Data Article

Dataset on the expression level of the genes involved in the synthesis of structural molecules in carbon-deficient microalgae

Cheng Qilu \textsuperscript{a}, Xu Ligen \textsuperscript{b}, Cheng Fangmin \textsuperscript{b}, Pan Gang \textsuperscript{b}, Zhou Qifa \textsuperscript{a,*}

\textsuperscript{a} College of Life Sciences, Zhejiang University, Hangzhou 310058, China
\textsuperscript{b} College of Agriculture and Biotechnology, Zhejiang University, Hangzhou 310058, China

\textbf{A R T I C L E   I N F O}

Article history:
Received 11 August 2018
Received in revised form 14 August 2018
Accepted 17 September 2018
Available online 20 September 2018

\textbf{A B S T R A C T}

The data presented in this article are related to the research article entitled “Bicarbonate-rich wastewater as a carbon fertilizer for culture of Dictyosphaerium sp. of a giant pyrenoid” (Cheng et al., 2018) [1]. This article provides data about the expression levels of the genes involved in the synthesis of structural molecules in the carbon-deficient algal cell and the carbon-treated algal cell, which can be helpful for analyzing the observed disruption of the structural integrity in the carbon-deficient microalgae at molecular level.

& 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

\textbf{Specifications table}

Subject area Environmental science and Biology
More specific subject area Wastewater; Algal culture and Algal physiology
Type of data Table, image and figure

DOI of original article: https://doi.org/10.1016/j.jclepro.2018.08.066
* Corresponding author.
E-mail addresses: qiluch316@163.com (C. Qilu), xuligeng@126.com (X. Ligen), chengfm@zju.edu.cn (C. Fangmin), pagang@zju.edu.cn (P. Gang), zzqqqq@zju.edu.cn (Z. Qifa).

https://doi.org/10.1016/j.dib.2018.09.045
2352-3409/© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
| How data was acquired | Microscope and RNA-sequence |
|-----------------------|-----------------------------|
| Data format           | Raw and filtered.           |
| Experimental factors  | The microalgae was isolated from wastewater. |
| Experimental features | Algae culture in an incubator, microscopic observation and RNA-sequence. |
| Data source location  | Hangzhou City, Zhejiang Province, China |
| Data accessibility    | The data is available within this article |
| Related research article | Cheng et al. 2018. Bicarbonate-rich wastewater as a carbon fertilizer for culture of Dictyosphaerium sp. of a giant pyrenoid. Journal of Cleaner Production (in press). |

### Value of the data

- The data provided comparative transcriptomic analysis for the related Dictyosphaerium sp. genes encoding synthesis of structural molecules in the algae cultured in the modified Hoagland + Wastewater medium and the modified Hoagland medium.
- The obtained data from the RNA-seq analysis can be used for analysis of the disruption of structural integrity in the carbon-deficient microalgae at molecular level.

### 1. Data

The dataset of this article provides information on the expression levels of the genes involved in synthesis of structural molecules in the microalgae. Tables 1–3 show the expression levels of extensin protein genes, proline-rich structural protein genes, gline-rich structural protein genes, and soluble starch synthase genes and granule-bound starch synthase genes, respectively. Figs. 1 and 2 show the expression levels of lipid synthase genes and cellulose synthase catalytic subunit genes and cellulose synthase genes, respectively.

### 2. Experimental design, materials, and methods

The green algae strain Dictyosphaerium sp. was isolated from wastewater originating from the Experimental Farm at Zhejiang University. Prior to cultivation experiments, the species was cultured in BG11 medium until reaching log phase. A batch of algal culture experiments were conducted in a model ACM–168 algal incubator (Jiangnan Instrument Co., Ningbo, China). Two treatments with three replications were conducted with 0 and 0.25 L/L of the autoclaved swine wastewater to modified Hoagland solution [2], and the wastewater-added medium contained 0.8 g/L bicarbonate. Microalgae seeds were added to different culture media in 250 mL triangular glass flasks with 200 mL of working solution, and with initial optical density values adjusted to between 0.05 and 0.07 absorbance at a wavelength of 680 nm (OD680).

