



**International Ph.D. Program in Environmental
Sciences and Engineering
XXVII cycle**

University
of Cagliari

Universidade
de Aveiro

University
of Pushchino

Ecole Nationale de
l'Industrie
Minerale

**GENETIC AND PHYLOGENETIC
CHARACTERIZATION OF
MICROALGAE STRAINS IN VIEW
OF THEIR EXPLOITATION FOR
CO₂ CAPTURE AND BIOFUEL
PRODUCTION**

Ph.D candidate: Cristina Costelli

**Advisors: Prof. Dr. Giacomo Cao, Prof. Roberto Orrù
Dr. Alessandro Concas**



Università degli Studi di Cagliari
Centro Interdipartimentale di Ingegneria e Scienze Ambientali (C.I.N.S.A.)
via San Giorgio, 12 – 09124 Cagliari
Academic Year 2013 – 2014

*Completing this thesis is a great goal
that I would never have achieved without
the encouragement, support and love
of my mother and my sister.
I dedicate this thesis to you.*

Acknowledgements

Many people supported this research, and I wish to thank Prof. Giacomo Cao without whom this research would have not been possible.

I gratefully acknowledge the Sardinia Regional Government for the financial support through the PhD scholarship (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2007–2013 - Axis IV Human Resources, Objective 1.3, Line of Activity 1.3.1).

A special acknowledgment to Ing. Alessandro Concas, for his useful comments, remarks and engagement through the learning process of this PhD.

I am grateful to all my colleagues for their advice, help and friendship.

Thanks to those who shared with me the joys and sufferings, and with their presence are helpful to color my life. Thank you for your love.

Last but not least, I want to thank my lovely mother and my fantastic sister, who are the reason I am here. They have always understood and encouraged me to realize my dreams, and this goal is also yours.

Table of contents

LIST OF TABLES	xi
LIST OF FIGURES	xii
LIST OF ABBREVIATIONS	xiv
CHAPTER 1. Introduction	11
1.1 Microalgae as a biofuel feedstock	12
1.2 Microalgae: importance, photosynthesis and lipid production	15
1.3 Parameters affecting microalgae growth	19
1.3.1 Effect of light intensity	20
1.3.2 Effect of nutrients	22
1.3.3 Effects of pH	23
1.3.4 Effects of temperature	23
1.4 Production of biodiesel from microalgae	24
1.5 Cultivation of algae in open ponds and closed systems	26
1.6 Brief overview of the current state of the microalgae based technology	27
1.7 Improving biofuel production by phototrophic microorganisms through system biology	29
1.7.1 Nuclear genome	30
1.7.2 Chloroplast genome	32
1.7.3 Mitochondrial genome	33
1.8 Genetic tools	33
CHAPTER 2. Aim of the work	36
CHAPTER 3. The phylum Chlorophyta	39
3.1. Trebouxiophyceae	40
3.2. The Chlorellales family	41
3.2.1. <i>Chlorella sorokiniana</i>	43
3.2.2. <i>Chlorella variabilis</i> NC64A	45
CHAPTER 4. Kinetic characterization of <i>Chlorella sorokiniana</i>	48
4.1. Strain and culture conditions	49

4.2.	Growth kinetics of <i>C. sorokiniana</i> in batch photobioreactors	49
4.3.	Grow and kinetics of <i>C. sorokiniana</i> in photobioreactor	53
4.4.	Qualitative lipid characterization	55
4.5.	Lipid extraction	58
4.6.	Fatty acid methyl esters analysis	60
CHAPTER 5. Genetic characterization of oleaginous algae belonging to the Chlorellaceae family		62
5.1.	Isolation, sequencing and annotation of DNA	63
5.2.	<i>Chlorella sorokiniana</i> organelles genomes	64
5.2.1.	Chloroplast DNA	64
5.2.2.	Mitochondrial DNA	67
5.3.	<i>Chlorella variabilis</i> NC64A organelles genomes	69
5.3.1.	Chloroplast DNA	69
5.3.2.	Mitochondrial DNA	71
CHAPTER 6. Phylogenetic analysis		73
6.1.	Introduction	74
6.2.	Gene content and structure	75
6.3.	Gene order	75
6.4.	Ancestral gene clusters	77
6.5.	Phylogenetic analysis	79
6.5.1.	Parameters and data set	79
6.5.2.	Phylogenetic reconstruction	80
6.6.	Concluding remarks	83
CHAPTER 7. Conclusions and suggestions for further research		84
REFERENCES		87

LIST OF TABLES

TABLE 1. Biomass productivities, lipid content and lipid productivities of different microalgae species	17
TABLE 2. Photonic features of major pigments in microalgae	21
TABLE 3. Comparison of next-generation sequencing platforms	30
TABLE 4. List of published algal-genome projects	31
TABLE 5. Chemical composition of the BBM growth medium	50
TABLE 6. Fatty acid methyl esters profile of lipids extracted after using the different culture condition	60
TABLE 7. Gene list for the <i>Chlorella sorokiniana</i> chloroplast chromosome	66
TABLE 8. Gene list for the <i>Chlorella sorokiniana</i> mitochondrial genome	68
TABLE 9. Gene list for the <i>Chlorella varibilis</i> chloroplast genome	70
TABLE 10. Gene list for the <i>Chlorella varibilis</i> mitochondrial genome	72
TABLE 11. General features of cp DNAs from <i>Chlorella sorokiniana</i> and other Chlorellaceae	74

LIST OF FIGURES

FIGURE 1. Classification of biofuels	12
FIGURE 2. Conceptual scheme for the production of biofuels and CO ₂ capture through microalgae	14
FIGURE 3. Simplified scheme of photosynthesis in microalgae and lipid production metabolic pathways (adapted from Yonghua, 2012).	16
FIGURE 4. Schematic representation of biomass growth in a batch culture	20
FIGURE 5. Schematic representation of the "algae to biodiesel" process	25
FIGURE 6. Next-generation sequencing and its applications	29
FIGURE 7. Overview phylogeny of the green lineage	40
FIGURE 8. Phylogeny of Trebouxiophyceae	42
FIGURE 9. <i>Chlorella sorokiniana</i> .	44
FIGURE 10. Distribution of total lipid extracts as fatty acids and TAG in three <i>Chlorella</i> strains	45
FIGURE 11. Paramecium bursaria with symbiotic <i>Chlorella variabilis</i>	46
FIGURE 12. Calibration line showing the correlation between wet weight of biomass pellet subjected to disruption and the corresponding dry weight content	51
FIGURE 13. <i>Chlorella sorokiniana</i> grown in standard BBM	52
FIGURE 14. <i>Chlorella sorokiniana</i> grown in nitrogen supplement	52
FIGURE 15. <i>Chlorella sorokiniana</i> grown in nitrogen starvation	53
FIGURE 16. Growth of <i>Chlorella sorokiniana</i> in helical tubular photobioreactor in fedbatch mode	54
FIGURE 17. Schematic representation of colorimetric method for lipids quantification which employs sulpho-phospho-vanillin	55
FIGURE 18. Calibration curve using of lipid standard solutions (canola oil)	56

FIGURE 19. Lipids amount in <i>Chlorella sorokiniana</i> growth in standard BBM medium in batch photobioreactor	57
FIGURE 20. Lipids amount in <i>Chlorella sorokiniana</i> growth in modified BBM with nitrogen supplementation in batch photobioreactor	57
FIGURE 21. Lipids amount in <i>Chlorella sorokiniana</i> growth in modified BBM with nitrogen starvation in batch photobioreactor	58
FIGURE 22. Illumina's HiSeq2000 sequencing system and data quality over two libraries (300 – 800 bp) for each investigated strain	63
FIGURE 23. Gene map of the chloroplast genome of <i>C. sorokiniana</i>	64
FIGURE 24. Gene map of the mtDNA of <i>C. sorokiniana</i>	67
FIGURE 25. Gene map of the chloroplast genome of <i>C. varibilis</i>	69
FIGURE 26. Gene map of the mitochondrial genome of <i>C. varibilis</i>	71
FIGURE 27. Comparison between the <i>C. sorokiniana</i> , <i>C. variabilis</i> and <i>Chlorella</i> sp..	76
FIGURE 28. Comparison between the <i>C. sorokiniana</i> and <i>C. vulgaris</i>	77
FIGURE 29. Conserved gene clusters in the Chlorellales algae	78
FIGURE 30. Phylogeny of 12 Chlorellaceae inferred using maximum likelihood (ML) analyses of nuclear 18S rRNA	81
FIGURE 31. Phylogeny of 12 Chlorellaceae inferred using chloroplast rRNA operon data sets (16S-tRNA Ile-Ala-23S)	82
FIGURE 32. Phylogeny of Chlorellaceae inferred using 71 chloroplast genes	83

LIST OF PAPERS

This thesis is based on the following papers:

- **Costelli C.**, Malavasi V., Orsini M., Angius A., Concas A., Cao G. (2015) *C. sorokiniana* SAG 2111-8k chloroplast sequence confirm the position of the *Chlorella* clade inside the Chlorellaceae family (Chlorophyta). Journal of Molecular Biology. Manuscript in preparation.
- Concas A., **Costelli C.**, Malavasi V., Cao G. (2015) Effects of nutrients on the kinetic growth and lipids amount of *Chlorella sorokiniana*. Advances in Environmental Research. Manuscript in preparation.
- Orsini M., **Costelli C.**, Malavasi V., Cusano R., Concas A., Angius A., Cao G. (2015) Complete sequence and characterization of mitochondrial and chloroplast genome of *Chlorella variabilis* NC64A. Mitochondrial DNA. doi: 10.3109/19401736.2015.1007297.
- Orsini M., **Costelli C.**, Malavasi V., Cusano R., Concas A., Angius A., Cao G. (2014) Complete genome sequence of mitochondrial DNA (mtDNA) of *Chlorella sorokiniana*. Mitochondrial DNA. doi: 10.3109/19401736.2014.953128.
- Orsini M., Cusano R., **Costelli C.**, Malavasi V., Concas A., Angius A., Cao G. (2014) Complete genome sequence of chloroplast DNA (cpDNA) of *Chlorella sorokiniana*. Mitochondrial DNA. doi: 10.3109/19401736.2014.919466.

LIST OF ABBREVIATIONS AND DEFINITION

BBM	Bold Basal Medium
BEN	Bilancio Energetico Nazionale
BLAST	Basic Local Alignment Search Tool
dsDNA	double strand DNA
EIA	Energy Information Administration
GC-MS	Gas Chromatography-Mass Spectrometry
GpC	Cytosine-phosphate-Guanine
NCBI	National Center for Biotechnology Information
NO _x	Nitrogen oxides
ORFs	Open Reading Frames
PBRs	Photobioreactors
rRNA	ribosomal Ribonucleic Acid
SAG	Sammlung von Algenkulturen Collection Gottingen
SO _x	Sulphur oxides
TAE	Tris Acetate-EDTA
tRNA	transfer ribonucleic acid

CHAPTER 1.

Introduction

1.1 Microalgae as a biofuel feedstock

The identification of biofuels feedstocks characterized by environmental, economic and ethical sustainability is one of the main challenges to which the scientific community involved in the biofuels sector, is called to find a solution. Moreover, the biofuels industry, while spreading worldwide at a high rate, is still facing a growing dilemma, i.e. the finding of a feedstock source capable of keeping up with demand of fuels for the transportation sector. In fact, as energy demands and oil reserves begin to falter in their stability, the need for a reliable renewable fuel source grows. In Italy, 92% of the energy consumed by the transportation sector is today obtained from fossil oil and only the 3.4% from renewable sources (BEN 2013). The data are similar to the American consumption in the same field for the same year: 95% comes from petroleum and 4.6 % from renewable source (EIA 2013). Biofuels deriving from feedstocks such as plants, organic wastes or algae could help to reduce the world's oil dependence (Naik et al., 2010). Moreover, biofuels would mitigate global warming problems since all the CO₂ emitted during their burning can be fixed by plants used as biomass feedstock, through photosynthetic mechanisms. To date there are three different generations of biofuels (cf. Figure 1).

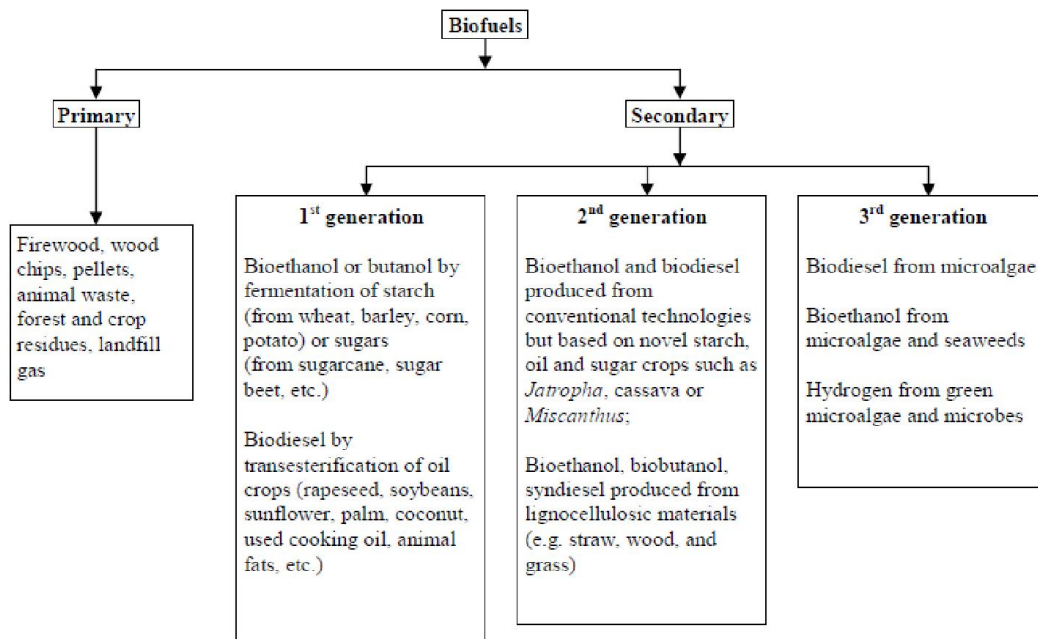


Figure 1. Classification of biofuels.

First generation biofuels are mainly biodiesel (or bio-esters in general) ethanol and biogas produced starting from biomass feedstocks such as rapeseed, sugarcane, maize, soybeans, palm oil and sunflower. They have now attained economic levels of

production and their use is continuously increasing all over the world. On the other hand, the large scale diverting of farmlands or crops for biofuels can provoke the detriment of the food supply as well as a dramatic increase of food price which mainly impacts on the most vulnerable regions of the world. Another important concern is related to the low areal productivity of energetic crops which results in the need of huge land's areas for meeting the global demand of biofuels (Chisti, 2007). This aspect determines important concerns related to the conversion of forests into energetic crops which can, in turn, contribute to enhance problems such as soil erosion, loss of habitat and reduction of valuable biodiversity. Finally, life cycle assessments of first generation biofuels clearly demonstrated that their production and transportation can result in the emission of CO₂ amounts that approach those ones typical of traditional fossil fuels.

The food versus fuel competition triggered by first generation biofuels has stimulated a greater interest for the development of biofuels produced from non-food biomass, commonly referred to as second generation biofuels (Timilsina et al., 2010). Biomasses to produce the latter ones are basically organic wastes and byproducts that may result from different manufacturing processes. Among the others, feedstocks can include for example agricultural residues, domestic wastes, wood/forestry residues, exhausted cooking oils and food wastes, even if, also non edible oil crops might be accounted among the suitable feedstock (Ahmad et al., 2011; Brennan and Owende, 2010). The most promising second generation biofuels are those produced starting from residual of non-edible lignocellulosic biomasses derived from plants. Lignocellulosic biomass is actually one of the most abundant and less exploited biological resources on the planet and thus today there is a great interest in developing suitable technologies for transforming it into liquid biofuels (Naik et al., 2010). When compared to first generation ones, second generation biofuels do not involve 'food vs. fuel' antagonism, are more efficient and more environmentally friendly, require less farmland and can be grown in lands that are not suitable for food crops. However, since their production is strictly related to byproducts availability, second generation biofuels may not be abundant enough to meet the fuel demand on a large scale (Ahmad et al., 2011).

Third generation biofuels derived from microalgae are considered to be a viable alternative energy resource to avoid the major drawbacks associated with first and second generation biofuels. When compared to conventional crop plants, which are

usually harvested once or twice a year, microalgae have a very short harvesting cycle, allowing multiple or continuous harvests with significantly higher yields. They can be cultivated under hard agro-climatic conditions and are able to produce a wide range of commercially interesting byproducts such as fats, oils, sugars and functional bioactive compounds. Several microalgae are exceedingly rich in oil, which can be extracted and subsequently converted to biodiesel using existing technologies (Chisti, 2008). When compared to crops used for first generation biofuels, microalgae display superior biomass growth rates. Moreover, the corresponding oil content is higher than the one of terrestrial crops since it can exceed 80% of the dry weight of biomass. For these reasons the oil productivity of microalgae exceeds that of terrestrial crops even 10-100 times (Chisti, 2008). Microalgae, differently from crops, are cultured in aquatic environments. For these reasons, cultivation of microalgae can be carried out in less extended and lower-quality lands, thus avoiding the exploitation of arable ones. In addition, cultivation of microalgae might be coupled with the direct bio-capture of CO₂ emitted by industrial activities. Certain microalgae are effective in the production of hydrogen and oxygen through the process of biophotolysis while others naturally manufacture hydrocarbons which are suitable for direct use as high-energy liquid fuels. Therefore, the potential use of microalgae as renewable feedstock for the massive production of liquid biofuels is receiving a rising interest mostly driven by the global concerns related to the depletion of fossil fuels supplies and the increase of CO₂ levels in the atmosphere. From a conceptual point of view the process shown in Figure 2 can be carried out for producing biofuels and capturing CO₂ through microalgae.

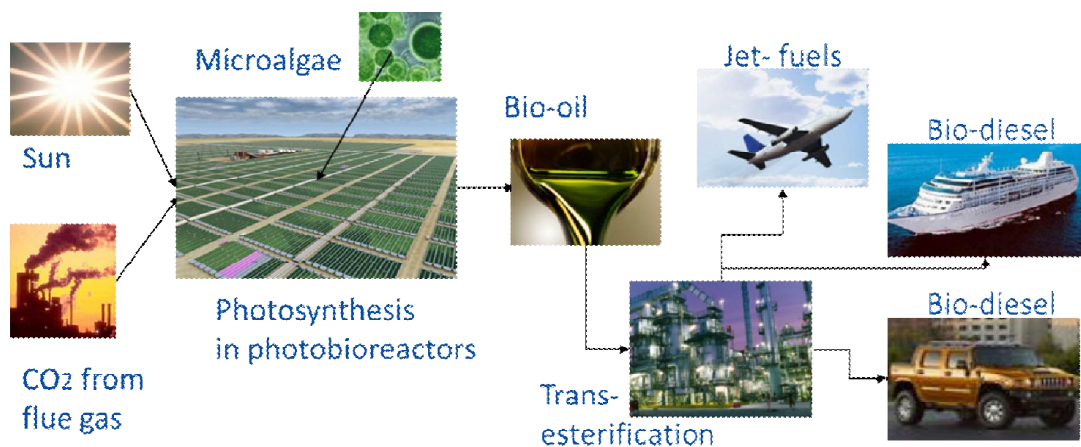


Figure 2. Conceptual scheme for the production of biofuels and CO₂ capture through microalgae (Concas et al., 2014).

Despite the apparent simplicity of the process, its implementation to the industrial scale is still not widespread since it is characterized by technical and economic constraints that might hinder its full scale-up.

1.2 Microalgae: importance, photosynthesis and lipid production

Microalgae, today well recognized to be as one of the oldest living organisms, are thallophytes having chlorophyll a as their primary photosynthetic pigment. The diversity among algae is enormous, not only with respect to the size and shape of the organisms, but also with respect to the various plastids that carry out photosynthesis in these eukaryotic cells. Those eukaryotic phototrophic microorganisms, are crucial for life on earth. Planktonic algae or phytoplankton, living in the oceans perform nearly half of the global photosynthesis (Behrenfeld and Falkowski, 1997). While the mechanism of photosynthesis in these microorganisms is similar to that of higher plants, they are generally more efficient converters of solar energy because of their simple cellular structure. In addition, since the cells grow in aqueous suspension, they have more efficient access to water, CO₂, and other nutrients. Microalgae contain also numerous bioactive compounds that can be harnessed for commercial use (astaxanthin, lutein, carotenoids, ω3 and ω6). According to recent researches, algal protein is considered also as an optimal source for animal feed since algal protein has been estimated to have a relevant profile of amino acid (Gross, 2013). Chlorophyll is one of the valuable bioactive compounds that can be extracted from microalgal biomass. There are two main types of chlorophyll, chlorophyll a and chlorophyll b. Such molecule allows to selectively absorb light in the red and blue regions and therefore emits a green colour. Photosynthesis is a process which uses the light energy harvested by pigments to convert water and carbon dioxide into oxygen and carbohydrates. Ultimately, it converts solar energy into chemical energy. The basic chemical equation can be summarized as: $6\text{H}_2\text{O} + 6\text{CO}_2 \rightarrow \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2$. The products from this chemical process reflects its significance, with carbohydrates being the primary building block for plants and oxygen being necessary for the survival of animal kingdom (Humphrey, 2004). The importance of photosynthesis for life on earth is further highlighted by plants forming the basis of all food chains

(Hosikian et al., 2010). The photosynthetic process (cf. Figure 3) can be divided into two sets of reactions: the light-dependent (light) reactions and the light-independent (dark) reactions. The first ones, which convert the energy of light into chemical energy, take place within the thylakoid membranes of the chloroplasts, whereas the dark reactions, which use the produced chemical energy to fix CO_2 into organic molecules, occur in the stroma of the chloroplast. During the light reactions, the energy transported by incident photons is captured by specific pigments and then used to "split water" into molecular oxygen, two H^+ ions and into one pair of electrons, respectively. The energy of light is thus transferred to these electrons and is, finally, used to generate adenosine triphosphate (ATP) and the electron carrier nicotinamide adenine dinucleotide phosphate (NADPH). These two compounds carry the energy and the electrons generated during the light reactions to the stroma, where they are used by the enzymatic dark reactions related to the Calvin cycle to synthesize sugars from CO_2 . The main sugar synthesized during the Calvin cycle is glyceraldehyde 3-phosphate (G3P). Therefore the net result of the photosynthesis is the conversion of light, water and CO_2 into G3P and molecular oxygen.

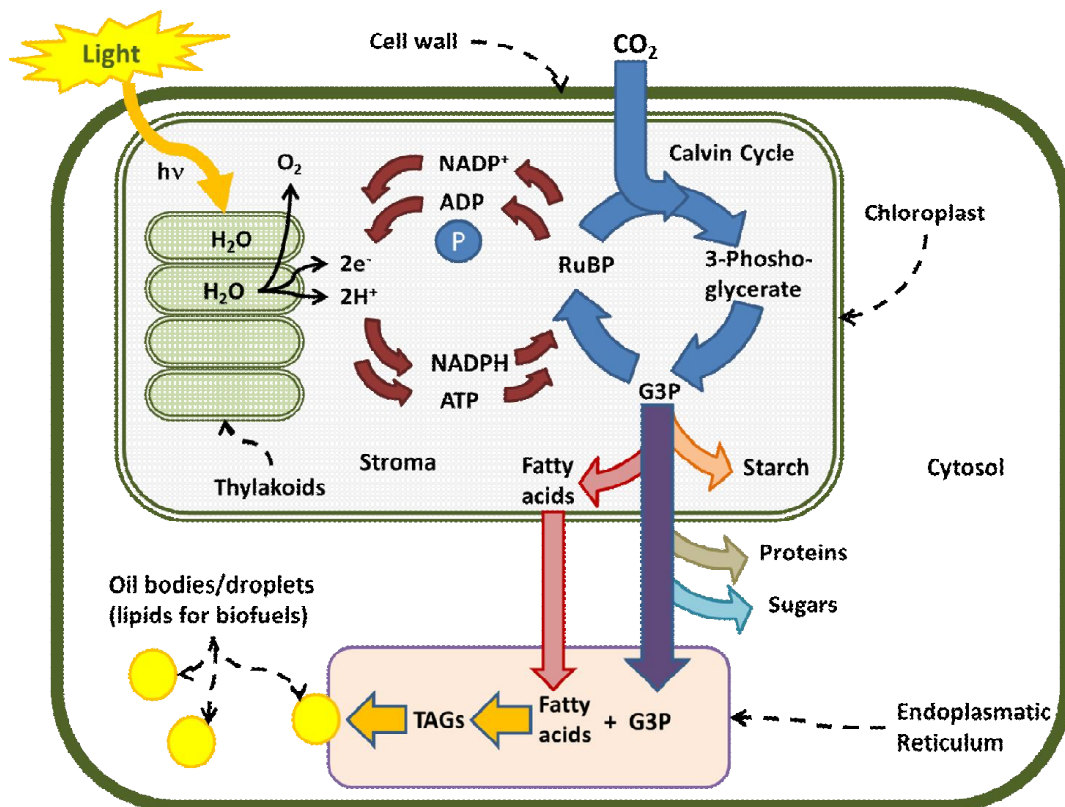


Figure 3. Simplified scheme of photosynthesis in microalgae and lipid production metabolic pathways (adapted from Yonghua, 2012).

The synthesized G3P finally passes into the cytosol where it will be involved as intermediate in the central metabolic pathways of the cell that lead to the production of several macromolecules among which starch, proteins and sugars. In the chloroplast also free fatty acids are synthesized starting from G3P. Fatty acids, along with G3P, are then transferred to the endoplasmatic reticulum where they are further converted into non polar storage lipids, such as triacylglycerides (TAGs), through a number of enzymatic reactions. Finally, TAGs are packaged into oil bodies that bud off into the cytosol (Sakthivel, 2011). These oil bodies have a fatty acid composition comparable to vegetable oils and thus can be extracted from the microalgae cell and subsequently converted to useful biofuels (Klok et al., 2013). Specifically, oils from algae can yield biodiesel through transesterification, and gasoline (petrol) or jet fuels through distillation and cracking, respectively (Georgianna and Mayfield, 2012). When compared to first generation biomass feedstocks, microalgae have been found to contain higher concentrations of lipids. The average lipid content varies between 1 and 70% while under specific operating conditions certain species can reach 90% of oil weight by weight of dry biomass (Mata et al., 2010). Depending on the specific strain considered, microalgae can be characterized by high biomass growth rates which, coupled with the intrinsic high lipid content, can lead to very high oil productivity. Table 1 shows lipid content as well as lipid and biomass productivities of different microalgae species.

Table 1. Biomass productivities, lipid content and lipid productivities of different microalgae species (Concas et al., 2014).

