

SEQUENCE LISTING

SEQ ID NO: 1

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

GATCAGACGGGCTGACCTGCGAGATAATCAAGTGCTCGTAGGCAACCAACTCA
GCAGCTGCTTGGTGTGTTGGGTCTGCAGGATAGTGTTCAGGGCCCCAAGGACAGC
AGGGGAACTTACACCTTGTCCCCGACCCAGTTTTATGGAGTGCATTGCCTCAAGA
GCCTAGCCGGAGCGCTAGGCTACATACTTGCCGCACCGGTATGAGGGGATATAG
TACTCGCACTGCGCTGTCTAGTGAGATGGGCAGTGCTGCCCATAAACAACCTGGCT
GCTCAGCCATTGTGTTGGCGGACCATTTCTGGGGGGGCCAGCAATGCCTGACTTTCG
GGTAGGGTGAAAACCTGAACAAAGACTACCAAAACAGAATTTCTTCCTCCTTGGA
GGTAAGCGCAGGCCCGGCCGCTGCGCCACATGGCGCTCCGAACACCTCCATA
GCTGTAAGGGCGCAAACATGGCCGGACTGTTGTCAGCACTCTTTCATGGCCATAC
AAGGTCATGTCGAGATTAGTGCTGAGTAAGACACTATCACCCCATGTTTCGATTGA
AGCCGTGACTTCATGCCAACCTGCCCTGGGCGTAGCAGACGTATGCCATCATGA
CCACTAGCCGACATGCGCTGTCTTTTGCCACCAAAACAACCTGGTACACCGCTCGA
AGTCGTGCCGCACACCTCCGGGAGTGAGTCCGGCGACTCCTCCCCGGCGGGCCG
CGGCCCTACCTGGGTAGGGTCGCCATACGCCACGACCAAAACGACGCAGGAGGG
GATTGGGGTAGGGAATCCCAACCAGCCTAACCAAGACGGCACCTATAATAATAG
GTGGGGGGGACTAACAGCCCTATATCGCAAGCTTTGGGTGCCTATCTTGAGAAGCA
CGAGTTGGAGTGGCTGTGTACGGTCGACCCTAAGGTGGGTGTGCCGCAGCCTGA
AACAAAGCGTCTAGCAGCTGCTTCTATAATGTGTCAGCCGTTGTGTTTCAGTTAT
ATTGTATGCTATTGTTTGTTCGTGCTAGGGTGGCGCAGGCCACCTACTGTGGCG
GGCCATTGGTTGGTGCTTGAATTGCCTCACCATCTAAGGTCTGAACGCTCACTCA
AACGCCTTTGTACAACCTGCAGAACTTTCCTTGGCGCTGCAACTACAGTGTGCAAA
CCAGCACATAGCACTCCCTTACATCACCCAGCAGTACAACA

SEQ ID NO: 2

Chlorella ellipsoidea nitrate reductase promoter from AY307383

CGCTGCGCACCAAGGGCCGCGAGCTCGCTGATGTCGCTCCAAATGCGGTCCCCCGA
TTTTTTGTTCTTCATCTTCTCCACCTTGGTGGCCTTCTTGCCAGGGCCTTCAGCTG
CATGCGCACAGACCGTTGAGCTCCTGATCAGCATCCTCAGGAGGCCCTTTGACAA
GCAAGCCCCTGTGCAAGCCCATTACGGGGTACCAGTGGTGCTGAGGTAGATGG
GTTTGAAAAGGATTGCTCGGTTCGATTGCTGCTCATGGAATTGGCATGTGCATGCA
TGTTTACAATATGCCACCAGGCTTTGGAGCAAGAGAGCATGAATGCCTTCAGGC
AGGTTGAAAGTTCTTGGGGGTGAAGAGGCGAGGGCCGAGGATTGGAGGAGGAAA
GCATCAAGTCGTCGCTCATGCTCATGTTTTTCAGTCAGAGTTTGCCAAGCTCACAG
GAGCAGAGACAAGACTGGCTGCTCAGGTGTTGCATCGTGTGTGTGGTGGGGGGG
GGGGGGTTAATACGGTACGAAATGCACTTGGAAATCCACCTCATGCCAGCGGA
CCCACATGCTTGAATTCGAGGCCTGTGGGGTGAGAAATGCTCACTCTGCCCTCGT
TGCTGAGGTACTTCAGGCCGCTGAGCTCAAAGTCGATGCCCTGCTCGTCTATCAG
GGCCTGCACCTCTGGGCTGACCGGCTCAGCCTCCTTCGCGGGCATGGAGTAGGCG
CCGGCAGCGTTCATGTCCGGGCCAGGGCAGCGGTGGTGCCATAAATGTCCGTG
ATGGTGGGGAGGGGGGCGCTCGCCACACCATTGCCGTTGCTGGCTGACGCATGC
ACATGTGGCCTGGCTGGCACCGGCAGCACTGGTCTCCAGCCAGCCAGCAAGTGG
CTGTTTCAAGGAAAGCGGCCATGTTGTTGGTCCCTGCGCATGTAATTCGCCAGATCA
AAGGAGGGAACAGCTTGGATTTGATGTAGTGCCCAACCGGACTGAATGTGCGAT

GGCAGGTCCCTTTGAGTCTCCCGAATTACTAGCAGGGCACTGTGACCTAACGCAG
CATGCCAACCGCAAAAAAATGATTGACAGAAAATGAAGCGGTGTGTCAATATTT
GCTGTATTTATTCGTTTTAATCAGCAACCAAGTTCGAAACGCAACTATCGTGGTG
ATCAAGTGAACCTCATCAGACTTACCTCGTTCGGCAAGGAAACGGAGGCACCAA
ATTCCAATTTGATATTATCGCTTGCCAAGCTAGAGCTGATCTTTGGGAAACCAAC
TGCCAGACAGTGGACTGTGATGGAGTGCCCCGAGTGGTGGAGCCTCTTCGATTTCG
GTTAGTCATTACTAACGTGAACCCTCAGTGAAGGGACCATCAGACCAGAAAGAC
CAGATCTCCTCCTCGACACCGAGAGAGTGTTCGCGCAGTAGGACGACAAG

SEQ ID NO: 3

Yeast sucrose invertase

MTNETSDRPLVHFTPKNKGWMNDPNGLWYDEKDAKWHL YFQYNPNDTVWGTPLF
WGHATSDDL TNWEDQPIAIAPKRND SGAFSGSMVVDYNNTSGFFNDTIDPRQRCVAI
WTYNTPESSEEQYISYSLDGGYTFTEYQKNPVLAANSTQFRDPKVFWEPSQKWIMT
AAKSQDYKIEIYSSDDLKSWKLES AFANEGFLGYQYECPLIEVPTEQDPSKSYWVM
FISINPGAPAGGSFNQYFVGSFNGTHFEAFDNQSRVVDFGKDYYALQTFENTDPTYGS
ALGIAWASNWEYS AFVPTNPWRSSMSLVRKFSLNTEYQANPETELINLKAEPILNISN
AGPWSRFATNTTLTKANSYNVDLSNSTGTLEFELVYAVNTTQTISKSVFADLSLWFK
GLEDPEEYLRMGFEVSASSFFLDRGNSKVKFVKENPYFTNRMSVNNQPFKSENDLSY
YKVYGLLDQNILELYFNDGDVVSTNTYFMTTGNALGSVNMTTGVDNLFYIDKFQVR
EVK

SEQ ID NO: 4

Yeast secretion signal

MLLQAFLFLLAGFAAKISAS

SEQ ID NO: 5

Higher plants secretion signal

MANKSLLLLLLLGLSLASG

SEQ ID NO: 6

Consensus eukaryotic secretion signal

MARLPLAALG

SEQ ID NO: 7

Combination higher plant/eukaryotic secretion signal

MANKLLLLLLLLLPLAASG

SEQ ID NO: 8

S. cerevisiae sucrose invertase NP_012104

GAATTCCCCAACATGGTGGAGCACGACACTCTCGTCTACTCCAAGAATATCAAAG
ATACAGTCTCAGAAGACCAAAGGGCTATTGAGACTTTTCAACAAAGGGTAATAT
CGGGAAACCTCCTCGGATTCCATTGCCAGCTATCTGTCACTTCATCAAAAGGAC
AGTAGAAAAGGAAGGTGGCACCTACAAATGCCATCATTGCGATAAAGGAAAGGC
TATCGTTCAAGATGCCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACG
AGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGAT

SEQ ID NO: 11*UTEX 329 Prototheca kruegani*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGAAAATACT
CTGGAGCCATAGCGAAAGCAAGTTTAGTAAGCTTAGGTCATTCTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTAGTAAAATAACATGGAG
GCCCCAACCGACTAATGTTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAGTACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGACTAAGGATGTGGGTATGTCAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 12*UTEX 1440 Prototheca wickerhamii*

TGTTGAAGAATGAGCCGGCGACTTAAAAATAAATGGCAGGCTAAGAGATTTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCTTAATAGGGCGTCAATTTAACAAAACCT
TAAATAAATTATAAAGTCATTTATTTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC
TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACCAGTTAAGGCCCCAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 13*UTEX 1442 Prototheca stagnora*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAAAATGGCATGGTTAAAGATATTTCT
CTGAAGCCATAGCGAAAGCAAGTTTACAAGCTATAGTCATTTTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTCAAATAACATGGAG
GCCCCAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGACTAAGGACGTGAGTATGTCAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 14*UTEX 288 Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGATAATTCT
CTGGAGCCATAGCGAAAGCAAGTTTAAACAAGCTAAAGTCACCCTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG

GCCCGAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGACTAAGGATGTGGGTATGTTAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 15

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

TGTTGAAGAATGAGCCGGCGACTTAAAAATAAATGGCAGGCTAAGAGAATTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCCTTAATAGGGCGCTAATTTAACAAAACAT
TAAATAAAAATCTAAAGTCATTTATTTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC
TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACCAGTTAAGGCCCCAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 16

UTEX 1533 *Prototheca wickerhamii*

TGTTGAAGAATGAGCCGTCGACTTAAAAATAAATGGCAGGCTAAGAGAATTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCCTTAATAGGGCGCTAATTTAACAAAACAT
TAAATAAAAATCTAAAGTCATTTATTTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC
TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACCAGTTAAGGCCCCAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 17

UTEX 1434 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCGTGGTTAAAGAAAATTCT
CTGGAACCATAGCGAAAGCAAGTTTAACAAGCTTAAGTCACTTTTCTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG
GCCCCAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTCCGAAAG
TTGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGA
CCTTGGGGGATAAGCTCCTTGGTCGAAAGGGAAACAGCCCAGATCACCAGTTAA

GGCCCCAAAATGAAAATGATAGTGAAGGATGTGAGTATGTCAAAACCTCCA
GCAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 18

UTEX 1438 *Prototheca zopfii*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGAAAATTCT
CTGGAGCCATAGCGAAAGCAAGTTTAAACAAGCTTAAGTCACTTTTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG
GCCCGAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTCGTGGGCTTCGAAAG
TTGTACCTCAAAGTGGAACACTCTGAATACTCTATTTAGATATCTACTAGTGAGA
CCTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAA
GGCCCCAAAATGAAAATGATAGTGAAGGATGTGAGTATGTCAAAACCTCCA
GCAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 19

UTEX 1436 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGACTTAGAAAAGGTGGCATGGTTAAGGAAATATTC
CGAAGCCGTAGCAAAAGCGAGTCTGAATAGGGCGATAAAATATATTAATATTTA
GAATCTAGTCATTTTTTCTAGACCCGAACCCGGGTGATCTAACCATGACCAGGAT
GAAGCTTGGGTGATACCAAGTGAAGGTCCGAACCGACCGATGTTGAAAAATCGG
CGGATGAGTTGTGGTTAGCGGTGAAATACCAGTCGAACCCGGAGCTAGCTGGTT
CTCCCCGAAATGCGTTGAGGCGCAGCAGTACATCTAGTCTATCTAGGGGTAAAGC
ACTGTTTTCGGTGCGGGCTGTGAGAACGGTACCAAATCGTGGCAAACCTCTGAATAC
TAGAAATGACGATGTAGTAGTGAGACTGTGGGGGATAAGCTCCATTGTCAAGAG
GGAAACAGCCCAGACCACCAGCTAAGGCCCCAAAATGGTAATGTAGTGACAAAG
GAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGCCATCCTTTA
AAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 20

Chicorium intybus invertase: Genbank Accession No. Y11124

MSNSSNASESLFPATSEQPYRTAFHFQPPQNWMDPNPGPMCYNGVYHLFYQYNPFG
PLWNLRLMYWAHSVSHDLINWIHLDLAFAPTEPFDINGCLSGSATVLPGNKPIMLYTG
IDTENRQVQNLAVPKDLSDPYLREWVKHTGNPIISLPEEIQPDDFRDPTTTWLEEDGT
WRLLVGSQKDKTGIAFLYHSGDFVNWTKSDSPLHKVSGTGMWECVDFFPVWVDST
NGVDTSIINPSNRVKHVLKLGIDHKGDCYLIGKYSADKENYVPEDELTLSTLRLDY
GMY YASKSFFDPVKNRRIMTAWVNESDSEADVIARGWSGVQSFPRSLWLDKNQKQ
LLQWPIEEEIEMLHQNEVSFHNKKLDGGSSLEVLGITASQADV KISFKLANLEEAEELD
PSWVDPQLICSENDASKKGKFGPFGLLALASSDLREQTAFRRVFRKNGRYVVL MCS
DQSRSSMKNGIEKRTYGA FVDIDPQQDEISLRTLIDHSIVESFGGRGKTCITTRVYPTL
AIGEQARLFAFNHGTESVEISELSAWSMKKAQMKVEEP

SEQ ID NO: 21

Schizosaccharomyces pombe Invertase: Genbank Accession No. AB011433

MFLKYILASGICLVSLLSSTNAAPRHLVVKRYPVIYNASNITEVSNSTTVPPPPFVNNT
APNGTCLGNYNEYLP SGYYNATDRPKIHFT PSSGFMNDPNGLVYTGGVYHMFQYS
PKTLTAGEVHWGHTVSKDLIHWENYPIAIYPDEHENGVLSPFSGSAVVDVHNSSGL
FSNDTIPEERIVLIYTDHWTGVAERQAIAYTTDGGYTFFKKYSGNPVLDINSLQFRDPK
VIWDFDANRWVMIVAMSQNYGIAFYSSYDLIHWTELSVFSTSGYLGLQYECPGMAR
VPVEGTDEYKWVLFISINPGAPLGGSVVQYFVG DWNGTNFVPDDGQTRFVDLGKDF
YASALYHSSSANADVIGVGWASNWQYTNQAPTQVFRSAMTVARKFTLRDVPQNPM
TNLTSLIQTPLNVSLLRDETFTAPVINSSSSSLSGSPITLPSNTAFEFNVTL SINYTEGCT
TGYCLGRIIDSDDPYRLQ SISVDVDFAASTLVINRAKAQMGWFNSLFTPSFANDIYY
GNVTLYGIVDNGLLELYVNNGEKT YTNDFFFLQGATPGQISFAAFQGVSFNNVTVTP
LKTWNC

SEQ ID NO: 22

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

MIQLSPLLLLPLFSVFNSIADASTEYLRPQIHLTPDQGWMNDPNGMFYDRKDKLWHV
YFQHNPDKKSIWATPVTWGHSTSKDLLTWDYHGNALPENDDEGIFSGSVVDRNN
TSGFFNDSTDPEQRIVAIYTNNAQLQTQEIAYSLDKGYSEFIKYDQNPVINVNSSQQRD
PKVLWHDESQWIMVVAKTQEFKVQIYGSPDLKKWDLKSNFTSNGYLG FQYECPG
LFLKLIENPLNDTVTSKWVLLLAINPGSPLGG SINEYFIGDFDGTTFHPDDGATRFMDI
GKDFYAFQSF DNTEPEDGALGLAWASNWQYANTVPTENWRSSMSLVRNYTLKYVD
VNPENYGLTLIQKPVYDTKETRLNETLKTLETINEYE VNDLKLDKSSFVATDFNTERN
ATGVFEFDLKTQTDLKMGYSNMTTQFGLYIHSQTVKGSQETLQLVFDTLSTTWYID
RTTQHSFQRNSPVFTERISTYVEKIDTTDQGNVYTLYG VVDRNILELYFNDGSIAMTN
TFFFREGKIPTSFEVVCDSEKSFITIDELSVRELARK

SEQ ID NO: 23

Debaryomyces occidentalis Invertase: Genbank Accession No. X17604

MVQVLSVLVIPLLTFFGYVASSSIDLSVDTSEYNRPLIHFTPEKGWMNDPNGLFYDK
TAKLWHL YFQYNPNATAWGQPLYWGHATSNDLVHWDEHEIAIGPEHDNEGIFSGSI
VVDHNNTSGFFNSSIDPNQRIVAIYTN NIPDLQTQDIAFSLDGGYTFTKYENNPNVIDVS
SNQFRDPKVFWHEREFKSMDHGCSEIARVKIQIFGSANLKNWVLNSNFSSGYYGNQY
GMSRLIEVPIENS DKSKWVMFLAINPGSPLGG SINQYFVGDFDGFQFVPDDSQTRFVD
IGKDFYAFQTFSEVEHGV LGLAWASNWQYADQVPTNPWRSSTSLARNYTLRYVIQM
LKL TANIDKSVLPDSINVVDKLKKKNVKL TNKKPIKTNFKGSTGLDFDNITFKVLNLN
VSPGKTHFDILINSQELNSSVDSIKIGFDSSQSLFYIDRHIPNVEFPRKQFFTDKLAAYL
EPLDYDQDLRVFSLYGIVDKNIIELYFNDGTVAMTNTFFMGE GKYPHDIQIVTDTEEP
LFELESVIIRELNK

SEQ ID NO: 24

Oryza sativa Invertase: Genbank Accession No. AF019113

MATSRLTPAYDLKNA AAAVYTPLPEQPHSAEVEIRDRKPFKIISAILSSLLLLALILVA
VNYQAPPSHSSGDNSQPAAVMPPSRGVSQGVSEKA FRGASGAGNGV SFAWSNLMLS
WQRTSYHFQPVKNWMNDPNGLYYKGWYHLFYQYNPDSAVWGNITWGHAVSTD
LINWLHLPFAMVPDQWYDVNGVWTGSATILPDGRIVMLYTGD TDDYVQDQNLAFP
ANLSDPLLVDWVKYPNNPVIYPPPGIGVKDFRDPTTAGTAGMQNGQRLVTIGSKVG
KTGISLVYETTNFTTFKLLYGVLHAVPGTGMWECVDLYPVSTTGENGLDTSVNLG
VKHVLKTSLDDDKHDYYALGTYPDVK NKWTPDNPDL DVGIGLRLDYGYKYYAARTF

YDQNKQRRILWGWIGETDLEAVDLMKGWASLQAIPRTIVFDKKTGTNVLQRPEEEV
ESWSSGDPITQRRIFEPGSVVIHVSGATQLDITASFEVDETLETTSESH DAGYDCSN
SGGAGTRGSLGPFGLLVVADEKLSELTPVYLYVAKGGDGKAKAHL CAYQTRSSMAS
GVEKEVYGSAPVLDGENYSARILIDHSIVESFAQAGRTCVRSDYPTKDIYGAARCF
FFNNATEASVRASLKAWQMKSFIRPYPFIPDQKS

SEQ ID NO: 25

Allium cepa Invertase: Genbank Accession No. AJ006067

MSSDDLESPPSSYLPIPPSDEFHDQPPPLRSWLRLLSIPLALMFLFLATFLSNLESPPSD
SGLVSDPVTFDVNPAVVRGKDAGVSDKTSVSDSGFVLDPVAVDANSVVVHRGKD
AGVSDKTSVSDSGLLKDSPLGPYPWTNQMLSWQRTGFHFQPVKNWMNDPNGPLY
KGWYHFFYQYNPEGAVWGNIAWGHAVSRDLVHWTHLPLAMVPDQWYDINGVWT
GSATILPDGQIVMLYTGATNESVQVQNLAVPADQSDTLLLRWKKSEANPILVPPP
DKDFRDPTTAWYEPSDDTWRIVIGSKDSSSHSGIAIVYSTKDFINYKLIPGILHA
VERVGMWECVDFYPVATADSSHANHGLDPSARPSAVKHVLKASMDDDRHDY
YAIGTYDPAQNTWVPDDASVDVGIGLRYDWGKFYASKTFYDHAKKRRILWSWIGETD
SETADIAKGWASLQGVPRTVLLDVKTGSNLITWPVVEIESLRTRPRDFSGITV
DAGSTFKLDVGGAAQLDIEAEFKISSELEAVKEADVSYNCSSSGGAAERGV
LGPFGLLVLANQDLTEQTATYFYVSRGMDGGLNTHFCQDEKRSSKASDIV
KRIVGHSVPVLDGESFALRILVDHSIVESFAQGGRASATSRVYPT
EAIYNNARVVFVNNATGAKVTAQSLKVWHMSTAI NEIYDPATSV

SEQ ID NO: 26

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

LFYQYNPNNGVIWGPVWGHSTSKDLVNWVPQPLTMEPEMAANINGSWSGSATILPG
NKPAILFTGLDPKYEQVQVLAYPKDTS DPNLKEWFLAPQNPVMFPTPQNQINAT
SFRDPTTAWRLPDGVWRLLIGSKRGQRGLSLLFRSRDFVHWVQAKHPLYSDKL
SGMWECPDFFPVYANGDQMGVDTSIIGSHVKHVLKNSLDITKHDIYTIGDYN
IKKDAYTPDIGYMNDSSLRYDYGKYYASKTFFDDAKKERILLGWANESSSV
EDDIKKGWSGIHTIPRKIWLDKLGKQLIQWPIANIEKLRQKPVNIYRKVLKGG
SQIEVSGITAAQADVEISFKIKDLKNVEKFDASWTSPQLLCSKKGASVKGGL
GPFGLLTLASXGLEEYTAFFRIFKAYDNKFVVL MCSDQSRSSLNPTNDKTTY
GTFVDVNPIREGLSLRVLIDHSVVESFGAKGKNVITARVYPTLAINEKAHLY
VFNRGTSNVEITGLTAWSMKKANIA

SEQ ID NO: 27

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

MTDFTPETPVLTPIRDHAAELAKAEAGVAEMA AKRNNRWYPKYHIASNGGWINDPN
GLCFYKGRWHVIFYQLHPYGTQWGPMHWGHVSSTDMLNWKREPIMFAPSLEQEKD
GVFSGSAVIDDNGDLRFYYTGHRWANGHDNTGGDWQVQMTALPDNDELTSATKQ
GMIIDCPTDKVDHHRDPKVWKTGDTWYMTFGVSSDKRGQM WLFSSKDMVRWE
YERVL FQHPDPDVFMLECPDFFPIKDKDGNEKWVIGFSAMGSKPSGFMNRNVN
NAGYMIGTWEPGGEFKPETEFRLWDCGHNY YAPQSFNVDGRQIVYGWMSPFVQ
PIPMEDDGWCGQLTLPREITLDDDGDVV TAPVAEMEGLREDTLDHGSITLDM
DGEQVIADDAEAVEIEMTIDLA ASTADRAGLKI HATEDGAYTYVAYDDQIGRV
VVD RQAMANGDHGYRAAPLTDAELASGKLDLRV FVDRGSVEVYVNGGHQVLSS
YSYASEGPRAIKLVA EFGNLKVESLKLHMHMSIGLE

SEQ ID NO: 28

Saccharomyces cerevisiae Invertase: Genbank Accession No. NP_012104

MLLQAFLFLLAGFAAKISASMTNETSDRPLVHFTPNKGWMNDPNGLWYDEKDAKW
HLYFQYNPNDTVWGTPLFWGHATSDDL TNWEDQPIAIAIPKRND SGAFSGSMVVDY
NNTSGFFNDTIDPRQRCVAIWTYNTPESSEEQYISYSLDGGYTFTTEYQKNPVLAANSTQ
FRDPKVFWEPSQKWIMTAAKSQDYKIEIYSSDDLKSWKLESANEGFLGYQYECF
GLIEVPTEQDPSKSYWVMFISINPGAPAGGSFNQYFVGSFNGTHFEAFDNQSRVVDG
KDYYALQTFNTDPTYGSALGIAWASNWEYSASFVPTNPWRSSMSLVRKFSLNTEYQ
ANPETELINLKAEPILNISNAGPWSRFATNTTLTKANSYNVDLSNSTGTLEFELVYAV
NTTQTISKSVFADLSLWFKGLEDPPEYLRMGFEVSASSFFLDRGNSKVVKFVKENPYFT
NRMSVNNQPFKSENDLSYYKVYGLLDQNILELYFNDGDVSTNTYFMTTGNALGSV
NMTTGVDNLFYIDKFQVREVK

SEQ ID NO: 29

Zymomonas mobilis Invertase A: Genbank Accession No. AY171597

MESPSYKNLIKAEDAQKKAGKRLLSSEWYPGFHVTPLTGWMNDPNGLIFFKGEYHL
FYQYYPFAPVWGPMPHWHAKSRDLVHWETLPVALAPGDLFDRDGCFSGCAVDNN
GVLTLIYTGHI VLSNDSPDAIREVQCMATSIDGIHFQKEGIVLEKAPMPQVAHFRDPR
VWKENDHWFMMVVG YRTDDEKHQGHGVALYRSENLDWIFVKTL LGDNSQLPLGK
RAFMWECPDFFS LGNRSVLMFSPQGLKASGYKNRNLFQNGYILGKWQAPQFTPETS
FQELDYGHDFYAAQRFEAKDGRQILIAWFDMWENQKPSQRD GWAGCMTLPRKLDL
IDNKIVMTPVREMEILRQSEKIESVVTLSDAEHPFTMDSP LQEIELIFDLEKSSAYQAG
LALRCNGKGQETLLYIDRSQNRILDRNRSGQNVKGIRSCPLPNTSKVRLHIFLDRSSIE
IFVGDDQTQGLYSSSRIFPDKDSLKGRLFAIEGYAVFDSFKRWTLQDANLAAFSSDA
C

SEQ ID NO: 30

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

MAIKTNRQPVEKPPFTIGTLRKAIPAHCFERSALRGRAPDWSMLFAVITTIFSAAEKQ
WTNLEWKPKPNPPQLDDHFGPHGLVFRRTFAIRSYEVGPDRSTSIVAVMNLQEAAL
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
CGACACGTGCTTGGGCGCCTGCCGCCTGCCGTCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCA TGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA

TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTCAGCAAAGGTAGCAGTTT
 GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
 GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
 AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTC
 ATCCAACGTAACCATGAGCTGATCAAACTGCAATCATCGGGCGGGCGTGATGC
 AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
 GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
 CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCTCT
 GGCTGCTTGCTGGGTGTCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
 GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCAACTGT
 CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
 GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
 AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
 GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCATT
 TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
 AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
 TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTCGTGTT
 GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
 CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
 TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
 TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
 CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
 CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
 ACGAGGCCGCCCGCCTGTCCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
 GCTGGACGTGGTGACCGAGGCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
 CGGCCAGGACCTGCTGTCCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
 GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCTTCGACC
 ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
 TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
 TCGCCCGCCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
 GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
 CGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
 CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCGGACCGCTTCCTGGTGCT
 GTACGGCATCGCCGCCCCCGACTCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
 GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACAACTCTGGAC
 GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
 CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
 TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATAACCACCCCGAGCATCCCCCTCC
 CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCC
 CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
 CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
 CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
 GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTTCGCACCTCAGCGCGGCATACAC
 CACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG
 GTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
 GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
 ACTTCCAGCAAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
 ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
 ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGAC
 ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC

ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
 GTTTCAGTCACAACCCGCAAACACTAGTATGGCTATCAAGACGAACAGGCAGCC
 TGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCA
 CTGTTTCGAGCGCTCGGCGCTTCGTGGGCGCGCCCCCGACTGGTCCATGCTGTTT
 GCCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGACCAACCTGGAGT
 GGAAGCCCAAGCCCAACCCCCCCCAGCTGCTGGACGACCACTTCGGCCCCCACG
 GCCTGGTGTTCGCGCGCACCTTCGCCATCCGCAGCTACGAGGTGGGCCCCGACCG
 CTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCCGCCCTGAACCA
 CGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCCTGGAGATGTCC
 AAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCCGTGGAGCGCTAC
 CCCGCCTGGGGCGACACCGTGGAGGTGGAGTGTCTGGGTGGGCGCCTCCGGCAAC
 AACGGCCGCCGCCACGACTTCCTGGTGC GCGACTGCAAGACCGGCCGAGATCCTG
 ACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCGACCCGCCGCCTGAGCA
 AGATCCCCGAGGAGGTGCGCGGCCGAGATCGGCCCCCGCCTTCATCGACAACGTGG
 CCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGACTCCACCGCCC
 ACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACATCAACCAGC
 ACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACCGTGCCCGACAGCA
 TCTTCGAGAGCCACCACATCTCCTCCTTCACCATCGAGTACCGCCGCGAGTGCAC
 CATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCGGCTCCTCCGAGGC
 CGGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCG
 CGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGT
 GATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCACGACGGCGACTACAA
 GGACCACGACATCGACTACAAGGACGACGACGACAAGTGATGACTCGAGGCAGC
 AGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTG
 CCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCC
 TCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCT
 ATTTGCGAATACCACCCCCAGCATCCCCCTTCCTCGTTTCATATCGCTTGCATCCC
 AACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCT
 CACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCA
 ACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACA
 CAAATGGAAAGCTT

SEQ ID NO: 32

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

MAIKTNRQPVEKPPFTIGTLRKAIPAHCFERSALRGRAQLPDWSRLLTAITTVFVSK
 RPDMDHRKSKRPDMLVDSFGLESTVQDGLVFRQSFSIRS YEIGTDRTASIELTMNHLQ
 ETSLNHCKSTGILLDGFGRITLEMCKRDLI WVVIKMQIKVNRYPAWGDTVEINTRFSR
 LGKIGMGRDWLISDCNTGEILVRATSAYAMMNQKTRRLSKLPYEVHQEIVPLFVDS
 P VIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVS NVKYIGWILESMPTVLET
 QELCSLALEYRRECGRDSVLESVTAMDPSKVG VRSQYQHLLRLEDGTAIVNGATEW
 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
 TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTGAGCAAAGGTAGCAGTTT
 GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
 GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
 AGTGTGTGCGAGCCAGCATCGCCGCTGGCTGTTTCAGCTACAACGGCAGGAGTC
 ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
 AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
 GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
 CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCTCT
 GGCTGCTTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
 GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCAGTGTCCACACTGT
 CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
 GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
 AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGGCCACCTTCTTGCG
 GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTTCATT
 TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
 AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCCTACTCCGCATCGATG
 TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGGCGGCTGTTTCTTTCTGTGT
 GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
 CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
 TGCTCGCCCTCTCTCCGTCCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
 TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
 CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCCAGGGC
 CGCCCCGTGCTGTTCTGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
 ACGAGGCCCGCCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
 GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
 CGGCCAGGACCTGCTGTCTTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
 GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCCCTTCGACC
 ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
 TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
 TCGCCCGCCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
 GCGACGCCTGCCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
 CGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
 CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTTGGTGCT
 GTACGGCATCGCCGCCCCCGACTCCCAGCGCATCGCCTTCTACCGCTGTCTGGAC
 GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
 GCTGGTCTGTGTGATGGACTGTTGCCGCCACACTTGTCTGCCTTGACCTGTGAATAT
 CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
 TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCC
 CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCC
 CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
 CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
 CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
 GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
 CACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG
 GTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
 GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
 ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA

ACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
 ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCCGATTGCAAAGAC
 ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
 ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
 GTTTCAGTCACAACCCGCAAACACTAGTATGGCTATCAAGACGAACAGGCAGCC
 TGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCA
 CTGTTTCGAGCGCTCGGCGCTTCGTGGGCGCGCCCAGCTGCCCGACTGGAGCCGC
 CTGCTGACCGCCATCACCACCGTGTTCGTGAAGTCCAAGCGCCCCGACATGCACG
 ACCGCAAGTCCAAGCGCCCCGACATGCTGGTGGACAGCTTCGGCCTGGAGTCCA
 CCGTGCAGGACGGCCTGGTGTTCGCGCAGTCCTTCTCCATCCGCTCCTACGAGAT
 CGGCACCGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAGAC
 CTCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTGGACGGCTTCGGCCGCAACC
 CTGGAGATGTGCAAGCGCGACCTGATCTGGGTGGTGAATGCAAGATGCAGATCAAG
 GTGAACCGCTACCCCGCCTGGGGGCGACACCGTGGAGATCAACACCCGCTTCAGC
 CGCCTGGGCAAGATCGGCATGGGCCGCGACTGGCTGATCTCCGACTGCAACACC
 GGCGAGATCCTGGTGCGCGCCACCAGCGCCTACGCCATGATGAACCAGAAGACC
 CGCCGCCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTGCCCCTGTTTCG
 TGGACAGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGA
 AGACCGGCGACAGCATCCAGAAGGGCCTGACCCCCGGCTGGAACGACCTGGACG
 TGAACCAGCACGTGTCCAACGTGAAGTACATCGGCTGGATCCTGGAGAGCATGC
 CCACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCCCTGGCCCTGGAGTACCGCC
 GCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGGACCCAGCA
 AGGTGGGCGTGCGCTCCAGTACCAGCACCTGCTGCGCCTGGAGGACGGCACCG
 CCATCGTGAACGGCGCCACCGAGTGGCGCCCCAAGAACGCCGGCGCCAACGGCG
 CCATCTCCACCGGCAAGACCAGCAACGGCAACTCCGTGTCCATGGACTACAAGG
 ACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGAC
 AAGTGACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGG
 TCGTGTGATGGACTGTTGCCGCCACACTTGTGCTTACCTGTGAATATCCCTGC
 CGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAG
 TTGCTAGCTGCTTGTGCTATTTGCGAATAACACCCCCAGCATCCCCCTCCCTCGTT
 TCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGC
 GCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCTG
 TATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAA
 GTAGTGGGATGGGAACACAAATGGAAAGCTT

SEQ ID NO: 34

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

MAIKTNRPVEKPPFTIGTLRKAIPAHCFERSALRGRAPDWSMLFAVITTIFSAAEKQ
 WTNLEWKPKPLPQLDDHFGLHGLVFRRTFAIRSYEVGPDRSTSILAVMNHMQEA
 TLNHAKSVGILGDGFGTTLEMSKRDLMWVRRTHVAVERYPTWGDVEVECWIGA
 SGNNGMRRDFLVRDCKTGEILTRCTSLVLMNTRTRRLSTIPDEVGRGEIGPAFIDNVA
 VKDDEIKKLQKLNSTADYIQGGLTPRWNDLDVNQHVNNLKYVAWVFETVPDSIFE
 SHHISFTLEYRRECTRDSVLRSLTTVSGGSSEAGLVCDHLLQLEGGSEVLRARTEWR
 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.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCCTGCCGCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAAGCTTATGTG
TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTGAGCAAAGGTAGCAGTTT
GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTC
ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCTCT
GGCTGCTTGTGCTGGGTTCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCAGTGTCCACACTGT
CACCAGGCCCCGAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCATT
TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCGCAATGG
AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTCGTGTT
GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCGCCCGCCTGTCCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCCCTTCGACC
ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCGGACCGCTTCCCTGGTGCT
GTACGGCATCGCCGCCCCCGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATAACCACCCCGAGCATCCCCCTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
CACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG

GTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
 GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
 ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
 ACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
 ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCGGATTGCAAAGAC
 ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
 ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
 GTTTCAGTCACAACCCGCAAACACTAGTATGGCTATCAAGACGAACAGGCAGCC
 TGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCA
 CTGTTTCGAGCGCTCGGCGCTTCGTGGGCGCGCCCCGACTGGTCCATGCTGTTT
 GCCGTGATCACCACCATCTTCAGCGCCGCCGAGAAGCAGTGGACCAACCTGGAG
 TGGAAGCCCAAGCCCAAGCTGCCCCAGCTGCTGGACGACCACTTCGGCCTGCAC
 GGCTGGTGTTCGCGCCACCTTCGCCATCCGCTCCTACGAGGTGGGCCCCGACC
 GCAGCACCTCCATCCTGGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACC
 ACGCCAAGAGCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCCTGGAGATGT
 CCAAGCGCGACCTGATGTGGGTGGTGCGCCGCACCCACGTGGCCGTGGAGCGCT
 ACCCCACCTGGGGCGACACCGTGGAGGTGGAGTGTGGATCGGCGCCAGCGGCA
 ACAACGGCATGCGCCGCGACTTTCCTGGTGCGCCGACTGCAAGACCGGCGAGATCC
 TGACCCGCTGCACCTCCCTGAGCGTGTGATGAACACCCGCACCCGCCGCTGAG
 CACCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCCGCTTCATCGACAACGT
 GGCCGTGAAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGC
 GACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACGTGAACCA
 GCACGTGAACAACCTGAAGTACGTGGCCTGGGTGTTTCGAGACCGTGCCCCGACAG
 CATCTTCGAGTCCCACCACATCAGCTCCTTCACCCTGGAGTACCGCCGCGAGTGC
 ACCCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCAGCTCCGAG
 GCCGGCCTGGTGTGCGACCACTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTG
 CGCGCCCCGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCCGCGGCATCAGC
 GTGATCCCCGCCGAGCCCCGCGTGTGACTACAAGGACCACGACGGCGACTAC
 AAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGATGACTCGAGGC
 AGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACT
 GTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAAC
 AGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTG
 TGCTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCA
 TCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCC
 TGCTCACTGCCCCTCGCACAGCCTTGTTTGGGCTCCGCCTGTATTCTCCTGGTAC
 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.

MAIKTNRQPVEKPPFTIGTLRKAIPAHCFERSALRGRAPANGSAVTLKSGSLNTQEDT
 LSSSPPPRAFFNQLPDWSMLLTAITTVFVAPEKRWTMFDRKSKRPNMLMDSFGLERV
 VQDGLVFRQSFSIRSIEICADRTASIEVTMNHVQETSLNQCKSIGLLDDGFGRSP EMC
 KRDLIWVVTRMKIMVNRYPWTGDTIEVSTWLSQSGKIGMGRDWLISDCNTGEILVR
 ATSVYAMMNQKTRRFSLPHEVRQEFAPHFLDSPPAIEDNDGKLQKFDVKTGDSIRK
 GLTPGWYDLVDNQHVS NVKYIGWILESMPTEVLETQELCSLTLEYRRECGRDSVLES
 VTSMDPSKVGD RFQYRHLLRLEDGADIMKGRTEWRPKNAGTNGAISTGKT

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.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCCTGCCTGCCGCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTCAGCAAAGGTAGCAGTTT
GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTC
ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCCTCT
GGCTGCTTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCAGAGTGTCACACTGT
CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCATT
TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCCTACTCCGCATCGATG
TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCGGCTGTTTCTTTCTGTGT
GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTTCGGCTACGACTGGGC
CCAGCAGACCATCGGCTGCTCCGACGCCGCGCTGTTCCGCTGTCCGCCCAGGGC
CGCCCCGTGCTGTTCTGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCCGCCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCTTCGACC
ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCCTGGTGCT
GTACGGCATCGCCGCCCCCGACTCCAGCGCATCGCCTTCTACCGCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATAACACCCCCAGCATCCCCTTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA

CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
 GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTTCGCACCTCAGCGCGGCATACAC
 CACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG
 GTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
 GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
 ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
 ACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
 ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCGGATTGCAAAGAC
 ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
 ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
 GTTTCAGTCACAACCCGCAAACACTAGTATGGCTATCAAGACGAACAGGCAGCC
 TGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCA
 CTGTTTCGAGCGCTCGGCGCTTCGTGGGCGCGCCCCCGCGAACGGCAGCGCGGTG
 ACCCTGAAGTCGGGCTCCCTGAACACCCAGGAGGACACGCTGAGCTCGTCCCCC
 CCCCCCGCGCGTTCTTCAACCAGCTGCCCCACTGGAGCATGCTGCTGACCGCGA
 TCACCACGGTCTTCGTGGCGCCCCGAGAAGCGCTGGACCATGTTTCGACCGCAAGTC
 GAAGCGCCCCAACATGCTGATGGACTCCTTCGGCCTGGAGCGCGTGGTCCAGGA
 CGGCCTGGTGTTCGCGCCAGAGCTTCTCGATCCGCTCCTACGAGATCTGCGCGGAC
 CGCACCGCGAGCATCGAGACGGTGATGAACCACGTCCAGGAGACCTCGCTGAAC
 CAGTGCAAGTCCATCGGCCTGCTGGACGACGGCTTCGGCCGCGACCCCCGAGATG
 TGCAAGCGCGACCTGATCTGGGTGGTCACCCGCATGAAGATCATGGTGAACCGC
 TACCCACGTGGGGCGACACCATCGAGGTCTCGACGTGGCTGTCCAGAGCGGC
 AAGATCGGCATGGGGCCGCGACTGGCTGATCTCGGACTGCAACACCGGCGAGATC
 CTGGTGC GCGGACGTCCGTCTACGCGATGATGAACCAGAAGACCCGCCGCTTC
 AGCAAGCTGCCCCACGAGGTGCGCCAGGAGTTCGCGCCCCACTTCCTGGACTCGC
 CCCCCGCGATCGAGGACAACGACGGCAAGCTGCAGAAGTTCGACGTCAAGACGG
 GCGACTCCATCCGCAAGGGCCTGACCCCCGGCTGGTACGACCTGGACGTGAACC
 AGCACGTGAGCAACGTCAAGTACATCGGCTGGATCCTGGAGTCGATGCCACCG
 AGGTCCTGGAGACGCGAGGAGCTGTGCTCCCTGACCCTGGAGTACCGCCGCGAGT
 GCGGCCGCGACTCGGTGCTGGAGAGCGTCACCAGCATGGACCCCTCGAAGGTGG
 GCGACCGCTTCAGTACCGCCACCTGCTGCGCCTGGAGGACGGCGCGGACATCA
 TGAAGGGCCGCAACGAGTGGCGCCCCAAGAACGCGGGCACGAACGGCGCGATCT
 CCACCGGCAAGACGTGACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACT
 CTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGT
 GAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACG
 CGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCC
 CCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTG
 CTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGTTTT
 GGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCT
 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.

MAIKTNRPVEKPPFTIGTLRKAIPAHCFERSALRGRAPANGSAVNLSGSLNTQEDT
 SSSPPPRAFNLNQLPDWSMLLTAITTVFVAAEKQWTMLDRKSKRPDMLVDSVGLKSIV
 RDGLVSRQSFLIRSYEIGADRTASIEITLMNHLQETSINHCKSLGLLNDGFGRTPGMCK
 NDLIWVLTKMQIMVNRYPWTGDTVEINTWFSQSGKIGMASDWLISDCNTGEILIRAT
 SVWAMMNQKTRRFSRLPYEVRQELTPHFVDSPHVIEDNDQKLHKFDVKTGDSIRKG

LTPRWNDLDVNQHVSNNVKYIGWILESMPIEVLETQELCSLTVEYRRECGMDSVLESV
TAVDPSENGGRSQYKHLRLLEDGTDIVKSRTTEWRPKNAGTNGAISTSTAKTSNGNSA
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.

ACTAGTATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTACAG
ATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTC
GTGGGCGCGCCCCCGCGAACGGCAGCGCGGTGAACCTGAAGTCGGGGCTCCCTGA
ACACCCAGGAGGACACGAGCTCGTCCCCCCCCCCCCCGCGCGTTCCTGAACCAGCT
GCCCCGACTGGAGCATGCTGCTGACCGCGATCACCACCGTCTTCGTGGCGGCGGA
GAAGCAGTGGACGATGCTGGACCGCAAGTCGAAGCGCCCCGACATGCTGGTGGA
CTCCGTCGGCCTGAAGAGCATCGTGCGCGACGGCCTGGTCTCGCGCCAGTCCTTC
CTGATCCGCGAGCTACGAGATCGGCGCGGACCGCACCGCGTCGATCGAGACCCTG
ATGAACCACCTGCAGGAGACGTCCATCAACCACTGCAAGAGCCTGGGCCTGCTG
AACGACGGCTTCGGCCGCACCCCCGGCATGTGCAAGAACGACCTGATCTGGGTG
CTGACCAAGATGCAGATCATGGTCAACCGCTACCCACGTTGGGGCGACACCGTC
GAGATCAACACGTGGTTCTCGCAGTCCGGCAAGATCGGCATGGCGAGCGACTGG
CTGATCTCGGACTGCAACACCGGCGAGATCCTGATCCGCGCGACCTCCGTGTGGG
CGATGATGAACCAGAAGACGCGCCGCTTCAGCCGCCTGCCCTACGAGGTCCGCC
AGGAGCTGACCCCCCACTTCGTGGACTCGCCCCACGTCATCGAGGACAACGACC
AGAAGCTGCACAAGTTCGACGTGAAGACCGGCGACTCCATCCGCAAGGGCCTGA
CGCCCCGCTGGAACGACCTGGACGTCAACCAGCACGTGTGCAACGTGAAGTACA
TCGGCTGGATCCTGGAGTCCATGCCATCGAGGTCTTGAGACCCAGGAGCTGTG
CTCGCTGACCGTGGAGTACCGCCGCGAGTGCGGCATGGACTCCGTGCTGGAGTC
GGTCACGGCGGTGGACCCAGCGAGAACGGCGGCCGCGAGCCAGTACAAGCACCT
GCTGCGCCTGGAGGACGGCACCGACATCGTCAAGTCGCGCACCGAGTGGCGCCC
CAAGAACGCGGGCACGAACGGCGCGATCTCCACCAGCACCGCGAAGACGTGAA
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.

MATASTFSAFNARCGDLRRSAGSGPRRPAPLPVRGRAAQAATRVNGSKVGLKTD
NKLEDAPFIPSSAPRTFYNQLPDWSVLLAAITTIFFLAEEKQWTLIDWKRGGPDMLSDA
FGLPKIENGLLYRQKFSIRSIEIGADQTASIEITLMNHLQETALNHVKCAGLLGNGFG
STPEMSKMNLIVVVTMQLVVEHYPSWGDVIEVDTWAAASGKNGMRRDWHVRD
WQTGQTIMRASSNWVMMNQNRRLSKFPPEEVRAEIEPYFMERAPVIDDDNRKLPKL
DDDTADHVRNGLTPRWSLDVNQHVKNVKYIGWILESAPISILESHELASMTLEYRR
ECGRDSVLQSLTSVSNCTDGSEELPIECQHLLRNEGGSEIVKGRTEWRPKKCGPFGA
GRP

SEQ ID NO: 41

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

ACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGGCGACC
TGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCCCTCCCCGTGCG
CGGGCGCGCCGCCAGGCGGCCACCCGCGTGAACGGCAGCAAGGTGGGCCTGAA
GACCGACACCAACAAGCTGGAGGACGCGCCCTTCATCCCCCTCGTCCGCCCCCGC
ACCTTCTACAACCAGCTGCCCCGACTGGAGCGTCCTGCTGGCGGGCCATCACCACCA
TCTTCCTGGCGGCCGAGAAGCAGTGGACCCTGATCGACTGGAAGCGCGGGCGGCC
CCGACATGCTGTGCGGACGCGTTTCGGCCTGCCCAAGATCATCGAGAACGGCCTGCT
GTACCGCCAGAAGTTCTCCATCCGCAGCTACGAGATCGGCGCCGACCAGACCGC
CTCGATCGAGACCCTGATGAACCACCTGCAGGAGACCGCGCTGAACCACGTCAA
GTGCGCCGGCCTGCTGGGCAACGGCTTCGGCTCCACCCCCGAGATGAGCAAGAT
GAACCTGATCTGGGTGGTCACCAAGATGCAGGTGCTGGTTCGAGCACTACCCCTCG
TGGGGCGACGTGATCGAGGTGGACACCTGGGCGGCCGCGTCCGGCAAGAACGGC
ATGCGCCGCGACTGGCACGTCCGCGACTGGCAGACCGGCCAGACCATCATGCGC
GCCAGCTCGAACTGGGTGATGATGAACCAGAACACCCGCCGCTGTCCAAGTTC
CCCGAGGAGGTCCGCGCCGAGATCGAGCCCTACTTCATGGAGCGCGCCCCCGTG
ATCGACGACGACAACCGCAAGCTGCCCAAGCTGGACGACGACACCGCGGACCAC
GTGCGCAACGGCCTGACCCCCCGCTGGAGCGACCTGGACGTGAACCAGCACGTC
AAGAACGTGAAGTACATCGGCTGGATCCTGGAGTCGGCCCCCATCTCCATCCTGG
AGAGCCACGAGCTGGCCTCGATGACCCTGGAGTACCGCCGCGAGTGC GGCCGCG
ACTCCGTCCTGCAGAGCCTGACCTCGGTGTCCAACAACCTGCACCGACGGCAGCG
AGGAGCTGCCCATCGAGTGCCAGCACCTGCTGCGCAACGAGGGCGGCTCGGAGA
TCGTCAAGGGCCGCACCGAGTGGCGCCCCAAGAAGTGC GGCCCCCTTCGGCGCCG
GCCGCCCCCTGACTCGAG

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
ESTKDEDIVAAPTVA PKRTFINQLPDWSMLLAAITTIFLAAEKQWTNLDWKPRRPDM
LVDFDPFSLGRFVQDGLIFRQNFIRSIEIGADRTASIEITLMNHLQETALNHVRCIGLL
DDGFGSTPEMTRRDLIWVVTRMQVLVD RYPSWGDVIEVDSWVTPSGKNGMKREWF
LRDCKTGEILTRATSVWMMNKRTRRLSKIPEEVRVEIEPYFVEHGVLEDEDSRKLPK
LNDNTANYIRRGLAPRWSDLDVNQHVN NVKYIGWILESVPSSLLESHELYGMTLEY
RKECGKDGLQLSLTAVASDYGGGSLEAGVECDHLLRLEDGSEIMRGKTEWRPKRAA
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.

ACTAGTATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTACAG
ATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTC
GTGGGCGCGCCGCCAACGCCACACCGTGCCCAAGATCAACGGCAACAAGGCCG
GCCTGCTGACCCCCATGGAGAGCACCAAGGACGAGGACATCGTTCGCGGCCCCCA
CCGTGGCGCCCAAGCGCACCTTCATCAACCAGCTGCCCCGACTGGTTCGATGCTGCT
GGCCGCGATCACCACCATCTTCCTGGCGGCCGAGAAGCAGTGGACCAACCTGGA
CTGGAAGCCCCGCCGCCCGACATGCTGGTTCGACTTCGACCCCTTCTCCCTGGGC
CGCTTCGTGCAGGACGGCCTGATCTTCGCCAGAACTTCAGCATCCGCTCGTACG
AGATCGGCGCGGACCGCACCGCCTCCATCGAGACCCTGATGAACCACCTGCAGG

AGACCGCGCTGAACCACGTCCGCTGCATCGGCCTGCTGGACGACGGCTTCGGCA
GCACCCCGAGATGACCCGCCGCGACCTGATCTGGGTGGTCACCCGCATGCAGG
TCCTGGTGGACCGCTACCCCTCGTGGGGCGACGTGATCGAGGTCGACTCCTGGGT
GACCCCGAGCGGCAAGAACGGCATGAAGCGCGAGTGGTTCCTGCGCGACTGCAA
GACCGGCGAGATCCTGACCCGCGCCACCTCGGTCTGGGTGATGATGAACAAGCG
CACCCGCCGCTGTCCAAGATCCCCGAGGAGGTCCGCGTGGAGATCGAGCCCTA
CTTCGTCGAGCACGGCGTGCTGGACGAGGACTCGCGCAAGCTGCCCAAGCTGAA
CGACAACACCGCCAACACTACATCCGCCGCGGCCTGGCGCCCCGCTGGTCCGACCTG
GACGTCAACCAGCACGTGAACAACGTCAAGTACATCGGCTGGATCCTGGAGAGC
GTGCCCAGCAGCCTGCTGGAGTCGCACGAGCTGTACGGCATGACCCTGGAGTAC
CGCAAGGAGTGCGGCAAGGACGGCCTGCTGCAGTCCCTGACCGCCGTCGCCAGC
GACTACGGCGGCGGCTCGCTGGAGGCGCGGTGGAGTGCGACCACCTGCTGCGC
CTGGAGGACGGCTCCGAGATCATGCGCGGCAAGACCGAGTGGCGCCCCAAGCGC
GCCGCGAACACCACCTACTTCGGCAGCGTCGACGACATCCCCCCCCGCCAACAAAC
GCGTGA CT CGAG

SEQ ID NO: 44

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

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRASMLLSAVTTVFGVAEKQW
PMLDRKSKRPDMLVEPLGVDRIVYDGVSRQSFIRSIEIGADRTASIELMNMFQET
SLNHCKIIGLLNDGFGRTPEMCKRDLIWVVTKMQIEVNRYPYTWGDTIEVNTWVSASG
KHGMGRDWLISDCHTGEILIRATSVWAMMNQKTRRLSKIPYEVQRQIEPQFVDSAPV
IVDDRKFHKLDLKTGDSICNGLTPRWTDLDVNQHVNNVKYIGWILQSVPTVEVFETQE
LCGLTLEYRRECCRDSVLESVTAMDPSKEGDRSLYQHLLRLEDGADIVKGRTEWRP
KNAGAKGAILTGKTSNGNSIS

SEQ ID NO: 45

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

ACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACC
TGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCG
CGGGCGCGCCAGCATGCTGCTGTGCGGCGGTGACCACGGTCTTCGGCGTGGCCGA
GAAGCAGTGGCCCATGCTGGACCGCAAGTCCAAGCGCCCCGACATGCTGGTCTGA
GCCCCCTGGGCGTGACCGCATCGTCTACGACGGCGTGAGCTTCCGCCAGTCGTTC
TCCATCCGCAGCTACGAGATCGGCGCCGACCGCACCGCCTCGATCGAGACGCTG
ATGAACATGTTCCAGGAGACCTCCCTGAACCACTGCAAGATCATCGGCCTGCTGA
ACGACGGCTTCGGCCGCACGCCCAGATGTGCAAGCGCGACCTGATCTGGGTCTG
TGACCAAGATGCAGATCGAGGTGAACCGCTACCCACGTGGGGCGACACCATCG
AGGTCAACACGTGGGTGAGCGCCTCGGGCAAGCACGGCATGGGCCGCGACTGGC
TGATCTCCGACTGCCACACCGGCGAGATCCTGATCCGCGCGACGAGCGTCTGGGC
GATGATGAACCAGAAGACCCGCCGCTGTGCAAGATCCCCACGAGGTGCGCCA
GGAGATCGAGCCCCAGTTCGTCTGACTCCGCCCCCGTGATCGTGGACGACCGCAA
GTTCCACAAGCTGGACCTGAAGACGGGCGACAGCATCTGCAACGGCCTGACCCC
CCGCTGGACGGACCTGGACGTGAACCAGCACGTCAACAACGTGAAGTACATCGG
CTGGATCCTGCAGTCGGTCCCCACCGAGGTGTTTCGAGACGCAGGAGCTGTGCGG
CCTGACCCTGGAGTACCGCCGCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCGT
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.

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAQLPDWSMLLAITTLFLAA
EKQWMMLDWKPKRPDMLVDPFGLGRFVQDGLVFRNNFSIRS YEIGADRTAS IETLM
NHLQETALNHVKS VGLLEDGLGSTREMSLRNLIWVVTKMQVAVD RYPTWGDDEVQV
SSWATAIGKNGMRREWIVTDFRTGETLLRATSVWVM MNKLTRRISKIPEEVWHEIGP
SFIDAPPLPTVEDDGRKLTRFDESSAD FIRKGLTPRWSDLDINQHVNNVKYIGWLLES
APPEIHESHEIASLTLEYRRECGRDSVLNSATKVS DSSQLGKSAVECNHLVRLQNGGE
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
TGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCG
CGGGCGCGCCCAGCTGCCCCGACTGGAGCATGCTGCTGGCCGCGATCACCACCCT
GTTCTGGCGGCCGAGAAGCAGTGGATGATGCTGGACTGGAAGCCCAAGCGCCC
CGACATGCTGGTGGACCCCTTCGGCCTGGGCGCGTTCGTGCAGGACGGCCTGGTG
TTCCGCAACAACCTTCAGCATCCGCAGCTACGAGATCGGCGCGGACCGCACCGCC
AGCATCGAGACCCTGATGAACCACCTGCAGGAGACCGCCCTGAACCACGTGAAG
AGCGTGGGCCTGCTGGAGGACGGCCTGGGCAGCACCCGCGAGATGAGCCTGCGC
AACCTGATCTGGGTGGTGACCAAGATGCAGGTGGCGGTGGACCGCTACCCACC
TGGGGCGACGAGGTGCAGGTGAGCAGCTGGGCGACCGCCATCGGCAAGAACGG
CATGCGCCGCGAGTGGATCGTGACCGACTTCCGCACCGGCGAGACCCTGCTGCG
CGCCACCAGCGTGTGGGTGATGATGAACAAGCTGACCCGCCGCATCAGCAAGAT
CCCCGAGGAGGTGTGGCACGAGATCGGCCCCAGCTTCATCGACGCGCCCCCCT
GCCACCGTGGAGGACGACGGCCGCAAGCTGACCCGCTTCGACGAGAGCAGCGC
CGACTTCATCCGCAAGGGCCTGACCCCCCGCTGGAGCGACCTGGACATCAACCA
GCACGTGAACAACGTGAAGTACATCGGCTGGCTGCTGGAGAGCGCGCCCCCGA
GATCCACGAGAGCCACGAGATCGCCAGCCTGACCCTGGAGTACCGCCGCGAGTG
CGGCCGCGACAGCGTGCTGAACAGCGCCACCAAGGTGAGCGACAGCAGCCAGCT
GGGCAAGAGCGCCGTGGAGTGCAACCACCTGGTGCGCCTGCAGAACGGCGGCGA
GATCGTGAAGGGCCGCACCGTGTGGCGCCCCAAGCGCCCCCTGTACAACGACGG
CGCCGTGGTGGACGTGCCCCGCCAAGACCAGCTGACTCGAG

SEQ ID NO: 48

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

ACTAGTATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTACG
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

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

GCTCTTCGGCCGCCGCGCCACTCCTGCTCGAGCGCGCCCCGACTCGCGCTCCGCCTGC
GCCCCGCGCTGCGCCGCCAGCGCCTTGGCCTTTTCGCCGCGCTCGTGCGCGTTCG
TGATGTCCATCACCAGGTCCATGAGGTCTGCCTTGCGCCGGCTGAGCCACTGCTT
CGTCCGGGCGGCCAAGAGGAGCATGAGGGAGGACTCCTGGTCCAGGGTCCCTGAC
GTGGTCGCGGCTCTGGGAGCGGGCCAGCATCATCTGGCTCTGCCGCACCGAGGC
CGCCTCCAACCTGGTCTCCAGCAGCCGCAGTCGCCGCCGACCCTGGCAGAGGAA
GACAGGTGAGGGGGGTATGAATTGTACAGAACAACCACGAGCCTTGTCTAGGCA
GAATCCCTACCAGTCATGGCTTTACCTGGATGACGGCCTGCGAACAGCTGTCCAG
CGACCCTCGCTGCCGCCGCTTCTCCCGCACGCTTCTTTCCAGCACCGTGATGGCG
CGAGCCAGCGCCGCACGCTGGCGCTGCGCTTCGCCGATCTGAGGACAGTCGGGG
AACTCTGATCAGTCTAAACCCCCCTTGCGCGTTAGTGTTGCCATCCTTTGCAGACC
GGTGAGAGCCGACTTGTGTGCGCCACCCCCACACCACCTCCTCCCAGACCAAT
TCTGTACCTTTTTTGGCGAAGGCATCGGCCTCGGCCTGCAGAGAGGACAGCAGTG
CCCAGCCGCTGGGGGTGGGCGGATGCACGCTCAGGTACCCTTTCTTGCGCTATGA
CACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGC
AACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCC
GATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGA
CATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTA
CACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTT
CGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTC
CTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCC
GACCGCCCCCTGGTGCACTTCACCCCCAACAAGGGCTGGATGAACGACCCCAAC
GGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAAC
CCGAACGACACCGTCTGGGGGACGCCCTTGTCTGGGGCCACGCCACGTCCGAC
GACCTGACCAACTGGGAGGACCAGCCATCGCCATCGCCCCGAAGCGCAACGAC
TCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCT
TCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACA
CCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTT
CACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGA
CCGAAGGTCTTCTGGTACGAGCCCTCCAGAAAGTGGATCATGACCGCGGCCAA
GTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAGTCTTGAA
GCTGGAGTCCGCGTTCGCCAACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCC
GGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCTTACTGGGTGATG
TTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCG
TCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGT
GGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACC
TACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCTTCG
TGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAAC

ACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCG
ATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACG
TTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACAGCACCCGGCACCCCTG
GAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGT
TCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCG
CATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCTTCTGGACCGCGGGAACAGCAAG
GTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAAC
CAGCCCTTCAAGAGCGAGAACGACCTGTCTTACTACAAGGTGTACGGCTTGCTGG
ACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACA
CCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGG
TGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATT
GGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTTCGTGTGATGG
ACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCA
AACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGC
TTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTT
GCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGC
TCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCTGTATTCTCCTGG
TACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGCGAT
GGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGCGCGCAGAGGAACGCT
GAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACACCACAATAACCACCTGA
CGAATGCGCTTGTTCTTCGTCCATTAGCGAAGCGTCCGGTTCACACACGTGCCA
CGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCTGATGGTCGAAACGTTT
ACAGCCTAGGGATATCGAATTCCTTTCTTTCGCTATGACACTTCCAGCAAAAGGT
AGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTC
GACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCG
AGCGCTGTTTAAATAGCCAGGCCCCCCGATTGCAAAGACATTATAGCGAGCTACC
AAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCT
TGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCC
GCAAACACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCG
GCGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCC
CCGTGCGCGGGCGCGCCCAGCTGCCCGACTGGAGCATGCTGCTGGCCGCGATCA
CCACCCTGTTCTTGGCGGCCGAGAAGCAGTGGATGATGCTGGACTGGAAGCCCA
AGCGCCCCGACATGCTGGTGGACCCCTTCGGCCTGGGCGGCTTCGTGCAGGACGG
CCTGGTGTTCGCAACAACCTCAGCATCCGCAGCTACGAGATCGGCGCGGACCGC
ACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAGACCGCCCTGAACCAC
GTGAAGAGCGTGGGCCTGCTGGAGGACGGCCTGGGCAGCACCCGCGAGATGAGC
CTGCGCAACCTGATCTGGGTGGTGACCAAGATGCAGGTGGCGGTGGACCGCTAC
CCCACCTGGGGCGACGAGGTGCAGGTGAGCAGCTGGGCGACCGCCATCGGCAAG
AACGGCATGCGCCGCGAGTGGATCGTGACCGACTTCCGCACCGGCGAGACCCTG
CTGCGCGCCACCAGCGTGTGGGTGATGATGAACAAGCTGACCCGCCGATCAGC
AAGATCCCCGAGGAGGTGTGGCACGAGATCGGCCCCAGCTTCATCGACGCGCCC
CCCCTGCCACCGTGGAGGACGACGGCCGCAAGCTGACCCGCTTCGACGAGAGC
AGCGCCGACTTCATCCGCAAGGGCCTGACCCCCCGCTGGAGCGACCTGGACATC
AACCAGCACGTGAACAACGTGAAGTACATCGGCTGGCTGCTGGAGAGCGCGCCC
CCCGAGATCCACGAGAGCCACGAGATCGCCAGCCTGACCTGGAGTACCGCCGC
GAGTGCGGCCGCGACAGCGTGCTGAACAGCGCCACCAAGGTGAGCGACAGCAGC
CAGCTGGGCAAGAGCGCCGTGGAGTGCAACCACCTGGTGCGCTGCAGAACGGC
GGCGAGATCGTGAAGGGCCGCACCGTGTGGCGCCCCAAGCGCCCCCTGTACAAC
GACGGCGCCGTGGTGGACGTGCCCGCCAAGACCAGCGATGACGATGACAAGCTG
GGATGACTCGAGTTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCGACAC

ACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACC
TGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGT
ACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGC
ATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGT
CCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTG
GTTTGGGCTCCGCCTGTATTTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAA
TGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTGTAGAGCT
CCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGATAACC
TCCAAAGCCGCTCTAATTGTGGAGGGGGTTTGAATTTAAAAGCTTGGAATGTTGG
TTCGTGCGTCTGGAACAAGCCCAGACTTGTGCTCACTGGGAAAAGGACCATCAG
CTCCAAAAAATTGCCGCTCAAACCGCGTACCTCTGCTTTCGCGCAATCTGCCCT
GTTGAAATCGCCACCACATTCATATTTGTGACGCTTGAGCAGTCTGTAATTGCCTC
AGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCGCGGGC
GAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAGTTCATA
ACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAGTGGC
CACCCCCCGGCCCTGGTGCTTGCGGAGGGGCAGGTCAACCGGCATGGGGCTACCG
AAATCCCCGACCGGATCCCACCACCCCCGCGATGGGAAGAATCTCTCCCCGGGA
TGTGGGCCCACCACCAGCACAACTGCTGGCCCAGGCGAGCGTCAAACCATAACC
ACACAAATATCCTTGGCATCGGCCCTGAATTCCTTCTGCCGCTCTGCTACCCGGT
GCTTCTGTCCGAAGCAGGGGTTGCTAGGGATCGCTCCGAGTCCGCAAACCCCTGT
CGCGTGGCGGGGCTTGTTTCGAGCTTGTTTCGAGCTTGAAGAGCCTCTAGAGTCGAC
CTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAAT
TGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA
GCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGC
CCGCTTTCAGTCGGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACG
CGCGGGGAGAGGCGGTTTTCGCTATTGGGCGCTCTTCC

SEQ ID NO: 51

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

GAATTCGCCCTCCCGTGATCACACAGGTGCCTTGCGAGCGTGATCACACTATTTT
GGGGGTCCTACAGTACTGAAATGGTGAGAAGTCGTACTGAAATCAAGGATGAAC
AATGAAAATGGTGCTGTGGTGGCTTCTCAAAGGTCAAGAATCAGTCGCTCGCGTC
AGGAAATCGCGGCGTCAACCAGCGTGGGCGCGGTCACTGGCCCCGCACTGGTCA
CCATAGCCTCTCCTGCCACAGTAGCGATCCCCTGGGCGTTCCTCTCAGCAGCGG
CTGTACTGCCTCCAGATTTTCTTCTTCTGGACCTGCGGGCGTGAGAGGATGAGC
AGGGTGGGCCAAGGGCTCAATCCTGAACGGGCCCTCATTCGGTTTCCAATCCACA
ACACATACCCACAGCAGGTGAGACCACGCATTCCACCATGCGCACCAATAACGT
GTCCTTACCTGATTGGGTGTGGCAGGCTCCGTGGACAGGAGTGCTCGTCCCCCG
CCCAGACCCGCTCCCCCGTCACGGCGGCGTCCGGGACCCGCGAGCGGCTCCACCG
CGGTGTGATCCGCGTTGGCGGCGCAGAGCAGCATCCCAGCCGATTTGACCCCGC
GCATGCTCCGAGGCTTGAGGTTGGCCAGCACCACCACCCGCCGGCCGACAAGGT
CCTCCAGGGTACGTGCCGGACCAGGCCACTCACGATGGTGCGAGGGCCCCCT
CCTCGCCGAGGTGATCTGCTCGACGTACAGACTGCGACATGCGTGCGGAGTGGT
CATCAGAAGGAAGCAGGTGTGAGAAGGGGCACGTGGTTGGTATTGAGAGTAGC
CAAAGCTTTGTGCCAATCAGAAAGTCAACGCAGCTGCCTGCCTGGCTCGCGTACA
ATTCCTTTCTTGCCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACG
GCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTC
GGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCC

AGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCATATTCAAACACC
TAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAA
GGGGGCGCCTCTTCCCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTG
CTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCA
TGACGAACGAGACGTCCGACCGCCCCCTGGTGCACCTTCACCCCCAACAAAGGGCT
GGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACC
TGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTCTGGGG
CCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGC
CCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAAC
AACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCA
TCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGG
ACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACT
CCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGAT
CATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGA
CCTGAAGTCCTGGAAGCTGGAGTCCGCGTTTCGCCAACGAGGGCTTCCTCGGCTAC
CAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAG
TCCTACTGGGTGATGTTTCATCTCCATCAACCCCGCGCCCCGGCCGGCGGCTCCT
TCAACCAGTACTTCGTCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAA
CCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTC
AACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGG
GAGTACTCCGCCTTCGTGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGC
GCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCA
ACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGT
TCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCA
ACAGCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGA
CGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGA
CCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCCTCCTTCTTCCTG
GACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAAC
CGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGAGAACGACCTGTCTACTAC
AAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGC
GACGTGCTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCG
TGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCCAGGTGC
GCGAGGTCAAGTGATTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCGAC
AACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGA
CCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGT
GTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCA
GCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTG
TCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTT
GGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCA
ATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTC
CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACACTAGTATGACGTTCCG
GGTCGCCCTCCCGGCCATGGGCCGCGGTGTCTCCCTTCCCCGGCCAGGGTCGCG
GTGCGCGCCAGTCGGCGAGTCAGGTTTTGGAGAGCGGGCGCGCCCCCGACTGG
TCCATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGA
CCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCAGCTGCTGGACGACCACT

TCGGCCCCACGGCCTGGTGTTCGCGCCGACCTTCGCCATCCGCAGCTACGAGGT
 GGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGC
 CGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCAC
 CCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGC
 CGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTGGAGTGCTGGGTGGG
 CGCCTCCGGCAACAACGGCCGCCGACACTTCCTGGTGC GCGACTGCAAGAC
 CGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACC
 CGCCGCCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCCGCCTTC
 ATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAA
 CGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCGCTGGAACGACCT
 GGACATCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGAC
 CGTGGCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTACCATCGAGTAC
 CGCCGCGAGTGACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGC
 GGCTCCTCCGAGGCCGGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGC
 AGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTC
 CGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCAC
 GACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTG
 ACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTG
 TGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTT
 TTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCT
 AGCTCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATAT
 CGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCT
 CCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTGTTGGGCTCCGCCTGTATTCT
 CCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGT
 GGGATGGGAACACAAATGGAAAGCTGGTACCCGTACCCATCAGCATCCGGGTGA
 ATCTTGGCCTCCAAGATATGGCCAATCCTCACATCCAGCTTGGCAAAATCGACTA
 GACTGTCTGCAAGTGGGAATGTGGAGCACAAGGTTGCTTGTAGCGATCGACAGA
 CTGGTGGGGTACATTGACAGGTGGGCAGCGCCGCATCCATCGTGCCTGACGCGA
 GCGCCGCCGGTTGCTCGCCCGTGCTGCGGTCAAAGAGCGGCAGAGAAATCGGG
 AACCAGAAAACGTCACATTGCCTGATGTTGTTACATGCTGGACTAGACTTTCTTG
 CGTGGGTCTGCTCCTCGCCAGGTGCGCGACGCCTCGGGGCTGGGTGCGAGGGAG
 CCGTGCGGCCACGCATTTGACAAGACCCAAAGCTCGCATCTCAGACGGTCAACC
 GTTCGTATTATACATTCAACATATGGTACATACGCAAAAAGCATGCCAACGATGA
 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.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAAC
 CGACACGTGCTTGGGCGCCTGCCGCCTGCCTGCCGCATGCTTGTGCTGGTGAGGC
 TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAAGCTTATGTG
 TGTGCTGGGCTTGCAATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
 TTGCTGGAGCTGCCGCGGTGCCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
 TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTGAGCAAAGGTAGCAGTTT
 GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
 GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCA
 AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTC
 ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
 AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT

GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCTCT
GGCTGCTTGCTGGGTTCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCACACTGT
CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCAAT
TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCGTGTT
GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
CGCCCCGTGCTGTTCTGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCGCCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCTTCGACC
ACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCGGACCGCTTCCTGGTGCT
GTACGGCATCGCCGCCCGGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATAACCACCCCGCATCCCCCTCC
CTCGTTTCATATCGCTTGCATCCCAAACCGCAACTTATCTACGCTGTCTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
CACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG
GTTACACACAGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGAC
ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
GTTTCAGTCACAACCCGCAAACACTAGTATGGCCACCGCATCCACTTTCTCGGCG
TTCAATGCCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCC
CAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCAGCTGCCCGACTGGAGCCGCC
TGCTGACCGCCATCACCACCGTGTTCTGTGAAGTCCAAGCGCCCCGACATGCACGA
CCGCAAGTCCAAGCGCCCCGACATGCTGGTGGACAGCTTCGGCCTGGAGTCCAC

CGTGCAGGACGGCCTGGTGTTCGCCAGTCCTTCTCCATCCGCTCCTACGAGATC
GGCACCGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAGACC
TCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTGGACGGCTTCGGCCGCACCC
TGGAGATGTGCAAGCGCGACCTGATCTGGGTGGTGTATCAAGATGCAGATCAAGG
TGAACCGCTACCCCGCCTGGGGCGACACCGTGGAGATCAACACCCGCTTCAGCC
GCCTGGGCAAGATCGGCATGGGCCGCGACTGGCTGATCTCCGACTGCAACACCG
GCGAGATCCTGGTGC GCGCCACCAGCGCCTACGCCATGATGAACCAGAAGACCC
GCCGCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTGCCCTGTTCGT
GGACAGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGAA
GACCGGCGACAGCATCCAGAAGGGCCTGACCCCGGCTGGAACGACCTGGACGT
GAACCAGCACGTGTCCAACGTGAAGTACATCGGCTGGATCCTGGAGAGCATGCC
CACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCCCTGGCCCTGGAGTACCGCCG
CGAGTGC GGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGGACCCCAAGCAA
GGTGGGCGTGC GCTCCCAGTACCAGCACCTGCTGCGCCTGGAGGACGGCACCGC
CATCGTGAACGGCGCCACCGAGTGGCGCCCCAAGAACGCCGGCGCCAACGGCGC
CATCTCCACCGGCAAGACCAGCAACGGCAACTCCGTGTCCATGGACTACAAGGA
CCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACA
AGTGAATCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGT
CGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGC
CGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAG
TTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCAAGCATCCCCTTCCCTCGTT
TCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGC
GCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCTG
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.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCCTGCCTGCCGCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTCAGCAAAGGTAGCAGTTT
GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
AGTGTGTGCGAGCCAGCATCGCCGCTGGCTGTTTCAGCTACAACGGCAGGAGTC
ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCCTCT
GGCTGCTTGTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCACACTGT
CACCAGGCCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCATT

TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
TTCCATAGTGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCTGTGTT
GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
TCCACGCCGGCTCCCCCGCCGCCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
CCAGCAGACCATCGGCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGC
CGCCCCGTGCTGTTCTGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCCGCCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCCGCCACCTGCCCCCTTCGACC
ACCAGGCCAAGCACCGCATCGAGCGCGCCCCGCACCCGCATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCCCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCCTGGTGCT
GTACGGCATCGCCGCCCCCCGACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
CACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG
GTTACACACAGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTCTTTCGCTATGAC
ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCCGATTGCAAAGAC
ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
GTTTCAGTCACAACCCGCAAACACTAGTATGGCCACCGCATCCACTTTCTCGGCG
TTCAATGCCCCGCTGCGGCGACCTGCGTTCGCTCGGCGGGCTCCGGGGCCCCGGCGCC
CAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCCCCGACTGGTCCATGCTGTTTCGC
CGTGATCACCACCATCTTCAGCGCCGCCGAGAAGCAGTGGACCAACCTGGAGTG
GAAGCCCAAGCCCAAGCTGCCCCAGCTGCTGGACGACCACTTCGGCCTGCACGG
CCTGGTGTTCGCGCCGACCTTCGCCATCCGCTCCTACGAGGTGGGCCCCGACCGC
AGCACCTCCATCCTGGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACCAC
GCCAAGAGCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCCTGGAGATGTCC
AAGCGCGACCTGATGTGGGTGGTGCGCCGCACCCACGTGGCCGTGGAGCGCTAC
CCCACCTGGGGCGACACCGTGGAGGTGGAGTGCTGGATCGGCGCCAGCGGCAAC
AACGGCATGCGCCGCGACTTCCTGGTGCGCGACTGCAAGACCGGCGAGATCCTG
ACCCGCTGCACCTCCCTGAGCGTGCTGATGAACACCCGCGACCCGCCGCCTGAGCA
CCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCCCGCCTTCATCGACAACGTGG

CCGTGAAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCC
 ACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACGTGAACCAGC
 ACGTGAACAACCTGAAGTACGTGGCCTGGGTGTTTCGAGACCGTGCCCGACAGCA
 TCTTCGAGTCCCACCACATCAGCTCCTTCACCCTGGAGTACCGCCGCGAGTGCAC
 CCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCAGCTCCGAGGC
 CGGCCTGGTGTGCGACCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCG
 CGCCCGCACCAGGTGGCGCCCCAAGCTGACCGACTCCTTCCGCGGCATCAGCGTG
 ATCCCCGCCGAGCCCCGCGTGATGGACTACAAGGACCACGACGGCGACTACAAG
 GACCACGACATCGACTACAAGGACGACGACGACAAGTGAAGTGAAGGACGACGAGC
 AGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCG
 CCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCA
 GTGTGTTTGATCTTGTGTGTACGCGCTTTTGCAGAGTTGCTAGCTGCTTGTGCTATT
 TGCGAATACCAACCCCGAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAAC
 CGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCAC
 TGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCGCTGTATTCTCCTGGTACTGCAACC
 TGTAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAA
 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.

GGTACCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAAC
 CGACACGTGCTTGGGCGCCTGCCGCTGCTGCGCATGCTTGTGCTGGTGAGGC
 TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAGCTTATGTG
 TGTGCTGGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
 TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
 TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTCAGCAAAGGTAGCAGTTT
 GCCAGAATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
 GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCA
 AGTGTGTGCGAGCCAGCATCGCCGCTGGCTGTTTCAGCTACAACGGCAGGAGTC
 ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
 AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
 GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
 CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCCTCT
 GGCTGCTTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
 GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCACACTGT
 CACCAGGCCCCGAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
 GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
 AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
 GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCATT
 TGCATGCCTGGAGAAGCGAGGCTGGGGGCTTTGGGCTGGTGCAGCCCGCAATGG
 AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
 TTCCATAGTGATTGGAAGTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTCGTGTT
 GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
 CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
 TGCTCGCCCTCTCTCCGTCCCGCCTCTAGAATATCAATGATCGAGCAGGACGGCC
 TCCACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGC
 CCAGCAGACCATCGGCTGCTCCGACGCCGCGCTGTTCCGCCTGTCCGCCAGGGC

CGCCCCGTGCTGTTTCGTGAAGACCGACCTGTCCGGCGCCCTGAACGAGCTGCAGG
ACGAGGCCCGCCCGCCTGTCCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGT
GCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCGAGGTGCC
CGGCCAGGACCTGCTGTCCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATG
GCCGACGCCATGCGCCGCCTGCACACCCTGGACCCCGCCACCTGCCCCCTTCGACC
ACCAGGCCAAGCACCCGCATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGG
TGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCCCGCCGAGCTGT
TCGCCCCGCCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCACG
GCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCTTCAT
CGACTGCGGCCGCCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCAC
CCGCGACATCGCCGAGGAGCTGGGCGGGCGAGTGGGCCGACCGCTTCTTGGTGCT
GTACGGCATCGCCGCCCCCGACTCCAGCGCATCGCCTTCTACCGCCTGCTGGAC
GAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGC
GCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACAC
CACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCG
GTTACACACAGTGCCACGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCT
GATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTCTTGCGCTATGAC
ACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCA
ACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCG
ATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCCGATTGCAAAGAC
ATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTAC
ACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTC
GTTTCAGTCACAACCCGCAAACACTAGTATGGCCACCGCATCCACTTTCTCGGCG
TTCAATGCCCCGCTGCGGCGACCTGCGTTCGCTCGGCGGGCTCCGGGCCCCGGCGCC
CAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCAGCTGCCCCGACTGGAGCATGC
TGCTGGCCGCGATCACCACCCTGTTCTTGGCGGCCGAGAAGCAGTGGATGATGCT
GGACTGGAAGCCCAAGCGCCCCGACATGCTGGTGGACCCCTTCGGCCTGGGCCG
CTTCGTGCAGGACGGCCTGGTGTTCCGCAACAACCTTCAGCATCCGCAGCTACGAG
ATCGGCGCGGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAG
ACCGCCCTGAACCACGTGAAGAGCGTGGGCCTGCTGGAGGACGGCCTGGGCAGC
ACCCGCGAGATGAGCCTGCGCAACCTGATCTGGGTGGTGACCAAGATGCAGGTG
GCGGTGGACCGCTACCCACCTGGGGCGACGAGGTGCAGGTGAGCAGCTGGGCG
ACCGCCATCGGCAAGAACGGCATGCGCCGCGAGTGGATCGTGACCGACTTCCGC
ACCGGCGAGACCCTGCTGCGCGCCACCAGCGTGTGGGTGATGATGAACAAGCTG
ACCCGCCGCATCAGCAAGATCCCCGAGGAGGTGTGGCACGAGATCGGCCCCAGC
TTCATCGACGCGCCCCCCTGCCCACCGTGGAGGACGACGGCCGCAAGCTGACC
CGCTTCGACGAGAGCAGCGCCGACTTCATCCGCAAGGGCCTGACCCCCCGCTGG
AGCGACCTGGACATCAACCAGCACGTGAACAACGTGAAGTACATCGGCTGGCTG
CTGGAGAGCGCGCCCCCGAGATCCACGAGAGCCACGAGATCGCCAGCCTGACC
CTGGAGTACCGCCGCGAGTGCGGCCGCGACAGCGTGTGAACAGCGCCACCAAG
GTGAGCGACAGCAGCCAGCTGGGCAAGAGCGCCGTGGAGTGCAACCACCTGGTG
CGCCTGCAGAACGGCGGCGAGATCGTGAAGGGCCGACCGTGTGGCGCCCCAAG
CGCCCCCTGTACAACGACGGCGCCGTGGTGGACGTGCCCCGCAAGACCAGCGAT

GACGATGACAAGCTGGGATGACTCGAGGCAGCAGCAGCTCGGATAGTATCGACA
CACTCTGGACGCTGGTTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGAC
CTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTG
TACGCGCTTTTGCAGATTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGC
ATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGT
CCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCTCGCACAGCCTTG
GTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAA
TGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTC

SEQ ID NO: 55

Prototheca moriformis FatB/A promoter/5'UTR

CCTGTCGATCGAAGAGAAGGAGACATGTGTACATTATTGGTGTGAGGGCGCTGA
ATCGGCCATTTTTTAAATGATCACGCTCATGCCAATAGACGCGGCACATAACGA
CGTTCAAACCCCCGCCAAAGCCGCGGACAACCCCATCCCTCCACACCCCCACAC
AAAGAACCCGCCACCGCTTACCTTGCCACGAGGTAGGCCTTTCGTTGCGCAAAA
CCGGCCTCGGTGATGAATGCATGCCCCGTTCTGACGAGCGCTGCCCCGGGCAACA
CGCTCTTTTGCTGCGTCTCCTCAGGCTTGGGGGCTCCTTGGGCTTGGGTGCCGCC
ATGATCTGCGCGCATCAGAGAAACGTTGCTGGTAAAAAGGAGCGCCCGGCTGCG
CAATATATATATAGGCATGCCAACACAGCCCAACCTCACTCGGGAGCCCGTCCCA
CCACCCCCAAGTCGCGTGCCTTGACGGCATACTGCTGCAGAAGCTTCATGAGAAT
GATGCCGAACAAGAGGGGCACGAGGACCCAATCCCGGACATCCTTGTGATAAT
GATCTCGTGAGTCCCCATCGTCCGCCCCGACGCTCCGGGGAGCCCGCCGATGCTCA
AGACGAGAGGGCCCTCGACCAGGAGGGGCTGGCCCCGGGCGGGCACTGGCGTCG
AAGGTGCGCCCCGTCGTTGCGCTGCAGTCCTATGCCACAAAACAAGTCTTCTGACG
GGGTGCGTTTGGCTCCCGTGCGGGCAGGCAACAGAGGTATTCACCTGGTCATGGG
GAGATCGGCGATCGAGCTGGGATAAGAGATACTTCTGGCAAGCAATGACAACTT
GTCAGGACCGGACCGTGCCATATATTTCTCACCTAGCGCCGCAAAACCTAACAAT
TTGGGAGTCACTGTGCCACTGAGTTCGACTGGTAGCTGAATGGAGTCGCTGCTCC
ACTAAACGAATTGTCAGCACCGCCAGCCGGCCGAGGACCCGAGTCATAGCGAGG
GTAGTAGCGCGCC

SEQ ID NO: 56

Prototheca moriformis NRAMP metal transporter promoter/5'UTR

ACTAATTGCAATCGTGACGTAATCATCGATATGGTCACAAGTAGATCCCCTACTG
ACACCCTCTCGTACATGTAGGCAATGTCATCGGCGCCGTCCTGCTGACCGATGCC
GACGTAGCAGAGCAGACCCGGGCCGATCTGGGATACGAGCCGGCCCTCCACCTG
CGCTCGAGGTGGAATCAAGTAAATAACCAATACACTTTTCGACACCACACAGAG
TTGCACGGACGGTGGCGTACCTCTACGCTCGCGCTCTTCACGCGCTGGACGACCG
CACGCATGAGCCCCGGGTGGCTTGGTCTGGGCTGCAAAAATGCACAACAAACAAG
TATCAGACGCTCATGGATGCACACGCGCTCCCAAGCACGCTCAGACTAAATATTA
CAGTAGCTCGTATCTGATAAGATATCGAGACATAACCGCTCAACTACCCGCAAAC
TGCGCCCCGCCAGGTGATGCGCACAGGGCCCCACCATGCGATCCATCGCATCGCT
CCTCGAGGGGCGCTATCACGTGGCCGGAGAGCGTTCACAGCGTACGCCACTGTATC
TGGGCGGTATGCGGTCCGTCAACATGGAGACAGATAACCCGCACCACCACCTTGC
AAGCTCTTCCATATTGGAAGTAGAAAATTGTAATTGTATCATCGCACGAGGGGCC
AACTTGCCGTCGGCGAGCTGGGCGACGAACACCACCTGGACGTTGTGAGACTC
GCTCGTGCCGTGCGCCGGGCCGCTGGGTATCCAGACCGTCCGCC

SEQ ID NO: 57

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

CAACGACAACCAGCAGGCAACTCGGTCAGCGACCCAACACGCGAGTCAAATTGT
TGCGTGTTCTTGCCTTGTCTATTTACTGTGATAGCAAGACTGTCGGTCAGTCAATA
CCGCGGTGCGCACGTGCGGGGTGCCAAGCCTAGCAGAGCACGGGACGGCTGGTGC
TGTGCGCCAGCTCAGCTCGCTTCGCGACCAATTGTAGGACCGGCAAAGTCACCAA
AACATGCCAGCGGTGCGATTCAATTGGTCATGAGCTCTACAAAATTGTTTTGTGC
GTCGCGCAGGTATCCAACGGCGCGGCAGAGAAAAGTTTGACAGCTCTCGATTTCAT
CTCGGAAAAATGGGGAGAATTTATGACACACAAGTGCGCAGGCGGCCCAGGCGG
CCAGCATATTCTGGCGTGACCTGGGCGGCCACAAAATGCTTGATGCACTCTAA
AATAATTATATTTGCCATGAACAAGGGAAGAGTTACCGCACCCAGCCCTAGACTT
GGGCGCCCGAGCAAGGTTACGTCAAGCCACCTTCGCCCATCGCCCAACTCCGTAT
TCCCCGACAGCCGCACGTGGCCCTCGCCGGAATGAACCCTGAATCGGCATCACG
CCACGCGTTCGCCAATCGTTCGCTCTCTGGCTTCATCGGCCTGCGCCTTCACGTC
GTGGTCACGACAGTGCATTCATACTTCCATTTGCACCTCGGCACACACTTTTACG
CATCGCCTACCCTTGCTGCGGCAGTCTAGGGTCACTTTGCAGCCATGGGACAGTG
CTACACCACCGTCGGTGCGCAAAGCTATTTCAAGTGAACCGTGGGCGGAAAAAA
GGAATGTACACTGTCTCAACCGACTCCTACAATTGTTTACCATGCAGATCAGAGC
TCGACGGCCATCATCGAGCAGGTGTGGGGCCTTGGTGGCGCGGCGCGGGGCCCC
AGGGCGTCGCAGGCATTGATGGCACTCTGAGACTTTCGCACGCGCATGAGGGAC
CCCATCAAGAGAAGAGTGTGTCTTTATGTCCCCATTCATGATGATGTATCTTGTG
ATTGTCGCAGTTTGGCAAGTTTAACCGGATCGCCGCTCCAGGTGTGGCGTGGCGG
ATTTTCTAGGGGTGCTTGAGCAGTCG

SEQ ID NO: 58

Prototheca moriformis SulfRed Sulfite reductase promoter/5'UTR

GGCCCAGGGCCCTGCGGATGGCCACACCAGATCTAGCCTCTCTTATGCCATGCC
CGCCTCGCTGCCCCGTCGTATCCCCCGCCGATCCGCGCGTAGGGGACCGCGGCCT
GACCCACGCCACGAAAGAGCTTTGCTCCTCAATTTCTCGCCAACAGAACCGTATC
AAACGCTCAACGCCTATCCCGAACAATCCGTATTACACCAAATCGAGTATACCG
GACTGGTTTGCCTAGTCTTGAAGGAAATGATCCCGTCCATGCTCGGAAGGGGGA
GCGGGCGGAGGATCCTACTCATCTCTGAAATGGGATTGGTCCGAAGATGGGTTG
GGCAAGCACGTGCCAAACCCAGCGAGTTGCTGACGAGCAGGCTCATCCAATCC
CCCGGCGAATCCTCCCTCACGCCCCGCATGCATACAAGTCCCTCCACACGCCCC
CTCCCATCCATTTTCGCTGGTCCGAACGCGAGCGGCGTCGAGGCGGACCCTTG
CTCCGCAGCGCCGTCTGGGTCTCCACCCACAGCGGCTTTGCTGCCAGAGGCACC
CCCCTTGCCCCACCTCCTTGCAGCC

SEQ ID NO: 59

Prototheca moriformis SugT Sugar transporter promoter/5'UTR

CCAGGCAGGCGGTAGGGTTGCCGATTGCTTGAGCGAATTGGAAGATATAATTTTT
TGTGGTGTCCCTGGACGCTGTTTGTGGCGCTCCTTTTTGGAGAAGATTGCGTGGG
GGAGCTTTCCATGTACCACGCTTCCTTCTGAAAGGATTCTGGCCGAGTCCTGATG
AGCCCAAAGAAAACACCTGCCTTTCAGTGCTGGCACTCTGAAAACGTCAACAGA
TGATTATACATGTCACAAAAGGCAGCCGATTAGGAACGGGAGCTCTGGCCGTTT
GTTTGGCTGCCTGGGCTGATTGAAGTGATCCACCCTGTTTGAATGAAGGCGGTTCG
AGTCGAATTATCGACCGGAGCTGTGCGGAAGGCGTCCGGGGCAGAGTGAGGTGC

TGCGGCCTGGTTGTCGTTCAAAAAGACCCCGGTAGCCCAACAATCACGAACGAA
AGGAATATAATTGCTTGCATACTATACATTTCAGTTTCTATGTGGCGGGTAGACAA
GTCTCATGGGCTTCTAAAGGCTGTCCCTTGAAGGCTACTTATAAAAACTTGCTGC
GCCATGGCACGGATCGCGCTTGCGCAGGCTGCAACCCTGCGCGCAAGGTCAAAT
ACACAGCAAAAAGATACTAACAGAATTTCTAAAAACATTTAAATATTTGTTTCGAC
CAGCCAATTGTGGTCGTAGGCACGCAAAAAGACTTTGTTTTGCGCCACCGAGCAT
CCACGCTGGCAGTCAAGCCAGTCCGATGTGCATTGCGTGGCAGCATCGAGGAGC
ATCAAAAACCTCGTGCACGCTTTTTCTGTCAATCATCATCAACCCTCCACCATGT
ATACCCGATGCATCGCGGTGCGCAGCGCGCCACGCGTCCCAGACCCGCCCAAAA
ACCCAGCAGCGGCGAAAGCAAATCTTCACTTGCCCGAAACCCCGAGCAGCGGCA
TTCACACGTGGGCGAAAACCCCACTTGCCCTAACAGGCGTATGTCTGCTGTCACG
ATGCTTGACAACGGTATTATAGATATACACTGATTAATGTTTGAGTGTGTGCGAG
TCGCGAATCAGGAATGAATTGCTAGTAGGCACTCCGACCGGGCGGGGGCCGAGG
GACCA

SEQ ID NO: 60

Prototheca moriformis Amt03-Ammonium transporter promoter/5'UTR

GGCCGACAGGACGCGCGTCAAAGGTGCTGGGCGTGTATGCCCTGGTCGGCAGGT
CGTTGCTGTTGCTGCGCTCGTGGTTCCGCAACCCTGATTTTGGCGTCTTATTCTGG
CGTGGAAGCGCTGACGCCCCGCGAGCCGGGCGGCGGCGATGCGGTGTCTCACG
GCTGCCGAGCTCCAAGGGAGGCAAGAGCGCCCGGATCAGCTGAAGGGCTTTACA
CGCAAGGTACAGCCGCTCCTGCAAGGCTGCGTGGTGGACTTGAACCTGTAGGTCC
TCTGCTGAAGTTCCCTCCACTACCTCACCAGGCCCAGCAGACCAAAGCACAGGCTT
TTCAGGTCCGTGTTCATCCACTCTAAAACACTCGACTACGACCTACTGATGGCCCT
AGATTCTTCATCAACAATGCCTGAGACACTTGCTCAGAATTGAAACTCCCTGAAG
GGACCACCAGAGGCCCTGAGTTGTTCCCTTCCCCCGTGGCGAGCTGCCAGCCAGG
CTGTACCTGTGATCGAGGCTGGCGGGGAAAATAGGCTTCGTGTGCTCAGGTCATGG
GAGGTGCAGGACAGCTCATGAAACGCCAACAATCGCACAATTCATGTCAAGCTA
ATCAGCTATTTCTCTTTCACGAGCTGTAATTGTCCCAAAATTCTGGTCTACCGGGG
GTGATCCTTCGTGTACGGGGCCCTTCCCTCAACCCTAGGTATGCGCGCATGCGGTC
GCCGCGCAACTCGCGCGAGGGCCGAGGGTTTGGGACGGGCGCTCCCGAAATGCA
GTTGCACCCGGATGCGCGGCGCCTTTCTTGCGATAATTTATGCAATGGACTGCTC
TGCAAATTTCTGGGTCTGTGCGCAACCCTAGGATCAGCGGCGTAGGATTTCTGTA
TCATTCGTCTGATGGGGAGCTACCGACTACCTAATATCAGCCCGGCTGCCTGA
CGCCAGCGTCCACTTTTGCGTACACATTCCATTCTGTCGCCAAGACATTTATTGTG
GTGCGAAGCGTCCCCAGTTACGCTCACCTGTTTCCCGACCTCCTTACTGTTCTGTC
GACAGAGCGGGCCACAGGCCGGTCGCGAGCC