RNA sampling and extraction was performed according to methods described in [3]. On the 13th day, 10 mL of cultures were collected by centrifugation (10,000 g, 7 min) from each replicate. The supernatant was discarded and the resulting cell pellets were immediately flash-frozen with liquid nitrogen and stored at −80 °C. Prior to RNA extraction, cell pellets were resuspended in lysis buffer and then ground using a micropestle. Total RNA was then extracted following the manufacturer’s instructions. Total RNA was maintained as replicates, rather than pooling, and then stored at −80 °C. An Agilent 2100 Bioanalyzer (Agilent RNA 6000 Nano Kit) was used to determine QC:RNA concentrations, RIN values, 28 S/18 S, and fragment length distributions. A NanoDropTM spectrophotometer was used to assess the purity of the RNA. Aliquots from mRNA samples were used to construct cDNA libraries.
| GI       | Length(bp) | Protein       | Cwc-EL | C0-EL | log2FC | Padj   |
|----------|------------|---------------|--------|-------|--------|--------|
| XP_011043443.1 | 502        | L-R repeat E-P 4 | 0.43   | 1.50  | 1.81   | 0.44   |
| XP_014660371.1 | 502        | E-P-like       | 8.81   | 0.26  | -5.07  | 0.016  |
| NP_001147655.1 | 303        | E-P            | 0.26   | 1.15  | 2.13   | 0.40   |
| XP_010691715.1 | 1029       | L-R repeat E-P 3 | 210.29 | 1447.15 | 2.78 | 2.76E-10 |
| XP_008224469.1 | 733        | L-R repeat E-P 3 | 0.16   | 0.11  | -0.62  | 0.76   |
| XP_009113077.1 | 2139       | L-R repeat E-P 2 isoform X1 | 1239.01 | 234.56 | 5.07 | 0.016 |
| XP_010691715.1 | 485        | L-R repeat E-P 3 | 241.40 | 385.51 | 0.68  | 0.12   |
| XP_014660371.1 | 841        | E-P-like       | 0.47   | 39.44 | 6.40   | 5.98E-07 |
| XP_010691715.1 | 1067       | L-R repeat E-P 3 | 240.82 | 6.79  | 4.27   | 0.025  |
| XP_014660371.1 | 2004       | E-P-like       | 226.18 | 11.70 | 4.27   | 0.06   |
| XP_014660371.1 | 1889       | E-P-like       | 203.96 | 364.59 | 0.84  | 0.23   |
| XP_010257306.1 | 485        | pollen-specific L-R repeat E-P 3 isoform X1 | 0.15 | 3.59 | 4.62 | 0.032 |

C. Qilu et al. / Data in Brief 20 (2018) 1870–1876
| GI            | Length(bp) | Protein                  | Cw-EL | Cq-EL | log2FC       | Padj     |
|--------------|------------|--------------------------|-------|-------|--------------|----------|
| XP_007013165.1 | 346        | PR-P E-P-like receptor kinase 1 | 0.17  | 13.30 | 6.30         | 0.00075  |
| XP_014660371.1 | 3836       | E-P-like                 | 0.71  | 33.83 | 5.58         | 0.00067  |
| XP_010448887.1 | 727        | L-R repeat E-P 4         | 0.86  | 0.26  | -1.72        | 0.50     |
| XP_009118855.1 | 319        | E-P-like                 | 0.13  | 2.11  | 3.97         | 0.081    |
| XP_015057744.1 | 831        | E-P-like                 | 0.15  | 5.85  | 5.26         | 0.0099   |