Strain	Biomass productivity (g/L/day)	Lipid content (% biomass)	Lipid productivity (mg/L/day)	Reference
Botryococcus braunii	0,35	17,9	61,8	(Orpez et al., 2009)
Botryococcus braunii	0,29	17,9	51,4	(Orpez et al., 2009)
Botryococcus braunii	0,03	36,1	12,3	(Sydney et al., 2011)
Botryococcus braunii*	0,02	50,0	10,0	(Mata et al., 2010)
Botryococcus braunii	0,04	22,0	9,5	(Dayananda et al., 2007)
Chaetoceros calcitrans	0,04	39,8	17,6	(Rodolfi et al., 2009)
Chaetoceros muelleri	0,07	33,6	21,8	(Rodolfi et al., 2009)
Chaetoceros muelleri*	0,07	33,6	21,8	(Mata et al., 2010)
Chlamydomonas reinhardtii	2,00	25,3	505,0	(Kong et al., 2010)
Chlorella*	0,00	37,5	18,7	(Mata et al., 2010)
Chlorella emersonii*	0,04	44,0	30,2	(Mata et al., 2010)
Chlorella protothecoides	7,30	50,3	3671,9	(Xiong et al., 2008)
Chlorella protothecoides	4,10	43,0	1763,0	(Cheng et al., 2009)
Chlorella protothecoides	4,85	36,2	1214,0	(Mata et al., 2010)
Chlorella protothecoides	2,02	55,2	1115,0	(Xu et al., 2006)
Chlorella pyrenoidosa*	3,27	2,0	65,4	(Mata et al., 2010)
Chlorella sorokiniana IAM-212	0,23	19,3	44,7	(Rodolfi et al., 2009)

<i>Chlorella</i> sp.	0,00	32,6	110,0	(Hsieh and Wu, 2009)
<i>Chlorella</i> sp.	0,08	66,1	51,0	(Hsieh and Wu, 2009)
<i>Chlorella</i> sp.*	1,26	29,0	42,1	(Mata et al., 2010)
<i>Chlorella vulgaris</i>	0,35	42,0	147,0	(Feng et al., 2011)
<i>Chlorella vulgaris</i>	0,35	42,0	147,0	(Feng et al., 2011)
<i>Chlorella vulgaris</i>	0,20	18,4	36,9	(Rodolfi et al., 2009)
<i>Chlorella vulgaris</i>	0,15	23,0	35,0	(Liang et al., 2009)
<i>Chlorella vulgaris</i>	0,17	19,2	32,6	(Rodolfi et al., 2009)
<i>Chlorella vulgaris</i>	0,09	34,0	31,0	(Liang et al., 2009)
<i>Chlorella vulgaris</i> *	0,11	31,5	25,6	(Mata et al., 2010)
<i>Chlorella vulgaris</i>	0,10	22,0	22,0	(Liang et al., 2009)
<i>Chlorococcum</i> sp.	0,28	19,3	53,7	(Rodolfi et al., 2009)
<i>Chlorococcum</i> sp.*	0,28	19,3	53,7	(Mata et al., 2010)
<i>Crypthecodinium cohnii</i> *	10,00	35,6	3555,0	(Mata et al., 2010)
<i>Dunaliella primolecta</i> *	0,09	23,1	20,8	(Mata et al., 2010)
<i>Dunaliella salina</i> *	0,28	15,5	116,0	(Mata et al., 2010)
<i>Dunaliella</i> sp.*	0,00	42,3	33,5	(Mata et al., 2010)
<i>Dunaliella tertiolecta</i> *	0,12	43,9	52,6	(Mata et al., 2010)
<i>Ellipsoidion</i> sp.	0,17	27,4	47,3	(Rodolfi et al., 2009)
<i>Euglena gracilis</i> *	7,70	17,0	1309,0	(Mata et al., 2010)
<i>Haematococcus pluvialis</i> *	0,06	25,0	13,8	(Mata et al., 2010)
<i>Isochrysis galbana</i> *	0,96	23,5	225,6	(Mata et al., 2010)
<i>Isochrysis galbana</i>	0,17	22,3	38,0	(Su et al., 2007)
<i>Isochrysis galbana</i>	0,12	14,3	17,2	(Su et al., 2007)
<i>Isochrysis</i> sp.	0,14	27,4	37,8	(Rodolfi et al., 2009)
<i>Isochrysis</i> sp.*	0,13	20,1	37,8	(Mata et al., 2010)
<i>Isochrysis</i> sp.	0,17	22,4	37,7	(Rodolfi et al., 2009)
<i>Monallanthus salina</i> *	0,08	21,0	16,8	(Mata et al., 2010)
<i>Monodus subterraneus</i>	0,19	16,1	30,4	(Liang et al., 2009)
<i>Monodus subterraneus</i> *	0,19	16,0	30,4	(Mata et al., 2010)
<i>Nannochloris</i> sp.*	0,34	38,0	68,7	(Mata et al., 2010)
<i>Nannochloropsis</i>	0,17	29,2	49,7	(Rodolfi et al., 2009)
<i>Nannochloropsis</i> sp.	0,21	29,6	61,0	(Rodolfi et al., 2009)
<i>Nannochloropsis</i> sp.	0,20	24,4	48,2	(Rodolfi et al., 2009)
<i>Nannochloropsis</i> sp.	0,17	21,6	37,6	(Rodolfi et al., 2009)
<i>Neochloris oleabundans</i>	0,31	40,0	125,0	(Li et al., 2008a)
<i>Neochloris oleabundans</i>	0,63	15,0	98,0	(Li et al., 2008a)
<i>Neochloris oleabundans</i>	0,15	28,0	37,8	(Gouveia et al., 2009)
<i>Neochloris oleabundans</i>	0,03	52,0	14,4	(Gouveia et al., 2009)
<i>Neochloris oleabundans</i> *	0,00	47,0	112,0	(Mata et al., 2010)
<i>Pavlova lutheri</i>	0,14	35,5	50,2	(Rodolfi et al., 2009)
<i>Pavlova salina</i>	0,16	30,9	49,4	(Rodolfi et al., 2009)
<i>Phaeodactylum tricorutum</i>	0,24	18,7	44,8	(Rodolfi et al., 2009)
<i>Porphyridium cruentum</i>	0,37	9,5	34,8	(Rodolfi et al., 2009)
<i>Scenedesmus obliquus</i> *	0,37	33,0	122,8	(Mata et al., 2010)
<i>Scenedesmus quadricauda</i>	0,19	18,4	35,1	(Rodolfi et al., 2009)
<i>Scenedesmus</i> sp.	0,26	21,1	53,9	(Rodolfi et al., 2009)
<i>Scenedesmus</i> sp.*	0,15	20,4	47,4	(Mata et al., 2010)
<i>Scenedesmus</i> sp.	0,21	19,6	40,8	(Rodolfi et al., 2009)
<i>Skeletonema costatum</i>	0,08	21,1	17,4	(Rodolfi et al., 2009)
<i>Skeletonema</i> sp.	0,09	31,8	27,3	(Rodolfi et al., 2009)
<i>Skeletonema</i> sp.*	0,09	22,6	27,3	(Mata et al., 2010)
<i>Spirulina maxima</i> *	0,23	6,5	15,0	(Mata et al., 2010)
<i>Spirulina platensis</i> *	2,18	10,3	224,5	(Mata et al., 2010)
<i>Tetraselmis</i> sp.*	0,30	13,7	43,4	(Mata et al., 2010)
<i>Tetraselmis suecica</i>	0,28	12,9	36,4	(Rodolfi et al., 2009)
<i>Tetraselmis suecica</i> *	0,22	15,8	31,7	(Mata et al., 2010)
<i>Tetraselmis suecica</i>	0,32	8,5	27,0	(Rodolfi et al., 2009)
<i>Thalassiosira pseudonana</i>	0,08	20,6	17,4	(Rodolfi et al., 2009)

* average values are reported

The volumetric lipid productivity of microalgae is extremely variable depending upon the specific strain considered, and goes from 0.01 to 3,67 g oil L⁻¹ day⁻¹. However, it is worth noting that lipid productivity can be strongly affected by the specific culturing conditions adopted, i.e. growth medium composition, light regime, photobioreactor configuration and operation mode etc.. Consequently, it can be argued that, by suitably choosing the best performing strains, very high volumetric productivities of lipids can be achieved by using algae.

1.3 Parameters affecting microalgae growth

It is well known that algae growth in batch cultures proceeds according to the five main phases depicted in Figure 4 and described in what follows (Jalalizadeh, 2012):

- a lag phase, where a growth delay takes place when cultivation starts due to physiological adjustments of the inoculum to changes in nutrient concentration, light intensity and other culture conditions;
- an exponential phase, where cells grow and replicate exponentially with time, as long as all the conditions affecting algae growth are optimized, i.e. nutrients and light availability, optimal temperature, pH, etc.;
- a linear growth phase, where biomass concentration grows linearly as a function of time;
- a stationary growth phase, where the biomass concentration remains constant as a result of the reduced availability of nutrients and light that lead the death rate to equal the growth one;
- a decline or death phase, where the decrease in the concentration of nutrients and/or the accumulation of toxic waste products lead the death rate to overcome the growth one.

Such a growth behavior can be well described by the mass balance for microalgae biomass reported in what follows:

$$\frac{dX}{dt} = (\mu - k_d) X$$

where X is the microalgal biomass concentration (mass/volume), μ is the specific growth rate (1/time) and k_d is the specific mass loss rate (1/time) which accounts for all the phenomena that are responsible of biomass depletion, i.e. cell catabolism,

apoptosis, lysis, etc.. The term $(\mu - k_d)$ is the net growth rate. While k_d is usually considered to be constant, the growth rate μ depends upon several factors which can affect microalgae growth. Among them, light, nutrient concentration, pH and temperature (T) are quite important.

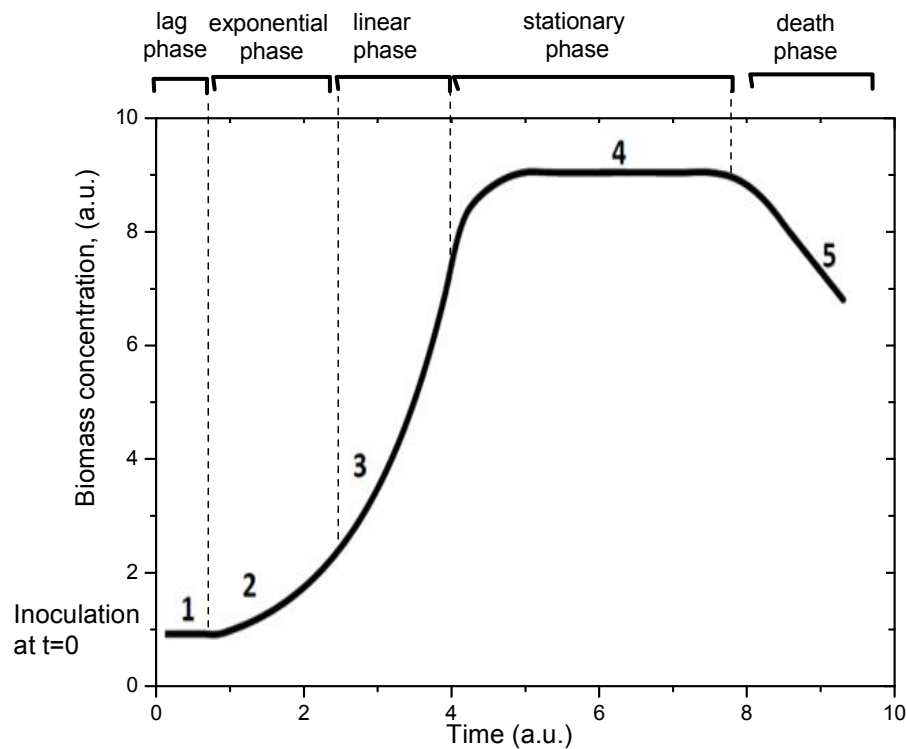


Figure 4. Schematic representation of biomass growth in a batch culture (adapted from Jalalizadeh, 2012).

1.3.1 Effect of light intensity

Light is essential for the phototropic growth of microalgae. Indeed, although photosynthesis has been optimized over three billion years of evolution, it remains inefficient at converting solar energy into chemical energy and biomass. The theoretical photoconversion efficiency of about 27% drops to 6% due to reductions in the efficiency of photon utilization and biomass accumulation (Weyer et al., 2010). Indeed, although the theoretical maximum productivity of microalgae is estimated to be around 170 to 190 g DW m⁻² d⁻¹ (Weyer et al., 2010), the reported efficiencies in ponds or PBRs ranged from 20 to 35 g DW m⁻² d⁻¹ (Rodolfi et al., 2009) with the present technology and available strains. The reasons for this efficiency drop in dense culture conditions are rooted in the very structural

organization of the photosynthetic apparatus. Oxygenic photosynthesis is performed by four multisubunit membrane-protein complexes in the thylakoid membrane: two photosystems (PSI and PSII), cytochrome b6f, and ATPase (Nelson et al., 2004). Each photosystem includes a core complex that binds cofactors involved in electron transport together with additional chlorophyll (Chl) a and β -carotene as antenna pigments. Associated with the photosystems is an array of antenna complexes called light harvesting complexes (LHCs) which bind Chl a, b and xanthophylls, and enhance photon absorption and transfer excitation energy for photochemical reactions (Croce et al., 2013; Van Amerongen et al., 2013). LHCs also have essential roles in photoprotection, through the dissipation of excess light as heat (non-photochemical energy quenching, NPQ), and in reactive oxygen species (ROS) scavenging. Photosynthetically active radiation (PAR) designates the spectral range (wave band) of solar radiation from 400 to 700 nm that microalgae are able to use during the process of photosynthesis (cf. Table 2). It should be noted that photons at shorter wavelengths (<400 nm) carry a very high energy content that can damage microalgal cells, while at longer wavelengths (> 700 nm) the energy carried does not allow photosynthesis to take place.

Table 2. Photonic features of major pigments in microalgae (Carvalho et al., 2011).

Pigment group	Color	Ranges of absorption bands (nm)		Pigments
Chlorophylls	Green	450–475 630–675	Hydrophobic	Chlorophyll a Chlorophyll b Chlorophyll c1, c2, d
Phycobilins	Blue, red	500–650	Hydrophilic	Phycocyanin Phycoerythrin Allophycocyanin
Carotenoids	Yellow, orange	400–550	Hydrophobic	β -Carotene α -Carotene Lutein Violaxanthin Fucoxanthin

When light penetrates in an optically dense medium such as a microalgal culture it experiences attenuation phenomena due to absorption by the medium as well by the pigments of microalgal cells. Indeed, the surface layers of microalgae easily reach saturation (and hence, photoinhibition) of photosynthesis (Powles, 1984) while the inner layers are light limited (Neale et al., 1986). This non-homogeneous light penetration results in a low productivity of the system. A reduction in pigment

content per cell and a reduction in antenna size are targets for optimizing the photosynthetic yield of unicellular algae under mass culture (Melis, 2009).

1.3.2 Effect of nutrients

The medium where microalgae grow basically consists of water enriched by macro (C, N, P, S) and micro (Mg, Zn, Fe, K, Na etc.) nutrients as well as by the CO₂ transferred from the gas phase (i.e flue gas or air). Besides CO₂, whose role in photosynthesis has been already discussed in paragraph 1.3.1, nitrogen and phosphorous are key elements for algae metabolism. Their suitable balancing in the growth medium is thus critical for an effective process design (Mandalam and Palsson, 1998). Ammonia, urea and nitrate are often selected as nitrogen source for the mass cultivation of microalgae. Although ammonia and urea are often used in mass cultivation owing to their relatively low-cost, selecting proper nitrogen source for each algal species is important in improving biomass and oil productivity (Li et al., 2008b). Urea and nitrate were found to be better nutrients than ammonia for the growth and lipid accumulation when considering *Chlorella* sp., *Chlorella vulgaris*, *Neochloris oleoabundans* and *Scenedesmus rubescens* (Li et al., 2008a; Hsieh and Wu, 2009). On the contrary, for different strains, the use of ammonia has been demonstrated to provoke higher biomass and lipid content than urea and nitrate (Xu et al., 2001). It should be noted that the optimal concentration of nitrogen to be assured in the growth medium depends upon two counteracting effects. Specifically, while a high availability of nitrogen typically leads to a high biomass productivity, a decrease of nitrogen concentration in the cultivation broth typically results in higher lipid contents counteracted by lower growth rates. Such behavior depends upon the fact that, under starvation conditions, nitrogen concentration is not enough for activating the metabolic pathways leading to protein synthesis required by growth so that the excess of carbon coming from photosynthesis is channeled into storage molecules such as triacylglycerides or starch (Scott et al., 2010). This inverse relationship between biomass productivity and lipid content makes the choice of the suitable nitrogen concentration not straightforward since a trade-off value should be assured in order to maximize lipid productivity (Concas et al., 2013). When considering phosphorus, microalgae are capable of metabolizing it mainly in polyphosphate form. Orthophosphate is generally considered the main limiting

nutrient for freshwater strains but also in this case its optimal concentration depends upon contrasting effects. In fact, phosphorus starvation can result in higher lipid productivity, as reported for *Monodus subterraneus*, while may provoke changes in fatty acids composition for *Phaeodactylum tricornerutum* and *Dunaliella tertiolecta* (Liu et al., 2007). For all these reasons the preparation of the culture broth is a critical step for the entire process of biofuels production through microalgae.

1.3.3 Effects of pH

The time evolution of medium's pH during algal growth is a significant indicator of how well are evolving photosynthetic processes. In fact, as algae grow, dissolved CO₂ is consumed by photosynthesis and, consequently, pH increases. However, pH variation not only represent a fundamental indicator of the evolution of photosynthetic activity but can also, in turn, strongly affects the growth kinetics of microalgae influencing the distribution of carbon dioxide species and carbon availability causing direct physiological effects (Cornet et al., 1995; Chen and Durbin, 1994). Moreover, in microalgal cultures, the hydrogen ion is recognized to be a non-competitive inhibitor near neutral conditions, while it can limit photosynthetic growth and substrate utilization rates at very low or very high pH levels (Mayo, 1997). Furthermore, pH can affect the enzymatic activity of intra and extra-cellular carbonic anhydrase, thus influencing the carbon capture mechanism of some microalgal strains (Concas et al., 2012).

1.3.4 Effects of temperature

Temperature is one of the main factors which regulate cellular, morphological and physiological responses of microalgae (Mayo, 1997; Durmaz et al., 2007). High temperatures generally accelerate the metabolic rates of microalgae, whereas low ones lead to inhibition of microalgal growth (Munoz and Guieysse, 2006). Under optimal temperature condition, the enzymes of microalgal cells show the highest activity. The optimal temperature range for microalgal growth depends on the specific strain considered but in general, it typically goes from a minimum of 5°C to a maximum of 35°C (Abu-Rezq et al., 1999). The control of temperature is a key

factor for cultivating microalgae outdoors. Actually, temperature can vary depending upon the geographic region of cultivation. Seasonal and even daily fluctuations in temperature can interfere with algae production. The internal temperature in photobioreactors can reach values that are 30°C higher than ambient one if suitable temperature control equipment is not used. To overcome this problem evaporation, cooling or shading techniques are successfully employed.

1.4 Production of biodiesel from microalgae

Microalgae cultivation systems are very different from those ones typically used for producing biomass feedstock for first and second generation biofuels. Therefore, when compared to terrestrial crops, the production of microalgae requires specific cultivation, harvesting and processing techniques which should be correctly implemented to the aim of viably produce biodiesel (Mata et al., 2010). Figure 5 shows a schematic representation of the process for the biodiesel production through microalgae. As it can be seen, the process starts with the CO₂ capture and its conveying in the cultivation system where microalgae grow exploiting the sunlight and the nutrients suitably provided. Then, it follows the biomass harvesting, downstream processing and oil extraction to supply the biodiesel production unit. Cultivation of microalgae can be performed in open systems (ponds, raceways, lakes) or in closed ones, i.e. photobioreactors. Whatever the system being used, a suitable source of CO₂ must be supplied to microalgae. To this aim, atmospheric air (0.03 %v/v of CO₂), flue gas (9-15% of CO₂) or pure concentrated CO₂ (100%v/v) can be used. Atmospheric air as CO₂ source, significantly simplifies the lay-out and the operation of the plant while, because of the lower CO₂ concentration in air, high volumes of air are required in order to sustain microalgae growth at an acceptable rate. This can result in very large cultivation systems that require a high land availability. On the contrary, when flue gases are used as carbon source, lower flow rates of gases should be pumped into the cultivation system for supplying the necessary amounts of carbon to sustain microalgae growth.

Moreover, the use of costless feedstocks such as flue gases as source of CO₂ might greatly improve the economic feasibility of the microalgae-based technology while, simultaneously, producing a positive impact on significant environmental concerns

such as air pollution and climate changes. For this reason the potential exploitation of CO₂ from flue gases is one of the main targets of scientists and technicians operating in this field (Concas and Cao, 2011; Francisco et al., 2010). However, the use of flue gas as carbon source might raise specific concerns related to the toxicity of some of its constituents with respect to algae. For this reason the flue gas should be pre-treated before feeding it in the cultivation system. A further challenge in the carbon capture through microalgae is the use of pure concentrated CO₂ (100 %v/v) obtained from flue gas. In this case in fact, beside the lesser volumes of photobioreactors that are needed, the potential poisoning effects provoked by other compounds in flue gas (NO_x, SO_x etc.) could be reduced thus increasing the net growth rate of microalgae.

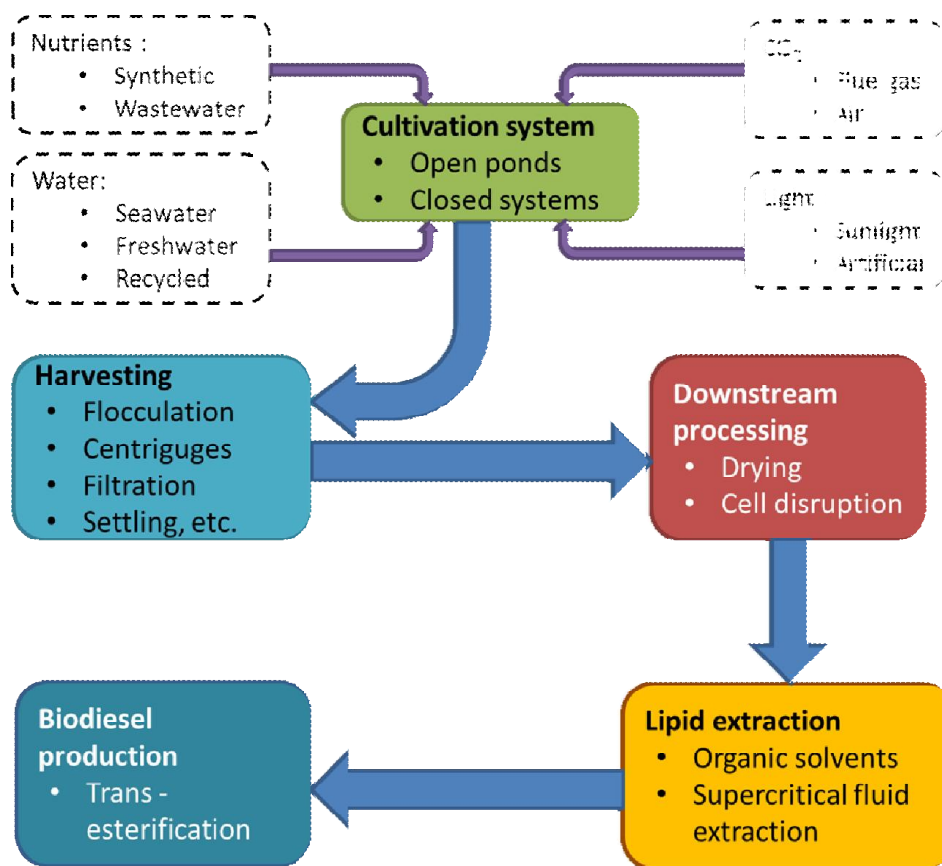


Figure 5. Schematic representation of the "algae to biodiesel" process (Concas et al., 2014).

Besides CO₂, several micro and macro nutrients must be supplied to the culture in order to sustain microalgal growth. It is noteworthy to underline that the exploitation of costless feedstocks such as wastewaters as sources of macronutrients, might greatly improve the economic feasibility of the microalgae-based technology while simultaneously producing a positive impact on important environmental concerns such as water pollution. In fact, wastewaters, even if pre-treated, may contain residual concentration of nitrogen and phosphorus which are capable to sustain

microalgal growth (Concas and Cao, 2011). In particular, industrial and agricultural wastewater and secondary sewage treated effluent can be used as source of nitrogen and phosphorus (Devi et al., 2012). For this reason the operation step of medium preparation can involve a pre-mixing with wastewater.

1.5 Cultivation of algae in open ponds and closed systems

Open pond systems are cheap to construct and different designs have been proposed for open ponds, natural or artificial ones, operating at large scale. Typical examples are the unstirred ponds (lakes and natural ponds), the inclined ones, central pivot and the raceway ponds. Among the others, the most widespread typology of open pond is the so called “raceway pond”. However, low-cost open pond systems struggle to overcome several problems. These systems are more susceptible to contaminations by competing organisms such as mushrooms, bacteria and protozoa. Furthermore, since atmospheric carbon dioxide is used as carbon source, its transfer rate is very low and consequently carbon starvation phenomena could take place. Finally, sunlight is available only at the surface of the pond and hence, in the deeper strata of the liquid bulk, light limitation phenomena can arise. In addition open systems do not offer control over temperature and lighting. The growing season is largely dependent on location and, aside from tropical areas, is limited to the warmer months.

To overcome limitations related to open system and in the meantime keeping their low operating cost, the potential use of closed raceway ponds are currently under study.

Photobioreactors (PBR) are closed systems having no direct exchange of gases and contaminants with the environment where culture broth and microalgae are exposed to a photonic energy flux which triggers photosynthetic phenomena hence allowing biomass growth. Since they are closed reactors the crucial operating parameters such as temperature, pH, nutrient concentration, light intensity distribution, mixing, gas mass transfer rate can be suitably controlled and optimized. As a result photobioreactors typically have higher biomass productivities than open ponds. On the contrary, photobioreactors are more expensive and complicated to operate than open ponds.

1.6 Brief overview of recent large-scale projects regarding the microalgae based technology

In November 2006, the U.S. Green Energy Technology Company and the Arizona Public Service Company signed a joint agreement aimed to conduct an industrial research activity whose target was to develop a microalgae production system in Arizona State. The microalgae production system was fed with the CO₂ of a flue gas from a 1040 MW power plant. In fact, the main target of the project was to verify whether the flue gas could be used as source of carbon dioxide to sustain microalgae growth in a large-scale plant. The microalgae were then subjected to a procedure aimed to extract the lipids which, subsequently were converted to biofuels through classical transesterification procedures. The achieved biofuels productivity was about 5000-10000 gallon per acre per year (Keune, 2012).

In 2007, The National Energy Board of the United States launched the so called “Mini-Manhattan Project” whose aim was to verify whether the use of microalgae might help the United States to get rid of oil dependence concerns. In this regard, a study by Keune (2012) confirmed that in 2010 U.S. had the potential capability to produce millions of microalgal bio-oil barrels per day through a commercially viable industrial scale production system. In the light of the consideration above, the U. S. Department of Energy in 2010 allocated \$ 6 million funds to the Arizona State University with the aim of establishing sustainable algal biofuel consortium (SABC). Furthermore, \$ 9 million U.S. dollars were allocated to University of California (San Diego) from the Department of Energy to fund a research activity aimed to create an algae biofuel commercialization consortium (CABC). Also the Cellana LLC consortium of Kailua-Kona and Hawaii was funded with 9 M\$ by the US Department of Energy in order to support algae development on finance.

The president Barack Obama made a speech at the University of Miami wherein he said that the renewable energy positioning in 2012 is algae bioenergy. In particular, he declared that his government would have invested 24 M\$ in biofuel development, and claimed that “Whether you believe it or not, we can rely on the United States home-grown biofuel to replace 17% of imported petroleum fuels” (Keune, 2012).

China is also paying attention to the microalgae bioenergy development. In the large-scale microalgae cultivation field, Chinese Academy of Sciences has successfully developed a large “S” shaped pipe of closed photo biological reactor. In 2008, Shandong University of Science and Technology used the thermal power plant and

chemical plant flue gas as sources of carbon dioxide to feed microalgae cultivation in the tower dimensional cultivation reactor. Subsequently, mature microalgae are transferred to energy conversion equipment to product bio-oil. This research results improved microalgae production technology.

In 2010, the Chinese Academy of Sciences and China Petroleum Chemical Cooperation held a meeting about “Microalgae Biodiesel Technology Projects” where they decide to co-operate in developing microalgae biodiesel technology. Till now the outcomes of this cooperation are represented by a pilot study and an outdoor middle scale of microalgae energy plant will be established within 2015 (Wei, 2010). Hainan Greenbelt microalgae Biotechnology Company realized a microalgae cultivation experimental base in Ledong County, the microalgae oil content can reach to 28% - 32% .This Company planned to invest 29.8 M\$ to develop a microalgae project in Hainan province. The target of this project is to produce about 30 Mtons/year of biodiesel (Li, 2009).

The new Austrian technology company is brewing a huge plan. They realized a microalgae ecological base in the Daqi of Inner Mongolia. It consists of 280 hectares of microalgal photobioreactors which have been completed in 2013. Microalgae ecological base officially reached industrialization level in 2014 (Li, 2009).

In 2007, an algae production plant in the north part of Sweden was set to be built by Umeå Energi. This project is headed by Swedish University of Agricultural Science, and the development funding from Umeå Energi, Umeva AB, Ragn-Sells AB and Energi myndigheten. The project is focusing on the wastewater reclamation of algae production process, CO₂ sequestration from flue gas, and valuable algae biomass production. (Aylott, 2011).

In September 6, 2011, the EU launched an algal bioenergy development action plan that will be set to the Algae Development Project (EnAlgae) during the next four years. This action focused on the growth and yield information of microalgae and giant algae which are growing in the North Western Europe. The National Non-Food Crops Centre (NNFCC) in the U.K. and a number of major EU institutions participated in this project and provided 14 million euro funds (Wang, 2013).

1.7 Improving biofuel production by phototrophic microorganisms through systems biology

First generation Sanger DNA sequencing (1997) revolutionized science over the past three decades and the current next-generation sequencing (NGS, 2005) technology has opened the doors to the next phase in the sequencing revolution (cf. Figure 6).

Using NGS, scientists are able to sequence entire genomes and to generate extensive transcriptome data from diverse organisms in a timely and cost-effective manner (cf. Table 3). The last developments, consisting of rapid and simplified library preparation methods as well as high throughput sequencing technologies, have extended the application of genomics, once limited to human and several model species, to many other species including photosynthetic eukaryotic algae. In particular, genome data shed light on the complicated evolutionary history of algae that form the basis of the food chain in many environments. In the Eukaryotic Tree of Life, the fact that photosynthetic lineages are positioned in four supergroups has important evolutionary consequences (Kim et al., 2014).

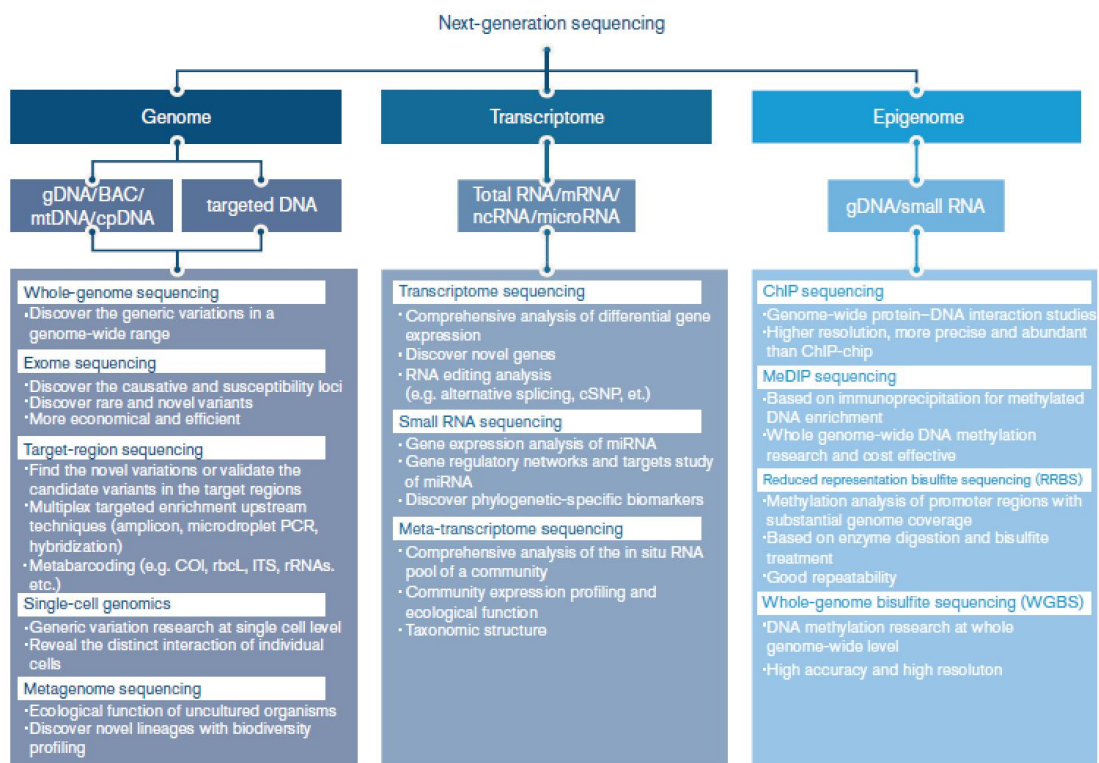


Figure 6. Next-generation sequencing and its applications. The three sources of genetic information: genomes, transcriptomes and epigenomes are shown in different columns with starting materials (i.e. DNA or RNA). Each application is highlighted in a white box with a short description (Kim et al., 2014).

