

Supplementary information

Engineering of methionine-auxotroph *Escherichia coli* via parallel evolution of two enzymes from *Corynebacterium glutamicum*'s direct-sulfurylation pathway enables its recovery in minimal medium

Matan Gabay^{1,#}, Inbar Stern^{1,#}, Nadya Gruzdev², Adi Cohen¹, Lucia Adriana-Lifshits¹, Tamar Ansbacher^{1,4}, Itamar Yadid^{2,3,*} and Maayan Gal^{1,*}

¹Department of Oral Biology, Goldschleger School of Dental Medicine, Faculty of Medicine, Tel Aviv University, Tel Aviv, 6997801, Israel

²Migal - Galilee Research Institute, Kiryat Shmona, 11016, Israel.

³Tel-Hai College, Upper Galilee, 1220800, Israel.

⁴Hadassah Academic College, 91010 Jerusalem, Israel

Equal contribution

*** Corresponding authors**

Itamar Yadid; Email: itamarya@migal.org.il

Maayan Gal; Email: mayyanga@tauex.tau.ac.il

Chroococcidiopsis/1-340 1 MPEQSQTEIFDLGDFQLSTGFISNAKLAYKTHKLNAA...KDNAILFPFLGSAPE... 55
Corynebacterium glutamicum/1-377 1 MPTLAPSGQLEIQAIGDVSTEAIIITNAEIAVHRWGEYRVDKGRSNVLLIEHALTQDSN... 61
Mycobacterium hassiacum/1-374 1 MTMADAEIARLPAEGELGIVDIGALTLSEAVIDNQIAYVERWGEELSPS...RDVVVVVHALTQDSHVAGP 69
Mycobacterium smegmatis/1-380 1 MTIIIEERATDTGMATVPLPAEGEIGLVHIGALTLLENITVFDVTIAVQRWGEELAPD...RGVVMMVHALTQDSHVTPG 78
Anabaena variabilis/1-359 1 MNYQDFISEQTEYYHLVPVFELEGGVITGVQVAYRTWGEKLNLSA...GDGVGLICHALTQDSAD... 60
Anabaena/1-359 1 MNYQDFISEQTEYYHLVPVFELEGGVITGVQVAYRTWGEKLNLSA...GDGVGLICHALTQDSAD... 60
Salinibacter ruber/1-352 1 MSQ-TLTVPTLLENSTTRDVPVAYRTWGEELNAT...GTNAVLVCHALTQDSAD... 50
Thermus thermophilus/1-380 1 MSEIALEAWGEHEALLKPPRSPISIPPPKPRTAVALFPRREGFYTELGGYIPEVRLRFETYGTLSRR...RDNAVLVFHALTQDSAHLAGT 87
Leptospira interrogans/1-366 1 MNETSIGIIE-E...TKYAEFKELILNNSVLSVPIVIAVETYSGLSSS...KNNAILICHALSQDAHAAGY 64
Chloroflexus aurantiacus/1-378 1 MEAIVQAPTPEGVIV-R...TQRMHWTPLTLTSSATGAPITLAVIETYSGLAPD...RSNAILILHALSQDAHAAGF 71
Geobacter lovleyi/1-367 1 MSIGVV-H...EQTITFEAGIRLESRIIAPITLVVLYETSMNAD...CSNIVIMVEHAWTQDAHLAG 81
Deferribacter desulfuricans/1-377 1 MKENSVLV-K...TKYVTFKDDFYFESRIISPIITVAVIETYSKLNK...KDNAILICHALTSAHAAGY 64

