

Fig. 1

Translation initiation factor 4E (eIF4E) - CDC33

SEQ ID NO:1

Komagataella phaffii

PP7435_Chr3-0197

MSETENAPHEKSLDDQVDFTVKHPLNSEWTLWYTKPAVHEQESWSDLLRPIISFNSVEEFWGFNSIPKASDLPV
KSDYHLFRDGIKPEWEDERNANGGKLTQFAYRKVDINELWSRGLLSVIGETIQDDDDGETEVNGIVLSIRRAALKI
ALWTKSKDEAVLRPIGERFKKVLKLTARDHIDFQPHKSGAGSKNVPSFSI

SEQ ID NO:2

Komagataella pastoris

Translation initiation factor eIF4E

Accession number (NCBI): CCA39169.1

1 msetenaphe ksllddqvdf tvkhplnsew tlwytkpavh eqeswsdllr piisfnsvee
61 fwgifnsipk asdlpvksdy hlfrdgikpe wedernangg kltyqfayrk vdelnlwsrg
121 llsvigetiq dddgetevng ivlsirraal kialwtkskd eavlrpiger fkkvklktar
181 dhidfqphks gagsknvpsf si

SEQ ID NO:3

Homo sapiens

Eukaryotic translation initiation factor 4E [Homo sapiens]

Accession number (NCBI): AAH12611.1

1 matvepettp tnpptteee ktesnqevan pehyikhplq nrwalwffkn dskstwgqnl
61 rliskfdtve dfwalynhig lssnlmpgcd yslfkdgiep mwedeknkrq grwlitlnkq
121 qrrsdlnrfw letllclige sfddysddvc gavvnvrakg dkiaiwttcc enreavthig
181 rvykerlglp pkivigyqsh adtatksqst tknrfvv

SEQ ID NO:4

Aspergillus niger

eukaryotic translation initiation factor 4E-1 [Aspergillus niger CBS 513.88]

Accession number (NCBI): XP_001395221.1

1 maavaengha pvadenaapv aaaveqkeqt ektngdsvtv fhdpennfvk hplmhewtlw
61 ftkppsgkgd nwndllkev tvfnsveefwg iynnispstse lglkadyhlf kkgvvpewed
121 pqnkhgkws ysfkdkrsvp iddlwlhaql aaigetlend gdnevngvvv nvrkgfyrvg
181 lwtrtvvgksi pgdkngtrtpa qgkeilenig rrfkevlrlk eadavefsgh tdsahsgstr
241 akakytv

SEQ ID NO:5

Cricetulus griseus

eukaryotic translation initiation factor 4E isoform X3 [Cricetulus griseus]

Accession number (NCBI): XP_027251622.1

1 mettpntnpp pveeektesn qevanpehyi khplqnrwal wffkndkskt wqanlrlisk
61 fdtvedfwal ynhiqlssnl mpgcdyslfk dgiepmwede knkrqgrwli tlnkqqrnsd
121 ldrfwletll cligesfddy sddvcgavvn vrakgdkiai wttecenrda vthigrvyke
181 rlglppkivi gyqshadtat ksgsttknrf vv

SEQ ID NO:6

Kluyveromyces lactis

Cdc33 [Kluyveromyces lactis]

Accession number (NCBI): QEU61328.1

1 msveevtqkt qdltvdeakp atvlstdskef qlkhlplsskw tlwytkppvd sseswsdllr
61 pvtsltdtvee fwaiqngipk pfdlpiksyd hlfrndirpe wedeansegg klsyqfknkr
121 hniddlwlrt llavigetid eeeseingcv lnvrkavfki alwtrttnln altniatkfk
181 qvlqlgegek lefynhrtsn sghapptall

Fig. 1 (continued)

SEQ ID NO:7

Ogataea polymorpha

uncharacterized protein OGAPODRAFT_15637 [Ogataea polymorpha]

Accession number (NCBI): XP_018212144.1

```
1 mseelnqaak dlsldekkdv tvldsreefn vkhplnswt lwytkpavds seswsdllkp
61 vvafdtveef wgifnsipka delpklsdyh lfrenikpew edpanskgr fsfqfkgkrl
121 ehgineiwtr allsvigeti eddgnevngv vlnvrkgyr vclwtksedr aalvpigerl
181 kkilmrege svefishkds ndrgakpafy v
```

SEQ ID NO:8

Saccharomyces cerevisiae

translation initiation factor eIF4E [Saccharomyces cerevisiae S288C]

Accession number (NCBI): NP_014502.1

```
1 msveevskkf eenvsyddt atpktvlsds ahfdvkhpln tkwtlwytkp avdkseswsd
61 llrpvtstfqt veefwaiiqn ipephelpk sdyhvfrndv rpewedeana kggkwsfqlr
121 gkgadidelw lrtllavige tideddsqin gvvlsirkgg nkfalwtkse dkepllrigg
181 kfkqvlkltd dghleffphs sangrhpqps itl
```

SEQ ID NO:9

Schizosaccharomyces pombe

translation initiation factor eIF4E 4F complex subunit [Schizosaccharomyces pombe]

Accession number (NCBI): NP_595451.1

```
1 madaedsrsh knegfpntsl itekldldl fgspkvkter egrparlleg lsavnaetaf
61 vkthplqhef tlwflkpptg glewsdlle iisfktveef wgifktiska smlpaksdys
121 yflkgirpew edpqnmggk wayqskhkg nldelwlymv laaigetldp tgkevtgvvc
181 nmrkgfyria vwtrncndkd vlekiglrk evlgisdket ieysahedss kagsmraktr
241 msl
```

SEQ ID NO:10

Trichoderma reesei

translation initiation factor Eif4e [Trichoderma reesei RUT C-30]

Accession number (NCBI): ETS06971.1

```
1 masaavdgan mdqqvdltsi pispngdeha getkenkddg kpvtvfhdkd nftvkhplqs
61 kwtlwfktpp sgkgdnwndl lkevitfdsv eefwgvynnv apvselalks dyhlfkagvr
121 pewedpqnkh ggkwsyqyk krnvdvdlw lqvmmmaige tleeeedgev mgvvvnvrkg
181 fyrigvwttr igksipgrgd gdvagkggrs nekgkeilmn igrfrketle lpaneqiefs
241 ghtdsahags trakakyvv
```

SEQ ID NO:11

Yarrowia lipolytica

translation initiation factor eIF 4e-like domain-containing protein

[Yarrowia lipolytica]

Accession number (NCBI): RDW25147.1

```
1 mstmaigtav svistdialq lsqmaattdn vteavanlni keeapvasap aevtetveee
61 vtpaepfstv fdsktefatv hkltpwtlw yikppspava whdlikevvs fdtvegfwav
121 ynnipklvel pirsdyaffr kgirpewecd snkaggkwfy qfkpkasnef deiwlhtvls
181 vigesletde eiakegkyn nvdgsasvvn gifvnirkag vrfnvwtkts evtdelrqmg
241 vkikqvlrlp ekeevqfqsh eesasknvks qgktrlav
```

Fig. 1 (continued)

Translation initiation factor 4A (eIF4A) - TIF2a

SEQ ID NO:12

Komagataella phaffii

PAS_chr3_0595 (sequence from GS115 strain)

MSEGIIEIDTNIETNYDQVSTFDELNLKEQILRAVFAYGYEFPSAIQSRAILPIVAGRDVLAQAQSGTGKTATF
 AISALQRLDEKLKAAQVLILAPTRELALQIKKVVLAALGMHLNISVHACIGGVDPALDVKAFNSGVQIVVGTGPRGVQ
 DMIERGAFTENVKMVLDEADEMLSTGFKEQIYTLFTLLPSTTQVVLLSATMPQEVLEVTTKFMNNPIRILVKKD
 QLTLEGIKQFYVDVGEEQYKFDCLCDLYDSISVTQAVIFCNTRRKVEELTERLTAQKFTVSAIHSELSQTERDTIM
 QEFRTGSSRILISTDLLARGIDVQQVSLVINYLPSNKENYIHRIGRGGRFGRKGVAINFVTSKDVEMMKEIEKFY
 STQIDELPSSITDLFETEEVNES

SEQ ID NO:13

Komagataella pastoris

BA75_00634T0 [Komagataella pastoris]

Accession number (NCBI): ANZ74112.1

1 msegileidt nlietnydqv vstfdelnlk eqilravfay gyefpsaiqs railpivagr
 61 dvlacaqsgt gktatfaisa lqrldeklla aqvlilaptr elalqikkvv lalgmhlis
 121 vhaciggvdp aldvkafnsg vqivvgtpgr vqdmiergaf rtenvkmvil deadeamlstg
 181 fkeqiytlft llpsttqvvl lsatmpqevl evttkfmnnp irilvkkdql tlegikqfyv
 241 dvgeeqykfd clcdlydsis vtqavifcnt rrkveelter ltaqkftvsa ihselsqser
 301 dtimqefrtg ssrilistdl largidvqv slvinydlps nkenyihrig rggrfgrkgv
 361 ainfvtskdv emmkeiekfy stqidelpss itdlfeteev nes

SEQ ID NO:14

Homo sapiens

Chain C, Probable Atp-dependent Rna Helicase Ddx48 [Homo sapiens]

Accession number (NCBI): 2HYI_C

1 mattatmats gsarkrllke edmtkvefet seevdtptf dtmglredll rgiyaygfek
 61 psaiqqraik qiikgrdvia qsqsgtgkta tfsisvlqcl diqvretqal ilaptrelav
 121 qiqkgllalg dymnvqchac iggtnvgedi rkldygghvv agtpgrvfdm irrrslrtra
 181 ikmlvldead emlnkgfkeq iydvyrylpp atqvvlisat lpheilemtn kfmdtpiril
 241 vkrdeltleg ikqffvaver eewkfdtlcd lydtltitqa vifcntkrkv dwltekmea
 301 nftvssmhgd mpqkeresim kefrsgasrv listdvwarg ldvpqvslii nydlpnnrel
 361 yihrigrgr ygrkgvainf vknddirilr diegyystqi dempmnvadl ile

SEQ ID NO:15

Aspergillus niger

ATP-dependent RNA helicase eIF4A [Aspergillus niger CBS 513.88]

Accession number (NCBI): XP_001400296.1

1 masndkglee ipeggiety devtdsfds dlpellrgv yaygferpsa iqqraikpii
 61 aghdviaaq sgtgktatfs isalqkidqe lkacqaliva ptrelaqqiq kvvvaigdfm
 121 niechacigg tnvrddmnl ragpqvvvgp pgrihdmier rvlktdqmkl fildeademl
 181 srgftegiyd ifqllpqstq vtllsatmpq dvlevttkfm rdpirilvkk qeltlegikq
 241 fyiavekeew kldtldlye tvttitqavif cntrrkvdwl tdkltardft vsamhgmeg
 301 qgrdvimkef rsgssrvlia tdllargidv qqvslvinyd lpanrenyih rigrggrfgr
 361 kgvainfvta ddvrmmreie qfystqieem pmnvadli

SEQ ID NO:16

Cricetulus griseus

eukaryotic initiation factor 4A-III isoform X1 [Cricetulus griseus]

Accession number (NCBI): XP_007645429.1

1 maaaatmats gsarkrllke edmtkvefet seevdtptf dtmglredll rgiyaygfek
 61 psaiqqraik qiikgrdvia qsqsgtgkta tfsisvlqcl diqvretqal ilaptrelav
 121 qiqkgllalg dymnvqchac iggtnvgedi rkldygghvv agtpgrvfdm irrrslrtra
 181 ikmlvldead emlnkgfkeq iydvyrylpp atqvvlisat lpheilemtn kfmdtpiril
 241 vkrdeltleg ikqffvaver eewkfdtlcd lydtltitqa vifcntkrkv dwltekmea
 301 nftvssmhgd mpqkeresim kefrsgasrv listdvwarg ldvpqvslii nydlpnnrel
 361 yihrigrgr ygrkgvainf vknddirilr diegyystqi dempmnvadl i

Fig. 1 (continued)

SEQ ID NO:17

Kluyveromyces lactis

uncharacterized protein KLLA0_A05731g [Kluyveromyces lactis]

Accession number (NCBI): XP_451255.1

```

1 msdgitdide nqiqtstydkv vysfddklk eellrgifgy gfvepsaiqq railpiiegk
61 dvlaqaqsgt gktgtfsiaa lqnidekika pggililaptr elalqiqkvv malaihmdvk
121 vhaciggtsl qedsealrgg aqiivgtpgr vfdmidrrif ktdnikmfil deademlstg
181 fkeqiynift mlpptsqvv llsatmpgdvl evtskfmkdp vrilvkkdel tlegigqyyv
241 nveeeqkykd cltdlydsis vtqavifcnt rrkveelter lrennftvsa iysdlqqqer
301 dtinkefrsg ssrilistdl largidvqqv slvinydlps nkenyihrig rggrfgrkgi
361 ainfvtnkdi gamrelerfy stqieelpss iselfd

```

SEQ ID NO:18

Ogataea polymorpha

ATP-dependent RNA helicase eIF4A [Ogataea polymorpha]

Accession number (NCBI): XP_018210402.1

```

1 maseniqevd snliesnydn vvytfdelnl kpeiirgifa ygyeqpsaiq qrailpiteg
61 rdvlaqaqsg tgktatftis alqrideklk ktgalilapt relalqiqkv vlaigmhldl
121 tvhasiggka vqedmealrn gaqivvgtpg rvydmiergy fktdaikmfi mdeadmllss
181 gfkeqiynvf rflptttqvv llsatmpgdvl levttkfmrn pvrilvkkde ltlegikqfy
241 idvdeeekyf dclcdlydsi svtqavifcn trrkvetltq qltennftvs aihsdlsqqd
301 rdtimnefrt gssrilistd llargidvqq vslvinydlp snkenyihri grgrfgrkgi
361 vainlitred temlreiemf ystqivemps sindlfd

```

SEQ ID NO:19

Saccharomyces cerevisiae

translation initiation factor eIF4A [Saccharomyces cerevisiae S288C]

Accession number (NCBI): NP_012397.1

```

1 msegitdiee sqiqtnydkv vykfddmeld enllrgvfgy gfeepsaiqq raimpiiegh
61 dvlaqaqsgt gktgtfsiaa lqridtsvka pqalmlaptr elalqiqkvv malafhmdik
121 vhaciggtstf vedaeglrra qivvgtpgrv fdniqrtrfr tdkikmfild eademlssgf
181 keqiygiftl lppttqvll satmpndvle vttkfmrnpv rilvkkdelt legikqfyvn
241 veeeeekyec ltdlydsis vtqavifcntr rkveelttkl rndkftvsai ysdlpqqrer
301 timkefrsgs srilistdll argidvqqvs lvinydlpan kenyihrigr grgrfgrkgva
361 infvtnedvg amrelekyfs tqieelpsd atlln

```

SEQ ID NO:20

Schizosaccharomyces pombe

putative translation initiation factor eIF4A [Schizosaccharomyces pombe]

Accession number (NCBI): NP_594854.1

```

1 mvdqledsvi etnydevitd fddmnlkpel lrgiyaygfe rpsaiqqrai mpilgerdvl
61 aqaqsgtgkt atfsisvlqk idtslkqcqa lilaptrrela qqiqkvvval gdlmnvecha
121 ciggtlvrrd maalqagvhv vvgtpgrvhd miqrralptd avqmfvldea demlsgfkd
181 qiydifqlp ptaqvllsa tmpgdvlevt tkfmrpripri lvkkdeltle gikqfyvave
241 keewkldtlc dlyetvtvtg avifcntrrk vdwlteqlte rdftvssmhg dmdqaqrtdl
301 mhefrtgssr ilittdllar gidvqqvslv inydlpanre nyihrigrgg rfgrkgvsn
361 fvtnddvrrm reieqfynth ieempniad li

```

Fig. 1 (continued)

SEQ ID NO:21

Trichoderma reesei

uncharacterized protein TRIREDRAFT_46619 [Trichoderma reesei QM6a]

Accession number (NCBI): XP_006963588.1

```

1 madkgledvp egqiesnyde tvdsfddmnl ksellrgvya ygferpsaiq qraimpvikg
61 hdviaqaqsg tgktatfsis vlqkidpsvk qcqalilapt relaqqiqkv vvaigdfmnv
121 echaciggtg vredmkalqd gpqvvtgtpg rvhdmigrf lktadamkmfv ldeademlsr
181 gftdqiydif qlpqtstqvv llsatmpqdv levtrfrmd pvrilvkkde ltlegikqfy
241 iavekeewkl dtlsdlyetv titgavifcn trrkvdwldt kltardftvs amhgmdmdaq
301 rdlimkefrs gssrvliatd llargidvqq vslvinydlp anrenyihri grgrfgrkgv
361 vainfvtadd vrmreieqf ystqieempn nvadli

```

SEQ ID NO:22

Yarrowia lipolytica

YALI0B20922p [Yarrowia lipolytica CLIB122]

Accession number (NCBI): XP_501158.1

```

1 madgltdids sqittnydev vtsfddlgik dellrgiyyg gfenpssiqg railpvikgn
61 dvlaqaqsgt gktatfsisa lgnidekikk pqalilaptr elahqiqkvv lafgeymkie
121 chaciggtsv aedirviqeg vlvivgtgpr ihdmieril ktdlikmfil deademlsre
181 fkdpiydift tlpettqtv llsatmpaevl ditgkfmrdp vrilvkkdel tlegikqfyi
241 dveqesykfe vlcdlyetin vsqavifcn trrkvdyltga lteadftvss mhgeteqsqr
301 dvimkaftrg ssrilittld largidvqqv slvinfdlps nrenyihrig rggrfgrkgv
361 ainfvtsedh gmlkelerfy steivemtpn iadli

```

Translation initiation factor 4A (eIF4A) - TIF2b

SEQ ID NO:23

Komagataella phaffii

PP7435_Chr3-0610 (sequence from GS115 strain)

```

MHPYTVVFARLLLGVFSTAAPNSPPPFCTSCHLGDQRLWNLAQNLAQNLACTLPPQKNFPCLVARRSKIKSFSTTFVPNK
SPFYLPIMSEGIIEIDTNLIETNYDQVVSTFDELNLKEQILRAVFAYGYEFPSAIQSRALPIVAGRDVLAQAQSG
TGKTATFAISALQRLDEKLKAAQVLILAPTRELALQIKKVVLAALGMHLNISVHACIGGVDPALDVKAENSGVQIVV
GTPGRVQDMIERGAFRTENVKMLVILDEADEMLSTGFKEQIYTLFTLLPSTTQVVLLSATMPQEVLEVTTKFMNNPI
RILVKKDQLTLEGIKQFYVDVGEEQYKFDCLCDLYDSISVTQAVIFCNTRRKVEELTERLTAQKFTVSAIHSELSQ
TERDTIMQEFRTGSSRILISTDLLARGIDVQQVSLVINYLPSNKENYIHRIGRGRFGRKGVAINFVTSKDVEMM
KEIEKFYSTQIDELPSSITDLFETEEVNES

```

SEQ ID NO:24

Komagataella pastoris

BA75_00634T0 [Komagataella pastoris]

Accession number (NCBI): ANZ74112.1

```

1 msegileidt nlietnydqv vstfdehnlk eqilravfay gyefpsaiqs railpivagr
61 dvlaqaqsgt gktatfaisa lqrldekka aqvlilaptr elalqikkvv lalgmhlis
121 vhaciggvdp aldvkafnsg vqivvgtpgr vqdmiergaf rtenvkmvil deademlstg
181 fkeqiytlft llpsttqvvl llsatmpqevl evttkfmnp irilvkkdql tlegikqfyv
241 dvgeeqykfd clcdlydsis vtqavifcn trrkveelter ltaqkftvsa ihselsqser
301 dtimqefrtg ssrilistdl largidvqqv slvinydlps nkenyihrig rggrfgrkgv
361 ainfvtskdv emmkeiekfy stqidelpss itdlfeteen nes

```

SEQ ID NO:25

Homo sapiens

translation initiation factor [Homo sapiens]

Accession number (NCBI): CAA56074.1

```

1 mattatmats gsarkrllke edmtkvefet seevdvtpf dtmgredll rgiyaygfek
61 psaiqqraik qiikgrdvia qsqsgtgkta tfsisvlqcl diqvretqal ilaptrelav
121 qiqkgllalq dymnvqchac iggtngvedi rkldygghvv agtpgrvfdm irrrslrtra
181 ikmlvldead emlnkgfkeq iydvyrylps atqvvlisat lpheilemtn kfntdpiril
241 vkrdeltleg ikqffvaver eewkfdtldc lydtltitqa vifcntkrkv dwltekmlrea
301 nftvssmhgd mpqkeresim kefrsgasrv listdvwarg ldvpqvslii nydlpnrel
361 yihrigrsgq ygrkgvainf vknddirilr dieqyystqi dempmnvadl i

```

Fig. 1 (continued)

SEQ ID NO:26

Aspergillus niger

ATP-dependent RNA helicase eIF4A [Aspergillus niger CBS 513.88]

Accession number (NCBI): XP_001400296.1

```

1 masndkglee ipegqietny devtdsfds dlpellrgv yaygferpsa iqgraikpii
61 aghdviagaq sgtgktatfs isalqkidqe lkacqaliva ptrelaqqiq kvvvaigdfm
121 niechacigq tnvrddmnaI ragpqvvvgt pgrihdmier rvlkt dqmkl fildeademl
181 srgfteqiya ifqlpqstq vtllsatmpq dvlevttkfm rdpirilvkk geltlegikq
241 fyiavekeew kldtIsdlye tvttitqavif cntrrkvdwl tdkltardft vsamhgdmeg
301 gqrdvimkef rsgssrvlia tdllargidv qqvslvinyd lpanrenyih rigrggrfgr
361 kgvainfvta ddvrmreie qfystqieem pmnvadli

```

SEQ ID NO:27

Cricetulus griseus

eukaryotic initiation factor 4A-III isoform X1 [Cricetulus griseus]

Accession number (NCBI): XP_007645429.1

```

1 maaaaatmats gsarkrllke edmtkvefet seevdtptf dtmgIredll rgiyaygfek
61 psaiqqraik qiikgrdvia qsqsqtgkta tfsisvlqcl diqvretqal ilaptrelav
121 qiqkgllalg dymnvqchac iggtnvgedi rkldygghv agtpgrvfdm irrrslrtra
181 ikmlvldead emlnkgfkeq iydvyrylpp atqvvlisat lpheilemtn kfmdtpiril
241 vkrdeltleg ikqffvaver eewkfdtlcd lydtltitga vifcntkrkv dwltekmea
301 nftvssmhgd mpqkeresim kefrsgasrv listdvwarg ldvpqvslii nydlpnrel
361 yihrigrgrg ygrkgvainf vknddirilr diegyystqi dempmnvadl i

```

SEQ ID NO:28

Kluyveromyces lactis

uncharacterized protein KLLA0_A05731g [Kluyveromyces lactis]

Accession number (NCBI): XP_451255.1

```

1 msdgIttdide nqiqtstydkv vysfddklk eellrgifgy gfvepsaiqq railpiiegk
61 dvlaqaqsgt gktgtfsiaa lqnidekika pqglilaptr elalqiqkv malaihmdvk
121 vhaciggtsl qedsealrgg aqiivgtprg vfdmidrrif ktdnikmfil deademlstg
181 fkeqiynift mlpptsqvvI lsatmpgdvl evtskfmkdp vrilvkkdel tlegigqyyv
241 nveeeykykd cltdlydsis svtqavifcn trkveelter lrennftvsa iysdlqqqr
301 dtimkefrsg ssrilstdl largidvqqv slvinydlps nkenyihrig rggrfgrkgi
361 ainfvtnkdi gamrelerfy stqieelpss iselfd

```

SEQ ID NO:29

Ogataea polymorpha

ATP-dependent RNA helicase eIF4A [Ogataea polymorpha]

Accession number (NCBI): XP_018210402.1

```

1 maseniqevd snliesnydn vvytfdeInI kpeiirgifa ygyeqpsaiq qrailpitem
61 rdvlaqaqsg gktgtfsiaa algrideklk ktqalilapt relalqiqkv vlaigmhldl
121 tvhasiggka vqedmealrn gaqivgtprg rvydmiergy fktDaikmfi mdeademlss
181 gfkeqiynvf rflptttqvv llsatmpgdv levttkfmrn pvrilvkkde ltlegikqfy
241 idvdeeqykf dclcdlydsi svtqavifcn trrkvetltq qltennftvs aihsdlsqqd
301 rdtimnefrt gssrilstd llargidvqq vslvinydlp snkenyihri grgrfgrkg
361 vainlitred temlreiemf ystqivemps sindlfd

```

SEQ ID NO:30

Saccharomyces cerevisiae

translation initiation factor eIF4A [Saccharomyces cerevisiae S288C]