Albeit oleaginous microalgae are promising feedstock for biofuels, the genetic diversity, origin and evolution of oleaginous traits still remains largely unknown. For

these reasons Whole-genome sequence analyses have become critical in the framework of algal research. In addition, several microalgae achieve maximal lipid yields only under stress conditions hindering their growth and providing compositions not ideal for biofuel applications.

Table 3. Comparison of next-generation sequencing platforms (Kim et al., 2014).

Platform/Current company	Sequencing method	Amplification method	Read length (bp)	Capacity for paired reads?	Error rate for raw reads (%)	Reagent cost per Mb	Yield (Mb) / run
Sanger	Chain terminator	PCR	600–900	Yes	0.001–1.0	\$1500	0.01
454 GS FLX +/Roche	Pyrosequencing	Emulsion PCR	600–800	Yes	3–4	\$12.4	700
Solexa GAIIx/Illumina	Synthesis	Solid-phase PCR	150	Yes	0.3–3.8	\$0.12	96 000
SOLID4/Life Technologies	Ligation	Emulsion PCR	50	Yes	0.1–1.0	\$0.12	71 400
HeliScope/Helicos	Synthesis	Single-molecule template	25–35	Yes	3–7	N/A	28 000
Ion Torrent (316 chip)/Life Technologies	Synthesis	Emulsion PCR	100–400	No	~1	\$7.5	>1 000
PacBio/Pacific Biosciences	Synthesis	Real-time single-molecule template	964 average	No	0.001	~\$180	5–10
Starlight/Life Technologies	Synthesis	Single-molecule sequencing with quantum dots	N/A	N/A	N/A	N/A	N/A

1.7.1 Nuclear genome

In spite of their evolutionary importance, algae have been poorly studied due to the relatively small community of dedicated researchers and the limited genetic and molecular tools available for assessing gene function for these highly diverse groups (Tirichine and Bowler, 2011). One exception is the green alga *Chlamydomonas reinhardtii*, a model organism to study photosynthesis and flagella origin using whole genome data (Pröschold et al., 2005; Merchant et al., 2007).

Several whole-genome sequences have become available over the last decade as a result of the growing interest in algae as buffers against global warming and as an alternative source of biofuel and food. The first algal genome of 16 Mbp in size was sequenced in 2004 from the extremophilic red alga *Cyanidioschyzon merolae* (Matsuzaki et al., 2004), followed by the 34 Mbp draft genome from the marine diatom *Thalassiosira pseudonana* (Armbrust et al., 2004).

Over the last few years, different genomes of important algal groups have been sequenced and a large number of other sequencing projects are currently under way. The results are providing new information about the evolution of these organisms and their dominant functions in marine ecosystems. It is also envisaged that there will be an economic benefit deriving from these findings.

From 2005 onwards, the introduction of NGS has dramatically accelerated genome sequencing of algae as shown in Table 4.

Table 4. List of published algal-genome projects (March 2015).

Genome	Size (Mb)	Reference	NGS platform
Chlorarachniophyta			
<i>Bigelowiella natans</i>	94.7	Curtis et al., 2012	Sanger, 454
Chlorophyta			
<i>Chlamydomonas reinhardtii</i>	121	Merchant et al., 2007	WGS
<i>Chlorella variabilis</i> NC64A	46	Blanc et al., 2010	WGS
<i>Chlorella protothecoides</i>	22.9	Gao et al., 2014	WGS
<i>Coccomyxa subellipsoidea</i>	48.83	Blanc et al., 2012	WGS
<i>Volvox carteri</i>	137.68	Prochnik et al., 2010	WGS
<i>Micromonas pusilla</i> CCMP1545	21.9	Worden et al., 2009	WGS
<i>Micromonas</i> sp. RCC299	20.9	Worden et al., 2009	WGS
<i>Ostreococcus tauri</i>	13.2	Palenik et al., 2007	WGS
<i>Ostreococcus lucimarinum</i>	12.6	Derelle et al., 2006	WGS
Cryptophyta			
<i>Guillardia theta</i> CCMP2712	87.2	Curtis et al., 2012	Sanger, 454
Glaucophyta			
<i>Cyanophora paradoxa</i>	70.2	Price et al., 2012	Sanger, 454, GAIIX
Haptophyta			
<i>Emiliana huxleyi</i> CCMP1516	167.68	Read et al., 2013	Sanger
Heterokonta			
<i>Aureococcus anophagefferens</i>	56.66	Gobler et al., 2011	WGS
<i>Ectocarpus siliculosus</i>	214	Cock et al., 2010	
<i>Nannochloropsis gaditana</i>	33.99	Qingdao Inst. Bioe. Biop. Tech.	
<i>Thalassiosira oceanica</i>	92.04	Lommer et al., 2012	454
<i>Thalassiosira pseudonana</i>	32.44	Armbrust et al., 2004	WGS
<i>Phaeodactylum tricornutum</i>	27.45	Bowler et al., 2008	
Rhodophyta			
<i>Chondrus crispus</i>	104.98	Collén et al., 2013	Sanger
<i>Cyanidioschyzon merolae</i>	16.55	Matsuzaki et al., 2004	WGS
<i>Galdieria sulphuraria</i>	13.71	Schönknecht et al., 2013	Sanger, GS20
<i>Porphyra yezoensis</i>	43	Kasuga DNA Res. Inst.	454, GAIIX
<i>Porphyridium purpureum</i>	19.7	Bhattacharya et al., 2013	GAIIX

1.7.2 Chloroplast genome

The endosymbiosis hypothesis was proposed to explain the origin of cellular organelles such as mitochondria and plastids (Mereschkowski, 1905; Sagan, 1967). In evolutionary biology, endosymbiosis specifically refers to a permanent irreversible symbiotic relationship in which a cell lives within another cell. Three major types of endosymbiosis (primary, secondary and tertiary) are recognized in photosynthetic eukaryotes based on the type of endosymbiont: i.e. a prokaryote gives rise to a primary plastid, and a eukaryotic endosymbiont gives rise to a secondary or tertiary plastid (Bhattacharya et al., 2004). Photosynthetic organelles called plastids, however, ultimately trace back to the single primary endosymbiosis where a cyanobacterium resided permanently within a heterotrophic eukaryotic host (Bhattacharya et al., 2004). This transition from a heterotroph to a photoautotrophic eukaryote altered ecosystems significantly because prior to primary endosymbiosis only cyanobacteria were able to conduct oxygenic photosynthesis. The first alga subsequently diverged into three phyla, the Glaucophyta, Rhodophyta and Viridiplantae (green algae and land plants). These three lineages contain two-membrane-bounded plastids, however, glaucophyte plastids retain the peptidoglycan layer between the organelle membranes as in cyanobacteria. Both plastids of rhodophytes and glaucophytes contain chlorophyll a and phycobiliproteins with unstacked thylakoids, whereas chloroplasts of Viridiplantae contain chlorophylls a and b as photopigments with thylakoid stacks (granum).

Plastid-genome size varies greatly across algal groups from 41 kbp (*Micromonas pusilla* CCMP1545, Worden et al., 2009) to 521 kbp (*Floydiella terrestris*, Brouard et al., 2010), but in general ranges from 100–200 kbp. The small size and highly conserved structure of plastid genomes makes the assembly and completion easier than for nuclear genomes (Harrison and Kidner, 2011). Sixty-five plastid genomes are available in the GenBank database including the euglenophyte *Euglena gracilis* (Hallick et al., 1993), which was the first to be sequenced. Approximately 42% of the published plastid genomes are from green algae, 21% are from stramenopiles, and 14% are from red algae (Kim et al., 2014).

The chloroplast genome of algae belonging to the Chlorophyta, displays extraordinary variability in terms of quadripartite structure, global gene organization and intron composition (de Cambiaire et al., 2006).

The chloroplast genomes are particularly useful for phylogenetic reconstruction because of their relatively high and condensed gene content, in comparison to nuclear genomes (Leliaert et al., 2012) and represent a useful tool to better understand the evolution of algae. Indeed, the last few decades have seen a rising interest in chloroplast genome characterization and, as a result, a growing number of algae chloroplasts have been sequenced.

1.7.3 Mitochondrial genome

Mitochondrial genes provide valuable information about closely related species as well as about populations due to their fast evolutionary rate (Brown et al., 1979; Wilson et al., 1985; Zhang et al., 2012). For instance, the *cox1* and *cox2-3* genes are frequently used as mtDNA markers for species identification (Conklin et al., 2009). Despite its utility, complete mitochondrial-genome data from algae are still limited. To date the mitochondrial genomes sequenced are: 20 green algae, 38 stramenopiles, 10 red algae, 2 cryptophytes, 2 glaucophytes and 1 haptophytes (Kim et al., 2014). Mitochondrial genome sequence data from the Chlorophyta, which includes most of the green algal diversification in the Plantae (Keeling et al., 2005), i.e., the chlorophyceans, trebouxiophyceans, ulvophyceans, and prasinophyceans (Lewis and McCourt, 2004), have revealed a fivefold variation in gene content and a genome architecture that varies from compact to expanded (Bullerwell and Gray, 2004; Gray et al., 2004; Pombert et al., 2004, 2006; Popescu and Lee, 2007).

1.8 Genetic tools

After the gene annotations on the sequences, used during the assembly process to determine where the genes start and end, a major part of the work when analyzing the algal assembled sequences consists of performing BLAST searches in the NCBI database. In fact, it helps determining whether an unknown algal sequence, matches with an already known one or, on the contrary, if the obtained sequence refers to genes coding for new proteins. For example, the *Chlamydomonas reinhardtii* genome is available within the NCBI and it can be compared with an unknown

sequence of any algal species. There are different kinds of Blasts according to what type of sequence has to be compared; for a transcript the Blast t is considered while the Blast n is referred to nucleotides and finally Blast p as far as proteins are concerned. The BLAST search in algae bioinformatics is generally used for the functional and phylogenetic comparison of the newly sequenced organisms, with the already known sequences available in the database. The researcher examines the similarities between the sequences, and some of the sequences are downloaded and aligned with the data. For the conserved regions, annotations are transferred to the sequenced data, providing valuable information for the next step in the analysis which is the functional and phylogenetic characterization of the organism.

Sequence analysis is a set of tools allowing one to carry out further and more detailed analyses on the query sequence. In particular, evolutionary analysis, identification of mutations, CpG islands and compositional biases, can be performed. The identification of these and other biological properties are all clues that aid the search to elucidate the specific function of your sequence.

Functional characterization of organism is an advanced application of algae bioinformatics which is carried out to learn more about the algae. It is generally done using programming languages such as R.

A phylogenetic tree is a mathematical structure which is used to model the actual evolutionary history of a group of sequences or organisms. The task of molecular phylogeny is to convert information available in the form of sequences into an evolutionary tree. A great (and ever increasing) number of methods have been described to perform this task. The most commonly used methods can be classified into three major groups; parsimony methods, likelihood methods, and distance methods. In maximum parsimony (MP) analysis, the tree(s) that requires the fewest character state changes is considered the best representation of the true phylogenetic tree (Kitching, 1998). In maximum likelihood (ML) methods, the likelihood of observing a given set of sequence data for a specific substitution model is maximized for each tree topology, and the topology that gives the highest maximum likelihood is chosen as the final tree (Nei and Kumar, 2000). MrBayes is another approach for reconstructing phylogeny and is based on Bayes' theorem. Bayesian methods are closely related to other likelihood methods e.g. the ML analysis which searches for the tree that maximizes the likelihood of the data given an evolutionary model. In distance methods, evolutionary distances are computed for all pairs of taxa, and a

phylogenetic tree is constructed by considering the relationships among these distance values.

CHAPTER 2.

Aim of the work

The exploitation of microalgae as renewable feedstock for the production of biofuels is receiving a rising interest in response to the global concerns regarding the depletion of fossil fuels supplies and the increase of CO₂ levels in the atmosphere. In fact microalgae can be cultivated through relatively simple systems while using costless wastewaters as source of macro-nutrients and finally are capable to accumulate high amounts of lipids when cultivated under suitable operating conditions. In spite of such interest, the existing microalgae-based technology for CO₂ sequestration and biofuels production is still not widespread since it is affected by economic and technical constraints that might have limited the development of industrial scale production systems. Therefore, in view of industrial scaling-up, the current technology should be optimized in terms of lipid productivities. In particular, the creation of new microalgal strains intrinsically characterized by high lipid productivities as well as by a good tolerance to high CO₂ levels is an ambitious goal which might be achieved through genetic manipulations of strains available in nature. Most efforts have focused on stimulating fatty acid biosynthesis by increasing the availability of glycerol as a backbone of lipids, and overexpressing the genes involved in the triacylglycerol (TAG) biosynthesis pathway (Radakovits et al., 2010). However, these efforts have shown limited success, typically yielding less than a 50% increase in lipid content. Another approach adopted to increase cellular lipid content is to block starch synthesis pathways. Several studies have reported the overaccumulation of lipid bodies in starch-devoid mutants of *Chlamydomonas reinhardtii* (Li et al., 2010a, 2010b; Wang et al., 2009; Work et al., 2010). However, impairing starch synthesis does not necessarily result in higher lipid content (Siaut et al., 2011). Therefore, the mechanisms involved in rerouting carbon metabolism from carbohydrates to fatty acids are still unknown.

To this aim, the identification of genes involved in the lipid biosynthetic pathways or in the CO₂ capture mechanisms represents an essential preliminary step to identify the most suitable genetic engineering strategy to improve lipid productivity of existing strains. For these reasons, sequencing of microalgae genomes has started to capture the interest of researchers during the last decade. However, while there are hundreds of strains potentially useful for producing biofuels, only 23 strains have been, so far, completely characterized from the genomic point of view.

The aim of this thesis was to genetically characterize some among the most promising algal strains in view of their commercial application for the production of

biofuels or CO₂ capture. In particular 6 microalgal strains have been considered in this work: *Chlorella sorokiniana* SAG 211-8k, *Pseudochloris wilhelmii* SAG 55.87, *Monodus subterraneus* SAG 848.1, *Scenedesmus obliquus* SAG 276-1, *Chlorella variabilis* NC64A and *Tetraspora* sp. SCCA024. However, to date, only the genetic characterization of chloroplast and mitochondrial DNA of the strains *Chlorella sorokiniana* and *Chlorella variabilis* has been completed. The latter ones are very promising strains and in particular *C. sorokiniana* may be exploited for wastewater treatment because of its high specific growth rate and high affinity for acetic acid, an important contaminant found in wastewaters. In addition, it has been demonstrated in the literature that this strain is capable to fix CO₂ with a good efficiency thus being potentially useful for CO₂ capture from flue gases (Wan et al., 2011). Finally, *C. sorokiniana* was found to be rich in proteins, several important minerals, lipids and anti-oxidants such as astaxanthin and β -carotene.

On the other hand, *C. variabilis* is a model system to investigate virus/algal interactions and is potentially useful to study plant hormones (and receptors). A deeper understanding of the role of plant hormone molecules in green algae as well as of their synthesis and perception would possibly lead to the selection and improvement of better algal strains that could improve agricultural practices in developing countries (Stirk et al., 2002), result in better production of biodiesel, and enhance the quality and quantity of nutrient supplements (proteins, vitamins, etc.) (Blank et al., 2010).

Subsequently, similarities and differences between strains belonging to green algae will be assessed with a specific focus on the identification of genes over/under expressed in those strains which are capable of accumulating high amounts of lipids. Finally, a suitable genetic manipulation strategy will be first identified and then implemented in order to improve the lipid productivity of wild strains.

CHAPTER 3.

The phylum Chlorophyta

3.1. Trebouxiophyceae

The green algae represent one of the most successful groups of photosynthetic eukaryotes, but compared to their land plant relatives, surprisingly little is known about their evolutionary history. This is in great part due to the difficulty of recognizing species diversity behind morphologically similar organisms (Lemieux et al., 2014). Early hypotheses on green algal phylogeny were based on morphology and ultrastructural data derived from the flagellar apparatus and processes of mitosis and cell division. These ultrastructural features, which apply to most green algae, supported the existence of the Streptophyta and Chlorophyta. The Chlorophyta includes the most investigated species of green algae. The Streptophyta consist of charophytes, a paraphyletic assemblage of freshwater algae, and land plants (Leliaert et al., 2012). On the other hand, four distinct classes can be recognized within the Chlorophyta, i.e. the predominantly marine, unicellular, Prasinophyceae; the predominantly marine and morphologically diverse Ulvophyceae; and the freshwater or terrestrial, morphologically diverse Trebouxiophyceae (=Pleurostrophyceae) and Chlorophyceae (cf. Figure 7).

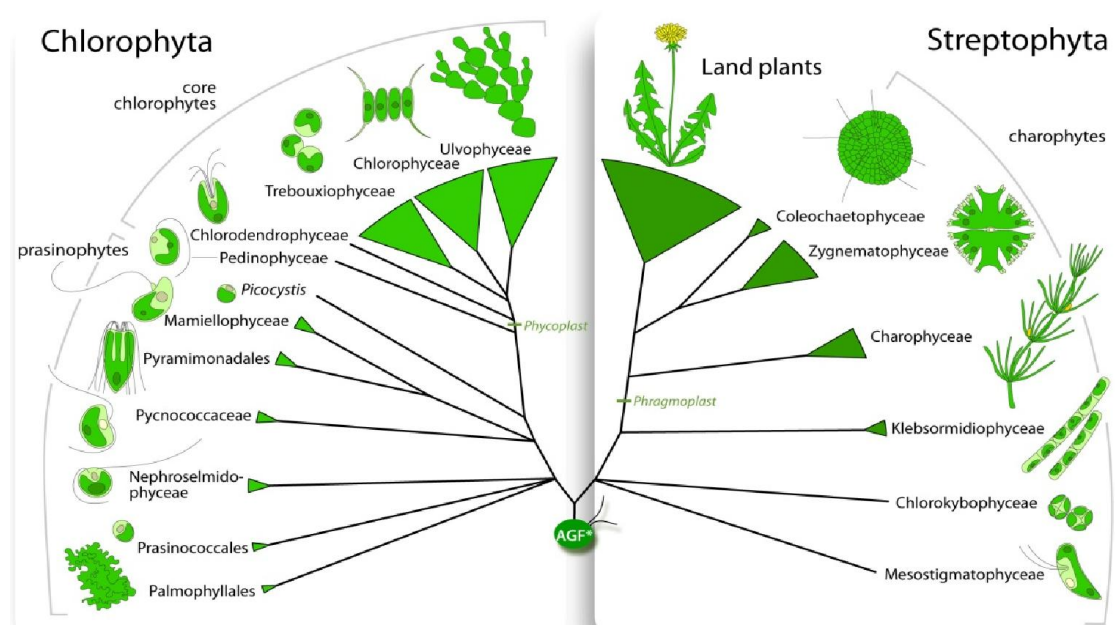


Figure 7. Overview phylogeny of the green lineage (Leliaert et al., 2012).

It has been postulated that the Prasinophyceae have given rise to the Ulvophyceae, Trebouxiophyceae and Chlorophyceae (UTC). Later, phylogenetic analyses based on the nuclear-encoded small subunit rRNA gene (18S rDNA) largely corroborated these hypotheses. Moreover, the class Trebouxiophyceae “sensu stricto” is not a monophyletic group (Lemieux et al., 2014). This species-rich class displays

remarkable variation in both morphology (comprising unicells, colonies, filaments and blades) and ecology (occurring in diverse terrestrial and aquatic environments). Moreover, there are some genera that include unicellular nonflagellated parasites/pathogens that still retain vestigial plastids (Figueroa-Martinez et al., 2014). Several species (e.g. *Trebouxia*, *Myrmecia* and *Prasiola*) participate in symbioses with fungi to form lichens and others (e.g. *Chlorella*, *Coccomyxa*, and *Elliptochloris*) occur as photosynthetic symbionts in ciliates, metazoa and plants.

Members of Trebouxiophyceae reproduce asexually by autospores or zoospores. Sexually reproductive stages have not been observed directly in any of the trebouxiophyte algae (Lewis and McCourt, 2004). The Trebouxiophyceae also involves species that have lost photosynthetic capacity and have evolved towards free-living or parasitic heterotrophic lifestyles (e.g. *Prototheca* and *Helicosporidium*). Aside from their intrinsic biological interest, trebouxiophycean algae have drawn the attention of the scientific community because of their potential utility in a variety of biotechnological applications such as the production of biofuels or other molecules of high economic value (Hannon et al., 2010, Mata et al., 2010).

3.2. The Chlorellales family

The Trebouxiophyceae class includes the order Chlorellales (cf. Figure 8) which mostly involves unicellular coccoids of minute size that thrive in an extremely broad range of habitats (Friedl and Rybalka, 2012). This order brings together the Chlorellaceae family (Brunnthaler, 1915), that have been recently reorganized into some sister groups: the *Chlorella* and the *Parachlorella* clades (Krienitz et al., 2004), the *Auxenochlorella* clade (Pröschold and Leilaert, 2007) and the *Marvania* clade (Lemieux et al., 2014).

Is interesting to note that this order includes photosynthetic and non photosynthetic organisms. The loss of photosynthesis has occurred several independent times. The genera *Prototheca* and *Helicosporidium* include unicellular nonflagellated parasites/pathogens that still retain vestigial plastids. Members of the genus *Prototheca* are ubiquitous opportunistic animal pathogens that can be found in diverse habitats, such as soil detritus, fresh and brackish water, and plant- and animal-derived foods for human consumption. *Helicosporidium* infections are

common in insects, mites, trematodes and cladocerans (Tartar, 2013). The loss of photosynthesis probably occurred in the ancestors of *Prototheca* and *Helicosporidium* during their shift from mixotrophy to parasitism (Pombert et al., 2014). It is unclear, however, if these two closely related genera lost their photosynthetic abilities independently.

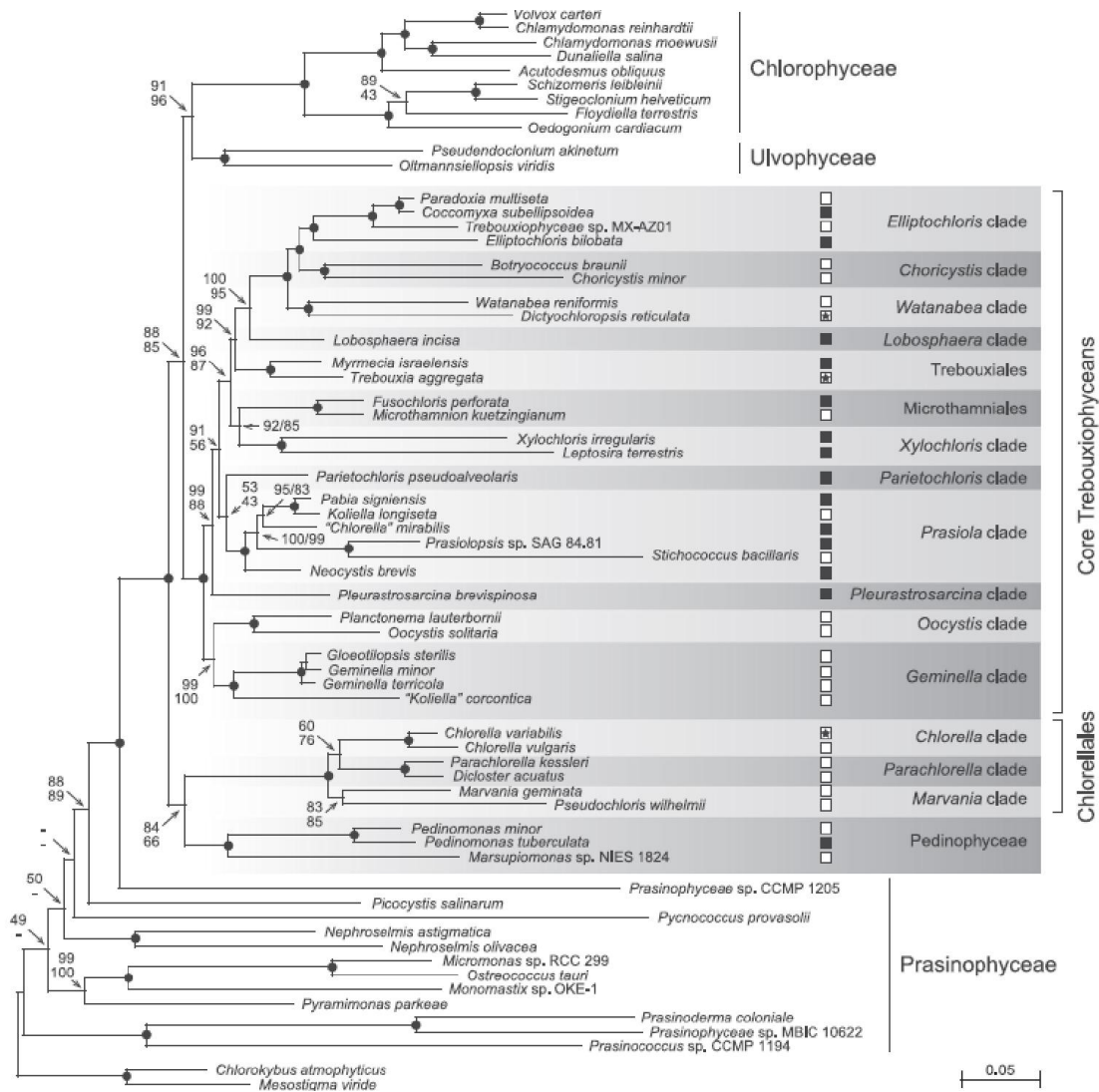


Figure 8. Phylogeny of Trebouxiophyceae (Lemieux et al., 2014).

Recent phylogenetic analyses of nuclear 18S rRNA and β -tubulin data have shown that some *Prototheca wickerhamii* isolates are more closely related to photosynthetic taxa (e.g. *Chlorella* spp.) than to other *Prototheca* species (Mancera et al., 2012). These data suggest that the loss of photosynthesis has occurred at least twice in the evolution of parasitic/pathogenic Chlorellales. Moreover, the mixotrophic capabilities of various *Chlorella* species (Lee et al., 1996), which are able to use different organic compounds (e.g. glucose, glycerol, ethanol, acetate, and butyrate) as carbon sources, imply that nonphotosynthetic Chlorellales probably evolved from

commensals (e.g. saprophytes; similar to *Prototheca* species living in animal integumentary tissues) that ultimately harnessed their heterotrophic abilities to invade novel ecological niches (Figueroa-Martinez et al., 2014).

The genus *Chlorella* was first delineated by Beyerinck in 1890 as a genus of four species, two of which were previously described by Brandt (1881, 1882) as *Zoochlorella*. Common in nature, the genus is one of the most conspicuous of those green unicells showing no motility during reproduction. Over a hundred algal isolates were originally assigned to the genus *Chlorella*, but their taxonomy classification has long remained unreliable because of their lack of conspicuous morphological characters (Blanc et al., 2010).

Among the different microalgae, *Chlorella* species are of interest because of their high productivity, high lipid content, and resistance to the high light conditions typical of photobioreactors. Several freshwater species of *Chlorella* have been extensively used commercially over the past 40 years as a food and feed supplement on account of their rapid growth and tolerance over a wide range of temperature and culture conditions. However, the economic feasibility of growing algae at an industrial scale is yet to be realized, in part because of biological constraints that limit biomass yield. On the basis of biochemical and molecular data, presently consists of five “true” *Chlorella* species: *C. vulgaris* Beyerinck, the archetype (lectotype) of coccoid green algal ‘balls’ (Krienitz et al., 2015), *C. lobophora* Andreyeva, *C. sorokiniana* Shihira et Krauss, *C. heliozoae* Pröschold et Darienko and *C. variabilis* Shihira et Krauss (Huss et al., 1999; Krienitz et al., 2004; Bock et al., 2011).

3.2.1. *Chlorella sorokiniana*

Chlorella sorokiniana is a non-motile unicellular alga (cf. Figure 9). This species is named after its isolator, Dr. Constantine Sorokin, who has published more detailed information for this species of *Chlorella*. The suitable temperature for the growth of this alga was found to be 30 °C. However, in the range of 30–40 °C, it had nearly similar growth profile. Temperature below the optimum was found to decrease the growth rate of the microorganism. The pH 7.5 was found favourable for the growth of this strain, while below pH 6.5, growth of the cell declines significantly.

