Laboratory of mathematical methods and models
in bioinformatics
Institute for Information Transmission Problems
Russian Academy of Sciences

Transcription regulation of plastid genes *cysT* and *cysA* in Viridiplantae

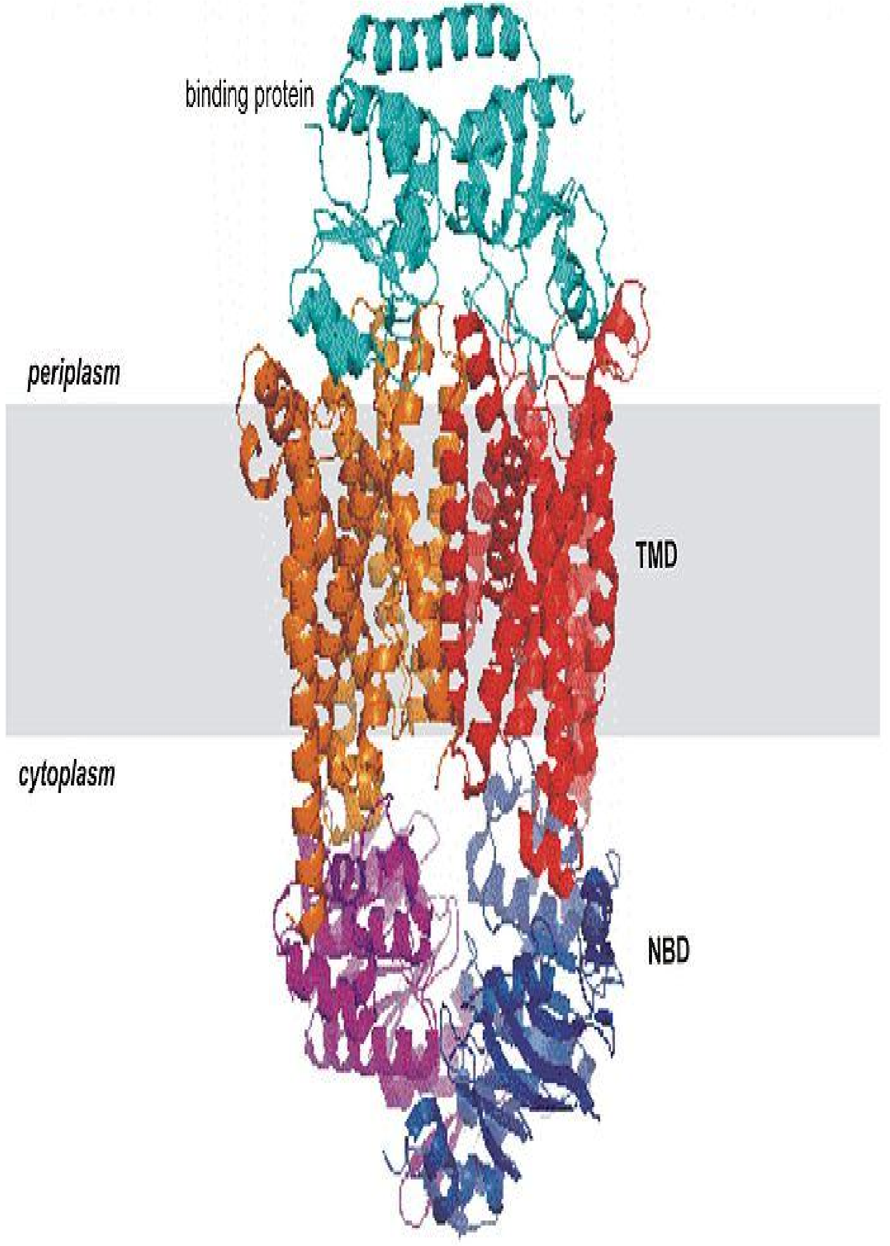
O. Zverkov, S. Korolev,
A. Seliverstov, V. Lyubetsky
http://lab6.iitp.ru/

Sulfate transport is realized by ABC transporter, in cyanobacteria it consists of 4 subunits. Among them:

CysT is a transmembrane protein.

Protein **CysA** contains ATP-binding domain and often contains TOBE domain, probably related to the **specificity to the sulfate**.

(Nuclear encoded protein CysP again determines specifity to the sulfate – out of discussion)



Firstly about *E. coli* regulation of sulfate transport and synthesis of cysteine:

DNA binds complex of **factor** CysB **with acetylserine** (ligand, sulfurless precursor of cysteine).

Regulation in ***E. coli*** takes place at the **transcription** level:

activation of **sulfate transport operon** ***cys****PT****W****A****M*.

Our <u>main result</u>: sulfate transport in plastids is regulated by a **factor** at the **transcription** level:

repression of **sulfate transport genes** ***cysT*** and ***cysA***.

We have studied expression regulation of subunits **CysT** and **CysA**, those are encoded in **plastids of Viridiplantae species**

(green algae, Anthocerotophyta, Marchantiophyta)

Anthoceros formosae and Marchantia polymorpha

Results:

1) **Most of the species have at least one suitable candidate** for **bacterial type promoter** in the 5'-leader regions of ***cysT*** or ***cysA*** genes. Mainly, in **Trebouxiophyceae** algae.

