Supplemental Information

## Materials and Methods

### 1 RNA Analyses

Total RNA was extracted by using TRIzol reagent (Invitrogen) and the RNeasy Plus Mini Kit (Qiagen). Samples were resuspended in 1.2 mL TRIzol reagent (Thermo Fisher Scientific) in lysing matrix D tubes (MP Biomedicals), homogenized by a FastPrep-24 machine (MP Biomedicals, 3 cycles, 8.0 m $s^{-1}$, 30 s, 3 min ice-chilling at interval), followed by centrifugation at 12,000 × g for 3 min at 4 °C (Eppendorf 5430R). RNA was extracted using a standard phenol-chloroform method (Chomczynski and Sacchi 1987). For more detail information, please see Zhang et al. (2021).

### 2 Bioinformatics

The quality of the raw reads was assessed by Fastqc v.0.11.8 (Andrews, 2010) and Fastq screen v.0.13.0 (Wingett and Andrews, 2018) and summarized using Multiqc (Ewels et al., 2016). Trimming of the raw reads was performed to remove low-quality bases and adapter sequences with Trimmomatic v.0.38 (Bolger et al., 2014). *De novo* transcriptome assembly was performed with the Trinity's version 2.11.0 (Grabherr et al., 2011), and the low-quality assembly were removed with CD-HIT (Li and Godzik, 2006). A preliminary assessment of *de novo* assembly quality was performed with Transrate (Smith-Unna et al., 2016) and Busco (Hara et al., 2015), and the completeness assessment yielded high scores for all assemblies. Open reading frames (ORFs) were then predicted using TransDecoder version v5.5.0 (Haas et al., 2013), and were then annotated by Blastx, Hmmpress, Signalp, Rnammer, PFam (Lagesen et al., 2007). All annotations were loaded and integrated with Trinotate v3.0.0 (Haas et

al., 2013). ORFs were further functionally annotated and assigned to the KEGG and GhostKOALA (Moreno-Santillán et al., 2019). Cleaned and trimmed reads of each sample were mapped to the assembled transcriptomes by salmon v0.9.1 (Patro et al., 2017). The differential expression was them calculated by DESeq2 v1.24.0 (Love et al., 2014) with a Benjamini-Hochberg adjusted $p$-value < 0.05. Data analysis and visualizations were made using R v.3.6.1 (Team, 2020), packages ggplot2 v.3.2.0 (Wickham, 2016) and Pheatmap v.1.0.12 (Kolde, 2015).

**Table S1.** Results of three-way ANOVAs of the effects of light intensity (L), dissolved inorganic phosphorus ($PO_4^{3-}$) concentration (P), $pCO_2$ level (C), and their interaction on growth rate, cellular contents of POC, PON, POP, PIC, carbohydrate and protein, and the ratios of POC : PON, POC : POP, PON : POP and PIC : POC, and the percentages of POC allocated to carbohydrate (carbohydrate–C : POC) and protein (protein–C : POC), and the percentage of PON allocated to protein (protein–N : PON). Please see figure 1 for more detailed information.

