

Figure 1

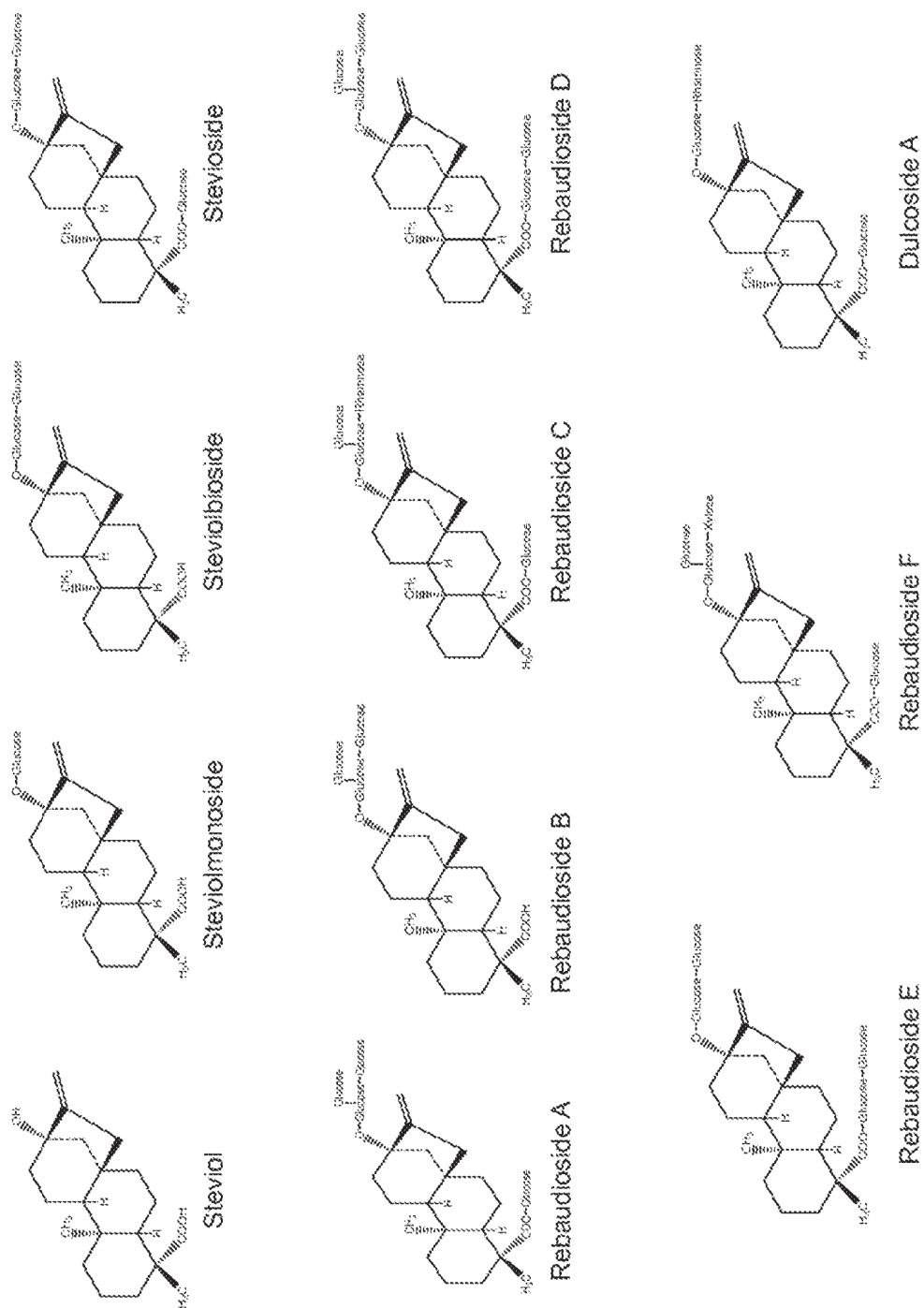


FIG 2A

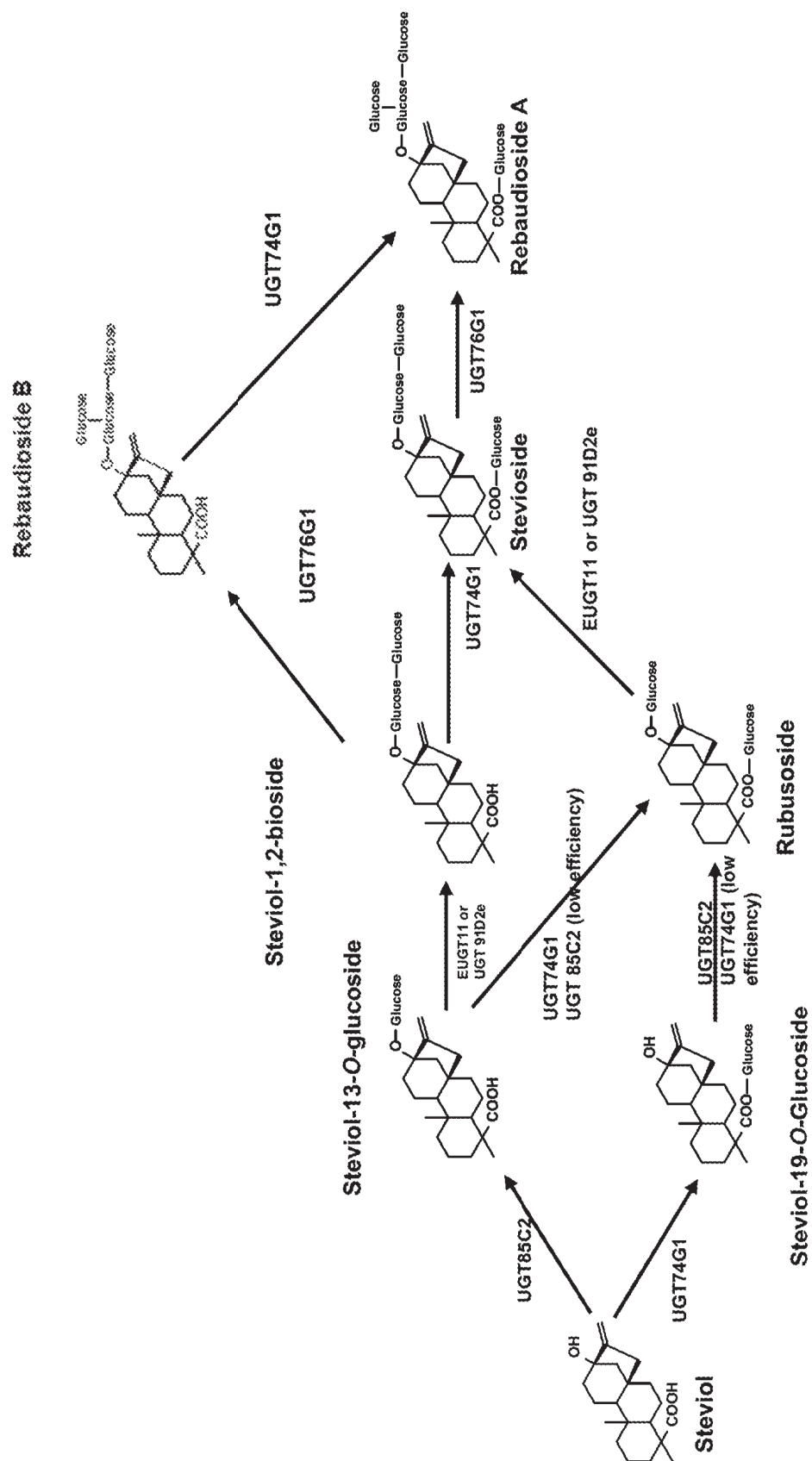


FIG 2B

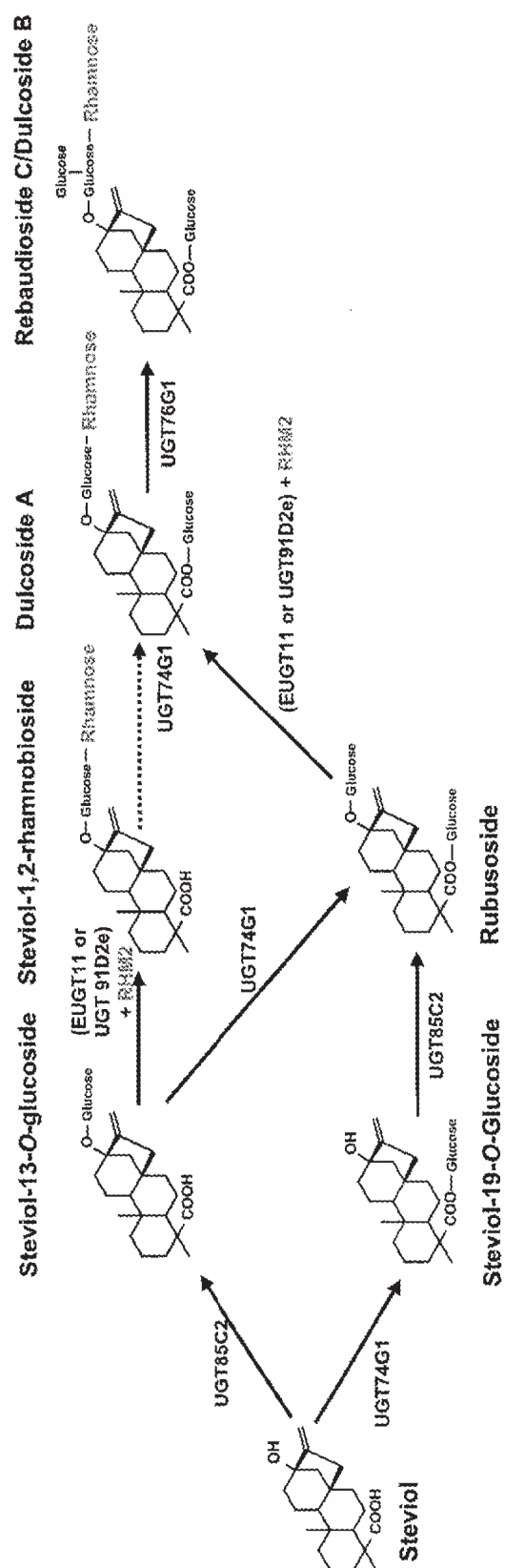


FIG. 2C

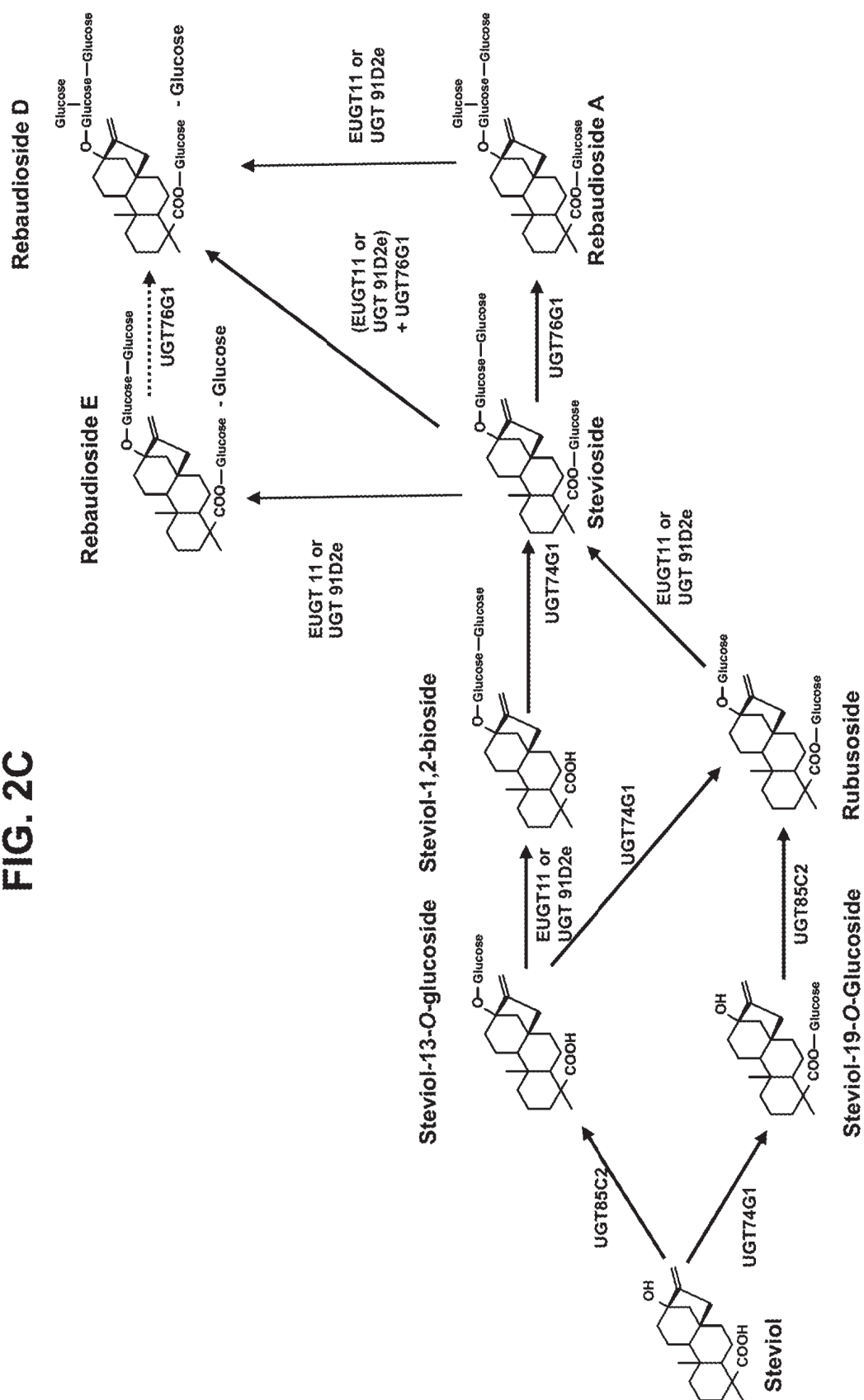


FIG. 2D

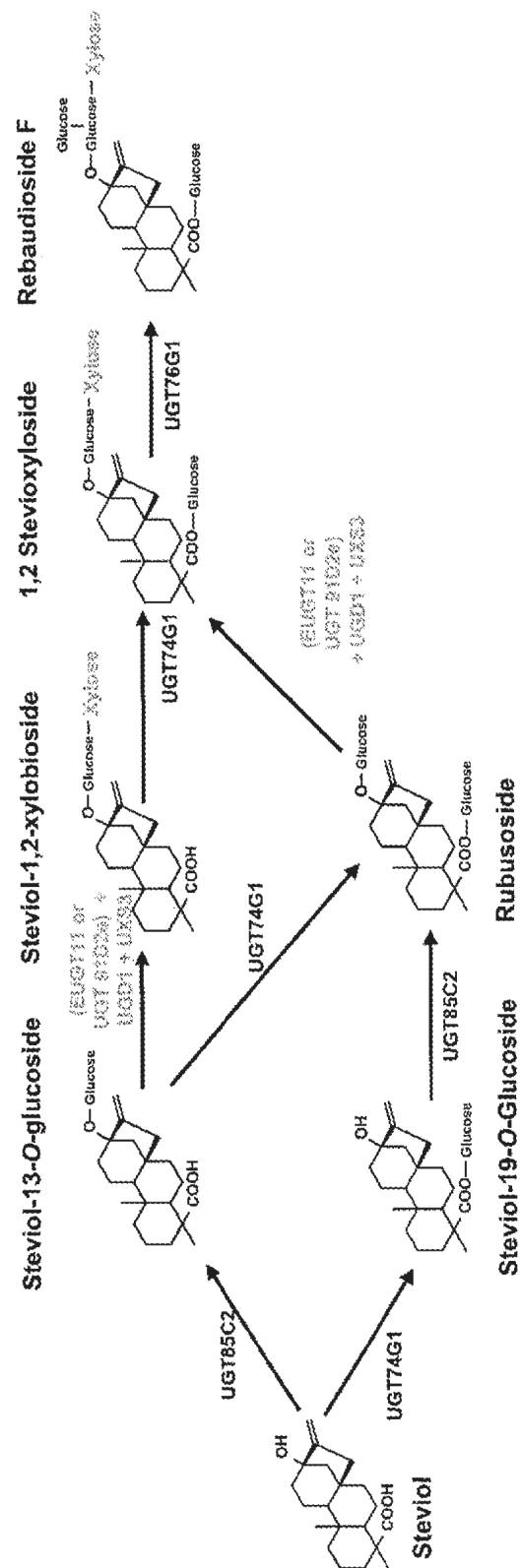


FIGURE 3

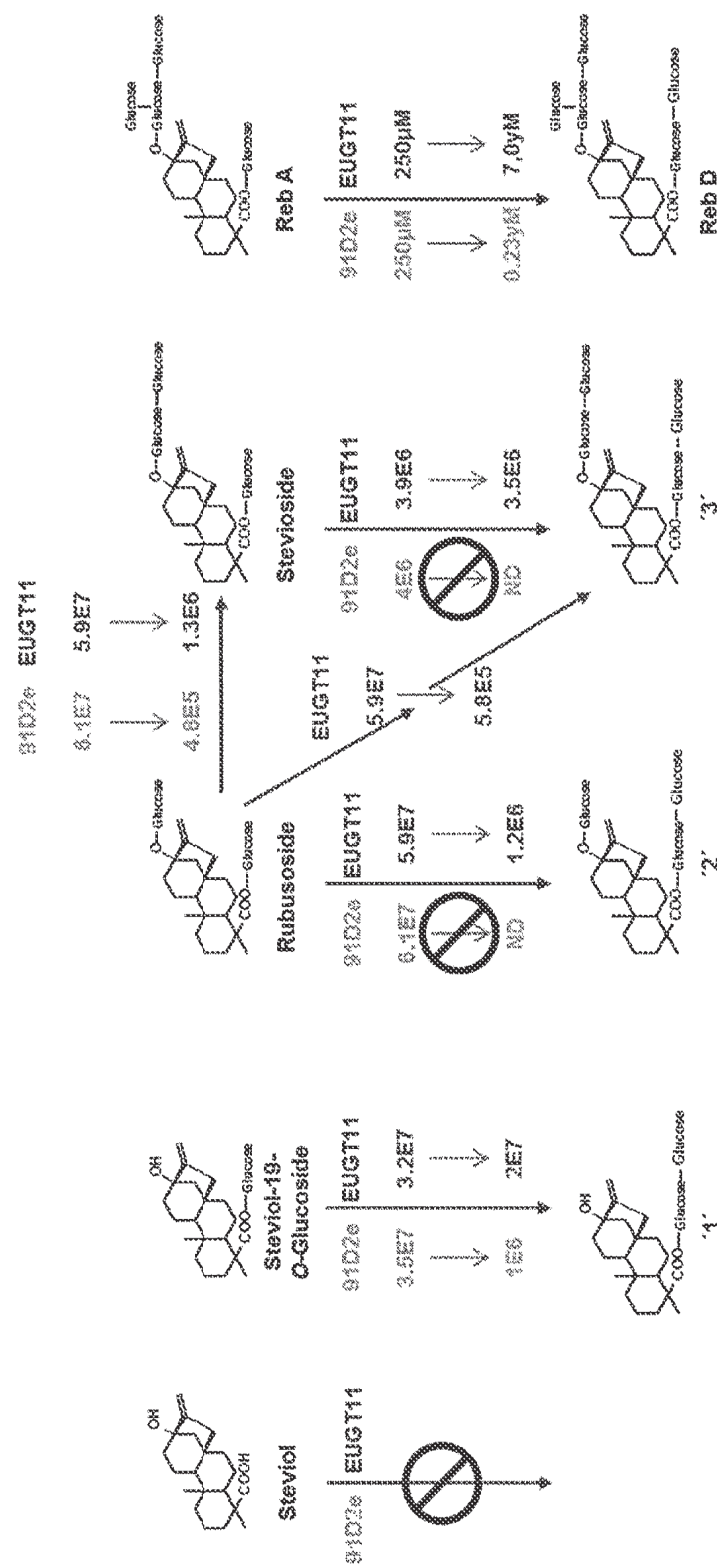


FIGURE 4

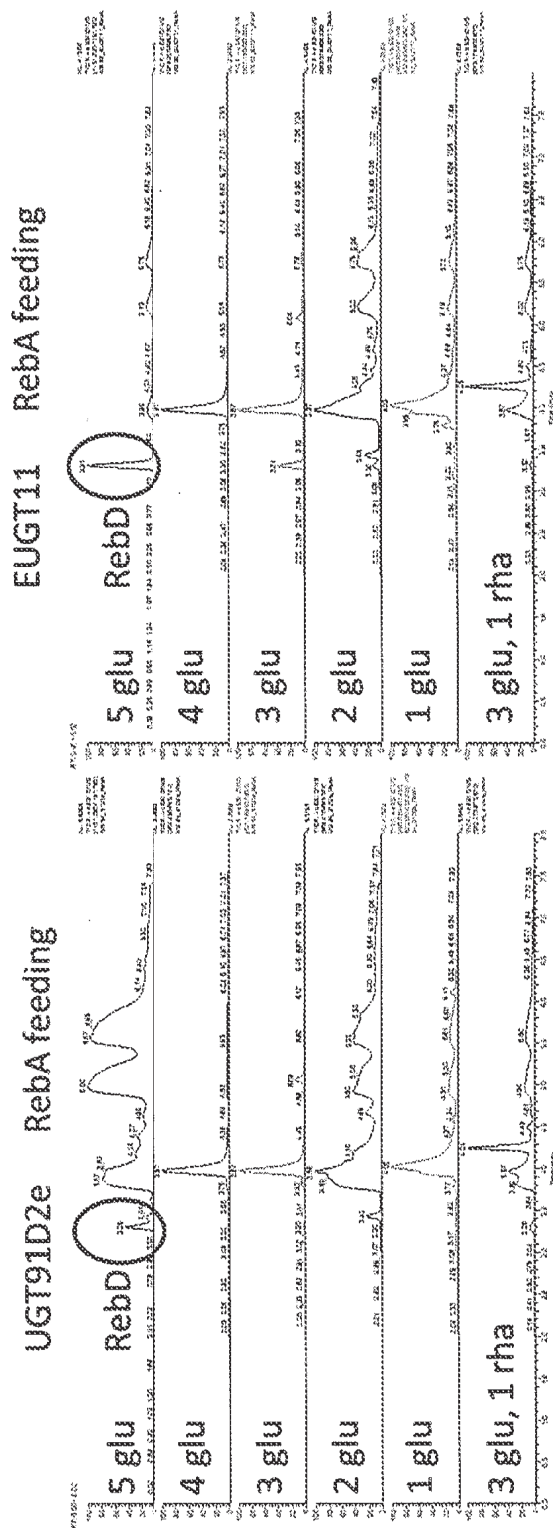


FIGURE 5

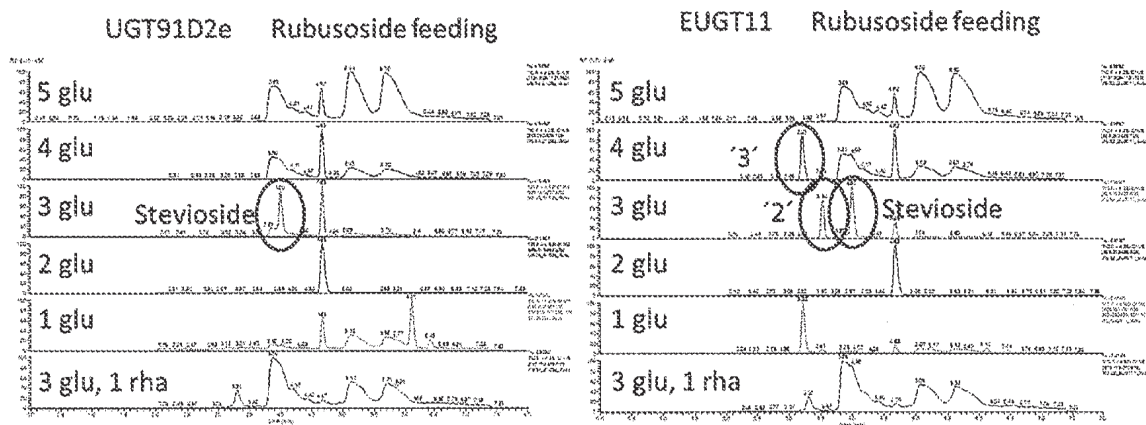


FIGURE 6

```
#fasta36 -p -q -w 80 -m 6 -l -Z 10000 -f 10 -g -2 TMP.q TMP.q22
FASTA searches a protein or DNA sequence data bank
version 36.3.5a Jun, 2011 (preload8)
```

Please advise

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= 0.1566; K=0.007058

statistics sampled from $\mathbf{I}^{(1)}$ 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:

XEROX

```

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

```

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

>>QUERY

inltn: 1037 inlt1: 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

	10	20	30	40	50	60	70	80
QUERY	MDSGYSSSSYAAAACNMHWV	ICPMIAEGHII	IPCTDI	ACPLASEGHRVSEKVS	TPTNTSR	LPVVRPALAPLVA	FTVALPL	PRVEG

[illegible]

QUERY	10	20	30	40	50	60	70
MATSDSIVDBRKQDHVATFFPLAFCHILPFLYQLSKLLIAEKGHKVSEFLSTIKNIQRLES---HISPLINVVQLIDLPKRVQ							

	90	100	110	120	130	140	150
--	----	-----	-----	-----	-----	-----	-----

QUERY LPDGAESTNDV-PHDRPDWVELHRRAFDGLAAPSEFGLGTACADWIVDVFNH--AAAAAL-.....EH---KVPCCAMLL

```

QUERY LPEDARATDVHPEDIPYL ---KKASDGLQPEVTRFLEQHSPDWLYDYTHYWLFPASLGISRAHFSVTTPWAIAYM

```

80 90 100 110 120 130 140 150

[illegible]

462 residues in 1 query sequences

473 residues in 1 library sequences

Tcomplib [36, 3.5a Jun, 2011 (preload8)] (16 proc in memory [0G])

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.

WYNSTIEGLMFGHPJLMIPIFGDQGPNNARJLEAKNAGLOVARNDGDSFDREGVAAAIRAVAVEEESKVFQAKAKKLQEAIVADMACHERYIDGFIQOLRSYKD

>EUGT-11 Native gene.

ATGGACTCGGCTACTCTCTACGCCGCCGCCGGGATGACAGTCGTGATCTGCCGTGGCTCGCTTGGGCACCTGCTCCGCTGCTCGACCTCGCCACGCGCTCGC
TTCGCGGGGCCACCGGTGTCTCTCACGCCGCCGGAACATATCCGGCTCCCGCGGTGCGCCCGCTCGCCGCTCGTTCGCTTCGTGGCGCTGCGCTCCCGCGC
GTGAGGGGCTCCCGACGGCGGAGTCCACCAACGACGTCCCGACGACAGGCCGACATGGTCAGCTCCACCGAGGGCTTCGACGGGCTCGCCGCGCTTCGCGGAG
TTCCTTGGACCGGTGCGCGGACTGGGTATGCTGACGTCTTCCACCACTGGGCCGACGCCGCTCTCGAGCACAAAGTGCATGTGCAATGATGTTGTGGGCTCTGCAC
ATATGATCGCTTCATAGCAGACAGACGGCTCGAGCGCGCGAGACAGATCGCTCGCGGTGCGGGCAGGACGCCAGCGGGGCCAACGTTCGAGGTGGCGAGGAT
GGAAGTTGATACGAACAAAGCTCATCGGGAATGTCTCTCGCGAGCGCTTCTCTTGACGCTTCGAGGACAGCTCTGCTCGGGGGAGCTGCTGGAGTTCGAGCCGGA
GACCGTCCCGCTCTGTGACGCTCCGCGTAAAGCTATTACCTTCTCTGGCTTATGCCGCCGTTCGATGAAGGCCGCCGCGAGGACGGCGAGGATGCCACCGCTCGCTGGCTC
GACCGCGACCGCCCAAGTCGCGCTTAGCGACGAGGTGCCACTGGAGTGGAGAAGTCCACGAGCTCGCGCTCGGGCTGGAGCTCGCCGGGACGCGCTT
CCTCTGGGCTTTAGGAAGCCCACTGGGCTCCGACGCCGACCTCTCCCGCGGCTTCGAGGAGCGCACGCGCGGCCGCGCTGTGGGACGAGATGGGTTCCTCAGAT
GAGCATCTGGCGCAGCGCGCGTGGCGGCTTCTGACCACTCGGCTGGAATCGACCATCGAGGGCTCATGTCGGCCACCCGCTTATCATGCTGCCGATCTTCGGCGCAG
CAGGGACCGAACCGCGGCTAATCGAGGGGAAGACCGCCGATTCAGGTGGCAAGAACGACGCGCATGGATGTTTCGACCCGAGAGCGCTCGCGCGCGGATTCGTGCGAG
TCGCGGTGGAGGAAGAAAGCAGCAAGTGTTCAGGCCAAAGCAAGACTGCAGGAGATGTCGCGGACATGGCCTGCCATGAGAGGTACATCGACGGATTCAATTCAGCAA