2) **No such promoters** were found in 5'-leader regions of both *cysT* and *cysA* in **green algae** *Nephroselmis olivacea, Pycnococcus provasolii, Bryopsis hypnoides, Leptosira terrestris* as well as in some Jungermanniopsida species; in the 5'-leader regions of *cysA* in green algae *Chlorella variabilis* and *Chlorokybus atmophyticus*; in the 5'-leader regions of *cysT* in green algae *Zygnema circumcarinatum*

cysA and *cysT* genes promoters in plastids:

Species	*cysA*	*cysT*
Nephroselmis olivacea	*(trnE)--cysA*	*rpl32--cysT*
Pycnococcus provasolii	no	*trnP--cysT*
Bryopsis hypnoides	*ccsA-!-cysA*	*rpl12--cysT*
Chlorella variabilis	*accD*-P-*cysA*	*rpl32*-P-*cysT*
Chlorella vulgaris	*accD--cysA*	*rpl32*-P-*cysT*
Coccomyxa subellipsoidea	*accD*-P-*cysA*	*rpl32*-P-*cysT*
Helicosporidium sp.	no	*ftsH*-P-*cysT*
Leptosira terrestris	*orf96--cysA*	*orf67--cysT*
Parachlorella kessleri	*accD*-P-P-*cysA*	*rpl32*-P-P-*cysT*
Mesostigma viride	*(trnE)*-P-*cysA*	*rpl32*-P-P-*cysT*
Chlorokybus atmophyticus	*(trnR)--cysA*	*(rpl32)*-P-*cysT*
Zygnema circumcarinatum	*(trnE)*-P-*cysA*	*trnV--cysT*
Anthoceros formosae	*trnE*-P-P-P-*cysA*	*rpl32*-P-*cysT*
Aneura mirabilis	*(trnE)--cysA* *	*rpl32--cysT* *
Marchantia polymorpha	*(trnE)*-P-*mbpX*	*rpl32*-P-*cysT*
Ptilidium pulcherrimum	*(trnE)--cysA* *	*rpl32--cysT* *

! - short intergenic region, * - pseudogen; Prasinophyceae, Ulvophyceae, Trebouxiophyceae (Chlorophyta) and Streptophyta

Helicosporidia group consisting of **parasites of invertebrates** is of particular interest.

***Helicosporidium* genus** belongs to **Trebouxiophyceae** class of **green algae**.

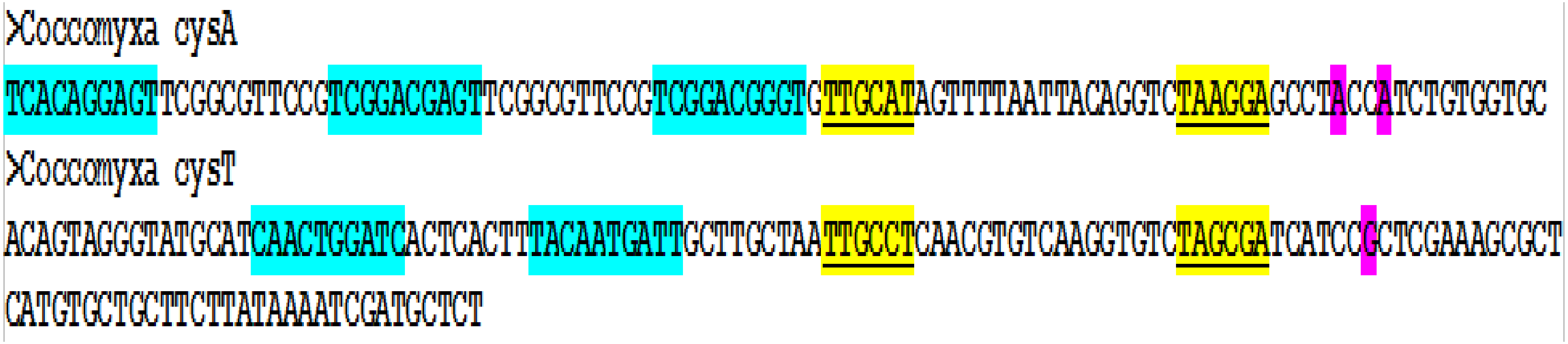
Helicosporidium sp. ex Simulium jonesii plastom is strongly reduced. Most genes of its plastom encode tRNA, rRNA, ribosomal proteins and bacterial type RNA polymerase subunits.
Subunit CysT of **sulfate ABC transporter** is **one of two proteins** which are not directly related to household

Result: **near each promoter** (except those in *Chlorokybus atmophyticus* and *Marchantia polymorpha*) **there is a site of conservative motif with consensus TAAWATGATT**, sometimes directly repeated, W={A,T}. This site is **for binding of a repression factor with DNA in the 5'-leader region of *cysT*** or ***cysA* in plastids**.

The frequencies of the nucleotides for each position in the motif multiplied by 28 are shown on the right. Values greater than 45% threshold are colored.

	A	G	C	T
1	7	3	5	**13**
2	**16**	2	3	7
3	**17**	2	3	6
4	9	5	4	10
5	**16**	0	0	12
6	2	9	3	**14**
7	0	**28**	0	0
8	**24**	1	2	1
9	4	3	1	**20**
10	6	1	3	**18**

Because: the site is often located near the -35 box of the promoter or overlaps that box. However a significant variability of distance between the motif and the promoter box allows us to assume that the motif is a **site of repressor binding**:



```
>Coccomyxa cysA
TCACAGGAGTTCGGCGTTCCGTCGGACGAGTTCGGCGTTCCGTCGGACGGGTGTTGCATAGTTTTAATTACAGGTCTAAGGAGCCTACCATCTGTGGTGC
>Coccomyxa cysT
ACAGTAGGGTATGCATCAACTGGATCACTCACTTTACAATGATTGCTTGCTAATTGCCTCAACGTGTCAAGGTGTCTAGCGATCATCCGCTCGAAAGCGCT
CATGTGCTGCTTCTTATAAAATCGATGCTCT
```

Binding sites are cyan, promoter boxes are yellow, transcription and initiation sites in the 5‘-leader region of *cysA* or *cysT* in *Coccomyxa* plastids are magenta.

Thus, **predicted regulation in plastids** in contrast **to *E.coli* regulation of sulfate transport is repression**

Are there other sulfate transport genes in plastids?
Orthologous genes to either ***cysA*** or ***cysT*** ?

Results: **clustering of proteins** encoded in plastids was performed, and it indicates that **there are no orthologous genes to *cysA*, *cysT*, *cysW*, *cysP* in other plastids**.

Cluster databases can be found at:

http://lab6.iitp.ru/ppc/chlorophyta/, http://lab6.iitp.ru/ppc/redline/

Proteins, annotated as CysA in two red algae (*Cyanidium caldarium*, *Cyanidioschyzon merolae*) and in *Cyanophora paradoxa*, are ATP-binding **subunits of Zn or Mn transporter**. Our type regulation is **not found in these species**.