SEQ ID NO: 61

Prototheca moriformis Amt02-Ammonium transporter promoter/5'UTR

TCACCAGCGGACAAAGCACCGGTGTATCAGGTCCGTGTCATCCACTCTAAAGAG
CTCGACTACGACCTACTGATGGCCCTAGATTCTTCATCAAAAACGCCTGAGACAC
TTGCCAGGATTGAAACTCCCTGAAGGGACCACCAGGGGCCCTGAGTTGTTCCCT
CCCCCGTGGCGAGCTGCCAGCCAGGCTGTACCTGTGATCGGGGCTGGCGGGAA
AACAGGCTTCGTGTGCTCAGGTTATGGGAGGTGCAGGACAGCTCATTAACGCC
AACAAATCGCACAATTCATGGCAAGCTAATCAGTTATTTCCCATTAACGAGCTATA
ATTGTCCCAAAATTCTGGTCTACCGGGGGTGATCCTTCGTGTACGGGGCCCTTCCCT
CAACCCTAGGTATGCGCACATGCGGTGCGCGCGCAACGCGCGCGAGGGCCGAGG

GTTTGGGACGGGCCGTCCCGAAATGCAGTTGCACCCGGATGCGTGGCACCTTTTT
TGCGATAATTTATGCAATGGACTGCTCTGCAAAATTCTGGCTCTGTTCGCCAACCC
TAGGATCAGCGGTGTAGGATTTTCGTAATCATTTCGTCCTGATGGGGAGCTACCGAC
TGCCCTAGTATCAGCCCGACTGCCTGACGCCAGCGTCCACTTTTGTGCACACATT
CCATTCGTGCCCAAGACATTTTCATTGTGGTGCGAAGCGTCCCCAGTTACGCTCAC
CTGATCCCCAACCTCCTTATTGTTCTGTTCGACAGAGTGGGCCCAGAGGCCGGTCG
CAGCC

SEQ ID NO: 62

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

CGAAGGGGTCTGCATCGATTTCGCGCGGTCTGGAGGCCAGCGTGACTGCTCGCGA
AAATGCTCTGCCGTGTTCGGGCTCTGGCTGGGGCGGCCAGAGATCTCACCGTGCCA
CACGCAACTGCCGCACTCTGTGCCCCGCCACCTGGCGCGCACATGCGACCTCTTCC
CCGTCATACCCTCTCCTCATGTGATCTTTCCACACGAGTGACGCAGGTGCGCGGA
GTGGAGGGAATCAGGACGTTTTCAAGGTACCTGCTCGAGCCGTACCAACAGCTG
CCGCCCGGCAAGGAAGAGATCGAGGCAGAGATTGCCCGGCTGGAGGCCCGGATA
ACGGAGCTCAAGAGCAAGCTGTCCGAGTGAGACCGCCCAGGTGCACGTGTTCGAC
TCGCTATGACATGTACTCGACACAACATGAGGAATTCATCGAATTTGTAGGAAGC
GGGCATTGGTACGGGAGTGGGAAAGCGAAAAAACCTCCCTCCGGCAGTGCCATC
TGCCGGAGTCGAACGTTGATAGGGTTCTCGTGACAGGGTGTGACCTCTCAGCCTT
GCATCAATTAACGCTATAGACATTATCAGTAACCGTGAATCCCGCATTGGATGC
CACCCGCGCGACCAATTGGGGACCTGCATTACAGATCTAGGTGAGATGACAGCGA
GGCAACTTCGGCCCCGCGGCCAGCTTGCGGCGCACCAATATTGGTTCACGGGAAG
CCACACACCGACCATAAATGAATACTTGTAAGCTATGTCAACCGATCAATGGCGT
CGAAAGTGTGCCACGAGGATCCATCTGGCGGGGCGGCGTGGCGCACAAGCGCAG
TCGCAATTTCTCGGACCCATCTGACCTAGGCCAGCGCCGCGGGAGAAATCCCCG
GCGGGTCTTCCACGCAGTAACCCTAATGAGTATCGAGCGCCGACCATTTACACCA
TCGCCCCCGAAATCCTTCCGACATTATTATTATCTTTTAGATCTTGGAACAGACTC
TGCCAACC

SEQ ID NO: 63

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

AGAGAGCGGAGGTGGGGTTGTGAGGTGGGGTTGCTGACCAGGAGCTCGCGTCGC
CGAGCGGCACTCGCACACGGTCCAGTTACCCCCCTCCGCCAAACGCAAGCCT
CCCATCTTGATGCCTTTCCGGCCACCTATACTATTTCTTAGTTTCGCTGTAACATCC
AGACCGTCCTGAATAATAACAATGCCCTGTGTCAAGTGCATTCCTAAAAAATTC
TGTCCCAACCAACAATCCCACCTGAAATACCACCAGCCCTGCCAGTACACTCTT
CCAATACCATCTCCCTACCTCCACGCGCAAGCGACCCCCATGCGCGACCAAGGCTC
GAAAGTGATTTATGACTTGAGACGAGCGAGTGGCGGCGCGGTTCGACTGCCTTTTC
ATCACGTGCCGTACGTGCGCGACCGCTAGGGCTTTGCACGGCAACGCACGGCTTC
GCCAACCCGACCAGCCAGGACCTCGACTACTCTACCGCGAATTCGCCTCAAGAA
GTCCGCCAAATGTGCCATACACCATTCCTTACAGCACTGTTCAAACCTTGATGCCAA
TTTTGACATTCGGGTTGCTCGTTGGCTGCGCCCACATCGGCCGTGAGTGCAGCAG
GCGGGATCGGACACGGAGGACGCGGCGTACGCCCCGAACGCAGCCCGTAACCTC
TACATCAACACGACGTGTTGCGTAATCCCGCCCGGCTGCGCATCGTGCCAACCCA
TTCGCGATGGATGGTCGGAAAAATGGTGTGCCAACTGCCCTGAGGGAGGCTCTCG
CGAAACGGGCACGTCCCTGAAACCGAAACTGTGGCCTTGTCGTGCGGCCACGCAA
GCACGTGGACCCTAAACACCAAGAAAATCAGTAAACAAGGTTGACATCCTCTAC

GGGCGAATTGTTTGCCCAACCCTTCATCGCACACTGCCATTATAATGCATCTAGC
TCGGCGACAAGTTTAGAAAAGGCAGGCTGCATTGTTCCATTTTCGCCGTGGCGGCG
TGGGTGCCCATTTTACGAGGTTTGGGCTCCCGGGCAGCGACCGAGCCAGGTCGA
GTCCCTCTCGCCCGTCGACAATGTTGCGAACCCCAAGCGGCTAACAACAACCTT
GATGGTACCTGTACACTGCCAATTCCTTCTTCCCGGCCGAGGTTTACACGTGAT
GGCCATGGCTTCGCATTTCAGGCCGACTTCCCATTCGACTTTCCAGAGGGTCCGC
GGACGCTGGGGGTGGCTGCCTGAGGCCACCCCTTTGTTCCCGCGTCCCGACAA
ACACAATTGCGTTACATAAGGGGGAGCCGCCCGGTTTCAGAGTGCAGAAATCTTT
CACTATATTTTCCAGTCGTCAGCGAAATCAAGT

SEQ ID NO: 64

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

GATGGTGGGGTGTCTGCCTTGGGCTGGGTGATGGAGGCTGGTGGTGC GCGGGTTT
CCTGATGCATTCTATCTACGCAGTGT CATGGTGTCCATTCCACACACCAGTACAC
CCTTACACTAAGGATCCATCCCTCCTTCCCTCTTCAGGACTACATGGACCCACG
AGCTACCGACCGGGCTTTCTCAAAAACGTCAAGGTCATGTTTGACATGCGGGACG
TGGTGGACGACGTGCAAGGTGCGTCCGGAGTGC GCGCAAATGAGCAAGTCGGGC
AATGTGTGCGGGTGGGCACCGGGGCTGGAGATCCGCGATCCCCGAGAAAACGCC
GTACCACCCCCCGCGCTATTCCCTCGATTGCGCGCAGATGTGGTGACCGACACGG
GGGACAACCTGGCGGACATGGGGCGCCGGACCTGGAAGCACGCCAAGTCGCACA
CGGGGAGGCTCGTGCAAGTCCCCCCCATCGTACCTCAAGGGTCTCTTTGGTCGCGA
TCCAAAGTACGCTGGTGGCATGGCATGCCCGAAATGAACATCATGTGTGATCTCC
GATTGCCAATGGCCACCTCCACGGACCACTTGCAGGCGGAAGCGCAATCCAGG
GCCCAGCCTGACGAGGACGGAGACTCCTCGTCCAGCGCGGGGTCCCCGACCCG
ACGCAGCAGCCGACCCCTGCTAACCCGGCAACGATCGGACCAGCAACCTTGCTG
TAGTTCCGATCCGTGATGACGGGCATTGCCGCCGCTCGATCCGCTTTGATGACTG
TCTATTATTTGCGCGGAGCCCCCTCGGAACCCTACCCCGCTCTTGCAAGCCCCTTG
CATCGGAGATCCTCGTGCGCCCGCCATGACCCCACTGGATTGCCCAACATCCTTC
TTTATCGTGTA AAAATGTGATTCTCGGCTGCAATCGACTGGCCTTCGCTTCTGGCC
CCAAGAGGGCTCGAACGTGCGGCAGCGAGGGCGCTGACACACCCAAGCCCTAGG
GCTTTCAACGTGCGCTGCCAGGCCGGATAGGGGGATCGCCTCCTTTCCACCACCC
ACCTACGAGGGATTTCGAGTCGGCTTCCAGCTCAGCTATTTCGGCCGCGCCCCCGGC
CCTGCAGACGTCTCCAGTTTCCGAACAGGTCGCTCTCAGAACACCTGCCGCGGC
TGCGATACGGCAGGCTCTCAAAGCGTCGAC

SEQ ID NO: 65

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

CGCGTGGAGCGGTGCGTGCGGATGCCGCGCGCCTGCCAAGGCCTTTTGTATGCCT
GGCCTGGGAAGTTTCCTGACTGAAGCATCTTCAAGATGCTCTCTCACGACCAGCG
ACACCAACACCGTCACTTTTTGCCCTCCTGCCGAGGTGCCACTTTCTACTTTGA
CGTCTTCTCCAGGCGGTACATTGCGGGACTGAGCGCCAATTTCGGCCAAGAACAG
CGCTGTGCACTTGAGGAGGCAGGGGTCCGTCGACTCTGCCGAGTGACACGCCTTC
GACCCGACTGTACTACGGCCTGCTGAAGAGTGGGTCTCGCCGGCCGGCGTGACC
GGCCCTGTGCCACAATCGACCATCTATTGCTCCTTGTCATCTGGCGCCGTCAAT
TGCCCGCGACTTGACGGCAACTGGCTCGATCGAGTCGATTGAAAAAGCACGTTT
TGTCTACAGGGCCGCGGTCCGTTACCAACGTGGTTCTCGTTAGGTTTTCGTCGG
GCGGTGGTGC GCGAACTGTCCGATGCCATCCCGGCAAACCCAGCAAGGTCGCC
AGTCTGGTTCTGACGCAATAGAGTGCGTTTTTGGGCCAGTCTAAAAATTCGTCTGG

CATGACGTGGCTCCACATCGTACCCGGAGCCTGCCTTGGTAATGTGAGGCACCGG
TGCCAACCTCCATTATGGCAGGCATCGAGCGCGCAGGTGAGTACATGACCTTCCGT
GAATTGGGAAGGCGAGCTTGTGTAACGCCTGCGATCGTGCCAGTGAGGCATCGT
AAACTCAAAATATTTTGTAGAAAGTGTCTGATGCCTGGTGAGGCTGCGTAGGGCA
AGGGCAAGCCCTTGGCAGATGGGTAATGGGTCCGGACCTCACAACAGCAACCCC
GCGTCCCCCTTAGGGCCCCCTGAGGCTCGATGGCAGGGCCAGCGAGCCCGCGGCC
AAAGGGCGCCATCCCACGGTCGCCCCAACGACTCCACGGGTCCTATACCTCATCTT
GAATGGCACTAAAACTATAGAATATCGGGCACTGGTGGGCGTCTGGGGTACAG
CTGGCCGAGCGCAGTGGCAAACCCTAGGTCCCGCCTCAAGGGCGATTCCCGGGT
CAATGACACGCAAGCAAGATCACATGGCGCGGTCCGCCTCGCGGCTCCACACCC
AGGCCCTAGTTTCGCAACCCATAAATATCGCCCCGATACCATCATAAGCCAGCAA
ATAATTTTTTATCAGAGTTCCAAACCTCCTCAGCTGTGGGAAACCAGCCCCTCT
GAACG

SEQ ID NO: 66

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

CCGAGCAGTTCATGGCCAAGTACAAGGACTAGAGACCGGAGGTTCGGTAGGCTGA
ATGGAGCTGGCGTCGTGCGCGACGTGCACGCGATGCGATACTACGACCCCA
CAAACGCATGCCTCCCATCTTGATGCCTTTCCGGCCATTTATACTATTTCTCATTT
CGCTGTAACATCTTGAATAATAGAATTGCCCTGTGTCAAGTGGATTCCAAGAAAT
ATTCTGTCCCAACAAAACAACCCAACCTGAAAACAACCTCAAATACCACCAGCC
CTGCCCACCTGCCCAGTACACTTTTCCAATACCATCTCCCTACCTTCACGCGCAAG
CGGCACCCATGCGCGACCAAGGCTCGAAAAGGATTTACGACTCAGGACGAGCGAG
TGGCGGCGCGACCGCCTGCCTGTTTCGTACGTGCCGTACGTTCGGCGACCGCTAGA
GCTTTGCCTGGCAACCCCCGGCTTCGTCAACCCGGCCAGCCAGGATCTCGACCAC
TCTACCGCGAAATCGCCTCAAGAAGTCGCCAAAAGTGCCGTACACCATGCTTCGC
AGCGCTGTTCAAACCTTGATGCCAATCTTGACAATCAGGTTGCTCGTTGGCTGCGT
CCACATCGGCCGTGATTGCAGCAGGCGGGGATCGGACACGGAGGACGCGGGCGTC
ACGCCGGAACGCAGCCCGTAACTCTACATCAACGCGATATGTTGCGTAATCCCG
CCCGGCTGCGCATTTGTGACAACCCATTTCGCGATGGATGGTCGGAAAATGGTGTGC
CAACTGCCCTGAGGGACTCTCTCGCGAAACGGGCACGTCCCTGTATCCGAAACTG
TGGCATGGCCTTGTCGACCACGCAAGCACGTGGACCCTAACACCACGAAAATAA
GTAAAAAAGGTTGACATCCTCTACGAGCGAATTGTTTGCTCGACCCTTCATCGCA
CACTGTCAATTATAATGCATCTAGCTCGGCGACAAGTTTAAAAAAGGCAGGCTGCA
TTATTCCATTTTGCCGTGGCGGCATGGGTGCCATTTTATGAGGTTTGGGCTCTTG
GGCAGCGACCGAGCCAGGTTGAGTCCCTCTCGCCCGTCGACAACGTTCCAAAGC
CCATAAGTGGCTAATAAACAACCTTGATGGTACCTGTACACTGCCAGTTCCTTCTT
CCCCGGCCGAGGTTTACACGTGATGGCCATGGCTTCGCGTTTCAGGCTGACTTCC
CATTCCGACTTTCCAGAGGGTCCGCGGACGCCGGGGGTTGGCTGCGTGAGGCCC
ACCCCTTGTTCCCCGCGTCCCGACAAACACAATTGCGTTACATAAGGGGGAAGCC
GCCCCCGTTTCAGAGTGCAAACATCTTTCATTATATTTTTCAGTCGTCAGCGAAAT
CAAGTATGTCGCTGACAGGCATGAAGGCC

SEQ ID NO: 67

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

GCTCTTCGGCCGCCGCCACTCCTGCTCGAGCGCGCCCGACTCGCGCTCCGCCTGC
GCCCGCGCGTGCGCCGCCAGCGCCTTGGCCTTTTCGCCGCGCTCGTGCGCGTCGC

TGATGTCCATCACCAGGTCCATGAGGTCTGCCTTGCGCCGGCTGAGCCACTGCTT
 CGTCCGGGCGGCCAAGAGGAGCATGAGGGAGGACTCCTGGTCCAGGGTCCTGAC
 GTGGTCGCGGCTCTGGGAGCGGGCCAGCATCATCTGGCTCTGCCGCACCGAGGC
 CGCCTCCAACCTGGTCTCCAGCAGCCGCAGTCGCCGCCGACCCTGGCAGAGGAA
 GACAGGTGAGGGGGGTATGAATTGTACAGAACAACCACGAGCCTTGTCTAGGCA
 GAATCCCTACCAGTCATGGCTTTACCTGGATGACGGCCTGCGAACAGCTGTCCAG
 CGACCCTCGCTGCCGCCGCTTCTCCCGCACGCTTCTTTCCAGCACCCTGATGGCG
 CGAGCCAGCGCCGCACGCTGGCGCTGCGCTTCGCCGATCTGAGGACAGTCGGGG
 AACTCTGATCAGTCTAAACCCCCCTTGCGCGTTAGTGTTGCCATCCTTTGCAGACC
 GGTGAGAGCCGACTTGTGTGCGCCACCCCCACACCACCTCCTCCCAGACCAAT
 TCTGTACCTTTTTTGGCGAAGGCATCGGCCTCGGCCTGCAGAGAGGACAGCAGTG
 CCCAGCCGCTGGGGGTGGCGGATGCACGCTCAGGTACCCTTTCTTGCGCTATGA
 CACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGC
 AACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGGCTGCATGGGCGCTCC
 GATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGA
 CATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTA
 CACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTT
 CGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTC
 CTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCC
 GACCGCCCCCTGGTGCCTTCACCCCCAACAAGGGCTGGATGAACGACCCCCAAC
 GGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAAC
 CCGAACGACACCGTCTGGGGGACGCCCTTGTTCCTGGGGCCACGCCACGTCCGAC
 GACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGAC
 TCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCT
 TCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACA
 CCCC GGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTT
 CACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGA
 CCCGAAGGTCTTCTGGTACGAGCCCTCCAGAAAGTGGATCATGACCGCGGCCAA
 GTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAA
 GCTGGAGTCCGCGTTCGCCAACGAGGGCTTCTCGGCTACCAGTACGAGTGCCCC
 GGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCCTACTGGGTGATG
 TTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCG
 TCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGT
 GGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACC
 TACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCTTCG
 TGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAAC
 ACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCG
 ATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACG
 TTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACAGCACCCGGCACCCCTG
 GAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGT
 TCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCG
 CATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCTGACCGCGGGGAACAGCAAG
 GTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAAC
 CAGCCCTTCAAGAGCGAGAACGACCTGTCTACTACAAGGTGTACGGCTTGCTGG
 ACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACA
 CCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGG
 TGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATT
 GGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGG
 ACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCA
 AACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGC

TTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATATCGCTT
 GCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGC
 TCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGG
 TACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGAT
 GGGAAACACAAATGGAGGATCCCGCGTCTCGAACAGAGCGCGCAGAGGAACGCT
 GAAGGTCTCGCCTCTGTGCGCACCTCAGCGCGGCATACACCACAATAACCACCTGA
 CGAATGCGCTTGGTTCTTCGTCCATTAGCGAAGCGTCCGGTTCACACACGTGCCA
 CGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCTGATGGTCGAAACGTTT
 ACAGCCTAGGGATATCGAATTCATCGAATTTGTAGGAAGCGGGCATTGGTACGG
 GAGTGGGAAAGCGAAAAAACCTCCCTCCGGCAGTGCCATCTGCCGGAGTCGAAC
 GTTGATAGGGTTCTCGTGACAGGGTGTGACCTCTCAGCCTTGCAATTAACG
 CTATAGACATTATCAGTAACCGTGAATCCCGCATTGGATGCCACCCGCGCGACCA
 TTGGGGACCTGCATTACAGATCTAGGTGAGATGACAGCGAGGCAACTTCGGCCC
 GCGGCCAGCTTGCGGCGCACCAATATTGGTACCGGAAGCCACACACCGACCA
 TAAATGAATACTTGTAAGCTATGTCAACCGATCAATGGCGTCGAAAGTGTGCCAC
 GAGGATCCATCTGGCGGGGCGGCGTGGCGCACAAGCGCAGTCGAATTTCTCGG
 ACCCATCTGACCTAGGCCAGCGCCGCGGGAGAAATCCCCGGCGGGTCTCCAC
 GCAGTAACCCTAATGAGTATCGAGCGCCGACCATTTACACCATCGCCCCCGAAA
 TCCTTCCGACATTATTATTATCTTTTAGATCTTGGAACAGACTCTGCCAACCCTA
 GTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCG
 TCGCTCGGCGGGCTCCGGGCCCCGCGGCCAGCGAGGCCCCCTCCCCGTGCGCGG
 GCGCGCCCCGACTGGTCCATGCTGTTGCGCGTGATCACCACCATCTTCTCCGCC
 GCCGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCAG
 CTGCTGGACGACCACTTCGGCCCCCACGGCCTGGTGTTCCGCGCACCTTCGCCA
 TCCGCAGCTACGAGGTGGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGA
 ACCACCTGCAGGAGGCCGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCG
 ACGGCTTCGGCACCACCCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGA
 AGCGCACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGG
 TGGAGTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCCTGG
 TGCGCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGAT
 GATGAACACCCGCACCCGCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGA
 GATCGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAA
 GCCCCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCC
 CCGCTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGA
 CTGGATCCTGGAGACCGTGCCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCC
 TTCACCATCGAGTACCGCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGA
 CCACCGTGAGCGGCGGCTCCTCCGAGGCCGGCCTGGTGTGCGAGCACCTGCTGC
 AGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGC
 TGACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGA
 CTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGA
 CGACGACAAGTGACTCGAGTTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTA
 TCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGC
 CTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCT
 TGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACC
 CCCAGCATCCCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTA
 CGCTGTCTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGACA
 GCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCA
 CTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTGT
 AGAGCTCCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCG
 ATAACCTCCAAAGCCGCTCTAATTGTGGAGGGGGTTTGAATTTAAAAGCTTGGA

TGTTGGTTCGTGCGTCTGGAACAAGCCCAGACTTGTTGCTCACTGGGAAAAGGAC
CATCAGCTCCAAAAAACTTGCCGCTCAAACCGCGTACCTCTGCTTTCGCGCAATC
TGCCCTGTTGAAATCGCCACCACATTCATATTGTGACGCTTGAGCAGTCTGTAAT
TGCCTCAGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCG
CGGGCGAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAG
TTCATAACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTG
AGTGGCCACCCCCCGGCCCTGGTGCTTGCGGAGGGCAGGTCAACCGGCATGGGG
CTACCGAAATCCCCGACCGGATCCCACCACCCCCGCGATGGGAAGAATCTCTCCC
CGGGATGTGGGCCCCACCACCAGCACAACTGCTGGCCCAGGCGAGCGTCAAACC
ATACCACACAAATATCCTTGGCATCGGCCCTGAATTCCTTCTGCCGCTCTGCTACC
CGGTGCTTCTGTCCGAAGCAGGGGTTGCTAGGGATCGCTCCGAGTCCGCAAACCC
TTGTCGCGTGGCGGGGCTTGTTTCGAGCTTGTTTCGAGCTTGAAGAGCCTCTAGAGT
CGACCTGCAGGCATGCAAGCTTGCGTAATCATGGTCATAGCTGTTTCCTGTGTG
AAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGT
AAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCAC
TGCCCGCTTTCAGTCGGGAAACCTGTCTGCGCCAGCTGCATTAATGAATCGGCCA
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
AGGTGTGTGTGAGGGTTGTGGTTGCCCCGTATCGAGGTCCTGGTGGCGCGCATGGG
GGAGAAGGCGCCTGTCCCGCTGACCCCCCGGCTACCCTCCCGGCACCTTCCAGG
GCGCCTTCGGCACGGGTGTCTGGGTGTGCGCGCACGAGTGCGGCCACCAGGCCTT

TTCCTCCAGCCAGGCCATCAACGACGGCGTGGGCCTGGTGTTCACAGCCTGCTG
CTGGTGCCCTACTACTCCTGGAAGCACTCGCACCGCCGCCACCACTCCAACACGG
GGTGCCTGGACAAGGACGAGGTGTTTGTGCCGCCGCACCGCGCAGTGGCGCACG
AGGGCCTGGAGTGGGAGGAGTGGCTGCCCATCCGCATGGGCAAGGTGCTGGTCA
CCCTGACCCTGGGCTGGCCGCTGTACCTCATGTTCAACGTCGCCTCGCGGCCGTA
CCCCGCGCTTCGCCAACCACCTTTGACCCGTGGTCGCCCATCTTCAGCAAGCGCGAG
GTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAAGGTAGGGCGGGCTGCGAGAC
GGCTTCCCGGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCTT
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
ATCAAACAGCCTCAGTGTGTTTGTATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAG
CTGCTTGTGCTATTTGCGAATAACACCCCCAGCATCCCCCTCCCTCGTTTCATATC
GCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTC
CTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTC
CTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTG
GGATGGGAACACAAATGGAGCATCGAGGTGGTCATCTCCGACCTCGCGTTGGTG
GCGGTGCTCAGCGGGCTCAGCGTGCTGGGCCGCACCATGGGCTGGGCCTGGCTG
GTCAAGACCTACGTGGTGCCCTACATGATCGTGAACATGTGGCTGGTGCTCATCA
CGCTGCTCCAGCACACGCACCCGGCCCTGCCGCACTACTTCGAGAAGGACTGGG
ACTGGCTACGCGGCGCCATGGCCACCGTCGACCGCTCCATGGGCCCCGCCCTTCAT
GGACAGCATCCTGCACCACATCTCCGACACCCACGTGCTGCACCACCTCTTCAGC
ACCATCCCGCACTACCACGCCGAGGAGGCCTCCGCGCCCATCCGGCCCCATCCTGG
GCAAGTACTACCAATCCGACAGCCGCTGGGTGCGCCGCGCCCTGTGGGAGGACT
GGCGCGACTGCCGCTACGTGCTCCCCGACGCGCCCGAGGACGACTCCGCGCTCTG
GTTCCACAAGTGAGCGCGCCTGCGCGAGGACGCAGAACAACGCTGCCGCCGTGT
CTTTTGCACGCGCGACTCCGGCGCTTCGCTGGTGGCACCCCCATAAAGAAACCCT
CAATTCTGTTTGTGGAAGACACGGTGTACCCCCACCCACCCACCTGCACCTCTAT
TATTGGTATTATTGACGCGGGAGTGGGCGTTGTACCCTACAACGTAGCTTCTCTA
GTTTTTCAGCTGGCTCCCACCATTGTAAAGAGCCTCTAGAGTCGACCTGCAGGCAT
GCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCT
CACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGC
CTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCCGCTTTCCAG
TCGGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGA
GGCGGTTTTCGTATTGGGCGCTCTTCC

SEQ ID NO: 74

Prototheca moriformis delta 12 FAD knockout homologous recombination targeting construct

GCTCTTCGGGTTTGCTCACCCGCGAGGTCGACGCCAGCATGGCTATCAAGACGA
ACAGGCAGCCTGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCA

TCCCCGCGCACTGTTTCGAGCGCTCGGGCGCTTCGTAGCAGCATGTACCTGGCCTT
TGACATCGCGGTCAATGTCCCTGCTCTACGTCGCGTCGACGTACATCGACCCTGCG
CCGGTGCCTACGTGGGTCAAGTATGGCGTCATGTGGCCGCTCTACTGGTTCTTCC
AGGTGTGTGTGAGGGTTGTGGTTGCCCCGTATCGAGGTCCTGGTGGCGCGCATGGG
GGAGAAGGCGCCTGTCCCGCTGACCCCCCGGCTACCCTCCCGGCACCTTCCAGG
GCGCCTTCGGCACGGGTGTCTGGGTGTGCGCGCACGAGTGCGGGCCACCAGGCCTT
TTCCTCCAGCCAGGCCATCAACGACGGCGTGGGCCTGGTGTTCACAGCCTGCTG
CTGGTGGCCTACTACTCCTGGAAGCACTCGCACCCGCCGCCACCCTCCAACACGG
GGTGCCTGGACAAGGACGAGGTGTTTGTGCCGCCGCACCGCGCAGTGGCGCACG
AGGGCCTGGAGTGGGAGGAGTGGCTGCCCATCCGCATGGGCAAGGTGCTGGTCA
CCCTGACCCTGGGCTGGCCGCTGTACCTCATGTTCAACGTCGCCTCGCGGCCGTA
CCCGCGCTTCGCCAACCCTTTGACCCGTGGTCGCCCATCTTCAGCAAGCGCGAG
GTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAAGGTAGGGCGGGCTGCGAGAC
GGCTTCCCGGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTT
CGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGC
CAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACA
CCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCT
AAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATG
CTGCTGCAGGCCTTCCTGTTCCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCT
CCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCACCTTCACCCCCAACAAGG
GCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGC
ACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCG
GGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCAT
CGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTAC
AACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGG
CCATCTGGACCTACAACACCCCCGAGTCCGAGGAGCAGTACATCTCCTACAGCCT
GGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAA
CTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGG
ATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGAC
GACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGGCTTCTTCGGCT
ACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCA
AGTCTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTC
CTTCAACCAGTACTTCGTCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGAC
AACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCT
TCAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAAC
GGGAGTACTCCGCCTTCGTGCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGT
GCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGAT
CAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCG
GTTCCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTGCACCTGTC
CAACAGCACCGGCACCCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCA
GACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAG
GACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCTCTTCTTCC
TGGACCGCGGGAACAGCAAGGTGAAGTTTCGTGAAGGAGAACCCCTACTTCACCA
ACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGACCTGTCCCTACT
ACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACG
GCGACGTCTGTTCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTC
CGTGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGT
GCGCGAGGTCAAGTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTC
TGGACGCTGGTCTGTGATGGACTGTTGCCGCCCACTTGCTGCCTTGACCTGTG
AATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGC

GCTTTTGCAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCC
 CTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCCTGC
 TATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCTCGCACAGCCTTGGTTTG
 GGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTG
 ATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAGCATCGAGGTGGTCATCT
 CCGACCTCGCGTTGGTGGCGGTGCTCAGCGGGCTCAGCGTGCTGGGCCGCACCAT
 GGGCTGGGCCTGGCTGGTCAAGACCTACGTGGTGCCCTACATGATCGTGAACATG
 TGGCTGGTGCTCATCACGCTGCTCCAGCACACGCACCCGGCCCTGCCGCACTACT
 TCGAGAAGGACTGGGACTGGCTACGCGGCGCCATGGCCACCGTCGACCGCTCCA
 TGGGCCCCGCCCTTCATGGACAGCATCCTGCACCACATCTCCGACACCCACGTGCT
 GCACCACCTCTTCAGCACCATCCCGCACTACCACGCCGAGGAGGCCTCCGCCGCC
 ATCCGGCCCATCCTGGGCAAGTACTACCAATCCGACAGCCGCTGGGTGCGCCGC
 GCCCTGTGGGAGGACTGGCGCGACTGCCGCTACGTGCTCCCCGACGCGCCCGAG
 GACGACTCCGCGCTCTGGTTCCACAAGTGAGCGCGCCTGCGCGAGGACGCAGAA
 CAACGCTGCCGCCGTGTCTTTTGCACGCGCGACTCCGGCGCTTCGCTGGTGGCAC
 CCCATAAAGAAACCCTCAATTCTGTTTGTGGAAGACACGGTGTACCCCCACCCA
 CCCACCTGCACCTCTATTATTGGTATTATTGACGCGGGAGTGGGCGTTGTACCCT
 ACAACGTAGCTTCTCTAGTTTTTCAGCTGGCTCCCACCATTGTAAAGAGCCTCTAG
 AGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
 GTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAA
 GTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGC
 TACTGCCCCTTCCAGTCGGGAAACCTGTGCTGCCAGCTGCATTAATGAATCG
 GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 75

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

GCTCTTCCGCCTGGAGCTGGTGCAGAGCATGGGTGAGTTTGCGGAGGAGAGGGT
 GCTCCCCGTGCTGCACCCCGTGGACAAGCTGTGGCAGCCGCAGGACTTCCTGCCC
 GACCCCGAGTCGCCCGACTTCGAGGACCAGGTGGCGGAGCTGCGCGCGCGCGCC
 AAGGACCTGCCCCGACGAGTACTTTGTGGTGTGGTGGGCGACATGATCACGGAG
 GAGGCGCTGCCGACCTACATGGCCATGCTCAACACCTTGGACGGTGTGCGCGAC
 GACACGGGCGCGGCTGACCACCCGTGGGCGCGCTGGACGCGGCAGTGGGTGGCC
 GAGGAGAACCGGCACGGCGACCTGCTGAACAAGTACTGTTGGCTGACGGGGCGC
 GTCAACATGCGGGCCGTGGAGGTGACCATCAACAACCTGATCAAGAGCGGCATG
 AACCCGCAGACGGACAACAACCTTACTTGGGCTTCGTCTACACCTCCTTCCAGG
 AGCGCGCCACCAAGTAGGTACC

SEQ ID NO: 76

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

CAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTG
 ATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTT
 ATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAG
 CTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATATC
 GCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTC
 CTGCTCCTGCTCACTGCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTC
 CTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTG

GGATGGGAACACAAATGGAAGGATCGTAGAGCTCCAGCCACGGCAACACCGCGC
GCCTGGCGGCCGAGCACGGCGACAAGGGCCTGAGCAAGATCTGCGGGCTGATCG
CCAGCGACGAGGGCCGGCACGAGATCGCCTACACGCGCATCGTGGACGAGTTCT
TCCGCCTCGACCCCGAGGGCGCCGTCGCCGCCTACGCCAACATGATGCGCAAGC
AGATCACCATGCCCCGCGCACCTCATGGACGACATGGGCCACGGCGAGGCCAACC
CGGGCCGCAACCTCTTCGCCGACTTCTCCGCCGTCGCCGAGAAGATCGACGTCTA
CGACGCCGAGGACTACTGCCGCATCCTGGAGCACCTCAACGCGCGCTGGAAGGT
GGACGAGCGCCAGGTCAGCGGCCAGGCCGCCGCGGACCAGGAGTACGTTCTGGG
CCTGCCCCAGCGCTTCCGAAACTCGCCGAGAAGACCGCCGCCAAGCGCAAGCG
CGTCGCGCGCAGGCCCGTCGCCCTTCTCCTGGAGAGAAGAGCCTCTAGAGTCGACC
TGCAGGCATGCAAGCTTGCGCTAATCATGGTCATAGCTGTTTCTGTGTGAAATT
GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAG
CCTGGGGTGCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCC
CGCTTTCCAGTCGGGAAACCTGTCTGTGCCAGCTGCATTAATGAATCGGCCAACGC
GCGGGGAGAGGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 77

Prototheca moriformis SAD2A knockout homologous recombination targeting construct

GCTCTTCCGCCTGGAGCTGGTGCAGAGCATGGGTGAGTTTGCGGAGGAGAGGGT
GCTCCCCGTGCTGCACCCCGTGGACAAGCTGTGGCAGCCGCAGGACTTCCTGCCC
GACCCCGAGTCGCCCAGCTTCGAGGACCAGGTGGCGGAGCTGCGCGCGCGCGCC
AAGGACCTGCCCCGACGAGTACTTTGTGGTGCTGGTGGGCGACATGATCACGGAG
GAGGCGCTGCCGACCTACATGGCCATGCTCAACACCTTGGACGGTGTGCGCGAC
GACACGGGCGCGGCTGACCACCCGTGGGCGCGCTGGACGCGGCAGTGGGTGGCC
GAGGAGAACCGGCACGGCGACCTGCTGAACAAGTACTGTTGGCTGACGGGGCGC
GTCAACATGCGGGCCGTGGAGGTGACCATCAACAACCTGATCAAGAGCGGCATG
AACCCGCAGACGGACAACAACCTTACTTGGGCTTCGTCTACACCTCCTTCCAGG
AGCGCGCCACCAAGTAGGTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGT
AGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTC
GACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCG
AGCGCTGTTTAAATAGCCAGGCCCGCGATTGCAAAGACATTATAGCGAGCTACC
AAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCT
TGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCC
GCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGC
CGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCA
CTTCACCCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGA
GAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTG
GGGGACGCCCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGA
GGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGC
TCCATGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACC
CGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGC
AGTACATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGA
ACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTA
CGAGCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGAT
CGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTCCGCC
AACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCA
CCGAGCAGGACCCAGCAAGTCTACTGGGTGATGTTTCATCTCCATCAACCCCGG
CGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTGCGCAGCTTCAACGGCACC
CACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACT

ACGCCCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCAT
 CGCGTGGGCTTCCAACCTGGGAGTACTCCGCCTTCGTGCCACCAACCCCTGGCGC
 TCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAAC
 CGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACG
 CCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCT
 ACAACGTCGACCTGTCCAACAGCACCCGGCACCCCTGGAGTTCGAGCTGGTGTACG
 CCGTCAACACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTG
 GTTCAAGGGCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTC
 CGCGTCCTCCTTCTTCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGA
 GAACCCCTACTTCACCAACCGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGA
 GAACGACCTGTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGA
 GCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACC
 GGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTAC
 ATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATTGGCAGCAGCAGCTCGG
 ATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACT
 TGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTT
 TGATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAAT
 ACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACT
 TATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCT
 CGCACAGCCTTGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAA
 CCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAA
 GGATCGTAGAGCTCCAGCCACGGCAACACCGCGCGCCTGGCGGCCGAGCACGGC
 GACAAGGGCCTGAGCAAGATCTGCGGGCTGATCGCCAGCGACGAGGGCCGGCAC
 GAGATCGCCTACACGCGCATCGTGGACGAGTTCTTCCGCCTCGACCCCGAGGGCG
 CCGTCGCCGCTACGCCAACATGATGCGCAAGCAGATCACCATGCCCCGCGCACCT
 CATGGACGACATGGGCCACGGCGAGGCCAACCCGGGCCGCAACCTCTTCGCCGA
 CTTCTCCGCCGTCGCCGAGAAGATCGACGTCTACGACGCCGAGGACTACTGCCGC
 ATCCTGGAGCACCTCAACGCGCGCTGGAAGGTGGACGAGCGCCAGGTGAGCGGC
 CAGGCCGCCGCGGACCAGGAGTACGTTCTGGGCCTGCCCCAGCGCTTCCGAAA
 CTCGCCGAGAAGACCGCCGCCAAGCGCAAGCGCGTCGCGCGCAGGCCCGTCGCC
 TTCTCCTGGAGAGAAGAGCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGT
 AATCATGGTCATAGCTGTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCACAC
 AACATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGC
 TAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGT
 CGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCCTA
 TTGGGCGCTCTTCC

SEQ ID NO: 78

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

GCTCTTCCCGCCTGGAGCTGGTGCAGAGCATGGGGCAGTTTGCGGAGGAGAGGG
 TGCTCCCCGTGCTGCACCCCGTGGACAAGCTGTGGCAGCCGACGACTTCCTGCC
 CGACCCCGAGTCGCCCGACTTCGAGGACCAGGTGGCGGAGCTGCGCGCGCGCGC
 CAAGGACCTGCCCGACGAGTACTTTGTGGTGCTGGTGGGCGACATGATCACGGA
 GGAGGCGCTGCCGACCTACATGGCCATGCTCAACACCTTGGACGGTGTGCGCGA
 CGACACGGGCGCGGCTGACCACCCGTGGGCGCGCTGGACGCGGCAGTGGGTGGC
 CGAGGAGAACCGGCACGGCGACCTGCTGAACAAGTACTGTTGGCTGACGGGGCG
 CGTCAACATGCGGGCCGTGGAGGTGACCATCAACAACCTGATCAAGAGCGGCAT

GAACCCGCAGACGGACAACAACCCTTACTTGGGCTTCGTCTACACCTCCTTCCAG
GAGCGCGCCACCAAGTAGGTACC

SEQ ID NO: 79

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

CAGCCACGGCAACACCGCGCGCCTTGGCGGCCGAGCACGGCGACAAGAACCTGAG
CAAGATCTGCGGGCTGATCGCCAGCGACGAGGGCCGGCACGAGATCGCCTACAC
GCGCATCGTGGACGAGTTCTTCCGCCTCGACCCCGAGGGCGCCGTCGCCGCCTAC
GCCAACATGATGCGCAAGCAGATCACCATGCCCCGCGCACCTCATGGACGACATG
GGCCACGGCGAGGCCAACCCGGGCGCAACCTCTTCGCCGACTTCTCCGCGGTCCG
CCGAGAAGATCGACGTCTACGACGCCGAGGACTACTGCCGCATCCTGGAGCACC
TCAACGCGCGCTGGAAGGTGGACGAGCGCCAGGTCAGCGGCCAGGCCGCCGCGG
ACCAGGAGTACGTCCTGGGCCTGCCCCAGCGCTTCCGGAAACTCGCCGAGAAGA
CCGCCGCCAAGCGCAAGCGCGTCGCGCGCAGGCCCGTCGCCTTCTCCTGGAGAA
GAGCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGC
TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGG
AAGCATAAAGTGTAAGCCTGGGGTGCTAATGAGTGAGCTAACTCACATTAAT
TGCGTTGCGCTCACTGCCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCAT
TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 80

Prototheca moriformis SAD2B knockout homologous recombination targeting construct

GCTCTTCCCGCCTGGAGCTGGTGCAGAGCATGGGGCAGTTTGCGGAGGAGAGGG
TGCTCCCCGTGCTGCACCCCGTGGACAAGCTGTGGCAGCCGCAGGACTTCCTGCC
CGACCCCGAGTCGCCCCGACTTCGAGGACCAGGTGGCGGAGCTGCGCGCGCGCGC
CAAGGACCTGCCCCGACGAGTACTTTGTGGTGCTGGTGGGCGACATGATCACGGA
GGAGGCGCTGCCGACCTACATGGCCATGCTCAACACCTTGGACGGTGTGCGCGA
CGACACGGGCGCGGCTGACCACCCGTGGGCGCGCTGGACGCGGCGAGTGGGTGGC
CGAGGAGAACC GGACGCGACCTGCTGAACAAGTACTGTTGGCTGACGGGGCG
CGTCAACATGCGGGCCGTGGAGGTGACCATCAACAACCTGATCAAGAGCGGCAT
GAACCCGCAGACGGACAACAACCCTTACTTGGGCTTCGTCTACACCTCCTTCCAG
GAGCGCGCCACCAAGTAGGTACCCTTTCTTGCCTATGACACTTCCAGCAAAAGG
TAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTT
CGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGC
GAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTAC
CAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGC
TTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCC
GCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGC
CGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCA
CTTCACCCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGA
GAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTG
GGGACGCCCTTGTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGA
GGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGC
TCCATGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACC
CGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGC
AGTACATCTCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGA
ACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTA

CGAGCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGAT
CGAGATCTACTCCTCCGACGACCTGAAGTCTGGAAGCTGGAGTCCGCGTTCGCC
AACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCA
CCGAGCAGGACCCCAGCAAGTCTACTGGGTGATGTTTCATCTCCATCAACCCCGG
CGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTCGGCAGCTTCAACGGCACC
CACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACT
ACGCCCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCAT
CGCGTGGGCCTCCAACCTGGGAGTACTCCGCTTTCGTGCCACCAACCCCTGGCGC
TCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCC
CGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACG
CCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCT
ACAACGTCGACCTGTCCAACAGCACCGGCACCCTGGAGTTCGAGCTGGTGTACG
CCGTCAACACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTG
GTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTC
CGCGTCCCTCCTTCTTCTGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGA
GAACCCCTACTTCACCAACCGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGA
GAACGACCTGTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGA
GCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACC
GGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTAC
ATCGACAAGTTCAGGTGCGCGAGGTCAAGTGACAATTGGCAGCAGCAGCTCGG
ATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACT
TGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTT
TGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAAT
ACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACT
TATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCT
CGCACAGCCTTGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAA
CCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAC
AGCCACGGCAACACCGCGCGCCTTGCGGCCGAGCACGGCGACAAGAACCTGAGC
AAGATCTGCGGGCTGATCGCCAGCGACGAGGGCCGGCACGAGATCGCCTACACG
CGCATCGTGGACGAGTTCTTCCGCCTCGACCCCGAGGGCGCCGTCGCCGCTACG
CCAACATGATGCGCAAGCAGATCACCATGCCCGCGCACCTCATGGACGACATGG
GCCACGGCGAGGCCAACCCGGGCCGCAACCTCTTCGCCGACTTCTCCGCGGTTCG
CGAGAAGATCGACGTCTACGACGCCGAGGACTACTGCCGCATCCTGGAGCACCT
CAACGCGCGCTGGAAGGTGGACGAGCGCCAGGTCAGCGGCCAGGCCGCCGCGG
ACCAGGAGTACGTCCTGGGCCTGCCCCAGCGCTTCCGGAACTCGCCGAGAAGA
CCGCCGCCAAGCGCAAGCGCGTCGCGCGCAGGCCCGTCGCCTTCTCCTGGAGAA
GAGCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGC
TGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGG
AAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAAT
TGCGTTGCGCTCACTGCCCGCTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCAT
TAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC

SEQ ID NO: 81

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

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

AGAGAGCGGAGGTGGGGTTGTGAGGTGGGGTTGCTGACCAGGAGCTCGCGTCGC
CGAGCGGCACTCGCACACGGTCCAGTTACCCCCCTCCGCCAAACGCAAGCCT
CCCATCTTGATGCCTTTCGGGCCACCTATACTATTTCTTAGTTCGCTGTAACATCC
AGACCGTCCTGAATAATAACAATGCCCTGTGTCAAGTGCATTCCTAAAAAATTC

TGTCCCAACCAACAATCCCACCTGAAATACCACCAGCCCTGCCCAGTACACTCTT
CCAATACCATCTCCCTACCTCCACGCGCAAGCGACCCCCATGCGCGACCAGGCTC
GAAAGTGATTTATGACTTGAGACGAGCGAGTGGCGGCGCGGTCGACTGCCTTTTC
ATCACGTGCCGTACGTGCGCGACCGCTAGGGCTTTGCACGGCAACGCACGGCTTC
GCCAACCCGACCAGCCAGGACCTCGACTACTCTACCGCGAATTGCGCTCAAGAA
GTCGCCAAATGTGCCATACACCATTCCTTACAGCACTGTTCAAACCTTGATGCCAA
TTTTGACATTCGGGTTGCTCGTTGGCTGCGCCCACATCGGCCCGTGAGTGCAGCAG
GCGGGATCGGACACGGAGGACGCGGCGTCAACGCCCCGAACGCAGCCCGTAACCTC
TACATCAACACGACGTGTTGCGTAATCCCGCCCCGGCTGCGCATCGTGCCAACCCA
TTCGCGATGGATGGTCGGAAAAATGGTGTGCCAACTGCCCTGAGGGAGGCTCTCG
CGAAACGGGCACGTCCCTGAAACCGAAACTGTGGCCTTGTCGTGCGGCCACGCAA
GCACGTGGACCCTAAACACCAAGAAAAATCAGTAAACAAGGTTGACATCCTCTAC
GGGCGAATTGTTTGCCCAACCCCTTCATCGCACACTGCCATTATAATGCATCTAGC
TCGGCGACAAGTTTAGAAAAAGGCAGGCTGCATTGTTCCATTTTCGCCGTGGCGGCG
TGGGTGCCCATTTTACGAGGTTTGGGCTCCCGGGCAGCGACCGAGCCAGGTGCA
GTCCCTCTCGCCCGTCGACAATGTTGCGAACCCCAAGCGGCTAACAACAACCTT
GATGGTACCTGTACACTGCCAATTCCCTTCTTCCCCGGCCGAGGTTTACACGTGAT
GGCCATGGCTTCGCATTCAGGCCGACTTCCCATTCCGACTTTCCAGAGGGTCCGC
GGACGCTGGGGGTTGGCTGCCTGAGGCCACCCCTTTGTTCCCCGCGTCCCGACAA
ACACAATTGCGTTACATAAGGGGGAGCCGCCCCCGTTTCAGAGTGCAGAAATCTTT
CACTATATTTTCCAGTCGTCAGCGAAATCAAGTACTAGTATGGCCACCGCATCCA
CTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGG
GCCCCGGCGCCAGCGAGGCCCTCCCGTGCGCGGGCGCGCCCCCGACTGGTC
CATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCGCCGAGAAGCAGTGGACC
AACCTGGAGTGGAAGCCCAAGCCCAACCCCCCAGCTGCTGGACGACCACTTC
GGCCCCACGGCCTGGTGTTCGCGCGACCTTCGCCATCCGCAGCTACGAGGTGG
GCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCCG
CCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCTT
GGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCCGT
GGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTGGAGTGCTGGGTGGGCGC
CTCCGGCAACAACGGCCGCGCCACGACTTCCTGGTGCGCGACTGCAAGACCGG
CGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGACCCCG
CGCCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCCGCGCTTCATC
GACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGA
CTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCGCTGGAACGACCTGGA
CATCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACCGT
GCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTCACCATCGAGTACCGC
CGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCGGC
TCCTCCGAGGCCGCGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGC
GAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCCGC
GGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCACGAC
GGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGACTC
GAGTTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCC
CTCGTTTCATATCGCTTGATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTT

SEQ ID NO: 82

5' 6S rRNA genomic donor sequence

GCTCTTCGCCGCCGCCACTCCTGCTCGAGCGCGCCCGCGCGTGCGCCGCCAGCGC
CTTGGCCTTTTCGCCGCGCTCGTGCGCGTGCCTGATGTCCATCACCAGGTCCATG
AGGTCTGCCTTGCGCCGGCTGAGCCACTGCTTCGTCCGGGCGGCCAAGAGGAGC
ATGAGGGAGGACTCCTGGTCCAGGGTCCTGACGTGGTTCGCGGCTCTGGGAGCGG
GCCAGCATCATCTGGCTCTGCCGCACCGAGGCCGCCTCCAACCTGGTCTCCAGCA
GCCGCAGTCGCCGCCGACCCTGGCAGAGGAAGACAGGTGAGGGGGGTATGAATT
GTACAGAACAAACCAGAGCCTTGTCTAGGCAGAATCCCTACCAGTCATGGCTTTA
CCTGGATGACGGCCTGCGAACAGCTGTCCAGCGACCCTCGCTGCCGCCGCTTCTC
CCGCACGCTTCTTTCCAGCACCGTGATGGCGCGAGCCAGCGCCGCACGCTGGCGC
TGCGCTTCGCCGATCTGAGGACAGTCGGGGAACCTCTGATCAGTCTAAACCCCTT
GCGCGTTAGTGTTGCCATCCTTTGCAGACCGGTGAGAGCCGACTTGTTGTGCGCC
ACCCCCACACCACCTCCTCCCAGACCAATTCTGTACCTTTTTTGCGAAGGCAT
CGGCCTCGGCCTGCAGAGAGGACAGCAGTGCCAGCCGCTGGGGGTGGCGGAT
GCACGCTCAGGTACC

SEQ ID NO: 83

Relevant expression construct for *Cinnamomum camphora* thioesterase

(β tub::neo::nitred:: β tub::C.camphora TE::nitred)

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACCTCTAGAATATCAATGATC
GAGCAGGACGGCCTCCACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTC
GGCTACGACTGGGCCCAGCAGACCATCGGCTGCTCCGACGCCGCGCTGTTCGCC
TGTCGCCCCAGGGCCGCCCGCTGCTGTTCTGTGAAGACCGACCTGTCCGGCGCCCT
GAACGAGCTGCAGGACGAGGCCGCCCGCTGTCTTGGCTGGCCACCACCGGCGT
GCCCTGCGCCGCGGTGCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCT
GCTGGGCGAGGTGCCCGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAG
AAGGTGTCCATCATGGCCGACGCCATGCGCCGCTGCACACCCTGGACCCCGCCA
CCTGCCCTTCGACCACCAGGCCAAGCACCGCATCGAGCGCGCCCGCACCCGCAT
GGAGGCCGGCCTGGTGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGG
CCCCCGCCGAGCTGTTCCGCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCTGCCTGCCCAACATCATGGTGGAGAACGGCC
GCTTCTCCGGCTTCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGA
CATCGCCCTGGCCACCCGCGACATCGCCGAGGAGCTGGGCGGCGAGTGGGCCGA
CCGCTTCCTGGTGCTGTACGGCATCGCCGCCCCCGACTCCCAGCGCATCGCCTTC
TACCGCCTGCTGGACGAGTTCTTCTGACAATTGGCAGCAGCAGCTCGGATAGTAT
CGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCC
TTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTT
GTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCC
CCAGCATCCCCTTCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTAC
GCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAG
CCTTGGTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACT

GCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCCG
GTCTCGAACAGAGCGCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTGCGACCTC
AGCGCGGCATACACCACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCAT
TAGCGAAGCGTCCGGTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAA
TGATCGGTGGAGCTGATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTT
TCTTGCGCTATGACACTTCCAGCAAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCC
GGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTG
CATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCC
GATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCA
CTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCG
CCTCTTCTCTTCGTTTCAGTCACAACCCGCAAACTAGTATGGCCACCGCATCC
ACTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCG
GGCCCCGGCGCCAGCGAGGCCCCCTCCCCGTGCGCGGGCGCGCCCCCGACTGGT
CCATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGAC
CAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCAGCTGCTGGACGACCACTT
CGGCCCCCACGGCCTGGTGTTCGCCCGCACCTTCGCCATCCGCAGCTACGAGGTG
GGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCC
GCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCACC
CTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCC
GTGGAGCGCTACCCCGCTGGGGCGACACCGTGGAGGTGGAGTGCTGGGTGGGC
GCCTCCGGCAACAACGGCCGCCGCCACGACTTCCTGGTGCGCGACTGCAAGACC
GGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCC
GCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGGCGAGATCGGCCCCGCCTTCA
TCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAAC
GACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTG
GACATCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACC
GTGCCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTCACCATCGAGTACC
GCCGCGAGTGCAACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCG
GCTCCTCCGAGGCCCGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCA
GCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCC
GCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCACG
ACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGA
CTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGT
GATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTT
TATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTA
GCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATAT
CGCTTGCAATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCT
CCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTGTTGGGCTCCGCCTGTATTCT
CCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGT
GGGATGGGAACACAAATGGAAAGCTT

SEQ ID NO: 84

3'6S rRNA genomic donor sequence

GAGCTCCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGA
TAACCTCCAAAGCCGCTCTAATTGTGGAGGGGGTTTCAATTTAAAAGCTTGGAAT
GTTGGTTTCGTGCGTCTGGAACAAGCCCAGACTTGTGCTCACTGGGAAAAGGACC
ATCAGCTCCAAAAAACTTGCCGCTCAAACCGCGTACCTTGCTTTCGCGCAATCT
GCCCTGTTGAAATCGCCACCACATTCATATTGTGACGCTTGAGCAGTCTGTAATT
GCCTCAGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTGCG

GGGCGAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAGTT
CATAACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAG
TGGCCACCCCCCGGCCCTGGTGCTTGC GGAGGGCAGGTCAACCGGCATGGGGCT
ACCGAAATCCCCGACCGGATCCCACCACCCCCGCGATGGGAAGAATCTCTCCCC
GGGATGTGGGCCCCACCACCAGCACAACCTGCTGGCCCAGGCGAGCGTCAAACCA
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
GTGGCAAACGCTGGCGCCCGCGAGCCGGGCGGCGGCGATGCGGTGCCCCACGG
CTGCCGGAATCCAAGGGAGGCAAGAGCGCCCGGGTCAGTTGAAGGGCTTTACGC
GCAAGGTACAGCCGCTCCTGCAAGGCTGCGTGGTGGAATTGGACGTGCAGGTCC
TGCTGAAGTTCTCCACCGCCTCACCAGCGGACAAAGCACCGGTGTATCAGGTCC
GTGTCATCCACTCTAAAGAGCTCGACTACGACCTACTGATGGCCCTAGATTCTTC
ATCAAAAACGCCTGAGACACTTGCCCAGGATTGAAACTCCCTGAAGGGACCACC
AGGGGCCCTGAGTTGTTCCCTTCCCCCGTGGCGAGCTGCCAGCCAGGCTGTACCT
GTGATCGAGGCTGGCGGGAAAATAGGCTTCGTGTGCTCAGGTTCATGGGAGGTGC
AGGACAGCTCATGAAACGCCAACAATCGCACAATTCATGTCAAGCTAATCAGCT
ATTTCTCTTACGAGCTGTAATTGTCCCAAATTCCTGGTCTACCGGGGGTGATCC
TTCGTGTACGGGCCCTTCCCTCAACCCTAGGTATGCGCGCATGCGGTGCCCGCGC
AACTCGCGCGAGGGCCGAGGGTTTGGGACGGGCGTCCCGAAATGCAGTTGCAC
CCGGATGCGTGGCACCTTTTTTTCGATAATTTATGCAATGGACTGCTCTGCAAAA
TTCTGGCTCTGTGCGCAACCCTAGGATCAGCGGCGTAGGATTTTCGTAATCATTCG
TCCTGATGGGGAGCTACCGACTACCCTAATATCAGCCCGACTGCCTGACGCCAGC
GTCCACTTTTGTGCACACATTCCATTTCGTGCCCAAGACATTTTCATTGTGGTGCGAA

GCGTCCCCAGTTACGCTCACCTGTTTCCCGACCTCCTTACTGTTCTGTGCGACAGAG
CGGGCCACAGGCCGGTCGCAGCC

SEQ ID NO: 90

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

GCTCTTCGGCCGCGCCACTCCTGCTCGAGCGCGCCCGACTCGCGCTCCGCCTGC
GCCCCGCGCTGCGCCGCCAGCGCCTTGGCCTTTTCGCCGCGCTCGTGCGCGTCGC
TGATGTCCATCACCAGGTCCATGAGGTCTGCCTTGCGCCGGCTGAGCCACTGCTT
CGTCCGGGCGGCCAAGAGGAGCATGAGGGAGGACTCCTGGTCCAGGGTCCTGAC
GTGGTCGCGGCTCTGGGAGCGGGCCAGCATCATCTGGCTCTGCCGCACCGAGGC
CGCCTCCAACCTGGTCCCTCCAGCAGCCGCAGTCGCCGCCGACCCTGGCAGAGGAA
GACAGGTGAGGGGGGTATGAATTGTACAGAACAACCACGAGCCTTGTCTAGGCA
GAATCCCTACCAGTCATGGCTTTACCTGGATGACGGCCTGCGAACAGCTGTCCAG
CGACCCTCGCTGCCGCCGCTTCTCCCGCACGCTTCTTTCCAGCACCGTGATGGCG
CGAGCCAGCGCCGCACGCTGGCGCTGCGCTTCGCCGATCTGAGGACAGTCGGGG
AACTCTGATCAGTCTAAACCCCCCTTGC GCGTTAGTGTGTCATCCTTTGCAGACC
GGTGAGAGCCGACTTGTGTGCGCCACCCCCACACCACCTCCTCCCAGACCAAT
TCTGTCACCTTTTTTGGCGAAGGCATCGGCCTCGGCCTGCAGAGAGGACAGCAGTG
CCCAGCCGCTGGGGGTTGGCGGATGCACGCTCAGGTACCCTTTCTTGCGCTATGA
CACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGC
AACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCC
GATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCGCGATTGCAAAGA
CATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTA
CACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTT
CGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTC
CTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCC
GACCGCCCCCTGGTGCACCTTCACCCCCAACAAGGGCTGGATGAACGACCCCAAC
GGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAAC
CCGAACGACACCGTCTGGGGGACGCCCTTGTTCCTGGGGCCACGCCACGTCCGAC
GACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGAC
TCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCT
TCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACA
CCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTT
CACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGA
CCCGAAGGTCTTCTGGTACGAGCCCTCCAGAAAGTGGATCATGACCGCGGCCAA
GTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAA
GCTGGAGTCCGCGTTCGCCAACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCC
GGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCCTACTGGGTGATG
TTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCG
TCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGT
GGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACC
TACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCCTTCG
TGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAAC
ACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCG
ATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACG
TTGACGAAGGCCAACAGCTACAACGTGACCTGTCCAACAGCACCGGCACCCTG
GAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGT
TCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCG
CATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCTTCTGGACCGCGGGAACAGCAAG

GTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAAC
CAGCCCTTCAAGAGCGAGAACGACCTGTCCTACTACAAGGTGTACGGCTTGCTGG
ACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACA
CCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGG
TGGACAACCTGTTCTACATCGACAAGTTCCAGGTGCGCGAGGTCAAGTGACAATT
GGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGG
ACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCA
AACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGGCAGTTGCTAGCTGC
TTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTCCCTCGTTTCATATCGCTT
GCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGC
TCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGG
TACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGAT
GGGAACACAAATGGAGGATCCCGCGTCTCGAACAGAGCGCGCAGAGGAACGCT
GAAGGTCTCGCCTCTGTGCGACCTCAGCGCGGCATACACCACAATAACCACCTGA
CGAATGCGCTTGTTCTTCGTCCATTAGCGAAGCGTCCGGTTCACACACGTGCCA
CGTTGGCGAGGTGGCAGGTGACAATGATCGGTGGAGCTGATGGTCGAAACGTTT
ACAGCCTAGGGATATCGAATTCGGCCGACAGGACGCGCGTCAAAGGTGCTGGTC
GTGTATGCCCTGGCCGGCAGGTGCTTGTGCTGCTGGTTAGTGATTCCGCAACCC
TGATTTTGGCGTCTTATTTTGGCGTGGCAAACGCTGGCGCCCGCGAGCCGGGCCG
GCGGCGATGCGGTGCCCCACGGCTGCCGGAATCCAAGGGAGGCAAGAGCGCCCCG
GGTCAGTTGAAGGGCTTTACGCGCAAGGTACAGCCGCTCCTGCAAGGCTGCGTG
GTGGAATTGGACGTGCAGGTCTGCTGAAGTTCCTCCACCGCCTCACCAGCGGAC
AAAGCACCGGTGTATCAGGTCCGTGTCATCCACTCTAAAGAGCTCGACTACGACC
TACTGATGGCCCTAGATTCTTCATCAAAAACGCCTGAGACACTTGCCCAGGATTG
AAACTCCCTGAAGGGACCACCAGGGGCCCTGAGTTGTTCTTCCCCCGTGGCGA
GCTGCCAGCCAGGCTGTACCTGTGATCGAGGCTGGCGGGAAAATAGGCTTCGTG
TGCTCAGGTCATGGGAGGTGCAGGACAGCTCATGAAACGCCAACAATCGCACAA
TTCATGTCAAGCTAATCAGCTATTTCTCTTCACGAGCTGTAATTGTCCAAAATT
CTGGTCTACCGGGGGTGATCCTTCGTGTACGGGGCCCTTCCCTCAACCCTAGGTAT
GCGCGCATGCGGTGCGCGGCAACTCGCGCGAGGGCCGAGGGTTTGGGACGGGC
CGTCCCGAAATGCAGTTGCACCCGGATGCGTGGCACCTTTTTTGGCGATAATTTAT
GCAATGGACTGCTCTGCAAAATTCTGGCTCTGTGCGCAACCCTAGGATCAGCGGC
GTAGGATTTCTGTAATCATTCGTCCTGATGGGGAGCTACCGACTACCCTAATATCA
GCCCCACTGCCTGACGCCAGCGTCCACTTTTGTGCACACATTCCATTCGTGCCA
AGACATTTCAATTGTGGTGCGAAGCGTCCCCAGTTACGCTCACCTGTTTCCCGACC
TCCTTACTGTTCTGTGCGACAGAGCGGGCCACAGGCCGGTCGCGAGCCACTAGTAT
GGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGGCGACCTGCGTCGC
TCGGCGGGCTCCGGGGCCCCGGCGCCCCAGCGAGGCCCTCCCCGTGCGCGGGGCGC
GCCCCCGACTGGTCCATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCCGCCG
AGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCCAGCTGC
TGGACGACCACTTCGGCCCCCACGGCCTGGTGTTCGCGCGCACCTTCGCCATCCG
CAGCTACGAGGTGGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCA
CCTGCAGGAGGGCCGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGG
CTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCG
CACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTGGA
GTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCCTGGTGCG
CGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGAT
GAACACCCGCACCCGCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGCCGAGAT
CGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCC
CCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCG

CTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGACTG
GATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTC
ACCATCGAGTACCGCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGACC
ACCGTGAGCGGCGGCTCCTCCGAGGCGGCTTGGTGTGCGAGCACCTGCTGCAG
CTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTG
ACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACT
ACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGAC
GACGACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT
CGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCC
TTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTT
GTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCC
CCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTAC
GCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAG
CCTTGGTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACT
GCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTGTAG
AGCTCCTTGTTCAGAAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGAT
AACCTCCAAAGCCGCTCTAATTGTGGAGGGGGTTCGAATTTAAAGCTTGGAAATG
TTGGTTTCGTGCGTCTGGAACAAGCCAGACTTGTGCTCACTGGGAAAAGGACCA
TCAGCTCCAAAAAACTTGCCGCTCAAACCGCGTACCTCTGCTTTTCGCGCAATCTG
CCCTGTTGAAATCGCCACCACATTTCATATTGTGACGCTTGAGCAGTCTGTAATTG
CCTCAGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCGCG
GGCGAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAGTTC
ATAACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAGT
GGCCACCCCCCGGCCCTGGTGCTTGCAGGAGGGCAGGTCAACCGGCATGGGGCTA
CCGAAATCCCCGACCGGATCCCACCACCCCCGCGATGGGAAGAATCTCTCCCCG
GGATGTGGGCCCACCACCAGCACAACTGCTGGCCCAGGCGAGCGTCAAACCAT
ACCACACAAATATCCTTGGCATCGGCCCTGAATTCCCTTCTGCCGCTCTGCTACCC
GGTGCTTCTGTCCGAAGCAGGGGTGCTAGGGATCGCTCCGAGTCCGCAAACCCCT
TGTCGCGTGCGGGGCTTGTTCGAGCTTGTTCGAGCTTGAAGAGCCTCTAGAGTC
GACCTGCAGGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCCTGTGTGA
AATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA
AAGCCTGGGGTGCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACT
GCCCCGCTTTCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAA
CGCGCGGGGAGAGGCGGTTTTCGTATTGGGCGCTCTTCC

SEQ ID NO: 91

Codon-optimized *C. camphora* thioesterase sequence

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTC
GCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCCCTCCCCGTGCGCGGGC
GCGCCCCCGACTGGTCCATGCTGTTCCGCGTGATCACCACCATCTTCTCCGCCGC
CGAGAAGCAGTGGAACCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCAGCT
GCTGGACGACCCTTCGGCCCCCACGGCCTGGTGTTCGCGCGCACCTTCGCCATC
CGCAGCTACGAGGTGGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAAC
CACCTGCAGGAGGCCGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGAC
GGCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAG
CGACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTG
GAGTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCCTGGTG
CGCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATG
ATGAACACCCGCACCCGCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAG

ATCGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAG
CCCCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCC
CGCTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGAC
TGGATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCT
TCACCATCGAGTACCGCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGAC
CACCGTGAGCGGCGGCTCCTCCGAGGCCGCGCCTGGTGTGCGAGCACCTGCTGCA
GCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCT
GACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGAC
TACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGAC
GACGACAAGTGA

SEQ ID NO: 92

Codon-optimized *U. californica* thioesterase sequence

GGCGCGCCCCCGACTGGTCCATGCTGTTTCGCCGTGATCACCACCATCTTCAGCGC
CGCCGAGAAGCAGTGGACCAACCTGGAGTGGGAAGCCCAAGCCCAAGCTGCCCA
GCTGCTGGACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCCCGCACCTTCGCC
ATCCGCTCCTACGAGGTGGGCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGA
ACCACATGCAGGAGGCCACCCTGAACCACGCCAAGAGCGTGGGCATCCTGGGCG
ACGGCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATGTGGGTGGTGC
GCCGACCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTGGAGG
TGGAGTGTCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGG
TGCGCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCT
GATGAACACCCGCACCCGCCGCTGAGCACCATCCCCGACGAGGTGCGCGGCGA
GATCGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGACGAGATCAAGAA
GCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCC
CCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGC
CTGGGTGTTCGAGACCGTGCCCGACAGCATCTTCGAGTCCCACCACATCAGCTCC
TTCACCCTGGAGTACCGCCGCGAGTGCACCCGCGACTCCGTGCTGCGCAGCCTGA
CCACCGTGAGCGGCGGCAGCTCCGAGGCCGCGCCTGGTGTGCGACCACTGCTGC
AGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCCGACCGAGTGGCGCCCCAAGC
TGACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGCGTGATGGA
CTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGA
CGACGACAAGTGACTCGAG

SEQ ID NO: 93

Codon-optimized *U. americana* thioesterase sequence

GGCGCGCCCAGCTGCCCCGACTGGAGCATGCTGCTGGCCGCGATCACCACCCTGTT
CCTGGCGGCCGAGAAGCAGTGGATGATGCTGGACTGGAAGCCCAAGCGCCCCGA
CATGCTGGTGGACCCCTTCGGCCTGGGCCGCTTCGTGCAGGACGGCCTGGTGTTC
CGCAACAACCTCAGCATCCGCAGCTACGAGATCGGCGCGGACCGCACCGCCAGC
ATCGAGACCCTGATGAACCACCTGCAGGAGACCGCCCTGAACCACGTGAAGAGC
GTGGGCCTGCTGGAGGACGGCCTGGGCAGCACCCGCGAGATGAGCCTGCGCAAC
CTGATCTGGGTGGTGACCAAGATGCAGGTGGCGGTGGACCGCTACCCACCTGG
GGCGACGAGGTGCAGGTGAGCAGCTGGGCGACCGCCATCGGCAAGAACGGCAT
GCGCCGCGAGTGGATCGTGACCGACTTCCGCACCGGCGAGACCCTGCTGCGCGC
CACCAGCGTGTGGGTGATGATGAACAAGCTGACCCGCCGCGATCAGCAAGATCCC
CGAGGAGGTGTGGCACGAGATCGGCCCCAGCTTCATCGACGCGCCCCCCTGCC
CACCGTGGAGGACGACGGCCGCAAGCTGACCCGCTTCGACGAGAGCAGCGCCGA

CTTCATCCGCAAGGGCCTGACCCCCCGCTGGAGCGACCTGGACATCAACCAGCA
CGTGAACAACGTGAAGTACATCGGCTGGCTGCTGGAGAGCGCGCCCCCGAGAT
CCACGAGAGCCACGAGATCGCCAGCCTGACCCTGGAGTACCGCCGCGAGTGCGG
CCGCGACAGCGTGCTGAACAGCGCCACCAAGGTGAGCGACAGCAGCCAGCTGGG
CAAGAGCGCCGTGGAGTGCAACCACCTGGTGCGCCTGCAGAACGGCGGCGAGAT
CGTGAAGGGCCGACCGTGTGGCGCCCCAAGCGCCCCCTGTACAACGACGGCGC
CGTGGTGGACGTGCCCCGCAAGACCAGCGATGACGATGACAAGCTGGGATGACT
CGAG

SEQ ID NO: 94

Codon-optimized *C. hookeriana* thioesterase sequence

ACTAGTATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTACG
ATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTC
GTGGGCGCGCCCAGCTGCCCCGACTGGAGCCGCCTGCTGACCGCCATCACCACCGT
GTTCTGTGAAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGA
CATGCTGGTGGACAGCTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTC
CGCCAGTCCTTCTCCATCCGCTCCTACGAGATCGGCACCGACCGCACCGCCAGCA
TCGAGACCCTGATGAACCACCTGCAGGAGACCTCCCTGAACCACTGCAAGAGCA
CCGGCATCCTGCTGGACGGCTTCGGCCGCACCCTGGAGATGTGCAAGCGCGACCT
GATCTGGGTGGTGATTAAGATGCAGATCAAGGTGAACCGCTACCCCGCCTGGGG
CGACACCGTGGAGATCAACACCCGCTTCAGCCGCCTGGGCAAGATCGGCATGGG
CCGCGACTGGCTGATCTCCGACTGCAACACCGGCGAGATCCTGGTGC GCGCCACC
AGCGCCTACGCCATGATGAACCAGAAGACCCGCCGCCTGTCCAAGCTGCCCTAC
GAGGTGCACCAGGAGATCGTGCCCCTGTTCGTGGACAGCCCCGTGATCGAGGAC
TCCGACCTGAAGGTGCACAAGTTCAAGGTGAAGACCGGCGACAGCATCCAGAAG
GGCCTGACCCCCGGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGTG
AAGTACATCGGCTGGATCCTGGAGAGCATGCCCACCGAGGTGCTGGAGACCCAG
GAGCTGTGCTCCCTGGCCCTGGAGTACCGCCGCGAGTGCGGCCGCGACTCCGTGC
TGGAGAGCGTGACCGCCATGGACCCCAGCAAGGTGGGCGTGCGCTCCAGTACC
AGCACCTGCTGCGCCTGGAGGACGGCACCGCCATCGTGAACGGCGCCACCGAGT
GGCGCCCCAAGAACGCCGGCGCCAACGGCGCCATCTCCACCGGCAAGACCAGCA
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
ATGTAGTCACCCCCCTCACCAATTGTGCGCCAGTTTGC GCAATCCATAAACTCA
AAACTGCAGCTTCTGAGCTGCGCTGTTCAAGAACACCTCTGGGGTTTGCTCACCC
GCGAGGTCGACGCCAGCATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAG
CCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGC
GCTCGGCGCTTCGTAGCAGCATGTACCTGGCCTTTGACATCGCGGTCATGTCCCT
GCTCTACGTCGCGTCGACGTACATCGACCCTGCGCCGGTGCCTACGTGGGTCAAG
TATGGCGTCATGTGGCCGCTCTACTGGTTCTTCCAGGTGTGTGTGAGGGTTGTGG
TTGCCCGTATCGAGGTCCTGGTGGCGCGCATGGGGGAGAAGGCGCCTGTCCCGCT
GACCCCCCGGCTACCCTCCCGGCACCTTCCAGGGCGCCTTCGGCACGGGTGTCT
GGGTGTGCGCGCACGAGTGCGGCCACCAGGCCTTTTCCTCCAGCCAGGCCATCAA
CGACGGCGTGGGCCTGGTGTTCACAGCCTGCTGCTGGTGGCCTACTACTCCTGG
AAGCACTCGCACCGGGTACC

SEQ ID NO: 104

pSZ1124 (FAD2B) 3' genomic targeting sequence

CCGCCACCACTCCAACACGGGGTGCCTGGACAAGGACGAGGTGTTTGTGCCGCC
GCACCGCGCAGTGGCGCACGAGGGCCTGGAGTGGGAGGAGTGGCTGCCCATCCG
CATGGGCAAGGTGCTGGTCACCCTGACCCTGGGCTGGCCGCTGTACCTCATGTTC

AACGTCGCCTCGCGGCCGTACCCGCGCTTCGCCAACCACCTTTGACCCGTGGTCGC
CCATCTTCAGCAAGCGCGAGCGCATCGAGGTGGTCATCTCCGACCTGGCGCTGGT
GGCGGTGCTCAGCGGGCTCAGCGTGCTGGGCCGCACCATGGGCTGGGCCTGGCT
GGTCAAGACCTACGTGGTGCCCTACCTGATCGTGAACATGTGGCTCGTGCTCATC
ACGCTGCTCCAGCACACGCACCCGCGCTGCCGCACTACTTCGAGAAGGACTGG
GACTGGCTGCGCGGCGCCATGGCCACCGTGGACCGCTCCATGGGCCCGCCCTTCA
TGGACAACATCCTGCACCACATCTCCGACACCCACGTGCTGCACCACCTCTTCAG
CACCATCCCGCACTACCACGCCGAGGAGGCCTCCGCCGCCATCAGGCCCATCCTG
GGCAAGTACTACCAGTCCGACAGCCGCTGGGTGCGCCGCGCCCTGTGGGAGGAC
TGGCGCGACTGCCGCTACGTGTCGCCGACGCGCCCGAGGACGACTCCGCGCTCT
GGTTCACAAGTGAGTGAGTGAGAAGAGC

SEQ ID NO: 105

S. cerevisiae suc2 cassette

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTG
CAGGCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGA
CGAACGAGACGTCCGACCGCCCCCTGGTGCACCTTCACCCCCAACAAGGGCTGGA
TGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGT
ACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCTGGGGCCA
CGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCC
GAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAA
CACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATC
TGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGAC
GGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCC
ACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGATCA
TGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCT
GAAGTCCTGGAAGCTGGAGTCCGCGTTTCGCCAACGAGGGCTTCCTCGGCTACCA
GTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTC
CTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTC
AACCAGTACTTCGTGCGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACC
AGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAA
CACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGA
GTACTCCGCCTTCGTGCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCA
AGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACC
TGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCG
CCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACA
GCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGA
TCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCC
CGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCCTCCTTCTTCCTGGAC
CGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGC
ATGAGCGTGAACAACCAGCCCTTCAAGAGCGAGAACGACCTGTCTACTACAAG
GTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGAC
GTCGTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGA
ACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCG

AGGTCAAGTGACAATTGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGTGCCTTGACCTGTGAATAT
CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
TGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATAACACCCCCAGCATCCCCTTCC
CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCC
CTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTC
CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
CGGGAAGTAGTGGGATGGGAACACAAATGGA

SEQ ID NO: 106

pSZ1125 (FAD2C) 5' genomic targeting sequence

GCTCTTCGAGGGGCTGGTCTGAATCCTTCAGGCGGGTGTACCCGAGAAAGAAA
GGGTGCCGATTTCAAAGCAGACCCATGTGCCGGGCCCTGTGGCCTGTGTTGGCGC
CTATGTAGTACACCCCCCTCACCAATTGTCGCCAGTTTGCACACTCCATAAACTC
AAAACAGCAGCTTCTGAGCTGCGCTGTTCAAGAACACCTCTGGGGTTTGTCTACC
CGCGAGGTTCGACGCCCAGCATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAA
GCCTCCGTTTCACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAG
CGCTCGGCGCTTCGTAGCAGCATGTACCTGGCCTTTGACATCGCGGTTCATGTCCC
TGCTCTACGTGCGCTCGACGTACATCGACCCTGCACCGGTGCCTACGTGGGTCAA
GTACGGCATCATGTGGCCGCTCTACTGGTTCTTCCAGGTGTGTTTGAGGGTTTTGG
TTGCCCCGTATTGAGGTCCTGGTGGCGCGCATGGAGGAGAAGGCGCCTGTCCCGCT
GACCCCCCGGCTACCCTCCCGGCACCTTCCAGGGCGCCTTCGGCACGGGTGTCT
GGGTGTGCGCGCACGAGTGCGGCCACCAGGCCTTTTCTCCAGCCAGGCCATCAA
CGACGGCGTGGGCCTGGTGTTCACAGCCTGCTGCTGGTGGCCTACTACTCCTGG
AAGCACTCGCACCGGGTACC

SEQ ID NO: 107

pSZ1125 (FAD2C) 3' genomic targeting sequence

CCGCCACCACTCCAACACGGGGTGCCTGGACAAGGACGAGGTGTTTGTGCCGCC
GCACCGCGCAGTGGCGCACGAGGGCCTGGAGTGGGAGGAGTGGCTGCCCATCCG
CATGGGCAAGGTGCTGGTCACCCTGACCCTGGGCTGGCCGCTGTACCTCATGTTC
AACGTGCGCTCGCGGCCGTACCCGCGCTTCGCCAACCACCTTTGACCCGTGGTCGC
CCATCTTCAGCAAGCGCGAGCGCATCGAGGTGGTCATCTCCGACCTGGCGCTGGT
GGCGGTGCTCAGCGGGCTCAGCGTGCTGGGCCGCACCATGGGCTGGGCCTGGCT
GGTCAAGACCTACGTGGTGCCCTACCTGATCGTGAACATGTGGCTCGTGCTCATC
ACGCTGCTCCAGCACACGCACCCGGCGCTGCCGCACTACTTCGAGAAGGACTGG
GACTGGCTGCGCGGCGCCATGGCCACCGTGGACCGCTCCATGGGCCCCGCCCTTCA
TGGACAACATCCTGCACCACATCTCCGACACCCACGTGCTGCACCACCTCTTCAG
CACCATCCCGCACTACCACGCCGAGGAGGCCTCCGCCGCCATCAGGCCCATCCTG
GGCAAGTACTACCAGTCCGACAGCCGCTGGGTGCGCCGCGCCCTGTGGGAGGAC
TGGCGCGACTGCCGCTACGTGCTCCCGGACGCGCCCCGAGGACGACTCCGCGCTCT
GGTTCCACAAGTGAGTGAGTGAGAAGAGC

SEQ ID NO: 108

S. carlbergensis MEL1 amino acid sequence

MFAFYFLTACISLKGVFGVSPSYNGLGLTPQMWDNWNTFACDVSEQLLLDTADRI
SDLGLKDMGYKYIILDDCWSSGRSDSDFLVADEQKFPNGMGMHVADHLHNSFLFG

MYSSAGEYTCAGYPGSLGREEEDAQFFANNRVDYLKYDNCYNKGQFGTPEISYHRY
KAMSDALNKTGRPIFYSLCNWGQDLTFYWGSGIANSWRMSGDVTAEFTRPDSRCP
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
GCGTCCTCAACGGAGACCGCACCCACACCACCACCATCGACTCCTCCCCGGCGG
GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCCGGGGGCCCTTGGGGCGCTGCTGGAGCGTGGTGAGAGCGCGGC
GGACGTGCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTGCG
GCCCCGATCCCCCACTGCCCCCATCATCTTGCCCTGTTGTCGTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTCCTTTCTTGCGCTATGACAC
TTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAAC
ACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGAT
GCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCGCGATTGCAAAGACAT
TATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACAC
AGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGT
TTCAGTCACAACCCGCAAACACTAGTATGTTGCGGTTCTACTTCCTGACGGCCTG
CATCTCCCTGAAGGGCGTGTTGCGGCTCTCCCCCTCCTACAACGGCCTGGGCCTG
ACGCCCCAGATGGGCTGGGACAACCTGGAACACGTTGCGCTGCGACGCTCTCCGAG
CAGCTGCTGCTGGACACGGCCGACCGCATCTCCGACCTGGGCCTGAAGGACATG
GGCTACAAGTACATCATCCTGGACGACTGCTGGTCCTCCGGCCGCGACTCCGACG
GCTTCCTGGTTCGCCGACGAGCAGAAGTTCCCCAACGGCATGGGCCACGTCGCCG
ACCACCTGCACAACAACCTCCTTCCTGTTGCGCATGTACTCCTCCGCGGGCGAGTA
CACGTGCGCCGGCTACCCCGGCTCCCTGGGCCGCGAGGAGGAGGACGCCAGTT
CTTCGCGAACAACCGCGTGGACTACCTGAAGTACGACAACCTGCTACAACAAGGG
CCAGTTTCGGCACGCCCCGAGATCTCCTACCACCGCTACAAGGCCATGTCCGACGCC
CTGAACAAGACGGGCGCCCCATCTTCTACTCCCTGTGCAACTGGGGCCAGGACC
TGACCTTCTACTGGGGCTCCGGCATCGCGAACTCCTGGCGCATGTCCGGCGACGT
CACGGCGGAGTTACGCGCCCCGACTCCCGCTGCCCCTGCGACGGCGACGAGTA
CGACTGCAAGTACGCCGGCTTCCACTGCTCCATCATGAACATCCTGAACAAGGCC
GCCCCCATGGGCCAGAACGCGGGCGTCCGGCGGCTGGAACGACCTGGACAACCTG
GAGGTTCGGCGTCCGCAACCTGACGGACGACGAGGAGAAGGCGCACTTCTCCATG
TGGGCCATGGTGAAGTCCCCCTGATCATCGGCGCAACGTGAACAACCTGAAG
GCCTCCTCCTACTCCATCTACTCCAGGCGTCCGTCATCGCCATCAACCAGGACT
CCAACGGCATCCCCGCCACGCGCGTCTGGCGCTACTACGTGTCCGACACGGACG

AGTACGGCCAGGGCGAGATCCAGATGTGGTCCGGCCCCCTGGACAACGGCGACC
 AGGTCGTGGCGCTGCTGAACGGCGGCTCCGTGTCCCGCCCCATGAACACGACCCT
 GGAGGAGATCTTCTTCGACTCCAACCTGGGCTCCAAGAAGCTGACCTCCACCTGG
 GACATCTACGACCTGTGGGCGAACC GCGTCGACAACCTCCACGGCGTCCGCCATCC
 TGGGCCGCAACAAGACCGCCACCGGCATCCTGTACAACGCCACCGAGCAGTCCT
 ACAAGGACGGCCTGTCCAAGAACGACACCCGCCTGTTTCGGCCAGAAGATCGGCT
 CCCTGTCCCCCAACGCGATCCTGAACACGACCCTCCCGGCCACGGCATCGCGTT
 CTACCGCCTGCGCCCCCTCCTCCTGACAATTGAAGCAGCAGCAGCTCGGATAGTAT
 CGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCC
 TTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTT
 GTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCC
 CCAGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTAC
 GCTGTCTGTCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAG
 CTTTGGTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACT
 GCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAG
 CTCAGAATAGTATCGGGTGATGCGAAGTCAGAACCAGGCAGGGCCTGTGCGCTG
 AGGTGGCAACGATGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATCCCG
 TGATCTCCCCCACCACACCTATCCCCGCCATCCCGGCCACCCTTTCAGTCCCC
 TCAGCATGCATTGTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGGTTCAGGC
 CACGGCGCACCGAGTCCTCGCCCTTCGCAGAGTGATCACCTCCCCGGAACCAGC
 CACGCTCGCTGCTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGGACA
 ACTCAGGCAACCACGCGCCTCACAAGCATGGCCGCCGTGGCATCCAACCCGCTC
 GTGACGGTGGGTGCGCAAGTGCCAGGGGCTCGTCGTACGGCGTGCATCCTCG
 AGGGATGCGATCCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGATGA
 GCCCCGTGTCTAGA

SEQ ID NO: 111

Codon-optimized *S. carbergensis* MEL1 sequence

ATGTTTCGCGTTCTACTTCCTGACGGCCTGCATCTCCCTGAAGGGCGTGTTTCGGCGT
 CTCCCCCTCCTACAACGGCCTGGGCTGACGCCCCAGATGGGCTGGGACAACCTGG
 AACACGTTTCGCTGCGACGTCTCCGAGCAGCTGCTGCTGGACACGGCCGACCGC
 ATCTCCGACCTGGGCCTGAAGGACATGGGCTACAAGTACATCATCCTGGACGACT
 GCTGGTCTCCGCGCGGACTCCGACGGCTTCCTGGTCGCCGACGAGCAGAAGTT
 CCCCACGGCATGGGCCACGTCGCCGACCACCTGCACAACAACCTCCTTCCTGTTT
 GGCATGTACTCCTCCGCGGGCGAGTACACGTGCGCCGGCTACCCCGGCTCCCTGG
 GCCGCGAGGAGGAGGACGCCCAGTTCTTCGCGAACAACCGCGTGGAACCTGA
 AGTACGACAACCTGCTACAACAAGGGCCAGTTTCGGCACGCCCCGAGATCTCCTACC
 ACCGCTACAAGGCCATGTCCGACGCCCTGAACAAGACGGGGCCGCCCATCTTCTA
 CTCCCTGTGCAACTGGGGCCAGGACCTGACCTTCTACTGGGGCTCCGGCATCGCG
 AACTCCTGGCGCATGTCCGGCGACGTACGGCGGAGTTCACGCGCCCCGACTCCC
 GCTGCCCCCTGCGACGGCGACGAGTACGACTGCAAGTACGCCGGCTTCCACTGCTC
 CATCATGAACATCCTGAACAAGGCCGCCCCCATGGGCCAGAACGCGGGCGTCGG
 CGGCTGGAACGACCTGGACAACCTGGAGGTCGGCGTCGGCAACCTGACGGACGA
 CGAGGAGAAGGCGCACTTCTCCATGTGGGCCATGGTGAAGTCCCCCTGATCATC
 GCGCGAACGTGAACAACCTGAAGGCCTCCTCCTACTCCATCTACTCCAGGCGT
 CCGTCATCGCCATCAACCAGGACTCCAACGGCATCCCCGCCACGCGCGTCTGGCG
 CTACTACGTGTCCGACACGGACGAGTACGGCCAGGGCGAGATCCAGATGTGGTC
 CGGCCCCCTGGACAACGGCGACCAGGTCGTGGCGCTGCTGAACGGCGGCTCCGT
 GTCCCGCCCCATGAACACGACCCTGGAGGAGATCTTCTTCGACTCCAACCTGGGC

TCCAAGAAGCTGACCTCCACCTGGGACATCTACGACCTGTGGGCGAACCGCGTC
GACAACTCCACGGCGTCCGCCATCCTGGGCCGCAACAAGACCGCCACCGGCATC
CTGTACAACGCCACCGAGCAGTCCTACAAGGACGGCCTGTCCAAGAACGACACC
CGCCTGTTTCGGCCAGAAGATCGGCTCCCTGTCCCCAACGCGATCCTGAACACGA
CCGTCCCCGCCACGGCATCGCGTTCTACCGCCTGCGCCCCCTCTCCTGA

SEQ ID NO: 112

5' Clp homologous recombination targeting sequence

GCGGCCGCGTGGACGAGGGCTACAACCCCGCCTACGGCGCGCGCCCCGCTGCGCC
GCGCCATCATGCGCCTGCTGGAGGACGCGCTGGCCGAGCGCATGCTCGCCGGCG
ACGTCAAGGAGGGCGACTCGGTCATCATGGACGTCGATGGCGATGGCGCCATCA
GCGTCCTCAACGGAGACCGCACCCACACCACCACCATCGACTCCTCCCCGGCGG
GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCCGGGGGCCTTGGGGCGCTGCTGGAGCGTGGTGAGAGCGCGGC
GGACGTGCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTTCG
GCCCCGATCCCCCACTGCCCCCATCATCTTGCCTGTTGTCGTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTC

SEQ ID NO: 113

3' Clp homologous recombination targeting sequence

GAGCTCAGAATAGTATCGGGTGATGCGAAGTCAGAACCAGGCAGGGCCTGTGCG
CTGAGGTGGCAACGATGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATC
CCGTGATCTCCCCACCGACACCTATCCCCGCCATCCCGGCCACCCCTTTCAGTC
CCCTCAGCATGCATTGTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGGTTCA
GGCCACGGCGCACCGAGTCCTCGCCCTTCGCAGAGTGATCACCTCCCCGGAACC
AGCCACGCTCGCTGCTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGG
ACAACTCAGGCAACCACGCGCCTCACAAGCATGGCCGCCGTGGCATCCAACCCG
CTCGTGACGGTGGGTGCGCAAGTGCCAGGGGCCTCGTCGTCACGGCGTGATCCT
CGAGGGATGCGATCCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGAT
GAGCCCCTGTCTAGA

SEQ ID NO: 114

Chlamydomonas reinhardtii TUB2 promoter/5' UTR

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAAC

SEQ ID NO: 115

Chlorella vulgaris nitrate reductase 3'UTR

GCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGA
CTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAA

ACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCT
TGTGCTATTTGCGAATACCACCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTG
CATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCT
CCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGT
ACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATG
GGAACACAAATGGAAAGCTT

SEQ ID NO: 116

A. niger AlgC amino acid sequence

MIGSSHAVVALGLFTLYGHSAAAPAIGASNSQTIVTNGTSFALNGDNVSYRFHVNSS
TGDLSIDHFGGVVSGTIPSPVEPAVNGWVGMPGRIRREFPDQGRGDFRIPAVRIRESA
GYTVSDLQYVSHEVIEGKYALPGLPATFGDAQDATTLLVVHLYDNYSSVAADLSYSIF
PKYDAIVRSVNVNTNQPGNITIEALASISIDFPYEDLDMVSLRGDWAREANVQRSKV
QYGVQFGSSTGYSSHLHNPFLAIVDPATTESQGEAWGFNLVYTGSFSAQVEKGSQG
FTRALLGFNPDQLSWNLGPGETLTSPECVAVYSDKGLGSVSRKFHRLYRNHLMKSK
FATSDRPVLLNSWEGVYFDYNQSSIETLAESAALGVHLFVMDDGWFGDKYPRVSD
NAGLGDWMPNPARFPDGLTPVVQDITNLTVNGTESTKLRFGIWVEPEMVNPNSTLY
HEHPEWALHAGPYPRTERRNQLVLNLALPAVQDFIIDFMTNLLQDTGISYVKWDNN
RGIHETPSPSTDHQMGLGLYRVFDLTLTRFPDVLWEGCASGGGRFDAGMLQYVPQI
WTSNDNTDAIDRITIQFGTSLAYPPSAMGAHLSAVPNAQTGRTPFTFRAHVAMMGGS
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

ATGATCGGCTCCTCCCACGCGGTCGTCGCCCTGGGCCTGTTACCCCTGTACGGCC
ACTCCGCCGCCGCGCCCGCCATCGGCGCCTCCAACCTCCCAGACCATCGTCACGAA
CGGCACCTCCTTCGCCCTGAACGGCGACAACGTGTCCTACCGCTTCCACGTGAAC
TCCTCCACGGGCGACCTGATCTCCGACCACTTCGGCGGCGTGGTGTCCGGCACCA
TCCCCTCCCCCGTGGAGCCCCGCGGTCAACGGCTGGGTGGGCATGCCCGGCCGCGAT
CCGCCGCGAGTTCCCCGACCAAGGCGCGCGGCGACTTCCGCATCCCCGCGGTGCGC
ATCCGCGAGTCCGCCGGCTACACCGTCTCCGACCTGCAGTACGTGTCCACGAGG
TGATCGAGGGCAAGTACGCGCTGCCCGGCCTGCCCGCCACGTTCGGCGACGCCC
AGGACGCCACCACCCTGGTGGTGCACCTGTACGACAACCTACTCCTCCGTCGCGGC
CGACCTGTCTACTCCATCTTCCCCAAGTACGACGCGATCGTCCGCTCCGTGAAC
GTGACCAACCAGGGCCCCGGCAACATCACCATCGAGGCGCTGGCCTCCATCTCC
ATCGACTTCCCCTACGAGGACCTGGACATGGTGTCCCTGCGCGGCGACTGGGCCC
GCGAGGCGAACGTGCAGCGCTCCAAGGTCCAGTACGGCGTGCAGGGCTTCGGCT
CCTCCACCGGCTACTCCTCCCACCTGCACAACCCCTTCCTGGCGATCGTCGACCC
CGCGACCACCGAGTCCCAGGGCGAGGCCTGGGGCTTCAACCTGGTCTACACCGG
CTCCTTCTCCGCCAGGTTCGAGAAGGGCTCCAGGGCTTCACGCGCGCCCTGCTG
GGCTTCAACCCCGACCAAGCTGTCTTGGAACCTGGGCCCCGGCGAGACGCTGACG

