

Figure 1

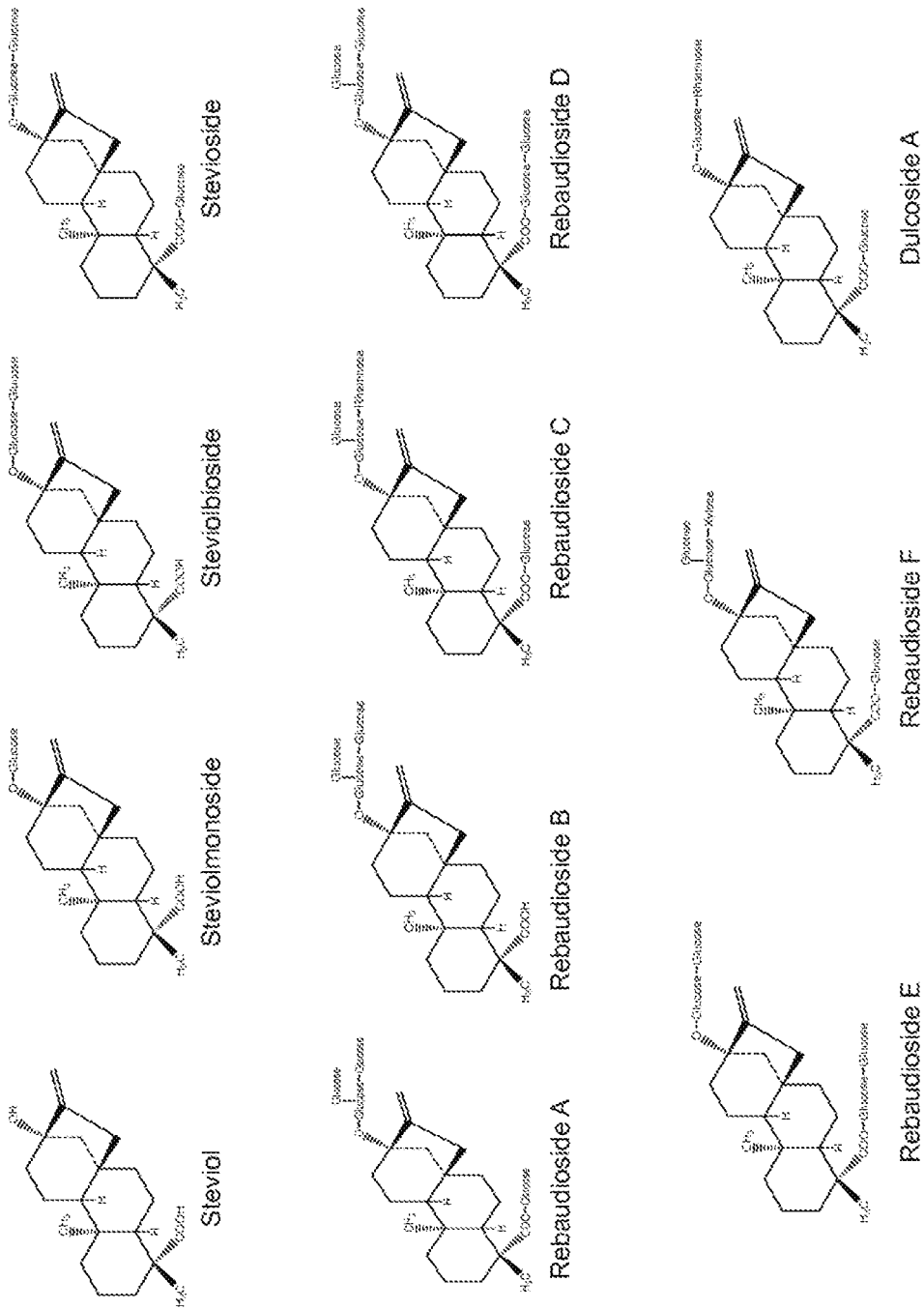


FIG 2A

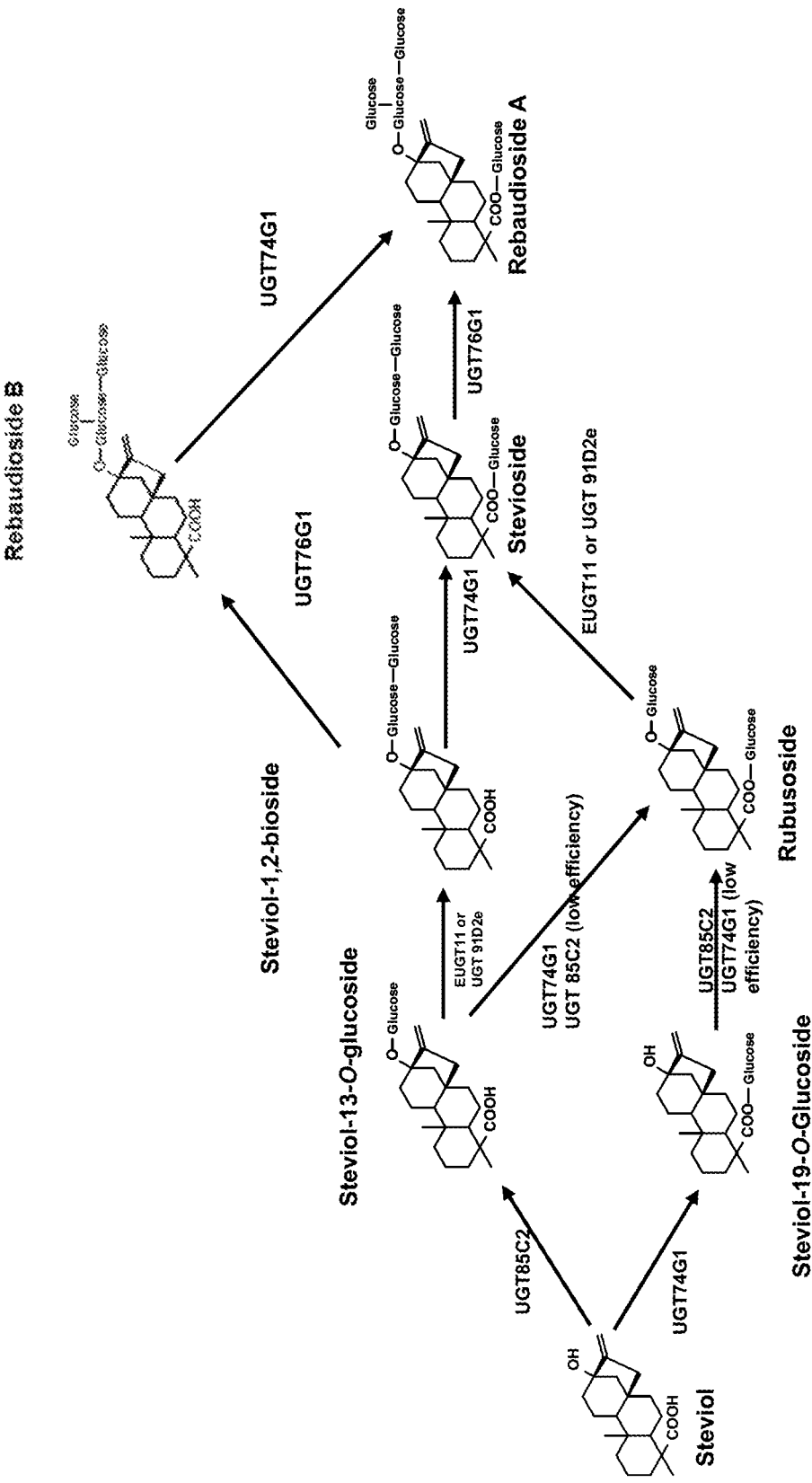


FIG 2B

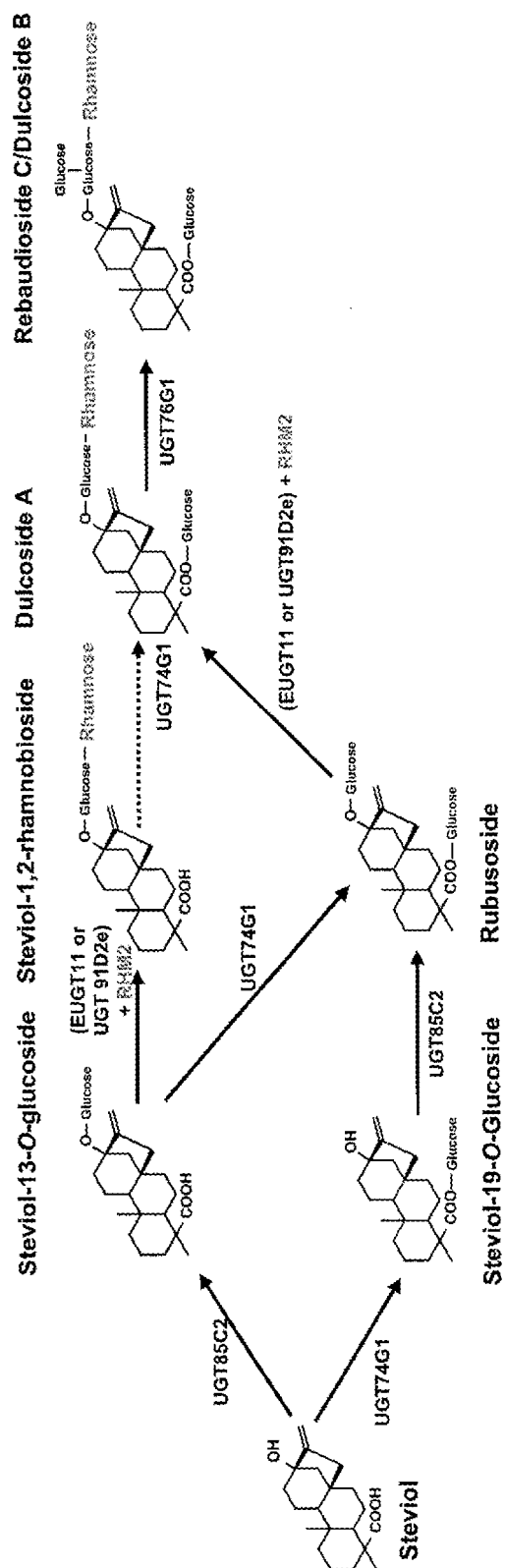


FIG. 2C

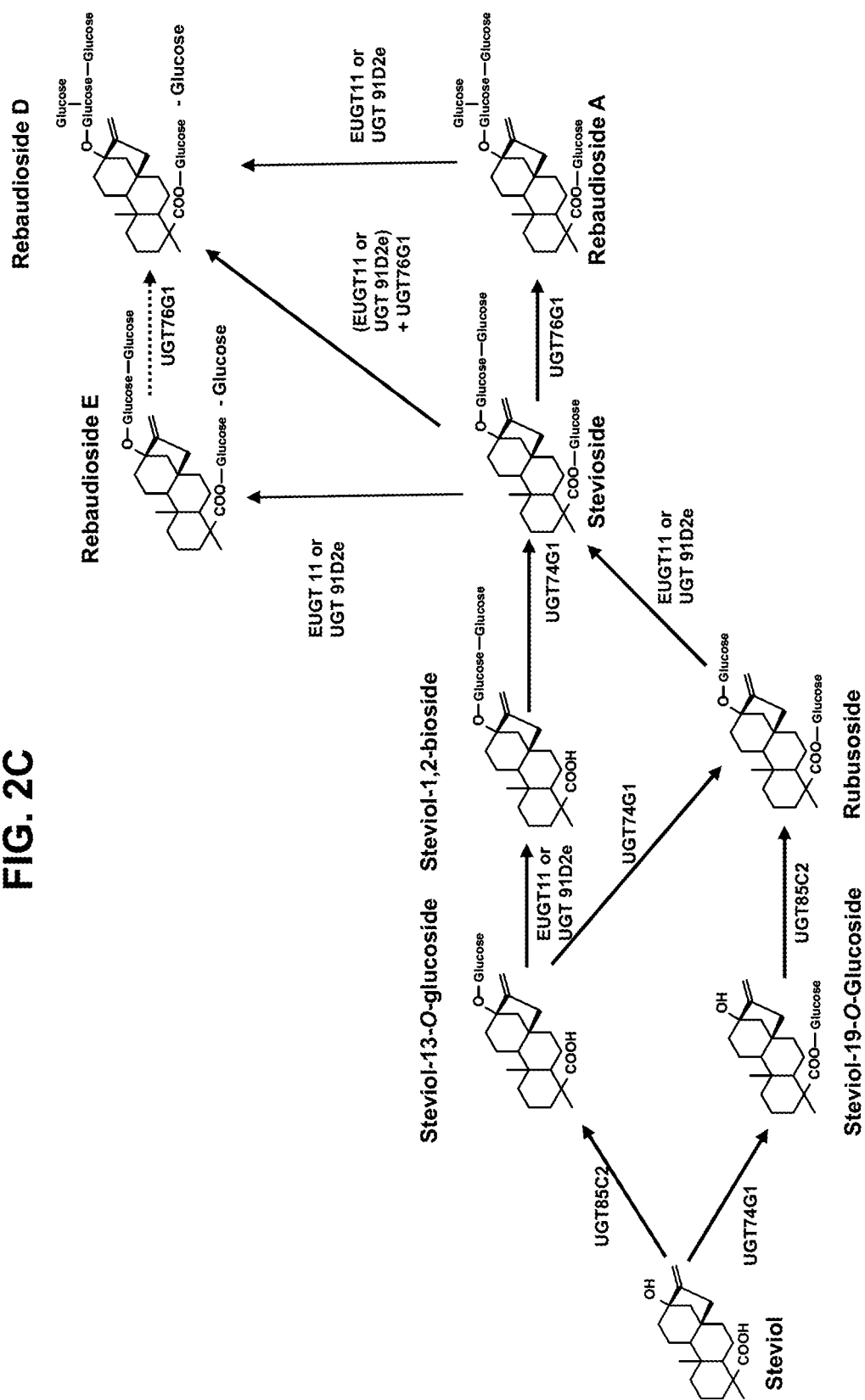


FIG. 2D

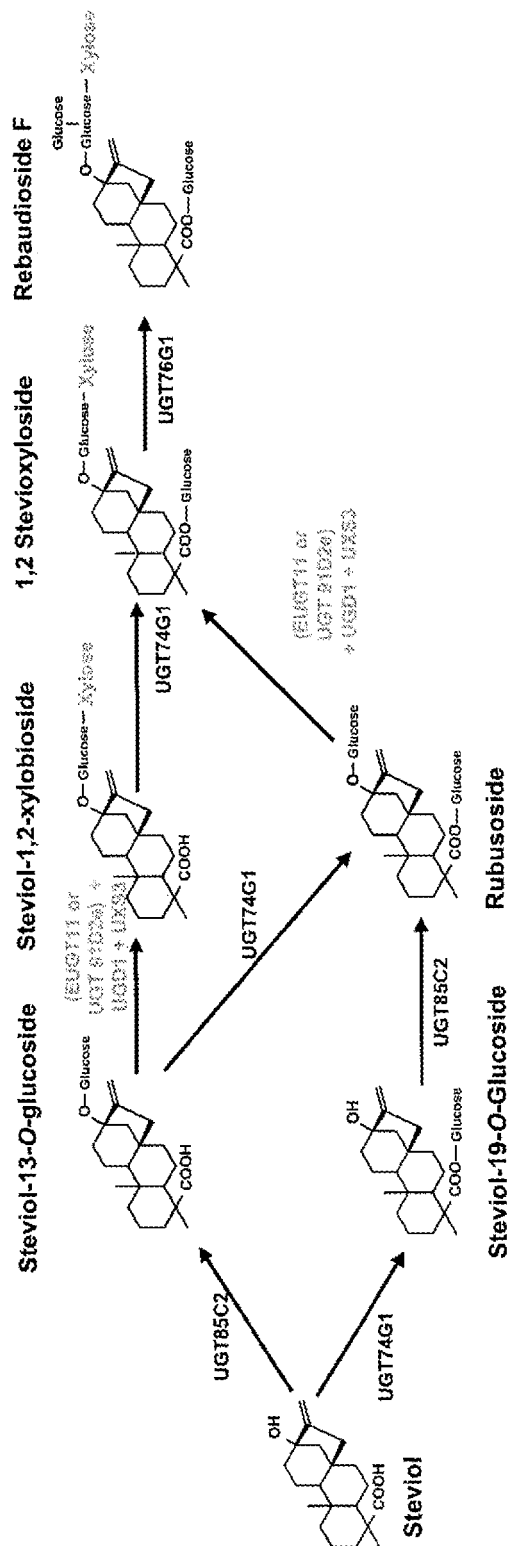


FIGURE 4

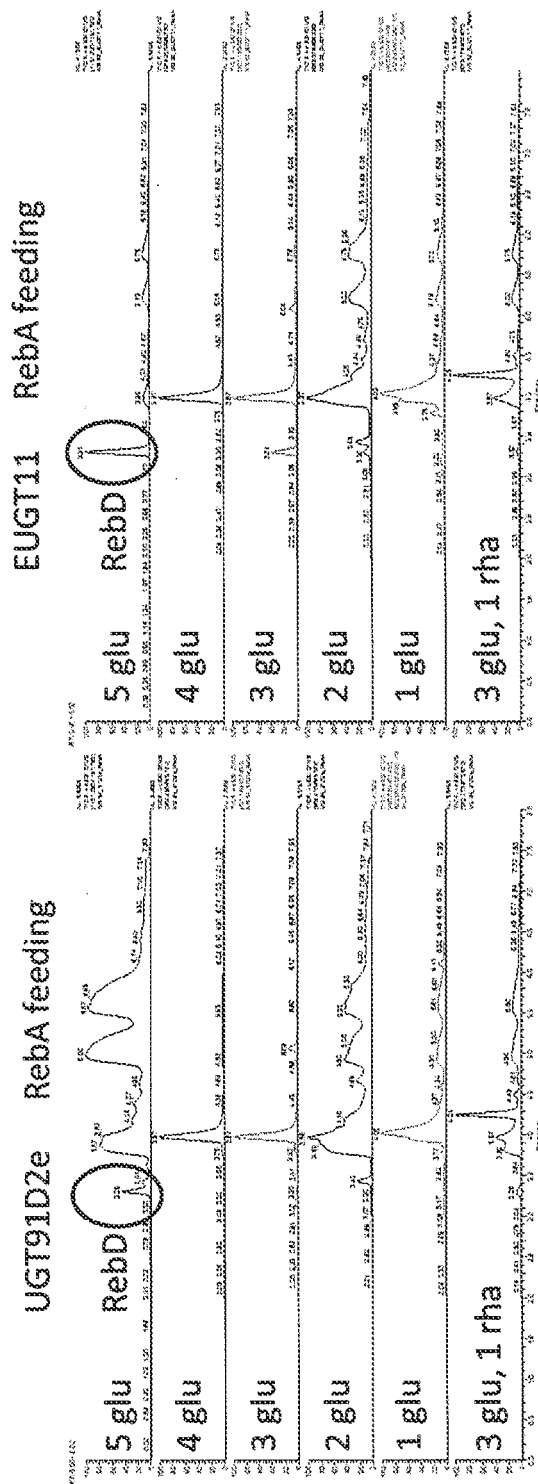


FIGURE 5

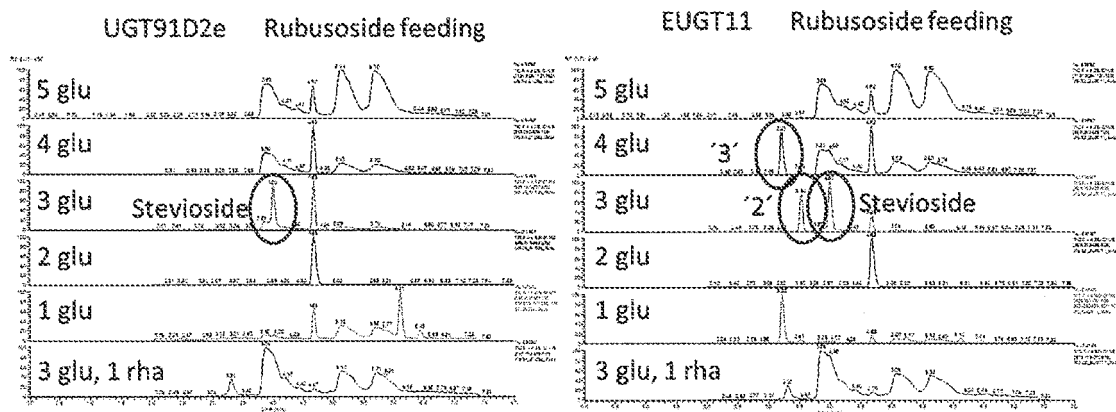


FIGURE 6

```
# fasta36-p-q-w 80-m 6-z 11-Z 10000-f-10-g-2 TMP.q TMP.q2.2
FASTA searches a protein or DNA sequence data bank
version 36.3.5a Jun, 2011 (preload8)
Please cite:
W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448

Query: TMP.q
1>>>QUERY - 462 aa
Library: TMP.q2
473 residues in 1 sequences
Statistics: (shuffled [128]) MLE statistics: Lambda= 0.1568; K=0.007058
statistics sampled from 1 (1) to 128 sequences
Algorithm: FASTA (3.7 Nov 2010) [optimized]
Parameters: BL50 matrix (15:-5), open/ext: -10/-2
ktup: 2, E-join: 1 (1), E-opt: 0.2 (1), width: 16
Scan time: 0.030
```

The best scores are:
QUERY

(473) 1003 236.0 2e-62 align
opt bits E(10000)

>>>QUERY, 462 aa vs TMP.q2 library

>>QUERY

initn: 1037 init1: 443 opt: 1003 Z-score: 1225.1 bits: 236.0 E(10000): 2e-62
Smith-Waterman score: 1104; 42.7% identity (67.6% similar) in 457 aa overlap (15-457:14-460)

Entrez Lookup Re-search database General re-search

QUERY	MEGYSYSSYAAAGMHVVICPWLAFGHLPLCLDLAQRLASRGHRVSEVSTPENISRLPFVRPALAPLVAFVALPLPRVEG	10	20	30	40	50	60	70	80
	...	...	...	...	...	...	...	...	...
QUERY	MATSDSIVDRKQLHVATFEPWLAFGHILPYQLSKLIAEKHKVSEFLSTTRNIQRLSS---HISPLINVVQLTLPRVOE	10	20	30	40	50	60	70	
	...	...	...	...	...	...	...	...	...
QUERY	LPDGAESTNDV-PHDPDPMVELHRRRAFDGLAAPFSEFLGTACADWIVDVETHW---AAAAAL-----EH---KVPCCAMMLL	90	100	110	120	130	140	150	
	...	...	...	...	...	...	...	...	...
QUERY	LPEDAEATIDVHPEDIPYL-----KKASDGLQPEVTRFLQHSQSPDWIIVDYTHYWLFLPSIAASLGISRAHPEVTTTPWAIAYM	80	90	100	110	120	130	140	150
	...	...	...	...	...	...	...	...	...

```

462 residues in 1 query sequences
473 residues in 1 library sequences
TcompLib [36.3.5a Jun, 2011(preload8)] (16 proc in memory [OG])
start: Thu Jul 7 13:19:40 2011 done: Thu Jul 7 13:19:40 2011
Scan time: 0.030 Display time: 0.010

Function used was FASTA [36.3.5a Jun, 2011(preload8)]

```

FIGURE 7

>EUGT-11 Protein.
 MD5GVSSSYAAAAAGMHVVICPWLAFGHLLPCLDLAQRASGRHVSFVSTPRNISRLPPVPALAPLVAFLPLPRVEGLPDGAESTNDVPHDRPDMVELHRRAFDGLAAPPSEFLGTA
 CADWVIVDFVHHWAAAAALEHKVPCAMMILGSAHMIASJADRRLEAETESPAAGQGRPAAPTFEVARMKLIRTKSSGMSLAERFSLTSSSLVVGRCVFEFEPETVPLSLTRGK
 PITLGLMPPLHEGRREDGEDATVRWLDAQPAKSVVVAIGSEVPLGVKEVHELALGLELAGTRFLWALRKPTGVSDADLLPAGFEERTGRGVWATRWVPMQMSILAHAAVGAFLTHCG
 WNSTIEGLMFGHPLMLPIFGDQGNARLIEAKNAGLQVARNDGDSFDREGVAAAIRAVAVEESSKVFOAKAKKLQEIADVMACHERYIDGFIQQLRSYKO

>EUGT-11 Native gene.
 ATGGACTCGGGTACTCTCTCTCTACGGCGCGCGCGGGATGCACGTGCTGATCTGCCCGTGGCTGGCTTCCGGCCACCTGCTCCCGTGCCTCGACCTCGCCAGCGGCTCGC
 GTGCGGGGCCACCGGTGCTGCTCGTCTCCACGCGCGGGAACATATCCCGCTCCCGCGGTGCGCCCGCGCTCGCGCGCTCGCTCGCTTCTGTCGGCTCGCGCTCGCGCGC
 GTCGAGGGCTCCCGACGCGCGGATCCCAACGACGTCGCCACGACGCGGACATGGTCGAGCTCCACCGAGGGCTTCGACGGCTCGCGCGCTCGCGCGCTTCCTGGAG
 TTCTTGGGCACTGGCTGCTGACGCTTCCACCACTGGCGCGAGCGCTCGAGCACAAAGTGCCATGTGCAATGATGTTGTTGGGCTCTGCAC
 ATATGATCGCTTCATAGCAGACAGCGCTCGAGCGCGGAGACAGAGTCGCTCGGCTCGCGGCGAGGACGCCAGCGCGCGCCCAACGTTGAGGTGGCGAGGAT
 GAAAGTTGATACGAACCAAGGCTCATCGGGAATGTCCTCGCGAGCGCTTCTCTGACGCTCTCGAGGACGCTCGTCTCGGGCGGAGCTCGTGGAGTTGAGCGCGGA