Accession number (NCBI): NP_012397.1

```

1 msegitdiee sqiqtnydkv vykfdmld enllrgvfgy gfeepsaiqq raimpiiegh
61 dvlaqaqsgt gktgtfsiaa lqridtsvka pqalmIaptr elalqiqkv malafhmdik
121 vhaciggtsf vedaeglrd a qivvgtpgrv fdniqrfrf tdkikmfild eademlssgf
181 keqiygiftl lppttqvvlI satmpndvle vttkfmrnpv rilvkkdel legikqfyvn
241 veeeykyec ltdlydsis tqavifcn trkveeltkl rndkftvsai ysdlpqqr
301 timkefrsgs srilistdlI argidvqqv lvinydlpan kenyihrigr grgrfgrkgv
361 infvtnedvg amrelekyfys tqieelpsdI atlln

```

Fig. 1 (continued)

SEQ ID NO:31

Schizosaccharomyces pombe

putative translation initiation factor eIF4A [Schizosaccharomyces pombe]

Accession number (NCBI): NP_594854.1

```

1 mvdqledsvi etnydevitd fddmnlkpel lrgiyaygfe rpsaiqqrai mpilgerdvl
61 aqaqsgtgkt atfsisvlqk idtslkqcqa lilaptrrela qqiqkvvval gdlmnvecha
121 ciggtlvrdd maalqagvhv vvgtpgrvhd miqrralptd avqmfvldea demlsrgfkf
181 qiydifqllp ptaqvllsa tmpqdvlevt tkfmrdrpiri lvkkdeltle gikqfyvave
241 keewkldtlc dlyetvtvtq avifcntrrk vdwlteqlte rdftvssmhg dmdqaqrtdl
301 mhefrtgssr ilittdllar gidvqqvslv inydlpanre nyihrigrgg rfggrkgvsin
361 fvtnddvrrm reieqfynth ieempniad li

```

SEQ ID NO:32

Trichoderma reesei

uncharacterized protein TRIREDRAFT_46619 [Trichoderma reesei QM6a]

Accession number (NCBI): XP_006963588.1

```

1 madkgledvp egqieshyde tvdsfddmnl ksellrgvya ygferpsaiq qraimpvikg
61 hdviaqaqsg tgktatfsis vlqkidpsvk qcqalilapt relaqqiqkv vvaigdfmnv
121 echaciggtv vredmkalqd gpqvvgtpg rvhdmigrf lktadamkmfv ldeademlsr
181 gftdqiydif qlpqstqvv llsatmpqdv levttfrmrp pvrilvkkde ltlegikqfy
241 iavekeewkl dtlsdlyetv titqavifcn trrkvdwlt dktardftvs amhgmdmdaq
301 rdlimkefrs gssrvliatd llargidvqq vslvinydlp anrenyihri grggrfgrkg
361 vainfvtadd vrmmreieqf ystqieempn nvadli

```

SEQ ID NO:33

Yarrowia lipolytica

YALI0B20922p [Yarrowia lipolytica CLIB122]

Accession number (NCBI): XP_501158.1

```

1 madgltddids sqittnydev vtsfddlglk dellrgiygy gfenpssiqg railpvikgn
61 dvlagaqsgt gktatfsisa lqnidekikk pqalilaptr elahqiqkvv lafgeymkie
121 chaciggtsv aedirviqeg vlvivgtpr ihdmierri lktlikmfil deademlsr
181 fkdpiydift tlpettqtv llsatmpaevl ditgkfmrdp vrilvkkdel tlegikqfyi
241 dveqesykfe vlcdlyetin vsqavifcn rrvkdyltqa lteadftvs mhgeteqsq
301 dvimkaftrg ssrilittl largidvqqv slvinfdlps nrenyihrig rggrfgrkgv
361 ainfvtsedh gmlkelerfy steivemtpn iadli

```

Translation initiation factor 4G - eIF4G - TIF4632

SEQ ID NO:34

Komagataella phaffii

PP7435_Chr1-0352

```

MSNKNVDTPDNGSGGNSPASVPHYPQQHRANYNPHKNHTNHGGNNGGAQGYMQHQQQQQMYYNQMGYPMPYPNA
GGNFNQNVYLVYNTMVPPTPYGAGAPGVPPSYVSPSPAFASPQKTHVPVSLRTPSGEPINLVPAGQTSSSSSNA
SPHVKSTSIPPSAVTSPAASIAPTVAAPTAAAGSEDAKTRDAFKQMVQALAAKKKAAAAAKAEGKTDKKESVE
SKPKSAVQPETKPKVEVKLEPKPKVEVKLEPKPKVEIKLEPKPVSEAKPETKTVPEVEEPSTVATIKSESTPIEEL
KPKVEVTEPEVASKPEQVESAEQEPANDVKPSEVSAKEVQESKEESLKQPKEALNDIKSTEESENKKFSLPEEESNDHQ
TSLQIQGVKDTEQSQEFEEPPQGTQETQETQETQETAKPLEAQAPQEIQETQEPPEEVQETPETSETVSSVTEEVSP
APAPGLPLSKFLDRIRKAPAINPFFKHEYSGDFHGPDESLSGLKTIKYDPGFLIQFATFNHIDEDFINDVGSKVV
IPVNTRRDNFSKSMSTSNFKSGSRNMGSNFRNDFGSRNNSKQGSKKKNSQKSQSRRGPKYSNEEPAEPAVPAEL
PNRAASKWVPRSLRNKEAEVKTAPDGSIIIVSEEDIKRKTKSLLNKLTLFFDDISNDIIALTQKQWEDDVKTLKQ
VIESIFAKACDEPYWSSMYAKLCAKMKCDTPPEIKETNEKGNFTTGGDLVRRVLINRCQTEYQKGWIDQLPTNEDG
SPLEPELMSDEYYQAAAKRRGLGLVKFIGQLYQLNLLSDNVIFHCVLGQSKETDNPS EDTMENLVQLEAVGYRM
EMSSEPKIRRVLDLAFANIKAVDANKIPSRIRFMLMDLMDLRNGWATTEENAGPKTIQEIHLEAEIKRSEEEERE
KERDRHSRRGRYNDKSPNSRQKSSWGSSVSPNDIRNAGVVRNSDKSLGPVNSFSRGKSSRQASKQNSNDFTSVPI
TSSGSRTALNHSNGDSSTSLNEPQSREYSKKSES RPAPSNIFAALEDHHDHDEHGHEQEENKHSKDLONGEAEAPED
DNVGEIAQADENVQGEQEEREEREEREQEVE

```

Fig. 1 (continued)

SEQ ID NO:35, *Komagataella pastoris*
 BA75_01499T0 [Komagataella pastoris]
 Accession number (NCBI): ANZ73389.1

```

1 msnknvdtap dngsggsnsp asvphypqgh rsynphknh tnhggnnga ggymqhqqqg
61 gmyynqmgyp mpypnaggnf nqnvylypyn tmvpgtpyga gapgvppsyp vsptpafasp
121 qkthvpvslr tpsgepinlv paggtsssss nasphvksts ipasavtspa asiaptvasp
181 akpaagseda gktrdafkqm vqalaakkka aaakaegktd dkkesveskp kaavqpeakp
241 kvevklepkp kvevklepkp kveiklepkp vseakletkt apeveepstv atiksetkpi
301 eelkptvevs lqveaepkqf esagepaadl nssevsakev ketkeealnq pkealndiks
361 aelnekstp qeeesndqgt lsqsqeaket eqsggseepq etgeaqevqv tqttpeggeg
421 ketqepgeaq eaqevqeqtp etsetvssvt eevspapapg lplstfldri rkspaisnfp
481 khqysgdhfg pdeslsglkt ikypdgfliq fasfnfhide dfindvgskv vipvntrrdn
541 fksmstsnf kgsgrnmgsn frndfgsrnn skqgskknk sqksqsrngp kysneeeepae
601 pavpaelpnr aaskwvprsl rnkeaevkta pdgsvivsee dikrktksll nkltleffdd
661 isndiialtk qaqweddvt lkqviesifa kacdepywss myaklcakmc kdtpeiket
721 nekgnftftg dlvrrelvlnr cqteyqkwi dqlptnedgs plepelmde yykqaaakrr
781 glglvkfigq lyqlnllsdn vifhcvlgqs ketdnpsedt menlvqllea vgyrmmsse
841 pkirrvldla fanikaivda nkipsrirmf lmdlmdlrn gwatteenag pktiqeihle
901 aeikrseer ekerdrhsrr gryndskpns rqksswgssv vspndirnag yvrnsdkslg
961 pvnsfsrgks srqaskqns dftsvpitss gsrtalnhsg dsstslnepq sreyskkses
1021 rpapsnifaa ledhdhdheh gheeenkhs kdlqngeaet penvnveeia qadenvggqe
1081 qdeeqdeeqd eereqeeeqe v

```

SEQ ID NO:36, *Homo sapiens*
 eukaryotic translation initiation factor 4 gamma, 3 variant [Homo sapiens]
 Accession number (NCBI): BAD92356.1

```

1 vkkihhrhrls patekgaass gegdkdpppp ahediavkgt ataagtgppt gaaaaaaaaa
61 ppphpaaaaa aaaaaaaaaa avppphpnir alqtqapqqi prgpvqqple driftpavsa
121 vystvtqvar qpqtptpspy saheinkghp nlaatppgha sspglsqtpy psgqnagptt
181 lvypqtptm nsqpttrspg gfrpiqffqr pqiqpprati pnsspsirpg aqtptavyqa
241 nqhimvnhl pmpypvpqgp qycipqyrhs gppyvpppqq ypvqpppggp fypggpggdf
301 pnaygtpfyp sqpvysapi ivptqqqppp akrekktiri rdpnqggkdi teeimsgggs
361 rnptppigrp tstptppqlp sqvpehsppv ygtvesahla astpvtasnd qkgeekpkpd
421 pvlkspspvl rlvlsgekke qeggtsetta ivsiaelplp pspttvssva rstiaaptss
481 alssqpiftt aiddrcelss predtipips ltsctetsdp lptnendddi ckkpcsvapn
541 diplvssstnl ineingvsek lsatesivei vkqevlpltl eleilenppe emklecipap
601 itpstvpstf ptpptppasp phtpvivpaa attvsspsaa itvqrvleed esirtclsed
661 akeiqnkief eadgqteeil dsqnlssrrs pvpqaiatv pktwkkpkdr trtteemlea
721 elklkaeel sidkvleseq dkmsqgfhe rdpsdlkkvk aveengeeae pvrngaesvs
781 egegidansg stdssgdgvt ffpkpeswkp tdteggkkyd reflldfqfm paciqlkegl
841 ppisdvvldk inqpklpmt ldprilprgp dftpfadfg rqtppgrgvp llngvssrsq
901 pgqrreprki itvsvkedvh lkkaenawkp sqkrdsqadd peniktqelf rkvrslnlkl
961 tpqmfnqlmk qvsgltvde erlkgvidlv fekaidepsf svayanmcrc lvtlkvpmad
1021 kpgntvnfrk lllnrcqkef ekdkaddvfv ekkqkelea sapeertrlh deleekadka
1081 rrrsignikf igelfklkml teaimhdcvv kllknhdees leclrltlt igkdldfeka
1141 kprmdqyfnq mekivkerkt ssrirlmlqd vidlrlcnwv srradqgpkp iegihkeaki
1201 eegeeqrkqv qlmtkekrp gvqrvdegw ntvqgaknsr vldpskflki tkptidekiq
1261 lvpkaqlgsw gkgssggaka setdalrssa sslnrfsalq ppapsgstps tpvefdrst
1321 ltsrgsmgre kndkplpsat arpntfmrsg sskdlldnqs qeeqrremle tvkqltggvd
1381 vernsteaer nktresakpe isamsahdka alseeelerk sksiideflh indfkeamqc
1441 veelnaagll hvfvrvgves tlersqitrd hmgqllyqlv gseklskqdf fkgfsetlel
1501 addmaidiph iwlylaelvt pmlkeggism reltiefskp llpvgragvl lseilhllck
1561 qmshkkvgal wreadlswkd flpegedvhn fllegkldfi esdpscssea lskkelsaee
1621 lykrleklii edkandeqif dwveanldei qmssptflra lmtavckaaai iadsstfrvd
1681 tavikqrvi llkyldsde kelqalyalq asivklqpa nllrmffdc l ydeeviseda
1741 fykwesskdp aeqngkgval ksvtafftwl reaeeseedn

```

Fig. 1 (continued)

SEQ ID NO:37

*Aspergillus niger*eukaryotic translation initiation factor subunit eIF-4F [*Aspergillus niger* CBS 513.88]

Accession number (NCBI): XP_001397987.2

```

1 mssippksgl qpqqgststg aassahtssp svvgksptps ststtaprsy anatkksatd
61 staapvtvvg psqhgkstta spvsgkpmqq sqtpgvtivn gapapasqgd htrkpsvtit
121 sagtsgfipn ggqgrpnsl qfgfaanqqs spnmgnpavl anqpqsglgv tppmnpvrvts
181 pqtspspipq passggrppp ssyqsqgnvp nfgsfgdagd nmrapqapl gpgaqrsthlr
241 ressqsthsd msnhmgggpg grggyhhggg rgrgfsqssy qgqmpyspna sfrstpnqpr
301 ggpnmapqfp nqprmpfpn sphqasrspa lanahpatpq mnqvpmahpq mapqpyyggp
361 mppqpvrphp lqkparrgga lnrkkgsprt pitprssvtp nphitlppnl apesghfery
421 ltmnksgqgy ptfdpaygfy pnynmggvpv mppspqprpg ipynaapym qnqypvqpaa
481 gatalrsrps hsndrpgssl gggqppagap gpahahnasr ssnsapqks sfviptakkn
541 avvikdpgsg avktfekapa sparvtpspv kiatptstpp prtgsadht rtdskstkt
601 eekkkelrda vhgkieqdri ekeeaerrka eeeaaqkkke eeeaaarkka eedaaqkkke
661 eeeaaqkkaa deeaarkgle dlsldkdpae tkaeepskpa aapapadddd didydaiere
721 maeieakera aeeayyarkk aekeekerke keereayean mknaereaea leearakkre
781 agveeeskdk dlfaslkkgg fpateastpg dsgaatpasd msmgppskpa sankrekpaa
841 lklettkave ppqpsaamka lhsarfvdld skitypasiv spnpalnasa padrkfhynk
901 efillqfqsdf kekpsidwda rvretvgdsd ssrpqsartp mmggrnssrp gaapfqmgnf
961 gqtpsrasnl ppgttsaerf amsnaartas mgnpfgnfgr pglgigapgi srtnsasamh
1021 pgmpssprvg ssnrsgtrtg skrdkqnakk eeemaksmp tagmevkalq vsstgwkprs
1081 iggapaaaag ptpggathmp pdmvqrkvka alnkmtpenf driagglllei vsqskdetdg
1141 rtlrqviqlt fekatdeahw asiyakfckr mlesmsaeik denihdkngn vvaggsflrk
1201 yllnrcqeef ergwkvnlpp kpegeteeaa mmsdeyykaa aakrrglglv kfigelyklg
1261 mlterimhec vkklvdyeqm pdeaeveslt slrtigasl dvsekghafm nayfhriglm
1321 iqtpglpsrl rfmlmdivdl raarwqskda dkgpktiqqi reeaaraqge aemerlrqqa
1381 nrgsrpqlgr gdarsysgyg nqappqdfas skvgsddlrr lrtsrqtnqp msfgpssmlg
1441 srsnsgrknl gpqgnlvrgs ddsaastrtg tptgkkekke aassinafsa lasledrdnm
1501 atspnsnpts plltksqpav errpsktpsk dgeas

```

Fig. 1 (continued)

SEQ ID NO:38

*Cricetulus griseus*eukaryotic translation initiation factor 4 gamma 3 [*Cricetulus griseus*]

Accession number (NCBI): ERE82257.1

```

1 mgrpqgfrp rggpvqqpled rifpptvsav ystvtqvarq pggptpspys aheinkghpn
61 laattppghas spglsgttyp sggnagpttl vvpqapqtmn sqpqtspys rtpvihatdn
121 wkrrkvleqt pvyrslagrg wikycifaag prpahhqqffq rpqigpprat ipnsspsirp
181 gaqtptavyq anghimmvnh lmpypvpvpgg pqycipqyrh sgppvyvgppg qypvqpppgg
241 pfyppgpggd fanaygtphy psqpvyqsap iivptqqqpp pakrekktir irdpnqggkd
301 iteeimsggg srnptppigr ptstptppqq lpsqvpehsp vvygtvesah laastpvtaa
361 sdqkgeekpk pdpvfqsps vlrlvlsgk keqasvpvet pageptpelp rtsspttlap
421 larssvpspt caalssqplf ttaledkrdl sssgedalpv psptsctvts dpsltddkni
481 ckkpcsvaph diqlistpnl mnemngvgek lpagesivdm akgetlpltl eleilehage
541 emkvectpta itpsmltsfs papltppssp capvisatta rspvaatvvq rvaeegeesir
601 tcisedanet qykteaeag qteeivdsqn lnsrrspgv qtailapktw kkakdrsrtp
661 devleaeap kadeelpddr vlepetdkms qgfherdss alkrvkagee ngeeveplrn
721 gaespsegeg gdansgstds sgdgipfpfk aeswkpadte gkkqydrefl ldiqfmpaci
781 kqpeglppis dvvldkingp rlpmtldpr ilprgpdftp afadfgqrtp ggrgvpllnv
841 gprsqpgqr reprkiitvs vkedvhlrka enawkpsqkr dshaddpesi ktqelfrkvr
901 silnkltpqm fnqlmkqvsg ltvdtterlk gvidlvfeka idepsfsvay anmcrcvltl
961 kvpmaekpgn tvnfrklilln rcqkefekdk adddvfekkq keleasape ertlhldele
1021 eakdkarrs ignikfigel fklkmltea mhdcvklk nhdeeslecl crllttigkd
1081 ldfekakprm dqyfnmeki vkerktssri rfmlqgvidl rlcnwvsrra dggpktieqi
1141 hkeakieege eqrkvvqlmt kekrrpgvqr vdeggwntvq gaknsrvldp skflkitkpt
1201 idekiqlvpk aqlgswgkgs sggakasets alrssassln rfsplqppap sgspsttple
1261 fdsrraltsr gsmgreksdk plptgaarpn tflrgsskd ldnqsqeeqr remletvkql
1321 tggvdteras seadrsktre lvksemcavs ppdkpalsee everksksii deflhndfk
1381 eatqcieels agglplhfvk vgveftlers qitrhdmghl lyqlvgsekl skqdfkkgfs
1441 etleladdma idiphilwyl aelvtplmke ggismreliv efskpllpvg ragvllseil
1501 rllckqmskh kvgalwread lswkdflpeg edvhffileq kldfietess cssealskre
1561 lsaeelsqrl eklimedkad derifdwvea nldesgmssp tflralmtav ckaaiiadcp
1621 tfrvdtavik qrvpillkyl dsdtekelqa lyalqasivk ldqpanllrm ffdcllydeev
1681 isedafykwe sskdpaeqng kgvalksvta fftwleaeae esedn

```

SEQ ID NO:39

*Kluyveromyces lactis*Tif4632/Tif4631 [*Kluyveromyces lactis*]

Accession number (NCBI): QEU58435.1

```

1 mgeptseqqp aveapvvqee tisspqknsq yvknatagsa prngkydgnr knsrpynqrg
61 nnnnnnngss snkhyqkynq paygvpaygi pnyavsaeyn plyynqyqqq qglyaaaayqt
121 pmsgggyvap vvspaavsak pakveitnks gehidiasia hphthshsq hsravpvvvp
181 panaavtaaa sssvpsasp avkvqspaan dkeqspakpe epkkttlivn dfleqvkrk
241 aalaakkave ergseeqkep vvgadtdanv daktgstaad stkseeaqse sqektkeeap
301 aepkpltlae klrlkrmeaa kqasakteel kteeskpeel kteepkaeep kaeepkleep
361 kpeepkteep kteepkteep kseepkseep ksdepksdep kaeepkteel kteepkteel
421 kteepkteet kpeepktees ateanakege papagsvetp advetkpee teveddgkit
481 mtdflqklke vspvddiysf qypsdiptpn drykktisiky aygpdflyqf kdkvdvkydp
541 awmaemtski vipppkpgss grgedrfskg kvgsrsegr sgrrsnskks skrddrksnr
601 sytsrkdrrer freeeveepk vevaplvpsa nrwvpkskmk ktevklapdg telydaeeas
661 rkmksllnkl tlemfepisd dimkiansk weekgetlki viqqifnkac dephwssmya
721 qlcgkvvkdL ddsikdsetp dktgshlvh ylvqrcqtef qtgwtdqlpt nedgtplqpe
781 mmsdeyykma aakrrglglv rfigflyrsn lltsrmvffc fkrmlkdign sptedtlesv
841 celletigeq fegariqvta eaviegsll dtlfdqiknv iengdissri kfkldidivl
901 rekrnwnskn kndgpkitia iheeealkra leererendr hgsrggsrrm nsernsrrd
961 fssshshnq nrdgftttrs ssvrysepkk eeqaptptks sagaanmfda lmdaedd

```

Fig. 1 (continued)

SEQ ID NO:40

Ogataea polymorpha

uncharacterized protein OGAPODRAFT_16130 [Ogataea polymorpha]

Accession number (NCBI): XP_018211346.1

```

1 msnpnsvppt gnahhhynsn yhknnrnngs ysnysqshpq ppyyqqggwt gymtqpyqaa
61 apgyaysyya pspggvsgga yyappppgai aisptlpnsa thtpvtkpvnsa lsiknkngea
121 vqfngkghskt aslssssssss tpvlskvnsa lpavppstss tasvpmsspsk kdssspakpaq
181 pevkkaneef kakfleqlrl knskatsekk eeksqptkev ekeskpdass kkseppkelp
241 geiskeapkp aekpetskee pkepaesakp akpavpsasi eqqkeetkei ektqpnkqqa
301 ekpeelqavp espvesiqpk pvaasapsvd evkeeaqptp vkeepinads eqteggkssp
361 vadsqedtkd etaepleaas pgidktfdmt rffaqiskat pikdiyaafpy keggekpdlr
421 lqnsapknyys ppflqlqfqp cqfpvddewd akvrsiidlv gakesrdrgr fsrsgsqfgr
481 ggagssnamp rsatlrgefgr srngsrqgsk kkmglgsrek sqrggskkpr rlddekdnkd
541 eeteapkvqv evkpleksan rwwprsqrnk ekevryapdg vtvilleddv eakvkshlnk
601 ltldnfdtis aeiiqianqa rweddvktlk qvieltfaka cdehlwssvy akcltltlctd
661 isedvedksq prpeqvprkr yssglvrkl1 lsrgcaeyek gwsdklptne dgtplepemmm
721 sdeyyklaaa krrglglvrf igelyl1nll sekiiikclv dqlsdsdps edtletltql
781 ltttgaklds snrgpldfaf ekinhiinn1 qlssrikfkl mdlcdlranr wadakinagp
841 ktiaaehqea eakrlqeerd airrrkadpr sgssrsnssr snskwgnsnq issndiskvg
901 vvrnssersl gpfnkqkskl sssklsqqqf ttvpsagsan ssgassgqkr agtppdepre
961 asrnifaalg dhededsdre aehhepepeh egadissahh egvsavadse sq

```

SEQ ID NO:41

Saccharomyces cerevisiae

translation initiation factor eIF4G [Saccharomyces cerevisiae S288C]

Accession number (NCBI): NP_011678.3

```

1 mtdetahptq saskqesaal kqtgddqges qqrgytnyn ngsnytqkqp ynsnrphqqr
61 ggkfgpnryn nrgnyngggs frgghmgans snvpwtgyyn nypvyyqppq maaagsapan
121 pipveekspv ptkieittks gehldlkeqh kaklqsqers tvspqpeskl ketsdststs
181 tptptpstnd skasseenis eaektrrfi eqvklrkaal ekkkrkeqleg ssgnnnmpmk
241 ttpenveekg sdkpevtekt kpaeksaep evkqetpae geggekgqik eestpkvltf
301 aerlklkkqg kereekteg enkevvpqee tksaiesapv ppseqvkeet evaeteqnsi
361 desattpaip tksdeaeaev eaeagdagtk igleaieett tdetddgtnt vshilnvlkd
421 atpiedvfsf nypegiegpd ikykkehvky tygptfllqf kdklnvkada ewvgstaski
481 vippgmgrgn rsrdsgfrgn nssrgdhfrn tsvrnmddra nsrtsskrrs krmnddrrsn
541 rsytsrrdre rgsyrneekr eddkpkeeva plvpsanrww pkfkskktek klapdgkte1
601 ldkdeverkm ksl1nkl1tle mfdaisseil aianisvwet ngetlkavie qiflkacdep
661 hwssmyaq1c gkvvkelnpd itdetnegkt gpk1vlhylv archaefdkg wtdklptned
721 gtplepemms eeyyaasak rrglglvr1fi gflyrln1lt gkmmfecfrr lmkdltdsps
781 eetlesvvel lntvgeqfet dsfrtgqatl egsqll1d1slf gildniiqta kissrikfkl
841 idikelrhdk nwnsdkkdng pktiqqihee eerqrqlknn srsnsrtrnn ssnrhsfrd
901 appaskdsfi ttrtysqrns qrapppkeep aaptstatnm fsalmgesdd ee

