

SEQUENCE LISTING

SEQ ID NO: 1

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

gatcagacgggcctgacctgcgagataatcaagtgtcgtaggcaaccaactcagcagctgcttggtgtgggtctgcaggatagtgtt
gcaggggcccaaggacagcaggggaacttacacctgtccccgaccagtttatggagtgcattgcctcaagagcctagccggagc
gctaggctacatacttcccgcaccgggtatgaggggatatagctcgcactgcgtgtctagttagatgggcagtgctgccataaac
aactggctgtcagccatttgttggcggaccattctggggggccagcaatgcctgactttcgggtagggtgaaaactgaacaaagac
taccaaacagaatttctcctccttggaggtaagcgcaggccggcccgctgcgcccacatggcgctccgaacacctccatagctgt
aagggcgcaaacatggccggactgtgtcagcactcttcatggccatacaaggctcatgtcagattagtgtgagtaagacactatca
ccccatgttcgattgaagccgtgacttcatgcaacctgccccggcgtagcagacgtatgccatcatgacctagccgacatgcg
ctgtcttttgcacaaaacaactgttacccgctcgaagtcgtgccgcacacctccgggagttagtccggcgactcctccccggcg
ggccgcggccctacctgggtagggtcgcatacggccagacaaacgacgcaggagggttggggtagggaatcccaaccag
cctaaccaagacggcacctataataataggtggggggactaacgacctatcgaagcttgggtgcctatcttgagaagcacgag
ttggagtggctgtgtacggcgtgacctaaaggtgggtgtgccgcagcctgaaacaaagcgtctagcagctgcttctataatgtgtcagcc
gttgtgttcagttatattgtatgctattgtttgtcgtctagggtggcgagggccacctactgtggcgggccattgggtggtgcttgaatt
gcctcaccatctaaggtctgaacgtcactcaaacgcctttgtacaactgcagaactttccttggcgctgcaactacagtggtgcaacca
gcacatagcactccctacatcacccagcagtacaaca

SEQ ID NO: 2

Chlorella ellipsoidea nitrate reductase promoter from AY307383

cgctgcgcaccaggcccgccagctcgtgatgtcgtccaaatgcgggtcccccgaattttgttcttcatcttctccaccttggtggccttc
ttggccaggggccttcagctgcatgcgcacagaccgttgagctcctgatcagcctcagcagggcccttgacaagcaagccctgtg
caagcccaatcacgggggtaccagtggtgtgaggtagatgggtttgaaaaggattgctcggtcgattgctgctcatggaattggcatgt
gcatgcatgttcacaatatgccaccaggcttggagcaagagagcatgaatgccttcaggcaggttgaagttcctgggggtgaagag
gcaggggccgaggattggaggaggaaagcalcaagtcgtcgtcatgctcatgttttcagtcagagtttgccaagctcacaggagcag
agacaagactggctgctcaggtgttcacgtgtgtgtgtggggggggggggttaatacggtagcaaatgcaattggaattccac
ctcatgccagcggaccacatgcttgaattcgaggcctgtgggggtgagaaatgctcactctgcctcgttgcgtgaggtacttcaggccg
ctgagctcaaagtcgatgccctgctcgtatcagggcctgcacctctgggctgaccggctcagcctccttcgcgggcatggagtagg
cgccggcagcgttcgtccggggccagggcagcgggtgtgccataaatgtcggatggtggggaggggggccgtcggcacacc
attgccgttgcgtgacgcatgcacatgtggcctggctggcaccggcagcactggtctccagccagccagcaagtggctgttcag
gaaagcgcccatgttgttggccctgcgcgtgtaattccccagatcaaggagggaacagcttggatttgatgtagtcccaaccgga
ctgaatgtgcgatggcaggtccctttagtctccgaattactagcagggcactgtgacctaacgcagcatgccaaccgcaaaaaaat
gattgacagaaaaatgaagcgggtgtgtcaatattgtctatttattcgttttaatcagcaaccaagttcgaacgcaactatcgtggtgatca
agtgaacctcatcagacttacctcgttcggcaaggaaacggaggcaccaaattccaatttgatattatcgcttgccaagctagagctgat
cttgggaaaccaactgccagacagtggactgtgatggagtccccgagtggtggagcctcttcgattcgggttagtcattactaacgtg
aacctcagtggaaggaccatcagaccagaaagaccagatctcctcctcgacaccgagagagtgttgcggcagtaggacgacaag

SEQ ID NO: 3

Yeast sucrose invertase

MTNETSDRPLVHFTPKNKGWMNDPNGLWYDEKDAKWHLVYFQYNPNDTVWGTPLF
WGHATSDDL TNWEDQPIAIAPKRND SGAFSGSMVVDYNN TSGFFNDTIDPRQRCVAI
WTYNTPESEEQYISYSLDGGYTFT EYQKNPVLAANSTQFRDPKVF WYEPSQKWIMT
AAKSQDYKIEIYSSDDLKSWKLES AFANEGFLGYQYECPLIEVPTEQDPSKSYWVM
FISINPGAPAGGSFNQYFVGSFNGTHFEAFDNQSRVVD FGDYALQTFNTDPTYGS
ALGIAWASNWEYS AFVPTNPWRSSMSLVRKFSLNTEYQANPETELINLKAEPILNISN
AGPWSRFATNTTLTKANSYNVDLSNSTGTLEFELVYAVNTTQTISKSVFADLSLWFK

GLEDP EEYL R MGF EVS ASSFF LDR GNSKVKFVKENPYFTNRMSVNNQPFKSENDLSY
YKVYGLLDQNILELYFNDGDV VSTNTYFMTTGNALGSVNMTTGVDNLFYIDKFQVR
EVK

SEQ ID NO: 4

Yeast secretion signal

MLLQAFLFLLAGFAAKISAS

SEQ ID NO: 5

Higher plants secretion signal

MANKSLLLLLLLLGSLASG

SEQ ID NO: 6

Consensus eukaryotic secretion signal

MARLPLAALG

SEQ ID NO: 7

Combination higher plant/eukaryotic secretion signal

MANKLLLLLLLLLLLLPLAASG

SEQ ID NO: 8

S. cerevisiae sucrose invertase NP_012104

gaattccccaacatgggtggagcagcacactctcgtctactccaagaatatcaaagatacagtcacagaagaccaaagggtattgaga
ctttcaacaaagggttaataatcggaacacctcctcggttccattgccagctatctgtcacttcatcaaaaggacagtagaaaaggag
gtggcacctacaaatgccatcattgcgataaaggaaaggctatcgttcaagatgcctctgccgacagtggtcccaaagatggaccccc
accacgaggagcatcgtggaaaaagaagacgttccaaccacgtctcaaagcaagtggattgatgtgaacatgggtggagcagcagc
actctcgtctactccaagaatatcaaagatacagtcacagaagaccaaagggtattgagactttcaacaaagggttaataatcggaac
ctcctcgggttccattgccagctatctgtcacttcatcaaaaggacagtagaaaaggaggtggcacctacaaatgccatcattgcgat
aaaggaaagggtatcgttcaagatgcctctgccgacagtggtcccaaagatggacccccaccacgaggagcatcgtggaaaaaga
agacgttccaaccacgtctcaaagcaagtggattgatgtgatctccactgacgtaagggaatgacgcacaatcccaatccttgcga
agaccttctctatataaggaggtcatttcatttggagaggacacgctgaaatcaccagtctctctacaaatctatctctggcgcgc
atacaATGCTTCTTCAGGCCTTTCTTTTCTTCTTGCTGGTTTTGCTGCCAAGATCAG
CGCCTCTATGACGAACGAAACCTCGGATAGACCACTTGTGCACTTTACACCAAAC
AAGGGCTGGATGAATGACCCCAATGGACTGTGGTACGACGAAAAAGATGCCAAG
TGGCATCTGTACTTTCAATACAACCCGAACGATACTGTCTGGGGGACGCCATTGT
TTTGGGGCCACGCCACGTCCGACGACCTGACCAATTGGGAGGACCAACCAATAG
CTATCGCTCCGAAGAGGAACGACTCCGGAGCATTCTCGGGTTCCATGGTGGTTGA
CTACAACAATACTTCCGGCTTTTTCACGATACCATTGACCCGAGACAACGCTGC
GTGGCCATATGGACTTACAACACACCGGAGTCCGAGGAGCAGTACATCTCGTAT
AGCCTGGACGGTGGATACACTTTTACAGAGTATCAGAAGAACCCTGTGCTTGCTG
CAAATTCGACTCAGTTCCGAGATCCGAAGGTCTTTTGGTACGAGCCCTCGCAGAA
GTGGATCATGACAGCGGCAAAGTCAACAGGACTACAAGATCGAAATTTACTCGTC
TGACGACCTTAAATCCTGGAAGCTCGAATCCGCGTTTCGCAAACGAGGGCTTTCTC
GGCTACCAATACGAATGCCAGGCCTGATAGAGGTCCCAACAGAGCAAGATCCC
AGCAAGTCCTACTGGGTGATGTTTATTTCCATTAATCCAGGAGCACCGGCAGGAG
GTTCTTTTAATCAGTACTTCGTCCGAAGCTTTAACGGAACTCATTTTCGAGGCATTT
GATAACCAATCAAGAGTAGTTGATTTTGGAAGGACTACTATGCCCTGCAGACTT

TCTTCAATACTGACCCGACCTATGGGAGCGCTCTTGGCATTGCGTGGGCTTCTAA
CTGGGAGTATTCCGCATTTCGTTCTACAAACCCCTTGGAGGTCTCCATGTGCTC
GTGAGGAAATTCTCTCTCAACACTGAGTACCAGGCCAACCCGGAACCGAACTC
ATAAACCTGAAAGCCGAACCGATCCTGAACAATTAGCAACGCTGGCCCCCTGGAGC
CGGTTTGCAACCAACACCACGTTGACGAAAGCCAACAGCTACAACGTCGATCTTT
CGAATAGCACCAGGTACACTTGAATTTGAACTGGTGTATGCCGTCAATACCACCCA
AACGATCTCGAAGTCGGTGTTCGCGGACCTCTCCCTCTGGTTTAAAGGCCTGGAA
GACCCCGAGGAGTACCTCAGAATGGGTTTCGAGGTTTCTGCGTCCTCCTTCTTCCT
TGATCGCGGGAACAGCAAAGTAAAATTTGTTAAGGAGAACCCATATTTTACCAA
CAGGATGAGCGTTAACAACCAACCATTCAAGAGCGAAAACGACCTGTGCTACTA
CAAAGTGTATGGTTTGCTTGATCAAAATATCCTGGAACCTCTACTTCAACGATGGT
GATGTGCTGTCCACCAACACATACTTCATGACAACCGGGAACGCACTGGGCTCCG
TGAACATGACGACGGGTGTGGATAACCTGTTCTACATCGACAAATTCCAGGTGA
GGGAAGTCAAGTGAgaatctgcatcgacaagctcgagtttctccataataatgtgtgagtagtcccagataaggaatta
gggttcctatagggttcgctcatgtgtgagcatataagaacccttagtatgtattgtattgtaaaatacttctatcaataaaatttctaatt
cctaaaacaaaatccagtactaaaatccagatccccgaalttaa

SEQ ID NO: 9

TGTTGAAGAATGAGCCGGCGAC

SEQ ID NO: 10

CAGTGAGCTATTACGCACTC

SEQ ID NO: 11

UTEX 329 *Prototheca kruegani*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGAAAATACT
CTGGAGCCATAGCGAAAGCAAGTTTAGTAAGCTTAGGTCAATTCITITTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTAGTAAATAACATGGAG
GCCCCGAACCGACTAATGTTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGGCGCAGC
AGTAGCAGTACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAGGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGACTAAGGATGTGGGTATGTCAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 12

UTEX 1440 *Prototheca wickerhamii*

TGTTGAAGAATGAGCCGGCGACTTAAAATAAATGGCAGGCTAAGAGATTTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCTTAATAGGGCGTCAATTTAACAAAACCTT
TAAATAAATTATAAAGTCAATTTATTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC
TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACAGTTAAGGCCCCAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 13UTEX 1442 *Prototheca stagnora*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAAATGGCATGGTTAAAGATAATTTCT
CTGAAGCCATAGCGAAAGCAAGTTTTACAAGCTATAGTCATTTTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTCAAATAACATGGAG
GCCCCAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAAGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGAATAAGGACGTGAGTATGTCAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 14UTEX 288 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGATAATTCT
CTGGAGCCATAGCGAAAGCAAGTTTAAACAAGCTAAAGTCACCCTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG
GCCCCAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTTGTGGGCTTCGAAAGT
TGTACCTCAAAGTGGCAAACCTCTGAATACTCTATTTAGATATCTACTAGTGAGAC
CTTGGGGGATAAGCTCCTTGGTCAAAAAGGGAAACAGCCCAGATCACCAGTTAAG
GCCCCAAAATGAAAATGATAGTGAATAAGGATGTGGGTATGTTAAAACCTCCAG
CAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 15UTEX 1439, UTEX 1441, UTEX 1435, UTEX 1437 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGACTTAAATAAAATGGCAGGCTAAGAGAATTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCTTAATAGGGCGCTAATTTAACAAAACAT
TAAATAAAATCTAAAGTCATTTATTTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC
TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACAGTTAAGGCCCAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 16UTEX 1533 *Prototheca wickerhamii*

TGTTGAAGAATGAGCCGTCGACTTAAATAAAATGGCAGGCTAAGAGAATTAATA
ACTCGAAACCTAAGCGAAAGCAAGTCTTAATAGGGCGCTAATTTAACAAAACAT
TAAATAAAATCTAAAGTCATTTATTTTAGACCCGAACCTGAGTGATCTAACCATG
GTCAGGATGAAACTTGGGTGACACCAAGTGGAAGTCCGAACCGACCGATGTTGA
AAAATCGGCGGATGAACTGTGGTTAGTGGTGAAATACCAGTCGAACTCAGAGCT
AGCTGGTTCTCCCCGAAATGCGTTGAGGCGCAGCAATATATCTCGTCTATCTAGG
GGTAAAGCACTGTTTCGGTGCGGGCTATGAAAATGGTACCAAATCGTGGCAAAC

TCTGAATACTAGAAATGACGATATATTAGTGAGACTATGGGGGATAAGCTCCAT
AGTCGAGAGGGAAACAGCCCAGACCACCAGTTAAGGCCCAAAATGATAATGAA
GTGGTAAAGGAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGC
CATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 17

UTEX 1434 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCGTGGTTAAAGAAAATTCT
CTGGAACCATAGCGAAAGCAAGTTTAAACAAGCTTAAGTCACTTTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG
GCCCCAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTGTGGGCTCCGAAAG
TTGTACCTCAAAGTGGCAAACTCTGAATACTCTATTTAGATATCTACTAGTGAGA
CCTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAA
GGCCCCAAAATGAAAATGATAGTGACTAAGGATGTGAGTATGTCAAAACCTCCA
GCAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCCTAATAGCTCACTG

SEQ ID NO: 18

UTEX 1438 *Prototheca zopfii*

TGTTGAAGAATGAGCCGGCGAGTTAAAAAGAGTGGCATGGTTAAAGAAAATTCT
CTGGAGCCATAGCGAAAGCAAGTTTAAACAAGCTTAAGTCACTTTTTTTAGACCCG
AAACCGAGTGATCTACCCATGATCAGGGTGAAGTGTTGGTAAAATAACATGGAG
GCCCCAACCGACTAATGGTGAAAAATTAGCGGATGAATTGTGGGTAGGGGCGAA
AAACCAATCGAACTCGGAGTTAGCTGGTTCTCCCCGAAATGCGTTTAGGCGCAGC
AGTAGCAACACAAATAGAGGGGTAAAGCACTGTTTCTTTCTGTGGGCTTCGAAAG
TTGTACCTCAAAGTGGCAAACTCTGAATACTCTATTTAGATATCTACTAGTGAGA
CCTTGGGGGATAAGCTCCTTGGTCAAAAGGGAAACAGCCCAGATCACCAGTTAA
GGCCCCAAAATGAAAATGATAGTGACTAAGGATGTGAGTATGTCAAAACCTCCA
GCAGGTTAGCTTAGAAGCAGCAATCCTTTCAAGAGTGCCTAATAGCTCACTG

SEQ ID NO: 19

UTEX 1436 *Prototheca moriformis*

TGTTGAAGAATGAGCCGGCGACTTAGAAAAGGTGGCATGGTTAAGGAAATATTC
CGAAGCCGTAGCAAAAGCGAGTCTGAATAGGGCGATAAAATATATTAATATTTA
GAATCTAGTCATTTTTTCTAGACCCGAACCCGGGTGATCTAACCATGACCAGGAT
GAAGCTTGGGTGATACCAAGTGAAGGTCCGAACCGACCGATGTTGAAAAATCGG
CGGATGAGTTGTGGTTAGCGGTGAAATACCAGTCGAACCCGGAGCTAGCTGGTT
CTCCCCGAAATGCGTTGAGGCGCAGCAGTACATCTAGTCTATCTAGGGGTAAAGC
ACTGTTTTCGGTGCGGGCTGTGAGAACGGTACCAAATCGTGGCAAACCTCTGAATAC
TAGAAATGACGATGTAGTAGTGAGACTGTGGGGGATAAGCTCCATTGTCAAGAG
GGAAACAGCCCAGACCACCAGCTAAGGCCCAAAATGGTAATGTAGTGACAAAG
GAGGTGAAAATGCAAATACAACCAGGAGGTTGGCTTAGAAGCAGCCATCCTTTA
AAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 20

Chicorium intybus invertase: Genbank Accession No. Y11124

MSNSSNASESLFPATSEQPYRTAFHFQPPQNWMNDPNGPMCYNGVYHLFYQYNPFG
PI.WNLRMWYWAHSVSHDLINWIHLDLAFAPTEPFDINGCLSGSATVLPGNKPIMLYTG

IDTENRQVQNLA VPKDLSDPYLREWVKHTGNPIISLP EEIQPDDFRDPTTTWLEEDGT
WRLLVGSQKDKTGIAFLYHSGDFVNWTKSDSPLHKVSGTGMWECVDFFPVWVDST
NGVDTSIINPSNRVKHVLKLG IQDHGKDCYLIGKYSADKENYVPEDELTSTLR LDY
GMY YASKSFFDPVKNR RIMTAWVNESDSEADVIARGWSGVQSFP RSLWLDKNQKQ
LLQWP IEEIEM LHQNEVSFHNKKLDGGSSLEV LGITASQADV KISFKLANLEEAEELD
PSWVDPQLICSENDASKKGKFGPFGLLALASSDLREQT AIFFRVFRKNGRYVVL MCS
DQSRSSMKNGIEKR TYGAFVDIDPQQDEISLRTLIDHSIVESFGGRGKTCITTRVYPTL
AIGE QARLFAFNHGTESVEISELSAWSMKKAQMKVEEP

SEQ ID NO: 21

Schizosaccharomyces pombe Invertase: Genbank Accession No. AB011433

MFLKYILASGICLV SLLSSTNAAPRHL YVKRYPVIYNASNITEVSNSTTVPPPPFVNTT
APNGTCLGNY NEYLPSGYYNATDRPKIHFTPSSGFMNDPNGLVYTGGVYHMFQYS
PKTLTAGEVHWGHTVSKDLIHWENYPIAIYPDEHENGVL SLPFSGSAVVDVHNSSGL
FSNDTIPEERIVLIYTDHWTGVAERQAIA YTTDGGYTFFKKYSGNPVLDINSLQFRDPK
VIWDFDANRWVMIVAMSQNYGIAFYSSYDLIHWTELSVFSTSGYLGLQYECPGMAR
VPVEGTDEYK WVLFI SINPGAPLGGSVVQYFVG DWNGTNFVPDDGQTRFVDLGKDF
YASALYHSSSANADVIGVGWASNWQYTNQAPTQVFRSAMTVARKFTLRDVPQNPM
TNLTSLIQ TPLNV SLLRDETLFTAPVINSSSSLSGSPITLPSNTAFEFNVTL SINYTEGCT
TGYCLGR IIIIDSDDPYRLQSISVDVDFAASTLVINRAKAQMGWFNSLFTPSFANDIYIY
GNVTLYGIVDNGLLELYVNNGEKTYTNDFFFLQGATPGQISFAAFQGV SFNNVTVTP
LKTIWNC

SEQ ID NO: 22

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

MIQLSPLLLLPLFSVFNSIADASTEYLRPQIHLTPDQGWMNDPNGMFYDRKDKLWHV
YFQHNPDKKSIWATPVTWGHSTSKDLLTW DYHGNAL EPENDD EGIFSGSVVDRNN
TSGFFNDSTDPEQRIVAIYTNN AQLQTQEIAYS LDKGYSFIKYDQNPVINVNSSQQRD
PKVLWHDESNQWIMVVAKTQEFKVQIYGSPDLKKWDLKSNFTSNGYLG FQYECPG
LFLKLP IENPLNDVTTSKWVLLLAINPGSPLGGSINEYFIGDFDGTTFHPDDGATRFMDI
GKDFYAFQSF DNTEPEDGALGLAWASNWQYANTVPTENWRSSMSLVRNYTLKYVD
VNPENYGLTLIQKPVYDTKETRLNETLKTLETINEYEVNDLKL DKSSFVATDFNTERN
ATGVFEFDLKTQTDLKMGYSNM TTFGLYIHSQTVKGSQETLQLVFDTLSTTWYID
RTTQHSFQRNSPVFTERISTYVEKIDTTDQGNVYTLYGVVDRNILELYFNDGSIAMTN
TFFFREGKIPTSFEVVCDSEKSFTITIDELSVRELARK

SEQ ID NO: 23

Debaryomyces occidentalis Invertase: Genbank Accession No. X17604

MVQVLSVLVIPLLT LFFGYVASSSIDLSVDTSEYNRPLIHFTPEKGWMNDPNGLFYDK
TAKLWHLYFQYNPNATAWGQPLYWGHATSNDLVHWDEHEIAIGPEHDNEGIFSGSI
VVDHNNTSGFFNSSIDPNQRIVAIYTNNIPDLQTQDIAFSLDGGYTFTKYENNPVIDVS
SNQFRDPKVFWHERFKSMDHGCSEIARVKIQIFGSANLKNWVLNSNFSSGYGNYQY
GMSRLIEVPIENS DKSKWVMFLAINPGSPLGGSINQYFVGDFDGFQFVPDDSQTRFVD
IGKDFYAFQTFSEVEHGV LGLAWASNWQYADQVPTNPWRSSSTSLARNYTLRYVIQM
LKL TANIDKSVLPDSINVVDK LKKKNVKL TNKKPIKTNFKGSTGLFDFNITFKVLNLN
VSPGKTHFDILINSQELNSSVDSIKIGFDSSQSLFYIDRHIPNVEFPRKQFFTDKLAAYL

EPLDYDQDLRVFSLYGIVDKNIIELYFNDGTVAMTNTFFMGEGKYPHDIQIVTDTEEP
LFELESVIIRELNK

SEQ ID NO: 24

Oryza sativa Invertase: Genbank Accession No. AF019113

MATSRLTPAYDLKNAAAAVYTPLPEQPHSAEVEIRD RKPFKIISAILSSLLLLLALILVA
VNYQAPPSHSSGDN SQPAAVMPPSRGVSQGVSEKA FRGASGAGNGVSFAWSNLMLS
WQRTSYHFQPVKNWMNDPNGPLYKGYHLYFYQYNPDSAVWGNITWGHAVSTD
LINWLHLPFAMVPDQWYDVNGVWTGSATILPDGRIVMLYTGDTDDYVQDQNLAFP
ANLSDLLVDWVKYPNNPVIYPPPGIGVKDFRDPTTAGTAGMQNGQRLVTIGSKVG
KTGISLVYETTNFTTFKLLYGVLHAVPGTGMWECVDLYPVSTTGENGLDTSVNLG
VKHVLKTSLDDDKHDYYALGTYPVKNKWTPDNPDLVDVGIGLRIDYGKYAARTF
YDQNKQRRILWG WIGETDLEAVDLMKGWASLQAIPRTIVFDKKTGTNVLQRPEEEV
ESWSSGDPITQRRIFEPGSVVIHVSGATQLDITASFEVDETLETTSESH DAGYDCSN
SGGAGTRGSLGPFGLLVVADEKLSFLTPVYLYVAKGGDGKAKAHLCA YQTRSSMAS
GVEKEVYGSAPVLDGENYSARILIDHSIVESFAQAGRTCVR SRDYPTKDIYGAARCF
FFNNATEASVRASLKA WQMKSFIRYPPIPDQKS

SEQ ID NO: 25

Allium cepa Invertase: Genbank Accession No. AJ006067

MSSDDI.ESPSSYLPIPPSDEFHDQPPPLRSWLRLLSIPLALMFLFLATFLSNLESPPSD
SGLVSDPVTFDVNPAVVRGKDAGVSDKTSGVDSGFVLDPVAVDANSVVVHRGKD
AGVSDKTSGVDSGLLKDSPLGPYPWTNQMLSWQRTGFHFQPVKNWMNDPNGPLY
KGYWHFFYQYNPEGAVWGNIWGHAVSRDLVHWTHLPLAMVPDQWYDINGVWT
GSATILPDGQIVMLYTGATNESVQVQNLA VPADQSDTLLLRWKKSEANPILVPPPGIG
DKDFRDPTTAWYEPSDDTW RIVIGSKDSSHSGIAIVYSTKDFINYKLIPGILHAVERVG
MWECVDFYPVATADSSHANHGLDPSARPSPAVKHVLKASMDDDRHIDYYAIGTYDP
AQNTWVPDDASVDVGIGLRYDWGKFYASKTFYDHAKKRRILWSWIGETDSETADIA
KGWASLQGVPRTVLLDVKTGSNLITWPVVEIESLRTRPRDFSGITVDAGSTFKLDVG
GAAQLDIEAEFKISSELEAVKEADVSYNCS SSGGAAERGVLPFGLLVLANQDLTE
QTATYFYVSRGMDGGLNTHFCQDEKRSSKASDIVKRIVGHSVPVLDGESFALRILVD
HSIVESFAQGGRASATSRVYPT EAIYNNARVVFVNNATGAKVTAQSLKVWHMSTAI
NEIYDPATSV M

SEQ ID NO: 26

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

LFYQYNPNGVIWGPVWGHSTSKDLVNWVPQPLTMEPEMAANINGSWSGSATILPG
NKPAILFTGLDPKYEQVQVLA YPKDTS DPNLKEWFLAPQNPVMFPTPQNQINATSFR
DPTTAWRLPDGVWRLLIGSKRGQRGLSLLFRSRDFVHWVQAKHPLYSDKLSGMWE
CPDFFPVYANGDQMGVDTSIIGSHVKHVLKNSLDITKHDIYTIGDYNIKKDAYTPDIG
YMNDSSLRYDYGYASKTFFDDAKKERILLGWANESSSVEDDIKKGWSGIHTIPRK
IWLDKLGKQLIQWP IANIEKL RQKPVNIYRKVLKGGSQIEVSGITAAQADVEISFKIKD
LKNVEKFDASWTS PQLLCSKKGASVKGGLGPFGLLTLASXGLEEYTA VFFRIFKAYD
NKFVVL MCDQSRSSI.NPTNDKTTYGT FVDVNPIREGLSLRVLIDHSVVESFGAKGK
NVITARVYPTLAINEKAHLYVFNRGTSNVEITGLTAWSMKKANIA

SEQ ID NO: 27

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

MTDFTPETPVLTPIRDHAAELAKAEAGVAEMAANKRNNRWYPKYHIASNGGWINDPN
GLCFYKGRWHVIFYQLHPYGTQWGPMIHWGHVSSTDMLNWKREPIMFAPSLEQEKD
GVFSGSAVIDDNGDLRFYYTGHRWANGHDNTGGDWQVQMTALPDNDELTSATKQ
GMIIDCPTDKVDHHRDPKPVWKTGDTWYMTFGVSSSEDKRGQMWFSSKDMVRWE
YERVLQHPDPDVFMLECPDFFPIKDKDGNEKWVIGFSAMGSKPSGFMNRNVNNAG
YMITWEPGGEFKPETEFRLWDCGHNYAPQSFNVDRQIVYGWMSPFVQPIPMED
DGWCGQLTLPREITLDDGDVVTAPVAEMEGLREDTLDHGSITLMDMDGEQVIADDA
EAVEIEMTIDLAASTADRAGLKIATEDGAYTYVAYDDQIGRVVVDQRQAMANGDH
GYRAAPLTDAELASGKLDLRFVDRGSVEVYVNGGHQVLSSYSYASEGPRAIKLVA
EFGNLKVESLKLHMKSIGLE

SEQ ID NO: 28

Saccharomyces cerevisiae Invertase: Genbank Accession No. NP_012104

MLLQAFLLAGFAAKISASMTNETSDRPLVHFTPNGKGMNDPNGLWYDEKDAKW
IILYFQYNPNDTVWGTPLFWGHATSDDL TNWEDQPIAIAPKRND SGAFSGSMVVDY
NNTSGFFNDTIDPRQRCVAIWTYNTPESSEEQYISYSLDGGYTFTFYQKNPVLAANSTQ
FRDPKVFWYEPSQKWIMTAAKSQDYKIEIYSSDDLKSWKLESANEGFLGYQYECF
GLIEVPTEQDPSKSYWVMFISINPGAPAGGSFNQYFVGSFNGTHFEAFDNQSRVVDFFG
KDYIALQTFFNTDPTYGSALGIASWNEYSFVPTNPWRSSMSLVKFSNLTEYQ
ANPETELINLKAEPILNISNAGPWSRFATNTTLTKANSYNVDLSNSTGTLEFELVYAV
NTTQTISKSVFADLSLWFKGLEDP EYL RMGFEVSASSFFLDRGNSKV KVFVKENPYFT
NRMSVNNQPKSENDLSYYKVYGLLDQNI LELYFNDGDVVSTNTYFMTTGNALGSV
NMTTGVDNLFYIDKFQVREVK

SEQ ID NO: 29

Zymomonas mobilis Invertase A: Genbank Accession No. AY171597

MESPSYKNLIKAEDAQKKAGKRLLSSEWYPGFHVTPLTGWMNDPNGLIFFKGEYHL
FYQYYPFAPVWGMHWHGAKSRDLVHWETLPVALAPGDLFDRDGCFSGCAVDNN
GVLTLIYTGHI VLSNDSPDAIREVQCMATSIDGIHFQKEGIVLEKAPMPQVAHFRDPR
VWKENDHWF MVVGYRTDDEKHQGHVVALYRSENLDWIFVKTL LGDNSQLPLGK
RAFMWECPDFFSLGNRSVLMFSPQGLKASGYKNRNL FQNGYILGWQAPQFTPETS
FQELDYGHDFYAAQRFEAKDGRQILIAWFDMWENQKPSQRD GWAGCMTLPRKLDL
IDNKIVMTPVREMEILRQSEKIESVVTLSDAEHPFTMDSPLQEIELIFDLEKSSAYQAG
LALRCNGKGQETLLYIDRSQNRILDRNRSGQNVKGIRSCPLPNTSKVRLHIFLDRSSIE
IFVGDDQTQGLYSSSRIFPDKDSLKGRLFAIEGYAVFDSFKRWTLQDANLAAFSSDA
C

SEQ ID NO:30

SAG 2214 *Chlorella luteoviridis*

TGTTGAAGAATGAGCCGCGGACTTATAGGAAGTGGCTTGGTTAAGGATACTTTCC
GAAGCCTAAGCGAAAGCAAGTTGTAACAATAGCGATATACCTCTTTGTAGGTCA
GTCACCTTCTTATGGACCCGAACCCGGGTGATCTAACCATGACCAGGATGAAGCTT
GGGTAACACCAAGTGAAGGTCCGAACCTCTTCGATCTTTAAAAATCGTGAGATGA
GTTATGGTTAGGGGTAAATCTGGCAGTTTTGCCCGCAAAAGGGTAACCTTTTGT
AATTACTGACTCATAACGGTGAAGCCTAAGGCGTTAGCTATGGTAATACCGTGGG

AAGTTTCAATACCTTCTTGCAATTTTTTTATTTGCACCTTTAGTGCAAACAGTGTA
AAGAAAGCGTTTTGAAACCCCTTAACGACTAATTTTTTGCTTTTGCAAGAACGTC
AGCACTCACCAATACACTTTCCGTTTTTTTTCTTTTATTAATTAAAGCAACATAAAA
ATATATTTTATAGCTTTAATCATAAAACTATGTTAGCACTTCGTGCTAATGTGCTA
ATGTGCTAATCAAATGAAAAGTGTTCTTAAAAGTGAGTTGAAGGTAGAGTCTAAT
CTTGCCTGAAAGGGCAAGCTGCACATTTTTTTTTTTGAATGTGCAACAATGGAAATG
CCAATCGAACTCGGAGCTAGCTGGTTCTCCCCGAAATGTGTGAGGGCGCAGCGAT
TCATGATTAGTACGGTGTAGGGGTAAAGCACTGTTTCGGTGCGGGCTGTGAAAAC
GGTACCAAATCGTGGCAAATAAGAATACTACGCTTGTATACCATGGATCAGTG
AGACTATGGGGGATAAGCTCCATAGTCAAGAGGGAAACAGCCCAGATCACCAGT
TAAGGCCCCAAAATGACAGCTAAGTGGCAAAGGAGGTGAAAAGTGCAAGAAACAA
CCAGGAGGTTTGCCCAGAAGCAGCCATCCTTTAAAGAGTGCGTAATAGCTCACTG

SEQ ID NO: 31

Endogenous *Chlorella protothecoides* actin promoter and 5'UTR

GAATTCGAGTTTAGGTCCAGCGTCCGTGGGGGGGGACGGGCTGGGAGCTTGGGC
CGGAAGGGCAAGACGATGCAGTCCCTCTGGGGAGTCACAGCCGACTGTGTGTG
TTGCACTGTGCGGCCCCGAGCACTCACACGCAAAATGCCTGGCCGACAGGCAGG
CCCTGTCCAGTGCAACATCCACGGTCCCTCTCATCAGGCTCACCTTGCTCATTGAC
ATAACGGAATGCGTACCGCTCTTTCAGATCTGTCCATCCAGAGAGGGGAGCAGG
CTCCCCACCGACGCTGTCAAACCTTGCTTCCTGCCCAACCGAAAACATTATTGTTT
GAGGGGGGGGGGGGGGGGGGGGAGATTGCATGGCGGGATATCTCGTGAGGAACAT
CACTGGGACACTGTGGAACACAGTGAGTGCAGTATGCAGAGCATGTATGCTAGG
GGTCAGCGCAGGAAGGGGGGCCTTTCCAGTCTCCCATGCCACTGCACCGTATCCA
CGACTCACCAGGACCAGCTTCTTGATCGGCTTCCGCTCCCGTGGACACCAGTGTG
TAGCCTCTGGACTCCAGGTATGCGTGCACCGCAAAGGCCAGCCGATCGTGCCGAT
TCCTGGGGTGGAGGATATGAGTCAGCCAACCTGGGGGCTCAGAGTGCACACTGGG
GCACGATACGAAACAACATCTACACCGTGTCTCCATGCTGACACACCACAGCTT
CGCTCCACCTGAATGTGGGCGCATGGGCCCCGAATCACAGCCAATGTGCTGCTGC
CATAATGTGATCCAGACCCTCTCCGCCCAGATGCCGAGCGGATCGTGGGCGCTGA
ATAGATTCCCTGTTTTGATCACTGTTTGGGTCTTTTCTTTTCGTCTCGGATGCGCG
TCTCGAAACAGGCTGCGTGGGCTTTTCGGATCCCTTTTGCTCCCTCCGTCACCATC
CTGCGCGCGGGCAAGTTGCTTGACCCTGGGCTGGTACCAGGGTTGGAGGGTATTA
CCGCGTCAGGCCATTCCCAGCCCGGATTCAATTCAAAGTCTGGGCCACCACCCTC
CGCCGCTCTGTCTGATCACTCCACATTCGTGCATACACTACGTTCAAGTCCTGATC
CAGGCGTGTCTCGGGACAAGGTGTGCTTGAGTTTGAATCTCAAGGACCCACTCCA
GCACAGCTGCTGGTTGACCCCGCCCTCGCAACTCCCTACCATGTCCTGCTGtaggtcca
gggatcttggcatgcacacaggaccccggttgtgggggtccccggtgcatgctgctgtgcaggcgccggtgtggggcctgggc
cccgcgaggctcaactcctcccatatgctgcccgtccctccaccaccgagacctggcccccttgcagAGGAAGGCGA
AGTCAGCGCCATCGTGTGCGATAATGGATCCGG

SEQ ID NO: 32

Endogenous *Chlorella protothecoides* EF1a promoter and 5'UTR

GAATTCGCCCTTGAGTTTAGGTCCAGCGTCCGTGGGGGGGGCGTGAGACTCCCCC
CTGACCTTCGTATGGCAGGGACTCCTACTTGCCAAGTAATCAGTTGACAATGCCA
CTTCAATGCTCGTTGTGGTACACTGACGCGGGTCTAACATACTGGGAAGCATGAA
TTGCCGACATGGACTCAGTTGGAGACAGTAACAGCTCTTTGTGTTCTATCTTCAG
GAACACATTTGGCAGCGCACCCATACAGTGGCGCACACGCAGCTGTACCTGATG

TGGCTCTATTCCCACATGTTTCAACTTGATCCAAAAGTCACTCAGACTCTCAGCA
GCTAGACTTGATCGCATCTTTGGCCATGAAGATGCTTGCGCAACTCTAGGAATGG
GACGAGAAAAGAGCCTGCTCTGATCGGATATTTCCATTCTCTGGATGGGACTGAG
ATGATTCTGAAGAAATGCTGCTCGACTTATTTGGAAGAACAGCACCTGACGCATG
CTTTGAGGCTGCTGTGGCTGGGATGTGCTGTATTTGTCAGCATTGAGCATCTACG
GGTAGATGGCCATAACACGCGCTGCCTATCATGCGGTGGGTGTGTGGAAAAC
GTACAATGGACAGAAATCAATCCCATTGCGAGCCTAGCGTGCAGCCATGCGCTC
CCTCTGTAGCCCCGCTCCAAGACAAAGCCAGCCAATGCCAGAACCCACATAGAG
AGGGTATCTTCCTAATGACCTCGCCCATCATTTCCCTCCAAATTA ACTATAATGCCT
TGATTGTGGAGTTGGCTTTGGCTTGACAGCTGCTCGCGCTGGCACTTTTGTAGGCA
GCACAGGGTATGCCAGCGCCGAACCTTTGTGCCCTTGAGCAGGCCACAAGGGCAC
AAGACTACACCATGCAGCTGGTATACTTGGAAGTATACCATTTCTTACCAAGCAA
GGCACAGCACAGCCTGCACCGACTCACTTTGCTTGAGCGGGGCACAGCGCCGCG
ACTGATCCTGCGAGCTGTGGGGAGTTCCGACTGTTCTGGACCTCGGTCTCTGAAA
GATGTGTACGATGGGATCAAGTCATTCAAGTATGCTCTTCACATGAGCAATCGGG
GGAGACACGGTGGCCCTAAAGGTGTTTCATCTGATTCAAGTGTAGTGGGGGGGTG
CTGTTTGTCCCGGGGCGCCCCCGCTCCCCGACCCCGGAGAAGGGCCCCAGAGG
ACTCGGCCGCCACAGAGGAATAACCGGGCGTGGCTCGGCCCTGCGCCTCCCTCT
TTCAATATTTACCTGGTGTTCAGTGCACGGACACGTAAAGAACTAGATACAATG
GCCGAGGGAAAGACGgtgagagcttggcgttggggaccgggcagcatcagaaactcctctccccgccccgccttgaaa
ctcaactgaactcctcctctccccctcgcagCATCTGTCTATCGTTATCgtgagtgaagggaactgccatgtgtcgggt
cgttgaccacggctcggcgcgtgctgccccgcgcgaacgttccctgcaaacgccgcgcagccgtcccttttctgccgcg
ccccccccctcgtcccccttcaatcacaccgcagTGCGGACATGTCGATTCCGGCAAGTCCACC

SEQ ID NO: 33

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

GAATTCCTGCAGGAAGAAGGCCGGCAGCAGCTGGTACTTGTCCTTCACCTCCTT
GATCGGCTGGGTGAGCTTGGCCGGGTGCGAGTCGTCGATGCCGGCATCGCCAG
CACGCTGTGCGGGGAGCCGGCATCGACAACCTTGGCACTGCTCACCTTGGTCACC
GGCATGGGGTCATGGCGCTGCAGACCAGCGGCCTGTCAGCATGCTGCAGGCATC
TGTGTTTTGTAGTAGATACTTTCTGATGCATCACACACGTTTGGAAGGTCCCCA
AGCCCCTTCAACAGTCTCGACATATGACACTCGCGCCCTCTTCCTCGTCCCGTGG
CCTGATGAGGGTACGCAGGTACCGCAGCTGCGCCCCGTCCCGCCAGTTGCCCTGG
CCCCGCCGGGCCCAATCTGTTCATTGCCGCTCCCTGGCAGCCGTGAACCTTCACAC
TACCGCTCTCTGTGACCTTCAGCACAGCAGGAATCGCCATTTACCGGCGGTCTGT
TGCTGCGGAGCCTCAGCTGATCTCGCCTGCGAGACCCACAGTTTGAATTTGCGG
TCCCCACACAACCTCTGACGCC

SEQ ID NO: 34

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

GAATTCCTCAGGAAGAAGGCCGGCAGCAGCTGGTACTTGTCCTTCACCTCCTTG
ATCGGCTGGGTGAGCTTCGCAGGATCGCAGTCGTCGATGCCGGCATCGCCAGC
ACGCTGTGCGGGGAGCCGGCATCNACAACCTTGGCACTGCTCCCCCTTGGTCACCG
GCATGGGGTCATGGCGCTGCAGCCCAGCGGCCTGTCAGCATGCTGCAGGCATCT
GTGTATTGTAGTAGGTACTTCCTGATGCATCAACACACGTTTGGAAGCTCCCCAA
GCCCCCTTCAACAGTCTCGACGTATGACACTCGCGCCCTCTTCCTCGCCCCGTGGC
CTGATGAGGGTACGCAGGTACCACAGCTGCGCCCCGTCCCGCCAGTTGCCCTGGC
CCGGCCGGGCCCAATCTGTTTCATTGCCGCTCCCTGGTAGCCGTGAACCTCACATTA
CCGCTCTCTGTGACCTTCAGCACAGCAGGAATCGCCATTTACCGGCGGTCTGTTG

CTGCGGAGCCTCAGCTGATCTCGCCTGCGAGACCCACAGTTTGAATTTGCGGTC
CCCACACAACCTCTGACGCC

SEQ ID NO: 35

TGACCTAGGTGATTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCG

SEQ ID NO: 36

CTACGAGCTCAAGCTTTCCATTTGTGTTCCCATCCCCTACTTCC

SEQ ID NO: 37

GATCAGAATTCCGCCTGCAACGCAAGGGCAGC

SEQ ID NO: 38

GCATACTAGTGGCGGGACGGAGAGAGGGCG

SEQ ID NO: 39

Beta-tubulin::neo::nit

gaattcctttcttgcgctatgacacttccagcaaaaggtaggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatgat
gcttcgacccccgaagctccttcggggctgcatggcgctccgatgccgctccaggcgagcgctgtttaaatagccaggcccc
gattgcaaagacattatagcgagctaccaaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatc
gcactccgctaagggggcgctccttcttcttcttllcagtcacaacccgcaaacggcgcccatatcaattgattgaacaagatggattg
cacgcagggttctccggcgcttgggtggagaggctattcggctatgactgggcacaacagacaatcggctgctctgatccgccgtgt
tccggctgtcagcgaggggcgccccggttcttttgcgaagaccgacctgtccgggtgccctgaatgaactgcaggacgaggcagcg
ggctatcgtggctggccacgacggcggttcttgcgcagctgtgctcgacgttgcactgaagcgggaagggaactggctgctattgg
gcgaagtgcggggcaggatctcctgtcatctcaccttgccttgcggagaaagtatccatcatggctgatgcaatgcggcgctgca
tacgcttgatccggctacctgcccattcgaccaccaagcgaacatcgcacgcagcgagcacgtactggatggaagccggcttctg
gatcaggatgatctggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcg
aggatctcgtcgtgacctatggcgatgcctgcttggcgaatatcatggtggaaaatggccgctttctggattcatcactgtgcccggc
tgggtgtggcgaccgctatcaggacatagcgttggctaccctgafattgctgaagagcttggcggcgaatgggtgaccgcttct
cgtgctttacggatcgcggctcccgttcgcagcgcacgccttctatgccttcttgacgagttcttcaagatctgtcgatcgacaagt
gactcgaggcagcagcagctcgatagatcgacacactctggacgctggtcgtgtgatgactgttccgccacacttgccttg
acctgtgaatatccctgccgttttatcaaacagcctcagtggtttgatctgtgtgtacgcgttttgcgagttgctagctgcttgcctatt
tgcgaataaccacccccagcatcccccttccctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcg
ctgctcctgctcctgctcactgcccctcgcacagccttgggttgggctccgctgtattctcctggtagtgaacctgtaaaccagcactg
caatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagctt**gagctc**

SEQ ID NO: 40

Glutamate dehydrogenase::neo::nit

gaattccgcctgcaacgaaggcagccacagccgctcccacccgccgctgaaccgacacgtgcttggcgccctgccgctgcct
gccgcatgcttgtgctggtgaggctgggcagtgctgccatgctgattgaggttgggtcatcgggtggaagcttatgtgtgctgggct
tgcattgccgggcaatgcgcatggtggaagaggcgccagcacttgcggacgtgccgggtgcttccagggtgttcaatcgcggc
agccagagggttcatgatcgcgcgtacaggtgagcagcagtgtagcaaaaggtagcagttgccagaatgatcggttcagctg
ttaatcaatgccagcaagagaagggtcaagtgaacacagggcatgccacagcacgggcaccggggagtgaatggcaccacc
aagtgtgtgcgagccagcatcggcgcttggctgttllcagctacaacggcaggagtcaccaactaacatagctgatcaacactgcaa
tcatcggcggtgatgcaagcatcctgcaagacacatgctgtgcgatgctgcgtgctgctgctgcgcacggcgttgattggcagc
agctcagccatgactggatcaggctgggtgccactgcaatgtggtggataggatgcaagtggagcgaaataccaaacccctctggt
gcttgcgtgggttgcattggcatcgccatcagcaggagcgcatgcgaagggaactggccccatgcacgccatgccaaaccggagcg
caccgagtgtccacactgtcaccaggccccgaagcttgcgaacctgctcatggacgcattagcgctgacgtcccttgacggcg

ctcctctcgggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcgagcagagcgggcgagcagcggcgggggcc
 acccttcttgcggggcgcgccccagccagcggtgagcgtgatcnnnccaaacgagttcacattcatttgcagcctggagaagcg
 aggctggggcctttgggtgtgtgcagcccgaatggaatcggggaccgccaggctagcagcaaggcgctcccctactccgcat
 cgtatgttccatagtgcatlggactgcatttgggtggggcgccggctgtttcttctgtgttgcaaaacgcgccacgtcagcaacctgtcc
 cgtgggtccccggtccgatgaaatcgtgtgcacgccgatcagctgattgccggctcgcgaagtaggcgccctcttctgtcgcgcc
 tctctccgtcccgccactagt**ggcgcgcc**atatca**atg**attgaacaagatggattgcagcaggttctccggccgclgggtggagag
gctattcggctatgactgggcacacagacaatcggtgctctgatgccggcggtgtccggctgtcagcgcagggcgcccggttctt
ttgtcaagaccgacctgtccggtgccctgaatgaactgcaggacgaggcagcgccgctatcgtggctggccacgacggcggttctt
ggcgagctgtgctcgcaggtgtcactgaagcggggaaggactggctgctattggcggaagtggcggggcaggatctcctgtcactc
acctgtcctcgtccgagaaagtatccatcatggtgatgcaatgcggcggtgcatacgttgatccggctacctgccattcgaccac
caagcgaaacatgcacgcagcgagcacgtactcgatggaagccggcttctgtcagcaggtatgactggacgaagagcatcagg
gctcgcggccagccgaactgttccgagcgaagcgccgatccccagcgcgaggatctcgtcgtgacctatggcgatgcctgct
tgccgaatatcatggtggaaaatggccgcttttctgattcatcgaactgtggcggtgtgggtgtgcgggacctatcaggacatagcg
ttggctacccgtgatattgctgaagagcttggcgcggaatggcggtgaccgcttctcgtgctttacggatcgcgcctccgattcgag
cgcacgccttctatcgcttcttgcagagttcttaagatctgtcgcagacaagt**gactcgagg**gcagcagcagctcggatagatcg
 acacactctggacgctggtcgtgtgagactgttgcggccacacttgccttgacctgtgaataacctgcccgtlilatcaaacagc
 ctacgtgtgttgcattgtgtgtacgcgcttttgcaggtgtagctgctgtgtatattgcgaataccacccccagcatcccttccctcgt
 ttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgccctcgcacag
 ccttggttgggctccgctgtatttctccttggtactgcaacctgtaaaccagcactgcaatgctgatgcaggggaagtagtgggatggg
 aacacaaatggaagctt**gagctc**

SEQ ID NO: 41

Glutamate dehydrogenase::neo-opt::nit

gaattcgcctgcaacgcaaggcgagccacagccgctcccacccgccgctgaaccgacacgtgcttgggcgcctgccgctgctgccgcat
 gcttgtgctggtgaggctgggcagtgctgccatgctgattgagccttgggttcatcgggtggaaagcttatgtgtgctgggcttgcagccgggcaa
 tgcgcatggtggcaagagggcgccagcacttgcctggacgtgccggcggtgcctccagggtggttcaatcgcggcagccagagggttcatga
 tcgcgctacaggttgagcagcagtgtagcgaaggtagcagtttgcagaatgatcggttcagctgttaataatgccagcaagagaagggtc
 aagtgcacaacacgggcatgccacagcacgggcaccggggagtggaatggcaccaccaagtgtgtgcgagccagcatcgccgctggtgttt
 cagctacaacggcgaggatcaccactaaccatagctgatcaacactgcaatcgcggcgctgatgcagcalctgcaagacacatgctgtg
 cgatgctgcgctgctgctgctgcgcacgccgttgagttggcagcagctcagccatgactggatcaggctgggctgccactgcaatgtgtgg
 ataggatgcaagtggagcgaatacacaacctctggctgcttgcgtgggtgcatggcatgcaccatcagcaggagcgcagtggaaggactg
 gccccatgcacgccatgcaaaaccggagcgaccgagtgccacactgtcaccaggccccgaagctttgcagaacctgctcatggacgcagtg
 tagcgctgacgtcccttgacggcgctcctctcgggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcgagcagagcggggca
 gcagcgggcgggggccaccttcttgcgggtcgcgccccagccagcggtgatgcgctgatcnnnccaaacgagttcacattcatttgcagcct
 ggagaagcgaggctggggcctttgggtgtgtgcagcccgaatggaatcggggaccgccaggctagcagcaaggcgctcccctactccg
 catcgatgttccatagtgcatlggactgcatttgggtggggcgccggctgtttcttctgtgttgcaaaacgcgccacgtcagcaacctgtccgltg
 ggtccccggtgccgatgaaatcgtgtgcacgccgatcagctgattgccggctcgcgaagtaggcgccctcttctgtcgcctctctccgtccc
 gccactagt**ggcgcgcc**atatca**atg**atcgcagcaggacggcctccacggcggtcccccgccgctgggtggagcgccgtgttcggctacgact
gggccccagcagaccatcggtgctccgacggcgccgtgttccgcttccgcccaggggccgccccgtgctgttcgtgaagaccgacctgtccg
gcgcctgaacgagctgcaggacgagggcgccgctgtcctggctggccaccaccggcggtgcctgcgccgctgtcgtgacgtggtgac
cgaggccggcgcgactggtgctgctggcgaggtgcccggccaggacctgtgtcctccacctggcccccgccgagaaggtgtccatca
tggccgacgccatgcggcgctgcacacctggaccccgccacctgccccctgaccaccaggccaaagcaccgcaltcagcgcggccgcac
ccgcatggagggcgccgtgtgtgaccaggacacctggacgaggagcaccaggcgctgcccccgccgagctgttcgccgctgaaggc
ccgcatgcccgcagggcgaggacctggtgtgacccacggcgacgcctgcctgcccacatcatggttgagaacggcgcttctccgcttcat
cgactgcggccgctggcgctggccgaccgtaccaggacatcgccctggccaccggcgacatcgccgaggagctggcgggcgagtgggc
cgaccgcttctggtgctgtacggcatcgccgccccgactcccagcgcacgccttctaccgctgtggacgagtttct**gactcgagg**gcag
 cagcagctcggatagatcgacacactctggacgctggtcgtgtgatggactgttgcggccacacttgccttgacctgtgaatatccctgccg
 tttatcaaacagcctcagtggttgcattgtgtgtacgcgttttgcaggttgtagctgcttgtgtatattgcgaataccacccccagcatcccttc
 cctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgccctcgcacagc

CTCCACCGCCCCGCCTGGCGGTCGCCGCCGGGGACGAGGCCCTGGCCAAGATCTG
CGGGACCATTGCGCGGGACGAGTCGCGCCACGAGGCGGCGTACACGCGGACCAT
GGATGCCATCTTCCAGCGCGACCCCAGCGGGGCCATGGTGGCGTTTGCACACATG
ATGATGCGCAAGATCACCATGCCCCGCCACCTCATGGACGACGGCCAGCACGGC
GCGCGCAACGGGGGGGGCGCAACTTGTTTCGACGACTTTGCGGCAGTGGCGGAGCG
GGCAGGGGTGTACACCGCCGGCGACTACATCGGCATCCTGCGCCACCTCATCCG
GCGCTGGGACGTGGAGGG

SEQ ID NO: 46

Acyl ACP desaturase

MHQASVSTMKTIQSRHAGPIGHVQSGRRSAGRAGSRMATPVAAATVAAPRSALNLS
PTIIRQEVLSASAAQQLDCVASLAPVFESQILPLLTPVDEMWOPTDFLPASNSEAFFDQ
IGDLRARSAAIPDDLVLCLVGDMITEEALPTYMAMLNLTLDVVRDETGHSHQHPYAKW
TRAWIAEENRHGDLNLYMWLTGRVGHAGGGAHHPAMLAVERTIQRLISSGMDPG
TENHPYHAFVFTSFQERATKLSHGSTARLAVAAGDEALAKICGTIARDES RHEAAYT
RTMDAIFQRDPSGAMVAFAMMMMRKITMPAHLMDDGQHGARNNGGAQLVRRLCGS
GGAGRGVHRRRLHRHPAPHPALGRGG

SEQ ID NO: 47

Geranyl geranyl diphosphate synthase

attatacatcgccatcgctcaggtttcacgatctgcatgctatctatgggactgtgactccgccggccaggttgggtgcgcgagaatc
ctccccgctctgccttcatttcctgacgggagtcgccgctgagcaccggcgatcATGGGCGTCGGCACACT
CCAAACCCCATATACATGTGGTCTGTCATTACGCATAGCGCACGGTATGTCCCG
CGACGCGCGGCTCGAAGCCGTGGCCATCCGACGCGCTGCACGGCCGAGGTGAGG
GCACGCCCCCTCCGCCAATGGCGCGCAGCCCATGACCGCCTTCGACTTCCGGCAGT
ACATGCAGCAGCGCGCCGCGCTGGTGGACGCAGCGCTGGACCTGGCAGTGCCGC
TGCAGTACCCCGAGAAGATCAACGAGGCCATGCGGTACAGCCTGCTGGCCGGGG
GCAAGCGCGTGCGCCCGCGCTCTGCCTCGCTGCCTGCGAGCTCGTGGGCGGCCC
CCTGGAGGCGGCCATGCCCCCGCCTGCGCCATGGAGATGATCCACACCATGAG
CCTCATCCACGACGACCTCCCCGCCATGGACAACGACGACTTCCGGCGCGGCCA
GCCCCGCCAACCACAAGGCCTATGGCGAGGAGATTGCGATCCTGGCGGGCGACGC
GCTGCTGTGCTGAGCTTTGAGCACATCGCGCGCGAGACGCGAGGCGTGGACCC
GGTGC GCGTCTGGCCGCCATCTCGGAGTGGCGCGCGGTGGGCAGCCGCGGGCT
GGTGGCGGGGCAGGTGGTGGACCTGGGTTTTCGAGGGCGGCGGCGTGGGGCTGGC
CCCGCTGCGCTACATCCACGAGCACAAAACCGCGGCGCTGCTGGAGGCGGCGGT
GGTGTCCGGCGCGCTGCTGGGCGGCGCGGAGGAGGCGGACCTGGAGCGCCTGCG
CACCTACAACCGCGCCATCGGCCTCGCTTTCAGGTGGTGGGGGACATCCTGGAC
ATCCCGGGGACCAGCGAGGAGCTGGGCAAGACCGCGGGCAAGGACCTGAGCTCC
CCAAAACCCCTACCCGTCCCTGGTGGGGCTGGCCAGGTCCAAAAAATTGCG
GACGAAGTATTGAGGACGCGAAAACCCAACTACCCAGTACGAGCCGGCCCCGA
GCGGCGCCCCCTCGTAACCCTGGCCGAAAACATTTGAaaccggaagaactgactgggggcccccc
ctgccccagatacggcggggctctccatccagtttgggatgggaggagcgacaaccgaccccgtaaccctgtacgcgttgc
ttgcatacgtacgatgccttgaaacccatccatgacctcaacaatacctggttgtgttagcttggtcctgaaaaaaaaaaaaaa
aaaaaaaaaaa

SEQ ID NO: 48

Geranyl geranyl diphosphate synthase

MGVGT LQTPYTCGRAFTHSARYVPRRAARSRGHPTRCTAEVRARPSANGAQPM TAF
DFRQYMQQRAALVDAALDLAVPLQYPEKINEAMRYSLLAGGKRVRPALCLAACEL
VGGPLEAAMPAACAMEMIHTMSLIHDDL PAMDND DFRRGQPANHKAYGEEIAII AG
DALLSLSFEHIARETRGVDPVRVLA AISEWRAVGSRGLVAGQVVDLGFEGGGVGLA
PLRYIHEHKTAALLEAAVVS GALLGGAEADLERLR TYNRAIGLAFQVVG DILDIPGT
SEELGKTAGKDLSSPKTPYPSLVGI ARSKKIAD EIJEDAKTQLTQYEPARAAPLVTLA
ENI

SEQ ID NO: 49

Geranyl geranyl diphosphate synthase

cagatgccATGCGCCCTCGGGCCGCGGGCCTGAGGGTCCACGCAGCGTCCTCGGTGG
CCCAGACGCACCAGGCCGCCCCCCCCGGCGGACAGGAGGTTTCGACGACTACCAGC
CCCGCACCGCCATCCTCTTCCCCGGCCAGGGCGCGCAGAGCGTGGGCATGGC GG
GAGAGCTGGCGAAGGCCGTCCCCGCCGCGCGGGCGCTGTTTCGACGCCGCCTCCG
ACCAGCTCGGCTATGACCTGCTCCGCGTGTGCGTTGAGGGCCCCAAGGCGCGCCT
GGACAGCACCGCCGTCAGCCAGCCCGCCATCTACGTGGCCAGCCTGGCGGGCGGT
GGAGAAGCTGCGCGCGGAGGGCGGGGAGGAGGCACTGGCCGCCATCGACGTCTG
CTGCCGGTCTGTCTTGGGCGAGTACACCGCGCTGGCCTTTGCCGGCGCCTTCTC
CTTCGCCGACGGGCTGCGCCTGGTGGCCCTGCGCGGCGCCAGCATGCAGGCCGC
CGCCGACGCCGCACCCTCGGGCATGGTCTCCGTCATCGGTCTGCCCTCCGACGCG
GTGGCCGCGCTGTGCGAGGCCGCCAACC GCGCAGGTGGCCCCCGACCAGGCCGTG
CGCATCGCCAACTACCTCTGCGACGGCAACTACGCCGTCAGCGGTGGGCTGGAG
GGCTGCGCGGCGGTGGAGGGCCTGGCCAAGGCCACAAGGCGCGCATGACGGTG
CGCCTGGCGGTGGCGGGCGCCTTCCACACCCCTTCATGCAGCCGGCGGTGGAG
GCGCTGAGCGCGGGCGCTGGCGGACACGCCGCTGGTCGCGCCGCGCATCCCCGT
GGTCAGCAACGGGACGCC

SEQ ID NO: 50

Geranyl geranyl diphosphate synthase

MRPRAAGLRVHAASSVAQTHQAAPPADRRFDDYQPRTAILFPGQGAQSVGMAGEL
AKAVPAAAAALFDAASDQLGYDLLRV CVEGPKARLDSTAVSQPAIYVASLAAVEKLR
AEGGEEALAAIDVAAGLSLGEYTA LAFAGAFSFADGLRLVALRGASMQAAADAAPS
GMVSVIGLPSDAVAALCEAANAQVAPDQAVRIANYLCDGNYAVSGGLEGCAAVEG
LAKAHKARMTVRLAVAGAFHTPFMQPAVEALSAGAGGHAAGRAAHPRGQQRDA

SEQ ID NO: 51

Glyceraldehyde 3-phosphate dehydrogenase cDNA sequence

TGTCCATCTCCCCCACCCTCCATCCAACCATCGTCGACGGCATGCAGGCGCTGT
GTTCTCACCCCGCGTCCCTCACGGCGCGTGCGGTACCCCATGGGCGGGCCAGCCC
AGCACAGCGGGTGTCCAGCGCCGGCCCGGCCTACACCGGCCTGTCCCGGCACAC
CCTGGGCTGCCCCAGCACCCCCACCCTCCAGTCCC GCGCCGCGGTCCAGACCCGC
GGCTCCTCCTCCGGCTCCACCACGCGCATGACCACCACCGCCCAGCGCAAGATCA
AGGTGGCCATCAACGGGTTTCGGCCGCATCGGCCGCCAGTTCCTGCGCTGCGTGGA
GGGGCGCGAGGACTCGCTGCTGGAGATCGTGGCCGTGAACGACTCCGGCGGGCGT
GAAGCAGGCCAGCCACCTGCTCAAGTACGACTCCACCATGGGCACCTTCAACGC
CGACATCAAGATCTCGGGCGAGGGCACCTTCTCCGTCAACGGCCGCGACATCCG
CGTCGTCTCCTCCCGCGACCCCTTGGCCCTGCCCTGGGGCGAGCTGGGCGTGGAC
CTGGTGATCGAGGGGACGGGAGTGTTTGTGGACCGCAAGGGTGCCAGCAAGCAC

CTGCAGGCGGGGGCCAAAGAAGGTCATCATCACCGCGCCGGCCAAGGGCTCCGAC
GTGCCCACCTACGTATGCGCGTGAACGCGGACCAGTACTCCAACCTCCGACGAC
ATCATCTCCAACGCCTCCTGCACCACCAACTGCCTGGCGCCCTTTGTCAAGGTGC
TCAACGACCGCTTCGGCATCGTGA

SEQ ID NO: 52

Glyceraldehyde 3-phosphate dehydrogenase

MQALCSHPASLTARAVPHGRASPAQRVSSAGPAYTGLSRHTLGCPSTPTLQSRAAVQ
TRGSSSGSTIRMTTTAQRKIKVAINGFGRIGRQFLRCVEGREDSI LEIVAVNDSGGVK
QASHLLKYDSTMGTFNADIKISGEGTFSVNGRDIRVVSSRDPLALPWGELGVDLVIEG
TGVFVDRKGASKHLQAGAKKVIITAPAKGSDVPTYVMGVNADQYSNSDDIISNASCT
TNCLAPFVKVLNDRFGIV

SEQ ID NO: 53

Oxygen Evolving Complex protein (OEE33) cDNA sequence

gatgttgagaatagtagcttgctgcttgccatgcagagcgtgtgcgcgcagtcggttcatgcaagggggccttcaccagtcctc
tgcggacccccgatgcagcaggagccagctcgtctgccgggctgatggcaaggccggagccttcataagaccgtaaagagcgg
tgctgccgctctggtgctcctcctcctgcttgggggtgcgggagcactgacctttgatgagctgcagggcctgacctacctgcag
gtgaagggtcttgcatcgccaacacctgccccacctgcttgggggtcctccaacatcaaggacctgaagagcgggacctactcc
gtcaacaagatgtgcttgagccacgtcctcaaggtaaggaggaggcacagttcaagaacggcgaggccgactttgtcccac
caagctgctacgcgtctgacctacacctggacgagatcttgccagatgaagatcgacggcagcggcggtggagttcaagg
aggaggatggcatgactatgctgcagtcaccgtgcagctccggcggggagcgcgtgcccttctcttaccatcaaggagcttg
acgccaaggggactgcccagggcttcaaggcgagttaccgtgccctcctaccgtgggtctccttcttgacccaagggccgc
ggcgctccaccggctacgacaacgccgtggccctgccccgccggcgatccgaggagttggagaaggagaacaacaagtc
accaaggctctgaagggggaggccatcttccatcgccaagggtggacgccgggacaggggaggtggcgggcatctttgagt

SEQ ID NO: 54

Oxygen Evolving Complex protein (OEE33)

MQSVCAQSVSCKGAFTQSLRTPRCRSRLVCRADGKAGAFIKTVKSGAAALAASLL
LSGGAGALTFDLQGLTYLQVKGSGIANTCPTLSGGSSNIKDLKSGTYSVNKMCLEP
TSFKVKEEAQFKNGEADFVPTKL VTRLTYTLDEISGQMKIDGSGGVEFKEEDGIDYA
AVTVQLPGGERVPLFTIKELDAKGTADGFKGEFTVPSYRGSSFLDPKGRGASTGYD
NAVALPAAGDSEELEKENNKSTKALKGEAIFSIKVDAGTGEVAGIFE

SEQ ID NO: 55

Clp protease cDNA sequence

ataatcggaacccagctgcacgcaccatcagtgccgcagcatgcagaccgtgcagccagctatggcgattggcgccctccggct
ccagcgtgacccggggctcgaccagcagcaagcagcacttaccacctcactccttttccggcttcaggcgctgaatcatgtgga
tcggcgggggcaggcgggggtctggagccccagacctgcagcaggccgtgggcaaggccgtgcgccggtcgcggggcccgc
accaccagcgcctgctgcgtgacctgcatgatgttgagcgggtcaccgagaaggccatcaagggtggtcatgctgcgcaggagga
ggctgcgcgtctggccacaactctgtgggagcggagcaaatcctgctggggttgattggggagtccacaggcatcgccccaagg
tctcaagtcgatggcggtcacgctgaaagatgcgcgttgagggtcgagaagatcatcgccgggggagcggctttgtggcgtg
gagatcccccttaccccccgcgccaagcgtgtgttgagctgtccctggaggaggctcgccagctcgccacaactacattggcac
ggagcacatcctgctggcctgctgcgcgagggtgagggcgtggcctcccgctgctggagacctggcgccgacccccagaa
gatccgcactcaggtgtacgcatggtgggtgagtcgcaggagcccggtgggcaccacgggtggcgagggtccaccggctccaa
caagatgccaccctggaggagtaggcaccaacctgaccgccagggcg

SEQ ID NO: 56

Clp protease

MQTVAASYGV LAPSGSSVTRGSTSSKQHFTTLTPFSGFRRLNHVDRAGQAGSGSPQT
LQQAVGKAVRRSRGRITSAVRVTRMMFERFTEKAIKVVMLAQEEARRLGHN FVGT
EQILLGLIGESTGIAAKVLKSMGVTLKDARVEVEKIIIGRSGSFVAVEIPFTPRAKRVLE
LSLEEARQLGHNYIGTEHILLGLLREGEVVASRVLET LGADPQKIRTQVVRMVGESQ
EPVGTTVGGGSTGSNKMP TLEEYGTNLTAQA

SEQ ID NO: 57

Crβ-tub::NCO-suc2::CvNitRed

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATATCAATG
CTTCTTCAGGCCTTTCTTTTCTTCTTGCTGGTTTTGCTGCCAAGATCAGCGCCTCT
ATGACGAACGAAACCTCGGATAGACCACTTGTGCACTTTACACCAAACAAGGGC
TGGATGAATGACCCCAATGGACTGTGGTACGACGAAAAAGATGCCAAGTGGCAT
CTGTACTTTCAATACAACCCGAACGATACTGTCTGGGGGACGCCATTGTTTTGGG
GCCACGCCACGTCCGACGACCTGACCAATTGGGAGGACCAACCAATAGCTATCG
CTCCGAAGAGGAACGACTCCGGAGCATTCTCGGGTTCCATGGTGGTTGACTACAA
CAATACTTCCGGCTTTTTCAACGATACCATTGACCCGAGACAACGCTGCGTGGCC
ATATGGACTTACAACACACCCGGAGTCCGAGGAGCAGTACATCTCGTATAGCCTG
GACGGTGGATACACTTTTACAGAGTATCAGAAGAACCCTGTGCTTGCTGCAAATT
CGACTCAGTTCGGAGATCCGAAGGTCTTTTGGTACGAGCCCTCGCAGAAGTGGAT
CATGACAGCGGCAAAGTCACAGGACTACAAGATCGAAATTTACTCGTCTGACGA
CCTTAAATCCTGGAAGCTCGAATCCGCGTTCGCAAACGAGGGCTTTCTCGGCTAC
CAATACGAATGCCAGGCCTGATAGAGGTCCCAACAGAGCAAGATCCCAGCAAG
TCCTACTGGGTGATGTTTATTTCCATTAAATCCAGGAGCACCGGCAGGAGGTTCTT
TTAATCAGTACTTCGTCGGAAGCTTTAACGGAACCTCATTTTCGAGGCATTTGATAA
CCAATCAAGAGTAGTTGATTTTGGAAAGGACTACTATGCCCTGCAGACTTTCTTC
AATACI'GACCCGACCTATGGGAGCGCTCTTGGCATTGCGTGGGCTTCTAACTGGG
AGTATTCCGCATTTCGTTCTACAACCCCTTGGAGGTCCTCCATGTGCTCGTGAG
GAAATTCTCTCTCAACACTGAGTACCAGGCCAACCCGGAACCGAACTCATAAA
CCTGAAAGCCGAACCGATCCTGAACATTAGCAACGCTGGCCCCCTGGAGCCGGTTT
GCAACCAACACCACGTTGACGAAAGCCAACAGCTACAACGTCGATCTTTCGAAT
AGCACCGGTACACTTGAATTTGAACTGGTGTATGCCGTCAATACCACCCAAACGA
TCTCGAAGTCGGTGTTTCGCGGACCTCTCCCTCTGGTTTAAAGGCCTGGAAGACCC
CGAGGAGTACCTCAGAAATGGGTTTTCGAGGTTTTCTGCGTCCTCCTTCTTCTTGATC
GCGGGAACAGCAAAGTAAAATTTGTTAAGGAGAACCCATATTTTACCAACAGGA
TGAGCGTTAACAACCAACCATTCAAGAGCGAAAACGACCTGTGCTACTACAAAG
TGTATGGTTTTGCTTGATCAAAATATCCTGGAACCTCTACTTCAACGATGGTGATGT
CGTGTCCACCAACACATACTTCATGACAACCGGGAACGCACTGGGCTCCGTGAA
CATGACGACGGGTGTGGATAACCTGTTCTACATCGACAAATTCCAGGTGAGGGA
AGTCAAGTGAGATCTGTGATCGACAAGCTCGAGGCAGCAGCAGCTCGGATAGT
ATCGACACACTCTGGACGCTGGTTCGTGTGATGGACTGTTGCCGCCACACTTGCTG
CCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATC
TTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCAC

CCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCT
ACGCTGTCCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCAC
AGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGC
ACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTT

SEQ ID NO: 58Crβ-tub::CO-suc2::CvNitRed

CTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTT
CCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCCGAAGCTCCTTCGGGG
CTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCC
CCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGA
TCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGG
GCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTG
CAGGCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGA
CGAACGAGACGTCCGACCGCCCCCTGGTGCATTCACCCCCAACAAGGGCTGGA
TGAACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGT
ACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCTGGGGCCA
CGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCC
GAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAA
CACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATC
TGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGAC
GGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCC
ACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGATCA
TGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCT
GAAGTCCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGGCTTCCTCGGCTACCA
GTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTC
CTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTC
AACCAGTACTTCGTCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACC
AGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAA
CACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAAGTGGGA
GTACTCCGCTTTCGTGCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCA
AGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACC
TGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTTCG
CCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACA
GCACCGGCACCTTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGA
TCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCC
CGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCTGAC
CGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGC
ATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGACCTGTCTACTACAAG
GTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGAC
GTCGTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGA
ACATGACGACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCG
AGGTCAAGTGATTAATTAACCTCGAGGCAGCAGCTCGGATAGTATCGACACA
CTCTGGACGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCT
GTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTA
CGCGCTTTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCAT
CCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCC
TGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCTCGCACAGCCTTGGT
TTGGGCTCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATG
CTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGA

SEQ ID NO: 59

Cinnamomum camphorum FatB1 (Q39473)

MATTSLASAFCSMKAVMLARDGRGMKPRSSDLQLRAGNAQTSCLKMINGTKFSYTES
LKKLPDWSMLFAVITITIFSAAEKQWTNLEWKPKPNPPQLDDHFGPHGLVFRRTFAI
RSYEVGPDRSTSIVAVMNLQEAALNHAKSVGILGDGFGTTLEMSKRDLIWVVKRT
HVAVERYPAWGDVEVECWVGASGNNRRHDFLVRDCKTGEILTRCTSLSVMMNT
RTRRLSKIPEEVRGEIGPAFIDNVAVKDEEIKKPQKLNDSTADYIQGGLTPRWNDLID
NQHVNNIKYVDWILETVPDSIFESHHSSTFIEYRRECTMDSVLQSLTTVSGGSSEAGL
VCEHLLQLEGGSEVLRAKTEWRPKLTDSFRGISVIPAESSV

SEQ ID NO: 60

Cinnamomum camphorum FatB1 codon optimized for Prototheca

ggcgcgccatggccaccacccctccctggcctccgcttctgcagcatgaaggccgtgatgctggcccgcgacggcc
gcgccatgaagcccgctccagcgacctgcagctgcgcgccggcaacgcccagacctccclgaagalgatcaacg
gcaccaagttctctacaccgagagcctgaagaagctgcccgaactggccatgctgttcgccgtgatcaccacca
tcttctccgcccgcgagaagcagtggaaccaacctggagtggaagcccaagcccaacccccccagctgctggacg
accacttcggcccccacggcctgggtgttcgcccgcaccttcgccatccgcagctacgaggtgggccccgacggct
ccaccagcatcgctggccgtgatgaaccacctgcaggaggccgcccctgaaccaagcccaagtcgctgggcatcctgg
gcgacggcttcggcaccacccctggagatgtccaagcgcgacctgatctgggtgggtgaagcgacccacgtggccg
tggagcgctacccccgctggggcgacacgctggaggtggagtgtctgggtggcgccgccggcaacaacggccgccc
gccacgacttccctgggtgcgcgactgcaagacggcgagatcctgacccgctgcacctccctgagcgctgatgatga
acacccgcacccgcccgcctgagcaagatccccgaggaggtgcgcggcgagatcgccccgccttcacgcacaacg
tggccgtgaaggacgaggagatcaagaagccccagaagctgaacgactccacgcgcgactacatccaggcgccg
tgaccccccgctggaacgacctggacatcaaccagcagtggaacaacatcaagtaacgtggactggatcctggaga
ccgtgcccgcacagcatcttcgagagccaccacatctcctccttcaccatcgagtaccgcccgcgagtgaccatgg
acagcgctgctgcagtcctgaccacgctgagcggcggtcctccgaggccggcctgggtgtgcgagcacctgctgc
agctggaggggcggcagcgaggtgctgcgcgccaagaccgagtgggcgcccaagctgaccgactccttcgcgccga
tcagcgctgatccccgcccagtcacgctgatggactacaaggaccacgacgagcgacgactacaaggaccacgacatcg
actacaaggacgacgacgacaagtgactcgagtltaattaa

SEQ ID NO: 61

Cuphea hookeriana FatB2 (GenBank AAC49269)

MVAAAASSAFFPVPAPGASPKPGKFGNWPSSLSPSFKPKSIPNGGFQVKANDSAHPK
ANGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLLTAITTVFVKSKRPDMHDRK
SKRPDMLVDSFGLESTVQDGLVFRQSFSIRS YEIGTDRTASIELTMNHLQETSLNHCK
STGILLDGFGRITLEMCKRDLIWVVIKMQIKVNRYPWGDVEINTRFSRLGKIGMGR
DWLISDCNTGEILVRATSAYAMMNQKTRRLSKI.PYEVHQEIVPLFVDSPIEDSDLK
VHKFKVKTGDSIQKGLTPGWNDLDVNQHVS NVKYIGWILESMPTVEVLETQELCSLA
LEYRRECGRDSVLESVTAMDPSKVGVR SQYQHLLRLEDGTAIVNGATEWRPKNAGA
NGAISTGKTSNGNSVS

SEQ ID NO: 62

Cuphea hookeriana FatB2 codon optimized for Prototheca

GGCGCGCCATGGTGGCCGCCGCCGCTCCAGCGCCTTCTTCCCCGTGCCCCCCCC
CGGCGCCTCCCCCAAGCCCGGCAAGTTCGGCAACTGGCCCTCCAGCCTGAGCCCC
TCCTTCAAGCCCAAGTCCATCCCCAACGGCGGCTTCCAGGTGAAGGCCAACGAC
AGCGCCACCCCAAGGCCAACGGCTCCGCCGTGAGCCTGAAGAGCGGCAGCCTG
AACACCCAGGAGGACACCTCCTCCAGCCCCCCCCCGCACCTTCTTGCACCAGC
TGCCCGACTGGAGCCGCTGCTGACCGCCATCACCACCGTGTTCTGTGAAGTCCAA
GCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGACATGCTGGTGGACAG

CTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTCGCCAGTCCTTCTCC
ATCCGCTCCTACGAGATCGGCACCGACCGCACCAGCATCGAGACCCTGATG
AACCACCTGCAGGAGACCTCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTG
GACGGCTTCGGCCGCACCCTGGAGATGTGCAAGCGCGACCTGATCTGGGTGGTG
ATCAAGATGCAGATCAAGGTGAACCGCTACCCCGCCTGGGGCGACACCGTGGAG
ATCAACACCCGCTTCAGCCGCCTGGGCAAGATCGGCATGGGGCCGCGACTGGCTG
ATCTCCGACTGCAACACCGGCGAGATCCTGGTGC GCGCCACCAGCGCCTACGCC
ATGATGAACCAGAAGACCCGCCGCCTGTCCAAGCTGCCCTACGAGGTGCACCAG
GAGATCGTGCCCTGTTCGTGGACAGCCCCGTGATCGAGGACTCCGACCTGAAG
GTGCACAAGTTCAAGGTGAAGACCGGCGACAGCATCCAGAAGGGCCTGACCCCC
GGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGTGAAGTACATCGGC
TGGATCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCC
CTGGCCCTGGAGTACCGCCGCGAGTGCGGGCCGCGACTCCGTGCTGGAGAGCGTG
ACCGCCATGGACCCCAGCAAGGTGGGCGTGCGCTCCAGTACCAGCACCTGCTG
CGCCTGGAGGACGGCACCGCCATCGTGAACGGCGCCACCAGTGGCGCCCCAAG
AACGCCGGCGCCAACGGCGCCATCTCCACCGGCAAGACCAGCAACGGCAACTCC
GTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGAC
TACAAGGACGACGACGACAAGTGACTCGAGTTAATTAA

SEQ ID NO: 63

Umbellularia FatB1 (GenBank Q41635)