Next conclusions

There is a **conservative motif** **with consensus TAAWATGATT**.

Our prediction of the transcription regulation confirms that ***Helicosporidium* sp. belongs to the class** Trebouxiophyceae. It emphasizes **importance of sulfate transport** **for the parasite** and can serve as a key for understanding roles of their plastids.

Absence of such regulation in *Leptosira terrestris* allows us to consider the **change of specificity** of the transporter to a substratum. The conjecture is also corroborated by **change of both proteins** CysT and CysA in this species

Other conclusions

Regulation absence in **green algae** *Nephroselmis olivacea*, *Pycnococcus provasolii*, *Bryopsis hypnoides*, and in *Marchantia polymorpha* (liverwort) can be associated with the **reducing of the corresponding proteins significance**, that is consistent with the full loss of these genes or with the pseudogenes emergence in **related species**.

We remind that two liverworts *Aneura mirabilis* и *Ptilidium pulcherrimum* have pseudogenes.

Difference between **enterobacteriaceae (*E. coli*) regulation** and **regulation in plastids** raises a question of the evolution of such regulation

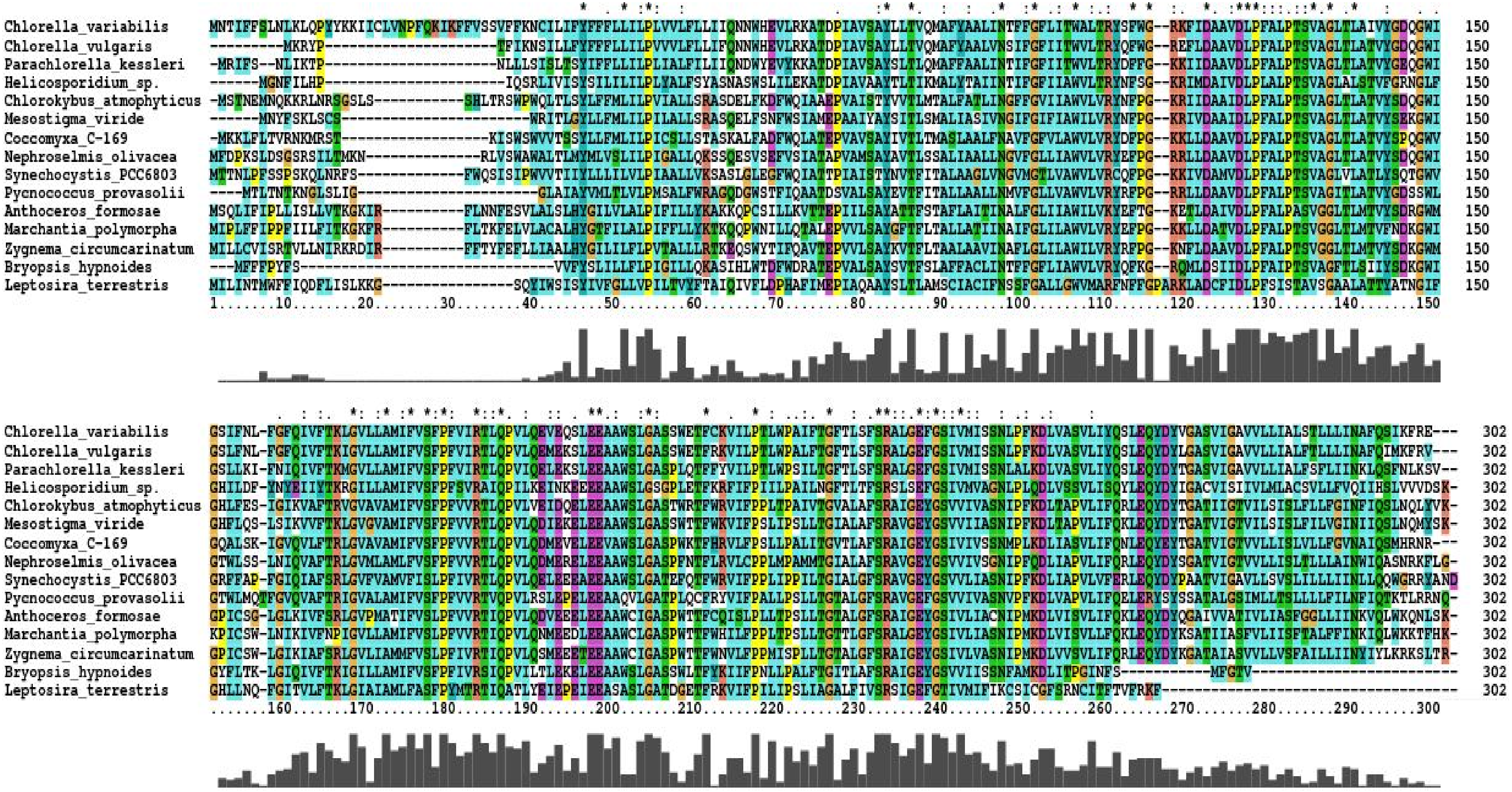
Other conclusions

In ***E. coli*** both proteins CysT and CysW consist of transmembrane domains, that are very close to each other. Their genes belong to sulfate transport **operon *cysPTWAM***.

But protein CysW is absent in plastids. Thus we **hypothesized** that in plastids **CysW is replaced by the second CysT copy**.