```

Fig. 1 (continued)

SEQ ID NO:42

Schizosaccharomyces pombe

translation initiation factor eIF4G [Schizosaccharomyces pombe]

Accession number (NCBI): NP_594602.1

```

1 msskppsntp kfsyaralas sqsnksnstk asenntatae kgavkpsgve ptntsranag
61 kkttestgkit seadtekyns skspvnkegs vekkssekss tnnkpwrgrdn tskspsanssa
121 ertssqhqqp etsssqigkdn aapvenvnek stsgetappv stvpiqfgsi trnaaipskp
181 kvsgnmqns gvssysksq svnsstsnph phteepvaak peasstatkg prpttsasnt
241 ntspangapt nkpsdintt dpatqttqvs asnspslsgs stpsntssrs nrqnhgnfse
301 krhydrgns hpsynkyshy qhgfnynnsq nnrnesghpr frnsrrnynn ggayptymsn
361 grsangsprn npqnvngst piqipvslqt pyggvygqpq yivdpnmvqy gpilqpgyvp
421 qyyvyhqtp ytnfnpmnsr sgsvsdqvv espnsstlsp rngfapivkq qkkssalkiv
481 npvthtevvv pqknasspnp setnsraetp taappqisee easqrkdaik laiqqriqek
541 aeaeakrkae ekarleaeen akreaeeqak reaeekakre aeekakreae ekakreaeen
601 akreaeeekak reaeekakre aeekakreae ekakreaeek akreaeeekak reaeekakre
661 aeekakreae ekakreaeen akreaeeekv reteenakrk aeekgkread knpeikssap
721 lasseanvdt skqtnatepe vvdktkvekl kasegkstss lsspshtstss krdllsgles
781 lslktnpkse qclesllnsq fitdfsalyv pstikppste ealkagkyey dvpfllqfqs
841 vytdkpmkgw dermketvas afsdkssrgm ysssrqssrs gsnthshagp gfggppserkg
901 isrlgidrgf sssgagfgsg snyksapsrg vshhghggms gshrgsgrgs rrgggerdkp
961 dpssltipvd qvapllqslan rwqpkkltek paetkgedee allppevvqr kvkgslnkmt
1021 lekfdkisdq ileiamqsrk endgrtlkqv iqltfekatd epnfsnmyar farkmmdsid
1081 dsirdegvld knngpvrggl lfrkyllsrc qedfergwka nlpsgkagea eimsdeyyva
1141 aaikrrglgl vrfigelfkl smlsekimhe cikrllgnvt dpeeeeeiesl crllmtvgvn
1201 idatekghaa mdvyvlrmet itkipnlpsr ikfmlmdvmd srkngrwavkn evekgpktia
1261 eiheeaerkk alaesqrps grmhgrdmnr gdsrmggrgs nppfsssdws nnkdgyarlq
1321 qggrglksqt qgshgptsls smlksgsvr tpsrqnsalr reqsvrapps nvavtsansf
1381 elleehdhdn dggqkdsnsk tss

```

SEQ ID NO:43

Trichoderma reesei

hypothetical protein M419DRAFT_88939 [Trichoderma reesei RUT C-30]

Accession number (NCBI): ETR98443.1

```

1 mtspanqgtp ipatttnapt apsyasaaga pkkpaqapvv asgsnaplvv gssaptasag
61 nakaasspsp vngkqpvpaa vptvarssfn gsgpdhfrkg svtmaangps gfvpnggsia
121 gaksqigqfy dspamghstp hvghsapipv pdannhhrvp spahspapip qpsasggrpp
181 sslqqptnqm tfgslgsdgd rhmrqgsvpp npnlangsga hfrresshvg qsdssrgnfg
241 pqnnrnnsf nphatqfnng mgfppqngfr tpgqnrgmmp sfqpnrgyp nspqparspa
301 lanlvpslpg tpmmpatmq pnmampnapa yhypppipas hqqqvqypfp avaakmhskp
361 sykksksrre ydrhfsqgsp ldtahdapt aawpqqqqq qhhqrrfsgs ggradtqwra
421 hdknnmrpis sgikdgpss pslplvdapl shpnsldlsp enggfelwlt wknqtygfpp
481 qgmnyggqpm tvpyyggamp ymgppqpaga aynqqfappt fhqgshmsr tpsqperpas
541 asqpsqpaiv ssnpkalna anpnsfvkpr ksaavqiknl agevvdstf kqpaspvpaq
601 trtppvias tppqqkldap shartdsqsp sktakeiqee lkakiqaeir aasekaakqe
661 agakagaaed asakkaeaae akveekteak aeagpevkaq peatpdkpe epkaepaetk
721 aeakaepgpe aaakaeapap ekaaekpae paekatekaa ekpaetpaea pkaeppaks
781 eepaksdepe edeleriiere meaeakrea qeehrkkre aeakaqkae eanqknaee
841 adrklreger emerleeeke rrreaaessg ktlsvaeala qmrsgnapdk kdepasveav
901 adkladfsia gdksgdaag kekrvakpaa lnlapiktsa vepqpasaal qslksarflq
961 vmdqdiypeg ikspnpalna avakkgtkf ydaaflqlfq kvfteqpslq fhqqikqlig
1021 dgdgnrsasr gqtpssarqg srggggfgpg sfgvpggrtl ppgttsadrm aiasgalprp
1081 qvnpmasfqr pgaafpass msrtssqnmr nnlpnsprqs srstrgsrrn dmaakeaqaa
1141 ktmpltagqe vkpitvsasg wkasslgkap saatpggql mdpemvqrkv kaalnkmtp
1201 nferiseqil aiasqskdet dgrtlrqiql ltfekatdea hwasmyakfk krmltmspd
1261 irdenikdkn gnvvsqgnlf rkyllnrcqk efergwtnl peapkdteeg ekktegeaml
1321 sdeyyaaaaa krrglglvqf igelfklqml terimhecvl klvdykgmpd eaeieslckl
1381 lrtiganlde tdkgrplmdi yfgrigqmad lpelqsrnrf mlmdvidlrk arwvskegnk
1441 gpktleevra eaeaqsaaka qeaartsqrg gpgggrrpag rgdargfsgs fnqpsnqvqm
1501 dillrrlkgs vntsssnvt lgptsmssr snsgrrrfpg galgrggdds nsssratpg
1561 ptssnafsl1 atmddghpas ppsagpsplm akavpatdkk dne

```

Fig. 1 (continued)

SEQ ID NO:44

Yarrowia lipolytica

Conserved hypothetical protein [Yarrowia lipolytica]

Accession number (NCBI): VBB85048.1

```

1 msytnngay sprnnqpgtp nqqqqhmays paamynggyg yvdmtdmqgg ypppqqegmsp
61 gaqyaqpygs phhthsplti mdpvthrpi vtpgsntgtpt gpnatlaspk pnsprvahrr
121 tdssnpagds ikanfmaqv1 aakakkeaae kgtekkeevk eekkeepkpe tkpeeevkke
181 epkkeepked qkeevkeelk etpkedakpv eaekpkedvk eeakpveaek paetkleees
241 sapaqeekve kkveiveppk epsaegepte dasestpnpv ekiaetmrk krfahswna
301 tpltaeieitt kvtypesvtk slmevspapt frysldfllq fqplvteqpp sglegakait
361 edckgskapr assrgsgggp nqqmgfsfnf grggpgggpsg rgmprmgsss nmrgdhgghh
421 qshrggdrgd regggrstrn nssrnrnps nreryppqga ngeiieyvp ppksenswaa
481 krlaakkets tpppaesepa eeetprippd vvqkkvnsm nkmtlekfdk isdqileias
541 gskhehdgrt lrqvleltfa katdeahwas myarfckktm vtvdpdidydd dvkdkgneyv
601 rggalyrkyl lnrcqdefek gwedklptne dgspmdpdna tpeydkavva krrglglmvf
661 igelfmnlml senimrsclt rlispkdgas eetieslckl tstvglrmdq spnskefmdm
721 ffqrltdtkn sgtlpsrlkf kildllemra ggwkqsndqg pktiseiqde aaaaakekqqq
781 qqdrdrshrg gnrrsghpmg grsesyrngn erdrntnsln tgdlkqfgki rdsstntfgp
841 pssnmgrggs nasrnssqts lqqqgrqns spgiplskds srasskrgni fehlmgdgeg
901 een

```

Poly(A) binding protein - PAB1

SEQ ID NO:45

Komagataella phaffii

PF7435_Chr2-1212

```

MSVDTKÉVQESLDQLNISESPATAVTEETPATSTEAEEESNESSTQASETLASLYVGELDPTVTESDLYEFFSPIG
SVNSIRVCRDAVTKRSLGYGVNFHSSQAAGERALEELNYAEIKGVRCLMWSQRDPSLRSGSGNIFIKNLDPALIE
NKTLDHDTFSSFGKVLSCKVATDENGNSKGFGFVHYESDEAAQAAIENINGMLLNGREIYVGPPLAKKDRESRFQEM
IKNYTNVFNKFDTESTEDELRELFESYGPITSIHLQVDSEGHNGKGFVNFVAEHDDAVKAVEALNDKEYKKGKPLY
VGRAQKKNERVHELTKKYEADRLKQLKYQSVNLFIKNLDESIDDARLEEEFKPFGTITSAKVMLDENGKSRGFGF
VCLSTPEEATKAISEMNQRMVANKPLYVALAQPKAIRRSQLAQQIQARNQMRMQQAGPGIPNQFVQPIFYGQQPG
MLPPGARVPPMGNIPOFAGMPRPFGFPQGFPRMAPNGQPMFVYGQPVFNNGPQQRGYYPNNRQNKNRNQKEEG
NSLAAILPQLPLEQQKRVLGEEELYPKVATNKTQDPEAAGKITGMILDLNQEILQLEDEELFNTHFNEALQAYE
EYKSKTVEQQQQAQAN

```

SEQ ID NO:46

Komagataella pastoris

BA75_05166T0 [Komagataella pastoris]

Accession number (NCBI): ANZ77385.1

```

1 msvdtkevqe sldqlnises patavteetp atsteaaees nesstqaset laslyvgeld
61 ptvtesdlye ffsfigsvns irvcrdavtk rslgygvnf hsqaaageral eelnyaiek
121 vrcrlmwsqr dpslrsgsg nifiknlpa ienktldhtf ssfgkvlsc vatdengnsk
181 gfgfvhyesd eaaqaaeni ngmlngrei yvgphlakd resrfqemik nytnvfvknf
241 dtestedelr elfesygpit sihlqvdsge hnkfgfvnf aehddavkav ealndkefkg
301 kplyvgraqk knervhelk kyeadrlek qkyqsvnlfi knldesiddt rleeeekpfg
361 titsakvml dengkrgfgf vclstpeeat kaisemnqrm vankplyval agpkairrsq
421 laqqiqarnq mrmqqqagpg ipnqfvqipf ygqpggmpp garvppmgng ipqfagmprp
481 gpfpqqgfpr mapngqmpv ygqpvfngng pqqrgyypnn rqnknrnqke egnsllaailp
541 qlpieqqkrv lgeelypkvv atnkaqdp eaagkitgmild ldnqeilql1 edeelfnthf
601 nealqayeey ksktveqqq adaqan

```

Fig. 1 (continued)

SEQ ID NO:47

Homo sapiens

polyadenylate-binding protein 1 [Homo sapiens]

Accession number (NCBI): NP_002559.2

```

1 mnpsapsypm aslyvgdlhp dvteamlyek fspagpilsr rvcrdmtrr slgyayvnfq
61 qpadaerald tmnfdvikgk pvrimsqrd psrlksgvgn ifiknldksi dnkalydtfs
121 afgnilsckv vcdengskgy gfvhfetgea aeraiekmng mllndrkvvf grfksrkere
181 aelgarakef tnvyiknfge dmdderlkd1 fgkfgpalsv kvmtdesgks kgfgfvsvfer
241 hedaqkavde mngkelngkq iyvgragkkv erqtelkrkf eqmkqdritr yggvnlvkn
301 lddgidderl rkefspfgti tsakvmegg rskgfgfvf sspeatkav temngrivat
361 kplyvalaqr keeqahlt nymqmasvr avnpviny qpappsgyfm aaipqtqnra
421 ayyppsquia lrpsprwtaq garphpfqnm pgairpaapr pfstmrpas sqvprvmstq
481 rvantstqtm gprpaaaaaa atpavrtvpq ykyaagvrnp qghlnaqqv tmqqpavhvq
541 gqepltasml asappqeqkq mlgerlfpli qamhptlagk itgmleiden sellhmlesp
601 eslrskvdea vavlqahqak eaaqkavnsa tgvptv

```

SEQ ID NO:48

Aspergillus niger

polyadenylate-binding protein [Aspergillus niger]

Accession number (NCBI): GAQ33421.1

```

1 msadvsttpa aenvngaaea spaaaapsat tpevtavens tpapaanqph saslyvgeld
61 psvteamlye lfssigqvas irvcrdavtr rslgyayvny nntadgeral edlnytlikg
121 kpcrimwsqr dpalrktggg nvfiknlds idnkaldhtf aafgnilsck vaqdefgnsk
181 gygfvyheta eaannaikhv ngmllndkkv fvgghiskkd rskfeemka nftnvyiknl
241 dseidddefr kmfekfgeit satlsrdqeg ksrgfgfvnf sthesaqaav eemndkeirs
301 qklyvgragk khereeelrk qyeaarleka skyqgvnlv knltddidde klrelfgyg
361 titsakvmrd tnveresspe sagkekeank endkeatpea ekaeekaek paesseekdk
421 eakksdkkpf gkskgfgfvf fsspdaska vtemnqrmvn gkplyvala rkdvrssqle
481 asiqarntir qqaaaaaagm ppqymqpavf yggpgqgfip ggqrgmafp qpvmvmgipg
541 grpgqypgpf pgqgggrgm pqqqippnfa qgipmgmgp gaipngmgyp qmaqvfgrg
601 agrgqvpqm pmqggmrqgp ygqggrrgvp vqggqirpqq ggrgqnaaap agpdegaagg
661 vnaqtlgaap paqqkqmlge alypkiaqq pelagkitgm llemndtell sileddeaalr
721 akvdealsvy deymknkgt seapaaeaak peakeeaste enka

```

SEQ ID NO:49

Cricetulus griseus

polyadenylate-binding protein 1 isoform X2 [Cricetulus griseus]

Accession number (NCBI): XP_027287290.1

```

1 mnpsapsypm aslyvgdlhp dvteamlyek fspagpilsr rvcrdmtrr slgyayvnfq
61 qpadaerald tmnfdvikgk pvrimsqrd psrlksgvgn ifiknldksi dnkalydtfs
121 afgnilsckv vcdengskgy gfvhfetgea aeraiekmng mllndrkvvf grfksrkere
181 aelgarakef tnvyiknfge dmdderlkd1 fgkfgpalsv kvmtdesgks kgfgfvsvfer
241 hedaqkavde mngkelngkq iyvgragkkv erqtelkrkf eqmkqdritr yggvnlvkn
301 lddgidderl rkefspfgti tsakvmegg rskgfgfvf sspeatkav temngrivat
361 kplyvalaqr keeqahlt nymqmasvr avnpviny qpappsgyfm aaipqtqnra
421 ayyppsquia lrpsprwtaq garphpfqnm pgairpaapr pfstmrpas sqvprvmstq
481 rvantstqtm gprpaaaaaa atpavrtvpq ykyaagvrnp qghlnaqqv tmqqpavhvq
541 gqepltasml asappqeqkq mlgerlfpli qamhptlagk itgmleiden sellhmlesp
601 eslrskvdea vavlqahqak eaaqkavnsa agvpsv

```

Fig. 1 (continued)

SEQ ID NO:50

Kluyveromyces lactis

uncharacterized protein KLLA0_C17600g [Kluyveromyces lactis]

Accession number (NCBI): XP_452986.1

```

1 msditektae qlenlqindd qppaqsasap stsaseseas svskvennna slyvgeldpn
61 iteallydvf splgpissir vcrdavtkas lgyayvnytd yeagkkaiqe lnyaeingrp
121 crimwserdp airkkgsgni fiknlhpaid nkalhetfst fgevlscvka ldengnsrgf
181 gfvhfkeesd akdaieavng mlmnglevyv amhvpkkdri skleeakanf tniyvknidv
241 ettdeefeql fsqygeivsa alekdaegkp kgfgfvnfvd hnaaakavee lngkefksqa
301 lyvgraqkky eraeelkkqy eqyrleklak fggvnlfikn lddsiddek1 keefapygti
361 tsarvmrdqe gnskfgfvc fsspeeatka mteknqqiva gkplyvaiaq rkdvrssqla
421 qqiqarnqir fqqqqqqqaa aaaagmpggy mpqmfygvma prgfpgpnpg mngpmgagip
481 kngmvpqpqq fagrpngpmy qgmppqngfp rhqqghyiqq qkqrqalgeq lykkvsakid
541 denaagkitg mildlppqqv iqllndndeq eqqfgealaa yenfkkeqea qa

```

SEQ ID NO:51

Ogataea polymorpha

uncharacterized protein OGAPODRAFT_16196 [Ogataea polymorpha]

Accession number (NCBI): XP_018211502.1

```

1 mssaeinqit eqaeklsvad paaaetpatt aeesstaest qnsetlasly vgeldpsvte
61 sdifevfspi gqvstirvcr davskqslgy ayvnfqshad gekaleelny tpikgkacri
121 mwsqrpslr rngsgnifik nlhpaidnkt lhdtsafgk ilsckiatde ngnskfgfv
181 hyedsesaka aienvngmll ndhevyvgph lakkdirqskm relianftnv yvkninlnwd
241 edklretfsp fgaissifls kdesgksrgf gfvnfekhed avkaveelnn keidgqklyv
301 graqkkserm eslkqhyeaa rgeqlnkyqg ynlfvknldd siddakleee fkyggtitsa
361 kvmllddagks kgfgfvcyss peeatkaite mhqrmvagkp lyvalaqrke vrrsqslsqqi
421 qarngmrqqq aaaqggmqgf vapmfyggnp gflppgarga pfaapggqmm mqqgiprpgq
481 gvpvspgqfr vgpngqpvpv ymqpvfneyy qngrmpqgry yqggrpgqqp pqagprgkse
541 dgavpslaqv lpqfppeqqk rmlgeelypl ivatgkaqdp eaagkitgmm lemdnqqila
601 lledrdlfn qfeealsaye eyknkeasa

```

SEQ ID NO:52

*Saccharomyces cerevisiae*Chain D, Polyadenylate-binding protein, cytoplasmic and nuclear
[Saccharomyces cerevisiae S288C]

Accession number (NCBI): 6R5K_D

```

1 gpdsmditd ktateqlenl iqddqkqaat gsesqsvens saslyvgdle psvseahlyd
61 ifspigsvss irvcrdaik tslgyayvnf ndheagrakai eqlnytpikg rlcrimwsqr
121 dpslrkkgsq nifiknlhpd idnkalydtf svfgdilssk iatdengksk gfgfvhfeee
181 gaakeaidal ngmllngqei yvaphlsrke rdsqleetka hytnlyvkni nsettdeqfq
241 elfakfgpiv saslekdadg klkgfgfvny ekhedavkav ealndslneg eklyvgraqk
301 knermhvlkk qyeayrlekm akyqgvnlfv knlddsvdde kleefapyg titsakvmrt
361 engkskgfgf vcfstpeeat kaiteknnqi vagkplyvai aqrkdvrssq laqqiqarnq
421 mryqqataaa aaaaagmpgq fmpmfygvn pprgvpfngp npqgmnpmgg mpkngmppqf
481 rngpvvgvpp qggfprnand nnqfyqqkqr qalgeqlykk vsaktsneea agkitgmild
541 lppqevfpll esdelfeqhy keasaayesf kkeqeqqteq a

```

Fig. 1 (continued)

SEQ ID NO:53

Schizosaccharomyces pombe

poly(A)-binding protein [Schizosaccharomyces pombe]

Accession number (NCBI): AAA35320.1

```

1  mpstdlkkqa daavesdvnt nneavesstk eessntpste tqpekkaeep eaaaepsest
61  stptnassva tpsgtaptsa slyvgeldps vteamlfelf nsigpvasir vcrdavtrrs
121 lgyayvnfhm medgekalde lnytlikgrp crimwsqrdp slrkmgtnv fiknldpaid
181 nkalhdtfss fgkilsckva vdelgnakgy gfvhfdsvs anaaiehvng mllndkkvyy
241 ghvsvrrerq skvealkanf tnvyiknldt eiteqefsd1 fgqfgeitsl slvkdqndkp
301 rgfgfvnyan hecaqkavde lndkeykgkk lyvgraqkhh ereeelrkry eqmklekmnk
361 yqgvnlfikn lqdevdderl kaefsafgti tsakimtdq gkskgfgfvc yttpeeanka
421 vtemnqrmla gkplyvala qkevrssle aqiqarnqfr lqqqvaaaag ipavqygatg
481 pliyqpggyp ipaavngrgm pmvpgnhgpm pmypgmptqf paggpapgy gmnnargpvp
541 qgrpmmpgs vpsagpaeae avpavpgmpe rftaadlaav peesrkqvlq ellypkvfvr
601 eeklsqkitg mlemnpnsel lelledsal nervneagv lqefvdqepg fte

```

SEQ ID NO:54

Trichoderma reesei

polyadenylate-binding protein [Trichoderma reesei QM6a]

Accession number (NCBI): XP_006966916.1

```

1  maanttsaav dqlasdlnt slngsgdvka paidtsvsag aedasaptpt aaphpqnsas
61  lyvgeldpsv teamlfelfs qigavasirv crdavtrrs1 gyayvynst pdgekalael
121 nytlikgrpc rimwsqrdpa lrktgggnvf iknldvaidn kalhdtfaaf gnlsckvaq
181 dengnskgyg fvhyetdeaa aqaikhvngm llnekkvyyg yhipkkdrqs kfeemkanft
241 nvyvkniapd vtdeedfrqlf ekfgdvtsss lardqegktr gfgfvnftth eaafkaveel
301 ngkdfgrgdl yvgraqkkhe reeelrksye aarlekanky qgvnlyikn1 dddvddeklr
361 qmfaefgpit sakvmrdvpq egeeeakdqe kdkenqkege kegesaegae kktekksdkk
421 lgkskgfgfv cfsnpddatk avaemnqrmi nnkplyvala qrkdvrrksq1 easiqarnql
481 rmqaaaaaag mpqqymqppv fyapggqpgf ipqgggrgipf pqpgmpmpnv qggrpgqfpg
541 ypqqgggrnmp qglppnmysi pgqfppqfqq pgtpqfmaam qqaqqaala ggrggapqgg
601 rgamagmppn vqggmpgypp nnrqmgrnnn ggngnnagr1 gnfpngaprd sntasmlqng
661 laaaqpaqqk qilgelifpk iqainaselag kitgmllemd nselvnlied eaalkakvde
721 alavydeyvk sqatgsekke eetka

```

SEQ ID NO:55

Yarrowia lipolytica

Polyadenylate-binding protein, cytoplasmic and nuclear [Yarrowia lipolytica]

Accession number (NCBI): AOW02148.1

```

1  msadaslvaa ekalekmtle nkaeaspatk eettteaapa egeaktesse ekgskedqgd
61  naslyvgeld psvteamlfe ifnpigpvts vrvcrdaitr rslgyayvnf hnqadgiral
121 eelnyspike rpcrimwsqr dpalrktgag niyiknldpa idnkalhdtf safggilsck
181 iatdefgnsr gfgfvhyesa esaesaighv ngmllndkkv fvgphvpksd rmqsfqeqkn
241 sftnvfikn1 gteiteaefe elvnkfgets svhlstndeg kptgfgfvd1 kehadvavkai
301 dglsetefkg nk1fagrakk kyeradelrk qyeasrlekl nkyqgvnlyi knlddtidd
361 klraefaphg titsakvmvd eagkskgfgf vcysseeat kavtemnhrl vagkplyvvl
421 aqrkdvrssq lqqqiqakng mrlqqqaaag glpgqymgnp gvfyppgpgf mppgrggmpf
481 ganpqmmmrp pmppqnqfpp rgvpggpnmy gappqgyqqg gfppqgpmrg gqpprrsgqpg
541 pqgqfrgapr rkdgesrvad sisnalenap eeqhkqlvge alypkvlaek aidgnaefag
601 kitgmllemp ikeileidd eeglqaqind aitayneyn1 sqkee

```

Fig. 1 (continued)