 GACCGTCCGCTCTGTCGACGCTCCGCGGTAAAGCTTATACCTTCTTGGCTTATGCCCGCTTGCATGAAGCGCGCGAGGACGGGAGGATGCCACCGTCCGCTGGCTC
 GACGCGCAGCGCGCAAGTCGTCGTACGTGCGCTAGGCGAGGTGCCACTGGAGTGGAGAAGTCCACGAGCTCGGCTCGGCTGGAGCTCGCGGGACGCGCTT
 CCTCTGGGCTCTAGGAAGCCCACTGGCTCTCCGACGCGGCTCTCCGCGCGCTTCGAGGAGCGCACGCGCGCGCTCGTGGCGACGAGATGGGTTCTCTCAGAT
 GAGCATACTGGCGACGCGCGTGGCGGCTTCTGACCCACTGGGCTGGAACTCGACCATCGAGGGCTCATGTTGGGCCACCGCTTATCATGCTGCCATCTTCGCGGAC
 CAGGACCGAACGCGCGCTAATCGAGCGCGAAGAACGCGGATTCAGAGTGGCAAGAAACGACGCGCATGGATGTTCCGACGAGAGGCTCGCGCGCGGATTCGTGCAG
 TCGCGTGGAGGAAGAACGACGCAAGTGTTCAGGCCAAAGCTGACGAGAGATCGTCGCGGACATGGCTGCCATGAGAGGTACATCGACGGATTCAATCAGCAA
 TTGAGATCTTACAAGGATTGA

>EUGT-11 Optimized DNA.
 ATGGATAGTGGTACTCTCTATCTATGCTGCTCGCGTGGTATGACAGTTGTGATCTGCCCTTGGTGGCTTGGTCACTGTTACACTGTTAGCCCAAGACTGGCC
 TCAAGAGGCCATAGAGTATCAITTTGTGCTACTCTAGAAATATCTCGTTTACCACGATCAGACTGCTAGCTCTCTAGTTGCAITTCGTTGCTCTTCCACTTCCAAGAGTAG
 AAGGATTGCCAGACGGCGTGAATCTAATGACGTACCACATGATAGACTGACATGGTGAATTCATAGAGAGCTTTGATGAGATGGCAGCTCCATTTCAGAGTTCTCTG
 GGCACAGCATGTGCAGACTGGTTATAGTCGATGTTTTCATCACTGGGCTGCTGCAGCGCATTTGGAACATAAGTGCCTTGTCTATGATGTTGTTAGGGTCAGCACACATGA
 TCGCATCCATAGCTGATAGAGATTGGAAAGAGCTGAACAGAAATCCCGCGCAGCAGGACAAGGTAGGCCAGCTGCCGCCCAACCTTTGAAGTGGCTAGAAATGA
 TTCTACTAAAGGTAGTTCAGGGATGAGTCTGCTGAAAGGTTTCTGACATTAATCTAGATCAATATGTTAGGTAGATCTCGCTCGAGTTCGAACCTGAAACAGTACCTT
 TACTATCTACTTTGAGAGGCAACCTATTACTTCTTGGTCTAATGCTTCCATTACATGAAGGAAGGAGAGAAGTGGTGAAGATGCTACTTGTAGGTGTTAGTATGCCAACCT
 GCTAAGTCTGTTTACGTTGCAATGGGTTCTGAGTACCCTAGGGTGGGAAAGGTGCAATGAATTAGCATTAGGACTTGAGCTGGCGGGAACAAGATTCCTTGGGCTTGA
 GAAACCAACCGGTGTTCTGACGCGGACTTGCTACCGAGTGGTTCGAGAGAGAACAGAGGCGCTGGTGTGTTGCTACTAGATGGGTCACAAATGAGTATCTAGCTCA
 TGCAGCTGAGGGGCTTCTAACCCATTGCGGTGGAACTCAACAATAGAGGAGGATGTTTGGTCACTCCATTATGTTACCAATCTTTGGCGATCAGGGACCTAACCGCAA
 GATTGATTGAGGCAAGAACGAGGCTGCAGGTTGCAGGTAATGATGGTGTGTTCTTGTATAGAGAAGGCGTTGCAGCTGCCATCAGAGCAGTCCCGTTGAGGAAGAGT
 CATCTAAAGTTTCCAAAGCTAAGGCCAAAAAATTACAAGAGATTGTGGGTGACATGGCTTGTTCACGAAAGATACATCGATGGTTTCATCCAAACAATTGAGAAAGTTATAAGACTAA

Figure 8

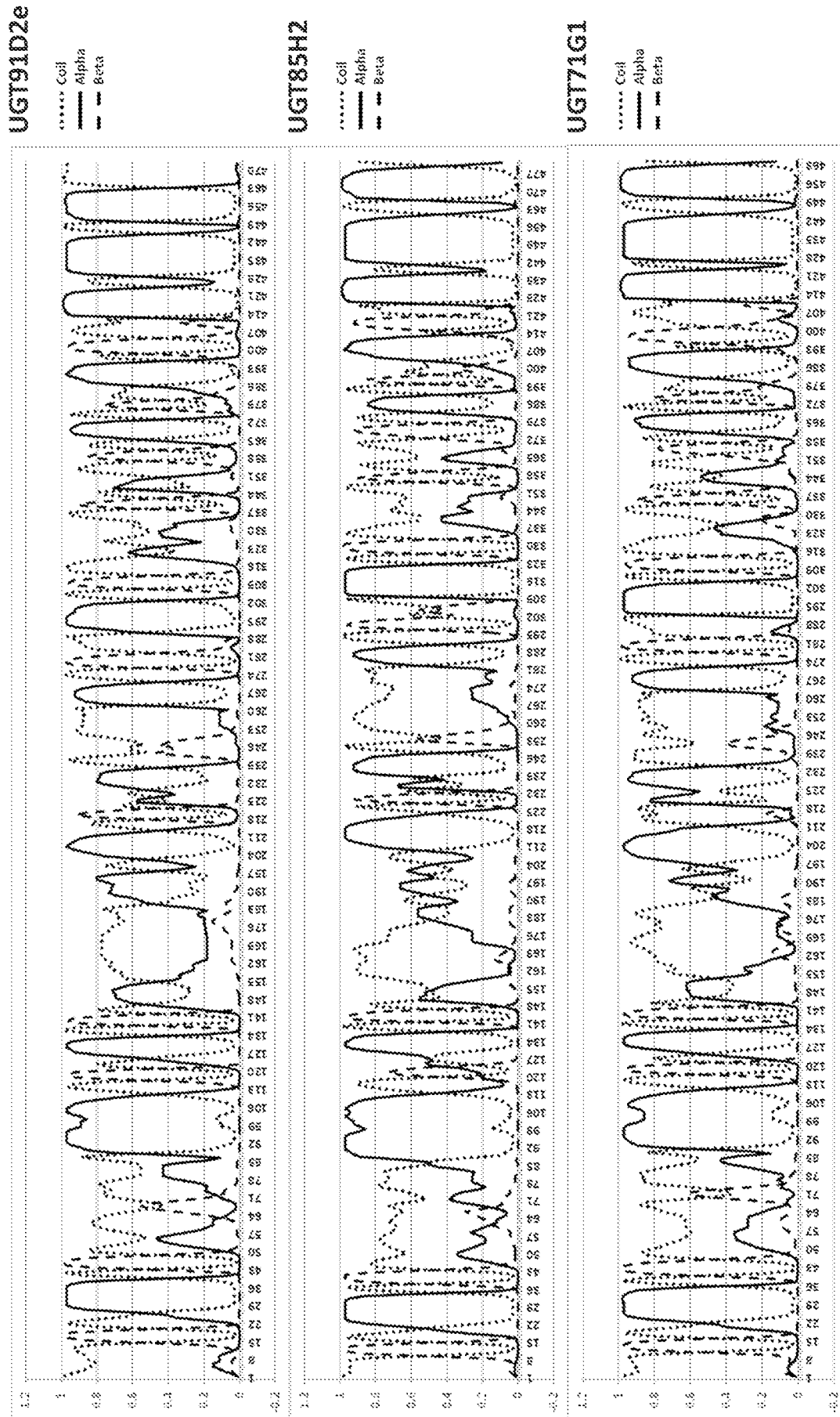


FIGURE 9

UGT91D1	MATSDSI VDD	RKQLHVATFP	WLAFGHILPF	LQLSKLIAEK	GHKVSFLSTT	RNIQRLSSHI	60
UGT91D2e	MATSDSI VDD	RKQLHVATFP	WLAFGHILPY	LQLSKLIAEK	GHKVSFLSTT	RNIQRLSSHI	60
			***			***	120
UGT91D1	SPLINVVQELT	LPRVQELPED	AEATTDVHPE	DIIYLKKAVD	GLQPEVTRFL	EQHSPDWIYY	120
UGT91D2e	SPLINVVQELT	LPRVQELPED	AEATTDVHPE	DIIYLKKAVD	GLQPEVTRFL	EQHSPDWIYY	120
			***			***	180
UGT91D1	DETHYWLPSI	AASLGISRAV	FCVTPWTIA	YLA PSSDAMI	NDSGRITVE	DLTTPPKWFP	180
UGT91D2e	DYTHYWLPSI	AASLGISRAH	FSVTPWAI A	YMGPSADAMI	NDSGRITVE	DLTTPPKWFP	180
			***			***	240
UGT91D1	FPTKVCWRKH	DLARMEPYEA	PGISDGYRMG	MVFKGSDCLL	EKCYHEFGTQ	WLPLETLHQ	240
UGT91D2e	FPTKVCWRKH	DLARLVYPYKA	PGISDGYRMG	LVFKGSDCLL	SKCYHEFGTQ	WLPLETLHQ	240
			***			***	300
UGT91D1	VPVVPVGLLP	PEIPGDEKDE	TWVSIKKWLD	GKQKGSVVYV	ALGSEALVSD	TEVVELALGL	300
UGT91D2e	VPVVPVGLLP	PEIPGDEKDE	TWVSIKKWLD	GKQKGSVVYV	ALGSEMLVSD	TEVVELALGL	300
			***			***	360
UGT91D1	ELSGLPFVWA	YRKPKGPAKS	DSVELPDGFV	ERTDRGLVW	TSWAPQLRIL	SHESVCGFLT	360
UGT91D2e	ELSGLPFVWA	YRKPKGPAKS	DSVELPDGFV	ERTDRGLVW	TSWAPQLRIL	SHESVCGFLT	360
			***			***	420
UGT91D1	HCGSGSIVEG	LMFGHPLIML	PIFGDQPLNA	RLLEDKQVGI	EIPRNEEDGC	LTKEVARSL	420
UGT91D2e	HCGSGSIVEG	LMFGHPLIML	PIFGDQPLNA	RLLEDKQVGI	EIPRNEEDGC	LTKEVARSL	420
			***			***	480
UGT91D1	RSVVVENEGE	IYKANARALS	KIYNDTKVEK	EYVSQFVDYL	EKNARAVAI D	HES	474
UGT91D2e	RSVVVEKEGE	IYKANARELS	KIYNDTKVEK	EYVSQFVDYL	EKNARAVAI D	HES	474