Preconditions: such situation is observed in other ABC transporters;
cysT and *cysA* **do not form an operon** in plastids, it allows us to consider possible **protein CysT abundance over protein CysA**. These are **concomitant** with the previous hypothesis

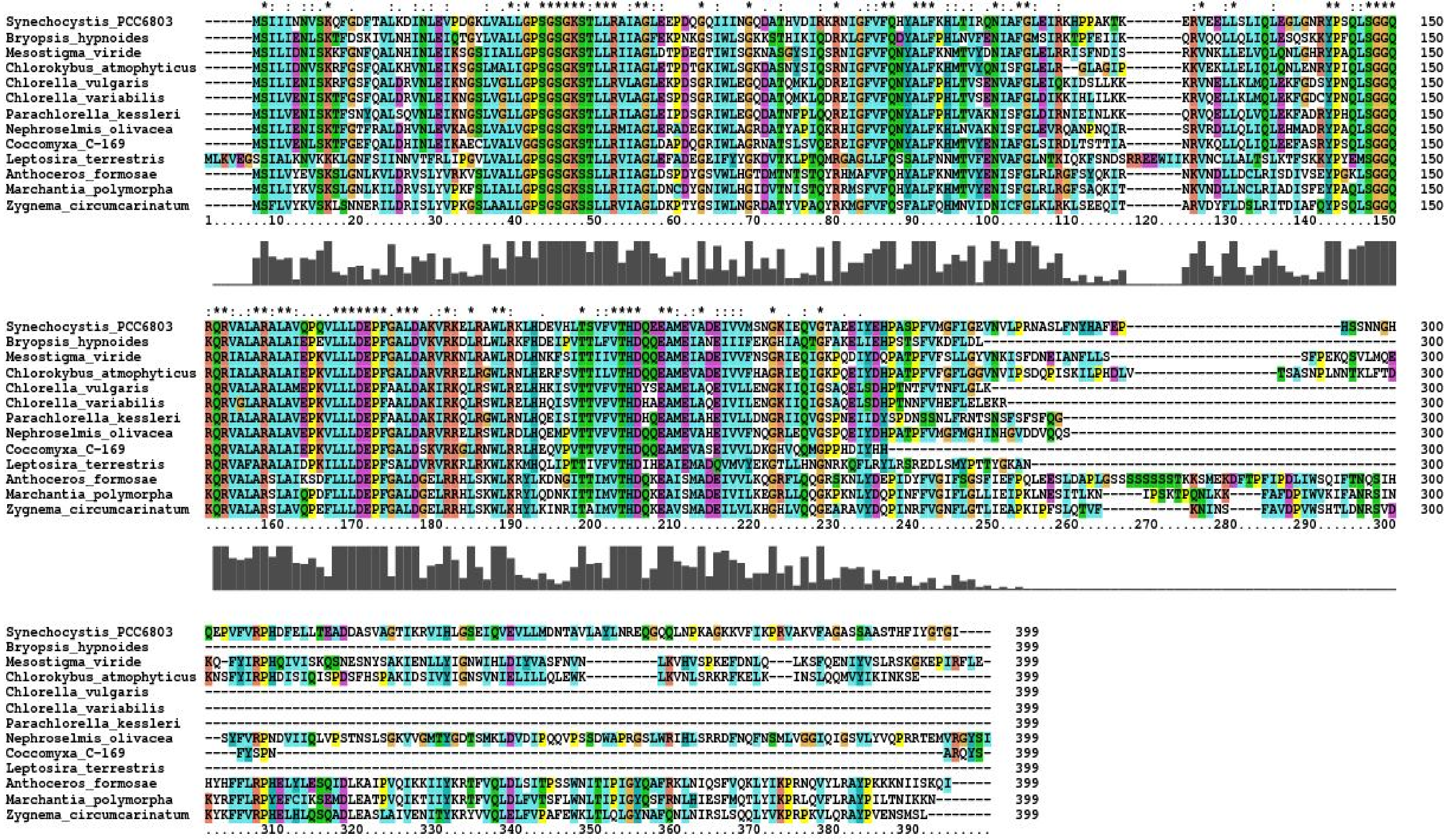
Multiple alignment of **orthologous** CysT proteins of *Synechocystis* sp. PCC 6803 cyanobacteria and in Viridiplantae plastids. Only *Bryopsis hypnoides* and *Leptosira terrestris* have shorter C-terminus:



```
Chlorella_variabilis      MNTIFFSLNLKLQPYYKKIICLVNPFQKIKFFVSSVFFKNCILIFYFFFLLILPLVVLFLLIIQNNWHEVLRKATDPIAVSAYLLTVQMAFYAALINTFFGFLITWALTRYSFWG--RKFIDAAVDLPFALPTSVAGLTLAIVYGDQGWI  150
Chlorella_vulgaris        ---------MKRYP---------------------TFIKNSILLFYFFFLLILPVVVLFLLIFQNNWHEVLRKATDPIAVSAYLLTVQMAFYAALVNSIFGFIITWVLTRYQFWG--REFLDAAVDLPFALPTSVAGLTLATVYGDQGWI  150
Parachlorella_kessleri    -MRIFS--NLIKTP---------------------NLLLSISLTSYIFFLLILPLIALFILIIQNDWYEVYKKATDPIAVSAYSLTLQMAFFAALINTIFGFIITWVLTRYDFFG--KKIIDAAVDLPFALPTSVAGLTLATVYGEQGWI  150
Helicosporidium_sp.       ------MGNFILHP----------------------IQSRLIVISYSILILILPLYALFSYASNASWSLILEKATDPIAVAAYTLTIKMALYTAIINTIFGFIIAWVLTRYNFSG--KRIMDAIVDLPLALPTSVAGLALSTVFGRNGLF  150
Chlorokybus_atmophyticus  -MSTNEMNQKKRLNRSGSLS-----------SHLTRSWPWQLTLSYLFFMLILPVIALLSRASDELFKDFWQIAAEPVAISTYVVTLMTALFATLINGFFGVIIAWVLVRYNFPG--KRIIDAAIDLPFALPTSVAGLTLATVYSDQGWI  150
Mesostigma_viride         ------MNYFSKLSCS-----------------------WRITLGYLLFMLILPILALLSRASQELFSNFWSIAMEPAAIYAYSITLSMALIASIVNGIFGIFIAWILVRYNFPG--KRIVDAAIDLPFALPTSVAGLTLATVYSEKGWI  150
Coccomyxa_C-169           -MKKLFLTVRNKMRST------------------KISWSWVVTSSYLLFMLILPICSLLSTASKALFADFWQLATEPVAVSAYIVTLTMASLAAFNAVFGFVLAWVLVRYDFPG--KKLLDAAVDLPFALPTSVAGLTLATVYSPQGWV  150
Nephroselmis_olivacea     MFDPKSLDSGSRSILTMKN--------------RLVSWAWALTLMYMLVSLILPIGALLQKSSQESVSEFVSIATAPVAMSAYAVTLSSALIAALLNGVFGLLIAWVLVRYEFPG--RRLLDAAVDLPFALPTSVAGLTLATVYSDQGWI  150
Synechocystis_PCC6803     MTTNLPFSSPSKQLNRFS-------------FWQSISIPWVVTIIYLLLILVLPIAALLVKSASLGLEGFWQIATTPIAISTYNVTFITALAAGLVNGVMGTLVAWVLVRCQFPG--KKIVDAMVDLPFALPTSVAGLVLATLYSQTGWV  150