MATTSLASAFCSMKAVMLARDGRGMKPRSSDLQLRAGNAPTSLKMINGTKFSYTES
LKRLPDWSMLFAVITITIFSAAEKQWTNLEWKPKLPQLLDDHFGLHGLVFRRTFAI
RSYEVGPDRSTSILAVMNHMQEATLNHAKSVGILGDGFGTTLEMSKRDLMWVVR
THVAVERYPTWGD TVEVECWIGASGNNGMRRDFLVRDCKTGEILTRCTSLSVLMNT
RTRRLSTIPDEV RGEIGPAFIDNVAVKDDEIKKLQKLNDSADYIQGGLTPRWNDLDV
NQHVNNLKYVAWVFETVPDSIFESHHSFTLEYRRECTRDSVLRSLTTVSGGSSEAG
LVCDHLLQLEGGSEVLRARTEWRPKLTDSFRGISVIPAEPRV

SEQ ID NO: 64

Umbellularia FatB1 codon optimized for *Prototheca*

GGCGCGCCATGGCCACCACCAGCCTGGCCTCCGCCTTCTGCTCCATGAAGGCCGT
GATGCTGGCCCGCGACGGCCGCGGCATGAAGCCCCGCGAGCTCCGACCTGCAGCT
GCGCGCCGGCAACGCCCCACCTCCCTGAAGATGATCAACGGCACCAAGTTTACAG
CTACACCGAGAGCCTGAAGCGCCTGCCGACTGGTCCATGCTGTTTCGCCGTGATC
ACCACCATCTTCAGCGCCGCGGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCC
AAGCCCAAGCTGCCCCAGCTGCTGGACGACCACTTCGGCCTGCACGGCCTGGTGT
TCCGCCGCACCTTCGCCATCCGCTCCTACGAGGTGGGCCCCGACCGCAGCACCTC
CATCCTGGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACCACGCCAAGAG
CGTGGGCATCCTGGGCGACGGCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGA
CCTGATGTGGGTGGTGCGCCGCACCCACGTGGCCGTGGAGCGCTACCCACCTGG
GGCGACACCGTGGAGGTGGAGTGCTGGATCGGCGCCAGCGGCAACAACGGCATG
CGCCGCGACTTCCTGGTGC GCGACTGCAAGACCGGCGAGATCCTGACCCGCTGC
ACCTCCCTGAGCGTGCTGATGAACACCCGCAACCCGCCGCTGAGCACCATCCCCG
ACGAGGTGCGCGGCGAGATCGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGG
ACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCC
AGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACA
ACCTGAAGTACGTGGCCTGGGTGTTTCGAGACCGTGCCCGACAGCATCTTCGAGTC

CCACCACATCAGCTCCTTCACCCTGGAGTACCGCCGCGAGTGCACCCGCGACTCC
GTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCAGCTCCGAGGCCGGCCTGGTG
TGCAGCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCGCACC
GAGTGGCGCCCCAAGCTGACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCG
AGCCCCGCGTGATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACA
TCGACTACAAGGACGACGACGACAAGTGACTCGAGTTAATTAA

SEQ ID NO: 65

NeoR forward primer

CCGCCGTGCTGGACGTGGTG

SEQ ID NO: 66

NeoR reverse primer

GGTGGCGGGGTCCAGGGTGT

SEQ ID NO: 67

Yeast invertase *suc2* forward primer

CGGCCGGCGGCTCCTTCAAC

SEQ ID NO: 68

Yeast invertase *suc2* reverse primer

GGCGCTCCCGTAGGTCGGGT

SEQ ID NO: 69

5'UTR/promoter *Chlorella sorokiniana* glutamate dehydrogenase

CGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAACCGACAC
GTGCTTGGGCGCCTGCCGCCTGCCTGCCGCATGCTTGTGCTGGTGAGGCTGGGCA
GTGCTGCCATGCTGATTGAGGCTTGGTTTCATCGGGTGGAAGCTTATGTGTGTGCT
GGGCTTGCATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCACTTGCTG
GAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGATTTTCAG
ATGATCGCGCGTACAGGTTGAGCAGCAGTGTCAGCAAAGGTAGCAGTTTGCCAG
AATGATCGGTTTCAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGTGCAAA
CACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCACCAAGTGTG
TGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTCATCCAA
CGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGCAAGCAT
GCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCTGCGCAC
GCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGGCTGCCA
CTGCAATGTGGTGGATAGGATGCAAGTGGAGCGAATACCAAACCTCTGGCTGC
TTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGGGACTGG
CCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCACACTGTCACCAG
GCCCCGAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGACGTCCCTT
GACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGCAGAGGCG
GCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCITCTTGCGGGGTGCG
GCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCATTTGCATGC
CTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGGAATGCGG

GACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATGTTCCATAG
TGCATTGGACTGCATTTGGGTGGGGCGGCCGGCTGTTTCTTTTCGTGTTGCAAAAC
GCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAATCGTGTGCA
CGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTCTGCTCGCC
CTCTCTCCGTCCCGCCACTAGTGGCGCGCC

SEQ ID NO: 70

Umbellularia californica C12 thioesterase coding region

ATGGCCACCACCAGCCTGGCCTCCGCCTTCTGCTCCATGAAGGCCGTGATGCTGG
CCCGCGACGGCCGCGGCATGAAGCCCCGAGCTCCGACCTGCAGCTGCGCGCCG
GCAACGCCCCACCTCCCTGAAGATGATCAACGGCACCAAGTTCAGCTACACCG
AGAGCCTGAAGCGCCTGCCCCGACTGGTCCATGCTGTTTCGCCGTGATCACCACCAT
CTTCAGCGCCGCGGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAA
GCTGCCCCAGCTGCTGGACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCGCCG
ACCTTCGCCATCCGCTCCTACGAGGTGGGCCCCGACCGCAGCACCTCCATCCTGG
CCGTGATGAACCACATGCAGGAGGCCACCCTGAACCACGCCAAGAGCGTGGGCA
TCCTGGGCGACGGCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATGTG
GGTGGTGCGCCGACCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACAC
CGTGGAGGTGGAGTGCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGA
CTTCCTGGTGCGCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTG
AGCGTGCTGATGAACACCCGCACCCGCCGCCTGAGCACCATCCCCGACGAGGTG
CGCGGCGAGATCGGCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGACGAG
ATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGC
CTGACCCCCCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAG
TACGTGGCCTGGGTGTTTCGAGACCGTGCCCGACAGCATCTTCGAGTCCCACCACA
TCAGCTCCTTCACCCTGGAGTACCGCCGCGAGTGCACCCGCGACTCCGTGCTGCG
CAGCCTGACCACCGTGAGCGGCGGCAGCTCCGAGGCGCGCCTGGTGTGCGACCA
CCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCGCACCGAGTGGCG
CCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGC
GTG

SEQ ID NO: 71

Cinnamomum camphora C14 thioesterase coding region

ATGGCCACCACCTCCCTGGCCTCCGCCTTCTGCAGCATGAAGGCCGTGATGCTGG
CCCGCGACGGCCGCGGCATGAAGCCCCGCTCCAGCGACCTGCAGCTGCGCGCCG
GCAACGCCCAGACCTCCCTGAAGATGATCAACGGCACCAAGTTCCTACACCG
AGAGCCTGAAGAAGCTGCCCCGACTGGTCCATGCTGTTTCGCCGTGATCACCACCAT
CTTCTCCGCCGCCGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAA
CCCCCCCCAGCTGCTGGACGACCACTTCGGCCCCCACGGCCTGGTGTTCGCGCCG
ACCTTCGCCATCCGACGCTACGAGGTGGGCCCCGACCGCTCCACCAGCATCGTGG
CCGTGATGAACCACCTGCAGGAGGCCGCCCTGAACCACGCCAAGTCCGTGGGCA
TCCTGGGCGACGGCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATCTG
GGTGGTGAAGCGCACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACAC
CGTGGAGGTGGAGTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCGCCACGA
CTTCCTGGTGCGCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTG
AGCGTGATGATGAACACCCGCACCCGCCGCCTGAGCAAGATCCCCGAGGAGGTG
CGCGGCGAGATCGGCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGAGGAG
ATCAAGAAGCCCCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGC

CTGACCCCCCGCTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAG
TACGTGGACTGGATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCAC
ATCTCCTCCTTCACCATCGAGTACCGCCGCGAGTGCACCATGGACAGCGTGCTGC
AGTCCCTGACCACCGTGAGCGGCGGCTCCTCCGAGGCCGGCCTGGTGTGCGAGC
ACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGC
GCCCCAAGCTGACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAG
CGTG

SEQ ID NO: 72

FLAG-tag sequence

ATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAG
GACGACGACGACAAGTGA

SEQ ID NO: 73

3' UTR *Chlorella vulgaris* nitrate reductase

CTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGT
GATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTT
TATCAAACAGCCTCAGTGTGTGTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTA
GCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCCCTCGTTTCATAT
CGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCT
CCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCT
CCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGT
GGGATGGGAACACAAATGGA

SEQ ID NO: 74

C12 thioesterase forward primer

5' CTGGGCGACGGCTTCGGCAC 3'

SEQ ID NO: 75

C12 thioesterase reverse primer

5' AAGTCGCGGCGCATGCCGTT 3'

SEQ ID NO: 76

C14 thioesterase forward primer

5' TACCCCGCCTGGGGCGACAC 3'

SEQ ID NO: 77

C14 thioesterase reverse primer

5' CTTGCTCAGGCGGCGGGTGC 3'

SEQ ID NO: 78

Codon optimized *Cuphea hookeriana* C8-10 thioesterase coding region with native plastid targeting sequence with 3x FLAG tag

ATGGTGGCCGCCGCCGCTCCAGCGCCTTCTTCCCCGTGCCCCGCCCCCGGCGCCT
 CCCCCAAGCCCGGCAAGTTCGGCAACTGGCCCTCCAGCCTGAGCCCCTCCTTCAA
 GCCAAGTCCATCCCCAACGGCGGCTTCCAGGTGAAGGCCAACGACAGCGCCCA
 CCCCCAAGGCCAACGGCTCCGCCGTGAGCCTGAAGAGCGGCAGCCTGAACACCCA
 GGAGGACACCTCCTCCAGCCCCCCCCCCCCGCACCTTCCTGCACCAGCTGCCCCGAC
 TGGAGCCGCCTGCTGACCGCCATCACCACCGTGTTCGTGAAGTCCAAGCGCCCCG
 ACATGCACGACCGCAAGTCCAAGCGCCCCGACATGCTGGTGGACAGCTTCGGCC
 TGGAGTCCACCGTGCAGGACGGCCTGGTGTTCGCCAGTCCTTCTCCATCCGCTC
 CTACGAGATCGGCACCGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCT
 GCAGGAGACCTCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTGGACGGCTT
 CGGCCGCACCCTGGAGATGTGCAAGCGCGACCTGATCTGGGTGGTGAATCAAGAT
 GCAGATCAAGGTGAACCGCTACCCCGCCTGGGGCGACACCGTGGAGATCAACAC
 CCGCTTCAGCCGCTGGGCAAGATCGGCATGGGCCGCGACTGGCTGATCTCCGAC
 TGCAACACCGGCGAGATCCTGGTGCAGCGCCACCAGCGCCTACGCCATGATGAAC
 CAGAAGACCCGCCGCCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTG
 CCCCTGTTTCGTGGACAGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGT
 TCAAGGTGAAGACCGGCGACAGCATCCAGAAGGGCCTGACCCCCGGCTGGAACG
 ACCTGGACGTGAACCAGCACGTGTCCAACGTGAAGTACATCGGCTGGATCCTGG
 AGAGCATGCCCACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCCCTGGCCCTGG
 AGTACCGCCGCGAGTGCAGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGG
 ACCCCAGCAAGGTGGGCGTGCGCTCCCAGTACCAGCACCTGCTGCGCCTGGAGG
 ACGGCACCGCCATCGTGAACGGCGCCACCGAGTGGCGCCCCAAGAACGCCGGCG
 CCAACGGCGCCATCTCCACCGGCAAGACCAGCAACGGCAACTCCGTGTCCATGG
 ACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACG
 ACGACGACAAGTGA

SEQ ID NO: 79

Cuphea hookeriana C8-10 thioesterase coding region and 3x FLAG tag, with *Chlorella* *protothecoides* stearoyl ACP desaturase plastid targeting sequence

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTC
 GCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGAGGCCCTCCCCGTGCGCGGGC
 GCGCCCAGCTGCCCCACTGGAGCGCCTGCTGACCGCCATCACCACCGTGTTCGT
 GAAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGACATGCT
 GGTGGACAGCTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTCGCCA
 GTCCTTCTCCATCCGCTCCTACGAGATCGGCACCGACCGCACCGCCAGCATCGAG
 ACCCTGATGAACCACCTGCAGGAGACCTCCCTGAACCACTGCAAGAGCACCGGC
 ATCCTGCTGGACGGCTTCGGCCGACCCCTGGAGATGTGCAAGCGCGACCTGATCT
 GGGTGGTGAATCAAGATGCAGATCAAGGTGAACCGCTACCCCGCCTGGGGCGACA
 CCGTGGAGATCAACACCCGCTTCAGCCGCCTGGGCAAGATCGGCATGGGCCGCG
 ACTGGCTGATCTCCGACTGCAACACCGGCGAGATCCTGGTGCAGCGCCACCAGCG
 CCTACGCCATGATGAACCAGAAGACCCGCCGCTGTCCAAGCTGCCCTACGAGG
 TGCACCAGGAGATCGTGCCCCTGTTTCGTGGACAGCCCCGTGATCGAGGACTCCGA
 CCTGAAGGTGCACAAGTTCAAGGTGAAGACCGGCGACAGCATCCAGAAGGGCCT
 GACCCCCGGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGTGAAGTA
 CATCGGCTGGATCCTGGAGAGCATGCCCACCGAGGTGCTGGAGACCCAGGAGCT
 GTGCTCCCTGGCCCTGGAGTACCGCCGCGAGTGCAGGCCGCGACTCCGTGCTGGAG
 AGCGTGACCGCCATGGACCCAGCAAGGTGGGCGTGCGCTCCAGTACCAGCAC
 CTGCTGCGCCTGGAGGACGGCACCGCCATCGTGAACGGCGCCACCGAGTGGCGC
 CCAAGAACGCCGGCGCCAACGGCGCCATCTCCACCGGCAAGACCAGCAACGGC

AACTCCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGAC
ATCGACTACAAGGACGACGACGACAAGTGA

SEQ ID NO: 80

Cuphea hookeriana C8-10 thioesterase coding region and 3x FLAG tag, with *Prototheca moriformis* delta 12 fatty acid desaturase plastid targeting sequence

ATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTACGATCGGG
ACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTCGTGGGC
GCGCCCAGCTGCCCCGACTGGAGCCGCTGCTGACCGCCATCACCACCGTGTTCGT
GAAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGACATGCT
GGTGGACAGCTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTCGCCA
GTCTTCTCCATCCGCTCCTACGAGATCGGCACCGACCGCACC GCCAGCATCGAG
ACCCTGATGAACCACCTGCAGGAGACCTCCCTGAACCACTGCAAGAGCACCGGC
ATCCTGCTGGACGGCTTCGGCCGCAACCCTGGAGATGTGCAAGCGCGACCTGATCT
GGGTGGTGATCAAGATGCAGATCAAGGTGAACCGCTACCCCGCCTGGGGCGACA
CCGTGGAGATCAACACCCGCTTCAGCCGCTGGGCAAGATCGGCATGGGCCGCG
ACTGGCTGATCTCCGACTGCAACACCGGCGAGATCCTGGTGCGCGCCACCAGCG
CCTACGCCATGATGAACCAGAAGACCCGCCGCTGTCCAAGCTGCCCTACGAGG
TGCACCAGGAGATCGTGCCCCCTGTTTCGTGGACAGCCCCGTGATCGAGGACTCCGA
CCTGAAGGTGCACAAGTTCAAGGTGAAGACCGGCGACAGCATCCAGAAGGGCCT
GACCCCCGGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGTGAAGTA
CATCGGCTGGATCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAGGAGCT
GTGCTCCCTGGCCCTGGAGTACCGCCGCGAGTGCGGCCGCGACTCCGTGCTGGAG
AGCGTGACCGCCATGGACCCCAGCAAGGTGGGCGTGCGCTCCAGTACCAGCAC
CTGCTGCGCCTGGAGGACGGCACCGCCATCGTGAACGGCGCCACCGAGTGGCGC
CCCAAGAACGCCGGCGCCAACGGCGCCATCTCCACCGGCAAGACCAGCAACGGC
AACTCCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGAC
ATCGACTACAAGGACGACGACGACAAGTGA

SEQ ID NO: 81

Cuphea hookeriana C8-10 thioesterase coding region and 3x FLAG tag, with *Prototheca moriformis* isopentenyl diphosphate plastid targeting sequence

ATGACGTTTCGGGGTTCGCCCTCCCGGCCATGGGCGCGGTGTCTCCCTTCCCCGGC
CCAGGGTTCGCGGTGCGCGCCAGTCGGCGAGTCAGGTTTTGGAGAGCGGGCGCG
CCCAGCTGCCCCGACTGGAGCCGCTGCTGACCGCCATCACCACCGTGTTCGTGAA
GTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGACATGCTGGT
GGACAGCTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTCGCCAGTCC
TTCTCCATCCGCTCCTACGAGATCGGCACCGACCGCACCGCCAGCATCGAGACCC
TGATGAACCACCTGCAGGAGACCTCCCTGAACCACTGCAAGAGCACCGGCATCC
TGCTGGACGGCTTCGGCCGCAACCCTGGAGATGTGCAAGCGCGACCTGATCTGGGT
GGTGATCAAGATGCAGATCAAGGTGAACCGCTACCCCGCCTGGGGCGACACCGT
GGAGATCAACACCCGCTTCAGCCGCTGGGCAAGATCGGCATGGGCCGCGACTG
GCTGATCTCCGACTGCAACACCGGCGAGATCCTGGTGCGCGCCACCAGCGCCTAC
GCCATGATGAACCAGAAGACCCGCCGCTGTCCAAGCTGCCCTACGAGGTGCAC
CAGGAGATCGTGCCCCCTGTTTCGTGGACAGCCCCGTGATCGAGGACTCCGACCTGA
AGGTGCACAAGTTCAAGGTGAAGACCGGCGACAGCATCCAGAAGGGCCTGACCC
CCGGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGTGAAGTACATCG
GCTGGATCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAGGAGCTGTGCT

CCCTGGCCCTGGAGTACCGCCGCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCG
TGACCGCCATGGACCCCAGCAAGGTGGGCGTGCGCTCCCAGTACCAGCACCTGC
TGCGCCTGGAGGACGGCACCGCCATCGTGAACGGCGCCACCGAGTGGCGCCCCA
AGAACGCCGGCGCCAACGGCGCCATCTCCACCGGCAAGACCAGCAACGGCAACT
CCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCG
ACTACAAGGACGACGACGACAAGTGA

SEQ ID NO: 82

Umbellularia californica C12 thioesterase coding region and 3x FLAG tag, with *Prototheca moriformis* isopentenyl diphosphate plastid targeting sequence

ATGACGTTTCGGGGTTCGCCCTCCCGGCCATGGGGCCGCGGTGTCTCCCTTCCCCGGC
CCAGGGTCGCGGTGCGCGCCCAGTCGGCGAGTCAGGTTTTGGAGAGCGGGCGCG
CCCCGACTGGTCCATGCTGTTCCGCGTGATCACCACCATCTTCAGCGCCGCCGA
GAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCAGCTGCT
GGACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCGCCGACCTTCGCCATCCGC
TCCTACGAGGTGGGCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGAACCACA
TGCAGGAGGCCACCCTGAACCACGCCAAGAGCGTGGGCATCCTGGGCGACGGCT
TCGGCACCAACCTGGAGATGTCCAAGCGCGACCTGATGTGGGTGGTGCGCCGCA
CCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTGGAGGTGGAGT
GCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGGTGC CG
ACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGATGA
ACACCCGCACCCGCCGCTGAGCACCATCCCCGACGAGGTGCGCGGCGAGATCG
GCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGACGAGATCAAGAAGCTGC
AGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCGCT
GGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGCCTGGG
TGTTTCGAGACCGTGCCCGACAGCATCTTCGAGTCCCACCACATCAGCTCCTTCAC
CCTGGAGTACCGCCGCGAGTGCACCCGCGACTCCGTGCTGCGCAGCCTGACCACC
GTGAGCGGCGGCAGCTCCGAGGCCGCGCCTGGTGTGCGACCACTGCTGCAGCTG
GAGGGCGGCAGCGAGGTGCTGCGCGCCCGCACCGAGTGGCGCCCCAAGCTGACC
GACTCCTTCCGCGGCATCAGCGTGATCCCCGCGGAGCCCCGCGTGATGGACTACA
AGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGAC
GACAAGTGA

SEQ ID NO: 83

Umbellularia californica C12 thioesterase coding region and 3x FLAG tag, with *Prototheca moriformis* stearoyl ACP desaturase plastid targeting sequence

ATGGCTTCCGCGGCATTACCATGTTCGGCGTGCCCCGCGATGACTGGCAGGGCCC
CTGGGGCACGTTCGCTCCGGACGGCCAGTCGCCACCCGCCTGAGGGGGCGCGCCC
CCGACTGGTCCATGCTGTTCCGCGTGATCACCACCATCTTCAGCGCCGCGGAGAA
GCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCAGCTGCTGGA
CGACCACTTCGGCCTGCACGGCCTGGTGTTCGCGCCGACCTTCGCCATCCGCTCC
TACGAGGTGGGCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGAACCACATG
CAGGAGGCCACCCTGAACCACGCCAAGAGCGTGGGCATCCTGGGCGACGGCTTC
GGCACCAACCTGGAGATGTCCAAGCGCGACCTGATGTGGGTGGTGCGCCGCACC
CACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTGGAGGTGGAGTGC
TGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGGTGC CGC
TGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGATGAAC
ACCCGCACCCGCCGCTGAGCACCATCCCCGACGAGGTGCGCGGCGAGATCGGC

CCCGCCTTCATCGACAACGTGGCCGTGAAGGACGACGAGATCAAGAAGCTGCAG
AAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCGCTGG
AACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGCCTGGGTG
TTCGAGACCGTGCCCGACAGCATCTTCGAGTCCCACCACATCAGCTCCTTCACCC
TGGAGTACCGCCGCGAGTGCACCCGCGACTCCGTGCTGCGCAGCCTGACCACCGT
GAGCGGCGGCAGCTCCGAGGCCGCGCCTGGTGTGCGACCACTGCTGCAGCTGGA
GGGCGGCAGCGAGGTGCTGCGCGCCCGCACCGAGTGGCGCCCCAAGCTGACCGA
CTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGCGTGATGGACTACAAG
GACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGA
CAAGTGA

SEQ ID NO: 84

Umbellularia californica C12 thioesterase coding region and 3x FLAG tag, with *Chlorella protothecoides* stearyl ACP desaturase plastid targeting sequence

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTC
GCTCGGCGGGCTCCGGGCCCCGCGCCAGCGAGGCCCTCCCCGTGCGCGGGC
GCGCCCCCGACTGGTCCATGCTGTTCCGCGTGATCACCACCATCTTCAGCGCCGC
CGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCAGCT
GCTGGACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCGCGACCTTCGCCATC
CGCTCCTACGAGGTGGGCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGAACC
ACATGCAGGAGGCCACCCTGAACCACGCCAAGAGCGTGGGCATCCTGGGCGACG
GCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATGTGGGTGGTGCGCC
GCACCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTGGAGGTGG
AGTGCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGGTGC
GCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGAT
GAACACCCCGCACCCGCCGCTGAGCACCATCCCCGACGAGGTGCGCGGCGAGAT
CGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGACGAGATCAAGAAGCT
GCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCG
CTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGCCTG
GGTGTTCGAGACCGTGCCCGACAGCATCTTCGAGTCCCACCACATCAGCTCCTTC
ACCCTGGAGTACCGCCGCGAGTGCACCCGCGACTCCGTGCTGCGCAGCCTGACC
ACCGTGAGCGGCGGCAGCTCCGAGGCCGCGCCTGGTGTGCGACCACTGCTGCAG
CTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCGCACCGAGTGGCGCCCCAAGCTG
ACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGCGTGATGGACT
ACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGAC
GACGACAAGTGATGA

SEQ ID NO: 85

Umbellularia californica C12 thioesterase coding region and 3x FLAG tag, with *Prototheca moriformis* delta 12 fatty acid desaturase plastid targeting sequence

ATGGCTATCAAGACGAACAGGCAGCCTGTGGAGAAGCCTCCGTTACGATCGGG
ACGCTGCGCAAGGCCATCCCCGCGCACTGTTTCGAGCGCTCGGCGCTTCGTGGGC
GCGCCCCCGACTGGTCCATGCTGTTCCGCGTGATCACCACCATCTTCAGCGCCGC
CGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCAGCT
GCTGGACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCGCGACCTTCGCCATC
CGCTCCTACGAGGTGGGCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGAACC
ACATGCAGGAGGCCACCCTGAACCACGCCAAGAGCGTGGGCATCCTGGGCGACG
GCTTCGGCACCAACCCTGGAGATGTCCAAGCGCGACCTGATGTGGGTGGTGCGCC

GCACCCACGTGGCCCGTGGAGCGCTACCCACCTGGGGCGACACCGTGGAGGTGG
 AGTGCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGGTGC
 GCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGAT
 GAACACCCGACCCGCGCCTGAGCACCATCCCCGACGAGGTGCGCGGCGAGAT
 CGGCCCCGCTTTCATCGACAACGTGGCCGTGAAGGACGACGAGATCAAGAAGCT
 GCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCG
 CTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGCCTG
 GGTGTTTCGAGACCGTGCCCGACAGCATCTTCGAGTCCCACCACATCAGCTCCTTC
 ACCCTGGAGTACCGCCGCGAGTGCACCCGCGACTCCGTGCTGCGCAGCCTGACC
 ACCGTGAGCGGCGGCAGCTCCGAGGCGGCGCTGGTGTGCGACCACCTGCTGCAG
 CTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCGCACCGAGTGGCGCCCCAAGCTG
 ACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCGAGCCCCGCGTGATGGACT
 ACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGAC
 GACGACAAGTGA

SEQ ID NO: 86

Cinnamomum camphorum C14 thioesterase coding region and 3x FLAG tag, with *Chlorella protothecoides* stearoyl ACP desaturase plastid targeting sequence

ATGGCCACCGCATCCACTTTCTCGGCGTTCAATGCCCCGCTGCGGCGACCTGCGTC
 GCTCGGCGGGCTCCGGGCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCGCGGGC
 GCGCCCCCGACTGGTCCATGCTGTTTCGCCGTGATCACCACCATCTTCTCCGCCGC
 CGAGAAGCAGTGGACCAACCTGGAGTGGAAAGCCCAAGCCCAACCCCCCCCCAGCT
 GCTGGACGACCACTTCGGCCCCACGGCCTGGTGTTCGCGCGACCTTCGCCATC
 CGCAGCTACGAGGTGGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAAC
 CACCTGCAGGAGGCGGCCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGAC
 GGCTTCGGGACCAACCTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAG
 CGCACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTG
 GAGTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCCTGGTG
 CGCGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATG
 ATGAACACCCGCACCCGCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAG
 ATCGGCCCCGCCTTCATCGACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAG
 CCCCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCC
 CGCTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGAC
 TGGATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCT
 TCACCATCGAGTACCGCCGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGAC
 CACCGTGAGCGGCGGCTCCTCCGAGGCCGCGCCTGGTGTGCGAGCACCTGCTGCA
 GCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCT
 GACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCGAGTCCAGCGTGATGGAC
 TACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGAC
 GACGACAAGTGATGA

SEQ ID NO: 87

C. reinhardtii β -tubulin:: *C. camphora* C14 thioesterase with *C. protothecoides* stearoyl ACP desaturase plastid targeting sequence::3xFLAG:: *C. vulgaris* NitRed expression construct element

gaattccttcttgcgctatgacacttcagcaaaaggtaggcgggctgcgagacggctcccgcgctgcatgcaacaccgatgat
gcttcgacccccgaagctcttcggggctgcatggcgctccgatccgctccagggcgagcgctgtttaatagccaggccccc
gattgcaaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatc

gcactccgctaagggggcgccctcttctcttctgtttcagtcacaacccgcaaactagtATGGCCACCGCATCCACT
TTCTCGGCGTTCAATGCCCCGCTGCGGCGACCTGCGTTCGCTCGGCGGGCTCCGGGC
CCCGGCGCCCAGCGAGGCCCCCTCCCCGTGCGCGGGCGCGCCCCCGACTGGTCCAT
GCTGTTCGCCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGACCAAC
CTGGAGTGGAAGCCCAAGCCCAACCCCCCCCAGCTGCTGGACGACCACTTCGGC
CCCCACGGCCTGGTGTTCGCGCGCACCTTCGCCATCCGCAGCTACGAGGTGGGCC
CCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCCGCC
TGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCCTGG
AGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCCGTGG
AGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTGGAGTGTGGGTGGGCGCCT
CCGGCAACAACGGCCGCCGCCACGACTTCCTGGTGC GCGACTGCAAGACCGGCG
AGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCCGCCG
CCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCCCGCCTTCATCGA
CAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGACT
CCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACA
TCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACCGTGC
CCGACAGCATCTTCGAGAGGCCACCACATCTCCTCCTTACCATCGAGTACCGCCG
CGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCGGCTC
CTCCGAGGCCGGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGCGA
GGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCCGCGG
CATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCACGACGG
CGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGATGActc
gaggcagcagcagctcgatagatcgacacactctggacgctggcgtgtgatggactgtlccgcccacactlgtlgecllgacclgl
gaataccctgcccgttttatcaaacagcctcagtggtttgatctgtgtgtacgcgttttgcgagttgctagctgctgtgtctatttgcga
ataccacccccagcatcccccttccctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtctgtctatccctcagcgtgtct
cctgtctgtctcactgccccctgcacagccttggttlgggctccgctgtatttctctgtactgcaacctgtaaaccagcactgcaatg
ctgatgcacgggaagtagtgggatgggaacacaaatgaaagctt

SEQ ID NO: 88

C. reinhardtii β -tubulin::*C. camphora* C14 thioesterase with *P.moriformis* stearyl ACP desaturase plastid targeting sequence::3xFLAG::*C.vulgaris* NitRed expression construct element

gaattcctttcttgcgctatgacacttcagcaaaaggtaggggcggtgcgagacggcttcccggcgctgcatgcaacaccgatgat
gcttcgacccccgaagctccttggggctgcatggcgctccgatgccgtccagggcgagcgtgtttaaatagccaggcccc
gattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcaactaccacttctacacaggccactcgagcttgtgatc
gcactccgctaagggggcgccctcttctcttctgtttcagtcacaacccgcaaactagtATGGCTTCCGCGGCATTC
ACCATGTTCGCGGTGCCCCGCGATGACTGGCAGGGCCCCCTGGGGCACGTTCGCTCC
GGACGGCCAGTCGCCACCCGCCTGAGGGGGCGCGCCCCCGACTGGTCCATGCTG
TTCGCCGTGATCACCACCATCTTCTCCGCCGCCGAGAAGCAGTGGACCAACCTGG
AGTGGAAGCCCAAGCCCAACCCCCCCCAGCTGCTGGACGACCACTTCGGCCCCC
ACGGCCTGGTGTTCGCGCGCACCTTCGCCATCCGCAGCTACGAGGTGGGCCCCGA
CCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCCGCCCTGAA
CCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCCTGGAGAT
GTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCCGTGGAGCG
CTACCCCGCCTGGGGCGACACCGTGGAGGTGGAGTGCTGGGTGGGCGCCTCCGG
CAACAACGGCCGCCGCCACGACTTCCTGGTGC GCGACTGCAAGACCGGCGAGAT
CCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCCGCCGCCTG
AGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCCCGCCTTCATCGACAAC
GTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGACTCCACC

GCCGACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACATCAAC
CAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACCGTGCCCCGAC
AGCATCTTCGAGAGCCACCACATCTCCTCCTTCACCATCGAGTACCGCCGCGAGT
GCACCATGGACAGCGTGTGCTGCAGTCCCTGACCACCGTGAGCGGCGGCTCCTCCG
AGGCCGGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGC
TGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCCGCGGCATCA
GCGTGATCCCCGCGGAGTCCAGCGTGATGGACTACAAGGACCACGACGGCGACT
ACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGATGActcgaggcag
cagcagctcggatagatcgacacactctggacgctggcgtgtgtgalggactgttgcgcccacacttgctgacctgacctgtgaatatcc
ctgccgcttttatcaaacagcctcagtggtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacc
cccagcatcccccttccctcgtttcatatcgcttgcateccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcct
gtcactgccccctgcacagccttggttgggctccgctgtattctcctggtactgcaacctgtaaccagcactgcaatgctgatgca
cggaagtagtgggatgggaacacaaatggaaagctt

SEQ ID NO: 89

Arabidopsis thaliana invertase sequence (with targeting signal from yeast suc2: *C. sorokiniana* glutamate dehydrogenase promoter and *C. vulgaris* nitrate reductase 3'UTR:3xFLAG)

GAATTCCGCCTGCAACGCAAGGGCAGCCACAGCCGCTCCCACCCGCCGCTGAAC
CGACACGTGCTTGGGCGCCTGCCGCCTGCCTGCCGCATGCTTGTGCTGGTGAGGC
TGGGCAGTGCTGCCATGCTGATTGAGGCTTGGTTCATCGGGTGGAAGCTTATGTG
TGTGCTGGGCTTGCAATGCCGGGCAATGCGCATGGTGGCAAGAGGGCGGCAGCAC
TTGCTGGAGCTGCCGCGGTGCCTCCAGGTGGTTCAATCGCGGCAGCCAGAGGGA
TTTCAGATGATCGCGCGTACAGGTTGAGCAGCAGTGTCAGCAAAGGTAGCAGTTT
GCCAGAATGATCGGTTGAGCTGTTAATCAATGCCAGCAAGAGAAGGGGTCAAGT
GCAAACACGGGCATGCCACAGCACGGGCACCGGGGAGTGGAATGGCACCCACCA
AGTGTGTGCGAGCCAGCATCGCCGCCTGGCTGTTTCAGCTACAACGGCAGGAGTC
ATCCAACGTAACCATGAGCTGATCAACACTGCAATCATCGGGCGGGCGTGATGC
AAGCATGCCTGGCGAAGACACATGGTGTGCGGATGCTGCCGGCTGCTGCCTGCT
GCGCACGCCGTTGAGTTGGCAGCAGGCTCAGCCATGCACTGGATGGCAGCTGGG
CTGCCACTGCAATGTGGTGGATAGGATGCAAGTGAGCGAATACCAAACCCTCT
GGCTGCTTGCTGGGTTGCATGGCATCGCACCATCAGCAGGAGCGCATGCGAAGG
GACTGGCCCCATGCACGCCATGCCAAACCGGAGCGCACCGAGTGTCCAACTGT
CACCAGGCCCGCAAGCTTTGCAGAACCATGCTCATGGACGCATGTAGCGCTGAC
GTCCCTTGACGGCGCTCCTCTCGGGTGTGGGAAACGCAATGCAGCACAGGCAGC
AGAGGCGGCGGCAGCAGAGCGGCGGCAGCAGCGGCGGGGGCCACCCTTCTTGCG
GGGTCGCGCCCCAGCCAGCGGTGATGCGCTGATCCCAAACGAGTTCACATTCAAT
TGCATGCCTGGAGAAGCGAGGCTGGGGCCTTTGGGCTGGTGCAGCCCGCAATGG
AATGCGGGACCGCCAGGCTAGCAGCAAAGGCGCCTCCCCTACTCCGCATCGATG
TTCCATAGTGCAATTGGACTGCAATTGGGTGGGGCGGCCGGCTGTTTCTTTCGTGTT
GCAAAACGCGCCAGCTCAGCAACCTGTCCCGTGGGTCCCCCGTGCCGATGAAAT
CGTGTGCACGCCGATCAGCTGATTGCCCGGCTCGCGAAGTAGGCGCCCTCCTTTC
TGCTCGCCCTCTCTCCGTCCCGCCACTAGTATGCTGCTGCAGGCCTTCTGTTCCT
GCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCCGA
CCGCCCCCTGGTGCATTCACCCCCAACAAGGGCTGGGGGCGCGCCAGCCACCA
CGTGTACAAGCGCCTGACCCAGAGCACCAACACCAAGTCCCCCAGCGTGAACCA
GCCCTACCGCACCGGCTTCCACTTCCAGCCCCCAAGAACTGGATGAACGACCCC
AACGGCCCCATGATCTACAAGGGCATCTACCACCTGTTCTACCAGTGGAACCCCA
AGGGCGCCGTGTGGGGCAACATCGTGTGGGGCCACTCCACCAGCACCGACCTGA

TCAACTGGGACCCCCACCCCCCGCCATCTTCCCCAGCGCCCCCTTCGACATCAA
CGGCTGCTGGTCCGGCAGCGCCACCATCCTGCCCCAACGGCAAGCCCCGTGATCCTG
TACACCGGCATCGACCCCAAGAACCAGCAGGTGCAGAACATCGCCGAGCCCAAG
AACCTGTCCGACCCCTACCTGCGCGAGTGGAAGAAGAGCCCCCTGAACCCCTG
ATGGCCCCCGACGCCGTGAACGGCATCAACGCCTCCAGCTTCCGCGACCCACCA
CCGCCTGGCTGGGCCAGGACAAGAAGTGGCGCGTGATCATCGGCTCCAAGATCC
ACCGCCGCGGCCTGGCCATCACCTACACCAGCAAGGACTTCCTGAAGTGGGAGA
AGTCCCCCGAGCCCCTGCACTACGACGACGGCAGCGGCATGTGGGAGTGCCCCG
ACTTCTTCCCCGTGACCCGCTTCGGCAGCAACGGCGTGAGACCTCCAGCTTCGG
CGAGCCCAACGAGATCCTGAAGCACGTGCTGAAGATCTCCCTGGACGACACCAA
GCACGACTACTACACCATCGGCACCTACGACCGCGTGAAGGACAAGTTCGTGCC
CGACAACGGCTTCAAGATGGACGGCACCGCCCCCCGCTACGACTACGGCAAGTA
CTACGCCAGCAAGACCTTCTTCGACTCCGCCAAGAACCGCCGCATCCTGTGGGGC
TGGACCAACGAGTCCTCCAGCGTGAGGACGACGTGGAGAAGGGCTGGTCCGGC
ATCCAGACCATCCCCGCAAGATCTGGCTGGACCGCAGCGGCAAGCAGCTGATC
CAGTGGCCCCGTGCGCGAGGTGGAGCGCCTGCGCACCAAGCAGGTGAAGAACCTG
CGCAACAAGGTGCTGAAGTCCGGCAGCCGCCTGGAGGTGTACGGCGTGACCGCC
GCCCAGGCCGACGTGGAGGTGCTGTTCAAGGTGCGCGACCTGGAGAAGGCCGAC
GTGATCGAGCCCTCCTGGACCGACCCCCAGCTGATCTGCAGCAAGATGAACGTGT
CCGTGAAGTCCGGCCTGGGCCCCCTTCGGCCTGATGGTGCTGGCCAGCAAGAACCT
GGAGGAGTACACCTCCGTGTACTTCCGCATCTTCAAGGCCCGCCAGAACAGCAA
CAAGTACGTGGTGCTGATGTGCTCCGACCAGTCCCGCAGCTCCCTGAAGGAGGA
CAACGACAAGACCACCTACGGCGCCTTCGTGGACATCAACCCCCACCAGCCCCT
GAGCCTGCGCGCCCTGATCGACCACTCCGTGGTGGAGAGCTTCGGCGGCAAGGG
CCGCGCCTGCATCACCTCCCGCGTGTACCCCAAGCTGGCCATCGGCAAGTCCAGC
CACCTGTTGCCTTCAACTACGGCTACCAGTCCGTGGACGTGCTGAACCTGAACG
CCTGGAGCATGAACTCGGCCCAGATCAGCATGGACTACAAGGACCACGACGGCG
ACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGATTAATTA
ACCGGCTCGAGGAGCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCG
TGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCG
CTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTIGCGAGTT
GCTAGCTGCTTGTGCTATTTGCGAATAACACCCCCAGCATCCCCTTCCCTCGTTTC
ATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATCCCTCAGCGC
TGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTA
TTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGT
AGTGGGATGGGAACACAAATGGAAAGCTTGAGCTC

SEQ ID NO: 90

Translated amino acid sequence of the *Arabidopsis thaliana* invertase expression construct

MLLQAFLFLLAGFAAKISASMTNETSDRPLVHFTPKNKGWGRASHHVYKRLTQSTNT
KSPSVNQPYRTGFHFQPPKNWMNDPNGPMIYKGIYHLFYQWNPKGAVWGNIVWAH
STSTDLINWDPHPPAIFPSAPFDINGCWSGSATILPNGKPVILYTIDPKNQQVQNI AEP
KNLSDPYLREWKKSPNLPLMAPDAVNGINASSFRDPTTAWLGQDKKWRVIIGSKIHR
RGLAITYTSKDFLKWEKSPEPLHYDDGSGMWECPDFFPVTRFGSNGVETSSFGEPNEI
LKHVLKISLDDTKHDY YTIGTYDRVKDKFVPDNGFKMDGTAPRYDYGKYASKTFF
DSAKNRRILWGWTNESSSVEDDVEKGWSGIQTIPRKIWLDRSGKQLIQWPVREVERL
RTKQVKNLNRNKVLKSGSRLEVYGVTAQAADVEVLFKVRDLEKADVIEPSWIDPQLI
CSKMNVSVKSGLGPFGMLV LASKNLEEYTSVYFRIFKARQNSNKYVVL MCSDQSR S

SLKEDNDKTTYGAFVDINPHQPLSLRALIDHSVVESFGGKGRACITSRVYPKLAIGKS
SHLFAFNYYGYQSVLDVNLNLAWSMNSAQISMDYKDHDGDYKDHDIDYKDDDDK

SEQ ID NO: 91

Prototheca moriformis FatB/A promoter/5'UTR

CCTGTGATCGAAGAGAAGGAGACATGTGTACATTATTGGTGTGAGGGCGCTGA
ATCGGCCATTTTTTAAAATGATCACGCTCATGCCAATAGACGCGGCACATAACGA
CGTTCAAACCCCGCCAAAGCCGCGGACAACCCCATCCCTCCACACCCCCACAC
AAAGAACCCGCCACCGCTTACCTTGCCACGAGGTAGGCCTTTCGTTGCGCAAAA
CCGGCCTCGGTGATGAATGCATGCCCCTTCCTGACGAGCGCTGCCCGGGCCAACA
CGCTCTTTTGCTGCGTCTCCTCAGGCTTGGGGGCTCCTTGGGCTTGGGTGCCGCC
ATGATCTGCGCGCATCAGAGAAACGTTGCTGGTAAAAAGGAGCGCCCCGGCTGCG
CAATATATATATAGGCATGCCAACACAGCCCAACCTCACTCGGGAGCCCGTCCCA
CCACCCCAAGTCGCGTGCCTTGACGGCATACTGCTGCAGAAGCTTCATGAGAAT
GATGCCGAACAAGAGGGGACGAGGACCCAATCCCGGACATCCTTGTGATAAT
GATCTCGTGAGTCCCCATCGTCCGCCCACGCTCCGGGGAGCCCGCCGATGCTCA
AGACGAGAGGGCCCTCGACCAGGAGGGGCTGGCCCGGGCGGGCACTGGCGTCG
AAGGTGCGCCCGTCGTTGCGCTGCAGTCCATGCCACAAAACAAGTCTTCTGACG
GGGTGCGTTTGCTCCCGTGCGGGCAGGCAACAGAGGTATTCACCCTGGTCATGGG
GAGATCGGCGATCGAGCTGGGATAAGAGATACTTCTGGCAAGCAATGACAACCT
GTCAGGACCGGACCGTGCCATATATTTCTCACCTAGCGCCGAAAACCTAACAAAT
TTGGGAGTCACTGTGCCACTGAGTTCGACTGGTAGCTGAATGGAGTCGCTGCTCC
ACTAAACGAATTGTGAGCACCGCCAGCCGGCCGAGGACCCGAGTCATAGCGAGG
GTAGTAGCGCGCC

SEQ ID NO: 92

Prototheca moriformis NRAMP metal transporter promoter/5'UTR

ACTAATTGCAATCGTGACGTAATCATCGATATGGTCACAAGTAGATCCCCTACTG
ACACCTCTCGTACATGTAGGCAATGTCATCGGCGCCGTCCTGCTGACCGATGCC
GACGTAGCAGAGCAGACCCGGGCCGATCTGGGATACGAGCCGGCCCTCCACCTG
CGCTCGAGGTGGAATCAAGTAAATAACCAATACACTTTTCGACACCACACAGAG
TTGCACGGACGGTGGCGTACCTCTACGCTCGCGCTCTTCACGCGCTGGACGACCG
CACGCATGAGCCCGGGTGGCTTGGTCTGGGCTGCAAAAATGCACAACAAACAAG
TATCAGACGCTCATGGATGCACACGCGCTCCCAAGCACGCTCAGACTAAATATTA
CAGTAGCTCGTATCTGATAAGATATCGAGACATACCGCTCAACTACCCGCAAAAC
TGCGCCCCGCCAGGTGATGCGCACAGGGCCCCACCATGCGATCCATCGCATCGCT
CCTCGAGGGCGCTATCACGTGGCCGGAGAGCGTTCACAGCGTACGCCACTGTATC
TGGGCGGTATGCGGTCCGTCAACATGGAGACAGATACCCGCACCACCACCTTGC
AAGCTCTTCATATTGGAAGTAGAAAATTGTAATTGTATCATCGCACGAGGGGCC
AACTTGCCGTCGGCGAGCTGGGCGACGAACACCACCTGGACGTTGTCGAGACTC
GCTCGTGCCGTGCGCCGGGCCGCTGGGTATCCAGACCGTCGCC

SEQ ID NO: 93

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

CAACGACAACCAGCAGGCAACTCGGTACGCGACCCAACACGCGAGTCAAATTGT
TGCGTGTTCTTGCTTGTCTATTTACTGTGATAGCAAGACTGTCGGTCAGTCAATA
CCGCGGTGCGCACGTCGGGGTGCCAAGCCTAGCAGAGCACGGGACGGCTGGTGC
TGTGCGCCAGCTCAGCTCGCTTCGCGACCAATTGTAGGACCGGCAAAGTCACCAA
AACATGCCAGCGGTGCGATTCAATTGGTCATGAGCTCTACAAAATTGTTTTGTGC
GTCGCGCAGGTATCCAACGGCGCGGCAGAGAAAGTTTGACAGCTCTCGATTTTCAT

CTCGGAAAAATGGGGAGAATTTATGACACACAAGTGCGCAGGCGGCCAGGCGG
 CCAGCATATTCTGGCGTGACCTGGGCGGCCACAAAATGCTTGGATGCACTCTAA
 AATAATTATATTTGCCATGAACAAGGGAAGAGTTACCGCACCCAGCCCTAGACTT
 GGGCGCCCCGAGCAAGGTTACGTCAAGCCACCTTCGCCCATCGCCCAACTCCGTAT
 TCCCCGACAGCCGCACGTGGCCCTCGCCGGAATGAACCCTGAATCGGCATCACG
 CCACGCGTTCGCCAATCGTTCGCTCTCTGGCTTCATCGGCCTGCGCCTTCACGTC
 GTGGTCACGACAGTGCATTCTACTTCCATTTCGACCTCGGCACACACTTTTACG
 CATCGCCTACCCTTGCTGCGGCAGTCTAGGGTCACTTTGCAGCCATGGGACAGTG
 CTACACCACCGTCGGTGCGCAAAGCTATTTCAAGTGAACCGTGGGCGGAAAAAA
 GGAATGTACACTGTCTCAACCGACTCCTACAATTGTTTACCATGCAGATCAGAGC
 TCGACGGCCATCATCGAGCAGGTGTGGGGCCTTGGTGGCGCGGCGCGGGGCCCC
 AGGGCGTCGCAGGCATTGATGGCACTCTGAGACTTTCGCACGCGCATGAGGGAC
 CCCATCAAGAGAAGAGTGTGTCTTTATGTCCCATTCATGATGATGTATCTTGTG
 ATTGTGCGAGTTTGGCAAGTTTAACCGGATCGCCGCTCCAGGTGTGGCGTGGCGG
 ATTTTTCTAGGGGTGCTTGAGCAGTCG

SEQ ID NO: 94

Prototheca moriformis SulfRed Sulfite reductase promoter/5'UTR

GGCCCAGGGCCCTGCGGATGGCCACACCAGATCTAGCCTCTCTTATGCCATGCC
 CGCCTCGCTGCCCCGTCGTATCCCCCGCCGATCCGCGCGTAGGGGACCGCGGCCT
 GACCCACGCCACGAAAGAGCTTTGCTCCTCAATTTCTCGCCAACAGAACCGTATC
 AAACGCTCAACGCCTATCCCGAACAATCCGTATTCACACCAAATCGAGTATACCG
 GACTGGTTTGCCTAGTCTTGAAGGAAATGATCCCGTCCATGCTCGGAAGGGGGA
 GCGGGCGGAGGATCCTACTCATCTCTGAAATGGGATTGGTCCGAAGATGGGTG
 GGCAAGCACGTGCCAAACCCCGAGCGAGTTGCTGACGAGCAGGCTCATCCAATCC
 CCCGGCGAATCCTCCCTCACGCCCCGCATGCATACAAGTCCCTCCCACACGCCCC
 CTCCCATCCATTTCGCTTGGTCCGAACGCGAGCGGCGTCGAGGCGGACCACTTG
 CTCCGCAGCGCCGTCTGGGTCTCCACCCACAGCGGCTTTGCTGCCAGAGGCACC
 CCCCTTGCCCCACCTCCTCTTGACGCC

SEQ ID NO: 95

Prototheca moriformis SugT Sugar transporter promoter/5'UTR

CCAGGCAGGCGGTAGGGTTGCCGATTGCTTGAGCGAATTGGAAGATATAATTTTT
 TGTGGTGTCCCTGGACGCTGTTTGTGGCGCTCCTTTTTGGAGAAGATTGCGTGGG
 GGAGCTTTCCATGTACCACGCTTCCTTCTGAAAGGATTCTGGCCGAGTCCTGATG
 AGCCCAAAGAAAACACCTGCCTTTCAGTGCTGGCACTCTGAAAACGTCAACAGA
 TGATTATACATGTCACAAAAGGCAGCCGATTAGGAACGGGAGCTCTGGCCGTTC
 GTTTGGCTGCCTGGGCTGATTGAAGTGATCCACCCTGTTTGAATGAAGGCGGTGCG
 AGTCGAATTATCGACCGGAGCTGTCGGGAAGGCGTCCGGGGCAGAGTGAGGTGC
 TGCGGCCTGGTTGTCGTTCAAAAAGACCCCGGTAGCCCAACAATCACGAACGAA
 AGGAATATAATTGCTTGACATACTATACATTTCAGTTTCTATGTGGCGGGTAGACAA
 GTCTCATGGGCTTCTAAAGGCTGTCCCTTGAAGGCTACTTATAAAAACCTTGCTGC
 GCCATGGCACGGATCGCGCTTGCGCAGGCTGCAACCCTGCGCGCAAGGTCAAAT
 ACACAGCAAAAAGATACTAACAGAATTTCTAAAAACATTTAAATATTTGTTTCGAC
 CAGCCAATTGTGGTCGTAGGCACGCAAAAAGACTTTGTTTTGCGCCCACCGAGCAT
 CCACGCTGGCAGTCAAGCCAGTCCGATGTGCATTGCGTGGCAGCATCGAGGAGC
 ATCAAAAACCTCGTGCACGCTTTTCTGTCAATCATCATCAACCACTCCACCATGT
 ATACCCGATGCATCGCGGTGCGCAGCGCGCCACGCGTCCAGACCCGCCCAAAA
 ACCCAGCAGCGGCGAAAGCAAATCTTCACTTGCCCGAAACCCCGAGCAGCGGCA
 TTCACACGTGGGCGAAAACCCCACTTGCCCTAACAGGCGTATGTCTGCTGTCACG

ATGCCTGACAACGGTATTATAGATATACACTGATTAATGTTTGAGTGTGTGCGAG
TCGCGAATCAGGAATGAATTGCTAGTAGGCACTCCGACCGGGCGGGGGCCGAGG
GACCA

SEQ ID NO: 96

Prototheca moriformis Amt03-Ammonium transporter promoter/5'UTR

GGCCGACAGGACGCGCGTCAAAGGTGCTGGGCGTGTATGCCCTGGTCGGCAGGT
CGTTGCTGTTGCTGCGCTCGTGGTTCCGCAACCCTGATTTTGGCGTCTTATTCTGG
CGTGGCAAGCGCTGACGCCCCGCGAGCCGGGCGGCGGCGATGCGGTGTCTCACG
GCTGCCGAGCTCCAAGGGAGGCAAGAGCGCCCGGATCAGCTGAAGGGCTTTACA
CGCAAGGTACAGCCGCTCCTGCAAGGCTGCGTGGTGGACTTGAACCTGTAGGTCC
TCTGCTGAAGTTCCTCCACTACCTCACCAGGCCAGCAGACCAAAGCACAGGCTT
TTCAGGTCCGTGTCATCCACTCTAAAACACTCGACTACGACCTACTGATGGCCCT
AGATTCTTCATCAACAATGCCTGAGACACTTGCTCAGAATTGAAACTCCCTGAAG
GGACCACCAGAGGCCCTGAGTTGTTCTTCCCCCGTGGCGAGCTGCCAGCCAGG
CTGTACCTGTGATCGAGGCTGGCGGGAAAATAGGCTTCGTGTGCTCAGGTCTATGG
GAGGTGCAGGACAGCTCATGAAACGCCAACAATCGCACAATTCATGTCAAGCTA
ATCAGCTATTTCTCTTCACGAGCTGTAATTGTCCCAAAATTCTGGTCTACCGGGG
GTGATCCTTCGTGTACGGGCCCTTCCCTCAACCCTAGGTATGCGCGCATGCGGTC
GCCGCGCAACTCGCGCGAGGGCCGAGGGTTTGGGACGGGCGGTCCCGAAATGCA
GTTGCACCCGGATGCGCGGCGCCTTTCTTGCGATAATTTATGCAATGGACTGCTC
TGCAAATTTCTGGGTCTGTGCGCAACCCTAGGATCAGCGGCGTAGGATTTCTGTAA
TCATTCGTCCTGATGGGGAGCTACCGACTACCCTAATATCAGCCCGGCTGCCTGA
CGCCAGCGTCCACTTTTGCCTACACATTCCATTCTGTGCCCAAGACATTTCAATTGTG
GTGCGAAGCGTCCCCAGTTACGCTCACCTGTTTCCCGACCTCCTTACTGTTCTGTC
GACAGAGCGGGCCACAGGCCGGTTCGCAGCC

SEQ ID NO: 97

Prototheca moriformis Amt02-Ammonium transporter promoter/5'UTR

TCACCAGCGGACAAAGCACCGGTGTATCAGGTCCGTGTCATCCACTCTAAAGAG
CTCGACTACGACCTACTGATGGCCCTAGATTCTTCATCAAAAACGCCTGAGACAC
TTGCCCAGGATTGAAACTCCCTGAAGGGACCACCAGGGGCCCTGAGTTGTTCTT
CCCCCGTGGCGAGCTGCCAGCCAGGCTGTACCTGTGATCGGGGCTGGCGGGAA
AACAGGCTTCGTGTGCTCAGGTTATGGGAGGTGCAGGACAGCTCATTAAACGCC
AACAATCGCACAATTCATGGCAAGCTAATCAGTTATTTCCCATTAACGAGCTATA
ATTGTCCCAAAATTCTGGTCTACCGGGGGTGATCCTTCGTGTACGGGCCCTTCCCT
CAACCCTAGGTATGCGCACATGCGGTGCGCCGCGCAACGCGCGCGAGGGCCGAGG
GTTTGGGACGGGCCGTCCCGAAATGCAGTTGCACCCGGATGCGTGGCACCTTTTT
TGCGATAATTTATGCAATGGACTGCTCTGCAAAATTCTGGCTCTGTGCGCAACCC
TAGGATCAGCGGTGTAGGATTTCTGTAATCATTCTGCTCCTGATGGGGAGCTACCGAC
TGCCCTAGTATCAGCCCGACTGCCTGACGCCAGCGTCCACTTTTGTGCACACATT
CCATTCTGTGCCAAGACATTTCAATTGTGGTGCGAAGCGTCCCCAGTTACGCTCAC
CTGATCCCCAACCTCCTTATTGTTCTGTGCGACAGAGTGGGCCCGAGAGGCCGGTCG
CAGCC

SEQ ID NO: 98

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

CGAAGGGGTCTGCATCGATTGCGCGGTCTGGAGGCCAGCGTGAAGTCTCGCGA
AAATGCTCTGCCGTGTGCGGCTCTGGCTGGGGCGGCCAGAGATCTCACCGTGCCA
CACGCAACTGCCGCACTCTGTGCCCCGCCACCTGGCGCGCACATGCGACCTCTTCC

CCGTCATACCCTCTCCTCATGTGATCTTTCCACACGAGTGACGCAGGTGCGCGGA
 GTGGAGGGAATCAGGACGTTTTCAAGGTACCTGCTCGAGCCGTACCAACAGCTG
 CCGCCCGGCAAGGAAGAGATCGAGGCAGAGATTGCCCGGCTGGAGGCCCGGATA
 ACGGAGCTCAAGAGCAAGCTGTCCGAGTGAGACCGCCCAGGTGCACGTGTGAC
 TCGCTATGACATGTACTCGACACAACATGAGGAATTCATCGAATTTGTAGGAAGC
 GGGCATTGGTACGGGAGTGGGAAAGCGAAAAAACCTCCCTCCGGCAGTGCCATC
 TGCCGGAGTCGAACGTTGATAGGGTTCTCGTGACAGGGTGTGACCTCTCAGCCTT
 GCATCAATTAAACGCTATAGACATTATCAGTAACCGTGAATCCCGCATTGGATGC
 CACCCGCGCGACCATTGGGGACCTGCATTACAGATCTAGGTGAGATGACAGCGA
 GGCAACTTCGGCCCCGCGGCCAGCTTGCGGCGCACCAATATTGGTCACGGGAAG
 CCACACACCGACCATAAATGAATACTTGTAAAGCTATGTCAACCGATCAATGGCGT
 CGAAAGTGTGCCACGAGGATCCATCTGGCGGGGCGGCGTGGCGCACAAGCGCAG
 TCGCAATTTCTCGGACCCATCTGACCTAGGCCAGCGCCGCGGGAGAAATCCCCG
 GCGGGTCTCCACGCAGTAACCCTAATGAGTATCGAGCGCCGACCAATTACACCA
 TCGCCCCCGAAATCCTTCCGACATTATTATTATCTTTTAGATCTTGGAACAGACTC
 TGCCAACC

SEQ ID NO: 99

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

AGAGAGCGGAGGTGGGGTTGTGAGGTGGGGTTGCTGACCAGGAGCTCGCGTCGC
 CGAGCGCGACTCGCACACGGTCCAGTTACCCCCCTCCGCCCAAACGCAAGCCT
 CCCATCTTGATGCCTTTCCGGCCACCTATACTATTTCTTAGTTCGCTGTAACATCC
 AGACCGTCCTGAATAATAACAATGCCCTGTGTCAAGTGCATTCTAAAAAAATTC
 TGTCCTCAACCAACAATCCCACCTGAAATACCACCAGCCCTGCCCAGTACACTCTT
 CCAATACCATCTCCCTACCTCCACGCGCAAGCGACCCCCATGCGCGACCAGGCTC
 GAAAGTGATTTATGACTTGAGACGAGCGAGTGGCGGCGCGGTGCGACTGCCTTTTC
 ATCACGTGCCGTACGTGCGCGACCGCTAGGGCTTTGCACGGCAACGCACGGCTTC
 GCCAACCCGACCAGCCAGGACCTCGACTACTCTACCGCGAATTCGCCTCAAGAA
 GTCGCCAAATGTGCCATACACCATTCCTTACAGCACTGTTCAAACCTTGATGCCAA
 TTTTGACATTCGGGTTGCTCGTTGGCTGCGCCCACATCGGCCGTGAGTGCAGCAG
 GCGGGATCGGACACGGAGGACGCGGCGTCACGCCCCGAACGCAGCCCGTAATCTC
 TACATCAACACGACGTGTTGCGTAATCCCGCCCCGGCTGCGCATCGTGCCAACCCA
 TTCGCGATGGATGGTTCGGAAAAATGGTGTGCCAACTGCCCTGAGGGAGGCTCTCG
 CGAAACGGGCACGTCCCTGAAACCGAAACTGTGGCCTTGTCGTGCGCCACGCAA
 GCACGTGGACCCTAAACACCAAGAAAATCAGTAAACAAGGTTGACATCCTCTAC
 GGGCGAATTGTTTGCCCAACCTTCATCGCACACTGCCATTATAATGCATCTAGC
 TCGGCGACAAGTTTAGAAAAGGCAGGCTGCATTGTTCCATTTGCGCGTGGCGGCG
 TGGGTGCCCATTTTACGAGGTTTGGGCTCCCGGGCAGCGACCGAGCCAGGTCTGA
 GTCCCTCTCGCCCGTCGACAATGTTGCGAACCCCAACAAGCGGCTAACAACAACCTT
 GATGGTACCTGTACACTGCCAATTCCTTCTTCCCCGGCCGAGGTTTACACGTGAT
 GGCCATGGCTTCGCATTCAGGCCGACTTCCCATTCGACTTTCCAGAGGGTCCGC
 GGACGCTGGGGGTTGGCTGCCTGAGGCCACCCCTTTGTTCCCCGCGTCCCGACAA
 ACACAATTGCGTTACATAAGGGGGAGCCGCCCCCGTTTCAGAGTGCAGAAATCTTT
 CACTATATTTCCAGTCGTCAGCGAAATCAAGT

SEQ ID NO: 100

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

GATGGTGGGGTGTCTGCCTTGGGCTGGGTGATGGAGGCTGGTGGTGCGCGGGTTT
 CCTGATGCATTCTATCTACGCAAGTGTGATGGTGTCCATTCCACACACCAGTACAC
 CCTTACACTAAGGATCCATCCCTCCTTCCCTCTTCAGGACTACATGGACCCACG

AGCTACCGACCGGGCTTTCTCAAAAACGTCAAGGTCATGTTTGACATGCGGGACG
TGGTGGACGACGTGCAAGGTGCGTCCGGAGTGCGCGCAAATGAGCAAGTCGGGC
AATGTGTCGGGGTGGGCACCGGGGCTGGAGATCCGCGATCCCCGAGAAAACGCC
GTACCACCCCCCGCGCTATTCCCTCGATTGCGCGCAGATGTGGTGACCGACACGG
GGGACAACCTGGCGGACATGGGGCGCCGGACCTGGAAGCACGCCAAGTCGCACA
CGGGGAGGCTCGTGAGTCCCCCCCCATCGTACCTCAAGGGTCTCTTTGGTCGCGA
TCCAAAGTACGCTGGTGGCATGGCATGCCCCGAAATGAACATCATGTGTGATCTCC
GATTGCCAATGGCCACCTCCACGGACCACCTTGCAGGCGGAAGCGCAATCCAGG
GCCCCGAGCCTGACGAGGACGGAGACTCCTCGTCCAGCGCGGGGTCCCCGACCCG
ACGCAGCAGCCGACCCCTGCTAACCCGGCAACGATCGGACCAGCAACCTTGCTG
TAGTTCGGATCCGTGATGACGGGCATTGCCGCCGCTCGATCCGCTTTGATGACTG
TCTATTATTTGCGCGGAGCCCCCTCGGAACCCTACCCCGCTCTTGCAAGCCCCCTG
CATCGGAGATCCTCGTGCGCCCCGCCATGACCCCACTGGATTGCCCAACATCCTTC
TTTATCGTGTA AAAATGTGATTCTCGGCTGCAATCGACTGGCCTTCGCTTCTGGCC
CCAAGAGGGCTCGAACGTGCGGCAGCGAGGGCGCTGACACACCCAAGCCCTAGG
GCTTTCAACGTGCGGTGCCAGGCCGGATAGGGGGATCGCCTCCTTTCCACCACCC
ACCTACGAGGGATTGAGTCGGCTTCCAGCTCAGCTATTCGGCCGCGCCCCCGGC
CCTGCAGACGTCCTCCAGTTTCCGAACAGGTCGCTCTCAGAACACCTGCCGCGGC
TGCGATACGGCAGGCTCTCAAAGCGTCGAC

SEQ ID NO: 101

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

CGCGTGGAGCGGTGCGTGCGGATGCCGCGCGCCTGCCAAGGCCTTTTGTATGCCT
GGCCTGGGAAGTTTCTGACTGAAGCATCTTCAAGATGCTCTCTCACGACCAGCG
ACACCAACACCGTCACTTTTTGCCCCCTCCTGCCGCAGGTGCCACTTTCTACTTTGA
CGTCTTCTCCAGGCGGTACATTGCGGGACTGAGCGCCAATTCGGCCAAGAACAG
CGCTGTCGACTTGAGGAGGCAGGGGTCCGTCGACTCTGCCGAGTGACACGCCTTC
GACCCGACTGTACTACGGCCTGCTGAAGAGTGGGTCTCGCCGGCCGGCGTGACC
GGCCCTGTGCCACAATCGACCATCTATTGCTCCTTGTCATCTGGCGCCGTCAAT
TGCCCCGCGACTTGACGGCAACTGGCTCGATCGAGTCGTATTGAAAAAGCACGTTT
TGTCCTACAGGGCCGCGGTCCGTTACCAACGTGGTTCTCGTTAGGTTTTCGTCGG
GCGGTGGTGCGCGAACTGTCCGATGCCATCCCGGCAAACCCAGCAAGGTGCGC
AGTCTGGTTCTGACGCAATAGAGTGCGTTTTGGGCCAGTCTAAAAATTCGTCTGG
CATGACGTGGCTCCACATCGTACCCGGAGCCCTGCTTGGTAATGTGAGGCACCGG
TGCCAACTCCATTATGGCAGGCATCGAGCGCGCAGGTGAGTACATGACCTTCCGT
GAATTGGGAAGGCGAGCTTGTGTAACGCCTGCGATCGTGCCAGTGAGGCATCGT
AAACTCAAAATATTTTGTAGAAAGTGTCTGATGCCTGGTGAGGCTGCGTAGGGCA
AGGGCAAGCCCTTGGCAGATGGGTAATGGGTCCGGACCTCACAAACAGCAACCCC
GCGTCCCCCTTAGGGCCCCCTGAGGCTCGATGGCAGGGCCAGCGAGCCCCGCGGCC
AAAGGGCGCCATCCACGGTCGCCCCAACGACTCCACGGGTCTTATACCTCATCTT
GAATGGCACTAAAACTATAGAATATCGGGCACTGGTGGGCGTCTGGGGTACAG
CTGGCCGAGCGCAGTGGC AAACCTAGGTCCCGCCTCAAGGGCGATTCCCGGGT
CAATGACACGCAAGCAAGATCACATGGCGCGGTCCGCCTCGCGGCTCCACACCC
AGGCCCTAGTTTCGCAACCCATAAATATCGCCCCGATACCATCATAAGCCAGCAA
ATAATTTTTTATCAGAGTTCCAAACCTCCTCAGCTGTGGGAAACCAGCCCACTCT
GAACG

SEQ ID NO: 102

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

CCGAGCAGTTCATGGCCAAGTACAAGGACTAGAGACCGGAGGTCGGTAGGCTGA
 ATGGAGCTGGCGTCGTCTGTGCGCGACGTGCACGCGATGCGATACTACGACCCCA
 CAAACGCATGCCTCCCATCTTGATGCCTTTCCGGCCATTTATACTATTTCTCATTT
 CGCTGTAACATCTTGAATAATAGAATTGCCCTGTGTCAAGTGGATTCCAAGAAAT
 ATTCTGTCCCAACAAAACAACCCAACTGAAAAACAACCTCAAATACCACCAGCC
 CTGCCCACCTGCCCAGTACACTTTTCCAATACCATCTCCCTACCTTCACGCGCAAG
 CGGCACCCATGCGCGACCAGGCTCGAAAGGATTTACGACTCAGGACGAGCGAG
 TGGCGGCGCGACCGCCTGCCTGTTTCGTACGTGCCGTACGTGCGGCGACCGCTAGA
 GCTTTGCCTGGCAACCCCCGGCTTCGTCAACCCGGCCAGCCAGGATCTCGACCAC
 TCTACCGCGAAATCGCCTCGAAGAAGTCGCCAAAAAGTGCCGTACACCATGCTTCGC
 AGCGCTGTTCAAACCTTGATGCCAATCTTGACAATCAGGTTGCTCGTTGGCTGCGT
 CCACATCGGCCGTGATTGCAGCAGGCGGGGATCGGACACGGAGGACGCGGGCGTC
 ACGCCGCGAACGCAGCCCGTAACCTCTACATCAACGCGATATGTTGCGTAATCCCG
 CCGGGCTGCGCATTGTGACAACCCATTTCGCGATGGATGGTCGGAAAAATGGTGTGC
 CAACTGCCCTGAGGGACTCTCTCGCGAAACGGGACAGTCCCTGTATCCGAAACTG
 TGGCATGGCCTTGTCGACCACGCAAGCACGTGGACCCTAACACCACGAAAAATAA
 GTAAAAAAGGTTGACATCCTCTACGAGCGAATTGTTTGCTCGACCCTTCATCGCA
 CACTGTCAATTATAATGCATCTAGCTCGGCGACAAGTTTAAAAAAGGCAGGCTGCA
 TTATTCCATTTTGGCGTGGCGGCATGGGTGCCCATTTTATGAGGTTTGGGCTCTTG
 GGCAGCGACCGAGCCAGGTTGAGTCCCTCTCGCCCGTCGACAACGTTCCAAAGC
 CCATAAGTGGCTAATAAACAACCTTGATGGTACCTGTACACTGCCAGTTCCTTCTT
 CCGCGGCCGAGGTTTACACGTGATGGCCATGGCTTCGCGTTTCAGGCTGACTTCC
 CATTCCGACTTTCCAGAGGGTCCGCGGACGCCGGGGGTTGGCTGCGTGAGGCCC
 ACCCTTGTTCCTCCGCGTCCCGACAAACACAATTGCGTTACATAAGGGGGAAGCC
 GCGCCCCGTTTCAGAGTGCAAACATCTTTCATTATATTTTTCAGTCGTCAGCGAAAT
 CAAGTATGTGCTGACAGGCATGAAGGCC

SEQ ID NO: 103

KE858 Homologous recombination construct SZ725

GCCCTTTGTCATCGTTGGCATGCTTTTTGCGTATGTACCATATGTTGAATGTATAA
 TACGAACGGTTGACCGTCTGAGATGCGAGCTTTGGGTCTTGTCAAATGCGTGGCC
 GCACGGCTCCCTCGCACCCAGCCCCGAGGCGTCGCGCACCTGGCGAGGAGCAGA
 CCCACGCCAAGAAAGTCTAGTCCAGCATGTAACAACATCAGGCAATGTGACGTT
 TTCGGTTCCCGATTTCTCTGCCGCTCTTTGACGGCAGGCACGGGCGAGCAACCGG
 CGGCGCTCGCGTCAGGCACGATGGATGCGGCGCTGCCACCTGTCAATGTACCCC
 ACCAGTCTGTGATCGCTACAAGCAACCTTGTGCTCCACATTCCCACCTTGACAGAC
 AGTCTAGTCGATTTTGCCAAGCTGGATGTGAGGATTGGCCATATCTTGAGGGCCA
 AGATTCACCCGGATGCTGATGGGTACGTACGCGAGCCAGGCAGGCAGCTGCGTT
 GACTTTCTGATTGGCACAAAGCTTTGGCTACTCTCAATACCAACCACGTGCCCCCT
 TCTGCACACCTGCTTCCTTCTGATGACCACTCGCCACGCATGTCGCAGTCTGTACG
 TCGAGCAGATCGACCTCGGCGAGGAGGGGGGCCCTCGCACCATCGTGAGTGGCC
 TGGTCCGGCACGTGACCCTGGAGGACCTTGTGCGCCGGCGGGTGGTGGTGTGCTGG
 CCAACCTCAAGCCTCGGAGCATGCGCGGGGTCAAATCGGCTGGGATGCTGCTCT
 GCGCCGCCAACGCGGATCACACCGCGGTGGAGCCGCTGCGGGTCCCGGACGCCG
 CCGTGACGGGGGAGCGGGTCTGGGCGGGGGACGAGGCACTCCTGTCCACGGAGC
 CTGCCACACCCAATCAGGTAAGGACACGTTATTGGTGCGCATGGTGCGAATGCGT
 GGTCTGACCTGCTGTGGGTATGTGTTGTGGGATTGGAAACCGAATGAGGGCCGTT
 CAGGATTGAGCCCTTGGCCCCACCTGCTCATCCTCTCACGCCCCGAGGTCCAGA
 AGAAGAAAATCTGGGAGGCAGTACAGCCGCTGCTGAGAGTGAACGCCCCAGGGG

ATCGCTACTGTGGCAGGAGAGGCTATGGTGACCAGTGCGGGGCCACTGACCGCG
CCCACGCTGGTTGACGCCGCGATTTCCTGACGCGAGCGACTGATTCTTGACCTTT
GAGAAGCCACCACAGCACCATTTTCATTGTTTCATCCTTGATTTCAGTACGACTTCT
CACCATTTTCAGTACTGTAGGACCCCCAAAATAGTGTGATCACGCTCGCAAGGCAC
CTGTGTGATCACGGGGAAGGGCGAATTCCTTTCTTGCGCTATGACACTTCCAGCA
AAAGGTAGGGCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGA
TGCTTCGACCCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCC
AGGGCGAGCGCTGTTTAAATAGCCAGGCCCGCGATTGCAAAGACATTATAGCGA
GCTACCAAAGCCATATTCAAACACCTAGATCACTACCACTTCTACACAGGCCACT
CGAGCTTGTGATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCAC
AACCCGCAAACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGC
TTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTG
GTGCACTTCACCCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTAC
GACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACC
GTCTGGGGGACGCCCTTGTCTGGGGGCCACGCCACGTCCGACGACCTGACCAACT
GGGAGGACCAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCT
CCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCAT
CGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGA
GGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCA
GAAGAACCCCGTGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTC
TGGTACGAGCCCTCCCAGAAAGTGGATCATGACCGCGGCCAAGTCCAGGACTAC
AAGATCGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGT
TCGCCAACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGT
CCCCACCGAGCAGGACCCAGCAAGTCTACTGGGTGATGTTTCATCTCCATCAAC
CCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTCGGCAGCTTCAACG
GCACCCACTTCGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGG
ACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCT
GGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCTTCGTGCCACCAACCCC
TGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGG
CCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCA
GCAACGCCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCA
ACAGCTACAACGTCGACCTGTCCAACAGCACCGGCACCCTGGAGTTTCGAGCTGG
TGTACGCCGTCAACACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTC
CCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGA
GGTGTCCGCGTCTCCTTCTTCCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTG
AAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAG
AGCGAGAACGACCTGTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATC
CTGGAGCTGTACTTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGA
CCACCGGGAACGCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGT
TCTACATCGACAAGTTCCAGGTGCGCGAGGTCAAGTGATTAATTAACCTCGAGGCA
GCAGCAGCTCGGATAGTATCGACACACTCTGGACGCTGGTCGTGTGATGGACTGT
TGCCGCCACACTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAG
CCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTTTGCGAGTTGCTAGCTGCTTGTG
CTATTTGCGAATAACACCCCCAGCATCCCCTTCCTCGTTTCATATCGCTTGCATC
CCAACCGCAACTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTG
CTCACTGCCCCCTCGCACAGCCTTGGTTTGGGCTCCGCCTGTATTCTCCTGGTACTG
CAACCTGTAAACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAA
CACAAATGGAAAGCTT

SEQ ID NO: 104

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

TTTGGCCCCGCTTTCCAGCTCCGGATCTGCTGGCGTCCGCCGCGAGACGTGACAT
CGCACGTGCGCCGGGAGCGCCAGCTTGATCACTTGGCAGGGGGCCGTGCTCTACA
AATACCAGGCCCCGCGGCGGTTCAGTTTCGCACATCCAATACCTGCCGAGCCATCTT
GCCTACACTTTTTATCGACTCCTCTACTCTGTTCGCGAGAGCGCTCGGTCCAGGCT
TGGAATTCGCCGAATTCAGCTCGATCAGTCGCTTCTGCAACTGATCTCGGCCGTT
CGCAGACTGCCTTTTCTCAGCTTGTGAGGTAGCGAGTTGTTGTTTTATATTTATTC
GATTTTCATCTGTGTTGCATGTCITGTTTCGTGCTGTGCGTTCTTTCTGGGCCGCGCT
GTCGGGTGCGATGGGCTAGCTGTACTCATGTTAGTCATGCCGGTCCGACCTTGTT
CGAGGAAGGCCCCACACTGAGCGTGCCCTCTTTCTACACCCCTTGTCAGAAATT
AGATAGAAAGCAGAATTCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGG
GCGGGCTGCGAGACGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGAC
CCCCCGAAGCTCCTTCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGC
GCTGTTTAAATAGCCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAA
GCCATATTCAAACACCTAGATCACTACCACITCTACACAGGCCACTCGAGCTTGT
GATCGCACTCCGCTAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCA
AACGGCGCGCCATGCTGCTGCAGGCCTTCCTGTTCTGCTGGCCGGCTTCGCCGC
CAAGATCAGCGCCTCCATGACGAACGAGACGTCCGACCGCCCCCTGGTGCACCTC
ACCCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACGACGAGAAG
GACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGG
ACGCCCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGAC
CAGCCCATCGCCATCGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCA
TGGTGGTGGACTACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCG
CCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTA
CATCTCCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCC
GTGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGAGC
CCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGA
TCTACTCCTCCGACGACCTGAAGTCTTGGAAGCTGGAGTCCGCGTTCGCCAACGA
GGGCTTCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAGGTCCCCACCGAG
CAGGACCCAGCAAGTCTACTGGGTGATGTTTCATCTCCATCAACCCCGGCGCCC
CGGCCGGCGGCTCCTTCAACCAAGTACTTCGTGCGCAGCTTCAACGGCACCCACTT
CGAGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGC
CCTGCAGACCTTCTTCAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCG
TGGGCCTCCAACCTGGGAGTACTCCGCCTTCGTGCCACCAACCCCTGGCGCTCCT
CCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGTACCAGGCCAACCCGGA
GACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACGCCGG
CCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAA
CGTCGACCTGTCCAACAGCACCGGCACCCTGGAGTTCGAGCTGGTGTACGCCGT
AACACCACCCAGACGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCA
AGGGCCTGGAGGACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGT
CCTCCTTCTTCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACC
CCTACTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACG
ACCTGTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTA
CTTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGAAC
GCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACATCGAC
AAGTTCCAGGTGCGCGAGGTCAAGTGATTAATTAACTCGAGGCAGCAGCAGCTC
GGATAGTATCGACACACTCTGGACGCTGGTTCGTGTGATGGACTGTTGCCGCCACA
CTTGCTGCCTTGACCTGTGAATATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTG

TTTGATCTTGTGTGTACGCGCTTTTGGCAGTTGCTAGCTGCTTGTGCTATTTGCGA
 ATACCACCCCCAGCATCCCCTTCCCTCGTTTCATATCGCTTGCATCCCAACCGCAA
 CTTATCTACGCTGTCTGCTATCCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCC
 CTCGCACAGCCTTGGTTTGGGCTCCGCTGTATTCTCCTGGTACTGCAACCTGTAA
 ACCAGCACTGCAATGCTGATGCACGGGAAGTAGTGGGATGGGAACACAAATGGA
 CCGACACGCCCCCGGCCAGGTCCAGTTCTCCTGGGTCTTCCAGAGGCCCGTCGC
 CATGTAAAGTGGCAGAGATTGGCGCCTGATTTCGATTTGGATCCAAGGATCTCCAA
 TCGGTGATGGGGACTGAGTGCCCAACTACCACCCTTGCACTATCGTCCTCGCACT
 ATTTATTCCACCTTCTGCTCGCCCTGCCGGGCGATTGCGGGCGTTTCTGCCCTTG
 ACGTATCAATTTGCCCCCTGCTGGCGCGAGGATTCTTCATTCTAATAAGAACTCA
 CTCCCGCCAGCTCTGTACTTTTCTGCGGGGCCCTGCATGGCTTGTTCCTCAATGC
 TTGCTCGATCGACGGCGCCCATTTGCCACGGCGCTGCCGCATCCATGTGAAGAAA
 CACGGAAGAGTGCGAAGACTGGAAGTGAATTAAGAGTATAAGAAGAGGTACCA
 AGGGATTCTCAGGTGCTCTTAGGAACGGCTTTTCTTCGCCAAGAGAACTGCTA
 CTGCTCGTGTGCGCCACGGTGGTCAAGCCGCCCCATCTGCGATCCACCAGGCCCAT
 CCGCGGACTCGCGATCAGCCTGCTGGATCCGGACTGCCGACCTGACCGCTCGCAT
 CCACCATTACAACCTTCCAATTGGACACCACTCCACGTCCTAAAGTTCACCATG
 CAAGCTGATCGATCGCATTTCGCCGATGCACTCGCCTGCCACAGAGGTGTGCGCTT
 CGGACTAGCGTGCAGGCGCCCCGAGGCCACCAGCATGCACCGATGGAAGCGGGC
 ACGGCCGCTGCTCCAGGTGCTGGCTCGCTCAGACCCATAGCAACCTCCGCTGCG
 TCCCTAAATGTCACACAGAGCGTCTTTGATGGGTACGGATGGGAGAGAATCTGAT
 TGGGCATTGCTGGTGCAGTGCAGGAAGATGGCAAGTGCACAGTCAGTCATGCTG
 TACAACTGGTGCCTCGTAGTATTGACTCGTATAGTGCATAGTATCATGCATGGT
 CGTTACTTGCAA

SEQ ID NO: 105

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

TTTGGCCCCGCTTTCAGCTCCGGATCTGCTGGCGTCCGCCGCGAGACGTGA
 CATCGCACGTCGCCGGGAGCGCCAGCTTGATCACTTGGCAGGGGGGCCGTGCT
 CTACAAATACCAGGCCCGCGGGCGGTTCAGTTCGCACATCCAATACCTGCCGA
 GCCATCTTGCCCTACACTTTTATTCGACTCCTCTACTCTGTTTCGCGAGAGCGCTC
 GGTCCAGGCTTGGAATTCGCCGAATTCAGCTCGATCAGTCGCTTCTGCAACT
 GATCTCGGCCGTTTCGCAGACTGCCTTTTCTCAGCTTGTTCAGGTAGCGAGTTGT
 TGTTTTATATTTATTCGATTTCATCTGTGTGTCATGTCTTGTTCGTGCTGTGCGT
 TCTTTCTGGGCCGCGCTGTCCGGTTCGCATGGGCTAGCTGTACTCATGTTAGTC
 ATGCCGGTCCGACCTTGTTCGAGGAAGGCCCCACACTGAGCGTGCCCTCTTT
 CTACACCCCTTGTGCAGAAATTAGATAGAAAGCAATGCTGCTGCAGGCCTTC
 CTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACG
 AGACGTCCGACCGCCCCCTGGTGCATTTACCCCCAACAAGGGCTGGATGA
 ACGACCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGT
 ACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCCTGGGG
 CCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCAT
 CGCCCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGAC
 TACAACAACACCTCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCT
 GCGTGGCCATCTGGACCTACAACACCCCGGAGTCCGAGGAGCAGTACATCT

CCTACAGCCTGGACGGCGGCTACACCTTCACCGAGTACCAGAAGAACCCCC
TGCTGGCCGCCAACTCCACCCAGTTCCGCGACCCGAAGGTCTTCTGGTACGA
GCCCTCCCAGAAGTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGAT
CGAGATCTACTCCTCCGACGACCTGAAGTCCTGGAAGCTGGAGTCCGCGTTC
GCCAACGAGGGGCTTCCTCGGCTACCAGTACGAGTGCCCCGGCCTGATCGAG
GTCCCCACCGAGCAGGACCCCAGCAAGTCCTACTGGGTGATGTTTCATCTCCA
TCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCAGTACTTCGTCCGGCAG
CTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCCCCGCGTGGTGGAC
TTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCGACCCGACCT
ACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGCCTT
CGTGCCCAACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCC
TCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGG
CCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCA
CCAACACCACGTTGACGAAGGCCAACAGCTACAACGTTCGACCTGTCCAACA
GCACCGGCACCCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGA
CGATCTCCAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAG
GACCCCGAGGAGTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCTCTCTTCT
CCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTT
CACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGACCT
GTCCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTAC
TTCAACGACGGCGACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGA
ACGCCCCTGGGCTCCGTGAACATGACGACGGGGGTGGACAACCTGTTCTACAT
CGACAAGTTCCAGGTGCGCGAGGTCAAGTGACCGACACGCCCCCGGCCAG
GTCCAGTTCTCCTGGGTCTTCCAGAGGCCCCGTCCGCATGTAAAGTGGCAGAG
ATTGGCGCCCTGATTTCGATTTCGGATCCAAGGATCTCCAATCGGTGATGGGGAC
TGAGTGCCCAACTACCACCCTTGCACTATCGTCTCTCGCACTATTTATTCCAC
CTTCTGCTCGCCCTGCCGGGCGATTGCGGGCGTTCCTGCCCTTGACGTATCAA
TTTCGCCCCTGCTGGCGCGAGGATCTTTCATTCTAATAAGAACTCACTCCCGC
CAGCTCTGTACTTTTCTGCGGGGCCCCCTGCATGGCTTGTTCCCAATGCTTGCT
CGATCGACGGCGCCCATTTGCCACGGCGCTGCCGCATCCATGTGAAGAAAC
ACGGAAGAGTGCGAAGACTGGAAGTGAATTAAGAGTATAAGAAGAGGTAC
CAAGGGATTCTCAGGTGCTCTTAGGAACGGCTTTTCTTCGCCAAGAGAAAC
TGCTACTGCTCGTGTGCCACGGTGGTCAAGCCGCCCCATCTGCGATCCACC
AGGCCCATCCGCGGACTCGCGATCAGCCTGCTGGATCCGGACCTGCCGACCCTG
ACCGCTCGCATCCACCATTACAACCCTCCAATTGGACACCCTCCACGTCC
TAAAGTTCACCATGCAAGCTGATCGATCGCATTCGCCGATGCACTCGCCTGC
CACAGAGGTGTGCGCTTCGGACTAGCGTGCAGGCGCCCCGAGGCCACCAGC
ATGCACCGAATGGAAGCGGGCACGGCCGCTGCTCCAGGTGCTGGCTCGCTC
AGACCCATAGCAACCTCCGCTGCGTCCCTAAATGTACACAGAGCGTCTTTG
ATGGGTACGGATGGGAGAGAATCTGATTGGGCATTGCTGGTGCAGTGCAGG
AAGATGGCAAGTGCACAGTCAGTCACTGCTGTACAACTGGTGCCTCGTAGTA
TTGACTCGTATAGTGCATAGTATCATGCATGGTTCGTTACTTGCAA

SEQ ID NO: 106

UTEX 21, 22, 28, 257

TGTTGAAGAATGAGCCGGCGACTTATAGGAAGTGGCGTGGTTAAGGAATTTTCC
GAAGCCCAAGCGAAAGCAAGTTTTTAAAAATAGCGATATTTGTCACTTTTTATGGA
CCCGAACCCGGGTGATCTAACCGTGACCAGGATGAAGCTTGGGTAACACCAAGT
GAAGGTCCGAACCTCTTCGATCTTTAAAAATCGTGAGATGAGTTGCGGTAGTAGG
TGAAATGCCAATCGAACTCGGAGCTAGCTGGTTCTCCCCGAAATGTGTTGAGGCG
CAGCGATGAATGACAAAACAAATAGTACGGTGTAGGGGTAAAGCACTGTTTCGG
TGCGGGCTGCGAAAGCGGTACCAAATCGTGGCAAACCTCAGAATACTACGCTTGT
ATACCATTTCATCAGTGAGACTGTGGGGGATAAGCTCCATAGTCAAGAGGGAAAC
AGCCCAGATCACCAGTTAAGGCCCAAAATGACAGCTAAGTGGCAAAGGAGGTG
AAAGTGAGAAACAACAGGAGGTTTGCCCAGAAGCAGCCATCCTTTAAAGAGT
GCGTAATAGCTCACTG

SEQ ID NO: 107*Cuphea palustris* thioesterase expression cassette

GAATTCCTTTCTTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGA
CGGCTTCCCGGCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCCT
TCGGGGCTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGC
CAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACA
CCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCT
AAGGGGGCGCCTCTTCTCTTCGTTTCAGTCACAACCCGCAAACACTAGTATGGC
CACCGCATCCACTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCGCTCG
GCGGGCTCCGGGGCCCCGGCGCCCAGCGAGGGCCCCCTCCCCGTGCGCGGGCGCGCC
AGCATGCTGCTGTGCGCGGTGACCACGGTCTTCGGCGTGGCCGAGAAGCAGTGG
CCCATGCTGGACCGCAAGTCCAAGCGCCCCGACATGCTGGTCGAGCCCCCTGGGC
GTGGACCGCATCGTCTACGACGGCGTGAGCTTCCGCCAGTCGTTCTCCATCCGCA
GCTACGAGATCGGCGCCGACCGCACCGCCTCGATCGAGACGCTGATGAACATGT
TCCAGGAGACCTCCCTGAACCACTGCAAGATCATCGGCCTGCTGAACGACGGCTT
CGGCCGCACGCCCCGAGATGTGCAAGCGCGACCTGATCTGGGTCTGTGACCAAGAT
GCAGATCGAGGTGAACCGCTACCCACGCTGGGGCGACACCATCGAGGTCAACAC
GTGGGTGAGCGCCTCGGGCAAGCACGGCATGGGCCGCGACTGGCTGATCTCCGA
CTGCCACACCGGCGAGATCCTGATCCGCGCGACGAGCGTCTGGGCGATGATGAA
CCAGAAGACCCGCCGCTGTGCAAGATCCCCTACGAGGTGCGCCAGGAGATCGA
GCCCCAGTTCGTGCTGACTCCGCCCCCGTGATCGTGGACGACCGCAAGTTCCACAAG
CTGGACCTGAAGACGGGCGACAGCATCTGCAACGGCCTGACCCCCCGCTGGACG
GACCTGGACGTGAACAGCACGTCAACAACGTGAAGTACATCGGCTGGATCCTG
CAGTCGGTCCCCACCGAGGTGTTTCGAGACGCAGGAGCTGTGCGGCCTGACCCTG
GAGTACCGCCGCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCGTCACGGCCATG
GACCCCTCGAAGGAGGGCGACCGCTCCCTGTACCAGCACCTGCTGCGCCTGGAG
GACGGCGCGGACATCGTGAAGGGCCGCACCGAGTGGCGCCCCAAGAACGCCGGC
GCCAAGGGCGCCATCCTGACGGGCAAGACCAGCAACGGCAACTCGATCTCCTGA
CTCGAGTTAATTAACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGG
ACGCTGGTCTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAAT
ATCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCT
TTTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTT
CCCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTAT
CCCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCCCTCGCACAGCCTTGGTTTGGGC

TCCGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATG
CACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTT

SEQ ID NO: 108

Codon-optimized *Ulmus americana* thioesterase

GGCGCGCCAGCTGCCCCACTGGAGCATGCTGCTGGCCGCGATCACCACCCTGTT
CCTGGCGGCCGAGAAGCAGTGGATGATGCTGGACTGGAAGCCCAAGCGCCCCGA
CATGCTGGTGGACCCCTTCGGCCTGGGCCGCTTCGTGCAGGACGGCCTGGTGTTC
CGCAACAACCTTCAGCATCCGCAGCTACGAGATCGGCGCGGACCGCACCGCCAGC
ATCGAGACCCTGATGAACCACCTGCAGGAGACCGCCCTGAACCACGTGAAGAGC
GTGGGCCTGCTGGAGGACGGCCTGGGCAGCACCCGCGAGATGAGCCTGCGCAAC
CTGATCTGGGTGGTGACCAAGATGCAGGTGGCGGTGGACCGCTACCCACCTGG
GGCGACGAGGTGCAGGTGAGCAGCTGGGCGACCGCCATCGGCAAGAACGGCAT
GCGCCGCGAGTGGATCGTGACCGACTTCCGCACCGGCGAGACCCTGCTGCGCGC
CACCAGCGTGTGGGTGATGATGAACAAGCTGACCCGCCGCATCAGCAAGATCCC
CGAGGAGGTGTGGCACGAGATCGGCCCCAGCTTCATCGACGCGCCCCCCTGCC
CACCGTGGAGGACGACGGCCGCAAGCTGACCCGCTTCGACGAGAGCAGCGCCGA
CTTCATCCGCAAGGGCCTGACCCCCCGCTGGAGCGACCTGGACATCAACCAGCA
CGTGAACAACGTGAAGTACATCGGCTGGCTGCTGGAGAGCGCGCCCCCGAGAT
CCACGAGAGCCACGAGATCGCCAGCCTGACCCTGGAGTACCGCCGCGAGTGC GG
CCGCGACAGCGTGTGAACAGCGCCACCAAGGTGAGCGACAGCAGCCAGCTGGG
CAAGAGCGCCGTGGAGTGCAACCACCTGGTGCGCCTGCAGAACGGCGGCGAGAT
CGTGAAGGGCCGCACCGTGTGGCGCCCCAAGCGCCCCCTGTACAACGACGGCGC
CGTGGTGGACGTGCCCCGCAAGACCAGCTGACTCGAG

SEQ ID NO: 109

Dual transformation construct: *suc2* sucrose invertase:: *Cuphea hookeriana* C8-10
thioesterase

ggtaccctttcttgcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatg
atgcttcgacccccgaagctccttcggggctgcatggcgctccgatccgctccagggcgagcgctgtttaaataagccaggccc
ccgattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacagggccactcgagctt
gtgatcgactccgctaagggggcgctccttcttcgtttcagtcacaaccgcaaaactctagaatatcaATGCTGCTGCAG
GCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGAACGAGACGTCCG
ACCGCCCCCTGGTGCACCTTCACCCCAACAAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACG
ACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGC
CCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGC
CCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGC
TTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGAGT
CCGAGGAGCAGTACATCTCTACAGCCTGGACGGCGGCTACACCTTACCGAGTACCAGAAGAACC
CCGTGCTGGCCGCCAACTCCACCCAGTTCGCGGACCCGAAGGTCTTCTGGTACGAGCCCTCCAGAA
GTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCTCCGACGACCTGAA
GTCCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGGCTTCTCGGCTACCAGTACGAGTGCCCCGG
CCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCTACTGGGTGATGTTTCATCTCCATCAAC
CCCGGCGCCCCGGCGGCTCCTTCAACCACTTTCGTCGGCAGCTTCAACGGCACCCACTTCG
AGGCCTTCGACAACCACTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTT
CAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACCTGGGAGTACTCCGC
CTTCGTGCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGT

ACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACG
CCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACC
TGTCCAACAGCACCAGGACCCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTC
CAAGTCCGTGTTTCGGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCGC
ATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTG
AAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGGAAACGAC
CTGTCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGCG
ACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGAC
GGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGAAcaattggcagcagc
agctcggatagtagacacactctggacgctggtcggtgtgatggactgttgcgccacacttgccttgacctgtgaatc
cctgcgcgttttatcaaacagcctcagtggttgcctgtgtgtacgcgctttgcgagttgtagctgcttgccttgcgaa
taccacccccagcatcccccttcctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccccagcgt
gctcctgctcctgctcactgcccctgcacagccttgggttgggctccgctgtattctcctggtactgcaacctgtaaacagcact
gcaatgctgatgcacgggaagttagggatgggaacacaatggaggtccgcgctcgaacagagcgcgcagaggaa
cgctgaagggtcgcctctgtgcacctcagcgcggcataccacaataaccacctgacgaatgcgcttgggttcttcgtccatt
agcgaagcgtccggttcacacacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgtt
cacagcctagggatcgaattccgctgcaacgcaagggcagccacagccgctcccacccgctgaaccgacacgtgcttgg
gcgctgccgctgctgccatgcttgcctggtgagggctgggcagtgctgcatgctgattgaggcttgggtcatcggtggaa
gcttatgtgtgctggtgctgcatgccgggcaatgcgcatggtggcaagagggcggcagcacttgcctggagctgccgctgct
ccaggtggttcaatgcggcagcagagggttgcagatgcgcgctacaggtgagcagcagtgctgcaaaaggtagcagtt
tgccagaatgatcggtcagctgttaataatgccagcaagagaaggggtcaagtgcacaacgggcatgccacagcagggga
ccggggagtggaatggcaccaccaagtgtgtgcgagccagcatgcgcgctggctgttcagctacaacggcaggagtcatcaa
cgtaacctagagctgatcaactgcaatcatcgggcgggcgctgatgcaagcatgcctggcgaagacacatggtgtgcggatgct
gccggctgctgctgctgcgcacgcttgagttggcagcaggtcagccatgcactggatggcagctgggctgccaactgcaatgt
ggtggataggatgcaagtggagcgaataccaaacctctggctgcttgcctgggtgcatggcatgcacatcagcaggagcgca
tgcaaggggactggcccatgcacgcatgcaaacgggagcgcacccaggtgtccacactgtcaccaggcccgcaagcttgcag
aacctgctcatggacgcatgtagcgtgacgtcccttgacggcgctcctctcggtgtgggaaacgcaatgcagcacaggcagc
agaggcgggcgagcagagcggcggcagcagcggcgggggccaccttctgcggggtcgcgccccagccagcggatgcgct
gatccaaacaggttcacattcatttgcagcctggagaagcgaggctggggccttgggctggtgcagcccgaatggaatgcg
ggaccgccaggctagcagaaaggcgctccccactccgcatcgatgttccatagtgattggactgcatttgggtggggcgcc
ggctgtttcttctgttgcacaaacgcgcccagctcagcaacctgtccgtgggtccccgctgccgatgaaatcggtgcacgccgat
cagctgattgcccggctcggaagtaggcgcctcttctgctgcctctctcgtcccgccactagtATGGCCACCGCATC
CACTTTCTCGGCGTTCAATGCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGC
CCAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCAGCTGCCCCACTGGAGCCGCTGCTGACCGCC
ATCACCACCGTTCGTGAAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGACA
TGCTGGTGGACAGCTTCGGCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTCGCCAGTCTTCTC
CATCCGCTCCTACGAGATCGGCACCGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCTGCAG
GAGACCTCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTGGACGGCTTCGGCCGACCCCTGGAG
ATGTGCAAGCGCGACCTGATCTGGGTGGTGTATCAAGATGCAGATCAAGGTGAACCGCTACCCCGCC
TGGGGCGACACCGTGGAGATCAACACCCGCTTCAGCCGCTGGGCAAGATCGGCATGGGCGCGA
CTGGCTGATCTCGACTGCAACACCGGCGAGATCCTGGTGCAGCGCCACCAGCGCCTACGCCATGATG
AACCAGAAGACCCGCCGCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTGCCCTGTTC
GTGGACAGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGAAGACCGGCGA
CAGCATCCAGAAGGGCCTGACCCCGGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGT
GAAGTACATCGGCTGGATCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAGGAGCTGTGCTC
CCTGGCCCTGGAGTACCGCCGCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGG
ACCCAGCAAGGTGGGCGTGCGCTCCAGTACCAGCACCTGCTGCGCCTGGAGGACGGCACCGCCA

TCGTGAACGGCGCCACCGAGTGGCGCCCCAAGAACGCCGGCGCCAACGGCGCCATCTCCACCGGCA
 AGACCAGCAACGGCAACTCCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACCACG
 ACATCGACTACAAGGACGACGACGACAAGTGActcgaggcagcagcagctcggtatgtatcgacacactctgga
 cgctggctgtgtatggactgttgccgccacacttgctgccttgacctgtgaatatccctgccgctttatcaaacagcctcagtgt
 gtttgatcttggtgtacgcgcttttgcgagttgctagctgcttggtctatttggaataccacccccagcatcccccttcctcgttc
 atatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgccccctgcaca
 gccttggttgggctccgctgtattctcctggtactgcaacctgtaaacccagcactgcaatgctgatgcacgggaagtagtggg
 atgggaacacaaatggaaagctt

SEQ ID NO: 110

Dual transformation construct: *suc2* sucrose invertase:: *Umbellularia californica* C12
 thioesterase

gglaccctttcttgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatg
 atgcttcgacccccgaagctccttcggggctgcatggcgctccgatccgctccagggcgagcgctgtttaaatagcaggccc
 ccgattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagctt
 gtgatcgactccgctaagggggcgcttctccttctgtttcagtcacaacccgcaaacctctagaatatcaATGCTGCTGCAG
 GCCTTCCTGTTCTGCTGGCCGGCTTCGCCGCAAGATCAGCGCCTCCATGACGAACGAGACGTCCG
 ACCGCCCCCTGGTGCATTCACCCCCAACAGGGCTGGATGAACGACCCCAACGGCCTGTGGTACG
 ACGAGAAGGACGCCAAGTGGCACCTGTACTTCAGTACAACCCGAACGACACCGTCTGGGGGACGC
 CTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGC
 CCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGC
 TTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGT
 CCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCGGCTACACCTTACCGAGTACCAGAAGAACC
 CCGTGCTGGCCGCAACTCCACCCAGTTCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAA
 GTGGATCATGACCGCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAA
 GTCCTGGAAGCTGGAGTCCGCGTTGCCAACGAGGGCTTCCTCGGCTACCAGTACGAGTGCCCCGG
 CCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCTACTGGGTGATGTTTCATCTCCATCAAC
 CCCGGCGCCCCGGCGGCTCCTTCAACCAGTACTTCGTGCGCAGCTTCAACGGCACCCACTTCG
 AGGCCTTCGACAACCAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTT
 CAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGCCTCCAAGTGGGAGTACTCCGC
 CTTCGTGCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTCTCCCTCAACACCGAGT
 ACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACG
 CCGGCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTGACCC
 TGTCCAACAGCACCCGGCACCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTC
 CAAGTCCGTGTTGCGGACCTCTCCCTCTGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCGC
 ATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCTGGACCGCGGGAACAGCAAGGTGAAGTTCGTG
 AAGGAGAACCCCTACTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGAC
 CTGTCTACTACAAGGTGTACGGCTTGCTGGACGAGAACATCCTGGAGCTGTACTTCAACGACGGCG
 ACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGAC
 GGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGAcaattggcagcag
 cagctcggtatgtatcgacacactctggacgctggtcgtgtatggactgttgccgccacacttgctgccttgacctgtgaatat
 ccctgccgctttatcaaacagcctcagtgtgtttgatcttggtgtacgcgcttttgcgagttgctagctgcttggtctatttgga
 ataccacccccagcatcccccttcctcgttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgc
 tgctcctgctcctgctcactgccccctgcacagccttggttgggctccgctgtattctcctggtactgcaacctgtaaacccagcac
 tgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaggatcccgctctcgaacagcgcgagagggaac
 gctgaaggctctgcctctgtgcacctcagcgcgccatacaccacaataaccacctgacgaatgcgcttggttcttcgtccattagcg

aagcgtccgggtcacacacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggcgaaacgttcacagcct
 agggatatcgaattcgcctgcaacgcaagggcagccacagccgctccacccgccgctgaaccgacacgtgcttgggcgcctgc
 cgctgcctgccgcatgcttgtgctggtagggctgggcagtgctgcatgctgattgaggcttgggtcatcggttggagcttatgtg
 tgtgctgggcttgcagtcggggcaatgcgcatggtggcaagagggcggcagcacttgctggagctccgcggtgcctccaggtggt
 tcaatcgcggcagccagagggatttcagatgatcgcgctacaggttgagcagcagtggtcagcaaaggttagcagtttgcagaat
 gatcggttcagctgttaataatgccagcaagagaaggggtcaagtgcaaacacgggcatgccacagcacgggcaccggggag
 tggaatggcaccaccaagtgtgtgcgagccagcatcgccgctggctgtttcagctacaacggcaggagtcacacgtaacca
 tgagctgatcaacactgcaatcatcgggcgggcggtgatgcaagcatgcctggcgaagacacatgggtgtgcggatgctgccggctg
 ctgctgctgcgcacggcttgagttggcagcaggctcagccatgactggatggcagctgggctgcaactgcaatgtgttgata
 ggatgcaagtggagcgaataacaaacccctggtgctgttgcgggttgcagtcgacccatcagcaggagcgcatgcaagg
 gactggccccatgcacgcatgccaaacgggagcgacccagtggtccacactgtcaccaggccccgaagcttgcagaacatgc
 tcatggagcagtagcgctgacgtcccttgacggcgctcctcgggtgtgggaaacgcaatgcagcacaggcagcagaggcg
 cggcagcagagcgggcggcagcagcgggggggccacccttctgcggggtcgcgcccagccagcggtgatgcgctgatccaa
 acgagttcacattcatttgcagcctggagaagcgaggctggggccttgggctggtgcagcccgcaatggaatcggggacggcc
 aggctagcagcaaggcgctccctactccgcatcgatgttccatagtgactggactgcatttgggtggggcgccgctgtttc
 tttcgtgttgcgaacgcgcagctcagcaacctgtccgtgggtccccgtgccgatgaaatcgtgtgcagcgcgatcagctgatt
 gcccggctcggaagtagggcgcctccttctgctgcctcctcctgcccgcactagtATGACGTTGCGGGTCGCCCTC
 CCGGCCATGGGCCGCGGTGTCTCCCTCCCCGCGCCAGGGTCGCGGTGCGCGCCAGTCGGCGAGT
 CAGGTTTTGGAGAGCGGGCGCGCCCCGACTGGTCCATGCTGTTGCGCGTGATCACCACCATCTTCA
 GCGCCGCGGAGAAGCAGTGGACCAACCTGGAGTGGAAAGCCCAAGCCCAAGCTGCCCCAGCTGCTG
 GACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCGGCACCTTCGCCATCCGCTCCTACGAGGTGG
 GCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACCACG
 CCAAGAGCGTGCGCATCCTGGGCGACGGCTTCGGCACCAACCTGGAGATGTCCAAGCGCGACCTGA
 TGTGGGTGGTGCGCCGACCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTGGAG
 GTGGAGTGCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGGTGCGCGACTG
 CAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGATGAACACCCGACCCGCCGC
 CTGAGCACCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCGCCTTCATCGACAACGTGGCCGTG
 AAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGG
 CCTGACCCCCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGCCTG
 GGTGTTGAGACCGTGCCCCGACAGCATCTTCGAGTCCCACCACATCAGCTCCTTACCCTGGAGTAC
 CGCCGCGAGTGACCCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGACGCTCCGAG
 GCCGGCCTGGTGTGCGACCACTGCTGCAGCTGGAGGGCGGCGAGCGAGGTGCTGCGCGCCCGCAC
 CGAGTGCGCCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGCGT
 GATGGACTACAAGGACACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGA
 CAAGTGActcaggcagcagcagctcggtatgtatcgacacactctggacgctggtcgtgtgatggactgttgcggccacac
 ttgctgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtggtttgatcttgtgtgtacgcgttttgcgagttg
 ctgctgcttgtgtctatttgcgaataccacccccagcatccccctccctcggttcatatcgcttgcacccaaccgcaacttatctac
 gctgtcctgtatccctcagcgctgtcctgctcctgctcactgcccctcgacacagccttggttgggctccgctgtatttctcgtgt
 actgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagctt

SEQ ID NO: 111

Dual transformation construct: *suc2* sucrose invertase:: *Cinnamomum camphora* C14
 thioesterase

ggtagcccttcttgcgctatgacacttcagcaaaaggtagggcgggctgcgagacggcttccggcgctgcatgcaacaccgatg
 atgcttcgacccccgaagctccttggggctgcatgggcgctccgatgccgtccaggcgagcgctgtttaaatagccaggccc
 ccgattgcaaagacattatagcgagctacaaagccatattcaaacacctagatcactaccattctacacaggccactcgagctt

gtgatcgactccgctaagggggcgctcttctcttctgtttcagtcacaacccgcaaactctagaatatcaATGCTGCTGCAG
GCCTTCCTGTTCTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCTCCATGACGAACGAGACGTCCG
ACCGCCCCCTGGTGCACCTCACCCCCAACAGGGCTGGATGAACGACCCCCAACGGCCTGTGGTACG
ACGAGAAGGACGCCAAGTGGCACCTGTACTTCCAGTACAACCCGAACGACACCGTCTGGGGGACGC
CCTTGTTCTGGGGCCACGCCACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGC
CCCGAAGCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACCTCCGGC
TTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGACCTACAACACCCCGGAGT
CCGAGGAGCAGTACATCTCTACAGCCTGGACGGCGGCTACACCTTACCCGAGTACCAGAAGAACC
CCGTGCTGGCCGCCAACTCCACCCAGTTCGCGGACCCGAAGGTCTTCTGGTACGAGCCCTCCAGAA
GTGGATCATGACCGCGGCCAAGTCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAA
GTCCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGGCTTCTCGGCTACCAGTACGAGTCCCCGG
CCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCTACTGGGTGATGTTTCATCTCCATCAAC
CCCGGCGCCCCGGCGCGGCTCCTTCAACAGTACTTTCGTCGGCAGCTTCAACGGCACCCACTTCG
AGGCCTTCGACAACCAAGTCCCGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTT
CAACACCGACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCCTCCAACTGGGAGTACTCCGC
CTTCGTGCCCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCAGTTCCTCCTCAACACCGAGT
ACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAGGCCGAGCCGATCCTGAACATCAGCAACG
CCGGCCCCCTGGAGCCGGTTCGCCACCAACACCACGTTGACGAAGGCCAACAGCTACAACGTGACC
TGCCAACAGCACCGGCACCCCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTC
CAAGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGAGTACCTCCGC
ATGGGCTTCGAGGTGTCCGCGTCTCCTTCTTCTGGACCGCGGAACAGCAAGGTGAAGTTCGTG
AAGGAGAACCCTACTTCACCAACCGCATGAGCGTGAACAACAGCCCTTCAAGAGCGAGAACGAC
CTGTCTACTACAAGGTGTACGGCTTGCTGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGCG
ACGTCGTGTCCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGACGAC
GGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAAGTGAcaattggcagcagc
agctcggatagtagtacacacactctggacgctggtcgtgtgatggactgttgccgccacacttgctgccttgacctgtgaatct
cctgcgcttttatcaaacagcctcagtggtgtgtgatctgtgtgacgcgttttcgaggtgctagctgctgtgtatttgcaa
taccacccccagcatccccctccctcgtttcatatcgcttgcatcccaacccgcaacttatctacgctgtcctgtatccctcagcgt
gctcctgctcctgctcactgcccctcgacagccttggtttgggctccgctgtattctcctggtactgcaacctgtaaaccagcact
gcaatgtgtgatgcaggggaagttagtggtgggaacacaaatggaggaatcccgcgctctgaacagagcgcgagaggaa
cgctgaagggtctcgctctgtcgacactcagcgcgcatcaccaacaataaccacctgacgaatgcgcttggttcttcgtccatt
agcgaagcgctccggttcacacacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtgaaacgtt
cacagcctagggtatccgaattccgctgcaacgcaaggcgagccacagcgctccaccccgctgaaaccgacacgtgcttgg
gcgctgcgctgctgctgcatgctgtgctggtgaggtgggcagtgctgcatgctgattgaggttggttcatcggttgaa
gcttatgtgtgtgctgggcttgcagtcggggcaatgcgcatggtggcaagagggcgagcacttgctggagctgcgcggtgcct
ccaggtggttcaatcgcggcagccagagggttcatgatcgcgctacaggttgagcagcagtgctcagaaaggtagcagtt
tgccagaatgatcggttcagctgttaatcaatgccagcaagagaagggtcaagtgcacacgggcatgccacagcacgggca
ccggggagtggaatggcaccaccaagtgtgtgcgagccagcatcgcgctggtgttgcagctacaacggcaggagtcacaa
cgtaaccatgagctgatcaacactgcaatcatcgggcgggcgatgcaagcatgcctggcgaagacacatggtgtgcggatgct
gcccgtgctgctgctgctgcacgcgcttgagttggcagcaggtcagccatgactggatggcagctgggctgacctgcaatgt
ggtggataggatgcaagtggagcgaataccaaacctctggctgctgctgggtgcatggcatcgacacatcagcaggagcgca
tgcaaggaggtggccccatgcacgcatgcaaacggagcgcaaccaggtgtccacactgtcaccaggccccgaagctttgag
aacatgctcatggacgcatgtagcgtgacgtcccttgacggcgctcctctcggtgtgggaaacgaatgcagcacaggcagc
agaggcgggcgagcagagcgggcgagcagcgggggccaccttcttcgggggtcgcgccccagccagcggtgatgcgt
gatcccaaacgagttcacattcatttgatgctggagaagcgaggttggggccttgggctggtgcagccgcaatggaatgcg
ggaccgaggttagcagcaaggcgctccctactccgcatcgatgttccatagtgattggactgcatttgggtggggcgcc
ggctgtttcttctgtgtgcaaacgcgccagctcagcaacctgtccgtgggtccccgtgccgatgaaatcgtgtgcacggcat

cagctgattgcccggtcggaagtagggcgccctcctttctgctcgccctctctccgtcccgccactagtATG GCCACCGCATC
CACTTTCTCGGCGTTCAATGCCCCGCTGCGGCGACCTGCGTGCCTCGGCGGGCTCCGGGCCCCGGCGC
CCAGCGAGGCCCCCTCCCCGTGCGCGGGCGCGCCCCGACTGGTCCATGCTGTTCCGCGTGATCACCA
CCATCTTCTCCGCCGCCGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCA
GCTGCTGGACGACCACTTCGGCCCCCACGGCCTGGTGTTCGCGCGACCTTCGCCATCCGCAGCTAC
GAGGTGGGCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCGCCCTG
AACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCAACCTGGAGATGTCCAAGCGC
GACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACC
GTGGAGGTGGAGTGCTGGGTGGGCGCCTCCGGCAACAACGGCCCGCCACGACTTCCTGGTGGC
CGACTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCAC
CCGCCGCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCGCCTTCATCGACAACGT
GGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGACTCCACCGCCGACTACATCC
AGGGCGGCCTGACCCCCCGCTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACG
TGGACTGGATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTCACCATC
GAGTACCGCCGCGAGTGACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCGGCTCC
TCCGAGGCCGCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGC
CAAGACCGAGTGCGCCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCCGAGTCC
AGCGTGATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGAC
GACGACAAGTGActcgaggcagcagcagctcgatagtagtacacactctggacgctggtcggtgatggactgttgc
gccacacttgctgcttgacctgtgaatatccctgccgcttttatcaaacagcctcagtggtttgatcttggtgtacgcgcttttg
cgagttgtagtgcctgtgtctatttggaataccacccccagcatccccctccctcgtttcatacgcttgcatcccaaccgcaact
tatctacgctgtcctgtatccctcagcgtgctcctgctcctgctcactgccccctgcacagccttggtttgggctccgcctgtattc
tcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagctt

SEQ ID NO: 112

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

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

AGAGAGCGGAGGTGGGGTTGTGAGGTGGGGTTGCTGACCAGGAGCTCGCGTTCGC
CGAGCGCGACTCGCACACGGTCCAGTTACCCCCCCTCCGCCCAAACGCAAGCCT
CCCATCTTGATGCCTTTCGGGCCACCTATACTATTTCTTAGTTTCGCTGTAACATCC
AGACCGTCCTGAATAATAACAATGCCCTGTGTCAAGTGCATTTCCTAAAAAATTC
TGTCCCAACCAACAATCCACCTGAAATACCACCAGCCCTGCCCAGTACACTCTT
CCAATACCATCTCCCTACCTCCACGCGCAAGCGACCCCCATGCGCGACCAGGCTC
GAAAGTGATTTATGACTTGAGACGAGCGAGTGGCGGGCGCGGTGCGACTGCCTTTTC
ATCACGTGCCGTACGTGCGCGACCGCTAGGGCTTTGCACGGCAACGCACGGCTTC
GCCAACCCGACCAGCCAGGACCTCGACTACTCTACCGCGAATTCGCCTCAAGAA
GTCGCCAAATGTGCCATACACCATTCCTTACAGCACTGTTCAAACCTTGATGCCAA
TTTTGACATTTCGGGTGCTCGTTGGCTGCGCCCCACATCGGCCGTGAGTGCAGCAG
GCGGGATCGGACACGGAGGACGCGGCGTACGCCCCGAACGCAGCCCGTAACCTC
TACATCAACACGACGTGTTGCGTAATCCCGCCCGGCTGCGCATCGTGCCAACCCA
TTCGCGATGGATGGTCGGAAAATGGTGTGCCAACTGCCCTGAGGGAGGCTCTCG
CGAAACGGGCACGTCCCTGAAACCGAAACTGTGGCCTTGTCGTGCGCCACGCAA
GCACGTGGACCCTAAACACCAAGAAAATCAGTAAACAAGGTTGACATCCTCTAC
GGGCGAATTGTTTGCCCAACCTTCATCGCACACTGCCATTATAATGCATCTAGC
TCGGCGACAAGTTTAGAAAAGGCAGGCTGCATTGTTCCATTTCCCGCTGGCGGCG
TGGGTGCCCATTTTACGAGGTTTGGGCTCCCGGGCAGCGACCGAGCCAGGTCGA
GTCCCTCTCGCCCGTTCGACAATGTTGCGAACCCCAACAAGCGGCTAACAACAACCTT
GATGGTACCTGTACACTGCCAATTCCTTCTTCCCCGGCCGAGGTTTACACGTGAT

GGCCATGGCTTCGCATTCAGGCCGACTTCCCATTTCCGACTTTCCAGAGGGTCCGC
 GGACGCTGGGGGTTGGCTGCCTGAGGCCACCCCTTTGTTCCCCGCGTCCCGACAA
 ACACAATTGCGTTACATAAGGGGGAGCCGCCCCCGTTTCAGAGTGCAGAAATCTTT
 CACTATATTTTCCAGTCGTCAGCGAAATCAAGTACTAGTATGGCCACCGCATCCA
 CTTTCTCGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGG
 GCCCCGGCGCCCAGCGAGGCCCTCCCCGTGCGCGGGCGCGCCCCCGACTGGTC
 CATGCTGTTTCGCCGTGATCACCACCATCTTTCTCCGCCGCCGAGAAGCAGTGGACC
 AACCTGGAGTGGAAGCCCAAGCCCAACCCCCCCCAGCTGCTGGACGACCACTTC
 GGCCCCACGGCCTGGTGTTCGCGCGCACCTTCGCCATCCGCAGCTACGAGGTGG
 GCCCCGACCGCTCCACCAGCATCGTGGCCGTGATGAACCACCTGCAGGAGGCCG
 CCCTGAACCACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCAACCT
 GGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAGCGCACCCACGTGGCCGT
 GGAGCGCTACCCCGCCTGGGGCGACACCGTGGAGGTGGAGTGTGGGTGGGCGC
 CTCCGGCAACAACGGCCGCCGCCACGACTTCCTGGTGCGCGACTGCAAGACCGG
 CGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCCGC
 CGCCTGAGCAAGATCCCCGAGGAGGTGCGCGGGCGAGATCGGCCCGCCTTCATC
 GACAACGTGGCCGTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGA
 CTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCCGCTGGAACGACCTGGA
 CATCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACCGT
 GCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTCACCATCGAGTACCGC
 CGCGAGTGCACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGGCGGC
 TCCTCCGAGGCCGGCCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGC
 GAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCCGC
 GGCAATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGACCACGAC
 GCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGACTC
 GAGTTAATTAACTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGAC
 GCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATAT
 CCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTTT
 TCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTCC
 CTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCTGCTATCC
 CTCAGCGCTGCTCCTGCTCCTGCTCACCTGCCCTCGCACAGCCTTGTTTGGGCTC
 CGCCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGCA
 CGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTT

SEQ ID NO: 113

C.reinhardtii β -tubulin promoter::native PTS::*Cuphea hookeriana* C8-10
 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

ggtaccgcctgcaacgcaagggcagccacagccgctccacccgcccgtgaaccgacacgtgcttgggcgcctgccgcctgctg
 ccgcatgcttgtgctggtgaggctgggcagtgctgcatgctgattgaggcttggttcatcggtggaagcttatgtgtgtgctggg
 cttgcatgccgggcaatgcatggtggcaagagggcggcagcacttgctggagctgccggtgcctccaggtggttcaatcgc
 ggcagccagagggatttcagatgatcgcgctacaggttgagcagcagtgctcagcaaaggttagcagtttgccagaatgatcggt
 cagctgttaataatgccagcaagagaaggggtcaagtgcaaacacgggcatgccacagcacgggcaacggggagtggaatgg
 caccaccaagtgtgtgcgagccagcatcgccgctggtgtttcagctacaacggcaggagtcaccaacgtaaccatgagctga
 tcaacactgcaatcatcgggcggtgatgcaagcatgcctggcgaagacacatggtgtgcggatgctgcgggctgctgctgct
 gcgcacgcccgttgagttggcagcaggctcagccatgcactggatggcagctgggctgactgcaatgtggtggataggatgcaa
 gtggagcgaataccaaaccctctggtgcttgcctgggttgatggcatcgaccatcagcaggagcgcatgcgaagggactggcc
 ccatgcacgcatgccaacccggagcgaccagtggtccacactgtcaccagggcccgaagcttgagaaccatgctcatggac
 gcatgtagcgctgacgtcccttgacggcgctcctctcggtgtgggaaacgcaatgcagcacaggcagcagaggcgcgccgagc

agagcggcggcagcagcggcggggccacccttcttgcggggtcgcgccccagccagcggtgatgcgctgatccaaaacgagttc
acattcatttgcagcctggagaagcgaggtcggggcctttgggctggcgagcccgcaatggaatgcgggaccgcccaggctagc
agcaaaggcgctcccctactccgcatcgatgttccatagtgcattggactgcatttgggtggggcgccggctgtttcttctgtgtt
gcaaaacgcgccagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcacgccgatcagctgattgccggct
cggaagtagggcgccctcttctgtcgcctctctccgtcccgcctctagaatatcaATGATCGAGCAGGACGGCCTCC
ACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGCCCAGCAGACCATCG
GCTGCTCCGACGCCCGCTGTTCCGCCTGTCCGCCAGGGCCGCCCGTGTGTTCTGTAAGACCGA
CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCTGTCTGGCTGGCCACCACCGG
CGTGCCCTGCGCCGCCGTGCTGGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCG
AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
ACGCCATGCGCCGCTGCACACCTGGACCCCGCCACCTGCCCTTCGACCACCAGGCCAAGCACCG
CATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGACCGAGGACGACCTGGACGAGGAGC
ACCAGGGCCTGGCCCCCGCCGAGCTGTTGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCTGCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCT
TCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGTGTACGGCATCGCCGCCCGCG
ACTCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAcaattggcagcagcagctcgatagt
atcgacacactctggacgctggtcggtgatggactgttgcgccacacttgccttgacctgtgaatatccctgccgctttat
caaacagcctcagtggttgccttgcgtgtacgcgctttgcgagtgctagctgctgtgtctatttgcgaataaccacccccagc
atccccctccctcgcttcataatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgct
cactgccccctgcacagccttgggttgggctccgctgtatttctcctggtactgcaacctgtaaacagcactgcaatgctgatgc
acgggaagtagtgggatgggaacacaaatggaggatcccgctctcgaacagagcgcgagaggaacgctgaaggtctcgcc
tctgtcgacactcagcgcgccatacaccacaataaacacctgacgaatcgcttgggttctcgtccattagcgaagcgtccgggtcac
acacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatcgaaattc
ctttcttgcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcagcaaacaccgatgatgctt
cgacccccgaagctccttgcgggctgcagggcgctccgatgccgtccagggcgagcgctgtttaaataagccaggcccccgatt
gcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatc
gcactccgctaagggggcgcttcttcttcttcttctcagtcacaacccgcaaacggcgccgcccATGGTGGCCGCCGCCGCT
CCAGCGCCTTCTTCCCCGTGCCCGCCCCCGCGCCTCCCCCAAGCCCGGCAAGTTCCGCAACTGGCC
CTCCAGCCTGAGCCCCCTCCTTCAAGCCCAAGTCCATCCCCAACGGCGGCTTCCAGGTGAAGGCCAAC
GACAGCGCCACCCCAAGGCCAACGGCTCCGCCGTGAGCCTGAAGAGCGGCAGCCTGAACACCCAG
GAGGACACCTCCTCCAGCCCCCCCCCGCACCTTCTGACCAAGCTGCCCCACTGGAGCCGCTGCT
GACCGCCATCACCACTGTGTTCTGTAAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCG
CCCCGACATGCTGGTGGACAGCTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTCGCCA
GTCCTTCTCCATCCGCTCCTACGAGATCGGCACCGACCGCACCGCCAGCATCGAGACCCTGATGAAC
CACCTGCAGGAGACCTCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTGGACGGCTTCGGCCGC
ACCCTGGAGATGTGCAAGCGGACCTGATCTGGGTGGTGTATCAAGATGCAGATCAAGGTGAACCGC
TACCCCGCCTGGGGCGACACCGTGGAGATCAACACCCGCTTACGCCCTGGGCAAGATCGGCATG
GGCCGCGACTGGCTGATCTCCGACTGCAACACCGCGGAGATCCTGGTGCGCGCCACCAGCGCCTAC
GCCATGATGAACCAGAAGACCCGCCGCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTG
CCCCGTTCGTGGACAGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGAAG
ACCGGCGACAGCATCCAGAAGGGCCTGACCCCGGCTGGAACGACCTGGACGTGAACCAGCACGT
GTCCAACGTGAAGTACATCGGCTGGATCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAGGA
GCTGTGCTCCCTGGCCCTGGAGTACCGCCGCGAGTGGCGCCGCGACTCCGTGCTGGAGAGCGTGAC
CGCCATGGACCCAGCAAGGTGGGCGTGCCTCCAGTACCAGCACCTGCTGCGCCTGGAGGACGG
CACCGCCATCGTGAACGGCGCCACCGAGTGGCGCCCCAAGAAGCCGGCGCCAACGGCGCCATCTC
CACCGGCAAGACCAGCAACGGCAACTCCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAA

GGACCACGACATCGACTACAAGGACGACGACGACAAGTGAActcgaggcagcagcagctcggatagtatcgac
acactctggacgctggtcgtgtgatggactgttgcgccacacttgctgccttgacctgtgaatatccctgccgcttttatcaaac
agcctcagtggtgttgcctgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcatcccc
ttccctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactg
ccctcgcacagccttggttgggctccgctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacggg
aagtagtgggatgggaacacaaatggaaagctt

SEQ ID NO: 114

C.reinhardtii β -tubulin promoter::*P.moriformis* isopentenyl diphosphate synthase

PTS::*Cuphea hookeriana* C8-10 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

ggtaccgcctgcaacgcaagggcagccacagccgctcccacccgccgtgaaccgacacgtgcttgggcgctgccgctgcctg
ccgcatgcttgtgctggtgaggctgggcagtgctgcatgctgattgaggcttgggtcatcggggtggaagcttatgtgtgtgctggg
cttgcacgcccgaatgcgcatggtggcaagagggcgccagcacttgcctggagctgccggtgctccagggtggtcaatgcg
ggcagccagagggattcagatgatcgcgctacaggttgagcagcagtgctcagcaaaggttagcagtttgcagaatgatcggtt
cagctgttaatcaatgccagcaagagaaggggtcaagtgcacaacacgggcagccacagcaggggcacggggagtggaatgg
caccaccaagtgtgtgcgagccagcatcgccgctggctgtttcagctacaacggcaggagtcaccaacgtaaccatgagctga
tcaactgcaatcatcgggcgggcgctgatgcaagcatgctggcgaagacacatgggtgtcgggatgtcgccgctgctgctgctgct
gcgacgcccgttgagttggcagcaggctcagccatgactggtggcagctgggctgccactgcaatgtggtggataggatgcaa
gtggagcgaataccaaaccctctggctgcttgcctgggttcagtgcatcgccatcagcaggagcgcatgcaagggactggcc
ccatgcacgcatgccaacccgagcgacccagtgctccactgtcaccaggcccgcaagcttgcagaaccatgctcatggac
gcatgtagcgctgacgtccctgacggcgctcctcctgggtgtgggaaacgcaatgcagcacaggcagcagagggcgggcagc
agagcgggcgccagcagcgccgggggcccacccttcttgcggggctcgcgcccagccagcggtgatgcgctgacccaaacgagttc
acattcatttgcacgctggagaagcgaggctggggcctttgggctggtgcagcccgcaatggaatgcccgaaccgagggtagc
agcaaaggcgctcccctactcgcacatgattccatagtgacttgagctgatttgggtggggcgccggctgttcttctgtgtt
gcaaaacgcgcagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcacgccgatcagctgattgccggct
cgcgaaagtaggcgcccctcttctgctcgccctctcctcgctccccttagaataatcaATGATCGAGCAGGACGGCCTCC
ACGCCGGCTCCCCGCCGCTGGGTGGAGCGCTGTCGGCTACGACTGGGCCAGCAGACCATCG
GCTGCTCCGACGCCGCGTGTCCGCTGTCCGCCAGGGCCGCCCGTGTGCTGTTCTGTAAGACCGA
CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCCTGTCTGGCTGGCCACCACCGG
CGTGCCCTGCGCCGCGTGTGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCG
AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
ACGCCATGCGCCGCTGCACACCTGGACCCCGCCACCTGCCCCTTCGACCACCAGGCCAAGCACCG
CATCGAGCGCGCCCGCACCCGCATGGAGGCCGCGCTGGTGACCGAGGACGACCTGGACGAGGAGC
ACCAGGGCCTGGCCCCCGCCGAGCTGTTGCCCGCCTGAAGGCCCGCATGCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCTGCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCT
TCATCGACTGCGGCCGCTGGGCGTGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
TCGCCGAGGAGCTGGGCGGCGAGTGGGCGGACCGCTTCTGGTGTGTACGGCATCGCCGCCCCCG
ACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAcaattgcagcagcagctcggatagt
atcgacacactctggacgctggtcgtgtgatggactgttgcgccacacttgctgccttgacctgtgaatatccctgccgcttttat
caaacagcctcagtggttgcctgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagc
atcccttccctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgct
cactgccccctgcacagccttggttgggctccgctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgc
acgggaagtagtgggatgggaacacaaatggaggatcccgcgtctgaacagagcgcgagaggaacgctgaaggtctcgcc
tctgtcgacctcagcgcgccatacaccacaataaccacctgacgaatgcgttgggttcttctgctcattagcgaagcgtccggttcac
acacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggcgaaacgttcacagcctagggatatcgaattc
cttcttgcgctatgacacttcagcaaaaggtagggcgggctgcgagacggctcccggcgctgcatgcaacaccgatgatgctt

ACGCCGGCTCCCCGCCGCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGCCCAGCAGACCATCG
 GCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGCCGCCCGTGTCTGTTCTGTAAGACCGA
 CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCTGTCTGGCTGGCCACCACCGG
 CGTGCCCTGCGCCGCCGTGCTGGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCG
 AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
 ACGCCATGCGCCGCTGCACACCCTGGACCCCGCCACCTGCCCCCTCGACCACCAGGCCAAGCACCG
 CATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGGACCAGGACGACCTGGACGAGGAGC
 ACCAGGGCCTGGCCCCCGCCGAGCTGTTCCGCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACC
 TGGTGGTGACCCACGGCGACGCCTGCCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCT
 TCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
 TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGTGTACGGCATCGCCGCCCCCG
 ACTCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAAcaattggcagcagcagctcggtatagt
atcgacacactctggacgctggtcggtgatgactgttgcgccacacttgctgccttgacctgtgaatatccctgccgttttat
caaocagcctcagtgtgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgtatttgcgaataccacccccagc
atccccctccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgt
cactgccccctgcacagccttggtttgggctccgctgtatttctctggtactgcaacctgtaaacagcactgcaatgctgatgc
acgggaagtagtggttggaacacaaatggaggatcccgctctcgaacagagcgcgagaggaacgctgaaggtctcgcc
tctgtcgacactcagcgggcatacaccacaataaccacctgacgaatgcgcttggttcttctgctcattagcgaagcgtccggttcac
acacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatcgaattc
ctttcttgcgctatgacacttcagcaaaaggtaggggcggtgcgagacggcttccggcgctgcatgcaacaccgatgatgctt
cgacccccgaagctccttcggggctgcatggggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggcccccgatt
gcaaagacattatagcgagctaccaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatc
gcactccgctaagggggcgcttcttcttctgtttcagtcacaaccgcaaacactagtATGGCTATCAAGACGAACAG
 GCAGCCTGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTT
 CGAGCGCTCGGCGCTTCGTGGGCGCGCCAGCTGCCCCACTGGAGCCGCCTGCTGACCGCCATCAC
 CACCGTGTTCGTGAAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGACATGCT
 GGTGGACAGCTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTTCGCCAGTCCTTCTCCATC
 CGCTCTACGAGATCGGCACCGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAG
 ACCTCCCTGAACCACTGCAAGAGCACCGGCATCCTGCTGGACGGCTTCGGCCGACCCCTGGAGATGT
 GCAAGCGCGACCTGATCTGGGTGGTGATCAAGATGCAGATCAAGGTGAACCGCTACCCCGCCTGGG
 GCGACACCGTGGAGATCAACACCCGCTTACGCCGCTGGGCAAGATCGGCATGGGCGCGACTGGC
 TGATCTCCGACTGCAACACCGGCGAGATCCTGGTGCAGGCCACCAGCGCTACGCCATGATGAACCA
 GAAGACCCGCGCCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTGCCCCTGTTCTGTGGA
 CAGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGAAGACCGGCGACAGCAT
 CCAGAAGGGCCTGACCCCGGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGTGAAGTA
 CATCGGCTGGATCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCCCTGGC
 CCTGGAGTACCGCCGCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGGACCCCA
 GCAAGGTGGGCGTGCCTCCAGTACCAGCACCTGCTGCGCCTGGAGGACGGCACCGCCATCGTGA
 ACGGCGCCACCGAGTGGCGCCCCAAGAACGCCGGCGCCAACGGCGCCATCTCCACCGGCAAGACCA
 GCAACGGCAACTCCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCG
 ACTACAAGGACGACGACGACAAGTGAActcgaggcagcagcagctcggtatgtatcgacacactctggacgctggtc
gtgtgatggaactgttgcgccacacttgctgccttgacctgtgaatatccctgccgttttatcaaocagcctcagtgtgtttgatc
ttgtgtgtacgcgcttttgcgagttgctagctgcttgtgtatttgcgaataccacccccagcatccccctccctcgtttcatatcgct
tgcaccccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgccccctgcacagccttggt
ttgggctccgctgtatttctctggtactgcaacctgtaaacagcactgcaatgctgatgcacgggaagtagtggttgga
acacaaatggaagctt

SEQ ID NO: 116

C.reinhardtii β -tubulin promoter::*C.protothecoides* stearyl ACP desaturase PTS::*Cuphea hookeriana* C8-10 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

ggtagccgcctgcaacgcaagggcagccacagccgctccacccgccgctgaaccgacacgtgcttggcgccctgccgctgctg
ccgcatgcttgctggtgaggctgggagctgctgcatgctgattgaggcttggttcacgggtggaagcttatgtgtgctggg
cttgcagtcgggcaatgcgcatggtggcaagaggcgagcacttgcaggagctgccggtgctccagggtggtcaatcgc
ggcagccagagggttgcagatgacgctgacaggtgagcagcagtgtagcaaaaggtagcagtttgccagaatgatcggtt
cagctgttaataatgccagcaagagaaggggtcaagtgcacacgggcatgccacagcagggcaccggggagtggaatgg
caccaccaagtgtgtgcgagccagcatgcgcgctggtgtttcagctacaacggcaggagtcacacgtaaccatgagctga
tcaacactgcaatcatcgggcggtgtagcaagcatgctggcgaagacacatggtgtgcggatgctgccggtgctgctgctg
gcgcacgcttgagttggcagcaggctcagcatgactggatggcagctgggctgcaatgtggtggataggatgcaa
gtggagcgaatacacaacctctggtgcttgcgtgggtgcatggcatgcacacatcagcaggagcgcatgcgaagggactggcc
ccatgcacgcatgccaaaccggagcgacaggtgtccacactgtcaccaggcccgcaagctttgcagaacctgctcatggac
gcatgtagcgtgacgtccctgacggcgctcctcctgggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcggcagc
agagcgggcgagcagcgggggggccaccttcttgcgggtgcgcggccagccagcggtgatgcgctgatccaaacgagttc
acattcatttgcagctgctggagaagcgaggctggggcctttgggtggtgcagccgcaatggaatgcgggaccgaccaggctagc
agcaaaggcgctcccactcgcacatcagtggtccatagtgactggactgcatttgggtggggcggtgctttcttctggtt
gcaaaacgcgcagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcacggcatcagctgattgccggct
cggaagtaggcgcccctcttctgctcgccctcctcctgccccgctctagaatatcaATGATCGAGCAGGACGGCCTCC
ACGCCGGCTCCCCGCCGCTGGGTGGAGCGCTGTTCCGGCTACGACTGGGCCAGCAGACCATCG
GCTGCTCCGACGCCGCGTGTTCGCCTGTCCGCCAGGGCCGCCCGTGTGTTCTGTAAGACCGA
CCTGTCCGGCGCCTGAACGAGCTGCAGGACGAGGCCGCCGCTGTCTGGCTGGCCACCACCGG
CGTGCCCTGCGCCGCGTGTGGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCG
AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
ACGCCATGCGCCGCTGCACACCTGGACCCGCCACCTGCCCTTCGACCACCAGGCCAAGCACCG
CATCGAGCGCGCCCGCACCCGCATGGAGGCCGCTGGTGACCGAGGACGACCTGGACGAGGAGC
ACCAGGGCTGGCCCCGCCGAGCTGTTCCGCCGCTGAAGGCCCGCATGCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCTGCCTGCCAACATCATGGTGAGAACGGCCGCTTCTCCGGCT
TCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGTGTACGGCATCGCCGCCCCCG
ACTCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAcaattggcagcagcagctcgatag
atcgacacactctggacgctggtgctgtgatgactgttgcggccacacttgccttgacctgtgaatatccctgcccgtttat
caaacagcctcagtggttgcctgtgtgtacgcgttttgcgagttgtagctgctgtgtctatttgcgaataaccaccccgagc
atccccctccctcgttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgct
cactccccctgcacagccttggttgggtccgctgtatttctcctggtactgcaacctgtaaccagcactgcaatgctgatgc
acgggaagtagtgggatgggaacacaatggaggatcccgctctcgaacagagcgcgagaggaacgctgaaggctcgc
tctgtgcacctcagcgcgccatacaccacaataaacacctgacgaatgcgcttgggtctcgtccattagcgaagcgtccggttcac
acacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggctgaaacgttcacagcctagggatatcgaattc
ctttcttgcgctatgacacttcagcaaaaggtaggcggtgtagagacggcttccggcgctgcatgcaacaccgatgatgctt
cgacccccgaagctccttcggggctgcatggcgctccgatgccgctccaggcgagcgctgtttaaatagccaggccccgatt
gcaaagacattatagcgagctacaaagccatattcaaacacctagatcactaccattctacacaggccactcgagcttgcgatc
gcactccgctaaggggcgctcttctccttgccttcagtcacaacccgcaaacactagtATGGCCACCGCATCCACTTCT
CGGCGTTCAATGCCCGCTGCGGCGACCTGCGTCGCTCGGCGGGCTCCGGGCCCCGGCGCCAGCGA
GGCCCCCTCCCGTGCAGCGGGCGCGCCAGCTGCCGACTGGAGCCGCTGCTGACCGCCATACCA
CCGTGTTCTGTAAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGACATGCTGG
TGGACAGCTTCGGCCTGGAGTCCACCGTGCAGGACGGCCTGGTGTCCGCCAGTCTTCTCCATCCG

CTCCTACGAGATCGGCACCGACCGCACCAGCATCGAGACCCTGATGAACCACCTGCAGGAGAC
CTCCCTGAACCACTGCAAGAGCACCAGCATCCTGCTGGACGGCTTCGGCCGACCCCTGGAGATGTGC
AAGCGCGACCTGATCTGGGTGGTATCAAGATGCAGATCAAGGTGAACCGCTACCCCGCTGGGGC
GACACCGTGGAGATCAACACCCGCTTCAGCCGCTGGGCAAGATCGGCATGGGCCGCGACTGGCTG
ATCTCCGACTGCAACACCGGCGAGATCCTGGTGCAGCGCCACCAGCGCCTACGCCATGATGAACCAG
AAGACCCGCGCCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTGCCCTGTTCTGTGGAC
AGCCCCGTGATCGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGAAGACCGGCGACAGCATC
CAGAAGGGCCTGACCCCGGCTGGAACGACCTGGACGTGAACCAGCACGTGTCCAACGTGAAGTAC
ATCGGCTGGATCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCCCTGGCC
CTGGAGTACCGCCGCGAGTGCAGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGGACCCAG
CAAGGTGGGCGTGCCTCCAGTACCAGCACCTGCTGCGCCTGGAGGACGGCACCGCCATCGTGAA
CGGCGCCACCGAGTGGCGCCCCAAGAACGCCGCGCCAACGGCGCCATCTCCACCGGCAAGACCA
GCAACGGCAACTCCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCG
ACTACAAGGACGACGACGACAAGTGAActcgaggcagcagcagctcgatagtagcacacactctggacgctggtc
gtgtgtatggactgttgcgcccacacttgccttgacctgtgaatatccctgcgcttttatcaacagcctcagtgtgtttgatc
ttgtgtgtacgcgcttttgcgagttgtagctgtgtgtctatttgcgaataccacccccagcatcccttcctcgtttcatatcgct
tgcatcccaaccgcaacttatctacgtgtctgtctatccctcagcgctgtctgtctcgtcactgcccctgcacagccttgg
ttgggctccgctgtatttctcgtgtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatggga
acacaaatggaaagctt

SEQ ID NO: 117

C.reinhardtii β -tubulin promoter::*P.moriformis* stearyl ACP desaturase PTS::*Cuphea
hookeriana* C8-10 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

ggtaccgcctgcaacgcaagggcagccacagccgctccacccgccgctgaaccgacacgtgcttgggcgctgccgctgctg
ccgcatgcttgtgtggtgaggctgggcagtgctgcatgctgattgaggcttggttcatcggtggaagcttatgtgtgtgctggg
cttgcagtcggggaatgcgcatggtggcaagagggcgagcacttgcctggagctgcccgggtgcctccaggtggttcaatcgc
ggcagccagagggatttcagatgatcgcgctacaggttgagcagcagtgctcagcaaaggttagcagtttgccagaatgatcggtt
cagctgttaataatgccagcaagagaaggggtcaagtgcacaacagggcatgccacagcacggggcaccggggagtggaatgg
caccaccaagtgtgtgcgagccagcatcgccgctggctgtttcagctacaacggcaggagtcacacgtaaccatgagctga
tcaacactgcaatcatcgggcgggcgctgatgcaagcatgcctggcggaagacacatggtgtgcggatgctgcggctgctgctgct
gcgcacgcccgttgagttggcagcaggctcagccatgactggatggcagctgggctgccactgcaatgtggtggataggatgcaa
gtggagcgaatacacaacccctctggctgcttgcctgggttgcagtcacccatcagcaggagcgcatgcgaagggactggcc
ccatgcacgcatgccaaaccggagcgacccagtgctcacactgtcaccagggcccgaagcttgcagaacctgctcatggac
gcatgtagcgtgacgtcccttgacggcgctcctcctgggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcggcagc
agagcgggcgagcagcgggcgggggccacccttctgcggggtgcgccccagccagcggtgatgcctgatcccaaacgagttc
acattcatttgcagcctggagaagcgaggctggggcctttgggctggtgcagcccgaatggaatgcgggaccgccaggctagc
agcaaaggcgctccctactccgcatcgatgttccatagtgacttggtgagctgatttgggtggggcgggcggtgtttcttctgtgtt
gcaaaacgcgcagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcacgccgatcagctgattgccggct
cgcgaaagtagggcgccctcttctgctcgccctctcctgctcccgctcagaatatcaATGATCGAGCAGGACGGCCTCC
ACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGCCAGCAGACCATCG
GCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGCCGCCCGCTGCTGTTCTGTGAAGACCGA
CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCCGCCCGCTGTCTGGCTGGCCACCACCGG
CGTGCCCTGCGCCGCCGTGCTGGACGTGGTGACCGAGGCCGGCCGCGACTGGCTGCTGCTGGGCG
AGGTGCCCGGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
ACGCCATGCGCCGCTGCACACCTGGACCCCGCCACCTGCCCTTCGACCACCAGGCCAAGCACCG
CATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGGACCAGGACGACCTGGACGAGGAGC

ACCAGGGCCTGGCCCCGCGAGCTGTTGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCCTGCCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCT
TCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGTGTACGGCATCGCCGCCCCCG
ACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAcaattggcagcagcagctcgatagt
atcgacacactctggacgctggtcggtgtgatggactgttgcgcccacacttgctgccttgacctgtgaatatccctgccgtttat
caaacagcctcagtgtgtttgatcttgtgtgtacgcgttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagc
atccccctccctcgtttcataatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctctgctcctgt
cactgccccctgcacagccttggttgggtccgctgtatttctctggtactgcaacctgtaaacagcactgcaatgctgatgc
acgggaagtagtgggatgggaacacaaatggaggatcccgctctcgaacagagcgcgagaggaaacgtgaaggtctgcc
tctgtcgacactcagcgcgccatacaccacaataaccacctgacgaatgcgcttggttcttctgctccattagcgaagcgtccggttcac
acacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatcgaattc
ctttcttgcgtatgacacttcagcaaaaggtagggcgggctgcgagacggcttccggcgctgcatgcaacacccgatgatgctt
cgacccccgaagctccttgcgggctgcatggcgctccgatgccgctccaggcgagcgctgtttaaatagccaggccccgatt
gcaaagacattatagcgagctacaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgac
gcactccgctaagggggcgcttcttcttcttcttctcagtcacaaccgcaaacactagtATGGCTTCCGCGGCATTACCC
ATGTCGGCGTGCCCCGCGATGACTGGCAGGGCCCTGGGGCACGTCGCTCCGGACGGCCAGTCGCC
ACCCGCTGAGGGGGCGCGCCCCGACTGGAGCCGCTGCTGACCGCCATCACACCGTGTTCTGTG
AAGTCCAAGCGCCCCGACATGCACGACCGCAAGTCCAAGCGCCCCGACATGCTGGTGGACAGCTTC
GGCCTGGAGTCCACCGTGACGACGGCCTGGTGTTCGCCAGTCCTTCTCCATCCGCTCCTACGAGA
TCGGCACCGACCGCACCGCCAGCATCGAGACCCTGATGAACCACCTGCAGGAGACCTCCCTGAACC
ACTGCAAGAGCACCGGCATCCTGCTGGACGGCTTCGGCCGACCCCTGGAGATGTGCAAGCGCGACC
TGATCTGGGTGGTGATCAAGATGCAGATCAAGGTGAACCGCTACCCCGCTGGGGCGACACCGTGG
AGATCAACACCCGCTTCAGCCGCTGGGCAAGATCGGCATGGGCCGCGACTGGCTGATCTCCGACT
GCAACACCGGCGAGATCCTGGTGC GCGCCACCAGCGCCTACGCCATGATGAACCAGAAGACCCGCC
GCCTGTCCAAGCTGCCCTACGAGGTGCACCAGGAGATCGTGCCCCCTGTTCTGTGGACAGCCCCGTGAT
CGAGGACTCCGACCTGAAGGTGCACAAGTTCAAGGTGAAGACCGGCGACAGCATCCAGAAGGGCC
TGACCCCCGCTGGAACGACCTGGACGTGAACCAGCACGTGTCAACGTGAAGTACATCGGCTGGA
TCCTGGAGAGCATGCCACCGAGGTGCTGGAGACCCAGGAGCTGTGCTCCCTGGCCCTGGAGTACC
GCCGCGAGTGCGGCCGCGACTCCGTGCTGGAGAGCGTGACCGCCATGGACCCAGCAAGGTGGGC
GTGCGTCCAGTACCAGCACCTGCTGCGCCTGGAGGACGGCACCGCCATCGTGAACGGCGCCACC
GAGTGGCGCCCCAAGAACGCCGGCGCCAACGGCGCCATCTCCACCGGCAAGACCAGCAACGGCAA
CTCCGTGTCCATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGA
CGACGACGACAAGTGActcgaggcagcagcagctcgatagtatcgacacactctggacgctggtcggtgtgatggactg
ttgcccccacacttgctgccttgacctgtgaatatccctgccgtttatcaaacagcctcagtgtgtttgatcttgtgtgtacgcgc
ttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcatccccctccctcgtttcataatcgcttgcacccaaccgc
aacttatctacgctgtcctgctatccctcagcgctgctctgctcctgctcactgccccctgcacagccttggttgggtccgctgt
attctcctggtactgcaacctgtaaacagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaag
ctt

SEQ ID NO: 118

C.reinhardtii β -tubulin promoter::native PTS::Umbellularia californica C12
thioesterase::C.vulgaris nitrate reductase 3'UTR

ggtaccgctgcaacgaaggcgacccagcgcctccacccgcccgtgaaccgacacgtgcttggcgccctgccgctgctg
ccgcatgcttgtgctggtgaggctgggagtgctgcatgctgattgaggcttggttcacgggtggaagcttatgtgtgtgctggg
cttgcacgcccgaatgcgcatggtggcaagagggcgagcacttgcggagctgccggtgcctccaggtggttcaatcgc
ggcagccagagggatttcagatgatcgcgctacaggttgagcagcagtgctcagcaaaggtagcagtttgccagaatgatcggtt

cagctgttaatcaatgccagcaagagaaggggtcaagtgcaaacacgggcatgccacagcacgggcaccggggagtggaaatgg
caccaccaagtgtgtgcgagccagcatcgccgctggctgtttcagctacaacggcaggagtcatccaacgtaaccatgagctga
tcaacactgcaatcatcgggcgggcggtgatgcaagcatgcctggcgaagacacatgggtgtcgggatgctgccggctgtgcctgt
gcgacgcccgttgagttggcagcaggctcagccatgcactggatggcagctgggctgccactgcaatgtggatagggatgcaa
gtggagcgaataccaaaccctctggctgcttgcgtgggtgcatggcatcgacccatcagcaggagcgcatgcaagggactggcc
ccatgcacgcatgccaaaccggagcgacccagtggtccacactgtcaccaggcccgaagctttgcagaacctgctcatggac
gcatgtagcgctgacgtcccttgacggcgctcctctcggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcggcagc
agagcgggcggcagcagcgggggggccacccttctgcggggctgcgccccagccagcggtgatgcgtgatcccaaagagttc
acattcatttgcagcctggagaagcgaggctggggcctttgggctgggtgcagccgcaatggaatggggaccgccaggctagc
agcaaaggcgctcccctactccgcatgatgtccatagtgcattggactgcatttgggtggggcgggcggtgtttcttctgtgtt
gcaaaacgcgcagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcacgccgatcagctgattgccggct
cggaagtaggcgcctccttctgctgcctcctcctgcctcccgctctagaatatcaATGATCGAGCAGGACGGCCTCC
ACGCCGGCTCCCCGCCGCTGGGTGGAGCGCCTGTCGGCTACGACTGGGCCCAGCAGACCATCG
GCTGCTCCGACGCCGCGTGTCCGCCTGTCCGCCAGGGCCGCCCGTGTGTTCTGTGAAGACCGA
CCTGTCCGGCGCCTGAACGAGCTGCAGGACGAGGCCGCCCGCCTGTCTGGCTGGCCACCACCGG
CGTGCCCTGCGCCGCGTGTGGACGTGGTGACCGAGGCCGCGCGACTGGCTGCTGCTGGGCG
AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
ACGCCATGCGCCGCTGCACACCCTGGACCCCGCACCTGCCCTTCGACCACCAGGCCAAGCACCG
CATCGAGCGCGCCCGCACCCGCATGGAGGCCGCGCTGGTGACCAGGACGACCTGGACGAGGAGC
ACCAGGGCCTGGCCCCCGCCGAGCTGTTGCCCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCCTGCCTGCCAACATCATGGTGAGAACGGCCGCTTCTCCGGCT
TCATCGACTGCGGCCGCTGGGCGTGCCGACCGCTACCAGGACATCGCCTGGCCACCCGCGACA
TCGCCGAGGAGCTGGGCGGCGAGTGGGCGGACCGCTTCTGGTGCTGTACGGCATCGCCGCCCCCG
ACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAcaattggcagcagcagctcggatagt
atcgacacactctggagcgtggtcgtgtgatggactgttgcgcacacttgccttgacctgtgaatatccctgccgtttat
caaacagcctcagtggtttgatcttgtgtgtacgcgttttgcgagttgtagctgcttgtgctatttgcgaataccacccccagc
atccccctccctcgtttcatatcgcttgcacccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgct
cactgccccctgcacagccttggttgggctccgcctgtatttctcgtgactgcaacctgtaaacagcactgcaatgctgatgc
acgggaagtagtggttggttggaacacaatggaggatcccgctcctgaacagagcgcgagaggaacgctgaaggtctcgcc
tctgtcgacctcagcgcgccatacaccacaataaccacctgacgaatgcgcttgggttctcgtccattagcgaagcgtccggtcac
acacgtgccacgttgccgaggtggcaggtgacaatgatcggtggagctgatggctgaaacgttcacagcctagggatatcgaattc
ctttcttgcgtatgacacttcagcaaaaggtaggggcggtgcgagacggcttccggcgctgcatgcaacaccgatgatgctt
cgacccccgaagctccttcggggctgcatggggcgctccgatgccgctccaggcgagcgctgtttaaatagccaggccccgatt
gcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatc
gcactccgctaaggggcgcccttcttctcgtttcagtcacaaccgcaaacggcgcgccATGGCCACCACCAGCCTGG
CCTCCGCTTCTGCTCCATGAAGGCCGTGATGCTGGCCCCGCGACGGCCGCGCATGAAGCCCCGCA
GCTCCGACCTGCAGCTGCGCGCCGCAACGCCCCACCTCCCTGAAGATGATCAACGGCACCAAGTT
CAGCTACACCGAGAGCCTGAAGCGCCTGCCGACTGGTCCATGCTGTTCCGCGTGATCACCACCATC
TTCAGCGCCGCGGAGAAGCAGTGGAACCACTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCAGCTG
CTGGACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCCGACCTTCGCCATCCGCTCCTACGAGG
TGGGCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACC
ACGCCAAGAGCGTGGGCATCCTGGGCGACGGCTTCGGCACCACTGGAGATGTCCAAGCGCGACC
TGATGTGGGTGGTGCGCCGACCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTG
GAGGTGGAGTGCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCTGGTGCGCGA
CTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGATGAACACCCGACCCGC
CGCCTGAGCACCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCCGCCTTCATCGACAACGTGGCC
GTGAAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGG

CGGCCTGACCCCCCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGC
 CTGGGTGTTTCGAGACCGTGCCCGACAGCATCTTCGAGTCCACACATCAGCTCCTTCACCTGGAG
 TACCGCCGCGAGTGACCCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCAGCTCC
 GAGGCCGGCCTGGTGTGCGACCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCG
 CACCGAGTGCGCCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGC
 GTGATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGAC
 GACAAGTGATGAActcgaggcagcagcagctcgatagtagtacacactctggacgctggtcggtgatggactgttgc
gccacacttgctgccttgacctgtgaatatccctgccgttttatcaaacagcctcagtggtttgatcttggtgtacgcgctttg
cgagttgctagctgcttgctatttgcgaataccacccccagcatccccctccctcgttcatatcgcttgatcccaaccgcaact
tatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgcccccgacacagccttggtttgggctccgctgtattc
tcttggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtggttggaacacaaatggaagctt

SEQ ID NO: 119

C.reinhartii β -tubulin promoter::*P.moriformis* isopentenyl diphosphate synthase