FIGURE 10

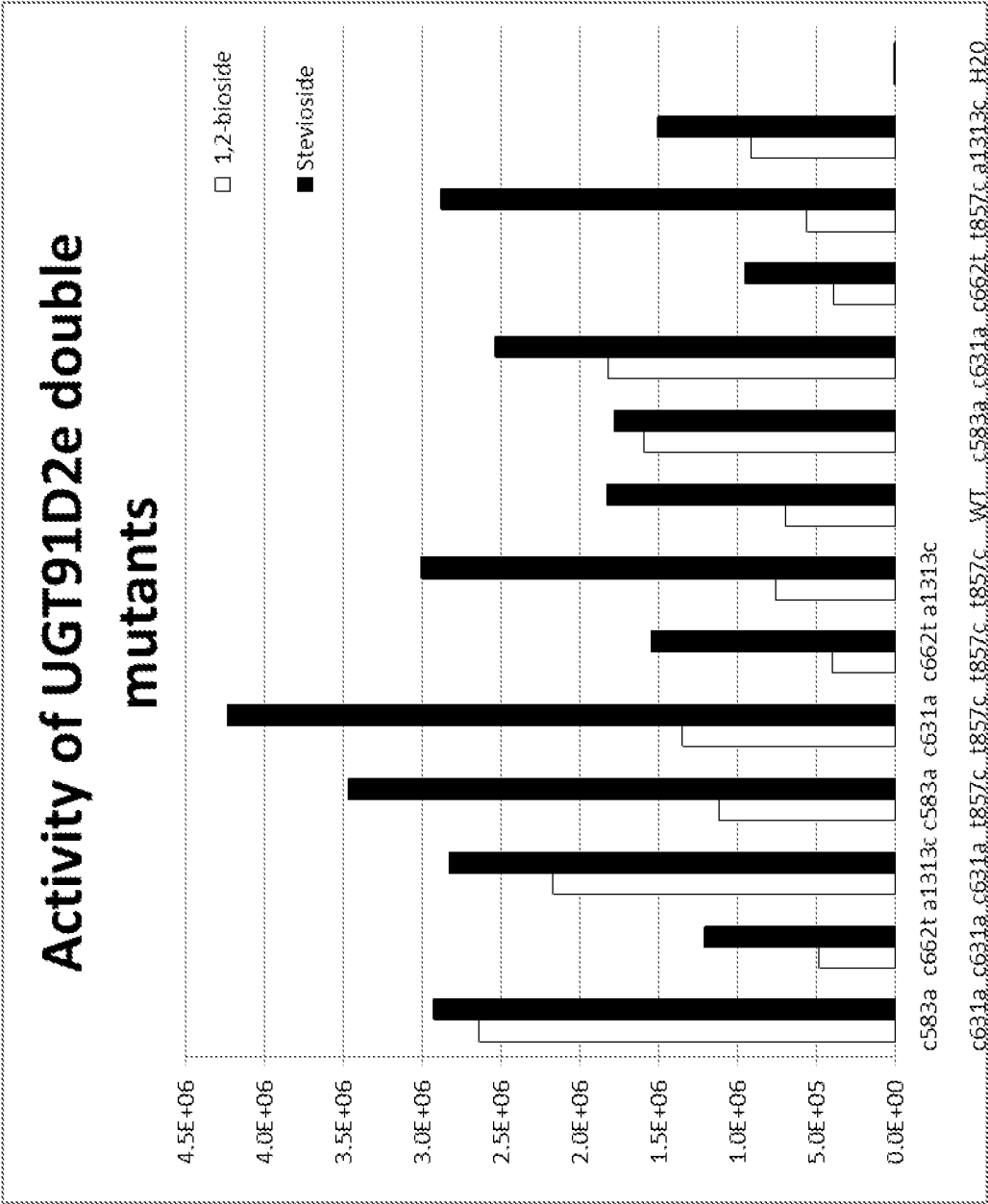


FIGURE 11

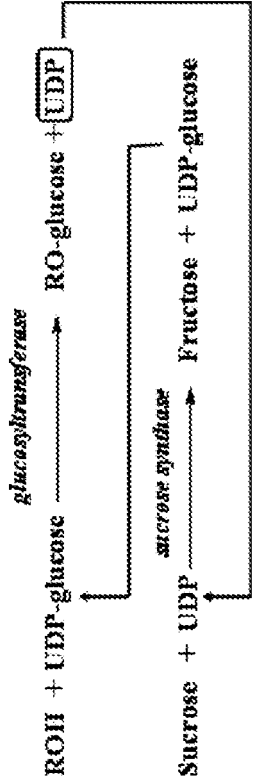


FIGURE 12

FIG. 12A

ATGGAGGCTTCATATCTATACATTTCCATTCTTCTGCTTCTAGCTTCTGTAACCTTTCAACCACCCAACTTCGTCTGTAATCCGCCAATCT
 ACCGCCGACGGTGTTCCTCATCCATCCCATAAATCCGTCACTCTACCTCCTCAAAAAACCACTCTATAGAACACTAGCCAAAAATTGCCG
 CCAATACGGCCCTATCTCCAACTCCAACCTAGGTACCGCGTGTCTCGTAATCTCTCCCTTCCGCCGCCGAAGAATGCTTCACC
 AACACGACGTTATCTTCGCCAACCGTCCGAAGACGCTATTTCGGAAAAATTGTAGGGGTACCAAGCTCGGGTCGCTGCTACGGCGA
 CCAGTGGCGCAACCTCCGCCGCTTGCAATCCATCGAGATTCTATCAGTCCACCGGCTCAACGAGTTTCAAGACATACGTGTTGACGAAA
 ACCGGCTTCTGATCCGTAAACTAAGATCCAGTCTCTCCGGTGACTCTGATAACGGTGTTTTACGCACCTAACGTTAAATGTGATTATG
 AGAATGATCTCCGGAAAGAGGTATTTCCGACTCGGGTGATCGGGAATTGGAGGAGGAAGGAAGCGATTCCGGGAGATACCTCGATGAGAC
 ACTTTTGCTCGCGGTGCTTCTAAATGTTGGGATTAATTGCCGATTCTGAATTGGTTGGGGTGAAGAGCTTGGAGAAAGAACTAATCG
 CATTGCAGAAAAAGAGAGATGATTTCTTTACGGGACTGATCGAGCAAGTTCGGAAATCTAGAGGGCTAAAGTGGGAAAAGGAAGGAAG
 ACGATGATCGAGTTGTTGTTATCCCTACAAGAACTGAACCTGAGTATTACACTGACGCCATGATCCGATCATTTTGCTGGGTTTATT
 AGCAGCAGGAGTGATACATCGGCTGGAACCTATGGAATGGCGATGTCTCTTTTGGTAACCAACCCGCACGTATTAATAAAGGCACAGG
 CTGAAAATTGATCGAGTCAATCGGCAACAACCGTCTAATTGATGAGTCCGACATAGGGAATATACCTTACATTGGTTGCATCATAAACGAG
 ACGCTCAGATTGTACCTCGGGGCCGTTGCTATTTCCTCATGAGTCATCAGCGGACTGTGTATCAGCGGGTACAACATCCCTCGTGG
 GACGATGCTTATTGTCAACCAATGGCGGATACATCATGACCCAAAGGTGTGGACGACCCGTGAGACATTCAAAACCGGAAAGATTCCAAG
 GGCTTGAAGGACACGAGACGGGTTTAAGCTGATGCCTTTGGGTCCGGAAGGAGGGGTTGTCCAGGGGAGGATTGGCGATTCTGTTTG
 CTTGGGATGACTCTCGGGTCGGTTATCCAATGCTTTGATTGGGAACGAGTCGGTGACGAGATGGTTGATATGACCGAAAGGTCCTTGGGGT
 CACGTTGCCTAAAGCTGTACCATTAGTCGCCAAGTCAAGCCGCTTCCGAAATGACGAATCTACTCTCTGAGCTATGA

FIG. 12B

ATGGAAGCCTCTTACCTATACATTTCTATTTTGTCTTTTACTGGCATCATACCTGTTCACCACTCAACTTAGAAGGAAGAGCGCTAATCT
 ACCACCAACCGTGTTCCTCATCAATACCAATCATTTGGACACTTATACTTACTCAAAAAGCCTCTTTATAGAACTTTAGCAAAAAATTGCCG
 CTAAGTACGGACCAATACTGCAATTACAACCTCGGTACAGACGTGTTCTGGTGATTTCTCTCACCATCAGCAGCAGAAGAGTGCTTTACC
 AATAACGATGTAACTCTCGCAAAATAGACCTAAGACATTGTTTGGCAAAATAGTGGGTGGAACATCCCTTGGCAGTTTATCCTACGGCGA

TCAATGGCGTAATCTAAGGAGAGTAGCTTCTATCGAAAATCCTATCAGTTCATAGGTTGAACGAATTTTCATGATATCAGAGTGGATGAGA
 ACAGATTGTTAATTAGAAAACCTTAGAAGTTTCATCTTCTCTGTTACTCTTATAACAGTCTTTTATGCTCTAACATTGAACGTCATTATG
 AGAATGATCTCTGGCAAAAGATAATTCGACAGTGGGATAGAGAATTGGAGGAGGAAGTAAGAGATTTCGAGAAAATCTTAGACGAAAC
 GTTGCTTCTAGCCGGTGCTTCTAATGTTGGCGACTACTTACCAATATTTGAACGTGTTGGAGTTAAGTCTCTTGAAAAAGAAATTGATCG
 CTTTGCAGAAAAAGAGAGATGACTTTTCCAGGGTTTGAATTGAACAGGTTAGAAAATCTCGTGGTGTAAAGTAGGCAAGGTAGAAAA
 ACGATGATCGAACTCTTATTATCTTTGCAAGAGTCAGAACCTGAGTACTATACAGATGCTATGATAAGATCTTTTGCTCCTAGTCTGCT
 GGCTGCAGGTAGTGATACTTCAGCGGGCACTATGGAATGGCCATGAGCTTACTGGTCAATCACCCACATGTATTGAAGAAAAGCTCAAG
 CTGAAAATCGATAGAGTTATCGGTAATAACAGATTGATTGACGAGTCAGACATTGGAATAATATCCCTTACATCGGGTGTTATCAATGAA
 ACTCTAAGACTCTATCCAGCAGGGCCATTGTTGTTCCACATGAAAAGTTCTGCCGACTGCGTTATTTCCGGTTACAATATACCTAGAGG
 TACAATGTTAATCGTAAACCAATGGCGGATTCAATCAGGATCCCTAAAGTCTGGGATGATCCTGAAACCTTTAAACCTGAAAGATTTCAAG
 GATTAGAGGAACTAGAGATGGTTTCAAACTTATGCCATTTCGGTCTGGGAGAGAGAGATGTCCAGGTGAAGGTTTGGCAATAAGGCTG
 TTAGGGATGACACTAGGCTCAGTGATCCAATGTTTGTGTTGGGAGAGAGTAGGAGATGAGATGGTTGACATGACAGAAAGGTTTGGGTGT
 CACACTTCCTAAGGCCGTTCCATTAGTTGCCAAATGTAAAGCCACGTTCCGAAATGACTAATCTCCTATCCGAACTTTTAA

FIG. 12C

MEASYLYISILLLLASYLFTTQLRRKSANLPPTVFPSIPIIGHLYLLKKPLYRTLAKIAA
 KYGPILQLGLYRRVLVSSPSAAEECFNNNDVIFANRPKTLFGKIVGGTSLGSLSYGDQ
 WRNLRVASIEILSVHRLNEFHDIRVDENRLLIRKLRSSSSPVTLITVFYALTINVMRM
 ISGKRYFDSGDRELEEEGKRFRFRIIDETLLLAGASNVGDYLPILNLGVKSLEKKLIAAQ
 KKRDDFFQGLIEQVRKSRGAKVGKGRKTMIELLLSLQSESEPEYTTDAMIRSFVLGLLAAG
 SDTSAGTMEWAMSLLVNHPHVLKKAQAEIDRVIGNNRLIDESDIGNIPIYIGCIINETLRL
 YPAGPLLPHESSADCVISGYNIPRGTMILVNQWAIHHDPKVWDDPETFKPERFQGLEGT
 RDGFKLMPFGSGRRGCPGEGLAIRLLGMTLGSVIQCQFDWERVGDENVDMTEGLGVTLPKA
 VPLVAKCKPRSEMTNLLSEL

FIGURE 13A

MPFGIDNTDFTVLAVLLVYVKNRSIKELLMSDDGDITAVS
 SGNRDIAQVVTENNKNYLVLVYASQTGTAEDYAKKFSKELVAKFNLVNMCADVENYDFF
 SLNDVPVIVSIFISTYGEFDPDGAVNFEDFICNAEAGALSRLRYNMFGLGNSTYEFF
 NGAAKKAEKHLAAGAIRLGLGEADDGAGTTDEDYMAWKDSILEVLKDELHLDEQEA
 KFTSQFQYTVLINEITDSMSLGEPSAHYLPHQNLNRNADGIQLGPFDLSPYIAPIVKS
 RELFSSNDRNCIHSEFDLSGSNIKYSTGDHLAVWPSNPLEKVEQFLSIFNLDPETIFD
 LKPLDPTVKVPFPTTTIGAAIKHYLEITGPVSRQLFSSLIQFAPNADVKEKLTLLSK
 DKDQFAVEITSKYFNIADALKYLSDGAKWDTVPMQFLVESVPQMTPRYSSSSLSSE
 KQTVHVTSIVENFPNPELPDAPPVVGVTNLLRNIQLAQNNVNIAETNLPVHYDLNGP
 RKL FANYKLPVHVRRSNFRLPSNPSTPVMIGPGTGVPFRGFIRERVAFLSQQKKG
 NNVSLGKHILFYGSRNTDDELYQDEWPEYAKKLDGSFEMVVAHSRLPNTKKVYVQDKL
 KDYEDQVFEMINNGAFIYVCGDAKMAKGVSTALVGILSRGKSITTDEATELIKMLKT
 SGRYQEDVW (***S. cerevisiae*** **CPR polypeptide encoded by NCP1**).
 1 mtsalyasdl fkqlksimgt dslsddvvlv iattslalva gfvllwktk tadrsgelkp
 61 lmipkslmak dedddldlgs gktrvsiffq tqgtaeafa kalseeikar yekaavkvid
 121 lddyaadddq yeeklkketl affcvatygd geptdnaarf ykwfteener diklqqalay
 181 vfaalnqrqe hfnkigivld eelckkgakr lievglgddd qsiedddfnaw keslwseldk
 241 llkdeddksv atpytavipe yrvvthdprf ttqksmesnv angnttidih hpcrvdvavq
 301 kelhthesdr scihlefdis rtgityetgd hvgvyaenhv eiveeagkll ghsldlvfsi
 361 hadkedgspl esavpppfpg pctlgtglar yadllnpprk salvalaaya tepseaeklk
 421 hltspdgkde ysqwivasqr sllevmaafp sakpplgvff aaiaprlqpr ysysssprl
 481 apsrhvhtsa lvygptptgr ihkgvcstwm knavpaeksh ecsgapifir asnfklpsnp
 541 stpivmvgpg tglapfrgfl qermalkedg eelgssllff gcrnrgmdfi yedelnnfvd
 601 ggviselima fsregaakey vqhkmekaa qvwdlikeeg ylyvcgdakg mardvhrtlh
 661 tivqegeqvs sseaaiavkk lqtegrylrd vw (***Arabidopsis thaliana*** **CPR polypeptide encoded by**
ATR1)

1 msssssts midlmaaiik gepvivsdpa nasayesvaa elssmlienr qfamivttsi
 61 avligciavl vmrrsgsgns krveplkplv ikpreeeidd grkkvtiffg tqtgtaegfa
 121 kalgeear yektrfkivd lddyaaddde yeeklkedv afflatygd geptdnaarf
 181 ykwfitegnr gewlknkyg vfglgnrque hfnkvakvvd dilveggaqr lvqvglgddd
 241 qcieddftaw realwpeldt ilreegdtav atpytaavle yrvsihdsed akfnditlan
 301 gngytfvdaq hpykanvavk relhtpesdr scihlefdia gsgltmklgd hvglcdnls
 361 etvdealrll dmspdtyfsl haekedgtpi ssslpppfpp cnlrtaltry acllsspkks
 421 alvalaahas dpteaerlkx laspagkdey skwvvesqrs lilevmaefps akpplgvffa
 481 gvaprlqprf ysissppkia etrihvtcal vyekmptgri hkgvcstwmk navpyeksek
 541 lflgrpifvr qsnfklpsds kvpiimigpg tglapfrgfl qerlalvesg velgpsvlff
 601 gcrnrrmdfi yeeelqrfev sgalaelsva fsregptkey vqhkmmdkas diwnmisqga
 661 ylyvcgdakg mardvhrslh tiaqeqgsmd stkaegfvkn lqtsgrylrd vw (**A. thaliana CPR**
polypeptide encoded by ATR2.)