Pycnococcus_provasolii    ----MTLTNTKNGLSLIG----------------------GLAIAYVMLTLVLPMSALFWRAGQDGWSTFIQAATDSVALSAYEVTFITALLAALLNMVFGLLVAWVLVRYRFPG--RRLLDAAVDLPFALPTSVAGITLATVYGDSSWL  150
Anthoceros_formosae       MSQLIFIPLLISLLVTKGKIR----------FLNNFESVLALSLHYGILVLALPIFILLYKAKKQPCSILLKVTTEPIILSAYATTFSTAFLAITINALFGLIIAWILVKYEFTG--KETLDAIVDLPFALPASVGGLTLMTVYSDRGWM  150
Marchantia_polymorpha     MIPLFFIPPFIILFITKGKFR----------FLTKFELVLACALHYGTFILALPIFFLLYKTKQQPWNILLQTALEPVVLSAYGFTFLTALLATIINAIFGLILAWVLVRYEFPG--KKLLDATVDLPFALPTSVGGLTLMTVFNDKGWI  150
Zygnema_circumcarinatum   MILLCVISRTVLLNIRKRDIR----------FFTYFEFLLIAALHYGILILFLPVTALLLRTKEQSWYTIFQAVTEPVVLSAYKVTFLTAALAAVINAFLGLILAWILVRYRFPG--KNFLDAAVDLPFALPTSVGGLTLMTVYSDKGWM  150
Bryopsis_hypnoides        ---MFFFPYFS-------------------------------VVFYSLILLFLPIGILLQKASIHLWTDFWDRATEPVALSAYSVTFSLAFFACLINTFFGFLIAWVLVRYQFKG--RQMLDSIIDLPFAIPTSVAGFTLSIIYSDKGWI  150
Leptosira_terrestris      MILINTMWFFIQDFLISLKKG----------------SQYIWSISYIVFGLLVPILTVYFTAIQIVFLDPHAFIMEPIAQAAYSLTLAMSCIACIFNSSFGALLGWVMARFNFFGPARKLADCFIDLPFSISTAVSGAALATTYATNGIF  150
                          1.......10........20........30........40........50........60........70........80........90.......100.......110.......120.......130.......140.......150

Chlorella_variabilis      GSIFNL-FGFQIVFTKLGVLLAMIFVSFPFVIRTLQPVLQEVEQSLEEAAWSLGASSWETFCKVILPTLWPAIFTGFTLSFSRALGEFGSIVMISSNLPFKDLVASVLIYQSLEQYDYVGASVIGAVVLLIALSTLLLINAFQSIKFRE---  302
Chlorella_vulgaris        GSLFNL-FGFQIVFTKIGVLLAMIFVSFPFVIRTLQPVLQEMEKSLEEAAWSLGASSWETFRKVILPTLWPALFTGFTLSFSRALGEFGSIVMISSNLPFKDLVASVLIYQSLEQYDYLGASVIGAVVLLIALFTLLLINAFQIMKFRV---  302
Parachlorella_kessleri    GSLLKI-FNIQIVFTKMGVLLAMIFVSFPFVIRTLQPVIQELEKSLEEAAWSLGASPLQTFFYVILPTLWPSILTGFTLSFSRALGEFGSIVMISSNLALKDLVASVLIYQSLEQYDYTGASVIGAVVLLIALFSFLIINKLQSFNLKSV--  302
Helicosporidium_sp.       GHILDF-YNYEIIYTKRGILLAMIFVSFPFSVRAIQPILKEINKEEEEAAWSLGSGPLETFKRFIFPIILPAILNGFTLTFSRSLSEFGSIVMVAGNLPLQDLVSSVLISQYLEQYDYIGACVISIIVLMLACSVLLFVQIIHSLVVVDSK-  302
