Development of genetic tools from a novel cyanobacterium *Synechococcus elongatus* PCC 11801 using transcriptome data

Swati Madhu[1], Annesha Sengupta[1], Bratati Mukherjee[1,2], Pramod P. Wangikar[1,2,3*]

[1]Department of Chemical Engineering, [2]DBT-Pan IIT Centre for Bioenergy, [3]Wadhwani Research Centre for Bioengineering, Indian Institute of Technology, Bombay, Mumbai-400076, India

[*]Corresponding author. Email: pramodwangikar@gmail.com

Background

Cyanobacteria holds great potential for the production of biofuel, other essential platform chemicals and feedstocks[1]. However, like other photosynthetic microorganisms, cyanobacterial productivity is often lagging their heterotrophic counterparts like *E. coli* because of the myriad reasons, prominent among which is slow growth rate. To overcome this drawback, there is a recent surge in engineering fast growing strains such as *Synechococcus elongatus* strains, UTEX 2973, PCC 11801, PCC 11802 and *Synechococcus* sp. PCC 11901 which has comparable growth rate with yeast [2], [3], [4], [5]). Metabolic engineering of these fast-growing strains might be a break-through in the developing cyanobacteria-derived compounds.

Synechococcus elongatus PCC 11801 has been recently reported as a fast growing, robust cyanobacterial strain which exhibits high tolerance towards toxic solvents such as butanol, 2,3BD, ethanol, pentanol, etc.[3] Its ability to tolerate high solvent stresses has been further augmented through adaptive evolution showing tolerance and cross-tolerance levels higher than other cyanobacteria. Besides its improved physiological characteristics, rational engineering of PCC 11801 have shown significantly higher yield of succinate (0.9 g/L) [7]. Demonstrating the potential of this strain as a candidate for photosynthetic cell factory. Additionally, the implementation of the marker less genome editing technique, CRISPR-Cpf1 on this strain for the co-production of ethylene and succinate with efficient carbon partitioning[8] is a step forward towards developing this fast growing PCC 11801 strain as a production host grown under outdoor conditions.

A photosynthetic cell factory necessitates the characterization of genetic toolbox as it ensures probable control of gene expression. In addition, transcriptomic study under different environmental growth condition will help to accelerate strain improvement

through systemic metabolic engineering[9]. Therefore, in this study, we are demonstrating our approach of promoter selection based on the expression levels of genes obtained from transcriptomic analysis of the strain grown under different growth conditions.

Methods

Growth conditions of *Synechococcus elongatus* PCC 11801

S. elongatus PCC 11801 was inoculated at an initial OD_{730} of 0.1 in BG-11 medium under culture conditions (control-38 °C, 185 µmol photons/m^2/s constant illumination). To study the transcriptomic responses, it was grown to an exponential OD_{730} under three different conditions: (i) high carbon (HC), 1% CO_2; (ii) high temperature (HT), 43°C; and (iii) high salt (HS), 400mM NaCl;

For transcriptomic sequencing, 50 mL of cultures with or without treatments was collected and after pelleting down preserved in RNAlater in -80°C and sent to MOgene, (St. Louis, MO) for next generation sequencing.

Read mapping, differential expression and enrichment analysis

The trimmed reads in the form of FASTQ files received from the MOgene were processed and aligned to the reference genome CP030139.1 and further expression and DSEQ2 analysis were done using the commercial software platform Geneious. Genes showing $\log_2$fold change ≥ 1 and p value≤ 0.05 were considered differentially expressed genes (DEGs) for which non-hierarchical clustering was performed in R[10]. Gene enrichment was implemented for each cluster using BiNGO tool in Cytoscape[11].

Characterization of promoter

According to transcript expression data the genes has been categorized into three categories which defines the promoter's strength: Switch, Tunable and Constitutive promoters. In order to corroborate this approach some native well reported promoters (P_{cpcB}, P_{rbcL}, P_{psbaI}, P_{psbaII}, $P_{psbaIII}$, P_{CCM326}, P_{CMP500}, P_{CMP893})[12] have been successfully chosen from different categories of transcript data and characterized based on the relative fluorescence of eYFP under four different growth conditions which can form a baseline for the new promoter characterization.