>Stevia rebaudiana CPR7

MQSEVEASTIDLMTAVLKDTVIDTANASDNGDSKMPPALAMMFEIRDLLILTTSAVL
 VGCFFVLVWKRSSGKKSGKELEPPKIVVPKRRLEQEVDDGKKKVTIFFGTGTAEQFAK
 ALFEEAKARYEKAFFKVIDDDYAADLDEYAEKLEKETYAFFFLATYGDGEPTDAAKFY
 KWFTEGDEKGVWLQKLQYGVFGLGNRQYEHFNKIGIVVDDGLTEQGAKRIVPVGLGDDDDQ
 SIEDDFSARKELVWPELDLLRDEDDKAAATPYTAAIPEYRVVFHDKPDAFSDDDHTQTNG
 HAVHDAQHPCRSNVAVKKELHTPESDRSCTHLEFDISHTGLSYETGDHVGVCENLIEVV
 EEAGKLLGLSTDYFSLHIDNEDGSLGGPSLQPPFPCTLRKALTNYADLLSSPKKSTL
 LALAAHASDPTAERLRFASREGKDEYAENWVANQRSLEVMEAFPSARPPLGVFFAAV
 APRLQPRYYSISSSPKMEPNRIHVTCALVYEKTPAGRIHKGICSTWMKNAVPLTESQDCS
 WAPIFVRTSNFRLPIDPKVPVIMIGPGTGLAPFRGFLQERLALKESGTELGSSILFFGCR
 NRKVDYIYENELNFFVENGALSELDVAFSRDGPTKEYVQHKMTQKASEIWNMLSEGAYLY
 VCGDAKGMADVHRTLHTIVQEQGSLDSSKAELVYKNLQMSGRYL RDVW

> Stevia rebaudiana CPR8

MQSNSVKISPLDLVTALFSGKVLDTSNASESGESAMLP TIAMENRELLMILTTSVAVL
 IGCVVVLVWRRSSTKKSALEPPVIVVPKRQVEEVDGKKKVTVFFGTGTAEQFAKAL

VEEAKARYEKAVFKVIDLDDYAADDEYEEKLKKESLAFFFLATYGDGEPTDNAARFYKW
FTEGDAKGEWLNKLQYGVFGLGNRQYEHFNKIAKVVDGGLVEQQAKRLVPVGLGDDDDQCI
EDDFTAWKELVWPELDQLLRDEDDTTVATPYTAAVAEYRVVFHEKPDALSEDYSYTNNGHA
VHDAQHPCRSNVAVKKELHSPESDRSCTHLEFFDISNTGLSYETGDHVGVCENLSEVVND
AERLVGLPPDTYSSIHDTSEDCSPLGGASLPPFPFCTLRKALTCYADVLSSPKKSALLA
LAAHATDPSEADRLKFLASPAGKDEYSQWIVASQRSLLLEVMEAFPSAKPSLGVFFASVAP
RLQPRYYSISSSPKMAPDRIHVTICALVYEKTPAGRIHKGVCSTWMKNAVPMTESQDCSWA
PIYVRTSNFRLPSPDKVPVIMIGPGTGLAPFRGFLQERLALKEAGTDLGLSILFFGCRNR
KVDFIYENELNNFVETGALSELIVAFSREGPTKEYVQHKMSEKASDIWNLLSEGAYLYVC

FIGURE 13B

>ATR1 codon optimized by DNA2.0.

ATGACTTCTGCACCTTTATGCCTCCGATCTTTTCAAACAATTGAAAAGTATCATGGGAACGGATTC
TTTGTCGGATGATGTTGTATTAGTTATTGCTACAACCTTCTCTGGCACTGGTTGCTGGTTTCGTTGT
CTTATTGTGGAATAAGACCAACGGCAGATCGTTCCGGCGAGCTAAAGCCACTAATGATCCCTAAG
TCTCTGATGGCGAAAGATGAGGATGATGACTTAGATCTAGGTTCTGGAATAACGAGAGTCTCT
ATCTTCTTCGGCACACAACCCGGAACAGCCGAAGGATTCGCTAAAGCACTTTCAGAAGAGATCA
AAGCAAGATACGAAAAGGGCGCTGTAAAAGTAATCGATTTCGATGATTACGCTGCCGATGATG
ACCAATATGAGGAAAAGTTGAAAAGGAAACATTTGGCTTTCTTTTGTGTAGCCACGTAATGGTGA
TGGTGAACCAACCGATAACGCCGCAAGATTCTACAAGTGGTTTACTGAAGAGAACGAAAAGAGA
TATCAAGTTGCAGCAACTTGCTTACGGCGTTTTCGCTTAGGTAACAGACAATACGAGCACTTTA
ACAAGATAGGTATTGCTTTAGATGAAGAGTTATGCAAAAAGGGTGCGAAGAGATTGATTGAAG
TCGGTTTAGGAGATGATCAATCTATCGAGGATGACTTTAATGCATGGAAGGAATCTTTGTG
GTCTGAATTAGATAAGTTACTTAAGGACGAAGATGATAAATCCGTTGCCACTCCATACACAGCC
GTCAATCCAGAAATATAGAGTAGTTACTCATGATCCAAAGATTCAACAACACAGAAATCAATGGAAA
GTAATGTGGCTAATGGTAATACTACCATCGATAATTCATCATCCATGTAGAGTAGACGTTGCAGTT
CAAAAGGAATTGCACACTCATGAATCAGACAGATCTTGCATACATCTTGAATTTGATATATCAG
TACTGGTATCACTTACGAAACAGGTGATCACGTGGGTGTCTACGCTGAAAACCATGTTGAAAT
GTAGAGGAAGCTGGAAAAGTTGTTGGGCCATAGTTTAGATCTTGTCTTCTCAATTCAATGCCGATA
AAGAGGATGGCTCACCACTAGAAAAGTGCAGTGCCTCCACCATTTCCAGGACCATGCACCCCTAGG
TACCGGTTTAGCTCGTTACGGGATCTGTTAAATCCTCCACGTAATCAAGCTCTAGTGGCCTTGG
CTGCGTACGCCACAGAACCTTCTGAGGCAGAAAACCTGAAACATCTAACTTCACCAGATGGTAA
GGATGAATACTCAATGGATAGTAGCTAGTCAACGTTCTTTACTAGAAGTTATGGCTGCTTTCC
CATCCGCTAAACCTCCTTTGGGTGTTTCTTCGCCGCAATAGCGCCTAGACTGCAACCAAGATAC
TATTCAATTTTCATCCTCAGCTAGACTGGCACCATCAAGAGTTTCATGTCAATCCGCTTAGTGATC
GGTCCAACTCCTACTGGTAGAATCCATAGGGCGTTTGTTCACATGGATGAAAACGCGGTTTC
CAGCAGAGAAGTCTCACGAATGTTCTGTGTCTCCAAATCTTTATCAGAGCCTCCAACCTCAAACTG
CCTTCCAATCCTTCTACTCCTATTGTTCATGTTGTCGGTCTGTGACAGGTCCTTGTCCATTTCAGAGGT
TTCTTACAAGAGAGAAATGGCCTTAAAGGAGGATGGTGAAGAGTTGGGATCTTCTTTGTGTTT
TCGGCTGTAGAAAACAGACAAAATGGATTTTCATCTACGAAGATGAACCTGAATAAATTTGTAGATCA
AGGAGTTATTTACAGAGTTGATAATGGCTTTTCTAGAGAAGGTGCTCAGAAGGAGTACGTCCAA

CACAAAATGATGAAAAAGCCGCACAAAGTTTGGGACTTAATCAAAGAGGAAGGCTATCTATAT
 GTCTGTGTGATGCAAAAGGGTATGGCAAGAGATGTTTACAGAAACACTTCATACACTATAGTCCAG
 GAACAGGAAGCGGTTAGTTCTTCTGAAGCGGAAGCAATTGTGAAAAAGTTACAAACAGAGGGA
 AGATACTTGAGAGATGTGTGGTAA
 >ATR2 codon optimized by Genscript
 ATGCTTCCCTCTCCTCTTCCAGTACCTCTATGATTGATGGCTGCTATTATTAAAGGTGAA
 CCAGTTATCGTCTCCGACCCAGCAAAATGCCTCTGCTTATGAATCAGTTGCTGCAGAAATGTCTTC
 AATGTTGATCGAAAAACAGACAAATTCGCCATGATCGTAACTACATCAATCGCTGTTTTGATCGGTT
 GTATTGTCATGTTGGTATGGAGAAAGATCCGGTAGTGGTAAATCTAAAAAGAGTCGAACCTTTGAA
 ACCATTAGTAATTAAGCCAAGAGAAGAAATAGATGACGGTAGAAAAAGAAAGTTACAATATT
 TTTCCGGTACCCAAACTGGTACAGCTGAAGGTTTTGCAAAAGCCTTAGGTGAAGAAGCTAAGGC
 AAGATACGAAAAGACTAGATTCAAGATAGTCGATTGATGACTATGCCGCTGATGACGATGA
 ATACGAAGAAAAAGTTGAAGAAAGAGATGTGCATTTTCTTTTGGCAACCTATGGTGACGGT
 GAACCAACTGACAAATGCAGCCAGATTCTACAAATGGTTTACAGAGGGTAAATGATCGTGGTGAAT
 GGTTGAAAAAAGTTAAAGTACGGTGTTTTCGGTTTGGGTAAACAGACAATAACGAACATTTCAACAA
 AGTTGCAAAAGGTTGTCGACGATAATTTTGGTCGAACAAGGTGCTCAAAAGATTAGTCCAAAGTAGGT
 TTGGGTGACGATGACCAATGTATAGAAGATGACTTTACTGCCCTGGAGAGAAAGCTTTGTGGCCTG
 AATTAGACACAATCTTGAGAGAAGAAAGGTGACACCGCGTTGCTACCCCATATACTGCTGCAGT
 ATTAGAAATACAGAGTTTCCATCCATGATAGTGAAGACGCAAAAGTTTAAATGATATCACTTTGGCC
 AATGGTAACGGTTATACAGTTTTCGATGCACAACACCCCTTACAAGCTAACGTTGCAGTCAAGA
 GAGAAATTACATACACCAGAAATCCGACAGAAGTTGTATACACTTGAAATTTGATATCGCTGGTTC
 CGGTTTAAACCATGAAGTTGGGTGACCATGTAGGTGTTTTATGCGACAAATTTGTCTGAAACTGTT
 GATGAAGCAATTGAGATTGTTGGATATGTCCCTGACACTTATTTTAGTTTGCACGCTGAAAAAG
 AAGATGGTACACCAATTTCCAGTTCTTTACCACCTCCATTCCTCCATGTAACCTTAAGAACACGCCT
 TGACCAGATACCGCTTGTGTTATCATCCCCTAAAAAGTCCGCCCTTGGTTGCTTTAGCCGCTCAT
 GCTAGTGATCCTACTGAAGCAGAAAGATTGAAACACTTAGCATCTCCAGCCGGTAAAGATGAAT
 ATTCAAAGTGGGTAGTTGAATCTCAAAGATCATTTGTTAGAAGTTATGGCAGAAATTTCCATCTGCC
 AAGCCTCCATTAGGTGCTTCTTTGCTGGGTAGCACCTAGATTGCAACCAAGATTCTACTCAAT
 CAGTTCTTCACCTAAGATCGCTGAAACTAGAAATTCATGTTACATGTGCATTAGTCTACGAAAAAGA
 TGCCAAACCGGTAGAATTCACAAGGGTGTATGCTCTACTTCTGATGAAAAAATGCTGTTCCTTACGA

AAAATCAGAAAAAGTTGTTCTTAGGTAGACCACCAATCTTCGTAAGACAAATCAAACTTCAAGTTGCCCTT
 CTGATTCAAAGGTTCCAATAATCATGATAGGTCCTGGTACAGGTTTAGCCCCCATTCAGAGGTTTTC
 TTGCAAGAAAGATTGGCTTTAGTTGAATCTGGTGTGGAATTAGGTCCCTCAGTTTGTCTCTTGG
 TTGTAGAAACAGAAAGAAATGGATTTCATCTATGAAGAAGAATTGCAAGATTTCGTCGAATCTGGT
 GCATTGGCCGAATTATCTGTAGCTTTTCAAGAGAAGGTCCAACTAAGGAATACGTTCAACATA
 AGATGATGGATAAGGCATCCGACATATGGAACATGATCAGTCAAGTGCTTATTTGTACGTTTG
 CCGTGACGCAAAAGGTATGGCCAGAGATGTCATAGATCTTTGCACACAATTGCTCAAGAACA
 AGGTTCCATGGATAGTACCAAAGCTGAAGGTTTCGTAAAGAACTTACAAACTTCCGGTAGATAC
 TTGAGAGATGCTGGTGA

>Stevia rebaudiana CPR7 (CPR cloned from Stevia rebaudiana cDNA)
 ATGCAATCGGAATCCGTTGAAGCATCGACGATTGATTTGATGACTGCTGTTTGAAGGACACAG
 TGATCGATACAGCAACGCATCTGATAACGGAGACTCAAGATGCCCGCGGTTGGCGATGA
 TGTTCCGAAATTCGTGATCTGTTGCTGATTTTGACTACGTACGTGCTGTTTGGTCGGATGTTTC
 GTTGTTTGGTGTGGAAGAGATCGTCCGGGAAGAAGTCCGGCAAGGAATTGGAGCCGCCGAA
 GATCGTTGTGCCGAAGAGCGGCTGGAGCAGGAGTTGATGATGTTAAGAAGAGGTTACGA
 TTTTCTTCGGAAACACAAACTGGAAACGGCTGAAAGGTTTCGCTAAGGCACTTTTCGAAGAAGCGAA
 AGCGGATATGAAAAGGCAGCGTTTAAAGTGATTGATTGGATGATTATGCTGCTGATTGGAT
 GAGTATGCAGAGAAGCTGAAGAAGGAACATATGCTTTCTTCTTGGCTACATATGGAGATG
 GTGAGCCAACTGATAATGCTGCCAAATTTTATAAATGGTTTACTGAGGGAGACGAGAAAGGCG
 TTTGGCTTCAAAAACCTTCAATATGGAGTATTTGGTCTTGGCAACAGACAATATGAACATTTCAAC
 AAGATTGGAATAGTGGTTGATGATGGTCTCACCAGCAGGGTGCAAAACGCATTTGTTCCCGTT
 GGCTTTGGAGACGACGATCAATCAATTGAAGACGATTTTTCGGCATGGAAGAGTTAGTGTGG
 CCCGAATTGGATCTATTGCTTCGGATGAAGATGACAAAGCTGCTGCAACTCCTTACACAGCTG
 CAATCCCTGAATACCGGTCGTATTTTCATGACAAACCCGATGCGTTTTCGTGATGATCACTCAA
 ACCAATGGTCATGCTGTTCATGATGCTCAACATCCATGCAGATCCAAATGTGGCTGTAAAAAAG
 AGCTTCATACTCCTGAATCCGATCGTTTCATGCACACATCTTGAATTTGACATTTCTCACACTGGAT
 TATCTTATGAAACTGGGGATCATGTTGGTGATACCTGTGAAAACCTAAATTGAAGTAGTGAAGA
 AGCTGGGAAATTGTTAGGATTATCAACAGATACCTTATTTCTCGTTACATATTGATAACGAAGATG
 GTTACCCACTTGGTGGACCTTCAATACAACTCCTTTTCCCTCCTTGTACTTTAAGAAAAGCATTGA
 CTAATTATGCAGATCTGTTAAGCTCTCCCAAAAAGTCAACTTTGCTTGTCTCTAGCTGCTCATGCTT

CCGATCCCACTGAAGCTGATCGTTTAAAGATTCTTGCACTCTCGGAGGGCAAGGATGAATATGC
 TGAATGGGTTGTTGCAAAACCAAGAGTCTTCTTGAAGTCATGGAAGCTTCCCGTCAGCTAGA
 CCGCCACTTGGTGTCTTTTGACAGCGTTGCACCGGTTTACAGCCCTCGTTACTACTCTATTTCT
 TCCTCCCCAAAGATGGAACCAACAGGATTCACTGTTACTTGGCGTTGGTTTATGAAAAAACTCC
 CGCAGGTCGTATCCACAAAGGAATCTGCTCAACCTGGATGAAGAACGCTGTACCTTTGACCGAA
 AGTCAAGATTGCAGTTGGGCACCGATTTTGTGTTAGAACATCAAACTTCAGACTTCCAAATTGACCC
 GAAAGTCCCGGTTATCATGATTGGTCCCTGGAACCGGTTGGCTCCATTAGGGGTTTCTCTCAA
 GAAAGATTGGCTCTTAAAGAAATCCGGAACCGAACTCGGTCATCTATTTTATTTCTTCGGTTGTAG
 AAACCGCAAAGTGGATTACATATATGAGAATGAACCTCAACAACTTTGTTGAAAAATGGTGGCGTT
 TCTGAGCTTGATGTTGCTTTCTCCCGGATGGCCCGACGAAAGAAATACGTGCAACATATAAATGA
 CCCAAAAGGCTTCTGAAATATGGAATATGCTTTCTGAGGGAGCATATTTATATGTATGTGGTGA
 TGCTAAAGGCATGGCTAAAGATGTACACCGTACACTTCACACCATTTGTGCAAGAACAGGGAAGT
 TTGGACTCGTCTAAAGCGGAGTTGTATGTGAAGAAATCTACAAATGTCAGGAAGATACCTCCGTG
 ATGTTTGGTAA

> Stevia rebaudiana CPR8 (CPR cloned from Stevia rebaudiana cDNA)
 ATGCAATCTAACTCCGTGAAGATTTCGCCGCTTGATCTGGTAACCTGCGTGTTAGCGGCAAGG
 TTTTGGACACATCGAACGCATCGGAATCGGGAGAACTCTGCTATGCTGCCGACTATAGCGATGAT
 TATGGAGAAATCGTGAGCTGTTGATGATACCTCAACACGTCGGTTGCTGTATTGATCGGATGCGTT
 GTCGTTTGGTGTGGCGGAGATCGTCTACGAAGAAGTCGGCGTTGGAGCCACCCTGATTGTG
 GTTCCGAAGAGAGTGCAAGAGGAGGAAGTTGATGATGTTAAGAAAGAAAGTTACGGTTTCTTC
 GGCACCCAAACTGGAACACAGCTGAAGGCTTCGCTAAGGCCTTGTGAGGAAGCTAAAGCTCGA
 TATGAAAAGGCTGTCTTTAAAGTAATTGATTTGGATGATTATGCTGCTGATGACGATGAGTATG
 AGGAGAAACTAAAGAAAGAAATCTTTGGCCCTTTTCTTTTGGCTACGTATGGAGATGGTGAGCC
 AACAGATAATGCTGCCAGATTTTATAAATGGTTTACTGAGGGAGATGCCAAAGGAGAAATGGCT
 TAATAAGCTTCAATATGGAGTATTGGTTTGGGTAACAGACAAATATGAACATTTTAACAAGATC
 GCAAAAGTGGTTGATGATGGTCTTGTAGAACAGGGTGCAAAAGCGTCTTGTCTCTGTTGGACTTG
 GAGATGATGATCAATGTATTGAAGATGACTTCACCGCATGGAAGAGTTAGTATGGCCGGAGT
 TGGATCAATTACTTTCGTGATGAGGATGACACAACTGTTGCTACTCCATACACAGCTGCTGTGCA
 GAATATCGCGTTGTTTTCATGAAAAACCAGACGCGCTTTCTGAAGATTATAGTTATACAAATGG

