctgaccccgctgctgctggtggcgccg
ggaaggagcgcgccaagacctggacgtcgctcctctccaaactcttggcttggcgggcaccaattcgctgctggtacccttctt

gcgctatgacacttcagcaaaaggtagggcgggctgcgagacggctcccggcgtgcatgcaacaccgatgatgcttcgacccc
 ccgaagctccttcggggctgcatggcgctccgatgccgctccagggcgagcgctgtttaaataagccagggccccgattgcaaaga
 cattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacagggcactcgagcttgatcgactccg
 ctaagggggcgctcttctcttctgttcagtcacaacccgcaaacggcgcgccATGctgctgcaggccttctgttctgtggtg
 ggcttcgcccgaagatcagcgctccatgacgaacgagacgtccgacggccccctgggtgcacttcacccccacaagggtg
 gatgaacgacccccacggcctgtggtacgacgagaaggacgccaagtggcacctgtacttcagtaacaacccaacgacac
 cgtctgggggacgccttgttctggggccacgcctcgacgacctgaccaactgggaggaccagcccatcgccatcgccc
 cgaagcgcaacgactccggcgcttctccggtccatggtggtggactacaacaacacctccggcttcttaacgacaccatcg
 acccgcgccagcgctgctggtgccatctggacctacaacaccccgagtcggaggagcagtaatatctctacagcctggacggc
 ggctacaccttcaccgagtaccagaagaaccccgctgtggcgcccaactccaccagttccgacacccgaaggcttctgtgtac
 gagccctcccagaagtggatcatgaccgcgccaagtcccaggactacaagatcgagatctactctccgacgacctgaagtc
 ctggaagctggagtcgcgttcgccaacgagggcttctcggtaccagtagcagtgccccggcctgacgaggtccccaccg
 agcaggacccccagcaagtctactgggtgatgttatctccatcaacccggcgccccggcggggctcttcaaccagtaatt
 cgtcggcagcttcaacggcaccacacttcgaggccttcgacaacacagtcctcggtgggtgacttcggcaaggactactacgcct
 gcagaccttcttaacacccgacccgacctacgggagcgccctgggcatcgctgggctccaactgggagtactccgcttctgt
 gcccaccaacccctggcgctctctcatgtccctcgtcgcaagttctccctcaacacccgagtaccaggccaacccggagacgga
 gctgatcaacctgaaggccgagccgatcctgaacatcagcaacgcgggccctggagccggttcgccaccaacaccacgttg
 acgaaggccaacagctacaacgtcgacctgtcaacagcaccggcaccctggagttcgagctggtgtacgcgtcaacacca
 ccagacgatctccaagtcggttcgaggacctctccctctggttcaaggcgctggaggaccccgaggagtacctccgcatgg
 gcttcgaggtgtccgctctcttcttctggaccgcggaacagcaaggtgaagttcgtgaaggagaacccctacttcacca
 accgcatgagcgtgaacaacacgacctcaagagcgagaacgacctgtctactacaaggtgtacggcttctgtggaccagaa
 catcttgagctgtacttcaacgacggcgacgtcgtgtccaccaacacacttctcatgaccacgggaacgccttgggtccgt
 gaacatgacgacgggggtggacaacctgttctacatcgacaagttccaggtgcgaggtcaagTGAcaattggcagcagc
agctcggatagtatcgacacactctggacgctggtcgtgtgatggactgttccgcccacacttgctgccttgacctgtgaatatccctg
ccgcttttatcaaacagccctcagtggttgccttctgtgtacgcgcttttgcgagttgctagctgcttgcctatttgcgaataccacc
cccagcatcccttccctcgttcatatcgcttgcacccaacggcaacttatctacgctgtcctgctatccctcagcgctgctcctgctc
ctgctcactgccccctcgacagccttgggttgggctccgctgtattctctggtactgcaacctgtaaacacagcactgcaatgtgat
gcacgggaagtagtgggatgggaacacaaatggaggatcgtagagctcatcttccgaaagtacgacgagtgagcgagctgattct
ctttgagcggggtcggttgggttcgggagagtgcgcgaaaggcgagagacgtgcggccggccgtgtccctccgtcttccctgg
ttggtgctatagtaacctgcctgtgtcgcgtgcgctcggaagagc

