

SEQUENCE LISTING

<110> BASF Plant Science GmbH

<120> New ketolases for the production of ketocarotenoids in Tagetes

<130> PF 0000058583

<160> 11

<170> PatentIn version 3.1

<210> 1

<211> 1153

<212> DNA

<213> Chlorella zofingiensis

<220>

<221> CDS

<222> (209)..(1147)

<223> beta carotene ketolase

<400> 1

ggatccaagc tttgcaattc atacagaagt gagaaaaatg gcttctatga tatcctcttc	60
agctgtgact acagtcagcc gtgcttctac ggtgcaatcg gccgcggtgg ctccattcgg	120
cggcctcaaa tccatgactg gattcccagt taagaaggtc aacctgaca ttacttccat	180
tacaagcaat ggtggaagag taaagtgc atg gcg cca gat gtg aca cat gtg	232
Met Ala Pro Asp Val Thr His Val	
1 5	
cag cca cgt gta cag tcc ccg gct ggc ccc gat gat gag gat gac gcg	280
Gln Pro Arg Val Gln Ser Pro Ala Gly Pro Asp Asp Glu Asp Asp Ala	
10 15 20	
tta agc ttg tgg aaa gcc caa tat cct atg ccg gag gag aag ggt aca	328
Leu Ser Leu Trp Lys Ala Gln Tyr Pro Met Pro Glu Glu Lys Gly Thr	
25 30 35 40	
gta tcc aaa cct caa gcc gca ctc aaa tac agg cca cca cgc agt gac	376
Val Ser Lys Pro Gln Ala Ala Leu Lys Tyr Arg Pro Pro Arg Ser Asp	
45 50 55	
tgg aag ggt gta tca att gca tgc act gtc atc acc cta tgg aca gct	424
Trp Lys Gly Val Ser Ile Ala Cys Thr Val Ile Thr Leu Trp Thr Ala	
60 65 70	
gtc ttt tac cat ggc tgc tgg caa atc aaa ctc aca ggc cct gat aag	472
Val Phe Tyr His Gly Cys Trp Gln Ile Lys Leu Thr Gly Pro Asp Lys	
75 80 85	
tca gcc tgg tgg gac gtt gtt gca acg ttt ctg gca ctg gag ttc ctc	520
Ser Ala Trp Trp Asp Val Val Ala Thr Phe Leu Ala Leu Glu Phe Leu	
90 95 100	
aac act ggg ctt ttc atc acc acg cat gat gcc atg cat ggg act att	568
Asn Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly Thr Ile	
105 110 115 120	

gcc atc agg aac cgt cgt ttg aat gac cta ctt ggc aat ata gcc atc Ala Ile Arg Asn Arg Arg Leu Asn Asp Leu Leu Gly Asn Ile Ala Ile 125 130 135	616
agc cta tat gcc tgg ttt gac tat gac atg ctg cac aag aag cac tgg Ser Leu Tyr Ala Trp Phe Asp Tyr Asp Met Leu His Lys Lys His Trp 140 145 150	664
gag cat cac aac ttc act ggg tta cca cat aaa gac cca gac ttc cat Glu His His Asn Phe Thr Gly Leu Pro His Lys Asp Pro Asp Phe His 155 160 165	712
cga ggc gat cct gcg cta cat aag tgg ttt ggc agg ttt atg tgg gag Arg Gly Asp Pro Ala Leu His Lys Trp Phe Gly Arg Phe Met Trp Glu 170 175 180	760
tat gca aca cca ctc cag ttt gcc aag atc ttc gca tac ccc ttc ttc Tyr Ala Thr Pro Leu Gln Phe Ala Lys Ile Phe Ala Tyr Pro Phe Phe 185 190 195 200	808
cta caa tcc tta cgg gtg caa tac ccc aat tta tgc gtc ttt ctg gcg Leu Gln Ser Leu Arg Val Gln Tyr Pro Asn Leu Cys Val Phe Leu Ala 205 210 215	856
gct gca ccc ctg gtc agt gcg ttc cga ttg ttc tat ttt ggc acc tat Ala Ala Pro Leu Val Ser Ala Phe Arg Leu Phe Tyr Phe Gly Thr Tyr 220 225 230	904
tta ccc cac ctc ccc tcc aat gct cag gag aca atg ccc tgg gag aaa Leu Pro His Leu Pro Ser Asn Ala Gln Glu Thr Met Pro Trp Glu Lys 235 240 245	952
tct cac agt gct gat gac cct cgg ccg ctg tca ttc ttg aaa tgt tat Ser His Ser Ala Asp Asp Pro Arg Pro Leu Ser Phe Leu Lys Cys Tyr 250 255 260	1000
cac ttt gat tat cac tgg gag cat cac agg tgg cct tat gcc cct tgg His Phe Asp Tyr His Trp Glu His His Arg Trp Pro Tyr Ala Pro Trp 265 270 275 280	1048
tgg gag tta ccc gtg tgt aag cgc atc aca aag aca ctg gat gct gca Trp Glu Leu Pro Val Cys Lys Arg Ile Thr Lys Thr Leu Asp Ala Ala 285 290 295	1096
gtt cca gga gta cag tca gac ggc acg aag aag agt cag ttg gtg aac Val Pro Gly Val Gln Ser Asp Gly Thr Lys Lys Ser Gln Leu Val Asn 300 305 310	1144
taa gaattc	1153