CCATGCTGTTCAATGCTCAACATCCATGCAGATCCAAACGTGGCTGTCAAAAAGGAACITTCAT
AGTCTGAAATCTGACCGGTCTTGCACTCATCTTGAATTTGACATCTCGAACACCGGACTATCATA
TGAAACTGGGACCATGTTGGAGTTTACTGTGAAAACCTTGAGTGAAGTTGTGAATGATGCTGA
AAGATTAGTAGGATTACCAACAGACACTTACTCCTCCATCCACACTGATAGTGAAGACGGGTG
CCACTTGGCGGAGCTCATTTGCCGCTCCTTTCCCGCCATGCACTTTAAGGAAAGCATTGACGTG
TTATGCTGATGTTTGAAGTTCTCCAAAGAGTCGGCTTGTCTGCACCTAGCTCATGCCCACCG
ATCCCAGTGAAGCTGATAGATTGAAATTTCTTGCATCCCGCCGAAAGGATGAATATTCTCA
ATGGATAGTTGCAAGCCAAAGAGTCTCCTTGAAGTCATGGAAGCATTCCCCGCAGCTAAGCCT
TCACCTTGGTGTCTTCTTTGCATCTGTTGCCCGCGCTTACAACCAAGATACTACTCTATTTCTTCCT
CACCCAAGATGGCACCGGATAGGATTCAATGTACATGTGCATTAGTCTATGAGAAAACACCTGC
AGGCCGATCCACAAGGAGTTTGTTCAACTTGGATGAAGAACGCAGTGCCTATGACCGAGAG
TCAAGATTGCAGTTGGGCCCCAATATACGTCCGAACATCCAAATTCAGACTACCATCTGACCCCTA
AGTCCCCGGTTATCATGATTGGACCTGGCACTGGTTTGGCTCCTTTTAGAGGTTTCCCTTCAAGAG
CGTTAGCTTTAAAGGAAGCCGGAACCTGACCTCGGTTTATCCATTTTATCTTCGGATGTAGGAA
TCGCAAGTGGATTTCATATATGAAAACGAGCTTAACAACTTTGTGGAGACTGGTGTCTTTCTG
AGCTTATTGTTGCTTCTCCCGTGAAGCCCCGACTAAGGAATATGTGCAACACACAAGATGAGTGA
GAAGGCTTCGGATATCTGGAACTTGCTTTCTGAAGGAGCATATTTATACGTATGTGGTGTGCC
AAAGGCATGGCCAAAGATGTACATCGAACCTCCACACAATTTGTCAAGAACACAGGGATCTCTTG
ACTCGTCAAAGGCAGAACTCTACGTGAAGAAATCTACAAATGTCAGGAAGATACCTCCGTGACGT
TTGGTAA

FIGURE 14

A.	atggttttgt	cttctttcttg	tactacagta	ccacacttat	cttcattagc	tgctgtgcaa	60
	cttggctcctt	ggagcagtag	gattaaaaag	aaaaccgata	ctgttgagc	accagccgct	120
	<u>gcaggaaggt</u>	<u>ggagaagggc</u>	<u>cttggctaga</u>	gcacagcaca	catcagaatc	cgcagctgtc	180
	gcaaagggca	gcagtttgac	ccctatagt	agaactgacg	ctgagtcaag	gagaacaaga	240
	tggccaacccg	atgacgatga	cgcggaacct	ttagtggatg	agatcagggc	aatgcttact	300
	tocatgtctg	atggtgacat	ttccgtgagc	gcatacagata	cagcctgggt	cggattgggt	360
	ccaagattag	acggcgggtga	aggtccctcaa	tttccagcag	ctgtgagatg	gataaagaaat	420
	aaccagttgc	ctgacggaag	ttggggcgat	gccgcattat	tctctgccta	tgacaggctt	480
	atcaataccc	ttgcctgcgt	tgtaactttg	acaagggtggt	ccttagaacc	agagatgaga	540
	gtagaggac	tatctttttt	gggtagggaac	atgtggaaat	tagcaactga	agatgaagag	600
	tcaatgccta	ttggcttcga	attagcattt	ccatctttga	tagagcttgc	taagagccta	660
	ggtgtccatg	acttccctta	tgatcaccag	gccctacaag	gaatctactc	ttcaagagag	720
	atcaaaaatga	agaggattcc	aaaagaagtg	atgcataccg	ttccaaacatc	aatattgcac	780
	agtttggagg	gtatgcctgg	cctagattgg	gctaaaactac	ttaaaactaca	gagcagcgac	840
	ggaagttttt	tgttctcacc	agctgccact	gcataatgctt	taatgaatac	cggagatgac	900
	aggtgtttta	gctacatcga	tagaacagta	aagaaaattca	acggcggcgt	ccctaattgtt	960
	tatccagtg	atctatttga	acatatattg	gccgttgata	gacttgaaag	attaggaatc	1020
	tccaggtaact	tccaaaagga	gacggaacaa	tgcatggatt	atgtaaacag	gcattggact	1080
	gaggacggta	tttgttgggc	aaggaactct	gatgtcaaa	aggtggacga	cacagctatg	1140
	gccttttagac	ttcttagggt	gcacggctac	agcgtcagtc	ctgatgtgtt	taaaaacttc	1200
	gaaaaggacg	gtgaattttt	cgcatttgtc	ggacagtcta	atcaagctgt	taccgggatg	1260
	tacaacttaa	acagagcaag	ccagatatcc	ttcccaggcg	aggatgtgct	tcatagagct	1320
	ggtgccttct	catatgagtt	cttgaggaga	aaagaagcag	agggagcttt	gagggacaag	1380
	tggatcattt	ctaaagatct	acctggtgaa	gttgtgtata	ctttggattt	tccatgggtac	1440
	ggcaacttac	ctagagtoga	ggccagagac	tacctagagc	aatacggagg	tggtgatgac	1500
	gtttggattg	gcaagacatt	gtataggatg	ccacttgttaa	acaatgatgt	atatattggaa	1560
	ttggcaagaa	tggatttcaa	ccactgccag	gctttgcatc	agttagagtg	gcaaggacta	1620
	aaaagatggt	atactgaaaa	taggttgatg	gaatttgggtg	tgcgcccaaga	agatgccctt	1680

agagcttatt ttcttgcagc cgcattctgtt tacgagcctt gtagagctgc cgagaggctt 1740
 gcatgggcta gagccgcaat actagctaac gccgtgagca ccacttaag aaatagccca 1800
 tcattcagag aaagggttaga gcatttctctt aggtgtagac ctagtgaaga gacagatggc 1860
 tcctgggtta actcctcaag tggctctgat gcagtttttag taaaggctgt cttaagactt 1920
 actgattcat tagccaggga agcacagcca atccatggag gtgacccaga agatattata 1980
 cacaagtgt taagatctgc ttgggccgag tgggttaggg aaaaggcaga cgctgccgat 2040
 agcgtgtgca atggtagtct tgcagtagaa caagagggat caagaatggt ccatgataaa 2100
 cagacctgtc tattattggc tagaatgac gaaatttctg ccggtagggc agctggtgaa 2160
 gcagccagtg aggacggcga tagaagaata attcaattaa caggctccat ctgcgacagt 2220
 cttaagcaaa aaatgctagt ttcacaggac cctgaaaaaa atgaagagat gatgtctcac 2280
 gtggatgacg aattgaagtt gaggattaga gagttcgttc aatatctgct tagactaggt 2340
 gaaaaaaaaga ctggatctag cgaaccagg caaacatttt taaagtatagt gaaatcatgt 2400
 tactatgctg ctcattgccc acotcatgtc gttgatagac acattagtag agtgattttc 2460
 gagccagtaa gtgccgcaaa gtaaccggcg 2490

B.

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

115	120	125
Pro Gln Phe Pro Ala Ala Val Arg Trp Ile Arg Asn Asn Gln Leu Pro		
130	135	140
Asp Gly Ser Trp Gly Asp Ala Ala Leu Phe Ser Ala Tyr Asp Arg Leu		
145	150	155
Ile Asn Thr Leu Ala Cys Val Val Thr Leu Thr Arg Trp Ser Leu Glu		
	165	170
Pro Glu Met Arg Gly Arg Gly Leu Ser Phe Leu Gly Arg Asn Met Trp		
	180	185
Lys Leu Ala Thr Glu Asp Glu Glu Ser Met Pro Ile Gly Phe Glu Leu		
	195	200
Ala Phe Pro Ser Leu Ile Glu Leu Ala Lys Ser Leu Gly Val His Asp		
210	215	220
Phe Pro Tyr Asp His Gln Ala Leu Gln Gly Ile Tyr Ser Ser Arg Glu		
225	230	235
Ile Lys Met Lys Arg Ile Pro Lys Glu Val Met His Thr Val Pro Thr		
	245	250
Ser Ile Leu His Ser Leu Glu Gly Met Pro Gly Leu Asp Trp Ala Lys		
	260	265
Leu Leu Lys Leu Gln Ser Ser Asp Gly Ser Phe Leu Phe Ser Pro Ala		
275	280	285
Ala Thr Ala Tyr Ala Leu Met Asn Thr Gly Asp Asp Arg Cys Phe Ser		
290	295	300
Tyr Ile Asp Arg Thr Val Lys Lys Phe Asn Gly Gly Val Pro Asn Val		
305	310	315
Tyr Pro Val Asp Leu Phe Glu His Ile Trp Ala Val Asp Arg Leu Glu		
	325	330
Arg Leu Gly Ile Ser Arg Tyr Phe Gln Lys Glu Ile Glu Gln Cys Met		
	340	345
Asp Tyr Val Asn Arg His Trp Thr Glu Asp Gly Ile Cys Trp Ala Arg		
355	360	365

Asn Ser Asp Val Lys Glu Val Asp Asp Thr Ala Met Ala Phe Arg Leu
 370
 Leu Arg Leu His Gly Tyr Ser Val Ser Pro Asp Val Phe Lys Asn Phe
 385
 Glu Lys Asp Gly Glu Phe Phe Ala Phe Val Gly Gln Ser Asn Gln Ala
 405
 Val Thr Gly Met Tyr Asn Leu Asn Arg Ala Ser Gln Ile Ser Phe Pro
 420
 Gly Glu Asp Val Leu His Arg Ala Gly Ala Phe Ser Tyr Glu Phe Leu
 435
 Arg Arg Lys Glu Ala Glu Gly Ala Leu Arg Asp Lys Trp Ile Ile Ser
 450
 Lys Asp Leu Pro Gly Glu Val Val Tyr Thr Leu Asp Phe Pro Trp Tyr
 465
 Gly Asn Leu Pro Arg Val Glu Ala Arg Asp Tyr Leu Glu Gln Tyr Gly
 485
 Gly Gly Asp Val Trp Ile Gly Lys Thr Leu Tyr Arg Met Pro Leu
 500
 Val Asn Asn Asp Val Tyr Leu Glu Leu Ala Arg Met Asp Phe Asn His
 515
 Cys Gln Ala Leu His Gln Leu Glu Trp Gln Gly Leu Lys Arg Trp Tyr
 530
 Thr Glu Asn Arg Leu Met Asp Phe Gly Val Ala Gln Glu Asp Ala Leu
 545
 Arg Ala Tyr Phe Leu Ala Ala Ala Ser Val Tyr Glu Pro Cys Arg Ala
 565
 Ala Glu Arg Leu Ala Trp Ala Arg Ala Ala Ile Leu Ala Asn Ala Val
 580
 Ser Thr His Leu Arg Asn Ser Pro Ser Phe Arg Glu Arg Leu Glu His
 595

Ser	Leu	Arg	Cys	Arg	Pro	Ser	Glu	Glu	Thr	Asp	Gly	Ser	Trp	Phe	Asn
610						615					620				
Ser	Ser	Ser	Gly	Ser	Asp	Ala	Val	Leu	Val	Lys	Ala	Val	Leu	Arg	Leu
625					630					635					640
Thr	Asp	Ser	Leu	Ala	Arg	Glu	Ala	Gln	Pro	Ile	His	Gly	Gly	Asp	Pro
				645					650					655	
Glu	Asp	Ile	Ile	His	Lys	Leu	Leu	Arg	Ser	Ala	Trp	Ala	Glu	Trp	Val
								665					670		
Arg	Glu	Lys	Ala	Asp	Ala	Ala	Asp	Ser	Val	Cys	Asn	Gly	Ser	Ser	Ala
		675					680					685			
Val	Glu	Gln	Glu	Gly	Ser	Arg	Met	Val	His	Asp	Lys	Gln	Thr	Cys	Leu
						695					700				
Leu	Leu	Ala	Arg	Met	Ile	Glu	Ile	Ser	Ala	Gly	Arg	Ala	Ala	Gly	Glu
705					710					715					720
Ala	Ala	Ser	Glu	Asp	Gly	Asp	Arg	Arg	Ile	Ile	Gln	Leu	Thr	Gly	Ser
				725					730					735	
Ile	Cys	Asp	Ser	Leu	Lys	Gln	Lys	Met	Leu	Val	Ser	Gln	Asp	Pro	Glu
								745					750		
Lys	Asn	Glu	Glu	Met	Met	Ser	His	Val	Asp	Asp	Glu	Leu	Lys	Leu	Arg
		755					760					765			
Ile	Arg	Glu	Phe	Val	Gln	Tyr	Leu	Leu	Arg	Leu	Gly	Glu	Lys	Lys	Thr
						775					780				
Gly	Ser	Ser	Glu	Thr	Arg	Gln	Thr	Phe	Leu	Ser	Ile	Val	Lys	Ser	Cys
785					790					795					800
Tyr	Tyr	Ala	Ala	His	Cys	Pro	Pro	His	Val	Val	Asp	Arg	His	Ile	Ser
				805					810						815
Arg	Val	Ile	Phe	Glu	Pro	Val	Ser	Ala	Ala	Lys					
								825							

FIGURE 15

A.
 ATGCCGTGTAATAATTGAAATGTGTACCCCAAGGACCTCAAGACTGGAAATGATTTTGTCTGCTGCTAAGAGTTTACTAGATCGAGC
 TTTCAAAAAGTCATCATTCCTACTACGGATTATGCTCAACTTCATGTCAAGTTTATGATACAGCTTGGGTTGCAATGATTCCAAAAACAA
 GAGATAATGTAAAACAGTGGTTGTTCCAGAAATGTTTCCATTACCTCTTAAAAACACAGCCGAGATGGCTCATGGGTTTCATTGGCCT
 ACAACACAGACAGCGGGTATCCTAGATACAGCCTCAGCTGTGCTGGCATTATTGTGCCACGCACAAAGAGCCTTTACAAAATATTGGATGT
 ATCTCCAGATGAAATGGGGTTGAGAAATAGAAACACGGTGTACATCCTTGAAACGTCAATTAGCAGTTTGGAAATGATGTGGAGGACACCA
 ACCATATTGGCGTCGAGTTTATCATACCAAGCCTTACTTTCCATGCTAGAAAAAGGAAATTAGATGTTCCATCTTTTGAATTTCCATGTAGG
 TCCATCTTAGAGAGAAATGCACGGGAGAAATTAGGTCAATTCGACCTGGAACAAGTTTACGGCAAGCCAAGCTCATTTGTTGCACTCAT
 GGAAGCATTTCTCGGTAAGCTAGATTTTGATCGACTATCACATCACTATACCAACGGCAGTATGATGGCATCTCCATCTTCAACGGCTG
 CTTATCTTATTGGGGCTACAAAATGGGATGACGAAGCCGAAGATTACCTAAGACATGTAATGCCGTAATGGTGCAGGACATGGGAATGGA
 GGATTTCTGGTACATTTCCAACACTACTCATTTTCGAATGTAGCTGGATTATAGCAACGTTGTTAAAGGTTGGCTTTACTTTGAAGCAAAAT
 TGACGGCGATGGCTTAAGAGGTTTATCAACCATCTTACTTGAGCGCTTTCGTGATGAGAATGGTGCATAGGCTTTGCCCCCTAGAACAG
 CAGATGTAGATGACACAGCCAAAGCTCTATTGGCCCTTGTCAATTGGTAAACAGCCAGTGTCACCTGATATCATGATTAAGGCTTTTGAG
 GGCAAGACCAATTTTACCACCTTTTGGTTCAGAAAGAGATCCATCATTGACTTCCAACCTGCACGTCCTTTTATCTTTACTTAAACAATC
 TAACCTGTCTCAATACCATCCCTCAAAATCCTCAAAACCAACATTAATCACTTGTAGATGGTGGGTTCCGATCATTTGTTGTCAAAGACA
 AATGGAATTTGAGTCACCTATATCCAACTATGTTGTTGTTGAAGCCTTCACGTGAAGTGTCCATCTCATTTGACGGTGGTGAATTTGTCT
 AGTCTGTTTGATGAATCCCTTTAAGTGAAGATTGGTCTTAGCATCTTTCAAGCGGTACTTAGAATAATCCTCACCCCAAGACAACGACGG
 CTCCTGGAGAGGATACAGAGAACAGACGCTGTTACGCAATATTGGCTTTAGTTCAAGCGAGACATGTATGCTTTTTCACATCACATGGTTG
 ACAGACTGCAATCATGTGTTGATCGAGGTTTCTCATGTTGAAATCTTGTCTTTTTCATTCTCAAGACCTGACTTGGACCTCTAAAAACA
 GCTTATGAAGTGGGTTTCGTAGCTGAAGCATATAAACTAGCTGCTTTACAATCTGCTTCCCTGGAGGTTCCCTGCTGCCACCATTTGGACA
 TTCGTACAGCTCGCCGTTCCATCAAGTGATCTTGAAAAAATACATGAGATTGGTGAGAAAAAACTGCGTTATTCTCTCCACTGGATGAGT
 GGGTCTAATGGCTTCTATCATCGAATCTTCAATTTTTCGTACCATTAATGTCAGGCACAAAGAGTTGAAATATACCCCTAGAGATAATATC
 AAGGTGGACGAAGATAAGTACTTGTCATTATTATCCCATTCACATGGGTGGATGCAATAATAGGTTAGAACCTTCGCAAGTAACAGATG
 GCTATACGATATGATGTAACCTTTTCACTCGGCTATCAAAACGACGAGTACATGGAAGCTGTAGCTGGGCCAGTGTGTTGGGATGTTT
 CCTTGTTACATCAACAATTGATAAGGTGATTGATAATAACAATGGGTAACTTGGAGAGCCCAATGGAACAGTACACAGTGGTAATGGA
 CATCAGCACGAATCTCCTAATATAGGTCAAGTCAGGACACCTTGACTCGTTTTCACAAAATTCAGTCTTGAATCACAAGACGTCCTTAA
 CTCTAGCTCATCTGATCAAGATACCTTTGAGAAAGAGAGTTTAGAACATTCATGACCGCTCATATAACAAAAATCGAAGATAACTCACGAT
 TCAGTAAGCAAGCCTCATCCGATCGGTTTTCCTCTCTCCGTGAACAACTTACTTTCAATGGGTGAACCTCAACTGGTGGCTCACATGTCCGT
 TGGCCCTATTCAATTTGCCCTTCTCTAATTGCCCTCATGTCTGCAAAATTTGTTGCAGGGTAAAGACGCTATTCCAAAGCGGAACGCAAAAGTA

CTTAATCTCCTCTGTTATGAGACATGCCACAAACATGTGTAGAAATGTATAACGACTTTGGCTCTATTGCCAGAGACAAACGCTGAGAGAA
 ATGTTAATAGTATTTCCTGAGTTTACTCTCTGTACGGAACCTTCTCAAAACCTAGATGAAAGGAAGAAAGACTTCTGAAAAATC
 GCAACTTACGAAACAAGGTATTGATAGAGCACTAGAGGCCTTGGAAAGACAGAGTAGAGATGATGCCGGAGACAGAGCTGGATCTAA
 AGATATGAGAAAAGTTGAAAATCGTTAAGTTATTCTGTGATGTACGGACTTATACGATCAGCTCTACGTTATCAAGATTTTGTCTCCT
 CTATGAAGTAA

B.