Chlorokybus_atmophyticus  GHLFES-IGIKVAFTRVGVAVAMIFVSFPFVVRTLQPVLVEIDQELEEAAWSLGASTWRTFWRVIFPPLTPAIVTGVALAFSRAIGEYGSVVIVASNIPFKDLTAPVLIFQRLEQYDYTGATIIGTVILSISLFLLFGINFIQSLNQLYVK-  302
Mesostigma_viride         GHFLQS-LSIKVVFTKLGVGVAMIFVSFPFVVRTLQPVLQDIEKELEEAAWSLGASSWTTFWKVIFPSLIPSLLTGIALAFSRAVGEYGSVVIIASNIPFKDLTAPVLIFQKLEQYDYTGATVIGTVILSISLFILVGINIIQSLNQMYSK-  302
Coccomyxa_C-169           GQALSK-IGVQVLFTRLGVAVAMIFVSFPFVVRTLQPVLQDMEVELEEVAWSLGASPWKTFHRVLFPSLLPALITGVTLAFSRAIGEYGSIVIVSSNMPLKDLIASVLIFQNLEQYEYTGATVIGTVVLLISLVLLFGVNAIQSMHRNR---  302
Nephroselmis_olivacea     GTWLSS-LNIQVAFTRLGVMLAMLFVSFPFVVRTLQPVLQDMERELEEAAWSLGASPFNTFLRVLCPPLMPAMMTGIALAFSRAVGEYGSVVIVSGNIPFQDLIAPVLIFQRLEQYDYSGATVIGTVVLLISLTLLLAINWIQASNRKFLG-  302
Synechocystis_PCC6803     GRFFAP-FGIQIAFSRLGVFVAMVFISLPFIVRTLQPVLQELEEEAEEAAWSLGATEFQTFWRVIFPPLIPPILTGIALGFSRAVGEYGSVVLIASNIPFKDLIAPVLVFERLEQYDYPAATVIGAVLLSVSLILLLIINLLQQWGRRYAND  302
Pycnococcus_provasolii    GTWLMQTFGVQVAFTRIGVALAMIFVSFPFVVRTVQPVLRSLEPELEEAAQVLGATPLQCFRYVIFPALLPSLLTGTALGFSRAVGEFGSVVIVASNVPFKDLVAPVLIFQELERYSYSSATALGSIMLLTSLLLLFILNFIQTKTLRRNQ-  302
Anthoceros_formosae       GPICSG-LGLKIVFSRLGVPMATIFVSLPFVVRTIQPVLQDVEEELEEAAWCIGASPWTTFCQISLPLLTPSLLTGTALGFSRAIGEYGSIVLIACNIPMKDLVISVLIFQKLEQYDYQGAIVVATIVLIASFGGLLIINKVQLWKQNLSK-  302
Marchantia_polymorpha     KPICSW-LNIKIVFNPIGVLLAMIFVSLPFVVRTIQPVLQNMEEDLEEAAWCLGASPWTTFWHILFPPLTPSLLTGTTLGFSRALGEYGSIVLIASNIPMKDLVISVLLFQKLEQYDYKSATIIASFVLIISFTALFFINKIQLWKKTFHK-  302
Zygnema_circumcarinatum   GPICSW-LGIKIAFSRLGVLIAMMFVSLPFIVRTIQPVLQSMEEETEEAAWCIGASPWTTFWNVLFPPMISPLLTGTALGFSRAIGEYGSIVLVASNIPMKDLVVSVLIFQRLEQYDYKGATAIASVVLLVSFAILLIINYIYLKRKSLTR-  302
Bryopsis_hypnoides        GYFLTK-LGIQIVFTKIGILLAMIFVSFPFIVRSIQPVILTLEKELEEAAWSLGASSWLTFYKIIFPNLLPALFTGITLAFSRAIGEYGSVVIISSNFAMKDLITPGINFS-----------MFGTV-------------------------  302
Leptosira_terrestris      GHLLNQ-FGITVLFTKLGIAIAMLFASFPYMTRTIQATLYEIEPEIEEASASLGATDGETFRKVIFPILIPSLIAGALFIVSRSIGEFGTIVMIFIKCSICGFSRNCITFTVFRKF------------------------------------  302
                          .......160.......170.......180.......190.......200.......210.......220.......230.......240.......250.......260.......270.......280.......290.......300
```