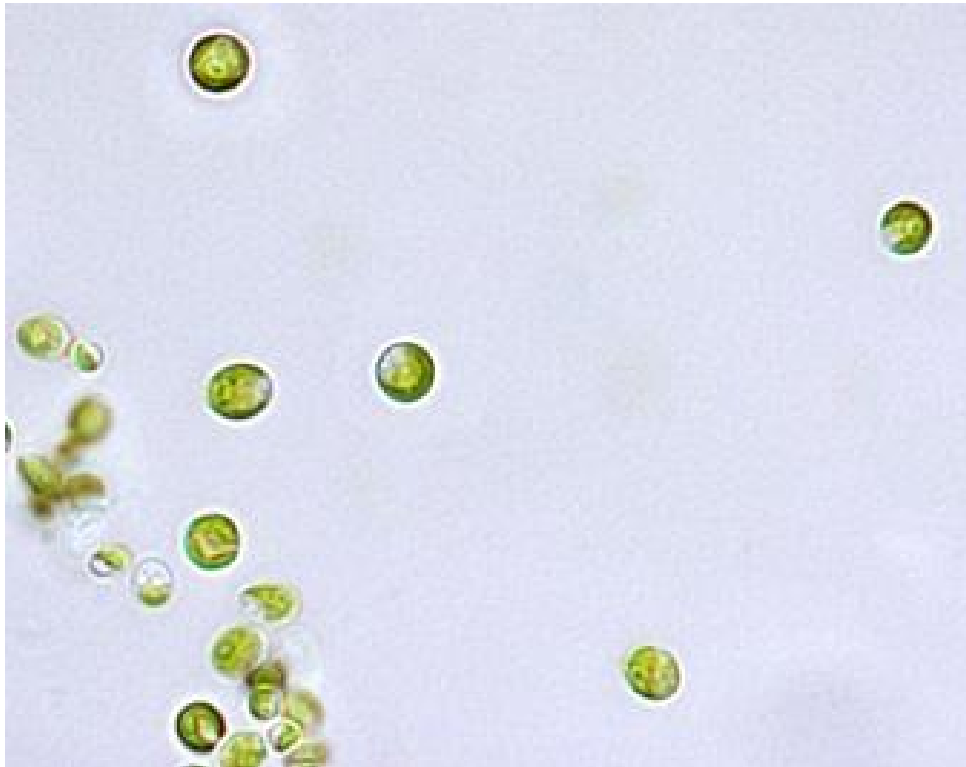


Figure 9. *Chlorella sorokiniana*.

C. sorokiniana has demonstrated good nutrient removal capability (Kim et al., 2013; Ogbonna et al., 2000) as well as good lipid accumulation potential (Qiao and Wang, 2009; Zheng et al., 2013, Lizzul et al., 2014). In addition, de-Bashan et al. (2008) reported the ability of this thermo-tolerant alga to grow in wastewaters hostile to other algal species. Griffiths and Harrison (2009) have also compared the lipid productivity amongst 55 algal species and identified *C. sorokiniana* as a potential candidate for achieving higher lipid productivity. This strain can take advantage of organic carbon sources, in addition to carbon dioxide, to grow under mixotrophic conditions with higher biomass yields than the corresponding ones observed under photoautotrophic conditions (Wan et al., 2011). Moreover, a recent work by Rosenber et al. (2014) investigated the effects of heterotrophic and mixotrophic growth on lipid biochemistry of *Chlorella* species, thus demonstrating that different lipid compositions can be achieved by cultivating *C. sorokiniana* under the trophic conditions above (cf. Figure 10). Ultimately, the desired lipid composition and content might be varied by suitably tuning the cultivation conditions.

In addition, *C. sorokiniana*, is playing an important role as food and feed because of the multiuse of its biomass previously known to be a rich source of carbohydrate, vitamins, and proteins. The high protein content makes it a suitable raw material for the production of single cell protein (Mahasneh, 1997) while the high vitamin

content makes it a suitable feed for aquaculture systems (Gapasin et al., 1998). Moreover, when cultivated under sulfur deprived condition (Chader et al., 2009) *C. sorokiniana* is capable to produce biohydrogen which in turn represents a clean source of energy. Additionally, it has been used for the production of commercially important antioxidants like lutein, α/β carotene, α/β tocopherol, zeaxanthin (Matsukawa et al., 2000).

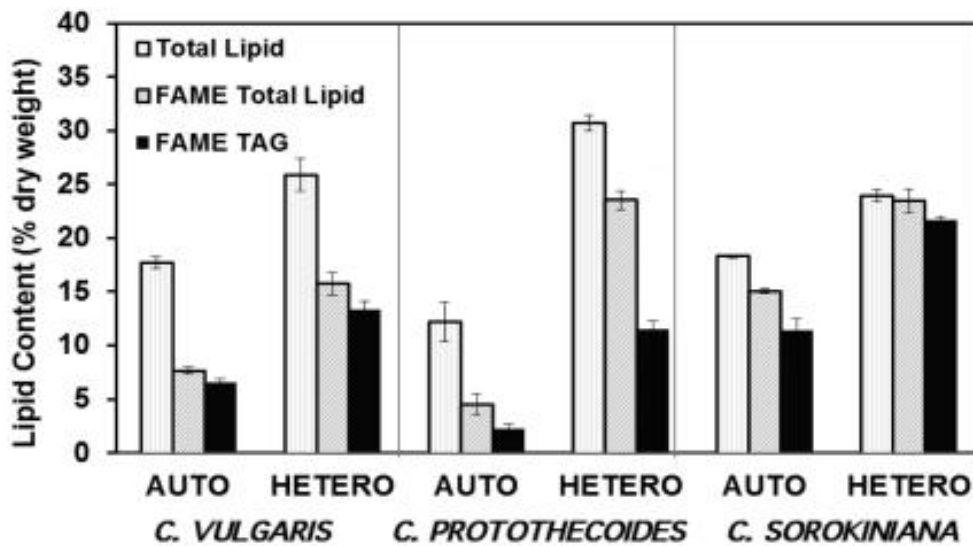


Figure 10. Distribution of total lipid extracts as fatty acids and TAG in three *Chlorella* strains (Rosenberg et al., 2014).

These characteristics make this strain, as well as most of the species belonging to the *Chlorella* genus, a particularly promising feedstock in the biotechnological, environmental and energy sectors.

In spite of this, the commercial exploitation of *C. sorokiniana* is still not widespread since its large scale production might be affected by technical constraints mainly arising from the still low lipid productivity achievable through the current cultivation technologies (Concas et al., 2012). In this regard, the knowledge of *C. sorokiniana* genome represents the first step towards the identification of suitable genetic engineering strategies aimed to increase its lipid productivity and thus to overcome the limitations described above.

3.2.2. *Chlorella variabilis* NC64A

Chlorella sp. NC64A, renamed *Chlorella variabilis* (Ryo et al., 2010), is a unicellular photosynthetic green alga member of the true *Chlorella* genus (Trebouxiophyceae) (Huss et al., 1999; Krienitz et al., 2004; Pröschold et al., 2011;

Bock et al., 2011). *Chlorella variabilis* Shihira et R. Krauss (cf. Figure 11) is an intracellular photobiont of *Paramecium bursaria* (Karakashian and Karakashian, 1965) but, as well as the other endosymbiotic *Chlorella* species, retains its ability to grow independently (Kodama and Fujishima, 2009). Although *Chlorella* species have been assumed to be asexual and nonmotile, the NC64A genome encodes all the known meiosis-specific proteins and a subset of proteins found in flagella. We hypothesize that *Chlorella* might have retained a flagella-derived structure that could be involved in sexual reproduction.

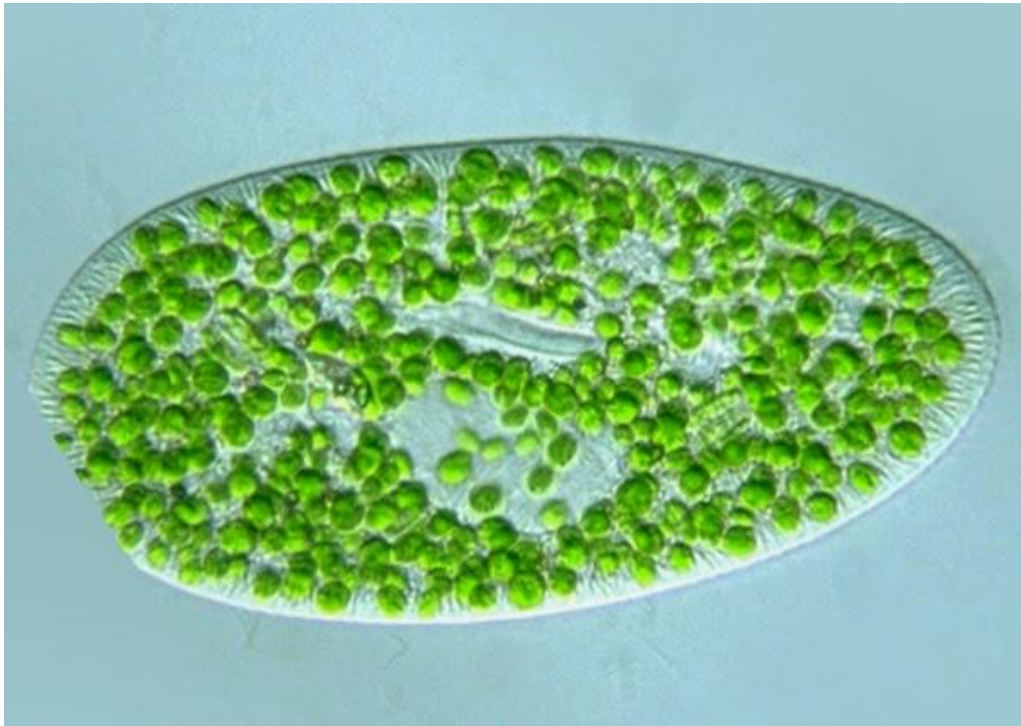


Figure 11. *Paramecium bursaria* with symbiotic *Chlorella variabilis*.

The microalga *C. variabilis* is particularly interesting since it is the host for a family of large (>310 kb), plaque-forming, dsDNA viruses (called *Chlorella* viruses) that are found in freshwater throughout the world. For this reason, this symbiotic alga represents a model system for the study of virus/algal interactions (Blanc et al., 2010). This model system is relevant because the about 56% of the worldwide photosynthetically fixed CO₂ is used by phytoplankton, including cyanobacteria and eukaryotic microalgae. However, about 20% of these photosynthetic organisms are typically infected by a virus. Thus, while viruses play a significant role in global carbon/nitrogen cycles, their role has been relatively neglected by scientists who model carbon/nitrogen cycles. Moreover, most known viruses infecting microalgae are large dsDNA viruses evolutionarily related to the *Chlorella* viruses. The knowledge of the *C. variabilis* genome will, for the first time, allow to identify in

which genes from both a dsDNA virus exceeding 300 kb and its host are available for analysis.

In addition, similar to *Chlamydomonas reinhardtii*, *Chlorella variabilis* exhibits [FeFe]-hydrogenase (HYDA) activity during anoxia. In contrast to *C. reinhardtii* and other chlorophycean algae, which contain hydrogenases with only the HYDA active site (H-cluster), *C. variabilis* is the only known green alga containing HYDA genes encoding accessory FeS cluster binding domains (F-cluster) (Meuser et al., 2011).

Moreover, it is likely that fermentative product secretion plays a significant role in providing reduced carbon substrates and or H₂ to other organisms within a microbial community or symbiosis.

In conclusion, like other microalgae, there is an increasing interest in using *Chlorella* in a variety of biotechnological applications, such as biofuels (Schenk et al., 2008), sequestering CO₂ (Chelf et al., 1993), producing molecules of high economic value, or removing heavy metals from wastewaters (Rajamani et al., 2007).

CHAPTER 4.

Kinetic characterization of the growth and lipid accumulation of *Chlorella sorokiniana*

4.1. Strain and culture conditions

Chlorella sorokiniana (SAG 211-8k; authentic strain) was obtained from Sammlung von Algenkulturen, University of Göttingen (SAG), Germany (<http://www.epsag.unigoettingen.de/html/sag.html>). Unialgal culture was maintained under axenic conditions in shaken flasks illuminated by a photon flux of 90-100 $\mu\text{E m}^{-2} \text{ s}^{-1}$ and containing modified WARIS-H culture medium (McFadden and Melkonian, 1986) without soil extract or Bold Basal Medium (Bischoff and Bold, 1963). A light/dark photoperiod of 12 h was assured during microalgae growth which was prolonged up to log phase.

The inoculation procedure has been carried out as follows. First a single colony of microalgae cells has been taken from agar tube and then placed in a 50 mL flask containing 25 mL of standard BBM growth medium. Subsequently, the resulting cultures were maintained at 25 ± 1 °C, and illuminated by a photon flux density of $\sim 80 \mu\text{mol m}^{-2} \text{ s}^{-1}$ for 2 weeks using a 12:12 h light–dark photoperiod. Finally, a suitable amount of the culture above has been transferred into batch photobioreactors containing 1 or 2 L of BBM medium in order to obtain culture characterized by an optical density of about 0.05. The batch growth experiments have been carried out by using the obtained culture and providing a photon flux of about $\sim 80 \mu\text{mol m}^{-2} \text{ s}^{-1}$ with a photoperiod equal to 12:12 h light/dark, under continuous stirring and an air flux. As far as the growth experiments in the Biocoil photobioreactor, the cultures grown in the batch PBRs have been used as the inoculum while the remaining operating conditions are kept similar to the ones already specified for the batch experiments.

4.2. Growth kinetics of *C. sorokiniana* in batch photobioreactors

Unialgal culture was maintained under axenic conditions in shaken flasks containing Bold Basal Medium whose composition is reported in the following Table 5.

Table 5. Chemical composition of the BBM growth medium

Component	Concentration [g L⁻¹]
KH ₂ PO ₄	0.175
CaCl ₂ · 2H ₂ O	0.025
MgSO ₄ · 7H ₂ O	0.075
NaNO ₃	0.25
K ₂ HPO ₄	0.075
NaCl	0.025
FeSO ₄ · 7H ₂ O	0.0049
Na ₂ -EDTA	0.05
KOH	0.031
H ₃ BO ₃	0.01142

Trace Metal Solution	Concentration [g L⁻¹]
MoO ₃	0.71
MnCl ₂ · 4H ₂ O	1.44
ZnSO ₄ · 7H ₂ O	8.82
CuSO ₄ · 5H ₂ O	1.57
Co(NO ₃) · 2.6H ₂ O	0.49

Currently, a number of studies are being carried out in order to evaluate the effect of different culture medium compositions on the lipid yield and biomass production. In particular, since nitrogen starvation has been recognized to trigger lipid metabolism while reducing microalgae growth rate (Xu et al., 2001, Hu et al., 2008; Lin and Lin, 2011), one of the most challenging tasks is to identify the concentration of nitrogen which represents the optimal compromise among lipid accumulation and growth rate lowering. For this reason, two different concentrations of nitrate, with respect to standard BBM, have been tested in order to identify the optimal one and to determine the influence of this element in the production of lipids.

The initial cell density of each experiment was standardized at 0.06 optical density (O.D.). *C. sorokiniana* was cultured in 2 L bottles containing 1,8 L sterilized culture medium, with agitation of 500 rpm with constant aeration. The growth was monitored through spectrophotometric measurements (Genesys 20 spectrophotometer, Thermo Fisher Scientific Inc. Waltham) of the culture media O.D. with 1 cm light path. The wavelengths corresponding to the absorption of Chlorophyll a (663 nm), Chlorophyll b (643 nm) and carotenoids (439 nm) have been analyzed since they are good indicators of algal biomass concentration at least in the range of cell densities usually observed in laboratory assays. The structural difference between chlorophyll a and chlorophyll b is that the former is a blue/green pigment with maximum absorbance from 660 to 665 nm while the latter one is a green/yellow pigment characterized by a maximum absorbance from 642 to 652 nm (Hosikian et al., 2010).

The biomass concentration C_b ($\text{g}_{\text{dw}} \text{L}^{-1}$) was calculated from O.D. measurements using a suitable C_b vs. O.D. calibration curve (cf. Figure 12) which was obtained by gravimetrically evaluating the biomass concentration of known culture medium volumes that were previously centrifuged at 4000 rpm for 10 minutes and dried at 60 °C for 24 hours.

The pH was daily measured by pH-meter basic 20 (Crison, Barcelona, Spain). For the sake of reproducibility, each experimental condition was investigated at least in duplicate.

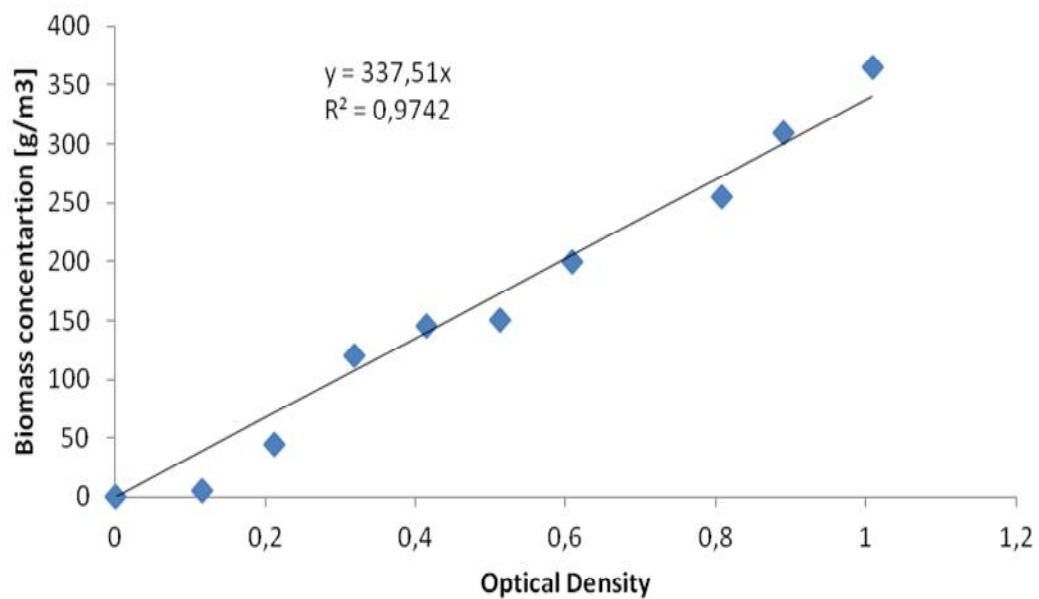


Figure 12. Calibration line showing the correlation between wet weight of biomass pellet subjected to disruption and the corresponding dry weight content.

By using the above calibration curve it was possible to follow the time evolution of the growth of *C. sorokiniana* in terms of biomass concentration.

In particular, as it can be observed from Figure 13, *C. sorokiniana* grows in standard BBM until a stationary phase is reached after about 25 days. After this period of time a stationary biomass concentration of about 350 g m^{-3} is achieved.

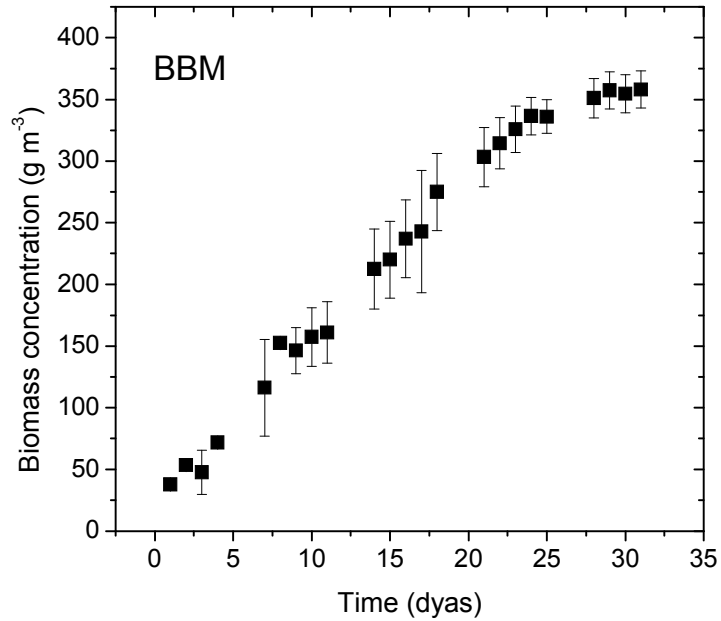


Figure 13. *Chlorella sorokiniana* grown in standard BBM.

On the contrary, when the alga is grown under an increased content of nitrogen, i.e., 5N (cf. Figure 14), the stationary phase is achieved after about 40 days, when the biomass concentration is about 550 g m⁻³ as a result of the higher concentration of nitrogen available in solution. Therefore, under such operating condition, a greater biomass concentration can be achieved with respect to corresponding one attained when using standard BBM medium.

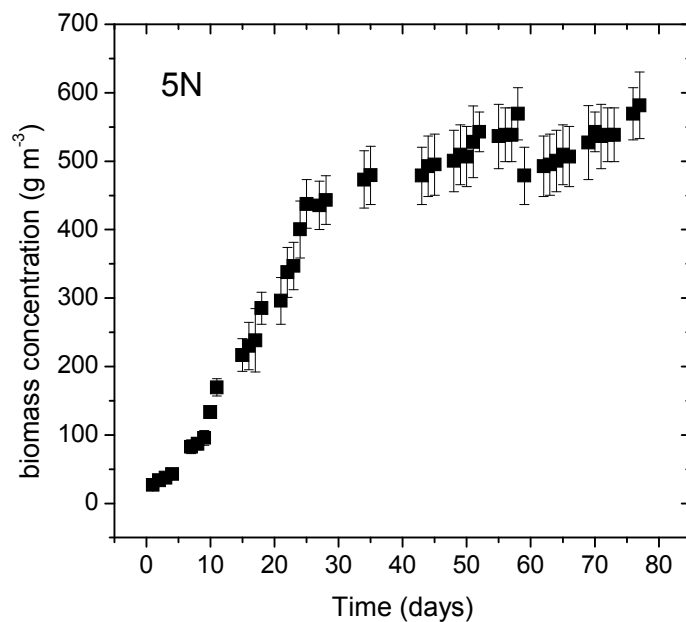


Figure 14. *Chlorella sorokiniana* grown in nitrogen supplement.

Finally as it can be observed from Figure 15, when the nitrogen concentration is reduced five times with respect to typical one of BBM standard medium, growth of

C. sorokiniana takes place until a stationary phase is reached after about 10 days. After this period of time, the biomass concentration starts to decrease due to a probable onset of cell death phenomena. As it can be observed, the biomass concentration achieved at the stationary phase is very low, i.e. about 200 g m^{-3} , as a result of the lower nitrogen concentration available in solution.

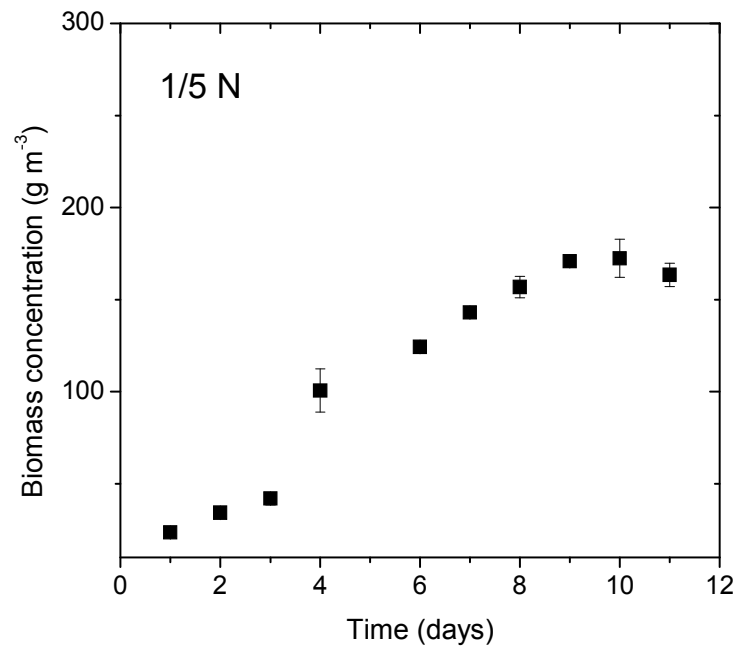


Figure 15. *Chlorella sorokiniana* grown in nitrogen starvation.

4.3. Growth kinetics of *C. sorokiniana* in the BIOCOIL photobioreactor

Growth of *C. sorokiniana* was carried out in a 6 L helical tubular photobioreactor coupled with a degasser system, as described in the literature (Concas et al., 2010). Briefly, the light collector of the photobioreactor consisted of 66 m transparent polyurethane tubing arranged around a circular metal frame. It was internally illuminated by three 60W white fluorescent lamps providing a light intensity of $100 \mu\text{E m}^{-2} \text{ s}^{-1}$ for a light-dark photoperiod of 12 h. Liquid circulation in the light collector was assured by a peristaltic pump. The degasser unit was a 1 L bubble column which allowed to remove photosynthetic oxygen by exposing the broth to atmosphere. Once the culture reached the stationary growth phase the photobioreactor was operated in fed-batch mode. The evolution of microalgae concentration during cultivation of *C. sorokiniana* in the BIOCOIL photobioreactor

fed with air is shown in Figure 16. At regular intervals the withdrawals made during the operation in fedbatch mode.

It can be observed that, after an exponential growth of 13 days, the culture reached the stationary phase when the biomass concentration was about 2 g L^{-1} . Once the steady state was attained, the photobioreactor was operated in fed-batch mode. In fact, starting from the 15th day of culture, suitable amounts of culture were periodically withdrawn and then replaced by an equal volume of fresh medium.

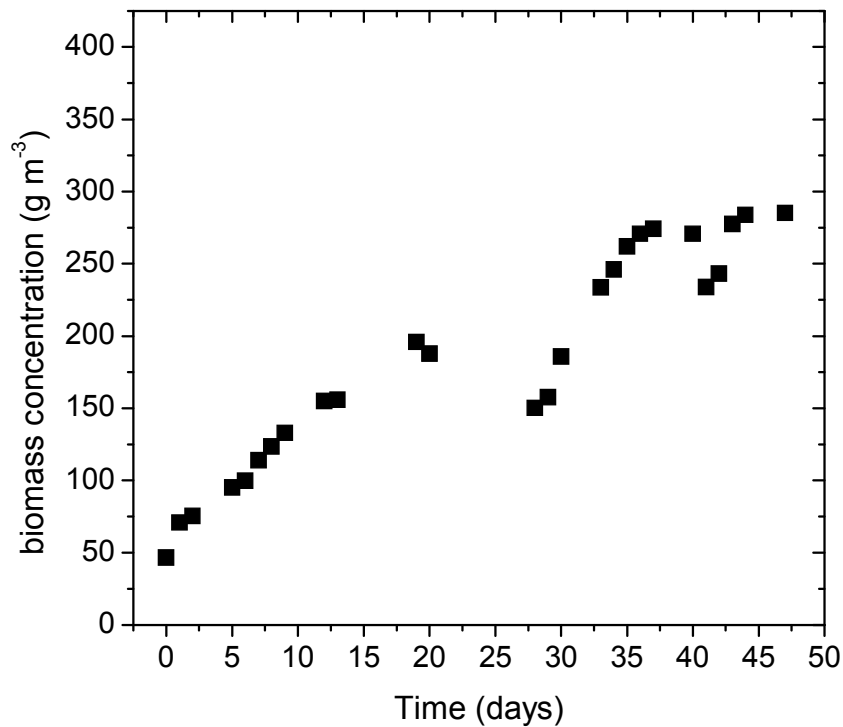


Figure 16. Growth of *Chlorella sorokiniana* in helical tubular photobioreactor in fedbatch mode.

As shown in Figure 16, after each withdrawal, the biomass concentration decreased and then started to increase as a result of the higher nutrient availability and the diminished concentration of toxic catabolites. The wet biomass harvested during each withdrawal cycle was centrifuged and then subjected to the different experiments.

4.4. Spectrophotometric analysis of lipid content during growth.

A rapid colorimetric method has been adopted to quantify the lipid content of microalgae during their growth. It is based on the capability of sulpho-phospho-vanillin (SPV) to react with lipids to generate a reaction product which is characterized by a typical pink color whose intensity can be quantified using spectrophotometric methods, i.e. by measuring absorbance at 530 nm (cf. Figure 17). Phosphovanillin reagent was prepared by initially dissolving 0.6 g vanillin (Sigma-Aldrich, St. Louis, MO, USA) in 10 ml absolute ethanol; 90 ml deionized water and stirred continuously.

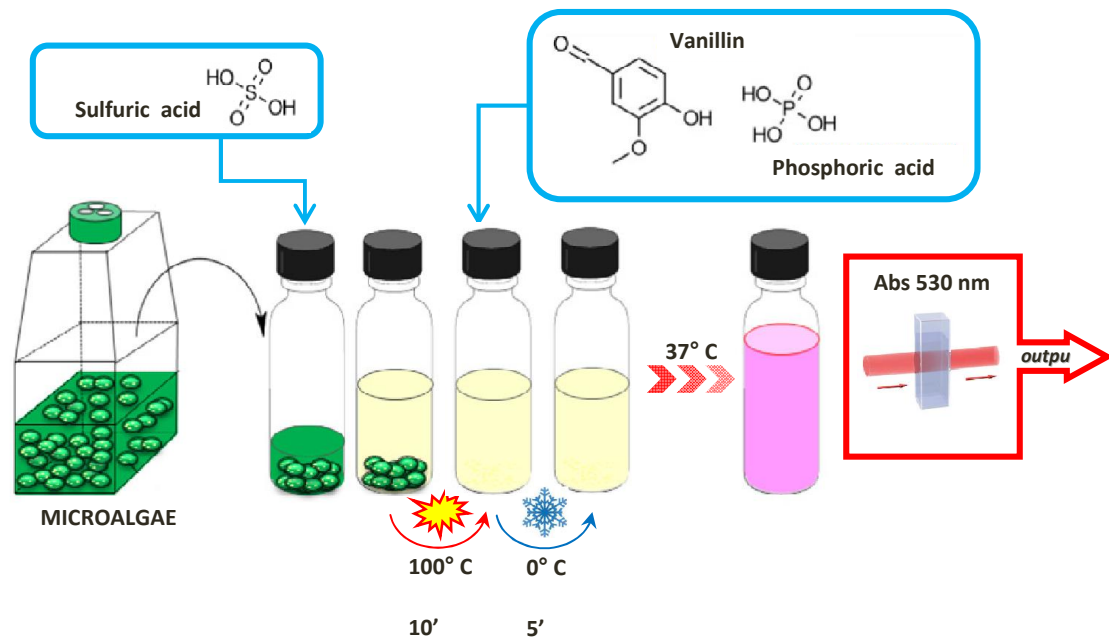


Figure 17. Schematic representation of colorimetric method for lipids quantification which employs sulpho-phospho-vanillin (Mishra et al., 2013).

Subsequently 400 ml of concentrated phosphoric acid was added to the mixture, and the resulting reagent was stored in the dark until use. To ensure high activity, fresh phospho-vanillin reagent was prepared shortly before every experiment run. For SPV reaction of the algal culture for lipid quantification, a known dry amount of biomass was re-suspended in 100 µl deionized water. The samples were sonicated for 30 minutes and then 2 mL of concentrated (98%) sulfuric acid was added to the sample and was heated for 10 minutes at 100 °C, and was cooled for 5 minutes in ice bath. 5 mL of freshly prepared phospho-vanillin reagent was then added, and the sample was incubated for 15 minutes at 37 °C incubator shaker at 200 rpm. Absorbance reading at 530 nm was taken in order to quantify the lipid within the sample.

The calibration curve was obtained as follows. About 20 mg of commercial canola oil was dissolved in 10 ml chloroform (final concentration, 2 mg/ml) and the resulting solution was stored at $-20\text{ }^{\circ}\text{C}$ before use.

A known amount of lipids (oil) was added in the empty tube which was then kept at $60\text{ }^{\circ}\text{C}$ for 10 minutes in order to evaporate the solvent. Subsequently, $100\text{ }\mu\text{l}$ of water was added to the lipid standard and the resulting solution was then subjected to the spectrophotometric analysis above described. Several samples where different amounts of lipids were used, were assessed thus allowing to obtain the calibration curve shown in Figure 18.

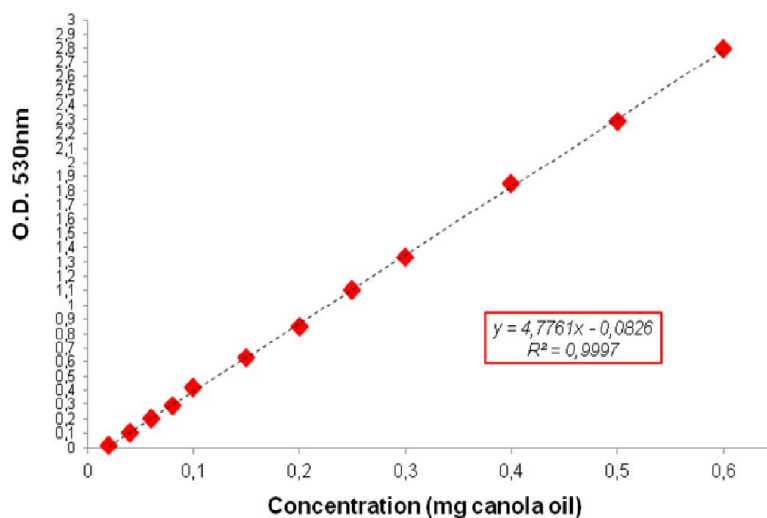


Figure 18. Calibration curve using of lipid standard solutions (canola oil).

As it can be observed a quite good correlation is achieved between optical density at 530 nm and lipid content of the sample. Therefore, the use of this calibration line allows to confidently assessing the lipid content of a sample by using spectrophotometric techniques.

The SPV method was then used to evaluate the lipid content of *Chlorella sorokiniana* during its batch growth. Figures 19-20-21 show the time evolution of lipid content of *C. sorokiniana* when it was cultivated under normal, five folds increased and five folds reduced, respectively nitrogen concentration in the growth medium. As it can be observed, the lipids content of algae decreases as the nitrogen content in solution is increased. In particular, when cultivated in standard BBM medium (cf. Figure 19), the lipid content of *C. sorokiniana* increases from an initial value of about 12 % wt to about 26 % wt. The time trend of the lipid content growth is quite similar to the one already observed for the corresponding biomass growth (cf. Figure 13). In fact, it can be observed that after an initial growth, the lipid

accumulation rate start to decrease and finally reaches a kind of stationary phase after about 25 days of cultivation.

On the other hand, when *C. sorokiniana* is cultivated under increased concentration of nitrogen, its lipid content increases with a smaller rate reaching a final stationary value of about 14 %wt (cf. Figure 20). It thus confirms that high nitrogen contents result in a kind of inhibition of lipid synthesis. Such a phenomenon is probably due to the fact that algal metabolism, when high amounts of nitrogen are available, is preferably shifted towards the protein synthesis rather than the synthesis of energy storage compounds such as lipids.

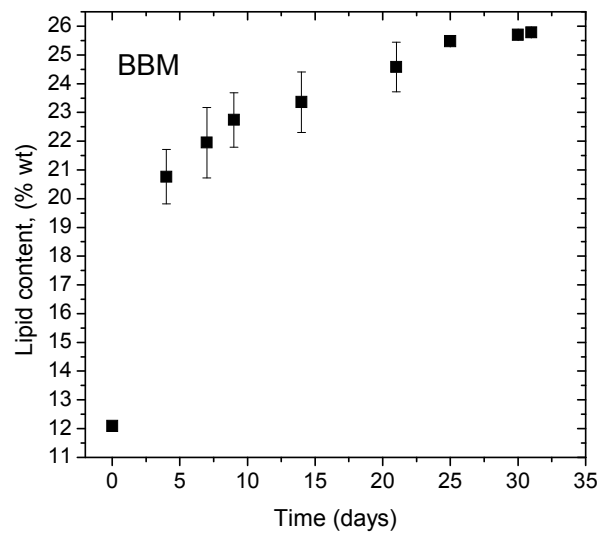


Figure 19. Lipids amount in *Chlorella sorokiniana* growth in standard BBM medium in batch photobioreactor.

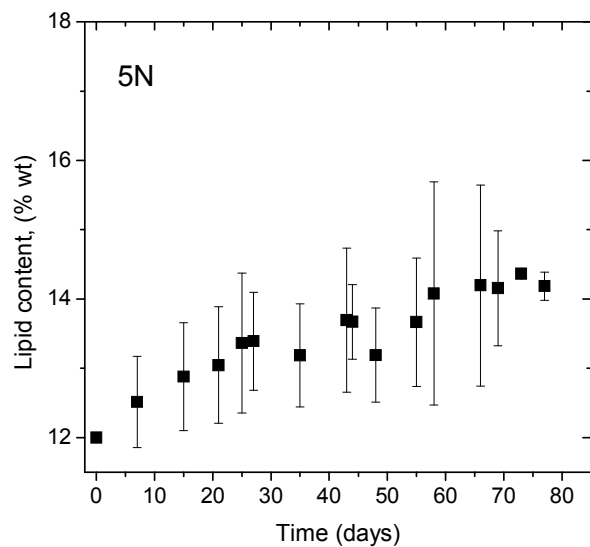


Figure 20. Lipids amount in *Chlorella sorokiniana* growth in modified BBM with nitrogen supplementation in batch photobioreactor.

Such result is somehow confirmed by the results obtained when growing *C. sorokiniana* under a reduced concentration of nitrogen. In fact, under these conditions, a significant lipid content, i.e. about 23 %wt, is achieved after only 10 days of cultivation (cf. Figure 21).

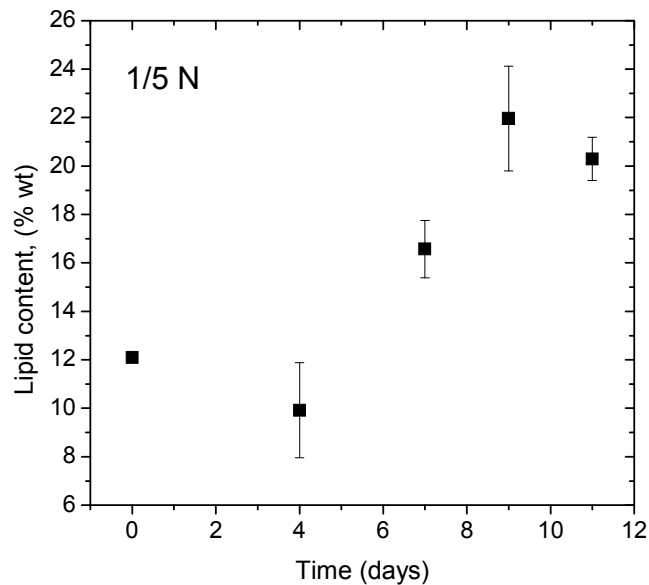


Figure 21. Lipids amount in *Chlorella sorokiniana* growth in modified BBM with nitrogen starvation in batch photobioreactor.

4.5. Lipid extraction

Several methods for lipid extraction from microalgae are currently under investigation at the laboratory scale but solvent extraction appears to be, so far, the only viable way for performing lipid extraction at the industrial scale (Chisti, 2007). Typically, solvent extraction is carried out by contacting microalgal biomass with an organic eluting solvent which diffuses through the cell wall/membrane into the cytoplasm and interacts, through van der Waals type bindings, with the neutral lipids by forming organic solvent-lipids complexes. The latter ones, driven by a concentration gradient, counter-diffuses across the cell wall towards the bulk solvent from which they can be collected to be further processed (Halim et al., 2011; 2012). Solvent extraction of algal lipids can be performed starting from both wet and dry microalgal biomass and, depending upon which option is chosen, specific pre-treatments should be carried out. In fact, lipid extraction from untreated wet biomass

is characterized by low yields due to the immiscibility of water with the organic solvents. Therefore, when solvent extraction is applied to wet biomass, the microalgal cells tend to remain in the water phase due to their surface charges and thus they cannot contact the organic solvent phase which is able to extract lipids (Kim et al., 2013). Fortunately, this phenomenon can be prevented by breaking the cell wall of microalgae to provoke the release of intracellular lipids into the extracting mixture, thus facilitating the access of solvent to lipids. Therefore, once released from the algal cell, lipids are able to pass to the solvent phase from which they can be collected after evaporation of the solvent. We use a simple and low energy consuming technique for cell disruption, based on the use of low toxicity and cheap reactants such as H₂O₂.

The wet biomass was resuspended in 1 ml of a 1/40 (v/v) solution of H₂O₂ for 4 minutes under agitation. Hence, neutral lipid extraction was performed directly on the wet disrupted biomass according to a method that represents a slight modification of the one proposed by Fajardo et al. (2007). The method consists firstly of diluting 1/10 the mixture of wet-disrupted biomass and disruption solution with ethanol (96% v/v) while assuring the contact for 18 hours under continuous stirring. As mentioned above, this step allowed also stopping the disruption reaction. The resulting hydro-alcoholic solution was then subjected to centrifugation at 4000 rpm for 10 minutes in order to separate solid residuals (i.e. pieces of broken cells, organelles, etc.) from the supernatant liquid where lipids were transferred. The lipid-rich supernatant was then suitably stored while the residual solid was further contacted with ethanol for 1 hour under stirring in order to extract residual lipids remained in the solid phase. The extracted oil was evaporated to dryness in a rotary evaporator flask.

At the end of each experiment growth in different conditions, was assessed the concentration of lipids in the algal biomass by extraction. The maximum value is obtained with growth in standard BBM (24.31%), the second one in nitrogen starvation (14.8%) and the last one in nitrogen supplement (12.75%). The cultivation of algae under nitrogen limitation conditions were reported to increase the accumulation of storage lipids such as triacylglycerols (TAG) while it had only limited effect on *C. sorokiniana*.

4.6. Fatty acid methyl esters analysis

As described before, one objective of this work was to select the best culture conditions with the aim of obtaining the maximum lipid productivity for biodiesel production. For this reason, after biomass recovery, lipids were extracted following the procedure described above, and fatty acids separated using GC–MS for their quantification.

The fatty acid methyl esters (FAMES) composition of extracted lipids was determined according to the European regulation/commission regulation EEC n° 2568 (1991) after transesterification with methanol-acetyl chloride is performed. To this aim gas chromatographic analysis was carried using a flame ionization detector (FID) 67 (Thermo Trace Ultra, GC-14B) and a RTX-WAX column T (fused silica, 0.25 mm x 60 m x 0.25 µm) maintained at 180 °C. Helium was used as carrier gas at a flow rate of 1 ml min⁻¹.

The effect of nitrate on fatty acid composition of *C. sorokiniana* were evaluated at early stationary growth phase when cells growth became plateau. The comparison among FAMES profiles of three different grow condition (standard BBM, modified with 5N o 1/5 N) is reported in Table 6 in terms of weight percentage of each fatty acid with respect to the total amount of FAMES identified.

Table 6. Fatty acid methyl esters profile of lipids extracted after using the different culture condition.

Carbon number	Fatty acid name	BBM	BBM modified with 5N
C14:0	Myristic	0.49	0.12
C16:0	Palmitic	23.45	8.66
C16:1	Palmitoleic	4.45	2.59
C17:0	Heptadecanoic	0.20	0.12
C17:1	Heptadecenoic		
C18:0	Stearic	3.22	1.18
C18:1	Oleic	8.21	4.70
C18:2	Linoleic	21.41	23.57
C18:3	Linolenic	16.78	37.11
C20:0	Arachidic	0	0.09
C22:0	Behenic	0.15	0
-	Other polyunsat.	0.1	0

The European Standard for Biodiesel (UNE-EN 14214, 2003) limits the content of C18:3 and PUFAs in a quality biodiesel at 12% and 1% respectively. It can be observed that FAMES obtained from biomass grow in modified BBM displayed an

high content of linolenic acid (C18:3) that reduces the quality of the biodiesel product.

Palmitic acid (C16:0) was the most abundance type of SFA that depicted similar accumulation trends as total SFAs in all samples. The content of C16:0 showed a sharp drop when nitrate was increased. Nitrogen limitation in the medium had been shown to increase the C18:1, while the C16:0 decreased.

On the contrary, a low content of total saturated (20 %wt/wt), monounsaturated (7.4 %wt/wt) and linoleic (8.4 % wt/wt) acids, which are the most useful fatty acids for producing biodiesel, was observed.

The results clearly showed that the ability for the accumulation of fatty acid composition in the different sample of *C. sorokiniana* was differentially regulated in respond to various nitrate concentrations at the early stationary growth phase.

CHAPTER 5.

Genetic characterization of oleaginous algae belonging to the Chlorellaceae family

5.1. Isolation, sequencing and annotation of DNA

The cells were harvested by centrifugation for 10 minutes at 4,000 g and 4 °C, and the cell pellet was ground to powder under liquid nitrogen using mortar and pestle. Total cellular DNA was isolated from fresh algal pellets with a DNeasy plant mini kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. The DNA concentration for each extraction was measured with a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). The quality of the genomic DNA was assessed by electrophoresis on a 1% 1X TAE agarose gel. Whole-genome sequencing of *Chlorella sorokiniana* SAG 211-8k, *Pseudochloris wilhelmii* SAG 55.87, *Monodus subterraneus* SAG 848.1, *Scenedesmus obliquus* SAG 276-1, *Chlorella variabilis* NC64A and *Tetraspora* sp. SCCA024 strains were performed using an Illumina HiSeq2000 platform. Libraries of 300 and 800 base pair (bp) fragments were constructed using the TrueSeq DNA Sample Prep Kit (Illumina, San Diego, CA, USA) and paired-end reads were generated on the Illumina HiSeq 2000 (100-bp reads) sequencing platform located at CRS4 in Pula (CA, Sardinia, Italy). Quantification and size distribution of the libraries were determined using the Agilent DNA 1000 kit (Agilent Technologies, Germany; cf. Figure 22). Each barcode sample was sequenced from both ends for 100 cycles and achieved 35 Gb of raw DNA sequence data each, which enabled an average of 1000-fold estimated coverage.

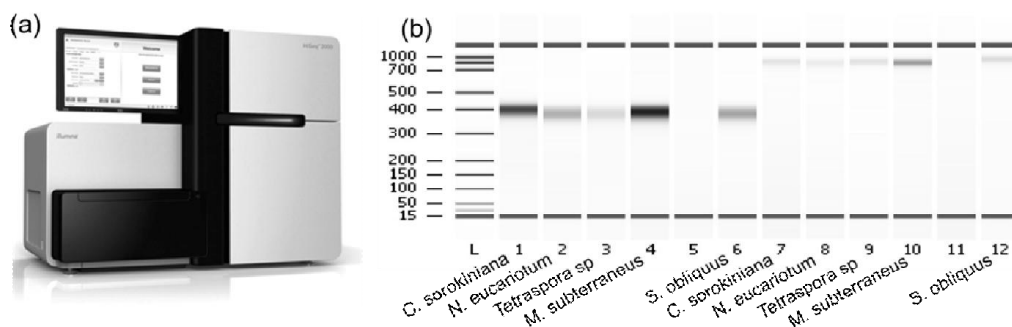


Figure 22. Illumina's HiSeq2000 sequencing system (a) and data quality over two libraries (300 – 800 bp) for each investigated strain.

Reads were assembled by using Velvet (Zerbino and Birney, 2008), scaffolded and finished by some ad-hoc developed python packages. Genome annotation was based on Dogma (Wyman et al., 2004), RNAmmer (Lagesen et al., 2007) and T-rna scan (Schattner et al., 2005) followed by manual inspection. Genes and ORFs were

compared by Blast homology searches against the non-redundant database of National Center for Biotechnology and Information (NCBI).

5.2. *Chlorella sorokiniana* organelles genomes

5.2.1. Chloroplast DNA

The chloroplast DNA (cpDNA) sequence of *C. sorokiniana* (GenBank accession number KJ397925) assembles as a circular map, obtained using the CGView software (Stothard and Wishart, 2005), of 109,811 bp which encodes a total of 111 genes (cf. Figure 23, Table 7).

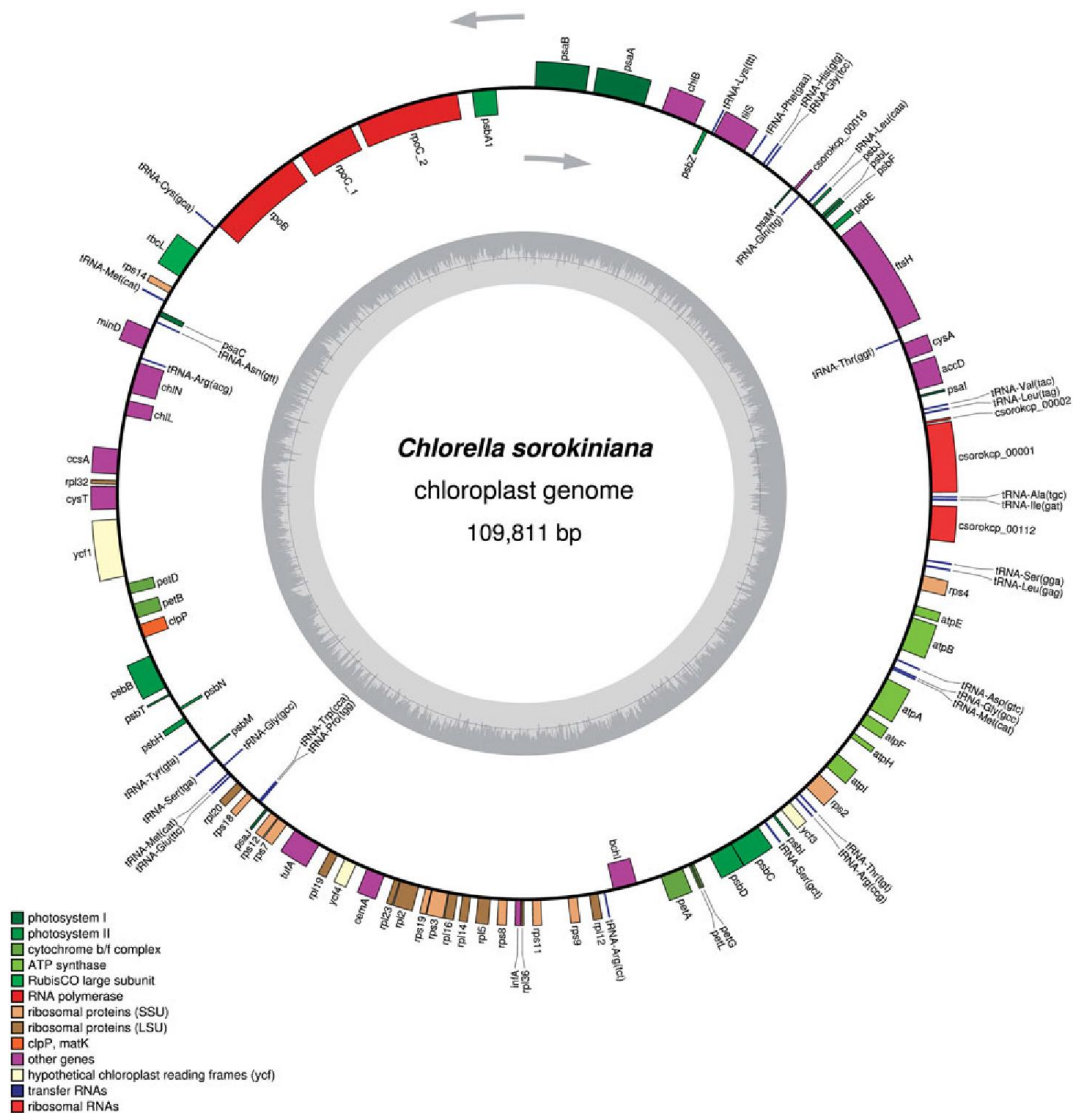


Figure 23. Gene map of the chloroplast genome of *C. sorokiniana* (Orsini et al., 2014a).

The chloroplast genes include 76 protein coding genes, 3 rRNAs and 31 tRNAs; the protein coding genes can be functionally categorized in the following classes: genes directly involved in photosynthesis, those involved in transcription, in translation and in division (cf. Table 7).

One of typical features of *C. sorokiniana* chloroplast genome is the lacking of rpoA subunit of the RNA polymerase (rpo) complex, while other Chlorellaceae belong the complete set of genes (rpoA, rpoB, rpoC1 and rpoC2).

The *C. sorokiniana* plastid also encodes for tRNA(Ile)-lysidine synthetase (tilS).

This enzyme is responsible for modifying the CAU anticodon of a unique tRNA that is cognate for isoleucine. Inside of algae group, however, tilS is generally encoded in the nuclear genome and targeted to the organelle (de Koning and Keeling, 2006). Conversely it lacks the full set of *ndh* genes likewise the large part of algal chloroplasts (Stoebe et al., 1998) and the conserved ORF ycf20.

Eleven small hydrophobic subunits proteins (psbH, psbI, psbJ, psbK, psbL, psbM, psbTc, psbX, psbY, psbZ and ycf12) are conserved among cyanobacteria and higher plants (Shi and Schroder, 2004; Iwai et al., 2010). We observed the absence of psbX and psbY.

A gene involved in starch and lipid metabolism, i.e. acetyl-CoA carboxylase beta subunit (accD gene), has been found in the chloroplast genome. This enzyme catalyzes the irreversible conversion of acetyl-CoA to malonyl-CoA during the fatty-acid synthesis (Cronan and Waldrop, 2002).

The four genes (chlB, chlI, chlL, chlN) involved in chlorophyll biosynthesis were found through similarity to the corresponding genes identified. In fact, *C. sorokiniana* shows the largest cytochrome f (petA gene product), cytochrome b6 (petB), and subunit IV (petD) and the smallest petG, petL subunits. The functions of the small subunits above are unknown.

The tRNA repertoire consists of 31 tRNAs, 24 of which occupying one strand while the remaining 7 genes occupying the other one.

The trnL (uaa) has not been identified while it is conserved in the other Chlorellaceae except in *P. wickerhamii*.

In addition, *C. sorokiniana* seems to have lost the trnI (cau) gene. In the genome, which doesn't show introns, all genes are present in single copy. Moreover, as well as for other *Chlorella* strains, the genome is lacking large inverted repeat (IR).

Table 7. Gene list for the *Chlorella sorokiniana* chloroplast chromosome

RNA Genes				
Ribosomal RNAs				
rrn23	rrn16	rrn5		
Transfer RNAs				
trnA(UGC)	trnC(GCA)	trnD(GUC)	trnE(UUC)	trnF(GAA)
trnG(GCC) X2	trnG(UCC)	trnH(GUG)	trnI(GAU)	trnK(UUU)
trnL(CAA)	trnL(GAG)	trnL(UAG)	trnM(CAU) X2	trnM(CAU)
trnN(GUU)	trnP(UGG)	trnQ(UUG)	trnR(ACG)	trnR(CCG)
trnR(UCU)	trnS(GCU)	trnS(GGA)	trnS(UGA)	trnT(GGU)
trnT(UGU)	trnV(UAC)	trnW(CCA)	trnY(GUA)	
tilS				
Photosynthesis				
Photosystem I				
psaA	psaB	psaC	psaI	psaJ
psaM				
Photosystem II				
psbA1	psbB	psbC	psbD1	psbE
psbF	psbH	psbI	psbJ	psbK
psbL	psbM	psbN	psbT	psbZ
Cytochrome				
petA	petB	petD	petG	petL
ATP synthase				
atpA	atpB	atpE	atpF	atpH
atpI				
Chlorophyll biosynthesis				
chlB	chlI	chlL	chlN	
Rubisco				
rbcL				
Ribosomal proteins				
Large subunits				
rpl2	rpl5	rpl12	rpl14	rpl16
rpl19	rpl20	rpl23	rpl32	rpl36
Small subunits				
rps2	rps3	rps4	rps7	rps8
rps9	rps11	rps12	rps14	rps18
rps19				
Trascription/translation				
RNA polymerase				
rpoB	rpoC_1	rpoC_2		
Translation factors				
infA	tufA			
Division				
minD	ftsH			
Miscellaneous proteins				
accD	clpP	cysA	cysT	CemA
ccsA				
Conserved ORFs				
ycf1	ycf3	ycf4	ycf12	ycf47

C. sorokiniana displays the standard architecture of the rRNA operon of Chlorophyta structure (16S rDNA – tRNA Ile – tRNA Ala – 23S rDNA). The overall AT content

of the cpDNA is 65.9% and the coding sequence is 59.1% (54,279 bp). The sequence analysis highlights that this plastid genome does not display a quadripartite structure.

5.2.2. Mitochondrial DNA

The mitochondrial DNA (mtDNA) sequence of *C. sorokiniana* (GenBank accession number KM241869) assembles as a circular map of 52528 bp which encodes a total of 32 protein coding genes, 3 rRNAs and 25 tRNAs (cf. Figure 24, Table 8), while accounting for 97.4 % of the total genome.

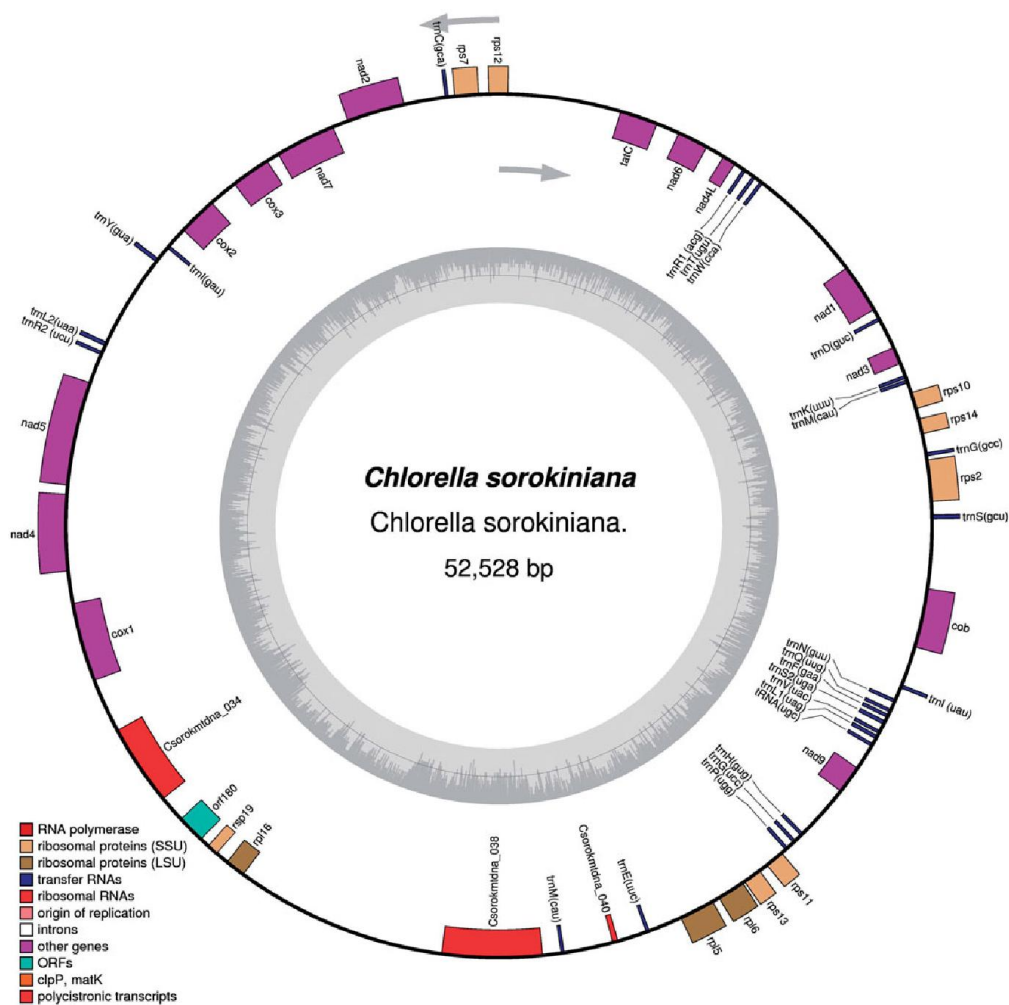


Figure 24. Gene map of the mtDNA of *C. sorokiniana* (Orsini et al., 2014b).

The AT content is 70.89%, all genes are present as single copies and do not contain introns. Seventeen conserved genes code for respiratory proteins of mitochondrial complexes I, III, IV, and V, but lack rps3 and rps19. The genome carries three protein coding genes, for subunits I and III of cytochrome c oxidase (cox1 and cox3, respectively) and cytochrome b (cob).

Table 8. Gene list for the *Chlorella sorokiniana* mitochondrial genome

RNA Genes				
Ribosomal RNAs				
rrn23	rrn16	rrn5		
Transfer RNAs				
trnA(ugc)	trnC(gca)	trnD(gcu)	trnE(uuc)	trnF(gaa)
trnG(gcc)	trnG(ucc)	trnH(gug)	trnI(gau)	trnK(uuu)
trnL(uag)	trnL(uaa)	trnM(cau)x2	trnM(cau)	trnN(guu)
trnP(ugg)	trnQ(uug)	trnR(acg)	trnR(ucu)	trnS(uga)
trnT(ugu)	trnV(uac)	trnW(cca)	trnY(gua)	
Genes for respiration and oxidative phosphorylation				
NADH dehydrogenase complex				
nad1	nad2	nad3	nad4	nad4L
nad5	nad6	nad7	nad9	
Complex III				
cob				
Cytochrome oxidase phosphorylation				
cox1	cox2	cox3		
ATP synthase complex				
atp1	atp4	atp6	atp8	atp9
Ribosomal protein genes				
Large subunits				
rpl5	rpl6	rpl16		
Small subunits				
rps2	rps3	rps4	rps7	rps10
rps11	rps12	rps13	rps14	rps19
Assembly, membrane insertion				
tatC				

In addition, *C. sorokiniana* presents the *tatC* gene which is found in Trebouxiophyceae but not in Chlorophyceae algae. In the genome, 37 genes occupy the forward strand while 22 genes occupying the other one. The gene distribution over the two DNA strands is slightly biased. By comparing of the mt genome of *C. sorokiniana* with the corresponding ones, so far sequenced, of other Trebouxiophyceae, a substantial rearrangements among strains, including large inversion, is observed.

These results indicate that important changes occurred at the levels of genome size, gene order, and intron content within the Trebouxiophyceae, while more similarity to the Chlorellaceae group are detected.

5.3. *Chlorella variabilis* NC64A organelles genomes

5.3.1. Chloroplast DNA

The cpDNA sequence of *C. variabilis* (GenBank accession number KP271969) assembles as a circular map of 124793 bp which encodes a total of 114 genes (cf. Figure 25, Table 9).

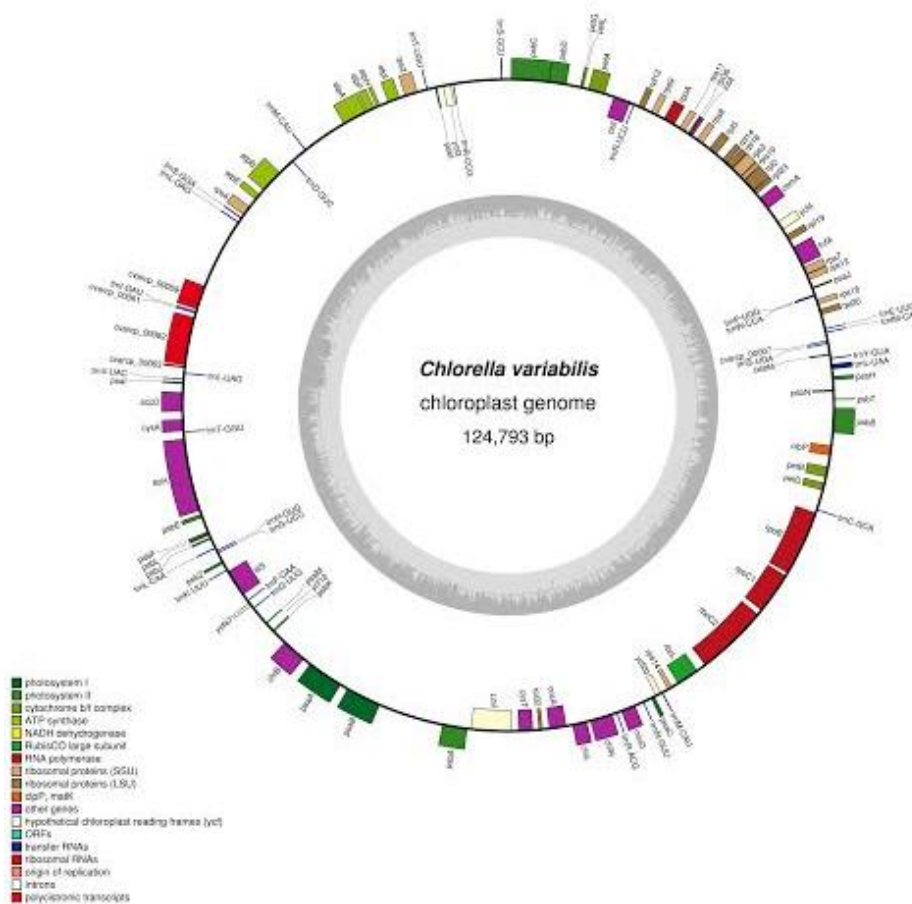


Figure 25. Gene map of the chloroplast genome of *C. variabilis* (Orsini et al., 2015).

This chloroplast gene repertoire includes 79 protein coding genes (50% of the entire genome with an average length of 790 bp), 3 rRNAs and 32 tRNAs. One tRNA coding gene (tRNA-Leu) and two protein coding genes (both belonging to the photosystem II) contained introns.

The genome shows strongly biased in gene distribution over the two DNA strands with 78 genes occupying the forward strand and 36 genes occupying the other. Like the others members of the *Chlorella* genus, the cp genome of *C. variabilis* not display a quadripartite structure. The GC content (33,9%) of the *C. variabilis* cp genome is close to those of *Chlorella* sp. ArM0029B (33,2%) (Jeong et al., 2014), *C. sorokiniana* (34,1%) (Orsini et al., 2014a) and *C. vulgaris* (31.6%) (Wakasugi et al., 1997).

Table 9. Gene list for the *Chlorella varibilis* chloroplast genome.

RNA Genes				
Ribosomal RNAs				
rrn23	rrn16	rrn5		
Transfer RNAs				
trnA(UGC)	trnC(GCA)	trnD(GUC)	trnE(UUC)	trnF(GAA)
trnG(GCC) X2	trnG(UCC)	trnH(GUG)	trnI(GAU)	trnK(UUU)
trnL(CAA)	trnL(GAG)	trnL(UAG)	trnM(CAU) X2	trnM(CAU)
trnN(GUU)	trnP(UGG)	trnQ(UUG)	trnR(ACG)	trnR(CCG)
trnR(UCU)	trnS(GCU)	trnS(GGA)	trnS(UGA)	trnT(GGU)
trnT(UGU)	trnV(UAC)	trnW(CCA)	trnY(GUA)	
tilS				
Photosynthesis				
Photosystem I				
psaA	psaB	psaC	psaI	psaJ
psaM				
Photosystem II				
psbA1	psbB	psbC	psbD1	psbE
psbF	psbH	psbI	psbJ	psbK
psbL	psbM	psbN	psbT	psbZ
Cytochrome				
petA	petB	petD	petG	petL
ATP synthase				
atpA	atpB	atpE	atpF	atpH
atpI				
Chlorophyll biosynthesis				
chlB	chlI	chlL	chlN	
Rubisco				
rbcL				
Ribosomal proteins				
Large subunits				
rpl2	rpl5	rpl12	rpl14	rpl16
rpl19	rpl20	rpl23	rpl32	rpl36
Small subunits				
rps2	rps3	rps4	rps7	rps8
rps9	rps11	rps12	rps14	rps18
rps19				
Transcription/translation				
RNA polymerase				
rpoA	rpoB	rpoC_1	rpoC_2	
Translation factors				
infA	tufA			
Division				
minD	ftsH			
Miscellaneous proteins				
accD	clpP	cysA	cysT	CemA
ccsA				
Conserved ORFs				
ycf1	ycf3	ycf4	ycf12	ycf20
ycf47				

The only difference with respect to the *C. sorokiniana* chloroplast genome, is the presence of the genes rpoA e ycf20. This demonstrates a close relation between the

two true *Chlorellae* species. The sequences of the symbiotic alga NC64A genomes presented here will help in the optimization of various processes, while further documenting the evolution of the green lineage.

5.3.2. Mitochondrial DNA

The mtDNA sequence of *C. variabilis* (GenBank accession number KP271968) assembles as a circular map of 78500 bp which encodes a total of 61 genes, all present as single copies (cf. Figure 26, Table 10).

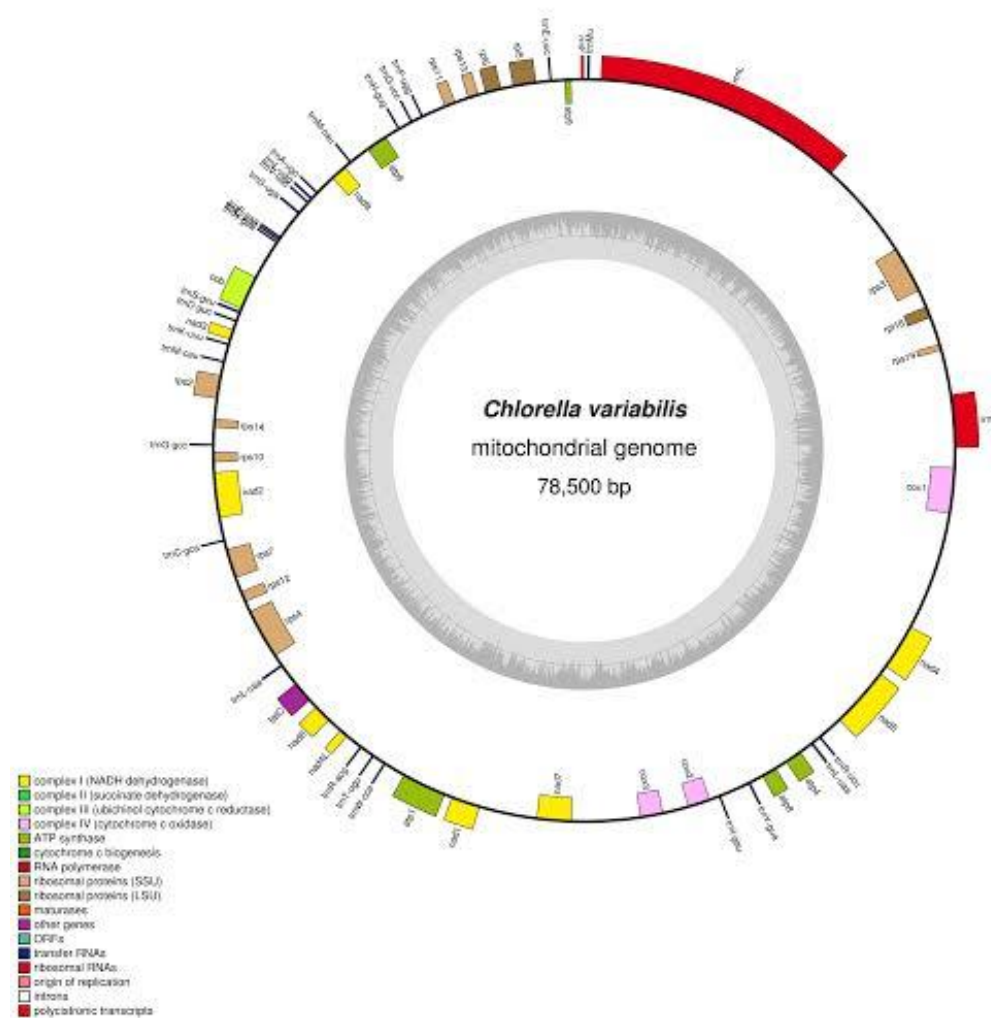


Figure 26. Gene map of the mitochondrial genome of *C. variabilis*

This genome encodes mainly for 27 tRNAs (*trn* genes), 13 ribosomal proteins (*rps* and *rpl*), and 3 subunits involved in cytochrome c biogenesis (*cox* genes). The overall GC content is 28.2% and the coding sequence is 34% (average length: 835 bp).

Like the *C. sorokiniana*'s one, the mt genome of *C. variabilis* exhibits the tatC gene. The gene distribution over the two DNA strains is strongly biased (46 genes occupy the forward strand while 16 genes occupying the other one). It is interesting to note that all tRNA is detected in the forward strand, while the cox gene in reverse strand.

Table 10. Gene list for the *Chlorella variabilis* mitochondrial genome.

RNA Genes				
Ribosomal RNAs				
rrn23	rrn16	rrn5		
Transfer RNAs				
trnA(ugc)	trnC(gca)	trnD(guc)	trnE(uuc)	trnF(gaa)
trnG(gcc)	trnG(ucc)	trnH(gug)	trnI(gau)	trnK(uuu)
trnL(caa)	trnL(uag)	trnL(uaa)	trnM(cau)x2	trnM(cau)
trnN(guu)	trnP(ugg)	trnQ(uug)	trnR(acg)	trnR(ucu)
trnS(gcu)	trnS(uga)	trnT(ugu)	trnV(uac)	trnW(cca)
trnY(gua)				
Genes for respiration and oxidative phosphorylation				
NADH dehydrogenase complex				
nad1	nad2	nad3	nad4	nad4L
nad5	nad6	nad7	nad9	
Complex III				
cob				
Cytochrome oxidase phosphorylation				
cox1	cox2	cox3		
ATP synthase complex				
atp1	atp4	atp6	atp8	atp9
Ribosomal protein genes				
Large subunits				
rpl5	rpl6	rpl16		
Small subunits				
rps2	rps3	rps4	rps7	rps10
rps11	rps12	rps13	rps14	rps19
Assembly, membrane insertion				
tatC				

Although the mitochondrial genome is characterized by a higher mutation rate, the only differences between *C. variabilis* and *C. sorokiniana* DNA consists of two tRNAs, i.e. trnL(caa) and trnS(ugu), that are not detected in the latter strain.

CHAPTER 6.
Phylogenetic analysis

6.1. Introduction

The chloroplast genome is particularly useful for phylogenetic reconstruction because of its relatively high and condensed gene content, in comparison to nuclear genomes (Leliaert et al., 2012) and thus its knowledge can be a useful to better understand the evolution of algae. To date, several complete chloroplast genomes of green algae have been sequenced and assembled (Leliaert et al., 2012; Hamaji et al., 2013; Lemieux et al., 2014) including 13 complete Chlorellaceae plastids (Wakasugi et al., 1997; de Koning and Keeling, 2006; Turmel et al., 2009; Jeong et al., 2014; Orsini et al., 2014a; Orsini et al., 2015). To perform a comprehensive comparative analysis of complete chloroplast genome evolution at different levels (strain-, species-, genus- and family) 11 members of this family, included two non-photosynthetic organisms, have been included in the study: *C. sorokiniana* SAG 211-8k; *C. sorokiniana* 1230; *C. vulgaris* C-27; *C. variabilis* NC64A; *Chlorella* sp. ArM0029B; *Parachlorella kessleri*, *Auxenochlorella protothecoides* 0710; *Helicosporidium* sp. ATCC 50920; *Prototheca wickerhamii* SAG 263-11; *Marvania geminata* SAG:12.88; *Pseudochloris wilhelmii* SAG:1.80 and *Dicloster acuatus* SAG:41.98 (cf. Table 11).

Table 11. General features of cp DNAs from *Chlorella sorokiniana* SAG 211-8k (*C. sor) - 1230 (*C. sor*) and the Chlorellaceae: *Auxenochlorella protothecoides* (*A. prot*), *Chlorella* sp. ArM0029B (*C. sp*), *C. variabilis* (*C. var*), *C. vulgaris* (*C. vul*), *Dicloster acuatus* (*D. acu*), *Helicosporidium* sp. (*H. sp*), *Marvania geminate* (*M. gem*), *Parachlorella kessleri* (*P. kess*), *Pseudochloris wilhelmii* (*P. wil*) and *Prototheca wickerhamii* (*P. wic*).**

Feature	<i>C. sor</i> 211-8k	<i>C. sor</i> 1230	<i>C. vul</i>	<i>C. var</i>	<i>C. sp.</i>	<i>P. kess</i>	<i>A. prot</i>	<i>P. wil</i>	<i>M. gem</i>	<i>D. acu</i>	<i>P. wic</i>	<i>H. sp.</i>
Accession	NC_023833	KJ742376	AB001684	KJ718922	KF554427	NC_012978	KC943975	KM462886	KM462888	NC_025546	KJ001761	NC_008100
Size (bp)												
Total	109811	109803	150613	124579	119989	123994	84580	109775	10847	169201	55636	37454
IR	-	-	-	-	-	10.913	-	-	-	-	-	-
LCS	-	-	-	-	-	13.871	-	-	-	-	-	-
SSC	-	-	-	-	-	88.297	-	-	-	-	-	-
A+T (%)	66.0	66.0	68.44	66.07	66.08	70.0		63.3	61.8	70.0	68.8	73.1
Coding (%)	59.1		60.9	53.4								
Genes (annotated)	11109	111	115*	114	114	119	110	126	114	128	70	54
rRNA	3	3	3	3	3	6	3	6	3	6	3	3
tRNA	31	31	33	32	32	36	30	36	32	35	27	25
Introns (n°)	0	0	3	0	1	1	0	1	1	10	0	1
Group I	0	0	3	0	0	1	0	1	1	10	0	1
Group II	0	0	0	0	0	0	0	0	0	0	0	0
Coding density	63.1	60.9	42.8	56.3	57.4	62.3		71.5	62.4	50.7	79.4	89.7
Average size (bp)	636.6	835.9	787.8	836.3	831.0	858.4		872.5	825.0	922.7	1027.9	1158.9

The loss of photosynthesis is typically associated with ptDNA reduction and the erosion of photosynthesis-related genes (de Koning and Keeling, 2006). The ptDNAs of *Helicosporidium* and *Prototheca* species have shorter intergenic regions, fewer introns, and reduced coding capacities as compared with photosynthetic Chlorellales (de Koning and Keeling, 2006).

6.2. Gene content and structure

The *C. sorokiniana* cpDNA shows the smallest plastid genome when compared to the other strains belonging to the *Chlorella* species. On the other hand, the AT content of 65.9% is close to one of the other *Chlorella* species while is far from the ones of *P. kessleri*, *A. protothecoides*, *D. acuatius* and *Helicosporidium* sp.. The *C. sorokiniana* cp genome with 87 (79.8%) genes occupying the forward strand and 22 genes occupying the other one, shows a gene distribution over the two DNA strands which is strongly biased, also when compared to the other Chlorellaceae. In fact, corresponding values equal to 68.4%, 65.2% have been reported for *C. variabilis* and *Chlorella* sp. respectively while the *C. vulgaris* strain shows a value of 49.2% (cf. Table 11). Among the genes involved in chlorophyll biosynthesis, similarly to other Chlorellaceae, *C. sorokiniana* cp genome lacks of psbX and psbY while, the four genes chlB, chlI, chlL, chlN are well conserved. Moreover, eleven small hydrophobic proteins (psbH, psbI, psbJ, psbK, psbL, psbM, psbTc, psbX, psbY, psbZ and ycf12) well conserved in cyanobacteria (Inoue-Kashino et al., 2011) and higher plants (Shi and Schroder, 2004; Iwai et al., 2010; Kashino et al. 2007), are present in all genomes of Chlorellaceae.

6.3. Gene order

The conserved gene order of cp genomes among *C. sorokiniana*, *Chlorella* sp and *C. variabilis* were compared in Figure 27. To avoid complexity, in this section we did not extend the analysis to all other organisms included in the study but we limited comparison to the *Chlorella* clade organisms only; the *C. sorokiniana* versus *C. vulgaris* comparison was considered separately.

The gene order between *C. sorokiniana* and *Chlorella* sp. is generally conserved; major rearrangements between them were found in three regions (cf. Figure 27). The cluster psbI-ycf3-trnR was observed in inverse orientation in the *C. sorokiniana* cp genome respect to both *Chlorella* sp. and *C. variabilis* strains. A similar rearrangement was observed in the trnQ-psaM-ycf12-psbK *C. sorokiniana* cluster, in particular it was remarked in different order respect to genome region comprising trnG-(present in two copies in the *C. sorokiniana* genome)-trnH-ycf47-trnF-trnC-psbZ gene cluster, which is itself in inverse orientation in the *Chlorella* sp. and *C.*

variabilis genomes. Moreover, in the latter strain this region shows an inversion (psbZ-trnK-tilS-trnF-ycf47 instead of ycf47-trnF-trnK-psbZ).

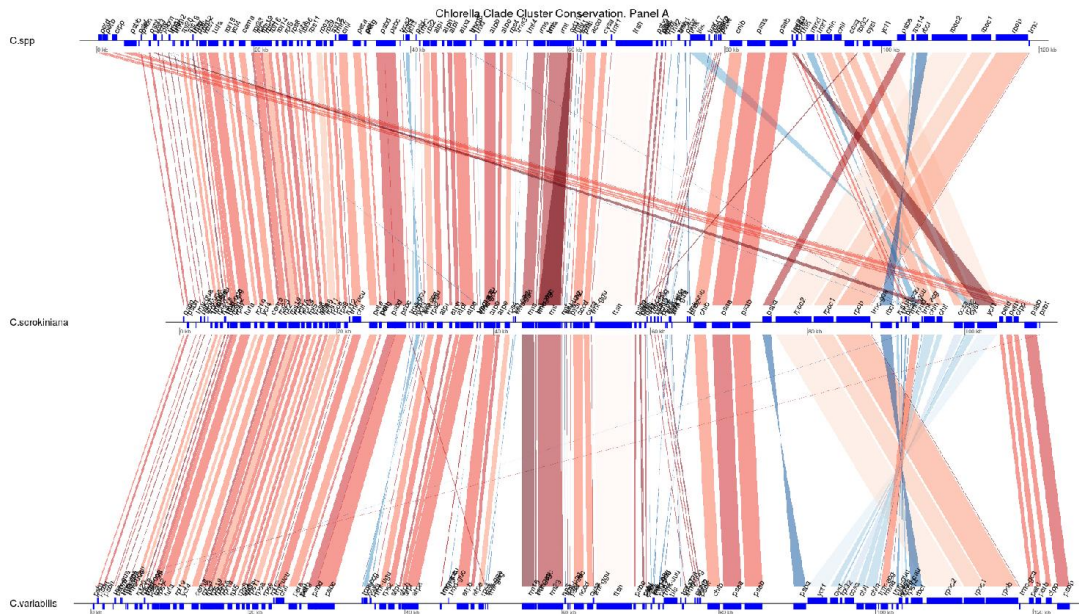


Figure 27. Comparison between the *C. sorokiniana*, *C. variabilis* and *Chlorella* sp..

A more complex rearrangement was observed in the *C. sorokiniana* *psaB-rbl32* region (around positions 72000-10000), the cluster *rpoC2-rpoC1-rpoB-trnC* which precede this region shares conserved orientation among the three considered strains. The *C. sorokiniana* *rbcL-rpl32* region can be considered as composed by several sub clusters: the *trnR(acg)-rpl32* cluster that is conserved in orientation between *C. sorokiniana* and *Chlorella* sp. but shows an inversion between *C. sorokiniana* and *C. variabilis*; the *minD* gene, which is present with the same orientation in *C. sorokiniana* and *C.variabilis* but in opposite direction between *C. sorokiniana* and *Chlorella* sp.; the *trnM-psaC-trnN* cluster shows the same orientation between *C. sorokiniana* and *Chlorella* sp. (even if in the latter strain the *ycf20* gene is inserted between the *trnN* and *psaC* genes) but is found in inverse orientation between the *C. sorokiniana* and *C. variabilis* genomes; the *rbcL-rps14* cluster is upside down among the *C. sorokiniana* and the other two organisms. This region can be then considered as a large cluster *rbcL-rps14-trnN-psaC-trnN* inverted between *C. sorokiniana* and *C. variabilis*: Moreover, the cluster itself, if compared between *C. sorokiniana* and *Chlorella* sp., can be divided in two subcluster: *rbcL-rps14* which is inverted in *Chlorella* sp. and similarly to the *C. variabilis* genome, and a subcluster *trnN-psaC-trnM* which is conserved between *C. sorokiniana* and *Chlorella* sp..

The gene order comparison between *C. sorokiniana* and *C. vulgaris* is shown in Figure 28. With respect to the two above mentioned comparisons, the *C. sorokiniana* versus *C. vulgaris* contrast presents a higher occurrence of rearrangements, each of them involving a larger number of genes. A small amount of clusters were found in same orientation, even if in different order, between the two strains: trnE-rpl20-rps18-trnW-trnP-psaJ-rps12-rps7-tnfA-rpl19-ycf4-psbI-ycf3-trnR, rns-trnI-trnA-rnL-rn5 (in the *C. vulgaris* the rnL gene is interrupted and contains an intron plus the I-cvuI gene), trnk-psbZ-chlB-psaA-psaB (this cluster in the *C. vulgaris* is interrupted by the insertion of multiple genes), rbcL-rps14, chlN-chlL-ycf5-rpl32-cysT, petD-petB-clpP.

Other clusters, on the contrary, are present in inverse orientation and different order between the two strains: psbN-psbH, rpl12-trnR-chlI-petA-petL-petG, trnT-rps2-atpI-atpH-atpF-atpA, atpB-atpE-rps4, psaI-accD-cysA, psbE-psbF-psbI-psbJ, trnG-trnH, trnN-minD-trnR (in the *C. vulgaris* the minE gene interrupts the cluster). Some other isolated genes, trnS in particular, were observed in inverse orientation between the two algae strains.

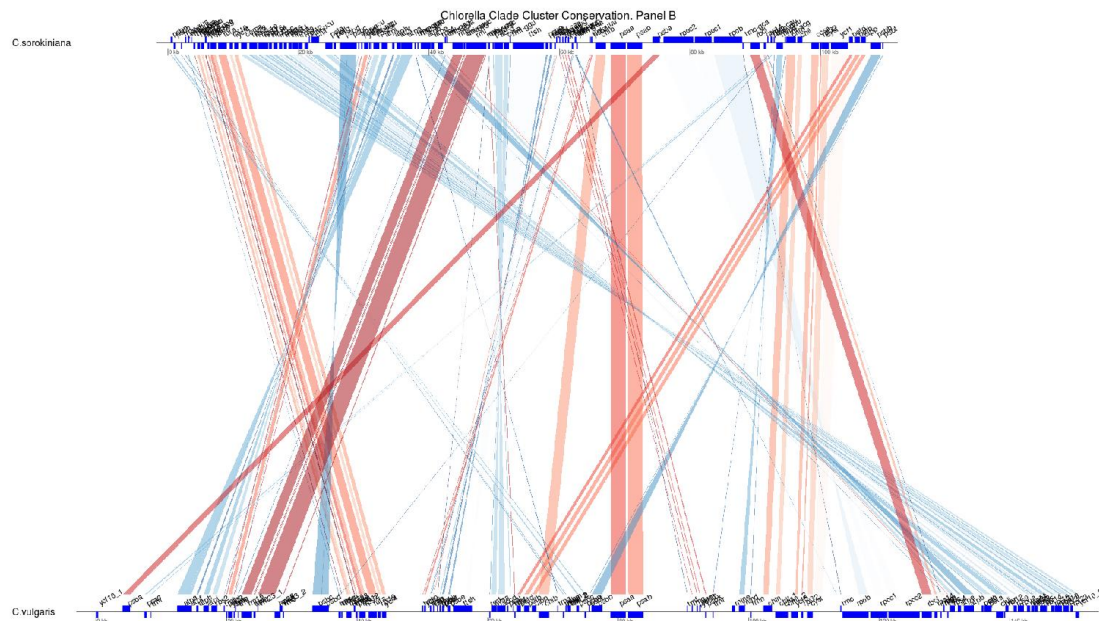


Figure 28. Comparison between the *C. sorokiniana* and *C. vulgaris*.

6.4. Ancestral gene clusters

Some ancestral gene clusters were preserved during the evolution of Chlorellales. These clusters provide important information about the evolution of the organisms and in addition help the understanding of the phylogenetic relationship between them. To identify the ancestral clusters carried by *C. sorokiniana* cpDNA as well as the

derived clusters that are shared with others Chlorellaceae cpDNAs, we investigated the 19 gene clusters (cf. Figure 29).

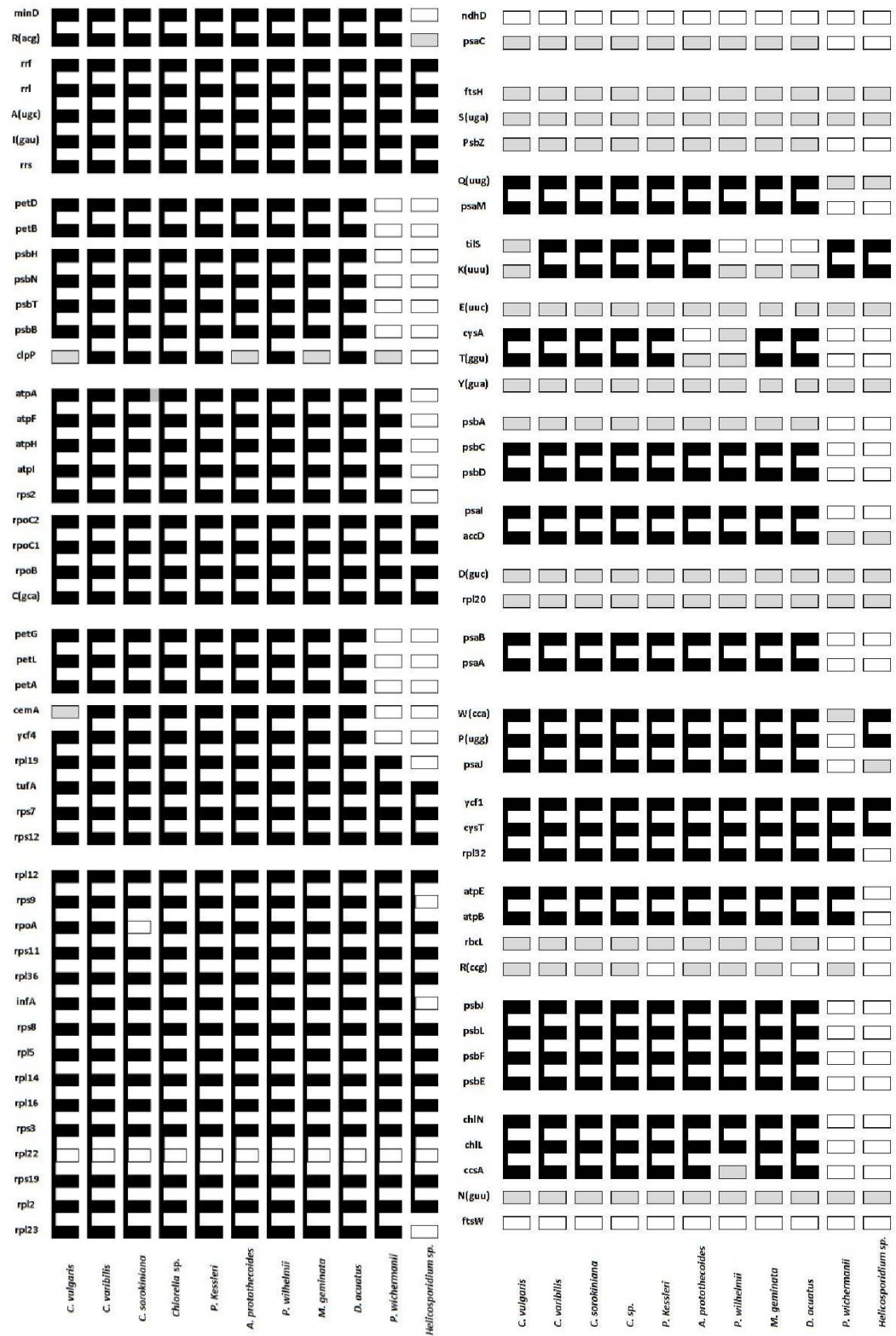


Figure 29. Conserved gene clusters in the Chlorellales algae.

Regarding the presence of derived gene clusters, the *C. sorokiniana* chloroplast genome clearly bears more similarity with *C. variabilis* and *Chlorella* sp. ArM0029B than with the other cpDNAs compared. All the photosynthetic strains possess six intact conserved ancestral gene clusters (psbE-psbF-psbL-psbJ; rpl32-cysT-ycf1; psaJ-trnP(ugg)-trnW(cca); psaA-psaB; accD-psaI; psaM-trnQ(uug) and eight partially conserved blocks diverging and conserving contemporary (petA-petL-petG; trnC(gca)-rpoB-rpoC1-rpoC2; rps2-atpI-atpH-atpF-atpA; petB-petD; rrs-trnI(gau)-trnA(ugc)-rrl-rrf; trnR(acg)-minD; atpB-atpE; psbD-psbC).

Moreover, *P. wickerhamii* shows only one of the six intact clusters (rpl32-cysT-ycf1) and three of the eight partially conserved blocks (trnC(gca)-rpoB-rpoC1-rpoC2; rps2-atpI-atpH-atpF-atpA; petB-petD; rrs-trnI(gau)-trnA(ugc)-rrl-rrf; trnR(acg)-minD; atpB-atpE). It is interesting to note that *P. wilhemlii* is the only photosynthetic Chlorellaceae that diverges from the others due to the absence of two clusters (ccsA-chlL-chlN; trnT(ggu)-cysA). Turmel et al. (2009) showed that 10 gene linkages unique to *Parachlorella* and *Chlorella* are consistent with the notions that these green algae belong to the same monophyletic group (Chlorellaceae). In our study we used more than one species for the *Chlorella* genus and we observed that the 10 clusters are preserved in all of them except for *C. sorokiniana* which doesn't show the gene rpoA. In addition we found two new cluster (rrs-trnI(gau)-trnA(ugc)-rrl-rrf/trnR(acg)-minD; accD-psaI) that are conserved between the species of the two genus.

6.5. Phylogenetic analysis

6.5.1. Parameters and data set

The chloroplast genomes of 14 green algae were considered for phylogenomic analysis. The GenBank accession numbers of these green algal genomes are the following: *C. sorokiniana* SAG 211-8k (NC_023835.1); *C. sorokiniana* 1230 (KJ742376.1); *C. vulgaris* C-27 (AB001684.1); *C. variabilis* NC64A (KP271969); *Chlorella* sp. ArM0029B (KF554427.1); *Parachlorella kessleri* (NC_012978.1), *Auxenochlorella protothecoides* (KC843975.1); *Helicosporidium* sp. ATCC 50920 (NC_008100.1), *Prototheca wickerhamii* SAG 263-11 (KJ001761.1), *Marvania*

geminata SAG:12.88 (KM462888.1), *Pseudochloris wilhelmii* SAG:1.80 (KM462886.1), *Dicloster acuatus* SAG:41.98 (NC_025546.1), *Coccomyxa subellipsoidea* C-169 (NC_015084) and *Pedinomonas minor* (NC_016733.1). We selected seventy-one protein-coding genes on the basis of the work by Lemieux et al., (2014): *accD*, *atpA*, B, E, F, H, I, *ccsA*, *cemA*, *chlB*, I, L, N, *clpP*, *ftsH*, *infA*, *petA*, B, D, G, L, *psaA*, B, C, I, J, M, *psbA*, B, C, D, E, F, H, I, J, K, L, M, N, T, Z, *rbcL*, *rpl2*, 5, 14, 16, 20, 23, 32, 36, *rpoA*, B, C1, C2, *rps2*, 3, 4, 7, 8, 9, 11, 12, 14, 18, 19, *tufA*, *ycf1*, 3, 4, 12. The above indicated aminoacids sequences were aligned by MAFFT, the ambiguously aligned regions were removed using TRIMAL (Capella-Gutierrez et al., 2009) with same parameters indicated in Lemieux et al. 2014 , Trimmed alignments were concatenated using Phyutility (Smith and Dunn, 2008). Phylogenesis was inferred by Maximum Likelihood method implemented in MEGA6 using cpREV + Gamma4 models of sequence evolution. Confidence of branch points was evaluated by fast-bootstrap analysis with 500 replicates.

Maximum likelihood tree from nuclear 18S gene (operon) sequences aligned by MAFFT with parameters optimized for RNA secondary structure, was inferred by MEGA package using GRT + Gamma models of sequence evolution. Similarly, 16S operons from chloroplast sequences were aligned by MAFFT as above described, and maximum likelihood tree was inferred using cpREV + Gamma models of sequence evolution. Confidence of branch points was evaluated for each tree by fast-bootstrap analyses with 500 replicates.

The complete cp genome of *C. sorokiniana*, was compared to the above mentioned organisms by using Mauve Alignment tool; chloroplast genomes were artificially rearranged when necessary for graphical reasons. Conserved cluster genes among *Chlorella* clade organisms were identified by BLASTN searches, while genoPlotR was used to prepare figures.

6.5.2. Phylogenetic reconstruction

The aim of this work was to improve the knowledge about the evolution of the *Chlorella*-like organisms within the Chlorellaceae family (Krienitz et al., 2004, 2010; Luo et al. 2006, 2010; Bock et al. 2011a, b; Skaloud et al. 2014) using a new data set of genes. The ML phylogenetic trees were inferred using respectively: nuclear sequences (18S rDNA) (cf. Figure 30), chloroplast rRNA operon data sets

(16S-tRNA Ile-Ala-23S) (cf. Figure 31) and 71 chloroplast genes (cf. Figure 32) from the most comprehensive taxa sampling to date of the strains compared in this study. Trees were rooted with the class Trebouxiophyceae and focused on the Chlorellaceae family with an emphasis on the *Chlorella* clade. The taxa set includes photosynthetic and colorless Chlorellales.



Figure 30. Phylogeny of 12 Chlorellaceae inferred using maximum likelihood (ML) analyses of nuclear 18S rRNA

The trees are consistent with previous analyses (Huss et al., 1999; Ueno et al., 2003; Krienitz et al., 2004; Pröschold and Leilaert, 2007; Lemieux et al., 2014; Figueroa-Martinez et al., 2014). Using the *Coccomyxa* and *Pedinomonas* genera as outgroup, the phylogenetic analyses demonstrated five distinct clades of Chlorellaceae: the *Chlorella*, *Parachlorella*, *Auxenochlorella*, *Marvania* and *Helicosporidium* clades. In particular, *C. sorokiniana*, *C. variabilis*, *C. vulgaris* and *Chlorella* sp. ArM0029B are depicted as members of a strongly supported clade.

In fact, the monophyly of the four species of *Chlorella* is well showed in the trees in Figure 32 and are characterized by their high sequence similarities with respectively 97% between *C. sorokiniana* and *Chlorella* sp. ArM0029B and 98% between *Chlorella* sp. and *C. variabilis* bootstrap values. Furthermore, relative to the distance between *Chlorella* sp. ArM0029B and *C. variabilis* and *Chlorella* sp. ArM0029B and *C. vulgaris* C-27 we have had analogous results to a recently study (Jeong et al., 2014). Indeed, the distance between the first and the second strain is more shorter

compared or the distance between the first and the third strain, this result proved the close relationships of *Chlorella* sp. ArM0029B to *C. variabilis*. The monophyletic *Chlorella* clade appears as a sister clade to *Auxenochlorella* and *Marvania* clade, and this relationship is strongly supported by bootstrap analyses. The *Dicloster* genus belonging to the *Parachlorella*-clade was highly supported in all bootstraps. In addition, in the 16S operon tree the *Parachlorella*, *Dicloster* subclade and *Chlorella* *Auxenochlorella*/*Marvania* groups are separated in two branches with 100% bootstrap supports. *Chlorella* clade is then separated by *Auxenochlorella*/*Marvania* clade with a very low bootstrap support (39%).



Figure 31. Phylogeny of 12 Chlorellaceae inferred using chloroplast rRNA operon data sets (16S-tRNA Ile-Ala-23S)

Our maximum likelihood (ML) analyses of nuclear 18S rRNA (cf. Figure 30), chloroplast rRNA operon data sets (16S-tRNA Ile-Ala-23S) (cf. Figure 31) and 71 chloroplast genes (cf. Figure 32) sequences from the taxon included in this study depict *P. wickerhamii* SAG 263-11 as a sister clade to *M. geminata* with low bootstrap support while to *A. protothecoides* there is a strongly supported, but the taxon sampling remains relatively poor to confirm this relationship. Interestingly, the relationships among the genera, and even the monophyly of some of them, differed significantly across these reconstructions. In view of that, 16S rRNA sequences from various trebouxiophytes depict *P. wickerhamii* SAG 263-11 as a nonsister lineage to the other *Prototheca* and *Helicosporidium* species. These data suggest that the loss of photosynthesis has occurred at least twice in the evolution of parasitic/pathogenic

Chlorellales (Figueroa-Martinez et al., 2014). The loss of photosynthesis has occurred several independent times among chlorophytes: at least twice in the order Chlorellales (Trebouxiophyceae) (Figueroa-Martinez et al., 2014).

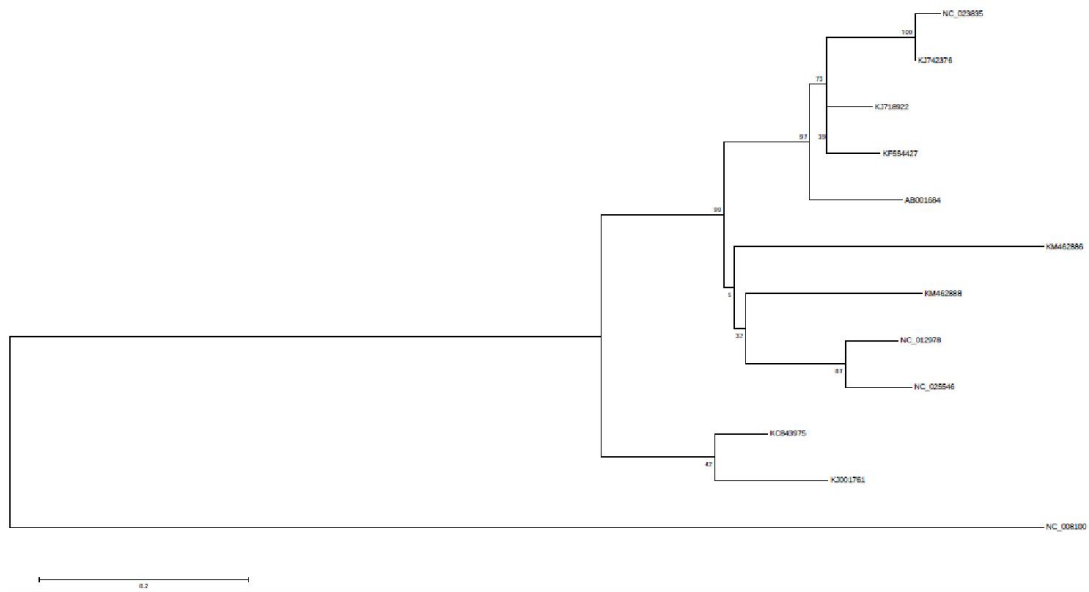


Figure 32. Phylogeny of Chlorellaceae inferred using 71 chloroplast genes.

6.6. Concluding remarks

Although significant efforts have been given to characterize growth phenotypes and the fatty acid content within the genus *Chlorella*, knowledge of the genetic and genomic basis that defines and controls their physiological behavior are still lacking; critical information is required to support effective genetic engineering strategies. Recently, an analysis of the mitochondrial and chloroplast genomes of two strains revealed that the genomic content was highly conserved between these species yet, evolutionarily divergent ‘hotspots’ were present, enabling an accurate phylotyping of these closely related species.

CHAPTER 7.

Conclusions and suggestions for further research

In the light of for sustainable development, microalgal biodiesel, as a renewable and sustainable energy type, has enjoyed a surge in popularity (Chisti, 2007; Chen et al., 2012; Feng et al., 2012; Ito et al., 2012). In fact, differently from the first generation biofuels, the use of microalgae to produce bioenergy does not involve the triggering of "food for fuel" competitiveness and thus represents a sustainable mean to face significant concerns, such as wars and political instabilities deriving from oil reserves shortage. Moreover, the high oil yields and less land use are the main advantages of microalgae. However, in order to make the current technology viable at the large-scale, several limitations should be overcome. In particular, biomass and lipid productivities should be further increased and all the downstream processes, from harvesting to lipid extraction, should be optimized. To these aims, high efforts involving high investments should be done in order to implement an intensive multidisciplinary research activity both at the laboratory and the industrial scale.

The microalgae cultivation is the base of biofuel development and suitable genetic engineering strategies have to be developed in order to augment the microalgae oil content and their growth rate so that biofuels production could be performed in a sustainable way. In particular, the creation of new microalgal strains intrinsically characterized by high lipid productivities as well as by a good tolerance to high CO₂ levels is an ambitious goal which might be achieved only once their genome is known. The results presented in this thesis represent the first step needed to design a genetic approach which may eventually facilitate large-scale production of algae. The use of transgenic microalgae for the production of bioproducts represents an enormous economic and biotechnology promise, because algal production combines the simplicity and speed of haploid, single-cell genetics in an organism with elaborate biosynthetic potential, and with the associated economic benefit of using photosynthesis to drive product formation.

As technology continues to be progressed and algae production industrialization continues to be improved, microalgae energy as the third generation biofuel will contribute their own strength to relieve the tense situation of resources.

The contribution of the present work, to this general target can be briefly summarized as follows. The growth kinetics of *C. sorokiniana* has been investigated along with their corresponding lipid content, both batch and helical photobioreactors. The main results achieved during this activity are the knowledge of the effect of nitrogen concentration in solution on the growth rate and lipid content of *C.*

sorokiniana. These informations represent the first step towards the development of a nitrogen based strategy for the optimization of lipid productivity of *C. sorokiniana* cultures.

As far as the genetic characterization activity, the chloroplast and mitochondrial DNAs of two strains, i.e. *C. sorokiniana* and *C. variabilis*, respectively, have been sequenced for the first time in the literature. The obtained results allowed to perform a phylogenetic assessment involving different microalgae strains belonging to the *Chlorella* clade. Such results represent the first important step towards the development of genetic engineering strategies aimed to improve the current microalgae based systems for the production of biofuels and the capture of CO₂.

References

- Abu-Rezq, T.S., Al-Musallam, L., Al-Shimmari, J., Dias, P. (1999) Optimum production conditions for different high-quality marine algae. *Hydrobiologia*. 403, 97-107.
- Ahmad, A., Yasin, N., Derek, C., Lim, J. (2011) Microalgae as a sustainable energy source for biodiesel production: a review. *Renewable and Sustainable Energy Reviews*. 15 (1), 584-593.
- Armbrust, E.V., Berges, J.A., Bowler, C., Green, B.R., Martinez, D., Putnam, N.H., Zhou, S., Allen, A.E., Apt, K.E. and other authors (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science*. 306, 79–86.
- Aylott, M. (2011). European EnAlgae partnership to unlock the potential of algal bioenergy.NNFCC.<http://www.nnfcc.co.uk/news/energetic-algae-project-to-investigate-potential-of-algae-energyand-fuels>.
- Behrenfeld, M.J. and Falkowski, P.G. (1997) Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*. 42(1), 1997. 1-20.
- Beyerinck, M.W. (1890) Culturversuche mit Zoochlorellen, Lichenengonidien und anderen niederen Algen. *Botanische Zeitung*. 47: 725-739, 741-754, 757-768, 781-785.
- Bhattacharya, D., Price, D.C., Chan, C.X., Qiu, H., Rose, N., Ball, S., Weber, A.P. M., Arias, M.C., Henrissat, B. and other authors (2013) Genome of the red alga *Porphyridium purpureum*. *Nature Communication*. 4, 1941.
- Bhattacharya, D., Yoon, H.S., Hackett, J.D. (2004) Photosynthetic eukaryotes unite: endosymbiosis connects the dots. *Bioessays*. 26, 50–60.
- Bischoff, H.W. and Bold, H.C. (1963) Phycological studies. IV. Some soil algae from Enchanted Rock and related algal species. University of Texas Publications. 6318: 1-95.

Blanc, G., Agarkova, I., Grimwood, J., Kuo, A., Brueggeman, A., Dunigan, D.D., Gurnon, J., Ladunga, I., Lindquist, E. and other authors (2012) The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biology*. 13, R39.

Blanc, G., Duncan, G., Agarkova, I., Borodovsky, M., Gurnon, J., Kuo, A., Lindquist, E., Lucas, S., Pangilinan, J. and other authors (2010) The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell*. 22, 2943–2955.

Bock, C., Pröschold, T., Krienitz, L. (2011) Updating the genus *Dictyosphaerium* and description of *Mucidosphaerium* gen. nov. (Trebouxiophyceae) based on morphological and molecular data *Journal of Phycology*. 47 (3), pp. 638-652. doi: 10.1111/j.1529-8817.2011.00989.

Bock, C., Krienitz, L., Pröschold, T. (2011) Taxonomic reassessment of the genus *Chlorella* (Trebouxiophyceae) using molecular signatures (barcodes), including description of seven new species *Fottea* 11(2): 293–312.

Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F. and other authors (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature*. 456, 239–244.

Brandt, K. (1881) Ueber das Zusammenleben von Thieren und Algen. *Archiv für Anatomie, Physiologie*. 1881: 570–4.

Brandt, K. (1882) Über die morphologische und physiologische Bedeutung des Chlorophylls bei Thieren. *Archiv für Anatomie, Physiologie*. 1882: 125–51.

Brennan, L. and Owende, P. (2010) Biofuels from microalgae: a review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable and Sustainable Energy Reviews*. 14 (2), 557-577.

Brouard, J.S., Otis, C., Lemieux, C., Turmel, M. (2010) The exceptionally large chloroplast genome of the green alga *Floydiella terrestris* illuminates the evolutionary history of the Chlorophyceae. *Genome Biology and Evolution*. 2, 240–256.

Brown, W.M., George, M.Jr., Wilson, A.C. (1979) Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA*. 76, 1967–1971.

Brunnthaler, J. (1915). Chlorophyceae 2. In: Pascher A, editors. *Sübwasserflora Deutschlands, Österreichs und der Schweiz*. Jena: Fischer. p 183–6.

Bullerwell, C.E. and Gray, M.W. (2004) Evolution of the mitochondrial genome: protist connections to animals, fungi and plants. *Current Opinion in Microbiology*. 7: 528–534.

Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 25(15):1972–1973.

Carvalho, P.A., Silva, O.S., Baptista, M.Jo., Malcata, F.X. (2011) Light Requirements in microalgal photobioreactors: An overview of biophotonic aspects. *Applied Microbiology and Biotechnology*. 89 p.1275–1288.

Chader, S., Hacene, H., Agathos, S.N. (2009) Study of hydrogen production by three strains of *Chlorella* isolated from the soil in the Algerian Sahara. *International Journal of Hydrogen Energy*. 34, 4941–4946.

Chelf, P., Brown, L.M., Wyman, C.E. (1993) Aquatic biomass resources and carbon dioxide trapping. *Biomass Bioenergy*. 4:175–183.

Chen, R., Li, R., Deitz, L., Liu, Y., Stevenson, R.J., Liao, W. (2012) Freshwater algal cultivation with animal waste for nutrient removal and biomass production. *Biomass and Bioenergy*. 39, 128–138.

Chen, C. and Durbin, E.G. (1994) Effects of pH on the growth and carbon uptake of marine phytoplankton. *Marine Ecology-Progress Series*. 109, 83-83.

Cheng, Y., Zhou, W., Gao, C., Lan, K., Gao, Y., Wu, Q. (2009) Biodiesel production from Jerusalem artichoke (*Helianthus Tuberosus* L.) tuber by heterotrophic microalgae *Chlorella protothecoides*. *Journal of Chemical Technology and Biotechnology*. 84 (5), 777-781.

Chisti, Y. (2008) Biodiesel from microalgae beats bioethanol. *Trends in biotechnology*. 26 (3), 126-131.

- Chisti, Y. (2007) Biodiesel from microalgae. *Biotechnology Advances*. 25 (3), 294-306.
- Cock, J.M., Sterck, L., Rouzé, P., Scornet, D., Allen, A.E., Amoutzias, G., Anthouard, V., Artiguenave, F., Aury, J.-M. and other authors (2010). The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature*. 465, 617–621.
- Collén, J., Porcel, B., Carré, W., Ball, S.G., Chaparro, C., Tonon, T., Barbeyron, T., Michel, G., Noel, B. and other authors (2013). Genome structure and metabolic features in the red seaweed *Chondrus crispus* shed light on evolution of the Archaeplastida. *Proc Natl Acad Sci USA*. 110, 5247–5252.
- Concas, A., Pisu, M., Cao, G. (2014) Engineering aspects related to the use of microalgae for biofuel production and CO₂ capture from flue gases. *Current Environmental Issues and Challenges*. pp 73-111.
- Concas, A., Lutz, G.A., Locci, A.M., Cao, G. (2013) *Nannochloris eucaryotum* growth in batch photobioreactors: kinetic analysis and use of 100% (v/v) CO₂. *Advances in Environmental Research*. 2 (1), 19-33.
- Concas, A., Lutz, G.A., Pisu, M., Cao, G. (2012) Experimental analysis and novel modeling of semi-batch photobioreactors operated with *Chlorella vulgaris* and fed with 100%(v/v) CO₂. *Chemical Engineering Journal*. 213, 203-213.
- Concas, A. and Cao, G. (2011) Process for bio-oil production involving the use of CO₂. Patent 2371940.
- Concas, A., Pisu, M., Cao, G. (2010) Novel simulation model of the solar collector of BIOCOIL photobioreactors for CO₂ sequestration with microalgae. *Chemical Engineering Journal*. 157 (2), 297-303.
- Conklin, K.Y., Kurihara, A., Sherwood, A.R. (2009) A molecular method for identification of the morphologically plastic invasive algal genera *Eucheuma* and *Kappaphycus* (Rhodophyta, Gigartinales) in Hawaii. *Journal of Applied Phycology*. 21, 691–699.
- Cornet, J.-F., Dussap, C.G., Gros, J.-B., Binois, C., Lasseur, C. (1995) A simplified monodimensional approach for modeling coupling between radiant light transfer and

growth kinetics in photobioreactors. *Chemical Engineering Science*. 50 (9), 1489-1500.

Croce, R. and Van Amerongen, H. (2013) Light-harvesting in photosystem I. *Photosynthesis Research*. 116:153–166.

Cronan, Jr J.E. and Waldrop, G.L. (2002) Multi-subunit acetyl-CoA carboxylases. *Progress in Lipid Research*. 41:407–35.

Curtis, B.A., Tanifuji, G., Burki, F., Gruber, A., Irimia, M., Maruyama, S., Arias, M. C., Ball, S.G., Gile, G.H. and other authors (2012) Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature*. 492, 59–65.

Dayananda, C., Sarada, R., Kumar, V., Ravishankar, G.A. (2007) Isolation and characterization of hydrocarbon producing green alga *Botryococcus braunii* from Indian freshwater bodies. *Electronic Journal of Biotechnology*. 10 (1), 78-91.

de-Bashan, L.E., Trejo, A., Huss, V.A.R., Hernandez, J.-P., Bashan, Y. (2008) *Chlorella sorokiniana* UTEX 2805, a heat and intense sunlight-tolerant microalga with potential for removing ammonium from wastewater. *Bioresource Technology*. 99 (11), 4980–4989.

de Cambiaire, J.-C., Otis, C., Lemieux, C., Turmel, M. (2006) The complete chloroplast genome sequence of the chlorophycean green alga *Scenedesmus obliquus* reveals a compact gene organization and a biased distribution of genes on the two DNA strands. *BMC Evolutionary Biology*. 6:37.

de Koning, A.P. and Keeling, P.J. (2006) The complete plastid genome sequence of the parasitic green alga *Helicosporidium* is highly reduced and structured. *BMC Biology*. 4, 12.

Derelle, E., Ferraz, C., Rombauts, S., Rouzé, P., Worden, A.Z., Robbens, S., Partensky, F., Degroeve, S., Echeynié, S. and other authors (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc Natl Acad Sci USA*. 103, 11647–11652.

Devi, M.P., Subhash, G.V., Mohan, S.V. (2012) Heterotrophic cultivation of mixed microalgae for lipid accumulation and wastewater treatment during sequential

growth and starvation phases: effect of nutrient supplementation. *Renewable Energy*. 43, 276-283.

Durmaz, Y., Monteiro, M., Bandarra, N., Gökpinar, S., Işık, O. (2007) The effect of low temperature on fatty acid composition and tocopherols of the red microalga, *Porphyridium cruentum*. *Journal of Applied Phycology*. 19 (3), 223-227.

Fajardo, A.R., Cerdan, L.E., Medina, A.R., Fernandez, F.G.A., Moreno, P.A.G., Grima, E.M. (2007) Lipid extraction from the microalga *Phaeodactylum tricorutum*. *European Journal of Lipid Science and Technology*. 109 (2), 120-126.

Feng, P., Deng, Z., Fan, L., Hu, Z. (2012) Lipid accumulation and growth characteristics of *Chlorella zofingiensis* under different nitrate and phosphate concentrations. *Journal of Bioscience and Bioengineering*. 114, 405–410.

Feng, Y., Li, C., Zhang, D. (2011) Lipid production of *Chlorella vulgaris* cultured in artificial wastewater medium. *Bioresource technology*. 102 (1), 101-105.

Figueroa-Martinez, F., Nedelcu, M.A, Smith, D.R., Reyes-Prieto A. (2014) When the lights go out: the evolutionary fate of free-living colorless green algae. *New Phytologist*. 206: 972–982 doi: 10.1111/nph.13279.

Francisco, E.C., Neves, D.B., Jacob-Lopes, E., Franco, T.T. (2010) Microalgae as feedstock for biodiesel production: carbon dioxide sequestration, lipid production and biofuel quality. *Journal of Chemical Technology and Biotechnology*. 85 (3), 395-403.

Friedl, T. and Rybalka, N. (2012) Systematics of the green algae: a brief introduction to the current status. *Progress in Botany*. 73: 259–280.

Gao, C., Wang, Y., Shen, Y., Yan, D., He, X., Dai, J., Wu, Q. (2014) Oil accumulation mechanisms of the oleaginous microalga *Chlorella protothecoides* revealed through its genome, transcriptomes, and proteomes. *BMC Genomics*, 15:582.

Gapasin, R.S.J., Bombeo, R., Lavens, P., Sorgeloos, P., Nelis, H. (1998) Enrichment of live food with essential fatty acids and vitamin C: effects on milkfish (*Chanos chanos*) larval performance. *Aquaculture*. 162, 269–286.

Georgianna, D.R. and Mayfield, S.P. (2012) Exploiting diversity and synthetic biology for the production of algal biofuels. *Nature*. 488 (7411), 329-335.

Gobler, C. J., Berry, D.L., Dyhrman, S.T., Wilhelm, S.W., Salamov, A., Lobanov, A. V., Zhang, Y., Collier, J.L., Wurch, L.L. and other authors (2011) Niche of harmful alga *Aureococcus anophagefferens* revealed through ecogenomics. *Proc Natl Acad Sci USA*. 108, 4352–4357.

Gouveia, L., Marques, A.E., da Silva, T.L., Reis, A. (2009) *Neochloris oleabundans* UTEX 1185: a suitable renewable lipid source for biofuel production. *Journal of industrial microbiology and biotechnology*. 36 (6), 821-826.

Gray, M.W., Lang, B.F., Burger, G. (2004) Mitochondria of protists. *Annual Review of Genetics*. 38: 477–524.

Gray, M.W., Cedergren, R., Abel, Y., Sankoff, D. (1989) On the evolutionary origin of the plant mitochondrion and its genome. *Proc Natl Acad Sci USA* 86:2267–71.

