



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Biologia

SCUOLA DI DOTTORATO DI RICERCA IN: BIOSCIENZE E BIOTECNOLOGIE

INDIRIZZO: BIOCHIMICA E BIOFISICA

CICLO XXVIII

# Biotechnological optimization of microalgae for the sustainable production of biocommodities

**Direttore della Scuola:** Ch.mo Prof. Paolo Bernardi

**Coordinatore d'indirizzo:** Ch.mo Prof. Fabio Di Lisa

**Supervisore:** Ch.mo Prof. Tomas Morosinotto

**Dottorando:** Giorgio Perin

*Ai miei genitori*



# Table of contents

<b>Abstract</b>	<b>III</b>
<b>Riassunto</b>	<b>IX</b>
<b>CHAPTER I</b>	<b>1</b>
<b>Introduction</b>	
<i>An algae-based economy has the potential to mitigate the environmental impact of fossil fuels exploitation.</i>	<b>3</b>
<i>Microalgae as feedstock for high-value natural products.</i>	<b>7</b>
<i>Microalgae as promising proteins over-expression hosts.</i>	<b>8</b>
<i>Microalgae genetic engineering.</i>	<b>11</b>
<i>Biotechnological optimization of the light energy to biomass conversion efficiencies.</i>	<b>11</b>
<i>Biotechnological optimization of metabolic pathways.</i>	<b>16</b>
<i>Nannochloropsis gaditana, a promising target for industrial applications.</i>	<b>21</b>
<b>CHAPTER II</b>	<b>35</b>
<b>Generation of Random Mutants to Improve Light-Use Efficiency of <i>Nannochloropsis gaditana</i> Cultures for Biofuel Production</b>	
Perin G. et al. Biotechnol Biofuels (2015) 8:161 - Doi 10.1186/S13068-015-0337-5	
<b>CHAPTER III</b>	<b>55</b>
<b>Connection Between Photosynthetic Properties and Biomass Productivity in <i>Nannochloropsis gaditana</i> Cultures in Industrially Relevant Conditions</b>	
<b>CHAPTER IV</b>	<b>77</b>
<b>Forward Genetics in <i>Nannochloropsis gaditana</i> Reveal a Splicing Variant of Protein Lhcx1 as Likely Responsible of a Mutant NPQ Inactivation</b>	

<b>CHAPTER V</b>	<b>99</b>
<b>Light Remodels Lipid Biosynthesis in <i>Nannochloropsis gaditana</i> by Modulating Carbon Partitioning Between Organelles</b>	
<b>CHAPTER VI</b>	<b>167</b>
<b>Investigation of Transcription Regulatory Sequences in <i>Nannochloropsis gaditana</i></b>	
<b>CHAPTER VII</b>	<b>191</b>
<b>LHCX3, the First Stress-Related LHC Fused to a Fasciclin 1-Like Domain, Identified in <i>Nannochloropsis Gaditana</i></b>	
<b>CHAPTER VIII</b>	<b>209</b>
<b>A Novel Micro-Photobioreactor Design for Assessing Microalgae Response to Light Intensity: the Case of <i>Nannochloropsis gaditana</i></b>	
<b>APPENDIX I</b>	<b>233</b>
<b>An Identifiable State Model to Describe Light Intensity Influence on Microalgae Growth</b> Bernardi A. et al. Ind Eng Chem Res. (2014); 53(16): 6738–6749 - Doi: 10.1021/ie500523z	
<b>APPENDIX II</b>	<b>247</b>
<b>Response of the Photosynthetic Apparatus to Salt Stress in a Resistant and Susceptible <i>Oryza sativa</i> variety</b>	
<b>APPENDIX III</b>	<b>261</b>
<b><i>In-Situ</i> Solid state NMR spectroscopy on intact photosynthetic thylakoid membranes</b>	
<b>Abbreviations</b>	<b>269</b>

## Abstract

The current global economy development trends strongly rely on fossil fuels exploitation, which are responsible for a net greenhouse gases (e.g. CO<sub>2</sub>) release in the atmosphere. In 2014, the Intergovernmental Panel on Climate Change (IPCC) stated that this net atmospheric CO<sub>2</sub> increase is anthropogenic, and it will lead to the rising of the global temperature of >2 °C, before the end of this century. The latter will strongly contribute to change the behavior of climate and oceans (e.g. leading to their acidification and oxygenation) in a permanent way, with a consequent magnification of the demographic pressures on food and water security, as well as on several ecosystems functional biodiversity.

To avoid this apocalyptic scenario, the development of renewable and clean energy sources to sustain a consistent part of the global economy is an unavoidable challenge for our society. A plant biomass-based economy could meet this need, but several studies predicted a food prices inflation and a concrete carbon debt, as consequence of this scenario. On the contrary, the exploitation of microalgae biomass could indeed avoid these issues and bring a positive effect on atmospheric CO<sub>2</sub> levels, leading to its sequestration and fixation in organic carbon. Currently, microalgae are indeed the best CO<sub>2</sub> sequestering organisms, thanks to higher photosynthetic rates with respect to plants. Their ability to grow on marginal lands and use wastewaters opens the doors for their application as environment-impact mitigating agents of current industrial processes. Their exploitation could be indeed the key for the development of an integrate process in which the biomass would be used to convert the majority of the current economy toward environmental-friendly processes.

Despite this promising scenario, mainly wild type microalgae species are currently available for these purposes. Their evolution in a natural environment, different from the artificial one exploited during their intensive industrial cultivation, strongly impairs their theoretical biomass productivities, leading to unsatisfactory values. The development of an algae-based economy indeed depends on the efficient conversion of light energy into biomass and the optimization of metabolic pathways to maximize the synthesis of the products of interest. Still far from the development of an economically competitive and energetically sustainable microalgae industrial cultivation, these organisms therefore need to undergo a biotechnological optimization to achieve the competitiveness threshold.

Although there isn't an ideal species that could serve to meet all human needs, we focused on the seawater species *Nannochloropsis gaditana*, which is a promising candidate for both basic biological and applied investigations. The following PhD thesis was conceived to provide a molecular insight on

*N. gaditana* photosynthetic efficiency and metabolic regulation, to provide the molecular targets for its biotechnological improvement. The prerequisite of this experimental work was the optimization of the currently available molecular toolkit for *N. gaditana* genetic manipulation and also the available molecular information was therefore significantly improved.

In chapter II random mutagenesis approaches were performed in order to isolate mutant strains with photosynthetic phenotypes, likely more suitable for intensive growth conditions. Selected strains indeed showed an improved photosynthesis in limiting growth conditions, also serving as biological tools to improve the available information on the underlying molecular elements controlling light-use efficiency in this organism. When tested in lab-scale growth conditions, the strain E2, selected for a reduction in its Chl content, indeed showed an improved biomass productivity, therefore representing a proof of concept for the developed biotechnological approaches, to be able to improve the photosynthetic performances of this organism in the artificial environment of a photobioreactor. Among the isolated mutant strains, in chapter III two major photosynthetic phenotypes, the reduction in the Chl content (strain E2) and the inability to activate NPQ (strain I48), were chosen as selection criteria to improve light penetration and energy conversion mechanisms, respectively, in an intensive culture. Both indeed showed an enhanced biomass productivity in industrial-simulating cultures, proving the theoretical advantage underneath their exploitation. However, when the growth conditions changed, leading to altering the light availability, the selected mutants altered their behavior. They weren't able of performing better than the WT in all the tested conditions. This would explain why the data published to date for the same mutants in other species often provided contrasting results. We concluded that photosynthetic mutants can modulate their phenotype in relation to the growth conditions and some of the latter could indeed highlight their drawbacks rather than their benefits, therefore the genetic engineering efforts have to be tailored properly to the growth conditions used.

The forward genetics strategy here developed could open the doors toward the identification of the molecular basis regulating photosynthesis in this promising species. In chapter IV mutant strain I48 was further investigated for the identification of the genetic basis of its phenotype. Thanks to the whole genome re-sequencing we identified a splicing variant in the 5'-donor splicing site of the 4<sup>th</sup> intron of the gene *Naga\_100173g12*, which encodes for the LHCX1 protein. This mutation caused the retention of the intron sequence, leading to a truncated protein product, which is likely degraded. The absence of the LHCX1 protein strongly correlates with inability to activate NPQ, since this protein's clade is well known to be involved in the activation of this mechanism. However, the future complementation of

the phenotype will serve to validate this conclusion. Moreover, the LHCX1 protein was found co-localized with the PSI in *N. gaditana*, therefore strain I48 could also serve as an optimal tool to investigate further its biological role.

To develop an efficient biotechnological optimization strategy, the information on the metabolism regulation of *N. gaditana* has to be highly enriched. Understanding the metabolic fluxes direction could lead to specifically affect those involved in a specific product accumulation, without affecting other pathways, leading to a possible negative impact on growth. In chapter V an integrated analysis of genome-wide, biochemical and physiological approaches helped us in deciphering the metabolic remodeling of *N. gaditana* that switches its metabolism toward a greater lipid production in excess light conditions. The latter indeed induced the accumulation of DAGs and TAGs, together with the up-regulation of genes involved in their biosynthesis. We saw the induction of cytosolic fatty acids synthase (*FAS1*) genes and the down-regulation of those of the chloroplast (*FAS2*). Lipid accumulation is accompanied by the regulation of triose phosphate/inorganic phosphate transport across the chloroplast membranes, tuning the carbon metabolic allocation between cell compartments and favoring the cytoplasm and endoplasmic reticulum at the expense of the chloroplast. This highlighted the flexibility of *N. gaditana* metabolism to respond to environmental needs.

In chapter VI the information gained from the latter work was exploited to test the potentiality of this prokaryotic species also as protein expression platform. We built up a modular system for protein overexpression in which the regulatory sequences were chosen among those which revealed to induce a high level of transcription or to be highly regulated by light availability. *N. gaditana* revealed to be a very promising host for protein expression, given the higher luciferase activity monitored with respect to the reference species for these applications, *C. reinhardtii*. A method to test the efficacy of several regulatory sequences in driving proteins expression was developed, as well as several expression vectors, which are ready to be tested.

The investigation of the *N. gaditana* metabolism regulation in chapter V, showed a fine tuning of its photosynthetic apparatus components, in different light conditions. Focusing on LHC proteins we identified a new LHCX protein in this species, called LHCX3 (GENE ID: Naga\_101036g3), whose gene coding sequence wasn't annotated correctly. In chapter VII the correct coding sequence of this gene was further investigated and experimentally validated with molecular techniques. The LHCX3 protein revealed to be fused with an N-terminal fasciclin I-like domain and a sequence analysis together with a preliminary evolution study was performed to infer the biological role of this association.



Since algae metabolism entirely relies on light availability, the importance of investigating the light intensity effect on growth is seminal for their industrial application. In chapter VIII we developed a micro-scale platform, that we called micro-photobioreactor, to easily investigate the impact of light intensity on *N. gaditana* growth. We were able to test the effect of different light regimes, simultaneously, also on the photosynthetic performances in an integrated system which could be merged with nutrients availability studies, speeding up the *N. gaditana* characterization process. Three appendix sections are also included in the thesis in which some of the experimental techniques exploited in this work were applied to different organisms toward the common target of investigating light-use efficiency and the molecular elements involved in its regulation.

In appendix I, the development of a mathematical prediction model for growth and fluorescence data of the species *Nannochloropsis salina* was described. The work was carried out in collaboration with Prof. Fabrizio Bezzo of the industrial engineering department of the University of Padova.

The development of behavior prediction models representing the phenomena affecting algae growth, could be very helpful in designing and optimizing the production systems at industrial level.

The developed model well represented *N. salina* growth over a wide range of light intensities, and could be further implemented to describe also the influence on growth of other parameters, such as nutrients availability and mixing.

In appendix II, the monitoring of the *in vivo* chlorophyll fluorescence was exploited to study the photosynthetic features of rice plants exposed to salt stress conditions. The presented results are part of a wider project (in collaboration with Prof. Fiorella Lo Schiavo from the biology department of the University of Padova), aiming to depict the physiological, biochemical and molecular remodeling, undergoing in one of the major food crop in the world, in response to salt stress conditions.

Depicting the impact of environmental stresses on photosynthesis is seminal to control biomass productivity since plants metabolism strongly relies on the former for growth. We showed the activation of the NPQ mechanism in salt tolerant plants, highlighting the importance of photosynthetic features monitoring to predict plants performances, directly on the field.

In appendix III, *Chlamydomonas reinhardtii*  $^{13}\text{C}$ – $^{15}\text{N}$  labeled thylakoids were isolated from the *cw15* and *npq2* mutant strain in order to study their structure and dynamics in terms of protein and lipid components *in situ*, by applying the solid-state NMR technique, in collaboration with Prof. Anjali Pandit group from the Leiden Institute of Chemistry. These analyses will serve to investigate the photosynthetic membranes remodeling that undergoes from an active (*cw15* strain) to a photo-

protective state (*npq2* mutant strain), during the switch toward excess light conditions, with the final aim to understand the biochemical processes regulating this event.



## Riassunto

Le attuali tendenze di sviluppo economico mondiale si basano fortemente sullo sfruttamento di combustibili fossili, che sono responsabili di un netto rilascio di gas serra (ad esempio CO<sub>2</sub>) nell'atmosfera.

Nel 2014, il gruppo intergovernativo di esperti sui cambiamenti climatici (IPCC), ha dichiarato che questo incremento netto di CO<sub>2</sub> atmosferica è di origine antropogenica, e che porterà all'aumento della temperatura globale di > 2 °C, prima della fine di questo secolo. Quest'ultimo contribuirà fortemente alla modifica del comportamento del clima e degli oceani (ad esempio portando alla loro acidificazione ed ossigenazione) in modo permanente, con conseguente aumento delle pressioni demografiche sulla sicurezza alimentare e idrica, nonché sulla bio-diversità funzionale di diversi ecosistemi.

Per evitare questo scenario apocalittico, lo sviluppo di fonti energetiche rinnovabili e pulite per sostenere una parte consistente dell'economia globale è una sfida inevitabile per la nostra società. Un'economia basata sulla biomassa vegetale potrebbe soddisfare questa esigenza, ma diversi studi hanno previsto un'inflazione dei prezzi alimentari e un concreto debito di carbonio, come conseguenza di questo scenario. Al contrario, lo sfruttamento della biomassa delle microalghe potrebbe in realtà evitare questi problemi e portare ad un effetto positivo sui livelli di CO<sub>2</sub>, portando al suo sequestro e fissazione in carbonio organico. Attualmente, le microalghe sono infatti i migliori organismi capaci di sequestrare la CO<sub>2</sub>, grazie a tassi fotosintetici più elevati rispetto alle piante. La loro capacità di crescere su terreni marginali e sfruttando le acque di scarto apre le porte per la loro applicazione come agenti in grado di mitigare l'impatto ambientale degli attuali processi industriali. Il loro impiego potrebbe essere infatti la chiave per lo sviluppo di un processo integrato in cui la biomassa dovrebbe essere usata per convertire la maggior dell'economia attuale verso processi rispettosi dell'ambiente.

Nonostante questo scenario promettente, attualmente soprattutto specie di microalghe di tipo selvatico sono impiegate a questi scopi. La loro evoluzione in un ambiente naturale, diverso da quello artificiale sfruttato durante la coltivazione industriale intensiva, ostacola fortemente le produttività teoriche della biomassa, portando a valori insoddisfacenti. Lo sviluppo di un'economia basata sulle alghe infatti dipende dall'efficiente conversione dell'energia luminosa in biomassa e dall'ottimizzazione delle vie metaboliche per massimizzare la sintesi dei prodotti di interesse.

Ancora lontani dallo sviluppo di una coltivazione industriale di microalghe economicamente competitiva ed energeticamente sostenibile, questi organismi devono quindi subire un'ottimizzazione biotecnologica per raggiungere la soglia di competitività. Sebbene non ci sia una specie ideale che potrebbe servire a soddisfare tutti i bisogni umani, noi ci siamo concentrati sulla specie di acqua marina *Nannochloropsis gaditana*, che è un candidato promettente sia per indagini biologiche di base che applicative. La seguente tesi di dottorato è stata concepita per fornire un'investigazione molecolare sull'efficienza fotosintetica e la regolazione metabolica di *N. gaditana*, al fine di fornire i bersagli molecolari per il suo miglioramento biotecnologico. Il presupposto di questo lavoro sperimentale era l'ottimizzazione degli strumenti molecolari attualmente disponibili per la manipolazione genetica di *N. gaditana* e anche le informazioni molecolari disponibili sono state quindi significativamente migliorate.

Nel capitolo II è stato eseguito un approccio di mutagenesi casuale al fine di isolare ceppi mutanti con fenotipi fotosintetici, possibilmente più adatti alle condizioni di crescita intensiva. I ceppi selezionati in effetti hanno mostrato un miglioramento della fotosintesi in condizioni di crescita limitanti, rappresentando anche degli strumenti biologici per migliorare le informazioni disponibili sugli elementi molecolari alla base del controllo dell'efficienza dell'uso della luce in questo organismo. Quando testato in condizioni di crescita su scala di laboratorio, il ceppo E2, selezionato per una riduzione del contenuto in clorofilla, ha infatti mostrato un miglioramento della produttività della biomassa, rappresentando quindi una conferma per gli approcci biotecnologici qui sviluppati, di essere in grado di migliorare le prestazioni fotosintetiche di questo organismo, nell'ambiente artificiale di un fotobioreattore. Tra i ceppi mutanti isolati, nel capitolo III, due principali fenotipi fotosintetici, la riduzione nel contenuto di clorofilla (ceppo E2) e l'impossibilità di attivare il meccanismo di NPQ (ceppo I48), sono stati scelti come criteri di selezione per migliorare, rispettivamente, i meccanismi di penetrazione della luce e di conversione dell'energia luminosa in una cultura intensiva. Entrambi infatti hanno mostrato una produttività della biomassa maggiore nelle culture che simulano le condizioni industriali, dimostrando il vantaggio teorico dato dal loro sfruttamento. Tuttavia, quando le condizioni di crescita sono state cambiate, variando la disponibilità di luce, i mutanti selezionati hanno alterato il loro comportamento. Non erano in grado di essere più produttivi del WT in tutte le condizioni testate. Questo spiegherebbe perché i dati pubblicati fino ad oggi per gli stessi mutanti, isolati in altre specie, spesso forniscano risultati contrastanti.

Abbiamo quindi concluso che i mutanti fotosintetici possono modulare il loro fenotipo in relazione alle condizioni di crescita e alcune di queste potrebbero infatti evidenziare i loro svantaggi piuttosto

che i loro benefici. Pertanto, i futuri approcci di ingegneria genetica dovranno essere adattati adeguatamente alle condizioni di crescita utilizzate.

La strategia di genetica diretta qui sviluppata potrebbe aprire le porte verso l'individuazione delle basi molecolari che regolano la fotosintesi in questa specie promettente. Nel capitolo IV il ceppo mutante I48 è stato ulteriormente indagato per individuare le basi genetiche responsabili del suo fenotipo. Grazie al ri-sequenziamento del suo intero genoma abbiamo identificato una variante di *splicing* nel sito di *splicing* donatore al 5' del 4° introne del gene *Naga\_100173g12*, che codifica per la proteina LHCX1. Questa mutazione ha causato la ritenzione della sequenza dell'introne, portando ad un prodotto proteico tronco, che viene probabilmente degradato. L'assenza della proteina LHCX1 si correla fortemente con l'incapacità di attivare l'NPQ, dal momento che questo gruppo di proteine è ben noto per essere coinvolto nell'attivazione di tale meccanismo. Tuttavia, la futura complementazione del fenotipo servirà per validare questa conclusione. Inoltre, la proteina LHCX1 è stata trovata co-localizzata con il PSI in *N. gaditana*, quindi il ceppo I48 potrebbe anche servire come strumento ottimale per indagare ulteriormente il ruolo biologico di questa associazione.

Per sviluppare un'efficace strategia di ottimizzazione biotecnologica, le informazioni relative alla regolazione del metabolismo di *N. gaditana* devono essere notevolmente arricchite. Capire la direzione dei flussi metabolici potrebbe permettere di colpire in particolare solo quelli che sono coinvolti nell'accumulo di un determinato prodotto, senza influenzare altre vie metaboliche, cosa che potrebbe portare ad un possibile impatto negativo sulla crescita. Nel capitolo V un'analisi integrata con approcci genomici, biochimici e fisiologici ci ha aiutato a decifrare il rimodellamento metabolico di *N. gaditana*, che porta il suo metabolismo verso una maggiore produzione di lipidi in condizioni di luce in eccesso. Quest'ultima condizione infatti induce l'accumulo di di-acilgliceroli (DAG) e tri-acilgliceroli (TAG), insieme alla sovra-regolazione di geni coinvolti nella loro biosintesi. Abbiamo visto l'induzione di geni del complesso citosolico dell'acido grasso sintasi (FAS1) e la sub-regolazione di quelli del cloroplasto (FAS2). L'accumulo di lipidi è accompagnato dalla regolazione di trasportatori di trioso fosfati / fosfato inorganico attraverso le membrane del cloroplasto, inducendo la ripartizione metabolica del carbonio tra compartimenti cellulari e favorendo il citoplasma ed il reticolo endoplasmatico a spese del cloroplasto. Ciò ha evidenziato la flessibilità del metabolismo di *N. gaditana* al fine di rispondere alle esigenze ambientali.

Nel capitolo VI le informazioni acquisite da quest'ultimo capitolo sono state sfruttate per testare le potenzialità di questa specie promettente, anche come piattaforma di espressione di proteine. Abbiamo costruito un sistema modulare per la sovra-espressione di proteine in cui le sequenze

regolatrici sono state scelte tra quelle che inducono un elevato livello di trascrizione o che sono altamente regolate dalla disponibilità di luce. *N. gaditana* si è rivelata essere un organismo molto promettente per l'espressione di proteine, data la maggiore attività luciferasica osservata, rispetto alla specie di riferimento per tali applicazioni, *C. reinhardtii*. È stato sviluppato un metodo per testare l'efficacia di diverse sequenze regolatrici nel guidare l'espressione di proteine così come sono stati preparati diversi vettori di espressione, pronti per essere testati.

L'indagine della regolazione del metabolismo di *N. gaditana*, svolta nel capitolo V, ha mostrato una fine regolazione dei suoi componenti dell'apparato fotosintetico, in diverse condizioni di luce. Concentrandosi sulle proteine antenna, coinvolte nella cattura della luce (*light-harvesting complex* (LHC) proteins) abbiamo identificato una nuova proteina LHCX in questa specie, chiamata LHCX3 (ID del gene: Naga\_101036g3), la cui sequenza codificante non era annotata in modo corretto.

Nel capitolo VII, la corretta sequenza codificante di questo gene è stata studiata ulteriormente e convalidata sperimentalmente con tecniche molecolari. La proteina LHCX3 ha rivelato la presenza di un dominio simile a quello di tipo fasciclina I, all'N-terminale, e sono stati eseguiti un'analisi di sequenza insieme ad uno studio evolutivo preliminare per dedurre il ruolo biologico di questa associazione.

Poiché il metabolismo delle alghe si basa interamente sulla disponibilità di luce, è fondamentale indagare l'effetto dell'intensità della luce sulla crescita per studiare la loro applicazione industriale. Nel capitolo VIII, abbiamo sviluppato una piattaforma su micro-scala, che abbiamo chiamato micro-fotobioreattore, per indagare facilmente l'impatto dell'intensità della luce sulla crescita di *N. gaditana*. Siamo stati in grado di testare simultaneamente l'effetto di diversi regimi di luce, anche sulle prestazioni fotosintetiche, in un sistema integrato che potrebbe essere associato a studi sull'impatto delle sostanze nutritive, accelerando il processo di caratterizzazione di *N. gaditana*.

Nella tesi sono state incluse anche tre sezioni di appendice, in cui alcune delle tecniche sperimentali sfruttate in questo lavoro sono state applicate a diversi organismi, con l'obiettivo comune di indagare l'efficienza dell'uso della luce e gli elementi molecolari coinvolti nella sua regolazione.

Nell'appendice I, è stato descritto lo sviluppo di un modello matematico per la previsione dei dati di crescita e di fluorescenza, nella specie *Nannochloropsis salina*. Il lavoro è stato svolto in collaborazione con il Prof. Fabrizio Bezzo del dipartimento di ingegneria industriale dell'università di Padova.

Lo sviluppo di modelli di previsione del comportamento che rappresentino i fenomeni che influenzano la crescita delle alghe, potrebbe essere molto utile nella progettazione e ottimizzazione dei sistemi di produzione a livello industriale. Il modello sviluppato descrive bene la crescita di *N. salina* in un ampio

intervallo di intensità di luce, e potrebbe essere ulteriormente implementato per descrivere anche l'influenza sulla crescita di altri parametri, come la disponibilità di nutrienti ed il mescolamento. Nell'appendice II, il monitoraggio della fluorescenza della clorofilla *in vivo* è stato sfruttato per studiare le caratteristiche fotosintetiche di piante di riso, esposte a condizioni di stress salino. I risultati presentati sono parte di un progetto più ampio (in collaborazione con la Prof. Fiorella Lo Schiavo del dipartimento di biologia dell'università di Padova), con l'obiettivo di rappresentare il rimodellamento fisiologico, biochimico e molecolare, subito da una delle principali colture destinate ad usi alimentari nel mondo, in risposta a condizioni di stress salino. La descrizione dell'impatto degli stress ambientali sulla fotosintesi è fondamentale per controllare la produttività della biomassa delle piante poiché il loro metabolismo dipende interamente dalla fotosintesi per la crescita. Abbiamo osservato l'attivazione del meccanismo di NPQ in piante tolleranti il sale, sottolineando l'importanza del monitoraggio delle caratteristiche fotosintetiche per prevedere le prestazioni delle piante, direttamente sul campo.

Nell'appendice III, sono stati isolati i tilacoidi di *Chlamydomonas reinhardtii* marcati con atomi  $^{13}\text{C}$  -  $^{15}\text{N}$ , sia dal ceppo *cw15* che dal ceppo mutante *npq2*, al fine di studiare la loro struttura e la dinamica delle proteine e dei lipidi costituenti, *in situ*, applicando la tecnica dell'NMR allo stato solido, in collaborazione con il gruppo del Prof. Anjali Pandit dell'Istituto di chimica di Leiden. Queste analisi serviranno per studiare il rimodellamento delle membrane fotosintetiche che passano da uno stato attivo (ceppo *cw15*) ad uno stato foto-protettivo (ceppo mutante *npq2*), durante il passaggio verso condizioni di luce in eccesso, con lo scopo finale di comprendere i processi biochimici che regolano quest'evento.





# CHAPTER I

## Introduction

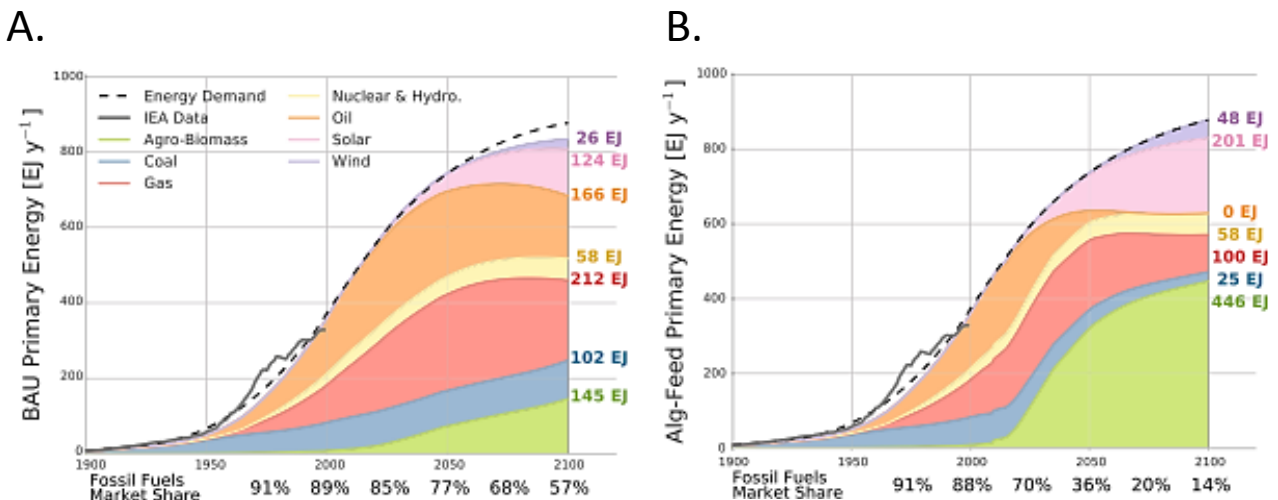


An algae-based economy has the potential to mitigate the environmental impact of fossil fuels exploitation.

The massive use of fossil fuels is responsible for the accumulation of greenhouse gases and other pollutants in the atmosphere (Moss et al., 2010). The latter are responsible of serious environmental consequences, such as oceans acidification and deoxygenation and global temperature increase, leading to possible extreme weather events (Levy and Patz; Allison and Bassett, 2015).

Based on currently available economical models, if the future economy is assumed to develop along historical trends and without any kind of perturbations (BAU scenario; Business-As-Usual), there will be a net increase in anthropogenic atmospheric emissions of the 30 % from 2010 to 2100, with a consequent rise in the mean global temperature of up to 3 °C (Walsh et al., 2015).

Despite this scenario, the market share of fossil fuels will however decrease from 90 %, at the beginning of this century (2000), to the 57 % in 2100, mainly for supply limitations. This would lead to a consequent increase in other energy sources exploitation (Figure 1A), but this fact will however not be sufficient to mitigate the environmental effect of the remaining 57% of fossil fuels impact in global economy.



**Figure 1. Total annual primary energy production through 2100 in the BAU (A) and Alg-Feed (B) scenarios.** The BAU (Business-As-Usual) plot represents the scenario in which the future global population growth, energy profiles and agricultural yields are assumed to develop along historical trends and without any kind of perturbations. The Alg-Feed plot represents the scenario in which algae biomass will be used for energy and animal feed, meeting the 40 % of the global demand for the latter. The numbers on the right of the plots display the primary supply of each energy source. Data for the historical trends were taken from IEA (International Energy Agency - <http://www.iea.org>)

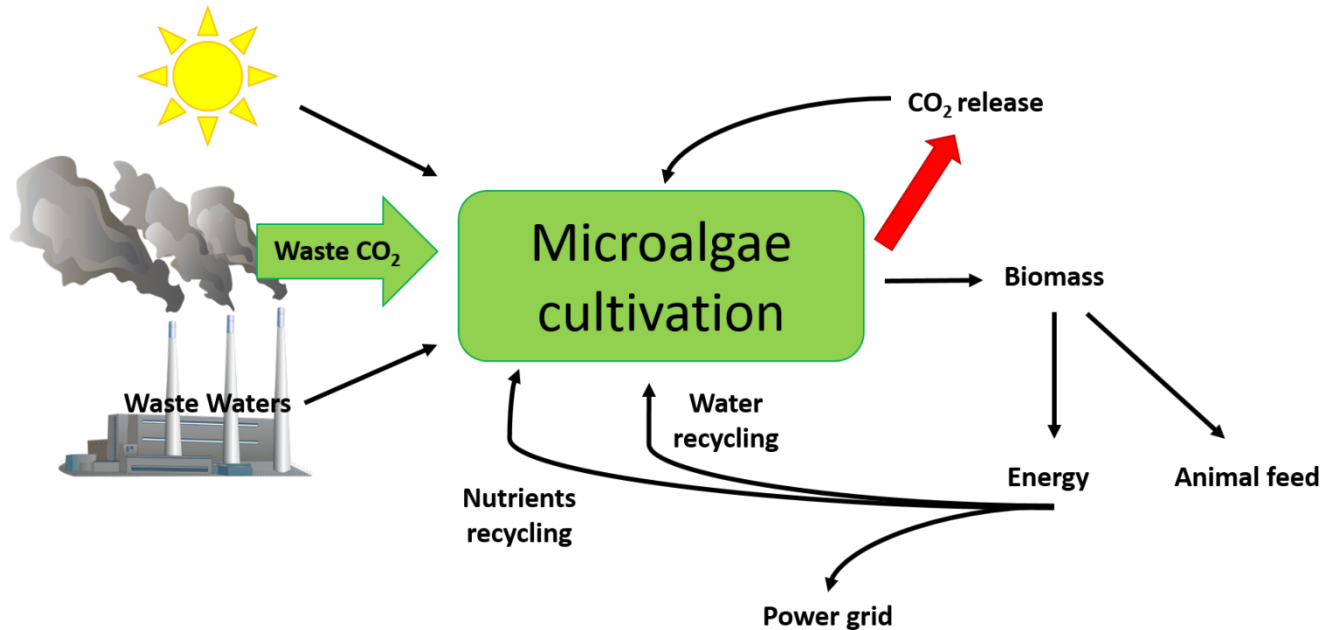
- Accessed 1 Oct 2015), while the expected energy demand was calculated from the population rising and energy needs predictions. Primary energy production is reported in Exajoules per year (EJ y<sup>-1</sup>); 1 EJ = 10<sup>18</sup> J. Adapted from (Walsh et al., 2015).

Therefore, the BAU scenario foresees a strong impact on global climate, leading to the magnification of the demographic pressures on food and water security, as well as on forests and other ecosystems and functional bio-diversity. To avoid this apocalyptic scenario, recently the Intergovernmental Panel on Climate Change (IPCC) stated that the net anthropogenic carbon emissions must be eliminated before the end of this century in order to limit the world mean temperature increase to less than 2 °C (<http://mitigation2014.org/report/>).

Therefore, the development of renewable and clean energy sources to sustain a consistent fraction of the global economy is an unavoidable challenge for our society to reduce the environmental consequences of massive fossil fuels exploitation. Among the different possible alternatives, plants biomass represents a promising possibility. However, its large exploitation for energy demands is in competition with food and feed productions for agricultural land use, risking the inflation of food prices (Kazamia and Smith, 2014). To respond to higher prices, farmers worldwide would convert present forests and grasslands to new croplands, with the final aim to replace the biomass diverted to biofuels (Fargione et al., 2008). However, since soils and plant biomass are the two largest biologically active stores of terrestrial carbon (Schlesinger and Bernhardt, 2013) converting native habitats to croplands would release a huge amount of CO<sub>2</sub> in the atmosphere, as a result of burning and consequent biomass microbial decomposition. Applying worldwide agricultural models it was estimated that this land-use change would therefore not save the current situation but would double the current greenhouse gas emissions over 30 years (e.g. for corn-based ethanol), therefore leading to a strong carbon debt only repairable in a long-term perspective (Searchinger et al., 2008).

Therefore, only the biomass coming from marginal lands exploitation would lead to save the current CO<sub>2</sub> emissions. In this context, algae biomass is emerging as a promising alternative. Microalgae, as well as plants, are photosynthetic organisms supporting their metabolism with the energy from sunlight and fix atmospheric CO<sub>2</sub>. Among the known carbon sequestering technologies, microalgae photosynthesis is indeed the cheapest and most efficient alternative. Microalgae CO<sub>2</sub> capturing efficiencies were shown to be indeed as high as 90 % in open ponds (Salih, 2011) and therefore their exploitation could lead to a positive balance on greenhouse gas emissions. Moreover, unlike plants, algae are also able to extract nutrients from civil and industrial wastewaters and are particularly efficient in removing nitrates and phosphates (Ramos Tercero et al., 2014). All these features make

them promising tools for the development of an environmental-friendly and energetically sustainable economy (Figure 2).



**Figure 2. Schematic representation of microalgae cultivation to achieve an environmental-friendly economy.** The process includes energy production from algae biomass and recycling of different industrial and civil wastes. However, despite the microalgae cultivation process still induces a CO<sub>2</sub> release, mainly for down-stream harvesting processes or for the temperature control, it can be recycled, leading to a no net CO<sub>2</sub> increase in the atmosphere.

In fact, in an algae-based economy prediction, in which algae biomass will serve to satisfy a fraction of the energy demands and the 40 % of the global animal feed (Alg-Feed scenario), fossil fuels market share would be reduced to the 14 % in 2100 (Figure 1B). However, this scenario is expected to lead to increase the global temperature of 2.3 °C as well. Therefore the development of an algae-based economy is not alone enough to stay below the 2.0 °C of global temperature increase (red and yellow boxes in Table 1), imposed as threshold from IPCC in 2014 (<http://mitigation2014.org/report/>).

There is indeed the need to couple it with Carbon Capture and Storage technologies (CCS), able to strongly reduce the energy sector carbon emissions (Table 1).

The latter include geological, ocean storage and mineral carbonation ([https://www.ipcc.ch/pdf/special-reports/srccs/srccs\\_wholereport.pdf](https://www.ipcc.ch/pdf/special-reports/srccs/srccs_wholereport.pdf)). Only in this scenario it would be possible to reduce the net CO<sub>2</sub> concentration in the atmosphere to pre-industrial levels before the end of the century, limiting the world mean temperature increase to values lower than 2

°C (green boxes in Table 1). The extent of the benefit given by the integration between these two technologies will depend on the fraction of the global economy that will rely on algae-based processes (Table 1). However, the conversion of the global economy toward the exploitation of algae-based processes is still in its infancy and strong biological efforts are needed to let it become concrete.

**Table 1. Projected CO<sub>2</sub> concentrations in the atmosphere for year 2100.** CO<sub>2</sub> concentrations are calculated from different scenarios in which energy sector emissions are mitigated in different percentages (through Carbon Capture and Storage technologies - CCS), coupling it with a business-as-usual (BAU) economy, or with the development of an algae-based economy (Alg-feed), applied for energy and animal feed purposes, to cover up to the 40 % of global demand. The CO<sub>2</sub> concentrations are expressed in ppm. Boxes in red, yellow and green indicate scenarios in which the mean world temperature will rise of > 2.0, 2.0 ± 0.2 or < 1.8 °C, respectively. Adapted from (Walsh et al., 2015).

		Energy sector emissions mitigation (CCS)						
		0 %	25 %	35 %	45 %	55 %	65 %	75 %
<b>BAU</b>		624	563	537	514	491	470	445
<b>Alg-Feed 10 %</b>		513	458	435	415	395	376	355
<b>Alg-Feed 20 %</b>		496	442	418	398	378	359	337
<b>Alg-Feed 30 %</b>		484	429	405	385	364	345	323
<b>Alg-Feed 40 %</b>		473	417	393	372	351	332	310

### *Microalgae as feedstock for high-value natural products.*

Besides their worldwide known potential as platforms for developing environmental-friendly processes, microalgae are presently on the market only for high added value molecules production, such as food and animal feed additives, cosmetics, pigments and fatty acids (Vanthoor-Koopmans et al., 2013; Gimpel et al., 2015). To date, the most exploited microalgae species, as feedstock for high-value products, is *Chlorella vulgaris*, which belongs to the green lineage. Together with the cyanobacterium *Arthrospira platensis*, also known as *Spirulina platensis*, the former represents the most common microalgae species to be used as diet supplement (Gimpel et al., 2015). Both are being used for years for this purpose, since able to provide natural food and feed additives in replacement of those of synthetic origin (Yaakob et al., 2014). Moreover, several recent reports discussed the application of microalgae for health purposes, demonstrating the positive effect of microalgae biomass as dietary supplement to reduce cancer and cardiovascular diseases occurrence in humans (Raposo and de Morais, 2015; Luo et al., 2015), by increasing natural killer cells titers and anti-inflammatory response (Kwak et al., 2012). Besides the biomass direct exploitation, microalgae can be also processed to purify the high-value molecules they synthesize. The market share of products enriched in the desired high-value molecules is very large (the global carotenoid market value was \$ 1.5 billion in 2014 and it is expected to reach nearly \$ 1.9 billion in 2019 - <http://www.bccresearch.com/market-research/food-and-beverage/carotenoids-global-market-report-fod025e.html> - Accessed 13/01/2016) and their production is expected to be very promising (Table 2).

**Table 2. Estimated annual production of microalgae high value products.** The estimation assumes that the microalgal biomass production will be 175 Mt and that the 10 % of it will be dedicated for each product. Mt = million tons, Kt = kilotons (one thousand tons). The content is expressed as the best biomass percentage currently achievable in the desired product. Adapted from (Markou and Nerantzis, 2013).

High-value product	Content	Annual Production (Kt)
<b><math>\beta</math>-carotene</b>	3.0 %	525
<b>Astaxanthin</b>	3.0 %	525
<b>Lutein</b>	0.5 %	87.5
<b>Bioplastics</b>	30.0 %	5250
<b>Glycerol</b>	30.0 %	5250
<b>Vitamins</b>	1.0 %	175



For instance, several methods have been developed so far for carotenoids or fatty acids purification from microalgae biomass (Kim et al., 2015). Among carotenoids,  $\beta$ -carotene is the most abundant antioxidant of microalgae origin, to be currently commercialized (Raposo et al., 2015). Its synthetic analog comprises only the all-*trans* isomer, which was recently reported to be responsible of negative effects on plasma membrane cholesterol levels in mice studies (Harari et al., 2008). The  $\beta$ -carotene purified from *Dunaliella salina*, which is the species currently mainly used for natural  $\beta$ -carotene production, is instead highly enriched in the 9-*cis* isomer, without negative health effects reported so far. A second example of carotenoid of microalgae origin, currently commercialized worldwide, is astaxanthin from the green microalgae *Haematococcus pluvialis*. The latter finds application as natural dye for salmon and shrimps pigmentation (Zhang and Wang, 2015). Besides carotenoids, microalgae biomass is an established sustainable feedstock also for fatty acids production.

Microalgae were shown to be highly enriched in omega-3 and omega-6 long chain polyunsaturated fatty acids, such as docosahexaenoic acid (DHA, C22:6; e.g. from *Isochrysis strain T-iso* and *Pavlova lutheri*), eicosapentaenoic acid (EPA, C20:5; e.g. *Nannochloropsis gaditana*, *Nannochloropsis oculata*), and alpha-linolenic acid (e.g. *Rhodomonas salina*, *Tetraselmis uecica*) (Tsai et al., 2016; Kagan et al.).

Given the high content of oils in microalgae biomass their commercialization could be a valuable alternative to those of fish origin, considering that intensive fishing is strongly depleting current oceans fish stocks (Ryckebosch et al., 2014).

Besides their established applications for nutraceutical purposes, microalgae fatty acids show huge impact on the biofuels market share, since easily convertible in biodiesel through simple chemical transesterification processes (Park et al., 2015; Medipally et al., 2015).

#### *Microalgae as promising proteins over-expression hosts.*

At present many molecules commercialized by biopharmaceutical and chemical industries are obtained through polluting chemical reactions, which often has also unsatisfactory final yields. Therefore the development of enzymatic processes, alternative to current chemical ones, is seminal to meet new competitiveness and safety thresholds (Ullrich et al., 2015).

This objective requires the availability of efficient expression platforms for recombinant enzymes production. To date, different expression systems have indeed been employed to produce a large variety of proteins, with particular interest for the production of health-promoting compounds.

The final aim is to obtain the desired product in the safest and cheapest way, together with convenient production yields (Rasala and Mayfield, 2015). Among the different alternatives, the use of

mammalian cells is advantageous to produce very complex molecules. However their cultivation is very expensive and it can be carried out only at very limited scale. Expression levels can be quite low and cell lines are often unstable, strongly susceptible to contamination constrains. Taken together these considerations limit mammalian cells applications mostly to the research field or proteins of extremely high value (Spadiut et al., 2014).

Bacteria are instead the best characterized expression system for proteins expression at industrial level, but they lack the molecular equipment for the major post-translational modifications, which are extremely important for the biological activity of complex enzymes (Spadiut et al., 2014).

Bacteria generally don't carry out disulphide bonds formation and glycosylation and often are inefficient in entire protein folding and secretion processes. Often complex recombinant proteins end up in insoluble inclusion bodies that necessitate expensive downstream processes for further applications. Concerning all these above mentioned constrains bacteria are best suited just for producing small and simple macromolecules.

Eukaryotic organisms, such as plants, in the past years have demonstrated to be valuable organisms for recombinant proteins production (Ullrich et al., 2015). However they require a long time from the transformation event to the obtainment of the mature genetically modified plant; it can indeed take between 1 and 3 years to obtain industrial relevant amounts of protein, as reported for tobacco or corn transgenic plants, respectively (Peters and Stoger, 2011). Moreover, genetically modified plants must comply complex containment procedures, to prevent the dispersion of the newly introduced genes through interspecific breeding events.

In this context, eukaryotic microorganisms emerged to meet both timescale and containment needs. The latter are safe and easily genetically modifiable through transformation, mutagenesis or breeding techniques (Böer et al., 2007). Furthermore they combine the advantages of an eukaryotic expression system, such as the ability to correctly fold complex proteins and perform post-translational modifications, with high yields and time-saving scalability, typical of microbial growing systems. Currently, the most widespread eukaryotic microorganisms used for these purposes are the yeasts *S. cerevisiae* and *P. pastoris*. However, they tend to hyper-glycosylate recombinant proteins, possibly affecting the correct protein folding leading to immunogenic effects.

Recently microalgae have attracted attention as platforms suitable for recombinant proteins production, since showing many desirable traits for an efficient, safe and cheap process development. As previously discussed, they are able to develop environmental-friendly processes, since avoiding expensive growth media and exploiting free and abundant energy and carbon sources.

Genetically engineered microalgae are easily containable in closed cultivation systems, therefore the risk of contamination of human food or animal feed supplies is extremely low.

Among them, *C. reinhardtii* owns the most advanced molecular toolkit for genetic manipulation. Therefore, it was the first species to be studied as candidate for recombinant proteins expression platform development. All its three genomes (nuclear, chloroplast and mitochondrial) have been fully sequenced and are transformable (Specht et al., 2010). However when the nuclear genome is transformed, the achieved recombinant proteins titres are very low, due to transgene silencing and positional effects, which are still poorly understood (Scranton et al., 2015). While the chloroplast transformation is not affected by these constrains, leading to higher accumulation levels (up to 1 % of total protein - (Young and Purton, 2015)).

However, when recombinant proteins are expressed in the chloroplast they aren't glycosylated and cannot be directed toward the secretory pathway or specific sub-cellular compartments (Scaife et al., 2015). Whether the glycosylation is necessary for their stability and enzymatic activity, also chloroplast transformation techniques thus show limited applications.

In the past years, *C. reinhardtii* nuclear transformation technologies have been improved, increasing the recombinant proteins titres (Gimpel et al., 2015). For instance, the fusion of the target coding sequence to a resistance gene, through the foot-and-mouth disease-virus 2A (FMDV 2A) self-cleavage peptide, led to a 100-fold increase in the target protein accumulation level (Rasala et al., 2012). Moreover, the recent development of a versatile and expandable vector toolkit for easy and time-saving manipulation approaches, improved the targeting efficiencies of a recombinant protein to ER / secretory pathway or any other sub-cellular compartment (chloroplast, nucleus, mitochondria, and intracellular micro-bodies - (Lauersen et al., 2013, 2015b)).

However, despite the strong experimental efforts devoted to *C. reinhardtii*, these positive results cannot be moved to any target recombinant protein, reducing therefore *C. reinhardtii* application to the expression of few industrial relevant proteins (e.g. xylanases (Rasala et al., 2012) or ice structuring proteins (Lauersen et al., 2015a)). This fact strongly limits the current industrial application of microalgae as expression platforms. However, these days are emerging studies on other microalgae species, able to accumulate industrially relevant enzymes to similar or higher levels than reported for *C. reinhardtii* (Gimpel et al., 2013), without having applied huge experimental efforts for expression system optimization, yet. Therefore, several other microalgal genera (*Chlorella*, *Scenedesmus* and *Dunaliella*) are currently under investigation, paving the way for the identification of more productive and robust protein expression systems.

### *Microalgae genetic engineering.*

Despite promising dry biomass productivities are indeed achievable in open cultivation systems in the southern of United States or also in closed plants operating in climates of the Benelux and Maghreb regions (Orfield et al., 2014; Slegers et al., 2015), these values are still below those needed to develop a cost-effective process based on microalgae biomass. In fact, even if the above discussed microalgae potentialities are now widely recognized, the cost-effective algae industrial applications are currently still limited to few high-value molecules production (carotenoids and keto-carotenoids, (Markou and Nerantzis, 2013)).

Currently, mainly wild type strains are used for industrial purposes. The latter are indeed not optimized for these applications, leading to the development of unsustainable processes. Therefore there is indeed a strong need to investigate on microalgae genetic improvement to meet the large-scale cultivation needs (Gimpel et al., 2015)(Gangl et al., 2015). It is seminal to select improved strains which perform better in the artificial environment of the photobioreactors, showing higher yields in the desired products. In fact, the competitive production of biofuels and of any other molecule of interest from algae biomass indeed depends on two factors: the efficient conversion of light energy into biomass and the optimization of metabolic pathways to maximize the fraction of the biomass consisting of the product of interest (Kruse, 2015). Both these factors needs to be improved through biotechnological approaches (Wobbe et al., 2015; Wobbe and Remacle, 2014; Gimpel et al., 2015). The improvements achieved in both fields will be later combined to manipulate several traits, to achieve the maximal productivity and finally exploit the potential yield for the organism under analysis.

### *Biotechnological optimization of the light energy to biomass conversion efficiencies.*

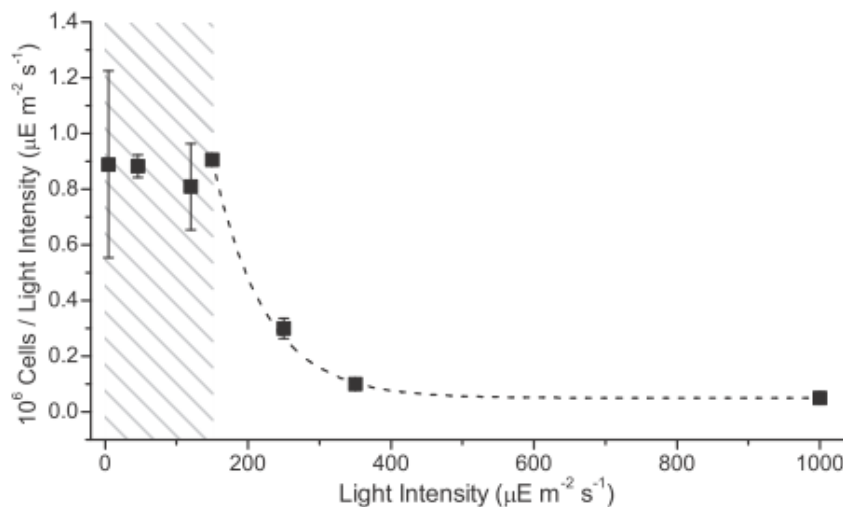
Microalgae are photosynthetic organisms, therefore the sustainability of their industrial application relies on how efficiently they exploit sunlight for growth. The light amount reaching the planet is very high, but just a fraction can be exploited by photosynthesis, the so-called photosynthetic active radiation (PAR - between 400 and 700 nm), which roughly accounts for the 46 % of the total sunlight energy (Wobbe and Remacle, 2014). Moreover, this final sunlight fraction is efficiently exploitable by the photosynthesis only if not exceeding an optimal intensity threshold, which differs from species to species. When light intensity exceeds such values, it has inhibitory effects on growth. For instance, in species belonging to the *Nannochloropsis* genus, light shows this negative effect on growth when its intensity exceeds  $150 \mu\text{mol photons} / (\text{m}^2 \text{ s})$  (Figure 3 – (Simionato et al., 2013a)).

In fact, despite light energy is fundamental for photosynthesis, when provided in excess it is harmful, leading to reactive oxygen species (ROS) formation. Photons constitutively induce the formation of singlet chlorophyll ( $^1\text{Chl}^*$ ) than can de-excite transferring energy to drive photochemistry.

However, when photons are provided in excess this phenomenon comes together with the formation of triplet chlorophyll ( $^3\text{Chl}^*$ ) which in turn can generate singlet oxygen ( $^1\text{O}_2$ ), that can oxidize pigments, proteins and lipids, leading to photo-inhibition and growth limitations. To prevent ROS formation, the excess of singlet chlorophyll can de-excite not through the triplet form but by dissipating heat, through the phenomenon of non-photochemical quenching (NPQ) activation.

Photosynthetic organisms are equipped with a set of light-harvesting complex (LHC) proteins, which serve to the biological role of capturing light and transfer it to photosynthesis reaction centers.

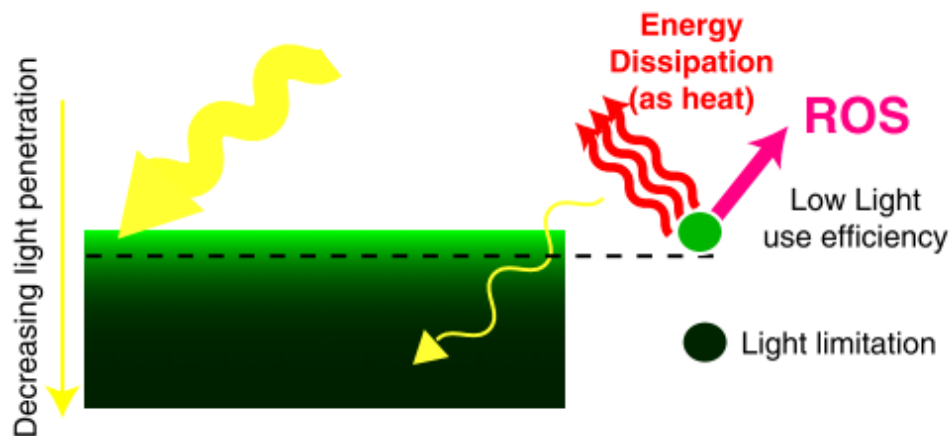
A peculiar class, widely known as LHCX (stress-related light-harvesting complex) modulate the response to the light excess, triggering the NPQ activation (Peers et al., 2009; Erickson et al., 2015).



**Figure 3. Light intensity impact on microalgae growth (Simionato et al., 2013a).** *Nannochloropsis* was cultivated in a flat-bed PBR in order to expose all the cells to the same irradiation, reducing as much as possible the cells' self-shading effect.  $\text{CO}_2$  and nutrients were provided in excess to highlight the influence of light regime on growth kinetics. Data were taken from (Sforza et al., 2012).

However, the majority of microalgae species are equipped with large antenna systems, including several LHC proteins binding hundreds of chlorophyll molecules per reaction center. These pigments maximize the light-harvesting efficiency as an evolutionary adaptation to a natural environment where solar radiation is often limiting for growth (Kirk, 1994) and where the competition with other

photosynthetic organisms is essential to thrive (Kirst et al., 2014). When microalgae are cultivated in high cell densities for industrial purposes, this peculiarity induces an inhomogeneous light distribution through the mass culture volume (Figure 4). The cells populating the external layers of the culture absorb all the radiation, saturating their photosynthetic ability and therefore reducing the efficiencies in energy conversion mechanisms. In these cells, up to 80 % of the absorbed energy is indeed wasted as heat through non-photochemical quenching (NPQ) activation (Peers et al., 2009; Melis, 2009), for preventing reactive oxygen species (ROS) formation.



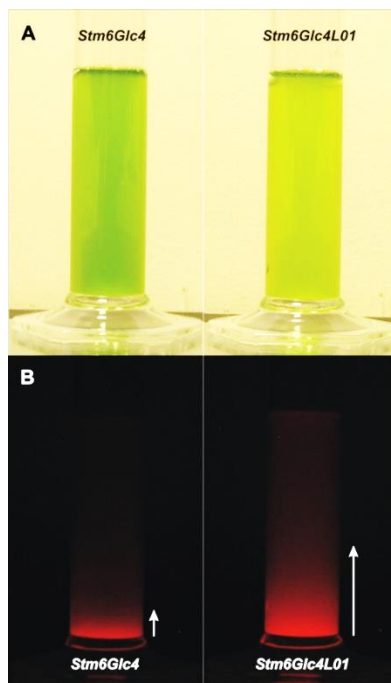
**Figure 4. Light distribution in a photobioreactor when applying wild type strains ((Simionato et al., 2013a).** The outer layers of the cultures are exposed to saturating light intensities, leading to energy dissipation mechanisms activation (Non-photochemical quenching - NPQ), which are however not enough to prevent reactive oxygen species (ROS) formation. The remaining inner culture layers are instead strongly light-limited. For these reasons, both cell populations show productivity constrains.

The absorption of the major fraction of light by the external layers of culture leaves the inner ones in the dark (Figure 4). Underneath layers are in fact strongly light limited and for this reason cells work below the point of compensation, where energy used for cells maintenance is higher than that directed to photosynthesis for growth. The constitutive qE activation (the fastest and reversible component of NPQ—(de Bianchi et al., 2010)) generally limits the amount of absorbed energy directed to photosynthesis. When light intensity is strongly reduced, like in the case of the culture inner layers, this phenomenon indeed further reduces the amount of light directed to photosynthesis, and cells growth rates are even more notably reduced for strong light availability constrains.

This configuration induces a strong impairment in the light energy to biomass conversion efficiencies of microalgae industrial intensive cultures.

In fact, the actual values for microalgae are far below (between the 2 and 3% - (Wobbe et al., 2015; Melis, 2009)) the theoretical estimations, expected to be around 10 % (Wobbe and Remacle, 2014), and this fact strongly limit any economically sustainable industrial microalgae application. Consequently, in the past years several attempts to enhance photosynthetic conversion efficiencies, when microalgae are applied to photobioreactors, were performed, by improving the efficiency of the enzymes (e.g. the carboxylation efficiency of the Rubisco - (Durão et al., 2015)) belonging to the major electron sink, the Calvin-Benson cycle, or the energy transfer mechanisms through the photosynthetic electron transport chain (Radakovits et al., 2010).

The reduction of the optical density of microalgae intensive cultures, through cell pigment content reduction, was postulated as viable biotechnological approach to increase the light penetration in a photobioreactor (Figure 5 - (Formighieri et al., 2012)).



**Figure 5. Improved light penetration in a culture of microalgae with a reduced pigment content (Oey et al., 2013).** The *Stm6Glc4* *C. reinhardtii* strain (mutant showing none alteration in photosynthesis but an improved H<sub>2</sub> production) is compared to the *Stm6Glc4L01* mutant, affected in LHCMB1, 2 and 3 antenna expression, by their simultaneous knock-down through RNAi techniques. In A. the cultures are illuminated with a white light while in B. with a blue light to highlight the improved light penetration (roughly 4 times) in the antenna truncated mutant strain.

So far several studies have demonstrated the possibility to reduce the pigment content of microalgae, on one hand directly affecting genes involved in the pigment biosynthetic pathways (Polle et al., 2000) and on the other one by striking genes encoding LHC proteins. In fact, the pigment biosynthesis in the chloroplast (e.g. the chlorophyll a) is strictly coordinated with the expression of both nucleus and chloroplast-encoded chlorophyll binding proteins (Formighieri et al., 2012). However, when the pigment reduction is reflected in antenna truncation, it showed some negative consequences when mutants are applied in light-saturating conditions (Formighieri et al., 2012), since affecting also the fundamental response to excess light. In fact, it was suggested that only LHC proteins with a main light harvesting function should be affected, while those involved in photo-protection should not be down-regulated. The development of more selective approaches thus allowed to obtain specific antenna truncation, for instance by striking those molecular factors involved in their co-translational or post-translational regulation, such as the NAB1 repressor (Beckmann et al., 2009; Wobbe et al., 2009) or some chloroplast signal recognition particles (CpSRPs) (Kirst et al., 2012a, 2012b; Kirst and Melis, 2014). An improved selectivity was also achieved by RNAi techniques application (Oey et al., 2013; Mitra et al., 2012).

Also the targeting of the NPQ mechanism is a viable alternative to improve the light energy conversion efficiencies (Wobbe et al., 2015; Wobbe and Remacle, 2014). As described above, the constitutive qE activation strongly limits the amount of light energy directed toward the photosynthesis, and this fact strongly limits the growth rates due to energy losses via thermal dissipation, mainly in limiting-light conditions. Therefore, when the NPQ activation is deleted, such as in the *npq4* mutant of *A. thaliana* (lacking PSBS), an improved growth is achieved in low light while a strong growth inhibition is reported for high light (Dall'Osto et al., 2005). However, in several microalgae species, the qE activation is not constitutive (such as in *C. reinhardtii*) and the NPQ deletions wouldn't enhance growth rate in low light but would just account for a higher sensitivity in high light conditions. A detailed understanding at the molecular level, of the response of the photosynthetic apparatus to different light conditions will be fundamental to clarify the regulation of NPQ activation in different species, in order to properly tune the biotechnological approaches to improve photosynthesis efficiency.

The above described photosynthesis biotechnological approaches need to be explored further in order to achieve a concrete improved light energy to biomass conversion efficiencies when microalgae are applied in intensive cultures. Their application to new promising species is seminal to achieve this purpose.



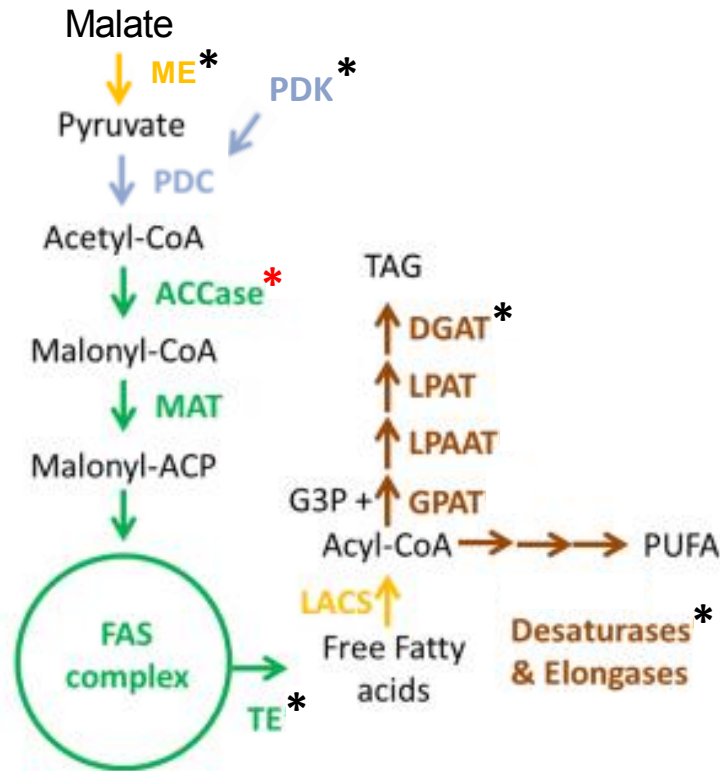
### *Biotechnological optimization of metabolic pathways.*

As above discussed, the increasing interest that the industry developed toward microalgae exploitation relies on their ability to naturally synthesize high-value molecules. However, this property often depends on the exposition to environmental stress conditions, from nutrients limitation (nitrogen and phosphorous) to high light exposure (Simionato et al., 2011; Yin-Hu et al., 2012; Chen et al., 2011; Alcántara et al., 2015; Chen et al., 2015; Menon et al., 2013).

For instance, several studies reported a great lipid accumulation when microalgae are exposed to nitrogen (N) limitation conditions (Simionato et al., 2013b; Jerez et al., 2015). However, the latter induces a cell growth impairment (Sun et al., 2013) since N shortage has a huge impact on the whole cell metabolism, shutting down proteins and nucleic acids biosynthesis. Therefore, the need to induce N starvation is not suitable with industrial operating conditions, since it would impair process continuity, leading to the need for huge culture volumes or repeated runs, which would increase the overall production costs (Rawat et al., 2013).

In this context, the need to find key master regulators, whose modification switches the metabolism toward the lipids accumulation, also during a maximal growth stage, is seminal to avoid the need for environmental stress conditions, which necessarily untie the two processes.

In the past years, for instance, different enzymatic steps of the lipid biosynthesis pathway of several diatoms, expected to be key regulators of the whole lipid metabolism, were targeted through genetic engineering in order to enhance / modify their lipid content / lipid composition profile (Figure 6), independently from stress conditions exposure. For instance, the malic enzyme (ME) catalyzes the decarboxylation of malate into pyruvate, producing also NADH used to feed the lipid biosynthesis with reducing power. ME overexpression in *P. tricornutum* induced a 2.5 fold increase in total lipids with respect to the parental strain (Xue et al., 2015). The pyruvate dehydrogenase complex (PDC) catalyzes the decarboxylation of pyruvate to acetyl-CoA and it is deactivated by a phosphorylation event catalyzed by the pyruvate dehydrogenase kinase (PDK). When the latter was knocked-down, the total neutral lipid content increased of up to 82% in *P. tricornutum* (Ma et al., 2014). In the latter organism, an endogenous fatty acids acyl-carrier protein thioesterase (TE) was also expressed leading to a 72% increase in total fatty acids (Gong et al., 2011). Finally, the overexpression of the endogenous DGAT2, coding for an Acyl-CoA : diacylglycerol acyltransferase (DGAT), in the *P. tricornutum* was shown to induce a 35% increase of neutral lipids accumulation (Niu et al., 2013).



**Figure 6. Simplified scheme for fatty acids (FA), triacylglycerols (TAG) and poly-unsaturated fatty acids (PUFA) synthesis in microalgae.** Enzymes are colored according to their localization. Blue – mitochondria; Green - plastid; Yellow - cytoplasm; Brown - endoplasmic reticulum. Asterisks highlight enzymes which have been targeted with genetic engineering, leading to a positive (black asterisks) or none (red asterisks) effect on lipid content or composition. ME, malic enzyme; PDC, pyruvate dehydrogenase complex; PDK, pyruvate carboxylase kinase; ACCase, acetyl-CoA carboxylase; MAT, malonyl-CoA/ACP transacylase; FAS, FA synthase; TE, fatty acyl-ACP thioesterase; LACS, long-chain acyl-CoA synthetase; GPAT, glycerol-3-phosphate acyltransferase; LPAAT, lyso-phosphatidic acid acyltransferase; LPAT, lyso-phosphatidylcholine acyltransferase; DGAT, diacylglycerol acyltransferase; CoA, coenzyme A; G3P, glycerate-3-phosphate. Adapted from (Gimpel et al., 2015).

Given the increasing market share covered by the poly-unsaturated fatty acids (PUFA) commercialization, increasing interest is gaining the application of these biotechnological approaches for the PUFA content improvement of microalgae biomass.

In this context, recently the long chain poly-unsaturated fatty acids (LC-PUFA) content was improved in *N. oceanica* by the overexpression of an endogenous  $\Delta 12$ -desaturase (Kaye et al., 2015). Moreover, the previously described DGAT2 overexpression in *P. tricornutum* was shown to induce also a greater omega-3 eicosapentaenoic acid (EPA) accumulation, of up to the 76 % than the parental strain (Niu et al., 2013).

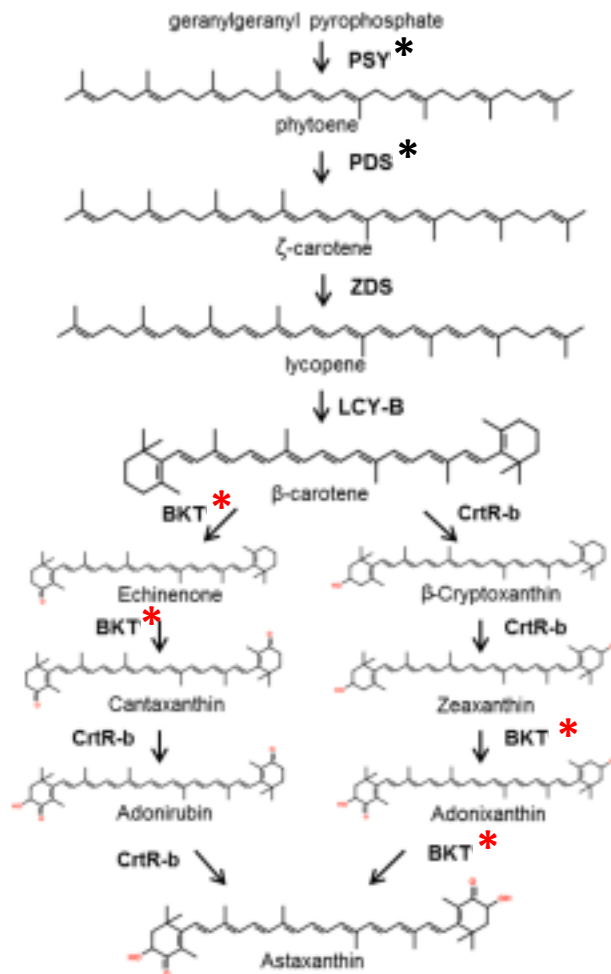
Moreover, the transcription factors (TF) are emerging as viable targets for the biotechnological optimization of microalgae metabolism since serving as master regulators for several molecular responses (Hu et al., 2014). Recently the overexpression of TF containing the basic helix-loop-helix (bHLH) motif in *N. salina* was shown to have a strong impact on the microalgae biomass accumulation as well as on their lipid production (Kang et al., 2015).

Despite the examples above presented show a positive impact on the lipid accumulation, when the same approaches are directed to different organisms, the results are contrasting.

In fact, when the overexpression of the same DGAT2 is induced in *C. reinhardtii*, none total lipid increase was reported (La Russa et al., 2012). Also when the same engineering approaches are directed toward other enzymes belonging to the lipid biosynthesis pathway, the results are different than expected, such as for the diatom *Cyclotella cryptica* in which the overexpression of the acetyl-CoA carboxylase (ACCase) enzyme didn't induce a total lipid accumulation (Figure 6 - (Dunahay et al., 1996)). Therefore, strong experimental efforts are still needed to reveal the complexity of the microalgae lipid metabolism with the final aim to explain the relation elapsing between the currently tested molecular targets in regulating the lipid metabolism.

Besides the lipid accumulation, microalgae are industrially relevant organisms also for their property of accumulating high amount of carotenoids. Biotechnological approaches are also being directed toward the carotenoids biosynthesis pathway, in order to improve their accumulation. However, since their biosynthesis pathway is still poorly characterized in microalgae, the biotechnological efforts devoted to enhance their accumulation are limited to very few attempts (Figure 7).

When the  $\beta$ -carotene ketolase genes from *H. pluvialis* (*bkt3* and *btk1*) and the endogenous gene of *C. reinhardtii* were expressed in the latter, none improvement in carotenoids accumulation was detected (León et al., 2007). When instead the genes coding for phytoene synthase (PSY) and phytoene desaturase (PDS) were expressed in *Haematococcus pluvialis*, *D. salina*, *C. zofingiensis* and *C. reinhardtii*, a significant increase in their carotenoids content was reported (Cordero et al., 2011; Couso et al.; Liu et al., 2013, 2014). The carotenoids biotechnological optimization has therefore not yielded consistent and definitive results in the elucidation of the rate-limiting reactions in the pathway. The targeted enzymes possibly do not represent the bottleneck steps and further investigations are therefore required to understand the complex regulation of their biosynthesis pathway.



**Figure 7. Simplified scheme for carotenoids biosynthesis in microalgae.** Asterisks highlight enzymes which have been targeted with genetic engineering, leading to a positive (black asterisks) or none (red asterisks) effect on carotenoids content or composition. PSY, phytoene synthase; PDS, phytoene desaturase; ZDS, ζ-carotene desaturase; LYC-B, lycopene β-cyclase; BKT, β-carotene ketolase; CrtR-b, β-carotene hydroxylase. Adapted from (Gimpel et al., 2015).

In fact, the regulation of the biosynthesis and catabolism of carotenoids and lipids, as well as of the metabolic processes involved in the modification of the length and saturation rate of the fatty acids, has not been widely investigated in algae yet. Filling this gap is of extreme importance since, for instance, both quantity and quality of diesel precursors depend on the regulation of this metabolism (Park et al., 2015; Medipally et al., 2015).

Several genome wide and random mutagenesis approaches have been applied in the past years to emerging industrial relevant species in order to fill the current incomplete knowledge of algae metabolism regulation (Martin et al., 2014; Carrier et al., 2014; Winck et al., 2013)(Lv et al.,

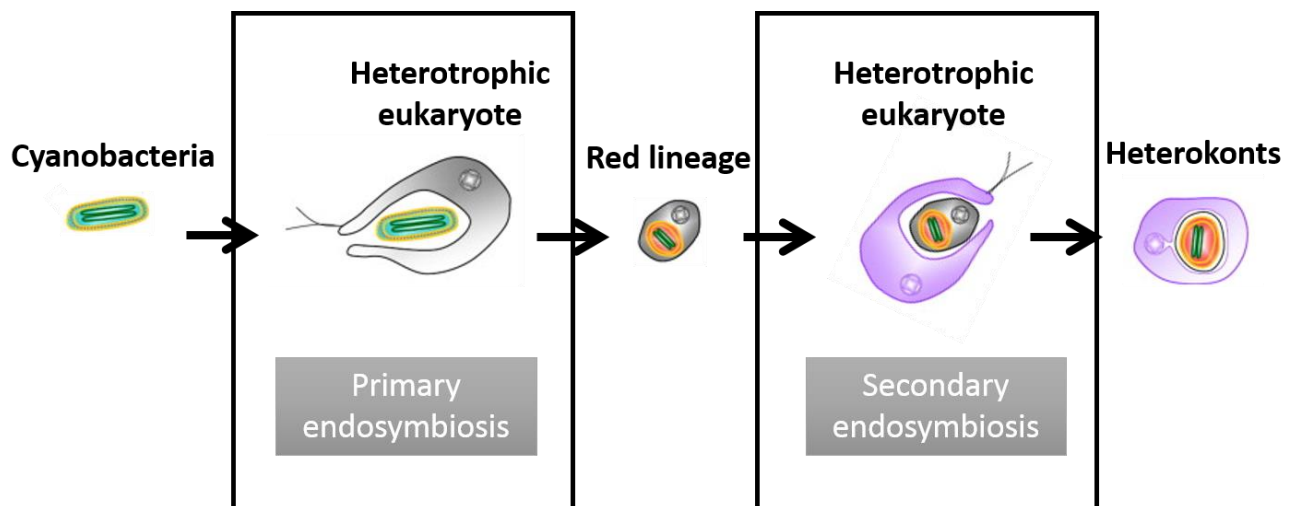
2013)(Bogen et al., 2013; Beacham et al., 2015). The main drawback is that microalgae show a strong metabolic divergence, which makes necessary the development of species-specific investigations and therefore huge experimental efforts are often required (Gimpel et al., 2015). However, the gained information will serve to direct the future genetic engineering efforts not toward specific enzymatic targets as widely done in the past and above discussed, but toward real key master regulators of the metabolism, with the final aim to redirect cell carbon and nitrogen fluxes toward the greater accumulation in the desired products.

In the near future, the investigation of the impact that the above described biotechnological optimization efforts have also on photosynthetic efficiency will be seminal to understand the relation between the planned metabolic fluxes redirection and microalgae growth. Only through this integrated approach it would be possible to develop microalgae as a biotechnological tool to concretely solve industrial problems.

*Nannochloropsis gaditana*, a promising target for industrial applications.

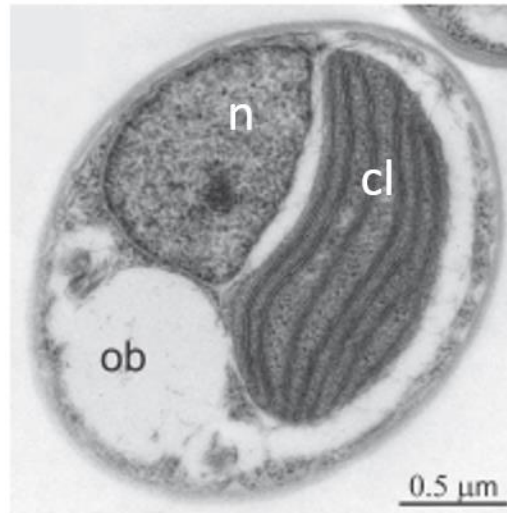
Microalgae are an extremely heterogeneous group of photosynthetic micro-organisms, which comprise species of eukaryotic algae spread among several phyla (Barra et al., 2014; Kiuru et al., 2014). Despite the estimated number of microalgae species in nature is between 40,000 and 70,000 (Appeltans et al., 2012), very few are those exploited for industrial purposes (6/7 - (Lebeau and Robert, 2003)), and even fewer those extensively characterized. Strong research efforts are therefore still needed to find the best candidates for industrial purposes and to domesticate them to optimally perform in intensive growth conditions. Although there isn't an ideal organism to fit all the industrial applications for the development of an algae-based economy, we focused our research efforts on the microalga *Nannochloropsis gaditana*.

This organism belongs to the kingdom of Chromista, which includes organisms whose chloroplast is surrounded by four membranes, as a consequence of a secondary endosymbiotic event, through which an ancestral unicellular red alga was engulfed by an eukaryotic host cell (Figure 8 - (Cavalier-Smith, 2004; Petroutsos et al., 2014)). Within this kingdom, *Nannochloropsis gaditana* belongs to the infra-kingdom of Heterokonta, which also includes diatoms and brown algae, and to the class of Eustigmatophyceae (Cavalier-Smith, 1995; Riisberg et al., 2009).



**Figure 8. Simplified schematic representation of Heterokonts evolution.** The two events of symbiosis which took place during evolution are here schematized in the black rectangles. Adapted from (Petroutsos et al., 2014).

*Nannochloropsis gaditana* also belongs to a genus which includes 6 species, one living in fresh water and the others in saline environments, all sharing common genomic oleaginous traits (Wang et al., 2014). It has a reduced cell size (2 - 3.5  $\mu\text{m}$ ) and the majority of the cell volume is occupied by the chloroplast (Figure 9 - (Simionato et al., 2013b)).

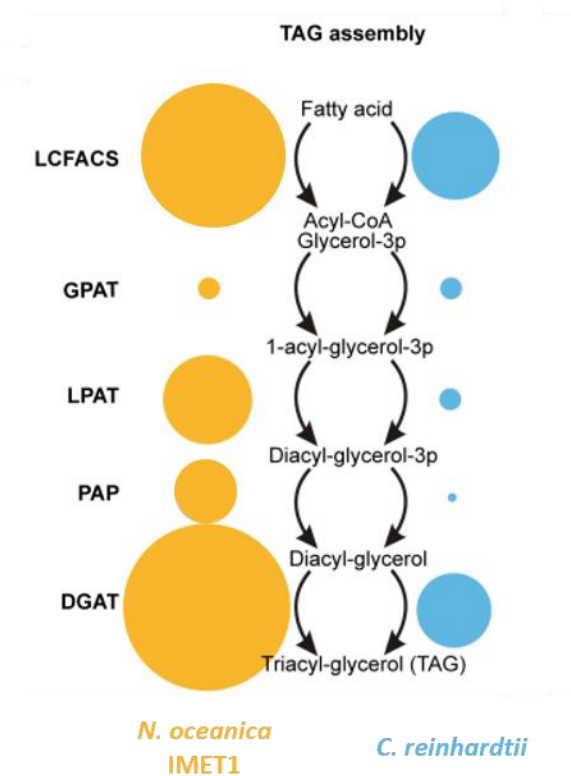


**Figure 9. Transmission microscopy image of a *Nannochloropsis gaditana* cell.** Organelles as indicated as, n, nucleus; cl, chloroplast; ob, oil body. The image was taken in N-replate conditions. Adapted from (Simionato et al., 2013b).

*Nannochloropsis gaditana* shows a special pigment composition, counting the presence of just chlorophyll a and lacking other accessory chlorophylls molecules, while the two major carotenoids detected in the cell are violaxanthin and vaucherixanthin esters (Basso et al., 2014). This feature makes it a valuable tool for studies concerning the evolution of the photosynthetic machinery, toward the identification of specific regulatory mechanisms that could be targeted to enhance photosynthetic efficiencies.

Moreover, in the past years, this species emerged also as a promising organism for the development of environmental-friendly industrial processes and as a model of study for biofuels and biocommodities production (e.g. animal feed). *Nannochloropsis gaditana* indeed shows a high biomass productivity in multiple light conditions (Simionato et al., 2011) and the ability to accumulate large amount of lipids (up to 38 % of its dry weight) and long chain poly-unsaturated fatty acids (LC-PUFA), at least in specific environmental conditions (Simionato et al., 2013b; Radakovits et al., 2012).

This organism shows indeed a prominent expansion in the gene copy number of specific reaction nodes belonging to the TAG biosynthesis pathway (Figure 10), despite showing one-fourth genome dimension with respect to *C. reinhardtii* (Corteggiani Carpinelli et al., 2014).



**Figure 10. Enrichment in the gene copy number for enzymes involved in the triacylglycerol (TAG) assembly pathway, in species belonging to the *Nannochloropsis* genus. *Nannochloropsis oceanica* IMET1 was taken as representative of the *Nannochloropsis* genus. The reaction nodes which show a gene copy number expansion are, Long-chain fatty acyl-CoA synthetase (LC-FACS, 11–12 vs. 7), lysophosphatidyl acyltransferase (LPAT, 7-8 vs. 1), phosphatidic acid phosphatase (PAP, 5 vs. 1) and diacylglycerol acyltransferase (DGAT, 11 vs. 6). Adapted from (Wang et al., 2014).**

The expansion in the gene dose in these selective enzymatic steps highlights the higher degree of carbon channeling toward TAG accumulation in these species, possibly uncovering a peculiar and more complex regulation of the lipid metabolism with respect to other organisms. Therefore, this feature makes it a valuable tool also for basic biological investigations of the molecular mechanisms that direct the whole metabolism toward an oleaginous trait.

However, strong experimental efforts still need to be devoted to improve our knowledge on its metabolism regulation. With this purpose, in the past years, huge steps forward were done to provide



to the scientific community the molecular information to speed up its biotechnological optimization. The genome-wide approaches led us to investigate the genome molecular organization of this species, and the transcription profile alterations in different growth conditions (Wang et al., 2014; Corteggiani Carpinelli et al., 2014; Radakovits et al., 2012; Vieler et al., 2012; Poliner et al., 2015), to focus the biotechnological efforts on specific metabolic pathways.

However, the molecular toolkit for its genetic manipulation still needs to be strongly improved, in order to fully exploit the information gained through these genome-wide approaches.

The work presented in the following PhD thesis will serve to improve the available molecular information *Nannochloropsis gaditana*. The former will be further used to optimize the molecular toolkit available for its efficient biotechnological optimization toward its cost-effective industrial application.

## References.

- Alcántara, C., Fernández, C., García-Encina, P.A., and Muñoz, R.** (2015). Mixotrophic metabolism of *Chlorella sorokiniana* and algal-bacterial consortia under extended dark-light periods and nutrient starvation. *Appl. Microbiol. Biotechnol.* **99**: 2393–404.
- Allison, E.H. and Bassett, H.R.** (2015). Climate change in the oceans: Human impacts and responses. *Science* (80-. ). **350**: 778–782.
- Appeltans, W. et al.** (2012). The magnitude of global marine species diversity. *Curr. Biol.* **22**: 2189–202.
- Barra, L., Chandrasekaran, R., Corato, F., and Brunet, C.** (2014). The Challenge of Ecophysiological Biodiversity for Biotechnological Applications of Marine Microalgae. *Mar. Drugs* **12**: 1641–1675.
- Basso, S., Simionato, D., Gerotto, C., Segalla, A., Giacometti, G.M., and Morosinotto, T.** (2014). Characterization of the photosynthetic apparatus of the Eustigmatophycean *Nannochloropsis gaditana*: evidence of convergent evolution in the supramolecular organization of photosystem I. *Biochim. Biophys. Acta* **1837**: 306–14.
- Beacham, T.A., Macia, V.M., Rooks, P., White, D.A., and Ali, S.T.** (2015). Altered lipid accumulation in *Nannochloropsis salina* CCAP849/3 following EMS and UV induced mutagenesis. *Biotechnol. reports (Amsterdam, Netherlands)* **7**: 87–94.
- Beckmann, J., Lehr, F., Finazzi, G., Hankamer, B., Posten, C., Wobbe, L., and Kruse, O.** (2009). Improvement of light to biomass conversion by de-regulation of light-harvesting protein translation in *Chlamydomonas reinhardtii*. *J. Biotechnol.* **142**: 70–7.
- de Bianchi, S., Ballottari, M., Dall’osto, L., and Bassi, R.** (2010). Regulation of plant light harvesting by thermal dissipation of excess energy. *Biochem. Soc. Trans.* **38**: 651–60.
- Böer, E., Steinborn, G., Kunze, G., and Gellissen, G.** (2007). Yeast expression platforms. *Appl. Microbiol. Biotechnol.* **77**: 513–23.
- Bogen, C., Al-Dilaimi, A., Albersmeier, A., Wichmann, J., Grundmann, M., Rupp, O., Lauersen, K.J., Blifernez-Klassen, O., Kalinowski, J., Goesmann, A., Mussgnug, J.H., and Kruse, O.** (2013). Reconstruction of the lipid metabolism for the microalga *Monoraphidium neglectum* from its genome sequence reveals characteristics suitable for biofuel production. *BMC Genomics* **14**: 926.
- Carrier, G., Garnier, M., Le Cunff, L., Bougaran, G., Probert, I., De Vargas, C., Corre, E., Cadoret, J.-P., and Saint-Jean, B.** (2014). Comparative transcriptome of wild type and selected strains of the

microalgae *Tisochrysis lutea* provides insights into the genetic basis, lipid metabolism and the life cycle. *PLoS One* **9**: e86889.