Recycling factor 1 - *RLI1* - PP7435_Ch2-1213

SEQ ID NO:56

Komagataella phaffii

PP7435_Ch2-1213

MSEKNTRIAIVSDDRCKPKKCRQECRKSCPVVKTGKLCIEVEAYS KIAFISETLCIGCGICVKKCPFGAINIINLP
 TNLEKEVTHRYSANSFKLHRLPTPRPGQVLGLVGTNGIGKSTALKILAGKQKPNLGRYDAPPDWEEILKYFRGSEL
 QNYFTKVLLEDNIKAIKPKQYVDNLPRAIKGPVNKVGELLKAKLERKEEGKHIIQVLELKHVLNREVAKLSSGGELQR
 FAIGMTCVQDANVYMFDEPSSYLDVKQRLNAAQIIRSLHATCYVIVVEHDLSDVLDYLSDFVCILYGVPSVYGVVT
 LPASVREGINIFLDGHIPTENLRFRITESLKFKLTDASDSVLIEQSRAFEYPKMVKTQGDFFKLTVNAGDFSDEIIV
 MMGENG TGKTTLLKLLAGAIEGDEGTVKPKLVSMKPKQKIAPKFTGTVRQLFLKKIRSSFLHPQFNTDVMKPLKVD
 DVIDQEVQHLSSGGELQRVAIVLALGKEADIIYLIDEPSAYLDSEQRIICSKVIRRFILHSSKKTAFIVEHDFIMATY
 LADKVMVFEGQPSKNAVANRPETLLTG CNRFLKLNLTFRDPNSFRPRINKLDSQMDREQKLSGNYFFLENTL

SEQ ID NO:57

Komagataella pastoris

BA75_05167T0 [Komagataella pastoris]

Accession number (NCBI): ANZ77330.1

1 msekntriai vsddrckpkk crqecrkscp vvktgklcie veayskiafi setlcigcgi
 61 cvkkcpfgai niinlptnle kevthrysan sfklhrlptp rpgqvlglvg tngigkstal
 121 kilagkqkpn lgrydappdw eelkyfrgs elqnyftkvl ednikaiikp qyvdlprai
 181 kgpvnkvgel lkaklerkee gkhiiqvllel khvlnrevak lsggelqrfa igmtcvqdan
 241 vymfdepssy ldvkqrlnaa qiirslhat cyvivvehdl svldylsdfv cilygvpsvy
 301 gvytlpasvr eginifldgh iptenlrftr eslkfklt da sdsvlieqsr afeypkmvkt
 361 qgdfkltvna gdfsdseiiv mmgengtgkt tlckllagai egdegtkvpk lnvsmkppki
 421 apkftgtvrq lflkkirssf lhpqfntdvm kplkvddvid gevqhlsgge lqrvaivlal
 481 gkeadiylid epsayldseq riicskvirr filhskktaf ivehdfimat yladkvmvfe
 541 gqpsknavan rpetlltgc n rflknlntf rrdpnsfrpr inklds qmdr eqklsgnyff
 601 lentel

SEQ ID NO:58

Homo sapiens

ATP-binding cassette sub-family E member 1 [Homo sapiens]

Accession number (NCBI): NP_001035809.1

1 madkltriai vnhdkckpkk crqeckkscp vvmgklcie vtpqskiawi setlcigcgi
 61 cikkcpgfal sivnlpsnle ketthrycan afklhrlpip rpgevlglvg tngigkstal
 121 kilagkqkpn lgkyddppdw qeiltlyfrgs elqnyftkil eddlkaiikp qyvdaqipkaa
 181 kgtvgsildr kdetktgaiv cqqlldthlk ernvedlsgg elqrfacavv ciqkadifmf
 241 depssyldvk qrlkaaitir slinpdryii vvehdlsvld ylsdficclly gvpsaygvvt
 301 mpfsvregin ifldgyvpte nlrfrdaslv fkvaetanee evkkmcmyky pgmkkkmgef
 361 elaiavageft dseimvmlge ngtgkttfir mlagrlkpde ggevpvlnvs ykpqkispks
 421 tgsvrqllhe kirdaythpq fvt dvmkplq ieniidqevq tlgsggelqr alalclgkpa
 481 dvylidpsa yldseqrlma arvvrkfilh akktafvveh dfimatylad rvivfdgvps
 541 kntvanspqt llagmnkfls qleitfrdp nnyrprinkl nsikdveqkk sgnyffldd

SEQ ID NO:59

Aspergillus niger

translation initiation factor RLI1 [Aspergillus niger CBS 513.88]

Accession number (NCBI): XP_001392338.2

1 mkkcrqeckk scpvvrtgkl cievtpeski afiserlcig cgicpkkcpf gaihiinlpt
 61 nletqvthry sansfklhrl pmprpgqvlv lgvtngigks talkilsgkl kpnlgrfdnp
 121 pdweeilkyf rgseelqnyft kvleddlkav vkpqqvqip ravkgpvqnv galikaraq
 181 dnldhildtl elrqvqdrdi ghlgsggelqr faiglvcvqk advymfdeps syl dvkqrla
 241 aarsirellr pddyvivveh dlsvldylsd fvcvlygrpa lygvvtlpas vreginifld
 301 ghiptenlrf reesltfirla eagddfmvdr arafaypsme ktlnfnhlki dagsftdsei
 361 ivmmgengtg kttfckmlag aekpdggksv prlnismkpq kitpkfggtv rqlffkrika
 421 aflspqfqt dvykplkidd idgevgnlsg gelqrvaivl algipadiyl idepsaylds
 481 eqrivasrvi krhimhtkkt afivehdfim atyladrviv fdgqpsvdah anapeslvtg
 541 cntflknldv tfrdpnsyr prinkyqsgm dgeqklagny ffleek

Fig. 1 (continued)

SEQ ID NO:60

*Cricetulus griseus*ATP-binding cassette sub-family E member 1 [*Cricetulus griseus*]

Accession number (NCBI): XP_003500814.1

```
1  madkltriaai vnhdkckpkk crqeckkscp vvrngklcie vtpgskiawi setlcigcgi
61  cikkkcpfgal sivnlpsnle ketthrycan afklhrlpip rpgevlglvq tngigkstal
121 kilagkqkpn lgkyddppdw qeiltyfrgs elqnyftkil eddikaiikp qyvdaqipkaa
181 kgtvgsildr kdetktgaiv cqqldlthlk ernvedlsgg elqrfacavv ciqkadifmf
241 depssyldvk qrlkaaitir slinpdryii vvehdlsvld ylsdficclly gvpsaygvvt
301 mpfsvregin ifldgyvpte nlrfrdaslv fkvaetanee evkkmcmyky pemkkkmgef
361 elaiavageft dseimvmlge ngtgkttfir mlagrlkpde ggavpilnvs ykpqkispks
421 tgsvrqllhe kirdaythpq fvtdvmkplq ieniidqevq tlgggelqrv alalclgkpa
481 dvylidepsa yldsegrlma arvvrkrilh akktafvveh dfimatylad rvivfdgvps
541 kntvanspqt llagmnkfls qleitfrrdp nnyrprinkl nsikdveqkk sgnyffldd
```

SEQ ID NO:61

*Kluyveromyces lactis*Rli1 [*Kluyveromyces lactis*]

Accession number (NCBI): QEU59573.1

```
1  msskntriaai vsedkckpkk crqeckrscp vvktgklcie vmptskiafi senlcigcgi
61  cvkkcpfdai qiinlptnld sqvthrysan sfklhrlptp rpggvlglvq tngigkstal
121 kilagkqkpn lgrfddppew qeiiyfrgs elqnyftkml eddikaiikp qyvsniprai
181 kgpvqkvgei lklrlekskd daknyidilq lknvlnrevs alsggelqrf aigmscvgea
241 dvymfdepss yldvkqrlna aliirsllkp ttyvicvehd lsvldylsdf vciiygvpsv
301 ygvvtlpsv reginifldg hipsenlrfr tealqfmad agdeleadas raftypdmkr
361 tqgdfsltve agefsdsqil vmmgengtqk ttlikllaga ipaddgqddp klnvsmkpqk
421 iapkftgtvr qlffkrirag flnpqfqtvd vkplkiddii dgevqhlsgg elqrvaivls
481 lgnnadiyli deptsayldse qriicskvir rfilhnkhta fivehdfima tyladkvivf
541 dgtpskqafa rapeslltgc nkflknlnvt frrdptsfrp rinkldsqud keqkssgnyf
601 fldhngv
```

SEQ ID NO:62

*Saccharomyces cerevisiae*Fe-S cluster-binding ribosome biosynthesis protein [*Saccharomyces cerevisiae* S288C]

Accession number (NCBI): NP_010376.3

```
1  msdknsriai vsadkckpkk crqeckrscp vvktgklcie vtptskiafi seilcigcgi
61  cvkkcpfdai qiinlptnle ahvthrysan sfklhrlptp rpggvlglvq tngigkstal
121 kilagkqkpn lgrfddppew qeiiyfrgs elqnyftkml eddikaiikp qyvsniprai
181 kgpvqkvgei lklrmekspe dvkryikilq lenvlkrdie klsggelqrf aigmscvgea
241 dvymfdepss yldvkqrlna aqiirsllap tkyvicvehd lsvldylsdf vciiygvpsv
301 ygvvtlpasv reginifldg hipaenlrfr tealqfriad atedlqnds srafsypslk
361 ktqgdfvlv eegefdsdei lvmmgengtq kttlikllag alkpdegqdi pklnvsmkpq
421 kiapkfpgtv rqlffkkirg qflnpqfqtvd vvkplridi idgevqhlsg gelqrvaivl
481 algipadiyl ideptsaylds eqriicskvi rrfilhnkkt afivehdfim atyladkviv
541 fegipsknah arapeslltg cnrflknlnv tfrrdpnsfr prinkldsqu dkeqkssgny
601 fflidntgi
```

Fig. 1 (continued)

SEQ ID NO:63

Schizosaccharomyces pombe

putative Fe-sulfur ATPase Rli1 [Schizosaccharomyces pombe]

Accession number (NCBI): NP_596732.1

```

1 msesltriali vsedkcrpkk crqecrrscp vvrtgklcie vnptdriafi setlcigcgi
61 cvkckpfgai niinlptnle sevthrysan sfklhrlptp rpgqvlglvg tngigkstal
121 kilsgkmkpn lgrydnppdw aevvkyfrgs elqnfftkvv ednikalikp qyvधिप्रा
181 ktgdktvsgl ikaranannnf eevmdhtdlq nllnrevghl sggelqrfaai aavatqkad
241 ymfdepssyl dikqrlkagr virsllattn yvivvehdls vldylsdfvc vlygvpsmyg
301 vvtlpysvre ginifldghi ptenlrfrse altfrladas deitadrtae ynypdhvieq
361 gdfkltiksg gfsdaeiiiv lgengtgtkt fckwmaknsd lkismkpqti apkfggtvrm
421 lflkkiraaf lngkfqsevc kplsidiid qevlnlsgge lqrvaiclal gmpadvylid
481 epsayldseq riiaskvirr fivnsrktaf ivehdfimat yladrvilfe gqpsrdarcn
541 ppqsliltgm tflknldvtf rrdpntlrpr inkfdsqmdq eqknagnyff len

```

SEQ ID NO:64

Trichoderma reesei

P-loop containing nucleoside triphosphate hydrolase protein [Trichoderma reesei RUT C-30]

Accession number (NCBI): ETS04026.1

```

1 mektlgnfrl ridageftds eiivmmgeng tgkttfcorll agalkpdskr svpdmkismk
61 pqtitpkfeg tvrqflffkki kaaflspqfq tdivkplkld dfidqevknl sggelqrvai
121 vlalgpadi ylidepsayl dseqriiasr vikrfimhsk ktafvvehdf imatyladrv
181 ivfdgkpgid ahanrpesll tgcntflknl dvtfrrdptn yrprinklgs qldqeqklsg
241 nfffleekpe eta

```

SEQ ID NO:65

Yarrowia lipolytica

YALIOA17941p [Yarrowia lipolytica CLIB122]

Accession number (NCBI): XP_500182.1

```

1 mppknknevk dgdkntriali vnvdkckpkk crqeckkscp vvrtgklcie vtpdskiafi
61 setlcigcgi cpkkcpfgal qiinlptnld sevthrysan sfklhrlptp rpgqvlglvg
121 tngigkstal kilagkqkpn lgryedpptw qeilthfrgs elqnyftrvl ednlkaiikp
181 qyvधिप्रा kgplqkvge lkkaklerkdd fenivntlel kgvkwrevsl lsggelqrfa
241 igmscvqlaq vymfdepssy ldvkqrlhaa amirslidst gyiivvehdl svldylsdfv
301 cilygvpsvy gvvtlpssvr eginifldgh iptenlrfrt eslqfrlada teiipldqsr
361 sfkyptmsrt qgdfklvnes geftdseiiv mmgengtgtkt tlcrllagav apdegqkapk
421 lnvsmkpqki apkftgtvrq lflkkirssm hhpqfqtdiv kplriddiid qevqnlsgge
481 lqrvaiavlal gqpadiylid epsayldseq riicakvirr filhskktaf ivehdfimat
541 yladrvivfd gqpsveafar rpeslltgc n rflkslnvtf rrdptsfrpr inkkdsqmdk
601 eqkahgnyff ldtaea

```

Fig. 1 (continued)

Translation initiation factor 4E (eIF4E) - CDC33 - PP7435_Ch3-0197
Nucleotide sequences encoding SEQ ID NO:1

SEQ ID NO:66

Komagataella phaffii

Original nucleotide sequence

ATGTCAGAGACTGAAAACGCTCCTCACGAAAAGTCACTGCTTGACGACCAAGTTGATTCTACTGTTAAGCATCCGC
TCAACTCTGAGTGGACACTATGGTATACGAAGCCAGCCGTTTCATGAACAAGAAAGTTGGTCAGATTTGCTACGTCC
TATCATTTTCATTCAACTCGGTAGAGGAGTTCTGGGGAATCTTCAACTCCATTCCAAAGGCCTCCGATTTGCCAGTC
AAGTCGGACTATCATCTTTTATAGAGATGGAATCAAACCTGAATGGGAAGACGAGCGCAATGCTAACGGAGGAAAAAC
TGACATACCAAGTTTGCTACAGAAAGGTTGATATCAATGAATTGTGGTCTCGCGGTCTGTTGAGTGTGATTGGTGA
AACGATTACAGGATGATGACGGAGAACTGAGGTGAACGGTATTGTTCTGAGCATTCGTCGTGCCGCTTTGAAGATT
GCCCTCTGGACCAAGTCCAAGGATGAGGCAGTTTGAAGACCAATTGGCGAGAGATTCAAGAAAGTTTGAAGTTGA
CAGCCCGTGATCATATTGACTTTCAACCTCATAAGAGTGGTGTCTGGTAGCAAGAACGTACCTTCTTTCAGCATTTA
G

SEQ ID NO:67

Golden gate optimized nucleotide sequence, i.e. SEQ ID NO:66, comprising
point mutations A276G, T354C, G492A; positions underlined

ATGTCAGAGACTGAAAACGCTCCTCACGAAAAGTCACTGCTTGACGACCAAGTTGATTCTACTGTTAAGCATCCGC
TCAACTCTGAGTGGACACTATGGTATACGAAGCCAGCCGTTTCATGAACAAGAAAGTTGGTCAGATTTGCTACGTCC
TATCATTTTCATTCAACTCGGTAGAGGAGTTCTGGGGAATCTTCAACTCCATTCCAAAGGCCTCCGATTTGCCAGTC
AAGTCGGACTATCATCTTTTATAGAGATGGAATCAAACCTGAATGGGAGGACGAGCGCAATGCTAACGGAGGAAAAAC
TGACATACCAAGTTTGCTACAGAAAGGTTGATATCAATGAATTGTGGTCTCGCGGTCTGTTGAGTGTGATTGGTGA
AACGATTACAGGATGATGACGGAGAACTGAGGTGAACGGTATTGTTCTGAGCATTCGTCGTGCCGCTTTGAAGATT
GCCCTCTGGACCAAGTCCAAGGATGAGGCAGTTTGAAGACCAATTGGCGAGAGATTCAAGAAAGTTTGAAGTTGA
CAGCCCGTGATCATATTGACTTTCAACCTCATAAGAGTGGTGTCTGGTAGCAAGAACGTACCTTCTTTCAGCATTTA
G

Translation initiation factor 4A (eIF4A) - TIF2
TIF2a - PAS_chr3_0595 (sequence from GS115 strain)
Nucleotide sequences encoding SEQ ID NO:12

SEQ ID NO:68

Komagataella phaffii

Original nucleotide sequence

ATGTCTGAAGGTATTATTGAAATCGACACTAACTTAATCGAGACCAACTACGACCAGGTTGTCAGCACCTTCGATG
AGCTGAACCTGAAGGAGCAAATCCTGAGAGCCGTTTTTGTCTACGGTTACGAGTTCCCATCTGCCATTCAATCTCG
TGCCATTTTGCCAAATTGTTGCTGGTCGTGATGTCTTGCCCAAGCCAGTCCGGTACCGGTAAGACTGCTACCTTT
GCCATCTCCGCTTTGCAGAGATTGGATGAGAAGCTGAAGGCTGCCCAAGTTTTGATCTTGCCCCAACAGAGAGT
TGGCCCTGCAGATCAAGAAGGTTGTTTTGGCCCTCGGTATGCACTTGAACATCTCTGTCCATGCCGTATCGGTGG
TGTGACCCAGCTTTGGATGTCAAGGCCCTTCAACTCCGGTGTTCAAATCGTCGTTGGTACCCAGGTAGAGTCCAA
GACATGATCGAACGTGGAGCTTTCAGAACTGAGAAGCTCAAGATGGTCAATTTGGATGAGGCCGATGAGATGTTGT
CCACTGGATTCAAGGAGCAAATCTACACCTTGTTCACATTGTTACCATCCACTACACAGGTTGTTCTTCTATCCGC
CACCATGCCTCAAGAGGTGTTGGAGGTTACCACCAAGTTCATGAACAACCTATTCGTATCTTGGTCAAGAAGGAC
CAATTGACCTTGGAAGGTATTAAGCAATTCACGTCGACGTCGGTGAAGAACAATACAAGTTCGACTGTCTGTGTG
ATCTGTACGACTCCATCTCCGTCACCCAGGCCGTTATCTTCTGTAAACACCAGACGAAAGGTCGAGGAGTTGACTGA
GCGTTTGACCGCCCAAGATTACCGTCTCTGCTATCCACTCTGAGCTGTCACAAACCGAGAGAGACACCATCATG
CAAGAGTTCAGAACCGGTTCTTCCAGAATCTTAATCTCCACCGACTTGCTGGCCAGAGGTATCGATGTCCAACAAG
TCTCATTGGTTATCAACTACGACTTACCATCTAACAAGGAAAACCTACATCCACAGAAATCGGTAGAGGTGGTCTGTT
CGGTAGAAAGGGTGTGCTATCAACTTCGTACCCAGCAAGGACGTTGAGATGATGAAGGAGATCGAGAAGTTCTAC
TCTACCCAAATCGACGAGTTGCCATCTTCCATCACTGACTTGTGTTGAGACTGAGGAAGTTAATGAGTCTTAG

Fig. 1 (continued)

SEQ ID NO:69

Golden gate nucleotide sequence, i.e. SEQ ID NO:68, comprising point mutation C45A, position underlined

ATGCTCTGAAGGTATTATTGAAATCGACACTAACTTAATCGAGACAAACTACGACCAGGTTGTCAGCACCTTCGATG
AGCTGAACCTGAAGGAGCAAATCCTGAGAGCCGTTTTTGGCTTACGGTTACGAGTTCCCATCTGCCATTCAATCTCG
TGCCATTTTGCCAAATTGTTGCTGGTTCGTGATGCTTGGCCCAAGCCAGTCCGGTACCGGTAAGACTGCTACCTTT
GCCATCTCCGCTTTCAGAGATTGGATGAGAAGCTGAAGGCTGCCCAAGTTTTGATCTTGGCCCCAACAGAGAGT
TGGCCCTGCAGATCAAGAAGGTTGTTTTGGCCCTCGGTATGCACTTGAACATCTCTGTCCATGCCGTATCGGTGG
TGTTGACCCAGCTTTGGATGTCAAGGCCCTCAACTCCGGTGTTCAAATCGTTCGTTGGTACCCAGGTAGAGTCCAA
GACATGATCGAACGTGGAGCTTTCAGAACTGAGAAGCTCAAGATGGTCATTTTGGATGAGGCCGATGAGATGTTGT
CCACTGGATTCAAGGAGCAAATCTACACCTTGTTCACATTGTTACCATCCACTACACAGGTTGTTCTTCTATCCGC
CACCATGCCTCAAGAGGTGTTGGAGGTTACCACCAAGTTCATGAACAACCCATTCGTATCTTGGTCAAGAAGGAC
CAATTGACCTTGAAGGTATTAAGCAATTCACGTCGACGTCGGTGAAGAACAATACAAGTTCGACTGTCTGTGTG
ATCTGTACGACTCCATCTCCGTCACCCAGGCCGTTATCTTCTGTAACACCAGACGAAAGGTCGAGGAGTTGACTGA
GCGTTTGACCGCCAGAAAGTTCACCGTCTCTGCTATCCACTCTGAGCTGTCAAAAACCGAGAGAGACACCATCATG
CAAGAGTTCAGAACCGGTTCTTCCAGAATCTTAATCTCCACCGACTTGTCTGGCCAGAGGTATCGATGTCCAACAAG
TCTCATTGGTTATCAACTACGACTTACCATCTAACAAGGAAAACCTACATCCACAGAATCGGTAGAGGTGGTTCGTTT
CGGTAGAAAGGGTGTGCTATCAACTTCGTCACCAGCAAGGACGTTGAGATGATGAAGGAGATCGAGAAGTCTAC
TCTACCCAAATCGACGAGTTGCCATCTTCCATCACTGACTTGTGTTGAGACTGAGGAAGTTAATGAGTCTTAG

Translation initiation factor 4A (eIF4A) - TIF2**TIF2b - PP7435_Chr3-0610****Nucleotide sequences encoding SEQ ID NO:23**

SEQ ID NO:70

Komagataella phaffii

Original nucleotide sequence

ATGCATCCATACACCGTAGTATTTGCGCGCCTCCTCCTGGGTGTTTTCTCAACTGCCGCCCCCAACTCGCCCCCCC
CCTTTTGCCTTCTTGCCACCTTGGAGATCAGAGATTGTGGAACCTTGCTCAAAATCTGGCATGCACACTACCTCC
TCAAAAAAATTTTCTTGCCTTGTGCAAGGCGCTCAAAAATCAAAAGTTTTTCTACAACCTTTTGTTCACAAAA
TCACCTTTTTTACTTACCCATTATGTCTGAAGGTATTATTGAAATCGACACTAACTTAATCGAGACCAACTACGACC
AGGTTGTCAGCACCTTCGATGAGCTGAACCTGAAGGAGCAAATCCTGAGAGCCGTTTTTGGCTTACGGTTACGAGTT
CCCATCTGCCATTCAATCTCGTGCCATTTTGCCAATTGTTGCTGGTTCGTGATGTCTTGGCCCAAGCCAGTCCGGT
ACCGGTAAGACTGCTACCTTTGCCATCTCCGCTTTCAGAGATTGGATGAGAAGCTGAAGGCTGCCCAAGTTTTGA
TCTTGGCCCCAACAGAGAGTTGGCCCTGCAGATCAAGAAGGTTGTTTTGGCCCTCGGTATGCACTTGAACATCTC
TGTCATGCCTGTATCGGTGGTGTGACCCAGCTTTGGATGTCAAGGCCTTCAACTCCGGTGTTCAAATCGTCGTT
GGTACCCAGGTAGAGTCCAAGACATGATCGAACGTGGAGCTTTCAGAACTGAGAAGCTCAAGATGGTCATTTTGG
ATGAGGCCGATGAGATGTTGTCCACTGGATTCAAGGAGCAAATCTACACCTTGTTCACATTGTTACCATCCACTAC
ACAGGTTGTTCTTCTATCCGCCACCATGCCTCAAGAGGTGTTGGAGGTTACCACCAAGTTCATGAACAACCCATAT
CGTATCTTGGTCAAGAAGGACCAATTGACCTTGGAAAGGTATTAAGCAATCTACGTCGACGTCGGTGAAGAACAAT
ACAAGTTCGACTGTCTGTGTGATCTGTACGACTCCATCTCCGTCACCCAGGCCGTTATCTTCTGTAACACCAGACG
AAAGGTCGAGGAGTTGACTGAGCGTTTGACCGCCAGAAAGTTCACCGTCTCTGCTATCCACTCTGAGCTGTCAAA
ACCGAGAGAGACACCATCATGCAAGAGTTCAGAACCGGTTCTTCCAGAATCTTAATCTCCACCGACTTGTGGCCA
GAGGTATCGATGTCCAACAAGTCTCATTGGTTATCAACTACGACTTACCATCTAACAAGGAAAACCTACATCCACAG
AATCGGTAGAGGTGGTTCGTTTCGGTAGAAAGGGTGTGCTATCAACTTCGTCACCAGCAAGGACGTTGAGATGATG
AAGGAGATCGAGAAGTCTACTCTACCCAAATCGACGAGTTGCCATCTTCCATCACTGACTTGTGTTGAGACTGAGG
AAGTTAATGAGTCTTAG