<210> 2
 <211> 312
 <212> PRT
 <213> *Chlorella zofingiensis*

<400> 2

Met Ala Pro Asp Val Thr His Val Gln Pro Arg Val Gln Ser Pro Ala 1 5 10 15
--

Gly Pro Asp Asp Glu Asp Asp Ala Leu Ser Leu Trp Lys Ala Gln Tyr
20 25 30

Pro Met Pro Glu Glu Lys Gly Thr Val Ser Lys Pro Gln Ala Ala Leu
35 40 45

Lys Tyr Arg Pro Pro Arg Ser Asp Trp Lys Gly Val Ser Ile Ala Cys
50 55 60

Thr Val Ile Thr Leu Trp Thr Ala Val Phe Tyr His Gly Cys Trp Gln
65 70 75 80

Ile Lys Leu Thr Gly Pro Asp Lys Ser Ala Trp Trp Asp Val Val Ala
85 90 95

Thr Phe Leu Ala Leu Glu Phe Leu Asn Thr Gly Leu Phe Ile Thr Thr
100 105 110

His Asp Ala Met His Gly Thr Ile Ala Ile Arg Asn Arg Arg Leu Asn
115 120 125

Asp Leu Leu Gly Asn Ile Ala Ile Ser Leu Tyr Ala Trp Phe Asp Tyr
130 135 140

Asp Met Leu His Lys Lys His Trp Glu His His Asn Phe Thr Gly Leu
145 150 155 160

Pro His Lys Asp Pro Asp Phe His Arg Gly Asp Pro Ala Leu His Lys
165 170 175

Trp Phe Gly Arg Phe Met Trp Glu Tyr Ala Thr Pro Leu Gln Phe Ala
180 185 190

Lys Ile Phe Ala Tyr Pro Phe Phe Leu Gln Ser Leu Arg Val Gln Tyr
195 200 205

Pro Asn Leu Cys Val Phe Leu Ala Ala Ala Pro Leu Val Ser Ala Phe
210 215 220

Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Leu Pro Ser Asn Ala
225 230 235 240

Gln Glu Thr Met Pro Trp Glu Lys Ser His Ser Ala Asp Asp Pro Arg
245 250 255

Pro Leu Ser Phe Leu Lys Cys Tyr His Phe Asp Tyr His Trp Glu His
260 265 270

His Arg Trp Pro Tyr Ala Pro Trp Trp Glu Leu Pro Val Cys Lys Arg
275 280 285

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

Thr Lys Lys Ser Gln Leu Val Asn
305 310

<210> 3
<211> 24
<212> DNA
<213> Artificial

<400> 3
gcgcatatgg ctcccaggcg gcaa 24

<210> 4
<211> 26
<212> DNA
<213> Artificial

<400> 4
cggtcgactt actccactac tgctcc 26

<210> 5
<211> 1010
<212> DNA
<213> Scenedesmus vacuolatus

<220>
<221> CDS
<222> (7)..(1002)
<223>

<400> 5
gcgcat atg gct ccc agg cg g caa tca acg ctg ccg cag cag acc aaa 48
Met Ala Pro Arg Arg Gln Ser Thr Leu Pro Gln Gln Thr Lys
1 5 10

gct ggc tct cca acc agt ggc tca gat gct gcc atc cct gag ccc gat 96
Ala Gly Ser Pro Thr Ser Gly Ser Asp Ala Ala Ile Pro Glu Pro Asp
15 20 25 30

gtc atc gac gtg tgg aaa gcg caa tac cct ctg ccg gat gaa aat gta 144
Val Ile Asp Val Trp Lys Ala Gln Tyr Pro Leu Pro Asp Glu Asn Val
35 40 45

gca ggg agc atg aat gag gtc aag cag ttg tac agg cca cct cgc aat 192
Ala Gly Ser Met Asn Glu Val Lys Gln Leu Tyr Arg Pro Pro Arg Asn

50					55					60						
gat	gtg	aag	ggc	ata	agc	att	gcc	ttg	ggc	ctg	att	gca	gcc	tgg	tgc	240
Asp	Val	Lys	Gly	Ile	Ser	Ile	Ala	Leu	Gly	Leu	Ile	Ala	Ala	Trp	Cys	
		65					70					75				
gtg	ctg	ttt	tac	cac	ggc	tgc	tgg	cag	atc	cag	ctg	tct	ggc	agt	cag	288
Val	Leu	Phe	Tyr	His	Gly	Cys	Trp	Gln	Ile	Gln	Leu	Ser	Gly	Ser	Gln	
	80					85					90					
cgc	tcc	tgg	tgg	att	gac	att	gct	ggc	aca	ttt	att	ttg	ttg	gag	ttc	336
Arg	Ser	Trp	Trp	Ile	Asp	Ile	Ala	Gly	Thr	Phe	Ile	Leu	Leu	Glu	Phe	
95					100					105					110	
gtc	aac	aca	ggc	ctt	ttc	atc	acc	acg	cac	gat	gcc	atg	cat	ggc	act	384
Val	Asn	Thr	Gly	Leu	Phe	Ile	Thr	Thr	His	Asp	Ala	Met	His	Gly	Thr	
				115					120					125		
gtt	tgt	tac	agg	aac	cgc	aag	ctg	aac	gat	ctg	ttg	ggt	cgt	ata	gcc	432
Val	Cys	Tyr	Arg	Asn	Arg	Lys	Leu	Asn	Asp	Leu	Leu	Gly	Arg	Ile	Ala	
			130					135					140			
atc	aca	ctg	tac	gcc	tgg	ttt	gac	tac	gac	atg	ctt	cac	agg	aag	cac	480
Ile	Thr	Leu	Tyr	Ala	Trp	Phe	Asp	Tyr	Asp	Met	Leu	His	Arg	Lys	His	
		145					150					155				
tgg	gag	cat	cac	aac	tac	aca	gga	cag	aag	ggc	aaa	gac	cct	gac	ttc	528
Trp	Glu	His	His	Asn	Tyr	Thr	Gly	Gln	Lys	Gly	Lys	Asp	Pro	Asp	Phe	
	160					165					170					
cac	agg	ggc	aac	cct	gca	ctg	cca	gtg	tgg	tat	gcc	agg	ttc	atg	tgg	576
His	Arg	Gly	Asn	Pro	Ala	Leu	Pro	Val	Trp	Tyr	Ala	Arg	Phe	Met	Trp	
175					180					185					190	
gaa	tac	tcc	acc	ccc	ttg	cag	ttt	gcc	aaa	atc	atc	ctg	gtg	agt	cag	624
Glu	Tyr	Ser	Thr	Pro	Leu	Gln	Phe	Ala	Lys	Ile	Ile	Leu	Val	Ser	Gln	
				195					200					205		
gtg	ctg	caa	gcc	ctg	gga	gtg	ccc	tac	aac	aac	ctg	tgt	gtc	tac	atg	672
Val	Leu	Gln	Ala	Leu	Gly	Val	Pro	Tyr	Asn	Asn	Leu	Cys	Val	Tyr	Met	
			210					215					220			
gct	gct	gcg	ccc	ctg	gtg	gcc	gcc	ttc	agg	ctg	ttc	tat	ttt	ggc	acc	720
Ala	Ala	Ala	Pro	Leu	Val	Ala	Ala	Phe	Arg	Leu	Phe	Tyr	Phe	Gly	Thr	
		225					230					235				
tac	ctg	ccg	cac	ttg	ccc	ccc	aac	gcc	cag	gag	gtg	atg	gtg	tgg	cag	768
Tyr	Leu	Pro	His	Leu	Pro	Pro	Asn	Ala	Gln	Glu	Val	Met	Val	Trp	Gln	
	240					245					250					
aag	agt	cac	tct	agt	gat	gct	ccc	tcc	tgg	ctg	tct	ttc	ctc	aag	tgt	816
Lys	Ser	His	Ser	Ser	Asp	Ala	Pro	Ser	Trp	Leu	Ser	Phe	Leu	Lys	Cys	
255					260					265					270	
tac	cac	ttt	gat	tat	cat	tgg	gaa	cac	cac	aga	tgg	cca	tat	gct	ccc	864
Tyr	His	Phe	Asp	Tyr	His	Trp	Glu	His	His	Arg	Trp	Pro	Tyr	Ala	Pro	
				275					280					285		
tgg	tgg	gag	ttg	ccg	aag	gcg	aag	aaa	att	aca	caa	caa	act	cag	cat	912
Trp	Trp	Glu	Leu	Pro	Lys	Ala	Lys	Lys	Ile	Thr	Gln	Gln	Thr	Gln	His	
			290					295					300			

