Turk. J. Fish.& Aquat. Sci. 22(11), TRJFAS21341
https://doi.org/10.4194/TRJFAS21341

RESEARCH PAPER

Turkish Journal of FISHERIES and AQUATIC SCIENCES

De Novo Assembly and Annotation of Microalga *Tetradesmus obliquus* Transcriptome

Ömer Can Ünüvar[1], Ercan Selçuk Ünlü[1,*]

[1]Bolu Abant İzzet Baysal University, Faculty of Arts and Science, Department of Chemistry, Bolu, 14030, Türkiye.

How to cite

Ünüvar, Ö.C., Ünlü, E.S. (2022). *De Novo* Assembly and Annotation of Microalga *Tetradesmus obliquus* Transcriptome. *Turkish Journal of Fisheries and Aquatic Sciences*, *22(11), TRJFAS21341*. https://doi.org/10.4194/TRJFAS21341

Article History
Received 26 January 2022
Accepted 27 June 2022
First Online 05 July 2022

Corresponding Author
Tel.: +905073215482
E-mail: esunlu06@gmail.com

Keywords
De novo assembly
Bioinformatics
Microalgae
Microbial biotechnology
Lutein

Abstract

Tetradesmus obliquus is a unicellular green microalga and considered as a potential source for biotechnological production of pigments such as lutein. No genome-related data is available for *T. obliquus* that would increase the ability to develop new approaches in biotechnological applications. We present the first transcriptome data for *T. obliquus*. The aim of this study is to provide a comprehensive transcriptome annotation and identification of conserved genes involved in lutein pigment biosynthesis in *Tetradesmus obliquus* cells by analyzing pooled RNA-Seq data. Next-Generation Sequencing was applied for the pooled cDNAs library prepared by combining the cell cultures collected from samples exposed to dark and high light intensity conditions. Transcripts were assembled by the d*e novo* assembly approach. Trinotate software was used for functional annotation of assembled transcripts. We also carried out BLAST analysis comparing the transcriptome data against known lutein biosynthesis genes. The 49.15% of the assembled sequences were functionally annotated, providing a total of 21490 unigenes. Our data also revealed the transcript sequences for ten conserved genes required for lutein biosynthesis. The data produced in this study can be used for molecular approaches in biotechnological applications related to *T. obliquus*, such as increasing the yield of pigment production.

Introduction

Algae species belonging to the Chlorophyta (green algae) branch are one or multicellular organisms that can live in both fresh and saltwater. They are regarded by many scientists as the ancestors of species in the plant kingdom (Delaux et al., 2015; Turmel et al., 2007). Like soil plants, they also produce chlorophylls a and b and store starch as a carbohydrate reservoir. Algae can be found in unicellular and multicellular forms in nature. *Tetradesmus obliquus* (formerly known as *Scenedesmus obliquus*) is a green alga and observed as single or colonized cells (Geng et al., 2014). Wide range of biotechnology studies including microscopy, chromatographic characterization, genetic manipulations, and cell sorting have been carried out for several species, especially for genera *Chlamydomonas*, *Chlorella*, *Dunaliella*, *Haematococcus,* and *Scenedesmus* (Ben-Amotz & Avron, 1983; Mishra et al., 2019).

Today, *Tetradesmus obliquus* is one of the most productive species in industrial biotechnology studies due to their high pigment contents and their ability to live at different temperatures, pH, and salinity (Sánchez, Fernández, et al., 2008; Wiltshire et al., 2000). *T. obliquus* has been used by the food and pharmaceutical industries as a source for carotenoid production in addition to wastewater treatment, carbohydrate, biomass, and lipid production (Caprio et al., 2018; Geng et al., 2014; Y. S. Kim et al., 2019; Md Nadzir et al., 2021). In addition to *T. obliquus*, several microalga species have been cultivated to produce carotenoids. Microalgae can produce not only the common carotenoids such as zeaxanthin, lutein, and antheraxanthin but also more specific carotenoids such as astaxanthin, alloxanthin, fucoxanthin, and peridinin.

Carotenoids have indispensable roles for the host organisms since they are involved in several biological processes such as maintaining the proper

photosynthetic function by joining the complexes in photosystems as well as their roles in the regulation of growth and development. Lutein is a photo-synthetic carotenoid classified under the xanthophyll family. It is one of the most well-known carotenoids found in fruits and vegetables. In addition to its role in the photosystem in producing organisms, lutein is a health-friendly natural molecule that has been used against cardiovascular diseases, cancer, and aging-related diseases (Dwyer et al., 2001; Granado et al., 2007; Heber & Lu, 2002). Increasing demand in natural therapeutic approaches mandates finding faster and cheaper alternatives to produce lutein. Marigold species have been used in the production of lutein. However, extraction of lutein from the higher plants requires extensive maintenance efforts, workload, and arable lands. The production of lutein at commercial scale using microalgae can be sustained using large photobioreactors minimizing the cost, time, and complexity of the parameters for maintenance of the cultures compared to the plant-based production (Ho et al., 2014; Sánchez, Fernández-Sevilla, et al., 2008). Studies with microalgae have shown that lutein production provides 10 times more lutein compared to marigold.