Griffiths, M. and Harrison, S.L. (2009) Lipid productivity as a key characteristic for choosing algal species for biodiesel production. *Journal of Applied Phycology*. 21 (5), 493–507.

Gross, M. (2013) Development and optimization of novel algal cultivation systems. Iowa State University. Food Science and Human Nutrition Department.

Hallick, R.B., Hong, L., Drager, R.G., Favreau, M.R., Monfort, A., Orsat, B., Spielmann, A., Stutz, E. (1993) Complete sequence of *Euglena gracilis* chloroplast DNA. *Nucleic Acids Research*. 21, 3537–3544.

Halim, R., Danquah, M. K., Webley, P.A. (2012) Extraction of oil from microalgae for biodiesel production: a review. *Biotechnology advances*. 30 (3), 709-732.

Halim, R., Gladman, B., Danquah, M.K., Webley, P.A. (2011) Oil extraction from microalgae for biodiesel production. *Bioresource technology*. 102 (1), 178-185.

Hamaji, T., Smith, D.R., Noguchi, H., Toyoda, A., Suzuki, M., et al. (2013) Mitochondrial and plastid genomes of the colonial green alga *Gonium pectorale* give insights into the origins of organelle DNA architecture within the volvocales. *PLoS ONE*. 8(2): e57177. doi:10.1371/journal.pone.0057177.

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

Harrison, N. and Kidner, C.A. (2011) Next-generation sequencing and systematics: What can a billion base pairs of DNA sequence data do for you? *Taxon*. 60, 1552–1566.

Hosikian, A., Lim, S., Alim, R. (2010) Chlorophyll extraction from microalgae: A review on the process engineering aspects. *International Journal of Chemical Engineering*. vol. 2010, Article ID 391632, 11 pages <http://dx.doi.org/10.1155/2010/391632>.

Hsieh, C.-H. and Wu, W.-T. (2009) Cultivation of microalgae for oil production with a cultivation strategy of urea limitation. *Bioresource technology*. 100 (17), 3921-3926.

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., Darzins, A. (2008) Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal*. 54, pp. 621–639.

Humphrey, A.M. (2004) Chlorophyll as a color and functional ingredient. *Journal of Food Science*. vol. 69, no. 5, pp. 422–425.

Huss, V.A.R., Frank, C., Hartmann, E.C., Hirmer, M., Kloboucek, A., Seidel, B.M., Wenzeler, P., Kessler, E. (1999) Biochemical taxonomy and molecular phylogeny of the genus *Chlorella* sensu lato (Chlorophyta). *Journal of Phycology*. 35: 587-598.

Inoue-Kashino, N., Kashino, Y., Takahashi, Y. (2011) Psb30 is a photosystem II reaction center subunit and is required for optimal growth in high light in *Chlamydomonas reinhardtii*. *Journal of Photochemistry and Photobiology B: Biology*. Vol. 104, Issues 1–2, July–August, Pages 220–228.

Ito, T., Tanaka, M., Shinkawa, H., Nakada, T., Ano, Y., Kurano, N., Soga, T., Tomita, M. (2012) Metabolic and morphological changes of an oil accumulating trebouxiophycean alga in nitrogen-deficient conditions. *Metabolomics*. DOI 10.1007/s11306-012-0463-z.

Iwai, M., Suzuki, T., Kamiyama, A., Sakurai, I., Dohmae, N., Inoue, Y., Ikeuchi, M. (2010) The PsbK subunit is required for the stable assembly and stability of other

small subunits in the PSII complex in the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. *Plant and Cell Physiology*. Apr; 51(4) :554-60.

Jalalizadeh, M. (2012) Development of an integrated process model for algae growth in a photobioreactor, Ph.D. dissertation, University of South Florida.

Jeong, H., Lim, J.M., Park, J., Sim, Y.M., Choi, H.G., Lee, J., Jeong, W.J. (2014) Plastid and mitochondrion genomic sequences from Arctic *Chlorella* sp. ArM0029B. *BMC Genomics*. 15(1):286.

Karakashian, S.J. and Karakashian, M.W. (1965) Evolution and symbiosis in the genus *Chlorella* and related algae. *Evolution*. 19:368–377.

Kashino, Y., Takahashi, T., Inoue-Kashino, N., Ban, A., Ikeda, Y., Satoh, K., Sugiura, M. (2007) Ycf12 is a core subunit in the photosystem II complex. *Biochimica et Biophysica Acta*. 1767 1269–1275

Keeling, P.J., Burger, G., Durnford, D.G., Lang, B.F., Lee R.W., et al. (2005) The tree of eukaryotes. *Trends in Ecology & Evolution*. 20 (12): 670–676.

Keune, N. Algae: Fuel of the Future? *National Review Online of the U.S.* 08.03.2012.<http://www.nationalreview.com/articles/292913/algae-fuel-future-nash-keune>.

Kim, M.K., Park, J.H, Bhattacharya, D., Yoon, H.S. (2014) Applications of next-generation sequencing to unravelling the evolutionary history of algae. *International Journal of Systematic and Evolutionary Microbiology*. 64, 333–345.

Kim, J., Yoo, G., Lee, H., Lim, J., Kim, K., Kim, C.W., Park, M.S., Yang, J.-W. (2013) Methods of downstream processing for the production of biodiesel from microalgae. *Biotechnology advances*.
<http://dx.doi.org/10.1016/j.biotechadv.2013.04.006>.

Kithching, I.J. (1998) *Cladistics* 2nd Edition. Oxford University Press, Oxford, UK.

Klok, A.J., Martens, D.E., Wijffels, R.H., Lamers, P.P. (2013) Simultaneous growth and neutral lipid accumulation in microalgae. *Bioresource Technology*. 134, 233–243.

- Kodama, Y. and Fujishima, M. (2009) Infection of *Paramecium bursaria* by Symbiotic *Chlorella* Species in Endosymbionts in *Paramecium*, Microbiology Monographs 12, 31 DOI: 10.1007/978-3-540-92677-1_2.
- Kong, Q.-X., Li, L., Martinez, B., Chen, P., Ruan, R. (2010) Culture of microalgae *Chlamydomonas reinhardtii* in wastewater for biomass feedstock production. Applied biochemistry and Biotechnology. 160 (1), 9-18.
- Krienitz, L., Huss, V.A.R., Bock, C. (2015) Chlorella: 125 years of the green survivalist. Home: Trends in Plant Science. doi: 10.1016/j.tplants.2014.11.005.
- Krienitz, L., Bock, C., Luo, W., Pröschold, T. (2010) Polyphyletic origin of the dictyosphaerium morphotype within Chlorellaceae (Trebouxiophyceae). Journal of Phycology. Vol. 46, Issue 3, pages 559–563.
- Krienitz, L., Hegewald, E., Hepperle, D., Huss, V., Rohr, T., Wolf, M. (2004) Phylogenetic relationship of Chlorella and Parachlorella gen. nov. (Chlorophyta, Trebouxiophyceae). Phycologia. 43:529–542.
- Lagesen, K., Hallin, P.F., Rødland, E., Stærfeldt, H.H., Rognes, T., Ussery, DW. (2007) RNAMmer: consistent annotation of rRNA genes in genomic sequences Nucleic Acids Research. 35:3100-8.
- Lee, Y.-K., Ding, S.-Y., Hoe, C.-H., Low, C.-S. (1996) Mixotrophic growth of *Chlorella sorokiniana* in outdoor enclosed photobioreactor. Journal of Applied Phycology. 8: 163–169.
- Leliaert, F., Smith, D.R., Moreau, H., Herron, M.D., Verbruggen, H., Delwiche, C.F., De Clerck, O. (2012) Phylogeny and molecular evolution of the green algae. CRC Critical Reviews in Plant Sciences. 31:1–46.
- Lemieux, C., Otis, C., Turmel, M. (2014) Six newly sequenced chloroplast genomes from prasinophyte green algae provide insights into the relationships among prasinophyte lineages and the diversity of streamlined genome architecture in picoplanktonic species. BMC Genomics.15:857.
- Lewis, L.A. and McCourt, R.M. (2004) Green algae and the origin of land plants. American Journal of Botany. 91: 1535–1556.

- Li, Y., Han, D., Hu, G., Dauvillee, D., Sommerfeld, M., Ball, S., Hu, Q. (2010a) *Chlamydomonas* starchless mutant defective in ADP-glucose pyrophosphorylase hyper-accumulates triacylglycerol. *Metabolic engineering*. 12(4), 387-391.
- Li, Y., Han, D., Hu, G., Sommerfeld, M., Hu, Q. (2010b) Inhibition of starch synthesis results in overproduction of lipids in *Chlamydomonas reinhardtii*. *Biotechnology and Bioengineering*, 107(2), 258-268.
- Li, N. (2009) Algae biofuel development suggestions, *Science Times*, 09.02.2009 29 <http://wenku.baidu.com/view/42c97ed049649b6648d74748.html> Accessed on 09.04.2013
- Li, Y., Horsman, M., Wang, B., Wu, N., Lan, C.Q. (2008a) Effects of nitrogen sources on cell growth and lipid accumulation of green alga *Neochloris oleoabundans*. *Applied microbiology and biotechnology*. 81 (4), 629-636.
- Li, G., Xiao, P., Webley, P., Zhang, J., Singh, R., Marshall, M. (2008b) Capture of CO₂ from high humidity flue gas by vacuum swing adsorption with zeolite 13X. *Adsorption*. 14 (2-3), 415-422.
- Liang, Y., Sarkany, N., Cui, Y. (2009) Biomass and lipid productivities of *Chlorella vulgaris* under autotrophic, heterotrophic and mixotrophic growth conditions. *Biotechnology letters*. 31 (7), 1043-1049.
- Lin, Q., Lin, J. (2011) Effect of nitrogen source and concentration on biomass and oil production of a *Scenedesmus rubescens* like microalga. *Bioresource Technology*. 102, pp. 1615–1621.
- Liu, J., Huang, J., Chen, F. (2007) Microalgae as feedstocks for biodiesel production. *Biodiesel-Feedstocks and Processing*, InTech. 133-160.
- Lizzul, A.M., Hellier, P., Purton, S., Baganz, F., Ladommatos, N., Campos, L. (2014) Combined remediation and lipid production using *Chlorella sorokiniana* grown on wastewater and exhaust gases. *Bioresource Technology*. 151:11–18.
- Lommer, M., Specht, M., Roy, A.S., Kraemer, L., Andreson, R., Gutowska, M.A., Wolf, J., Bergner, S.V., Schilhabel, M.B. and other authors (2012) Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biology*. 13, R66.

Luo, W., Pröschold, T., Bock, C., Krienitz, L. (2010). Generic concept in *Chlorella*-related coccoid green algae (Chlorophyta, Trebouxiophyceae). *Plant Biology*. May 1;12(3):545-53.

Luo, W., Pflugmacher, S., Pröschold, T., Walz, N., Krienitz, L. (2006) Genotype versus phenotype variability in *Chlorella* and *Micractinium* (Chlorophyta, Trebouxiophyceae). *Protist* 157: 315–33.

Mahasneh, I.A. (1997) Production of single cell protein from five strains of the microalga *Chlorella* spp. (Chlorophyta). *Cytobios*. 90, 153–161.

Mancera, N., Douma, L.G., James, S., Liu, S., Van, A., Boucias, D.G., Tartar, A. (2012) Detection of *Helicosporidium* spp. in metagenomic DNA. *Journal of Invertebrate Pathology*. 111: 13–19.

Mandalam, R.K. and Palsson, B. (1998) Elemental balancing of biomass and medium composition enhances growth capacity in high-density *Chlorella vulgaris* cultures. *Biotechnology and Bioengineering*. 59 (5), 605-611.

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

Matsukawa, R., Hotta, M., Masuda, Y., Chihara, M., Karube, I. (2000) Antioxidants from carbon dioxide fixing *Chlorella sorokiniana*. *Journal of Applied Phycology*. 12, 263–267.

Matsuzaki, M., Misumi, O., Shin-I, T., Maruyama, S., Takahara, M., Miyagishima, S.Y., Mori, T., Nishida, K., Yagisawa, F. and other authors (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature*. 428, 653–657.

Mayo, A.W. (1997) Effects of temperature and pH on the kinetic growth of unialga *Chlorella vulgaris* cultures containing bacteria. *Water Environment Research*. 69 (1), 64-72.

McFadden, G.I. and Melkonian, M. (1986) Culture Medium Waris-H. *Phycologia*. 25:551–7.

- Melis, A. (2009) Solar energy conversion efficiencies in photosynthesis: Minimizing the chlorophyll antennae to maximize efficiency. *Plant Science*. 177:272–280.
- Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K. and other authors (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*. 318, 245–250.
- Mereschkowski, C. (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol Centralbl.* 25, 593–604.
- Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K. and other authors (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*. 318, 245–250.
- Meuser, J.E., Boyd, E.S., Ananyev, G., Karns, D., Radakovits, R., Narayana, Murthy, U.M., Ghirardi M.L., Dismukes, G.C., Peters, J.W., Posewitz M.C. (2011) Evolutionary significance of an algal gene encoding an [FeFe]-hydrogenase with F-domain homology and hydrogenase activity in *Chlorella variabilis* NC64A. *Planta*. 234:829–843 DOI 10.1007/s00425-011-1431-y.
- Mishra, S.K., Suh, I.W., Farooq, W., Moon, M., Shrivastav, A., Park, M.S., Yang J.-W. (2013) Rapid quantification of microalgal lipids in aqueous medium by a simple colorimetric method. *Bioresource Technology*. 155, 330–333.
- Munoz, R. and Guieysse, B. (2006) Algal-bacterial processes for the treatment of hazardous contaminants: a review. *Water research*. 40 (15), 2799-2815.
- Naik, S., Goud, V.V., Rout, P.K., Dalai, A.K. (2010) Production of first and second generation biofuels: a comprehensive review. *Renewable and Sustainable Energy Reviews*. 14 (2), 578-597.
- Neale, P.J. and Melis, A. (1986) Algal photosynthetic membrane complexes and the photosynthesis-irradiance curve: a comparison of light-adaptation response in *Chlamydomonas reinhardtii* (Chlorophyta). *Journal of Phycology*, 22:531–538.
- Nei, M. and Kumar, S. (2000) *Molecular evolution and phylogenetics*. Oxford University Press, New York. 333 pp.

- Nelson, N. and Ben Shem, A. (2004) The complex architecture of oxygenic photosynthesis. *Nature*. 5:1–12.
- Ogbonna, J.C., Yoshizawa, H., Tanaka, H. (2000) Treatment of high strength organic wastewater by a mixed culture of photosynthetic microorganisms. *Journal of Applied Phycology*. 12 (3–5), 277–284.
- Orpez, R., Martinez, M. E., Hodaifa, G., El Yousfi, F., Jbari, N., Sanchez, S. (2009) Growth of the microalga *Botryococcus braunii* in secondarily treated sewage. *Desalination*. 246 (1), 625-630.
- Orsini, M., Costelli, C., Malavasi, V., Cusano, R., Concas, A., Angius, A., Cao, G. (2015) Complete sequence and characterization of mitochondrial and chloroplast genome of *Chlorella variabilis* NC64A. *Mitochondrial DNA*. doi:10.3109/19401736.2015.1007297.
- Orsini, M., Cusano, R., Costelli, C., Malavasi, V., Concas, A., Angius, A., Cao, G. (2014a) Complete genome sequence of chloroplast DNA (cpDNA) of *Chlorella sorokiniana*. *Mitochondrial DNA*. doi: 10.3109/19401736.2014.919466.
- Orsini, M., Cusano, R., Costelli, C., Malavasi, V., Concas, A., Angius, A., Cao, G. (2014b) Complete genome sequence of chloroplast DNA (cpDNA) of *Chlorella sorokiniana*. *Mitochondrial DNA*. doi: 10.3109/19401736.2014.919466.
- Palenik, B., Grimwood, J., Aerts, A., Rouzé, P., Salamov, A., Putnam, N., Dupont, C., Jorgensen, R., Derelle, E. and other authors (2007). The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc Natl Acad Sci USA*. 104, 7705–7710.
- Pombert, J.-F., Blouin, N.A., Lane, C., Boucias, D., Keeling, P.J. (2014) A lack of parasitic reduction in the obligate parasitic green alga *Helicosporidium*. *PLoS Genetics*. 10:e1004355.
- Pombert, J.F., Beauchamp, P., Otis, C., Lemieux, C., Turmel, M. (2006) The complete mitochondrial DNA sequence of the green alga *Oltmannsiellopsis viridis*: evolutionary trends of the mitochondrial genome in the Ulvophyceae. *Current Genetics*. 50: 137–147.

- Pombert, J.F., Otis, C., Lemieux, C., Turmel, M. (2004) The complete mitochondrial DNA sequence of the green alga *Pseudendoclonium akinetum* (Ulvophyceae) highlights distinctive evolutionary trends in the Chlorophyta and suggests a sister-group relationship between the Ulvophyceae and Chlorophyceae. *Molecular Biology and Evolution*. 21 (5): 922–935.
- Popescu, C.E. and Lee, R.W. (2007) Mitochondrial genome sequence evolution in *Chlamydomonas*. *Genetics Society of America*. DOI: 10.1534/genetics.106.063156.
- Powles, S.B. (1984) Photoinhibition of photosynthesis induced by visible light. *Annual Review of Plant Physiology and Plant Molecular Biology*. 35:14–44.
- Price, D.C., Chan, C.X., Yoon, H.S., Yang, E.C., Qiu, H., Weber, A.P.M., Schwacke, R., Gross, J., Blouin, N.A. and other authors (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. *Science*. 335, 843–847.
- Prochnik, S.E., Umen, J., Nedelcu, A.M., Hallmann, A., Miller, S.M., Nishii, I., Ferris, P., Kuo, A., Mitros, T. and other authors (2010) Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. *Science*. 329, 223–226.
- Pröschold, T., Darienko, T., Silva, P.C., Reisser, W., Krienitz, L. (2011) The systematic of “*Zoochlorella*” revisited employing an integrative approach. *Environmental Microbiology*. 13:350–364.
- Pröschold, T. and Leliaert, F. (2007) Systematics of the green algae: conflict of classic and modern approaches. In: *Unravelling the algae: The past, present, and future of algal systematics*, pp. 123–153. Brodie, J. and Lewis, J., Eds., CRC Press, Boca Raton, FL.
- Pröschold, T., Harris, E.H., Coleman, A.W. (2005) Portrait of a species: *Chlamydomonas reinhardtii*. *Genetics*. 170, 1601–1610.
- Qiao, H. and Wang, G. (2009) Effect of carbon source on growth and lipid accumulation in *Chlorella sorokiniana* GXNN01. *Chinese Journal of Oceanology and Limnology*. 27 (4), 762–768.

- Radakovits, R., Jinkerson, R.E., Darzins, A., Posewitz, M.C. (2010) Genetic engineering of algae for enhanced biofuel production. *Eukaryotic Cell*. 9, pp. 486–501.
- Rajamani, S., Siripornadulsil, S., Falcao, V., Torres, M., Colepicolo, P., Sayre, R. (2007) Phycoremediation of heavy metals using transgenic microalgae. *Advances in Experimental Medicine and Biology*. 616:99–109.
- Read, B.A., Kegel, J., Klute, M.J., Kuo, A., Lefebvre, S.C., Maumus, F., Mayer, C., Miller, J., Monier, A. and other authors (2013) Pangenome of the phytoplankton *Emiliana underpins* its global distribution. *Nature*. 499, 209–213.
- Rodolfi, L., Chini Zittelli, G., Bassi, N., Padovani, G., Biondi, N., Bonini, G., Tredici, M.R. (2009) Microalgae for oil: Strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnology and bioengineering*. 102 (1), 100-112.
- Rosenberg, J.N., Kobayashi, N., Barnes, A., Noel, E.A., Betenbaugh, M.J., Oyler G.A. (2014) Comparative analyses of three chlorella species in response to light and sugar reveal distinctive lipid accumulation patterns in the Microalga *C. sorokiniana*. *PLoS One*. 9:e92460 doi:10.1371/journal.pone.0092460.
- Ryo, H., Mitsunori, I., Nobutaka, I. (2010) *Chlorella variabilis* and *Micractinium reisseri* sp. nov. (Chlorellaceae, Trebouxiophyceae): Redescription of the endosymbiotic green algae of *Paramecium bursaria* (Peniculia, Oligohymenophorea) in the 120th year. *Phycological Research*. 58:188–201.
- Sagan, L. (1967) On the origin of mitosing cells. *Journal of Theoretical Biology*. 14, 225–274.
- Sakthivel, R. (2011) Microalgae lipid research, past, present: a critical review for biodiesel production, in the future. *Journal of Experimental Sciences*. 2 (10), 29-49.
- Schattner, P., Brooks, A.N., Lowe, T.M. (2005) The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Research*. Jul 1;33 doi: 10.1093/nar/gki366.

Scott, S.A., Davey, M.P., Dennis, J.S., Horst, I., Howe, C.J., Lea-Smith, D.J., Smith, A.G. (2010) Biodiesel from algae: challenges and prospects. *Current Opinion in Biotechnology*. 21 (3), 277-286.

Schönknecht, G., Chen, W.H., Ternes, C.M., Barbier, G.G., Shrestha, R.P., Stanke, M., Bräutigam, A., Baker, B.J., Banfield, J.F. and other authors (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science*. 339, 1207–1210.

Schenk, P.M., Thomas-Hall, S.R., Stephens, E., Marx, U.C., Mussgnug, J.H., Posten, C., Kruse, O., Hankamer, B. (2008) Second generation biofuels: High-efficiency microalgae for biodiesel production. *BioEnergy Research*. 1:20–43.

Servín-Garcidueñas, L.E. and Martínez-Romero, E. (2012) Complete mitochondrial and plastid genomes of the green microalga *Trebouxiophyceae* sp. strain MX-AZ01 isolated from a highly acidic geothermal lake. *Eukaryotic Cell* 11:1417–18.

Shi, L.X. and Schroder, W.P. (2004) The low molecular mass subunits of the photosynthetic supracomplex, photosystem II . *Biochimica et Biophysica Acta*. 1608: 75 – 96.

Siaut, M., Cuine, S., Cagnon, C., Fessler, B., Nguyen, M., Carrier, P., Beyly, A., Beisson, F., Triantaphylides, C., Li-Beisson, Y., Peltier, G. (2011) Oil accumulation in the model green alga *Chlamydomonas reinhardtii*: characterization, variability between common laboratory strains and relationship with starch reserves. *BMC Biotechnology*. 11, 7.

Skaloud, P., Skaloudová, M., Procházková, A. & Nemcová, Y. (2014). Morphological delineation and distribution patterns of four newly described species within the *Synura petersenii* species complex (Chrysophyceae, Stramenopiles). *European Journal of Phycology* 49(2): 213-229.

Smith, D.R., Burki, F., Yamada, T., Grimwood, J., Grigoriev, I.V., van Etten, J.L., Keeling, P.J. (2011) The GC-rich mitochondrial and plastid genomes of the green alga *Coccomyxa* give insight into the evolution of organelle DNA nucleotide landscape. *PLoS One* 6:e23624.

- Smith, S.A., Dunn, C.W. (2008) Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics*, 24(5):715–716.
- Stirk, W.A., Ördög, V., Van Staden J., Jäger, K. (2002) Cytokinin and auxin-like activity in Cyanophyta and microalgae. *Journal of Applied Phycology*. 14: 215–221.
- Stoebe, B., Martin, W., Kowallik, K.V. (1998) Distribution and nomenclature of protein-coding genes in 12 sequenced chloroplast genomes. *Plant Molecular Biology Reporter*. 16: 243–255.
- Stothard, P. and Wishart, D.S. (2005) Circular genome visualization and exploration using CGView. *Bioinformatics* 21:537–9.
- Su, C.-H., Giridhar, R., Chen, C.-W., Wu, W.-T. (2007) A novel approach for medium formulation for growth of a microalga using motile intensity. *Bioresource technology*. 98 (16), 3012-3016.
- Sydney, E., Da Silva, T., Tokarski, A., Novak, A., De Carvalho, J., Woiciechowski, A., Larroche, C., Soccol, C. (2011) Screening of microalgae with potential for biodiesel production and nutrient removal from treated domestic sewage. *Applied Energy*. 88 (10), 3291-3294.
- Tartar, A. (2013) The non-photosynthetic algae *Helicosporidium* spp.: emergence of a novel group of insect pathogens. *Insects*. 4: 375–391.
- Tirichine, L. and Bowler, C. (2011) Decoding algal genomes: tracing back the history of photosynthetic life on Earth. *The Plant Journal*. 66, 45–57.
- Timilsina, G.R., Du, X., Carriquiry, M.A. (2010) Second-Generation Biofuels. Policy Research Working Paper. 5406, Doi: 10.1596/1813-9450-5406.
- Turmel, M., Otis, C., Lemieux, C. (2009) The chloroplast genomes of the green algae *Pedinomonas minor*, *Parachlorella kessleri*, and *Oocystis solitaria* reveal a shared ancestry between the Pedinomonadales and Chlorellales. *Molecular Biology and Evolution*. 26, 2317–2331.
- Turmel, M., Lemieux, C., Burger, G., Lang, B.F., Otis, C., Plante, I., Gray, M.W. (1999) The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*. Two radically different evolutionary patterns within green algae. *Plant Cell* 11:1717–30.

Ueno, R., Urano, N., Suzuki, M. (2003) Phylogeny of the non-photosynthetic green micro-algal genus *Prototheca* (Trebouxiophyceae, Chlorophyta) and related taxa inferred from SSU and LSU ribosomal DNA partial sequence data. FEMS Microbiology Letters. 223(2):275–280.

Van Amerongen, H. and Croce, R. (2013) Light-harvesting in photosystem II. Photosynthesis Research. 116:251–263.

Wakasugi, T., Nagai, T., Kapoor, M., Sugita, M., Ito, M., Ito, S., Tsudzuki, J., Nakashima, K., Tsudzuki, T., Suzuki, Y., Hamada, A., Ohta, T., Inamura, A., Yoshinaga, K., Sugiura, M. (1997) Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division. PNAS. 94, 5967–5972.

Wan, M.X., Liu, P., Xia, J., Rosenberg, J.N., Oyler, G.A., Betenbaugh, M.J., Nie, Z., Qiu, G. (2011) The effect of mixotrophy on microalgal growth, lipid content, and expression levels of three pathway genes in *Chlorella sorokiniana*. Applied Microbiology and Biotechnology. 91:835–44.

Wang, Y. (2013) Microalgae as the Third generation biofuel : production, usage, challenges and prospects. Ph.D. dissertation, University of Uppsala.

Wang, Z.T., Ullrich, N., Joo, S., Waffenschmidt, S., Goodenough, U. (2009) Algal lipid bodies: stress induction, purification, and biochemical characterization in wild-type and starchless *Chlamydomonas reinhardtii*. Eukaryotic Cell. 8(12), 1856-186.

Wei, Z. (2010) China Academy of Science and China Petroleum Chemical Cooperation cooperate to develop microalgae biofuel technology. Chemical Technology Market. Vol.33-No.9 pp. 60.

Weyer, K.M., Bush, D.R., Darzins, A., Willson, B.D. (2010) Theoretical maximum algal oil production. BioEnergy Research. 3:204–213.

Wilson, A.C., Cann, R.L., Carr, S.M., Matthew, G., Gyllensten, U.B., Helm-Bychowski, K.M., Higuchi, R.G., Palumbi, S.T., Prager, E.M. and other authors (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. Biological Journal of the Linnean Society. 26, 375–400.

- Wolff, G., Plante, I., Lang, B.F., Kück, U., Burger, G. (1994) Complete sequence of the mitochondrial DNA of the chlorophyte alga *Prototheca wickerhamii*. Gene content and genome organization. *J Mol Biol* 237: 75–86.
- Worden, A.Z., Lee, J.H., Mock, T., Rouzé, P., Simmons, M.P., Aerts, A.L., Allen, A.E., Cuvelier, M.L., Derelle, E. and other authors (2009) Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*. 324, 268–272.
- Work, V.H., Radakovits, R., Jinkerson, R.E., Meuser, J.E., Elliott, L.G., Vinyard, D.J., Laurens, L.M., Dismukes, G.C., Posewitz, M.C. (2010) Increased lipid accumulation in the 142 *Chlamydomonas reinhardtii* sta7-10 starchless isoamylase mutant and increased carbohydrate synthesis in complemented strains. *Eukaryotic Cell*. 9(8), 1251-1261.
- Wyman, S.K., Jansen, R.K., Boore, J.L. (2004) Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 20:3252-3255.
- Zerbino, D.R. and Birney, E. (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*. 18:821-829.
- Xiong, W., Li, X., Xiang, J., Wu, Q. (2008) High-density fermentation of microalga *Chlorella protothecoides* in bioreactor for microbio-diesel production. *Applied Microbiology and Biotechnology*. 78 (1), 29-36.
- Xu, H., Miao, X., Wu, Q. (2006) High quality biodiesel production from a microalga *Chlorella protothecoides* by heterotrophic growth in fermenters. *Journal of Biotechnology*. 126 (4), 499-507.
- Xu, N., Zhang, X., Fan, X., Han, L., Zeng, C. (2001) Effects of nitrogen source and concentration on growth rate and fatty acid composition of *Ellipsoidion* sp. (Eustigmatophyta). *Journal of applied phycology*. 13 (6), 463-469.
- Yonghua, L.-B. (2012) Triacylglycerol biosynthesis in eukaryotic microalgae. http://lipidlibrary.aocs.org/plantbio/tag_algae/index.htm.
- Zheng, Y., Li, T., Yu, X., Bates, P.D., Dong, T., Chen, S. (2013). High-density fed-batch culture of a thermotolerant microalga *Chlorella sorokiniana* for biofuel production. *Applied Energy*. 108, 281–287.

Zhang, L., Wang, X., Qian, H., Chi, S., Liu, C., Liu, T. (2012) Complete sequences of the mitochondrial DNA of the wild *Gracilariopsis lemaneiformis* and two mutagenic cultivated breeds (Gracilariaceae, Rhodophyta). PLoS ONE. 7, e40241.