cac	caa	caa	acc	aag	cag	cag	cag	ccc	atg	cag	cag	gca	aaa	gcg	cag	960
His	Gln	Gln	Thr	Lys	Gln	Gln	Gln	Pro	Met	Gln	Gln	Ala	Lys	Ala	Gln	
		305					310					315				

gtt	gtc	tcc	cag	ctg	gcc	cct	gca	gga	gca	gta	gtg	gag	taa	gtc	gaccg	1010
Val	Val	Ser	Gln	Leu	Ala	Pro	Ala	Gly	Ala	Val	Val	Glu				
		320				325					330					

<210> 6
 <211> 331
 <212> PRT
 <213> Scenedesmus vacuolatus

<400> 6

Met	Ala	Pro	Arg	Arg	Gln	Ser	Thr	Leu	Pro	Gln	Gln	Thr	Lys	Ala	Gly
1				5					10					15	

Ser	Pro	Thr	Ser	Gly	Ser	Asp	Ala	Ala	Ile	Pro	Glu	Pro	Asp	Val	Ile
		20					25						30		

Asp	Val	Trp	Lys	Ala	Gln	Tyr	Pro	Leu	Pro	Asp	Glu	Asn	Val	Ala	Gly
		35					40					45			

Ser	Met	Asn	Glu	Val	Lys	Gln	Leu	Tyr	Arg	Pro	Pro	Arg	Asn	Asp	Val
	50					55					60				

Lys	Gly	Ile	Ser	Ile	Ala	Leu	Gly	Leu	Ile	Ala	Ala	Trp	Cys	Val	Leu
65					70					75					80

Phe	Tyr	His	Gly	Cys	Trp	Gln	Ile	Gln	Leu	Ser	Gly	Ser	Gln	Arg	Ser
				85					90					95	

Trp	Trp	Ile	Asp	Ile	Ala	Gly	Thr	Phe	Ile	Leu	Leu	Glu	Phe	Val	Asn
			100					105						110	

Thr	Gly	Leu	Phe	Ile	Thr	Thr	His	Asp	Ala	Met	His	Gly	Thr	Val	Cys
		115					120					125			

Tyr	Arg	Asn	Arg	Lys	Leu	Asn	Asp	Leu	Leu	Gly	Arg	Ile	Ala	Ile	Thr
	130					135					140				

Leu	Tyr	Ala	Trp	Phe	Asp	Tyr	Asp	Met	Leu	His	Arg	Lys	His	Trp	Glu
145					150					155					160

His	His	Asn	Tyr	Thr	Gly	Gln	Lys	Gly	Lys	Asp	Pro	Asp	Phe	His	Arg
				165					170					175	

Gly Asn Pro Ala Leu Pro Val Trp Tyr Ala Arg Phe Met Trp Glu Tyr
180 185 190

Ser Thr Pro Leu Gln Phe Ala Lys Ile Ile Leu Val Ser Gln Val Leu
195 200 205

Gln Ala Leu Gly Val Pro Tyr Asn Asn Leu Cys Val Tyr Met Ala Ala
210 215 220

Ala Pro Leu Val Ala Ala Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu
225 230 235 240

Pro His Leu Pro Pro Asn Ala Gln Glu Val Met Val Trp Gln Lys Ser
245 250 255

His Ser Ser Asp Ala Pro Ser Trp Leu Ser Phe Leu Lys Cys Tyr His
260 265 270

Phe Asp Tyr His Trp Glu His His Arg Trp Pro Tyr Ala Pro Trp Trp
275 280 285

Glu Leu Pro Lys Ala Lys Lys Ile Thr Gln Gln Thr Gln His His Gln
290 295 300

Gln Thr Lys Gln Gln Gln Pro Met Gln Gln Ala Lys Ala Gln Val Val
305 310 315 320

Ser Gln Leu Ala Pro Ala Gly Ala Val Val Glu
325 330

<210> 7
<211> 5905
<212> DNA
<213> Artificial

<400> 7
gattgtcggt tccgccttc agtttaaact atcagtgttt gacaggatat attggcgggt 60
aaacctaaga gaaaagagcg tttattagaa taatcggata tttaaaaggg cgtgaaaagg 120
tttatccggt cgtccatttg tatgtccatg gaacgcagtg gcggttttca tggcttggtta 180
tgactgtttt tttgggttac agtctatgcc tcgggcatcc aagcagcaag cgcgttacgc 240
cgtgggtcga tgtttgatgt tatggagcag caacgatgtt acgcagcagg gcagtcgccc 300
taaaacaaag ttaaacatca tgggggaagc ggtgatcgcc gaagtatcga ctcaactatc 360
agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg ttgctggccg tacatttgta 420
cggctccgca gtggatggcg gcctgaagcc acacagtgat attgatttgc tggttacggt 480