Results

Global transcriptomic analysis and Promoter selection

A total of 19,629,885 to 37,814,467 sequence reads were obtained in all the four abiotic conditions. Considering the criteria of both p-value ≤ 0.01 and $\log_2$ fold change (log|FC|) ≥ 1, among 802 DEGs 452 genes were upregulated and 350 genes were downregulated in HC condition (Figure1). Similarly, in HT conditions 138 genes were upregulated and 124 genes were downregulated and in HS conditions 97 upregulated and 135 downregulated genes were present. Since, production of most platform chemicals is favored in HC condition and it has also shown more DEGs as compared to others, so, we would like to explain our approach considering this condition.

Figure 1 shows the distribution of TPM in all the four different conditions and DEGs in HC which imparts the possibility to select promoters from different strength categories.

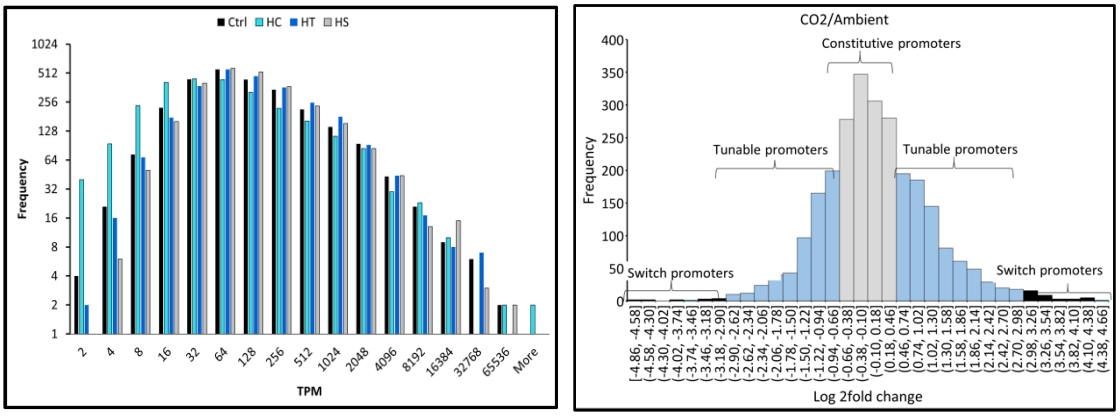


Figure 1: a. Distribution of transcript count of all genes across four conditions b. Distribution of differentially expressed genes in high carbon (HC) condition.

In order to group genes showing similar level of expression into a single cluster *k*-means clustering was implemented across the conditions for differentially expressed genes (Figure 2). Then GO enrichment analysis was performed which showed that the genes upregulated under HC condition were mainly enriched in molecular function of sulfur metabolism and downregulated for bicarbonate transporters. Sulfur metabolism is necessary for photosynthetic organisms for coenzyme biosynthesis, protein formation and for sulfolipid biosynthesis which is predominantly present in photosynthetic membranes[13]. Similarly, GO enrichment in HS condition showed genes upregulated are

enriched in starch and sucrose metabolism. Sucrose is compatible solute to tolerate high salt condition and cyanobacteria synthesize them to maintain osmotic balance and avoid membrane desiccation.