| XP_014660371.1 | 1605       | E-P-like                 | 1.30  | 0.31  | -2.05        | 0.42     |
| XP_014660371.1 | 1672       | E-P-like                 | 3.69  | 86.21 | 4.55         | 1.19e-14 |
| XP_010448982.1 | 477        | L-R repeat E-P 5         | 0.40  | 5.55  | 3.81         | 0.020    |
| XP_013743627.1 | 950        | pollen-specific L-R repeat E-P 1 | 148.85 | 286.64 | 0.95        | 0.18     |
| XP_008224469.1 | 1413       | L-R repeat E-P 3         | 306.51| 31.40 | -3.29        | 8.98e-26 |
| XP_007013165.1 | 346        | PR-P E-P-like receptor kinase 1 | 0.17  | 13.30 | 6.30         | 0.00075  |
| XP_014660371.1 | 711        | E-P-like                 | 42.98 | 3.81  | -3.50        | 0.15     |
| XP_010448982.1 | 972        | L-R repeat E-P 4         | 111.37| 30.23 | -1.83        | 2.20e-07 |
| XP_006480071.1 | 249        | pollen-specific L-R repeat E-P 4 | 0.48  | 0.20  | -1.25        | 0.61     |
| XP_008224469.1 | 2210       | L-R repeat E-P 3         | 3.76  | 626.49| 7.38         | 2.57e-16 |
| XP_014660371.1 | 779        | E-P-like                 | 0.68  | 2.66  | 1.97         | 0.36     |
| XP_010647090.1 | 377        | E-P-like                 | 0.19  | 16.62 | 6.42         | 0.00067  |
| XP_009348443.1 | 1192       | L-R repeat E-P 4         | 178.48| 127.98| -0.48        | 0.081    |
| XP_014660371.1 | 1192       | E-P-like                 | 42.98 | 3.81  | -3.50        | 0.15     |
| XP_006480071.1 | 249        | pollen-specific L-R repeat E-P 4 | 0.48  | 0.20  | -1.25        | 0.61     |
| XP_008224469.1 | 2210       | L-R repeat E-P 3         | 3.76  | 626.49| 7.38         | 2.57e-16 |
| XP_014660371.1 | 779        | E-P-like                 | 0.68  | 2.66  | 1.97         | 0.36     |
| XP_010647090.1 | 377        | E-P-like                 | 0.19  | 16.62 | 6.42         | 0.00067  |
| XP_009348443.1 | 1192       | L-R repeat E-P 4         | 178.48| 127.98| -0.48        | 0.081    |
| XP_014660371.1 | 1192       | E-P-like                 | 42.98 | 3.81  | -3.50        | 0.15     |
| XP_006480071.1 | 249        | pollen-specific L-R repeat E-P 4 | 0.48  | 0.20  | -1.25        | 0.61     |
| XP_008224469.1 | 2210       | L-R repeat E-P 3         | 3.76  | 626.49| 7.38         | 2.57e-16 |
| XP_014660371.1 | 779        | E-P-like                 | 0.68  | 2.66  | 1.97         | 0.36     |
| XP_010647090.1 | 377        | E-P-like                 | 0.19  | 16.62 | 6.42         | 0.00067  |
| XP_009348443.1 | 1192       | L-R repeat E-P 4         | 178.48| 127.98| -0.48        | 0.081    |
| XP_014660371.1 | 1192       | E-P-like                 | 42.98 | 3.81  | -3.50        | 0.15     |
After quality filtering sequence reads, clean reads were mapped to the genomic reference using Bowtie2 [2], and gene expression levels were calculated with RSEM [4]. Differentially expressed genes (DEGs) were determined with DEseq. 2 [5].