gaccgtaagg cttgatgaaa caacgcggcg agctttgatc aacgaccttt tggaaacttc	540
ggcttccccct ggagagagcg agattctccg cgctgtagaa gtcaccattg ttgtgcacga	600
cgacatcatt ccgtggcggtt atccagctaa ggcggaactg caatttggag aatggcagcg	660
caatgacatt cttgcaggta tcttcgagcc agccacgatc gacattgatc tggctatctt	720
gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg aggaactctt	780
tgatccggtt cctgaacagg atctatttga ggcgctaaat gaaaccttaa cgctatggaa	840
ctcgccgccc gactgggctg gcgatgagcg aaatgtagtg cttacgttgt cccgcatttg	900
gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga	960
gcgcctgccg gccagtatc agcccgcat acttgaagct agacaggctt atcttggaca	1020
agaagaagat cgcttggcct cgcgcgcaga tcagttggaa gaatttgtcc actacgtgaa	1080
aggcgagatc accaaggtag tcggcaaata atgtctagct agaaattcgt tcaagccgac	1140
gccgcttcgc ggcgcgggctt aactcaagcg ttagatgcac taagcacata attgctcaca	1200
gccaaactat cagggtcaagt ctgcttttat tttttttaag cgtgcataat aagccctaca	1260
caaattggga gatatatcat gcatgaccaa aatcccttaa cgtgagtttt cgttccactg	1320
agcgtcagac ccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt	1380
aatctgctgc ttgcaaaaa aaaaaccacc gctaccagcg gtggtttgtt tgccggatca	1440
agagctacca actctttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac	1500
tgtccttcta gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac	1560
atacctcgct ctgctaatac tgttaccagt ggctgctgcc agtggcgata agtcgtgtct	1620
taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg	1680
gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga gatactaca	1740
gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt	1800
aagcggcagg gtcggaacag gagagcgac gagggagctt ccagggggaa acgcctggta	1860
tctttatagt cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc	1920
gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg gcctttttac ggttctggc	1980
cttttgctgg ctttttgctc acatgttctt tcctgcgtta tcccctgatt ctgtggataa	2040
ccgtattacc gcctttgagt gagctgatac cgctcgccgc agccgaacga ccgagcgag	2100
cgagtcagtg agcgaggaag cggaagagcg cctgatgcgg tattttctcc ttacgcctct	2160
gtgcggtatt tcacaccgca tatggtgcac tctcagtaca atctgctctg atgccgcata	2220
gttaagccag tatacactcc gctatcgcta cgtgactggg tcatggctgc gccccgacac	2280

ccgccaacac	ccgctgacgc	gccctgacgg	gcttgtctgc	tcccggcatc	cgcttacaga	2340
caagctgtga	ccgtctccgg	gagctgcatg	tgtcagaggt	tttcaccgtc	atcaccgaaa	2400
cgcgcgaggc	aggggtgcctt	gatgtgggcg	ccggcggtcg	agtggcgacg	gcgcggcttg	2460
tccgcgccct	ggtagattgc	ctggccgtag	gccagccatt	tttgagcggc	cagcggccgc	2520
gataggccga	cgcggaagcg	cggggcgtag	ggagcgcagc	gaccgaaggg	taggcgcttt	2580
ttgcagctct	tcggtctgtc	gctggccaga	cagttatgca	caggccaggc	gggttttaag	2640
agttttaata	agttttaaag	agttttaggc	ggaaaaatcg	ccttttttct	cttttatatc	2700
agtcacttac	atgtgtgacc	ggttcccaat	gtacggcttt	gggttcccaa	tgtacggggt	2760
ccggttccca	atgtacggct	ttgggttccc	aatgtacgtg	ctatccacag	gaaagagacc	2820
ttttcgacct	ttttcccctg	ctagggcaat	ttgccctagc	atctgctccg	tacattagga	2880
accggcggat	gcttcgccct	cgatcagggt	gcggtagcgc	atgactagga	tcggggccagc	2940
ctgccccgcc	tctctcttca	aatcgtactc	cggcagggtca	tttgaccoga	tcagcttgcg	3000
cacggtgaaa	cagaacttct	tgaactctcc	ggcgctgcca	ctgcgttcgt	agatcgtctt	3060
gaacaaccat	ctggcttctg	ccttgcctgc	ggcgcgggct	gccaggcggt	agagaaaacg	3120
gccgatgccg	ggatcgatca	aaaagtaatc	gggggtgaacc	gtcagcacgt	ccgggttctt	3180
gccttctgtg	atctcgcggt	acatccaatc	agctagctcg	atctcgatgt	actccggccg	3240
cccggtttcg	ctctttacga	tcttgtagcg	gctaatacaag	gcttcaccct	cggataaccgt	3300
caccaggcgg	ccgttcttgg	ccttcttcgt	acgctgcatg	gcaacgtgcg	tggtgtttaa	3360
ccgaatgcag	gtttctacca	ggtcgtcttt	ctgctttccg	ccatcggtc	gccggcagaa	3420
cttgagtacg	tccgcaacgt	gtggacggaa	cacgcggccg	ggcttgtctc	ccttcccttc	3480
ccggtatcgg	ttcatggatt	cggttagatg	ggaaaccgcc	atcagtacca	ggtcgtaatc	3540
ccacacactg	gccatgccgg	ccggccctgc	ggaaacctct	acgtgcccg	ctggaagctc	3600
gtagcggatc	acctcgccag	ctcgtcggtc	acgttctgac	agacggaaaa	cggccacgtc	3660
catgatgctg	cgactatcgc	gggtgcccac	gtcatagagc	atcggaacga	aaaaatctgg	3720
ttgctcgtcg	cccttggggc	gcttccta	cgacggcgca	ccggtcgccg	gcggttgccg	3780
ggattctttg	cggattcgat	cagcggccgc	ttgccacgat	tcaccggggc	gtgcttctgc	3840
ctcgatgcgt	tgccgctggg	cggcctgcgc	ggccttcaac	ttctccacca	ggtcatcacc	3900
cagcgccgcg	ccgatttgta	ccgggcccga	tggtttgcga	ccgctcacgc	cgattcctcg	3960
ggcttggggg	ttccagtgcc	attgcagggc	cggcagacaa	cccagccgct	tacgcctggc	4020
caaccgcccg	ttcctccaca	catggggcat	tcacggcggt	cggtgccctg	ttgttcttga	4080
ttttccatgc	cgccctcttt	agccgctaaa	attcatctac	tcattttattc	atttgctcat	4140