TCCCCCGAGTGCCTCGCCGTCTACTCCGACAAGGGCCTGGGCTCCGTCTCCCGCA
 AGTTCCACCGCCTGTACCGCAACCACCTGATGAAGTCCAAGTTCGCCACGTCCGA
 CCGCCCCGTGCTGCTGAACTCCTGGGAGGGCGTCTACTTCGACTACAACCAGTCC
 TCCATCGAGACGCTGGCGGAGGAGTCCGCCGCCCTGGGCGTGCACCTGTTTCGTCA
 TGGACGACGGCTGGTTTCGGCGACAAGTACCCCCGCGTGTCCGACAACGCCGGCC
 TGGGCGACTGGATGCCCAACCCCGCCCGCTTCCCCGACGGCCTGACGCCCGTCGT
 GCAGGACATCACCAACCTGACCGTCAACGGCACCGAGTCCACCAAGCTGCGCTT
 CGGCATCTGGGTGGAGCCCCGAGATGGTGAACCCCAACTCCACCCTGTACCACGA
 GCACCCCGAGTGGGCGCTGCACGCGGGCCCCCTACCCCGCACCGAGCGCCGCAA
 CCAGCTGGTCTGAACCTGGCCCTGCCCGCGGTCCAGGACTTCATCATCGACTTC
 ATGACCAACCTGCTGCAGGACACCGGCATCTCCTACGTCAAGTGGGACAACAAC
 CGCGGCATCCACGAGACGCCCTCCCCCTCCACGGACCACCAGTACATGCTGGGCC
 TGTACCGCGTGTTCGACACGCTGACCACGCGCTTCCCCGACGTCTGTGGGAGGG
 CTGCGCGTCCGGCGGGCGGCCGGTTCGACGCCGGCATGCTGCAGTACGTCCCCCAG
 ATCTGGACGTCCGACAACACGGACGCGATCGACCGCATCACCATCCAGTTCGGC
 ACCTCCCTGGCCTACCCCCCTCCGCCATGGGCGCCACCTGTCCGCGGTGCCCA
 ACGCCCAGACCGGCCGCACGGTGCCTTTCACCTTCCGCGCCCACGTGCGCATGAT
 GGGCGGCTCCTTCGGCCTGGAGCTGGACCCCGCCACCGTGGAGGGCGACGAGAT
 CGTCCCCGAGCTGCTGGCGCTGGCCGAGAAGGTGAACCCCATCATCCTGAACGG
 CGACCTGTACCGCCTGCGCCTGCCCCAGGACTCCAGTGGCCCCGCGGCCCTGTTT
 GTGTCCAGGACGGCGCCCCAGGCCGTCTGTTCTACTTCCAGGTGCAGCCCAACG
 TCAACCACGCCGTCCCCTGGGTCCGCCTGCAGGGCCTGGACCCCAAGGCCGACTA
 CACGGTGGACGGCGACCAGACGTACTCCGGCGCGACCCTGATGAACCTGGGCCT
 GCAGTACTCCTTCGACACCGAGTACGGCTCCAAGGTGGTGTTCCTGGAGCGCCAG
 TAA

SEQ ID NO: 119

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

GCGGCCGCGTGGACGAGGGCTACAACCCCGCCTACGGCGCGCGCCCCGCTGCGCC
 GCGCCATCATGCGCCTGCTGGAGGACGCGCTGGCCGAGCGCATGCTCGCCGGCG
 ACGTCAAGGAGGGCGACTCGGTCATCATGGACGTCGATGGCGATGGCGCCATCA
 GCGTCCTCAACGGAGACCGCACCCACACCACCACCATCGACTCCTCCCCGGCGG
 GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
 CGAGGACTGGGCCGGGGGCCCTTGGGGCGCTGCTGGAGCGTGGTGGAGAGCGCGG
 GGACGTGCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTGCG
 GCCCGCATCCCCCACTGCCCCCATCATCTTGCCCTGTTGTCTGTTGGCACTGACATA
 AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
 TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTCTTTCTTGGCGCTATGACAC
 TTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAAC
 ACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGAT
 GCCGCTCCAGGGCGAGCGCTGTTAAATAGCCAGGCCCGGATTGCAAAGACAT
 TATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACAC
 AGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCTCTTTCGT
 TTCAGTCACAACCCGCAAACACTAGTATGATCGGCTCCTCCCACGCGGTTCGTGCG
 CCTGGGCCTGTTACCCCTGTACGGCCACTCCGCCGCCGCGCCCGCCATCGGCGCC
 TCCAACCTCCAGACCATCGTCACGAACGGCACCTCCTTCGCCCTGAACGGCGACA
 ACGTGTCTTACCGCTTCCACGTGAACCTCCTCCACGGGCGACCTGATCTCCGACCA
 CTTCGGCGGGCGTGGTGTCCGGCACCATCCCCTCCCCCGTGGAGCCCGCGGTCAAC
 GGCTGGGTGGGCATGCCCGGCCGCATCCGCCGCGAGTTCCCCGACCAGGGCCGC

GGCGACTTCCGCATCCCCGCGGTGCGCATCCGCGAGTCCGCCGGCTACACCGTCT
CCGACCTGCAGTACGTGTCCACGAGGTGATCGAGGGCAAGTACGCGCTGCCCG
GCCTGCCCCGCCACGTTTCGGCGACGCCCAGGACGCCACCACCCTGGTGGTGCACCT
GTACGACAACACTACTCCTCCGTGCGGGCCGACCTGTCCTACTCCATCTTCCCCAAG
TACGACGCGATCGTCCGCTCCGTGAACGTGACCAACCAGGGCCCCGGAACATC
ACCATCGAGGCGCTGGCCTCCATCTCCATCGACTTCCCCTACGAGGACCTGGACA
TGGTGTCCCTGCGCGGGCGACTGGGCCCCGCGAGGCGAACGTGCAGCGCTCCAAGG
TCCAGTACGGCGTGCAGGGCTTTCGGCTCCTCCACCGGCTACTCCTCCACCTGCA
CAACCCCTTCCTGGCGATCGTCGACCCCGCGACCACCGAGTCCCAGGGCGAGGC
CTGGGGCTTCAACCTGGTCTACACCGGCTCCTTCTCCGCCCAGGTCGAGAAGGGC
TCCCAGGGCTTCACGCGCGCCCTGCTGGGCTTCAACCCCGACCAGCTGTCCTGGA
ACCTGGGGCCCCGGCGAGACGCTGACGTCCCCCGAGTGCGTCGCCGTCTACTCCGA
CAAGGGCCTGGGCTCCGTCTCCCGCAAGTTCCACCGCCTGTACCGCAACCACCTG
ATGAAGTCCAAGTTCGCCACGTCCGACCGCCCCGTGCTGCTGAACTCCTGGGAGG
GCGTCTACTTCGACTACAACCAGTCCCTCCATCGAGACGCTGGCGGAGGAGTCCGC
CGCCCTGGGCGTGCACCTGTTTCGTATGGACGACGGCTGGTTCGGCGACAAGTAC
CCCCGCGTGTCCGACAACGCCGGCCTGGGCGACTGGATGCCCAACCCCGCCCGCT
TCCCCGACGGCCTGACGCCCGTCGTGCAGGACATCACCACCTGACCGTCAACG
GCACCGAGTCCACCAAGCTGCGCTTCGGCATCTGGGTGGAGCCCGAGATGGTGA
ACCCCAACTCCACCCTGTACCACGAGCACCCCGAGTGGGCGCTGCACGCGGGGCC
CCTACCCCGCACCGAGCGCCGCAACCAGCTGGTCCTGAACCTGGCCCTGCCCGC
GGTCCAGGACTTCATCATCGACTTCATGACCAACCTGCTGCAGGACACCGGCATC
TCCTACGTCAAGTGGGACAACAACCGCGGCATCCACGAGACGCCCTCCCCCTCCA
CGGACCACCAGTACATGCTGGGCCTGTACCGCGTGTTCGACACGCTGACCACGCG
CTTCCCCGACGTCTGTGGGAGGGCTGCGCGTCCGGCGGGCGGCGGTTCGACGCC
GGCATGCTGCAGTACGTCCCCCAGATCTGGACGTCCGACAACACGGACGCGATC
GACCGCATCACCATCCAGTTCGGGCACCTCCCTGGCCTACCCCCCTCCGCCATGG
GCGCCACCTGTCCGCGGTGCCCAACGCCAGACCGGCCGACGCTGCCCTTCAC
CTTCCGCGCCACGTCGCGATGATGGGCGGCTCCTTCGGCCTGGAGCTGGACCCC
GCCACCGTGGAGGGCGACGAGATCGTGCCCGAGCTGCTGGCGCTGGCCGAGAAG
GTGAACCCCATCATCCTGAACGGCGACCTGTACCGCCTGCGCCTGCCCCAGGACT
CCCAGTGGCCCCGCGGCCCTGTTTCGTGTCCCAGGACGGCGGCCAGGCCGTCTGTT
CTACTTCCAGGTGCAGCCCAACGTCAACCACGCCGTCCCCTGGGTCCGCCTGCAG
GGCCTGGACCCCAAGGCCGACTACACGGTGGACGGCGACCAGACGTACTCCGGC
GCGACCCTGATGAACCTGGGCCTGCAGTACTCCTTCGACACCGAGTACGGCTCCA
AGGTGGTGTTCCTGGAGCGCCAGTAACAATTGAAGCAGCAGCAGCTCGGATAGT
ATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTG
CCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATC
TTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCAC
CCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCT
ACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCAC
AGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGC
ACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTT
GAGCTCAGAATAGTATCGGGTGATGCGAAGTCAGAACCAGGCAGGGCCTGTGCGC
CTGAGGTGGCAACGATGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATC
CCGTGATCTCCCCACCGACACCTATCCCCGCCATCCCGGCCACCCTTTCAGTC
CCCTCAGCATGCATTGTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGGTTCA
GGCCACGGCGCACCGAGTCTCGCCCTTCGCGAGAGTGATCACCTCCCCGGAACC
AGCCACGCTCGCTGCTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGG
ACAACTCAGGCAACCACGCGCCTCACAAGCATGGCCGCCGTGGCATCCAACCCG

CTCGTGACGGTGGGTGCGCAAGTGCCAGGGGCCTCGTCGTCACGGCGTGATCCT
CGAGGGATGCGATCCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGAT
GAGCCCCTGTCTAGA

SEQ ID NO: 120

C. tetragonobola α -galactosidase amino acid sequence

MATHYSIIIGGMIIVVLLMIIGSEGGRLLEKKNRTSAEAEHYNVRRYLAENGLGQTPPM
GWNSWNHFGCDINENVRETADAMVSTGLAALGYQYINLDDCWAELNRDSEGNM
VPNAAAFPSGIKALADYVHSKGLKLGVSADAGNQTC SKRMPGSLGHEEQDAKTFAS
WGV DYLKYDNCENLGISVKERYPPMGKALLSSGRPIFFSMCEWGWEDPQIWAKSIG
NSWRTTGDIEDNWSMTSIADSNDKWASYAGPGGWNDPDMLEVGNNGGMTTEEYR
SHFSIWALAKAPLLVGCDIRAMDDTTHELISNAEVIAVNQDKLGVQGGKVKSTNDLE
VWAGPLSDNKVAVILWNRSSSRATVTASWSDIGLQQGTTVDARDLWEHSTQSLVSG
EISAEIDSHACKMYVLTPRS

SEQ ID NO: 121

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

ATGGCCACCCACTACTCCATCATCGGCGGCATGATCATCGTCGTCCTGCTGATGA
TCATCGGCTCCGAGGGCGGCCGGCTGCTGGAGAAGAAGAACCGCACCTCCGCCG
AGGCGGAGCACTACAACGTGCGCCGCTACCTGGCCGAGAACGGCCTGGGCCAGA
CCCCCCCCATGGGCTGGAACCTCCTGGAACCACTTCGGCTGCGACATCAACGAGA
ACGTCGTCGCGAGACGGCGGACGCCATGGTGTCCACCGGCCTGGCCGCCCTGG
GCTACCAGTACATCAACCTGGACGACTGCTGGGCGGAGCTGAACCGCGACTCCG
AGGGCAACATGGTGCCCAACGCCGCCGCGTTCCCCTCCGGCATCAAGGCGCTGG
CGGACTACGTCCACTCCAAGGGCCTGAAGCTGGGCGTCTACTCCGACGCGGGCA
ACCAGACCTGCTCCAAGCGCATGCCCGGCTCCCTGGGCCACGAGGAGCAGGACG
CGAAGACGTTTCGCTCCTGGGGCGTCGACTACCTGAAGTACGACAACCTGCGAGA
ACCTGGGCATCTCCGTGAAGGAGCGCTACCCCCCATGGGCAAGGCGCTGCTGTC
CTCCGGCCGCCCATCTTCTTCTCCATGTGCGAGTGGGGCTGGGAGGACCCCCAG
ATCTGGGCCAAGTCCATCGGCAACTCCTGGCGCACCAACCGGCGACATCGAGGAC
AACTGGAACCTCATGACCTCCATCGCCGACTCCAACGACAAGTGGGCCTCCTACG
CCGGCCCCGGCGGCTGGAACGACCCCGACATGCTGGAGGTTCGGCAACGGCGGCA
TGACGACGGAGGAGTACCGCTCCCACTTCTCCATCTGGGCCCTGGCCAAGGCGCC
CCTGCTGGTGGGCTGCGACATCCGCGCGATGGACGACACGACGACGACGAGCTGAT
CTCCAACGCCGAGGTGATCGCGGTGAACCAGGACAAGCTGGGCGTGCAGGGCAA
GAAGGTCAAGTCCACGAACGACCTGGAGGTCTGGGCCGGCCCCCTGTCCGACAA
CAAGGTGGCGGTGATCCTGTGGAACCGCTCCTCCTCCCGCGCCACCGTCACCGCG
TCCTGGTCCGACATCGGCCTGCAGCAGGGCACCAACCGTCGACGCGCGCGACCTGT
GGGAGCACTCCACGCAGTCCCTGGTGTCCGGCGAGATCTCCGCCGAGATCGACTC
CCACGCCTGCAAGATGTACGTCCTGACGCCCCGCTCCTAA

SEQ ID NO: 122

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

GCGGCCGCGTGGACGAGGGCTACAACCCCGCCTACGGCGCGCGCCCGCTGCGCC
GCGCCATCATGCGCCTGCTGGAGGACGCGCTGGCCGAGCGCATGCTCGCCGGCG
ACGTCAAGGAGGGCGACTCGGTCATCATGGACGTCGATGGCGATGGCGCCATCA
GCGTCCTCAACGGAGACCGCACCCACACCACCACCATCGACTCCTCCCCGGCGG

GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCGGGGGGCCTTGGGGCGCTGCTGGAGCGTGGTGAGAGCGCGGC
GGACGTGCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTGCG
GCCCCGATCCCCCACTGCCCCCATCATCTTGCCTGTTGTCTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTCCTTTCTTGGCTATGACAC
TTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAAC
ACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGAT
GCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACAT
TATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACAC
AGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGT
TTCAGTCACAACCCGCAAACACTAGTATGGCCACCCACTACTCCATCATCGGCGG
CATGATCATCGTCGTCCTGCTGATGATCATCGGCTCCGAGGGCGGGCCGGCTGCTG
GAGAAGAAGAACCGCACCTCCGCCGAGGCGGAGCACTACAACGTGCGCCGCTAC
CTGGCCGAGAACGGCCTGGGCCAGACCCCCCATGGGCTGGAACCTCTGGAAC
CACTTCGGCTGCGACATCAACGAGAACGTCTGTCGCGAGACGGCGGACGCCATG
GTGTCCACCGGCCTGGCCGCCCTGGGCTACCAGTACATCAACCTGGACGACTGCT
GGGCGGAGCTGAACCGCGACTCCGAGGGCAACATGGTGCCCAACGCCGCCGCGT
TCCCTCCGGCATCAAGGCGCTGGCGGACTACGTCCACTCCAAGGGCCTGAAGCT
GGGCGTCTACTCCGACGCGGGCAACCAGACCTGCTCCAAGCGCATGCCCGGCTC
CCTGGGCCACGAGGAGCAGGACGCGAAGACGTTTCGCCTCCTGGGGCGTCGACTA
CCTGAAGTACGACAACTGCGAGAACCTGGGCATCTCCGTGAAGGAGCGCTACCC
CCCCATGGGCAAGGCGCTGCTGTCTCCGGCCGCCCATCTTCTTCTCCATGTGC
GAGTGGGGCTGGGAGGACCCCAAGTCTGGGCCAAGTCCATCGGCAACTCCTGG
CGCACCACCGGCGACATCGAGGACAACCTGGAACCTCCATGACCTCCATCGCCGAC
TCCAACGACAAGTGGGCCTCCTACGCCGGCCCCGGCGGGCTGGAACGACCCCGAC
ATGCTGGAGGTCGGCAACGGCGGCATGACGACGGAGGAGTACCGCTCCCACTTC
TCCATCTGGGCCCTGGCCAAGGCGCCCCCTGCTGGTGGGCTGCGACATCCGCGCGA
TGGACGACACGACGACGAGCTGATCTCCAACGCCGAGGTGATCGCGGTGAACC
AGGACAAGCTGGGCGTGCAGGGCAAGAAGGTCAAGTCCACGAACGACCTGGAG
GTCTGGGCCGGCCCCCTGTCCGACAACAAGGTGGCGGTGATCCTGTGGAACCGCT
CCTCCTCCCGCGCCACCGTCACCGCGTCTTGGTCCGACATCGGCCTGCAGCAGGG
CACCACCGTCGACGCGCGCGACCTGTGGGAGCACTCCACGCAAGTCCCTGGTGTCC
GGCGAGATCTCCGCCGAGATCGACTCCCACGCCTGCAAGATGTACGTCCTGACGC
CCCGCTCCTAACAATTGAAGCAGCAGCAGCTCGGATAGTATCGACACACTCTGG
ACGCTGGTTCGTGTGATGGACTGTTGCCGCCCACTTGCTGCCTTGACCTGTGAAT
ATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCT
TTTGCAGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTT
CCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGTGCTAT
CCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGC
TCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATG
CACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTCAGAATAGTA
TCGGGTGATGCGAAGTCAGAACCAGGCAGGGCCTGTCGCTGAGGTGGCAACGA
TGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATCCCGTGATCTCCCCCAC
CGACACCTATCCCCGCCCATCCCCGGCCACCCCTTTCAGTCCCCTCAGCATGCATT
GTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGAGGTCAGGCCACGGCGCACC
GAGTCCTCGCCCTTCGCAGAGTGATACCCCTCCCCGGAACCAGCCACGCTCGCTG
CTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGGACAACCTCAGGCAAC
CACGCGCCTCACAAGCATGGCCGCCGTGGCATCCAACCCGCTCGTGACGGTGGG
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
RLQQTDAFAELKALSSRQSVNRPQKGELSFRQSPTFQDCFPGSEKCYREVEHDGKTL
KVPFRRVHLQDDNGHFDLYDTS GPQGVNPREGLPKIRSSWVEPREARGDKVQTQQY
YAKQGIITEEMAFCAARERMDPEFIRSEVARGRAIIPANKRHLELEPTVVGRNFLVKV
NANIGNSAVSSSIEEEVEKLQWSTIWGADTAMDLS TGNNIHETREWVMRNSPVPVGT
VPIYQCLEKAGGIVENITWELFRETLEQAEQGVDFTHAGVLLRYIPLTANRVTGIV
SRGGSIAKLCLLDHTENFAYMHWDEILDICAQYDITLSIGDGLRPGCIADANDAAQF
AELKTQGELTRRAWAKDVQVMNEGPVPLHKIPENMEKQLDWCSEAPFYTLGPL
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

GCTCTTCGCCCGCCGCACTCCTGCTCGAGCGCGCCCGCGCGTGCGCCGCCAGCGC
CTTGGCCTTTTCGCCGCGCTCGTGCGCGTGCTGATGTCCATCACCAGGTCCATG
AGGTCTGCCTTGCGCCGGCTGAGCCACTGCTTCGTCCGGGCGGCCAAGAGGAGC
ATGAGGGAGGACTCCTGGTCCAGGGTCCTGACGTGGTCGCGGCTCTGGGAGCGG
GCCAGCATCATCTGGCTCTGCCGCACCGAGGCCGCCTCCAACCTGGTCCTCCAGCA
GCCGCAGTCGCCGCCGACCCTGGCAGAGGAAGACAGGTGAGGGGGGTATGAATT
GTACAGAACAAACCACGAGCCTTGCTAGGCAGAAATCCCTACCAGTCATGGCTTTA
CCTGGATGACGGCCTGCGAACAGCTGTCCAGCGACCCCTCGCTGCCGCCGCTTCTC
CCGCACGCTTCTTTCCAGCACCGTGATGGCGCGAGCCAGCGCCGCACGCTGGCGC
TGCGCTTCGCCGATCTGAGGACAGTCGGGGAACTCTGATCAGTCTAAACCCCTT
GCGCGTTAGTGTTGCCATCCTTTGCAGACCGGTGAGAGCCGACTTGTTGTGCGCC
ACCCCCACACCACCTCCTCCCAGACCAATTCTGTACCTTTTTGGCGAAGGCAT
CGGCCTCGGCCTGCAGAGAGGACAGCAGTGCCAGCCGCTGGGGGTGGCGGAT

GCACGCTCAGGTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGG
 GCTGCGAGACGGCTTCCCGGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCC
 GAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGT
 TAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCAT
 ATTCAAACACCTAGATCACTACCCTTCTACACAGGCCACTCGAGCTTGTGATCG
 CACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGG
 CGCGCCATGCTGCTGCAGGCCTTCCTGTTTCCTGCTGGCCGGCTTCGCCGCCAAGA
 TCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCCTTCACCCC
 CAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGC
 CAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCC
 CTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCC
 CATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTG
 GTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGC
 GCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCT
 CCTACAGCCTGGACGGCGGCTACACCTTCAACGAGTACCAGAAGAACCCCGTGC
 TGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTC
 CCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTA
 CTCCTCCGACGACCTGAAGTCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGG
 CTTCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAG
 GACCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGG
 CCGGCGGCTCCTTCAACCAGTACTTCGTCGGCAGCTTCAACGGCACCCACTTCGA
 GGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTG
 CAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGG
 GCCTCCAACCTGGGAGTACTCCGCCTTCGTGCCCACCAACCCCTGGCGCTCCTCCA
 TGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGAC
 GGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCC
 CTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGT
 CGACCTGTCCAACAGCACCCGGCACCCCTGGAGTTCGAGCTGGTGTACGCCGTCAAC
 ACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGG
 GCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCT
 CTTCTTCCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTA
 CTTACCAACCCGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGAGAACGACCT
 GTCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTC
 AACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGAACGCC
 CTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAG
 TTCCAGGTGCGCGAGGTCAAGTGACAATTGGCAGCAGCAGCTCGGATAGTATCG
 ACACACTCTGGACGCTGGTTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTT
 GACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGT
 GTGTACGCGCTTTTGGCAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCC
 AGCATCCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGC
 TGTCTTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCC
 TTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGC
 AATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAGGATCCCGCGT
 CTCGAACAGAGCGCGCAGAGGAACGCTGAAGGTCTCGCCTCTGTTCGCACCTCAG
 CGCGGCATACACCACAATAACCACCTGACGAATGCGCTTGGTTCTTCGTCCATTA
 GCGAAGCGTCCGGTTCACACACGTGCCACGTTGGCGAGGTGGCAGGTGACAATG
 ATCGGTGGAGCTGATGGTCGAAACGTTACAGCCTAGGGATATCGAATTCCTTTC
 TTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCG
 GCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGC
 ATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCC

GATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCA
 CTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCG
 CCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACACTAGTATGACCACCAACCT
 GGCTAAGCTGCCTCTCGGCGGCCTCTCTTCTCGGTCGTCCATCGCCGGGGCTCCTC
 TGCGCGTGAACAGCCACAGCCGCGAGCGCCACCTGGGCGCCAAGACCACCGCGA
 TCGCCGCTCCCGAGCGGCTGGACTACCTCGACAACGCGGAGGAGGCGCGGCTGC
 AGCAGACGGACGCCTTCGCCGAGCTGAAGGCGCTGTCTGAGCCGCCAGTCCGTGA
 ACCGGCCCCAAAAGGGCGAGCTGTCTTCCGGCAGTCGCCGACCTTTCAGGACT
 GCTTCCCTGGCTCGGAGAAGTGCTACCGCGAGGTGGAGCACGATGGCAAGACGC
 TCAAGGTGCCTTTTCGCCGGGTGCACCTGCAGGACGACAACGGCCATTTTCGACCT
 GTACGACACCTCTGGCCCCCAGGGCGTGAACCCCCGCGAGGGTCTGCCCAAGAT
 CCGCTCCTCCTGGGTGGAGCCGCGCGAGGCGCGGGGGCGACAAGGTGCAAACGCA
 GCAGTACTACGCCAAGCAAGGCATCATCACCGAGGAGATGGCCTTCTGCGCTGC
 GCGCGAGCGCATGGACCCCGAGTTCATCCGGTCTGAGGTGGCCCGGGGTCTGGGC
 CATCATCCCTGCGAACAAGCGCCATCTGGAGCTGGAGCCGACCGTCTGTTGGGCCG
 CAACTTTCTGGTCAAGGTGAACGCCAACATCGGCAACTCTGCGGTGAGCTCGAGC
 ATCGAGGAGGAGGTGGAGAAGCTGCAGTGGAGCACCATCTGGGGTGCCGACACG
 GCCATGGACCTGAGCACCGGCAACAACATCCACGAGACGCGCGAGTGGGTGATG
 CGCAACAGCCCGGTCCCTGTGGGTACGGTCCCTATCTACCAGTGCCTGGAGAAGG
 CGGGCGGCATCGTGGAGAACATCACCTGGGAGCTGTTCCGGGAGACGCTGATCG
 AGCAGGCCGAGCAGGGCGTCGACTACTTTACCATCCACGCGGGGGTCCTGCTCC
 GCTACATCCCCCTGACGGCGAACC CGCTCACCGGCATCGTCAGCCGCGGCGGCTC
 TATCCATGCCAAGCTCTGCCTGCTGGACCACACCGAGAACTTCGCCTACATGCAC
 TGGGACGAGATCCTGGACATCTGCGCCCAGTACGACATCACCTGTCTGATCGGCG
 ACGGCCTGCGCCCCCGGTGCATCGCGGACGCCAACGACGCTGCCCAGTTTGTCTGA
 GCTGAAGACCCAGGGCGAGCTGACCCGCCGCGCCTGGGGCGAAGGACGTGCAGGT
 GATGAACGAGGGTCCCGGCCATGTGCCGCTGCACAAGATCCCCGAGAACATGGA
 GAAGCAGCTGGACTGGTGCAGCGAGGCGCCCTTCTACACGCTCGGCCCCCTGGC
 GACGGACATCGCGCCTGCGTACGACCACATCACCTCCGCGATCGGCGCCGCCAC
 GATCGGCGCGCTGGGGACCGCGCTGCTGTGTTACGTGACCCCGAAGGAGCACCT
 GGGCCTCCCCGATCGCGACGACGTGAAGGCCGGGGTGATCGCCTACAAGATCGC
 GCGCACGCGGCTGACCTGGCGAAGGGTACCCCTACGCTCAGGAGTGGGATAA
 CGCCCTCTCTAAGGCGCGCTTCGAGTTCCGGTGGTACGACCAGTTCCACCTGTCTG
 CTGGACCCCGTCACCGCCCGCCTGTTCCATGACGCGACCCTGCCGCAGGAGCCCG
 CCAAGACCGCCCACTTCTGCTCCATGTGCGGCCCAAGTTCTGCTCCATGCAGAT
 CACCCAGGACCTCCGCGAGTACGCCAAGAACCATCAAATGGAGGAGGACGAGGC
 GATCCAAACCGGCATGGCCGAGATGTCCGAGCAGTTCAAGAGCATCGGCGCGGA
 GGTCTACCTCGACGAGGCGGTGCTGGAGCACGCGTAAGCAGCAGCAGCTCGGAT
 AGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTG
 CTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTG
 ATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATAC
 CACCCCCAGCATCCCCTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTA
 TCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCG
 CACAGCCTTGTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAAACC
 AGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAA
 GCTTGAGCTCTTGTTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCT
 CGATAACCTCCAAAGCCGCTCTAATTGTGGAGGGGGTTCGAATTTAAAAGCTTGG
 AATGTTGGTTCTGTGCGTCTGGAACAAGCCCAGACTTGTGCTCACTGGGAAAAGG
 ACCATCAGCTCCAAAAAACTTGCCGCTCAAACCGCGTACCTCTGCTTTCGCGCAA
 TCTGCCCTGTTGAAATCGCCACCACATTCATATTGTGACGCTTGAGCAGTCTGTA

ATTGCCTCAGAATGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGT
CGCGGGCGAGGACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAAT
AGTTCATAACAGTGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTC
TGAGTGGCCACCCCCCGGCCCTGGTGTCTGCGGAGGGCAGGTCAACCGGCATGG
GGCTACCGAAATCCCCGACCGGATCCCACCACCCCCGCGATGGGAAGAATCTCT
CCCCGGGATGTGGGCCCACCACCAGCACAACCTGCTGGCCCAGGCGAGCGTCAA
ACCATAACCACACAAATATCCTTGGCATCGGCCCTGAATTCCTTCTGCCGCTCTGCT
ACCCGGTGCTTCTGTCCGAAGCAGGGGTGCTAGGGATCGCTCCGAGTCCGCAAA
CCCTTGTCGCGTGGCGGGGCTTGTTGAGCTGAAGAGC

SEQ ID NO: 128

Codon-optimized *Coccomyxa* C-169 THIC sequence

ATGACCACCAACCTGGCTAAGCTGCCTCTCGGCGGCCCTCTCTTCTCGGTTCGTCCA
TCGCCGGGGCTCCTCTGCGCGTGAACAGCCACAGCCGCGAGCGCCACCTGGGCG
CCAAGACCACCGCGATCGCCGCTCCCGAGCGGCTGGACTACCTCGACAACGCGG
AGGAGGCGCGGCTGCAGCAGACGGACGCCTTCGCCGAGCTGAAGGCGCTGTCGA
GCCGCCAGTCCGTGAACCGGCCCAAAAGGGCGAGCTGTCGTTCCGGCAGTCGC
CGACCTTTCAGGACTGCTTCCCTGGCTCGGAGAAGTGCTACCGCGAGGTGGAGCA
CGATGGCAAGACGCTCAAGGTGCCTTTTCGCCGGGTGCACCTGCAGGACGACAA
CGGCCATTTTCGACCTGTACGACACCTCTGGCCCCCAGGGCGTGAACCCCCGCGAG
GGTCTGCCCAAGATCCGCTCCTCCTGGGTGGAGCCGCGCGAGGCGCGGGGCGAC
AAGGTGCAAACGCAGCAGTACTACGCCAAGCAAGGCATCATCACCGAGGAGATG
GCCTTCTGCGCTGCGCGCGAGCGCATGGACCCCGAGTTCATCCGGTCTGAGGTGG
CCCGGGGTTCGGGCCATCATCCCTGCGAACAAGCGCCATCTGGAGCTGGAGCCGA
CCGTCGTGGGCCGCAACTTTCTGGTCAAGGTGAACGCCAACATCGGCAACTCTGC
GGTCAGCTCGAGCATCGAGGAGGAGGTGGAGAAGCTGCAGTGGAGCACCATCTG
GGGTGCCGACACGGCCATGGACCTGAGCACCGGCAACAACATCCACGAGACGCG
CGAGTGGGTGATGCGCAACAGCCCGGTCCCTGTGGGTACGGTCCCTATCTACCAG
TGCCTGGAGAAGGCGGGCGGCATCGTGGAGAACATCACCTGGGAGCTGTTCCGG
GAGACGCTGATCGAGCAGGCCGAGCAGGGCGTTCGACTACTTTACCATCCACGCG
GGGGTCCCTGCTCCGCTACATCCCCCTGACGGCGAACC GCGTCACCGGCATCGTCA
GCCGCGGCGGCTCTATCCATGCCAAGCTCTGCCTGCTGGACCACACCGAGAACTT
CGCCTACATGCACTGGGACGAGATCCTGGACATCTGCGCCCAGTACGACATCACC
CTGTTCGATCGGCGACGGCCTGCGCCCCGGTGCATCGCGGACGCCAACGACGCT
GCCAGTTTGCTGAGCTGAAGACCCAGGGCGAGCTGACCCGCCGCGCCTGGGCG
AAGGACGTGCAGGTGATGAACGAGGGTCCCGGCCATGTGCCGCTGCACAAGATC
CCCGAGAACATGGAGAAGCAGCTGGACTGGTGCAGCGAGGCGCCCTTCTACACG
CTCGGCCCCCTGGCGACGACATCGCGCCTGCGTACGACCACATCACCTCCGCGA
TCGGCGCCGCCACGATCGGCGCGCTGGGGACCGCGCTGCTGTGTTACGTGACCCC
GAAGGAGCACCTGGGCCTCCCCGATCGCGACGACGTGAAGGCCGGGGTGATCGC
CTACAAGATCGCGGCGCACGCGGCTGACCTGGCGAAGGGTCACCCCTACGCTCA
GGAGTGGGATAACGCCCTCTCTAAGGCGCGCTTCGAGTTCCGGTGGTACGACCA
GTTCCACCTGTGCTGGACCCCGTCACCGCCCGCCTGTTCCATGACGCGACCCTG
CCGCAGGAGCCCGCCAAGACCGCCCACTTCTGCTCCATGTGCGGCCCAAGTTCT
GCTCCATGCAGATCACCCAGGACCTCCGCGAGTACGCCAAGAACCATCAAATGG
AGGAGGACGAGGCGATCCAAACCGGCATGGCCGAGATGTCCGAGCAGTTCAAGA
GCATCGGCGCGGAGGTCTACCTCGACGAGGCGGTGCTGGAGCACGCGTAA

SEQ ID NO: 129

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

GGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAA
GATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCAC TTCACC
CCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGAC
GCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACG
CCCTTGTTCTGGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAG
CCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGG
TGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCA
GCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACAT
CTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCGT
GCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCC
TCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATC
TACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTTCGCCAACGAGG
GCTTCCTCGGCTACCACTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCA
GGACCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCG
GCCGGCGGCTCCTTCAACCAGTACTTCGTCGGCAGCTTCAACGGCACCCACTTCG
AGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCT
GCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTG
GGCCTCCAACCTGGGAGTACTCCGCTTTCGTGCCACCAACCCCTGGCGCTCCTCC
ATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGA
CGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCC
CCTGGAGCCGGTTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACG
TCGACCTGTCCAACAGCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTCAA
CACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAG
GGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCT
CCTTCTTCTGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCT
ACTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGACC
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
GCGTCCTCAACGGAGACCGCACCCACACCACCACCATCGACTCCTCCCCGGCGG
GCATCTCGTAGACGCGAGAGGGAGGGAGGGAGGATTTTCAGAGGGAGATGAGA
CGAGGACTGGGCCGGGGGCCCTTGGGGCGCTGCTGGAGCGTGGTGAAGCGCGGC
GGACGTGCCTTTTCTTCTTCCGTGCGCGCGCTCTTGGCCATTGATCCCCGATTTCG
GCCCGCATCCCCCACTGCCCCCATCATCTTGCTGTTGTCGTGGCACTGACATA
AACCCCTGCGCTGCGCTGCTCCGCTACTATTGATATAGGTCTCACGCGCCAATC
TTTTTTGCTCCGGGTAACCGTCTGGACGCCAGAATTCCTTTCTTTCGCTATGACAC
TTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAAC
ACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGAT
GCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCGCCGATTGCAAAGACAT

TATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACAC
AGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCTCTTTCGT
TTCAGTCACAACCCGCAAACACTAGTATGGCCACCGCATCCACTTTCTCGGCGTT
CAATGCCCCGCTGCGGGGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCC
AGCGAGGCCCCCTCCCCGTGCGCGGGGCGCGCCGCGATCGCCGCTCCCGAGCGGCT
GGACTACCTCGACAACGCGGAGGAGGCGCGGCTGCAGCAGACGGACGCCTTCGC
CGAGCTGAAGGCGCTGTTCGAGCCGCCAGTCCGTGAACCGGCCCAAAAGGGCGA
GCTGTTCGTTCCGGCAGTCGCCGACCTTTTCAAGGACTGCTTCCCTGGCTCGGAGAAG
TGCTACCGCGAGGTGGAGCACGATGGCAAGACGCTCAAGGTGCCTTTTCGCCGG
GTGCACCTGCAGGACGACAACGGCCATTTTCGACCTGTACGACACCTCTGGCCCCC
AGGGCGTGAACCCCCGCGAGGGTCTGCCCAAGATCCGCTCCTCCTGGGTGGAGC
CGCGCGAGGCGCGGGGCGACAAGGTGCAAACGCAGCAGTACTACGCCAAGCAA
GGCATCATCACCGAGGAGATGGCCTTCTGCGCTGCGCGCGAGCGCATGGACCCC
GAGTTCATCCGGTCTGAGGTGGCCCCGGGGTTCGGGCCATCATCCCTGCGAACAAG
CGCCATCTGGAGCTGGAGCCGACCGTCGTGGGCCGCAACTTTCTGGTCAAGGTGA
ACGCCAACATCGGCAACTCTGCGGTCAGCTCGAGCATCGAGGAGGAGGTGGAGA
AGCTGCAGTGGAGCACCATCTGGGGTGCCGACACGGCCATGGACCTGAGCACCG
GCAACAACATCCACGAGACGCGCGAGTGGGTGATGCGCAACAGCCCGGTCCCTG
TGGGTACGGTCCCTATCTACCAGTGCCTGGAGAAGGCGGGCGGCATCGTGGAGA
ACATCACCTGGGAGCTGTTCCGGGAGACGCTGATCGAGCAGGCCGAGCAGGGCG
TCGACTACTTTACCATCCACGCGGGGGTCTGCTCCGCTACATCCCCCTGACGGC
GAACCGCGTCACCGGCATCGTCAGCCGCGGCGGCTCTATCCATGCCAAGCTCTGC
CTGCTGGACCACACCGAGAACTTCGCCTACATGCACTGGGACGAGATCCTGGAC
ATCTGCGCCCAGTACGACATCACCTGTTCGATCGGCGACGGCCTGCGCCCCGGCT
GCATCGCGGACGCCAACGACGCTGCCCAGTTTGTGAGCTGAAGACCCAGGGCG
AGCTGACCCGCGCGCCTGGGCGAAGGACGTGCAGGTGATGAACGAGGGTCCCG
GCCATGTGCCGCTGCACAAGATCCCCGAGAACATGGAGAAGCAGCTGGACTGGT
GCAGCGAGGCGCCCTTCTACACGCTCGGCCCCCTGGCGACGGACATCGCGCCTGC
GTACGACCACATCACCTCCGCGATCGGCGCCGCCACGATCGGCGCGCTGGGGAC
CGCGCTGCTGTGTTACGTGACCCCGAAGGAGCACCTGGGCCTCCCCGATCGCGAC
GACGTGAAGGCCGGGGTGTTCGCCTACAAGATCGCGGCGCACGCGGCTGACCTG
GCGAAGGGTACCCCTACGCTCAGGAGTGGGATAACGCCCTCTCTAAGGCGCGC
TTCGAGTTCCGGTGGTACGACCAGTTCACCTGTTCGCTGGACCCCGTCACCGCCC
GCCTGTTCCATGACGCGACCCTGCCGCAGGAGCCCGCCAAGACCGCCCACTTCTG
CTCCATGTGCGGCCCCAAGTTCTGCTCCATGCAGATCACCCAGGACCTCCGCGAG
TACGCCAAGAACCATCAAATGGAGGAGGACGAGGCGATCCAAACCGGCATGGCC
GAGATGTCCGAGCAGTTCAAGAGCATCGGCGCGGAGGTCTACCTCGACGAGGCG
GTGCTGGAGCACGCGTAACAATTGAAGCAGCAGCAGCTCGGATAGTATCGACAC
ACTCTGGACGCTGGTTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACC
TGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGT
ACGCGCTTTTTCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGC
ATCCCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGT
CCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTG
GTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAA
TGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTCA
GAATAGTATCGGGTGTGCGAAGTCAGAACCGAGGAGGCGCTGTCGCCTGAGGT
GGCAACGATGGGAAGCAATCAATCTGGGTACAGTCGTCCGCACGATCCCGTGAT
CTCCCCACCGACACCTATCCCCGCCCATCCCGGCCACCCTTTCAGTCCCCCTCAG
CATGCATTGTGCACCGCGACAAAGCATGTCTGCTCGTGCCTGGTTTCAGGCCACG
GCGCACCGAGTCCCTCGCCCTTCGCAGAGTGATCACCCCTCCCCGGAACCAGCCACG

CTCGCTGCTGCGGGCCGATCAGCCGCGCGCACTCCCTGCAACTAGGGACAACCTCA
GGCAACCACGCGCCTCACAAGCATGGCCGCCGTGGCATCCAACCCGCTCGTGAC
GGTGGGTGCGCAAGTGCCAGGGGCTCGTCGTACGGCGTGCATCCTCGAGGGA
TGCGATCCGGCAACTATATGTCGTTTATCTCCCCACCAATCACAGGATGAGCCCC
TGTCTAGA

SEQ ID NO: 131

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

ACTAGTATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGCGACC
TGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCG
CGGGCGCGCCGCGATCGCCGCTCCCGAGCGGCTGGACTACCTCGACAACGCGGA
GGAGGCGCGGCTGCAGCAGACGGACGCCTTCGCCGAGCTGAAGGCGCTGTGAG
CCGCCAGTCCGTGAACCGGCCCAAAAGGGCGAGCTGTGTTCCGGCAGTCGCC
GACCTTTCAGGACTGCTTCCCTGGCTCGGAGAAGTGCTACCGCGAGGTGGAGCAC
GATGGCAAGACGCTCAAGGTGCCTTTTCGCCGGGTGCACCTGCAGGACGACAAC
GGCCATTTTCGACCTGTACGACACCTCTGGCCCCCAGGGCGTGAACCCCCGCGAGG
GTCTGCCCAAGATCCGCTCCTCCTGGGTGGAGCCGCGCGAGGCGCGGGGCGACA
AGGTGCAAACGCAGCAGTACTACGCCAAGCAAGGCATCATCACCGAGGAGATGG
CCTTCTGCGCTGCGCGCGAGCGCATGGACCCCGAGTTTCATCCGGTCTGAGGTGGC
CCGGGGTCGGGCCATCATCCCTGCGAACAAGCGCCATCTGGAGCTGGAGCCGAC
CGTCGTGGGCCGCAACTTTCTGGTCAAGGTGAACGCCAACATCGGCAACTCTGCG
GTCAGCTCGAGCATCGAGGAGGAGGTGGAGAAGCTGCAGTGGAGCACCATCTGG
GGTGCCGACACGGCCATGGACCTGAGCACCGGCAACAACATCCACGAGACGCGC
GAGTGGGTGATGCGCAACAGCCCGGTCCCTGTGGGTACGGTCCCTATCTACCAGT
GCCTGGAGAAGGCGGGCGGCATCGTGGAGAACATCACCTGGGAGCTGTTCCGGG
AGACGCTGATCGAGCAGGCCGAGCAGGGCGTCGACTACTTTACCATCCACGCGG
GGGTCTGCTCCGCTACATCCCCCTGACGGCGAACC CGCTCACCGGCATCGTCAG
CCGCGGCGGCTCTATCCATGCCAAGCTCTGCCTGCTGGACCACACCGAGAACTTC
GCCTACATGCACTGGGACGAGATCCTGGACATCTGCGCCCAGTACGACATCACCC
TGTCGATCGGCGACGGCCTGCGCCCCGGCTGCATCGCGGACGCCAACGACGCTG
CCCAGTTTGCTGAGCTGAAGACCCAGGGCGAGCTGACCCGCCGCGCCTGGGCGA
AGGACGTGCAGGTGATGAACGAGGGTCCCGGCCATGTGCCGCTGCACAAGATCC
CCGAGAACATGGAGAAGCAGCTGGACTGGTGCAGCGAGGCGCCCTTCTACACGC
TCGGCCCCCTGGCGACGGACATCGCGCCTGCGTACGACCACATCACCTCCGCGAT
CGGCGCCGCCACGATCGGCGCGCTGGGGACCGCGCTGCTGTGTTACGTGACCCC
GAAGGAGCACCTGGGCCTCCCCGATCGCGACGACGTGAAGGCCGGGGTGATCGC
CTACAAGATCGCGGCGCACGCGGCTGACCTGGCGAAGGGTCACCCCTACGCTCA
GGAGTGGGATAACGCCCTCTCTAAGGCGCGCTTCGAGTTCCGGTGGTACGACCA
GTTCCACCTGTCGCTGGACCCCGTCACCGCCCGCCTGTTCCATGACGCGACCCTG
CCGCAGGAGCCCGCCAAGACCGCCCACTTCTGCTCCATGTGCGGCCCAAGTTCT
GCTCCATGCAGATCACCCAGGACCTCCGCGAGTACGCCAAGAACCATCAAATGG
AGGAGGACGAGGCGATCCAAACCGGCATGGCCGAGATGTCCGAGCAGTTCAAGA
GCATCGGCGCGGAGGTCTACCTCGACGAGGCGGTGCTGGAGCACGCGTAACAAT
TG

SEQ ID NO: 132

Chlorella protothecoides actin promoter/5'UTR

ACTAGAGAGTTTAGGTCCAGCGTCCGTGGGGGGGGACGGGCTGGGAGCTTGGGC
 CGGGAAGGGCAAGACGATGCAGTCCCTCTGGGGAGTCACAGCCGACTGTGTGTG
 TTGCACTGTGCGGCCCCGAGCACTCACACGCAAAATGCCTGGCCGACAGGCAGG
 CCCTGTCCAGTGCAACATCCACGGTCCCTCTCATCAGGCTCACCTTGCTCATTGAC
 ATAACGGAATGCGTACCGCTCTTTTCAGATCTGTCCATCCAGAGAGGGGAGCAGG
 CTCCCCACCGACGCTGTCAAACCTTGCTTCCTGCCCCAACCGAAAAACATTATTGTTT
 GAGGGGGGGGGGGGGGGGGGGCAGATTGCATGGCGGGATATCTCGTGAGGAACAT
 CACTGGGACACTGTGGAACACAGTGAGTGCAGTATGCAGAGCATGTATGCTAGG
 GGTCAGCGCAGGAAGGGGGCCCTTTCCAGTCTCCCATGCCACTGCACCGTATCCA
 CGACTCACAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGTGGACACCAGTGTG
 TAGCCTCTGGACTCCAGGTATGCGTGCACCGCAAAGGCCAGCCGATCGTGCCGAT
 TCCTGGGGTGGAGGATATGAGTGCAGCAACTTGGGGCTCAGAGTGCACACTGGG
 GCACGATACGAAACAACATCTACACCGTGTCCCTCCATGCTGACACACCACAGCTT
 CGCTCCACCTGAATGTGGGCGCATGGGCCCCGAATCACAGCCAATGTGCTGCTGC
 CATAATGTGATCCAGACCCTCTCCGCCAGATGCCGAGCGGATCGTGGGCGCTGA
 ATAGATTCTGTTTCGATCACTGTTTGGGTCCCTTTCCCTTTTCGTCTCGGATGCGCG
 TCTCGAAACAGGCTGCGTCGGGCTTTTCGGATCCCTTTTGCTCCCTCCGTCACCATC
 CTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACCAGGGTTGGAGGGTATTA
 CCGCGTCAGGCCATTCCCAGCCCGGATTCAATTCAAAGTCTGGGCCACCACCCTC
 CGCCGCTCTGTCTGATCACTCCACATTCGTGCATACACTACGTTCAAGTCCTGATC
 CAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAATCTCAAGGACCCACTCCA
 GCACAGCTGCTGGTTGACCCCGCCCTCGCAA

SEQ ID NO: 133

Chlorella protothecoides EF1A 3'UTR

ACGGAGCGTCGTGCGGGAGGGAGTGTGCCGAGCGGGGAGTCCCGGTCTGTGCGA
 GGCCCCGGCAGCTGACGCTGGCGAGCCGTACGCCCCGAGGGTCCCCCTCCCCCTGC
 ACCCTCTTCCCCTTCCCTCTGACGGCCGCGCCTGTTCTTGCATGTTTCAGCGAC

SEQ ID NO: 134

Relevant sequence of *A. thaliana* THIC expression cassette

GCTCTTCGCCGCCGCCACTCCTGCTCGAGCGCGCCCGCGCGTGCGCCGCCAGCGC
 CTTGGCCTTTTCGCCGCGCTCGTGCGCGTGCCTGATGTCCATCACCAGGTCCATG
 AGGTCTGCCTTGCGCCGGCTGAGCCACTGCTTCGTCCGGGCGGCCAAGAGGAGC
 ATGAGGGAGGACTCCTGGTCCAGGGTCCTGACGTGGTCGCGGCTCTGGGAGCGG
 GCCAGCATCATCTGGCTCTGCCGCACCGAGGCCGCCTCCAACTGGTCCTCCAGCA
 GCCGCACTCGCCGCCGACCCTGGCAGAGGAAGACAGGTGAGGGGGGTATGAATT
 GTACAGAACAAACCACGAGCCTTGTCTAGGCAGAATCCCTACCAGTCATGGCTTTA
 CCTGGATGACGGCCTGCGAACAGCTGTCCAGCGACCCTCGCTGCCGCCGCTTCTC
 CCGCACGCTTCTTTCCAGCACCGTGATGGCGCGAGCCAGCGCCGCACGCTGGCGC
 TGCGCTTCGCCGATCTGAGGACAGTCGGGGAACCTCTGATCAGTCTAAACCCCTT
 GCGCGTTAGTGTTGCCATCCTTTGCAGACCGGTGAGAGCCGACTTGTTGTGCGCC
 ACCCCCCACACCACCTCCTCCCAGACCAATTCTGTACCTTTTTGGCGAAGGCAT
 CGGCCTCGGCCTGCAGAGAGGACAGCAGTGCCAGCCGCTGGGGGTGGCGGAT
 GCACGCTCAGGTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGG
 GCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCC
 GAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGT
 TAAATAGCCAGGCCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCAT

ATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCG
CACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACCTC
TAGAATATCAATGATCGAGCAGGACGGCCTCCACGCCGGCTCCCCCGCCGCCTG
GGTGGAGCGCCTGTTTCGGCTACGACTGGGCCAGCAGACCATCGGCTGCTCCGA
CGCCGCCGTGTTCCGCCTGTCCGCCAGGGCCGCCCGTGCTGTTTCGTGAAGACC
GACCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCCTGTCTTG
CTGGCCACCACCGGCGTGCCCTGCGCCGCCGTGCTGGACGTGGTGACCGAGGCC
GGCCGCGACTGGCTGCTGCTGGGCGAGGTGCCCGGCCAGGACCTGCTGTCTTCCC
ACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCGACGCCATGCGCCGCCTGC
ACACCCTGGACCCCGCCACCTGCCCCCTTCGACCACCAGGCCAAGCACCGCATCGA
GCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGGACCAGGACGACCTGGACGA
GGAGCACCAGGGCCTGGCCCCCGCCGAGCTGTTTCGCCCGCCTGAAGGCCCGCAT
GCCCAGCGGCGAGGACCTGGTGGTGACCCACGGCGACGCCTGCCTGCCCAACAT
CATGGTGGAGAACGGCCGCTTCTCCGGCTTCATCGACTGCGGCCGCCTGGGCGTG
GCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACATCGCCGAGGAGCTG
GGCGGCGAGTGGGCCGACCGCTTCTGCTGTACGGCATCGCCGCCCGGACT
CCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGACAATTGGCAGC
AGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTG
CCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCC
TCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCT
ATTTGCGAATACCACCCCGAGCATCCCTTCCCTCGTTTCATATCGCTTGCATCCC
AACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCT
CACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCA
ACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACA
CAAATGGAGGATCCACTAGAGAGTTTAGGTCCAGCGTCCGTGGGGGGGGACGGG
CTGGGAGCTTGGGCCGGGAAGGGCAAGACGATGCAGTCCCTCTGGGGAGTCACA
GCCGACTGTGTGTGTTGCACTGTGCGGCCCGCAGCACTCACACGCAAAATGCCTG
GCCGACAGGCAGGCCCTGTCCAGTGCAACATCCACGGTCCCTCTCATCAGGCTCA
CCTTGCTCATTGACATAACGGAATGCGTACCGCTCTTTCAGATCTGTCCATCCAG
AGAGGGGAGCAGGCTCCCCACCGACGCTGTCAAACCTTGCTTCCTGCCCAACCGA
AAACATTATTGTTTGAGGGGGGGGGGGGGGGGGGGCAGATTGCATGGCGGGATATC
TCGTGAGGAACATCACTGGGACACTGTGGAACACAGTGAGTGACAGTATGCAGAG
CATGTATGCTAGGGGTCAGCGCAGGAAGGGGGCCTTTCCCAGTCTCCCATGCCAC
TGCACCGTATCCACGACTCACCAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGT
GGACACCAAGTGTGTAGCCTCTGGACTCCAGGTATGCGTGACCGCAAAGGCCAG
CCGATCGTGCCGATTCTGGGGTGGAGGATATGAGTCAGCCAACTTGGGGCTCA
GAGTGACACTGGGGCACGATACGAAACAACATCTACACCGTGTCTCCATGCT
GACACACCACAGCTTCGCTCCACCTGAATGTGGGCGCATGGGGCCGAATCACAG
CCAATGTGCTGCTGCCATAATGTGATCCAGACCCTCTCCGCCAGATGCCGAGC
GGATCGTGGGCGCTGAATAGATTCTGTTTCGATCACTGTTTGGGTCTTTCTTTT
TCGTCTCGGATGCGCGTCTCGAAACAGGCTGCGTCGGGCTTTCGGATCCCTTTTG
CTCCCTCCGTCACCATCCTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACC
AGGGTTGGAGGGTATTACCGCGTCAGGCCATTCCCAGCCCGGATTCAATTCAAAG
TCTGGGCCACCACCCTCCGCCGCTCTGTCTGATCACTCCACATTCGTGCATACACT
ACGTTCAAGTCTGATCCAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAAT
CTCAAGGACCCACTCCAGCACAGCTGCTGGTTGACCCCGCCCTCGCAAACCTAGTA
TGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCG
CTCGGCGGGCTCCGGGGCCCCGGCGCCAGCGAGGCCCTCCCCGTGCGCGGGCG
CGCCGTCCAGGCCGCGGCCACCCGCTTCAAGAAGGAGACGACGACCACCCGCGC
CACGCTGACGTTTCGACCCCCCACGACCAACTCCGAGCGCGCCAAGCAGCGCAA

GCACACCATCGACCCCCTCCTCCCCCGACTTCCAGCCCATCCCCCTCCTTCGAGGAG
TGCTTCCCCAAGTCCACGAAGGAGCACAAGGAGGTGGTGCACGAGGAGTCCGGC
CACGTCCTGAAGGTGCCCTTCCGCCGCGTGCACCTGTCCGGCGGCGAGCCCGCCT
TCGACAACCTACGACACGTCCGGCCCCCAGAACGTCAACGCCCACATCGGCCTGG
CGAAGCTGCGCAAGGAGTGGATCGACCGCCGCGAGAAGCTGGGCACGCCCCGCT
ACACGCAGATGTACTACGCGAAGCAGGGCATCATCACGGAGGAGATGCTGTACT
GCGCGACGCGCGAGAAGCTGGACCCCCGAGTTCGTCCGCTCCGAGGTGCGCGGG
GCCGCGCCATCATCCCCCTCCAACAAGAAGCACCTGGAGCTGGAGCCCATGATCG
TGGGCCGCAAGTTCCTGGTGAAGGTGAACGCGAACATCGGCAACTCCGCCGTGG
CCTCCTCCATCGAGGAGGAGGTCTACAAGGTGCAGTGGGCCACCATGTGGGGCG
CCGACACCATCATGGACCTGTCCACGGGCGCCACATCCACGAGACGCGCGAGT
GGATCCTGCGCAACTCCGCGGTCCCCGTGGGCACCGTCCCCATCTACCAGGCGCT
GGAGAAGGTGGACGGCATCGCGGAGAACCTGAACTGGGAGGTGTTCCGCGAGAC
GCTGATCGAGCAGGCCGAGCAGGGCGTGGACTACTTCACGATCCACGCGGGCGT
GCTGCTGCGCTACATCCCCCTGACCGCCAAGCGCCTGACGGGCATCGTGTCCCGC
GGCGGCTCCATCCACGCGAAGTGGTGCCTGGCCTACCACAAGGAGAACTTCGCC
TACGAGCACTGGGACGACATCCTGGACATCTGCAACCAGTACGACGTCGCCCTGT
CCATCGGCGACGGCCTGCGCCCCGGCTCCATCTACGACGCCAACGACACGGCCC
AGTTCGCCGAGCTGCTGACCCAGGGCGAGCTGACGCGCCGCGCGTGGGAGAAGG
ACGTGCAGGTGATGAACGAGGGCCCCGGCCACGTGCCCATGCACAAGATCCCCG
AGAACATGCAGAAGCAGCTGGAGTGGTGC AACGAGGCGCCCTTCTACACCCTGG
GCCCCCTGACGACCGACATCGCGCCCCGGCTACGACCACATCACCTCCGCCATCGG
CGCGGCCAACATCGGCGCCCTGGGCACCGCCCTGCTGTGCTACGTGACGCCAA
GGAGCACCTGGGCCTGCCCAACCGCGACGACGTGAAGGCGGGCGTCATCGCCTA
CAAGATCGCCGCCACGCGGGCCGACCTGGCCAAGCAGCACCCCCACGCCCAGGC
GTGGGACGACGCGCTGTCCAAGGCGCGCTTCGAGTTCGCTGGATGGACCAGTTC
GCGCTGTCCCTGGACCCCATGACGGCGATGTCCTTCCACGACGAGACGCTGCCCG
CGGACGGCGCGAAGGTGCGCCACTTCTGCTCCATGTGCGGCCCAAGTTCTGCTC
CATGAAGATCACGGAGGACATCCGCAAGTACGCCGAGGAGAACGGCTACGGCTC
CGCCGAGGAGGCCATCCGCCAGGGCATGGACGCCATGTCCGAGGAGTTCAACAT
CGCCAAGAAGACGATCTCCGGCGAGCAGCACGGCGAGGTGCGCGGCGAGATCTA
CCTGCCCCGAGTCCTACGTCAAGGCCGCGCAGAAGTGATACGTACTCGAGACGGA
GCGTCGTGCGGGAGGGAGTGTGCCGAGCGGGGAGTCCCGGTCTGTGCGAGGCC
GGCAGCTGACGCTGGCGAGCCGTACGCCCCGAGGGTCCCCCTCCCCTGCACCCTC
TTCCCCTTCCCTCTGACGGCCGCGCCTGTTCTTGCATGTTACGCGACGAGCTCTTG
TTTTCCAGAAGGAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGATAACCTCCAA
AGCCGCTCTAATTGTGGAGGGGGTTTCAATTTAAAAGCTTGGAATGTTGGTTTCGT
GCGTCTGGAACAAGCCCAGACTTGTTGCTCACTGGGAAAAGGACCATCAGCTCC
AAAAAACTTGCCGCTCAAACCGCGTACCTCTGCTTTCGCGCAATCTGCCCTGTTG
AAATCGCCACCACATTTCATATTGTGACGCTTGAGCAGTCTGTAATTGCCTCAGAA
TGTGGAATCATCTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCGCGGGCGAGG
ACACCCGCCACTCGTACAGCAGACCATTATGCTACCTCACAATAGTTCATAACAG
TGACCATATTTCTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAGTGGCCACC
CCCCGGCCCTGGTGTCTGCGGAGGGCAGGTCAACCGGCATGGGGCTACCGAAAT
CCCCGACCGGATCCCACCACCCCCGCGATGGGAAGAATCTCTCCCCGGGATGTG
GGCCACCAACCAGCACAACTGCTGGCCCAGGCGAGCGTCAAACCATAACCACAC
AAATATCCTTGGCATCGGCCCTGAATTCTTCTGCCGCTCTGCTACCCGGTGCTTC
TGTCCGAAGCAGGGGTGCTAGGGATCGCTCCGAGTCCGCAAACCCTTGTGCGGT
GGCGGGGCTTGTTTCGAGCTTGAAGAGC

SEQ ID NO: 135

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

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGCGACCTGCGTC
GCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCGCGGGC
GCGCCGTCCAGGCCGCGGCCACCCGCTTCAAGAAGGAGACGACGACCACCCGCG
CCACGCTGACGTTTCGACCCCCCACGACCAACTCCGAGCGCGCCAAGCAGCGCA
AGCACACCATCGACCCCTCCTCCCCGACTTCCAGCCCATCCCCCTCCTTCGAGGA
GTGCTTCCCCAAGTCCACGAAGGAGCACAAGGAGGTGGTGCACGAGGAGTCCGG
CCACGTCTGAAGGTGCCCTTCCGCCGCGTGCACCTGTCCGGCGGGCAGCCCCGCC
TTCGACAACCTACGACACGTCCGGCCCCCAGAACGTCAACGCCACATCGGCCTG
GCGAAGCTGCGCAAGGAGTGGATCGACCGCCGCGAGAAGCTGGGACGCCCCCG
TACACGCAGATGTACTACGCGAAGCAGGGCATCATCACGGAGGAGATGCTGTAC
TGCGCGACGCGCGAGAAGCTGGACCCCGAGTTCGTCCGCTCCGAGGTCGCGCGG
GGCCGCGCCATCATCCCCTCCAACAAGAAGCACCTGGAGCTGGAGCCCATGATC
GTGGGCGCGCAAGTTCCTGGTGAAGGTGAACGCGAACATCGGCAACTCCGCCGTG
GCCTCCTCCATCGAGGAGGAGGTCTACAAGGTGCAGTGGGCCACCATGTGGGGC
GCCGACACCATCATGGACCTGTCCACGGGCCGCCACATCCACGAGACGCGCGAG
TGGATCCTGCGCAACTCCGCGGTCCCCGTGGGCACCGTCCCCATCTACCAGGCGC
TGGAGAAGGTGGACGGCATCGCGGAGAACCTGAACTGGGAGGTGTTCCGCGAGA
CGCTGATCGAGCAGGCCGAGCAGGGCGTGGACTACTTCACGATCCACGCGGGCG
TGCTGCTGCGCTACATCCCCCTGACCGCCAAGCGCCTGACGGGCATCGTGTCCCG
CGGCGGCTCCATCCACGCGAAGTGGTGCCTGGCCTACCACAAGGAGAACTTCGC
CTACGAGCACTGGGACGACATCCTGGACATCTGCAACCAGTACGACGTGCGCCT
GTCCATCGGCGACGGCCTGCGCCCCGGCTCCATCTACGACGCCAACGACACGGC
CCAGTTCGCCGAGCTGCTGACCCAGGGCGAGCTGACGCGCCGCGCGTGGGAGAA
GGACGTGCAGGTGATGAACGAGGGGCCCGGCCACGTGCCCATGCACAAGATCCC
CGAGAACATGCAGAAGCAGCTGGAGTGGTGCAACGAGGCGCCCTTCTACACCCT
GGGCCCCCTGACGACCGACATCGCGCCCGGCTACGACCACATCACCTCCGCCATC
GGCGCGGCCAACATCGGCGCCCTGGGCACCGCCCTGCTGTGCTACGTGACGCC
AAGGAGCACCTGGGCCTGCCCAACCGCGACGACGTGAAGGCGGGCGTCATCGCC
TACAAGATCGCCGCCACGCGGGCCGACCTGGCCAAGCAGCACCCCCACGCCAG
GCGTGGGACGACGCGCTGTCCAAGGCGCGCTTCGAGTTCGCTGGATGGACCAG
TTCGCGCTGTCCCTGGACCCCATGACGGCGATGTCTTCCACGACGAGACGCTGC
CCGCGGACGGCGCGAAGGTGCGCCACTTCTGCTCCATGTGCGGCCCAAGTTCTG
CTCCATGAAGATCACGGAGGACATCCGCAAGTACGCCGAGGAGAACGGCTACGG
CTCCGCCGAGGAGGCCATCCGCCAGGGCATGGACGCCATGTCCGAGGAGTTCAA
CATCGCCAAGAAGACGATCTCCGGCGAGCAGCACGGCGAGGTCGGCGGGCAGAT
CTACCTGCCCCGAGTCCTACGTCAAGGCCGCGCAGAAAGTGA

SEQ ID NO: 136

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

MAASVHCTLMSVVCNNKNHSARPKLPNSSLLPGFDVVVQAAATRFKKETTTTRATL
TFDPPTTNSERAKQRKHTIDPSSPDFQPIPSFEFCFPKSTKEHKEVVHEESGHVLKVPF
RRVHLSGGEPAFDNYDTSGPQNVNAHIGLAKLRKEWIDRREKLGTPTYTQMYYAKQ
GIITEEMLYCATREKLDPEFVRSEVARGRAIIPSNKKHLELEPMIVGRKFLVKVNANIG
NSAVASSIEEEVYKVQWATMWGADTIMDLSTGRHIHETREWILRNSAVPVGTVPIYQ
ALEKVDGIAENLNWEVFRETLEQAEQGVDFYTIHAGVLLRYIPLTAKRLTGIVSRGG
SIHAKWCLAYHKENFAYEHWDILDICNQYDVALSIGDGLRPGSIYDANDTAQFAEL

LTQGELTRRAWEKDVQVMNEGPGHVPMHKIPENMQKQLEWCNEAPFYTLGPLTTD
IAPGYDHITSAIGAANIGALGTALLCYVTPKEHLGLPNRDDVKAGVIAKYIAAHAAD
LAKQHPHAQAWDDALSKARFEFRWMDQFALSLDPMTAMSFHDETLPADGAKVAH
FCSMCGPKFCSMKITEDIRKYAEENGYGSAEEAIRQGM DAMSEEFNIAKKTISGEQH
GEVGGEIYLPESYVKAQK

SEQ ID NO: 137

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

GCTCTTCGCCGCCGCCACTCCTGCTCGAGCGCGCCCGCGCGTGCGCCGCCAGCGC
CTTGCCCTTTTCGCCCGCGCTCGTGCGCGTGCCTGATGTCCATCACCAGGTCCATG
AGGTCTGCCTTGCGCCGGCTGAGCCACTGCTTCGTCCGGGCGGCCAAGAGGAGC
ATGAGGGAGGACTCCTGGTCCAGGGTCCTGACGTGGTTCGCGGCTCTGGGAGCGG
GCCAGCATCATCTGGCTCTGCCGCACCGAGGCCGCCTCCAACTGGTCTCCAGCA
GCCGCAGTCGCCGCCGACCCTGGCAGAGGAAGACAGGTGAGGGGGGTATGAATT
GTACAGAACAAACCACGAGCCTTGCTAGGCAGAAATCCCTACCAGTCATGGCTTTA
CCTGGATGACGGCCTGCGAACAGCTGTCCAGCGACCCCTCGCTGCCGCCGCTTCTC
CCGCACGCTTCTTTCCAGCACCGTGATGGCGCGAGCCAGCGCCGCACGCTGGCGC
TGCGCTTCGCCGATCTGAGGACAGTCGGGGAACCTCTGATCAGTCTAAACCCCTT
GCGCGTTAGTGTTGCCATCCTTTGCAGACCGGTGAGAGCCGACTTGTTGTGCGCC
ACCCCCACACCACCTCCTCCCAGACCAATTCTGTACCTTTTTGGCGAAGGCAT
CGGCCTCGGCCTGCAGAGAGGACAGCAGTGCCAGCCGCTGGGGGTGGCGGAT
GCACGCTCAGGTACCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGG
GCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCC
GAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGT
TAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCAT
ATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCG
CACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACTC
TAGAATATCAATGATCGAGCAGGACGGCCTCCACGCCGGCTCCCCCGCCGCCTG
GGTGGAGCGCCTGTTCCGGCTACGACTGGGCCAGCAGACCATCGGCTGCTCCGA
CGCCGCCGTGTTCCGCCTGTCCGCCAGGGCCGCCCGCTGCTGTTCTGTGAAGACC
GACCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCCTGTCTTG
CTGGCCACCACCGGCGTGCCCTGCGCCGCCGTGCTGGACGTGGTGACCGAGGCC
GGCCGCGACTGGCTGCTGCTGGGCGAGGTGCCCGGCCAGGACCTGCTGTCTCTCC
ACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCGACGCCATGCGCCGCCTGC
ACACCCTGGACCCCGCCACCTGCCCCCTTCGACCACCAGGCCAAGCACCGCATCGA
GCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGGACCAGGACGACCTGGACGA
GGAGCACCAGGGCCTGGCCCCCGCCGAGCTGTTCCGCCCGCTGAAGGCCCGCAT
GCCCGACGGCGAGGACCTGGTGGTGACCCACGGCGACGCCTGCCTGCCCAACAT
CATGGTGGAGAACGGCCGCTTCTCCGGCTTCATCGACTGCGGCCGCTGGGCGTG
GCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACATCGCCGAGGAGCTG
GGCGGCGAGTGGGCCGACCGCTTCTTGGTGCTGTACGGCATCGCCGCCCGGACT
CCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGACAATTGGCAGC
AGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTG
CCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCC
TCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCT
ATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCATCCC
AACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCT
CACTGCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCTGGTACTGCA
ACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACA

CAAATGGAGGATCCACTAGAGAGTTTAGGTCCAGCGTCCGTGGGGGGGGGACGGG
 CTGGGAGCTTGGGCCGGGAAGGGCAAGACGATGCAGTCCCTCTGGGGAGTCACA
 GCCGACTGTGTGTGTTGCACTGTGCGGCCCGCAGCACTCACACGCAAAATGCCTG
 GCCGACAGGCAGGCCCTGTCCAGTGCAACATCCACGGTCCCTCTCATCAGGCTCA
 CCTTGCTCATTGACATAACGGAATGCGTACCGCTCTTTCAGATCTGTCCATCCAG
 AGAGGGGAGCAGGCTCCCCACCGACGCTGTCAAACCTTGCTTCCTGCCCCAACCGA
 AAACATTATTGTTTGAGGGGGGGGGGGGGGGGGGGGCAGATTGCATGGCGGGATATC
 TCGTGAGGAACATCACTGGGACACTGTGGAACACAGTGAGTGCAGTATGCAGAG
 CATGTATGCTAGGGGTGAGCGCAGGAAGGGGGCCTTTCAGTCTCCCATGCCAC
 TGCACCGTATCCACGACTCACCAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGT
 GGACACCAAGTGTGTAGCCTCTGGACTCCAGGTATGCGTGCACCGCAAAGGCCAG
 CCGATCGTGCCGATTCCCTGGGGTGGAGGATATGAGTCAGCCAACTTGGGGCTCA
 GAGTGCACACTGGGGCACGATACGAAACAACATCTACACCGTGTCTCCATGCT
 GACACACCACAGCTTCGCTCCACCTGAATGTGGGCGCATGGGCCCCGAATCACAG
 CCAATGTGCTGCTGCCATAATGTGATCCAGACCCTCTCCGCCAGATGCCGAGC
 GGATCGTGGGCGCTGAATAGATTCTGTTCGATCACTGTTTGGGTCTTTCTTTT
 TCGTCTCGGATGCGCGTCTCGAAACAGGCTGCGTGGGGCTTTCGGATCCCTTTTG
 CTCCCTCCGTCACCATCCTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACC
 AGGGTTGGAGGGTATTACCGCGTCAGGCCATTCCCAGCCCCGGATTCAATTCAAAG
 TCTGGGCCACCACCCTCCGCCGCTCTGTCTGATCACTCCACATTCGTGCATACACT
 ACGTTCAAGTCCTGATCCAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAAT
 CTCAAGGACCCACTCCAGCACAGCTGCTGGTTGACCCCGCCCTCGCAAACCTAGTA
 TGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCG
 CTCGGCGGGCTCCGGGGCCCCGGCGCCAGCGAGGCCCCCTCCCGTGCAGCGGGCG
 CGCCATGCGCACCGCGTGGGTCGCCAAGCGCCAGGGCCAGACGAACGTCTCCCA
 GATGCACTACGCGCGCAAGGGCGTGATCACCGAGGAGATGGACTACGTGGCGAA
 GCGCGAGAACCTGCCCCGTGGAGCTGATCAAGGACGAGGTGGCGCGGGGGCCGCAT
 GATCATCCCCGCCAACATCAACCACACGAACCTGGAGCCCATGGCCATCGGCAT
 CGCCTCCAAGTGCAAGGTCAACGCGAACATCGGCGCGTCCCCCAACTCCTCCAAC
 ATCGACGAGGAGGTGGAGAAGCTGCTGCTGTCCGTGAAGTACGGCGCGGACACG
 GTGATGGACCTGTCCACCGGCGGGCGGCGACCTGGACGTGATCCGCACCGCCATC
 ATCAACGCCTCCCCCGTGCCCATCGGCACGGTGCCCATCTACCAGGCCCTGGAGT
 CCGTGACGGCTCCATCGAGAACCTGACGCCCCGACGACTTCCTGCACATCATCGA
 GAAGCACGCCCAGCAGGGCGTGGACTACATGACGATCCACGCGGGCCTGCTGAT
 CGAGTACCTGCCCCCTGGTGAAGTCCCGCATCACGGGCATCGTGTCCCGCGGGCGG
 GGCATCATCGCCAAGTGGATGCTGCACCACCACAAGCAGAACCCCTGTACACG
 CACTTCGACGAGATCATCGAGATCTTCAAGAAGTACGACGTGTCTTCTCCCTGG
 GCGACTCCCTGCGCCCCGGCTGCACCCACGACGCCTCCGACGACGCCCAGCTGTC
 CGAGCTGAAGACGCTGGGCCAGCTGACGCGCCGCGCCTGGGAGCACGACGTCCA
 GGTCATGGTGGAGGGCCCCGGCCACGTCCCCATCGACCAGATCGAGTTCAACGT
 CAAGAAGCAGATGGAGGAGTGCTCCGAGGCCCCCTTCTACGTGCTGGGCCCCCT
 GGTCACCGACATCGCGCCCGGCTACGACCACATCACCTCCGCGATCGGCGCCGCC
 ATCGCGGGCTGGCACGGCACCGCGATGCTGTGCTACGTGACGCCCAAGGAGCAC
 CTGGGGCCTGCCCAACGCGGAGGACGTGCGCAACGGCCTGATCGCCTACAAGATC
 GCCGCGCACGCCGCGGACATCGCCCGCCACCGCCCCGGCGCGCGGACCGCGAC
 GACGAGCTGTCCAAGGCGCGCTACAACCTTCGACTGGAACCGCCAGTTCGAGCTG
 TCCCTGGACCCCGAGCGCGCGAAGGAGTACCACGACGAGACGCTGCCCGCGGAC
 ATCTACAAGACCGCGGAGTTCTGCTCCATGTGCGGCCCAAGTTCTGCCCCATGC
 AGACCAAGGTGGACGCCGAGATGCTGGAGGAGCTGGAGGTCTTCTGGCCAAGG
 ACAAGGAGATGGTGTCCAGCGCTGATACGTACTCGAGACGGAGCGTCTGTGCGG

GAGGGAGTGTGCCGAGCGGGGAGTCCCCGGTCTGTGCGAGGCCCCGGCAGCTGACG
 CTGGCGAGCCGTACGCCCCGAGGGTCCCCCTCCCCTGCACCTCTTCCCCCTTCCCT
 CTGACGGCCGCGCCTGTTCTTGCATGTTTCAGCGACGAGCTCTTGTTTTCCAGAAG
 GAGTTGCTCCTTGAGCCTTTCATTCTCAGCCTCGATAACCTCCAAAGCCGCTCTAA
 TTGTGGAGGGGGTTCGAATTTAAAAGCTTGGAATGTTGGTTCGTGCGTCTGGAAC
 AAGCCCAGACTTGTTGCTCACTGGGAAAAGGACCATCAGCTCCAAAAAACTTGC
 CGCTCAAACCGCGTACCTCTGCTTTCGCGCAATCTGCCCTGTTGAAATCGCCACC
 ACATTCAATATTGTGACGCTTGAGCAGTCTGTAATTGCCTCAGAATGTGGAATCAT
 CTGCCCCCTGTGCGAGCCCATGCCAGGCATGTCGCGGGCGAGGACACCCGCCAC
 TCGTACAGCAGACCATTATGCTACCTCACAATAGTTCATAACAGTGACCATATTT
 CTCGAAGCTCCCCAACGAGCACCTCCATGCTCTGAGTGGCCACCCCCCGGCCCTG
 GTGCTTGCGGAGGGCAGGTCAACCGGCATGGGGCTACCGAAATCCCCGACCGGA
 TCCCACCACCCCCGCGATGGGAAGAATCTCTCCCCGGGATGTGGGGCCCACCACCA
 GCACAACCTGCTGGCCCAGGCGAGCGTCAAACCATAACCACACAAATATCCTTGG
 CATCGGCCCTGAATTCCTTCTGCCGCTCTGCTACCCGGTGCTTCTGTCCGAAGCAG
 GGGTTGCTAGGGATCGCTCCGAGTCCGCAAACCTTGTGCGGTGGCGGGGCTTGT
 TCGAGCTTGAAGAGCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGCGTAATC
 ATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAAC
 ATACGAGCCGGAAGCATAAAGTGTAAGCCTGGGGTGCCTAATGAGTGAGCTAA
 CTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCAGTCGGGAAACCTGTGCT
 GCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTG
 GGCGCTCTTC

SEQ ID NO: 138

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

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTC
 GCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCCCTCCCCGTGCGCGGGC
 GCGCCATGCGCACCGCGTGGGTTCGCCAAGCGCCAGGGCCAGACGAACGTCTCCC
 AGATGCACTACGCGCGCAAGGGCGTGATCACCGAGGAGATGGACTACGTGGCGA
 AGCGCGAGAACCTGCCCCGTGGAGCTGATCAAGGACGAGGTGGCGCGGGGCGCGCA
 TGATCATCCCCGCCAACATCAACCACACGAACCTGGAGCCCATGGCCATCGGCAT
 CGCCTCCAAGTGCAAGGTCAACGCGAACATCGGCGCGTCCCCCAACTCCTCCAAC
 ATCGACGAGGAGGTGGAGAAGCTGCTGCTGTCCGTGAAGTACGGCGCGGACACG
 GTGATGGACCTGTCCACCGGCGGCGGCGACCTGGACGTGATCCGCACCGCCATC
 ATCAACGCCTCCCCCGTGCCCATCGGCACGGTGCCCATCTACCAGGCCCTGGAGT
 CCGTGACGGCTCCATCGAGAACCTGACGCCCCGACGACTTCCTGCACATCATCGA
 GAAGCACGCCCAGCAGGGCGTGGAATACATGACGATCCACGCGGGCCTGCTGAT
 CGAGTACCTGCCCCCTGGTGAAGTCCCGCATCACGGGCATCGTGTCCCGCGGGCGGC
 GGCATCATCGCCAAGTGGATGCTGCACCACCACAAGCAGAACCCCTGTACACG
 CACTTCGACGAGATCATCGAGATCTTCAAGAAGTACGACGTGTCTTCTCCCTGG
 GCGACTCCCTGCGCCCCGGCTGCACCCACGACGCCTCCGACGACGCCAGCTGTC
 CGAGCTGAAGACGCTGGGCCAGCTGACGCGCCGCGCCTGGGAGCACGACGTCCA
 GGTCATGGTGGAGGGCCCCGGCCACGTCCCCATCGACCAGATCGAGTTCAACGT
 CAAGAAGCAGATGGAGGAGTGCTCCGAGGCCCCCTTCTACGTGCTGGGCCCCCT
 GGTCACCGACATCGCGCCCGGCTACGACCACATCACCTCCGCGATCGGCGCCGCC
 ATCGCGGGCTGGCACGGCACCGCGATGCTGTGCTACGTGACGCCCAAGGAGCAC
 CTGGGCCTGCCAACGCGGAGGACGTGCGCAACGGCCTGATCGCCTACAAGATC
 GCCGCGCACGCCGCGGACATCGCCCGCCACCGCCCCGGCGCGCGGACCGCGAC
 GACGAGCTGTCCAAGGCGCGCTACAACCTTCGACTGGAACCGCCAGTTCGAGCTG

TCCCTGGACCCCGAGCGCGCGAAGGAGTACCACGACGAGACGCTGCCCCGCGGAC
ATCTACAAGACCGCGGAGTTCTGCTCCATGTGCGGCCCAAGTTCTGCCCCATGC
AGACCAAGGTGGACGCCGAGATGCTGGAGGAGCTGGAGGTCTTCCTGGCCAAGG
ACAAGGAGATGGTGTCCCAGCGCTGA

SEQ ID NO: 139

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

MRTAWVAKRQGQTNVSQMHYARKGVITEEMDYVAKRENLPVELIKDEVARGRMII
PANINHTNLEPMAIGIASKCKVNANIGASPNSSNIDEEVEKLLLSVKYGADTVMDLST
GGGDLDVIRTAIINASPVPITVPIYQALSVHGSIENTPDDFLHIIIEKHAQQGVNDYM
TIHAGLLIEYLPLVKSRITGIVSRGGGIIAKWMLHHHKQNPLYTHFDEIIIEIFKKYDVSF
SLGDSLRLPGCTHDASDDAQLSELKTLGQLTRRAWVQVQVMVEGPGHVPIDQIEFN
VKKQMEECSEAPFYVLGPLVTDIAPGYDHITSAGAAIAGWHGTAMLCYVTPKEHLG
LPNAEDVRNGLIAYKIAAHAADIARHRPGARDRDELKARYNFDWNRQFELSLDP
ERAKEYHDETLPADIIYKTAEFCSMCGPKFCPMQTKVDAEMLEELEVFLAKDKEMVS
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

EVHVQVTHSMPPQKIEIFKSLDNWAEENILVHLKPVEKC'WQPQDFLPDPASDGFDEQVRE
LRERAKEIPDDYFVVLVGDMITEEALPTYQTMNLTDGVRDET'GASPTSWAIWTRAWTAE
ENRHGDLNLYL'LSGRVDMRQIEKTIQY'LGSGMDPRTENSPYLGFIYTSFQERATFIS
HGNTARQAKEHGDIKLAQICGTIAADEKRHETAYTKIVEKLFEIDPDGTVLAFADMMRKK
ISMPAHLMYDGRDDNLFDFHSAVAQRLGVYAKDYADILEFLVGRWKVDKLTGLSAEGQK
AQDYVCRLPPRIRLRLEERAQGRAKEAPTMPFSWIFDRQVKLMDYKDHDGDYKDHDIDYKD
DDDK

SEQ ID NO: 149

Nucleotide sequence of transforming DNA contained in pSZ1503_[KASII_btub-y.inv-
nr_KASII]

gctttcccgcaaccggtgggtccacccaacttgaacctcgagaaccccgcgctggcgctgaccccgctgctgctggggc
cgcggaaggagcgcgccgaagacctggacgtcgctctccaactcctttggctttggcgggcacaattcgctgctggtacc
ctttctgcgctatgacactccagcaaaaggtagggcgggctgcgagacggcttccggcgctgcatgcaacaccgatgatgcttcg
acccccgaagctccttcggggctgcatggcgctccgatccgctccaggcgagcgctgtttaaatagccaggcccccgattgc
aaagacattatagcgagctaccaaaagccatattcaaacacctagatcactaccactttacacaggccactcgagcttgtgatcgactc
cgctaagggggcgctcttctcttctgttcagtcacaaccgcaaac**ggcgcgccATG**ctgtctgaggccttctgttctgtctgg
ccggttcgccccaagatcagcgctccatgacgaacgagacgtccgaccgccccctggcgacttcaccccaacaagggt
ggatgaacgaccccaacggcctgtggtacgacgagaaggacgccaagtggcacctgtacttccagtacaacccgaacgacac
cgtctgggggacgcccttgttctggggccacgccacgtccgacgacctgaccaactgggaggaccagcccatgccatgccccc
gaagcgcaacgactccggcgcttctccggctccatggtggtggactacaacaacacctccggttcttcaacgacaccatcgac
ccgcgccagcgctgcgtggccatctggacctacaacaccccgagtcggaggagcagtacatctctacagcctggacggcgg
ctacaccttcaccgagtaccagaagaaccccgctggcgcccaactccacccagttccgacccgaaggcttctggtacgag
ccctcccagaagtggatcatgaccgcggccaagtcccaggactacaagatcgagatctactctccgacgacctgaagtcttg
aagctggagtcgcgttcgccaacgagggcttctcggtaccagtacgagtccccggcctgatcgaggtccccaccgagcag
gaccccgcaagtcctactgggtgatgttcattccataaccccgcgccccggcgccggtccttcaaccagtacttctcg
cagcttcaacggcaccacttcgaggccttcgacaaccagtcgccggtggtgacttcggcaaggactactacgccctcgagacc
tttctcaacaccgacccgacctacgggagcgccctgggcatcgcgtggcctccaactgggagtactccgccttcgtgccacca
acccctggcgcttctccatgtccctgtgcgaagtttccctcaacaccgagtaccaggccaacccggagacggagctgatcaa
cctgaaggccgagccgatcctgaacatcagcaacgccggccctggagccggttcgccaccaacaccacgttgacgaaggcc
aacagctacaacgtcgacctgtccaacgacccggcaccctggagttcgagctggtgtacgccgtcaacaccacccagacgatc
tccaagtcgtgttcgggaccttctccctgtgttcaaggcgctggaggaccccgaggagtacctccgcatgggttcgaggtgtc
cgcgtctcttcttcttgaccgcgggaacagcaaggtgaagttcgtgaaggagaacccctacttaccdaaccgatgagcgtg
aacaaccagccctcaagagcgagaacgacctgtctactacaaggtgtacggcttctggaccagaacatcctggagctgtact
tcaacgacggcgacgtcgtgtccaccaacacctacttcatgaccaccgggaacgccctgggctccgtgaacatgacgacgggg
gtggacaacctgttctacatcgacaagttccaggtgcgcgaggtcaag**TGA****caattggcagcagcagctcggatagtatcgac**
acactctggacgctggtcgtgtgatggactgttgcgccacacttgccttgacctgtgaatatccctgccgctttatcaaacagcct
cagtgtgtttgatctgtgtacgcgcttttgcgagttgctagctgcttgcctatttgcgaataccacccccagcatcccttccctcgtt
catatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgccctcgcacagc
cttggtttgggctccgctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtggatggga
acacaatggaggatcgtagagctcatcttccgaaagtacgacgagtgagcgagctgattctctttgagcggggtcgggtgggttc
ggggagagtgcgcggaaggcgagagacgtgcggccggcggttccctccgttccctgggttggtgctatagtaacctgc
ctgtgtcgctgcgcgtcgggaagagc

SEQ ID NO: 150

***B. braunii* malate dehydrogenase 5'UTR**

aattgaaaccccgcgcaagaccgggtgtttggccgcctgaccggaaaggggggcctgtcccgaagggggctatctcttgggg
gatgtcgggcgcggaaagtcgatgttgatggaccttcttcgacctatgtcggggtcgaaggccaagacccgctccatttcgccagt

SEQ ID NO: 152

Chlamydomonas RBCS2 promoter

[illegible]

ctcatgttaccgatgtattcggagaacggcaactaagctgccgggttgaaacacggatgatctcgcggaggtagcatgttgattg
taacgatgacagagcgttgctgcctgtgatcaccgcgg

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
CTCCCCACCGACGCTGTCAAACCTTGCTTCCCTGCCCCAACCGAAAACATTATTGTTT
GAGGGGGGGGGGGGGGGGGGGCAGATTGCATGGCGGGATATCTCGTGAGGAACAT
CACTGGGACACTGTGGAACACAGTGAGTGCAGTATGCAGAGCATGTATGCTAGG
GGTCAGCGCAGGAAGGGGGCCTTTCCAGTCTCCCATGCCACTGCACCGTATCCA
CGACTCACCAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGTGGACACCAGTGTG
TAGCCTCTGGACTCCAGGTATGCGTGCACCGCAAAGGCCAGCCGATCGTGCCGAT
TCCTGGGGTGGAGGATATGAGTCAGCCAACCTTGGGGCTCAGAGTGCACACTGGG
GCACGATACGAAACAACATCTACACCGTGTCTCCATGCTGACACACCACAGCTT
CGCTCCACCTGAATGTGGGCGCATGGGCCCGAATCACAGCCAATGTCGCTGCTGC
CATAATGTGATCCAGACCCTCTCCGCCAGATGCCGAGCGGATCGTGGGCGCTGA
ATAGATTCTGTTTCGATCACTGTTTGGGTCTTTCTTTTCGCTCTCGGATGCGCG
TCTCGAAACAGGCTGCGTCGGGCTTTCGGATCCCTTTTGCTCCCTCCGTCACCATC
CTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACCAGGGTTGGAGGGTATTA
CCGCGTCAGGCCATTCCCAGCCCGGATTCAATTCAAAGTCTGGGCCACCACCCTC
CGCCGCTCTGTCTGATCACTCCACATTCGTGCATACACTACGTTCAAGTCCTGATC
CAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAATCTCAAGGACCCACTCCA
GCACAGCTGCTGGTTGACCCCGCCCTCGCAACTCCCTACC**ATGTCTGCTG**taggtcca
gggatctttgcatgcacacaggaccccggttggtgggggtccccgggtgcatgctgctgctgcagggcgccgggtgtggggcctgggc
cccggggagctcaactcctccccatgctgcccgtccctccccaccgcgacctggcccccttgcag**AGGAAGGCGA**
AGTCAGCGCCATCGTGTGCGATAATGGATCCGG

SEQ ID NO: 156

Endogenous *Chlorella protothecoides* EF1a promoter and 5'UTR

GAATTCGCCCTTGAGTTTAGGTCCAGCGTCCGTGGGGGGGGCGTGAGACTCCCCC
CTGACCTTCGTATGGCAGGGACTCCTACTTGCCAAGTAATCAGTTGACAATGCCA
CTTCAATGCTCGTTGTGGTACACTGACGCGGGTCTAACATACTGGGAAGCATGAA
TTGCCGACATGGACTCAGTTGGAGACAGTAACAGCTCTTTGTGTTCTATCTTCAG
GAACACATTTGGCAGCGCACCCATACAGTGGCGCACACGCAGCTGTACCTGATG
TGGCTCTATTCCACATGTTTCAACTTGATCCAAAAGTCACTCAGACTCTCAGCA
GCTAGACTTGATCGCATCTTTGGCCATGAAGATGCTTGCGCAACTCTAGGAATGG

CTGCGGAGCCTCAGCTGATCTCGCCTGCGAGACCCCACAGTTTGAATTTGCGGTC
CCCACACAACCTCTGACGCC

SEQ ID NO: 159

Acyl ACP desaturase

atcaaaggcatagattcacattgttgccattgcagagcaatcatcgcgaggacgaacatcgctcaccaagcacgtactgggcatcc
ggaggcctccgcaaattctgcaacaggactcgctgatcagttcgcccaagggtctacgacgctccctatcgcgctagacttcaacac
atatctcactgtcacagcctcggcATGCATCAGGCCTCAGTCTCCACCATGAAGACCATCCAGTC
TCGGCACGCCGGTCCCATCGGACATGTGCAGTCGGGTCGCCGATCGGCGGGGCG
CGCGGGATCCCGCATGGCGACCCCCGTGGCCGCAGCTACCGTCGCAGCCCCCTCGC
TCGGCCCTCAACCTCTCCCCCACCATCATTCGACAGGAGGTGCTCCACTCCGCCA
GCGCCCAGCAACTAGACTGCGTGGCCTCCCTGGCGCCCGTCTTCGAGTCCCAGAT
CCTCCCCCTCCTGACGCCCCGTGGACGAGATGTGGCAGCCCACCGACTTCCTCCCC
GCCTCGAACTCGGAGGCATTCTTCGACCAGATCGGCGACCTGCGGGGCGCGATCG
GCGGCCATCCCCGACGACCTGCTGGTCTGCCTGGTGGGGGACATGATCACGGAG
GAGGCCCTGCCACCTACATGGCCATGCTGAACACCCTGGACGTCGTGCGCGATG
AGACAGGGCACAGCCAGCACCCCTACGCCAAGTGGACCAGGGCTTGGATCGCGG
AGGAGAACCGCCATGGCGACCTGCTGAACAAGTACATGTGGCTGACGGGGCGGG
TGGGACATGCTGGCGGTGGAGCGCACCATCCAGCCATGCTGGCGGTGGAGCGCA
CCATCCAGCGCCTCATCTCATCGGGCATGGACCCGGGCACGGAGAACCACCCCT
ACCACGCCTTTGTGTTACACAGCTTCCAGGAGCGCGCCACCAAGCTGAGCCACGG
CTCCACCGCCCGCCTGGCGGTGCGCGCCGGGGACGAGGCCCTGGCCAAGATCTG
CGGGACCATTGCGCGGGACGAGTCGCGCCACGAGGCGGCGTACACGCGGACCAT
GGATGCCATCTTCCAGCGCGACCCCAGCGGGGCCATGGTGGCGTTTGCGCACATG
ATGATGCGCAAGATCACCATGCCCCGCCACCTCATGGACGACGGCCAGCACGGC
GCGCGCAACGGGGGGGCGCAACTTGTTTCGACGACTTTGCGGCAGTGGCGGAGCG
GGCAGGGGTGTACACCGCCGGCGACTACATCGGCATCCTGCGCCACCTCATCCG
GCGCTGGGACGTGGAGGG

SEQ ID NO: 160

Acyl ACP desaturase

MHQASVSTMKTIQSRHAGPIGHVQSGRRSAGRAGSRMATPVAAATVAAPRSALNLS
PTIIRQEVLSASAQQLDCVASLAPVFESQILPLLTPVDEMWPQPTDFLPASNSEAFDQ
IGDLRARSAAIPDDLVLVGDMMITEALPTYMAMLNLTLDVVRDETGHSHQPHYAKW
TRAWIAEENRHGDLLNKYMWLTGRVGHAGGGAHHPAMLAVERTIQRLISSGMDPG
TENHPYHAFVFTSFQERATKLSHGSTARLAVAAGDEALAKICGTIARDESRHEAAYT
RTMDAIFQRDPSGAMVAFAHMMMRKITMPAHLMDDGQHGARNGGAQLVRRLCGS
GGAGRGVHRRRLHRHPAPHPALGRGG