Biotechnology applications using molecular genetics are often preferred to improve the production efficiency of different pigment types at commercial-scale (Lagarde et al., 2000; Mann et al., 2000; Pons et al., 2014). However, genome data of the species must be available to apply these approaches. *T. obliquus*, as a promising candidate for lutein production can produce up to 0.5% of the dry weight (Ho et al., 2014), but there is limited data for *T. obliquus* in this sense, yet. There is no available whole genome data on the species except the published draft genome (Carreres et al., 2017). For understanding and controlling the lutein production in *T. obliquus*, more genome-based information is required. In this study, we present the first transcriptome data for *T. obliquus*. To provide a comprehensive study for future experiments, we collected pooled samples for the RNA-Seq experiment to cover more genes possible.

Material and Methods

Culture Conditions

For normal growth conditions, *T. obliquus* cells were cultured in BG11 medium (17.6 mM, $NaNO_3$; 0.23 mM, K_2HPO_4; 0.3 mM, $MgSO_4{\cdot}7H_2O$; 0.24 mM, $CaCl_2{\cdot}2H_2O$; 0.031 mM, $C_6H_8O_7{\cdot}H_2O$; 0.021 mM, $C_6H_8FeNO_7$; 0.0027 mM, $Na_2EDTA{\cdot}2H_2O$; 0.19 mM, Na_2CO_3; 0.0046 mM, H_3BO_3; 0.009 mM, $MnCl_2{\cdot}4H_2O$; 0.00077 mM, $ZnSO_4{\cdot}7H_2O$; 0.0016 mM, $Na_2MoO_4{\cdot}2H_2O$; 0.0003 mM, $CuSO_4{\cdot}5H_2O$; 0.00017 mM, $Co(NO_3)_2{\cdot}6H_2O$; 1 mM, $Na_2S_2O_3$). The cultures were grown under 14 hours of light (150 µmol $m^{-2}s^{-1}$) and 10 hours of dark cycle by continuously mixing with 0.22 µm filtered air until the end of exponential phase.

Since previous studies showed that lutein biosynthesis can be enhanced by the exposure to light in range of the 1700-1900 µmol photon $m^{-2}s^{-1}$ after dark cycle (Del Campo et al., 2001; Sánchez, Fernández, et al., 2008), the cultures were exposed to ~1800 µmol $m^{-2}s^{-1}$ light intensity for approximately 10 hours after final dark-phase period to force the expression of the genes that can be related with lutein synthesis in cDNA samples. Then, 300 mg dry algae samples were separated from these cultures by centrifugation and stored in -80°C freezer for RNA isolation.

cDNA Library Construction and Next-generation Sequencing

Total RNA extractions were carried out using a TRIZol® Plus RNA purification kit (Cat. no: 12183555; Thermo Fisher Scientific, USA). For the extraction of mRNAs from total RNA poly-oligoT beads (Promega, USA) were used. Since no transcriptome-based studies are available on *T. obliquus*, we wanted to increase the efficiency of the library by creating a pooled library as previous studies demonstrated the approach of pooled library on detecting transcripts with low copy number (Wu et al., 2014; Yang et al., 2009). Thus, 10ng mRNA extracts from each sample were pooled at 10 ng/µl concentration and the cDNA library was prepared using TruSeq Stranded mRNA Library Prep kit (Illumina, USA). cDNA library was sequenced using Illumina HiSeq™ 3000 platform.

De novo Assembly and Annotation of *T. obliquus* RNA-Seq Data

Before the assembly, adapter and quality trimming was carried out using Skewer (v0.1) (Jiang et al., 2014) with default parameters. *De novo* assembly sequence assembly was carried out using the Trinity software package (v2014-07-17) with default parameters (k-mer=25) (Haas et al., 2013).

Annotation of the assembly was obtained processing the sequences through SwissProt (Database: UniRef90 database), PFAM (Database: Pfam-A.hmm), eggNOG (Database: COG.funccat) databases via Trinotate tool (v3.0.0) (http://trinotate.github.io/). The results were collected and merged into a comprehensive annotation report using Trinotate software using default parameters. For detailed analysis of the data, custom Perl codes (available upon request) were used for removal of recursive sequences and further clustering and representation of data according to GO, COG, and KEGG annotation distributions for the unigenes that each unigene represents a unique transcript expressed from the genome. The final assembly file contains several copies of similar transcripts. The annotation of the raw assembly data may include overrepresented transcripts which would project false distribution patterns for the annotations. Thus, the data obtained from BLAST analysis using the

longest open reading frames through the SwissProt/UniProt database (UniRef90) was reduced by selecting the matches with the highest hit number (E-value >0.01) and excluding the repeating sequences from the assembly file. For clustering the unigenes based on their relation to KEGG pathways, first the UniProt IDs were converted to KEGG IDs for each unigene from https://www.uniprot.org/. Distribution data was obtained by processing the KEGG IDs through the KEGG database from https://www.kegg.jp/.

Prediction of Conserved Genes Involved in Lutein Biosynthesis

To predict the genes related with the lutein synthesis pathway, peptide sequence data for conserved genes responsible for lutein synthesis in plants and alga species were downloaded from the NCBI database (Table 1). Before the comparative analysis, reference databases were prepared using the NCBI database formatting tool Formatdb for each pool of sequences downloaded for the specific genes. The assembled contigs and their complementary counterparts were converted to putative peptide sequences for every three possible frames. The sequences were analyzed through the reference databases using BLAST.