MPGKIENGTPKDLKTGNDFVSAAKSLLDRAFKSHSYGLCSTSCQVYDTAWVAMIPKTR
 DNVKQWLFPECFHYLLKTQAADGSWGLPPTQTAGILDTASAVLALLCHAQEPQLIILDVS
 PDEMGLRIEHGVTSLKRQLAVWNVEDTNHIGVEFIIPALLSMLEKELDVPSFEFFCRSI
 LERMHGEKLGHFLEQVYKPSLLHSLEAFLGKLDLDFDRLSHHLHYHGSMMASPSSTAAYL
 IGATKWDDEAEDYLRHVMRNGAGHNGGISGTFPTTHFECSWIIATLLKVGFLLKQIDGD
 GLRGLSTILLEALRDENGVI GFAPRTADVDDTAKALLALSLVNQPVSPDIMIKVFEGKDH
 FTTFGSERDPSLTSNLHVLSSLKQSNLSQYHPQILKTTFTCRWWGSDHCVKDKWNLS
 HLYPTMLLVEAFTEVLHLIDGGELESLFDESFKCKIGLSIFQAVLRIILTQDNDGSRGY
 REQTCYAILALVQARHVCFFTHMVDRLQSCVDRGFSWLKSCSFHSQDLTWTSTAYEVGF
 VAEAYKLAALQASLEVPAAATIGHSVTS AVPSSDLEKYMRLVRKTALEPLDEWGLMASI
 IESSFFVPLLQAQRVEIYPRDNIKVDEDKYLSIIPFTWVGCNNRSRTFASNRWLYDMMYL
 SLLGYQTDEYMEAVAGPVFGDVSLHQTIDKVIDNTMGNLARANGTVHSGNGHQHESPNI
 GQVEDTLTRFTNSVLNKKDVLNSSSDQDTLRRREFRTFMHAHITQIEDNSRFSKQASSDA
 FSSPEQSYFQWVNSTGGSHVACAYSFAFNSCLMSANLLQKDAFPSTQKYLISSVMRHA
 TNMCRMYNDFGSIARDNAERNVNSIHFEFTLCNGTSQNLDERKERLLKIATYEQGYLDR
 ALEALERQSRDDAGDRAGSKDMRKLKIVKLCFCDVTDLYDQLYVIKDLSSSMK

FIGURE 16

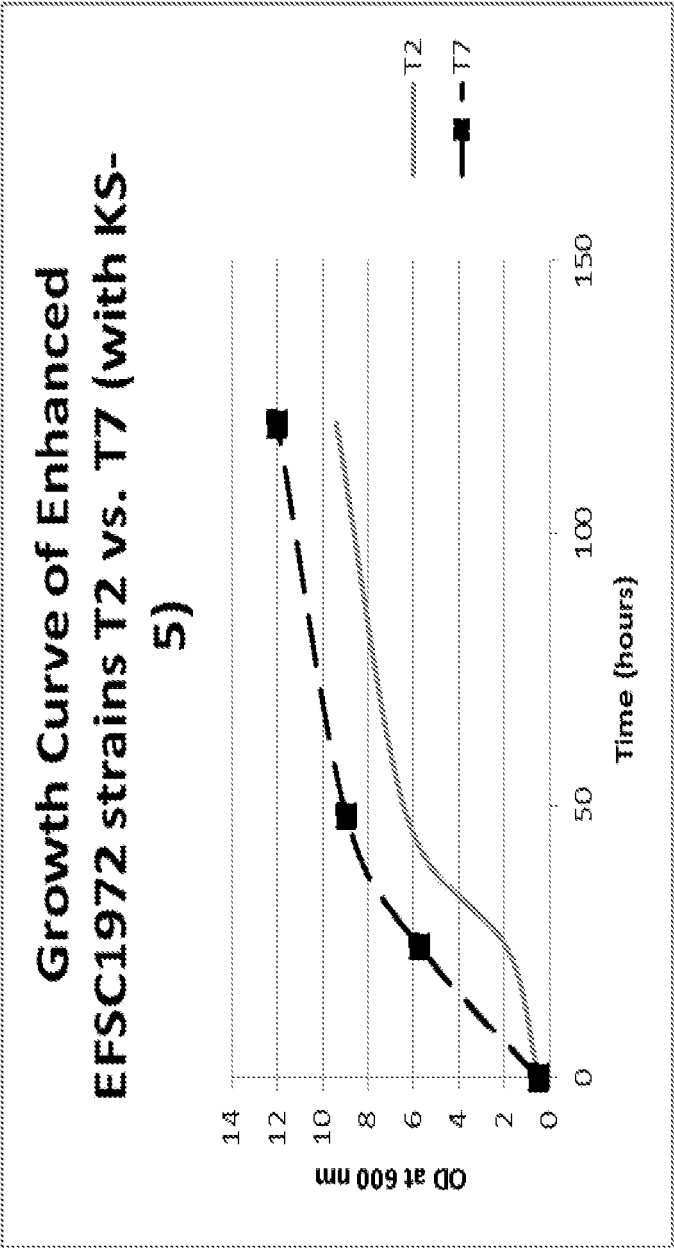


FIGURE 17

A. > Native AtSUS1
ATGGCAAAACGCTGAACGTATGATAACCGCGGTCCACAGCCAAACGTGAGCGTTTGAACGAAACGCTTGTTCTGAGAGAAACGAAAGTCCT
 TGCCTTGCTTTCCAGGGTTGAAGCCAAAGGTAAAGGTATTTACAACAAACAGATCATTTGCTGAATTCGAAAGCTTTGCCTGAACAAA
 CCGGAAGAAAACCTGAAGGTGGTCTTTCTTTGACCTTCTCAAATCCACTCAGGAAGCAATTTGTGTGCTGTTGCCACCATGGGTTGCTCTAGCT
 GTGAGCCCAAGGCTGGTGTGGGAATACTTACGAGTCAATCTCCATGCTCTTGTCTGTTGAAGAACTCCAACCTGCTGAGTTTCTTCA
 TTCAAGGAAGAACTCGTTGATGGAGTTAAGAAATGGTAATTTCACTCTTGAGCTTGATTTTCGAGCCATTCAATGCGTCTATCCCTCGTC
 CAACACTCCACAAATACATTTGGAATGGTGTGACTTCTTAACCGTCAATTTATCGGCTAAGCTCTTCCATGACAAGGAGAGTTTGCTT
 CCATTGCTTAAGTTCCCTTCGTCTTCACAGCCACCAGGGCAAGAACCTGATGTTGAGCGAGAAGATTCAGAACTTCAACACTCTGCAACA
 CACCTTGAGGAAAGCAGAAAGATCTAGCAGAGCTTAAGTCCGAAACACTGTATGAAGAGTTTGAGGCCAAGTTTGAGGAGATTGGTC
 TTGAGAGGGATGGGAGACAAATGCAGAGCGTGTCCCTTGACATGATACGTCTCTTTTGGACCTTCTTGAGGCGCTGATCCTTGCACT
 CTTGAGACTTTTCTTGAAGAGTACCAATGGTGTCAACGTTGTGATCCCTCTCCACATGGTTACTTTGCTCAGGACAATGTTCTTGG
 TTACCTTGACACTGGTGGACAGGTGTTTACATTTCTTGATCAAGTTCGTGCTCTGGAGATAGAGATGCTTCAACGTATTAAGCAACAAG
 GACTCAACATTAACCAAGGATTTCTCAATCTCACTCGACTTCTACCTGATGCGGTAGGAACATACTGCGGTGAACGCTCTCGAGAGAGTT
 TATGATTTCTGAGTACTGTGATATTCTTCGTGTGCCCTTCAGAAACAGAGAAGGATTTGTTCCGCAAAATGGATCTCAAGGTTCCGAGTCTG
 GCCATATCTAGAGACTTACACCGAGGATGCTGCGGTTGAGCTATCGAAAGAAATGAATGGCAAGCCTGACCTTATCATTTGGTAACTACA
 GTGATGGAAATCTTGTGCTTTTATTGGCTCACAAACTTGGTGTCACTCAGTGTAACCATTTGCTCATGCTCTTGAGAAACAAAGTAC
 CCGGATTTCTGATATCTACTGGAAAGAGCTTGACGACAAAGTACCATTTCTCATGCCAGTTCACCTGCGGATATTTTCGCAATGAACCCAC
 TGATTTCAATCATCACTAGTACTTTCCAAGAAATTGCTGGAAGCAAGAAACTGTTGGGCAGTATGAAAGCCACACAGCCTTTACTCTTC
 CCGGATTGTATCGAGTTGTTACGGGATTGATGTGTTTGTATCCCAAGTTCAACATTTGTCTCTCCTGGTGTGATATGAGCATCTACTTC
 CCTTACACAGAGAGAAAGCGTAGATTGACTAAGTTCCACTCTGAGATCGAGGAGCTCCTCTACAGCGATGTTGAGAACAAAGAGCACTT
 ATGTGTGCTCAAGGACAAAGAACGCCGATTCTCTTCACAATGGCTAGGCTTGATCGTGTCAAGAACTTGTCAAGGCTTGTGAGTGGT
 ACGGGAAGAACACCCGCTTGGGTGAGCTAGCTAACTTGGTGTGTTGGAGGAGACAGGAGGAAAGAGTCAAAAGGACAATGAAGAGAAA
 GCAGAGATGAAGAAAATGTATGATCTCATTGAGGAATACAAGCTAAACGGTCAGTTCAAGGTGATCTCCTCTCAGATGGACCGGGTAAG
 GAACGGTGAGCTGTACCGGTACATCTGTGACACCAAGGTGCTTTTGTCCAACTGCAATATATGAAGCCTTTGGGTTAACTGTTGTGG
 AGGCTATGACTTGTGGTTTACCGACTTTTCGCCACTTGCAAGGTGGTCCAGCTGAGATCATTTGTGCACGGTAAATCGGGTTTCCACATT
 GACCTTACCATGGTGATCAGGCTGCTGATACCTCTTGCTGATTTCTTCAACCAAGTGAAGGAGATCCATCTCACTGGGATGAGATCTC
 AAAAGGAGGGCTTCAGAGGATTGAGGAGAAAATACACTTGGCAAAATCTATTACAGAGGCTCTTGACATTGACTGGTGTGATGATTTCT

GGAAGCATGCTCGAAACCTTGACCGTCTTGAGGCTCGCCGTTACCTTGAAATGTTCTATGCATTGAAGTATCGCCCATTTGGCTCAGGCT
 GTTCCTCTTGCACAAGATGATTGA

B.

> SrSUS contig10573 selection_ORF_S11E

ATGGCGGAACGTGTACTCACTCGTGTTCAC**gaa**CTTCGTGAGCGTCTCGATTCAACTCTCGCAACTCATCGTAATGAAATCCTCTTGTT
 TCTTTCAGGATTGAAAAGCCATGGAAGGAATATTGAAGCCTCATCAAGTTATGACTGAATTTGAAGCTATCTGCAAGAAAGATCAGA
 GCAAACTCTCTGATGGTCTTTTATGAAGTTCTTAAATGCACAGGAAGCAATAGTGCAACCTCCATGGTTGCACCTCGCATCCGT
 CTTGACCCGGTGTGGGAATATGTTAGAGTCAATGTTAATGTTTGGTGGTTGAAGAAATTAAGTGTTCCTGAATATCTTCACTTCAA
 AGAAGAAATTGGTTAATGGAACATCGAATGGCAACTTCGTGTTGGAACCTTTTACCGCATCGTTTCCTCGACCAACTT
 TAACCAAGTCTATTGGTAATGGTGTGAGTTTCTAAACAGACATTTATCTGCTAAAATGTTTCATGATAAGGATAGCATGCACCCTCTT
 CTTGATTTCCCTACGGACTCACCACATATAAGGGAAGACAAATGATGTTGAATGATAGAAATCCAAACCTCAATGCTCTACAATCGGTGTT
 GCGAAAGCGCTCAGAGTACTTATCAACACTCGACCGGCAACACCGTACTCTCAGTTTGAACATAAGTTTCAAGAAATCGGGTTGGAGA
 GAGGTTGGGTGATAAAAGCGGAGGTCGTAATGGAGATGATCCACATGCTTCTAGACCTTCTAGAAGCACCCGACGCATGCACACTCGAG
 AAGTTTCTCGGAAGAAATCCCAATGGTTTCAATGTTGTCTATCTTTTCGCCCTCATGGCTACTTCGCGCAAGAAAATGTTGGGATATCC
 CGACACTGGCGGTGAGGTTGTTTACATCTTGGATCAAGTTCCCGCTCTGGAACGCGAGATGCTCAAAAGGATTAAGGAGCAAGACTCG
 ATATCATCTCCTCGTATATTGATTGTTACGAGGCTTCTTCTGACGCGGTTGGGACCACATCGGGCAACGTTTAGAGAAAGTGTGTTGGA
 GCCGAACACTCGCATATTCTTCGGTCCCCTTTAGAACCAGAAAGGTATTCTTCGTAATGGATCTCTCGTTTGGAGTGTGSCCTTA
 CATCGAGACITTCACCGAGGATGTTGCTAAAGAAAGTTACAGCAGAGTTGCAAGCAAAACAGATTTGATCATTTGGAACATAAGTGAAG
 GAAATTTGGTTGCATCTTTGCTAGCTCACAAAGTTGGGTGTCACAGTGTAACCATTGCTCATGCTTTGGAGAAACATAAATACCCGGAT
 TCTGATATCTACTGGAAGAACTTTGAGGAGAAATATCATTTCTTTCGCAGTTTACCGCTGATCTTATCGCTATGAACCATACCGACTT
 CATCATCACAGTACTTCCAAAGAAATGCTGGAAGTAAGGACACGCTTGGACAGTACGAGAGTCATACCGCTTCACAATGCCGGGAT
 TGATTCGGGTGGTTCACGGGATCGATGTTTTTGACCCCAATTCATAATTTGTTTACCCGGGCGGATATGGGAATTTACTACTCGTAT
 ACCGAGAAAGAAAAGAGGCTCACTGCGCTTCACCTGAAATCGATGAACTTCTTTAGTTCCGTCGAAAACGAACACTTATGTGT
 GTTGAAGGATAAGAGTAAACCAATCTTGTTCACAATGCGCGGATTGGATAATGTGAAGAATTTAACCGGACTGGTTGAATGGTAGCCTA
 AAAACGACCGCTTCGTGAGCTCGTGAACCTCGTGGTCTGCGTGGTGACCGAAGGAAGAGTCGAAAGATCTTGAAGAGCAAGCTCAG
 ATGAAGAAGATGCATGAACCTTATCGAAACCTACAAACTCAACGGTCAGTTCAGGTGGATATCCTCACAAATGAACCGGTGAGGAACGG
 TGAGTTGTATCGCGTTATTGCTGACACGCGAGGTGCGTTTATCCAGCCTGCGTTTACGAGGCGTTTGGGTGACGGTTGTGAGGCGCA
 TGACTTGTGGCTGCCGACATTCGCGACACTTCATGGTGGCCCGCTGAGATTATTGTTTACCGGGAATCCGGGTTCATATTGACCCG
 TATCACGGTGACCAAGTCAACCGAGTTGCTGGTCAATTTCTTTGAGAAAACATAACAGACCCGGGTCAATGGGAGGCCATTTCCAAGG
 TGGTCTGCAACGTATTACGAGAAATACACGTGGCAGATTATTTCAGATAGGTTGTTGACGCTTGCCGGAGTTTATGATTTCTGGAAGC

ATGTGTCGAAGCTTGACAGGCTCGAGATCCGTCGTTATCTTGAAATGTTTTACGCGCTCAAGTATCGCAAACTGGCTGAATCTGTTCCCA
TTGGCTGTTGATGAGTGA

C.