PTS::*Umbellularia californica* C12 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

ggtaccgcctgcaacgcaagggcagccacagccgctccacccgcccgtgaaccgacacgtgcttggggcctgccgctgctg
ccgcatgcttgtgctggtgaggctgggcagtgctgcatgctgattgaggcttgggtcatcggttggaagcttatgtgtgctggg
cttgcagtcgggcaatgcgcatggtggcaagaggggcgagcacttgctggagctgccggtgctccagggtggtcaatgc
ggcagccagagggttgcagatgcgctgacaggttgagcagcagtgctgcaaaaggtagcagtttgcagaatgatcggtt
cagctgttaataatgccagcaagagaaggggtcaagtgcacacacgggcagtcacacagcaggggcacggggagtggaatgg
caccaccaagtgtgtgcgagccagcatgccgctggctgtttcagctacaacggcaggagtcaccaacgtaaccatgagctga
tcaacactgcaatcatcgggcggtgatgcaagcatgctggcgaagacacatggtgtgcggatgctgccggtgctgctgct
gcgcacgcccgttgagttggcagcaggctcagccatgcactggatggcagctgggctgcaatgtggtggataggatgcaa
gtggagcgaatacacaacctctggctgcttgcgtgggtgcatggcatgcacatcagcaggagcgcatgcgaaggagctggcc
ccatgcacgcatgccaaaccggagcgacccagtgctccacactgtcaccaggcccgaagcttgcagaaccatgctcatggac
gcatgtagcgtgacgtcccttgacggcgctcctctcggtgtgggaaacgcaatgcagcacaggcagcagaggcgccgagc
agagcgggcgagcagcgggggggccaccttcttgcggggtcgccccagccagcggtgatgcgctgatccaaacgagttc
acattcatttgcagctgctggagaagcgaggctggggccttgggctggtgcagcccgaatggaatgcgggaccgccaggctagc
agcaaaggcgctcccctactccgcatcgatgttccatagtgattggactgcatttgggtggggcgccggtgtttcttctggtt
gcaaaacgcgccagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcacgccgatcagctgattgccggct
cggaagtaggcgcccctcttctgctcgccctctctccgtccccgctctagaatatcaATGATCGAGCAGGACGGCCTCC
 ACGCCGGCTCCCCGCCGCTGGGTGGAGCGCCTGTTCCGCTACGACTGGGCCCAGCAGACCATCG
 GCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGCCGCCCGTGTGTTCTGTAAGACCGA
 CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCGCCCTGTCTGGCTGGCCACCACCGG
 CGTGCCCTGCGCCGCCGTGCTGGACGTGGTGACCGAGGCCGCCGCCGCGACTGGCTGCTGCTGGGCG
 AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
 ACGCCATGCGCCGCCTGCACACCTGGACCCCGCCACCTGCCCCCTCGACCACCAGGCCAAGCACCG
 CATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGACCGAGGACGACCTGGACGAGGAGC
 ACCAGGGCCTGGCCCCCGCCGAGCTGTTCCGCCGCTGAAGGCCCGCATGCCGACGGCGAGGACC
 TGGTGGTGACCCACGGCGACGCTGCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCT
 TCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
 TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGTGTACGGCATCGCCGCCCCCG
 ACTCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAcaattggcagcagcagctcgatagtag
atcgacacactctggacgctggtcggtgatggactgttgcgccacacttgctgccttgacctgtgaatatccctgccgtttat
caaacagcctcagtggtttgatcttggtgtgtacgcgcttttgcgagttgctagctgcttgctatttgcgaataccacccccagc
atccccctccctcgttcatatcgcttgatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgct

cactgccccgcacagccttggtttgggctccgctgtattctcctggtactgcaacctgtaaacagcactgcaatgctgatgc
 acgggaagtagtgggatgggaacacaaaatggaggatcccgcgtctcgaacagagcgcgagaggaacgctgaaggtctcgcc
 tctgtcgacactcagcgggcataccacaataaccacctgacgaatgcgcttggttctcgtccattagcgaagcgtccggttac
 acacgtgccacgttgccgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatcgaattc
 ctttcttgcgctatgacacttccagcaaaaggtagggcggtcgagacggcttcccggcgctgcatgcaacaccgatgatgctt
 cgacccccgaagctccttcggggctgcatggcgctccgatgccgtccagggcgagcgctgtttaaataagccaggccccgatt
 gcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatc
 gcactccgctaagggggcgcttcttcttctgtttcagtcacaacccgcaaacactagtATGACGTTCCGGGTGCGCCCTC
 CCGGCCATGGGCGCGGTGTCTCCCTCCCCGGCCAGGGTTCGCGGTGCGCGCCAGTCGGCGAGT
 CAGGTTTTGGAGAGCGGGCGCGCCCCGACTGGTCCATGCTGTTCCGCGTGATACCACCATCTTCA
 GCGCCGCCGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCAGCTGCTG
 GACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCGGCACCTTCGCCATCCGCTCCTACGAGGTGG
 GCCCCGACCGCAGCACCTCCATCTGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACCACG
 CCAAGAGCGTGGGCATCCTGGGCGACGGCTTCGGCACACCCTGGAGATGTCCAAGCGCGACCTGA
 TGTGGGTGGTGCGCCGACCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACCCGTGGAG
 GTGGAGTGCTGGATCGGCGCCAGCGCAACAACGGCATGCGCCGCGACTTCCTGGTGCAGCGACTG
 CAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGATGAACACCCGCACCCGCCG
 CTGAGCACCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCGCCTTCATCGACAACGTGGCCGTG
 AAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGG
 CCTGACCCCCCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGCCTG
 GGTGTTGAGACCGTGCCCGACAGCATCTTCGAGTCCACCACATCAGCTCCTTACCCTGGAGTAC
 CGCCGCGAGTGACCCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCGAGCTCCGAG
 GCCGGCCTGGTGTGCGACCACTGCTGCAGCTGGAGGGCGGCGAGCGAGGTGCTGCGCGCCCGCAC
 CGAGTGGCGCCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGCGT
 GATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGA
 CAAGTGATGAGctcaggcagcagcagctcgatagtagtacacactctggacgctggtcgtgtgatggactgttgcgc
 acacttgctgcttgacctgtgaatatccctgcgcttttatcaaacagcctcagtggtttgatctgtgtgtacgcgcttttgcga
 gttgtgtagctgctgtgtctatttgcgaataccacccccagcatccccttccctcgttcatatcgcttgcatcccaaccgcaacttat
 ctacgctgtcctgtatccctcagcgctgtcctgtcctgctcactgccccgcacagccttggtttgggctccgctgtattctcc
 tggtagtgcacacctgtaaacagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaaatggaaagctt

SEQ ID NO: 120

C.reinhardtii β -tubulin promoter::*P.moriformis* delta 12 fatty acid desaturase

PTS::*Umbellularia californica* C12 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

ggtaccgcctgcaacgcaagggcagccacagccgtcccacccgcgctgaaccgacacgtgcttgggcgcttgcgcctgctg
 ccgcatgcttgtgctggtaggctgggcagtgctgcatgctgattgaggcttggttcatcgggtggaagcttatgtgtgtgctggg
 cttgcatgccgggaatgcgcatggtggcaagaggcggcagcacttctgagctgcccgggtgcctccagggtggttcaatcgc
 ggcagccagagggttcatgatgatcgcgctacaggttgagcagcagtgtagcaaaaggttagcagtttgcagaatgatcggtt
 cagctgttaataatgccagcaagagaaggggtcaagtgcacaacgggcatgccacagcacgggcacggggagtggaatgg
 caccaccaagtgtgtgcgagccagcatcgccgctggctgtttcagctacaacggcaggagtcataacgtaaccatgagctga
 tcaactgcaatcatcgggcggtgctgatcaagcatgcctggcgaagacacatggtgtgcggatgctgccggctgctgcctgct
 gcgcacgccgttgagttggcagcaggtcagccatgactggtggcagctgggtgccaactgcaatgtggtggataggatgcaa
 gtggagcgaataccaaacctctggctgcttgcgtgggttcagtgcatcgccatcagcaggagcgcatgcaagggactggcc
 ccattgcacgccatgccaaaccggagcgacccagtgctcacactgtcaccaggcccgaagctttgcagaacctgctcatggac
 gcatgtagcgtgacgtcccttgacggcgctcctcgggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcggcagc
 agagcggcggcagcagcggcgggggccacccttcttcgggggtcgcgccccagccagcggtgatgcgctgatcccaaacgagttc

acattcatttgcctggagaagcgaggctggggcctttgggctgggtgcagcccgcaatggaatcggggaccgccaggctagc
agcaaaggcgctcccactccgcctcgcctgctggtccatagtgctggtgactgcatttgggtggggcgccggctgtttcttcgtgtt
gcaaaacgcgcagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcagccgatcagctgattgccggct
cggaagtaggcgccccctttctgctgcctctctccgtccgcctctagaatatcaATGATCGAGCAGGACGGCCTCC
ACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGCCAGCAGACCATCG
GCTGCTCCGACGCCGCCGTGTTCCGCCTGTCCGCCAGGGCCGCCCGTGTGTTCTGTAAGACCGA
CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCTGTCTGGCTGGCCACCACCGG
CGTGCCCTGCGCCCGCTGCTGGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCG
AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
ACGCCATGCGCCGCTGCACACCTGGACCCCGCCACCTGCCCCCTTCGACCACCAGGCCAAGCACCG
CATCGAGCGCGCCCGCACCCGCATGGAGGCCGCCCTGGTGACCAGGACGACCTGGACGAGGAGC
ACCAGGGCCTGGCCCCCGCCGAGCTGTTCCGCCCTGAAGGCCCGCATGCCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCCTGCCTGCCAACATCATGGTGAGAACGGCCGCTTCTCCGGCT
TCATCGACTGCGGCCGCTGGGCGTGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGCTGTACGGCATCGCCGCCCCCG
ACTCCAGCGCATCGCTTCTACCGCTGCTGGACGAGTTCTTCTGAcaattggcagcagcagctcgatagt
atcgacacactctggacgctggtcgtgtgatggactgttgcgcacacttgccttgacctgtgaatatccctgccgtttat
caaacagcctcagtgtgtttgatcttgtgtgtacgcgttttgcgagttgctagctgcttgtgtatttgcgaataccacccccagc
atccccctccctgcttcatatcgcttgcacccaaccgaacttatctacgctgtcctgctatccctcagcgtgctcctgctcctgct
cactgccccctgcacagccttgggttgggctccgcctgtattctcctggtactgcaacctgtaaacagcactgcaatgctgatgc
acgggaagtgtggttgggaacacaaatggaggatcccgctctcgaacagagcgcgaggaacgctgaaggtctcgcc
tctgtgcacctcagcgccgataccacaataaccacctgacgaatgcgcttgggttcttcgtccattagcgaagcgtccggttcac
acacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatcgaattc
cttcttgcgtatgacacttcagcaaaaggtagggcgggctgcgagacggcttccggcgctgcatgcaacaccgatgatgctt
cgacccccgaagctccttcggggctgcatgggcgtccgatgccgtccaggcgagcgctgtttaaatagccaggccccgatt
gcaaagacattatagcgagctacaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatc
gcactccgctaaggggcgcttcttcttcttcttcagtcaaacccgcaaacactagtATGGCTATCAAGACGAACAG
GCAGCCTGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTT
CGAGCGCTCGGCGCTTCGTGGGCGCGCCCCGACTGGTCCATGCTGTTCCGGTGATACCACCATC
TTCAGCGCCGCCGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCAGCTG
CTGGACGACCACTTCGGCCTGCACGGCCTGGTGTTCCGCCGACCTTCGCCATCCGCTCCTACGAGG
TGGGCCCGACCGCAGCACCTCCATCCTGGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACC
ACGCCAAGAGCGTGGGCATCCTGGGCGACGGCTTCGGCACCAACCTGGAGATGTCCAAGCGCGACC
TGATGTGGGTGGTGCGCCGACCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTG
GAGGTGGAGTGCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGGTGCGCGA
CTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGATGAACACCCGACCCGC
CGCCTGAGCACCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCCGCCCTTCATCGACAACGTGGCC
GTGAAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGG
CGGCCTGACCCCCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGC
CTGGGTGTTGAGACCGTGCCCGACAGCATCTTCGAGTCCACACATCAGCTCCTTACCCTGGAG
TACCGCCGCGAGTGACCCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCAGCTCC
GAGGCCGGCCTGGTGTGCGACCACTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCG
CACCGAGTGCGCCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGC
GTGATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGAC
GACAAGTGATGActcaggcagcagcagctcgatagtatcgacacactctggacgctggtcgtgtgatggactgttgc
gccacacttgccttgacctgtgaatatccctgccgttttatcaaacagcctcagtgtgtttgatcttgtgtacgcgttttg
cgagttgctagctgcttgtgctatttgcgaataccacccccagcatccccctccctgcttcatatcgcttgcacccaaccgaact

tatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgccccgcacagccttggttgggctccgcctgtattc
tcctggtagctgaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagctt

SEQ ID NO: 121

C.reinhardtii β -tubulin promoter::*C.protothecoides* stearyl ACP desaturase

PTS::*Umbellularia californica* C12 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

ggtagccgcctgcaacgaagggcagccacagccgctccaccccgctgaaccgacacgtgcttgggcgcctgccgcctgcctg
ccgcatgcttgtgctggtgaggctgggcagtgctgcatgctgattgaggcttggttcacgggtggaagcttatgtgtgtgctggg
cttgcatgccgggcaatgcgcatggtggcaagagggcggcagcacttgctggagctgccggtgctccaggttggttcaatcgc
ggcagccagagggatttcagatgatcgcgctacaggttgagcagcagtgctagcaaggttagcagtttgcagaatgatcggtt
cagctgttaataatgccagcaagagaaggggtcaagtgcacaacacgggcatgccacagcacgggcacccggggagtggatgg
caccaccaagtgtgtgcgagccagcatcgccgctggtgttcagctacaacggcaggagtcaccaacgtaacctagagctga
tcaacactgcaatcatcgggcgggcggtgatgaagcatgcctggcgaagacacatggtgtgcggatgctgccggctgctgctgct
gcgcacgcccgttgagttggcagcaggctcagccatgcactggatggcagctgggctgcccactgcaatgtggtggataggatgcaa
gtggagcgaataccaaacctctggctgcttgcctgggttgatggcatcgccatcagcaggagcgcatgcgaagggaactggcc
ccatgcacgcatgccaaaccggagcgacccaggtgtccacactgtcaccaggcccgcaagcttgcagaacctgctcatggac
gcatgtagcgctgacgtccctgacggcgctcctctcggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcggcagc
agagcgggcggcagcagcgggcgggggccaccttcttgcggggtcgcgccccagccagcggtgatgcgtgatcccaaagagttc
acattcatttgatgcctggagaagcgaggctggggccttgggctggtgcagcccgaatggaatgcgggaccgcaggctagc
agcaaaggcgctccctactccgcacatgatgttccatagtgcatggtgactgcatttgggtggggcgccggctgttcttctgtgt
gcaaaacgcgcagctcagcaacctgtccgtgggtccccgtgccgatgaaatcgtgtgcacgccgatcagctgattgccggct
cggaagtaggcgcccctcttctgctcgccctctctcgtcccgctctagaatatcaATGATCGAGCAGGACGGCCTCC
ACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTCCGCTACGACTGGGCCAGCAGACCATCG
GCTGCTCCGACGCCGCGTGTCCGCCTGTCCGCCAGGGCCGCCCGTGTGTTCTGTGAAGACCGA
CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCTGTCTGGCTGGCCACCACCGG
CGTGCCCTGCGCCGCGTGTGGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCG
AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCGAGAAGGTGTCCATCATGGCCG
ACGCCATGCGCCGCTGCACACCCTGGACCCCGCCACCTGCCCCCTCGACCACCAGGCCAAGCACCG
CATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGACCGAGGACCTGGACGAGGAGC
ACCAGGGCCTGGCCCCCGCCGAGCTGTTCCGCCCTGAAGGCCCGCATGCCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCTGCCTGCCAACATCATGGTGGAAGACGGCCGCTTCTCCGGCT
TCATCGACTGCGGCCGCTGGGCGTGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGCTGTACGGCATCGCCGCCCGG
ACTCCCAGCGCATCGCCTTCTACCGCTGCTGGACGAGTTCTTCTGAAaattggcagcagcagctcgatagt
atcgacacactctggacgtggtgctgtgatggactgttgccgcacacttgctgccttgacctgtgaatatccctgccgttttat
caaacagcctcagtgtgtttgatcttgtgttacgcgttttgcgagttgctagctgcttgtgtatttgcgaataccacccccagc
atccccctccctcgtttcataatcgcttgcatcccaaccgcaacttatctacgtgtcctgctatccctcagcgctgctcctgctcctgct
cactgccccctgcacagccttggttgggctccgcctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgc
acgggaagtagtgggatgggaacacaaatggaggaccccgcgtcgaacagagcgcgaggaacgctgaaggctcgc
ctgtgcacctcagcgcgccatacaccacaataaccacctgacgaatgcgcttggttcttctgctccattagcgaagcgtccggttac
acacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatcgaattc
ctttcttgcgctatgacacttcagcaaaaggtaggggcgggctgcgagacggcttccggcgctgcatgcaacaccgatgatgctt
cgacccccgaagctccttcggggctgcatggggctccgatgccgtccaggcgagcgctgtttaaatagccaggccccgatt
gcaaaagacattatagcgagctacaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttggatc
gcactccgctaaggggcgcccttctcttcttcttcagtcaaacccgcaaacactagATGGCCACCGCATCCACTTTCT
CGGCGTTCAATGCCCGCTGCGGCGACCTGCGTGCCTCGGCGGGCTCCGGGCCCGGCGCCAGCGA

GGCCCCCTCCCCGTGCGCGGGCGCGCCCCGACTGGTCCATGCTGTTGCGCGTGATCACCACCATCTT
 CAGCGCCGCGGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAAGCTGCCCCAGCTGCT
 GGACGACCACTTCGGCCTGCACGGCCTGGTGTTCGCGCGACCTTCGCCATCCGCTCCTACGAGGTG
 GGCCCCGACCGCAGCACCTCCATCCTGGCCGTGATGAACCACATGCAGGAGGCCACCCTGAACCAC
 GCCAAGAGCGTGGGCATCCTGGGCGACGGCTTCGGCACCACCCTGGAGATGTCCAAGCGCGACCTG
 ATGTGGGTGGTGCGCCGACCCACGTGGCCGTGGAGCGCTACCCACCTGGGGCGACACCGTGGA
 GGTGGAGTGCTGGATCGGCGCCAGCGGCAACAACGGCATGCGCCGCGACTTCCTGGTGCGCGACT
 GCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGCTGATGAACACCCGACCCGCC
 GCCTGAGCACCATCCCCGACGAGGTGCGCGGCGAGATCGGCCCGCCTTCATCGACAACGTGGCCG
 TGAAGGACGACGAGATCAAGAAGCTGCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGC
 GGCTGACCCCCCGCTGGAACGACCTGGACGTGAACCAGCACGTGAACAACCTGAAGTACGTGGCC
 TGGGTGTTGAGACCGTGGCCGACAGCATCTTCGAGTCCACCACATCAGTCTCTTCACCTGGAGT
 ACCGCCGCGAGTGACCCGCGACTCCGTGCTGCGCAGCCTGACCACCGTGAGCGGCGGCAGCTCCG
 AGGCCGCGCTGGTGTGCGACCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCCGC
 ACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCGCGGCATCAGCGTGATCCCCGCCGAGCCCCGC
 GTGATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGAC
 GACAAGTGATGActcgaggcagcagcagctcgatagtagacacactctggacgctggtcgtgtgatggactgttgc
gccacacttgctgcttgacctgtgaatatccctgccgttttatcaaacagcctcagtggtttgatcttgtgtacgcgcttttg
cgagttgctagctgctgtgtatttgcgaataccacccccagcatcccttccctcgttcatatcgcttgcatcccaaccgcaact
tatctacgctgtctgtatccctcagcgctgctcgtctcgtcactgcccctgcacagccttggttgggctccgctgtattc
tcttggtactgcaacctgtaaaccagcactgcaatgctgtatgcacgggaagtagtgggatgggaacacaaatggaagctt

SEQ ID NO: 122

C.reinhardtii β -tubulin promoter::*P.moriformis* stearoyl ACP desaturase PTS::*Umbellularia californica* C12 thioesterase::*C.vulgaris* nitrate reductase 3'UTR

ggtaccgctgcaacgcaagggcagccacagccgctccacccgctgaaccgacacgtgcttgggcgctgccgctgctg
ccgcatgcttgtgctggtgaggctgggcagtgctgcatgctgattgaggcttggttcatcggttggaagcttatgtgtgtgctggg
cttgcagtcgggcaatgcatggtggcaagagggcggcagcacttgctggagctgccggtgctccagtggttcaatgcg
ggcagccagagggatttcagatgatcgcgctacaggttgagcagcagtgctcagcaaaggttagcagtttgccagaatgatcggtt
cagctgttaataatgccagcaagagaaggggtcaagtgcaaacgggcatgccacagcacgggacacggggagtggaatgg
caccaccaagtgtgtgcgagccagcatcgccgctggctgttgcagctacaacggcaggagtcacaaacgtaacctgagctga
tcaactgcaatcatcgggcgggcgctgatgcaagcatgctggcgaagacacatggtgtgcggatgctgcggctgctgctgct
gcgcacgcccgttgagttggcagcaggctcagccatgactggatggcagctgggctgccactgcaatgtggtggataggatgcaa
gtggagcgaataccaaacctctggctgcttgcgtgggtgcatggcatcgaccatcagcaggagcgcatgcgaagggactggcc
ccatgcacgcatgcaaacggagcgacccagtgctccactgtcaccaggcccgaagcttgcagaacctgctcatggac
gcatgtagcgctgacgtcccttgacggcgctcctctcggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcggcagc
agagcgggcggcagcagcgggggggccaccttcttgcggggtcgcgccccagccagcggtgatgcgtgatcccaaacgagttc
acattcatttgcatgctggagaagcgaggctggggccttgggctggtgcagcccgaatggaatcgggaccgaccaggctagc
agcaaaggcgctccctactccgcatcgatgttccatagtgacttgactgatttggtggggcgccggctgttcttctgtgtt
gcaaaacgcgcagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcacgccgatcagctgattgccggct
cgcgaagtaggcgcccctcttctgctgcctctctcgtccccgctctagaatatcaATGATCGAGCAGGACGGCCTCC
 ACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCTGTTGCGCTACGACTGGGCCAGCAGACCATCG
 GCTGCTCCGACGCCGCGCTGTTCCGCTGTCCGCCAGGGCCGCCCGCTGCTGTTCTGTAAGACCGA
 CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCTGCTGCTGGCTGGCCACCACCGG
 CGTGCCCTGCGCCGCGCTGCTGGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCG
 AGGTGCCCCGCCAGGACCTGCTGTCCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCG

63

tgc atgccgggcaatgcgc atggtggcaagagggcggcagcacttgctggagctgccg ggtgcctccaggtggttcaatcgcggc
agccagagggatttcagatgatcgcgctacaggttgagcagcagtgctagcaaggtagcagtttgccagaatgatcggttcagc
tgtaataatgccagcaagagaaggggtcaagtcaaacacgggcatgccacagcacgggcacccggggagtggaatggcacca
ccaagtgtgtgcgagccagcatcgccgctggctgtttcagctacaacggcaggagtcaccaacgtaacatgagctgatcaaca
ctgcaatcatcgggcgggcggtgatgcaagcatgcctggcgaagacacatggtgtgcggatgctgccggctgctgctgctgcgcac
gccgttgagttggcagcaggctcagccatgcactggatggcagctgggctgccactgcaatgtggtggataggatgcaagtggagc
gaataccaaaccctctggctgcttgctgggttgcatggcatcgaccatcagcaggagcgcatgcaagggactggccccatgcac
gccatgccaaaccggagcgacccagtgctccacactgtcaccaggccccgaagctttgcagaacatgctcatggacgc atgtagc
gctgacgtcccttgacggcgctcctctcggtgtgggaaacgaatgcagcacaggcagcagaggcgggcggcagcagagcgggcg
cagcagcgggcgggggccacccttctcggggtcgcgccccagccagcggtgatgcgctgatcccaaacgagttcacattcatttgc
atgcctggagaagcgaggctggggcctttgggctggtgcagcccgaatggaatgcgggacccgaggttagcagcaaaaggcgcc
tcccactactcgcatcgatgttccatagtgacttgacttgacttggtggggcgccggctgtttcttctgtgtgcaaacgcgcca
gctcagcaacctgtcccgtgggtccccgtgccgatgaaatgctgtgcacgcgatcagctgattgccccgctcggaagtagcggc
cctcctttctgctcgccctctctccgtccccgctctagaatatcaATGATCGAGCAGGACGGCCTCCACGCCGGCTCCC
CCGCCGCTGGGTGGAGCGCCTGTTCCGCTACGACTGGGCCAGCAGACCATCGGCTGCTCCGACG
CCGCCGTGTTCCGCTGTCCGCCAGGGCCGCCCGTGCTGTTCTGTAAGACCGACCTGTCCGGCGC
CCTGAACGAGCTGCAGGACGAGGCCGCCGCCTGTCTGGCTGGCCACCACCGGCGTGCCCTGCGC
CGCCGTGCTGGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCGAGGTGCCCGGCC
AGGACCTGCTGTCTCCACCTGGCCCCCGCCGAGAAGGTGTCCATCATGGCCGACGCCATGCGCCG
CCTGCACACCCTGGACCCCGCCACCTGCCCTTCGACCACCAGGCCAAGCACCGCATCGAGCGCGCC
CGCACCCGCATGGAGGCCGGCCTGGTGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGC
CCCCGCCGAGCTGTTCCGCCGCTGAAGGCCCGCATGCCCGACGGCGAGGACCTGGTGGTGACCCA
CGGCGACGCCTGCCTGCCCAACATCATGGTGAGAACGGCCGCTTCTCCGGCTTCATCGACTGCGGC
CGCCTGGGCGTGCCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACATCGCCGAGGAGCTG
GGCGGCGAGTGGGCCGACCGCTTCTGGTGCTGTACGGCATCGCCGCCCCGACTCCCAGCGCATC
GCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAcaattggcagcagcagctcggaatagtatgcacacactctggac
gctggtcgtgtgatggactgttgccgccacacttgctgccttgacctgtgaatatccctgccgttttatcaaacagcctcagtgtg
tttgatcttgtgtgacgcgttttgcgagttgctagctgcttgtgctatttgcgaataccccccagcatcccccttccctcgtttca
tatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgcccctgcacag
ccttggtttgggctccgctgtattctctggtactgcaacctgtaaacagcactgcaatgctgtgacgggaagtagtgga
tggaacacaaatggagggatcccgctctcgaacagagcgcgagaggaacgctgaaggctcgcctctgtgcacctcagcgc
ggcatacaccacaataaccacctgacgaatgcgcttggttcttgcctcattagcgaagcgctccggttcacacacgtgccacgttggc
gaggtggcaggtgacaatgatcggtggagctgatggtcgaacggttcacagcctagggatatcgaattccttcttgcgctatgaca
cttcagcaaaaaggtaggggcggtgcgagacggcttccggcgctgc atgcaacaccgatgatgcttcgacccccgaagctcct
cggggctgc atggcgctccgatgccgctccaggcgagcgctgtttaatagccaggcccccgattgcaaagacattatagcgag
ctaccaaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtatcgactccgctaaggggcg
ctcttctcttctgttcagtcacaaccgcaaacggcgcgccATGGCCACCACCTCCCTGGCCTCCGCTTCTGCAGC
ATGAAGGCCGTGATGCTGGCCCGCGACGGCCGCGCATGAAGCCCCGCTCCAGCGACCTGCAGCTG
CGCGCCGCAACGCCAGACCTCCCTGAAGATGATCAACGGCACCAAGTTCTCTACACCGAGAGCC
TGAAGAAGCTGCCGACTGGTCCATGCTGTTCCGGTGATCACCACCATCTTCTCCGCCGCCGAGAA
GCAGTGACCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCAGCTGCTGGACGACCACTTCGG
CCCCACGGCCTGGTGTTCGCCGACCTTCGCCATCCGCAGCTACGAGGTGGGCCCGACCGTCC
ACCAGCATCGTGCCGTGATGAACCACCTGCAGGAGGCCGCCCTGAACCACGCCAAGTCCGTGGGC
ATCCTGGGCGACGGCTTCGGCACCACTGGAGATGTCCAAGCGCGACCTGATCTGGGTGGTGAAG
CGACCCACGTGGCCGTGGAGCGCTACCCCGCTGGGGCGACACCGTGGAGGTGGAGTGCTGGGT
GGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCTGGTGCGGACTGCAAGACCGGCGAGAT
CCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCCGCCGCTGAGCAAGATCCCC

GAGGAGGTGCGCGGCGAGATCGGCCCCGCTTCATCGACAACGTGGCCGTGAAGGACGAGGAGAT
 CAAGAAGCCCCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGGCGGCCTGACCCCCGCTG
 GAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGACTGGATCCTGGAGACCGT
 GCCCGACAGCATCTTCGAGAGCCACCACATCTCTCTTCACCATCGAGTACCGCCGCGAGTGCACC
 ATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCGGCTCCTCCGAGGCCGCGCTGGTGTGC
 GAGCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAGACCGAGTGGCGCCCCAA
 GCTGACCGACTCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAGCGTGATGGACTACAAGGA
 CCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGACGACAAGTGATGAActcgag
gcagcagcagctcgatagtagtcgacacactctggacgctggtcgtgtgatggactgttgccgccacacttgccttgacct
gtgaatatccctgccgttttatcaaacagcctcagtggtttgatcttggtgtacgcgcttttgaggttgctagctgcttggt
atttgcaataccacccccagcatcccttccctcgtttcatactgcttgcatcccaaccgcaacttatctacgtgtcctgctatcc
ctcagcgctgctcctgctcctgctcactgccccctgcacagccttggtttgggctccgctgtattctcctggtactgcaacctgtaa
accagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaagctt

SEQ ID NO: 124

C.reinhardtii β -tubulin promoter::*P.moriformis* isopentenyl diphosphate synthase

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

ggtaccgcctgcaacgcaagggcagccacagccgctcccacccgccgtgaaccgacacgtgcttgggcgctgccgctgctg
ccgcatgcttgctggtgaggtggtgagctgctgcatgctgattgaggttggttcatcggttggaagcttatgtgtgctggg
cttgcagtcgggcaatgcgcatggtggcaagagggcgagcacttgcaggagctgccggtgctccaggtggttcaatgcg
ggcagccagagggatttcagatgatcgcgctacaggttgagcagcagtgctagcaaaaggttagcagtttgccagaatgatcggtt
cagctgttaataatgccagcaagagaaggggtcaagtgcacacagggcatgccacagcagggcaccggggagtggaatgg
caccaccaagtgtgtgcgagccagcatcgccgctggctgtttcagctacaacggcaggagtcacacgtaaccatgagctga
tcaacactgcaatcatcgggcggtgctgatcaagcatgcctggcgaagacacatggtgtgcggatgctgccggtgctgctgct
gcgacgcccgttgagttggcagcaggtcagccatgactggtggcagctgggtgccactgcaatgtggtggataggtgcaa
gtggagcgaataccaaacctctggctgcttgctgggttcagtcgcatcgccatcagcaggagcgatgcgaagggactggcc
ccatgcacgcatgccaaaccggagcgaccgagtgctcacactgtcaccaggcccgaagctttgcagaacctgctcatggac
gcatgtagcgtgacgtcccttgacggcgctcctcgggtgtgggaaacgcaatgcagcacaggcagcagggcgggcggcagc
agagcgggcgagcagcggcgggggccacccttctgcggggtgcgccccagccagcggtgatgcgctgatccaaacgagttc
acattcatttgcatgcctggagaagcgaggctggggcctttgggtggtgcagcccgaatggaatggggaccgaccaggtagc
agcaaaggcgctcccctactccgcatcgatgttccatagtgactggactgcatttgggtggggcgccggtgtttcttctgtgtt
gcaaaacgcgcagctcagcaacctgtcccgtgggtccccgtgccgatgaaatcgtgtgcacgccgatcagctgattgccggct
cgcgaagtagggcgccctctttctgctcgccctctctccgctccgctctagaatatcaATGATCGAGCAGGACGGCCTCC
 ACGCCGGCTCCCCCGCCGCTGGGTGGAGCGCCTGTTCCGGCTACGACTGGGCCAGCAGACCATCG
 GCTGCTCCGACGCCGCGTGTTCGCTGTCCGCCAGGGCCGCCCGTGTCTGCTGAAGACCGA
 CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCTGTCTGGCTGGCCACCACCGG
 CGTGCCCTGCGCCGCGTGTGGACGTGGTGACCGAGGCCGCGCGACTGGCTGCTGCTGGGCG
 AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGGCCCCCGCGAGAAGGTGTCCATCATGGCCG
 ACGCCATGCGCCGCTGCACACCTGGACCCCGCCACCTGCCCCCTCGACCACCAGGCCAAGCACCG
 CATCGAGCGCGCCCGCACCCGCATGGAGGCCGCGCTGGTGACCGAGGACGACCTGGACGAGGAGC
 ACCAGGGCCTGGCCCCCGCGAGCTGTTCCGCCGCTGAAGGCCCGCATGCCGACGGCGAGGACC
 TGGTGGTGACCCACGGCGACGCTGCCTGCCAACATCATGGTGGAGAACGGCCGCTTCTCCGGCT
 TCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
 TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGTGTACGGCATCGCCGCCCCCG
 ACTCCAGCGCATCGCTTCTACCGCCTGCTGGACGAGTTCTTCTGAaattggcagcagcagctcgatag
atcgacacactctggacgctggtcgtgtgatggactgttgccgccacacttgccttgacctgtgaatatccctgccgttttat

66

gcatgtagcgctgacgtcccttgacggcgctcctctcgggtgtgggaaacgcaatgcagcacaggcagcagaggcgggcggcagc
agagcgggcggcagcagcgggggggccacccttcttgcggggtcgcgccccagccagcggtgatgcgtgatcccaaacgagttc
acattcatttgcagtcctggagaagcgaggctggggcctttgggctggtgcagcccgcaatggaatcggggaccgccaggctagc
agcaaaggcgctccctactccgcatcgatgttccatagtcattggactgcatttgggtggggcgggcggtgtttcttctggtt
gcaaaacgcgcagctcagcaacctgtcccgtgggtccccgtccgatgaaatcgtgtgcacgcgatcagctgattgccggct
cggaagtaggcgcctccttctgctcgccctctcctcgtcccgctctagaatatcaATGATCGAGCAGGACGGCCTCC
ACGCCGGCTCCCCGCCGCTGGGTGGAGCGCCTGTTCGGCTACGACTGGGCCAGCAGACCATCG
GCTGCTCCGACGCCGCGTGTTCGCCTGTCCGCCAGGGCCGCCCGTGTGTTCTGTAAGACCGA
CCTGTCCGGCGCCCTGAACGAGCTGCAGGACGAGGCCGCCCGCTGTCTGGCTGGCCACCACCGG
CGTGCCCTGCGCCGCGTGTGACGTGGTGACCGAGGCCGCCGCGACTGGCTGCTGCTGGGCG
AGGTGCCCCGCCAGGACCTGCTGTCTCCACCTGCCCCCGCCGAGAAGGTGTCCATCATGGCCG
ACGCCATGCGCCGCTGCACACCTGGACCCGCCACCTGCCCTTCGACCACCAGGCCAAGCACCG
CATCGAGCGCGCCCGCACCCGCATGGAGGCCGGCCTGGTGACCAGGACGACCTGGACGAGGAGC
ACCAGGGCCTGGCCCCCGCCGAGCTGTTCCGCCGCTGAAGGCCCGCATGCCGACGGCGAGGACC
TGGTGGTGACCCACGGCGACGCTGCCTGCCAACATCATGGTGAGAACGGCCGCTTCTCCGGCT
TCATCGACTGCGGCCGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACA
TCGCCGAGGAGCTGGGCGGCGAGTGGGCCGACCGCTTCTGGTGCTGTACGGCATCGCCGCCCCCG
ACTCCCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTGAcaattggcagcagcagctcgatagt
atcgacacactctggacgtggtcgtgtgatggactgttccgccacacttgctgccttgacctgtgaatatccctgccgtttat
caaacagcctcagtggtttgatcttgtgtgtacgcgttttgcgagttgtagctgcttgtgctatttgcgaataccacccccagc
atccccctccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgtgctcctgctcgt
cactgccccctgcacagccttggtttgggctccgctgtatttctctggtactgcaacctgtaaacagcactgcaatgctgatgc
acgggaagttagtggttggaacacacaaatggaggatcccgctctcgaacagagcgcgagaggaaacgtgaaggctcgc
tctgtcgacctcagcgcgccatacaccacaataaacacctgacgaatgcgcttgggtcttctgctcattagcgaagcgtccggtcac
acacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggcgaaacgttcacagcctagggatatcgaaatc
ctttcttgcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccgcgctgcatgcaacaccgatgatgctt
cgacccccgaagctccttggggctgcagggcgctccgatccgctccagggcgagcgctgtttaaatagccaggccccgatt
gcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatc
gcactccgctaaggggcgcccttctccttctcgtttcagtcacaaccgcaaacactagtATGGCTATCAAGACGAACAG
GCAGCCTGTGGAGAAGCCTCCGTTACGATCGGGACGCTGCGCAAGGCCATCCCCGCGCACTGTTT
CGAGCGCTCGGCGCTTCGTGGGCGCGCCCCGACTGGTCCATGCTGTTCCGCGTGATCACCACCATC
TTCTCCGCCGCCGAGAAGCAGTGGACCAACCTGGAGTGGAAGCCCAAGCCCAACCCCCCAGCTG
CTGGACGACCACTTCGGCCCCACGGCCTGGTGTTCGCCGACCTTCGCCATCCGAGCTACGAGG
TGGGCCCCGACCGCTCCACCAGCATCGTGCCGTGATGAACCACCTGCAGGAGGCCGCCCTGAACC
ACGCCAAGTCCGTGGGCATCCTGGGCGACGGCTTCGGCACCAACCTGGAGATGTCCAAGCGCGACC
TGATCTGGGTGGTGAAGCGCACCCACGTGGCCGTGGAGCGCTACCCCGCCTGGGGCGACACCGTG
GAGGTGGAGTGCTGGGTGGGCGCCTCCGGCAACAACGGCCGCCGCCACGACTTCTGGTGCGCGA
CTGCAAGACCGGCGAGATCCTGACCCGCTGCACCTCCCTGAGCGTGATGATGAACACCCGCACCCGC
CGCCTGAGCAAGATCCCCGAGGAGGTGCGCGGCGAGATCGGCCCCGCTTCATCGACAACGTGGCC
GTGAAGGACGAGGAGATCAAGAAGCCCCAGAAGCTGAACGACTCCACCGCCGACTACATCCAGGG
CGGCCTGACCCCCGCTGGAACGACCTGGACATCAACCAGCACGTGAACAACATCAAGTACGTGGA
CTGGATCCTGGAGACCGTGCCCGACAGCATCTTCGAGAGCCACCACATCTCCTCCTTACCATCGAGT
ACCGCCGCGAGTGACCATGGACAGCGTGCTGCAGTCCCTGACCACCGTGAGCGGCGGCTCCTCCG
AGGCCGGCTGGTGTGCGAGCACCTGCTGCAGCTGGAGGGCGGCAGCGAGGTGCTGCGCGCCAAG
ACCGAGTGGCGCCCCAAGCTGACCGACTCCTTCCGCGGCATCAGCGTGATCCCCGCCGAGTCCAGC
GTGATGGACTACAAGGACCACGACGGCGACTACAAGGACCACGACATCGACTACAAGGACGACGAC
GACAAGTGATGAActgaggcagcagcagctcgatagtatcgacacactctggacgctggtcgtgtgatggactgttgc

*gccacacttgctgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtggtttgatcttgtgtgtacgcgcttttg
cgagttgctagctgcttgtgtctatttgcgaataccacccccagcatccccttcctcgcttcatatcgcttgcatcccaaccgcaact
tatctacgctgtcctgtatccctcagcgctgtcctgtcctgtcactgcccccgacacagccttggttgggctccgcctgtattc
tcctggtactgcaacctgtaaacacagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagctt*

SEQ ID NO: 126

Homologous recombination construct SZ726

CCCGTGATCACACAGGTGCCTTGCGAGCGTGATCACACTATTTTGGGGGTCCTAC
AGTACTGAAATGGTGAGAAAGTCGTAAGTCAAGGATGAACAATGAAAATGG
TGCTGTGGTGGCTTCTCAAAGGTCAAGAATCAGTCGCTCGCGTCAGGAAATCGCG
GCGTCAACCAGCGTGGGCGCGGTTCAGTGGCCCCGCACTGGTCAACATAGCCTCTC
CTGCCACAGTAGCGATCCCCTGGGCGTTCCTCTCAGCAGCGGCTGTACTGCCTC
CCAGATTTTCTTCTTCTGGACCTGCGGGCGTGAGAGGATGAGCAGGGTGGGGCCA
AGGGCTCAATCCTGAACGGCCCTCATTCGGTTTCCAATCCCACAACACATACCCA
CAGCAGGTTCAGACCACGCATTCGCACCATGCGCACCAATAACGTGTCCITACCT
GATTGGGTGTGGCAGGCTCCGTGGACAGGAGTGCCTCGTCCCCCGCCAGACCC
GCTCCCCCGTTCACGGCGGCGTCCGGGACCCGCAGCGGCTCCACCGCGGTGTGATC
CGCGTTGGCGGCGCAGAGCAGCATCCAGCCGATTTGACCCCGCGCATGCTCCG
AGGCTTGAGGTTGGCCAGCACCAACCCGCCGGCCGACAAGGTCTCTCCAGGGT
CACGTGCCGGACCAGGCCACTCACGATGGTGCGAGGGCCCCCTCTCGCCGAG
GTCGATCTGCTCGACGTACAGACTGCGACATGCGTGGCGAGTGGTCATCAGAAG
GAAGCAGGTGTGCAGAAAGGGGCACGTGGTTGGTATTGAGAGTAGCCAAAGCTTT
GTGCCAATCAGAAAGTCAACGCAGCTGCCTGCCTGGCTCGCGTACAATTCCTTTC
TTGCGCTATGACACTTCCAGCAAAAGGTAGGGCGGGCTGCGAGACGGCTTCCCG
GCGCTGCATGCAACACCGATGATGCTTCGACCCCCGAAGCTCTTCGGGGCTGC
ATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAGCCAGGCCCCC
GATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAACACCTAGATCA
CTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGCTAAGGGGGCG
CCTCTTCTCTTCGTTTCAGTCACAACCCGCAAACGGCGCGCCATGCTGCTGCAG
GCCTTCCTGTTCTCTGCTGGCCGGCTTCGCCGCCAAGATCAGCGCCTCCATGACGA
ACGAGACGTCCGACCGCCCCCTGGTGCATTCACCCCCAACAAAGGGCTGGATGA
ACGACCCCCAACGGCCTGTGGTACGACGAGAAGGACGCCAAGTGGCACCTGTACT
TCCAGTACAACCCGAACGACACCGTCTGGGGGACGCCCTTGTTCCTGGGGCCACGC
CACGTCCGACGACCTGACCAACTGGGAGGACCAGCCCATCGCCATCGCCCCGAA
GCGCAACGACTCCGGCGCCTTCTCCGGCTCCATGGTGGTGGACTACAACAACACC
TCCGGCTTCTTCAACGACACCATCGACCCGCGCCAGCGCTGCGTGGCCATCTGGA
CCTACAACACCCCGGAGTCCGAGGAGCAGTACATCTCCTACAGCCTGGACGGCG
GCTACACCTTCACCGAGTACCAGAAGAACCCCGTGCTGGCCGCCAACTCCACCCA
GTTCCGCGACCCGAAGGTCTTCTGGTACGAGCCCTCCCAGAAGTGGATCATGACC
GCGGCCAAGTCCCAGGACTACAAGATCGAGATCTACTCCTCCGACGACCTGAAG
TCCTGGAAGCTGGAGTCCGCGTTCGCCAACGAGGGCTTCCTCGGCTACCAGTACG
AGTGCCCCGGCCTGATCGAGGTCCCCACCGAGCAGGACCCAGCAAGTCCTACT
GGGTGATGTTTCATCTCCATCAACCCCGGCGCCCCGGCCGGCGGCTCCTTCAACCA
GTACTTCGTCGGCAGCTTCAACGGCACCCACTTCGAGGCCTTCGACAACCAGTCC
CGCGTGGTGGACTTCGGCAAGGACTACTACGCCCTGCAGACCTTCTTCAACACCG
ACCCGACCTACGGGAGCGCCCTGGGCATCGCGTGGGCTCCAAGTGGGAGTACT
CCGCCTTCGTGCCACCAACCCCTGGCGCTCCTCCATGTCCCTCGTGCGCAAGTTC
TCCCTCAACACCGAGTACCAGGCCAACCCGGAGACGGAGCTGATCAACCTGAAG
GCCGAGCCGATCCTGAACATCAGCAACGCCGGCCCCCTGGAGCCGGTTCGCCACC

AACACCACGTTGACGAAGGCCAACAGCTACAACGTCGACCTGTCCAACAGCACC
GGCACCTGGAGTTCGAGCTGGTGTACGCCGTCAACACCACCCAGACGATCTCCA
AGTCCGTGTTTCGCGGACCTCTCCCTCTGGTTCAAGGGCCTGGAGGACCCCGAGGA
GTACCTCCGCATGGGCTTCGAGGTGTCCGCGTCCCTCCTTCTTCCTGGACCGCGGG
AACAGCAAGGTGAAGTTCGTGAAGGAGAACCCCTACTTCACCAACCGCATGAGC
GTGAACAACCAGCCCTTCAAGAGCGAGAACGACCTGTCCTACTACAAGGTGTAC
GGCTTGCITGGACCAGAACATCCTGGAGCTGTACTTCAACGACGGCGACGTCGTGT
CCACCAACACCTACTTCATGACCACCGGGAACGCCCTGGGCTCCGTGAACATGAC
GACGGGGGTGGACAACCTGTTCTACATCGACAAGTTCAGGTGCGCGAGGTCAA
GTGATTAATTAACCTCGAGGCAGCAGCAGCTCGGATAGTATCGACACACTCTGGA
CGCTGGTCGTGTGATGGACTGTTGCCGCCACACTTGCTGCCTTGACCTGTGAATA
TCCCTGCCGCTTTTATCAAACAGCCTCAGTGTGTTTGATCTTGTGTGTACGCGCTT
TTGCGAGTTGCTAGCTGCTTGTGCTATTTGCGAATACCACCCCCAGCATCCCCTTC
CCTCGTTTCATATCGCTTGCATCCCAACCGCAACTTATCTACGCTGTCCTGCTATC
CCTCAGCGCTGCTCCTGCTCCTGCTCACTGCCCTCGCACAGCCTTGGTTTGGGCT
CCGCTGTATTCTCCTGGTACTGCAACCTGTAAACCAGCACTGCAATGCTGATGC
ACGGGAAGTAGTGGGATGGGAACACAAATGGAAAGCTTGAGCTCGGTACCCGTA
CCCATCAGCATCCGGGTGAATCTTGGCCTCCAAGATATGGCCAATCCTCACATCC
AGCTTGGCAAAATCGACTAGACTGTCTGCAAGTGGGAATGTGGAGCACAAGGTT
GCTTGTAGCGATCGACAGACTGGTGGGGTACATTGACAGGTGGGCAGCGCCGCA
TCCATCGTGCCTGACGCGAGCGCCGCCGGTTGCTCGCCCGTGCCTGCCGTCAAAG
AGCGGCAGAGAAATCGGGAACCGAAAACGTCACATTGCCTGATGTTGTTACATG
CTGGACTAGACTTTCTTGGCGTGGGTCTGCTCCTCGCCAGGTGCGCGACGCCTCG
GGGCTGGGTGCGAGGGAGCCGTGCGGCCACGCATTTGACAAGACCCAAAGCTCG
CATCTCAGACGGTCAACCGTTCGTATTATACATTCAACATATGGTACATACGCAA
AAAGCATG

SEQ ID NO: 127

P.moriformis isopentenyl diphosphate synthase plastid targeting sequence

MTFGVALPAMGRGVSLPRPRVAVRAQSASQVLESGRAQL

SEQ ID NO: 128

P.moriformis delta 12 fatty acid desaturase plastid targeting sequence

MAIKTNRQPVEKPPFTIGTLRKAIPAHCFERSALRGRAQL

SEQ ID NO: 129

P.moriformis stearoyl ACP desaturase plastid targeting sequence

MASAAFTMSACPAMTGRAPGARRSGRPVATRLRGRA

SEQ ID NO: 130

C.protothecoides stearoyl ACP desaturase plastid targeting sequence

MATASTFSAFNARCGDLRRSAGSGPRRPARPLPVRGRAQL

SEQ ID NO: 131

Cuphea hookeriana fatty acyl-ACP thioesterase (C8-10) plastid targeting sequence

MVAAAASSAFFPVPAPGASPKPGKFGNWPSSLSPSFKPKSIPNGGFQVKANDSAHPK
ANGSAVSLKSGSLNTQEDTSSSPPPRTFLH

SEQ ID NO: 132

Umbellularia californica fatty acyl-ACP thioesterase (C12) plastid targeting sequence

MATTSLASAFCSMKAVMLARDGRGMKPRSSDLQLRAGNAPTSCLKMINGTKFSYTES
LKRL

SEQ ID NO: 133

Cinnamomum camphora fatty acyl-ACP thioesterase (C14) plastid targeting sequence

MATTSLASAFCSMKAVMLARDGRGMKPRSSDLQLRAGNAQTSCLKMINGTKFSYTES
LKKL

SEQ ID NO: 134

P.moriformis fatty acyl-ACP thioesterase-1 cDNA sequence

ATGGCACCGACCGAGCCTGCTTGCCAGTACTGGCGTCTCTTCCGCTTCTCTGTGGTC
CTCTGCGCGCTCCAGCGCGTGCGCTTTTCCGGTGGATCATGCGGTCCGTGGCGCA
CCGCAGCGGCCGCTGCCCATGCAGCGCCGCTGCTTCCGAACAGTGGCGGTGAGG
GCCGCACCCGCGGTAGCCGTCCGTCCGGAACCCGCCCAAGAGTTTGGGAGCAG
CTTGAGCCCTGCAAGATGGCGGAGGACAAGCGCATCTTCCTGGAGGAGCACCGC
ATTTCGGGGCAACGAGGTGGGCCCCCTCGCAGCGGCTGACGATCACGGCGGTGGCC
AACATCCTGCAGGAGGCGGCGGGCAACCACGCGGTGGCCATGTGGGGGCCGGAGC
TCGGAGGGTTTCGCGACGGACCCGGAGCTGCAGGAGGCGGGTCTCATCTTTGTG
ATGACGCGCATGCAGATCCAAATGTACCGCTACCCGCGCTGGGGCGACCTGATG
CAGGTGGAGACCTGGTTCCAGACGGCGGGCAAGCTAGGCGCGCAGCGCGAGTGG
GTGCTGCGCGACAAGCTGACCGGCGAGGCGCTGGGCGCGGCCACCTCCAGCTGG
GTCATGATCAACATCCGCACGCGCCGCGCGTGCCGCATGCCCCAGCTCGTCCGCG
TCAAGTCGGCCTTCTTCGCGCGCGAGCCGCGCGCCTGGCGCTGCCGCCACCGT
CACGCGCGCAAGCTGCCCAACATCGCGACGCCGGCGCCGCTGCGCGGGCACCG
CCAGGTGCGCGCGCCGACCGACATGGACATGAACGGGCACGTGAACAACGTGGC
CTACCTGGCCTGGTGCCTGGAGGCCGTGCCCGAGCACGTCTTCAGCGACTACCAC
CTCTACCAGATGGAGATCGACTTCAAGGCCGAGTGCCACGCGGGCGACGTATC
TCCTCCCAGGCCGAGCAGATCCCGCCCCAGGAGGCGCTCACGCACAACGGCGCC
GGCCGCAACCCCTCCTGCTTCGTCCATAGCATTCTGCGCGCCGAGACCGAGCTCG
TCCGCGCGCGAACCACATGGTCGGCCCCCATCGACGCGCCCGCCGCCAAGCCGC
CCAAGGCGAGCCACTGA

SEQ ID NO: 135

P.moriformis fatty acyl-ACP thioesterase-2 cDNA sequence

ATGGCACCGACCGAGCCTGCTTGCCCGTACTGGCGTCTCTTCCGCTTCTCTGTGCTC
CTCTACGCGCTCCGGCGCGTGCGCTTTTCCGGTGGATCATGCGGTCCGTGGCGCA
CCGCAGCGGCCGCTGCCCATGCAGCGCCGCTGCTTCCGAACAGTGGCTGTGAGG
GCCGCACCCGCGAGTAGCCGTCCGTCCGGAACCCGCCCAAGAGTTTGGGAGCAG
CTTGAGCCCTGCAAGATGGCGGAGGACAAGCGCATCTTCCTGGAGGAGCACCGC
ATTTCGTGGCAACGAGGTGGGCCCCCTCGCAGCGGCTGACGATCACGGCGGTGGCC
AACATCCTGCAGGAGGCGGCGGGCAACCACGCGGTGGCCATGTGGGGTTCGGAGC

TCGGAGGGTTTCGCGACGGACCCGGAGCTGCAGGAGGCGGGCCTCATCTTTGTG
ATGACGCGCATGCAGATCCAAATGTACCGCTACCCGCGCTGGGGCGACCTGATG
CAGGTGGAGACCTGGTTCCAGACGGCGGGCAAGCTAGGCGCGCAGCGCGAGTGG
GTGCTGCGCGACAAGCTGACCGGCGAGGCGCTGGGCGCGGGCCACCTCCAGCTGG
GTCATGATCAACATCCGCACGCGCCGGCCGTGCCGCATGCCCCGAGCTCGTCCGCG
TCAAGTCGGCCTTCTTCGCGCGCGAGCCGCCGCGCCTGGCGCTGCCGCCCGCGGT
CACGCGTGCCAAGCTGCCCAACATCGCGACGCCGGCGCCGCTGCGCGGGCACCG
CCAGGTGCGCGCGCCGCACCGACATGGACATGAACGGCCACGTGAACAACGTTGC
CTACCTGGCCTGGTGCCTGGAGGCCGTGCCCGAGCACGTCTTCAGCGACTACCAC
CTCTACCAGATGGAGATCGACTTCAAGGCCGAGTGCCACGCGGGCGACGTCATC
TCCTCCCAGGCCGAGCAGATCCCGCCCCAGGAGGCGCTCACGCACAACGGCGCC
GGCCGCAACCCCTCCTGCTTCGTCCATAGCATTCTGCGCGCCGAGACCGAGCTCG
TCCGCGCGCGAACCACATGGTCGGCCCCCATCGACGCGCCCGCCGCAAGCCGC
CCAAGGCGAGCCACTGA

SEQ ID NO: 136

P.moriformis fatty acyl-ACP thioesterase-1 amino acid sequence

MAPTSLLASTGVSSASLWSSARSSACAFVVDHAVRGAPQRPLPMQRRCFRTVAVRA
APAVAVRPEPAQEFWEQLEPCKMAEDKRIFLEEHRIRGNEVGPSQRLTITAVANILQE
AAGNHAVAMWGRSSEGFATDPELQEAGLIFVMTRMQIQMYRYPRWGDLMQVETW
FQTAGKLGQAQREWVLRLDKLTGEALGAATSSWVMINIRTRRPCRMPPELVRVKSAPFA
REPRLALPPTVTRAKLPNIATPAPLRGHRQVARRTDMDMNGHVNNVAYLAWCLE
AVPEHVFSDYHLYQMEIDFKAECCHAGDVISSQAEQIPPQEALTHNGAGRNPSCFVHSI
LRAETELVRARTTWSAPIDAPAAKPPKASH

SEQ ID NO: 137

P.moriformis fatty acyl-ACP thioesterase-2 amino acid sequence

MAPTSLLARTGVSSASLCSSTRSGACAFVVDHAVRGAPQRPLPMQRRCFRTVAVRA
APAVAVRPEPAQEFWEQLEPCKMAEDKRIFLEEHRIRGNEVGPSQRLTITAVANILQE
AAGNHAVAMWGRSSEGFATDPELQEAGLIFVMTRMQIQMYRYPRWGDLMQVETW
FQTAGKLGQAQREWVLRLDKLTGEALGAATSSWVMINIRTRRPCRMPPELVRVKSAPFA
REPRLALPPAVTRAKLPNIATPAPIRGHRQVARRTDMDMNGHVNNVAYLAWCLE
AVPEHVFSDYHLYQMEIDFKAECCHAGDVISSQAEQIPPQEALTHNGAGRNPSCFVHSI
LRAETELVRARTTWSAPIDAPAAKPPKASH

SEQ ID NO: 138

C.hookeriana fatty acyl-ACP thioesterase amino acid sequence without PTS

QLPDWSRLITAITTVFVKS KRPDMDHDRKSKRPDMLVDSFGLESTVQDGLVFRQSFSI
RSYEIGTDRTASIEITLMNHLQETSLNHCKSTGILLDGFGRRTLEMCKRDLIWVVIKMQI
KVNRYPAWGDVTEINTRFSRLGKIGMGRDWLISDCNTGEILVRATSAYAMMNQKTR
RLSKLPYEVHQEIVPLFVDSPVIEDSDLKVHKFKVKTGDSIQKGLTPGWNLDLVNQH
VSNVKYIGWILESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVRQS
YQHLLRLEDGTAIVNGATEWRPKNAGANGAISTGKTSNGNSVS

SEQ ID NO: 139

U.californica fatty acyl-ACP thioesterase amino acid sequence without PTS

PDWSMLFAVITTIFSAAEKQWTNLEWKPKPKLPQLDDHFGLHGLVFRRTFAIRSYE
VGPDRSTSILAVMNIHQEATLNHAKSVGILGDGFGTTLEMSKRDLMWVVRTHVA
VERYPTWGDVEVECWIGASGNNGMRDRLVRDCKTGEILTRCTSLVLMNTRTRR
LSTIPDEVGEIGPAFIDNVAVKDDEIKKLQKLNDSTADYIQGGLTPRWNDLDVNQH
VNNLKYVAWVFETVPDSIFESHHSFTLEYRRECTRDSVLRSLTTVSGGSSEAGLVC
DHLLQLEGGSEVLRARTEWRPKLTDSFRGISVIPAEPRV

SEQ ID NO: 140

Cinnamomum camphora fatty acyl-ACP thioesterase amino acid sequence without PTS

PDWSMLFAVITTIFSAAEKQWTNLEWKPKPNPPQLDDHF^GPHGLVFRRTFAIRSYE
VGPDRSTSIVAVMNIHLQEALNHAKSVGILGDGFGTTLEMSKRDLIWVVKRTHVAV
ERYPAWGDVEVECWVGASGNNGRRHDFLVRDCKTGEILTRCTSLVMMNTRTRR
LSKIP^EEVGEIGPAFIDNVAVKDDEIKKPQKLNDSTADYIQGGLTPRWNDLDINQHV
NNIKYVDWILETVPDSIFESHHSFTIEYRRECTMDSVLQSLTTVSGGSSEAGLVCEH
LLQLEGGSEVLR^AKTEWRPKLTDSFRGISVIPAE^{SS}VMDYKDHDGDYKDHDIDYKD
DDDK

026172-005020PC - Sequence Listing_1
SEQUENCE LISTING

<110> SOLAZYME, INC.

<120> MANUFACTURING OF TAILORED OILS IN RECOMBINANT HETEROTROPHIC MICROORGANISMS

<130> 026172-005020PC

<140> PCT/US2009/066141

<141> 2009-11-30

<150> 61/219,525

<151> 2009-06-23

<150> 61/174,357

<151> 2009-04-30

<150> 61/118,994

<151> 2008-12-01

<150> 61/118,590

<151> 2008-11-28

<160> 143

<170> PatentIn version 3.5

<210> 1

<211> 1187

<212> DNA

<213> Chlorella sp.

<400> 1

gatcagacgg gcctgacctg cgagataatc aagtgtctgt aggcaaccaa ctcagcagct	60
gcttggtgtt ggggtctgcag gatagtgttg cagggcccca aggacagcag ggggaacttac	120
accttgtccc cgacctcagtt ttatggagtg cattgcctca agagcctagc cggagcgccta	180
ggctacatac ttgccgcacc ggtatgaggg gatatagtac tcgcactgctg ctgtctagtgt	240
agatgggagcag tgctgccccat aaacaactgg ctgctcagcc atttgttggc ggaccattct	300
ggggggggcca gcaatgcctg acttttcgggt aggggtgaaaa ctgaacaaag actacaaaaa	360
cagaatttct tcctccttgg aggttaagcgc agggccggccc gcctgcgccc acatggcgct	420
ccgaacacct ccatagctgt aaggggcgcaa acatggccgg actgttgtca gcactctttc	480
atggccatac aagggtcatgt cgagattagt gctgagtaag aactatcac cccatgttcg	540
attgaagccg tgacttcattg ccaacctgcc cctgggcgta gcagacgtat gccatcatga	600
ccactagccg acatgcgctg tcttttgcca ccaaaacaac tggtagaccg ctcgaagtcg	660
tgccgcacac ctccgggagt gagtccggcg actcctcccc ggcggggccgc ggccctacct	720
gggtaggggtc gccatacgcc cagcaccaaa cgacgcagga ggggattggg gtagggaatc	780
ccaaccagcc taaccaagac ggcacctata ataataggtg gggggactaa cagccctata	840
tcgcaagctt tgggtgccta tcttgagaag cagagtttg agtggctgtg tacggctcgac	900
cctaaggtgg gtgtgccgca gcctgaaaca aagcgtctag cagctgcttc tataatgtgt	960
cagccgttgt gtttcagtta tattgtatgc tattgtttgt tcgtgctagg gtggcgagc	1020

026172-005020PC - Sequence Listing_1

cccacctact gtggcgggcc attggttggg gcttgaattg cctcaccatc taagggtctga	1080
acgctcactc aaacgccttt gtacaactgc agaactttcc ttggcgctgc aactacagtg	1140
tgcaaaccag cacatagcac tcccttacat caccagcag tacaaca	1187

<210> 2
 <211> 1414
 <212> DNA
 <213> *Chlorella ellipsoidea*

<400> 2	
cgctgcgcac cagggccgcc agctcgctga tgctgctcca aatgcggtcc cccgattttt	60
tgttcttcat cttctccacc ttggtggcct tcttggccag ggcccttcagc tgcatgcgca	120
cagaccgttg agctcctgat cagcatcctc aggaggccct ttgacaagca agccccctgtg	180
caagcccatt cacggggtac cagtgggtgct gaggtagatg ggtttgaaaa ggattgctcg	240
gtcgattgct gctcatggaa ttggcatgtg catgcatgtt cacaatatgc caccaggctt	300
tggagcaaga gagcatgaat gccttcaggc aggttgaaaag ttcctggggg tgaagaggca	360
gggccgagga ttggaggagg aaagcatcaa gtcgctgctc atgctcatgt tttcagtcag	420
agtttgccaa gctcacagga gcagagacaa gactggctgc tcagggtgtg catcgtgtgt	480
gtggtggggg ggggggggtt aatacggtag gaaatgcact tggaattccc acctcatgcc	540
agcggaccca catgcttgaa ttcgaggcct gtggggtgag aaatgctcac tctgccctcg	600
ttgctgaggt acttcaggcc gctgagctca aagtcgatgc cctgctcgtc tatcagggcc	660
tgcacctctg ggctgaccgg ctcagcctcc ttcgcgggca tggagtaggc gccggcagcg	720
ttcatgtccg ggcccagggc agcgggtggtg ccataaatgt cggatgatggg ggggaggggg	780
gccgtcgcca caccattgcc gttgctggct gacgcatgca catgtggcct ggctggcacc	840
ggcagcactg gtctccagcc agccagcaag tggctgttca ggaaagcggc catgttggtg	900
gtccctgcgc atgtaattcc ccagatcaaa ggagggaaca gcttggattt gatgtagtgc	960
ccaaccggac tgaatgtgcg atggcaggtc cttttgagtc tcccgaatta ctagcagggc	1020
actgtgacct aacgcagcat gccaaaccga aaaaaatgat tgacagaaaa tgaagcgggtg	1080
tgtcaatatt tgctgtatatt attcgtttta atcagcaacc aagttcgaaa cgcaactatc	1140
gtggtgatca agtgaacctc atcagactta cctcgttcgg caaggaaacg gaggcaccaa	1200
attccaattt gatattatcg cttgccaaag tagagctgat ctttgggaaa ccaactgcca	1260
gacagtggac tgtgatggag tgccccgagt ggtggagcct cttcgattcg gttagtcatt	1320
actaacgtga accctcagtg aagggaacct cagaccagaa agaccagatc tcctcctcga	1380
caccgagaga gtgttgccgc agtaggacga caag	1414

<210> 3
 <211> 512
 <212> PRT
 <213> Unknown

<220>

<223> Description of Unknown: Yeast sequence

<400> 3

Met Thr Asn Glu Thr Ser Asp Arg Pro Leu Val His Phe Thr Pro Asn
 1 5 10 15

Lys Gly Trp Met Asn Asp Pro Asn Gly Leu Trp Tyr Asp Glu Lys Asp
 20 25 30

Ala Lys Trp His Leu Tyr Phe Gln Tyr Asn Pro Asn Asp Thr Val Trp
 35 40 45

Gly Thr Pro Leu Phe Trp Gly His Ala Thr Ser Asp Asp Leu Thr Asn
 50 55 60

Trp Glu Asp Gln Pro Ile Ala Ile Ala Pro Lys Arg Asn Asp Ser Gly
 65 70 75 80

Ala Phe Ser Gly Ser Met Val Val Asp Tyr Asn Asn Thr Ser Gly Phe
 85 90 95

Phe Asn Asp Thr Ile Asp Pro Arg Gln Arg Cys Val Ala Ile Trp Thr
 100 105 110

Tyr Asn Thr Pro Glu Ser Glu Glu Gln Tyr Ile Ser Tyr Ser Leu Asp
 115 120 125

Gly Gly Tyr Thr Phe Thr Glu Tyr Gln Lys Asn Pro Val Leu Ala Ala
 130 135 140

Asn Ser Thr Gln Phe Arg Asp Pro Lys Val Phe Trp Tyr Glu Pro Ser
 145 150 155 160

Gln Lys Trp Ile Met Thr Ala Ala Lys Ser Gln Asp Tyr Lys Ile Glu
 165 170 175

Ile Tyr Ser Ser Asp Asp Leu Lys Ser Trp Lys Leu Glu Ser Ala Phe
 180 185 190

Ala Asn Glu Gly Phe Leu Gly Tyr Gln Tyr Glu Cys Pro Gly Leu Ile
 195 200 205

Glu Val Pro Thr Glu Gln Asp Pro Ser Lys Ser Tyr Trp Val Met Phe
 210 215 220

Ile Ser Ile Asn Pro Gly Ala Pro Ala Gly Gly Ser Phe Asn Gln Tyr
 225 230 235 240

Phe Val Gly Ser Phe Asn Gly Thr His Phe Glu Ala Phe Asp Asn Gln
 245 250 255

Ser Arg Val Val Asp Phe Gly Lys Asp Tyr Tyr Ala Leu Gln Thr Phe

026172-005020PC - Sequence Listing_1

260

265

270

Phe Asn Thr Asp Pro Thr Tyr Gly Ser Ala Leu Gly Ile Ala Trp Ala
 275 280 285

Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn Pro Trp Arg Ser
 290 295 300

Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr Glu Tyr Gln Ala
 305 310 315 320

Asn Pro Glu Thr Glu Leu Ile Asn Leu Lys Ala Glu Pro Ile Leu Asn
 325 330 335

Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe Ala Thr Asn Thr Thr Leu
 340 345 350

Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu Ser Asn Ser Thr Gly Thr
 355 360 365

Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr Gln Thr Ile Ser
 370 375 380

Lys Ser Val Phe Ala Asp Leu Ser Leu Trp Phe Lys Gly Leu Glu Asp
 385 390 395 400

Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu Val Ser Ala Ser Ser Phe
 405 410 415

Phe Leu Asp Arg Gly Asn Ser Lys Val Lys Phe Val Lys Glu Asn Pro
 420 425 430

Tyr Phe Thr Asn Arg Met Ser Val Asn Asn Gln Pro Phe Lys Ser Glu
 435 440 445

Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu Asp Gln Asn Ile
 450 455 460

Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val Val Ser Thr Asn Thr Tyr
 465 470 475 480

Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn Met Thr Thr Gly
 485 490 495

Val Asp Asn Leu Phe Tyr Ile Asp Lys Phe Gln Val Arg Glu Val Lys
 500 505 510

<210> 4

<211> 20

<212> PRT

<213> Unknown

026172-005020PC - Sequence Listing_1

<220>

<223> Description of Unknown: Yeast sequence

<400> 4

Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys
1 5 10 15

Ile Ser Ala Ser
20

<210> 5

<211> 18

<212> PRT

<213> Unknown

<220>

<223> Description of Unknown: Higher plant secretion signal

<400> 5

Met Ala Asn Lys Ser Leu Leu Leu Leu Leu Leu Gly Ser Leu Ala
1 5 10 15

Ser Gly

<210> 6

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic consensus sequence

<400> 6

Met Ala Arg Leu Pro Leu Ala Ala Leu Gly
1 5 10

<210> 7

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic peptide

<400> 7

Met Ala Asn Lys Leu Leu Leu Leu Leu Leu Leu Leu Leu Pro Leu
1 5 10 15

Ala Ala Ser Gly
20

<210> 8

<211> 2615

<212> DNA

<213> Saccharomyces cerevisiae

<400> 8

gaattcccca acatggtgga gcacgacact ctcgtctact ccaagaatat caaagataca

60

026172-005020PC - Sequence Listing_1

gtctcagaag accaaagggc tattgagact tttcaacaaa gggtaatatc gggaaacctc	120
ctcggattcc attgcccagc tatctgtcac ttcataaaaa ggacagtaga aaaggaaggt	180
ggcacctaca aatgccatca ttgcgataaa ggaaaggcta tcgttcaaga tgcctctgcc	240
gacagtggtc ccaaagatgg acccccaccc acgaggagca tcgtggaaaa agaagacgtt	300
ccaaccacgt cttcaaagca agtggattga tgtgaacatg gtggagcacg acactctcgt	360
ctactccaag aatatcaaag atacagtctc agaagaccaa agggctattg agacttttca	420
acaaagggta atatcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat	480
caaaaggaca gtagaaaagg aagggtggcac ctacaaatgc catcattgcg ataaaggaaa	540
ggctatcgtt caagatgcct ctgccgacag tgggtcccaa gatggacccc caccacgag	600
gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga	660
tatctccact gacgtaaggg atgacgcaca atcccactat ctttcgcaag acccttcctc	720
tatataagga agttcatttc atttggagag gacacgctga aatcaccagt ctctctctac	780
aaatctatct ctggcgcgcc atatcaatgc ttcttcaggc ctttcttttt cttcttgctg	840
gttttgctgc caagatcagc gcctctatga cgaacgaaac ctcggataga ccacttgctc	900
actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaaag	960
atgccaaagt gcatctgtac tttcaataca acccgaacga tactgtcttg gggacgccat	1020
tgttttgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta	1080
tcgctccgaa gaggaacgac tccggagcat tctcgggttc catggtgggt gactacaaca	1140
atacttccgg ctttttcaac gataccattg acccgagaca acgctgcgtg gccatatgga	1200
cttacaacac accggagtcg gaggagcagt acatctcgta tagcctggac ggtggataca	1260
cttttacaga gtatcagaag aaccctgtgc ttgctgcaaa ttcgactcag ttccgagatc	1320
cgaaggctct ttggtacgag ccctcgcaag agtggatcat gacagcggca aagtcacagg	1380
actacaagat cgaaatttac tcgtctgacg accttaaadc ctggaagctc gaatccgcgt	1440
tcgcaaacga gggctttctc ggctaccaat acgaatgccc aggcctgata gaggtcccaa	1500
cagagcaaga tcccagcaag tcctactggg tgatgtttat ttccattaat ccaggagcac	1560
cggcaggagg ttcttttaat cagtacttcg tcggaagctt taacggaact catttcgagg	1620
catttgataa ccaatcaaga gtagttgatt ttggaaagga ctactatgcc ctgcagactt	1680
tcttcaatac tgacccgacc tatgggagcg ctcttgccat tgcgtgggct tctaactggg	1740
agtattccgc attcgttcct acaaaccctt ggagggtcctc catgtcgtc gtgaggaaat	1800
tctctctcaa cactgagtac caggccaacc cggaaccga actcataaac ctgaaagccg	1860
aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgcaacc aacaccacgt	1920
tgacgaaagc caacagctac aacgtcgatc ttctgaatag caccggtaca ctggaatttg	1980
aactggtgta tgccgtcaat accacccaaa cgatctcgaa gtcggtgttc gcggacctct	2040
ccctctgggt taaaggcctg gaagaccccg aggagtagct cagaatgggt ttcgaggttt	2100

026172-005020PC - Sequence Listing_1

ctgcgtcctc cttcttcctt gatcgcggga acagcaaagt aaaatttggt aaggagaacc	2160
catattttac caacaggatg agcgttaaca accaaccatt caagagcgaa aacgacctgt	2220
cgtactacaa agtgtatggt ttgcttgatc aaaatatcct ggaactctac ttcaacgatg	2280
gtgatgtcgt gtccaccaac acatacttca tgacaaccgg gaacgcactg ggctccgtga	2340
acatgacgac ggggtgtggat aacctgttct acatcgacaa attccagggtg aggggaagtca	2400
agtgagatct gtcgatcgac aagctcgagt ttctccataa taatgtgtga gtagttccca	2460
gataagggaa ttagggttcc tatagggttt cgctcatgtg ttgagcatat aagaaaccct	2520
tagtatgtat ttgtatttgt aaaatacttc tatcaataaa atttctaatt cctaaaacca	2580
aaatccagta ctaaaatcca gatccccga attaa	2615

<210> 9
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 9	
tgttgaagaa tgagccggcg ac	22

<210> 10
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 10	
cagtgaagctt ttacgcactc	20

<210> 11
 <211> 541
 <212> DNA
 <213> Prototheca kruegani

<400> 11	
tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa tactctggag	60
ccatagcgaa agcaagttta gtaagcttag gtcattcttt ttagaccgga aaccgagtga	120
tctacccatg atcagggtga agtgtttagta aaataacatg gaggcccgaa ccgactaatg	180
ttgaaaaatt agcggatgaa ttgtgggttag gggcgaaaaa ccaatcgaac tcggagttag	240
ctggttctcc ccgaaatgag tttaggcgca gcagtagcag tacaaataga ggggtaaagc	300
actgtttctt ttgtgggctt cgaaagtgtt acctcaaagt ggcaaactct gaatactcta	360
tttagatatc tactagttag accttggggg ataagctcct tgggtcaaaag ggaaacagcc	420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggatatgtca	480

026172-005020PC - Sequence Listing_1

aaacctccag caggtagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541

<210> 12
<211> 573
<212> DNA
<213> *Prototheca wickerhamii*

<400> 12
tggtgaagaa tgagccggcg acttaaaata aatggcaggc taagagattt aataactcga 60
aacctaagcg aaagcaagtc ttaatagggc gtcaatttaa caaaacttta aataaattat 120
aaagtcattt atttttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc 300
gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat 360
ggtaccaa at cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420
gggataagct ccatagtcga gagggaaaaca gcccagacca ccagttaagg ccccaaatg 480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggagggttg cttagaagca 540
gccatccttt aaagagtgcg taatagctca ctg 573

<210> 13
<211> 541
<212> DNA
<213> *Prototheca stagnora*

<400> 13
tggtgaagaa tgagccggcg agttaaaaaa aatggcatgg ttaaagatat ttctctgaag 60
ccatagcgaa agcaagtttt acaagctata gtcatttttt ttagacccga aaccgagtga 120
tctacccatg atcaggggta agtgttggc aaataacatg gagggccgaa ccgactaatg 180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
actgtttctt ttgtgggctt cgaaagtgtt acctcaaagt ggcaactct gaatactcta 360
tttagatatt tactagttag accttggggg ataagctcct tggcaaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatatgta ctaaggacgt gagtatgtca 480
aaacctccag caggtagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541

<210> 14
<211> 541
<212> DNA
<213> *Prototheca moriformis*

<400> 14
tggtgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagataa ttctctggag 60
ccatagcgaa agcaagttta acaagctaaa gtcacccttt ttagacccga aaccgagtga 120

026172-005020PC - Sequence Listing_1

tctacccatg atcaggggtga agtgtttggtg aaataacatg gagggcccgaa ccgactaatg	180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag	240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc	300
actgtttctt ttgtgggctt cgaaagtgtg acctcaaagt ggcaaactct gaatactcta	360
tttagatatc tactagttag accttggggg ataagctcct tgggtcaaaag ggaaacagcc	420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggatatgtta	480
aaacctccag cagggttagct tagaagcagc aatcctttca agagtgcgta atagctcact	540
g	541

<210> 15
 <211> 573
 <212> DNA
 <213> Prototheca moriformis

<400> 15	
tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga	60
aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct	120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg	180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt	240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc	300
gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat	360
ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg	420
gggataagct ccatagtcga gagggaaaca gccagacca ccagttaagg ccccaaatg	480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggagggttg cttagaagca	540
gccatccttt aaagagtgcg taatagctca ctg	573

<210> 16
 <211> 573
 <212> DNA
 <213> Prototheca wickerhamii

<400> 16	
tgttgaagaa tgagccgtcg acttaaaata aatggcaggc taagagaatt aataactcga	60
aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct	120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg	180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt	240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc	300
gcagcaatat atctcgtcta tctaggggta aagcactggt tcggtgcggg ctatgaaaat	360
ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg	420
gggataagct ccatagtcga gagggaaaca gccagacca ccagttaagg ccccaaatg	480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggagggttg cttagaagca	540

026172-005020PC - Sequence Listing_1

gccatccttt aaagagtgcg taatagctca ctg 573

<210> 17
 <211> 541
 <212> DNA
 <213> Prototheca moriformis

<400> 17
 tgttgaagaa tgagccggcg agttaaaaag agtggcgtag ttaaagaaaa ttctctggaa 60
 ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagacccga aaccgagtga 120
 tctacccatg atcagggtag agtggttgta aaataacatg gagggccgaa ccgactaatg 180
 gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaa tcggagttag 240
 ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
 actgtttctt ttgtgggctc cgaaagtgtg acctcaaagt ggcaaactct gaatactcta 360
 tttagatata tactagttag accttggggg ataagctcct tggtcgaaag ggaaacagcc 420
 cagatcacca gttaaggccc caaaatgaaa atgatatgta ctaaggatgt gagtatgtca 480
 aaacctccag caggtttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
 g 541

<210> 18
 <211> 541
 <212> DNA
 <213> Prototheca zopfii

<400> 18
 tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa ttctctggag 60
 ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagacccga aaccgagtga 120
 tctacccatg atcagggtag agtggttgta aaataacatg gagggccgaa ccgactaatg 180
 gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaa tcggagttag 240
 ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
 actgtttctt tcgtgggctt cgaaagtgtg acctcaaagt ggcaaactct gaatactcta 360
 tttagatata tactagttag accttggggg ataagctcct tggtcgaaag ggaaacagcc 420
 cagatcacca gttaaggccc caaaatgaaa atgatatgta ctaaggatgt gagtatgtca 480
 aaacctccag caggtttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
 g 541

<210> 19
 <211> 565
 <212> DNA
 <213> Prototheca moriformis

<400> 19
 tgttgaagaa tgagccggcg acttagaaaa ggtggcatgg ttaaggaaat attccgaagc 60
 cgtagcaaaa gcgagtctga atagggcgat aaaatatatt aatatttaga atctagtcac 120

026172-005020PC - Sequence Listing_1

tttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaagctt gggtgatacc	180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcgggtg	240
aaataccagt cgaacccgga gctagctggt tctccccgaa atgcgttgag gcgcagcagt	300
acatctagtc tatctagggg taaagcactg tttcgggtgcg ggctgtgaga acggtaccaa	360
atcgtggcaa actctgaata ctagaaatga cgatgtagta gtgagactgt gggggataag	420
ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta	480
gtgacaaagg aggtgaaaat gcaaatacaa ccaggagggtt ggcttagaag cagccatcct	540
ttaaagagtg cgtaatagct cactg	565

<210> 20
 <211> 550
 <212> PRT
 <213> Cichorium intybus

<400> 20
 Met Ser Asn Ser Ser Asn Ala Ser Glu Ser Leu Phe Pro Ala Thr Ser
 1 5 10 15