Results and Discussion

Sequencing, Assembly, and Annotation of *T. obliquus* Transcriptome

Sequencing of the pooled cDNA library resulted in a total of 73,109,956.00 raw sequence reads from the mRNA extracts of *T. obliquus*. *De novo* assembly process resulted in 27737 individual contigs (transcripts). Raw sequence reads and related information are uploaded to the NCBI SRA archive and available with the SRX5809206 accession number. Table 2 represents the summary statistics for *de novo* assembly.

Trinotate scan of Swiss-Prot (BlastX/BlastP), PFAM, and EggNOG databases was used for the annotation of unigenes based on similarities of the sequences to known proteins. We were able to predict 14111 transcripts (Table 3) and provided a description for 10563 unigenes among 14111 contigs. Length distributions of the contigs are presented in Figure 1(a).

Table 1. Conserved genes required for lutein biosynthesis in plants and algae

Gene Name	Symbol	EC number
Phytoene synthase	*PSY*	2.5.1.32
Phytoene desaturase	*PDS*	1.3.5.5
15-cis-zeta-carotene isomerase	*ZISO*	5.2.1.12
ζ-carotene desaturase	*ZDS*	1.3.5.6
Carotenoid isomerase	*CRTISO*	5.2.1.13
Lycopene ε-cyclase	*LCYE*	5.5.1.18
Lycopene β-cyclase	*LCYB*	5.5.1.19
Beta-ring hydroxylase	*LUT5*	1.14.-.-
β-Carotene hydroxylase	*CRTZ*	1.14.13.129
Carotene ε-monooxygenase	*LUT1*	1.14.99.45

Table 2. Summary statistics for *Tetradesmus obliquus* transcriptome

Definition	Values
Unigenes	21490
Contigs	27737
Average contig length	557.6
Median contig length	401
Percent GC	61.65
N50	704
N70	472
N90	262

Table 3. Database distributions summary for transcriptome annotated

Database	Number of Annotated Sequences	
	Contigs	*Unigenes*
Swiss-Prot (BlastX)	12547	9359
Swiss-Prot (BlastP)	10863	8171
Pfam	9705	7265
EggNOG	7901	6027
TOTAL	14111	10563

The Venn diagram provided in Figure 1(b) shows the unigene hit distributions among the scanned databases.

Clustering of GO terms resulted in a total of 47076 GO term hits for 7223 unigenes. Annotated unigenes are available through the NCBI TSA archive under GHLU00000000 accession number. In summary, GO terms were distributed as 18720 GO terms (5757 unigenes) under biological process, 14561 GO terms (6163 unigenes) under cellular component and 13795 GO terms (5941 unigenes) under molecular function (Figure 1c and Figure 2).

Among the assembled unigenes, 49.15% of the sequences were associated with the proteins. In addition, at least one GO term was assigned for 72.06% of the annotated unigenes through the Swiss-Prot database.

4282 EggNOG entries were matched to 7223 annotated unigenes. Classification of the unigenes resulted in a total of 5570 orthologous matches for COG, KOG, and NOG orthologous groups. COG analysis showed that 5336 matches were distributed for 4045 unigenes (Figure 3).

KEGG analysis showed that 2969 KO entries were matched for 4771 unigenes. Detailed analysis showed that a total of 9636 hits represented 397 KEGG pathways. The metabolism category showed the highest portion with 51.43% of the total hits while environmental information processing showed the lowest distribution with 8.47% of the total hits (Figure 4). In addition, a total of 1084 enzymes (977 unigenes) were represented with 232 oxidoreductases, 405 transferases, 230 hydrolases, 80 lyases, 54 isomerases, 66 ligases, and 17 translocases.

Prediction of Conserved Genes Involved in Lutein Biosynthesis

We screened assembly data for putative genes conserved for lutein biosynthesis. To predict the genes involved in the lutein synthesis pathway, sequences showing similarities to ten conserved genes responsible for lutein synthesis in other plant and algae species were analyzed. BLAST analysis results revealed that all 10 genes in the lutein biosynthesis pathway were captured through the transcriptome sequencing (Figure 5). Data suggest that the genes involved in lutein synthesis in other plant and algae species are also expressed in *T. obliquus*. The results were confirmed by cross-comparison of query IDs in the annotation report, and a detailed list of match results for the BLAST analysis is provided in Table 4. Our data suggest that *T. obliquus* uses the primary route for lutein biosynthesis as described for other plants and microalgae (J. Kim & DellaPenna, 2006; Sun et al., 2016).

The annotation data can be used as a source for further studies and biotechnological applications related to *T. obliquus*. The unigene sequences can be accessed through the NCBI database anytime and they can be used for testing expression levels of the ten genes related with the lutein biosynthesis. In addition, it is also possible to understand the contribution of each gene during lutein biosynthesis by manipulating these genes via gene knock-out or know-down experiments which would provide the necessary tools to carry out effective modifications to control the lutein biosynthesis.