Chroococcidiopsis/1-340 56 ALEIYIGEDRPLPRKYFIIILPGLLNGVSSSPSNTAPF...DRGAF...PQTHIA...DIVIAQHRLVTEKF 121
Corynebacterium glutamicum/1-377 62 AADWADLLCPGKAINDIRYCVICTNVIGGNGSTGPGSMHFD...GNFWGNRFPATSIROQVNAEKQF-LDALG 132
Mycobacterium hassiacum/1-374 70 AGPNY...PTPGWDDGVVPGAAIDTRRWCAIATNVLGGGSGSTGPGSLHFD...GKAWGSRFPVAVTRQVRRADLAA-LNAMG 146
Mycobacterium smegmatis/1-380 77 AGDGH...PTAGWDDVAGPGAPIIDTWHCAIATNVLGGGSGSTGPGSLAFD...GKPGWGRFPQITIRQVVAADRAA-LAALG 153
Anabaena variabilis/1-359 61 ADEWDEGLLANKALSDRDFIICSNILGSGYGTGATSINRQTGIPYGAISPAITIRDMVRLQAAL-IOHLG 132
Anabaena/1-359 61 ADEWDEGLLANKALSDRDFIICSNILGSGYGTGATSINRQTGIPYGAISPAITIRDMVRLQAAL-IOHLG 132
Salinibacter ruber/1-352 51 VADWDEGLLPGRALSDRDFVVLVVPVGGSPYGSVAVPTVNDTQERYAGFPFITTRRLRLHRAA-LETLG 122
Thermus thermophilus/1-380 88 YDEETFRSLSPLEQAFREGWDSLVGPRILDPALVYVVSANHLGSGYSGTGPLSLDPLHTGRPYGRDFPPLTIRRLARAQARL-DHLG 176
Leptospira interrogans/1-366 65 HSGSD...KKPGWDDYIIGPKSFTNQYFICSNVIGGKSGSGPLSIHPETSTPYGRSFPVSIQDMVKAQKLL-VESLG 142
Chloroflexus aurantiacus/1-378 72 HSPDT...RKPGWDDAMICGPRPFITNRYFVICSNVIGGKSGSGPLSIHPETSTPYGRSFPVSIQDMVKAQKLL-VESLG 149
Geobacter lovleyi/1-367 62 RREDD...PKPGWDDAIVGPRLLTDYRYCVLCSNVIGGKSGSGPLSIHPETSTPYGRSFPVSIQDMVKAQKLL-LDLHG 139
Deferribacter desulfuricans/1-377 85 NSPDD...QKPGWDDMIIGPKAFITDXYFICSNVIGGKSGYGTGATGASIDESTGPKYGLKFPVFTVMKMLQKLL-IDYLG 142

Chroococcidiopsis/1-340 122 DELYLWLVGALQTYEAWVRF...HMVKRAASIGAPKPSPTWLLWLRTAIEEP...LADPANNFYTD-PQAVQABL...RQAH 203
Corynebacterium glutamicum/1-377 133 ITTVAALVGGSMGARTLEWAAMYF...ETVGAAVLAVSARASAWQIGIQSAQ-IKAIENDHHHEHNYYESGCFATGLGAARRIAH 217
Mycobacterium hassiacum/1-374 147 HQVAAVGGSMGARAELWVIGH...ETVRAGLILAVGARATADQIGTQSTQ-VAATKADPNHQNDDYGTGLKPDVGLQIARRFAH 231
Mycobacterium smegmatis/1-380 154 ITEVAAVGGSMGARAELWVTHF...DDVRAGLVAVGARATADQIGTQSTQ-VAATKADPNHQNDDYGTGLKPDVGLQIARRFAH 238
Anabaena variabilis/1-359 133 IKSLLVIGGSLGGMQVLEWALLYE...EIVQAIAPITSGRHSAWCIGLSEAAQ-ROAIYADPNWKGNYTK-EQPPSQGLAVARMMAM 216
Anabaena/1-359 133 IKSLLVIGGSLGGMQVLEWALLYE...EIVQAIAPITSGRHSAWCIGLSEAAQ-ROAIYADPNWKGNYTK-EQPPSQGLAVARMMAM 216
Salinibacter ruber/1-352 123 QVRVACAVGGSMGMHVLEWAEATDDGAFVRSVLPVAVGGHATAWQIGWAAQ-ROAIFADPKKRDITYP-DDPPTNGLATARMMAM 210
Thermus thermophilus/1-380 177 VEKAI...VIGGSLGGMVALEFALMYP...ERVKKLVVLAAPARHGPAWAFNHL-S-ROALLODPEYQKSN...PAPKSMALARGIAM 254
Leptospira interrogans/1-366 143 EKLLCVAAGSMGMQALEWVIAYP...NSLSNCLVMASTAESHAMQIAFNEVG-ROAILSDPNKKNLKY-D-ENSFRKGLALARVMVG 225
Chloroflexus aurantiacus/1-378 150 IDTLVAAGSMGMQALEWVIEYF...QRVRAIILLATSPSSPQTVAWNYIG-RRAIMADPRMRGDYD-SDAFRDGLAVARMLGH 233
Geobacter lovleyi/1-367 140 IRRLLCVMGGSMGMQALEWATQYF...ERRASVVALATTPRSPQASLNLAVA-RWAIYNDPTFKKKEYK...HNPKDGALARGIGH 221
Deferribacter desulfuricans/1-377 143 EKLLCVAAGSMGMQALEWAVTF...EKTYSIIPITAGRIITPMIAFNTIG-RFAIMKDPNMMNDYVG-KTFPRDGLAARMAHG 226