- Cavalier-Smith, T.** (2004). Only six kingdoms of life. *Proc. Biol. Sci.* **271**: 1251–62.
- Cavalier-Smith, T.** (1995). Zooflagellate phylogeny and classification. *Tsitologiia* **37**: 1010–1029.
- Chen, H., Hu, J., Qiao, Y., Chen, W., Rong, J., Zhang, Y., He, C., and Wang, Q.** (2015). Ca<sup>2+</sup>-regulated cyclic electron flow supplies ATP for nitrogen starvation-induced lipid biosynthesis in green alga. *Sci. Rep.* **5**: 15117.
- Chen, M., Tang, H., Ma, H., Holland, T.C., Ng, K.Y.S., and Salley, S.O.** (2011). Effect of nutrients on growth and lipid accumulation in the green alga *Dunaliella tertiolecta*. *Bioresour. Technol.* **102**: 1649–55.
- Cordero, B.F., Couso, I., León, R., Rodríguez, H., and Vargas, M.A.** (2011). Enhancement of carotenoids biosynthesis in *Chlamydomonas reinhardtii* by nuclear transformation using a phytoene synthase gene isolated from *Chlorella zofingiensis*. *Appl. Microbiol. Biotechnol.* **91**: 341–51.
- Corteggiani Carpinelli, E., Telatin, A., Vitulo, N., Forcato, C., D'Angelo, M., Schiavon, R., Vezzi, A., Giacometti, G.M., Morosinotto, T., and Valle, G.** (2014). Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion. *Mol. Plant* **7**: 323–35.
- Couso, I., Vila, M., Rodriguez, H., Vargas, M.A., and León, R.** Overexpression of an exogenous phytoene synthase gene in the unicellular alga *Chlamydomonas reinhardtii* leads to an increase in the content of carotenoids. *Biotechnol. Prog.* **27**: 54–60.
- Dall'Osto, L., Caffarri, S., and Bassi, R.** (2005). A mechanism of nonphotochemical energy dissipation, independent from PsbS, revealed by a conformational change in the antenna protein CP26. *Plant Cell* **17**: 1217–32.
- Dunahay, T.G., Jarvis, E.E., Dais, S.S., and Roessler, P.G.** (1996). Manipulation of microalgal lipid production using genetic engineering. *Appl. Biochem. Biotechnol.* **57-58**: 223–231.
- Durão, P., Aigner, H., Nagy, P., Mueller-Cajar, O., Hartl, F.U., and Hayer-Hartl, M.** (2015). Opposing effects of folding and assembly chaperones on evolvability of Rubisco. *Nat. Chem. Biol.* **11**: 148–55.
- Erickson, E., Wakao, S., and Niyogi, K.K.** (2015). Light stress and photoprotection in *Chlamydomonas reinhardtii*. *Plant J.* **82**: n/a–n/a.
- Fargione, J., Hill, J., Tilman, D., Polasky, S., and Hawthorne, P.** (2008). Land clearing and the biofuel

carbon debt. *Science* **319**: 1235–8.

- Formighieri, C., Franck, F., and Bassi, R.** (2012). Regulation of the pigment optical density of an algal cell: filling the gap between photosynthetic productivity in the laboratory and in mass culture. *J. Biotechnol.* **162**: 115–23.
- Gangl, D., Zedler, J.A.Z., Rajakumar, P.D., Martinez, E.M.R., Riseley, A., Włodarczyk, A., Purton, S., Sakuragi, Y., Howe, C.J., Jensen, P.E., and Robinson, C.** (2015). Biotechnological exploitation of microalgae. *J. Exp. Bot.*
- Gimpel, J.A., Henríquez, V., and Mayfield, S.P.** (2015). In *Metabolic Engineering of Eukaryotic Microalgae: Potential and Challenges Come with Great Diversity*. *Front. Microbiol.* **6**: 1376.
- Gimpel, J.A., Specht, E.A., Georgianna, D.R., and Mayfield, S.P.** (2013). Advances in microalgae engineering and synthetic biology applications for biofuel production. *Curr. Opin. Chem. Biol.* **17**: 489–95.
- Gong, Y., Guo, X., Wan, X., Liang, Z., and Jiang, M.** (2011). Characterization of a novel thioesterase (PtTE) from *Phaeodactylum tricornutum*. *J. Basic Microbiol.* **51**: 666–72.
- Harari, A., Harats, D., Marko, D., Cohen, H., Barshack, I., Kamari, Y., Gonen, A., Gerber, Y., Ben-Amotz, A., and Shaish, A.** (2008). A 9-cis beta-carotene-enriched diet inhibits atherogenesis and fatty liver formation in LDL receptor knockout mice. *J. Nutr.* **138**: 1923–30.
- Hu, J., Wang, D., Li, J., Jing, G., Ning, K., and Xu, J.** (2014). Genome-wide identification of transcription factors and transcription-factor binding sites in oleaginous microalgae *Nannochloropsis*. *Sci. Rep.* **4**: 5454.
- Jerez, C.G., Malapascua, J.R., Sergejevová, M., Figueroa, F.L., and Masojídek, J.** (2015). Effect of Nutrient Starvation under High Irradiance on Lipid and Starch Accumulation in *Chlorella fusca* (Chlorophyta). *Mar. Biotechnol.* (NY).
- Kagan, M.L., Sullivan, D.W., Gad, S.C., and Ballou, C.M.** Safety assessment of EPA-rich polar lipid oil produced from the microalgae *Nannochloropsis oculata*. *Int. J. Toxicol.* **33**: 459–74.
- Kang, N.K., Jeon, S., Kwon, S., Koh, H.G., Shin, S.-E., Lee, B., Choi, G.-G., Yang, J.-W., Jeong, B.-R., and Chang, Y.K.** (2015). Effects of overexpression of a bHLH transcription factor on biomass and lipid production in *Nannochloropsis salina*. *Biotechnol. Biofuels* **8**: 200.
- Kaye, Y., Grundman, O., Leu, S., Zarka, A., Zorin, B., Didi-Cohen, S., Khozin-Goldberg, I., and Boussiba, S.** (2015). Metabolic engineering toward enhanced LC-PUFA biosynthesis in *Nannochloropsis oceanica*: Overexpression of endogenous  $\Delta 12$  desaturase driven by stress-inducible promoter leads to enhanced deposition of polyunsaturated fatty acids in TAG. *Algal*

Res. **11**: 387–398.

- Kazamia, E. and Smith, A.G.** (2014). Assessing the environmental sustainability of biofuels. *Trends Plant Sci.* **19**: 615–8.
- Kim, D.-Y., Vijayan, D., Praveenkumar, R., Han, J.-I., Lee, K., Park, J.-Y., Chang, W.-S., Lee, J.-S., and Oh, Y.-K.** (2015). Cell-wall disruption and lipid/astaxanthin extraction from microalgae: *Chlorella* and *Haematococcus*. *Bioresour. Technol.* **199**: 300–10.
- Kirk, J.** (1994). Light and photosynthesis in aquatic ecosystems.
- Kirst, H., Formighieri, C., and Melis, A.** (2014). Maximizing photosynthetic efficiency and culture productivity in cyanobacteria upon minimizing the phycobilisome light-harvesting antenna size. *Biochim. Biophys. Acta* **1837**: 1653–64.
- Kirst, H., García-Cerdán, J.G., Zurbriggen, A., and Melis, A.** (2012a). Assembly of the light-harvesting chlorophyll antenna in the green alga *Chlamydomonas reinhardtii* requires expression of the TLA2-CpFTSY gene. *Plant Physiol.* **158**: 930–45.
- Kirst, H., Garcia-Cerdan, J.G., Zurbriggen, A., Ruehle, T., and Melis, A.** (2012b). Truncated photosystem chlorophyll antenna size in the green microalga *Chlamydomonas reinhardtii* upon deletion of the TLA3-CpSRP43 gene. *Plant Physiol.* **160**: 2251–60.
- Kirst, H. and Melis, A.** (2014). The chloroplast signal recognition particle (CpSRP) pathway as a tool to minimize chlorophyll antenna size and maximize photosynthetic productivity. *Biotechnol. Adv.* **32**: 66–72.
- Kiuru, P., D’Auria, M.V., Muller, C.D., Tammela, P., Vuorela, H., and Yli-Kauhaluoma, J.** (2014). Exploring marine resources for bioactive compounds. *Planta Med.* **80**: 1234–46.
- Kruse, O.** (2015). Prospects and challenges for the development of algal biotechnology. *J. Biotechnol.* **215**: 1.
- Kwak, J.H., Baek, S.H., Woo, Y., Han, J.K., Kim, B.G., Kim, O.Y., and Lee, J.H.** (2012). Beneficial immunostimulatory effect of short-term *Chlorella* supplementation: enhancement of natural killer cell activity and early inflammatory response (randomized, double-blinded, placebo-controlled trial). *Nutr. J.* **11**: 53.
- Lauersen, K.J. et al.** (2015a). Investigating the dynamics of recombinant protein secretion from a microalgal host. *J. Biotechnol.* **215**: 62–71.
- Lauersen, K.J., Berger, H., Mussnug, J.H., and Kruse, O.** (2013). Efficient recombinant protein production and secretion from nuclear transgenes in *Chlamydomonas reinhardtii*. *J. Biotechnol.* **167**: 101–10.

- Lauersen, K.J., Kruse, O., and Mussnug, J.H.** (2015b). Targeted expression of nuclear transgenes in *Chlamydomonas reinhardtii* with a versatile, modular vector toolkit. *Appl. Microbiol. Biotechnol.* **99**: 3491–503.
- Lebeau, T. and Robert, J.-M.** (2003). Diatom cultivation and biotechnologically relevant products. Part II: current and putative products. *Appl. Microbiol. Biotechnol.* **60**: 624–32.
- León, R., Couso, I., and Fernández, E.** (2007). Metabolic engineering of ketocarotenoids biosynthesis in the unicellular microalga *Chlamydomonas reinhardtii*. *J. Biotechnol.* **130**: 143–52.
- Levy, B.S. and Patz, J.A.** Climate Change, Human Rights, and Social Justice. *Ann. Glob. Heal.* **81**: 310–22.
- Liu, J., Gerken, H., Huang, J., and Chen, F.** (2013). Engineering of an endogenous phytoene desaturase gene as a dominant selectable marker for *Chlamydomonas reinhardtii* transformation and enhanced biosynthesis of carotenoids. *Process Biochem.* **48**: 788–795.
- Liu, J., Sun, Z., Gerken, H., Huang, J., Jiang, Y., and Chen, F.** (2014). Genetic engineering of the green alga *Chlorella zofingiensis*: a modified norflurazon-resistant phytoene desaturase gene as a dominant selectable marker. *Appl. Microbiol. Biotechnol.* **98**: 5069–79.
- Luo, X., Su, P., and Zhang, W.** (2015). Advances in Microalgae-Derived Phytosterols for Functional Food and Pharmaceutical Applications. *Mar. Drugs* **13**: 4231–54.
- Lv, H., Qu, G., Qi, X., Lu, L., Tian, C., and Ma, Y.** (2013). Transcriptome analysis of *Chlamydomonas reinhardtii* during the process of lipid accumulation. *Genomics* **101**: 229–37.
- Ma, Y.-H., Wang, X., Niu, Y.-F., Yang, Z.-K., Zhang, M.-H., Wang, Z.-M., Yang, W.-D., Liu, J.-S., and Li, H.-Y.** (2014). Antisense knockdown of pyruvate dehydrogenase kinase promotes the neutral lipid accumulation in the diatom *Phaeodactylum tricornutum*. *Microb. Cell Fact.* **13**: 100.
- Markou, G. and Nerantzis, E.** (2013). Microalgae for high-value compounds and biofuels production: a review with focus on cultivation under stress conditions. *Biotechnol. Adv.* **31**: 1532–42.
- Martin, G.J.O., Hill, D.R.A., Olmstead, I.L.D., Bergamin, A., Shears, M.J., Dias, D.A., Kentish, S.E., Scales, P.J., Botté, C.Y., and Callahan, D.L.** (2014). Lipid profile remodeling in response to nitrogen deprivation in the microalgae *Chlorella* sp. (Trebouxiophyceae) and *Nannochloropsis* sp. (Eustigmatophyceae). *PLoS One* **9**: e103389.
- Medipally, S.R., Yusoff, F.M., Banerjee, S., and Shariff, M.** (2015). Microalgae as sustainable renewable energy feedstock for biofuel production. *Biomed Res. Int.* **2015**: 519513.
- Melis, A.** (2009). Solar energy conversion efficiencies in photosynthesis: Minimizing the chlorophyll antennae to maximize efficiency. *Plant Sci.* **177**: 272–280.

- Menon, K.R., Balan, R., and Suraishkumar, G.K.** (2013). Stress induced lipid production in *Chlorella vulgaris*: relationship with specific intracellular reactive species levels. *Biotechnol. Bioeng.* **110**: 1627–36.
- Mitra, M., Kirst, H., Dewez, D., and Melis, A.** (2012). Modulation of the light-harvesting chlorophyll antenna size in *Chlamydomonas reinhardtii* by TLA1 gene over-expression and RNA interference. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **367**: 3430–43.
- Moss, R.H. et al.** (2010). The next generation of scenarios for climate change research and assessment. *Nature* **463**: 747–56.
- Niu, Y.-F., Zhang, M.-H., Li, D.-W., Yang, W.-D., Liu, J.-S., Bai, W.-B., and Li, H.-Y.** (2013). Improvement of neutral lipid and polyunsaturated fatty acid biosynthesis by overexpressing a type 2 diacylglycerol acyltransferase in marine diatom *Phaeodactylum tricornutum*. *Mar. Drugs* **11**: 4558–69.
- Oey, M., Ross, I.L., Stephens, E., Steinbeck, J., Wolf, J., Radzun, K.A., Kügler, J., Ringsmuth, A.K., Kruse, O., and Hankamer, B.** (2013). RNAi knock-down of LHCBM1, 2 and 3 increases photosynthetic H<sub>2</sub> production efficiency of the green alga *Chlamydomonas reinhardtii*. *PLoS One* **8**: e61375.
- Orfield, N.D., Keoleian, G.A., and Love, N.G.** (2014). A GIS based national assessment of algal bio-oil production potential through flue gas and wastewater co-utilization. *Biomass and Bioenergy* **63**: 76–85.
- Park, J.-Y., Park, M.S., Lee, Y.-C., and Yang, J.-W.** (2015). Advances in direct transesterification of algal oils from wet biomass. *Bioresour. Technol.* **184**: 267–75.
- Peers, G., Truong, T.B., Ostendorf, E., Busch, A., Elrad, D., Grossman, A.R., Hippler, M., and Niyogi, K.K.** (2009). An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* **462**: 518–21.
- Peters, J. and Stoger, E.** (2011). Transgenic crops for the production of recombinant vaccines and anti-microbial antibodies. *Hum. Vaccin.* **7**: 367–74.
- Petroutsos, D. et al.** (2014). Evolution of galactoglycerolipid biosynthetic pathways--from cyanobacteria to primary plastids and from primary to secondary plastids. *Prog. Lipid Res.* **54**: 68–85.
- Poliner, E., Panchy, N., Newton, L., Wu, G., Lapinsky, A., Bullard, B., Zienkiewicz, A., Benning, C., Shiu, S.-H., and Farré, E.M.** (2015). Transcriptional coordination of physiological responses in *Nannochloropsis oceanica* CCMP1779 under light/dark cycles. *Plant J.* **83**: 1097–113.

- Polle, J.E.W., Benemann, J.R., Tanaka, A., and Melis, A.** (2000). Photosynthetic apparatus organization and function in the wild type and a chlorophyll b-less mutant of *Chlamydomonas reinhardtii*. Dependence on carbon source. *Planta* **211**: 335–344.
- Radakovits, R., Jinkerson, R.E., Darzins, A., and Posewitz, M.C.** (2010). Genetic engineering of algae for enhanced biofuel production. *Eukaryot. Cell* **9**: 486–501.
- Radakovits, R., Jinkerson, R.E., Fuerstenberg, S.I., Tae, H., Settlage, R.E., Boore, J.L., and Posewitz, M.C.** (2012). Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. *Nat. Commun.* **3**: 686.
- Ramos Tercero, E.A., Sforza, E., Morandini, M., and Bertucco, A.** (2014). Cultivation of *Chlorella protothecoides* with urban wastewater in continuous photobioreactor: biomass productivity and nutrient removal. *Appl. Biochem. Biotechnol.* **172**: 1470–85.
- Raposo, M.F. de J. and de Morais, A.M.M.B.** (2015). Microalgae for the prevention of cardiovascular disease and stroke. *Life Sci.* **125**: 32–41.
- Raposo, M.F. de J., de Morais, A.M.M.B., and de Morais, R.M.S.C.** (2015). Carotenoids from Marine Microalgae: A Valuable Natural Source for the Prevention of Chronic Diseases. *Mar. Drugs* **13**: 5128–55.
- Rasala, B.A., Lee, P.A., Shen, Z., Briggs, S.P., Mendez, M., and Mayfield, S.P.** (2012). Robust expression and secretion of Xylanase1 in *Chlamydomonas reinhardtii* by fusion to a selection gene and processing with the FMDV 2A peptide. *PLoS One* **7**: e43349.
- Rasala, B.A. and Mayfield, S.P.** (2015). Photosynthetic biomanufacturing in green algae; production of recombinant proteins for industrial, nutritional, and medical uses. *Photosynth. Res.* **123**: 227–39.
- Rawat, I., Ranjith Kumar, R., Mutanda, T., and Bux, F.** (2013). Biodiesel from microalgae: A critical evaluation from laboratory to large scale production. *Appl. Energy* **103**: 444–467.
- Riisberg, I., Orr, R.J.S., Kluge, R., Shalchian-Tabrizi, K., Bowers, H.A., Patil, V., Edvardsen, B., and Jakobsen, K.S.** (2009). Seven gene phylogeny of heterokonts. *Protist* **160**: 191–204.
- La Russa, M., Bogen, C., Uhmeyer, A., Doebbe, A., Filippone, E., Kruse, O., and Mussgnug, J.H.** (2012). Functional analysis of three type-2 DGAT homologue genes for triacylglycerol production in the green microalga *Chlamydomonas reinhardtii*. *J. Biotechnol.* **162**: 13–20.
- Ryckebosch, E., Bruneel, C., Termote-Verhalle, R., Goiris, K., Muylaert, K., and Foubert, I.** (2014). Nutritional evaluation of microalgae oils rich in omega-3 long chain polyunsaturated fatty acids as an alternative for fish oil. *Food Chem.* **160**: 393–400.



- Salih, F.M.** (2011). Microalgae Tolerance to High Concentrations of Carbon Dioxide: A Review. *J. Environ. Prot. (Irvine, Calif)*. **02**: 648–654.
- Scaife, M.A., Nguyen, G.T.D.T., Rico, J., Lambert, D., Helliwell, K.E., and Smith, A.G.** (2015). Establishing *Chlamydomonas reinhardtii* as an industrial biotechnology host. *Plant J.* **82**: 532–46.
- Schlesinger, W.H. and Bernhardt, E.S.** (2013). *Biogeochemistry* (Elsevier).
- Scranton, M.A., Ostrand, J.T., Fields, F.J., and Mayfield, S.P.** (2015). *Chlamydomonas* as a model for biofuels and bio-products production. *Plant J.* **82**: 523–31.
- Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D., and Yu, T.-H.** (2008). Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science* **319**: 1238–40.
- Sforza, E., Simionato, D., Giacometti, G.M., Bertucco, A., and Morosinotto, T.** (2012). Adjusted light and dark cycles can optimize photosynthetic efficiency in algae growing in photobioreactors. *PLoS One* **7**: e38975.
- Simionato, D., Basso, S., Giacometti, G.M., and Morosinotto, T.** (2013a). Optimization of light use efficiency for biofuels production in algae. *Biophys. Chem.*
- Simionato, D., Block, M.A., La Rocca, N., Jouhet, J., Maréchal, E., Finazzi, G., and Morosinotto, T.** (2013b). The response of *Nannochloropsis gaditana* to nitrogen starvation includes de novo biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. *Eukaryot. Cell* **12**: 665–76.
- Simionato, D., Sforza, E., Corteggiani Carpinelli, E., Bertucco, A., Giacometti, G.M., and Morosinotto, T.** (2011). Acclimation of *Nannochloropsis gaditana* to different illumination regimes: effects on lipids accumulation. *Bioresour. Technol.* **102**: 6026–32.
- Slegers, P.M., Leduc, S., Wijffels, R.H., van Straten, G., and van Boxtel, A.J.B.** (2015). Logistic analysis of algae cultivation. *Bioresour. Technol.* **179**: 314–22.
- Spadiut, O., Capone, S., Krainer, F., Glieder, A., and Herwig, C.** (2014). Microbials for the production of monoclonal antibodies and antibody fragments. *Trends Biotechnol.* **32**: 54–60.
- Specht, E., Miyake-Stoner, S., and Mayfield, S.** (2010). Micro-algae come of age as a platform for recombinant protein production. *Biotechnol. Lett.* **32**: 1373–83.
- Sun, D., Zhu, J., Fang, L., Zhang, X., Chow, Y., and Liu, J.** (2013). De novo transcriptome profiling uncovers a drastic downregulation of photosynthesis upon nitrogen deprivation in the nonmodel green alga *Botryosphaerella sudeticus*. *BMC Genomics* **14**: 715.
- Tsai, H.-P., Chuang, L.-T., and Chen, C.-N.N.** (2016). Production of long chain omega-3 fatty acids and

carotenoids in tropical areas by a new heat-tolerant microalga *Tetraselmis* sp. DS3. *Food Chem.* **192**: 682–90.

**Ullrich, K.K., Hiss, M., and Rensing, S.A.** (2015). Means to optimize protein expression in transgenic plants. *Curr. Opin. Biotechnol.* **32**: 61–7.

**Vanthoor-Koopmans, M., Wijffels, R.H., Barbosa, M.J., and Eppink, M.H.M.** (2013). Biorefinery of microalgae for food and fuel. *Bioresour. Technol.* **135**: 142–9.

**Vieler, A. et al.** (2012). Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779. *PLoS Genet.* **8**: e1003064.

**Walsh, B.J., Ryzak, F., Palazzo, A., Kraxner, F., Herrero, M., Schenk, P.M., Ciaï, P., Janssens, I.A., Peñuelas, J., Niederl-Schmidinger, A., and Obersteiner, M.** (2015). New feed sources key to ambitious climate targets. *Carbon Balance Manag.* **10**: 26.

**Wang, D. et al.** (2014). *Nannochloropsis* Genomes Reveal Evolution of Microalgal Oleaginous Traits. *PLoS Genet.* **10**.

**Winck, F.V., Páez Melo, D.O., and González Barrios, A.F.** (2013). Carbon acquisition and accumulation in microalgae *Chlamydomonas*: Insights from “omics” approaches. *J. Proteomics* **94**: 207–18.

**Wobbe, L., Bassi, R., and Kruse, O.** (2015). Multi-Level Light Capture Control in Plants and Green Algae. *Trends Plant Sci.*

**Wobbe, L., Blifernéz, O., Schwarz, C., Mussnug, J.H., Nickelsen, J., and Kruse, O.** (2009). Cysteine modification of a specific repressor protein controls the translational status of nucleus-encoded LHClI mRNAs in *Chlamydomonas*. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 13290–5.

**Wobbe, L. and Remacle, C.** (2014). Improving the sunlight-to-biomass conversion efficiency in microalgal biofactories. *J. Biotechnol.*

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

**Yaakob, Z., Ali, E., Zainal, A., Mohamad, M., and Takriff, M.S.** (2014). An overview: biomolecules from microalgae for animal feed and aquaculture. *J. Biol. Res. (Thessalonikē, Greece)* **21**: 6.

**Yin-Hu, W., Yin, Y., Xin, L., Hong-Ying, H., and Zhen-Feng, S.** (2012). Biomass production of a *Scenedesmus* sp. under phosphorous-starvation cultivation condition. *Bioresour. Technol.* **112**: 193–8.

**Young, R.E.B. and Purton, S.** (2015). Codon reassignment to facilitate genetic engineering and biocontainment in the chloroplast of *Chlamydomonas reinhardtii*. *Plant Biotechnol. J.*

**Zhang, L. and Wang, H.** (2015). Multiple Mechanisms of Anti-Cancer Effects Exerted by Astaxanthin.  
Mar. Drugs **13**: 4310–30.

# CHAPTER II

## **Generation of Random Mutants to Improve Light-Use Efficiency of *Nannochloropsis gaditana* Cultures for Biofuel Production**

### **Author names and affiliations**

Giorgio Perin<sup>1</sup>, Alessandra Bellan<sup>1</sup>, Anna Segalla<sup>1</sup>, Andrea Meneghesso<sup>1</sup>, Alessandro Alboresi<sup>1</sup> and Tomas Morosinotto<sup>1</sup>

<sup>1</sup>PAR-Lab\_Padua Algae Research Laboratory, Department of Biology, University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy

**THIS CHAPTER WAS PUBLISHED IN BIOTECHNOLOGY FOR BIOFUELS**

Perin G. et al. *Biotechnol Biofuels* (2015) 8:161 - Doi 10.1186/S13068-015-0337-5






RESEARCH

Open Access



# Generation of random mutants to improve light-use efficiency of *Nannochloropsis gaditana* cultures for biofuel production

Giorgio Perin<sup>1</sup>, Alessandra Bellan<sup>1,2</sup>, Anna Segalla<sup>1</sup>, Andrea Meneghesso<sup>1</sup>, Alessandro Alboresi<sup>1</sup> and Tomas Morosinotto<sup>1\*</sup> 

## Abstract

**Background:** The productivity of an algal culture depends on how efficiently it converts sunlight into biomass and lipids. Wild-type algae in their natural environment evolved to compete for light energy and maximize individual cell growth; however, in a photobioreactor, global productivity should be maximized. Improving light use efficiency is one of the primary aims of algae biotechnological research, and genetic engineering can play a major role in attaining this goal.

**Results:** In this work, we generated a collection of *Nannochloropsis gaditana* mutant strains and screened them for alterations in the photosynthetic apparatus. The selected mutant strains exhibited diverse phenotypes, some of which are potentially beneficial under the specific artificial conditions of a photobioreactor. Particular attention was given to strains showing reduced cellular pigment contents, and further characterization revealed that some of the selected strains exhibited improved photosynthetic activity; in at least one case, this trait corresponded to improved biomass productivity in lab-scale cultures.

**Conclusions:** This work demonstrates that genetic modification of *N. gaditana* has the potential to generate strains with improved biomass productivity when cultivated under the artificial conditions of a photobioreactor.

**Keywords:** Photosynthesis, *Nannochloropsis*, Mutants, EMS, Insertional mutagenesis, Biodiesel, Algae

## Background

Biodiesel production using microalgae biomass represents one interesting alternative for replacing fossil liquid fuels and reducing emission of greenhouse gasses into the atmosphere. Several thousand different algae species exist in nature, and thus far, few of them have been identified as being particularly suitable for these types of applications [1]. Species of the *Nannochloropsis* genus are included in the list of biodiesel production candidates

due to their ability to accumulate large amounts of lipids, especially during nutrient starvation [2, 3].

Algae are photosynthetic organisms and thus depend on sunlight for energy to support their metabolism [4]. Solar light is an abundant resource, but it is also distributed on a very large surface, making the average amount of energy available per unit area relatively low ( $1084 \times 10^{16}$  J/km<sup>2</sup> [5]). Consequently, algae growing in outdoor ponds and photobioreactors (PBR) are commonly limited by light availability, and their efficiency in converting light into biomass substantially influences overall productivity [6, 7]; this limitation is particularly true for photoautotrophs, such as the species belonging to the *Nannochloropsis* genus [8]. To make algae-based products competitive on the market, it is, therefore,

\*Correspondence: tomas.morosinotto@unipd.it

<sup>1</sup> Dipartimento di Biologia, Università di Padova, Via U. Bassi 58/B, 35121 Padua, Italy

Full list of author information is available at the end of the article

imperative that biomass production be maximized by optimizing the light use efficiency of these organisms in industrial-scale cultivation systems [6, 7].

One of the major problems associated with growing algae in any large-scale PBR (or pond) is that cultures have high optical densities, resulting in strongly inhomogeneous light distributions [9, 10]. Consequently, most of the available light is absorbed by superficial cells, which easily absorb energy exceeding their photochemical capacity. This surplus excitation leads to oxidative damage and photoinhibition, which can be avoided in part by dissipating a fraction of the absorbed energy as heat through a photoprotective mechanism known as Non-Photochemical Quenching (NPQ) [11]. NPQ effectively protects cells from photoinhibition, but it can drive the dissipation of up to 80 % of the absorbed energy as heat, strongly decreasing the light use efficiency of cells exposed to excess irradiation [12–14]. While the external cells absorb most of the available energy and use it with low efficiency, cells underneath are strongly light-limited, reducing the overall accumulation of biomass. Due to this inhomogeneous light distribution, algae grown in large-scale PBRs are significantly less productive than algae grown under laboratory conditions [15–17].

The large antenna system of the algae photosynthetic apparatus binds hundreds of chlorophyll molecules (Chl) per reaction center. These pigments maximize light-harvesting efficiency as an evolutionary adaptation to a natural environment where solar radiation is often limiting for growth, and competition with other organisms for light is essential [5, 18]. In contrast, in a large-scale PBR, such large antenna systems limit light penetration into algal cultures and the capacity for competition exists at the expense of overall productivity. In this artificial environment, an “ideal” organism should harvest only the amount of light that it can use with high efficiency, without removing useful energy that other cells could use; such a domesticated strain would have reduced fitness in a natural context but improved productivity in the artificial environment of a PBR [16, 19].

Following on this hypothesis, genetically engineered strains have been generated in past years with altered composition and regulation of the photosynthetic apparatus to reduce their competitive capacity, for example with reduced cell pigment content [6, 16, 20]. Reduction of culture light harvesting efficiency is indeed conceivable and it has been suggested that only approximately 50/350 and 90/300 of the chlorophyll molecules found in Photosystems II (PSII) and PSI (PSI), respectively, in *Chlamydomonas reinhardtii* are strictly necessary, while the rest are bound to light-harvesting complexes (LHC)

and are, in principle, dispensable [19, 21]. The mutations responsible for Chl content reduction should bear two positive effects: first, a cell will harvest less light when exposed to strong irradiation, reducing the need for the activation of energy dissipation mechanisms, and second, light will be better distributed in the culture volume, making more energy available to support the growth of biomass in cells in the internal light-limited layers [16].

Several recent reports have demonstrated that it is possible to isolate algae mutant strains with reduced Chl content and/or reduced antenna size, defined as TLA (truncated light-harvesting) mutant strains, especially in the model organism *C. reinhardtii* [15, 22–25]. These modifications were achieved by altering genes that are involved in the chlorophyll biosynthetic pathway [26] and genes that encode LHC proteins [27], as well as factors involved in their co- or post-translational regulation [28] or in their import into the chloroplast [18, 29]. Some of these TLA mutant strains have been shown increased maximum photosynthetic rates and a greater photosynthesis light saturation [28]; however, strains with reduced antenna systems might also exhibit enhanced photosensitivity under high light conditions because the pigment-binding complexes function not only in harvesting light but also in photoprotection and NPQ activation [11, 30, 31].

Despite promising results obtained in lab-scale experiments, TLA mutant strains produced contrasting results with regard to the improvement of biomass productivity when tested in a larger scale PBR [25, 32, 33], likely because the mutations can carry negative consequences, such as the aforementioned increased sensitivity to light stress. To effectively improve productivity, it is, therefore, necessary to find the optimal compromise between the benefits of reducing light harvesting and the disadvantages associated with the potential reduction in photoprotection [34].

Here, we isolated and characterized mutant strains with photosynthetic apparatus alterations that were generated by chemical and insertional random mutagenesis in *Nannochloropsis gaditana*, an interesting model organism for biofuel production and large-scale cultivation [35]. Such a collection of strains will facilitate the search for mutant strains with reduced competitive capacity that could potentially provide positive advantages in the artificial environment of a PBR. We also characterized those strains with reduced cellular Chl contents in greater detail, identifying strain E2 as having a reduced PSII antenna size, improved photosynthetic activity and improved productivity in lab-scale fed-batch cultures, representing a promising starting point in the search for improved productivity outdoors.

## Results and discussion

### Isolation of a collection of *Nannochloropsis* mutant strains with alterations in the photosynthetic apparatus

A *N. gaditana* culture in the late exponential growth phase was treated with ethyl methane sulfonate (EMS) to induce random genomic mutations. EMS is an alkylating agent that causes a high frequency of nucleotide substitutions and has been largely employed to induce mutations in various organisms. In recent years, EMS mutagenesis has also been employed in algae, e.g., to induce the over-production of metabolites such as astaxanthin, carotenoids and eicosapentaenoic acid [36]. Unlike other approaches, chemical mutagenesis avoids the issues connected with the outdoor cultivation of genetically modified organisms (GMOs) and generates strains that are readily applicable to large-scale systems; however, chemical mutagenesis readily induces multiple mutations (typically point mutations), making the identification of the gene(s) responsible for the phenotype complex and time-consuming. For this reason, we also generated a second insertional mutant collection in parallel by transforming *N. gaditana* cells by electroporation and selecting transformants on plates containing zeocin [35]. Colonies surviving the antibiotic treatment contained at least one copy of the resistance gene integrated in the genome, and this was verified using PCR (Additional file 1: Figure S1). In principle, the insertion of exogenous DNA occurs randomly throughout the genome and can thus alter the expression of a wild-type (WT) gene to generate a mutant. The probability of gene disruption is expected to be very high in *N. gaditana* due to the high density of genes in its nuclear genome (10,646 genes in 27.5 Mbp, approximately 387 genes/Mbp) compared with that of *C. reinhardtii* (17,741 genes in 111.1 Mbp, approximately 160 genes/Mbp [37]).

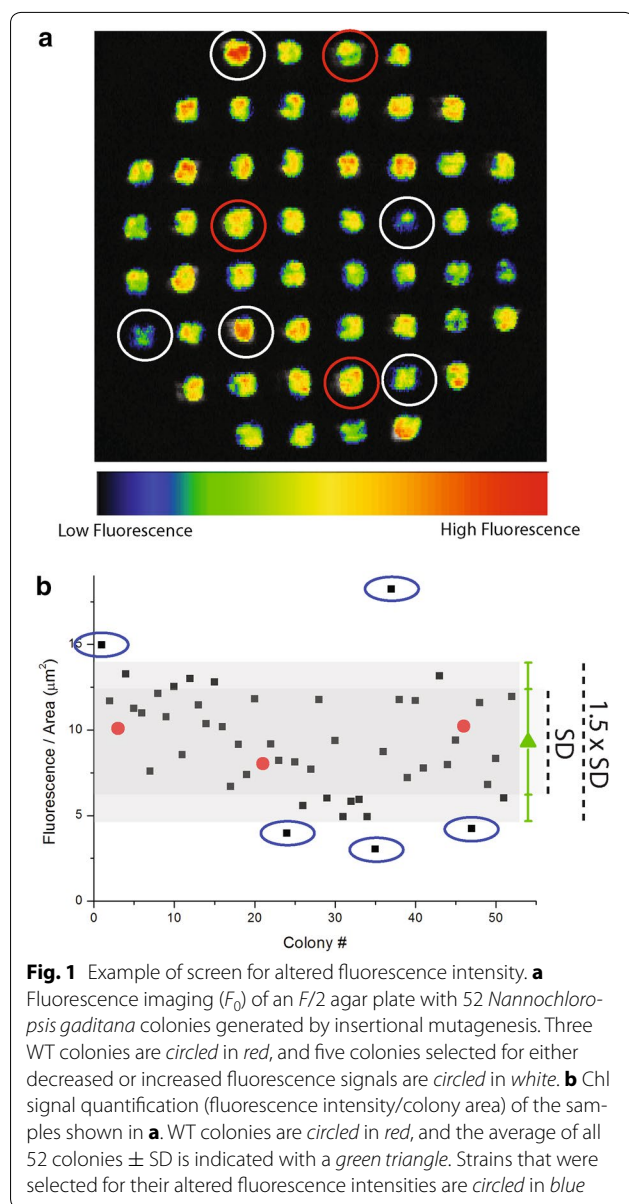
Following EMS and insertional mutagenesis, independent colonies ( $7$  and  $5 \times 10^3$ , respectively) were screened by measuring in vivo Chl fluorescence, which is a powerful approach for assessing the presence of changes to the composition and/or regulation of the photosynthetic apparatus [38]. This widely used technique is based on the principle that once light is absorbed by a chlorophyll molecule within an algal cell, excitation energy can take three alternative pathways: it can drive photochemical reactions, be dissipated as heat, or be re-emitted as fluorescence. These three processes compete for the same excited states and although only a relatively small amount of energy is re-emitted as fluorescence (1–2 % of the absorbed light), this fluorescence can be measured and used to gather indirect information about the other pathways [38].

To gather information about possible changes in photosynthesis, we measured the Chl fluorescence of algal

mutant collections on agar plates using a video-imaging apparatus (Fig. 1). Several parameters were monitored during screening to facilitate the selection of various phenotypes (Additional file 1: Table S1). The first parameter was the fluorescence of dark-adapted cells ( $F_0$ ); when normalized to the colony area, this metric provides an estimate of the cellular Chl content, allowing the identification of mutant strains with potentially reduced pigment contents. As shown by the false color image in Fig. 1a, a few colonies had fluorescence signals that were either more or less intense than the others. Quantification of the fluorescence signal per area (Fig. 1b) facilitated the selection of potentially interesting mutant strains with altered pigment contents. Because most of the cells in the plate are not likely to have an altered photosynthetic apparatus, the average value for all of the colonies on the plate was used as a reference. This method was preferable over direct comparison with WT colonies present on the plate as it allowed a more reliable evaluation of variability and this choice was validated by the fact that WT colonies fell close to the average value of the screened population, as shown in Fig. 1b. Fluorescence intensity signals have substantial variability, and we observed that the standard deviations for each plate were between 20 and 35 % of the average (33 % for the plate shown in Fig. 1a). We used the average value  $\pm 1.5$  SD as a threshold and selected mutant strains falling outside of that interval for further investigation (e.g., the five mutant strains highlighted in Fig. 1a). An incontrovertible problem with this screening approach is the selection of several false positive strains during the first round analysis for simple statistical reasons. Using more stringent selection criteria would have reduced their number but would have also yielded only those strains with the strongest phenotypes, which are undesirable as extensive alterations in the photosynthetic apparatus would most likely negatively affect productivity.

Other parameters determined from in vivo Chl fluorescence measurements were used to identify mutant strains with additional alterations in the photosynthetic apparatus (Additional file 1: Table S1). The maximum quantum yield of PSII ( $\Phi_{PSII}$ , or  $F_v/F_m$ ) estimates the fraction of absorbed light used for photochemistry [38] and, therefore, quantifies the maximum photochemical capacity of dark-adapted cells. When deviating from WT values, this metric suggests the existence of alterations in the cell's photochemical reactions. The same parameter can also be monitored in cells that have been exposed to an actinic light ( $\Phi_{PSII}'$ , or  $F'_m - F/F'_m$ ); in this case, its value decreases in part due to the oxidation of PS reaction centers, making them unavailable for photochemistry (the so-called "closed" state). Light-sensitive strains will show a greater decrease in PSII quantum yield, while





strains with higher potential photochemical capacity will have greater  $\Phi_{PSII}'$  values.

The same fluorescence measurements after cell illumination can also be used to evaluate the mutant's capacity to regulate photosynthesis through NPQ. In natural environments, the loss of heat dissipation mechanisms would be detrimental [11]; however, these strains are potentially interesting for use in PBRs as they do not waste energy under light-limited conditions [19].

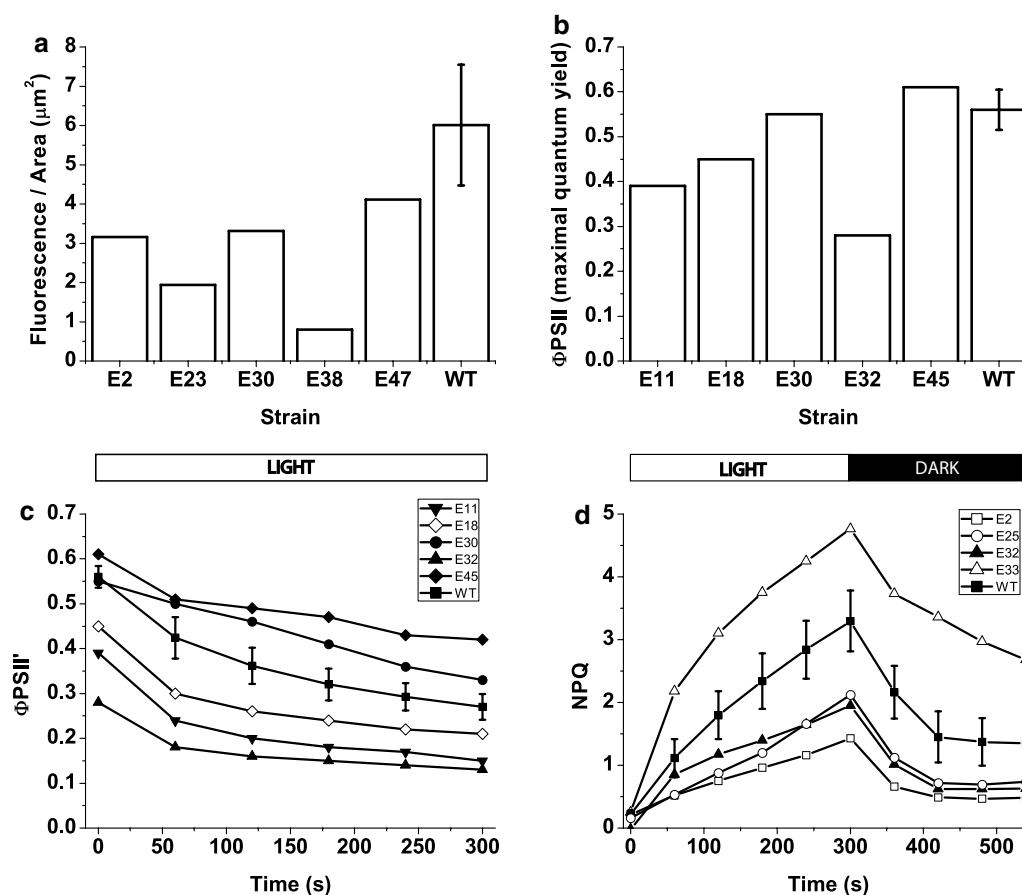
As mentioned above, all of the strains selected for one or more parameters (approximately 10 % of the initial number) were subjected to a second round of screening to identify those showing reproducible alterations

to reduce the number of false positives. At this stage, the relative number of mutant strains was theoretically enriched and WT colonies grown on the same plate were used as a reference instead of the average value of all colonies.

In several cases, differences in colony size or density affected fluorescence measurements, leading to the use of a third round of screening. All strains were homogeneously spotted onto solid media (Additional file 1: Figure S2) at equal cellular concentrations ( $OD_{750} = 0.2$ ). Although homogenizing the inoculum is time-consuming for a large number of colonies, it is extremely helpful for increasing screening sensitivity, as colony heterogeneity often prevents the detection of mutant strains with weak phenotypes [24]. Following this third round of screening, the strains exhibiting confirmed phenotypes were retained for further characterization; examples of the identified phenotypes are shown in Fig. 2. In total, 48 were selected from the collection of insertional mutant strains (named I1-48) and 49 were selected from the EMS mutant strains (named E1-49) (Additional file 1: Figure S3). This screen identified mutant strains with altered fluorescence density, PSII quantum yield in dark-adapted and illuminated cells and NPQ kinetics (Fig. 2a–d, respectively). Selection according to the first parameter led to the isolation of mutant strains exhibiting potential alterations in their cellular Chl contents compared with the WT strain that could be advantageous in an artificial growth system through improved light penetration in the entire culture volume. The second parameter was used to isolate strains with either improved or unaltered PSII maximum quantum yield compared with WT. The third parameter identifies strains that can saturate photosynthesis at higher irradiances than WT. The NPQ kinetics were instead useful for identifying those strains that were unable to constitutively activate excess energy dissipation, which may be valuable when grown in photobioreactor regions that are heavily light-limited. It is worth underlining that at this stage all mutants showing reproducible phenotypes were retained even if their phenotypes, such in the case of strains exhibiting larger NPQ than the WT, are likely not advantageous for industrial applications. In some cases (Fig. 2), the same strain (e.g., clone E32) was altered in more than one parameter; this feature was expected, as the parameters are not entirely independent.

#### Isolation of *Nannochloropsis* strains with reduced pigment contents

The work described thus far produced a collection of mutant strains carrying a broad spectrum of alterations in the photosynthetic apparatus that was useful for isolating strains with a specific phenotype of interest. For



**Fig. 2** Examples of selected mutant strains with altered fluorescence parameters. Cells were diluted to the same cell density ( $\text{OD}_{750} = 0.2$ ) prior to measurements, and examples of EMS-induced mutant strains exhibiting significant variation from WT are shown in each panel. **a**  $F_0/\text{Area}$ ; **b** PSII quantum yield in dark-adapted colonies; **c** PSII quantum yield in colonies illuminated with  $450 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ; **d** NPQ induction and recovery (light was switched off after 300 s). WT  $n = 9$

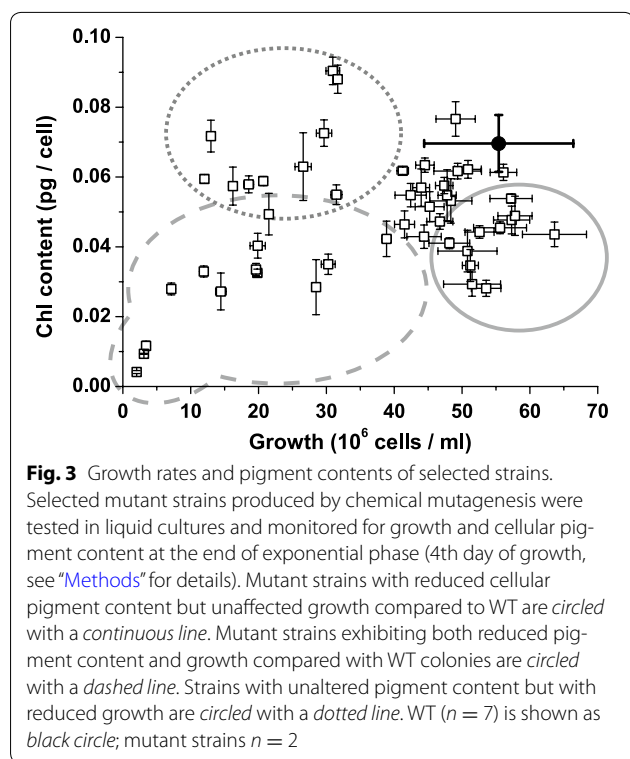
all mutant strains, the mutation's effect on cell growth is key piece of information. *Nannochloropsis gaditana* is a strict photoautotroph and we can assume that some photosynthesis mutant strains may also have growth defects that would prohibit their use in large-scale cultivation applications.

For this reason, we tested all selected strains for growth in liquid media; it should be noted that the cells were grown in flasks containing  $F/2$  media, under conditions in which the duplication rate is largely limited by  $\text{CO}_2$  and nutrient availability rather than light use efficiency [39]. Consequently, these tests did not facilitate the identification of mutant strains with improved biomass productivity but rather the elimination of those with growth defects.

Cellular Chl contents were monitored during the late exponential phase (4th day of growth) to highlight alterations in the composition of the photosynthetic apparatus. The overall results obtained for strains produced

by chemical mutagenesis are reported in Fig. 3. There was a group of mutant strains with altered growth but not altered cellular Chl content; for these strains, the reduced growth rate caused a reduction in cell concentration and a consequent decrease in fluorescence signal, leading to their selection during the screening stage even without a change in the cellular pigment content (e.g., Additional file 1: Figure S2, E1 and E4). This result is reasonable considering that *N. gaditana* is an obligate phototroph and its growth depends entirely on photosynthesis [8]; therefore, any growth impairment, even if independent of photosynthesis, will result in reduced fluorescence.

Another group of mutant strains exhibited a reduced growth rates that were accompanied by a reduced cellular pigment contents. These mutant strains, whose phenotypes are particularly severe in some cases (e.g., Additional file 1: Figure S2, I1 and E3), likely carry mutations that not only impair the photosynthetic apparatus but



also drastically affect growth and are, therefore, of limited interest for biotechnological applications.

A third group of strains exhibited growth rates and pigment contents that were indistinguishable from WT. There were a few residual false positives within this group; however, this group also contained mutant strains with altered fluorescence parameters (e.g., NPQ and/or Fv/Fm) without significant effects on either growth rates or pigment contents.

Finally, the fourth group includes mutant strains with reduced cellular pigment contents but not growth defects (Fig. 3; Additional file 1: Figure S2, E2, I2 and I3). These mutant strains are interesting for applied research as they could potentially drive improved light distribution and productivity in large-scale cultures.

Similar analysis of the insertional mutant strains produced comparable results even though the growth conditions used in the two experiments were slightly different (Additional file 1: Figure S4). One difference between the two strategies was that for the insertional mutant strains, strains with increased fluorescence levels were retained, leading to the isolation of strains with increased Chl contents. Another difference between the two experiments is that no insertional mutant strains had dramatically impaired growth as observed for the chemical mutant set (Fig. 3). One possible explanation for this observation is that in the case of chemical mutagenesis, mutant strains

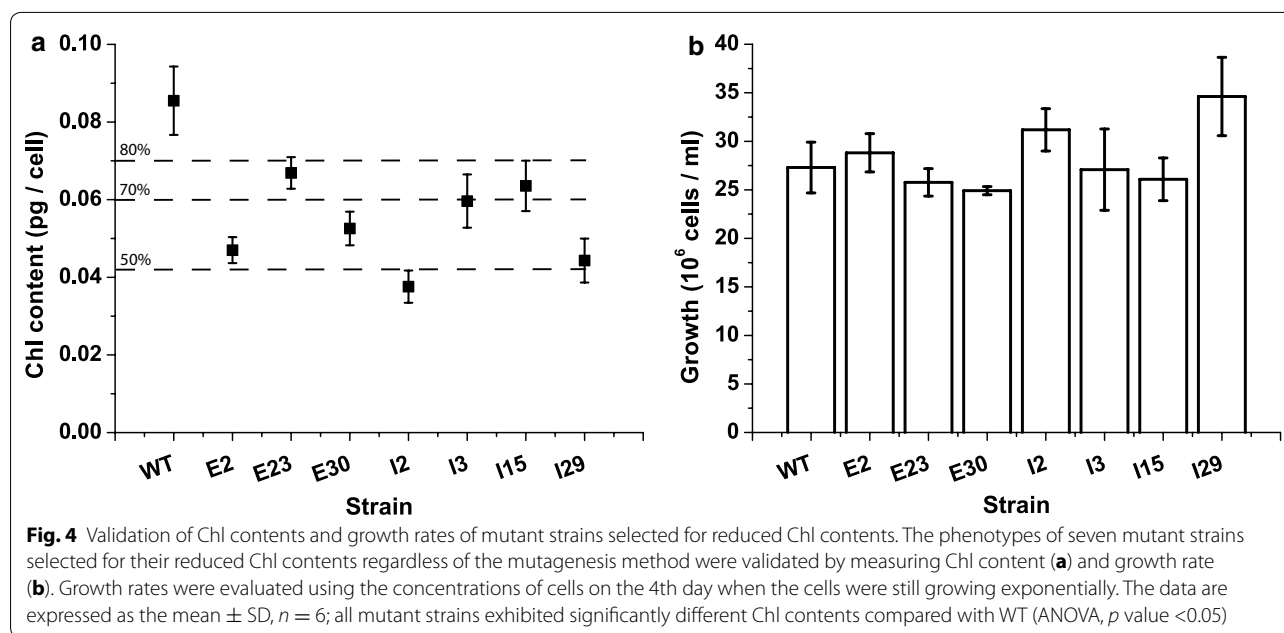
were chosen for screening over a longer interval (up to 8 weeks); therefore, even slowly growing mutant strains could form detectable colonies and be screened. In contrast, the insertional mutant colonies were all collected between 4 and 5 weeks to standardize the fluorescence imaging screening, leading to the elimination of the slowest growing colonies. Furthermore, the EMS-induced mutant strains likely carry many mutations and their combined effects can cause dramatic growth phenotypes.

#### Characterization of mutants with reduced pigment contents

The collection of mutant strains described thus far includes strains with various photosynthetic properties from which smaller groups of mutant strains with the desired phenotypes can be selected. We focused on selecting mutant strains with reduced cellular Chl contents that may improve light distribution in mass cultures [19]. To this end, 7 strains from both screens were identified as being the most promising due to their unchanged growth rates and reduced pigment contents; these mutant strains were then analyzed in greater detail. Repetition of the growth curve analysis confirmed that these strains have between 45 and 80 % of the WT cellular Chl content (Fig. 4a), with no change in growth rate under the tested conditions (Fig. 4b). As previously stated, the cultures were grown in flasks with CO<sub>2</sub> and nutrient-limiting conditions, potentially minimizing the growth differences between WT and mutant strains; regardless, this step was useful in confirming the reduced cellular Chl contents of the selected mutant strains.

First, we used western blotting to test whether two highly expressed pigment-binding proteins present in the *N. gaditana* photosynthetic apparatus were responsible for the selected phenotype. We used antibodies recognizing LHCf1 and LHCX1, the major antenna proteins of *N. gaditana*, which are expected to bind a significant fraction of the Chl molecules within the cell [40, 41]. Western blotting analysis revealed that five of the seven mutant strains exhibited consistently reduced LHCf1 levels compared with WT (E2, E23, E30, I2 and I29; Fig. 5a). In some cases (I2 and I29), this reduction was also accompanied by a reduction in a major PSII core subunit, D2, suggesting that PSII was reduced overall in these cells. For two other strains, E23 and E30, D2 levels were instead close to WT levels, suggesting a specific decrease in the LHCf1 antenna. For strain E2, D2 levels appeared to be greater than in WT, emphasizing the reduction in LHCf1 levels. In contrast, the two other strains selected for their reduced Chl contents (I3 and I15) exhibited increased LHCf1 levels.

The levels of another major antenna protein, LHCX1, were also reduced in five of the seven strains (E2, E23,



I15, I2 and I29). For strains E2 and I15, D2 levels were maintained or even increased compared with WT, suggesting specific reductions in LHCX1 subunits (Fig. 5b).

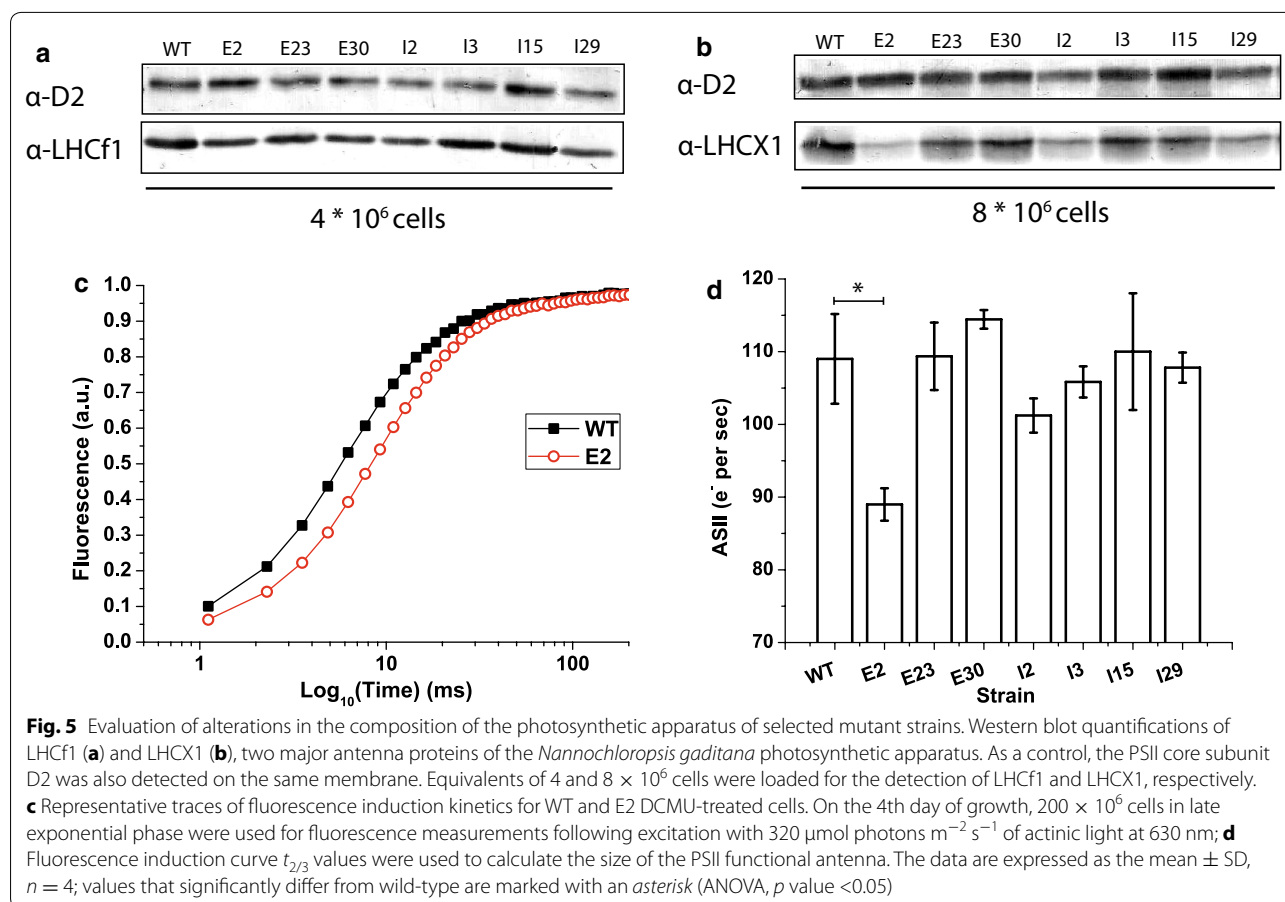
Together, these results clearly suggest that reductions in the pigment contents of various strains are caused by different molecular alterations to the photosynthetic apparatus. Among the identified strains, I2 and I29, exhibited a clear general decrease in the levels of all three tested proteins, suggesting a global reduction in the accumulation of all of the photosynthetic complexes. Two other strains, E2, and less obviously, E30, exhibited reduced LHCf1 and LHCX1 levels while the reaction center contents were more stable, suggesting specific effects on the antenna proteins. In contrast, strains I15 and I3 exhibited more specific effects on LHCX1 while the other two proteins were present at levels that were similar to or even greater than in WT.

Complementary information on the light-harvesting efficiency of the photosynthetic apparatus can be obtained using fluorescence induction kinetics to estimate PSII functional antenna size *in vivo*. In cells treated with 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU), electron transport from PSII is inhibited, and PSII reaction centers are saturated upon illumination, resulting in an increase in the emitted fluorescence over time (Fig. 5c). The time required to completely saturate all of the reaction centers depends on the number of pigments that are harvesting light and transferring energy; consequently, fluorescence kinetics provide an estimate of the number of antenna complexes associated with each PS [2, 24]. In the case of strain E2, the fluorescence rise was

clearly slower than for WT (Fig. 5c), suggesting a specific reduction in antenna proteins in this strain and confirming the above western blotting results. In contrast, the functional antenna sizes of the other mutant strains, e.g., E23, E30, I2, I3, I15 and I29, are very similar to WT, suggesting that the pigment content reductions in these strains are not due to specific effects on the antenna complexes, or at least on those associated with PSII (Fig. 5d). This result is also consistent with our biochemical data, with the exception of strain E30, for which it is possible that other antennas compensate for the decrease in LHCf1 levels.

#### Photosynthetic efficiency and productivity of strains with reduced pigment contents

Photosynthesis functionality in all strains was verified using a PAM fluorimeter to measure Chl fluorescence kinetics (see “Methods” for details). First, we analyzed PSII maximum quantum yield [38]; this parameter was not reduced for strains with reduced cellular Chl contents compared to the WT strain, suggesting that their mutations were not deleterious to photosynthesis. Four strains (E2, I2, I15 and I29) even exhibited increased PSII maximum quantum yield compared with the WT strain (Fig. 6a). Second, we analyzed the photosynthetic electron transport rate (ETR), which can be estimated by monitoring the PSII quantum yield of cells treated with increasingly intense light [38]. For WT *Nannochloropsis*, ETR progressively increased with increased illumination up to  $540 \mu\text{mol photons m}^{-2} \text{s}^{-1}$  (Fig. 6b, c). At this point, photosynthesis reached saturation and further increases



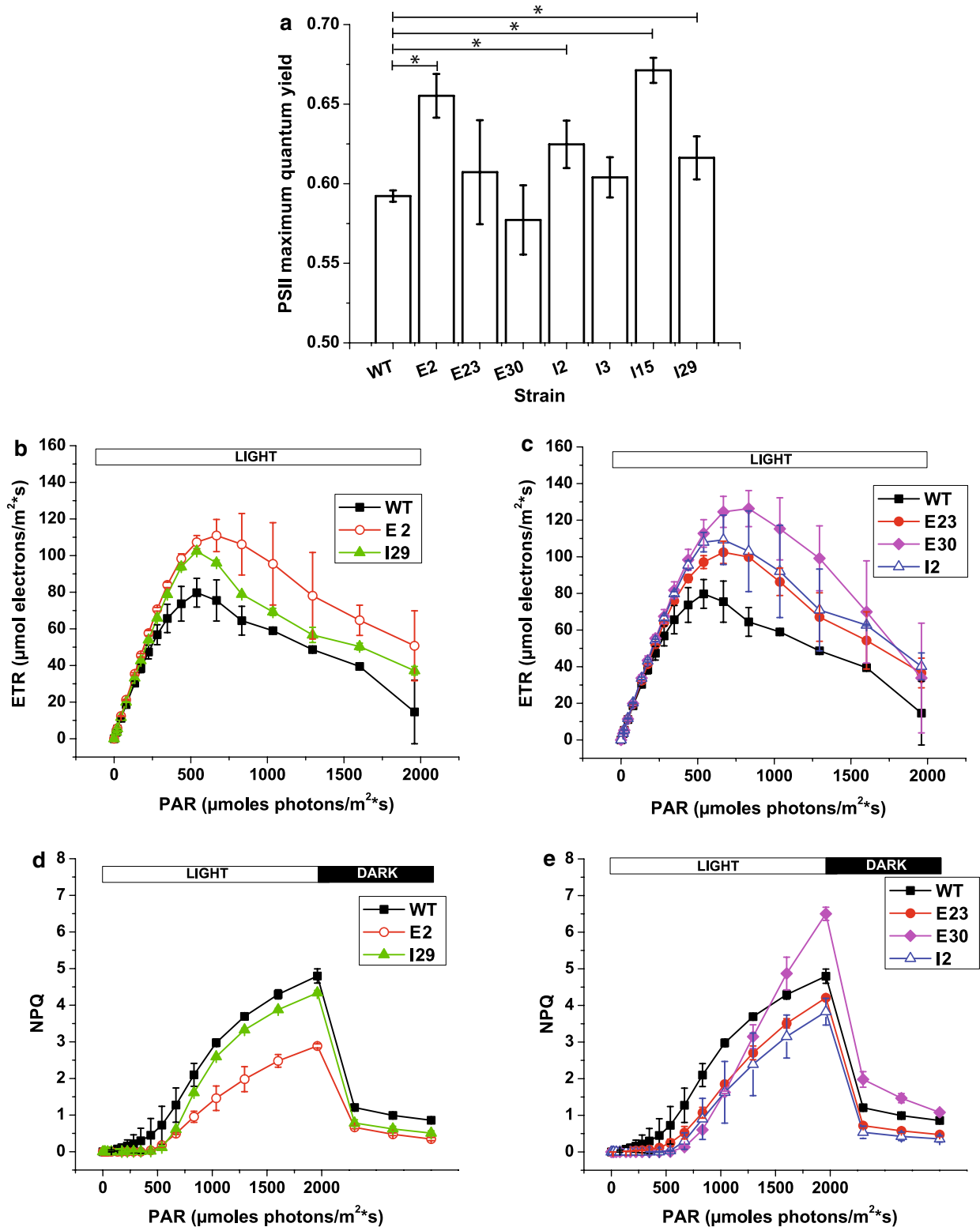
in light intensity did not induce an increase in ETR; in contrast, ETR began to decline (Fig. 6b, c). Most mutant strains behaved like the WT strain, while some showed an increase in maximal ETR (E2, E30, E23, I2 and I29), and two also reached saturation at higher light intensities (E2 and E30), which is a desired phenotype that could lead to increased productivity (Fig. 6b, c).

Another important feature of photosynthetic organisms is their ability to dissipate light energy as heat, a property that strongly influences algal light use efficiency [6]. This ability was quantified using an NPQ parameter, and the data revealed that most of the mutant strains did not significantly differ from WT, with the exception of E23, I2 and I29, which exhibited a small but reproducible reduction in NPQ activation. Strain E2 exhibited the strongest reduction in NPQ activation kinetics (Fig. 6d) enhancing its potential utility in large-scale cultivation systems as it could also use light much more efficiently in the inner PBR layers that are often light-limited. It is worth emphasizing that this reduction in NPQ agrees with our western blotting results (Fig. 5b) showing a reduction in LHCX1 protein levels in E2 compared with the WT strain. This analysis suggests a role for LHCX

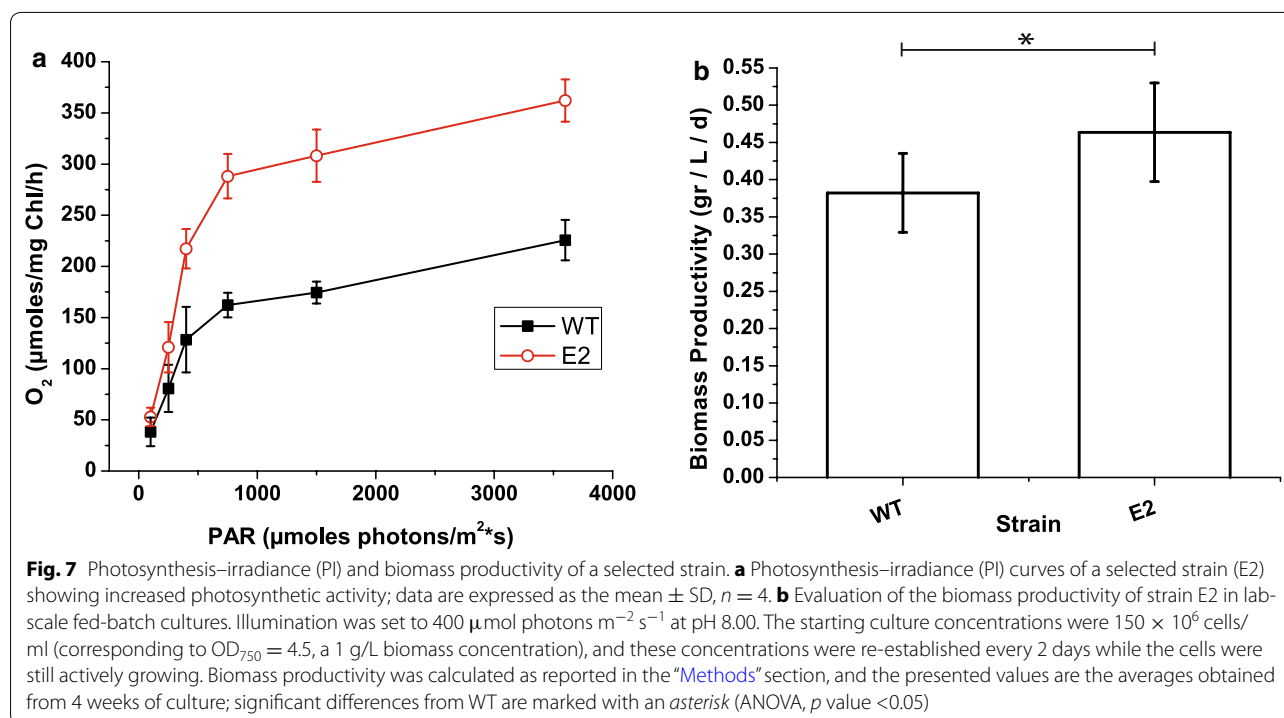
proteins in protection from light stress, similar to what has been observed for other algae species [42]. In contrast, strain E30 showed greater NPQ when exposed to strong illumination; in this case, it is worth noting that strong NPQ did not result from photoinhibition, as it was rapidly deactivated after the light was switched off (Fig. 6e).

For the mutant strain showing the largest increase in photosynthetic efficiency (ETR + PSII maximum quantum yield, E2), we also measured the photosynthesis–irradiance (PI) curves to evaluate the dependence of the oxygen evolution rate on light intensity (see “Methods” section for details). Similar to our observations of ETR measures, mutant strains E2 exhibited greater photosynthetic activity, with maximal activity that was ( $P_{\max}$ ) 44 % greater than the wild-type strain (Fig. 7a).

E2 strain productivity was also tested in fed-batch cultures, in conditions optimized for biomass productivity by providing excess CO<sub>2</sub> and enriching the media for nitrogen, phosphorus and iron (see “Methods” for details). The light intensity was set to 400 μmol photons m<sup>-2</sup> s<sup>-1</sup> to reach light saturation in WT cells exposed to direct illumination and leave the cells deeper within the



**Fig. 6** Measurements of photosynthetic activity in strains with reduced Chl contents. Photosynthetic activity in liquid cultures was monitored using in vivo Chl fluorescence (see “Methods” for details) at the end of exponential phase (4th day of growth). **a** PSII maximum quantum yield; **b, c** photosynthetic electron transport rate (ETR); **d, e** NPQ activation in cells exposed to increasing intensities of actinic light followed by 3 min of recovery in the dark. Only those mutant strains that significantly differed from WT are shown. The data are expressed as the mean ± SD, *n* = 4; values that significantly differed from wild-type are marked with an asterisk (ANOVA, *p* value <0.05)



culture light-limited. The pH of the fed-batch culture was set to 8.0 and fresh media was added every other day to restore the initial cell biomass concentration to 1 g/L to reproduce the high cell density conditions in industrial photobioreactors. As shown in Fig. 7b, the E2 strain exhibited a 21 % increase of biomass productivity with respect to the WT strain under these conditions. It is worth emphasizing that this strain is the one that exhibited a reduced PSII antenna size, confirming the hypothesis that a specific decrease in these pigment-binding complexes without altering the reaction center contents could indeed be beneficial [19]. Although they do not ensure increased productivity in industrial-scale PBRs, these results are proof of concept that the photosynthetic productivity of WT algae can indeed be improved by genetic modification in an industrially relevant species such as *N. gaditana*.

## Conclusions

Industrial cultivation of algae for biofuel production is highly promising, but will likely require the genetic optimization of these organisms. Here we described the generation and selection of a collection of *N. gaditana* strains with mutations affecting the photosynthetic apparatus, which is a valuable tool for achieving this objective.

Several mutant strains exhibiting reduced cellular chlorophyll contents were investigated in greater detail and they indeed exhibited improved photosynthetic activity,

which in one case resulted in an improved biomass productivity. This work shows that the genetic approaches described are indeed capable of generating strains with potentially improved productivity in an obligatory phototroph such as *N. gaditana*. Moreover, it is worth emphasizing that the isolated mutant strains contain a pool of photosynthetic alterations that the entire scientific community can use to better understand photosynthesis regulation in such a new, promising, model organism.

## Methods

### Microalgae growth

*Nannochloropsis gaditana* (strain 849/5) from the Culture Collection of Algae and Protozoa (CCAP) were used as the WT strain for the generation of mutant strains. Cells were grown in sterile *F/2* media with sea salts (32 g/L, Sigma Aldrich), 40 mM Tris-HCl (pH 8) and Guillard's (*F/2*) marine water enrichment solution (Sigma Aldrich) in Erlenmeyer flasks with  $100 \mu\text{mol photons m}^{-2} \text{s}^{-1}$  illumination and 100 rpm agitation at  $22 \pm 1^\circ \text{C}$  in a growth chamber. Fed-batch cultures were grown in 5 cm diameter Drechsel's bottles with a 250-ml working volume and bubbled using air enriched with 5 %  $\text{CO}_2$  (v/v); in this case, *F/2* growth media was enriched with added nitrogen, phosphate and iron sources (0.75 g/L  $\text{NaNO}_3$ , 0.05 g/L  $\text{NaH}_2\text{PO}_4$  and 0.0063 g/L  $\text{FeCl}_3 \cdot 6 \text{H}_2\text{O}$  final concentrations). Illumination at  $400 \mu\text{E m}^{-2} \text{s}^{-1}$  was provided by a LED Light Source SL 3500 (Photon Systems

Instruments, Brno, Czech Republic). The pH of the fed-batch cultures was set to 8.00 and fresh media added every other day to restore the starting cell biomass concentration of 1 g/L. Algal growth was measured by the change in optical density at 750 nm ( $OD_{750}$ ; Cary Series 100 UV-VIS spectrophotometer, Agilent Technologies) and cells were counted using a cell counter (Cellometer Auto X4, Nexcelom Bioscience). The biomass concentration was measured gravimetrically as a dry weight (DW) by filtering 5 ml of culture that had been diluted 1:5 to dissolve salts using 0.45  $\mu\text{m}$  pore size cellulose acetate filters. The filters were then dried at 70 °C for at least 24 h and the dry weights were measured in grams per liter. Biomass productivity was then calculated as  $([C_f] - [C_i]) / (t_f - t_i)$ , where  $C$  is the final or initial biomass concentration of the culture and  $t$  is the day number.

#### Chemical mutagenesis

To generate the mutant collection, mutagenesis conditions were set to induce 90 % cell mortality, ensuring a high mutation frequency. Microalgae suspensions ( $2 \times 10^7$  cells/mL) in the late exponential growth phase were mutagenized using 70 mM EMS (Sigma Aldrich) for 1 h in darkness at room temperature with mild agitation. The treated cells were then centrifuged at 5000g for 8 min, washed four times with sterile *F/2* media to remove excess EMS and then plated on *F/2* agar dishes. Plates were cultured at  $22 \pm 1$  °C with 20  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  illumination until algae colonies emerged (5 and 8 weeks after mutagenesis).

#### Insertional mutagenesis

A library of insertional mutant strains was generated via transformation of the *N. gaditana* strain CCAP 849/5 with the *Sh-ble* gene (from *Streptoalloteichus hindustanus*, kindly provided by Prof. Matthew Posewitz [35]) conferring resistance to zeocin. The DNA cassette (approximately 1.3 kb) containing the *Sh-ble* gene with an endogenous UBIQUITIN promoter and the FCPA terminator [35] was digested from the pPha-T1-UEP vector and purified on an agarose gel [0.8 % (w/v)]. For transformations,  $5 \times 10^8$  *N. gaditana* cells were washed four times with 375 mM sorbitol at 4 °C and resuspended in 100  $\mu\text{l}$  375 mM cold sorbitol. Cells were then incubated with 5  $\mu\text{g}$  DNA for 10 min on ice and then electroporated in a 2-mm cuvette using an ECM630 BTX electroporator set (500  $\Omega$ , 50  $\mu\text{F}$  and 2400 V). Following electroporation, the cells were maintained on ice for 5 min, resuspended in 10 ml *F/2* media and recovered for 24 h at  $22 \pm 1$  °C with agitation and 20  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  before plating onto *F/2* plates containing 3.5  $\mu\text{g/ml}$  zeocin. Resistant colonies were detected after 3 weeks and picked after 4–5 weeks.

#### Mutant selection and in vivo fluorescence-based high-throughput screening

Approximately 5–6 weeks after mutagenesis or transformation, colonies were collected and transferred. Colonies were then analyzed by in vivo fluorescence using a FluorCam FC 800 video-imaging apparatus (Photon Systems Instruments, Brno, Czech Republic) to identify those with mutations affecting the regulation of the photosynthetic apparatus. The selected mutant strains were retained if they were positively selected in at least three successive screening rounds.

#### Pigment content analysis

Chlorophyll a and total carotenoids were extracted from cells after 4 days of growth at the end of exponential phase using a 1:1 biomass to solvent ratio of 100 % *N,N*-dimethylformamide (Sigma Aldrich) [43]. Pigments were extracted at 4 °C in the dark for at least 24 h. Absorption spectra were determined between 350 and 750 nm using a Cary 100 spectrophotometer (Agilent Technologies) to spectrophotometrically determine pigment concentrations using specific extinction coefficients [43].

Absorption values at 664 and 480 nm were used to calculate the concentrations of Chlorophyll a and total carotenoids, respectively.

#### Measurements of photosynthetic activity and PSII functional antenna size

Oxygen evolution activity was recorded using a Clark-type  $\text{O}_2$  electrode (Hansatech, Norfolk, UK) at 25 °C. Cell suspensions (1.3 ml) containing  $325 \times 10^6$  cells were illuminated with a halogen lamp (KL 1500, Schott, Germany) after a 5-min dark adaptation. The concentrations of Chl a were determined after pigment extraction. PSII functionality was assessed in vivo by measuring Chl fluorescence using a PAM 100 fluorimeter (Heinz-Walz, Effeltrich, Germany) after a 20' dark adaptation. Samples were treated with increasingly intense light up to 2000  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ , and the light was then switched off to evaluate NPQ relaxation. NPQ and ETR values were calculated as previously described [38].

PSII antenna sizes were determined at the end of exponential phase (4th day of growth) using a JTS10 spectrophotometer. Samples ( $200 \times 10^6$  cells/ml final concentration) were adapted to the dark for 20 min and then incubated with 80  $\mu\text{M}$  DCMU for 10 min. Fluorescence induction kinetics were then monitored upon excitation with 320  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  of actinic light at 630 nm. The  $t_{2/3}$  values obtained from the induction curves were then used to calculate the size of the PSII functional antenna.



### SDS-PAGE electrophoresis and western blotting

Samples for western blotting analysis were collected from 4-day-old cultures in the late exponential phase. Cells were lysed using a Mini Bead Beater (Biospec Products) at 3500 RPM for 20 s in the presence of glass beads (150–212  $\mu\text{m}$  diameter), B1 buffer (400 mM NaCl, 2 mM  $\text{MgCl}_2$ , and 20 mM Tricine–KOH, pH 7.8), 0.5 % milk powder, 1 mM PMSE, 1 mM DNP- $\epsilon$ -amino-n-caproic acid and 1 mM benzamidine. Broken cells were then solubilized in 10 % glycerol, 45 mM Tris (pH 6.8), 30 mM dithiothreitol and 3 % SDS at RT for 20 min. Western blot analysis was performed by transferring the proteins to nitrocellulose (Bio Trace, Pall Corporation) and detecting them with alkaline phosphatase-conjugated antibodies. The antibodies recognizing D2, LHCf1 and LHCX1 proteins were produced by immunizing New Zealand rabbits with purified spinach protein [UniProt: P06005 for D2] or recombinant proteins obtained from cDNA overexpression in *E. coli* [UniProt: W7T4V5 for LHCf1, and UniProt: K8YWB4 for LHCX1].

### Additional file

**Additional file 1. Table S1.** In vivo Chl fluorescence parameters used for screening mutant strains. **Figure S1.** Verification of the successful integration of the *Sh-ble* gene in *N. gaditana* colonies. **Figure S2.** Example of fluorescence screening using homogenized inoculum. **Figure S3.** Outline of the three-step screening used to isolate *N. gaditana* photosynthetic mutant strains. **Figure S4.** Growth rates and pigment contents of selected insertional mutant strains. **Figure S5.** Western blot analysis of strains with reduced Chl contents that were selected from both screens.

### Abbreviations

Chl: chlorophyll; DCMU: 3-(3,4-dichlorophenyl)-1,1-dimethylurea; EMS: ethyl methane sulfonate; ETR: electron transport rate; GMO: genetically modified organism; LHC: light harvesting complex; NPQ: non-photochemical quenching; PBR: photobioreactor; PS: photosystem; ROS: reactive oxygen species; TLA: truncated light harvesting; WT: wild-type.

### Authors' contributions

TM designed the research; GP, AB and AS performed the mutagenesis screenings; GP, AB and AM physiologically characterized the mutant strains; GP, AA and TM wrote the manuscript. All authors read and approved the final manuscript.

### Author details

<sup>1</sup> Dipartimento di Biologia, Università di Padova, Via U. Bassi 58/B, 35121 Padua, Italy. <sup>2</sup> Centro studi di economia e tecnica dell'energia Giorgio Levi Cases, Università di Padova, Padua, Italy.

### Acknowledgements

This work was supported by the ERC starting Grant BIOLEAP nr 309485 and PRIN 2012XSAWYM to TM. AB is grateful to the Centro studi di economia e tecnica dell'energia Giorgio Levi Cases of the University of Padova for support. The authors also thank Prof. Matthew Charles Posewitz of the Colorado School of Mines (USA) for sharing the transformation construct.

### Compliance with ethical guidelines

### Competing interests

The authors declare that they have no competing interests.