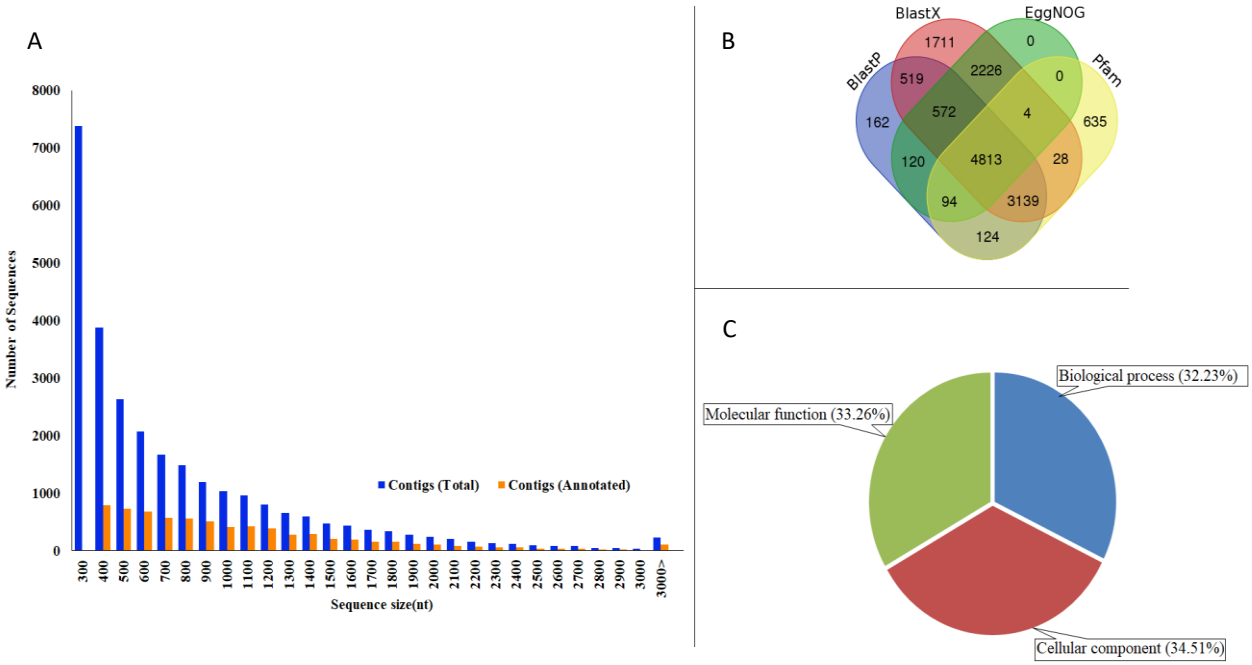

Figure 1. Summary for *Tetradesmus obliquus* transcriptome assembly data. The length distribution was provided before and after filtering of the contigs (A). Annotations were carried out after removing the recursive sequences. Venn distribution of database matches under BlastP, BlastX, EggNOG, and Pfam analysis (B), and Distribution of GO term matches (C) for annotated unigenes are presented.

