

***Synechococcus elongatus* UTEX 2973, a fast growing  
cyanobacterial chassis for biosynthesis using light and CO<sub>2</sub>**

Jingjie Yu<sup>1</sup>, Michelle Liberton<sup>1</sup>, Paul F. Cliften<sup>2</sup>, Richard D. Head<sup>2</sup>, Jon M.  
Jacobs<sup>3</sup>, Richard D. Smith<sup>3</sup>, David W. Koppenaal<sup>3</sup>, Jerry J. Brand<sup>4</sup> & Himadri B.  
Pakrasi<sup>1,\*</sup>

<sup>1</sup>Department of Biology, Washington University, St. Louis, MO 63130

<sup>2</sup>Genome Technology Access Center, Washington University School of Medicine, St.  
Louis, MO 63110

<sup>3</sup>Pacific Northwest National Laboratory, Richland, WA 99352

<sup>4</sup>UTEX The Culture Collection of Algae, University of Texas, Austin, TX 78712

\*Correspondence and request for materials should be addressed to H.B.P.

Phone – 314-935-6853; Fax – 314-935-8818; Email - pakrasi@wustl.edu

**Table S1.** ORFs in the *Synechococcus* 7942 genome that are missing in the *Synechococcus* 2973 genome

Locus tag (Synpcc7942)	Nucleotide location	protein	Notes
0767	759413-759847	Hypothetical protein	
0768	760522-761178	Hypothetical protein	confirmed by proteomics data (19 unique peptides)
0769	761226-762128	Hypothetical protein	confirmed by proteomics data (5 unique peptides)
0770	762159-763643	Hypothetical protein	confirmed by proteomics data (25 unique peptides)
0771	763640-764854	Hypothetical protein	confirmed by proteomics data (5 unique peptides)
0772	764984-765577	Chromophore Lyase CpcT/CpeT	confirmed by proteomics data (4 unique peptides)

**Table S2.** Peptide sequence comparisons of genes on the edges of the region of inversion between UTEX 2973 and PCC 7942. Dots represent the amino acids that are identical to the corresponding ones in the middle line.

Edge 1 peptide alignment

Upper line: UTEX 2973 M744\_08310

Middle line: Amino acid sequence

Bottom line: PCC 7942 synpcc7942\_1463

```

      10      20      30      40      50      60      70      80      90
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
MKNLFKVMLAGPLAASVAFAGAAQAESLNRDITNQVNAIANEELPAAQVTSITQLSDVKPTDWAYQALQSLVERYGCIVGYPDRTYRGSR
----MH.PA.LWMGTGLWLGL.PATIAFDAS..R.....D.....R.....
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      100     110     120     130     140     150     160     170     180
ALTRYEFAAGLNACLDKVIEFAASKEDLDTLKKLMEEFQAEELATLRGRVDSLEARVTELEATQFSTTTKLQGEVILSLDRAFGPSGLDTG
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      190     200     210     220     230     240     250     260     270
TSFSQRVSLNLTSTFTGKDLLKTRLRNSITSPLRDFNGGVPGSGIISPAALDYDNATNAPNANFFLDTLLYTFPVGDVKEFTVGTSNL
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      280     290     300     310     320     330     340     350     360
QVEDILSTGTFYAPYISYFFSTPIPGIYGDADTTNSAGAGFNWQLSKNFNFGAAYINQNGPTTGGNLASDPTSGVFGADSQTTAQLAFK
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      370     380     390     400     410     420     430     440     450
NDSGTFIGALAYAYRKGPDLATDLSFLPSQFGGVNFGTFPRALAGLGLGSSGSDIVSTNNIGLSLGWAVSDGFTISGSYGISFNGFGANN
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      460     470     480     490     500     510     520     530     540
TVQSWTVGFTFPNLFADGNEFGFAVQAPYVISDSRGAAFQDSGNFAFEVYVKLQVTDNISITPALYAITNAGGGLSFPDSANGDFIVPV
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
.....|..
IKTVFTF
.....