> Coffea arabica sus1 gene

ATGGCCGAACGTGTTCTGACCCGTTGTTACAGCCCTCCGGAAACGCCCTTGATGCTACTTTGGCTGCCCAACGCCAACGATGTTTTGCTGTT
 TATGTCGAGGCTTGAAACCCCATGGGAAAGGATCCTGAACCCACCAACTTTTGGCTGAGTTTGAAAGAAATTAACAAAGGATGGTAAAC
 AAAAAATTCATGATCATGCCCTTTGAAGAAGTCTGAAGTCCACACAGGAAGCAATTTGTGTGCCCCCTGGGTGACACTTGCTATTCTGT
 CTCAGACCTGGTGTCTGGGAGTATGTTTCGAGTCAATGTCCATGCACTCGTTGTGAGGAGTTAACCGTGCCAGAGTACCTGCATTTTCAA
 GGAAGAACTCGTTGATGGAAGCAAAAAATGGGAAATTTGTTTGGAACTGGACTTCGAACCATTTACAGCCTCTTTTCCCAAGCCAACTC
 TAACTAAGTACATAGCGACGGAGTTGAGTTCTCAACAGGCACCTCTCTGCCAAAAATGTTCCATGACAAGGAGAGCATGGCCCTCTC
 CTTGATTTTCTCCGTGTTACCAATAACAAGGCAAGACAATGATGCTTAACGACAGGATCAAGGACCTTAACACTCTCCAAGCAGTTCT
 GAGGAAGGCAGAGGAGTACCTAACAACTCTCTGCAGATACACCATACTCTGAATTCGAGCACAAAATCCAAAGAAATTGGACTGGAGA
 GAGGTTGGGTGATACTGCTGAGCGTGTCTTGAAATGATCTGCATGCTTCTGGGCTCTCTGACTCGTGACACACTAGAG
 AAATTCCTAGGGAGAAATCCCTATGGTATTCAATGTTGTTATTTCTTTCCCCCATGGATACTTTGCCCAAGGAAAAACGTTATGGGTTATCC
 TGATACCGGTGGCCAGGTTGTTTACATATTGGATCAAGTTCTTGCTTGAGCGTGAGATGCTGAAGAGGATAAAGGAACAAGGACTTG
 ATGTCAAAGCCACGCATTCTAATTATAACTAGGCTGCTTCTGTGATGCCCTGGAAACCACTTGTGTCAAACGGCTTGAGAAAGTATATGGA
 TCAGAGTACTCCCATATACCTCAGAGTCCCTTCAGAACTGAGAAGGAGTTGTTGCAAAATGGATCTCTCGCTTTGAAAGTTTGGCCCTA
 CATGGAACAATTTACTGAGGATGTTGCAAAAGAAAGTCACTGCAGAAATTACAGGCAAAAGCCAGATTTGGTTATTGGTAACTACAGTGAGG
 GTAACTTGTGCTTGTGCTTGTCTCACAAGTTAGGTGTAACACAGTGTACCATTGCTCATGCTTTGGAAAAAACCAAGTATCCTGAT
 TCTGATATTATTGAGCAAAATTGATGAAAAAGTACCACTTCTCATGTCAAGTTCAGTGGGATCTTATCGCAATGAACCATACAGATT
 CATTATCACTAGCACTTCCAAAGAAATAGCTGGAAGCAAGGACACTGTGGGCAATATGAAAGCCATATGGCCCTTCACAAATGCCAGGAT
 TATACAGAGTTGTACATGGCATTGATGTTTTTGTATCCAAATTCAAACATTGTCTCACCCTGGAGCTGATACAAACCTCTACTTCCCACAC
 ACAGAGAAAGGAAAAGAGATTGACATCTTTCCATCCTGAAATTGAGGAGTTGCTTTTTCAGCGATGTGGAGAAATGAGGAAACACCTATGTGT
 GCTAAAAGACAAAAAGAGCCCTATCTTATTCACCATGCAAGACTGGATCGCGTAAAGAAATTTGACAGGGCTTGTGAAATGTATGCTA
 AGAACCCAAAATAAGGGAATTGGTCAATCTTGTGCTGGTTGGTGGAGACCGAAGGAAATCCAAAGATTTGGAAGAAACAAGCTGAG
 ATGAAGAAAAATGTATTCAATTGATAGAAAATTAACAATTGAACGGCCAAATTCAGATGGAATTTCTCTCAGATGAACAGGGTTAGAAAATGG
 TGAACCTCTATCGGTACATTGCTGACACCAGGGGAGCATTCGTGCAACCTGCATTTTATGAGGCATTTGGGTGACTGTGTCGAGGCCA
 TGACATGTGGTTTGCCCAACATTGCAACCAACCATGGTGGTCTCTGAGATCATTAATCATGGGAAATCTGGTTTCCACATTGATCCA
 TACCATGTGTAGCAGGTCAGCGAGCTCCTTGCCCAATTTCTTTGAAAGGTGCAAGAAAGAGCCTTCTTACTGGACACCATTTTCAGCCGG
 TGGCTTGAAGCGTATCCAGGAAAAGTACACCTGGCAAAATTTACTCAGATCGGTTGCTGACGCTGGCTGGAGTTTATGATTTCTGGAAAT

GTGTTTCCAAGCTTGATCGGCCAGGAGATCCGCCGTTATCTGGAAATGTTTATGCTCTCAAGTATCGCAAGTTGGCTGAAGCTGTTCCA
TTGGCTGTTGATCAGTAA

FIGURE 18

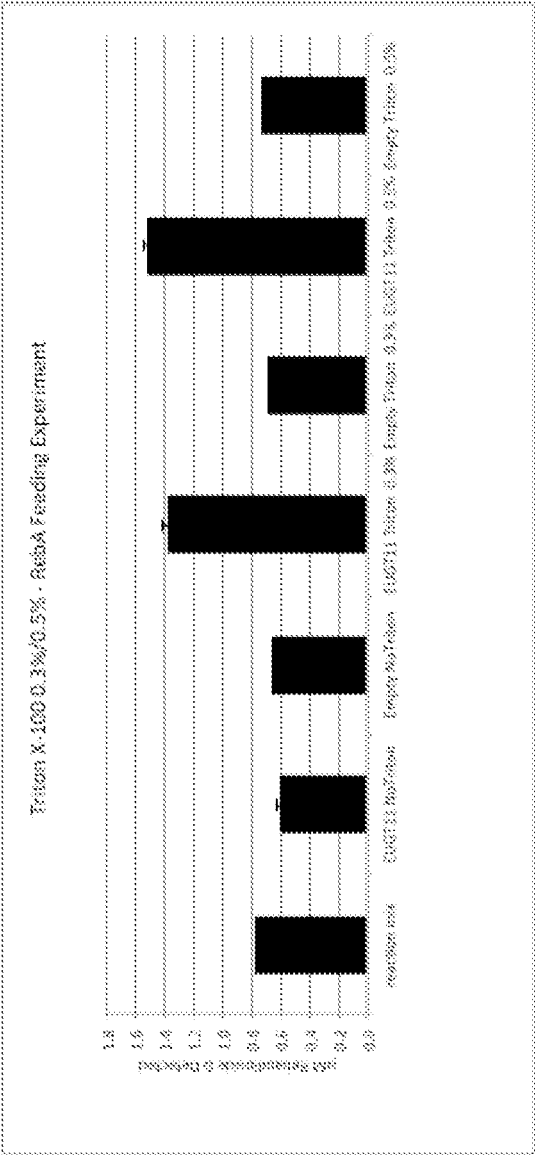


FIGURE 19

A.

MHITKPHAAMFSSPGMGHVIPVIELGKRLSANNGFHVTVFVLETDAAASQSKFLNSTGVDIVKLPSPDIYGLVDPDDHVVTKIGVIMRAAVPALRSKIAAM
 HQKPTALIVDLFGTDALCLAKEFNMLSVFIPTNARFLGVSYYPNLDKDIKEHTVQRNPLAIPGCEPVRFEDTLDAYLVPDEPVYRDFVRHGLAYPKADGIL
 VNTWEEEMEPKSLKLLNPILLGRVARVPVPIGPLCRPIQSSETHPVLDWLNQPNESVLYISFGSGGCLSAKQLTELAWGLEQSQQRVWVVRPPVDG
 SCCSEYVSANGGGTEDNTPEYLPEGFVSRTSDRGFVPSWAPQAEILSHRAVGGFLTHCGWSSTLESVVGVPVMIAWPLFAEQNMNAALLSDELGIAVRL
 DDPKEDISRWKIEALVRKVMTEKEGEAMRRKVKKL RDSAEMSLSIDGGGLAHESLCRVTKCQRFLEVVVDSL RGA

B.

MGAYETKPTKDAAALETQSPEDFDQPSPLRKIISVASIAAGVQFGWALQLSLTPYVQLLGIPIHKWSSLIWLCGPVSGMIVQPIVGFHSRCSRKFGRRRPF
 IATGAALVAVAVFLIGYAADFGYKMGDKLEEKVKVRAIGIFALGFWILDVANNTLQGPCRAFLADLAAGDAKTRVRANAFSFFMAVGNVLGVAAAGSYTNL
 HKMFPFTMTKACDIYCANLKTCFFLSITLLIVTVTSLWVYNDKQWSPPRNADDDKTSVPLFGEIGAFKVMKRPMWMLLIVTALNWIWIAWFPFLFDT
 DWMGREVFGGSDGNERSKKIYSLGVQSGAMGLMFNSIVLGFMSLGVWIGRKLGGAKRLWGVNFILAAGLAMTVLTKFAEDHRKTAGDLAGPSAS
 VKAGALSILFAVLGIPLAITFTPFALASIFSSCSGAGQGLSLGVNLAIVIPQMIVSLGGGPFDAIFGGGNLPAFIVAIAAISGVLTALTVLPSPPDPAPKATTM
 GGFH

C.

MAERVLTRVHSLRERLDATLAAHRNDVLLFMSRLETHGKGILKPHQLLAEFEEINKDGKQKIHDFHAFEEVLKSTQEAIVLPPWVALAIRLRPGVWEYVRVN
 VHALVVEELTVPEYLHFKEELVDGSKNGNFVLELDFEPFTASFPKPTLTKYIGDGVFELNRHLSAKMFHDKESMAPLLDFLRVHQYKGKTMMLNDRIKDLNT
 LQAVLRKAAEEYLTLSADTPYSEFEHKFQEIGLERGWGDTAERVLEMICMLLDLLGAPDSCCTLEKFLGRIPMVFNVVILSPHGYFAQENVLGYDPTGGQVVI
 LDQVPALEREMLKRIKEQGLDVKPRILITRLLPDAPGTCGQRLEKVGSEYSHILRVPFRTEKGVVRKWISRFVWPYMETFTEDVAKEVTAELQAKPDLVI
 GNYSEGNLVASLLAHKLGVTQCTIAHALEKTKYPDSDIYLSKFDEKYHFSCQFTADLIAMNHTDHIITSTFQEIAGSKDVTGQYESHMAFTMPGLYRVVHIGD
 VFDPKFNIVSPGADTNLYFPHTEKEKRLTSFHPEIEELLFSDVENEELCVLKDKKPILFTMARLDRVKNLTGLVELYAKNPKLRELNLVVVGGDRRKESKDL
 EEQAEEMKMYSLIETYNLNGQFRWISSQMNVRNNGELYRIADTRGAFVQPAFYEAFTGLTVVEAMTCGLPTFATNHGGPAEIIHIGSKSGFHIDPYHGEQVS
 ELLANFFERCKEPSYWDTISAGGLKRIQEKYTWQIYSDRLLTLAGVYGFWKCVSKLDRQEIRRYLEMFYALKYRKLAEAVPLAVDQ

Fig. 20

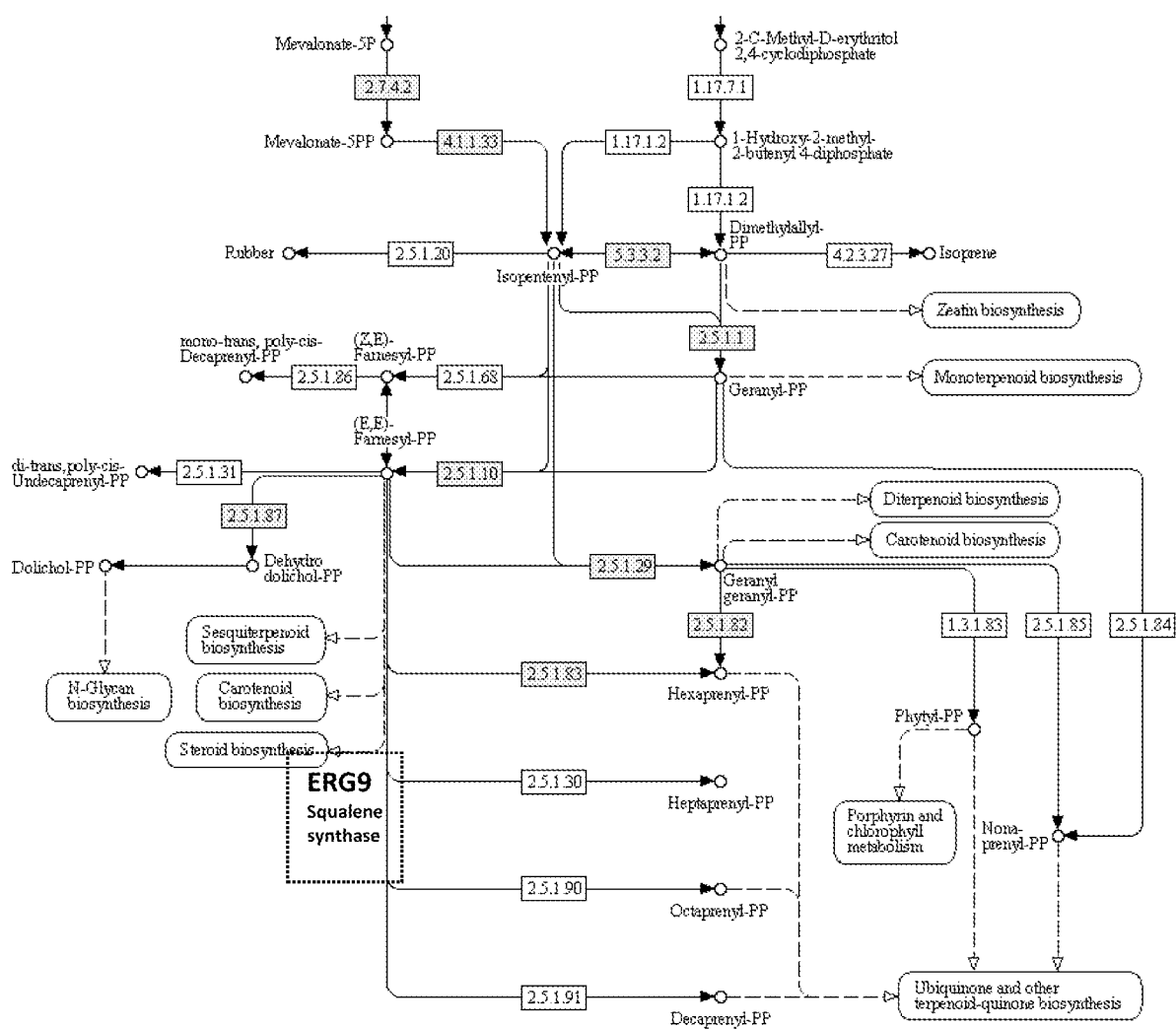


FIGURE 21

SEQ ID NO: 185 (ScCyc1 promoter)

gcgttggtggatcaagcccacgcgtaggcaatcctcgagcagatccgccaggcgtgtatatatagcgtggatggccaggcaacttttagtgcigacacatacaggcatatatatgtgtgc
gacgacacatgatcatatggcatgcatgtgtgctctgtatgtatataaaactctgttttcttctctaaatattcttctctatacattaggacctttgcagcataaattactacttctatagaca
cacaacacaaaatacacacactaaattaata

SEQ ID NO: 186 (ScKex2 promoter)

ggtcagcagctctgatgtagatacacgtatctcgacatgttttttactatacatataaagaataaaaaatgataacgtgtatatattattcataataatcaatgagggtcattttcigaa
acgcaaaaaaacggtaaatgggaaaaataaagatagaaaaaagaaaaacaaaggaaggtagcatattataaataactgagctgatacttcaacagcatcgctgaagagaacagattatt
gaaaccgaaacatttttctaaaggcaaacaggtaactcctatttgcggacgtgttctctcgtttcctatgcatataattctgtcataagcctgttttttctggcttaaacatcccgtttttaa
aagagaaatctattccacatatttcattcattcggctaccatactaaggataaaactaatccogtgttttggcctcgtcacataattataaaactaaccattatcaga