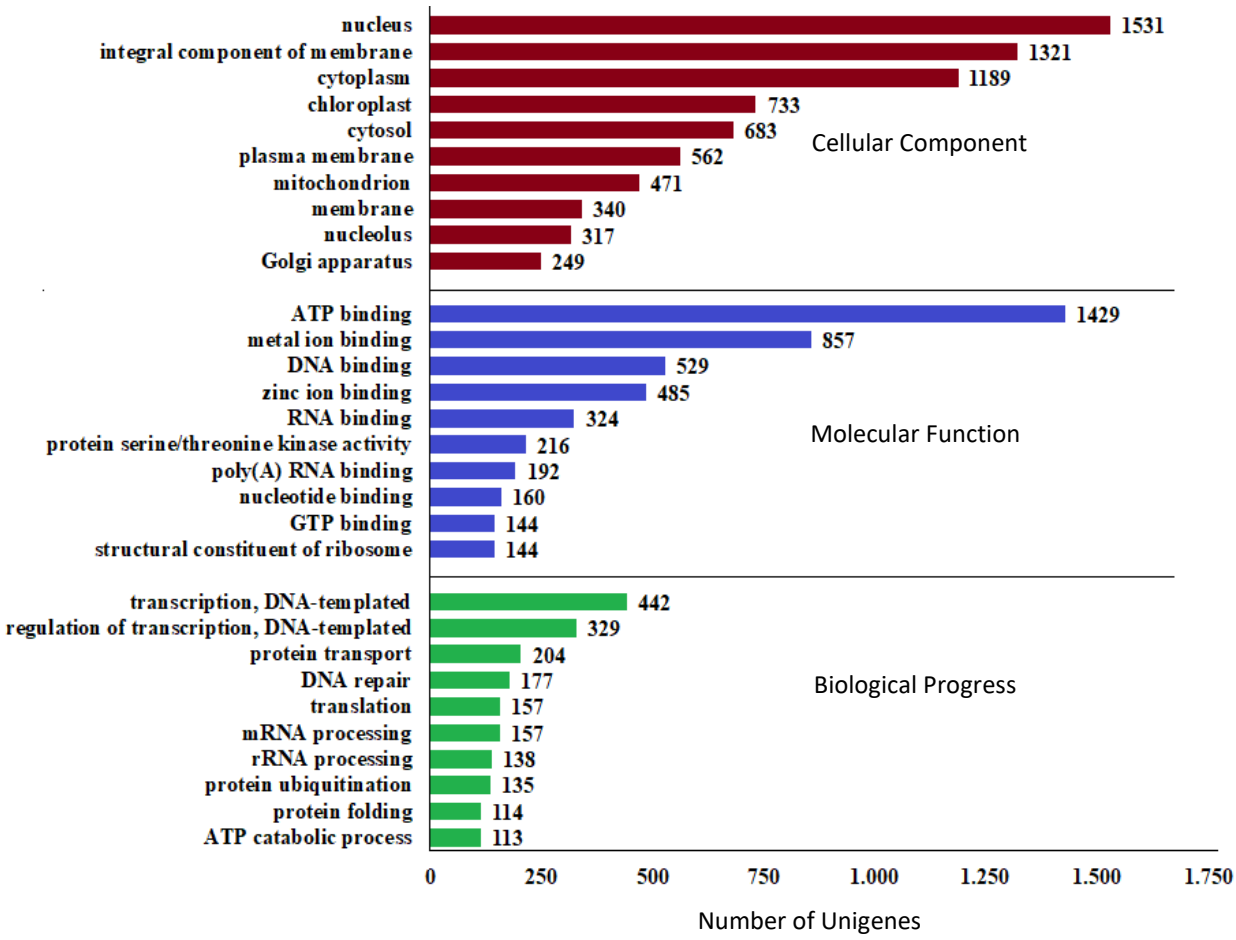

Figure 2. Detailed GO distribution of the unigenes. Distributions were based on the matches of unigenes through the Swiss-Prot database. Distribution of a total of 47076 GO term hits for 7223 unigenes were provided under “Cellular component”, “Molecular Function”, "Biological Process” terms. Detailed list of unigenes is available through the NCBI TSA archive under GHLU00000000 accession number.

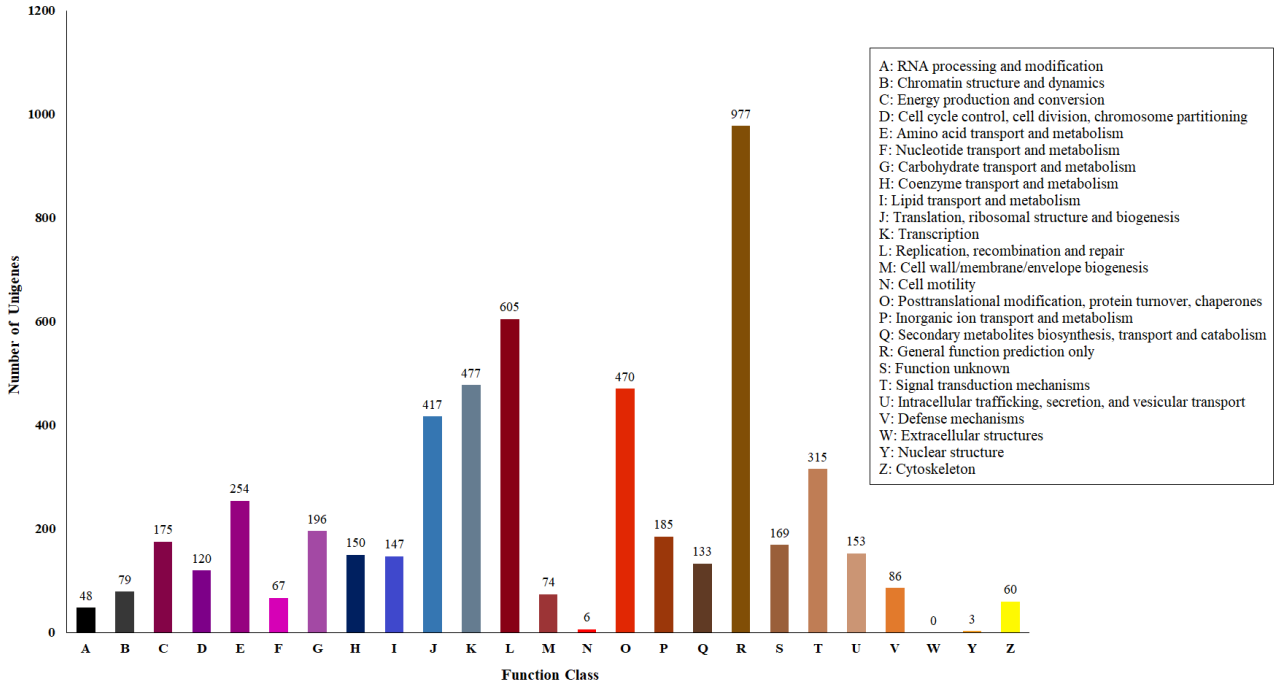

Figure 3. Function-based distribution of Cluster of Orthologous Groups of proteins (COG)for *Tetradesmus obliquus* transcriptome. Data represents the distribution of a total of 5336 matches for 4045 unigenes. eggNOG “*COG.funccat database”* was used for the analysis.