SEQ ID NO: 161

Geranyl geranyl diphosphate synthase

attatacatcgccatcgctcaggtttcacgatctgcatgctatctatgggactgtgactccgcccggccaggttggtgctgcgcgagaatc
ctccccgctcctgctctcatttccctgacggggagtcgcccgtgagcaccggcgccgacATGGGCGTTCGGCACACT
CCAAACCCCATATACATGTGGTTCGTGCATTACGCATAGCGCACGGTATGTCCCG
CGACGCGCGGCTCGAAGCCGTGGCCATCCGACGCGCTGCACGGCCGAGGTGAGG
GCACGCCCCCTCCGCCAATGGCGCGCAGCCCATGACCGCCTTCGACTTCCGGCAGT
ACATGCAGCAGCGCGCCGCGCTGGTGGACGCAGCGCTGGACCTGGCAGTGCCGC

TGCAGTACCCCGAGAAGATCAACGAGGCCATGCGGTACAGCCTGCTGGCCGGGG
GCAAGCGCGTGCGCCCCGCGCTCTGCCTCGCTGCCTGCGAGCTCGTGGGCGGCCC
CCTGGAGGCGGCCATGCCCCGCCGCTGCGCCATGGAGATGATCCACACCATGAG
CCTCATCCACGACGACCTCCCCGCCATGGACAACGACGACTTCCGGCGCGGCCA
GCCCCGCAACCACAAGGCCTATGGCGAGGAGATTGCGATCCTGGCGGGCGACGC
GCTGCTGTCGCTGAGCTTTGAGCACATCGCGCGCGAGACGCGAGGCGTGGACCC
GGTGC GCGTCTTGGCCGCCATCTCGGAGTGGCGCGCGGTGGGCAGCCGCGGGCT
GGTGGCGGGGCAGGTGGTGGACCTGGGTTTCGAGGGCGGCGGCGTGGGGCTGGC
CCCGCTGCGCTACATCCACGAGCACAAAACCGCGGCGCTGCTGGAGGCGGCGGT
GGTGTCCGGCGCGCTGCTGGGCGGCGCGGAGGAGGCGGACCTGGAGCGCCTGCG
CACCTACAACCGCGCCATCGGCCTCGCTTTCAGGTGGTGGGGGACATCCTGGAC
ATCCCGGGGACCAGCGAGGAGCTGGGCAAGACCGCGGGCAAGGACCTGAGCTCC
CCCAAACCCCCCTACCCGTCCCTGGTGGGGCTGGCCAGGTCCAAAAAAATTGCG
GACGAACTGATTGAGGACGCGAAAACCCAACCTACCCAGTACGAGCCGGCCCCGA
GCGGCGCCCCCTCGTAACCCTGGCCGAAAACATTTGAaaccggaagaactgactggggggccccc
ctgccccagatacggcggggctcctccatccagtttgggatgggaggagcgacaaccgaccccgtaaccctgtgacgcgttgc
ttgcatacgtacgcatgccttgaaacccatccatgacctcaacaatacctggttgtgtgtagcttggtcctgaaaaaaaaaaaaaaaa
aaaaaaaaaaaa

SEQ ID NO: 162

Geranyl geranyl diphosphate synthase

MGVGTLPYTCGRAFTHSARYVPRRAARSRGHPTRCTAEVRARPSANGAQPM TAF
DFRQYMQQRAALVDAALDLAVPLQYPEKINEAMRYSLLAGGKRVRPALCLAACEL
VGGPLEAAMPAACAMEMIHTMSLIHDDLPAAMDNDDFRRGQPANHKAYGEEIAILAG
DALLSLSFEHIARETRGVDPVRVLA AIS EWRAVGSRLVAGQVVDLGFEGGGVGLA
PLRYIHEHKTAALLEAAVVS GALLGGAEEADLERLRTYNRAIGLAFQVVGDILDIPGT
SEELGKTAGKDLSSPKTPYPSLVGLARSKKIADELIEDAKTQLTQYEPARAAPLVTLA
ENI

SEQ ID NO: 163

Geranyl geranyl diphosphate synthase

cagatgccATGCGCCCTCGGGCCGCGGGCCTGAGGGTCCACGCAGCGTCCTCGGTGG
CCCAGACGCACCAGGCCGCCCCCGGCGGACAGGAGGTTGACGACTACCAGC
CCCGCACCGCCATCCTCTTCCCCGGCCAGGGCGCGCAGAGCGTGGGCATGGCGG
GAGAGCTGGCGAAGGCCGTCCCCGCCGCGCGGCGCTGTTCGACGCCGCCTCCG
ACCAGCTCGGCTATGACCTGCTCCGCGTGTGCGTTGAGGGCCCCAAGGCGCGCCT
GGACAGCACCGCCGTCAGCCAGCCCGCCATCTACGTGGCCAGCCTGGCGGCGGT
GGAGAAGCTGCGCGCGGAGGGCGGGGAGGAGGCACTGGCCGCCATCGACGTCG
CTGCCGGTCTGTCCTTGGGCGAGTACACCGCGCTGGCCTTTGCCGGCGCCTTCTC
CTTCGCCGACGGGCTGCGCCTGGTGGCCCTGCGCGGCGCCAGCATGCAGGCCGC
CGCCGACGCCGCACCCTCGGGCATGGTCTCCGTTCATCGGTCTGCCCTCCGACGCG
GTGGCCGCGCTGTGCGAGGCGGCCAACGCGCAGGTGGCCCCCGACCAGGCCGTG
CGCATCGCCAACTACCTCTGCGACGGCAACTACGCCGTCAGCGGTGGGCTGGAG
GGCTGCGCGGCGGTGGAGGGCCTGGCCAAGGCCACAAAGGCGCGCATGACGGTG
CGCCTGGCGGTGGCGGGCGCCTTCCACACCCCCCTTCATGCAGCCGGCGGTGGAG
GCGCTGAGCGCGGGCGCTGGCGGACACGCCGCTGGTCGCGCCGCGCATCCCCGT
GGTCAGCAACGGGACGCC

SEQ ID NO: 164

Geranyl geranyl diphosphate synthase

MRPRAAGLRVHAASSVAQTHQAAPPADRRFDDYQPRTAILFPGQGAQSVGMAGEL
AKAVPAAAAALFDAASDQLGYDLLRVCVEGPKARLDSTAVSQPAIYVASLAAVEKLR
AEGGEEALAAIDVAAGLSLGEY TALAFAGAFSFADGLRLVALRGASMQAAADAAPS
GMVSVIGLPSDAVAALCEAANAQVAPDQAVRIANYLCDGNYAVSGGLEGCAAVEG
LAKAHKARMTVRLAVAGAFHTPFMQPAVEALSAGAGGHAAGRAAHPRGQQRDA

SEQ ID NO: 165

Glyceraldehyde 3-phosphate dehydrogenase cDNA sequence

TGTCCATCTCCCCCACCTCCATCCAACCATCGTCGACGGCATGCAGGCGCTGT
GTTCTCACCCCGCGTCCCTCACGGCGCGTGCAGTACCCCATGGGCGGGCCAGCCC
AGCACAGCGGGTGTCCAGCGCCGGCCCGGCCTACACCGGCCTGTCCCGGCACAC
CCTGGGCTGCCCCAGCACCCCCACCCTCCAGTCCCGCGCCGCGGTCCAGACCCGC
GGCTCCTCCTCCGGCTCCACCACGCGCATGACCACCACCGCCCAGCGCAAGATCA
AGGTGGCCATCAACGGGTTTCGGCCGCATCGGCCGCCAGTTCCTGCGCTGCGTGGA
GGGGCGCGAGGACTCGCTGCTGGAGATCGTGGCCGTGAACGACTCCGGCGGGCGT
GAAGCAGGCCAGCCACCTGCTCAAGTACGACTCCACCATGGGCACCTTCAACGC
CGACATCAAGATCTCGGGCGAGGGCACCTTCTCCGTCAACGGCCGCGACATCCG
CGTCGTCTCCTCCCGCGACCCCTGGCCCTGCCCTGGGGCGAGCTGGGCGTGGAC
CTGGTGATCGAGGGGACGGGAGTGTGTTGTGGACCGCAAGGGTGCCAGCAAGCAC
CTGCAGGCGGGGGCCAAGAAGGTCATCATCACCGCGCCGGCCAAGGGCTCCGAC
GTGCCACCTACGTCATGGGCGTGAACGCGGACCAGTACTCCAACCTCCGACGAC
ATCATCTCCAACGCCTCCTGCACCACCAACTGCCTGGCGCCCTTTGTCAAGGTGC
TCAACGACCGCTTCGGCATCGTGA

SEQ ID NO: 166

Glyceraldehyde 3-phosphate dehydrogenase

MQALCSHPASLTARAVPHGRASPAQRVSSAGPAYTGLSRHTLGCPSTPTLQSRRAVQ
TRGSSSGSTTRMTTTAQRKIKVAINGFRIGRQFLRCVEGREDSLLEIVAVNDSGGVK
QASHLLKYDSTMGTFNADIKISGEGTFSVNGRDIRVVSSRDPLALPWGELGVDLVIEG
TGVFVDRKGASKHLQAGAKKVIITAPAKGSDVPTYVMGMVNADQYSNSDDIISNASCT
TNCLAPFVKVLNDRFGIV

SEQ ID NO: 167

Oxygen Evolving Complex protein (OEE33) cDNA sequence

gatgttgagaatagtagcttgccttgcgccatgcagagcgtgtgcgcgcagtcggttcatgaagggggccttaccagtccc
tgcggacccccgatgcagcaggagccagctcgtctgccgggctgatggcaaggccggagccttcatcaagaccgtaagagcgg
tgetgccgtctggctgcctccctcctcctgtctgggggtgcgggcgcactgaccttgcagagctgcagggcctgacctacgtcag
gtgaagggtctggcatcgccaacacctgccccacctgtctggcggctcctccaacatcaaggacctgaagagcgggacctactcc
gtcaacaagatgtgcctggagcccacgtcctcaaggtcaaggaggagcacagttcaagaacggcgaggccgacttgcgccac
caagctcgtcacgcgttgacctacacctggacgagatctctggccagatgaagatcgacggcagcggcggtggagtcaagg
aggaggtggcatcgactatgctgcagtcaccgtgcagcttccgggcggggagcgcgtgcccttctcttaccatcaaggagcttg
acgccaaggggactgccgacggctcaaggcgagttaccgtgccctcctaccgtgggtcctccttctggacccaagggccgc
ggcgctccaccggctacgacaacgccgtggccctgcccgccggcgattccgaggagttggagaaggagaacaacaagtc
accaaggctctgaagggggaggccatcttccatcgccaaggtggacgccgggacaggggaggtggcgggcatcttgagt

SEQ ID NO: 168

Oxygen Evolving Complex protein (OEE33)

MQSVCAQSVSCKGAFTQSLRTPRCRSRLVCRADGKAGAFIKTVKSGAAALAASLL
LSGGAGALTFDELQGLTYLQVKGSGIANTCPTLSGGSSNIKDLKSGTYSVNKMCLEP
TSFKVKEEAQFKNGEADFVPTKLVTRLTYTLDEISGQMKIDGSGGVFKEEDGIDYA
AVTVQLPGGERVPFLFTIKELDAKGTADGFKGEFTVPSYRGSSFLDPKGRGASTGYD
NAVALPAAGDSEELEKENNKSTKALKGEAIFSIKVDAGTGEVAGIFE

SEQ ID NO: 169

Clp protease cDNA sequence

ataatcggaaccagctgcacgcaccatcagtcgggcagcatgcagaccgtcgcagccagctatggcgtattggcgccctccggct
ccagcgtgaccggggctcgaccagcagcaagcagcacttcaccaccctcactccctttccggcttcaggcgctgaatcatgtgga
tcgggcggggcaggcggggctctgggagccccagaccctgcagcaggccgtgggcaaggccgtgcgccgtcggggccgc
accaccagcgccgtgcgcgtgaccgcgatgatttgagcgggtcaccgagaaggccatcaagggtgcatgctcgcgcaggagga
ggctcgccgtctgggccacaacttcgtgggacggagcaaatcctgctggggttgattggggagtcacagggcatgcccgaagg
tctcaagtcgatgggcgtcacgctgaaagatgcgcgtgtggaggtcgagaagatcgcggccggggagcggctttgtggcgtg
gagatcccttcacccccgcgccaagcgtgtgctggagctgtccctggaggaggctcgcagctcggccacaactacattggcac
ggagcacatcctgctgggctgctgcgcgagggtgagggcgtggcctcccgcgtgctggagaccctgggcgcccagccccagaa
gatccgcactcaggtgtacgcatggtgggtgagtcgcaggagcccgtgggcaccacggtgggcggagggtccaccggctccaa
caagatgcccaccctggaggagtacggcaccaacctgaccgccagggcg

SEQ ID NO: 170

Clp protease

MQTVAASYGVLAPSGSSVTRGSTSSKQHFTTLTPFSGFRRLNHVDRAGQAGSGSPQT
LQQAVGKAVRRSRGRTTSAVRVTRMMFERFTEKAIKVVMLAQEEARRLGHNFGVT
EQILLGLIGESTGIAAKVLKSMGVTLKDARVEVEKIIIGRSGGFVAVEIPFTPRAKRVLE
LSLEEARQLGHNHYIGTEHILLGLLREGEQVSRVLETLGADPQKIRTQVVRMVGESQ
EPVGTTVGGGSTGSNKMPTLEEYGTNLTAQA

SEQ ID NO: 171

TCCGGTGTGTTGTAAGTCCA

SEQ ID NO: 172

TTGTCTGGAATGTCATATCAA

SEQ ID NO: 173

AACGCCTTTGTACAACTGCA

SEQ ID NO: 174

CTGACCCGACCTATGGGAGCGCTCTTGCC

SEQ ID NO: 175

CTTGACTTCCCTCACCTGGAATTTGTCG

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
CGGCGCTCGCGTCAGGCACGATGGATGCGGCGCTGCCACCTGTCAATGTACCCC
ACCAGTCTGTGATCGCTACAAGCAACCTTGTGCTCCACATTCCCACCTTGCAGAC
AGTCTAGTTCGATTTTGCCAAGCTGGATGTGAGGATTGGCCATATCTTGGAGGCCA
AGATTCACCCGGATGCTGATGGGTACGTACGCGAGCCAGGCAGGCAGCTGCGTT
GACTTTCTGATTGGCACAAAGCTTTGGCTACTCTCAATACCAACCACGTGCCCT
TCTGCACACCTGCTTCCTTCTGATGACCACTCGCCACGCATGTGCGAGTCTGTACG
TCGAGCAGATCGACCTCGGCGAGGAGGGGGGCCCTCGCACCATCGTGAGTGGCC
TGGTCCGGCACGTGACCCTGGAGGACCTTGTGCGCCGGCGGGTGGTGGTGTCTGG
CCAACCTCAAGCCTCGGAGCATGCGCGGGGTCAAATCGGCTGGGATGCTGCTCT
GCGCCGCCAACGCGGATCACACCGCGGTGGAGCCGCTGCGGGTCCCGGACGCCG
CCGTGACGGGGGAGCGGGTCTGGGCGGGGGACGAGGCACTCCTGTCCACGGAGC
CTGCCACACCCAATCAGGTAAGGACACGTTATTGGTGCGCATGGTGCGAATGCGT
GGTCTGACCTGCTGTGGGTATGTGTTGTGGGATTGGAAACCGAATGAGGGCCGTT
CAGGATTGAGCCCTTGGCCCCACCCTGCTCATCCTCTCACGCCCCGAGGTCCAGA
AGAAGAAAATCTGGGAGGCAGTACAGCCGCTGCTGAGAGTGAACGCCCAGGGG
ATCGTACTGTGGCAGGAGAGGCTATGGTGACCAGTGCGGGGCCACTGACCGCG
CCCACGCTGGTTGACGCCGCGATTTCTGACGCGAGCGACTGATTCTTGACCTTT
GAGAAGCCACCACAGCACCATTTTCATTGTTTCATCCTTGATTTTCAGTACGACTTCT
CACCATTTTCAGTACTGTAGGACCCCCAAAATAGTGTGATCACGCTCGCAAGGCAC
CTGTGTGATCACGGGGAAGGGCGAATTCCTTTCTTTCGCTATGACACTTCCAGCA
AAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGA
TGCTTCGACCCCCCGAAGCTCCTTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCC
AGGGCGAGCGCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGA
GCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACT
CGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCAC
AACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCCTTCCTGTTCCTGCTGGCCGGC
TTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTG

GTGCACTTCACCCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTAC
GACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACC
GTCTGGGGGACGCCCTTGTCTCTGGGGCCACGCCACGTCCGACGACCTGACCAACT
GGGAGGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCT
CCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCAT
CGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGA
GGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCA
GAAGAACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTC
TGGTACGAGCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTAC
AAGATCGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGT
TCGCCAACGAGGGCTTCTTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGT
CCCCACCGAGCAGGACCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAAC
CCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTGCGCAGCTTCAACG
GCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGG
ACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCT
GGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCTTCGTGCCACCAACCCC
TGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGG
CCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCA
GCAACGCCGGCCCCCTGGAGCCGTTTCGCCACCAACACCACGTTGACGAAGGCCA
ACAGCTACAACGTCGACCTGTCCAACAGCACCGGCACCCTGGAGTTCGAGCTGG
TGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTC
CCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGA
GGTGTCCGCGTCTCTCTTCTTCTTCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTG
AAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAACCAGCCCTTCAAG
AGCGAGAACGACCTGTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATC
CTGGAGCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGA
CCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGT
TCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGATTAATTAACCTCGAGGCA
GCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGT
TGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAG
CCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTTCGAGTTGCTAGCTGCTTGTG
CTATTTGCGAATAACCACCCCGAGCATCCCTTCCCTCGTTTCATATCGCTTGCATC
CCAACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTG
CTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTG
CAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAA
CACAAATGGAAAGCTT

SEQ ID NO: 180

Homologous recombination construct SZ726

CCCGTGATCACACAGGTGCCTTTCGAGCGTGATCACACTATTTTGGGGGTCCTAC
AGTACTGAAATGGTGAGAAGTCGTACTGAAATCAAGGATGAACAATGAAAATGG
TGCTGTGGTGGCTTCTCAAAGGTCAAGAATCAGTCGCTCGCGTCAGGAAATCGCG
GCGTCAACCAGCGTGGGCGCGGTCAAGTGGCCCCGCACTGGTCACCATAGCCTCTC
CTGCCACAGTAGCGATCCCCTGGGCGTTCACTCTCAGCAGCGGCTGTACTGCCTC
CCAGATTTTCTTCTTCTGGACCTGCGGGCGTGAGAGGATGAGCAGGGTGGGGCCA
AGGGCTCAATCCTGAACGGCCCTCATTCGGTTTCCAATCCCACAACACATACCCA
CAGCAGGTGAGACCACGCATTCGCACCATGCGCACCAATAACGTGTCCTTACCT
GATTGGGTGTGGCAGGCTCCGTGGACAGGAGTGCCCTCGTCCCCCGCCAGACCC
GCTCCCCCGTACGGCGGCGTCCGGGACCCGCAGCGGCTCCACCGCGGTGTGATC

CGCGTTGGCGGCGCAGAGCAGCATCCCAGCCGATTTGACCCCGCGCATGCTCCG
AGGCTTGAGGTTGGCCAGCACCACCACCCGCCGGCCGACAAGGTCCTCCAGGGT
CACGTGCCGGACCAGGCCACTCACGATGGTGCGAGGGCCCCCTCCTCGCCGAG
GTCGATCTGCTCGACGTACAGACTGCGACATGCGTGGCGAGTGGTCATCAGAAG
GAAGCAGGTGTGCAGAAGGGGACGTGGTTGGTATTGAGAGTAGCCAAAGCTTT
GTGCCAATCAGAAAGTCAACGCAGCTGCCTGCCTGGCTCGCGTACAATTCTTTTC
TTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCG
GCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCTTCGGGGCTGC
ATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCC
GATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCA
CTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCG
CCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAG
GCCTTCCTGTTCCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGA
ACGAGACGTCCGACCGCCCCCTGGTGCACCTTACCCCCAACAAGGGGCTGGATGA
ACGACCCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACT
TCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCTGGGGGCCACGC
CACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCCGAA
GCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACC
TCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGA
CCTACAACACCCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCG
GCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCA
GTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGATCATGACC
GCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAG
TCCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGGCTTCCTCGGCTACCAGTACG
AGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCCAGCAAGTCCTACT
GGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCA
GTACTTCGTTCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAAGTCC
CGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCG
ACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAAGTGGGAGTACT
CCGCCCTTCGTGCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTC
TCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAG
GCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACC
AACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACAGCACC
GGCACCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCA
AGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGA
GTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCCTCCTTCTTCCTGGACCGCGGG
AACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGC
GTGAACAACCAGCCCTTCAAGAGCGAGAACGACCTGTCTACTACAAGGTGTAC
GGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTCGTGT
CCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGAC
GACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCCAGGTGCGCGAGGTCAA
GTGATTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGA
CGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATA
TCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTT
TTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCCTC
CCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATC
CCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCT
CCGCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGC
ACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTCGGTACCCGTA
CCCATCAGCATCCGGGTGAATCTTGGCCTCCAAGATATGGCCAATCCTCACATCC

AGCTTGGCAAAATCGACTAGACTGTCTGCAAGTGGGAATGTGGAGCACAAGGTT
GCTTGTAGCGATCGACAGACTGGTGGGGTACATTGACAGGTGGGCAGCGCCGCA
TCCATCGTGCCTGACGCGAGCGCCGCCGGTTGCTCGCCCGTGCCTGCCGTCAAAG
AGCGGCAGAGAAATCGGGAACCGAAAACGTCACATTGCCTGATGTTGTTACATG
CTGGACTAGACTTTTCTTGGCGTGGGTCTGCTCCTCGCCAGGTGCGCGACGCCTCG
GGGCTGGGTGCGAGGGAGCCGTGCGGCCACGCATTTGACAAGACCCAAAGCTCG
CATCTCAGACGGTCAACCGTTTCGTATTATACATTCAACATATGGTACATACGCAA
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
CGCACGTCGCCGGGAGCGCCAGCTTGATCACTTGGCAGGGGGCCGTGCTCTACA
AATACCAGGCCCGCGCGCGGTTCAGTTCGCACATCCAATACCTGCCGAGCCATCTT
GCCTACACTTTTTATCGACTCCTCTACTCTGTTTCGCGAGAGCGCTCGGTCCAGGCT
TGGAATTCGCCGAATTCAGCTCGATCAGTCGCTTCTGCAACTGATCTCGGCCGTT
CGCAGACTGCCTTTTCTCAGCTTGTGAGGTAGCGAGTTGTTGTTTATATTTATTC
GATTTTCATCTGTGTTGCATGTCTTGTTCGTGCTGTGCGTTCTTTCTGGGCCGCGCT
GTCGGGTGCGATGGGCTAGCTGTACTCATGTTAGTCATGCCGGTCCGACCTTGTT
CGAGGAAGGCCCCACACTGAGCGTGCCCTCTTTCTACACCCCTTGTGCAGAAATT
AGATAGAAAGCAGAATTCTTTCTTTCGCGCTATGACACTTCCAGCAAAAGGTAGG
GCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGAC
CCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGC
GCTGTTTAAATAGCCAGGCCCGCGATTGCAAAGACATTATAGCGAGCTACCAAA
GCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGT
GATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCA
AACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGCCGC
CAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCACCTTC
ACCCCCAACAAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAG
GACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGG
ACGCCCTTGTTCCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGAC
CAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCA
TGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCG
CCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTA
CATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCC
GTGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGC
CCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGA
TCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTTCGCCAACGA
GGGCTTCTTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAG
CAGGACCCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCC
CGGCCGGCGGCTCCTTCAACCAGTACTTCGTGCGCAGCTTCAACGGCACCCACTT
CGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGC
CCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCG
TGGGCCTCCAACCTGGGAGTACTCCGCCTTCGTGCCACCAACCCCTGGCGCTCCT
CCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGA
GACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGG
CCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAA
CGTCGACCTGTCCAACAGCACCCGGCACCCCTGGAGTTTCGAGCTGGTGTACGCCGTC

AACACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCA
AGGGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGT
CCTCCTTCTTCCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACC
CCTACTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACG
ACCTGTCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTA
CTTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGAAC
GCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGAC
AAGTTCAGGTGCGCGAGGTCAAGTGATTAATTAACCTCGAGGCAGCAGCAGCTC
GGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACA
CTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTG
TTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGA
ATACCACCCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAA
CTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCC
CTCGCACAGCCTTGGTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAA
ACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGA
CCGACACGCCCCCGGCCAGGTCCAGTTCTCCTGGGTCTTCCAGAGGGCCCGTCGC
CATGTAAAGTGGCAGAGATTGGCGCCTGATTTCGATTTGGATCCAAGGATCTCCAA
TCGGTGATGGGGACTGAGTGCCCAACTACCACCTTGCCTATCGTCTCGCACT
ATTTATTCCACCTTCTGCTCGCCCTGCCGGGCGATTGCGGGCGTTTCTGCCCTTG
ACGTATCAATTTTCGCCCTGCTGGCGCGAGGATTCTTCATTCTAATAAGAACTCA
CTCCCGCCAGCTCTGTACTTTTCTGCGGGGCCCTGCATGGCTTGTTCCTAATGC
TTGCTCGATCGACGGCGCCCATTTGCCACGGCGCTGCCGCATCCATGTGAAGAAA
CACGGAAGAGTGCGAAGACTGGAAGTGAATTAAGAGTATAAGAAGAGGTACCA
AGGGATTCTCAGGTGCTCTTAGGAACGGCTTTTCTTCGCCAAGAGAACTGCTA
CTGCTCGTGTGCCACGGTGGTCAAGCCGCCCATCTGCGATCCACCAGGCCCCAT
CCGCGGACTCGCGATCAGCCTGCTGGATCCGGACTGCCGACCTGACCGCTCGCAT
CCACCATTACAACCCTCCAATTGGACACCACTCCACGTCCTAAAGTTCACCATG
CAAGCTGATCGATCGCATTCGCCGATGCACTCGCCTGCCACAGAGGTGTGCGCTT
CGGACTAGCGTGACGGCGCCCCGAGGCCACCAGCATGCACCGATGGAAGCGGGC
ACGGCCGCTGCTCCAGGTCGCTGGCTCGCTCAGACCCATAGCAACCTCCGCTGCG
TCCCTAAATGTCACACAGAGCGTCTTTGATGGGTACGGATGGGAGAGAATCTGAT
TGGGCATTGCTGGTGCAGTGCAGGAAGATGGCAAGTGCACAGTCAGTCATGCTG
TACAAACTGGTGCCTCGTAGTATTGACTCGTATAGTGCATAGTATCATGCATGGT
CGTTACTTGCAA

SEQ ID NO: 182

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

TTTGGCCCCGCTTTCCAGCTCCGGATCTGCTGGCGTCCGCCGCGAGACGTGA
CATCGCACGTCGCCGGGAGCGCCAGCTTGATCACTTGGCAGGGGGCCGTGCT
CTACAAATACCAGGCCCCGCGGCGGTGAGTTTCGCACATCCAATACCTGCCGA
GCCATCTTGCCTACACTTTTTATCGACTCCTCTACTCTGTTTCGCGAGAGCGCTC
GGTCCAGGCTTGAATTCGCCGAATTCAGCTCGATCAGTCGCTTCTGCAACT
GATCTCGGCCGTTTCGCAGACTGCCTTTTCTCAGCTTGTGAGGTAGCGAGTTGT
TGTTTTATATTTATTCGATTTTCATCTGTGTTGCATGTCTTGTTCGTGCTGTGCGT
TCTTTCTGGGCCGCGCTGTGCGGTGCGATGGGCTAGCTGTACTCATGTTAGTC
ATGCCGGTCCGACCTTGTTCGAGGAAGGCCCCACACTGAGCGTGCCCTCTTT

CTACACCCCTTGTGCAGAAATTAGATAGAAAGCAATGCTGCTGCAGGCCTTC
CTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACG
AGACGTCCGACCGCCCCCTGGTGCACCTTACCCCCAACAAGGGCTGGATGA
ACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGT
ACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTCTGGGG
CCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCAT
CGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGAC
TACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCT
GCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCT
CCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCG
TGCTGGCCGCCAACTCCACCCAGTTCGCGACCCGAAGGTCTTCTGGTACGA
GCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGAT
CGAGATCTACTCCTCCGACGACCTGAAGTCTTGAAGCTGGAGTCCGCGTTC
GCCAACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAG
GTCCCCACCGAGCAGGACCCAGCAAGTCTACTGGGTGATGTTTCATCTCCA
TCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTCGGCAG
CTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGAC
TTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACCT
ACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACTGGGAGTACTCCGCCTT
CGTGCCCAACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCC
TCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGG
CCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCA
CCAACACCACGTTGACGAAGGCCAACAGTACAACGTCGACCTGTCCAACA
GCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGA
CGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAG
GACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCTCTTCTT
CCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTT
CACCAACCGCATGAGCGTGAACAACCAGCCCTTCAAGAGCGAGAACGACCT
GTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTAC
TTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGA
ACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACAT
CGACAAGTTCCAGGTGCGCGAGGTCAAGTGACCGACACGCCCCCGGCCAG
GTCCAGTTCTCCTGGGTCTTCCAGAGGCCCGTCGCCATGTAAAGTGGCAGAG
ATTGGCGCCTGATTTCGATTTGGATCCAAGGATCTCCAATCGGTGATGGGGAC
TGAGTGCCCAACTACCACCCTTGCACTATCGTCCTCGCACTATTTATTCCCAC
CTTCTGCTCGCCCTGCCGGGCGATTGCGGGCGTTTCTGCCCTTGACGTATCAA
TTTCGCCCCTGCTGGCGCGAGGATTCTTCATTCTAATAAGAACTCACTCCCGC
CAGCTCTGTACTTTTCTGCGGGGCCCTGCATGGCTTGTTCCCAATGCTTGCT
CGATCGACGGCGCCCATTTGCCACGGCGCTGCCGCATCCATGTGAAGAAAC
ACGGAAGAGTGCGAAGACTGGAAGTGAATTAAGAGTATAAGAAGAGGTAC
CAAGGGATTCTCAGGTGCTCTTAGGAACGGCTTTTCTTCGCCAAGAGAAAC
TGCTACTGCTCGTGTGCGCACGGTGGTCAAGCCGCCCCATCTGCGATCCACC
AGGCCCATCCGCGGACTCGCGATCAGCCTGCTGGATCCGGACTGCCGACCTG

ACCGCTCGCATCCACCATTACAACCCTCCAATTGGACACCACTCCCACGTCC
TAAAGTTCACCATGCAAGCTGATCGATCGCATTGCCGATGCACTCGCCTGC
CACAGAGGTGTGCGCTTCGGACTAGCGTGCAGGCGCCCCGAGGCCACCAGC
ATGCACCGATGGAAGCGGGCACGGCCGCTGCTCCAGGTGCTGGCTCGCTC
AGACCCATAGCAACCTCCGCTGCGTCCCTAAATGTCACACAGAGCGTCTTTG
ATGGGTACGGATGGGAGAGAATCTGATTGGGCATTGCTGGTGCAGTGCAGG
AAGATGGCAAGTGCACAGTCAGTCATGCTGTACAACTGGTGCCTCGTAGTA
TTGACTCGTATAGTGCATAGTATCATGCATGGTCGTTACTTGCAA

SEQ ID NO: 183

Cuphea wrightii acyl-ACP thioesterase Genbank ACCESSION U56103

VERSION U56103.1 GI:1336005

MVAAAASSAFFSVPTPGTSPKPGKFGNWPSSLSVFPKPDNGGFH
VKANASAHPKANGSAVNLSKSGSLETPPRSFINQLPDLSMLLSKITTVFGAAEKQWKRP
GMLVEPFVDRIFQDGVFFRQSFIRSIEIGVDRTASIELMNIFQETSLNHCKSIGL
LNDGFGRTPEMCKRDLIWVVTKIQVEVNRYPTWGDITIEVNTWVSESGKNGMGRDWLIS
DCRTGEILIRATSVWAMMNQNTLRSLKFPYEVQRQEIAPHFVDSAPVIEDDRKLHKLDV
KTGDSIRDGLTPRWNDLDVNQHVNNVKYIGWILKSVPIEVFETQELCGVTLEYRRECG
RDSVLESVTTMDPAKEGDRVCVYQHLLRLEDGADITIGRTEWRPKNAGANGAISSGKTS
NGNSVS

SEQ ID NO: 184

Cuphea wrightii acyl-ACP thioesterase Genbank ACCESSION U56104

VERSION U56104.1 GI:1336007

MVAAAASSAFFPVPAPRPTPKPGKFGNWPSSLSQPFKPKSNPN
GRFQVKANVSPHPKANGSAVSLKSGSLNTLEDPPSSPPRPTFLNQLPDWSRLRTAITT
VFVAAEKQFTRLDRKSKRPDMLVDWFGSETIVQDGLVFRERFSIRSIEIGADRTASIE
TLMNHLQDTSLNHCKSVGLLNDGFGRTPEMCTRDLIWVLTKMQIVVNRYPTWGDITVEI
NSWFSQSGKIGMGREWLI SDCNTGEILVRATSAWAMMNQKTRRF SKLPCEVRQEIAPH
FVDAPPVIEDNDRKLHKFDVKTGDSICKGLTPGWNDFDVNQHVSNNVKYIGWILESMPT
EVLETQELCSLTLEYRRECGRESVSVTSMNPSKVGDRSQYQHLLRLEDGADIMKGR
TEWRPKNAGTNRAIST

SEQ ID NO: 185

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

VERSION U67317.1

MAAAASMVASPFCTWLIVASCMSTSFNDNDRSPSVKRFPRRKRVL
SQRGSTFQCLVASCIDPCDQYRSSASLSFLGDNGFASLFGSKPFMSNRGHRRLRRASH
SGEAMAVALQPAQEAGTKKKPVIKQRRVVVTGMGVVTPLGHEPDVFYNNLLDGVSGIS
EIETFDCTQFPTRIAGEIKSFSTDGWVAPKLSKRMDKFMLYLLTAGKKALADGGITDE
VMKELDKRKCGVLIGSGMGGMKVFNDIAEALRVSYKKMNPFCVPFATTNMGSAAMLAM
LGWMGPNYSISTACATSNFCILNAAHHIRGEADMMLCGGSDAVIIPIGLGGFVACRA
LSQRNSDPTKASRPWDSNRDGFVMGEGAGVLLLEELHAKKRGATIYAEFLGGSFTCD
AYHMTPEHPGAGVILCIEKALAQAGVSKEDVNYINAHATSTSAGDIKEYQALARCFG
QNSLRVNSTKSMIGHLLGAAGGVEAVTVVQAIRTGWIHPNLNLEDPDKAVDAKLLVG
PKKERLNVKVGLSNSFGFGGHNSILFAPCNV

SEQ ID NO: 186

Cuphea hookeriana 3-ketoacyl-ACP synthase (Kas4)

Genbank ACCESSION AF060519 VERSION AF060519.2 GI:30995500

MATASCMVASPFCTWLVAACMPTSSDNDPRSLSHKRLRLSRRRR
 TLSSHCSLRGSTFQCLDPCNQQRFLGDNGFASLFGSKPLRSNRGHLRLGRTSHSGEVM
 AVAMQPAQEVSTNKKPATKQRRVVVTGMGVVTPLGHDPDVYYNNLLDGISGISEIENF
 DCSQFPTRIAGEIKSFSTDGWVAPKFSERMDKFMLYMLTAGKKALADGGITEDAMKEL
 NKRKCGVLIGSGLGGMKVFSDSIEALRTSYKKISPFVFPFSTTNMGSAIAMD LGWMG
 PNYSISTACATSNFCILNAANHIKGEADMMLCGGSDAAVLPVGLGGFVACRALSQRN
 NDPTKASRPWDSNRDGFVMGEGAGVLLLEELEHAKKRGATIYAEFLGGSFTCDAYHMT
 EPHPEGAGVILCIEKALAQSGVSREDVNYINAHATSTPAGDIKEYQALAHCFGQNSSEL
 RVNSTKSMIGHLLGGAGGVEAVAVVQAI RTGWIHPNINLEDPDEGVDAKLLVGPKKEK
 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 EK FVIRCYEV GINKTATIET
 IANLLQEVGG NHAQGVGFST DGFATTTTMR KLHLIWTAR MHIEIYRYPA WSDVIEIETW
 VQGEKGKVGTR RDWILKDYAN GEVIGRATSK WMMNEDTRR LQKVSDDVRE EYLVFCPTL
 RLAFPEENNN SMKKIPKLED PAEYSRLGLV PRRSDLDMNK HVNNVTYIGW ALESIPPEII
 DTHELQAITL 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 NTKSPSVNQ P YRTGFHFQPP
 KNWMNDPNGP MIYKGIYHLF YQWNPKGAVW GNIVWAHSTS TDLINWDPHP PAIFPSAPFD
 INGCWGS SAT ILPNGKPVIL YTGIDPKNQ VQNI AEPKNL SDPYLREWK SPLNPLMAPD
 AVNGINASSF RDPTTAWLGQ DKKWRVIIGS KIHRRGLAIT YTSKDFLKWE KSPEPLHYDD
 GSGMWECPDF FVPTRFGSNG VETSSFGE PN EILKHVLKIS LDDTKHDYIT IGTYDRVKDK
 FVPDNGFKMD GTAPRYDYGK YYASKTFEFS AKNRRILGW TNESSSVEDD VEKGWSGIQT
 IPRKIWLDRS GKQLIQWPVR EVERLR TKQV KNLRNKVLKS GSRLEVYGV T AAQADVEVLF
 KVRDLEKADV IEP SWTD PQ L ICSKMNVS VK SGLGPFGLMV LASKNLEEY T SVYFRIFKAR
 QNSNKYVVL M CSDQSRSSLK EDNDKTTYGA FVDINPHQPL SLRALIDHSV VESFGGKGRA
 CITSRVYPKL AIGKSSHLFA FNYGYQSDV LNLNAWSMNS AQIS

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

1 ACTAGTATGG TGGCCTCCGT GGCCGCTCC TCCTCCTTCT TCCCCGTGCC CTCCTCCTCC
 61 TCCTCCGCCT CCGCCAAGGC CTCCGCGGC ATCCCCGACG GCCTGGACGT GCGCGGCATC
 121 GTGGCCAAGC CCGCCTCCTC CTCCGCTGG ATGCAGGCCA AGGCCTCCGC CCGCGCCATC
 181 CCAAGATCG ACGACACCAA GGTGGGCTG CGCACCGACG TGGAGGAGGA CGCCGCCTCC
 241 ACCGCCCCGC GCACCTCCTA CGGGCGCGCC AACCAGCTGC CCGACTGGTC CATGCTGCTG
 301 GCCGCCATCC GCACCATCTT CTCCGCGCC GAGAAGCAGT GGACCCTGCT GGACTCCAAG
 361 AAGCGCGGCG CCGACGCCGT GGCCGACGCC TCCGGCGTGG GCAAGATGGT GAAGAACGGC
 421 CTGGTG TACC GCCAGAACTT CTCCATCCGC TCCTACGAGA TCGGCGTGGA CAAGCGCGCC
 481 TCCGTGGAGG CCCTGATGAA CCACTTCCAG GAGACCTCCC TGAACCACTG CAAGTGCATC
 541 GGCTGATGC ACGGCGGCTT CGGCTGCACC CCCGAGATGA CCCGCCGCAA CCTGATCTGG
 601 GTGGTGGCCA AGATGCTGGT GCACGTGGAG GCCTACCCCT GGTGGGGCGA CGTGGTGCAG
 661 ATCAACACCT GGATCTCCTC CTCCGGCAAG AACGGCATGG GCCGCGACTG GCACGTGCAC

```

721 GACTGCCAGA CCGGCCTGCC CATCATGCGC GGCACCTCCG TGTGGGTGAT GATGGACAAG
781 CACACCCGCC GCCTGTCCAA GCTGCCCGAG GAGGTGCGCG CCGAGATCAC CCCCTTCTTC
841 TCCGAGCGCG ACGCCGTGCT GGACGACAAAC GGCCGCAAGC TGCCCAAGTT CGACGACGAC
901 TCCGCCGCCG ACGTGCGCCG CGGCCTGACC CCCCCTGGC ACGACTTCGA CGTGAACCAG
961 CACGTGAACA ACGTGAAGTA CGTGGGCTGG ATTCTGGAGT CCGTGCCCGT GTGGATGCTG
1021 GACGGCTACG AGGTGGCCAC CATGTCCCTG GAGTACCGCC GCGAGTGCCG CATGGACTCC
1081 GTGGTGCACT CCCTGACCGC CGTGTCTCTC GACCACGCCG ACGGCTCCCC CATCGTGTGC
1141 CAGCACCTGC TCGCCTGGA GGACGGCACC GAGATCGTGC GCGGCCAGAC CGAGTGGCGC
1201 CCAAAGCAGC 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 AWWSGADSGI
VVASPDNDP DYFYTWTRDS GLVLKTLVDL FRNGDTSLLS TIENYISAQA IVQGISNPSG
DLSSGAGLGE PKFNVDETAY TGSWGRPQRD GPALRATAMI GFGQWLLDNG YTSTATDIVW
PLVRNDLSYV AQYWNQTGYD LWEEVNGSSF FTIAVQHRAL VEGSAFATAV GSSCSWCDSQ
APEILCYLQS FWTGSFILAN FDSSRSRKDA NTLLGSIHTF DPEAACDDST FQPCSPRALA
NHKEVVDSEFR SIYTLNDGLS DSEAVAVGRY PEDTYYNGNP WFLCTLAAAE QLYDALYQWD
KQGSLEVTDV SLDFFKALYS DAATGTYSST SSTSIVDA VKTFADGFVS IVETHAASNG
SMSEQYDKSD GEQLSARDLT WSYAALLTAN NRRNSVVPAS WGETSASSVP GTCAATSAIG
TYSSVTVTWS PSIVATGGTT TTATPTGSGS VTSTSKTTAT ASKTSTSTSS TSCTTPTAVA
VTFDLTATTT YGENIYLVGS ISQLGDWETS DGIALSADKY TSSDPLWYVT VTLPAGESFE
YKFIRIESDD SVEWESDPNR EYTVPQACGT STATVTDWTR

```

SEQ ID NO: 191

Rhizopus oryzae (Rhizopus delemar) Glucoamylase 1;

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

VERSION P07683.2

```

MQLFNLPLKV SFFVLVLSYFS LLVSAASIPS SASVQLDSYN YDGSTFSGKI YVKNIAYSCK
VTVIYADGSD NWNNGNTIA ASYSAPISGS NYEYWTFSAS INGIKEFYIK YEVSCKTYD
NNNSANYQVS TSKPTTTTAT ATTTTAPSTS TTPPSRSEP ATFTGNSTI SSWIKKQEGI
SRFAMLRNIN PPGSATGFIA ASLSTAGPDY YYAWTRDAAL TSNVIVYEYN TTLSGNKTIL
NVLKDYVTFV VKTQSTSTVC NCLGEPKFNPD DASGYTGAWG RPQNDGPAER ATTFILFADS
YLTQTKDASY VTGTLKPAIF KDLDYVNVW SNGCFDLWEE VNGVHFYTLM VMRKGLLLGA
DFAKRNQDST RASTYSSTAS TIANKISSFW VSSNNWQVS QSVTGGVSKK GLDVSTLLAA
NLGSVDGDFP TPGSEKILAT AVAVEDSFAS LYPINKNLPS YLGNSIGRYP EDTYNGNGNS
QGNISWFLAVT GYAELYRAI KEWINGGGVT VSSISLPFFK KFDSSATSGK KYTVGTSDFN
NLAQNIALAA DRFLSTVQLH AHNNGSLAEE FDRTTGLSTG ARDLTWSHAS LITASYAKAG
APAA

```

SEQ ID NO: 192

Chlamydomonas reinhardtii hydroxymethylpyrimidine phosphate synthase, Genbank accession no.

XP_001697756, VERSION XP_001697756.1

```

MQLFARSPAG VRVQQAQRK SAKVAGRGR LNVRAEAIAMP PAPAEQEKEL QRLEKGNAFE
ELKAMATSPR QSVNRPPQAE DLTFRQAPTI ADCFPDSEKM YKEVKYDDEI TLRVPFRRIH
LTTGTDFDVY DTSQPNQNDP RVGLPKLRKP WVERREANGT HARAGVTQMA LARAGIISEE
MRFAAEREGL RPEFVRELA RGRAIIPANK CHLELEPCVI GRNFMKTVNS NFGNSAVTSS
IEEEVEKLQW STIWGADTVM DLSTGHNIFE TREWVMRNSP VPGTVPIYE ALERADGQVE
GITWELFRQV LLDQAEQGVY YWTIHAGVLL RHVPLTANRI TGIVSRGGS I HAKLCLMEHK
ENFAYEHWDE ILDICAKYDI TLSIGDGLRP GCIADANDAA QFAELKTQGE LTRRAWKENV
QVMNEGPGHV PLNKIPENMA KQLEWCSEAP FYTLGLPLTD IAPGYDHITS AIGAATIGAL

```

GTALLCYVTP KEHLGLPNRD DVKAGVIAYK IAAHAADLAK GHPAAASWDL ELSKARFEFR
WRDQFALSLD PVTAQAYHDA TLPQEPAKTA HFCSMCGPKF CSMNITQELR QMVQAEQAQE
AAAAAAAADD ADLLAAAAEG 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
LRHQHPGMNF GARLPENITP EFVRDEVAAG RAIIPANINH PESEPMIIGR NFLVKVNANI
GNSAVTSSIE EEVEKLVWST RWGADTVMDL STGRYIHETR EWILRNSPVP IGTVPYIYQAL
EKVNGIAEDL TWEAFRD TLL EQAEQGVDFY TIHAGVLLRY VPMTAKRLTG IVSRGGSIMA
KWCLSHHKEN FLFEHFREIC EICAAVDVSL SLGDGLRPGS IQDANDEAQF SELHTLGELT
KIAWEYDVQV MIEGPGHVPM HMIQRNMTEE LESCHEAPFY TLGPLTTDIA PGYDHFTSGI
GAAMIGWFGC AMLCYVTPKE HLGLPNKEDV KQGLITYKIA AHAADLAKGH PGAQIRDNAM
SKARFEFRWE DQFNALDPF TARAYHDETL PQESGKVAHF CSMCGPKFCS MKISQEVDRDY
AAAQTIEIGM ADMSEDFRAK GGEIYLKREE A

SEQ ID NO: 194

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

MIEQDGLHAG SPAAWVERLF GYDWAQQTIG CSDAAVFRSL AQGRPVLFVK TDLSGALNEL
QDEAARLSWL ATTGVPCAAV LDVVTEAGRDL WLLLGEPVPGQ DLLSSHLAPA EKVSIAMADAM
RRLHTLDPAT CPFQDHQAKHR IERARTRMEA GLVDQDDLDE EHQGLAPAE FARLKARMPD
GEDLVVTHGD ACLPNIMVEN GRFSGFIDCG RLGVDADRYQD IALATRDIAE ELGGEWADRF
LVLYGIAAPD SQRIAFYRL DEFF

SEQ ID NO: 195

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

MLKLSCNVTN NLHTFSFFSD SSLFIPVNR TIAVSSSQLR KPALDPLRAV ISADQGSISP
VNSCTPADRL RAGRLMEDGY SYKEKFIVRS YEVGINKTAT VETIANLLQE VACNHVQKCG
FSTDGFATTL TMRKLHLIIV TARMHIEIYK YPAWSDVVEI ETWCQSEGRI GTRRDWILRD
SATNEVIGRA TSKWVMNQD TRRLQRV TDE VRDEYLVFCP REPRLAFPEE NNSSLKKIPK
LEDPAQYSML ELKPRRADLD MNQHVNNVTY IGWVLESIPQ EIIDTHELVQ ITLDYRRECQ
QDDIVDSLTT SEIPDDPI SK FTGTNGSAMS SIQGHNESQF LHMLRLSENG QEINRGRTQW
RKKSSR

SEQ ID NO: 196

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

MVATAATSSF FPLTSPSGDA KSGNPGKGSV SFGSMKSKSA ASSRGLQVKA NAQAPTKING
STDDAQLPAP RTFINQLPDW SMLLAAITTV FLAAEKQWMM LDWKPRRPDM LIDTFGLGRI
VQDGLVFRQN FSIRS YEIGA DRTAS IETVM NHLQETALNH VKTAGLLGDG FGSTPEMSKR
NLIWVVTM Q VEVD RYPTWG DVVQVD TWVS ASGKNGMRRD WLLRDGNTGE TLTRASSVWV
MMNKLTRRLS KIPEEVREEI GSYFVN SDPV VEEDGRKVTK LDDNTADFVR KGLTPKWN DL
DINQHVNNVK YIGWILESAP QPILE TRELS AVTLEYRREC GRDSVLRLSLT 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 PPRTFLHQLP DWSRLLTAIT TVFVKSKRPD MHDRKSKRPD
MLVDSFGLES TVQDGLVFRQ SFSIRSIEIG TDRTASIELT MNHLQETSLN HCKSTGILLD
GFGRTLEMCK RDLIWVVIKM QIKVNRYPAP GDTVEINTRF SRLGKIGMGR DWLISDCNTG
EILVRATSAY AMMNQKTRRL SKLPYEVHQE IVPLFVDSPV IEDSDLKVHK FKVKGTGDSIQ
KGLTPGWNDL DVNQHVSNVK YIGWILESMP TEVLETQELC SLALEYRREC GRDSVLESVT
AMDPSKVGVR SQYQHLLRLE DGTAVNGAT EWRPNKAGAN GAISTGKTSN GNSVS
```

SEQ ID NO: 198

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

```
MLKVPCCNAT DPIOQLSSQC RFLTHFNNRP YFTRRPSIPT FFSSKNSSAS LQAVVSDISS
VESAAACDCLA NRLRLGKLTE DGFYSYKEFI VRSYEVGINK TATVETIANL LQEVGCNHAQ
SVGFSTDGFA TTSMRKMHL IWVTARMHIE IYKPAWSDV VEVETWCQSE GRIGTRDWI
LTDYATGQII GRATSKWMM NQDTRRLQKV TDDVREEYLV FCPRELRLAF PEENNRSSKK
ISKLEDPAQY SKLGLVPRRA DLDMNQHVNN VTYIGWVLES IPQEIIDTHE LQTITLDYRR
ECQHDDIVDS LTSVEPSENL EAVSELRGTN GSATTAGDE DCRNFLHLLR LSGDGLLEINR
GRTEWRKKSAR
```

SEQ ID NO: 199

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

```
FRQITKNQKK KVRKKTMAK LNPFLSQTK LPSFALPPMA STRSPKFYMA STLKSGSKEV
ENLKKPFMPP REVHVQVTHS MPPQKIEIFK SLDNWAEENI LVHLKPVEKC WQPQDFLPDP
ASDGFDEQVR ELRERAKEIP DDYFVVLVGD MITEEALPTY QTMLNTLDGV RDETASPTS
WAIWTRAWTA EENRHGDLN KYLYLSGRVD MRQIEKTIQY LIGSGMDPRT ENSPYLGFYI
TSFQERATFI SHGNTARQAK EHGDIKLAQI CGTIAADEKR HETAYTKIVE KLFEIDPDGT
VLAFAADMMRK KISMPAHLMY DGRDDNLFDF FSAVAQRLGV YTAQDYADIL EFLVGRWKVD
KLTGLSAEGQ KAQDYVCRLP PRIRRLEERA QGRAKEAPTM PFSWIFDRQV KL
```

SEQ ID NO: 200

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

```
MALKLCFPPH KMPSPFDARI RSHRVFMAS IHSPSMEVGK VKKPFPTPRE VHVQVTHSLA
PEKREIFNSL NNWAQENILV LLKDVDKWCQ PSDFLPDSAS EGFDEQVMEL RKRCKEIPDD
YFIVLVGDMI TEEALPTYQT MLNTLDGVRD ETGASLTPWA IWTRAWTAEE NRHGDLNLY
LYLSGRVDMK 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 SLSPKPKST PNGGLQVKAN ASAPPKINGS
PVGLKSGGLK TQEDAHSAAP PRTFINQLPD WSMLLAIAIT VFLAAEKQWM MLDWKPKRPD
MLVDPFGLGS IVQDGLVFRQ NFSIRSIEIG ADRTASIELT MNHLQETALN HVKIAGLSND
```

GFGRTPEMYK RDLIWVVAKM QVMVNRYPTW GDTVEVNTWV AKSGKNGMRR DWLISDCNTG
 EILTRASSVW VMMNQKTRRL SKIPDEVNRNE IEPHFVDSPP VIEDDDRKLKLP KLDEKTADSI
 RKGLTPRWND LDVNQHVNNV KYIGWILEST PPEVLETQEL CSLTLEYRRE CGRESVLES
 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 WSMLLAAITT VFLLAAEQWM MLDWKPKRPD
 MLVDPFGLGS IVQHGLVFRQ NFSIRSIEIG ADRTASIEIV MNHLQETALN HVKSAGLMND
 GFGRTPEMYK KDLIWVVAKM QVMVNRYPTW GDTVEVNTWV AKSGKNGMRR DWLISDCNTG
 EILTRASSVW VMMNQKTRRL SKIPDEVNRNE IEPHFVDSPP VIEDDDRKLKLP KLDDKTADSI
 RKGLTPRWND LDVNQHVNNV KYIGWILEST PPEVLETQEL CSLTLEYRRE CGRESVLES
 TAADPSGKGF GSQFQHLLRL EDGGEIVKGR TEWRPKTAGI NGAIPSGETS PGDS

SEQ ID NO: 203

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

MVTTSLASAF FSMKAVMLAP DSGSIKPRSS GLQVRAGKEQ NSCKMINGTK VKDTEGLKGR
 STLHGWSMPL ELITTIFSAA EKQWTNLVSK PPQLLDDHLG LHGLVFRRTF AIRCSEVGPD
 RSTSIVAVMN YLQEAACNHA ESLGLLGDGF GETLEMSRRD LIWVVRTHV VVERYPAWGD
 TVEVEAWIGA AGNIGMRRHF LVRDCKTGHI LARCTSVSVM MNMRTRRLSK IPQEVGEID
 PLFIEKFAVK EGEIKKLQKF NDSTADYIQG GWTPRWNLDD VNQHVNNIKY VGWIFKSVPD
 SIYENHLLSS ITLEYRRECT RGRALQSLTT VCGGSSEAGI ICEHLLQLED GSEVLRGRD
 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 MQGKVMAQAV
 SKINGAKVGL KAESQKAEED AAPSSAPRTF YNQLPDWSVL LAAVTTIFLA AEKQWTLLDW
 KPRRPDMLTD AFSLGKIVQD GLVFKQNFSE RSYEIGADRT ASIETLMNHL QETALNHVRS
 AGLMGDGFGA TPEMSKRNL I WVTKVRLV EHYPWGDV EVDTWVGPA KNGMRDWHV
 RDHRTGQTIL RATRVWVMN KNTRKLSKVP EEVRAEIGPY FVERAAIVDE DSRKLPKLDE
 DTTDYVKGL 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 KQGKVMAQAV
 PKINGAKVGL KAESQKAEED AAPSSAPRTF YNQLPDWSVL LAAVTTIFLA AEKQWTLLDW
 KPRRPDMLTG AFSLGKIVQD GLVFRQNFSE RSYEIGADRT ASIETLMNHL QETALNHVRN
 AGLMGDGFGA TPEMSKRNL I WVTKMQLV EHYPWGDV EVDTWVGASG KNGMRDWHV
 RDYRTGQTIL RATSIWVMM 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 AAEKQWTLTD WKPRRPDMLA DAFGLGKIVQ DGLVFKQNF'S
IRSYEIGADR TASIETLMNH LQETALNHVR SAGLMGDGFG ATPEMSKRNL IWVVTKMRVL
IKHYPSWGDV VEVDTWVGPT GKNMRRDWH VRDHRTGQTI LRATSVWVMM NKNTRKLSKV
PEEVRAELGP YFVERAAIVD EDSRKLPKLD EDTTDYIKKG LTPRWSDLVD 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 FVPSPPPGDA KSSKVGSGSA SLGGIKSKSA SSGALQVKAN AQAPPKINGS
PVGLTASVET AKKEDVVSSP APRTFINQLP DWSMLLAIT TMFLAAEKQW MMLDWKPKRA
DMLIDPFGIG RIVQDGLVFS QNFSIRSIEI GADRTASIEI LMNHLQETAL NHVKTAGLLG
DGFGSTPEMS KRNLIVVVTR MQILVDRYPT WGDVVHVDTW VSASGKNGMR RDWLVRDAKT
GETLTRASSL WVMMNKVTRR LSKIPEDVRG EIEPYFLNSD PVVNEDSTKL PKLDDKTADY
IRKGLTPRWV DLDVNQHVNN VKYIGWILES APPPILESHE LAAITLEYRR ECGRDSVLQS
LTAVSGAGIG NLGGPGKVEC QHLLRHEDGA EIVRGRTWR 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 FVPSSSLDP NGKGNKIGST NLAGLNSAPN SGRMKVKPNA QAPPKINGKK
VGLPGSVDIV RTDTETSSHP APRTFINQLP DWSMLLAIT TIFLAAEKQW MMLDWKPRRS
DMLVDPFGIG RIVQDGLVFR QNFSIRSIEI GADRSASIEI VMNHLQETAL NHVKTAGLLG
DGFGSTPEMF KKNLIWVVTR MQVVVDKYPT WGDVVEVDTW VSQSGKNGMR RDWLVRDCNT
GETLTRASSV WVMMNKLTRR LSKIPPEVRG 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 FVPSSSLDP NGKGNKIGST NLAGLNSAPN SGRMKVKPNA QAPPKINGKR
VGLPGSVDIV RTDTETSSHP APRTFINQLP DWSMLLAIT TIFLAAEKQW MMLDWKPRRS
DMLVDPFGIG RIVQDGLVFR QNFSIRSIEI GADRSASIEI VMNHLQETAL NHVKTAGLLG
DGFGSTPEMF KKNLIWVVTR MQVVVDKYPT WGDVVEVDTW VSQSGKNGMR RDWLVRDCNT
GETLTRASSV WVMMNKLTRR LSKIPPEVRG 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 FVPSSSLDP NGKGNKIGST NLAGLNSTPN SGRMKVKPNA QAPPKINGKR
VGLPGSVDIV RTDTETSSHP APRTFINQLP DWSMLLAIT TIFLAAEKQW MMLDWKPRRS
DMLVDPFGIG RIVQDGLVFR QNFSIRSIEI GADRSASIEI VMNHLQETAL NHVKTAGLLG
DGFGSTPEMF KKNLIWVVTR MQVVVDKYPT WGDVVEVDTW VSQSGKNGMR RDWLVRDCNT
```

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

SEQ ID NO: 211

Gossypium hirsutum Genbank ACCESSION Q9SQI3

MVATAVTSFAF FTVTSSPDSS DSKNKKLGS I KSKPSVSSGS LQVKANAQAP PKINGTVAST
 TPVEGSKNDD GASSPPPTF INQLPDWSML LAATTITFLA AEKQWMLDW KPRRPDMVID
 PFGIGKIVQD GLVFSQNF SI RSYEIGADQT ASIETLMNHL QETAINHCRS AGLLGEGFGA
 TPMECKKNLI WVVTRMQVVV DRYPTWGDV QVDTWVSASG KNGMRRDWLV SNSETGEILT
 RATSVMVMN KLTRRLSKIP EEVRGEIEPF FMNSDPVLAE DSQKLVKLDD STAEHVCKGL
 TPKWSDLDVN QHVNNVKYIG WILESAPLPI LESHELSALT LEYRRECGRD SVLQSLTTVS
 DSENTENAVNV 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 WSMLLAATT VFLAAEKQWM MLDWKPKRPD
 MLVDPFGLGS IVQGGVFRQ NFSIRSYEIG ADRTASIE TV MNHLQETALN HVKSAGLLND
 GFGRTPPEMFK RDLIWVAKM QVMVNRYP TW GDTVEVNTWV AKSGKNGMRR DWLISDCNTG
 EILTRASSV VMNQKTRKL SKIPDEV RHE IEPHFIDCAP VIEDDDRKL R KLDEKTADSI
 RKGLTPKWN DLDVNQHVNN 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 WSMLLAATT VFLAAEKQWM MLDWKPKRPD
 MLVDPFGLGR IVQDGLVFRQ NFSIRSYEIG ADRTASIE TV MNHLQETALN HVKSAGLLND
 GFGRTPPEMYK RDLIWVAKM QVMVNRYP TW GDTVEVNTWV AKSGKNGMRR DWLISDCNTG
 EILTRASSV VMNQKTRRL SKIPDEV RHE IEPHFVDSAP VIEDDDRKL R KLDEKTADSI
 RKGLTPKWN DLDVNQHVNN 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 QWMLDWKPR RPDVIDPFG
 IGKIVQDGLV FSQNF SIRS Y EIGADQTASI ETLMNHLQET AINHCRSAGL LGEGFGATPE
 MCKKNLIWV TRMQVVVD RY PTWGDVQVD TWVSASGKNG MRRDWLVNS ETGEILTRAT
 SVWVMNKL T RRLSKIPEEV RGEIEPFMN SDPVLAE DSQ KLVKLDDSTA EHVCKGLTPK
 WSDLDVNQHV NNVKYIGWIL ESAPLPILES HELSALTLEY RRECGRDSVL QSLTTVSDSN
 TENAVNVGEF NCQHLLRLDD GAEIVRGRTR WRPKHAKSSA NMDQITAKRA

SEQ ID NO: 215

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

MVATAATSAF	FAVASPSSDP	DAKPSTKPGX	GSAILRGIKS	RNAPSGSLQV	KANAQAPPKI
NGTTVGYTSS	AEGVKIEDDM	SSPPPRTFIN	QLPDWSMLLA	AITTIFLAAE	KQWMLLDWKP
RRSDMLIDPF	GLGKIVQDGL	VFRQNFSIRS	YEIGADRTAS	IETLMNHLQE	TALNHVRTAG
LLGDGFGSTP	EMSIRNLIWV	VTRMQVVVDR	YPTWGDVVQV	DTWVCASGKN	GMRRDWIIRD
CKTGETLTRA	SSVWVMNMQ	TRRLSKIPDA	VRAEIEPYFM	DSAPIVDEDG	RKLPKLDDST
ADYIRTGLTP	RWSDLDVNQH	VNNVKYIGWI	LESAPLPILE	SHELSSMTLE	YRRECGRDSV
LQSLTAVSGT	GVGNLLDCGN	VECQHLLRLE	EGAEIVEGKD		

SEQ ID NO: 216

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

MVATSATSFL	FPLPSSSLDP	NXKTNNRVTS	TNFAGLGPTP	NSGGRMKVKP	NAQAPPKING
KKVGLPGSVE	IETSQQQQA	PRTFINQLPD	WSMLLAAITT	VFLAAEKQWM	MLDWKPRRS
MIMEPFGLGR	IVQDGLVFRQ	NFSIRSYEIG	ADRSASIVTV	MNHLQETALN	XVKTAGLLGD
GFGSTPEMVK	KXLIWVVTRM	QVVVDITYPTW	GDVVEVDTWV	SKSGKNGMRR	DWLVRDGNITG
QILTRASSVW	VMMNKLTRRL	SKIPEEVRGE	IEPYFVDFDP	VLAEDSRKLT	KLDDKTADYV
RSGLTWRSD	LDVNQHVNNV	KYIGWILESA	PVGMMESQKL	KSMTLEYRRE	CGRDSVLQSL
TAVSGCDIGN	LGTAGEVECC	HLLRLQDGAE	VVRGRTEWSS	KTPTTTWDIT	P

SEQ ID NO: 217

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

INQLPDWSML	LAAITTIFLA	AEKQWMLLDW	KPKRPDMLID	PFGLGKIVQD	GLVFRQNFSI
RSYEIGADRT	ASIVETLMNHL	QETALNHVKT	AGLLGDGFGV	TPEMCKKNLI	WVAKMQVLV
DRYPTWGDVV	QVDTWVAASG	KNGMRDRLV	RDFETGDILT	KASSVWVMN	RETRRLSKIP
DEVRLKIGSY	FVDSPPVLDE	DGRKLPLDE	STADHIRTGL	TPRWNDLDVN	QHVNNVKYIG
WILESAPQPI	LENHELASMT	LEYRRECCKD	SVLQSLTGVT	SGGVGGLADS	GHVECHLLR
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	EETVPYSAPK	TFYNQLPDWS	MLLAAITTIF	LAAEKQWTL
DWKPKKPDML	VDTFGFGRII	QDGMVFRQNF	MIRSYEIGAD	RTASIVETLMN	HLQETALNHV
RTAGLLGDGF	GATPEMSKRN	LIWVVSQIQL	LVEQYPAWGD	TVQVDTWVAA	AGKNGMRD
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	EETVPYSAPK	TFYNQLPDWS	MLLAAITTIF	LAAEKQWTL
DWKPKKPDML	VDTFGFGRII	QDGMVFRQNF	MIRSYEIGAD	RTASIVETLMN	HLQETALNHV
RTAGLLGDGF	GATPEMSKRN	LIWVVSQIQL	LVEQYPAWGD	TVQVDTWVAA	AGKNGMRD
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
LPKVNGTKVN LKTTSPDKED IIPYTAPKTF YNQLPDWSML LAAVTTIFLA AEKQWTLDDW
KPKKPDMLAD TFGFGRIIQD GLVFRQNFLI RSYEIGADRT ASIETLMNHL QETALNHVKT
AGLLGDGFGA TPMSKRNL I WVSQIQLLV ERYPSWGMV QVDTWVAAAG KNGMRRDWHV
RDYNSGQTIL RATSVMVMN KNTRRLSKMP DEVRAEIGPY FNGRSAISEE QGEKLPKPGT
TFDGAATKQF TRKGLTPKWS DLDVNQHVNN VKYIGWILES APISILEKHE LASMTLDYRK
ECGRDSVLQS LTTVSGECDD GNTESSIQCD 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 GADRTASLET LMNHLQETSI NHCKSLGLLN
DGFGRTPGMC KNDLIWVLTK MQIMVNRYP T WGDTEINTW FSQSGKIGMA SDWLISDCNT
GEILIRATSV WAMMNQKTRR FSRLPYEVRQ ELTPHFVDSP HVEDNDQKL HKFDVKTGDS
IRKGLTPRW N DLDVNQHVS N VKYIGWILES MPEVLETQE LCSLTVEYRR ECGMDSVLES
VTAVDPSENG GRSQYKHLR LEDGTDIVKS RTEWRPKNAG TNGAISTSTA KTSNGNSAS

SEQ ID NO: 222

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

AAB71730.1

PDWSMLLAAI TTIFLAAEKQ WTNLDWKPRR PDMLVDFDPF SLGRFVQDGL IFRQNFIRS
YEIGADRTAS IETLMNHLQE TALNHVRCIG LLDDGFGSTP EMTRDLI WV VTRMQVLVDR
YPSWGDVIEV DSWVTPSGKN GMKREWF LRD CKTGEILTRA TSVWMMNKR TRRLSKIPEE
VRVEIEPYFV EHGVLDEDSR KLPKLNDNTA NYIRRLAPR WSDLDVNQHV NNKYIGWIL
ESVPSSLLES HELYGMTLEY RKECGKDGLL QSLTAVASDY GGSLEAGVE CDHLLRLEDG
SEIMRGKTEW RPKRAANTTY FGSVDDIPPH PIYIYIYIYI YIYIYWGSS CSGSSTMSR
TR

SEQ ID NO: 223

Cinnamomum camphora acyl-ACP thioesterase. Genbank ACCESSION U31813 VERSION

U31813.1

MATTSLASAFCSMKAVMLARDGRGMKPRSSDLQLRAGNAQTSLK
MINGTKFSYTESLKKLPDWSMLFAVITTTIFSAEKQWNTNLEWKPKPNPPQLDDHFGP
HGLVFRRTFAIRSYEVGPDRSTSIVAVMNLHQAALNHAKSVGILGDGFGTTLEMSKR
DLIWWVKRTHVAVERYPAGDTEVEECWVGASGNNRRHDFLVRDCKTGEILTRCTSL
SVMNTRTRRLSKIPEEVRGEIGPAFIDNVAVKDEEIKKPQKLNSTADYIQGGLTPR
WNDLDINQHVNNIKYVDWILETVPDSIFESHHSFTIEYRRECTMDSVLQSLTTVSG
GSSEAGLVCEHLLQLEGGSEVLRAKTEWRPKLTDSEFRGISVIPAESSV

SEQ ID NO: 224

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

VERSION AAB71729.1

MVATSAASAF FPVASPSPVK PSMMLGGGGG SDNLDARGIK SRPASSGGLQ VKANAHTVPK
INGNKAGLLT PMESTKDEDI VAAPTVAAPKR TFINQLPDWS MLLAAITTIF LAAEKQWTNL
DWKPRRPDML VDFDPFSLGR FVQDGLIFRQ NFSIRSYEIG ADRTASIELT MNHLQETALN
HVRICIGLLDD GFGSTPEMTR RDLIWVVTM QVLVDRYPSW GDVIEVDSWV TPSGKNGMKR
EWFLRDCKTG EILTRATSVW VMMNKRTRRL SKIPEEVRVE 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 GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAGAAA GAGTTGGTAG CTCTTGATCC GGCAACAAA
181 CCACCGCTGG TAGCGGTGGT TTTTGTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGACT CCCCGTCGTG TAGATACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTGGTATG GCTTCATTCA
781 GCTCCGTTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGTCTCCT CGGTCTCTCC ATCGTTGTCA GAAGTAAGTT GGCCGCGAGT TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCAATGCC ATCCGTAAGA TGCTTTTCTG
961 TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCCGC GTCAATACGG GATAATACCG GCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC 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 CTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGTCCCCG AGACGGTCAC AGCTTGTCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGGTG CAGCGGGTGT TGCGGGTGTG CGGGGCTGGC
1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT CAGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGATGCTC TTCGCCGCC CCACTCCTGC TCGAGCGCGC
1921 CCGCGGTGTC GCCGCCAGCG CTTGGCCTT TTCGCCGCG TCGTGCGCGT CGCTGATGTC
1981 CATACCAGG TCCATGAGGT CTGCCTTGCG CCGGCTGAG CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGT CAGGGTCTG ACGTGGTCG GGTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAACCTGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCTT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGCTTA GGCAGAATCC CTACCAGTCA TGGCTTTACC TGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCTCGCT GCCGCCGCTT CTCCGCACG CTTCTTTCCA
2341 GCACCGTGAT GCGCGAGGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCC ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGCCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTACCTTT TTGGCGAAG CATCGGCCTC GGCTGCGAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTT CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCG GGTGTCATGG GCGCTCCGAT GCCGCTCCAG GCGAGCGCT

2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCCTCT TCGTTTCAGT CACAACCCGC AAATCTTAGA ATATCAATGC
2941 TGCTGCAGGC CTTCCCTGTTT CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCAAACGG CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCCGAACGA CACCGTCTGG GGGACGCCCT TGTTCGTTGGG CCACGCCACG TCCGACGACC
3181 TGACCAACTG GGAGGACCAG CCCATCGCCA TCGCCCCGAA GCGCAACGAC TCCGGCGCCT
3241 TCTCCGCTC CATGGTGGTG GACTACAACA ACACCTCCGG CTCTTCAAC GACACCATCG
3301 ACCCGCGCCA GCGCTGCGTG GCCATCTGGA CCTACAACAC CCCGGAGTCC GAGGAGCAGT
3361 ACATCTCCTA CAGCCTGGAC GCGGCTTACA CTTTACCCGA GTACCAGAAG AACCCCGTGC
3421 TGGCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCCGA
3481 AGTGGATCAT GACCGCGGCC AAGTCCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCGCGT TCGCCAACGA GGGCTTCCTC GGCTACCAGT
3601 ACGAGTGCCC CGGCCTGATC GAGGTCCCCA CCGAGCAGGA CCCAGCAAG TCCTACTGGG
3661 TGAITTCAT CTCCATCAAC CCCGCGCCCC CGGCCGGCGG CTCCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CTTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCCTGGGCAT CGCGTGGGCC TCCAACCTGGG AGTACTCCGC TTCTGTGCC ACCAACCCCT
3901 GCGCTCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCCG AGCCGATCCT GAACATCAGC AACGCCGACC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTCGACC
4081 TGTCCAACAG CACCGGCACC CTGGAGTTTC AGCTGGTGTG CGCCGTCAAC ACCACCCAGA
4141 CGATCTCAA GTCCGTGTTT GCGGACCTCT CCCTCTGGTT CAAGGGCCTG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCTC TTCTTCTCTG 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 CGCGAGGTGA AGTGACAATT GGCAGCAGA CACTGGATAG
4561 TATCGACACA CTCTGGACGC TGGTCGTGTG ATGGACTGTT GCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTTGTGTGT
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 CACACGTGCC ACGTTGGCGA
5101 GGTGGCAGGT GACAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 CGAATTCCCTT TCTTGCGCTA TGACACTTCC AGCAAAAGGT AGGGCGGGCT GCGAGACGGC
5221 TTCCCGCGCG TGCAATGCAAC ACCGATGATG CTTTCGACCCC CCGAAGCTCC TTCGGGGCTG
5281 CATGGGCGCT CCGATGCCGC TCCAGGGCGA GCGCTGTTTA AATAGCCAGG CCCCCGATTG
5341 CAAAGACATT ATAGCGAGCT ACCAAAGCCA TATTCAAACA CCTAGATCAC TACCACTTCT
5401 ACACAGGCCA CTCGAGCTTG TGATCGCACT CCGCTAAGGG GCGCCTCTTT CCTCTTCGTT
5461 TCAGTCACAA CCCGCAAACA CTAGTATGGC CACCGCATCC ACTTCTTCGG CGTTCAATGC
5521 CCGCTGCGGC GACCTGCGTC GCTCGGCGGG CTCCGGGCCC GGGCGCCCC CCGGCCCCCT
5581 CCCCCTGCGC GGGCGCGCCA CCCAGGAGGA CCCCCACTCC GCCCCCCCC CCGGCCCCCT
5641 CATCAACCAG CTGCCCCACT GGTCCATGCT GCTGGCCGCC ATCACCACCG TGTTCTGGC
5701 CGCCGAGAAG CAGTGGATGA TGCTGGACTG GAAGCCCAAG CGCCCCGACA TGCTGGTGGG
5761 CCCCTTCGGC CTGGGCTCCA TCGTGCAGGA CGGCCTGGTG TTCCGCCAGA ACTTCTCCAT
5821 CCGCTCCTAC GAGATCGGCG CCGACCGCAC CGCTTCCATC GAGACCGTGA TGAACCACCT
5881 GCAGGAGACC GCCCTGAACC ACGTGAAGAT CGCCGGCCTG TCCAACGACG GCTTCGGCCG
5941 CACCCCCGAG ATGTACAAGC GCGACCTGAT CTGGGTGGTG GCCAAGATGC AGGTGATGGT
6001 GAACCGCTAC CCCACCTGGG GCGACACCGT GGAGGTGAAC ACCTGGGTGG CCAAGTCCGG
6061 CAAGAACGGC ATGCGCCGCG ACTGGCTGAT CTCCGACTGC AACACCGGCG AGATCTGAC
6121 CCGCGCCTCC TCCGTGTGGG TGATGATGAA CCAGAAGACC CGCCGCTGT CCAAGATCCC
6181 CGACGAGGTG CGCAACGAGA TCAGACCCCA CTTCTGTGGT TCCCCCCCCG TGATCGAGGA
6241 CGACGACCGC AAGCTGCCCA AGCTGGACGA GAAGACCGCC GACTCCATCC GCAAGGGCCT
6301 GACCCCCCGC TGGAACGACC TGGACGTGAA CCAGCACGTG AACACGTGA AGTACATCGG
6361 CTGGATCCTG GAGTCCACCC CCCCCGAGGT GCTGGAGACC CAGGAGCTGT GCTCCCTGAC

6421	CCTGGAGTAC	CGCCGCGAGT	GCGGCCGCGA	GTCCGTGCTG	GAGTCCCTGA	CCGCCATGGA
6481	CCCCTCCGGC	GGCGGCTACG	GCTCCCAGTT	CCAGCACCTG	CTGCGCCTGG	AGGACGGCGG
6541	CGAGATCGTG	AAGGGCCGCA	CCGAGTGGCG	CCCCAAGAAC	GGCGTGATCA	ACGGCGTGTT
6601	GCCCACCGGC	GAGTCCTCCC	CCGGCGACTA	CTCCATGGAC	TACAAGGACC	ACGACGGCGA
6661	CTACAAGGAC	CACGACATCG	ACTACAAGGA	CGACGACGAC	AAGTGAATCG	ATAGATCTCT
6721	TAAGGCAGCA	GCAGCTCGGA	TAGTATCGAC	ACACTCTGGA	CGCTGGTCGT	GTGATGGACT
6781	GTTGCCGCCA	CAC TTGCTGC	CTTGACCTGT	GAATATCCCT	GCCGCTTTTA	TCAAACAGCC
6841	TCAGTGTGTT	TGATCTTGTT	TGTACGCGCT	TTTGCGAGTT	GCTAGCTGCT	TGTGCTATTT
6901	GCGAATACCA	CCCCCAGCAT	CCCCTTCCCT	CGTTTCATAT	CGCTTGCATC	CCAACCGCAA
6961	CTTATCTACG	CTGTCTTGCT	ATCCCTCAGC	GCTGCTCCTG	CTCCTGCTCA	CTGCCCTCG
7021	CACAGCCTTG	GTTTGGGCTC	CGCCTGTATT	CTCCTGGTAC	TGCAACCTGT	AAACCAGCAC
7081	TGCAATGCTG	ATGCACGGGA	AGTAGTGGGA	TGGGAACACA	AATGGAAGC	TTAATTAAGA
7141	GCTCTTGTTT	TCCAGAAGGA	GTTGCTCCTT	GAGCCTTTCA	TTCTCAGCCT	CGATAACCTC
7201	CAAAGCCGCT	CTAATTGTGG	AGGGGGTTCG	AATTTAAAAG	CTTGGAATGT	TGTTTCGTGC
7261	GTCTGGAACA	AGCCCAGACT	TGTTGCTCAC	TGGGAAAAGG	ACCATCAGCT	CCAAAAAACT
7321	TGCCGCTCAA	ACCGCGTACC	TC TGCTTTTCG	CGCAATCTGC	CCTGTTGAAA	TCGCCACCAC
7381	ATT CATATTG	TGACGCTTGA	GCAGTCTGTA	ATTGCCTCAG	AATGTGGAAT	CATCTGCCCC
7441	CTGTGCGAGC	CCATGCCAGG	CATGTCGCGG	GCGAGGACAC	CCGCCACTCG	TACAGCAGAC
7501	CATTATGCTA	CCTCACAATA	GTT CATAACA	GTGACCATAT	TTCTCGAAGC	TCCCCAACGA
7561	GCACCTCCAT	GCTCTGAGTG	GCCACCCCCC	GGCCCTGGTG	CTTGCGGAGG	GCAGGTCAAC
7621	CGGCATGGGG	CTACCGAAAT	CCCCGACCGG	ATCCCACCAC	CCCCGCGATG	GGAAGAACTC
7681	CTCCCCGGGA	TGTGGGCCCA	CCACCAGCAC	AACCTGCTGG	CCCAGGCGAG	CGTCAAACCA
7741	TACCACACAA	ATATCCTTGG	CATCGGCCCT	GAATTCCTTC	TGCCGCTCTG	CTACCCGGTG
7801	CTTCTGTCCG	AAGCAGGGGT	TGCTAGGGAT	CGCTCCGAGT	CCGCAAACCC	TTGTCGCGTG
7861	GCGGGGCTTG	TTCGAGCTTG	AAGAGCCTCT	AGAGTCGACC	TGCAGGCATG	CAAGCTTGGC
7921	GTAATCATGG	TCATAGCTGT	TTCTGTGTG	AAATTGTTAT	CCGCTCACAA	TTCCACACAA
7981	CATACGAGCC	GGAAGCATAA	AGTGTAAAGC	CTGGGGTGCC	TAATGAGTGA	GCTAACTCAC
8041	ATTAATTGCG	TTGCGCTCAC	TGCCC GCTTT	CCAGTCGGGA	AACCTGTCTG	GCCAGCTGCA
8101	TTAATGAATC	GGCCAACGCG	CGGGGAGAGG	CGGTTTGCGT	ATTGGGCGCT	CTTCCGCTTC
8161	CTCGTCACT	GACTCGCTGC	GCTCGGTCGT	TCGGCTGCGG	CGAGCGGTAT	GAGCTCACTC
8221	AAAGGCGGTA	ATACGGTTAT	CCACAGAATC	AGGGGATAAC	GCAGGAAAGA	ACATGTGAGC
8281	AAAAGGCCAG	CAAAAGGCCA	GGAACCGTAA	AAAGGCCGCG	TTGCTGGCGT	TTTTCCATAG
8341	GCTCCGCCCC	CCTGACGAGC	ATCACA AAAA	TCGACGCTCA	AGTCAGAGGT	GGCGAAACCC
8401	GACAGGACTA	TAAAGATAAC	AGGCGTTTCC	CCCTGGAAGC	TCCCTCGTGC	GCTCTCCTGT
8461	TCCGACCCTG	CCGCTTACCG	GATACCTGTC	CGCCTTTCTC	CCTTCGGGAA	GCGTGGCGCT
8521	TTCTCATAGC	TCACGCTGTA	GGTATCTCAG	TTGCGTGTAG	GTCGTTGCTG	CCAAGCTGGG
8581	CTGTGTGCAC	GAACCCCCCG	TTCAGCCCCG	CCGCTGCGCC	TTATCCGGTA	ACTATCGTCT
8641	TGAGTCCAAC	CCGGTAAGAC	ACGACTTATC	GC		

SEQ ID NO: 226
pSZ1417

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	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACTTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT
421	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTCTG	TCATCCATAG
481	TTGCCTGACT	CCCCGTCGTG	TAGATAACTA	CGATACGGGA	GGGCTTACCA	TCTGGCCCCA
541	GTGCTGCAAT	GATACGCGA	GACCCACGCT	CACCGGCTCC	AGATTTATCA	GCAATAAAC
601	AGCCAGCCGG	AAGGGCCGAG	CGCAGAAAGT	GTCTTGCAAC	TTTATCCGCC	TCCATCCAGT
661	CTATTAATTG	TTGCCGGGAA	GCTAGAGTAA	GTAGTTCGCC	AGTTAATAGT	TTGCGCAACG
721	TTGTTGCCAT	TGCTACAGGC	ATCGTGTTGT	CACGCTCGTC	GTTTGATATG	GTTTCAATCA
781	GCTCCGGTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC	CATGTTGTGC	AAAAAAGCGG
841	TTAGTCTCTT	CGGTCTCTCG	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAAGT	TTATCACTCA
901	TGGTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTCAATGC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTA CTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA

1081	TCATTGAAAA	ACGTTCTTCG	GGGCGAAAAAC	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCGATGTA	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	AGCTTGCTCG	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCCG	CAGGCGCGCT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATACCGCA	TCAGGCGCCA	TTCGCCATTG	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAGGGG
1801	GATGTGCTGC	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TCCCCAGTCA	CGACGTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCGCCGCCG	CCACTCCTGC	TCGAGCGCGC
1921	CCGCGCGTGC	GCCGCCAGCG	CCTTGCCCTT	TTCGCCGCGC	TCGTGCGCGT	CGCTGATGTC
1981	CATCACCAGG	TCCATGAGGT	CTGCCTTGCG	CCGGCTGAGC	CACTGCTTCG	TCCGGGCGGC
2041	CAAGAGGAGC	ATGAGGGAGG	ACTCCTGGTC	CAGGGTCCTG	ACGTGGTTCG	GGCTCTGGGA
2101	GCGGGCCAGC	ATCATCTGGC	TCTGCCGCAC	CGAGGCCGCC	TCCAACCTGG	CCTCCAGCAG
2161	CCGCAGTCGC	CGCCGACCCT	GGCAGAGGAA	GACAGGTGAG	GGGGGTATGA	ATTGTACAGA
2221	ACAACCACGA	GCCTTGTCTA	GGCAGAATCC	CTACCAGTCA	TGGCTTTTAC	TGGATGACGG
2281	CCTGCGAACA	GCTGTCCAGC	GACCTTCGCT	GCCGCCGCTT	CTCCCCAGC	CTTCTTTTCA
2341	GCACCGTGAT	GGCGCGAGCC	AGCGCCGCAC	GCTGGCGCTG	CGCTTCGCCG	ATCTGAGGAC
2401	AGTCGGGGAA	CTCTGATCAG	TCTAAACCCC	CTTGCGCGTT	AGTGTTGCCA	TCCTTTGCAG
2461	ACCGGTGAGA	GCCGACTTGT	TGTGCGCCAC	CCCCACACC	ACCTCCTCCC	AGACCAATTC
2521	TGTCACCTTT	TTGGCGAAGG	CATCGGCCCT	GGCCTGCAGA	GAGGACAGCA	GTGCCAGGCC
2581	GCTGGGGGTT	GGCGGATGCA	CGCTCAGGTA	CCCTTTCTTG	CGCTATGACA	CTTCCAGCAA
2641	AAGGTAGGGC	GGGCTGCGAG	ACGGCTTCCC	GGCGCTGCAT	GCAACACCGA	TGATGCTTCG
2701	ACCCCCCGAA	GCTCCTTCGG	GGCTGCATGG	GCGCTCCGAT	GCCGCTCCAG	GGCGAGCGCT
2761	GTTTTAAATAG	CCAGGCCCCC	GATTGCAAAAG	ACATTATAGC	GAGCTACCAA	AGCCATATTC
2821	AAACACCTAG	ATCACTACCA	CTTCTACACA	GGCCACTCGA	GCTTGTGATC	GCACTCCGCT
2881	AAGGGGGCGC	CTCTTCCTCT	TCGTTTCAGT	CACAACCCGC	AAACGGCGCG	CCATGCTGCT
2941	GCAGGCCCTT	CTGTTCTTGC	TGGCCGGCTT	CGCCGCCAAG	ATCAGCGCCT	CCATGACGAA
3001	CGAGACGTCC	GACCGCCCCC	TGGTGCACCT	CACCCCCAAC	AAGGGCTGGA	TGAACGACCC
3061	CAACGGCCTG	TGGTACGACG	AGAAGGACGC	CAAGTGGCAC	CTGTACTTCC	AGTACAACCC
3121	GAACGACACC	GTCTGGGGGA	CGCCTTGTTT	CTGGGGCCAC	GCCACGTCCG	ACGACCTGAC
3181	CAACTGGGAG	GACCAGCCCA	TCGCCATCGC	CCCGAAGCGC	AACGACTCCG	GCGCCTTCTC
3241	CGGCTCCATG	GTGGTGGACT	ACAACAACAC	CTCCGGCTTC	TTCAACGACA	CCATCGACCC
3301	GCGCCAGCGC	TGCGTGGCCA	TCTGGACCTA	CAACACCCCG	GAGTCCGAGG	AGCAGTACAT
3361	CTCCTACAGC	CTGGACGGCG	GCTACACCTT	CACCGAGTAC	CAGAAGAACC	CCGTGCTGGC
3421	CGCCAACTCC	ACCCAGTTCC	GCGACCCGAA	GGTCTTCTGG	TACGAGCCCT	CCGAGAAGTG
3481	GATCATGACC	GCGGCCAAGT	CCCAGGACTA	CAAGATCGAG	ATCTACTCCT	CCGACGACCT
3541	GAAGTCCTGG	AAGCTGGAGT	CCGCGTTTCG	CAACGAGGGC	TTCTTCGGCT	ACCAGTACGA
3601	GTGCCCCGGC	CTGATCGAGG	TCCCCACCGA	GCAGGACCCC	AGCAAGTCCT	ACTGGGTGAT
3661	GTTCACTCTC	ATCAACCCCG	GCGCCCCGGC	CGGCGGCTCC	TTCAACCAGT	ACTTCGTCGG
3721	CAGCTTCAAC	GGCACCCACT	TCGAGGCCCT	CGACAACCAG	TCCCGCGTGG	TGGACTTCGG
3781	CAAGGACTAC	TACGCCCTGC	AGACCTTCTT	CAACACCGAC	CCGACCTACG	GGAGCGCCCT
3841	GGGCATCGCG	TGGGCCCTCA	ACTGGGAGTA	CTCCGCCTTC	GTGCCCACCA	ACCCCTGGCG
3901	CTCCTCCATG	TCCCTCGTGC	GCAAGTTCTC	CTCAACACC	GAGTACCAGG	CCAACCCGGA
3961	GACGGGAGCT	ATCAACCTGA	AGGCCGAGCC	GATCCTGAAC	ATCAGCAACG	CCGGCCCCTG
4021	GAGCCGGTTC	GCCACCAACA	CCACGTTGAC	GAAGGCCAAC	AGCTACAACG	TCGACCTGTC
4081	CAACAGCACC	GGCACCCCTG	AGTTCGAGCT	GGTGTACGCC	GTCAACACCA	CCCAGACGAT
4141	CTCCAAGTCC	GTGTTTCGCG	ACCTCTCCCT	CTGGTTCAAG	GGCCTGGAGG	ACCCCGAGGA
4201	GTACCTCCGC	ATGGGCTTCG	AGGTGTCCGC	GTCTCTCCTC	TTCTTGGACC	GCGGGAACAG
4261	CAAGGTGAAG	TTCGTGAAGG	AGAACCCCTA	CTTCACCAAC	CGCATGAGCG	TGAACAACCA
4321	GCCCTTCAAG	AGCGAGAACG	ACCTGTCTTA	CTACAAGGTG	TACGGCTTGC	TGGACCAGAA
4381	CATCTGGAG	CTGTACTTCA	ACGACGGCGA	CGTCGTGTCC	ACCAACACCT	ACTTCTAGAC
4441	CACCGGAAC	GCCCTGGGCT	CCGTGAACAT	GACGACGGGG	GTGGACAACC	TGTTCTACAT
4501	CGACAAGTTC	CAGGTGCGCG	AGGTCAAGTG	ACAATTGGCA	GCAGCAGCTC	GGATAGTATC
4561	GACACACTCT	GGACGCTGGT	CGTGTGATGG	ACTGTTGCCG	CCACACTTGC	TGCCTTGACC
4621	TGTGAATATC	CCTGCCGCTT	TTATCAAACA	GCCTCAGTGT	GTTTGATCTT	GTGTGTACGC
4681	GCTTTTTCGA	GTTGCTAGCT	GCTTGTGCTA	TTTGCGAATA	CCACCCCCAG	CATCCCCTTC

4741 CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTATCT ACGCTGTCTT GCTATCCCTC
4801 AGCGCTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG CTCCGCTGTG
4861 ATTCTCTTGG TACTGCAACC TGTAACCAG CACTGCAATG CTGATGCACG GGAAGTAGTG
4921 GGATGGGAAC ACAAATGGAG GATCCCGCGT CTCGAACAGA GCGCGCAGAG GAACGCTGAA
4981 GGTCTCGCCT CTGTGCGACC TCAGCGCGGC ATACACCACA ATAACCACCT GACGAATGCG
5041 CTTGGTTCTT CGTCCATTAG CGAAGCGTCC GGTTACACACA CGTGCCACGT TGGCGAGGTG
5101 GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG GGATATCGAA
5161 TTCGGCCGAC AGGACGCGCG TCAAAGGTGC TGGTCGTGTA TGCCCTGGCC GGCAGGTCTG
5221 TGCTGCTGCT GGTTAGTGAT TCCGCAACCC TGATTTTGGC GTCTTATTTT GGCGTGGCAA
5281 ACGCTGGCGC CCGCGAGCCG GGCCGGCGGC GATGCGGTGC CCCACGGCTG CCGGAATCCA
5341 AGGGAGGCAA GAGCGCCCGG GTCAGTTGAA GGGCTTTACG CGCAAGGTAC AGCCGCTCCT
5401 GCAAGGTGTC GTGGTGGAAT TGGACGTGCA GGTCTGTCTG AAGTTCTCTC ACCGCTCAG
5461 CAGCGGACAA AGCACCGGTG TATCAGGTCC GTGTCATCCA CTCTAAAGAA CTCGACTACG
5521 ACCTACTGAT GGCCCTAGAT TCTTCATCAA AAACGCCTGA GACACTTGCC CAGGATTGAA
5581 ACTCCCTGAA GGGACCACCA GGGGCCCTGA GTTGTTCCTT CCCCCGTGG CGAGCTGCCA
5641 GCCAGGTGT ACCTGTGATC GAGGCTGGCG GGAAAATAGG CTTCTGTGTC TCAGGTCAATG
5701 GGAGGTGACG GACAGCTCAT GAAACGCCAA CAATCGCACA ATTCATGTCA AGCTAATCAG
5761 CTATTTCCTC TTCACGAGCT GTAATTGTCC CAAAATTCTG GTCTACCGGG GGTGATCCTT
5821 CGTGACGGG CCCTTCCCTC AACCTTAGGT ATGCGCGCAT GCGGTGCGCG CGCAACTCGC
5881 GCGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGACCCG GATGCGTGGC
5941 ACCTTTTTTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT CTGTGCCCAA
6001 CCGTAGATC AGCGGCGTAG GATTTCGTAA TCATTCTGTC TGATGGGGAG CTACCGACTA
6061 CCCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC ATTCCATTCTG
6121 TGCCCAAGAC ATTTTCATTGT GTTGCGAAGC GTCCCCAGTT ACGCTCACCT GTTTCCCGAC
6181 CTCCTTACTG TTCTGTGAC AGAGCGGGCC CACAGGCCGG TCGCAGCCAC TAGTATGGTG
6241 GCCACCGCCG CCTCCTCCGC CTTCTTCCCC CTGCCCTCCG CCGACACCTC CTCCCGCCCC
6301 GGCAAGCTGG GCAACAAGCC CTCTCCCTG TCCCCCTGA AGCCCAAGTC CACCCCAAC
6361 GGCGGCTGC AGGTGAAGGC CAACGCCTCC GCCCCCCCCA AGATCAACGG CTCCCCCGTG
6421 GGCTGAAGT CCGGCGGCCT GAAGGGGCGC GCCACCCAGG AGGACGCCCA CTCCGCCCCC
6481 CCCCCCGCA CCTTCATCAA CCAGTGCCC GACTGGTCCA TGCTGTGGC CGCCATCACC
6541 ACCGTGTTC TGGCCGCGGA 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 GCCGCACCCC CGAGATGTAC AAGCGCGACC TGATCTGGGT GGTGGCCAAG
6841 ATGCAGGTGA TGGTGAACCG CTACCCACC TGGGGCGACA CCGTGAGGT GAACACCTGG
6901 GTGGCCAAGT CCGGCAAGAA CGGCATGCGC CGCGACTGGC TGATCTCCGA CTGCAACACC
6961 GGCGAGATCC TGACCCGCGC CTCTCCGTG TGGGTGATGA TGAACCAGAA GACCCGCGCG
7021 CTGTCCAAGA TCCCCGACGA GGTGCGCAAC GAGATCGAGC CCCACTTCGT GGACTCCCCC
7081 CCGTGATCG AGGACGACGA CCGCAAGCTG CCAAGCTGG ACGAGAAGAC CGCCGACTCC
7141 ATCCGCAAGG GCCTGACCCC CCGCTGGAAC GACCTGGACG TGAACCAGCA CGTGAACAAC
7201 GTGAAGTACA TCGGCTGGAT CCTGGAGTCC ACCCCCCCGG AGGTGCTGGA GACCCAGGAG
7261 CTGTGCTCCC TGACCCTGGA GTACCGCCGC GAGTGC GGCC GCGAGTCCGT GCTGGAGTCC
7321 CTGACCGCCA TGGACCCCTC CGGCGGCGGC TACGGCTCCC AGTTCCAGCA CCTGTGCGC
7381 CTGGAGGACG GCGGCGAGAT CGTGAAGGGC CGCACCAGT GCGCCCCCAA GAACGGCGTG
7441 ATCAACGGCG TGGTGCCAC CGGCGAGTCC TCCCCCGCG ACTACTCCAT GGACTACAAG
7501 GACCACGACG GCGACTACAA GGACCACGAC ATCGACTACA AGGACGACGA CGACAAGTGA
7561 ATCGATAGAT CTCTTAAGGC AGCAGCAGCT CGGATAGTAT CGACACATC TGGACCTTGG
7621 TCGTGTGATG GACTGTTGCC GCCACACTTG CTGCCTTGAC CTGTGAATAT CCCTGCCGCT
7681 TTTATCAAAC AGCCTCAGTG TGTGTGATCT TGTGTGTACG CGCTTTTGCG AGTTGCTAGC
7741 TGCTTGTGCT ATTTGCGAAT ACCACCCCCA GCATCCCCTT CCCTCGTTTC ATATCGCTTG
7801 CATCCCAACC GCAACTTATC TACGCTGTCC TGCTATCCCT CAGCGCTGCT CCTGCTCCTG
7861 CTCACGCCCC CTCGCACAGC CTTGGTTTGG GCTCCGCTG TATTCTCCTG GTACTGCAAC
7921 CTGTAACCA GCACTGCAAT GCTGATGCAC GGGAAGTAGT GGGATGGGAA CACAAATGGA
7981 AAGCTTAATT AAGAGCTCTT GTTTTCCAGA AGGAGTTGCT CCTTGAGCCT TTCATTCTCA
8041 GCCTCGATAA CCTCCAAGC CGCTCTAATT GTGGAGGGGG TTCGAATTTA AAAGCTTGA
8101 ATGTTGGTTC GTGCGTCTGG AACAAGCCCA GACTTGTTC TCACTGGGAA AAGGACCATC
8161 AGCTCCAAAA AACTTGCCGC TCAAACCGCG TACCTCTGCT TTCGCGCAAT TGCCCTGTT
8221 GAAATCGCCA CCACATTATC ATTGTGACGC TTGAGCAGTC TGTAATTGCC TCAGAATGTG
8281 GAATCATCTG CCCCCTGTGC GAGCCCATGC CAGGCATGTC GCGGGCGAGG ACACCCGCCA
8341 CTCGTACAGC AGACCATTAT GCTACCTCAC AATAGTTCAT AACAGTGACC ATATTTCTCG