Multiple alignment of **orthologous** CysA proteins of *Synechocystis* sp. PCC 6803 cyanobacteria and in Viridiplantae plastids. TOBE domain is often absent:



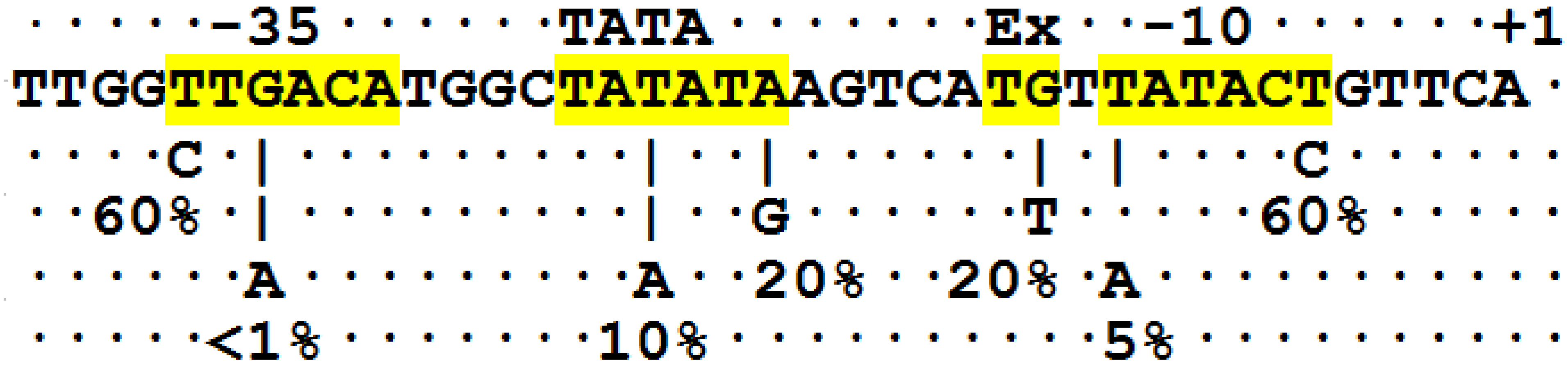
```
                                *: : ::.*  .    :. :.: :    : .*:* ******:*** :**::    * :   * . :    : *  .::**  *** ::.: .*:.**:  :                  :*  :*    :      ** ::****
Synechocystis_PCC6803    ------MSIIINNVSKQFGDFTALKDINLEVPDGKLVALLGPSGSGKSTLLRAIAGLEEPDQGQIIINGQDATHVDIRKRNIGFVFQHYALFKHLTIRQNIAFGLEIRKHPPAKTK-------ERVEELLSLIQLEGLGNRYPSQLSGGQ  150
Bryopsis_hypnoides       ------MSILIENLSKTFDSKIVLNHINLEIQTGYLVALLGPSGSGKSTLLRIIAGFEKPNKGSIWLSGKKSTHIKIQDRKLGFVFQDYALFPHLNVFENIAFGMSIRKTPFEIIK-------QRVQQLLQLIQLESQSKKYPFQLSGGQ  150
Mesostigma_viride        ------MSILIDNISKKFGNFQALNHINLEIKSGSIIALLGPSGSGKSTLLRIIAGLDTPDEGTIWISGKNASGYSIQSRNIGFVFQNYALFKNMTVYDNIAFGLELRRISFNDIS-------RKVNKLLELVQLQNLGHRYPAQLSGGQ  150
Chlorokybus_atmophyticus ------MSILIDNVSKRFGSFQALKHVNLEIKSGSLMALIGPSGSGKSTLLRIIAGLETPDTGKIWLSGKDASNYSIQSRNIGFVFQNYALFKHMTVYQNISFGLELR--GLAGIP-------KKVEKLLELIQLQNLENRYPIQLSGGQ  150
Chlorella_vulgaris       ------MSILIENISKRFGSFQALDRVNLEIKNGSLVGLLGPSGSGKSTLLRVLAGLEKPDSGRIWLEGQDATQMKLQDREIGFVFQNYALFPHLTVSENVAFGLEIQKIDSLLKK-------KRVNELLKLMQLEKFGDSYPNQLSGGQ  150
Chlorella_variabilis     ------MSILVENISKTFGSFQALDRVNLEIKNGSLVGLLGPSGSGKSTLLRVLAGLESPDSGRIWLEGQDATQMKLQDREIGFVFQNYALFPHLTVSENIAFGLDIKKIHLILKK-------KRVQELLKLMQLEKFGDCYPNQLSGGQ  150
Parachlorella_kessleri   ------MSILVENISKTFSNYQALSQVNLEIKNGSLVGLLGPSGSGKSTLLRILAGLEIPDSGRIWLEGQDATNFPLQQREIGFVFQNYALFPHLTVAKNISFGLDIRNIEINLKK-------QRVQELLQLVQLEKFADRYPHQLSGGQ  150
Nephroselmis_olivacea    ------MSILIENISKTFGTFRALDHVNLEVKAGSLVALVGPSGSGKSTLLRMIAGLERADEGKIWLAGRDATYAPIQKRHIGFVFQNYALFKHLNVAKNISFGLEVRQANPNQIR-------SRVRDLLQLIQLEHMADRYPAQLSGGQ  150
Coccomyxa_C-169          ------MSILVENLSKTFGEFQALDHINLEIKAECLVALVGGSGSGKSTLLRIIAGLDAPDQGRIWLAGRNATSLSVQEREIGFVFQNYALFKHMTVYENIAFGLSIRDLTSTTIA-------NRVKQLLQLIQLEEFASRYPSQLSGGQ  150
Leptosira_terrestris     MLKVEGSSIALKNVKKKLGNFSIINNVTFRLIPGVLVALLGPSGSGKSTLLRVIAGLEFADEGEIFYYGKDVTKLPTQMRGAGLLFQSSALFNNMTVFENVAFGLNTKIQKFSNDSRREEWIIKRVNCLLALTSLKTFSKKYPYEMSGGQ  150
Anthoceros_formosae      ------MSILVYEVSKSLGNLKVLDRVSLYVRKVSLVALLGPSGSGKSSLLRIIAGLDSPDYGSVWLHGTDMTNTSTQYRHMAFVFQHYALFKNMTVYENISFGLRLRGFSYQKIR-------NKVNDLLDCLRISDIVSEYPGKLSGGQ  150
Marchantia_polymorpha    ------MSILIYKVSKSLGNLKILDRVSLYVPKFSLIALLGPSGSGKSSLLRIIAGLDNCDYGNIWLHGIDVTNISTQYRRMSFVFQHYALFKHMTVYENISFGLRLRGFSAQKIT-------NKVNDLLNCLRIADISFEYPAQLSGGQ  150
Zygnema_circumcarinatum  ------MSFLVYKVSKLSNNERILDRISLYVPKGSLAALLGPSGSGKSSLLRVIAGLDKPTYGSIWLNGRDATYVPAQYRKMGFVFQSFALFQHMNVIDNICFGLKLRKLSEEQIT-------ARVDYFLDSLRITDIAFQYPSQLSGGQ  150
                         1.......10........20........30........40........50........60........70........80........90.......100.......110.......120.......130.......140.......150

                         :**:.:**:**:....*******.*** .:*: *  **:   .    * :::****  **:.:* ::::   * . : *  . .
Synechocystis_PCC6803    RQRVALARALAVQPQVLLLDEPFGALDAKVRKELRAWLRKLHDEVHLTSVFVTHDQEEAMEVADEIVVMSNGKIEQVGTAEEIYEHPASPFVMGFIGEVNVLPRNASLFNYHAFEP---------------------------HSSNNGH  300
Bryopsis_hypnoides       KQRVALARALAIEPEVLLLDEPFGALDVKVRKDLRLWLRKFHDEIPVTTLFVTHDQQEAMEIANEIIIFEKGHIAQTGFAKELIEHPSTSFVKDFLDL----------------------------------------------------  300
Mesostigma_viride        RQRIALARALAIEPKVLLLDEPFGALDARVRKNLRAWLRDLHNKFSITTIIVTHDQQEAMEIADEIVVFNSGRIEQIGKPQDIYDQPATPFVFSLLGYVNKISFDNEIANFLLS------------------------SFPEKQSVLMQE  300
Chlorokybus_atmophyticus RQRIALARALAIEPKVLLLDEPFGALDARVRRELRGWLRNLHERFSVTTILVTHDQQEAMEVADEIVVFHAGRIEQIGKPQEIYDHPATPFVFGFLGGVNVIPSDQPISKILPHDLV------------------TSASNPLNNTKLFTD  300
Chlorella_vulgaris       RQRVALARALAMEPKVLLLDEPFAALDAKIRKQLRSWLRELHHKISVTTVFVTHDYSEAMELAQEIVLLENGKIIQIGSAQELSDHPTNTFVTNFLGLK---------------------------------------------------  300
Chlorella_variabilis     RQRVGLARALAVEPKVLLLDEPFAALDAKIRKQLRSWLRELHHQISVTTVFVTHDHAEAMELAQEIVILENGKIIQIGSAQELSDHPTNNFVHEFLELEKR-------------------------------------------------  300