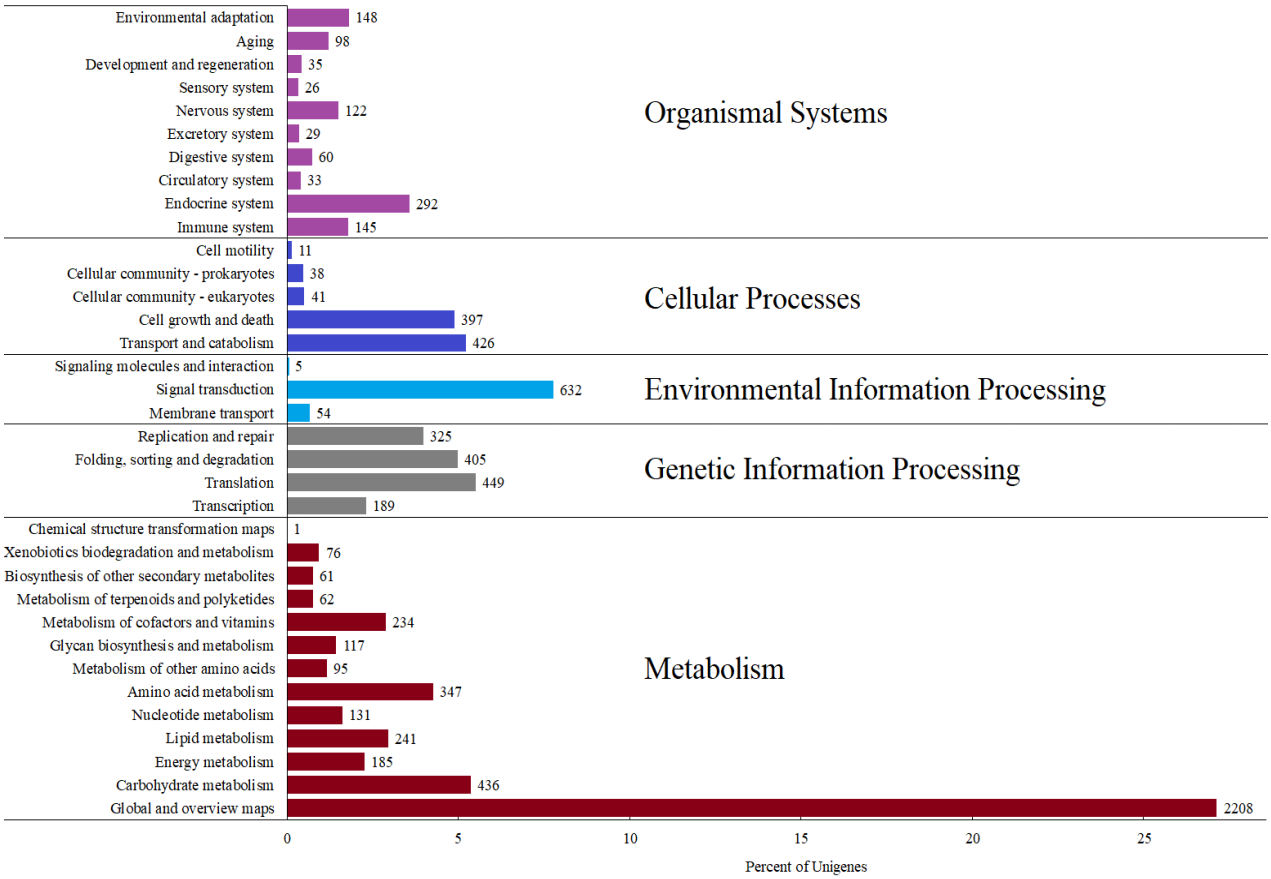

Figure 4. KEGG pathway distribution summary for *Tetradesmus obliquus* transcriptome. The graph shows the distribution of a total of 9636 hits represented 397 KEGG pathways. KEGG IDs were obtained for each unigene converting annotated UniProt IDs to KEGG IDs from https://www.uniprot.org/. KEGG IDs were processed through the KEGG database from https://www.kegg.jp/.

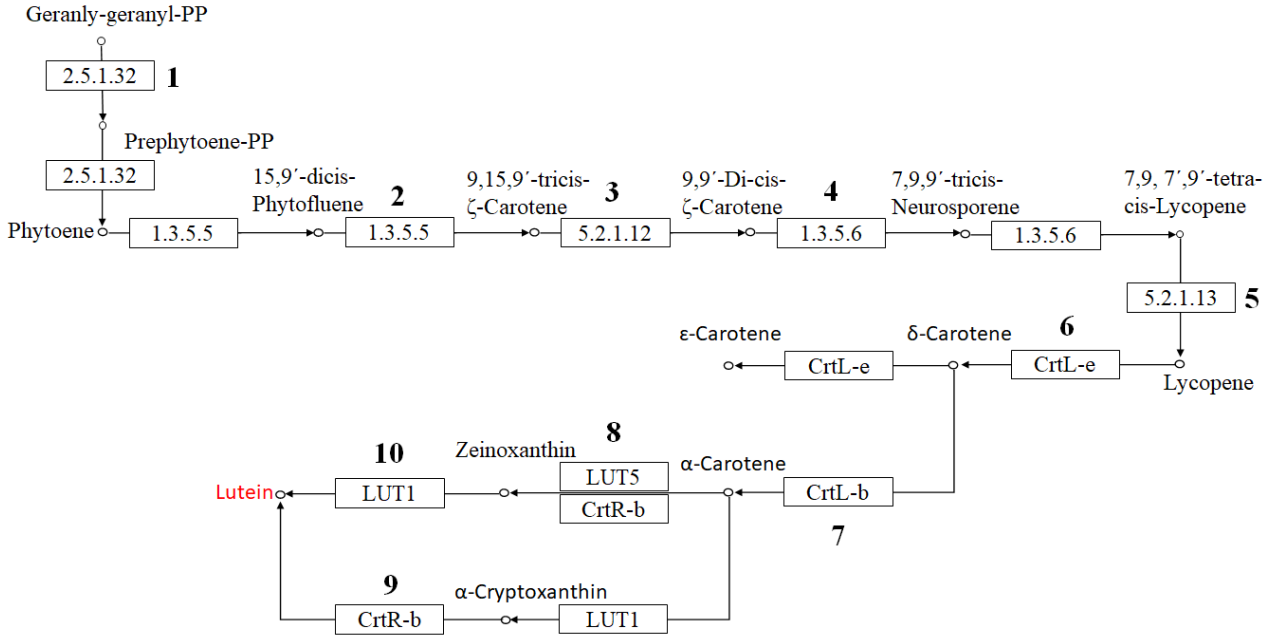

Figure 5. Predicted pathway for lutein biosynthesis in *Tetradesmus obliquus*. The pathway was manually cured based on KEGG-Carotenoid biosynthesis reference pathway (Map id: map00906)