8401	AAGCTCCCCA	ACGAGCACCT	CCATGCTCTG	AGTGGCCACC	CCCCGGCCCT	GGTGCTTGCG
8461	GAGGGCAGGT	CAACCGGCAT	GGGGCTACCG	AAATCCCCGA	CCGGATCCCCA	CCACCCCCGC
8521	GATGGGAAGA	ATCTCTCCCC	GGGATGTGGG	CCCACCACCA	GCACAACCTG	CTGGCCCAGG
8581	CGAGCGTCAA	ACCATACCAC	ACAAATATCC	TTGGCATCGG	CCCTGAATTC	CTTCTGCCGC
8641	TCTGCTACCC	GGTGCTTCTG	TCCGAAGCAG	GGGTTGCTAG	GGATCGCTCC	GAGTCCGCAA
8701	ACCCTTGTCTG	CGTGGCGGGG	CTTGTTCTGAG	CTTGAAGAGC	CTCTAGAGTC	GACCTGCAGG
8761	CATGCAAGCT	TGGCGTAATC	ATGGTCATAG	CTGTTTCCTG	TGTGAAATTG	TTATCCGCTC
8821	ACAATTCCAC	ACAACATACG	AGCCGGAAGC	ATAAAGTGTA	AAGCCTGGGG	TGCCTAATGA
8881	GTGAGCTAAC	TCACATTAAT	TGCGTTGCGC	TCACTGCCCC	CTTCCAGTC	GGGAAACCTG
8941	TCGTGCCAGC	TGCATTAATG	AATCGGCCAA	CGCGCGGGGA	GAGGCGGTTT	GCGTATTGGG
9001	CGCTCTTCCG	CTTCCTCGCT	CACTGACTCG	CTGCGCTCGG	TCGTTCGGCT	GCGGCGAGCG
9061	GTATCAGCTC	ACTCAAAGGC	GGTAATACGG	TTATCCACAG	AATCAGGGGA	TAACGCAGGA
9121	AAGAACATGT	GAGCAAAAGG	CCAGCAAAAG	GCCAGGAACC	GTAAAAAAGG	CGCGTTGCTG
9181	GCGTTTTTCC	ATAGGCTCCG	CCCCCTGAC	GAGCATCACA	AAAAATCGAC	CTCAAGTCAG
9241	AGGTGGCGAA	ACCCGACAGG	ACTATAAAGA	TACCAGGCGT	TTCCCCCTGG	AAGCTCCCTC
9301	GTGCGCTCTC	CTGTTCCGAC	CCTGCCGCTT	ACCGGATACC	TGTCCGCCTT	TCTCCCTTCG
9361	GGAAGCGTGG	CGCTTTCTCA	TAGCTCACGC	TGTAGGTATC	TCAGTTCGGT	GTAGGTCGTT
9421	CGCTCCAAGC	TGGGCTGTGT	GCACGAACCC	CCCGTTCAGC	CCGACCGCTG	CGCCTTATCC
9481	GGTAACATATC	GTCTTGAGTC	CAACCCGGTA	AGACACGACT	TATCG	

SEQ ID NO: 227
pSZ1119

1	CACITGGCAGC	AGCCACTGGT	AACAGGATTA	GCAGAGCGAG	GTATGTAGGC	GGTGCTACAG
61	AGTTCCTGAA	GTGGTGGCCT	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	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	CGCAGAAAGT	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	CGGTCCCTCC	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAGTG	TTATCACTCA
901	TGTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTATGCCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCAGAAAC	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCGATGTA	ACCCACTCGT	GCACCCAACT	GATCTTCAGC	ATCTTTTACT	TTCACCAGCG
1201	TTTCTGGGTG	AGCAAAAACA	GGAAGGCAAA	ATGCCGCAAA	AAAGGGAATA	AGGGCGACAC
1261	GGAAATGTTG	AATACTCATA	CTCTTCCTTT	TTCAATATTA	TTGAAGCATT	TATCAGGGTT
1321	ATTGTCCTCAT	GAGCGGATAC	ATATTTAGAT	GTATTTAGAA	AAATAAACAA	ATAGGGGTTT
1381	CGCGCACATT	TCCCCGAAAA	GTGCCACATG	ACGTCTAAGA	AACCATTATT	ATCATGACAT
1441	TAACCTATAA	AAATAGGCGT	ATCACGAGGC	CCTTTCGTCT	CGCGCGTTTC	GGTGATGACG
1501	GTGAAAACCT	CTGACACATG	CAGCTCCCGG	AGACGGTCAC	AGCTTGTCTG	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCCT	CAGGGCGCGT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATACCGCA	TCAGGCGCCA	TTCCGCATTTC	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAGGGG
1801	GATGTGCTGC	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TTCCAGCTCA	CGACGTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCCGCGCGG	CCACTCCTGC	TCGAGCGCGC
1921	CCGCGCGTGC	GCCGCCAGCG	CCTTGCCCTT	TTCCGCGCGC	TGCTGCGCGT	TCGTGATGTC
1981	CATCACCAGG	TCCATGAGGT	CTGCCTTGCG	CCGGCTGAGC	CACTGCTTCG	TCCGGGCGGC
2041	CAAGAGGAGC	ATGAGGGAGG	ACTCCTGGTC	CAGGGTCCTG	ACGTGGTCTG	GGCTCTGGGA
2101	GCGGGCCAGC	ATCATCTGGC	TCTGCCGCAC	CGAGGCCGCC	TCCAACCTGGT	CCTCCAGCAG
2161	CCGCAGTCGC	CGCCGACCCT	GGCAGAGGAA	GACAGGTGAG	GGGGGTATGA	ATTGTACAGA

2221 ACAACCACGA GCCTTGTCTA GGCAGAAATCC CTACCAGTCA TGGCTTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCTCGCT GCCGCCGCTT CTCCCGCACG CTTCTTTTCCA
2341 GCACCGTGAT GCGCGGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTTGCCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACTTTT TTGGCGAAGG CATCGGCCCTC GGCCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGTGCATGG GCGCTCCGAT GCCGCTCCAG GCGAGCGCT
2761 GTTTAAATAG CCAGGCCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGCGC CTCTTCTCT TCGTTTCAGT CACAACCCGC AAACGGCGCG CCATGCTGCT
2941 GCAGGCCTTC CTGTTCCCTG TGGCCGGCTT CGCCGCCAAG ATCAGCGCCT CCATGACGAA
3001 CGAGACGTCC GACCGCCCCC TGGTGCACTT CACCCCCAAC AAGGGCTGGA TGAACGACCC
3061 CAACGGCCTG TGGTACGACG AGAAGGACGC CAAGTGGCAC CTGTACTTCC AGTACAACCC
3121 GAACGACACC GTCTGGGGGA CGCCCTTGTT CTGGGGCCAC GCCACGTCCG ACGACCTGAC
3181 CAACTGGGAG GACCAGCCCA TCGCATCGC CCCGAAGCGC AACGACTCCG GCGCCTTCTC
3241 CGGCTCCATG GTGGTGGACT ACAACAACAC CTCCGGCTTC TTCAACGACA CCATCGACCC
3301 GCGCCAGCGC TGCGTGGCCA TCTGGACCTA CAACACCCCG GAGTCCGAGG AGCAGTACAT
3361 CTCTACAGC CTGGACGGCG GCTACACCTT CACCGAGTAC CAGAAGAACC CCGTGCTGGC
3421 CGCCAACTCC ACCCAGTTCC GCGACCCGAA GGTCTTCTGG TACGAGCCTT CCCAGAAGTG
3481 GATCATGACC GCGGCCAAGT CCCAGGACTA CAAGATCGAG ATCTACTCCT CCGACGACCT
3541 GAAGTCTTGG AAGCTGGAGT CCGCGTTTCG CAACGAGGGC TTCTTCGGCT ACCAGTACGA
3601 GTGCCCCGGC CTGATCGAGG TCCCCACCGA GCAGGACCCC AGCAAGTCCT ACTGGGTGAT
3661 GTTCATCTCC ATCAACCCCG GCGCCCCGGC CGGCGGCTCC TTCAACCAGT ACTTCGTCGG
3721 CAGCTTCAAC GGCACCCACT TCGAGGCCCT CGACAACCAG TCCCGCGTGG TGGACTTCGG
3781 CAAGGACTAC TACGCCCTGC AGACCTTCTT CAACACCGAC CCGACCTACG GGAGCGCCCT
3841 GGGCATCGCG TGGGCCTCCA ACTGGGAGTA CTCCGCCTTC GTGCCCACCA 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 CTCCAAGTCC GTGTTTCGGG ACCTCTCCCT CTGGTTCAAG GGCCTGGAGG ACCCCGAGGA
4201 GTACCTCCGC ATGGGCTTCG AGGTGTCCGC GTCTCTCTTC TTCCTGGACC GCGGGAACAG
4261 CAAGGTGAAG TTCGTGAAGG AGAACCCTTA CTTACCAAC CGCATGAGCG TGAACAACCA
4321 GCCCTTCAAG AGCGAGAACG ACCTGTCTTA CTACAAGGTG TACGGCTTGC TGGACCAGAA
4381 CATCTGGAG CTGTACTTCA ACGACGGCGA CGTCGTGTCC ACCAACACCT ACTTCATGAC
4441 CACCGGGAAC GCCCTGGGCT CCGTGAACAT GACGACGGGG GTGGACAACC TGTTCTACAT
4501 GCACAAGTTC CAGGTGCGCG AGGTCAAGTG ACAATTGGCA GCAGCAGCTC GGATAGTATC
4561 GACACACTCT GGACGCTGGT CGTGTGATGG ACTGTTGCCG CCACACTTGC TGCCCTGACC
4621 TGTGAATATC CCTGCCGCTT TTATCAAACA GCCTCAGTGT GTTTGATCTT GTGTGTACGC
4681 GCTTTTGC GA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCCAG CATCCCCTTC
4741 CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTATCT ACGCTGTCTT GCTATCCCTC
4801 AGCGTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG CTCCGCCTGT
4861 ATTCTCCTGG TACTGCAACC TGTAACCAG CACTGCAATG CTGATGCACG GGAAGTAGTG
4921 GGATGGGAAC ACAAATGGAG GATCCCGCGT CTCGAACAGA GCGCGCAGAG GAACGCTGAA
4981 GTTCTCGCCT CTGTCGCACC TCAGCGCGG ATACACCACA ATAACCACCT GACGAATGGC
5041 CTTGGTTCTT CGTCCATTAG CGAAGCGTCC GGTTCACACA CGTGCCACCT TGGCGAGGTG
5101 GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG GGATATCGAA
5161 TTCGGCCGAC AGGACGCGCG TCAAAGGTGC 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 TGACGTGCA GGTCTGTCTG AAGTTCCTCC ACCGCTCAC
5461 CAGCGGACAA AGCACCGGTG TATCAGGTCC GTGTATCCA CTCTAAAGAG CTCGACTACG
5521 ACCTACTGAT GGCCCTAGAT TCTTCATCAA AAACGCCTGA GACACTTGCC CAGGATTGAA
5581 ACTCCCTGAA GGGACCACCA GGGGCCCTGA GTTGTTCTTT CCCCCGTGG CAGAGTCCCA
5641 GCCAGGCTGT ACCTGTGATC GAGGCTGGCG GGAAAATAGG CTTCTGTGTC TCAGGTCATG
5701 GGAGGTGCAG GACAGCTCAT GAAACGCCAA CAATCGCACA ATTCATGTCA AGCTAATCAG
5761 CTATTTCTCT TTCACGAGCT GTAATTGTCC CAAAATTCTG GTCTACCGGG GGTGATCCTT
5821 CGTGTACGGG CCCTTCCCTC AACCTAGGT ATGCGCGCAT GCGGTCGCCG CGCAACTCGC

5881 GCGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGCAACCG GATGCGTGGC
5941 ACCTTTTTTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT CTGTCGCCAA
6001 CCCTAGGATC AGCGGCGTAG GATTTCGTAA TCATTCTGTC TGATGGGGAG CTACCGACTA
6061 CCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC ATTCCATTCTG
6121 TGCCCAAGAC ATTTTCATTGT GGTGCGAAGC GTCCCCAGTT ACGCTCACCT GTTTCCCGAC
6181 CTCCTTACTG TTCTGTGCGC AGAGCGGGCC CACAGGCCGG TCGCAGCCAC TAGTATGGCC
6241 ACCGCATCCA CTTTCTCGGC GTTCAATGCC CGCTGCGGCG ACCTGCGTCG CTCGGCGGGC
6301 TCCGGGCCCC GCGGCCCAGC GAGGCCCTC CCCGTGCGCG GCGCGCCCA GCTGCCCGAC
6361 TGGAGCATGC TGCTGGCCGC GATCACCACC CTGTTCCCTG 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 GCGTGGGCCT GCTGGAGGAC GGCCTGGGCA GCACCCGCGA GATGAGCCTG
6661 CGCAACCTGA TCTGGGTGGT GACCAAGATG CAGGTGGCGG TGGACCGCTA CCCCACCTGG
6721 GCGGACGAGG TGCAGGTGAG CAGCTGGGCG ACCGCCATCG GCAAGAACGG CATGCGCCGC
6781 GAGTGGATCG TGACCGACTT CCGCACCAGC GAGACCCTGC TCGCGCCAC CAGCGTGTGG
6841 GTGATGATGA ACAAGCTGAC CCGCCGCATC AGCAAGATCC CCGAGGAGGT GTGGCACGAG
6901 ATCGGCCCCA GCTTCATCGA CGCGCCCCC CTGCCACCAG TGGAGGACGA CGGCCGCAAG
6961 CTGACCCGCT TCGACGAGAG CAGCGCCGAC TTCATCCGCA AGGGCCTGAC CCCCCGCTGG
7021 AGCGACCTGG ACATCAACCA GCACGTGAAC AACGTGAAGT ACATCGGCTG GCTGCTGGAG
7081 AGCGCGCCCC CCGAGATCCA CGAGAGCCAC GAGATCGCCA GCCTGACCCT GGAGTACCGC
7141 CGCGAGTGCG GCGCGACAG CGTGCTGAAC AGCGCCACCA AGGTGAGCGA CAGCAGCCAG
7201 CTGGGCAAGA GCGCCGTGGA GTGCAACCAC CTGGTGCGCC TGCAGAACGG CGGCGAGATC
7261 GTGAAGGGCC GCACCGTGTG GCGCCCCAAG CGCCCCCTGT ACAACGACGG CGCCGTGGTG
7321 GACGTGCCCG CCAAGACCAG CGATGACGAT GACAAGCTGG GATGACTCGA GGCAGCAGCA
7381 GCTCGGATAG TATCGACACA CTCTGGACGC TGGTCGTGTG ATGGACTGTT GCCGCCACAC
7441 TTGCTGCCCT GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA
7501 TCTTGTGTGT ACGCGCTTTT GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC
7561 CCAGCATCCC CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG
7621 TCTGTCTATC CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGTTT
7681 TGGGCTCCGC CTGTATTCTC CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGATG
7741 CACGGGAAGT AGTGGGATGG GAACACAAAT GGAAAGCTTG AGCTCTTGTT TTCCAGAAGG
7801 AGTTGCTCCT TGAGCCTTTC ATTCTCAGCC TCGATAACCT CCAAAGCCGC TCTAATTGTG
7861 GAGGGGGTTC GAATTTAAAA GCTTGAATG TTGGTTTCGT CGTCTGGAAC AAGCCCAGAC
7921 TTGTTGCTCA CTGGGAAAAG GACCATCAGC TCCAAAAAAC TTGCCGCTCA AACCAGCTAC
7981 CTCTGCTTTC GCGCAATCTG CCCTGTTGAA ATCGCCACCA CATTATATT GTGACGCTTG
8041 AGCAGTCTGT AATTGCCTCA GAATGTGGAA TCATCTGCCC CCTGTGCGAG CCCATGCCAG
8101 GCATGTCGCG GCGGAGGACA CCCGCCACTC GTACAGCAGA CCATTATGCT ACCTCACAAT
8161 AGTTCTAATC AGTGACCATA TTTCTCGAAG CTCCCCAACG AGCACCTCCA TGCTCTGAGT
8221 GGCCACCCCC CGGCCCTGGT GCTTGCAGGAG GGCAGGTCAA CCGGATGGG GCTACCGAAA
8281 TCCCCGACCG GATCCCACCA CCCCCGCGAT GGGAAGAATC TCTCCCCGGG ATGTGGGCCC
8341 ACCACCAGCA CAACCTGCTG GCCCAGGCGA GCGTCAAACC ATACCACACA AATATCCTTG
8401 GCATCGGCCC TGAATTCTT CTGCCGCTCT GCTACCCGGT GCTTCTGTCC GAAGCAGGGG
8461 TTGCTAGGGA TCGCTCCGAG TCCGCAAACC CTGTGTCGCT GCGGGGGCTT GTTCGAGCTT
8521 GAAGAGCCTC TAGAGTCGAC CTGCAGGCAT GCAAGCTTGG CGTAATCATG GTCATAGCTG
8581 TTTCTGTGT GAAATTGTTA TCCGCTCACA ATTCCACACA ACATACGAGC CGGAAGCATA
8641 AAGTGTAAG CCTGGGGTGC CTAATGAGTG AGCTAACTCA CATTAATTGAT GTTGCGTCA
8701 CTGCCCCTT TCCAGTCGGG AAACCTGTG TGCCAGCTGC ATTAATGAAT CCGCCAACGC
8761 GCGGGGAGAG GCGGTTTTCG TATTGGGCGC TCTTCCGCTT CCTCGCTCAC TGAAGCTGCTG
8821 CGCTCGGTG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT CAAAGGCGGT AATACGGTTA
8881 TCCACAGAA CAGGGGATAA CGCAGGAAAG AACATGTGAG CAAAAGGCCA GCAAAAGGCC
8941 AGGAACCGTA AAAAGGCCGC GTTGCTGGCG TTTTTCATA GGCTCCGCC CCCTGACGAG
9001 CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC
9061 CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCTT GCCGCTTACC
9121 GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC TTTCTCATAG CTCACGCTGT
9181 AGGTATCTCA GTTCGGTGTA GGTCGTTCGC TCCAAGCTGG GCTGTGTGCA CGAACCCCCC
9241 GTTCAGCCCG ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA
9301 CACGACTTAT CGC

SEQ ID NO: 228

pSZ944

```

1 CTAAATTGTA AGCGTTAATA TTTTGTAAAA ATTTCGCGTTA AATTTTTGTT AAATCAGCTC
61 ATTTTTTAAC CAATAGGCCG AAATCGGCCA AATCCCTTAT AAATCAAAAAG AATAGACCGA
121 GATAGGGTTG AGTGTGTGTT CAGTTTGGAA CAAGAGTCCA CTATTAAAGA ACGTGGACTC
181 CAACGTCAAA GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCATCACC
241 CTAATCAAGT TTTTGGGGT CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG
301 CCCCCGATTT AGAGCTTGAC GGGGAAAGCC GGCGAACGTG GCGAGAAAGG AAGGGAAGAA
361 AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC GCGTAACCAC
421 CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCC CATTGCCCAT TCAGGCTGCG
481 CAACTGTTGG GAAGGGCGAT CGGTGCGGGC CTCTTCGCTA TTACGCCAGC TGGCGAAAGG
541 GGGATGTGCT GCAAGGCGAT TAAGTTGGGT AACGCCAGGG TTTCCCAGT CACGACGTTG
601 TAAAACGACG GCCAGTGAGC GCGCGTAATA CGACTCACTA TAGGGCGAAT TGCAGCTTTT
661 GTTCCCTTTA GTGAGGGTTA ATTGCGCGCT TGGCGTAATC ATGGTCATAG CTGTTTCCTG
721 TGTGAAATTG TTATCCGCTC ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA
781 AAGCCTGGGG TGCCTAATGA GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG
841 CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA
901 GAGGCGGTGC GTATTGGGCG CTCTTCGCT CCCGTGATCA CACAGGTGCC TTGCGAGCGT
961 GATCACACTA TTTTGGGGT CCTACAGTAC TGAAATGGTG AGAAGTCGTA CTGAAATCAA
1021 GGATGAACAA TGAAAATGGT GCTGTGGTGG CTTCTCAAAG GTCAAGAATC AGTCGCTCGC
1081 GTCAGGAAAT CGCGGCGTCA ACCAGCGTGG GCGCGGTGAG TGGCCCCGCA CTGGTCACCA
1141 TAGCCTCTCC TGCCACAGTA GCGATCCCCT GGGCGTTTAC TCTCAGCAGC GGCTGTACTG
1201 CCTCCAGAT TTTCTTCTTC TGACCTGCG GGGCTGAGAG GATGAGCAGG GTGGGGCCAA
1261 GGGCTCAATC CTGAACGGCC CTCATTGCTG TTCCAATCCC ACAACACATA CCCACAGCAG
1321 GTCAGACAC GCATTTCGAC CATGCGCACC AATAACGTGT CTTACCTGA TTGGGTGTGG
1381 CAGGTCCGT GGACAGGAGT GCCTCGTCCC CCGCCAGAC CCGCTCCCC GTACAGCGCG
1441 CGTCCGGGAC CCGCAGCGGC TCCACCGCGG TGTGATCCGC GTTGGCGGCG CAGAGCAGCA
1501 TCCCAGCCGA TTTGACCCCG CGCATGCTCC GAGGCTTGAG GTTGGCCAGC ACCACCACCC
1561 GCCGGCCGAC AAGGTCCTCC AGGGTCACGT GCCGGACCAG GCCACTCACG ATGGTGCAGG
1621 GGCCCCCTC CTCGCCGAGG TCGATCTGCT CGACGTACAG ACTGCGACAT GCGTGGCGAG
1681 TGGTCATCAG AAGGAAGCAG GTGTGCAGAA GGGGCACGTG GTTGGTATTG AGAGTAGCCA
1741 AAGCTTTGTG CCAATCAGAA AGTCAACGCA GCTGCCTGCC TGGCTCGCGT ACGGTACCTT
1801 TTCTTGCGTG ATGACACTTC CAGCAAAAGG TAGGGCGGGC TGCGAGACGG CTTCCCGGCG
1861 CTGCATGCAA CACCGATGAT GCTTCGACCC CCCGAAGCTC CTTGCGGGCT CTATGGGCGC
1921 TCCGATGCCG CTCCAGGGCG AGCGCTGTTT AAATAGCCAG GCCCCCGATT GCAAAGACAT
1981 TATAGCGAGC TACCAAAGCC ATATTCAAAC ACCTAGATCA CTACCACTTC TACACAGGCC
2041 ACTCGAGCTT GTGATCGCAC TCCGCTAAGG GGGCGCTCT TCTCTTCGT TTCAGTCACA
2101 ACCCGCAAAC TCTAGAATAT CAATGCTGCT GCAGGCCTTC CTGTTCCTGC TGGCCGGCTT
2161 CGCCGCCAAG ATCAGCGCCT CCATGACGAA CGAGACGTCC GACCGCCCC TGGTGCACCT
2221 CACCCCCAAC AAGGGCTGGA TGAACGACCC CAACGGCCTG TGGTACGACG AGAAGGACGC
2281 CAAGTGGCAC CTGTACTTCC AGTACAACCC GAACGACACC GTCTGGGGGA CGCCCTTGTT
2341 CTGGGGCCAC GCCACGTCCG ACGACCTGAC CAACTGGGAG GACCAGCCCA TCGCCATCGC
2401 CCCGAAGCGC AACGACTCCG GCGCCTTCTC CGGCTCCATG GTGGTGGACT AGACAACAC
2461 CTCCGGCTTC TTCAACGACA CCATCGACCC GCGCCAGCGC TCGTGGCCA TCTGGACCTA
2521 CAACACCCCG GAGTCCGAGG AGCAGTACAT CTCCTACAGC CTGGACGGCG GCTACACCTT
2581 CACCGAGTAC CAGAAGAACC CCGTGCTGGC CGCCAACCTC ACCCAGTTCC GCGACCCGAA
2641 GGTCTTCTGG TACGAGCCCT CCCAGAAGTG GATCATGACC GCGGCCAAGT CCCAGGACTA
2701 CAAGATCGAG ATCTACTCCT CCGACGACCT GAAGTCCTGG AAGCTGGAGT CCGCGTTGCG
2761 CAACGAGGGC TTCCTCGGCT ACCAGTACGA GTGCCCGGCG CTGATCGAGG TCCCCACCGA
2821 GCAGGACCCC AGCAAGTCCT ACTGGGTGAT GTTCATCTCC ATCAACCCCG GCGCCCCGGC
2881 CGGCGGTCC TTCAACCACT ACTTCGTCGG CAGTTTCAAC GGCACCCACT TCGAGGCCTT
2941 CGACAACCAG TCCCGCGTGG TGGACTTCGG CAAGGACTAC TACGCCCTGC AGACCTTCTT
3001 CAACACCGAC CCGACCTACG GGAGCGCCCT GGGCATCGCG TGGGCCTCCA ACTGGGAGTA
3061 CTCCGCCTTC GTGCCCACCA ACCCCTGGCG CTCCTCCATG TCCCTCGTGC GCAAGTTCTC
3121 CCTCAACACC GAGTACCAGG CCAACCCGGA GACGGAGCTG ATCAACCTGA AGGCCGAGCC
3181 GATCCTGAAC ATCAGCAACG CCGGCCCTG GAGCCGGTTC GCCACCAACA CCACGTTGAC
3241 GAAGGCCAAC AGCTACAACG TCGACCTGTC CAACAGCACC GGCACCCCTG AGTTCGAGCT
3301 GGTGTACGCC GTCAACACCA CCCAGACGAT CTCCAAGTCC GTGTTCGCGG ACCTCTCCCT
3361 CTGGTTCAAG GGCCTGGAGG ACCCCGAGGA GTACCTCCGC ATGGGCTTCG AGGTGTCCGC
3421 GTCCTCCTTC TTCCTGGACC GCGGGAACAG CAAGGTGAAG TTCGTGAAGG AGAACCCTA
3481 CTTACCAAC CGCATGAGCG TGAACAACCA GCCCTTCAAG AGCGAGAACG ACCTGTCTTA

```

3541	CTACAAGGTG	TACGGCTTGC	TGGACCAGAA	CATCCTGGAG	CTGTACTTCA	ACGACGGCGA
3601	CGTCGTGTCC	ACCAACACCT	ACTTCATGAC	CACCGGGAAC	GCCCTGGGCT	CCGTGAACAT
3661	GACGACGGGG	GTGGACAACC	TGTTCTACAT	CGACAAGTTC	CAGGTGCGCG	AGGTCAAGTG
3721	ACAATTGGCA	GCAGCAGCTC	GGATAGTATC	GACACACTCT	GGACGCTGGT	CGTGTGATGG
3781	ACTGTTGCCG	CCACACTTGC	TGCCTTGACC	TGTGAATATC	CCTGCCGCTT	TTATCAAACA
3841	GCCTCAGTGT	GTTTGATCTT	GTGTGTACGC	GCTTTTGCGA	GTTGCTAGCT	GCTTGTGCTA
3901	TTTGCGAATA	CCACCCCCAG	CATCCCCTTC	CCTCGTTTCA	TATCGCTTGC	ATCCCAACCG
3961	CAACTTATCT	ACGCTGTCCT	GCTATCCCTC	AGCGCTGCTC	CTGCTCCTGC	TCACTGCCCC
4021	TCGCACAGCC	TTGGTTTGGG	CTCCGCCTGT	ATTCTCCTGG	TACTGCAACC	TGTAAACCAG
4081	CACTGCAATG	CTGATGCACG	GGAAGTAGTG	GGATGGGAAC	ACAAATGGAG	GATCCCAGCT
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	TTCTTTTCTT	GCGCTATGAC	ACTTCCAGCA
4381	AAAGGTAGGG	CGGGCTGCGA	GACGGCTTCC	CGGCGCTGCA	TGCAACACCG	ATGATGCTTC
4441	GACCCCCCGA	AGCTCCTTCG	GGGCTGCATG	GGGCTCCGA	TGCCGCTCCA	GGGCGAGCGC
4501	TGTTTAAATA	GCCAGGCCCC	CGATTGCAAA	GACATTATAG	CGAGCTACCA	AAGCCATATT
4561	CAAACACCTA	GATCACTACC	ACTTCTACAC	AGGCCACTCG	AGCTTGTGAT	CGCACTCCGC
4621	TAAGGGGGCG	CCTCTTCCTC	TTCTGTTTCT	TCACAACCCG	CAAACACTAG	TATGGCCACC
4681	GCATCCACTT	TCTCGGCGTT	CAATGCCCGC	TGCGGCGACC	TGCGTCGCTC	GGCGGGCTCC
4741	GGGCCCCGGC	GCCCAGCGAG	GCCCCCTCCC	GTGCGCGGGC	GCGCCCCGGA	CTGGTCCATG
4801	CTGTTTCGCC	TGATCACCAC	CATCTTCTCC	GCCGCCGAGA	AGCAGTGGAC	CAACCTGGAG
4861	TGGAAGCCCA	AGCCCAACCC	CCCCCAGCTG	CTGGACGACC	ACTTCGGCCC	CCACGGCCTG
4921	GTGTTCCGCC	GCACCTTCGC	CATCCGCGAG	TACGAGGTGG	GCCCCGACCG	CTCCACCAGC
4981	ATCGTGGCCG	TGATGAACCA	CCTGCAGGAG	GCCGCCCTGA	ACCACGCCAA	GTCCGTGGGC
5041	ATCCTGGGCG	ACGGCTTCGG	CACCACCCTG	GAGATGTCCA	AGCGCGACCT	GATCTGGGTG
5101	GTGAAGCGCA	CCCACGTGGC	CGTGAGCGCG	TACCCCGCCT	GGGGCGACAC	CGTGGAGGTG
5161	GAGTGCTGGG	TGGGCGCCTC	CGGCAACAAC	GGCCGCGGCC	ACGACTTCCT	GGTGCGCGAC
5221	TGCAAGACCG	GCGAGATCCT	GACCCGCTGC	ACCTCCCTGA	GCGTGATGAT	GAACACCCGC
5281	ACCCGCCGCG	TGAGCAAGAT	CCCCGAGGAG	GTGCGCGGGC	AGATCGGCCC	CGCCTTCATC
5341	GACAACGTGG	CCGTGAAGGA	CGAGGAGATC	AAGAAGCCCC	AGAAGCTGAA	CGACTCCACC
5401	GCCGACTACA	TCCAGGGCGG	CCTGACCCCC	CGCTGGAACG	ACCTGGACAT	CAACCAGCAC
5461	GTGAACAACA	TCAAGTACGT	GGACTGGATC	CTGGAGACCG	TGCCCCACAG	CATCTTCGAG
5521	AGCCACCACA	TCTCCTCCTT	CACCATCGAG	TACCGCCGCG	AGTGCACCAT	GGACAGCGTG
5581	CTGCAGTCCC	TGACCACCGT	GAGCGGCGGC	TCCTCCGAGG	CCGGCCTGGT	GTGCGAGCAC
5641	CTGCTGCAGC	TGGAGGGCGG	CAGCGAGGTG	CTGCGCGCCA	AGACCGAGTG	GCGCCCCAAG
5701	CTGACCGACT	CCTTCCGCGG	CATCAGCGTG	ATCCCCGCGG	AGTCCAGCGT	GATGGACTAC
5761	AAGGACCACG	ACGGCGACTA	CAAGGACCAC	GACATCGACT	ACAAGGACGA	CGACGACAAG
5821	TGACTCGAGG	CAGCAGCAGC	TCGGATAGTA	TCGACACACT	CTGGACGCTG	GTCTGTGTAT
5881	GGACTGTTGC	CGCCACACTT	GCTGCCCTGA	CCTGTGAATA	TCCCTGCCGC	TTTTATCAAA
5941	CAGCCTCAGT	GTGTTTGATC	TTGTGTGTAC	GCGCTTTTGC	GAGTTGTCTAG	CTGCTTGTGC
6001	TATTTGCGAA	TACCACCCCC	AGCATCCCCT	TCCCTCGTTT	CATATCGCTT	GCATCCCAAC
6061	CGCAACTTAT	CTACGCTGTC	CTGCTATCCC	TCAGCGCTGC	TCCTGCTCCT	GCTCACTGCC
6121	CCTCGCACAG	CCTTGGTTTG	GGCTCCGCCT	GTATTCTCCT	GGTACTGCAA	CCTGTAAACC
6181	AGCACTGCAA	TGCTGATGCA	CGGGAAGTAG	TGGGATGGGA	ACACAAATGG	AAAGCTTGAG
6241	CTCGTACCCA	TCAGCATCCG	GGTGAATCTT	GGCCTCCAAG	ATATGGCCAA	TCCTCACATC
6301	CAGCTTGGCA	AAATCGACTA	GACTGTCTGT	AAGTGGGAAT	GTGGAGCACA	AGATTGCTTG
6361	TAGCGATCGA	CAGACTGGTG	GGGTACATTG	ACAGGTGGGC	AGCGCCGCAAT	CCGTGTGCC
6421	TGACGCGAGC	GCCGCCGGTT	GCTCGCCCCG	GCCTGCCGTC	AAAGAGCGGC	AGAGAAATCG
6481	GGAACCGAAA	ACGTCACATT	GCCTGATGTT	GTTACATGCT	GGACTAGACT	TTCTTGCGCT
6541	GGGTCTGCTC	CTCGCCAGGT	GCGCGACGCC	TCGGGGCTGG	GTGCGAGGGA	GCCGTGCGGC
6601	CACGCATTTG	ACAAGACCCA	AAGCTCGCAT	CTCAGACGGT	CAACCGTTCG	TATTATACAT
6661	TCAACATATG	GTACATACGC	AAAAAGCATG	CCAACGATGA	CAGCTCTTCG	CCCACATGTG
6721	AGCAAAAAGG	CAGCAAAAAG	CCAGGAACCG	TAAAAAGGCC	GCGTTGTCTG	CGTTTTTCCA
6781	TAGGTCCGCG	CCCCCTGACG	AGCATCACAA	AAATCGACGC	TCAAGTCAGA	GGTGGCGAAA
6841	CCCGACAGGA	CTATAAAGAT	ACCAGGCGTT	TCCCCCTGGA	AGCTCCCTCG	TGCGCTCTCC
6901	TGTTCCGACC	CTGCCGCTTA	CCGATACCTT	GTCCGCCTTT	CTCCCTTCGG	GAGCGGTGGC
6961	GCTTTCTCAT	AGCTCACGCT	GTAGGTATCT	CAGTTCGGTG	TAGGTCGTTT	GCTCCAAGCT
7021	GGGCTGTGTG	CACGAACCCC	CCGTTTCAGC	CGACCGCTGC	GCCTTATCCG	GTAACATATC
7081	TCTTGAGTCC	AACCCGGTAA	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 TTAAATCAAT
7501 CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC
7561 TATCTCAGCG ATCTGTCTAT TTCGTTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT
7621 AACTACGATA CGGGAGGGCT TACCATCTGG CCCAGTGCT GCAATGATAC CGCGAGACCC
7681 ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG
7741 AAGTGGTCC TCAACTTTAT CCGCTCCAT CCAGTCTATT AATTGTTGCC GGAAGCTAG
7801 AGTAAGTAGT TCGCCAGTTA ATAGTTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT
7861 GGTGTCACGC TCGTCGTTTG GTATGGCTTC ATTCAGCTCC GGTTCCTCAAC GATCAAGGCG
7921 AGTTACATGA TCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCGATCGT
7981 TGTCAGAAAGT AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTC
8041 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
8101 ATTCTGAGAA TAGTGTATGC GCGCACCAGG TTGCTCTTGC CCGGCGTCAA TACGGGATAA
8161 TACCGCGCCA CATAGCAGAA CTTTAAAGT GCTCATCATT GGAAAACGTT CTTGGGGGCG
8221 AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCC ATGTAACCCA CTCGTGCACC
8281 CAACTGATCT TCAGCATCTT TACTTTTAC CAGCGTTTCT GGTGAGCAA AAACAGGAAG
8341 GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTGAATAC TCATACTCTT
8401 CCTTTTCAA TATTATTGAA GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT
8461 TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCGCGC ACATTTCCCC GAAAAGTGCC
8521 AC

```

SEQ ID NO: 229
pSZ1283

```

1 CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG
61 AGTTCTTGAA GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAA
181 CCACCGCTGG TAGCGGTGGT TTTTTGTITT GCAAGCAGCA GATTACGCGC AGAAAAAAG
241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAACT
301 CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACCTAG ATCCTTTTAA
361 ATTAAAAATG AAGTTTTTAA TCATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
481 TTGCTGCAAT CCCCCTCGTG TAGATACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACC
601 AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT
661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTTCGC AGTTAATAGT TTGCGCAACG
721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGTTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGTCTCTT CGGTCTCTCG ATCGTTGTCA GAAGTAAGTT GGCCGCAAGT TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTATGCC ATCCGTAAGA TGCTTTTCTG
961 TGA CTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTGCCCCGGC GTC AATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT CTTACCGCTG TTGAGATCCA
1141 GTTCGATGTA ACCCACTCGT GCACCAAAC 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 TCCCGAAAA 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 TTAACATGCG GCATCAGAG CAGATTGTAC TGAGAGTGCA CCAATATGCG GTTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT CAGGTCGCGA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCG CCACTCTGCT TCGAGCGCGC
1921 CCGCGCTGTC GCCGCCAGCG CCTTGCCCTT TTCGCCGCG TCGTGCCTG CTGATGTC

```

1981	CATCACCAGG	TCCATGAGGT	CTGCCTTGCG	CCGGCTGAGC	CACTGCTTCG	TCCGGGCGGC
2041	CAAGAGGAGC	ATGAGGGAGG	ACTCCTGGTC	CAGGGTCCTG	ACGTGGTCGC	GGCTCTGGGA
2101	GCGGGCCAGC	ATCATCTGGC	TCTGCCGCAC	CGAGGCCGCC	TCCAACCTGGT	CCTCCAGCAG
2161	CCGCAGTCGC	CGCCGACCCT	GGCAGAGGAA	GACAGGTGAG	GGGGGTATGA	ATTGTACAGA
2221	ACAACCACGA	GCCTTGCTA	GGCAGAATCC	CTACCAGTCA	TGGCTTTACC	TGGATGACGG
2281	CCTGCGAACA	GCTGTCCAGC	GACCCTCGCT	GCGCCGCTT	CTCCGCACG	CTTCTTTCCA
2341	GCACCGTGAT	GGCGCGAGCC	AGCGCCGCAC	GCTGGCGCTG	CGCTTCGCCG	ATCTGAGGAC
2401	AGTCGGGGAA	CTCTGATCAG	TCTAAACCCC	CTTGCGCGTT	AGTGTGCCA	TCCTTTGCAG
2461	ACCGGTGAGA	GCCGACTTGT	TGTGCGCCAC	CCCCACACC	ACCTCCTCCC	AGACCAATTC
2521	TGTCACCTTT	TTGGCGAAGG	CATCGGCCTC	GGCCTGCAGA	GAGGACAGCA	GTGCCAGCC
2581	GCTGGGGGTT	GGCGGATGCA	CGCTCAGGTA	CCCTTTCTTG	CGCTATGACA	CTTCCAGCAA
2641	AAGGTAGGGC	GGGCTGCGAG	ACGGCTTCCC	GGCGCTGCAT	GCAACACCGA	TGATGCTTCG
2701	ACCCCCGAA	GCTCCTTCGG	GGCTGCATGG	GCGCTCCGAT	GCCGCTCCAG	GGCGAGCGCT
2761	GTTTAAATAG	CCAGGCCCCC	GATTGCAAAG	ACATTATAGC	GAGCTACCAA	AGCCATATTC
2821	AAACACCTAG	ATCACTACCA	CTTCTACACA	GGCCACTCGA	GCTTGTGATC	GCACTCCGCT
2881	AAGGGGCGC	CTCTTCCTCT	TCGTTTCACT	CACAACCCGC	AAACGGCGCG	CCATGCTGCT
2941	GCAGGCCTTC	CTGTTCCCTG	TGGCCGGCTT	CGCCGCCAAG	ATCAGCGCCT	CCATGACGAA
3001	CGAGACGTCC	GACCGCCCCC	TGGTGCACCT	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	TCAACGACA	CCATCGACCC
3301	GCGCCAGCGC	TGCGTGGCCA	TCTGGACCTA	CAACACCCCG	GAGTCCGAGG	AGCAGTACAT
3361	CTCCTACAGC	CTGGACGGCG	GCTACACCTT	CACCGAGTAC	CAGAAGAACC	CCGTGCTGGC
3421	CGCCAACTCC	ACCCAGTTCC	GCGACCCGAA	GGTCTTCTGG	TACGAGCCCT	CCCAGAAGTG
3481	GATCATGACC	GCGGCCAAGT	CCCAGGACTA	CAAGATCGAG	ATCTACTCCT	CCGACGACCT
3541	GAAGTCTTGG	AAGCTGGAGT	CCGCGTTTCG	CAACGAGGGC	TTCTTCGGCT	ACCAGTACGA
3601	GTGCCCCGGC	CTGATCGAGG	TCCCCACCGA	GCAGGACCCC	AGCAAGTCCT	ACTGGGTGAT
3661	GTTCACTCTC	ATCAACCCCG	GCGCCCCGGC	CGGCGGCTCC	TTCAACCAGT	ACTTCGTCGG
3721	CAGTTCAAC	GGCACCCTAC	TCGAGGCCCT	CGACAACCAG	TCCCGCTTGG	TGGACTTCGG
3781	CAAGGACTAC	TACGCCCTGC	AGACCTTCTT	CAACACCGAC	CCGACCTACG	GAGACGCCCT
3841	GGGCATCGCG	TGGGCCTCCA	ACTGGGAGTA	CTCCGCCTTC	GTGCCCACCA	ACCCCTGGCG
3901	CTCCTCCATG	TCCCTCGTGC	GCAAGTTCTC	CCTCAACACC	GAGTACCAGG	CCAACCCGGA
3961	GACGGAGCTG	ATCAACCTGA	AGGCCGAGCC	GATCCTGAAC	ATCAGCAACG	CCGGCCCCTG
4021	GAGCCGTTTC	GCCACCAACA	CCACGTTGAC	GAAGGCCAAC	AGCTACAACG	TCGACCTGTC
4081	CAACAGCACC	GGCACCCCTG	AGTTCGAGCT	GGTGTACGCC	GTCAACACCA	CCCAGACGAT
4141	CTCCAAGTCC	GTGTTTCGCG	ACCTCTCCCT	CTGGTTCAAG	GGCCTGGAGG	ACCCCGAGGA
4201	GTACCTCCGC	ATGGGCTTCG	AGGTGTCCCG	GTCTCTCTTC	TTCTTGGACC	GCGGGAACAG
4261	CAAGGTGAAG	TTCGTGAAGG	AGAACCCCTA	CTTCAACCAAC	CGCATGAGCG	TGAACAACCA
4321	GCCCTTCAAG	AGCGAGAACG	ACCTGTCCCTA	CTACAAGGTG	TACGGCTTGC	TGGACCGAGAA
4381	CATCTGGAG	CTGTACTTCA	ACGACGGCGA	CGTCGTGTCC	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	CATCCCCCTC
4741	CCTCGTTTCA	TATCGCTTGC	ATCCCAACCC	CAACTTATCT	ACGCTGTCTC	GCTATCCCTC
4801	AGCGCTGCTC	CTGCTCCTGC	TCACTGCCCG	TCGCACAGCC	TTGGTTTGGG	TCCCGCTTGT
4861	ATTCTCCTGG	TACTGCAACC	TGTAAACCAG	CACTGCAATG	CTGATGCACG	GGAAGTAGTG
4921	GGATGGGAAC	ACAAATGGAG	GATCCCCGCT	CTCGAACAGA	GCGCGCAGAG	GAACGCTGAA
4981	GGTCTCGCCT	CTGTGCGACC	TCAGCGCGGC	ATACACCACA	ATAACCACCT	GACGAATGCG
5041	CTTGGTTCTT	CGTCCATTAG	CGAAGCGTCC	GGTTACACA	CGTGCCACGT	TGGCGAGGTG
5101	GCAGGTGACA	ATGATCGGTG	GAGCTGATGG	TCGAAACGTT	CACAGCCTAG	GGATATCGAA
5161	TTCGGCCGAC	AGGACGCGCG	TCAAAGGTGC	TGGTCGTGTA	TGCCCTGGCC	GGCAGGTCGT
5221	TGCTGCTGCT	GGTTAGTGAT	TCCGCAACCC	TGATTTTGGC	GTCTTATTTT	GGCGTGGCAA
5281	ACGCTGGCGC	CCGCGAGCCG	GGCCGGCGGC	GATGCGGTGC	CCCACGGCTG	CCGGAATCCA
5341	AGGGAGGCAA	GAGCGCCCCG	GTCAGTTGAA	GGGCTTTACG	CGCAAGGTAC	AGCCGCTCCT
5401	GCAAGGCTGC	GTGGTGGAAT	TGGACGTGCA	GGTCCTGCTG	AAGTTCCTCC	ACCGCTCAC
5461	CAGCGGACAA	AGCACCGGTG	TATCAGGTCC	GTGTATCCCA	CTCTAAAGAG	CTCGACTACG
5521	ACCTACTGAT	GGCCCTAGAT	TCTTCATCAA	AAACGCCTGA	GACACTTGCC	CAGGATTGAA
5581	ACTCCTGAA	GGGACCACCA	GGGGCCCTGA	GTTGTTCTT	CCCCCGTGG	CGAGCTGCCA

5641 GCCAGGCTGT ACCTGTGATC GAGGCTGGCG GGAAAATAGG CTTCGTGTGC TCAGGTCATG
5701 GGAGGTGCAG GACAGCTCAT GAAACGCCAA CAATCGCACA ATTCATGTCA AGCTAATCAG
5761 CTATTTCTC TTCACGAGCT GTAATTGTCC CAAAATTCTG GTCTACCGGG GGTGATCCTT
5821 CGTGTACGGG CCTTCCCTC AACCTAGGT ATGCGCGCAT GCGGTCGCCG CGCAACTCGC
5881 GCGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGCAACCG GATGCGTGGC
5941 ACCTTTTTTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT CTGTCGCCAA
6001 CCCTAGGATC AGCGGCGTAG GATTTCGTAA TCATTCTGCC TGATGGGGAG CTACCGACTA
6061 CCCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC ATTCCATTCTG
6121 TGCCCAAGAC ATTTTCATTGT GGTGCGAAGC GTCCCCAGTT ACGCTCACCT GTTTCCCGAC
6181 CTCCTTACTG TTCTGTGCGC AGAGCGGGCC CACAGGCCGG TCGCAGCCAC TAGTATGGTG
6241 GTGGCCGCCG CCGCCAGCAG CGCCTTCTTC CCCGTGCCCG CCCCCCGCCC CACCCCCAAG
6301 CCCGGCAAGT TCGGCAACTG GCCAGCAGC CTGAGCCAGC CCTCAAGCC CAAGAGCAAC
6361 CCCAACGGCC GCTTCCAGGT GAAGGCCAAC GTGAGCCCCC ACGGGCGCGC CCCCAGGCC
6421 AACGGCAGCG CCGTGAGCCT GAAGTCCGGC AGCCTGAACA CCCTGGAGGA CCCCCCAGC
6481 AGCCCCCCCC CCCGCACCTT CCTGAACCAG CTGCCCCACT GGAGCCGCTT GCGCACCGCC
6541 ATCACCACCG TGTTCTGTGGC CGCCGAGAAG CAGTTCACCC GCCTGGACCG CAAGAGCAAG
6601 CGCCCCGACA TGCTGGTGGG CTGGTTCGGC AGCGAGACCA TCGTGCAGGA CGGCCTGGTG
6661 TTCCGCGAGC GCTTCAGCAT CCGCAGCTAC GAGATCGGCG CCGACCGCAC CGCCAGCATC
6721 GAGACCTGA TGAACCACCT GCAGGACACC AGCCTGAACC ACTGCAAGAG CGTGGGCCCTG
6781 CTGAACGACG GCTTCGGCCG CACCCCCGAG ATGTGCACCC GCGACCTGAT CTGGGTGCTG
6841 ACCAAGATGC AGATCGTGGT GAACCGCTAC CCCACCTGGG GCGACACCGT GGAGATCAAC
6901 AGCTGGTTCA CCGAGAGCGG CAAGATCGGC ATGGGCCGCG AGTGGCTGAT CAGCATGCTG
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 TGGAACGACT TCGACGTGAA CCAGCACGTG
7201 AGCAACGTGA AGTACATCGG CTGGATTCTG GAGAGCATGC CCACCGAGGT GCTGGAGACC
7261 CAGGAGCTGT GCAGCCTGAC CCTGGAGTAC CGCCGCGAGT GCGGCCGCGA GAGCGTGGTG
7321 GAGAGCGTGA CCAGCATGAA CCCAGCAAG GTGGGCGACC GCAGCCAGTA CCAGCACCTG
7381 CTGCGCCTGG AGGACGGCGC CGACATCATG AAGGGCCGCA CCGAGTGGCG CCCCAGAAGC
7441 GCCGGCACCA ACCGCGCCAT CAGCACCTGA TTAATTAACT CGAGGCAGCA CAGCTCGGA
7501 TAGTATCGAC ACACTCTGGA CGCTGGTCTG GTGATGGACT GTTGCCGCCA CACTTGCTGC
7561 CTTGACCTGT GAATATCCCT GCCGCTTTTA TCAAACAGCC TCAGTGTGTT TGATCTTGTG
7621 TGTACGCGCT TTTGCGAGTT GCTAGCTGCT TGTGCTATTT GCGAATACCA CCCCAGCAT
7681 CCCCTTCCCT CGTTTCATAT CGCTTGATC CCAACCGCAA CTATCTACG CTGTCTGTCT
7741 ATCCCTCAGC GCTGCTCCTG CTCTGCTCA CTGCCCCCTG CACAGCCTTG GTTTGGGCTC
7801 CGCCTGTATT CTCCTGGTAC TGCAACCTGT AAACCAGCAC TGCAATGCTG ATGCACGGGA
7861 AGTAGTGGGA TGGGAACACA AATGGAAAGC TTGAGCTCTT GTTTCCAGA AGGAGTTGCT
7921 CCTTGAGCCT TTCATTCTCA GCCTCGATAA CCTCCAAAGC CGCTCTAATT GTGGAGGGGG
7981 TTCGAATTTA AAAGCTTGA ATGTTGGTTC GTGCGTCTGG AACAAAGCCA CACTTGTTGC
8041 TCACTGGGAA AAGGACCATC AGCTCCAAAA AACTTGCCGC TCAAACCGCG TACCTCTGCT
8101 TTCGCGCAAT CTGCCCTGTT GAAATCGCCA CCACATTCAT ATTGTGACGC TTGAGCAGTC
8161 TGTAATTGCC TCAGAATGTG GAATCATCTG CCCCCTGTGC GAGCCCATGC CAGGCATGTC
8221 GCGGGCGAGG ACACCCGCCA CTCGTACAGC AGACCATTAT GCTACCTCAC AATAGTTCAT
8281 AACAGTGACC ATATTTCTCG AAGCTCCCCA ACGAGCACCT CCATGCTCTG AGTGGCCACC
8341 CCCCGGCCCT GGTGCTTGCG GAGGGCAGGT CAACCGGCAT GGGGCTACCG AAATCCCCGA
8401 CCGGATCCCA CCACCCCGC GATGGGAAGA ATCTCTCCC GGGATGTGGG CCCACCACCA
8461 GCACAACCTG CTGGCCCGAG CGAGCGTCAA ACCATACCAC ACAAATATCC TTGGCATCGG
8521 CCCTGAATTC CTTCTGCCGC TCTGCTACCC GGTGCTTCTG TCCGAAGCAG GGGTTGCTAG
8581 GGATCGCTCC GAGTCCGCAA ACCCTTGTCG CGTGGCGGGG CTTGTTGAG CTTGAAGAGC
8641 CTCTAGAGTC GACCTGCAGG CATGCAAGCT TGGCGTAATC ATGGTCATAG CTGTTTCTG
8701 TGTGAAATTG TTATCCGCTC ACAAATCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA
8761 AAGCTGGGGG TGCCTAATGA GTGAGCTAAC TCACATTAAT TGC GTTGCGC TCACTGCCCG
8821 CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA
8881 GAGGCGGTTT GCGTATTGGG CGCTCTTCCG CTTCTCTGCT CACTGACTCG CTGCGCTCGG
8941 TCGTTCGGCT GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG TTTATCCACG
9001 AATCAGGGGA TAACGCAGGA AAGAACATGT GAGCAAAAGG CAGCAAAAGG CCAGGAAGACC
9061 GTAAAAAGGC CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG CCCCCTGAC GAGCATCACA
9121 AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT
9181 TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCGCCTT ACCGGATACC
9241 TGTCCGCTT TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC

9301 TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTTCAGC
 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 AACTAGAAAG GACAGTATTT GGTATCTGCG
 121 CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAAACAAA
 181 CCACCGCTGG TAGCGGTGGT TTTTGTGTTT GCAAGCAGCA GATTACGCGC AGAAAAAAG
 241 GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA CGCTCAGTGG AACGAAAACCT
 301 CACGTAAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT CTTACACCTAG ATCCTTTTAA
 361 ATTAAAAATG AAGTTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
 421 ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCTG TCATCCATAG
 481 TTGCTGACT CCCCCTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA
 541 GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACCC
 601 AGCCAGCCGG AAGGGCCGAG CGCAGAAAGT GTCTTGCAAC TTTATCCGCC TCCATCCAGT
 661 CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
 721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTGGGTATG GCTTCATTCA
 781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
 841 TTAGCTCCTT CGGTCTCTCG ATCGTTGTCA GAAGTAAGTT GGCCGCAAGT TTATCACTCA
 901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTATGCC ATCCGTAAGA TGCTTTTCTG
 961 TGAATGTTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
 1021 CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
 1081 TCATTGGAAG ACGTTCTTCG GGGCGAAAAA TCTCAAGGAT CTTACCGCTG TTGAGATCCA
 1141 GTTCGATGTA ACCCACTCGT GCACCCAACG GATCTTCAGC ATCTTTTACT TTCACGACCG
 1201 TTTCTGGGTG AGCAAAACAA GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGCGCAGAC
 1261 GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCAAT TTGACGGTT
 1321 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTT
 1381 CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA AACCATTATT ATCATGACAT
 1441 TAACCTATAA AAATAGGCGT ATCAGGAGGC CCTTTCGTCT CGCGCGTTTC GGTGATGACG
 1501 GTGAAACCTT CTGACACATG CAGCTCCCGG AGACGGTCAC AGCTTGCTCT TAAGCGGATG
 1561 CCGGGAGCAG ACAAGCCCGT CAGGGCGCGT CAGCGGGTGT TGGCGGGTGT CCGGGCTGGC
 1621 TTAACATATG GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
 1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATTG AGGCTGCGCA
 1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTCGTATT ACGCCAGCTG CGGAAAGGGG
 1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
 1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCCGCCTGG AGCTGGTGCA GAGCATGGGG
 1921 CAGTTTGGCG 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 GCGCGACGAC
 2161 ACGGGCGCGG CTGACCACCC GTGGGCGCGC TGGACGCGG AGTGGGTGGC CGAGGAGAAC
 2221 CGGCACGGCG ACCTGCTGAA CAAGTACTGT TGGCTGACGG GGCGCGTCAA CATGCGGGCC
 2281 GTGGAGGTGA CCATCAACAA CCTGATCAAG AGCGGCATGA ACCCGAGAG GACAAACAA
 2341 CCTTACTTGG GCTTCGTCTA CACTCTCTTC CAGGAGCGCG CCACCAAGTA GGTACCTTT
 2401 CTTGCGCTAT GACACTTCCA GCAAAAAGTA GGGCGGGCTG CGAGACGGCT TCCCGGCGCT
 2461 GCATGCAACA CCGATGATGC TTCGACCCCG CGAAGCTCCT TCGGGGCTGC ATGGGCGCTC
 2521 CGATGCCGCT CCAGGGCGAG CGCTGTTTAA ATAGCCAGGC CCCCATTGTC AAAGACATTA
 2581 TAGCGAGCTA CCAAAGCCAT ATTCAAACAC CTAGATCACT ACCACTTCTA CACAGGCCAC
 2641 TCGAGCTTGT GATCGCACTC CGCTAAGGGG GCGCCTCTTC CTCTTCGTTT CAGTCACAAC
 2701 CCGCAAACCT TAGAATATCA ATGCTGCTGT AGGCCTTCCT GTTCTGTCTG GCCGGCTTCG
 2761 CCGCCAAGAT CAGCGCCTCC ATGACGAACG AGACGTCCGA CCGCCCCCTG GTGCACTTCA
 2821 CCCCACAA GGGCTGGATG AACGACCCCA ACGGCTGTG GTACGAGAG AAGGAGGCCA
 2881 AGTGGACCTT GACTTCCAG TACAACCCGA ACGACACCGT CTGGGGGACG CCTTGTCTCT
 2941 GGGGCCACGC CACGTCCGAC GACCTGACCA ACTGGGAGGA CCAGCCCATC GCCATCGCCC
 3001 CGAAGCGCAA CCACTCCGGC GCCTTCTCCG GCTCCATGGT GGTGGACTAC AACACACCT
 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 GCGTTGCCA
3361 ACGAGGGCTT CCTCGGCTAC CAGTACGAGT GCCCGGCCT GATCGAGGTC CCCACCGAGC
3421 AGGACCCAG CAAGTCCTAC TGGGTGATGT TCATCTCCAT CAACCCCGGC GCCCGGCCG
3481 GCGGCTCCTT CAACCAGTAC TTCGTCGGCA GCTTCAACGG CACCCACTTC GAGGCCTTCG
3541 ACAACCAGTC CCGCGTGGTG GACTTCGGCA AGGACTACTA CGCCCTGCAG ACCTTCTTCA
3601 ACACCGACCC GACCTACGGG AGCGCCCTGG GCATCGCGTG GGCTTCCAAC TGGGAGTACT
3661 CCGCCTTCGT GCCCACC AACCTGGCGCT CCTCCATGTC CCTCGTGC GC AAGTTCTCCC
3721 TCAACACCGA GTACCAGGCC AACC CGGAGCTGAT CAACCTGAAG GCCGAGCCGA
3781 TCCTGAACAT CAGCAACGCC GGCCCTTGGG GCGGTTTCGC CACCAACACC AC GTTGACGA
3841 AGGCCAACAG CTACAACGTC GACCTGTCCA ACAGCACC GG CACCTGGAG TTCGAGCTGG
3901 TGTACGCCGT CAACACCACC CAGACGATCT CCAAGTCCGT GTTCGCGGAC CTCTCCCTCT
3961 GGTTCAAGGG CCTGGAGGAC CCCGAGGAGT ACCTCCGCAT GGGCTTCGAG GTGTCCGCGT
4021 CCTCCTTCTT CCTGGACCGC GGGAACAGCA AGGTGAAGTT CGTGAAGGAG AACCCCTACT
4081 TCACCAACCG CATGAGCGTG AACCAACGAG CTTTCAAGAG CGAGAACGAC CTGTCTTACT
4141 ACAAGGTGTA CGGCTTGCTG GACCAGAACA TCCTGGAGCT GTACTTCAAC GACGGCGACG
4201 TCGTGTCCAC CAACACCTAC TTCTATGACCA CCGGGAACGC CCTGGGCTCC GTGAACATGA
4261 CGACGGGGGT GGACAACCTG TTCTACATCG ACAAGTTCCA GGTGCGCGAG GTCAAGTGAC
4321 AATTGGCAGC AGCAGCTCGG ATAGTATCGA CACACTCTGG ACGCTGGTCG TGTGATGGAC
4381 TGTTCGCCGC ACACCTTGCT CTTGACCTG TGAATATCCC TGCCGCTTTT ATCAAACAGC
4441 CTCAGTGTGT TTGATCTTGT GTGTACGCGC TTTTGCGAGT TGCTAGCTGC TTGTGCTATT
4501 TGCGAATACC ACCCCCAGCA TCCCCTTCCC TCGTTTCATA TCGCTTGCAT CCCAACCAGCA
4561 ACTTATCTAC GCTGTCTTGC TATCCCTCAG CGCTGCTCCT GCTCCTGCTC ACTGCCCTC
4621 GCACAGCCTT GGTTCGGGCT CCGCCTGTAT TCTCCTGGTA CTGCAACCTG TAAACCAGCA
4681 CTGCAATGCT GATGCACGGG AAGTAGTGGG ATGGGAACAC AAATGGAGGA TCCCGCGTCT
4741 CGAACAGAGC GCGCAGAGGA ACCTGGAAGG TCTCGCCTCT GTCGCACCTC AGCGCGGCAT
4801 ACACCACAAT AACCACCTGA CGAATGCGCT TGGTTCCTCG TCCATTAGCG AAGCGTCCGG
4861 TTCACACACG TGCCACGTTG GCGAGGTGGG AGGTGACAAT GATCGGTGGA GCTGATGGTC
4921 GAAACGTTCA CAGCCTAGGG ATATCGAATT CGGCCGACAG GACGCGCGTC AAAGGTGCTG
4981 GTCGTGTATG CCCTGGCCGG CAGGTCTGTTG CTGCTGCTGG TTAGTGATTG CAGAACCTG
5041 ATTTTGCGCT CTTATTTTGG CGTGGCAAAC GCTGGCGCCC GCGAGCCGGG CCGGCGGCGA
5101 TGCGGTGCCC CACGGCTGCC GGAATCCAAG GGAGGCAAGA GCGCCCGGGT CAGTTGAAGG
5161 GCTTTACGCG CAAGGTACAG CCGCTCCTGC AAGGCTGCGT GGTGGAATTG GACGTGCAGG
5221 TCCTGCTGAA GTTCTCTCCAC CGCTCACC ACGGACAAAG CACCGGTGTA TCAGGTCCGT
5281 GTCATCCACT CTAAAGAACT CGACTACGAC CTACTGATGG CCCTAGATTG TTCATCAAAA
5341 ACGCTGAGA CACTTGCCCA GGATTGAAAC TCCCTGAAGG GACCACCAGG GGCCCTGAGT
5401 TGTTCTTCC CCCCCTGGCG AGCTGCCAG CAGGCTGTAC CTGTGATCGA GGCTGGCGGG
5461 AAAATAGGCT TCGTGTGCTC AGGTCAATGG AGGTGACGGA CAGCTCATGA AAGCCTAACA
5521 ATCGACAAT TCATGTCAAG CTAATCAGCT ATTTCTCTT CACGAGCTGT AATTGTCCCA
5581 AAATTCTGGT CTACCGGGG TGATCCTTCG TGTACGGGCC CTTCCCTCAA CCCTAGGTAT
5641 GCGCGCATGC GGTGCGCGC CAACTCGCGC GAGGGCCGAG GGTTCGGGAC GGGCCGTCCC
5701 GAAATGCAGT TGCACCCGGA TCGTGGCAC CTTTTTTGCG ATAATTTATG CAATGGACTG
5761 CTCTGCAAAA TTCTGGCTCT GTCGCCAACC CTAGGATCAG CGGCGTAGGA TTTCGTAATC
5821 ATTCGTCTTG ATGGGGAGCT ACCGACTACC CTAATATCAG CCCGACTGCC TGACGCCAGC
5881 GTCCACTTTT GTGCACACAT TCCATTCTGT CCCAAGACAT TTCATTGTGG TGCGAAGCGT
5941 CCCAGTTAC GCTCACCTGT TTCCCGACCT CTTTACTGTT CTGTGACAG AGCGGGCCCCA
6001 CAGCCGGTC GCAGCCACTA GTATGGTGGT GGCCGCCGCC CCTTCTTCCC
6061 CGTGCCCGCC CCCC GCCCAAGCC CGGCAAGTTC GGCAACTGGC CGCAGAGCCT
6121 GAGCCAGCCC TTCAAGCCCA AGAGCAACCC CAACGGCCGC TTCCAGGTGA AGGCCAACGT
6181 GAGCCCCCAC GGGCGCGCCC CCAAGGCCAA CGGCAGCGCC GTGAGCCTGA AGTCCGGCAG
6241 CCTGAACACC CTGGAGGACC CCCCAGCAG CCCCCCCCC CGCACCTTCC TGAACCAGCT
6301 GCGCGACTGG AGCCGCCTGC GCACCGCCAT CACCACCGTG TTCGTGGCCG CCGAGAAGCA
6361 GTTACCCCGC CTGGACCGCA AGAGCAAGCG CCCCAGCATG CTGGTGGACT GGTTCGGCAG
6421 CGAGACCATC GTGCAGGACG GCCTGGTGTT CCGCGAGCGC TTCAGCATCC GCAGCTACGA
6481 GATCGGCGCC GACCGACCG CCAGCATCGA GACCTGATG AACCACCTGC AGGACACAG
6541 CCTGAACACC TGCAAGAGCG TGGGCTTGCT GAACGACGGC TTCGGCCGCA CCCCCAGAT
6601 GTGCACCCGC GACCTGATCT GGGTGTGAC CAAGATGCAG ATCGTGGTGA ACCGCTACCC
6661 CACCTGGGGC GACACCGTGG AGATCAACAG CTGGTTTACG CAGAGCGGCA AGATCGGCAT
6721 GGGCCGCGAG TGGCTGATCA GCGACTGCAA CACCGGCGAG ATCCTGGTGC GCGCCACCAG
6781 CGCCTGGGCC ATGATGAACC AGAAGACCCG CCGCTTCAGC AAGCTGCCCT GCGAGGTGCG

6841	CCAGGAGATC	GCCCCCACT	TCGTGGACGC	CCCCCCCCTG	ATCGAGGACA	ACGACCGCAA
6901	GCTGCACAA	TTTCGACGTGA	AGACCGGCGA	CAGCATCTGC	AAGGGCCTGA	CCCCCGGCTG
6961	GAACGACTTC	GACGTGAACC	AGCACGTGAG	CAACGTGAAG	TACATCGGCT	GGATTCTGGA
7021	GAGCATGCCC	ACCGAGGTGC	TGGAGACCCA	GGAGCTGTGC	AGCCTGACCC	TGGAGTACCG
7081	CCGCGAGTGC	GGCCGCGAGA	GCGTGGTGGA	GAGCGTGACC	AGCATGAACC	CCAGCAAGGT
7141	GGGCGACCGC	AGCCAGTACC	AGCACCTGCT	GCGCCTGGAG	GACGGCGCCG	ACATCATGAA
7201	GGGCCGACCC	GAGTGGCGCC	CCAAGAACGC	CGGCACCAAC	CGCGCCATCA	GCACCTGATT
7261	AATTAACCTG	AGGCAGCAGC	AGCTCGGATA	GTATCGACAC	ACTCTGGACG	CTGGTCGTGT
7321	GATGGACTGT	TGCCGCCACA	CTTGCTGCCT	TGACCTGTGA	ATATCCCTGC	CGCTTTTATC
7381	AAACAGCCTC	AGTGTGTTTG	ATCTTGTTTG	TACGCGCTTT	TGCGAGTTGC	TAGCTGCTTG
7441	TGCTATTTGC	GAATACCACC	CCCAGCATCC	CCTTCCCTCG	TTTCATATCG	CTTGTCATCCC
7501	AACCGCAACT	TATCTACGCT	GTCTTGCTAT	CCCTCAGCGC	TGCTCCTGCT	CCTGCTCACT
7561	GCCCCTCGCA	CAGCCTTGGT	TTGGGCTCCG	CCTGTATTCT	CCTGGTACTG	CAACCTGTAA
7621	ACCAGCACTG	CAATGCTGAT	GCACGGGAAG	TAGTGGGATG	GGAACACAAA	TGAAAGCTT
7681	GAGCTCCAGC	CACGGCAACA	CCGCGCGCCT	TGCGGCCGAG	CACGGCGACA	AGAACCTGAG
7741	CAAGATCTGC	GGGCTGATCG	CCAGCGACGA	GGGCCGGCAC	GAGATCGCCT	ACACGCGCAT
7801	CGTGGACGAG	TTCTTCCGCC	TCGACCCCGA	GGGCGCCGTC	GCCGCCCTACG	CCAACATGAT
7861	GCGCAAGCAG	ATCACCATGC	CCGCGCACCT	CATGGACGAC	ATGGGCCACG	GCGAGGCCAA
7921	CCCGGGCCGC	AACCTCTTCG	CCGACTTCTC	CGCGGTCGCC	GAGAAGATCG	ACGTCTACGA
7981	CGCCGAGGAC	TACTGCCGCA	TCCTGGAGCA	CCTCAACGCG	CGCTGGAAGG	TGGACGAGCG
8041	CCAGTCCAGC	GGCCAGGCCG	CCGCGGACCA	GGAGTACGTC	CTGGGCCTGC	CCCAGCGCTT
8101	CCGGAAGCTC	GCCGAGAAGA	CCGCGGCCAA	GCGCAAGCGC	GTCCGCGCGA	GGCCCGTCGC
8161	CTTCTCCTGG	AGAAGAGCCT	CTAGAGTCGA	CCTGCAGGCA	TGCAAGCTTG	GCGTAATCAT
8221	GGTCATAGCT	GTTTCTGTG	TGAAATTGTT	ATCCGCTCAC	AATTCCACAC	AACATACGAG
8281	CCGGAAGCAT	AAAGTGTAAG	GCCTGGGGTG	CCTAATGAGT	GAGCTAACTC	ACATTAATTG
8341	CGTTGCGCTC	ACTGCCCCTG	TTCCAGTCGG	GAAACCTGTC	GTGCCAGCTG	CATTAATGAA
8401	TCGGCCAACG	CGCGGGGAGA	GGCGTTTTCG	GTATTGGGCG	CTCTTCCGCT	TCCTCGCTCA
8461	CTGACTCGCT	GCGCTCGGTC	GTTCGGCTGC	GGCGAGCGGT	ATCAGCTCAC	TCAAAGGCGG
8521	TAATACGGTT	ATCCACAGAA	TCAGGGGATA	ACGCAGGAAA	GAACATGTGA	GCAAAAGGCC
8581	AGCAAAAGGC	CAGGAACCGT	AAAAAGGCCG	CGTTGCTGGC	GTTTTTCCAT	AGGCTCCGCC
8641	CCCCTGACGA	GCATCACAAA	AATCGACGCT	CAAGTCAGAG	GTGGCGAAAC	CCGACAGGAC
8701	TATAAAGATA	CCAGGCGTTT	CCCCCTGGAA	GCTCCCTCGT	GCGCTCTCCT	GTTCCGACCC
8761	TGCCGCTTAC	CGGATACCTG	TCCGCCTTTC	TCCCTTCGGG	AAGCGTGGCG	CTTTCTCATA
8821	GCTCAGCTG	TAGGTATCTC	AGTTCGGTGT	AGGTCGTTTC	CTCCAAGCTG	GGCTGTGTGC
8881	ACGAACCCCC	CGTTCAGCCC	GACCCTGCG	CCTTATCCGG	TAACATATCGT	CTTGAGTCCA
8941	ACCCGGTAAG	ACACGACTTA	TCGC			

SEQ ID NO: 231
pSZ1413

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	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	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	CGCAGAAAGT	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	CGGTCTCCCT	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAATG	TTATCACTCA
901	TGTTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTCATGCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTAATCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAC	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTCGATGTA	ACCCACTCGT	GCACCCAACT	GATCTTCAGC	ATCTTTTACT	TTCACCAGCG

1201	TTTCTGGGTG	AGCAAAAACA	GGAAGGCAAA	ATGCCGCAAA	AAAGGGAATA	AGGGCGACAC
1261	GGAATGTTG	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	CAGTCCCCGG	AGACGGTCAC	AGCTTGTCTG	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGGCGCGT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATACCGCA	TCAGGCGCCA	TTCGCCATTG	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAGGGG
1801	GATGTGCTGC	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TTCCCAGTCA	CGACGTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCCGCCTGG	AGCTGGTGCA	GAGCATGGGG
1921	CAGTTTGCGG	AGGAGAGGGT	GCTCCCCGTG	CTGCACCCCG	TGGACAAGCT	GTGGCAGCCG
1981	CAGGACTTCC	TGCCCCGACCC	CGAGTCGCCC	GACTTCGAGG	ACCAGGTGGC	GGAGCTGCGC
2041	GCGCGCGCCA	AGGACCTGCC	CGACGAGTAC	TTTGTGGTGC	TGGTGGGCGA	CATGATCACG
2101	GAGGAGGCGC	TGCCGACCTA	CATGGCCATG	CTCAACACCT	TGGACGGTGT	GCGCGACGAC
2161	ACGGGCGCGG	CTGACCACCC	GTGGGCGCGC	TGGACGCGGC	AGTGGGTGGC	CGAGGAGAAC
2221	CGGCACGGCG	ACCTGCTGAA	CAAGTACTGT	TGGCTGACGG	GGCGCGTCAA	CATGCGGGCC
2281	GTGGAGGTGA	CCATCAACAA	CCTGATCAAG	AGCGGCATGA	ACCCGCGAGC	GGACAACAAC
2341	CCTTACTTGG	GCTTCGTCTA	CACCTCCTTC	CAGGAGCGCG	CCACCAAGTA	GGTACCCCTT
2401	CTTGCGTAT	GACACTTCCA	GCAAAAGGTA	GGGCGGGCTG	CGAGACGGCT	TCCCGCGCT
2461	GCATGCAACA	CCGATGATGC	TTCGACCCCG	CGAAGCTCCT	TCGGGGCTGC	ATGGGCGCTC
2521	CGATGCCGCT	CCAGGGCGAG	CGCTGTTTTA	ATAGCCAGGC	CCCCGATTGC	AAAGACATTA
2581	TAGCGAGCTA	CCAAAGCCAT	ATTCAAACAC	CTAGATCACT	ACCACTTCTA	CACAGGCCAC
2641	TCGAGCTTGT	GATCGCACTC	CGCTAAGGGG	GCGCCTCTTC	CTCTTCGTTT	CAGTCACAAC
2701	CCGCAAACCTC	TAGAATATCA	ATGTGCTGTC	AGGCCTTCCT	GTTCCTGCTG	GCCGGCTTCG
2761	CCGCCAAGAT	CAGCGCCTCC	ATGACGAACG	AGACGTCCGA	CCGCCCCCTG	GTGCACTTCA
2821	CCCCAACAAA	GGGCTGGATG	AACGACCCCA	ACGGCCTGTG	GTACGACGAG	AAGGACGCCA
2881	AGTGGCACCT	GTACTTCCAG	TACAACCCGA	ACGACACCGT	CTGGGGGACG	CCCTTGTTCT
2941	GGGGCCACGC	CACGTCCGAC	GACCTGACCA	ACTGGGAGGA	CCAGCCCATC	CCCATCGCCC
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	AGGACCCAG	CAAGTCCTAC	TGGGTGATGT	TCATCTCCAT	CAACCCCGGC	GCCCCGGCCG
3481	GCGGCTCCTT	CAACCAGTAC	TTCGTCGGCA	GCTTCAACGG	CACCCACTTC	GAGGCCCTTCG
3541	ACAACCAAGTC	CCGCGTGGTG	GACTTCGGGA	AGGACTACTA	CGCCCTGCAG	ACCTTCTTCA
3601	ACACCGACCC	GACCTACGGG	AGCGCCCTGG	GCATCGCGTG	GGCCTCCAAC	TGGGAGTACT
3661	CCGCCTTCGT	GCCCACCAAC	CCCTGGCGCT	CCTCCATGTC	CCTCGTGCGC	AAGTTCTCCC
3721	TCAACACCGA	GTACCAGGCC	AACCCGGAGA	CGGAGCTGAT	CAACCTGAAG	GCCGAGCCGA
3781	TCCTGAACAT	CAGCAACGCC	GGCCCTTGGA	GCCGGTTCGC	CACCAACACC	ACGTTGACGA
3841	AGGCCAACAG	CTACAACGTC	GACCTGTCCA	ACAGCACCGG	CACCCCTGGAG	TTCGAGCTGG
3901	TGTACGCCGT	CAACACCACC	CAGACGATCT	CCAAGTCCGT	GTTGCGGGAC	CTCTCCCTCT
3961	GTTTCAAGGG	CCTGGAGGAC	CCCGAGGAGT	ACCTCCGCAT	GGGCTTCGAG	GTGTCCGCGT
4021	CCTCCTTCTT	CCTGGACCGC	GGGAACAGTA	AGGTGAAGTT	CGTGAAGGAG	AATCCCTACT
4081	TCACCAACCG	CATGAGCGTG	AACAACCAGC	CCTTCAAGAG	CGAGAACGAC	CTGTCTACT
4141	ACAAGGTGTA	CGGCTTGCTG	GACCAGAACA	TCCTGGAGCT	GTACTTCAAC	GACGGCGACG
4201	TCGTGTCCAC	CAACACCTAC	TTCATGACCA	CCGGAACGC	CCTGGGCTCC	GTGAACATGA
4261	CGACGGGGGT	GGACAACCTG	TTCTACATCG	ACAAGTTCCA	GGTGCGCGAG	GTCAAGTGAC
4321	AATTGGCAGC	AGCAGCTCGG	ATAGTATCGA	CACACTCTGG	ACGCTGGTCG	TGTGATGGAC
4381	TGTTGCCGCC	ACACTTGCTG	CCTTGACCTG	TGAATATCCC	TGCCGCTTTT	ATCAAACAGC
4441	CTCAGTGTGT	TTGATCTTGT	GTGTACGCGC	TTTTGCGAGT	TGCTAGCTGC	TTGTGCTATT
4501	TGCGAATACC	ACCCCCAGCA	TCCCCTTCCC	TCGTTTCATA	TCGCTTGCTC	CCCAACCCGA
4561	ACTTATCTAC	GCTGTCTGAC	TATCCCTCAG	CGCTGCTCCT	GCTCCTGCTC	ACCAACCCCTC
4621	GCACAGCCTT	GGTTTGGGCT	CCGCCTGTAT	TCTCCTGGTA	CTGCAACCTG	TAAACCAGCA
4681	CTGCAATGCT	GATGCACGGG	AAGTAGTGGG	ATGGGAACAC	AAATGGAGGA	TCCGCGCTCT
4741	CGAACAGAGC	GCGCAGAGGA	ACGCTGAAGG	TCTCGCCTCT	GTCGCACCTC	AGCGCGGCAT
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 CGCAACCTTG
5041 ATTTTGGCGT CTTATTTTGG CGTGCCAAAC GCTGGCGCCC GCGAGCCGGG CCGGCGGCGA
5101 TCGGTGCCCC CACGGCTGCC GGAATCCAAG GGAGGCAAGA GCGCCCGGGT CAGTTGAAGG
5161 GCTTTACGCG CAAGGTACAG CCGCTCCTGC AAGGCTGCGT GGTGGAATTG GACGTGCAGG
5221 TCCTGCTGAA GTTCCTCCAC CGCTCACCA GCGGACAAAG CACCGGTGTA TCAGGTCCGT
5281 GTCATCCACT CTAAAGAACT CGACTACGAC CTACTGATGG CCCTAGATTG TTCATCAAAA
5341 ACGCCTGAGA CACTTGCCCA GGATTGAAAC TCCCTGAAGG GACCACCAGG GGCCCTGAGT
5401 TGTTCTTCC CCCCCTGGCG AGCTGCCAGC CAGGCTGTAC CTGTGATCGA GGCTGGCGGG
5461 AAAATAGGCT TCGTGTGCTC AGGTCATGGG AGGTGCAGGA CAGCTCATGA AACGCCAACA
5521 ATCGACAAT TCATGTCAAG CTAATCAGCT ATTTCTCTT CACGAGCTGT AATTGTCCCA
5581 AAATTCTGGT CTACCGGGGG TGATCCTTCG TGTACGGGCC CTTCCTCAA CCCTAGGTAT
5641 GCGCGCATGC GGTGCGCGCG CAACTCGCGC GAGGGCCGAG GGTTTGGGAC GGGCCGTCCC
5701 GAAATGCAGT TGCACCCGGA TCGTGGCAC CTTTTTTGCG ATAATTTATG CAATGGACTG
5761 CTCTGCAAAA TTCTGGCTCT GTCGCCAACC CTAGGATCAG CGGCGTAGGA TTTCTGAATC
5821 ATTCTGCTCG ATGGGGAGCT ACCGACTACC CTAATATCAG CCCGACTGCC TGACGCCAGC
5881 GTCCACTTTT GTGCACACAT TCCATTCTGT CCCAAGACAT TTCATTGTGG TGCGAAGCGT
5941 CCCCAGTTAC GCTCACCTGT TTCCCGACCT CCTTACTGTT CTGTGACAG AGCGGGCCCA
6001 CAGGCCGGTC GCAGCCACTA GTATGGCTAT CAAGACGAAC AGGCAGCCTG TGGAGAAGCC
6061 TCCGTTACG ATCGGGACGC TGCGBAAGC CATCCCCGCG CACTGTTTCG AGCGCTCGGC
6121 GCTTCGTGGG CGCGCCCCCA AGGCCAACGG CAGCGCCGTG AGCCTGAAGT CCGGCGCCCT
6181 GAACACCCTG GAGGACCCCC CCAGCAGCCC CCCCCCCGCG ACCTTCTGA ACCAGCTGCC
6241 CGACTGGAGC CGCTGCGCA CCGCCATCAC CACCGTGTTG GTGGCCGCCG AGAAGCAGTT
6301 CACCCGCTTG GACCGCAAGA GCAAGCGCCC CGACATGCTG GTGGACTGGT TCGGCAGCGA
6361 GACCATCGTG CAGGACGGCC TGGTGTTCG CGAGCGCTTC AGCATCCGCA GCTACGAGAT
6421 CGGCGCCGAC CGCACCGCCA GCATCGAGAC CCTGATGAAC CACCTGCAGG ACACCAGCCT
6481 GAACCACTGC AAGAGCGTGG GCCTGCTGAA CGACGGCTTC GGCCGCACCC CCGAGATGTG
6541 CACCCGCGAC CTGATCTGGG TGCTGACCAA GATGCAGATC GTGGTGAACC GCTACCCAC
6601 CTGGGGCGAC ACCGTGGAGA TCAACAGCTG GTTCAGCCAG AGCGGCAAGA TCGGCATGGG
6661 CCGCGAGTGG CTGATCAGCG ACTGCAACAC CGGCGAGATC CTGGTGCGCG CCACCAGCGC
6721 CTGGGCCATG ATGAACCAGA AGACCCGCGG CTTAGCAAG CTGCCCTGCG AGGTGCGCCA
6781 GGAGATCGCC CCCCCTTCG TGACGCCCC CCCCCTGATC GAGGACAACG ACCGCAAGCT
6841 GCACAAGTTC GACGTGAAGA CCGGCGACAG CATCTGCAAG GGCCTGACCC CCGGCTGGAA
6901 CGACTTCGAC GTGAACCAGC ACGTAGCAA CGTGAAGTAC ATCGGCTGGA TTCTGGAGAG
6961 CATGCCACC GAGGTGCTGG AGACCCAGGA GCTGTGCAGC CTGACCCTGG AGTACCGCCG
7021 CGAGTGGCGG CGCGAGAGCG TGGTGGAGAG CGTGACCAGC ATGAACCCCA GCAAGGTGGG
7081 CGACCGCAGC CAGTACCAGC ACCTGCTGCG CCTGGAGGAC GGCGCCGACA TCATGAAGGG
7141 CCGCACCGAG TGGCGCCCCA AGAACGCCGG CACCAACCGC GCCATCAGCA CCTGATTAA
7201 TAACTCGAGG CAGCAGCAGC TCGGATAGTA TCGACACACT CTGGACGCTG CTCGTGTGAT
7261 GGA CTGTTGC CGCCACACTT GCTGCCCTGA CCTGTGAATA TCCCTGCCGC TTTTATCAAA
7321 CAGCCTCAGT GTGTTTGATC TTGTGTGTAC GCGCTTTTGC GAGTTGCTAG CTGCTTGTGC
7381 TATTTGCGAA TACCACCCCC AGCATCCCCT TCCCTCGTTT CATATCGCTT GCATCCCAAC
7441 CGCAACTTAT CTACGCTGTC CTGCTATCCC TCAGCGCTGC TCCTGCTCCT GCTCACTGCC
7501 CCTCGCACAG CCTTGGTTTG GGCTCCGCCT GTATTCTCCT GGTACTGCAA CCTGTAAACC
7561 AGCACTGCAA TGCTGATGCA CGGGAAGTAG TGGGATGGGA ACACAAATGG AAAGCTTGAG
7621 CTCAGCCAC GGCAACACCG CGCGCCTTGG GGCCGAGCAC GGCGACAAGA ACCTGAGCAA
7681 GATCTGCGGG CTGATCGCCA GCGACGAGGG CCGGCACGAG ATCGCCTACA CGCGCATCGT
7741 GGACGAGTTC TTCCGCCTCG ACCCCGAGGG CGCCGTGCGC GCCTACGCCA ACATGATGCG
7801 CAAGCAGATC ACCATGCCCG CGCACCTCAT GGACGACATG GGCCACGGCG AGGCCAACCC
7861 GGGCCGCAAC CTCTTCGCCG ACTTCTCCGC GGTGCGCCGAG AAGATCGACG TCTACGACGC
7921 CGAGGACTAC TGCCGCATCC TGGAGCACCT CAACGCGCGC TGGAAGGTGG ACGAGCGCCA
7981 GGTCAGCGGC CAGGCCGCCG CGGACCAGGA GTACGTCTTG GGCTTCCCC AGCGCTTCCG
8041 GAAACTCGCC GAGAAGACCG CCGCCAAGCG CAAGCGCGTC GCGCGCAGGC CCGTCGCCCTT
8101 CTCTGGAGA AGAGCCTCTA GAGTCGACCT GCAGGCATGC AAGCTTGGCG TAATCATGGT
8161 CATAGCTGTT TCCTGTGTGA AATTGTTATC CGCTCACAAT TCCACACAAC ATACGAGCCG
8221 GAAGCATAAA GTGTAAAGCC TGGGTGCCT AATGAGTGAG CTAACCTACA TTAATGTCGT
8281 TGCGTCACT GCCCGCTTTC CAGTCGGGAA ACCTGTCGTG CCAGCTGCAT TAATGAATCG
8341 GCCAACGCGC GGGGAGAGGC GGTGTGCGTA TTGGGCGCTC TTCCGCTTCC TCGCTCACTG
8401 ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC GAGCGGTATC AGCTCACTCA AAGGCGGTAA
8461 TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC

8521	AAAAGGCCAG	GAACCGTAAA	AAGGCCGCGT	TGCTGGCGTT	TTTCCATAGG	CTCCGCCCCC
8581	CTGACGAGCA	TCACAAAAAT	CGACGCTCAA	GTCAGAGGTG	GCGAAACCCG	ACAGGACTAT
8641	AAAGATACCA	GGCGTTTCCC	CCTGGAAGCT	CCCTCGTGCG	CTCTCCTGTT	CCGACCCCTG
8701	CGCTTACCGG	ATACCTGTCC	GCCTTTCTCC	CTTCGGGAAG	CGTGCGCCTT	TCTCATAGCT
8761	CACGCTGTAG	GTATCTCAGT	TCGGTGTAGG	TCGTTTCGCTC	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	TTCGGA AAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT
421	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTCTG	TCATCCATAG
481	TTGCCGTGACT	CCCCGTCTGT	TAGATAACTA	CGATACGGGA	GGGCTTACCA	TCTGGCCCCA
541	GTGCTGCAAT	GATACCGCGA	GACCCACGCT	CACCGGCTCC	AGATTTATCA	GCAATAAACCC
601	AGCCAGCCGG	AAGGGCCGAG	CGCAGAAGTG	GTCTTGCAAC	TTTATCCGCC	TCCATCCAGT
661	CTATTAATTG	TTGCCGGGAA	GCTAGAGTAA	GTAGTTCGCC	AGTTAATAGT	TTGCGCAACG
721	TTGTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGGTATG	GCTTCATTCA
781	GCTCCGGTTC	CCAACGATCA	AGGCAGGTTA	CATGATCCCC	CATGTTGTGC	AAAAAAGCGG
841	TTAGTCTCCT	CGGTCTCTCC	ATCGTTGTCA	GAAGTAAGTT	GGCCGCGAGT	TTATCACTCA
901	TGGTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTATGCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAAC	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCGATGTA	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	CCTTTCTGCT	CGCGCGTTTC	GGTGATGACG
1501	GTGAAAACCT	CTGACACATG	CAGCTCCCGG	AGACGGTCAC	AGCTTGCTCG	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGCGCGCT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATACCGCA	TCAGGCGCCA	TTCGCCATTG	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAGGGG
1801	GATGTGCTGC	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TTCCCAGTCA	CGACGTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCGCCGCCG	CCACTCCTGC	TCGAGCGCGC
1921	CCGCGCGTGC	GCCGCCAGCG	CCTTGGCCTT	TTCGCCGCGC	TCGTGCGCGT	CGCTGATGTC
1981	CATCACCAGG	TCCATGAGGT	CTGCCTTGCG	CCGGCTGAGC	CACTGCTTCG	TCCGGGCGGC
2041	CAAGAGGAGC	ATGAGGGAGG	ACTCTTGGTC	CAGGGTCTTG	ACGTGGTCTG	GGCTCTGGGA
2101	GCGGGCCAGC	ATCATCTGGC	TCTGCCGCAC	CGAGGCCGCC	TCCAACCTGG	CCTCCAGCAG
2161	CCGCAGTCGC	CGCCGACCCT	GGCAGAGGAA	GACAGGTGAG	GGGGGTATGA	ATTGTACAGA
2221	ACAACCACGA	GCCTTGCTTA	GGCAGAATCC	CTACCAGTCA	TGGCTTTACC	TGGATGACGG
2281	CCTGCGAACA	GCTGTCCAGC	GACCTTCGCT	GCCGCCGCTT	CTCCGCGACG	CTTCTTTCCA
2341	GCACCGTGAT	GGCGCGAGCC	AGCGCCGCAC	GCTGGCGCTG	CGCTTCGCCG	ATCTGAGGAC
2401	AGTCGGGGAA	CTCTGATCAG	TCTAAACCCC	CTTGCGCGTT	AGTGTTGCCA	TCCTTTGCAG
2461	ACCGGTGAGA	GCCGACTTGT	TGTGCGCCAC	CCCCACACC	ACCTCCTCCC	AGACCAATTC
2521	TGTCACCTTT	TTGGCGAAGG	CAICGGCCTC	GGCCTGCAGA	GAGGACAGCA	GTGCCGACCC
2581	GCTGGGGGTT	GGCGGATGCA	CGCTCAGGTA	CCCTTTCTTG	CGCTATGACA	CTTCCAGCAA
2641	AAGGTAGGGC	GGGCTGCGAG	ACGGCTTCCC	GGCGCTGCAT	GCAACACCGA	TGATGCTTCG
2701	ACCCCCCGAA	GCTCCTTCGG	GGCTGCATGG	GCGCTCCGAT	GCCGCTCCAG	GGCGAGCGCT
2761	GTTTAAATAG	CCAGGCCCCC	GATTGCAAAG	ACATTATAGC	GAGCTACCAA	AGCCATATTC
2821	AAACACCTAG	ATCACTACCA	CTTCTACACA	GGCCACTCGA	GCTGTGATC	GCACTCCGCT
2881	AAGGGGGCGC	CTCTTCCTCT	TCGTTTCAGT	CACAACCCGC	AAACTCTAGA	ATATCAATGC

2941 TGCTGCAGGC CTTCTGTTC CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCAACGG CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCCGAACGA CACCGTCTGG GGGACGCCCT TGTCTGGGG 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 GCGGCTACA CCTTCACCGA GTACCAGAAG AACCCCGTGC
3421 TGGCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCCAGA
3481 AGTGGATCAT GACCGCGGCC AAGTCCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCGCGT TCGCCAACGA GGGCTTCCTC GGCTACCAGT
3601 ACGAGTGCCC CGGCCGTGATC GAGGTCCCCA CCGAGCAGGA CCCAGCAAG TCCTACTGGG
3661 TGATGTTTCA CTCCATCAAC CCCGCGGCC CCGCCGGCGG CTCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CCTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCTGGGCAT CGCGTGGGCC TCCAACGGG AGTACTCCGC CTTCTGCCCC ACCAACCCT
3901 GCGCTCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCC AGCCGATCCT GAACATCAGC AACGCCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTCGACC
4081 TGTCCAACAG CACCGGCACC CTGGAGTTCG AGCTGGTGTA CGCCGTCAAC ACCACCCAGA
4141 CGATCTCAA GTCCGTGTTT CCGGACCTCT CCCTCTGGTT CAAGGGCTG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCTC CTCTTCTG CACCGCGGGA
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 GGGGTGGAC AACCTGTTCT
4501 ACATCGACAA GTTCCAGGTG CGCGAGGTCA AGTGACAATT GGCAGCAGCA GCTCGGATAG
4561 TATCGACACA CTCTGGACGC TGGTCGTGTG ATGGACTGTT GCCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTTGTGTGT
4681 ACGCGCTTTT CCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC CCAGCATCCC
4741 CTTCCCTGCT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGTCTATC
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 TGCGCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGGTTCA CACACGTGCC ACCTTGGCGA
5101 GGTGGCAGGT GACAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 CGAATTCGGC CGACAGGACG CGCGTCAAAG GTGCTGGTCTG TGTATGCCCT GGCCGGCAGG
5221 TCGTTGCTGC TGCTGGTTAG TGATCCGCA ACCCTGATTT TGGCGTCTTA TTTTGGCGTG
5281 GCAAACGCTG CGGCCCGCGA GCCGGGCCGG CGGCGATGCG GTGCCCCACG GCTGCCGGA
5341 TCCAAGGGAG GCAAGAGCGC CCGGTCAGT 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 TGAAGTCCC TGAAGGGACC ACCAGGGGCC CTGAGTTGTT CCTTCCCCC GTGGCGAGCT
5641 GCCAGCCAGG CTGTACCTGT GATCGAGGCT GGCGGGAAAA TAGGCTTCGT GTGCTCAGGT
5701 CATGGGAGGT GCAGGACAGC TCATGAAACG CCAACAATCG CACAATTCAT GTCAAGCTAA
5761 TCAGCTATTT CCTCTTACG AGCTGTAATT TCCCAAAT TCTGGTCTAT CGGGGTGAT
5821 CCTTCGTGTA CGGGCCCTT CCTCAACCT AGGTATGCGC GCATGCGGTC GCCGCGCAAC
5881 TCGCGCAGG GCGGAGGTT TGGGACGGGC CGTCCCGAAA TGCAGTTGCA CCCGATGCG
5941 TGGCACCTTT TTTGCGATAA TTTATGCAAT GGACTGCTCT GCAAAATTCT GGCTCTGTCTG
6001 CCAACCCTAG GATCAGCGGC GTAGGATTTT GTAATCATTG GTCCTGATGG GGAGCTACCG
6061 ACTACCTTAA TATCAGCCCG ACTGCTGAC GCCAGCGTCC ACTTTTGTGC ACACATTCCA
6121 TTCGTGCCCC AGACATTTCA TTGTGGTGCG AAGCGTCCCC AGTTACGCTC ACCTGTTTCC
6181 CGACCTCCTT ACTGTTCTGT CGACAGAGCG GGCCACAGG CCGGTCGCG CCCATATGGC
6241 TTCCGCGGCA TTCACCATGT CGGCGTGCCC CGCGATGACT GGCAGGGCCC CTGGGGCAG
6301 TCGCTCCGGA CGGCCAGTCG CCACCCGCT GAGGTACGTA TTCCAGTGCC TGGTGGCCAG
6361 CTGCATCGAC CCCTGCGACC AGTACCGCAG CAGCGCCAGC CTGAGCTTCC TGGGCGACAA
6421 CGGCTTCGCC AGCCTGTTCG GCAGCAAGCC CTTTATGAGC AACCAGCGCC ACCGCCGCT
6481 GCGCCGCGCC AGCCACAGCG GCGAGGCCAT GGCCGTGGCC CTGCAGCCCG CCCAGGAGGC
6541 CGGCACCAAG AAGAAGCCCG 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 CCGCCGGCAA 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 TCGCCTGGG CGGCTTCGTG GCCTGCCGCG CCCTGAGCCA
7201 GCGCAACAGC GACCCACCA AGGCCAGCCG CCCCTGGGAC AGCAACCGCG ACGGCTTCGT
7261 GATGGGCGAG GCGCGCGGCG TGCTGTGCT GGAGGAGCTG GAGCAGCCCA AGAAGCGCGG
7321 CGCCACCATC TACGCCGAGT TCCTGGGCGG CAGCTTCACC TCGCAGCCCT ACCACATGAC
7381 CGAGCCCCAC CCCGAGGGCG CCGGCGTGAT CCTGTGCATC GAGAAGGCC TGGCCAGGC
7441 CGGCGTGAGC AAGGAGGACG TGAATACAT CAACGCCAC GCCACCAGCA CCAGCGCCGG
7501 CGACATCAAG GAGTACCAGG CCCTGGCCCG CTGCTTCGGC CAGAACAGCG AGCTGCGCGT
7561 GAACAGCACC AAGAGCATGA TCGCCACCT GCTGGGCGCC GCCGGCGGCG TGGAGGCCGT
7621 GACCTGGGTG CAGGCCATCC GCACCGGCTG GATTACCCC AACCTGAACC TGGAGGACCC
7681 CGACAAGGCC GTGGACGCCA AGCTGCTGGT GGGCCCCAAG AAGGAGCGCC TGAACGTGAA
7741 GGTGGGCGTG AGCAACAGCT TCGGCTTCGG CGGCCACAAC AGCAGCATCC TGTTCCGCCC
7801 CTGCAACGTG TGAATGCATA CGGAGCGTCG TCGGGGAGGG AGTGTGCCGA GCGGGGAGTC
7861 CCGTCTGTG CGAGGCCCGG CAGCTGACGC TGGCGAGCCG TACGCCCCGA GGGTCCCCCT
7921 CCCCTCCCCC CTCTTCCCCT TCCCTCTGAC GGCCGCGCCT GTTTTTCAT GTTCAGCGAC
7981 CTTAAGGATC TAAGTAAGAT TCGAAGCGCT CGACCGTGCC GGACGGACTG CAGCCCCATG
8041 TCGTAGTGAC CGCCAATGTA AGTGGGCTGG CGTTTCCCTG TACGTGACTC AACGTCACCTG
8101 CACGCGCACC ACCCTCTCGA CCGGAGGAC CAGGCATCGC GAGATACAGC GCGAGCCAGA
8161 CACGGAGTGC CGAGCTATGC GCACGCTCCA ACTAGATATC GAATTCGGCC GACAGGACGC
8221 GCGTCAAAGG TGCTGGTTCG GTATGCCCTG GCCGGCAGGT CGTTGCTGCT GCTGGTTAGT
8281 GATTCCGCAA CCCTGATTTT GCGCTCTTAT TTTGGCGTGG CAAACGCTGG CGCCCGCGAG
8341 CCGGGCCGCG GCGGATGCGG TGCCCCACGG CTGCCGGAAT CCAAGGAGGG CCAAGAGGCC
8401 CCGGTACGTT GAAGGGCTTT ACGCGCAAGG TACAGCCGCT CCTGCAAGGC TCGTGGTTGG
8461 AATTGGACGT GCAGGTCTCT CTGAAGTTCC TCCACCGCCT CACCAGCGGA CAAAGCACCG
8521 GTGTATCAGG TCCGTGTCTT CCACTCTAAA GAATCGACT ACGACCTACT GATGGCCCTA
8581 GATTCTTCAT CAAAAACGCC TGAGACACTT GCCCAGGATT GAAACTCCCT GAAGGGACCA
8641 CCAGGGGCCC TGAGTTGTTC CTTCCCCCGG TGGCGAGCTG CCAGCCAGGC TGTACCTGTG
8701 ATCGAGGCTG GCGGGAAAAT AGGCTTCGTG TGCTCAGGTC ATGGGAGGTG CAGGACAGCT
8761 CATGAAACGC CAACAATCGC ACAATTCATG TCAAGCTAAT CAGCTATTTT CTCTTCACGA
8821 GCTGTAATTG TCCCAAAATT CTGGTCTACC GGGGGTGATC CTTCTGTGAT GGGCCCTTCC
8881 CTCAACCCTA GGTATGCGCG CATGCGGTCTG CCGCGCAACT CGCGCGAGGG CCGAGGGTTT
8941 GGGACGGGCC GTCCCGAAAT GCAGTTGCAC CCGGATGCGT GGCACCTTTT TTGCGATAAT
9001 TTATGCAATG GACTGCTCTG CAAAATTCTG GCTCTGTCTG CAACCCTAGG ATCAGCGGCG
9061 TAGGATTTCT TAATCATTCG TCCTGATGGG GAGCTACCGA CTACCCTAAT ATCAGCCCGA
9121 CTGCCTGACG CCAGCGTCCA CTTTTGTGCA CACATTCCAT TCGTGCCCAA GACATTTTCT
9181 TGTGGTGCGA AGCGTCCCCA GTTACGCTCA CCTGTTTCCC GACCTCCTTA CTGTTCTGTC
9241 GACAGAGCGG GCCCACAGGC CGGTCTGACG CACTAGTATG GCCACCGCAT CCACCTTCTC
9301 GCGTTCAAT GCCCGCTGCG GCGACCTGCG TCGTCTGGCG GGCTCCGGGC CCCGCGCCCC
9361 AGCAGAGGCC CTCCCCGTGC GCGGGCGCGC CCCCAGGCC AACGGCAGCC CCGTGAGCCT
9421 GAAGTCCGGC AGCCTGAACA CCTGGAGGCA CCCCCCAGC AGCCCCCCC CCCGTACCTT
9481 CCTGAACAG CTGCCCCACT GGAGCCGCGT GCGCACCGCC ATCACCACCG TGTTCTGTGG
9541 CGCCGAGAAG CAGTTCACCC GCCTGGACCG CAAGAGCAAG CGCCCCGACA TGCTGGTGGA
9601 CTGGTTCGGC AGCGAGACCA TCGTGCAGGA CGGCCTGGTG TTCCGCGAGC GCTTCAGCAT
9661 CCGCAGCTAC GAGATCGGCG CCGACCGCAC CGCCAGCATC GAGACCTGA TGAACCACCT
9721 GCAGGACACC AGCCTGAACC ACTGCAAGAG CGTGGGCGCT CTGAACGACG GCTTCGGCCG
9781 CACCCCCGAG ATGTGCACCC GCGACCTGAT CTGGGTGCTG ACCAAGATGC AGATCGTGGT
9841 GAACCGCTAC CCCACCTGGG GCGACACCGT GGAGATCAAC AGCTGGTTCA GCCAGAGCGG
9901 CAAGATCGGC ATGGGCGCGG AGTGGCTGAT CAGCGACTGC AACACCGCGG AGATCTGGT
9961 GCGCGCCACC AGCGCCTGGG CCATGATGAA CCAGAAGACC CGCCGCTTCA CGAAGCTGCC
10021 CTGCGAGGTG GCCCAGGAGA TCGCCCCCA CTTCTGTGGC GCCCCCCCCG TGATCGAGGA
10081 CAACGACCGC AAGCTGCACA AGTTCGACGT GAAGACCGGC GACAGCATCT GCAAGGGCCT
10141 GACCCCCGGC 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	CCCCAAGAAC	GCCGGCACCA	ACCGCGCCAT
10441	CAGCACCTGA	TTAATTAAC	CGAGGCAGCA	GCAGCTCGGA	TAGTATCGAC	ACACTCTGGA
10501	CGCTGGTCGT	GTGATGGACT	GTTGCCGCCA	CACTTGCTGC	CTTGACCTGT	GAATATCCCT
10561	GCCGCTTTTA	TCAAACAGCC	TCAGTGTGTT	TGATCTTGTTG	TGTACGCGCT	TTTGCAGATT
10621	GCTAGCTGCT	TGTGCTATTT	GCGAATACCA	CCCCCAGCAT	CCCCTTCCCT	CGTTTCATAT
10681	CGCTTGCAATC	CCAACCGCAA	CTTATCTACG	CTGTCCTGCT	ATCCCTCAGC	GCTGCTCCTG
10741	CTCCTGCTCA	CTGCCCCCTG	CACAGCCTTG	GTTTGGGCTC	CGCCTGTATT	CTCCTGGTAC
10801	TGCAACCTGT	AAACCAGCAC	TGCAATGCTG	ATGCACGGGA	AGTAGTGGGA	TGGGAACACA
10861	AATGGAAAGC	TTGAGCTCTT	GTTTTCCAGA	AGGAGTTGCT	CCTTGAGCCT	TTCATTCTCA
10921	GCCTCGATAA	CCTCCAAAGC	CGCTCTAATT	GTGGAGGGGG	TTCGAATTTA	AAAGCTTGGA
10981	ATGTTGGTTC	GTGCGTCTGG	AACAAGCCCA	GACTTGTGTC	TCACTGGGAA	AAGGACCATC
11041	AGCTCCAAAA	AACTTGCCGC	TCAAACCGCG	TACCTCTGCT	TTCGCGCAAT	CTGCCCTGTT
11101	GAAATCGCCA	CCACATTTCAT	ATTGTGACGC	TTGAGCAGTC	TGTAATTGCC	TCAGAATGTG
11161	GAATCATCTG	CCCCCTGTGC	GAGCCCATGC	CAGGCATGTC	GCGGGCGAGG	ACACCCGCCA
11221	CTCGTACAGC	AGACCATTAT	GCTACCTCAC	AATAGTTCAT	AACAGTGACC	ATATTTCTCG
11281	AAGCTCCCCA	ACGAGCACCT	CCATGCTCTG	AGTGGCCACC	CCCCGGCCCT	GGTGCTTGCG
11341	GAGGGCAGGT	CAACCGGCAT	GGGGCTACCG	AAATCCCCGA	CCGGATCCCCA	CCACCCCCGC
11401	GATGGGAAGA	ATCTCTCCCC	GGGATGTGGG	CCCACCACCA	GCACAACCTG	CTGGCCCAGG
11461	CGAGCGTCAA	ACCATAACCAC	ACAAATATCC	TTGGCATCGG	CCCTGAATTC	CTTCTGCCGC
11521	TCTGCTACCC	GGTGCTTCTG	TCCGAAGCAG	GGGTTGCTAG	GGATCGCTCC	GAGTCCGCAA
11581	ACCCTTGTCG	CGTGCGGGGG	CTTGTTTCGAG	CTTGAAGAGC	CTCTAGAGTC	GACCTGCAGG
11641	CATGCAAGCT	TGGCGTAATC	ATGGTCATAG	CTGTTTCCTG	TGTGAAATTG	TTATCCGCTC
11701	ACAATTCAC	ACAACATACG	AGCCGGAAGC	ATAAAGTGTA	AAGCCTGGGG	TGCCTAATGA
11761	GTGAGCTAAC	TCACATTAAT	TGCGTTGCGC	TCCTGCCCCG	CTTCCAGTC	GGGAAACCTG
11821	TCGTGCCAGC	TGCATTAATG	AATCGGCCAA	CGCGCGGGGA	GAGGCGGTTT	GCGTATTGGG
11881	CGCTCTTCCG	CTTCTCTCGT	CACTGACTCG	CTGCGCTCGG	TCGTTCCGGT	GCGGCGAGCG
11941	GTATCAGCTC	ACTCAAAGGC	GGTAATACGG	TTATCCACAG	AATCAGGGGA	TAACGCAGGA
12001	AAGAACATGT	GAGCAAAAGG	CCAGCAAAAG	GCCAGGAACC	GTAAAAAGGC	CGCGTTTGCTG
12061	GCGTTTTTCC	ATAGGCTCCG	CCCCCTTGAC	GAGCATCACA	AAAATCGACG	CTCAAGTCAG
12121	AGGTGGCGAA	ACCCGACAGG	ACTATAAAGA	TACCAGGCGT	TTCCCCCTGG	AAGCTCCCTC
12181	GTGCGCTCTC	CTGTTCCGAC	CCTGCCGCTT	ACCGGATACC	TGTCCGCCTT	TCTCCCTTCG
12241	GGAAGCGTGG	CGCTTTCTCA	TAGCTCACGC	TGTAGGTATC	TCAGTTCGGT	GTAGGTCGTT
12301	CGCTCCAAGC	TGGGCTGTGT	GCACGAACCC	CCCGTTCAGC	CCGACCGCTG	CGCCTTATCC
12361	GGTAACATATC	GTCTTGAGTC	CAACCCGGTA	AGACACGACT	TATCGC	

SEQ ID NO: 233

pSZ1500

1	CACTGGCAGC	AGCCACTGGT	AACAGGATTA	GCAGAGCGAG	GTATGTAGGC	GGTGCTACAG
61	AGTTCCTGAA	GTGGTGGCCT	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAAC
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTACCTTAG	ATCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAT	GTATATATGA	GTAAACCTTG	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	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	TTAGTTCCTT	CGGTCCCTCC	ATCGTTGTCA	GAAGTAAGTT	GGCCGAGTG	TTCATCTCA
901	TGGTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTCATGCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTAATCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAC	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCATGTA	ACCCACTCGT	GCACCCAACT	GATCTTCAGC	ATCTTTTACT	TTCACCGAGC

1201	TTTCTGGGTG	AGCAAAAACA	GGAAGGCAAA	ATGCCGCAAA	AAAGGGAATA	AGGGCGACAC
1261	GGAATGTTG	AATACTCATA	CTCTTCCTTT	TTCAATATTA	TTGAAGCATT	TATCAGGGTT
1321	ATTGTCTCAT	GAGCGGATAC	ATATTTGAAT	GTATTTAGAA	AAATAAACAA	ATAGGGGTTT
1381	CGCGCACATT	TCCCCGAAAA	GTGCCACCTG	ACGTCTAAGA	AACCATTTAT	ATCATGACAT
1441	TAACCTATAA	AAATAGGCGT	ATCACGAGGC	CCTTTTCGTCT	CGCGCGTTTC	GGTGATGACG
1501	GTGAAAACCT	CTGACACATG	CAGTCCCCGG	AGACGGTCAC	AGCTTGTCTG	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGGCGCGT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATACCGCA	TCAGGCGCCA	TTCGCCATTG	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAGGGG
1801	GATGTGCTGC	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TTCCCAGTCA	CGACGTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCGGGCTGG	TCTGAATCCT	TCAGCGGGT
1921	GTTACCCGAG	AAAGAAAGGG	TGCCGATTTC	AAAGCAGACC	CATGTGCCGG	GCCCTGTGGC
1981	CTGTGTTGGC	GCCTATGTAG	TCACCCCCCC	TCACCCAATT	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	ATCGACCCTG	CACCGGTGCC	TACGTGGGTC	AAGTACGGCA	TCATGTGGCC
2341	GCTCTACTGG	TTCTTCCAGG	TGTGTTTGAG	GGTTTTGGTT	GCCCGTATTG	AGGTCTTGGT
2401	GGCGCGCATG	GAGGAGAAGG	CGCCTGTCCC	GCTGACCCCC	CCGGTACCC	TCCCGGCACC
2461	TTCCAGGGCG	CCTTCGGCAC	GGGTGCTTGG	GTGTGCGCGC	ACGAGTCCGG	CCACCAGGCC
2521	TTTTCTCCA	GCCAGGCCAT	CAACGACGGC	GTGGGCCTGG	TGTTCCACAG	CCTGCTGCTG
2581	GTGCCCTACT	ACTCCTGGAA	GCACTCGCAC	CGGGTACCCT	TTCTTGCGCT	ATGACACTTC
2641	CAGCAAAAGG	TAGGGCGGGC	TGCGAGACGG	CTTCCCGGCG	CTGCATGCAA	CACCGATGAT
2701	GCTTCGACCC	CCCGAAGCTC	CTTCGGGGCT	GCATGGGCGC	TCCGATGCCG	CTCCAGGGCG
2761	AGCGCTGTTT	AAATAGCCAG	GCCCCCGATT	GCAAAGACAT	TATAGCGAGC	TACCAAAGCC
2821	ATATTCAAAC	ACCTAGATCA	CTACCACTTC	TACACAGGCC	ACTCGAGCTT	GTGATCGCAC
2881	TCCGCTAAGG	GGGCGCCTCT	TCCTCTTCGT	TTCACTCACA	ACCCGCAAAC	TCTAGAATAT
2941	CAATGCTGCT	GCAGGCCTTC	CTGTTCCTGC	TGGCCGGCTT	CGCCGCCAAG	ATCAGCGCCT
3001	CCATGACGAA	CGAGACGTCC	GACCGCCCCC	TGGTGCACTT	CACCCCCAAC	ATCAGCTTGA
3061	TGAACGACCC	CAACGGCCTG	TGGTACGACG	AGAAGGACGC	CAAGTGGCAC	CTGTACTTCC
3121	AGTACAACCC	GAACGACACC	GTCTGGGGGA	CGCCCTTGTT	CTGGGGCCAC	GCCACGTCCG
3181	ACGACCTGAC	CAACTGGGAG	GACCAGCCCA	TCGCCATCGC	CCCGAAGCGC	AACGACTCCG
3241	GCGCCTTCTC	CGGCTCCATG	GTGGTGGACT	ACAACAACAC	CTCCGGCTTC	TTCAACGACA
3301	CCATCGACCC	GCGCCAGCGC	TGCGTGGCCA	TCTGGACCTA	CAACACCCCG	GAGTCCGAGG
3361	AGCAGTACAT	CTCCTACAGC	CTGGACGGCG	GCTACACCTT	CACCGAGTAC	CAGAAGAACC
3421	CCGTGCTGGC	CGCCAACTCC	ACCCAGTTCC	GCGACCCGAA	GGTCTTCTG	TACGAGCCCT
3481	CCCAGAAATG	GATCATGACC	GCGGCCAAGT	CCAGGACTA	CAAGATCGAG	ATCTACTCCT
3541	CCGACGACCT	GAAGTCCTGG	AAGCTGGAGT	CCGCGTTTCG	CAACGAGGGC	TTCTCTGGCT
3601	ACCAGTACGA	GTGCCCCGGC	CTGATCGAGG	TCCCCACCGA	GCAGGACCCC	AGCAAGTCCT
3661	ACTGGGTGAT	GTTTATCTCC	ATCAACCCCG	GCGCCCCGGC	CGGCGGCTCC	TTCAACCAGT
3721	ACTTCGTGCG	CAGCTTCAAC	GGCACCCACT	TCGAGGCCTT	CGACAACCAG	TCCCGCGTGG
3781	TGGACTTCGG	CAAGGACTAC	TACGCCCTGC	AGACCTTCTT	CAACACCGAC	CCGACCTACG
3841	GGAGCGCCCT	GGGCATCGCG	TGGGCCCTCA	ACTGGGAGTA	CTCCGCCTTC	GTGCCACCA
3901	ACCCCTGGCG	CTCCTCCATG	TCCCTCGTGC	GCAAGTTCTC	CCTCAACACC	GAGTACCAGG
3961	CCAACCCGGA	GACGGAGCTG	ATCAACCTGA	AGGCCGAGCC	GATCCTGAAC	ATCAGCAACG
4021	CCGCCCCCTG	GAGCCGGTTC	GCCACCAACA	CCACGTTGAC	GAAGGCCAAC	AGCTACAACG
4081	TCGACCTGTC	CAACAGCACC	GGCACCCCTG	AGTTCGAGCT	GGTGTTACGC	GTCAACACCA
4141	CCCAGACGAT	CTCCAAGTCC	GTGTTGCGGG	ACCTCTCCCT	CTGGTTCAAG	GGCCTGGAGG
4201	ACCCCGAGGA	GTACCTCCGC	ATGGGCTTCG	AGGTGTCCGC	GTCTCTCTTC	TTCTTGGACC
4261	GCGGGAACAG	CAAGGTGAAG	TTCGTGAAGG	AGAACCCCTA	CTTACCAAC	CGCATGAGCG
4321	TGAACAACCA	GCCCTTCAAG	AGCGAGAACG	ACCTGTCCCTA	CTACAAGGTG	TACGGCTTGC
4381	TGGACCAGAA	CATCCTGGAG	CTGTACTTCA	ACGACGGCGA	CGTCGTGTCC	ACCAACACCT
4441	ACTTCATGAC	CACCGGGAAC	GCCCTGGGCT	CCGTGAACAT	GACGACGGGG	GTGGACAACC
4501	TGTTTCTACAT	CGACAAGTTC	CAGGTGCGCG	AGGTCAAGTG	ACAATTGGCA	GCAGCAGTTC
4561	GGATAGTATC	GACACACTCT	GGACGCTGGT	CGTGTGATGG	ACTGTTGCCG	CCACACTTGC
4621	TGCCTTGACC	TGTGAATATC	CCTGCCGCTT	TTATCAAACA	GCCTCAGTGT	GTTTGATCTT
4681	GTGTGTACGC	GCTTTTGCAG	GTTGCTAGCT	GCTTGTGCTA	TTTGCGAATA	CCACCCCCAG
4741	CATCCCCTTC	CCTCGTTTCA	TATCGCTTGC	ATCCCAACCG	CAACTTATCT	ACGCTGTCTT
4801	GCTATCCCTC	AGCGCTGCTC	CTGCTCCTGC	TCACTGCCCC	TCGCACAGCC	TTGGTTTGGG

4861	CTCCGCCTGT	ATTCTCCTGG	TACTGCAACC	TGTAAACCAG	CACTGCAATG	CTGATGCACG
4921	GGAAGTAGTG	GGATGGGAAC	ACAAATGGAG	GATCCCGCGT	CTCGAACAGA	GCGCGCAGAG
4981	GAACGCTGAA	GGTCTCGCCT	CTGTGCGACC	TCAGCGCGGC	ATACACCACA	ATAACCACCT
5041	GACGAATGCG	CTTGTTCTT	CGTCCATTAG	CGAAGCGTCC	GGTTCACACA	CGTGCCACGT
5101	TGGCGAGGTG	GCAGGTGACA	ATGATCGGTG	GAGCTGATGG	TCGAAACGTT	CACAGCCTAG
5161	GGATATCGAA	TTCGGCCGAC	AGGACGCGCG	TCAAAGGTGC	TGGTCGTGTA	TGCCCTGGCC
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	GGTCCTGCTG	AAGTTCCTCC
5461	ACCGCCTCAC	CAGCGGACAA	AGCACCGGTG	TATCAGGTCC	GTGTCATCCA	CTCTAAAGAA
5521	CTCGACTACG	ACCTACTGAT	GGCCCTAGAT	TCTTCATCAA	AAACGCCTGA	GACACTTGCC
5581	CAGGATTGAA	ACTCCCTGAA	GGGACCACCA	GGGGCCCTGA	GTGTTCCTT	CCCCCGTGG
5641	CGAGCTGCCA	GCCAGGCTGT	ACCTGTGATC	GAGGCTGGCG	GGAAAAATAGG	CTTCGTGTGC
5701	TCAGGTCATG	GGAGGTGCAG	GACAGCTCAT	GAAACGCCAA	CAATCGCACA	ATTATGTCA
5761	AGCTAATCAG	CTATTTCTC	TTCACGAGCT	GTAATTGTCC	CAAAATTCTG	GTCTACCGGG
5821	GGTGATCCTT	CGTGACGGG	CCCTTCCCTC	AACCCTAGGT	ATGCGCGCAT	GCGGTGCGCG
5881	CGCAACTCGC	GCGAGGGCCG	AGGGTTTGGG	ACGGGCCGTC	CCGAAATGCA	GTTGCACCCG
5941	GATGCGTGCC	ACCTTTTTTG	CGATAATTTA	TGCAATGGAC	TGCTCTGCAA	AATTCTGGCT
6001	CTGTCGCCAA	CCCTAGGATC	AGCGGCGTAG	GATTTTCGTAA	TCATTCGTCC	TGATGGGGAG
6061	CTACCGACTA	CCCTAATATC	AGCCCGACTG	CCTGACGCCA	GCGTCCACTT	TTGTGCACAC
6121	ATTCCATTCTG	TGCCCAAGAC	ATTCTATTGT	GGTGCGAAGC	GTCCCCAGTT	ACGCTACCT
6181	GTTTCCCGAC	CTCCTTACTG	TTCTGTCGAC	AGAGCGGGCC	CACAGGCCGG	TCGCAGCCAC
6241	TAGTATGGCC	ACCGCATCCA	CTTTCTCGGC	GTTCAATGCC	CGCTGCGGCG	ACCTGCGTCG
6301	CTCGGCGGGC	TCCGGGCCCC	GGCGCCAGC	GAGGCCCTC	CCCGTGC GCG	GGCGCGCCGC
6361	CACCGGCGAG	CAGCCCTCCG	GCGTGCCCTC	CCTGCGCGAG	GCCGACAAGG	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	GTGACGCCCC	CGATGCACAT
6661	CGAGATCTAC	CGCTACCCCG	CCTGTGTCCGA	CGTGATCGAG	ATCGAGACCT	GGGTGCAGGG
6721	CGAGGGCAAG	GTGGGCACCC	GCCGCGACTG	GATCCTGAAG	GACTACGCCA	ACGGCGAGGT
6781	GATCGGCCGC	GCCACCTCCA	AGTGGGTGAT	GATGAACGAG	GACACCCGCC	GCCTGCAGAA
6841	GGTGTCGAC	GACGTGCGCG	AGGAGTACCT	GGTGTCTGCG	CCCCGCACCC	TGCGCCTGGC
6901	CTTCCCCGAG	GAGAACAACA	ACTCATGAA	GAAGATCCCC	AAGCTGGAGG	ACCCCGCCGA
6961	GTACTCCCGC	CTGGGCCTGG	TGCCCCGCCG	CTCCGACCTG	GACATGAACA	AGCACGTGAA
7021	CAACGTGACC	TACATCGGCT	GGGCCCTGGA	GTCCATCCCC	CCCAGATCA	TCGACACCCA
7081	CGAGCTGCAG	GCCATCACCC	TGGACTACCG	CCGCGAGTGC	CAGCGCGACG	ACATCGTGGA
7141	CTCCCTGACC	TCCCGCGAGC	CCCTGGGCAA	CGCCGCCGGC	GTGAAGTTCA	AGGAGATCAA
7201	CGGCTCCGTG	TCCCCAAGA	AGGACGAGCA	GGACCTGTCC	CGCTTCATGC	ACCTGCTGCG
7261	CTCCGCCGGC	TCCGGCCTGG	AGATCAACCG	CTGCCGCACC	GAGTGGCGCA	AGAAGCCCCG
7321	CAAGCGCATG	GACTACAAGG	ACCACGACGG	CGACTACAAG	GACCACGACA	TCGACTACAA
7381	GGACGACGAC	GACAAGTGAA	TCGATAGATC	TCTTAAGGCA	GCAGCAGCTC	GGATAGTATC
7441	GACACACTCT	GGACGCTGGT	CGTGTGATGG	ACTGTTGCCG	CCACACTTGC	TGCCTTGACC
7501	TGTGAATATC	CCTGCCGCTT	TTATCAAAACA	GCCTCAGTGT	GTTTGATCTT	GTGTGTACGC
7561	GCTTTTGCGA	GTTGCTAGCT	GCTTGTGCTA	TTTGCGAATA	CCACCCCCAG	CATCCCCCTC
7621	CCTCGTTTCA	TATCGCTTGC	ATCCCAACCG	CAACTTATCT	ACGCTGTCTT	GCTATCCCTC
7681	AGCGCTGCTC	CTGCTCCTGC	TCACTGCCCC	TCGCACAGCC	TTGGTTTGGG	CTCCGCTGCT
7741	ATTCTCCTGG	TACTGCAACC	TGTAAACCAG	CACTGCAATG	CTGATGCACG	GGAAGTAGTG
7801	GGATGGGAAC	ACAAATGGAA	AGCTTAATTA	AGAGCTCCCC	CCACCACTCC	AACACGGGGT
7861	GCCTGGACAA	GGACGAGGTG	TTTGTGCCGC	CGCACCGCGC	AGTGGCGCAC	GAGGGCCTGG
7921	AGTGGGAGGA	GTGGCTGCCC	ATCCGCATGG	GCAAGGTGCT	GGTACCCCTG	ACCTTGGGCT
7981	GGCCGCTGTA	CCTCATGTTT	AACGTCGCCT	CGCGGCCGTA	CCCGCGCTTC	GCCAACCACT
8041	TTGACCCGTG	GTCGCCCATC	TTAGCAAGC	GCGAGCGCAT	CGAGGTGGTC	ATCTCCGACC
8101	TGGCGCTGGT	GGCGGTGCTC	AGCGGGCTCA	GCGTGTGTTG	CCGCACCATG	GGCTGGGCCT
8161	GGCTGGTCAA	GACCTACGTG	GTGCCCTACC	TGATCGTGAA	CATGTGGCTC	GTGCTCATCA
8221	CGCTGCTCCA	GCACACGCAC	CCGGCGCTGC	CGCACTACTT	CGAGAAGGAC	TGGACTGGGC
8281	TGCGCGGCGC	CATGGCCACC	GTGGACCGCT	CCATGGGCCC	GCCCTTCATG	GACAACATCC
8341	TGCACCACAT	CTCCGACACC	CACGTGCTGC	ACCACCTCTT	CAGCACCATC	CCGCACTACC
8401	ACGCCGAGGA	GGCCTCCGCC	GCCATCAGGC	CCATCCTGGG	CAAGTACTAC	CAGTCCGACA
8461	GCCGCTGGGT	CGGCCGCGCC	CTGTGGGAGG	ACTGGCGCGA	CTGCCGCTAC	GTCGTCCCGG