| Parameter | | L | P | C | L×P | L×C | P×C | L×P×C |
|---|---|---|---|---|---|---|---|---|
| Growth rate | F | 4588.5 | 198.3 | 225.7 | 6.7 | 14.7 | 100.9 | 8.6 |
| | *p* | <0.01 | <0.01 | <0.01 | =0.02 | <0.01 | <0.01 | <0.01 |
| POC | F | 864.7 | 0.6 | 0.6 | 36.4 | 8.1 | 0.2 | 2.6 |
| | *p* | <0.01 | =0.46 | =0.46 | <0.01 | <0.01 | =0.65 | =0.12 |
| PON | F | 172.2 | 70.6 | 17.7 | 1.2 | 6.3 | 10.0 | 0.2 |
| | *p* | <0.01 | <0.01 | <0.01 | =0.33 | =0.02 | <0.01 | =0.65 |
| POP | F | 188.3 | 724.8 | 8.2 | 6.6 | 0.1 | 0.2 | 0.2 |
| | *p* | <0.01 | <0.01 | <0.01 | =0.02 | =0.89 | =0.65 | =0.65 |
| PIC | F | 329.4 | 109.1 | 47.2 | 17.7 | 8.9 | 1.9 | 0.1 |
| | *p* | <0.01 | <0.01 | <0.01 | <0.01 | <0.01 | =0.17 | =0.98 |
| POC : PON | F | 207.6 | 396.4 | 0.1 | 14.9 | 0.2 | 9.3 | 2.1 |
| | *p* | <0.01 | <0.01 | =0.96 | <0.01 | =0.65 | <0.01 | =0.16 |
| POC : POP | F | 7.7 | 351.3 | 18.7 | 0.1 | 0.3 | 2.6 | 1.0 |
| | *p* | =0.01 | <0.01 | <0.01 | =0.85 | =0.62 | =0.12 | =0.32 |
| PON : POP | F | 29.4 | 183.1 | 29.9 | 1.9 | 0.1 | 0.3 | 0.1 |
| | *p* | <0.01 | <0.01 | <0.01 | =0.18 | =0.97 | =0.61 | =0.94 |
| PIC : POC | F | 90.2 | 71.9 | 77.6 | 0.6 | 6.6 | 2.4 | 0.3 |
| | *p* | <0.01 | <0.01 | <0.01 | =0.46 | =0.02 | =0.14 | =0.61 |
| Carbohydrate | F | 925.2 | 37.7 | 23.6 | 40.9 | 7.9 | 0.9 | 0.9 |
| | *p* | <0.01 | <0.01 | <0.01 | <0.01 | <0.01 | =0.34 | =0.34 |
| Protein | F | 782.4 | 77.0 | 1.0 | 21.1 | 0.8 | 2.0 | 0.8 |
| | *p* | <0.01 | <0.01 | =0.34 | <0.01 | =0.38 | =0.17 | =0.37 |
| Carbohydrate–C : POC | F | 793.3 | 76.2 | 17.1 | 4.9 | 0.7 | 0.6 | 0.1 |
| | *p* | <0.01 | <0.01 | <0.01 | =0.04 | =0.41 | =0.43 | =0.95 |
| Protein–C : POC | F | 6.3 | 120.0 | 0.7 | 2.4 | 0.6 | 2.2 | 2.7 |
| | *p* | =0.02 | <0.01 | =0.42 | =0.13 | =0.45 | =0.15 | =0.11 |
| Protein–N : PON | F | 36.3 | 4.3 | 2.8 | 9.2 | 0.7 | 0.7 | 0.2 |
| | *p* | <0.01 | =0.05 | =0.10 | <0.01 | =0.41 | =0.41 | =0.63 |