ttactctggt agctgcgcga tgtattcaga tagcagctcg gtaatggtct tgccttggcg	4200
taccgcgtac atcttcagct tgggtgtgatc ctccgccggc aactgaaagt tgacccgctt	4260
catggctggc gtgtctgcc a ggctggccaa cgttgccagcc ttgctgctgc gtgcgctcgg	4320
acggccggca cttagcgtgt ttgtgctttt gctcattttc tctttacctc attaaactcaa	4380
atgagttttg atttaatttc agcggccagc gcctggacct cgcgggcagc gtcgccctcg	4440
ggttctgatt caagaacggt tgtgccggcg gcggcagtgc ctgggtagct cagcgctgc	4500
gtgatacggg actcaagaat gggcagctcg taccggcca gcgcctcggc aacctcaccg	4560
ccgatgcgcg tgcctttgat cccccgcgac acgacaaagg ccgcttgtag ccttccatcc	4620
gtgacctcaa tgcgctgctt aaccagctcc accaggctcg cgggtggcca tatgtcgtaa	4680
gggcttggct gcaccggaat cagcacgaag tcggctgcct tgatcgcgga cacagccaag	4740
tccgccgcct ggggcgctcc gtcgatcact acgaagtcgc gccggccgat ggccttcacg	4800
tcgcggtcaa tcgtcggggc gtcgatgccg acaacggtta gcggttgatc ttcccgcacg	4860
gccgcccaat cgcgggcact gccctgggga tcggaatcga ctaacagaa atcggccccg	4920
gcgagttgca gggcgcgggc tagatgggtt gcgatggtcg tcttgctga cccgcctttc	4980
tggttaagta cagcgataac cttcatgcgt tccccttgcg tatttgttta ttactcatc	5040
gcacatata cgcagcgacc gcatgacgca agctgtttta ctcaaataca catcaccttt	5100
ttagacggcg gcgctcgggt tcttcagcgg ccaagctggc cggccaggcc gccagcttgg	5160
catcagacaa accggccagg atttcatgca gccgcacggt tgagacgtgc gcgggcggct	5220
cgaacacgta cccggccgcg atcatctccg cctcgatctc ttcggtaatg aaaaacggtt	5280
cgtcctggcc gtccctggtgc ggtttcatgc ttgttctct tggcgttcat tctcggcggc	5340
cgccagggcg tcggcctcgg tcaatgcgtc ctacggaag gcaccgcgcc gcctggcctc	5400
ggtgggcgtc acttcctcgc tgcgtcaag tgcgcggtac agggtcgagc gatgcacgcc	5460
aagcagtgca gccgcctctt tcacggtgcg gccttctctg tcgatcagct cgcgggcgtg	5520
cgcgatctgt gccggggtga gggtagggcg ggggccaaac ttcacgcctc gggccttggc	5580
ggcctcgcgc ccgctccggg tcggtcgat gattagggaa cgctcgaact cggcaatgcc	5640
ggcgaacacg gtcaacacca tgcggccggc cggcgtggtg gtaacgcgtg gtgattttgt	5700
gccgagctgc cggtcgggga gctgttggt ggctggtggc aggatatatt gtggtgtaaa	5760
caaattgacg cttagacaac ttaataaacac attgcggacg tctttaatgt actgaattaa	5820
catccgtttg atacttgtct aaaattggct gatttcgagt gcacttatgc ataaaaacaa	5880
tctaatagaca attattacca agcag	5905

<210> 8
<211> 5027
<212> DNA
<213> Artificial

<400> 8
ctgatcatga gcggagaatt aaggagtgca cgttatgacc cccgccgatg acgcgggaca 60
agccgtttta cgtttggaac tgacagaacc gcaacgttga aggagccact cagccgcggg 120
tttctggagt ttaatgagct aagcacatac gtcagaaacc attattgogc gttcaaaagt 180
cgcctaaggt cactatcagc tagcaaatat ttcttgtcaa aaatgctcca ctgacgttcc 240
ataaattccc ctcggtatcc aattagagtc tcatattcac tctcaatcca aataatctcg 300
acatgtctcc ggagaggaga ccagttgaga ttaggccagc tacagcagcc gatatggccg 360
cggtttgtga catcgttaac cattacattg agacgtctac agtgaacttt aggacagagc 420
cacaaacacc acaagagtgg attgatgacc tagagagggt gcaagataga tacccttggt 480
tggttgctga ggttgagggg gttgtggctg gtattgctta cgctgggccc tggaaggcta 540
ggaacgctta cgattggaca gttgagagta ctgtttacgt gtcacatagg catcaaaggt 600
tgggcctagg atctacattg tacacacatt tgcttaagtc tatggaggcg caaggtttta 660
agtctgtggt tgctgttata ggccttccaa acgatccatc tgttagggtg catgaggctt 720
tgggatacac agcgcggggg acattgcgcg cggttgata caagcatggg ggatggcatg 780
atgttggttt ttggcaaagg gattttgagt tgccagctcc tccaaggcca gttaggccag 840
ttaccagat ctgagtcgat cgaccgatct tgctgcgttc ggatattttc gtggagttcc 900
cgccacagac ccgcatgatc cccgatcggt caaacatttg gcaataaagt ttcttaagat 960
tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat tacgttaagc 1020
atgtaataat taacatgtaa tgcacgacgt tatttatgag atgggttttt atgattagag 1080
tcccgaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca aactaggata 1140
aattatcgcg cgcggtgtca tctatgttac tagatcgggc ctgagtcgtt gtaaaacgac 1200
ggccagtga ttatccaact ttgtataata aagttgccat gattacgcca agcttgcatg 1260
ccgtcgacca gatctgatat ctgcggccgc ctgcagcata tgctagagga tcccgggta 1320
ccccactttg tacaagaaag ctgggtggta cggggcccc cctcgagttt aaaccactcc 1380
gccctacact cgtatatata tgcctaaacc tgccccgttc ctcatatgtg atattattat 1440
ttcattatta ggtataagat agtaaacgat aaggaaagac aattttattga gaaagccatg 1500
ctaaaatata gatagatata ccttagcagg tgtttatttt acaacataac ataacatagt 1560
agctagccag caggcaggct aaaacatagt atagtctatc tgcaggggaa ttcttagttc 1620