Fig. 1 (continued)

SEQ ID NO:71

Golden gate nucleotide sequence, i.e. SEQ ID NO:70, comprising point mutation C294A, position underlined

ATGCATCCATACACCGTAGTATTTGCGCGCCTCCTCCTGGGTGTTTTCTCAACTGCCGCCCCCAACTCGCCCCCCC
CCTTTTGCACCTTCTTGCCACCTTGGAGATCAGAGATTGTGGAACCTTGCTCAAAATCTGGCATGCACACTACCTCC
TCAAAAAAATTTTCCTTGCCCTGTTGCAAGGCGCTCAAAAAATCAAAAGTTTTTCTACAACCTTTTGTCCCAACAAA
TCACCTTTTTACTTACCCATTATGTCTGAAGGTATTATTGAAATCGACACTAACTTAATCGAGACAAACTACGACC
AGGTTGTCAGCACCTTCGATGAGCTGAACCTGAAGGAGCAAATCCTGAGAGCCGTTTTTGCTTACGGTTACGAGTT
CCCATCTGCCATTCAATCTCGTGCCATTTTGCCAATTGTTGCTGGTCGTGATGTCTTGGCCCAAGCCCAGTCCGGT
ACCGGTAAGACTGCTACCTTTGCCATCTCCGCTTTGCAGAGATTGGATGAGAAGCTGAAGGCTGCCCAAGTTTTGA
TCTTGGCCCCAACAGAGAGTTGGCCCTGCAGATCAAGAAGGTTGTTTTGGCCCTCGGTATGCACCTGAACATCTC
TGTCATGCCTGTATCGGTGGTGTGACCCAGCTTTGGATGTCAAGGCCTTCAACTCCGGTGTTCAAATCGTCGTT
GGTACCCCAGGTAGAGTCCAAGACATGATCGAACGTGGAGCTTTCAGAACTGAGAACGTCAAGATGGTCATTTTGG
ATGAGGCCGATGAGATGTTGTCCACTGGATTCAAGGAGCAAATCTACACCTTGTTTACATTGTTACCATCCACTAC
ACAGGTGTTCTCTATCCGCCACCATGCCCAAGAGGTGTTGGAGGTTACCACCAAGTTCATGAACAACCCATT
CGTATCTTGGTCAAGAAGGACCAATTGACCTTGAAGGTATTAAGCAATCTACGTCGACGTCCGGTGAAGAACAAT
ACAAGTTCGACTGCTGTGTGATCTGTACGACTCCATCTCCGTCACCCAGGCCGTTATCTTCTGTACACCAGACG
AAAGGTCGAGGAGTTGACTGAGCGTTTGACCGCCCAGAAGTTCACCGTCTCTGCTATCCACTCTGAGCTGTACAA
ACCGAGAGAGACACCATCATGCAAGAGTTCAGAACCGGTTCTTCCAGAATCTTAATCTCCACCGACTTGCTGGCCA
GAGGTATCGATGTCCAACAAGTCTCATTGGTTATCAACTACGACTTACCATCTAACAAGGAAAACTACATCCACAG
AATCGGTAGAGGTGGTTCGTTTCGGTAGAAAGGGTGTGCTATCAACTTCGTCACCAGCAAGGACGTTGAGATGATG
AAGGAGATCGAGAAGTCTACTCTACCCAAATCGACGAGTTGCCATCTTCCATCACTGACTTGTTTGAGACTGAGG
AAGTTAATGAGTCTTAG

Fig. 1 (continued)

Translation initiation factor 4G - eIF4G - *TIF4632* - PP7435_Chrl-0352
Nucleotide sequences encoding SEQ ID NO:34

SEQ ID NO:72

Komagataella phaffii

Original nucleotide sequence

ATGTCCAATAAGAACGTGGATACAGCTCCAGATAATGGTTCTGGTGGCTCCAATCCCCAGCTTCTGTGCCGCATT
ATCCACAGCAACACCGTGCTAATTATAATCCTCATAAGAATCATACTAATCATGGGGGCAATAATGGAGGCGCCCA
GGGTACATGCAGCATCAGCAACAGCAACAAATGTATTACAATCAGATGGGCTACCTATGCCGTATCCGAATGCT
GGTGGCAATTTCAACCAGAATGTATACTTGGTTCTTTATAATACTATGGTTCCAGGTACTCCATATGGTGCAGGTG
CCCCGGGCGTGCTCCAAGCTACCCAGTGTCTCCTTCTCCAGCATTCGCCTCTCCACAAAAGACTCATGTGCCAGT
AAGTCTGCGTACCCCTTCTGGAGAACCTATAAATTTGGTTCTGCTGGGCAACGAGTTCAGTTCATCCAATGCC
TCTCCTCATGTTAAGTCTACCTCTATTCCTCCTTCTGCTGTCACTAGTCTGCTGCTTCCATTGCTCCAAGTGTG
CATCCCGACTAAGGCCGCGGTGGATCTGAAGACGCTGGCAAGACTAGAGATGCATTCAAGCAGATGGTTCAGGC
TTTAGCTGCTAAAAAGAAGGCTGCAGCTGCTGCCAAGGCTGAAGGTAAAAACAGATGATAAAAAAGGAGTCTGTGGAG
TCTAAGCCTAAGTCTGCTGTACAGCCGAGACAAAGCCAAAAGTTGAGGTAAAGTTAGAGCCAAAACCAAAAGTTG
AGGTCAAATTAGAGCCAAAGCCAAAGGTTGAGATTAACTAGAGCCCAACAGTATCGGAGGCTAAACCAGAGAC
AAAAACAGTACCTGAGGTTGAAGAACCTTCAACAGTTGCAACGATCAAAATCAGAATCCACACCAATTGAGGAGCTC
AAACCCAAAGTTGAGGTAACCCAGAAGTTGCATCCAAGCCTGAACAGGTTGAGTCTGCACAGGAGCCGGCTAACG
ATGTAAAACCATCTGAAGTCTCAGCTAAAGAGGTTCAAGAATCCAAGGAGGAGTCTTTGAAACAACCTAAGGAAGC
TTTGAATGACATTAAATCCACCGAGGAGTCAATAAAAAAGTTTCTCTTCCAGAAGAGGAGTCTAACGATCACCAG
ACACTTTCCCAAATTCAGGAGTTAAAGATACAGAACAATCGCAAGAATTTGAAGAACCTCAAGGAACCCAGGAAA
CTCAGGAACTCAGGAAACGCAGGAAACCGCAAAGCCTTTGGAAGCTCAAGCACCTCAAGAAATTCAGGAAACCCA
AGAACCTGAGGAAGTTCAAGAACAACTCCTGAAACATCTGAGACTGTTTCTTCTGTACAGAAGAAGTGTCTCCA
GCTCCAGCACCTGGATTGCCCTTTCAAAGTTCTTGGACAGAATCAGAAAGGCTCCCGCCATTTCAAACCCATTCA
AGCATGAGTACTCAGGTGATTTCCATGGTCCAGACGAAAGTTTAAAGTGGTCTGAAAACCATCAAATATGATCCGGG
TTTCTTGATCCAATTCGCCACCTTCAACTTCCATATTGATGAAGATTTCAACGATGTCGGTTCTAAGGTTGTG
ATCCCTGTCAACACCAGAAGAGACAATTTTCCAAGAGCATGTCTACCAGCAATTTCAAAGGTTCTGGTAGAAATA
TGGGTTCTAATTTCAAGAAATGATTTTGGATCTAGAAACAACCTCCAAGCAAGGTTCCAAGAAAAAGATAAATCTCA
AAAGAGCCAATCTAGGAGGGGTCCAAAGTACTCCAATGAAGAAGAACCTGCTGAACCAAGCCGTGCCTGCTGAATTG
CCTAATAGAGCTGCAAGCAAGTGGGTACCTCGTTCTGTTAAGAAATAAGGAGGCTGAAGTCAAGACTGCTCCAGACG
GATCTATAATAGTATCAGAAGAAGACATCAAAGGAAACTAAATCCCTTTTGAATAAGTTGACGTTGGAATCTCT
TGATGATATCTCAAACGATATAATTGCTTTGACCAAGCAAGCTCAATGGGAAGATGACGTGAAGACTTTGAAACAA
GTTATTGAGTCTATATTTGCAAAGGCTTGTGACGAACCTTACTGGTCTTCTATGTACGCTAAATTATGCGCCAAAA
TGTGCAAGGACACCCACCTGAGATCAAGGAACTAATGAGAAGGAAATACTTTCACCGGTGGTGATTTGGTGAG
AAGAGTGTTGATTAATAGATGTCAAACCGAATATCAGAAAGGATGGATTGACCAGCTACCAACCAACGAGGATGGT
TCACCTCTAGAGCCAGAACTGATGTCTGACGAATATTACAACAAGCTGCTGCTAAACGTAGAGGCTTGGTCTGG
TTAAGTTCAATTGGTCAATTGTACCAAGTTGAACTTATTATCAGACAATGTCAATTTTCCACTGTGTCCTGGTCAAGT
CAAAGAAACAGACAATCCATCAGAAGATACTATGGAGAACCTGGTTCAATTATTGGAAGCTGTAGGTTATAGAATG
GAGATGTCCAGTGAACCTAAGATTAGAAGGTTCTTGATCTCGCCTTTGCCAACATCAAGGCCATTGTTGACGCTA
ACAAGATTCCTTCAAGAATAAGATTTATGCTCATGGATTTGATGGATCTAAGAAGAAATGGCTGGGCTACAACAGA
AGAGAACGCTGGTCCAAAGACCATTCAAGAGATTAAGTTAGAGGCTGAGATCAAGAGATCAGAGGAGGAGAGAGAA
AAAGAAAGAGACAGGCATTCGTAGAGGACGTTACAACGACTCCAAACCTAACTCTCGTCAAAAGTCACTTTGGG
GAAGTTCAGTGGTTTACCTAATGATATTAGAAATGCAGGCTATGTTAGAAACTCTGACAAATCTCTGGGGCCTGT
CAACAGTTTTTCCAGGGGCAAGTCATCCCGTCAAGCTTCAAAGCAATCAAACAGTGATTTTACCAGTGTTCCTATC
ACTTCCAGTGGTAGTAGAAGTCTCTAAACCACTCTGGCGACAGCTCAACGTCTCTGAACGAACCTCAATCAAGAG
AATATTTCAAAAAATCTGAATCTAGGCCCGCTCCTTCCAACATTTTTGCTGCATTAGAGGATCACGATCATGATCA
TGAACATGGCCACGAACAAGAAGAAAACAAGCATTCCAAGGATCTCCAAAACGAGAGGCGGAAGCACCTGAGGAC
GACAATGTGGGAGAAATAGCCCAAGCTGACGAAAATGTGCAAGGGCAAGAGCAAGAGGAAGAACGAGAGGAAGAGC
GAGAGCAAGAAGAGGAACAGGAAGTTTAA

Fig. 1 (continued)

SEQ ID NO:73

Golden gate nucleotide sequence, i.e. SEQ ID NO:72, comprising point mutations A564G, A1923G, G2037C, T2100C, positions underlined

ATGTCCAATAAGAACGTGGATACAGCTCCAGATAATGGTTCTGGTGGCTCCAATCCCCAGCTTCTGTGCCGCATT
ATCCACAGCAACACCGTGCTAATTATAATCCTCATAAGAATCATACTAATCATGGGGGCAATAATGGAGGCGCCCA
GGGTACATGCAGCATCAGCAACAGCAACAAATGTATTACAATCAGATGGGCTACCTATGCCGTATCCGAATGCT
GGTGGCAATTTCAACCAGAATGTATACTTGGTTCTTTATAATACTATGGTTCCAGGTACTCCATATGGTGCAGGTG
CCCCGGGCGTGCCCTCCAAGCTACCCAGTGTCTCCTTCTCCAGCATTGCTCTCCACAAAAGACTCATGTGCCAGT
AAGTCTGCGTACCCCTTCTGGAGAACCTATAAATTTGGTTCTGCTGGGCAAACGAGTTCAGTTTCATCCAATGCC
TCTCCTCATGTTAAGTCTACCTCTATTCTCTCTCTGCTGTCCTAGTCTGCTGCTTCCATTGCTCCAAGTGTG
CATCACCGACTAAGGCCGCGGCTGGATCTGAGGACGCTGGCAAGACTAGAGATGCATTCAAGCAGATGGTTCAGGC
TTTAGCTGCTAAAAAGAAGGCTGCAGCTGCTGCCAAGGCTGAAGGTAACAGATGATAAAAGGAGTCTGTGGAG
TCTAAGCCTAAGTCTGCTGTACAGCCGAGACAAAGCCAAAGTTGAGGTTAAGTTAGAGCCAAACCAAAAGTTG
AGGTCAAATTAGAGCCAAAGCCAAAGGTTGAGATTAACCTAGAGCCCAACCAAGTATCGGAGGCTAAACAGAGAC
AAAAACAGTACCTGAGGTTGAAGAACCTTCAACAGTTGCAACGATCAAATCAGAATCCACACCAATTGAGGAGCTC
AAACCCAAAGTTGAGGTAACCCAGAAGTTGCATCCAAGCCTGAACAGGTTGAGTCTGCACAGGAGCCGGCTAACG
ATGTAAAACCATCTGAAGTCTCAGCTAAAGAGGTTCAAGAATCCAAGGAGGAGTCTTTGAAACAACCTAAGGAAGC
TTTGAATGACATTAATCCACCGAGGAGTCAAATAAAAGTTTCTCTTCCAGAAGAGGAGTCTAACGATCACCAG
ACACTTTCCAAATTCAGGAGTTAAAGATACAGAACAATCGCAAGAATTTGAAGAACCTCAAGGAACCCAGGAAA
CTCAGGAAACTCAGGAAACGCAGGAAACCGCAAAGCCTTTGGAAGCTCAAGCACCTCAAGAAATTCAGAAACCCA
AGAACCTGAGGAAGTTCAGGAACAAACTCCTGAAACATCTGAGACTGTTTCTTCTGTCACAGAAGAAGTGTCTCCA
GCTCCAGCACCTGGATTGCCCTTTCAAAGTTCTTGGACAGAATCAGAAAGGCTCCCGCCATTTCAAACCCATTCA
AGCATGAGTACTCAGGTGATTTCCATGGTCCAGACGAAAGTTTAAGTGGTCTGAAAACCATCAAATATGATCCGGG
TTTCTTGATCCAATTCGCCACCTTCAACTTCCATATTGATGAAGATTTTCATCAACGATGTCGGTTCTAAGGTTGTG
ATCCCTGTCAACACCAGAAGAGACAATTTTCCAAGAGCATGTCTACCAGCAATTTCAAAGGTTCTGGTAGAAATA
TGGGTTCTAATTTCAGAAATGATTTTGGATCTAGAAACAACCTCCAAGCAAGGTTCCAAGAAAAAGATAAATCTCA
AAAGAGCCAATCTAGGAGGGGTCCAAAGTACTCCAATGAAGAAGAACCTGCTGAACCAAGCCGTGCTGCTGAATTG
CCTAATAGAGCTGCAAGCAAGTGGGTACCTCGTTTCTTAAGAAATAAGGAGGCTGAAGTCAAGACTGCTCCAGACG
GATCTATAATAGTATCAGAAGAGGACATCAAAGGAAAACTAAATCCCTTTTGAATAAGTTGACGTTGGAATCTT
TGATGATATCTCAAACGATATAATTGCTTTGACCAAGCAAGCTCAATGGGAAGATGACGTCAAGACTTTGAAACAA
GTTATTGAGTCTATATTTGCAAAGGCTTGTGACGAACCTACTGGTCTCTATGTACGCTAAATTATGCGCCAAAA
TGTGCAAGGACACCCACCTGAGATCAAGGAACTAATGAGAAGGGAATACTTTACCCGGTGGTGATTTGGTGAG
AAGAGTGTGATTAATAGATGTCAAACCGAATATCAGAAAGGATGGATTGACCAGCTACCAACCAACGAGGATGGT
TCACCTCTAGAGCCAGAAGTATGTCTGACGAATATTACAACAAGCTGCTGCTAAACGTAGAGGCTTGGTCTGG
TTAAGTTTCATTGGTCAATTGTACCAGTTGAACTTATTATCAGACAATGTCATTTTCCACTGTGTCTTGGTCAGTC
CAAAGAAACAGACAATCCATCAGAAGATACTATGGAGAACCTGGTTCAATTATTGGAAGCTGTAGGTTATAGAATG
GAGATGTCCAGTGAACCTAAGATTAGAAGGTTCTTGATCTCGCCTTTGCCAACATCAAGGCCATTGTTGACGCTA
ACAAGATTCTTCAAGAATAAGATTTATGCTCATGGATTGATGGATCTAAGAAGAAATGGCTGGGCTACAACAGA
AGAGAACGCTGGTCCAAAGACCATTCAAGAGATTCACTTAGAGGCTGAGATCAAGAGATCAGAGGAGGAGAGAGAA
AAAGAAAGAGACAGGCATTCTCGTAGAGGACGTTACAACGACTCCAAACCTAACTCTCGTCAAAGTCTATCTTGGG
GAAGTTCAGTGGTTTACCTAATGATATTAGAAATGCAGGCTATGTTAGAACTCTGACAAATCTCTGGGGCCTGT
CAACAGTTTTTCCAGGGGCAAGTCATCCCGTCAAGCTTCAAAGCAATCAAACAGTGATTTTACCAGTGTTCCTATC
ACTTCCAGTGGTAGTAGAACTGCTCTAAACCACTCTGGCGACAGCTCAACGTCTCTGAACGAACCTCAATCAAGAG
AATATTTCAAAAAATCTGAATCTAGGCCGCTCCTTCCAACATTTTTGCTGCATTAGAGGATCACGATCATGATCA
TGAACATGGCCACGAACAAGAAGAAAACAAGCATTCCAAGGATCTCCAAACGAGAGGCGGAAGCACCTGAGGAC
GACAATGTGGGAGAAATAGCCCAAGCTGACGAAAATGTGCAAGGGCAAGAGCAAGAGGAAGAACGAGAGGAAGAGC
GAGAGCAAGAAGAGGAACAGGAAGTTTAA

Fig. 1 (continued)

Poly(A) binding protein - *PAB1* - PP7435_Ch2-1212
Nucleotide sequences encoding SEQ ID NO:45

SEQ ID NO:74

Komagataella phaffii

Original nucleotide sequence

ATGTCTGTCGATACCAAGGAAGTTCAAGAATCCCTTGATCAACTTAACATTTCAGAATCACCAGCTACTGCTGTGA
CTGAAGAGACTCCAGCTACCTCTACCGAGGCCGCTGAAGAGTCCAACGAGTCATCGACTCAAGCTTCCGAGACCTT
GGCTTCATTGTATGTTGGTGAGCTGGACCCCTACCGTTACTGAGAGTGACTTGTATGAGTTCTTCTCTCCAATCGGT
AGTGTTAATTCCATCCGTGTGTGTCGTGATGCAGTCACCAAGCGTTCGTTGGGTTATGGATACGTTAACTTTTACA
GCCAAGCTGCCGCGAGCGTGCCCTTGGAGGAGCTGAACCTACGCTGAGATCAAAGGTGTCCGTTGTAGATTGATGTG
GTCTCAACGTGACCCCTTCTTGCCTGCTAGTGGTTCCGGTAAACATCTTCATCAAGAACTTAGACCCCTGCTATTGAG
AACAAGACTTTGCACGATACCTTCTCCTCTTTGGAAAAGTTTGTCTGTAAAGTTGCCACTGATGAGAATGGAA
ACTCAAAGGGATTTCGGATTTCATTACGAAAGCGATGAAGCAGCTCAAGCTGCCATTGAGAACATCAACGGTAT
GCTGTTGAATGGTCGTGAAATCTACGTTGGACCTCACTTGGCCAAGAAGGACCGTGAATCTAGATTCCAGGAGATG
ATCAAGAACTACACCAATGTCTTCGTCAAGAACTTTGACACTGAGTCCACTGAGGATGAAGTGAAGAGAGCTGTTTG
AGTCATACGGTCCAATCACTTCGATCCACTTGCAAGTTGATTCGTGAGGGTCACAACAAGGGTTTTGGTTTCGTCAA
CTTTGCTGAACACGACGATGCTGTTAAGGCCGTTGAAGCCTTGAACGACAAGGAGTACAAGGGCAAGCCATTGTAC
GTTGGACGTGCTCAGAAGAAAAACGAGAGAGTCCACGAACTTACCAAGAAGTACGAGGCTGATCGTCTTGAAAAAT
TGCAAAAGTATCAATCAGTTAACTTGTTTCATCAAGAACTTGGATGAGTCCATTGATGATGCCCGTCTTGAGGAAGA
GTTCAAACCATTTGGAACCATCACCTCTGCCAAGGTTCATGTTGGACGAAAATGGAATCTAGAGGTTTCGGTTTC
GTTTGTCTCTCTACTCCAGAGGAAGCCACTAAGGCTATCAGTGAGATGAACCAACGTATGGTTGCCAACAAGCCAT
TATACGTTGCTTTGGCTCAACCAAAAGCCATCAGACGCTCTCAATTGGCTCAGCAGATCCAAGCTAGAAACCAAT
GAGAATGCAACAACAAGCTGGTCCAGGTATACCAAAACAGTTCGTTCAACCTATCTTTACGGGCAACAGCCAGGT
ATGTTGCCACCAGGTGCTAGAGTCCCACCAATGGGTAACCAGATTCCCTCAATTTGCTGGTATGCCAAGACCAGGCC
CATTTCCCAAGGCCAATTCCCTAGAATGGCTCCTAACGGTCAGCCAATGCCTGTGTATGGCCAACCTGTATTCAA
CGGTAACGGTCTCTCAGCAGCGTGGCTACTACCCAAACAACAGACAGAACAAGAACAGGAATCAAAGGAGGAAGGA
AACTCTTTGGCTGCAATCTTGCCTCAGTTGCCATTAGAGCAACAAAAGAGAGTACTGGGTGAGGAGCTTTATCCAA
AGGTGTTGCTACTAACAAGACCCAAGATCCAGAAGCTGCAGGTAAAATCACTGGTATGATCTTGGATTTGGACAA
CCAAGAAATTTTGAACCTGTTGGAAGATGAAGAATTTTCAACACTCATTTCAACGAAGCTCTCCAAGCTTACGAG
GAGTACAAGAGTAAGACTGTTGAACAGCAACAGCAAGCGGATGCACAAGCAAACCTAG

Fig. 1 (continued)

SEQ ID NO:75

Golden gate nucleotide sequence, i.e. SEQ ID NO:74, comprising point mutations C150A, T384C, C707T; positions underlined