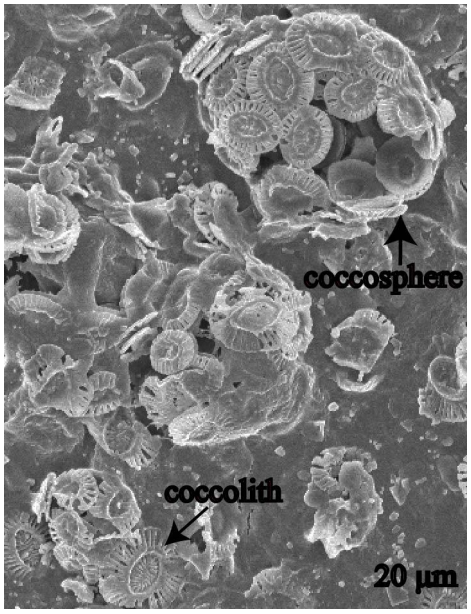


**Figure S1.** A scanning electron microscopy (SEM) picture of the coccosphere and coccolith of *E. huxleyi* RCC 1266.

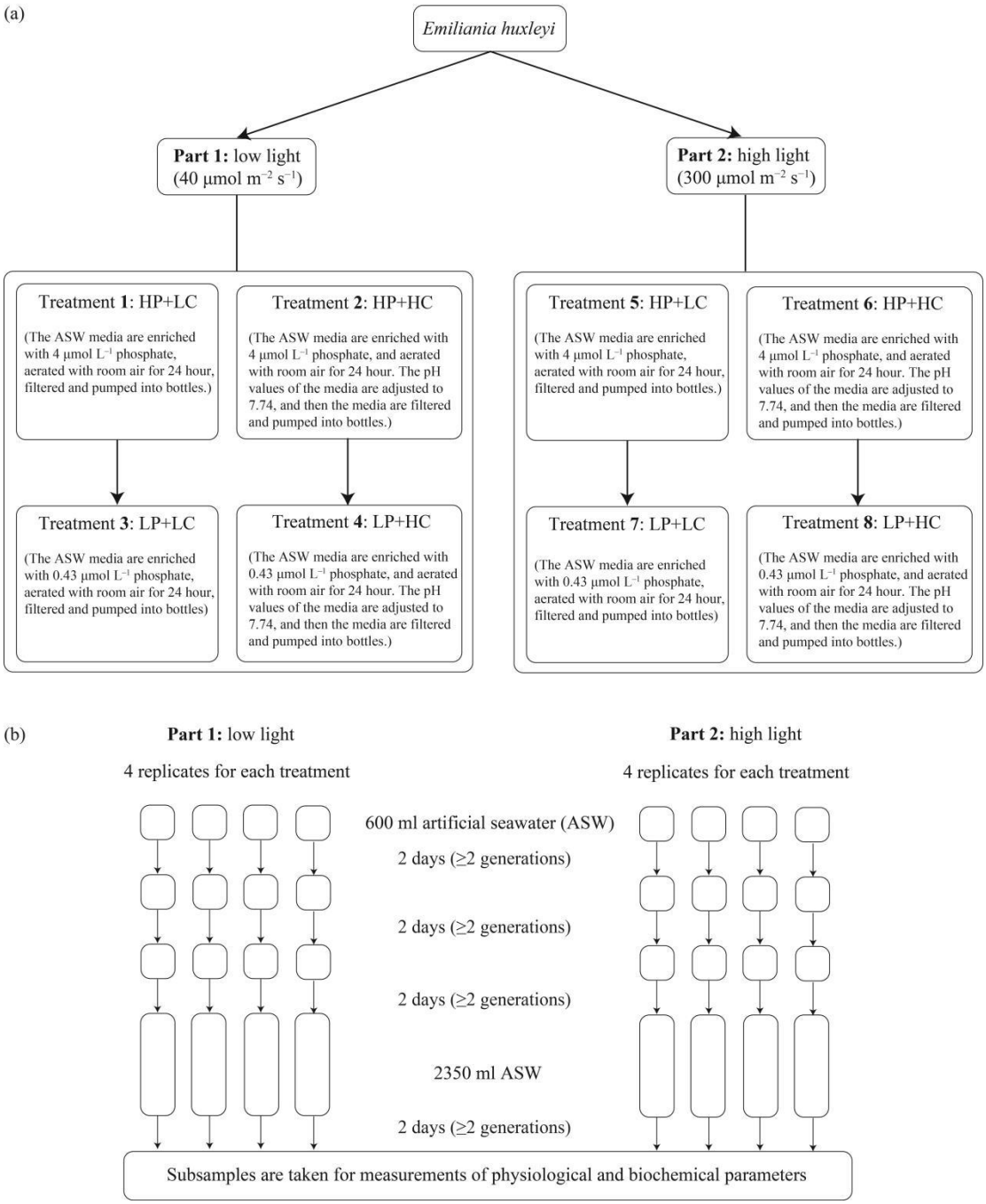


**Figure S2.** Flow chart of the experimental processes (**a**) and flow chart of the pre-culture and experimental cultures in each treatment (**b**). At the beginning of the incubations, HP and LP represent 4 and 0.43 µmol $L^{-1}$ $PO_4^{3-}$, and LC and HC represent low (about 426 µatm) and high $CO_2$ (about 946 µatm) concentrations, respectively.