```

Edge 2 peptide alignment

Upper line: UTEX 2973 M744\_09255

Middle line: Amino acid sequence

Bottom line: PCC 7942 synpcc7942\_1635

```

      10      20      30      40      50      60      70      80      90
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
MS--MHVPAALWMTGLWLGLPATIAFDASTIRQVNAIANEDLPAAQVTSITQLSDVRPTDWAYQALQSLVERYGCIVGYPDRTYRGSR
.KNLFK.ML.GPLAASVAFAG.AQAESLNRD..N.....E.....K.....
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      100     110     120     130     140     150     160     170     180
ALTRYEFAAGLNACLDKVIEFAASKEDLDTLKKLTEEFQSEELATLRGRVDGLEARVTELEATQFSTTTKLQGDVLFSLDRAFGPSGLDTG
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      190     200     210     220     230     240     250     260     270

```

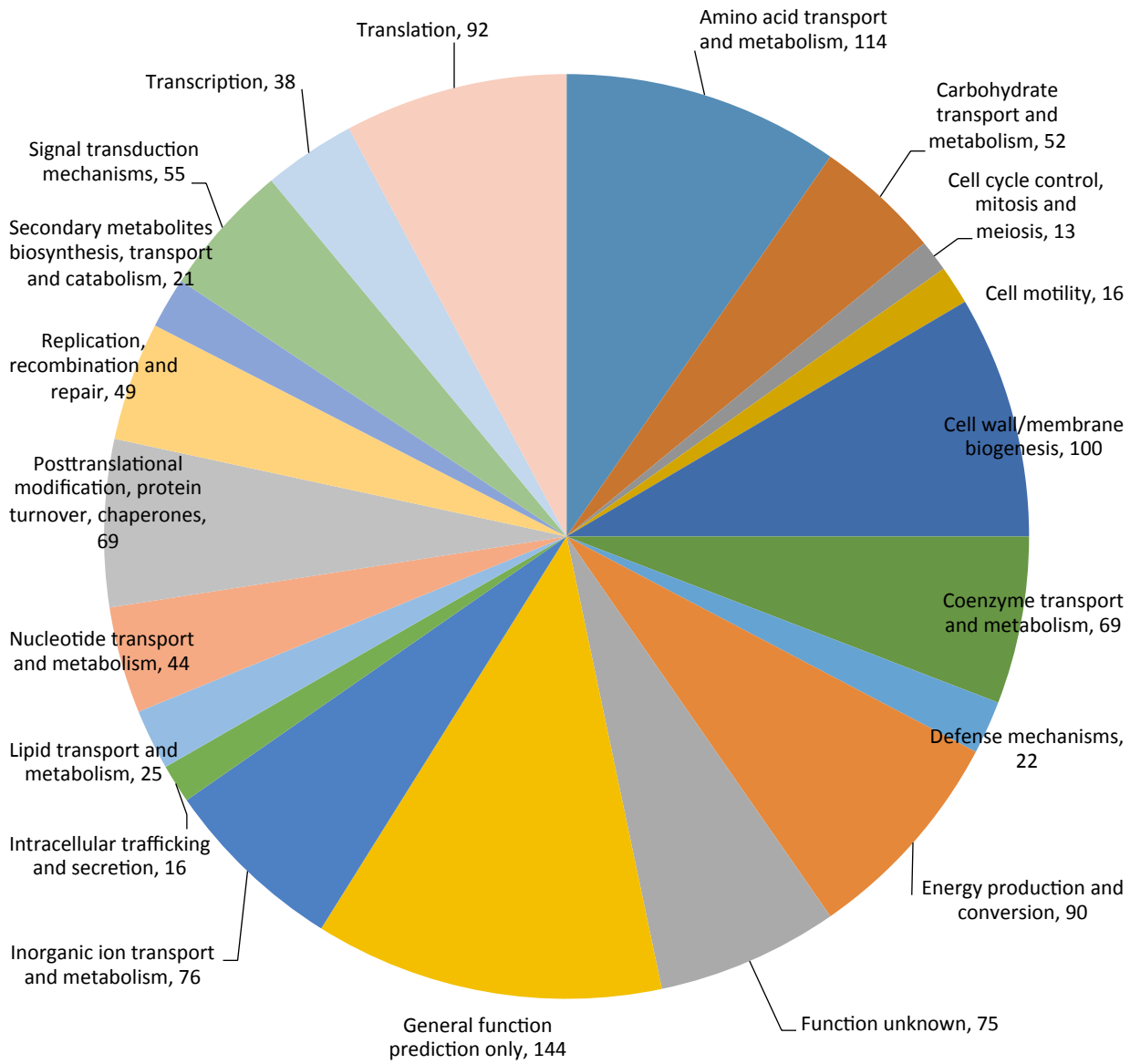
```

.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
TSFSQRVGLDFNTSFTGKDLLKTRLETNSITSPLRDLNGFAAGAGIPSPGAALAYDNAATGADPNFTLNTLIYQFPVGSVNFTVGTSNV
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
280      290      300      310      320      330      340      350      360
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
QVDDVFSTNGSFYAPELSYFFNNPVPGIYNDADTSNSAGAGFNWQINPNFNLGLAYINQQGPTSGQNLEADPTAGVFGADSQTTAQLAFK
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
370      380      390      400      410      420      430      440      450
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
NDDGTFICALAYAYRKGPAFDPVNPDGVGLPAQFGGVSYGTVRSLIGPSNSPTDLISSNNIGLSLGWAVSENF TISGSYGISFLSGSAGS