TTGAGATCTTACAAGGATTGA

>EUGT-11 Optimized DNA.

ATGGGATAGTGGGTACTCTCATCTTATGCTGCTGCCGTGGTATGCACGTTTGATCTGCCCTTGGTCACTGTTACCATGTCTGGATTAGCCCAAGACTGGCC
TCAAGAGGCCATAGAGTATCATTTGTGCTACTCTAGAAATATCTCTGTTTACCACCAAGTCAGAACCTGCTCTAGCTCCTCTAGTTGCATTGCTTCCACTTCCAAGAGTAG
AAAGGATTGCCAGACGGCGCTGAATCTACTAATGACGTACCACATGATAGACCTGCACTGGTCSAATTGCATAGAAGAGCCTTTGATGGATTGGCAGCTCCATTTTCTGAGTTCCCTG
GGCACAGCATGTGCAGACTGGGTTATAGTCGATGTTTCATCACTGGGCTGTCAGCCGCATTGGAAACATAAAGGTGCCTTGTCCTATGATGTTGTTAGGGTCAGCACACATGA
TTCGCATCCATAGCTGATAGAAGATTGGAAGAGAGCTGAAACAGAAATCCCGAGCCGACGAGGACAAGGTAGGCCAGCTGCCGCCCAACCTTTGAAGTGGCTAGAAATGAAATTGA
TTTCGTAATAAGGTAGTTTCAGGGATGAGTCTGCTGAAAGGTTTTCCTGACATTTATCTAGATCATCATTTAGGTAGATCTCGCTCGAGTTCGAACTGAAACAGTACCTTT
TACTATCTACTTTGAGAGGCAAACTTATTACTTTCCCTTGGCTTAATGCCTCCATTACATGAAGGAAGGAGAGAAAGATGGTGAAGATGCTACTGTTAGGTGTTAGATGCCCAACT
GGCTAAGTCTGTTTACGTTGCAATGGGTTCTGAGGTACCACTAGGGGTGGAAGAGGTGCATGAATTAGCAATTAGCACTTGAGCTGGCCGGAAACAAGATTCTTTGGGCTTTGA
GAAACCAACACCGGTGTTCTGACGCCGACTTGCTACCAGCTGGGTTTCTGAAGAGAGAACAAGAGCGCGTGGTCTGCTACTAGATGGGTCCCAACAATGAGTATTTCTAGCTCA
TTCGAGCTGTAGGGGCCCTTCTAACCCATTGCGGTTGGAACTCAACAATAGAAGGACTGATGTTTGGTCACTCCACTTATATGTTACCAATCTTTGGCGATCAGGGACCTAACGCCAA
GATTGATTGAGGCAAGAAGACGAGGTCTGCAGGTTGCACGTTAATGATGTTGATGGTCCCTTTGATAGAGAAGGCGTTGCAGCTGCCATCAGAGCAGCTCGCCGTTGAGGAAGAGT
CATCTAAAAGTTTCCAGGCTAAGGCCAAAAAAATTACAAGAGATTGTGGCTGACATGGCTGTGTCACGAAAGATACATCGATGGTTTCATCCCAACAATTCAGAAGTTATAAAGACTAA

Figure 8

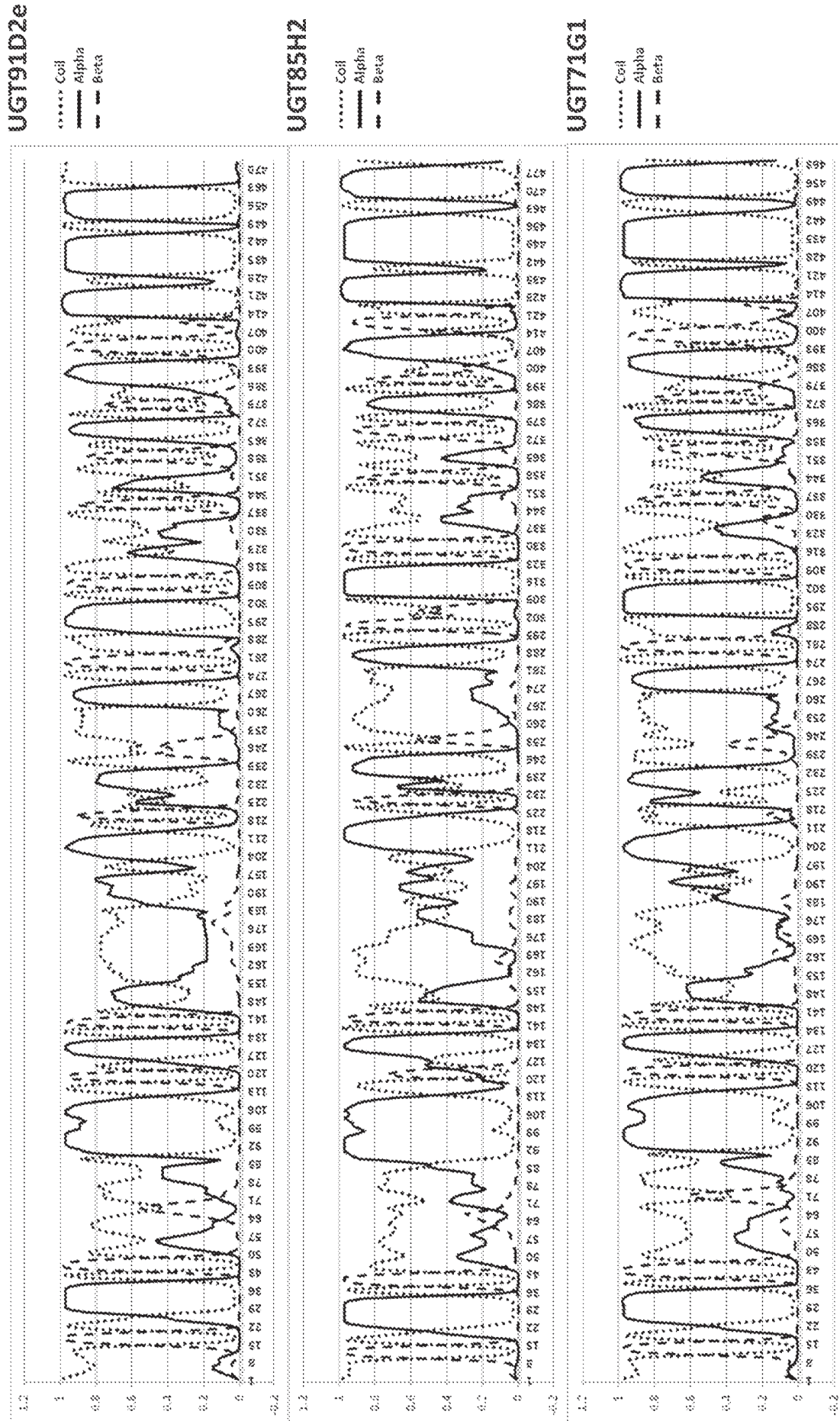


FIGURE 9

UGT91D1	20	40	60	80	100	120	140	160	180	200	220	240	260	280	300	320	340	360	380	400	420	440	460	480	500
UGT91D1	MATSDSIVDD	RKQLHVATFP	WLAFGHILPF	LQLSKLIAEK	GHKVSFLSTT	RNIQRLSSHI	60																		
UGT91D2e	MATSDSIVDD	RKQLHVATFP	WLAFGHILPF	LQLSKLIAEK	GHKVSFLSTT	RNIQRLSSHI	60																		
UGT91D1	SPLINVVQLT	LPRVQELPED	AEATTDVHPE	DIIQYLKKA	VDGLQPEVTRFL	EQHSPDWIYY	120																		
UGT91D2e	SPLINVVQLT	LPRVQELPED	AEATTDVHPE	DIIQYLKKA	VDGLQPEVTRFL	EQHSPDWIYY	120																		
UGT91D1	DETHYWLPSI	AASLGISRAH	FCVITPWTIA	YLA	PSADAMI	NDSDGRITVE	DLTTPPKWFP	180																	
UGT91D2e	DETHYWLPSI	AASLGISRAH	FCVITPWTIA	YLA	PSADAMI	NDSDGRITVE	DLTTPPKWFP	180																	
UGT91D1	FPTKVCWRKH	DLARMEPYEA	PGISDGYRMG	MVF	KGSDCLL	FKCYHEFGTQ	WLPLLETLHQ	240																	
UGT91D2e	FPTKVCWRKH	DLARMEPYEA	PGISDGYRMG	MVF	KGSDCLL	FKCYHEFGTQ	WLPLLETLHQ	240																	
UGT91D1	VPVVPVGLLP	PEIPGDEKDE	TWVSIKKWLD	GKQKGSVVYV	ALGSEALVSD	TEVVELALGL	300																		
UGT91D2e	VPVVPVGLLP	PEIPGDEKDE	TWVSIKKWLD	GKQKGSVVYV	ALGSEALVSD	TEVVELALGL	300																		
UGT91D1	ELSGLPFVWA	YRKPKGPAKS	DSVELPDGFV	ERTDRGLVW	TSWAPQLRIL	SHESVCGFLT	360																		
UGT91D2e	ELSGLPFVWA	YRKPKGPAKS	DSVELPDGFV	ERTDRGLVW	TSWAPQLRIL	SHESVCGFLT	360																		
UGT91D1	HCGSGSIVEG	LMFGHPLIML	PIFGDQPLNA	RLLEDKQVGI	EIPRNEEDGC	LTKESVARSL	420																		
UGT91D2e	HCGSGSIVEG	LMFGHPLIML	PIFGDQPLNA	RLLEDKQVGI	EIPRNEEDGC	LTKESVARSL	420																		
UGT91D1	RSVVVENEGE	IYKANARALS	KIYNDTKVEK	EYVSQFVDYL	EKNARAVAID	HES	474																		
UGT91D2e	RSVVVENEGE	IYKANARELS	KIYNDTKVEK	EYVSQFVDYL	EKNARAVAID	HES	474																		