Received: 29 May 2015 Accepted: 9 September 2015

Published online: 25 September 2015

### References

- Malcata FX. Microalgae and biofuels: a promising partnership? *Trends Biotechnol.* 2011;29:542–9.
- Simionato D, Block MA, La Rocca N, Jouhet J, Maréchal E, Finazzi G, Morosinotto T. The response of *Nannochloropsis gaditana* to nitrogen starvation includes de novo biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. *Eukaryot Cell.* 2013;12:665–76.
- Biondi N, Bassi N, Chini Zittelli G, De Faveri D, Giovannini A, Rodolfi L, Allevi C, Macrì C, Tredici MR. *Nannochloropsis* sp. F&M-M24: oil production, effect of mixing on productivity and growth in an industrial wastewater. *Environ Prog Sustain Energy.* 2013;32:846–53.
- Rodolfi L, Chini Zittelli G, Bassi N, Padovani G, Biondi N, Bonini G, Tredici MR. Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnol Bioeng.* 2009;102:100–12.
- Kirk J. Light and photosynthesis in aquatic ecosystems. 1994.
- Simionato D, Basso S, Giacometti GM, Morosinotto T. Optimization of light use efficiency for biofuel production in algae. *Biophys Chem.* 2013;182:71–8.
- Moody JW, McGinty CM, Quinn JC. Global evaluation of biofuel potential from microalgae. *Proc Natl Acad Sci USA.* 2014;111:8691–6.
- Sforza E, Cipriani R, Morosinotto T, Bertuccio A, Giacometti GM. Excess  $\text{CO}_2$  supply inhibits mixotrophic growth of *Chlorella protothecoides* and *Nannochloropsis salina*. *Bioresour Technol.* 2012;104:523–9.
- Zou N, Richmond A. Light-path length and population density in photoacclimation of *Nannochloropsis* sp. (Eustigmatophyceae). *J Appl Phycol.* 2000;12:349–54.
- Carvalho AP, Silva SO, Baptista JM, Malcata FX. Light requirements in microalgal photobioreactors: an overview of biophotonic aspects. *Appl Microbiol Biotechnol.* 2011;89:1275–88.
- Peers G, Truong TB, Ostendorf E, Busch A, Elrad D, Grossman AR, Hippler M, Niyogi KK. An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature.* 2009;462:518–21.
- Erickson E, Wakao S, Niyogi KK. Light stress and photoprotection in *Chlamydomonas reinhardtii*. *Plant J.* 2015;82:449–65.
- Derks A, Schaven K, Bruce D. Diverse mechanisms for photoprotection in photosynthesis. Dynamic regulation of photosystem II excitation in response to rapid environmental change. *Biochim Biophys Acta Bioenerg.* 2015;1847:468–85.
- Tian W, Chen J, Deng L, Yao M, Yang H, Zheng Y, Cui R, Sha G. An irradiation density dependent energy relaxation in plant photosystem II antenna assembly. *Biochim Biophys Acta.* 2015;1847:286–93.
- Melis A, Neidhardt J, Benemann JR. *Dunaliella salina* (Chlorophyta) with small chlorophyll antenna sizes exhibit higher photosynthetic productivities and photon use efficiencies than normally pigmented cells. *J Appl Phycol.* 1998;10:515–25.
- Melis A. Solar energy conversion efficiencies in photosynthesis: minimizing the chlorophyll antennae to maximize efficiency. *Plant Sci.* 2009;177:272–80.
- Cuaresma M, Janssen M, Vilchez C, Wijffels RH. Productivity of *Chlorella sorokiniana* in a short light-path (SLP) panel photobioreactor under high irradiance. *Biotechnol Bioeng.* 2009;104:352–9.
- Kirst H, Melis A. The chloroplast signal recognition particle (CpSRP) pathway as a tool to minimize chlorophyll antenna size and maximize photosynthetic productivity. *Biotechnol Adv.* 2014;32:66–72.
- Formighieri C, Franck F, Bassi R. Regulation of the pigment optical density of an algal cell: filling the gap between photosynthetic productivity in the laboratory and in mass culture. *J Biotechnol.* 2012;162:115–23.
- Wobbe L, Remeacle C. Improving the sunlight-to-biomass conversion efficiency in microalgal biofactories. *J Biotechnol.* 2015;201:28–42.
- Melis A. Dynamics of photosynthetic membrane composition and function. *BBA Bioenerg.* 1991;1058:87–106.
- Mussgnug JH, Wobbe L, Elles I, Claus C, Hamilton M, Fink A, Kahmann U, Kapazoglou A, Mullineaux CW, Hippler M, Nickelsen J, Nixon PJ, Kruse

- O. NAB1 is an RNA binding protein involved in the light-regulated differential expression of the light-harvesting antenna of *Chlamydomonas reinhardtii*. *Plant Cell*. 2005;17:3409–21.
23. Nakajima Y, Ueda R. The effect of reducing light-harvesting pigment on marine microalgal productivity. *J Appl Phycol*. 2000;12:285–90.
  24. Bonente G, Formighieri C, Mantelli M, Catalanotti C, Giuliano G, Morosinotto T, Bassi R. Mutagenesis and phenotypic selection as a strategy toward domestication of *Chlamydomonas reinhardtii* strains for improved performance in photobioreactors. *Photosynth Res*. 2011;108:107–20.
  25. Cazzaniga S, Dall'Osto L, Szaub J, Scibilia L, Ballottari M, Purton S, Bassi R. Domestication of the green alga *Chlorella sorokiniana*: reduction of antenna size improves light-use efficiency in a photobioreactor. *Biotechnol Biofuels*. 2014;7:157.
  26. Polle JEW, Benemann JR, Tanaka A, Melis A. Photosynthetic apparatus organization and function in the wild type and a chlorophyll b-less mutant of *Chlamydomonas reinhardtii*. Dependence on carbon source. *Planta*. 2000;211:335–44.
  27. Mussgnug JH, Thomas-Hall S, Rupprecht J, Foo A, Klassen V, McDowall A, Schenk PM, Kruse O, Hankamer B. Engineering photosynthetic light capture: impacts on improved solar energy to biomass conversion. *Plant Biotechnol J*. 2007;5:802–14.
  28. Wobbe L, Blifernoz O, Schwarz C, Mussgnug JH, Nickelsen J, Kruse O. Cysteine modification of a specific repressor protein controls the translational status of nucleus-encoded LHClI mRNAs in *Chlamydomonas*. *Proc Natl Acad Sci USA*. 2009;106:13290–5.
  29. Kirst H, García-Cerdán JG, Zurbriggen A, Melis A. Assembly of the light-harvesting chlorophyll antenna in the green alga *Chlamydomonas reinhardtii* requires expression of the TLA2-CpFTSY gene. *Plant Physiol*. 2012;158:930–45.
  30. Elrad D. A major light-harvesting polypeptide of photosystem II functions in thermal dissipation. *Plant Cell Online*. 2002;14:1801–16.
  31. Horton P, Ruban A. Molecular design of the photosystem II light-harvesting antenna: photosynthesis and photoprotection. *J Exp Bot*. 2005;56:365–73.
  32. De Mooij T, Janssen M, Cerezo-Chinarro O, Mussgnug JH, Kruse O, Ballottari M, Bassi R, Bujaldon S, Wollman FA, Wijffels RH. Antenna size reduction as a strategy to increase biomass productivity: a great potential not yet realized. *J Appl Phycol*. 2014;27(3):1063–77.
  33. Huesemann MH, Hausmann TS, Bartha R, Aksoy M, Weissman JC, Benemann JR. Biomass productivities in wild type and pigment mutant of *Cyclotella* sp. (Diatom). *Appl Biochem Biotechnol*. 2009;157:507–26.
  34. Bonente G, Ballottari M, Truong TB, Morosinotto T, Ahn TK, Fleming GR, Niyogi KK, Bassi R. Analysis of LHC3R3, a protein essential for feedback de-excitation in the green alga *Chlamydomonas reinhardtii*. *PLoS Biol*. 2011;9:e1000577.
  35. Radakovits R, Jinkerson RE, Fuerstenberg SI, Tae H, Settlage RE, Boore JL, Posewitz MC. Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. *Nat Commun*. 2012;3:686.
  36. Doan T, Obbard J. Enhanced intracellular lipid in *Nannochloropsis* sp. via random mutagenesis and flow cytometric cell sorting. *Algal Res*. 2012;1(1):17–21.
  37. Blaby IK, Blaby-Haas CE, Tourasse N, Hom EFY, Lopez D, Aksoy M, Grossman A, Umen J, Dutcher S, Porter M, King S, Witman GB, Stanke M, Harris EH, Goodstein D, Grimwood J, Schmutz J, Vallon O, Merchant SS, Prochnik S. The *Chlamydomonas* genome project: a decade on. *Trends Plant Sci*. 2014;19:672–80.
  38. Maxwell K, Johnson GN. Chlorophyll fluorescence—a practical guide. *J Exp Bot*. 2000;51(345):659–68.
  39. Simionato D, Sforza E, Corteggiani Carpinelli E, Bertuccio A, Giacometti GM, Morosinotto T. Acclimation of *Nannochloropsis gaditana* to different illumination regimes: effects on lipids accumulation. *Bioresour Technol*. 2011;102:6026–32.
  40. Vieler A, Wu G, Tsai CHH, Bullard B, Cornish AJ, Harvey C, Reza IBB, Thornburg C, Achawanantakun R, Buehl CJ, Campbell MS, Cavalier D, Childs KL, Clark TJ, Deshpande R, Erickson E, Armenia-Ferguson A, Handee W, Kong Q, Li X, Liu B, Lundback S, Peng C, Roston RL, Sanjaya Simpson JP, TerBush A, Warakanont J, Zäuner S, Farre EM, et al. Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779. *PLoS Genet*. 2012;8:e1003064.
  41. Basso S, Simionato D, Gerotto C, Segalla A, Giacometti GM, Morosinotto T. Characterization of the photosynthetic apparatus of the Eustigmatophyceae *Nannochloropsis gaditana*: evidence of convergent evolution in the supramolecular organization of photosystem I. *Biochim Biophys Acta*. 2014;1837:306–14.
  42. Niyogi KK, Truong TB. Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr Opin Plant Biol*. 2013;16:307–14.
  43. Wellburn AR. The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *J Plant Physiol*. 1994;144:307–13.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)





## SUPPLEMENTARY MATERIAL.

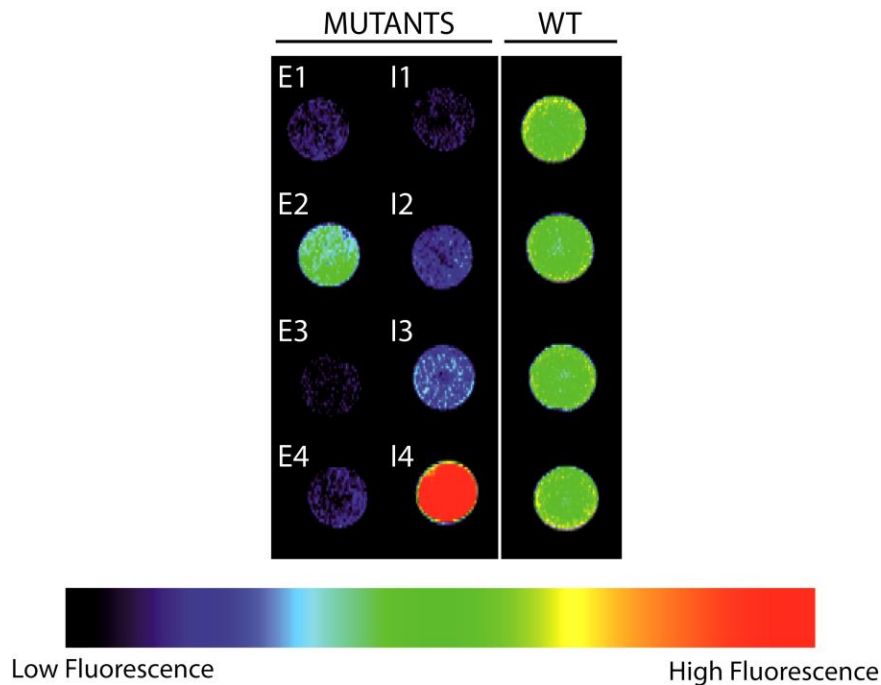
**Table S1. *In vivo* Chl fluorescence parameters used for screening mutant strains.** The parameters reported below were employed to select strains with potential alterations in photosynthetic apparatus.

Fluorescence parameter	Phenotype
$F_0/\text{Area}$	Fluorescence intensity per area, indication of alteration in pigment contents
$\Phi_{\text{PSII}}$ in dark adapted cells	Alteration in cells photochemical ability
$\Phi_{\text{PSII}}$ in light adapted cells	Strong decrease / higher values in quantum yield indicate light sensitivity / improved photochemical ability.
NPQ in light adapted cells	Alteration in cell ability of heat dissipation

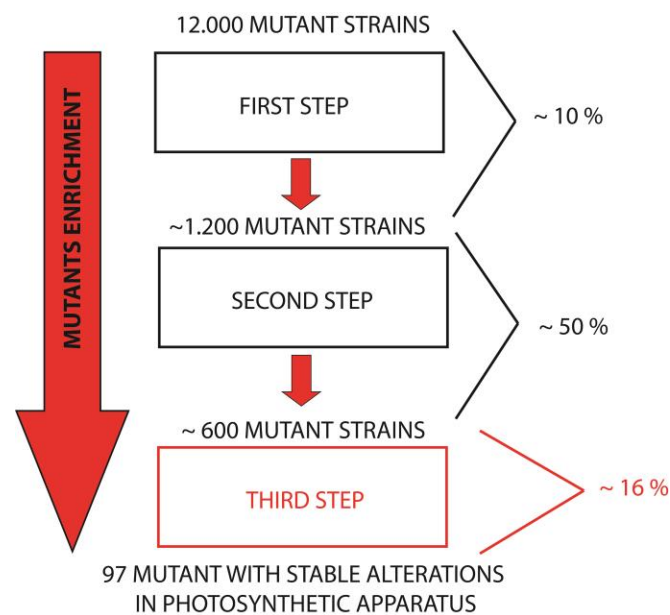
**Figure S1. Verification of the successful integration of the *Sh-ble* gene in *N. gaditana* colonies.** Colony PCR in *N. gaditana* insertional mutant strains confirming the presence of the transgene in six independent transformants (lanes 1-6), and its absence in a WT control (C-). Colony PCR templates were obtained with the Chelex-100 (BioRad) method (Cao et al., 2009) and were performed using the following primers: ZEO\_FOR: ATGGCCAAGTTGACCAAGTGC, ZEO\_REV: TCAGTCCTGCTCCTCGGCC. C-, WT *N. gaditana* strain electroporated with water; C+, plasmid DNA (pPha-T1-UEP vector). (Cao, M., Fu, Y., Guo, Y., and Pan, J. (2009). *Chlamydomonas* (Chlorophyceae) colony PCR. *Protoplasma* **235**: 107–10.)



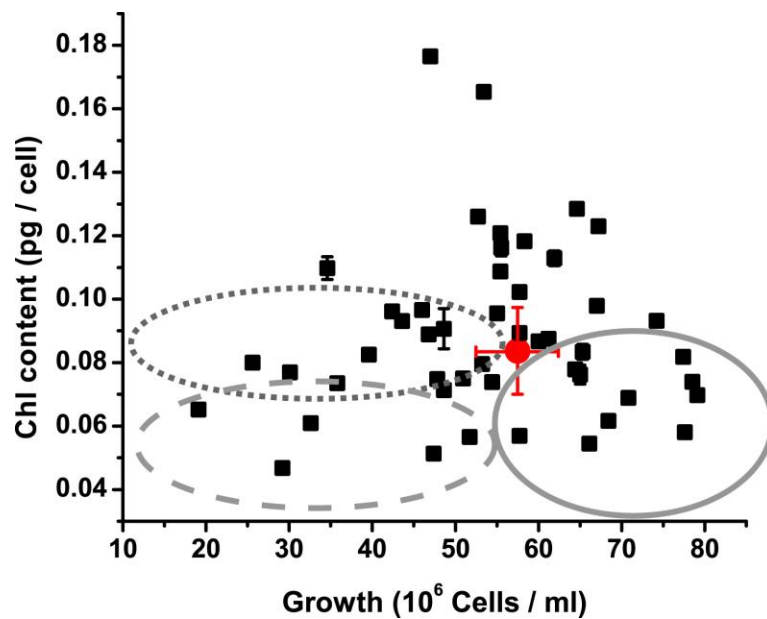
**Figure S2. Example of fluorescence screening using homogenized inoculum.** Fluorescence imaging ( $F_0$ ) of F/2 agar plates in which all colonies were spotted homogenously, at equal cellular concentration ( $OD_{750}=0.2$ ). The first two lanes on the left show mutant strains coming from the chemical (E1-4) and the insertional mutagenesis (I1-4), respectively. The third lane shows 4 replicas of the WT strain.



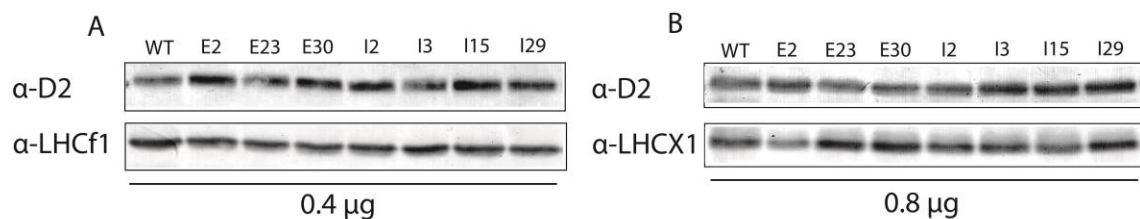
**Figure S3. Outline of the three-step screening used to isolate *N. gaditana* photosynthetic mutant strains.** In red is underlined the third round of screening after colonies inoculation at the same starting cellular concentration.



**Figure S4. Growth rates and pigment contents of selected insertional mutant strains.** Selected mutant strains produced by insertional mutagenesis were tested in liquid cultures and monitored for growth and cellular pigment content at the end of exponential phase (4<sup>th</sup> day of growth, see methods for details). Mutant strains with reduced cellular pigment content but unaffected growth compared to WT are circled with a continuous line. Mutant strains exhibiting both reduced pigment content and growth compared with WT colonies are circled with a dashed line. Strains with unaltered pigment content but with reduced growth are circled with a dotted line. WT (n = 7) is shown as red circle; mutant strains n = 2.



**Figure S5. Western blot analysis of strains with reduced Chl contents that were selected from both screens.** Western blot quantification of LHCf1 (A) and LHCX1 (B), two major antenna proteins of *Nannochloropsis gaditana* photosynthetic apparatus. As a control, the PSII core subunit D2 was also detected on the same membrane. Total cells extracts corresponding to the same total Chl content were loaded (0.4  $\mu$ g for LHCf1 and 0.8  $\mu$ g for LHCX1).





# CHAPTER III

## **Connection Between Photosynthetic Properties and Biomass Productivity in *Nannochloropsis gaditana* Cultures in Industrially Relevant Conditions**

### **Author names and affiliations**

Giorgio Perin<sup>1</sup>, Davide Dal Bo<sup>1</sup>, Tomas Morosinotto<sup>1</sup>

<sup>1</sup>PAR-Lab\_Padua Algae Research Laboratory, Department of Biology,  
University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy



**Abstract.**

As discussed in previous chapters, microalgae cultivated in a photobioreactor or pond are exposed to artificial environmental conditions, different from the ones they adapted to, during evolution.

In chapter II, we described the generation of a collection of *Nannochloropsis gaditana* random mutants, with the final aim to isolate strains with a tuned composition / regulation of the photosynthetic apparatus, to meet the needs of the photobioreactor artificial environment. In that work, we proved that genetic engineering efforts directed toward the generation of photosynthetic phenotypes can indeed lead to an improved biomass productivity. Here we want to deeply investigate the impact of these phenotypes on overall photosynthetic performances, to better understand the connection between photosynthesis and biomass productivity, in industrially relevant growth conditions. We focused on two major phenotypes as selection criteria, the reduction in Chl content and in the ability to activate heat dissipation mechanisms. In this chapter, we describe their application in industrial-simulating growth conditions, to prove their theoretical potential of enhancing overall productivity by improving light distribution and energy conversion mechanisms.

## Introduction.

The photosynthetic features that have been positively selected during evolution in wild type species, are indeed detrimental for algae industrial cultivation and, as a consequence, the direct exploitation of parental strains is economically unsustainable (Formighieri et al., 2012). The development of any industrial process based on algae biomass (e.g. biofuels or bio-commodities production) requires therefore their biotechnological optimization or domestication (Radakovits et al., 2010). To overcome the current limitations in light to biomass conversion efficiencies, the generation of strains with a tuned composition / regulation of the photosynthetic apparatus is seminal (Simionato et al., 2013a; Wobbe and Remacle, 2014).

Strains with alterations in their photosynthetic properties were already described as showing a higher maximum photosynthetic rate, as well as light saturation efficiencies. This work was performed especially in the model genus *Chlamydomonas* (Nakajima and Ueda, 2000; Nakajima et al., 2001; Bonente et al., 2011b; Wobbe et al., 2009; Beckmann et al., 2009; Huesemann et al., 2009), but also for some other species with a major industrial interest (Melis et al., 1998; Lea-Smith et al., 2014; Cazzaniga et al., 2014).

However, despite their theoretical benefits, when tested in industrial-scale cultivation systems, these strains presented contrasting results, and some of the mutants showed no improvement in biomass productivity when tested in photobioreactors (de Mooij et al., 2014). This demonstrates that we still have a limited knowledge on which phenotypes are really driving to an improved biomass productivity and particularly in which conditions. In chapter II, we described the forward genetics approach followed to isolate *Nannochloropsis gaditana* random mutants affected in their photosynthetic properties. Here we would like to investigate further the connection between phenotype and biomass productivity in industrially relevant conditions. For this reason, we selected two strains with photosynthetic phenotypes with a theoretical beneficial effect on productivity, namely a reduction in Chl content and in the ability to activate heat dissipation mechanisms. Here we describe their exploitation in industrial-simulating growth conditions, to evaluate whether in this promising organisms the selected phenotypes indeed show an improved biomass productivity and in which operational conditions. The global photosynthetic performances of strains performing in intensive cultivation conditions was evaluated.

## Materials and Methods.

### *Microalgae growth.*

*Nannochloropsis gaditana*, strain 849/5, from the Culture Collection of Algae and Protozoa (CCAP) were used as the WT strain. A random mutagenesis was performed as previously described (Perin et al., 2015), and the 2 mutants (E2 and I48) here tested were isolated among the pool of strains showing altered photosynthetic performances.

Mutants were maintained in F/2 solid media, with sea salts (32 g/L, Sigma Aldrich), 40 mM Tris-HCl (pH 8), Guillard's (F/2) marine water enrichment solution (Sigma Aldrich), 1 % agar (Duchefa Biochemie), supplied with 3.5 µg/ml zeocin for strain I48. Cells were pre-cultured in sterile F/2 liquid media in Erlenmeyer flasks with 100 µmol photons m<sup>-2</sup> s<sup>-1</sup> illumination and 100 rpm agitation at 22 ± 1 °C in a growth chamber.

Fed-batch cultures were performed in 5 cm diameter Drechsel's bottles, illuminated from one side, with a 250 ml working volume and bubbled using air enriched with 5 % CO<sub>2</sub> (v/v) (for bubbling apparatus set-up see (Sforza et al., 2012a)); in this case, F/2 growth media was enriched with added nitrogen, phosphate and iron sources (0.75 g/L NaNO<sub>3</sub>, 0.05 g/L NaH<sub>2</sub>PO<sub>4</sub> and 0.0063 g/L FeCl<sub>3</sub>•6H<sub>2</sub>O final concentrations). Constant illumination at 100 or 400 µE m<sup>-2</sup> s<sup>-1</sup> was provided by daylight fluorescent lamps. The pH of the fed-batch cultures was set to 8.00 and fresh media was added every day to restore the starting cell biomass concentration of 1 or 1.7 gr/L. Algal growth was monitored using a cell counter (Cellometer Auto X4, Nexcelom Bioscience). The biomass concentration (dry weight) was measured gravimetrically as previously reported in (Perin et al., 2015). Biomass productivity was calculated as  $([C_f] - [C_i]) / (t_f - t_i)$ , where C is the final or initial biomass concentration of the culture and t is the days number.

### *Pigment extraction and analyses.*

Pigments were extracted from cells grown in fed-batch conditions, during active growth. Chlorophyll a and total carotenoids were extracted from intact cells, using a 1:1 biomass to solvent ratio of 100 % N,N-dimethylformamide (Sigma Aldrich) (Moran and Porath, 1980), at 4 °C in the dark for at least 24 h. Absorption spectra were registered between 350 and 750 nm using a Cary 100 spectrophotometer (Agilent Technologies) to spectrophotometrically determine pigment concentrations, using specific extinction coefficients (Porra et al., 1989; Wellburn, 1994).

Absorption values at 664 and 480 nm were used to calculate the concentrations of Chlorophylla and total carotenoids, respectively.

High-pressure liquid chromatography (HPLC) analyses were carried out in order to determine the individual carotenoids cells content. It was performed using a Beckman System Gold instrument equipped with a UV-VIS detector and a C-18 column (25 cm by 4.6 mm; Zorbax octyldecyl silane).

Pigments were extracted from cells lysed using a Mini Bead Beater (Biospec Products) at 3500 RPM for 20 s in the presence of glass beads (150–212  $\mu\text{m}$  diameter) and 80 % acetone.

Runs were performed using 86.7 % acetonitrile, 9.6 % methanol and 3.6 % Tris-HCl, pH 7.8 and 80 % methanol and 20 % hexane as elution solvent.

The peaks of each sample were identified through the retention time and absorption spectrum (Jeffrey et al., 1997) and quantified as described previously (Färber and Jahns, 1998). The vaucheriaxanthin retention factor was estimated by correcting that of violaxanthin for their different absorption at 440 nm.

#### *Spectroscopic analysis.*

*Photosynthetic performances estimation.* Photosynthesis monitoring was performed by measuring *in vivo* Chl fluorescence using a PAM 100 fluorimeter (Heinz-Walz, Effeltrich, Germany). PSII functionality was expressed as PSII maximum quantum yield, according to (Maxwell and Johnson, 2000). Samples were treated with increasingly intense light up to 2000  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ , and then it was switched off to evaluate NPQ relaxation kinetic. NPQ and ETR values were calculated as previously described (Maxwell and Johnson, 2000).

*PSII functional antenna size evaluation.* PSII antenna sizes were determined for fed-batch cultures in active growth conditions, using a JTS-10 spectrophotometer (Biologic, France). Samples ( $200 \times 10^6$  cells/ml final concentration) were dark-adapted for 20 minutes and incubated with 80  $\mu\text{M}$  DCMU for 10 minutes. Fluorescence induction kinetics were then monitored upon excitation with 320  $\mu\text{mol photons / (m}^2 \text{s)}$  of actinic light at 630 nm. The  $t_{2/3}$  values obtained from the induction curves were used to calculate the size of the PSII functional antenna.

*P<sub>700</sub> and total electron flow estimation.* The P<sub>700</sub> and the total electron flow (TEF) spectroscopic quantifications were performed measuring P<sub>700</sub> (the primary electron donor to PSI) absorption at 705 nm in intact cells. Analyses were conducted exposing the samples ( $300 \times 10^6$  cells/ml final concentration) to saturating actinic light (2050  $\mu\text{mol of photons / (m}^2 \text{s)}$ , at 630 nm) for 15 seconds to maximize P<sub>700</sub> oxidation (P<sub>700</sub><sup>+</sup>) and reach a steady state. Then the light was switched off to follow the

$P_{700}^+$  re-reduction kinetic in the dark for 5 seconds. The total electron flow (TEF) was calculated from the monitoring of the  $P_{700}^+$  re-reduction rates after illumination in untreated cells. The electron transport rate was calculated assuming a single exponential decay of  $P_{700}^+$ . This allowed calculating the rate constant of  $P_{700}^+$  reduction as  $1/\tau$ . By multiplying the rate constant with the fraction of oxidized  $P_{700}$ , and considering this value as 1 in DCMU- and DBMIB-treated cells (see below) we could evaluate the number of electrons transferred per unit of time per PSI unit (Simionato et al., 2013b). The PSI content was evaluated from the maximum change in the absorption of  $P_{700}^+$  in cells treated with DCMU (80  $\mu$ M) and DBMIB (dibromothymoquinone, 300  $\mu$ M) at a saturating actinic light (2050  $\mu$ mol of photons / ( $m^2$  s), at 630 nm). Under these conditions, re-reduction of  $P_{700}^+$  through photosynthetic electron flow is largely slowed down, allowing to evaluate the full extent of photo-oxidizable  $P_{700}$  (Simionato et al., 2013b; Alric, 2010).

#### *Lipid content evaluation.*

The lipid content analyses and their classification were carried out by the DAFNAE of Padova University, applying the method of (Matyash et al., 2008). Lipids were extracted from dried cells samples.

## Results and discussion.

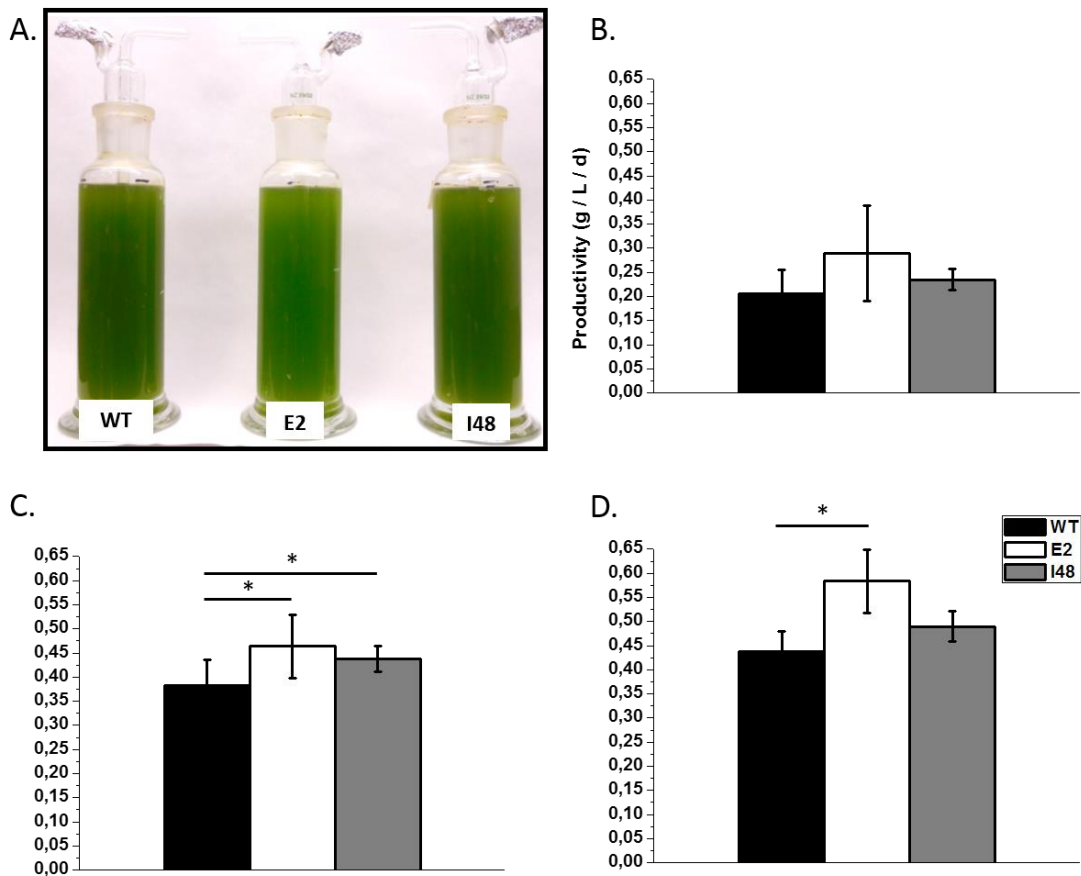
In chapter II, *Nannochloropsis gaditana* strain CCAP 849/5, was used as wild type strain to generate a collection of random mutants, further selected for alterations in their photosynthetic performances (Perin et al., 2015). In order to investigate further the connection between photosynthetic phenotypes and biomass productivity (Formighieri et al., 2012), we selected two strains with potentially interesting phenotypes, the reduction in Chl content and in the ability to activate heat dissipation mechanisms, E2 and I48.

### *Biomass productivity in intensive fed-batch cultures.*

To evaluate whether the above described photosynthetic phenotypes are indeed advantageous when *Nannochloropsis gaditana* is applied for intensive cultivation, we tested strains E2 and I48 in fed-batch cultures (Figure 1A). Fed-batch cultures operates in semi-continuous mode and everyday the starting biomass concentration value was restored by adding fresh media (see Methods for details). This operational strategy, although more complex, was chosen since it meets the needs of industrial-scale plants where it offers a reduction in downtime than batch processes, resulting in a more efficient use of equipment and lower investment costs (Benvenuti et al., 2015). Moreover, this cultivation mode, when maximal biomass productivity is achieved, ensures the culture stability over prolonged time periods. A further advantage is that with this system it is possible to work at high biomass concentration, that was in fact set to a value of 1 g / L, corresponding to  $OD_{750} = 4.5$  to simulate the growth conditions achieved in industrial cultivation systems (Benvenuti et al., 2015). Nutrients (nitrogen, phosphate and iron sources) were provided in excess (see Methods for details on their concentrations), as well as the CO<sub>2</sub> supply, according to the experimental set-up described in the methods section. These growth conditions thus ensured that microalgae growth rates were not limited by nutrients or CO<sub>2</sub> availability, but only by light-use efficiency, in order to highlight the theoretical benefits of these photosynthetic phenotypes.

Two light conditions were chosen, 100 and 400  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ). The former is just below the optimal light intensity for *Nannochloropsis* (Sforza et al., 2012b). Furthermore, considering that here we are working with a dense culture in 5 cm diameter (see Methods for system geometry description), this light intensity should be limiting for growth. On the other side, the latter should be in excess, at least for the first cells layers, which absorb most of the light, leaving the cells populating the inner layers of culture in light limitation (Formighieri et al., 2012; Ort and Melis, 2011). In fact, when light

intensity was switched from 100  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ) to 400  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ), leaving unaltered the starting biomass concentration to 1 g / L, the parental strain showed an improvement in biomass productivity of the 86 %, proving that 100  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ) were indeed limiting for growth. When the starting biomass concentration was instead increased to 1.7 g / L, leaving unaltered the light intensity to 400  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ) and consequently improving light inhomogeneity, the parental strain still improved its biomass productivity of the 15 %, highlighting that light was in excess in the previous condition.



**Figure 1. Biomass productivity in fed-batch cultures.** A. Image of the Drechsel's bottles in which the fed-batch cultivation for wild type and mutants strains was performed. The picture was taken for cultures at the same starting cellular concentration. B. Biomass productivity obtained from fed-batch cultures at 100  $\mu\text{mol}$  photons / ( $\text{m}^2 \text{s}$ ), for starting biomass concentration of 1 g/L (corresponding to  $\text{OD}_{750} = 4,5$ ; starting cellular concentration of  $150 \times 10^6$  cells/ml). C. Biomass productivity obtained for cultures illuminated with 400  $\mu\text{mol}$  photons / ( $\text{m}^2 \text{s}$ ), with a starting biomass concentration of 1 g/L. D. Biomass productivity obtained from cultures at 400  $\mu\text{mol}$  photons / ( $\text{m}^2 \text{s}$ ), for starting biomass concentration of 1,7 g/L (corresponding to  $\text{OD}_{750} = 6,5$ ; starting cellular concentration of  $250 \times 10^6$  cells/ml). The starting cellular concentrations were re-established every day, while the cells were

actively growing. The data are expressed as the average values obtained from 2 weeks of culture  $\pm$  SD; statistically significant values from wild type are marked with an asterisk (one-way ANOVA,  $p$ -value  $< 0.05$ ).

However, when cultures with a starting biomass concentration of 1 g / L were exposed to 100  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ), both mutant strains showed the same biomass productivity than the WT (Figure 1B). When light intensity was increased up to 400  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ), using the same starting cellular concentrations, both mutants instead showed a significant improvement in biomass productivity with respect to the parental strain of 21 and 15 % for E2 and I48 strains, respectively (Figure 1C). When starting cellular concentrations were increased even further, up to 1.7 g / L, while leaving unaltered light intensity conditions, strain E2 still performed better than WT, enhancing its difference from the latter and producing 33 % more biomass than the parental strain, in these experimental conditions (Figure 1D).

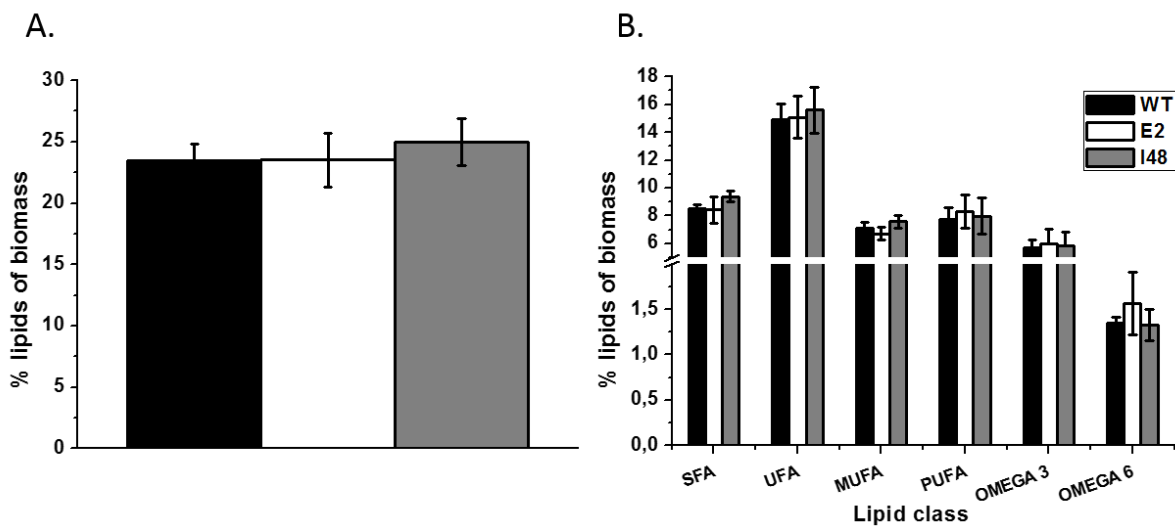
These data show that the selected photosynthetic phenotypes, in *Nannochloropsis gaditana* are indeed capable of enhancing biomass productivity, when exploited in intensive cultivation conditions, as those here described. However, the most important conclusion is that the productivity increase strongly correlates with the operational conditions. In fact, the latter have a strong impact on several growth parameters, such as the light path and intensity that reach cells populating different culture layers. These parameters could also have a strong influence on the strains photosynthetic performances and could affect the balance between positive and negative effects of the mutation, when applied in intensive growth conditions. The cultivation condition in which both mutant strains showed an improved biomass productivity with respect to the parental strain (1 g / L starting concentration with 400  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ) irradiance) was selected for further characterization.

*Both mutant strains are not impaired in the lipid content.*

Given their possible future industrial application, we evaluated whether these mutant strains showed an impairment in the lipid metabolism as a result of the mutagenesis. Mutant strains were thus evaluated for their lipid content in the selected operational conditions. Lipids were extracted from dried cells samples as described in the methods section. These growth conditions led all the tested strains accumulate up to 25 % (w/w of biomass) of lipids, even if cells were not exposed to environmental stress conditions. These values are close to those achieved in pilot industrial-scale plants as reported in (Benvenuti et al., 2015) and therefore highlights that the operational conditions, chosen to test whether mutant strains are more productive than parental strain, are close to those



commonly exploited in the field. Both mutant strains didn't show any impairment in the lipid content with respect to the WT (Figure 2A), meaning that the mutations at the level of photosynthetic apparatus did not affect lipid biosynthetic pathways. Also lipids composition showed no major differences (Figure 2B). When the single lipid classes abundance was monitored in detail (Table 1), we showed that C16:0 are the main lipid class also in the selected mutant strains, followed by C16:1, C20:5 and C14:0 in order of abundance, respectively. Therefore E2 and I48 strains are not impaired in the accumulation of the main molecular products interesting for their nutraceutics (C20:5, EPA – eicosapentaenoic acid) or biodiesel production applications.



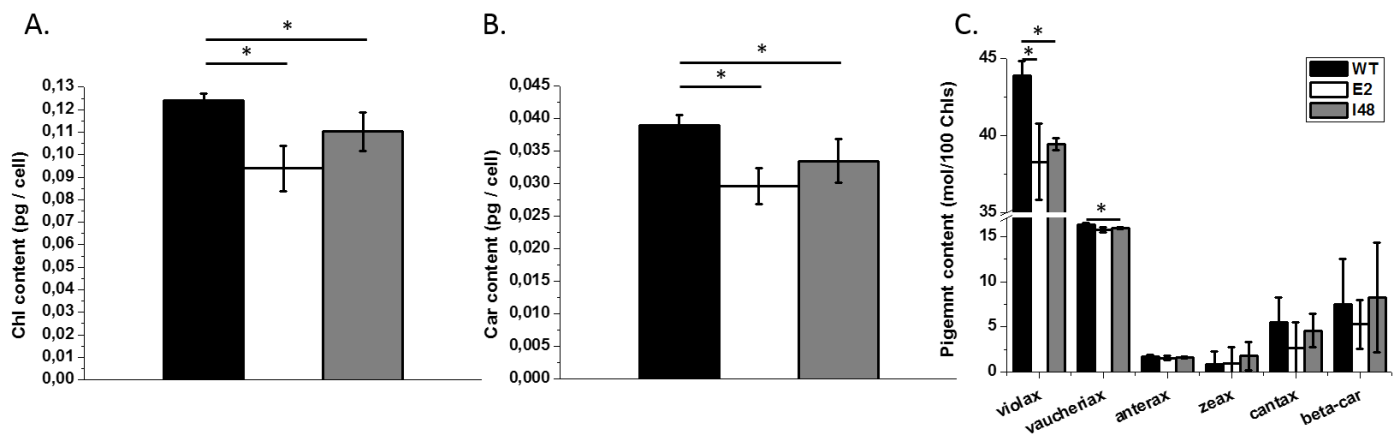
**Figure 2. Lipid content of *N. gaditana* wild type (WT) and mutant (E2 and I48) strains.** A. Lipid content of wild type and mutant strains cells in fed-batch growth conditions reported as percentage (%) of total biomass. B. Classification of the lipid content in different lipid clusters. The sum of the different classes doesn't match with the total lipid content, because they aren't independent one from the other; SFA, saturated fatty acids; UFA, unsaturated fatty acids; MUFA, mono-unsaturated fatty acids; PUFA, poly-unsaturated fatty acids; OMEGA 3 ( $\omega$ -3); OMEGA 6 ( $\omega$ -6). The data are expressed as mean  $\pm$  SD, n = 3.

**Table 1. Fatty acids composition of *N. gaditana* wild type (WT) and mutant (E2 and I48) strains. The data are expressed as mean  $\pm$  Standard Deviation (SD), n = 3.**

Lipid class	WT	E2	I48
	% of biomass		
<b>C8:0</b>	0,05 $\pm$ 0,01	0,05 $\pm$ 0,02	0,05 $\pm$ 0,01
<b>C9:0</b>	0,01 $\pm$ 0,01	0,01 $\pm$ 0,01	0,01 $\pm$ 0,00
<b>C10:0</b>	0,03 $\pm$ 0,01	0,04 $\pm$ 0,02	0,03 $\pm$ 0,01
<b>C12:0</b>	0,11 $\pm$ 0,01	0,10 $\pm$ 0,02	0,12 $\pm$ 0,02
<b>C14:0</b>	1,15 $\pm$ 0,13	1,04 $\pm$ 0,15	1,39 $\pm$ 0,13
<b>C15:0</b>	0,07 $\pm$ 0,00	0,09 $\pm$ 0,02	0,07 $\pm$ 0,00
<b>C16:0</b>	6,90 $\pm$ 0,28	6,89 $\pm$ 0,84	7,50 $\pm$ 0,41
<b>C17:0</b>	0,06 $\pm$ 0,00	0,06 $\pm$ 0,01	0,07 $\pm$ 0,01
<b>C18:0</b>	0,11 $\pm$ 0,01	0,11 $\pm$ 0,01	0,12 $\pm$ 0,02
<b>C19:0</b>	0	0	0
<b>C20:0</b>	0	0	0
<b>C24:0</b>	0,03 $\pm$ 0,01	0,05 $\pm$ 0,01	0,03 $\pm$ 0,02
<b>C12:1</b>	0,01 $\pm$ 0,01	0,02 $\pm$ 0,01	0,02 $\pm$ 0,01
<b>C14:1</b>	0,01 $\pm$ 0,00	0,01 $\pm$ 0,01	0,02 $\pm$ 0,00
<b>C15:1</b>	0,01 $\pm$ 0,00	0,01 $\pm$ 0,01	0,01 $\pm$ 0,00
<b>C16:1</b>	6,66 $\pm$ 0,34	6,13 $\pm$ 0,55	7,12 $\pm$ 0,41
<b>C17:1</b>	0,03 $\pm$ 0,00	0,04 $\pm$ 0,00	0,03 $\pm$ 0,00
<b>C18:1</b>	0,41 $\pm$ 0,03	0,53 $\pm$ 0,16	0,41 $\pm$ 0,05
<b>C19:1</b>	0	0	0
<b>C20:1</b>	0	0	0
<b>C16:2</b>	0,38 $\pm$ 0,12	0,42 $\pm$ 0,23	0,42 $\pm$ 0,14
<b>C16:3</b>	0,04 $\pm$ 0,01	0,04 $\pm$ 0,03	0,05 $\pm$ 0,01
<b>C17:2</b>	0,02 $\pm$ 0,01	0,03 $\pm$ 0,02	0,02 $\pm$ 0,01
<b>C18:2</b>	0,21 $\pm$ 0,01	0,27 $\pm$ 0,07	0,19 $\pm$ 0,02
<b>C18:3</b>	0,16 $\pm$ 0,02	0,15 $\pm$ 0,03	0,15 $\pm$ 0,03
<b>C18:4</b>	0,15 $\pm$ 0,06	0,15 $\pm$ 0,09	0,16 $\pm$ 0,07
<b>C18:5</b>	0	0	0
<b>C19:2</b>	0,01 $\pm$ 0,00	0,01 $\pm$ 0,01	0,01 $\pm$ 0,00
<b>C19:3</b>	0	0	0
<b>C19:5</b>	0,05 $\pm$ 0,02	0,06 $\pm$ 0,03	0,05 $\pm$ 0,02
<b>C20:2</b>	0	0,01 $\pm$ 0,01	0
<b>C20:3</b>	0,08 $\pm$ 0,00	0,12 $\pm$ 0,05	0,08 $\pm$ 0,02
<b>C20:4</b>	0,97 $\pm$ 0,06	1,10 $\pm$ 0,21	0,97 $\pm$ 0,12
<b>C20:5</b>	5,72 $\pm$ 0,54	5,98 $\pm$ 1,08	5,86 $\pm$ 0,99
<b>C22:2</b>	0	0,01 $\pm$ 0,00	0

### Pigments and photosynthetic apparatus composition in intensive growth conditions.

Pigments are fundamental for photosynthesis since their content determines the amount of light that can be absorbed and efficiently exploited for growth. Mutant E2 was previously selected for a strong reduction in the Chl a content (55 % of the WT value, when analyzed in a diluted culture in limiting growth conditions (Perin et al., 2015)), while in the same screening conditions mutant I48 never showed such a pigment content reduction. Nevertheless, in the intensive growth conditions here experienced, both E2 and I48 mutant strains showed a significant reduction in both the chlorophyll a and total carotenoids content, with respect to the parental strain (Figure 3A and B – E2 shows a 24 % reduction for both and I48 shows a 11 and 14 % reduction for Chl a and total Car, respectively).

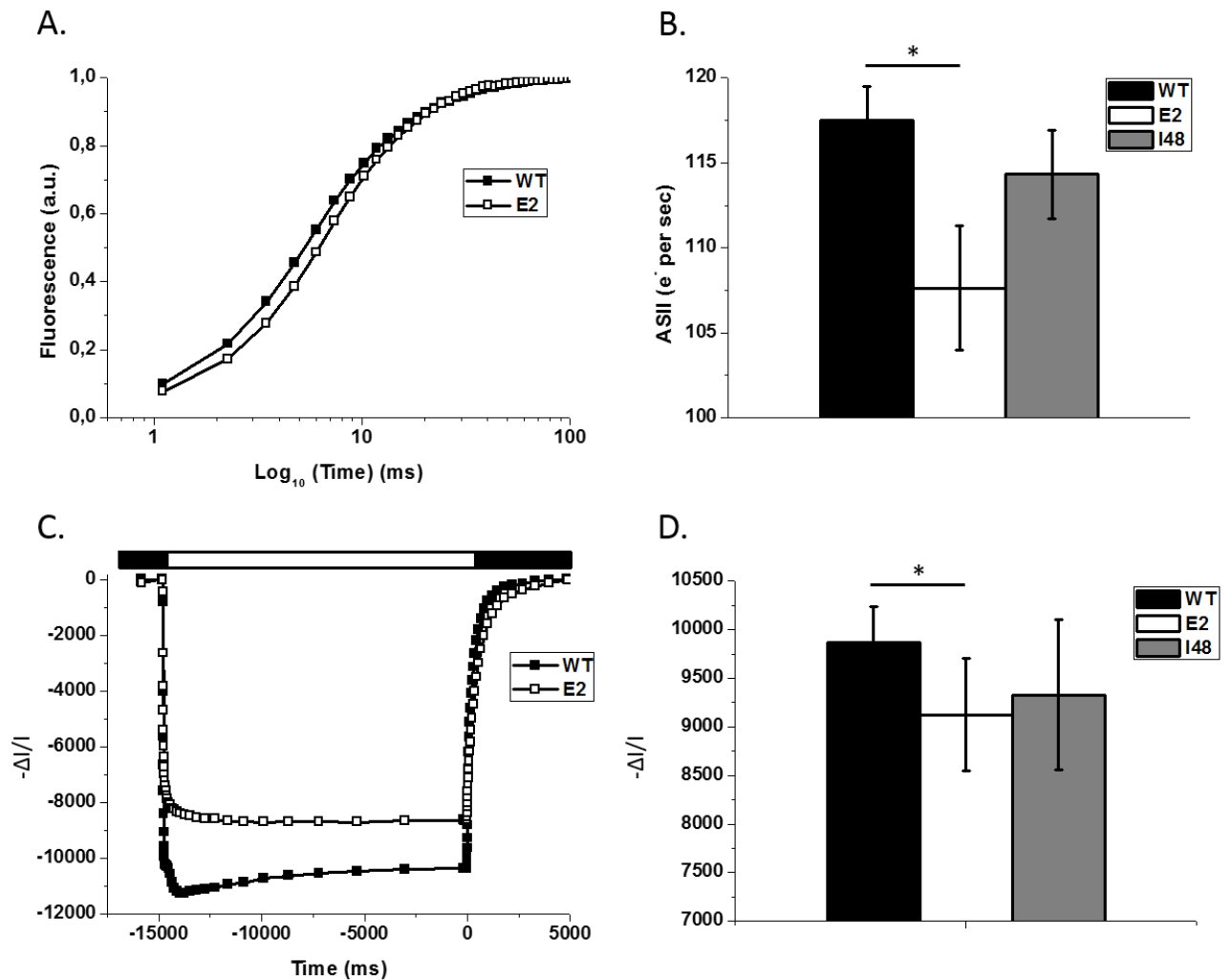


**Figure 3. Pigment content of *N. gaditana* wild type (WT) and mutant (E2 and I48) strains in fed-batch growth conditions.** A. Chlorophyll a content for *N. gaditana* WT and mutant strains (E2 and I48) in fed-batch growth conditions. B. Total carotenoids content for the same strains reported in A. C. Individual carotenoids data, expressed as mol/100 Chls, measured for the same growth conditions. violax, violaxanthin; vaucheriax, vaucheriaxanthin; anterax, anteraxanthin; zeax, zeaxanthin; cantax, cantaxanthin; beta-car, beta-carotene. The data are expressed as mean  $\pm$  SD, n = 4. Statistically significant values from wild type are marked with an asterisk (one-way ANOVA, p-value < 0,05).

Moreover, in the selected operational conditions the WT strain indeed showed a 33 % increment in Chl a content with respect to the limiting growth conditions, used in the screening step (Figure 4A of chapter II), as well as the two mutant strains, which almost double its content. These data highlight that operational conditions used for growth have a strong impact also on the cell pigments content, as expected (Sforza et al., 2015), but also that mutant strains can respond differently than the WT. HPLC analyses were run to investigate in more detail individual carotenoids composition. Violaxanthin and vaucheriaxanthin are the most abundant carotenoids in all strains. However, violaxanthin showed a 13 and 10 % reduction, for E2 and I48 respectively, with respect to the parental strain, while leaving

unaltered the other carotenoids (except for vaucheriaxanthin of I48 strain which showed a slight but significant decrease with respect to the WT - Figure 3C). This difference is interesting because violaxanthin and vaucheriaxanthin are preferentially associated to the light harvesting complexes in *Nannochloropsis gaditana* (Basso et al., 2014).

In order to investigate it further, we measured the PSII functional antenna size (ASII) from the fluorescence induction kinetics of DCMU-treated cells at 320  $\mu\text{mol}$  of photons / ( $\text{m}^2 \text{s}$ ) (Figure 4A – see methods for details). Strain E2 showed a 8 % reduction in functional PSII antenna size with respect to the WT, while strain I48 showed a slight but not significant reduction (Figure 4B). E2 strain was already previously shown to have a reduced ASII than parental strain (Perin et al., 2015). Therefore, its phenotype remains stable even when it is applied to intensive growth conditions.



**Figure 4. PSII functional antenna size and PSI content for fed-batch cultures of *N. gaditana* wild type and mutant (E2 and I48) strains.** A. Representative traces of fluorescence induction kinetics for WT and mutant E2 DCMU-treated cells. Fluorescence induction curves  $t_{2/3}$  values were used to calculate

the PSII functional antenna size, as reported in B. C. The PSI content was evaluated from the maximum absorption of  $P_{700}^+$  at 705 nm, in the presence of DCMU and DBMIB and a saturating light of 2050  $\mu\text{moles of photons} / (\text{m}^2 \text{ s})$ ;  $300 \times 10^6$  cells/ml were employed for the measurement. D. Average values for PSI content in *N. gaditana* wild type and mutant (E2 and I48) strains, grown in fed-batch cultures, are shown; white box, actinic light on; black box, actinic light off. The data are expressed as mean  $\pm$  SD,  $n = 3$ . Statistically significant values from wild type are marked with an asterisk (one-way ANOVA,  $p$ -value  $< 0,05$ ).

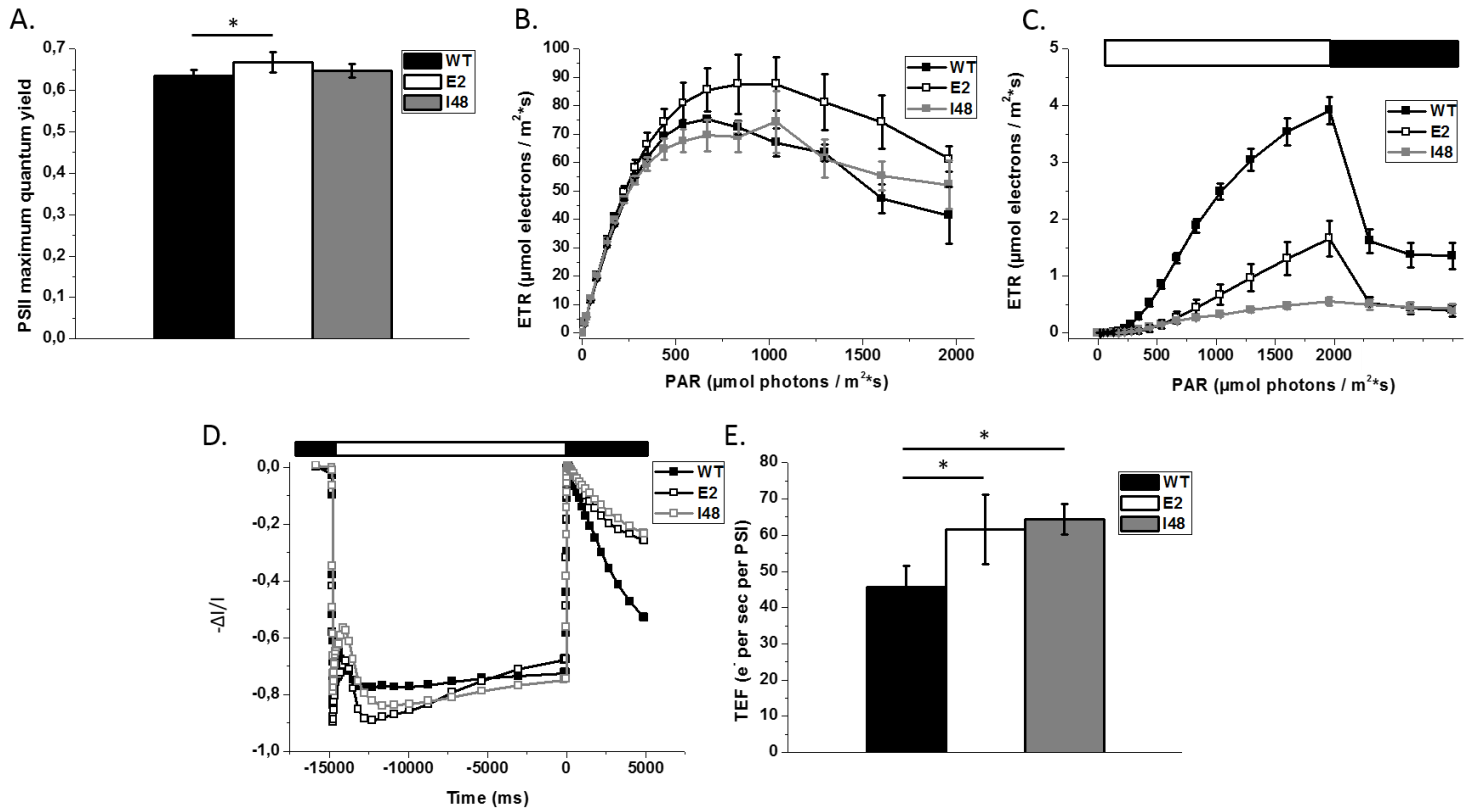
To assess whether the above described alterations in pigments content affected other photosynthetic components, also the PSI content was evaluated, after monitoring the  $P_{700}^+$  absorption signal in cells exposed to saturating light conditions, in the presence of DCMU and DBMIB to prevent PSI re-reduction (Figure 4C). The maximal  $P_{700}$  signal was exploited to quantify the PSI in all the tested strains. Mutant strain E2 showed a significant reduction in PSI content per cell with respect to the parental strain, while strain I48 showed a slight but not significant reduction (Figure 4D). Despite these data will be integrated with the further investigation on other photosynthetic components, the current data already highlighted that the mutants, when applied to intensive growth conditions, show a remodeling of some of the photosynthetic apparatus components, toward its advantageous tuning to meet the artificial growth environment needs.

*Both phenotypes show improved photosynthetic ETR in intensive growth conditions.*

The photosynthetic apparatus components remodeling so far described in these intensive growth conditions could show a strong impact on photosynthetic performances. For instance, antenna proteins play a complex biological role (Floris et al., 2013) and their content alteration may have unexpected secondary effects (Liberton et al., 2013). In fact, LHC are also involved in NPQ activation (Avenso et al., 2008; Bonente et al., 2011a; Elrad, 2002; Horton and Ruban, 2005; Peers et al., 2009) and their reduction could consequently affect photo-protection ability and might lead to increased photo-inhibition in culture external layers. Secondly, the sole reduction in the photosystem II antenna proteins could decrease the photon-use efficiency because a portion of the light-energy absorbed by the photosystem I could not be utilized in the linear photosynthetic electron transport process, if this antenna reduction is not balanced by an adjustment in photosystem II/ photosystem I ratio (Polle et al., 2000).

Therefore, to assess whether the photosynthesis components alterations bring a positive advantage to mutant strains photosynthetic performances, we monitored photosynthesis functionality exploiting *in vivo* Chl fluorescence measurements. PSII functionality was monitored through PSII

maximum quantum yield, which was calculated as  $(F_m - F_0) / F_m$ , according to (Maxwell and Johnson, 2000). All the tested strains show a high PSII maximum quantum yield (between 0.6 and 0.7 – Figure 5A), highlighting that in the selected growth conditions there is no light-induced damage to PSII.



**Figure 5. Photosynthetic performances for fed-batch cultures of *N. gaditana* wild type and mutant (E2 and I48) strains.** A. PSII maximum quantum yield as indicator of PSII functionality. B. and C. Photosynthetic electron transport rate (B) and NPQ activation in cells exposed to increasing light intensities, followed by 3 min in the dark (C). D.  $P_{700}$  redox kinetics in wild type and mutant strains cells, used for the total electron transport quantification experiment (see Methods for details). Total electron flow (TEF) was calculated from the  $P_{700}^+$  re-reduction rates after illumination. D. Average values of  $P_{700}^+$  re-reduction rates after illumination; white box, actinic light on; black box, actinic light off. The data are expressed as mean  $\pm$  SD,  $n = 5$ . Statistically significant values from wild type are marked with an asterisk (one-way ANOVA,  $p$ -value  $< 0,05$ ).

Moreover for mutant E2 this value is even significantly higher than for WT (Figure 5A), suggesting a possible improved photosynthetic efficiency in these conditions. The latter indeed showed a higher electrons transportation rate (ETR) than WT, with the ability to saturate it at higher light irradiances (Figure 5B). Mutant I48 instead didn't show any improvement in both these parameters. *In vivo* Chl fluorescence measurements provided also information on the activation of heat dissipation mechanisms (non-photochemical quenching - NPQ). Both mutant strains are affected in NPQ

activation with respect to the WT, with strain I48 showing a stronger impairment (Figure 5C). However, the higher NPQ values for the parental strain didn't completely relax during the dark recovery, highlighting a photo-inhibition event during the measurement.

During the selection steps, both mutant strains already showed to be affected in NPQ activation and this fact prove the stability of their phenotypes also when exploited in intensive growth conditions. Moreover, both mutant strains were strong impaired in NPQ relaxation kinetic, showing a strong significant reduction in qE value (Table 2).

**Table 2. Fastest and reversible component of NPQ (qE) values for strains tested in this work.** The values were calculated as the difference between the NPQ values in the last point of light exposition and in the second point of the dark recovery phase, of the NPQ kinetics of Figure 4C. The data are expressed as mean  $\pm$  SD, n = 3; all the values significantly differ from each other (one-way ANOVA, p-value < 0.05).

WT	E2	I48
2,54 $\pm$ 0,03	1,22 $\pm$ 0,21	0,1 $\pm$ 0,02

The data so far presented reflect the PSII functional activity, therefore for a comprehensive analysis of photosynthetic functionality, the evaluation of electron transport rates at the level of PSI were also performed. The total electron flow (TEF) was measured through P<sub>700</sub> re-reduction kinetics after a saturating light treatment (Figure 5D). Both mutant strains showed a higher TEF with respect to the parental strain (Figure 4E - 35 and 41 % for E2 and I48 strains, respectively). While strain E2 also showed enhanced electrons flow also at the PSII level, strain I48 instead showed the same behavior than parental strain. In addition to the difference due to the measurement method, TEF increment in I48 strain could be explained with a possible improvement in the cyclic electron flow around PSI and or / reduction of alternative electron flows, like PTOX dependent water oxidation (Chaux et al., 2015). Overall, both strains are shown to be able to improve biomass productivity in intensive growth conditions and these is reflected by an increase in photosynthetic activity, improving its efficiency through the whole culture volume. Both mutants have in common a reduction of Chl content and NPQ, suggesting the two properties are connected.

Among the two selected phenotypes, E2 strain performed better than WT in 2 out of 3 tested operational conditions (Figure 1) therefore it is likely characterized by a better compromise between the benefits of reducing light harvesting and the disadvantages of the reduced photo-protection ability, than strain I48 (Bonente et al., 2011b; Simionato et al., 2013a).

## **Conclusions.**

*Nannochloropsis gaditana* mutant strains, selected for alterations in their photosynthetic properties, perform indeed better than parental strain, when exploited in industrially relevant conditions. Mutants phenotypes are shown to respond to the environmental conditions, pushing for the necessity of their characterization in a significant environment. The theoretical benefits brought by these phenotypes are indeed reflected positively on their photosynthetic performances, and they show a higher photosynthetic activity than the parental strain. However, they are not able of performing better than WT in all tested circumstances. We conclude that photosynthesis genetic modifications must be tailored with the specifics of the PBR operational conditions and that this process is not trivial. Moreover, a better understanding of this artificial environment is seminal to drive further genetic engineering efforts.

The selected strains are indeed eligible for pilot industrial-scale cultivation tests, starting from the E2 strain which was generated through a chemical mutagenesis approach and its employment is thus not limited by the European legislation regarding the outdoor release of transgenic organisms.



## References.

- Alric, J.** (2010). Cyclic electron flow around photosystem I in unicellular green algae. *Photosynth. Res.* **106**: 47–56.
- Avenson, T.J., Ahn, T.K., Zigmantas, D., Niyogi, K.K., Li, Z., Ballottari, M., Bassi, R., and Fleming, G.R.** (2008). Zeaxanthin radical cation formation in minor light-harvesting complexes of higher plant antenna. *J. Biol. Chem.* **283**: 3550–8.
- Basso, S., Simionato, D., Gerotto, C., Segalla, A., Giacometti, G.M., and Morosinotto, T.** (2014). Characterization of the photosynthetic apparatus of the Eustigmatophycean *Nannochloropsis gaditana*: evidence of convergent evolution in the supramolecular organization of photosystem I. *Biochim. Biophys. Acta* **1837**: 306–14.
- Beckmann, J., Lehr, F., Finazzi, G., Hankamer, B., Posten, C., Wobbe, L., and Kruse, O.** (2009). Improvement of light to biomass conversion by de-regulation of light-harvesting protein translation in *Chlamydomonas reinhardtii*. *J. Biotechnol.* **142**: 70–7.
- Benvenuti, G., Bosma, R., Klok, A.J., Ji, F., Lamers, P.P., Barbosa, M.J., and Wijffels, R.H.** (2015). Microalgal triacylglycerides production in outdoor batch-operated tubular PBRs. *Biotechnol. Biofuels* **8**: 100.
- Bonente, G., Ballottari, M., Truong, T.B., Morosinotto, T., Ahn, T.K., Fleming, G.R., Niyogi, K.K., and Bassi, R.** (2011a). Analysis of LHCSR3, a protein essential for feedback de-excitation in the green alga *Chlamydomonas reinhardtii*. *PLoS Biol.* **9**: e1000577.
- Bonente, G., Formighieri, C., Mantelli, M., Catalanotti, C., Giuliano, G., Morosinotto, T., and Bassi, R.** (2011b). Mutagenesis and phenotypic selection as a strategy toward domestication of *Chlamydomonas reinhardtii* strains for improved performance in photobioreactors. *Photosynth. Res.* **108**: 107–20.
- Cazzaniga, S., Dall’Osto, L., Szaub, J., Scibilia, L., Ballottari, M., Purton, S., and Bassi, R.** (2014). Domestication of the green alga *Chlorella sorokiniana*: reduction of antenna size improves light-use efficiency in a photobioreactor. *Biotechnol. Biofuels* **7**: 157.
- Chaux, F., Peltier, G., and Johnson, X.** (2015). A security network in PSI photoprotection: regulation of photosynthetic control, NPQ and O<sub>2</sub> photoreduction by cyclic electron flow. *Front. Plant Sci.* **6**: 875.
- Elrad, D.** (2002). A Major Light-Harvesting Polypeptide of Photosystem II Functions in Thermal Dissipation. *PLANT CELL ONLINE* **14**: 1801–1816.

- Färber, A. and Jahns, P.** (1998). The xanthophyll cycle of higher plants: influence of antenna size and membrane organization. *Biochim. Biophys. Acta - Bioenerg.* **1363**: 47–58.
- Floris, M., Bassi, R., Robaglia, C., Alboresi, A., and Lanet, E.** (2013). Post-transcriptional control of light-harvesting genes expression under light stress. *Plant Mol. Biol.* **82**: 147–54.
- Formighieri, C., Franck, F., and Bassi, R.** (2012). Regulation of the pigment optical density of an algal cell: filling the gap between photosynthetic productivity in the laboratory and in mass culture. *J. Biotechnol.* **162**: 115–23.
- Horton, P. and Ruban, A.** (2005). Molecular design of the photosystem II light-harvesting antenna: photosynthesis and photoprotection. *J. Exp. Bot.* **56**: 365–73.
- Huesemann, M.H., Hausmann, T.S., Bartha, R., Aksoy, M., Weissman, J.C., and Benemann, J.R.** (2009). Biomass productivities in wild type and pigment mutant of *Cyclotella* sp. (Diatom). *Appl. Biochem. Biotechnol.* **157**: 507–26.
- Jeffrey, S.W., Mantoura, R.F.C., and Wright, S.W.** (1997). Phytoplankton pigments in oceanography: guidelines to modern methods. *Monogr. Oceanogr. Methodol.*
- Lea-Smith, D.J., Bombelli, P., Dennis, J.S., Scott, S.A., Smith, A.G., and Howe, C.J.** (2014). Phycobilisome-Deficient Strains of *Synechocystis* sp. PCC 6803 Have Reduced Size and Require Carbon-Limiting Conditions to Exhibit Enhanced Productivity. *Plant Physiol.* **165**: 705–714.
- Liberton, M., Collins, A.M., Page, L.E., O'Dell, W.B., O'Neill, H., Urban, V.S., Timlin, J.A., and Pakrasi, H.B.** (2013). Probing the consequences of antenna modification in cyanobacteria. *Photosynth. Res.* **118**: 17–24.
- Matyash, V., Liebisch, G., Kurzchalia, T. V, Shevchenko, A., and Schwudke, D.** (2008). Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *J. Lipid Res.* **49**: 1137–46.
- Maxwell, K. and Johnson, G.N.** (2000). Chlorophyll fluorescence - A practical guide. *J. Exp. Bot.* **51**: 659–668.
- Melis, A., Neidhardt, J., and Benemann, J.R.** (1998). *Dunaliella salina* (Chlorophyta) with small chlorophyll antenna sizes exhibit higher photosynthetic productivities and photon use efficiencies than normally pigmented cells. *J. Appl. Phycol.* **10**: 515–525.
- de Mooij, T., Janssen, M., Cerezo-Chinarro, O., Mussnug, J.H., Kruse, O., Ballottari, M., Bassi, R., Bujaldon, S., Wollman, F.-A., and Wijffels, R.H.** (2014). Antenna size reduction as a strategy to increase biomass productivity: a great potential not yet realized. *J. Appl. Phycol.*
- Moran, R. and Porath, D.** (1980). Chlorophyll determination in intact tissues using *n,n*-dimethylformamide. *Plant Physiol.* **65**: 478–9.

- Nakajima, Y., Tsuzuki, M., and Ueda, R.** (2001). Improved productivity by reduction of the content of light-harvesting pigment in *Chlamydomonas perigranulata*. *J. Appl. Phycol.* **13**: 95–101.
- Nakajima, Y. and Ueda, R.** (2000). The effect of reducing light-harvesting pigment on marine microalgal productivity. In *Journal of Applied Phycology*, pp. 285–290.
- Ort, D.R. and Melis, A.** (2011). Optimizing antenna size to maximize photosynthetic efficiency. *Plant Physiol.* **155**: 79–85.
- Peers, G., Truong, T.B., Ostendorf, E., Busch, A., Elrad, D., Grossman, A.R., Hippler, M., and Niyogi, K.K.** (2009). An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* **462**: 518–21.
- Perin, G., Bellan, A., Segalla, A., Meneghesso, A., Alboresi, A., and Morosinotto, T.** (2015). Generation of random mutants to improve light-use efficiency of *Nannochloropsis gaditana* cultures for biofuel production. *Biotechnol. Biofuels* **8**: 161.
- Polle, J.E., Benemann, J.R., Tanaka, A., and Melis, A.** (2000). Photosynthetic apparatus organization and function in the wild type and a chlorophyll b-less mutant of *Chlamydomonas reinhardtii*. Dependence on carbon source. *Planta* **211**: 335–44.
- Porra, R.J., Thompson, W.A., and Kriedemann, P.E.** (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta - Bioenerg.* **975**: 384–394.
- Radakovits, R., Jinkerson, R.E., Darzins, A., and Posewitz, M.C.** (2010). Genetic engineering of algae for enhanced biofuel production. *Eukaryot. Cell* **9**: 486–501.
- Sforza, E., Bertucco, A., Morosinotto, T., and Giacometti, G.M.** (2012a). Photobioreactors for microalgal growth and oil production with *Nannochloropsis salina*: From lab-scale experiments to large-scale design. *Chem. Eng. Res. Des.* **90**: 1151–1158.
- Sforza, E., Calvaruso, C., Meneghesso, A., Morosinotto, T., and Bertucco, A.** (2015). Effect of specific light supply rate on photosynthetic efficiency of *Nannochloropsis salina* in a continuous flat plate photobioreactor. *Appl. Microbiol. Biotechnol.* **99**: 8309–18.
- Sforza, E., Simionato, D., Giacometti, G.M., Bertucco, A., and Morosinotto, T.** (2012b). Adjusted light and dark cycles can optimize photosynthetic efficiency in algae growing in photobioreactors. *PLoS One* **7**: e38975.
- Simionato, D., Basso, S., Giacometti, G.M., and Morosinotto, T.** (2013a). Optimization of light use efficiency for biofuels production in algae. *Biophys. Chem.*

- Simionato, D., Block, M.A., La Rocca, N., Jouhet, J., Maréchal, E., Finazzi, G., and Morosinotto, T.** (2013b). The response of *Nannochloropsis gaditana* to nitrogen starvation includes de novo biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. *Eukaryot. Cell* **12**: 665–76.
- Wellburn, A.R.** (1994). The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *J. Plant Physiol.* **144**: 307–313.
- Wobbe, L., Blifernéz, O., Schwarz, C., Mussgnug, J.H., Nickelsen, J., and Kruse, O.** (2009). Cysteine modification of a specific repressor protein controls the translational status of nucleus-encoded LHClI mRNAs in *Chlamydomonas*. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 13290–5.
- Wobbe, L. and Remacle, C.** (2014). Improving the sunlight-to-biomass conversion efficiency in microalgal biofactories. *J. Biotechnol.*



# CHAPTER IV

## **Forward Genetics in *Nannochloropsis gaditana* Reveal a Splicing Variant of Protein Lhcx1 as Likely Responsible of a Mutant NPQ Inactivation**

### **Author names and affiliations**

Giorgio Perin<sup>1</sup>, Alessandra Bellan<sup>1</sup>, Alessandro Alboresi<sup>1</sup> and  
Tomas Morosinotto<sup>1</sup>

<sup>1</sup>PAR-Lab\_Padua Algae Research Laboratory, Department of Biology,  
University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy

**Abstract.**

In this chapter we describe the investigation of the genetic basis responsible for the phenotype of *Nannochloropsis gaditana* insertional mutant strain I48. Since TAIL-PCR failed to provide a clear result on the resistance cassette insertion *locus*, the genome of the strain I48 was fully re-sequenced and the identified mutations were mapped on the reference genome. We identified the gene Naga\_100173g12 (coding for protein LHCX1) as carrying a deletion of the second nucleotide of the 5'-donor splicing site of the 4<sup>th</sup> intron (GT → G). Through RT-PCR we confirmed the retention of the 4<sup>th</sup> intron of the gene in the mature LHCX1 mRNA, which led to a truncated protein. We validated this result using western blot analysis on *N. gaditana* isolated thylakoids. Sucrose gradients, used to separate the photosynthetic apparatus components of WT and I48 strains, showed the presence of the LHCX1 protein also in association with PSI in WT. The application of the latter strain as molecular tool to investigate further the biological implications of this association is here discussed.

## Introduction.

Forward genetics approaches concern the generation of random mutations in a target genome, and the consequent investigation of the genetic basis responsible of the selected phenotypes.

These strategies, applied to photosynthetic eukaryotes, would open doors to investigate the molecular bases of the biological networks involved in photosynthesis regulation, especially in emerging species. The mutagenesis event (s) can be induced through the random insertion of a DNA cassette, conferring resistance to an antibiotic, or through chemical (e.g. ethyl methane sulfonate - EMS) and physical agents (e.g. UV radiation), having a strong effect on the DNA integrity.

The identification of the genomic *locus* of insertion or of the genetic variants, cause of a target phenotype, is a very complex and time-consuming process. In this context, an insertional mutagenesis approach is preferable since several molecular techniques have been developed in the past years to rescue the insertion genomic *locus* (e.g. PCR-based techniques, Inverse PCR (Ochman et al., 1988), thermal asymmetric interlaced (TAIL)-PCR (Dent et al., 2005; Liu et al., 1995), single nucleotide nested (SON)-PCR (Antal et al., 2004), restriction enzyme site-directed amplification (RESDA)-PCR (González-Ballester et al., 2005), Mmel-based strategy (Zhang et al., 2014) and plasmid rescue (Gumpel and Purton, 1994) which isn't instead based on PCR). Through their application the success in rescuing the flanking regions of insertion isn't however fully guaranteed, and several studies described in literature indeed showed positive (Polle et al., 2003; Kirst et al., 2012) and negative (Bonente et al., 2011) results. The situation strongly get worse when chemical or physical mutagenesis events are exploited, since such above described molecular techniques aren't applicable. In this case, the most recent advanced approaches are essentially based on two strategies: meiotic mapping by bulked segregant analysis in combination with a whole genome-sequencing or unlinked mutation removal by backcrossing of the mutant to the WT strain, prior to sequencing. However, if the target species don't have a described reproductive life cycle (e.g. for the recently emerging species such as those belonging to the *Nannochloropsis* genus) the strategies based on crossing steps aren't exploitable.

Recently, a new highly time-saving technique was developed by (Schierenbeck et al., 2015) thanks to the direct sequencing of the mutant genomes and the subsequent subtraction of common variants between the mutants and the WT strain. This procedure omits crossing steps and saves a lot of time and efforts needed for genetic variants identification.

In this chapter we described the procedure followed to identify the genetic basis of the photosynthetic phenotype of *N. gaditana* mutant strain I48, exploiting its whole genome re-sequencing and the



subtraction of common variants with the WT strain, thanks to the availability of a reference information (Corteggiani Carpinelli et al., 2014). This procedure led us to identify a single point mutation in the Naga\_100173g12 gene, coding for protein LHCX1. Validation of the mutation effect at the protein level was provided using western blot analysis. The impact of this mutation on the photosynthetic apparatus organization of *N. gaditana* mutant strain I48 will be here presented and discussed.

## Materials and Methods.

### *Microalgae growth.*

*Nannochloropsis gaditana*, strain 849/5, from the Culture Collection of Algae and Protozoa (CCAP) were used as the WT strain, while the I48 mutant strain was isolated from a random insertional mutagenesis approach, performed as previously described (Perin et al., 2015). The latter contains a resistance cassette able to provide resistance to zeocin. Both strains were maintained in F/2 solid media, with sea salts (32 g/L, Sigma Aldrich), 40 mM Tris-HCl (pH 8), Guillard's (F/2) marine water enrichment solution (Sigma Aldrich) and 1% agar (Duchefa Biochemie), supplied with 3.5 µg/ml zeocin for strain I48. Cells were pre-cultured in sterile F/2 liquid media in Erlenmeyer flasks with 100 µmol photons m<sup>-2</sup> s<sup>-1</sup> illumination and 100 rpm agitation at 22 ± 1 °C in a growth chamber. The same growth conditions were used for genomic DNA extraction. The material to perform the molecular and biochemical analysis was collected from fed-batch cultures. The latter were performed in Drechsel's bottles, bubbled using air enriched with 5% CO<sub>2</sub> (v/v) (for experimental apparatus set-up see chapter III); F/2 media was enriched with added nitrogen, phosphate and iron sources (0.75 g/L NaNO<sub>3</sub>, 0.05 g/L NaH<sub>2</sub>PO<sub>4</sub> and 0.0063 g/L FeCl<sub>3</sub>•6H<sub>2</sub>O final concentrations) to speed up *Nannochloropsis gaditana* growth. Constant illumination at 400 µE m<sup>-2</sup> s<sup>-1</sup> was provided by daylight fluorescent lamps.

### *DNA extraction, TAIL-PCR and whole genome re-sequencing.*

*N. gaditana* strain I48 genomic DNA was extracted from a 4-day-old culture, performed in Erlenmeyer flasks using F/2 liquid media. Cells were lysed using a Mini Bead Beater (Biospec Products) at 3500 RPM for 20 s in the presence of glass beads (150–212 µm diameter). Genomic DNA was then purified using the *EUROGOLD™ Plant DNA Mini Kit* (Euroclone), applying minor modifications. DNA concentration and purity was determined by 100 UV–VIS spectrophotometer (Cary Series, Agilent Technologies).

Thermal asymmetric interlaced (TAIL)-PCR utilizes nested-specific primers in successive reactions together with a shorter arbitrary degenerate primer, so that the relative amplification efficiencies of specific and non-specific products can be thermally controlled (Dent et al., 2005; Liu et al., 1995). The nested-specific primers used in the present study are reported in table 1 (#1, #2 and #3 sequences). The latter were designed to be specific for the 3'-end of the gene conferring resistance to zeocin (Perin et al., 2015). The short degenerate oligonucleotides chosen are called RMD227 and

RMD228 (#4 and #5 sequences in table 1) and they were already successfully used by (Dent et al., 2005). Details of the PCR reaction mix and thermal cycles have been described in (Dent et al., 2005).

**Table 1. Primer sequences used to perform TAIL-PCR and RT-PCR reactions.**

Primer #	Name	Primer Sequence (5' – 3')
1	Primer I	GAGATCGGCGAGCAGCCGTG
2	Primer II	GCCCTGCGCGACCCGGCCGG
3	Primer III	GTGGCCGAGGAGCAGGACTG
4	RMD227	NTCGWGWTSNAGC
5	RMD228	WGNTCWGNCANGCG
6	LHCX1_FOR	CATCATCACCACCATCACCGTGTACTCTCTTTCTTGCTATTA
7	LHCX1_REV	GTGGCGGCCGCTCTATTATTAGAAGAAGAAGTTGTAGACGTCC

The whole genome re-sequencing for mutant strain I48 was performed by IGA Technologies Services (Udine, Italy), with the Illumina MiSeq platform, using 100bp pair-end sequences leading to a 70-fold genome coverage. The *N. gaditana* B-31 genome published in (Corteggiani Carpinelli et al., 2014), was used as reference. Genetic variants annotation and their prediction effects was carried out using the SnpEff software (Cingolani et al.).

#### *Total RNA extraction and cDNA preparation.*

Total RNA was extracted from the *N. gaditana* fed-batch cultures. Cells were lysed using a Mini Bead Beater (Biospec Products) at 3500 RPM for 20 s in the presence of glass beads (150–212 µm diameter). Total RNA was thus purified using the *RNeasy™ Plant Mini Kit* (Qiagen), applying minor modifications to the manufacturer's instruction. Total RNA concentration and purity was determined by 100 UV–VIS spectrophotometer (Cary Series, Agilent Technologies).

cDNA was prepared from 2 µg of total RNA-template with the RevertAid Reverse Transcriptase cDNA kit (Thermo Fisher Scientific, Epsom, UK).

#### *RT-PCR.*

The cDNA was used as template for the RT-PCR reactions, in order to amplify the LHCX1 (GENE ID: Naga\_100173g12) gene from the WT and I48 strains templates. The primer sequences reported in table 1 (sequences #6 and #7) and KAPA HIFI high-fidelity DNA polymerase (KAPA Biosystems) were used for this purpose.

### *Sequence analysis.*

*Multiple sequence alignments and three-dimensional modeling.* The light-harvesting protein sequences from *N. gaditana* were retrieved from [www.nannochloropsis.org](http://www.nannochloropsis.org) (Corteggiani Carpinelli et al., 2014). Multiple sequence alignments of the chlorophyll binding proteins were generated using the Clustal W platform integrated in BioEdit 7.2.5 (Hall, 1999). They were manually refined using the latter, on the basis of the crystal structure of the chlorophyll a/b binding protein from spinach (PDB code 1RWT, (Liu et al., 2004)). The structure of the 1RWT protein from the PDB was used to generate a three dimensional (3D) model of the LHCX1 (GENE ID: Naga\_100173g12) protein (from 46 to 231 aa), using the Phyre2 software (Kelley et al., 2015), with default parameters. The multiple sequence alignment and the 3D model were displayed using ESPript (Robert and Gouet, 2014) and PyMOL (DeLano, 2005), respectively.

### *Thylakoids isolation*

*N. gaditana* cells were collected from the fed-batch cultures. They were harvested by 10 minutes of centrifugation at 4000 x g and washed twice in B1 buffer (0.4 M NaCl, 2 mM MgCl<sub>2</sub>, and 20 mM Tricine-KOH [pH 7.8]). After centrifugation, glass beads (diameter of 150-212 µm) and B1 buffer with 0.5 % milk powder and 1 mM PMSF, 1 mM DNP-ε-amino-n-caproic acid and 1 mM benzamidine were added. Cells were then disrupted using a Mini Bead Beater (Biospec Products) for 20s at 3500 RPM. B1 buffer containing milk and proteases inhibitors, as previously described, was added and the pellet resuspended. Unbroken cells were then separated by a centrifugation step (2500 x g, 15 minutes), and the supernatant was collected. The latter contained the thylakoids and was centrifuged at 15000 x g for 20 min, and the pellet was washed twice with B2 buffer (0.15 M NaCl, 5 mM MgCl<sub>2</sub>, and 20 mM Tricine-KOH [pH 7.8]). The thylakoids were resuspended in B4 buffer (0.4 M sorbitol, 15 mM NaCl, 5 mM MgCl<sub>2</sub> and 10 mM HEPES-KOH [pH 7.5]) and immediately frozen in liquid nitrogen and stored at -80 °C until use. All steps were performed at 4 °C and in dim light. The total pigments were extracted with 80 % acetone, and the chlorophyll a concentrations were determined spectrophotometrically using specific extinction coefficients (Porra et al., 1989) and the acetone spectra fitting previously described (Croce et al., 2002), which were modified to account for the *N. gaditana* peculiar pigment content.