```

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

```

SEQ ID NO: 234

pSZ1501_FADc-inv

```

1  CACTGGCAGC  AGCCACTGGT  AACAGGATTA  GCAGAGCGAG  GTATGTAGGC  GGTGCTACAG
61  AGTTCCTGAA  GTGGTGGCCT  AACTACGGCT  ACACTAGAAG  GACAGTATTT  GGTATCTGCG
121  CTCTGCTGAA  GCCAGTTACC  TTCGGAAAAA  GAGTTGGTAG  CTCTTGATCC  GGCAAAACAA
181  CCACCGCTGG  TAGCGGTGGT  TTTTTGTTT  GCAAGCAGCA  GATTACGCGC  AGAAAAAAG
241  GATCTCAAGA  AGATCCTTTG  ATCTTTCTA  CGGGGTCTGA  CGCTCAGTGG  AACGAAAAC
301  CACGTTAAGG  GATTTTGGTC  ATGAGATTAT  CAAAAGGAT  CTTACCTAG  ATCCTTTTAA
361  ATTAATAATG  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  GTTGGTATG  GCTTCATTCA
781  GCTCCGTTTC  CCAACGATCA  AGGCAGTTA  CATGATCCCC  CATGTTGTGC  AAAAAAGCGG
841  TTAGTCCCTT  CGGTCTCCG  ATCGTTGTCA  GAAGTAAGTT  GGCCGCAAGT  TTATCACTCA
901  TGGTTATGGC  AGCACTGCAT  AATTCTCTTA  CTGTCAATGC  ATCCGTAAGA  TGCTTTTCTG
961  TGACTGGTGA  GTACTCAACC  AAGTCATTCT  GAGAATAGTG  TATGCGGCGA  CCGAGTTGCT
1021  CTTGCCCGGC  GTCAATACGG  GATAATACCG  GCCACATAG  CAGAAGTTTA  AAAGTGCTCA
1081  TCATTGGAAA  ACGTTCTTCG  GGGCGAAAAA  TCTCAAGGAT  CTACCGCTG  TTGAGATCCA
1141  GTTCGATGTA  ACCCACTCGT  GCACCCAACT  GATCTTCAGC  ATCTTTTACT  TTCACCAGCG
1201  TTTCTGGGTG  AGCAAAAACA  GGAAGGCAAA  ATGCCGCAAA  AAAGGGAATA  AGGGCGACAC
1261  GGAAATGTTG  AATACTCATA  CTCTTCCTTT  TTCAATATTA  TTGAAGCATT  TATCAGGGTT
1321  ATTGTCAT  GAGCGGATAC  ATATTGAAT  GTATTTAGAA  AAATAAACAA  ATAGGGGTTT
1381  CGCGCACATT  TCCCCGAAAA  GTGCCACCTG  ACGTCTAAGA  AACCATTATT  ATCATGACAT
1441  TAACCTATAA  AAATAGGCGT  ATCAGGAGGC  CCTTTCGTCT  CGCGCGTTTC  GGTGATGACG
1501  GTGAAAACCT  CTGACACATG  CAGTCCCCGG  AGACGGTCAC  AGCTTGTCTG  TAAGCGGATG
1561  CCGGAGCAG  ACAAGCCCGT  CAGGCGCGT  CAGCGGGTGT  TGCGGGGTG  CGGGGCTGGC
1621  TTAACATATG  GCATCAGAG  CAGATTGTAT  TGAGAGTGCA  CCATATGCGG  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  TTCGGGCTGG  TCTGAATCCT  TCAGGCGGGT
1921  GTTACCCGAG  AAAGAAAGGG  TGCCGATTTC  AAAGCAGACC  CATGTGCCGG  GCCCTGTGGC
1981  CTGTGTTGGC  GCCTATGTAG  TCACCCCCC  TCACCCAATT  GTCGCCAGTT  TGCGCACTCC
2041  ATAAACTCAA  AACAGCAGCT  TCTGAGCTGC  GCTGTTCAAG  AACACCTCTG  GGGTTTGCTC
2101  ACCCGCAGG  TCGACGCCCA  GCATGGCTAT  CAAGACGAAC  AGGCACCTG  TGGAGAAGCC
2161  TCCGTTACG  ATCGGGACGC  TCGCAAGGC  CATCCCCGCG  CACTGTTTCG  ACGGCTCGGC
2221  GCTTCGTAGC  AGCATGTACC  TGGCCTTTGA  CATCGCGGTC  ATGTCCCTGC  TCTACGTCGC
2281  GTCGACGTAC  ATCGACCCTG  CACCGGTGCC  TACGTGGGTC  AAGTACGGCA  TCATGTGGCC
2341  GCTCTACTGG  TTCTTCCAGG  TGTGTTTGG  GGTTTTGGTT  GCCCGTATTG  AGGTCTGGT
2401  GCGCGCATG  GAGGAGAAGG  CGCTGTCCC  GCTGACCCCC  CCGGCTACCC  TCCCGGCACC

```

2461 TTCCAGGGCG CCTTCGGCAC GGGTGTCTGG GTGTGCGCGC ACGAGTGC GG CCACCAGGCC
2521 TTTTCTCCA GCCAGGCCAT CAACGACGGC GTGGGCCTGG TGTTCCACAG CCTGCTGCTG
2581 GTGCCCTACT ACTCCTGGAA GCACTCGCAC CGGGTACCCT TTCTTGCGCT ATGACACTTC
2641 CAGCAAAAGG TAGGGCGGGC TGCAGACGG CTTCCCGGCG CTGCATGCAA CACCGATGAT
2701 GCTTCGACCC CCCGAAGCTC CTTCCGGGCT GCATGGGCGC TCCGATGCCG CTCCAGGGCG
2761 AGCGCTGTTT AAATAGCCAG GCCCCGATT GCAAAGACAT TATAGCGAGC TACCAAAGCC
2821 ATATTCAAAC ACCTAGATCA CTACCACTTC TACACAGGCC ACTCGAGCTT GTGATCGCAC
2881 TCCGTAAGG GGGCGCCTCT TCCTCTTCGT TTCAGTCACA 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 CGCCCTTGTT CTGGGGCCAC GCCACGTCCG
3181 ACGACCTGAC CAACTGGGAG GACCAGCCCA TCGCCATCGC CCCGAAGCGC AACGACTCCG
3241 GCGCTTCTC CGGCTCCATG GTGGTGGACT ACAACAACAC CTCCGGCTTC TTCAACGACA
3301 CCATCGACCC GCGCCAGCGC TCGTGGCCA TCTGGACCTA CAACACCCCG GAGTCCGAGG
3361 AGCAGTACAT CTCTACAGC CTGGACGGCG GCTACACCTT CACCGAGTAC CAGAAGAACC
3421 CCGTGTGGC CGCCAACCTC ACCCAGTTCC GCGACCCGAA GGTCTTCTGG TACGAGCCCT
3481 CCCAGAAGTG GATCATGACC GCGGCCAAGT CCCAGGACTA CAAGATCGAG ATCTACTCCT
3541 CCGACGACCT GAAGTCCTGG AAGCTGGAGT CCGCGTTCGC CAACGAGGGC TTCCTCGGCT
3601 ACCAGTACGA GTGCCCCGGC CTGATCGAGG TCCCCACCGA GCAGGACCCC AGCAAGTCTT
3661 ACTGGGTGAT GTTCATCTCC ATCAACCCCG GCGCCCCGGC CGCGGGCTCC TTCAACAGT
3721 ACTTCTCGG CAGCTTCAAC GGCACCCACT TCGAGGCCTT CGACAACCAG TCCCGCTGG
3781 TGGACTTCGG CAAGGACTAC TACGCCCTGC AGACCTTCTT CAACACCGAC CCGACCTACG
3841 GGAGCGCCCT GGGCATCGCG TGGGCTCCA ACTGGGAGTA CTCCGCTTC GTGCCACCA
3901 ACCCTGGCG CTCTCCATG TCCCTCGTGC GCAAGTTCTC CCTCAACACC GAGTACCAGG
3961 CCAACCCGGA GACGGAGCTG ATCAACCTGA AGGCCGAGCC GATCCTGAAC ATCAGCAACG
4021 CCGGCCCTG GAGCCGGTTC GCCACCAACA CCACGTTGAC GAAGGCCAAC AGCTACAACG
4081 TCGACCTGTC CAACAGCACC GGCACCCTGG AGTTCGAGCT GGTGTACGCC GTCAACACCA
4141 CCCAGACGAT CTCCAAGTCC GTGTTGCGC ACCTCTCCCT CTGGTTCAAG GGCCTGGAGG
4201 ACCCGAGGA GTACCTCCGC ATGGGCTTCG AGGTGTCCGC GTCTCCCTT TTCTTGACC
4261 GCGGGAACAG CAAGGTGAAG TTCGTGAAGG AGAACCCTA CTTCAACAC CGCATGAGCG
4321 TGAACAACCA GCCCTTCAAG AGCGAGAACG ACCTGTCCTA CTACAAGGTG TACGGCTTGC
4381 TGGACCAGAA CATCCTGGAG CTGTACTTCA ACGACGGCGA CGTCGTGTCC ACCAACACCT
4441 ACTTCATGAC CACCGGGAAC GCCCTGGGCT CCGTGAACAT GACGACGGGG GTGGACAACC
4501 TGTTCTACAT CGACAAGTTC CAGGTGCGCG AGGTCAAGTG ACAATTGGCA GCAGCAGCTC
4561 GGATAGTATC GACACACTCT GGACGCTGGT CGTGTGATGG ACTGTTGCCG CCACACTTGC
4621 TGCCTTGACC TGTGAATATC CTTGCCGCTT TTATCAAACA GCCTCAGTGT GTTTGATCTT
4681 GTGTGTACGC GCTTTTGCGA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCACG
4741 CATCCCTTC CCTCGTTTCA TATCGTTTGC ATCCAACCG CAACTTATCT ACGCTGTCTT
4801 GCTATCCCTC AGCGCTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG
4861 CTCCGCTGT ATTCTCCTGG TACTGCAACC TGTAACCCAG 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 TTCCTTTCTT GCGCTATGAC ACTTCCAGCA AAAGGTAGGG CGGGCTGCGA
5221 GACGGCTTCC CGGCGCTGCA TGCAACACCG ATGATGCTTC GACCCCCGA AGCTCCTTCG
5281 GGGCTGCAATG GCGCTCCGA TGCCGCTCCA GGGCGAGCGC TGTTTAAATA GCCACGCCC
5341 CGATTGCAAA GACATTATAG CGAGCTACCA AAGCCATATT CAAACACCTA GATCACTACC
5401 ACTTCTACAC AGGCCACTCG AGCTTGTGAT CGCACTCCGC TAAGGGGGCG CCTCTTCTC
5461 TTCGTTTCAG TCACAACCCG CAAACACTAG TATGGCCACC GCATCCACTT TCTCGGCGTT
5521 CAATGCCCCG TCGGGCGACC TCGTCTGCTC GCGGGGCTCC GGGCCCCGGC GCCAGCGAG
5581 GCCCCTCCCC GTGCGCGGGC GCGCGCCAC CGGCGAGCAG CCCTCCGGCG TGGCCTCCCT
5641 GCGCGAGGCC GACAAGGAGA AGTCCCTGGG CAACCGCCTG CGCCTGGGCT CCCTGACCGA
5701 GGACGGCCTG TCCTACAAGG AGAAGTTCTG GATCCGCTGC TACGAGGTGG GCATCAACAA
5761 GACCGCCACC ATCGAGACCA TCGCCAACCT GCTGCAGGAG GTGGGCGGCA ACCACGCCC
5821 GGGCTGGGC TTCTCCACCG ACGGCTTCGC CACCACCACC ACCATGCGCA AGCTGCACCT
5881 GATCTGGGTG ACCGCCCCGA TGACATCGA GATCTACCGC TACCCCGCTT GGTCCGACGT
5941 GATCGAGATC GAGACCTGGG TGCAAGGCGA GGGCAAGGTG GGCACCCGCC GCGACTGGAT
6001 CCTGAAGGAC TACGCCAACG GCGAGGTGAT CGGCCGCGCC ACCTCCAAGT GGGTGTATGAT
6061 GAACGAGGAC ACCCGCCGCC TGCAGAAGGT GTCCGACGAC GTGCGCGAGG AGTACCTGGT

6121	GTTCTGCCCC	CGCACCTGTC	GCCTGGCCTT	CCCCGAGGAG	AACAACAACCT	CCATGAAGAA
6181	GATCCCCAAG	CTGGAGGACC	CCGCCGAGTA	CTCCCGCCTG	GGCCTGGTGC	CCCGCCGCTC
6241	CGACCTGGAC	ATGAACAAGC	ACGTGAACAA	CGTGACCTAC	ATCGGCTGGG	CCCTGGAGTC
6301	CATCCCCCCC	GAGATCATCG	ACACCCACGA	GCTGCAGGCC	ATCACCTTGG	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	AGCCCGCCAA	GCGCATGGAC	TACAAGGACC	ACGACGGCGA
6601	CTACAAGGAC	CACGACATCG	ACTACAAGGA	CGACGACGAC	AAGTGAATCG	ATAGATCTCT
6661	TAAGGCAGCA	GCAGCTCGGA	TAGTATCGAC	ACACTCTGGA	CGCTGGTCTG	GTGATGGACT
6721	GTTGCCGCCA	CAC TTGCTGC	CTTGACCTGT	GAATATCCCT	GCCGCTTTTA	TCAAACAGCC
6781	TCAGTGTGTT	TGATCTTGTG	TGTACGCGCT	TTTGCGAGTT	GCTAGCTGCT	TGTGCTATTT
6841	GCGAATACCA	CCCCAGCAT	CCCTTCCCT	CGTTTCATAT	CGCTTGCATC	CCAACCGCAA
6901	CTTATCTACG	CTGTCTTGCT	ATCCCTCAGC	GCTGCTCCTG	CTCCTGCTCA	CTGCCCTCTG
6961	CACAGCCTTG	GTTTGGGCTC	CGCCTGTATT	CTCCTGGTAC	TGCAACCTGT	AAACCAGCAC
7021	TGCAATGCTG	ATGCACGGGA	AGTAGTGGGA	TGGGAACACA	AATGGAAGC	TTAATTAAGA
7081	GCTCCGCCCA	CCACTCCAAC	ACGGGGTGCC	TGGACAAGGA	CGAGGTGTTT	GTGCCGCCGC
7141	ACCGCGCAGT	GGCGCACGAG	GGCTTGGAGT	GGGAGGAGTG	GCTGCCCATC	CGCATGGGCA
7201	AGGTGCTGGT	CACCCTGACC	CTGGGCTGGC	CGCTGTACCT	CATGTTCAAC	GTCGCCTCGC
7261	GGCCGTACCC	GCGCTTCGCC	AACCACTTTG	ACCCGTGGTC	GCCCATCTTC	AGCAAGCGCG
7321	AGCGCATCGA	GGTGGTCATC	TCCGACCTGG	CGCTGGTGGC	GGTGCTCAGC	GGGCTCAGCG
7381	TGCTGGCCCG	CACCATGGGC	TGGGCTTGGC	TGGTCAAGAC	CTACGTGGTG	CCCTACCTGA
7441	TCGTGAACAT	GTGGCTCGTG	CTCATCACGC	TGCTCCAGCA	CACGCACCCG	GCGCTGCCGC
7501	ACTACTTCGA	GAAGGACTGG	GACTGGCTGC	GCGGCGCCAT	GGCCACCGTG	GACCGCTCCA
7561	TGGGCCCCGCC	CTTCATGGAC	AACATCCTGC	ACCACATCTC	CGACACCCAC	GTGCTGCACC
7621	ACCTCTTCAG	CACCATCCCG	CACTACCACG	CCGAGGAGGC	CTCCGCCGCC	ATCAGGCCCA
7681	TCCTGGGCAA	GTACTACCAG	TCCGACAGCC	GCTGGGTCGG	CCGCGCCCTG	TGGGAGGACT
7741	GGCGCGACTG	CCGCTACGTC	GTCCCGGACG	CGCCCGAGGA	CGACTCCGCG	CTCTGGTTCC
7801	ACAAGTGAGT	GAGTGAGAAG	AGCCTCTAGA	GTCGACCTGC	AGGCATGCAA	GCTTGGCGTA
7861	ATCATGGTCA	TAGCTGTTTC	CTGTGTGAAA	TTGTTATCCG	CTCACAATTC	CACACAACAT
7921	ACGAGCCGGA	AGCATAAAGT	GTAAAGCCTG	GGGTGCCTAA	TGAGTGAGCT	AACTCACATT
7981	AATTGCGTTG	CGCTCACTGC	CCGCTTTCCA	GTCGGGAAAC	CTGTCTGTCC	AGCTGCATTA
8041	ATGAATCGGC	CAACGCGCGG	GGAGAGGCGG	TTTGCCTATT	GGGCGCTCTT	CCGCTTCCTC
8101	GCTCACTGAC	TCGCTGCGCT	CGGTCGTTCC	GCTGCGGCGA	GCGGTATCAG	CTCACTCAAA
8161	GGCGGTAATA	CGGTTATCCA	CAGAATCAGG	GGATAACGCA	GGAAAGAACA	TGTGAGCAAA
8221	AGGCCAGCAA	AAGGCCAGGA	ACCGTAAAAA	GGCCGCGTTG	CTGGCGTTTT	TCCATAGGCT
8281	CCGCCCCCCT	GACGAGCATC	ACAAAAATCG	ACGCTCAAGT	CAGAGGTGGC	GAAACCCGAC
8341	AGGACTATAA	AGATACCAGG	CGTTTCCCCC	TGGAAGCTCC	CTCGTGCGCT	CTCCTGTTCC
8401	GACCTTGCCG	CTTACCGGAT	ACCTGTCCGC	CTTTCTCCCT	TCGGGAAGCG	TGGCGCTTTC
8461	TCATAGCTCA	CGCTGTAGGT	ATCTCAGTTC	GGTGTAGGTC	GTTTCGTCCA	AGCTGGGCTG
8521	TGTGCACGAA	CCCCCGTTTC	AGCCCGACCG	CTGCGCCTTA	TCCGGTAACT	ATCGTCTTGA
8581	GTCCAACCCG	GTAAGACACG	ACTTATCGC			

SEQ ID NO: 235
pSZ1570

1	CACTGGCAGC	AGCCACTGGT	AACAGGATTA	GCAGAGCGAG	GTATGTAGGC	GGTGCTACAG
61	AGTTCTTGAA	GTGGTGCCCT	AAC TACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAAACAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT
421	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTCTG	TCATCCATAG
481	TTGCCTGACT	CCCCGTCTGT	TAGATAACTA	CGATACGGGA	GGGCTTACCA	TCTGGCCCCA
541	GTGCTGCAAT	GATACCGCGA	GACCCACGCT	CACCGGCTCC	AGATTATATCA	GCAATAAACC
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	CGGTCTCCG	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAGTG	TTATCACTCA
901	TGTTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTATGCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAC	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCATGTA	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	ATCACGAGGG	CCTTTTCGTCT	CGCGCGTTTC	GGTGATGACG
1501	GTGAAAACCT	CTGACACATG	CAGCTCCCGG	AGACGGTCAC	AGCTTGCTGT	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGGCGCGT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	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	CCACTCCTGC	TGAGCGCGC
1921	CCGCGCGTGC	GCCGCCAGCG	CCTTGCCCTT	TTCGCCGCGC	TCGTGCGCGT	CGCTGATGTC
1981	CATCACCAGG	TCCATGAGGT	CTGCCTTGCG	CCGGCTGAGC	CACTGCTTCG	TCCGGGCGGC
2041	CAAGAGGAGC	ATGAGGGAGG	ACTCTTGGTC	CAGGGTCCTG	ACGTGGTCGC	GGCTCTGGGA
2101	GCGGGCCAGC	ATCATCTGGC	TCTGCCGCAC	CGAGGCCGCC	TCCAACCTGT	CCTCCAGCAG
2161	CCGCAGTCGC	CGCCGACCCT	GGCAGAGGAA	GACAGGTGAG	GGGGGTATGA	ATTGTACAGA
2221	ACAACCACGA	GCCTTGCTTA	GGCAGAATCC	CTACCAGTCA	TGGCTTTACC	TGGATGACGG
2281	CCTGCGAACA	GCTGTCCAGC	GACCCTCGCT	GCCGCCGCTT	CTCCGCACG	CTTCTTTCCA
2341	GCACCGTGAT	GGCGCGAGCC	AGCGCCGCAC	GCTGGCGCTG	CGCTTCGCCG	ATCTGAGGAC
2401	AGTCGGGGAA	CTCTGATCAG	TCTAAACCCC	CTTGCGCGTT	AGTGTTGCCA	TCCTTTGCAG
2461	ACCGTGAGAG	GCCGACTTGT	TGTGCGCCAC	CCCCACACC	ACCTCCTCCC	AGACCAATTC
2521	TGTCACCTTT	TTGGCGAAGG	CATCGGCCTC	GGCCTGCAGA	GAGGACAGCA	GTGCCACGCC
2581	GCTGGGGGTT	GGCGGATGCA	CGCTCAGGTA	CCCTTTCTTG	CGCTATGACA	CTTCCAGCAA
2641	AAGGTAGGGC	GGGCTGCGAG	ACGGCTTCCC	GGCGCTGCAT	GCAACACCGA	TGATGCTTCG
2701	ACCCCCCGAA	GCTCCTTCGG	GGCTGCATGG	GCGCTCCGAT	GCCGCTCCAG	GGCGAGCGCT
2761	GTTTAAATAG	CCAGGCCCCC	GATTGCAAAG	ACATTATAGC	GAGCTACCAA	AGCCATATTC
2821	AAACACCTAG	ATCACTACCA	CTTCTACACA	GGCCACTCGA	GCTTGATGATC	GCACTCCGCT
2881	AAGGGGCGCG	CTCTTCCTCT	TCGTTTCAGT	CACAACCCGC	AAACGGCGCG	CCATGCTGCT
2941	GCAGGCCCTT	CTGTTTCCTG	TGGCCGGCTT	CGCCGCCAAG	ATCAGCGCCT	CCATGACGAA
3001	CGAGACGTCC	GACCGCCCCC	TGGTGCACTT	CACCCCAAC	AAGGGCTGGA	TGAACGACCC
3061	CAACGGCCTG	TGGTACGACG	AGAAGGACGC	CAAGTGGCAC	CTGTACTTCC	AGTACAACCC
3121	GAACGACACC	GTCTGGGGGA	CGCCCTTGTT	CTGGGGCCAC	GCCACGTCGG	ACGCACTGAC
3181	CAACTGGGAG	GACCAGCCCC	TCGCCATCGC	CCCGAAGCGC	AACGACTCCG	GCGCCTTCTC
3241	CGGCTCCATG	GTGGTGGAAT	ACAACAACAC	CTCCGGCTTC	TTCAACGACA	CCATCGACCC
3301	GCGCCAGCGC	TGCGTGCCCA	TCTGGACCTA	CAACACCCCG	GAGTCCGAGG	AGCAGTACAT
3361	CTCCTACAGC	CTGGACGGCG	GCTACACCTT	CACCGAGTAC	CAGAAGAACC	CCGTGCTGGC
3421	CGCCAACTCC	ACCCAGTTCC	GCGACCCGAA	GGTCTTCTGG	TACGAGCCCT	CCCAGAAGTG
3481	GATCATGACC	GCGGCCAAGT	CCCAGGACTA	CAAGATCGAG	ATCTACTCCT	CCGACGACCT
3541	GAAGTCCTGG	AAGCTGGAGT	CCGCGTTTCG	CAACGAGGGC	TTCTTCGGCT	ACCAGTACGA
3601	GTGCCCCGGC	CTGATCGAGG	TCCCCACCGA	GCAGGACCCC	AGCAAGTCCT	ACTGGGTGAT
3661	GTTTCATCTC	ATCAACCCCG	GCGCCCCGGC	CGGCGGCTCC	TTCAACCACT	ACTTCGTGCG
3721	CAGCTTCAAC	GGCACCCACT	TCGAGGCCCT	CGACAACCAG	TCCCGCGTGG	TGGACTTCGG
3781	CAAGGACTAC	TACGCCCTGC	AGACCTTCTT	CAACACCGAC	CCGACCTACG	GGAGCGCCCT
3841	GGGCATCGCG	TGGGCCCTCA	ACTGGGAGTA	CTCCGCCTTC	GTGCCCACCA	ACCCCTGGCG
3901	CTCTCCATG	TCCCTCGTGC	GCAAGTTCTC	CCTCAACACC	GAGTACCAGG	CCAACCCGGA
3961	GACGGAGCTG	ATCAACCTGA	AGGCCGAGCC	GATCCTGAAC	ATCAGCAACG	CCGGCCCCCTG
4021	GAGCCGGTTC	GCCACCAACA	CCACGTTGAC	GAAGGCCAAC	AGCTACAACG	TCGACCTGTC
4081	CAACAGCACC	GGCACCCCTG	AGTTCGAGCT	GGTGTACGCC	GTCAACACCA	CCCAGACGAT
4141	CTCCAAGTCC	GTGTTTCGCG	ACCTCTCCCT	CTGGTTCAAG	GGCCTGGAGG	ACCCCGAGGA
4201	GTACCTCCGC	ATGGGCTTCG	AGGTGTCCGC	GTCCTCCTTC	TTCTTGAGCC	CGGGGAACAG
4261	CAAGGTGAAG	TTCTGTGAAG	AGAACCCCTA	CTTACCAAC	CGCATGAGCG	TGAACAACCA
4321	GCCCTTCAAG	AGCGAGAACG	ACCTGTCCCTA	CTACAAGGTG	TACGGCTTGC	TGGACCAGAA
4381	CATCCTGGAG	CTGTACTTCA	ACGACGGCGA	CGTCGTGTCC	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 TTATCAAACA GCCTCAGTGT GTTGTATCTT GTGTGTACGC
4681 GCTTTTGC GA GTTGCTAGCT GCTTGTGCTA TTTGCGAATA CCACCCCAG CATCCCCCTC
4741 CCTCGTTTCA TATCGCTTGC ATCCCAACCG CAACTTATCT ACGCTGTCTT GCTATCCCTC
4801 AGCGCTGCTC CTGCTCCTGC TCACTGCCCC TCGCACAGCC TTGGTTTGGG CTCCGCCTGT
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 GGTTACACACA CGTGCCACGT TGGCGAGGTG
5101 GCAGGTGACA ATGATCGGTG GAGCTGATGG TCGAAACGTT CACAGCCTAG GGATATCGAA
5161 TTCGGCCGAC AGGACGCGCG TCAAAGGTG TGGTCGTGTA TGCCCTGGCC GGCAGGTCTG
5221 TGCTGCTGCT GGTTAGTGAT TCCGCAACCC TGATTTTGGC GTCTTATTTT GGC GTGCGCAA
5281 ACGCTGGCGC CCGCGAGCCG GGCCGGCGGC GATGCGGTGC CCCACGGCTG CCGGAATCCA
5341 AGGGAGGCAA GAGCGCCCGG GTCAGTTGAA GGGCTTTACG CGCAAGGTAC AGCCGCTCCT
5401 GCAAGGCTGC GTGGTGGAAT TGGACGTGCA GGTCTCTGCT AAGTCTCTCC ACCGCTCAC
5461 CAGCGGACAA AGCACCGGTG TATCAGGTCC GTGTCATCCA CTCTAAAGAG CTCGACTACG
5521 ACCTACTGAT GGCCCTAGAT TCTTCATCAA AAACGCCTGA GACACTTGCC CAGGATTGAA
5581 ACTCCCTGAA GGGACCACCA GGGGCCCTGA GTTGTTCCTT CCCCCGTGG CGAGCTGCCA
5641 GCCAGGCTGT ACCTGTGATC GAGGCTGGCG GGAAAATAGG CTTCTGTGTC TCAGGTCTAG
5701 GGAGGTGACG GACAGCTCAT GAAACGCCAA CAATCGCACA ATTCATGTCA AGCTAATCAG
5761 CTATTTCTCT TTCACGAGCT GTAAATTCTC CAAAATTCTG GTCTATCCGG GTGATCTCTT
5821 CGTGTACGGG CCCTTCCCTC AACCTTAGGT ATGCGCGCAT GCGGTCGCCG CGCAACTCGC
5881 GCGAGGGCCG AGGGTTTGGG ACGGGCCGTC CCGAAATGCA GTTGACCCG GATGCGTGGC
5941 ACCTTTTTTG CGATAATTTA TGCAATGGAC TGCTCTGCAA AATTCTGGCT CTGTCGCCAA
6001 CCCTAGGATC AGCGGCGTAG GATTTCGTAA TCATTCTGCC TGATGGGGAG CTACCGACTA
6061 CCCTAATATC AGCCCGACTG CCTGACGCCA GCGTCCACTT TTGTGCACAC ATTCCATTCT
6121 TGCCCAAGAC ATTTTCATTGT GTGCGAAGC GTCCCCAGTT ACGCTCACCT GTTTCCCGAC
6181 CTCCTTACTG TTCTGTGAC AGAGCGGGCG CACAGGCCGG TCGCAGCCAC TAGTATGGCT
6241 TCCGCGGCAT TCACCATGTC GGCGTGCCCC GCGATGACTG GCAGGCCCCC TGGGGCAGCT
6301 CGCTCCGGAC GGCCAGTCGC CACCCGCTG AGGGGGCGCG CCCCCAAGGC CACCGGCTCC
6361 GCCGTGTCCC TGAAGTCCGG CTCCTTGGAC ACCCAGGAGG ACACCTCCTC CTCCTCCTCC
6421 CCCCCCGCA CCTTCATCAA CCAGCTGCCC GACTGGTCCA TGCTGCTGTC CGCCATCACC
6481 ACCGTGTTCTG TGGCCGCCGA GAAGCAGTGG ACCATGCTGG ACCGCAAGTC CAAGCGCCCC
6541 GACACCTGA TGGACCCCTT CGGCGTGGAC CGCGTGGTGC AGGACGGCGT GGTGTTCCGC
6601 CAGTCTTCT CCATCCGCTC CTACGAGATC GGCGCCGACC GCACCGCCTC CATCGAGACC
6661 CTGATGAACA TCTTCCAGGA GACCTCCCTG AACCCTGCA AGTCCATCGG CCTGTGAAC
6721 GACGGCTTCG GCCGCACCCC CGAGATGTG AAGCGCGACC TGATCTGGGT GGTGACCAAG
6781 ATGCACATCG AGGTGAACCG CTACCCACC TGGGGCGACA CCATCGAGGT GAACACCTGG
6841 GTGTCCGAGT CCGGCAAGAC CGGCATGGGC CGCGACTGGC TGATCTCCGA CTGCCACACC
6901 GGCGAGATCC TGATCCGCGC CACCTCCGTG TGCGCCATGA TGAACCAGAC CACCCGCCGC
6961 TTCTCCAAGT TCCCCTACGA GGTGCGCCAG GAGCTGGCCC CCCACTTCGT GGACTCCGCC
7021 CCCGTGATCG AGGACTACCA GAAGCTGCAC AAGCTGGACG TGAAGACCGG CGACTCCATC
7081 TGCAACGGCC TGACCCCCCG CTGGAACGAC CTGGACGTGA ACCAGCACGT GAACAACGTG
7141 AAGTACATCG GCTGGATTCT GGAGTCCGTG CCCAAGGAGG TGTTCCGAGAC CCAGGAGCTG
7201 TGCGGCTTGA CCCTGGAGTA CCGCCGCGAG TGCGGCCGCG ACTCCGTGCT GAAGTCCGTG
7261 ACCGCCATGG ACCCCTCCAA GGAGGGCGAC CGCTCCCTGT ACCAGCACCT GCTGCGCCTG
7321 GAGCAGGCA CCGACATCGC CAAGGGCCGC ACCAAGTGGC GCCCAAGAA CTGCCGCACC
7381 AACGGCGCCA TCTCCACCGG CAAGACCTCC AACGGCAACT CCATCTCCTG ATTAATTAAC
7441 TCGAGGCAGC AGCAGCTCGG ATAGTATCGA CACACTCTGG ACGCTGGTCG TGTGATGGAC
7501 TGTTGCCGCC ACACTTGCTG CCTTGACCTG TGAATATCCC TGCCGCTTTT ATCAAACAGC
7561 CTCAGTGTGT TTGATCTTGT GTGTACGCGC TTTTGCGAGT TGCTAGCTGC TTGTGCTATT
7621 TGCGAATACC ACCCCAGCA TCCCCTTCCC TCGTTTCATA TCGCTTGCAT CCAACCGCA
7681 ACTTATCTAC GCTGTCCTGC TATCCCTCAG CGCTGCTCCT GCTCCTGCTC ACTGCCCTC
7741 GCACAGCCTT GGTGTGGGCT CCGCCTGTAT TCTCCTGGTA CTGCAACCTG TAAACCAGCA
7801 CTGCAATGCT GATGCACGGG AAGTAGTGGG ATGGGAACAC AAATGGAAAG CTTGAGCTCT
7861 TGTTTTCCAG AAGGAGTTGC TCCTTGAGCC TTTCATTCTC AGCCTCGATA ACCTCCAAAG
7921 CCGCTCTAAT TGTGGAGGGG GTTCGAATTT AAAAGCTTGG AATGTTGGTT CGTGCCTCTG
7981 GAACAAGCCC AGACTTGTTG CTCACTGGGA AAAGGACCAT CAGCTCCAAA AAACCTGCCC
8041 CTCAAACCGC GTACCTCTGC TTTGCGCAA TCTGCCCTGT TGAAATCGCC ACCACATTCA
8101 TATTGTGACG CTTGAGCAGT CTGTAATTGC CTCAGAATGT GGAATCATCT GCCCCTGTG

8161	CGAGCCCATG	CCAGGCATGT	CGCGGGCGAG	GACACCCGCC	ACTCGTACAG	CAGACCATTA
8221	TGCTACCTCA	CAATAGTTCA	TAACAGTGAC	CATATTTCTC	GAAGCTCCCC	AACGAGCACC
8281	TCCATGCTCT	GAGTGGCCAC	CCCCCGGCC	TGGTGTTCG	GGAGGGCAGG	TCAACCGGCA
8341	TGGGGCTACC	GAAATCCCCG	ACCGGATCCC	ACCACCCCG	CGATGGGAAG	AATCTCTCCC
8401	CGGGATGTGG	GCCCACCACC	AGCACAACT	GCTGGCCAG	GCGAGCGTCA	AACCATACCA
8461	CACAAATATC	CTTGGCATCG	GCCCTGAATT	CCTTCTGCCG	CTCTGCTACC	CGGTGCTTCT
8521	GTCCGAAGCA	GGGGTTGCTA	GGGATCGCTC	CGAGTCCGCA	AACCCTTGTC	GCGTGGCGGG
8581	GCTTGTTTCA	GCTTGAAGAG	CCTCTAGAGT	CGACCTGCAG	GCATGCAAGC	TTGGCGTAAT
8641	CATGGTCATA	GCTGTTTCCT	GTGTGAAATT	GTTATCCGCT	CACAATTCCA	CACAACATAC
8701	GAGCCGGAAG	CATAAAGTGT	AAAGCCTGGG	GTGCCTAATG	AGTGAGCTAA	CTCACATTAA
8761	TTGCGTTGCG	CTCACTGCCC	GCTTTCCAGT	CGGGAAACCT	GTCGTGCCAG	CTGCATTAAAT
8821	GAATCGGCCA	ACGCGCGGGG	AGAGGCGGTT	TGCGTATTGG	GCGCTCTTCC	GCTTCCTCGC
8881	TCACTGACTC	GCTGCGCTCG	GTCGTTCGGC	TGCGGCGAGC	GGTATCAGCT	CACTCAAAGG
8941	CGGTAATACG	GTTATCCACA	GAATCAGGGG	ATAACGCAGG	AAAGAACATG	TGAGCAAAAG
9001	GCCAGCAAAA	GGCCAGGAAC	CGTAAAAAGG	CCGCGTTGCT	GGCGTTTTTC	CATAGGCTCC
9061	GCCCCCTGA	CGAGCATCAC	AAAAATCGAC	GCTCAAGTCA	GAGGTGGCGA	AACCCGACAG
9121	GACTATAAAG	ATACCAGGCG	TTTCCCCCTG	GAAGCTCCCT	CGTGCGCTCT	CCTGTTCCGA
9181	CCCTGCCGCT	TACCGGATAC	CTGTCCGCCT	TTCTCCCTTC	GGGAAGCGTG	GCGCTTTCTC
9241	ATAGCTCACG	CTGTAGGTAT	CTCAGTTCCG	TGTAGGTCGT	TCGCTCCAAG	CTGGGCTGTG
9301	TGCACGAACC	CCCCGTTTCA	CCCGACCGCT	GCGCCTTATC	CGGTAACAT	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	CTCTGTGAA	GCCAGTTACC	TTCGAAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAACAAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAACCTTGG	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	CGCAGAAAGT	GTCTTGCAAC	TTTATCCGCC	TCCATCCAGT
661	CTATTAATTG	TTGCCGGGAA	GCTAGAGTAA	GTAGTTTCGCC	AGTTAATAGT	TTGCGCAACG
721	TTGTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGGTATG	GCTTCATTCA
781	GCTCCGGTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC	CATGTTGTGC	AAAAAAGCGG
841	TTAGCTCCTT	CGGTCTCTCG	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAAGT	TTATCACTCA
901	TGTTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTATGCTC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAA	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCGATGT	ACCCACTCGT	GCACCCAAT	GATCTTCAGC	ATCTTTTACT	TTACCAGCG
1201	TTTCTGGGTG	AGCAAAAAACA	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	CCTTTCTGCT	CGCGCGTTTC	GGTGATGACG
1501	GTGAAAACCT	CTGACACATG	CAGCTCCCGG	AGACGGTTCAC	AGCTTGCTCG	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGGCGCGT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAATATATG	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATACCGCA	TCAGGCGCCA	TTCGCCATTG	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAGGGG
1801	GATGTTGCTG	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TTCCAGTCA	CGACGTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCCGCGCCG	CCACTCCTGC	TCGAGCGCGC
1921	CCGCGCGTGC	GCCGCCAGCG	CCTTGCCCTT	TTCCGCGCGC	TCGTGCGCGT	CGCTGATGTC
1981	CATACCAGG	TCCATGAGGT	CTGCCTTGCG	CCGGCTGAGC	CACTGCTTCG	TCCGGGCGGC
2041	CAAGAGGAGC	ATGAGGGAGG	ACTCCTGGTC	CAGGGTCTCT	ACGTGGTTCG	GGCTCTGGGA

2101	GCGGGCCAGC	ATCATCTGGC	TCTGCCGCAC	CGAGGCCGCC	TCCAACCTGGT	CCTCCAGCAG
2161	CCGCAGTCGC	CGCCGACCCT	GGCAGAGGAA	GACAGGTGAG	GGGGGTATGA	ATTGTACAGA
2221	ACAACCACGA	GCCTTGTCTA	GGCAGAATCC	CTACCACTCA	TGGCTTTACC	TGGATGACGG
2281	CCTGCCAACA	GCTGTCCAGC	GACCTTCGCT	GCCGCCGCTT	CTCCCGCACG	CTTCTTTCCA
2341	GCACCGTGAT	GGCGCGAGCC	AGCGCCGCAC	GCTGGCGCTG	CGCTTCGCCG	ATCTGAGGAC
2401	AGTCGGGGAA	CTCTGATCAG	TCTAAACCCC	CTTGCGCGTT	AGTGTTGCCA	TCCTTTGCAG
2461	ACCGGTGAGA	GCCGACTTGT	TGTGCGCCAC	CCCCCACACC	ACCTCCTCCC	AGACCAATTC
2521	TGTCACCTTT	TTGGCGAAGG	CATCGGCCCTC	GGCCTGCAGA	GAGGACAGCA	GTGCCCAGCC
2581	GCTGGGGGTT	GGCGGATGCA	CGCTCAGGTA	CCCTTTCTTG	CGCTATGACA	CTTCCAGCAA
2641	AAGGTAGGGC	GGGCTGCGAG	ACGGCTTCCC	GGCGCTGCAT	GCAACACCGA	TGATGCTTCG
2701	ACCCCCCGAA	GCTCCTTCGG	GGCTGCATGG	GCGCTCCGAT	GCCGCTCCAG	GGCGAGCGCT
2761	GTTTAAATAG	CCAGGCCCCC	GATTGCAAAG	ACATTATAGC	GAGCTACCAA	AGCCATATTC
2821	AAACACCTAG	ATCACTACCA	CTTCTACACA	GGCCACTCGA	GCTGTGTATC	GCACTCCGCT
2881	AAGGGGGCGC	CTCTTCCTCT	TCGTTTCACT	CACAACCCGC	AAACTCTAGA	ATATCAATGA
2941	TCGAGCAGGA	CGGCCTCCAC	GCCGGCTCCC	CCGCCGCCTG	GGTGGAGCGC	CTGTTCCGGCT
3001	ACGACTGGGC	CCAGCAGACC	ATCGGCTGCT	CCGACGCCGC	CGTGTTCGCG	CTGTCCGCCC
3061	AGGGCCGCC	CGTGCTGTTC	GTGAAGACCG	ACCTGTCCGG	CGCCCTGAAC	GAGCTGCAGG
3121	ACGAGGCCGC	CCGCCTGTCC	TGGCTGGCCA	CCACCGGCGT	GCCCTGCGCC	GCCGTGCTGG
3181	ACGTGGTGAC	CGAGGCCGGC	CGCGACTGGC	TGCTGCTGGG	CGAGGTGCC	GGCCAGGACC
3241	TGCTGTCTC	CCACCTGGCC	CCCGCCGAGA	AGGTGTCCAT	CATGGCCGAC	GCCATGCGCC
3301	GCCTGCACAC	CCTGGACCCC	GCCACCTGCC	CCTTCGACCA	CCAGGCCAAG	CACCGCATCG
3361	AGCGCGCCCG	CACCCGCATG	GAGGCCGGCC	TGGTGGACCA	GGACGCACTG	GACGAGGAGC
3421	ACCAGGGCCT	GGCCCCCGCC	GAGCTGTTCG	CCCGCCTGAA	GGCCCCGCATG	CCCGACGGCG
3481	AGGACCTGGT	GGTGACCCAC	GGCGACGCCT	GCCTGCCCAA	CATCATGGTG	GAGAACGGCC
3541	GCTTCTCCGG	CTTCATCGAC	TGCGGCCGCC	TGGGCGTGGC	CGACCGCTAC	CAGGACATCG
3601	CCCTGGCCAC	CCGCACATC	GCCGAGGAGC	TGGGCGGCCA	GTGGGCCGAC	CGCTTCCTGG
3661	TGCTGTACGG	CATCGCCGCC	CCCGACTCCC	AGCGCATCGC	CTTCTACCGC	CTGCTGGACG
3721	AGTTCTTCTG	ACAATTGGCA	GCAGCAGCTC	GGATAGTATC	GACACACTCT	GGACGCTGGT
3781	CGTGIGATGG	ACTGTTGCCG	CCACACTTGG	TGCCTTGACC	TGTGAATATC	CCTGCCGCTT
3841	TTATCAAACA	GCCTCAGTGT	GTTTGATCTT	GTGTGTACGC	GCTTTTGCGA	GTGTGCTAGT
3901	GCTTGTGCTA	TTTGCGAATA	CCACCCCCAG	CATCCCCTTC	CCTCGCTTCA	TATCGCTTGC
3961	ATCCCAACCG	CAACTTATCT	ACGCTGTCTC	GCTATCCCTC	AGCGCTGCTC	CTGCTCCTGC
4021	TCCTGCCCC	TCGCACAGCC	TTGGTTTGGG	CTCCGCCTGT	ATTCTCCTGG	TACTGCAACC
4081	TGTAAACCAG	CACTGCAATG	CTGATGCACG	GGAAGTAGTG	GGATGGGAAC	ACAAATGGAG
4141	GATCCCGCGT	CTCGAACAGA	GCGCGCAGAG	GAACGCTGAA	GGTCTCGCCT	CTGTCGCACC
4201	TCAGCGCGGC	ATACACCACA	ATAACCACCT	GACGAATGCG	CTTGGTTCTT	CGTCCATTAG
4261	CGAAGCGTCC	GGTTCACACA	CGTGCCACGT	TGGCGAGGTG	GCAGGTGACA	ATGATCGGTG
4321	GAGCTGATGG	TCGAAACGTT	CACAGCCTAG	GGATATCGAA	TTCTTTTCTT	GCGCTATGAC
4381	ACTTCCAGCA	AAAGGTAGGG	CGGGCTGCGA	GACGGCTTCC	CGGCGCTGCA	TGCAACACCG
4441	ATGATGCTTC	GACCCCCCGA	AGCTCCTTCG	GGGCTGCATG	GCGCGTCCGA	TGCCGCTCCA
4501	GGGCGAGCGC	TGTTTAAATA	GCCAGGCCCC	CGATTGCAAA	GACATTATAG	CGAGCTACCA
4561	AAGCCATATT	CAAACACCTA	GATCACTACC	ACTTCTACAC	AGGCCACTCG	AGCTTGTGAT
4621	CGCACTCCGC	TAAGGGGGCG	CCTCTTCTCT	TTGTTTTTCA	TCACAACCCG	CAAACACTAG
4681	TATGGCCACC	GCATCCACTT	TCTCGGCGTT	CAATGCCCGC	TGCGGCGACC	TGCGTGCCTC
4741	GGCGGGCTCC	GGGCCCCCGC	GCCCAGCGAG	GCCCCCTCCC	GTGCGCGGGC	GCGCCCCCGA
4801	CTGGTCCATG	CTGTTTCGCC	TGATCACCAC	CATCTTCTCC	GCCGCCGAGA	AGCAGTGGAC
4861	CAACCTGGAG	TGGAAGCCCA	AGCCCAACCC	CCCCCAGCTG	CTGGAGGACC	ACTTTCGGCC
4921	CCACGGCCTG	GTGTTCCGCC	GCACCTTCGC	CATCCGCAGC	TACGAGGTGG	GCCCCGACCG
4981	CTCCACCAGC	ATCGTGGCCG	TGATGAACCA	CCTGCAGGAG	GCCGCCCTGA	ACCACGCCAA
5041	GTCCGTGGGC	ATCCTGGGCG	ACGGCTTCGG	CACCACCCTG	GAGATGTCCA	AGCGCGACCT
5101	GATCTGGGTG	GTGAAGCGCA	CCACAGTGGC	CGTGGAGCGC	TACCCCGCCT	GGGGCGACAC
5161	CGTGGAGGTG	GAGTGTCTGG	TGGGCGCCTC	CGGCAACAAC	GGCCGCCGCC	ACGACTTCCT
5221	GGTGC GCGAC	TGCAAGACCG	GCGAGATCCT	GACCCGCTGC	ACCTCCCTGA	GCGTGATGAT
5281	GAACACCCGC	ACCCGCCGCC	TGAGCAAGAT	CCCCGAGGAG	GTGCGCGGGC	AGATCGGCCC
5341	CGCCTTCATC	GACAACGTGG	CCGTGAAGGA	CGAGGAGATC	AAGAAGCCCC	AGAAGCTGAA
5401	CGACTCCACC	GCCGACTACA	TCCAGGGCGG	CCTGACCCCC	CGCTGGAACG	ACCTGGACAT
5461	CAACCAGCAC	GTGAACAACA	TCAAGTACGT	GGACTGGATC	CTGGAGACCG	TGCCCCGACG
5521	CATCTTCGAG	AGCCACCACA	TCTCTCTCTT	CACCATCGAG	TACCGCCGCG	AGTGCACCAT
5581	GGACAGCGTG	CTGCAGTCCC	TGACCACCGT	GAGCGGCGGC	TCTCCGAGG	CCGGCCTGGT
5641	GTGCGAGCAC	CTGCTGCAGC	TGGAGGGCGG	CAGCGAGGTG	CTGCGCGCCA	AGACCGAGTG
5701	GCGCCCCAAG	CTGACCGACT	CCTTCCGCGG	CATCAGCGTG	ATCCCCGCCG	AGTCCAGCGT

5761	GATGGACTAC	AAGGACCACG	ACGGCGACTA	CAAGGACCAC	GACATCGACT	ACAAGGACGA
5821	CGACGACAAG	TGACTCGAGG	CAGCAGCAGC	TCGGATAGTA	TCGACACACT	CTGGACGCTG
5881	GTCGTGTGAT	GGACTGTTGC	CGCCACACTT	GCTGCCTTGA	CCTGTGAATA	TCCCTGCCGC
5941	TTTTATCAAA	CAGCCTCAGT	GTGTTTGATC	TTGTGTGTAC	GCGCTTTTGC	GAGTTGCTAG
6001	CTGCTTGTGC	TATTTGCGAA	TACCACCCCC	AGCATCCCCT	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	AGCCTTTCAT	TCTCAGCCTC
6301	GATAACCTCC	AAAGCCGCTC	TAATTGTGGA	GGGGGTTCGA	ATTTAAAAGC	TTGGAATGTT
6361	GGTTCGTGCG	TCTGGAACAA	GCCCAGACTT	GTTGCTCACT	GGGAAAAGGA	CCATCAGCTC
6421	CAAAAACTT	GCCGCTCAAA	CCGCGTACCT	CTGCTTTCGC	GCAATCTGCC	CTGTTGAAAT
6481	CGCCACCACA	TTCATATTGT	GACGCTTGAG	CAGTCTGTAA	TTGCCTCAGA	ATGTGGAATC
6541	ATCTGCCCCC	TGTGCGAGCC	CATGCCAGGC	ATGTCGCGGG	CGAGGACACC	CGCCACTCGT
6601	ACAGCAGACC	ATTATGCTAC	CTCACAATAG	TTCATAACAG	TGACCATATT	TCTCGAAGCT
6661	CCCCAACGAG	CACCTCCATG	CTCTGAGTGG	CCACCCCCCG	GCCCTGGTGC	TTGCGGAGGG
6721	CAGGTCAACC	GGCATGGGGC	TACCGAAATC	CCCAGCCGGA	TCCCACCACC	CCCGCGATGG
6781	GAAGAATCTC	TCCCCGGGAT	GTGGGCCCCAC	CACCAGCACAC	ACCTGCTGGC	CCAGGCGAGC
6841	GTCAAACCAT	ACCACACAAA	TATCCTTGGC	ATCGGCCCTG	AATTCCCTTCT	GCCGCTCTGC
6901	TACCCGGTGC	TTCTGTCCGA	AGCAGGGGTT	GCTAGGGATC	GCTCCGAGTC	CGCAAACCCCT
6961	TGTCGCGTGG	CGGGGCTTGT	TCGAGCTTGA	AGAGCCTCTA	GAGTCGACCT	GACGGCATGC
7021	AAGCTTGGCG	TAATCATGGT	CATAGCTGTT	TCCTGTGTGA	AATTGTTATC	CGCTCACAAT
7081	TCCACACAAC	ATACGAGCCG	GAAGCATAAA	GTGTAAAGCC	TGGGGTGCCCT	AATGAGTGAG
7141	CTAACTCACA	TTAATTGCGT	TGCGCTCACT	GCCCGCTTTC	CAGTCGGGAA	ACCTGTCGTG
7201	CCAGCTGCAT	TAATGAATCG	GCCAACGCGC	GGGAGAGGC	GGTTTGCGTA	TTGGGCGCTC
7261	TTCCGCTTCC	TCGCTCACTG	ACTCGCTGCG	CTCGGTGCTT	CGGCTGCGGC	GAGCGGTATC
7321	AGCTCACTCA	AAGGCGGTAA	TACGGTTATC	CACAGAATCA	GGGGATAACG	CAGGAAAGAA
7381	CATGTGAGCA	AAAGGCCAGC	AAAAGGCCAG	GAACCGTAAA	AAGGCCGCGT	TGCTGGCGTT
7441	TTTCCATAGG	CTCCGCCCCC	CTGACGAGCA	TCACAAAAAT	CGACGCTCAA	GTCAGAGGTG
7501	GCGAAACCCG	ACAGGACTAT	AAAGATACCA	GGCGTTTCCC	CCTGGAAGCT	CCCTCGTGCG
7561	CTCTCCGTGT	CCGACCCTGC	CGCTTACCGG	ATACCTGTCC	GCCTTTCTCC	CTTCGGGAAG
7621	CGTGCGGCTT	TCTCATAGCT	CACGCTGTAG	GTATCTCAGT	TCGGTGTAGG	TCGTTTCGCTC
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	GTGGTGCCCT	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTCGAAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAAC
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAACCTTGG	TCTGACAGTT
421	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTCTG	TATCCATAG
481	TTGCCCTGACT	CCCCGTCGTG	TAGATAACTA	CGATACGGGA	GGGCTTACCA	TCTGGCCCCA
541	GTGCTGCAAT	GATACGCGA	GACCCACGCT	CACCGGCTCC	AGATTTATCA	GCAATAAACC
601	AGCCAGCCGG	AAGGGCCGAG	CGCAGAAGTG	GTCTTGCAAC	TTTATCCGCC	TCCATCCAGT
661	CTATTAATTG	TTGCCGGGAA	GCTAGAGTAA	GTAGTTTCGCC	AGTTAATAGT	TTGCGCAACG
721	TTGTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGGTATG	GCTTCATTCA
781	GCTCCGGTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC	CATGTTGTGC	AAAAAAGCGG
841	TTAGTCTCCT	CGGTCTCTCG	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAAGT	TTATCACTCA
901	TGTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTATGCCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTACTCAACC	AAGTCAATTCT	GAGAATAGTG	TATGCGGCCA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACCTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAAC	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCGATGTA	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 CTTTTCGTCT CGCGCGTTTC GGTGATGACG
1501 GTGAAAACCT CTGACACATG CAGCTCCCG AGACGGTCAC AGCTTGCTG TAAGCGGATG
1561 CCGGGAGCAG ACAAGCCCGT CAGGCGCGT CAGCGGGTGT TGGCGGGTGT CGGGGCTGGC
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 TCCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTC TTCGCCGCCG CCACTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCGC TCGTGCCTG CGCTGATGTC
1981 CATCACCAGG TCCATGAGGT CTGCCTTGGC CCGGCTGAGC CACTGCTTCG TCCGGGCGGC
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCCTG ACGTGGTCGC GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAACTGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCCT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGTCTA GGCAGAATCC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCTCGCT GCCGCCGCTT CTCCGCGACG CTTCTTTCCA
2341 GCACCGTGAT GCGCGGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGCCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTACCTTT TTGGCGAAGG CATCGGCCCT GGCCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GGCGGATGCA CGCTCAGGTA CCCTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GCGAGCGCT
2761 GTTTAAATAG CCAGGCCCC GATTGCAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTGTGATC GCACTCCGCT
2881 AAGGGGCGC CTCTTCCTCT TCGTTTCAGT CACAACCCGC AAACCTCTAGA ATATCAATGC
2941 TGCTGCAGGC CTTCTGTTC CTGCTGGCCG GCTTCGCCG CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTG ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCAAACGG CCTGTGGTAC GACGAGAAGG ACGCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCCGAACGA CACCGTCTGG GGGACGCCCT TGTCTGGGG 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 CAGCCTGGAC GGCGCTACA CTTTACCGA GTACCAGAAG AACCCCGTGC
3421 TGGCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCCAGA
3481 AGTGGATCAT GACCGCGGCC AAGTCCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCCGT TCGCCAACGA GGGCTTCCTC GGCTACCACT
3601 ACGAGTGCCC CGGCCTGATC GAGGTCCCCA CCGAGCAGG CCCCAGCAAG TCCTACTGGG
3661 TGATGTTTCA CTCCATCAAC CCCGCGGCC CGGCCGCGCG CTCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CTTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCTGGGCAT CGCGTGGGCC TCCAAGTGG AGTACTCCGC CTTGCTGCC ACCAACCCT
3901 GGCGTCCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCC AGCCGATCCT GAACATCAGC AACCGCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTCGACC
4081 TGTCCAACAG CACCGGCACC CTGGAGTTC AGCTGGTGTA CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT GCGGACCTCT CCCTCTGGTT CAAGGGCTG GAGGACCCG
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 GTTCAGGTG CCGGAGGTCA AGTGACAATT GGCAGCAGCA GCTCGGATAG
4561 TATCGACACA CTCTGGACGC TGTCGTGTG ATGGACTGTT GCCGCCACAC TTGCTGCCTT
4621 GACCTGTGAA TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTGA TCTGTGTGT
4681 ACGCGCTTTT GCGAGTTGCT AGCTGCTGTG GCTATTTGCG AATACCACC CACGACTCCC
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	GCCTCTGTCTG	CACCTCAGCG	CGGCATACAC	CACAATAAACC	ACCTGACGAA
5041	TGCGCTTGGT	TCTTCGTCCA	TTAGCGAAGC	GTCCGGTTCA	CACACGTGCC	ACGTTGGCGA
5101	GGTGGCAGGT	GACAATGATC	GGTGGAGCTG	ATGGTCGAAA	CGTTCACAGC	CTAGGGATAT
5161	CGAATTCCTT	TCTTGCGCTA	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	CCCCGTGCGC	GGGCGCGCCG	CCTCCGCCAA	GGCCTCCAAG	ACCATCGGCG	AGGGCTCCGA
5641	GAACCTGAAC	GTGCGCGGCA	TGATCGCCAA	GCCCACCTCC	TCCTCCGCCG	CCAAGCAGGG
5701	CAAGGTGATG	GCCCAGGCCG	TGCCCAAGAT	CAACGGCGCC	AAGGTGGGCC	TGAAGGCCGA
5761	GTCCCAAGAA	GCCGAGGAGG	ACGCCGCCCC	CTCCTCCGCC	CCCCGCACCT	TCTACAACCA
5821	GCTGCCCCGAC	TGGTCCGTGC	TGCTGGCCGC	CGTGACCACC	ATCTTCCTGG	CCGCCGAGAA
5881	GCAGTGGACC	CTGCTGGACT	GGAAGCCCCG	CCGCCCCGAC	ATGCTGACCG	GCGCCTTCTC
5941	CCTGGGCAAG	ATCGTGCAGG	ACGGCTGGT	GTTCCGCCAG	AACTTCTCCA	TCCGCTCCTA
6001	CGAGATCGGC	GCCGACCGCA	CCGCTCCAT	CGAGACCCTG	ATGAACCACC	TGCAGGAGAC
6061	CGCCCTGAAC	CACGTGCGCA	ACGCCGGCCT	GCTGGGCGAC	GGCTTCGGCG	CCACCCCCGA
6121	GATGTCCAAG	CGCAACCTGA	TCTGGGTGGT	GACCAAGATG	CAGGTGCTGA	TCGAGCACTA
6181	CCCCTCCTGG	GGCGACGTGG	TGGAGGTGGA	CACCTGGGTG	GGCGCCTCCG	GCAAGAACGG
6241	CATGCGCCGC	GACTGGCACG	TGCGCGACTA	CCGCACCGGC	CAGACCATCC	TGCGCGCCAC
6301	CTCCATCTGG	GTGATGATGG	ACAAGCACAC	CCGCAAGCTG	TCCAAGATGC	CCGAGGAGGT
6361	GCGCGCCGAG	ATCGGCCCT	ACTTCATGGA	GCACGCCGCC	ATCGTGGACG	AGGACTCCCCG
6421	CAAGCTGCCC	AAGCTGGACG	ACGACACCGC	CGACTACATC	AAGTGGGGCC	TGACCCCCCG
6481	CTGGTCCGAC	CTGGACGTGA	ACCAGCACGT	GAACAACGTG	AAGTACATCG	GCTGGATCCT
6541	GGAGTCCGCC	CCCATCTCCA	TCCTGGAGAA	CCACGAGCTG	GCCTCCATGA	CCCTGGAGTA
6601	CCGCCGCGAG	TGCGGCCGCG	ACTCCGTGCT	GCAGTCCCTG	ACCGCCGTGG	CCAACGACTG
6661	CACCGCGGGC	CTGCCCGAGG	CCTCCATCGA	GTGCCAGCAC	CTGCTGCACG	TGGAGTGGCG
6721	CGCCGAGATC	GTGCGCGGCC	GCACCCAGTG	GCGCCCCCGC	GCGCCTCCG	GCCCCACCTC
6781	CGCCGGCTCC	GCCATGGACT	ACAAGGACCA	CGACGGCGAC	TACAAGGACC	ACGACATCGA
6841	CTACAAGGAC	GACGACGACA	AGTGAATCGA	TAGATCTCTT	AAGGCAGCAG	CAGCTCGGAT
6901	AGTATCGACA	CACCTCTGGAC	GCTGGTCGTG	TGATGGACTG	TTGCCGCCAC	ACTTGCTGCC
6961	TTGACCTGTG	AATATCCCTG	CCGCTTTTAT	CAAACAGCCT	CAGTGTGTTT	GATCTTGTGT
7021	GTACGCCTT	TTGCGAGTTG	CTAGCTGCTT	GTGCTATTTG	CGAATACCAC	CCCCAGCATC
7081	CCCTTCCCTC	GTTTCATATC	GCTTGCACTC	CAACCGCAAC	TTATCTACGC	TGTCCTGCTA
7141	TCCCTCAGCG	CTGCTCCTGC	TCCTGCTCAC	TGCCCCCTCG	ACAGCCTTGG	TTTGGGCTCC
7201	GCCTGTATTC	TCCTGGTACT	GCAACCTGTA	AACCAGCACT	GCAATGCTGA	TGCACGGGAA
7261	GTAGTGGGAT	GGAACACAA	ATGGAAGAGT	TAATTAAGAG	CTCTTGTTTT	CCAGAAGGAG
7321	TTGCTCCTTG	AGCCTTTTCAT	TCTCAGCCTC	GATAACCTCC	AAAGCCGCTC	TAATTGTGGA
7381	GGGGGTTTCA	ATTTAAAAGC	TTGGAATGTT	GGTTTCGTGC	TCTGGAACAA	GCCCAGACTT
7441	GTTGCTCACT	GGGAAAAGGA	CCATCAGCTC	CAAAAACTT	GCCGCTCAAA	CCGCGTACCT
7501	CTGCTTTTCG	GCAATCTGCC	CTGTTGAAAT	CGCCACCACA	TTCATATTGT	GACGCTTGAG
7561	CAGTCTGTAA	TTGCCTCAGA	ATGTGGAATC	ATCTGCCCCC	TGTGCGAGCC	CATGCCAGGC
7621	ATGTCGCGGG	CGAGGACACC	CGCCACTCGT	ACAGCAGACC	ATTATGCTAC	CTCACAATAG
7681	TTCATAACAG	TGACCATATT	TCTCGAAGCT	CCCCAACGAG	CACCTCCATG	CTCTGAGTGG
7741	CCACCCCCCG	GCCCTGGTGC	TTGCGGAGGG	CAGGTCAACC	GGCATGGGGG	TACCGAAATC
7801	CCCAGCCGGA	TCCCACCACC	CCCGCATGGG	GAAGAATCTC	TCCCCGGGAT	GTCGGCCCAT
7861	CACCAGCACA	ACCTGCTGGC	CCAGCGGAGC	GTCAAACCAT	ACCACACAAA	TATCCTTGGC
7921	ATCGGCCCTG	AATTCTTTCT	GCCGCTCTGC	TACCCGGTGC	TTCTGTCCGA	AGCAGGGGTT
7981	GCTAGGGATC	GCTCCGAGTC	CGCAAACCTT	TGTCGCGTGG	CGGGGCTTGT	TCGAGCTTGA
8041	AGAGCCTCTA	GAGTCGACCT	GCAGGCATGC	AAGCTTGGCG	TAATCATGGT	CATAGCTGTT
8101	TCCTGTGTGA	AATTGTTATC	CGCTCACAAT	TCCACACAAC	ATACGAGCCG	GAAGCATAAA
8161	GTGTAAAGCC	TGGGGTGCCT	AATGAGTGAG	CTAACTCACA	TTAATTGCGT	TGCGCTCACT
8221	GCCCGCTTTC	CAGTCGGGAA	ACCTGTCGTG	CCAGCTGCAT	TAATGAATCG	GCCAACGCGC
8281	GGGGAGAGGC	GGTTTGCGTA	TTGGGCGCTC	TTCCGCTTCC	TCGCTCACTG	ACTCGTGGC
8341	CTCGTCTGTT	CGGCTGCGGC	GAGCGGTATC	AGCTCACTCA	AAGGCGGTAA	TACGGTTATC
8401	CACAGAATCA	GGGGATAACG	CAGGAAAGAA	CATGTGAGCA	AAAGGCCAGC	AAAAGGCCAG
8461	GAACCGTAAA	AAGGCCGCGT	TGCTGGCGTT	TTTCCATAGG	CTCCGCCCCC	CTGACGAGCA
8521	TCACAAAAAT	CGACGCTCAA	GTCAGAGGTG	GCGAAACCCG	ACAGGACTAT	AAAGATACCA
8581	GGCGTTTCCC	CCTGGAAGCT	CCCTCGTGCG	CTCTCCTGTT	CCGACCCTGC	CGCTTACCGG

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

SEQ ID NO: 238
pSZ1449

1	CACTGGCAGC	AGCCACTGGT	AACAGGATTA	GCAGAGCGAG	GTATGTAGGC	GGTGCTACAG
61	AGTTCTTGAA	GTGGTGGCCT	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTCGAAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT
421	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTTCGT	TCATCCATAG
481	TTGCCTGACT	CCCCGTCGTG	TAGATAACTA	CGATACGGGA	GGGCTTACCA	TCTGGCCCCA
541	GTGCTGCAAT	GATACCGCGA	GACCCACGCT	CACCGGCTCC	AGATTTATCA	GCAATAAACC
601	AGCCAGCCGG	AAGGGCCGAG	CGCAGAAGTG	GTCTTGCAAC	TTATCCGCC	TCCATCCAGT
661	CTATTAATTG	TTGCCGGGAA	GCTAGAGTAA	GTAGTTCGCC	AGTTAATAGT	TTGCGCAACG
721	TTGTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGGTATG	GCTTCATTCA
781	GCTCCGGTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC	CATGTTGTGC	AAAAAAGCGG
841	TTAGTCTCTT	CGGTCTCTCG	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAGTG	TTATCACTCA
901	TGGTTATGGC	AGCACTGCAT	AATTCCTTTA	CTGTCATGCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAG	ACGTTCTTCG	GGGCGAAAAA	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCGATGA	ACCCACTCGT	GCACCCAATC	GATCTTCAGC	ATCTTTTACT	TTCAACGCGC
1201	TTTCTGGGTG	AGCAAAAAACA	GGAAGGCAAA	ATGCCGCAAA	AAAGGGGAATA	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	CAGCTCCCCG	AGACGGTCAC	AGCTTGTCTG	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGGCGCGT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATACCGCA	TCAGGCGCCA	TTCGCCATTC	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAAGGG
1801	GATGTGCTGC	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TTCCCAGTCA	CGACGTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCGCCGCCG	CCACTCCTGC	TCGAGCGCGC
1921	CCGCGCTGTC	GCCGCCAGCG	CCTTGGCCTT	TTCGCCGCGC	TCGTGCGCGT	CGCTGATGTC
1981	CATACCAGG	TCCATGAGGT	CTGCCTTGCG	CCGGCTGAGC	CACTGCTTCG	TCCGGGCGGC
2041	CAAGAGGAGC	ATGAGGGAGG	ACTCCTGGTC	CAGGGTCCTG	ACGTGGTTCG	GGCTCTGGGA
2101	GCGGGCCAGC	ATCATCTGGC	TCTGCCGCAC	CGAGGCCGCC	TCCAACCTGG	CCTCCAGCAG
2161	CCGCAGTCGC	CGCCGACCCT	GGCAGAGGAA	GACAGGTGAG	GGGGGTATGA	ATTGTACAGA
2221	ACAACCAACA	GCCTTGTCCTA	GGCAGATACC	CTACCACTCA	TGGCTTTTACC	TGGATGACGG
2281	CCTGCCAACA	GCTGTCCAGC	GACCTTCGCT	GCCGCCGCTT	CTCCCGCAGC	CTCTTTTCCA
2341	GCACCGTGAT	GGCGCGAGCC	AGCGCCGCAC	GCTGGCGCTG	CGCTTCGCCG	ATCTGAGGAC
2401	AGTCGGGGAA	CTCTGATCAG	TCTAAACCCC	CTTGCGCGTT	AGTGTTGCCA	TCCTTTGCAG
2461	ACCGGTGAGA	GCCGACTTGT	TGTGCGCCAC	CCCCACACC	ACCTCCTCCC	AGACCAATTC
2521	TGTCACCTTT	TTGGCGAAGG	CATCGGCCTC	GGCCTGCAGA	GAGGACAGCA	GTGCCAGGCC
2581	GCTGGGGGTT	GGCGGATGCA	CGCTCAGGTA	CCCTTTCTTG	CGCTATGACA	CTTCCAGCAA
2641	AAGGTAGGGC	GGGCTGCGAG	ACGGCTTCCC	GGCGCTGCAT	GCAACACCGA	TGATGCTTCG
2701	ACCCCCCGAA	GCTCCTTCGG	GGCTGCATGG	GCGCTCCGAT	GCCGCTCCAG	GGCGAGCGCT
2761	GTTTAAATAG	CCAGGCCCCC	GATTGCAAGG	ACATTATAGC	GAGCTACCAA	AGCCATATTTC
2821	AAACACCTAG	ATCACTACCA	CTTCTACACA	GGCCACTCGA	GCTTGTGATC	GCACTCCGCT
2881	AAGGGGGCGC	CTCTTCCTCT	TCGTTTTCAGT	CACAACCCGC	AAACTCTAGA	ATATCAATGC
2941	TGCTGCAGGC	CTTCCTGTTC	CTGCTGGCCG	GCTTCGCCGC	CAAGATCAGC	GCCTCCATGA
3001	CGAACGAGAC	GTCCGACCGC	CCCTTGGTGC	ACTTCACCCC	CAACAAGGGC	TGGATGAACG
3061	ACCCCAACGG	CCTGTGGTAC	GACGAGAAGG	ACGCCAAGTG	GCACCTGTAC	TTCCAGTACA

3121	ACCCGAACGA	CACCGTCTGG	GGGACGCCCT	TGTTCTGGGG	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	GGCGGCTACA	CCTTCACCGA	GTACCAGAAG	AACCCCGTGC
3421	TGGCCGCCAA	CTCCACCCAG	TTCCGCGACC	CGAAGGTCTT	CTGGTACGAG	CCCTCCCAGA
3481	AGTGGATCAT	GACCGCGGCC	AAGTCCAGG	ACTACAAGAT	CGAGATCTAC	TCCTCCGACG
3541	ACCTGAAGTC	CTGGAAGCTG	GAGTCCGCGT	TCGCCAACGA	GGGCTTCCTC	GGCTACCAGT
3601	ACGAGTGCCC	CGGCCTGATC	GAGGTCCCCA	CCGAGCAGGA	CCCCAGCAAG	TCCTACTGGG
3661	TGATGTTTCA	CTCCATCAAC	CCCGGCGCCC	CGGCCGGCGG	CTCCTTCAAC	CAGTACTTCG
3721	TCGGCAGCTT	CAACGGCACC	CACTTCGAGG	CCTTCGACAA	CCAGTCCCAG	GTGGTGGACT
3781	TCGGCAAGGA	CTACTACGCC	CTGCAGACCT	TCTTCAACAC	CGACCCGACC	TACGGGAGCG
3841	CCCTGGGCAT	CGCGTGGGCC	TCCAACCTGG	AGTACTCCGC	CTTCGTGCC	ACCAACCCCT
3901	GGCGTCCCTC	CATGTCCCTC	GTGCGCAAGT	TCTCCCTCAA	CACCGAGTAC	CAGGCCAAC
3961	CGGAGACGGA	GCTGATCAAC	CTGAAGGCCG	AGCCGATCCT	GAACATCAGC	AACGCCGGCC
4021	CCTGGAGCCG	GTTTCGCCACC	AACACCACGT	TGACGAAGGC	CAACAGCTAC	AACGTCGACC
4081	TGTCCAACAG	CACCGGCACC	CTGGAGTTCG	AGCTGGTGTA	CGCCGTCAAC	ACCACCCAGA
4141	CGATCTCCAA	GTCCGTGTTC	GCGGACCTCT	CCCTCTGGTT	CAAGGGCCTG	GAGGACCCCG
4201	AGGAGTACCT	CCGCATGGGC	TTCGAGGTGT	CCGCGTCCTC	CTTCTTCCTG	GACCGCGGGA
4261	ACAGCAAGGT	GAAGTTCGTG	AAGGAGAACC	CCTACTTCAC	CAACCCGATG	AGCGTGAACA
4321	ACCAGCCCTT	CAAGAGCGAG	AACGACCTGT	CCTACTACAA	GGTGTACGGC	TTGCTGGACC
4381	AGAACATCCT	GGAGCTGTAC	TTCAACGACG	GCGACGTTCG	GTCCACCAAC	ACCTACTTCA
4441	TGACCACCGG	GAACGCCCTG	GGTCCGTGTA	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	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	GCCTCTGTGC	CACCTCAGCG	CGGCATACAC	CACAATAACC	ACCTGACGAA
5041	TGCGCTTGGT	TCTTCGTCCA	TTAGCGAAGC	GTCCGGTTCA	CACACGTGCC	ACGTTGGCGA
5101	GGTGGCAGGT	GACAATGATC	GGTGGAGCTG	ATGGTCGAAA	CGTTCACAGC	CTAGGGATAT
5161	CGAATTCGGC	CGACAGGACG	CGCGTCAAAG	GTGCTGGTCG	TGTATGCCCT	GGCCGGCAGG
5221	TCGTTGCTGC	TGCTGGTTAG	TGATTCCGCA	ACCCTGATTT	TGGCGTCTTA	TTTTGGCGTG
5281	GCAAACGCTG	GCGCCCCGCA	GCCGGGCCGG	CGGCGATGCG	GTGCCCCACG	GCTGCCGGAA
5341	TCCAAGGGAG	GCAAGAGCGC	CCGGGTTCAG	TGAAGGGCTT	TACGCGCAAG	GTACAGCCGC
5401	TCCTGCAAGG	CTGCGTGGTG	GAATTGGACG	TGCAGGTCCT	GCTGAAGTTC	CTCCACCGCC
5461	TCACCAGCGG	ACAAAGCACC	GGTGTATCAG	GTCCGTGTCA	TCCACTCTAA	AGAACTCGAC
5521	TACGACCTAC	TGATGGCCCT	AGATTCTTCA	TCAAAAACGC	CTGAGACACT	TGCCCAGGAT
5581	TGAAACTCCC	TGAAGGGACC	ACCAGGGGCC	CTGAGTTGTT	CCTTCCCCCC	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	CCTCAACCTT	AGGTATGCGC	GCATGCGGTC	GCCGCGCAAC
5881	TCGCGCGAGG	GCCGAGGGTT	TGGGACGGGC	GGTCCCGAAA	TGCAGTTGCA	CCCGGATGCG
5941	TGGCACCTTT	TTTGCGATAA	TTTATGCAAT	GGACTGCTCT	GCAAAAATTCT	GGCTGTGTCG
6001	CCAACCCTAG	GATCAGCGGC	GTAGGATTTT	GTAATCATTC	GTCTGATGG	GGAGCTACCG
6061	ACTACCTTAA	TATCAGCCCG	ACTGCCTGAC	GCCAGCGTCC	ACTTTTGTGC	ACACATTCCA
6121	TTCGTGCCCA	AGACATTTCA	TTGTGGTGCG	AAGCGTCCCC	AGTTACGCTC	ACCTGTTTCC
6181	CGACCTCCTT	ACTGTTCTGT	CGACAGAGCG	GGCCACAGG	CCGGTCGCAG	CCACTAGTAT
6241	GGCCACCGCA	TCCACTTTCT	CGGCGTTCAA	TGCCCCGCTG	GGCGACCTGC	GTGCTCGGC
6301	GGGCTCCGGG	CCCCGGCGCC	CAGCGAGGCC	CCTCCCCGTG	CGCGGGCGCG	CCTCCCAGCT
6361	GCGCAAGCCC	GCCCTGGACC	CCCTGCGCGC	CGTGATCTCC	GCCGACCAGG	GCTCCATCTC
6421	CCCCGTGAAC	TCCTGCACCC	CCGCCGACCG	CCTGCGCGCC	GGCCGCTGTA	TGAGGACGG
6481	CTACTCCTAC	AAGGAGAAGT	TCAATCGTGC	CTCCTACGAG	GTGGGACTCA	ACAAGACCGC
6541	CACCGTGGAG	ACCATCGCCA	ACCTGCTGCA	GGAGGTGGCC	TGCAACCACG	TGCAGAAGTG
6601	CGGCTTCTCC	ACCGACGGCT	TCGCCACCAC	CCTGACCATG	CGCAAGCTGC	ACCTGATCTG
6661	GGTGACCGCC	CGCATGCACA	TCGAGATCTA	CAAGTACCCC	GCCTGGTCCG	ACGTGGTGGA
6721	GATCGAGACC	TGGTGCCAGT	CCGAGGGCCG	CATCGGCACC	CGCCGCGACT	GGATCCTGCG

6781	CGACTCCGCC	ACCAACGAGG	TGATCGGCCG	CGCCACCTCC	AAGTGGGTGA	TGATGAACCA
6841	GGACACCCGC	CGCCTGCAGC	GCGTGACCGA	CGAGGTGCGC	GACGAGTACC	TGGTGTTCTG
6901	CCCCCGCGAG	CCCCGCCTGG	CCTTCCCCGA	GGAGAACAAAC	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	CAAGTTCACC	GGCACCAACG	GCTCCGCCAT	GTCTCTCCATC	CAGGGCCACA	ACGAGTCCCA
7261	GTTCTGCGAC	ATGCTGCGCC	TGTCCGAGAA	CGGCCAGGAG	ATCAACCGCG	GCCGCACCCA
7321	GTGGCGCAAG	AAGTCCTCCC	GCATGGACTA	CAAGGACCAC	GACGGCGACT	ACAAGGACCA
7381	CGACATCGAC	TACAAGGACG	ACGACGACAA	GTGAATCGAT	AGATCTCTTA	AGGCAGCAGC
7441	AGCTCGGATA	GTATCGACAC	ACTCTGGACG	CTGGTCGTGT	GATGGACTGT	TGCCGCCACA
7501	CTTGCTGCCT	TGACCTGTGA	ATATCCCTGC	CGCTTTTATC	AAACAGCCTC	AGTGTGTTTG
7561	ATCTTGTTGT	TACGCGCTTT	TGCGAGTTGC	TAGCTGCTTG	TGCTATTTGC	GAATACCACC
7621	CCCAGCATCC	CCTTCCCTCG	TTTCATATCG	CTTGCATCCC	AACCGCAACT	TATCTACGCT
7681	GTCTTGCTAT	CCCTCAGCGC	TGCTCCTGCT	CCTGCTCACT	GCCCCTCGCA	CAGCCTTGGT
7741	TTGGGCTCCG	CCTGTATTCT	CCTGGTACTG	CAACCTGTAA	ACCAGCACTG	CAATGCTGAT
7801	GCACGGGAAG	TAGTGGGATG	GGAACACAAA	TGGAAAGCTT	AATTAAGAGC	TCTTGTTTTT
7861	CAGAAGGAGT	TGCTCCTTGA	GCCTTTCATT	CTCAGCCTCG	ATAACCTCCA	AAGCCGCTCT
7921	AATTGTGGAG	GGGGTTCGAA	TTTAAAAGCT	TGGAATGTTG	GTTCTGCGT	CTGGAACAAAG
7981	CCCAGACTTG	TTGCTCACTG	GGAAAAGGAC	CATCAGCTCC	AAAAAACTTG	CCGCTCAAAAC
8041	CGCGTACCTC	TGCTTTCGCG	CAATCTGCCG	TGTTGAAATC	GCCACCACAT	TCATATTGTG
8101	ACGCTTGAGC	AGTCTGTAAT	TGCTCAGAA	TGTGGAATCA	TCTGCCCCCT	GTGCGAGCCC
8161	ATGCCAGGCA	TGTCGCGGGC	GAGGACACCC	GCCACTCGTA	CAGCAGACCA	TTATGCTACC
8221	TCACAATAGT	TCATAACAGT	GACCATATTT	CTCGAAGCTC	CCCAACGAGC	ACCTCCATGC
8281	TCTGAGTGGC	CACCCCCCGG	CCCTGGTGCT	TGCGGAGGGC	AGGTCAACCG	GCATGGGGCT
8341	ACCGAAATCC	CCGACCGGAT	CCCACCACCC	CCGCGATGGG	AAGAATCTCT	CCCCGGGATG
8401	TGGGCCCACC	ACCAGCACAA	CCTGCTGGCC	CAGGCGAGCG	TCAAACCATA	CCACACAAAT
8461	ATCCTTGGCA	TCGGCCCTGA	ATTCTTCTG	CCGCTCTGCT	ACCCGGTGCT	TCTGTCCGAA
8521	GCAGGGGTTG	CTAGGGATCG	CTCCGAGTCC	GCAAACCCTT	GTCGCGTGCG	GGGGCTTGTT
8581	CGAGCTTGAA	GAGCCTCTAG	AGTCGACCTG	CAGGCATGCA	AGCTTGCGGT	AATCATGGTC
8641	ATAGCTGTTT	CCTGTGTGAA	ATTGTTATCC	GCTCACAATT	CCACACAACA	TACGAGCCGG
8701	AAGCATAAAG	TGTAAAGCCT	GGGGTGCCTA	ATGAGTGAGC	TAACTCACAT	TAATTGCGTT
8761	GCGCTCACTG	CCCGCTTTCC	AGTCGGGAAA	CCTGTCTGTC	CAGCTGCATT	AATGAATCGG
8821	CCAACGCGCG	GGGAGAGGCG	GTTTGCGTAT	TGGGCGCTCT	TCCGCTTCCT	CGCTCACTGA
8881	CTCGTGCCTG	TCGGTTCGTT	GGCTGCGGCG	AGCGGTATCA	GCTCACTCAA	AGGCGGTAAT
8941	ACGGTTATCC	ACAGAATCAG	GGGATAACGC	AGGAAAGAAC	ATGTGAGCAA	AAGGCCAGCA
9001	AAAGGCCAGG	AACCGTAAAA	AGGCCGCGTT	GCTGGCGTTT	TTCCATAGGC	TCCGCCCCCC
9061	TGACGAGCAT	CACAAAAATC	GACGCTCAAG	TCAGAGGTGG	CGAAACCCGA	CAGGACATATA
9121	AAGTATACCAG	GCGTTTCCCC	CTGGAAGCTC	CCTCGTGCGC	TCTCCTGTTC	CGACCTTGCC
9181	GCTTACCGGA	TACCTGTCCG	CCTTCTCCCC	TTGCGGAAGC	GTGGCGCTTT	CTCATAGCTC
9241	ACGCTGTAGG	TATCTCAGTT	CGGTGTAGGT	CGTTTCGCTC	AAGCTGGGCT	GTGTGCACGA
9301	ACCCCCCGTT	CAGCCCGACC	GCTGCGCCTT	ATCCGGTAAC	TATCGTCTTG	AGTCCAACCC
9361	GGTAAGACAC	GACTTATCGC				

SEQ ID NO: 239
pSZ1451

1	CAC TGGCAGC	AGCCACTGGT	AACAGGATTA	GCAGAGCGAG	GTATGTAGGC	GGTGCTACAG
61	AGTTCTTGAA	GTGGTGCCCT	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT
421	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTCTG	TCATCCATAG
481	TTGCCTGACT	CCCCGTCTGT	TAGATAACTA	CGATACGGGA	GGGCTTACCA	TCTGGCCCCA
541	GTGCTGCAAT	GATACCGCGA	GACCCACGCT	CACCGGCTCC	AGATTTATCA	GCAATAAACC
601	AGCCAGCCGG	AAGGGCCGAG	CGCAGAAAGT	GTCCTGCAAC	TTTATCCGCC	TCCATCCAGT
661	CTATTAATTG	TTGCCGGGAA	GCTAGAGTAA	GTAGTTTCGCC	AGTTAATAGT	TTGCGCAACG

721 TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG GCTTCATTCA
781 GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC CATGTTGTGC AAAAAAGCGG
841 TTAGTCTCCTT CGGTCTCTCCG ATCGTTGTCA GAAGTAAGTT GGCCGCAGTG TTATCACTCA
901 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG
961 TGA CTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
1021 CTTGCCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA
1081 TCATTGGAAA ACGTTCTTCG GGGCGAAAAC 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 TCCCCGAAA 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 TTA ACTATGC GGCATCAGAG CAGATTGTAC TGAGAGTGCA CCATATGCGG TGTGAAATAC
1681 CGCACAGATG CGTAAGGAGA AAATACCGCA TCAGGCGCCA TTCGCCATT CAGGCTGCGCA
1741 ACTGTTGGGA AGGGCGATCG GTGCGGGCCT CTTGCTATT ACGCCAGCTG GCGAAAGGGG
1801 GATGTGCTGC AAGGCGATTA AGTTGGGTAA CGCCAGGGTT TTCCAGTCA CGACGTTGTA
1861 AAACGACGGC CAGTGAATTG ATGCATGCTT TTCGCCGCG CCCTCCTGC TCGAGCGCGC
1921 CCGCGCGTGC GCCGCCAGCG CCTTGGCCTT TTCGCCGCG TCGTGCCTG CGCTGATGTC
1981 CATCCACAGG TCCATGAGGT CTGCTTTCG CCGGCTGAGC CACTGCTTCG TCCGGGCGCG
2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCCTG ACGTGGTTCG GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCG TCCA ACTGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCCT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGTCTA GGCAGAATCC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCTCGCT GCCGCCGCTT CTCCCGCACG CTTCTTTCCA
2341 GCACCGTGAT GCGCGGAGCC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGCA TCCTTTGCAG
2461 ACCGTTGAGA GCCGACTTGT TGTGCGCCAC CCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACTTTT TTGGCGAAGG CATCGGCCTC GGCCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GCGCTGCGT 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 CTCTTCCTCT TCGTTTCAGT CACAACCCGC AA ACTCTAGA ATATCAATGC
2941 TGCTGCAGGC CTTCTGTTC CTGCTGGCCG GCTTCGCCG CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCCAACGG CTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCCGAACGA CACCGTCTGG GGGACGCCCT TGTTCTGGGG 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 GGCGGCTACA CCTTCACCGA GTACCAGAAG AACCCCGTGC
3421 TGGCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCAGCA
3481 AGTGGATCAT GACCGCGGCC AAGTCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCGCT TCGCCAACGA GGGCTTCCTC TCGTACCAGT
3601 ACGAGTGCCC CGGCCTGATC GAGGTCCCCA CCGAGCAGGA CCCCAGCAAG TCCTACTGGG
3661 TGATGTTTCA CTCCATCAAC CCCGGCGCCC CGGCCGGCGG CTCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CTTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TACGGGAGCG
3841 CCCTGGGCAT CGCGTGGGCC TCCA ACTGGG AGTACTCCGC CTTCGTGCC ACCAACCCT
3901 GCGCTCCTC CATGTCCCTC GTGCGCAAAG TCTCCCTCAA CACCGAGTAC CAGGCCAAC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCCG AGCCGATCCT GAACATCAGC AACCGCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCAGT TGACGAAGGC CAACAGCTAC AACCTCGACC
4081 TGTCCAACAG CACCGGCACC CTGGAGTTCT AGCTGGTGTG CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT GCGGACCTCT CCCTCTGGTT CAAGGGCCTG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCTC 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 CGCGAGGTGA 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 CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGCTATC
4801 CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGCTCCGC
4861 CTGTATTCTC CTGGTACTGC AACC TGTA AA CCAGCACTGC AATGCTGATG CACGGGAAGT
4921 AGTGGGATGG GAACACAAAT GGAGGATCCC GCGTCTCGAA CAGAGCGCGC AGAGGAACGC
4981 TGAAGGTCTC GCCTCTGTCT CACCTCAGCG CGGCATACAC CACAATAACC ACCTGACGAA
5041 TGCCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGGTTCA CACACGTGCC ACGTTGGCGA