ATGTCGTGTCGATACCAAGGAAGTTCAGAATCCCTTGATCAACTTAACATTTCAGAATCACCAGCTACTGCTGTGA
CTGAAGAGACTCCAGCTACCTCTACCGAGGCCGCTGAAGAGTCCAACGAGTCATCGACTCAAGCTCCGAGACATT
GGCTTCATTGTATGTTGGTGAGCTGGACCCCTACCGTTACTGAGAGTGACTTGTATGAGTTCTTCTCTCCAATCGGT
AGTGTTAATTCCATCCGTGTGTGTCGTGATGCAGTCACCAAGCGTTCGTTGGGTTATGGATACGTTAACTTTTACA
GCCAAGCTGCCGGCGAGCGTGCCTTGGAGGAGCTGAACCTACGCTGAGATCAAAGGTGTCCGTTGTAGATTGATGTG
GTCCCAACGTGACCCTTCCCTTGCCTCGTAGTGGTTCCGGTAACATCTTCATCAAGAACTTAGACCCTGCTATTGAG
AACAAGACTTTGCACGATACCTTCTCCTCTTTGGAAAAGTTTTGTCTGTAAAGGTTGCCACTGATGAGAATGGAA
ACTCAAAGGGATTTCGGATTCTGTTTCATTACGAAAGCGATGAAGCAGCTCAAGCTGCCATTGAGAACATCAACGGTAT
GCTGTTGAATGGTCGTGAAATCTACGTTGGACCTCACTTGGCCAAGAAGGACCGTGAATCTAGATTCCAGGAGATG
ATCAAGAACTACACCAATGTCTTTGTCAAGAACTTTGACACTGAGTCCACTGAGGATGAACTGAGAGAGCTGTTTG
AGTCATACGGTCCAATCACTTCGATCCACTTGCAAGTTGATTCTGAGGGTCACAACAAGGGTTTTGGTTTCGTCAA
CTTTGCTGAACACGACGATGCTGTTAAGGCCGTTGAAGCCTTGAACGACAAGGAGTACAAGGGCAAGCCATTGTAC
GTTGGACGTGCTCAGAAGAAAAACGAGAGAGTCCACGAACTTACCAAGAAGTACGAGGCTGATCGTCTTGAAAAAT
TGCAAAAGTATCAATCAGTTAACTTGTTCATCAAGAACTTGGATGAGTCCATTGATGATGCCCGTCTTGAGGAAGA
GTTCAAACCATTTGGAACCATCACCTCTGCCAAGGTGATGTTGGACGAAAATGGAAAATCTAGAGGTTTCGGTTTC
GTTTGTCTCTCTACTCCAGAGGAAGCCACTAAGGCTATCAGTGAGATGAACCAACGTATGGTTGCCAACAAGCCAT
TATACGTTGCTTTGGCTCAACCAAAAGCCATCAGACGCTCTCAATTGGCTCAGCAGATCCAAGCTAGAAACCAAT
GAGAATGCAACAACAAGCTGGTCCAGGTATACCAAAACAGTTCGTTCAACCTATCTTTTACGGGCAACAGCCAGGT
ATGTTGCCACCAGGTGCTAGAGTCCCACCAATGGGTAACCAGATTCCCTCAATTTGCTGGTATGCCAAGACCAGGCC
CATTTCCCAAGGCCAATTCCCTAGAATGGCTCCTAACGTCAGCCAATGCCTGTGTATGGCCAACCTGTATTCAA
CGGTAACGGTCTCTCAGCAGCGTGGCTACTACCCAAACAACAGACAGAACAAGAACAGGAATCAAAAGGAGGAAGGA
AACTCTTTGGCTGCAATCTTGCTCAGTTGCCATTAGAGCAACAAAAGAGAGTACTGGGTGAGGAGCTTTATCCAA
AGGTTGTTGCTACTAACAAGACCCAAGATCCAGAAGCTGCAGGTAAAATCACTGGTATGATCTTGGATTTGGACAA
CCAAGAAATTTTGCAACTGTTGGAAGATGAAGAATTTTCAACACTCATTTCAACGAAGCTCTCCAAGCTTACGAG
GAGTACAAGAGTAAGACTGTTGAACAGCAACAGCAAGCGGATGCACAAGCAAACCTAG

Fig. 1 (continued)

Recycling factor 1 - *RLI1* - PP7435_Chr2-1213
Nucleotide sequences encoding SEQ ID NO:56

SEQ ID NO:76

Komagataella phaffii

ATGAGTGAGAAAAACACACGTATTGCCATTGTTTCAGACGATAGATGCAAACCTAAGAAATGTCGTCAGGAATGCA
GAAAATCCTGTCTGTTGTTAAGACAGGAAAGCTTTGTATTGAAGTTGAAGCATACTCCAAAATTGCCTTCATCAG
TGAGACTCTGTGTATAGGTTTGGGTATTTGTGTGAAAAAGTGTCCATTTGGAGCAATTAAACATTATCAATCTGCCC
ACTAACTTGGAGAAGGAGGTCACCCACAGATACAGTGCCAACTCATTCAAATTGCACAGATTGCCAACACCTAGGC
CAGGTCAAGTCCTAGGTCTTGTGGAACCAATGGTATTGGGAAATCCACCGCATTGAAAATTCCTGCAGGTAAGCA
GAAACCAAATTTAGGTAGATACGACGCGCCACCTGACTGGGAAGAAATTTTGAAATACCTTCAGAGGAAGTGAATTG
CAAAATTATTTTACCAAAGTTTGGAGGACAATATTAAAGCAATTATTAAGCCTCAATATGTTGACAACCTTCCAA
GGGCAATCAAAGGTCCTGTTAATAAGGTTGGTGAGCTTTTGAAAGCCAACTGGAAAGAAAGGAAGGGAAAGCA
TATCATCCAAGTTTTGGAATTGAAACACGTATTGAACAGAGAAGTCGCTAAGCTTTCTGGTGGTGAGCTTCAAAGA
TTTGCAATTGGTATGACTTGTGTCCAAGATGCTAATGTGTACATGTTTGATGAACCTTCATCCTATTTGGATGTTA
AGCAACGTTTGAATGCTGCTCAGATTATTAGATCTTTGTTGCATGCCACCTGTTATGTTATTGTTGTTGAGCACGA
TCTGTCTGTTTTGGATTATCTTCTGATTTTGTGTTTGTATTTGTATGGTGTACCATCGGTTTACGGTGTGTGCTACT
TTGCCAGCCTCTGTGCAGAGAAGGTATCAATATTTTCTTGTATGGTGCATATCCCTACCGAGAATTTGAGGTTTCAGAA
CTGAGTCTTTAAAAATTCAAACCTACAGATGCTTCGGACTCTGTATTAATTGAACAAAGTAGAGCTTTTGAATACCC
CAAAATGGTTAAGACTCAAGGTGATTTCAAGTTAACCCTCAACGCTGGTGACTTCTCAGACTCGGAAATTATTGTG
ATGATGGGTGAAAATGGTACCGGTAAGACCACCTTTATGTAAGCTGTTAGCTGGTGCCATTGAAGGTGATGAAGGCA
CCAAGGTACCCAAATTGAATGTTTCTATGAAGCCCCAAAAGATTGCTCCAAAATTCACAGGAACAGTTAGACAAC
GTTCTTGAAAAAGATTAGATCATCTTTCTTACATCCACAATTCAACACCGATGTGATGAAACCTTTAAAAGTTGAC
GATGTCATTGATCAAGAAGTCCAGCATTGTTCTGGTGGTGAATTGCAAAGAGTTGCCATCGTCTTAGCTTTAGGAA
AAGAGGCAGACATCTACTTGATTGACGAGCCATCTGCTTACTTGGACTCTGAACAGCGTATTATCTGTTCTGAAGGT
TATTAGAAGATTATCCTGCATTCCTCAAGAAAACCTGCTTTCATTGTGGAACACGATTTTCATCATGGCCACATACTTA
GCCGATAAAGTGATGGTATTTGAAGGTCAGCCTTCCAAGAATGCTGTTGCTAACCGGCCAGAAACCCCTTTGACTG
GTTGTAACAGGTTCTTGAAAACCTGAACGTCACCTTTCAGAAGAGATCCAACTCTTTTAGACCCAGAATCAACAA
ACTGGACTCCCAGATGGATAGAGAGCAAAAACCTGAGTGGTAATTACTTCTTCTTGGAGAACACTGAGTTATAA

Fig. 1 (continued)

SEQ ID NO:77

vHH, used in the Examples as model protein (with alphaMF-leader (underlined) and His6 Tag (bold)) Amino acid sequence

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYSDLEGDFDVAVLPFSNSTNNGLLFINTTIASIAAK
EEGVSLEKRQVQLQESGGGLVQAGGSLRLSCAASGRTFTSFAMGWFRQAPGKEREFVASISRSGTLTRYADSAKGR
 FTISVDNAKNTVSLQMDNLNPDDTAVYYCAADLHRPYGPGTQRSDEYDSWGQGTQVTVSSGGSGGGSGGGSGGGSGG
 GSGSGGGSGGGEVQLVESGGALVQPGGSLRLSCAASGFPNRYSMRWYRQAPGKEREWVAGMSSAGDRSSYEDSVK
 GRFTISRDDARNTVYLQMNLSLKPEDTAVYYCNVNVGF~~EW~~QGTQVTVSSGG**HHHHHH**

SEQ ID NO:78

vHH, used in the Examples as model na (with alphaMF-leader (underlined) and His6 Tag (bold)) Nucleotide Sequence

ATGAGATTCCCATCTATTTTCACGCTGTCTTGTTCGCTGCCTCCTCTGCATTGGCTGCCCTGTTAACTACCA
CTGAAGACGAGACTGCTCAAATCCAGCTGAAGCAGTTATCGGTTACTCTGACCTTGAGGGTGATTCGACGTCGC
TGTTTTCCTTTCTCTAACTCCACTAACAACGGTTTGTGTTCATTAACACCACTATCGCTTCCATTGCTGCTAAG
GAAGAGGGTGTCTCTCTCGAGAAGAGACAGGTCAGCTGCAGGAGTCCGGTGGTGGTCTGGTTCAAGCCGGTGGTT
CATTAAGATTGTCTGTCTGCCTCTGGTAGAACTTTCACTTCTTTTCGCAATGGGTGGTTTAGACAAGCACCTGG
 AAAAGAGAGAGAGTTTGTGTCTCTATCTCCAGATCCGGTACTTTAACTAGATACGCTGACTCTGCCAAGGGTAGA
 TTCACTATTTCTGTTGACAACGCCAAGAACACTGTTTCTTTGCAAATGGACAACCTTAACCCAGATGACACCGCAG
 TCTATTACTGTCCGCTGACTTGACACAGACCATACGGTCCAGGAACCCAAAGATCCGATGAGTACGATTCTTGGGG
 TCAGGGAACCTCAAGTCACTGTCTCTTCAGGTGGTGGATCTGGTGGTGGAGGTTCAAGTGGTGGAGGATCCGGTGGT
 GGTGGTCTCTGGTGGTGGTGGATCTGGTGGAGGTGAAGTTCAACTTGTCGAATCCGGTGGTGCACCTGTCCAACCTG
 GTGGATCTCTTAGACTTTCTGTGCCGCCTCCGGTTTTCTGTTAACCGTTACTCTATGCGTTGGTACAGACAAGC
 CCCTGGAAAAGAACGTGAATGGGTTGCCGGAATGTCCTCAGCTGGTGACAGATCCTCCTACGAAGATTCTGTGAAG
 GGACGTTTCACCATCTCCAGAGATGACGCCCGTAACACCGTTTACCTTCAAATGAACCTCCCTTAAGCCTGAGGATA
 CTGCCGTCTACTATTGTAACGTGAATGTCGGATTGTAATACTGGGGACAGGGAACCCAAAGTTACTGTCTCTTCCGG
 TGGACAT**CACCACCACCATCAC**TAATAG

SEQ ID NO:79

Human serum albumin, used in the Examples as model protein - HSAopt (native human secretion signal (underlined)), Amino acid sequence

MKWVTFISLLFLFSSAYSSRGVFRRDAHKSEVAHRFKDLGEENFKALVLIAFAYLQQCPFEDHVKLVNEVTEFAKT
CVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLQHKDDNPNLPRLVRPEVDVMCTAFH
 DNEETFLKKLYEIAARRHPYFYAPELLFFAKRYKAAFTCCQAADKAACLLPKLDELRLDEGKASSAKQRLKCASLQ
 KGERAFKAWAVARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCE
 KPILLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLY~~EY~~ARRHPDYSVVLLRLAKTYET
 LEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCLEFELGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGK
 VGSKCKKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFT
 FHADICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCKADDKETCF~~AE~~EGKKLV~~AA~~SQAALG
 L

Fig. 1 (continued)

SEQ ID NO:80

Human serum albumin, used in the Examples as model na - HSAopt (native human secretion signal (underlined)), Nucleotide Sequence

ATGAAGTGGGTTACTTTTCATCTCCTTGTTGTTCTTGTCTCCTCAGCTTACTCCAGAGGTGTTTTTCAGAAGAGATG
CTCACAAAGTCCGAGGTTGCTCACAGATTCAAGGACTTGGGTGAAGAGAACTTCAAGGCTTTGGTTTTGATCGCTTT
CGCTCAGTACTTGCAGCAGTGTCCATTTCGAGGACCACGTTAAGTTGGTTAACGAGGTTACTGAGTTCGCTAAGACT
TGTGTTGCTGACGAATCCGCTGAGAACTGTGATAAGTCCCTGCACACTTTGTTCCGGTGACAAGTTGTGTACTGTTG
CTACTTTGAGAGAACTTACGGTGAGATGGCTGACTGTTGTGCTAAGCAAGAGCCTGAGAGAAACGAGTGTTCCTT
GCAACACAAGGACGACAACCCAACTTGCCAAGATTGGTTAGACCAGAGGTTGACGTTATGTGTACTGCTTTCCAC
GACAACGAAGAGACTTTCTTGAAGAAGTACTTGTACGAGATCGCTAGAAGACACCCATACTTCTACGCTCCAGAGT
TGTTGTTCTTCGCTAAGAGATACAAGGCTGCTTTCACTGAGTGTGTCAGGCTGCTGATAAGGCTGCTTGTGTTGTT
GCCAAAGTTGGACGAGTTGAGAGATGAGGGTAAGGCTTCTTCCGCTAAGCAGAGATTGAAGTGTGCTTCCTTGCA
AAGTTCGGAGAGAGAGCTTTTAAGGCTTGGGCTGTTGCTAGATTGTCCAGAGATTCCCAAAGGCTGAGTTCGCTG
AGGTTTCCAAGTTGGTTACTGACTTGACTAAGGTTCCACACAGAGTGTGTCACGGTGACTTGTGGAATGTGCTGA
TGACAGAGCTGACTTGGCTAAGTACATCTGTGAGAACCAGGATTCCATCTCCTCCAAGTTGAAAGAATGTTGTGAG
AAGCCTTGTGTTGGAAGTCCCACTGTATCGCTGAGGTTGAAAACGACGAAATGCCAGCTGACTTGCCATCTTTGG
CTGCTGACTTCGTTGAATCCAAGGACGTCTGCAAGAATACGCTGAGGCTAAGGACGTTTTCTTGGGTATGTTCTT
GTATGAGTACGCTAGAAGACATCCAGACTACTCCGTTGTTTTGTTGTTGAGATTGGCTAAGACTTACGAGACTACT
TTGGAGAAGTGTGTGCTGCTGCTGACCCACATGAGTGTACGCTAAGGTTTTTCGACGAGTTCAAGCCATTGGTTG
AGGAACCACAGAACTTGATCAAGCAGAACTGTGAGTGTTCGAGCAGTTGGGTGAGTACAAGTTCCAGAACGCTTT
GTTGGTTAGATACACTAAGAAGGTTCCACAGGTTTCCACTCCAACCTTTGGTTGAGGTTTCCAGAACTTGGGTAAG
GTTGGTTCCAAGTGTGTGAAGCACCAGAGGCTAAGAGAATGCCATGTGCTGAGGACTACTTGTCTGTTGTTTTGA
ACCAGTTGTGTGCTTGCACGAAAAGACACCAGTTTCCGACAGAGTTACTAAGTGTGTACTGAATCCTTGGTTAA
CAGAAGACCTTGTCTTCTCCGCTTTGGAGGTTGACGAGACTTACGTTCCAAAAGAGTTCAACGCTGAGACTTTCACT
TTCCACGCTGACATCTGTACTTTGTCCGAGAAAAGAGAGACAGATCAAGAAGCAGACTGCTTTGGTTGAGTTGGTTA
AGCACAAGCCAAAGGCTACAAAAGAGCAGTTGAAGGCTGTTATGGACGACTTCGCTGCTTTCGTTGAGAAATGTTG
TAAGGCTGACGACAAAGAGACTTGTTCGCTGAAGAGGGTAAGAAGTTGGTTGCTGCTTCCCAAGCTGCTTTGGGT
CTGTAA

SEQ ID NO:81, pAOX1 promoter P. pastoris/ Komagataella phaffii, CBS7435:

CATGTTGGTATTGTGAAATAGACGCAGATCGGGAACACTGAAAAATAACAGTTATTATTCGAGATCTAACATCCAA
AGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCACAGGTTCCATTCTCACACATAAGTGCCAAACGCAA
CAGGAGGGGATACACTAGCAGCAGACCGTTGCAACGCAGGACCTCCACTCCTCTCTCCTCAACACCCCACTTTTG
CCATCGAAAAACCAGCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTTCCAATTCCTTCTATTAGGCTACTAACA
CCATGACTTTATTAGCCTGTCTATCCTGGCCCCCTGGCGAGGTTTCATGTTTGTGTTATTTCCGAATGCAACAAGCT
CCGATTACACCCGAACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCAAATG
GCCCCAAACTGACAGTTTAAACGCTGTCTTGAACCTAATATGACAAAAGCGTGATCTCATCCAAGATGAACCTAAG
TTTGGTTCGTTGAAATGCTAACGGCCAGTTGGTCAAAAAGAACTTCCAAAAGTCGGCATACCGTTTGTCTTGTGTT
GGTATTGATTGACGAATGCTCAAAAATAATCTCATTAAATGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCGGT
GCACCTGTGCCGAAACGCAAATGGGGAAACACCCGCTTTTTGGATGATTATGCATTGTCTCCACATTGTATGCTTC
CAAGATTCTGGTGGGAATACTGCTGATAGCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCTTACTTGACA
GCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTTATCATCATTATTAGCTTACTTTCATAAT
TGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAAAGACTTTTAAACGACAACTTGAGAAGATCAAAAAACAATA
ATTATTTCGAAACG

Fig. 1 (continued)

SEQ ID NO:82, pAOX2 promoter *P. pastoris*/ *Komagataella phaffii*, CBS7435:
 GCTTAAAGGACTCCATTTTCCTAAAAATTTCAAGCAGTCCTCTCAACTAAATTTTTTCCATTCTCTGCACCCAGCC
 CTCTTCATCAACCGTCCAGCCTTCTCAAAAGTCCAATGTAAAGTAGCCTGCAAATTCAGGTTACAACCCCTCAATTT
 TCCATCCAAGGGCGATCCTTACAAAGTTAATATCGAACAGCAGAGACTAAGCGAGTCATCATCACCACCAACGAT
 GGTGAAAAACTTTAAGCATAGATTGATGGAGGGTGTATGGCACTTGGCGGCTGCATTAGAGTTTGAAACTATGGGG
 TAATACATCACATCCGGAACGATCCGACTCCGAGATCATATGCAAAGCACGTGATGTACCCCGTAAACTGCTCGG
 ATTATCGTTGCAATTCATCGTCTTAAACAGTACAAGAACTTTATTTCATGGGTCATTGGACTCTGATGAGGGGCAC
 ATTTCCCCAATGATTTTTTGGGAAAGAAAGCCGTAAGAGGACAGTTAAGCGAAAGAGACAAGACAACGAACAGCAA
 AAGTGACAGCTGTCAGCTACCTAGTGGACAGTTGGGAGTTTCCAATTGGTTGGTTTTGAATTTTTACCCATGTTGA
 GTTGTCTTGGCTTCTCCTTGCAAACAATGCAAGTTGATAAGACATCACCTTCCAAGATAGGCTATTTTTGTGCGCAT
 AAATTTTTGTCTCGGAGTGAAAACCCCTTTTATGTGAACAGATTACAGAAGCGTCTACCCCTTACCCTTACCCGGTTGAGAT
 GGGGAGAAAATTAAGCGATGAGGAGACGATTATTGGTATAAAAGAAGCAACCAAAATCCCTTATTGTCTTTTCTG
 ATCAGCATCAAAGAAATATTGTCTTAAACGGGCTTTTAACTACATTGTCTTACACATTGCAAACCTCTTCTTCT
 ATTTTCGGATCAACTGTATTGACTACATTGATCTTTTTTAACGAAGTTTACGACTTACTAAATCCCCACAACAAAT
 CAACTGAGAAAA

SEQ ID NO:83-96, Methanol inducible promoters and their respective chromosomal positions in the strain *P. pastoris* CBS7435 (Gasser, Steiger, & Mattanovich, 2015)

P _{SHB17}	PP7435_chr2 (340616...341606)	(SEQ ID NO:83) GCAAGGCAACTGAGAAATTGAATAGTGGTTTCAAGCCCCGCTGACTTTTT GTATTATCTCAATGTCGGTGTTCACAGTCCCCAGAAGGGGGCTTTGCC TTCAAGGGAGACGGAAGAGACATCGTCAACCTGGGGAGAAGTATTTCA AATGGCGCAAGTTCGCTAATTTTTACGATTAAGCAGTGCTGTATGGGGT AGTTAATAAATCGGAATATCCTTCTGACGTGACTGTACAAATCTCTT TTTACGTGGTGCGCATACTGGACAGAGGCAGAGTCTCAATTTCTTCTTT TGAGACAGGCTACTACAGCCTGTGATTCTCTTGGTACTTGGATTTGCT TTTATCTGGCTCCGTTGGGAAGTGTGCTGGGTTTTGAAGTATCTTGTG GATGTGTTTTCTAACACTTTTTCAATCTTCTTGGAGTGAGAAATGCAGGAC TTTGAACATCGTCTAGCTCGTTGGTAGGTGAACCGTTTTACCTTGCATG TGTTTAGGAGTTTTCTGGAGTAACCAAGACCGTCTTATCATCGCCGTAA AATCGCTCTTACTGTGCTAATAATCCCGCTGGAAGAGAGTTTCAACA GAAGTAGCACGCAAAGCTCTTGTCAAATGAGAATTGTTAATCGTTTGAC AGGTCACACTCGTGGGCTATGTACGATCAACTTGCCGGCTGTTGCTGGA GAGATGACACCAGTTGTGGCATGGCCAATTGGTATTCAGCCGTACCCT GTATGGAAAATGAGATTATCTTGTCTTGTATCTAGTTTTCTTGCATTTT AGAGTTGCCACATTCGTAGGTTTCAGTACCAATAATGGTAACTTCCAAA CTTCCAACGCAGATACCAGAGATCTGCCGATCCTTCCCCAACAATAGGA GCTTACTACGCCATACATATAGCCTATCTATTTTCACTTTCGCGTGGGT GCTTCTATATAAACGGTTCCTCCATCTTCCGTTTCATACTACTTGAATTT TAAGCACTAAA
P _{ALD4}	PP7435_chr2 (1466285...1467148)	(SEQ ID NO:84) CTTTTCTTTGGGCAAGGAAAAATCAAGAAAAAGCAGAGGTTAAAGTTTT CAGGGGAATGGCAATTGCTTTATATATGGGAGAAAAGTTAACTACGTCGG TGCTGTAGGCGTAGAGAGCGACTGGAGAATGCGTGATGAGGTGCTCTCT TTTCGCCCCCTTGGCGGGGTAAAAATTGCACTACTGCAGAATTACTA CACCCCTATTCCGAGGAGACGAGTGCGACAAAAATGGTAAAGTTCACC CTAGTCTGCGACTTTTAATTGACGGACACCGGCGTTTACATGCGAAAAA AACTAAAGTGCGCGCATTTACGGCCGAGGGGGTCCCCTTGGGACTG AGAGGGGGTGGGATCTGAAATCGAGGAGGTATCAAGACCCCCCGTTTCT CAACTCCCTAATCAAAAATTACGAAGTCCCTCGTTGGAAGGAGTTAAAA TAATTAAGCGGGGTGCGACGCCATACCGAGGTATCTTGCAGGCATTTT ACTAATATTGGAATTCGAGCTCAACTTGAACACAGGCAGGGTTTAGCT ATGTAATCAATGTAATCAATATAATAAAGCACTACCACATCGAAGGTTT GGGAGGGAGGCCAATAGTGTCCCCACAGGGTGTGATATCGCGATTCT TGGGTGAGGAGACACATATTCACTCCTCTCACCACCAACCAACCGGC TCCTCGCAAGATGATTTATCCGATTATCCGGACACTATACTCCCATCCA GTTTGATGCCGATTCATCGATTGTCCTAAATAATCCTTAAATATGTAT