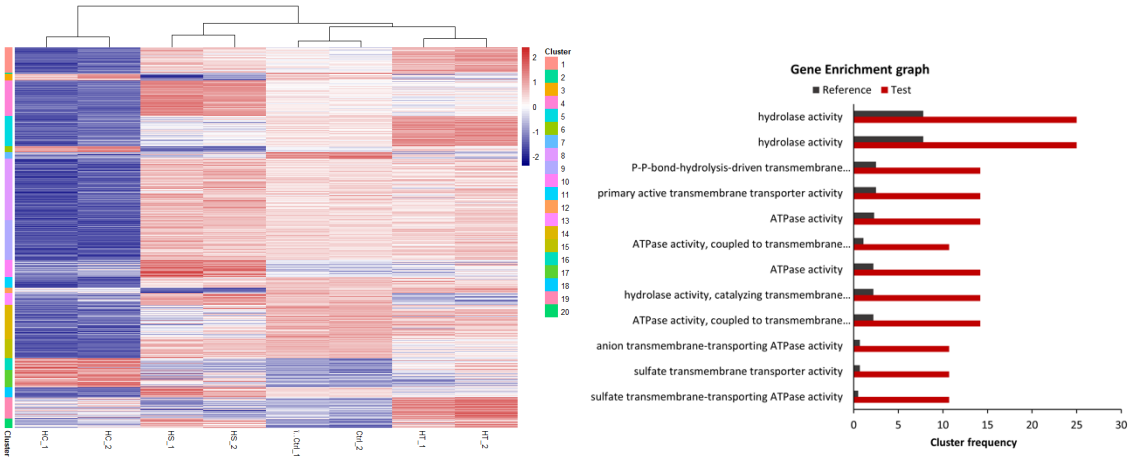


Figure 2: a. Non-hierarchical clustering**.** Genes in the data set were subjected to average-linkage, non-hierarchical clustering using a Euclidean distance metric. b. GO enrichment analysis of DEGs of cluster16.

Characterization of promoters

By analyzing fluorescence intensities, the strongest expression was shown by P_{cpcB} in low light with 1% CO_2. Additionally P_{cpcB} and P_{rbcl} has also shown higher expression than the respective promoter in PCC 7942 [14] (Figure 3). Considering the transcript count of analyzed promoters *cpcB* gene shows log_2FC -2.04 which makes this a tunable promoter (HC repressor) and it resonates with the eYFP fluorescence data. Similarly, the insignificant repression in transcript levels observed for the rbcL gene was validated by the eYFP fluorescence, making P_{rbcl} a strong constitutive promoter (Figure 3). P_{psbAI} and $P_{psbAIII}$ falls into the category of tunable promoter having log_2 FC 1.6 (HC activator) and 0.6 (HC repressor) under HC condition respectively. Moreover, the low transcript count of genes of P_{psbAII} , P_{CCM326}, P_{CMP500}, P_{CMP893} is reflected in the low eYFP fluorescence.

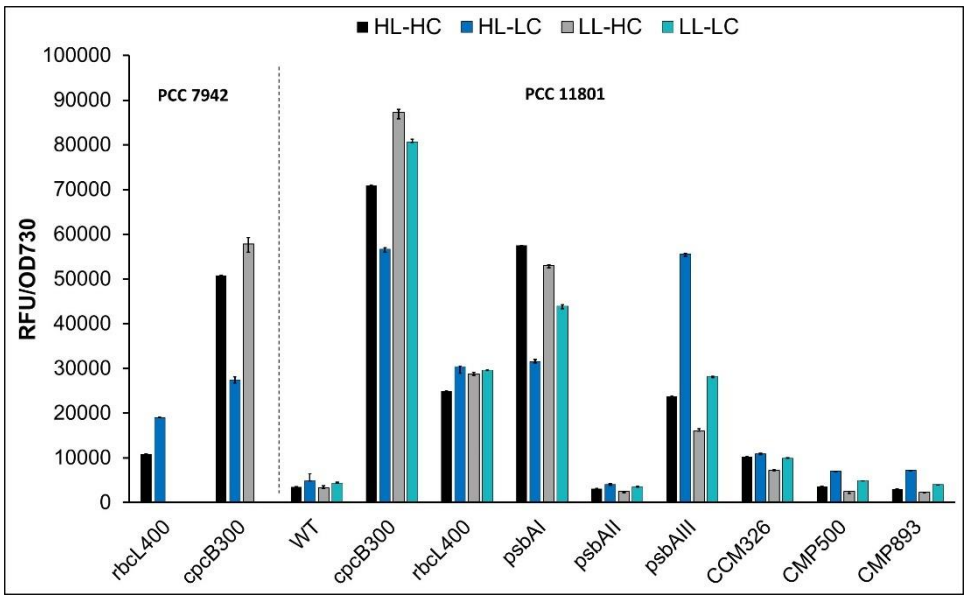


Figure 3: Comparison of native promoter activities in *S. elongatus* PCC 11801. Activities were measured as eYFP fluorescence per abs_{730} for cell cultures in BG11medium.