5101 GGTGGCAGGT GACAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 CGAATTCGGC CGACAGGACG CGCGTCAAAG GTGCTGGTCG TGTATGCCCT GGCCGGCAGG
5221 TCGTTGCTGC TGCTGGTTAG TGATTCCGCA ACCCTGATTT TGGCGTCTTA TTTTGGCGTG
5281 GCAAACGCTG GCGCCCGCGA GCCGGGCCGG CGGCGATGCG GTGCCCCACG GCTGCCGGAA
5341 TCCAAGGGAG GCAAGAGCGC CCGGTCAGT TGAAGGGCTT TACGCGCAAG GTACAGCCGC
5401 TCCTGCAAGG CTGCGTGGTG GAATTGGACG TGCAGGTCCT GCTGAAGTTC CTCCACCGCC
5461 TCACCAGCGG ACAAAGCACC GGTGTATCAG GTCCGTGTCA TCCACTCTAA AGAACTCGAC
5521 TACGACCTAC TGATGGCCCT AGATTCTTCA TCAAAAACGC CTGAGACACT TGCCCAGGAT
5581 TGAAACTCCC TGAAGGGACC ACCAGGGGCC CTGAGTTGTT CCTCCCCCCT GTGGCGAGCT
5641 GCCAGCCAGG CTGTACCTGT GATCGAGGCT GGCGGGAAAA 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 GCAAAAATTCT GGCTCTGTCT
6001 CCAACCCTAG GATCAGCGGC GTAGGATTTT GTAATCATTC GTCTGATGG GGAGCTACCG
6061 ACTACCCTAA TATCAGCCCG ACTGCCTGAC GCCAGCGTCC ACTTTTGTGC ACACATTCCA
6121 TTCGTGCCAA AGACATTTCA TTGTGGTGCG AAGCGTCCCC AGTTACGCTC ACCTGTTTCC
6181 CGACCTCCCT ACTGTTCTGT CGACAGAGCG GGCCACAGG CCGGTCGCGC CCACTAGTAT
6241 GGCCACCGCA TCCACTTTCT CGGCGTTCAA TGCCCGCTGC GGCGACCTGC GTCGCTCGGC
6301 GGGCTCCGGG CCCC GGCGCC CAGCGAGGCC CCTCCCCGTG CGCGGGCGCG CCGCCACCGG
6361 CGAGCAGCCC TCCGGCGTGG CCTCCCTGCG CGAGGCCGAC AAGGAGAAGT CCCTGGGCAA
6421 CCGCTGCGC CTGGGCTCCC TGACCGAGGA CGGCCTGTCC TACAAGGAGA AGTTCGTGAT
6481 CCGCTGCTAC GAGGTGGGCA TCAACAAGAC CGCCACCATC GAGACCATCG CCAACCTGCT
6541 GCAGGAGGTG GGCGGCAACC ACGCCCAGGG CGTGGGCTTC TCCACCGACG GCTTCGCCAC
6601 CACCACCACC ATGCGCAAGC TGCACCTGAT CTGGGTGACC GCCCGCATGC ACATCGAGAT
6661 CTACCCTAC CCGCCTGGT CCGACGTGAT CGAGATCGAG ACCTGGGTGC AGGGCGAGGG
6721 CAAGGTGGGC ACCCGCCGCG ACTGATCCT GAAGGACTAC GCCAACGGCG AGGTGATCGG
6781 CCGCGCCACC TCCAAGTGGG TGATGATGAA CGAGGACACC CGCCGCTGC AGAAGGTGTC
6841 CGACGACGTG CGCGAGGAGT ACCTGGTGTT CTGCCCCCGC ACCCTGCGCC TGGCCTTCCC
6901 CGAGGAGAAC AACAAC TCCA TGAAGAAGAT CCCC AAGCTG GAGGACCCCG CCGAGTACTC
6961 CCGCTGGGCT CTGGTGCCCC GCGCTCCGA CCTGGACATG AACAAGCACG TGAACAACGT
7021 GACCTACATC GGCTGGGCCC TGAGTCCAT CCCCCCGGAG ATCATCGACA CCCACGAGCT
7081 GCAGGCCATC ACCCTGGACT ACCGCCGCGA GTGCCAGCGC GACGACATCG TGGACTCCCT
7141 GACCTCCCGC GAGCCCCCTG GCAACGCCCG CTGCGTGAAG TTCAAGGAGA TCAACGGCTC
7201 CGTGTCCCCC AAGAAGGACG AGCAGGACCT GTCCCGCTTC ATGCACCTGC TGCCTCCGC
7261 CGGCTCCGGC CTGGAGATCA ACCGCTGCCG CACCGAGTGG CGCAAGAAGC CCGCAAGCG
7321 CATGGACTAC AAGGACCACG ACGCGACTA CAAGGACCAC GACATCGACT ACAAGGACGA
7381 CGACGACAAG TGAATCGATA GATCTCTTAA GGCAGCAGCA GCTCGGATAG TATCGACACA
7441 CTCTGGACGC TGGTCGTGTG ATGGACTGTT GCCGCCACAC TTGCTGCCTT GACCTGTGAA
7501 TATCCCTGCC GCTTTTATCA AACAGCCTCA GTGTGTTTGA TCTTGTGTGT ACGCGCTTTT
7561 GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC CCAGCATCCC CTTCCCTCGT
7621 TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGCTATC CCTCAGCGCT
7681 GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGCTCCGC CTGTATTCTC
7741 CTGGTACTGC AACCTGTAAA CCAGCACTGC AATGCTGATG CACGGGAAGT AGTGGGATGG
7801 GAACACAAAT GGAAAGCTTA ATTAAGAGCT CTTGTTTTCC AGAAGGAGTT GCTCCTTGAG
7861 CCTTTCATTC TCAGCCTCGA TAACCTCCAA AGCCGCTCTA ATTGTGGAGG GGGTTCGAAT
7921 TTAAGGCTT GGAATGTTGG TTCGTGCGTC TGGAACAAGC CCAGACTTGT TGCTCACTGG
7981 GAAAAGGACC ATCAGCTCCA AAAA ACTTGC CGCTCAAACC GCGTACCTCT GCTTTCGCGC

8041	AATCTGCCCT	GTTGAAATCG	CCACCACATT	CATATTGTGA	CGCTTGAGCA	GTCTGTAATT
8101	GCCTCAGAA	GTGGAATCAT	CTGCCCCCTG	TGCGAGCCCA	TGCCAGGCAT	GTGCGGGGCG
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	CCAGCACAAC
8401	CTGCTGGCCC	AGGCGAGCGT	CAAACCATAC	CACACAAATA	TCCTTGGCAT	CGGCCCTGAA
8461	TTCTTTCTGC	CGCTCTGCTA	CCCGGTGCTT	CTGTCCGAAG	CAGGGGTTGC	TAGGGATCGC
8521	TCCGAGTCCG	CAAACCCTTG	TCGCGTGGCG	GGGCTTGTTT	GAGCTTGAAG	AGCCTCTAGA
8581	GTCGACCTGC	AGGCATGCAA	GCTTGGCGTA	ATCATGGTCA	TAGCTGTTTC	CTGTGTGAAA
8641	TTGTTATCCG	CTCACAATTC	CACACAACAT	ACGAGCCGGA	AGCATAAAGT	GTAAAGCCTG
8701	GGGTGCCTAA	TGAGTGAGCT	AACTCACATT	AATTGCGTTG	CGCTCACTGC	CCGCTTTCCA
8761	GTCGGGAAAC	CTGTCGTGCC	AGCTGCATTA	ATGAATCGGC	CAACGCGCGG	GGAGAGGCGG
8821	TTTGCGTATT	GGGCGCTCTT	CCGCTTCTCT	GCTCACTGAC	TCGCTGCGCT	CGGTCGTTTC
8881	GCTGCGGCGA	GCGGTATCAG	CTCACTCAAA	GGCGGTAATA	CGGTTATCCA	CAGAATCAGG
8941	GGATAACGCA	GGAAAGAACA	TGTGAGCAAA	AGGCCAGCAA	AAGGCCAGGA	ACCGTAAAAA
9001	GGCCGCGTTG	CTGGCGTTTT	TCCATAGGCT	CCGCCCCCCT	GACGAGCATC	ACAAAAATCG
9061	ACGCTCAAGT	CAGAGGTGGC	GAAACCCGAC	AGGACTATAA	AGATAACCAGG	CGTTTCCCCC
9121	TGGAAGCTCC	CTCGTGCCT	CTCCTGTTCC	GACCCTGCCG	CTTACCGGAT	ACCTGTCCGC
9181	CTTTCTCCCT	TCGGGAAGCG	TGGCGCTTTC	TCATAGCTCA	CGCTGTAGGT	ATCTCAGTTC
9241	GGTGTAGGTC	GTTTCGCTCCA	AGCTGGGCTG	TGTGCACGAA	CCCCCGTTT	AGCCCGACCG
9301	CTGCGCCTTA	TCCGGTAACT	ATCGTCTTGA	GTCCAACCCG	GTAAGACACG	ACTTATCGC

SEQ ID NO: 240
pSZ1452

1	CACTGGCAGC	AGCCACTGGT	AACAGGATTA	GCAGAGCGAG	GTATGTAGGC	GGTGCTACAG
61	AGTTCTTGAA	GTGGTGCCCT	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTTCGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAAACAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAAC
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAACCTTGG	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	TTATCCGCC	TCCATCCAGT
661	CTATTAAATT	TTGCCGGGAA	GCTAGAGTAA	GTAGTTTCGCC	AGTTAATAGT	TTGCGCAACG
721	TTGTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGGTATG	GCTTCATTCA
781	GCTCCGGTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC	CATGTTGTGC	AAAAAAGCGG
841	TTAGTCTCTT	CGGTCTCTCG	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAGTG	TTATCACTCA
901	TGGTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTCAATGC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTAATCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAA	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCGATGA	ACCCACTCGT	GCACCCAAC	GATCTTCAGC	ATCTTTTACT	TTACCCAGCG
1201	TTTCTGGGTG	AGCAAAAACA	GGAAGGCAAA	ATGCCGCAAA	AAAGGGAATA	AGGCGGACAC
1261	GGAAATGTTG	AATACTCATA	CTCTTCCTTT	TTCAATATTA	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	CAGTCCCCGG	AGACGGTCAC	AGCTTGTCTG	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGGCGCGT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	GGCATCAGAG	CAGATTGTAT	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	TTCCCGCCG	CCACTCCTGC	TCGAGCGCGC
1921	CCGCGCGTGC	GCCGCCAGCG	CCTTGCCCTT	TTCCCGCGCG	TCGTGCGCGT	CGCTGATGTC
1981	CATACCAGG	TCCATGAGGT	CTGCCTTGCG	CCGGCTGAGC	CAGTCTTCG	TCCGGGCGGC

2041 CAAGAGGAGC ATGAGGGAGG ACTCCTGGTC CAGGGTCCTG ACGTGGTCCG GGCTCTGGGA
2101 GCGGGCCAGC ATCATCTGGC TCTGCCGCAC CGAGGCCGCC TCCAACCTGGT CCTCCAGCAG
2161 CCGCAGTCGC CGCCGACCCT GGCAGAGGAA GACAGGTGAG GGGGGTATGA ATTGTACAGA
2221 ACAACCACGA GCCTTGTCCTA GGCAGAATCC CTACCAGTCA TGGCTTTACC TGGATGACGG
2281 CCTGCGAACA GCTGTCCAGC GACCCTCGCT GCCGCCGCTT CTCCCGCACG CTTCTTTCCA
2341 GCACCGTGAT GCGCGAGACC AGCGCCGCAC GCTGGCGCTG CGCTTCGCCG ATCTGAGGAC
2401 AGTCGGGGAA CTCTGATCAG TCTAAACCCC CTTGCGCGTT AGTGTGCCA TCCTTTGCAG
2461 ACCGGTGAGA GCCGACTTGT TGTGCGCCAC CCCCCACACC ACCTCCTCCC AGACCAATTC
2521 TGTCACCTTT TTGGCGAAGG CATCGGCCCT GGCCTGCAGA GAGGACAGCA GTGCCAGCC
2581 GCTGGGGGTT GCGGATGCA CGCTCAGGTA CCCTTTCTTG CGCTATGACA CTTCCAGCAA
2641 AAGGTAGGGC GGGCTGCGAG ACGGCTTCCC GGCGCTGCAT GCAACACCGA TGATGCTTCG
2701 ACCCCCCGAA GCTCCTTCGG GGCTGCATGG GCGCTCCGAT GCCGCTCCAG GGCAGCGCT
2761 GTTTAAATAG CCAGGCCCC GATTGCAAAAG ACATTATAGC GAGCTACCAA AGCCATATTC
2821 AAACACCTAG ATCACTACCA CTTCTACACA GGCCACTCGA GCTTGTGATC GCACTCCGCT
2881 AAGGGGGCGC CTCTTCCTCT TCGTTTCAGT CACAACCCGC AAACCTCTAGA ATATCAATGC
2941 TGCTGCAGGC CTTCTGTTC CTGCTGGCCG GCTTCGCCGC CAAGATCAGC GCCTCCATGA
3001 CGAACGAGAC GTCCGACCGC CCCCTGGTGC ACTTCACCCC CAACAAGGGC TGGATGAACG
3061 ACCCAACCGC CCTGTGGTAC GACGAGAAGG ACGCCAAGTG GCACCTGTAC TTCCAGTACA
3121 ACCGGAACGA CACCGTCTGG GGGACGCCCT TGTTCTGGGG CCACGCCACG TCCGACGACC
3181 TGACCAACTG GGAGGACCAG CCCATCGCCA TCGCCCCGAA GCGCAACGAC TCCGGCGCCT
3241 TCTCCGGCTC CATGGTGGTG GACTACAACA ACACCTCCGG CTTCTTCAAC GACACCATCG
3301 ACCCGGCCCA GCGCTGCGTG GCCATCTGGA CCTACAACAC CCCGGAGTCC GAGGAGCAGT
3361 ACATCTCCTA CAGCCTGGAC GCGGCTACA CTTTCACCGA GTACCAGAAG AACCCCGTGC
3421 TGGCCGCCAA CTCCACCCAG TTCCGCGACC CGAAGGTCTT CTGGTACGAG CCCTCCCAGA
3481 AGTGGATCAT GACCGCGGCC AAGTCCAGG ACTACAAGAT CGAGATCTAC TCCTCCGACG
3541 ACCTGAAGTC CTGGAAGCTG GAGTCCGCGT TCGCCAACGA GGGCTTCCTC GGCTACCAGT
3601 ACGAGTGCCC CGGCCTGATC GAGGTCCCCA CCGAGCAGGA CCCCAGCAAG TCCTACTGGG
3661 TGATGTTTAT CTCCATCAAC CCCGGCGCCC CGGCCGGCGG TCCTTCAAC CAGTACTTCG
3721 TCGGCAGCTT CAACGGCACC CACTTCGAGG CTTTCGACAA CCAGTCCCGC GTGGTGGACT
3781 TCGGCAAGGA CTACTACGCC CTGCAGACCT TCTTCAACAC CGACCCGACC TCGGAGCGC
3841 CCTGGGCAT CGCGTGGGCC TCCAACCTGG AGTACTCCGC CTTCGTGCC ACCAACCCT
3901 GCGCTCCTC CATGTCCCTC GTGCGCAAGT TCTCCCTCAA CACCGAGTAC CAGGCCAACC
3961 CGGAGACGGA GCTGATCAAC CTGAAGGCCG AGCCGATCCT GAACATCAGC AACGCCGGCC
4021 CCTGGAGCCG GTTCGCCACC AACACCACGT TGACGAAGGC CAACAGCTAC AACGTCGACC
4081 TGTCCACAG CACCGGCACC CTGGAGTTCG AGCTGGTGTA CGCCGTCAAC ACCACCCAGA
4141 CGATCTCCAA GTCCGTGTTT GCGGACCTCT CCCTCTGGTT CAAGGGCCTG GAGGACCCCG
4201 AGGAGTACCT CCGCATGGGC TTCGAGGTGT CCGCGTCTC CTTCTTCCTG GACCGCGGGA
4261 ACAGCAAGGT GAAGTTCGTG AAGGAGAACC CTTACTTCAC CAACCGCATG AGCGTGAACA
4321 ACCAGCCCTT CAAGAGCGAG AACGACCTGT CTTACTACAA GGTGTACGG TTTGTGGACC
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 GTGTGTTGA TCTTGTGTGT
4681 ACGCGCTTTT GCGAGTTGCT AGCTGCTTGT GCTATTTGCG AATACCACCC CCAGCATCCC
4741 CTTCCCTCGT TTCATATCGC TTGCATCCCA ACCGCAACTT ATCTACGCTG TCCTGCTATC
4801 CCTCAGCGCT GCTCCTGCTC CTGCTCACTG CCCCTCGCAC AGCCTTGGTT TGGGTCCTCG
4861 CTGTATTCTC CTGGTACTGC AACCTGTAAA CAGCAGCTGC AATGCTGATG TAGGGGAAGT
4921 AGTGGGATGG GAACACAAAT GGAGGATCCC GCGTCTCGAA CAGAGCGCGC AGAGGAACGC
4981 TGAAGGTCTC GCCTCTGTCT CACCTCAGCG CGGCATACAC CACAATAACC ACCTGACGAA
5041 TGCGCTTGGT TCTTCGTCCA TTAGCGAAGC GTCCGTTTCA CACACGTGCC ACGTTGGCGA
5101 GGTGGCAGGT GACAATGATC GGTGGAGCTG ATGGTCGAAA CGTTCACAGC CTAGGGATAT
5161 CGAATTCGGC CGACAGGACG CGCGTCAAAG GTGCTGGTCG TGTATGCCCT GGCCGGCAGG
5221 TCGTTGCTGC TGCTGGTTAG TGATTCCGCA ACCCTGATTT TGGCGTCTTA TTTTGGCGTG
5281 GCAAACGCTG GCGCCCGCGA GCCGGGCCGG CGGCGATGCG GTGCCCCACG GCTGCCGGAA
5341 TCCAAGGGAG GCAAGAGCGC CCGGGTCAGT TGAAGGGCTT TACGCGCAAG GTACAGCCGC
5401 TCCTGCAAGG CTGCGTGGTG GAATTGGACG TGCAGGTCCT GCTGAAGTTT CTCCACCGCC
5461 TCACCAGCGG ACAAAGCACC GGTGTATCAG GTCCGTGTCA TCCACTTAA 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 GTCCCAAAAT TCTGGTCTAC CGGGGGTGAT
5821 CCTTCGTGTA CGGGCCCTTC CCTCAACCCT AGGTATGCGC GCATGCGGTC GCCGCGCAAC
5881 TCGCGCGAGG GCCGAGGGTT TGGGACGGGC CGTCCCGAAA TGCAGTTGCA CCCGGATGCG
5941 TGGCACCTTT TTTGCGATAA TTTATGCAAT GGACTGCTCT GCAAAAATTCT GGCTCTGTCTG
6001 CCAACCCTAG GATCAGCGGC GTAGGATTTT GTAATCATTG GTCTGATGG GGAGCTACCG
6061 ACTACCTTAA TATCAGCCCG ACTGCCTGAC GCCAGCGTCC ACTTTTGTGC ACACATTCCA
6121 TTCGTGCCCA AGACATTTCA TTGTGGTGCG AAGCGTCCCC AGTTACGCTC ACCTGTTTCC
6181 CGACCTCCTT ACTGTTCTGT CGACAGAGCG GGCCACAGG CCGGTCGCAG CCACTAGTAT
6241 GGCCACCGCA TCCACTTTCT CGGCGTTCAA TGCCCGCTGC GGCGACCTGC GTCGCTCGGC
6301 GGGCTCCGGG CCCC GGCGCC CAGCGAGGCC CCTCCCCGTG CGCGGGCGCG CCGCCAACGC
6361 CCAGGCCCCC ACCAAGATCA ACGGCTCCAC CGACGACGCC CAGCTGCCCG CCCCTCGCAC
6421 CTTTCATCAAC CAGCTGCCCG ACTGGTCCAT GCTGCTGGCC GCCATCACCA CCGGTTCCT
6481 GGCCGCCGAG AAGCAGTGGA TGATGCTGGA CTGGAAGCCC CGCCGCCCG ACATGCTGAT
6541 CGACACCTTC GGCTGGGCC GCATCGTGCA GGACGGCCTG GTGTTCCGCC AGAATTCTC
6601 CATCCGCTCC TACGAGATCG GCGCCGACCG CACCGCCTCC ATCGAGACCG TGATGAACCA
6661 CCTGCAGGAG ACCGCCCTGA ACCACGTGAA GACCGCCGGC CTGCTGGGCG ACGGCTTCGG
6721 CTCCACCCCC GAGATGTCCA AGCGCAACCT GATCTGGGTG GTGACCAAGA TGCAGGTGGA
6781 GGTGGACCGC TACCCACCT GGGCGACGT GGTGCAGGTG GACACCTGGG TGTCCGCCCTC
6841 CGGCAAGAAC GGCATGCGCC GCGACTGGCT GCTGCGCGAC GGCAACACCG GCGAGACCT
6901 GACCCGCGCC TCCTCCGTGT GGGTATGAT GAACAAGCTG ACCCGCCGCC TGTCCAAGAT
6961 CCCC GAGGAG GTGCGCGAGG AGATCGGCTC CTACTTCGTG AACTCCGACC CCGTGGTGGA
7021 GGAGGACGGC CGCAAGGTGA CCAAGCTGGA CGACAACACC GCCGACTTCG TGC GCAAGGG
7081 CCTGACCCCC AAGTGGAACG ACCTGGACAT CAACCAGCAC GTGAACAACG TGAAGTACAT
7141 CGGCTGGATC CTGGAGTCCG CCCCCAGCC CATCCTGGAG ACCCGCGAGC TGTCCGCCGT
7201 GACCTGGAG TACCGCCGCG AGTGGGCGCG CGACTCCGTG CTGCGCTCCC TGACCGCCGT
7261 GTCCGCGGCG GCGTGGGCG ACCTGGGCCA CGCCGGCAAC GTGGAGTGCC AGCAGTGCT
7321 GCGCTGGAG GACGCGCCG AGATCGTGCG CGGCCGCACC GAGTGGCGCC CCAAGTACAT
7381 CAACAACCTC TCCATCATGG GCCAGATCCC CACCGACGCC TCCATGGACT ACAAGGACCA
7441 CGACGGCGAC TACAAGGACC ACGACATCGA CTACAAGGAC GACGACGACA AGTGAATCGA
7501 TAGATCTCTT AAGGCAGCAG CAGCTCGGAT AGTATCGACA CACTCTGGAC CTGGTCTGTG
7561 TGATGGACTG TTGCCGCCAC ACTTGCTGCC TTGACCTGTG AATATCCCTG CCGCTTTTAT
7621 CAAACAGCCT CAGTGTGTTT GATCTTGTTT GTACGCGCTT TTGCGAGTTG CTAGCTGCTT
7681 GTGCTATTTG CGAATACCAC CCCCAGCATC CCCTTCCCTC GTTTCATATC GCTTGCATCC
7741 CAACCGCAAC TTATCTACGC TGTCCTGCTA TCCCTCAGCG CTGCTCCTGC TCCTGCTCAC
7801 TGCCCTCGC ACAGCCTTGG TTTGGGCTCC GCCTGTATTC TCCTGGTACT GCAACCTGTA
7861 AACCAGCACT GCAATGCTGA TGCACGGGAA GTAGTGGGAT GGGAAACAAA ATGGAAGCT
7921 TAATTAAGAG CTCTTGTTTT CCAGAAGGAG TTGCTCCTTG AGCCTTTCAT TCTCAGCCTC
7981 GATAACCTCC AAAGCCGCTC TAATGTGGA GGGGTTTGA ATTTAAAGC TTGGAATGTT
8041 GGTTCGTGCG TCTGGAACAA GCCCAGACTT GTTGCTCACT GGGAAAAGGA CCATCGACTC
8101 CAAAAA ACTT GCCGCTCAAA CCGCGTACCT CTGCTTTTCG GCAATCTGCC CTGTTGAAAT
8161 CGCCACCACA TTCATATTGT GACGCTTGAG CAGTCTGTAA TTGCCTCAGA ATGTGGAATC
8221 ATCTGCCCCC TGTGCGAGCC CATGCCAGGC ATGTGCGGGG CGAGGACACC CGCCACTCGT
8281 ACAGCAGACC ATTATGCTAC CTCACAATAG TTCATAACAG TGACCATATT TCTCGAAGCT
8341 CCCCACGAG CACCTCCATG CTCTGAGTGG CCACCCCCCG GCCCTGGTGC TTGCGGAGGG
8401 CAGGTCAACC GGCATGGGGC TACCGAAATC CCCGACCGGA TCCCACCACC CCCGCGATGG
8461 GAAGAATCTC TCCCCGGGAT GTGGGCCACA CACCGACACA ACCTGCTGGC CCGGCGAGC
8521 GTCAAAACCA TATCCTTGGC ATCGGCCCTG AATTCCTTCT CCGGCTCTGC
8581 TACCCGTGTC TTCTGTCCGA AGCAGGGGTT GCTAGGGATC GCTCCGAGTC CGCAAACCT
8641 TGTCGCTGG CGGGGCTTGT TCGAGCTTGA AGAGCCTCTA GAGTCGACCT GCAGGCATGC
8701 AAGCTTGGCG TAATCATGGT CATAGCTGTT TCCTGTGTGA AATTGTTATC CGCTCACAAT
8761 TCCACACAAC ATACGAGCCG GAAGCATAAA GTGTAAAGCC TGGGGTGCCT AATGAGTGAG
8821 CTAACACACA TTAATTGCGT TGCGCTCACT GCCCGCTTTC CAGTCGGGAA ACCTGTCGTG
8881 CCAGTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTTGCGTA TTGGGCGCTC
8941 TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCTGT CGGCTGCGGC GAGCGGTATC
9001 AGCTCACTCA AAGGCGGTAA TACGTTTAT CACAGAATCA GGGGATAACG CAGGAAGAA
9061 CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGCTGGCGTT
9121 TTTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG
9181 GCGAAACCCG ACAGGACTAT AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG
9241 CTCTCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCCGGGAG
9301 CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTGCTC

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	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAACAAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	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	CGCAGAAAGT	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	TTAGTCCCTT	CGGTCCCTCC	ATCGTTGTCA	GAAGTAAGTT	GGCCGCAAGT	TTATCACTCA
901	TGGTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTTCATGCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA	AAAGTGCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAC	TCTCAAGGAT	CTTACCGCTG	TTGAGATCCA
1141	GTTTCGATGTA	ACCCACTCGT	GCACCCAAC	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	AGCTTGCTCT	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGCGCGCT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAATCATGC	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATACCGCA	TCAGGCGCCA	TTCGCCATTTC	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAGGGG
1801	GATGTGCTGC	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TTCCAGTGCA	CGAGCTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCCGCCTGG	AGCTGGTGCA	GAGCATGGGG
1921	CAGTTTGCGG	AGGAGAGGGT	GCTCCCCGTG	CTGCACCCCG	TGGACAAGCT	GTGGCAGCCG
1981	CAGGACTTCC	TGCCCGACCC	CGAGTCGCCC	GACTTCGAGG	ACCAGGTGGC	GGAGCTGCGC
2041	GCGCGCGCCA	AGGACCTGCC	CGACGAGTAC	TTTGTGGTGC	TGGTGGGCGA	CATGATCACG
2101	GAGGAGCGCG	TGCCGACCTA	CATGGCCATG	CTCAACACCT	TGGACGGTGT	GCGCGACGAC
2161	ACGGGCGCGG	CTGACCACCC	GTGGGCGCGC	TGGACGCGGC	AGTGGGTGGC	CGAGGAGAAC
2221	CGGCACGGCG	ACCTGCTGAA	CAAGTACTGT	TGGCTGACGG	GGCGCGTCAA	CATGCGGGCC
2281	GTGGAGGTGA	CCATCAACAA	CCTGATCAAG	AGCGGCATGA	ACCCGACAGC	GGACAACAAC
2341	CCTTACTTGG	GCTTCGTCTA	CACCTCCTTC	CAGGAGCGCG	CCACCAAGTA	TGTACCCCTT
2401	CTTGCGCTAT	GACACTTCCA	GCAAAAGGTA	GGGCGGGCTG	CGAGACGGCT	TCCCGGCGCT
2461	GCATGCAACA	CCGATGATGC	TTCGACCCCC	CGAAGCTCCT	TCGGGGCTGC	ATGGGCGCTC
2521	CGATGCCGCT	CCAGGGCGAG	CGCTGTTTAA	ATAGCCAGGC	CCCCGATTGC	AAAGACATTA
2581	TAGCGAGCTA	CCAAAGCCAT	ATTCAAACAC	CTAGATCACT	ACCACTTCTA	CACAGGCCAC
2641	TCGAGCTTGT	GATCGCACTC	CGCTAAGGGG	GCGCCTCTTC	CTCTTCGTTT	CAGTCACAAC
2701	CCGCAAACCTC	TAGAATATCA	ATGCTGCTGC	AGGCCTTCCT	GTTCTTGCTG	GCCGGCTTCG
2761	CCGCCAAGAT	CAGCGCCTCC	ATGACGAACG	AGACGTCCGA	CCGCCCCCTG	GTGCACTTCA
2821	CCCCAACAA	GGGCTGGATG	AACGACCCCA	ACGGCCTGTG	GTACGACGAG	AAGGACGCCA
2881	AGTGGACCTT	GTACTTCCAG	TACAACCCGA	ACGACACCGT	CTGGGGGACG	CCCTTGTCTT
2941	GGGGCGACGC	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	CCAACCTCCAC	CCAGTTCCGC	GACCCGAAGG

3241	TCTTCTGGTA	CGAGCCCTCC	CAGAAGTGGA	TCATGACCGC	GGCCAAGTCC	CAGGACTACA
3301	AGATCGAGAT	CTACTCCTCC	GACGACCTGA	AGTCCTGGAA	GCTGGAGTCC	GCGTTCGCCA
3361	ACGAGGGCTT	CCTCGGCTAC	CAGTACGAGT	GCCCCGGCCT	GATCGAGGTC	CCCACCGAGC
3421	AGGACCCCAG	CAAGTCCTAC	TGGGTGATGT	TCATCTCCAT	CAACCCCGGC	GCCCCGGCCG
3481	GCGGCTCCTT	CAACCAGTAC	TTCGTGCGCA	GCTTCAACGG	CACCCACTTC	GAGGCCCTTCG
3541	ACAACCAGTC	CCGCGTGGTG	GACTTCGGCA	AGGACTACTA	CGCCCTGCAG	ACCTTCTTCA
3601	ACACCGACCC	GACCTACGGG	AGCGCCCTGG	GCATCGCGTG	GGCCTCCAAC	TGGGAGTACT
3661	CGGCCTTCGT	GCCCACCAAC	CCCTGGCGCT	CCTCCATGTC	CCTCGTGCGC	AAGTTCTCCC
3721	TCAACACCGA	GTACCAGGCC	AACCCGGAGA	CGGAGCTGAT	CAACCTGAAG	GCCGAGCCGA
3781	TCCTGAACAT	CAGCAACGCC	GGCCCTTGGA	GCCGGTTCGC	CACCAACACC	ACGTTGACGA
3841	AGGCCAACAG	CTACAACGTC	GACCTGTCCA	ACAGCACCAG	CACCCTGGAG	TTCGAGCTGG
3901	TGTACGCCGT	CAACACCACC	CAGACGATCT	CCAAGTCCGT	GTTCGCGGAC	CTCTCCCTCT
3961	GGTTCAAGGG	CCTGGAGGAC	CCCAGGAGT	ACCTCCGCAT	GGGCTTCGAG	GTGTCCGCGT
4021	CCTCCTTCTT	CCTGGACCGC	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	TGTTGCCGCC	ACACTTGCTG	CCTTGACCTG	TGAATATCCC	TGCCGCTTTT	ATCAAACAGC
4441	CTCAGTGTGT	TTGATCTTGT	GTGTACGCGC	TTTTGCGAGT	TGCTAGCTGC	TTGTGCTATT
4501	TGCGAATACC	ACCCCCAGCA	TCCCTTTCCC	TCGTTTTATA	TCGCTTGATC	CCCAACGCCA
4561	ACTTATCTAC	GCTGTCCTGC	TATCCCTCAG	CGCTGCTCCT	GCTCCTGCTC	ACTGCCCTTC
4621	GCACAGCCTT	GGTTTGCGCT	CCGCCTGTAT	TCTCCTGGTA	CTGCAACCTG	TAAACCAGCA
4681	CTGCAATGCT	GATGCACGGG	AAGTAGTGGG	ATGGGAACAC	AAATGGAGGA	TCCCGCGTCT
4741	CGAACAGAGC	GCGCAGAGGA	ACGCTGAAGG	TCTCGCCTCT	GTCGCACCTC	AGCGCGGCAT
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	CTGTGCTGG	TTAGTGATTG	CGCAACCCTG
5041	ATTTTGGCGT	CTTATTTTGG	CGTGGCAAAC	GCTGGCGCCC	GCGAGCCGGG	GCGCGCGCGA
5101	TGCGGTGCC	CACGGCTGCC	GGAATCCAAG	GGAGGCAAGA	GCGCCCGGGT	CAGTTGAAGG
5161	GCTTTACGCG	CAAGGTACAG	CCGCTCCTGC	AAGGCTGCGT	GGTGAATTG	GACGTGCAGG
5221	TCCTGCTGAA	GTTCTCTCCAC	CGCCTACCA	GCGGACAAAG	CACCGGTGTA	TCAGGTCCGT
5281	GTCATCCACT	CTAAAGAACT	CGACTACGAC	CTACTGATGG	CCCTAGATTG	TTCATCAAAA
5341	ACGCCTGAGA	CACCTGCCCA	GGATTGAAAC	TCCCTGAAGG	GACCACCAGG	GGCCCTGAGT
5401	TGTTCTTCC	CCCCGTGGCG	AGCTGCCAGC	CAGGCTGTAC	CTGTGATCGA	GGCTGGCGGG
5461	AAAATAGGCT	TCGTGTGCTC	AGGTCATGGG	AGGTGCAGGA	CAGCTCATGA	AACGCCAACA
5521	ATCGCACAA	TCATGTCAAG	CTAATCAGCT	ATTTCTCTTT	CACGAGCTGT	AATTGTCCCA
5581	AAATTCTGGT	CTACCGGGGG	TGATCCTTCG	TGTACGGGCC	CTTCCCTCAA	CCCTAGGTAT
5641	GCGCGCATGC	GGTCGCCGCG	CAACTCGCGC	GAGGGCCGAG	GGTTTGGGAC	GGGCCGTCCC
5701	GAAATGCAGT	TGCACCCGGA	TGCGTGGCAC	CTTTTTTGCG	ATAATTTATG	CAATGGACTG
5761	CTCTGCAAAA	TTCTGGCTCT	GTCGCCAACC	CTAGGATCAG	CGGCGTAGGA	TTTCGTAATC
5821	ATTCTGCTCG	ATGGGGAGCT	ACCGACTACC	CTAATATCAG	CCCAGCTGCC	TGACGCCAGC
5881	GTCCACTTTT	GTGCACACAT	TCCATTCTGT	CCCAAGACAT	TTCATTGTGG	TGCGAAGCGT
5941	CCCCAGTTAC	GCTCACCTGT	TTCCCGACCT	CCTTACTGTT	CTGTGACAG	AGCGGGCCCCA
6001	CAGGCCGGTC	GCAGCCACTA	GTATGGCCAC	CGCATCCACT	TTCTCGGCGT	TCAATGCCCG
6061	CTGCGGCGAC	CTGCGTCGCT	CGGCGGGCTC	CGGGCCCCGG	CGCCACAGCA	CGCCCTCCCC
6121	CGTGCGCGGG	CGCGCCCCCG	ACTGGTCCAT	GCTGTTTCGCC	GTGATCACCA	CCATCTTCAG
6181	CGCCGCCGAG	AAGCAGTGGA	CCAACCTGGA	GTGGAAGCCC	AAGCCCAAGC	TGCCCCAGCT
6241	GCTGGACGAC	CACCTTCGGC	TGCACGGCCT	GGTGTTCGCG	CGCACCTTCG	CCATCCGCTC
6301	CTACGAGGTG	GGCCCCGACC	GCAGCACCTC	CATCCTGGCC	GTGATGAACC	ACATGCAGGA
6361	GGCCACCCTG	AACCACGCCA	AGAGCGTGGG	CATCCTGGGC	GACGGCTTCG	GCACCACCCT
6421	GGAGATGTCC	AAGCGCGACC	TGATGTGGGT	GGTGCGCCGC	ACCCACGTGG	CCGTGGAGCG
6481	CTACCCCACC	TGGGGCGACA	CCGTGGAGGT	GGAGTGCTGG	ATCGGCGCCA	GCGGCAACAA
6541	CGGCATGCGC	CGCGACTTCC	TGGTGCGCGA	CTGCAAGACC	GGCGAGATCC	TGACCCGCTG
6601	CACCTCCCTG	AGCGTGCTGA	TGAACACCCG	CACCCGCCGC	CTGAGACCCA	TCCCCGACGA
6661	GGTGCGCGGC	GAGATCGGCC	CCGCCTTCAT	CGACAACGTG	GCCGTGAAGG	ACGACGAGAT
6721	CAAGAAGCTG	CAGAAGCTGA	ACGACTCCAC	CGCCGACTAC	ATCCAGGGCG	GCCTGACCCC
6781	CCGCTGGAAC	GACCTGGACG	TGAACCAGCA	CGTGAACAAC	CTGAAGTACG	TGGCCTGGGT
6841	GTTTCGAGACC	GTGCCCCACA	GCATCTTCGA	GTCCCACCAC	ATCAGCTCCT	TCACCCTGGA

6901	GTACCGCCGC	GAGTGCACCC	GCGACTCCGT	GCTGCGCAGC	CTGACCACCG	TGAGCGGCGG
6961	CAGTCCGAG	GCCGGCCTGG	TGTGCGACCA	CCTGCTGCAG	CTGGAGGGCG	GCAGCGAGGT
7021	GCTGCGCGCC	CGCACCGAGT	GGCGCCCCAA	GCTGACCGAC	TCCTTCCGCG	GCATCAGCGT
7081	GATCCCCGCC	GAGCCCCGCG	TGATGGACTA	CAAGGACCAC	GACGGCGACT	ACAAGGACCA
7141	CGACATCGAC	TACAAGGACG	ACGACGACAA	GTGACTCGAG	GCAGCAGCAG	CTCGGATAGT
7201	ATCGACACAC	TCTGGACGCT	GGTCGTGTGA	TGGACTGTTG	CCGCCACACT	TGCTGCCTTG
7261	ACCTGTGAAT	ATCCCTGCCG	CTTTTATCAA	ACAGCCTCAG	TGTGTTTGAT	CTTGTGTGTA
7321	CGCGCTTTTG	CGAGTTGCTA	GCTGCTTG TG	CTATTTGCGA	ATACCACCCC	CAGCATCCCC
7381	TTCCCTCGTT	TCATATCGCT	TGCATCCCCA	CCGCAACTTA	TCTACGCTGT	CCTGCTATCC
7441	CTCAGCGCTG	CTCCTGCTCC	TGCTCACTGC	CCCTCGCACA	GCCTTGGTTT	GGGCTCCGCC
7501	TGTATTCTCC	TGGTACTGCA	ACCTGTAAAC	CAGCACTGCA	ATGCTGATGC	ACGGGAAGTA
7561	GTGGGATGGG	AACACAAATG	GAAAGCTTGA	GCTCCAGCCA	CGGCAACACC	CGCGCCTTGG
7621	CGGCCGAGCA	CGGCGACAAG	AACCTGAGCA	AGATCTGCGG	GCTGATCGCC	AGCGACGAGG
7681	GCCGGCACGA	GATCGCCTAC	ACGCGCATCG	TGGACGAGTT	CTTCCGCCTC	GACCCCGAGG
7741	GCGCCGTCGC	CGCCTACGCC	AACATGATGC	GCAAGCAGAT	CACCATGCCC	GCGCACCTCA
7801	TGGACGACAT	GGGCCACGGC	GAGGCCAACC	CGGGCCGCAA	CCTCTTCGCC	GACTTCTCCG
7861	CGGTCGCCGA	GAAGATCGAC	GTCTACGACG	CCGAGGACTA	CTGCCGCATC	CTGGAGCACC
7921	TCAACGCGCG	CTGGAAGGTG	GACGAGCGCC	AGGTCAAGCG	CCAGGCCGCC	GCGGACCAGG
7981	AGTACGTCTT	GGGCCTGCCC	CAGCGCTTCC	GGAAACTCGC	CGAGAAGACC	GCCGCCAAGC
8041	GCAAGCGCGT	CGCGCGCAGG	CCCGTCGCCT	TCTCCTGGAG	AAGAGCCTCT	AGAGTCGACC
8101	TGCAGGCATG	CAAGCTTGCG	GTAATCATGG	TCATAGCTGT	TTCTGTGTGT	AAATTGTTAT
8161	CCGCTCACAA	TTCCACACAA	CATACGAGCC	GGAAGCATAA	AGTGTAAGAG	CTGGGGTGCC
8221	TAATGAGTGA	GCTAACTCAC	ATTAATTGCG	TTGCGCTCAC	TGCCCCGCTT	CCAGTCGGGA
8281	AACCTGTCGT	GCCAGCTGCA	TTAATGAATC	GGCCAACGCG	CGGGGAGAGG	CGGTTTGCGT
8341	ATTGGGCGCT	CTTCCGCTTC	CTCGCTCACT	GACTCGCTGC	GCTCGGTCGT	TCGGCTGCGG
8401	CGAGCGGTAT	CAGCTCACTC	AAAGCGGGTA	ATACGGTTAT	CCACAGAATC	AGGGGATAAC
8461	GCAGGAAAGA	ACATGTGAGC	AAAAGGCCAG	CAAAAGGCCA	GGAACCGTAA	AAAGGCCGCG
8521	TTGCTGGCGT	TTTTCCATAG	GCTCCGCCCC	CCTGACGAGC	ATCACAAAAA	TCGACGCTCA
8581	AGTCAGAGGT	GGCGAAACCC	GACAGGACTA	TAAAGATACC	AGGCGTTTCC	CCCTGGAAGC
8641	TCCCTCGTGC	GCTCTCCTGT	TCCGACCCTG	CCGCTTACCG	GATACCTGTC	CGCCTTTCTC
8701	CCTTCGGGAA	GCGTGGCGCT	TTCTCATAGC	TCACGCTGTA	GGTATCTCAG	TTCGGTGTAG
8761	GTCGTTTCGT	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	AGTTCTTGAA	GTGGTGCCCT	AACTACGGCT	ACACTAGAAG	GACAGTATTT	GGTATCTGCG
121	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC	GGCAAAACAA
181	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA	GATTACGCGC	AGAAAAAAAG
241	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT
301	CACGTTAAGG	GATTTTGGTC	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA
361	ATTAAAAATG	AAGTTTTTAA	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT
421	ACCAATGCTT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTCTG	TCATCCATAG
481	TTGCTGACT	CCCCGTCTGT	TAGATACTA	CGATACGGGA	GGGCTTACCA	TGAGGCCCA
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	TTAGTCCCTT	CGGTCCCTCC	ATCGTTGTCA	GAAGTAAGTT	GGCCGCGAGT	TTATCACTCA
901	TGTTTATGGC	AGCACTGCAT	AATTCTCTTA	CTGTATGCCC	ATCCGTAAGA	TGCTTTTCTG
961	TGACTGGTGA	GTAATCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCCG	CCGAGTTGCT
1021	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAAGTTTA	TAAGGCTCTCA
1081	TCATTGGAAA	ACGTTCTTCG	GGGCGAAAAA	TCTCAAGGAT	CTTACCGCTG	TTAGATCCCA
1141	GTTTCGATGTA	ACCCACTCGT	GCACCCAACT	GATCTTCAGC	ATCTTTTACT	TTCACCAGCG
1201	TTTCTGGGTG	AGCAAAAAACA	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	GTGAAAACTT	CTGACACATG	CAGCTCCCGG	AGACGGTCAC	AGCTTGCTGT	TAAGCGGATG
1561	CCGGGAGCAG	ACAAGCCCGT	CAGGCGCGGT	CAGCGGGTGT	TGGCGGGTGT	CGGGGCTGGC
1621	TTAACTATGC	GGCATCAGAG	CAGATTGTAC	TGAGAGTGCA	CCATATGCGG	TGTGAAATAC
1681	CGCACAGATG	CGTAAGGAGA	AAATAACGCA	TCAGGCGCCA	TTCGCCATTG	AGGCTGCGCA
1741	ACTGTTGGGA	AGGGCGATCG	GTGCGGGCCT	CTTCGCTATT	ACGCCAGCTG	GCGAAAGGGG
1801	GATGTGCTGC	AAGGCGATTA	AGTTGGGTAA	CGCCAGGGTT	TTCCAGTCA	CGACGTTGTA
1861	AAACGACGGC	CAGTGAATTG	ATGCATGCTC	TTCCGCCTGG	AGCTGGTGCA	GAGCATGGGG
1921	CAGTTTGCGG	AGGAGAGGGT	GCTCCCCGTG	CTGCACCCCG	TGGACAAGCT	GTGGCAGCCG
1981	CAGGACTTCC	TGCCCCGACC	CGAGTCGCCC	GACTTCGAGG	ACCAGGTGGC	GGAGCTGCGC
2041	GCGCGCGCCA	AGGACCTGCC	CGACGAGTAC	TTTGTGGTGC	TGGTGGGCGA	CATGATCACG
2101	GAGGAGGCGC	TGCCGACCTA	CATGGCCATG	CTCAACACCT	TGGACGGTGT	GCGCGACGAC
2161	ACGGGCGCGG	CTGACCACCC	GTGGGCGCGC	TGGACGCGGC	AGTGGGTGGC	CGAGGAGAAC
2221	CGGCACGGCG	ACCTGCTGAA	CAAGTACTGT	TGGCTGACGG	GGCGCGTCAA	CATGCGGGCC
2281	GTGGAGGTGA	CCATCAACAA	CCTGATCAAG	AGCGGCATGA	ACCCGCGAGC	GGACAACAAC
2341	CCTTACTTGG	GCTTCGTCTA	CACCTCCTTC	CAGGAGCGCG	CCACCAAGTA	GGTACCCTTT
2401	CTTGCGCTAT	GACACTTCCA	GCAAAAGGTA	GGGCGGGCTG	CGAGACGGCT	TCCCGGCGCT
2461	GCATGCAACA	CCGATGATGC	TTCGACCCCC	CGAAGCTCCT	TCGGGGCTGC	ATGGGCGCTC
2521	CGATGCCGCT	CCAGGGCGAG	CGCTGTTTAA	ATAGCCAGGC	CCCCGATTGC	AAAGACATTA
2581	TAGCGAGCTA	CCAAAGCCAT	ATTCAAACAC	CTAGATCACT	ACCACTTCTA	CACAGGCCAC
2641	TCGAGCTTGT	GATCGCACTC	CGCTAAGGGG	GCGCCTCTTC	CTCTTCGTTT	CAGTCACAAC
2701	CCGCAAACTC	TAGAATATCA	ATGCTGCTGC	AGGCCTTCCT	GTTCTTGCTG	GCCGGCTTCG
2761	CCGCCAAGAT	CAGCGCCTCC	ATGACGAACG	AGACGTCCGA	CCGCCCCCTG	GTGCACTTCA
2821	CCCCCAACAA	GGGCTGGATG	AACGACCCCA	ACGGCCTGTG	GTACGACGAG	AAGGACGCCA
2881	AGTGGCACCT	GTACTTCCAG	TACAACCCGA	ACGACACCGT	CTGGGGGACG	CCCTTGTTCT
2941	GGGGCCACGC	CACGTCCGAC	GACCTGACCA	ACTGGGAGGA	CCAGCCCATC	GCCATCGCCC
3001	CGAAGCGCAA	CGACTCCGGC	GCCTTCTCCG	GCTCCATGGT	GGTGGACTAC	AACAACACCT
3061	CCGGCTTCTT	CAACGACACC	ATCGACCCCG	GCCAGCGCTG	CGTGGCCATC	TGGACCTACA
3121	ACACCCCGGA	GTCCGAGGAG	CAGTACATCT	CCTACAGCCT	GGACGGCGGC	TACACCTTCA
3181	CCAGGTACCA	GAAGAACCCC	GTGCTGGCCG	CCAACCTCCAC	CCAGTTCCGC	GACCCGAAGG
3241	TCTTCTGGTA	CGAGCCCTCC	CAGAAGTGGA	TCATGACCGC	GGCCAAGTCC	CAGGACTACA
3301	AGATCGAGAT	CTACTCCTCC	GACGACCTGA	AGTCTTGGA	GCTGGAGTCC	GCGTTCGCCA
3361	ACGAGGGCTT	CCTCGGCTAC	CAGTACGAGT	GCCCCGGCCT	GATCGAGGTC	CCCACCGAGC
3421	AGGACCCAG	CAAGTCCTAC	TGGGTGATGT	TCATCTCCAT	CAACCCCGGC	GCCCCGCGCG
3481	GCGGCTCCTT	CAACCAGTAC	TTCTGCGGCA	GCTTCAACGG	CACCCACTTC	GAGGCCTTCG
3541	ACAACCAGTC	CCGCGTGGTG	GACTTCGGCA	AGGACTACTA	CGCCCTGCAG	ACCTTCTTCA
3601	ACACCGACCC	GACCTACGGG	AGCGCCCTGG	GCATCGCGTG	GGCCTCCAAC	TGGGAGTACT
3661	CGCCTTCGT	GCCCACCAAC	CCCTGGCGCT	CCTCCATGTC	CCTCGTGC	AAGTTCTCCC
3721	TCAACACCGA	GTACCAGGCC	AACCCGGGAG	CGGAGCTGAT	CAACCTGAAG	GCCGACCCGA
3781	TCCTGAACAT	CAGCAACGCC	GGCCCTTGGA	GCCGGTTTCG	CACCAACACC	ACGTTGACGA
3841	AGGCCAACAG	CTACAACGTC	GACCTGTCCA	ACAGCACCGG	CACCCTGGAG	TTCGAGCTGG
3901	TGTACGCCGT	CAACACCACC	CAGACGATCT	CCAAGTCCGT	GTTGCGGGAC	CTCTCCCTCT
3961	GGTTCAAGGG	CCTGGAGGAC	CCCAGGAGT	ACCTCCGCAT	GGGCTTCGAG	GTGTCCGCGT
4021	CCTCCTTCTT	CCTGGACCGC	GGGAACAGCA	AGGTGAAGTT	CGTGAAGGAG	AACCCCTACT
4081	TCACCAACCG	CATGAGCGTG	AACAACCAGC	CCTTCAAGAG	CGAGAACGAC	CTGTCTTACT
4141	ACAAGGTGTA	CGGCTTGCTG	GACCAGAACA	TCCTGGAGCT	GTACTTCAAC	GATGGCGACG
4201	TCGTGTCCAC	CAACACCTAC	TTCATGACCA	CCGGGAACGC	CCTGGGCTCC	GACGAACATGA
4261	CGACGGGGGT	GGACAACCTG	TTCTACATCG	ACAAGTTCCA	GGTGCGCGAG	GTCAAGTGAC
4321	AATTGGCAGC	AGCAGCTCGG	ATAGTATCGA	CACACTCTGG	ACGCTGGTCG	TGTGATGGAC
4381	TGTTGCCGCC	ACACTTGCTG	CCTTGACCTG	TGAATATCCC	TGCCGCTTTT	ATCAAACAGC
4441	CTCAGTGTGT	TTGATCTTGT	GTGTACGCGC	TTTTGCGAGT	TGCTAGCTGC	TTGTGCTATT
4501	TGCGAATACC	ACCCCCAGCA	TCCCCTTCCC	TCGTTTCATA	TCGCTTGCAT	CCCAACCGCA
4561	ACTTATCTAC	GCTGTCCTGC	TATCCCTCAG	CGCTGCTCCT	GCTCCTGCTC	ACTGCCCTC
4621	GCACAGCCTT	GGTTTGGGCT	CCGCCTGTAT	TCTCCTGGTA	CTGCAACCTG	TAAACCAGCA
4681	CTGCAATGCT	GATGCACGGG	AAGTAGTGAG	ATGGGAACAC	AAATGGAGGA	TCCCGCTCT
4741	CGACAGAGAG	ACGCTGAAGG	TCTCGCCTCT	TCTCGCCTCT	GTCGCACCTC	GTCGCGGCAT
4801	ACACCACAAT	AACCACCTGA	CGAATGCGCT	TGGTTCTTCG	TCCATTAGCG	AAGCGTCCGG
4861	TTCACACACG	TGCCACGTTG	GCGAGGTGGC	AGGTGACAAT	GATCGGTGGA	GCTGATGGTC
4921	GAAACGTTCA	CAGCCTAGGG	ATATCGAATT	CGGCCGACAG	GACGCGCGTC	AAAGGTGCTG
4981	GTCGTGTATG	CCCTGGCCGG	CAGGTGCTTG	CTGCTGCTGG	TTAGTGATTG	CGCAACCCTG

5041 ATTTTGGCGT CTTATTTTGG CGTGGCAAAC GCTGGCGCCC GCGAGCCGGG CCGGCGGCGA
5101 TGCGGTGCC CACGGCTGCC GGAATCCAAG GGAGGCAAGA GCGCCCGGT CAGTTGAAGG
5161 GCTTTACGCG CAAGGTACAG CCGCTCCTGC AAGGCTGCGT GGTGGAATTG GACGTGCAGG
5221 TCCTGCTGAA GTTCTCCAC CGCCTCACC GCGGACAAAG CACCGGTGTA TCAGGTCCGT
5281 GTCATCCACT CTAAAGAACT CGACTACGAC CTACTGATGG CCCTAGATTG TTCATCAAAA
5341 ACGCCTGAGA CACTTGCCCA GGATTGAAAC TCCCTGAAGG GACCACCAGG GGCCCTGAGT
5401 TGTTCTTCC CCCCCTGGCG AGCTGCCAGC CAGGCTGTAC CTGTGATCGA GGCTGGCGGG
5461 AAAATAGGCT TCGTGTGCTC AGGTCATGGG AGGTGCAGGA CAGCTCATGA AACGCCAACA
5521 ATCGCACAAT TCATGTCAAG CTAATCAGCT ATTTCTCTT CACGAGCTGT AATTGTCCCA
5581 AAATTCTGGT CTACCGGGGG TGATCCTTCG TGTACGGGCC CTTCCCTCAA CCCTAGGTAT
5641 GCGCGCATGC GGTGCGCGCG CAACTCGCGC GAGGGCCGAG GGTTTGGGAC GGGCCGTCCC
5701 GAAATGCAGT TGCACCCGGA TCGTGGCACC CTTTTTTCG ATAATTTATG CATTGGACTG
5761 CTCTGCAAAA TTCTGGCTCT GTCGCCAACC CTAGGATCAG CGGCGTAGGA TTTCGTAATC
5821 ATTCGTCTTG ATGGGGAGCT ACCGACTACC CTAATATCAG CCCGACTGCC TGACGCCAGC
5881 GTCCACTTTT GTGCACACAT TCCATTCGTG CCCAAGACAT TTCATTGTGG TGCGAAGCGT
5941 CCCAGTTAC GCTCACCTGT TTCCCGACCT CTTACTGTT CTGTGACAG AGCGGGCCCA
6001 CAGGCCGCTC GCAGCCACTA GTATGGCCAC CGCATCCACT TTCTCGGCGT TCAATGCCCG
6061 CTGCGCGGAC CTGCGTCGCT CGGCGGGCTC CGGGCCCCGG CGCCAGCGA GGCCCTCCC
6121 CGTGC CGGG CGCGCCAGC TGCCCGACTG GAGCATGCTG CTGGCCGCGA TCACCACCT
6181 GTTCTGGCG GCCGAGAAGC AGTGGATGAT GCTGGACTGG AAGCCCAAGC GCCCCGACAT
6241 GCTGGTGGAC CCCTTCGGCC TGGGCGCTT GTGCGAGGAC GGCTTGGTGT GGCCTGACAA
6301 CTTCAAGCAT CGCAGCTACG AGATCGGCGC GGACCGCACC GCGCATCG AGACCTGAT
6361 GAACCACCTG CAGGAGACCG CCCTGAACCA CGTGAAGAGT GTGGGCTGCG TGAGGACGG
6421 CCTGGGCAGC ACCCGCGAGA TGAGCTGCG CAACCTGATC TGGGTGGTGA CCAAGATGCA
6481 GGTGGCGGTG GACCGCTACC CCACCTGGGG CGACGAGGTG CAGGTGAGCA GCTGGGCGAC
6541 CGCCATCGGC AAGAACGGCA TGCGCCGCGA GTGGATCGTG ACCGACTTCC GCACCGGCGA
6601 GACCTGCTG CGCGCCACCA GCGTGTGGGT GATGATGAAC AAGCTGACCC GCCGCATCAG
6661 CAAGATCCCC GAGGAGGTGT GGCACGAGAT CGGCCCCAGC TTCATCGACG CGCCCCCCT
6721 GCCCACCCTG GAGGACGACG GCCGCAAGCT GACCCGCTT GACGAGAGCA GCGCCGACTT
6781 CATCCGCAAG GCCTTGACCC CCCGCTGGAG CGACCTGGAC ATCAACGACG ACGTGAACAA
6841 CGTGAAGTAC ATCGGCTGGC TGCTGGAGAG CGCGCCCCCG GAGATCCACG AGAGCCACGA
6901 GATCGCCAGC CTGACCCTGG AGTACCGCCG CGAGTGCAGC CGCGACAGCG TGCTGAACAG
6961 CGCCACCAAG GTGAGCGACA GCAGCCAGCT GGGCAAGAGC GCCGTGGAGT GCAACCACCT
7021 GGTGCGCTG CAGAACGGCG GCGAGATCGT GAAGGGCCGC ACCGTGTGGC GCCCCAAGCG
7081 CCCCCTGTAC AACGACGGCG CCGTGGTGA CGTGCCCGCC AAGACCAGCT GACTCGAGGC
7141 AGCAGCAGCT CGGATAGTAT CGACACACTC TGGACGCTGG TCGTGTGATG GACTGTTGCC
7201 GCCCACTTG CTGCCTTGAC CTGTGAATAT CCCTGCCGCT TTTATCAAA AGCCTCAGTG
7261 TGTTTGATCT TGTGTGTACG CGCTTTTGG AGTTGCTAGC TGCTGTGTCT ATTTGCGAAT
7321 ACCACCCCA GCATCCCCTT CCCTCGTTT ATATCGCTTG CATCCCAACC GCAACTTATC
7381 TACGCTGTCC TGCTATCCCT CAGCGCTGCT CCTGCTCCTG CTCACGCCC CTCGCACAGC
7441 CTTGTTTGG GCTCCGCTG TATTCTCTG GTACTGCAAC CTGTAAACCA GCACTGCAAT
7501 GCTGATGCAC GGAAGTAGT GGGATGGGAA CACAAATGGA AAGCTTGAGC TCCAGCCACG
7561 GCAACACCGC GCGCCTTGGC GCCGAGCAG GCGACAAGAA CCTGAGCAAG ATCTGCGGGC
7621 TGATCGCCAG CGACGAGGGC CGGCACGAGA TCGCCTACAC GCGCATCGTG GACGAGTTCT
7681 TCCGCTCGA CCCCAGGGG GCCGTGCGCG CCTACGCCAA CATGATGCGC AAGCAGATCA
7741 CCATGCCCGC GCACCTCATG GACGACATGG GCCACGGCGA GGCCAACCCG GGCCGCAACC
7801 TCTTCGCCGA CTTCTCCGCG GTCGCCGAGA AGATCGACGT CTACGACGCC GAGGACTACT
7861 GCCGCATCCT GGAGCACCTC AACGCGCGT GGAAGGTGGA CGAGCCGAG GTCAGCGCC
7921 AGGCCGCCG GGACCAGGAG TACGTCTTGG GCCTGCCCA GCGCTTCCG AAACTCGCCG
7981 AGAAGACCGC CGCCAAGCGC AAGCGCGTCG CGCGCAGGCC CGTCGCTTTC TCCTGGAGAA
8041 GAGCCTCTAG AGTCGACCTG CAGGCATGCA AGCTTGGCGT AATCATGGTC ATAGCTGTTT
8101 CCTGTGTGAA ATTGTTATCC GTCACAATT CCACACAACA TACGAGCCGG AAGCATAAAG
8161 TGTAAAGCCT GGGGTGCCTA ATGAGTGAGC TAACTCACAT TAATTGCGTT GCGCTCACTG
8221 CCCGCTTCC AGTCGGGAAA CCTGTGCTG CAGCTGCATT AATGAATCGG CCAACGCGCG
8281 GGGAGAGGCG GTTTGCGTAT TGGGCGCTCT TCCGCTTCTT CGTCACTGA CTCGCTGCGC
8341 TCGGTCGTTT GGTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT AGGTTATCC
8401 ACAGAAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAGGCCAGG
8461 AACCGTAAAA AGGCGCGTTC GTGCGCTTT TTCCATAGGC TCCGCCCCC TGACGAGCAT
8521 CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAG
8581 GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTT CGACCCTGCC GCTTACCGGA
8641 TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCATAGCTC ACGCTGTAGG

[illegible]

SEQ ID NO: 245

[illegible]

SEQ ID NO: 244

[illegible]

cctcgtttcatatcgcttgcacccaaccgcaacttatctacgtgtcctgctatccctcagcgctgctcctgctcactgcccct
cgacacagccttggttgggctccgctgtattctctggtagctgaacctgtaaacagcactgcaatgctgatgcacgggaagtagt
gggatgggaacacaaatggaagcctgagctc

SEQ ID NO: 246

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

ATGgcgaccgccagctgcatggtggcgctgcccttctgcacctggctggcgccgctgcatgccacctccagcgacaacga
ccccgcctgctgtccacaagcgctgcgctgagccgccgcccgcacctgagctgcactgctcctgcgccgagcacc
ttccagtgcctggaccctgcaaccagcagcgcttctggcgacaacggcttcgctgctgttcggctccaagccccctgcga
gcaaccgcggccacctgcgctgggcccgcacctgcactccggcgaggtgatggcgtcgcatgcagcccgccaggaggt
gagcacaacaagaagccccgcaccaagcagcgccgctggctgacggcatgggctgctgacccccctgggccacga
ccccgacgtgtattataacaacctgctggacggcatctcgggcatctccgagatcgagaacttcgactgcagccagttccccc
ccgcatcgccggcgagatcaagtcgttctccaccgacggctgggtcgcgcccaagttcagcgagcgcatggacaagttcatgc
tgtatatgctgacggccggcaagaaggcgctggccgacggcgccatcaccgaggacgcatgaaggagctgaacaagcgc
aagtggcgctgctgatcggtcgggctgggcccgcaggaaggtcttctccgacagcatcgaggccctgcgcacctcgataa
gaagatctcccccttctgctgcccttcagcaccaccaaatgggctcgccgatcctggcgatggacctgggctggatgggccc
caactattccatcagcaccgctgcgccacctcgaacttctgcactcctgaacgcggccaaccacatcatcaagggcgaggcgg
acatgatgctgtgcccggctccgacgccggtgctgccgctggcggttcgtggcctgccgcgctgagccag
cgacaacagaccccccaaggcctcgccccctgggactccaaccgcgacggcttcgtcatggcgaggcgcgggcgctgc
tgctgctggaggagctggagcacgccaagaagcgcgcgccatctatgccgagttctgggcccagcttcacctgcga
cgctatcacatgaccgagccccaccccgaggcgccgctatctgtgcatcgagaaggcgctggcccagtcggcgctgt
ccccgagggagctgaactatatcaacgcgcacgccaccagcaccgccggcgacatcaaggagtatcagggcctggcgc
actgcttcggccagaactcggagctgcgctcaactccaccaagagcatgatcgccacctgctggcgcgccggcgcgct
ggaggcggtcgccgtggtccaggcgatccgcacggctggatccacccaacatcaacctggaggaccccgacgagggcgt
ggacgccaagctgctggtcgccccaagaaggagaagctgaaggtgaaggtcgccctgtcgaactccttcggcttcggcgcc
cacaacagctcgatcctgttcgcccctgcaac**TGA**

SEQ ID NO: 247

C. hookeriana KAS IV enzyme

MATASCMVASPFCTWLVAACMPTSSDNDPRSLSHKRLRLSRRRRTLSSHCSLRGSTFQCL
DPCNQQRFLGDNGFASLFGSKPLRSNRGHLRLGRTSHSGEVMAMQPAQEVSTNKKPAT
KQRRVVVTGMGVVTPLGHDPDVYYNNLLDGISGISEIENFDCSQFPTRIAGEIKSFSTDG
WVAPKFSEMRMDKFMLYMLTAGKKALADGGITEDAMKELNKRKCGVLIGSGLGGMKVFSDS
IEALRTSYKKISPFVFPSTTNMGSAIAMD LGWMGPYNYSISTACATS NFILNAANHII
KGEADMMLCGGSDAAVL PVGLGGFVACRALSRNNDPTKASRPWDSNRDGFVMGEGAGVL
LLEELEHAKKRGATIAEFLGGSFTCDAYHMTPEHPGAGVILCIEKALAQSGVSREDVN
YINAHATSTPAGDIKEYQALAHCFGQNSELRVNSTKSMIGHLLGGAGGVEAVAVVQAI RT
GWIHPNINLEDPDEGVDAKLLVGPKEKLVKVGLSNSFGFGGHNSSILFAPCN

SEQ ID NO: 248

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

ggtaccgcctgcaacgcaagggcagccaagcgctcccaccgcgctgaaccgacacgtgcttggcgccctgccgctgctt
gccgcatgcttgctggtgaggtgggcagtgctgcatgctgattgaggcttggtcatcggggtgaagcttatgtgtgctgg
gcttgcatgccgggcaatgcgcatggtggcaagagggcggcagcactgctggagctgccgggtgctccaggtggttcaatcg
cggcagccagagggatttcagatgatcgcgctacaggttgagcagcagtgtagcaaggttagcagtttccagaatgatcggt

tcagctgttaatcaatgccagcaagagaaggggtcaagtgcacacacgggcatgccacagcacgggcaccggggagtggaaatg
 gcaccaccaagtgtgtgcagccagcatcgccgctggctgtttcagctacaacggcaggagtcaccaacgtaaccatgagctg
 atcaacactgcaatcatcgggcgggcgtgatgaagcatgctggcgaagacacatgggtgtcggtatgctccggctgctgctg
 ctgcgcaacgcttgagttggcagcaggctcagccatgcaactggatggcagctgggctgcaactgcaatgtggtggaataggatgc
 aagtggagcgaataacaaacccctctggctgcttgcctgggtgcatggcatcgacacatcagcaggagcgcatgcgaaggagctgg
 ccccatgcacgcatgcaaaacgggagcgacacgagtgccacactgtcaccaggcccgcaagctttgcagaacatgctcatgg
 acgcatgtagcgtgacgtcccttgacggcgctcctctcggtgtgggaaaacgcaatgcagcacaggcagcagaggcgggcgga
 gcagagcgggcgagcagcgggggggccacccttctgcggggtcgccccagccagcggtgatgcgtgatccaaacgag
 ttacattcatttgcagcctggagaagcaggctggggccttgggctggtgcagcccgaatggaatgcgggacggccaggcta
 gcagcaaaggcgctccctactcctcgcatgatgttccatagtcatttggatgctcatttgggtggggcgggcggtgtttcttctg
 ttgaaaaacggccagctcagcaacctgtcccgtgggtccccgtccgatgaaatcgtgtgcagccgatcagctgattgccgg
 ctgcgaagtagggcctccttctgctgcctctctcctgctccgccttagatctagaatatcaATGatcgagcaggacggcc
 tccacgcccgtccccgcgctgggtggagcgctgttcggtacgactgggcccagcagaccatcggtgctccgacgccc
 ccgtgttcgctgtccgcccagggcgccccgtgctgttctgaagaccgacctgtccggcgccctgaacgagctgcaggacg
 agggcgcccgtgctggtggtggccaccacggcgctgctgcgcccgtgctggacgtggtgacggagggcgccgacgac
 tggctgctgctgggaggtgcccggccaggacctgctgtctccacactggccccgcgagaaggtgtccatcatggccgac
 gccatgcgcccgtgcacacccctggaccccgcacactgccccttgaccaccaggccaagcaccgcatcgagcgcccgcacc
 cgcatggaggccggcctggtggaccaggacgacctggacgaggagaccaggcctggccccgcgagctgttcgcccgc
 ctgaaggcccgatgcccgcgagggacctggtggtgacccacggcgacgctgctgcccacatcatggtggagaac
 ggccgcttctccggttcatcgactgcggccgctggcggtggccgacggctaccaggacatcgccctggccacccgcgacatc
 gccgaggagctgggcccaggtgggcccagcgttctggtgctgtacggcatcgccgccccgactccagcgcatcgctt
 ctaccgctgctggacgagttcttTGAcattggcagcagcagctcgatagtatgcacactctggacgctggtcgtgtgatg
 gactgttgcgcccacacttgccttgacctgtgaatatccctgccgctttatcaaacagcctcagtggttgcctgtgtgtacg
 cgctttgcgagttgctagctgcttgcctatttgcgaataccaccccagcatcccctccctcgttcatatcgcttgcacccaac
 gcaacttatctacgtgtcctgctatccctcagcgctgctcctgctcactgcccctgcacagccttgggttgggctccgctg
 tattctctggtactgcaacctgtaaacaggcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaggatcc
 cgctctcgaacagagcgcgagaggaaacgtgaaggtctgcctctgtgcacactcagcgccgcatacaccacaataaccactg
 acgaatgcgcttgggttcttgcctattagcgaagcgtccggttcacacacgtgccacgcttggcgaggtggcaggtgacaatgatcg
 tggagctgatggtcgaacgttcacagcctaggagtttaggtccagcgtccgtggggggggacgggctgggagcttgggcccggg
 aagggaagacgatgcagctccctctggggagtcacagccgactgtgtgtgtgactgtgcccgcgagcactcacacgcaaaat
 gcctggccgacaggcaggccctgtccagtgaacatccacggctccctctcatcaggctcaccttgcctcattgacataacggaatgcg
 taccgctctttcagatctgtccatccagagggggagcaggctccccaccgacgctgtcaaaactgcttctgcccacccgaaaaca
 ttattgtttgagggggggggggggggggcagattgcattgcgggatctcgtgaggaacatcactgggacactgtggaacacagt
 gagtgcagtatgcagagcatgtatgctaggggtcagcgaggaagggggccttccagctcctccatgccactgcacgcatccacg
 actcaccaggaccagcttcttgatcggttccgctcccggtggacaccagtgtgtgacctctggactccaggatgcgtgcaccgcaaa
 ggccagccgatcgtgccgattcctgggtggaggatatgagtcagccaacttggggctcagagtgacactggggcacgatacga
 acaacatctacacgctgtcctccatgtgcacaccacagcttgcctccacctgaatgtggcgcatgggcccgaatcacagccaat
 gtcgctgctgcataatgtgatccagacccctctccgcccagatgccgagcggatcgtggcgctgaatagattcctgtttcagatcact
 gtttgggtccttcttcttctgctcggatgcgctctcgaacaggctgcgtcgggcttccggtcccttctcctccctgaccatcc
 tgcgcggggcaagttgcttgaccctgggctggtaggggttggagggtattaccgctcaggccattccagcccggattcaatt
 caaagtctggggccaccacccctccgcccctgtgtgatcactccacattcgtgcatacactacgttcaagtctgatccaggcgtgtct
 cgggacaaggtgtgcttgaatttgaatctcaaggaccactccagcacagctgctggttgaccccgccctcgcaacatATGgcgac
 cgccagctgcatggtggctgccttctgcacttggtgctgcgctgcatgcccactccagcgacaacgacccccgctcg
 ctgtcccaacagcgctgcgctgagccgcccgcgcccacccctgagctgcactgctcctgcgcccagcacttccagtgc
 tggacccctgcaaccagcagcgttcttggcgacaacggcttcgctgctgcttgcgctccaagccccgcgagcaaccgcg
 gccacctgcgctgggcccacctgcactccggcgaggtgatggcgctgcgagtcagcccgcgagggaggtgagaccaaa
 caagaagcccgcgaccaagcagcgccgctggtgctgacggcatgggctgctgacccccctggggccagaccccgacgtg

SEQ ID NO: 249
coding region for *C. lunceolata* C10:0 specific thioesterase with UTTEX 1435 Δ12 fatty acid desaturase

Acting

[illegible]

SEQ ID NO: 252

G. mangostana C16:0 thioesterase construct

gctcttcgccgcccaactcctgctcagcgcccgcgctgccccgagcgcttggccttttcgccgctcgtgcgctgct
gatgtccatcaccaggtccatgaggtctgccttgcgccggtgagccactgcttcgtccggcgccaagaggagcatgaggag
gactcctggtccagggtcctgacgtggctcgcgctctgggagcgggcccagcatcatctggtctgccgcaccgaggccctccaa
ctggtcctccagcagccgagtcgccgccgacctggcagaggaagacaggtgaggggggtatgaattgtacagaacaaccacg
agccttgtctaggcagaatccctaccagtcaggtttacctggatgacggcctgcgaacagctgtccagcgacctcgtgccgcc
gcttctccgcacgcttcttccagcaccgtgatggcgcgagccagcgccgcacgtggcgctgcgcttcgccatctgaggacagt
cggggaactctgatcagctctaaaccccttgcgcttagtgttgccatccttgcagaccggtgagagccgacttgttgccac
ccccacaccacctcctccagaccaattctgtcaccttttgggaaggcatcgccctcgccctgcagagaggacagcagtgccc
agccgctgggggttggcggtgacgctcaggtacccttcttgcgctatgacacttcagcaaaaggtagggcggtgcgagac
ggcttccggcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctccgatgccgtcca
ggcgagcgctgttaaatagccaggccccgattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcac
taccacttctacacaggccactcgagcttgatcgactcgcctaaggggcgctcctcctcttcgtttcagtcacaacccgaaa
ctctagaatatca**ATG**ctgctgcaggccttctgttctgctgcccgttcgccgccaagatcagcgctccatgacgaacga
gacgtccgacggccccctggtgcacttcccccaacaagggtggatgaacgacccccacggcctgtggtacgacgagaag
gacgccaagtggcactgtacttccagtacaacccgaacgacacccgtctggggagcgcccttgttggggccacgccacgtc
cgacgacctgaccaactgggaggaccagccccatcgccatcgccccgaagcgcaacgactccggcgcttctccggctccatgg
tggtggactacaacaacacctccggcttctcaacgacaccatcgacccgcagcgctgcgtggccatctggacctacaaca
ccccggagtccgaggagcagtacatctctacagcctggacggcggtacaccttaccgagtaccagaagaaccccgctgtg
gccgccaactccaccagttccgcgacccgaaggtcttctggtacgagccctccagaagtggatcatgaccgcgccaagt
ccaggactacaagatcgagatctactcctccgacgacctgaagtctggaagctggagtccgcgttcgccaacgagggcttcc
tcggctaccagtacgagtgcgccggcctgatcgaggtccccaccgagcaggaccccagcaagtctactgggtgatgttcatct
ccatcaacccccggcgccccggcggtccttcaaccagtaactctgctggcagcttcaacggcaccacttcgaggccttcga
caaccagtcccgcgtggtgacttcggcaaggactactacgcccgcagaccttctcaacaccgacccgacctacgggagcg
ccctgggcatcgctgggctccaactgggagtactccgcttcgtgccaccaacccctggcgctcctcatgtccctcgtgcgc
aagttctccctcaacaccgagtaccaggccaacccggagacggagctgatcaacctgaaggccgagccgatcctgaacatca
gcaacgcccggccccctggagccggttcgccaccaacaccacgttgacgaaggccaacagctacaacgtcgacctgtccaacag
caccggcaccctggagttcgagctggtgtacgcgtcaacaccaccagacgatctccaagtccgtgttcgaggacctctccctc
tggttaagggcctggaggaccccgaggagtacctccgatgggcttcgaggtgtccgcgtcctccttcttctggaccgagg
aacagcaaggtgaagttcgtgaaggagaacccctacttcaacaaccgatgagcgtgaacaaccagcccttcaagagcgag
aacgacctgtctactacaaggtgtacggcttctggaccagaacatcctggagctgtacttcaacgacggcgacgtcgtgtcc
accaacacctacttcatgaccacgggaacgcctgggctccgtgaacatgacgacgggggtggacaacctgttctacatcga
caagttccaggtgcgcgaggtcaag**TGA**caattggcagcagcagctcggaatgatcgacacactctggacgctggctgtgtga
tggactgttgcgccacacttgctgccttgacctgtgaatatccctgccgctttatcaaacagcctcagtggtttgatctgtgtac
gcgcttttcgagttgctagctgcttgtgtatttgcgaataccacccccagcatccccttccctcgtttcatatcgcttgcacccaacc
gcaacttatctacgctgtcctgtatccctcagcgctgctcctgctcactgcccctcgacagccttgggttgggctccgctgt
attctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaggatccc
gcgtctgaacagagcgcgagaggaacgtgaaggtctgcctctgtgcacctcagcgcgcatacaccacaataaacacctga
cgaatgcgttgggttcttctccattagcgaagcgtccggttcacacacgtgccacgttggcgaggtggcaggtgacaatgatcggt
ggagctgatggtcgaaacgttcacagcctagggatatc**gaattc**ggccgacaggacgcgctcaaaaggtgctggtcgtgtatgc
cctggccggcaggtcgtgctgctggttagtgattccgcaacccctgattttggcgcttattttggcggtggcaaacgctggcg
cccgcgagccggggcgggcgatgcggtgccccacggctgcccgaatccaagggaggcaagagcgccccgggtcagttga
agggctttacgcgaaggtacagccgctcctgcaaggctgcgtggtggaattggacgtgcaggtcctgctgaagttcctccac
cgctcaccagcggaacaaagcaccggtgtatcaggtccgtgtcatccactctaaagaactcgactacgacctactgatggcct
agattcttcatcaaaaaacgcctgagacacttgcacaggattgaaactccctgaagggaaccaccaggggacctgattgttctt

SEQ ID NO: 253

[illegible]

1. 2023.01.01 - 2023.01.31 2. 2023.02.01 - 2023.02.28 3. 2023.03.01 - 2023.03.31 4. 2023.04.01 - 2023.04.30 5. 2023.05.01 - 2023.05.31 6. 2023.06.01 - 2023.06.30 7. 2023.07.01 - 2023.07.31 8. 2023.08.01 - 2023.08.31 9. 2023.09.01 - 2023.09.30 10. 2023.10.01 - 2023.10.31 11. 2023.11.01 - 2023.11.30 12. 2023.12.01 - 2023.12.31	1. 2023.01.01 - 2023.01.31 2. 2023.02.01 - 2023.02.28 3. 2023.03.01 - 2023.03.31 4. 2023.04.01 - 2023.04.30 5. 2023.05.01 - 2023.05.31 6. 2023.06.01 - 2023.06.30 7. 2023.07.01 - 2023.07.31 8. 2023.08.01 - 2023.08.31 9. 2023.09.01 - 2023.09.30 10. 2023.10.01 - 2023.10.31 11. 2023.11.01 - 2023.11.30 12. 2023.12.01 - 2023.12.31
--	--

gcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcatccccttccctcgtttcatatcgcttgcacccaacc
 gcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcactgcccctcgacagccttggtttgggctccgctgt
 attctcctggtactgcaacctgtaaacagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaggatccc
 gcgtctgaacagagcgcgagaggaacgctgaaggctcgcctctgtcgacactcagcgcgacatacaccacaataaacacctga
 cgaatgcgcttggttcttctgctcattagcgaagcgctccggttcacacacgtgccacgttggcgaggtggcaggtgacaatgatcggt
 ggagctgatggtcgaaacgttcacagcctagggatatgaattcggccgacaggacgcgctcaaaaggtgctggtcgtgtatgc
 cctggccggcaggtcggtgctgctggttagtgattccgcaacccctgattttggcgcttattttggcggtggcaaacgctggcg
 cccgcgagccggggccggcgatgcggtgccccacggctgccggaatccaagggaggcaagagcgccccgggtcagttga
 agggctttacgcgaaggtacagccgctcctgcaaggctgctggtggaattggacgtgcaggtcctgctgaagttcctccac
 cgcctcaccagcggaacaaagcaccgggtgatcagggtcgtcatcactctaaagaactcgactacgacactactgatggcct
 agattcttcatcaaaaaacgctgagacactgcccaggattgaaactccctgaagggaccaccaggggcccgtgagttgttctt
 ccccggtggcgagctgccagccaggctgtacctgtgatcgaggctggcgggaaaataaggctcgtgtgctcaggtcatggg
 aggtgcaggacagctcatgaaacgccaacaatcgcaaatcatgtcaagctaatacagctatttcttctcagagctgtaatt
 gtccaaaattctggtctaccgggggtgatccttctgttacgggcccctccctcaaccctaggtatgcgcgatgcggtcgcgc
 gcaactcgcgcgagggccgaggggtttgggacgggcccgtcccgaatgcagttgcacccggatgcgtggcacctttttgcgat
 aatttatgcaatggactgctctgcaaaattctggtctgtgcgcaaccctaggtatcagcggcgtaggatttctgaatcattcgtc
 ctgatggggagctaccgactaccctaataatcagcccagctgctgacgcagcgctccacttttgtgcacacattccattcgtgcc
 caagacatttattgtgtggaagcgctcccagttacgctcacctgtttcccagcctccttactgttctgtcgacagagcggggc
 cacaggccgggtcgcagccactagtATGctgaagctgtctgcaacgtgaccaacaacctgcacaccttctcttcttctccgac
 tctcctctgttcatcccgtgaaccgcccgcacatcgccgtgtcctccggcgccctccagctgcgcaagcccgcctggacc
 ccttgcgcgctgatctccgcccagcagggtccatctccccgtgaactctctgacccccgcgacccgctgcgcgcccggcc
 gcctgatggaggacggctactctacaaggagaagttcatcgtgctcctacgaggtgggcatcaacaagaccgcccaccg
 tggagaccatcgccaacctgtcgcaggaggtggcctgcaaccacgtgcagaagtgcggcttctccaccgacggcttcgccac
 caccctgaccatgcgaagctgcacctgatctgggtgaccgcccgcacatcgagatctacaagtaccccgctggtccg
 acgtggtggagatcgagacctggtgccagtcggaggccgcacccgcccgcgactggatctgcgcgactccgcca
 ccaacgaggtgatcgccgcgacacctccaagtgggtgatgatgaaccaggacacccgcccgtcagcgcgtgaccgac
 gaggtgcgcgacgagtacctggtgttctgccccgcgagccccgcctggccttcccagaggagaacaactctcctctgaaga
 agatcccaagctggaggaccccgcccagctactccatgctggagctgaagccccgcgcgacccctggacatgaaccagc
 acgtgaacaacgtgacctacatcggtggtgctggagtccatccccaggagatcatcgacacccacgagctgcaggtga
 taccctggactaccgcccgcgagtgccagcaggacgacatcggtgactccctgaccacctccgagatccccgacgaccccat
 ctcaagttaccggcaccacaggtccgcatgtcctcatccaggggccacaacgagctccagttcttgcacatgctgcgcct
 gtccgagaacggccaggagatcaaccgcccgcacccagtggcgaagaagtctcccgatggactacaaggaccac
 gacggcgactacaaggaccagacatcgactacaaggacgacgacgacaagTGAatcgatagatctcttaaggcagcag
 cagctcggatagtatcgacacacttggacgctggtcgtgtgatggactgttgcggccacacttgccttgacctgtgaatatccct
 gccgctttatcaaacagcctcagtggtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgcgaataaccac
 cccagcatccccttccctggttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgtgctcctgct
 cctgctcactgcccctcgacagccttggtttgggctccgctgtatttctcgtgactgcaacctgtaaacagcactgcaatgctga
 tgcacgggaagtagtgggatgggaacacaaatggaaagcttaattaagagctcttggtttccagaaggagttgctccttgagcctt
 cattctcagcctcgataacctccaaagccgctctaattgtggaggggttcgaatttaaaagcttggaatgttggttcgtgctg
 gaacaagcccagacttgttgcactgggaaaaggaccatcagctcaaaaaaacttgccgctcaaacccgctacctctgcttgc
 gcaatctgcctgttgaaatcgccaccacattcatattgtgacgcttgagcagctgtgaattgcctcagaatgtggaatcatctgcc
 cctgtgcgagcccatgccaggcatgtcgcggcgaggacacccgcaactcgtacagcagaccattatgtacctcacaatagtca
 taacagtgacctatcttctgaagctcccaacgagcacctccatgctctgagtgccacccccggccctgggtgctgcggagggc
 aggtcaaccggcatggggctaccgaaatccccgacccgatccaccacccccgcgatgggaagaatctctcccgggatgtgggc
 ccaccacagcacaacctgctggcccaggcgagcgtcaaacataccacacaaatatccttggcatcgccctgaattccttctgc

SEQ ID NO: 256

[illegible]

SEQ ID NO: 255

[illegible]

caaccagtcctcggtggacttcggcaaggactactacgccctgcagaccttcttaacaccgacccgacctacgggagcg
 ccctgggcatcgctgggctccaactgggagtactccgccttcgtgccaccaaccctggcgctcctcatgtccctcgctgc
 aagttctccctcaacaccgagtaccaggccaaccggagacggagctgatcaacctgaaggccgagccgacgtgaacatca
 gcaacgccggccccctggagccggttcgccaccaacaccacgttgacgaaggccaacagctacaacgtcgacctgtccaacag
 caccggcaccctggagttcgagctggtgtacgccgtcaacaccaccagacgatctccaagtccgtgttcgaggacctctccctc
 tggttcaaggccctggaggaccccgaggagtacctccgcatgggcttcgaggtgtccgctcctccttctcctggaccgagg
 aacagcaaggtgaagttcgtaaggagaacccctacttcaccaaccgcatgagcgtgaacaaccagcccttaagagcgag