Parachlorella_kessleri   RQRIALARALAVEPKVLLLDEPFAALDAKIRKQLRGWLRNLHQEISITTVFVTHDHQEAMELAHEIVLLDNGRIIQVGSPNEIIDYSPDNSSNLFRNTSNSFSFSFQG------------------------------------------  300
Nephroselmis_olivacea    RQRVALARALAVEPKVLLLDEPFGALDARVRRELRSWLRDLHQEMPVTTVFVTHDQQEAMEVAHEIVVFNQGRLEQVGSPQEIYDHPATPFVMGFMGHINHGVDDVQQS-----------------------------------------  300
Coccomyxa_C-169          RQRVALARALAIEPKVLLLDEPFGALDSKVRKGLRNWLRRLHEQVPVTTVFVTHDQQEAMEVASEIVVLDKGHVQQMGPPHDIYHH----------------------------------------------------------------  300
Leptosira_terrestris     RQRVAFARALAIDPKILLLDEPFSALDVRVRKRLRKWLKKMHQLIPTTIVFVTHDIHEAIEMADQVMVYEKGTLLHNGNRKQFLRYLRSREDLSMYPTTYGKAN----------------------------------------------  300
Anthoceros_formosae      KQRVALARSLAIKSDFLLLDEPFGALDGELRRHLSKWLKRYLKDNGITTIMVTHDQKEAISMADEIVVLKQGRFLQQGRSKNLYDEPIDYFVGIFSGSFIEFPQLEESLDAPLGSSSSSSSSTKKSMEKDFTPFIPDLIWSQIFTNQSIH  300
Marchantia_polymorpha    KQRVALARSLAIQPDFLLLDEPFGALDGELRRHLSKWLKRYLQDNKITTIMVTHDQKEAISMADEIVILKEGRLLQQGKPKNLYDQPINFFVGIFLGLLIEIPKLNESITLKN-----IPSKTPQNLKK----FAFDPIWVKIFANRSIN  300
Zygnema_circumcarinatum  KQRVALARSLAVQPEFLLLDEPFGALDGELRRHLSKWLKHYLKINRITAIMVTHDQKEAVSMADEILVLKHGHLVQQGEARAVYDQPINRFVGNFLGTLIEAPKIPFSLQTVF-----------KNINS----FAVDPVWSHTLDNRSVD  300
                         .......160.......170.......180.......190.......200.......210.......220.......230.......240.......250.......260.......270.......280.......290.......300

Synechocystis_PCC6803    QEPVFVRPHDFELLTEADDASVAGTIKRVIHLGSEIQVEVLLMDNTAVLAYLNREQGQQLNPKAGKKVFIKPRVAKVFAGASSAASTHFIYGTGI----  399
Bryopsis_hypnoides       ---------------------------------------------------------------------------------------------------  399
Mesostigma_viride        KQ-FYIRPHQIVISKQSNESNYSAKIENLLYIGNWIHLDIYVASFNVN---------LKVHVSPKEFDNLQ---LKSFQENIYVSLRSKGKEPIRFLE-  399
Chlorokybus_atmophyticus KNSFYIRPHDISIQISPDSFHSPAKIDSIVYIGNSVNIELILLQLEWK---------LKVNLSRKRFKELK---INSLQQMVYIKINKSE---------  399
Chlorella_vulgaris       ---------------------------------------------------------------------------------------------------  399
Chlorella_variabilis     ---------------------------------------------------------------------------------------------------  399
Parachlorella_kessleri   ---------------------------------------------------------------------------------------------------  399
Nephroselmis_olivacea    --SYFVRPNDVIIQLVPSTNSLSGKVVGMTYGDTSMKLDVDIPQQVPSSDWAPRGSLWRIHLSRRDFNQFNSMLVGGIQIGSVLYVQPRRTEMVRGYSI  399
Coccomyxa_C-169          ----FYSPN------------------------------------------------------------------------------------ARQYS-  399
Leptosira_terrestris     ---------------------------------------------------------------------------------------------------  399
Anthoceros_formosae      HYHFFLRPHELYLESQIDLKAIPVQIKKIIYKRTFVQLDLSITPSSWNITIPIGYQAFRKLNIQSFVQKLYIKPRNQVYLRAYPKKKNIISKQI-----  399
Marchantia_polymorpha    KYRFFLRPYEFCIKSEMDLEATPVQIKTIIYKRTFVQLDLFVTSFLWNLTIPIGYQSFRNLHIESFMQTLYIKPRLQVFLRAYPILTNIKKN-------  399
Zygnema_circumcarinatum  KYKFFVRPHELHLQSQADLEASLAIVENITYKRYVVQLELFVPAFEWKLTLQLGYNAFQNLNIRSLSQQLYVKPRPKVLQRAYPVENSMSL--------  399
                         .......310.......320.......330.......340.......350.......360.......370.......380.......390.........
```

Thank you

Bacterial type promoters were detected by the data of the mutations in *psbA*-77 promoter of the *Sinapis alba* mustard plastids



```
·····-35······TATA·······Ex··-10······+1
TTGGTTGACATGGCTATATAAGTCATGTTATACTGTTCA·
····C·|·········|··|······|·|····C······
··60%·|·········|··G······T·····60%·····
······A·········A··20%··20%·A···········
·····<1%·······10%··········5%··········
```

Most species have at least one suitable candidate for bacterial type promoters in the 5'-leader regions of these genes. No promoters were found in 5'-leader regions of both *cysT* and *cysA* in green algae *Nephroselmis olivacea, Pycnococcus provasolii, Bryopsis hypnoides, Leptosira terrestris* as well as in some Jungermanniopsida species; in the 5'-leader regions of *cysA* in green algae *Chlorella variabilis* and *Chlorokybus atmophyticus*; in the 5'-leader regions of *cysT* in green algae *Zygnema circumcarinatum*.