*Thylakoids solubilization and sucrose gradients.*

Thylakoid membranes corresponding to 500 µg of Chl were washed with 50 mM EDTA and then solubilized for 20 minutes on ice in 1 ml of final 0.4 % n-Dodecyl α-D-maltoside (α-DM) and 10 mM HEPES (pH7.5), after vortexing for 1min. The solubilized samples were centrifuged at 15000 x g for 20 min to eliminate any unsolubilized material, and the supernatant with the photosynthetic complexes was then fractionated by ultracentrifugation in a 0.1 – 1 M sucrose gradient containing 0.06 % α-DM and 10 mM HEPES (pH7.5) (280000 x g, 18 hours, 4°C).

*SDS-page and western blot.*

A 12% SDS-page analysis for both the sucrose gradient fractions and thylakoids extracts was performed using a TRIS-glycine buffer system, as previously described (Laemmli, 1970). The samples were solubilized for 20 minutes at RT in 10 % glycerol, 45 mM TRIS (pH 6.8), 0.03 M dithiothreitol and 3 % SDS. Western blot analyses were performed after transferring the proteins to nitrocellulose (Bio Trace, Pall Corporation), using alkaline phosphatase conjugated secondary antibodies. The antibody against D2 [UniProt: P06005 for D2] was generated by immunizing New Zealand rabbits with the spinach protein, whereas the recombinant proteins were used for the antibodies against VCP and LHCX1, which were obtained by cloning the two cDNA into pETite N-HIS (Lucigen-Expresso T7 Cloning and Expression System), expressing the proteins in *Escherichia coli* BL21 (DE3, Invitrogen), and purifying as inclusion bodies [UniProt: W7T4V5 for LHCf1, and Uniprot: K8YWB4 for LHCX1].

## Results and discussion.

In chapter II, random mutants of *Nannochloropsis gaditana* were generated and further selected for alterations in their photosynthetic apparatus regulation and composition. Among them, mutant strain I48 showed a strong reduction in NPQ mechanism activation. In chapter III its phenotype was shown to be stable in intensive growth conditions, leading to an improved electrons transportation rate and a consequent higher biomass productivity. In this chapter we investigated the genetic basis of its phenotype, to dissect the molecular bases involved in photosynthesis regulation in this organism.

*Whole genome re-sequencing of strain I48 revealed a splicing-donor variant in gene Naga\_100173g12.* *Nannochloropsis gaditana* mutant strain I48 was generated after a mutagenesis approach in which a transformation cassette (see chapter II for details on the molecular elements of the cassette), conferring resistance to zeocin, was randomly inserted in *Nannochloropsis gaditana* strain CCAP 849/5 parental genome (Perin et al., 2015). It was selected for a strong reduction in the activation of the mechanism of NPQ (see figure 5C of chapter III). Mutant strain I48 was kept in F/2 solid medium for 2 years (~ 240 generations), alternating 5 generations in selective medium and 5 without selection, constantly monitoring its photosynthetic phenotype and the growth. Its phenotype indeed turned out to be stable over time, as well as its resistance to zeocin (data not shown). Thermal asymmetric interlaced (TAIL)-PCR was performed, according to (Dent et al., 2005; Liu et al., 1995), to rescue the genomic *locus* regions flanking the insertion site of the zeocin resistance cassette. However, this technique failed to identify the genomic *locus* of insertion and the obtained results weren't able to provide a clear insight on the insertion event (s).

The mutant strain was therefore subjected to the whole genome re-sequencing for the systematic identification of the mutation (s), responsible for its photosynthetic phenotype. The latter was performed on an *Illumina MiSeq* to obtain a theoretical 70-fold genome coverage (see Methods for details). The software SnpEff (Cingolani et al.) was then run to annotate the genetic variants identified in strain I48 on the *N. gaditana* reference genome (Corteggiani Carpinelli et al., 2014) and later predict the effects of these genetic variants on annotated genes (such as deletions and insertions leading to frame shifting or amino acid changes, etc.). Given its photosynthetic phenotype, among the identified genetic variants, we focused on those affecting those genes involved in the photosynthetic process. Among them, the SnpEff software identified only in position 1510841 bp of chromosome 01 a high impact genetic variant, falling in the genomic *locus* occupied by the *Naga\_100173g12* gene. The latter

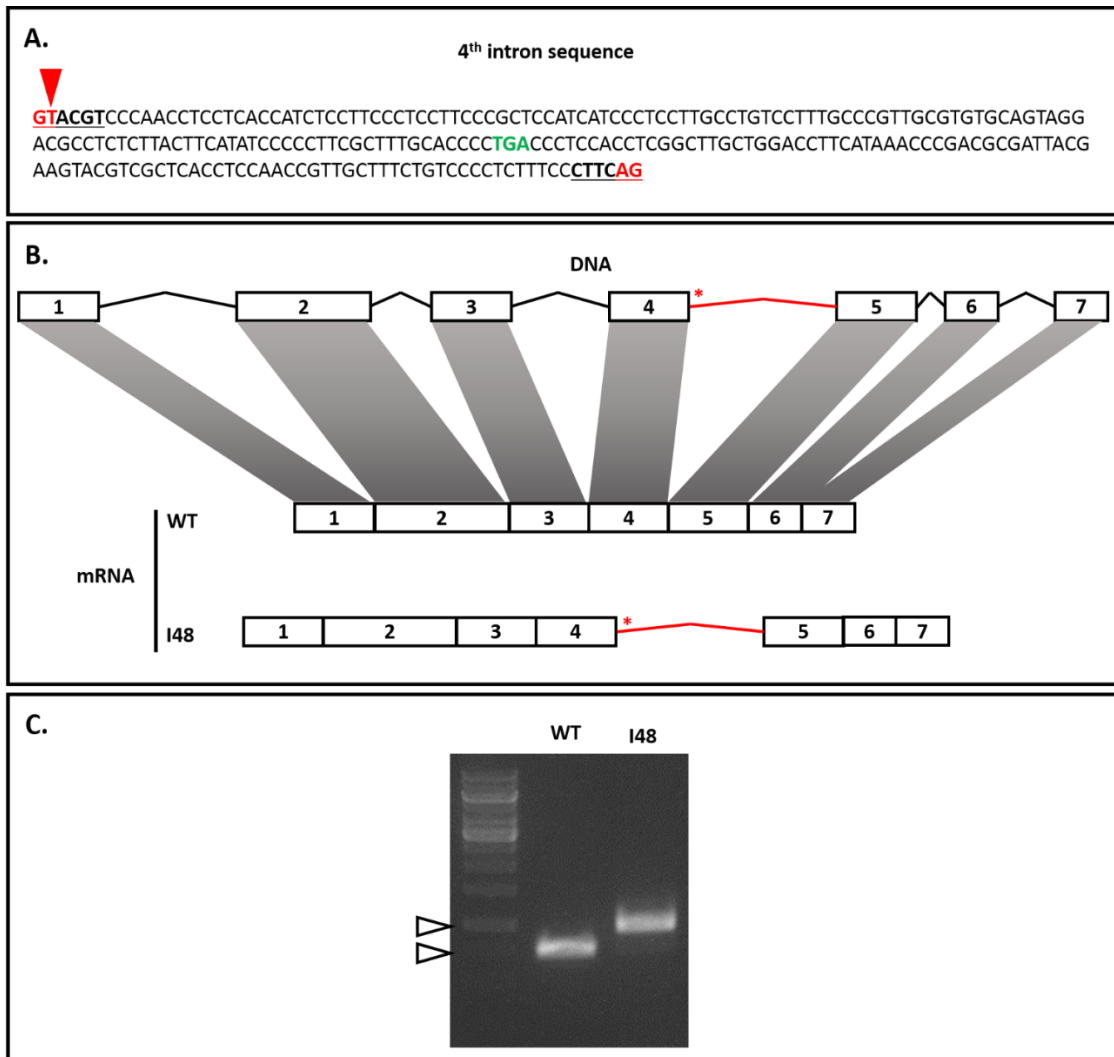
encodes for the LHCX1 protein, a light-harvesting complex (LHC) protein belonging to the sub-clade of LHCX, which are well known to be stress-related proteins, responsible for the non-photochemical quenching (NPQ) process activation in excess light conditions (Niyogi and Truong, 2013).

The Naga\_100173g12 gene ATG is localized in 1509688 bp position of the chromosome 01, therefore the identified high impact genetic variant fell 1153 bp from the gene translational start site.

This means that the latter didn't fall inside the gene coding sequence, as annotated in the *N. gaditana* genome browser ([www.nannochloropsis.org](http://www.nannochloropsis.org)), but in the 4<sup>th</sup> intron sequence, by altering its 5'-donor splicing site. The identified genetic variant involves indeed a deletion of a nucleotide, leading to change the first two nucleotides of the 5'-donor splicing site of the 4<sup>th</sup> intron from a GT to a G (Figure 1A). After the discovery of the splicing event, it was observed that all intron sequences contain two highly conserved dinucleotides, a GT at 5'-end and an AG at the 3'-end (Bursset et al., 2000; Mount and Steitz, 1981), which are responsible for the efficient binding of the spliceosome (Papasaikas and Valcárcel, 2015). With the accumulation of the genomic sequences data, in the past years it was concluded that this "GT-AG rule", was always obeyed, to induce an efficient splicing event (Bursset et al., 2000). Therefore the identified deletion of the thymine of the dinucleotide GT of the 5'-donor splicing site could have a huge effect on the efficiency of the splicing event for the 4<sup>th</sup> intron of this gene, leading to the possible retention of its sequence in the mature mRNA of the Naga\_100173g12 gene (Figure 1B).

*The sequence of the 4<sup>th</sup> intron of the Naga\_100173g12 gene is indeed retained in the mature mRNA.*

To verify the hypothesis of the retention of the 4<sup>th</sup> intron sequence of the Naga\_100173g12 gene in the mature mRNA, we performed a RT-PCR reaction to amplify the LHCX1 coding sequence from a cDNA template (see methods for details). The RT-PCR performed on the sample coming from the strain I48 showed a single but higher band, than that observed for the parental strain (Figure 1C). While the band obtained from the parental strain sample was around 700 bp, the former was instead close to 1000 bp, and therefore compatible with the retention of the 4<sup>th</sup> intron sequence (231 bp). Both bands were sequenced and, if on the one hand the band coming from the parental strain template showed the expected sequence, the second one, coming from the I48 strain, confirmed instead the retention of the 4<sup>th</sup> intron, according to the scheme of figure 1B. The latter also showed the deletion of the thymine in the 5'-donor splicing site of the retained sequence, as previously observed by the whole genome re-sequencing.



**Figure1. Retention of the 4<sup>th</sup> intron in the sequence of the mature mRNA of the LHCX1 protein in the I48 *Nannochloropsis gaditana* strain.** A. In the 4<sup>th</sup> intron sequence the 5'-donor and 3'-acceptor splicing sites are underlined and shown in bold. The most conserved splicing sequences in eukaryotes (GT, 5'-site and AG, 3'-site) are highlighted in red. The thymine missing in the version of the Naga\_100173g12 gene found in strain I48 is highlighted with a red arrow. The new stop codon identified in the intron sequence is highlighted in green. B. The Naga\_100173g12 gene includes 7 exons (rectangles) and 6 introns (black and red lines), according to the annotation published in (Corteggiani Carpinelli et al., 2014). Scheme of the splicing event that leads to obtain the mRNA of the LHCX1 protein, both in the WT and in the I48 strain. The localization of the point mutation found in the Naga\_100173g12 gene in the I48 strain is highlighted with a red asterisks. The exons are indicated with sequential numbers from 1 to 7. C. Agarose gel in which the RT-PCR products coming from the amplification of the LHCX1 cDNA for the WT and I48 protein versions were loaded. The two arrows indicate 750 (bottom) and 1000 bp (top), respectively.



The retention of the 4<sup>th</sup> intron modifies the LHCX1 protein sequence, consequently affecting its accumulation.

The retention of the sequence of the 4<sup>th</sup> intron of the Naga\_100173g12 gene in the mature mRNA changes the amino acids sequence of the LHCX1 protein in the strain I48. The LHCX1 proteins found in the parental and I48 mutated strain are identical for the first 150 aa, as a consequence of the shared nucleic acids sequence for the first 450 bp, while the remaining part of the proteins is instead different (Figure 2).

Protein\_LHCX1\_WT (231 aa)

MRVLSFLAIIGTAAAFVKPTLPAAGSRTRAGALRMNLAIELEAGKTSPPFDGFDPLGLSKDKSFKELKKWREAELKHGRVAM  
α1  
α2  
LAVLGTAVQENFHLWGFNEKEMDGAIFHFQEIQNVYPLFWTALLFIIGIIEARTISTGW DENMAGSSQIAGVKEDYICGNLG  
α3  
 LDPLKIIENDDEEAFLSYRNKELNNGRLAMIAAAGITVQEKFVTNGLPEFEFHRFALSDVYNFFF

Protein\_LHCX1\_I48 (193 aa)

MRVLSFLAIIGTAAAFVKPTLPAAGSRTRAGALRMNLAIELEAGKTSPPFDGFDPLGLSKDKSFKELKKWREAELKHGRVAM  
α1  
α2  
LAVLGTAVQENFHLWGFNEKEMDGAIFHFQEIQNVYPLFWTALLFIIGIIEARTISTGW DENMAGSDVPTSSPSPSLLPAPSS  
 LLACPLPVACAVGRLSYFISPFALHP

**Figure 2. LHCX1 protein sequences as found in the WT and I48 *Nannochloropsis gaditana* strains (from the N to the C-terminal).** The amino acids sequences are identical for the first 150 aa from the N-terminal (underlined and highlighted in bold), while the remaining part of the sequence is different. The residues necessary for the binding of the chlorophyll a molecules are highlighted in red. The sequences belonging to the three membrane-spanning  $\alpha$ -helices are highlighted in green.

In fact, the presence of the 4<sup>th</sup> intron sequence is responsible, starting from the amino acid number 151, for the synthesis of a new amino acids sequence in the strain I48, which is however truncated prematurely, due to the presence of a new stop codon in the retained intron sequence (Figure 1A). When a prediction model of the LHCX1 protein 3D structure was investigated (see methods for details), we showed indeed the loss of the third membrane-spanning alpha helix, in the mutated version of the protein, found in the I48 strain (Figure 2 and Figure 3A). Since the three membrane-spanning alpha helices form a common structural motif shared among all the LHC proteins (Green and

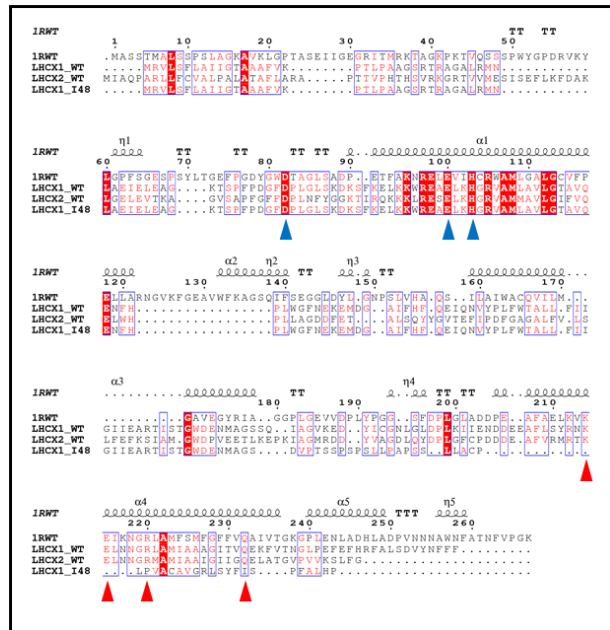
Pichersky, 1994), the absence of one of them could have a strong impact on the mutated protein 3D structure.

Moreover, since involved in the high light response, LHCX proteins include a series of amino acids residues involved in chlorophyll molecules coordination (Dal'Osto et al., 2015). In the wild type form of the LHCX1 protein, 7 residues involved in the chlorophyll a binding were identified (Figure 2), according to (Dittami et al., 2010). In the mutated LHCX1 form, the retention of the intron sequence caused the loss of 4 out of 7 of these chlorophyll a coordination sites, as shown in figure 2 and figure 3B. Since the coordination of the chlorophyll a molecules is also seminal to achieve a stable conformational structure in the LHC proteins (Dal'Osto et al., 2015; Hooper et al.), we concluded that the substitution of the 22 % of the protein sequence and its consequent truncation induced also a structural change of the protein, affecting its stability, likely directing the latter toward degradation.

A.



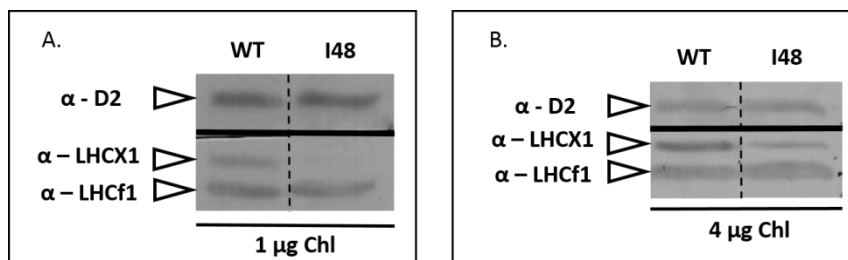
B.



**Figure 3. Predicted structural 3D model for the LHCX1 protein.** In A. the predicted structural 3D model for the LHCX1 protein (from the 46 to the to the 231 aa) is shown in a stereo ribbon representation. In blue are depicted the conserved amino acids involved in chlorophyll a binding. In red is depicted the region between the 151 and the 231 aa of the WT protein which misses in the mutated protein version, found in the I48 strain. B. Structure-based sequence alignment of the crystallized chlorophyll a/b binding protein from spinach (PDB code: 1RWT) with the LHCX1\_WT and I48 version and the LHCX2 proteins from *Nannochloropsis gaditana*. The secondary structure of the spinach protein is shown above the alignment. Identical amino acids are highlighted by a red

background while in red letters are depicted those similar. Alpha helices are represented as helices and  $\beta$ -turns are marked with TT. Blue triangles indicate the conserved residues involved in the binding of chlorophyll a molecules, still present in the LHCX1\_I48 mutated version, while red triangles indicate those missing in the latter.

This hypothesis was validated with a western blot analysis on *Nannochloropsis gaditana* thylakoids (see methods for details). Thylakoids from both the parental and I48 strain were loaded on the same gel and then detected with an antibody specific for the parental LHCX1 protein. The second major LHC protein of the *Nannochloropsis gaditana* photosynthetic apparatus, LHCf1, was detected on the same membrane, as well as the PSII core subunit D2 protein, used as a control. As shown in figure 4A, the protein LHCX1 is not detectable in strain I48 thylakoids, if samples corresponding to 1  $\mu\text{g}$  of Chl were loaded. However, both LHCf1 and D2 proteins instead didn't show any alteration in content in strain I48 thylakoids, with respect to the parental strain, highlighting that the alteration in the LHCX1 protein content didn't trigger other alterations in the photosynthetic apparatus, at least for the two major proteins here investigated. When the gel was instead loaded with 4  $\mu\text{g}$  of Chl per each sample and the alkaline phosphatase signal was saturated, the LHCX1 protein was also detectable in strain I48 thylakoids (Figure 4B).



**Figure 4. Western blot analysis.** Western blot quantification of the LHCX1 protein from *N. gaditana* thylakoids. As a control, the PSII core subunit D2 was also detected on the same membrane. Also the LHCf1 protein was detected on the same membrane, to evaluate alterations in its content. In A. thylakoids corresponding to 1  $\mu\text{g}$  of total Chl were loaded while in B. the samples correspond to 4  $\mu\text{g}$  of total Chl. For both images, WT and I48 samples come from the same membrane and were exposed to the enzymatic reaction for the same time.

The detected protein cannot be the shortest mutated variant of LHCX1, since its predicted molecular weight is around 21 KDa, therefore visibly different from the 25.5 KDa corresponding to the parental LHCX1. Therefore, the detected protein could be the LHCX2 (its sequence is shown in figure 3B), which

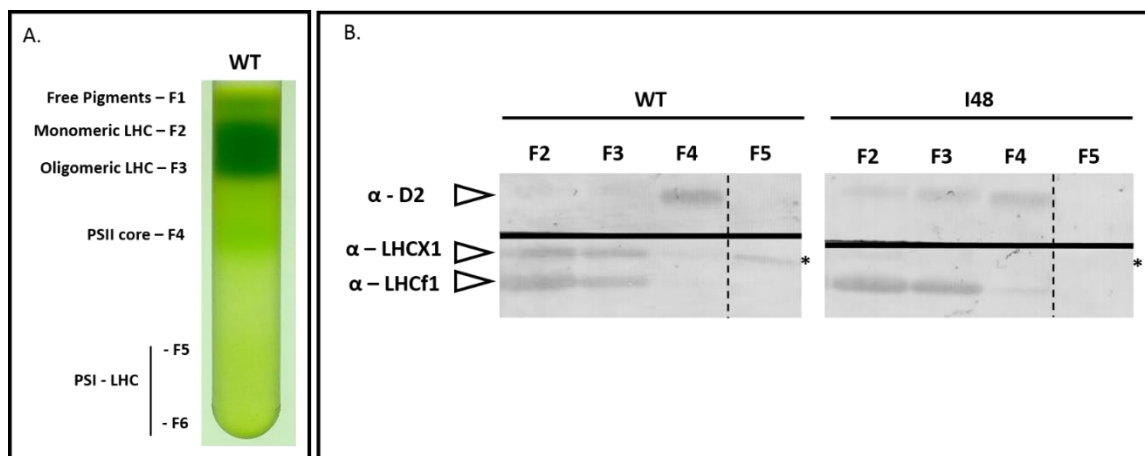
shows a 45 % amino acids sequence identity with LHCX1, and has a molecular weight of 26 KDa. Otherwise we could hypothesize that the splicing event of the 4<sup>th</sup> intron of the gene could indeed take place also if the GT dinucleotide of the 5'-donor splicing site is converted in a GA (Figure 1A), but with a strongly reduced efficiency with respect to the WT *locus*, therefore leading to the accumulation of the parental form of LHCX1 also in the I48 strain, but in strongly minor amount. Indeed, in literature several non-canonical splicing sites (GC-AG, GG-AG, GT-TG, GT-CG or CT-AG) have been already described as likely involved in alternative-splicing events (Zhang et al., 2007). In this case, the thymine deletion in the dinucleotide GT could have indeed induced the formation of an alternative 5'-donor splicing site, recognized by the spliceosome as well but with a low efficiency. Further investigations will however be required to better elucidate the molecular mechanism involved.

The absence (or the strong reduction in accumulation) of LHCX1, found in the *N. gaditana* mutant strain I48, is likely responsible for the inability to activate the NPQ mechanism in this strain. However, it is worth remembering that the mutant I48 was selected following an insertional mutagenesis event, therefore at least one copy of the resistance cassette is integrated in its genome. However, the latter wasn't found integrated in any gene involved in photosynthesis and its position in the genome hasn't been identified, yet, as well as its copy number. Mutant I48 complementation is indeed seminal to exclude the participation of the resistance cassette insertion *locus (i)* to this photosynthetic phenotype. Rescuing the parental phenotype through the complementation of the I48 genome with a wild type copy of the Naga\_100173g12 gene will be therefore necessary to correlate the identified genetic variant with the inability to activate NPQ. The occurrence of a nucleotide deletion could have been caused by the prolonged exposition to the antibiotic zeocin, during the selection event. It is indeed known the effect of the latter in inducing DNA strands breaks, with some studies showing its effect in inhibiting the incorporation of thymidine into DNA strands (Delacôte et al., 2007; Krol et al., 2015; Chen and Umeda, 2015).

*In Nannochloropsis gaditana, LHCX1 is associated to the PSI.*

To investigate further the impact of the absence (or the strong reduction in accumulation) of LHCX1 in the photosynthetic apparatus complexes of strain I48, the *N. gaditana* WT and the I48 strain thylakoids were solubilized and their pigment-binding complexes were separated by ultracentrifugation in a sucrose gradient (see methods for details). Five distinct fractions, showing different migration rates in the sucrose gradient (F1-F5 in Figure 5A), were separated. The sixth fraction was instead collected from the bottom of the tube. Thanks to spectroscopic analyses and the

data previously reported in (Basso et al., 2014), the sucrose gradient bands were identified as reported in figure 5A. Thanks to further western blot analysis, the fractions identity was validated (Figure 5B). The absence (or the strong reduction in accumulation) of the LHCX1 protein in strain I48 didn't induce any alteration in the migration profile of its major pigment-binding complexes, therefore only the WT sucrose gradient was reported in figure 5A. When the major protein components of the *N. gaditana* photosynthetic apparatus were detected through western blot, we showed that, while the LHCf1 protein, mainly involved in the light harvesting process, is only localized in the F2 and F3 fractions (monomeric and oligomeric antenna proteins), the LHCX1 protein is instead present in the latter, but also in the F5 band, corresponding to the PSI core, associated to its antenna complexes.



**Figure 5. Isolation of *Nannochloropsis gaditana* pigment binding proteins by sucrose gradient ultracentrifugation.** A. Sucrose gradient for the *Nannochloropsis gaditana* wild type strain after mild solubilization with 0.4%  $\alpha$ -DM. Five bands are clearly visible (F1–F5) while the sixth is present in the bottom of the tube (F6). The identification of the bands is here reported, according to western blot and spectroscopic analysis. B. Western blot analysis of the proteins composition for some of the sucrose gradient bands reported in A. In order to assess the protein distribution in the sucrose gradient, an equal volume of each band was loaded (40  $\mu$ l). For both the WT and I48 strain, the F5 sample comes from the same membrane. A black asterisks highlights the presence of the LHCX1 protein in the F5 band of the WT and its missing in the I48 strain.

This is the first documentation of LHCX1 co-localization with the PSI in *Nannochloropsis gaditana*. In other photosynthetic organisms, such as the moss *P. patens* and two marine diatoms (*T. pseudonana* and *P. tricornutum*) the co-localization of LHCX proteins and the PSI was however already

documented, and their possible role as excess energy quenchers to protect the latter was indeed already hypothesized (Pinnola et al., 2015; Grouneva et al., 2011), but never demonstrated. Therefore, the absence of the LHCX1 protein in the I48 strain could serve to demonstrate its biological role when found close to the PSI.

### **Conclusions.**

*N. gaditana* mutant strain I48 was previously isolated for a strong reduction in the activation of the process of NPQ. Thanks to its whole genome re-sequencing, a genetic variant in the 4<sup>th</sup> intron 5'-donor splicing site of the Naga\_100173g12 gene (coding for the LHCX1 protein) was here identified. The latter was shown to retain the intron sequence also in the mature mRNA, affecting the LHCX1 translation event with a consequent impairment in the protein accumulation titer. This work represents a step forward in the elucidation of the genetic basis, cause of the I48 strain photosynthetic phenotype. The validation that the identified mutation in the Naga\_100173g12 gene is the only cause of this phenotype will be however provide through the complementation of its genome with the wild type Naga\_100173g12 gene. However, the absence of the LHCX1 protein in the strain I48 makes the latter a good experimental tool for further investigations on the biological role of this protein in relation to the PSI in *N. gaditana*.

## References.

- Antal, Z., Rasclé, C., Fèvre, M., and Bruel, C.** (2004). Single oligonucleotide nested PCR: a rapid method for the isolation of genes and their flanking regions from expressed sequence tags. *Curr. Genet.* **46**: 240–6.
- Basso, S., Simionato, D., Gerotto, C., Segalla, A., Giacometti, G.M., and Morosinotto, T.** (2014). Characterization of the photosynthetic apparatus of the Eustigmatophycean *Nannochloropsis gaditana*: evidence of convergent evolution in the supramolecular organization of photosystem I. *Biochim. Biophys. Acta* **1837**: 306–14.
- Bonente, G., Formighieri, C., Mantelli, M., Catalanotti, C., Giuliano, G., Morosinotto, T., and Bassi, R.** (2011). Mutagenesis and phenotypic selection as a strategy toward domestication of *Chlamydomonas reinhardtii* strains for improved performance in photobioreactors. *Photosynth. Res.* **108**: 107–20.
- Burset, M., Seledtsov, I.A., and Solovyev, V. V** (2000). Analysis of canonical and non-canonical splice sites in mammalian genomes. *Nucleic Acids Res.* **28**: 4364–75.
- Chen, P. and Umeda, M.** (2015). DNA double-strand breaks induce the expression of flavin-containing monooxygenase and reduce root meristem size in *Arabidopsis thaliana*. *Genes Cells* **20**: 636–46.
- Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M.** A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. **6**: 80–92.
- Corteggiani Carpinelli, E., Telatin, A., Vitulo, N., Forcato, C., D'Angelo, M., Schiavon, R., Vezzi, A., Giacometti, G.M., Morosinotto, T., and Valle, G.** (2014). Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion. *Mol. Plant* **7**: 323–35.
- Croce, R., Canino, G., Ros, F., and Bassi, R.** (2002). Chromophore Organization in the Higher-Plant Photosystem II Antenna Protein CP26. *Biochemistry* **41**: 7334–7343.
- Dall'Osto, L., Bressan, M., and Bassi, R.** (2015). Biogenesis of light harvesting proteins. *Biochim. Biophys. Acta* **1847**: 861–71.
- Delacôte, F., Deriano, L., Lambert, S., Bertrand, P., Saintigny, Y., and Lopez, B.S.** (2007). Chronic exposure to sublethal doses of radiation mimetic Zeocin selects for clones deficient in homologous recombination. *Mutat. Res.* **615**: 125–33.

- DeLano, W.L.** (2005). The case for open-source software in drug discovery. *Drug Discov. Today* **10**: 213–7.
- Dent, R.M., Haglund, C.M., Chin, B.L., Kobayashi, M.C., and Niyogi, K.K.** (2005). Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of *Chlamydomonas reinhardtii*. *Plant Physiol.* **137**: 545–56.
- Dittami, S.M., Michel, G., Collén, J., Boyen, C., and Tonon, T.** (2010). Chlorophyll-binding proteins revisited—a multigenic family of light-harvesting and stress proteins from a brown algal perspective. *BMC Evol. Biol.* **10**: 365.
- González-Ballester, D., de Montaigu, A., Galván, A., and Fernández, E.** (2005). Restriction enzyme site-directed amplification PCR: a tool to identify regions flanking a marker DNA. *Anal. Biochem.* **340**: 330–5.
- Green, B.R. and Pichersky, E.** (1994). Hypothesis for the evolution of three-helix Chl a/b and Chl a/c light-harvesting antenna proteins from two-helix and four-helix ancestors. *Photosynth. Res.* **39**: 149–162.
- Grouneva, I., Rokka, A., and Aro, E.-M.** (2011). The thylakoid membrane proteome of two marine diatoms outlines both diatom-specific and species-specific features of the photosynthetic machinery. *J. Proteome Res.* **10**: 5338–53.
- Gumpel, N.J. and Purton, S.** (1994). Playing tag with *Chlamydomonas*. *Trends Cell Biol.* **4**: 299–301.
- Hall, T.** (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**: 95 – 98.
- Hooper, J.K., Eggink, L.L., and Chen, M.** Chlorophylls, ligands and assembly of light-harvesting complexes in chloroplasts. *Photosynth. Res.* **94**: 387–400.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E.** (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**: 845–858.
- Kirst, H., Garcia-Cerdan, J.G., Zurbriggen, A., Ruehle, T., and Melis, A.** (2012). Truncated photosystem chlorophyll antenna size in the green microalga *Chlamydomonas reinhardtii* upon deletion of the TLA3-CpSRP43 gene. *Plant Physiol.* **160**: 2251–60.
- Krol, K., Brozda, I., Skoneczny, M., Bretner, M., Bretne, M., and Skoneczna, A.** (2015). A genomic screen revealing the importance of vesicular trafficking pathways in genome maintenance and protection against genotoxic stress in diploid *Saccharomyces cerevisiae* cells. *PLoS One* **10**: e0120702.
- Laemmli, U.K.** (1970). Cleavage of structural proteins during the assembly of the head of



- bacteriophage T4. *Nature* **227**: 680–5.
- Liu, Y.G., Mitsukawa, N., Oosumi, T., and Whittier, R.F.** (1995). Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.* **8**: 457–63.
- Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X., and Chang, W.** (2004). Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428**: 287–292.
- Mount, S.M. and Steitz, J.A.** (1981). Sequence of U1 RNA from *Drosophila melanogaster*: implications for U1 secondary structure and possible involvement in splicing. *Nucleic Acids Res.* **9**: 6351–68.
- Niyogi, K.K. and Truong, T.B.** (2013). Evolution of flexible non-photochemical quenching mechanisms that regulate light harvesting in oxygenic photosynthesis. *Curr. Opin. Plant Biol.* **16**: 307–14.
- Ochman, H., Gerber, A.S., and Hartl, D.L.** (1988). Genetic applications of an inverse polymerase chain reaction. *Genetics* **120**: 621–3.
- Papasaikas, P. and Valcárcel, J.** (2015). The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends Biochem. Sci.* **41**: 33–45.
- Perin, G., Bellan, A., Segalla, A., Meneghesso, A., Alboresi, A., and Morosinotto, T.** (2015). Generation of random mutants to improve light-use efficiency of *Nannochloropsis gaditana* cultures for biofuel production. *Biotechnol. Biofuels* **8**: 161.
- Pinnola, A., Cazzaniga, S., Alboresi, A., Nevo, R., Levin-Zaidman, S., Reich, Z., and Bassi, R.** (2015). Light-Harvesting Complex Stress-Related Proteins Catalyze Excess Energy Dissipation in Both Photosystems of *Physcomitrella patens*. *Plant Cell* **27**: 3213–27.
- Polle, J.E.W., Kanakagiri, S.-D.D., and Melis, A.** (2003). *tla1*, a DNA insertional transformant of the green alga *Chlamydomonas reinhardtii* with a truncated light-harvesting chlorophyll antenna size. *Planta* **217**: 49–59.
- Porra, R.J., Thompson, W.A., and Kriedemann, P.E.** (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta - Bioenerg.* **975**: 384–394.
- Robert, X. and Guet, P.** (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**: W320–4.
- Schierenbeck, L., Ries, D., Rogge, K., Grewe, S., Weisshaar, B., and Kruse, O.** (2015). Fast forward genetics to identify mutations causing a high light tolerant phenotype in *Chlamydomonas reinhardtii* by whole-genome-sequencing. *BMC Genomics* **16**: 57.

**Zhang, C., Hastings, M.L., Krainer, A.R., and Zhang, M.Q.** (2007). Dual-specificity splice sites function alternatively as 5' and 3' splice sites. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 15028–33.

**Zhang, R., Patena, W., Armbruster, U., Gang, S.S., Blum, S.R., and Jonikas, M.C.** (2014). High-Throughput Genotyping of Green Algal Mutants Reveals Random Distribution of Mutagenic Insertion Sites and Endonucleolytic Cleavage of Transforming DNA. *Plant Cell* **26**: 1398–1409.



# CHAPTER V

## Light Remodels Lipid Biosynthesis in *Nannochloropsis gaditana* by Modulating Carbon Partitioning Between Organelles

### Author names and affiliations

Alessandro Alboresi<sup>1\*</sup>, Giorgio Perin<sup>1\*</sup>, Nicola Vitulo<sup>3</sup>, Gianfranco Diretto<sup>4</sup>, Maryse Block<sup>2</sup>, Juliette Jouhet<sup>2</sup>, Andrea Meneghesso<sup>1</sup>, Giorgio Valle<sup>3</sup>, Giovanni Giuliano<sup>4</sup>, Eric Maréchal<sup>2</sup> and Tomas Morosinotto<sup>1</sup>

\* Equal contribution

<sup>1</sup>PAR-Lab\_Padua Algae Research Laboratory, Department of Biology, University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy

<sup>2</sup>Laboratoire de Biologie Cellulaire et Végétale, UMR 5168 CNRS - CEA - INRA - Université Grenoble Alpes, BIG, CEA-Grenoble, 17 rue Des Martyrs, 38054 Grenoble, Cedex 9, France

<sup>3</sup>Innovative Biotechnologies Interdepartmental Research Center (CRIBI), University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy

<sup>4</sup>Italian National Agency for New Technologies, Energy, and Sustainable Development (ENEA), Casaccia Research Centre, Via Anguillarese 301, 00123 Roma, Italy

THIS CHAPTER WAS SUBMITTED TO PLOS BIOLOGY

**Abstract.**

The seawater microalga *Nannochloropsis gaditana* is capable of accumulating a large fraction of reduced carbon as lipids. To clarify the molecular bases of this metabolic feature, we investigated light-driven lipid biosynthesis in *Nannochloropsis gaditana* cultures combining the analysis of photosynthetic functionality with transcriptomic, lipidomic and metabolomic approaches.

Light-dependent alterations are observed in amino acid, isoprenoid, nucleic acid and vitamin biosynthesis, suggesting a deep remodeling in the microalgal metabolism triggered by photoadaptation. In particular, high light intensity is shown to affect lipid biosynthesis, inducing the accumulation of diacylglyceryl-N,N,N-trimethylhomoserine (DGTS) and triacylglycerols (TAGs), together with the up-regulation of genes involved in their biosynthesis. Concurrently, chloroplast polar lipids are instead decreased. This situation correlates with the induction of genes coding for a putative cytosolic fatty acid synthase of type 1 (*FAS1*) and polyketide synthase (*PKS*) and the downregulation of the chloroplast fatty acid synthase of type 2 (*FAS2*). Lipid accumulation is accompanied by the regulation of triose phosphate/inorganic phosphate transport across the chloroplast membranes, tuning the carbon metabolic allocation between cell compartments, favoring the cytoplasm and endoplasmic reticulum at the expense of the chloroplast.

These results highlight the high flexibility of lipid biosynthesis in *Nannochloropsis gaditana* and also indicate a mechanism for the regulation of primary carbon partitioning by controlling metabolite allocation at the sub-cellular level.

## Introduction

More than 45% of our planet's annual net primary biomass is obtained by the photosynthetic activity of microalgae that convert light energy into chemical energy in the chloroplast (Wang et al., 2011; Falkowski et al., 2004). The fixation of CO<sub>2</sub> within the Calvin-Benson cycle plays a central role in the microalgal primary metabolism and originates fundamental precursors feeding the synthesis of starch, proteins or fatty acids (Zhao and Su, 2014; Ho et al., 2014). In most photosynthetic organisms, starch and other polysaccharides are the primary form of temporary accumulated macromolecules with high energetic content (Stitt and Zeeman, 2012; Vítová et al., 2015). However, some species of microalgae are highly efficient in accumulating triacylglycerols (TAG) in lipid bodies in the cytosol by a metabolic pathway that is highly coordinated across the chloroplast, the endoplasmic reticulum and the cytosol itself (Radakovits et al., 2010; Bogen et al., 2013; Poliner et al., 2015). Some of these microalgae species can accumulate lipids up to 60% of their dry weight (Hu et al., 2008), and for this reason, they have recently attracted attention as a possible feedstock for the production of biofuels. In this context, particular interest is raised by species of the *Nannochloropsis* genus, unicellular photosynthetic microalgae belonging to the Eustigmatophyceae class within Heterokonts and distributed worldwide in marine, fresh and brackish waters, capable of massive lipid accumulation, especially under nutrient deprivation (Boussiba et al., 1987; Rodolfi et al., 2009).

Because of these properties, *Nannochloropsis* is an interesting model for the study of metabolism regulation and understanding how different organisms regulate metabolic fluxes towards lipid or carbohydrate synthesis. In the case of *Nannochloropsis*, the ability to synthesize lipids is supported by a genomic enrichment in biosynthesis genes and glycoside hydrolases that should play a role in shifting carbon fluxes toward TAG metabolism. An example of this expansion is the *Diacylglycerol acyltransferase-2 (DGAT-2)* genes, which in *Nannochloropsis* species investigated showed 11 copies. As a comparison, diatomssuch as *Thalassiosira pseudonana* or *Phaeodactylum tricorutum* have 4 *DGAT-2* copies, while the green alga *Chlamydomonas reinhardtii* has 5 *DGAT-2* genes, despite all having much larger genomes than *Nannochloropsis* (Wang et al., 2014; Radakovits et al., 2012; Vieler et al., 2012; Corteggiani Carpinelli et al., 2014).

Lipid accumulation in several algal species is triggered by nitrogen (N) depletion, and these conditions have been heavily investigated to understand lipid metabolism regulation (Li et al., 2014; Lu et al., 2014a; Martin et al., 2014; Jia et al., 2015; Abida et al., 2015; Simionato et al., 2013; Miller et al., 2010). However, N depletion also causes a major impairment of the basic metabolism, shutting down nucleic acid and protein biosynthesis and limiting cell division and growth

(Simionato et al., 2013). Upon nitrogen shortage, the fatty acids required for TAG accumulation in *Nannochloropsis gaditana* are mostly produced *de novo* biosynthesis with, nevertheless, a significant contribution of galactolipids recycled from degraded thylakoid membranes (Simionato et al., 2013). Investigating other stimuli capable of inducing TAG biosynthesis could thus be highly valuable to disentangle the nitrogen depletion response and identify metabolic switches that induce lipid accumulation without depressing biomass growth. It has been shown that light can induce lipid accumulation in *Nannochloropsis salina* (Sforza et al., 2012), and as photosynthesis the primary source of all reduced carbon, it is understandable that light is a master regulator of metabolism. Broadly speaking, lipid biosynthesis and management depend on the production and recycling of photosynthates. Investigating the metabolic remodeling following light treatment and understanding the crosstalk between the metabolic pathways in response to light changes can thus clarify how carbon partitioning is modulated.

To this end, in this work we have investigated the light response of *N. gaditana* cultures, combining a genome-wide transcriptional analysis with lipidomic, metabolomic and functional approaches. The results showed that the lipid accumulation induced by strong illumination is consistent with the coordinated activation of the putative fatty acid biosynthesis in the cytoplasm/ER and the parallel inhibition of the chloroplast fatty acid biosynthesis. This process is accompanied by increased carbon fluxes out of the chloroplast, shifting the balance towards ER and TAG accumulation.

## Materials and Methods

### *Strains and culture conditions*

*N. gaditana* wild-type from CCAP (Culture Collection of Algae and Protozoa), strain 849/5, was cultivated in sterile F/2 medium (Guillard and Ryther, 1962), using sea salts 32 g/L (S9883, Sigma Aldrich), 40 mM TRIS-HCl pH 8 and Guillard's (f/2) marine water enrichment solution (G9903, Sigma Aldrich). Cultures were generally kept in Erlenmeyer flasks with magnetic stirring or in plates on the same F/2 medium supplemented by agar 10 g/L, under continuous light at 100  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  at 21 °C.

Before starting the experiment, cultures were treated with an antibiotic cocktail of Ampicillin (100  $\mu\text{g/mL}$ ), Streptomycin sulphate (100  $\mu\text{g/mL}$ ) and Kanamycin sulfate (100  $\mu\text{g/mL}$ ) (all from Sigma Aldrich) for 48 h to obtain axenic cultures.

A Multi-Cultivator MC 1000-OD device (Photon Systems Instruments, Czech Republic) was used to grow *N. gaditana* cells for transcriptomics, metabolomics and lipidomics. The same mother culture was grown in bottle inflated by 5 % CO<sub>2</sub> and then used to inoculate independent Multi-Cultivator tubes (three for low/high and two for standard light condition). The starting OD<sub>720nm</sub> was about 0.2 and the volume per tube was 80 mL.

The full experiment was repeated three times in order to collect independent biological repetitions. Optical density at 720 nm was measured every hour by the instrument and an air pump was used to aerate the culture and to keep the cells in suspension. A water bath stabilized the culture to the constant temperature of 21 °C. After 5 days of growth, cell density was measured using a cell counter (Cellometer Auto X4 – Nexcelom Bioscience).

### *Pigment content analysis and chlorophyll fluorescence measurement*

Chlorophyll a and total carotenoids were extracted from cells after 5 days of growth using 100 % N,N'-dimethylformamide (Sigma Aldrich). Pigment extraction was carried out at 4°C, for at least 24h, in the dark. Pigment concentration was determined spectrophotometrically using a Cary 100 spectrophotometer (Agilent Technologies). PSII function was monitored through *in vivo* Chl fluorescence determination directly on cell suspensions at room temperature, with a PAM 100 fluorimeter (Heinz-Walz, Effeltrich, Germany). A saturating light pulses of 0,6 s at 6000  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  was used to measure Fm. Fv/Fm was measured as (Fm-F0)/Fm.

PSII functional antenna size (ASII) was determined using a JTS10 spectrophotometer (Bio-Logic, France). Cells after 5 days of growth were kept in recovery in low light conditions (LL, 10  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ ) to minimize the contribution of closed-state reaction centers to the measurement,



a phenomenon otherwise observed for ML and HL conditions. Once brought to a concentration of  $200 \times 10^6$  cells/ml, samples were incubated with 80  $\mu$ M DCMU for 10 minutes. Fluorescence induction kinetics were monitored upon excitation with  $320 \mu\text{mol photons m}^{-2} \text{s}^{-1}$  of actinic light at 630 nm.  $t_{2/3}$  values were used to calculate the functional antenna size of PSII.

Samples, for both Fv/Fm and ASII determination, were dark-adapted for 20 minutes before starting the measurements.

#### *Transcriptome sampling and sequencing*

After 5 days of growth, independent photo-bioreactor tubes were pulled together to get 20 mL of LL culture (about  $2 \times 10^7$  cells/mL) and 10 mL of ML/HL culture (about 5 and  $3.5 \times 10^7$  cells/mL respectively). Cells were collected at 5000 g for 10 min at room temperature (RT) and the pellet was suspended in 1 mL of F/2 medium and then transferred in a 2 mL tubes. The cells were centrifuged again at 5000 g for 5 min at RT and the pellet was frozen in liquid nitrogen. Less than 20 min passed since the end of the experiment (5 days of growth) and the freezing of the collected samples. RNA was extracted from frozen pellets using the RNeasy Plant Mini kit (Qiagen), by adapting the cell-grinding step to the need of *N. gaditana*. RNA quality was checked with a Bioanalyzer (Agilent). All the samples had a RIN of 7.2-7.8. Samples for Ligation Sequencing were prepared according to the SOLiD Whole transcriptome library preparation protocol (pn 4452437 Rev.B). Samples were sequenced using a SOLiD System 5500XL. The average number of RNA-seq reads per sample was approximately 45M and they ranged between 30M to 55M (Supplemental Table 2). The reads of each of the 9 samples (3 biological replicates for each of the 3 light intensities) were mapped to the *N. gaditana* B-31 genome (Corteggiani Carpinelli et al., 2014) using PASS aligner (Campagna et al., 2013). The percentage identity was set to 90% with one gap allowed whereas the quality filtering parameters were set automatically by PASS. Moreover, a minimum reads length cut-off of 50 and 30 nt was set for the forward sequences and reverse reads, respectively. The spliced reads were identified using the procedure described in PASS manual (<http://pass.cribi.unipd.it>). Forward and reverse reads were aligned independently on the reference genome while the pairing between forward and reverse reads was performed using PASS-pair program from the PASS package. The abundance of each gene was quantified using htseq-counts program (<http://www-huber.embl.de/users/anders/HTSeq/doc/count.html>). The quantification was performed using only the uniquely mapped reads and providing as reference prediction the gene set available (Corteggiani Carpinelli et al., 2014).

### *Estimation of differential gene expression*

Differentially expressed gene analysis was performed using EdgeR software (Robinson et al., 2010). The raw read counts were normalized taking into consideration both the different depth of sequencing among the samples and the gene GC content. The normalization was performed using EDASeq package (Risso et al., 2011). We considered as differentially expressed all the genes with a p-value < 0.05 after FDR correction.

### *Hierarchical clustering*

Cluster analysis of differentially expressed genes was performed using the TMeV suite (TMeV 4.3 - Significance Analysis of Microarray; multiclass analysis with 5% FDR accepted) (Tusher et al., 2001; Saeed et al., 2003). HCL was performed on each cluster to represent gene relationships in dendrograms (TMeV), with Pearson's correlation distance as the metric. Classes considered for clustering the differentially expressed genes were (a) cells grown in LL conditions, (b) cells grown in ML conditions and (c) cells grown in HL conditions (3 samples per each class). Gene Ontology analysis of the differentially expressed genes was performed according to Blast2GO web tool (Conesa et al., 2005).

### *Phylogenetic analysis*

Protein sequences were chosen among three sub-classes of the large LHC superfamily that are LHCf, mainly devoted to light harvesting, LHCX/LHCSR, involved in photoprotection, and LHCa/b, the main antenna complexes of PSI and PSII of *Chlamydomonas reinhardtii* and *Arabidopsis thaliana*. The multiple alignment of protein sequences were performed using MUSCLE (Multiple Sequence Comparison by Log-Expectation) (Edgar, 2004) and manually adjusted in BioEdit (Supplemental Figure 5). Phylogeny reconstruction was performed using Maximum Likelihood as statistical method and the reliability of the interior branches of the unrooted phylogenetic trees was assessed through 1000-iteration bootstrap resampling using Mega6.0 (Tamura et al., 2013). Results on Naga\_101036g3 were also confirmed by performing a Maximum Likelihood analysis using PhyML (Guindon et al., 2010) and a Neighbor-Joining using Mega6.0 (Tamura et al., 2013).

### *Metabolomic analysis*

For the determination of isoprenoids, 5mg of lyophilized microalgal cells were extracted with 0.5ml methanol, followed by sonication (3 cycles of 10'' at 10Hz) and lysis with an Ultraturrax homogenizer at full speed; subsequently, 1ml of chloroform containing 25 mg/l of  $\alpha$ -tocopherol acetate (as internal standard; SigmaAldrich) was added and samples were shaken using a

MixerMill 300 (Qiagen) for 30' at 25 Hz. Subsequently, 0.5ml of 50 mM Tris-HCl pH 7.5 containing 1M NaCl were added to each sample. After vortexing, samples were centrifugated at 20,000 g for 5' at 4°C. The chloroform fractions were collected and dried using a speedvac. The pellets were resuspended in 1ml ethylacetate and samples were vortexed and centrifugated at 20,000 g for 15' at 4°C. Finally 20 µl for each sample were injected to the LC-APCI-MS. During the whole extraction, the samples were kept on ice and in dim light conditions in order to prevent pigment degradation. For each experimental point, at least two independent extractions from three independent cell cultures were performed.

LC-MS analysis was carried out using an Accela U-HPLC-DAD system coupled to an LTQ-Orbitrap Discovery mass spectrometer (Thermo Fischer Scientific) operating in positive mode-atmospheric pressure chemical ionization (APCI); chromatographic conditions have been previously described (Liu et al., 2014). The APCI-MS ionization parameters were as follows: 25 Units of nitrogen (sheath gas) and 10 units of auxiliary gas were used, respectively; the vaporizer temperature was 250°C, the capillary temperature was 200°C, the discharge current was 4.0 µA, the capillary voltage and tube lens settings were 60 V and 150 V, respectively. All the Chemicals and solvents were LC-MS grade quality from Sigma-Aldrich (CHROMASOLV®). Carotenoids have been quantified as previously described (Liu et al., 2014; Su et al., 2015) while tocochromanols and quinones have been identified on the basis of the comparison with chromatographic and spectral properties with authentic standards and reference spectra, literature data, m/z accurate masses obtained from the Pubchem database (<http://pubchem.ncbi.nlm.nih.gov/>) for native compounds, or from the Metabolomics Fiehn Lab Mass Spectrometry Adduct Calculator (<http://fiehnlab.ucdavis.edu/staff/kind/Metabolomics/MSAdduct-Calculator/>) for adducts. Subsequently, a relative quantification has been performed on the basis of the level of the internal standard, and of the dry weight.

LC-ESI(+)-MS analysis of fruit primary and secondary semi-polar metabolome has been performed as previously described (De Vos et al., 2007; Iijima et al., 2008) with slight modifications. 10 mg of freeze-dried microalgal cells were extracted with 0.5 mL cold 75% (v/v) methanol, 0.1% (v/v) formic acid, spiked with 10 µg/ml formononetin. After grinding and sonication, the extracts have been shaken for 40' at 20 Hz using a Mixer Mill 300 (Qiagen) and centrifuged for 15' at 20,000 g at 4°C. 0.3 mL of supernatants were removed and transfer to HPLC tubes. For each experimental point, at least two independent extractions from three independent cell cultures were performed. LC-MS analyses were carried out using a LTQ-Orbitrap Discovery mass spectrometry system (Thermo Fisher Scientific) operating in positive electrospray ionization (ESI), coupled to an Accela U-HPLC system (Thermo Fisher Scientific, Waltham, MA). Liquid chromatography was carried out

using a Phenomenex C18 Luna column (150 x 2.0 mm, 3  $\mu$ m) and mobile phase was composed by water–0.1% Formic Acid (A) and acetonitrile –0.1% Formic Acid (B). The gradient was: 95%A:5%B (one minute), a linear gradient to 25%A:75%B over 40 minutes, 2 minutes isocratic, before going back to the initial LC conditions in 18 minutes. Ten  $\mu$ l of each sample were injected and a flow of 0.2 mL was used during the whole LC runs. Detection was carried out continuously from 230 to 800 nm with an online Accela Surveyor photodiode array detector (PDA, Thermo Fischer Scientific, Waltham, MA). All solvents used were LC-MS grade quality (CHROMASOLV<sup>®</sup> from Sigma-Aldrich). Metabolite identification and quantification have been performed as previously described.

#### *Glycerolipid analysis*

Glycerolipids (i.e. membrane glycerolipids containing 2 fatty acids and triacylglycerols containing 3 fatty acids) were extracted from 10 mg of freeze-dried *N. gaditana* cells, according to (Folch et al., 1957). Cells were frozen in liquid nitrogen immediately after harvest. The freeze-dried cell pellet was resuspended in 4 mL of boiling ethanol for 5 minutes followed by the addition of 2 mL of methanol and 8 mL of chloroform at room temperature. The mixture was then saturated with argon and stirred for 1h at room temperature. After filtration through glass wool, cell remains were rinsed with 3 mL of chloroform/methanol 2:1, v/v. In order to initiate biphasic formation, 5 mL of NaCl 1% was then added to the filtrate. The chloroform phase was dried under argon before re-solubilization of the lipid extract in pure chloroform. To isolate TAG, lipids were run on silica gel plates (Merck) with hexane:diethylether:acetic acid (70:30:1, v/v). To isolate polar glycerolipids, lipids were analyzed on silica gel plates (Merck) by two-dimensional thin-layer chromatography (2D-TLC). The first solvent was chloroform:methanol:water (65:25:4, v/v) while the second one was chloroform:acetone:methanol:acetic acid:water (50:20:10:10:5, v/v). Lipids were then visualized under UV light after pulverization of 8-anilino-1-naphthalenesulfonic acid at 2% in methanol. Lipid classes were then assessed based on position on the 2D-TLC, according to previous studies (Simionato et al., 2013), adding a DGTS standard that was not used in this previous study. Carboxymethyl phosphatidylethanolamine was identified by mass spectrometry as described by (Shoji et al., 2010). Glycerolipids were then scraped off from the 2D-TLC plates for further analyses. For lipid quantification, fatty acids were methylated using 3 mL of 2.5% H<sub>2</sub>SO<sub>4</sub> in methanol during 1 h at 100°C (including standard amounts of 21:0). The reaction was stopped by the addition of 3 mL of water and 3 mL of hexane. The hexane phase was analyzed by gas liquid chromatography (Perkin Elmer) on a BPX70 (SGE) column. Methylated fatty acids were identified by comparison of their retention times with those of standards and quantified by surface peak method using 21:0 for calibration. Extraction and quantification were done at least 3 times.

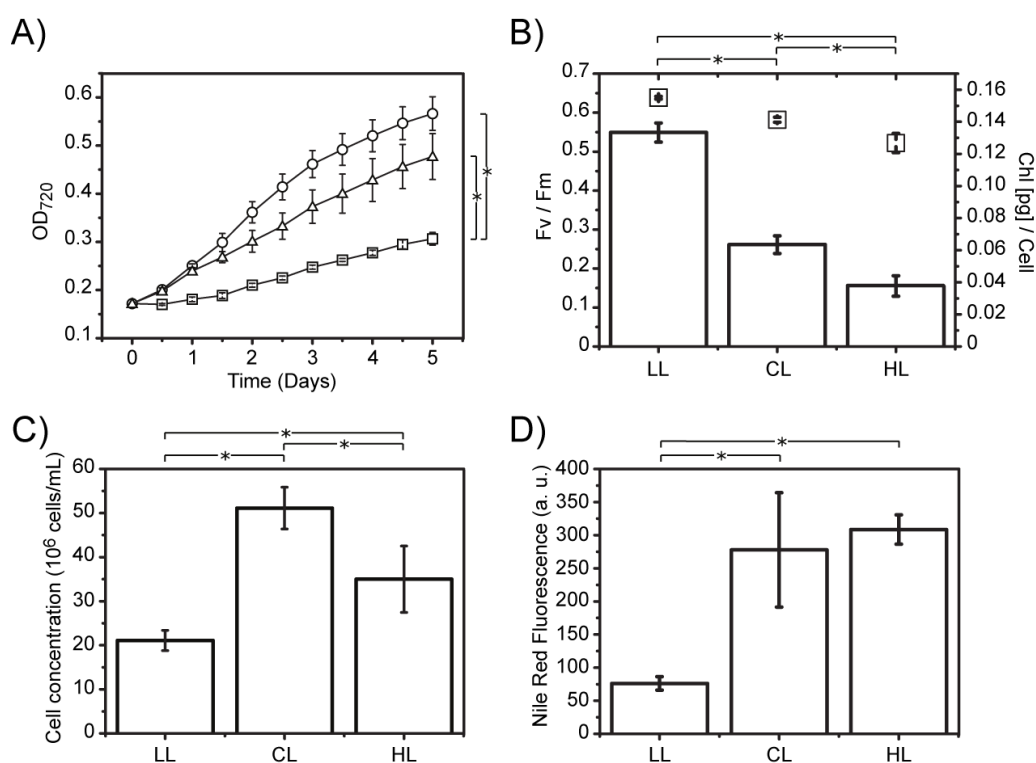
## Accession Numbers

Sequence data used for phylogenetic reconstruction can be found in the GenBank data library. AtLhca2, AT3G61470. AtLhca3, AT1G61520. AtLhca4, AT3G47470. AtLhcb1.1, AT1G29930. AtLhcb2.1, AT2G05070. AtLhcb5, AT4G10340. CrLhca3, XP\_001701405. CrLhca4, XP\_001698519. CrLhcb5, XP\_001695927. CrLhcbm2, XP\_001700243. CrLhcbm2, XP\_001694115. CeLHCSR, X60721. SoLHCSR, DQ394297. MvLHCSR, DQ370082. OtLHCSR, AY954739. CrLHCSR3, XM\_001696073. CrLHCSR1, XM\_001696086. PpLHCSR1, XM\_001776900. PpLHCSR2, XM\_001768019. TpLHCX6, XM\_002295147. EsiLHCX1, CBJ48499. EsiLHCX2, CBN73961. EsiLHCX3, CBJ27268. EsiLHCx4, CBJ27787. EsiLHCX5, CBJ27788. EsiLHCX6, CBJ27803. EsiLHCX7, CBJ27803. NgLhcx1, EWM30071. NgLHCf1, EWM21572.1. NgLHCX2, EWM27041. PtLHCf12, XM\_002176699. TpLHCf7, XM\_002294544. PtLHCf11, XM\_002184727. PtLHCf10, XM\_002183345. TpLHCf10, XM\_002295583. EsiLHCf1, CBN76797. EsiLHCf2, CBN74050. EsiLHCf3, CBJ30555. EsiLHCf4, CBN76797. PtLHCf13, XP\_002182219. PtLHCf14, XP\_002182305. CycLHCf1, CAA04178. TpLHCf1, XP\_002294743. TpLHCf3, XP\_002295015. TpLHCf4, XP\_002294845. TpLHCf5, XP\_002294608. TpLHCf6, XP\_002288729. TpLHCf8, XP\_002286436.

## Results and Discussion.

### *Nannochloropsis gaditana* cell growth in limiting and excess light conditions

Cells of *Nannochloropsis gaditana* strain CCAP 849/5 were cultivated in small-scale photobioreactors at three different light intensities: low (LL,  $10 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ), medium (ML,  $100 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ) and high (HL  $1000 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ) light. Cell duplication was continuously monitored throughout the 5 days of growth by measuring  $\text{OD}_{720\text{nm}}$  (Figure 1A). Cultures exposed to LL showed reduced growth compared to ML because of the light limitation (Simionato et al., 2011; Sforza et al., 2012). Cells under HL conditions also showed slower growth compared to ML, suggesting a photo-inhibitory effect (Figure 1A) (Simionato et al., 2011; Sforza et al., 2012), as supported by the reduction in PSII quantum efficiency ( $\Phi_{\text{PSII}}$ ) observed under the higher light regime (Figure 1B). Growth differences in response to light availability were also confirmed by measuring cell concentration (Figure 1C) and dry weight (Supplemental Figure 1) after 5 days, towards the end of the exponential phase. Taken together, these results clearly show that cells in LL are in limiting light conditions, while in HL they experience light excess. Nile-Red staining showed that after 5 days of light treatment, neutral lipid accumulation was induced by four fold in the ML and HL cells compared to LL (Figure 1D).



**Figure 1. Light effect on the growth and lipid accumulation of *Nannochloropsis gaditana* cells.** A) Growth curves of wild-type cells during the five days of differential light treatment. Squares, low light; circles, medium light; triangles, high light. At day 5, cells were harvested to measure the

following parameters: B) Photosystem II quantum yield (Fv/Fm, Empty squares) and total chlorophyll content (White bars) by absorption spectra analysis of DMF total extracts, C) Cell concentration and D) Neutral lipid accumulation determined by fluorescence spectra analysis of Nile-red stained cells. LL (Low Light), ML (Medium Light) and HL (High Light). Asterisks indicate statistical significant differences - one-way analysis of variance (ANOVA), p-value < 0.05.

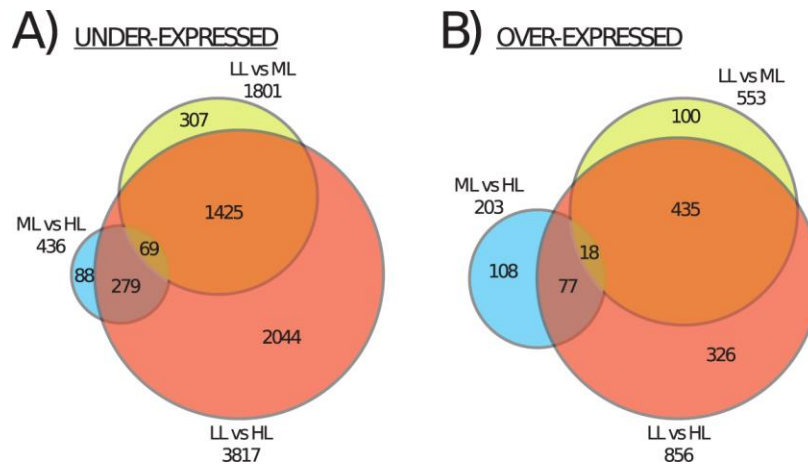
This result is consistent with the literature showing enhanced TAG accumulation for *Nannochloropsis* in response to increased light intensity (Sforza et al., 2012; Pal et al., 2011; Teo et al., 2014). To better understand how light intensity affects the *N. gaditana* metabolism and induces heightened lipid accumulation, cells grown for 5 days under different light intensities were subjected to transcriptomic, metabolomic, lipidomic and functional analyses, using the same biological samples.

#### *Genome-wide changes in mRNA abundance in response to light*

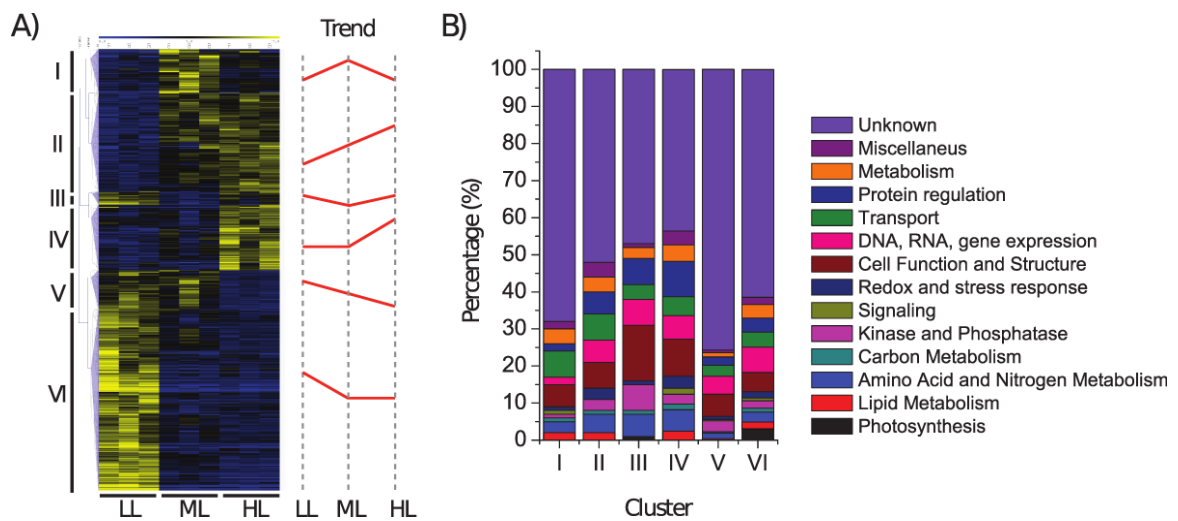
An mRNA-Seq experiment was performed with three biological replicates for each sample to measure the response of *N. gaditana* to the three different light intensities. Out of 10,919 total predicted nuclear genes (Corteggiani Carpinelli et al., 2014), transcripts of 10,456 genes were detected under the three conditions. For the functional classification of light response, we chose those genes showing statistically significantly different expression, as described in the Methods section. A total of 2,354 genes were differentially expressed between LL- and ML-grown cells and 4,673 between LL- and HL-grown cells. A lower number of genes (639) were differentially expressed between ML and HL-grown cells, suggesting that overall the ML transcriptome was closer to HL than to LL conditions, an observation supported by the corresponding MDS plot (Supplemental Figure 2). Out of the 639 genes differentially regulated between ML and HL, 146 (23%) were specific to this comparison, while the others were also differentially expressed between LL and HL. Similarly, out of the 2,354 LL vs ML regulated genes, most were also differentially expressed between LL and HL, and only 398 (16%) were specific to this comparison, suggesting that most of the genes differentially regulated in LL vs ML and ML vs HL were subgroups of the LL vs HL response (Figure 2). Only 69 genes were under-expressed in both conditions, representing the core of genes progressively activated by light availability (LL<ML<HL; Figure 2A, Supplemental Data Set 1A) and 18 genes were over-expressed in both conditions, representing genes gradually turned off by increasing light intensity (LL>ML>HL; Figure 2B, Supplemental Data Set 1B).

The differentially expressed genes were clustered into 6 groups on the basis of their expression levels (Figure 3). Clusters I, IV and VI contains genes specifically over-expressed only in ML, HL and LL, respectively. Genes highly expressed in both LL and HL are found in cluster III, while genes

progressively induced from LL to HL or progressively repressed from LL to HL form clusters II and V (Figure 3A).



**Figure 2. Venn diagrams of differentially expressed genes identified in the three light conditions.** A) Genes under-expressed in LL vs ML are compared with genes under-expressed in ML vs HL and LL vs HL. B) Genes over-expressed in LL vs ML are compared with genes over-expressed in ML vs HL and LL vs HL. LL, low light. ML, medium light. HL, high light.



**Figure 3. Hierarchical clustering of light regulated genes.** A) A transcriptional profile dendrogram was created by using TMeV 4.3 software. Six groups have been highlighted on the left side of the picture. The red lines on the right side represent schematically the overall trend of light response of the genes belonging to the corresponding cluster. B) Manual annotation and category distribution among clusters. Manual annotation was done on the base of gene description and gene ontology (biological process). The genes of panel A were grouped according to their function in 14 categories: photosynthesis, lipid metabolism, amino acid and nitrogen metabolism, carbon metabolism, kinase and phosphatase, signaling, redox and stress response, cell function and structure, DNA/RNA/gene expression, transport, protein regulation, metabolism, miscellaneous, unknown. Protein domains that could not be classified in any specific category were grouped into



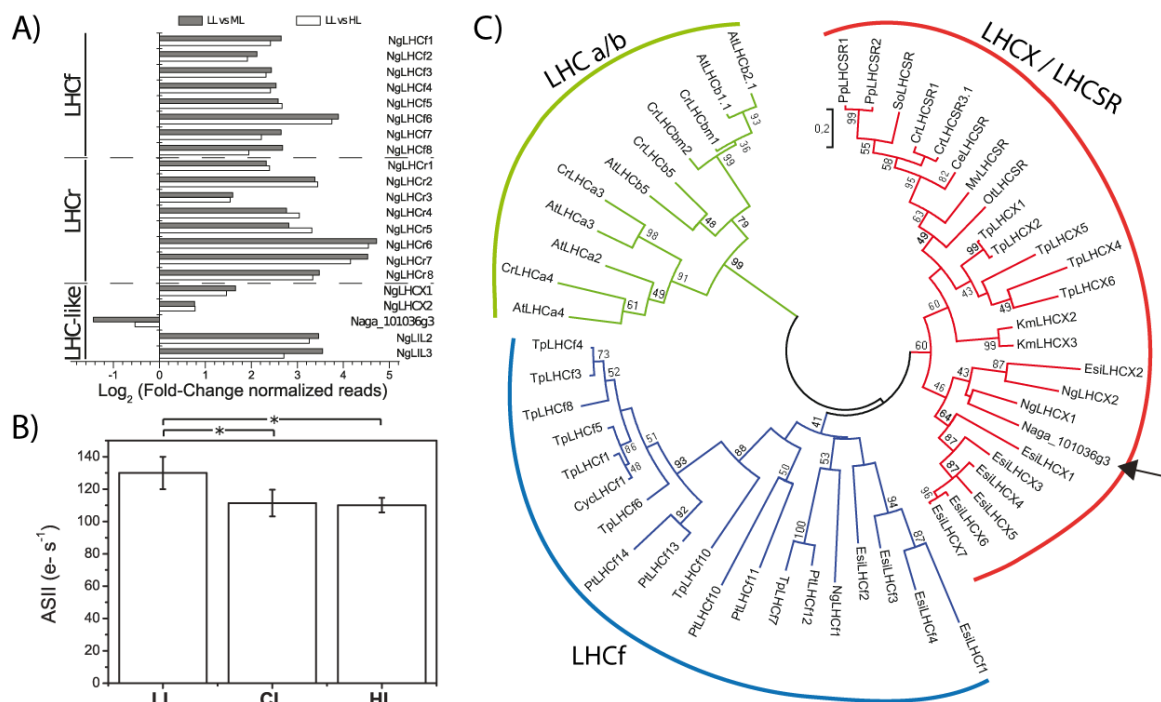
the “miscellaneous” category. The “unknown” category refers to protein sequences for which no consensus was reachable through annotation. Percentage of each functional category is represented in the total numbers of differentially expressed transcripts from each individual set of data. LL, low light. ML, medium light. HL, high light.

For clusters II and VI, we obtained results after a GO enrichment analysis using the blast2go software (Fisher’s test, FDR < 5%). Cluster II is enriched in GO:0019643 (reductive tricarboxylic acid cycle), GO:0006767 (water-soluble vitamin metabolic process) and GO:0046364 (monosaccharide biosynthetic process) while the genes of cluster VI are well represented by GO:0015074 (DNA integration), GO:0015995 (chlorophyll biosynthetic process), GO:0016117 (carotenoid biosynthetic process), GO:0006636 (unsaturated fatty acid biosynthetic process) and GO:0009765 (photosynthesis, light harvesting). All other clusters included genes involved in various functions and are not enriched in any specific category. To determine if there was a global biological function associated with a given expression pattern, we implemented this analysis by manual annotation, distributing the genes of each cluster into fourteen major functional groups (Figure 3B; Supplemental Figure 3). In all cases, genes with unknown function were the most represented class, constituting between 44 and 73% of the genes in each cluster, showing that the function of a large fraction of *N. gaditana* genes cannot be easily deduced from annotation. For other functional groups with a given annotation, useful information could be inferred from the identification of the gene clusters where they were relatively enriched. Genes involved in photosynthesis were enriched in cluster VI and therefore up-regulated in LL compared to ML and HL. Genes involved in lipid metabolism were instead more abundant in clusters I, II and IV, meaning they are up-regulated in ML and HL compared to LL, in accordance with the higher lipid accumulation observed in Figure 1D. Genes involved in redox and stress response were more abundant in clusters II and IV, suggesting the activation of an oxidative stress response at the transcriptional level in *N. gaditana* exposed to excess light (Figure 3; Supplemental Figure 3). Starting from the results of cluster analysis, a selection of gene families and metabolic pathways were analyzed in greater detail, complementing the gene expression data with metabolic and functional analyses.

#### *Modulation of photosynthetic apparatus in response to different light intensities*

Photosynthetic organisms react to different light regimes by modulating their photosynthetic apparatus both in protein and pigment content (Maxwell et al., 1994; Bailey et al., 2001; Ballottari et al., 2007; Dietz, 2015). As mentioned above, in ML and HL conditions, we observed a general down-regulation of genes involved in photosynthesis and light harvesting (GO:0009765, photosynthesis and light harvesting; Figure 3). We also consistently observed a strong and

generalized decrease in the expression of genes of the Light Harvesting Complex (LHC) superfamily, which encodes the proteins constituting the antenna system of the photosynthetic apparatus in eukaryotes. The only exception to this general behavior was the gene *Naga\_101036g3*, which was significantly up-regulated in ML and HL with respect to LL conditions (Figure 4A; Supplemental Data Set 2A). The biological function of LHC proteins can be measured by determining their efficiency in harvesting light and transferring it to the reaction center by monitoring the functional antenna size of PSII (ASII) based on chlorophyll fluorescence. ASII decreased in cells grown in ML or HL compared to LL conditions (Figure 4B), confirming that the down-regulation of the LHC superfamily observed at the transcriptome level corresponds to a functional decrease in light harvesting efficiency. Overall, this observation suggests that LHC expression is enhanced in LL-treated cells of *N. gaditana* to maximize light harvesting capacities.



**Figure 4. Regulation of LHC complexes.** A) Gene expression regulation of LHC superfamily members. Grey and white bars represent the fold change of normalized reads of LL vs ML and LL vs HL, respectively. For each sample, the average of normalized reads of three repetitions was used to calculate the fold change, which is expressed in  $\log_2$  scale. All genes except *Naga\_100056g42* (*NgLHX2*) are significantly regulated in response to light in the reported comparisons. B) Functional antenna size of PSII measured by fluorescence induction. C) Phylogenetic tree of LHC superfamily. *Naga\_101036g3* gene model was translated to the corresponding protein sequence for a tBLASTn analysis in the NCBI nucleotide collection. To attribute *Naga\_101036g3* to a specific subgroup of LHC, its primary sequence was compared to a larger number of LHC protein sequences, notably 11 LHC proteins of chlorophyll a/b-containing organisms, 19 LHCf proteins and 23 LHCX/LHCSR photoprotective proteins from different algae. A multiple alignment of these 52 full-length protein sequences was performed using MUSCLE

(Multiple Sequence Comparison by Log-Expectation) and the aligned protein sequences were used to construct an unrooted maximum likelihood phylogenetic tree utilizing MEGA6.0. Sequence alignment and details are given in Supplemental Figure 5.

LHC proteins from *Nannochloropsis* can be classified into three subgroups (Vieler et al., 2012), with LHCF and LHCR composing the main antenna proteins, as in diatoms and brown algae, and found associated predominantly but not exclusively with the PSII and PSI complexes (De Martino et al., 2000; Beer et al., 2006; Ikeda et al., 2013; Lepetit et al., 2007). A third group, LHCSR or LHCX, instead consists of antennas involved in photoprotection and stress response in several algae species (Peers et al., 2009; Zhu and Green, 2010; Richard et al., 2000) and in the moss *Physcomitrella patens* (Alboresi et al., 2010). Naga\_101036g3 was the only sequence down-regulated in ML and HL. To infer its putative function from its primary sequence, protein BLAST searches retrieved sequences annotated as LHCSR in the top hits (e-value  $< 10^{-34}$ ) (first accession was XP\_001768071 from *Physcomitrella patens*). This assignment is confirmed by the phylogenetic tree, showing that Naga\_101036g3 appeared to be part of the LHCX/LHCSR subgroup together with two other putative LHCX proteins of *N. gaditana* (i.e., Naga\_100173g12 and Naga\_100056g42, respectively NgLHCX1 and NgLHCX2) (Figure 4C). The similarity of Naga\_101036g3 to the LHCX proteins and its peculiar up-regulation in stronger light (Figure 4A) suggest that this gene may have a specific role in response to strong light intensities. Interestingly, however, two other LHCX genes have been identified in the *N. gaditana* genome, but they exhibit similar expression profiles to other antennas (Figure 4A; Supplementary Data Set 2A).

#### *Influence of light regime on isoprenoid composition*

Isoprenoids are a group of highly diverse metabolites, including carotenoids, tocochromanols, quinones, prenyl groups and sterols, derived from a common precursor, isopentenyl-pyrophosphate (IPP). Some isoprenoids, such as carotenoids or tocochromanols, are important antioxidant molecules and are expected to be critical compounds in stress such as like high illumination.

Plants are usually able to synthesize IPP both in the cytosol *via* the action of a 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase), producing mevalonate, or by a non-mevalonate pathway in the stroma of chloroplasts, using pyruvate and glyceraldehyde-3-phosphate (GAP), this latter intermediate deriving from the isomerization of dihydroxyacetone phosphate (DHAP). The genome of *Nannochloropsis* does not contain any HMG-CoA reductase gene (Lu et al., 2014b), and the production of isoprenoids therefore should only rely on IPP and geranylgeranyl pyrophosphate

(GGPP) produced in the plastid and thus on the availability of GAP and DHAP in this organelle (see below).

Genes involved in the carotenoid biosynthetic pathway (GO:0016117) were identified, starting from sequences already annotated in *Solanum lycopersicum* (Sato et al., 2012) and in *N. oceanica* (Lu et al., 2014a) (Supplemental Data Set 2B, Supplemental Figure 4A). Genes of this pathway are enriched in cluster VI (Figure 3), and most of them are significantly up-regulated in LL compared to ML, except for *carotenoid isomerase 3* (*CRTISO3*, Naga\_100023g33) ( $FC_{LL\ vs\ ML} = -0.50$ ) and *carotenoid cleavage dioxygenase 4* (*CCD4*, Naga\_100050g23), which are instead up-regulated in ML compared to LL ( $FC_{LL\ vs\ ML} = -2.03$ ). A similar general up-regulation was also observed for LL vs HL (Supplemental Data Set 2B), with two additional carotenoid cleavage dioxygenases up-regulated in HL (*CCD1*, Naga\_100022g32 and *CCD2*, Naga\_100083g 22), suggesting an intriguing role for apocarotenoids in response to strong light. Another indication of the regulation of the carotenoid pathway is given by the up-regulation in ML and HL of plastid terminal oxidase (*PTOX*, Naga\_100066g11) (Supplemental Data Set 2B), a protein involved in carotenoid biosynthesis (Wang et al., 2009) but also reported to play a role in chlororespiration and possibly photoprotection (Cournac et al., 2000; Houille-Vernes et al., 2011).

In agreement with the general reduction of gene expression in the biosynthetic pathway, the total amount of carotenoids decreased in proportion to increasing light intensity (Table I).

**Table I. Changes in chlorophyll and carotenoid content in *Nannochloropsis gaditana* cells at different light regimes.** LL, low light. ML, medium light. HL, high light. Each sample was significantly different from the other (one-way analysis of variance (ANOVA) plus Tukey's t-test, p-value < 0.05).

Sample	Chl [pg]/Cell	Car [pg]/Cell	Chl/Car
LL	0.133 ± 0.006	0.0389 ± 0.0012	3.42 ± 0.05
ML	0.063 ± 0.005	0.0239 ± 0.0017	2.65 ± 0.04
HL	0.038 ± 0.006	0.0184 ± 0.0018	2.04 ± 0.14

This result is consistent with the fact that carotenoids are also structural elements of the photosynthetic apparatus and are thus more abundant in LL than ML and HL. Nevertheless, the ratio between chlorophyll and carotenoid content after 5 days of LL, ML and HL treatment revealed a decrease in the chlorophyll/carotenoid (Chl/Car) ratio in response to light (Table I), suggesting a relative enrichment in carotenoids per photosynthetic unit, consistent with the carotenoid role in protection from excess light (Li et al., 2009; Cazzonelli and Pogson, 2010).

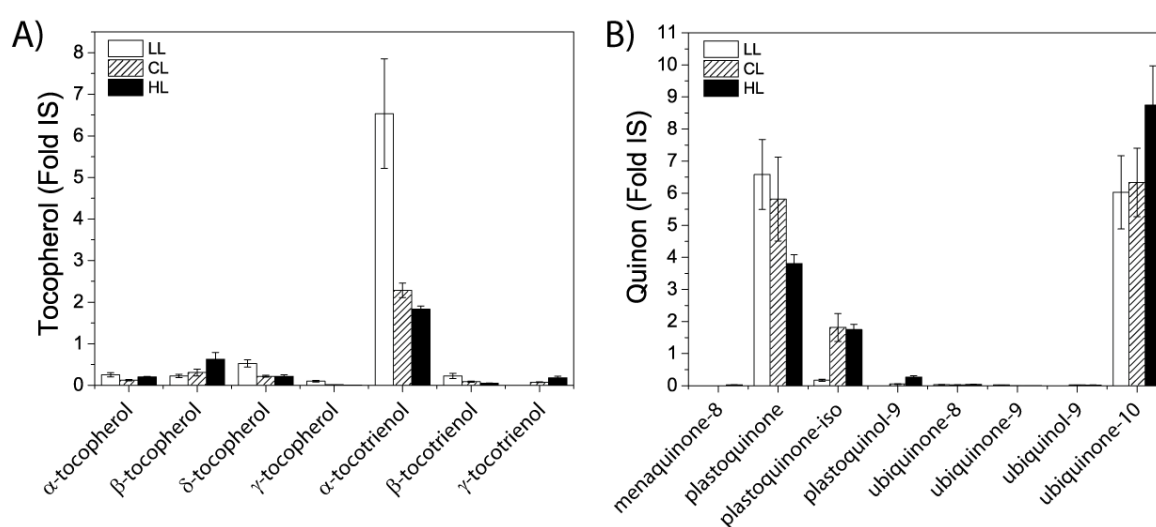
Analysis by HPLC coupled with high-resolution mass spectrometry (LC-HRMS) revealed a general decrease in accumulation for major carotenoids (violaxanthin, vaucherixanthin,  $\beta$ -carotene) per cell in response to increasing intensity of light treatment, in agreement with the repression of carotenoid pathway genes described above (Supplemental Figure 4B). Certain carotenoid species, such as antheraxanthin, zeaxanthin, astaxanthin and canthaxanthin (Supplemental Figure 4B), were more abundant in ML and/or HL than LL, which is consistent with a role in protection from oxidative stress (Baroli et al., 2003; Lemoine and Schoefs, 2010).

Chlorophyll synthesis also relies on isoprenoid precursors for the synthesis of the phytol side chain linked to a magnesium-containing tetrapyrrole deriving from protoporphyrin IX. Similarly to carotenoids, chlorophyll biosynthetic pathway genes (GO:0015995) were highly represented in cluster IV, including genes down-regulated in ML and HL conditions compared to LL. Important genes in the early part of the pathway were in fact down-regulated in higher light regimes (Supplemental Data Set 2C), e.g., glutamyl-tRNA reductase (*GluTR*, Naga\_100125g4;  $FC_{LL\ vs\ ML} = 1.11$ ), 5-aminolevulinate dehydratase (*ALAD*, Naga\_100010g65;  $FC_{LL\ vs\ ML} = 1.00$ ), porphobilinogen deaminase (*PBGD*, Naga\_100003g83;  $FC_{LL\ vs\ ML} = 1.05$ ) and coproporphyrinogen-III oxidase (*CPOX-1*, Naga\_100665g2;  $FC_{LL\ vs\ ML} = 1.31$ ). This result is in agreement with the lower chlorophyll content in ML and HL samples compared to LL, as well as with the increased content of the 5-aminolevulinate precursor. Additionally, several chlorophyll catabolites such as pheophytin a (not detectable in LL cells), pheophorbide a and the primary fluorescent chlorophyll catabolite (limited to ML/LL) were increased at higher light intensities (Supplemental Table 1). The regulation of the lower part of the pathway showed instead a marginal dependence on light intensity. Two chlorophyll synthase genes were even more highly expressed in HL than in LL (*CHLG*, Naga\_101224g2 and Naga\_100989g1) (Supplemental Data Set 2C).