Conclusion

In this study we successfully demonstrate the methodology adopted for the selection of novel promoters exhibiting different strength under different conditions based on its normalized mRNA expression value is affirmative. So far, P_{cpcB} has shown the highest activity among all the native promoters and discovered P_{rbcl} as a constitutive promoter unlike other model cyanobacteria. We are working towards increasing the repertoire of promoters from this strain. Therefore, this study also opens the avenues to select and characterize more genetic tools like RBSs and neutral sites for this cyanobacterium and will help establish this non model cyanobacterium as a green chassis.

Funding The authors acknowledge research grant from Department of Biotechnology (DBT), Government of India, awarded to P.P.W. towards DBT-PAN IIT Centre for Bioenergy (Grant No: BT/EB/PAN IIT/2012).

References

1. Knoot, C. J., Ungerer, J., Wangikar, P. P. & Pakrasi, H. B. Cyanobacteria: Promising biocatalysts for sustainable chemical production. *J. Biol. Chem.* **293**, 5044–5052 (2018).
2. Yu, J. *et al. Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO_2. *Sci. Rep.* **5**, 8132 (2015).

3. Jaiswal, D. *et al.* Genome Features and Biochemical Characteristics of a Robust, Fast Growing and Naturally Transformable Cyanobacterium *Synechococcus elongatus* PCC 11801 Isolated from India. *Sci. Rep.* **8**, 1–13 (2018).
4. Jaiswal, D. *et al.* OPEN A Novel Cyanobacterium *Synechococcus elongatus* PCC 11802 has Distinct Genomic and Metabolomic Characteristics Compared to its Neighbor PCC. 1–15 (2020) doi:10.1038/s41598-019-57051-0.
5. Włodarczyk, A., Selão, T. T., Norling, B. & Nixon, P. J. Newly discovered Synechococcus sp. PCC 11901 is a robust cyanobacterial strain for high biomass production. *Commun. Biol.* **3**, (2020).
6. Artur Wlodarczyk, Tiago Toscano Selao, Birgitta Norling, P. J. N. Unprecedented biomass and fatty acid production by the newly discovered cyanobacterium Synechococcus sp. PCC 11901. *BioRxiv* (2019).
7. Sengupta, S. *et al.* Metabolic engineering of a fast-growing cyanobacterium *Synechococcus elongatus* PCC 11801 for photoautotrophic production of succinic acid. *Biotechnol. Biofuels* **13**, 1–18 (2020).
8. Sengupta, A. *et al.* Photosynthetic Co-Production of Succinate and Ethylene in A Fast-Growing Cyanobacterium, *Synechococcus elongatus* PCC 11801. *Metabolites* **10**, 250 (2020).
9. Tan, X. *et al.* The primary transcriptome of the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Biotechnol. Biofuels* **11**, 1–17 (2018).
10. Quackenbush, J. Computational analysis of microarray data. *Nat. Rev. Genet.* **2**, 418–427 (2001).
11. Maere, S., Heymans, K. & Kuiper, M. BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. *Bioinformatics* **21**, 3448–3449 (2005).
12. Sengupta, A., Sunder, A. V., Sohoni, S. V. & Wangikar, P. P. Fine-Tuning Native Promoters of *Synechococcus elongatus* PCC 7942 to Develop a Synthetic Toolbox for Heterologous Protein Expression. *ACS Synth. Biol.* **8**, 1219–1223 (2019).
13. Hamilton, T. L., Klatt, J. M., De Beer, D. & Macalady, J. L. Cyanobacterial photosynthesis under sulfidic conditions: Insights from the isolate Leptolyngbya sp. strain hensonii. *ISME J.* **12**, 568–584 (2018).
14. Sengupta, A., Madhu, S., Wangikar, P. P. & Accepted, J. A library of tunable , portable and inducer-free promoters derived from cyanobacteria. (2020) doi:10.1021/acssynbio.0c00152.