accaactgac	tcttcttcgt	gccgtctgac	tgtactcctg	gaactgcagc	atccagtgtc	1680
tttgtgatgc	gcttacacac	gggtaactcc	caccaagggg	cataaggcca	cctgtgatgc	1740
tcccagtgat	aatcaaagtg	ataacatttc	aagaatgaca	gcggccgagg	gtcatcagca	1800
ctgtgagatt	tctcccaggg	cattgtctcc	tgagcattgg	aggggaggtg	gggtaaatag	1860
gtgccaaaat	agaacaatcg	gaacgcactg	accaggggtg	cagccgccag	aaagacgcat	1920
aaattggggg	attgcacccg	taaggattgt	aggaagaagg	ggtatgcgaa	gatcttggca	1980
aactggagtg	gtgttgcata	ctccacata	aacctgccaa	accacttatg	tagcgcagga	2040
tcgcctcgat	ggaagtctgg	gtctttatgt	ggtaaccag	tgaagttgtg	atgctcccag	2100
tgcttcttgt	gcagcatgtc	atagtcaaac	caggcatata	ggctgatggc	tatattgcca	2160
agtaggtcat	tcaaacgacg	gttcctgatg	gcaatagtcc	catgcatggc	atcatgcgtg	2220
gtgatgaaaa	gccagtggtt	gaggaaactcc	agtgccagaa	acgttgcaac	aacgtcccac	2280
caggctgact	tatcagggcc	tgtgagtttg	atttgccagc	agccatggta	aaagacagct	2340
gtccataggg	tgatgacagt	gcatgcaatt	gatacacctt	tccagtcact	gcgtggtggc	2400
ctgtatttga	gtgcggcttg	aggtttggat	actgtaccct	tctcctccgg	cataggatat	2460
tgggctttcc	acaagcttaa	cgcgatcatc	tcacatcggg	ggccagccgg	ggactgtaca	2520
cgtggctgca	catgtgtcac	atctggcgcc	atgcacttta	ctcttcacc	attgcttgta	2580
atggaagtaa	tgatcagtgtt	gaccttctta	actgggaatc	cagtcatgga	tttgaggccg	2640
ccgaatggag	ccaccgcggc	cgattgcacc	gtagaagcac	ggctgactgt	agtcacagct	2700
gaagaggata	tcatagaagc	catttttctc	acttctgtat	gaattgcaaa	gcttggatcc	2760
tgaaatcgaa	attcagagtt	ttgatagtga	gagcaaagag	ggacggactt	atgaggattt	2820
cgagtatttc	aagagatggg	acttgttgat	cggacggcta	cgatgatctc	gatttggtta	2880
atccagtatc	tcgcggtgta	tggagttatg	gtaggggttaa	tgggtcaattt	catctaacgg	2940
tagagaatga	tgtaattaga	taagaatctt	gagatactgg	tttagattgg	atgagtgtag	3000
ggtccatctt	atcttgataa	gtggatggtt	tttagagaca	cagtgaatat	tagccaatcg	3060
aagttccata	tcaccatcat	catctgtata	attttgtttt	tttggagat	aataatgatt	3120
gaaattttgg	tagattttat	ttttcattat	ttaccttgta	tgttgagtgg	tcttcaaatt	3180
attgaacgtg	acagattcac	aagaaagtag	attttttata	aatgaaattt	tacttatttt	3240
aaaggatatc	ctattttaatt	tcttttgttt	atggttgtct	gtcagcattt	gacttgcagt	3300
ttcatgctca	tagtcatata	cgttattcta	ggcttttttg	aatatcttat	tacttttttc	3360
gtaatacaat	tttataattt	tatcaaagtt	atacaactat	aactaaaatt	agggttttct	3420
acaaaacaaa	aaaatcttct	aatttttttt	gttgtagcca	gtttactcgt	aagttacaaa	3480

aaaatacaaa tgaaccacaca tgtattatgc gtttaactag gattaccatg tactttcatg	3540
tactcaattc accctatact cttttttttt ttttttctag ttccacccaa tctataaaat	3600
tctgtccatt tgaccaaatt caattaatth ctgtaattgc gattttaaht taatattaca	3660
tgttcactat ttctcgattt gagggaaacc gagtttaaht atgataaaaa tgttgacca	3720
tcactacaaa tatgtttatag ttatacttta atagtgggtg ttttggggat aattgatgaa	3780
ttaagtaaac atgattcttc ttatgaagtt gattgagtga ttattgtatg taaacctatg	3840
tgattgatgt tattggttga ttgagtgatt attgtattag tatgtaagca aagatgattg	3900
ttcttatgag gtaatttggt actcattcat ccttttgcat atgagaaatt gtgttagcgt	3960
acgcaaaaca atagagaaca taaaagatat gtgtatttat ttaaggtgac ttttgttaat	4020
gatattgtag tatctataca ttatatata acttgttgaa tttagagtata agctatcatt	4080
ttctaattat tggttgaatt tatactttta aattaggtgc aaatatatta gatgagtagt	4140
ccattgatct aaatatthcg cataataatt ttaagtacta aaatatgtt gacatattht	4200
ataatthtta aactthaaht tthttatata agctthtggt gacacttaht tcgththaaht	4260
ttggttaata ctgtaataht atgtactthg ttgcctatta gagagacacc ctctaahtcg	4320
tccagagact cttctctthg cagctaaagt ctcccatthc ataaacccta gccaaactg	4380
gtgagcgatt cttctcggc aatttgatta accttctagt tcgtaacgag thtcattgtt	4440
gcttccttca cthttggaac tcgtatgttt tgaagatcaa atgattthtc tggctctgtg	4500
thctthcatt tggthtggtc ctatcgaaht cgtthtatgca cacaacctgt tcgacgtagt	4560
gcgtagctga gattcttctc tataaaaagg tgtattctth gctattgttg gaaggtatca	4620
ttgtthtagta atgggtaaca aagtgatgat gttcaagtgg tcaaagaaca tcacgccttc	4680
acaggttatc aaactgatgc gagctgaaaa agatgtggag aagtctatgg cagththtga	4740
thccgcaacg gcggaatacg caaatgggtg cgtacatgat cagagthctt ttggttatga	4800
gatcagatct ggtcgacggc atgcaagctt ggcgtaatca tggagcctgc thththgtac	4860
aaacttgcca tgattacgcc aagcttgcat gccgtcgacc agatctgata tctgcggccg	4920
cctcgagcat atgctagagg atccccgggt acccaactth tctatacaaa gttggatggc	4980
atgcaagctt ggcgtaatca tggcatagc tgtthctac tagatct	5027