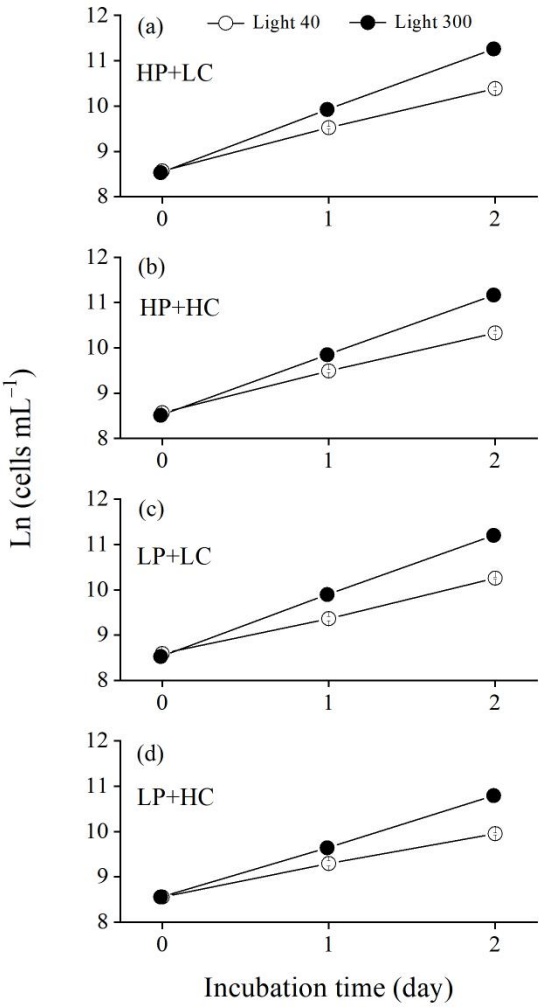


**Figure S3.** Cell density of *E. huxleyi* RCC1266 in the treatments of HP+LC, HP+HC, LP+LC and LP+HC under low light (empty) and high light (fill) intensities during the experimental cultures. The cells were inoculated to achieve an initial density of about 5000 cell mL$^{-1}$, and cultured in each experimental condition for 2 days and then diluted to the initial cell density again. This process was repeated four times. The data represents the means and standard deviation of four independent cultures in the fourth incubation. Please see figure 1 for more detailed information.