FIGURE 10

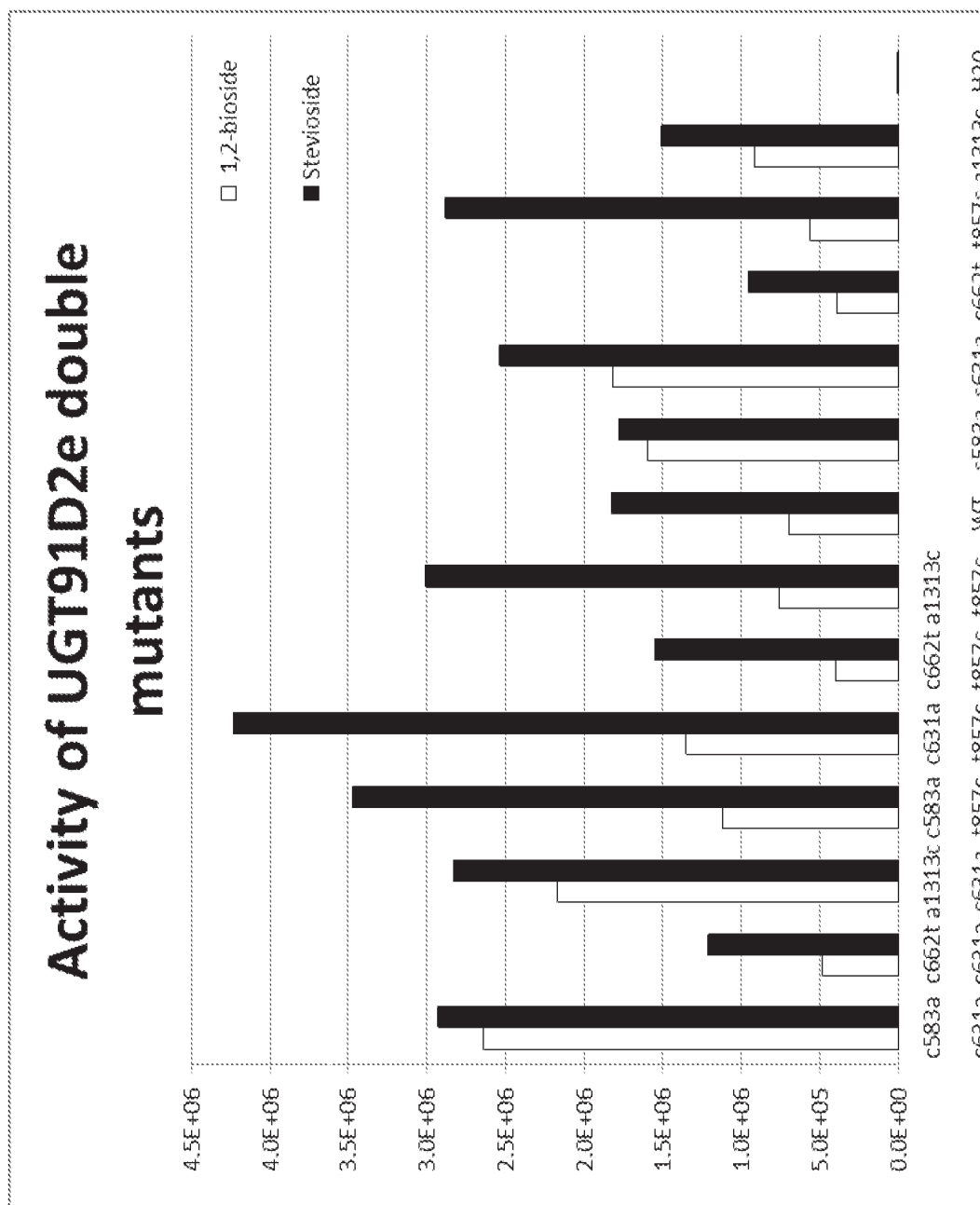


FIGURE 11

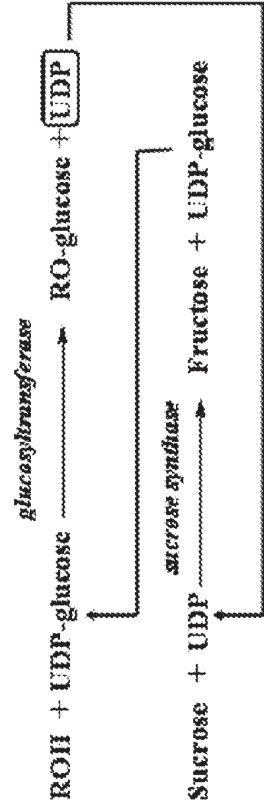


FIGURE 12

FIG. 12A

ATGAGGCTTCATATACATTTCCATTCTTCTGCTTCTAGCTTCGTACCTCTTACCACCCAACTTCGTCTGTAATCCGCCAATCT
 ACCGCCGACGGTGTCCCATCCCATTAATCGGTCACTCTACCTCCTCAAAAACCACTCTATAGAACACTAGCCAAAATTGCCG
 CCAAAATACGGCCCTATCCTCAACTCCAACTAGGTACGCCGTGCTCGTAATCTCCTCCCTTCCGCCGCCGAAGAAATGCTTCACC
 AACAAACGACGTTATCTTCGCCAACCGTCCGAAGACGCTATTCCGAAAAAATTGTAGGGGTACCAGCCTCGGGTCGCTGTCGTACGGCGA
 CCAGTGGCGCAACCTCCGCCGGTTCATCCATCGAGATTCTATCAGTCCACCGGCTCAACGAGTTTCACGCACATACGTGTTGACGAAA
 ACCGGCTTCTGATCCGTAAACTAAGATCCAGTTCTTCTCCGGTGACTCTGATAACGGTGTTTTACGCACTAACGTTAAATGTGATTATG
 AGAATGATCTCCGGAAAGAGGTATTTCCGACTCGGGTGATCGGGAATTGGAGGAGGAAGGAAGGATTCCGGGAGATACCTCGATGAGAC
 ACTTTTGCTCGCGGGTGCTTCTAAATGTTGGGATTACTTGGCGATTCTGAAATTGGTTGGGGTGAAGAGCTTGGAGAAAGACTAATCG
 CATTGCAGAAAAAGAGAGATGATTTCTTTTCAGGGACTGATCGAGCAAGTTCGGAATCTAGAGGGCTAAAGTGGGAAAAAGGAAGGAAG
 ACGATGATCGAGTTGTTTATCCCTACAAGAATCTGAACCTGAGTATTACACTGACGCCATGATCCGATCATTTTGTGCTGGGTTTATT
 AGCAGCAGGAGTGATACATCGGCTGGAACATATGGAATGGCGATGTCCTTTTGGTAAACCAACCGCACGTATTAAAAAAGGCACAGG
 CTGAAATTGATCGAGTCACTCGGCAACAACCGTCTAATTGATGAGTCCGACATAGGGAATATACCTTACATTGGTTGCATCATAAAACGAG
 ACGCTCAGATTGTACCCTCGGGGCCGTTGCTATTTCGCCATGAGTCAATCAGCGGACTGTGTATCAGCGGTACAACATCCCTCGTGG
 GACGATGCTTATTGTCAACCAATGGCGGATACATCATGACCCAAAGGTGTGGGACGACCCCTGAGACATTCAAACCGGAAAGATTCCAAG
 GGCTTGAAGGGACACGAGACGGGTTTAAGCTGATGCCCTTTTGGGTCCGGAAGGAGGGGTTGTCCAGGGAGGGATTGGCGATTCTGTTTG
 CTTGGGATGACTCTCGGGTCGGTTATCCAATGCTTTGATTGGGAACGAGTCGGTGACGAGATGGTTGATATGACCGAAGGTCTTGGGGT
 CACGTTGCCCTAAAGCTGTACCATTAGTCGCCAAGTGCAAGCCCGCTTCCGAAATGACGAAATCTACTCTCTGAGCTATGA

FIG. 12B

ATGGAAGCCCTTACCTATACATTTCTATTTTGTCTTTTACTGGCATCATACCTGTTCAACCACTCAACTTAGAAGGAAGAGCGCTAATCT
 ACCACCAACCGTGTTCATCAATACCAATCATTTGGACACTTACTACTCAAAAAGCCCTCTTTATAGAACTTTAGCAAAAATTGCCG
 CTAAGTACGGACCAATACCTGCAATTACAACCTCGGCTACAGACGTGTTCTGGTGATTTCCTCACCATCAGCAGCAGAGTGCTTTACC
 AATAACGATGTAATCTTCGCAAAATAGACCTAAGACATTGTTTGGCAAAATAGTGGGTGGAAACATCCCTTGGCAGTTTATCTCTACGGCGA

TCAATGGCGTAATCTAAGGAGAGTAGCTTCTATCGAAAATCCTATCAGTTCATAGGTTGAACGAAATTTTCATGATATCAGAGTGGATGAGA
 ACAGATTGTTAATTAGAAAACCTAGAAAGTTTCATCTTCTCCTGTACTCTTATAACACAGTCTTTTATGCTCTAACATTGAACGTCATTATG
 AGAATGATCTCTGGCAAAAGATATTTTCGACAGTGGGATAGAGAATTGGAGGAGGAAGGTAAGAGATTTTCGAGAAAATCTTAGACGAAAC
 GTTGCTTCTAGCCGGTGCTTCTAATGTTGGCGACTACTTACCAATATTGAACCTGGTTGGGAGTTAAGTCTCTTGAAGAAAATTGATCG
 CTTTGCAGAAAAGAGAGATGACTTTTCCAGGGTTTGATTGAACAGGTTAGAAAATCTCGTGGTAAAGTAGGCAAAAGGTAGAAAAA
 ACGATGATCGAACTCTTATTATCTTTGCAAGAGTCAGAACCTGAGTACTATACAGATGCTATGATAAGATCTTTTGTCTCCTAGGCTGCT
 GGCTGCAGGTAGTGATACTTCAGCGGGCACTATGGAATGGGCCATGAGCTTACTGCTCAATCACCCACATGTATTGAAGAAAAGCTCAAG
 CTGAAATCGATAGAGTTATCGGTAATAACAGATTGATTGACGAGTCAGACATTTGGAATAATATCCCTTACATCGGGTGATTATCAATGAA
 ACTCTAAGACTCTATCCAGCAGGGCCATTGTTGTTCCCATGAAAGTTCTGCCGACTGCGTTATTTCCGGTTACAAATATACCTAGAGG
 TACAAATGTTAATCGTAAACCAATGGCGATTTCATCACGATCCTAAAGCTGGGATGATCCTGAAACCTTAAACCTGAAAGATTTCAAG
 GATTAGAAGAACTAGAGATGGTTTCAAACTTATGCCATTTCGGTCTGGGAGAAAGAGGATGTCCAGGTGAAGGTTTGGCAATAAAGGCTG
 TTAGGGATGACACTAGGCTCAGTGATCCCAATGTTTGTATTGGAGAGAGTAGGAGATGAGATGGTTGACATGACAGAAAGGTTTGGGTGT
 CACACTTCCTAAGGCCGTTCCATTAGTTGCCAAAATGTAAAGCCACGTTCCGAAATGACTAATCTCCTATCCGAACTTTAA

FIG. 12C

MEASYLYISILLLLASYLFTTQLRRKKSANLPPTVFPSIPIIGHLYLLKKPLYRTLAKIAA
 KYGPILQLQGYRRVLVISSPSAAEEECFTNNDVIFANRPKTLFGKIVGGTSLGSLSYGDQ
 WRNLRRVASIEILSVHRLNEFHDIRVDENRLLIRKLRSSSPVTLITVFYALTINVMRM
 ISGKRYFDSGDRELEEEGKRFRFRIIDETLLLAGASNVGDYLPILNWLGVKSLEKKLIAAQ
 KKRDDFFQGLIEQVRKSRGAKVGKGRKTMIELLLSLQSESEPEYTDAMIRSFVLGLLAAG
 SDTSAGTMEWAMSLLVNHPVLKKAQAEIDRVIGNNRLLIDESDIGNIPYIGCIINETLRL
 YPAGPLLFPHESSADCVISGYNIPRGTMILIVNQWAIHHDPKVDDPETFKPERFQGLEGT
 RDGFKLMPFFSGRRGCPGEGLAIRLLGMTLGSVIQCFDWERNVGDENVDMTEGLGVTLPKA
 VPLVAKCKPRSEMTNLLSEL

FIGURE 13A

MPFGIDNTDFTVLAVLLVYKRNSIKELMSDDGDITAVS
 SGNRDIAQVVTENNNKYLVLASQTGTAEDYAKKFSKELVAKFNLVNMCADVENDFE
 SLNDVPVIVSIFISTYGEFDPDGA VNFEDFICNAEAGALS NLRNMFGLGNSTYEFF
 NGAAKKAEKHLAAGAIRLKLGEADDGAGTTDEDYMAWKDSILEVLKDELHLDEQEA
 KFTSQFYQYTVLNEITDSMSLGEPSAHYLP SHQLNRNADG IQLGPFDL SQPYIAPIVKS
 RELFSSNDRNCIHSEFDLSGSNIKYSTGDHLAVWPSNPLEKVEQFLSIFNLDPETIFD
 LKPLDPTVKVPFPTTTIGAAIKHYLEITGPVSRQLFSSLIQFAPNADVKEKLTLLSK
 DKDQFAVEITSKYFNIAADALKYLSDGAKWDTVPMQFLVESVPQMTPRYYSISSSLSE
 KQTVHVTSIVENFPNPELPDAPPVVGVTNLLRNITQLAQNNVNIAETNLPVHYDLNGP
 RKL FANYKLPVHVRRSNFRLPSNPSTPVMIGPGTG VAPFRGFIRERVAFL ESQKKGG
 NNVSLGKHILFYGSRNTDDFLYQDEWPEYAKKLDGSEFEMVVAHSRLPNTKKVYVQDKL
 KDYEDQVFEMINNGAFIYVCGDAKGMAGVSTALVGILSRGKSITTDEATELIKMLKT
 SGRYQEDVW (**S. cerevisiae** **CPR polypeptide encoded by NCP1**).

1 mtsalyasdl fklqlksimgt dslsddvvlv iattslalva gfvvllwkk tadrsgelkp
 61 lmipkslmak dedddldlgs gktrvsiffg tqtgtaegfa kalseeikar yekaavkvid
 121 lddyaadddq yeeklkketl affcvatygd geptdnaarf ykwfteener diklqqlayg
 181 vfallgnrqye hfnkigivld eelckkgakr lievglgddd qsieddfnaw keslwseldk
 241 llkdeddksv atpytavipe yrvvthdprf ttqksmesnv angnttidih hpcrvdvavq
 301 kelhthesdr scihlefdis rtgityetgd hvgvyaenhv eiveeagkll ghsldlvfsi
 361 hadkedgspl esavppfpfg pctltgtglar yadllnpprk salvalaaya tepseaeclk
 421 hltspdgkde ysqwivasqr sllevmaafp sakpplgvff aaiaprlqpr ysysssprl
 481 apsrhvhtsa lvygptptgr ihkgvcstwm knavpaeksh ecsgapifir asnfklpsnp
 541 stpivmvgpg tglapfrgl qermalkedg eelgssllff gcrnrgmdfi yedelnnfvd
 601 qgviselima fsregaqkey vqhkmmeaa qvwdlikeeg ylyvcgdakg mardvhrtlh
 661 tivqeqegvs sseaeaiavkk lqtegrylrd vw (**Arabidopsis thaliana** **CPR polypeptide encoded by**
ATR1)

1 msssssst midlmaalik gepvivsdpa nasayesvaa elssmlienr qfamivttsi
 61 avligciavl vwrsgsgns krveplkplv ikpreeeidd grkkvtiffg tqtgtaegfa
 121 kalgeeakar yektrfkivd lddyaaddde yeeklkedv affflatygd geptdnaarf
 181 ykwftegndr gewlknlyg vfglgnrqye hfnkvakvvd dilveqgaqr lvqvlgdddd
 241 qcieddftaw realwpeldt ilreeegdtav atpytaavle yrvsihdsed akfnditlan
 301 gngyvtvfdaq hpykanvavk relhtpesdr scihlefdia gsgltmklgd hvglcdnls
 361 etvdealrll dmspdtyfsl haekedgtpi ssslpppfpp cnlrtaltary acllsspkks
 421 alvalaahas dpteaerlkh laspagkdey skwvvesqrs llevmaefps akpplgvffa
 481 gvaprlqprf ysissspkia etrihvtcal vyeemptgri hkgvcstwmk navpyeksek
 541 lflgrpifvr qsnfkllpsds kvpiimigpg tglapfrgfl qerlalvesg velgpsvlf
 601 gcrnrrmdfi yeeelqrfe sgalaelsva fsregptkey vqhkmmdkas diwnmisqga
 661 ylyvcgdakg mardvhrslh tiaqeqgsmd stkaegfvkn lqtsgrylrd vw (**A. thaliana CPR**
polypeptide encoded by ATR2.)