Chroococcidiopsis/1-340 204 VMALTLPLPGFYREGEEVWRTIGFASMD...DFVSRFWEAFMLPQDPNNLVNQRKTRAA...PSA-GGDLSAALS 274
Corynebacterium glutamicum/1-377 218 L...TYRGELEIDERFGTKAQKNENPLGYPYKPDQRFVAVESLDYQADKLVQRFDAAGSVLLTDALNRHDIRDRGGLNKAL 298
Mycobacterium hassiacum/1-374 232 L...TYRGEVELDTRFGNAPQDDENPLL...GGRYAVESVLEYQGRKLVDFDAGTYVLTDSLSSHVGRGRRGVEAALS 307
Mycobacterium smegmatis/1-380 239 L...TYRGEELDDRANTPQDDDPLT...GGRYAVESVLEYQGGKLVDFDAGTYVLTDSLSSHVGRGRRGVEAALS 314
Anabaena variabilis/1-359 217 S...ARSWQSF TARGRQYDVA...ADQFAIASVQLHHGQKLVQRFDAANTYITLTQAMDSDVAQGR-DYKSVLQS 286
Anabaena/1-359 217 S...ARSWQSF TARGRQYDVA...ADQFAIASVQLHHGQKLVQRFDAANTYITLTQAMDSDVAQGR-DYKSVLQS 286
Salinibacter ruber/1-352 211 V...SYRSRPSLDGRGRDAMPEQD...GTPYAVESVLLHHGNKLVDFDANCVALTRQMDSDVARGRGGYAKVLR 283
Thermus thermophilus/1-380 255 M...SYRAPEGF EARWGAEP...GETYLDYQGEKFLRRFHESVLLSRAMDTHVGRGRRGVEEALKR 319
Leptospira interrogans/1-366 226 I...TSLSDMKREKFRNPPR-GNI...LSTDFAVGSYLIYQGESVDFDANSYIYVTKLHDHSLGKGG-ELTAALS 298
Chloroflexus aurantiacus/1-378 224 I...TFLCEEKLEQRRVRRV-DGALD...LGRFAIEHLEHQAAFRNDRFDANSYLVI TRAMPNDL TARYGSLTAAFDL 308
Geobacter lovleyi/1-367 232 I...TFLSDSMMWQKERRRSKDGDFD...FFGQFEVERLRYNGYNFVDFDANCEFLYRAKALDLYVAVWGYESMTDAF 298
Deferribacter desulfuricans/1-377 227 I...TMSDKSFHKKGRRYATFGIYD...FFGYFEVENLRYNGYKFTEDANS...LYI IKAMDIFLDSYGYGSEEAIGR 303

Chroococcidiopsis/1-340 275 IKA-QMVFVAF TGRMFP...PEEQQRDAERIPNA...KYWEVNSIGGHLTTFSLTEQDRQAMDVLRQVLT... 340
Corynebacterium glutamicum/1-377 299 IKV-PVLVAGVDTDIILYFYHQEHLRNLGNL...LAMAKIVSPVGHDAF...TESRQMDRIVRNFFSLISPEDNPSTYIEFYI 377
Mycobacterium hassiacum/1-374 308 CEV-PVVVGGFTSDRLYFLRLQEEELALMPGC...RGLNVVESIYGHDFLIETEAAGKLI RQTELEASRP... 374
Mycobacterium smegmatis/1-380 315 CPV-PVVVGGFTSDRLYFLRLQEEELALMPGC...RGLNVVESIYGHDFLIETEAAGKLI RQTELEASRP... 380
Anabaena variabilis/1-359 287 IKQ-PALVVAIDSDILYFPTEQEQELADFI PDA...QLGWLQSYGHDAFLIDIATLSQLVINFRQSLSLKFTSDVTT... 359
Anabaena/1-359 287 IKQ-PALVVAIDSDILYFPTEQEQELADFI PDA...QLGWLQSYGHDAFLIDIATLSQLVINFRQSLSLKFTSDVTT... 359
Salinibacter ruber/1-352 284 IEQ-PSLVVGDSDVLYRLSEQEQELAEHLPSA...TLEVLSPHGDFTLIELDANLVLSTWRANICSSVAA... 352
Thermus thermophilus/1-380 320 LRAIPSLFVGDITDILLYFAWEVRQAAKAA...GARYREIKSPHGHDAFLIETDQVEEILDALF... 380
Leptospira interrogans/1-366 299 ATC-RFLVVYSYSDWLYFPAQSREIVKSLAADKRVFYVELQGECHDSFLLKPKQIEILKGFLENPN... 366
Chloroflexus aurantiacus/1-378 309 TRA-RFLALAYSYSDWLYFPAETTYQMAAAQAAGRSFTTHLITDADHAF...LTDVAAQSELIRDFLNRMLTE... 378
Geobacter lovleyi/1-367 299 ITA-PIQFFAFSSDWLYFPPYQTEEMVTCQLGGLKEVEYHLIQSAYGHDAFLLEHETFTPMVRSLLERVAP... 367
Deferribacter desulfuricans/1-377 304 IEA-DSLIFITFTSDFLFSYQTEEIVNIMKNHGNPEWVNI...DYGHDAFLLEFDTQTSCKIEFLSKYINNVANQ... 377