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
460      470      480      490      500      510      520      530      540
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
STVQSWNIGLTFPNLFADGNELGLAIGQIPYVTSDSRGAFAFADSGPFAFEVYKQLTDNIAITPALYAVTNAGGGLTLGNLANGDYWVP
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

.....|...
VLKTEFLF
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

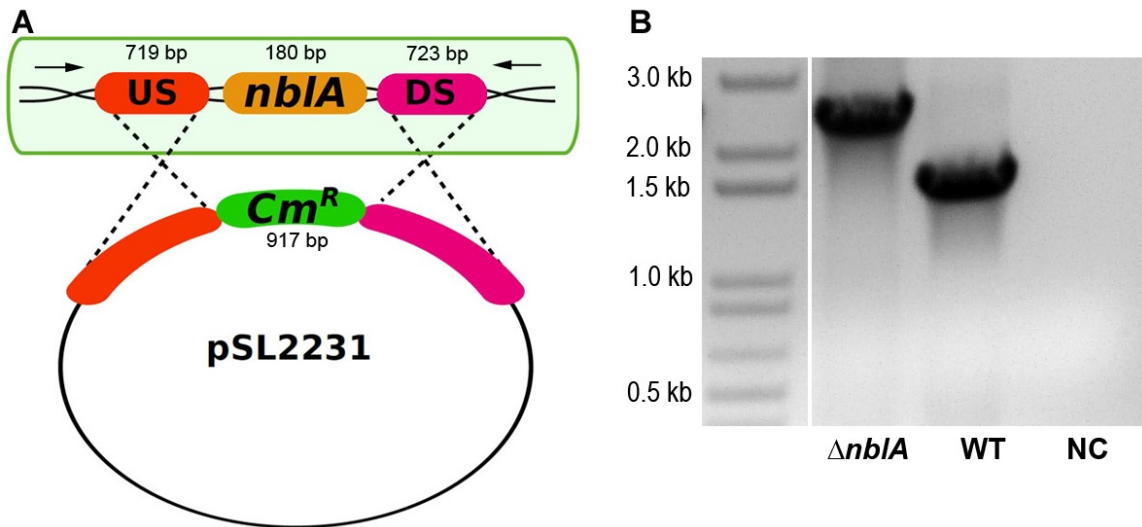
```

**Figure S1.**



**Figure S1. Categories of functional proteins of UTEX 2973 detected in global protein mass spectra.**

Figure S2.



**Figure S2. Verification of gene replacement in the *Synechococcus* UTEX 2973  $\Delta nblA$  mutant strain.** **A.** Expected double reciprocal homologous recombination. US: upstream sequence; DS: downstream sequence;  $Cm^R$ : chloramphenicol resistance gene cassette; pSL2231: a plasmid derivative of pBR322. Sizes of all fragments are indicated. Arrows represent primers used to verify recombination results. **B.** PCR products from  $\Delta nblA$  mutant and WT cells using primers shown in (A). The expected size of the PCR product from WT cells is 1622 bp; while the expected size of the PCR product from the  $\Delta nblA$  mutant with the coding region of *nblA* replaced by  $Cm^R$  gene is 2359 bp. Absence of the 1.6 kb WT band in  $\Delta nblA$  lane indicated that the mutant strain had completely segregated. NC: negative control (without template).