>Stevia rebaudiana CPR7
 MQSESVEASTIDLMTAVLKDTVIDTANASDNGDSKMPPALAMMFEIRDLLLLLTTSVAVL
 VGCFFVLWKRSSGKSGKELEPPKIVPKRRLEQEVDDGKKKVTIFFGTGTGAEQFAK
 ALFEEAKARYEKAFFVIDLDDYAADLDEYAEEKLKETYAFFFLATYGDGEPTDNAAKFY
 KWFTEGDEKGVWLQKLYGVFGLGNRQYEHFNKIGIVVDDGLTEQGAKRIVPVGLGDDDDQ
 SIEDDFSARKELVWPELDLLRDEDDKAAATPYTAAIPEYRVVFHDKPDADFDDHTQTNG
 HAVHDAQHPCRSNVAVKKELHTPESDRSCTHLEFFDISHTGLSYETGDHVGVCENLIEVV
 EEAGKLLGLSTDYFSLHIDNEDGSLGGPSLQPPFPFCTLRKALTNVADLLSSPKKSTL
 LALAAHASDPTAADRLRFLASREGKDEYAEWVANQRSLLLEVMEAFPSARPPGLGVFFAAV
 APRLQPRYYSSISSPKMEPNRIHVTICALVYEKTPAGRIHKGICSTWMKNAVPLTESQDCS
 WAPIFVRTSNFRLPIDPKVPVIMIGPTGLAPFRGLQERLALKESGTELGSSILFFGCR
 NRKVDYIYENELNFFVENGALSELVDVAFSRDGPPTKEYVQHKMTQKASEIWNMLSEGAYLY
 VCGDAKGMKADVHRTLHTIVQEQGLDSSKAELVVKNLQMSGRYL RDVW

> Stevia rebaudiana CPR8
 MQSNSVKISPLDLVTALFSGKVLDTSNASESGESAMLP TIAMINENRELLMILTTSVAVL
 IGCVVVLVWRRSSTKKSALPPVIVVPKRVQEEVDDGKKKVTVFFGTGTGAEQFAKAL

VEEAKARYEKAVFKVIDLDDYAADDDYEYEEKLKKESLAFFFLATYGDGEPTDNAARFYKW
FTEGDAKGEWLNKLQYGVFGLGNRQYEHFNKIAKVVDGGLVEQGAKRLLVPVGLGDDDDQCI
EDDFTAWKELVWPELDQLLRDEDDTTVATPYTAAVAEYRVVFHEKPDALSEDYSYTNNGHA
VHDAQHPCRSNVAVKKELHSPESDRSCTHLEFDISNTGLSYETGDHVGVCENLSEVND
AERLVGLPPDTYSSIHDTSEDGSPGLGASLPPPPCTLRKALTCYADVLSLSPKKSALLA
LAAHATDPSEADRLKFLASPAGKDEYSQWIVASQRSLLLEVMEAFPSAKPSLGVFFASVAP
RLQPRYYSISSSPKMAPDRIHVTICALVYEKTPAGRIHKGVCSTWMKNAVPMTESQDCSWA
PIYVRTSNFRLPSDPKVPVIMIGPGTGLAPFRGFLQERLALKEAGTDGLSILFFGCRNR
KVDFIYENELNNFVETGALSELIVAFSREGPTKEYVQHKMSEKASDIWNLLSEGAYLYVC

FIGURE 13B

>ATR1 codon optimized by DNA2.0.
ATGACTTCTGCACCTTTATGCCCTCCGATCTTTTCAAAACAATTGAAAAGTATCATGGGAACGGATTC
TTTGTCCGATGATGTTGTATTAGTTATTGCTACAACTTCTCTGGCACTGGTTGCTGGTTTCGTTGT
CTTATTGTGAAAAAGACCACGGCAGATCGTTCCGGCAGCTAAAGCCACTAATGATCCCTAAG
TCTCTGATGCGGAAAGATGAGGATGATGACTTAGATCTAGGTTCTGGAAAAACGAGAGTCTCT
ATCTTCTTCGGCACACAAACCGGAACAGCCGAAGGATTGCTAAAGCACTTTCAGAAGAGATCA
AAGCAAGATACGAAAAGCGGCTGTAAAAGTAATCGATTGATGATTACGCTGCCGATGATG
ACCAATATGAGGAAAAGTTGAAAAAGGAAACATTGGCTTCTTTTGTGTAGCCACGTAATGGTGA
TGGTGAACCAACCGATAAACGCCGCAAGATTCTACAAGTGGTTTACTGAAGAGAACGAAAAGAGA
TATCAAGTTGCAGCAACTTGCTTACGGCGTTTTTCCTTAGGTAACAGACAAATACGAGCACCTTA
ACAAAGATAGGTATTGCTTAGATGAAGAGTTATGCAAAAAGGTGCGAAGAGATTGATTGAAG
TCGGTTTAGGAGATGATGATCAATCTATCGAGGATGACTTTAATGCAATGGAAAGGAATCTTTGTG
GTCTGAATTAGATAAGTTACTTAAGGACGAAGATGATAAATCCGTTGCCACTCCATACACAGCC
GTCAATCCAGAAATATAGAGTAGTTACTCATGATCCAAGATTCAACAACACAGAAATCAATGGAAA
GTAAATGTGGCTAATGGTAATACTACCATCGATATTCATCATCCATGTAGAGTAGACGTTGCAGTT
CAAAAAGGAATTGCACACTCATGAATCAGACAGATCTTGCCATACATCTTGAATTTGATATATCACG
TACTGGTATCACTTACGAAACAGGTGATCACGTGGTGCTACGCTGAAAAACCATGTTGAAAATT
GTAGAGGAAGCTGGAAGTTGTTGGGCCATAGTTTAGATCTTGTCTTCTCAATTCAATGCCGATA
AAGAGGATGGCTCACCACTAGAAAGTGCAGTGCCCTCCACCATTTCAGGACCATGCACCCTAGG
TACCGGTTTAGCTCGTTACGCGGATCTGTTAAATCCTCCACGTAAATCAGCTCTAGTGGCCTTGG
CTGCGTACGCCACAGAACCTTCTGAGGCAGAAAAACTGAACATCTAACTTCAACCATGGTAA
GGATGAATACTCACAATGGATAGTAGCTAGTCAACGTTCTTTACTAGAAGTTATGGCTGCTTTCC
CATCCGCTAAACCTCCTTTGGGTGTTTCTTCGCCGCAATAGCGCCTAGACTGCACCAACCAAGATAC
TATTCAATTTCACTCCTACCTAGACTGGCAACCATCAAGAGTTCAATGTCACATCCGCTTTAGTGTA
GGTCCAACTCCTACTGTTAGAAATCCATAAGGGCGTTTGTTCAAACATGATGAAAAACGCGGTTT
CAGCAGAGAAAGTCTACGAAATGTTCTGGTGCTCCAATCTTTATCAGAGCCTCCAACCTTCAAACTG
CCTTCCAAATCCTTCTACTCCTATTGTCAATGGTCCGTCCTGGTACAGGTCCTTGTCTCCATTCAGAGGT
TTCTTACAAGAGAGAATGGCCTTAAAAGGAGGATGGTGAAGAGTTGGGATCTTCTTTGTTGTTTT
TCGGCTGTAGAAAACAGACAAAATGGATTTCATCTACGAAGATGAACCTGAATAACTTTGTAGATCA
AGGAGTTATTCAGAGTTGATAATGGCTTTTCTAGAGAAGGTGCTCAGAAAGGAGTACGTCCAA

CACAAAATGATGGAAGGCGCACAAAGTTTGGGACTTAATCAAAGAGGAAGGCTATCTATAT
GTCTGTGGTGATGCAAAAGGTATGGCAAGAGATGTTACAGAAACACTTCATATACTAGTCCAG
GAACAGGAAGGCGTTAGTTCCTCTGAAGCGGAAGCAATTGTGAAAAAGTTACAAACAGAGGGA
AGATACTTGAGAGATGTGTGGTAA

>ATR2 codon optimized by Genscript
ATGCTTTCCTCTTCCTCTTCCAGTACCTCTATGATTGATTGGCTGCTATTATTAAAGGTGAA
CCAGTTATCGTCTCCGACCCAGCAAAATGCCCTCTGCTTATGAATCAGTTGCTGCAGAAATTGCTCTC
AATGTTGATCGAAAAACAGACAAATTCGCCATGATCGTAACTACATCAATCGCTGTTTGTGATCGGTT
GTATTGTCATGTTGGTATGGAGAAGATCCGGTAGTGGTAATCTAAAAGAGTCGAAACCTTTGAA
ACCATTAGTAATTAAGCCAAGAGAAGAAAGAAATAGATGACGGTAGAAAAGAAAGTTACAATATT
TTTCGGTACCCAAACTGTTACAGCTGAAGGTTTTGCAAAAGCCTTAGGTGAAGAAGCTAAGGC
AAGATACGAAAAGACTAGATTCAAGATAGTCGATTGGATGACTATGCCCGCTGATGACGATGA
ATACGAAGAAAAGTTGAAAGAAAAGATGTTGCATTTTCTTTTGGCAACCTATGGTGACGGT
GAACCAACTGACAAATGCAGCCAGATTCTACAAATGGTTTACAGAGGTAATGATCGTGGTGAAT
GGTTGAAAAAATTAAGTACGGTGTTCGGTTTGGGTAAACAGACAAATACGAACATTTCAACAA
AGTTGCAAAAGGTTGTCGACGATATTTTGGTCGAACAAGGTGCTCAAAAGATTAGTCCAAGTAGGT
TTGGGTGACGATGACCAATGTATAGAAGATGACTTTACTGCCCTGGAGAGAAGCTTTGTGGCCTG
AATTAGACACAATCTTGAGAGAGAAGAGGTGACACCGCCGTTGCTACCCCATATACTGCTGCAGT
ATTAGAATACAGAGTTTCATCCATGATAGTGAAGACGCAAAAGTTTAATGATATCACCTTTGGCC
AATGGTAACGGTTATACAGTTTTCGATGCACAACACCCCTTACAAGCTAACGTTGCAGTCAAGA
GAGAAATTACATACACCAGAAATCCGACAGAAAGTTGTATACACTTGGAATTTGATATCGCTGGTTC
CGGTTTAACCATGAAGTTGGGTGACCATGTAGGTGTTTTATGCGACAAATTTGCTGAAAACGTGT
GATGAAGCATTGAGATTGTTGGATATGTCCCCTGACACTTATTTAGTTTGCACGCTGAAAAAG
AAGATGGTACACCAATTTCCAGTTCTTTACCACCTCCAATCCCTCCATGTAACCTAAGAACAGCCT
TGACCAAGATACGCTTGCTTGTTATCATCCCTAAAAAGTCCGCCCTTGTTGCTTTAGCCGCTCAT
GCTAGTGATCCTACTGAAGCAGAAAAGATTGAAACACTTAGCATCTCCAGCCGGTAAAGATGAAT
ATTCAAAGTGGGTAGTTGAATCTCAAAGATCATTTGTTAGAAGTTATGGCAGAAATTTCCATCTGCC
AAGCCCTCCATTAGGTGCTTCTTTGCTGGTGTAGCACCTTAGATTGCAACCAAGATTCTACTCAAT
CAGTTCTTCACCTAAGATCGCTGAAACTAGAATTCAATGTTACATGTGCATTAGTCTACGAAAAAGA
TGCCAAACCCGGTAGAAATTCACAAAGGGTGTATGCTCTACTTGGATGAAAAAATGCTGTTCTCTTACGA

AAAAACAGAAAAAGTTGTTCTTAGGTAGACCAATCTTCGTAAGACAAATCAAACTTCAAGTTGCCTT
CTGATTCAAAGGTTCCAAATAATCATGATAGTCTCTGGTACAGGTTTAGCCCCATTACAGAGTTTC
TTGCAAGAAAGATTGGCTTTAGTTGAATCTGGTGTGAAATTAGTCCCTTCAGTTTGTGTTCTTTGG
TTGTAGAAAACAGAAAGAAATGGATTTTCATCTATGAAGAAGAAATTGCAAAGATTCTGTCGAATCTGGT
GCATTGGCCGAATTATCTGTAGCTTTTCAAGAGAAAGGTCCTCAACTAAGGAATACGTTCAACATA
AGATGATGGATAAGGCATCCGACATATGGAACATGATCAGTCAAGTGCTTATTGTACGTTTG
CGGTACGCCAAAGGATATGGCCAGAGATGTCCATAGATCTTTGCACACAATTGCTCAAGAACAA
AGGTTCCATGGATAGTACCAAGCTGAAGGTTTCGTAAGAAGCTTACAAACTTCCGGTAGATAC
TTGAGAGATGTCCTGGTGA

>Stevia rebaudiana CPR7 (CPR cloned from Stevia rebaudiana cDNA)
ATGCAATCGGAATCCGTTGAAGCATCGACGATTGATTTGATGACTGCTGTTTGAAGGACACAG
TGATCGATACAGCGAACGCATCTGATAACGGAGACTCAAAGATGCCGCCGGCTTGGCGATGA
TGTTCCGAAATTCCGTGATCTGTTGCTGATTTTGACTACGTACGTGCTGTTTGGTCGGATGTTTC
GTTGTTTTGGTGTGGAAGAGATCGTCCGGGAAGAAAGTCCGGCAAGGAATTGGAGCCGCCGAA
GATCGTTGTGCCGAAGAGGCGGTGGAGCAGGAGTTGATGATGGTAAGAAGAGGTTACGA
TTTTCTTCGGAACACAAAACCTGGAACGGCTGAAGGTTTCGCTAAGGCACCTTTCGAAAGAAGCGAA
AGCGCGATATGAAAAGGCAGCGTTTAAAGTGATTGATTTGGATGATTATGCTGCTGATTTGGAT
GAGTATGCAGAGAAGCTGAAGAAGGAACATATGCTTTCTTCTTGGCTACATATGGAGATG
GTGAGCCAACTGATAATGCTGCCAAAATTTTATAAATGGTTTACTGAGGGAGACGAGAAAGCG
TTTGGCTTCAAAAACCTTCAATATGGAGTATTGGTCTTGGCAACAGACAATATGAACATTTCAAC
AAGATTGGAATAGTGGTTGATGATGGTCTCACCGAGCAGGGTGCAAAACGCAATTGTTCCCGTT
GGTCTTGAGACGACGATCAATCAATTGAAGACGATTTTCGGCATGGAAGAGTTAGTGTTGG
CCCGAATTGGATCTATTGCTTCGCGATGAAGATGACAAAAGCTGCTGCAACTCCTTACACAGCTG
CAATCCCTGAATACCGCGTCGTATTTTCATGACAAAACCCGATGCCGTTTCTGTGATGATCATACTCAA
ACCAATGGTCATGCTGTTCATGATGCTCAACATCCATGCAGATCCAAATGTGGCTGTAAAAAAG
AGCTTCATACTCCGAAATCCGATCGTTCATGCACACATCTTGAATTTGACATTTCTCACACTGGAT
TATCTTATGAAACTGGGGATCATGTTGGTGATACTGTGAAAAACCTAATTGAAGTAGTGGAGA
AGCTGGGAAATTTAGGATTATCAACAGATACTTATTTCTCGTTACATATTGATAACGAAGATG
GTTCAACCACTTGGTGACCTTCATTACAACCTCCTTTTCCCTCCTTGTACTTTAAGAAAAAGCATTGA
CTAATTATGCAGATCTGTTAAGCTCTCCCCAAAAAGTCAACTTTTGCTTGTCTCTAGCTGCTCATGCTT

CCGATCCCACTGAAGCTGATCGTTTAAAGATTTCCTTGCACTCTCGCGAGGGCAAGGATGAATATGC
 TGAATGGGTGTTGCAAAACCAAGAGTCTTCTTGAAGTCATGGAAGCTTTCCTCGTCAGCTAGA
 CCGCCACTTGGTGTCTTTTGCAGCGGTTGCACCGGTTTACAGCCTCGTTACTACTCTATTTCT
 TCCTCCCCAAAGATGGAAACCAACAGGATTTCATGTTACTTGCGCGTTGGTTTATGAAAAAATCC
 CGCAGTCTGATCCACAAAGGAATCTGCTCAACCTGGATGAAGAACGCTGTACCTTTGACCGAA
 AGTCAAGATTGCAGTTGGCACCGATTCTTTGTTAGAACATCAAACTTCAGACTTCCAATTGACCC
 GAAAGTCCCGGTTATCATGATTGGTCTTGGAAACCGGTTGGCTCCATTAGGGGTTTTCCTTCAA
 GAAAGATTGGCTCTTAAAGAAATCCGGAACCGAACTCGGGTCATCTATTTATTTCTTCGGTTGTAG
 AAACCGCAAAGTGGATTACATATATGAGAAATGAACTCAACAACTTTGTTGAAAAATGGTGGCTT
 TCTGAGCTTGATGTTGCTTCTCCCGGATGGCCCGACGAAAGAAATACGTGCAACATAAAAATGA
 CCCAAAGGCTTCTGAAAATATGGAATATGCTTTCTGAGGGAGCATATTTATATGTATGTGTGA
 TGCTAAAGGCATGGCTAAAGATGTACACCGTACACTTCACACCATTTGTCAAGAACAGGGAAGT
 TTGGACTCGTCTAAAGCGGAGTTGTATGTGAAGAAATCTACAAATGTCAGGAAGATACCTCCGTG
 ATGTTTGGTAA