Figure S1. Sequence alignment of MetX of the following bacterial strains: *Geobacter lovleyi* (B3E278), *Leptospira interrogans* (Q8F4I0), *Thermus thermophilus* (Q9RA51), *Anabaena variabilis* (Q3M5Q6), *Deferribacter desulfuricans* (D3P9D1), *Chloroflexus aurantiacus* (A9WKM8), *Salinibacter ruber* (Q2S5A6), *Mycobacterium hassiacum* (K5B926), *Mycobacterium smegmatis* (A0QSZ0), *Anabaena variabilis* (Q3M5Q6) and *Chroococcidiopsis thermalis* (F9GTW70).

Table S1. Protein sequences.

Protein name	Sequence
MetX CG	MPTLAPSGQLEIQAIGDVSTEAGAIKNAEIAHYHRWGEYRVDKEGRSNVVLIEHALTGDSN AADWWADLLGPGKAINTDIYCVICTNVIGGCNGSTGPGSMHPDGNFWGNRFPATSIRDQ VNAEKQFLDALGITTVA AVLGGSMGGARTLEWAAMYPEIVGAAAVLAVSARASAWQIGIQ SAQIKAIENDHHWHEGNYYESGCNPATGLGAARRIAHLTYRGELEIDERFGTKAQKNEN PLGPYRKPDQRFAVESYLDYQADKLVQRF DAGSYVLLTDALNRHDIGRDRGGLNKALES I KVPVLVAGVDTDILYLYHQQEHL SRNLGNLLAMAKIVSPVGHDAFLTESRQMDRIVRNFF SLISPDENNPSTYIEFYI
MetX CG R46H, L315P	MPTLAPSGQLEIQAIGDVSTEAGAIKNAEIAHYHRWGEYRVDKEGHSNVVLIEHALTGDSN AADWWADLLGPGKAINTDIYCVICTNVIGGCNGSTGPGSMHPDGNFWGNRFPATSIRDQ VNAEKQFLDALGITTVA AVLGGSMGGARTLEWAAMYPEIVGAAAVLAVSARASAWQIGIQ SAQIKAIENDHHWHEGNYYESGCNPATGLGAARRIAHLTYRGELEIDERFGTKAQKNEN PLGPYRKPDQRFAVESYLDYQADKLVQRF DAGSYVLLTDALNRHDIGRDRGGLNKALES I KVPVLVAGVDTDILYPYHQQEHL SRNLGNLLAMAKIVSPVGHDAFLTESRQMDRIVRNFF SLISPDENNPSTYIEFYI
MetY CG	MGPKYDNSNADQWGFETR SIHAGQSVDAQTSARNLPIYQSTAFVFDSAEHAKQRFALE DLGPVYSRLTNPTVEALENRIASLEGGVHAVAFSSGQAATTNAILNLAGAGDHIVTSPRLY GGTETLFLITLNLRLGIDVSFVENPDDPESWQAAVQPNTKAFFGETFANPQADVLDIPAVAE VAHRNSVPLIIDNTIATAALVRPLELGADV VVASLTKFYTGNGSGLGGVLIDGGKFDWTV E KDGKSVFPYFVTPDAAYHGLKYADLGAPAFGLKVRVGLLRDTGSTLSAFNAWAAVQGID TLSRLRERHNENAIKVAEFLNNHEKVEKVNFAGLKDSPWYATKEKLGLKYTG SVLTFEIKG GKDEAWAFIDALKLHSNLANIGDVRSLVHPATTTSHSQSDEAGLARAGVTQSTVRLSVGI ETIDDIADLEGGFAAI
MetY-del	MGPKYDNSNADQWGFETR SIHAGQSVDAQTSARNLPIYQSTAFVFDSAEHAKQRFALE DLGPVYSRLTNPTVEALENRIASLEGGVHAVAFSSGQAATTNAILNLA
CAB0498630	MPTKYDNSNANKWGFETR SIHAGQSVSDTGARNLPIYLTSSYVFND AEHAANRFNLSD AGPVYSRLTNPTVA AVEERLANLEGGVHAVLFASGMAAETAAILNIARAGSHIVSSPRIYG GTETLFAVTLHAWASKPLS
Link to sequencing results of Naive and enriched library	https://drive.google.com/drive/folders/1t8xtPBNrBfUSPrZRAVRsKONzSCJ5YzMx?usp=sharing