Glu Gln Pro Tyr Arg Thr Ala Phe His Phe Gln Pro Pro Gln Asn Trp
 20 25 30

Met Asn Asp Pro Asn Gly Pro Met Cys Tyr Asn Gly Val Tyr His Leu
 35 40 45

Phe Tyr Gln Tyr Asn Pro Phe Gly Pro Leu Trp Asn Leu Arg Met Tyr
 50 55 60

Trp Ala His Ser Val Ser His Asp Leu Ile Asn Trp Ile His Leu Asp
 65 70 75 80

Leu Ala Phe Ala Pro Thr Glu Pro Phe Asp Ile Asn Gly Cys Leu Ser
 85 90 95

Gly Ser Ala Thr Val Leu Pro Gly Asn Lys Pro Ile Met Leu Tyr Thr
 100 105 110

Gly Ile Asp Thr Glu Asn Arg Gln Val Gln Asn Leu Ala Val Pro Lys
 115 120 125

Asp Leu Ser Asp Pro Tyr Leu Arg Glu Trp Val Lys His Thr Gly Asn
 130 135 140

Pro Ile Ile Ser Leu Pro Glu Glu Ile Gln Pro Asp Asp Phe Arg Asp
 145 150 155 160

Pro Thr Thr Thr Trp Leu Glu Glu Asp Gly Thr Trp Arg Leu Leu Val
 165 170 175

Gly Ser Gln Lys Asp Lys Thr Gly Ile Ala Phe Leu Tyr His Ser Gly

026172-005020PC - Sequence Listing_1

180

185

190

Asp Phe Val Asn Trp Thr Lys Ser Asp Ser Pro Leu His Lys Val Ser
 195 200 205

Gly Thr Gly Met Trp Glu Cys Val Asp Phe Phe Pro Val Trp Val Asp
 210 215 220

Ser Thr Asn Gly Val Asp Thr Ser Ile Ile Asn Pro Ser Asn Arg Val
 225 230 235 240

Lys His Val Leu Lys Leu Gly Ile Gln Asp His Gly Lys Asp Cys Tyr
 245 250 255

Leu Ile Gly Lys Tyr Ser Ala Asp Lys Glu Asn Tyr Val Pro Glu Asp
 260 265 270

Glu Leu Thr Leu Ser Thr Leu Arg Leu Asp Tyr Gly Met Tyr Tyr Ala
 275 280 285

Ser Lys Ser Phe Phe Asp Pro Val Lys Asn Arg Arg Ile Met Thr Ala
 290 295 300

Trp Val Asn Glu Ser Asp Ser Glu Ala Asp Val Ile Ala Arg Gly Trp
 305 310 315 320

Ser Gly Val Gln Ser Phe Pro Arg Ser Leu Trp Leu Asp Lys Asn Gln
 325 330 335

Lys Gln Leu Leu Gln Trp Pro Ile Glu Glu Ile Glu Met Leu His Gln
 340 345 350

Asn Glu Val Ser Phe His Asn Lys Lys Leu Asp Gly Gly Ser Ser Leu
 355 360 365

Glu Val Leu Gly Ile Thr Ala Ser Gln Ala Asp Val Lys Ile Ser Phe
 370 375 380

Lys Leu Ala Asn Leu Glu Glu Ala Glu Glu Leu Asp Pro Ser Trp Val
 385 390 395 400

Asp Pro Gln Leu Ile Cys Ser Glu Asn Asp Ala Ser Lys Lys Gly Lys
 405 410 415

Phe Gly Pro Phe Gly Leu Leu Ala Leu Ala Ser Ser Asp Leu Arg Glu
 420 425 430

Gln Thr Ala Ile Phe Phe Arg Val Phe Arg Lys Asn Gly Arg Tyr Val
 435 440 445

Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser Met Lys Asn Gly Ile

026172-005020PC - Sequence Listing_1

450

455

460

Glu Lys Arg Thr Tyr Gly Ala Phe Val Asp Ile Asp Pro Gln Gln Asp
465 470 475 480

Glu Ile Ser Leu Arg Thr Leu Ile Asp His Ser Ile Val Glu Ser Phe
485 490 495

Gly Gly Arg Gly Lys Thr Cys Ile Thr Thr Arg Val Tyr Pro Thr Leu
500 505 510

Ala Ile Gly Glu Gln Ala Arg Leu Phe Ala Phe Asn His Gly Thr Glu
515 520 525

Ser Val Glu Ile Ser Glu Leu Ser Ala Trp Ser Met Lys Lys Ala Gln
530 535 540

Met Lys Val Glu Glu Pro
545 550

<210> 21

<211> 581

<212> PRT

<213> Schizosaccharomyces pombe

<400> 21

Met Phe Leu Lys Tyr Ile Leu Ala Ser Gly Ile Cys Leu Val Ser Leu
1 5 10 15

Leu Ser Ser Thr Asn Ala Ala Pro Arg His Leu Tyr Val Lys Arg Tyr
20 25 30

Pro Val Ile Tyr Asn Ala Ser Asn Ile Thr Glu Val Ser Asn Ser Thr
35 40 45

Thr Val Pro Pro Pro Pro Phe Val Asn Thr Thr Ala Pro Asn Gly Thr
50 55 60

Cys Leu Gly Asn Tyr Asn Glu Tyr Leu Pro Ser Gly Tyr Tyr Asn Ala
65 70 75 80

Thr Asp Arg Pro Lys Ile His Phe Thr Pro Ser Ser Gly Phe Met Asn
85 90 95

Asp Pro Asn Gly Leu Val Tyr Thr Gly Gly Val Tyr His Met Phe Phe
100 105 110

Gln Tyr Ser Pro Lys Thr Leu Thr Ala Gly Glu Val His Trp Gly His
115 120 125

Thr Val Ser Lys Asp Leu Ile His Trp Glu Asn Tyr Pro Ile Ala Ile
130 135 140

026172-005020PC - Sequence Listing_1

Tyr Pro Asp Glu His Glu Asn Gly Val Leu Ser Leu Pro Phe Ser Gly
 145 150 155 160
 Ser Ala Val Val Asp Val His Asn Ser Ser Gly Leu Phe Ser Asn Asp
 165 170 175
 Thr Ile Pro Glu Glu Arg Ile Val Leu Ile Tyr Thr Asp His Trp Thr
 180 185 190
 Gly Val Ala Glu Arg Gln Ala Ile Ala Tyr Thr Thr Asp Gly Gly Tyr
 195 200 205
 Thr Phe Lys Lys Tyr Ser Gly Asn Pro Val Leu Asp Ile Asn Ser Leu
 210 215 220
 Gln Phe Arg Asp Pro Lys Val Ile Trp Asp Phe Asp Ala Asn Arg Trp
 225 230 235 240
 Val Met Ile Val Ala Met Ser Gln Asn Tyr Gly Ile Ala Phe Tyr Ser
 245 250 255
 Ser Tyr Asp Leu Ile His Trp Thr Glu Leu Ser Val Phe Ser Thr Ser
 260 265 270
 Gly Tyr Leu Gly Leu Gln Tyr Glu Cys Pro Gly Met Ala Arg Val Pro
 275 280 285
 Val Glu Gly Thr Asp Glu Tyr Lys Trp Val Leu Phe Ile Ser Ile Asn
 290 295 300
 Pro Gly Ala Pro Leu Gly Gly Ser Val Val Gln Tyr Phe Val Gly Asp
 305 310 315 320
 Trp Asn Gly Thr Asn Phe Val Pro Asp Asp Gly Gln Thr Arg Phe Val
 325 330 335
 Asp Leu Gly Lys Asp Phe Tyr Ala Ser Ala Leu Tyr His Ser Ser Ser
 340 345 350
 Ala Asn Ala Asp Val Ile Gly Val Gly Trp Ala Ser Asn Trp Gln Tyr
 355 360 365
 Thr Asn Gln Ala Pro Thr Gln Val Phe Arg Ser Ala Met Thr Val Ala
 370 375 380
 Arg Lys Phe Thr Leu Arg Asp Val Pro Gln Asn Pro Met Thr Asn Leu
 385 390 395 400
 Thr Ser Leu Ile Gln Thr Pro Leu Asn Val Ser Leu Leu Arg Asp Glu
 405 410 415

026172-005020PC - Sequence Listing_1

Thr Leu Phe Thr Ala Pro Val Ile Asn Ser Ser Ser Ser Leu Ser Gly
420 425 430

Ser Pro Ile Thr Leu Pro Ser Asn Thr Ala Phe Glu Phe Asn Val Thr
435 440 445

Leu Ser Ile Asn Tyr Thr Glu Gly Cys Thr Thr Gly Tyr Cys Leu Gly
450 455 460

Arg Ile Ile Ile Asp Ser Asp Asp Pro Tyr Arg Leu Gln Ser Ile Ser
465 470 475 480

Val Asp Val Asp Phe Ala Ala Ser Thr Leu Val Ile Asn Arg Ala Lys
485 490 495

Ala Gln Met Gly Trp Phe Asn Ser Leu Phe Thr Pro Ser Phe Ala Asn
500 505 510

Asp Ile Tyr Ile Tyr Gly Asn Val Thr Leu Tyr Gly Ile Val Asp Asn
515 520 525

Gly Leu Leu Glu Leu Tyr Val Asn Asn Gly Glu Lys Thr Tyr Thr Asn
530 535 540

Asp Phe Phe Phe Leu Gln Gly Ala Thr Pro Gly Gln Ile Ser Phe Ala
545 550 555 560

Ala Phe Gln Gly Val Ser Phe Asn Asn Val Thr Val Thr Pro Leu Lys
565 570 575

Thr Ile Trp Asn Cys
580

<210> 22
<211> 550
<212> PRT
<213> Pichia anomala

<400> 22
Met Ile Gln Leu Ser Pro Leu Leu Leu Leu Pro Leu Phe Ser Val Phe
1 5 10 15

Asn Ser Ile Ala Asp Ala Ser Thr Glu Tyr Leu Arg Pro Gln Ile His
20 25 30

Leu Thr Pro Asp Gln Gly Trp Met Asn Asp Pro Asn Gly Met Phe Tyr
35 40 45

Asp Arg Lys Asp Lys Leu Trp His Val Tyr Phe Gln His Asn Pro Asp
50 55 60

Lys Lys Ser Ile Trp Ala Thr Pro Val Thr Trp Gly His Ser Thr Ser

65					70					75					80
Lys	Asp	Leu	Leu	Thr 85	Trp	Asp	Tyr	His	Gly 90	Asn	Ala	Leu	Glu	Pro 95	Glu
Asn	Asp	Asp	Glu 100	Gly	Ile	Phe	Ser	Gly 105	Ser	Val	Val	Val	Asp 110	Arg	Asn
Asn	Thr	Ser 115	Gly	Phe	Phe	Asn	Asp 120	Ser	Thr	Asp	Pro	Glu 125	Gln	Arg	Ile
Val	Ala 130	Ile	Tyr	Thr	Asn	Asn 135	Ala	Gln	Leu	Gln	Thr 140	Gln	Glu	Ile	Ala
Tyr 145	Ser	Leu	Asp	Lys	Gly 150	Tyr	Ser	Phe	Ile	Lys 155	Tyr	Asp	Gln	Asn	Pro 160
Val	Ile	Asn	Val	Asn 165	Ser	Ser	Gln	Gln	Arg 170	Asp	Pro	Lys	Val	Leu 175	Trp
His	Asp	Glu	Ser 180	Asn	Gln	Trp	Ile	Met 185	Val	Val	Ala	Lys	Thr 190	Gln	Glu
Phe	Lys	Val 195	Gln	Ile	Tyr	Gly	Ser 200	Pro	Asp	Leu	Lys	Lys 205	Trp	Asp	Leu
Lys	Ser 210	Asn	Phe	Thr	Ser	Asn 215	Gly	Tyr	Leu	Gly	Phe 220	Gln	Tyr	Glu	Cys
Pro 225	Gly	Leu	Phe	Lys	Leu 230	Pro	Ile	Glu	Asn	Pro 235	Leu	Asn	Asp	Thr	Val 240
Thr	Ser	Lys	Trp	Val 245	Leu	Leu	Leu	Ala	Ile 250	Asn	Pro	Gly	Ser	Pro 255	Leu
Gly	Gly	Ser	Ile 260	Asn	Glu	Tyr	Phe	Ile 265	Gly	Asp	Phe	Asp	Gly 270	Thr	Thr
Phe	His	Pro 275	Asp	Asp	Gly	Ala	Thr 280	Arg	Phe	Met	Asp	Ile 285	Gly	Lys	Asp
Phe	Tyr 290	Ala	Phe	Gln	Ser	Phe 295	Asp	Asn	Thr	Glu	Pro 300	Glu	Asp	Gly	Ala
Leu 305	Gly	Leu	Ala	Trp	Ala 310	Ser	Asn	Trp	Gln	Tyr 315	Ala	Asn	Thr	Val	Pro 320
Thr	Glu	Asn	Trp	Arg 325	Ser	Ser	Met	Ser	Leu 330	Val	Arg	Asn	Tyr	Thr 335	Leu
Lys	Tyr	Val	Asp	Val	Asn	Pro	Glu	Asn	Tyr	Gly	Leu	Thr	Leu	Ile	Gln

026172-005020PC - Sequence Listing_1

340

345

350

Lys Pro Val Tyr Asp Thr Lys Glu Thr Arg Leu Asn Glu Thr Leu Lys
 355 360 365

Thr Leu Glu Thr Ile Asn Glu Tyr Glu Val Asn Asp Leu Lys Leu Asp
 370 375 380

Lys Ser Ser Phe Val Ala Thr Asp Phe Asn Thr Glu Arg Asn Ala Thr
 385 390 395 400

Gly Val Phe Glu Phe Asp Leu Lys Phe Thr Gln Thr Asp Leu Lys Met
 405 410 415

Gly Tyr Ser Asn Met Thr Thr Gln Phe Gly Leu Tyr Ile His Ser Gln
 420 425 430

Thr Val Lys Gly Ser Gln Glu Thr Leu Gln Leu Val Phe Asp Thr Leu
 435 440 445

Ser Thr Thr Trp Tyr Ile Asp Arg Thr Thr Gln His Ser Phe Gln Arg
 450 455 460

Asn Ser Pro Val Phe Thr Glu Arg Ile Ser Thr Tyr Val Glu Lys Ile
 465 470 475 480

Asp Thr Thr Asp Gln Gly Asn Val Tyr Thr Leu Tyr Gly Val Val Asp
 485 490 495

Arg Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Ser Ile Ala Met Thr
 500 505 510

Asn Thr Phe Phe Phe Arg Glu Gly Lys Ile Pro Thr Ser Phe Glu Val
 515 520 525

Val Cys Asp Ser Glu Lys Ser Phe Ile Thr Ile Asp Glu Leu Ser Val
 530 535 540

Arg Glu Leu Ala Arg Lys
 545 550

<210> 23

<211> 533

<212> PRT

<213> Debaryomyces occidentalis

<400> 23

Met Val Gln Val Leu Ser Val Leu Val Ile Pro Leu Leu Thr Leu Phe
 1 5 10 15

Phe Gly Tyr Val Ala Ser Ser Ser Ile Asp Leu Ser Val Asp Thr Ser
 20 25 30

026172-005020PC - Sequence Listing_1

Glu Tyr Asn Arg Pro Leu Ile His Phe Thr Pro Glu Lys Gly Trp Met
35 40 45

Asn Asp Pro Asn Gly Leu Phe Tyr Asp Lys Thr Ala Lys Leu Trp His
50 55 60

Leu Tyr Phe Gln Tyr Asn Pro Asn Ala Thr Ala Trp Gly Gln Pro Leu
65 70 75 80

Tyr Trp Gly His Ala Thr Ser Asn Asp Leu Val His Trp Asp Glu His
85 90 95

Glu Ile Ala Ile Gly Pro Glu His Asp Asn Glu Gly Ile Phe Ser Gly
100 105 110

Ser Ile Val Val Asp His Asn Asn Thr Ser Gly Phe Phe Asn Ser Ser
115 120 125

Ile Asp Pro Asn Gln Arg Ile Val Ala Ile Tyr Thr Asn Asn Ile Pro
130 135 140

Asp Leu Gln Thr Gln Asp Ile Ala Phe Ser Leu Asp Gly Gly Tyr Thr
145 150 155 160

Phe Thr Lys Tyr Glu Asn Asn Pro Val Ile Asp Val Ser Ser Asn Gln
165 170 175

Phe Arg Asp Pro Lys Val Phe Trp His Glu Arg Phe Lys Ser Met Asp
180 185 190

His Gly Cys Ser Glu Ile Ala Arg Val Lys Ile Gln Ile Phe Gly Ser
195 200 205

Ala Asn Leu Lys Asn Trp Val Leu Asn Ser Asn Phe Ser Ser Gly Tyr
210 215 220

Tyr Gly Asn Gln Tyr Gly Met Ser Arg Leu Ile Glu Val Pro Ile Glu
225 230 235 240

Asn Ser Asp Lys Ser Lys Trp Val Met Phe Leu Ala Ile Asn Pro Gly
245 250 255

Ser Pro Leu Gly Gly Ser Ile Asn Gln Tyr Phe Val Gly Asp Phe Asp
260 265 270

Gly Phe Gln Phe Val Pro Asp Asp Ser Gln Thr Arg Phe Val Asp Ile
275 280 285

Gly Lys Asp Phe Tyr Ala Phe Gln Thr Phe Ser Glu Val Glu His Gly
290 295 300

026172-005020PC - Sequence Listing_1

Val Leu Gly Leu Ala Trp Ala Ser Asn Trp Gln Tyr Ala Asp Gln Val
305 310 315 320

Pro Thr Asn Pro Trp Arg Ser Ser Thr Ser Leu Ala Arg Asn Tyr Thr
325 330 335

Leu Arg Tyr Val Ile Gln Met Leu Lys Leu Thr Ala Asn Ile Asp Lys
340 345 350

Ser Val Leu Pro Asp Ser Ile Asn Val Val Asp Lys Leu Lys Lys Lys
355 360 365

Asn Val Lys Leu Thr Asn Lys Lys Pro Ile Lys Thr Asn Phe Lys Gly
370 375 380

Ser Thr Gly Leu Phe Asp Phe Asn Ile Thr Phe Lys Val Leu Asn Leu
385 390 395 400

Asn Val Ser Pro Gly Lys Thr His Phe Asp Ile Leu Ile Asn Ser Gln
405 410 415

Glu Leu Asn Ser Ser Val Asp Ser Ile Lys Ile Gly Phe Asp Ser Ser
420 425 430

Gln Ser Leu Phe Tyr Ile Asp Arg His Ile Pro Asn Val Glu Phe Pro
435 440 445

Arg Lys Gln Phe Phe Thr Asp Lys Leu Ala Ala Tyr Leu Glu Pro Leu
450 455 460

Asp Tyr Asp Gln Asp Leu Arg Val Phe Ser Leu Tyr Gly Ile Val Asp
465 470 475 480

Lys Asn Ile Ile Glu Leu Tyr Phe Asn Asp Gly Thr Val Ala Met Thr
485 490 495

Asn Thr Phe Phe Met Gly Glu Gly Lys Tyr Pro His Asp Ile Gln Ile
500 505 510

Val Thr Asp Thr Glu Glu Pro Leu Phe Glu Leu Glu Ser Val Ile Ile
515 520 525

Arg Glu Leu Asn Lys
530

<210> 24

<211> 654

<212> PRT

<213> Oryza sativa

<400> 24

Met Ala Thr Ser Arg Leu Thr Pro Ala Tyr Asp Leu Lys Asn Ala Ala

026172-005020PC - Sequence Listing_1

```

1           5           10           15
Ala Ala Val Tyr Thr Pro Leu Pro Glu Gln Pro His Ser Ala Glu Val
      20           25           30
Glu Ile Arg Asp Arg Lys Pro Phe Lys Ile Ile Ser Ala Ile Ile Leu
      35           40           45
Ser Ser Leu Leu Leu Leu Ala Leu Ile Leu Val Ala Val Asn Tyr Gln
      50           55           60
Ala Pro Pro Ser His Ser Ser Gly Asp Asn Ser Gln Pro Ala Ala Val
      65           70           75           80
Met Pro Pro Ser Arg Gly Val Ser Gln Gly Val Ser Glu Lys Ala Phe
      85           90           95
Arg Gly Ala Ser Gly Ala Gly Asn Gly Val Ser Phe Ala Trp Ser Asn
      100          105          110
Leu Met Leu Ser Trp Gln Arg Thr Ser Tyr His Phe Gln Pro Val Lys
      115          120          125
Asn Trp Met Asn Asp Pro Asn Gly Pro Leu Tyr Tyr Lys Gly Trp Tyr
      130          135          140
His Leu Phe Tyr Gln Tyr Asn Pro Asp Ser Ala Val Trp Gly Asn Ile
      145          150          155          160
Thr Trp Gly His Ala Val Ser Thr Asp Leu Ile Asn Trp Leu His Leu
      165          170          175
Pro Phe Ala Met Val Pro Asp Gln Trp Tyr Asp Val Asn Gly Val Trp
      180          185          190
Thr Gly Ser Ala Thr Ile Leu Pro Asp Gly Arg Ile Val Met Leu Tyr
      195          200          205
Thr Gly Asp Thr Asp Asp Tyr Val Gln Asp Gln Asn Leu Ala Phe Pro
      210          215          220
Ala Asn Leu Ser Asp Pro Leu Leu Val Asp Trp Val Lys Tyr Pro Asn
      225          230          235          240
Asn Pro Val Ile Tyr Pro Pro Pro Gly Ile Gly Val Lys Asp Phe Arg
      245          250          255
Asp Pro Thr Thr Ala Gly Thr Ala Gly Met Gln Asn Gly Gln Arg Leu
      260          265          270
Val Thr Ile Gly Ser Lys Val Gly Lys Thr Gly Ile Ser Leu Val Tyr

```

026172-005020PC - Sequence Listing_1

275

280

285

Glu Thr Thr Asn Phe Thr Thr Phe Lys Leu Leu Tyr Gly Val Leu His
 290 295 300

Ala Val Pro Gly Thr Gly Met Trp Glu Cys Val Asp Leu Tyr Pro Val
 305 310 315 320

Ser Thr Thr Gly Glu Asn Gly Leu Asp Thr Ser Val Asn Gly Leu Gly
 325 330 335

Val Lys His Val Leu Lys Thr Ser Leu Asp Asp Asp Lys His Asp Tyr
 340 345 350

Tyr Ala Leu Gly Thr Tyr Asp Pro Val Lys Asn Lys Trp Thr Pro Asp
 355 360 365

Asn Pro Asp Leu Asp Val Gly Ile Gly Leu Arg Leu Asp Tyr Gly Lys
 370 375 380

Tyr Tyr Ala Ala Arg Thr Phe Tyr Asp Gln Asn Lys Gln Arg Arg Ile
 385 390 395 400

Leu Trp Gly Trp Ile Gly Glu Thr Asp Leu Glu Ala Val Asp Leu Met
 405 410 415

Lys Gly Trp Ala Ser Leu Gln Ala Ile Pro Arg Thr Ile Val Phe Asp
 420 425 430

Lys Lys Thr Gly Thr Asn Val Leu Gln Arg Pro Glu Glu Glu Val Glu
 435 440 445

Ser Trp Ser Ser Gly Asp Pro Ile Thr Gln Arg Arg Ile Phe Glu Pro
 450 455 460

Gly Ser Val Val Pro Ile His Val Ser Gly Ala Thr Gln Leu Asp Ile
 465 470 475 480

Thr Ala Ser Phe Glu Val Asp Glu Thr Leu Leu Glu Thr Thr Ser Glu
 485 490 495

Ser His Asp Ala Gly Tyr Asp Cys Ser Asn Ser Gly Gly Ala Gly Thr
 500 505 510

Arg Gly Ser Leu Gly Pro Phe Gly Leu Leu Val Val Ala Asp Glu Lys
 515 520 525

Leu Ser Glu Leu Thr Pro Val Tyr Leu Tyr Val Ala Lys Gly Gly Asp
 530 535 540

Gly Lys Ala Lys Ala His Leu Cys Ala Tyr Gln Thr Arg Ser Ser Met

026172-005020PC - Sequence Listing_1

545 550 555 560

Ala Ser Gly Val Glu Lys Glu Val Tyr Gly Ser Ala Val Pro Val Leu
565 570 575

Asp Gly Glu Asn Tyr Ser Ala Arg Ile Leu Ile Asp His Ser Ile Val
580 585 590

Glu Ser Phe Ala Gln Ala Gly Arg Thr Cys Val Arg Ser Arg Asp Tyr
595 600 605

Pro Thr Lys Asp Ile Tyr Gly Ala Ala Arg Cys Phe Phe Phe Asn Asn
610 615 620

Ala Thr Glu Ala Ser Val Arg Ala Ser Leu Lys Ala Trp Gln Met Lys
625 630 635 640

Ser Phe Ile Arg Pro Tyr Pro Phe Ile Pro Asp Gln Lys Ser
645 650

<210> 25
<211> 690
<212> PRT
<213> Allium cepa

<400> 25
Met Ser Ser Asp Asp Leu Glu Ser Pro Pro Ser Ser Tyr Leu Pro Ile
1 5 10 15

Pro Pro Ser Asp Glu Phe His Asp Gln Pro Pro Pro Leu Arg Ser Trp
20 25 30

Leu Arg Leu Leu Ser Ile Pro Leu Ala Leu Met Phe Leu Leu Phe Leu
35 40 45

Ala Thr Phe Leu Ser Asn Leu Glu Ser Pro Pro Ser Asp Ser Gly Leu
50 55 60

Val Ser Asp Pro Val Thr Phe Asp Val Asn Pro Ala Val Val Arg Arg
65 70 75 80

Gly Lys Asp Ala Gly Val Ser Asp Lys Thr Ser Gly Val Asp Ser Gly
85 90 95

Phe Val Leu Asp Pro Val Ala Val Asp Ala Asn Ser Val Val Val His
100 105 110

Arg Gly Lys Asp Ala Gly Val Ser Asp Lys Thr Ser Gly Val Asp Ser
115 120 125

Gly Leu Leu Lys Asp Ser Pro Leu Gly Pro Tyr Pro Trp Thr Asn Gln
130 135 140

026172-005020PC - Sequence Listing_1

Met	Leu	Ser	Trp	Gln	Arg	Thr	Gly	Phe	His	Phe	Gln	Pro	Val	Lys	Asn
145					150					155					160
Trp	Met	Asn	Asp	Pro	Asn	Gly	Pro	Leu	Tyr	Tyr	Lys	Gly	Trp	Tyr	His
				165					170					175	
Phe	Phe	Tyr	Gln	Tyr	Asn	Pro	Glu	Gly	Ala	Val	Trp	Gly	Asn	Ile	Ala
			180					185					190		
Trp	Gly	His	Ala	Val	Ser	Arg	Asp	Leu	Val	His	Trp	Thr	His	Leu	Pro
		195					200					205			
Leu	Ala	Met	Val	Pro	Asp	Gln	Trp	Tyr	Asp	Ile	Asn	Gly	Val	Trp	Thr
	210					215					220				
Gly	Ser	Ala	Thr	Ile	Leu	Pro	Asp	Gly	Gln	Ile	Val	Met	Leu	Tyr	Thr
225					230					235					240
Gly	Ala	Thr	Asn	Glu	Ser	Val	Gln	Val	Gln	Asn	Leu	Ala	Val	Pro	Ala
				245					250					255	
Asp	Gln	Ser	Asp	Thr	Leu	Leu	Leu	Arg	Trp	Lys	Lys	Ser	Glu	Ala	Asn
			260					265					270		
Pro	Ile	Leu	Val	Pro	Pro	Pro	Gly	Ile	Gly	Asp	Lys	Asp	Phe	Arg	Asp
		275					280					285			
Pro	Thr	Thr	Ala	Trp	Tyr	Glu	Pro	Ser	Asp	Asp	Thr	Trp	Arg	Ile	Val
	290					295					300				
Ile	Gly	Ser	Lys	Asp	Ser	Ser	His	Ser	Gly	Ile	Ala	Ile	Val	Tyr	Ser
305					310					315					320
Thr	Lys	Asp	Phe	Ile	Asn	Tyr	Lys	Leu	Ile	Pro	Gly	Ile	Leu	His	Ala
				325					330					335	
Val	Glu	Arg	Val	Gly	Met	Trp	Glu	Cys	Val	Asp	Phe	Tyr	Pro	Val	Ala
			340					345					350		
Thr	Ala	Asp	Ser	Ser	His	Ala	Asn	His	Gly	Leu	Asp	Pro	Ser	Ala	Arg
		355					360					365			
Pro	Ser	Pro	Ala	Val	Lys	His	Val	Leu	Lys	Ala	Ser	Met	Asp	Asp	Asp
	370					375					380				
Arg	His	Asp	Tyr	Tyr	Ala	Ile	Gly	Thr	Tyr	Asp	Pro	Ala	Gln	Asn	Thr
385					390					395					400
Trp	Val	Pro	Asp	Asp	Ala	Ser	Val	Asp	Val	Gly	Ile	Gly	Leu	Arg	Tyr
				405					410					415	

026172-005020PC - Sequence Listing_1

Asp Trp Gly Lys Phe Tyr Ala Ser Lys Thr Phe Tyr Asp His Ala Lys
420 425 430

Lys Arg Arg Ile Leu Trp Ser Trp Ile Gly Glu Thr Asp Ser Glu Thr
435 440 445

Ala Asp Ile Ala Lys Gly Trp Ala Ser Leu Gln Gly Val Pro Arg Thr
450 455 460

Val Leu Leu Asp Val Lys Thr Gly Ser Asn Leu Ile Thr Trp Pro Val
465 470 475 480

Val Glu Ile Glu Ser Leu Arg Thr Arg Pro Arg Asp Phe Ser Gly Ile
485 490 495

Thr Val Asp Ala Gly Ser Thr Phe Lys Leu Asp Val Gly Gly Ala Ala
500 505 510

Gln Leu Asp Ile Glu Ala Glu Phe Lys Ile Ser Ser Glu Glu Leu Glu
515 520 525

Ala Val Lys Glu Ala Asp Val Ser Tyr Asn Cys Ser Ser Ser Gly Gly
530 535 540

Ala Ala Glu Arg Gly Val Leu Gly Pro Phe Gly Leu Leu Val Leu Ala
545 550 555 560

Asn Gln Asp Leu Thr Glu Gln Thr Ala Thr Tyr Phe Tyr Val Ser Arg
565 570 575

Gly Met Asp Gly Gly Leu Asn Thr His Phe Cys Gln Asp Glu Lys Arg
580 585 590

Ser Ser Lys Ala Ser Asp Ile Val Lys Arg Ile Val Gly His Ser Val
595 600 605

Pro Val Leu Asp Gly Glu Ser Phe Ala Leu Arg Ile Leu Val Asp His
610 615 620

Ser Ile Val Glu Ser Phe Ala Gln Gly Gly Arg Ala Ser Ala Thr Ser
625 630 635 640

Arg Val Tyr Pro Thr Glu Ala Ile Tyr Asn Asn Ala Arg Val Phe Val
645 650 655

Phe Asn Asn Ala Thr Gly Ala Lys Val Thr Ala Gln Ser Leu Lys Val
660 665 670

Trp His Met Ser Thr Ala Ile Asn Glu Ile Tyr Asp Pro Ala Thr Ser
675 680 685

026172-005020PC - Sequence Listing_1

Val Met
690

<210> 26
<211> 501
<212> PRT
<213> Beta vulgaris

<220>
<221> MOD_RES
<222> (382)..(382)
<223> Any amino acid

<400> 26
Leu Phe Tyr Gln Tyr Asn Pro Asn Gly Val Ile Trp Gly Pro Pro Val
1 5 10 15

Trp Gly His Ser Thr Ser Lys Asp Leu Val Asn Trp Val Pro Gln Pro
20 25 30

Leu Thr Met Glu Pro Glu Met Ala Ala Asn Ile Asn Gly Ser Trp Ser
35 40 45

Gly Ser Ala Thr Ile Leu Pro Gly Asn Lys Pro Ala Ile Leu Phe Thr
50 55 60

Gly Leu Asp Pro Lys Tyr Glu Gln Val Gln Val Leu Ala Tyr Pro Lys
65 70 75 80

Asp Thr Ser Asp Pro Asn Leu Lys Glu Trp Phe Leu Ala Pro Gln Asn
85 90 95

Pro Val Met Phe Pro Thr Pro Gln Asn Gln Ile Asn Ala Thr Ser Phe
100 105 110

Arg Asp Pro Thr Thr Ala Trp Arg Leu Pro Asp Gly Val Trp Arg Leu
115 120 125

Leu Ile Gly Ser Lys Arg Gly Gln Arg Gly Leu Ser Leu Leu Phe Arg
130 135 140

Ser Arg Asp Phe Val His Trp Val Gln Ala Lys His Pro Leu Tyr Ser
145 150 155 160

Asp Lys Leu Ser Gly Met Trp Glu Cys Pro Asp Phe Phe Pro Val Tyr
165 170 175

Ala Asn Gly Asp Gln Met Gly Val Asp Thr Ser Ile Ile Gly Ser His
180 185 190

Val Lys His Val Leu Lys Asn Ser Leu Asp Ile Thr Lys His Asp Ile
195 200 205

026172-005020PC - Sequence Listing_1

Tyr Thr Ile Gly Asp Tyr Asn Ile Lys Lys Asp Ala Tyr Thr Pro Asp
 210 215 220
 Ile Gly Tyr Met Asn Asp Ser Ser Leu Arg Tyr Asp Tyr Gly Lys Tyr
 225 230 235 240
 Tyr Ala Ser Lys Thr Phe Phe Asp Asp Ala Lys Lys Glu Arg Ile Leu
 245 250 255
 Leu Gly Trp Ala Asn Glu Ser Ser Ser Val Glu Asp Asp Ile Lys Lys
 260 265 270
 Gly Trp Ser Gly Ile His Thr Ile Pro Arg Lys Ile Trp Leu Asp Lys
 275 280 285
 Leu Gly Lys Gln Leu Ile Gln Trp Pro Ile Ala Asn Ile Glu Lys Leu
 290 295 300
 Arg Gln Lys Pro Val Asn Ile Tyr Arg Lys Val Leu Lys Gly Gly Ser
 305 310 315 320
 Gln Ile Glu Val Ser Gly Ile Thr Ala Ala Gln Ala Asp Val Glu Ile
 325 330 335
 Ser Phe Lys Ile Lys Asp Leu Lys Asn Val Glu Lys Phe Asp Ala Ser
 340 345 350
 Trp Thr Ser Pro Gln Leu Leu Cys Ser Lys Lys Gly Ala Ser Val Lys
 355 360 365
 Gly Gly Leu Gly Pro Phe Gly Leu Leu Thr Leu Ala Ser Xaa Gly Leu
 370 375 380
 Glu Glu Tyr Thr Ala Val Phe Phe Arg Ile Phe Lys Ala Tyr Asp Asn
 385 390 395 400
 Lys Phe Val Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser Leu Asn
 405 410 415
 Pro Thr Asn Asp Lys Thr Thr Tyr Gly Thr Phe Val Asp Val Asn Pro
 420 425 430
 Ile Arg Glu Gly Leu Ser Leu Arg Val Leu Ile Asp His Ser Val Val
 435 440 445
 Glu Ser Phe Gly Ala Lys Gly Lys Asn Val Ile Thr Ala Arg Val Tyr
 450 455 460
 Pro Thr Leu Ala Ile Asn Glu Lys Ala His Leu Tyr Val Phe Asn Arg
 465 470 475 480

026172-005020PC - Sequence Listing_1

Gly Thr Ser Asn Val Glu Ile Thr Gly Leu Thr Ala Trp Ser Met Lys
485 490 495

Lys Ala Asn Ile Ala
500

<210> 27

<211> 518

<212> PRT

<213> Bifidobacterium breve

<400> 27

Met Thr Asp Phe Thr Pro Glu Thr Pro Val Leu Thr Pro Ile Arg Asp
1 5 10 15

His Ala Ala Glu Leu Ala Lys Ala Glu Ala Gly Val Ala Glu Met Ala
20 25 30

Ala Lys Arg Asn Asn Arg Trp Tyr Pro Lys Tyr His Ile Ala Ser Asn
35 40 45

Gly Gly Trp Ile Asn Asp Pro Asn Gly Leu Cys Phe Tyr Lys Gly Arg
50 55 60

Trp His Val Phe Tyr Gln Leu His Pro Tyr Gly Thr Gln Trp Gly Pro
65 70 75 80

Met His Trp Gly His Val Ser Ser Thr Asp Met Leu Asn Trp Lys Arg
85 90 95

Glu Pro Ile Met Phe Ala Pro Ser Leu Glu Gln Glu Lys Asp Gly Val
100 105 110

Phe Ser Gly Ser Ala Val Ile Asp Asp Asn Gly Asp Leu Arg Phe Tyr
115 120 125

Tyr Thr Gly His Arg Trp Ala Asn Gly His Asp Asn Thr Gly Gly Asp
130 135 140

Trp Gln Val Gln Met Thr Ala Leu Pro Asp Asn Asp Glu Leu Thr Ser
145 150 155 160

Ala Thr Lys Gln Gly Met Ile Ile Asp Cys Pro Thr Asp Lys Val Asp
165 170 175

His His Tyr Arg Asp Pro Lys Val Trp Lys Thr Gly Asp Thr Trp Tyr
180 185 190

Met Thr Phe Gly Val Ser Ser Glu Asp Lys Arg Gly Gln Met Trp Leu
195 200 205

Phe Ser Ser Lys Asp Met Val Arg Trp Glu Tyr Glu Arg Val Leu Phe

026172-005020PC - Sequence Listing_1

210

215

220

Gln His Pro Asp Pro Asp Val Phe Met Leu Glu Cys Pro Asp Phe Phe
225 230 235 240

Pro Ile Lys Asp Lys Asp Gly Asn Glu Lys Trp Val Ile Gly Phe Ser
245 250 255

Ala Met Gly Ser Lys Pro Ser Gly Phe Met Asn Arg Asn Val Asn Asn
260 265 270

Ala Gly Tyr Met Ile Gly Thr Trp Glu Pro Gly Gly Glu Phe Lys Pro
275 280 285

Glu Thr Glu Phe Arg Leu Trp Asp Cys Gly His Asn Tyr Tyr Ala Pro
290 295 300

Gln Ser Phe Asn Val Asp Gly Arg Gln Ile Val Tyr Gly Trp Met Ser
305 310 315 320

Pro Phe Val Gln Pro Ile Pro Met Glu Asp Asp Gly Trp Cys Gly Gln
325 330 335

Leu Thr Leu Pro Arg Glu Ile Thr Leu Asp Asp Asp Gly Asp Val Val
340 345 350

Thr Ala Pro Val Ala Glu Met Glu Gly Leu Arg Glu Asp Thr Leu Asp
355 360 365

His Gly Ser Ile Thr Leu Asp Met Asp Gly Glu Gln Val Ile Ala Asp
370 375 380

Asp Ala Glu Ala Val Glu Ile Glu Met Thr Ile Asp Leu Ala Ala Ser
385 390 395 400

Thr Ala Asp Arg Ala Gly Leu Lys Ile His Ala Thr Glu Asp Gly Ala
405 410 415

Tyr Thr Tyr Val Ala Tyr Asp Asp Gln Ile Gly Arg Val Val Val Asp
420 425 430

Arg Gln Ala Met Ala Asn Gly Asp His Gly Tyr Arg Ala Ala Pro Leu
435 440 445

Thr Asp Ala Glu Leu Ala Ser Gly Lys Leu Asp Leu Arg Val Phe Val
450 455 460

Asp Arg Gly Ser Val Glu Val Tyr Val Asn Gly Gly His Gln Val Leu
465 470 475 480

Ser Ser Tyr Ser Tyr Ala Ser Glu Gly Pro Arg Ala Ile Lys Leu Val

Ala Glu Phe Gly Asn Leu Lys Val Glu Ser Leu Lys Leu His His Met
500 505 510

Lys Ser Ile Gly Leu Glu
515

<210> 28
<211> 532
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 28
Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys
1 5 10 15

Ile Ser Ala Ser Met Thr Asn Glu Thr Ser Asp Arg Pro Leu Val His
20 25 30

Phe Thr Pro Asn Lys Gly Trp Met Asn Asp Pro Asn Gly Leu Trp Tyr
35 40 45

Asp Glu Lys Asp Ala Lys Trp His Leu Tyr Phe Gln Tyr Asn Pro Asn
50 55 60

Asp Thr Val Trp Gly Thr Pro Leu Phe Trp Gly His Ala Thr Ser Asp
65 70 75 80

Asp Leu Thr Asn Trp Glu Asp Gln Pro Ile Ala Ile Ala Pro Lys Arg
85 90 95

Asn Asp Ser Gly Ala Phe Ser Gly Ser Met Val Val Asp Tyr Asn Asn
100 105 110

Thr Ser Gly Phe Phe Asn Asp Thr Ile Asp Pro Arg Gln Arg Cys Val
115 120 125

Ala Ile Trp Thr Tyr Asn Thr Pro Glu Ser Glu Glu Gln Tyr Ile Ser
130 135 140

Tyr Ser Leu Asp Gly Gly Tyr Thr Phe Thr Glu Tyr Gln Lys Asn Pro
145 150 155 160

Val Leu Ala Ala Asn Ser Thr Gln Phe Arg Asp Pro Lys Val Phe Trp
165 170 175

Tyr Glu Pro Ser Gln Lys Trp Ile Met Thr Ala Ala Lys Ser Gln Asp
180 185 190

Tyr Lys Ile Glu Ile Tyr Ser Ser Asp Asp Leu Lys Ser Trp Lys Leu
195 200 205

026172-005020PC - Sequence Listing_1

Glu Ser Ala Phe Ala Asn Glu Gly Phe Leu Gly Tyr Gln Tyr Glu Cys
210 215 220

Pro Gly Leu Ile Glu Val Pro Thr Glu Gln Asp Pro Ser Lys Ser Tyr
225 230 235 240

Trp Val Met Phe Ile Ser Ile Asn Pro Gly Ala Pro Ala Gly Gly Ser
245 250 255

Phe Asn Gln Tyr Phe Val Gly Ser Phe Asn Gly Thr His Phe Glu Ala
260 265 270

Phe Asp Asn Gln Ser Arg Val Val Asp Phe Gly Lys Asp Tyr Tyr Ala
275 280 285

Leu Gln Thr Phe Phe Asn Thr Asp Pro Thr Tyr Gly Ser Ala Leu Gly
290 295 300

Ile Ala Trp Ala Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn
305 310 315 320

Pro Trp Arg Ser Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr
325 330 335

Glu Tyr Gln Ala Asn Pro Glu Thr Glu Leu Ile Asn Leu Lys Ala Glu
340 345 350

Pro Ile Leu Asn Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe Ala Thr
355 360 365

Asn Thr Thr Leu Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu Ser Asn
370 375 380

Ser Thr Gly Thr Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr
385 390 395 400

Gln Thr Ile Ser Lys Ser Val Phe Ala Asp Leu Ser Leu Trp Phe Lys
405 410 415

Gly Leu Glu Asp Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu Val Ser
420 425 430

Ala Ser Ser Phe Phe Leu Asp Arg Gly Asn Ser Lys Val Lys Phe Val
435 440 445

Lys Glu Asn Pro Tyr Phe Thr Asn Arg Met Ser Val Asn Asn Gln Pro
450 455 460

Phe Lys Ser Glu Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu
465 470 475 480

026172-005020PC - Sequence Listing_1

Asp Gln Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val Val Ser
485 490 495

Thr Asn Thr Tyr Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn
500 505 510

Met Thr Thr Gly Val Asp Asn Leu Phe Tyr Ile Asp Lys Phe Gln Val
515 520 525

Arg Glu Val Lys
530

<210> 29

<211> 512

<212> PRT

<213> Zymomonas mobilis

<400> 29

Met Glu Ser Pro Ser Tyr Lys Asn Leu Ile Lys Ala Glu Asp Ala Gln
1 5 10 15

Lys Lys Ala Gly Lys Arg Leu Leu Ser Ser Glu Trp Tyr Pro Gly Phe
20 25 30

His Val Thr Pro Leu Thr Gly Trp Met Asn Asp Pro Asn Gly Leu Ile
35 40 45

Phe Phe Lys Gly Glu Tyr His Leu Phe Tyr Gln Tyr Tyr Pro Phe Ala
50 55 60

Pro Val Trp Gly Pro Met His Trp Gly His Ala Lys Ser Arg Asp Leu
65 70 75 80

Val His Trp Glu Thr Leu Pro Val Ala Leu Ala Pro Gly Asp Leu Phe
85 90 95

Asp Arg Asp Gly Cys Phe Ser Gly Cys Ala Val Asp Asn Asn Gly Val
100 105 110

Leu Thr Leu Ile Tyr Thr Gly His Ile Val Leu Ser Asn Asp Ser Pro
115 120 125

Asp Ala Ile Arg Glu Val Gln Cys Met Ala Thr Ser Ile Asp Gly Ile
130 135 140

His Phe Gln Lys Glu Gly Ile Val Leu Glu Lys Ala Pro Met Pro Gln
145 150 155 160

Val Ala His Phe Arg Asp Pro Arg Val Trp Lys Glu Asn Asp His Trp
165 170 175

Phe Met Val Val Gly Tyr Arg Thr Asp Asp Glu Lys His Gln Gly Ile

026172-005020PC - Sequence Listing_1

180

185

190

Gly His Val Ala Leu Tyr Arg Ser Glu Asn Leu Lys Asp Trp Ile Phe
 195 200 205

Val Lys Thr Leu Leu Gly Asp Asn Ser Gln Leu Pro Leu Gly Lys Arg
 210 215 220

Ala Phe Met Trp Glu Cys Pro Asp Phe Phe Ser Leu Gly Asn Arg Ser
 225 230 235 240

Val Leu Met Phe Ser Pro Gln Gly Leu Lys Ala Ser Gly Tyr Lys Asn
 245 250 255

Arg Asn Leu Phe Gln Asn Gly Tyr Ile Leu Gly Lys Trp Gln Ala Pro
 260 265 270

Gln Phe Thr Pro Glu Thr Ser Phe Gln Glu Leu Asp Tyr Gly His Asp
 275 280 285

Phe Tyr Ala Ala Gln Arg Phe Glu Ala Lys Asp Gly Arg Gln Ile Leu
 290 295 300

Ile Ala Trp Phe Asp Met Trp Glu Asn Gln Lys Pro Ser Gln Arg Asp
 305 310 315 320

Gly Trp Ala Gly Cys Met Thr Leu Pro Arg Lys Leu Asp Leu Ile Asp
 325 330 335

Asn Lys Ile Val Met Thr Pro Val Arg Glu Met Glu Ile Leu Arg Gln
 340 345 350

Ser Glu Lys Ile Glu Ser Val Val Thr Leu Ser Asp Ala Glu His Pro
 355 360 365

Phe Thr Met Asp Ser Pro Leu Gln Glu Ile Glu Leu Ile Phe Asp Leu
 370 375 380

Glu Lys Ser Ser Ala Tyr Gln Ala Gly Leu Ala Leu Arg Cys Asn Gly
 385 390 395 400

Lys Gly Gln Glu Thr Leu Leu Tyr Ile Asp Arg Ser Gln Asn Arg Ile
 405 410 415

Ile Leu Asp Arg Asn Arg Ser Gly Gln Asn Val Lys Gly Ile Arg Ser
 420 425 430

Cys Pro Leu Pro Asn Thr Ser Lys Val Arg Leu His Ile Phe Leu Asp
 435 440 445

Arg Ser Ser Ile Glu Ile Phe Val Gly Asp Asp Gln Thr Gln Gly Leu

026172-005020PC - Sequence Listing_1

450

455

460

Tyr Ser Ile Ser Ser Arg Ile Phe Pro Asp Lys Asp Ser Leu Lys Gly
 465 470 475 480

Arg Leu Phe Ala Ile Glu Gly Tyr Ala Val Phe Asp Ser Phe Lys Arg
 485 490 495

Trp Thr Leu Gln Asp Ala Asn Leu Ala Ala Phe Ser Ser Asp Ala Cys
 500 505 510

<210> 30

<211> 987

<212> DNA

<213> *Chlorella luteoviridis*

<400> 30

```

tgttgaagaa tgagccggcg acttatagga agtggcttgg ttaaggatac tttccgaagc      60
ctaagcgaaa gcaagttgta acaatagcga tatacctctt tgtaggtcag tcacttctta      120
tggaacccgaa cccgggtgat ctaaccatga ccaggatgaa gcttgggtaa caccaagtga      180
aggtccgaac tcttcgatct ttaaaaatcg tgagatgagt tatggttagg ggtaaatctg      240
gcagttttgc cccgcaaaag ggtaaccttt tgtaattact gactcataac ggtgaagcct      300
aaggcgttag ctatggtaat accgtgggaa gtttcaatac cttcttgcac attttttatt      360
tgcaccttta gtgcaaacag tgtaaagaaa gcgttttgaa accccttaac gactaatttt      420
ttgcttttgc aagaacgtca gcactcacca atacactttc cgtttttttc ttttattaat      480
taaagcaaca taaaaatata ttttatagct ttaatcataa aactatgtta gcacttcgtg      540
ctaattgtgct aatgtgctaa tcaaatgaaa agtgttctta aaagtgagtt gaaggtagag      600
tctaattcttg cctgaaaggg caagctgcac attttttttt gaatgtgcaa caatggaaat      660
gccaatcgaa ctcgagagcta gctggttctc cccgaaatgt gttgaggcgc agcgattcat      720
gattagtacg gtgtaggggt aaagcactgt ttcggtgcgg gctgtgaaaa cggtaccaaa      780
tcgtggcaaa ctaagaatac tacgcttgta taccatggat cagtgagact atgggggata      840
agctccatag tcaagagggg aacagcccag atcaccagtt aaggcccca aatgacagct      900
aagtggcaaa ggaggtgaaa gtgcagaaac aaccaggagg ttgcccaga agcagccatc      960
ctttaagag tgcgtaatag ctactg      987

```

<210> 31

<211> 1412

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 31

```

gaattcgagt ttaggtccag cgtccgtggg gggggacggg ctgggagctt gggccgggaa      60

```

026172-005020PC - Sequence Listing_1

```

gggcaagacg atgcagtcct tctggggagt cacagccgac tgtgtgtgtt gcactgtgcg      120
gcccgcagca ctcacacgca aaatgcctgg ccgacaggca ggccctgtcc agtgcaacat      180
ccacggtccc tctcatcagg ctcaccttgc tcattgacat aacggaatgc gtaccgctct      240
ttcagatctg tccatccaga gaggggagca ggctccccac cgacgctgtc aaacttgctt      300
cctgcccacg cgaaaacatt attgtttgag gggggggggg gggggggcaga ttgcatggcg      360
ggatatctcg tgaggaacat cactgggaca ctgtggaaca cagtgagtgc agtatgcaga      420
gcatgtatgc taggggtcag cgcaggaagg gggcctttcc cagtctccca tgccactgca      480
ccgtatccac gactcaccag gaccagcttc ttgatcggct tccgctcccg tggacaccag      540
tgtgtagcct ctggactcca ggtatgcgtg caccgcaaag gccagccgat cgtgccgatt      600
cctgggggtg aggatatgag tcagccaact tggggctcag agtgcacact ggggcacgat      660
acgaaacaac atctacaccg tgtcctccat gctgacacac cacagcttcg ctccacctga      720
atgtgggagc atggggccga atcacagcca atgtcgtgct tgccataatg tgatccagac      780
cctctccgcc cagatgccga gcggatcgtg ggcgctgaat agattcctgt ttcatcact      840
gtttgggtcc ttctcttttc gtctcggatg cgcgtctcga aacaggctgc gtcgggcttt      900
cggatccctt ttgctccctc cgtcaccatc ctgcgcgcgg gcaagttgct tgaccctggg      960
ctggtaccag ggttgagggg tattaccgag tcaggccatt cccagcccgg attcaattca     1020
aagtctgggc caccaccctc cgccgctctg tctgatcact ccacattcgt gcataacta     1080
cgttcaagtc ctgatccagg cgtgtctcgg gacaagggtg gcttgagttt gaatctcaag     1140
gaccactccc agcacagctg ctggttgacc ccgccctcgc aactccctac catgtctgct     1200
ggtaggtcca gggatctttg ccatgcacac aggaccccggt ttgtgggggt ccccggtgca     1260
tgctgtcgtg gtgcaggcgc cgggtgtgggg cctggggccc gcgggagctc aactcctccc     1320
catatgcctg ccgtccctcc caccacccgc gacctggccc cctttgcaga ggaaggcgaa     1380
gtcagcgcca tcgtgtgcga taatggatcc gg                                     1412

```

<210> 32

<211> 1627

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 32

```

gaattcgccc ttgagtttag gtccagcgtc cgtggggggg gcgtgagact cccccctgac      60
cttcgtatgg cagggactcc tacttgccaa gtaatcagtt gacaatgcca cttcaatgct     120
cgttgtggta cactgacgag ggtctaacat actgggaagc atgaattgcc gacatggact     180
cagttggaga cagtaacagc tctttgtgtt ctatcttcag gaacacattt ggcagcgcac     240
ccatacagtg gcgcacacgc agctgtacct gatgtggctc tattcccaca tgtttcaact     300
tgatccaaaa gtcactcaga ctctcagcag ctagacttga tcgcatcttt ggccatgaag     360

```

026172-005020PC - Sequence Listing_1

atgcttgatg	aactctagga	atgggacgag	aaaagagcct	gctctgatcg	gatatttcca	420
ttctctggat	gggactgaga	tgattctgaa	gaaatgctgc	tcgacttatt	tggaagaaca	480
gcacctgacg	catgctttga	ggctgctgtg	gctgggatgt	gctgtatttg	tcagcattga	540
gcatctacgg	gtagatggcc	ataaccacgc	gctgcctatc	atgcggtggg	ttgtgtggaa	600
aacgtacaat	ggacagaaat	caatcccat	gcgagcctag	cgtgcagcca	tgcgctccct	660
ctgtagcccc	gctccaagac	aaagccagcc	aatgccagaa	cccacataga	gagggatatct	720
tcctaatagac	ctcgcccatc	atttcctcca	aattaactat	aatgccttga	ttgtggagtt	780
ggctttggct	tgcaagtgtc	cgcgctggca	ctttttagag	cagcacaggg	tatgccagcg	840
ccgaactttg	tgcccttgag	caggccacaa	gggcacaaga	ctacaccatg	cagctgggat	900
acttggaact	gataaccattc	ttaccaagca	aggcacagca	cagcctgcac	cgactcactt	960
tgcttgagcg	gggcacagcg	ccgcgactga	tcctgcgagc	tgtggggagt	tccgactgtt	1020
ctggacctcg	gtctctgaaa	gatgtgtacg	atgggatcaa	gtcattcaag	tatgctcttc	1080
acatgagcaa	tcgggggaga	cacggtggcc	ctaaaggtgt	tcacttgatt	caagtgtagt	1140
gggggggtgc	tgtttgtccc	ggggcgcccc	ccgctccccg	accccgagga	agggccccag	1200
aggactcggc	cgcccacaga	ggaataaccg	ggcgtggctc	ggccctgcgc	ctccctcttt	1260
caatatttca	cctggtgttc	agtgcacgga	cacgtaaaga	actagatata	atggccgagg	1320
gaaagacggt	gagagcttgg	cggttggtga	ccgggcagca	tcagaaactc	ctcttccccg	1380
cccgcttga	aactcactgt	aactccctcc	tcttccccct	cgcagcatct	gtctatcggt	1440
atcgtagagt	aaagggactg	ccatgtgtcg	ggtcggtgac	cacggtcggc	tcgggctgct	1500
ctgcccgcgt	cgcgaaacgt	ccctgcaaac	gccgcgcagc	cgtccctttt	tctgcccgcg	1560
ccccaccccc	tcgctcccc	cttcaatcac	accgcagtgc	ggacatgtcg	attccggcaa	1620
gtccacc						1627

<210> 33

<211> 570

<212> DNA

<213> Chlorella protothecoides

<400> 33

gaattccctg	caggaagaag	gccggcagca	gctggtactt	gtccttcacc	tccttgatcg	60
gctgggtgag	cttggccggg	tcgcagtcgt	cgatgccggc	atcgcccagc	acgctgtgcg	120
gggagccggc	atcgacaacc	ttggcactgc	tcaccttgg	caccggcatg	gggtcatggc	180
gctgcagacc	agcggcctgt	cagcatgctg	caggcatctg	tgttttgtag	tagatacttt	240
ctgatgcata	accacacgtt	tggaagggtc	ccaagccctt	tcaacagtct	cgacatatga	300
cactcgcgcc	ctcttcctcg	tcccgtggcc	tgatgaggg	acgcaggtac	cgcagctgcg	360
ccccgtcccg	ccagttgccc	tggccccgcc	gggccaatc	tgttcattgc	cgctccctgg	420
cagccgtgaa	cttcacacta	ccgctctctg	tgaccttcag	cacagcagga	atcgccattt	480

026172-005020PC - Sequence Listing_1

caccggcggt cgttgctgcg gagcctcagc tgatctcgcc tgcgagaccc cacagtttga 540
atttgcggtc cccacacaac ctctgacgcc 570

<210> 34
<211> 568
<212> DNA
<213> *Chlorella protothecoides*

<220>
<221> modified_base
<222> (133)..(133)
<223> a, c, t, g, unknown or other

<400> 34
gaattccctc aggaagaagg ccggcagcag ctggtacttg tccttcacct ccttgatcgg 60
ctgggtgagc ttcgcaggat cgcagtcgtc gatgccggca tcgcccagca cgctgtgcgg 120
ggagccggca tcnacaacct tggcactgct ccccttggtc accggcatgg ggtcatggcg 180
ctgcagccca gcggcctgtc agcatgctgc aggcattctgt gtattgtagt aggtacttcc 240
tgatgcatca acacacgttt ggaagctccc caagccccctt caacagtctc gacgtatgac 300
actcgcgccc tcttcctcgc cccgtggcct gatgagggta cgcaggtacc acagctgcgc 360
cccgccccgc cagttgccct ggcccggccg ggcccaatct gttcattgcc gctccctggt 420
agccgtgaac tcacattacc gctctctgtg accttcagca cagcaggaat cgccatttca 480
ccggcggtcg ttgctgcgga gcctcagctg atctcgctg cgagaccca cagtttgaat 540
ttgcggtccc cacacaacct ctgacgcc 568

<210> 35
<211> 50
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 35
tgacctaggc gattaattaa ctcgaggcag cagcagctcg gatagtatcg 50

<210> 36
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic primer

<400> 36
ctacgagctc aagctttcca tttgtgttcc catcccacta ctccc 45

<210> 37
<211> 32
<212> DNA
<213> Artificial Sequence

026172-005020PC - Sequence Listing_1

```

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 37
gatcagaatt ccgcctgcaa cgcaagggca gc 32

<210> 38
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      primer

<400> 38
gcatactagt ggcgggacgg agagagggcg 30

<210> 39
<211> 1568
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
      polynucleotide

<400> 39
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt 300
cagtcacaac ccgcaaacgg cgcgccatat caatgattga acaagatgga ttgcacgcag 360
gttctccggc cgcttggggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg 420
gctgctctga tgccgcccgtg ttccggctgt cagcgcaggg gcgcccgggtt ctttttgtca 480
agaccgacct gtccggtgcc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc 540
tggccacgac gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg 600
actggctgct attgggcgaa gtgcccggggc aggatctcct gtcattctac cttgctcctg 660
ccgagaaagt atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta 720
cctgcccatt cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag 780
ccggtcttgt cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac 840
tgttcgccag gctcaaggcg cgcattgccc acggcgagga tctcgtcgtg acccatggcg 900
atgcctgctt gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg 960
gccggctggg tgtggcggaac cgctatcagg acatagcggt ggctaccctg gatattgctg 1020
aagagcttgg cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccc 1080

```

026172-005020PC - Sequence Listing_1

attcgagcgc catcgcccttc tatcgcccttc ttgacgagtt cttctaagat ctgtcgatcg	1140
acaagtgact cgaggcagca gcagctcgga tagtatcgac acactctgga cgctggctgt	1200
gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct gccgctttta	1260
tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct tttgcgagtt gctagctgct	1320
tgtgctatatt gcgaatacca cccccagcat ccccttccct cgtttcatat cgcttgcatc	1380
ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg ctcttgctca	1440
ctgccccctcg cacagccttg gtttgggctc cgcttgatt ctcctgggtac tgcaacctgt	1500
aaaccagcac tgcaatgctg atgcacggga agtagtgga tgggaacaca aatggaaagc	1560
ttgagctc	1568

<210> 40

<211> 2571

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>

<221> modified_base

<222> (997)..(999)

<223> a, c, t, g, unknown or other

<400> 40

gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgctgccgc atgcttgctg tggtaggct gggcagtgt	120
gccatgctga ttgaggcttg gttcatcggg tggaaagctta tgtgtgtgct gggcttgcat	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcggtgcc	240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgct acaggttgag	300
cagcagtgtc agcaaaggta gcagtttgcc agaattgatc gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaactaac catagctgat caacactgca atcatcggcg gctgatgcaa	540
gcatcctgca agacacatgc tgtgcatg tgcgtgctg cctgctgctg acgccgttga	600
gttggcagca gctcagccat gcaactggatc aggctgggct gccactgcaa tgtggtggat	660
aggatgcaag tggagcgaat accaaaccct ctggctgctt gctgggttgc atggcatcgc	720
accatcagca ggagcgcatt cgaagggact ggccccatgc acgccatgcc aaaccggagc	780
gcaccgagtg tccacactgt caccaggccc gcaagctttg cagaacctat ctcatggacg	840
catgtagcgc tgacgtccct tgacggcgct cctctcgggt gtgggaaacg caatgcagca	900
caggcagcag aggcggcggc agcagagcgg cggcagcagc ggcgggggccc acccttcttg	960
cggggctcgc cccagccag cgggtgatgc ctgatchnnc caaacgagtt cacattcatt	1020

026172-005020PC - Sequence Listing_1

tgcagcctgg agaagcgagg ctggggcctt tgggctggtg cagcccgcaa tggaatgcgg	1080
gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcac	1140
tggactgcat ttgggtgggg cggccggctg tttctttcgt gttgcaaaac gcgccacgtc	1200
agcaacctgt cccgtgggtc ccccggtgccg atgaaatcgt gtgcacgccg atcagctgat	1260
tgcccggtc gcgaagtagg cgccctcttt ctgctcgccc tctctccgtc ccgccactag	1320
tggcgcgccca tatcaatgat tgaacaagat ggattgcacg caggttctcc ggccgcttgg	1380
gtggagaggc tattcggcta tgactgggca caacagacaa tcggctgctc tgatgccgcc	1440
gtgttccggc tgcagcgca ggggcgccc gttctttttg tcaagaccga cctgtccggt	1500
gcccctgaatg aactgcagga cgaggcagcg cggctatcgt ggctggccac gacgggcgtt	1560
ccttgcgcag ctgtgctcga cgttgtcact gaagcgggaa gggactggct gctattgggc	1620
gaagtgccgg ggcaggatct cctgtcatct caccttgctc ctgccgagaa agtatccatc	1680
atggctgatg caatgcggcg gctgcatacg cttgatccgg ctacctgccc attcgaccac	1740
caagcgaaac atcgcatcga gcgagcacgt actcggatgg aagccggtct tgtcgatcag	1800
gatgatctgg acgaagagca tcaggggctc gcgccagccg aactgttcgc caggctcaag	1860
gcgcgcatgc ccgacggcga ggatctcgtc gtgacccatg gcgatgcctg cttgccgaat	1920
atcatgggtg aaaatggccg cttttctgga ttcactgact gtggccggct ggggtgtggc	1980
gaccgctatc aggacatagc gttggctacc cgtgatattg ctgaagagct tggcggcgaa	2040
tgggctgacc gcttcctcgt gctttacggg atcgccgctc ccgattcgca gcgcatcgcc	2100
ttctatcgcc ttcttgacga gttcttctaa gatctgtcga tcgacaagtg actcgaggca	2160
gcagcagctc ggatagtatc gacacactct ggacgctggc cgtgtgatgg actgttgccg	2220
ccacacttgc tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt	2280
gtttgatctt gtgtgtacgc gcttttgcca gttgctagct gcttgtgcta tttgcgaata	2340
ccacccccag catccccttc cctcgtttca tatcgcttgc atcccaaccg caacttatct	2400
acgctgtcct gctatccctc agcgtgtcct ctgctcctgc tcaactgcccc tcgcacagcc	2460
ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg	2520
ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agcttgagct c	2571

<210> 41

<211> 2550

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<220>

<221> modified_base

<222> (997)..(999)

<223> a, c, t, g, unknown or other

026172-005020PC - Sequence Listing_1

<400> 41	
gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgatgc tggtagaggct gggcagtgct	120
gccatgctga ttgaggcttg gttcatcggg tggaaagctta tgtgtgtgct gggcttgcat	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggacgtg ccgcggtgcc	240
tccaggtggt tcaatcgcg gaggcagagg gatttcagat gatcgcgct acaggttgag	300
cagcagtgct agcaaaggta gcagtttgcc agaagatgcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcagtcac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaactaac catagctgat caacactgca atcatcggcg gctgatgcaa	540
gcatcctgca agacacatgc tgtgcgatgc tgcgctgctg cctgctgcgc acgccgttga	600
gttggcagca gctcagccat gcaactggatc aggcctgggt gccactgcaa tgtggtggat	660
aggatgcaag tggagcgaat accaaaccct ctggctgctt gctgggttgc atggcatcgc	720
accatcagca ggagcgcgtg cgaagggact ggcccatgc acgcatgcc aaaccggagc	780
gcaccgagtg tccacactgt caccaggccc gcaagctttg cagaaccatg ctcatggacg	840
catgtagcgc tgacgtccct tgacggcgct cctctcggtt gtgggaaacg caatgcagca	900
caggcagcag aggcggcgcc agcagagcgg cggcagcagc ggcgggggcc acccttcttg	960
cggggtcgcg cccagccag cggatgatgc ctgactnnnc caaacgagtt cacattcatt	1020
tgcagcctgg agaagcgagg ctggggcctt tgggctggtg cagcccgcaa tggaaatgcgg	1080
gaccgccagg ctagcagcaa aggcgcctcc cctactccgc atcgatgttc catagtgcac	1140
tggactgcat ttgggtgggg cggccggctg tttctttcgt gttgcaaac gcgccacgtc	1200
agcaacctgt cccgtgggtc ccccgctgcc atgaaatcgt gtgcacgccg atcagctgat	1260
tgcccggtc gcgaagtagg cggcctcttt ctgctcgccc tctctccgtc ccgccactag	1320
tggcgcgcca tatcaatgat cgagcaggac ggcctccacg ccggctcccc cgccgcctgg	1380
gtggagcgcc tgttcggcta cgactgggccc cagcagacca tcggctgctc cgacgccgcc	1440
gtgttcgcc tgtccgcca gggccgcccc gtgctgttcg tgaagaccga cctgtccggc	1500
gccctgaacg agctgcagga cgaggccgcc cgctgtcct ggctggccac caccggcgtg	1560
ccctgcgcc ccgtgctgga cgtggtgacc gaggccggcc gcgactggct gctgctgggc	1620
gaggtgccc gccaggacct gctgtcctcc cactggccc ccgccagaa ggtgtccatc	1680
atggccgacg ccatgcgcc cctgcacacc ctggaccccg ccacctgcc cttcgaccac	1740
caggccaagc accgcatcga gcgcgcccgc acccgcatgg aggcggcct ggtggaccag	1800
gacgacctgg acgaggagca ccagggcctg gccccgccg agctgttcgc ccgcctgaag	1860
gccccgatgc ccgacggcga ggacctggtg gtgaccacg gcgacgcctg cctgccaac	1920
atcatggtgg agaacggccg cttctccggc ttcacgact gcggccgcct gggcgtggcc	1980

026172-005020PC - Sequence Listing_1

gaccgctacc aggacatcgc cctggccacc cgcgacatcg ccgaggagct gggcggcgag	2040
tgggcccacc gcttcctggt gctgtacggc atcgccgccc ccgactccca gcgcatcgcc	2100
ttctaccgcc tgctggacga gttcttctga ctcgaggcag cagcagctcg gatagtatcg	2160
acacactctg gacgctggtc gtgtgatgga ctgttgccgc cacacttgct gccttgacct	2220
gtgaatatcc ctgccgcttt tatcaaacag cctcagtgtg ttgatcttg tgtgtacgcg	2280
cttttgcgag ttgctagctg cttgtgctat ttgcgaatac cacccccagc atccccttcc	2340
ctcgtttcat atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca	2400
gcgctgctcc tgctcctgct cactgcccct cgcacagcct tggtttgggc tccgcctgta	2460
ttctcctggt actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg	2520
gatgggaaca caaatggaaa gcttgagctc	2550

<210> 42

<211> 1547

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 42

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	60
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	120
atggggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	240
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	300
cagtcacaac ccgcaaacgg cgcgccatat caatgatcga gcaggacggc ctccacgccg	360
gctccccgc cgcctgggtg gagcgctgt tcggctacga ctgggcccag cagaccatcg	420
gctgctccga cgccgccgtg ttccgcctgt ccgcccaggg ccgccccgtg ctgttcgtga	480
agaccgacct gtccggcgcc ctgaacgagc tgcaggacga ggccgccgc ctgtcctggc	540
tggccaccac cggcgtgccc tgcgccgccg tgctggacgt ggtgaccgag gccggccgcg	600
actggctgct gctgggagag gtgcccggcc aggacctgct gtcctccac ctggcccccg	660
ccgagaaggt gtccatcatg gccgacgcca tgcgcgcct gcacaccctg gaccccgcca	720
cctgccccctt cgaccaccag gccaaagcacc gcatcgagcg cggccgcacc cgcattggagg	780
ccggcctggt ggaccaggac gacctggacg aggagcacca gggcctggcc cccgccgagc	840
tgttcgcccc cctgaaggcc cgcattgccg acggcgagga cctgggtggtg acccacggcg	900
acgcctgcct gcccaacatc atggtggaga acggccgctt ctccggcttc atcgactgcg	960
gccgcctggg cgtggccgac cgctaccagg acatcgccct ggccacccgc gacatcgccg	1020
aggagctggg cggcgagtgg gccgaccgct tcctgggtgct gtacggcatc gccgcccccg	1080
actcccagcg catcgcttc taccgcctgc tggacgagtt cttctgactc gaggcagcag	1140

026172-005020PC - Sequence Listing_1

cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac	1200
acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt	1260
gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac	1320
ccccagcatc cccttccttc gtttcataac gcttgcatcc caaccgcaac ttatctacgc	1380
tgtcctgcta tccctcagcg ctgctcctgc tctgctcac tgccccctgc acagccttgg	1440
tttgggctcc gcctgtattc tcctgggtact gcaacctgta aaccagcact gcaatgctga	1500
tgcacgggaa gtagtgggat gggaacacaa atggaaagct tgagctc	1547

<210> 43
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 43 gccgcgactg gctgctgctg g	21
-------------------------------------	----

<210> 44
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 44 aggtcctcgc cgtcgggcat g	21
-------------------------------------	----

<210> 45
 <211> 1292
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 45 atcaaaggca tagattcaca ttgtttggca ttgcagagca atcatcgcg aggacgaaca	60
tcgctcacca agcacgtact gggcatccgg aggcctccgc aaattcctgc aacaggactc	120
gctgatcagt tcgcccgaagg tctacgacgc tccctatcgg cgctagactt caacacatat	180
ttcactgtca cagcctcggc atgcatcagg cctcagtctc caccatgaag accatccagt	240
ctcggcacgc cggctccatc ggacatgtgc agtcgggtcg ccgatcggcg gggcgcgcg	300
gatcccgcat ggcgaccccc gtggccgcag ctaccgtcgc agccccctgc tcggccctca	360
acctctcccc caccatcatt cgacaggagg tgctccactc cgccagcgcc cagcaactag	420
actgcgtggc ctccctggcg cccgtcttcg agtcccagat cctccccctc ctgacgccc	480

026172-005020PC - Sequence Listing_1

tggacgagat gtggcagccc accgacttcc tccccgcctc gaactcggag gcattcttcg	540
accagatcgg cgacctgcgg gcgcgatcgg cggccatccc cgacgacctg ctggtctgcc	600
tgggtggggga catgatcacg gaggaggccc tgcccaccta catggccatg ctgaacaccc	660
tggacgtcgt gcgcgatgag acagggcaca gccagcacc ctagccaag tggaccagg	720
cttggatcgc ggaggagaac cgccatggcg acctgctgaa caagtacatg tggctgacgg	780
ggcgggtggg acatgctggc ggtggagcgc accatccagc catgctggcg gtggagcgca	840
ccatccagcg cctcatctca tcgggcatgg acccgggcac ggagaaccac ccctaccag	900
cctttgtgtt caccagcttc caggagcgcg ccaccaagct gagccacggc tccaccgccc	960
gcctggcggg cgccgccggg gacgaggccc tggccaagat ctgcgggacc attgcgcggg	1020
acgagtcgcg ccacgaggcg gcgtacacgc ggaccatgga tgccatcttc cagcgcgacc	1080
ccagcggggc catggtggcg tttgcgcaca tgatgatgcg caagatcacc atgcccgcgc	1140
acctcatgga cgacggccag cacggcgcgc gcaacggggg ggcgcaactt gttcgacgac	1200
tttgcggcag tggcggagcg ggcaggggtg tacaccgccg gcgactacat cggcatcctg	1260
cgccacctca tccggcgctg ggacgtggag gg	1292

<210> 46

<211> 364

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 46

Met	His	Gln	Ala	Ser	Val	Ser	Thr	Met	Lys	Thr	Ile	Gln	Ser	Arg	His
1				5					10					15	

Ala	Gly	Pro	Ile	Gly	His	Val	Gln	Ser	Gly	Arg	Arg	Ser	Ala	Gly	Arg
			20					25					30		

Ala	Gly	Ser	Arg	Met	Ala	Thr	Pro	Val	Ala	Ala	Ala	Thr	Val	Ala	Ala
		35					40					45			

Pro	Arg	Ser	Ala	Leu	Asn	Leu	Ser	Pro	Thr	Ile	Ile	Arg	Gln	Glu	Val
	50				55						60				

Leu	His	Ser	Ala	Ser	Ala	Gln	Gln	Leu	Asp	Cys	Val	Ala	Ser	Leu	Ala
65					70					75				80	

Pro	Val	Phe	Glu	Ser	Gln	Ile	Leu	Pro	Leu	Leu	Thr	Pro	Val	Asp	Glu
			85					90						95	

Met	Trp	Gln	Pro	Thr	Asp	Phe	Leu	Pro	Ala	Ser	Asn	Ser	Glu	Ala	Phe
			100					105					110		

Phe	Asp	Gln	Ile	Gly	Asp	Leu	Arg	Ala	Arg	Ser	Ala	Ala	Ile	Pro	Asp
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

026172-005020PC - Sequence Listing_1

115

120

125

Asp Leu Leu Val Cys Leu Val Gly Asp Met Ile Thr Glu Glu Ala Leu
130 135 140

Pro Thr Tyr Met Ala Met Leu Asn Thr Leu Asp Val Val Arg Asp Glu
145 150 155 160

Thr Gly His Ser Gln His Pro Tyr Ala Lys Trp Thr Arg Ala Trp Ile
165 170 175

Ala Glu Glu Asn Arg His Gly Asp Leu Leu Asn Lys Tyr Met Trp Leu
180 185 190

Thr Gly Arg Val Gly His Ala Gly Gly Gly Ala His His Pro Ala Met
195 200 205

Leu Ala Val Glu Arg Thr Ile Gln Arg Leu Ile Ser Ser Gly Met Asp
210 215 220

Pro Gly Thr Glu Asn His Pro Tyr His Ala Phe Val Phe Thr Ser Phe
225 230 235 240

Gln Glu Arg Ala Thr Lys Leu Ser His Gly Ser Thr Ala Arg Leu Ala
245 250 255

Val Ala Ala Gly Asp Glu Ala Leu Ala Lys Ile Cys Gly Thr Ile Ala
260 265 270

Arg Asp Glu Ser Arg His Glu Ala Ala Tyr Thr Arg Thr Met Asp Ala
275 280 285

Ile Phe Gln Arg Asp Pro Ser Gly Ala Met Val Ala Phe Ala His Met
290 295 300

Met Met Arg Lys Ile Thr Met Pro Ala His Leu Met Asp Asp Gly Gln
305 310 315 320

His Gly Ala Arg Asn Gly Gly Ala Gln Leu Val Arg Arg Leu Cys Gly
325 330 335

Ser Gly Gly Ala Gly Arg Gly Val His Arg Arg Arg Leu His Arg His
340 345 350

Pro Ala Pro Pro His Pro Ala Leu Gly Arg Gly Gly
355 360

<210> 47

<211> 1395

<212> DNA

<213> Artificial sequence

026172-005020PC - Sequence Listing_1

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 47

attatacatc ggcacgtctt caggtttcac gatctgcatg ctatctatgg gactgtgact	60
ccgccggcca ggttgtggtg cgcgagaatc ctccccgctc ctgccttctc atttccctga	120
cgggagtcgc cgctgagcac cgggcggatc atgggcgctc gcacactcca aaccccatat	180
acatgtggtc gtgcattcac gcatagcgca cggatatgtc cgcgacgcgc ggctcgaagc	240
cgtggccatc cgacgcgctg cacggccgag gtgagggcac gcccctccgc caatggcgcg	300
cagcccatga ccgccttcga cttccggcag tacatgcagc agcgcgccgc gctggtggac	360
gcagcgctgg acctggcagt gccgctgcag taccctcgaga agatcaacga ggccatgcgg	420
tacagcctgc tggccggggg caagcgctg cgcctcgctc tctgcctcgc tgcctgcgag	480
ctcgtgggcg gccccttgga ggcgcccatg cccgcccctt gcgccatgga gatgatccac	540
accatgagcc tcatccacga cgacctcccc gccatggaca acgacgactt ccggcgcggc	600
cagcccgcga accacaaggc ctatggcgag gagattgcga tcctggcggg cgacgcgctg	660
ctgtcgctga gctttgagca catcgcgcg cgcgacgcgag gcgtggaccc ggtgcgcgctc	720
ctggccgcga tctcggagtg gcgcgcggtg ggcagccgcg ggctggtggc ggggcaggtg	780
gtggacctgg gtttcgaggg cggcgcgctg gggctggccc cgctgcgcta catccacgag	840
cacaaaaccg cggcgctgct ggaggcgcg gtggtgtccg gcgcgctgct gggcggcgcg	900
gaggaggcgg acctggagcg cctgcgcacc tacaaccgcg ccatcggcct cgctttccag	960
gtggtggggg acatcctgga catcccgggg accagcgagg agctgggcaa gaccgcgggc	1020
aaggacctga gtcccccaa aacccctac ccgtccctgg tggggctggc caggtccaaa	1080
aaaattgcgg acgaactgat tgaggacgcg aaaaccacac tcaccagta cgagccggcc	1140
cgagcggcgc cctcgtaac cctggccgaa aacatttgaa accggaagaa ctgactgggg	1200
gccccccctg cccccagata cggcggggct cctccatcca gttttgggat gggaggagcg	1260
acaaccgacc ccgtaaccct gtgacgcgtt tgccttgcat acgtacgcat gccttgaaac	1320
ccatccatga cctcaacaa tacctggttg tgtgtagctt ggtcctgaaa aaaaaaaaaa	1380
aaaaaaaaaa aaaaa	1395

<210> 48

<211> 342

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 48

Met	Gly	Val	Gly	Thr	Leu	Gln	Thr	Pro	Tyr	Thr	Cys	Gly	Arg	Ala	Phe
1				5				10						15	

026172-005020PC - Sequence Listing_1

Thr His Ser Ala Arg Tyr Val Pro Arg Arg Ala Ala Arg Ser Arg Gly
 20 25 30
 His Pro Thr Arg Cys Thr Ala Glu Val Arg Ala Arg Pro Ser Ala Asn
 35 40 45
 Gly Ala Gln Pro Met Thr Ala Phe Asp Phe Arg Gln Tyr Met Gln Gln
 50 55 60
 Arg Ala Ala Leu Val Asp Ala Ala Leu Asp Leu Ala Val Pro Leu Gln
 65 70 75 80
 Tyr Pro Glu Lys Ile Asn Glu Ala Met Arg Tyr Ser Leu Leu Ala Gly
 85 90 95
 Gly Lys Arg Val Arg Pro Ala Leu Cys Leu Ala Ala Cys Glu Leu Val
 100 105 110
 Gly Gly Pro Leu Glu Ala Ala Met Pro Ala Ala Cys Ala Met Glu Met
 115 120 125
 Ile His Thr Met Ser Leu Ile His Asp Asp Leu Pro Ala Met Asp Asn
 130 135 140
 Asp Asp Phe Arg Arg Gly Gln Pro Ala Asn His Lys Ala Tyr Gly Glu
 145 150 155 160
 Glu Ile Ala Ile Leu Ala Gly Asp Ala Leu Leu Ser Leu Ser Phe Glu
 165 170 175
 His Ile Ala Arg Glu Thr Arg Gly Val Asp Pro Val Arg Val Leu Ala
 180 185 190
 Ala Ile Ser Glu Trp Arg Ala Val Gly Ser Arg Gly Leu Val Ala Gly
 195 200 205
 Gln Val Val Asp Leu Gly Phe Glu Gly Gly Gly Val Gly Leu Ala Pro
 210 215 220
 Leu Arg Tyr Ile His Glu His Lys Thr Ala Ala Leu Leu Glu Ala Ala
 225 230 235 240
 Val Val Ser Gly Ala Leu Leu Gly Gly Ala Glu Glu Ala Asp Leu Glu
 245 250 255
 Arg Leu Arg Thr Tyr Asn Arg Ala Ile Gly Leu Ala Phe Gln Val Val
 260 265 270
 Gly Asp Ile Leu Asp Ile Pro Gly Thr Ser Glu Glu Leu Gly Lys Thr
 275 280 285