		AGAACGGTACCCTGGGGTTACATAATCCTTATTTAATAATCCCTCCCC ACCGCTTTTCTTTTTTTTTTCTTCTTATTGTC
P _{FDH1}	PF7435_chr3 (423504...424503)	(SEQ ID NO:85) AAATGGCAGAAGGATCAGCCTGGACGAAGCAACCAGTTCCAACGTGCTAA GTAAAGAAGATGCTAGACGAAGGAGACTTCAGAGGTGAAAAGTTTGCAA GAAGAGAGCTGCGGGAAATAAATTTTCAATTTAAGGACTTGAGTGCGTC CATATTCGTGTACGTGTCCAACGTGTTTTCCATTACCTAAGAAAAACATA AAGATTAAAAAGATAAACCCCAATCGGGAACTTTAGCGTGCCGTTTCGG ATTCGGAATAAATTTTGGAGCGCCAGATGACTATGGAAGAGAGGAGTGTA CCAAAATGGCAAGTCGGGGGCTACTCACCGGATAGCCAATACATTCTCT AGGAACCAGGGATGAATCCAGGTTTTTGTGTGTCACGGTAGGTCAAGCAT TCACTTCTTAGGAATATCTCGTTGAAAGCTACTTGAAATCCCATTGGGT GCGGAACCAGCTTCTAATTAATAGTTCGATGATGTTCTCTAAGTGGGA CTCTACGGCTCAAACCTTCTACACAGCATCATCTTAGTAGTCCCTTCCCA AAACACCATTCTAGGTTTTCGAACGTAACGAAACAATGTTCCTCTCTTC ACATTGGGCCGTTACTCTAGCCTTCCGAAGAACCAATAAAAGGGACCGG CTGAAACGGGTGTGGAACTCCTGTCCAGTTTATGGCAAAGGCTACAGA AATCCCAATCTTGTGCGGATGTTGCTCCTCCCAACGCCATATTGTACT GCAGTTGGTGCGCATTTTAGGGAAAATTTACCCAGATGTCCTGATTTT CGAGGGCTACCCCAACTCCCTGTGCTTATACTTAGTCTAATTCTATTCT AGTGTGCTGACCTACACGTAATGATGTCGTAACCCAGTTAAATGGCCGA AAACTATTTAAGTAAGTTTATTTCTCCTCCAGATGAGACTCTCCTTCT TTTCTCCGCTAGTTATCAAACATAAACCTATTTTACCTCAAATACCTC CAACATCACCCACTTAAACA
P _{DAS1}	PF7435_chr3 (634140...634688)	(SEQ ID NO:86) AATGATATAAACAACAATTGAGTGACAGGTCTACTTTGTTCTCAAAAGG CCATAACCATCTGTTTGTCATCTCTTATCACCACACCATCCTCCTCATCT GGCCTTCAATTGTGGGGAACAACCTAGCATCCCAACACCAGACTAATCC ACCCAGATGAAACCAGTTGTGCTTACCAGTCAATGAATGTTGAGCTAA CGTTCCCTTGAAACTCGAATGATCCAGCCTTGCTGCGTATCATCCCTCC GCTATTCCGCCGCTTGCTCCAACCATGTTTCCGCCTTTTTTGAACAAGT TCAAATACCTATCTTTGGCAGGACTTTTCTCCTGCCTTTTTTAGCCTC AGGTCTCGGTTAGCCTCTAGGCAAATTCTGGTCTTCATACCTATATCAA CTTTTCATCAGATAGCCTTTGGGTTCAAAAAAGAACTAAAGCAGGATGC CTGATATATAAATCCCAGATGATCTGCTTTTGAACTATTTTCAGTATC TTGATTCTGTTTACTTACAAACAACCTATTGTTGATTTTATCTGGAGAATA ATCGAACAAA
P _{DAS2}	PF7435_chr3 (632201...633200)	(SEQ ID NO:87) ATTACTGTTTTGGGCAATCCTGTTGATAAGACGCATTCTAGAGTTGTTT CATGAAAGGGTTACGGGTGTTGATTGGTTTGAGATATGCCAGAGGACAG ATCAATCTGTGGTTTGCTAACTGGAAGTCTGGTAAGGACTCTAGCAAG TCCGTTACTCAAAAAGTCATACCAAGTAAGATTACGTAACACCTGGGCA TGACTTTCTAAGTTAGCAAGTCACCAAGAGGGTCCTATTTAACGTTTGG CGGTATCTGAAACACAAGACTTGCCATATCCCATAGTACATCATATTACC TGTCAGCTATGCTACCCACAGAAATACCCAAAAGTTGAAGTGAAAA AATGAAAATTACTGGTAACCTCACCCCATAACTTAATAATTTCTG TAGCCAATGAAAGTAAACCCATTCAATGTTCCGAGATTTAGTATACCTT GCCCTATAAGAAACGAAGGATTTAGCTTCCTTACCCCATGAACAGAA ATCTTCCATTTACCCCTTACTGGAGAGATCCGCCCAAACGAACAGATAA TAGAAAAAGAAAATTCGGACAAATAGAACACTTTCTCAGCCAATTAAAG TCATTCCATGCACTCCCTTTAGCTGCCGTTCCATCCCTTTGTTGAGCAA CACCATCGTTAGCCAGTACGAAAGAGGAACTTAACCGATACCTTGGAG AAATCTAAGGCGGAATGAGTTTAGCCTAGATATCCTTAGTGAAGGGTT GTTCCGATACTTCTCCACATTCAGTCATAGATGGGCAGCTTTGTTATCA TGAAGAGACGGAACGGGCATTAAGGGTTAACCGCCAAATTATATAAAG ACAACATGTCCCAGTTTAAAGTTTTTCTTTTCTTATTCTGTATCCTGA GTGACCGTTGTGTTAATATAACAAGTTCGTTTTTAACCTAAGACCAAAA CCAGTTACAACAAATTATAACCCCTCTAAACACTAAAGTTCACTCTTAT CAAACATCAACATCAAAA

[illegible]

P _{FLD}	PP7435_Ch3 (262519...263518)	(SEQ ID NO:91) CAGCCATTAATCTCACCTCAGTTTTTGAATCAGTAGAATTTTAAATGAA ACAAACGGTTGGTATATTATTTGATAGAGTTGCCAAATTTCCAAAGATA AATTTTTCATCAGGTAATATCCTGAATACCGTAACATAGTGACTATTGG AAGACACTGCTATCATATTATATTTCGGATAAAAAATCCAAACCCAGAC CGACCTCTTGAGTCTCAACTCCAAGTCAGCCGCAACTTTAATTATCCGT GGATTGGGAGCTAGTTTGGACAACGCATCAGTATAATATAACTTTACGG TTCCATTATCAGACGCTATTGCAAGAACTTCCTTTCCATTGATCTCGCC AATGCGGCAGTAATTGATATCGTAGGGTAGGTCTGGAAAGACGCTGGCG CTTGTGTCCCATTCTGCAGGAATCTCTGGCACGGTGCTAATGGTAGTTA TCCAACGGAGCTGAGGTAGTCGATATATCTGGATATGCCGCTATAGGA TAAAAACAGGAGAGGGTGAACCTTGCTTATGGCTACTAGATTGTTCTTG TACTCTGAATTCCTATTATGGGAACTAACTAATCTCATCTGTGTGTT GCAGTACTATTGAATCGTTGTAGTATCTACCTGGAGGGCATTCCATGAA TTAGTGAGATAACAGAGTTGGGTAAGTAGAGAGAATAATAGACGTATGC ATGATTACTACACAACGGATGTCGCACTCTTTCCTTAGTTAAAACTATC ATCCAATCACAAAGATGCGGGCTGGAAAGACTTGCTCCCGAAGGATAATC TTCTGCTTCTATCTCCCTTCCTCATATGGTTTCGCAAGGCTCATGCCCC TTCTTCCTTCGAACTGCCCCGATGAGGAAGTCCTTAGCCTATCAAAGAAT TCGGGACCATCATCGATTTTTAGAGCCTTACCTGATCGCAATCAGGATT TCACTACTCATATAAATACATCGCTCAAAGCTCCAACCTTTGCTTGTTC TACAATCTTGATATTCACA
P _{FGH1}	PP7435_Ch3 (555587...556586)	(SEQ ID NO:92) TGGTTCCTCTCGGTCCAATACCAAAAAATATTATCACCATACAGGTCTC CCTTCGATACCAGTGCAAAGTTGAACCGTGGGATTACCTTGAATCTAC AAAAATAGTGTCACTCACAAGTTTGTTCATCAACCACGCTGCCGCTTGCA AAGGAGAACTGAACATGAAGTTGTTAGGGTTTGTATATTGGAATAAG TGGTGGATTGTTGAAGGCGAACGCACCAAGCTACATCCGTCTCGAGC ACACTGTGAATTTGTACGGAATTGACCAAGAGGTGAGACGATCCTGTA TCCCATTGAGCCGTTATGCTTTGTGGGGGAAACCTATTTCTATCGTAC TAAGAAAACCAATGGTGAACATCATATTCGGTATCAATGGCGACGATTCC AGCATAGCCTGTAGACAGTAACAACACTAGGGCAACAGCAACTAACATA TCTTCATTGATGAAACGTTGTGATCGGTGTGACTTTTATAGTAAAAGCT ACAACCTGTTTGAAATACCAAGATATCATTTGTAATGGCTCAAAGGGTA ATACATCTGAAAAACCTGAAGTGTGGAAAAATCCGATGAGGCCAATCA TGATAACGCAGAAGTCCCATTTTGCCATCTTCTCTTGGTATGAAACGGT AGAAAAATGATCCGAGTATGCCAATTGATACTCTTGATTATGCCCATA GTTTGCCTAGGGTTTAATTGATCTCCTGGTCTATCGATCTGGGACGCAA TGTAGACCCCATTAGTGGAACACTGAAAGGGATCCAACACTCTAGGCG GACCCGCTCACAGTCATTTTCAAGCAATCACCACAGGAATCAACTACTT CTCCAGTCTTCTTGCCTGAAGCTTCAAGCCTACAACATAACACTCTT TACTTAATCTTTGATTCTCGAATTGTTTACCCAATCTTGACAACCTTAGC CTAAGCAATACTCTGGGGTTATATATAGCAATTGCTCTTCTCGCTGTA GCGTTCATTCCATCTTTCTA
P _{TAL1-2}	PP7435_Ch2 (644082...645082)	(SEQ ID NO:93) GATATCGATCTACACTTAATAGTAGATGACGAGGCATCTCTCCAATAGG TACCATATCTGGTGTCTTCTTGTAAATTTAAGAACTGTTGGTCTATGAAT GTAGATTTGTCTATGAACAATGATATATGGGTCAGGAGGACAAGATGGTT TCTCTGAGTTGGGTGTTGTTGAGGTGCCTGGCAAGACTTCGGAGCGTTGAT ATCCCCAAGACTTGTAGTGACCGATAGTTGAAGCGTGTGTTGCAGGAA CGGCACATCAATGCAACTTTTCGTAACCTTTGGAATTGAGAGTTGATGCAC TGATGACGATACCCGAAATTTTGACGATTTTACCAATATGACTTGAAGA CAAGTCTCTCATTTGAAACCTTATTATCGTTACTAAGCAAAACGAGCTGA CAAGAAGGGAAGGTGGTGGTATTTCTCGTTGTTCAAAATATATGATTC TCCTGGCAATATCTGTGATGGCTGTTCAAAAAGTGGAAATCATTTCTGC AGGATCATCTACCAACTTTTTATTGAGCTCCTCATTGAATACGATTAAG TGGTCATTTTGAATCGTCAGTAAGTACTTGTTTACAAGTAAATCTGTGTC TGAGTTGTTCTCTGTAGATGTACTGATTTTCCATACGAACTCCAAAAAT GAACGAACGGAATGCCTTAATGACCTCACCTGAAGTGGTCTGCTGCTGT TCTCCGGGAAGGACACTTGTGTTAAAGACTGATGCTCTATCAAAGGACA TTGCAACAAAGTATAAACGGTTGTGAGCGGGAAAAAGATGTGTAGGTAA

		TTGTCGTAGATGAGACTGATTTCAGTAGAAAAACGCGTCCTGCACTATTTT TTTCTTTCTTCATTACATTTCCCTAATCGGGACAAAATGAATCTAAAGAC GTGGTTATGTAGTACACGCATCGATAGGCTATCCCCATACCAAAACACT ATTTTACCCCATCCTTGACAGGTTATAAATATGCGATAGTATGAGTATC TTCAAATTCAGCTGAAATATC
P _{DAS2}	PP7435_Ch3 (633689-634688)	SEQ ID NO:94) AATAAAAAAACGTTATAGAAAAGAAATTGGACTACGATATGCTCCAATCC AAATTGTCAAAATTGACCACCGAAAAAGAACAAATTGGAATTTGACAAGA GGAACAACCTCACTAGATTCTCAAACGGAGCGTCACCTAGAGTCAGTTTC CAAGTCAATTACAGAAAGTTTGGAAACAGAAGAGGAGTATCTACAATTG AATTCAAACTTAAAGTCGAGCTGTCCGAATTCATGTCGCTAAGGCTTT CTTACTTGGACCCCATTTTTGAAAGTTTCATTAAAGTTTCAGTCAAAAA TTTCATGGACATTTATGACACATTAAAGAGCGGACTACCTTATGTTGAT TCTCTATCCAAAGAGGATTATCAGTCCAAGATCTTGGACTCTAGAATAG ATAACATTCTGTGAAAATGGAAGCGCTGAACCTTCAAGCTTACATTGA TGATTAGAGCAATGATATAAACAACAATTGAGTGACAGGCTACTTTGT TCTCAAAAGGCCATAACCATCTGTTTGCATCTCTTATCACCACACCATC CTCCTCATCTGGCCTTCAATTGTGGGGAACAACCTAGCATCCCCAACCCA GACTAACTCCACCCAGATGAAACCAGTTGTCGCTTACCAGTCAATGAAT GTTGAGCTAACGTTCTTGAACCTCGAATGATCCCAGCCTTGCTGCGTA TCATCCCTCCGCTATTCGCGCGCTTGCTCCAACCATGTTCCGCGCTTT TCGAACAAGTTCAAATACCTATCTTTGGCAGGACTTTTCTCCTGCGCTT TTTTAGCCTCAGGTCTCGGTTAGCCTCTAGGCAAATCTGGTCTTCATA CCTATATCAACTTTTCATCAGATAGCCTTTGGGTTCAAAAAAGAACTAA AGCAGGATGCCTGATATATAAATCCCAGATGATCTGCTTTTGAACTAT TTTCAGTATCTTGATTCGTTTACTTACAAACAACCTATTGTTGATTTTAT CTGGAGAATAATCGAACAAA
P _{CAM1}	PP7435_Ch3 (178828-179827)	SEQ ID NO:95) ATTGTTGTGAATACTCTCCTTCATTTGGATTTCTTGGACTTCGGACTCT CTTGATCTCTCTCGAAAAGTTTAACTCTGTTCATGTATAATTTTACCC GCTGTAGGTCGCTCATAATACCATGAGTATGCACATCTTTTACTCCATT AACTTTCAGGTATGCAAAATACAATGAAGATAGTATATAGCTCAAAGAA TTTAGCATTTTGCATTGATCTAATTGTGACATTTTCTCTATGATATCAT CTAGCTTCTTAAACTCGAGAACTCTCGTCCAACGAGGCAGAAACATTGTC CAGTCTTACGTCAAGATTATTCACGAGTTTCTGGACCGTATCAACGTTT TCCATCTTAAGATTACAGTAAGTATCGTCCTTTTGAACGTCAAAGGTAG AAAAGTTAATTTTGTATTTGGTAGTACACTATGAAACTTGCTCACCCCA ATCTTTCCTCCTGACAGGTTGATCTTTATCCCTCTACTAAATTGCCCA AGTGATATCAAGTAGACTAGATCTCGCGAAAAGAACAGCCTAATAAATCC GAAGCATGATGGCCTCTATCCGGAACGTTAAGAGATGTGGCAACAGG AGGGCAGATAGAATTTTTTAAAGACGCTGAAGAATGTATATAGTCCGT AAAAATGTGATAGTACTTTGTTTAGTGCGTACGCCACTTATTCGGGGCC AATAGCTAAACCCAGGTTTGTGCTGGCAGCAAATCAACTGTAGATTGAAT CTCTCTAACAATAATGGTGTCAATCCCTGGCTGGTCAAGGGGAGGAC TATCTTGCGTGATCCGCTTGGAAAATGTTGTGTATCCCTTCTCAATTG CGGAAAGCATCTGCTACTTCCCATAGGCACCAGTTACCCAATTGATATT TCCAAAAAAGATTACCATATGTTTCATCTAGAAGTATAAATACAAGTGGA CATTCAATGAATATTTTCATTCAATTAGTCATTGACACTTTCATCAACTT ACTACGTCTTATTCAACAAT
P PP7435_Ch1 -0336	PP7435_Ch1 (615194-615193)	SEQ ID NO:96) TATACGGTCTATCCACTTTGGAAACGATGTAGTTGAAACGGGGAAGTAA TAGTGGTTCCCAAACGACATGAAGAGGTTATATAAGTTTGCAAGAGGGT GACACCATTTTAGTTGTGGTTCCCGGGTATTTTTTAATCTTTTTTAGTC TAAGATAGCCTCCCCAGATATTACCGAGTTGGGCCATTTGGGGCGGTAT CGGTGGTATCTGATGGTAGCGGTTTACATGCCTGTGCATTGAACTG GCAAAGAGTATACTATCGTGGGGCCCTGAAGGAGGCAGCAAATGGACCG TCAATTGGTTGATCAGGGACTCAAGACAGGTATTGAGCTTTTCAAACAA AAAGAGTATAGGCGCTGCTACAAGGCATTTACTTCTACTATCAATTTCA TTGAGAATGATCCCGAGTTGGCCGCCAGCTGTGTATCTCAACTGATATC TCTGTTAGATTGTAGGGCAGCCTGTTTGGAAAAGCTAGATCAATTGAAT ATGGCCTTGAAAGATGGTCTTAAATGATCAAGAGAGAGTGCCACAAC

		GCAAGGGTTATTTGAGAACTTGCAAAATTTTAGACCTACAAGGGAAGAT CAGTGAGGCTTTGTCTACAGCAAGAGAAGGGATCTCCATAATAGAACT AGAAGAGATCAGGATAATCAATTTAGATATCCAAGGTTCTTTTGGAAAC AATTAAAGGAAGTGAAGAAATGCACTGAAATCAAATTTGGACAAGAAAA TCAGCTACACTTCAAAGTTTTAAAGTTTGACGCACCAAGTGCCTTGTACA AAGAACTAAGATTAGTCACCTCCAAGAACAATAGATCCCTCCATTTTTT TGCCGATAGAGCTAGTGAAGCTGATCTTTCGCCTGTTGAATTTCTCAGA CATGTATGCCTGTTTATTGGTCTCAACAAAATGGAACCTAATTATATCC TCATCACCAGGAAGTGTTCGAAAACCTCAGTTGAAATCCCAACTGTCCA ACAAGGCGTTAAACAATTGT

SEQ ID NO:97

pG1 (P_{GT1}) comprising the TA(T)_x motif, x=14-16

CAAACATTTGCTCCCCCTAGTCTCCAGGGAAATGTAAATATACTGCTAATAGAAAACAGTAAGACGCTCAGTTGT
CAGGATAATTACGTTTCGACTGTAGTAAACAGGAATCTGTATTGTTAGAAAAGAACGAGAGTTTTTTACGGCGCCGC
CATATTTGGGCCGTGTGAAAACAGCTTGAAACCCCACTACTTTCAAAGGTTCTGTTGCTATACACGAACCATGTTTA
ACCAACCTCGCTTTTGACTTGACTGAAGTCATCGGTTAAACAATCAAGTACCCTAGTCTGTCTGAATGCTCCTTTCC
ATATTCAGTAGGTGTTTCTTGCACTTTTGCAATGCACTGCGGAAGAATTAGCCAATAGCGCGTTTCATATGCGCTTT
TACCCCTCTTTTGTCAAGCGCAAAATGCCTGTAAGATTTGGTGGGGGTGTGAGCCGTTAGCTGAAGTACAACAGG
CTAATTCCTGAAAAAACTGCAGATAGACTTCAAGATCTCAGGGATTCCCACTATTTGGTATTCTGATATGTTTTT
CCTGATATGCATCAAACTCTAATCTAAAACCTGAATCTCCGCTAnGATGACCCCGTTTTTCGTGACAAATTAATTT
CCAACGGGGTCTTGTCCGGATAAGAGAATTTGTTTGATTATCCGTTTCGGATAAATGGACGCCTGCTCCATATTTT
TCCGTTATTACCCACCTGGAAGTGCCCGAATTTTCCGGGGATTACGGATAATACGGTGGTCTGGATTAATTAA
TACGCCAAGTCTTACATTTTGTGCACTCTCGTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGA
TTAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGTCTATGGGTGTTAGATGATGCACCTGGATGCAGT
GAGTTTTGGAGTATAAAGATCCTTAAATTTCCACCCTT

n...(T)_x wherein x is 14-16

pG1-x sequences Figure legend:

- Main regulatory region: **bold**
- Core regulatory region: ***bold, italic and underlined***, or ***double underlined***
- T motif: ***italic and underlined***, may be optionally extended (at the 5'-terminal end of the T motif) by a preceding TA sequence, or (at the 3'-terminal end of the T motif) by a succeeding AT sequence
- 3'-terminal region: **underlined with dotted line**

SEQ ID NO:98

pG1-3 (PG1-D1240-2xTA(T)₁₄): Example comprising two T motifs which are (T)_x (x=14), each extended by preceding TA (extending the T motif at its 5'-end to become a TA(T)_x motif)

CAAACATTTGCTCCCCCTAGTCTCCAGGGAAATGTAAATATACTGCTAATAGAAAACAGTAAGACGCTCAGTTGT
CAGGATAATTACGTTTCGACTGTAGTAAACAGGAATCTGTATTGTTAGAAAAGAACGAGAGTTTTTTACGGCGCCGC
CATATTTGGGCCGTGTGAAAACAGCTTGAAACCCCACTACTTTCAAAGGTTCTGTTGCTATACACGAACCATGTTTA
ACCAACCTCGCTTTTGACTTGACTGAAGTCATCGGTTAAACAATCAAGTACCCTAGTCTGTCTGAATGCTCCTTTCC
ATATTCAGTAGGTGTTTCTTGCACTTTTGCAATGCACTGCGGAAGAATTAGCCAATAGCGCGTTTCATATGCGCTTT
TACCCCTCTTTTGTCAAGCGCAAAATGCCTGTAAGATTTGGTGGGGGTGTGAGCCGTTAGCTGAAGTACAACAGG
CTAATTCCTGAAAAAACTGCAGCTCAGGGATTCCCACTATTTGGTATTCTGATATGTTTTTCTGATATGCATCA
AACTCTAATCTAAAACCTGAATCTCCGCTATTTTTTTTTTTTTTTGATGACCCCGTTTTTCGTGACAAATTAATTTT
CAACGGGGTCTTGT**CCGGATAAGAGAATTTTGTGTTGATTATCCGTTCCGGATAAAATGGACGCCTGCTCCATATTTT**
CCGGTTATTACCCACCTGGAAGTGCCCGAATTTTCCGGGGATTACGGATAATACGGTGGTCTGGATTAATTAAT
ACGAGATCTCAGGGATTCCCACTATTTGGTATTCTGATATGTTTTTCTGATATGCATCAAACTCTAATCTAAAA
CCTGAATCTCCGCTATTTTTTTTTTTTTTTGATGACCCCGTTTTTCGTGACAAATTAATTTCCAACGGGGTCTTGTCC
GGATAAGAGAATTTGTTTGATTATCCGTTCCGGATAAAATGGACGCCTGCTCCATATTTTCCGGTTATTACCCAC
CTGGAAGTGCCCGAATTTTCCGGGGATTACGGATAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACAT
TTTGTGCAGTCTCGTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGATTAGTTGCAGCTTGACC
CCGCCATAGCTAGGCATAGCCAAGTGTCTATGGGTGTTAGATGATGCACCTGGATGCAGTGAGTTTTGGAGTATAAA
AGATCCTTAAATTTCCACCCTT

Fig. 1 (continued)

SEQ ID NO:99

pG1-4 (pG1-D1427-2xTA(T)₁₄): Example comprising two T motifs which are (T)_x (x=14), each extended by preceding TA (extending the T motif at its 5'-end to become a TA(T)_x motif)

CAACATTTGCTCCCCCTAGTCTCCAGGGAAATGTAAATATACTGCTAATAGAAAACAGTAAGACGCTCAGTTGT
CAGGATAATTACGTTCTGACTGTAGTAAAACAGGAATCTGTATTGTTAGAAAGAACGAGAGTTTTTTACGGCGCCGC
CATATTTGGGCCGTGTGAAAACAGCTTGAAACCCCACTACTTTCAAAGGTTCTGTTGCTATACACGAACCATGTTTA
ACCAACCTCGCTTTTGACTTGACTGAAGTCATCGGTAAACAATCAAGTACCCTAGTCTGTCTGAATGCTCCTTTCC
ATATTCAGTAGGTGTTTCTTGCACTTTTGCATGCAGTGCAGGAAGAATTAGCCAATAGCGCGTTTCATATGCGCTTT
TACCCCTCTTTTGTCAAGCGCAAAATGCCGTGAAGATTTGGTGGGGGTGTGAGCCGTAGCTGAAGTACAACAGG
CTAATTCCTGAAAAAACTGCAGCTCAGGGATTCCCACTATTTGGTATTCTGATATGTTTTCTGATATGCATCA
AACTCTAATCTAAAACCTGAATCTCCGCTATTTTTTTTTTTTTTTGATGACCCCGTTTTCGTGACAAATTAATTC
CAACGGGGTCTTTGTCGGATAAGAGAATTTTGTGTTGATTATCCGTTCCGGATAAAATGGACGCCTGCTCCATATTTTT
CCGGTTATTACCCACCTGGAAGTGCCAGAAATTTCCGGGGATTACGGATAATACGGTGGTCTGGATTAATTAAT
ACGCCAAGTCTTACATTTTGTGTCAGTCTCGTGCGAGTATGTGCAATAATAAACAAGATGAGCCAAATTTATTGGAT
TAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACCTGGATGCAGTG
AGTTTTGGAGTATAAAGATCCTTAAAATTCACCCCTTAGATCTCAGGGATTCCCACTATTTGGTATTCTGATATG
TTTTTCCTGATATGCATCAAACTCTAATCTAAAACCTGAATCTCCGCTATTTTTTTTTTTTTTTGATGACCCCGTT
TTCGTGACAAATTAATTTCCAACGGGGTCTTGTCCGGATAAGAGAATTTTGTGTTGATTATCCGTTCCGGATAAAATGG
ACGCCTGCTCCATATTTTTCCGGTTATTACCCACCTGGAAGTGCCAGAAATTTCCGGGGATTACGGATAATACG
GTGCTGATTAATTAATACGCCAAGTCTTACATTTTGTGTCAGTCTCGTGCGAGTATGTGCAATAATAAACAAG
ATGAGCCAATTTATTGGATTAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGAT
GATGCACCTGGATGCAGTGAGTTTGGAGTATAAAGATCCTTAAAAATTCACCCCTT