 aacgacctgtctactacaaggtgtacggcttgctggaccagaacatcctggagctgtacttaacgacggcgacgtcgtgtcc
 accaacacctacttcatgaccacgggaacgcctgggctccgtgaacatgacgacgggggtggacaacctgttctacatcga
 caagttccaggtgctgaggtcaag**TGAcaattggcagcagcagctcgga**tagtatcgacacactctggacgctggctgtgtga
 tggactgttgcggcacacttgctgccttgacctgtgaatatccctgccgctttatcaaacagcctcagtggtttgatctgtgtac
 gcgcttttgcgagttgtagctgctgtgtatttgcgaataccaccccgacatcccttccctcgtttcatatcgcttgcatccaac
 gcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcactgcccctcgacagcccttggttgggctccgctgt
 attctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggag**gatccc**
 gcgtctgaacagagcgcgagaggaacgctgaaggtctgcctctgtcgacctcagcgcgacataccacaataaccacctga
 cgaatgcgttgggttctcgtccattagcgaagcgtccggttcacacagtgccacggttggcgaggtggcaggtgacaatgatcggt
 ggagctgatggtcgaaacgttcacagcctagggatc**gaattc**ctttcttgcgctatgacacttcagcaaaaggtagggcggg
 ctgcgagacggctcccgcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctcc
 gatgccgctccaggcgagcgtgtttaaatagacaggcccccgattgcaaagacattatagcgagctaccaagccatattc
 aaacacctagatcactaccacttctacacaggccactcgagcttgtgatcgactccgctaaggggcgctcttctcctcgttt
 cagtcacaacccgcaaac**actagtATG**gccaccgcatccactttctcgcggttcaatgcccgtgcggcgacctgcgtcgtc
 ggcgggctccgggccccggcgccagcgagggccctcccgctgcgcggcgccgaggtgcacgtgcaggtgacctactc
 cctggcccccgagaagcgcgagatcttaactccctgaacaactgggccagagagaacatcctggtgctgtaaggacgt
 ggacaagtgtgagcaccctcgacttctgcccgaactccgctccgagggcttcgacgagcaggtgatggagctgcgcaa
 gcgctgcaaggagatccccgacgactacttcatcgtgctggtggcgacatgatcaccgaggaggccctgccacctacca
 gacctgctgaacacctggacggcgtgcgcgacgagaccggcgctccctgacccctgggccatctggaccccgcgctgg
 accgcccaggagaaccgccaacggcagctgctgaacaagtacctgtacctgtccggccgctggacatgaagcagatcga
 gaagaccatccagtacctgatcggtccggcatggacccccgcaccgagaacaaccctacctgggcttcatctacacctcct
 tccaggagcgcgccaccttcatctcccacggcaacaccgcccgcctggccaaggagcacggcgacctgaagctggcccaga
 tctgcggcatcatcgccgacgagaagcgccacgagaccgctacaccaagatcgtggagaagctgttcgagatcgac
 cccgacggcaccgtgctggcctggccgacatgatgcgaagaaggtgtccatgcccgccacctgatgtacgacggccag
 gacgacaacctgttcgagaacttctcctcgtggcccagcgctggcggtgtacaccgccaaggactacggcgacatctgg
 agttcctggtgggcccgtgggacatcgagaagctgacggcctgtccggcgagggccgcaaggccaggactacgtgtgc
 acctgccccccgcatccgcccgtggaggagcgcgccagctcccgctgaagaaggcctccgcccccccttctcctggatc
 ttcggccgagatcaacctgatggactacaaggaccacgacggcgactacaaggaccacgacatcgactacaaggacg
 accgacgacaag**TGA**atcgatagatctctaaggcagcagcagctcggaatgatcgacacactctggacgctggctgtgtgatg
 gactgttgcggcacacttgctgccttgacctgtgaatatccctgccgctttatcaaacagcctcagtggtttgatcttgtgtacg
 cgcttttgcgagttgtagctgcttgtgtatttgcgaataccaccccgacatcccttccctcgtttcatatcgcttgcatccaaccg
 caacttatctacgctgtcctgctatccctcagcgctgctcctgctcactgcccctcgacagcccttggttgggctccgctgta
 ttctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagcttaa
 ttaa**gagctctt**gttttcagaaggagttgctccttgagccttattctcagcctcgataacctccaaagcgctctaattgtggag
 ggggttgaatttaaaagcttggaaatgttggctcgtcgtctggaacaagcccagactgttgcactgggaaaaggacatcag
 ctcaaaaaaacttgcgctcaaacgcgtacctctgcttgcgcaatctgcctgttgaaatcgccaccaattcatattgtgacgc
 ttgagcagctgtaattgctcagaatgtggaatcatctgccccctgtgcgagccatgcccaggcatgtcgggcgaggacccc
 gccactcgtacagcagaccattatgtacctcacaatagttcataacagtgaacatatttctgaagctcccaacgagcacctcc
 atgctctgagtgccacccccggcctgggtccttgcggaggcgaggtcaaccggcatggggctaccgaaatccccgaccggatc
 ccaccacccccgatgggaagaatctctccccgggatgtggggccaccaccagcacaacctgctggccaggcgagcgtcaaac

SEQ ID NO: 258

[illegible]

SEQ ID NO: 257

[illegible]

[illegible]

SEQ ID NO: 260

[illegible]

SEQ ID NO: 259

[illegible]

[illegible]

SEQ ID NO: 262

C. tinctorius ACP-thioesterase at $\Delta 12$ fatty acid desaturase locus construct

gctcttcgggctggtctgaatccttcaggcgggtgttacccgagaaagaaagggtgccgatttcaaagcagacccatgtgccgggc
cctgtggcctgtgttggcgctatgtatcaccctccacccaattgtcgccagtttgcgactccataaactcaaaacagcagct
tctgagctgcgtgttcaagaacacctctggggttgtcaccgcgaggtcgacgccagcatggctatcaagacgaacaggca
gcctgtggagaagcctccgttcacgatcgggacgtcgcaaggccatccccgcgactgtttcgagcgtcggcgcttcgtagca
gcatgtacctggcctttgacatcgcggtcatgtcctgtctacgtcggtcgacgtacatcgacctgcacgggtgcctacgtgggt
caagtacggcatcatgtggcgctctactgttcttcagggtgtgttgagggtttggttgcctgtattgaggtcctgtggcgcg
atggaggagaaggcgctgtccgctgacccccggctacccctccggcaccttcaggggcgcttcggcacgggtgtctgggtg
tgcgcgcacgagtgccgcaaccaggcctttctccagccaggccatcaacgacggcggtgggctggtgttcacagcctgtgct
ggtgcctactactcctggaagcactcgacccgggtaccctttcttcgctatgacacttcagcaaaaggtagggcggtgctgag
acggcttcccgcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctccgatgccgctc
cagggcgagcgctgtttaaatagccaggccccgattgcaaagacattatagcgagctacaaagccatattcaaacacctagatc
actaccattctacacaggccactcgagcttgatgcactccgctaagggggcgcttctctcttcgtttcagtcacaaccgcga
aactctagaatatca**ATG**ctgctgcaggccttctgttctgtggtgcccgttcgccccaagatcagcgctccatgacgaac
gagacgtccgacgccccctggtgcacttcaccccaacaagggtggatgaacgaccccaacggcctgtggtacgacgaga
aggacgccaagtggcactgtacttccagtacaacccgaacgacacccgtctgggggacgcccctgttctggggccacgccacg
tccgacgacctgaccaactgggaggaccagccatcgccatcgccccgaagcgcaacgactccggcgcttctccgggtccat
ggtggtggactacaacaacacctccggcttcttaacgacacccatcgacccgcgacgctgctggtggccatctggacctaca
acacccggagtcggaggagcagtagatctctacagcctggacggcggtacaccttcaccgagtaccagaagaaccccg
gctggccgcaactccacccagttccgcgacccgaaggtcttctggtacgagccctccagaagtggatcatgaccgcggcca
agtcccaggactacaagatcgagatctactctccgacgacctgaagtcctggaagctggagtccgcttcgccaacgagggc
ttctcggctaccagtacgagtccccggcctgatcgaggtccccaccgagcaggacccagcaagtcctactgggtgatgttc
atctccatcaacccggcgccccggcgcggtccttcaaccagtacttcgtcggcagcttcaacggcacccacttcgaggcct
tcgacaaccagtcgctggtggtgacttcggcaaggactactacgcctgcagaccttctcaacaccgacccgacctacggg
agcgccctgggcatcgctgggctccaactgggagtactccgcttcgtgccaccaacccctggcgctcctccatgtcctcgt
gcgcaagtctcctcaacaccgagtaccaggccaacccggagacggagctgatcaacctgaaggccgagccgatcctgaac
atcagcaacgcggccccctggagccggttcgccaccaacaccacgttgacgaaggccaacagctacaacgtcgacctgtcca
acagcaccggcacccctggagttcgagctggtgtacgcgtcaacaccacccagacgatctccaagtccgtgttcgaggacctt
ccctctggttaaggccctggaggaccccgaggagtacctccgatgggcttcgaggtgtccgctcctccttctcgtgaccg
cggaacagcaaggtgaagttcgtgaaggagaacccctacttcaccaaccgcatgagcgtgaacaaccagccctcaagag
cgagaacgacctgtctactacaaggtgtacggctgtgtggaccagaacatcctggagctgtacttcaacgacggcgacgtcg
tgtccaccaacacctacttcatgaccacgggaacgcctgggtccgtgaacatgacgacgggggtggacaacctgttctac
atcgacaagttccaggtgcgcgaggtcaag**TGA**aattggcagcagcagctcggatagtatcgacacactctggacgtggtc
gtgtgatggactgttgcgccaacttgctgccttgacctgtgaatatccctgccgctttatcaaacagcctcagtggtttgatctgt
gtgtacgcgcttttgcgagttgctagctgctgtgtctatttgcgaataccacccccagcatcccctccctcgttcatatcgcttgcac
ccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcactgcccctcgacagccttggttgggctc
cgctgtatttctcgtgactgcaacctgtaaacagcactgcaatgctgatgcagggaagtagtgggatgggaacacaaatgga
ggatcccgctctgaacagagcgcgagaggaacgtgaaggtctcgctctgtcgacctcagcgcgccatacaccacaataac
cacctgacgaatgcgcttggttcttctgctcattagcgaagcgtccggttcacacagtgccacgttggcgaggtggcaggtgacaat
gatcgggtggagctgatggtcgaaacgttcacagcctagggatat**gaattc**
cttcttcgctatgacacttcagcaaaaggtagggcggtgagacggcttccggcgctgcatgcaacaccgatgatg
cttcgacccccgaagctccttcggggctgcatggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggcccc
cgattgcaaagacattatagcgagctacaaagccatattcaaacacctagatcactaccattctacacaggccactcgagc
ttgtgatcgactccgctaagggggcgcttctccttctcgtttcagtcacaacccgcaaac**actagtATG****gccaccgcatccac**
tttctcggcggttcaatgcccgtgcggcgacctgcgtcgctcggcggttccggggcccgccagcgaggccctccccgt

[illegible][illegible]

SEQ ID NO: 268

SEQ ID NO: 267

SEQ ID NO: 266

[illegible]

[illegible]

atgaaccagcacgtgaacaacgtgacctacatcggtgggtgctggagtccatccccaggagatcatcgacacccacgag
 ctgcagaccatcacctggactaccgccgagtgccagcacgacgacatcgtggactccctgacctcgtggagccctccg
 agaacctggaggccgtgtccgagctgcgcccaccaacggctccgccaccaccacgcccggcgagaggactgccgcaac
 ttctgcacctgctgcgctgtccggcgacggcctggagatcaaccgcccgcaccgagtgccgcaagaagtcccccga
 tggactacaaggaccacgacggcgactacaaggaccacgacatcgactacaaggacgacgacgacaagTGAatcgata
 gatctcttaaggcagcagcagctcgatagtagtacacactctggacgctggctgctgtgatggactgttgccgccacacttgcctgcc
 ttgacctgtgaatatccctgccgcttttatcaaacagcctcagtggtgttgatcttgtgtgtacgctgttgccgagttgctagctgcttg
 tgctatttggaataaccaccccccagcatccccttccctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatc
 cctcagcgtgctcctgctcctgctcactgcccctcgacagccttggttgggctccgctgtatttctcctggactgcaacctgtaaa
 ccagcactgcaatgctgatgcaggggaagtagtgggatgggaacacaaatggaaagcttaattaagagctcttgtttccagaagg
 agttgctccttgagcctttcattctcagcctcgataacctcaaagccgctctaattgtggagggggttcgaatttaaaagcttgga
 atgttggtcgtgcgtcgtggaacaagcccagactgtgtcactgggaaaaggaccatcagctccaaaaaacttgccgctcaaac
 cgctacctctgctttcgcaatctgcccgttgaaatcgccaccaattcatattgtacgcttgagcagctgtaattgctcag
 aatgtggaatcatctgccccctgtgagcccatgccaggcatgtcgccggcgaggacacccgccactctacagcagacattat
 gctacctacaatagttcataacagtgaccatatttctgaagctcccaacgagcacctccatgctctgagtgccacccccggc
 cctggtgcttgccgagggcaggtcaaccggcatggggctaccgaaatccccgaccggatccaccacccccgcgatgggaagaa
 tctctccccgggatgtggggccaccacagcacaacctgctggcccaggcgagcgtcaaaaccataccacacaaatatccttggcat
 cggccctgaattccttctgcccgtctgctacccggtgcttctgccaagcaggggttgctagggatcgtccgagtcgcgaaccct
 tgtcgcgtggcgggctgttcgagcttgaagagc

SEQ ID NO: 269

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

ATGgccaccgcatccactttctcggcgttcaatgcccgtgcggcgacctgcgtcgtcggcgggtccgggccccggcgccca
 gcgaggccctccccgtgcgcccgcgcctcctacgagtggtggcatcaacaagaccgccaccgtggagaccatcgcca
 acctgctgcaggaggtgggtgcaaccacgcccagtcctggtggttctccaccgacggcttcgccaccaccacctccatgcgca
 agatgcacctgatctgggtgaccgccgcatgcacatcgagatctacaagtaccccgctggtccgacgtggtggaggtgga
 gacctggtgccagtccgagggccgcatcgccacccgcgactggatcctgaccgactacgccaccggccagatcatcggcc
 gcgccacctccaagtgggtgatgatgaaccaggacacccgcgctgcagaaggtgaccgacgacgtgcgcgagggagtacc
 tgggtgttctgccccgcgagctgcgctggccttccccgaggagaacaaccgctcctccaagaagatctccaagctggaggac
 cccgcccagtactccaagctgggctggtgccccgcgcccgcgactggacatgaaccagcacgtgaacaacgtgacctacat
 cggctgggtgctggagtccatccccaggagatcatcgacacccacgagctgcagaccatcacctggactaccgccgagat
 gccagcacgacgacatcgtggactccctgacctccgtggagccctccgagaacctggaggccgtgtccgagctgcgcggcacc
 aacggctccgccaccaccacggcggcgagggactgccgaacttctgacctgctgcgctgtccggcgacggcctgga
 gatcaaccgcccgcacccagtggtgcgaagaagtcgcccgcagtgactacaaggaccacgacggcgactacaaggacc
 acgacatcgactacaaggacgacgacaagTGA

SEQ ID NO: 270

C. camphora C14:0 thioesterase construct

gctcttcgccgcgccaactcctgctcagcgcgcccgcgctgcgccgacgacgcttggccttttcgccgctcgtgcgctcgt
 gatgtccatcacagggtccatgaggtctgccttgcgccggctgagccactgcttctcggggcggaagaggagcatgaggag
 gactcctggtccagggtcctgacgtggtcgcgctctgggagcgggacgacatcatctgctcgcgaccgaggccgctccaa
 ctggtcctccagcagccgagtcgcccgcgacctggcagaggaagacaggtgaggggggtatgaattgtacagaacaaccacg
 agccttgtctaggcagaatccctaccagtcagtgctttacctggatgacggcctgcgaacagctgtccagcgacccctcgtgcgccc
 gcttctccgcacgcttcttccagcaccgtgatggcgagccagcgccgcacgctggcgctgcgcttgcggatctgaggacagt
 cggggaactctgatcagctcaaaaccccttgcgcttagtgttgcatccttgcagaccggtgagagccgacttgtgtgcgccac
 cccacacacacctcctccagaccaattctgtcaccttttggcgaaggcatcggcctcgccctgcagagaggacagcagtgccc

agccgctgggggttgccgatgcacgctcaggtaccctttcttgcgtatgacacttccagcaaaaggtaggcgggctgcgagac
 ggcttccggcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctccgatgccgtcca
 gggcgagcgctgtttaaatagccaggccccgattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcac
 taccattctacacaggccactcgagcttgatgcactccgctaagggggcgctcttctcttctgttcagtcacaacccgcaaa
 ctctagaatatcaATGatcgagcaggacggcctccacgccggctccccgcccgtgggtggagcgctgttcggctacgact
 gggcccagcagaccatcggtgctccgacgcccgtgttcgctgtccgcccaggggcgccccgtgtgttcgtgaagaccg
 acctgtccggcgccctgaacgagctgcaggacgaggccgcccgtgtcctggctggccaccacggcggtgcctgtgcggcc
 gtgttgacgtggtgaccgaggccggccgactggctgctgtggcgaggtgcccggccaggacctgtgtcctccacct
 ggccccgcccagagaaggtgtccatcatggccgacgcatgcgcccgtgcacacctggaccccgccacctgccccctcgacca
 ccaggccaagcaccgcatcgagcgcccgcacccgcatggaggccggcctgtgtgaccaggacgacctggacgaggagc
 accagggcctggccccgcccagctgttcgcccgcctgaaggcccgcatgcccgacggcgaggacctggtggtgacccacgg
 cgacgctgctgcccacatcatggtggagaacggccgttctccggttcatgactgcccgcctgggctggccgaccg
 ctaccaggacatcgccctggcaccgacatcgccgaggagctgggcccgcgagtgggccgaccgcttctggtgtgtacg
 gcatcgccgccccgactcccagcgcatgccttctaccgctgtggacgagttctcTGAcaattggcagcagcagctcggat
 agtatcgacacactctggacgtggtcgtgtgatggactgttgcggccacacttctgccttgacctgtgaatatccctgccgtttat
 caaacagcctcagtggtttgatcttgtgtacgcgcttttgcgagttgctagctgcttgtgtatttgcgaataccacccccagcatc
 cccttccctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctact
 gccccctgcacagccttggtttgggctccgctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacggga
 agtagtgggatgggaacacaaatggaggatcccgctctcgaacagagcgccgagaggaaacgtgaaggtctgcctctgtcgca
 cctcagcgccgcataccacacaaataaccacctgacgaatgcgttgggttctcgtccattagcgaagcgtccggttcacacacgtgc
 cacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatcgaattcctttcttgc
 gctatgacacttccagcaaaaggtaggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatgatgcttcgaccccc
 gaagctccttcggggctgcatggcgctccgatgccgtccaggcgagcgctgtttaaatagccaggccccgattgcaaagaca
 ttatagcgagctaccaaagccatattcaaacacctagatcactaccattctacacaggccactcgagcttgatgcactccgcta
 agggggcgctcttctcttctgtttcagtcacaacccgcaaacactagtATGgccaccgcatccatttctcgggcttcaatgcc
 gctgcggcgacctgcgtcgtcggcggtcggggccccggcgcccagcagggccctcccgtgcggaagcagccccgac
 tggccatgctgttcgccgtgatcaccacatcttctccgcccggagaagcagtggaaccaacctggagtggaaagccaaagcc
 aacccccccagctgctggacgaccacttcggccccacggcctggtgttcgcccaccttcgcatccgacgtacgaggtg
 ggccccgaccgtccaccagcatcgtggcgtgatgaaccaacctgcaggaggccgacctgaaccaagcgaagtcctgggca
 tctggggcgacggcttcggcaccacctggagatgtcaagcgcgacctgatctgggtggtgaagcgacccacgtggcgtg
 gagcgctaccccgctggggcgacacctggaggtggagtgtgggtggcgccctccggcaacaacggccgcccacgac
 ttctggtgcgagctgcaagaccggcgagatcctgaccgctgcacctcctgagcgtgatgatgaacacccgacccgccc
 cctgagcaagatccccgaggaggtgcgcccgcgagatcgccccgcttcatcgacaacgtggcgtgaaggacgaggagat
 caagaagccccagaagctgaacgactccaccgcccactacacagggcgctgacccccgctggaacgacctggacatc
 aaccagcagctgaacaacatcaagtacgtggactggatcctggagaccgtgcccgacagcatcttcgagagccaccacatct
 cctcttcaccatcgagtaccgcccgcgagtgcacctggacagcgtgctgcagtcctgaccacctgagcgggcgtcctccg
 agggccgctggtgtgcgagcacctgtgcagctggaggcgccgagcaggtgctgcgccaagaccgagtggtgccccca
 agctgaccgactccttcgcccgcagcgtgatccccgcgagtcagcgtgatggactacaaggaccacgacggcgactac
 aaggaccacgacatcgactacaaggacgacgacgacaagTGAactcgaggcagcagcagctcgatagatcgacacactct
 ggacgctggtcgtgtgatggactgttgcggccacacttctgccttgacctgtgaatatccctgccgttttatcaaacagcctcagtg
 tgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgtatttgcgaataccacccccagcatcccttccctgcttcat
 atcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgccccctgcacagcct
 tggtttgggctccgctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatggga
 acacaaatggaaagcttgagctccttgtttccagaaggagttgctccttgagccttctcagcctcgataacctccaaagccg
 ctctaattgtggagggggttgaatttaaaagcttggatgttgggtcgtgcttggaaacaagccagacttgtgtcactggga
 aaaggacctcagctcaaaaaaacttgcgctcaaacccgctacctctgcttgcgcaatctgcctgttgaaatcgccaccaca

ttcatattgtacgcttgagcagctgtgaattgcctcagaatgtggaatcatctgccccctgtgcgagcccatgccaggcatgtcgc
gggagaggacacccgccactcgtacagcagaccattatgtacctcacaatagttcataacagtaccatatttctcgaagctccc
caacgagcacctccatgctctgagtgccacccccggccctggtgcttgaggaggcaggtcaaccggcatggggctaccgaaa
tccccgaccggatcccaccaccccgcatgggaagaatctctcccgggatgtgggcccaccaccagcacaaacctgtggccca
ggcagcgtcaaacataccacacaaatatccttgcatcgccctgaattccttctgcccgtctgtaccgggtcttctgtccga
agcaggggtgctagggatcgctccgagtcgcgaaccctgtcgctggcggggctgttcgagcttgaagagc

SEQ ID NO: 271

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

ATGgccaccgcatccactttctcggcgttcaatgcccgtgcggcgacctgcgtcgctcggcgggctccgggccccggcgccca
gagaggccctccccgtgcgcg**gacgaccc**ccgactggtccatgctgttcgcccgtgatcaccaccatcttctccgcccgcgaga
agcagtggaccaacctggagtggagcccaagcccaacccccccagctgctggacgaccacttcggccccacggcctggt
gttcgcccgcaccttcgccatccgagctacgaggtgggccccgaccgctccaccagcatcgtggccgtgatgaaccacctga
ggaggccgcccgaaccacgccaagtccgtgggcatcctgggcgacggcttcggcaccaccctggagatgtccaagcgcgac
ctgatctgggtggtgaagcgcacccacgtggccgtggagcgctaccccgcctggggcgacaccctggaggtggagtgtctggg
tgggcgctccggcaacaacggccgcccgcacgacttctggtgcgcgactgcaagaccggcgagatcctgaccgctgcacc
tcctgagcgtgatgatgaacacccgcacccgcccgtgagcaagatccccgaggaggtgcgcggcgagatcgccccgcct
tcacgcacaacgtggccgtgaaggacgaggagatcaagaagcccagaagctgaacgactccaccgcccactacatccagg
gcccgtgacccccgctggaacgacctggacatcaaccagcacgtgaacaacatcaagtacgtggactggatcctggagac
cgtgcccgcagcatcttcgagagccaccacatctctccttcaccatcgagtaccgcccgcgagtgacccatggacagcgtgt
gcagtcctgaccaccgtgagcggcggtcctccgaggccgcccgtgtgtgcgagcacctgctgcagctggagggcggcgagc
gaggtgtgcgcgccaagaccgagtgggcgcccaagctgaccgactcctccgcccgcagcgtgatccccgcgagtcagc
cgtgatggactacaaggaccacgacggcgactacaaggaccacgacatcgactacaaggacgacgacgacaag**TGA**

SEQ ID NO: 272

C. camphora C14:0 specific thioesterase construct

gaattcgccctcccgtgatcacacaggtgccttgcgagcgtgatcacactattttgggggtcctacagtactgaaatggtgagaagtc
gtactgaaatcaaggatgaacaatgaaaatggtgctgtggtggcttctcaaaggtcaagaatcagtcgctcgcgtcaggaaatcgc
ggcgtcaaccagcgtgggcgcggctcagtgccccgcactggtcaccatagccttctcctgccacagtagcgatccccctgggcgttac
tctcagcagcggctgtactgcctcccagattttcttcttggacctgcggcgctgagaggatgagcaggggtgggccaagggtcaat
cctgaacggccctcattcggtttccaatcccacaacatacccacagcaggtcagaccacgcattccaccatgcgcaccaataacg
tgtccttacctgattgggtgtggcaggctccgtggacaggagtgctcgtcccccgcccagaccgctcccccgctcacggcggcgtcc
gggaccgcagcggctccaccgcggtgtgatccgcttggcggcgagagcagcatcccagccgatttgaccccgcgcatgctccg
aggcttgaggttgccagcaccaccacccgcccggcgacaaggtctccagggtcacgtgcccggaccaggccactcacgatggtgc
gagggccccctcctcgcgaggtgatctgctgcagctacagactgcgacatcgctggcgagtggtcatcagaaggaagcaggtg
tgcagaagggggcacgtggttggtattgagagtagccaaagcttgtgccaatcagaaagtcaacgcagctgctgcctggctcgcgt
ac**aat**cttcttctgcgtatgacacttcagcaaaaggttagggcgggctgcgagacggctcccggcgctgcatgcaacaccgatg
atgcttcgacccccgaagctccttcggggctgcatgggcgtccgatgccgtccaggcgagcgtgtttaaatagccaggcccc
cgattgcaaagacattatagcgagctaccaaagcatattcaaacacctagatcactaccattctacacaggccactcgagcttgtg
atcgactccgctaaggggcgcttctccttcgtttcagtcacaaccgcgaac**ggcgcgccATG**ctgtgcaggccttctgtt
cctgtggcggttcgcccgaagatcagcgctccatgacgaacgagacgtccgaccgccccctggtgacttcacccccaa
caagggctggatgaacgaccccaacggcctgtgttacgacgagaaggacccaagtggcacctgtacttcagtacaacc
gaacgacaccgtctggggagcgccttgttctggggccacgccagctccgacgacctgaccaactgggaggaccagccatc
gcatcgccccgaagcgcaacgactccggcgcttctccggctccatggtggtggactacaacaacacctccggcttcttaac
gacaccatcgacccgcgacgctgcgtggccatctggacctacaacacccggagtcggaggagcagtagcatctcctacag

cctggacggcggtacaccttaccgagtagcagaagaaccccgctgctggccgccaactccaccagttccgagacccgaagg
 tcttctggtacgagccctccagaagtggatcatgaccgaggccaagtcccaggactacaagatcgagatctactctccgacg
 acctgaagtctggaagctggagtcgctgccaacgagggcttctcggtaccagtagcaggtgccccggcctgatcgag
 gtccccaccgagcaggacccagcaagtctactgggtgatgttcatctccatcaaccccgcgccccggcgggcctcttc
 aaccagtagtctgctgagcgttcaacggcaccacttcgaggccttcgacaaccagtagcccgctggtggacttcggaagga
 ctactacgacctgagaccttcttaacacggacccgacctacgggagcgccctgggcatcgctgggctccaactgggagta
 ctccgcttctgctccaccaacccctggcgctctccatgtccctcgctgcgcaagttctccctcaacaccgagtagcaggccaacc
 cggagacggagctgatcaacctgaaggccgagccgatcctgaacatcagcaacgccccgctggagccggttcgccaacca
 acaccagttgacgaaggccaacagctacaacgtcgacctgtccaacagcaccggcacctggagttcgagctggtgtacgc
 cgtcaacaccacccagacgatctcaagtccgtgttcgaggacctctccctggttcaagggcctggaggaccccgaggagta
 cctccgcatgggcttcgaggtgtccgctctcttcttctggaccgagggaacagcaaggtgaagttcgtaaggagaaccc
 ctacttaccacacccgatgagcgtgaacaaccagccctcaagagcgagaacgacctgtctactacaaggtgtacggcttgc
 tggaccagaacatcctggagctgtacttcaacgacggcgacgtcgtgtccaccaacacctacttcatgaccacgggaacgcc
 ctgggctccgtgaacatgacgacgggggtggacaacctgttctacatcgacaagttccagggtgcgaggtcaag**TGAttaa**
ttaactcgaggcagcagcagctcggaatgatcgacacactctggacgtggtcggtgatggactgttccgccacactgtgc
cttgacctgtgaatatccctgccgttttatcaaacagcctcagtggtttgatcttgtgtgtacgcgttttgcagttgtagctgt
tgtgtatttgcgaataccacccacagcatccctccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgtgtcctgct
atccctcagcgtgctcctgctcctgctcactgcccctcgacagccttgggttgggctccgctgtattctcctggtactgcaacctgt
aaaccagcactgcaatgctgatg
cacgggaagtagtgggatgggaacacaaatggaaagctt

gagctccttcttgcgctatgacacttccagcaaaaggtagggcgg
gctgcgagacggcttccggcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctccga
tgccgtccaggcgagcgtgtttaaatacgaggccccgattgcaaagacattatagcgagctacaaagccatattcaaac
cctagatcactaccacttctacacaggccactcgagcttgtgatcgactccgctaagggggcgcttctctctcgtttcagtcaca
accgcaaac

actagtagTGacgttcggggtcgccctccggccatgggcccgggtgtctccctccccggcccagggtcgcggt
 gcgcgccagtagtcggcagtagcaggttttgagagcgg**ggcgagc**cccgactggtccatgctgttcgccgtgatcaccacatctt
 ctccgcccggagaagcagtggaacacctggagtggaagcccaagcccaacccccagctgtggagcaccacttcggc
 cccacggcctggtgttcgcccaccttcgcatccgagctacgaggtgggccccgaccgctccaccagcatcgtggcctg
 atgaaccacctgcaggaggccgacctgaaccacgccaagtccgtgggcatcctgggacgagcgttcggcaccacccctggaga
 tgtcaagcgcgacctgatctgggtggtgaagcgacccacgtggcctggagcgctaccccgctggggcgacacctgga
 ggtggagtgctgggtggcgctccggcaacaacggccgcccacgacttctggtgcgagctgaagaccggcgagatc
 ctgacccgctgcacctccctgagcgtgatgatgaacacccgacccgcccctgagcaagatccccgaggaggtgcgcgcg
 agatcgccccgccttcatcgacaacgtggcctgaaggacgaggagatcaagaagcccagaagctgaacgactccaccg
 ccgactacatccagggcggtgacccccgctggaacgacctggacatcaaccagcagctgaacaacatcaagtacgtgga
 ctggatcctggagaccgtgcccagacg
 atcttcgagagccaccacatctctccttaccatcgagtagccgcccagtagcaccatggacagcgtgctgcagtccttgacca
 ccgtgagcggcggtcctccgaggccggcctggtgtgcgagcacctgctgcagctggagggcgagcagaggtgctgcgcgc
 caagaccgagtggcgccccaaagtgaccgactccttcgcgcatcagcgtgatccccgagtagcagcgtgatggactaca
 aggaccacgacggcgactacaaggaccacgacatcgactacaaggacgacgacgacaag**TGActcgaggcagcagcagc**
tcggatagtagtcacacactctggacgtggtcggtgatggactgttccgccacactgtgccttgacctgcaatatccctgcc
gcttttatcaaacagcctcagtggtttgatcttgtgtgtacgcgttttgcagttgtagctcttgtgtatttgcgaataccaccc
cagcatcccttctcctgtttcatatcgcttgcatcccaaccgcaacttatctacgtgtcctgctatccctcagcgtgctcctgctc
tgctcactgcccctcgacagccttgggttgggctccgctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgat
gcacgggaagtagtgggatgggaacacaaatggaaagctggtagccgtacccatcagcatccgggtgaatcttggcctcaagat
atggccaatcctcacatccagcttggcaaaatcgactagactgtctgaagtgggaatgtggagcaaaaggttgcctgtagcgatcg
acagactggtgggtacattgacaggtgggcagcgccgcatccatcgctgacgagcgagcggccgggttgcctgcccgctgcctg
ccgtcaaagagcggcagagaaatcgggaaccgaaaacgtcacattgcctgatgttgttacatgctggactagacttcttggcgtgg

gtctgctcctcgccaggtgcgcgacgcctcggggctgggtgcgagggagccgtgcggccacgcatttgacaagacccaaagctcgc
atctcagacgggtcaaccgttcgtattatacattcaacatatggtacatacgcaaaaagcatgccaacgatgacataggcgaattc

SEQ ID NO: 273

C. camphora C14:0 specific thioesterase construct

gctcttcggccgcccactcctgctcgagcgcggcgactcgctcgcctcgcccgcgtgcgcccagcgccttggccttt
tcgccgctcgtgcgcgtcgtgatgtccatcaccaggtccatgaggtctgccttgcgcccgtgagccactgcttcgtccggcg
gccaagaggagcatgaggaggactcctggtccaggtctgacgtggctgcggctctgggagcgggcccagcatcatctggtct
gccgacccgaggccgctccaactggtctccagcagccgcagtcgcccgaacctggcagaggaagacaggtgaggggggta
tgaattgtacagaacaaccagagccttgtctaggcagaatccctaccagtcaggtttacctggatgacggcctgcgaacagct
gtccagcagccctcgctgcccgcttctccgcacgcttcttcagcaccgtgatggcgagccagcgcgcaagctggcgctg
cgcttcgccgatctgaggacagtcggggaactctgatcagtctaaaccccttgcgcttagtggtccatccttgcagaccggtg
agagccgacttgttgcgccacccccacaccctctccagaccaattctgtcaccttttggcgaaggcatcgccctcgccct
gcagagaggacagcagtgcccagccgctgggggttgccggatgcacgctcaggtaccctttcttgcgctatgacacttcagcaaa
aggtagggcgggctgcgagacggcttccggcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgca
tgggcgctccgatgccgctccaggcgagcgctgttaaatagccaggccccgattgcaaagacattatagcgagctaccaaagc
catattcaaacctagatcactacccttctacacaggccactcgagcttgtgatcgactccgctaagggggcgcccttctctt
cgtttcagtcacaacccgcaaacggcgcgcc**ATG**ctgctgcaggccttctgttctgctggtggcggttcgcgccaagatcagc
gcctccatgacgaacgagacgtccgaccgccccctgggtgacttcacccccaaacagggctggatgaacgaccccaacggcct
gtgttacgacgagaaggacgccaagtggcacctgtacttcagtacaacccgaacgacacccgtctgggggacgcccctgttct
ggggccacgcccagtcgacgacctgaccaactgggaggaccagccccatcgccatcgccccgaagcgcaacgactccggcg
ccttctccggctccatggtggtgactacaacaacacctccggcttcttaacgacaccatcgacccgcgcagcgctgcgtggc
catctggacctacaacccccggagtcggaggagcagtcacatctctacagcctggacggcggtacaccttcaccgagtacc
agaagaaccccgctgctggcgccaactccaccagttccgcgacccgaaggcttctgtgtacgagccctccagaagtggatc
atgaccgcgccaagtcccaggactacaagatcgagatctactcctccgacgacctgaagtctggaagtggagtcggctt
cgccaacgagggcttctcggtaccagtacgagtccccggcctgatcgaggtccccaccgagcaggaccccagcaagtctt
actgggtgatgttcatctccatcaacccggcgccccggcgcggtccttcaaccagtaactcgtcggcagcttaacggcac
ccacttcgaggccttcgacaaccagtcggcggtggtgacttcggcaaggactactacgcccgcagaccttcttaacaccga
cccgacctacgggagcgccctgggcatcgctggtggtcctcaactgggagtaactccgcttcgtgcccaccaacccctggcgctc
ctccatgtccctgctgcgaagtctcctcaacaccgagtaccaggccaacccggagacggagctgatcaacctgaaggccg
agccgatcctgaacatcagcaacgcccggccccctggagccggttcgccaccaacaccacgttgacgaaggccaacagctaaa
cgtcgacctgtccaacagcaccggcacccctggagttcgagctggtgtacgccgtcaacaccaccagacgatctcaagtccgt
gttcgaggacctctcctctgttcaaggcctggaggaccccgaggagtacctccgcatgggcttcgaggtgtccgctcctcc
ttcttctggaccgcggaacagcaagggtgaagttcgtgaaggagaacccctacttcaccaaccgcatgagcgtgaacaacc
agcccttaagagcgagaacgacctgtctactacaaggtgtacggcttgcgtggaccagaacatcctggagctgtacttaac
gacggcgacgtcgtgtccaccaacacctacttcattgaccacggggaacgccctgggctcgtgaacatgacgacgggggtgg
acaacctgttctacatcgacaagttccaggtgcgcgaggtcaag**TGA**caattggcagcagcagctcggatagtatcgacacac
tctggacgctggtcgtgtgatggactgttgcgcccacacttgctgccttgacctgtgaatatccctgccgctttatcaaacagcctca
gtgtgtttgatcttgtgtacgcgcttttgcgagttgctagctgcttgcctatttgcgaataccacccccagcatccccttccctcgtt
catatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcgtcactgcccctcgacag
ccttgggttgggtccgctgtattctcctggtactgcaacctgtaaacagcagctgcaatgctgatgcacgggaagtagtgggatgg
gaacacaaatggaggatcccgcgtctcgaacagagcgcgcagaggaaacgctgaaggctcgcctctgtcgcacctcagcgggca
tacaccacaataaccacctgacgaatgcgcttggttctcgtccattagcgaagcgtccggttcacacacgtgccacgttggcgagg
ggcaggtgacaatgatcggtggagctgatggtcgaacgttcacagcctagggatcgaattcgccgacaggacgcgcgtca
aagggtgctggtcgtgtatgcctggccggcaggtcgttgcgtgctggttagtgattccgcaacctgattttggcgcttatt
ttggcggtggcaaacgctggcgcccgagccgggcccggcgatgcggtgccccacggctgcccgaatccaagggagg

caagagcgccgggtcagttgaagggctttacgcgaaggtacagccgctcctgcaaggctgctggtggaattggacgt
 gcaggtcctgctgaagttcctccaccgctcaccagcggacaaagcaccggtgtatcaggtccgtgtcatccactctaaaga
 gctcgactacgacctactgatggccctagattcttcatcaaaaacgcctgagacacttgcccaggattgaaactccctgaag
 ggaccaccagggggccctgagttgttcttcccccggtggcgagctgcagaccaggctgtacctgtgatcgaggctggcggga
 aaataggcttcgtgtgctcaggtcatgggaggtgcaggacagctcatgaaacgccaacaatcgacaattcatgtcaagct
 aatcagctatttcttcttcacgagctgtaattgtccaaaattctggtctaccgggggtgatccttcgtgtacgggcccctccctc
 aaccctaggtatgcgcgcatgcggtcgccgcgcaactcgcgaggggcccagggttgggacgggcccgtcccgaatgca
 gttgcacccggatgcgtggcacctttttgcgataatttatgcaatggactgctctgcaaaattctggtctgtcgccaacccta
 ggatcagcggcgtaggatttctgaatcattcgtcctgatggggagctaccgactaccctaatacagcccgactgctgacgc
 cagcgtccacttttgtgcacacattccattcgtgccaaagacatttcatttggtgcgaagcgtcccagttacgctcacctgtt
 cccgacctcttactgttctgtcgacagagcggggccacaggccggctgcagccactagtATGgccaccgcatccactttctcg
 ggttcaatgcccgtgcggcgacctgctgctcgggggtccgggccccggcgccagcgaggcccctcccgtgcgagg
gcacgcccccgactggtccatgctgttcgctgatcaccaccatcttccgcccgcgagagcagtggaaccaacctggagt
 gaagcccaagcccaacccccccagctgctggacgaccattcgccccacggcctggttccgcccacacctcgccatccgc
 agctacgaggtgggccccgaccgctccaccagcatcgtggcgtgatgaaccacctgcaggaggccgcccgaaccacgcca
 agtccgtgggcatcctggcgacggcttcggcaccacctggagatgtccaagcgcgacctgatctgggtggtgaagcgcacc
 cactggtggcgtggagcgctaccccgctggggcgacacccgtggaggtggagtgtgggtggcgccctccggcaacaacggc
 cgccgccagacttctggtgcgcgactgcaagaccggcgagatcctgaccgctgcacctccctgagcgtgatgatgaacac
 ccgaccccgccgctgagcaagatccccgaggaggtgcgcggcgagatcgccccgccttcacgacaacgtggccgtgaag
 gacgaggagatcaagaagccccagaagctgaacgactccaccgcccactacatccagggcgccctgacccccgctggaac
 gacctggacatcaaccagcacgtgaacaacatcaagtacgtggactggatcctggagaccgtgcccgcagacatcttcgaga
 gccaccacatctctcttcaccatcgagtaccgcccgcgagtgccatggacagcgtgtcgcagtcctgaccacctgagcg
 gcggctcctccgaggccggcctggtgtgcgagcacctgctgcagctggagggcggcagcgaggtgtcgcgcgaagaccg
 agtggcgcccaagctgaccgactccttcgcgcatcagcgtgatccccgcgagtcagcgtgatggactacaaggaccac
 gacggcgactacaaggaccacgacatcgactacaaggacgacgacgacaagTGActcagagttaattaactcgaggcagca
 gcagctcggatagtatcgacacactctggacgctggtcgtgtgatggactgttgcggccacacttgctgccttgacctgtgaatacc
 ctgcgcttttatcaaacagcctcagtggtttgatcttgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaatacc
 accccagcatccccttccctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctct
 gctctgctcactgcccctgcacagccttggttgggctccgctgtattctcctggtactgcaacctgtaaaccagcactgcaatgc
 tgatgcacgggaagtagtgggatgggaacacaaatggaaagctgtagagctcctgttttccagaaggagttgctccttgagccttt
 cattctcagcctcgataacctccaaagccgctctaattgtggagggggttcgaatttaaagcttggaatgttggttcgtgcgtctg
 gaacaagcccagacttgttgcactgggaaaaggaccatcagctcaaaaaaacttgccgctcaaacgcgtacctgtcttgcg
 gcaatctgccctgttgaaatcgccaccaattcatattgtgacgcttgagcagctgttaattgcctcagaatgtggaatcatctgcc
 cctgtgcgagcccatgccaggcatgtcgcggcgaggacacccgccactcgtacagcagaccattatgctacctcacaatagtta
 taacagtgacctatcttctgaagctcccaacgagcacctccatgctctgagtgccacccccggccctggtgcttgcggagggc
 aggtcaaccggcatggggctaccgaaatccccgaccggatccaccacccccgcgatgggaagaatctctccccgggatgtgggc
 ccaccaccagcacaactgctgcccaggcgagcgtcaaaaccataccacacaaatatccttggcatcgccctgaattccttctgc
 cgctctgctacccggtgcttctgtccgaagcaggggttgcaggatcgctccgagtcgcaaaccttgcgcgtggcggggcttg
 ttcgagcttctgcagcttgaagagccttagagtcgacctgcaggcatgcaagcttggcgtaacatggatcatagctgttctgtgt
 gaaattgttatccgctcacaattccacacaacatacagccgggaagcataaagtgtaaagcctggggtgcctaagtagtgagctaa
 ctacattaattgcgttgcgctcactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcgcccaacgcgcg
 ggagaggcggttgcgtattgggcgctcttcc

SEQ ID NO: 274

U. Americana C10:0-C16:0 specific thioesterase in a SAD locus

gctcttcgcctggagctggtgcagagcatggggcagtttgcggaggagagggtgctccccgtgctgcaccccgaggacaagctgt
ggcagccgcaggacttctgcccgaacccgagtcgcccacttcgaggaccaggtggcggagctgcgcgcgcgccaaggacct
gcccgcagagtactttgtggtgctggtgggcgacatgatcaaggaggaggcgctgacgacctacatggccatgctcaaaccttgg
acggtgtgcgcgacgacacgggcgcggtgaccacccgtgggcgcgctggacgcggcagtggtggcggaggagaacggcac
ggcgacctgtgaacaagtactgttggctgacggggcgctcaacatgcgggcccgtggaggtgacctcaacaacctgatcaag
agcggcatgaacccgcagacggacaacaacccctaattgggcttctctacacctcttccaggagcgcgccaccaagtaggtac
cctttcttgcgctatgacacttccagcaaaaggtaggcgggctgcgagacggcttccggcgctgcatgaacaccgatgatgctt
cgacccccgaagctccttcggggctgcatgggcgctccgatgccgtccaggcgagcgctgtttaaataggcaggccccgattg
caaagacattatagcgagctaccaaaagccatattcaaaccttagatcactaccatttctacacaggccactcgagcttgtgatcgc
actccgctaagggggcgcttcttcttctgttccagtcacaacccgcaaac**tctaga**aatatcaatatca**ATG**ctgctgcaggcctt
cctgttctgtgctggcggcttcgcccgaagatcagcgctccatgacgaacgagacgtccgacccgccccctggtgcaattcac
cccaacaagggctggatgaacgaccccaacggcctgtggtacgacgagaaggacgccaagtggcacctgtacttcagtac
aacccgaacgacacgtctgggggacgcccgtgttctggggccacgccaagtcgacgacctgaccaactgggaggaccagc
ccatcgccatcgccccgaagcgcaacgactccggcgcttctccggtccatggtggtggactacaacaacacctccggttctt
caacgacaccatcgacccgcgcccagcgctgcgtggccatctggacctacaacaccccgagtcgaggagcagtacatctct
acagcctggacggcggtacaccttcaccgagtaccagaagaacccgtgctggccgccaactccaccagttccgcgacccg
aaggtcttctgtacgagccctcccagaagtggatcatgaccgcccgaagtcccaggactacaagatcgagatctactctc
cgacgacctgaagtctggaagtggagtcggttcgccaacgaggggttctccggtaccagtagcagtgccccggcctga
tcgaggtccccaccgagcaggaccccagcaagtctactggtgatgttcatctccatcaaccccgcgccccggcgggcgt
ccttaaccagtagtctgtcgcgagcttaacggcaccacttcgaggccttcgacaaccagtcgccggtggtgacttcggca
ggactactacgcccgcagaccttcttaacaccgacccgacctacgggagcgccctgggcatcgctgggctccaactggg
agtactccgcttctgtgccaccaacccctggcgctcctcatgtcctcgtgcgcaagtctccctcaacaccgagtaccaggcc
aacccggagacggagctgatcaacctgaaggccgagccgatcctgaacatcagcaacgcccggccccctggagccggtcgc
accaacaccacgttgacgaaggccaacagctacaacgtcgacctgtccaacagcaccggcaccctggagttcgagctggtgt
acgccgtcaacacccaccagacgatctcaagtccgtgttcgcgacctctccctctggttcaagggcctggaggaccccgag
agtacctccgcatgggcttcgaggtgtccgctcctcttctctggaccgcccgaacagcaaggtgaagttcgtgaaggag
aaccctacttcaccaaccgcatgagcgtgaacaaccagccctcaagagcgagaacgacctgtcctactacaaggtgtacgg
cttctgtggaccagaacatcctggagctgtacttcaacgacggcgacgtcgtgtccaccaacacttactcatgaccacgggaa
cgccctgggctccgtgaacatgacgacgggggtggacaacctgttctacatcgacaagttccagggtgcgcgaggtcaag**TG**
Acaattgacggagcgtcgtgcgggagggagtgccgagcggggagtcgggtctgtgcgagggccggcagctgacgctggcga
gcccgtacgccccgaggggtccccctcccctgcaccccttcccctctgacggccgcccgtgttcttcatgttcagcgacggatc
ccgctctcgaacagagcgcgagaggaacgctgaaggctcgcctctgtgcacctcagcgcgccatacaccacaataaccacct
gacgaatgcgcttggttctctgctccattagcgaagcgtccggttcacacagctgccacgttggcgaggtggcaggtgacaatgatcg
gtggagctgatggctgaaacgttcacagcctagggatat**gaattc**ggccgacaggacgcgcgtcaaaaggtgctggtcgtgtat
gcccgtggccggcaggtcgttctgctgctggttagtgattccgcaaccctgattttggcgcttattttggcggtggcaaacgctgg
cgcccgcgagccgggcccggcgatgcggtgccccacggctgcgggaatccaaggagggaagagcgcccgggtcagtt
gaagggctttacgcgaaggtacagccgctcctgaaggctgcgtggtggaattggacgtgcaggtcctgctgaagttcctcc
accgctcaccagcggacaaagcaccggtgtatcaggtccgtgtcatccactctaagaactcgactacgacctactgatggc
cctagatttctcatcaaaaaacgctgagacacttgcacaggattgaaactccctgaagggaccaccaggggcccctgagttgttc
cttcccccggtggcgagctgccagccaggtgtacctgtgatcgaggctggcgggaaaataggcttcgtgtgctcaggtcatg
ggaggtgcaggacagctcatgaaacgccaacaatcgcaaatcatgtcaagtaaatcagctatttctcttcacgagctgta
attgtccaaaattctggtctaccgggggtgatccttctgtacgggcccctccctcaaccctaggtatgcgcgcatgcggtcgcc
gcgcaactcgcgaggggcccaggggttgggacgggcccgtcccgaatgcagttgcacccggatgcgtggcacctttttgcg
ataatttatcaatggactgctctgcaaaattctggctctgtgccaaacctaggatcagcggcgtaggatttcgtaattcattcg
tcctgatggggagctaccgactacctaataatcagcccagctgctgacgccagcgctccacttttgtgcacacattccattcgtgc
cgaagacatttcattgtggtgcgaagcgtccccagttacgctcacctgtttcccgaccttactgttctgtcgacagagcgggc

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

[illegible]

gctcttcgccgcccactcctgctcgagcgcccgcgctgcccagcgccctggcctttcgccgctcgtgcgctcgt
gatgtccatcaccaggtccatgaggtctgcttgcgcccgtgagccactgcttcgtccggcgcccaaggaggagcatgagggagg
actcctgggtccagggctcgtgacgtggctcgcgctctgggagcgggccagcatcatctggctctgccgaccgagggcgctccaact
ggctctccagcagccgagtcgcccggaccctggcagaggaagacaggtgaggggggtatgaattgtacagaacaaccacgag
ccttgtctaggcagaatccctaccagtcaggtttacctggatgacggcctgcgaacagctgtccagcgaccctcgtgccgcccgt
tctcccgacgcttcttccagcaccgtgatggcgagccagcgccgacgctggcgctgcgcttcgcccagatctgaggacagtcgg
ggaaactctgatcagctctaaaccccttgcgcgttagtggtgccatcctttgcagaccgggtgagagccgacttgttgcgccaccccc
acaccacctcctccagaccaattctgtcacctttttggcgaaggcatcgccctcgccctgcagagaggacagcagtgcccagccgc
tggggggttggcggtgacgctcaggtacccttctgcgctatgacacttcagcaaaaggtagggcgggctgcgagacggcttcc
cggcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctccgatgccgctccagggcga
gcgctgtttaatagccagggccccgattgcaaagacattatagcgagctaccaaagcatattcaaacacctagatcactaccactt
ctacacaggccactcgagcttgtgatcgactccgctaagggggcgcccttcccttcgtttcagtcacaacccgcaaac**ggcgcg**
catatcaATGtgtgtcagggccttctgttctgctggcggcttcgcccgaagatcagcgctccatgacgaacgagacgtc
cgaccgccccctgggtgacttcaccccccaacaagggttgatgaacgacccccacggcctgtggtacgacgagaaggacgcc
aagtggcacctgtacttcagtagcaacccgaacgacacccgtctgggggacgcccttgttctggggccacgcccagtcgcagca
cctgaccaactgggaggaccagccatcgccatcgccccgaagcgcaacgactccggcgcttctccggctccatggttggtg
actacaacaacacctccggttcttaacgacaccatcgacccgcccagcgctgctggtggccatctggacctacaacacccgg
agtccgaggagcagtagatctctacagcctggacggcggtacaccttcaccgagtagcagaagaaccccggtgctggccggc
aactccaccagttccgagaccgaaggtcttctggtacgagccctccagaagtgatcatgaccgcccgaagtcaccagga
ctacaagatcgagatctactctccgacgacctgaagtctggaagctggagtcgcttcgccaacgaggggttctccggtta
ccagtacgagtgccccggcctgatcgaggtccccaccgagcaggacccccgaagtcctactgggtgatgttcatctccatcaa
ccccggcgccccggcgccggtccttcaaccagtagtctcgtcgagcttcaacggcaccacttcgaggccttcgacaacca
gtcccgcgtgggtgacttcggcaaggactactacgcccgtcagaccttcttaacaccgaccgacctacgggagcgccctgg
gcatcgctgggctccaactgggagtagtccgcttcgtgccaccaacccctggcgctcctccatgtccctcgtgcgcaagttc
tccctcaaacaccgagtagcaggccaacccggagacggagctgatcaacctgaaggccgagccgatctgaacatcagcaac
gccggccccctggagccggttcgccaccaacaccacggttagcgaaggccaacagctacaacgtcgacctgtccaacagcacccg
gcacctggagttcgagctggtgtacgcccgaacaccaccagacgatctccaagtccgtgttcgaggaccttccctctggtt
caagggcctggaggacccccgaggagtagtcccgcatgggcttcgaggtgtccgctcctccttcttggtgaccgcggaaca
gcaaggtgaagttcgtgaaggagaacccctacttcaacaaccgcatgagcgtgaacaaccagccctcaagagcgagaacg
acctgtctactacaaggtgtacggcttgcgtggaccagaacatcctggagctgtacttcaacgacggcgacgtcgtgtccacca
aacctacttcatgaccacgggaacgcctgggctccgtgaacatgacgacgggggtggacaacctgttctacatcgacaag
ttccaggtgcgagaggtcaag**TGA**caattggcagcagcagctcggatagtagtcgacacactctggacgctggtcgtgtgatgga
ctgttggcccaacttgccttgacctgtgaatatccctgcgcttttatcaaacagcctcagtggtttgatcttgtgtgtacgc
gcttttgcgagttgtagctgttgccttatttgcgaataccacccccagcatcccttccctcgtttcatatcgcttgcatcccaaccg
caacttatctacgctgtcctgtatccctcagcgctgctcctgctcctgctcactgcccctcgacagccttggttgggctccgctgt
atttctcgtgtagtgaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaggatcc
cgctctcgaacagagcgcgagaggaacgctgaaggtctgcctctgtcgacactcagcgcgccatacaccacaataacca
cctgacgaatgcgcttggttcttctccattagcgaagcgtccggttcacacagtgccacgttggcgaggtggcaggtgacaa
tgatcggtggagctgatggtcgaacgttcacagcctagg**gatatcgaattc**ggcgctatgacacttcagcaaaaggtagggc
gggctgcgagacggcttccggcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatggcgctcc
gatgccgctccagggcgagcgctgtttaatagccagggccccgattgcaaagacattatagcgagctacaaagccatattcaa
cacctagatcactaccattctacacaggccactcgagcttgtatgcactccgctaagggggcgcccttctccttctcgtttcagtcg
gccgacaggacgcgctcaaaggtgctggtcgtgtatgcctggccggcaggtcgttgctgctgctggttagtgattccgaacct
gattttggcgcttattttggcgtggcaaacgctggcgcccgagccggcgccgagtgccgtgccccacggctgccggaatc
caagggaggcaagagcgccgggtcagttgaagggtttacgcgaaggtacagccgctcctgcaaggtcgtggtggaattgg
acgtgcaggtcctgctgaagttctccaccgctcaccagcggacaaagcaccggtgatcaggtcgtgtcatcactctaaagag

cagcccgactgcctgacgccagcgctccacttttgtcacacattccattcgtgcccaagacatttcatttggtgcaagcggtcccca
 gttacgctcacctgtttcccgacctcttactgttctgtcgacagagcgggccacagggccggtcgagcc**actagtATG**gccacc
 gcatccacttttctggcggttcaatgcccgtgcgggcagcctgctgctcgggcggtccgggccccggcgcccagcgaggccc
 ctccccgtgcgcg**ggcgcg**cccccaaggccaacggcagcgccgtgagcctgaagtccggcagcctgaacacctgagaggacc
 ccccccagcagcccccccccccgacacttctgaaccagctgcccgactggagccgctgcgacccgcatcaccacctgttctgt
 ggccgcccagagaagcagttcacccgctggaccgcaagagcaagcgccccgacatgctggtggaactggttcggcagcgagac
 catcgtgcaggacggcctggtgttccgagagcgttcagcatccgcagctacgagatcggcgcccagaccgaccgccaagcatcg
 agacctgatgaaccacctgcaggacaccagcctgaaccactgcaagagcgtgggctgtgaacgacggcttcggccgcac
 ccccgagatgtgcacccgagacatgctgggtgctgaccaagatgcagatcgtggtgaaccgctacccacctgtggggcgaca
 ccgtggagatcaacagctggttcagccagagcggaagatcgccatgggcccgcagtggtgatcagcgactgcaacaccg
 gcgagatcctggtgcgcgccaccagcgctgggcatgatgaaccagaagaccgcccgttcagcaagctgacctgcgaggt
 gcgcagagagatgcccccacttctgtggacgcccccccgctgatcgaggacaacgaccgcaagctgcacaagttcgacgtg
 aagaccggcgacagcatctgcaagggcctgacccccggtggaacgacttcgacgtgaaccgacagctgagcaacgtgaag
 tacatcggttgattctggagagcatgccaccgaggtgtgtagagaccagagctgtgcagcctgacctggagtaccgcc
 gcgagtgcgccgcgagagcgtggtggagagcgtgaccagcatgaacccagcaaggtgggagaccgagccagctacca
 gcacctgctgcgctggagagcgccgacatcatgaagggccgacccgagtggtgcgccccagaacgcccgcaccaaccg
 cgccatcagcacc**TGA**taattaactcgaggcagcagcagctcgtagtatcgacacactctggacgctggctggtgatggac
 tgttgcgcccaacttgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtggtttgatcttgtgtacgcg
 cttttgcgagttgctagctgcttgcctatttgcgaataccacccccagcatccccttccctcgtttcatatcgttgcacccaaccgc
 aacttatctacgctgctctgctatccctcagcgctgctcctgctcctgctcactgcccctcgacagccttgggttgggctccgctgta
 ttctcctggtactgcaacctgtaaacagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagctg
 agctcttgttttcagaaggagttgctccttgagcctttcattctagcctcgataacctccaaagccgctctaattgtggagggggttc
 gaatttaaaagcttgaatgttgggtcgtgctggaacaagcccagacttgttgcctactgggaaaaggaccatcagctccaaaa
 aacttgccgctcaaacggcgtacctctgctttcgcgcaatctgcctgttgaatcgccaccacattcatattgtgacgttgagcagt
 ctgaattgcctcagaatgtggaatcatctgccccctgtgcgagcccatgccaggcatgtcgccggcgaggacaccgccaactcgt
 cagcagaccattatgctacctcacaatagttcataacagtgaccatatttctgaagctcccaacgagcacctccatgctctgagtg
 gccacccccggcctggtgcttgcggagggcaggtcaaccggcatggggctaccgaaatccccgaccggatcccaccacccccgc
 gatgggaagaatctctccccgggatgtgggcccaccaccagcacaacctgctggcccaggcgagcgtcaaaccataccacacaaa
 tatccttggcatcgccctgaattccttctgcgctctgctaccgggtccttctgtccgaagcaggggttgcagggatcgctccgagt
 ccgcaaaccttgcgcgtggcggggcttgcgagcttgaagagc

SEQ ID NO: 277

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

ATGgccaccgcatccacttttctggcggttcaatgcccgtgcgggcagcctgctgctcgggcggtccgggccccggcgccca
 gcgagggccctccccgtgcg**tacgt**attccagtgcctggtggccagctgcatcgaccttgcgaccagtaccgcagcagcgcc
 agcctgagcttctggcgacaacggcttcgccagcctgttcggcagcaagcccttcagcaaccgcgccaccgcccgt
 gcgcgcgcagccacagcggcgagggccatggccgtggccctgcagcccgcagggaggccgacccaagaagaagcccg
 tgatcaagcagcgccgctggtggtgaccggcatggcggtggtgacccccctgggcccagagcccgcgtgttctacaaca
 cctgctggacggcgtgagcggcatcagcgagatcgagaccttcgactgcacccagttccccaccgcatcgccggcgagatca
 agagcttcagcaccgacggctgggtggccccaagctgagcaagcgcatggacaagttcatgctgtacctgtgaccgccgg
 caagaaggccctggccgacggcgccatcacccgacgaggtgatgaaggagctggacaagcgcaagtgcggcggtgctgatcg
 gcagcggcatggcgccatgaaggtgttaacgacgccatcgaggccctgcgcgtgagctacaagaagatgaacccctctg
 cgtgcccttcgccaccaccaacatgggcagcgccatgtggccatggacctgggctggatgggcccactacagcatcagca
 ccgctgcgccaccagcaacttctgcatcctgaacgcgcgaaccacatcatccgcggcgaggccgacatgatgctgtgcggc
 ggcagcgacgccgtgatccccatcgccctggcggttctgtggcctgccgcgcctgagccagcgcaacagcgacccccac
 caaggccagccgcccctgggacagcaaccgcgacggctcgtgatggcgagggcgccggcggtgctgctggtgaggagct

[illegible]

၁၆၅၆၇၈၉၁၀၁၁၁၂၁၃၁၄၁၅၁၆၁၇၁၈၁၉၂၀၂၁၂၂၂၃၂၄၂၅၂၆၂၇၂၈၂၉၃၀၃၁၃၂၃၃၃၄၃၅၃၆၃၇၃၈၃၉၄၀၄၁၄၂၄၃၄၄၄၅၄၆၄၇၄၈၄၉၅၀၅၁၅၂၅၃၅၄၅၅၅၆၅၇၅၈၅၉၆၀၆၁၆၂၆၃၆၄၆၅၆၆၆၇၆၈၆၉၇၀၇၁၇၂၇၃၇၄၇၅၇၆၇၇၇၈၇၉၈၀၈၁၈၂၈၃၈၄၈၅၈၆၈၇၈၈၈၉၉၀၉၁၉၂၉၃၉၄၉၅၉၆၉၇၉၈၉၉၁၀၀၁၀၂၀၃၀၄၀၅၀၆၀၇၀၈၀၉၁၀၁၁၁၂၁၃၁၄၁၅၁၆၁၇၁၈၁၉၂၀၂၁၂၂၂၃၂၄၂၅၂၆၂၇၂၈၂၉၃၀၃၁၃၂၃၃၃၄၃၅၃၆၃၇၃၈၃၉၄၀၄၁၄၂၄၃၄၄၄၅၄၆၄၇၄၈၄၉၅၀၅၁၅၂၅၃၅၄၅၅၅၆၅၇၅၈၅၉၆၀၆၁၆၂၆၃၆၄၆၅၆၆၆၇၆၈၆၉၇၀၇၁၇၂၇၃၇၄၇၅၇၆၇၇၇၈၇၉၈၀၈၁၈၂၈၃၈၄၈၅၈၆၈၇၈၈၈၉၉၀၉၁၉၂၉၃၉၄၉၅၉၆၉၇၉၈၉၉

SEQ ID NO: 280
KASII-allele_2_cDNA

ATGCAGACCGCGCACCCAGCGGGCCCCGACCGAGGGGCACTGCTTCGGTGCGAGGCTGCCACGGCGTC
GAGGCGGGCGGTGCGCCGGGCGTGGTCCCGCATCGCGCGCGCGGGCGGCCGCGGCCGACGCGACCCCCG
CCCGCCCTCCGCGCCGCGTGGTCTGTGACGGGCCAGGGCGTGGTGACCAGCCTGGGCCAGACGATCGAG
CAGTTTTACAGCAGCCTGCTGGAGGGCGTGAGCGGCATCTCGCAGATCCAAAAGTTTGACACCACGGG
CTACACGACGACGATCGCGGGCGAGATCAAGTCGCTGCAGCTGGACCCGTACGTGCCCAAGCGCTGGG
CCAAGCGCGTGGACGACGTCAAGTACGTCTACATCGCGGGCAAGCAGGCGCTGGAGAACCGGGGG
CTGCCGATCGAGGCGGGGGGCTGGCGGGCGCGGGGCTGGACCCGCGCTGTGCGGCGTGCTCATCGG
CACCGCCATGGCGGGCATGACGTCTTCGCGGGCGGGCGTGGAGGCGCTGACGCGCGGCGGCGTGCGCA
AGATGAACCCCTTTTGCATCCCTTCTCCATCTCCAACATGGGCGGCGCGATGCTGGCGATGGACATC
GGCTTCATGGGCCCCAACTACTCCATCTCCACGGCCTGCGCGACGGGCAACTACTGCATCCTGGGCGC
GGCGGACCACATCCGGCGCGGGCGACGCGGACGTGATGCTGGCCGGCGGCGCGGACGCGGCCATCATCC
CCTCGGGCATCGGCGGCTTCATCGCGTGCAAGGCGCTGAGCAAGCGCAACGACGAGCCCAGCGCGCG
AGCCGGCCCTGGGACGCCGACCGCGACGGCTTCGTATGGGCGAGGGCGCCGGCGTGCTGGTGTGGA
GGAGCTGGAGCACGCCAAGCGCCGCGGCGGCGGACCATCTGGCCGAATTCGTGCGGCGCGGGCCACCT
CGGACGCGCACCATGACCGAGCCGACCCGCGAGGGCCGCGGCGTGCGCCTCTGCCTCGAACGCGCG
CTCGAGCGCGCGCCTCGCGCCGAGCGCGTCTGGCTACGTCAACGCGCACGGCACAGCACGCCCCG
GGGCGACGTGGCCGAGTACCGCGCCATCCGCGCGTCTATCCCGCAGGACTCGCTGCGCATCAACTCCA
CCAAGTCCATGATCGGGCACCTGCTCGGCGGCGCGGCGCGGTGAGGCCGTGGCCGCCATCCAGGCC
CTGCGCACCGGCTGGCTCCACCCCAACCTCAACCTCGAGAACCCCGCACCCGGGTCGACCCCGTCTG
GCTCGTGGGGCCGCGCAAGGAGCGCGCCGAAGACCTCGACGTGCTCCTCTCCAACCTCCTTTGGCTTCG
CGGGCACAACTCGTGCGTCATCTTCCAAAAGTACGACGAGTGA

SEQ ID NO: 281
KASII-allele_1

MQTAHQRPPTGHCFCGARLPTASRRVRRAWSRIARAAAAADANPARPERRVITGQGVVTSLGQTIE
QFYSSLLEGVSGISQIQKFDTTGYTTTIIAGEIKSLQLDPYVPKRWAKRVDDVIKYVYIAGKQALESAG
LP IEAAGLAGAGLDPALCGVLIGTAMAGMTSFAAGVEALTRGGVRKMNPF CIPFSISNMGGAMLAMD I
GFMGPNYSISTACATGNYCILGAADHIRRGDANVMLAGGADAAIIPSGIGGFIACKALSKRNDEPERA
SRPWDADRDGFMGEGAGVVLVEELEHAKRRGATILAEVLVGAATSDAHHMTEPDQPQGRGVRLCLERA
LERARLAPERVGYVNAHGTSTPAGDVAEYRAIRAVIPQDSLRLNSTKSMIGHLLGGAGAVEAVAAIQAL
LRTGWLHPNLNLNENPAPGVDPVVLVGPRKERAEDLDVVLNSNFGFGGHNSCVIFRKYDE

SEQ ID NO: 282
KASII-allele_2

MQTAHQRPPTGHCFCGARLPTASRRVRRAWSRIARAAAAADATPARPPRRVVVTGQGVVT
SLGQTIEQFYSSLLEGVSGISQIQKFDTTGYTTTIIAGEIKSLQLDPYVPKRWAKRVDDVIKYVY
IAGKQALENAGLPIEAAGLAGAGLDPALCGVLIGTAMAGMTSFAAGVEALTRGGVRKMNPF
CIPFSISNMGGAMLAMDIGFMGPNYSISTACATGNYCILGAADHIRRGDADVMLAGGADAAI
IPSGIGGFIACKALSKRNDEPERASRPWDADRDGFMGEGAGVVLVEELEHAKRRGATILAEF
VGGAATSDAHHMTEPDQPQGRGVRLCLERALERARLAPERVGYVNAHGTSTPAGDVAEYRAI
RAVIPQDSLRLNSTKSMIGHLLGGAGAVEAVAAIQALRTGWLHPNLNLNENPAPGVDPVVLVG
PRKERAEDLDVVLNSNFGFGGHNSCVIFQKYDE

SEQ ID NO: 283
pSZ1503

gctcttccgcaccggctggctccacccaacttgaacctcgagaaccccgcgctggcgctgaccccgctgctgctggtggcgccg
ggaaggagcgcgccgaagacctggacgtcgctcctccaactcctttggctttggcgggcacaattcgcgctggtaccccttctt

gcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatgatgcttcgacccc
ccgaagctccttcggggctgcatggggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggccccgattgcaaaga
cattatagcgagctaccaaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgatcgactccg
ctaagggggcgctccttctcttcgttcagtcacaacccgcaaacggcgcgccATGctgctgcaggccttctgttctgctggcc
 ggcttcgcccgaagatcagcgctccatgacgaacgagacgtccgaccgccccctgggtgcacttcacccccaaacaagggctg
 gatgaacgacccccaacggcctgtggtacgacgagaaggacgccaagtggcacctgtacttccagtacaacccgaacgacac
 cgtctgggggacgcccctgttctggggccacgccacgtccgacgacctgaccaactgggaggaccagcccatcgccatcgccc
 cgaagcgcaacgactccggcgcttctccggctccatggtggtgactacaacaacacctccggcttcttaacgacaccatcg
 acccgccagcgctgcgtggccatctggacctacaacaccccgagtcaggaggagcagtacatctctacagcctggacggc
 ggctacacctcaccgagtaccagaagaaccccgctgctggccgcaactccaccagttccgcgacccgaaggcttctggtac
 gagccctcccagaagtggtatcatgaccgcgccaagtcccaggactacaagatcgagatctactcctccgacgacctgaagt
 ctggaagctggagtcgctgcgaacgagggcttctccggctaccagtacgagtgcggcgctgatcgaggctccccaccg
 agcaggaccccagcaagtctactgggtgatgttcatctccatcaaccccgcgccccggcgcggtccttcaaccagtactt
 cgtcggcagcttcaacggcaccacacttcgaggccttcgacaaccagtcccgctggtggacttcggcaaggactactacgcct
 gcagaccttcttaacaccgacccgacctacgggagcgccctgggcatcgctgggctccaactgggagtactccgccttctg
 gccaccaacccctggcgctctccatgtccctcgtgcgaagttctccctcaacaccgagtaccaggccaacccggagacgga
 gctgatcaacctgaaggccgagccgatcctgaacatcagcaacgcccggccctggagccggttcgccaccaacaccacgttg
 acgaaggccaacagctacaacgtcgacctgtccaacagcacggcacccctggagttcgagctggtgtacgccgtcaacacca
 ccagacgatctccaagtcggttcgaggacctctccctctggttaaggcctggaggaccccgaggagtacctccgcatgg
 gcttcgaggtgtccgctcctcttcttctggaccgcggaacagcaaggtgaagttcgtgaaggagaacccctacttcacca
 accgcatgagcgtgaacaaccagccctcaagagcgagaacgacctgtcctactacaaggtgtacggcttctggaccagaa
 catctggagctgtacttaacgacggcgacgtcgtgtccaccaacacacttctcatgaccaccgggaacgcctgggctccgt
 gaacatgacgacgggggtggacaacctgttctacatcgacaagttccaggtgcgagaggtcaagTGAcaattggcagcagc
agctcggatagtatcgacacactctggacgctggctgtgtaggactgttcggccacacttgctgccttgacctgtgaatatccctg
ccgcttttatcaaacagcctcagtggttgccttctgtgtgacgcgcttttgcgagttgctagctgcttgcctatttgcgaataaccac
cccagcatcccttccctcgttcatatcgcttgcacccaacccgcaactatctacgctgtcctgctatccctcagcgctgctcctgctc
ctgctcactgcccctcgacagccttgggttgggctccgctgtatttctcctggtactgcaacctgtaaaccagcactgcaatgctgat
gcacgggaagtagtgggatgggaacacaatggaggatcgtagagctcatcttccgaaagtacgacgagtgagcgagctgattct
ctttgagcggggtcgggtggttcggggagagtgcgcgaaaggcgacagacgtgcggcgccgctgtccctccgtcttccctgg
ttggtgctatagtaacctgcctgtgtcgtgcgctcggaagagc

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	AGTGAATTTC	GCGAATGTCG	GCTCACC CGC	ACGCGATTAC	GAATCCGCTA
ACTCTAACAC	TTTGCTGCGG	GGGATGTGCA	GGTCACATCA	ACCACTCCCT	CTTCTGGAAG
AACCTCGCCC	CAGCCAAGTC	TGAGGGCAAG	GGAAACGGCG	GCGCCATCAA	GGACGGCCCC
CTCAAGTCGG	CGATCGAGCA	GAAC TTTGGT	TCCGTCGACG	TCTTCATCAA	GGAGTTCAAC
GCGACCACCG	CGTCCATCCA	GGGCTCCGGC	TGGGGCTGGC	TCGTGAGTCT	TCTCTTCTGC
GATCGGAAGG	CGCATCCGGT	GTTATGGTCG	CGCGCTGCAC	GCGTGTCTCG	CTTCGGCTCC
GCCGTCGAAG	CGCTGATACC	GAGACCGATC	TGCTGCAAAT	CATT TGC GCT	GCGATAGAGC
TCCGGGGGCT	TTAACCTTTC	GTTCCATGTG	ATAAGACCGC	GTCTGACCGG	GTGGTGT TTT
CTATCGCAGG	GTCTCAACCC	GGCGACCAAG	CGCCTCGAGA	TCACGACGAC	GGCCAACCAG
GACCTCTCC	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 CCAGAAAGTGG GAAATCTTCC
CCATGAAATA ACGGATCCTC TTGCTACTGC TACTACTACT ACTACAAACT GTAGCCATTT
ATATAATTCT ATATAATTTT CAACATGGCC ACCACCTCTT TAGCTTCCGC TTTCTGCTCG
ATGAAAGCTG TAATGTTGGC TCGTGATGGC CGGGGCATGA AACCCAGGAG CAGTGATTTG
CAGCTGAGGG CGGGAAATGC GCCAACCTCT TTGAAGATGA TCAATGGGAC CAAGTTCAGT
TACACGGAGA GCTTGAAAAG GTTGCCTGAC TGGAGCATGC TCTTTGCAGT GATCACAACC
ATCTTTTCGG CTGCTGAGAA GCAGTGGACC AATCTAGAGT GGAAGCCGAA GCCGAAGCTA
CCCCAGTTGC TTGATGACCA TTTTGAGACT 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 TGGTGCATCT
GGAAATAATG GCATGCGACG TGATTTCCCT GTCCGGGACT GCAAAACAGG CGAAATTCTT
ACAAGATGTA CCAGCCTTTC GGTGCTGATG AATACAAGGA CAAGGAGGTT GTCCACAATC
CCTGACGAAG TTAGAGGGGA GATAGGGCCT GCATTCATTG ATAATGTGGC TGTCAAGGAC
GATGAAATTA AGAACTACA 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 TGTGTAAGTT
TCTCCCATGC TGTTCTGTGAG GATACTTTTT AGAAGCTGCA GTTTGCATTG CTTGTGCAGA
ATCATGGTCT GTGGTTTTAG ATGTATATAA AAAATAGTCC TGTAGTCATG AAACCTAATA
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 TTVFVAAEKQ WTMLDRKSKR
PDTLMDPFGV DRVVQDGVVF RQSFIRSIE IGADRTASIE TLMNIFQETS LNHCKSIGLL
NDGFGRTPEM CKRDLIWVVT KMHIEVNRYP TWGDTIEVNT WVSESGKTGM GRDWLISDCH
TGEILIRATS VCAMMNQTTR RFSKFPYEVF QELAPHFVDS APVIEDYQKL HKLDVKTGDS
ICNGLTPRWN 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 DWSMLLAIT
TLFLAAEKQW MMLDWKPKRP DMLVDPFGLG RFVQDGLVFR NNFSIRSIEI GADRTASIEI
LMNHLQETAL NHVKSGLLE DGLGSTREMS LRNLIVVVTQ MQVAVDRYPT WGDEVQVSSW
ATAIGKNGMR REWIVTDFRT GETLLRATSV WVMNKLTRR ISKIPPEVWH EIGPSFIDAP
PLPTVEDDGR KLTRFDESSA DFIRXGLTPR WSDLDINQHV NNVKYIGWLL ESAPPEIHES

```

HEIASLTLEY RRECGRDSVL NSATKVSDSS QLGKSAVECN HLVRLQNGGE IVKGRTVWRP
KRPLYNDGAV VDVXAKTS