<210> 9
 <211> 4157
 <212> DNA
 <213> Artificial

<400> 9	
ctgatcatga gcggagaatt aaggagtcg cgttatgacc cccgccgatg acgcgggaca	60

agccgtttta	cgtttggaac	tgacagaacc	gcaacgttga	aggagccact	cagccgcggg	120
tttctggagt	ttaatgagct	aagcacatac	gtcagaaacc	attattgcgc	gttcaaaagt	180
cgcctaaggt	cactatcagc	tagcaaatat	ttcttgtcaa	aaatgctcca	ctgacgttcc	240
ataaattccc	ctcggtatcc	aattagagtc	tcatattcac	tctcaatcca	aataatctcg	300
acatgtctcc	ggagaggaga	ccagttgaga	ttaggccagc	tacagcagcc	gatatggccg	360
cggtttgtga	catcgttaac	cattacattg	agacgtctac	agtgaacttt	aggacagagc	420
cacaaacacc	acaagagtgg	attgatgacc	tagagaggtt	gcaagataga	tacccttggg	480
tggttgctga	ggttgagggg	gttggtgctg	gtattgctta	cgctgggccc	tggaaggcta	540
ggaacgctta	cgattggaca	gttgagagta	ctgtttacgt	gtcacatagg	catcaaaggt	600
tgggcctagg	atctacattg	tacacacatt	tgcttaagtc	tatggaggcg	caaggtttta	660
agtctgtggt	tgctgttata	ggccttccaa	acgatccatc	tgtaggttg	catgaggctt	720
tgggatacac	agcgcggggg	acattgcgcg	cggctggata	caagcatggg	ggatggcatg	780
atgttggttt	ttggcaaagg	gattttgagt	tgccagctcc	tccaaggcca	gttaggccag	840
ttaccagat	ctgagtcgat	cgaccgatct	tgctgcgttc	ggatattttc	gtggagtcc	900
cgccacagac	ccggatgata	cccgatcggt	caaacatttg	gcaataaagt	ttcttaagat	960
tgaatcctgt	tgccggtctt	gcgatgatta	tcatataatt	tctgttgaat	tacgttaagc	1020
atgtaataat	taacatgtaa	tgcatgacgt	tatttatgag	atgggttttt	atgattagag	1080
tcccgcgaatt	atacatttaa	tacgcgatag	aaaacaaaat	atagcgcgca	aactaggata	1140
aattatcgcg	cgcggtgtca	tctatgttac	tagatcgggc	ctgagtcggt	gtaaaacgac	1200
ggccagtga	ttatccaact	ttgtataata	aagttgccat	gattacgcca	agcttgcatg	1260
ccgtcgacca	gatctgatat	ctgcggccgc	ctcgagcata	tgctagagga	tccccgggta	1320
ccccactttg	tacaagaaag	ctgggtggta	ccgggcccc	cctcgagttt	aaaccactcc	1380
gccctacact	cgtatatata	tgccataaac	tgccccgttc	ctcatatgtg	atattattat	1440
ttcattatta	ggtataagat	agtaaacgat	aaggaaagac	aattttattga	gaaagccatg	1500
ctaaaatata	gatagatata	ccttagcagg	tgtttatttt	acaacataac	ataacatagt	1560
agctagccag	caggcagggt	aaaacatagt	atagtctatc	tgcatgctcg	agcggccgcc	1620
agtgtgatgg	atatctgcag	aattcgcctt	tcggtcgact	tactccacta	ctgctcctgc	1680
aggggccagc	tgggagacaa	cctgcgcttt	tgctgtctgc	atgggctgct	gctgcttggt	1740
ttgttggtga	tgctgagttt	gttggtgaat	tttcttcgcc	ttcggaact	cccaccaggg	1800
agcatatggc	catctgtggt	gttcccaatg	ataatcaaag	tggtaacact	tgaggaaaga	1860

cagccaggag	ggagcatcac	tagagtgact	cttctgccac	accatcacct	cctgggcgtt	1920
ggggggcaag	tgcggcaggt	aggtgccaaa	atagaacagc	ctgaaggcgg	ccaccagggg	1980
cgcagcagcc	atgtagacac	acaggttggt	gtagggcact	cccagggctt	gcagcacctg	2040
actcaccagg	atgatttttg	caaactgcaa	gggggtggag	tattcccaca	tgaacctggc	2100
ataccacact	ggcagtgcag	ggttgcccct	gtggaagtca	gggtctttgc	ccttctgtcc	2160
tgtgtagttg	tgatgctccc	agtgtctcct	gtgaagcatg	tcgtagtcaa	accaggcgta	2220
cagtgtgatg	gctatacgac	ccaacagatc	gttcagcttg	cggttcctgt	aacaaacagt	2280
gccatgcatg	gcatcgtgcg	tggatgagaa	aaggcctgtg	ttgacgaact	ccaacaaaat	2340
aaatgtgcca	gcaatgtcaa	tccaccagga	gcgctgactg	ccagacagct	ggatctgcca	2400
gcagccgtgg	taaaacagca	cgcaccaggc	tgcaatcagg	cccaaggcaa	tgcttatgcc	2460
cttcacatca	ttgcgaggtg	gcctgtacaa	ctgcttgacc	tcattcatgc	tccctgctac	2520
atthttcatcc	ggcagagggg	attgcgcttt	ccacacgtcg	atgacatcgg	gctcagggat	2580
ggcagcatct	gagccactgg	ttggagagcc	agctttggtc	tgctgcggca	gcgttgattg	2640
ccgcctggga	gccatgcact	ttactcttcc	accattgctt	gtaatggaag	taatgtcagt	2700
gttgaccttc	ttacttgga	atccagtcac	ggatttgagg	ccgccgaatg	gagccaccgc	2760
ggccgattgc	accgtagaag	cacggctgac	tgtagtcaca	gctgaagagg	atatcataga	2820
agccattttt	ctcacttctg	tatgaattgc	aaagctatct	cgaagctagt	gaacaagaaa	2880
aatgtacata	cactgactac	taatttatag	ccatcattac	tgctgatttg	gcaaatggcg	2940
taatctcaaa	ccaaatcacg	aggggtaata	aatgttagcc	acataaattg	gtaagatcaa	3000
aggacgatat	tcagacatct	taagcatcgt	ggtgtgaggc	gataattggt	ttgggctttt	3060
tggctcttgt	ttatggcaag	aaaatccaat	ttgacacatg	ctggctaatt	ttcaaaaagc	3120
ataaaggatt	agtcagggat	tctgaaacta	ggagtgcaat	tttacacaca	gaaaaatcac	3180
ataatcgtaa	cctataatca	cgatgtgatt	caatgccatg	attatatctt	caaactgaaa	3240
tcatagagct	cattaaattg	ttgtcttttg	ttatatatta	atggtacata	atgcccttta	3300
aatgcttttg	tactttgcga	aattttgaaa	taaacaaaat	acattgtttc	tgtatgcaca	3360
tggaaagtac	taataaacga	aaaagaaagt	ttctctgcaa	aaaacaaaaa	aaaatatatt	3420
tagggataga	cagaaatcat	tgtttaggaa	agttgtctct	tattatctct	actacttctc	3480
cacgagaatt	aggcgccatg	caaatgacac	cgatgctaga	aacttcggca	gtgtactatt	3540
gtctggccgt	tttctgatta	tcgctcagtt	tgccaactac	tttggctcct	taacggtcga	3600
gatctgctac	catgcgataa	atatcttgac	agcgttttgg	aaagcgcgat	tgagaggaaa	3660
cgtccaaact	atthgtttat	tgggcttata	attgactaaa	aatgggcctt	gttaacagaa	3720