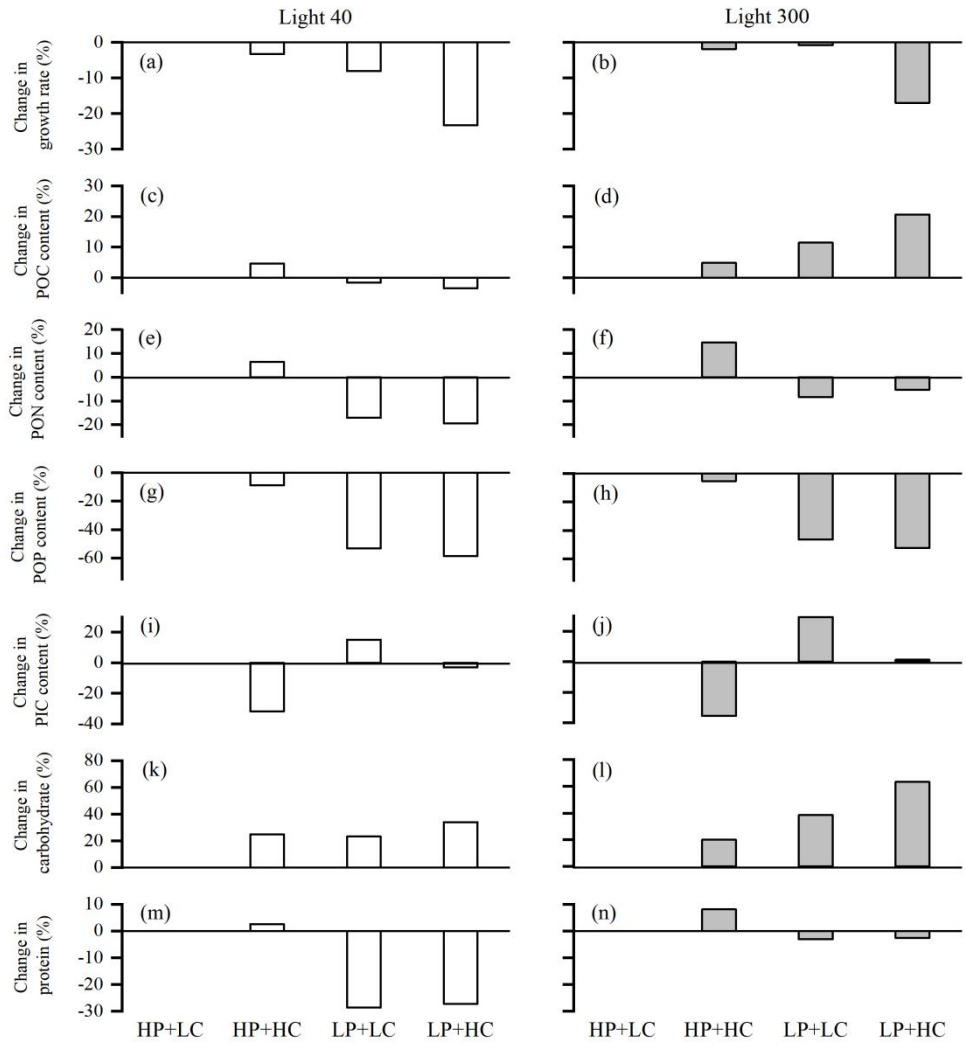


**Figure S4.** In comparison to the HP+LC treatment, changes in growth rate (**a**, **b**) and cellular contents of POC (**c**, **d**), PON (**e**, **f**), POP (**g**, **h**), PIC (**i**, **j**), carbohydrate (**k**, **l**) and protein (**m**, **n**) of *E. huxleyi* RCC1266 in the treatments of HP+HC, LP+LC and LP+HC under low light (empty) and high light (fill) intensities. The data represents the means of four independent cultures. Please see figure 1 for more detailed information.

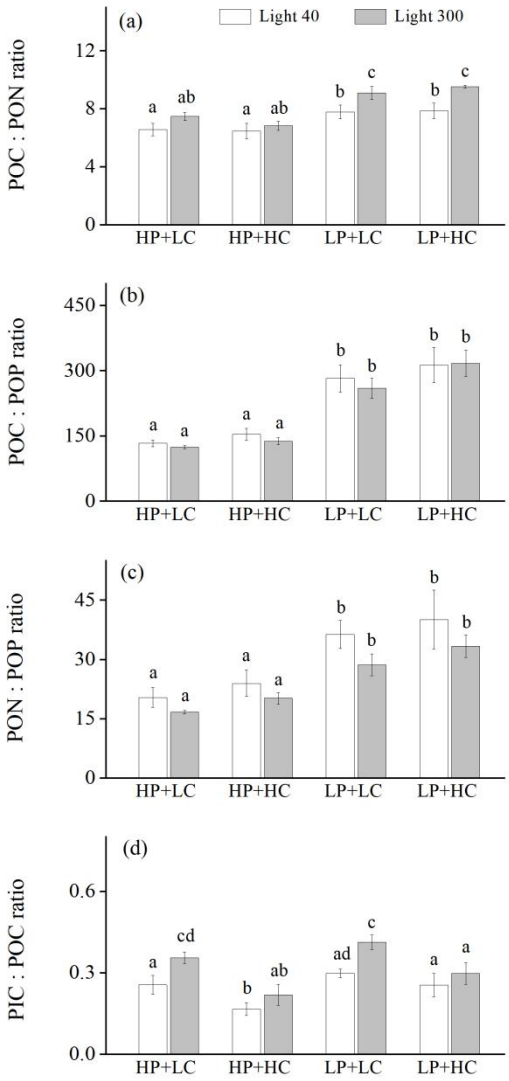


**Figure S5.** The ratios of POC : PON (**a**), POC : POP (**b**), PON : POP (**c**) and PIC : POC (**d**) of *E. huxleyi* RCC1266 in the treatments of HP+LC, HP+HC, LP+LC and LP+HC under low light (empty) and high light (fill) intensities. Different letters represent significant differences in each parameters between treatments ($p < 0.05$). The data represents the means and standard deviation of four independent cultures. Please see figure 1 for more detailed information.