SEQ ID NO:100, pGAP, P. pastoris/ Komagataella phaffii (GS115), originating from PPUZZLE vectors:

CTTTTTGTAGAAATGTCTTGGTGTCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCTCCGTTGCAACTC
CGAACGACCTGCTGGCAACGTAAAATCTCCGGGGTAAAACCTTAAATGTGGAGTAATGGAACCAGAAACGTCTCTT
CCCTCTCTCTCCTTCCACCGCCCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTCTACGGCCCC
CTTGCAAGCAATGCTCTTCCACGATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGACCTAGCAGCCAGGGAT
GGAAAAGTCCCGGCCGTCGCTGGCAATAATAGCGGGCGGACGCATGTATGAGATTATTGGAAACCACCAGAATCG
AATATAAAGGCGAACACCTTTCCCAATTTTGGTTTCTCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCT
ATTTCAATCAATTGAACAACATACCTGCAGGCC

SEQ ID NO:101: pGAP, P. pastoris/ Komagataella phaffii (GS115), originating from GoldenPiCS vectors

CTTTTTGTAGAAATGTCTTGGTGTCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCTCCGTTGCAACTC
CGAACGACCTGCTGGCAACGTAAAATCTCCGGGGTAAAACCTTAAATGTGGAGTAATGGAACCAGAAACGTCTCTT
CCCTCTCTCTCCTTCCACCGCCCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTCTACGGCCCC
CTTGCAAGCAATGCTCTTCCACGATTACGTTGCGGGTAAAACGGAGGTCGTGTACCCGACCTAGCAGCCAGGGAT
GGAAAAGTCCCGGCCGTCGCTGGCAATAATAGCGGGCGGACGCATGTATGAGATTATTGGAAACCACCAGAATCG
AATATAAAGGCGAACACCTTTCCCAATTTTGGTTTCTCTGACCCAAAGACTTTAAATTTAATTTATTTGTCCCT
ATTTCAATCAATTGAACAACATACAAACACA

SEQ ID NO:102, pCS1 P. pastoris/ Komagataella phaffii:

AGGGCATCATTTGAGGTTTCCACAAAAGGAAGAAACATGGATCCAGAGACATCAACAGAGAGGAAAGCGGGTAGTGA
AGCCGAAGCCACACACAGCCCGATTTGGAAGGGAGTTTCAATCAAGGTGAGTCCAGCCATTTTTTTTCTTTTTT
TTTTTTTTTATTCAGGTGAACCCACCTAATTTTTTAACTGGGATCCAGTGAGCTCGCTGGGTGAAAGCCAACCAT
CTTTTGTTCGGGGAACCGTGTCTGCCCCGTAAAGTTAATTTTTTTTTTCCCGCGCAGCTTTAATCTTTCCGCGAGAG
AAGCGTTTTTACGTAGCGTGGGAACAGAATAATCAGTTCATGTGCTATACAGGCACATGGCAGCAGTCACTATT
TTGCTTTTTTAACCTTAAAGTCGTTTCAATCAATTAATGACCAATCAGATTTTTTGCATTTGCCACTTATCTAAA
AATACTTTTGTATCTCGCAGATACGTTTCAAGTGGTTTCCAGGACAACACCCAAAAAAGGTATCAATGCCACTAGGC
AGTCGGTTTTATTTTTTGGTACCCACGCAAGAAGCACCCACCTCTTTTAGGTTTTAAGTTGTGGGAACAGTAACA
CCGCTAGAGCTTCAGGAAAAACAGTACCTGTGACCGCAATTCACCATGATGCAGAAATGTTAATTTAAACGAGTG
CCAAATCAAGATTCAACAGACAAATCAATCGATCCATAGTTACCATTTCCAGCCTTTTCGTGCTGAGCCCTGCTT
CATTCCTGCCTCAGGTGCATAACTTTGCATGAAAAGTCCAGATTAGGGCAGATTTTGAGTTTAAAAATAGGAAATAT
AAACAAATATACCGCAAAAAGGTTTGTTTATAGCTTTTCGCTGGTGGGTACGGTATAAATACATACTCTCCTC
CCCCCCTGGTTCTCTTTTTTCTTTTGTACTTACATTTTACCGTTCCGTCACCTCACTCAACAACAAAA

Fig. 1 (continued)

SEQ ID NO:103: pMDH3 promoter *P. pastoris*/ *Komagataella phaffii*, CBS7435
GTAGCTTGGGTAGGACTTGACAAGTACGGCTTCCGTGGTCATACCAAACGCCTTTGTTACCGTTGGCTATACCTAA
TGACCAAGGCATTTGTGGATTATAACGGTATCGTAGTTGAAAAATATGACGTAACCACTGGTACTAGCCCCACAA
GGTTGATGCTGAATACGGGAATCAAGGTGCCGATTTTAAAGGAGTAGCCACTGAAGGGTTTGGCTGGGTCAATGCC
TCTTTTATTTTGGGATTAACCTACTTAGATGTCCAAGGCATCCGTGCCGATAGGCGCCGTTACGTCCCCTGATGTAT
TTTTAGGAAGCTCAAACCTTGGGAACGCGCAAGTTATGGCCTAAGGCCATGTAACGAGATAGTCAAGTCAAACCTA
GAAGTATACGGTTTCCCCGAGAAATAGCAGAAATAGGCGACAAATACATACAACATTTTCATTGTGATAGGGGGC
GGCGGTTCCCTAGGAGGGACAACCCCCAGAAACCTTGTAGACTACGTTTTACGACGATGGGTTATTACTGTAAAGG
AAGAAATATACTACCCACCAGTTGAATGTTTGAACGGATCAAAGGTCGAAGGGAGTACACGGCCCAACCAACGTAGC
TACCGGAGAAAGCAAGACTTTCCCAACCAATAGCTCCGGGTTTCTTCTCCGGCAACCCGTCAGTTTTTGTGTGG
CCGGACAAAAATTGCAACCCCTCAGTCTAATTGAAAGGTCGGGCTCCGAGCTCTAGGCGTTTGCAGTGTAAATATTG
CATCCCCCTCCCATAGATAATACTGCGCGAACACAGGGTGCAAATTATGATGACCACACATGCCAGTGACCAAAACA
GTTTTTTAGTCTTTAAAAACCCCTCGGAACCTCTGAGTATATAAAGGCTTCTCATTTCTACAAGCAAAACAAGAAG
AACTTCCACTTTCTAACTTTTTATCTATAGACTTTAGAGTTACAACCAACGAACAATAACAA

SEQ ID NO:104: pPOR1 promoter *P. pastoris*/ *Komagataella phaffii*, CBS7435
CGCATTATTTGGGTGCGCTACTGTCTATTGGTTATGATGGCCTTACCGAGCACGAAGGTGGTTGGATAGACCACAAC
TGGAAACAGTTGTACAAACAGATTGCATTCATTTTTGCCTGTATTGGTACTCGATGGCCATCACCCCTCTTATCT
GTTTCATCCTCAACCGTATTCCATTTTTTGCACCTGCGAGCTTCAGAAGAGGCTGAGGAGAAAGGTATGGATGAGGA
TCAGATTGGAGAGTTCGCTTATGACTACGTGGAAGTACGTCGTGATTTTTTGGCTTGGGGATCAGGCCCAACAAT
GGCTTCAAGGAGCCGGAAGTTCTGGATCAGGTAGTTCCGGTTAATGATTTTCAGCAGTGACCAGAATGTGACTAATC
AGACCAACGAATCTGAGAAGCAGTAGAGTAAATATAGAGATGATATTTAGTGTATTCTAATGCTTATGTAATGTAT
TAAGCAAAAAGTTGTGTTTATGAGTTAGCATTTGTCTTAGCAAACATAAAATATGTCGACATTTGCAACCCGCGAT
GTCTAGTGTTTTTAGATCGATCTTCGATGTGTAGAATAATGCCTCCACGTGATGCCCCGCGATTTTGTGGGTGCG
AATGCCTCCAACATAAACCCATCACGTATAAAAAGCCCTCTTAACCCCTCCCCCTGTTTCGTTTGCCTTCATCACTT
AACCTGAACATATCAA

SEQ ID NO:105: pG3
GTAAATAGCGGCAGCAATCCAGTAACCTTTTCTGAATAGCAGAGCCTTAACTAAAATAATGGCCAGGGTAAAAAAT
TCGAAATTTGACACCAAAAAATAAGACTTGTGCTTATAAGTCTTAACAAAGTCCGCAATTTTGGAGCTAACGGTGG
CGGTTGCTGGGATATTCAATAATGGTAGAATGTTGCTGCGGGTATATGACAGAGCGTGAAACACACTGAACAAGGT
AAATGGAACAACAGCAATTGCAATATGGGGGAGGATAGTCAAGAACAAGCAGCAATGGCAAAGTACTGAATATTC
TCCAAAGCCAAAAGGTCCAGTGGTTTCAACGACAAAGTCTTGTGGTATAGCTTTGGAACAAAAGGACACCGAAAG
ACTCGACAGCGCCCACAATACAGCGTTGTAGAAGAACGAATTGATTGCTCCAGAGCTTCTAATAGTCAGAAGATA
CCCCAAACCTCCGAGCAACGTTAGCACATGACCTAAGAACCAGGCGAAGTGAAGAGTCTGGAATAACGACACCCAG
TCAGTTTTTCTGAGCTCCTGGTGGGATTTGTAGAAGCATTTGATTTGCTTGGAGTGGTTTTATTTGAAGATGGTG
TTGAAGCCATTGTTGCTAAAGAGTCCGAGTTTTGCTTTTAGGGTTTGTTAAGCAAAGGAGGAAAAACTGCGCCGTT
TGAAGTCCCAGGTAGTTTCGCGTGTGAGGCCAGCCAGGGAAAGCTTCCTTCGGTACTTTTTTTCTTTTGCAGGTT
CCGACGGATTAAAGCTTCGGTTATGAGGGGGCGGTAGCCAATCCGGACACAATATTGCGTCGCGAGCTAGTCAC
CCCGCCATAAATATACGCAGGATTGAGGTAATAACATCGATAGTCTTAGTAATTAACAATTCAGTGGCGAATTT
GGCAACATGACGTAAGGCCCACTGTTGTCTATAAAAGGGGATGAATTTTCATGTTTTTGGAGCCCTCCCGACAATT
TATTGAACATCAA

SEQ ID NO:106: pG4
TGGACTGTTCAATTTGAAGTCGATGCTGACGATGTCAAGAGAGATGCTCAATTATATTTGTCAATTTGCTGGTTACA
CTGGAAACGCTACTTTTGTGGCGGAACTCTACCAGTTTGGCCGTCCATGTAAACGATGTGCTTCTGGGCCGTGA
CCGTTTCAACACGAACATAACCAATGACAAATCCACTTACAGGTCTAGTTTATATGGAGGCAATTTGTACCTTACT
TCTTTGGATGTCCCAAGTGGGGCTTTAACGTCTGGTACTAACAATGTCTCGTTTGTCACTACAACTCCGAGGTAA
ATAAAGGATTCTTGTGGGATTTCTCAAGTTTGTGGAGTTGTAACAGGTTTATAAGCATATCGTGCCTTGTGTC
CACAAATGAATCATTTTATTTGCGAGATACATGAACAAAGTGTGAACCTGGGACCCATTACTACAATTTCCACGCA
ACCGTTGTTTTCAAGCCCATATTTTTTGACAATTGTTTCGTTACACCCCAAGTTTGTGTACATCGCTTGAATGA
TGTGTGTCCCGGAGTATTTTCCATATTCAGCTTGAATTCGTATACTCAACCAATATCTGGGGGTATACTTTTATGT
AACCTATACAAATCAACTATACTATTTTACCTTTTCGACCAATCATCTCCCATCTTGTAAAGTTTTGCTTCTATAT
CCCTGACCCTGACATCACCCATGATTCGGCTCAACGGTTCTCTCTACATCGTCCCTCTTTTGGAGAGGGTGTTC
GTTTGACATTTCAAATTAACCCCGCCATCACGCGCAACCGAGACCGCACCCCGAATTTTCAACAATTAACCCACA
CCCTACTCCACCCTATGAGGGTTATTAGAAGTATGATACGATATAAATACCACCGCAAGTTTCCCAAGGGATCGTG
TTCTTCTTCTCCAATTGCAATCATATTTCTGACTCTTTCTAGTTTCAAGTTAATTCCTTTTACACTTGTCTTTTCCC
TTACCTTTATCC

Fig. 1 (continued)

SEQ ID NO:107: pG5

AGACCAGCAGTTTAACTACGCCAAATCCACAGGAATTTCTACATCACAATACCAATGGTAATACCACGACGTCAAGG
AATGGAACGACGACTTGGAGGAAGACTTCGTCAACCTCTTGCGGAGTACCCGAGGCTAAGACAATAAGAAGAAAA
AAAAAAGAAAAGCGGTGGGGGAGGGATTATTAATAAGGATTATGTAACCCAGGGTACCGTTCTATACATATTTA
AGGATTATTTAGGACAATCGATGAAATCGGCATCAAACCTGGATGGGAGTATAGTGTCCGGATAATCGGATAAATCA
TCTTGCGAGGAGCGCTTGGTTGGTTGGTGAGAGGAGTGAATATGTGTCTCCTCACCAAGAATCGCGATATCAG
CACCTGTGGGGGACACTATTGGCCTCCCTCCCAAACCTTCGATGTGGTAGTGCTTTATTATATTGATTACATTGA
TTACATAGCTAAACCTGCCTGGTTGCAAGTTGAGCTCCGAATTCGAATATTAGTAAAATGCCTGCAAGATAACCT
CGGTATGGCGTCCGACCCCGCTTAATTATTTTAACTCCTTTCCAACGAGGACTTCGTAATTTTTGATTAGGGAGTT
GAGAAACGGGGGCTTTGATACCTCCTCGATTTGAGATCCACCCCTCTCAGTCCCAAGTGGGACCCCTCGGC
CGTGAAATGCGCGCACTTTAGTTTTTTTCGCATGTAACGCCGGTGTCCGTCAATTAAGTTCGCAGACTAGGGTG
AACTTTACCATTTTTGTGCGACTCCGTCTCCTCGGAATAGGGGTGTAGTAATTCTGCAGTAGTGCAATTTTTACCC
CGCCAAGGGGGGGCGAAAAGAGACGACCTCATCAGCATTTCTCCAGTCGCTCTCTACGCCACAGACCGACGTAG
TTAACTTTCTCCCATATATAAAGCAATTGCCATTCCCTGAAAACCTTAACCTCTGCTTTTTCTTGATTTTTCTT
GCCAAAGAAAAAG

SEQ ID NO:108: pG7

AATTGATTAAAGTTGAGTGAATTTCAAACCGCTATACACAACAGGACAACCTTTGAGTTTAGAAAAATCCGATGTAG
TGTAACGGCTAGCACGGTCCGCTTTCACCGGGCAGACCCGGGTTCGACTCCCGGCATCGGAGTTTATTTTTCCATT
TCGTTCTTTAGAGTATTCTCCTCAGCATGCCCCCTGAATTTTTCTTTTTTCCATGTGTCCATTTTTCCACTTT
TTTTACAGTTTTTCTCGTGATGTTAATTGGCTACACAAAAGCTGCCACACGAAACCTTAATCACGAAAACTATAC
AGCCTTCACTAATCCGTAGCCCCATAATATGTTGTCCACGTGCTGTTGGGTACTACCTGTAGACTCTCATACCCCA
CTCCGTCTTTCTCCAACAATTAACGCAGTACCGAGATTTATCAGCAGACTCAAATTTGGGCAAACCTCTGTATTTTT
CTTGCCCGCATATTTATGGGTCTCAGGCCCTCCACGTTTTCTGTTTACTTGAAGAATATTGGCTGCGAAAAAGTG
GTAAGGACAACCCCTTTTAATTGGATCCAGTTTTTCCGAAATGTTCCGATCCGTACGTCACTCTCCGAAGCCGTAC
ATTTTCACTCAATCTACGTAGCTTTGGACTCAGCGCTCCTGGAATTGCCAGGACAGTTTACTTGAGTTGATATTC
CTTGATAGATTGTGTGCTTCTTTTTCCAAAATTTGAGGCTTCGTTTGAAAAGTGGAATCTGGTCGCTAGATCACTTC
ATGCCTATTTTTTACGGAATAAAGTGGTACTATGCACCCCTTAAACCTAAAGAAAAACGGAATAATACCCCAA
AACCTGGTGATGTTTTTCGCCCTTTCTTTTTATCCGAGTTTTCTTTTTTCTTGCTGCCAAATTCCTCTCCTGA
CCTTAGCGTCCCCGGAATAAATAACTACTTAAGGACCGAATGAGCCCCAGCTTTTCCCTTCTCTTCATTATTC
CCATAATATAAT

SEQ ID NO:109: pG8

CTGCACAACCATTGCCAGTAAGGACGAAGAGAAGGCCCCACTACCCAAAATTCAGGATAACGTCTTCATACCATGC
AGCGACGCCTACAAGACGCTGTCAAGACATGCCAATTCACGAAGTGAACCTTAAACACATTGATCGGGAAATTGA
CCACCAAGGGAATGCTGGTTGAGGCTGGAAGCGTTGCCAGTGTCTGAGGGAACCTGGACCGAAAGTTTAGTAATGC
ATAAGAGGATATATAGGAATGCAGTAATAATATTAGTACCCATTAAGTGGGCTAAGCCATTGGAAGGCCGTCTG
ACTGATGGTGGTGTCTTCTCATTTAGATAGTGCATTTGCAACTACCGTCTGAGATTGAGTTTGATGTGAAGCTCC
AGCGCCAAAACAGTATAAGAACCTTATCTCCGCATTATTGTTCTTGCGTAAAAGTTTGTTGTGAAGAAACAGGGGTA
GTTGCGCAGATTAGTTGTAATATGCGCATAGGATGGGTCAATTGACTTCTTCTCGAAAGAGCCACACCGTTAGCT
AAAAAAGGACGCGCATCTACCCAAAATAGAATGTGGGGAATAGGACGCGCAACTTCCTCTCAATCACTGGACGT
CAGAAAAACAAATGCGCAATCGAGTCACCTCCCGTGATACCTCCGTGATACCCCTCTCCGTCTATTCTGACAGC
GTCTCCCATGACGTTTCAATCTACTTAGAAAAGATTTGTTTTTTTTCTTCAATTACACGATCTCATCTTCTG
CAAGGGTCTGGAGGACATCACCAATCTGCGACTCCATAACTTAGTCTGAGTTTATATTTACGCTTCATCTGATGA
GTAGGAAGAAAAAGTTTACGAAATTTCCCGCCAACTTGCCCTTCGGAATAAGCAGCCACTCTCCTTCTGCCCAT
AGTAAGCTTGCGCGAGGCCCCAACTTGCCAGAACTTTAAATATGCCAAACAATCTCCCCAATCTAAGTTCTCC
CTCTTCTAAAAA

Fig. 1 (continued)

SEQ ID NO:110: pG1a

GGAATCTGTATTGTTAGAAAGAACGAGAGTTTTTTACGGCGCCGCCATATTGGGCCGTGTGAAAACAGCTTGAAAC
CCCCTACTTTCAAAGGTTCTGTTGCTATACACGAACCATGTTTAACCAACCTCGCTTTTGACTTGACTGAAGTCA
TCGGTTAAACAATCAAGTACCTAGTCTGTCTGAATGCTCCTTTCCATATTCAGTAGGTGTTTCTTGCACTTTTGCA
TGCACTGCGGAAGAATTAGCCAATAGCGCGTTTCATATGCGCTTTTACCCCTCTTTTGTCAAGCGCAAAATGCCT
GTAAGATTTGGTGGGGGTGTGAGCCGTTAGCTGAAGTACAACAGGCTAATTCCTGAAAAAACTGCAGATAGACTT
CAAGATCTCAGGATTCCCCTATTTGGTATCTGATATGTTTTCTGATATGCATCAAAACTCTAATCTAAAAC
CTGAATCTCCGCTATTTTTTTTTTTTTTTTGGATGACCCCGTTTTCTGACAAATTAATTTCCAACGGGGTCTTGTC
CGGATAAGAGAAATTTGTTTGATTATCCGTTCCGATAAATGGACGCCTGCTCCATATTTTCCGGTTATTACCCCA
CCTGGAAGTGCCCGAATTTTCCGGGGATTACGGATAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACA
TTTTGTTGCAGTCTCGTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGATTAGTTGCAGCTTGAC
CCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACCTGGATGCAGTGAGTTTTGGAGTATAA
AAGATCCTTAAAAATCCACCCTT

SEQ ID NO:111: pG1b

CCATATTCAGTAGGTGTTTCTTGCACTTTTGCAATGCACCTGCGGAAGAATTAGCCAATAGCGCGTTTCATATGCGCT
TTTACCCCTCTTTTGTCAAGCGCAAAATGCCTGTAAGATTTGGTGGGGGTGTGAGCCGTAGCTGAAGTACAACA
GGCTAATTCCTGAAAAAACTGCAGATAGACTTCAAGATCTCAGGGATTCCCCTATTTGGTATTCTGATATGTTT
TTCTGATATGCATCAAAACTCTAATCTAAAACCTGAATCTCCGCTATTTTTTTTTTTTTTTTGGATGACCCCGTTT
TCGTGACAAATTAATTTCCAACGGGGTCTTGTCGGATAAGAGAATTTGTTTGATTATCCGTTCCGATAAATGGA
CGCTGCTCCATATTTTCCGGTTATTACCCACCTGGAAGTGCCCGAATTTTCCGGGGATTACGGATAATACGG
TGGTCTGGATTAATTAATACGCCAAGTCTTACATTTTGTGTCAGTCTCGTGCGAGTATGTGCAATAATAACAAGA
TGAGCCAATTTATTGGATTAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATG
ATGCACCTGGATGCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAAATCCACCCTT

SEQ ID NO:112: pG1c

CTGCAGATAGACTTCAAGATCTCAGGGATTCCCCTATTTGGTATTCTGATATGTTTTCTGATATGCATCAAAA
CTCTAATCTAAAACCTGAATCTCCGCTATTTTTTTTTTTTTTTTGGATGACCCCGTTTTCTGACAAATTAATTTCC
AACGGGGTCTTGTCGGATAAGAGAATTTTGTGTTGATTATCCGTTCCGATAAATGGACGCCTGCTCCATATTTTTC
CGGTATTACCCACCTGGAAGTGCCCGAATTTTCCGGGGATTACGGATAATACGGTGGTCTGGATTAATTAATA
CGCCAAGTCTTACATTTTGTGTCAGTCTCGTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGATT
AGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACCTGGATGCAGTGA
GTTTTGGAGTATAAAAGATCCTTAAAAATCCACCCTT

SEQ ID NO:113: pG1d

GACCCCGTTTTCTGACAAATTAATTTCCAACGGGGTCTTGTCGGATAAGAGAATTTGTTTGATTATCCGTTCCG
GATAAATGGACGCCTGCTCCATATTTTCCGGTTATTACCCACCTGGAAGTGCCCGAATTTTCCGGGGATTACG
GATAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTTTGTGTCAGTCTCGTGCGAGTATGTGCAATA
ATAAACAAGATGAGCCAATTTATTGGATTAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGG
GTGTTAGATGATGCACCTGGATGCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAAATCCACCCTT

SEQ ID NO:114: pG1e

CCGGATAAGAGAATTTGTTTGATTATCCGTTCCGATAAATGGACGCCTGCTCCATATTTTCCGGTTATTACCCC
ACCTGGAAGTGCCCGAATTTTCCGGGGATTACGGATAAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTAC
ATTTGTTGCAGTCTCGTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGATTAGTTGCAGCTTGGA
CCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACCTGGATGCAGTGAGTTTTGGAGTATA
AAAGATCCTTAAAAATCCACCCTT

SEQ ID NO:115: pG1f

GCCTGCTCCATATTTTTCCGGTTATTACCCACCTGGAAGTGCCCGAATTTTCCGGGGATTACGGATAAATACGGT
GGTCTGGATTAATTAATACGCCAAGTCTTACATTTTGTGTCAGTCTCGTGCGAGTATGTGCAATAATAACAAGAT
GAGCCAATTTATTGGATTAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGA
TGCACCTGGATGCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAAATCCACCCTT