CHEMICAL ENGINEERING TRANSACTIONS
VOL. 43, 2015
A publication of AIDIC
The Italian Association of Chemical Engineering
Online at www.aidic.it/cet
Chief Editors:SauroPierucci, JiříJ. Klemeš

ISBN 978-88-95608-34-1; **ISSN** 2283-9216
DOI: 10.3303/CET1543200

Comparative Genomic and Transcriptomic Analyses of Microalgae as Biofactories

Barbara Menin[a,b], Paola Branduardi[a], Paola Mariani*[b]

[a] Dipartimento di Biotecnologie e Bioscienze dell'Università degli Studi di Milano Bicocca, P.zza della Scienza 2, 20126 Milano, Italy
[b] Sezione Genomica Funzionale, Parco Tecnologico Padano , Via A. Einstein, Loc. Cascina Codazza, 26900 Lodi, Italy
paola.mariani@tecnoparco.org

Microalgae are unicellular photo- and heterotrophic organisms that, being efficient producers of a rich and complex biomass, have attracted global interest as biofactories for the production of high-value compounds (e.g. PUFAs, proteins and pigments) that can find application in nutraceutical, pharmaceutical and food industries. Microalgae have also been extensively studied in the last two decades, as potential producers of biofuels. However, based on current technologies, this production process is not economically sustainable yet. In this frame, the recent advances in genomic and transcriptomic characterization of microalgae can represent the key to understand the metabolic mechanisms underlying the growth regulation and the biosynthesis of those valuable metabolites, in order to improve the use of microalgae as biofactories.

1. Introduction

Energy crisis and freshwater deficit are two of the biggest challenges the world is facing, due to the constant increasing of world population and the deterioration of natural resources (Markou and Nerantzis, 2013). In the last two decades, microalgae have attracted a growing interest as biofactories for the production of high value compounds, for their potential to become an alternative source of biofuels, and their ability to remediate effluents. Microalgae is a large and heterogeneous group of eukaryotic and prokaryotic (Cyanobacteria) unicellular organisms that are able to growth under photo-autotrophic, heterotrophic and/or mixotrophic conditions. The capability of autotrophy is a direct way to convert light energy and inorganic nutrient, such as carbon dioxide, nitrogen and phosphorous, into valuable compounds (e.g. carbohydrates, proteins, lipids, pigments). Microalgae are ubiquitous in nature, since they can be found both in freshwater and marine systems. Compared to terrestrial crops, they have many advantages in terms of productivity and resource management (Georgianna et al., 2012). Despite several studies on the metabolic content, growing conditions and potential industrial application of microalgae, the understanding of main algal metabolic pathways and in-depth knowledge on their regulation is still lacking. To obtain this information, an essential requirement is the whole sequencing of genome and transcriptome of the microalgae species chosen for biomass production. Such a strategy will open the door to efficient metabolic regulation and genetic engineering of microalgae. This short review describes the emerging concept of microalgae as biofactories and summarizes the advance of genetic and transcriptomic characterization of microalgae as tool to improve their use as integrated biorefinery.

2. Microalgae as biofactories

Microalgae, as main constituent of the phytoplankton, represent the major agents for marine primary production, playing a key-role in the global food chain. Microalgae are extremely productive organisms and

Please cite this article as: Menin B., Branduardi P., Mariani P., 2015, Comparative genomic and transcriptomic analyses of microalgae as biofactories, Chemical Engineering Transactions, 43, 1195-1200 DOI: 10.3303/CET1543200

their biomass is rich and complex, feature that led microalgae to be considered as a promising biofactory for the production of commodities, chemicals and biofuels (Cuellar-Bermudez et al., 2014).
Microalgae biomass approximately contains 30-40% carbohydrates, 30-55% proteins, 20-30% lipids and 5-10% other compounds (mainly ash), but the overall composition and metabolism may significantly differ depending on the species and the growing conditions. Each fraction can be used for industrial purposes in different production processes, that have been the subject of numerous studies. The sugar-based and starch-based feedstock can be converted into fermentable sugars for bioethanol production via microbial fermentation (Chen et al., 2013). Neutral lipids may be used for biodiesel production, through the reaction of trasterification (Chisti 2007; Mata et al., 2010), while long-chain polyunsaturated fatty acids (PUFAs) have important nutraceutical applications (de Jesus Raposo et al., 2013). Enzymatic hydrolysis of proteins can lead to the production of free aminoacid concentrates (Romero García et al., 2012) and, finally, the great amount of residual biomass can be used for the generation of biogas (biomethane) by anaerobic digestion (Ramos-Suárez and Carreras, 2014).
The integration of the main production processes, discussed above, in a logic of microalgae-based biorefinery, may enhance the economic sustainability of the system. However, it is necessary to pay attention to their consequential succession, because each single extraction step implies losing part of the remaining fractions of the biomass . At present, there is no a definitive answer regarding the best strategy to be applied. Possible alternatives are to be considered according to the characteristics of the initial algal biomass, the final desired products and the plant structure. Critical to improve the economic feasibility of microalgae as biofactories is the advancement in genomics and transcriptomics. Novel know how in these fields will clarify the metabolic pathways and the factors that regulate the production of compounds of interest. Genetic engineering applied to microalgae biomass will allow to enhance both the efficiency of the production system and the quality and yield of the final products (Radakovits *et al*., 2010).

2.1 High-value products

Currently, microalgae are one of the most important feed sources in aquaculture and to a less extent in animal feeding. Over the recent years they have been also considered as valuable candidates for functional or health food. Some microalgae strains particularly rich in proteins, such as the green algae *Chlorella* sp. and the cyanobacteria *Spirulina* sp., are already present in the market for human nutrition (Koller et al., 2014). Other microalgae strains, instead, are cultivated for the extraction of high-value commercial products, especially PUFAs, pigments and vitamins.
Among PUFAs, eicosapentaenoic acid (EPA, 20:5) and docosahexaenoic acid (DHA, 22:6) are the most important bioactive compounds from the nutritional point of view, because humans need them but are not able to synthesize them. EPA and DHA have been demonstrated to also possess relevant therapeutical properties: they are anti-inflammatory agents, can reduce hypertension and problems related to cardiovascular and arthritis diseases (de Jesus Raposo et al., 2013). DHA is mostly extracted and commercialized from *Schizochytrium* sp and *Crypthecodinium cohnii* cultures, whereas the economic feasibility of the production of EPA is still under investigation, because of high purification costs. A possible manufacturing process for EPA have been proposed through monoseptic cultivation of the diatom *Phaedactylum tricornutum* (Molina Grima et al., 2003). Another species that recently attracted interest as possible PUFAs producer is the haptophyta *Pavlova lutheri* (Shah et al., 2014).
Pigments and vitamins are the products with the highest value on the market and thereby with the best potential for commercial success. They also possess significant pharmacological properties. The most common carotenoid, β-carotene, mainly produced by *Dunaliella salina*, is a strong antioxidant agent, able to protect protein and DNA from free radical damage proteins, lipids and DNA from oxidative damage caused by free radicals. Furthermore, this organism posses radical scavenging capacities and metal-chelating activity. Another valuable antioxidant compound is the astaxanthin, produced by *Haematococcus pluvialis*, which is a xantophills able to increase the immune response and reduce atherosclerosis and age-macular degeneration (de Jesus Raposo et al., 2013).

2.2 Cultivation under nutrient stress condition

The metabolism of microalgae is extremely dynamic, they are able to drastically change the composition of carbohydrates, proteins and lipids, under nutrient stress conditions. Nitrogen starvation is one of the most studied and used strategies to increase the accumulation of lipids and starch in microalgae cultures, since under nitrogen-depletion conditions many microalgal strains can transform proteins or peptides into lipids or carbohydrates as energy reserve components (Markou and Nerantzis, 2013).

Lipid accumulation is drastically increased and lipid profile remodeling in response to nitrogen deprivation. A detailed investigation of changes in lipid composition in response to nitrogen deprivation in *Chlorella* sp. and *Nannochloropsis* sp. have been reported by Martin et al. (2014). They observed that N-deprivation leads to a considerable accumulation of triacylglycerides (TAGs) in both species (340% in *Chlorella* sp. and 130% in *Nannochloropsis* sp.) and this increment is associated with a decrease in total phosphoglycerolipid and glycolipid content. Most important, it was also observed that the acyl composition of TAGs in *Nannochloropsis* sp. remains relatively constant, whereas Chlorella sp. shows greater variability following N-deprivation.
Also total carbohydrates content have been reported to significantly increase in response to nitrogen-depletion in many microalgal species, such as *C. vulgaris* (with a total carbohydrates content that reaches 55% of dry cell weight) and *Tetraselmis suecica*, whose cellular carbohydrate content raises from 10 to 57% (Chen et al., 2013).
Another nutrient stress condition exploited to increase fatty acids and carbohydrates content is the phosphate-depletion, which have been proved to be effective both in green algae (El-Kassas, 2013) and diatoms (Valenzuela et al., 2012). Since diatoms possess a characteristic cell wall made of silica, silicon starvation is also used to enhance lipids production in these microalgae (Trentacoste et al., 2013).
Several transcriptomic studies have been carried out in different algal species to investigate genes differentially expressed under nutrient limitation and therefore candidate to be involved in metabolites accumulation. These studies will eventually highlight the key enzymes involved in the biosynthetic pathway of interest. Changing in global gene expression (Valenzuela et al., 2012) and protein synthesis (Yang et al., 2014) have been extensively characterized during enhanced lipid production as a consequence of nutrient starvation in the diatoms *P. tricornutum*.

3. Genomic and transcriptomic analysis: a tool to improve microalgal biofactories

Genomic (DNA-sequencing) and transcriptomic (RNA-sequencing) studies play a crucial role in quantitative and qualitative improvement of microalgal biomass. The sequence analysis afford useful information about the evolutionary history of the different microalgae groups, thus providing scientific hints regarding the role plaid by specific genes and gene networks. More interestingly, whole genome sequence information can contribute to our understanding of the molecular mechanisms that microalgae use to adapt to environmental changes, as well as unlocking the potential to develop new and economically important products and technologies (Hannon et al., 2010). Genomic and transcriptomic information can be used to assess metabolic pathways and to perform more focused genetic engineering approaches, such as up-regulation or knock out of genes involved in the pathways of interest. Furthermore, a good annotation of the whole genome, alongside to a comparative transcriptomic approach, not only allows to gain insight into the metabolic pathways and their key-enzymes, but also to identify regulatory factors and promoters of gene expression.

3.1 Available genomic and transcriptomic data

The advent of next-generation sequencing (NGS) technologies has led to a great transformation in the way scientists extract genetic information from biological systems, providing a tremendous potential to genomic and transcriptomic studies. Timing and analyses costs have been significantly reduced over the last years, making relatively easy the whole sequencing of small size genomes.
As result, recently, an increasing number of genome and transcriptome data of different microalgae species has been released. At present the Department of Energy Joint Genome Institute (JGI), aiming at supporting genomic studies related to clean energy generation and environmental characterization, has more than 60 algae ongoing genome projects (http://www.algaeu.com/strains-of-algae-publications.html). The majority of these projects are focused on *Chlorella* sp. and on the model organism *Chlamidomonas reinhardtii*, due to their economic importance as biofactories. This number will undoubtedly increase in the near future, mainly due to the global growing interest in microalgae production.
Cyanobacteria have been studied much more thoroughly than the eukaryotic microalgae. Recently, 54 phylogenetically and phenotypically diverse cyanobacterial strains have been sequenced, 29 complete genomes and 25 assembled to draft genome status, and at present more than 120 genomic sequences of cyanobacteria strains is available (Shih et al., 2013).
Whole genome sequencing and *de novo* assembly can be difficult to carry out due to lack of reference genomes. Therefore, many research groups have chosen to focus only on the expressed genes by means of RNA sequencing performed by high throughput technologies -NGS transcriptomic analysis- even if the major drawback of this strategy is the loss of information related to non-transcribed sequences. A good example of this approach is the recent accomplished Marine Microbial Eukaryote Transcriptome Sequencing Project

(MMETS) (Kelling et al., 2014). This project was the first large-scale attempt to get reference data regarding marine eukaryotic plankton. The complete study created over 650 assembled, functionally annotated transcriptomes of more than 300 species. Transcriptomes have been generated *de novo* using the Illumina HiSeq 2000 platform and an open database is now available (http://marinemicroeukaryotes.org/project_organisms). Main purposes of MMETS were: i) to clarify the evolutionary relationships among marine microbial eukaryotic clades and within the overall eukaryotic tree of life and ii) to explore the physiology of diverse marine microbial eukaryotes. The microbial algal trancriptomes produced in this project will be used as reference in further comparative studies or as starting point for a deeper exploration of the microalgal metabolism.
However, despite the great advances in bacteria genetic engineering, a stable transformation system, specific for the majority of eukariotic microalgae, still needs to be developed. Currently, only about 20 different microalgal species can be transformed. Genetic engineering techniques applied to microalgae and a summary of bioproducts manufactured in microalgae for molecular farming were comprehensively reviewed elsewhere (Hannon et al., 2010, Rasala et al., 2014 and references therein). However, it seems likely that the rapid advances in genomic and transcriptomic studies will allow in the near future to develop targeted approaches for genetic improvement, that could really make feasible the advent of eukaryotic microalgae as third generation biofuel biofactories (Radakovits *et al*., 2010).

3.2 Improving traits for production

The majority of research is currently focused on the biosynthetic pathway of fatty acids in order to increase their accumulation. Further in-depth transcriptomic and proteomic studies are needed to identify and characterize the key genes involved in the lipid biosynthetic pathway. Characterization of these genes in different terrestrial crops may help to identify homologous proteins in microalgae. The first biosynthesis step in many organisms is the conversion of acetyl-coenzyme A (CoA) to malonyl-CoA, in a two step reaction catalyzed by acetyl-CoA carboxylase (ACCase) (Ohlrogge and Browse , 1995). Therefore, it might be possible to enhance fatty acid synthesis through the overexpression of ACCase. However, the main reason for the slow progress in genetically manipulating ACCase is the complexity of this enzymes. Recently, the full length ORF (Open Reading Frame) of ACCase from *P. tricornutum* was isolated. The ORF is long 7271 bp and encodes for five exons and four introns and so it is hard to be genetically modified. The first exon, containing the critical domain, have been successfully cloned and characterized in *E. coli* (Xie et al., 2013). Since overexpression of ACCase is still tricky, other possible strategies to improve lipid content have been essayed. An option is to block the metabolic pathways that lead to the accumulation of other energy-rich storage compounds, such as starch. Starch-deficient strains of *C. reinhardtii* and *Chlorella* have been shown to possess elevated PUFAs content (Cuellar-Bermudez et al., 2014).
In place of increasing the accumulation of total lipids, a valuable alternative strategy is represented by the quality improvement of specific lipids suitable as biodiesel feedstock, mainly TAGs. Ideal TAGs for biodiesel production should have a chain length of 12:0 and 14:0, whereas several microalgae accumulates mostly 16:1, 16:0, and 18:1 TAGs. It is therefore possible to manipulate the enzymes that determine the length of the fatty acids, which are the acyl-ACP thioesterases, to enhance the conversion of lipids into the desired shape. For example, the heterologous expression of two thioesterases, biased towards the production of lauric (C12:0) and myristic acid (C14:0), have been reported to cause an increased accumulation of shorter chain length fatty acids in *P. tricornutum*. (Radakovits et al., 2011). Other two examples of genetic engineering for the improvement of *P tricornutum* fatty acids have been recently published. Firstly, Hamilton and coworkers (2014) obtained a transgenic strain of *P. tricornutum* with enhanced DHA accumulation, this was achieved by the heterologous expression of the Δ5-elongase from the picoalga *Ostreococcus tauri* to augment the endogenous fatty acid biosynthetic pathway. Shortly after, the overexpression of *P. tricornutum* malic enzyme in modified *P. tricornutum* cells have been shown to significantly increase (2.5 folds) the total lipid content (Xue et al., 2015). These results indicate that metabolic engineering strategies are a real powerful tool for overproduction of biofuel compounds.

4. Biorefill project

The BIOREFILL - BIO-REFinery Integrated Lombardy Labs project, is an integrated research project aiming at connecting resources and expertise at both industrial and academic level to implement a biorefinery production system in the Lombardy Region.
At Parco Tecnologico Padano (PTP), in collaboration with Gruppo Ricicla (University of Milan) and with the Department of Biotechnology and Biosciences (University of Milano-Bicocca) both partners in the project, the

molecular characterization of the genetic mechanisms accounting for the production potential of four microalgae species: *Arthrospira platensis* (spirulina), *Pavlova lutherii*, *Phaedactilum tricornutum*, and *Chlorella spp.* is currently in progress. Whole genome *de novo* sequencing and functional annotation of *Pavlova lutherii* and an unknown *Chlorella spp*. strain, isolated from local livestock farm effluents, is carried out using the Illumina Miseq available at the PTP Genomic Platform (PGP).Side by side, the comparative transcriptomic analysis of the four species grown under standard conditions and in nitrogen deprivation (ND) is ongoing. RNA sequencing is performed by the Illumina HiSeq2000 at the PGP. Global transcriptional profiling of microalgal cells grown in standard medium and in response to ND will allow researchers to identify the gene regulatory networks involved in the metabolic pathways leading to the synthesis of TAGs and PUFAs, and adaptation and survival in presence of an environmental stress. The latter, together with the comparative transcriptomic data produced in standard growth condition, will allow gaining insight into the complex molecular mechanisms underlying the biosynthetic pathways involved in growth regulation and the storage of metabolites of interest.
In light of these results, the molecular data will be associated with the functional data and used to select suitable microalgal species and exploit their high added value products. In the logic of biorefinery and activities integration, the best candidate will be grown on nutrients media recovered from anaerobic digestion slurries. The final aim is to set up a low cost cultivation system capable of reducing production costs of microalgal biomass maintaining the ability to produce high-value products for nutraceutical and farmaceutical applications at industrial scale.

5. Conclusions

The identification of microalgal species suitable for biomass production in a logic of biorafinery is the starting point to achieve success in the entire production process. Although it may be difficult to find an “optimal wild type”, which combines rapid growth and high productivity rates of valuable metabolites (especially lipids and pigments) and which is resistant to infections and ease to be harvested, genomic and transcriptomic approaches make possible to work at the selection and of a candidate strain with the desired characteristics and its improvement. These molecular approaches are, indeed, powerful tools that can be used to characterize the metabolic pathways of interest in order to improve the synthesis of valuable metabolites from the qualitative and quantitative point of view. Finally, optimizing the growth rate and the maximum growth density at pilot scale, instead of laboratory condition, is the key step to reach economic viability and remains an open challenge for the industry.

Acknowledgements

The research activity is funded by Regione Lombardia and Fondazione Cariplo.

References

Chen C.Y, Zhao X.Q., Yen H.W., Ho S.H., Cheng C.L., Lee D.J., Bai F.W., Chang J.S., 2013, Microalgae-based carbohydrates for biofuel production, Biochemical Engineering Journal, 78, 1-10.

Chisti Y., 2007, Biodisel from microalgae, Biotechnology Advances, 25, 294-306.

Cuellar-Bermudez S.P., Aguilar-Hernandez I., Cardenas-Chavez D.L., Ornelas-Soto N., Romero-Ogawa M.A., Parra-Saldivar R., 2014, Extraction and purification of high-value metabolites from microalgae: essential lipids, astaxanthin and phycobiliproteins, Microbial Biotechnology, DOI: 10.1111/1751-7915.12167.

De Jesus Raposo M.F., de Moiras R.M., de Moiras A.M, 2013, Health applications of bioactive compounds from marine microalgae, Life Science, 93, 479-486.

El-Kassas H.Y., 2013, Growth and fatty acid profile of the marine microalga *Picochlorum Sp*. grown under nutrient stress conditions.,The Egyptian Journal of Aquatic Research, 39 (4): 233–239.

Georgianna D.R., Mayfield S.P., 2012, Exploiting diversity and synthetic biology for the production of algal biofuels. Nature, 488, 329–335.

Hamilton M.L., Haslam R.P., Napier J.A., Sayanova O., 2014, Metabolic engineering of *Phaeodactylum tricornutum* for the enhanced accumulation of omega-3 long chain polyunsaturated fatty acids, Metab. Eng., 22, 3–9.

Hannon M., Gimpel J.,Tran M., Rasala B., Mayfield S., 2010, Biofuels from algae: challenges and potential, Biofuels, 1, (5): 763–784.

Kelling P.J. *et al*., 2014, The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing, Plos Biology, 12(6), e1001889.

Koller M., Muhr A., Braunegg G., 2014, Microalgae as versatile cellular factories for valued products, Algal Research, 6:52-63.

Mata T.M., Martins A.A., Caetano N.S., 2010, Microalgae for biodiesel production and other applications: A review, Renewable and Sustainable Energy Reviews, 14 (1): 217-232.

Markou G., Nerantzis E., 2013, Microalgae for high-value compounds and biofuels production: A review with focus on cultivation under stress conditions, Biotechnology Advances, 31, 1532-1542.

Martin G.J.O., *et al*., 2014, Lipid Profile Remodeling in Response to Nitrogen Deprivation in the Microalgae Chlorella sp. (Trebouxiophyceae) and Nannochloropsis sp. (Eustigmatophyceae), PlosOne, 9 (8):1-10.

Molina Grima E., Belarbi E.H., Acién Fernández F.G., Robles Medina A., Chisti Y., 2003, Recovery of microalgal biomass and metabolites: process options and economics. Biotech. Advances, 20, 491-515.

Ohlrogge J. and Browse J., 1995, Lipid biosynthesis. Plant Cell,7, 957–970.

Radakovits R., Jinkerson R.E., Darzins A., Posewits M.C., 2010, Genetic engineering of algae for enhanced biofuel production. Eukaryotic Cell, 486-501.

Radakovits R., Eduafo P.M., Posewitz M.C., 2011, Genetic engineering of fatty acid chain length in *Phaeodactylum tricornutum.* Metab. Eng., 13, 89–95.

Ramos-Suárez J.L., Carreras N., 2014, Use of microalgae residues for biogas production, Chemical Engeneering Jornal, 242, 86-95.

Rasala B.A, Mayfield S.P., 2014, Photosynthetic biomanufacturing in green algae; production of recombinant proteins for industrial, nutritional, and medical uses. Photosynthesis Research, DOI 10.1007/s11120-014-9994-7.

Romero García J.M., Acién Fernández F.G., Fernández Sevilla J.M., 2012, Development of a process for the production of L-amino-acids concentrates from microalgae by enzymatic hydrolysis. Bioresources Technology, 112, 164-170.

Shah S.M.U., Ahmad A., Othman M.F., Abdullah M.A., 2014, Enhancement of lipid content in *Isochrysis galbana* and *Pavlova lutheri* using palm oil mill effluent as an alternative medium. Chemical Engeneering Transactions, 37, 733-738.

Shih P.M., Wu D.Y., Latifi A., Axen S.D., Fewer D.P., Talla E., Calteau A., Cai F., de Marsac N.T., Rippka R., 2013, Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. Proc Natl Acad Sci, 110:1053-1058.

Trentacoste E.M., Shrestha R.P., Smith S.R., Glé C., Hartmann A.C., Hildebrand M., Gerwick W.H, 2013, Metabolic engineering of lipid catabolism increases microalgal lipid accumulation without compromising growth.Proc Natl Acad Sci U S A, 110(49):19748-53.

Valenzuela J., Mazurie A., Carlson R., Gerlach R., Cooksey K.E., Peyton B.M., Fields M.W., 2012, Potential role of multiple carbon fixation pathways during lipid accumulation in *Phaeodactylum tricornutum*, Biotechnology for Biofuels, 5(1):40.

Xie W.H., Pang F., Niu Y.F., Zhang M.H., Yang W.D., Liu J.S., Yan D.G., Li H.Y., 2013, Functional characterization of an ACCase subunit from the diatom Phaeodactylum tricornutum expressed in Escherichia coli.Biotechnol Appl Biochem., 60 (3):330-5.

Xue J., Niu Y.F., Huang T., Yang W.D., Liu J.S., Li H.Y., 2015, Genetic improvement of the microalga *Phaeodactylum tricornutum* for boosting neutral lipid accumulation. Metabolic Engineering, 27:1-9.

Yang Z.K., Ma H.Y., Zheng J.W., Yang W.D., Liu J.S., Li H.Y., 2014, Proteomics to reveal metabolic network shifts towards lipid accumulation following nitrogen deprivation in the diatom *Phaeodactylum tricornutum*, Journal of Applied Phycology, 26 (1):73-82.