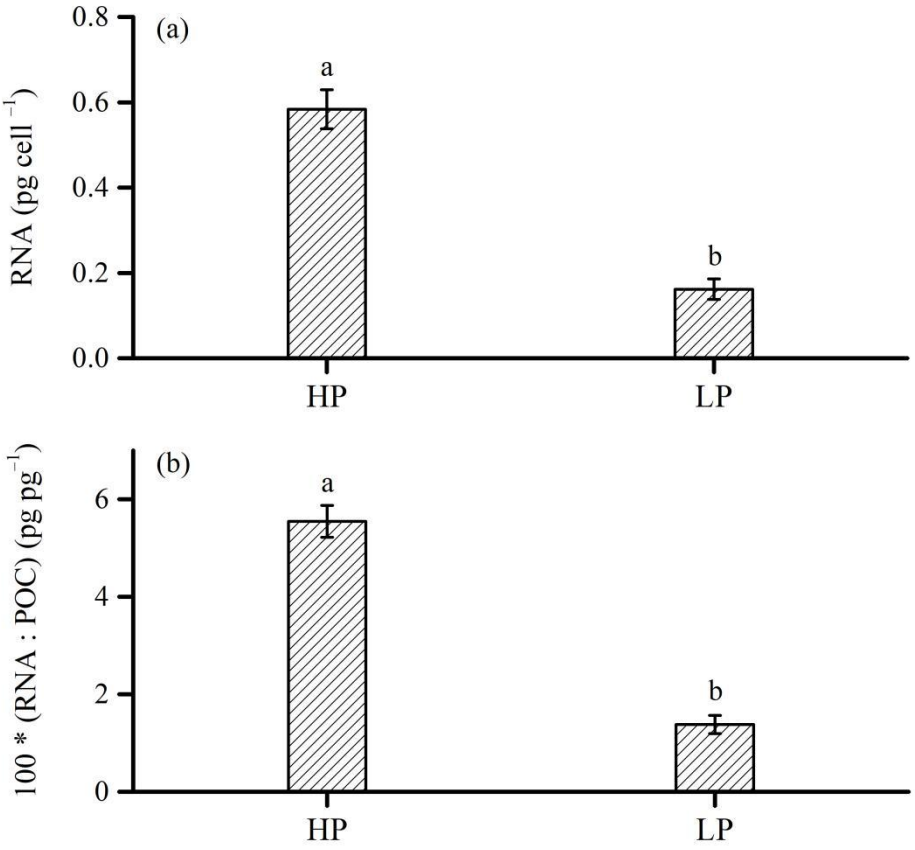


**Figure S6.** Cellular RNA content (**a**) and POC normalized RNA content (**b**) of *E. huxleyi* RCC1266 in the treatments of HP+LC and LP+LC under high light intensity. Different letters in each panel represent significant differences between treatments ($p < 0.05$). The data represents the means and standard deviation of four independent cultures. Please see figure 1 for more detailed information.