Table 4. BLAST analysis results for the unigenes associated with lutein biosynthesis

Enzyme	GenBank ID	Species name	NCBI GI Accession	Match length	Identity	E-value
Phytoene synthase	GHLU01006609	*C. reinhardtii*	47779181	301	74.75	1,00E-134
Phytoene desaturase	GHLU01001932	*O. tauri*	693501886	541	40.11	1,00E-97
15-cis-zeta-carotene isomerase	GHLU01003784	*G. biloba*	700256946	288	63.19	1,00E-111
Zeta-carotene desaturase	GHLU01004888	*A. protothecoides*	313870540	304	97.7	1,00E-178
Carotenoid isomerase	GHLU01004960	*C. subellipsoidea C-169*	384252368	358	71.51	1,00E-161
Lycopene ε-cyclase	GHLU01003590	*C. reinhardtii*	159476860	407	61.67	1,00E-143
Lycopene β-cyclase	GHLU01002482	*C. zofingiensis*	290454881	373	67.83	1,00E-148
β-ring hydroxylase	GHLU01001479	*Tetraselmis sp, GSL018*	654182978	510	48.04	1,00E-124
β-Carotene hydroxylase	GHLU01001840	*A. protothecoides*	760441531	210	66.19	5,00E-74
Carotene ε-monooxygenase	GHLU01001463	*C. sinensis*	743066329	385	61.56	1,00E-143

Conclusion

T. obliquus is one of the promising species to produce lutein on an industrial scale. In this study, we presented the transcriptome annotation for *T. obliquus*. Raw and annotated sequences are available through the NCBI database. We also presented the sequence information for the genes encoding conserved proteins that have roles in lutein biosynthesis in *T. obliquus*. Data shows that *T. obliquus* is using a similar pathway for lutein biosynthesis suggesting that these gene sequences can be modified to increase the efficiency within *T. obliquus* or other algae species. For instance, taking advantage of modern genetic engineering approaches, the gene(s) can be expressed in other algae with more efficient biomass yield such as *Haematococcus pluvialis*.

To better understand the dynamics of lutein metabolism in *T. obliquus*, expressional evaluations by DEG analysis and qPCR experiments can be carried out for various parameters. Both the industry and academia would benefit from the data gathered from further analysis of the expression data by selecting markers or finding more precise conditions for optimization purposes.

Ethical Statement

No ethical approval was required to finish this study.

Funding Information

This study was funded by Bolu Abant İzzet Baysal University (Grant number: 2015.03.03.819).

Author Contribution

Ömer Can Ünüvar: Sampling, data collection and writing/editing the manuscript.

Ercan Selçuk Ünlü: Conception and design, sampling, data collection and analysis, evaluation of results, and writing/editing/revising the manuscript.

Conflict of Interest

Ömer Can Ünüvar and Ercan Selçuk Ünlü declare that they have no conflict of interest.

References

Ben-Amotz, A., & Avron, M. (1983). Accumulation of metabolites by halotolerant algae and its industrial potential. *Annual Review of Microbiology*, *37*, 95–119. https://doi.org/10.1146/annurev.mi.37.100183.000523

Caprio, F. Di, Altimari, P., Iaquaniello, G., Toro, L., & Pagnanelli, F. (2018). T. obliquus mixotrophic cultivation in treated and untreated olive mill wastewater. *Chemical Engineering Transactions*, *64*. https://doi.org/10.3303/CET1864105

Carreres, B. M., de Jaeger, L., Springer, J., Barbosa, M. J., Breuer, G., van den End, E. J., Kleinegris, D. M. M., Schäffers, I., Wolbert, E. J. H., Zhang, H., Lamers, P. P., Draaisma, R. B., Martins dos Santos, V. A. P., Wijffels, R. H., Eggink, G., Schaap, P. J., & Martens, D. E. (2017). Draft genome sequence of the oleaginous green alga Tetradesmus obliquus UTEX 393. *Genome Announcements*, *5*(3). https://doi.org/10.1128/genomeA.01449-16

Del Campo, J. A., Rodríguez, H., Moreno, J., Vargas, M. Á., Rivas, J., & Guerrero, M. G. (2001). Lutein production by Muriellopsis sp. in an outdoor tubular photobioreactor. *Journal of Biotechnology*, *85*(3), 289–295. https://doi.org/10.1016/S0168-1656(00)00380-1

Dwyer, J. H., Navab, M., Dwyer, K. M., Hassan, K., Sun, P., Shircore, A., Hama-Levy, S., Hough, G., Wang, X., Drake, T., Merz, C. N., & Fogelman, A. M. (2001). Oxygenated carotenoid lutein and progression of early atherosclerosis: the Los Angeles atherosclerosis study. *Circulation*, *103*(24), 2922–2927. http://www.ncbi.nlm.nih.gov/pubmed/11413081