Table S2. Vitamin mix

Vitamin mix	500 mL
p-aminobenzoic acid	0.5 gr
Niacin	0.5 gr
Pyridoxine	0.5 gr
Riboflavin	0.5 gr
Thiamin HCl	0.5 gr
Choline HCl	0.5 gr
d-Biotin	1 mg
Store at 4 °C in the dark after autoclaving	

Table S3. Trace element solution

Trace elements solution	250 mL (H ₂ O)
H ₃ BO ₃	100 mg
CuSO ₄ 5H ₂ O	100 mg
Iron chloride 4H ₂ O	200 mg
MnCl ₂	200 mg
NaMoO ₄ 2H ₂ O	200 mg
ZnSO ₄ 7H ₂ O	2 gr
Add HCl until the brown murky solution turns clear and yellow, filter sterile through 0.2 μM, keep in the dark.	

Table S4. Primers used in this study

Primer name	Sequence	Use
cTrc_Ins_Fw	ATAAGGAGGAATAAACCATGGGTCCG	Error-prone PCR and sequencing
cTrc_Ins_Rev	GGTACCAGCTGCAGATCTCGAGTTA	Error-prone PCR and sequencing
metY_F_SEQ	AAAGTTCGTGTGGGCCTGCT	Sequencing
metY_R_SEQ	AGCAGGCCACACGAACTTT	Sequencing
metX_F_SEQ	TGGTTCTGATCGAACACGCGC	Sequencing