FIGURE 22

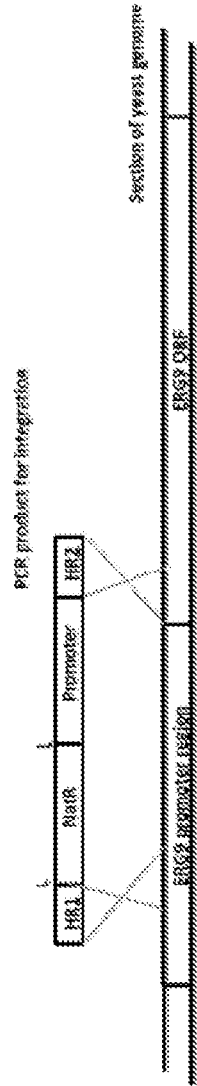


FIGURE 23

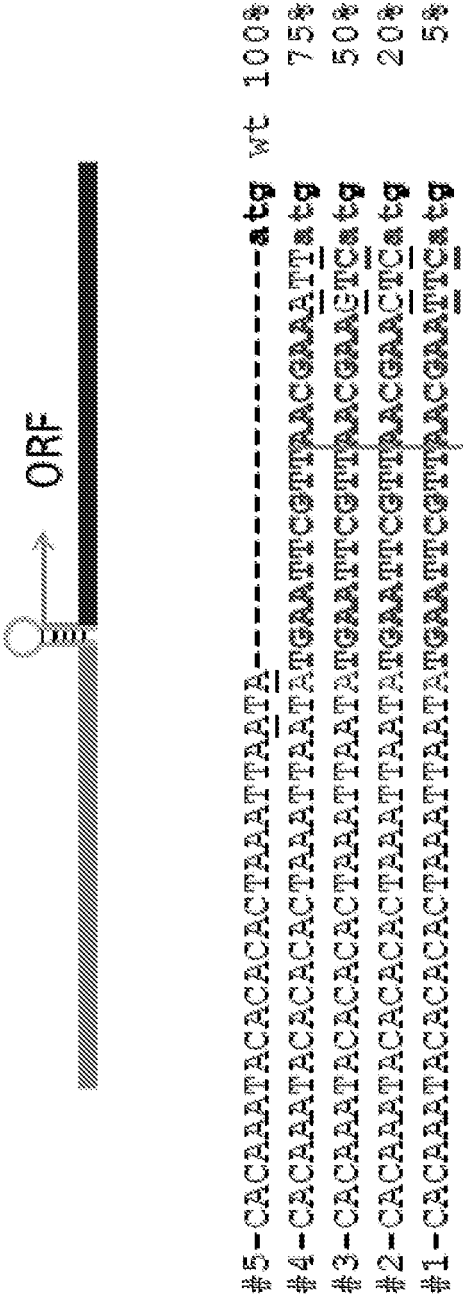


FIGURE 24

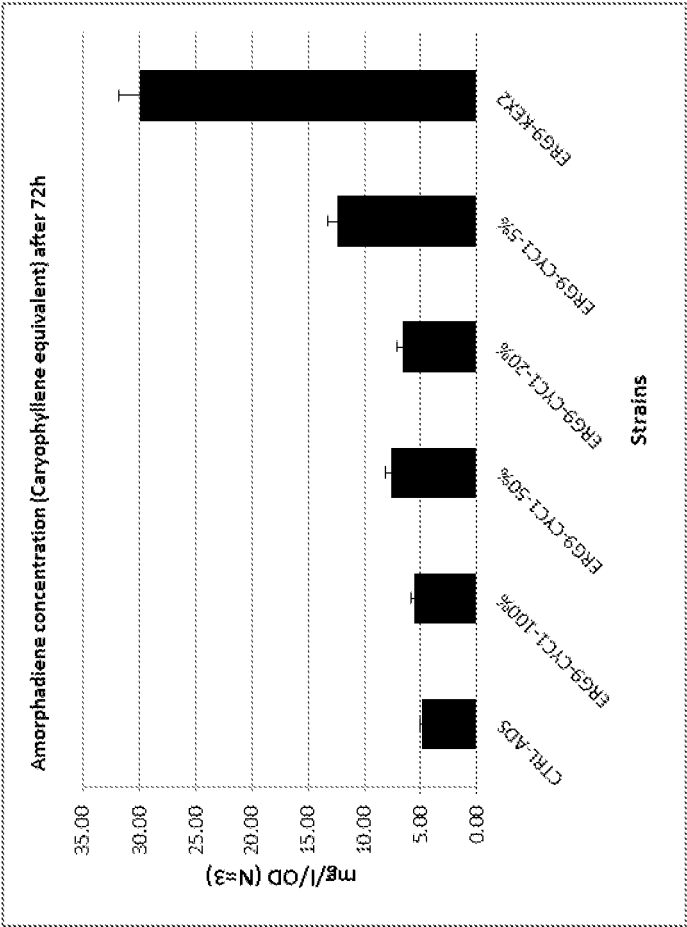


FIGURE 25

```

>sp|P29704|FDEFT_YEAST Squalene synthase OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=ERG9 PE=1 SV=2
MGKLLQLALHPVEMKAALKKFCRTPLFSIQDSTSPYLLHCFELLNLTSSFAAVIREL
HPELRNCVTLFYILRALDTIEDDMSIEHDLKIDLLRHFHEKLLLTWTFDGNAPDVKDR
AVLTFDESILJEHKLKPEYQEVKEITEKMGNGMADYILDENYNLNGLQTVHDYDVYCH
YVAGLVGDGLTRLJVIKAFANESLYSNEQLYESMGLFLQKTIIRDYNEDELVDGRSFWPK
EIWSQYAPQLKDFMKPENEQLGLDCINHLVLNALSHVIDVLTYLAGIHEQSTFQFCAIPQ
VMAIATLALVFNNREV LHGNVVKIRKGTTCYLILKSRTLRCVGEIFDYLYLRDIKSKLAVQD
PNFLKLNQISKIEQFMEEYQDKLPPNVKPNETPIFLKVKERSRYDDELVPTQQEEBYK
FNMVLSILSVLLGFYIYTLHRA

>sp|P36596|FDEFT_SCHPO Squalene synthase OS=Schizosaccharomyces pombe (strain ATCC 38366 / 972) GN=erg9 PE=1 SV=1
MSLANRIEIRCLCQYKLWNDLPSYGEDENVQNIIRRCYQLLDMTSRSFAVVIKELPNGI
REAVMIFYLVLRGLDVTEDDMTLPDLKKLPILRDFYKTIEVEGWTFNESGPNKDRQLLV
EFDVVIKEYLNLSEGYRNVISNITKEMGDGMAYYASLAEKNDGFSVETIEDFNKYCHYVA
GLVGIGLSRLFAQSKLEDPDLAHSQAISNSLGLFLQKVNIIRDYREDFDDNRHFWPREIW
SKYTSSEGDLCLPDNSEKALECLSDMTANALTHATDALVYLSQLKTQEIFNFCAPQVMA
IATLAAVFRNPDPVFQTNVKIRKQQA VQIILHSVNLKNVCDLFLRYTRDIHYKNTPKDPNF
LKISIECGKIEQVSESLFPRRFREMYEKAYVSKLSEQKKGGNGTKAILNDEQKELYRKDL
QKLGISILFVFFIILVCLA VIFYVFNIRIHWSDFKELNLF

>sp|Q9V753|FDEFT_YARLI Squalene synthase OS=Yarrowia lipolytica (strain CLIB 122 / E 150) GN=SQS1 PE=3 SV=1
MCKLIELLLHPSELSAAIHYKLRQPLHPRDLKSKESTELRRCYELLDVCSRSFAAVIREL
HPEVRDAVMLFYILRALDTIEDDMTL SRDIKIPILRDFTKCMKTPGWKFTDSDPNERDR
VVLQEFVVMTEFNKLPKYQEVYDITDRMGNGMADYVIDDDFNNGVDTIAAYDLYCH
HVAGIVGEGLTRITLAGFGTDVLHENPRLQESMGLFLQKVNIIRDYREDIDVNRFAWPR
EIWHKYAEEMRDFKPKYSKKALHCTSDLVANALGHATDCLDYLDNVNTPSTFTFCAIPQ

```

VMAIATLDLVYRNPVDFQKNVQLRKGTTSVLEASVSVGVCDIFTRYARKVYKKS DPND
 PNYFRVSVLCGKIEQHAALIKRQRGPPAKTIAQLEGKERKEMALSIVCLAVIFSMSG LMA
 YIAYVSGFRWSPREIFDSKMFPLRD

>sp|Q9HGZ6|FDFDFT_CANGA Squalene synthase OS=Candida glabrata (strain ATCC 2001 / CBS 138 / JCM 3761 / NBRC 0622 / NRRL Y-65) GN=ERG9 PE=3 SV=2
 MGKVLDLALHPLELRAALKLKFIHQPLFSTNDTRATPQLERCYELLNLTSRSF AAVIMEL
 HPELRNVIMVFYLILRALDVTEDDMTIDPQLKVKVLRFPDSKLDTTDWSFDGNDLKEKDR
 VVLTEFPCLGEYHKLKPEYQKVIKRTGLMGNGMADYILDENFNLNGVQTVKDYDKYCH
 YVAGLVGDGLTELIVLAGFGSDDL YHGKNSFQLYESMGLFLQKTNHIRDY AEDLDDGRSF
 WPKEIWSEYATKLTDFRDPKNTQKGVDCINHLVLNALTHVIDVLTYLSSHEQSSSFQFCA
 IPQVMAIATLAKVFNNPEVLRKNVIRKKGTTCDLILNSRTLKGCVDIFQYYLRDMKQRLP
 VEDPNYLKFNIVAKIEQFIEEMFQDNLPAGVEPRETMIVLYKVQERLKWDTQVIPRVQEE
 DYKFNMAHSVVFVCLLSFYFFTK

>sp|Q752X9|FDFDFT_ASHGO Squalene synthase OS=Ashbya gossypii (strain ATCC 10895 / CBS 109,51 / FGSC 9923 / NRRL Y-1056) GN=ERG9 PE=3 SV=1
 MGKVVQLFTHPLELKAALKLKFLREPLYPADDTQGS AELKRCYQLLQRTSRSF AAVIMEL
 HPELRNAVMLFYILRALDVTEDDMTISPVKVPLLRFPDQKLKLDTSFSGNAKTEKDR
 DVLVEFSTILAEFHKLKPEYQQVIADITHKMNGMGADYILDEKFNLSGLETIQDYDRYCH
 YVAGLVGDGLTHLIMLAKFSSPGLYYDSPDLYESMGLFLQKTNHIRDY AEDLADGRSFWP
 KEIWSHYADDLASFSKPENATAGVYCINHLVLNALGHVQHVLTYLASLREQSSSFQFCAIP
 QVMAIATLALVFGNERVLQTSVKIRKGTTCVYLILKSRTFQGCVEIFEHYLRDIRKRLTVA
 DPNYLKLNIEIAKLDKFIEMYQDKLPVGAKPQETETIYKKVRERSAYDLEVLPREQEEEF
 KFNVLLSILFTVFGALYWYAK

>sp|O74I65|FDFDFT_CYBJA Squalene synthase OS=Cyberlindnera jadinii GN=ERG9 PE=3 SV=1
 MGKLLQLALHPDELASIVQKLFKKNENARNPATESAELIRCYELLNLTSRSF AAVIEEL
 HPELRNVIMVFYLILRALDVTVEVDMISIENSVKLPVLRQFHEKLDTKDWTDFGNSPNEKDR
 CVLVEFDRIILGQYHELPQYQKVIKEITEKMGNGMGADYIENENFNSNGLLTIEDYDLYCY

YVAGLVGDGLTQLIVLAKFGNSELVSNKQLFKSMGLFLQKTNIRDYEEQVDGRAFWPK
 EIWKGYANELSDFMKPENQSGLWCISELVCNALDHVIDVLQYLALVEEQTSFNFC AIPQ
 VMAIATLELVFQNPQVLTHVKIRKGTTVSLJLESRTLGCARIFRRLKJHHKSHPSD
 PNYLRIGITIGKIEQELDGMYPHYVPKGIPTQTTSIRTQVVVKRLQLDEPMKRDIDEEILK
 TRILLLSLGVAVFGVVVYGVVRIH

>sp|P78589|FDFT_CANAL Squalene synthase OS=Candida albicans GN=ERG9 PE=3 SV=1
 MGKFLQLLSHPTTELKAVIQLFGRQPLHPGKRDVNDKELGRCYELLNLTSRSF AAVIEEL
 HPELRDAVMIFYLVLRALDTIEDDMTIKSSIKPLLRREFDTKLNKTKNWTFDGYGPNKDR
 TVLVEFDKILNVVHRLKPYQDIHKSTITKMGNGMADYILDEEFNVYGVATVEDYNLYCH
 YVAGLVGEGLTNLFVLANFGDKTLTENNAKADSMGLFLQKTNIRDYHEDLQDGRSFWP
 REIWSKYTENLQDFHKVKTPAKEFAGVSCINELVLNALGHVTDCLDYL SLVKDPSSFSC
 AIPQVMAVATLAEVYNNPKVLHG VVKIRKGTTCRLILESRTLPGVVVKIFKEYIQVINHKS
 SVRDPNYLKIGIKCGEIEQYCEMIVPNKQALPPSMKSLPENKFTKIVASRESIDLSVQRR
 IEPGNFNCNVVLFGIGALILSLIYFVLY

>sp|P38604|ERG7_YEAST Lanosterol synthase OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=ERG7 PE=1 SV=6
 MTEFYSDTIGLPKTDPRLWRLRTDELGRESWEYLT PQQAANDPPSTFTQWLLQDPKFPQP
 HPERNKHSPDFS AFDACHNGASFFKLLQEPDSGIFPCQYKGPMTGTGYVAVNYIAGIEI
 PEHERIELRYIVNTAHPVDGGWGLHSVDKSTVFGTVLNYVILRLGLPKDHPVCAKARS
 TLLRLGGAIGSPHWGKIWLSALNLYKWEGVNPAPPETWLLPYSLPMHPGRWWVHTRGVYI
 PVSYL SLVKFSCPMTPLEEELRNEIYTKPFDKINFSKNRNTVCGVDLYYPHSTTLNIANS
 LVVFYEKYLNRNFIYSLKKKVVDLIKTELQNTDLSLCIAPVNQAFCALVTLIEEGVDSEA
 FQRLQYRKDALFHGPQGMITMGTVQWDCAFAIQYFFVAGLAERPEFYNTIVSAYKF
 LCHAQFDTECVPGSYRDKRKGAWGFSK TQGYTVADCTAEAIAIMVKNSPVSEVHHM
 ISSERLFEIGIDVLLNLQNISFEYGSFATYEEKIAPLAME TLNPAEVFGNIMVEYPYVEC
 TDSSVLGLTYFHKYFDYRKEEIRIRIAIEFIKKSQLPDGSWYGSWGICFTYAGMFALE
 ALHTVGETYENSSTVRKGCDFLVSKQMKDGGWGESMKSSSELHSYVDSEKSLVVTAWALI

ALLFAEYPNKEVIDRGIDLLKNRQEESEGEWKFESEGVFNHSCAIEYPSYRFLFPIKALG
 MYSRAYETHTL

 >sp|P37268|FDF1_HUMAN Squalene synthase OS=Homo sapiens GN=FDF1 PE=1 SV=1
 MEFVKCLGHPEEFYNLVRFRIGGKRKVMKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA
 LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLNHFNHFLYQPDWRFMESKEKDRQ
 VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWKDYCHYVAGLVG
 IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNHIRDYLEDQGGREFWPQEVWSRYV
 KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVNFCAIPQVMAIATL
 AACYNNQQVFKGAVKIRKGGQAVTLMMDATNMPAVKAIYQYMEEIYHRIPSDPSSSKTR
 QIISTIRTQNLPCQLISRSHYSPIYLSFVMLLAALSWQYLTLSQVTEYVQTGEH

 >sp|P53798|FDF1_MOUSE Squalene synthase OS=Mus musculus GN=Fdf1 PE=2 SV=2
 MEFVKCLGHPEEFYNLLRFRMGRRNFIPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA
 LDGDIRHAICVFYLVLRALDTVEDDMSISVEKKIPLL CNFHFTFLYDPEWRFTESKEKDRQ
 VLEDFPTISLEFRNLAEKYQTVIDDICHRMGCGMAEFVDKDVTSKQDWDKDYCHYVAGLVG
 IGLSRLFSASEFEDPIVGEDIECANSMGLFLQKTNHIRDYLEDQGEGRKFWPQEVWGRIYI
 KKLEDFAKPENVDVAVQCLNELITNLQHIPPVITYLSRLRNQSVNFCAIPQVMAIATL
 AACYNNQQVFKGVVKIRKGGQAVTLMMDATNMPAVKAIYQYIEIYHRIPSDPSSSKTK
 QVISKIRTQNLPCQLISRSHYSPIYLSFIMLLAALSWQYLSLTSQVTEYVQREH

 >sp|Q02769|FDF1_RAT Squalene synthase OS=Rattus norvegicus GN=Fdf1 PE=2 SV=1
 MEFVKCLGHPEEFYNLLRFRMGRRNFIPKMDRNSLSNLKTCYKYLDQTSRSFAAVIQA
 LDGDIRHAVCVFYILRAMDTVEDDMAISVEKKIPLL RNHFTFLYEPWRFTESKEKHRV
 VLEDFPTISLEFRNLAEKYQTVIADICHRMGCGMAEFLNKDVTSKQDWDKDYCHYVAGLVG
 IGLSRLFSASEFEDPIVGEDIECANSMGLFLQKTNHIRDYLEDQGEGRQFWPQEVWGKYV
 KKLEDFVKPENVDVAVKCLNELITNALQHIPPVITYLSRLRNQSVNFCAIPQVMAIATL

AACYNNHQVFKGVVKIRKGQAVTLMMDATNMPAVKAHYQYIEIYHRVPNSDPASKAK
QLISNRTQSLPNCQLISRSHYSPYLSFIMLLAALSWQYLSLTSQVTEDEVVQREH

Aspergillus nidulans GGPPS-2

MTSDSHFPPHAIPPRISSNRMSGASTRDKAALMGNFEDWLSKGDKLQTNLTLTKRHTR
NQSSLDGTYKDKGWSQENEEVIMGPYDMLQHPGKDLRRQMINAFNVWLKVPSESLAI
TKVWAMLHTASLLIDVEDNSLLRRGIPVAHSIYGTAQTINSANYVYFLALQEVQKLKSP
AAIDYVQELNLHRGQGMDFWRDTLTCPSEDEYLEMVGNKTGGFLRLAVKLMQAESST
GKDCVALNVGLVFCDDYLNLSDTTYTQNKGLCEDLTEGKFSFPIIHSIRSNPGNHQ
LINILRQRTKDEEVKRYALQYMESTGSFKHTQDVVRQLRARALQLIEEIEENSENGEQPEE
HNDGTMVRAILDKITESTLADTNTTTRDINGNCATR

Saccharomyces cerevisiae GGPPS-3

MEAKIDELINNDPVWSSQNESLISKPNHILLKPGKNFRLNLIVQINRVNMLPKDQLAIV
SQIVELLHNSSLIDIEDNAPLRGQTTSHLIFGVPSTINTANYMYFRAMQLVSQLTTK
EPLYHNLTIFNEELINLHRGQGLDIYWRDFLPEIPTQEMYLNVMVNMKTGGFLRLTLRL
MEALSPSSHHGHSIVPFINLLGIYQIRDDYLNKDFQMSSEKGFADITEGKLSFPPIVH
ALNFTTKGQTEQHNEILRILLRTSDKDIKLLIQILEFDTNSLAYTKNFINQLVNMK
NDNENKYLFDLASHSDTATNLHDELLYIIDHLSEL

FIGURE 26

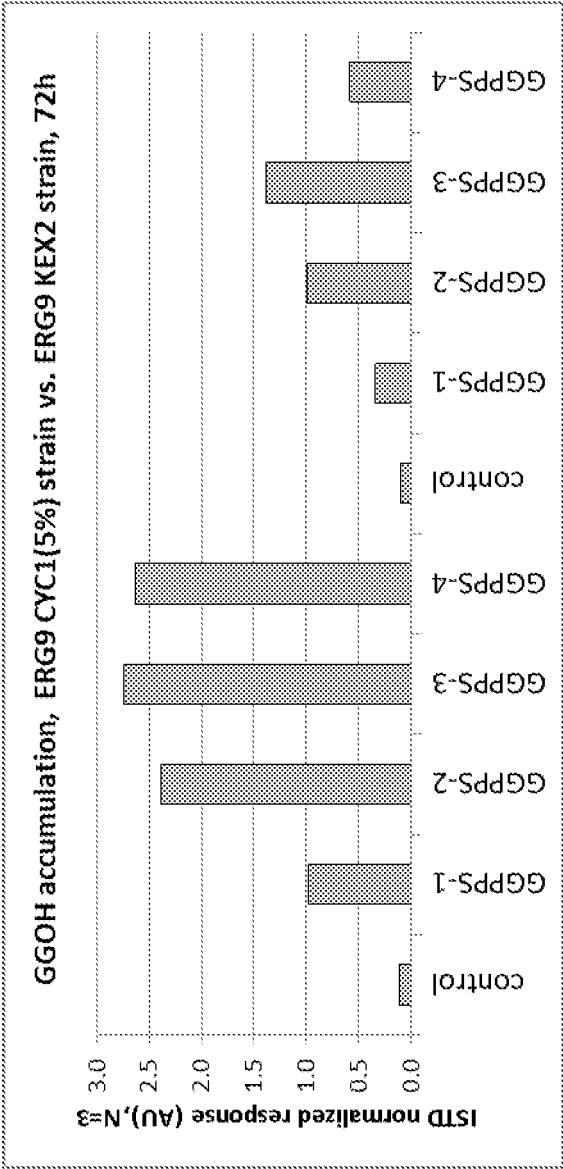


FIGURE 27