Geng, L., Zhang, J., Qin, B., & Yang, Z. (2014). No difference in colony formation of Scenedesmus obliquus exposed to lakes with different nutrient levels. *Biochemical Systematics and Ecology*, *57*, 178–182. https://doi.org/10.1016/j.bse.2014.08.024

Granado, F., Olmedilla, B., & Blanco, I. (2007). Nutritional and clinical relevance of lutein in human health. *British Journal of Nutrition*, *90*(03), 487. https://doi.org/10.1079/BJN2003927

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., William, T., Dewey, C. N., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

Heber, D., & Lu, Q. Y. (2002). Overview of mechanisms of action of lycopene. *Experimental Biology and Medicine*, *227*(10), 920–923. http://www.scopus.com/inward/record.url?eid=2-s2.0-0036846616&partnerID=tZOtx3y1

Ho, S.-H., Chan, M.-C., Liu, C.-C., Chen, C.-Y., Lee, W.-L., Lee, D.-J., & Chang, J.-S. (2014). Enhancing lutein productivity of an indigenous microalga Scenedesmus obliquus FSP-3 using light-related strategies. *Bioresource Technology*, *152*, 275–282. https://doi.org/10.1016/j.biortech.2013.11.031

Jiang, H., Lei, R., Ding, S.-W., & Zhu, S. (2014). Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics*, *15*(1), 182. https://doi.org/10.1186/1471-2105-15-182

Kim, J., & DellaPenna, D. (2006). Defining the primary route for lutein synthesis in plants: The role of Arabidopsis carotenoid β-ring hydroxylase CYP97A3. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(9), 3474–3479. https://doi.org/10.1073/pnas.0511207103

Kim, Y. S., Yi, J., Do, J. M., Chang, J., & Yoon, H. S. (2019). Characterization of fatty acid components from Tetradesmus obliquus KNUA019 (Chlorophyta, Scenedesmaceae) for a resource of biofuel production. *Revista Brasileira de Botanica*, *42*(3). https://doi.org/10.1007/s40415-019-00556-9

Lagarde, D., Beuf, L., & Vermaas, W. (2000). Increased Production of Zeaxanthin and Other Pigments by Application of Genetic Engineering Techniques to Synechocystis sp. Strain PCC 6803. *Applied and Environmental Microbiology*, *66*(1), 64–72. https://doi.org/10.1128/AEM.66.1.64-72.2000

Mann, V., Harker, M., Pecker, I., & Hirschberg, J. (2000). Metabolic engineering of astaxanthin production in tobacco flowers. *Nature Biotechnology*, *18*(8), 888–892. https://doi.org/10.1038/78515

Md Nadzir, S., Yusof, N., Nordin, N., Abdullah, H., & Kamari, A. (2021). Optimisation of carbohydrate, lipid and biomass productivity in Tetradesmus obliquus using response surface methodology. *Biofuels*, *12*(7). https://doi.org/10.1080/17597269.2018.1542568

Mishra, A., Medhi, K., Malaviya, P., & Thakur, I. S. (2019). Omics approaches for microalgal applications: Prospects and challenges. In *Bioresource Technology* (Vol. 291, p. 121890). Elsevier Ltd. https://doi.org/10.1016/j.biortech.2019.121890

Pons, E., Alquézar, B., Rodríguez, A., Martorell, P., Genovés, S., Ramón, D., Rodrigo, M. J., Zacarías, L., & Peña, L. (2014). Metabolic engineering of β-carotene in orange fruit increases its in vivo antioxidant properties. *Plant Biotechnology Journal*, *12*(1), 17–27. https://doi.org/10.1111/pbi.12112

Sánchez, J. F., Fernández-Sevilla, J. M., Acién, F. G., Cerón, M. C., Pérez-Parra, J., & Molina-Grima, E. (2008). Biomass and lutein productivity of Scenedesmus almeriensis: influence of irradiance, dilution rate and temperature. *Applied Microbiology and Biotechnology*, *79*(5), 719–729. https://doi.org/10.1007/s00253-008-1494-2

Sánchez, J. F., Fernández, J. M., Acién, F. G., Rueda, A., Pérez-Parra, J., & Molina, E. (2008). Influence of culture conditions on the productivity and lutein content of the new strain Scenedesmus almeriensis. *Process Biochemistry*, *43*(4), 398–405. https://doi.org/10.1016/j.procbio.2008.01.004

Sun, Z., Li, T., Zhou, Z. G., & Jiang, Y. (2016). Microalgae as a source of lutein: Chemistry, biosynthesis, and carotenogenesis. *Advances in Biochemical Engineering/ Biotechnology*, *153*, 37–58. https://doi.org/10.1007/10_2015_331

Wiltshire, K. H., Boersma, M., Möller, A., & Buhtz, H. (2000). *Extraction of pigments and fatty acids from the green alga Scenedesmus obliquus (Chlorophyceae)*. *1997*, 119–126.

Wu, B., Suo, F., Lei, W., & Gu, L. (2014). Comprehensive analysis of alternative splicing in Digitalis purpurea by strand-specific RNA-Seq. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0106001

Yang, W., Ying, D., & Lau, Y. L. (2009). In-depth cDNA Library Sequencing Provides Quantitative Gene Expression Profiling in Cancer Biomarker Discovery. *Genomics, Proteomics and Bioinformatics*, *7*(1–2), 1–12. https://doi.org/10.1016/S1672-0229(08)60028-5