Table 2
Comparative transcriptomic analysis for the related *Dictyosphaerium* sp. genes encoding gline-rich structural proteins (GR-Ps) in the algae cultured in the modified Hoagland + Wastewater medium (Cwc, Control treatment) and the Hoagland medium (C0). Data are means of three replications. GI-gene ID, length-gene length, log2FC-log2 transformed fold change between control and treat samples, Padj- Statistic of adjusted p value (DEseq. 2 method used).

| GI       | Length(bp) | Protein                  | Cwc-EL | CO-EL | log2FC | Padj  |
|----------|------------|--------------------------|--------|-------|--------|-------|
| XP_015574762.1 | 351     | GR-P-like                 | 0.13   | 0.17  | 0.42   | 0.83  |
| XP_006651317.1 | 1054    | GR-P-like                 | 1.41   | 2.92  | 1.05   | 0.70  |
| XP_008666361.1 | 444     | GR-P 1.0-like             | 0.25   | 0.27  | 0.11   | 0.97  |
| XP_010466892.1 | 514     | GR-P 1-like               | 0.23   | 0.39  | 0.77   | 0.77  |
| XP_010466892.1 | 497     | GR-P 1-like               | 0.60   | 0.81  | 0.42   | 0.87  |
| XP_013681405.1 | 480     | GR-P 1.8-like isoform X1 | 24.21  | 138.83| 2.52   | 0.33  |
| XP_015574762.1 | 841     | GR-P-like                 | 0.35   | 3.03  | 3.10   | 0.20  |
| XP_012570960.1 | 630     | GR-P-like isoform X6      | 0.24   | 2.46  | 3.34   | 0.15  |
| XP_015574762.1 | 4268    | GR-P-like                 | 0.23   | 0.39  | 0.77   | 0.77  |
| XP_015165590.1 | 510     | GR-P 1.8-like             | 0.13   | 0.17  | 0.42   | 0.83  |
| XP_010467102.1 | 722     | GR-P 1.0-like isoform X10 | 5.08   | 11.83 | 1.22   | 0.29  |
| XP_015167339.1 | 3234    | GR-P 1                    | 73.83  | 153.22| 1.05   | 0.43  |
| XP_010688958.1 | 854     | GR-P-like isoform X1      | 8.55   | 52.39 | 2.62   | 0.012 |
| XP_009143614.1 | 249     | GR-P 1                    | 1.09   | 1.83  | 0.75   | 0.80  |
| XP_009143614.1 | 594     | GR-P 1                    | 1.69   | 7.46  | 2.14   | 0.13  |
| XP_013668804.1 | 1666    | GR-P 1-like               | 0.97   | 1.97  | 1.02   | 0.66  |
| XP_013731827.1 | 1381    | GR-P-like                 | 349.42 | 289.25| 0.27   | 0.77  |
| XP_010424670.1 | 359     | GR-like                   | 0.22   | 0.37  | 0.76   | 0.77  |
| XP_008671409.1 | 1553    | GR-P 1.0-like             | 16.78  | 103.53| 2.63   | 0.00082|
| XP_008671490.1 | 385     | GR-P 1.8-like             | 0.18   | 1.01  | 4.25   | 0.35  |
| XP_010467104.1 | 890     | GR-P 1.0-like isoform X12 | 0.50   | 2.75  | 2.44   | 0.29  |
| XP_010466892.1 | 853     | GR-P 1-like               | 125.17 | 288.18| 1.20   | 0.15  |
| XP_014628139.1 | 392     | GR-P 1.8                  | 27.57  | 16.66 | 0.73   | 0.58  |
| XP_010467093.1 | 1047    | GR-P 1.8-like isoform X1  | 99.08  | 5.45  | 14.19  | 0.0052|
| XP_014628139.1 | 369     | GR-P 1.8                  | 57.90  | 52.71 | 0.14   | 0.88  |
| XP_015580618.1 | 923     | GR-P 1.8 isoform X2       | 0.21   | 0.34  | 0.71   | 0.78  |
| XP_011465792.1 | 254     | GR-P 1.8                  | 38.54  | 28.54 | 0.43   | 0.77  |
| XP_012570962.1 | 906     | GR-P-like isoform X10     | 61.95  | 77.82 | 0.33   | 0.77  |
| XP_010317299.1 | 478     | putative GR-P 1           | 0.37   | 0.88  | 1.25   | 0.66  |
| XP_015389986.1 | 418     | putative GR-P 1           | 0.16   | 0.41  | 1.39   | 0.63  |
| XP_006490141.1 | 952     | GR-P-like                 | 1173.51| 5692.52| 2.28  | 6.05E−07|
| XP_008229859.1 | 950     | GR-P                     | 1.82   | 2.83  | 0.64   | 0.77  |
| XP_010466892.1 | 554     | GR-P 1-like               | 45.86  | 40.86 | 0.17   | 0.85  |
| XP_010432984.1 | 328     | putative GR-P 1           | 0.21   | 0.68  | 1.71   | 0.54  |
| XP_010467093.1 | 1149    | GR-P 1.8-like isoform X1  | 29.54  | 0.36  | 6.35   | 0.00014|
| XP_010456984.1 | 1304    | GR-P 1-like               | 0.83   | 0.23  | 1.84   | 0.51  |
| XP_012570964.1 | 1671    | GR-P-like isoform X12     | 0.29   | 0.29  | 0.02   | 0.99  |
| XP_009336768.1 | 1071    | putative GR-P 1           | 186.65 | 605.46| 1.70   | 0.092 |
| XP_015574762.1 | 707     | GR-P-like                 | 0.16   | 0.41  | 1.39   | 0.63  |
| XP_010424670.1 | 454     | GR-P-like                 | 0.66   | 0.62  | 0.10   | 0.98  |
| XP_010680074.1 | 1244    | GR-P-like                 | 2.88   | 22.39 | 2.96   | 0.014 |
| XP_0104961.1  | 998     | GR-P 1.8; Short=GRP 1.8; Flags: Precursor | 282.59| 1011.12| 1.84  | 0.00066|
| XP_010424670.1 | 925     | GR-P-like                 | 0.73   | 3.77  | 2.36   | 0.25  |
| XP_012570964.1 | 473     | GR-P-like isoform X12     | 76.32  | 100.18| 0.39   | 0.65  |
Table 3
Comparative transcriptomic analysis for the related *Dictyosphaerium* sp. genes encoding soluble starch synthase (SSS) and granule-bound starch synthase (GBSS) in the algae cultured in the modified Hoagland+Wastewater medium (*C*wc, Control treatment) and the modified Hoagland medium (*C*0). Data are means of three replications. GI-gene ID, length-gene length, log2FC-log2 transformed fold change between control and treat samples, Padj- Statistic of adjusted pvalue (DEseq. 2 method used).