The lower chlorophyll content was consistent not only with the down-regulation of most chlorophyll synthesis genes but also with the up-regulation in HL of pheophorbide  $\alpha$  oxygenase (*PAO*, Naga\_100072g25), a gene involved in chlorophyll degradation. In addition, Fe-chelatase (Naga\_100010g10), a gene involved in the biosynthesis of an iron-containing heme rather than a magnesium-containing protoporphyrin IX, is up-regulated in HL (Supplemental Data Set 2C), possibly diverting the general synthesis of tetrapyrroles away from chlorophylls under these conditions.

Tocochromanols are lipid-soluble isoprenoids, including tocopherols and tocotrienols, belonging to a group of vitamin-E molecules that play, such as carotenoids, a protective role under oxidative stress. They are essential in the human diet, are synthesized only by photosynthetic organisms and may represent a precious by-product of mass cultures. In ML and HL, we observed the down-

regulation of p-hydroxyphenylpyruvate dioxygenase (*HPPD-1*, Naga\_100291g3) and homogentisate phytyltransferase (*HPT/VTE2*, Naga\_100225g6), two early genes in the synthesis of both tocopherols and tocotrienols, while 2-methyl-6-phytyl-1,4-benzoquinone methyltransferase (*VTE3*, Naga\_100062g16), which is responsible for the synthesis of  $\alpha$ -tocopherol, is up-regulated in ML and HL (Supplemental Data Set 2D). We measured the levels of tocopherols, and the most abundant species in all three light conditions is  $\alpha$ -tocotrienol. At the metabolite level, the increase in light intensity resulted in a reduced accumulation of all the tocopherols, with the exception of  $\beta$ -tocopherol, which was significantly increased in ML and HL compared with LL. Interestingly,  $\gamma$ -tocotrienol, which was not detectable in LL, showed a gradual accumulation in ML and HL. In some cases, we could observe a strict correlation between the accumulation of these metabolites and the expression levels of genes involved in their biosynthetic pathways. For example, *HPPD* and *HPT/VTE2* are down-regulated, in agreement with the decrease of most of the down-stream metabolites; in contrast, both  $\alpha$ -tocopherol and  $\alpha$ -tocotrienol respond negatively to light intensity, while *VTE3* is up-regulated (Figure 5A).



**Figure 5. LC-APCI(+)-MS analyses of tocopherols and quinones in *Nannochloropsis gaditana* cells grown in different light conditions.** A) tocopherols and B) quinones. Data are expressed as relative content with respect to the level of the internal standard ( $\alpha$ -tocopherol acetate). Data are the average of at least 4 extractions. For more details, see methods.

The antiparallel regulation of the main gene for  $\alpha$ -tocopherol biosynthesis and of its metabolic products might reflect the increased depletion of  $\alpha$ -tocopherol acting as ROS scavengers in HL conditions and the consequent enhancement of their biosynthesis through a hitherto undescribed feedback loop.

The two main quinone molecules in *N. gaditana* are benzoquinones: plastoquinone and ubiquinone 10, which are involved, respectively, in photosynthetic and mitochondrial electron

transport (Figure 5B) (Nowicka and Kruk, 2010). At high light intensities, the total pool of plastoquinone, the primary PSII electron acceptor, is decreased, in agreement with the global reduction of the photosynthetic apparatus, and its redox status is shifted more towards the reduced form, plastoquinol. It is therefore possible that, like in green algae and diatoms (Escoubas et al., 1995; Lepetit et al., 2013), also in *N. gaditana* the redox state of the plastoquinone pool may serve as a signal triggering acclimation to high light.

Ubiquinone, the primary electron acceptor for NADH dehydrogenase in mitochondria, is increased under HL conditions (Figure 5B). It is notable that the primary route for FA biosynthesis is the synthesis of Acetyl-CoA from pyruvate by the Pyruvate Dehydrogenase Complex (PDHC). *Nannochloropsis* contains both mitochondrial and chloroplast forms of PDHC and the latter has been suggested to be responsible for Acetyl-CoA biosynthesis under N starvation conditions (Li et al., 2014). It is interesting to evidence that this is not the case here where there is no specific regulation of chloroplast PDHC, while mitochondrial genes are up-regulated in HL (Supplemental Data Set 3A).

A consistent regulation of genes involved in quinone biosynthesis is observed in LL vs HL (Supplemental Data Set 2D). Menaquinone biosynthesis genes display the greater extent of alterations: naphthoate synthase (*MenE*, Naga\_100074g14) is down-regulated in response to light, while two long-chain-fatty-acid CoA ligases (*MenB*, Naga\_100021g41 and Naga\_100713g1) are up-regulated. Additional perturbations regard the over-expression of two ubiquinone biosynthesis genes, the 3-demethylubiquinone-9 3-methyltransferase (*Coq3*, Naga\_100398g3) and the ubiquinone biosynthesis hydroxylase (*Coq6*, Naga\_100006g73).

Overall, transcription and metabolic data concur in suggesting that the biosynthesis of isoprenoids, either acting under oxidative stresses such as carotenoids and tocopherols or as precursors for chlorophylls and electron carriers such as quinones, are strongly affected by the light regime. Most of the molecules are down-regulated in HL, especially ones playing a role correlated with photosynthesis. Some interesting exceptions nevertheless occur, suggesting a specific function in stress response (e.g., canthaxanthin is more abundant in HL).

#### *Remodeling in vitamin and nucleic acid metabolism*

Gene Ontology enrichment analysis indicated the presence of two classes of transcripts associated with vitamin (GO:0006767, water-soluble vitamin metabolic process) and nucleic acid (GO:0015074, DNA integration) metabolism. Genes in these categories are up-regulated in LL over ML and HL cells and include phosphomannomutase (Naga\_100096g21), L-galactono- $\gamma$ -lactone dehydrogenase (Naga\_100087g11) and aldo-keto reductase (Naga\_100011g89), involved in

ascorbic acid biosynthesis; dihydrofolate reductase (Naga\_101645g1), involved in folate metabolism; dephospho-CoA-kinase (Naga\_100004g90) and phosphopantothencysteine decarboxylase (Naga\_100093g19), playing a role in pantothenic acid catabolism; pyridoxamine 5-phosphate oxidase (Naga\_100044g35 and Naga\_100634g3) and phosphoserine aminotransferase (Naga\_100013g61), implicated in pyridoxine/pyridoxal synthesis; four soluble pyridine nucleotide transhydrogenases (Naga\_102355g1, Naga\_102104g1, Naga\_100100g18 and Naga\_101301g1), playing a role in nicotinamide metabolism; 6,7-dimethyl-8-ribityllumazine synthase (Naga\_100004g9), participating in riboflavin biosynthesis; and thiamine-phosphate diphosphorylase (Naga\_100104g3), coding for an enzyme of the catabolism of thiamine diphosphate, a cofactor of PDHC. Simultaneously, we measured the level of 11 vitamins and vitamin intermediates by LC-MS (Supplemental Data Set 4). ML and HL cells exhibited a general increase in all vitamin-related metabolisms, with the exceptions of nicotinamide and pyridoxal, which were reduced. In general, we could not demonstrate a strong correlation between gene expression and metabolite accumulation. Because many vitamins are strong antioxidants, we can hypothesize that the generalized increase in their content could serve to scavenge toxic ROS to reduce photooxidative stress.

In the DNA integration class, we did not find regulated genes involved in nucleic acid metabolism but rather transcripts using nucleotide-sugar substrates for a series of reactions (NAD-dependent epimerase dehydratase, Naga\_100235g1) and retrotransposons (ten gag-pol polyprotein: Naga\_101449g1, Naga\_100433g2, Naga\_100497g3, Naga\_100549g4, Naga\_100294g5, Naga\_100294g4, Naga\_101142g1, Naga\_100055g1, Naga\_101002g3, Naga\_101146g2; a copia-like retrotransposable element, Naga\_101291g1; retrovirus-related pol polyprotein from transposon tnt 1-94, Naga\_102348g1; retrotransposon unclassified, Naga\_101693g1; copia-type polyprotein, Naga\_100999g3). At the metabolite level, we observed a general over-accumulation of a series of molecules involved in nucleic acid metabolism, including nucleotides (adenine, cytosine, uracil), nucleosides (adenosine, cytidine, pseudo-uridine) and pyridine intermediates (picolinic acid). With certain exceptions, guanine and nucleoside phosphates (adenosine mono- and di-phosphates) and nucleic acid intermediates seem to exhibit stronger accumulation in ML, which is consistent with these conditions having the highest growth rate and similar to previous observations in other algae, where nucleic acid intermediates were strongly accumulated in cells with enhanced growth rates (Goyal et al., 1988; Allen et al., 2011). Similarly to the observations in these previous studies, we can explain these changes in the general framework of cell exposure to higher light conditions.



### *Additional alterations in the primary metabolome*

We also measured 59 nonpolar (Supplementary Table 1) and 156 polar metabolites (Supplementary Data Set 4). In addition to the classes previously discussed, many additional alterations were observed, mainly at the primary metabolism level. Overall, with few exceptions, we could identify distinct patterns of alterations within the primary metabolome of *N. gaditana* cells in the low and the medium/high light regimes: most amino acids and peptides were over-accumulated in ML/HL over LL cells. This group includes sixteen amino acids (agmatine, alanine, asparagine, carnitine, cystine,  $\gamma$ -butyrobetaine, glutamine, histidine, homospermidine, methionine, methionine sulfoxide, phenylalanine, threonine/homoserine, tryptophan, tryptophan derivative and tyrosine), 3 amines (N-acetylglucosamine, N-carboxyethyl- $\gamma$ -aminobutyric acid and spermidine), and 3 tripeptides (citrulline, gamma-guanidinobutyrate and glycine-betaine). A more variable response between ML/LL and HL/LL was instead revealed in organic acid metabolism. For example, 4 organic acids (5-amino-levulinate, citric acid, methyl-citric acid, phenylpyruvic acid) were over-accumulated in both ML and HL cells, while 2 others (allantoic and shikimic acids) were under-accumulated in ML and HL cells. GO enrichment analysis did not identify classes of genes involved in these biosyntheses, suggesting the possible occurrence of post-transcriptional regulation. It is relevant to mention that similar metabolic changes were observed in previous studies with *Chlamydomonas reinhardtii* under light stress (Davis et al., 2013), day/night transition (Willamme et al., 2015) or nitrogen starvation (Wase et al., 2014).

Within the semi-polar metabolome, we could identify some “metabolic markers” associated with lipid accumulation: citric acid, undetectable in LL cells, and methyl-citric acid, increased in both ML and HL: because citric acid can serve as a carrier of acetyl units for fatty acid synthesis, its accumulation could be explained on the basis of a cellular response to provide more precursors for lipid production and storage.

### *Lipid biosynthesis*

Most of the genes putatively involved in lipid biosynthetic pathways have been previously annotated for *N. gaditana* and *oceanica* (Corteggiani Carpinelli et al., 2014; Vieler et al., 2012; Li et al., 2014), showing that lipid biosynthesis in *Nannochloropsis* follows a complex pathway extended across several compartments. In photosynthetic eukaryotes, fatty acids (FA) are synthesized in the chloroplast stroma; thylakoid lipids (monogalactosyldiacylglycerol, MGDG; digalactosyldiacylglycerol, DGDG; sulfoquinovosyldiacylglycerol, SQDG and phosphatidylglycerol, PG) are synthesized in chloroplast envelope membranes; and major phospho-glycerolipids (e.g.,

phosphatidylcholine, PC and phosphatidylethanolamine, PE), betaine lipids (diacylglyceryltrimethylhomoserine, DGTS) and triacylglycerols (TAG) are synthesized in the ER (Petroutsos et al., 2014). Based on genomic evidence, the metabolic pathways for the production of glycerolipids are expected to follow this general scheme in *Nannochloropsis*.

However, *Nannochloropsis* also differs from other photosynthetic eukaryotes in a remarkable feature of FA synthesis, as genes coding for cytosolic proteins that could act as a fatty acid synthase of type 1 (FAS1) or a polyketide synthase (PKS) have also been identified (Vieler et al., 2012). Based on a phylogenetic analysis using three FAS1/PKS-like gene sequences from *N. oceanica*, it was recently proposed that these proteins were more similar to fungal polyketide synthases than to FAS1 (Poliner et al., 2015).

In cells exposed to different illumination intensities, we observed the statistically significant down-regulation of several lipid biosynthesis genes coding for proteins localized in the chloroplast. In fact, genes coding for components of the chloroplast fatty acid synthase of type 2 (FAS2) were less expressed in ML and HL (BXP1, Naga\_100594g3; MCT, Naga\_100046g34; KAR1, Naga\_100037g12; HAD1, Naga\_100113g7, ENR1, Naga\_101525g2) (Supplemental Data Set 3B). By the same token, the gene coding for cytosolic acetyl-CoA carboxylase (ACC1, Naga\_100605g1) were down-regulated under the ML and HL conditions. In contrast with the down-regulation of genes encoding plastid FAS2 components, we detected the up-regulation in HL of two beta-ketoacyl synthases (Naga\_102909g1 and Naga\_101811g1) and five polyketide synthases (Naga\_102303g1, Naga\_102722g1, Naga\_101380g1, Naga\_103117g1 and Naga\_100022g68), all predicted to be localized in the cytosol (Supplemental Data Set 3B), suggesting a re-modulation of lipid biosynthesis with cytosolic fatty acid synthesis, increasing its relevance at the expense of the chloroplast.

Second, we detected a general up-regulation of genes involved in thylakoid lipid assembly in the chloroplast, i.e., lyso-phosphatidic acid and phosphatidic acid (*GPAT*, also known as *ATS1*, Naga\_100562g3), diacylglycerol (*PAP*, Naga\_100075g9), PG (*CDS1*, Naga\_100056g35; *PGPS*, Naga\_100114g5), SQDG (*SQD1*, Naga\_100038g9) and galactolipids (*MGD1*, Naga\_100092g7) (Supplemental Data Set 3B). This feature is different from what was reported very recently in *N. oceanica* along a circadian rhythm (Poliner et al., 2015), in which the expression of the plastid glycerolipid genes (genes involved in the synthesis of DAG, MGDG, DGDG and SQDG) displayed high coordination with FA synthesis genes at the transcriptional level so that their expression preceded the observed lipid accumulation during the daylight period. Although we cannot exclude the existence of a previous unrecorded phase of upregulation of FAS2 genes, the HL regime seems to disrupt this coordination.

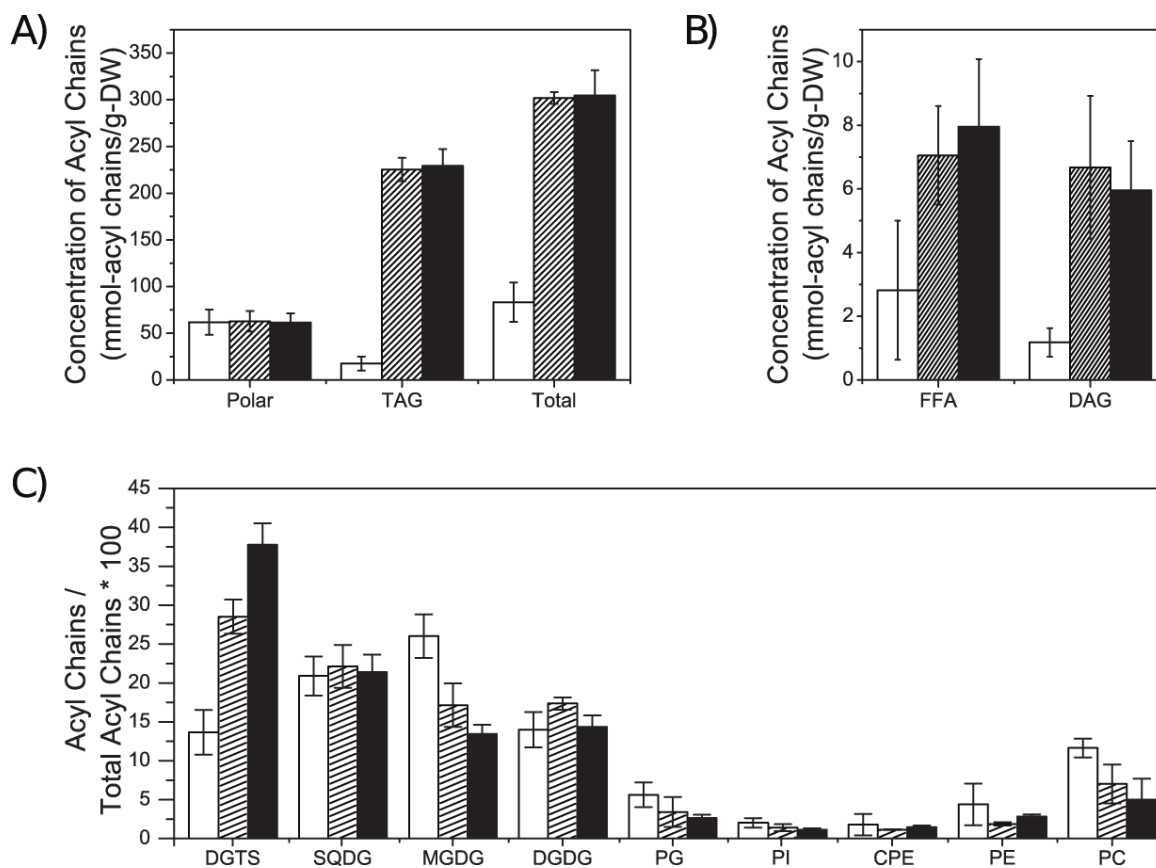
Third, in membrane glycerolipid biosynthesis in the ER, two putative genes coding for enzymes producing phospho-choline (Naga-100331g3 and Naga\_100090g21), the polar head used for PC synthesis, were less expressed in ML and HL. By contrast, two genes putatively encoding betaine lipid synthesizing enzymes (BTA1, Naga\_100016g36 and Naga\_100004g72) were up-regulated. In storage glycerolipid biosynthesis in the ER, instead, we detected a significant up-regulation of the expression of several diacylglycerol-acyltransferase (DGAT) genes in ML and HL compared to LL, suggesting increased synthesis of these lipids. More precisely, *DGAT1* (Naga\_101968g1) is the most over-expressed *DGAT* gene in response to increasing light ( $FC_{LL\ vs\ ML} = -1.12$ ) (Supplemental Data Set 3B). Of the 11 *DGAT2* genes present in the genome, three were more highly expressed in ML and HL than LL (*DGAT2a*, Naga\_100343g3; *DGAT2b*, Naga\_100010g31; *DGAT2f*, Naga\_100006g86), while two are more specifically induced in HL compared to LL (*DGAT2h*, Naga\_100010g4; *DGAT2k*, Naga\_100004g173). A phospholipid: diacylglycerol acyltransferase that acts by transferring an acyl group from a phospholipid to DAG, thus producing TAG, is also up-regulated in ML and HL (*PDAT*, Naga\_100065g17) (Supplemental Data Set 3B). Taken together, these data suggest a general increase in ER-localized TAG biosynthesis.

The sub-group of desaturases showed a peculiar behavior, with expression levels peaking in ML conditions and lower expression in both LL and HL (Supplemental Data Set 3B). Fatty acid desaturation is usually initiated in the stroma of chloroplasts by the palmitoyl/stearoyl-acyl carrier protein  $\Delta 9$ -desaturase, introducing one double bond into 16:0 or 18:0 fatty acids (producing 16:1 or 18:1, respectively). In *N. gaditana*, we detected one gene coding for such a desaturase (*FAD2*, 100013g52), showing a lower expression level in ML and HL. Other genes coding for putative desaturases show a general decrease in ML and HL (Naga\_100027g27, Naga\_100061g21, Naga\_100273g7 and Naga\_100545g1). By contrast, a gene coding for a putative palmitoyl- $\Delta 3$ -desaturase, (*FAD4*, 100028g29) acting specifically on PG, is up-regulated in ML and HL.

Several genes classified as involved in lipid catabolism are expressed more in HL than in LL but also in ML compared to LL (Supplemental Data Set 3B), with the strongest upregulation observed for a lipase ( $FC_{LL\ vs\ ML} = -2.37$ ,  $FC_{LL\ vs\ HL} = -3.07$ ; Naga\_100040g9). Of 76 putative genes involved in lipid catabolism, only 3 are expressed at higher levels in LL than ML or HL, namely one dehydratase (Naga\_100034g21), one AMP-dependent synthetase and ligase (Naga\_100012g66) and one phospholipase A2 (Naga\_100247g4). As recently reported for the oleaginous diatom *Fistulifera solaris* under N shortage (Tanaka et al., 2015), some genes involved in fatty acid beta-oxidation in the mitochondria have a higher expression level in conditions that trigger TAG accumulation, i.e., in ML and HL (*KAS4*, Naga\_100004g8, *HAD2*, Naga\_100486g4, *ENR2*, Naga\_100016g22).

In summary, gene expression analysis suggests that lipid biosynthesis is down-regulated in the chloroplast in response to increased light, while genes involved in cytoplasmic fatty acid synthesis as well as the insertion of fatty acids into TAG are induced.

A lipidomic analysis was performed to characterize how lipid accumulation correlated with gene regulation (Figure 6).



**Figure 6. Glycerolipid remodeling in response to light intensity.** Lipid profiling was performed on *Nannochloropsis* cells grown for 5 days under three different light conditions. A) Concentration of total lipids, triacylglycerol and polar lipids and B) concentration of diacylglycerol are reported on two independent panels with different scales. Their concentration is reported as moles of acyl chains on dry weight. C) Relative accumulation of each class of polar lipid on the total amount of polar lipids. White bars, Low Light. Striped bars, Medium Light. Black bars, High Light. Error bars correspond to standard deviation of three biological replicates. DAG, diacylglycerol; DGDG, digalactosyldiacylglycerol; DGTS, diacylglyceryltrimethylhomoserine; CPE, carboxymethyl phosphatidylethanolamine; MGDG, monogalactosyldiacylglycerol; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; TAG, triacylglycerol.

It is worth noting here that the same batches of cells were used for the mRNA-seq, lipidomic and metabolomic analyses to make all the results directly comparable. The total amount of acyl lipids was higher in ML and HL than in LL (Figure 6A), consistently with Nile-Red quantification (Figure

1D). The step between ML and HL did not trigger any significant change in overall lipid content (Figure 6A), suggesting that the light-dependent induction of TAG accumulation is already activated in ML cultures. Interestingly, while the content of polar lipids was not significantly affected, the acyl lipid increase was completely due to the larger TAG accumulation.

This increase in TAG suggests that fatty acid production is activated by increased light intensity, in contrast with the observed down regulation of genes encoding plastid FAS2. This result suggests that the increased production of fatty acids relies on the activity of the cytosolic FAS1 system, most notably on the FAS1/PKS-like protein (Naga\_100022g68), whose expression increases at higher light intensities (Supplemental Data Set 3B).

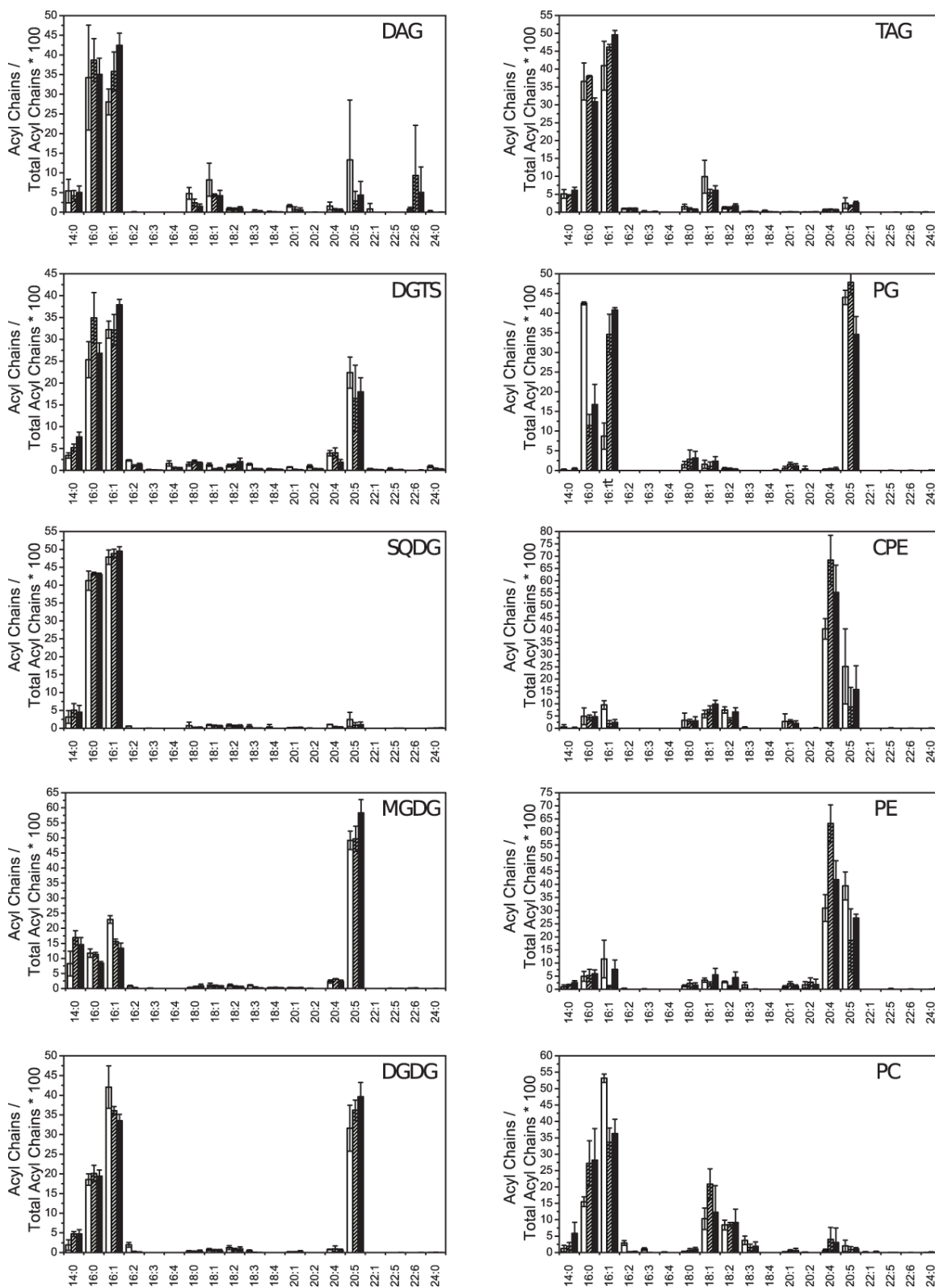
DAGs, the likely precursors of TAG biosynthesis, are accumulated fifty times less than TAG, but the two lipid classes show a similar response to light intensity with the induction of DAG accumulation in ML and HL (Figure 6B). This result suggests that precursors of TAG biosynthesis are accumulated together with the final product. Based on the transcriptome analyses, both the *de novo* synthesis of TAG (*via* DGAT enzymes) and phospholipid recycling (*via* PDAT) are likely to drive TAG accumulation.

We also analyzed the different classes of membrane lipids, observing some significant changes among the three light intensities (Figure 6C). The most notable one is a gradual and strong increase of DGTS from LL to ML and HL. In contrast, PC decreased in HL, possibly counterbalancing the variations in DGTS content (Figure 6C). Concerning plastid lipids, whereas MGDG decreased, DGDG and SQDG remained stable, and PG decreased moderately. This result is extremely interesting because it means that specific polar lipids are differentially accumulated in response to light, even though the total amount of the polar fraction is not altered. It is also worth emphasizing that there is a relationship between the regulation of polar lipid accumulation and the compartmentalization of the corresponding biosynthetic pathway. In fact, DGTS is likely synthesized in the endoplasmic reticulum from DAGs by a BTA enzyme that is up-regulated at the transcriptional level (Supplemental Data Set 3B). As synthesis of PC relies on phospho-choline, its decrease might also be correlated with a decline in the genes coding for enzymes that produce phospho-choline. The synthesis of chloroplast lipids might result from the combination of opposite trends, i.e., a decline in fatty acids synthesized by FAS2 components encoded by down-regulated genes (KAR1, HAD1, ENR1; Supplemental Data Set 3B) and a stable or increased expression of genes involved in thylakoid lipid assembly (GPAT/ATS1, PAP, CDS1, PGPS, SQD1, MGD1, DGD1; Supplemental Data Set 3B), resulting in an overall stability of SQDG and DGDG, a slight decrease in PG and a decline in MGDG.

The specific increase of the betaine lipid triggered by light intensity brings up the question on how DGTS accumulation in HL could be supported if plastid fatty acid biosynthesis is inhibited.

Again, a possible explanation for this finding lies in the identification of cytosolic FAS1/PKS-like genes (Naga\_100022g68, Naga\_100001g49, Naga\_100028g49; Supplemental Data Set 3B), which are also induced in HL. This result suggests that upon light stimulus, DGTS accumulation is also increased by the activation of cytosolic fatty acid biosynthesis, compensating for the decrease in plastid pathway activity. This possibility of supporting lipid biosynthesis in two different compartments provides a high flexibility to *N. gaditana* that would allow it to finely regulate fatty acid biosynthesis in response to environmental stimuli, as observed here. This organism also represents the first described photoautotroph capable of this separation (Petroutsos et al., 2014; Ramakrishnan et al., 2013).

When we analyzed the acyl profiles of each lipid class, some variations in 16:0, 16:1 and 20:5 could be observed in most membrane lipids, with the exception of SQDG.



**Figure 7. Estimated acyl chain composition as a function of lipid class.** DAG, diacylglycerol; DGDG, digalactosyldiacylglycerol; DGTS, diacylglyceryltrimethylhomoserine; CPE, carboxymethyl phosphatidylethanolamine; MGDG, monogalactosyldiacylglycerol; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; TAG, triacylglycerol.

Most lipids with 16:1 acyl-groups are generated through the neo-synthesis of fatty acids in the plastid by the stearoyl-ACP desaturase (actually acting as a palmitoyl-ACP desaturase).

An exception is 16:1 in PG, which is synthesized by an alternative pathway as a *trans*-double bond rather than a *cis*-double bond (=16:1t). Concerning light-dependent regulations, HL induces a relative decrease in 16:1 and an increase in 20:5 in plastid galactolipids (MGDG and DGDG), while LL triggers an opposite effect. We also observed significant changes in PC under HL, with a decrease in 16:1 in favor of 16:0. In PG, the most important change is observed in LL, with an increase in 16:0 at the expense of 16:1t. The 16:1t acyl group of PG was shown to be critical for association with photosystems, particularly the D1 core of PSII (Endo et al., 2015). The increase in PG with higher proportions of 16:1t therefore appears, like the increase in betaine lipid, to be another hallmark of the response to light increase, possibly playing a role in PSII protection.

This result is consistent with the evidence in figure 1B that HL cells suffer from significant PSII photoinhibition, increasing their need for PSII repair. In PE and CPE, which are present in lower proportions than the others (Figure 6), LL and HL both induce an increase of 20:5 content at the expense of 20:4 (Figure 7). It is interesting to note that in HL conditions, there are an excess of electrons and oxidative conditions (Figure 1A-B), but nevertheless, the MGDG and DGDG lipids, which due to being present in the thylakoids are the lipids most exposed to this stress, contain more unsaturated lipid species. Either these species play a role in response to oxidation, or a mechanism protecting these polyunsaturated lipids is established. It is worth mentioning that LL samples are characterized by an increase of 16:1 membrane lipids, probably because the demand for TAG biosynthesis is not strong and the newly synthesized 16:1 lipids are dedicated to feeding MGDG, DGDG and PC.

The acyl distribution of TAG followed the same trend observed in DAGs, with a limited amount of 20:5, suggesting that the fatty acids are not recycled from polar lipids but originate entirely from *de novo* synthesis. This finding is significantly different from the observations in nitrogen depleted cells, where even if neo-synthesis played a major role in sustaining TAG production, there was clear evidence of fatty acid recycling from the polar lipids coinciding with the degradation of thylakoids (Simionato et al., 2013).

#### *Triose, pentose and hexose profiling*

Light is the primary source of energy for photosynthetic organisms, used for the synthesis of all forms of reduced carbon. We therefore analyzed the regulation of the genes belonging to those pathways directly involved in primary carbon metabolism, which will necessarily feed all other metabolic pathways (Supplemental Data Set 3A). Glycolysis and the Calvin-Benson cycle are

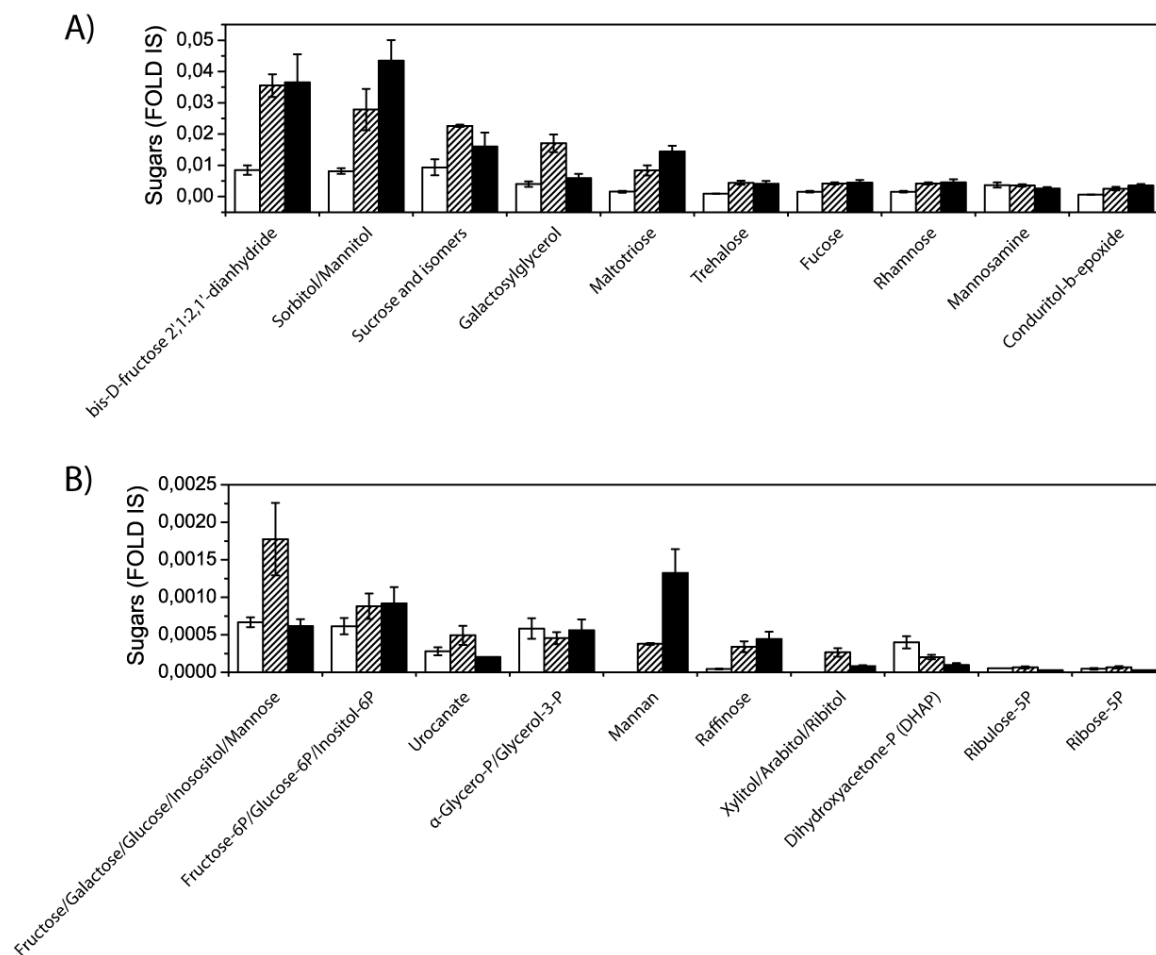


involved in the metabolism of hexose molecules, essential building blocks for all biomolecules. Deciphering the regulation of metabolic response in microalgae is extremely complex because of the overlap among pathways and because of their sophisticated compartmentalization (Johnson and Alric, 2013; Li et al., 2014). The collection of genome-wide data from different environmental conditions is essential to entangle this complex metabolism and its various crosstalks.

Among the significant regulations identified in the conditions analyzed here, glucokinases (*GK*; Naga\_100170g2, Naga\_100119g19) are up-regulated in ML and HL, while the cytosolic phosphoglucose isomerase is strongly up-regulated in LL (*GPI*; Naga\_100305g6). Phosphofructose kinase (*PFK*) and aldolases (*ALDO*) did not show any transcriptional regulation, but a strong response to light was found for the group of 8 fructose bisphosphatase. Two of them (*FBP*; Naga\_100005g122 and Naga\_100159g16) were up-regulated in HL, while three were down-regulated in HL (Naga\_101891g1, Naga\_102887g1, Naga\_100005g28). Interestingly, *in silico* analysis of their coding sequences suggested a cytosolic localization for up-regulated sequences and a plastid localization for down-regulated sequences. High expression of *FBP* in the cytosol could serve for the synthesis of storage carbohydrates (likely 1,3-beta-glucans, known as chrysolaminarin or laminarin in *Nannochloropsis*) (Jia et al., 2015), while its low expression in the plastid could avoid the removal of fructose 1,6-biphosphate from the glycolytic pathway.

Through glycolysis, hexose phosphates are broken down into three-carbon intermediates by the action of aldolases, producing DHAP and GAP, two compounds that are interconverted by the action of triose phosphate isomerases (*TPI*). DHAP and GAP are essential molecules in the metabolism being intermediates in various metabolic pathways located in different compartments of the cell, including glycolysis in the chloroplast, mitochondrion or cytosol, the Calvin-Benson cycle in the chloroplast or the pentose phosphate pathway (Chen and Thelen, 2010; Zaffagnini et al., 2014). As mentioned above, DHAP is also a main precursor for IPP biosynthesis *via* the non-mevalonate pathway in the chloroplast, namely the precursor for isoprenoids, which are severely down-regulated under HL. Interestingly, one of the major conclusions of a recently reported study on circadian transcriptional regulation in *N. oceanica* was that a coordination between the glycolysis and fatty acid synthesis occurred for lipid production (Poliner et al., 2015). Out of three triose phosphate isomerase sequences, one is down-regulated in ML (*TPI*, Naga\_100011g1), while the two others (Naga\_103287g1, Naga\_100001g84) are up-regulated in ML and HL compared to LL. The down-regulated enzyme, involved in the Calvin-Benson cycle, is localized in the chloroplast, while the up-regulated ones are cytosolic and participate in glycolysis. An altered coupling of photosynthesis and glycolysis therefore seems to be supported by these data, with a cytosolic energetic metabolism relying on hexose phosphate consumption.

To complement the expression data, the light response of primary carbon metabolites, trioses, pentoses and hexoses, were quantified by liquid chromatography coupled with high-resolution mass spectrometry (LC-HRMS; Figure 8, Supplemental Data Set 4).



**Figure 8. LC-ESI(+)-MS analyses of sugars in *Nannochloropsis gaditana* cells grown in different light conditions.** Data are expressed as relative content with respect to the level of the internal standard (formononetin). Data are the average of at least 4 extractions. For more details, see methods.

For most sugars, there is a generalized increase in accumulation with light intensity, and 13 out of 23 measured carbohydrates were significantly more abundant under ML compared to LL (Figure 8; Supplemental Data Set 4). In particular, maltotriose is the fifth most abundant sugar under ML (Figure 8A) and responds progressively to light (LL<ML<HL; Figure 8A). Similarly, sorbitol/mannitol and fructose-dianhydride are accumulated in response to increased light. The sugar accumulation signature in HL could negatively regulate photosynthesis and Calvin-Benson cycle gene expression (Couée et al., 2006) (Supplemental Data Set 3A), as reported in several plant and cyanobacteria species when exposed to high light conditions (Pego et al., 2000; Rolland et al., 2006).

In other photosynthetic organisms, sugar accumulation has also been correlated with the downregulation of peroxisomal lipid  $\beta$ -oxidation (Contento et al., 2004), carotenoid biosynthesis (Ryu et al., 2004) and ROS scavenging processes (Elbein et al., 2003; Lunn et al., 2014; Stoop et al., 1996), all phenomena also observed in this work, suggesting a possible role for sugars not only as carbon sinks but also as signaling molecules.

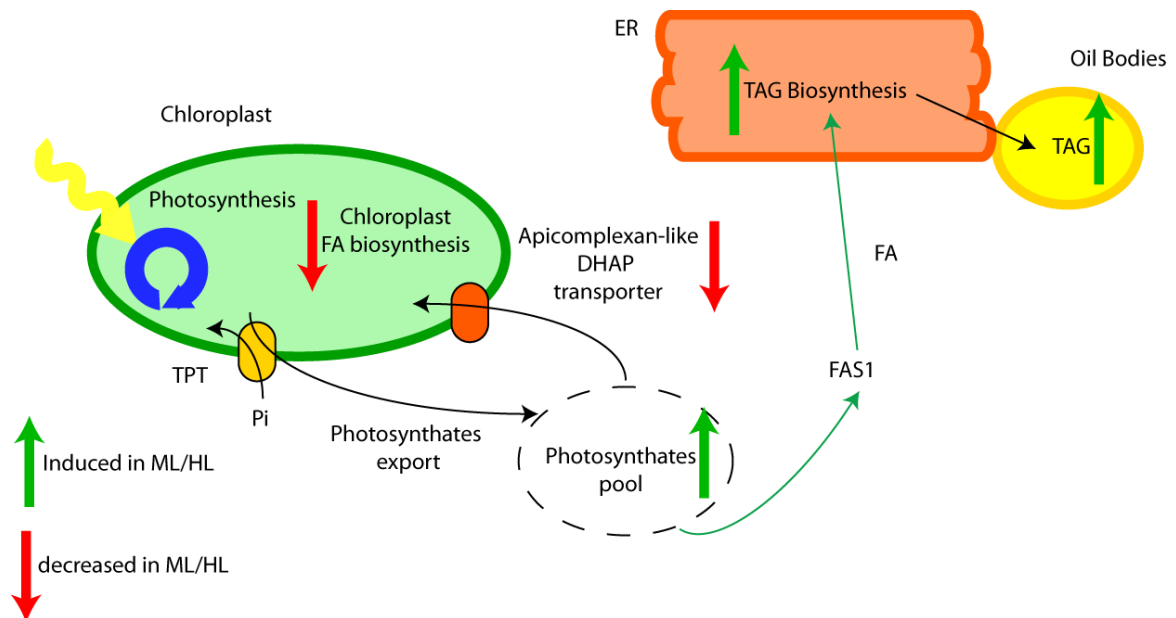
In this scenario, it is interesting to observe the presence of one carbohydrate intermediate more abundant in LL than ML, i.e., DHAP (LL vs ML = 1.97). Its accumulation is fully light-dependent in the framework of this experiment, as it is also more abundant in ML than in HL (ML/HL = 2.11) (Figure 8B). In plants, DHAP synthesis in the chloroplast is generally activated by light and acts as a precursor of starch biosynthesis. The reduced DHAP accumulation in *N. gaditana* is instead consistent with the depression of isoprenoid synthesis genes and FASII in the chloroplast.

It is intriguing to note that DHAP levels diminished in response to light in *N. gaditana*, indicating that in this species, DHAP could represent a checkpoint for metabolic reprogramming towards lipid biosynthesis instead of chrysolaminarin or laminarin biosynthesis.

Considering all data presented here, the subcellular compartmentalization of the enzymes producing and consuming DHAP and the trafficking of this triose phosphate from one compartment to another are therefore critical to comprehending carbon metabolic fluxes and carbon partitioning. In primary chloroplasts, DHAP is known to be exported to the cytosol by the triose phosphate/phosphate translocator (TPT) in exchange for inorganic phosphate (Pi) needed to regenerate ATP in the stroma. In the secondary plastid of Apicomplexa (organisms presenting a plastid limited by 4 membranes, as in Eustigmatophyceae, but lacking photosynthesis), TPT translocators act in the opposite way, importing DHAP to feed the non-mevalonate pathway of isoprenoid synthesis and FA synthesis (Mullin et al., 2006; Brooks et al., 2010; Botté et al., 2012, 2013). We identified 4 genes coding for TPTs in *N. gaditana* (Naga\_100100g11, Naga\_101308g1, Naga\_100006g37, Naga\_100007g104), and at least one of these genes has high homology with the Apicomplexa TPT (Naga\_100100g11), showing conserved features such as 10 membrane spanning domains and a chloroplast transit peptide, suggesting a likely similar role in importing DHAP into the plastid. The presence of an Apicomplexa-like TPT suggests that *N. gaditana* might import DHAP from the cytosol, possibly to feed IPP and FA syntheses. The four TPT genes of *N. gaditana* are down-regulated in HL, which could lead to decreased import of DHAP in the chloroplast under these conditions in favor of FA biosynthesis in the cytosol and in the ER.

This hypothesis is also consistent with the downregulation of isoprenoids and FA within the plastid.

In an organism that grows phototrophically, such as *N. gaditana*, all reduced carbon molecules are first synthesized in the chloroplast and can be directly utilized there or exported to the cytosol. This study shows that the adaptation to high light involves a strong reorganization of the roles of the different organelles and supports a scenario where reduced carbon molecules can be transported back to the chloroplast to fuel plastid biosynthetic pathways such as isoprenoids and FA biosynthesis, initially fed on plastid precursors. Following this hypothesis, a regulatory loop for carbon partitioning between the chloroplast and cytosol could be tuned by specific transporters. The modulation of metabolite fluxes in and out of the chloroplast could thus represent a major mechanism in the metabolic remodeling of the whole cell, increasing cytosolic pathways when needed and, as a consequence, ER-localized TAG biosynthesis (Figure 9).



**Figure 9. Model of regulation of carbon partitioning and lipid metabolisms in *Nannochloropsis gaditana* cells treated with high light.** A possible mechanism for the modulation of lipid biosynthesis involving the DHAP transporter is proposed. Pathways and metabolites that are down/up-regulated in ML/HL are shown in red/green, respectively. In ML/HL, chloroplast FA biosynthesis is inhibited. Photosynthates exported to the cytoplasm might be converted into FA by a putative cytosolic FAS1. Based on the presented data, DHAP transporters could be part of a regulatory loop controlling the carbon fluxes between the chloroplast and cytosol, determining a novel cellular metabolic status adapted to light intensity.

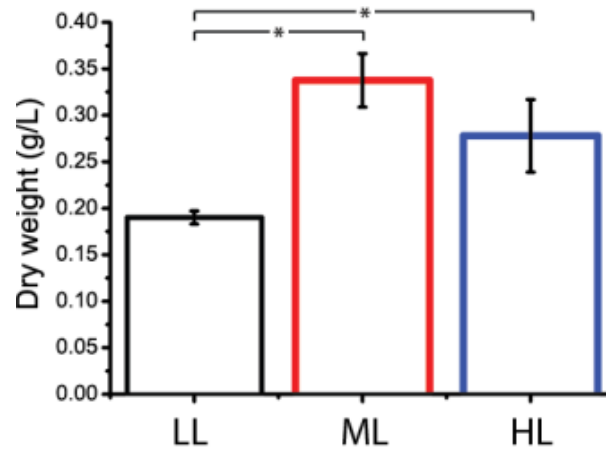
### **Acknowledgment**

A.A., G.P., A.M. and T.M. were supported by ERC starting grant BIOLEAP nr 309485 to T.M.. M.B., J.J. and E.M. were supported by grants from Agence nationale de la Recherche (ANR DiaDomOil, ChloroMitoLipid and Reglisse) and Infrastructure d’Avenir Océanomics. We thank Gino Schiano di Visconte for helping in the setup of the conditions of LC-MS for metabolomics analyses.

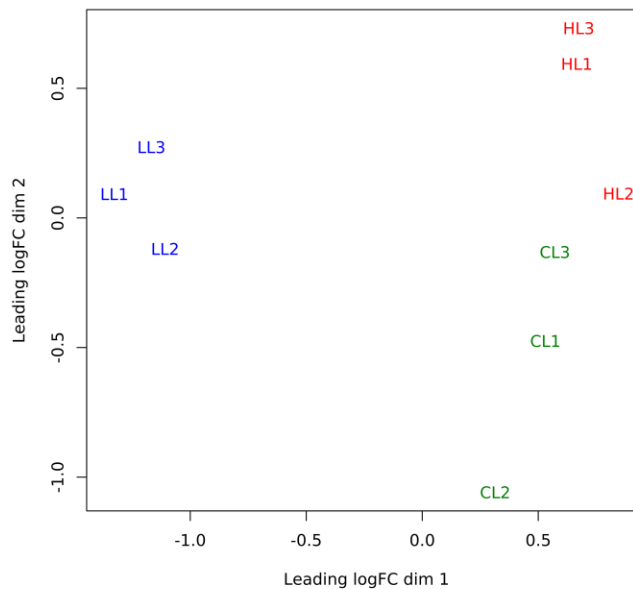
### **Author Contributions**

T.M., G.G. and E.M. designed the research; A.A and T.M. coordinated all experiments; G.P. and A.M. prepared biological material and performed physiological analysis; M.B., J.J. and E.M. performed lipidomic analysis and analyzed corresponding data; G.D. and G.G. performed metabolomic analysis and analyzed corresponding data; N.V. and G.V. performed transcriptomic analysis and analyzed corresponding data. G.P. and A.A. analyzed and compared all data. A.A, G.P., T.M., E.M., G.G. wrote the article.

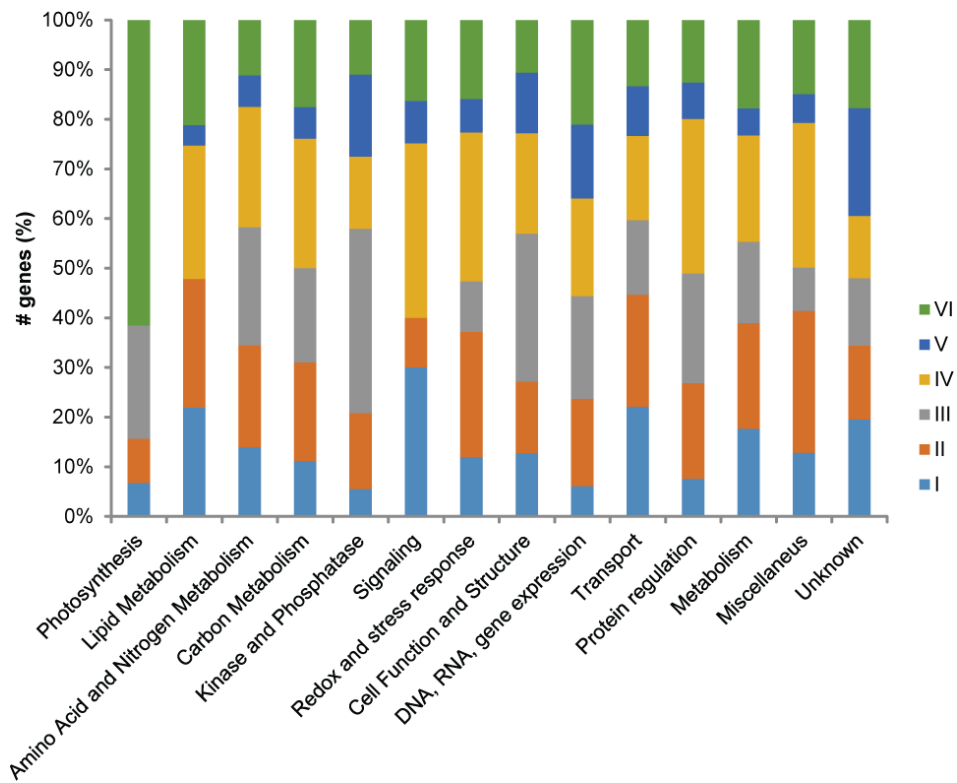
## Supplemental Data



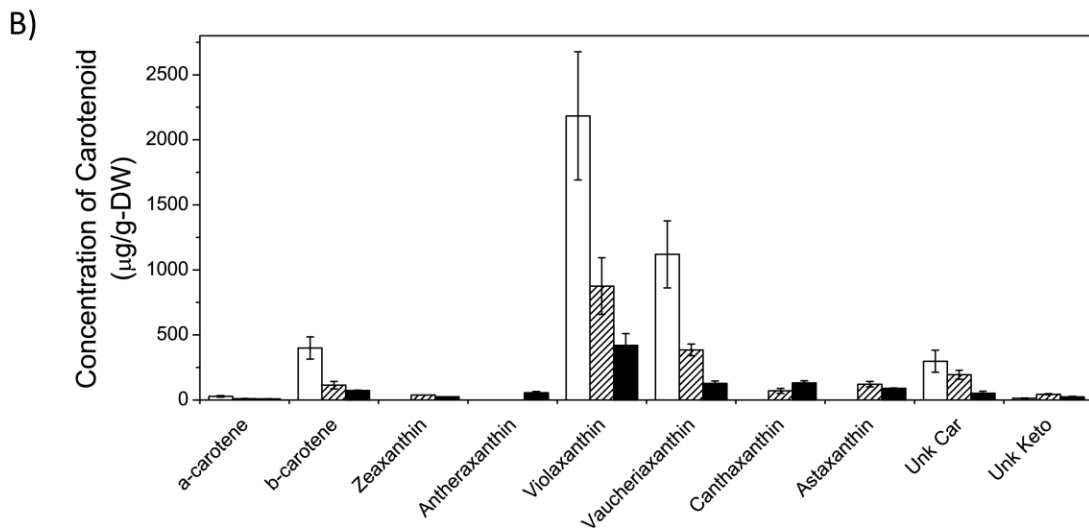
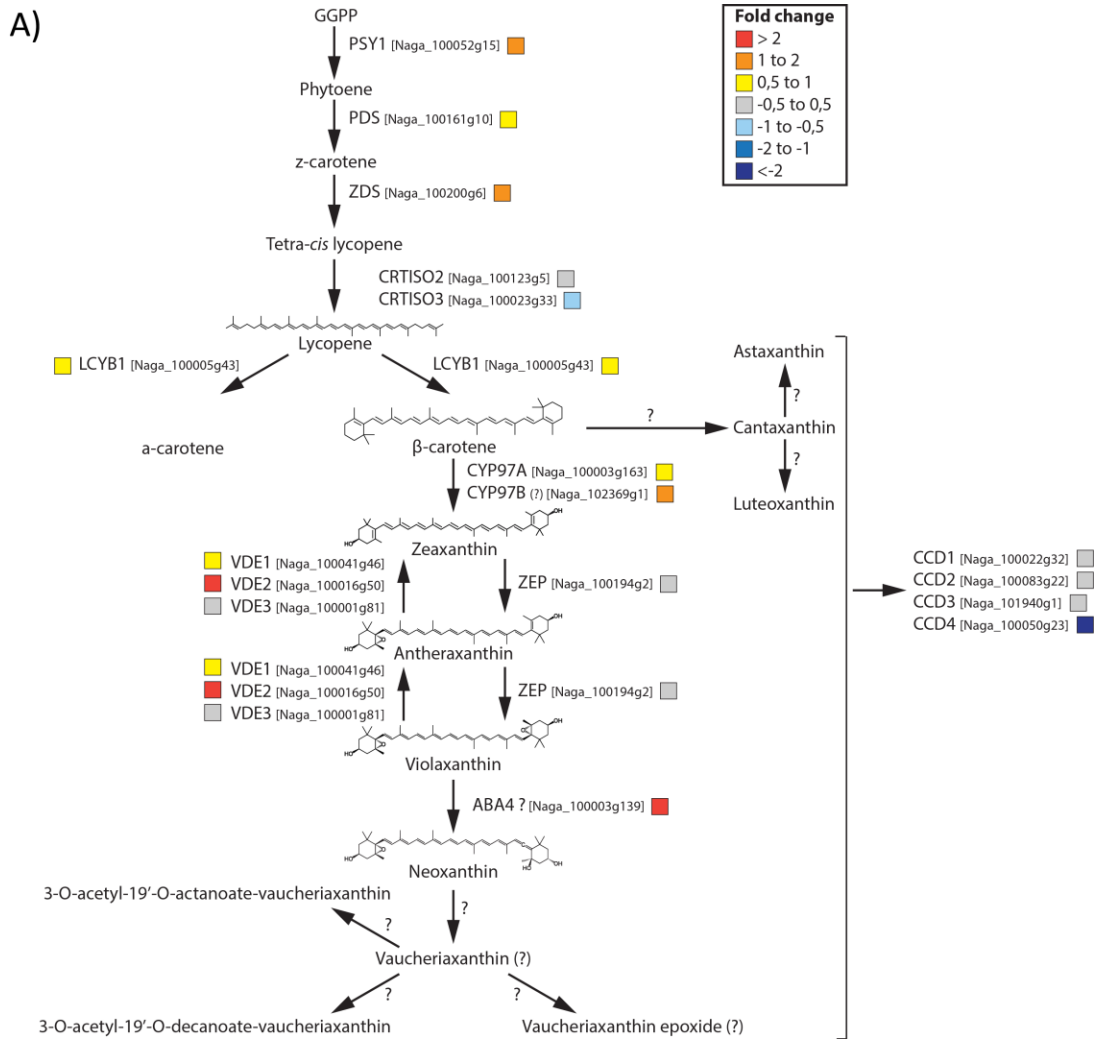
**Supplemental Figure 1. Light effect on the growth of *Nannochloropsis gaditana* cells.** Dry weight of the cultures after 5 days of differential light treatment. LL (Low Light), ML (Medium Light) and HL (High Light). Values represent means of three biological replicates  $\pm$  SD (n=3). \*P < 0.05, one-way analysis of variance (ANOVA).



**Supplemental Figure 2. Two dimensional (2D) MDS plot for the light transcriptome dataset.** LL (Low Light), CL (Control Light = Medium Light) and HL (High Light).



**Supplemental Figure 3. Distribution of genes in manual functional categories for the six groups identified by hierarchical clustering analysis.** For each category the number of genes was normalized to 100%.



**Supplemental Figure 4. Regulation of carotenoid biosynthetic pathway.** A) Hypothesized carotenoid biosynthetic pathway, colored squares indicate the regulation pattern of genes encoding putative enzymes in LL vs ML. The color scale for induction/repression ratios is reported in the inset. B) Carotenoid concentration. Data are the average of at least 4 extractions.







**Supplemental Table 1. Non-polar metabolites in *Nannochloropsis gaditana* cells grown in three different light conditions.** Data are expressed as  $\mu\text{g}$  of metabolite per g of dry weight and they are the average of at least 4 extractions.

Metabolite	RT	$\lambda_{\text{max}}$	LL	CL	HL
			average	average	average
Chlorophyllide a	2,58	422, 664	471,9 $\pm$ 53,6	555,1 $\pm$ 136,6	197,7 $\pm$ 18,7
Chlorophyllide a'	2,91	420, 664	79,3 $\pm$ 1,2	118,9 $\pm$ 6,8	74,3 $\pm$ 15,2
Chl cat.	3,02	422, 664	325,5 $\pm$ 0,5	285,5 $\pm$ 34,7	271,8 $\pm$ 73,0
DivinylChlA	3,13	421, 666	271,6 $\pm$ 4,2	194,3 $\pm$ 63,4	-
Chl cat.	3,21	422, 666	14,4 $\pm$ 5,6	537,8 $\pm$ 182,8	-
Chl cat.	3,38	418, 663	67,2 $\pm$ 6,6	68,8 $\pm$ 16,7	-
Chl cat.	3,55	419, 665	63,3 $\pm$ 19,9	438,6 $\pm$ 235,2	76,9 $\pm$ 38,3
Chl cat.	4,74	431, 667	231,5 $\pm$ 36,2	60,6 $\pm$ 30,2	296,5 $\pm$ 33,9
Chl cat.	5,19	432, 665	44,1 $\pm$ 10,3	-	-
Unk. Car	5,98	416, 437, 467	-	4,0 $\pm$ 0,5	5,4 $\pm$ 7,3
Unk. Car	6,54	416, 438, 468	53,8 $\pm$ 16,4	8,6 $\pm$ 0,7	-
Unk. Car	7,12	416, 439, 468	110,0 $\pm$ 21,4	30,4 $\pm$ 1,0	8,5 $\pm$ 1,9
Chl cat.	8	432, 663	-	7,4 $\pm$ 0,0	16,7 $\pm$ 6,2
All-trans-Neoxanthin	10,85	412, 434, 462	151,4 $\pm$ 13,0	56,5 $\pm$ 8,6	24,7 $\pm$ 3,2
All-trans-Violaxanthin	11,43	416, 439, 467	2184,1 $\pm$ 493,0	870,9 $\pm$ 217,0	419,6 $\pm$ 90,6
9-cis-Violaxanthin	12	416, 439, 468	-	4,8 $\pm$ 1,0	-
C8-vaucheriaxanthin	12,44	443, 471	699,2 $\pm$ 151,4	314,5 $\pm$ 35,6	104,3 $\pm$ 11,7
C10-vaucheriaxanthin	14,13	443, 471	420,7 $\pm$ 106,8	70,8 $\pm$ 9,0	24,2 $\pm$ 5,5
Antheraxanthin	14,93	443, 471	-	-	55,3 $\pm$ 10,2
Unk. Car	14,93	441, 472	98,9 $\pm$ 33,8	85,2 $\pm$ 24,1	-
MetoxyChlA	16,4	422, 662	993,0 $\pm$ 237,2	-	-
Astaxanthin	16,57	473	-	121,6 $\pm$ 21,0	84,7 $\pm$ 3,5
Chl cat.	18,12	429	31,5 $\pm$ 5,2	-	-
Unk. Keto	18,12	460	-	35,9 $\pm$ 1,9	19,0 $\pm$ 4,2
Unk. Car	18,68	436, 464	-	3,3 $\pm$ 1,3	-
Zeaxanthin	18,7	447, 470	-	38,7 $\pm$ 0,6	24,7 $\pm$ 0,6
ChlA	19,49	428, 664	12600 $\pm$ 2253	3004,3 $\pm$ 333,2	1009,0 $\pm$ 193,4
ChlA'	20,35	430, 665	2461,5 $\pm$ 359,2	529,9 $\pm$ 141,0	-
Canthaxanthin	20,6	474	-	70,2 $\pm$ 19,4	130,7 $\pm$ 17,1
Unk. Car	22,56	416, 438, 470	16,1 $\pm$ 4,0	13,0 $\pm$ 2,8	9,4 $\pm$ 1,7
Unk. Car	22,62	442, 473	0,4 $\pm$ 0,1	-	1,6 $\pm$ 0,1
Unk. Keto	24,4	459	13,4 $\pm$ 1,5	8,2 $\pm$ 2,0	4,5 $\pm$ 0,2
$\alpha$ -carotene	25,47	448, 472	27,9 $\pm$ 6,3	10,4 $\pm$ 1,1	7,8 $\pm$ 1,9
Astaxanthin ester	26,6	473	-	-	4,2 $\pm$ 0,7
$\beta$ -carotene	27,87	452, 478	337,6 $\pm$ 75,1	102,4 $\pm$ 25,4	60,3 $\pm$ 1,7
pheophytin a	28,38	410, 471	-	284,0 $\pm$ 5,6	217,0 $\pm$ 55,5
$\beta$ -carotene isomer	28,88	447, 473	22,8 $\pm$ 4,1	12,2 $\pm$ 2,8	4,2 $\pm$ 0,6
cis- $\beta$ -carotene	29,2	446, 474	39,7 $\pm$ 6,5	-	7,8 $\pm$ 1,2
Unk. Car	30,59	424, 446, 475	18,8 $\pm$ 2,6	5,6 $\pm$ 0,2	3,6 $\pm$ 0,6

**Supplemental Table 2.** Description of the SOLID whole transcriptome RNA-seq libraries.

Lib	Tag	Tot	Filtered	# Good	Aligned	%Aligned	# alignment
LL1	F3	42696548	9497539	33199009	17227348	51,89	18183687
LL1	F5	43360926	5360200	38000726	14767874	38,86	16766312
LL2	F3	29622042	5483332	24138710	14108857	58,45	14945420
LL2	F5	30968049	7122785	23845264	11236780	47,12	13094741
LL3	F3	40427259	8357765	32069494	16338164	50,95	17484940
LL3	F5	41355172	11915288	29439884	13854237	47,06	15948982
ML1	F3	44588099	13283483	31304616	18771256	59,96	19480959
ML1	F5	46553466	8858790	37694676	16086441	42,68	17375835
ML2	F3	45246570	14270841	30975729	16069468	51,88	16707982
ML2	F5	46096388	8142102	37954286	13362380	35,21	14552425
ML3	F3	49502792	8121220	41381572	17303781	41,82	17765379
ML3	F5	50098960	8619033	41479927	16353233	39,42	17619760
HL1	F3	55554459	8246561	47307898	20567044	43,47	21184336
HL1	F5	56413315	3321880	53091435	20517078	38,64	22068900
HL2	F3	52503656	11501665	41001991	20731896	50,56	21545882
HL2	F5	53413338	11321261	42092077	18790274	44,64	20307905
HL3	F3	48010702	10740740	37269962	18008835	48,32	18661418
HL3	F5	48712873	27066872	21646001	14203065	65,62	15331213

**Supplemental Data Set 1. Light dose regulation of gene expression in *Nannochloropsis gaditana*.** A) The mRNA-Seq data of genes that are more expressed both in HL as compared to CL and in CL as compared to LL. For each Gene ID, the predicted Gene Function is reported together with the fold changes in logarithmic scale (logFC) of the number of reads in the three light conditions. Only genes significantly differentially expressed were considered and data are the average of 3 replicates.

Gene ID	Gene Function	logFC LL vs CL	logFC CL vs HL	logFC LL vs HL
Naga_100107g9	hypothetical protein	-1,28	-2,19	-3,46
Naga_100040g9	Acyl transferase/acyl hydrolase/lysophospholipase	-2,62	-0,80	-3,42
Naga_100889g1	putative lipase	-2,07	-0,97	-3,04
Naga_101253g1	Plastid lipid-associated protein	-2,10	-0,93	-3,03
Naga_100028g29	Kua-ubiquitin conjugating enzyme hybrid, localisation	-1,04	-1,99	-3,03
Naga_100012g74	hypothetical protein	-1,66	-1,35	-3,01
Naga_100009g65	hypothetical protein	-2,20	-0,78	-2,98
Naga_100260g8	serine threonine-protein kinase nek 1	-1,49	-1,38	-2,87
Naga_100193g8	hypothetical protein	-0,71	-2,09	-2,80
Naga_100038g37	Protein of unknown function DUF1772	-1,34	-1,34	-2,68
Naga_100355g2	hypothetical protein	-1,72	-0,94	-2,66
Naga_100032g24	hypothetical protein	-1,13	-1,44	-2,57
Naga_100686g2	multicopper oxidase	-0,85	-1,67	-2,51
Naga_100126g11	myb-like dna-binding shaqkyf class family protein	-1,36	-1,09	-2,45
Naga_100091g4	alanine-glyoxylate aminotransferase 2-like 1	-1,43	-0,99	-2,42
Naga_100568g2	hypothetical protein	-1,54	-0,87	-2,41
Naga_100483g1	hypothetical protein	-1,47	-0,92	-2,38
Naga_100149g1	multicopper oxidase type 2	-1,06	-1,29	-2,35
Naga_100208g9	glycosyl transferase	-1,34	-0,99	-2,33
Naga_100356g4	hypothetical protein	-1,39	-0,93	-2,32
Naga_100321g1	cell division cycle 20	-0,94	-1,29	-2,23
Naga_100276g9	sugar phosphatase of the had superfamily	-0,90	-1,29	-2,19
Naga_100005g113	hypothetical protein	-1,30	-0,83	-2,12
Naga_100773g1	TRX domain-containing protein	-1,13	-0,97	-2,10

Naga_100751g1	hypothetical protein	-0,96	-1,14	-2,10
Naga_100253g2	mitochondrial carrier protein	-1,31	-0,78	-2,09
Naga_100024g37	succinate dehydrogenase flavoprotein subunit	-1,32	-0,76	-2,08
Naga_101539g1	glutathione s-transferase	-1,01	-1,02	-2,03
Naga_100131g6	Isochorismatase	-1,01	-1,02	-2,03
Naga_100187g2	hypothetical protein	-0,98	-1,02	-2,00
Naga_100051g32	gcn5-related n-acetyltransferase	-1,02	-0,96	-1,98
Naga_100568g3	alternative oxidase	-0,87	-1,06	-1,94
Naga_100317g3	hypothetical protein	-0,89	-1,04	-1,93
Naga_100190g5	pentatricopeptide repeat-containing protein	-1,16	-0,77	-1,93
Naga_100009g110	hypothetical protein	-0,91	-1,02	-1,93
Naga_100277g1	inositol oxygenase	-0,91	-1,02	-1,93
Naga_100202g5	nad-dependent deacetylase sirtuin-5	-0,89	-1,03	-1,93
Naga_100578g2	peptidase s8 and s53 subtilisin kexin sedolisin	-1,02	-0,90	-1,92
Naga_100485g4	r-snare vamp72-family	-0,90	-1,01	-1,92
Naga_100106g14	hypothetical protein	-0,99	-0,91	-1,90
Naga_100030g4	derlin-like protein	-1,10	-0,78	-1,88
Naga_100010g78	hsp90-like protein	-0,81	-1,06	-1,88
Naga_100022g33	conserved unknown protein	-0,86	-1,00	-1,86
Naga_100379g3	rhomboid domain-containing protein 1	-1,02	-0,82	-1,83
Naga_100159g16	fructose- -bisphosphatase	-0,94	-0,89	-1,83
Naga_100020g13	platelet-activating factor acetylhydrolase ib subunit alpha	-0,86	-0,97	-1,83
Naga_100104g3	hydroxymethylpyrimidine kinase phosphomethylpyrimidine kinase thiamine-phosphate diphosphorylase	-1,05	-0,77	-1,82
Naga_100008g101	family protein	-1,08	-0,73	-1,82
Naga_100005g129	aminopeptidase-like 1	-1,04	-0,77	-1,81
Naga_100150g10	phosphoacetylglucosamine mutase	-0,98	-0,81	-1,79
Naga_100019g30	hypothetical protein	-0,99	-0,80	-1,79

Naga_100889g4	Neuraminidase	-1,03	-0,75	-1,78
Naga_100002g102	cystathionine gamma-lyase	-0,74	-1,04	-1,78
Naga_100704g3	hypothetical protein	-0,95	-0,82	-1,77
Naga_100047g41	hypothetical protein	-0,87	-0,90	-1,77
Naga_100094g5	cof-like hydrolase	-1,00	-0,76	-1,76
Naga_100025g57	ribosomal protein l23	-0,85	-0,89	-1,75
Naga_101483g1	delta-4 fatty acid desaturase	-0,78	-0,95	-1,72
Naga_100254g4	Ankyrin repeat-containing protein	-0,84	-0,86	-1,69
Naga_100313g8	tyrosyl-dna phosphodiesterase	-0,92	-0,77	-1,69
Naga_100066g26	hypothetical protein	-0,90	-0,75	-1,65
Naga_100096g23	probable dolichyl pyrophosphate glc1man9 c2 alpha- -glucosyltransferase	-0,89	-0,76	-1,64
Naga_100214g7	putative oxidoreductase yrbe	-0,80	-0,83	-1,63
Naga_100004g98	solute carrier family 36 (proton amino acid symporter) member 1	-0,82	-0,78	-1,60
Naga_100063g18	arsenite inducible rna associated protein aip	-0,80	-0,74	-1,54
Naga_100077g12	hypothetical protein	-0,77	-0,75	-1,53
Naga_100565g2	hypothetical protein	-0,73	-0,79	-1,52
Naga_100001g31	alpha beta hydrolase	-0,78	-0,72	-1,50
Naga_100498g3	purple acid phosphatase isoform b2	-0,72	-0,75	-1,47

**Supplemental Data Set 1. Light dose regulation of gene expression in *Nannochloropsis gaditana*.** B) The mRNA-Seq data of genes that are more expressed both in LL as compared to CL and in CL as compared to HL. For each Gene ID, the predicted Gene Function is reported together with the fold changes in logarithmic scale (logFC) of the number of reads in the three light conditions. Only genes significantly differentially expressed were considered and data are the average of 3 replicates.

Gene ID	Gene Function	logFC LL vs CL	logFC CL vs HL	logFC LL vs HL
Naga_101992g1	hypothetical protein	3,44	2,06	5,51
Naga_100012g10	hypothetical protein	2,72	0,94	3,66
Naga_102720g1	hypothetical protein	2,08	1,52	3,60
Naga_101500g1	hypothetical protein	1,91	1,64	3,55
Naga_101500g2	hypothetical protein	2,02	1,28	3,30
Naga_100604g2	hypothetical protein	1,31	1,69	3,00
Naga_100668g5	hypothetical protein	2,08	0,81	2,89
Naga_101262g1	hypothetical protein	1,64	1,24	2,88
Naga_100145g6	hypothetical protein	1,31	1,25	2,57
Naga_100557g5	hypothetical protein	1,68	0,85	2,54
Naga_100497g4	hypothetical protein	1,28	1,02	2,30
Naga_100540g4	mitochondrial carrier domain-containing protein	1,46	0,83	2,29
Naga_100540g5	hypothetical protein	1,10	1,08	2,19
Naga_100038g16	hypothetical protein	1,29	0,80	2,09
Naga_100459g5	hypothetical protein	0,99	1,04	2,02
Naga_103833g1	hypothetical protein	1,08	0,78	1,86
Naga_100001g79	hypothetical protein	0,97	0,84	1,81
Naga_100453g4	hypothetical protein	0,83	0,83	1,66



**Supplemental Data Set 2. Light response of various group of genes related to plastid physiology and metabolism in *Nannochloropsis gaditana*.** A) The mRNA-Seq data of genes that belong to the light harvesting complex superfamily are reported. For each Gene ID, the predicted Gene Function and the Short name are reported. logFC indicates the fold changes in logarithmic scale of the number of reads in the three light conditions. OVER stands for significantly over-expressed genes while SUB stands for significantly down-regulated genes. Data are the average of 3 replicates.

Gene ID	Gene Function	Short name (I)	Short name (II)	LL vs CL	CL vs HL	LL vs HL	logFC LL vs CL	logFC CL vs HL	logFC LL vs HL
Naga_100012g50	Light harvesting - F-type	NgLHCf1	Naga 3	OVER		OVER	2,65	-0,23	2,42
Naga_100005g99	Light harvesting - F-like-type	NgLHCf2	Naga 13	OVER		OVER	2,12	-0,21	1,92
Naga_100157g5	Light harvesting - F-like-type	NgLHCf3	Naga 18	OVER		OVER	2,43	-0,12	2,32
Naga_100168g14	Light harvesting - F-like-type	NgLHCf4	Naga 12	OVER		OVER	2,54	-0,12	2,42
Naga_100017g59	Light harvesting - F-like-type	NgLHCf5	Naga 9	OVER		OVER	2,58	0,08	2,67
Naga_100004g86	Light harvesting - F-like-type	NgLHCf6	Naga 4	OVER		OVER	3,89	-0,14	3,75
Naga_100013g28	Light harvesting - F-like-type	NgLHCf7	Naga 17	OVER		OVER	2,65	-0,43	2,22
Naga_100027g19	Light harvesting - F-like-type	NgLHCf8	Naga 2	OVER	SUB	OVER	2,68	-0,73	1,95
Naga_100002g18	Light harvesting - R-like-type	NgLHCr1	Naga 16	OVER		OVER	2,33	0,07	2,39
Naga_100168g13	Light harvesting - R-like-type	NgLHCr2	Naga 19	OVER		OVER	3,38	0,07	3,45
Naga_100018g45	Light harvesting - R-like-type	NgLHCr3	Naga 14	OVER		OVER	1,60	-0,06	1,54
Naga_100092g17	Light harvesting - R-type	NgLHCr4	Naga 8	OVER		OVER	2,82	0,50	3,32
Naga_100434g4	Light harvesting - R-type	NgLHCr5	Naga 7	OVER		OVER	4,73	-0,16	4,54
Naga_100641g3	Light harvesting - R-like-type	NgLHCr6	Naga 10	OVER		OVER	4,53	-0,38	4,15
Naga_100017g83	Light harvesting - R-type	NgLHCr7	Naga 20	OVER		OVER	3,48	-0,14	3,34
Naga_100173g12	Light harvesting - X-Type	NgLHCx1	Naga 1	OVER		OVER	1,66	-0,20	1,46
Naga_100056g42	Light harvesting - X-Type	NgLHCx2	Naga 5				0,77	0,01	0,78
Naga_101036g3	Light harvesting - Lhc-like	NgLIL1	Naga 6	SUB	OVER	SUB	-1,44	0,91	-0,53
Naga_100056g15	Light harvesting - Lhc-like	NgLIL2	Naga 21	OVER		OVER	2,77	0,28	3,04
Naga_100030g5	Light harvesting - ELIP	NgLIL3	Naga 22	OVER		OVER	3,46	-0,20	3,26
Naga_101227g1	Light harvesting - LHC-like	NgLIL4	Naga 15	OVER	SUB	OVER	3,55	-0,84	2,71

**Supplemental Data Set 2. Light response of various group of genes related to plastid physiology and metabolism in *Nannochloropsis gaditana*.** B) The mRNA-Seq data of genes that belong to the carotenoid biosynthetic pathway are reported. For each Gene ID, the predicted Gene Function and the Short name are reported. logFC indicates the fold changes in logarithmic scale of the number of reads in the three light conditions. OVER stands for significantly over-expressed genes while SUB stands for significantly down-regulated genes. Data are the average of 3 replicates.

Gene ID	Gene Function	Short name	LL vs CL	CL vs HL	LL vs HL	Log FC LLvsCL	Log FC CLvsHL	Log FC LLvsHL
Naga_100052g15	phytoene syntase 1	PSY1	OVER		OVER	1,22	-0,17	1,04
Naga_100427g4	phytoene dehydrogenase	PDH (Crtl)				0,10	0,39	0,50
Naga_100161g10	phytoene desaturase	PDS				0,68	0,06	0,74
Naga_100200g6	z-carotene desaturase	ZDS	OVER			1,12	-0,24	0,88
Naga_100123g5	carotenoid isomerase 2	CRTISO2				0,04	0,12	0,16
Naga_100023g33	carotenoid isomerase 3	CRTISO3	SUB		SUB	-0,50	-0,33	-0,83
Naga_100066g11	plastid terminal oxydase	PTOX	SUB		SUB	-0,50	0,09	-0,40
Naga_100005g43	lycopene b cyclase 1	LCYB1				0,65	-0,09	0,56
Naga_100003g163	carotene hydroxylase	CYP97A				0,68	-0,08	0,62
Naga_102369g1	carotene hydroxylase	CYP97B	OVER		OVER	1,65	0,40	2,05
Naga_100194g2	zeaxanthin epoxydase	ZEP				0,36	0,10	0,46
Naga_100041g46	violaxanthin deoxygenase	VDE				0,65	-0,01	0,64
Naga_100016g50	violaxanthin deoxygenase	VDE	OVER		OVER	2,17	-0,15	2,02
Naga_100001g81	violaxanthin deoxygenase	VDE			SUB	-0,10	-0,16	-0,34
Naga_100022g32	carotenoid cleavage dioxygenase	CCD		SUB	SUB	0,11	-0,72	-0,60
Naga_100083g22	carotenoid cleavage dioxygenase	CCD			SUB	-0,40	0,17	-0,23
Naga_101940g1	carotenoid cleavage dioxygenase	CCD				0,24	0,38	0,13
Naga_100050g23	carotenoid cleavage dioxygenase	CCD	SUB		SUB	-2,03	0,44	-1,59
Naga_100003g139	NXY (?)	ABA4	OVER		OVER	1,77	0,32	2,09
Naga_100403g3	ABA 8'-hydroxylase	CYP707A1			SUB	0,14	-0,34	-0,20
Naga_100006g75	Cytochrome p450				SUB	0,00	-0,29	-0,30

**Supplemental Data Set 2. Light response of various group of genes related to plastid physiology and metabolism in *Nannochloropsis gaditana*.** C) The mRNA-Seq data of genes that belong to the chlorophyll biosynthetic pathway are reported. For each Gene ID, the predicted Gene Function and the Short name are reported. logFC indicates the fold changes in logarithmic scale of the number of reads in the three light conditions. OVER stands for significantly over-expressed genes while SUB stands for significantly down-regulated genes. Data are the average of 3 replicates.

Gene ID	Gene Function	Short name	LL vs CL	CL vs HL	LL vs HL	Log FC LLvsCL	Log FC CLvsHL	Log FC LLvsHL
Naga_100001g200	Glutamyltransferasi RNA	GiRNAS-1				0.52	0.01	0.53
Naga_100018g48	Glutamyltransferasi RNA	GiRNAR-2				0.16	-0.09	0.07
Naga_100125g4	Glutamyl-tRNA reductase	GluTR	OVER			1.11	-0.25	0.86
Naga_100010g65	5-aminolevulinat dehidratase	ALAD	OVER			1.00	-0.13	0.87
Naga_100003g83	porfobilinogen deaminase	PBGD	OVER		OVER	1.05	0.39	1.44
Naga_100040g15	Uroporfirinogen III sintase	UROS				0.52	0.12	0.65
Naga_100017g11	Uroporfirinogen dekarboksilase	UROD-1				0.60	-0.22	0.38
Naga_100096g24	Uroporfirinogen dekarboksilase	UROD-2				0.53	-0.01	0.52
Naga_100665g2	Coproporfirinogen III oksidase aerobik//Coproporfirinogen-III oksidase, kloroplastik	CPOX-1	OVER		OVER	1.31	0.06	1.37
Naga_100344g2	coproporfirinogen iii oksidase	CPOX-2				0.24	0.15	0.38
Naga_100646g3	Protoporfirinogen IX oksidase	PPO				0.46	0.05	0.51
Naga_1Chloroplast23	Magnesium chelatase ATPase subunit I//Magnesium-chelatase subunit chlI, kloroplastik//Mg chelatase	Mg-chelatase sub. -1						
Naga_100105g7	Magnesium chelatase H subunit//Magnesium-chelatase subunit H//Mg chelatase	Mg-chelatase sub. H				0.22	0.12	0.34
Naga_100008g68	Magnesium-protoporfirin O-metiltransferase//S-adenosil-L-metionine Mg-protoporfirin IX metiltransferase	Mg-PPMT	SUB		SUB	-0.54	0.19	-0.35
Naga_1Chloroplast29	Magnesium-protoporfirin IX monometil ester siklase	MCT						
Naga_100548g3	3,8-divinil protofloklorofilid 8-vinil reduktase	DVR	OVER			1.06	-0.18	0.88
Naga_100258g4	Protofloklorofilid reduktase	PORA-1			OVER	0.86	0.39	1.25
Naga_100014g48	Protofloklorofilid reduktase	PORA-2				0.39	-0.10	0.29
Naga_100207g6	Geranilgeranil difosfat reduktase	CHL P-1	OVER	SUB		1.26	-0.63	0.62
Naga_101224g2	Klorofil sintase	CHLG			SUB	0.06	-0.54	-0.48
Naga_100989g1	Klorofil sintase	CHLG			SUB	-0.07	-0.46	-0.54
Naga_100034g24	klorofilid a oksigenase	CAO				0.25	0.07	0.32
Naga_100072g25	Pheoforbida a oksigenase (Fragment)//Pheoforbida a oksigenase, kloroplastik//letal daun spot 1-like protein	PaO	SUB		SUB	-0.52	-0.17	-0.69
Naga_100016g32	Pheoforbida a oksigenase	PaO				0.36	0.16	0.51
Naga_100008g43	Pheofitinase	PPH				0.90	-0.33	0.57
Naga_100010g10	Fe-chelatase	Fe-chelatase 1			SUB	-0.05	-0.28	-0.33

**Supplemental Data Set 2. Light response of various group of genes related to plastid physiology and metabolism in *Nannochloropsis gaditana*.** D) The mRNA-Seq data of genes that belong to the tocopherol and tocochromanol biosynthetic pathway are reported. For each Gene ID, the predicted Gene Function and the Short name are reported. logFC indicates the fold changes in logarithmic scale of the number of reads in the three light conditions. OVER stands for significantly over-expressed genes while SUB stands for significantly down-regulated genes. Data are the average of 3 replicates.

Gene ID	Gene Function	Short name	LL vs CL	CL vs HL	LL vs HL	Log FC LLvsCL	Log FC CLvsHL	Log FC LLvsHL
Naga_100207g6	Geranylgeranyl diphosphate reductase	GGR	OVER	SUB		1,26	-0,63	0,62
Naga_100003g167	Tyrosine aminotransferase	Tyra			SUB	-0,18	-0,43	-0,57
Naga_100145g15	Tyrosine aminotransferase-like protein	TAT			OVER	0,51	0,52	1,03
Naga_100291g3	p-hydroxyphenylpyruvate dioxygenase	HPPD-1	OVER		OVER	1,30	-0,03	1,27
Naga_100113g9	p-hydroxyphenylpyruvate dioxygenase	HPPD-2				0,29	-0,26	0,03
Naga_100225g6	Homogentisate phytyltransferase	HPT (VTE2)	OVER		OVER	1,06	0,01	1,13
Naga_100062g16	Methyltransferase	VTE3	SUB		SUB	-0,79	-0,10	-0,89
Naga_100741g3	Menaquinone biosynthesis methyltransferase ubiE	VTE3				0,54	0,11	0,64
Naga_100005g102	Gamma-tocopherol methyltransferase	VTE4				0,09	-0,08	0,01
Naga_100014g6	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase	Phyllo (MenD,H,C) / MenF	OVER	SUB	OVER	2,25	-0,74	1,52
Naga_100021g41	Long-chain-fatty-acid CoA ligase	(ECHID) MenB			SUB	0,01	-0,27	-0,25
Naga_100713g1	Long-chain-fatty-acid CoA ligase	(ECHID) MenB			SUB	-0,40	0,17	-0,23
Naga_100074g14	Naphthoate synthase	(AAE14) Men E			OVER	0,86	0,38	1,24
Naga_100123g4	4-hydroxybenzoate octaprenyltransferase	UbiA				0,17	-0,21	-0,05
Naga_100398g3	3-demethylubiquinone-9 3-methyltransferase	Coq3			SUB	0,06	-0,31	-0,25
Naga_100006g73	Ubiquinone biosynthesis hydroxylase	Coq6			SUB	0,35	-0,55	-0,20
Naga_100741g3	Ubiquinone/menaquinone biosynthesis methyltransferase	Coq5				0,54	0,11	0,64

**Supplemental Data Set 3. Light response of various group of genes involved in primary carbon and lipid metabolism in *Nannochloropsis gaditana*.** A) The mRNA-Seq data of genes that belong to the glycolytic pathway and Calvin-Benson cycle are reported. For each Gene ID, the predicted Gene Function and the Short name are reported. logFC indicates the fold changes in logarithmic scale of the number of reads in the three light conditions. OVER stands for significantly over-expressed genes while SUB stands for significantly down-regulated genes. Data are the average of 3 replicates.

Gene ID	Gene Function	Short name	LL vs CL	CL vs HL	LL vs HL	Log FC LLvsCL	Log FC CLvsHL	Log FC LLvsHL	Subcellular localization
Naga_100170g2	glucokinase	GK			SUB	0.98	-0.40	-0.35	cytosol
Naga_100119g19	glucokinase	GK	SUB		SUB	-0.56	-0.28	-0.61	cytosol
Naga_100305g6	cytosolic phosphoglucose isomerase	GPI	OVER		OVER	2.48	0.38	2.86	cytosol
Naga_100268g2	pyrophosphate-dependent phosphofructose kinase	PFK				-0.11	0.20	0.09	cytosol
Naga_100138g9	pyrophosphate-fructose 6-phosphate 1-phosphotransferase subunit beta	PFK				0.38	0.04	0.42	cytosol
Naga_100049g16	fructose-bisphosphate aldolase					0.43	0.03	0.46	
Naga_100154g7	fructose-bisphosphate aldolase					0.19	0.25	0.44	
Naga_100119g3	Aldolase-type TIM barrel					0.75	-0.18	0.57	
Naga_100011g1	triose-phosphate isomerase	TPI	OVER			1.02	-0.43	0.59	Plastid
Naga_103287g1	triosephosphate isomerase	TPI			SUB	0.99	-0.32	-0.23	
Naga_100001g84	triosephosphate isomerase	TPI	SUB		SUB	-0.47	-0.28	-0.19	
Naga_100257g1	glyceraldehyde-3-phosphate dehydrogenase					-0.13	0.24	0.11	
Naga_100001g58	glyceraldehyde-3-phosphate dehydrogenase		SUB		SUB	-1.44	-0.01	-1.44	
Naga_100008g2	dihydrolipoamide s-acetyltransferase (e2 component of pyruvate dehydrogenase complex)	PDHC			SUB	-0.30	-0.13	-0.43	Plastid
Naga_100081g16	glyceraldehyde-3-phosphate dehydrogenase	PDHC	OVER		OVER	1.05	0.29	1.35	Plastid
Naga_100065g5	pyruvate dehydrogenase component x	PDHC			SUB	-0.93	-0.39	-0.42	Mitochondrion
Naga_100124g14	mitochondrial pyruvate dehydrogenase kinase	PDHC	OVER		OVER	1.33	0.52	1.84	Mitochondrion
Naga_100136g4	pyruvate dehydrogenase e1 component alpha subunit	PDHC			SUB	-0.26	-0.11	-0.39	Mitochondrion
Naga_101324g2	pyruvate dehydrogenase	PDHC				0.07	0.26	0.33	Mitochondrion
Naga_100064g27	pyruvate dehydrogenase	PDHC				0.94	0.27	0.32	Mitochondrion
Naga_100245g5	mitochondrial pyruvate dehydrogenase kinase	PDHC				0.42	0.17	0.60	Mitochondrion
Naga_100257g10	hypothetical protein	PDHC	SUB		SUB	-0.52	0.00	-0.52	Mitochondrion
Naga_100081g17	glyceraldehyde 3-phosphate	PDHC	SUB		SUB	-0.60	0.08	-0.51	Mitochondrion
Naga_100041g27	Signal transduction histidine kinase, core	PDHC				0.39	0.25	0.64	Mitochondrion
Naga_101009g1	protein phosphatase	PDHC			SUB	0.06	-0.40	-0.34	Mitochondrion
Naga_100064g4	phosphoglycerate kinase		SUB		SUB	-0.43	-0.30	-0.73	
Naga_100014g18	phosphoglycerate kinase					-0.15	0.26	0.11	
Naga_100410g3	phosphoglycerate kinase		OVER		OVER	1.08	0.02	1.10	
Naga_102489g1	phosphoglycerate mutase 1 family					-0.24	0.62	0.38	



**Supplemental Data Set 3. Light response of various group of genes involved in primary carbon and lipid metabolism in *Nannochloropsis gaditana*.** B) The mRNA-Seq data of genes that belong to lipid metabolism are reported. For each Gene ID, the predicted Gene Function and the Short name are reported. logFC indicates the fold changes in logarithmic scale of the number of reads in the three light conditions. OVER stands for significantly over-expressed genes while SUB stands for significantly down-regulated genes. Short name (I) and Short name (II) refer to the nomenclature given by Vieler et al 2012 and Li et al. 2014 respectively. Data are the average of 3 replicates.

Gene ID	Gene Function	Short name (I)	Short name (II)	LL vs CL	CL vs HL	LL vs HL	Log FC LLvsCL	Log FC CLvsHL	Log FC LLvsHL	Pathway	Localization
Naga_100002g147	pyruvate carboxylase	ACX1					-0.23	0.97	0.74	Acetyl-CoA carboxylase cc	Plastid
Naga_100061g16	carboxyl transferase	BCX					-0.15	0.46	-0.33	Acetyl-CoA carboxylase cc	Plastid
Naga_100053g29	methylcrotonoyl- carboxylase	BCR1					-0.27	0.43	-0.16	Acetyl-CoA carboxylase cc	Plastid
Naga_100079g2	propionyl- alpha subunit	BCR2					0.38	0.79	0.56	Acetyl-CoA carboxylase cc	Plastid
Naga_100594g3	branched-chain alpha-keto acid dehydrogenase subunit e2	BXP1		OVER		OVER	1.93	0.79	1.14	Acetyl-CoA carboxylase cc	Plastid
Naga_100046g34	malonyl-CoA:ACP transacylase	MCT	MCMT	OVER		OVER	1.45	-0.37	1.07	Acetyl-CoA carboxylase cc	Plastid
Naga_100037g12	3-oxoacyl-(acyl-carrier-protein) reductase	KAR1		OVER		OVER	1.87	0.71	2.04	Type II fatty acid synthase	Plastid
Naga_100113g7	beta-hydroxyacyl- acp dehydratase precursor	HAD1		OVER		OVER	1.26	-0.25	1.01	Type II fatty acid synthase	Plastid
Naga_101525g2	enoyl-acyl carrier reductase	ENR1				OVER	0.74	0.16	0.92	Type II fatty acid synthase	Plastid
Naga_101053g1	enoyl-acyl carrier reductase	ENR1					0.42	-0.09	0.33	Type II fatty acid synthase	Plastid
Naga_100059g6	rab gtpase e Thioesterase superfamily	FAT					-0.05	0.19	0.14	Type II fatty acid synthase	Plastid
Naga_100059g5	rab gtpase e Thioesterase superfamily	FAT					-0.17	0.15	0.02	Type II fatty acid synthase	Plastid
Naga_100004g8	beta-ketoacyl synthase	KAS4					-0.31	0.25	0.06	Type II fatty acid synthase	Mitochondrion
Naga_100486g4	adenylate kinase isoenzyme 6	HAD2					-0.04	0.16	0.12	Type II fatty acid synthase	Mitochondrion
Naga_100016g22	trans-2-enoyl- reductase	ENR2		SUB		SUB	-0.58	0.12	-0.46	Type II fatty acid synthase	Mitochondrion
Naga_100331g7	hypothetical protein	ACP2				SUB	-0.34	0.17	-0.17	Type II fatty acid synthase	Mitochondrion
Naga_100605g1	acetyl-coa carboxylase	ACC1		OVER		OVER	0.96	0.26	1.22		Cytosol
Naga_100086g24	beta-ketoacyl synthase						0.53	-0.16	0.58	Type I fatty acid synthase	Cytosol
Naga_100022g68	polyketide synthase				SUB	SUB	0.33	-0.74	-0.41	Type I fatty acid synthase	Cytosol
Naga_100001g49	type i polyketide synthase e polyketide synthase					SUB	-0.23	0.01	-0.22	Type I fatty acid synthase	Cytosol
Naga_100028g49	type i polyketide synthase e polyketide synthase					SUB	0.18	-0.13	0.05	Type I fatty acid synthase	Cytosol
Naga_102909g1	beta-ketoacyl synthase				SUB	SUB	0.47	-1.25	-0.78	Type I fatty acid synthase	Cytosol
Naga_101811g1	beta-ketoacyl synthase				SUB	SUB	0.84	-0.74	0.1	Type I fatty acid synthase	Cytosol
Naga_102303g1	polyketide synthase				SUB	SUB	0.59	-1.03	-0.43	Type I fatty acid synthase	Cytosol
Naga_102722g1	polyketide synthase				SUB	SUB	0.70	-1.19	-0.49	Type I fatty acid synthase	Cytosol
Naga_101380g1	polyketide synthase				SUB	SUB	0.61	-0.80	-0.19	Type I fatty acid synthase	Cytosol
Naga_103117g1	polyketide synthase				SUB	SUB	0.64	-0.68	0.04	Type I fatty acid synthase	Cytosol
Naga_100093g2	polyketide synthase						-0.15	-0.29	-0.14	Type I fatty acid synthase	Cytosol
Naga_100007g48	polyketide synthase						0.37	-0.31	0.06	Type I fatty acid synthase	Cytosol
Naga_100002g173	3-oxoacyl-(acyl-carrier-protein) synthase 2			OVER		OVER	1.45	-0.55	0.90	Type I fatty acid synthase	Cytosol





Naga_100013g52	Fatty acid desaturase type 2			OVER		OVER	1.75	0.34	2.08
Naga_100028g29	Kua-ubiquitin conjugating enzyme hybrid	DES3t		SUB	SUB	SUB	-0.76	-1.90	-2.66
Naga_100027g27	stearoyl- desaturase 5	DES9		OVER		OVER	2.58	-0.34	2.24
Naga_100092g5	monogalactosyldiacylglycerol synthase	DES12					0.60	0.01	0.79
Naga_100061g21	delta-6 fatty acid desaturase	DES6		OVER	SUB		1.47	-1.01	0.46
Naga_100273g7	delta 5 fatty acid desaturase	DES5		OVER	SUB		1.19	-0.73	0.46
Naga_100017g62	Fatty acid hydroxylase						0.39	0.01	0.40
Naga_100545g1	fatty acid desaturase			OVER	SUB	OVER	2.31	-0.82	1.49
Naga_100020g80	gns1 sur4 family protein	ELO					0.33	0.40	0.73
Naga_100003g61	26s protease regulatory subunit 6a-a			SUB		SUB	-0.64	-0.19	-0.83
Naga_100399g1	fatty-acyl e fatty-acyl elongase			SUB		SUB	-0.51	-0.09	-0.42
Naga_100004g102	fatty-acy						0.42	-0.41	-0.39
Naga_100017g49	elongation of very long chain fatty acids protein 6			SUB		SUB	-0.72	0.06	-0.64
Naga_100162g4	fatty-acyl elongase					SUB	-0.40	-0.02	-0.43
Naga_100083g23	fatty-acyl						0.77	-0.02	0.76
Naga_100608g3	fatty-acyl					OVER	0.81	-0.10	0.90
Naga_100162g5	fatty-acyl elongase			SUB			-0.44	0.44	-0.14
Naga_100649g1	long-chain acyl- synthetase 7					SUB	-0.31	-0.06	-0.37
Naga_100047g8	amp-dependent synthetase and ligase						-0.38	0.32	-0.06
Naga_100014g59	long chain acyl-coa synthetase					SUB	-0.05	-0.27	-0.32
Naga_101051g1	amp-dependent synthetase and ligase e long-chain acyl- synthetase						0.45	0.41	0.85
Naga_100158g20	amp-dependent synthetase and ligase e long-chain acyl- synthetase						0.38	0.32	0.71
Naga_100035g43	fatty-acid- ligase fadd9						0.19	-0.31	-0.12
Naga_100028g54	acetyl-coenzyme a synthetase					SUB	-0.36	-0.19	-0.54
Naga_100405g2	acetyl- synthetase			SUB		SUB	-0.76	0.06	-0.70
Naga_100012g66	amp-dependent synthetase and ligase			OVER		OVER	1.70	-0.28	1.42
Naga_100002g172	acyl- dehydrogenase			SUB		SUB	-1.55	0.65	-0.91
Naga_100017g26	acyl- dehydrogenase					SUB	-0.29	-0.16	-0.36
Naga_100096g20	isobutyryl- mitochondrial			SUB			-0.48	0.39	-0.09
Naga_100077g15	peroxisomal acyl-coenzyme a oxidase 1						-0.02	0.00	-0.02
Naga_100466g3	trifunctional enzyme subunit mitochondrial			SUB		SUB	-1.14	-0.11	-0.97
Naga_102524g1	beta-ketoacyl- thiolase e acetyl- acetyltransferase			SUB		SUB	-0.91	0.41	-0.50
Naga_100268g10	3-hydroxyacyl-CoA dehydrogenase nad-binding protein						-0.14	-0.26	-0.10
Naga_100042g8	3-ketoacyl- mitochondrial			SUB		SUB	-0.74	0.52	-0.22
Naga_100076g17	3-ketoacyl- thiolase peroxisomal						-0.41	0.52	0.11
Naga_100746g2	acetyl- acetyltransferase					SUB	-0.19	-0.24	-0.42
Naga_100009g108	delta( )-delta( )-dienoyl- mitochondrial-like protein						-0.30	0.27	-0.03
Naga_100267g8	methylglutaconyl- mitochondrial				SUB	SUB	-0.46	-1.12	-1.58
Naga_100196g7	enoyl- hydratase			SUB		SUB	-0.93	0.11	-0.74
Naga_100006g36	peroxisomal -trans-enoil- isomerase e 3-hydroxybutyryl- dehydratase						-0.07	0.09	0.02
Naga_100006g32	peroxisomal -trans-enoil- isomerase e 3-hydroxybutyryl- dehydratase			SUB		SUB	-0.61	-0.46	-1.08
Naga_100034g21	-like dehydratase					OVER	0.78	-0.16	0.62
Naga_100002g56	peroxisomal -dienoyl- reductase						0.29	-0.09	0.20
Naga_100172g8	peroxisomal trans-2-enoil- reductase						0.30	0.05	0.34
Naga_102434g1	peroxisomal hydratase-dehydrogenase-epimerase					SUB	-0.27	0.06	-0.21
Naga_100065g28	lysophospholipase ii				SUB	SUB	-0.21	-0.67	-0.88
Naga_100017g37	lysophospholipase-like 1					SUB	-0.06	-0.45	-0.53
Naga_100049g37	cysteine proteinase-like protein					SUB	-0.40	-0.16	-0.56
Naga_100247g4	Phospholipase A2			OVER		OVER	1.64	0.64	2.28
Naga_100055g28	Acyl transferase/acyl hydrolase/lysophospholipase						0.64	0.01	0.64
Naga_100677g1	patatin-like serine						0.46	-0.01	0.45
Naga_100020g21	sn1-specific diacylglycerol lipase alpha-like protein					SUB	-0.26	-0.24	-0.50
Naga_100948g2	hypothetical protein			SUB		SUB	-0.65	-0.01	-0.71
Naga_100171g1	Lipase						0.29	0.22	0.50
Naga_100180g7	myb family transcription factor			SUB		SUB	-0.44	0.24	-0.20
Naga_100011g88	lipase containing protein						-0.02	0.22	0.19
Naga_100057g25	triacylglycerol lipase						0.82	-0.19	0.72
Naga_100039g12	lipase family protein						0.34	0.06	0.40
Naga_102685g1	conserved unknown protein e patatin-like serine						-0.01	0.03	0.02
Naga_100012g87	Acyl transferase/acyl hydrolase/lysophospholipase			SUB		SUB	-0.97	-0.01	-1.01

Naga_100008g43	pheophytinase					0.90	-0.33	0.57
Naga_101000g2	alpha beta hydrolase fold protein					-0.91	0.00	-0.91
Naga_100072g8	hypothetical protein					0.03	0.22	0.25
Naga_100003g118	alpha beta hydrolase fold protein				SUB	-0.04	-0.27	-0.30
Naga_100013g71	hypothetical protein				SUB	0.17	-0.55	-0.38
Naga_100031g10	hypothetical protein			SUB	SUB	0.31	-0.70	-0.40
Naga_100321g3	serine protease family s09x				SUB	-0.33	-0.21	-0.44
Naga_100005g24	serine protease family e serine protease family					0.11	-0.05	0.06
Naga_100008g11	serine protease family e serine protease family					0.57	0.13	0.70
Naga_100044g12	crystal protein					0.58	0.03	0.60
Naga_100043g20	carboxylesterase type b					0.11	0.05	0.16
Naga_100414g2	hypothetical protein					-0.23	0.37	0.14
Naga_100391g2	carboxyl-terminal processing protease					0.50	-0.32	0.18
Naga_100039g23	cysteine proteinase				SUB	-0.16	-0.37	-0.54
Naga_100288g3	cysteine proteinase			SUB	SUB	-0.58	0.38	-0.20
Naga_100226g13	dual specificity protein phosphatase					-0.23	0.50	0.27
Naga_100156g4	Acyl transferase/acyl hydrolase/lysophospholipase				SUB	-0.26	-0.33	-0.60
Naga_100270g8	cysteine protease				SUB	-0.30	-0.12	-0.42
Naga_100324g4	dipeptidyl aminopeptidase			SUB	SUB	-0.64	-0.33	-0.97
Naga_100036g28	lysosomal pro-x				SUB	-0.21	-0.25	-0.46
Naga_100007g56	Cyclic nucleotide-binding protein			SUB	SUB	-0.50	-0.27	-0.77
Naga_100101g5	notchless protein				SUB	-0.35	-0.36	-0.71
Naga_100002g45	dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit dad1				SUB	-0.38	-0.37	-0.76
Naga_100040g9	Acyl transferase/acyl hydrolase/lysophospholipase			SUB	SUB	-2.37	-0.70	-3.07
Naga_100093g9	prolyl endopeptidase					0.12	-0.12	0.00
Naga_100012g53	hydrolase-like protein					0.55	0.21	0.76
Naga_100002g136	Peptidase S10 serine carboxypeptidase					0.29	0.14	0.44
Naga_100002g135	serine carboxypeptidase-like protein					0.58	0.35	0.45
Naga_100101g12	Lipase class 3					0.61	-0.37	0.24
Naga_100124g15	oligopeptidase b					0.21	-0.07	0.14
Naga_100008g60	pheromone processing carboxypeptidase					-0.08	0.08	0.00
Naga_100331g4	s-formylglutathione hydrolase			SUB	SUB	-0.75	-0.43	-1.18
Naga_100019g38	embryogenesis-associated protein emb8					-0.04	-0.05	-0.09

**Supplemental Data Set 4. LC-ESI-MS of semi-polar metabolites in *Nannochloropsis gaditana* cells grown in three different light conditions.** Data are expressed as fold on the level of the metabolite in the Low Light (LL) conditions, and the pValue of a Tukey's t-test is reported (\*= 0.005<pValue<0.05; \*\*= 0.001<pValue<0.005; \*\*\*= 0.001<pValue<0.005). Up- (red) and down- (green) accumulated metabolites are indicated. Data are the average of at least 4 extractions. For more details, see materials and methods.

	LL		ML/LL				HL/LL			
	AVG	ST.DEV.	AVG	ST.DEV.	STAT.	VARIAT.	AVG	ST.DEV.	STAT.	VARIAT.
<b>ALKALOIDS</b>										
2,4-Dimethylpyridine	1,00	0,20	<b>0,23</b>	0,02	**	DW	<b>0,52</b>	0,13	*	DW
<b>AMINES</b>										
N-Acetylglucosamine	1,00	0,18	<b>1,84</b>	0,31	**	UP	<b>3,50</b>	1,01	*	UP
N8-Acetylspermidine	1,00	0,15	<b>0,84</b>	0,22			<b>2,34</b>	0,07	***	UP
N-Carboxyethyl-γ-aminobutyric acid	1,00	0,20	<b>2,13</b>	0,55	***	UP	<b>1,60</b>	0,03	*	UP
Catechol	1,00	0,16	<b>1,66</b>	0,27	*	UP	<b>1,72</b>	0,43		
N-Methylethanolamine-phosphate	1,00	0,24	<b>0,75</b>	0,13			<b>0,48</b>	0,09	*	DW
Spermidine	1,00	0,27	<b>6,11</b>	1,42	***	UP	<b>10,90</b>	0,97	***	UP
Tyramine	1,00	0,25	<b>2,48</b>	0,43	***	UP	<b>0,40</b>	0,07	***	DW
<b>AMINOACIDS</b>										
N-Acetyl-b-alanine	1,00	0,18	<b>1,18</b>	0,11			<b>1,32</b>	0,37		
O-Acetylcarnitine	1,00	0,22	<b>0,37</b>	0,05	*	DW	<b>1,11</b>	0,26		
N-Acetylglutamate	1,00	0,05	<b>1,60</b>	0,17	**	UP	<b>1,20</b>	0,10		
N-e-Acetylysine	1,00	0,05	<b>1,32</b>	0,21			<b>2,29</b>	0,23	***	UP

N-Acetylnithine	1,00	0,14	<b>0,98</b>	0,18			<b>3,14</b>	0,69	**	<b>UP</b>
N-Acetylserine	1,00	0,22	<b>1,30</b>	0,17			<b>0,46</b>	0,06	***	<b>DW</b>
S-Adenosylhomocysteine	1,00	0,29	<b>2,09</b>	0,53	*	<b>UP</b>	<b>1,18</b>	0,32		
S-Adenosylmethionine	1,00	0,05	<b>2,31</b>	0,40	**	<b>UP</b>	<b>1,25</b>	0,07		
Agmatine	1,00	0,25	<b>1,87</b>	0,22	*	<b>UP</b>	<b>4,94</b>	1,35	*	<b>UP</b>
Alanine	nd	nd	<b>1257,07</b>	355,41			<b>958,96</b>	126,26		
Arginine	1,00	0,12	<b>0,94</b>	0,18			<b>1,65</b>	0,21	*	<b>UP</b>
Argininosuccinate	1,00	0,20	<b>0,63</b>	0,11			<b>0,22</b>	0,03	***	<b>DW</b>
Asparagine	1,00	0,27	<b>2,74</b>	0,20	***	<b>UP</b>	<b>3,61</b>	0,27	*	<b>UP</b>
Aspartate	1,00	0,07	<b>0,58</b>	0,16	**	<b>DW</b>	<b>1,65</b>	0,36		
gamma-Butyrobetaine	nd	nd	<b>33,75</b>	5,07			<b>62,40</b>	8,73		
Carbamoylaspartate	1,00	0,16	<b>0,73</b>	0,07			<b>0,74</b>	0,04		
N-Carbamylglutamate	1,00	0,12	<b>1,84</b>	0,12	**	<b>UP</b>	<b>1,56</b>	0,23		
Carnitine	1,00	0,08	<b>2,67</b>	0,66	***	<b>UP</b>	<b>4,95</b>	0,36	**	<b>UP</b>
Citrulline	1,00	0,16	<b>0,12</b>	0,01	***	<b>DW</b>	<b>0,18</b>	0,05	***	<b>DW</b>
Cystathionine	1,00	0,15	<b>0,72</b>	0,20			<b>0,35</b>	0,07	**	<b>DW</b>
Cysteine	nd	nd	<b>nd</b>	nd			<b>12,08</b>	0,41		
Cysteine sulfinate	1,00	0,24	<b>0,69</b>	0,05			<b>1,41</b>	0,12		
Cystine	nd	nd	<b>12,06</b>	0,36			<b>1049,93</b>	299,79		
N-Formylmethionine	nd	nd	<b>nd</b>	nd			<b>178,48</b>	36,33		
Gamma-guanidinobutyrate	1,00	0,19	<b>0,01</b>	0,00	***	<b>DW</b>	<b>nd</b>	nd		
Glutamic Acid	1,00	0,22	<b>1,33</b>	0,36			<b>0,46</b>	0,06	**	<b>DW</b>
Glutamate-1-semialdehyde	1,00	0,20	<b>0,79</b>	0,07			<b>0,67</b>	0,12		
Glutamine	1,00	0,12	<b>1,90</b>	0,43	**	<b>UP</b>	<b>4,43</b>	0,07	***	<b>UP</b>
Glycine-Betaine	1,00	0,01	<b>nd</b>	nd			<b>nd</b>	nd		
Histidine	1,00	0,07	<b>1,97</b>	0,41	***	<b>UP</b>	<b>2,37</b>	0,46	**	<b>UP</b>
Homospermidine	1,00	0,01	<b>1,96</b>	0,52	***	<b>UP</b>	<b>2,62</b>	0,60	**	<b>UP</b>
Hydroxyproline	1,00	0,12	<b>1,11</b>	0,16			<b>1,32</b>	0,30		
Hypotaurine	1,00	0,11	<b>0,81</b>	0,05			<b>0,73</b>	0,15		
Leucine/Isoleucine/Norleucine	1,00	0,26	<b>1,62</b>	0,03	**	<b>UP</b>	<b>1,86</b>	0,53		
Lysine	1,00	0,24	<b>3,83</b>	0,38	***	<b>UP</b>	<b>1,76</b>	0,43		
Methionine	1,00	0,20	<b>2,27</b>	0,45	***	<b>UP</b>	<b>3,05</b>	0,29	***	<b>UP</b>
Methionine sulfoxide	1,00	0,05	<b>6,83</b>	1,74	***	<b>UP</b>	<b>12,76</b>	0,34	***	<b>UP</b>
N-Methylglutamate	1,00	0,05	<b>1,26</b>	0,32			<b>0,89</b>	0,25		
Ornithine	1,00	0,07	<b>0,63</b>	0,11	**	<b>DW</b>	<b>1,41</b>	0,29		
5-Oxoproline	1,00	0,08	<b>1,27</b>	0,27			<b>1,00</b>	0,20		
Phenilalanine	1,00	0,23	<b>1,93</b>	0,25	**	<b>UP</b>	<b>4,21</b>	1,00	**	<b>UP</b>
Phospho-L-serine	1,00	0,12	<b>0,82</b>	0,19			<b>3,27</b>	0,96	**	<b>UP</b>

Proline	1,00	0,13	<b>0,47</b>	0,07	***	DW	<b>1,01</b>	0,28		
Propionyl-L-carnitine	1,00	0,17	<b>1,23</b>	0,04			<b>1,58</b>	0,45		
Succinic semialdehyde	1,00	0,02	<b>0,85</b>	0,09			<b>0,98</b>	0,18		
Threonine/homoserine	1,00	0,26	<b>2,35</b>	0,52	**	UP	<b>2,56</b>	0,27	**	UP
Trimethyllysine	1,00	0,26	<b>3,15</b>	0,88	***	UP	<b>1,41</b>	0,36		
Tryptophan	1,00	0,12	<b>2,72</b>	0,31	***	UP	<b>2,18</b>	0,61	*	UP
Tryptophan derivative	1,00	0,14	<b>3,06</b>	0,33	***	UP	<b>2,41</b>	0,59	*	UP
Tyrosine	1,00	0,18	<b>2,06</b>	0,33	**	UP	<b>2,28</b>	0,61	*	UP
3-Ureidopropionate	1,00	0,14	<b>0,81</b>	0,20			<b>0,58</b>	0,13	**	DW
Valine/Norvaline	1,00	0,19	<b>1,29</b>	0,11			<b>1,72</b>	0,48		
<b>NUCLEIC ACID METABOLISM</b>										
Adenine	1,00	0,04	<b>4,53</b>	0,20	***	UP	<b>3,70</b>	0,70	**	UP
Adenosine	1,00	0,09	<b>4,94</b>	0,85	***	UP	<b>4,66</b>	0,15	***	UP
Adenosine Diphosphate	1,00	0,01	<b>0,28</b>	0,08	***	DW	<b>0,27</b>	0,06	**	DW
Adenosine-5-monophosphate	1,00	0,07	<b>0,28</b>	0,06	**	DW	<b>0,29</b>	0,08	***	DW
Cytidine	1,00	0,13	<b>6,39</b>	1,70	***	UP	<b>3,09</b>	0,63	**	UP
Cytosine	1,00	0,13	<b>5,21</b>	0,88	***	UP	<b>2,66</b>	0,55	**	UP
5'-Deoxy-5'-methylthioadenosine	1,00	0,25	<b>0,91</b>	0,10			<b>3,17</b>	0,79	***	UP
Dihydroroate	1,00	0,20	<b>1,05</b>	0,26			<b>2,20</b>	0,22	*	UP
2,4-Dihydropyrimidine-5-carboxylate	1,00	0,16	<b>0,79</b>	0,11			<b>0,88</b>	0,16		
Guanine	1,00	0,12	<b>0,62</b>	0,13	**	DW	<b>0,60</b>	0,14	*	DW
Guanosine	1,00	0,15	<b>0,92</b>	0,09			<b>0,86</b>	0,24		
Inosine	1,00	0,09	<b>0,70</b>	0,12			<b>1,26</b>	0,28		
5-Methylthioadenosine	1,00	0,25	<b>0,91</b>	0,10			<b>3,17</b>	0,79	*	UP
Picolinic Acid	1,00	0,08	<b>1,63</b>	0,17	**	UP	<b>4,53</b>	0,79	**	UP
Pseudo-uridine	1,00	0,26	<b>2,85</b>	0,67	**	UP	<b>0,97</b>	0,15		
Uracil	1,00	0,08	<b>3,86</b>	0,93	**	UP	<b>2,27</b>	0,65	**	UP
Uridine	1,00	0,26	<b>2,85</b>	0,67	**	UP	<b>0,97</b>	0,15		
<b>ORGANIC ACIDS</b>										
Aconitic Acid	nd	nd	<b>16225,46</b>	3881,07			<b>14299,41</b>	2584,43		
Allantoic Acid	1,00	0,03	<b>0,09</b>	0,02	***	DW	<b>0,03</b>	0,01	**	DW
4-Amino-4-deoxychorismate	nd	nd	<b>134,76</b>	30,41			<b>91,40</b>	19,99		
5-Aminolevulinat	1,00	0,27	<b>4,49</b>	0,85	**	UP	<b>4,65</b>	0,87	**	UP
Azelaic Acid	1,00	0,23	<b>0,33</b>	0,05	***	DW	<b>0,74</b>	0,13		
Citraconate	1,00	0,26	<b>1,04</b>	0,23			<b>0,59</b>	0,04	**	DW
Citric Acid	nd	nd	<b>30448,70</b>	8525,69			<b>28927,63</b>	7593,27		
2,6-Dimethoxybenzoic acid	1,00	0,03	<b>nd</b>	nd			<b>1,19</b>	0,25		
3-Hydroxy-3-methylglutarate	1,00	0,05	<b>1,04</b>	0,04			<b>1,60</b>	0,31		
2-Methylcitric Acid	1,00	0,26	<b>8,36</b>	0,98	***	UP	<b>15,10</b>	3,17	***	UP

Mevalonate	1,00	0,13	<b>1,04</b>	0,25			<b>0,65</b>	0,08	*	<b>DW</b>
Monomethylphosphate	1,00	0,20	<b>0,64</b>	0,05	*	<b>DW</b>	<b>1,50</b>	0,36		
Phenylpyruvic Acid	1,00	0,20	<b>2,04</b>	0,07	**	<b>UP</b>	<b>4,76</b>	0,43	***	<b>UP</b>
Phosphoric Acid-propyl ester	1,00	0,09	<b>1,24</b>	0,32			<b>1,74</b>	0,33	*	<b>UP</b>
Pipecolic Acid	1,00	0,19	<b>1,75</b>	0,43	**	<b>UP</b>	<b>1,75</b>	0,41		
Pyroglutamic Acid	1,00	0,08	<b>1,27</b>	0,27			<b>1,00</b>	0,20		
Saccharic Acid	1,00	0,29	<b>1,85</b>	0,48	*	<b>UP</b>	<b>0,95</b>	0,10		
Shikimic Acid	1,00	0,27	<b>0,28</b>	0,07	***	<b>DW</b>	<b>0,30</b>	0,06	***	<b>DW</b>
Tartaric Acid	1,00	0,21	<b>0,70</b>	0,02			<b>0,50</b>	0,04	**	<b>DW</b>
<b>PEPTIDES</b>										
Cysteine-glutathione disulfide	1,00	0,01	<b>1,80</b>	0,33	**	<b>UP</b>	<b>15,06</b>	3,07	***	<b>UP</b>
Cysteinylglycine	1,00	0,24	nd	nd			nd	nd		
Dipeptide	1,00	0,26	<b>2,05</b>	0,27	*	<b>UP</b>	<b>4,74</b>	1,36	**	<b>UP</b>
Glutathione	1,00	0,28	<b>0,98</b>	0,22			<b>0,13</b>	0,03	***	<b>DW</b>
Glutathione Oxized	1,00	0,01	<b>4,72</b>	1,06	**	<b>UP</b>	<b>4,10</b>	1,11	**	<b>UP</b>
His Gln Cys Cys	1,00	0,09	<b>14,67</b>	4,19	**	<b>UP</b>	<b>5,77</b>	1,57	**	<b>UP</b>
Pro Asp Asn	1,00	0,20	<b>4,23</b>	0,34	**	<b>UP</b>	<b>4,26</b>	1,09	*	<b>UP</b>
Tripeptide	1,00	0,20	<b>0,42</b>	0,08	*	<b>DW</b>	<b>3,94</b>	0,73	***	<b>UP</b>
<b>PHENOLIC ACIDS</b>										
Chlorogenic Acid	1,00	0,16	<b>0,63</b>	0,16			<b>0,66</b>	0,15		
Cinnamic Acid	1,00	0,07	<b>0,82</b>	0,10			<b>0,86</b>	0,20		
Hydroxycinnamic acid	1,00	0,20	<b>2,02</b>	0,06	**	<b>UP</b>	<b>4,40</b>	1,08	*	<b>UP</b>
Syringate	1,00	0,28	<b>3,14</b>	0,35	***	<b>UP</b>	<b>3,19</b>	0,34	*	<b>UP</b>
<b>SUGARS</b>										
Amylopectin	nd	nd	nd	nd			<b>204,56</b>	54,13		
Bis-d-Fructose-2',1'-dianhydride	1,00	0,18	<b>4,20</b>	0,43	***	<b>UP</b>	<b>4,32</b>	1,06	***	<b>UP</b>
Conduritol-β-epoxide	1,00	0,15	<b>4,06</b>	0,91	***	<b>UP</b>	<b>5,65</b>	0,65	***	<b>UP</b>
Dihydroacetone-P (DHAP)	1,00	0,20	<b>0,51</b>	0,07	*	<b>DW</b>	<b>0,24</b>	0,06	***	<b>DW</b>
Fructose/Galactose/Glucose/Inositol/Mannose	1,00	0,10	<b>2,66</b>	0,72			<b>0,92</b>	0,14		
Fructose-6P/Glucose-6P/Inositol-6P	1,00	0,18	<b>1,43</b>	0,28			<b>1,50</b>	0,35		
Fucose	1,00	0,15	<b>2,72</b>	0,21	***	<b>UP</b>	<b>2,88</b>	0,55	**	<b>UP</b>
Galactosylglycerol	1,00	0,22	<b>4,27</b>	0,71	**	<b>UP</b>	<b>1,48</b>	0,34		

<b>α-Glycero-P/Glycerol-3-P</b>	1,00	0,24	<b>0,78</b>	0,14			<b>0,96</b>	0,25		
<b>Inositol-pentakisphosphate</b>	1,00	0,26	<b>0,85</b>	0,22			<b>2,21</b>	0,47	**	<b>UP</b>
<b>Maltotriose</b>	1,00	0,19	<b>5,34</b>	1,04	***	<b>UP</b>	<b>9,21</b>	1,18	***	<b>UP</b>
<b>Mannan</b>	nd	nd	<b>379,56</b>	12,86			<b>1324,93</b>	316,84		
<b>Mannosamine</b>	1,00	0,23	<b>0,97</b>	0,10			<b>0,71</b>	0,12		
<b>Raffinose</b>	1,00	0,12	<b>7,56</b>	1,61	***	<b>UP</b>	<b>9,84</b>	2,15	***	<b>UP</b>
<b>Rhamnose</b>	1,00	0,15	<b>2,72</b>	0,21	***	<b>UP</b>	<b>2,95</b>	0,62	**	<b>UP</b>
<b>Ribose-5P</b>	1,00	0,23	<b>1,40</b>	0,36			<b>0,62</b>	0,04	*	<b>DW</b>
<b>Ribulose-5P</b>	1,00	0,03	<b>1,23</b>	0,28			<b>0,55</b>	0,04	****	<b>DW</b>
<b>Sorbitol/Mannitol</b>	1,00	0,11	<b>3,41</b>	0,81	***	<b>UP</b>	<b>5,33</b>	0,81	***	<b>UP</b>
<b>Sucrose and isomers</b>	1,00	0,27	<b>2,42</b>	0,04	***	<b>UP</b>	<b>1,71</b>	0,48		
<b>Trehalose</b>	1,00	0,05	<b>4,65</b>	0,67	***	<b>UP</b>	<b>4,37</b>	0,92	***	<b>UP</b>
<b>Urocanate</b>	1,00	0,18	<b>1,76</b>	0,45	**	<b>UP</b>	<b>0,73</b>	0,00		
<b>Xylitol/Arabitol/Ribitol</b>	nd	nd	<b>265,61</b>	56,20			<b>82,04</b>	12,86		
<b>Xylose/Arabinose</b>	1,00	0,05	<b>1,47</b>	0,04	***	<b>UP</b>	<b>nd</b>	nd		
<b>VITAMINS</b>										
<b>Biotin</b>	1,00	0,23	<b>10,04</b>	1,98	***	<b>UP</b>	<b>4,89</b>	1,34	*	<b>UP</b>
<b>Dehydroascorbate</b>	nd	nd	<b>16225,67</b>	3881,10			<b>14305,44</b>	2568,44		
<b>Desthiobiotin</b>	1,00	0,24	<b>4,68</b>	1,10	***	<b>UP</b>	<b>7,73</b>	1,04	*	<b>UP</b>
<b>10-Formyltetrahydrofolate</b>	1,00	0,28	<b>11,56</b>	3,07	***	<b>UP</b>	<b>15,78</b>	4,48		<b>UP</b>
<b>Niacin</b>	1,00	0,10	<b>1,69</b>	0,11	***	<b>UP</b>	<b>4,70</b>	0,90	*	<b>UP</b>
<b>Nicotinamide</b>	1,00	0,09	<b>0,54</b>	0,16	*	<b>DW</b>	<b>0,56</b>	0,12		<b>DW</b>
<b>Pantothenic Acid</b>	1,00	0,25	<b>2,34</b>	0,14	**	<b>UP</b>	<b>4,53</b>	1,00	*	<b>UP</b>
<b>Pyridoxal</b>	1,00	0,19	<b>0,30</b>	0,01	***	<b>DW</b>	<b>0,54</b>	0,14	*	<b>DW</b>
<b>Riboflavin</b>	1,00	0,06	<b>1,45</b>	0,14	*	<b>UP</b>	<b>3,62</b>	0,92	*	<b>UP</b>
<b>Tetrahydrofolate</b>	nd	nd	<b>80,16</b>	22,16			<b>80,35</b>	20,71		
<b>Thiamine</b>	1,00	0,05	<b>1,53</b>	0,25	*	<b>UP</b>	<b>1,95</b>	0,49		<b>UP</b>

## References

- Abida, H. et al. (2015). Membrane Glycerolipid Remodeling Triggered by Nitrogen and Phosphorus Starvation in *Phaeodactylum tricornutum*. *Plant Physiol.* 167: 118–136.
- Alboresi, A., Gerotto, C., Giacometti, G.M., Bassi, R., and Morosinotto, T. (2010). *Physcomitrella patens* mutants affected on heat dissipation clarify the evolution of photoprotection mechanisms upon land colonization. *Proc. Natl. Acad. Sci. U. S. A.* 107: 11128–33.
- Allen, A.E., Dupont, C.L., Oborník, M., Horák, A., Nunes-Nesi, A., McCrow, J.P., Zheng, H., Johnson, D.A., Hu, H., Fernie, A.R., and Bowler, C. (2011). Evolution and metabolic significance of the urea cycle in photosynthetic diatoms. *Nature* 473: 203–7.
- Bailey, S., Walters, R.G., Jansson, S., and Horton, P. (2001). Acclimation of *Arabidopsis thaliana* to the light environment: the existence of separate low light and high light responses. *Planta* 213: 794–801.
- Ballottari, M., Dall’Osto, L., Morosinotto, T., and Bassi, R. (2007). Contrasting behavior of higher plant photosystem I and II antenna systems during acclimation. *J. Biol. Chem.* 282: 8947–8958.
- Baroli, I., Do, A.D., Yamane, T., and Niyogi, K.K. (2003). Zeaxanthin accumulation in the absence of a functional xanthophyll cycle protects *Chlamydomonas reinhardtii* from photooxidative stress. *Plant Cell* 15: 992–1008.
- Beer, A., Gundermann, K., Beckmann, J., and Büchel, C. (2006). Subunit composition and pigmentation of fucoxanthin-chlorophyll proteins in diatoms: evidence for a subunit involved in diadinoxanthin and diatoxanthin binding. *Biochemistry* 45: 13046–53.
- Bogen, C., Al-Dilaimi, A., Albersmeier, A., Wichmann, J., Grundmann, M., Rupp, O., Lauersen, K.J., Blifernez-Klassen, O., Kalinowski, J., Goesmann, A., Mussnug, J.H., and Kruse, O. (2013). Reconstruction of the lipid metabolism for the microalga *Monoraphidium neglectum* from its genome sequence reveals characteristics suitable for biofuel production. *BMC Genomics* 14: 926.
- Botté, C.Y. et al. (2013). Atypical lipid composition in the purified relict plastid (apicoplast) of malaria parasites. *Proc. Natl. Acad. Sci. U. S. A.* 110: 7506–11.
- Botté, C.Y., Dubar, F., McFadden, G.I., Maréchal, E., and Biot, C. (2012). *Plasmodium falciparum* apicoplast drugs: targets or off-targets? *Chem. Rev.* 112: 1269–83.
- Boussiba, S., Vonshak, A., Cohen, Z., Avissar, Y., and Richmond, A. (1987). Lipid and biomass production by the halotolerant microalga *Nannochloropsis salina*. *Biomass* 12: 37–47.
- Brooks, C.F., Johnsen, H., van Dooren, G.G., Muthalagi, M., Lin, S.S., Bohne, W., Fischer, K., and



- Striepen, B. (2010). The toxoplasma apicoplast phosphate translocator links cytosolic and apicoplast metabolism and is essential for parasite survival. *Cell Host Microbe* 7: 62–73.
- Campagna, D., Telatin, A., Forcato, C., Vitulo, N., and Valle, G. (2013). PASS-bis: a bisulfite aligner suitable for whole methylome analysis of Illumina and SOLiD reads. *Bioinformatics* 29: 268–70.
- Cazzonelli, C.I. and Pogson, B.J. (2010). Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci.* 15: 266–274.
- Chen, M. and Thelen, J.J. (2010). The plastid isoform of triose phosphate isomerase is required for the postgerminative transition from heterotrophic to autotrophic growth in *Arabidopsis*. *Plant Cell* 22: 77–90.
- Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674–6.
- Contento, A.L., Kim, S.-J., and Bassham, D.C. (2004). Transcriptome profiling of the response of *Arabidopsis* suspension culture cells to Suc starvation. *Plant Physiol.* 135: 2330–47.
- Corteggiani Carpinelli, E., Telatin, A., Vitulo, N., Forcato, C., D’Angelo, M., Schiavon, R., Vezzi, A., Giacometti, G.M., Morosinotto, T., and Valle, G. (2014). Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion. *Mol. Plant* 7: 323–35.
- Couée, I., Sulmon, C., Gouesbet, G., and El Amrani, A. (2006). Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. *J. Exp. Bot.* 57: 449–59.
- Cournac, L., Redding, K., Ravenel, J., Rumeau, D., Josse, E.M., Kuntz, M., and Peltier, G. (2000). Electron flow between photosystem II and oxygen in chloroplasts of photosystem I-deficient algae is mediated by a quinol oxidase involved in chlororespiration. *J. Biol. Chem.* 275: 17256–62.
- Davis, M.C., Fiehn, O., and Durnford, D.G. (2013). Metabolic acclimation to excess light intensity in *Chlamydomonas reinhardtii*. *Plant. Cell Environ.* 36: 1391–1405.
- Dietz, K.-J. (2015). Efficient high light acclimation involves rapid processes at multiple mechanistic levels. *J. Exp. Bot.* 66: 2401–2414.
- Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792–7.
- Elbein, A.D., Pan, Y.T., Pastuszak, I., and Carroll, D. (2003). New insights on trehalose: a multifunctional molecule. *Glycobiology* 13: 17R–27R.

- Endo, K., Mizusawa, N., Shen, J.-R., Yamada, M., Tomo, T., Komatsu, H., Kobayashi, M., Kobayashi, K., and Wada, H. (2015). Site-directed mutagenesis of amino acid residues of D1 protein interacting with phosphatidylglycerol affects the function of plastoquinone QB in photosystem II. *Photosynth. Res.*
- Escoubas, J.M., Lomas, M., LaRoche, J., and Falkowski, P.G. (1995). Light intensity regulation of cab gene transcription is signaled by the redox state of the plastoquinone pool. *Proc. Natl. Acad. Sci. U. S. A.* 92: 10237–41.
- Falkowski, P.G., Katz, M.E., Knoll, A.H., Quigg, A., Raven, J.A., Schofield, O., and Taylor, F.J.R. (2004). The evolution of modern eukaryotic phytoplankton. *Science* 305: 354–60.
- Folch, J., Lees, M., and Sloane Stanley, G.H. (1957). A simple method for the isolation and purification of total lipides from animal tissues. *J. Biol. Chem.* 226: 497–509.
- Goyal, A., Brown, A.D., and Lilley, R.M. (1988). The response of the green halotolerant alga *Dunaliella* to osmotic stress: effects on pyridine nucleotide contents. *Biochim. Biophys. Acta - Bioenerg.* 936: 20–28.
- Guillard, R.R. and Ryther, J.H. (1962). Studies of marine planktonic diatoms. I. *Cyclotella nana* Hustedt, and *Detonula confervacea* (Cleve) Gran. *Can. J. Microbiol.* 8: 229–39.
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59: 307–21.
- Ho, S.-H., Chan, M.-C., Liu, C.-C., Chen, C.-Y., Lee, W.-L., Lee, D.-J., and Chang, J.-S. (2014). Enhancing lutein productivity of an indigenous microalga *Scenedesmus obliquus* FSP-3 using light-related strategies. *Bioresour. Technol.* 152: 275–82.
- Houille-Vernes, L., Rappaport, F., Wollman, F.-A., Alric, J., and Johnson, X. (2011). Plastid terminal oxidase 2 (PTOX2) is the major oxidase involved in chlororespiration in *Chlamydomonas*. *Proc. Natl. Acad. Sci. U. S. A.* 108: 20820–5.
- Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., and Darzins, A. (2008). Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *Plant J.* 54: 621–39.
- Iijima, Y. et al. (2008). Metabolite annotations based on the integration of mass spectral information. *Plant J.* 54: 949–62.
- Ikedo, Y., Yamagishi, A., Komura, M., Suzuki, T., Dohmae, N., Shibata, Y., Itoh, S., Koike, H., and Satoh, K. (2013). Two types of fucoxanthin-chlorophyll-binding proteins I tightly bound to the photosystem I core complex in marine centric diatoms. *Biochim. Biophys. Acta* 1827: 529–39.

- Jia, J., Han, D., Gerken, H.G., Li, Y., Sommerfeld, M., Hu, Q., and Xu, J. (2015). Molecular mechanisms for photosynthetic carbon partitioning into storage neutral lipids in *Nannochloropsis oceanica* under nitrogen-depletion conditions. *Algal Res.* 7: 66–77.
- Johnson, X. and Alric, J. (2013). Central carbon metabolism and electron transport in *Chlamydomonas reinhardtii*: metabolic constraints for carbon partitioning between oil and starch. *Eukaryot. Cell* 12: 776–93.
- Lemoine, Y. and Schoefs, B. (2010). Secondary ketocarotenoid astaxanthin biosynthesis in algae: a multifunctional response to stress. *Photosynth. Res.* 106: 155–77.
- Lepetit, B., Sturm, S., Rogato, A., Gruber, A., Sachse, M., Falciatore, A., Kroth, P.G., and Lavaud, J. (2013). High light acclimation in the secondary plastids containing diatom *Phaeodactylum tricornutum* is triggered by the redox state of the plastoquinone pool. *Plant Physiol.* 161: 853–65.
- Lepetit, B., Volke, D., Szabó, M., Hoffmann, R., Garab, G., Wilhelm, C., and Goss, R. (2007). Spectroscopic and molecular characterization of the oligomeric antenna of the diatom *Phaeodactylum tricornutum*. *Biochemistry* 46: 9813–22.
- Li, J., Han, D., Wang, D., Ning, K., Jia, J., Wei, L., Jing, X., Huang, S., Chen, J., Li, Y., Hu, Q., and Xu, J. (2014). Choreography of Transcriptomes and Lipidomes of *Nannochloropsis* Reveals the Mechanisms of Oil Synthesis in Microalgae. *Plant Cell* 26: 1645–1665.
- Li, Z., Wakao, S., Fischer, B.B., and Niyogi, K.K. (2009). Sensing and responding to excess light. *Annu. Rev. Plant Biol.* 60: 239–60.
- Liu, M., Diretto, G., Pirrello, J., Roustan, J.-P., Li, Z., Giuliano, G., Regad, F., and Bouzayen, M. (2014). The chimeric repressor version of an Ethylene Response Factor (ERF) family member, SI-ERF.B3, shows contrasting effects on tomato fruit ripening. *New Phytol.* 203: 206–18.
- Lu, Y., Tarkowská, D., Turečková, V., Luo, T., Xin, Y., Li, J., Wang, Q., Jiao, N., Strnad, M., and Xu, J. (2014a). Antagonistic roles of abscisic acid and cytokinin during response to nitrogen depletion in oleaginous microalga *Nannochloropsis oceanica* expand the evolutionary breadth of phytohormone function. *Plant J.* 80: 52–68.
- Lu, Y., Zhou, W., Wei, L., Li, J., Jia, J., Li, F., Smith, S.M., and Xu, J. (2014b). Regulation of the cholesterol biosynthetic pathway and its integration with fatty acid biosynthesis in the oleaginous microalga *Nannochloropsis oceanica*. *Biotechnol. Biofuels* 7: 81.
- Lunn, J.E., Delorge, I., Figueroa, C.M., Van Dijck, P., and Stitt, M. (2014). Trehalose metabolism in plants. *Plant J.* 79: 544–67.
- Martin, G.J.O., Hill, D.R.A., Olmstead, I.L.D., Bergamin, A., Shears, M.J., Dias, D.A., Kentish, S.E., Scales, P.J., Botté, C.Y., and Callahan, D.L. (2014). Lipid Profile Remodeling in Response to

- Nitrogen Deprivation in the Microalgae *Chlorella* sp. (Trebouxiophyceae) and *Nannochloropsis* sp. (Eustigmatophyceae). *PLoS One* 9: e103389.
- De Martino, A., Douady, D., Quinet-Szely, M., Rousseau, B., Crépineau, F., Apt, K., and Caron, L. (2000). The light-harvesting antenna of brown algae: highly homologous proteins encoded by a multigene family. *Eur. J. Biochem.* 267: 5540–9.
- Maxwell, D.P., Falk, S., Trick, C.G., and Huner, N. (1994). Growth at Low Temperature Mimics High-Light Acclimation in *Chlorella vulgaris*. *Plant Physiol.* 105: 535–543.
- Miller, R. et al. (2010). Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. *Plant Physiol.* 154: 1737–52.
- Mullin, K.A., Lim, L., Ralph, S.A., Spurck, T.P., Handman, E., and McFadden, G.I. (2006). Membrane transporters in the relict plastid of malaria parasites. *Proc. Natl. Acad. Sci. U. S. A.* 103: 9572–7.
- Nowicka, B. and Kruk, J. (2010). Occurrence, biosynthesis and function of isoprenoid quinones. *Biochim. Biophys. Acta* 1797: 1587–605.
- Pal, D., Khozin-Goldberg, I., Cohen, Z., and Boussiba, S. (2011). The effect of light, salinity, and nitrogen availability on lipid production by *Nannochloropsis* sp. *Appl. Microbiol. Biotechnol.* 90: 1429–41.
- Peers, G., Truong, T.B., Ostendorf, E., Busch, A., Elrad, D., Grossman, A.R., Hippler, M., and Niyogi, K.K. (2009). An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* 462: 518–521.
- Pego, J. V., Kortstee, A.J., Huijser, C., and Smeekens, S.C. (2000). Photosynthesis, sugars and the regulation of gene expression. *J. Exp. Bot.* 51 Spec No: 407–16.
- Petroutsos, D. et al. (2014). Evolution of galactoglycerolipid biosynthetic pathways--from cyanobacteria to primary plastids and from primary to secondary plastids. *Prog. Lipid Res.* 54: 68–85.
- Poliner, E., Panchy, N., Newton, L., Wu, G., Lapinsky, A., Bullard, B., Zienkiewicz, A., Benning, C., Shiu, S.-H., and Farré, E.M. (2015). Transcriptional coordination of physiological responses in *Nannochloropsis oceanica* CCMP1779 under light/dark cycles. *Plant J.*
- Radakovits, R., Jinkerson, R.E., Darzins, A., and Posewitz, M.C. (2010). Genetic engineering of algae for enhanced biofuel production. *Eukaryot. Cell* 9: 486–501.
- Radakovits, R., Jinkerson, R.E., Fuerstenberg, S.I., Tae, H., Settledge, R.E., Boore, J.L., and Posewitz, M.C. (2012). Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. *Nat. Commun.* 3: 686.
- Ramakrishnan, S., Serricchio, M., Striepen, B., and Bütikofer, P. (2013). Lipid synthesis in protozoan

- parasites: a comparison between kinetoplastids and apicomplexans. *Prog. Lipid Res.* 52: 488–512.
- Richard, C., Ouellet, H., and Guertin, M. (2000). Characterization of the LI818 polypeptide from the green unicellular alga *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* 42: 303–16.
- Risso, D., Schwartz, K., Sherlock, G., and Dudoit, S. (2011). GC-content normalization for RNA-Seq data. *BMC Bioinformatics* 12: 480.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–40.
- Rodolfi, L., Chini Zittelli, G., Bassi, N., Padovani, G., Biondi, N., Bonini, G., and Tre dicit, M.R. (2009). Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnol. Bioeng.* 102: 100–12.
- Rolland, F., Baena-Gonzalez, E., and Sheen, J. (2006). Sugar sensing and signaling in plants: conserved and novel mechanisms. *Annu. Rev. Plant Biol.* 57: 675–709.
- Ryu, J.-Y., Song, J.Y., Lee, J.M., Jeong, S.W., Chow, W.S., Choi, S.-B., Pogson, B.J., and Park, Y.-I. (2004). Glucose-induced expression of carotenoid biosynthesis genes in the dark is mediated by cytosolic pH in the cyanobacterium *Synechocystis* sp. PCC 6803. *J. Biol. Chem.* 279: 25320–5.
- Saeed, A.I. et al. (2003). TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* 34: 374–8.
- Sato, S., Tabata, S., Hirakawa, H., Asamizu, E., Shirasawa, K., Isobe, S., Kaneko, T., Nakamura, Y., Shibata, D., Aoki, K., and et al (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485: 635–41.
- Sforza, E., Simionato, D., Giacometti, G.M., Bertucco, A., and Morosinotto, T. (2012). Adjusted light and dark cycles can optimize photosynthetic efficiency in algae growing in photobioreactors. *PLoS One* 7: e38975.
- Shoji, N., Nakagawa, K., Asai, A., Fujita, I., Hashiura, A., Nakajima, Y., Oikawa, S., and Miyazawa, T. (2010). LC-MS/MS analysis of carboxymethylated and carboxyethylated phosphatidylethanolamines in human erythrocytes and blood plasma. *J. Lipid Res.* 51: 2445–53.
- Simionato, D., Block, M.A., La Rocca, N., Jouhet, J., Maréchal, E., Finazzi, G., and Morosinotto, T. (2013). The response of *Nannochloropsis gaditana* to nitrogen starvation includes de novo biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. *Eukaryot. Cell* 12: 665–76.
- Simionato, D., Sforza, E., Corteggiani Carpinelli, E., Bertucco, A., Giacometti, G.M., and

- Morosinotto, T. (2011). Acclimation of *Nannochloropsis gaditana* to different illumination regimes: effects on lipids accumulation. *Bioresour. Technol.* 102: 6026–6032.
- Stitt, M. and Zeeman, S.C. (2012). Starch turnover: pathways, regulation and role in growth. *Curr. Opin. Plant Biol.* 15: 282–92.
- Stoop, J., Williamson, J., and Masonpharr, D. (1996). Mannitol metabolism in plants: a method for coping with stress. *Trends Plant Sci.* 1: 139–144.
- Su, L., Diretto, G., Purgatto, E., Danoun, S., Zouine, M., Li, Z., Roustan, J.-P., Bouzayen, M., Giuliano, G., and Chervin, C. (2015). Carotenoid accumulation during tomato fruit ripening is modulated by the auxin-ethylene balance. *BMC Plant Biol.* 15: 114.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30: 2725–9.
- Tanaka, T. et al. (2015). Oil accumulation by the oleaginous diatom *Fistulifera solaris* as revealed by the genome and transcriptome. *Plant Cell* 27: 162–76.
- Teo, C.L., Idris, A., Zain, N.A.M., and Taisir, M. (2014). Synergistic effect of optimizing light-emitting diode illumination quality and intensity to manipulate composition of fatty acid methyl esters from *Nannochloropsis* sp. *Bioresour. Technol.* 173: 284–90.
- Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U. S. A.* 98: 5116–21.
- Vieler, A. et al. (2012). Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779. *PLoS Genet.* 8: e1003064.
- Vítová, M., Bišová, K., Kawano, S., and Zachleder, V. (2015). Accumulation of energy reserves in algae: From cell cycles to biotechnological applications. *Biotechnol. Adv.*
- De Vos, R.C.H., Moco, S., Lommen, A., Keurentjes, J.J.B., Bino, R.J., and Hall, R.D. (2007). Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 2: 778–91.
- Wang, D. et al. (2014). *Nannochloropsis* genomes reveal evolution of microalgal oleaginous traits. *PLoS Genet.* 10: e1004094.
- Wang, J., Sommerfeld, M., and Hu, Q. (2009). Occurrence and environmental stress responses of two plastid terminal oxidases in *Haematococcus pluvialis* (Chlorophyceae). *Planta* 230: 191–203.
- Wang, Y., Duanmu, D., and Spalding, M.H. (2011). Carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*: inorganic carbon transport and CO<sub>2</sub> recapture. *Photosynth. Res.* 109: 115–22.
- Wase, N., Black, P.N., Stanley, B.A., and DiRusso, C.C. (2014). Integrated quantitative analysis of

- nitrogen stress response in *Chlamydomonas reinhardtii* using metabolite and protein profiling. *J. Proteome Res.* 13: 1373–96.
- Willamme, R., Alsafr, Z., Arumugam, R., Eppe, G., Remacle, F., Levine, R.D., and Remacle, C. (2015). Metabolomic analysis of the green microalga *Chlamydomonas reinhardtii* cultivated under day/night conditions. *J. Biotechnol.* 215: 20–6.
- Zaffagnini, M., Michelet, L., Sciabolini, C., Di Giacinto, N., Morisse, S., Marchand, C.H., Trost, P., Fermani, S., and Lemaire, S.D. (2014). High-resolution crystal structure and redox properties of chloroplastic triosephosphate isomerase from *Chlamydomonas reinhardtii*. *Mol. Plant* 7: 101–20.
- Zhao, B. and Su, Y. (2014). Process effect of microalgal-carbon dioxide fixation and biomass production: A review. *Renew. Sustain. Energy Rev.* 31: 121–132.
- Zhu, S.-H. and Green, B.R. (2010). Photoprotection in the diatom *Thalassiosira pseudonana*: role of LI818-like proteins in response to high light stress. *Biochim. Biophys. Acta* 1797: 1449–57.

# CHAPTER VI

## Investigation of Transcription Regulatory Sequences in *Nannochloropsis gaditana*

### Author names and affiliations

Giorgio Perin<sup>1</sup>, Mattia Storti<sup>1</sup>, Alessandro Alboresi<sup>1</sup> and  
Tomas Morosinotto<sup>1</sup>

<sup>1</sup>PAR-Lab\_Padua Algae Research Laboratory, Department of Biology,  
University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy



**Abstract.**

Heterologous expression is a seminal tool for industrial production of proteins. The development of molecular tools for protein expression in new, unexplored organisms, remains a high priority in biotechnology research to avoid present limitations and express recalcitrant proteins. Recently microalgae have attracted large attention for recombinant protein production. Their added value is sustainability since photosynthetic microorganisms use only sunlight, water, CO<sub>2</sub> and mineral nutrients to support their metabolism.

Among different algal species, *Chlamydomonas reinhardtii* is the only one extensively tested for the expression of recombinant proteins. Recently also microalgae species of the genus *Nannochloropsis* emerged as candidates for industrial applications, especially for the production of biofuels. To explore the possibility of other industrial uses for this species, we tested the *Nannochloropsis* potential as microalgal host for protein production. This however requires the molecular toolbox for its genetic manipulation to be highly enriched and in particular there is the need of identifying strong and inducible promoters. Here we present the development of a system for efficient protein expression in *Nannochloropsis gaditana*. It showed that *Nannochloropsis* is indeed capable of accumulating heterologous proteins to a much higher level than the reference species *C. reinhardtii*. The potentiality to use it as molecular tool to study transcription regulatory sequences in *Nannochloropsis gaditana* is here discussed.

## Introduction.

Different organisms are currently being tested and exploited as platforms for heterologous protein expressions for research and biotechnological applications. Each one of them presents specific strengths and limitations in terms of protein activity, yield and production cost (Rasala and Mayfield, 2015; Rosales-Mendoza et al., 2012)(Cardi et al., 2014; Ullrich et al., 2015).

Unicellular eukaryotic organisms are among the best models for the development of expression systems, suitable for protein production at industrial scale, because of their contained cultivation. Eukaryotic microorganisms are safe, often accessible to genetic modifications through transformation, mutagenesis and breeding techniques. Furthermore they combine advantages of eukaryotic expression system, such as the ability to correctly fold complex proteins and perform post-translational modifications, with the high productivity of prokaryotes. The most characterized and worldwide used members of this group are the yeasts *Saccharomyces cerevisiae* and *Pichia pastoris*. However these microorganisms tend to hyper-glycosylate recombinant proteins, affecting their correct folding and causing immunogenic effects on human beings (Mathieu-Rivet et al., 2014). Moreover their growth media, although cheaper than those for mammalian cells, still represent a substantial cost at the industrial scale (e.g. the need for expensive carbon sources to support their heterotrophic metabolism).

Despite the available alternatives, cases where a heterologous protein does not reach the market, because of an inefficient or expensive expression host, are still very common. This makes the development of new, unexplored, organisms as expression hosts a high priority in the biotechnology research, to find new strategies to produce large amounts of active proteins using a safe, cheap and environmentally sustainable process.

Recently other eukaryotic microorganisms like microalgae have attracted large attention as next generation platforms, suitable for recombinant proteins production, but also for other applications from bioenergy to biopharmaceutical/nutraceutical for animal and human care (Rasala et al., 2014)(Mathieu-Rivet et al., 2014). Microalgae are water unicellular microorganisms that lack roots, leaves and other tissues characteristic of higher plants, even though they also perform photosynthesis. In fact, they use sunlight to support their metabolism and their growth requires only CO<sub>2</sub> (Salih, 2011) and mineral nutrients such as nitrates and phosphates, which are available at very low cost. They have even been shown to be capable of exploiting nutrients from wastewaters for growth, providing an environment-friendly high-value cultivation (Ramos Tercero et al., 2014).

Algae exploitation thus have the potential of becoming a platform for molecules production integrated in a more sustainable process than traditional fermentation platforms, from both the energetic and environmental point of view (Nielsen et al., 2013; Lassen et al., 2014).

Among microalgae, *Chlamydomonas reinhardtii* is the most characterized algae species and the almost exclusive microalga tested for the expression of recombinant proteins (Lauersen et al., 2015b, 2015a; Plucinak et al., 2015; Rasala et al., 2012)(Gangl et al., 2015). This choice, however, was not due to characteristics of the algae itself, but only to the availability of sophisticated tools for its genetic modification (Lauersen et al., 2015b)(Rasala et al., 2014). On the other side, in recent years other species like *Chlorella*, *Phaeodactylum* and *Nannochloropsis* emerged for their potential industrial applications and for their ability to grow large biomass in industrial scale photobioreactors (Niu et al., 2013; Hamilton et al., 2015; Vanier et al., 2015)(Peng et al., 2014)(Zhang et al., 2014). *Nannochloropsis* spp. (members of the class Eustigmatophyceae) have been intensively studied during the past few years as promising candidates for biofuels production, thanks to their capacity of accumulating a high amount of lipids. Currently a molecular toolbox for *Nannochloropsis* genetic manipulation and metabolic engineering is being developed (Corteggiani Carpinelli et al., 2014; Perin et al., 2015; Radakovits et al., 2012), but efficient expression vectors provided with strong and inducible promoters are still not available. This limits *Nannochloropsis* exploitation as heterologous proteins expression host. To fill this gap, in this work we present the development of an efficient expression system for *N. gaditana*, based on the endogenous LDSP promoter driving the expression the *Renilla reniformis* luciferase gene. The work will end with the preliminary identification and cloning of 5' - UTR molecular regions of highly expressed / light-regulated genes, to assess, in the near future, their applicability as strong / inducible promoter regions.

## Materials and Methods.

### *Nannochloropsis gaditana* growth conditions

*Nannochloropsis gaditana*, strain 849/5, was obtained from the Culture Collection of Algae and Protozoa (CCAP). It was maintained in sterile F/2 medium plates, containing sea salts (32 g/L, Sigma Aldrich), 40 mM Tris-HCl (pH 8), Guillard's (F/2) marine water enrichment solution (Sigma Aldrich) and 1 % plant agar (Duchefa Biochemie), at 100  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$  illumination and  $22 \pm 1$  °C. Transformants are maintained in the same growth medium, supplied with 3.5  $\mu\text{g/ml}$  zeocin. For transformation experiments and bioluminescence assays, the strains were cultivated in liquid F/2 medium in Erlenmeyer flasks with 100 rpm agitation, in the same illumination and temperature conditions. Algal growth was measured counting the cells using a cell counter (Cellometer Auto X4, Nexcelom Bioscience).

### *Preparation of the transformation vectors*

**Genomic DNA extraction.** Up to -1000 bp from the translation-start site upstream regions of several genes were cloned as promoter regions. *N. gaditana* genomic DNA was used as template for the amplification of the 5'-UTR regions of different endogenous genes. Genomic DNA was extracted from 4-day-old *N. gaditana* cultures. Cells were lysed using a Mini Bead Beater (Biospec Products) at 3500 RPM for 20 s in the presence of glass beads (150–212  $\mu\text{m}$  diameter). Genomic DNA was then purified using the *EUROGOLD™ Plant DNA Mini Kit* (Euroclone), applying minor modifications.

**Cloning.** A cassette conferring resistance to the antibiotic zeocin (approximately 1.3 kb) was amplified (primers #1 and #2 of table 1), with KAPA HIFI high-fidelity DNA polymerase (KAPA Biosystems), from the pPha-T1-UEP vector, kindly provided by Prof. Matthew Posewitz (Radakovits et al., 2012).

It contains the *Sh-ble* gene (from *Streptoalloteichus hindustanus*) with an endogenous UBIQUITIN (Naga\_100009g87) promoter and the FCPA terminator. The promoter region of the Lipid Droplet Surface Protein gene (LDSP, Gene ID - Naga\_100086g4) was amplified from *N. gaditana* genomic DNA and fused to the coding sequence of the *Crluc* gene (Fuhrmann et al., 2004; Ferrante et al., 2008) and to the FCPA terminator, exploiting the PCR-driven overlap extension technique (Heckman and Pease, 2007) (primers #3, #4, #7, #8, #9 and #10 of table 1). The same technique was applied to clone the Ring-U-BOX domain containing protein (RUBOX, Gene ID - Naga\_100094g14) 5'-UTR region and the deleted regions of the LDSP promoter ( $\Delta 1$ ,  $\Delta 2$  e  $\Delta 3$ ) (primers #5, #6, #11, #12 and #13 of table 1).

Both cassettes were then cloned in series (Figure 1A) in a pBlueScript SKII (+) backbone vector, harboring ampicillin resistance and *E. coli* origin of replication. Promoters used to drive the *Crluc* expression were then substituted with the Transketolase (Tkl, Gene ID - Naga\_100001g128) promoter (primers #14 and #15 of table 1), through standard PCR and cloning techniques. Standard cloning was also used to insert endogenous antenna proteins promoters (LHCf1, Gene ID: 100012g50; LHCX1, Gene ID: 100173g12; LHCX2, Gene ID: 100056g42; gene Naga\_101036g3) to drive *Crluc* expression (primers #16, #17, #18, #19, #20, #21, #22 and #23 of table 1). Amplified sequences accuracy was assessed through DNA-seq.

**Table 1. Primer sequences used to build the transformation vectors.**

Primer #	Name	Primer Sequence (5' – 3')
1	UEP_FOR	GAGGCCTGAATTCAGCTGCTGCCCGACCGTATCTCCAAGTCAGACATGAA
2	UEP_REV	CTGCAACATATGACTCACTATAGGGAGACCGGCTCGAGAAAACATCATCCTG
3	LDSP_1for	ATCATATGCGTCCTCAAGACAATACA
4	LDSP_ <i>Crluc</i> 2rev	GGCTGCAGGAATTCGATATCCTTATCTTGAGATAGAGT
5	RUBOX_1for	ATCATATGGTATTAGTTTTCACTCCT
6	RUBOX_ <i>Crluc</i> 2rev	GGCTGCAGGAATTCGATATCCTTGTCAAAGCACGTGCT
7	<i>Crluc</i> _3for	GATATCGAATTCCTGCAGCC
8	<i>Crluc</i> _fcpA4rev	TAAATTTTTAAGGAAGGTAGATCTTTACGTATCGTT
9	<i>Crluc</i> _fcpA5for	AACGATACGTAAAGATCTACCTTCCTTAAAAATTTA
10	fcpA_6rev	ATCCGCGGAAAACATCCTGTGCCT
11	LDSP Δ3	TATCATATGATAGAATCTGATGCCACG
12	LDSP Δ2	TATCATATGCGTCTTTTTGACGGCTAT
13	LDSP Δ1	TATCATATGTGTGACGAAGCAATGTG
14	Tkl_NdeI	ATACATATGCGCCATGGAATTCATAGTCAGCT
15	Tkl_PstI	ATACTGCAGCTTCACTGTAGTTGTAAAGAGTGGAGT
16	LHCX1_fcpA	GATGAGTTTTCCGCGGATAAAGGCGGCAGAAAGATCG
17	LHCX1_BamHI	TAT GGATCCGTTCTGGGGGTGGGATGGAA
18	LHCf1_NdeI	TATCATATGATCCGAAAAGGCGGAGA
19	LHCf1_PstI	TATCTGCAGTATGAGAAAAGTTTTAAAGTGGAG
20	LHCX2_NdeI	TATCATATGTTTGACGCCGCTTCAGGCATG
21	LHCX2_BamHI	TATGGATCCGCTTCAGCTTTACAATGCGTGAT
22	Naga101036g3_NdeI	TATCATATGCTTCGTACGCATCATCTTTC
23	Naga101036g3_BamHI	TATGGATCCAAGTAGAAAAAGAAGGGAC

#### *Transformation of Nannochloropsis gaditana*

The DNA cassettes (approximately between 3 and 3.5 kb) were digested from the transformation vectors and the raw digested was used for *N. gaditana* transformation.  $5 \times 10^8$  *N. gaditana* cells were washed four times with 375 mM sorbitol at 4 °C and resuspended in 100 μl 375 mM cold sorbitol.

Cells were then incubated with 5 µg DNA on ice and then electroporated in a 2-mm cuvette using an ECM630 BTX electroporator set (500 Ω, 50 µF and 2400 V). Cells were thus recovered in the light for 24 h and plated onto F/2 plates containing 3.5 µg/ml zeocin. Resistant colonies were detected after 3 weeks and picked after 4–5 weeks. Transformants were then screened for the presence of the *Crluc* gene, exploiting the colony PCR protocol of (Cao et al., 2009), using templates obtained with the Chelex-100 method. Primers used for screening are #7 and #8 of table 1.

#### *Luciferase activity evaluation*

To assay luciferase activity of *N. gaditana* transformants we followed the protocols of (Shao and Bock, 2008) and (Lauersen et al., 2013), applying minor revisions, and using both a luminometer TD-20/20 or a microplates reader. For *N. gaditana* strains, protocols are here briefly described. For *C. reinhardtii* strains, we followed strictly the references protocols.

*Luciferase activity quantification.* Coelenterazine (ThermoFisher Scientific) was dissolved in methanol (0.05 mM stock solution). *N. gaditana* transformants were grown in liquid F/2 medium as previously described. After sampling, cells were lysed using a Mini Bead Beater (Biospec Products) at 3500 RPM for 20 s in the presence of glass beads (150–212 µm diameter) and 100 µl of sample buffer [1.5 mM Tris–HCl (pH 8.0), 1 mM EDTA]. After lysis, 125 µl of assay buffer [0.1 M K<sub>2</sub>HPO<sub>4</sub> (pH 7.0), 0.5 M NaCl, 1 mM EDTA] were added to each sample. Bioluminescence was thus assayed using a luminometer TD-20/20 (Turner BioSystems) after the manual injection of the substrate (coelenterazine 0.05 mM; 50 µl). Parameters set: delay, 3 sec; Integration, 60 sec; Replicate, 1. Luminescence was expressed as relative luminescence units (RLU). Luciferase activity was normalized to the total volume of cells in the sample (volume was calculated from the cell diameter obtained from the Cellometer Auto X4 cell counter).

*Luciferase activity screening.* *N. gaditana* transformants were grown in solid F/2 medium as previously described. Colonies were transferred in 100 µl of sample buffer in 96-well white microtiter plates. Bioluminescence was assayed using the EnVision multilabel Reader (Perkin Elmer), after 125 µl assay buffer addition and the manual injection of the substrate (coelenterazine 0.05 mM; 50 µl per well). Luciferase activity was normalized to the optical density at 540 nm, measured per each well after the assay buffer addition.

## Results and Discussion.

Here we describe the development of a novel vector system in order to bring out *Nannochloropsis* potentiality as host for proteins expression. It was designed to be a modular and expandable molecular tool, to be generally and intensively used with *Nannochloropsis*, for heterologous proteins expression and regulatory sequences studies. Once tested, it will be further expanded by using different combination of regulatory sequences, taking advantage of tools for gene regulatory sequences combination, like Gibson assembly.

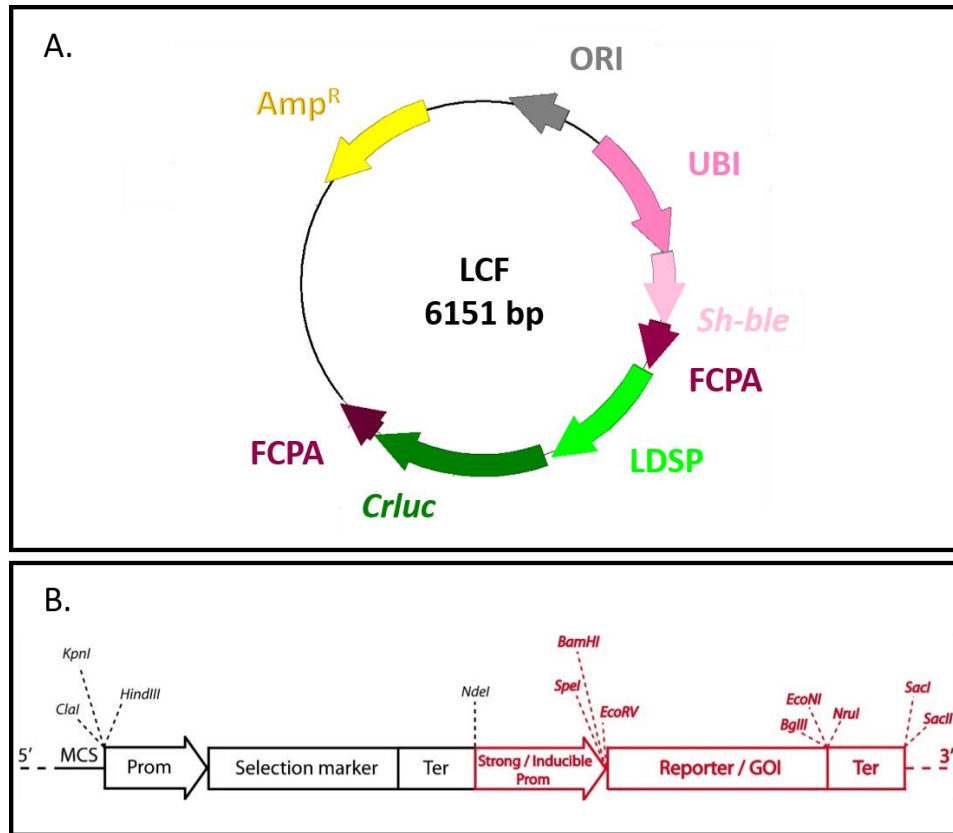
### *Vector design.*

In order to develop a standardized and modular expression vector, the latter was designed to carry a selection cassette together with an over-expression cassette. The former, necessary to select transformants, harbors the *Sh-ble* gene conferring resistance to the antibiotic zeocin, whose transcription is regulated by the endogenous UBIQUITIN promoter (UBI), together with the FCPA terminator (Radakovits et al., 2012). The overexpression cassette was instead designed to carry the *Renilla reniformis* luciferase gene (*Rluc*), whose expression was driven by a strong endogenous promoter and regulated by the FCPA terminator. Both cassettes were cloned sequentially using the pBlueScript SKII (+) as vector backbone, harboring an origin of replication for *E. coli* and the ampicillin resistance gene (Figure 1A and B).

Bioluminescent proteins are widely used as reporter genes to identify new transcriptional regulatory elements and often measure promoters strength. Luciferases are indeed bioluminescent nontoxic reporters, suitable to monitor gene expression quantitatively. Since being small proteins (37 kDa - (Loening et al., 2007)) they do not interfere much with the cellular metabolism and therefore are useful to compare different organisms capacity to be proteins expression hosts. The *Rluc* gene was codon optimized for *C. reinhardtii* nuclear genome (Fuhrmann et al., 2004), we will refer to it as *Crluc* gene (*C. reinhardtii* luciferase)). This gene was chosen to assess its functionality in *Nannochloropsis*, opening the possibility to move some of the molecular tools developed for *C. reinhardtii* to this organism, without modifications. To drive its expression we needed a strong promoter to be sure to accumulate the luciferase at high titer.

The Lipid Droplet Surface Protein (LDSP) gene promoter was described to be able to induce the efficient over-expression of endogenous proteins in *N. oceanica* (e.g. an endogenous  $\Delta$ 12-desaturase (Kaye et al., 2015)). In RNA-seq experiments, the corresponding gene in *N. gaditana* (LDSP, Gene ID -

Naga\_100086g4) was shown to be highly expressed (e.g. in medium light conditions, LDSP gene was 4 times more expressed than the UBIQUITIN gene, Figure 2A), supporting its suitability as a promoter to drive a strong gene expression.

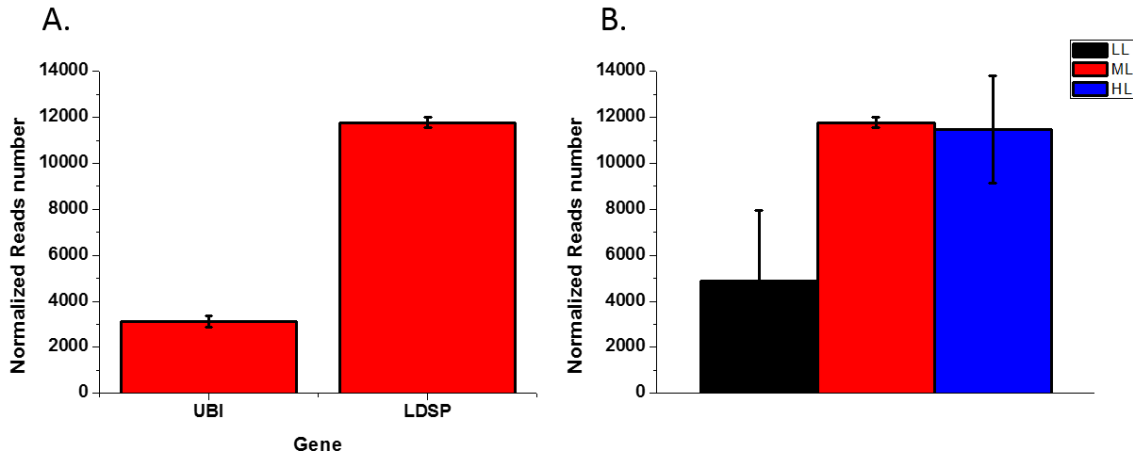


**Figure 1. Diagram of the over-expression vector (LCF – LDSP, *Crluc*, FCPA) for *Nannochloropsis gaditana*.** A. Vector designed for the proteins over-expression in *Nannochloropsis gaditana*. The backbone relies on the pBlueScript II SK (+) vector. The vector molecular components are depicted in different colors: yellow, Amp<sup>R</sup>; grey, *E. coli* ORI; pink, endogenous UBIQUITIN promoter (UBI); light-pink, *Sh-ble* gene conferring resistance to Zeocin; violet, FCPA terminator; light-green, LDSP promoter; green, *Crluc* gene. B. Schematic overview of the cassettes used for transformants selection (in black) and proteins over-expression (in red). Each element of the over-expression cassette was designed to be surrounded by unique restriction sites. This feature allowed the simple opening of the vector for investigating different promoter regions and for the future replacement of the coding sequence of target proteins. Prom, Promoter; Ter, Terminator; GOI, Gene Of Interest.

For this reason the 5'-UTR region of the LDSP gene was the first choice as regulatory region to drive the expression of the *Crluc* gene. In this case, the molecular sequence considered as a putative promoter was taken starting from 677 bp upstream of the corresponding gene, to avoid the coding

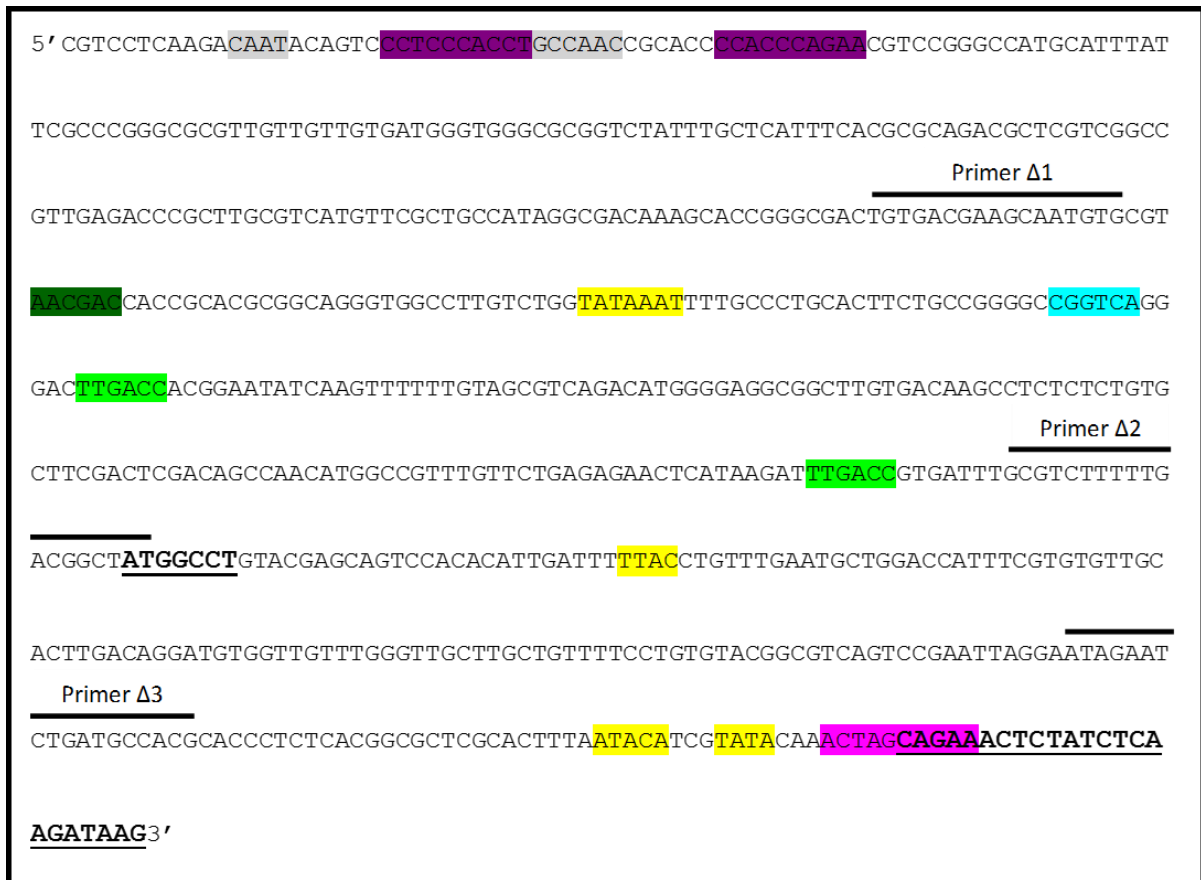


sequence of another gene (Gene ID - Naga\_100086g5), present in the negative strand (genome browser - [www.nannochloropsis.org](http://www.nannochloropsis.org)).



**Figure 2. Normalized reads number coming from the RNA-seq experiment of Chapter V.** A. Normalized reads number for the UBIQUITIN (UBI) and LIPID DROPLET SURFACE PROTEIN (LDSP) genes, coming from samples acclimated to medium light conditions ( $100 \mu\text{moles photons} / (\text{m}^2 \text{s})$ ). B. Normalized reads number for the LIPID DROPLET SURFACE PROTEIN gene, coming from samples acclimated to 3 light intensities, low, medium and high light conditions ( $10, 100$  and  $1000 \mu\text{moles photons} / (\text{m}^2 \text{s})$ , respectively). Data come from the average  $\pm$  SD of 3 biological replicates.

It is worth mentioning that several 5'-UTR regions in *N. gaditana* are expected to include bi-directional promoters, given the compactness of the genome (Corteggiani Carpinelli et al., 2014) and this possibility should be considered when evaluating promoters activities (Kilian et al., 2011). It is also worth noting that the cloned LDSP promoter region indeed includes the first exon, the first intron and also the first 23 bp of the second exon (Figure 3) of the LDSP gene. In fact when the LDSP promoter was previously cloned without these molecular elements (shorter version of the LDSP promoter, including thus only the 5'-UTR and accounting for 454 bp of length) to drive the expression of the *Sh-ble* gene in place of the UBIQUITIN promoter, after electroporation we never get any resistant microalgae strain. On the contrary, resistant transformants were indeed obtained from the cloning of the longer LDSP promoter version (data not shown), thus including the first intron and part of the coding sequence of the gene, as previously here described. As shown in the figure 3, the first LDSP version (the shortest) indeed misses 3 molecular elements, recognized as TATA-boxes by the PlantCARE software (Rombauts et al., 1999), necessary for the RNA-polymerase recruitment. This could be responsible for the absence of transcriptional activity, shown for the shortest LDSP promoter version.

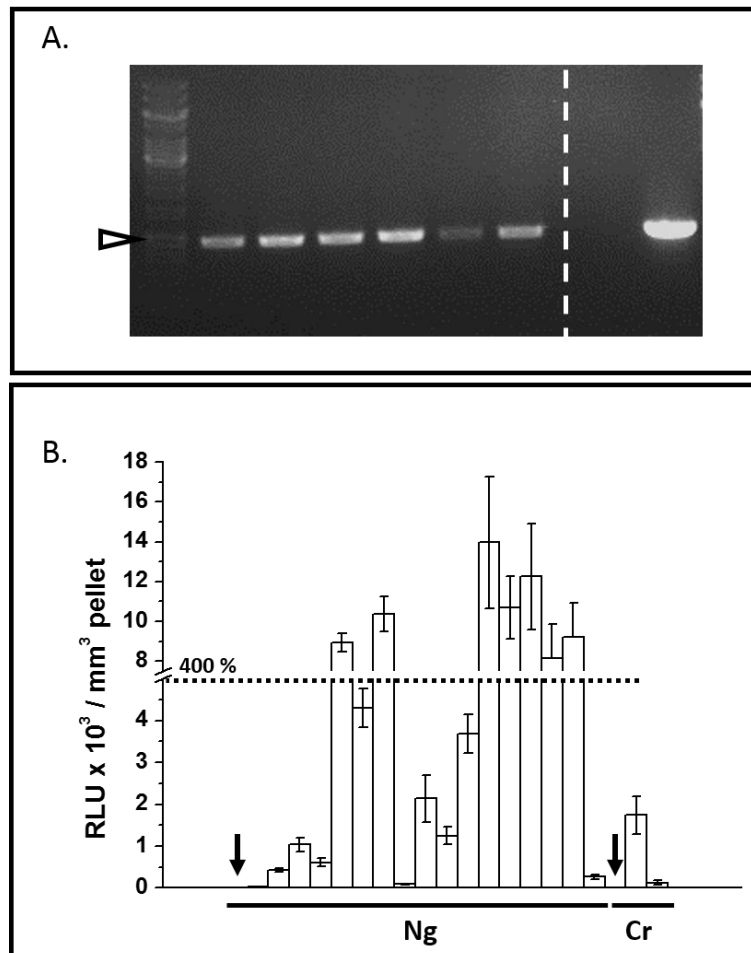


**Figure 3. Lipid Droplet Surface Protein (LDSP – Gene ID: Naga\_100086g4) promoter sequence.** The sequence cloned as promoter of the LDSP gene is here depicted. The sequence is 677 bp long. Nucleotides belonging to the first exon and the first 23 bases of the second one are underlined, according to the annotation reported in genome browser ([www.nannochloropsis.org](http://www.nannochloropsis.org)) (Corteggiani Carpinelli et al., 2014). The primers used to obtain the promoter deletions are depicted above the corresponding sequence; primer Δ1, promoter 474 bp long; primer Δ2, promoter 239 bp long, primer Δ3, promoter 87 bp long. Nucleotides belonging to cis-acting regulatory elements binding sites were identified running the PlantCARE software (Rombauts et al., 1999) (see Methods for details) and are highlighted in different colors: yellow, TATA-box; grey, CAAT-box; fuchsia, CTAG motif; light-blue, MYB binding site; green, W-box; violet, light-responsive elements.

To achieve the desired vector modularity, the regulatory elements chosen for the over-expression cassette were surrounded by unique restriction sites (Figure 1B), in order to facilitate their removal for further studies of regulatory sequences as well as reporter genes. The over-expression vector thus designed was named LCF (LDSP – *Crluc* - FCPA), from the names of the molecular components of the over-expression cassette (Figure 1A).

*Isolation of N. gaditana transformants expressing active luciferase.*

To determine the efficacy of the developed vector in driving heterologous proteins expression in *Nannochloropsis gaditana*, we transformed its nuclear genome with LCF through electroporation (see Methods for details). We verified by PCR the integration of the *Crluc* gene in the genome of zeocin resistant transformants (Figure 4A), supporting the efficacy of the transformation, even if the exogenous DNA inserted was more than 3 kb long.



**Figure 4. *Nannochloropsis gaditana* transformation with the LCF vector.** A. *Nannochloropsis gaditana* transformants were tested, through Colony-PCR, for the successful integration of the *Crluc* gene (936 bp) in the genome. Six independent transformants (lanes 1 – 6) confirmed the presence of the *Crluc* gene in their genome with respect to WT strains (C - ). C -, WT *N. gaditana* strain electroporated with water; C +, plasmid DNA (LCF vector). The left and the right part of the picture come from different regions of the same agarose gel. The arrow indicates 1000 bp. B. Relative luminescence units (RLU) obtained per mm<sup>3</sup> of cell pellet. The values of the left series belong to *N. gaditana* (Ng) strains transformed with the LCF vector, while the values of the right series belong to *C. reinhardtii* (Cr) strains transformed with the vector reported in (Fuhrmann et al., 2004), carrying

the same *Crluc* gene. For both series, negative control values (WT strains transformed with water) are highlighted by an arrow. A dotted line refers to the 400 % of the RLU/mm<sup>3</sup> value of the best *C. reinhardtii* over-expressing strain. Data are expressed as averages and SD of 2 biological replicates.

Furthermore, when transformants were left for 20 generations in non-selective growth conditions, they still showed the expression of the resistance cassette after switching them to the selective medium, highlighting the stable integration of the transformation vector in the genome.

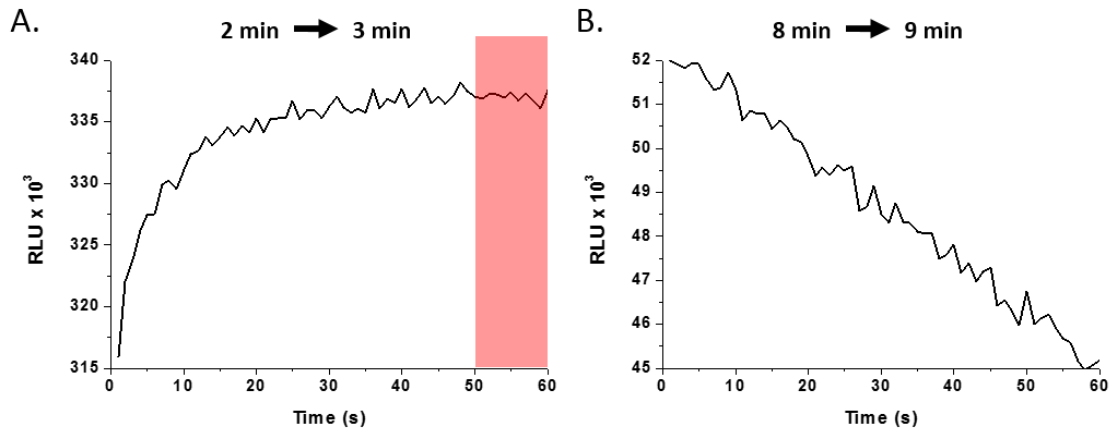
A selected number of resistant strains was further grown in liquid medium (see Methods for details) in order to assess luciferase activity, expressed as relative luminescence units (RLU), in comparison to the luciferase activity of *C. reinhardtii* nuclear transformants (Fuhrmann et al., 2004), used as reference expression platform for microalgae. The RLU values were normalized to the total cell volume of the processed samples (expressed in mm<sup>3</sup>), since *C. reinhardtii* cells are  $\approx$  50 times bigger than those of *N. gaditana* (Simionato et al., 2013; Harris, 2003).

No luminescence signal was detectable from parental strains of both species. Almost 50 % of the *N. gaditana* selected strains showed a luciferase activity 4 times higher than the *C. reinhardtii* luciferase expressing strains (Figure 4B).

#### *High-throughput screening for luciferase activity evaluation.*

As shown in the figure 4B, *N. gaditana* strains overexpressing the *Crluc* gene show a large variability in their luciferase activity. Several phenomena could explain this variability such as gene integration at genomic *loci* with different transcription level, or its integration in different copies, or the partial integration of the overexpression cassette. In order to expand the study to other regulatory sequences in *N. gaditana*, we thus needed the development of a reliable screening procedure to simply discriminate among hundreds of strains the best expressing ones (Lauersen et al., 2013). To quickly assess luciferase activity from multiple samples at a time, we developed a method based on a microplate reader (EnVision multilabel Reader, Perkin Elmer). The latter consists in directly placing the strains in a sample buffer in multiwell plates and, after the manual addition of the substrate, evaluate the luminescence signals (see Methods for details). To monitor the luciferase activity stability over time, the luminescence signals were tracked for 60 sec, after different times of incubation with the substrate. The signal reaches a plateau, 2 min and 50 sec after the addition of the substrate (Figure 5A), while strongly decreases after 8 minutes (Figure 5B). Therefore, to quickly and reliably monitor luminescence we decided to track the signal for 10 sec, 2 min and 50 sec after the addition of the substrate (red zone of the figure 5A). After having set up the method, to evaluate its reliability, we

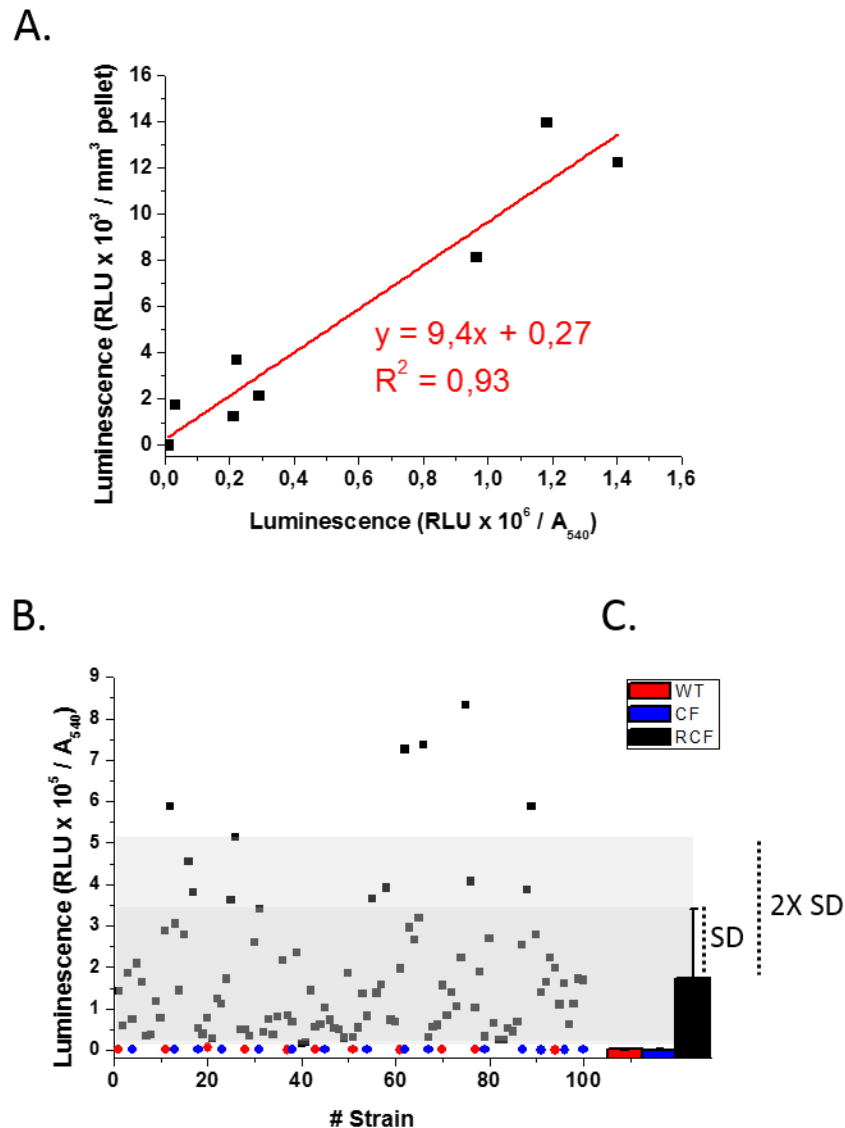
analyzed a selection of strains from the same pool presented in figure 4B (LCF transformed strains), that were previously measured with the luminometer (see Methods for details).



**Figure 5. Luminescence signal over time.** After addition of coelenterazine 0.05 mM the luminescence signal (relative luminescence units –RLU) was monitored over time. It was monitored after different times of incubation with the substrate; in A. luminescence was tracked for 60 sec after 2 minutes of coelenterazine incubation while in B. after 8 minutes. Luminescence signals come from different samples. The red zone in A highlights the time interval chosen to monitor luminescence signal.

Also the previous *C. reinhardtii* luciferase best expressing strain was evaluated with this method and parental strains too, as reference. The luciferase activity values obtained with this high-throughput method linearly correlate with those previously measured by the luminometer (Figure 6A). Both instruments identified the same strains as the top *Crluc* expressors (3 spots on the right in figure 6A). We can conclude that the high-throughput method is suitable for our purpose, significantly reducing the time required for best expressing strains identification (1 h vs. the 4 h needed using the luminometer to process 96 samples). Furthermore, almost 1 month passed between the first and the second evaluation of the luciferase activity for the LCF transformed strains. This means that CrLUC expression remained stable over time. This phenomenon highlights that *N. gaditana* is likely not affected by silencing phenomena, or at least, not as much as reported in literature for *C. reinhardtii* (Rasala et al., 2012, 2014). These considerations are true in the case of *Renilla reniformis* luciferase expression. It would be very interesting to investigate whether this peculiarity were true also for other proteins expression, opening the possibility for *Nannochloropsis gaditana* efficient exploitation as heterologous proteins expression host.

The developed method was thus tested with RCF transformants, obtained after electroporation with vectors where the *Crluc* gene expression was driven by the Ring-U-BOX domain containing protein (RUBOX, Gene ID - Naga\_100094g14) promoter (see methods for cloning details).



**Figure 6. High-throughput screening for luciferase activity.** A. Linear correlation between luminescence values obtained from the luminometer (y axis) and those obtained from the EnVision microplates reader (x axis). Both values were normalized on the amount of sample (volume of pellet in mm<sup>3</sup> in the first case and optical density at 540 nm in the second). WT values for *N. gaditana* and *C. reinhardtii* strains overlap in the same spot (bottom left of the series). Transformants with the LCF vector were used to validate the screening procedure. In B. and C., data obtained from the application of the developed screening are presented. Strains transformed with the RCF vector are depicted in

black; WT strains electroporated with water are in red; strains transformed with the CF vector, lacking the promoter region to drive *Crluc* gene transcription, are depicted in blu. In C. data are expressed as averages of 100, 11 and 15 independent strains for RCF, WT and CF transformants, respectively.

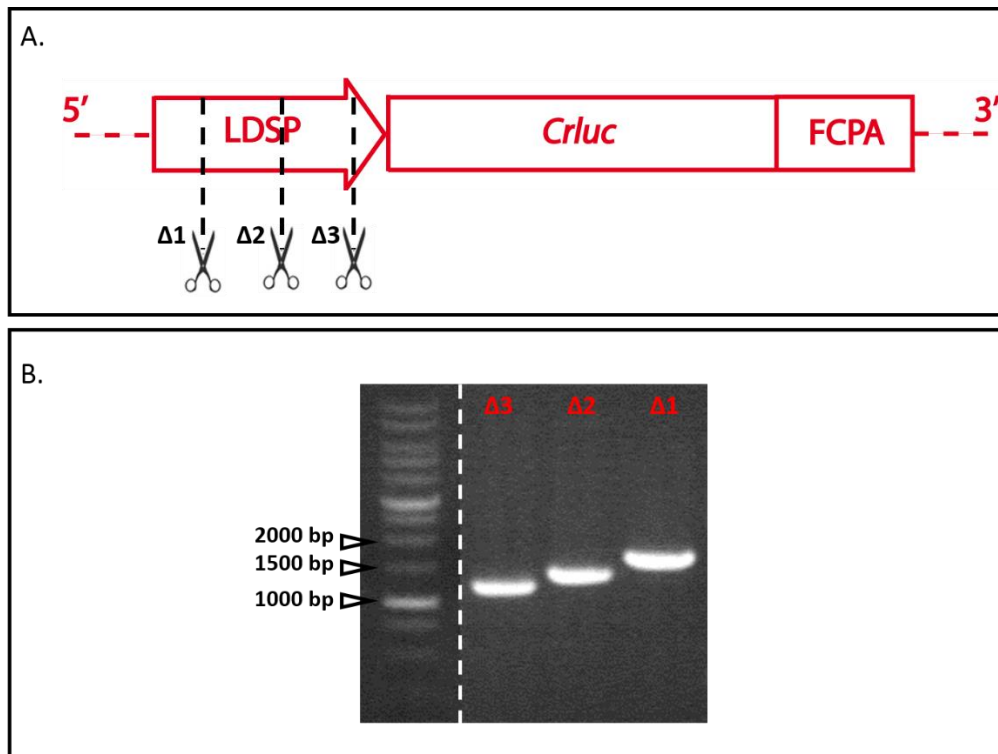
This promoter region was chosen since in RNA-seq experiments the related gene was shown to be transcribed at high levels (e.g. in medium light conditions, RUBOX gene was >4 times more expressed than the LDSP gene, Figure 8A), highlighting its suitability as molecular region to drive a strong gene expression. Thanks to this high-throughput strategy, we measured a hundred of strains simultaneously (Figure 6B). Among the tested strains, we decided to use the average value of luciferase activity  $\pm 2$  SD, as a threshold for selection of best performing strains. Therefore we retained only those transformants falling outside that interval. Thanks to this screening procedure we were indeed able to quickly isolate the best luciferase expressing strains (Figure 6B and C), also in this case.

*Identification of the minimal regulatory sequences and the most promising promoter regions to drive heterologous expression in N. gaditana.*

The development of an efficient over-expression platform requires a deep knowledge of the molecular elements necessary to drive transcription, as well as their regulation. Since this information is largely incomplete in *N. gaditana*, we decided to identify the optimal boundaries required for maximal promoter activity, using the LDSP promoter region as reference. 5' - end deletion analysis of this promoter region is being planned in order to remove from the 5' – promoter end 203, 437 and 590 bp to get the LDSP  $\Delta 1$ ,  $\Delta 2$  and  $\Delta 3$  deletions, respectively (Figure 7A).

These deletions were obtained amplifying the proper molecular regions from the full-length promoter sequence (primer sequences are shown in figure 3 and in table 1, #11, #12 and #13) and fusing them to the *Crluc* coding sequence (Figure 7B), through PCR-driven overlap extension technique (see Methods for details). The planned deletions are useful to remove specific molecular elements or reduce their number (Table 2), in order to assess their importance to drive transcription. Furthermore the LDSP promoter region was previously chosen since the associated gene shows a strong up-regulation in response to light intensity increase (Figure 2B). This means that this promoter region should include some light –responsive molecular elements, able to improve promoter activity when light intensity increases. Therefore the planned 5' – end deletion analysis will allow evaluating the activity of these *in silico* identified elements (Table 2) in order to assess their functionality as key transcriptional regulators. To date we concluded the vectors development (Figure 8B) and, as soon as *N. gaditana* transformants are isolated, this analysis will lead to a huge step forward in the

identification of the minimal molecular elements necessary to drive transcription, as well as the characterization of the light-responsive molecular elements in this promising organisms.



**Figure 7. Study of the minimal molecular region necessary to drive transcription in *Nannochloropsis gaditana*.** A. Promoter LDSP (677 bp) was cut to obtain 3 molecular regions ( $\Delta 1$ ,  $\Delta 2$  and  $\Delta 3$ ) progressively shorter (474, 239 and 87 bp respectively). The latter were used to drive the expression of the *Crluc* gene in order to test their functionality. LDSP promoter deleted regions were fused to the *Crluc* coding sequence by PCR-driven overlap extension. B. Agarose gel in which the LDSP deleted regions fused to the *Crluc* gene were loaded;  $\Delta 1 = 1410$  pb,  $\Delta 2 = 1175$  pb,  $\Delta 3 = 1023$  pb.

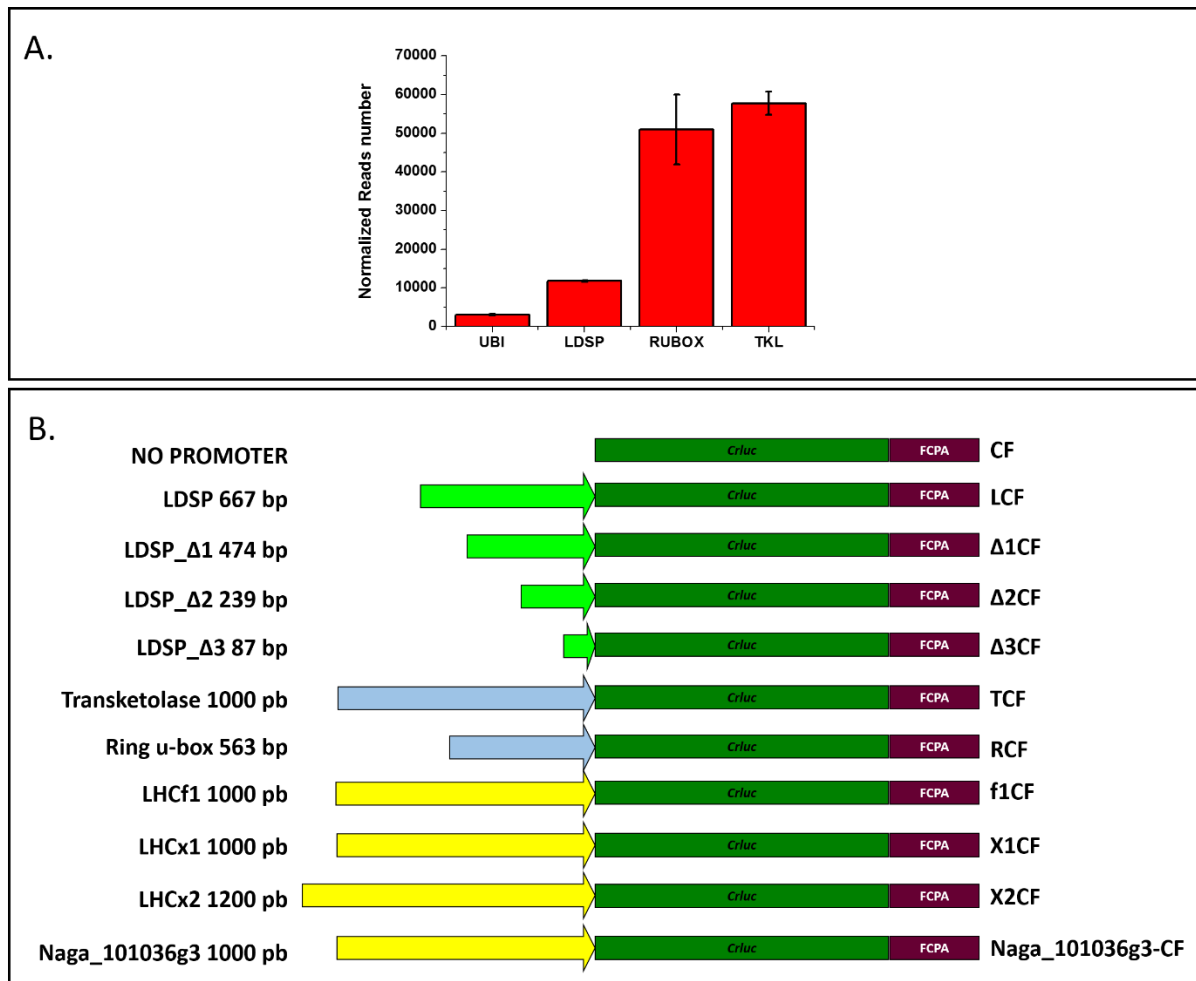


**Table 2. Relevant molecular elements present in full-length and LDSP promoter deletions.** The molecular elements here reported are among the most relevant identified with the PlantCARE software (Rombauts et al., 1999). CATT-motif and Sp1 are light-responsive elements; CGTCA-motif, TGACG-motif and TGA-element are hormone-responsive elements; TATA-box and CAAT-box are necessary for RNA-polimerase activity.

Molecular element	LDSP Full length	LDSP Δ1	LDSP Δ2	LDSP Δ3
	# identified motives			
TATA-box	5	5	3	2
CAAT-box	7	4	1	/
CATT-motif	1	1	1	/
CGTCA-motif	5	4	2	/
TGACG-motif	5	4	2	/
circadian	1	1	1	/
Sp1	5	2	/	/
TGA-element	1	1	/	/

The LCF transformation vector was also developed as a molecular tool to study regulatory sequences in *N. gaditana*. Therefore we started to clone several promoter regions upstream the *Crluc* gene in order to isolate the best one to achieve a high luciferase activity. The first chosen molecular regions are up to 1000 bp long and start from the translational-start site of 2 genes (RUBOX, Gene ID - Naga\_100094g14 and Tkl, Gene ID - Naga\_100001g128), whose transcription was found very high, as shown by the data get from the RNA-seq experiment described in Chapter V (Figure 8A). Particular attention was dedicated to the 5' – UTR regions of 4 genes coding for antenna proteins (LHCf1, Gene ID: 100012g50; LHCX1, Gene ID: 100173g12; LHCX2, Gene ID: 100056g42; gene Naga\_101036g3). These regulatory elements were chosen since the associated genes show a regulation strongly light-dependent (Figure 4 of Chapter V), and therefore their 5'-UTR regions could be studied for application as light-inducible promoters. During acclimation LHCf1, LHCX1 and LHCX2 show the same light intensity response, with an up-regulation in LL and a down-regulation in ML and HL conditions, while the protein encoded by the gene Naga\_101036g3 show the opposite trend (Figure 4 of Chapter V). This means that the study of these promoters could reveal the presence of molecular elements responding to light in an opposite way, therefore leading to improve the current knowledge on transcriptional regulation. To date we concluded the preparation of the vectors (all the current available vectors are depicted in Figure 8B) and their future employment will lead to the identification of the best promoters to efficiently drive protein expression in *N. gaditana*, as well as to the

improvement of the current knowledge on the light-responsive molecular elements in this promising organism.



**Figure 8. Currently available transformation vectors.** A. Normalized reads number coming from the RNA-seq experiment of chapter V, for some genes whose 5' – UTR region was cloned in this work. B. Several transformation vectors, carrying different promoters upstream of the *Crluc* gene (in dark green), were designed in order to study transcriptional regulatory sequences in *N. gaditana*. On the left of the overexpression cassette scheme there is the name of the promoter region and its length, expressed in bp. On the right, there is the name of the over-expression vectors thus obtained. The CF vector shows no promoter region to drive *Crluc* transcription and it will be used as negative control.

## Conclusion.

The development of new protein expression platforms is necessary to overcome the limitations widespread in the currently available expression systems. Here we developed a standardized and modular expression vector, harboring the *R. reniformis* luciferase reporter gene, in order to assess *N. gaditana* potentiality as expression host. *N. gaditana* showed to accumulate luciferase at higher levels than the current reference species, *C. reinhardtii*. For the first time we described the efficient use of this reporter gene in *N. gaditana* and the development of a reliable high-throughput assay to *in vivo* monitor bioluminescence from these cells. The developed tools will be thus used to evaluate the best regulatory sequences to drive transcription and to improve our knowledge on the light-inducible molecular elements regulating transcription in this promising organism.

## References.

- Cao, M., Fu, Y., Guo, Y., and Pan, J.** (2009). Chlamydomonas (Chlorophyceae) colony PCR. *Protoplasma* **235**: 107–10.
- Cardi, T., Lenzi, P., and Maliga, P.** (2014). Chloroplasts as expression platforms for plant-produced vaccines. *Expert Rev. Vaccines*.
- Corteggiani Carpinelli, E., Telatin, A., Vitulo, N., Forcato, C., D'Angelo, M., Schiavon, R., Vezzi, A., Giacometti, G.M., Morosinotto, T., and Valle, G.** (2014). Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion. *Mol. Plant* **7**: 323–35.
- Ferrante, P., Catalanotti, C., Bonente, G., and Giuliano, G.** (2008). An optimized, chemically regulated gene expression system for *Chlamydomonas*. *PLoS One* **3**: e3200.
- Fuhrmann, M., Hausherr, A., Ferbitz, L., Schödl, T., Heitzer, M., and Hegemann, P.** (2004). Monitoring dynamic expression of nuclear genes in *Chlamydomonas reinhardtii* by using a synthetic luciferase reporter gene. *Plant Mol. Biol.* **55**: 869–81.
- Gangl, D., Zedler, J.A.Z., Włodarczyk, A., Jensen, P.E., Purton, S., and Robinson, C.** (2015). Expression and membrane-targeting of an active plant cytochrome P450 in the chloroplast of the green alga *Chlamydomonas reinhardtii*. *Phytochemistry* **110**: 22–8.
- Hamilton, M.L., Warwick, J., Terry, A., Allen, M.J., Napier, J.A., and Sayanova, O.** (2015). Towards the Industrial Production of Omega-3 Long Chain Polyunsaturated Fatty Acids from a Genetically Modified Diatom *Phaeodactylum tricornutum*. *PLoS One* **10**: e0144054.
- Harris, E.H.** (2003). CHLAMYDOMONAS AS A MODEL ORGANISM.
- Heckman, K.L. and Pease, L.R.** (2007). Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat. Protoc.* **2**: 924–32.
- Kaye, Y., Grundman, O., Leu, S., Zarka, A., Zorin, B., Didi-Cohen, S., Khozin-Goldberg, I., and Boussiba, S.** (2015). Metabolic engineering toward enhanced LC-PUFA biosynthesis in *Nannochloropsis oceanica*: Overexpression of endogenous  $\Delta 12$  desaturase driven by stress-inducible promoter leads to enhanced deposition of polyunsaturated fatty acids in TAG. *Algal Res.* **11**: 387–398.
- Kilian, O., Benemann, C.S.E., Niyogi, K.K., and Vick, B.** (2011). High-efficiency homologous recombination in the oil-producing alga *Nannochloropsis* sp. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 21265–21269.

- Lassen, L.M., Nielsen, A.Z., Ziersen, B., Gnanasekaran, T., Møller, B.L., and Jensen, P.E.** (2014). Redirecting photosynthetic electron flow into light-driven synthesis of alternative products including high-value bioactive natural compounds. *ACS Synth. Biol.* **3**: 1–12.
- Lauersen, K.J. et al.** (2015a). Investigating the dynamics of recombinant protein secretion from a microalgal host. *J. Biotechnol.* **215**: 62–71.
- Lauersen, K.J., Berger, H., Mussnug, J.H., and Kruse, O.** (2013). Efficient recombinant protein production and secretion from nuclear transgenes in *Chlamydomonas reinhardtii*. *J. Biotechnol.* **167**: 101–10.
- Lauersen, K.J., Kruse, O., and Mussnug, J.H.** (2015b). Targeted expression of nuclear transgenes in *Chlamydomonas reinhardtii* with a versatile, modular vector toolkit. *Appl. Microbiol. Biotechnol.* **99**: 3491–503.
- Loening, A.M., Fenn, T.D., and Gambhir, S.S.** (2007). Crystal structures of the luciferase and green fluorescent protein from *Renilla reniformis*. *J. Mol. Biol.* **374**: 1017–28.
- Mathieu-Rivet, E., Kiefer-Meyer, M.-C., Vanier, G., Ovide, C., Burel, C., Lerouge, P., and Bardor, M.** (2014). Protein N-glycosylation in eukaryotic microalgae and its impact on the production of nuclear expressed biopharmaceuticals. *Front. Plant Sci.* **5**: 359.
- Nielsen, A.Z., Ziersen, B., Jensen, K., Lassen, L.M., Olsen, C.E., Møller, B.L., and Jensen, P.E.** (2013). Redirecting photosynthetic reducing power toward bioactive natural product synthesis. *ACS Synth. Biol.* **2**: 308–15.
- Niu, Y.-F., Zhang, M.-H., Li, D.-W., Yang, W.-D., Liu, J.-S., Bai, W.-B., and Li, H.-Y.** (2013). Improvement of neutral lipid and polyunsaturated fatty acid biosynthesis by overexpressing a type 2 diacylglycerol acyltransferase in marine diatom *Phaeodactylum tricorutum*. *Mar. Drugs* **11**: 4558–69.
- Peng, K.-T., Zheng, C.-N., Xue, J., Chen, X.-Y., Yang, W.-D., Liu, J.-S., Bai, W., and Li, H.-Y.** (2014). Delta 5 fatty acid desaturase upregulates the synthesis of polyunsaturated fatty acids in the marine diatom *Phaeodactylum tricorutum*. *J. Agric. Food Chem.* **62**: 8773–6.
- Perin, G., Bellan, A., Segalla, A., Meneghesso, A., Alboresi, A., and Morosinotto, T.** (2015). Generation of random mutants to improve light-use efficiency of *Nannochloropsis gaditana* cultures for biofuel production. *Biotechnol. Biofuels* **8**: 161.
- Plucinak, T.M., Horken, K.M., Jiang, W., Fostvedt, J., Nguyen, S.T., and Weeks, D.P.** (2015). Improved and versatile viral 2A platforms for dependable and inducible high-level expression of dicistronic nuclear genes in *Chlamydomonas reinhardtii*. *Plant J.* **82**: 717–29.

- Radakovits, R., Jinkerson, R.E., Fuerstenberg, S.I., Tae, H., Settlage, R.E., Boore, J.L., and Posewitz, M.C.** (2012). Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. *Nat. Commun.* **3**: 686.
- Ramos Tercero, E.A., Sforza, E., Morandini, M., and Bertuccio, A.** (2014). Cultivation of *Chlorella protothecoides* with urban wastewater in continuous photobioreactor: biomass productivity and nutrient removal. *Appl. Biochem. Biotechnol.* **172**: 1470–85.
- Rasala, B.A., Chao, S.-S., Pier, M., Barrera, D.J., and Mayfield, S.P.** (2014). Enhanced genetic tools for engineering multigene traits into green algae. *PLoS One* **9**: e94028.
- Rasala, B.A., Lee, P.A., Shen, Z., Briggs, S.P., Mendez, M., and Mayfield, S.P.** (2012). Robust expression and secretion of Xylanase1 in *Chlamydomonas reinhardtii* by fusion to a selection gene and processing with the FMDV 2A peptide. *PLoS One* **7**: e43349.
- Rasala, B.A. and Mayfield, S.P.** (2015). Photosynthetic biomanufacturing in green algae; production of recombinant proteins for industrial, nutritional, and medical uses. *Photosynth. Res.* **123**: 227–39.
- Rombauts, S., Déhais, P., Van Montagu, M., and Rouzé, P.** (1999). PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res.* **27**: 295–6.
- Rosales-Mendoza, S., Paz-Maldonado, L.M.T., and Soria-Guerra, R.E.** (2012). *Chlamydomonas reinhardtii* as a viable platform for the production of recombinant proteins: current status and perspectives. *Plant Cell Rep.* **31**: 479–94.
- Salih, F.M.** (2011). Microalgae Tolerance to High Concentrations of Carbon Dioxide: A Review. *J. Environ. Prot. (Irvine, Calif.)* **02**: 648–654.
- Shao, N. and Bock, R.** (2008). A codon-optimized luciferase from *Gussia princeps* facilitates the in vivo monitoring of gene expression in the model alga *Chlamydomonas reinhardtii*. *Curr. Genet.* **53**: 381–8.
- Simionato, D., Block, M.A., La Rocca, N., Jouhet, J., Maréchal, E., Finazzi, G., and Morosinotto, T.** (2013). The response of *Nannochloropsis gaditana* to nitrogen starvation includes de novo biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. *Eukaryot. Cell* **12**: 665–76.
- Ullrich, K.K., Hiss, M., and Rensing, S.A.** (2015). Means to optimize protein expression in transgenic plants. *Curr. Opin. Biotechnol.* **32**: 61–7.
- Vanier, G., Hempel, F., Chan, P., Rodamer, M., Vaudry, D., Maier, U.G., Lerouge, P., and Bardor, M.** (2015). Biochemical Characterization of Human Anti-Hepatitis B Monoclonal Antibody Produced

in the Microalgae *Phaeodactylum tricornutum*. PLoS One **10**: e0139282.

**Zhang, J. et al.** (2014). Overexpression of the soybean transcription factor GmDof4 significantly enhances the lipid content of *Chlorella ellipsoidea*. Biotechnol. Biofuels **7**: 128.

# CHAPTER VII

## LHCX3, the First Stress-Related LHC Fused to a Fasciclin 1-Like Domain, Identified in *Nannochloropsis Gaditana*

### Author names and affiliations

Giorgio Perin<sup>1</sup>, Mattia Storti<sup>1</sup>, Alessandro Alboresi<sup>1</sup> and  
Tomas Morosinotto<sup>1</sup>

<sup>1</sup>PAR-Lab\_Padua Algae Research Laboratory, Department of Biology,  
University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy

### Abstract.

Chapter V showed the regulation of *Nannochloropsis gaditana* transcriptome in response to different light intensities. Twenty-one genes encoding Light Harvesting complex (LHC) proteins were identifiable in the genome and they all showed an up-regulation in low light conditions. The only exception was the gene Naga\_101036g3, that was up-regulated in response to stronger light intensity, suggesting a specialized function for the encoded protein. Unfortunately an imperfect identification of the gene coding sequence impaired further analysis. In this chapter we report the experimental identification of its coding sequence and its consequent classification as LHCX3. This allowed observing that the latter indeed includes an unusual fasciclin 1-like domain fused to a LHC. A preliminary sequence and phylogenetic analysis was carried out to infer the evolution and function of the fusion between the two domains.





## Introduction.

Photosynthetic organisms entirely rely on light energy to drive oxygenic photosynthesis. The first step of this process is light energy capture, which is performed by pigment binding proteins, belonging to the multigenic family of light harvesting complex (LHC) proteins (Dittami et al., 2010). These are nuclear-encoded, integral thylakoid membrane proteins, coordinating a variable number of chlorophyll and xanthophyll molecules (Dall'Osto et al., 2015). They share a common structural organization, characterized by 3 membrane-spanning  $\alpha$  helices connected by stroma or lumen-exposed loops. Helix 1 and 3 are homologous, likely originated from internal sequence duplication (Green and Pichersky, 1994), and they show a conserved motif counting an Arg and Glu residue, which form a salt bridge stabilizing the two helices, in a left-handed supercoil (Kühlbrandt et al., 1994; Liu et al., 2004).

The LHC superfamily is composed by different members, playing a specialized role (Büchel, 2015). Among them there are the LHCX proteins, also named LHCSR in viridiplantae, which are found in several algal species but also in mosses, while they are missing from higher plants. These proteins have been shown to be more involved in photoprotection, dissipating excess energy, rather than harvesting light (Alboresi et al., 2010; Bailleul et al., 2010; Bonente et al., 2011).

In fact, when light energy is provided in excess, to prevent ROS formation the excited singlet chlorophyll can de-excite by dissipating heat, through the phenomenon of non-photochemical quenching (NPQ). LHCX proteins are essential for NPQ activation (Peers et al., 2009). In contrast to LHC proteins mainly involved in light harvesting, that are down-regulated in high light (Erickson et al., 2015), most LHCX/LHCSR genes are instead induced by high light (Lepetit et al., 2013; Peers et al., 2009; Bonente et al., 2011).

In the *Nannochloropsis gaditana* genome, two LHCX proteins were previously identified (LHCX1 - Gene ID: Naga\_100173g12 and LHCX2 - Gene ID: Naga\_100056g42 - (Basso et al., 2014), according to sequence analysis. However when *N. gaditana* cells were subjected to genome-wide analysis in response to light intensity increase, we found a third gene possibly falling in this clade, that we consequently called LHCX3 (Gene ID: Naga\_101036g3) (see Chapter V). However, the coding sequence of the gene wasn't predicted correctly, since missing common molecular elements necessary to drive transcription efficiently. Moreover it missed two out of the three conserved membrane-spanning  $\alpha$ -helices, typical of LHC proteins. In this chapter we present the improvement of the gene model for LHCX3, also verified by experimental evidences.

## Materials and Methods.

### *Nannochloropsis gaditana* cultivation.

*Nannochloropsis gaditana* strain 849/5 from Culture Collection of Algae and Protozoa (CCAP) was cultivated in sterile F/2 medium, containing sea salts (32 g/L, Sigma Aldrich), 40 mM Tris-HCl (pH 8), Guillard's (F/2) marine water enrichment solution (Sigma Aldrich), at 100  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  illumination and  $22 \pm 1$  °C with 100 rpm agitation. Algal growth was measured counting the cells using a cell counter (Cellometer Auto X4, Nexcelom Bioscience).

### *Genomic DNA, total RNA extraction and cDNA preparation.*

*N. gaditana* genomic DNA and total RNA were extracted from a 4-day-old culture, performed in Erlenmeyer flasks using F/2 liquid media. Cells were lysed using a Mini Bead Beater (Biospec Products) at 3500 RPM for 20 s in the presence of glass beads (150–212  $\mu\text{m}$  diameter). Genomic DNA was then purified using the *EUROGOLD™ Plant DNA Mini Kit* (Euroclone), applying minor modifications. Total RNA was instead purified using the *RNeasy™ Plant Mini Kit* (Qiagen), applying minor modifications to the manufacturer's instruction. Genomic DNA and total RNA concentration and purity were determined by 100 UV-VIS spectrophotometer (Cary Series, Agilent Technologies). cDNA was then prepared from 2  $\mu\text{g}$  of total RNA-template with the RevertAid Reverse Transcriptase cDNA kit (Thermo Fisher Scientific, Epsom, UK).

### *PCR on genomic DNA and RT-PCR.*

The genomic DNA was used as template for PCR in order to fill the gap between the two contigs in the genome, applying primer sequences of Table 1 (#1 and #2).

The cDNA was instead used as template for the RT-PCR, useful to amplify the new LHCX3 gene model with the primer sequences reported in table 1 (#3 and #4). For both PCR, the KAPA HIFI high-fidelity DNA polymerase (KAPA Biosystems) was used.

**Table 1. Primer sequences used to fill the gaps in the LHCX3 gene coding sequence through PCR on genomic DNA and RT-PCR.** Primer #1 binds to the upstream, while primer #2 to the downstream contig sequence. Primer #3 binds to the Naga\_100967g1 upstream gene from its ATG, while primer #4 binds to the Naga\_101036g3 downstream gene, reaching its stop codon.

Primer #	Name	Primer Sequence (5' – 3')
1	Up_contig	GAACGTGGTG GCGAC
2	Down_contig	CCCATCGGGCTC AGG
3	Naga6full	ATACTGCAGATGCGTGTCTAGTTTCTCT
4	Naga6BglII	TATAGATCTTTAGGAGAAGATTTTTGAGTTGG

*Sequence analysis.*

*Multiple sequence alignments and three-dimensional modeling.* Multiple sequence alignments of the chlorophyll binding proteins and the fasciclin 1-like domain containing proteins were generated using the Clustal W platform integrated in BioEdit 7.2.5 (Hall, 1999), and manually refined using the latter, on the basis of the crystal structure of the chlorophyll a/b binding protein from spinach (PDB code 1RWT, (Liu et al., 2004)), and of the solved NMR structure of the single-domain bacterial fasciclin I protein from the photosynthetic bacterium *Rhodobacter sphaeroides* (PDB code: 1W7E, (Moody and Williamson, 2013)), respectively. The structures of the 1RWT and of 1W7E proteins from the PDB were used to generate a three dimensional (3D) model of the C-terminal (from 191 to 370 aa) and the N-terminal (from 1 to 370 aa) of the LHCX3 protein, respectively, using the Phyre2 software (Kelley et al., 2015), with default parameters. The multiple sequence alignment and the 3D models were displayed and overlapped using ESPript (Robert and Gouet, 2014) and PyMOL (DeLano, 2005), respectively. *Phylogenetic analysis.* Phylogenetic trees were constructed after having aligned the sequences using Clustal W with default settings. Neighbor-Joining trees were generated using MEGA6 (Tamura et al., 2013).

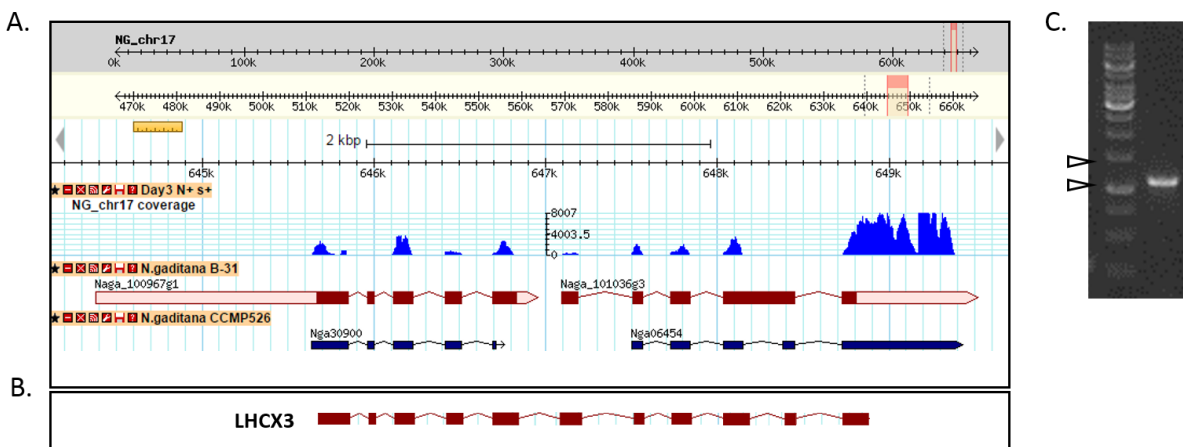
## Results and discussion.

In chapter V, after *N. gaditana* exposition for 5 days to increasing doses of light, we showed that among the 21 LHC proteins present in its genome, only the gene Naga\_101036g3 was found up-regulated in response to high light intensities, thus suggesting a specialized role for this gene in photoprotection and light acclimation, that deserved further investigation.

### LHCX3 sequence analysis.

Gene Naga\_101036g3 coding sequence in v1.0 genome was shown to conserve well only the first alpha helix typical of LHC proteins, while it missed the other 2. Moreover, the sequence also missed the usual translation-starting codon (ATG) and the typical N-terminal transit peptide found in nuclear-encoded chloroplastic proteins, according to common and Heterokonts specialized bioinformatic tools for prediction of transit peptide localization (SignalP (Petersen et al., 2011), ChloroP (Emanuelsson et al., 1999) and HECTAR (Gschloessl et al., 2008)), even by searching within 100 bp upstream and 100 bp downstream for an alternative translation start site.

This information suggested the possibility of the identification of a novel one-helix LHC protein (Neilson and Durnford, 2010b, 2010a) or more likely an imprecise identification of this gene in the genome. The latter hypothesis became more likely after the identification of a series of unidentified nucleotides (N) in the upstream region of Naga\_101036g3 gene, revealing a possible imperfect merging of two contigs in the step of genome assembling (Corteggiani Carpinelli et al., 2014) (Figure 1A).



**Figure 1.** LHCX3 (Gene ID: Naga\_101036g3) locus in *Nannochloropsis gaditana* genome. A. Magnification of the chromosome 17 (from 644,1 Kb to 649,7 kb) as reported in the *N. gaditana*

genome browser. The image of the genomic *locus* was taken from [www.nannochloropsis.org](http://www.nannochloropsis.org). In red rectangles are depicted the exons while in red lines the introns, according to the annotation published in (Corteggiani Carpinelli et al., 2014). In light-red are depicted the genes 5' or 3' – UTR regions. In blue rectangles are depicted the exons while in blue lines the introns, according to the annotation published in (Radakovits et al., 2012). In bright-blue are depicted the reads counts for *N. gaditana* strain B-31 in standard growing conditions. B. Agarose gel in which the RT-PCR product coming from the amplification of the merged coding sequences of two genes was loaded. The two arrows indicate 1000 (bottom) and 1500 bp (top), respectively. C. LHCX3 new coding sequence coming from the merging of the coding sequences of the Naga\_100967g1 and Naga\_101036g3 genes.

Using PCR on the genomic DNA (primer sequences of table 1, #1 binds to the upstream contig while #2 to the downstream contig) and sequencing the molecular product, we identified that the sequence was indeed complete and the two contigs were adjacent with no gap in between them. Thus we hypothesized that the starting codon could be found upstream of its original, wrongly predicted, position.

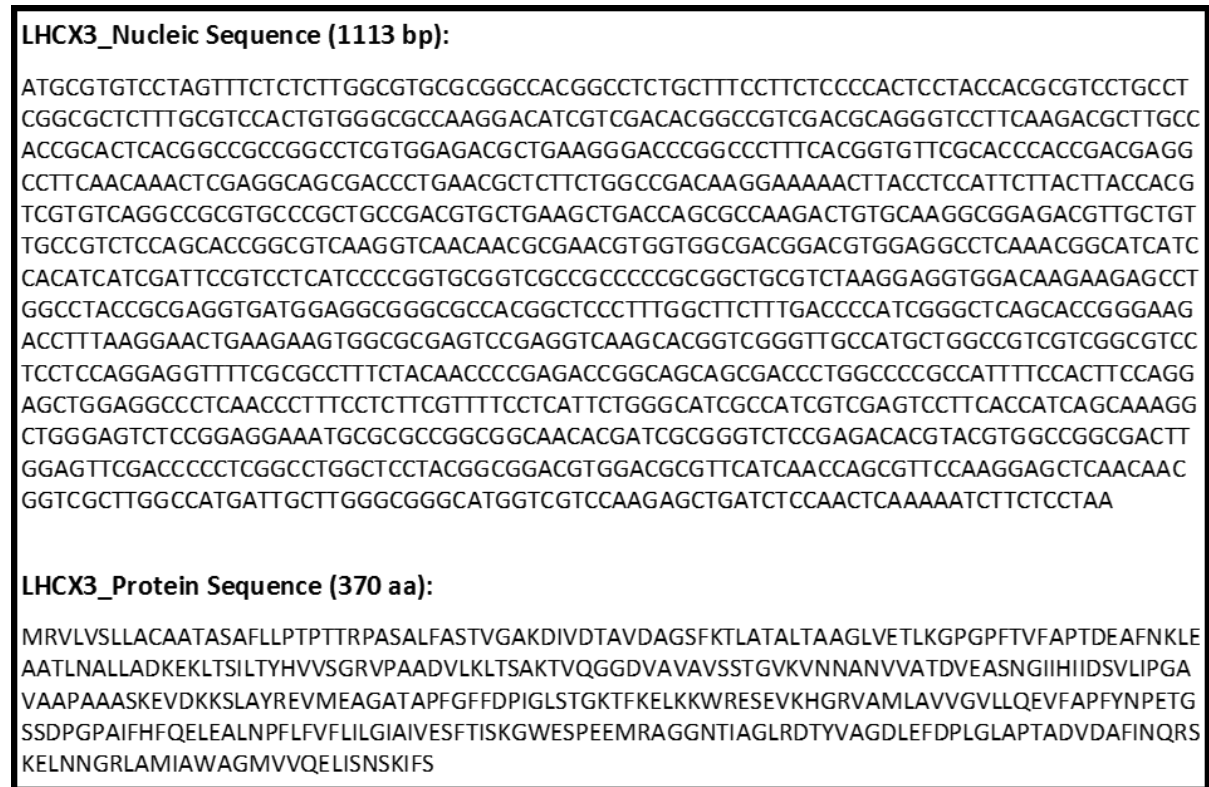
We then used RT-PCR to identify the starting codon of the Naga\_101036g3 gene. Using primer sequences #2 and #3 of table 1, we obtained a PCR product >1000 bp long (Figure 1B), that included also another encoded sequence that was annotated as a separate gene (Gene ID: Naga\_100967g1 in Figure 1A), upstream of the Naga\_101036g3 gene. We thus worked on a new hypothesis that Naga\_100967g1 and Naga\_101036g3 were a single gene. This was supported by their highly similar behavior at transcriptional level (Table 2) in response to the light intensity.

**Table 2. LogFC values for the Naga\_100967g1 and Naga\_101036g3 genes (see annotation of (Corteggiani Carpinelli et al., 2014)), according to the transcriptomics experiment described in chapter V.**

Gene ID	LogFC		
	LL vs ML	ML vs HL	LL vs HL
Naga_100967g1	-1.50	0.45	-1.05
Naga_101036g3	-1.44	0.91	-0.53

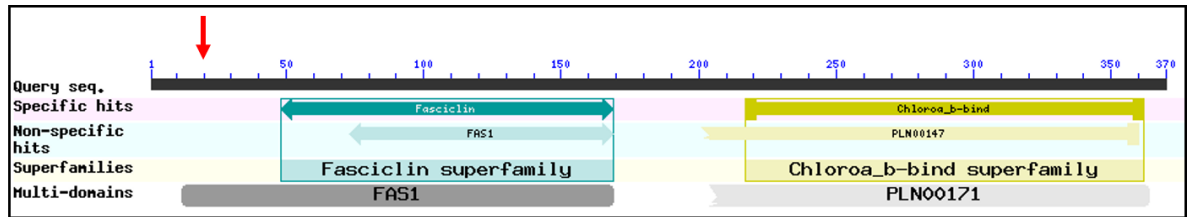
The PCR product amplified (from the ATG of the Naga\_100967g1 gene to the stop codon of the Naga\_101036g3 gene) was re-sequenced and Naga\_101036g3 gene was re-named LHCX3, after a new gene model was thus constructed (see figure 2 for the gene and protein sequence details). Differently than that expected from a simple merging between the two coding sequences, the new gene model

also showed modifications in exons and introns distribution, as well as in their length, with respect to the old gene models. LHCX3 in fact counts 11 exons and 10 introns (Figure 1C).



**Figure 2. LHCX3 nucleic and amino acids sequence, according to the new coding sequence model defined in this chapter.**

Accordingly with the transcriptional profile in figure 1A, the 4<sup>th</sup> exon, as previously annotated for the Naga\_101036g3 gene, was instead splitted in two (the 9<sup>th</sup> and the 10<sup>th</sup>). Moreover the last exon was shifted a few tens base pairs in length to meet the frame needs. The new LHCX3 coding sequence showed all conserved features of LHC proteins, with three transmembrane helices. Surprisingly, the sequence also showed an extra N-terminal domain, sharing high homology with the fasciclin 1-like domain (FAS1:PF02469) (Figure 3).



**Figure 3. Domains identified in the translated new coding sequence of the LHCX3 gene.** The fascidin 1-like and the chlorophyll a/b binding superfamily domains are highlighted in light-blue and in yellow, respectively. The output comes from the NCBI's conserved domains database (Marchler-Bauer et al., 2014). The red arrow represents the predicted cleavage site, according to the bioinformatic prediction tool HECTAR (Gschloessl et al., 2008), separating on the left the chloroplast transit peptide.

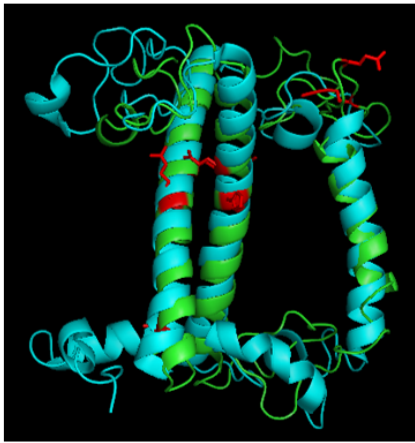
The coding sequence of the latter is upstream preceded by a chloroplastic transit peptide, according to the bioinformatic tool HECTAR (Gschloessl et al., 2008). Taken together these analyses, clearly show that *LHCX3* gene model indeed codifies for a fusion protein, counting these two molecular domains.

*LHCX3 C-terminal region is a chlorophyll binding domain.*

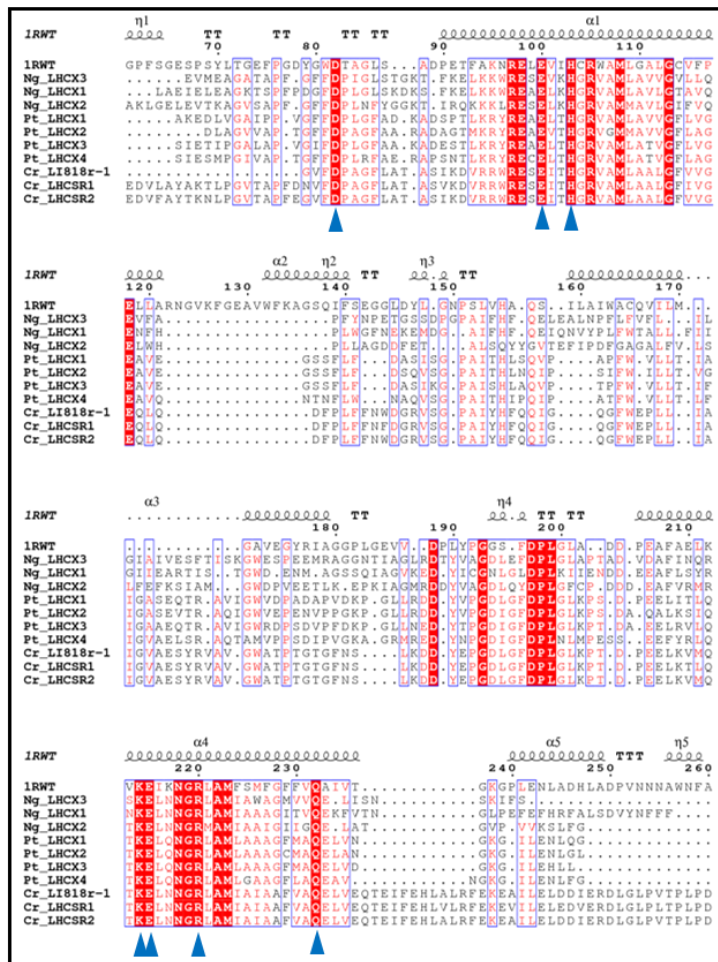
To confirm the reliability of the new LHCX3 protein model, we assessed 3D homology modeling by comparing its chlorophyll binding domain (from 191 to 370 aa) with the crystallized chlorophyll a/b binding protein from spinach (PDB code: 1RWT, (Liu et al., 2004)).



A.



B.



**Figure 4. LHCX3 chlorophyll binding domain.** A. Structural overlay between the 3D model of the chlorophyll binding domain of LHCX3 (from the 191 to the 370 aa, in green) with the crystallized structure of the chlorophyll a/b binding protein from spinach (in light-blue—PDB code: 1RWT). In red are depicted the conserved amino acids involved in chlorophyll a binding. The two proteins are depicted in a stereo ribbon representation. B. Structure-based sequence alignment of the crystallized chlorophyll a/b binding protein from spinach (PDB code: 1RWT) with the chlorophyll binding domain of proteins belonging to the LHCX clade. The secondary structure of the spinach protein is shown above the alignment. Identical amino acids are highlighted by a red background while in red letters are depicted those similar. Alpha helices are represented as helices and  $\beta$ -turns are marked with TT. Blue triangles indicate the conserved residues involved in the binding of chlorophyll a molecules. Ng – *Nannochloropsis gaditana*; Pt – *Phaeodactylum tricornutum*; Cr – *Chlamydomonas reinhardtii*.

The alpha helices  $\alpha 1$  and  $\alpha 4$  were well conserved between the two (Figure 4A), while the alpha helix  $\alpha 3$  shows the main differences.

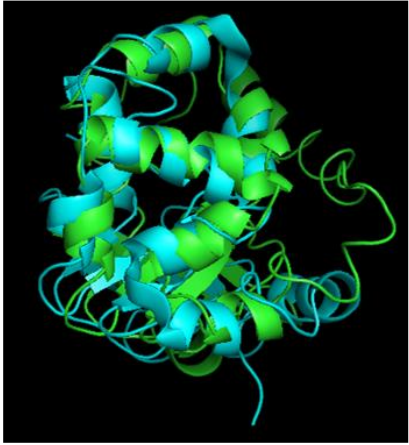
Moreover, the chlorophyll binding domain of LHCX3 shares common features with the LHCX from *P. tricornutum* and LHCSR from *C. reinhardtii*, as confirmed by the multiple sequence alignment in figure 4B. The main chlorophyll a binding sites are well conserved among all proteins considered, as shown by the blue arrows in figure 4B. Taken together these results clearly show that the C-terminal region of the LHCX3 protein can be reliably classified as a chlorophyll binding domain, falling in the LHCX clade.

*LHCX3 N-terminal region is a fasciclin 1-like domain.*

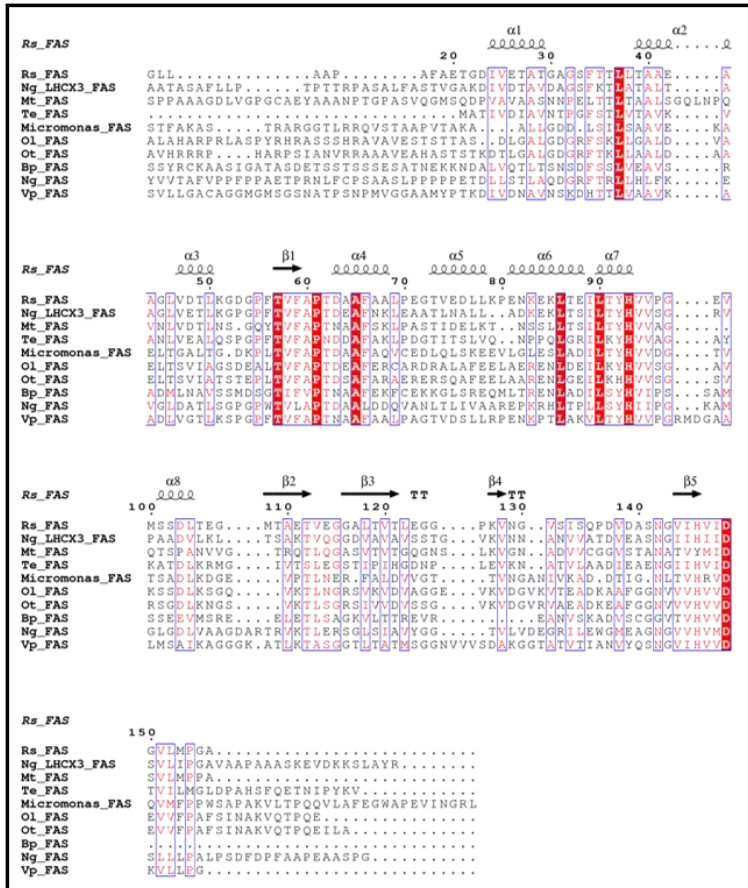
To infer the structural features characterizing the fasciclin 1-like domain of LHCX3 (from 1 to 190 aa) we compared the solved NMR structure of the single-domain bacterial fasciclin 1 protein from the photosynthetic bacterium *Rhodobacter sphaeroides* (PDB code: 1W7E, (Moody and Williamson, 2013)) with the 3D model of the former, constructed by homology modeling according to the methods section. The fasciclin 1-like domain of LHCX3 shows the 58 % of identity with the 1W7E protein, therefore it was sufficient to generate a reliable 3D model. The alpha helices  $\alpha 1$ ,  $\alpha 4$ ,  $\alpha 6$ ,  $\alpha 7$  and the beta sheet  $\beta 1$  are well conserved (Figure 5A), as confirmed by the multiple sequence alignment (Figure 5B) with other proteins showing a fasciclin 1-like domain. The main differences in the secondary structure fall in the alpha helices  $\alpha 3$  and  $\alpha 5$ . Taken together these results clearly show that the N-terminal region of the LHCX3 protein can be reliably classified as a fasciclin 1-like domain.

The fasciclin 1-like domain was mainly studied in mammals, so far, often belonging to proteins strongly induced in tumor cells in response to a transforming growth factor- $\beta$  (TGF- $\beta$ ) treatment (Kawamoto et al., 1998). Proteins showing this domain are involved in cell adhesion processes (LeBaron et al., 1995), since they mediate interactions with collagen or proteins belonging to the extracellular matrix, and are often secretory or membrane-anchored proteins. Proteins showing fasciclin 1-like domains were also identified in *A. thaliana* (Johnson et al., 2003) and in *Triticum aestivum* as well as in *Oryza sativa* (Faik et al., 2006), and were called arabinogalactan proteins (AGPs).

A.



B.



**Figure 5. LHCX3 fasciclin 1-like domain.** A. Structural overlay between the 3D model of the fasciclin 1-like domain of LHCX3 (from the 1 to the 190 aa, in green) with the deposited structure (NMR solved) of the bacterial fasciclin 1 protein from the photosynthetic bacterium *Rhodobacter sphaeroides* (in light-blue – PDB code: 1W7E). The two proteins are depicted in a stereo ribbon representation. B. Structure-based sequence alignment of the deposited NMR solved single-domain bacterial fasciclin 1 protein from the photosynthetic bacterium *Rhodobacter sphaeroides* (PDB code: 1W7E) and here called Re-FAS with the fasciclin domain of different proteins. The secondary structure of the fasciclin protein from *Rhodobacter sphaeroides* is shown above the alignment. Identical amino acids are highlighted by a red background while in red letters are depicted those similar. Alpha helices are represented as helices,  $\beta$ -turns are marked with TT and  $\beta$ -sheets are depicted with a black arrow. Rs - *Rhodobacter sphaeroides*; Ng - *Nannochloropsis gaditana*; Mt - *Mycobacterium tuberculosis*; Te - *Thermosynechococcus elongatus*; Micromonas - *Micromonas sp. (strain RCC299)*; Ol - *Ostreococcus lucimarinus*; Ot - *Ostreococcus tauri*; Bp - *Bathycoccus prasinos*; Vp - *Variovorax paradoxus*. FAS – fasciclin domain.

The latter are a class of highly glycosylated proteins found mainly in the plant cell wall and plasma membrane, possibly involved also here in cells interaction. In the green alga *Volvox carteri* a protein

called Algal-CAM contains a fasciclin 1-like domain and it was identified as having a fundamental role in maintaining cell adhesion during embryogenesis (Huber and Sumper, 1994).

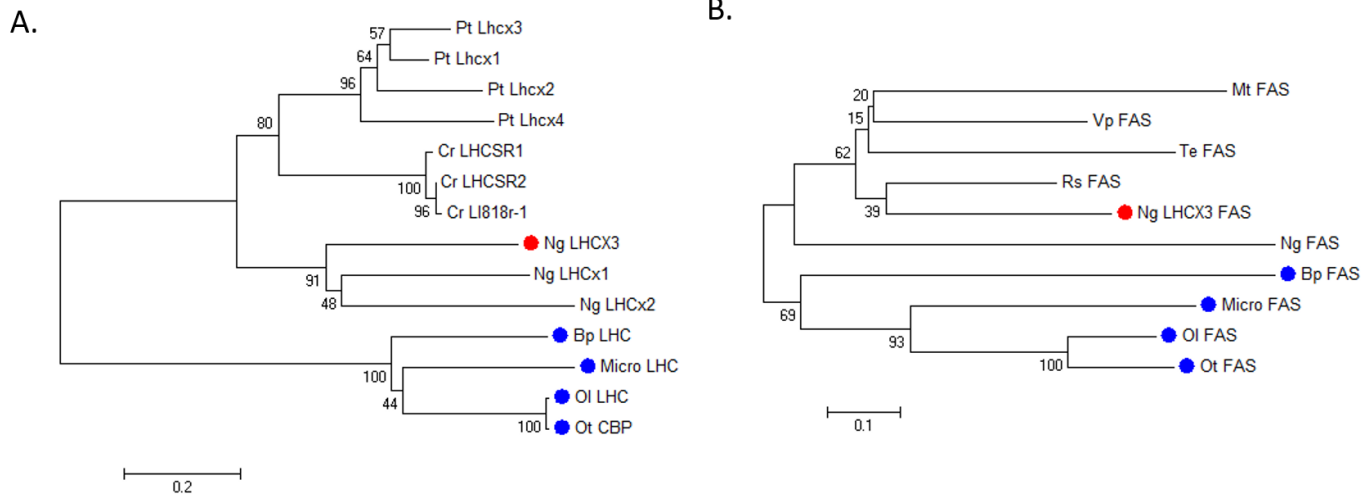
It is highly relevant to notice that a fasciclin 1-like domain was already found fused to a protein of the photosynthetic apparatus (Kawasaki et al., 2013), and in particular to a protein called AstaP, an orange carotenoids binding proteins (OCPs), mainly binding astaxanthin.

Sequence analysis also allowed identifying even LHC-like proteins (LIL) fused to a fasciclin 1-like domain, mainly in organisms belonging to the *Mamiellophyceae* class (Neilson and Durnford, 2010b, 2010a). The former are low-molecular-mass membrane proteins with a common evolutionary origin as LHC, and can have one, two, three or four alpha helices. Similar genes were then identified in organisms belonging to the *Micromonas* genus ((Worden et al., 2009) - GeneBank ID: ACO62775.1), in *Ostreococcus lucimarinus* ((Palenik et al., 2007) - GeneBank ID: ABO95507.1), in *Ostreococcus tauri* ((Derelle et al., 2006) - GeneBank ID: CAL51589.1) and in *Bathycoccus prasinus* ((Moreau et al., 2012) - GeneBank ID: CCO17162.1). The identification of other LHC proteins showing this fusion with a fasciclin 1-like domain highlights that this association isn't an isolated case but could reveal a possible common biological role. In our case, the presence of the fasciclin 1-like domain fused to a stress-related LHC protein raises questions about its possible biological role in photoprotection.

#### *Phylogenetic analysis.*

The annotated sequences of the LIL proteins containing a fasciclin 1-like domain (GeneBank IDs: ACO62775.1; ABO95507.1; CAL51589.1; CCO17162.1) were compared to the LHCX from *N. gaditana*, *P. tricornutum* and the LHCSR proteins from *C. reinhardtii*, using a multiple sequence alignment, after removal of the fasciclin 1-like domain (the N-terminal region taken from the 1 to the 190 aa). The multiple alignment was then used to build a phylogenetic tree (see Methods for details). As reported in figure 6A, two subgroups are clearly visible. The LHCX proteins from *N. gaditana* and *P. tricornutum* and the LHCSR proteins from *C. reinhardtii* are clearly forming a separate group, different from the LIL proteins. This suggests two conclusions: i) LHCX3 contains a genuine LHCX domain; ii) light-harvesting proteins acquired a fasciclin 1-like domain at least twice during evolution, since the LHC fraction is diverged before the fusion occurred. The gene fusion between the former and the chlorophyll binding domain took place independently in different organisms. This suggests some functional reason for such a convergent evolution.

A sequence analysis of the fasciclin 1-like domain is consistent with this hypothesis since again, the sequence fused with LHCs are not particularly similar (Figure 6B).



**Figure 6. Phylogenetic trees.** A. Phylogenetic tree of the chlorophyll binding domain for chlorophyll binding proteins, belonging to the LHCX clade (*P. tricornutum* and *C. reinhardtii*) and to the LIL clade (*B. prasinos*, *Ostreococcus* and *Micromonas*). B. Phylogenetic tree of the fasciclin 1-like domain for fasciclin containing proteins. In red is highlighted LHCX3 while in bright-blue other proteins showing the fasciclin domain fused to a chlorophyll binding domain. Abbreviations used are the same of figures 3 and 4.

### Conclusions.

In this work, *N. gaditana* LHCX3 gene model was re-calculated *in silico* and validated by PCR on genomic DNA and RT-PCR. We showed that it is indeed formed by the fusion of a N-terminal fasciclin 1-like domain and a C-terminal chlorophyll binding domain. Sequence analysis highlighted that LHCX3 is the first protein, belonging to the stress-related clade, fused to this domain, but other LHC-like proteins showed the same fusion. Phylogenetic analysis showed that the fusion between the two domains took place independently, at least twice during evolution. This speaks for an unclear function of the association between the fasciclin 1-like domains and LHC proteins, that open the doors for future investigation on the biological role of this association.

## References.

- Alboresi, A., Gerotto, C., Giacometti, G.M., Bassi, R., and Morosinotto, T.** (2010). Physcomitrella patens mutants affected on heat dissipation clarify the evolution of photoprotection mechanisms upon land colonization. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 11128–33.
- Bailleul, B., Rogato, A., de Martino, A., Coesel, S., Cardol, P., Bowler, C., Falciatore, A., and Finazzi, G.** (2010). An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 18214–9.
- Basso, S., Simionato, D., Gerotto, C., Segalla, A., Giacometti, G.M., and Morosinotto, T.** (2014). Characterization of the photosynthetic apparatus of the Eustigmatophycean *Nannochloropsis gaditana*: evidence of convergent evolution in the supramolecular organization of photosystem I. *Biochim. Biophys. Acta* **1837**: 306–14.
- Bonente, G., Ballottari, M., Truong, T.B., Morosinotto, T., Ahn, T.K., Fleming, G.R., Niyogi, K.K., and Bassi, R.** (2011). Analysis of LHCSR3, a protein essential for feedback de-excitation in the green alga *Chlamydomonas reinhardtii*. *PLoS Biol.* **9**: e1000577.
- Büchel, C.** (2015). Evolution and function of light harvesting proteins. *J. Plant Physiol.* **172**: 62–75.
- Corteggiani Carpinelli, E., Telatin, A., Vitulo, N., Forcato, C., D'Angelo, M., Schiavon, R., Vezzi, A., Giacometti, G.M., Morosinotto, T., and Valle, G.** (2014). Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion. *Mol. Plant* **7**: 323–35.
- Dall'Osto, L., Bressan, M., and Bassi, R.** (2015). Biogenesis of light harvesting proteins. *Biochim. Biophys. Acta* **1847**: 861–71.
- DeLano, W.L.** (2005). The case for open-source software in drug discovery. *Drug Discov. Today* **10**: 213–7.
- Derelle, E. et al.** (2006). Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci. U. S. A.* **103**: 11647–52.
- Dittami, S.M., Michel, G., Collén, J., Boyen, C., and Tonon, T.** (2010). Chlorophyll-binding proteins revisited--a multigenic family of light-harvesting and stress proteins from a brown algal perspective. *BMC Evol. Biol.* **10**: 365.
- Emanuelsson, O., Nielsen, H., and von Heijne, G.** (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**: 978–84.
- Erickson, E., Wakao, S., and Niyogi, K.K.** (2015). Light stress and photoprotection in *Chlamydomonas*

- reinhardtii. *Plant J.* **82**: n/a–n/a.
- Faik, A., Abouzouhair, J., and Sarhan, F.** (2006). Putative fasciclin-like arabinogalactan-proteins (FLA) in wheat (*Triticum aestivum*) and rice (*Oryza sativa*): identification and bioinformatic analyses. *Mol. Genet. Genomics* **276**: 478–94.
- Green, B.R. and Pichersky, E.** (1994). Hypothesis for the evolution of three-helix Chl a/b and Chl a/c light-harvesting antenna proteins from two-helix and four-helix ancestors. *Photosynth. Res.* **39**: 149–162.
- Gschloessl, B., Guerneur, Y., and Cock, J.M.** (2008). HECTAR: a method to predict subcellular targeting in heterokonts. *BMC Bioinformatics* **9**: 393.
- Hall, T.** (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**: 95–98.
- Huber, O. and Sumper, M.** (1994). Algal-CAMs: isoforms of a cell adhesion molecule in embryos of the alga *Volvox* with homology to *Drosophila* fasciclin I. *EMBO J.* **13**: 4212–22.
- Johnson, K.L., Jones, B.J., Bacic, A., and Schultz, C.J.** (2003). The fasciclin-like arabinogalactan proteins of *Arabidopsis*. A multigene family of putative cell adhesion molecules. *Plant Physiol.* **133**: 1911–25.
- Kawamoto, T., Noshiro, M., Shen, M., Nakamasu, K., Hashimoto, K., Kawashima-Ohya, Y., Gotoh, O., and Kato, Y.** (1998). Structural and phylogenetic analyses of RGD-CAP/beta ig-h3, a fasciclin-like adhesion protein expressed in chick chondrocytes. *Biochim. Biophys. Acta* **1395**: 288–92.
- Kawasaki, S., Mizuguchi, K., Sato, M., Kono, T., and Shimizu, H.** (2013). A novel astaxanthin-binding photooxidative stress-inducible aqueous carotenoprotein from a eukaryotic microalga isolated from asphalt in midsummer. *Plant Cell Physiol.* **54**: 1027–40.
- Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E.** (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**: 845–858.
- Kühlbrandt, W., Wang, D.N., and Fujiyoshi, Y.** (1994). Atomic model of plant light-harvesting complex by electron crystallography. *Nature* **367**: 614–21.
- LeBaron, R.G., Bezverkov, K.I., Zimmer, M.P., Pavelec, R., Skonier, J., and Purchio, A.F.** (1995). Beta IG-H3, a novel secretory protein inducible by transforming growth factor-beta, is present in normal skin and promotes the adhesion and spreading of dermal fibroblasts in vitro. *J. Invest. Dermatol.* **104**: 844–9.
- Lepetit, B., Sturm, S., Rogato, A., Gruber, A., Sachse, M., Falciatore, A., Kroth, P.G., and Lavaud, J.** (2013). High light acclimation in the secondary plastids containing diatom *Phaeodactylum*

- tricornutum is triggered by the redox state of the plastoquinone pool. *Plant Physiol.* **161**: 853–65.
- Liu, Z., Yan, H., Wang, K., Kuang, T., Zhang, J., Gui, L., An, X., and Chang, W.** (2004). Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428**: 287–292.
- Marchler-Bauer, A. et al.** (2014). CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**: D222–6.
- Moody, R.G. and Williamson, M.P.** (2013). Structure and function of a bacterial Fasciclin I Domain Protein elucidates function of related cell adhesion proteins such as TGFBIp and periostin. *FEBS Open Bio* **3**: 71–7.
- Moreau, H. et al.** (2012). Gene functionalities and genome structure in *Bathycoccus prasinos* reflect cellular specializations at the base of the green lineage. *Genome Biol.* **13**: R74.
- Neilson, J.A.D. and Durnford, D.G.** (2010a). Evolutionary distribution of light-harvesting complex-like proteins in photosynthetic eukaryotes. *Genome / Natl. Res. Counc. Canada = Génome / Cons. Natl. Rech. Canada* **53**: 68–78.
- Neilson, J.A.D. and Durnford, D.G.** (2010b). Structural and functional diversification of the light-harvesting complexes in photosynthetic eukaryotes. *Photosynth. Res.* **106**: 57–71.
- Palenik, B. et al.** (2007). The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci. U. S. A.* **104**: 7705–10.
- Peers, G., Truong, T.B., Ostendorf, E., Busch, A., Elrad, D., Grossman, A.R., Hippler, M., and Niyogi, K.K.** (2009). An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* **462**: 518–21.
- Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H.** (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**: 785–6.
- Radakovits, R., Jinkerson, R.E., Fuerstenberg, S.I., Tae, H., Settlage, R.E., Boore, J.L., and Posewitz, M.C.** (2012). Draft genome sequence and genetic transformation of the oleaginous alga *Nannochloropsis gaditana*. *Nat. Commun.* **3**: 686.
- Robert, X. and Gouet, P.** (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**: W320–4.
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A., and Kumar, S.** (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* **30**: 2725–9.
- Worden, A.Z. et al.** (2009). Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* **324**: 268–72.