> Stevia rebaudiana CPR8 (CPR cloned from Stevia rebaudiana cDNA)
 ATGCAATCTAACTCCGTGAAGATTTCGCCGCTTGATCTGGTAACCTGCGTGTCTTAGCGCAAGG
 TTTTGGACACATCGAACGCATCGGAATCGGGAGAACTCTGCTATGCTGCCGACTATAGCGATGAT
 TATGGAGAAATCGTGAGCTGTTGATGATACCTCAAAACGTCGGTTGCTGTATTGATCGGATGCGTT
 GTCGTTTTGGTGTGGCGGAGATCGTCTACGAAGAAAGTCGGCTTGGAGCCACCGGTGATTGTG
 GTTCCGAAGAGAGTGCAAGAGGAGGAAGTTGATGATGGTAAGAAAGTTACGGTTTTCTTC
 GGCACCCAAACTGGAACACAGCTGAAGGCTTCGCTAAGGCACTTGTGAGGAAGCTAAAGCTCGA
 TATGAAAAGGCTGCTTTAAAGTAATTGATTGGATGATTATGCTGCTGATGACGATGAGTATG
 AGGAGAAACTAAAGAAAGATCTTTGGCCTTTTCTTTTGGCTACGTATGGAGATGGTGAGCC
 AACAGATAATGCTGCCAGATTATATAAATGGTTTACTGAGGGAATGCGAAAAGGAGAATGGCT
 TAATAAGCTTCAATATGGAGTATTTGGTTTGGTAACAGACAAATATGAACATTTTAACAAGATC
 GCAAAAGTGGTTGATGATGGTCTTGTAGAACAGGTGCAAGCGTCTTGTTCCTGTTGGACTTG
 GAGATGATGATCAATGTATTGAAGATGACTTCACCGCATGGAAAGAGTTAGTATGCCCGGAGT
 TGGATCAATTACTTCGTGATGAGGATGACACAACGTGTTGCTACTCCATACACAGCTGCTGTTGCA
 GAATATCGCGTTGTTTTTCATGAAAAAACCCAGACCGGCTTTTCTGAAGATTATAGTTATACAAATGG

CCATGCTGTTCAATGCTCAACATCCATGCAGATCCAAACGTGGCTGTCAAAAAGGAACCTTCAT
AGTCCCTGAATCTGACCGGTCTTGCACTCATCTTGAATTTGACATCTCGAACACCCGGACTATCATA
TGAAACTGGGACCATGTTGGAGTTTACTGTGAAAACCTTGAGTGAAGTTGTGAATGATGCTGA
AAGATTAGTAGGATTACCAACAGACACTTACTCCTCCATCCACACTGATAGTGAAGACGGGTCCG
CCACTTGGCGGAGCCTCATTTGCCGCCCTCCTTCCCGCCATGCACCTTTAAGGAAAGCATTGACGTG
TTATGCTGATGTTTGTAGTTCTCCCAAGAAAGTCGGCTTTGCTTGCACTAGCTGCTCATGCCACCG
ATCCCCAGTGAAGCTGATAGATTGAAAATTTC TTGCATCCCCCGCGAAAGGATGAATATTCTCA
ATGGATAGTTGCAAGCCAAAGAAGTCTCCTTGAAGTCATGGAAAGCATTCCCGTCAGCTAAGCCT
TCACTTGGTGTCTTCTTTGCACTCTGTTGCCCGCGCTTACAACCAAGATACTACTCTATTCTTCCT
CACCCAAAGATGGCACCGGATAGGATTTCATGTTACATGTGCATTAGTCTATGAGAAAACACCTGC
AGGCCGCATCCACAAAGGAGTTTGTTCAACTTGGATGAAGAAGCAGTGCCCTATGACCCGAGAG
TCAAGATTGCAGTTGGGCCCCAATATACGTCCGAACATCCAATTTCAGACTACCATCTGACCCCTA
AGTCCCCGGTTATCATGATTGGACCTGGCACTGGTTTGGCTCCTTTTAGAGGTTTCCTTCAAGAG
CGGTTAGCTTTAAAAGGAAGCCGGAACCTGACCTCGGTTTATCCATTTTATTCTTCGGATGTAGGAA
TCGCAAAAGTGGATTTCATATATGAAAACGAGCTTAACAACCTTGTGGAGACTGGTGCTCTTTCTG
AGCTTATTGTTGCTTTCTCCCGTGAAAGGCCCGACTAAGGAATATGTGCAACACAAAGATGAGTGA
GAAGGCTTCGGATATCTGGAACCTTGCTTTCTGGAAGGAGCATATTTATACGTATGTGGTGATGCC
AAAGGCATGGCCAAAGATGTACATCGAACCCCTCCACACAATTGTGCAAGAACAGGGATCTCTTG
ACTCGTCAAAGGCAGAACTCTACGTGAAGAAATCTACAAAATGTCAAGGAAGATACCTCCGTGACGT
TTGGTAA

FIGURE 14

A.	atggttttgt	cttcttcttg	tactacagta	ccacacttat	cttcattagc	tgtcgtgcaa	60
	cttggtcctt	ggagcagtag	gattaaaaag	aaaaccgata	ctgttcagc	accagccgct	120
	gcaggaagg	gcaggaagg	cttggctaga	gcacagcaca	catcagaatc	cgcagctgtc	180
	gcaaaaggga	gcagtttgac	ccctatagtg	agaactgacg	ctgagtcacg	gagaacaaga	240
	tggccaaaccg	atgacgatga	cgcggaacct	ttagtggatg	agatcagggc	aatgcttact	300
	tccatgtctg	atggtgacat	ttccgtgagc	gcatacgata	cagcctgggt	cggattggtt	360
	ccaagattag	acggcgggtga	aggtcctcaa	tttccagcag	ctgtgagatg	gataagaaat	420
	aaccagttgc	ctgacggaag	ttggggcgat	gccgcattat	tctctgccta	tgacaggctt	480
	atcaataccc	ttgcctgcgt	tgtaactttg	acaagggtgt	cctagaacc	agagatgaga	540
	ggtagaggac	tatctttttt	gggtagggaac	atgtggaaat	tagcaactga	agatgaagag	600
	tcaatgccta	ttggcttcga	attagcattt	ccatctttga	tagagcttgc	taagagccta	660
	ggtgtccatg	acttccctta	tgatcaccag	gccctacaag	gaatctactc	ttcaagagag	720
	atcaaaaatga	agaggatcc	aaaagaagtg	atgcataccg	ttccaacatc	aatattgcac	780
	agtttggagg	gtatgcctgg	cctagattgg	gctaaactac	ttaaactaca	gagcagcgac	840
	ggaagttttt	tgttctcacc	agctgccact	gcataatgctt	taatgaatac	cggagatgac	900
	aggtgtttta	gctacatcga	tagaacagta	aagaaattca	acggcggcgt	cctaattgtt	960
	tatccagtg	atctatttga	acataatttg	gccgttgata	gacttgaaag	attaggaatc	1020
	tccaggtact	tccaaaagga	gatcgaacaa	tgcattggatt	atgtaaacag	gcattggact	1080
	gaggacggta	tttggtgggc	aaggaaactct	gatgtcaaa	aggtggacga	cacagctatg	1140
	gcctttagac	ttcttaggtt	gcacggctac	agcgtcagtc	ctgatgtgtt	taaaaacttc	1200
	gaaaaggacg	gtgaattttt	cgcattttgtc	ggacagtcta	atcaagctgt	taccggtatg	1260
	tacaacttaa	acagagcaag	ccagatatcc	ttcccaggcg	aggatgtgct	tcataagagct	1320
	ggtgccttct	catatgagtt	cttgaggaga	aaagaagcag	agggagcctt	gagggacaag	1380
	tggatcattt	ctaaagatct	acctgggtgaa	gttgtgtata	ctttggattt	tccatggtac	1440
	ggcaacttac	ctagagtoga	ggccagagac	tacctagagc	aatacggagg	tggtgatgac	1500
	gtttggattg	gcaagacatt	gtataggatg	ccacttgtaa	acaatgatgt	atatattggaa	1560
	ttgggcaagaa	tggattttcaa	ccactgccag	gctttgcatc	agttagagtg	gcaaggacta	1620
	aaaagatggt	atactgaaaa	taggttgatg	gactttgggtg	tcgccccaga	agatgccctt	1680

agagccttatt ttcttgccgc cgcattctgtt tacgagccctt gtagagctgc cgagaggctt 1740
 gcatgggcta gagccgcaat actagctaac gccgtgagca cccacttaag aaatagccca 1800
 tcattcagag aaaggttaga gcattctctt aggtgtagac ctagtgaaga gacagatggc 1860
 tcctggttta actcctcaag tggctctgat gcagttttag taaaggctgt cttaagactt 1920
 actgattcat tagccaggga agcacagcca atccatggag gtgacccaga agatattata 1980
 cacaagtgtg taagatctgc ttgggccgag tgggttaggg aaaaggcaga cgctgccgat 2040
 agcgtgtgca atggtagttc tgcagtagaa caagagggat caagaatggg ccatgataaa 2100
 cagacctgtc tattattggc tagaatgac gaaatttctg ccggtagggc agctggtgaa 2160
 gcagccagtg aggacggcga tagaagaata attcaattaa caggctccat ctgcgacagt 2220
 cttaagcaaa aatgcttagt ttcacaggac cctgaaaaaa atgaagagat gatgtctcac 2280
 gtggatgacg aattgaagtt gaggattaga gaggctcttc aatatctgct tagactaggt 2340
 gaaaaaaga ctggatctag cgaaccagg caaacatttt taagtatagt gaaatcatgt 2400
 tactatgctg ctcattgccc acctcatgtc gttgatagac acattagtag agtgattttc 2460
 gagccagtaa gtgccgcaaa gtaaccggcg 2490

B.

Met	Val	Leu	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	Thr
		20						25				30		
Asp	Thr	Val	Ala	Val	Pro	Ala	Ala	Ala	Gly	Arg	Trp	Arg	Arg	Ala
		35					40					45		
Ala	Arg	Ala	Gln	His	Thr	Ser	Glu	Ser	Ala	Ala	Val	Ala	Lys	Gly
	50					55					60			
Ser	Leu	Thr	Pro	Ile	Val	Arg	Thr	Asp	Ala	Glu	Ser	Arg	Arg	Thr
65				70						75				80
Trp	Pro	Thr	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
		100						105						Tyr
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	175
Pro Glu Met Arg Gly Arg Gly Leu Ser Phe Leu Gly Arg Asn Met Trp		
180	185	190
Lys Leu Ala Thr Glu Asp Glu Glu Ser Met Pro Ile Gly Phe Glu Leu		
195	200	205
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 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	255
Ser Ile Leu His Ser Leu Leu Glu Gly Met Pro Gly Leu Asp Trp Ala Lys		
260	265	270
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	335
Arg Leu Gly Ile Ser Arg Tyr Phe Gln Lys Glu Ile Glu Gln Cys Met		
340	345	350
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						375					380				
Leu	Arg	Leu	His	Gly	Tyr	Ser	Val	Ser	Pro	Asp	Val	Phe	Lys	Asn	Phe
385					390					395					400
Glu	Lys	Asp	Gly	Glu	Phe	Ala	Phe	Val	Val	Gly	Gln	Ser	Asn	Gln	Ala
				405					410					415	
Val	Thr	Gly	Met	Tyr	Asn	Leu	Asn	Arg	Ala	Ser	Gln	Ile	Ser	Phe	Pro
			420				425					430			
Gly	Glu	Asp	Val	Leu	His	Arg	Ala	Gly	Ala	Phe	Ser	Tyr	Glu	Phe	Leu
		435				440						445			
Arg	Arg	Lys	Glu	Ala	Glu	Gly	Ala	Leu	Arg	Asp	Lys	Trp	Ile	Ile	Ser
450						455					460				
Lys	Asp	Leu	Pro	Gly	Glu	Val	Val	Tyr	Thr	Leu	Asp	Phe	Pro	Trp	Tyr
465					470					475					480
Gly	Asn	Leu	Pro	Arg	Val	Glu	Ala	Arg	Asp	Tyr	Leu	Glu	Gln	Tyr	Gly
				485					490					495	
Gly	Gly	Asp	Asp	Val	Trp	Ile	Gly	Lys	Thr	Leu	Tyr	Arg	Met	Pro	Leu
			500					505					510		
Val	Asn	Asn	Asp	Val	Tyr	Leu	Glu	Leu	Ala	Arg	Met	Asp	Phe	Asn	His
		515					520					525			
Cys	Gln	Ala	Leu	His	Gln	Leu	Glu	Trp	Gln	Gly	Leu	Lys	Arg	Trp	Tyr
530						535					540				
Thr	Glu	Asn	Arg	Leu	Met	Asp	Phe	Gly	Val	Ala	Gln	Glu	Asp	Ala	Leu
545					550					555					560
Arg	Ala	Tyr	Phe	Leu	Ala	Ala	Ala	Ser	Val	Tyr	Glu	Pro	Cys	Arg	Ala
				565					570					575	
Ala	Glu	Arg	Leu	Ala	Trp	Ala	Arg	Ala	Ala	Ile	Leu	Ala	Asn	Ala	Val
			580					585					590		
Ser	Thr	His	Leu	Arg	Asn	Ser	Pro	Ser	Phe	Arg	Glu	Arg	Leu	Glu	His
							600						605		
		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
							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	Lys	Gln	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
							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
									810					815	
Arg	Val	Ile	Phe	Glu	Pro	Val	Ser	Ala	Ala	Lys					
								820							

FIGURE 15

A.

ATGCCCTGGTAAAAATTGAAATGGTACCCCAAGGACCTCAAGACTGGAAATGATTTTGTCTCTGCTAAGAGTTTACTAGATCGAGC
 TTTCAAAAGTCATCATCTCTACTACGGATTATGCTCAACTTCATGTCAAGTTTATGATACAGCTTGGTTGCAATGATTCCAAAAACAA
 GAGATAATGTAAAAACAGTGGTTGTTCCAGAAATGTTTCCATTACCTCTTAAAAACACAAGCCGAGATGGCTCATGGGGTTTCATTGGCCT
 ACAACACAGACAGCGGGTATCCTAGATACAGCCTCAGCTGTGGCATTATTGTGCCACGCACAAGAGCCTTTACAAAATATTGGATGT
 ATCTCCAGATGAAATGGGTTGAGAAATAGAACACGCGTGCACATCCTTGAAACGTCATATTAGCAGTTTGGAAATGATGTGGAGGACACCA
 ACCATATTGGCGTCGAGTTTATCATACACGCTTACTTTCCATGCTAGAAAAAGGAATTAGATGTTCCATCTTTTGAATTTCCATGTAGG
 TCCATCTTAGAGAGAAATGCACGGGAGAAAATTAGGTCAATTTCGACCTGGAAACAAGTTTACGGCAAGCCAAAGCTCATTTGTTGCACCTCAT
 GGAAGCATTTCTCGGTAAAGCTAGATTTTGATCGACTATCACATCACTATACACGCGCATGATGGCATCTCCATCTTCAACGGCTG
 CTATATCTTATTTGGGCTACAAAAATGGGATGACGAAAGCCGAAGATTAACCTAAGACATGTAATGGTGCAGGACATGGGAAATGGA
 GGTATTTCTGGTACATTTCCAACTACTCATTTTCGAATGTAGCTGGATTATAGCAACGTTGTTAAAGGTTGGCTTTACTTTGAAGCAAAAT
 TGACGGCGATGGCTTAAGAGGTTTATCAACCATCTTACTTGAGCGCTTTCGTGATGAGAAATGGTGTATAGGCTTTGCCCTAGAACAG
 CAGATGTAGATGACACAGCCAAAGCTCTATTGGCCTTGTCATTGGTAAACCAAGCCAGTGCACCTGATATCATGATTAAGGTCCTTTGAG
 GGCAAGACCAATTTTACCACCTTTTGGTTCAGAAAGAGATCCATCATTTGACTTCCAAACCTGCACGTCCTTTTATCTTTACTTAAACAATC
 TAACCTTGCTCAATACCATCCTCAAAATCCTCAAAACAACATTAATCACTTGATAGTGGTGGGTTCCGATCATTTGTGTCAAAAGACA
 AATGGAATTTGAGTCACCTATATCCAACTATGTTTGTGTTGAAGCCTTCACTGAAGTGCCTCCATCTCATTGACGGTGGTGAATTTGTCT
 AGTCTGTTTGATGAATCCTTTAAGTGAAGATTGGTCTTAGCATCTTTCAAGCGGTACTTAGAAATAATCCTCACCCCAAGACAACGACGG
 CTCTTGGAGAGGATACAGAGAAACAGACGTGTTACGCAATATTGGCTTTAGTTCGAAGGAGACATGTATGCTTTTCACTCACATGTTG
 ACAGACTGCAATCATGTGTTGATCGAGGTTTCTCATGGTTGAAATCTTGCTCTTTTTCATTCTCAAGACCTGACTTGGACCTCTAAACA
 GCTTATGAAGTGGGTTTTCGTAGCTGAAGCATATAAACTAGCTGCTTTACAACTCTGCTTCCCTGGAGGTTCTGCTGCCACCATTTGGACA
 TTCTGTCAAGTGTCCATCAAGTGAATCTTGAAAAATACATGAGATTGGTGAGAAAAAAGTGCCTTATTCTCTCCACTGGATGAGT
 GGGGCTAATGGCTTCTATCATCGAATCTTCATTTTTCGTACCATTAATGTCAGGCACAAAAGAGTTGAAATATACCTTAGAGATAATATC
 AAGGTGGACGAAGATAAGTACTTGTCTATTATCCCATTCACATGGGTCGGATGCAATAATAGGTCTAGAACTTTCGCAAGTAACAGATG
 GCTATACGATATGATGTACCTTTCATTAATCTCGGCTATCAAAACCGACGAGTACATGGAAGCTGTAGCTGGCCAGTGTGTTGGGGATGTTT
 CCTTGTACATCAAAACAATTGATAAGGTGATTGATAACAAATGGGTAACTTTCGAGAGCCAAATGGAACAGTACACAGTGGTAATGGA
 CATCAGCACGAATCTCCTAATAATAGGTCAAGTCGAGGACACCTTGACTCGTTTACAAAAATTCAGTCTTGAATCACAAAGACGTCCTTAA
 CTCTAGCTCATCTGATCAAGATACCTTTGAGAAGAGAGGTTTAGAACATTCATGCACGCTCATATACACAAAATCGAAGATAAATCAACGAT
 TCAGTAAGCAAGCCTCATCCGATGCGTTTTTCCCTCCTCGAACAAATCTTACTTTCAATGGGTGAACCTCAACTGGTGGCTCACATGTCGCT
 TGCGGCTATTCAATTTGGCTTCTCTAAATTGCGCTCATGTCTGCAAAATTTGTTGCAGGGTAAAGACGCTTTTCCAAAGCGGAACGCAAAAGTA

CTTAATCTCCTCTGTTATGAGACATGCCACAAACATGTGTAGAAATGTATAACGACTTTGGCTCTATTGCCAGAGACAAACGCTGAGAGAA
 ATGTTAATAGTATTCTCTGAGTTTACTCTCTGTACGGAACCTCTCAAAACCTAGATGAAAGGAAAGAAAGACTTCTGAAAATC
 GCAACTTACGAACAAGGTATTGATAGAGCACCTAGAGGCCTTGGAAGACAGAGTAGAGATGATGCCGGAGACAGAGCTGGATCTAA
 AGATATGAGAAAAGTTGAAAAATCGTTAAGTTATTCTGTGATGTTACGGACTTATACGATCAGCTCTACGTTATCAAAAGATTGTGTCATCCT
 CTATGAAGTAA

B.