| GI          | Length(bp) | Protein                           | Cwc-EL    | CO-EL     | log2FC | Padj    |
|-------------|------------|-----------------------------------|-----------|-----------|--------|---------|
| BAE79814.1  | 376        | GBSS                              | 1043.71   | 236.79    | −2.14  | 4.30E−12|
| XP_011398759.1 | 5924     | SSS 3, chloroplastic/amyloplastic | 2879.90   | 930.15    | −1.63  | 2.71E−15|
| BAE79814.1  | 357        | GBSS                              | 3.01      | 223.15    | 6.21   | 3.60E−17|
| XP_001697117.1 | 1674     | GBSS-I                           | 559.29    | 1833.07   | 1.71   | 0.014   |
| BAE79814.1  | 562        | GBSS                              | 0.95      | 12.94     | 3.76   | 0.0015  |
| BAE79814.1  | 292        | GBSS                              | 744.96    | 148.47    | −2.33  | 4.62E−17|
| BAE79814.1  | 292        | GBSS                              | 744.96    | 148.47    | −2.33  | 4.62E−17|
| BAE79814.1  | 2113       | GBSS                              | 4963.46   | 2460.34   | −1.01  | 4.80E−05|
| XP_011401562.1 | 2260    | putative GBSS 1, chloroplastic/amyloplastic | 2.83      | 82.37     | 4.86   | 4.57E−13|
| BAE79814.1  | 562        | GBSS                              | 0.95      | 12.94     | 3.76   | 0.0015  |
| XP_001697117.1 | 1030     | GBSS-I                           | 49.21     | 94.50     | 0.94   | 0.041   |
| BAE79814.1  | 376        | GBSS                              | 1043.71   | 236.79    | −2.14  | 4.30E−12|
| BAE79814.1  | 4024       | GBSS                              | 1101.80   | 2197.76   | 1.00   | 0.011   |
| BAE79814.1  | 537        | GBSS                              | 3.01      | 223.15    | 6.21   | 3.60E−17|
| XP_001697117.1 | 1674     | GBSS-I                           | 559.29    | 1833.07   | 1.71   | 0.014   |
| BAE79814.1  | 2113       | GBSS                              | 4963.46   | 2460.34   | −1.01  | 4.80E−05|
| XP_011398759.1 | 6692     | SSS 3, chloroplastic/amyloplastic | 12.72     | 384.67    | 4.92   | 2.36E−22|
| XP_005642568.1 | 375      | SSS                              | 14.16     | 3.84.67   | −3.02  | 0.0026  |
| XP_011398759.1 | 496      | SSS 3, chloroplastic/amyloplastic | 41.16     | 9.23      | −2.16  | 0.00011 |
| AAC17970.2  | 259        | SSS                               | 9.10      | 8.88      | −0.036 | 0.97    |
| XP_011398759.1 | 6692     | SSS 3, chloroplastic/amyloplastic | 12.72     | 384.67    | 4.92   | 2.36E−22|
| XP_005642568.1 | 375      | SSS                              | 14.16     | 3.84.67   | −3.02  | 0.0026  |
| XP_001695327.1 | 248      | SSS- III                         | 4.19      | 4.55      | 0.12   | 0.92    |
| XP_011398759.1 | 5924     | SSS 3, chloroplastic/amyloplastic | 2879.90   | 930.15    | −1.63  | 2.71E−15|
| XP_011398759.1 | 496      | SSS 3, chloroplastic/amyloplastic | 41.16     | 9.23      | −2.16  | 0.00011 |

Fig. 1. The expression level of the genes encoding the lipid synthases in the algae cultured in the modified Hoagland+Wastewater medium (*C*wc) and the Hoagland medium (*C*0). Data are means of three replications. SS- sulfolipid synthase, CFAPS- cyclopropane-fatty-acyl-phospholipid synthase, MFAPS- /methylene-fatty-acyl-phospholipid synthase.
Acknowledgments

This work is supported by the National Key Research and Development Program of China (Grant number, 2017YFD0300103) and Huzhou Municipal Science and Technology Bureau (Grant number, 2015GZ08).

Transparency document. Supporting information

Transparency data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.09.045.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2018.09.045.

References

[1] Q.L. Cheng, L.G. Xu, F.M. Cheng, G. pan, Q.F. Zhou, Bicarbonate-rich wastewater as a carbon fertilizer for culture of Dictyosphaerium sp. of a giant pyrenoid, J. Clean. Prod. 202 (2018) 439–443.
[2] J. Zhang, X. Wang, Q. Zhou, Co-cultivation of Chlorella spp and tomato in a hydroponic system, Biomass- Bioenergy 97 (2017) 132–138.
[3] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat. Methods 9 (2012) 357–359.
[4] B. Li, C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, BMC Bioinform. 12 (2011) 323.
[5] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq. 2, Genome Biol. 5 (2014) 550.