tctcttgtat	tttttttctt	ttgaaattta	catcactttt	ccataacaaa	cgaaatgcaa	3780
caaactgtaa	aaaaactttg	tgtagatgg	agaatttaca	gccattcgat	tggtttcaaa	3840
atatgtcaag	agtctcaatg	atcggacggg	tgcaaagagg	acacgagcct	ctgcagcgca	3900
agggcgaatt	ccagcacact	ggcggccggt	actagtggat	ccgagctcgg	taccaagctt	3960
ggcgtaatca	tggagcctgc	ttttttgtac	aaacttgcca	tgattacgcc	aagcttgcac	4020
gccgtcgacc	agatctgata	tctgcggccg	cctcgagcat	atgctagagg	atccccgggt	4080
acccaacttt	tctatacaaa	gttggatggc	atgcaagctt	ggcgtaatca	tggtcatagc	4140
tgtttcctac	tagatct					4157

<210> 10
 <211> 960
 <212> DNA
 <213> Brassica napus

<400> 10	
ggtaccaggc	tcgtgtcctc
tttgcaaccg	tccgatcatt
gagatcttag	acatattttg
60	
aaaccaatcg	aatggctgta
aattctcgat	ctaacacagt
ttatacattc	gtttgttatg
120	
ggaagatgat	gtaaatttca
aaaaaataca	agatattctg
ttcagaaggc	ccattaagtc
180	
aatttttaagc	ccaataaacc
aatagtttgg	acgtttcctc
tctatcgcg	ttttccgaaa
240	
cgccagaaaa	acgctgtcaa
gatatttatc	gcatggttagc
agatctcgac	cgttaaagga
300	
ccaaagtagt	tggcaaaactg
agcgataatc	agaaaacagc
cagacaatag	tacaccgccg
360	
aagtttctag	catcggtgca
tttgcattgg	gcctaattct
cgtcgagAAC	tagagataat
420	
aaagagcaac	tttccctaaac
aataatttct	gtctatccct
aaatattttt	ttttgttctt
480	
tgCagagaaa	ctttcttttt
catttattag	tactttccat
gtgcacagag	aaacaatgta
540	
ttttgtttac	ttcaaaattt
cacaaagtac	aaaagcattt
aaaggacatt	atgtgacaca
600	
cgtcatacaa	tttaattgagc
tctatgactt	tcagtttgaa
gatctaatta	tagcattgaa
660	
tcacatcgta	attacagatt
atgattatct	gtgtgttaaaa
ttgtactcct	agtttcagaa
720	
tcccttacta	atcaatcctt
tctacatttt	gaaaaattag
tcagcatgtg	tcaaattgga
780	
tttcttgcca	taaacaagag
ccaaaaaaa	cccaaaacaa
tgatcgccct	acaccacgat
840	
gcttaagatg	tctgtctgat
tatcgtgcct	ttgatcttac
caatttatgt	ggctaacatt
900	
tattaccct	cgtgatttgg
tttgagatta	cgccatttgc
caaataagca	gtaaaagctt
960	

<210> 11
 <211> 1043
 <212> DNA
 <213> Brassica napus

```

<400> 11
ggtaccaggc tcgtgtcctc ttgcaaccg tccgatcatt gagactcttg acatattttg 60
aaaccaatcg aatggctgta aattctccat ctaacacaaa gtttttttac agtttggtgc 120
atttcgtttg ttatggaaaa gtgatgtaaa tttcaaaaga aaaaaaatac aagagattct 180
gttaacaagg cccattttta gtcaattata agccaataa acaaatagtt tggacgtttc 240
ctctcaatcg cgcttttcaa aacgctgtca agatatttat cgcatggtag cagatctcga 300
ccgttaaagg accaaagtag ttggcaaact gagcgataat cagaaaacgg ccagacaata 360
gtacactgcc gaagtttcta gcatcgggtg catttgcatt gcgcctaatt ctctgtggaga 420
agtagtagag ataataaaga gcaactttcc taaacaatga tttctgtcta tccctaaata 480
tatttttttt tgttttttgc agagaaactt tctttttcgt ttatttagtac tttccatgtg 540
catacagaaa caatgtattt tgtttatttc aaaatttcgc aaagtacaaa agcattttaa 600
gggcattatg taccattaat atataacaaa agacaacaat ttaatgagct ctatgatttc 660
agtttgaaga tataatcatg gcattgaatc acatcgtgat tataggttac gattatgtga 720
tttttctgtg tgtaaaattg cactcctagt ttcagaatcc ctgactaatc ctttatgctt 780
tttgaaaatt agccagcatg tgtcaaattg gattttcttg ccataaaca gagccaaaaa 840
gcccaaaaca attatgcctt cacaccacga tgcttaagat gtctgaatat cgtcctttga 900
tcttaccaat ttatgtggct aacatttatt acccctcgtg atttggtttg agattacgcc 960
atttgccaaa tcagcagtaa tgatggctat aaattagtag tcagtgtatg tacatttttc 1020
ttgttcacta gcttcgaaag ctt 1043

```