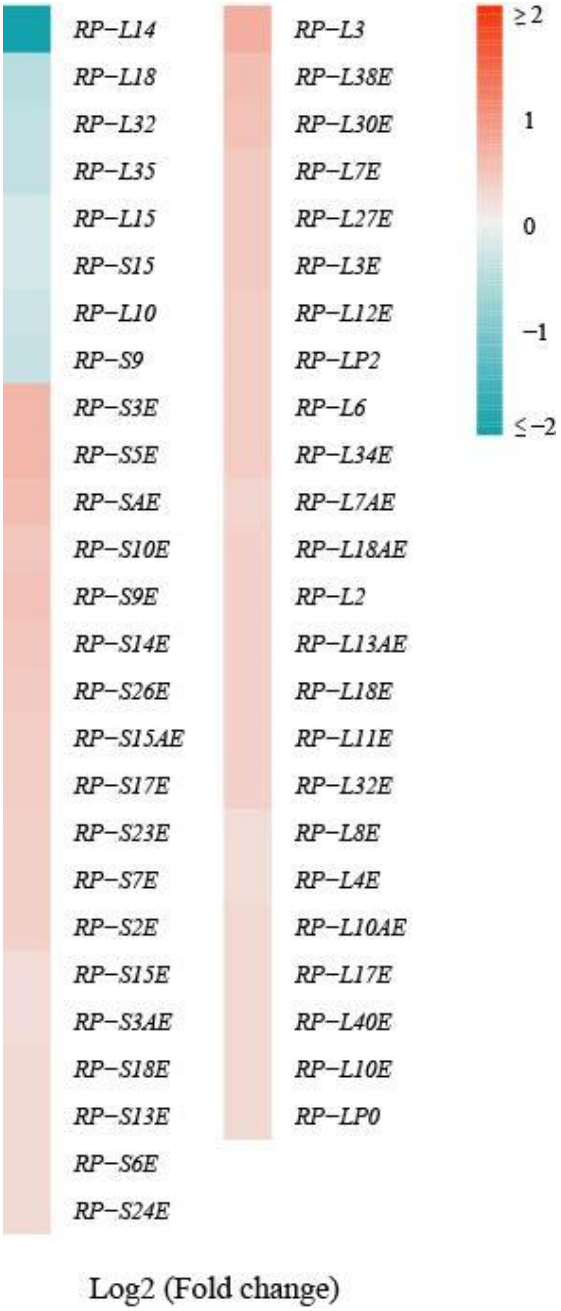


**Figure S7.** Heatmap of significant change in expression of genes linked to ribosome metabolism (Ribosome_ko03010) under the low pH treatment (ocean acidification) in comparison to the high pH treatment (present $CO_2$ level). Red indicates up-regulation of gene, and blue indicates down-regulation of gene. *RP-S* presents genes linked to small subunit ribosomal protein, and *RP-L* presents genes linked to large subunit ribosomal protein.

**References**

Andrews, S.: FASTQC: a quality control tool for high throughput sequence data, Cambridge, 2010. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bolger, A. M., Lohse, M., and Usadel, B.: TRIMMOMATIC : a flexible trimmer for illumina sequence data, Bioinformatics, 30, 2114–2120, 2014.

Chomczynski, P., and Sacchi, N.: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, Anal. Biochem., 162, 156–159, doi: 10.1006/abio.1987.9999, 1987.

Ewels, P., Magnusson, M., Lundin, S., and Käller, M.: MULTIQC: summarize analysis results for multiple tools and samples in a single report, Bioinformatics, 32, 3047–3048, 2016.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, e., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N, and Regev, A.: Full-length transcriptome assembly from RNA-Seq data without a reference genome, Nat. Biotechnol., 29, 644–652, 2011.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., Willian, T., Dewey, C. N., Henschet, R., LeDuc, R. D., Friedman, N., and Regev, A: *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis, Nat. Protoc., 8, 1494–1512, 2013.

Hara, Y., Tatsumi, K., Yoshida, M., Kajikawa, E., Kiyonari, H., and Kuraku, S.: Optimizing and benchmarking *de novo* transcriptome sequencing: from library

preparation to assembly evaluation, BMC Genomics, 16, 977, 2015.

Kolde, R.: Package 'pheatmap'. R. Package, 1, 790, 2015.

Lagesen, K., Hallin, P., Rodland, E. A., Staerfeldt, H. H., Rognes, T., and Ussery, D. W.: RNAmmer: Consistent and rapid annotation of ribosomal RNA genes, Nucleic Acids Res., 35, 3100–3108, 2007.

Li, W., and Godzik, A.: Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics, 22, 1658–1659, 2006.

Love, M. I., Huber, W., and Anders, S.: Moderated estimation of fold change and dispersion for RNA-seq data with DES EQ 2, Genome Biol., 15, 550, 2014.

Moreno-Santillán, D. D., Machain-Williams, C., Hernández-Montes, G., and Ortega, J.: *De novo* transcriptome assembly and functional annotation in five species of bats, Sci. Rep., 9, 6222, 2019.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C.: Salmon provides fast and bias-aware quantification of transcript expression, Nat. Methods., 14, 417–419, 2017.

R Development Core Team: R: a language and environment for statistical computing, v. 4.0.3. [WWW document] URL www.r-project.org [accessed 10 October 2020].

Smith-Unna, R., Boursnell, C., Patro, R., Hibberd, J. M. and Kelly, S.: TransRate: Reference-free quality assessment of *de novo* transcriptome assemblies, Genome Res., 26, 1134–1144, 2016.

Wickham, H.: ggplot2: elegant graphics for data analysis, R package v.3.3.2. [WWW document] URL. https://cran.r-project.org/web/packages/ggplot2/index.html [accessed 17 June 2020], 2016.

Wingett, S. W and Andrews, S.: FASTQ screen: a tool for multi-genome mapping and quality control, F1000Research, 7, 1338, 2018.

Zhang, Y., Li, Z. K., Schulz, K. G., Hu, Y., Irwin, A. J., and Finkel, Z. V.: Growth-dependent changes in elemental stoichiometry and macromolecular allocation in the coccolithophore *Emiliania huxleyi* under different environmental conditions, Limnol. Oceanogr., 66, 2999–3009, doi: 10.1002/lno.11854, 2021.