026172-005020PC - Sequence Listing_1

Ala Gly Lys Asp Leu Ser Ser Pro Lys Thr Pro Tyr Pro Ser Leu Val
290 295 300

Gly Leu Ala Arg Ser Lys Lys Ile Ala Asp Glu Leu Ile Glu Asp Ala
305 310 315 320

Lys Thr Gln Leu Thr Gln Tyr Glu Pro Ala Arg Ala Ala Pro Leu Val
325 330 335

Thr Leu Ala Glu Asn Ile
340

<210> 49
<211> 833
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polynucleotide

<400> 49
cagatgccat gcgccctcgg gccgcgggccc tgaggggtcca cgcagcgtcc tcggtggccc 60
agacgcacca ggccgcccc cggcgggaca ggaggttcga cgactaccag ccccgaccg 120
ccatcctctt ccccggccag ggcgcgcaga gcgtgggcat ggccggagag ctggcgaagg 180
ccgtccccgc cgccgcggcg ctgttcgacg ccgcctccga ccagctcggc tatgacctgc 240
tccgcgtgtg cgttgagggc cccaaggcgc gcctggacag caccgccgtc agccagcccg 300
ccatctacgt ggccagcctg gcggcggtgg agaagctgcg cgcggagggc ggggaggagg 360
cactggccgc catcgacgtc gctgccggtc tgctcttggg cgagtacacc gcgctggcct 420
ttgccggcgc cttctccttc gccgacgggc tgcgcctggt ggccctgcgc ggcgccagca 480
tgcaggccgc cgccgacgcc gcaccctcgg gcatgggtctc cgtcatcggc ctgccctccg 540
acgcggtggc cgcgctgtgc gagggcccca acgcgcaggc ggcccccgac caggccgtgc 600
gcatcgccaa ctacctctgc gacggcaact acgccgtcag cgggtgggctg gagggctgcg 660
cggcggtgga gggcctggcc aaggcccaaca aggcgcgcac gacggtgcgc ctggcggtgg 720
cgggcgcctt ccacaccccc ttcatgcagc cggcggtgga ggcgctgagc gcgggcgctg 780
gcggacacgc cgctgggtgc gccgcgcac cccgtggtca gcaacgggac gcc 833

<210> 50
<211> 275
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
polypeptide

<400> 50
Met Arg Pro Arg Ala Ala Gly Leu Arg Val His Ala Ala Ser Ser Val
1 5 10 15

026172-005020PC - Sequence Listing_1

Ala Gln Thr His Gln Ala Ala Pro Pro Ala Asp Arg Arg Phe Asp Asp
20 25 30

Tyr Gln Pro Arg Thr Ala Ile Leu Phe Pro Gly Gln Gly Ala Gln Ser
35 40 45

Val Gly Met Ala Gly Glu Leu Ala Lys Ala Val Pro Ala Ala Ala Ala
50 55 60

Leu Phe Asp Ala Ala Ser Asp Gln Leu Gly Tyr Asp Leu Leu Arg Val
65 70 75 80

Cys Val Glu Gly Pro Lys Ala Arg Leu Asp Ser Thr Ala Val Ser Gln
85 90 95

Pro Ala Ile Tyr Val Ala Ser Leu Ala Ala Val Glu Lys Leu Arg Ala
100 105 110

Glu Gly Gly Glu Glu Ala Leu Ala Ala Ile Asp Val Ala Ala Gly Leu
115 120 125

Ser Leu Gly Glu Tyr Thr Ala Leu Ala Phe Ala Gly Ala Phe Ser Phe
130 135 140

Ala Asp Gly Leu Arg Leu Val Ala Leu Arg Gly Ala Ser Met Gln Ala
145 150 155 160

Ala Ala Asp Ala Ala Pro Ser Gly Met Val Ser Val Ile Gly Leu Pro
165 170 175

Ser Asp Ala Val Ala Ala Leu Cys Glu Ala Ala Asn Ala Gln Val Ala
180 185 190

Pro Asp Gln Ala Val Arg Ile Ala Asn Tyr Leu Cys Asp Gly Asn Tyr
195 200 205

Ala Val Ser Gly Gly Leu Glu Gly Cys Ala Ala Val Glu Gly Leu Ala
210 215 220

Lys Ala His Lys Ala Arg Met Thr Val Arg Leu Ala Val Ala Gly Ala
225 230 235 240

Phe His Thr Pro Phe Met Gln Pro Ala Val Glu Ala Leu Ser Ala Gly
245 250 255

Ala Gly Gly His Ala Ala Gly Arg Ala Ala His Pro Arg Gly Gln Gln
260 265 270

Arg Asp Ala
275

026172-005020PC - Sequence Listing_1

<210> 51
 <211> 787
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 51
 tgtccatctc cccccaccct ccatccaacc atcgtcgacg gcatgcaggc gctgtgttct 60
 caccgccggt ccctcacggc gcgtgcggtta ccccatgggc gggccagccc agcacagcgg 120
 gtgtccagcg ccggcccggc ctacaccggc ctgtcccggc acaccctggg ctgccccagc 180
 acccccaccc tccagtcccg cgccgcggtc cagaccgcg gctcctctct cggtccacc 240
 acgcgcatga ccaccaccgc ccagcgcaag atcaaggtag ccatcaacgg gttcggccgc 300
 atcggccgcc agttcctgcg ctgctgtag gggcgcgagg actcgtgct ggagatcgtg 360
 gccgtgaacg actccggcgg cgtgaagcag gccagccacc tgctcaagta cgactccacc 420
 atgggcacct tcaacgccga catcaagatc tcgggcgagg gcaccttctc cgtcaacggc 480
 cgcgacatcc gcgtcgtctc ctcccgcgac cccctggccc tgccctgggg cgagctgggc 540
 gtggacctgg tgatcgaggg gacgggagtg tttgtggacc gcaagggtgc cagcaagcac 600
 ctgcaggcgg gggccaagaa ggtcatcatc accgcgccgg ccaagggtc cgacgtgccc 660
 acctacgtca tgggcgtgaa cgcggaccag tactccaact ccgacgacat catctccaac 720
 gcctcctgca ccaccaactg cctggcgccc tttgtcaagg tgctcaacga ccgcttcggc 780
 atcgtga 787

<210> 52
 <211> 248
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polypeptide

<400> 52
 Met Gln Ala Leu Cys Ser His Pro Ala Ser Leu Thr Ala Arg Ala Val
 1 5 10 15
 Pro His Gly Arg Ala Ser Pro Ala Gln Arg Val Ser Ser Ala Gly Pro
 20 25 30
 Ala Tyr Thr Gly Leu Ser Arg His Thr Leu Gly Cys Pro Ser Thr Pro
 35 40 45
 Thr Leu Gln Ser Arg Ala Ala Val Gln Thr Arg Gly Ser Ser Ser Gly
 50 55 60
 Ser Thr Thr Arg Met Thr Thr Thr Ala Gln Arg Lys Ile Lys Val Ala
 65 70 75 80

026172-005020PC - Sequence Listing_1

Ile Asn Gly Phe Gly Arg Ile Gly Arg Gln Phe Leu Arg Cys Val Glu
85 90 95

Gly Arg Glu Asp Ser Leu Leu Glu Ile Val Ala Val Asn Asp Ser Gly
100 105 110

Gly Val Lys Gln Ala Ser His Leu Leu Lys Tyr Asp Ser Thr Met Gly
115 120 125

Thr Phe Asn Ala Asp Ile Lys Ile Ser Gly Glu Gly Thr Phe Ser Val
130 135 140

Asn Gly Arg Asp Ile Arg Val Val Ser Ser Arg Asp Pro Leu Ala Leu
145 150 155 160

Pro Trp Gly Glu Leu Gly Val Asp Leu Val Ile Glu Gly Thr Gly Val
165 170 175

Phe Val Asp Arg Lys Gly Ala Ser Lys His Leu Gln Ala Gly Ala Lys
180 185 190

Lys Val Ile Ile Thr Ala Pro Ala Lys Gly Ser Asp Val Pro Thr Tyr
195 200 205

Val Met Gly Val Asn Ala Asp Gln Tyr Ser Asn Ser Asp Asp Ile Ile
210 215 220

Ser Asn Ala Ser Cys Thr Thr Asn Cys Leu Ala Pro Phe Val Lys Val
225 230 235 240

Leu Asn Asp Arg Phe Gly Ile Val
245

<210> 53

<211> 860

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 53

gatgttgaga atagtagctt gctgccttgt cgccatgcag agcgtgtgcg cgcagtcggt 60

ttcatgcaag ggggccttca cccagtcctt gcggaccccc cgatgcagca ggagccagct 120

cgtctgccgg gctgatggca aggccggagc cttcatcaag accgtaaaga gcggtgctgc 180

cgctctggct gcctccctcc tcctgtctgg ggggtgcgggc gcaactgacct ttgatgagct 240

gcagggcctg acctacctgc aggtgaaggg ctctggcatc gccaacacct gccccaccct 300

gtctggcggc tcctccaaca tcaaggacct gaagagcggg acctactccg tcaacaagat 360

026172-005020PC - Sequence Listing_1

gtgcctggag cccacgtcct tcaaggtcaa ggaggaggca cagttcaaga acggcgaggc	420
cgactttgtg cccaccaagc tcgtcacgcg tctgacctac accctggacg agatctctgg	480
ccagatgaag atcgacggca gcggcggcgt ggagttcaag gaggaggatg gcatcgacta	540
tgctgcagtc accgtgcagc ttccgggcgg ggagcgcgtg cccttcctct tcaccatcaa	600
ggagcttgac gccaaagggga ctgccgacgg cttcaagggc gagttcaccg tgccctccta	660
ccgtgggtcc tccttcctgg accccaaggg ccgcggcgcc tccaccggct acgacaacgc	720
cgtggccctg cccgccgccg gcgattccga ggagttggag aaggagaaca acaagtccac	780
caaggctctg aagggggagg ccatcttctc catcgccaag gtggacgccg ggacagggga	840
ggtggcgggc atctttgagt	860

<210> 54
 <211> 275
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polypeptide

<400> 54
 Met Gln Ser Val Cys Ala Gln Ser Val Ser Cys Lys Gly Ala Phe Thr
 1 5 10 15
 Gln Ser Leu Arg Thr Pro Arg Cys Ser Arg Ser Gln Leu Val Cys Arg
 20 25 30
 Ala Asp Gly Lys Ala Gly Ala Phe Ile Lys Thr Val Lys Ser Gly Ala
 35 40 45
 Ala Ala Leu Ala Ala Ser Leu Leu Leu Ser Gly Gly Ala Gly Ala Leu
 50 55 60
 Thr Phe Asp Glu Leu Gln Gly Leu Thr Tyr Leu Gln Val Lys Gly Ser
 65 70 75 80
 Gly Ile Ala Asn Thr Cys Pro Thr Leu Ser Gly Gly Ser Ser Asn Ile
 85 90 95
 Lys Asp Leu Lys Ser Gly Thr Tyr Ser Val Asn Lys Met Cys Leu Glu
 100 105 110
 Pro Thr Ser Phe Lys Val Lys Glu Glu Ala Gln Phe Lys Asn Gly Glu
 115 120 125
 Ala Asp Phe Val Pro Thr Lys Leu Val Thr Arg Leu Thr Tyr Thr Leu
 130 135 140
 Asp Glu Ile Ser Gly Gln Met Lys Ile Asp Gly Ser Gly Gly Val Glu
 145 150 155 160

026172-005020PC - Sequence Listing_1

Phe Lys Glu Glu Asp Gly Ile Asp Tyr Ala Ala Val Thr Val Gln Leu
165 170 175

Pro Gly Gly Glu Arg Val Pro Phe Leu Phe Thr Ile Lys Glu Leu Asp
180 185 190

Ala Lys Gly Thr Ala Asp Gly Phe Lys Gly Glu Phe Thr Val Pro Ser
195 200 205

Tyr Arg Gly Ser Ser Phe Leu Asp Pro Lys Gly Arg Gly Ala Ser Thr
210 215 220

Gly Tyr Asp Asn Ala Val Ala Leu Pro Ala Ala Gly Asp Ser Glu Glu
225 230 235 240

Leu Glu Lys Glu Asn Asn Lys Ser Thr Lys Ala Leu Lys Gly Glu Ala
245 250 255

Ile Phe Ser Ile Ala Lys Val Asp Ala Gly Thr Gly Glu Val Ala Gly
260 265 270

Ile Phe Glu
275

<210> 55

<211> 818

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 55

ataatcggaa cccagctgca cgcaccatca gtgcggcagc atgcagaccg tcgcagccag	60
ctatggcgta ttggcgccct ccggctccag cgtgacccgg ggctcgacca gcagcaagca	120
gcacttcacc accctcactc ccttttccgg cttcaggcgc ctgaatcatg tggatcgggc	180
ggggcaggcg ggggtctggga gccccagac cctgcagcag gccgtgggca aggccgtgcg	240
ccggtcgcgg ggccgcacca ccagcgccgt gcgcgtgacc cgcattgatgt ttgagcgggt	300
caccgagaag gccatcaagg tggcatgct cgcgcaggag gaggctcgcc gtctgggcca	360
caacttcgtg gggacggagc aaatcctgct ggggttgatt ggggagtcca caggcatcgc	420
cgccaaggtc ctcaagtcga tgggcgtcac gctgaaagat gcgcgtgtgg aggtcgagaa	480
gatcatcggc cgggggagcg gctttgtggc cgtggagatc cccttcaccc cccgcgcaa	540
gcgtgtgctg gagctgtccc tggaggaggc tcgccagctc ggccacaact acattggcac	600
ggagcacatc ctgctgggcc tgctgcgcga gggtaggggc gtggcctccc gcgtgctgga	660
gaccctgggc gccgacccc agaagatccg cactcagggtg gtacgcatgg tgggtgagtc	720
gcaggagccc gtgggcacca cggtagggcg aggggtccacc ggctccaaca agatgccac	780

cctggaggag tacggcacca acctgaccgc ccaggccg

818

<210> 56

<211> 259

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

<400> 56

Met Gln Thr Val Ala Ala Ser Tyr Gly Val Leu Ala Pro Ser Gly Ser
1 5 10 15Ser Val Thr Arg Gly Ser Thr Ser Ser Lys Gln His Phe Thr Thr Leu
20 25 30Thr Pro Phe Ser Gly Phe Arg Arg Leu Asn His Val Asp Arg Ala Gly
35 40 45Gln Ala Gly Ser Gly Ser Pro Gln Thr Leu Gln Gln Ala Val Gly Lys
50 55 60Ala Val Arg Arg Ser Arg Gly Arg Thr Thr Ser Ala Val Arg Val Thr
65 70 75 80Arg Met Met Phe Glu Arg Phe Thr Glu Lys Ala Ile Lys Val Val Met
85 90 95Leu Ala Gln Glu Glu Ala Arg Arg Leu Gly His Asn Phe Val Gly Thr
100 105 110Glu Gln Ile Leu Leu Gly Leu Ile Gly Glu Ser Thr Gly Ile Ala Ala
115 120 125Lys Val Leu Lys Ser Met Gly Val Thr Leu Lys Asp Ala Arg Val Glu
130 135 140Val Glu Lys Ile Ile Gly Arg Gly Ser Gly Phe Val Ala Val Glu Ile
145 150 155 160Pro Phe Thr Pro Arg Ala Lys Arg Val Leu Glu Leu Ser Leu Glu Glu
165 170 175Ala Arg Gln Leu Gly His Asn Tyr Ile Gly Thr Glu His Ile Leu Leu
180 185 190Gly Leu Leu Arg Glu Gly Glu Gly Val Ala Ser Arg Val Leu Glu Thr
195 200 205Leu Gly Ala Asp Pro Gln Lys Ile Arg Thr Gln Val Val Arg Met Val
210 215 220

026172-005020PC - Sequence Listing_1

Gly Glu Ser Gln Glu Pro Val Gly Thr Thr Val Gly Gly Gly Ser Thr
225 230 235 240

Gly Ser Asn Lys Met Pro Thr Leu Glu Glu Tyr Gly Thr Asn Leu Thr
245 250 255

Ala Gln Ala

<210> 57
<211> 2357
<212> DNA
<213> Artificial sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 57
ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120
gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc agggccccga ttgcaaagac 180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240
ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300
caacccgcaa acggcgcgcc atatcaatgc ttcttcaggc ctttcttttt cttcttgctg 360
gttttgctgc caagatcagc gcctctatga cgaacgaaac ctcggataga ccacttgctg 420
actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaaag 480
atgccaagtg gcatctgtac tttcaataca acccgaaacga tactgtctgg gggacgccat 540
tgttttgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta 600
tcgctccgaa gaggaacgac tccggagcat tctcgggttc catggtggtt gactacaaca 660
atacttccgg ctttttcaac gataccattg acccgagaca acgctgcgtg gccatatgga 720
cttacaacac accggagtcc gaggagcagt acatctcgta tagcctggac ggtggatata 780
cttttacaga gtatcagaag aaccctgtgc ttgctgcaaa ttcgactcag ttccgagatc 840
cgaaggctct ttggtacgag ccctcgaga agtggatcat gacagcggca aagtcacagg 900
actacaagat cgaaatttac tcgtctgacg accttaaate ctggaagctc gaatccgcgt 960
tcgcaaacga gggctttctc ggctaccaat acgaatgcc aggcctgata gaggtcccaa 1020
cagagcaaga tcccagcaag tcctactggg tgatgtttat ttccattaat ccaggagcac 1080
cggcaggagg ttcttttaat cagtacttcg tcggaagctt taacggaact catttcgagg 1140
catttgataa ccaatcaaga gtagttgatt ttggaaagga ctactatgcc ctgcagactt 1200
tcttcaatac tgacccgacc tatgggagcg ctcttgcatc tgcgtgggct tctaactggg 1260
agtattccgc attcgttcct acaaaccctt ggaggtcctc catgtcgcctc gtgaggaaat 1320

026172-005020PC - Sequence Listing_1

tctctctcaa cactgagtag caggccaacc cggaaccga actcataaac ctgaaagccg	1380
aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgcaacc aacaccacgt	1440
tgacgaaagc caacagctac aacgtcgatc tttcgaatag caccgggtaca cttgaatttg	1500
aactggtgta tgccgtcaat accacccaaa cgatctcgaa gtcggtgttc gcggacctct	1560
ccctctggtt taaaggcctg gaagaccccg aggagtacct cagaatgggt ttcgaggttt	1620
ctgcgtcctc cttcttcctt gatcgcgga acagcaaagt aaaatttggt aaggagaacc	1680
catattttac caacaggatg agcgttaaca accaaccatt caagagcgaa aacgacctgt	1740
cgtactacaa agtgtatggt ttgcttgatc aaaatatact ggaactctac ttcaacgatg	1800
gtgatgtcgt gtccaccaac acataacttca tgacaaccgg gaacgcactg ggctccgtga	1860
acatgacgac ggggtgtgat aacctgttct acatcgacaa attccagggtg agggaaagtca	1920
agtgagatct gtcgatcgac aagctcgagg cagcagcagc tcggatagta tcgacacact	1980
ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata	2040
tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc	2100
gagttgctag ctgcttgtgc tatttgcgaa taccacccc agcatccct tccctcgttt	2160
catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgtgc	2220
tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct	2280
ggtagtcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga	2340
acacaaatgg aaagctt	2357

<210> 58

<211> 2335

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 58

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg	60
cgctgcatgc aacaccgatg atgttcgac cccccgaagc tccttcgggg ctgcatgggc	120
gctccgatgc cgctccaggg cgagcgtgt taaatagcc agggccccga ttgcaaagac	180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg	240
ccactcgagc ttgtgatcgc actccgctaa gggggcgccct cttcctcttc gtttcagtca	300
caaccgcgaa acggcgcgcc atgtgtgtgc aggccttcct gttcctgctg gccggcttcg	360
ccgccaagat cagcgccctc atgacgaacg agacgtccga ccgccccctg gtgcacttca	420
cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag aaggacgcca	480
agtggcacct gtacttccag tacaaccgga acgacaccgt ctgggggacg cccttgttct	540
ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc gccatcgccc	600
cgaagcgcaa cgactccggc gccttctccg gctccatggg ggtggactac aacaacacct	660

026172-005020PC - Sequence Listing_1

ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc tggacctaca	720
acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc tacaccttca	780
ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc gacccgaagg	840
tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc caggactaca	900
agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc gcgttcgcca	960
acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc cccaccgagc	1020
aggacccccag caagtcctac tgggtgatgt tcatctccat caaccccggc gccccggccg	1080
gcggctcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc gaggccttcg	1140
acaaccagtc ccgctgggtg gacttcggca aggactacta cgccctgcag accttcttca	1200
acaccgaccc gacctacggg agcgccctgg gcatcgcgctg ggctccaac tgggagtact	1260
ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc aagttctccc	1320
tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag gccgagccga	1380
tcctgaacat cagcaacgcc ggcccttga gccggttcgc caccaacacc acgttgacga	1440
aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag ttcgagctgg	1500
tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac ctctccctct	1560
ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag gtgtccgcgt	1620
cctccttctt cctggaccgc gggaaacagca aggtgaagtt cgtgaaggag aaccctact	1680
tcaccaaccg catgagcgtg aacaaccagc ctttcaagag cgagaacgac ctgtcctact	1740
acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac gacggcgacg	1800
tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc gtgaacatga	1860
cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag gtcaagtgat	1920
taattaactc gaggcagcag cagctcggat agtatcgaca cactctggac gctggtcgtg	1980
tgatggactg ttgccgccac acttgctgcc ttgacctgtg aatatccctg ccgcttttat	2040
caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt	2100
gtgctatttg cgaataccac cccagcacc ccttccctc gtttcatatc gcttgcaccc	2160
caaccgcaac ttatctacgc tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac	2220
tgccccctgc acagccttgg ttggggctcc gcctgtattc tcctgggtact gcaacctgta	2280
aaccagcact gcaatgctga tgcacgggaa gtagtgggat gggaacacaa atgga	2335

<210> 59

<211> 382

<212> PRT

<213> Cinnamomum camphorum

<400> 59

Met	Ala	Thr	Thr	Ser	Leu	Ala	Ser	Ala	Phe	Cys	Ser	Met	Lys	Ala	Val
1				5					10					15	

026172-005020PC - Sequence Listing_1

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

Gln Leu Arg Ala Gly Asn Ala Gln Thr Ser Leu Lys Met Ile Asn Gly
35 40 45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Lys Leu Pro Asp Trp Ser
50 55 60

Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala Ala Glu Lys Gln
65 70 75 80

Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Asn Pro Pro Gln Leu Leu
85 90 95

Asp Asp His Phe Gly Pro His Gly Leu Val Phe Arg Arg Thr Phe Ala
100 105 110

Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr Ser Ile Val Ala
115 120 125

Val Met Asn His Leu Gln Glu Ala Ala Leu Asn His Ala Lys Ser Val
130 135 140

Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu Met Ser Lys Arg
145 150 155 160

Asp Leu Ile Trp Val Val Lys Arg Thr His Val Ala Val Glu Arg Tyr
165 170 175

Pro Ala Trp Gly Asp Thr Val Glu Val Glu Cys Trp Val Gly Ala Ser
180 185 190

Gly Asn Asn Gly Arg Arg His Asp Phe Leu Val Arg Asp Cys Lys Thr
195 200 205

Gly Glu Ile Leu Thr Arg Cys Thr Ser Leu Ser Val Met Met Asn Thr
210 215 220

Arg Thr Arg Arg Leu Ser Lys Ile Pro Glu Glu Val Arg Gly Glu Ile
225 230 235 240

Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp Glu Glu Ile Lys
245 250 255

Lys Pro Gln Lys Leu Asn Asp Ser Thr Ala Asp Tyr Ile Gln Gly Gly
260 265 270

Leu Thr Pro Arg Trp Asn Asp Leu Asp Ile Asn Gln His Val Asn Asn
275 280 285

026172-005020PC - Sequence Listing_1

Ile Lys Tyr Val Asp Trp Ile Leu Glu Thr Val Pro Asp Ser Ile Phe
290 295 300

Glu Ser His His Ile Ser Ser Phe Thr Ile Glu Tyr Arg Arg Glu Cys
305 310 315 320

Thr Met Asp Ser Val Leu Gln Ser Leu Thr Thr Val Ser Gly Gly Ser
325 330 335

Ser Glu Ala Gly Leu Val Cys Glu His Leu Leu Gln Leu Glu Gly Gly
340 345 350

Ser Glu Val Leu Arg Ala Lys Thr Glu Trp Arg Pro Lys Leu Thr Asp
355 360 365

Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Ser Ser Val
370 375 380

<210> 60

<211> 1240

<212> DNA

<213> Cinnamomum camphorum

<400> 60

ggcgcgccat ggccaccacc tccctggcct ccgccttctg cagcatgaag gccgtgatgc	60
tggcccgcga cggccgcggc atgaagcccc gctccagcga cctgcagctg cgcgccggca	120
acgcccagac ctccctgaag atgatcaacg gcaccaagtt ctctacacc gagagcctga	180
agaagctgcc cgactggtcc atgctgttcg ccgtgatcac caccatcttc tccgccgccg	240
agaagcagtg gaccaacctg gagtggaagc ccaagcccaa cccccccag ctgctggacg	300
accacttcgg ccccccacggc ctggtgttcc gccgcacctt cgccatccgc agctacgagg	360
tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag gaggccgccc	420
tgaaccacgc caagtccgtg ggcatcctgg gcgacggctt cggcaccacc ctggagatgt	480
ccaagcgcga cctgatctgg gtggtgaagc gcacccacgt ggccgtggag cgctacccccg	540
cctggggcga caccgtggag gtggagtgtt ggggtggcgc ctccggcaac aacggccgcc	600
gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgcg tgcacctccc	660
tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag gaggtgcgcg	720
gcgagatcgg ccccgcttc atcgacaacg tggccgtgaa ggacgaggag atcaagaagc	780
cccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc ccccgctgga	840
acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg atcctggaga	900
ccgtgcccga cagcatcttc gagagccacc acatctctc cttcaccatc gagtaccgcc	960
gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc ggctcctccg	1020
aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag gtgctgcgcg	1080
ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc gtgatccccg	1140

026172-005020PC - Sequence Listing_1

ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac caccacatcg 1200
actacaagga cgacgacgac aagtgactcg agttaattaa 1240

<210> 61
<211> 415
<212> PRT
<213> Cuphea hookeriana

<400> 61
Met Val Ala Ala Ala Ala Ser Ser Ala Phe Phe Pro Val Pro Ala Pro
1 5 10 15

Gly Ala Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu
20 25 30

Ser Pro Ser Phe Lys Pro Lys Ser Ile Pro Asn Gly Gly Phe Gln Val
35 40 45

Lys Ala Asn Asp Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Ser
50 55 60

Leu Lys Ser Gly Ser Leu Asn Thr Gln Glu Asp Thr Ser Ser Ser Pro
65 70 75 80

Pro Pro Arg Thr Phe Leu His Gln Leu Pro Asp Trp Ser Arg Leu Leu
85 90 95

Thr Ala Ile Thr Thr Val Phe Val Lys Ser Lys Arg Pro Asp Met His
100 105 110

Asp Arg Lys Ser Lys Arg Pro Asp Met Leu Val Asp Ser Phe Gly Leu
115 120 125

Glu Ser Thr Val Gln Asp Gly Leu Val Phe Arg Gln Ser Phe Ser Ile
130 135 140

Arg Ser Tyr Glu Ile Gly Thr Asp Arg Thr Ala Ser Ile Glu Thr Leu
145 150 155 160

Met Asn His Leu Gln Glu Thr Ser Leu Asn His Cys Lys Ser Thr Gly
165 170 175

Ile Leu Leu Asp Gly Phe Gly Arg Thr Leu Glu Met Cys Lys Arg Asp
180 185 190

Leu Ile Trp Val Val Ile Lys Met Gln Ile Lys Val Asn Arg Tyr Pro
195 200 205

Ala Trp Gly Asp Thr Val Glu Ile Asn Thr Arg Phe Ser Arg Leu Gly
210 215 220

Lys Ile Gly Met Gly Arg Asp Trp Leu Ile Ser Asp Cys Asn Thr Gly

026172-005020PC - Sequence Listing_1

225 230 235 240

Glu Ile Leu Val Arg Ala Thr Ser Ala Tyr Ala Met Met Asn Gln Lys
245 250 255

Thr Arg Arg Leu Ser Lys Leu Pro Tyr Glu Val His Gln Glu Ile Val
260 265 270

Pro Leu Phe Val Asp Ser Pro Val Ile Glu Asp Ser Asp Leu Lys Val
275 280 285

His Lys Phe Lys Val Lys Thr Gly Asp Ser Ile Gln Lys Gly Leu Thr
290 295 300

Pro Gly Trp Asn Asp Leu Asp Val Asn Gln His Val Ser Asn Val Lys
305 310 315 320

Tyr Ile Gly Trp Ile Leu Glu Ser Met Pro Thr Glu Val Leu Glu Thr
325 330 335

Gln Glu Leu Cys Ser Leu Ala Leu Glu Tyr Arg Arg Glu Cys Gly Arg
340 345 350

Asp Ser Val Leu Glu Ser Val Thr Ala Met Asp Pro Ser Lys Val Gly
355 360 365

Val Arg Ser Gln Tyr Gln His Leu Leu Arg Leu Glu Asp Gly Thr Ala
370 375 380

Ile Val Asn Gly Ala Thr Glu Trp Arg Pro Lys Asn Ala Gly Ala Asn
385 390 395 400

Gly Ala Ile Ser Thr Gly Lys Thr Ser Asn Gly Asn Ser Val Ser
405 410 415

<210> 62

<211> 1339

<212> DNA

<213> Cuphea hookeriana

<400> 62

ggcgcccat ggtggccgcc gccgcctcca gcgccttctt ccccggtgccc gcccccggcg	60
cctcccccaa gcccggaag ttcggcaact ggccctccag cctgagcccc tccttcaagc	120
ccaagtccat cccaacggc ggcttccagg tgaaggccaa cgacagcgcc caccccaagg	180
ccaacggctc cgccgtgagc ctgaagagcg gcagcctgaa caccaggag gacacctct	240
ccagcccccc ccccgccacc ttctgcacc agctgcccga ctggagccgc ctgctgaccg	300
ccatcaccac cgtgttcgtg aagtccaagc gcccgcacat gcacgaccgc aagtccaagc	360
gcccgcacat gctggtggac agcttcggcc tggagtcac cgtgcaggac ggcctggtgt	420
tccgccagtc cttctccatc cgctcctacg agatcggcac cgaccgcacc gccagcatcg	480

026172-005020PC - Sequence Listing_1

```

agaccctgat gaaccacctg caggagacct ccctgaacca ctgcaagagc accggcatcc      540
tgctggacgg cttcggccgc accctggaga tgtgcaagcg cgacctgata tgggtggtga      600
tcaagatgca gatcaaggtg aaccgctacc ccgcctgggg cgacaccgtg gagatcaaca      660
cccgcttcag ccgcctgggc aagatcggca tgggccgcga ctggctgata tccgactgca      720
acaccggcga gatcctggtg cgcgccacca gcgcctacgc catgatgaac cagaagaccc      780
gccgcctgtc caagctgccc tacgaggtgc accaggagat cgtgcccctg ttcgtggaca      840
gccccgtgat cgaggactcc gacctgaagg tgcacaagtt caaggtgaag accggcgaca      900
gcatccagaa gggcctgacc ccgggtgga acgacctgga cgtgaaccag cacgtgtcca      960
acgtgaagta catcggtctg atcctggaga gcatgccac cgaggtgctg gagaccagg      1020
agctgtgctc cctggccctg gactaccgcc gcgagtgcgg ccgcgactcc gtgctggaga      1080
gcgtgaccgc catggacccc agcaaggtgg gcgtgcgctc ccagtaccag cacctgctgc      1140
gcctggagga cggcaccgcc atcgtgaacg gcgccaccga gtggcgcccc aagaacgccg      1200
gcgccaacgg cgccatctcc accggcaaga ccagcaacgg caactccgtg tccatggact      1260
acaaggacca cgacggcgac tacaaggacc acgacatcga ctacaaggac gacgacgaca      1320
agtgactcga gttaattaa                                     1339

```

<210> 63
 <211> 382
 <212> PRT
 <213> Umbellularia sp.

<400> 63
 Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
 1 5 10 15
 Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
 20 25 30
 Gln Leu Arg Ala Gly Asn Ala Pro Thr Ser Leu Lys Met Ile Asn Gly
 35 40 45
 Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Arg Leu Pro Asp Trp Ser
 50 55 60
 Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala Ala Glu Lys Gln
 65 70 75 80
 Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Lys Leu Pro Gln Leu Leu
 85 90 95
 Asp Asp His Phe Gly Leu His Gly Leu Val Phe Arg Arg Thr Phe Ala
 100 105 110
 Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr Ser Ile Leu Ala
 115 120 125

026172-005020PC - Sequence Listing_1

Val Met Asn His Met Gln Glu Ala Thr Leu Asn His Ala Lys Ser Val
130 135 140

Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu Met Ser Lys Arg
145 150 155 160

Asp Leu Met Trp Val Val Arg Arg Thr His Val Ala Val Glu Arg Tyr
165 170 175

Pro Thr Trp Gly Asp Thr Val Glu Val Glu Cys Trp Ile Gly Ala Ser
180 185 190

Gly Asn Asn Gly Met Arg Arg Asp Phe Leu Val Arg Asp Cys Lys Thr
195 200 205

Gly Glu Ile Leu Thr Arg Cys Thr Ser Leu Ser Val Leu Met Asn Thr
210 215 220

Arg Thr Arg Arg Leu Ser Thr Ile Pro Asp Glu Val Arg Gly Glu Ile
225 230 235 240

Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp Asp Glu Ile Lys
245 250 255

Lys Leu Gln Lys Leu Asn Asp Ser Thr Ala Asp Tyr Ile Gln Gly Gly
260 265 270

Leu Thr Pro Arg Trp Asn Asp Leu Asp Val Asn Gln His Val Asn Asn
275 280 285

Leu Lys Tyr Val Ala Trp Val Phe Glu Thr Val Pro Asp Ser Ile Phe
290 295 300

Glu Ser His His Ile Ser Ser Phe Thr Leu Glu Tyr Arg Arg Glu Cys
305 310 315 320

Thr Arg Asp Ser Val Leu Arg Ser Leu Thr Thr Val Ser Gly Gly Ser
325 330 335

Ser Glu Ala Gly Leu Val Cys Asp His Leu Leu Gln Leu Glu Gly Gly
340 345 350

Ser Glu Val Leu Arg Ala Arg Thr Glu Trp Arg Pro Lys Leu Thr Asp
355 360 365

Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Pro Arg Val
370 375 380

<210> 64
<211> 1240

026172-005020PC - Sequence Listing_1

<212> DNA

<213> Umbellularia sp.

<400> 64

```

ggcgcgccat ggccaccacc agcctggcct ccgccttctg ctccatgaag gccgtgatgc      60
tggcccgcca cgcccgcggc atgaagcccc gcagctccga cctgcagctg cgcccgcca      120
acgcccccac ctccctgaag atgatcaacg gcaccaagtt cagctacacc gagagcctga      180
agcgccctgcc cgactggtcc atgctgttcg ccgtgatcac caccatcttc agcgccgccg      240
agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag ctgctggacg      300
accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc tcctacgagg      360
tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag gaggccaccc      420
tgaaccacgc caagagcgtg ggcatacctg gcgacggctt cggcaccacc ctggagatgt      480
ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag cgctacccca      540
cctggggcga caccgtggag gtggagtgtt ggatcggcgc cagcggcaac aacggcatgc      600
gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgcg tgcacctccc      660
tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac gaggtgcgcg      720
gcgagatcgg ccccgcttc atcgacaacg tggccgtgaa ggacgacgag atcaagaagc      780
tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc ccccgctgga      840
acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga      900
ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttaccctg gagtaccgcc      960
gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc ggcagctccg     1020
aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag gtgctgcgcg     1080
cccgacccga gtggcgcccc aagctgaccg actccttccg cggcatcagc gtgatccccg     1140
ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac cagcacatcg     1200
actacaagga cgacgacgac aagtgactcg agttaattaa                               1240

```

<210> 65

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

<400> 65

```

ccgccgtgct ggacgtggtg      20

```

<210> 66

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic primer

026172-005020PC - Sequence Listing_1

<400> 66
 ggtggcgggg tccaggggtgt 20

<210> 67
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 67
 cggccggcgg ctccttcaac 20

<210> 68
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 68
 ggcgctcccg taggtcgggt 20

<210> 69
 <211> 1335
 <212> DNA
 <213> *Chlorella sorokiniana*

<400> 69
 cgcctgcaac gcaagggcag ccacagccgc tcccacccgc cgctgaaccg acacgtgctt 60
 gggcgccctgc cgcctgcctg ccgcatgctt gtgctggtga ggctgggcag tgctgccatg 120
 ctgattgagg cttggttcac cgggtggaag cttatgtgtg tgctgggctt gcatgccggg 180
 caatgcgcat ggtggcaaga gggcggcagc acttgctgga gctgccgcgg tgcctccagg 240
 tggttcaatc gcggcagcca gagggatttc agatgatcgc gcgtacaggt tgagcagcag 300
 tgtcagcaaa ggtagcagtt tgccagaatg atcgggttcag ctgttaatca atgccagcaa 360
 gagaaggggt caagtgcaca cacgggcatg ccacagcacg ggcaccgggg agtggaatgg 420
 caccaccaag tgtgtgcgag ccagcatcgc cgcctggctg tttagctac aacggcagga 480
 gtcattcaac gtaacatga gctgatcaac actgcaatca tcgggcgggc gtgatgcaag 540
 catgcctggc gaagacacat ggtgtgcgga tgctgccggc tgctgcctgc tgcgcacgcc 600
 gttgagttgg cagcaggctc agccatgcac tggatggcag ctgggctgcc actgcaatgt 660
 ggtggatagg atgcaagtgg agcgaatacc aaaccctctg gctgcttgct gggttgcatg 720
 gcatgcgacc atcagcagga gcgcatgcga agggactggc cccatgcacg ccatgccaaa 780
 ccggagcgca ccgagtgtcc aactgtcac caggcccgcga agctttgcag aaccatgctc 840
 atggacgcat gtagcgctga cgtcccttga cggcgctcct ctgggtgtg ggaaacgcaa 900
 tgcagcacag gcagcagagg cggcggcagc agagcggcgg cagcagcggc gggggccacc 960

026172-005020PC - Sequence Listing_1

cttcttgcgg ggtcgcgccc cagccagcgg tgatgcgctg atcccaaacg agttcacatt	1020
catttgcatg cctggagaag cgaggctggg gcctttgggc tggcgcagcc cgcaatggaa	1080
tgccggaccg ccaggctagc agcaaaggcg cctcccctac tccgcatacg tgttccatag	1140
tgcatcggac tgcatttggg tggggcgggc ggctgtttct ttcgtgttgc aaaacgcgcc	1200
agctcagcaa cctgtcccgt ggggtccccg tgccgatgaa atcgtgtgca cgccgatcag	1260
ctgattgccc ggctcgcgaa gtaggcgccc tcctttctgc tcgccctctc tccgtcccg	1320
cactagtggc gcgcc	1335

<210> 70
 <211> 1146
 <212> DNA
 <213> Umbellularia californica

<400> 70	
atggccacca ccagcctggc ctccgccttc tgctccatga aggccgtgat gctggccccg	60
gacggccgcg gcatgaagcc ccgcagctcc gacctgcagc tgcgcgccgg caacgcccc	120
acctccctga agatgatcaa cggcaccaag ttcagctaca ccgagagcct gaagcgcctg	180
cccgactggg ccatgctgtt cgccgtgatc accaccatct tcagcgccgc cgagaagcag	240
tggaccaacc tggagtggaa gcccagccc aagctgcccc agctgctgga cgaccacttc	300
ggcctgcacg gcctgggtgtt ccgccgcacc ttcgccatcc gctcctacga ggtgggcccc	360
gaccgcagca cctccatcct ggccgtgatg aaccacatgc aggaggccac cctgaaccac	420
gccaagagcg tgggcatcct gggcgacggc ttcggcacca ccctggagat gtccaagcgc	480
gacctgatgt ggggtggcgc ccgcacccac gtggccgtgg agcgctaccc cacctggggc	540
gacaccgtgg aggtggagtg ctggatcggc gccagcggca acaacggcat gcgccgcgac	600
ttcttggcgc gcgactgcaa gaccggcgag atcctgaccc gctgcacctc cctgagcgtg	660
ctgatgaaca cccgcacccg ccgcctgagc accatccccg acgaggtgcg cggcgagatc	720
ggccccgcct tcacgacaa cgtggccgtg aaggacgacg agatcaagaa gctgcagaag	780
ctgaacgact ccaccgccga ctacatccag ggcggcctga cccccgctg gaacgacctg	840
gacgtgaacc agcacgtgaa caacctgaag tacgtggcct ggggtgttcga gaccgtgccc	900
gacagcatct tcgagtccca ccacatcagc tccttcaccc tggagtaccg ccgcgagtgc	960
acccgcgact ccgtgctgcg cagcctgacc accgtgagcg gcggcagctc cgaggccggc	1020
ctggtgtgcg accacctgct gcagctggag ggcggcagcg aggtgctgcg cggccgcacc	1080
gagtggcgcc ccaagctgac cgactccttc cgcggcatca gcgtgatccc cgccgagccc	1140
cgcgctg	1146

<210> 71
 <211> 1146
 <212> DNA
 <213> Cinnamomum camphora

026172-005020PC - Sequence Listing_1

<400> 71
 atggccacca cctccctggc ctccgccttc tgcagcatga aggccgtgat gctggcccgc 60
 gacggccgcg gcatgaagcc ccgctccagc gacctgcagc tgcgcgccgg caacgcccag 120
 acctccctga agatgatcaa cggcaccaag ttctcctaca ccgagagcct gaagaagctg 180
 cccgactggt ccatgctggt cgccgtgatc accaccatct tctccgccgc cgagaagcag 240
 tggaccaacc tggagtggaa gcccagccc aacccccccc agctgctgga cgaccacttc 300
 ggccccacg gcctggtggt ccgccgcacc ttcgccatcc gcagctacga ggtgggcccc 360
 gaccgctcca ccagcatcgt ggccgtgatg aaccacctgc aggaggccgc cctgaaccac 420
 gccaagtccg tgggcatcct gggcgacggc ttccggacca ccctggagat gtccaagcgc 480
 gacctgatct ggggtggtgaa ggcacccac gtggccgtgg agcgctaccc cgcctggggc 540
 gacaccgtgg aggtggagtg ctgggtgggc gcctccggca acaacggccg ccgccacgac 600
 ttcttggtgc gcgactgcaa gaccggcgag atcctgaccc gctgcacctc cctgagcgtg 660
 atgatgaaca ccgcacccg ccgcctgagc aagatccccg aggaggtgcg cggcgagatc 720
 ggccccgcct tcatcgacaa cgtggccgtg aaggacgagg agatcaagaa gccccagaag 780
 ctgaacgact ccaccgccga ctacatccag ggcggcctga cccccgctg gaacgacctg 840
 gacatcaacc agcacgtgaa caacatcaag tacgtggact ggatcctgga gaccgtgccc 900
 gacagcatct tcgagagcca ccacatctcc tccttcacca tcgagtaccg ccgcgagtgc 960
 accatggaca gcgtgctgca gtccctgacc accgtgagcg gcggctcctc cgaggccggc 1020
 ctggtgtgcg agcacctgct gcagctggag ggcggcagcg aggtgctgcg cgccaagacc 1080
 gagtggcgcc ccaagctgac cgactccttc cgcgcatca gcgtgatccc cgccgagtcc 1140
 agcgtg 1146

<210> 72
 <211> 72
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic
 oligonucleotide

<400> 72
 atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 60
 gacgacaagt ga 72

<210> 73
 <211> 408
 <212> DNA
 <213> *Chlorella vulgaris*

<400> 73
 ctcgaggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga 60
 ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag 120
 cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat 180

026172-005020PC - Sequence Listing_1

ttgcgaatac cccccccagc atcccccttcc ctcgtttcat atcgcttgca tcccaaccgc	240
aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgcccct	300
cgcacagcct tggtttgggc tccgcctgta ttctcctggg actgcaacct gtaaaccagc	360
actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatgga	408

<210> 74
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 74 ctgggcgacg gcttcggcac	20
-----------------------------------	----

<210> 75
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 75 aagtcgcggc gcatgccgtt	20
-----------------------------------	----

<210> 76
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 76 taccgccct ggggcgacac	20
----------------------------------	----

<210> 77
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic primer

<400> 77 cttgctcagg cggcgggtgc	20
-----------------------------------	----

<210> 78
 <211> 1317
 <212> DNA
 <213> Artificial Sequence

<220>

026172-005020PC - Sequence Listing_1

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 78

```

atggtggccg ccgcccctc cagcgccttc ttccccgtgc ccgcccccg cgcctcccc 60
aagcccggca agttcggcaa ctggccctcc agcctgagcc ctccttcaa gccaagtcc 120
atccccaacg gcggcttcca ggtgaaggcc aacgacagcg cccaccccaa ggccaacggc 180
tccgccgtga gcctgaagag cggcagcctg aacacccagg aggacacctc ctccagcccc 240
cccccccgca ctttctgca ccagctgccc gactggagcc gcctgctgac cgccatcacc 300
accgtgttcg tgaagtccaa gcgccccgac atgcacgacc gcaagtccaa gcgccccgac 360
atgctgggtg acagcttcgg cctggagtcc accgtgcagg acggcctggg gttccgccag 420
tcctttctca tccgctccta cgagatcggc accgaccgca ccgccagcat cgagaccctg 480
atgaaccacc tgcaggagac ctccctgaac cactgcaaga gcaccggcat cctgctggac 540
ggcttcggcc gcaccctgga gatgtgcaag cgcgacctga tctgggtggg gatcaagatg 600
cagatcaagg tgaaccgcta ccccgccctg ggcgacaccg tggagatcaa caccgccttc 660
agccgcctgg gcaagatcgg catgggccgc gactggctga tctccgactg caacaccggc 720
gagatcctgg tgcgcgccac cagcgcctac gccatgatga accagaagac ccgcccctg 780
tccaagctgc cctacgaggt gcaccaggag atcgtgcccc tgttcgtgga cagccccgtg 840
atcgaggact ccgacctgaa ggtgcacaag ttcaaggatga agaccggcga cagcatccag 900
aagggcctga ccccggtg gaacgacctg gacgtgaacc agcacgtgtc caacgtgaag 960
tacatcggct ggatcctgga gagcatgccc accgaggtgc tggagaccca ggagctgtgc 1020
tccctggccc tggagtaccg ccgcgagtgc ggccgcgact ccgtgctgga gagcgtgacc 1080
gccatggacc ccagcaaggt gggcgtgcgc tcccagtacc agcacctgct gcgcctggag 1140
gacggcaccg ccatcgtgaa cggcgccacc gagtggcgcc ccaagaacgc cggcgccaac 1200
ggcgccatct ccaccggcaa gaccagcaac ggcaactccg tgtccatgga ctacaaggac 1260
cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtga 1317

```

<210> 79

<211> 1170

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 79

```

atggccaccg catccacttt ctggcgcttc aatgcccgt gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cggccagctg 120
cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcccc 180
gacatgcacg accgcaagtc caagcgcccc gacatgctgg tggacagctt cggcctggag 240
tccaccgtgc aggacggcct ggtgttccgc cagtccttct ccatccgctc ctacgagatc 300

```

026172-005020PC - Sequence Listing_1

ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctccctg	360
aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcaccct ggagatgtgc	420
aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc	480
tggggcgaca ccgtggagat caacacccgc ttcagccgcc tgggcaagat cggcatgggc	540
cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tgggtgcgcgc caccagcgcc	600
tacgccatga tgaaccagaa gaccgcgcgc ctgtccaagc tgccctacga ggtgcaccag	660
gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac	720
aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac	780
ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg	840
cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgcgag	900
tgccggccgc actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggcgtg	960
cgctcccagt accagcacct gctgcgcctg gaggacggca ccgccatcgt gaacggcgcc	1020
accgagtggc gcccgaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc	1080
aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactacaa ggaccacgac	1140
atcgactaca aggacgacga cgacaagtga	1170

<210> 80

<211> 1170

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 80

atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg	60
cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttcgtgggcg cgcccagctg	120
cccgactgga gccgcctgct gaccgccatc accaccgtgt tcgtgaagtc caagcgcccc	180
gacatgcacg accgcaagtc caagcgcccc gacatgctgg tggacagctt cggcctggag	240
tccaccgtgc aggacggcct ggtgttcgcg cagtccttct ccatccgctc ctacgagatc	300
ggcaccgacc gcaccgccag catcgagacc ctgatgaacc acctgcagga gacctccctg	360
aaccactgca agagcaccgg catcctgctg gacggcttcg gccgcaccct ggagatgtgc	420
aagcgcgacc tgatctgggt ggtgatcaag atgcagatca aggtgaaccg ctaccccgcc	480
tggggcgaca ccgtggagat caacacccgc ttcagccgcc tgggcaagat cggcatgggc	540
cgcgactggc tgatctccga ctgcaacacc ggcgagatcc tgggtgcgcgc caccagcgcc	600
tacgccatga tgaaccagaa gaccgcgcgc ctgtccaagc tgccctacga ggtgcaccag	660
gagatcgtgc ccctgttcgt ggacagcccc gtgatcgagg actccgacct gaaggtgcac	720
aagttcaagg tgaagaccgg cgacagcatc cagaagggcc tgacccccgg ctggaacgac	780

026172-005020PC - Sequence Listing_1

ctggacgtga accagcacgt gtccaacgtg aagtacatcg gctggatcct ggagagcatg	840
cccaccgagg tgctggagac ccaggagctg tgctccctgg ccctggagta ccgccgcgag	900
tgcgggccgcg actccgtgct ggagagcgtg accgccatgg accccagcaa ggtgggcgtg	960
cgctcccagt accagcacct gctgcgcctg gaggacggca ccgccatcgt gaacggcgcc	1020
accgagtggc gcccgaagaa cgccggcgcc aacggcgcca tctccaccgg caagaccagc	1080
aacggcaact ccgtgtccat ggactacaag gaccacgacg gcgactacaa ggaccacgac	1140
atcgactaca aggacgacga cgacaagtga	1170

<210> 81
 <211> 1167
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 81	
atgacgttcg gggtcgccct cccggccatg ggccgcggtg tctcccttcc ccggcccagg	60
gtcgcggtgc gcgccagtc ggcgagtcag gttttggaga gcgggcgcgcc ccagctgccc	120
gactggagcc gcctgctgac cgccatcacc accgtgttcg tgaagtccaa gcgccccgac	180
atgcacgacc gcaagtccaa gcgccccgac atgctggtgg acagcttcgg cctggagtcc	240
accgtgcagg acggcctggt gttccgccag tccttctcca tccgctccta cgagatcggc	300
accgaccgca ccgccagcat cgagaccctg atgaaccacc tgcaggagac ctccctgaac	360
cactgcaaga gcaccggcat cctgctggac ggcttcggcc gcaccctgga gatgtgcaag	420
cgcgacctga tctgggtggt gatcaagatg cagatcaagg tgaaccgcta ccccgccctgg	480
ggcgacaccg tggagatcaa caccgccttc agccgcctgg gcaagatcgg catgggcccgc	540
gactggctga tctccgactg caacaccggc gagatcctgg tgcgcgccac cagcgcctac	600
gccatgatga accagaagac ccgccgcctg tccaagctgc cctacgaggt gcaccaggag	660
atcgtgcccc tgttcgtgga cagccccgtg atcgaggact ccgacctgaa ggtgcacaag	720
ttcaaggtga agaccggcga cagcatccag aagggcctga cccccggctg gaacgacctg	780
gacgtgaacc agcacgtgtc caacgtgaag tacatcggtt ggatcctgga gagcatgccc	840
accgaggtgc tggagaccca ggagctgtgc tccctggccc tggagtaccg ccgcgagtgc	900
ggccgcgact ccgtgctgga gagcgtgacc gccatggacc ccagcaaggt gggcgtgcgc	960
tcccagtacc agcacctgct ggcctggag gacggcaccg ccacgtgaa cggcgccacc	1020
gagtggcgcc ccaagaacgc cggcgccaac ggcgccatct ccaccggcaa gaccagcaac	1080
ggcaactccg tgtccatgga ctacaaggac cagcagggcg actacaagga ccacgacatc	1140
gactacaagg acgacgacga caagtga	1167

<210> 82
 <211> 1149

026172-005020PC - Sequence Listing_1

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 82

```

atgacgttcg gggtcgccct cccggccatg ggccgcggtg tctcccttcc ccggcccagg      60
gtcgcggtgc gcgccagtc ggcgagtcag gttttggaga gcgggcgcgc ccccgactgg      120
tccatgctgt tcgccgtgat caccaccatc ttacgcgcgc ccgagaagca gtggaccaac      180
ctggagtgga agcccaagcc caagctgccc cagctgctgg acgaccactt cggcctgcac      240
ggcctggtgt tcgcgcgcac cttcgccatc cgctcctacg aggtgggccc cgaccgcagc      300
acctccatcc tggcgcgtgat gaaccacatg caggaggcca ccctgaacca cgccaagagc      360
gtgggcatcc tgggcgcagg cttcggcacc accctggaga tgtccaagcg cgacctgatg      420
tgggtggtgc gccgcaccca cgtggccgtg gagcgctacc ccacctgggg cgacaccgtg      480
gaggtggagt gctggatcgg cgccagcggc aacaacggca tgcgccgcga cttcctggtg      540
cgcgactgca agaccggcga gatactgacc cgctgcacct ccctgagcgt gctgatgaac      600
acccgcaccc gccgcctgag caccatcccc gacgaggtgc gcggcgagat cggccccgcc      660
ttcatcgaca acgtggccgt gaaggacgac gagatcaaga agctgcagaa gctgaacgac      720
tccaccgccc actacatcca gggcggcctg acccccgct ggaacgacct ggacgtgaac      780
cagcacgtga acaacctgaa gtacgtggcc tgggtgttcg agaccgtgcc cgacagcatc      840
ttcgagtccc accacatcag ctctttcacc ctggagtacc gccgcgagtg caccgcgcac      900
tccgtgctgc gcagcctgac caccgtgagc ggcggcagct ccgaggccgg cctggtgtgc      960
gaccacctgc tgcagctgga gggcggcagc gaggtgctgc gcgcccgcac cgagtggcgc      1020
cccaagctga ccgactcctt ccgcggcatc agcgtgatcc ccgccgagcc ccgcgtgatg      1080
gactacaagg accacgacgg cgactacaag gaccacgaca tcgactacaa ggacgacgac      1140
gacaagtga                                     1149

```

<210> 83

<211> 1146

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 83

```

atggcttcg cgccattcac catgtcggcg tgccccgcga tgactggcag ggcccctggg      60
gcacgtcgct ccggacggcc agtcgccacc cgcttgaggg ggccgcgccc cgactggctc      120
atgctgttcg ccgtgatcac caccatcttc agcgcgcgcg agaagcagtg gaccaacctg      180
gagtggaagc ccaagcccaa gctgccccag ctgctggacg accacttcgg cctgcacggc      240
ctggtgttcc gccgcacctt cgccatccgc tcctacgagg tgggccccga ccgcagcacc      300

```

026172-005020PC - Sequence Listing_1

tccatcctgg ccgtgatgaa ccacatgcag gaggccaccc tgaaccacgc caagagcgtg	360
ggcatcctgg gcgacggctt cggcaccacc ctggagatgt ccaagcgcga cctgatgtgg	420
gtggtgcgcc gcacccacgt ggccgtggag cgctacccca cctggggcga caccgtggag	480
gtggagtgtt ggatcggcgc cagcggcaac aacggcatgc gccgcgactt cctggtgcgc	540
gactgcaaga ccggcgagat cctgaccgcg tgcacctccc tgagcgtgtt gatgaacacc	600
cgcacccgcc gcctgagcac catccccgac gaggtgcgcg gcgagatcgg ccccgccctt	660
atcgacaacg tggccgtgaa ggacgacgag atcaagaagc tgcagaagct gaacgactcc	720
accgccgact acatccaggg cggcctgacc ccccgctgga acgacctgga cgtgaaccag	780
cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga ccgtgcccga cagcatcttc	840
gagtcccacc acatcagctc ctccacctg gagtaccgcc gcgagtgac ccgcgactcc	900
gtgtgcgcga gcctgaccac cgtgagcggc ggcagctccg aggccggcct ggtgtgcgac	960
cacctgtgtc agctggaggg cggcagcgag gtgtgcgcg cccgcaccga gtggcgcccc	1020
aagctgaccg actccttcct cggcatcagc gtgatccccg ccgagccccg cgtgatggac	1080
tacaaggacc acgacggcga ctacaaggac cagcagatcg actacaagga cgacgacgac	1140
aagtga	1146

<210> 84

<211> 1155

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 84

atggccaccg catccacttt ctggcggttc aatgcccgt gcggcgacct gcgtcgctcg	60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg ccccccgac	120
tgggtccatgc tgttcgccgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc	180
aacctggagt ggaagcccaa gcccagctg cccagctgc tggacgacca cttcggcctg	240
cacggcctgg tgttcgcgcg caccttcgcc atccgtctct acgaggtggg ccccgaccgc	300
agcacctcca tcctggccgt gatgaaccac atgcaggagg ccacctgaa ccacgccaag	360
agcgtgggca tcctgggcga cggcttcggc accaccttg agatgtccaa gcgcgacctg	420
atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgtt accccacctg gggcgacacc	480
gtggagggtg agtgctggat cggcgccagc ggcaacaacg gcatgcgccg cgacttcctg	540
gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgtgatg	600
aacacccgca cccgccgctt gagcaccatc cccgacgagg tgcgcggcga gatcggtccc	660
gccttcacgc acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac	720
gactccaccg ccgactacat ccaggggcgc ctgaccccc cgtggaacga cctggacgtg	780

026172-005020PC - Sequence Listing_1

aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgcacagc	840
atcttcgagt cccaccacat cagctccttc accctggagt accgccgcga gtgcacccgc	900
gactccgtgc tgcgcagcct gaccaccgtg agcggcgcca gctccgaggc cggcctggtg	960
tgcgaccacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgcccgc caccgagtgg	1020
cgccccaaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gccccgcgtg	1080
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac	1140
gacgacaagt gatga	1155

<210> 85
 <211> 1152
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 85	
atggctatca agacgaacag gcagcctgtg gagaagcctc cgttcacgat cgggacgctg	60
cgcaaggcca tccccgcgca ctgtttcgag cgctcggcgc ttcgtgggcg ccccccgac	120
tggtccatgc tggtcggcgt gatcaccacc atcttcagcg ccgccgagaa gcagtggacc	180
aacctggagt ggaagcccaa gcccagctg cccagctgc tggacgacca cttcggcctg	240
cacggcctgg tggtccgccc caccctcgcc atccgctcct acgaggtggg ccccgaccgc	300
agcacctcca tcctggccgt gatgaaccac atgcaggagg ccaccctgaa ccacgccaag	360
agcgtgggca tcctgggcga cggcttcggc accaccctgg agatgtccaa gcgcgacctg	420
atgtgggtgg tgcgccgcac ccacgtggcc gtggagcgct accccacctg gggcgacacc	480
gtggaggtgg agtgctggat cggcgccagc ggcaacaacg gcatgcgccg cgacttcctg	540
gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgctgatg	600
aacaccgca cccgccgcct gagcaccatc cccgacgagg tgcgcggcga gatcggcccc	660
gccttcacatg acaacgtggc cgtgaaggac gacgagatca agaagctgca gaagctgaac	720
gactccaccg ccgactacat ccagggcggc ctgaccccc gctggaacga cctggacgtg	780
aaccagcacg tgaacaacct gaagtacgtg gcctgggtgt tcgagaccgt gcccgcacagc	840
atcttcgagt cccaccacat cagctccttc accctggagt accgccgcga gtgcacccgc	900
gactccgtgc tgcgcagcct gaccaccgtg agcggcgcca gctccgaggc cggcctggtg	960
tgcgaccacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgcccgc caccgagtgg	1020
cgccccaaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gccccgcgtg	1080
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac	1140
gacgacaagt ga	1152

<210> 86
 <211> 1155

026172-005020PC - Sequence Listing_1

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 86

```

atggccaccg catccacttt ctggtcggtt aatgcccgtt gcggcgacct gcgtcgctcg      60
gcgggctccg gggcccggtg cccagcgagg cccctccccg tgcgcgggctg cggcccgac      120
tggtccatgc tgttcgctgt gatcaccacc atctttctcc cgcgcgagaa gcagtggacc      180
aacctggagt ggaagcccaa gcccaccccc cccagctgct tggacgacca cttcggtccc      240
cacggcctgg tgttcgctgt caccttcgct atccgcagct acgaggtggg ccccgaccgc      300
tccaccagca tcgtggcctg gatgaaccac ctgcaggagg ccgcccctga ccacgccaag      360
tccgtgggca tcctgggctg cggcttcggc accaccctgg agatgtccaa gcgcgacctg      420
atctgggtgg tgaagcgac ccacgtggcc gtggagcgct accccgctg gggcgacacc      480
gtggaggtgg agtgcctggg gggcgccctc ggcaacaacg gccgcgcgca cgacttcctg      540
gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgatgatg      600
aacacccgca cccgcgcctt gagcaagatc cccgaggagg tgcgcggcga gatcggtccc      660
gccttcacat acaacgtggc cgtgaaggac gaggagatca agaagcccca gaagctgaac      720
gactccaccg ccgactacat ccagggcggt ctgacccccg gctggaacga cctggacatc      780
aaccagcacg tgaacaacat caagtacgtg gactggatcc tggagaccgt gcccgacagc      840
atcttcgaga gccaccacat ctctctcttc accatcgagt accgcgcgca gtgcaccatg      900
gacagcgtgc tgcagtcctt gaccaccgtg agcggcggtt cctccgaggc cggcctggtg      960
tgcgagcacc tgctgcagct ggagggcggt agcgaggtgc tgcgcgcca gaccgagtgg      1020
cgccccaagc tgaccgactc cttccgctgt atcagcgtga tccccgcca gtccagcgtg      1080
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac      1140
gacgacaagt gatga                                     1155

```

<210> 87

<211> 1893

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 87

```

gaattccttt cttgcgtat gacacttcca gcaaaaggta gggcgggctg cgagacggct      60
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc      120
atgggctgtc cgatgccgtt ccagggcgag cgctgtttta atagccaggc ccccgattgc      180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctatgactact accacttcta      240
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt      300

```

026172-005020PC - Sequence Listing_1

```

cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc      360
cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gagggccctc      420
cccgtgcgcg ggcgcgcccc cgactgggtcc atgctgttcg ccgtgatcac caccatcttc      480
tccgccgccc agaagcagtg gaccaacctg gagtggaagc ccaagcccaa cccccccag      540
ctgctggacg accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc      600
agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag      660
gaggccgccc tgaaccacgc caagtccgtg ggcacctctg gcgacggctt cggcaccacc      720
ctggagatgt ccaagcgcga cctgatctgg gtggtgaagc gcacccacgt ggccgtggag      780
cgctaccccc cctggggcga caccgtggag gtggagtgtt ggggtgggcgc ctccggcaac      840
aacggccgccc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgcg      900
tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag      960
gaggtgcgcg gcgagatcgg ccccgcttc atcgacaacg tggccgtgaa ggacgaggag     1020
atcaagaagc ccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc     1080
ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg     1140
atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcttc cttcaccatc     1200
gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc     1260
ggctcctccg aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag     1320
gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc     1380
gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac     1440
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg     1500
atagtatcga cacactctgg acgctgggtc tgtgatggac tgttgccgcc acacttgctg     1560
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt     1620
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcaataacc acccccagca     1680
tccccctccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc     1740
tatccctcag cgctgctcct gctcctgctc actgccccct gcacagcctt ggtttgggct     1800
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg     1860
aagtagtggg atgggaacac aaatggaaag ctt                                     1893

```

<210> 88

<211> 1887

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 88

```

gaattccttt cttgcgtat gacacttcca gcaaaaggta gggcgggctg cgagacggct      60

```

026172-005020PC - Sequence Listing_1

tccccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcgggggctgc	120
atggggcgctc cgatgccgct ccagggcgag cgctgtttta atagccaggc ccccgattgc	180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	240
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	300
cagtcacaac ccgcaaacac tagtatggct tccgcggcat tcacccatgtc ggcgtgcccc	360
gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc caccgcctg	420
agggggcgcg ccccgactg gtccatgtg ttcgccgtga tcaccaccat cttctccgc	480
gccgagaagc agtggaccaa cctggagtgg aagcccaagc ccaaccccc ccagctgctg	540
gacgaccact tcggccccca cggcctggtg ttccgcccga ctttcgccat ccgcagctac	600
gaggtgggccc ccgaccgctc caccagcatc gtggccgtga tgaaccacct gcaggaggcc	660
gccctgaacc acgccaagtc cgtgggcatc ctgggcgacg gcttcggcac caccctggag	720
atgtccaagc gcgacctgat ctgggtggtg aagcgcaccc acgtggccgt ggagcgctac	780
cccgcctggg gcgacaccgt ggaggtggag tgctgggtgg gcgcctccgg caacaacggc	840
cgccgccacg acttcctggt gcgcgactgc aagaccggcg agatcctgac ccgctgcacc	900
tccctgagcg tgatgatgaa caccgcacc cgccgcctga gcaagatccc cgaggagggtg	960
cgcgccgaga tcggccccgc cttcatcgac aacgtggccg tgaaggacga ggagatcaag	1020
aagccccaga agctgaacga ctccaccgcc gactacatcc agggcggcct gacccccgc	1080
tggaacgacc tggacatcaa ccagcacgtg aacaacatca agtacgtgga ctggatcctg	1140
gagaccgtgc ccgacagcat cttcgagagc caccacatct cctccttcac catcgagtac	1200
cgccgcgagt gcacccatgga cagcgtgctg cagtccctga ccaccgtgag cggcggctcc	1260
tccgaggccg gcctggtgtg cgagcacctg ctgcagctgg agggcggcag cgagggtgctg	1320
cgcgccaaga ccgagtggcg ccccaagctg accgactcct tccgcggcat cagcgtgac	1380
cccgccgagt ccagcgtgat ggactacaag gaccacgacg gcgactacaa ggaccacgac	1440
atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta	1500
tcgacacact ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga	1560
cctgtgaata tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac	1620
gcgcttttgc gagggtgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccc	1680
tccctcgttt catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc	1740
tcagcgctgc tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct	1800
gtattctcct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag	1860
tgggatggga acacaaatgg aaagctt	1887

<210> 89

<211> 3631

<212> DNA

<213> Artificial Sequence

026172-005020PC - Sequence Listing_1

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 89

gaattccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgctgccgc atgcttgctg tggtagaggct gggcagtgct	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccagggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgctg acagggttag	300
cagcagtgct agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgtcatgag acgcatgtag cgctgacgtc ccttgacggc gtcctctctg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccga	1080
atggaatgag ggaccgccag gctagcagca aaggcgccct ccctactccg catcgatggt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgcccctct ttctgctcgc cctctctccg	1320
tcccgccact agtatgctgc tgcaggcctt cctgttctctg ctggccggct tcgccgcaa	1380
gatcagcgcc tccatgacga acgagacgtc cgaccgcccc ctggtgcact tcacccccaa	1440
caagggtctg gggcgcgcca gccaccacgt gtacaagcgc ctgaccaga gcaccaacac	1500
caagtcccc agcgtgaacc agccctaccg caccggcttc cacttcagc ccccaagaa	1560
ctggatgaac gacccaacg gcccctgat ctacaagggc atctaccacc tgttctacca	1620
gtggaacccc aaggcgccg tgtggggcaa catcggtggt gccactcca ccagcaccga	1680
cctgatcaac tgggaccccc accccccgc catcttcccc agcggccctc tcgacatcaa	1740
cggctgctgg tccggcagcg ccaccatcct gcccacggc aagcccgtga tcctgtacac	1800
cggcatcgac cccaagaacc agcagggtgca gaacatcgcc gagcccaaga acctgtccga	1860
cccctacctg cgcgagtgga agaagagccc cctgaacccc ctgatggccc ccgacgccgt	1920

026172-005020PC - Sequence Listing_1

gaacggcatc aacgcctcca gcttccgcga cccaccacc gcctggctgg gccaggacaa	1980
gaagtggcgc gtgatcatcg gctccaagat ccaccgccgc ggcttgcca tcacctacac	2040
cagcaaggac ttcctgaagt gggagaagtc ccccgagccc ctgcactacg acgacggcag	2100
cggcatgtgg gagtgccccg acttcttccc cgtgacccgc ttcggcagca acggcgtgga	2160
gacctccagc ttcggcgagc ccaacgagat cctgaagcac gtgctgaaga tctccctgga	2220
cgacaccaag cagcactact acaccatcgg cacctacgac cgcgtgaagg acaagttcgt	2280
gccccgacaac ggcttcaaga tggacggcac cgccccccgc tacgactacg gcaagtacta	2340
cgccagcaag accttcttcg actccgcca gaaccgccgc atcctgtggg gctggaccaa	2400
cgagtcctcc agcgtggagg acgacgtgga gaagggtgg tccggcatcc agaccatccc	2460
ccgcaagatc tggctggacc gcagcggcaa gcagctgac cagtggcccc tgcgcgaggt	2520
ggagcgcctg cgcaccaagc aggtgaagaa cctgcgcaac aagggtgctga agtccggcag	2580
ccgcctggag gtgtacggcg tgaccgccgc ccaggccgac gtggagggtgc tgttcaaggt	2640
gcgcgacctg gagaaggccg acgtgatcga gccctcctgg accgaccccc agctgatctg	2700
cagcaagatg aacgtgtccg tgaagtccgg cctgggcccc ttcggcctga tgggtgctggc	2760
cagcaagaac ctggaggagt acacctccgt gtacttccgc atcttcaagg cccgccagaa	2820
cagcaacaag tacgtggtgc tgatgtgctc cgaccagtcc cgcagctccc tgaaggagga	2880
caacgacaag accacctacg gcgccttcgt ggacatcaac cccaccagc ccctgagcct	2940
gcgcgccccg atcgaccact ccgtgggtgga gagcttcggc ggcaagggcc gcgcctgcat	3000
cacctccgc gtgtacccca agctggccat cggcaagtcc agccacctgt tcgccttcaa	3060
ctacggctac cagtccgtgg acgtgctgaa cctgaacgcc tggagcatga actccgcca	3120
gatcagcatg gactacaagg accacgacgg cgactacaag gaccacgaca tcgactacaa	3180
ggacgacgac gacaagtgat taattaaccg gctcgaggca gcagcagctc ggatagtatc	3240
gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc tgccttgacc	3300
tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatctt gtgtgtacgc	3360
gctttttgca gttgctagct gcttgtgcta tttgcgaata ccacccccag catccccctc	3420
cctcgtttca tatcgcttgc atcccaaccg caacttatct acgctgtcct gctatccctc	3480
agcgtgctc ctgctcctgc tactgcccc tcgcacagcc ttggtttggg ctccgcctgt	3540
attctcctgg tactgcaacc tgtaaaccag cactgcaatg ctgatgcacg ggaagtagtg	3600
ggatgggaac acaaatggaa agcttgagct c	3631

<210> 90

<211> 621

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polypeptide

026172-005020PC - Sequence Listing_1

<400> 90

```

Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys
1      5      10      15

Ile Ser Ala Ser Met Thr Asn Glu Thr Ser Asp Arg Pro Leu Val His
      20      25      30

Phe Thr Pro Asn Lys Gly Trp Gly Arg Ala Ser His His Val Tyr Lys
      35      40      45

Arg Leu Thr Gln Ser Thr Asn Thr Lys Ser Pro Ser Val Asn Gln Pro
      50      55      60

Tyr Arg Thr Gly Phe His Phe Gln Pro Pro Lys Asn Trp Met Asn Asp
65      70      75      80

Pro Asn Gly Pro Met Ile Tyr Lys Gly Ile Tyr His Leu Phe Tyr Gln
      85      90      95

Trp Asn Pro Lys Gly Ala Val Trp Gly Asn Ile Val Trp Ala His Ser
      100      105      110

Thr Ser Thr Asp Leu Ile Asn Trp Asp Pro His Pro Pro Ala Ile Phe
      115      120      125

Pro Ser Ala Pro Phe Asp Ile Asn Gly Cys Trp Ser Gly Ser Ala Thr
      130      135      140

Ile Leu Pro Asn Gly Lys Pro Val Ile Leu Tyr Thr Gly Ile Asp Pro
145      150      155      160

Lys Asn Gln Gln Val Gln Asn Ile Ala Glu Pro Lys Asn Leu Ser Asp
      165      170      175

Pro Tyr Leu Arg Glu Trp Lys Lys Ser Pro Leu Asn Pro Leu Met Ala
      180      185      190

Pro Asp Ala Val Asn Gly Ile Asn Ala Ser Ser Phe Arg Asp Pro Thr
      195      200      205

Thr Ala Trp Leu Gly Gln Asp Lys Lys Trp Arg Val Ile Ile Gly Ser
      210      215      220

Lys Ile His Arg Arg Gly Leu Ala Ile Thr Tyr Thr Ser Lys Asp Phe
225      230      235      240

Leu Lys Trp Glu Lys Ser Pro Glu Pro Leu His Tyr Asp Asp Gly Ser
      245      250      255

Gly Met Trp Glu Cys Pro Asp Phe Phe Pro Val Thr Arg Phe Gly Ser
      260      265      270

```

026172-005020PC - Sequence Listing_1

Asn Gly Val Glu Thr Ser Ser Phe Gly Glu Pro Asn Glu Ile Leu Lys
275 280 285

His Val Leu Lys Ile Ser Leu Asp Asp Thr Lys His Asp Tyr Tyr Thr
290 295 300

Ile Gly Thr Tyr Asp Arg Val Lys Asp Lys Phe Val Pro Asp Asn Gly
305 310 315 320

Phe Lys Met Asp Gly Thr Ala Pro Arg Tyr Asp Tyr Gly Lys Tyr Tyr
325 330 335

Ala Ser Lys Thr Phe Phe Asp Ser Ala Lys Asn Arg Arg Ile Leu Trp
340 345 350

Gly Trp Thr Asn Glu Ser Ser Ser Val Glu Asp Asp Val Glu Lys Gly
355 360 365

Trp Ser Gly Ile Gln Thr Ile Pro Arg Lys Ile Trp Leu Asp Arg Ser
370 375 380

Gly Lys Gln Leu Ile Gln Trp Pro Val Arg Glu Val Glu Arg Leu Arg
385 390 395 400

Thr Lys Gln Val Lys Asn Leu Arg Asn Lys Val Leu Lys Ser Gly Ser
405 410 415

Arg Leu Glu Val Tyr Gly Val Thr Ala Ala Gln Ala Asp Val Glu Val
420 425 430

Leu Phe Lys Val Arg Asp Leu Glu Lys Ala Asp Val Ile Glu Pro Ser
435 440 445

Trp Thr Asp Pro Gln Leu Ile Cys Ser Lys Met Asn Val Ser Val Lys
450 455 460

Ser Gly Leu Gly Pro Phe Gly Leu Met Val Leu Ala Ser Lys Asn Leu
465 470 475 480

Glu Glu Tyr Thr Ser Val Tyr Phe Arg Ile Phe Lys Ala Arg Gln Asn
485 490 495

Ser Asn Lys Tyr Val Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser
500 505 510

Leu Lys Glu Asp Asn Asp Lys Thr Thr Tyr Gly Ala Phe Val Asp Ile
515 520 525

Asn Pro His Gln Pro Leu Ser Leu Arg Ala Leu Ile Asp His Ser Val
530 535 540

026172-005020PC - Sequence Listing_1

Val Glu Ser Phe Gly Gly Lys Gly Arg Ala Cys Ile Thr Ser Arg Val
545 550 555 560

Tyr Pro Lys Leu Ala Ile Gly Lys Ser Ser His Leu Phe Ala Phe Asn
565 570 575

Tyr Gly Tyr Gln Ser Val Asp Val Leu Asn Leu Asn Ala Trp Ser Met
580 585 590

Asn Ser Ala Gln Ile Ser Met Asp Tyr Lys Asp His Asp Gly Asp Tyr
595 600 605

Lys Asp His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys
610 615 620

<210> 91
<211> 997
<212> DNA
<213> Prototheca moriformis

<400> 91
cctgtcgatc gaagagaagg agacatgtgt acattattgg tgtgagggcg ctgaatcggc 60
catttttttaa aatgatcacg ctcatgccaa tagacgcggc acataacgac gttcaaacc 120
ccgccaaagc cgcggacaac cccatccctc cacaccccc acacaaagaa cccgccaccg 180
cttaccttgc ccacgaggta ggcctttcgt tgcgcaaaac cggcctcggg gatgaatgca 240
tgcccgttcc tgacgagcgc tgcccgggcc aacacgctct tttgctgcgt ctctcaggc 300
ttgggggcct ccttgggctt ggggtgccgc atgatctgcg cgcacagag aaacgttgct 360
ggtaaaaagg agcgcgggc tgcgcaatat atatataggc atgccaacac agcccaacct 420
cactcgggag cccgtccac ccccccaag tcgctgcct tgacggcata ctgctgcaga 480
agcttcatga gaatgatgcc gaacaagagg ggcacgagga cccaatccc gacatccttg 540
tcgataatga tctcgtgagt ccccatcgtc cgcccgacgc tccggggagc ccgccgatgc 600
tcaagacgag agggccctcg accaggagg gctggcccg gcgggactg gcgtcgaagg 660
tgcgcccgtc gttcgcctgc agtcctatgc caaaaacaa gtcttctgac ggggtgcgtt 720
tgctcccgtg cgggcaggca acagaggtat tcaccctggg catggggaga tcggcgatcg 780
agctgggata agagatactt ctggcaagca atgacaactt gtcaggaccg gaccgtgcca 840
tatatttctc acctagcgcc gcaaaccta acaatttggg agtcactgtg ccactgagtt 900
cgactggtag ctgaatggag tcgctgctcc actaaacgaa ttgtcagcac cgccagccgg 960
ccgaggaccc gagtcatagc gagggtagta gcgcgcc 997

<210> 92
<211> 753
<212> DNA
<213> Prototheca moriformis

026172-005020PC - Sequence Listing_1

<400> 92

```

actaattgca atcgtgcagt aatcatcgat atggtcacaa gtagatcccc tactgacacc      60
ctctcgtaca tgtaggcaat gtcatcggcg ccgtcctgct gaccgatgcc gacgtagcag      120
agcagacccg ggccgatctg ggatacgagc cggccctcca cctgcgctcg aggtggaatc      180
aagtaaataa ccaatacact ttctgacacc acacagagtt gcacggacgg tggcgtacct      240
ctacgctcgc gctcttcacg cgctggacga ccgcacgcat gagcccgggt ggcttggtct      300
gggctgcaaa aatgcacaac aaacaagtat cagacgctca tggatgcaca cgcgctccca      360
agcacgctca gactaaatat tacagtagct cgtatctgat aagatatcga gacataccgc      420
tcaactcacc cgcaaactgc gccccgccag gtgatgcgca cagggcccca ccatgcgata      480
catcgcatac ctctcagagg gcgctatcac gtggccggag agcgttcaca gcgtacgcca      540
ctgtatcttg gcggtatgca gtccgtcaac atggagacag ataccgcac caccaccttg      600
caagctcttc catattggaa gtagaaaatt gtaattgtat catcgcacga ggggccaact      660
tgccgtcggc gagctgggca acgaacacca cctggacggt gtcgagactc gctcgtgccg      720
tgcgcccggc cgctgggtat ccagaccgtc gcc                                753