MPGKIENGTPKDLKTGNDEVSAKSLLDRAFKSHHSYYGLCSTSCQVYDTAWVAMIPKTR
 DNVKQWLFPECFHYLLKTQAADGSWGLPTTQTAGILDTASAVLALLCHAQEPQLILDVS
 PDEMGLRIEHGVTSLKRQLAVWNDVEDTNHIGVEFIIPALLSMEKELDVPSFEPPCRSI
 LERMHGEKLGHFDFLEQVYGKPSLLHSLEAFLGKLDLDRSLSHHLYHGSMASPSSTAAYL
 IGATKWDDEAEDYLRHVVRNGAGHNGGIGSTFPTTHFECSWIIATLLKVGFTLKQIDGD
 GLRGLSTILLEALRDENGVI GFAPRTADVDDTAKALLALSLVNQVPSPDIMIKVFEGKDH
 FTTFGSERDPSLTSLNLHVLSSLKQSNLSQYHPQILKTTLFTCRWWWGSDHCVKDKWNLS
 HLYPTMLLVEAFTEVLHLIDGGELSSLFDESFKCKIGLSIFQAVLRIILTQDNDGSWRGY
 REQTCYAILALVQARHVCFFTHMVDRLQSCVDRGFSWLKSCSFHSQDLTWTSTAYEVGF
 VAEAYKLAALQASASLEVPAATIGHSVTSAVPSSDLEKYMRLVRKTALESPLDDEWGLMASI
 IESSFFVPLLQAQRVEIYPRDNIKVDKYLIIIPFTWVGNNRSRTFASNRWLYDMMYL
 SLLGYQTDYMEAVAGPVFGDVSLHQTIDKVIDNTMGNLARANGTVHSGNGHQHESPNI
 GQVEDTLTRFTNSVLNKKDVLNSSSDQDTLRREFRTFMHAHITQIEDNSRFSKQASSDA
 FSSPEQSYFQWVNSTGGSHVACAYSAFNSCLMSANLLQKDAFPGTQKYLISSVMRHA
 TNMCRMYNDFGSIARDNAERNVNSIHFFPEFTLCNGTSQNLDERKERLLKIATYEQGYLDR
 ALEALERQSRDDAGDRAGSKDMRKLKIVKLFCDVTDLYDQLYVIKDLSSSMK

FIGURE 16

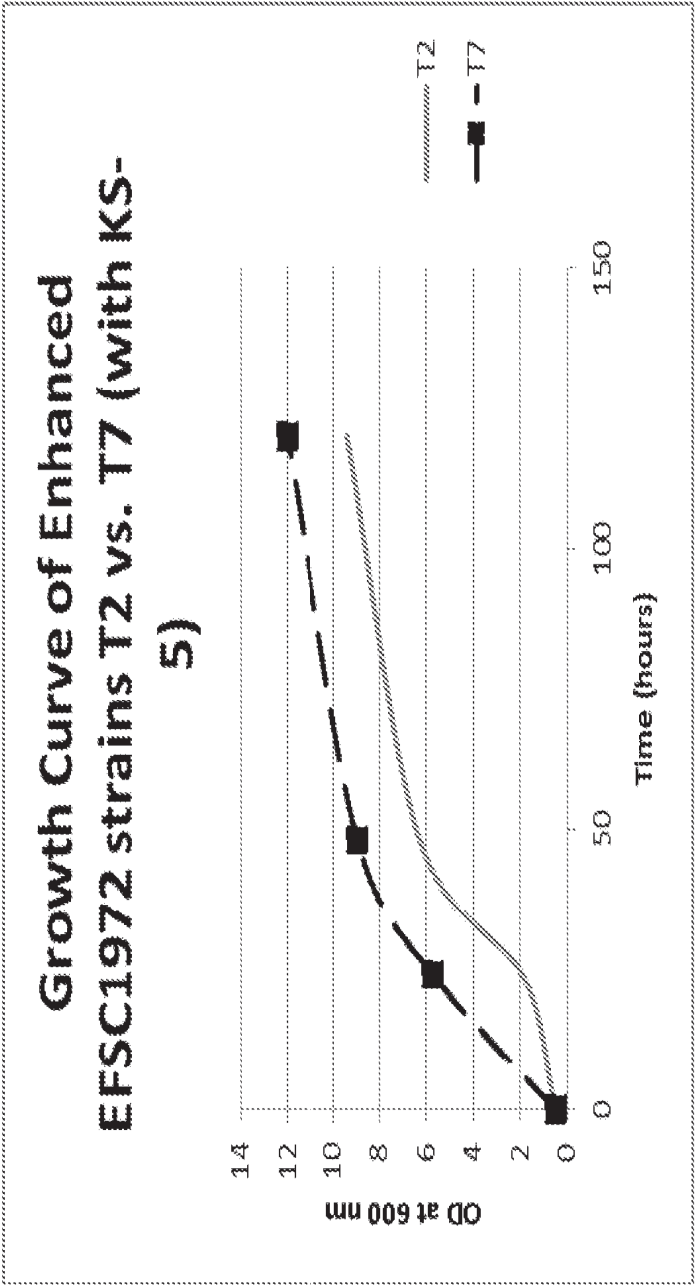


FIGURE 17

A.
> Native AtSUS1

ATGGCAAAACGCTGAACGTATGATAAACGGCGGTCCACAGCCAAACGTGAGCGTTTGAACGAAACGCTTGTCTTGAGAGAAACGAAGTCCT
 TGCCTTGCTTTCAGGGTTGAAGCCAAAGGTAAAGGTATTTTACAACAAAACAGATCATTGCTGAATTCGAAGCTTTGCCGTGAACAAA
 CCGGAAGAAACTTGAAGTGGTCCTTTCTTGACCTTCTCAAATCCACTCAGGAAGCAATTGTGTTGCCACCATGGGTTGCTCTAGCT
 GTGAGGCCAAGGCTGGTGTGGGAATACTTACGAGTCAATCTCCATGCTCTTGTGCTTGAAGAACTCCAACCTGCTGAGTTTCTTCA
 TTTCAAGGAAGAACTCGTTGATGGAGTTAAGAAATGTTAAATTTCACTCTTGAGCTTGATTTTCGAGCCATTCAAATGCGTCTATCCCTCGTC
 CAACACTCCACAAAATACATTGGAAATGGTGTGACTTCTTAACCGTCAATTATCGGCTAAGCTCTTCCATGACAAGGAGAGTTTGCTT
 CCATTGCTTAAGTTCCTTCACAGCCACCAGGGCAAGAACCTGATGTTGAGCGAGAAGATTTCAGAACCCTCAACACTCTGCAACA
 CACCTTGAGGAAAGCAGAAGATATCTAGCAGAGCTTAAGTCCGAAACACTGTATGAAGAGTTTGAGGCCAAAGTTTGAGGAGATTGGTC
 TTGAGAGGGATGGGGAGACAAATGCAGAGCGTGCTTGACATGATACGTCTTCTTTGGACCTTCTTGAGGCGCTGATCCTTGCACT
 CTTGAGACTTTTCTTGGAGAGTACCAATGGTGTTCAAACGTTGTGATCCTCTCTCCACATGGTTACTTTGCTCAGGACAAATGTTCTTGG
 TTACCCCTGACACTGGTGACAGGTTGTTTACATTTCTTGATCAAGTTCTGTGCTCTGGAGATAGAGATGCTTCAACGTATTAAGCAACAAG
 GACTCAACATTAAACCAAGGATCTCATTTCTAACTCGACTTCTACCTGATGCGGTAGGAACTACATGCGGTGAACGTCTCGAGAGAGTT
 TATGATTCTGAGTACTGTGATAATCTTCGTGTGCCCTTCAGAAACAGAGAAGGATATTGTTTCGCAAAATGGATCTCAAGGTTTCGAAGTCTG
 GCCATATCTAGAGACTTACACCGAGGATGCTGCGGTTGAGCTATCGAAAGAAATTGAATGGCAAGCCTGACCTTATCATTTGGTAACATACA
 GTGATGGAAATCTTGTTGCTTTTATTGGCTCACAAACTTGGTGCTCACTCAGTGTAACCATTTGCTCATGCTCTTGAGAAAAACAAAGTAC
 CCGGATTCTGATATCTACTGGAAGAAGCTTGACGACAAGTACCATTTCTCATGCCAGTTCACTGCGGATATTTTCGCAATGAACCCACAC
 TGATTTTCATCATCACTAGTACTTTCCAAAGAAATTGCTGGAAGCAAGAAACTGTTGGGCAGTATGAAAGCCACACAGCCTTTACTCTTC
 CCGGATTGTATCGAGTTGTCACGGATTTGATGTTTGTATCCCAAGTTCAACATTGTCTCTCCTGGTGTGATATGAGCATCTACTTC
 CCTTACACAGAGGAGAAGCGTAGATTGACTAAGTTCACCTCTGAGATCGAGGAGCTCCTCTACAGCGATGTTGAGAAACAAAGAGCACTT
 ATGTGTGCTCAAGGACAAAGAACCGGATTTCTCTTCACAATGGCTAGGCTTGATCGTGTCAAGAACTTGTCAAGTCTTGTGAGTGGT
 ACGGGAAGAACACCCGCTTGGCTGAGCTAGCTAAGTAACTTGGTTGTTGTTGGAGGAGACAGGAGGAAAGAGTCAAAGGACAAATGAAGAGAAA
 GCAGAGATGAAGAAAATGTATGATCTCATTTAGGAATACAAGCTAAACGGTCAAGTTTCAGGTGGATCTCCTCTCAGATGGACCGGGTAAG
 GAACGGTGAGCTGTACCGGTACATCTGTGACACCAAGGTGCTTTTGTCCAACTGCTATATGAAGCCTTTGGGTAACTGTTGTGG
 AGGCTATGACTTGTGGTTTACCGACTTTCGCCACTTGCAAGGTGGTCCAGCTGAGATCAATTGTGCACGGTAAATCGGGTTTCCACATT
 GACCCCTTACCATGTGTGATCAGGCTGCTGATACTCTTGTGATTTCTTCAACCAAGTGAAGGAGGATCCATCTCACTGGGATGAGATCTC
 AAAAGGAGGGCTTCAGAGGATTGAGGAGAAAATACACTTGGCAAAATCTATTCAACAGAGGCTCTTGACATTGACTGGTGTATGGATTCT

GGAAGCATGTCTCGAACCCTTGACCGTCTTGAGGCTCGCCGTTACCTTGAAATGTTCTATGCATTGAAGTATCGCCCATTTGGCTCAGGCT
 GTT**CCTCTTGACAAGATGATTGA**

B.

> SrSUS contig10573 selection_ORF_S11E

ATGGCGGAACGTTACTCACTCGTGTTCAC**gaac**TTTCGTGAGCGTCTCGATTCAACTCTCGCAACTCATCGTAATGAAATCCTCTTGTT
 TCCTTTCAAGGATTGAAAGCCATGGAAGAGGAATATTGAAGCCTCATCAAGTTATGACTGAATTTGAAGCTATCTGCAAGAAAGATCAGA
 GCAAACTCTCTGATGGTGCTTTTATGAAGTTCTTAAATGCACACAGGAAGCAATAGTGCAACCTCCATGGGTGCACTCGCGATCCGT
 CTTCCGACCCGGTGTTTGGGAATATGTTAGAGTCAATGTTAATGTTTGGTGGTTGAAGAAATTAAGTGTTCCTGAATATCTTCACTTCAA
 AGAAGAAATTGGTTAATGGAACATCGAATGGCAACTTCGTGTTGGAACCTTTTGAACCTTTTACCGCATCGTTTCTCGACCAACTT
 TAACCAAGTCTATTGGTAATGGTGTGAGTTTCTAAACAGACATTTATCTGCTAAAATGTTTCAATGATAAGGATAGCATGACCCCTCTT
 CTTGATTTCTACGGACTCACCACTATAAGGGGAAAGACAATGATGTTGAATGATAGAATCCAAAACCTCAATGCTCTACAATCGGTGTT
 GCGAAAGCGTCAGAGTACTTATCAACACTCGACGGCGCAACACCGTACTCTCAGTTTGAACATAAGTTTCAAGAAATCGGGTTGGAGA
 GAGGTTGGGTGATAAAGCGGAGGTCTGTAATGGAGATGATCCACATGCTTCTAGACCTTCTAGAAGCACCCGACGCAATGCACACTCGAG
 AAGTTTCTCGGAAGAAATCCCAATGGTTTTCATATGTTGTCAATCTTTTCGCCCTCATGGCTACTTCGCGCAAGAAAATGTGTTGGGATATCC
 CGACACTGGCGTCAGGTTGTTTACATCTTGGATCAAGTTCCCGCTCTGGAAACGCGAGATGCTCAAAAGGATTAAGGAGCAAGGACTCG
 ATATCATTCCTCGTATATTGATTGTTACGAGGCTTCTTCTGACCGGTTGGGACCACATGCGGGCAACGTTTAGAGAAAAGTGTGTTGGA
 GCCGAACACTCGCATATTCTTCGGGTCCTTTAGAACCCGAAAAGGATATTCTTCGTAATGGATCTCTCGTTTGAGGTGTGGCCCTTA
 CATCGAGACTTTCACCGAGGATGTTGCTAAAGAAGTTACAGCAGAGTTGCAAGCAAAACCAGATTTGATCATTTGGAAACTATAGTGAAG
 GAAATTTGGTTGCACTTTGCTAGCTCACAAAGTTGGGTGCACTCAGTGACCATTGCTCATGCTTTTGAGAGAAAACATAAATACCCGGAT
 TCTGATATCTACTGGAAGAACTTTGAGGAGAAATATCATTTCTTCGCAGTTTACCGCTGATCTTATCGCTATGAACCATACCGACTT
 CATCATCACCACTACTTTCCAAAGAAATTGCTGGAAAGTAAGGACACGGTTGGACAGTACGAGAGTCATACCGGTTTCACAATGCCGGGAT
 TGATACGGGTGGTTCACGGATCGATGTTTGTGACCCCAAATTCAATATTGTTTCAACCCGGGCGGATATGGGAATTTACTACTCGTAT
 ACCGAGAAAAGAAAGAGCTCACTGGCTTCACCCCTGAAATCGATGAACCTTCTCTTAGTTCCTCGTCAAAACGAAGAACACTTATGTGT
 GTTGAAGGATAAGAGTAACCAATCTTGTTCACAAATGGCGCGATTGGATAATGTGAAGAAATTAACCGGACTGGTTGAATGGTACGCTA
 AAAACGACCCGCTTCGTGAGCTCGTGAACCTTCGTGGTTCGTGGTGGTGAACGAAAGAGTCGAAAGATCTTGAAGAGCAAGCTCAG
 ATGAAGAAGATGCACTTATCGAAACCTACAAACTCAACGGTCAGTTCAGGTGGATATCTTCAAAATGAACCCGCTGAGGAACGG
 TGAGTTGATACGCGTTATTGCTGACACGCGAGGTGCGTTTATCCAGCTGCGTTTACGAGGCGTTTGGGTTGACGGTTGTTGGAGGCCA
 TGACTTGTGGCCTGCCGACATTCCGACACTTCAATGGTGGCCCGCTGAGATTATTGTTACGGGAAATCCGGGTTCCATATTGACCCG
 TATCACGGTGACCAAGGTACCGAGTTGCTGTCATTTCTTGAAGAAAACATAACAAAGACCCGGTCAATGGGAGGCCATTTCCAAGGG
 TGGTCTGCAACGTTATTCAGGAGAAAATACACGTGGCAGATTTTATTCAGATAGGTTGTTGACGCTTGCCCGAGTTTATGGATTCTGGAAGC

ATGTGTCGAAGCTTGACAGGCTCGAGATCCGTGTTATCTTGAAATGTTTTACGCGCTCAAGTATCGCAAACCTGGCTGAATCTGTCTCCA
TTGGCTGTTGATGAGTGA

C.