Table S5. Sequence of MetYX CG construct

Gene	Sequence
<p>Wild-type MetXY gene of CG that was cloned and transformed to the E. coli</p>	<pre> > metYX_CG_OQ291222 gaattcTTTATTCTTGACACTAGTCGGCCAAAATGATATAATACCTGAGTTT AACTTTAAGAGAGGTATATATTACCATGGGTCCGAAGTACGACAACAGC AACGCGGATCAGTGGGGCTTCGAGACCCGTAGCATCCACGCGGGTCA GAGCGTTGACGCGCAAACAGCGCGCGTAACCTGCCGATTTACCAGA GCACCGCGTTTCGTGTTTGACAGCGCGGAGCACGCGAAACAACGTTTC GCGCTGGAAGATCTGGGCCCGGTTTATAGCCGTCTGACCAACCCGAC CGTGGAGGCGCTGGAAAACCGTATTGCGAGCCTGGAGGGTGGCGTT CATGCGGTGGCGTTTAGCAGCGGTCAGGCGGCGACCACCAACGCGAT CCTGAACCTGGCGGGTGGCGGTGACCACATTGTTACCAGCCCGCGTC TGTATGGTGGCACCGAAACCCTGTTCTGATCACCTGAACCGTCTGG GCATTGATGTTAGCTTTGTGGAGAACCCGGATGATCCGGAAAGCTGGC AGGCGGCGGTTCAACCGAACACCAAGGCGTTCTTTGGCGAGACCTTT GCGAACCCGCAAGCGGACGTGCTGGATATCCCGGCGGTTGCGGAAG TGCGCACCGTAACAGCGTTCGCTGATCATTGACAACACCATTGCGA CCGCGGCGCTGGTGCCTCCGCTGGAGCTGGGTGCGGATGTGGTTGT GGCGAGCCTGACCAAGTTCTACACCGGTAACGGCAGCGGTCTGGGTG GCGTTCTGATCGACGGTGGCAAATTTGATTGGACCGTGGAAAAGGAC GGCAAAGCGTTTTCCCGTATTTTGTACCCCGGATGCGGCGTACCAC GGTCTGAAGTATGCGGATCTGGGTGCGCCGGCGTTTGGTCTGAAAGT TCGTGTGGGCCTGCTGCGTGACACCGGTAGCACCTGAGCGCGTTTA ACGCGTGGGCGGCGGTTCAAGGCATCGATACCCTGAGCCTGCGTCTG GAGCGTCACAACGAAAACGCGATTAAGGTGGCGGAGTTCCTGAACAA CCACGAGAAGGTTGAAAAAGTGAACTTTGCGGGTCTGAAGGATAGCC CGTGGTACGCGACCAAGGAAAAACTGGGCCTGAAATATACCGGTAGC GTGCTGACCTTCGAGATCAAGGGTGGCAAAGACGAAGCGTGGGCGTT TATTGATGCGCTGAAACTGCACAGCAACCTGGCGAACATCGGCGACGT TCGTAGCCTGGTTGTGCATCCGGCGACCACCACCCATAGCCAAAGCG ATGAGGCGGGCCTGGCGCGTGCGGGTGTGACCCAAAGCACCGTTTCG TCTGAGCGTGGGTATCGAGACCATTGACGATATCATTGCGGACCTGGA AGGTGGCTTCGCGGCGATTTAAGGATCCAGAGGTATATATTAatgCCGAC CCTGGCGCCGAGCGGTCAGCTGGAGATCCAAGCGATTGGTGACGTTA GCACCGAGGCGGGCGCGATCATTACCAACGCGGAAATTGCGTACCAC CGTTGGGGTGAGTATCGTGTGGACAAAGAAGGCCGTAGCAACGTGGT TCTGATCGAACACGCGCTGACCGGTGATAGCAACGCGGCGGACTGGT GGGCGGATCTGCTGGGTCCGGGCAAGGCGATCAACACCGACATTTAC TGCGTTATCTGCACCAACGTGATCGGTGGCTGCAACGGCAGCACCGG TCCGGGCAGCATGCACCCGGATGGTAACTTCTGGGGCAACCGTTTTTC CGGCGACCAGCATTCTGACCAGGTTAACGCGGAGAAACAATTCCTG GATGCGCTGGGTATTACCACCGTTGCGGCGGTGCTGGGTGGCAGCAT GGGTGGCGCGCGTACCCTGGAGTGGGCGGCGATGTATCCGGAAACC GTTGGTGCGGCGGCGGTGCTGGCGGTTAGCGCGCGTGCAGCGCGT GGCAGATCGGCATTCAGAGCGCGCAAATCAAGGCGATTGAAAACGATC ACCACTGGCACGAGGGTAACTACTATGAAAGCGGCTGCAACCCGGCG ACCGGTCTGGGTGCGGCGCGTCTGATTGCGCACCTGACCTACCGTGG TGAGCTGGAAATCGACGAGCGTTTTTGGCACCAAGGCGCAGAAAAACG AAAACCCGCTGGGTCCGTATCGTAAGCCGGATCAACGTTTCGCGGTTG </pre>

	AGAGCTACCTGGACTATCAGGCGGATAAACTGGTTCAACGTTTTGACG CGGGTAGCTACGTGCTGCTGACCGATGCGCTGAACCGTCACGACATT GGCCGTGATCGTGGTGGCCTGAACAAGGCGCTGGAGAGCATTAAAGT GCCGTTCTGGTGGCGGGCGTTGACACCGATATCCTGTACCCGTATCA CCAGCAAGAACACCTGAGCCGTAACCTGGGTAACCTGCTGGCGATGG CGAAAATCGTTAGCCCGGTGGGTCATGATGCGTTCCTGACCGAAAGCC GTCAAATGGATCGTATTGTGCGTAACTTCTTTAGCCTGATCAGCCCGGA CGAGGATAACCCGAGCACCTACATTGAATTTTATATCTAACTCGAGCAA CCTGGAGGCGGGCGCAGGCCCGCCTTTTaagctt
--	--