```

<210> 93

<211> 1122

<212> DNA

<213> *Prototheca moriformis*

<400> 93

```

caacgacaac cagcaggcaa ctcggtcagc gacccaacac gcgagtcaaa ttgttgctg      60
ttcttgccct gtctatttac tgtgatagca agactgtcgg tcagtcaata ccgcggtgcg      120
cacgtcgggg tgccaagcct agcagagcac gggacggctg gtgctgtgca ccagctcagc      180
tcgcttcgca accaattgta ggaccggcaa agtcacaaaa acatgccagc ggtgcgattc      240
aattgggtcat gagctctaca aaattgtttt gtgcgtcgcg caggtatcca acggcgcggc      300
agagaaagtt tgacagctct cgatttcata tcggaaaaat ggggagaatt tatgacacac      360
aagtgcgcag gcggcccagg cggccagcat attctggcgt gacctgggca gccacaaaaa      420
tgcttggaat cactctaaaa taattatatt tgccatgaac aaggggaagag ttaccgcacc      480
cagccctaga cttgggcgcc cgagcaaggc tacgtcaagc caccttcgcc catcgcccaa      540
ctccgtattc cccgacagcc gcacgtggcc ctgcgaggaa tgaaccctga atcggcatac      600
cgccacgcgt tcgccaatcg ttccgctctc tggcttcata ggcctgcgcc ttcacgtcgt      660
ggtcacgaca gtgcattcat acttccattt gcacctcggc acacactttt acgcatcgcc      720
tacccttgct gcggcagctt agggtcactt tgcagccatg ggacagtgtc acaccaccgt      780
cgggtgcgca agctatttca agtgaaccgt gggcggaata aaggaatgta cactgtctca      840
accgactcct acaattgttt accatgcaga tcagagctcg acggccatca tcgagcaggt      900
gtggggccct ggtggcgcgg cgcggggccc cagggcgctc caggcattga tggcactctg      960
agactttcgc acgcgcatga gggaccccat caagagaaga gtgtgtcttt atgtcccat      1020

```

026172-005020PC - Sequence Listing_1

tcatgatgat gtatcttgtg attgtcgcag ttggcaagt ttaaccggat cgccgctcca 1080
 ggtgtggcgt ggcggatttt tctaggggtg cttgagcagt cg 1122

<210> 94
 <211> 574
 <212> DNA
 <213> *Prototheca moriformis*

<400> 94
 ggcccagggc cctgcggatg gcccacacca gatctagcct ctcttatgcc atgcccgcct 60
 cgctgcccgt cgtatcccc cgccgatccg cgcgtagggg accgcggcct gaccacgcgc 120
 acgaaagagc ttgtctctc aatttctcgc caacagaacc gtatcaaacg ctcaacgcct 180
 atccccgaaca atccgtattc acaccaaatac gagtataccg gactggtttg cctagtcttg 240
 aaggaaatga tcccgatccat gctcgggaagg gggagcgggc ggaggatcct actcatctct 300
 gaaatgggat tggatccgaag atgggttggg caagcacgtg ccaaacccca gcgagtgtgt 360
 gacgagcagg ctcatccaat cccccggcga atcctccctc acgccccgca tgcatacaag 420
 tccctccac acgccccctc ccattccattt tcgcttggtc cgaacgcgag cggcgctcag 480
 gcggaccact tgctccgcag cgccgtcttg gtctccaccc cacagcggct ttgctgccag 540
 aggcaccccc cttgccccac ctctctttgc agcc 574

<210> 95
 <211> 1096
 <212> DNA
 <213> *Prototheca moriformis*

<400> 95
 ccaggcaggc ggtaggggtg ccgattgctt gagcgaattg gaagatataa ttttttgtgg 60
 tgtccctgga cgctgtttgt ggcgctcctt ttggagaag attgcgtggg ggagctttcc 120
 atgtaccacg ctctcttctg aaaggattct ggccgagtc tgatgagccc aaagaaaaca 180
 cctgcctttc agtgcctggc ctctgaaaac gtcaacagat gattatacat gtcacaaaag 240
 gcagccgatt aggaacggga gctctggccg ttctgttggc tgcctgggct gattgaagtg 300
 atccaccctg ttgaatgaa ggcggctcag tcgaattatc gaccggagct gtcgggaagg 360
 cgtccggggc agagtgaagt gctgcggcct ggttgtcgtt caaaaagacc ccggtagccc 420
 aacaatcacg aacgaaagga atataattgc ttgcatacta tacattcagt ttctatgtgg 480
 cgggtagaca agtctcatgg gcttctaaag gctgtccctt gaaggctact tataaaaact 540
 tgctgcgcca tggcacggat cgcgcttgcg caggctgcaa ccctgcgcgc aagggtcaaat 600
 acacagcaaa agatactaac agaatttcta aaaacattta aatatttggt tcgaccagcc 660
 aattgtggtc gtaggcacgc aaaagacttt gttttgcgcc caccgagcat ccacgctggc 720
 agtcaagcca gtccgatgtg cattgcgtgg cagcatcgag gagcatcaaa aacctcgtgc 780
 acgcttttct gtcaatcatc atcaaccact ccaccatgta taccgatgc atcgcggtgc 840
 gcagcgcgcc acgcgtccca gaccgcccc aaaaccagc agcggcgaaa gcaaatcttc 900

026172-005020PC - Sequence Listing_1

acttgcccga aaccccgagc agcggcattc acacgtgggc gaaaacccca cttgccctaa 960
caggcgatg tctgctgtca cgaatgcctga caacgggtatt atagatatatac actgattaat 1020
gtttgagtgt gtgcgagtcg cgaatcagga atgaattgct agtaggcact ccgaccgggc 1080
gggggccgag ggacca 1096

<210> 96
<211> 1075
<212> DNA
<213> *Prototheca moriformis*

<400> 96
ggccgacagg acgcgcgtca aagggtgctgg gcgtgtatgc cctgggtcggc aggtcgttgc 60
tgttgctgcg ctggtggttc cgcaaccctg attttggcgt cttattctgg cgtggcaagc 120
gctgacgccc gcgagccggg ccggcggcga tgcggtgtct cacggctgcc gagctccaag 180
ggaggcaaga gcgcccggat cagctgaagg gctttacacg caaggtagac ccgctcctgc 240
aaggctgctg ggtggacttg aacctgtagg tcctctgctg aagttcctcc actacctcac 300
cagggccagc agaccaaagc acaggctttt cagggtccgtg tcattccactc taaaacactc 360
gactacgacc tactgatggc cctagattct tcattcaaaa tgcctgagac acttgctcag 420
aattgaaact ccctgaaggg accaccagag gccctgagtt gttccttccc ccggtggcga 480
gctgccagcc aggtgtgacc tgtgatcgag gctggcgagg aaataggctt cgtgtgtcga 540
ggcatgagg ggtgcaggac agctcatgaa acgccaacaa tcgcacaatt catgtcaagc 600
taatcagcta tttcctcttc acgagctgta attgtcccaa aattctgggtc taccgggggt 660
gatccttcgt gtacgggccc ttcctcaac cctaggtatg cgcgcatgcg gtcgccgcgc 720
aactcgcgcg agggccgagg gtttgggacg ggccgtccc aaatgcagtt gcacccggat 780
gcgcggcgcc tttcttgca taatttatgc aatggactgc tctgcaaatt tctgggtctg 840
tcgccaaccc taggatcagc ggcgtaggat ttcgtaata ttcgtcctga tggggagcta 900
ccgactaccc taatatcagc ccggctgcct gacgccagcg tccacttttg cgtacacatt 960
ccattcgtgc ccaagacatt tcattgtggt gcgaagcgtc ccagattacg ctcacctgtt 1020
tcccgcacct cttactgttc tgtcgacaga gcggggccac aggccgggtc cagcc 1075

<210> 97
<211> 772
<212> DNA
<213> *Prototheca moriformis*

<400> 97
tcaccagcgg acaaagcacc ggtgtatcag gtccgtgtca tccactctaa agagctcgac 60
tacgacctac tgatggccct agattcttca tcaaaaacgc ctgagacact tgcccaggat 120
tgaaactccc tgaagggacc accaggggccc ctgagttgtt ccttcccccc gtggcgagct 180
gccagccagg ctgtacctgt gatcggggct ggcgggaaaa caggcttcgt gtgctcaggt 240
tatgggaggt gcaggacagc tcattaaacg ccaacaatcg cacaattcat ggcaagctaa 300

026172-005020PC - Sequence Listing_1

tcagttat	tttccc	at	taacg	agctata	aatt	gtcccaaa	aatt	tctgggt	ctac	cggggg	gtgat	360	
cc	ttcgtg	tga	cg	ggcccttc	cctcaacc	cct	aggtat	gcgc	acatgc	ggtc	gccgcg	caac	420
gcgcgc	gagg	gccgag	gggt	tgggac	gggc	cg	ccccg	aaa	tgcag	ttgca	cccgat	gcg	480
tggcac	cttt	tttgc	gataa	tttatg	caat	ggactg	ctct	gcaaaa	attct	ggctct	gtcg	540	
ccaacc	ctag	gatcag	cggg	gtagg	atttc	gtaatc	atttc	gtcctg	atgg	ggagct	accg	600	
actgcc	ctag	tatcag	ccccg	actgc	ctgac	gccagc	gtcc	actttt	gtgc	acacatt	cca	660	
ttcgtg	ccca	agacatt	tca	ttgtg	gtgcg	aagcgt	cccc	agttac	gctc	acctgat	ccc	720	
caacct	ctct	attgtt	ctgt	cgacag	agt	ggcccag	agg	ccggtc	gcag	cc		772	

<210> 98

<211> 991

<212> DNA

<213> *Protoheca moriformis*

<400> 98

cgaagggg	gtc	tgc	atc	gatt	cg	gcgc	gggtct	ggaggc	cagc	gtgact	gctc	gcgaaa	atgc	60
tctgcc	gtgt	cg	gggtct	gg	ctgggg	cggc	cagagat	ctc	accgtg	ccac	acgca	actgc		120
cg	actct	gt	gccc	gccacc	tggcgc	gcac	atgcg	acctc	ttcccc	gtca	tacctc	ctcc		180
tc	atgt	g	atc	tttcc	acacg	agt	gacgcag	gtgcgc	ggag	tggagg	gaat	caggac	gttt	240
tca	agg	tacc	tgctc	gagcc	gtacca	acag	ctgccg	ccccg	gcaagg	aaga	gatc	gagg	ga	300
gag	attg	ccc	ggctg	gaggc	ccggata	acg	gagct	caaga	gcaag	ctgtc	cgagt	gagac		360
cg	ccag	gtg	cag	gtgtc	ga	ctcgt	tatga	catgt	actcg	acaca	acatg	agga	attcat	420
cga	attt	gta	ggaag	cgggc	attgg	tacgg	gagt	gggaaa	gcgaaaaa	ac	ctccct	ccgg		480
cag	tgcc	atc	tgccg	gagtc	gaacg	ttgat	agggt	ttctc	tgacag	gggtg	tgacct	ctca		540
gc	cttg	catc	aattaa	acgc	tatag	acatt	atcag	taacc	gtga	atcccc	cattg	gatgc		600
ca	ccgcgc	g	accatt	gggg	acctg	catta	cagat	ctagg	tgag	atgaca	gcgagg	gaac		660
ttc	ggccc	gc	ggccc	agctt	gcggc	gcacc	aatatt	ggtc	acggg	aagcc	acacacc	gac		720
cata	aatgaa	tactt	gtaag	ctatg	tcaac	cgatc	aatgg	cg	tcgaa	agt	gtgcc	acgag		780
gat	ccat	ctg	gcgggg	cggc	gtggc	gcaca	agcgc	agtcg	caattt	ctc	gaccc	atctg		840
ac	ctag	gccc	agcgc	cgcg	gagaa	atccc	cggcg	gggtcc	tccac	gcag	aacc	cta	atg	900
ag	tatc	gagc	gccg	accatt	tacacc	atcg	cccc	gaaat	ccttcc	gaca	ttatt	attat		960
cttt	tag	atc	ttgga	acaga	ctctg	ccaac	c							991

<210> 99

<211> 1347

<212> DNA

<213> *Prototheca moriformis*

<400> 99

agagag	cgga	ggtggg	gttg	tgagg	tggg	ttgct	gacca	ggagct	gcg	tcgcc	gagcg	60
cgactc	gcac	acgg	tccag	t	accccccc	tccg	cccaaa	cgcaag	cctc	ccatct	tgat	120

026172-005020PC - Sequence Listing_1

gcctttccgg ccacctatac tttttcttag ttcgctgtaa catccagacc gtcctgaata	180
ataacaatgc cctgtgtcaa gtgcattcct aaaaaaatc tgtcccaacc aacaatccca	240
cctgaaatac caccagccct gcccagtaca ctcttccaat accatctccc tacctccacg	300
cgcaagcgac ccccatgcmc gaccaggctc gaaagtgatt tatgacttga gacgagcgag	360
tggcggcgcg gtcgactgcc ttttcatcac gtgccgtacg tcggcgaccg ctagggcttt	420
gcacggcaac gcacggcttc gccaaaccga ccagccagga cctcgactac tctaccgca	480
attcgctca agaagtcmc aaatgtgcca tacaccattc cttacagcac tgttcaaact	540
tgatgccaat tttgacattc gggttgctcg ttggctgcmc ccacatcmc cgtgagtga	600
gcaggcggga tcggacacgg aggacgcggc gtcacgcccc gaacgcagcc cgtaactcta	660
catcaacacg acgtgttgcg taatcccgcc cggctgcmc tcgtgccaac ccattcgca	720
tggatggctg gaaaatggcg tgccaactgc cctgagggag gctctcgca aacgggcacg	780
tccctgaaac cgaaactgtg gccttgctcg cggccacgca agcacgtga ccctaaacac	840
caagaaaatc agtaaacaag gttgacatcc tctacggcg aattgtttgc ccaacccttc	900
atcgcacact gccattataa tgcatttagc tcggcgacaa gtttagaaaa ggcaggctgc	960
attgttccat ttcgccgtgg cggcggtggg gccatttta cgagggttg gctcccgggc	1020
agcgaccgag ccaggctcag tccctctcgc ccgtcgacaa tgttgcaac cccacaagcg	1080
gctaacaaca acttgatgg acctgtacac tgccaattcc ttcttccccg gccgagggtt	1140
acacgtgatg gccatggctt cgcattcagg ccgacttccc attccgactt tccagagggt	1200
ccgcggacgc tgggggttg ctgcctgagg cccacccttt gttcccccg tcccgacaaa	1260
cacaattgcg ttacataagg gggagccgcc cccgttcaga gtgcagaaat ctttactat	1320
attttccagt cgtcagcgaa atcaagt	1347

<210> 100

<211> 1180

<212> DNA

<213> *Prototheca moriformis*

<400> 100

gatgggtggg tgtctgcctt gggctgggtg atggaggctg gtggtgcgcg ggtttctga	60
tgcatcttat ctacgcagtg tcatggtgtc cattccacac accagtacac ctttacta	120
aggatccatc cctcttccc tcttcaggac tacatggacc ccacgagcta ccgaccgggc	180
tttctcaaaa acgtcaagg catgtttgac atgcgggacg tgggtggacga cgtgcaagg	240
gcgtccggag tgcgcgcaaa tgagcaagtc gggcaatgtg tcgggggtgg caccggggct	300
ggagatccgc gatccccgag aaaacgccgt accaccccc gcgctattcc ctcgattgcg	360
cgcagatgtg gtgaccgaca cgggggacaa cctggcgac atggggcgcc ggacctggaa	420
gcacgccaag tcgcacacgg ggaggctcgt gcagtcccc ccacgtacc tcaagggtct	480
ctttggtcgc gatccaaagt acgctgggtg catggcatgc ccgaaatgaa catcatgtgt	540
gatctccgat tgccaatggc cacctccacg gaccacctg caggcggaag cgcaatccag	600

026172-005020PC - Sequence Listing_1

ggcccgagcc tgacgaggac ggagactcct cgtccagcgc ggggtccccg acccgacgca	660
gcagccgacc cctgctaacc cggcaacgat cggaccagca accttgctgt agttccgatc	720
cgtgatgacg ggcattgccg ccgctcgatc cgctttgatg actgtctatt atttgcgcgg	780
agccccctcg gaaccctacc ccgctcttgc aagccccctg catcggagat cctcgtgcgc	840
ccgccatgac cccactggat tgcccaacat ccttctttat cgtgtaaaat gtgattcctc	900
ggctgcaatc gactggcctt cgcttctggc cccaagaggg ctcgaacgtg cggcagcgag	960
ggcgtgaca cacccaagcc ctagggcttt caacgtcggc tgccaggccg gataggggga	1020
tcgcctcctt tccaccaccc acctacgagg gattcgagtc ggcttccagc tcagctattc	1080
ggccgcgccc ccggccctgc agacgtcctc cagtttccga acaggtcgct ctcagaacac	1140
ctgccgcggc tgcgatacgg caggctctca aagcgtcgac	1180

<210> 101

<211> 1263

<212> DNA

<213> *Prototheca moriformis*

<400> 101

cgcgtaggagc ggtgcgtgcg gatgccgcgc gcctgccaaag gccttttgta tgcctggcct	60
gggaagtttc ctgactgaag catcttcaag atgctctctc acgaccagcg acaccaacac	120
cgtcactttt tgccccctct gccgcagggtg ccactttcta ctttgacgtc ttctccaggc	180
ggtacattgc gggactgagc gccaattcgg ccaagaacag cgctgtcgac ttgaggaggc	240
aggggtccgt cgactctgcc gagtgacacg ccttcgaccc gactgtacta cggcctgctg	300
aagagtgggt ctgcgcggcc ggcgtgaccg gccctgtgcc cacaatcgac catctattcg	360
ctccttgatc tctggcgccg tcaattgccc gcgacttgac ggcaactggc tcgatcgagt	420
cgtattgaaa aagcacgttt tgtcctacag ggccgcggtc cgttaccaac gtggttctcg	480
ttaggttttc gtcgggcggg ggtgcgcgaa ctgtccgatg ccatcccggc aaaccccagc	540
aaggctcgcca gtctggttct gacgcaatag agtgcgtttt gggccagtct aaaaattcgt	600
ctggcatgac gtggctccac atcgtacccg gagcctgcct tggtaatgtg aggcaccggt	660
gccaactcca ttatggcagg catcgagcgc gcaggtagt acatgacctt ccgtgaattg	720
ggaaggcgag cttgtgtaac gcctgcgatc gtgccagtga ggcatcgtaa actcaaaata	780
ttttgtagaa agtgtctgat gcctggtgag gctgcgtagg gcaaggggcaa gcccttggca	840
gatgggtaat ggggtccggac ctcaacaacg caaccccgcg tcccccttag ggcccctgag	900
gctcgatggc agggccagcg agcccgcggc caaaggcgcg catcccacgg tcgcccacg	960
actccacggg tcctatacct catcttgaat ggcactaaaa actatagaat atcgggcact	1020
ggtgggcgtc tggggtacag ctggccgagc gcagtggcaa accctaggtc ccgcctcaag	1080
ggcgattccc ggggtcaatga cacgcaagca agatcacatg gcgcgggtccg cctcgcggct	1140
ccacacccag gccctagttt cgcaacccat aaatatcgcc ccgataccat cataagccag	1200

026172-005020PC - Sequence Listing_1

caaataatTTT tttatcagag ttccaaacct cctcagctgt gggaaaccag cccactctga 1260
acg 1263

<210> 102
<211> 1400
<212> DNA
<213> Prototheca moriformis

<400> 102
ccgagcagtt catggccaag tacaaggact agagaccgga ggtcggtagg ctgaatggag 60
ctggcgctcgt cgtgcgcgac gtgcacgcga tgcgatacta cgaccccaaca aacgcacgac 120
tcccatcttg atgcctttcc ggccatttat actatttctc atttcgctgt aacatcttga 180
ataatagaat tgccctgtgt caagtggatt ccaagaaata ttctgtccca acaaaacaac 240
ccaacctgaa aacaacctca aataccacca gccctgcccc cctgcccagt acacttttcc 300
aataccatct ccctaccttc acgcgcaagc ggcacccatg cgcgaccagg ctcgaaagga 360
tttcacgact caggacgagc gagtggcggc gcgaccgcct gcctgttcgt cacgtgccgt 420
acgtcggcga ccgctagagc tttgcctggc aacccccggc ttcgtcaacc cggccagcca 480
ggatctcgac cactctaccg cgaaatcgcc tcaagaagtc gccaaaagtg ccgtacacca 540
tgcttcgcag cgctgttcaa acttgatgcc aatcttgaca atcagggttg tcgttggtg 600
cgtccacatc ggccgtgatt gcagcaggcg gggatcggac acggaggacg cggcgtcacg 660
ccgcgaacgc agcccgtaac tctacatcaa cgcgatatgt tgcgtaatcc cgcccggctg 720
cgcattgtga caaccattc gcgatggatg gtcggaaaat ggtgtgccaa ctgccctgag 780
ggactctctc gcgaaacggg cacgtccctg tatccgaaac tgtggcatgg ccttgctcgc 840
cacgcaagca cgtggaccct aacaccacga aaataagtaa aaaagggtga catcctctac 900
gagcgaattg tttgctcgac ccttcacgc acactgtcat tataatgcat ctagctcggc 960
gacaagttaa aaaaaggcag gctgcattat tccattttgc cgtggcggca tgggtgccca 1020
ttttatgagg tttgggctct tgggcagcga ccgagccagg ttgagtcctc ctgcccgtc 1080
gacaacgttc caaagcccat aagtggctaa taaacaactt gatggtacct gtacactgcc 1140
agttccttct tccccggccg aggtttacac gtgatggcca tggcttcgcg tttcaggctg 1200
acttccatt ccgactttcc agaggggtccg cggacgccgg gggttggctg cgtgaggccc 1260
accccttggt ccccgctcc cgacaaacac aattgcgtta cataaggggg aagccgcccc 1320
ccgttcagag tgcaaacatc tttcattata tttttcagtc gtcagcgaat tcaagtatgt 1380
cgctgacagg catgaaggcc 1400

<210> 103
<211> 3681
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

026172-005020PC - Sequence Listing_1

<400> 103
gccctttgtc atcgtttggca tgcttttttgc gtatgtacca tatgttgaat gtataatacgc 60
aacggttgac cgtctgagat gcgagctttg ggtcttgtca aatgcgtggc cgcacggctc 120
cctcgcaccc agccccgagg cgtcgcgcac ctggcgaggga gcagacccac gccaaagaaag 180
tctagtccag catgtaacaa catcaggcaa tgtgacgttt tcggttccccg atttctctgc 240
cgctctttga cggcaggcac gggcgagcaa ccggcggcgc tcgcgtcagg cacgatggat 300
gcggcgctgc ccacctgtca atgtaccca ccagtctgtc gatcgctaca agcaaccttg 360
tgctccacat tcccacttgc agacagtcta gtcgattttg ccaagctgga tgtgaggatt 420
ggccatatct tggaggccaa gattcacccg gatgctgatg ggtacgtacg cgagccaggc 480
aggcagctgc gttgactttc tgattggcac aaagctttgg ctactctcaa taccaaccac 540
gtgccccctt tgcacacctg cttccttctg atgaccactc gccacgcatg tcgcagtctg 600
tacgtcgagc agatcgacct cggcgaggag gggggccctc gcaccatcgt gagtggcctg 660
gtccggcacg tgaccttggg ggaccttgtc ggccggcggg tgggtggtgct ggccaacctc 720
aagcctcgga gcatgcgcgg ggtcaaactc gctgggatgc tgctctgcgc cgccaacgcg 780
gatcacaccg cgggtggagcc gctgcgggtc ccggacgccg ccgtgacggg ggagcgggtc 840
tgggcggggg acgaggcact cctgtccacg gagcctgcca caccatca ggtaaggaca 900
cgttattggt gcgcatggtg cgaatgcgtg gtctgacctg ctgtgggtat gtgttgtggg 960
atttgaaacc gaatgagggc cgttcaggat tgagcccttg gccccaccct gctcatcctc 1020
tcacgcccgc aggtccagaa gaagaaaatc tgggaggcag tacagccgct gctgagagtg 1080
aacgcccagg ggatcgctac tgtggcagga gaggctatgg tgaccagtgc ggggccactg 1140
accgcgcca cgctggttga cgccgcgatt tcctgacgcg agcgactgat tcttgacctt 1200
tgagaagcca ccacagcacc attttcattg ttcatccttg atttcagtac gacttctcac 1260
catttcagta ctgtaggacc cccaaaatag tgtgatcacg ctcgcaaggc acctgtgtga 1320
tcacggggaa gggcgaattc ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg 1380
gctgcgagac ggcttccccg cgctgcatgc aacaccgatg atgcttcgac cccccgaagc 1440
tccttcgggg ctgcatgggc gctccgatgc cgctccaggg cgagcgtgt ttaaatagcc 1500
aggccccga ttgcaaagac attatagcga gctaccaaag ccatattcaa acacctagat 1560
cactaccact tctacacagg ccactcgagc ttgtgatcgc actccgctaa gggggcgcct 1620
cttcctcttc gtttcagtca caaccgcga acggcgcgcc atgctgctgc aggccttcct 1680
gttcctgctg gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga 1740
ccgccccctg gtgcacttca cccccaacaa gggctggatg aacgaccca acggcctgtg 1800
gtacgacgag aaggacgcca agtggcacct gtacttccag tacaaccga acgacaccgt 1860
ctgggggacg cccttgttct gggggccacgc cagctccgac gacctgacca actgggagga 1920
ccagcccatc gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt 1980

026172-005020PC - Sequence Listing_1

ggtggactac aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg	2040
cgtggccatc tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct	2100
ggacggcggc tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac	2160
ccagttccgc gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc	2220
ggccaagtcc caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa	2280
gctggagtcc gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct	2340
gatcgaggtc cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat	2400
caaccccggc gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg	2460
caccacttcc gaggccttcg acaaccagtc ccgcgtgggtg gacttcggca aggactacta	2520
cgccctgcag accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcggtg	2580
ggcctccaac tgggagtact ccgccttcgt gccaccaaac ccctggcgct cctccatgtc	2640
cctcgtgcgc aagttctccc tcaacaccga gtaccaggcc aaccggaga cggagctgat	2700
caacctgaag gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc	2760
caccaacacc acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg	2820
caccctggag ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt	2880
gttcgcggac ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat	2940
gggcttcgag gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt	3000
cgtgaaggag aaccctact tcaccaaccg catgagcggtg aacaaccagc ctttcaagag	3060
cgagaacgac ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct	3120
gtacttcaac gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc	3180
cctgggctcc gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca	3240
ggtgcgcgag gtcaagtgat taattaactc gaggcagcag cagctcggat agtatcgaca	3300
cactctggac gctggtcgtg tgatggactg ttgccgccac acttgctgcc ttgacctgtg	3360
aatatccctg ccgcttttat caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt	3420
ttgcgagttg ctagctgctt gtgctatttg cgaataccac ccccagcatc cccttccctc	3480
gtttcatatc gcttgcattc caaccgcaac ttatctacgc tgtcctgcta tccctcagcg	3540
ctgtcctgc tcctgctcac tgcccctgc acagccttgg tttgggctcc gcctgtattc	3600
tcctggtact gcaacctgta aaccagcact gcaatgctga tgcacgggaa gtagtgggat	3660
gggaacacaa atggaaagct t	3681

<210> 104
 <211> 3850
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

026172-005020PC - Sequence Listing_1

<400> 104

tttggccccc	ctttccagct	ccggatctgc	tggcgccgc	cgcgagacgt	gacatcgcac	60
gtcgccggga	gcgccagctt	gatacattgg	cagggggccg	tgctctacaa	ataccaggcc	120
ccgcggcggt	cagttcgcac	atccaatacc	tgccgagcca	tcttgccctac	actttttatc	180
gactcctcta	ctctgttcgc	gagagcgctc	gggccaggct	tggaattcgc	cgaattcagc	240
tcgatcagtc	gcttctgcaa	ctgatctcgg	ccgttcgcag	actgcctttt	ctcagcttgt	300
caggtagcga	gttggtgttt	tatatattatt	cgatttcata	tgtgttgcat	gtcttggttcg	360
tgctgtgcgt	tctttctggg	ccgcgctgtc	gggtcgcata	ggctagctgt	actcatgtta	420
gtcatgccgg	tccgaccttg	ttcgaggaag	gccccacact	gagcgtgccc	tctttctaca	480
cccccttgtc	agaaattaga	tagaaagcag	aatttcctttc	ttgcgctatg	acacttcag	540
caaaaggtag	ggcgggctgc	gagacggctt	ccggcgctg	catgcaacac	cgatgatgct	600
tcgaccccc	gaagctcctt	cggggctgca	tgggcgctcc	gatgccgctc	cagggcgagc	660
gctgtttaaa	tagccaggcc	cccgattgca	aagacattat	agcgagctac	caaagccata	720
ttcaaacacc	tagatcacta	ccacttctac	acaggccact	cgagcttggtg	atcgactcc	780
gctaaggggg	cgctctctcc	tcttcgtttc	agtcacaacc	cgcaaacggc	gcgccatgct	840
gctgcaggcc	ttcctgttcc	tgctggccgg	cttcgccgcc	aagatcagcg	cctccatgac	900
gaacgagacg	tccgaccgcc	ccctggtgca	cttcaccccc	aacaagggct	ggatgaacga	960
ccccaacggc	ctgtggtacg	acgagaagga	cgccaagtgg	cacctgtact	tccagtacaa	1020
cccgaaacgac	accgtctggg	ggacgcccc	gttctggggc	cacgccacgt	ccgacgacct	1080
gaccaactgg	gaggaccagc	ccatcgccat	cgccccgaag	cgcaacgact	ccggcgccctt	1140
ctccggctcc	atggtggtgg	actacaacaa	cacctccggc	ttcttcaacg	acaccatcga	1200
ccgcgccag	cgctgcgtgg	ccatctggac	ctacaacacc	ccggagtccg	aggagcagta	1260
catctcctac	agcctggacg	gcggctacac	cttcaccgag	taccagaaga	accccgtgct	1320
ggccgccaac	tccacccagt	tccgcgaccc	gaaggctctc	tggtagcagc	cctcccagaa	1380
gtggatcatg	accgcggcca	agtcccagga	ctacaagatc	gagatctact	cctccgacga	1440
cctgaagtcc	tggaaagctgg	agtccgcgtt	cgccaacgag	ggcttcctcg	gctaccagta	1500
cgagtgcgcc	ggcctgatcg	aggccccac	cgagcaggac	cccagcaagt	cctactgggt	1560
gatgttcata	tccatcaacc	ccggcgcccc	ggccggcggc	tccttcaacc	agtacttcgt	1620
cggcagcttc	aacggcaccc	acttcgaggc	cttcgacaac	cagtcccgcg	tgggtggactt	1680
cggcaaggac	tactacgccc	tgagacctt	cttcaacacc	gacccgacct	acgggagcgc	1740
cctgggcata	gcgtgggcct	ccaactggga	gtactccgcc	ttcgtgcca	ccaacccctg	1800
gcgctcctcc	atgtccctcg	tgcgcaagtt	ctccctcaac	accgagtacc	aggccaaccc	1860
ggagacggag	ctgatcaacc	tgaaggccga	gccgatcctg	aacatcagca	acgccggccc	1920
ctggagccgg	ttcgccacca	acaccacggt	gacgaaggcc	aacagctaca	acgtcgacct	1980
gtccaacagc	accggcaccc	tggagttcga	gctggtgtac	gccgtcaaca	ccaccagac	2040

026172-005020PC - Sequence Listing_1

gatctccaag tccgtgttcg cggacctctc cctctgggtc aagggcctgg aggaccccgga	2100
ggagtacctc cgcattgggt tgcagggtgc cgcgtcctcc ttcttcctgg accgcgggaa	2160
cagcaagggtg aagttcgtga aggagaaccc ctacttcacc aaccgcatga gcgtgaacaa	2220
ccagcccctc aagagcggga acgacctgtc ctactacaag gtgtacggct tgctggacca	2280
gaacatcctg gagctgtact tcaacgacgg cgacgtcgtg tccaccaaca cctacttcac	2340
gaccaccggg aacgcctctg gctccgtgaa catgacgacg ggggtggaca acctgttcta	2400
catcgacaag ttccagggtgc gcgagggtcaa gtgattaatt aactcgaggc agcagcagct	2460
cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg	2520
ctgccttgac ctgtgaatat ccctgccgct ttatcaaac agcctcagtg tgtttgatct	2580
tgtgtgtacg cgcttttgcg agttgctagc tgcttgtgct atttgcaaat accaccccca	2640
gcatccccct ccctcgcttc atatcgcttg catcccaacc gcaacttacc tacgctgtcc	2700
tgctatccct cagcgtgtct cctgctcctg ctactgccc ctgcacagc cttgggttgg	2760
gctccgcctg tattctcctg gtactgcaac ctgtaaacca gcactgcaat gctgatgcac	2820
gggaagtagt gggatgggaa cacaatgga ccgacacgcc cccggcccag gtccagttct	2880
cctgggtctt ccagaggccc gtcgccatgt aaagtggcag agattggcgc ctgattcgat	2940
ttggatccaa ggatctccaa tcggtgatgg ggactgagtg cccaactacc acccttgcac	3000
tatcgtctc gcactattta ttccacctt ctgctcgcct tgccgggcca ttgcgggcgt	3060
ttctgccctt gacgtatcaa ttctgcccct gctggcgcca ggattcttca ttctaataag	3120
aactcactcc cgccagctct gtacttttcc tgcggggccc ctgcatggct tgttcccaat	3180
gcttgctcga tcgacggcgc ccattgccc aaggcgtgcc gcatccatgt gaagaaacac	3240
ggaagagtgc gaagactgga agtgaattaa gagtataaga agaggtagca agggattctc	3300
agggtgctctt aggaacggct ttctcttcgc caagagaaac tgctactgct cgtgtcgcca	3360
cgggtgtcaa gccgcccct ctgcgatcca ccaggcccct ccgcggactc gcgatcagcc	3420
tgctggatcc ggactgccga cctgaccgct cgcatccacc attacaaccc tccaattgga	3480
caccactccc acgtcctaaa gttcaccatg caagctgacg gatcgattc gccgatgcac	3540
tcgcctgcca cagagggtgt cgcttcggac tagcgtgcag gcgccccgag gccaccagca	3600
tgcaccgatg gaagcgggca cggccgctgc tccaggctgc tggtctgctc agaccatag	3660
caacctccgc tgcgtcccta aatgtcacac agagcgtctt tgatgggtac ggatgggaga	3720
gaatctgatt gggcattgct ggtgcagtgc aggaagatgg caagtgcaca gtcagtcatg	3780
ctgtacaaac tgggtgcctcg tagtattgac tcgtatagtg catagtatca tgcattgtcg	3840
ttacttgcaa	3850

<210> 105

<211> 3108

<212> DNA

<213> Artificial Sequence

026172-005020PC - Sequence Listing_1

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 105

tttggccccg ctttccagct ccggatctgc tggcgtccgc cgcgagacgt gacatcgcac	60
gtcgccggga gcgccagctt gatcacttgg cagggggcgc tgctctacaa ataccaggcc	120
ccgcggcggt cagttcgcac atccaatacc tgccgagcca tcttgccctac actttttatc	180
gactcctcta ctctgttcgc gagagcgtc ggtccaggct tggaattcgc cgaattcagc	240
tcgatcagtc gcttctgcaa ctgatctcgg ccgttcgcag actgcctttt ctcagcttgt	300
caggtagcga gttgttgttt tatatttatt cgatttcac tgtgttgcac gtcttgttcg	360
tgctgtgcgt tctttctggg ccgcgctgtc gggtcgcacg ggctagctgt actcatgtta	420
gtcatgccgg tccgacctg ttcgaggaag gcccacact gagcgtgccc tctttctaca	480
ccccttgctg agaaattaga tagaaagcaa tgctgctgca ggccttcctg ttcctgctgg	540
ccggcttcgc cgccaagatc agcgcctcca tgacgaacga gacgtccgac cgccccctgg	600
tgcacttcac cccaacaag ggctggatga acgaccccaa cggcctgtgg tacgacgaga	660
aggacgcaa gtggcacctg tacttccagt acaaccgaa cgacaccgtc tgggggacgc	720
ccttgttctg gggccacgcc acgtccgacg acctgaccaa ctgggaggac cagcccatcg	780
ccatcgcccc gaagcgcaac gactccggcg ctttctccgg ctccatggtg gtggactaca	840
acaacacctc cggcttcttc aacgacacca tcgaccgcgc ccagcgtgc gtggccatct	900
ggacctacaa cccccggag tccgaggagc agtacatctc ctacagcctg gacggcggct	960
acaccttcac cgagtaccag aagaaccccg tgctggccgc caactccacc cagttccgcg	1020
acccgaaggt cttctggtac gagccctccc agaagtggat catgaccgcg gccaaagtccc	1080
aggactacaa gatcgagatc tactcctccg acgacctgaa gtcctggaag ctggagtcgc	1140
cgttcgccaa cgagggttc ctcggctacc agtacgagtg ccccggcctg atcgaggtcc	1200
ccaccgagca ggaccccgag aagtcctact ggggtgatgtt catctccatc aaccccggcg	1260
ccccggccgg cggctccttc aaccagtact tcgtcggcag cttcaacggc acccacttcg	1320
aggccttcga caaccagtcc cgctgggtgg acttcggcaa ggactactac gccctgcaga	1380
ccttcttcaa caccgacccg acctacggga gcgccctggg catcgctgg gcctccaact	1440
gggagtactc cgcttcctg cccaccaacc cctggcgctc ctccatgtcc ctcgtgcgca	1500
agttctccct caacaccgag taccaggcca acccgagac ggagctgatc aacctgaagg	1560
ccgagccgat cctgaacatc agcaacgccg gccctggag ccggttcgcc accaaccacca	1620
cgttgacgaa ggccaacagc tacaacgtcg acctgtccaa cagcaccggc accctggagt	1680
tcgagctggg gtacgccgtc aacaccaccc agacgatctc caagtccgtg ttcgcggacc	1740
tctccctctg gttcaagggc ctggaggacc ccgaggagta cctccgcatg ggcttcgagg	1800
tgtccgcgtc ctcttcttc ctggaccgcg ggaacagcaa ggtgaagttc gtgaaggaga	1860

026172-005020PC - Sequence Listing_1

acccctactt	caccaaccgc	atgagcgtga	acaaccagcc	cttcaagagc	gagaacgacc	1920
tgtcctacta	caaggtgtac	ggcttgctgg	accagaacat	cctggagctg	tacttcaacg	1980
acggcgacgt	cgtgtccacc	aacacctact	tcatgaccac	cgggaacgcc	ctgggctccg	2040
tgaacatgac	gacgggggtg	gacaacctgt	tctacatcga	caagttccag	gtgcgcgagg	2100
tcaagtgacc	gacacgcccc	cggcccaggt	ccagttctcc	tgggtcttcc	agaggccccgt	2160
cgccatgtaa	agtggcagag	attggcgcct	gattcgatit	ggatccaagg	atctccaatc	2220
ggatgatggg	actgagtgcc	caactaccac	ccttgcaacta	tcgtcctcgc	actatttatt	2280
cccaccttct	gctcgccctg	ccgggcgatt	gcgggctgtt	ctgcccttga	cgtatcaatt	2340
tcgccccctg	tggcgcgagg	attcttccatt	ctaataagaa	ctcactcccc	ccagctctgt	2400
acttttctct	cggggccccct	gcatggcttg	ttcccaatgc	ttgctcgatc	gacggcgccc	2460
attgcccacg	gcgctgccgc	atccatgtga	agaaacacgg	aagagtgcga	agactggaag	2520
tgaattaaga	gtataagaag	aggtaccaag	ggattctcag	gtgctcttag	gaacggcttt	2580
tccttcgcca	agagaaactg	ctactgctcg	tgtcgccacg	gtggtaagc	cgccccatct	2640
gcgatccacc	aggcccatcc	gcggactcgc	gatcagcctg	ctggatccgg	actgccgacc	2700
tgaccgctcg	catccaccat	tacaaccctc	caattggaca	ccactccac	gtcctaaagt	2760
tcaccatgca	agctgatcga	tcgcattcgc	cgatgcactc	gcctgccaca	gagggtgtcg	2820
cttcggacta	gcgtgcaggc	gccccgaggc	caccagcatg	caccgatgga	agcgggcacg	2880
gccgctgctc	caggctcgctg	gctcgctcag	acccatagca	acctccgctg	cgtccctaaa	2940
tgtcacacag	agcgtctttg	atgggtacgg	atgggagaga	atctgattgg	gcattgctgg	3000
tgcagtgcag	gaagatggca	agtgcacagt	cagtcattgt	gtacaaactg	gtgcctcgta	3060
gtattgactc	gtatagtgcg	tagtatcatg	catggctcgt	acttgcaa		3108

<210> 106

<211> 559

<212> DNA

<213> Chlorella luteoviridis

<400> 106

tgttgaagaa	tgagccggcg	acttatagga	agtggcgtgg	ttaaggaatt	ttccgaagcc	60
caagcgaaag	caagttttta	aaatagcgat	atttgtcact	ttttatggac	ccgaacccgg	120
gtgatctaac	cgtgaccagg	atgaagcttg	ggtaacacca	agtgaaggtc	cgaactcttc	180
gatctttaaa	aatcgtgaga	tgagttgcgg	ttagtaggtg	aaatgccaat	cgaactcgga	240
gctagctggg	tctccccgaa	atgtgttgag	gcgcagcgat	gaatgacaaa	acaaatagta	300
cgggtgtagg	gtaaagcact	gtttcggtgc	gggctgcgaa	agcggtagca	aatcgtggca	360
aactcagaat	actacgcttg	tataccattc	atcagtgcga	ctgtggggga	taagctccat	420
agtcaagagg	gaaacagccc	agatcaccag	ttaaggcccc	aaaatgacag	ctaagtggca	480
aaggaggtga	aagtgcagaa	acaaccagga	ggtttgccca	gaagcagcca	tcctttaaag	540
agtgcgtaat	agctcactg					559

026172-005020PC - Sequence Listing_1

<210> 107
 <211> 1841
 <212> DNA
 <213> *Cuphea palustris*

<400> 107
 gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
 tcccggcgct gcatgcaaca ccgatgatgc ttcgacccc cgaagctcct tcggggctgc 120
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 cacaggccac tcgagcttgt gatcgctcct cgctaagggg gcgcctcttc ctcttcgttt 300
 cagtcaaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 360
 cgctgcggcg acctgcgtcg ctccggcgggc tccgggccc ggcgcccagc gaggcccctc 420
 cccgtgcgcg ggcgcgccag catgctgctg tcggcggtga ccacggtctt cggcggtggc 480
 gagaagcagt ggcccatgct ggaccgcaag tccaagcgcc ccgacatgct ggtcgagccc 540
 ctgggctggt accgcatcgt ctacgacggc gtgagcttcc gccagtcgtt ctccatccgc 600
 agctacgaga tcggcgccga ccgcaccgcc tcgatcgaga cgctgatgaa catgtttccag 660
 gagacctccc tgaaccactg caagatcatc ggcctgctga acgacggctt cggccgcacg 720
 cccgagatgt gcaagcgcg cctgatctgg gtcgtgacca agatgcagat cgaggtgaac 780
 cgctacccca cgtggggcg caccatcgag gtcaacacgt ggggtgagcg ctccgggaag 840
 cacggcatgg gccgcgactg gctgatctcc gactgccaca ccggcgagat cctgatccgc 900
 gcgacgagcg tctgggcat gatgaaccag aagaccgcc gcctgtcgaa gatccccctac 960
 gaggtgcgcc aggagatcga gccccagttc gtcgactccg cccccgtgat cgtggacgac 1020
 cgcaagtctc acaagctgga cctgaagacg ggcgacagca tctgcaacgg cctgaccccc 1080
 cgctggacgg acctggacgt gaaccagcac gtcaacaacg tgaagtacat cggctggatc 1140
 ctgcagtcgg tccccaccga ggtgttcgag acgcaggagc tgtgcggcct gaccctggag 1200
 taccgcccg agtgcgggcg cgactccgtg ctggagagcg tcacggccat ggacccctcg 1260
 aaggagggcg accgctccct gtaccagcac ctgctgcgcc tggaggacgg cgcggacatc 1320
 gtgaagggcc gcaccgagt ggcggcccaag aacgcggcg ccaagggcg catcctgacg 1380
 ggcaagacca gcaacggcaa ctcgatctcc tgactcgagt taattaactc gaggcagcag 1440
 cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac 1500
 acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt 1560
 gatcttgtgt gtacgcgctt ttgcgagttg ctgctgctt gtgctatttg cgaataccac 1620
 ccccgatc ccttccctc gtttcatatc gcttgcatcc caaccgcaac ttatctacgc 1680
 tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg 1740
 tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga 1800

026172-005020PC - Sequence Listing_1
 tgcacgggaa gtagtgggat gggaacacaa atggaaagct t

1841

<210> 108
 <211> 1010
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 108
 ggcgcgcca gctgcccac tggagcatgc tgctggccgc gatcaccacc ctgttcctgg 60
 cggccgagaa gcagtggatg atgctggact ggaagcccaa gcgccccgac atgctggtgg 120
 accccttcgg cctggggccgc ttcgtgcagg acggcctggt gttccgcaac aacttcagca 180
 tccgcagcta cgagatcggc gcggaccgca ccgccagcat cgagaccctg atgaaccacc 240
 tgcaggagac cgccctgaac cacgtgaaga gcgtgggcct gctggaggac ggcctgggca 300
 gcacccgcga gatgagcctg cgcaacctga tctgggtggt gaccaagatg caggtggcgg 360
 tggaccgcta cccacactgg ggcgacgagg tgcaggtgag cagctgggcg accgccatcg 420
 gcaagaacgg catgcgccgc gagtggatcg tgaccgactt ccgcaccggc gagaccctgc 480
 tgcgcgccac cagcgtgtgg gtgatgatga acaagctgac ccgccgcac agcaagatcc 540
 ccgaggaggt gtggcacgag atcggcccca gttcatcga cgcgcccccc ctgcccaccg 600
 tggaggacga cggccgcaag ctgaccgct tcgacgagag cagcgccgac ttcacccgca 660
 agggcctgac cccccgtgg agcgacctgg acatcaacca gcacgtgaac aacgtgaagt 720
 acatcggtcg gctgctggag agcgcgcccc ccgagatcca cgagagccac gagatcgcca 780
 gcctgaccct ggagtaccgc cgcgagtgcg gccgcgacag cgtgctgaac agcgccacca 840
 aggtgagcga cagcagccag ctgggcaaga gcgccgtgga gtgcaaccac ctggtgcgcc 900
 tgcagaacgg cggcgagatc gtgaagggcc gcaccgtgtg gcgccccaaag cgcctctgt 960
 acaacgacgg cgccgtggtg gacgtgcccc ccaagaccag ctgactcgag 1010

<210> 109
 <211> 5472
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 109
 ggtacccttt cttgcgtat gacacttcca gcaaaaggta gggcgggctg cgagacggct 60
 tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 120
 atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 180
 aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 240
 cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt 300

026172-005020PC - Sequence Listing_1

cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg	360
gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg	420
gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag	480
aaggacgcca agtggcacct gtacttccag tacaaccgga acgacaccgt ctgggggacg	540
cccttgttct gggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatac	600
gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggg ggtggactac	660
aacaacacct ccggcttctt caacgacacc atcgaccgcg gccagcgctg cgtggccatac	720
tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc	780
tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc	840
gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc	900
caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc	960
gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc	1020
cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc	1080
gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc	1140
gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag	1200
accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcggt ggccctcaac	1260
tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc	1320
aagttctccc tcaacaccga gtaccaggcc aaccgggaga cggagctgat caacctgaag	1380
gccgagccga tcctgaacat cagcaacgcc ggccccctgga gccggttcgc caccaacacc	1440
acgttgacga aggccaaacg ctacaacgtc gacctgtcca acagcaccgg caccctggag	1500
ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac	1560
ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag	1620
gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag	1680
aacccctact tcaccaaccg catgagcggtg aacaaccagc ctttcaagag cgagaacgac	1740
ctgtcctact acaagggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac	1800
gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc	1860
gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag	1920
gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg	1980
tgtgatggac tgttgccgcc aacttgctg ccttgacctg tgaatatccc tgccgctttt	2040
atcaaacagc ctcaagtgtg ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc	2100
ttgtgctatt tgcgaatacc acccccagca tccccctccc tcgtttcata tcgcttgcat	2160
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc	2220
actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg	2280
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga	2340

026172-005020PC - Sequence Listing_1

tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgctct gtcgcacctc	2400
agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg	2460
aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcgggtga	2520
gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca	2580
gccacagccg ctcccacccg ccgctgaacc gacacgtgct tgggcgcctg ccgcctgcct	2640
gccgcatgct tgtgctgggtg aggctgggca gtgctgccat gctgattgag gcttggttca	2700
tcgggtggaa gcttatgtgt gtgctgggct tgcattgccg gcaatgcgca tgggtggcaag	2760
agggcggcag cacttgctgg agctgccgcg gtgcctccag gtggttcaat cgcggcagcc	2820
agagggattt cagatgatcg cgcgtacagg ttgagcagca gtgtcagcaa aggtagcagt	2880
ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa	2940
acacgggcat gccacagcac gggcaccggg gaggtggaatg gcaccaccaa gtgtgtgcga	3000
gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcattcaa cgtaaccatg	3060
agctgatcaa cactgcaatc atcggggcggg cgtgatgcaa gcatgcctgg cgaagacaca	3120
tgggtgtcgg atgctgccgg ctgctgcctg ctgcgcacgc cgttgagttg gcagcaggct	3180
cagccatgca ctggatggca gctgggctgc cactgcaatg tggtaggatag gatgcaagtg	3240
gagcgaatac caaaccctct ggctgcttgc tgggttgcat ggcatcgcac catcagcagg	3300
agcgcattcg aagggaactg ccccatgcac gccatgccaa accggagcgc accgagtgtc	3360
cacactgtca ccaggcccg cagctttgca gaaccatgct catggacgca tgtagcgtg	3420
acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag	3480
gcggcggcag cagagcggcg gcagcagcgg cgggggccac ctttcttgcg gggtcgcgcc	3540
ccagccagcg gtgatgcgct gatcccaaac gaggttcacat tcatttgcat gcctggagaa	3600
gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag	3660
cagcaaaggc gcctccccta ctccgcatcg atgttccata gtgcattgga ctgcatttgg	3720
gtggggcggc cggctgtttc tttcgtgttg caaaacgcgc cagctcagca acctgtcccg	3780
tgggtccccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga	3840
agtaggcgcc ctcttttctg ctgcacctct ctccgtcccg ccactagtat ggccaccgca	3900
tccactttct cggcgttcaa tgcccgctgc ggcgacctgc gtcgctcggc gggctccggg	3960
ccccggcgcc cagcgaggcc cctccccgtg cgcgggcgcg cccagctgcc cgactggagc	4020
cgctgtgta ccgccatcac caccgtgttc gtgaagtcca agcgccccga catgcacgac	4080
cgcaagtcca agcgcgccga catgctgggtg gacagcttcg gcctggagtc caccgtgcag	4140
gacggcctgg tgttccgcca gtccttctcc atccgctcct acgagatcgg caccgaccgc	4200
accgccagca tcgagaccct gatgaaccac ctgcaggaga cctccctgaa ccaactgcaag	4260
agcaccggca tctgctgga cggcttcggc cgcacctgg agatgtgcaa gcgcgacctg	4320
atctgggtgg tgatcaagat gcagatcaag gtgaaccgct accccgcctg gggcgacacc	4380

026172-005020PC - Sequence Listing_1

gtggagatca acacccgctt cagccgcctg ggcaagatcg gcatggggccg cgactggctg	4440
atctccgact gcaacaccgg cgagatcctg gtgcgcgcca ccagcgccta cgccatgatg	4500
aaccagaaga cccgccgcct gtccaagctg ccctacgagg tgcaccagga gatcgtgccc	4560
ctgttcgtgg acagccccgt gatcgaggac tccgacctga aggtgcacaa gttcaagggtg	4620
aagaccggcg acagcatcca gaagggcctg acccccggct ggaacgacct ggacgtgaac	4680
cagcacgtgt ccaacgtgaa gtacatcggc tggatcctgg agagcatgcc caccgagggtg	4740
ctggagaccc aggagctgtg ctccctggcc ctggagtacc gccgcgagtg cggccgcgac	4800
tccgtgctgg agagcgtgac cgccatggac ccagcaagg tgggcgtgcg ctcccagtac	4860
cagcacctgc tgcgcctgga ggacggcacc gccatcgtga acggcgccac cgagtggcgc	4920
cccaagaacg ccggcgccaa cggcgccatc tccaccggca agaccagcaa cggcaactcc	4980
gtgtccatgg actacaagga ccacgacggc gactacaagg accacgacat cgactacaag	5040
gacgacgacg acaagtgact cgaggcagca gcagctcggg tagtatcgac acactctgga	5100
cgctggtcgt gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct	5160
gccgctttta tcaaacagcc tcagtgtgtt tgatcttggt tgtacgcgct tttgcgagtt	5220
gctagctgct tgtgctatatt gcgaatacca ccccagcat ccccttccct cgtttcatat	5280
cgcttgcatc ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg	5340
ctcctgctca ctgccccctg cacagccttg gtttgggctc cgctgtatt ctcttggtac	5400
tgcaacctgt aaaccagcac tgcaatgctg atgcacggga agtagtgagg tgggaacaca	5460
aatggaaagc tt	5472

<210> 110

<211> 5451

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 110

ggtacccttt cttgcgtat gacacttcca gcaaaaggta gggcgggctg cgagacggct	60
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	120
atgggcgctc cgatgccgct ccaggcgag cgctgtttaa atagccaggc ccccgattgc	180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	240
cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt	300
cagtcaaac ccgcaaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg	360
gccggcttcg ccgccaagat cagcgctcc atgacgaacg agacgtccga ccgccccctg	420
gtgcacttca cccccaacaa gggctggatg aacgaccca acggcctgtg gtacgacgag	480
aaggacgcca agtggcacct gtacttccag tacaaccgga acgacaccgt ctgggggacg	540
cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc	600

026172-005020PC - Sequence Listing_1

gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac	660
aacaacacct ccggcttctt caacgacacc atcgaccgcg gccagcgctg cgtggccatc	720
tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc	780
tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc	840
gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc	900
caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc	960
gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc	1020
cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc	1080
gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc	1140
gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag	1200
accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgct ggccctcaac	1260
tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc	1320
aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag	1380
gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc	1440
acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag	1500
ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac	1560
ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag	1620
gtgtccgct cctccttctt cctggaccgc gggaaacagca aggtgaagtt cgtgaaggag	1680
aacccctact tcaccaaccg catgagcgctg aacaaccagc cttcaagag cgagaacgac	1740
ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac	1800
gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc	1860
gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag	1920
gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg	1980
tgtgatggac tgttgccgcc acacttgctg cttgacctg tgaatatccc tgccgctttt	2040
atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc	2100
ttgtgctatt tgcgaatacc acccccagca tccccctccc tcgtttcata tcgcttgcat	2160
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc	2220
actgccccct gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg	2280
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga	2340
tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgccctc gtcgcacctc	2400
agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg	2460
aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga	2520
gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca	2580
gccacagccg ctcccacccg ccgctgaacc gacacgtgct tgggcgcctg ccgcctgcct	2640

026172-005020PC - Sequence Listing_1

gccgcatgct tgtgctggtg aggctgggca gtgctgccat gctgattgag gcttggttca	2700
tcgggtggaa gcttatgtgt gtgctgggct tgcatgccgg gcaatgcgca tgggtggcaag	2760
agggcggcag cacttgctgg agctgccgag gtgcctccag gtggttcaat cgcggcagcc	2820
agagggattt cagatgatcg cggtacagg ttgagcagca gtgtcagcaa aggtagcagt	2880
ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa	2940
acacgggcat gccacagcac gggcaccggg gagtggaaatg gcaccaccaa gtgtgtgcga	3000
gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcatccaa cgtaaccatg	3060
agctgatcaa cactgcaatc atcgggcggg cgtgatgcaa gcatgcctgg cgaagacaca	3120
tgggtgtgcg atgctgccgg ctgctgcctg ctgcgcacgc cgttgagttg gcagcaggct	3180
cagccatgca ctggatggca gctgggctgc cactgcaatg tgggtggatag gatgcaagtg	3240
gagcgaatac caaacctctt ggctgcttgc tgggttgcat ggcatcgac catcagcagg	3300
agcgcatgag aagggaactg ccccatgcac gccatgccaa accggagcgc accgagtgtc	3360
cacactgtca ccaggcccg cagctttgca gaaccatgct catggacgca tgtagcgtg	3420
acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag	3480
gcggcggcag cagagcggcg gcagcagcgg cgggggccac ccttcttgcg gggtcgcgcc	3540
ccagccagcg gtgatgcgt gatcccaaac gagttcacat tcatttgcag gcctggagaa	3600
gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag	3660
cagcaaaggc gcctccccta ctccgcatcg atgttccata gtgcattgga ctgcatttgg	3720
gtggggcggc cggctgtttc tttcgtgttg caaaacgcgc cagctcagca acctgtccc	3780
tgggtccccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga	3840
agtaggcgcc ctctttctg ctgcctctt ctccgtccc cactagtagt gacgttcggg	3900
gtcgccttcc cggccatggg ccgcggtgtc tcccttcccc ggcccagggt cgcggtgcgc	3960
gcccagtcgg cgagtcagg tttggagagc gggcgcgccc ccgactggtc catgctgttc	4020
gccgtgatca ccaccatctt cagcgccgcc gagaagcagt ggaccaacct ggagtggaag	4080
cccaagccca agctgcccc gctgctggac gaccatttc gcctgcacgg cctggtgttc	4140
cgccgcacct tcgccatccg ctctacgag gtgggccccg accgcagcac ctccatcctg	4200
gccgtgatga accacatgca ggaggccacc ctgaaccacg ccaagagcgt gggcatcctg	4260
ggcgacggct tcggcaccac cctggagatg tccaagcgcg acctgatgtg ggtggtgcgc	4320
cgcacccacg tggccgtgga gcgtacccc acctggggcg acaccgtgga ggtggagtgc	4380
tggatcgggc ccagcggcaa caacggcatg cgccgcgact tctggtgcg cgactgcaag	4440
accggcgaga tcctgacctg ctgcacctc ctgagcgtgc tgatgaacac ccgcacccgc	4500
cgcctgagca ccatccccga cgagggtgcgc ggcgagatcg gccccgcctt catcgacaac	4560
gtggccgtga aggacgacga gatcaagaag ctgcagaagc tgaacgactc caccgccgac	4620
tacatccagg gcggcctgac cccccgctgg aacgacctgg acgtgaacca gcacgtgaac	4680

026172-005020PC - Sequence Listing_1

aacctgaagt acgtggcctg ggtgttcgag accgtgcCcg acagcatctt cgagtccac	4740
cacatcagct ccttcaccct ggagtaccgc cgcgagtGca cccgcgactc cgtgctgcgc	4800
agcctgacca ccgtgagcgg cggcagctcc gaggccggcc tgggtgtgca ccacctgctg	4860
cagctggagg gcggcagcga ggtgctgcgc gcccgcaCcg agtggcgccc caagctgacc	4920
gactccttcc gcggcatcag cgtgatcccc gccgagcccc gcgtgatgga ctacaaggac	4980
cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtgactc	5040
gaggcagcag cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg	5100
ttgccgccac acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct	5160
cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg	5220
cgaataccac ccccgacatc cccttccctc gtttcatatc gcttgcatcc caaccgaac	5280
ttatctacgc tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc	5340
acagccttgg tttgggctcc gcctgtattc tcctggtagt gcaacctgta aaccagcact	5400
gcaatgctga tgcacgggaa gtagtgggat gggaacacaa atggaaagct t	5451

<210> 111
 <211> 5454
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 111	
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	60
tccccggcgt gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	120
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	180
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	240
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	300
cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg	360
gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg	420
gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag	480
aaggacgcca agtggcacct gtacttccag tacaaccga acgacaccgt ctgggggacg	540
cccttgcttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccac	600
gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac	660
aacaacacct ccggcttctt caacgacacc atcgaccgc gccagcgctg cgtggccatc	720
tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc	780
tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc	840
gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc	900

026172-005020PC - Sequence Listing_1

caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc	960
gcgttcgcca acgaggggctt cctcggctac cagtacgagt gccccggcct gatcgaggctc	1020
cccaccgagc aggaccccag caagtcctac tgggtgatgt tcattctccat caaccccggc	1080
gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg caccacttc	1140
gaggccttcg acaaccagtc ccgcgtgggtg gacttcggca aggactacta cgccctgcag	1200
accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcggtg ggcctccaac	1260
tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc	1320
aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag	1380
gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc	1440
acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag	1500
ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac	1560
ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag	1620
gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag	1680
aacccctact tcaccaaccg catgagcggtg aacaaccagc ctttcaagag cgagaacgac	1740
ctgtcctact acaagggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac	1800
gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc	1860
gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag	1920
gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggctg	1980
tgtgatggac tgttgccgcc acacttgctg ctttgacctg tgaatatccc tgccgctttt	2040
atcaaacagc ctcaagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc	2100
ttgtgtctatt tgcaataacc acccccagca tccccctccc tcgtttcata tcgcttgcat	2160
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc	2220
actgccccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg	2280
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga	2340
tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgccctct gtcgcacctc	2400
agcgcggcat acaccacaat aaccacctga cgaatgcgtt tggttcttcg tccattagcg	2460
aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga	2520
gctgatggtc gaaacgttca cagcctaggg atatcgaatt ccgcctgcaa cgcaagggca	2580
gccacagccg ctcccacccg ccgctgaacc gacacgtgct tgggcgcctg ccgcctgcct	2640
gccgcatgct tgtgtcgggtg aggctgggca gtgctgccat gctgattgag gcttggttca	2700
tcgggtggaa gcttatgtgt gtgctgggct tgcattgccg gcaatgcgca tgggtggcaag	2760
agggcggcag cacttgctgg agctgccgcg gtgcctccag gtggttcaat cgcggcagcc	2820
agagggattt cagatgatcg cgcgtacagg ttgagcagca gtgtcagcaa aggtagcagt	2880
ttgccagaat gatcggttca gctgttaatc aatgccagca agagaagggg tcaagtgcaa	2940

026172-005020PC - Sequence Listing_1

acacgggcat gccacagcac gggcaccggg gagtggaatg gcaccaccaa gtgtgtgcga	3000
gccagcatcg ccgcctggct gtttcagcta caacggcagg agtcatccaa cgtaaccatg	3060
agctgatcaa cactgcaatc atcgggcggg cgtgatgcaa gcatgcctgg cgaagacaca	3120
tgggtgtgcg atgctgccgg ctgctgcctg ctgcgcacgc cgttgagttg gcagcaggct	3180
cagccatgca ctggatggca gctgggctgc cactgcaatg tggtaggatag gatgcaagtg	3240
gagcgaatac caaacctctt ggctgcttgc tgggttgcac ggcatcgac catcagcagg	3300
agcgcatcg aagggactgg ccccatgcac gccatgccaa accggagcgc accgagtgtc	3360
cacactgtca ccaggcccg c aagctttgca gaaccatgct catggacgca tgtagcgtg	3420
acgtcccttg acggcgctcc tctcgggtgt gggaaacgca atgcagcaca ggcagcagag	3480
gcggcggcag cagagcggcg gcagcagcgg cgggggccac ctttcttgcg gggtcgcgcc	3540
ccagccagcg gtgatgcgct gatcccaaac gagttcacat tcatttgcac gcctggagaa	3600
gcgaggctgg ggcctttggg ctggtgcagc ccgcaatgga atgcgggacc gccaggctag	3660
cagcaaaggc gcctccccta ctccgcatcg atgttccata gtgcattgga ctgcatttgg	3720
gtggggcggc cggtgtttt tttcgtgttg caaaacgcgc cagctcagca acctgtccc	3780
tgggtccccc gtgccgatga aatcgtgtgc acgccgatca gctgattgcc cggctcgcga	3840
agtaggcgcc ctctttctg ctccgctct ctccgtccc cactagtat ggccaccgca	3900
tccactttct cggcgttcaa tgcccgtgc ggcgacctgc gtcgctcggc gggctccggg	3960
ccccggcgcc cagcgaggcc cctccccgtg cgcgggcgcg ccccgactg gtccatgctg	4020
ttcgccgtga tcaccacat cttctccgcc gccgagaagc agtggacca cctggagtgg	4080
aagcccaagc ccaaccccc ccagctgctg gacgaccact tcggcccca cggcctggtg	4140
ttccgcgca ctttcgccat ccgcagctac gaggtgggccc ccgaccgctc caccagcatc	4200
gtggccgtga tgaaccacct gcaggaggcc gccctgaacc acgccaagtc cgtgggcatc	4260
ctgggcgacg gcttcggcac caccctggag atgtccaagc gcgacctgat ctgggtggtg	4320
aagcgcaccc acgtggccgt ggagcgctac cccgcctggg gcgacaccgt ggaggtggag	4380
tgctgggtgg gcgcctccgg caacaacggc cgcgccacg acttcctggt gcgcgactgc	4440
aagaccggcg agatcctgac ccgctgcacc tccctgagcg tgatgatgaa caccgcacc	4500
cgccgcctga gcaagatccc cgaggagggtg cgcggcgaga tcggccccgc cttcatcgac	4560
aacgtggccg tgaaggacga ggagatcaag aagccccaga agctgaacga ctccaccgcc	4620
gactacatcc agggcggcct gacccccgc tggaaacgacc tggacatcaa ccagcacgtg	4680
aacaacatca agtacgtgga ctggatcctg gagaccgtgc ccgacagcat cttcgagagc	4740
caccacatct cctccttcac catcgagtac cgcgcgagt gcaccatgga cagcgtgctg	4800
cagtccctga ccaccgtgag cggcggctcc tccgaggccg gcctggtgtg cgagcacctg	4860
ctgcagctgg agggcggcag cgagggtgctg cgcgccaaga ccgagtggcg cccaagctg	4920
accgactcct tccgcggcat cagcgtgac cccgccgagt ccagcgtgat ggactacaag	4980

026172-005020PC - Sequence Listing_1

gaccacgacg gcgactacaa ggaccacgac atcgactaca aggacgacga cgacaagtga	5040
ctcgaggcag cagcagctcg gatagtatcg acacactctg gacgctggc gtgtgatgga	5100
ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag	5160
cctcagtgtg tttgatcttg tgtgtacgcg cttttgagag ttgctagctg cttgtgctat	5220
ttgcgaatac cccccccagc atcccccttc ctcgtttcat atcgcttgca tcccaaccgc	5280
aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgcccct	5340
cgcacagcct tggtttgggc tccgcctgta ttctcctggt actgcaacct gtaaacagc	5400
actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatggaaa gctt	5454

<210> 112

<211> 2933

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 112

agagagcggg ggtggggttg tgaggtgggg ttgctgacca ggagctcgcg tcgccgagcg	60
cgactcgcac acggtccagt tcccccccc tccgccccaa cgcaagcctc ccatcttgat	120
gcctttcccg ccacctatac tttttcttag ttgctgttaa catccagacc gtcctgaata	180
ataacaatgc cctgtgtcaa gtgcattcct aaaaaaatc tgtcccaacc aacaatccca	240
cctgaaatac caccagccct gccagtgaca ctcttccaat accatctccc tacctccacg	300
cgcaagcgac ccccatgccc gaccaggctc gaaagtgatt tatgacttga gacgagcgag	360
tggcgggcgc gtcgactgcc ttttcatcac gtgccgtacg tcggcgaccg ctagggtctt	420
gcacggcaac gcacggcttc gccaaaccga ccagccagga cctcgactac tctaccgca	480
attcgctca agaagtcgcc aaatgtgcca tacaccattc cttacagcac tgttcaaact	540
tgatgccaat ttgacattc gggttgctcg ttggctgcgc ccacatcggc cgtgagtga	600
gcaggcgagg tcggacacgg aggacgcggc gtcacgcccc gaacgcagcc cgtaactcta	660
catcaacacg acgtgttgcg taatcccgc cggtgcgca tcgtgccaac ccattcgca	720
tggatggctg gaaaatggtg tgccaactgc cctgaggagg gctctcgca aacgggcacg	780
tccctgaaac cgaaactgtg gccttgctgt cggccacgca agcacgtgga ccctaaacac	840
caagaaaatc agtaaacaag gttgacatcc tctacgggcg aattgtttgc ccaacccttc	900
atcgcacact gccattataa tgcatttagc tcggcgacaa gtttagaaaa ggcaggctgc	960
attgttccat ttcgccgtgg cggcggtggg gccatttta cgagggttg gctcccgggc	1020
agcgaccgag ccaggtcgag tccctctcgc ccgtcgacaa tgttgcgaa cccacaagcg	1080
gctaacaaca acttgatggt acctgtacac tgccaattcc ttcttccccg gccgagggtt	1140
acacgtgatg gccatggctt cgcattcagg ccgacttccc attccgactt tccagagggt	1200
ccgcggacgc tgggggttgg ctgcctgagg cccacccttt gttccccgcg tcccgacaaa	1260

026172-005020PC - Sequence Listing_1

cacaattgcg ttacataagg gggagccgcc cccgttcaga gtgcagaaat ctttcactat	1320
attttccagt cgtcagcgaa atcaagtact agtatggcca ccgcatccac tttctcggcg	1380
ttcaatgccc gctgcggcga cctgcgtcgc tcggcgggct ccgggccccg gcgcccagcg	1440
aggccccctc ccgtgcgcgg gcgcgcccc gactgggtcca tgctgttcgc cgtgatcacc	1500
accatcttct ccgcccgcga gaagcagtgg accaacctgg agtggaagcc caagcccaac	1560
cccccccagc tgctggacga ccacttcggc cccacaggcc tgggtgttcg ccgcaccttc	1620
gccatccgca gctacgaggt gggccccgac cgctccacca gcctcgtggc cgtgatgaac	1680
cacctgcagg aggccgccct gaaccacgcc aagtccgtgg gcctcctggg cgacggcttc	1740
ggcaccaccc tggagatgtc caagcgcgac ctgatctggg tgggtgaagc caccacgtg	1800
gccgtggagc gctacccgc ctggggcgac accgtggagg tggagtgtg ggtgggcgcc	1860
tccggcaaca acggccgccg ccacgacttc ctggtgcgcg actgcaagac cggcgagatc	1920
ctgacccgct gcacctcct gagcgtgatg atgaacaccc gcacccgccg cctgagcaag	1980
atccccgagg aggtgcgcgg cgagatcggc cccgccttca tcgacaacgt ggccgtgaag	2040
gacgaggaga tcaagaagcc ccagaagctg aacgactcca ccgccgacta catccagggc	2100
ggcctgaccc cccgctggaa cgacctggac atcaaccagc acgtgaacaa catcaagtac	2160
gtggactgga tcctggagac cgtgcccgc agcatcttcg agagccacca catctcctcc	2220
ttcaccatcg agtaccgccg cgagtgcacc atggacagcg tgctgcagtc cctgaccacc	2280
gtgagcggcg gctcctccga ggccggcctg gtgtgcgagc acctgctgca gctggagggc	2340
ggcagcgagg tgctgcgcgc caagaccgag tggcgcccca agctgaccga ctcttccgc	2400
ggcatcagcg tgatccccgc cgagtccagc gtgatggact acaaggacca cgacggcgac	2460
tacaaggacc acgacatcga ctacaaggac gacgacgaca agtgactcga gtttaattaac	2520
tcgaggcagc agcagctcgg atagtatcga cacactctgg acgctggctg tgtgatggac	2580
tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt atcaaacagc	2640
ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt	2700
tgcaataacc acccccagca tccccctccc tcgtttcata tcgcttgcat cccaaccgca	2760
acttatctac gctgtcctgc tatccctcag cgctgtcctt gctcctgctc actgccccctc	2820
gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg taaaccagca	2880
ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggaaaag ctt	2933

<210> 113

<211> 4817

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 113

026172-005020PC - Sequence Listing_1

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtagaggct gggcagtgtc	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccagggtgt tcaatcgcgg cagccagagg gatttcagat gatcgcgctg acaggttgag	300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcaactga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgcctg gagaagcgag gctggggcct ttgggctggt gcagcccga	1080
atggaatgca ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
ccatagtga ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg	1320
tccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc	1380
ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgttccg ccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgccgcctg tcctggctgg ccaccaccgg	1560
cgtagccctg gccgccgtg tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaagggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggcccg atgcccagc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc cccccgcgac atcgccgagg agctgggcgg	2040

026172-005020PC - Sequence Listing_1

cgagtggggcc gaccgcttcc tgggtgctgta cggcatcgcc gcccccgact cccagcgcac	2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcgatagat	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcggt tcatacgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgctcc tgcactgc ccctgcaca gccttggtt gggctccgcc	2460
tgtattctcc tggtagtga acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatccc cgtctcgaac agagcgcgca gaggaacgt	2580
gaaggctctg cctctgctgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
gaattccttt cttgctgat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacgg cgcgccatgg tggccgccgc cgctccagc gccttcttcc	3120
ccgtgcccgc ccccggcgc tccccaaagc ccggcaagtt cggcaactgg ccctccagcc	3180
tgagccctc cttcaagccc aagtccatcc ccaacggcgg cttccagggtg aaggccaacg	3240
acagcgccca cccaaggcc aacggctccg ccgtgagcct gaagagcggc agcctgaaca	3300
cccaggagga cactcctcc agccccccc cccgcacctt cctgcaccag ctgcccagct	3360
ggagccgcct gctgaccgcc atcaccaccg tgttcgtgaa gtccaagcgc cccgacatgc	3420
acgaccgcaa gtccaagcgc cccgacatgc tggtagacag cttcggcctg gagtccaccg	3480
tgcaggacgg cctggtgttc cgccagtcct tctccatccg ctctacagag atcggcaccg	3540
accgcaccgc cagcatcgag acctgatga accacctgca ggagacctc ctgaaccact	3600
gcaagagcac cggcatcctg ctggacggct tcggccgcac cctggagatg tgcaagcgcg	3660
acctgatctg ggtggtgatc aagatgcaga tcaaggtgaa ccgctacccc gcctggggcg	3720
acaccgtgga gatcaacacc cgcttcagcc gcctgggcaa gatcggcatg ggccgcgact	3780
ggctgatctc gactgcaac accggcgaga tcctggtgcg cgccaccagc gcctacgcca	3840
tgatgaacca gaagacccgc cgctgtcca agctgcccta cgagggtgcac caggagatcg	3900
tgccccgtgt cgtggacagc cccgtgatcg aggactccga cctgaagggtg cacaagttca	3960
aggtgaagac cggcgacagc atccagaagg gcctgacccc cggctggaac gacctggacg	4020
tgaaccagca cgtgtccaac gtgaagtaca tcggctggat cctggagagc atgcccaccg	4080

026172-005020PC - Sequence Listing_1

aggtgctgga gacccaggag ctgtgctccc tggccctgga gtaccgccgc gagtgcggcc	4140
gcgactccgt gctggagagc gtgaccgcca tggaccccag caaggagggc gtgcgctccc	4200
agtaccagca cctgctgcgc ctggaggacg gcaccgccat cgtgaacggc gccaccgagt	4260
ggcgcccca gaacgccggc gccaacggcg ccatctccac cggcaagacc agcaacggca	4320
actccgtgtc catggactac aaggaccacg acggcgacta caaggaccac gacatcgact	4380
acaaggacga cgacgacaag tgactcgagg cagcagcagc tcggatagta tcgacacact	4440
ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata	4500
tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc	4560
gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccct tccctcgttt	4620
catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc	4680
tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct	4740
ggctactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga	4800
acacaaatgg aaagctt	4817

<210> 114

<211> 4665

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 114

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcttgcgc atgcttgtgc tggtagaggct gggcagtgtc	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgttgtgtgct gggcttgcac	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccagggtgt tcaatcgcg cagccagagg gatttcagat gatcgcgct acaggttgag	300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
gccaaaccgg agcgaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gtcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcggg	960

026172-005020PC - Sequence Listing_1

gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccgc	1080
atggaatgcg ggaccgccag gctagcagca aaggcgccct ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg	1320
tcccgccctc agaatatcaa tgatcgagca ggacggccct caccgccggct cccccgccgc	1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgcctgtccg ccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgccctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccccgcg agaagggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac ccgcccacct gccccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcg cgcacccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggccgc atgcccgcg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgcac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tgggtgctga cggcatcgcc gcccccgact ccagcgc	2100
cgcttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcgtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgctcc tgctactgc ccctgcaca gccttggttt gggctccgcc	2460
tgtattctcc tggactgca acctgtaaac cagactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggctctg cctctgtcgc acctcagcgc ggcatacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cggtggcgag	2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatt	2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgt gcattgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000

026172-005020PC - Sequence Listing_1

cacaggccac tcgagcttgt gatcgactc cgctaagggg ggcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatgacg ttctgggtcg ccctcccggc catgggccgc	3120
gggtgtctccc ttccccggcc caggggtcgc gtgcgcgcc agtcggcgag tcaggttttg	3180
gagagcgggc gcgcccagct gcccgactgg agccgcctgc tgaccgcat caccaccgtg	3240
ttcgtgaagt ccaagcgccc cgacatgcac gaccgcaagt ccaagcgccc cgacatgctg	3300
gtggacagct tcggcctgga gtccaccgtg caggacggcc tgggtgttcg ccagtccttc	3360
tccatccgct cctacgagat cggcaccgac cgcaccgcca gcatcgagac cctgatgaac	3420
cacctgcagg agacctcct gaaccactgc aagagcaccg gcatcctgct ggacggcttc	3480
ggccgcaccc tggagatgtg caagcgcgac ctgatctggg tggatgatcaa gatgcagatc	3540
aaggatgaacc gctaccccg cctggggcgac accgtggaga tcaacacccg cttcagccgc	3600
ctgggcaaga tcggcatggg ccgcgactgg ctgatctccg actgcaacac cggcgagatc	3660
ctggtgcgcg ccaccagcgc ctacgccatg atgaaccaga agaccgcgcg cctgtccaag	3720
ctgccctacg aggtgcacca ggagatcgtg cccctgttcg tggacagccc cgtgatcgag	3780
gactccgacc tgaaggtgca caagttcaag gtgaagaccg gcgacagcat ccagaagggc	3840
ctgacccccc gctggaacga cctggacgtg aaccagcacg tgtccaacgt gaagtacatc	3900
ggctggatcc tggagagcat gccacccgag gtgctggaga cccaggagct gtgctccctg	3960
gccctggagt accgcccga gtgcggccgc gactccgtgc tggagagcgt gaccgcatg	4020
gacccagca aggtgggcgt gcgctcccag taccagcacc tgctgcgcct ggaggacggc	4080
accgcatcg tgaacggcgc caccgagtgg cgcaccaaga acgcccgcgc caacggcgcc	4140
atctccaccg gcaagaccag caacggcaac tccgtgtcca tggactacaa ggaccacgac	4200
ggcgactaca aggaccacga catcgactac aaggacgacg acgacaagtg actcgaggca	4260
gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg	4320
ccacacttgc tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt	4380
gtttgatctt gtgtgtacgc gcttttgca gtgtgtagct gcttgtgcta tttgcgaata	4440
ccacccccag catccccctt cctcgtttca tatcgcttgc atcccaaccg caacttatct	4500
acgctgtcct gctatccctc agcgtgctc ctgctcctgc tcaactgcccc tcgcacagcc	4560
ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg	4620
ctgatgcacg ggaagtagtg ggatgggaac acaaatggaa agctt	4665