SEQ ID NO: 284

Ganoderma tsugae manganese-superoxide dismutase (Mn-SOD) gene, partial cds., Genbank

ACCESSION U56106 VERSION U56106.1

| | | | | | |
|------------|------------|------------|------------|------------|------------|
| CTCCACCACA | AGAAGCACCA | CCAGACCTAC | GTGAACGCGC | TCAACGCCGC | AGAGCAGGCC |
| TACGCCAAGG | CCACCACCCC | CAAGGAGCGC | ATCGCCCTCC | AGTCCGCCCT | CAAGTTCAAC |
| GGCGGTGGTG | AGTGAATTTT | GCGAATGTCT | GCTCACCAGC | ACGCGATTAC | GAATCCGCTA |
| ACTCTAACAC | TTTGCTGCGG | GGGATGTGCA | GGTCACATCA | ACCACTCCCT | CTTCTGGAAG |
| AACCTCGCCC | CAGCCAAGTC | TGAGGGCAAG | GGAAACGGCG | GCGCCATCAA | GGACGGCCCC |
| CTCAAGTCGG | CGATCGAGCA | GAACTTTGGT | TCCGTCGACG | TCTTCATCAA | GGAGTTCAAC |
| GCGACCACCG | CGTCCATCCA | GGGCTCCGGC | TGGGGCTGGC | TCGTGAGTCT | TCTCTTCTGC |
| GATCGGAAGG | CGCATCCGGT | GTTATGGTCT | CGCGCTGCAC | GCGTGTCTCG | CTTCGGCTCC |
| GCCGTCGAAG | CGCTGATACC | GAGACCGATC | TGCTGCAAAT | CATTTGCGCT | GCGATAGAGC |
| TCCGGGGGCT | TTAACCTTTC | GTTCCATGTG | ATAAGACCGC | GTCTGACCGG | GTGGTGTTTT |
| CTATCGCAGG | GTCTCAACCC | GGCGACCAAG | CGCCTCGAGA | TCACGACGAC | GGCCAACACG |
| GACCCTCTCC | TCACGCACGT | CCCCATCATC | GGCGTCGACA | TCTGGGAGCA | CGCGTTC |

SEQ ID NO: 285

Umbellularia californica (California bay) thioesterase mRNA, complete cds., Genbank ACCESSION
M94159 VERSION M94159.1

```

AGAGAGAGAG AGAGAGAGAG AGCTAAATTA AAAAAAAAAAC CCAGAAGTGG GAAATCTTCC
CCATGAAATA ACGGATCCTC TTGCTACTGC TACTACTACT ACTACAAACT GTAGCCATTT
ATATAATTCT ATATAATTTT CAACATGGCC ACCACCTCTT TAGCTTCCGC TTTCTGCTCG
ATGAAAGCTG TAATGTTGGC TCGTGATGGC CGGGGCATGA AACCCAGGAG CAGTGATTG
CAGCTGAGGG CGGGAAATGC GCCAACCTCT TTGAAGATGA TCAATGGGAC CAAGTTCAGT
TACACGGAGA GCTTGAAAAG GTTGCCTGAC TGGAGCATGC TCTTTGCAGT GATCACAACC
ATCTTTTCGG CTGCTGAGAA GCAGTGGACC AATCTAGAGT GGAAGCCGAA GCCGAAGCTA
CCCCAGTTGC TTGATGACCA TTTTGGACTG CATGGGTTAG TTTTCAGGCG CACCTTTGCC
ATCAGATCTT ATGAGGTGGG ACCTGACCGC TCCACATCTA TACTGGCTGT TATGAATCAC
ATGCAGGAGG CTACACTTAA TCATGCGAAG AGTGTGGGAA TTCTAGGAGA TGGATTCGGG
ACGACGCTAG AGATGAGTAA GAGAGATCTG ATGTGGGTTG TGAGACGCAC GCATGTTGCT
GTGGAACGGT ACCCTACTTG GGGTGATACT GTAGAAGTAG AGTGCTGGAT TGGTGCACTT
GGAAATAATG GCATGCGACG TGATTTTCCT GTCCGGGACT GCAAAAACAG CGAAATTCTT
ACAAGATGTA CCAGCCTTTC GGTGCTGATG AATACAAGGA CAAGGAGGTT GTCCACAATC
CCTGACGAAG TTAGAGGGGA GATAGGGCCT GCATTCATTG ATAATGTGGC TGCAAGGAC
GATGAAATTA AGAAACTACA GAAGCTCAAT GACAGCACTG CAGATTACAT CCAAGGAGGT
TTGACTCCTC GATGGAATGA TTTGGATGTC AATCAGCATG TGAACAACCT CAAATACGTT
GCCTGGGTTT TTGAGACCGT CCCAGACTCC ATCTTTGAGA GTCATCATAT TTCCAGCTTC
ACTCTTGAAT ACAGGAGAGA GTGCACGAGG GATAGCGTGC TGCGGTCCCT GACCACTGTC
TCTGGTGGCT CGTCGGAGGC TGGGTTAGTG TGCGATCACT TGCTCCAGCT TGAAGGTGGG
TCTGAGGTAT TGAGGGCAAG AACAGAGTGG AGGCCTAAGC TTACCGATAG TTTCAGAGGG
ATTAGTGTGA TACCCGCAGA ACCGAGGGTG TAACTAATGA AAGAAGCATC TGTTGAAGTT
TCTCCCATGC TGTTCTGTAG GATACTTTTT AGAAGCTGCA GTTTGCATTG CTTGTGCAGA
ATCATGGTCT GTGGTTTTAG ATGTATATAA AAAATAGTCC TGTAGTCATG AACTTAATA
TCAGAAAAAT AACTCAATGG GTCAAGGTTA TCGAAGTAGT CATTTAAGCT TTGAAATATG
TTTTGTATTC CTCGGCTTAA TCTGTAAGCT CTTTCTCTTG CAATAAAGTT CGCCTTTCAA
T