> Coffea arabica sus1 gene

ATGGCCGAACGTTCTGACCCCGTGTTCACAGCCTCCGGAACGCCCTTGATGCTACTTTGGCTGCCACCCGCAACGATGTTTGTGCTGTT
 TATGTCGAGGCTTGAAACCCATGGGAAAGGATCCTGAAACCCACCAACTTTTGGCTGAGTTTGAAGAAATTAACAAGGATGGTAAAC
 AAAAAATTCAATGATCATGCGCTTGAAAGAAGTCCCTGAAGTCCACACAGGAAGCAATTGTGTGCCCCCTGGGTGCACTTGCTATTCTGT
 CTCAGACCTGGTGTCTGGGAGTATGTTGAGTCAATGTCCATGCACTCGTTGTTGAGGAGTTAACCGTGCCAGAGTACCTGCATTTCAA
 GGAAGAACTCGTTGATGGAGCAAAAAATGGGAATTTGTGTTGGAACTGGACTTCGAACCAATTTACAGCCTCTTTTCCCAAGCCAACCTC
 TAACTAAGTACATAGGCGACGGAGTTGAGTTCCTCAACAGGCACCTCTCTGCCAAAATGTTCCATGACAAGGAGAGCATGGCCCCCTCTC
 CTTGATTTTCTCCGTGTTCAACCAATACAAAGGCAAGACAATGATGCTTAACGACAGGATCAAGGACCTTAACACTCTCCAAGCAGTTCT
 GAGGAAGCAGAGGAGTACCTAACAACTCTCTGCAGATACACCATACTCTGAAATTCGAGCACAAAATTCCAAGAAAATTTGGACTGGAGA
 GAGGTGGGTGATACTGCTGAGCGTGTCTTGAAATGATCTGCATGCTTCTGGATCTTCTGGGGCTCTGACTCGTGCACTAGAG
 AAATTCCTAGGGAGAATCCCTATGGTATTCAATGTTGTTATTTTCCCCCATGGATACTTTGCCAGGAAAACGTAATTTGGGTTATCC
 TGATACCGGTGGCCAGGTTGTATTACATATTGGATCAAGTTCCTGGCTGGAGCGTGAGATGCTGAAGAGGATAAAGGAACAAGGACTTG
 ATGTCAAGCCACGCATTCTAATTATAACTAGGCTGCTTCTGATGCCCCCTGGAACCACTTGTGGTCAACGGCTTGAGAAAAGTATATGGA
 TCAGAGTACTCCCATATACTCAGAGTCCCCTTCAGAACTGAGAAGGGAGTTGTTCCGAAAATGGATCTCTCGCTTTGAAAGTTTGGCCCTA
 CATGGAAACATTTACTGAGGATGTTGCCAAAAGAAAGTCACTGCAGAAATTACAGGCCAAAGCCAGATTGGTTATTGGTAACTACAGTGAGG
 GTAAACCTTGTGCTCCCTTGTGCTCACAAAGTTAGGTGAACACAGTGTAACCATGCTCATGCTTTTGAAAAAACCAAGTATCCCTGAT
 TCTGATATTATTGAGCAAAATTGATGAAAAGTACCACCTCTCATGTCACTGCGGATCTTATCGCAATGAACCATACAGATTT
 CATTATCACTAGCACTTTCCAAAGAAATAGCTGGAAAGCAAGGACACTGTTGGGCAATATGAAAGCCATATGGCCTTCACAAATGCCAGGAT
 TATACAGAGTTGTACATGGCATTGATGTTTGTGATCCAAAATTCAACATTGCTCACCTGGAGCTGATACAAAACCTCTACTTCCCCACAC
 ACAGAGAAAGGAAAAGAGATTGACATCTTTCCATCCTGAAAATTGAGGAGTTGCTTTTCAGCGATGTGGAGAAATGAGGAACACCTATGTGT
 GCTAAAAGACAAAAAGAACCTATCTTATTCACCATGGCAAGACTGGATCGCGTAAAGAAATTGACAGGGCTTGTGAAATTGTATGCTA
 AGAACCCAAAACCTAAGGGAATTGGTCAATCTTGTGCTGGTTGGTGAGACCCGAAGGAAGGAATCCAAAAGATTGGAAGAAACAAGCTGAG
 ATGAAGAAAAATGTATTGATAGAAAACCTACAACCTGAACGGCCAAATTCAGATGGATTTCTTCTCAGATGAACAGGGTTAGAAAATGG
 TGAACCTCTATCGGTACATTGCTGACACCCAGGGAGCATTCGTGCAACCTGCATTTTATGAGGCAATTTGGGTTGACTGTGGTCGAGGCCA
 TGACATGTGTTGCCAACATTTGCAACCAACCATGGTGTCTGCTGAGATCATATTATCATGGGAAAATCTGGTTTCCACATTTGATCCA
 TACCATGGTGAGCAGGTACGGAGCTCCTTGCCAAATTTCTTGAAAAGGTGCAAGAAAGAGCCTTCTTACTGGGACACCATTTTCAGCCGG
 TGGCTTGAAAGCGTATCCAGGAAAAGTACACCTGGCAAATTTTACTCAGATCGGTGCTGACGCTGGCTGGAGTTTATGGATTCTTGGAAT

GTGTTTCCAAGCTTGATCGCCAGGAGATCCGCCGTTATCTGGAAATGTTTTATGCTCTCAAGTATCGCAAGTTGGCTGAAGCTGTTCCA
TTGGCTGTTGATCAGTAA

FIGURE 18

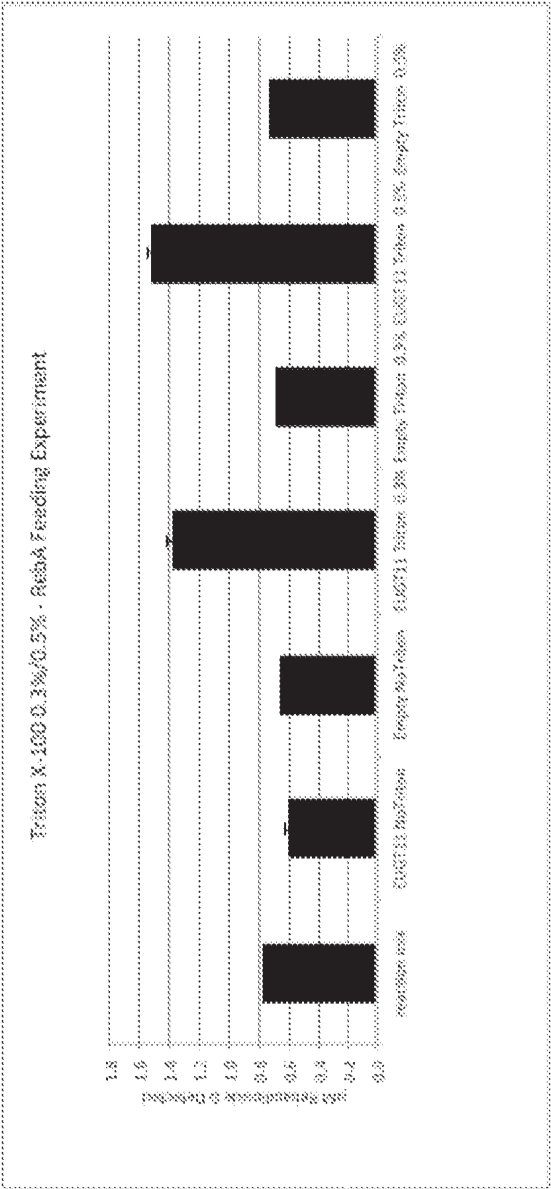


FIGURE 19

A.

MHITKPHAAMFSSPGMIGHVIPVIELGKRLSANNNGFHVTVFVLETDAAASQSKFLNSTGVDIVKLPSPDIYGLVDPDDHVVTKIGVIMRAAVPALRSKIAAM
 HQKPTALIVDLFGTDALCLAKEFNMLSVFIPTNARFLGVSIYYPNLDKIKEHTVQRNPLAIPGCEPVRFEEDTLDAYLVPDEPVYRDFVRHGLAYPKADGIL
 VNTWEEMEPKSLKLLNPKLLGRVARVPVYPIGPLCRPIQSSETDHPVLDWLNEQPNESVLYISFGSGGCLSAKQLTELAWGLEQSQQRFFVWVVRPPVDG
 SCCSEYVSANGGGTEDNTPEYLPEGFVSRSDRGFVPSWAPQAEILSHRAVGGFLTHCGWSSTLESVVGVPMAIWPLFAEQNMNAALLSDELGIIVRL
 DDPKEDISRWKIEALVRKVMTEKEGEAMRRKVKKLRDSAEMSLSIDGGGLAHESLCRVTKECQRFLEVRVVDLSRGA

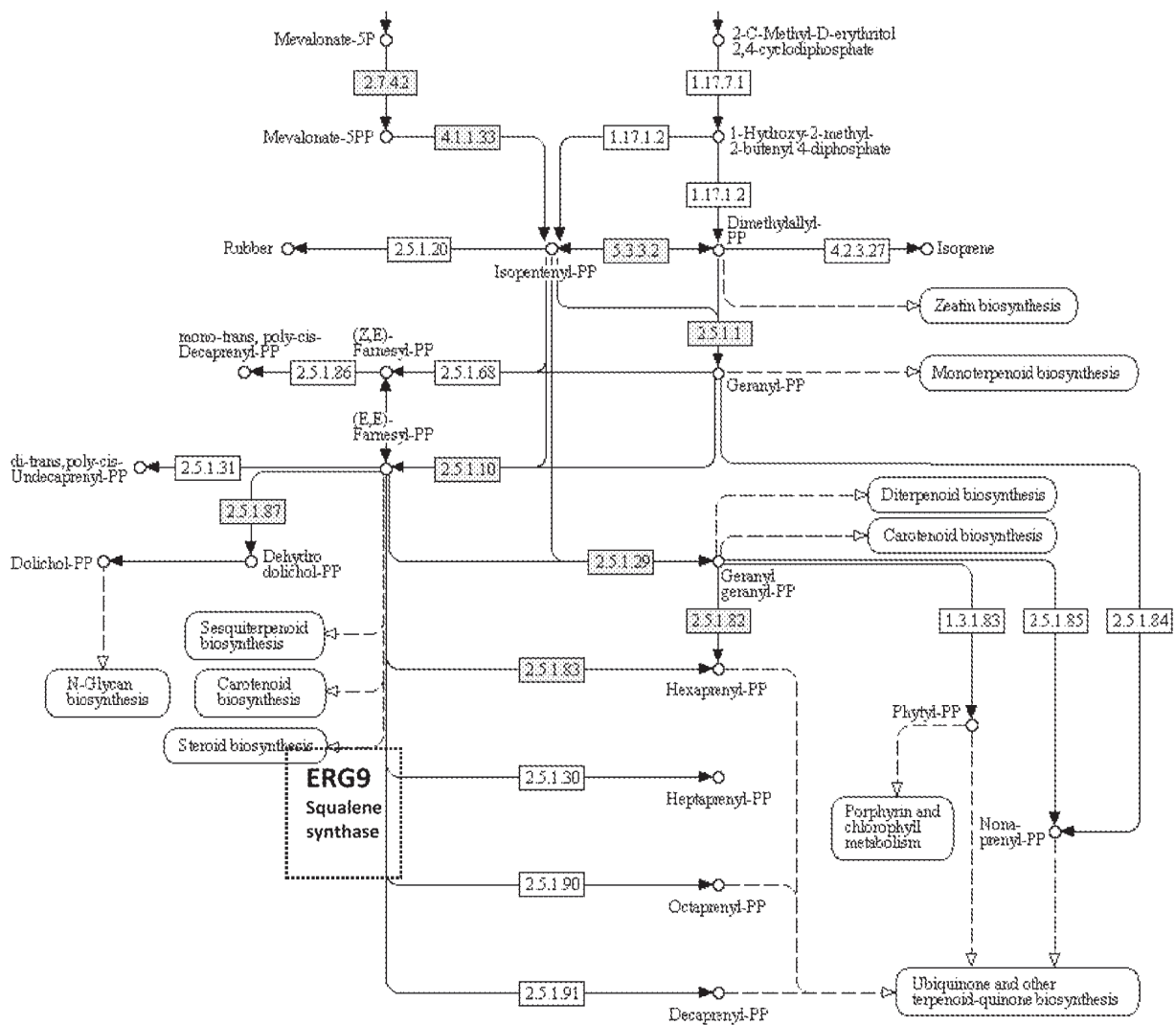
B.

MGAYETEKPTKDAAALETQSPEDFDQPSPLRKIISVASIAAGVQFGWALQLSLTPYQLLGIPIHKWSSLIWLCGPVSGMIVQPIVGFHSDRCRSKFGRRRPF
 IATGAALVAVAVFLIGYAADFGYKMGDKLEEKVKVRAIGIFALGFWILDVANNTLQGPCRAFLADLAAGDAKTRRVANAFSFFMAVGNVLGYAAGSYTNL
 HKMFPFTMTKACDIYCANLKTCCFLSITLLIVTVTSLWYVNDKQWSPPRNADDDKETSVPFGEIFGAFKVMKRPMWMLLIVTALNWIWFPFLFDT
 DWMGREVFGGSDGNERSKKIYSLGVQSGAMGLMFNSIVLGFMSLGVWIGRKLGGAKRLWGVNFILAAGLAMTVLTKFAEDHRKTAGDLAGPSAS
 VKAGALSFAVLGIPLAITFSTPFALASIFSSCSGAGQGSLGLVNLAIVIPQMIVSLGGGPFDALEFGGGLPAFIVAAIAAISGVLAITVLPSPPPDAPKATTM
 GGFH

C.

MAERVLT RVHSLRERLDATLAHRNDVLLFMSRLETHGKGILKPHQLLAEEFEEINKDGKQKIHDAFEVVKSTQEAIVLPPWVALAIRLRPGVWEYVRVN
 VHALVVEELTVPEYLHFKEELVDGSKNGNFVLELDFEPFTASFPKPTLTKYIGDGVFELNRHLSAKMFHDKESMAPLLDFLRVHQYKGKTMIMLNDRIKDLNT
 LQAVLRKAEYYLTLSADTPYSEFEHKFQEIGLERGWGDTAERVLEMICMLLDLLGAPDSCCTLEKFLGRIPMVFNVVILSPHGYFAQENVLGYPDTGGQVVYI
 LDQVPALEREMLKRIKEQGLDVKPRIIITRLLPDAGTTCCGQRLEKVYGEYSEYSHILRVPFRTKEGVVRKWISRFVWVWPMETFTEDVAKEVTAELQAKPDVI
 GNYSEGNLVASLLAHKLGVTQCTIAHALEKTKYPDSDIYLSKFDEKYHFSCQFTADLIAMNHTDFIITSTFQEIAGSKDTVQYESHMAFTMPGLYRVVHGID
 VFDPKFNIVSPGADTNLYFPHTKEKRLTSFHPHEIEELLFSDVENEELCVLKDKKPILFTMARLDRVKNLTGLVELYAKNPKRELNVLVVGGDRRKESKDL
 EEQAEMKKMYSIETYNLNGQFRWISSQMNVRNGELYRYIADTRGAFVQPAFYEAFLTVVEAMTCGLPTFATNHGGPAEIIHKGSGFHIDPYHGEQVS
 ELLANFFERCKEPSYWDTISAGGLKRIQEKYTWQIYSDRLITLAGVYGFWKCVSKLDRQERIRRYLEMFFYALKYRKLAEAVPLAVDQ

Fig. 20



SEQ ID NO: 185 (ScCyc1 promoter)

gcgtgggttgatgaagccacgcgtaggcaatcctcgagcagatccgcaggcgtatataatagcgtggaatgtagtgcacatacaggcataataatgtgtgc
gacgacacatgatggcatgcgtgctgtatataaaactctgtttcttctctaaatattcttctatacattaggaccttgacagcataaattactatacttatagaca
cacaacacaaaatacacacactaaattaata

SEQ ID NO: 186 (ScKex2 promoter)

ggctcagcagctctgtagatacacgtagtctcgacatgttttttttactatacatataaagaaataaaaaatgataacgtgtatatattattattcatataatcaatgagggctcatcttctgaa
acgcaaaaaaacggttaattggaaaaaaaataaagatagaaaaaagaaacaaacaaaggaagggttagcatattaaataactgagctgatacttcaacagcatcgctgaagagaaacagtatt
gaaacggaacatttctaaggcaacaaggtactccatatattgtcggacgtgtcttctctcgttttcataatgtgcataaattgtcataagccgtcttttctcgtggttaaacatcccggtttgttaa
aagagaaatctattccacatatttcatttcggtccacatactaaggataaaactaatccgtgtgttttttggccgtgcacataattaaactaataaccattatcaga