<210> 115

<211> 4668

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 115

026172-005020PC - Sequence Listing_1

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgctg tggtagaggct gggcagtgct	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgctg acaggttgag	300
cagcagtgct agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggttcagcc atgcaactga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgcacccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccga	1080
atggaatgcg ggaccgccag gctagcagca aaggcgccct ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccgctgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgcctctct ttctgctcgc cctctctccg	1320
tccgcctctc agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc	1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgcctgtccg cccaggggcg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgccctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcg cgcacccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggccgc atgcccagc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcacg gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040

026172-005020PC - Sequence Listing_1

cgagtggggcc gaccgcttcc tgggtgctgta cggcatcgcc gccccgact cccagcgcac	2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcgatagat	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctccctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc	2460
tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggctctg cctctgtgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata	2760
gaattccttt cttgcgtat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag	3120
cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgtcg	3180
gcgcttcgtg ggcgcgccc gctgcccga tggagccgcc tgctgaccgc catcaccacc	3240
gtgttcgtga agtccaagcg ccccgacatg cagcaccgca agtccaagcg ccccgacatg	3300
ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctgggtgtt ccgccagtcc	3360
ttctccatcc gtcctacga gatcggcacc gaccgcaccg ccagcatcga gacctgatg	3420
aaccacctgc aggagacct cctgaaccac tgcaagagca ccggcatcct gctggacggc	3480
ttcggccgca ccctggagat gtgcaagcgc gacctgatct ggggtggtgat caagatgcag	3540
atcaagggtga accgctaccc cgcttggggc gacaccgtgg agatcaaac ccgcttcagc	3600
cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag	3660
atcctgggtgc gcgccaccag cgcctacgcc atgatgaacc agaagacccg ccgcctgtcc	3720
aagctgccct acgaggtgca ccaggagatc gtgccccgtg tcgtggacag ccccgatgc	3780
gaggactccg acctgaaggt gcacaagttc aagggtgaaga ccggcgacag catccagaag	3840
ggcctgaccc ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac	3900
atcggctgga tcctggagag catgcccacc gaggtgctgg agaccagga gctgtgctcc	3960
ctggccctgg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc	4020
atggacccca gcaagggtgg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac	4080

026172-005020PC - Sequence Listing_1

ggcaccgcc	tcgtgaacgg	cgccaccgag	tgccgcccc	agaacgccgg	cgccaacggc	4140
gccatctcca	ccggcaagac	cagcaacggc	aactccgtgt	ccatggacta	caaggaccac	4200
gacggcgact	acaaggacca	cgacatcgac	tacaaggacg	acgacgacaa	gtgactcgag	4260
gcagcagcag	ctcggatagt	atcgacacac	tctggacgct	ggtcgtgtga	tggactgttg	4320
ccgccacact	tgtgccttg	acctgtgaat	atccctgccg	cttttatcaa	acagcctcag	4380
tgtgtttgat	cttgtgtgta	cgcgcttttg	cgagttgcta	gctgcttggt	ctatttgca	4440
ataccacccc	cagcatcccc	ttccctcggt	tcataatcgt	tgcattccaa	ccgcaactta	4500
tctacgctgt	cctgctatcc	ctcagcgtcg	ctcctgctcc	tgctcactgc	ccctcgacaa	4560
gccttggttt	gggctccgcc	tgtattctcc	tggtactgca	acctgtaaac	cagcactgca	4620
atgctgatgc	acgggaagta	gtgggatggg	aacacaaatg	gaaagctt		4668

<210> 116

<211> 4668

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 116

ggtacccgcc	tgcaacgcaa	gggcagccac	agccgctccc	acccgccgct	gaaccgacac	60
gtgcttgggc	gcctgccgcc	tgcctgccgc	atgcttggtc	tggtgaggct	gggcagtgtc	120
gccatgctga	ttgaggcttg	gttcacggg	tggaagctta	tgtgtgtgct	gggcttgcat	180
gccgggcaat	gcgcatggtg	gcaagagggc	ggcagcactt	gctggagctg	ccgcggtgcc	240
tccagggtgt	tcaatcgcg	cagccagagg	gatttcagat	gatcgcgcgt	acaggttgag	300
cagcagtgtc	agcaaaggta	gcagtttgcc	agaatgatcg	gttcagctgt	taatcaatgc	360
cagcaagaga	aggggtcaag	tgcaaacacg	ggcatgccac	agcacgggca	ccggggagtg	420
gaatggcacc	accaagtgtg	tgcgagccag	catcgccgcc	tggtgttttc	agctacaacg	480
gcaggagtca	tccaacgtaa	ccatgagctg	atcaacactg	caatcatcgg	gcgggcgtga	540
tgcaagcatg	cctggcgaag	acacatggtg	tgcggatgct	gccggctgct	gcctgctgcg	600
cacgccgttg	agttggcagc	aggctcagcc	atgactgga	tggcagctgg	gctgccactg	660
caatgtggtg	gataggatgc	aagtggagcg	aataccaaac	cctctggctg	cttgtctggg	720
tgcattggcat	cgcacccatc	gcaggagcgc	atgcgaaggg	actggcccca	tgcacgccat	780
gccaaaccgg	agcgcaccga	gtgtccacac	tgtcaccagg	cccgcaagct	ttgcagaacc	840
atgctcatgg	acgcatgtag	cgctgacgtc	ccttgacggc	gctcctctcg	gggtgtggaa	900
acgcaatgca	gcacaggcag	cagaggcggc	ggcagcagag	cggcggcagc	agcggcgggg	960
gccacccttc	ttgcgggggtc	gcgccccagc	cagcgggtgat	gcgctgatcc	caaacgagtt	1020
cacattcatt	tgcattgcctg	gagaagcgag	gctggggcct	ttgggctggg	gcagcccgc	1080
atggaatgcg	ggaccgccag	gctagcagca	aaggcgccctc	ccctactccg	catcgatgtt	1140

026172-005020PC - Sequence Listing_1

ccatagtgca ttggactgca tttgggtggg gcgccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccggtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaaagtag gcgcctctct ttctgctcgc cctctctccg	1320
tcccgctctt agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc	1380
ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgtccg cccagggccg ccccggtctg ttctggaaga ccgacctgtc	1500
cggcgccttg aacgagctgc aggacgaggc cgccgcctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
ggcgagggtg cccggccagg acctgctgtc ctccacctg gccccgccg agaaggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgcccgcct	1860
gaaggccgc atgcccgcg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc ccccgcgac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tgggtgctgta cggcatcgcc gccccgact cccagcgc	2100
cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacct tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgctcc tgctcactgc ccctcgaca gccttggttt gggctccgcc	2460
tgtattctcc tgggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgccg cgttggcgag	2700
gtggcagggt acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatggcc accgcattca ctttctcggc gttcaatgcc	3120
cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgccagc gagggccctc	3180

026172-005020PC - Sequence Listing_1

```

cccgtgcgcg ggcgcgcccc gctgcccgcac tggagccgcc tgctgaccgc catcaccacc 3240
gtgttcgtga agtccaagcg ccccgacatg cacgaccgca agtccaagcg ccccgacatg 3300
ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctgggtgtt ccgccagtcc 3360
ttctccatcc gtcctacga gatcggcacc gaccgcaccg ccagcatcga gaccctgatg 3420
aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480
ttcggccgca ccctggagat gtgcaagcgc gacctgatct ggggtggatg caagatgcag 3540
atcaagggtga accgctaccc cgcctggggc gacaccgtgg agatcaaacac ccgcttcagc 3600
cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag 3660
atcctgggtgc gcgccaccag cgcctacgcc atgatgaacc agaagaccg ccgcctgtcc 3720
aagctgccct acgaggtgca ccaggagatc gtgcccctgt tcgtggacag ccccgatgac 3780
gaggactccg acctgaaggt gcacaagttc aagggtgaaga ccggcgacag catccagaag 3840
ggcctgaccc ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac 3900
atcggctgga tcctggagag catgcccacc gaggtgctgg agaccagga gctgtgctcc 3960
ctggccctgg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc 4020
atggacccca gcaaggtggg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac 4080
ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc 4140
gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac 4200
gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgaaa gtgactcgag 4260
gcagcagcag ctcgatagt atcgacacac tctggacgct ggtcgtgtga tggactgtg 4320
ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag 4380
tgtgtttgat ctgtgtgtga cgcgcttttg cgagttgcta gctgcttggt ctatttgca 4440
ataccacccc cagcatcccc ttccctcgtt tcatatcgtt tgcattccaa ccgcaactta 4500
tctacgctgt cctgctatcc ctacgcgctg ctctgctcc tgctcactgc cctcgcaca 4560
gccttggttt gggctccgcc tgtattctcc tggtagtgca acctgtaaac cagcactgca 4620
atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt 4668

```

<210> 117

<211> 4656

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 117

```

ggtaccgcgc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtaggct gggcagtgc 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180

```

026172-005020PC - Sequence Listing_1

gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccagggtggt tcaatcgcg cagccagagg gatttcagat gatcgcgctg acagggttgag	300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctga	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagagggcgc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccga	1080
atggaatgcg ggaccgccag gctagcagca aaggcgccct ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaaagtag gcgcctctct ttctgctcgc cctctctccg	1320
tcccgctctt agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc	1380
ctgggtggag cgctgttctg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgcgcg agaagggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcg cgcacccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggcccg atgcccagc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggttctatc gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgcac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tggctgtgta cggcatcgcc gccccgact cccagcgcat	2100
cgcttctac cgctgtgtg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220

026172-005020PC - Sequence Listing_1

acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc	2340
ttccctcggtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgctg ctcttgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc	2460
tgtattctcc tgggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggctctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata	2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatggct tccgcggcat tcaccatgtc ggcgtgcccc	3120
gcgatgactg gcagggcccc tggggcacgt cgctccggac ggccagtcgc caccgcctg	3180
agggggcgcg ccccgactg gagccgcctg ctgaccgcca tcaccaccgt gttcgtgaag	3240
tccaagcgcc ccgacatgca cgaccgcaag tccaagcgcc ccgacatgct ggtggacagc	3300
ttcggcctgg agtccaccgt gcaggacggc ctgggtgttc gccagtcctt ctccatccgc	3360
tcctacgaga tcggcaccga ccgcaccgcc agcatcgaga ccctgatgaa ccacctgcag	3420
gagacctccc tgaaccactg caagagcacc ggcattctgc tggacggctt cggccgcacc	3480
ctggagatgt gcaagcgcg cctgatctgg gtggtgatca agatgcagat caaggtgaac	3540
cgctaccccc cctggggcga caccgtggag atcaacaccc gcttcagccg cctgggcaag	3600
atcggcatgg gccgcgactg gctgatctcc gactgcaaca ccggcgagat cctggtgcgc	3660
gccaccagcg cctacgccat gatgaaccag aagacccgcc gcctgtccaa gctgccctac	3720
gaggtgcacc aggagatcgt gcccctgttc gtggacagcc ccgtgatcga ggactccgac	3780
ctgaaggtgc acaagttcaa ggtgaagacc ggcgacagca tccagaaggg cctgaccccc	3840
ggctggaacg acctggacgt gaaccagcac gtgtccaacg tgaagtacat cggctggatc	3900
ctggagagca tgcccaccga ggtgctggag acccaggagc tgtgctccct ggccctggag	3960
taccgcccgc agtgccggccg cgactccgtg ctggagagcg tgaccgccat ggaccccagc	4020
aaggtgggcg tgcgctccca gtaccagcac ctgctgcgcc tggaggacgg caccgccatc	4080
gtgaacggcg ccaccgagtg gcgccccaa aacgccggcg ccaacggcgc catctccacc	4140
ggcaagacca gcaacggcaa ctccgtgtcc atggactaca aggaccacga cggcgactac	4200
aaggaccacg acatcgacta caaggacgac gacgacaagt gactcgaggc agcagcagct	4260

026172-005020PC - Sequence Listing_1

cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg	4320
ctgccttgac ctgtgaatat ccctgccgct tttatcaaac agcctcagtg tgtttgatct	4380
tgtgtgtacg cgctttttgcg agttgctagc tgcttgtgct atttgcgaat accaccccca	4440
gcatccccctt ccctcgtttc atatcgcttg catcccaacc gcaacttata tacgctgtcc	4500
tgctatccct cagcgtgct cctgctcctg ctactgccc ctgcacagc cttggtttgg	4560
gctccgctg tatttctctg gtactgcaac ctgtaaacca gcaactgcaat gctgatgcac	4620
gggaagtagt gggatgggaa cacaaatgga aagctt	4656

<210> 118

<211> 4721

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 118

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtagaggct gggcagtgct	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtagt tcaatcgcgg cagccagagg gatttcagat gatcgcgctg acaggttgag	300
cagcagtgct agcaaaggta gcagtttgcc agaataatgc gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctga	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcaactga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgcat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgct ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcgggggc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggg gcagcccgc	1080
atggaatgca ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg	1320

026172-005020PC - Sequence Listing_1

tcccgccctct agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc	1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgcctgtccg cccagggccg ccccgctgtg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgccctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgagggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaagggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcg cgcacccgc atggaggccg gcctgggtga	1800
ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgccccct	1860
gaaggccgc atgcccgcg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggttcatc gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tgggtctgta cggcatcgcc gccccgact cccagcgcat	2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgtctcc tgctcactgc ccctcgcaca gccttgggtt gggctccgcc	2460
tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggctctc cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacgg cgcgccatgg ccaccaccag cctggcctcc gccttctgct	3120
ccatgaaggc cgtgatgctg gcccgcgacg gccgcggcat gaagccccgc agctccgacc	3180
tgacgtgctg cgccggcaac gccccacct ccctgaagat gatcaacggc accaagttca	3240
gctacaccga gagcctgaag cgctgcccc actggtccat gctgttcgcc gtgatcacca	3300
ccatcttcag cgccgccgag aagcagtgga ccaacctgga gtggaagccc aagcccaagc	3360

026172-005020PC - Sequence Listing_1

tgccccagct gctggacgac cacttcggcc tgcacggcct ggtgttccgc cgcaccttcg	3420
ccatccgctc ctacgaggtg ggccccgacc gcagcacctc catcctggcc gtgatgaacc	3480
acatgcagga ggccaccctg aaccacgcca agagcgtggg catcctgggc gacggcttcg	3540
gcaccaccct ggagatgtcc aagcgcgacc tgatgtgggt ggtgcgccgc acccacgtgg	3600
ccgtggagcg ctacccccacc tggggcgaca ccgtggaggt ggagtgtctg atcggcgcca	3660
gcggcaacaa cggcatgcgc cgcgacttcc tgggtgcgcga ctgcaagacc ggcgagatcc	3720
tgacccgctg cacctccctg agcgtgtctga tgaacacccg caccgcccgc ctgagcacca	3780
tccccgacga ggtgcgcggc gagatcggcc ccgccttcat cgacaacgtg gccgtgaagg	3840
acgacgagat caagaagctg cagaagctga acgactccac cgccgactac atccagggcg	3900
gcctgacccc ccgctggaac gacctggacg tgaaccagca cgtgaacaac ctgaagtacg	3960
tggcctgggt gttcgagacc gtgcccgaca gcatcttcga gtcccaccac atcagctcct	4020
tcaccctgga gtaccgccgc gagtgcaccc gcgactccgt gctgcgcagc ctgaccaccg	4080
tgagcggcgg cagctccgag gccggcctgg tgtgcgacca cctgctgcag ctggagggcg	4140
gcagcgaggt gctgcgcgcc cgcaccgagt ggcgccccaa gctgaccgac tccttccgcg	4200
gcatcagcgt gatccccgcc gagccccgcg tgatggacta caaggaccac gacggcgact	4260
acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgatgactc gaggcagcag	4320
cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac	4380
acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt	4440
gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac	4500
ccccagcatc cccttccctc gtttcataatc gcttgcattc caaccgcaac ttatctacgc	4560
tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgccccctgc acagccttgg	4620
tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga	4680
tgcacgggaa gtagtgggat gggaaacaaa atggaaagct t	4721

<210> 119

<211> 4650

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 119

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtagggct gggcagtgtc	120
gccatgctga ttgaggcttg gttcatcggg tggaaagtta tgtgtgtgct gggcttgcac	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag	300

026172-005020PC - Sequence Listing_1

cagcagtggtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctga	540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcaactga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc cttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcag gctggggcct ttgggctggg gcagcccgc	1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgcccctct ttctgctcgc cctctctccg	1320
tcccgccctc agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc	1380
ctgggtggag cgctgtttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgccctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggg gaccgaggcc ggccgcgact ggctgctgct	1620
gggagagggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaagggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctgggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggcccgat atgcccagc gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcacg gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tgggtgctga cggcatcgcc gccccgact cccagcgcat	2100
cgcttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcgatagat	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340

026172-005020PC - Sequence Listing_1

ttccctcggtt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgctg ctctgctcc tgctcactgc ccctcgcaaca gccttggttt gggctccgcc	2460
tgtattctcc tggtagtgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgct	2580
gaagggtctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata	2760
gaattccttt cttgctctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccaggggcgag cgctgtttta atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatgacg ttcggggctg ccctcccggc catgggcccgc	3120
ggtgtctccc ttccccggcc cagggtcgcg gtgcgcgccc agtcggcgag tcaggttttg	3180
gagagcgggc gcgccccga ctggtccatg ctgttcgccc tgatcaccac catcttcagc	3240
gccgccgaga agcagtgagc caacctggag tggaaagcca agcccaagct gcccagctg	3300
ctggacgacc acttcggcct gcacggcctg gtgttcgccc gcaccttcgc catccgctcc	3360
tacgaggtgg gccccgaccg cagcacctcc atcctggccg tgatgaacca catgcaggag	3420
gccaccctga accacgcca gagcgtgggc atcctgggcg acggcttcgg caccaccctg	3480
gagatgtcca agcgcgacct gatgtgggtg gtgcgcccga cccacgtggc cgtggagcgc	3540
tacccacact ggggcgacac cgtggaggtg gagtgctgga tcggcgccag cggcaacaac	3600
ggcatgcgcc gcgacttcct ggtgcgcgac tgcaagaccg gcgagatcct gaccgcgtgc	3660
acctccctga gcgtgctgat gaacaccgcg acccgccgcc tgagcaccat ccccgacgag	3720
gtgcgcggcg agatcgggcc cgcccttcac gacaacgtgg ccgtgaagga cgacgagatc	3780
aagaagctgc agaagctgaa cgactccacc gccgactaca tccaggggcg cctgaccccc	3840
cgttggaacg acctggacgt gaaccagcac gtgaacaacc tgaagtacgt ggcctgggtg	3900
ttcgagaccg tgccccgacag catcttcgag tcccaccaca tcagctcctt caccctggag	3960
taccgcgcg agtgcacccg cgactccgtg ctgcgcagcc tgaccaccgt gagcggcggc	4020
agctccgagg ccggcctggt gtgcgaccac ctgctgcagc tggagggcgg cagcgaggtg	4080
ctgcgcgccc gcaccgagtg gcgccccaa ctgaccgact ccttcgcgcg catcagcgtg	4140
atccccgcg agccccgcgt gatggactac aaggaccacg acggcgacta caaggaccac	4200
gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata	4260
gtatcgacac actctggacg ctggtcgtgt gatggactgt tgccgccaca cttgtgcct	4320
tgacctgtga atatccctgc cgctttttatc aaacagcctc agtgtgtttg atcttgtgtg	4380

026172-005020PC - Sequence Listing_1

tacgcgcttt	tgcgagttgc	tagctgcttg	tgctatttgc	gaataccacc	cccagcatcc	4440
ccttccctcg	tttcatatcg	cttgcacccc	aaccgcaact	tatctacgct	gtcctgctat	4500
ccctcagcgc	tgctcctgct	cctgctcact	gccccctgca	cagccttggg	ttgggctccg	4560
cctgtattct	cctgggtactg	caacctgtaa	accagcactg	caatgctgat	gcacgggaag	4620
tagtgggatg	ggaacacaaa	tggaagctt				4650

<210> 120

<211> 4653

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 120

ggtaccgcgc	tgcaacgcaa	gggcagccac	agccgctccc	acccgccgct	gaaccgacac	60
gtgcttgggc	gcctgccgcc	tgcttgcgc	atgcttgtgc	tggtgaggct	gggcagtgtc	120
gccatgctga	ttgaggcttg	gttcatcggg	tggaagctta	tgtgtgtgct	gggcttgcac	180
gccgggcaat	gcgcatgggtg	gcaagagggc	ggcagcactt	gctggagctg	ccgcggtgcc	240
tccagggtgt	tcaatcgcg	cagccagagg	gatttcagat	gatcgcgct	acaggttgag	300
cagcagtgtc	agcaaaggta	gcagtttgcc	agaatgatcg	gttcagctgt	taatcaatgc	360
cagcaagaga	aggggtcaag	tgcaaacacg	ggcatgccac	agcacgggca	ccggggagtg	420
gaatggcacc	accaagtgtg	tgcgagccag	catcgccgcc	tggtgtttc	agctacaacg	480
gcaggagtca	tccaacgtaa	ccatgagctg	atcaacactg	caatcatcgg	gcgggcgtga	540
tgcaagcatg	cctggcgaag	acacatgggtg	tgcggatgct	gccggctgct	gcctgctgcg	600
cacgccgttg	agttggcagc	aggctcagcc	atgcactgga	tggcagctgg	gctgccactg	660
caatgtgggtg	gataggatgc	aagtggagcg	aataccaaac	cctctggctg	cttgcctgggt	720
tgcatggcat	cgcaccatca	gcaggagcgc	atgcgaaggg	actggcccca	tgcacgccat	780
gccaaaccgg	agcgcaccga	gtgtccacac	tgtcaccagg	cccgcgaagct	ttgcagaacc	840
atgctcatgg	acgcatgtag	cgctgacgtc	ccttgacggc	gctcctctcg	gggtgtgggaa	900
acgcaatgca	gcacaggcag	cagaggcggc	ggcagcagag	cggcggcagc	agcggcgggg	960
gccacccttc	ttgcgggggtc	gcgccccagc	cagcgggtgat	gcgctgatcc	caaacgagtt	1020
cacattcatt	tgcatgcctg	gagaagcgag	gctggggcct	ttgggctggg	gcagccccgca	1080
atggaatgcg	ggaccgccag	gctagcagca	aaggcgctc	ccctactccg	catcgatgtt	1140
ccatagtgca	ttggactgca	tttgggtggg	gcggccggct	gtttctttcg	tgttgcaaaa	1200
cgcgccagct	cagcaacctg	tcccgtgggt	cccccgctgc	gatgaaatcg	tgtgcacgcc	1260
gatcagctga	ttgcccggt	cgcgaagtag	gcgccctcct	ttctgctcgc	cctctctccg	1320
tcccgctct	agaatatcaa	tgatcgagca	ggacggcctc	cacgccggct	ccccgcggc	1380
ctgggtggag	cgcctgttcg	gctacgactg	ggcccagcag	accatcggct	gctccgacgc	1440

026172-005020PC - Sequence Listing_1

cgccgtgttc cgctgtccg cccagggccg ccccgctgtg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgccctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggg gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctgggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggcccg c atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcctc gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040
cgagtgggccc gaccgcttcc tgggtgctgta cggcatcgcc gccccgact cccagcgcct	2100
cgcttctac cgctgtctgg acgagttctt ctgacaattg gcagcagcag ctcgatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgcga ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgtctcc tgctcactgc ccctcgaca gccttggttt gggctccgcc	2460
tgtattctcc tggtagtga acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
gaagggtctg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt ctctgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatctc	2760
gaattccttt cttgcgtat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag	3120
cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgtcg	3180
gcgcttcgtg ggcgcgcccc cgactgggtc atgctgttcg ccgtgatcac caccatcttc	3240
agcgcgccc agaagcagtg gaccaacctg gagtggaaag ccaagcccaa gctgccccag	3300
ctgctggacg accacttcgg cctgcacggc ctgggtgttc gccgcacctt cgccatccgc	3360
tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag	3420
gaggccaccc tgaaccacgc caagagcgtg ggcacctgg gcgacggctt cggcaccacc	3480

026172-005020PC - Sequence Listing_1

ctggagatgt ccaagcgcga cctgatgtgg gtggtgccc gcacccacgt ggccgtggag	3540
cgctacccca cctggggcga caccgtggag gtggagtgtt ggatcggcgc cagcggcaac	3600
aacggcatgc gccgcgactt cctggtgccc gactgcaaga ccggcgagat cctgacccgc	3660
tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac	3720
gaggtgccc gcgagatcgg ccccgccctt atcgacaacg tggccgtgaa ggacgacgag	3780
atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc	3840
ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg	3900
gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcacctg	3960
gagtaccgcc gcgagtgcac ccgcgactcc gtgctgccc gcctgaccac cgtgagcggc	4020
ggcagctccg aggccggcct ggtgtgccc cacctgctgc agctggaggg cggcagcag	4080
gtgctgccc cccgcaccga gtggcgcccc aagctgaccg actccttcg cggcatcagc	4140
gtgatcccc cggagcccc cgtgatggac tacaaggacc acgacggcga ctacaaggac	4200
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg	4260
atagtatcga cacactctgg acgctggctg tgtgatggac tgttgcccgc acacttgctg	4320
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt	4380
gtgtacgccc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca	4440
tccccctccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc	4500
tatccctcag cgctgctcct gctcctgctc actgccccct gcacagcctt ggtttgggct	4560
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg	4620
aagtagtggg atgggaacac aaatggaaag ctt	4653

<210> 121

<211> 4653

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 121

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgctg tggtagggct gggcagtgct	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtggt tcaatcgagg cagccagagg gatttcagat gatcgcgctg acaggttgag	300
cagcagtgtc agcaaaggta gcagtttgcc agaattgatc gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480

026172-005020PC - Sequence Listing_1

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctga	540
tgcaagcatg cctggcgaag acacatgggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcaactgga tggcagctgg gctgccactg	660
caatgtgggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgtcatggt acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcggg	960
gccacccttc ttgcgggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcag gctggggcct ttgggctgggt gcagcccgc	1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgcccctcct ttctgctcgc cctctctccg	1320
tcccgccctc agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc	1380
ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgccctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtgggt gaccgaggcc ggccgcgact ggctgctgct	1620
ggcgaggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaagggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctgggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggccgc atgcccgcg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgcac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tgggtgctga cggcatgcc gccccgact cccagcgc	2100
cgccttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcgtt tcatatcgct tgcacccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgctcc tgctactgc ccctgcaca gccttgggtt gggctccgcc	2460
tgtattctcc tggtagtgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520

026172-005020PC - Sequence Listing_1

gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatata	2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttta atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcaaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc	3120
cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gagggccctc	3180
cccgtgcgcy ggcgcgcccc cgactggctc atgctgttcg ccgtgatcac caccatcttc	3240
agcgcgcccg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag	3300
ctgctggacg accacttcgg cctgcacggc ctgggtgttc gccgcacctt cgccatccgc	3360
tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag	3420
gaggccaccc tgaaccacgc caagagcgtg ggcacctcgg gcgacggctt cggcaccacc	3480
ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag	3540
cgctacccca cctggggcga caccgtggag gtggagtgtt ggatcggcgc cagcggcaac	3600
aacggcatgc gccgcgactt cctgggtgcgc gactgcaaga ccggcgagat cctgacccgc	3660
tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac	3720
gagggtgcgc gcgagatcgg ccccgccctt atcgacaacg tggccgtgaa ggacgacgag	3780
atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc	3840
ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg	3900
gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcaccctg	3960
gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc	4020
ggcagctccg aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag	4080
gtgctgcgcy cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc	4140
gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac	4200
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg	4260
atagtatcga cacactctgg acgctggctg tgtgatggac tgttgccgcc acacttgctg	4320
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctgagtggtt ttgatcttgt	4380
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcaataacc acccccagca	4440
tccccctccc tcgtttcata tcgcttgcac cccaaccgca acttatctac gctgtcctgc	4500
tatccctcag cgctgctcct gctcctgctc actgccccct gcacagcctt ggtttgggct	4560

026172-005020PC - Sequence Listing_1

ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620
aagtagtgagg atgggaacac aaatggaaag ctt 4653

<210> 122
<211> 4647
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 122
ggtagccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgccctgccgc atgcttggtg tggtagggct gggcagtgct 120
gccatgctga ttgaggcttg gttcatcggg tggaaagctta tgtgtgtgct gggcttgcat 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccagggtgg tcaatcgcgg cagccagagg gatttcagat gatcgcgctg acagggttgag 300
cagcagtgct agcaaaggta gcagtttgcc agaattgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtc tccaacgtaa ccatgagctg atcaaacactg caatcatcg gggggcgtag 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
cacgccgttg agttggcagc aggtcagcc atgactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg ccgcgaagct ttgcagaacc 840
atgtcatgag acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcggg 960
gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcattgctg gagaagcag gctggggcct ttgggctggg gcagcccgc 1080
atggaatgca ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatggt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
cgcgccagct cagcaacctg tcccgtgggt ccccgtgcc gatgaaatcg tgtgcacgcc 1260
gatcagctga ttgcccggct cgcgaagtag gcgcctcct ttctgctgc cctctctccg 1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc 1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
cgccgtgttc cgcctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc 1500
cggcgccctg aacgagctgc aggacgaggc cgccgcctg tcctggctgg ccaccaccgg 1560
cgtgccctgc gccgccgtgc tggacgtggg gaccgaggcc ggccgcgact ggctgctgct 1620

026172-005020PC - Sequence Listing_1

gggcgaggtg	cccggccagg	acctgctgtc	ctccacactg	gccccgccg	agaaggtgtc	1680
catcatggcc	gacgccatgc	gccgcctgca	caccttgga	cccgccacct	gcccccttca	1740
ccaccaggcc	aagcaccgca	tcgagcgcg	ccgcacccgc	atggaggccg	gcctggtgga	1800
ccaggacgac	ctggacgagg	agcaccagg	cctggcccc	gccgagctgt	tcgccccct	1860
gaaggcccc	atgccccgac	gagaggacct	ggtggtgacc	cacggcgacg	cctgcctgcc	1920
caacatcatg	gtggagaacg	gccgcttctc	cggcttcac	gactgcggcc	gcctgggctg	1980
ggccgaccgc	taccaggaca	tcgccctggc	caccgcgac	atcgccgagg	agctgggctg	2040
cgagtgggcc	gaccgcttcc	tgggtgctgt	cggcatcgcc	gcccccgact	cccagcgcat	2100
cgccttctac	cgcctgctgg	acgagttctt	ctgacaattg	gcagcagcag	ctcgatagtg	2160
atcgacacac	tctggacgct	ggtcgtgtga	tggactgttg	ccgccacact	tgctgccttg	2220
acctgtgaat	atccctgccg	cttttatcaa	acagcctcag	tgtgtttgat	cttgtgtgta	2280
cgcgcttttg	cgagttgcta	gctgcttggt	ctatttgca	ataccacccc	cagcatcccc	2340
ttccctcggt	tcatatcgct	tgcattccaa	ccgcaactta	tctacgctgt	cctgctatcc	2400
ctcagcgctg	ctcctgctcc	tgctcactgc	ccctcgca	gccttggttt	gggctccgcc	2460
tgtattctcc	tggtactgca	acctgtaaac	cagcactgca	atgctgatgc	acgggaagta	2520
gtgggatggg	aacacaaatg	gaggatcccc	cgtctcgaac	agagcgcgca	gaggaacgct	2580
gaaggctctg	cctctgtcgc	acctcagcgc	ggcatacacc	acaataacca	cctgacgaat	2640
gcgcttggtt	cttcgtccat	tagcgaagcg	tccggttcac	acacgtgcca	cggtggcgag	2700
gtggcaggtg	acaatgatcg	gtggagctga	tggtcgaaac	gttcacagcc	tagggatata	2760
gaattccttt	cttgcgctat	gacacttcca	gcaaaaggta	gggcgggctg	cgagacggct	2820
tcccgcgct	gcatgcaaca	ccgatgatgc	ttcgaccccc	cgaagctcct	tcggggctgc	2880
atgggcgctc	cgatgccgct	ccagggcgag	cgctgtttaa	atagccaggc	ccccgattgc	2940
aaagacatta	tagcgagcta	ccaaagccat	attcaaacac	ctagatcact	accacttcta	3000
cacaggccac	tcgagcttgt	gatcgactc	cgctaagggg	gcgcctcttc	ctcttcgttt	3060
cagtcacaac	ccgcaaacac	tagtatggct	tccgcggcat	tcaccatgtc	ggcgtgcccc	3120
gcgatgactg	gcagggcccc	tggggcacgt	cgctccggac	ggccagtcgc	caccgcctg	3180
agggggcgcg	ccccgactg	gtccatgctg	ttcgccgtga	tcaccacat	cttcagcgcc	3240
gccgagaagc	agtggaccaa	cctggagtgg	aagcccaagc	ccaagctgcc	ccagctgctg	3300
gacgaccact	tcggcctgca	cggcctggtg	ttccgccgca	ccttcgccat	ccgctcctac	3360
gaggtgggcc	ccgaccgcag	cacctccatc	ctggccgtga	tgaaccacat	gcaggaggcc	3420
accctgaacc	acgccaagag	cgtgggcata	ctgggcgacg	gcttcggcac	cacctggag	3480
atgtccaagc	gcgacctgat	gtgggtggtg	cgccgcaccc	acgtggccgt	ggagcgctac	3540
cccacctggg	gcgacaccgt	ggaggtggag	tgctggatcg	gcgccagcgg	caacaacggc	3600
atgcgccgcg	acttcctggt	gcgcgactgc	aagaccggcg	agatcctgac	ccgctgcacc	3660

026172-005020PC - Sequence Listing_1

tccctgagcg tgctgatgaa caccgcgacc cgccgcctga gcaccatccc cgacgaggtg	3720
cgcgggcgaga tcggccccgc cttcatcgac aacgtggccg tgaaggacga cgagatcaag	3780
aagctgcaga agctgaacga ctccaccgcc gactacatcc agggcggcct gacccccgc	3840
tggaacgacc tggacgtgaa ccagcacgtg aacaacctga agtacgtggc ctgggtgttc	3900
gagaccgtgc ccgacagcat cttcaggtcc caccacatca gtccttcac cctggagtac	3960
cgccgcgagt gaccccgca ctccgtgctg cgcagcctga ccaccgtgag cggcggcagc	4020
tccgaggccg gcctgggtgtg cgaccacctg ctgcagctgg agggcggcag cgaggtgctg	4080
cgcgccccga ccgagtggcg ccccaagctg accgactcct tccgcggcat cagcgtgac	4140
cccgccgagc cccgcgtgat ggactacaag gaccacgacg gcgactacaa ggaccacgac	4200
atcgactaca aggacgacga cgacaagtga tgactcgagg cagcagcagc tcggatagta	4260
tcgacacact ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga	4320
cctgtgaata tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac	4380
gcgcttttgc gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatccct	4440
tccctcgttt catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc	4500
tcagcgtgc tcctgtctct gctcactgcc cctgcacag ccttggtttg ggctccgcct	4560
gtattctcct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag	4620
tgggatggga acacaaatgg aaagctt	4647

<210> 123

<211> 4721

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 123

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtagaggct gggcagtgct	120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgct acaggttgag	300
cagcagtgtc agcaaaggta gcagtttgcc agaattgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaaacctg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaa acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg	660

026172-005020PC - Sequence Listing_1

caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcggggtc gcgccccagc cagcggatgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcattgctg gagaagcag gctggggcct ttgggctggg gcagcccga	1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt	1140
ccatagtgc tttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcaagtag gcgcccctct ttctgctcgc cctctctccg	1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc	1380
ctgggtggag cgctgttctg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgttccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggg gaccgaggcc ggccgcgact ggctgctgct	1620
gggagaggtg cccggccagg acctgctgtc ctcccacctg gccccgcgg agaaggtgtc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctgggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggcccc cccgagctgt tcgcccgcct	1860
gaaggccccg atgccccgac gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttctac gactgcggcc gcctgggctg	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgcac atcgccgagg agctgggctg	2040
cgagtgggcc gaccgcttcc tgggtgctgta cggcatcgcc gcccccgact cccagcgcat	2100
cgccttctac cgctgtctgg acgagttctt ctgacaattg gcagcagcag ctcgatagat	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgtg ctctgtctcc tgcactgc ccctgcaca gccttggttt gggctccgcc	2460
tgtattctcc tggtagtgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatccc cgtctcgaa agagcgcgca gaggaacgct	2580
gaaggctctc cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700

026172-005020PC - Sequence Listing_1

gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
gaatttccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttta atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacgg cgcgccatgg ccaccacctc cctggcctcc gccttctgca	3120
gcatgaaggc cgtgatgctg gcccgcgacg gccgcggcat gaagccccgc tccagcgacc	3180
tgcagctgcg cgccggcaac gcccagacct ccctgaagat gatcaacggc accaagtctt	3240
cctacaccga gagcctgaag aagctgcccc actggtccat gctgttcgcc gtgatcacca	3300
ccatcttctc cgccgccgag aagcagtggg ccaacctgga gtggaagccc aagcccaacc	3360
ccccccagct gctggacgac cacttcggcc cccacggcct ggtgttcgcg cgcaccttcg	3420
ccatccgcag ctacgaggtg ggccccgacc gctccaccag catcgtggcc gtgatgaacc	3480
acctgcagga ggccgcccctg aaccacgcca agtccgtggg catcctgggc gacggcttcg	3540
gcaccaccct ggagatgtcc aagcgcgacc tgatctgggt ggtgaagcgc acccacgtgg	3600
ccgtggagcg ctaccccgcc tggggcgaca ccgtggaggt ggagtgtctg gtgggcgcct	3660
ccggcaacaa cgcccgccgc cacttctcc tggcgcgca ctgcaagacc ggcgagatcc	3720
tgacccgctg cacctccctg agcgtgatga tgaacacccg caccgcccgc ctgagcaaga	3780
tccccgagga ggtgcgcggc gagatcggcc ccgccttcat cgacaacgtg gccgtgaagg	3840
acgaggagat caagaagccc cagaagctga acgactccac cgccgactac atccagggcg	3900
gcctgacccc ccgctggaac gacctggaca tcaaccagca cgtgaacaac atcaagtacg	3960
tggactggat cctggagacc gtgcccgaca gcatcttcga gagccaccac atctcctcct	4020
tcaccatcga gtaccgcccgc gagtgcacca tggacagcgt gctgcagtcc ctgaccaccg	4080
tgagcggcgg ctctccgag gccggcctgg tgtgcgagca cctgctgcag ctggagggcg	4140
gcagcgaggt gctgcgcgcc aagaccgagt ggcgccccaa gctgaccgac tccttcgcgcg	4200
gcatcagcgt gatccccgcc gagtccagcg tgatggacta caaggaccac gacggcgact	4260
acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgatgactc gaggcagcag	4320
cagctcggat agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac	4380
acttgctgcc ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt	4440
gatcttgtgt gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac	4500
ccccagcatc cccttccttc gtttcatatc gcttgcatcc caaccgcaac ttatctacgc	4560
tgtcctgcta tccctcagcg ctgctcctgc tcctgctcac tgccccctgc acagccttgg	4620
tttgggctcc gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga	4680
tgcacgggaa gtagtgggat gggaacacaa atggaaagct t	4721

026172-005020PC - Sequence Listing_1

<210> 124
 <211> 4650
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 124
 ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
 gtgcttgggc gcctgccgcc tgcctgccgc atgcttggtg tggtagaggct gggcagtgtc 120
 gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcac 180
 gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
 tccaggtggt tcaatcgccg cagccagagg gatttcagat gatcgcgctg acaggttgag 300
 cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
 cagcaagaga aggggtcaag tgcaaacacg ggcattgccac agcacgggca ccggggagtg 420
 gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
 gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggctga 540
 tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
 cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
 caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
 tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
 gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
 atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
 acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
 gccacccttc ttgcggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt 1020
 cacattcatt tgcattgctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
 atggaatgca ggaccgccag gctagcagca aaggcgccct ccctactccg catcgatgtt 1140
 ccatagtga ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
 cgcgccagct cagcaacctg tcccgtgggt ccccgctgcc gatgaaatcg tgtgcacgcc 1260
 gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320
 tccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct ccccgccgc 1380
 ctgggtggag cgctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
 cgccgtgttc cgctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc 1500
 cggcgccctg aacgagctgc aggacgaggc cgccgcctg tcctggctgg ccaccaccgg 1560
 cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
 gggcgagggtg cccggccagg acctgctgtc ctcccacctg gccccgcccg agaaggtgtc 1680
 catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga 1740

026172-005020PC - Sequence Listing_1

ccaccaggcc aagcaccgca tcgagcgcg cccgacccgc atggaggccg gcctggtgga	1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct	1860
gaaggcccg atgcccgcg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcac gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgcac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tgggtgtgta cggcatcgcc gccccgact cccagcgcat	2100
cgccttctac cgctgtctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat ctgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgcga ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgctg ctctgtctcc tgctcactgc cctcgcaca gccttggttt gggctccgcc	2460
tgtattctcc tggactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct	2580
gaagggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcagggt acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
gaattccttt cttgcgctat gacacttcca gaaaaggta gggcgggctg cgagacggct	2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc	2880
atgggcgctc cgatgccgct ccaggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatgacg ttcggggctg ccctcccggc catgggccgc	3120
ggtgtctccc tttcccggcc cagggtcgcg gtgcgcgcc agtcggcgag tcaggttttg	3180
gagagcgggc gcgccccga ctggtccatg ctgttcgccg tgatcaccac catcttctcc	3240
gccgccgaga agcagtggac caacctggag tggaaagcca agcccaaccc cccccagctg	3300
ctggacgacc acttcggccc ccacggcctg gtgttcgcc gcaccttcgc catccgcagc	3360
tacgaggtgg gccccgaccg ctccaccagc atcgtggccg tgatgaacca cctgcaggag	3420
gccgccctga accacgcaa gtccgtgggc atcctgggcg acggcttcgg caccaccctg	3480
gagatgtcca agcgcgacct gatctgggtg gtgaagcgca cccacgtggc cgtggagcgc	3540
taccccgctt ggggcgacac cgtggaggtg gagtgctggg tgggcgcctc cggcaacaac	3600
ggccgcggcc acgacttcct ggtgcgcgac tgcaagaccg gcgagatcct gacccgctgc	3660
acctccctga gcgtgatgat gaacacccgc acccgccgcc tgagcaagat ccccgaggag	3720
gtgcgcggcg agatcggccc cgccttcac gacaacgtgg ccgtgaagga cgaggagatc	3780

026172-005020PC - Sequence Listing_1

aagaagcccc agaagctgaa cgactccacc gccgactaca tccagggcgg cctgaccccc	3840
cgctggaacg acctggacat caaccagcac gtgaacaaca tcaagtacgt ggactggatc	3900
ctggagaccg tgcccgcacag catcttcgag agccaccaca tctcctcctt caccatcgag	3960
taccgccgcg agtgcaccat ggacagcgtg ctgcagtccc tgaccaccgt gagcggcggc	4020
tcctccgagg ccggcctggt gtgcgagcac ctgctgcagc tggagggcgg cagcgagggtg	4080
ctgcgcgcca agaccgagtg gcgccccaaag ctgaccgact ccttccgcgg catcagcgtg	4140
atccccgcgg agtccagcgt gatggactac aaggaccacg acggcgacta caaggaccac	4200
gacatcgact acaaggacga cgacgacaag tgatgactcg aggcagcagc agctcggata	4260
gtatcgacac actctggacg ctggtcgtgt gatggactgt tgccgccaca cttgctgcct	4320
tgacctgtga atatccctgc cgcttttatc aaacagcctc agtgtgtttg atcttgtgtg	4380
tacgcgcttt tgcgagttgc tagctgcttg tgctatttgc gaataccacc cccagcatcc	4440
ccttccctcg tttcatatcg cttgcatccc aaccgcaact tatctacgt gtctgtctat	4500
ccctcagcgc tgctcctgct cctgctcact gcccctcgca cagccttgggt ttgggctccg	4560
cctgtattct cctggtactg caacctgtaa accagcactg caatgctgat gcacgggaag	4620
tagtgggatg ggaacacaaa tggaaagctt	4650

<210> 125

<211> 4653

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 125

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac	60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtagaggct gggcagtgct	120
gccatgctga ttgaggcttg gttcatcggg tggaaagctta tgtgtgtgct gggcttgcac	180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc	240
tccaggctggt tcaatcgcgg cagccagagg gatttcagat gatcgcgct acaggttgag	300
cagcagtgct agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc	360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg	420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg	480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga	540
tgcaagcatg cctggcgaaag acacatggtg tgcggatgct gccggctgct gcctgctgcg	600
cacgccgttg agttggcagc aggtcagcc atgcactgga tggcagctgg gctgccactg	660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt	720
tgcattggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat	780

026172-005020PC - Sequence Listing_1

gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc	840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa	900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg	960
gccacccttc ttgcgggggtc gcgccccagc cagcgggtgat gcgctgatcc caaacgagtt	1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggg gcagcccgc	1080
atggaatgcg ggaccgccag gctagcagca aaggcgccct ccctactccg catcgatgtt	1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa	1200
cgcgccagct cagcaacctg tcccgtgggt ccccgtgcc gatgaaatcg tgtgcacgcc	1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg	1320
tccgcctct agaatatcaa tgatcgagca ggacggcctc cagcgggct cccccgcgc	1380
ctgggtggag cgctgttctg gctacgactg ggcccagcag accatcggct gctccgacgc	1440
cgccgtgttc cgctgtccg cccagggccg ccccgctgctg ttcgtgaaga ccgacctgtc	1500
cggcgccctg aacgagctgc aggacgaggc cgccgcctg tcctggctgg ccaccaccgg	1560
cgtgccctgc gccgccgtgc tggacgtggg gaccgaggcc ggccgcgact ggctgctgct	1620
gggcgagggtg cccggccagg acctgctgtc ctcccacctg gccccgccg agaagggtgc	1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccttcga	1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcaccgc atggaggccg gcctgggtga	1800
ccaggacgac ctggacgagg agcaccaggg cctggcccc gccgagctgt tcgcccgcct	1860
gaaggccccg atgcccagc gcgaggacct ggtggtgacc cagggcgacg cctgcctgcc	1920
caacatcatg gtggagaacg gccgcttctc cggcttcctc gactgcggcc gcctgggcgt	1980
ggccgaccgc taccaggaca tcgccctggc caccgcgac atcgccgagg agctgggcgg	2040
cgagtgggcc gaccgcttcc tgggtgctgta cggcatcgcc gccccgact cccagcgcat	2100
cgccttctac cgctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt	2160
atcgacacac tctggacgt ggtcgtgtga tggactgttg ccgccacact tgctgccttg	2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta	2280
cgcgcttttg cgagttgcta gctgcttggt ctatttgca ataccacccc cagcatcccc	2340
ttccctcggt tcatatcgct tgcattccaa ccgcaactta tctacgctgt cctgctatcc	2400
ctcagcgctg ctctgctcc tgctcactgc ccctcgaca gccttggttt gggctccgcc	2460
tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta	2520
gtgggatggg aacacaaatg gaggatcccc cgtctcgaac agagcgcgca gaggaacgt	2580
gaagggtctc cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat	2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag	2700
gtggcagggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc	2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct	2820

026172-005020PC - Sequence Listing_1

tccccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcgggggctgc	2880
atggggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc	2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta	3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt	3060
cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag	3120
cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg	3180
gcgcttcgtg ggcgcgcccc cgactgggtcc atgctgttcg ccgtgatcac caccatcttc	3240
tccgccgccg agaagcagtg gaccaacctg gagtgggaagc ccaagcccaa cccccccag	3300
ctgctggagc accacttcgg ccccccacggc ctgggtgttc gccgcacctt cgccatccgc	3360
agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag	3420
gaggccgccc tgaaccacgc caagtccgtg ggcaccttg gcgacggctt cggcaccacc	3480
ctggagatgt ccaagcgca cctgatctgg gtggtgaagc gcacccacgt ggccgtggag	3540
cgctaccccc cctggggcgca caccgtggag gtggagtgt ggggtgggcgc ctccggcaac	3600
aacggccgcc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgaccgc	3660
tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag	3720
gaggtgcgcg gcgagatcgg ccccgcttc atcgacaacg tggccgtgaa ggacgaggag	3780
atcaagaagc ccagaagct gaacgactcc accgcccact acatccaggg cggcctgacc	3840
ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg	3900
atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcttc cttcaccatc	3960
gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc	4020
ggctcctccg aggcgggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag	4080
gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc	4140
gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac	4200
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg	4260
atagtatcga cacactctgg acgctggctg tgtgatggac tgttgccgcc acacttgctg	4320
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt	4380
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcaataacc acccccagca	4440
tccccttccc tcgtttcata tcgcttgcac cccaaccgca acttatctac gctgtcctgc	4500
tatccctcag cgctgctcct gtcctgctc actgccccct gcacagcctt ggtttgggct	4560
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg	4620
aagtagtggg atgggaacac aaatggaaag ctt	4653

<210> 126

<211> 3669

<212> DNA

<213> Artificial Sequence

026172-005020PC - Sequence Listing_1

<220>

<223> Description of Artificial Sequence: Synthetic polynucleotide

<400> 126

```

cccgtgatca cacaggtgcc ttgcgagcgt gatcacacta ttttgggggt cctacagtac      60
tgaaatggtg agaagtcgta ctgaaatcaa ggatgaacaa tgaaaatggt gctgtggtgg      120
cttctcaaag gtcaagaatc agtcgctcgc gtcaggaaat cgcggcgtca accagcgtgg      180
gcgcggtcag tggccccgca ctggtcacca tagcctctcc tgccacagta gcgatcccct      240
gggcgttcac tctcagcagc ggctgtactg cctcccagat tttcttcttc tggacctgcg      300
ggcgtgagag gatgagcagg gtggggccaa gggctcaatc ctgaacggcc ctcatcgggt      360
ttccaatccc acaacacata cccacagcag gtcagaccac gcattcgcac catgcgcacc      420
aaataacgtg tccttacctg attgggtgtg gcaggctccg tggacaggag tgcctcgtcc      480
cccgcccaga cccgctcccc cgtcacggcg gcgtccggga cccgcagcgg ctccaccgcg      540
gtgtgatccg cgttggcggc gcagagcagc atcccagccg atttgacccc gcgcatgctc      600
cgaggcttga ggttggccag caccaccacc cgccggccga caaggctctc cagggtcacg      660
tgccggacca ggccactcac gatggtgcga gggccccctt cctcgccgag gtcgatctgc      720
tcgacgtaca gactgcgaca tgcgtggcga gtggtcatca gaaggaaagca ggtgtgcaga      780
aggggcacgt ggttgggtatt gagagtagcc aaagctttgt gccaatcaga aagtcaacgc      840
agctgcctgc ctggctcgcg tacaattcct ttcttgcgct atgacacttc cagcaaaaagg      900
tagggcgggc tgcgagacgg cttcccggcg ctgcatgcaa caccgatgat gcttcgaccc      960
cccgaagctc cttcgggggt gcatgggcgc tccgatgccg ctccaggcgg agcgtgttt      1020
aatagccag gccccgatt gcaaagacat tatagcgagc taccaaagcc atattcaaac      1080
acctagatca ctaccacttc tacacaggcc actcgagctt gtgatcgcac tccgctaagg      1140
gggcgcctct tcctcttcgt ttcagtcaca acccgcaaac ggcgcgccat gctgctgcag      1200
gccttcctgt tcctgctggc cggttcgcc gccaaagatca gcgcctccat gacgaacgag      1260
acgtccgacc gccccctggt gcacttcacc cccaacaagg gctggatgaa cgacccaac      1320
ggcctgtggt acgacgagaa ggacgccaag tggcacctgt acttcagta caaccgaac      1380
gacaccgtct gggggacgcc cttgttctgg ggccacgcca cgtccgacga cctgaccaac      1440
tgaggaggacc agcccatcgc catcgccccg aagcgcaacg actccggcgc cttctccggc      1500
tccatggtgg tggactacaa caacacctcc ggcttcttca acgacaccat cgacccgcgc      1560
cagcgtcgcg tggccatctg gacctacaac accccggagt ccgaggagca gtacatctcc      1620
tacagcctgg acggcggcta caccttcacc gagtaccaga agaaccccgt gctggccgcc      1680
aactccaccc agttccgcga cccgaaggct ttctggtacg agccctcca gaagtggatc      1740
atgaccgcgg ccaagtccca ggactacaag atcgagatct actcctccga cgacctgaag      1800
tcctggaagc tggagtccgc gttcgccaac gagggcttcc tcggctacca gtacgagtgc      1860
cccggcctga tcgagggtccc caccgagcag gacccagca agtcctactg ggtgatgttc      1920

```

026172-005020PC - Sequence Listing_1

atctccatca	acccccggcgc	cccggccggc	ggctccttca	accagtactt	cgtcggcagc	1980
ttcaacggca	cccacttcga	ggccttcgac	aaccagtccc	gcgtggtgga	cttcggcaag	2040
gactactacg	ccctgcagac	cttcttcaac	accgaccgga	cctacgggag	cgccctgggc	2100
atcgcgtagg	cctccaactg	ggagtactcc	gccttcgtgc	ccaccaaccc	ctggcgctcc	2160
tccatgtccc	tcgtgcgcaa	gttctccctc	aacaccgagt	accaggccaa	cccggagacg	2220
gagctgatca	acctgaaggc	cgagccgatc	ctgaacatca	gcaacgccgg	cccctggagc	2280
cggttcgcca	ccaacaccac	gttgacgaag	gccaacagct	acaacgtcga	cctgtccaac	2340
agcaccggca	ccctggagtt	cgagctggtg	tacgccgtca	acaccacca	gacgatctcc	2400
aagtccgtgt	tcgcggacct	ctccctctgg	ttcaagggcc	tggaggacct	cgaggagtac	2460
ctccgcatgg	gcttcgaggt	gtccgcgtcc	tccttcttcc	tggaccgcgg	gaacagcaag	2520
gtgaagttcg	tgaaggagaa	cccctacttc	accaaccgca	tgagcgtgaa	caaccagccc	2580
ttcaagagcg	agaacgacct	gtcctactac	aagggtgtacg	gcttgctgga	ccagaacatc	2640
ctggagctgt	acttcaacga	cggcgacgtc	gtgtccacca	acacctactt	catgaccacc	2700
gggaacgccc	tgggctccgt	gaacatgacg	acgggggtgg	acaacctgtt	ctacatcgac	2760
aagttccagg	tgcgcgaggt	caagtgatta	attaactcga	ggcagcagca	gctcggatag	2820
tatcgacaca	ctctggacgc	tggtcgtgtg	atggactgtt	gccgccacac	ttgctgcctt	2880
gacctgtgaa	tatccctgcc	gcttttatca	aacagcctca	gtgtgtttga	tcttgtgtgt	2940
acgcgctttt	gcgagttgct	agctgcttgt	gctatttgcg	aataccaccc	ccagcatccc	3000
cttccctcgt	ttcatatcgc	ttgcatccca	accgcaactt	atctacgctg	tcctgctatc	3060
cctcagcgct	gctcctgctc	ctgctcactg	cccctcgcac	agccttggtt	tgggctccgc	3120
ctgtattctc	ctgggtactgc	aacctgtaaa	ccagcactgc	aatgctgatg	cacgggaagt	3180
agtgggatgg	gaacacaaat	ggaaagcttg	agctcggtag	ccgtacccat	cagcatccgg	3240
gtgaatcttg	gcctccaaga	tatggccaat	cctcacatcc	agcttgga	aatcgactag	3300
actgtctgca	agtgggaatg	tggagcacia	ggttgcttgt	agcgatcgac	agactggtgg	3360
ggtacattga	caggtgggca	gcgccgcac	catcgtcct	gacgcgagcg	ccgccggttg	3420
ctcgcccggtg	cctgccgtca	aagagcggca	gagaaatcgg	gaaccgaaaa	cgtcacattg	3480
cctgatgttg	ttacatgctg	gactagactt	tcttggcggtg	ggtctgctcc	tcgccagggtg	3540
cgcgacgcct	cggggctggg	tgcgaggag	ccgtgcggcc	acgcatttga	caagacccaa	3600
agctcgcac	tcagacggtc	aaccgttcgt	attatacatt	caacatatgg	tacatacgca	3660
aaaagcatg						3669

<210> 127

<211> 39

<212> PRT

<213> *Prototheca moriformis*

<400> 127

026172-005020PC - Sequence Listing_1

Met Thr Phe Gly Val Ala Leu Pro Ala Met Gly Arg Gly Val Ser Leu
1 5 10 15

Pro Arg Pro Arg Val Ala Val Arg Ala Gln Ser Ala Ser Gln Val Leu
20 25 30

Glu Ser Gly Arg Ala Gln Leu
35

<210> 128
<211> 40
<212> PRT
<213> Prototheca moriformis

<400> 128
Met Ala Ile Lys Thr Asn Arg Gln Pro Val Glu Lys Pro Pro Phe Thr
1 5 10 15

Ile Gly Thr Leu Arg Lys Ala Ile Pro Ala His Cys Phe Glu Arg Ser
20 25 30

Ala Leu Arg Gly Arg Ala Gln Leu
35 40

<210> 129
<211> 36
<212> PRT
<213> Prototheca moriformis

<400> 129
Met Ala Ser Ala Ala Phe Thr Met Ser Ala Cys Pro Ala Met Thr Gly
1 5 10 15

Arg Ala Pro Gly Ala Arg Arg Ser Gly Arg Pro Val Ala Thr Arg Leu
20 25 30

Arg Gly Arg Ala
35

<210> 130
<211> 40
<212> PRT
<213> Chlorella protothecoides

<400> 130
Met Ala Thr Ala Ser Thr Phe Ser Ala Phe Asn Ala Arg Cys Gly Asp
1 5 10 15

Leu Arg Arg Ser Ala Gly Ser Gly Pro Arg Arg Pro Ala Arg Pro Leu
20 25 30

Pro Val Arg Gly Arg Ala Gln Leu
35 40

<210> 131
<211> 87

026172-005020PC - Sequence Listing_1

<212> PRT

<213> *Cuphea hookeriana*

<400> 131

Met Val Ala Ala Ala Ala Ser Ser Ala Phe Phe Pro Val Pro Ala Pro
1 5 10 15

Gly Ala Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu
20 25 30

Ser Pro Ser Phe Lys Pro Lys Ser Ile Pro Asn Gly Gly Phe Gln Val
35 40 45

Lys Ala Asn Asp Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Ser
50 55 60

Leu Lys Ser Gly Ser Leu Asn Thr Gln Glu Asp Thr Ser Ser Ser Pro
65 70 75 80

Pro Pro Arg Thr Phe Leu His
85

<210> 132

<211> 60

<212> PRT

<213> *Umbellularia californica*

<400> 132

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

Gln Leu Arg Ala Gly Asn Ala Pro Thr Ser Leu Lys Met Ile Asn Gly
35 40 45

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Arg Leu
50 55 60

<210> 133

<211> 60

<212> PRT

<213> *Cinnamomum camphora*

<400> 133

Met Ala Thr Thr Ser Leu Ala Ser Ala Phe Cys Ser Met Lys Ala Val
1 5 10 15

Met Leu Ala Arg Asp Gly Arg Gly Met Lys Pro Arg Ser Ser Asp Leu
20 25 30

Gln Leu Arg Ala Gly Asn Ala Gln Thr Ser Leu Lys Met Ile Asn Gly
35 40 45

026172-005020PC - Sequence Listing_1

Thr Lys Phe Ser Tyr Thr Glu Ser Leu Lys Lys Leu
 50 55 60

<210> 134
 <211> 1104
 <212> DNA
 <213> *Prototheca moriformis*

<400> 134
 atggcaccga ccagcctgct tgccagtact ggcgtctctt ccgcttctct gtggtcctct 60
 gcgcgtccca gcgcgtgcgc ttttccggtg gatcatgcgg tccgtggcgc accgcagcgg 120
 ccgctgcccc tgcagcgccg ctgcttccga acagtggcgg tcagggccgc acccgcggtg 180
 gccgtccgtc cggaacccgc ccaagagttt tgggagcagc ttgagccctg caagatggcg 240
 gaggacaagc gcatcttcct ggaggagcac cgcattcggg gcaacgaggt gggcccctcg 300
 cagcggctga cgatcacggc ggtggccaac atcctgcagg aggcggcggg caaccacgcg 360
 gtggccatgt ggggcccggag ctcgaggggg ttcgcgacgg acccgagct gcaggaggcg 420
 ggtctcatct ttgtgatgac gcgcgtgcag atccaaatgt accgctaccc gcgctggggc 480
 gacctgatgc aggtggagac ctggttccag acggcgggca agctaggcgc gcagcgcgag 540
 tgggtgctgc gcgacaagct gaccggcgag gcgctggggc cggccacctc cagctgggtc 600
 atgatcaaca tccgcacgcg ccggccgtgc cgcattgccc agctcgtccg cgtcaagtcg 660
 gccttcttcg cgcgcgagcc gccgcgcctg gcgctgccgc ccacggtcac gcgcgccaaag 720
 ctgccccaca tcgcgacgcc ggcgcgcctg cgcgggcacc gccaggtcgc gcgccgcacc 780
 gacatggaca tgaacgggca cgtgaacaac gtggcctacc tggcctggtg cctggaggcc 840
 gtgcccagac acgtcttcag cgactaccac ctctaccaga tggagatcga cttcaaggcc 900
 gagtgccacg cgggcgacgt catctcctcc caggccgagc agatcccgcc ccaggaggcg 960
 ctcacgcaca acggcgccgg ccgcaacccc tcctgcttcg tccatagcat tctgcgcgcc 1020
 gagaccgagc tcgtccgcgc gcgaaccaca tggtcggccc ccatcgacgc gcccgccgcc 1080
 aagccgcca aggcgagcca ctga 1104

<210> 135
 <211> 1104
 <212> DNA
 <213> *Prototheca moriformis*

<400> 135
 atggcaccga ccagcctgct tgcccgtact ggcgtctctt ccgcttctct gtgctcctct 60
 acgcgtcccg gcgcgtgcgc ttttccggtg gatcatgcgg tccgtggcgc accgcagcgg 120
 ccgctgcccc tgcagcgccg ctgcttccga acagtggctg tcagggccgc acccgcagta 180
 gccgtccgtc cggaacccgc ccaagagttt tgggagcagc ttgagccctg caagatggcg 240
 gaggacaagc gcatcttcct ggaggagcac cgcattcgtg gcaacgaggt gggcccctcg 300
 cagcggctga cgatcacggc ggtggccaac atcctgcagg aggcggcggg caaccacgcg 360
 gtggccatgt ggggtcggag ctcgaggggg ttcgcgacgg acccgagct gcaggaggcg 420

026172-005020PC - Sequence Listing_1

```

ggcctcatct ttgtgatgac ggcgatgcag atccaaatgt accgctaccc gcgctggggc 480
gacctgatgc aggtggagac ctggttccag acggcgggca agctaggcgc gcagcgcgag 540
tgggtgctgc gcgacaagct gaccggcgag gcgctggggc cggccacctc cagctgggtc 600
atgatcaaca tccgcacgcg ccggccgtgc cgcattgccc agctcgtccg cgtcaagtcg 660
gcctttcttc gcgcgcgagcc gccgcgcctg gcgctgccgc ccgcggtcac gcgtgccaaag 720
ctgcccacaaca tcgcgacgcc ggcgcgcctg cgcgggcacc gccaggctgc gcgccgcacc 780
gacatggaca tgaacggcca cgtgaacaac gttgcctacc tggcctgggtg cctggaggcc 840
gtgcccgcgc acgtcttcag cgactaccac ctctaccaga tggagatcga cttcaaggcc 900
gagtgccacg cgggcgacgt catctcctcc caggccgagc agatcccgcc ccaggaggcg 960
ctcacgcaca acgggcgccg ccgcaacccc tctgtcttcg tccatagcat tctgcgcgcc 1020
gagaccgagc tcgtccgcgc gcgaaccaca tggtcggccc ccatcgacgc gcccgccgcc 1080
aagccgccc aaggcagcca ctga 1104

```

<210> 136
 <211> 367
 <212> PRT
 <213> *Prototheca moriformis*

```

<400> 136
Met Ala Pro Thr Ser Leu Leu Ala Ser Thr Gly Val Ser Ser Ala Ser
1          5          10          15
Leu Trp Ser Ser Ala Arg Ser Ser Ala Cys Ala Phe Pro Val Asp His
20          25          30
Ala Val Arg Gly Ala Pro Gln Arg Pro Leu Pro Met Gln Arg Arg Cys
35          40          45
Phe Arg Thr Val Ala Val Arg Ala Ala Pro Ala Val Ala Val Arg Pro
50          55          60
Glu Pro Ala Gln Glu Phe Trp Glu Gln Leu Glu Pro Cys Lys Met Ala
65          70          75          80
Glu Asp Lys Arg Ile Phe Leu Glu Glu His Arg Ile Arg Gly Asn Glu
85          90          95
Val Gly Pro Ser Gln Arg Leu Thr Ile Thr Ala Val Ala Asn Ile Leu
100         105         110
Gln Glu Ala Ala Gly Asn His Ala Val Ala Met Trp Gly Arg Ser Ser
115         120         125
Glu Gly Phe Ala Thr Asp Pro Glu Leu Gln Glu Ala Gly Leu Ile Phe
130         135         140

```

026172-005020PC - Sequence Listing_1

Val Met Thr Arg Met Gln Ile Gln Met Tyr Arg Tyr Pro Arg Trp Gly
145 150 155 160

Asp Leu Met Gln Val Glu Thr Trp Phe Gln Thr Ala Gly Lys Leu Gly
165 170 175

Ala Gln Arg Glu Trp Val Leu Arg Asp Lys Leu Thr Gly Glu Ala Leu
180 185 190

Gly Ala Ala Thr Ser Ser Trp Val Met Ile Asn Ile Arg Thr Arg Arg
195 200 205

Pro Cys Arg Met Pro Glu Leu Val Arg Val Lys Ser Ala Phe Phe Ala
210 215 220

Arg Glu Pro Pro Arg Leu Ala Leu Pro Pro Thr Val Thr Arg Ala Lys
225 230 235 240

Leu Pro Asn Ile Ala Thr Pro Ala Pro Leu Arg Gly His Arg Gln Val
245 250 255

Ala Arg Arg Thr Asp Met Asp Met Asn Gly His Val Asn Asn Val Ala
260 265 270

Tyr Leu Ala Trp Cys Leu Glu Ala Val Pro Glu His Val Phe Ser Asp
275 280 285

Tyr His Leu Tyr Gln Met Glu Ile Asp Phe Lys Ala Glu Cys His Ala
290 295 300

Gly Asp Val Ile Ser Ser Gln Ala Glu Gln Ile Pro Pro Gln Glu Ala
305 310 315 320

Leu Thr His Asn Gly Ala Gly Arg Asn Pro Ser Cys Phe Val His Ser
325 330 335

Ile Leu Arg Ala Glu Thr Glu Leu Val Arg Ala Arg Thr Thr Trp Ser
340 345 350

Ala Pro Ile Asp Ala Pro Ala Ala Lys Pro Pro Lys Ala Ser His
355 360 365

<210> 137

<211> 367

<212> PRT

<213> Prototheca moriformis

<400> 137

Met Ala Pro Thr Ser Leu Leu Ala Arg Thr Gly Val Ser Ser Ala Ser
1 5 10 15

Leu Cys Ser Ser Thr Arg Ser Gly Ala Cys Ala Phe Pro Val Asp His
20 25 30

026172-005020PC - Sequence Listing_1

Ala Val Arg Gly Ala Pro Gln Arg Pro Leu Pro Met Gln Arg Arg Cys
35 40 45

Phe Arg Thr Val Ala Val Arg Ala Ala Pro Ala Val Ala Val Arg Pro
50 55 60

Glu Pro Ala Gln Glu Phe Trp Glu Gln Leu Glu Pro Cys Lys Met Ala
65 70 75 80

Glu Asp Lys Arg Ile Phe Leu Glu Glu His Arg Ile Arg Gly Asn Glu
85 90 95

Val Gly Pro Ser Gln Arg Leu Thr Ile Thr Ala Val Ala Asn Ile Leu
100 105 110

Gln Glu Ala Ala Gly Asn His Ala Val Ala Met Trp Gly Arg Ser Ser
115 120 125

Glu Gly Phe Ala Thr Asp Pro Glu Leu Gln Glu Ala Gly Leu Ile Phe
130 135 140

Val Met Thr Arg Met Gln Ile Gln Met Tyr Arg Tyr Pro Arg Trp Gly
145 150 155 160

Asp Leu Met Gln Val Glu Thr Trp Phe Gln Thr Ala Gly Lys Leu Gly
165 170 175

Ala Gln Arg Glu Trp Val Leu Arg Asp Lys Leu Thr Gly Glu Ala Leu
180 185 190

Gly Ala Ala Thr Ser Ser Trp Val Met Ile Asn Ile Arg Thr Arg Arg
195 200 205

Pro Cys Arg Met Pro Glu Leu Val Arg Val Lys Ser Ala Phe Phe Ala
210 215 220

Arg Glu Pro Pro Arg Leu Ala Leu Pro Pro Ala Val Thr Arg Ala Lys
225 230 235 240

Leu Pro Asn Ile Ala Thr Pro Ala Pro Leu Arg Gly His Arg Gln Val
245 250 255

Ala Arg Arg Thr Asp Met Asp Met Asn Gly His Val Asn Asn Val Ala
260 265 270

Tyr Leu Ala Trp Cys Leu Glu Ala Val Pro Glu His Val Phe Ser Asp
275 280 285

Tyr His Leu Tyr Gln Met Glu Ile Asp Phe Lys Ala Glu Cys His Ala
290 295 300

026172-005020PC - Sequence Listing_1

Gly Asp Val Ile Ser Ser Gln Ala Glu Gln Ile Pro Pro Gln Glu Ala
305 310 315 320

Leu Thr His Asn Gly Ala Gly Arg Asn Pro Ser Cys Phe Val His Ser
325 330 335

Ile Leu Arg Ala Glu Thr Glu Leu Val Arg Ala Arg Thr Thr Trp Ser
340 345 350

Ala Pro Ile Asp Ala Pro Ala Ala Lys Pro Pro Lys Ala Ser His
355 360 365

<210> 138

<211> 328

<212> PRT

<213> Cuphea hookeriana

<400> 138

Gln Leu Pro Asp Trp Ser Arg Leu Leu Thr Ala Ile Thr Thr Val Phe
1 5 10 15

Val Lys Ser Lys Arg Pro Asp Met His Asp Arg Lys Ser Lys Arg Pro
20 25 30

Asp Met Leu Val Asp Ser Phe Gly Leu Glu Ser Thr Val Gln Asp Gly
35 40 45

Leu Val Phe Arg Gln Ser Phe Ser Ile Arg Ser Tyr Glu Ile Gly Thr
50 55 60

Asp Arg Thr Ala Ser Ile Glu Thr Leu Met Asn His Leu Gln Glu Thr
65 70 75 80

Ser Leu Asn His Cys Lys Ser Thr Gly Ile Leu Leu Asp Gly Phe Gly
85 90 95

Arg Thr Leu Glu Met Cys Lys Arg Asp Leu Ile Trp Val Val Ile Lys
100 105 110

Met Gln Ile Lys Val Asn Arg Tyr Pro Ala Trp Gly Asp Thr Val Glu
115 120 125

Ile Asn Thr Arg Phe Ser Arg Leu Gly Lys Ile Gly Met Gly Arg Asp
130 135 140

Trp Leu Ile Ser Asp Cys Asn Thr Gly Glu Ile Leu Val Arg Ala Thr
145 150 155 160

Ser Ala Tyr Ala Met Met Asn Gln Lys Thr Arg Arg Leu Ser Lys Leu
165 170 175

026172-005020PC - Sequence Listing_1

Pro Tyr Glu Val His Gln Glu Ile Val Pro Leu Phe Val Asp Ser Pro
180 185 190

Val Ile Glu Asp Ser Asp Leu Lys Val His Lys Phe Lys Val Lys Thr
195 200 205

Gly Asp Ser Ile Gln Lys Gly Leu Thr Pro Gly Trp Asn Asp Leu Asp
210 215 220

Val Asn Gln His Val Ser Asn Val Lys Tyr Ile Gly Trp Ile Leu Glu
225 230 235 240

Ser Met Pro Thr Glu Val Leu Glu Thr Gln Glu Leu Cys Ser Leu Ala
245 250 255

Leu Glu Tyr Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Glu Ser Val
260 265 270

Thr Ala Met Asp Pro Ser Lys Val Gly Val Arg Ser Gln Tyr Gln His
275 280 285

Leu Leu Arg Leu Glu Asp Gly Thr Ala Ile Val Asn Gly Ala Thr Glu
290 295 300

Trp Arg Pro Lys Asn Ala Gly Ala Asn Gly Ala Ile Ser Thr Gly Lys
305 310 315 320

Thr Ser Asn Gly Asn Ser Val Ser
325

<210> 139
<211> 322
<212> PRT
<213> Umbellularia californica

<400> 139
Pro Asp Trp Ser Met Leu Phe Ala Val Ile Thr Thr Ile Phe Ser Ala
1 5 10 15

Ala Glu Lys Gln Trp Thr Asn Leu Glu Trp Lys Pro Lys Pro Lys Leu
20 25 30

Pro Gln Leu Leu Asp Asp His Phe Gly Leu His Gly Leu Val Phe Arg
35 40 45

Arg Thr Phe Ala Ile Arg Ser Tyr Glu Val Gly Pro Asp Arg Ser Thr
50 55 60

Ser Ile Leu Ala Val Met Asn His Met Gln Glu Ala Thr Leu Asn His
65 70 75 80

Ala Lys Ser Val Gly Ile Leu Gly Asp Gly Phe Gly Thr Thr Leu Glu
85 90 95

026172-005020PC - Sequence Listing_1

Met Ser Lys Arg Asp Leu Met Trp Val Val Arg Arg Thr His Val Ala
100 105 110

Val Glu Arg Tyr Pro Thr Trp Gly Asp Thr Val Glu Val Glu Cys Trp
115 120 125

Ile Gly Ala Ser Gly Asn Asn Gly Met Arg Arg Asp Phe Leu Val Arg
130 135 140

Asp Cys Lys Thr Gly Glu Ile Leu Thr Arg Cys Thr Ser Leu Ser Val
145 150 155 160

Leu Met Asn Thr Arg Thr Arg Arg Leu Ser Thr Ile Pro Asp Glu Val
165 170 175

Arg Gly Glu Ile Gly Pro Ala Phe Ile Asp Asn Val Ala Val Lys Asp
180 185 190

Asp Glu Ile Lys Lys Leu Gln Lys Leu Asn Asp Ser Thr Ala Asp Tyr
195 200 205

Ile Gln Gly Gly Leu Thr Pro Arg Trp Asn Asp Leu Asp Val Asn Gln
210 215 220

His Val Asn Asn Leu Lys Tyr Val Ala Trp Val Phe Glu Thr Val Pro
225 230 235 240

Asp Ser Ile Phe Glu Ser His His Ile Ser Ser Phe Thr Leu Glu Tyr
245 250 255

Arg Arg Glu Cys Thr Arg Asp Ser Val Leu Arg Ser Leu Thr Thr Val
260 265 270

Ser Gly Gly Ser Ser Glu Ala Gly Leu Val Cys Asp His Leu Leu Gln
275 280 285

Leu Glu Gly Gly Ser Glu Val Leu Arg Ala Arg Thr Glu Trp Arg Pro
290 295 300

Lys Leu Thr Asp Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Pro
305 310 315 320

Arg Val

<210> 140
<211> 345
<212> PRT
<213> Cinnamomum camphora
<400> 140

026172-005020PC - Sequence Listing_1

Pro	Asp	Trp	Ser	Met	Leu	Phe	Ala	Val	Ile	Thr	Thr	Ile	Phe	Ser	Ala
1				5					10					15	
Ala	Glu	Lys	Gln	Trp	Thr	Asn	Leu	Glu	Trp	Lys	Pro	Lys	Pro	Asn	Pro
			20					25					30		
Pro	Gln	Leu	Leu	Asp	Asp	His	Phe	Gly	Pro	His	Gly	Leu	Val	Phe	Arg
		35					40					45			
Arg	Thr	Phe	Ala	Ile	Arg	Ser	Tyr	Glu	Val	Gly	Pro	Asp	Arg	Ser	Thr
	50					55					60				
Ser	Ile	Val	Ala	Val	Met	Asn	His	Leu	Gln	Glu	Ala	Ala	Leu	Asn	His
65					70					75				80	
Ala	Lys	Ser	Val	Gly	Ile	Leu	Gly	Asp	Gly	Phe	Gly	Thr	Thr	Leu	Glu
				85					90					95	
Met	Ser	Lys	Arg	Asp	Leu	Ile	Trp	Val	Val	Lys	Arg	Thr	His	Val	Ala
			100					105					110		
Val	Glu	Arg	Tyr	Pro	Ala	Trp	Gly	Asp	Thr	Val	Glu	Val	Glu	Cys	Trp
		115					120					125			
Val	Gly	Ala	Ser	Gly	Asn	Asn	Gly	Arg	Arg	His	Asp	Phe	Leu	Val	Arg
	130					135					140				
Asp	Cys	Lys	Thr	Gly	Glu	Ile	Leu	Thr	Arg	Cys	Thr	Ser	Leu	Ser	Val
145					150					155					160
Met	Met	Asn	Thr	Arg	Thr	Arg	Arg	Leu	Ser	Lys	Ile	Pro	Glu	Glu	Val
				165					170					175	
Arg	Gly	Glu	Ile	Gly	Pro	Ala	Phe	Ile	Asp	Asn	Val	Ala	Val	Lys	Asp
			180					185					190		
Glu	Glu	Ile	Lys	Lys	Pro	Gln	Lys	Leu	Asn	Asp	Ser	Thr	Ala	Asp	Tyr
		195					200					205			
Ile	Gln	Gly	Gly	Leu	Thr	Pro	Arg	Trp	Asn	Asp	Leu	Asp	Ile	Asn	Gln
	210					215					220				
His	Val	Asn	Asn	Ile	Lys	Tyr	Val	Asp	Trp	Ile	Leu	Glu	Thr	Val	Pro
225					230					235					240
Asp	Ser	Ile	Phe	Glu	Ser	His	His	Ile	Ser	Ser	Phe	Thr	Ile	Glu	Tyr
				245					250					255	
Arg	Arg	Glu	Cys	Thr	Met	Asp	Ser	Val	Leu	Gln	Ser	Leu	Thr	Thr	Val
			260					265					270		

026172-005020PC - Sequence Listing_1

Ser Gly Gly Ser Ser Glu Ala Gly Leu Val Cys Glu His Leu Leu Gln
275 280 285

Leu Glu Gly Gly Ser Glu Val Leu Arg Ala Lys Thr Glu Trp Arg Pro
290 295 300

Lys Leu Thr Asp Ser Phe Arg Gly Ile Ser Val Ile Pro Ala Glu Ser
305 310 315 320

Ser Val Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp
325 330 335

Ile Asp Tyr Lys Asp Asp Asp Asp Lys
340 345

<210> 141
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic peptide

<400> 141
Lys Asp Glu Leu
1

<210> 142
<211> 6
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown: Higher plant fatty acyl-ACP thioesterase sequence

<400> 142
Leu Asp Met Asn Gln His
1 5

<210> 143
<211> 6
<212> PRT
<213> Unknown

<220>
<223> Description of Unknown: Algal fatty acyl-ACP thioesterase sequence

<400> 143
Met Asp Met Asn Gly His
1 5