```

SEQ ID NO: 286

Cuphea calophylla subsp. *Mesostemon*, seed-specific acyl-acyl carrier protein thioesterase, Genbank
ACCESSION ABB71579 VERSION ABB71579.1

```

MVAAEASSAL FSVRTPGTSP KPGKFRNWPS SLSVPFKSKS NHNGGFQVKA NASARPKANG
SAVSLKSGSL DTQEDTSSSS SPPRTFINQL PDWSMLLSAI TTVFVAEEKQ WTMLDRKSKR
PDTLMDPFGV DRVVDGTVVF RQSFIRSIE IGADRTASIE TLMNIFQETS LNHCKSIGLL
NDGFGRTPEM CKRDLIWVVT KMHIEVNRYP TWGDTIEVNT WVSESGKTGM GRDWLISDCH
TGEILIRATS VCAMMNQTTR RFSKFPYEVF QELAPHFVDS APVIEDYQKL HKLDVKTGDS
ICNGLTPRWV DLDVNQHVNN VKYIGWILES VPKEVFETQE LCGLTLEYRR ECGRDSVLKS
VTAMDPSKEG DRSLYQHLLR LEDGTDIAKG RTKWRPKNAG TNGAISTGKT SNGNSIS

```

SEQ ID NO: 287

Ulmus americana (American elm) acyl-ACP thioesterase, Genbank ACCESSION AAB71731
VERSION AAB71731.1

```

GSGALQVKAS SQAPPKLNGS NVGLVKSSQI VKKGDDTTSP PARTFINQLP DWSMLLAAIT
TLFLAAEKQW MMLDWKPKRP DMLVDPFGLG RFVQDGLVFR NNFSIRSIEI GADRTASIEI
LMNHLQETAL NHVKSIVGLLE DGLGSTREMS LRNLIVVVTQ MQVAVDRIPT WGDEVQVSSW
ATAIGKNGMR REWIVTDFRT GETLLRATSV WVMNKLTRR ISKIPPEVWH EIGPSFIDAP
PLPTVEDDGR KLTRFDESSA DFIRXGLTPR WSDLDINQHV NNVKYIGWLL ESAPPEIHES

```

HEIASLTLEY RRECGRDSVL NSATKVSDSS QLGKSAVECN HLVRLQNGGE IVKGRTVWRP
KRPLYNDGAV VDVXAKTS