FIGURE 22

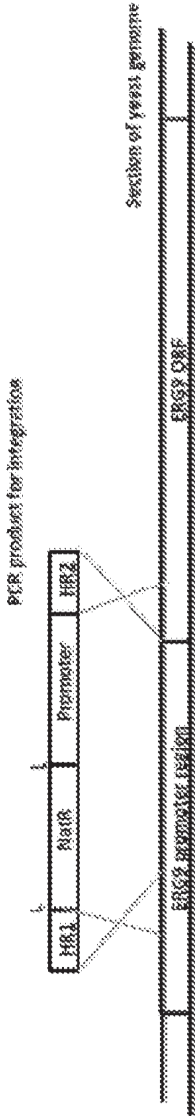


FIGURE 23

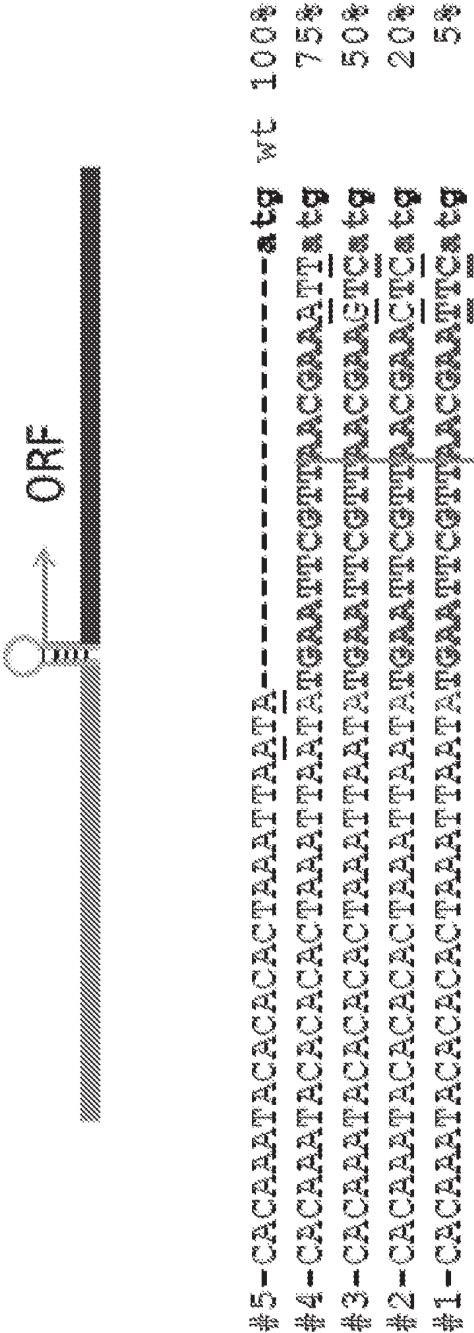


FIGURE 24

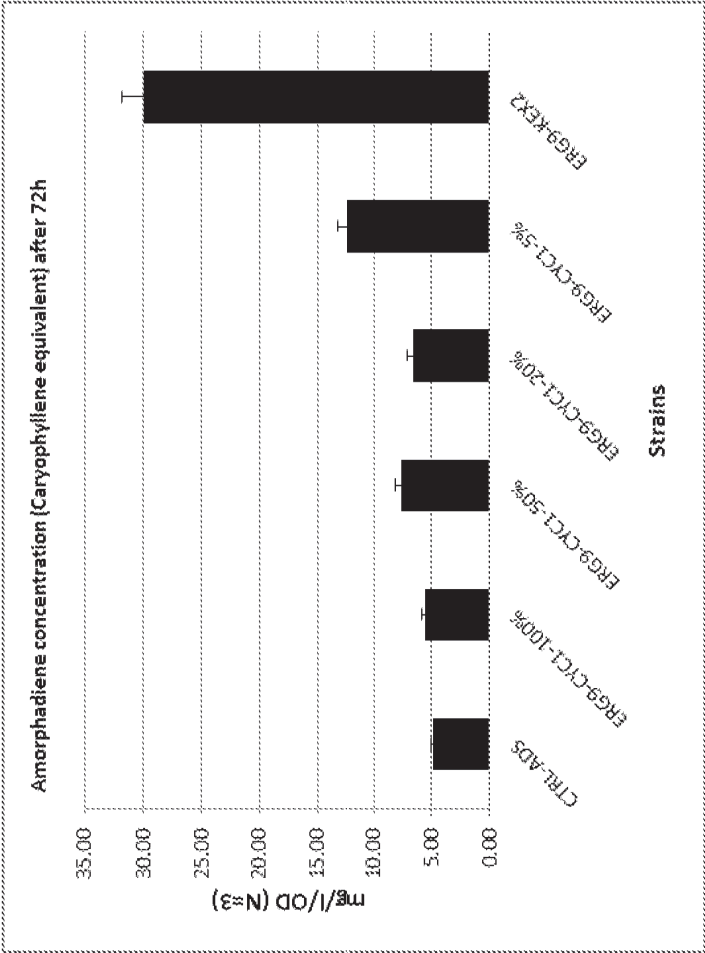


FIGURE 25

```

>sp|P29704|FDFT_YEAST Squalene synthase OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=ERG9 PE=1 SV=2
MGKLLQLALHPVEMKAALKKFCRTPLFSYDQSTSPYLLHCFELNLTSSRFAAVIREL
HPELRNCVTLFYILRALDTHEDDMSIEHDLKIDLLRHFHEKLLLTWKWDFDGNAPDVKDR
AVLTFDESILIEFHKLKPEYQEVKEITEKMGNGMADYILDENYNLGLQTVHDYDVYCH
YVAGLVGDGLTRLIVIAKPFANESLYSNEQLYESMGLFLQKTNIRDYNEDLVDGRSFWPK
EIWSQYAPQLKDFMKPENEQLGLDCINHLVLNALSHVIDVLTYLAGIHEQSTTFCAIPQ
VMAIATLALVFNNREVLHGNVKIRKGTTCYLILKSRTLRGCVEIFDYVYLRDIKSKLAVQD
PNFLKLNIOISKIEQFMEEMYQDKLPPNVKPNETPIFLKVKERSRYDDELVPTQQEEBYK
ENMVLISILSVLLGFYIYTLHRA

>sp|P26596|FDFT_SCHPO Squalene synthase OS=Schizosaccharomyces pombe (strain ATCC 38366 / 972) GN=erg9 PE=1 SV=1
MSLANRIIEJRCLCYQKLVNDLPSYGEDENVPQNIRRCYQLLDMTSRSFAVVIKELPNGI
REAVMIFYLVLRGLDTEDDMTPLDKKLPIRLDFYKTIEVEGWTFNESGPNKDRQLLV
EFDVVIKEYNLNSEGYNVISNITKEMMGDMAYYASLAEKNDGFSVETIEDFNKYCHYVA
GLVGIGLSRLFAQSKLEDPLAHSQAISNSLGLFLQKVNIIRDYREDFDDNRHFWPREIW
SKYTSSFGDLCLPDNSEKALECLSDMTANALTHAIDALVYLSQLKTQEIFNFCaipQVMA
IATLAAVFRNPVFTNVKIRKGQAVQHILHSVNLKNVCDLFLRYTRDIHYKNTPKDPNF
LKISIECGKIEQVSESLFPRRFRMEYKAYVSKLSEQKKGNGTQKAILNDEQKELYRKDL
QKLGISILFVFFHILVCLAVIFYVFNIRIHWSDFKELNLF

>sp|Q9Y753|FDFT_YARLI Squalene synthase OS=Yarrowia lipolytica (strain CLIB 122 / E 150) GN=SQS1 PE=3 SV=1
MGKJIELLLHPSELSAAIHVKLWRQPLHPRDLSKESTELRRCYELLDVCSRSFAAVIREL
HPEVRDAVMLFYILRALDTHEDDMLSRDIKIPILRDFTKCMKTPGWKFTDSDPNERDR
VVLQEPVVMTEFNKLPKYQEVYDITDRMGNGMADYVIDDDFNNGVDITIAAYDLYCH
HVAGIVGEGLTRITILAGFGTDVLHENPRLQESMGLFLQKVNIIRDYREDIDVNRWFWR
EIWHKYAEEMRDFKPKYSKKALHCTSDLVANALGHATDCLDYLDNVDPSTFTFCAIPQ

```

VMAIATLDELVYRNPDVFQKNVKLRKGTTSVLILEASNVSGVCDIFTRYARKVYKKSDFND
PNYFRVSVLCGKIEQHAALIKRQRPAAKTIAQLEGERKEMALSILVCLAVIFMSGMLMA
YIAYVSGFRWSPREIFDSKMFPLRD

>sp|Q9HGZ6|FDFT__CANGA Squalene synthase OS=Candida glabrata (strain ATCC 2001 / CBS 138 / JCM 3761 / NBRC 0622 / NRRL Y-65) GN=ERG9 PE=3 SV=2
MGKVLDLALHPLERAAALKLKFIROPLFSTNDTRATPQLERCYELLNLTSRSFAAVIMEL
HPELRNVIMVYFYLIRALDVTVEDDMTIDPQKVKVLREFDSKLDTTDWSFDGNDLKEKDR
VVLTFFPCILGEYHKLKPEYQKVIKRITGLMGNGMADYILDENFNLNGVQTVKDYDKYCH
YVAGLVGDGLTELIVLAGFGSDDL YHGKNSFQLYESMGLFLQKTNIIIRDY AEDLDDGRSF
WPKEIWSEYATKLTDFRDPKNTQKGVDCINHLVLNALTHVIDVLTYLSSIHQSFSQFCA
IPQVMAIATLAKVFNNPEVLKKNVKIRKGTTCDLILNSRTLKGCVDIFQYYLRDMKQRLP
VEDPNYLKFNIVAKIEQFIEEMFQDNLPAVEPRETMILKVVQERLKWDTQVIPRVQEE
DYKFNMAHSVFCVLLSFYFFTK

>sp|Q752X9|FDFT__ASHGO Squalene synthase OS=Ashbya gossypii (strain ATCC 10895 / CBS 109.51 / FGSC 9923 / NRRL Y-1056) GN=ERG9 PE=3 SV=1
MGKVVQLFTHPLELKAALKLKFLREPLYPADDTQGSAEKRCYQLLQRTSRSFAAVIMEL
HPELRNAVMLFYLIRALDVTVEDDMTISPVKVPLREFDQKLKLDTSFDGNAKTEKDR
DVLVEFSTILAEFHKLKPEYQQVIADITHKMNGMGMDYILDEKENLSGLETIQDYDRYCH
YVAGLVGDGLTHLIMLAKFSSPGLYYDSPDLYESMGLFLQKTNIIIRDY AEDLADGRSFWP
KEIWSHYADDLASFSPENATAGVYCINHLVLNALGHVQHVLTYLASLREQSSSFQFCAIP
QVMAIATLALVFGNERVLQTSVKIRKGTTCYLILKSRTFQGCVEFEHYLRDIRKRLTVA
DPNYLKNIEIAKLDKFIEEMYQDKLPVGAKPQETEIYKKVRSAYDLEVLPREQEEEF
KFNVLLSILFTVFGALYWYAK

>sp|O74165|FDFT__CYBJA Squalene synthase OS=Cyberlindnera jadinii GN=ERG9 PE=3 SV=1
MGKLLQALIPDELASIVQFKLPRKNENARNPATESAELIRCVELLNLTSRSFAAVIEEL
HPELRNVIMVYFYLIRALDVTVEVDMSIENSVKLPVLRQFHEKLDTKDWTDFGNSPNEKDR
CVLVEFDRILGQYHELKPQYQKVKEITEKMGNGMADYIENENFNSNGLLTIEDYDLYCY

YVAGLVGDGLTQLIVLAKFGNSELSVNKQLFKSMGLFLQKTNIRDYEEEDQVDGRAFWPK
 EIWKGYANELSDFMKPENQSQGLWCISELVCNALDHVIDVLQYLALVEEQTSFNFCaipQ
 VMAIATLELVQNPNQVLQHVVKIRKGGTTVSLILESRTLEGCARIFRRYLRRKIHHSHPSD
 PNYLRLGITIGKIEQFLDGMYPHYVPKGITPQTTSIRTQVVKRLQLDEPMKRDIDEEILK
 TRILLLSLGVAVFGVVYGVVRII

>sp|P78589|FDFT_CANAL Squalene synthase OS=Candida albicans GN=ERG9 PE=3 SV=1
 MGKFLQLLSHPTTELKAVIQLFGRQPLHPGKRVDNDKELGRCYELLNLTSRSFAAVIEEL
 HPELRDAVMIFYLVLRALDTIEDDMTKSSIKIPLREFDTKLNTKNWTFTDGYGPNKDR
 TVLVEFDKILNVYHRLKPQYQDIIKSITFKMGNGMADYILDEEFNVYGVATVEDYNLYCH
 YVAGLVGEGLTNLFVLANFGDKTLTENNFAKADSMGLFLQKTNIRDYHEDLQDGRSFWP
 REIWSKYTENLQDFHKVKTPAKEFAGVSCINELVLNALGHVTDCLDYLSLVKDPSSFSC
 AIPQVMAVATLAEVYNNPKVLHGvvKIRKGGTTCLRILESRTLPGVVKIFKEYIQVINHKS
 SVRDPNYLKIGIKCGEIEQYCEMIVPNKQALPPSMKSLPENKFKIVASRESIDLSVQRR
 IEPGNFNCNVVLFGIGALILSLIYFVLY

>sp|P38604|ERG7_YEAST Lanosterol synthase OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=ERG7 PE=1 SV=6
 MTEFYSDTIGLPKTDPRLWRLRTDELGRESWEYLTPQQAANDPPSTFTTQWLLQDPKFPQP
 HPERNKHSPDFSADFACHNGASFFKLLQEPDSGIFPCQYKGPMTIGYVAVNYIAGIEI
 PEHERIELRYIVNTAHPVDGGWGLHSVDKSTVFGTVLNVYLRLGLPKDHPVCAKARS
 TLLRLGGAIGSPHWGKIWLSALNLYKWEGVNPAPPETWLLPYSLPMHPGRWWVHTRGVYI
 PVSYLSLVKFCSPMTPLLEELRNEIYTKPFDKINFKNRNTVCGVDLYYPHSTTLNIANS
 LVVPYEKYLNRNFIYSLSKKVYDLIKTELQNTDSLCAIPVNAFCALVTLIEEGVDSEA
 FQRLQYRFKDALFHGPQGMIMGTNGVQTDCAFAIQYFFVAGLAERPEFYNTIVSAYKF
 LCHAQFDTECVPGSYRDKRKGAWGFSTKTQGYTVADCTAEAIAIMVKNSPVSEVHHM
 ISSERLFEgidvLLNLQNIQSFEYGFATYEKIKAPLAMETLNPAEVFGNIMVEYPYVEC
 TDSSVLGLTYFHKYFDYRKEEIRIRIAIEFIKKSQLPDGSWYGSWGICFTYAGMFALE
 ALHTVGETYENSSTVRKGCDFLVSKQMKDGGWGESMKSSSELHSYVVDSEKSLVVQTAWALI

ALLFAEYPNKEVIDRGIDLLKNRQEESEGEWKFESEGVFNHSCAIEYPSYRFLFPIKALG
MYSRAYETHTL

>sp|P37268|FDFDT_HUMAN Squalene synthase OS=Homo sapiens GN=FDFDT1 PE=1 SV=1
MEFVKCLGHPEEFYNLVRFRIGGKRKVMMPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA
LDGEMRNAVCIFYLVLRALDTLEDDMTISVEKKVPLLNHFHSFLYQPDWRFMESKEKDRQ
VLEDFPTISLEFRNLAEKYQTVIADICRRMGIGMAEFLDKHVTSEQEWDKYCHYVAGLVG
IGLSRLFSASEFEDPLVGEDTERANSMGLFLQKTNIRDYLEDQGGREFWPQEVVWSRYV
KKLGDFAKPENIDLAVQCLNELITNALHHIPDVITYLSRLRNQSVNFCAIPQVMAIATL
AACYNQVQVFKGAVKIRKGQAVTLMMDATNMPAVKAIHYQYMEIYHRIPSDSPSSKTR
QIQTIRTONLPNCQLISRSHYSPYLSFVFMLLAALSWQYLTLSQVTEDEVVQTGEH

>sp|P3798|FDFDT_MOUSE Squalene synthase OS=Mus musculus GN=Fdf1 PE=2 SV=2
MEFVKCLGHPEEFYNLLRFRMGRRNFIPKMDQDSLSSSLKTCYKYLNQTSRSFAAVIQA
LDGDIRHAICVFYLVLRALDTVEDDMSISVEKKIPLLCNFHTFLYDPEWRFTESKEKDRQ
VLEDFPTISLEFRNLAEKYQTVIDDICHRMGCGMAEFVDKDVTSKQDWDKYCHYVAGLVG
IGLSRLFSASEFEDPIVGEDIECANSMGLFLQKTNIRDYLEDQGEGRKFWPQEVWGRIY
KKLEDFAKPENVDVAVQCLNELITNTLQHIPPVLTLYLSRLRNQSVNFCAIPQVMAIATL
AACYNQVQVFKGAVKIRKGQAVTLMMDATNMPAVKAIHYQYIEIYHRIPNSDPSSKTK
QVISKIRTQNLPCQLISRSHYSPYLSFIMLLAALSWQYLSLTSQVTEDEVVQREH

>sp|Q02769|FDFDT_RAT Squalene synthase OS=Rattus norvegicus GN=Fdf1 PE=2 SV=1
MEFVKCLGHPEEFYNLLRFRMGRRNFIPKMDRNSLSNLKTCYKYLDDQTSRSFAAVIQA
LDGDIRHAVCVFYLLRAMDTVEDDMAISVEKKIPLLRNFHTFLYEPQWRFTESKEKHRV
VLEDFPTISLEFRNLAEKYQTVIADICHRMGCGMAEFLNKDVTSKQDWDKYCHYVAGLVG
IGLSRLFSASEFEDPIVGEDTECANSMGLFLQKTNIRDYLEDQGEGRQFWPQEVWGKYV
KKLEDFVKPENVDVAVKCLNELITNALQHIPPVITYLSRLRNQSVNFCAIPQVMAIATL

AACYNHHQVFKGVVKIRKGQAVTMMDATNMPAVKAIHQYHEIYHRVPNSDPSPASKAK
QLISNIRTSQSLPNCQLISRSHYSPIYLSFIMLLAALSWQYLSQVTEQVQREH

Aspergillus nidulans GGPPS-2

MTSDSHFPHPAIPPRISSNRMSGASTRDKAALMGNEFKDWLSKGDKLQTNLDLSKRHTR
NQSSLDGTYKDKGWSQENEEVIMGPYDYMLOHPGKDLRRQMINAFNVWLKVPSESLAII
TKVVAMLHTASLLIDDEDNSLLRRGIPVAHSIYGTAQTINSANYVYFLALQEVQKLKSP
AAIDIVVQELNLHRGQGMDFWRDITLTCPEDEYLEMVGNKGGGLFRLAVKLMQAESST
GKDCVALNVGLVFQICDDYLNLSDTTYTQNKGLCEDLTEGKFSFPIIHSIRSNPGNHQ
LINILRQRTKDEEVKRYALQYMESTGSFKHTQDVVRQLRARALQLIEEIEENSENGEQPEE
HNDGTMVRAILDKITESTLADTNTTTRDINGNCATR

Saccharomyces cerevisiae GGPPS-3

MEAKIDELINNDPVWSSQNESLISKYPYNHILLKPGKNFRLNLIVQINRVNMLPKDQLAIV
SQIVELLHNSSLLIDDEDNAPLRGQTTSHLIFGVPSTINTANYMYFRAMQLVSQLTK
EPLYHNLTITFNEELINLHRGQGLDIYWRDFLPEIIPTEMYLNVMVNMKTGGLFRLTLRL
MEALSPSSHHGHSVPFINLLGIIYQIRDDYLNKDFQMSSEKGFADITEGKLSFPPIVH
ALNFTKRGQTEQHNEILRILLRTSDKDIKLLIQILEFDTNSLAYTKNFINQLVNMK
NDNENKYLPLASHSDTATNLHDELLYIIDHLSEL

FIGURE 26

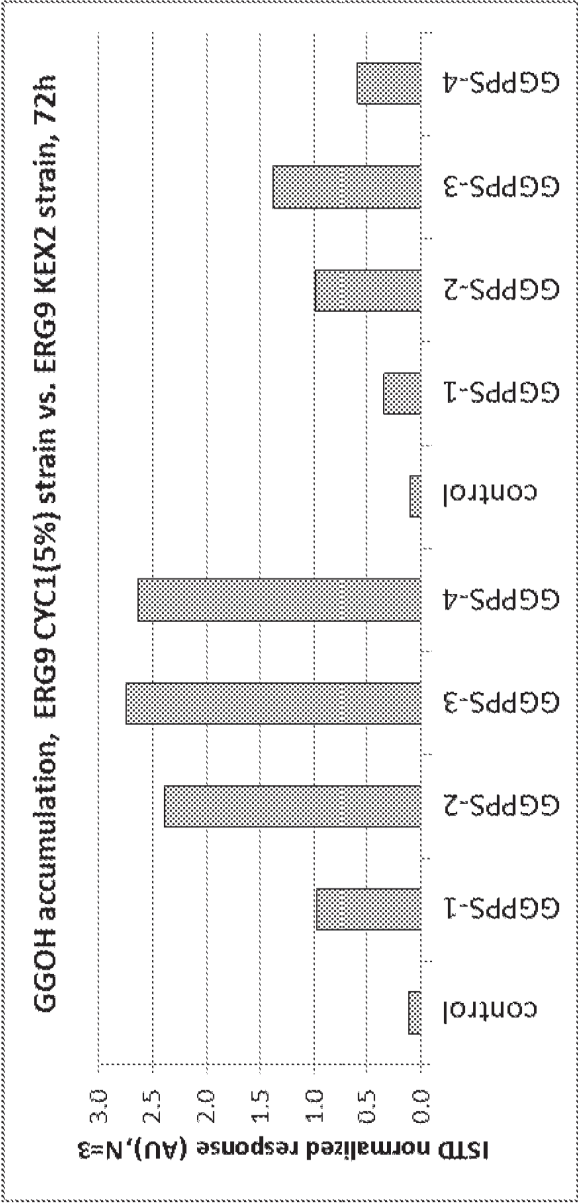


FIGURE 27