# CHAPTER VIII

## A Novel Micro-Photobioreactor Design for Assessing Microalgae Response to Light Intensity: the Case of *Nannochloropsis gaditana*

### Author names and affiliations

Giorgio Perin<sup>1,\*</sup>, Elisa Cimetta<sup>2,\*</sup>, Federico Monetti<sup>3</sup>, Tomas Morosinotto<sup>1</sup>, Fabrizio Bezzo<sup>3,4</sup>

\* Equal contribution

<sup>1</sup>PAR-Lab (Padua Algae Research Laboratory), Department of Biology, University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy

<sup>2</sup>Columbia University, 116<sup>th</sup> St & Broadway, New York, NY 10027, United States

<sup>3</sup>CAPE-Lab (Computer Aided Process Engineering Laboratory) and PAR-Lab (Padua Algae Research Laboratory), Department of Industrial Engineering, University of Padova, via Marzolo 9, 35121 Padova, Italy

<sup>4</sup>Centro studi di economia e tecnica dell'energia "Giorgio Levi Cases", University of Padova, via Marzolo 9, 35131, Padova, Italy

**Abstract.**

Microalgae are rising as promising feedstock for the sustainable production of biofuels and many other bio-commodities. Despite their established potential, the development of an economically competitive industrial process based on microalgae biomass is still far from being achieved. The identification of the best microalgae candidate as well as the optimization of the impact that several environmental factors have on microalgae growth still requires intense investigation. Since algae productivity is the result of the synergic influence of many parameters, mathematical models can be highly helpful for the optimization of the industrial cultivation processes. However, these models need to be fed with a large set of biological data in order to be reliable in their prediction. In this scenario, microscale technologies are emerging as a valuable tool to speed up the characterization process and improve data production, while maintaining high reliability standards. In this field, microalgae show additional challenges, for the necessity of dealing with light as a source of energy to support cells metabolism. Nevertheless, currently developed platforms are very complicated, often requiring a deep experience in the microfluidic field to monitor the device functionality. To overcome this drawback while keeping microscale benefits, here we describe the development of a simple micro-photobioreactor coupled with a quick and reliable growth evaluation method. The system is shown to be able to monitor impact of different light intensities on *Nannochloropsis gaditana* growth, also using fluorescence measurements to extract data on the cells photosynthetic functionality.

## **Introduction.**

The world population rising of the past decades, and the consequent increase in the energy demand to support modern life, led to a rise in fossil fuels exploitation, as the current primary energy source (Walsh et al., 2015; Moss et al., 2010). This phenomenon finds direct correlation with the net greenhouse gases increase in the atmosphere, which is posing serious threats for the future well-being of the society (Kazamia and Smith, 2014; Levy and Patz). Therefore, the development of renewable feedstock for the sustainable production of fuels and other oil-derived commodities is a strong priority. In this scenario, microalgae are emerging as promising candidate thanks to their ability to accumulate at high levels molecules such as sugars and lipids that can be easily converted into biofuels. Furthermore, they use sunlight and CO<sub>2</sub> as energy and carbon source, and they can thus be at the base of strongly environmental-friendly processes (Gimpel et al., 2015; Gangl et al., 2015; Hariskos and Posten, 2014). However, the realization of such potential still requires large research efforts, for instance concerning the identification of the best candidate species, having the highest biomass or lipid productivity. Despite the estimated number of microalgae species in nature is between 40,000 and 70,000 (Appeltans et al., 2012), the number of those extensively characterized and exploited for industrial purposes to date is extremely low (less than 10 (Lebeau and Robert, 2003)), pushing for the need of bioprospecting efforts (Kiuru et al., 2014; Barra et al., 2014). In fact a huge variety of species still needs to be explored (Sunagawa et al., 2015), looking for a potential better candidate for industrial purposes and that will require efficient experimental approaches to sort out potentially promising candidates.

Another major field of improvement is the exploration of growth-affecting parameters. Microalgae growth is in fact influenced by several parameters such as CO<sub>2</sub>, O<sub>2</sub> and nutrients availability (NH<sub>4</sub><sup>+</sup>, NO<sub>3</sub>, PO<sub>4</sub>) but also light intensity, quality and dynamic changes (Wobbe and Remacle, 2014; Wobbe et al., 2015; Carvalho and Malcata, 2005; Simionato et al., 2013b; Sforza et al., 2012a, 2012b; Cembella et al., 1984). The presence of a photosynthetic metabolism in fact strongly increases the complexity of the system with respect to other microorganism. Therefore, a detailed knowledge of the factors influencing microalgae growth is seminal for the optimization of productivity for any industrial application (Camacho-Rodríguez et al., 2013, 2015; Hariskos and Posten, 2014).

Since the system is highly complex, mathematical prediction models are potentially highly useful to quantify the impact of these parameters on microalgae growth (Zhang et al., 2015; Bernardi et al., 2014). The main purpose is to predict microalgae behavior in industrial cultivation systems which

show a growth environments that strongly differ from those at the lab-scale. On the other side, the synergic effect of the parameters can have multiple effects, difficult to be described, pushing for the test of multiple combinations. Using conventional laboratory equipment these kind of experiments takes indeed huge time and resources, limiting the amount of conditions testable.

In this scenario, microscale technologies have the potential to provide a huge step forward in understanding microorganisms biology. Their purpose is to scale-down the cultivation system, improving the high-throughput quality and reliability of data, with the final aim to cut down the time and costs required to evaluate the effect of several parameters on microorganisms growth. Despite mainly developed for human cells biological investigations (Villa-Diaz et al., 2009; Faley et al., 2008; Luni et al., 2016), recently the application of this technology also emerged in other biological fields, from eukaryotic microorganisms (Au et al., 2011) to plant cells cultivation (Grossmann et al., 2011; Sanati Nezhad, 2014).

Moreover, several microfluidic platforms are being developed in the past years, in order to simultaneously assess the impact of multiple parameters on microalgae growth (Holcomb et al., 2011; Juang and Chang, 2016; Zheng et al., 2013). The major established technology relies on static microdroplets arrays, produced to develop a finely tuned chemical environment in which microalgae growth takes place (Abalde-Cela et al., 2015; Pan et al., 2011; Dewan et al., 2012). However, all these microdroplets-based technologies don't represent dynamic platforms. The latter indeed often require to change the design of the cultivation system to fit the need to investigate the biological impact of other environmental parameters. It is indeed rare the application of the same device to respond to different biological investigations, and the few cases reported are not based on this technology (Grossmann et al., 2011, 2012; Lanquar et al., 2014).

Moreover, the microfluidic platforms so far developed were designed to investigate mainly the impact of nutrients availability on growth, and very poor consideration was instead given to the major factor influencing photosynthetic organisms growth, the light. So far only (Kim et al., 2014) provided to the scientific community a microscale platform to assess light intensity and light / dark cycles impact on microalgae growth. However, their system completely lacks the further integrated investigation of the light effect on photosynthetic performances. Since photosynthetic organisms, algae rely on photosynthesis to support their metabolism, therefore the investigation of the impact of the light conditions on the photosynthetic performances is a prerequisite for their future industrial cultivation (Simionato et al., 2013a).

Furthermore, the main drawback of currently adopted microfluidic platforms is their high level of complexity, and the necessity of highly skilled users in order to be efficiently exploited (Shih et al., 2014; Luke et al., 2016). The set up described is often very complicated, requiring a deep experience in electronics together with computer programming to monitor the device functionality (Graham et al., 2015).

In order to scale-down the growth platform, while leaving unaltered the simplicity in its use, we developed a simple microfluidic device and a quick and reliable growth evaluation method, based on the *in vivo* chlorophyll fluorescence evaluation. The latter was also used for integrated analyses on the light-driven alterations in the photosynthetic performances, directly *in situ*. The developed system proved capable of speeding up the evaluation of how different light regimes influence *Nannochloropsis gaditana* growth, a species chosen as a model organism for industrial purposes. In the microfluidic system, growth rates were as high as in continuous cultures. In addition, we showed that the developed platform is capable of assessing the impact of different illumination conditions at a time, entirely exploiting the high-throughputness of a microfluidic device and therefore opening the gates to a deep evaluation of the molecular basis involved in the light response of this promising species.

## Materials and Methods.

### *Microfluidic device fabrication.*

The platform comprises a matrix of wells, each representing an independent experimental condition. For our future studies in which fluid flow will be applied, the microalgae in culture will be protected from hydrodynamic shear forces and exposed to stable concentration gradients generated by an integrated microfluidic platform with independent rows of wells and microchannels connected to two lateral flow channels. Each micro-photobioreactor yields 45 data points, each represented by the cell cultures inside the 4 mm diameter wells. The platform was designed with the aid of 3D CAD software. The mold was produced via stereolithography rapid prototyping, and the final device was replica-molded in poly(dimethylsiloxane) (PDMS). The pre-polymer Sylgard 184 (Dow Corning) was mixed with a cross-linker in a ratio of 10 : 1 w/w and then poured onto the patterned silicon wafer and degassed in a dessicator. After curing at 65 °C for 3 h, the device is ready for use. The micro-photobioreactor is optically transparent, for microscopy observation. The microbioreactor and the clamping system were steam sterilized (20 min at 121 °C) prior to use.

### *Preparation of *Nannochloropsis gaditana* cells.*

*Nannochloropsis gaditana* wild type cells (strain 849/5), from the Culture Collection of Algae and Protozoa (CCAP), were cultured in sterile F/2 media containing sea salts (32 g/L, Sigma Aldrich), 40 mM Tris-HCl (pH 8) and Guillard's (F/2) marine water enrichment solution (Sigma Aldrich). Cells were routinely kept in pre-culture in Erlenmeyer flasks with 6, 60 or 360  $\mu\text{mol photons} / (\text{m}^2 \text{ s})$  illumination and 100 rpm agitation at  $22 \pm 1$  °C in a growth chamber.

In the middle of the exponential growth, cells were counted using a cell counter (Cellometer Auto X4, Nexcelom Bioscience) and then centrifuged at 3500 g for 5 minutes to remove the exhausted F/2 medium. Cells were used to inoculate the micro-photobioreactor at multiple starting cellular concentrations in fresh F/2 medium. Illumination was always kept constant and intensities ranged from 6 to 360  $\mu\text{mol photons} / (\text{m}^2 \text{ s})$ . The different light intensities were achieved thanks to daylight fluorescent lamps or a LED Light Source SL 3500 (Photon Systems Instruments). No external CO<sub>2</sub> supply was provided to the Erlenmeyer flasks and thus carbon dioxide to support growth came directly from the atmosphere.

*Microalgae growth monitoring and photosynthetic parameters estimation in the micro-photobioreactor.*

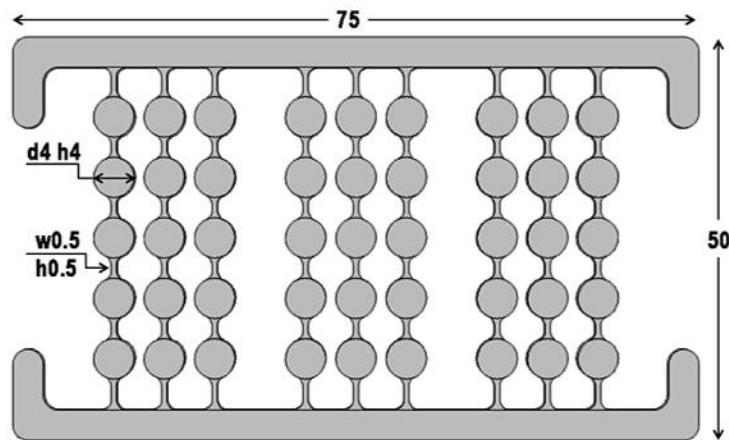
Microalgae growth in the micro-photobioreactor was monitored evaluating *in vivo* chlorophyll fluorescence using a FluorCam FC 800 video-imaging apparatus (Photon Systems Instruments, Brno, Czech Republic). After cells dark adaptation for 20 minutes, the instrument measured the fluorescence of dark-adapted cells ( $F_0$ ), and after a saturating light pulse with 3000  $\mu\text{mol photons} / (\text{m}^2 \text{ s})$  intensity and 800 ms duration, the maximum chlorophyll fluorescence value ( $F_m$ ) was evaluated. The latter was correlated to the cellular concentration after cells counting using a cell counter (Cellometer Auto X4, Nexcelom Bioscience). PSII functionality was expressed as PSII maximum quantum efficiency and was calculated as  $(F_m - F_0) / F_m$ . Cells were treated for 5 minutes with 500  $\mu\text{mol photons} / (\text{m}^2 \text{ s})$  light intensity to evaluate ETR and NPQ activation, then light was switched off to evaluate NPQ relaxation kinetic. NPQ and ETR values were calculated as previously described in (Maxwell and Johnson, 2000). The specific growth rate was calculated by the slope of logarithmic phase for the number of cells.



## Results and discussion.

### *Micro-photobioreactor array design*

In this paper, we investigated the growth of *N. gaditana* in a micro-photobioreactor, when exposed to different light intensities. The microfluidic platform included 45 wells (40  $\mu$ l of working volume each) linked in columns by a network of microfluidic channels (see figure 1 for dimensions details). This network was also connected to two lateral flow channels for the future fluid flow application. The design of the micro-photobioreactor followed specific criteria such as: (i) support of long-term culture of microalgae in a no-shear environment, (ii) high-throughput studies with biologically relevant numbers of replicates, (iii) compatibility with on-line imaging, (iv) ease of retrieval of the cell samples for additional post-processing, and (v), last but not least, capability to generate stable concentration gradients over the culture area in case fluid flow is applied (Somaweera et al., 2015).



**Figure 1. Microfluidic platform.** The micro-photobioreactor used comprises a matrix of 45 microwells (40  $\mu$ l), arranged in independent rows and connected to two lateral flow channels. The relevant dimensions (all in mm) are highlighted in the drawing. The corresponding mold was fabricated using standard stereolithography rapid prototyping techniques and used to obtain the final PDMS micro-photobioreactor thanks to replica molding.

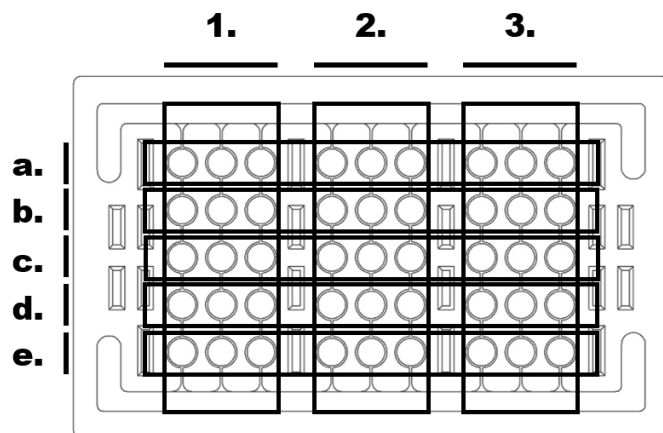
The device was fabricated in poly(dimethylsiloxane) (PDMS) which entirely fits the metabolic need of photosynthetic organisms, transparency and high gas permeability (Kuncová-Kallio and Kallio, 2006; McDonald and Whitesides, 2002). Both features are fundamental when dealing with photoautotrophs organisms such as *N. gaditana*, since the former is fundamental to feed the energy needs and the latter to feed the Calvin-Benson cycle, necessary to convert the light energy in carbon blocks used for

growth. However evaporation is a major threat when dealing with such minute volumes of samples (40  $\mu\text{l}$  per well). To avoid this problem we pre-saturated the PDMS with the same algae growth medium for 24 h, and covered the device with a thin ( $\sim 1$  mm) PDMS layer to turn the wells into real independent growth chambers.

#### *Growth monitoring based on in vivo Chl fluorescence*

To efficiently exploit the high-throughputness of a microscale platform for assessing the impact of several light conditions on growth simultaneously, we needed to be able to track microalgae growth over time. Therefore, our first goal was to build a simple and reliable method to evaluate *on site* the microalgae growth. The use of a live-imaging system concretely fits this purpose, since it allows the evaluation of the microalgae growth without perturbing the experimental conditions. Our device was however fabricated containing wells with 40  $\mu\text{l}$  working volume, which made impossible the application of optical microscope-based monitoring techniques, for the presence of multiple focal planes in each sample. Therefore, we developed a new method, based on the *in vivo* chlorophyll fluorescence evaluation, for monitoring quantitatively microalgae growth. In literature, chlorophyll fluorescence was already used for assessing microalgae growth, but never in a quantitative way (Bae et al., 2013).

The principle exploited in this work in fact relies on the direct proportionality between the chlorophyll content of a microalgae sample and the number of cells included, whether self-absorbance phenomena are avoided. Therefore, the first step concerned the building of a mathematical correlation between the Chl fluorescence values and the cells concentrations. To achieve this goal, cells were initially grown in bulk conditions as reported in the methods section. During the exponential growth, cells pre-cultured at 60  $\mu\text{mol photons} / (\text{m}^2 \text{s})$  were thus inoculated in the microfluidic device of figure 1, at 15 different starting concentrations, each in triplicate (see figure 2 for the experimental set up – data of the starting cells concentrations are reported in table 1).

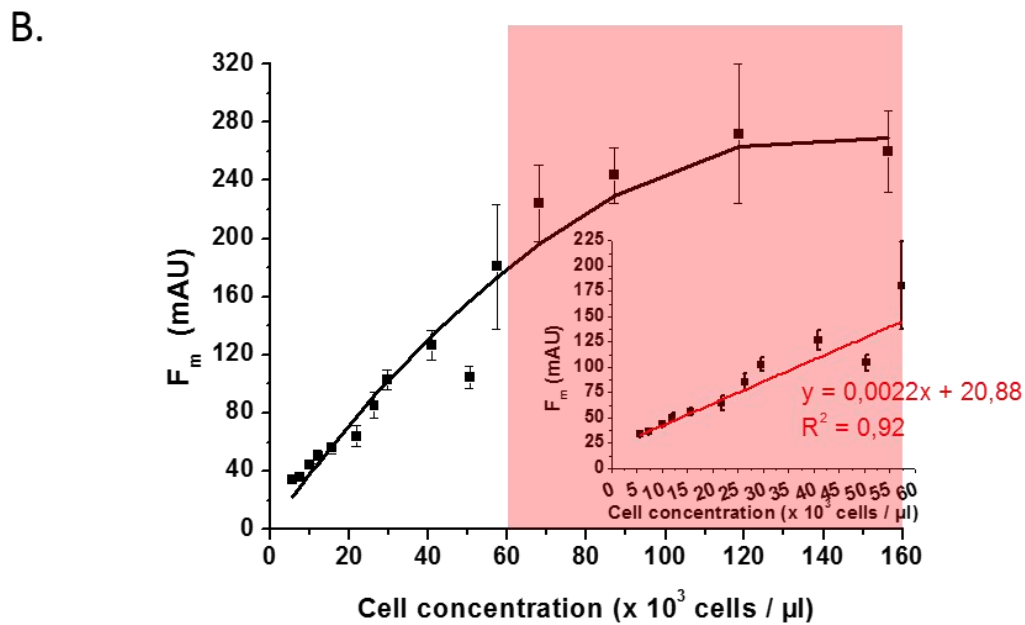
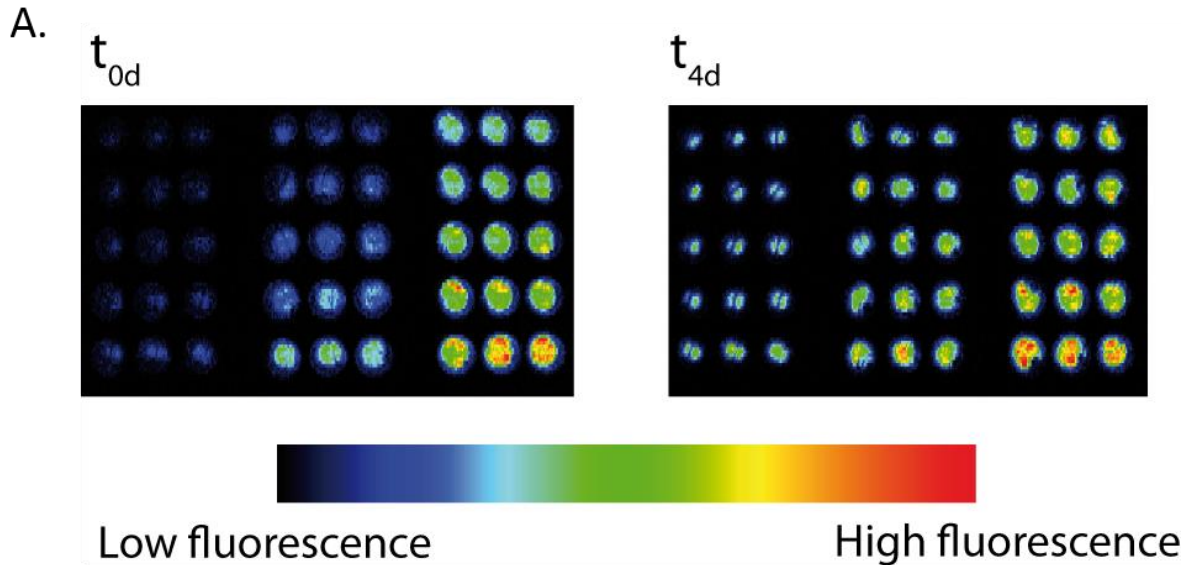


**Figure 2. Microwells classification in the microfluidic platform.** An identifier (ID) was associated to each subsection of the device using an alphanumeric code: alphabet letters for the rows (from letter a. for the first to letter e. for the fifth) and numbers for the 3 main growth sections (1., 2. and 3.), each including 15 microwells.

**Table 1. Cells concentration and  $F_m$  values used to build up the growth monitoring method.** The cells concentrations values used to inoculate the micro-photobioreactor are here reported (cells concentration<sub>i</sub>). These data refer to cells pre-cultured at 60  $\mu\text{mol photons} / (\text{m}^2 \text{s})$ . These values were chosen to cover a wide range of concentrations in the linearity region of the  $F_m$  – cells concentration correlation, before reaching the saturation points due to self-absorption phenomena. Except for the first, the cells concentration values were chosen with 30% increments. After cells sedimentation, cells were counted (cells concentration<sub>f</sub>) and  $F_m$  values were measured. SD comes from data collected from wells inoculated in triplicate. Microwells Identifiers (ID) come from the scheme of figure 2.

Microwell Identifier	Cells Concentration <sub>i</sub> x 10 <sup>3</sup> cells/ $\mu\text{l}$	Cells Concentration <sub>f</sub> x 10 <sup>3</sup> cells/ $\mu\text{l}$	$F_m$ mA.U.	$SD_{F_m}$ mA.U.
1.a.	3	5,6	33,8	1,7
1.b.	4	7,3	36,2	0,7
1.c.	5	9,9	44	1,6
1.d.	7	12	50	4,4
1.e.	9	15,6	55,9	4,1
2.a.	11	21,8	64,1	7,3
2.b.	15	26,3	85,3	8,6
2.c.	20	29,5	102,6	7
2.d.	25	40,8	126,6	10,1
2.e.	33	50,4	104,5	7,9
3.a.	43	57,3	180,6	43,1
3.b.	56	68	224,2	26
3.c.	73	87,1	243,5	19,2
3.d.	94	118,5	272	48
3.e.	123	156,5	260	28

Given the very low culture volume and the absence of active mixing, the microalgae showed in the first few hours some precipitation on the bottom of the wells, despite *N. gaditana* is a non-motile saltwater microalga species with a small diameter ( $\sim 3,5 \mu\text{m}$ ). We thus let the microalga settle for 20 h from the inoculum before evaluating the *in vivo* chlorophyll fluorescence for each well (see methods section). As shown in Figure 3A, the chlorophyll fluorescence image clearly highlights the 15 different starting cellular concentrations, inoculated in the microfluidic device.



**Figure 3. Growth monitoring method in the micro-photobioreactor device.** A. Images showing the chlorophyll maximum fluorescence value ( $F_m$ ) after cells exposure to a saturating light pulse (see methods). The image on the left was taken 20 h from the inoculum ( $t_{0d}$ ) and on the right after 96 h ( $t_{4d}$ ). The former was used to build the correlation  $F_m$  – cell concentration ( $\times 10^3$  cells/ $\mu$ l) represented in panel B. Data are expressed as average of 4 biological replicates  $\pm$  SD. Where the correlation is linear, data were fitted with a linear function ( $y = ax + b$ ) as shown in the magnification of B. The plot area, where the linearity between  $F_m$  and cellular concentration is lost, is highlighted in red.

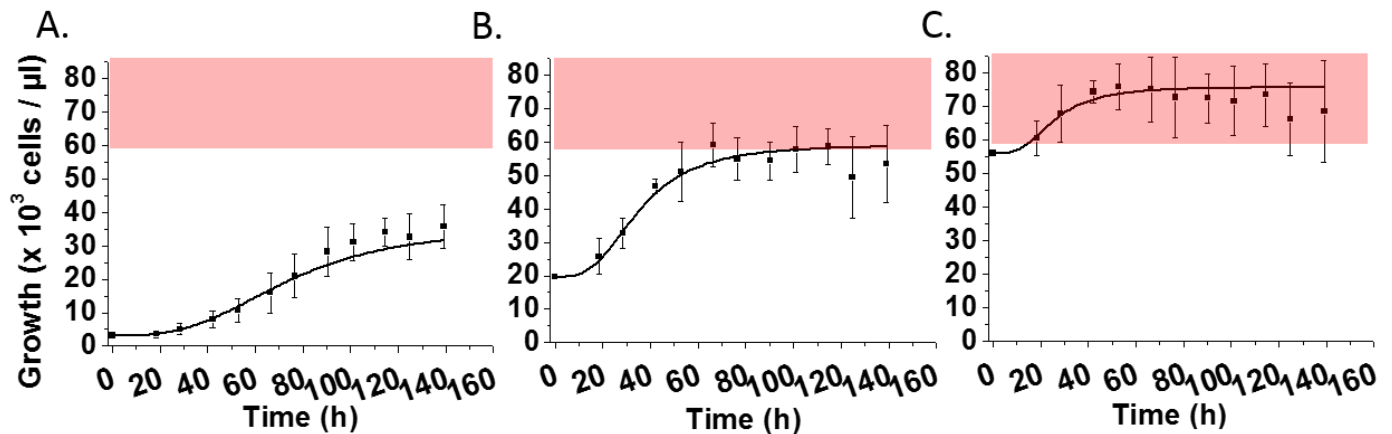
Cells in each well were then collected and counted (see methods section) in order to correlate the  $F_m$  values to the real cellular concentrations (expressed as cells/ $\mu$ l). We assume that the Chl content of a single cell oscillates around an average value over time -given the growth conditions are kept unaltered-, therefore when monitoring the chlorophyll fluorescence we can infer the corresponding value of cellular concentration. From the Chl fluorescence image of Figure 3A (first time point -  $t_{0d}$ ) and the cell counts, we built the mathematical correlation between the  $F_m$  values and the corresponding cellular concentration, for microalgae cells pre-cultured at 60  $\mu$ mol photons / ( $m^2$  s). As shown in Figure 3B, the linearity of the correlation between the two parameters is maintained up to a concentration of 60  $\times 10^3$  cells /  $\mu$ l; beyond that value, self-absorbance of emitted fluorescence starts playing a significant influence and the correlation deviates from linearity. Anyhow, within the linearity range, the correlation was satisfactory (see table 2 for the parameters values of the linear function for cells pre-cultured at 60  $\mu$ mol photons / ( $m^2$  s) – Medium light (ML)).

**Table 2. Linear function parameters values.** Parameters values for the linear functions ( $y = ax + b$ ), fitting the  $F_m$  – cells concentration correlations for the cells acclimated to 3 light intensities are here reported. The correlation for cells acclimated to the medium light intensity was obtained using the  $F_m$  and cells concentration values reported in Table 1.

Light Intensity	a	b
Low (LL)	0.0037	16
Medium (ML)	0.0022	20.88
High (HL)	0.0012	15

The microfluidic device was thus used to monitor *N. gaditana* cells growth when exposed to 60  $\mu$ mol photons / ( $m^2$  s), using fluorescence measurements and the correlation reported in Table 2 (ML parameters values) to calculate cellular concentrations. After 96 h of growth we obtained the Chl fluorescence image reported in Figure 3A (second time point –  $t_{4d}$ ). Based on the color scale of figure 3A, we can notice that the fluorescence intensity increased, demonstrating the microalgae growth in

all the inoculated wells. Figure 4 shows three growth curves derived for the wells inoculated with 3, 20 and 56 x 10<sup>3</sup> cells/μl (Figure 4A, B and C respectively); clearly in some cases cells growth pushed them out of the linearity range, impairing a reliable evaluation of the cells concentration. In some cases, however, as reported in figure 4A, microalgae growth curves remained in the linearity range. In other cases, as in figure 4B, 4 points of exponential growth are visible before the F<sub>m</sub> values reached saturation. Other cases, like example in figure 4C, were instead discarded because fluorescence was too high in all data points, impairing a reliable growth evaluation.



**Figure 3. Growth curves of *Nannochloropsis gaditana* cells in the micro-photobioreactor.** Light intensity was set at 60 μmol photons / (m<sup>2</sup> s). *N. gaditana* cells were inoculated at three different initial cellular concentrations: 3 x 10<sup>3</sup> cells/μl (A), 20 x 10<sup>3</sup> cells/μl (B), 56 x 10<sup>3</sup> cells/μl (C). Cells concentration data are expressed as average of 4 biological replicates ± SD. Data were fitted with a logistic function. The plot area, where the linearity between F<sub>m</sub> and cellular concentration is lost, is highlighted in red.

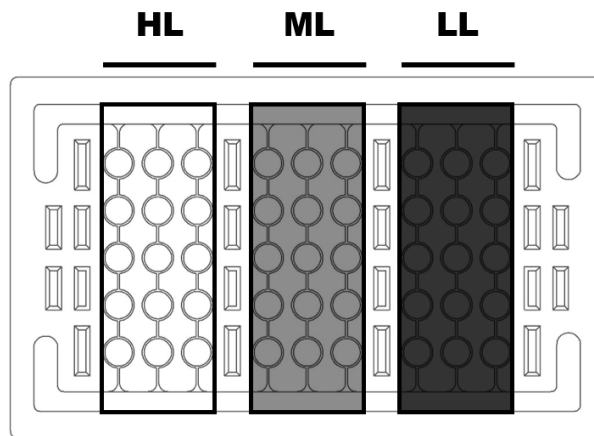
It was thus possible to calculate the specific growth rate (μ) of *N. gaditana* cells grown in the micro-photobioreactor, as in figure 4A (see methods). *N. gaditana* cells show a specific growth rate (μ) of 0.49 ± 0.08 (d<sup>-1</sup>), very close to that achievable for the same species in industrial-simulating cultivation systems (Table 3). These growth rates are obtained in cultures supplied with excess CO<sub>2</sub> and nutrients and thus in micro-PBR cells do not experience any significant nutrients or CO<sub>2</sub> limitations. Therefore, the applied micro-scaled platform revealed to be very useful for microalgae cultivation, supporting the original idea to be used to scale-down the cultivation process while reducing the time and costs needed for growth.

**Table 3. Specific growth rates in different cultivation systems.** Specific growth rates ( $\mu$ ) for *Nannochloropsis* cells, grown in different lab-scale or industrial cultivation systems, are here reported. (Y) stands for YES.

Microalgae species	Lab-scale	Simulating Industrial-scale	Light intensity	$\mu$	No Nutrients limitation	No CO <sub>2</sub> limitation	No Self-shading	Reference
	(Y)	(Y)	$\mu\text{mol photons m}^{-2} \text{s}^{-1}$	$\text{d}^{-1}$	(Y)	(Y)	(Y)	
<i>N. gaditana</i>	(Y)		100	0.2				(Simionato et al., 2011)
<i>N. salina</i>	(Y)		120	0.55	(Y)	(Y)	(Y)	(Sforza et al., 2012b)
<i>N. salina</i>		(Y)	150	0.92	(Y)	(Y)	(Y)	(Sforza et al., 2015)

*Nannochloropsis gaditana* response to light intensity.

Algae are photosynthetic organisms, therefore light has a huge impact on their overall metabolism (see chapter V), affecting photosynthetic performances and consequently their growth. Thus, in optimizing the cultivation conditions to achieve an optimal growth performance, it is seminal to assess the influence of this parameter. The set-up described above was useful to determine the proper operating conditions in order to assess *Nannochloropsis gaditana* response to light intensity. The idea is to improve the number of data gathered from a single experiment, while not affecting their reliability. Therefore the micro-photobioreactor was equipped with a light filter to precisely cut-off the light intensity to simultaneously get different irradiances on the same device (Figure 5).



**Figure 5. Simultaneous multi-light growth evaluation.** The micro-photobioreactor was covered with a custom made light-intensity controlling layer, able to precisely cut-off the light intensity to simultaneously get 6 (LL – low light), 60 (ML – medium light) and 360  $\mu\text{mol photons / (m}^2 \text{s)}$  (HL – high

light) on the same device. In the CAD drawing the device in the multi-light conformation is reported; 15 microwells are dedicated to 15 biological replicates for each light intensity.

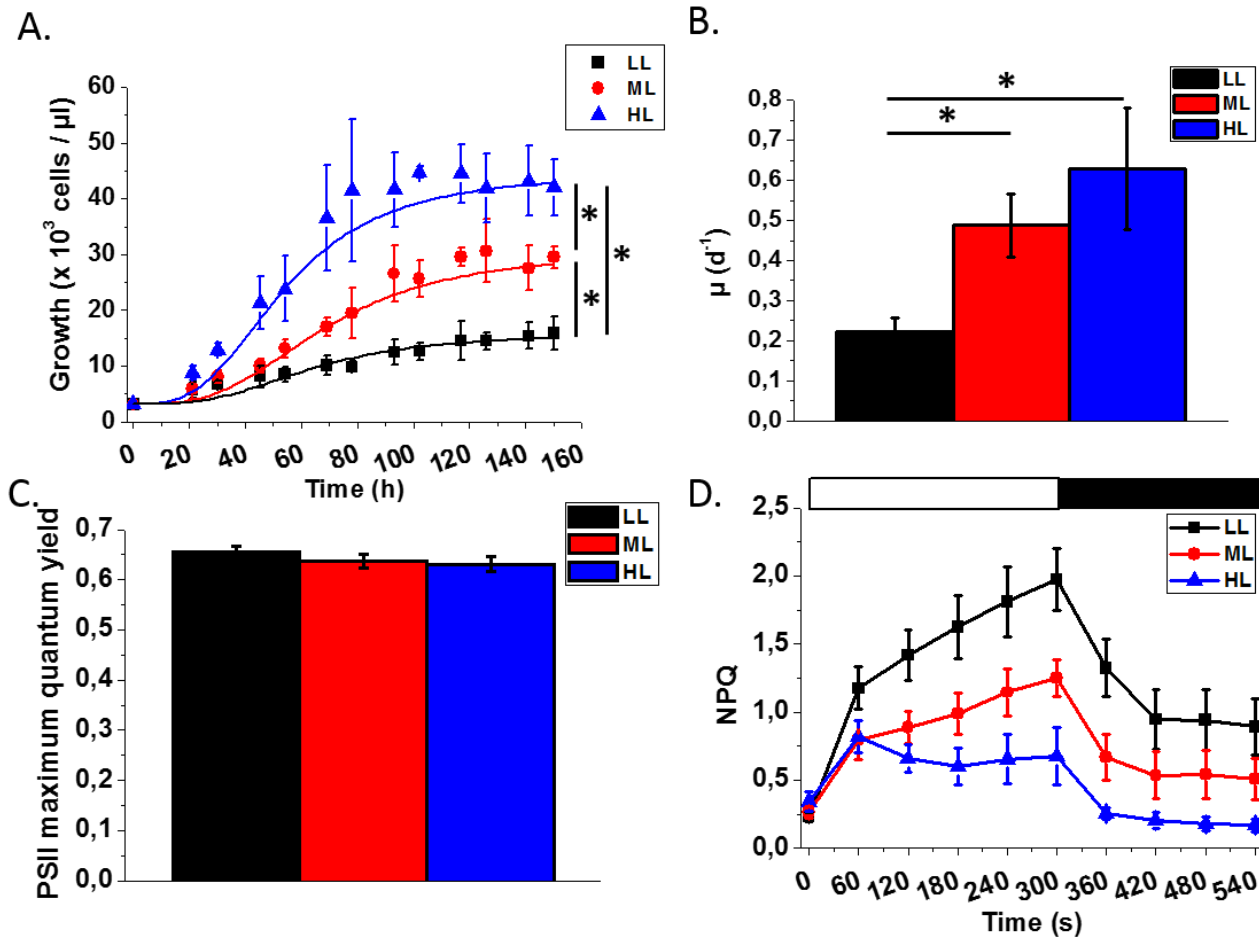
It was built overlaying some photo filters, with the final purpose to selectively lower the light intensity for specific micro-photobioreactor regions. The photo filters were properly chosen to cut between the 20 and the 80 % of the light intensity, while showing a constant absorption spectrum throughout the entire photosynthetically-active radiation (PAR) region of the electromagnetic spectrum. This design allowed obtaining three regions of the micro-PBR with 6 (LL – low light), 60 (ML – medium light) and 360 (HL – high light)  $\mu\text{mol photons} / (\text{m}^2 \text{ s})$  intensities. The micro-photobioreactor thus equipped (Figure 5) was inoculated with *N. gaditana* cells (starting cellular concentration:  $3 \times 10^3 \text{ cells} / \mu\text{l}$ ) after reaching the exponential growth phase in bulk conditions. The final configuration of the device allows following the microalgae growth, simultaneously exposed to 3 light intensities, with 15 biological replicates for each condition (Figure 5). Growth was monitored over time according to the  $F_m$ -cellular concentration correlation method described above.

Inoculated cells were pre-cultured at the same three different light intensities (low, medium and high light) and  $F_m$ -cellular concentration correlations were built also for cells pre-cultured at LL and HL conditions, calculating the parameters values reported in table 2. The latter shows indeed a reduction in the linear function slope, moving from LL to HL conditions, highlighting a change in the cells chlorophyll content with respect to the light intensity, a consideration already widely discussed for *Nannochloropsis* cells (Simionato et al., 2011).

The micro-PBR design and the growth monitoring method developed let us to obtain the growth curves reported in figure 6A. The data showed that light intensity has a huge impact on growth. Enhancing the light intensity, cultures reach higher cellular concentration in a proportional way. The cells grew on the device with different specific growth rates ( $\mu$  - Figure 6B) according to the light regimes to which they are exposed. Since they show increasing  $\mu$  values according to the light intensity increase, this confirms that cultures in the micro-PRB are not limited by the  $\text{CO}_2$  or nutrients availability. In fact, when light is provided in excess but cultures show  $\text{CO}_2$  limitation, growth rate always reaches a saturation threshold (Simionato et al., 2011), given by the inability of the Calvin-Benson cycle to stay ahead of the photosynthetic light energy capture. In this device,  $\text{CO}_2$  availability indeed feeds the Calvin-Benson cycle that can therefore build up photosynthates exploiting all the light energy provided to the cells. This means that the developed platform is very promising for integrated studies, in which the impact on growth of several environmental parameters could be assessed at a time. In fact, the basic biological investigations of the nutrients deprivation impact when



combined to light availability always require a lot of time, when conventional lab-scale devices are used. The developed micro-photobioreactor instead can answer to the request of speeding up this process.



**Figure 6. Simultaneous growth evaluation at three different light intensities.** A. Growth curves of *Nannochloropsis gaditana* cells at the three chosen light intensities. Data were fitted with a logistic function. B. Specific growth rates, expressed in in days<sup>-1</sup> (d<sup>-1</sup>), of *Nannochloropsis gaditana* cells grown at three different light intensities. C. and D. Photosynthetic parameters evaluation. C. Maximum quantum efficiency of PSII and D. NPQ kinetics. NPQ activation was monitored switching on the light at 500 μmol photons / (m<sup>2</sup> s). NPQ relaxation kinetic was evaluated switching off the light for 3 minutes. White box, light on; black box, light off. LL (low light – 6 μmol photons / (m<sup>2</sup> s), ML (medium light – 60 μmol photons / (m<sup>2</sup> s), HL (high light – 360 μmol photons / (m<sup>2</sup> s). Data are expressed as average of 3 biological replicates ± SD. Statistical significant differences are highlighted with an asterisk (One-way ANOVA, p -value < 0.05).

Although a number of microscale platforms were recently developed for microalgae growth characterization, none of them was never equipped with an efficient and reliable method for photosynthetic features estimation (Kim et al., 2014; Graham et al., 2015; Wang et al., 2013; Pan et

al., 2011). In fact so far chlorophyll fluorescence has been exploited just in two microfluidic devices, but, on one hand with the final aim of detecting microalgae cells viability (Wang et al., 2013), and on the other to verify microalgae growth, but just qualitatively (Bae et al., 2013).

Given the confirmed reliability of the micro-photobioreactor for the growth evaluation of photosynthetic microorganisms (Figure 6A), we decided to couple it with the samples *in vivo* Chl fluorescence monitoring, to also describe *on site* microalgae photosynthetic features when exposed to different light intensities. This may be an extremely valuable feature since it will allow to monitor not only growth but also to gather information on the photosynthetic efficiency of the cultures. Microalgae rely on the efficient light energy exploitation for supporting their metabolism, therefore understanding the impact that several environmental parameters have on photosynthesis is seminal to see the extent of microalgae industrial application. The efficiency of photosynthesis is indeed influenced by several environmental factors such as quality and quantity of light. We therefore focused on the latter and monitored PSII functionality and therefore the efficiency to exploit all the given light for growth by the cells. PSII functionality was described using the PSII maximum quantum yield and cells propensity to waste excess energy as heat was assessed evaluating non-photochemical quenching (NPQ) activation (see methods for details). Microalgae cells exposed to these 3 light intensities show a high PSII functionality (between 0.63 and 0.66), as shown in figure 6C. When light is provided in excess it can drive to the over-excitation of the photosystems and the formation of harmful reactive oxygen species (ROS), often leading to cells death. PSII functionality is considerably affected by excess light (Wobbe et al., 2015) and often it leads to its considerable decrease, due to photosynthesis saturation. In figure 6C instead the PSII maximum quantum yield of HL exposed cells is close to the one of LL and ML cells therefore highlighting that this light intensity is not in excess for cells when put in the micro-photobioreactor environment.

Photosynthetic organisms evolved NPQ mechanisms to alleviate such photo-oxidative stress (Erickson et al., 2015). Therefore, we monitored the NPQ activation kinetic of cells grown for 72 h at 3 different light intensities. We saw that cells grown in LL conditions, after switching on the light, activate NPQ much more rapidly than ML or HL cells. Cells grown in LL conditions reach a higher NPQ value than cells grown in ML, which in their turn show a higher NPQ than those grown in HL (NPQ activation trend, LL>ML>HL). It is worth noting that NPQ activation was monitored after cells exposition to the same light intensity ( $500 \mu\text{mol photons} / (\text{m}^2 \text{ s})$ ), therefore cells were able to respond differently to the same light experienced, according to the light to whom they were exposed for growth. When light is switched off, NPQ for HL cells completely relax while ML and LL cells aren't able to return to the

ground state (NPQ relaxation trend, LL<ML<HL), highlighting a possible photo-inhibitory event, underwent during the measurement. The same photosynthetic features were already described for *Nannochloropsis* cells, exploiting lab-scale facilities (Meneghesso et al., unpublished). This is however the first time that the monitoring of these parameters is performed in such microscale environment and this highlights its utility as integrated and high-throughput platform also for the biological investigation of light impact on photosynthesis.

## **Conclusion.**

In this work we have demonstrated that *Nannochloropsis gaditana* is a versatile microalgae species. For the first time it was efficiently cultured in a micro-photobioreactor, developing a simple and reliable micro-scale cultivation platform. A growth monitoring method based on *in vivo* Chl fluorescence evaluation was developed for the first time, to reliably follow its growth over time, and calculate its specific growth rates ( $\mu$ ). In the micro-photobioreactor, *N. gaditana* is not CO<sub>2</sub> limited therefore confirming its possible future exploitation for integrated biological studies. Growth monitoring in 3 different light regimes was used to assess the extent of the device applicability for the simultaneous evaluation of light intensity effects on growth. The specific growth rates estimation in function of light availability was completed in few days, with respect to the several weeks needed, exploiting lab-scale devices. Beyond the growth evaluation, we developed also a reliable method to assess light impact on photosynthetic performances over time, without any impairment of the growth micro-environments. Therefore given the benefits of the designed platform, in the near future it will be used for integrated biological studies in which nutrients availability together with the light impact on growth will be studied.

## References.

- Abalde-Cela, S., Gould, A., Liu, X., Kazamia, E., Smith, A.G., and Abell, C.** (2015). High-throughput detection of ethanol-producing cyanobacteria in a microdroplet platform. *J. R. Soc. Interface* **12**.
- Appeltans, W. et al.** (2012). The magnitude of global marine species diversity. *Curr. Biol.* **22**: 2189–202.
- Au, S.H., Shih, S.C.C., and Wheeler, A.R.** (2011). Integrated microbioreactor for culture and analysis of bacteria, algae and yeast. *Biomed. Microdevices* **13**: 41–50.
- Bae, S., Kim, C.W., Choi, J.S., Yang, J.-W., and Seo, T.S.** (2013). An integrated microfluidic device for the high-throughput screening of microalgal cell culture conditions that induce high growth rate and lipid content. *Anal. Bioanal. Chem.* **405**: 9365–74.
- Barra, L., Chandrasekaran, R., Corato, F., and Brunet, C.** (2014). The Challenge of Ecophysiological Biodiversity for Biotechnological Applications of Marine Microalgae. *Mar. Drugs* **12**: 1641–1675.
- Bernardi, A., Perin, G., Sforza, E., Galvanin, F., Morosinotto, T., and Bezzo, F.** (2014). An Identifiable State Model To Describe Light Intensity Influence on Microalgae Growth. *Ind. Eng. Chem. Res.* **53**: 6738–6749.
- Camacho-Rodríguez, J., Cerón-García, M.C., Fernández-Sevilla, J.M., and Molina-Grima, E.** (2015). Genetic algorithm for the medium optimization of the microalga *Nannochloropsis gaditana* cultured to aquaculture. *Bioresour. Technol.* **177**: 102–9.
- Camacho-Rodríguez, J., Cerón-García, M.C., González-López, C. V., Fernández-Sevilla, J.M., Contreras-Gómez, A., and Molina-Grima, E.** (2013). A low-cost culture medium for the production of *Nannochloropsis gaditana* biomass optimized for aquaculture. *Bioresour. Technol.* **144**: 57–66.
- Carvalho, A.P. and Malcata, F.X.** (2005). Optimization of  $\omega$ -3 fatty acid production by microalgae: crossover effects of CO<sub>2</sub> and light intensity under batch and continuous cultivation modes. *Mar. Biotechnol.* **7**: 381–388.
- Cembella, A.D., Antia, N.J., and Harrison, P.J.** (1984). The utilization of inorganic and organic phosphorous compounds as nutrients by eukaryotic microalgae: a multidisciplinary perspective: part 1. *Crit. Rev. Microbiol.* **10**: 317–91.
- Dewan, A., Kim, J., McLean, R.H., Vanapalli, S.A., and Karim, M.N.** (2012). Growth kinetics of microalgae in microfluidic static droplet arrays. *Biotechnol. Bioeng.* **109**: 2987–96.
- Erickson, E., Wakao, S., and Niyogi, K.K.** (2015). Light stress and photoprotection in *Chlamydomonas*

reinhardtii. *Plant J.* **82**: n/a–n/a.

- Faley, S., Seale, K., Hughey, J., Schaffer, D.K., VanCompernelle, S., McKinney, B., Baudenbacher, F., Unutmaz, D., and Wikswo, J.P.** (2008). Microfluidic platform for real-time signaling analysis of multiple single T cells in parallel. *Lab Chip* **8**: 1700–12.
- Gangl, D., Zedler, J.A.Z., Rajakumar, P.D., Martinez, E.M.R., Riseley, A., Włodarczyk, A., Purton, S., Sakuragi, Y., Howe, C.J., Jensen, P.E., and Robinson, C.** (2015). Biotechnological exploitation of microalgae. *J. Exp. Bot.*
- Gimpel, J.A., Henríquez, V., and Mayfield, S.P.** (2015). In *Metabolic Engineering of Eukaryotic Microalgae: Potential and Challenges Come with Great Diversity*. *Front. Microbiol.* **6**: 1376.
- Graham, P.J., Riordon, J., and Sinton, D.** (2015). Microalgae on display: a microfluidic pixel-based irradiance assay for photosynthetic growth. *Lab Chip* **15**: 3116–24.
- Grossmann, G., Guo, W.-J., Ehrhardt, D.W., Frommer, W.B., Sit, R. V, Quake, S.R., and Meier, M.** (2011). The RootChip: an integrated microfluidic chip for plant science. *Plant Cell* **23**: 4234–40.
- Grossmann, G., Meier, M., Cartwright, H.N., Sosso, D., Quake, S.R., Ehrhardt, D.W., and Frommer, W.B.** (2012). Time-lapse fluorescence imaging of *Arabidopsis* root growth with rapid manipulation of the root environment using the RootChip. *J. Vis. Exp.*
- Hariskos, I. and Posten, C.** (2014). Biorefinery of microalgae - opportunities and constraints for different production scenarios. *Biotechnol. J.* **9**: 739–52.
- Holcomb, R.E., Mason, L.J., Reardon, K.F., Cropek, D.M., and Henry, C.S.** (2011). Culturing and investigation of stress-induced lipid accumulation in microalgae using a microfluidic device. *Anal. Bioanal. Chem.* **400**: 245–53.
- Juang, Y.-J. and Chang, J.-S.** (2016). Applications of microfluidics in microalgae biotechnology: A review. *Biotechnol. J.*
- Kazamia, E. and Smith, A.G.** (2014). Assessing the environmental sustainability of biofuels. *Trends Plant Sci.* **19**: 615–8.
- Kim, H.S., Weiss, T.L., Thapa, H.R., Devarenne, T.P., and Han, A.** (2014). A microfluidic photobioreactor array demonstrating high-throughput screening for microalgal oil production. *Lab Chip* **14**: 1415–25.
- Kiuru, P., D’Auria, M.V., Muller, C.D., Tammela, P., Vuorela, H., and Yli-Kauhaluoma, J.** (2014). Exploring marine resources for bioactive compounds. *Planta Med.* **80**: 1234–46.
- Kuncová-Kallio, J. and Kallio, P.J.** (2006). PDMS and its suitability for analytical microfluidic devices. *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.* **1**:

2486–9.

- Lanquar, V., Grossmann, G., Vinkenborg, J.L., Merkx, M., Thomine, S., and Frommer, W.B.** (2014). Dynamic imaging of cytosolic zinc in Arabidopsis roots combining FRET sensors and RootChip technology. *New Phytol.* **202**: 198–208.
- Lebeau, T. and Robert, J.-M.** (2003). Diatom cultivation and biotechnologically relevant products. Part II: current and putative products. *Appl. Microbiol. Biotechnol.* **60**: 624–32.
- Levy, B.S. and Patz, J.A.** Climate Change, Human Rights, and Social Justice. *Ann. Glob. Heal.* **81**: 310–22.
- Luke, C.S., Selimkhanov, J., Baumgart, L., Cohen, S.E., Golden, S.S., Cookson, N.A., and Hasty, J.** (2016). A Microfluidic Platform for Long-Term Monitoring of Algae in a Dynamic Environment. *ACS Synth. Biol.* **5**: 8–14.
- Luni, C., Gagliano, O., and Elvassore, N.** (2016). Microfluidic technology enhances the potential of human pluripotent stem cells. *Biochem. Biophys. Res. Commun.*
- Maxwell, K. and Johnson, G.N.** (2000). Chlorophyll fluorescence - A practical guide. *J. Exp. Bot.* **51**: 659–668.
- McDonald, J.C. and Whitesides, G.M.** (2002). Poly(dimethylsiloxane) as a material for fabricating microfluidic devices. *Acc. Chem. Res.* **35**: 491–9.
- Moss, R.H. et al.** (2010). The next generation of scenarios for climate change research and assessment. *Nature* **463**: 747–56.
- Pan, J., Stephenson, A.L., Kazamia, E., Huck, W.T.S., Dennis, J.S., Smith, A.G., and Abell, C.** (2011). Quantitative tracking of the growth of individual algal cells in microdroplet compartments. *Integr. Biol. (Camb).* **3**: 1043–51.
- Sanati Nezhad, A.** (2014). Microfluidic platforms for plant cells studies. *Lab Chip* **14**: 3262–74.
- Sforza, E., Calvaruso, C., Meneghesso, A., Morosinotto, T., and Bertucco, A.** (2015). Effect of specific light supply rate on photosynthetic efficiency of *Nannochloropsis salina* in a continuous flat plate photobioreactor. *Appl. Microbiol. Biotechnol.* **99**: 8309–18.
- Sforza, E., Cipriani, R., Morosinotto, T., Bertucco, A., and Giacometti, G.M.** (2012a). Excess CO<sub>2</sub> supply inhibits mixotrophic growth of *Chlorella protothecoides* and *Nannochloropsis salina*. *Bioresour. Technol.* **104**: 523–9.
- Sforza, E., Simionato, D., Giacometti, G.M., Bertucco, A., and Morosinotto, T.** (2012b). Adjusted light and dark cycles can optimize photosynthetic efficiency in algae growing in photobioreactors. *PLoS One* **7**: e38975.

- Shih, S.C.C., Mufti, N.S., Chamberlain, M.D., Kim, J., and Wheeler, A.R.** (2014). A droplet-based screen for wavelength-dependent lipid production in algae. *Energy Environ. Sci.* **7**: 2366.
- Simionato, D., Basso, S., Giacometti, G.M., and Morosinotto, T.** (2013a). Optimization of light use efficiency for biofuels production in algae. *Biophys. Chem.*
- Simionato, D., Block, M.A., La Rocca, N., Jouhet, J., Maréchal, E., Finazzi, G., and Morosinotto, T.** (2013b). The response of *Nannochloropsis gaditana* to nitrogen starvation includes de novo biosynthesis of triacylglycerols, a decrease of chloroplast galactolipids, and reorganization of the photosynthetic apparatus. *Eukaryot. Cell* **12**: 665–76.
- Simionato, D., Sforza, E., Corteggiani Carpinelli, E., Bertucco, A., Giacometti, G.M., and Morosinotto, T.** (2011). Acclimation of *Nannochloropsis gaditana* to different illumination regimes: effects on lipids accumulation. *Bioresour. Technol.* **102**: 6026–32.
- Somaweera, H., Ibraguimov, A., and Pappas, D.** (2015). A Review of Chemical Gradient Systems for Cell Analysis. *Anal. Chim. Acta* **907**: 7–17.
- Sunagawa, S. et al.** (2015). Ocean plankton. Structure and function of the global ocean microbiome. *Science* **348**: 1261359.
- Villa-Diaz, L.G., Torisawa, Y., Uchida, T., Ding, J., Nogueira-de-Souza, N.C., O’Shea, K.S., Takayama, S., and Smith, G.D.** (2009). Microfluidic culture of single human embryonic stem cell colonies. *Lab Chip* **9**: 1749–55.
- Walsh, B.J., Rydzak, F., Palazzo, A., Kraxner, F., Herrero, M., Schenk, P.M., Ciais, P., Janssens, I.A., Peñuelas, J., Niederl-Schmidinger, A., and Obersteiner, M.** (2015). New feed sources key to ambitious climate targets. *Carbon Balance Manag.* **10**: 26.
- Wang, J., Sun, J., Song, Y., Xu, Y., Pan, X., Sun, Y., and Li, D.** (2013). A label-free microfluidic biosensor for activity detection of single microalgae cells based on chlorophyll fluorescence. *Sensors (Basel)*. **13**: 16075–89.
- Wobbe, L., Bassi, R., and Kruse, O.** (2015). Multi-Level Light Capture Control in Plants and Green Algae. *Trends Plant Sci.* **21**: 55–68.
- Wobbe, L. and Remacle, C.** (2014). Improving the sunlight-to-biomass conversion efficiency in microalgal biofactories. *J. Biotechnol.*
- Zhang, D., Chanona, E.A.D.-R., Vassiliadis, V.S., and Tamburic, B.** (2015). Analysis of green algal growth via dynamic model simulation and process optimization. *Biotechnol. Bioeng.* **112**: 2025–39.
- Zheng, G., Wang, Y., Wang, Z., Zhong, W., Wang, H., and Li, Y.** (2013). An integrated microfluidic

device in marine microalgae culture for toxicity screening application. *Mar. Pollut. Bull.* **72**: 231–43.





# APPENDIX I

## An Identifiable State Model to Describe Light Intensity Influence on Microalgae Growth

### Author names and affiliations

Andrea Bernardi<sup>1,3</sup>, Giorgio Perin<sup>2</sup>, Eleonora Sforza<sup>3</sup>, Federico Galvanin<sup>1</sup>, Tomas Morosinotto<sup>2</sup> and Fabrizio Bezzo<sup>1,3</sup>

<sup>1</sup>CAPE-Lab – Computer Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo 9, 35131 Padova, Padua, Italy

<sup>2</sup>PAR-Lab – Padova Algae Research Laboratory, Department of Biology, University of Padova, via U. Bassi 58 B, 35131 Padova, Padova, Italy

<sup>3</sup>PAR-Lab – Padova Algae Research Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo 9, 35131 Padova, Padua, Italy

**THIS CHAPTER WAS PUBLISHED IN INDUSTRIAL & ENGINEERING  
CHEMISTRY RESEARCH**

Bernardi A. et al. Ind Eng Chem Res. 2014 Apr 23; 53(16): 6738–6749 -

Doi: 10.1021/ie500523z



# An Identifiable State Model To Describe Light Intensity Influence on Microalgae Growth

A. Bernardi,<sup>†,§</sup> G. Perin,<sup>‡</sup> E. Sforza,<sup>§</sup> F. Galvanin,<sup>†</sup> T. Morosinotto,<sup>‡</sup> and F. Bezzo<sup>\*,†,§</sup>

<sup>†</sup>CAPE-Lab—Computer Aided Process Engineering Laboratory, Department of Industrial Engineering, University of Padova, via Marzolo 9, 35131 Padova, Padua, Italy

<sup>‡</sup>PAR-Lab—Padova Algae Research Laboratory, Department of Biology, University of Padova, via U. Bassi 58 B, 35131 Padova, Padova, Italy

<sup>§</sup>PAR-Lab—Padova Algae Research Laboratory, Department of Industrial Engineering, University of Padova via Marzolo 9, 35131 Padova, Padua, Italy

**ABSTRACT:** Despite the high potential as feedstock for the production of fuels and chemicals, the industrial cultivation of microalgae still exhibits many issues. Yield in microalgae cultivation systems is limited by the solar energy that can be harvested. The availability of reliable models representing key phenomena affecting algae growth may help designing and optimizing effective production systems at an industrial level. In this work the complex influence of different light regimes on seawater alga *Nannochloropsis salina* growth is represented by first principles models. Experimental data such as *in vivo* fluorescence measurements are employed to develop the model. The proposed model allows description of all growth curves and fluorescence data in a reliable way. The model structure is assessed and modified in order to guarantee the model identifiability and the estimation of its parametric set in a robust and reliable way.

## 1. INTRODUCTION

Microalgae-based processes are considered one of the most promising alternative technologies for the production of liquid fuels in the transport sector.<sup>1</sup> The main advantages of microalgae with respect to other possible feedstock are the high potential productivity and the absence of competition with traditional crops for arable land and clean water. However, this potential is still theoretical, and algae production on large scale is not profitable yet. Several issues need to be addressed to reach this objective, ranging from algae cultivation and harvesting as well as products extraction.<sup>1,2</sup>

The availability of reliable models would be greatly beneficial as the possibility to represent the fundamental physical, chemical, and biological phenomena would allow one to assess the interactions between equipment design and product yields and to scale-up and optimize the process design and operation.<sup>3</sup>

However, to be reliable a predictive model requires a specific set of parameters to be estimated in the most precise and accurate way. One important advantage of a model whose parameters have been reliably identified is that it can be used to predict the system response also in conditions (significantly) different from those tested during the identification experiments, which is the typical case when a model is exploited for process scale-up and optimization. To achieve that, however, model parameters must be identifiable, and this may present critical issues even in relatively simple models.<sup>4</sup>

In this work we will discuss a modeling structure to represent some key phenomena in algae growth by giving special attention to the identifiability of the proposed model. Algae growth is affected by several variables such as nutrient availability, temperature, and mixing, etc. However, being algae photosynthetic organisms, light is the key variable determining growth efficiency and kinetics: for this reason, the focus of this work will

be on the representation of its influence on growth. It is worth stressing that the correlation between light intensity and growth is a very complex one, and while low irradiation is limiting, its excess drives to the formation of reactive oxygen species and has an inhibitory effect.<sup>5</sup>

In the literature several models of photosynthetic biomass growth have been presented. It is possible to divide such models into two main groups: “physiological models” and “state models”. Physiological models attempt to describe the dynamic behavior of photosynthetic cells and propose approximations for the actual mechanisms involved in the cells’ growth. These models may try to represent the optimal allocation of energy and nutrients during cells’ activities (e.g., Ross and Geider<sup>6</sup>) or to represent a specific metabolic reaction (e.g., Marshall et al.,<sup>7</sup> where the damage and repair cycle of protein D1 is described). Usually these models are extremely detailed and involve a large amount of variables and parameters. The actual identification procedure may be extremely complex (sometimes even impossible) and require numerous, highly specific, and costly experiments.

State models are instead based on the concept of a photosynthetic unit (PSU) and are more instrumental for simulating and optimizing industrial cultivation systems. The PSU is defined as the sum of the light harvesting complex, the reaction center, and the associated apparatus, which are activated by a given amount of light energy to produce a certain amount of photoproduct.<sup>8</sup> These models are called state models because the PSU can be in different states of excitation. One of the first

**Received:** September 27, 2013

**Revised:** March 21, 2014

**Accepted:** March 23, 2014

**Published:** March 24, 2014

proposed state models was the model by Fasham and Platt,<sup>9</sup> whose objective was to describe photoinhibition. Several other modeling approaches have been proposed over the years.<sup>10–16</sup>

In this work we will consider the models by Camacho Rubio et al.<sup>10</sup> and by Eilers and Peeters (initially developed by Eilers and Peeters<sup>11</sup> and then improved by Wu and Merchuk<sup>16</sup>). These models are capable of representing the key phenomena of interest in this work, and they are reasonably simple so as to limit possible identifiability issues. In the Eilers and Peeters model, the authors assume that if an activated PSU absorbs an additional photon, it may become inhibited. For this reason they assume the rate of photoinhibition to be proportional to light intensity. It is also assumed that photosynthesis (and by consequence biomass growth) is proportional to the transition between the activated state and the resting state. Later Wu and Merchuk<sup>16</sup> modify the model of Eilers and Peeters, introducing a constant maintenance factor in the description of biomass growth.

In the model by Camacho Rubio et al.<sup>10</sup> both photo-inhibition and photoacclimation, which represent the cells' response to optimize their photosynthetic apparatus to different light intensities, are considered (note that in the original work, and in several others, photoacclimation is instead called photo-adaptation: photoacclimation is however a more accurate definition: in fact, adaptation refers to the organisms' modification during evolution to their environment, thus responses with time scales extremely longer than the ones considered here). Photoacclimation was at first represented as a steady state process but more recently extended by the same authors in order to represent its dynamics, together with the effect of nonphotochemical quenching (NPQ) and dark respiration (Camacho Rubio et al.<sup>12</sup>). This last model is indeed very flexible, but the number of model parameters to be estimated is very high and their precise identification may become a long and difficult task, especially if a limited amount of data is available, and for this reason it will not be discussed further in this work.

For model development and identification we consider experimental data referring to a particular species of microalgae of industrial interest (*Nannochloropsis salina* (*N. salina*)) grown in nonlimiting nutrients conditions and in a flat-plate photobioreactor.<sup>17</sup> These data sets were selected because experimental conditions were optimized to minimize all influences on algae growth other than light intensity. In fact, nutrients and CO<sub>2</sub> were provided in excess, but also the photobioreactor light path was minimized to reduce as much as possible light attenuation due to cells' shading and scattering. Sforza et al.<sup>17</sup> demonstrated that this assumption was an acceptable approximation, especially considering that we are interested in representing the exponential growth phase, where nutrient availability is high and cell concentration low. Accordingly, these data represent an accurate description of the influence of light alone on algae growth, minimizing the effect of other parameters. It is worth clarifying that other phenomena, which also play a major influence on algae productivity in industrial photobioreactors, such as the dark/light cycles due to mixing are not considered in this model.

There are two main contributions in the present work. On the methodological side, a step by step approach will be presented and applied to guarantee the identifiability of the growth model. Second, the utilization of both biomass concentration (growth curves) and multiple fluorescence measurements will allow shading light on some fundamental phenomena in the

correlation between illumination and growth; in particular, differently from previous contributions, measurements of the light profile of PSU saturation will be exploited.

The work is organized as follows: after outlining the general identification methodology, the available experimental data and the two candidate models will be introduced and discussed. The successive section is about model discrimination and the enhancement of the selected model. Then, an identifiability analysis and a reparameterization approach will allow setting up an identifiable model. The performance of the model in describing algal growth will be critically discussed. Some final remarks will conclude the work.

## 2. MODEL DEVELOPING APPROACH

Identifiability is a key issue to guarantee reliability and predictive capability in a model being developed. Figure 1 outlines the basic tasks and information flux required to achieve such a target. The preliminary step is to identify the phenomena that need describing and the fundamental mathematical laws that should be implemented to represent them. Here we assume that some modeling assumptions are already available. In other cases, preliminary experimental data may be needed to envisage the correlation among data and to set up a suitable physical interpretation through a mathematical model.

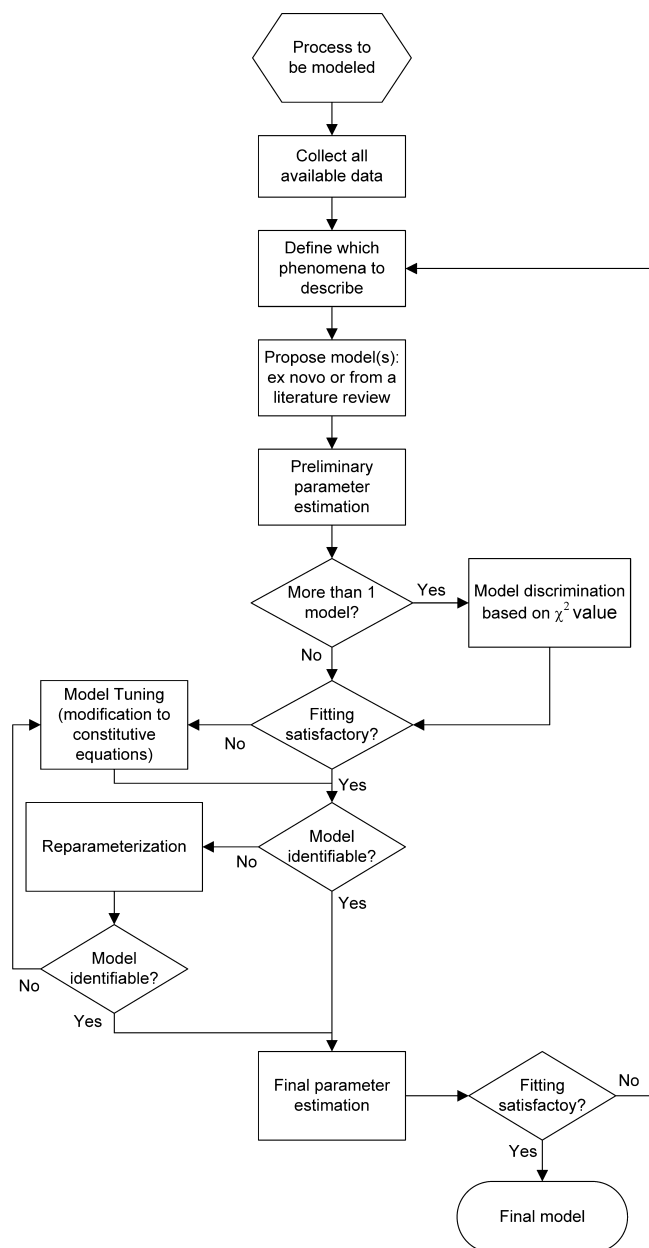
Typically, some available data may be exploited at this stage to choose among competitive modeling approaches through suitable discrimination techniques.<sup>18,19</sup> At least in the easiest cases a  $\chi^2$  test on experimental data may be sufficient to make the discrimination.<sup>20,21,22</sup> *Ad hoc* experiments can also specifically be designed to allow for a more effective and reliable discrimination among different candidates.<sup>21–23</sup> Once a suitable candidate model has been selected (and a preliminary estimation of its parameter has been carried out), the model may need upgrading to improve its capability of representing the phenomena being investigated (new experiments may be needed and possibly designed, and an estimation of all model parameters should be attempted).

Then it is extremely important to verify the model identifiability, i.e., to confirm that the optimal set of parameters values is unique and that their values can be determined in a precise way (and ideally in a physically meaningful way). If some identifiability issues arise, then a first approach to tackle the problem is to reparameterize the model.<sup>24</sup> If this is not enough, then the model structure should be modified.

Once the model is proved to be identifiable, the final parameter estimation can be performed. Note that even when a model is identifiable, measurements of noise and other uncertainty effects may still hinder its practical identifiability, although properly designed experiments may help in tackling the issue.<sup>25</sup> Once the final parameters estimation has been performed, new data, not involved in model calibration, should be used to validate the model. This approach has been applied to the specific case study and is discussed in the following sections. For simulation and parameter estimation purposes, the gPROMS software has been used.<sup>26</sup>

## 3. MODELING APPROACHES

The main advantage of state models is that they reduce the complexity of photosynthesis into a few possible states of the PSUs. This simple structure is also particularly effective in the use of fluorescence measurements, which can be exploited to monitor the PSU populations at different states. A PSU is



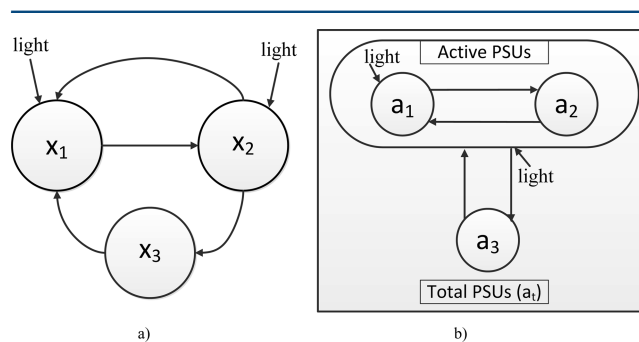
**Figure 1.** Information flux of the model identification procedure.

defined as the sum of the light harvesting complex, the reaction center, and the associated apparatus, which are activated by a given amount of light energy to produce certain quantities of photoproducts. The Eilers and Peeters model in the form proposed by Wu and Merchuck<sup>16</sup> (afterward denoted as EPM) and the Camacho Rubio model (Camacho Rubio et al.,<sup>10</sup> later called CRM) are two of the simplest models that can describe photosynthetic biomass growth as a function of light intensity. Both models consider that a PSU can assume three different states of excitation: (1) the resting (or open) state, which is the state of PSU before the light energy excites the reaction center; (2) the activated (or closed) state, that is the state of PSU excited by light energy; (3) the inhibited state, which is the state of PSU damaged by an excess of light energy. Both models do not consider any limitation on nutrients availability or mass transport of nutrients: i.e., only the exponential growth phase is described. CRM considers photoacclimation too, but without any representation of its dynamics. This is

a reasonable assumption also in our case study, since acclimation characteristic time scale (hours or days) is significantly larger than the time scales of the other phenomena being investigated.

In the work of Wu and Merchuck<sup>16</sup> EPM was used to fit the data of experiments carried out in a thin tubular loop reactor. Part of the reactor was kept in dark to simulate mixing; light intensities used for the experiments were 110, 220, and 550  $\mu\text{E}/(\text{m}^2 \text{ s})$ . The measurements used to calibrate the model were both biomass concentration and dark fluorescence measurements (fluorescence measurements will be described in section 4). Each experiment was carried out for 48 h, and measurements were taken every 12 h. In the work of Camacho Rubio et al.,<sup>10</sup> CRM was applied to a wider range of light intensities (ranging from 0 to 2000  $\mu\text{E}/(\text{m}^2 \text{ s})$ ) and to different light regimes (constant light, flashing light, and day–night cycle). Data used by the authors were growth rate constant and  $P$ – $I$  curves taken from the literature.

EPM assumes that the number of PSUs is constant with respect to light intensity and accordingly refers to the PSU  $x_1$ ,  $x_2$ , and  $x_3$  to represent the resting, activated, and inhibited states, respectively. Conversely, CRM assumes that the number of PSUs is a function of light intensity (indicated as  $a_i$ ), and the model equations are expressed as a function of the amount of the PSUs in each of the three states ( $a_1$ ,  $a_2$ , and  $a_3$ ). Parts a and b of Figure 2 illustrate the two model structures.



**Figure 2.** (a) Scheme of EPM.  $x_1$ ,  $x_2$ , and  $x_3$  represent the fractions of PSUs in resting, activated, and inhibited states, respectively. (b) Scheme of CRM.  $a_1$ ,  $a_2$ , and  $a_3$  are the numbers of PSUs in resting, activated, and inhibited states, respectively;  $a_i$  represents the total number of PSUs.

**3.1. Eilers–Peeters model.** In Figure 2a, the three PSU states are represented by circles and the possible state transitions are represented by the arrows. The resting state PSU can capture light energy and transfer it to an activated state. The PSU in the activated state can be damaged by light, or pass down the energy to start the dark phase of photosynthesis (and then return to a resting state). An inhibited PSU can be recovered and then return to the resting state. The reaction rate of the transitions involving the absorption of light (i.e.,  $x_1 \rightarrow x_2$  and  $x_2 \rightarrow x_3$ ) is assumed to be first order with respect to light intensity. The other two transitions are assumed to be zero order with respect to light intensity. Each transition is assumed to be first order with respect to the PSU fraction involved in the transition. The growth rate constant ( $\mu^{\text{EP}}$  ( $\text{h}^{-1}$ )) is assumed to be proportional to the state transition from activated to resting state, representing the photochemical reactions. Considering that the growth rate can be negative in the dark or at very low light intensity, a constant

maintenance ( $M^{EP}$  ( $h^{-1}$ )) factor is introduced. The model equations are as follows:

$$\frac{dx_1}{dt} = -k_a^{EP} I x_1 + k_d^{EP} x_2 + k_r^{EP} x_3 \quad (1)$$

$$\frac{dx_2}{dt} = k_a^{EP} I x_1 - k_d^{EP} x_2 - k_i^{EP} I x_2 \quad (2)$$

$$x_1 + x_2 + x_3 = 1 \quad (3)$$

$$\mu^{EP} = k_p^{EP} k_d^{EP} x_2 - M^{EP} \quad (4)$$

The set of parameters (whose physical meaning is summarized in Table 1) is represented by vector  $\hat{\theta}^{EP} = [k_a^{EP}, k_d^{EP}, k_i^{EP}, k_r^{EP}, k_p^{EP}, M^{EP}]$ .

**Table 1. Parameters of EPM Significance and Units**

param	significance	units
$k_a^{EP}$	kinetic constant of the activation reaction rate	$m^2/\mu E$
$k_d^{EP}$	kinetic constant of the deactivation reaction rate (photochemical quenching)	$s^{-1}$
$k_i^{EP}$	kinetic constant of inhibition reaction rate	$m^2/\mu E$
$k_r^{EP}$	kinetic constant of the recovery reaction rate	$s^{-1}$
$k_p^{EP}$	proportionality factor between photochemical quenching and biomass growth rate constant	s/h
$M_{EP}$	maintenance factor	$h^{-1}$

**3.2. Camacho Rubio Model.** In CRM, the photoinhibition rate is assumed to be proportional to the sum of resting state and activated PSUs, i.e., the active PSUs represented by  $a_1 + a_2$  [PSUs/cells] in Figure 2b. As in EPM, PSU activation reaction is assumed to be first order with respect to light intensity and to the amount of resting state PSUs. However, in CRM the transition from activated to resting state, related to biomass growth, is defined as a Michaelis–Menten kinetic, assuming an enzymatic reaction as the limiting step of this process. Also, from an analysis of experimental data, the authors<sup>10</sup> assume that the photoinhibition reaction rate is first order with respect to the square root of light intensity. As in EPM the recovery of damaged PSUs is assumed to be a first order reaction with respect to the number of damaged PSUs. Finally, as anticipated, CRM includes photoacclimation. The total amount of PSUs ( $a_t$  [PSUs/cells]) in CRM is thus assumed to be a hyperbolic decreasing function of light intensity. As for EPM, the growth rate constant ( $\mu^{CR}$  ( $h^{-1}$ )) is assumed to be proportional to the transition from activated to resting state and a constant maintenance factor ( $M^{CR}$  ( $h^{-1}$ )) is introduced. The model equations are as follows:

$$\frac{da_2}{dt} = k_a^{CR} I a_1 - \frac{r_m^{CR}}{K_S^{CR} + a_2} a_2 \quad (5)$$

$$\frac{da_3}{dt} = k_i^{CR} \sqrt{I} (a_1 + a_2) - k_r^{CR} a_3 \quad (6)$$

$$a_1 + a_2 + a_3 = a_t \quad (7)$$

$$a_t = \frac{r_m^{CR}}{k_c^{CR} + \frac{k_a^{CR} k_r^{CR}}{k_i^{CR}} \sqrt{I}} \quad (8)$$

$$\mu^{CR} = k_p^{CR} \frac{r_m^{CR}}{K_S^{CR} + a_2} a_2 - M^{CR} \quad (9)$$

As reported in the literature, it appears that the condition  $K_S^{CR} \gg a_2$  is true for the entire range of light intensities considered in the experiments (ranging from 50 to 1000  $\mu E/(m^2 s)$ ). Thus, the Michaelis–Menten kinetic may be well approximated by a first order kinetic as in EPM. Accordingly, the reduced set of parameters is represented by vector  $\hat{\theta}^{CR} = \{k_a^{CR}, k_d^{CR}, k_i^{CR}, k_r^{CR}, k_p^{CR}, k_c^{CR}, M^{CR}\}$  (parameters' physical meanings and units are reported in Table 2).

**Table 2. Parameters of CRM Significance and Units**

param	significance	units
$k_a^{CR}$	kinetic constant of the activation reaction rate	$m^2/\mu E$
$k_d^{CR}$	kinetic constant of the deactivation reaction rate (photochemical quenching)	$s^{-1}$
$k_i^{CR}$	kinetic constant of inhibition reaction rate	$m/(\mu E \cdot s)^{0.5}$
$k_r^{CR}$	kinetic constant of the recovery reaction rate	$s^{-1}$
$k_p^{CR}$	proportionality factor between photochemical quenching and biomass growth rate constant	s/h
$k_c^{CR}$	rate constant involved in the photoacclimation process	
$M^{CR}$	maintenance factor	$h^{-1}$

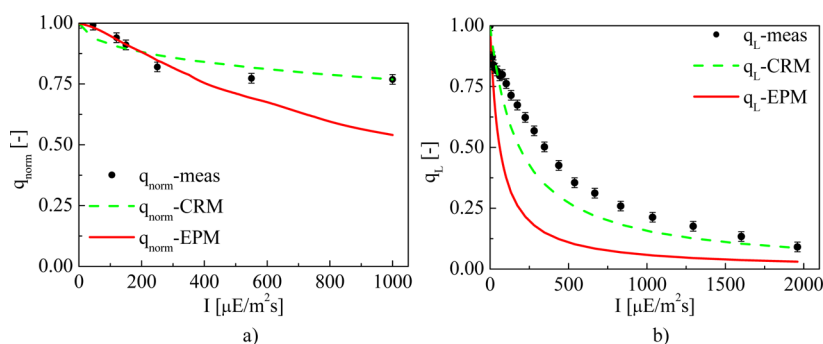
#### 4. EXPERIMENTAL SETUP AND AVAILABLE DATA

The aim of this work is to describe the growth of microalgae in nonlimiting nutrient conditions and according to the hypothesis that the light intensity is constant with respect to the culture time and depth. The fundamental phenomena to be described are as follows: (i) reaction centers oxidation/reduction cycle, to represent the photosynthesis, and (ii) the damaging effect of excess light on PSUs (photoinhibition).

Our data refer to algal cultures grown at different light intensities.<sup>17</sup> During the experiments, microalgae were acclimated to the light used and to the geometry of the photobioreactor. Each experiment was conducted in parallel at least twice and in two identical photobioreactors, in order to ensure its reproducibility.<sup>17</sup> Only the data in the exponential phase of the original growth curves were used for the parameters estimation, since our model represents only exponential growth and does not consider nutrients limitation. Experimental setup was built to limit as much as possible cells shading by decreasing as much as possible the light path and working at low cells concentration. This, together with the presence of nutrients and  $CO_2$  in non-limiting amounts, ensures that growth is dependent only from the light intensity reaching the culture.

Moreover, experimental measurements of fluorescence will be used as additional data for the parameters estimation.<sup>27</sup> These data are commonly available for the photosynthetic organisms and have been exploited to estimate photosynthetic efficiency in a large body of experimental literature (reviewed by Maxwell and Johnson<sup>27</sup>). In our case we considered in particular two fluorescence parameters, related to the oxidation state of the algal cells: the parameters  $q$  (or  $F_v/F_m$ ),<sup>28–31</sup> and  $q_L$ , whose physical meanings will be briefly detailed in the following.

Parameter  $q = F_v/F_m$  (with  $F_v = F_m - F_0$ ) was measured for each light intensity at which microalgae were grown.  $F_0$  represents the fluorescence in the dark when all reaction centers are open. Conversely, the maximum  $F_m$  is the fluorescence after a saturating light flash, when all reaction centers are closed. Here the fluorescence is higher than  $F_0$  because saturated PSUs are unable to perform photochemistry and therefore a larger amount of excitation energy is re-emitted as fluorescence.



**Figure 3.** Measurement (black circles) and predicted values of (a)  $q_{\text{norm}}$  and (b)  $q_L$ . Red solid lines represent the profiles according to EPM, while the dashed green lines represent the profiles according to CRM.

When PSUs are active, there is a large difference between  $F_m$  and  $F_0$  (and thus a larger  $F_v/F_m$ ), while if PSUs are photoinhibited,  $F_m$  is closer to  $F_0$ . For this reason, the  $F_v/F_m$  parameter is commonly used in the literature to quantify active PSUs *in vivo*. The precise value of  $F_v/F_m$  in the case of fully active PSUs is known to be variable between species, since it depends on specific properties such as the antenna size. Here it is set to be equal to 0.65 as this is a value commonly measured in several microalgae<sup>32</sup> and also in healthy *Nannochloropsis* cultures, exposed to low light.<sup>17,31</sup> In order to have a parameter bound between 0 and 1, it is possible to define the parameter  $q_{\text{norm}} = q/q_{\text{max}}$  where  $q_{\text{max}}$  is the value of  $F_v/F_m$  in the case of fully active PSUs.

In cultures exposed to different light intensities, a decrease in this value indicates the presence of photoinhibited PSUs, as normally experienced in high light conditions. For this reason, it can be used to estimate the content of photoinhibited PSUs and thus it can be correlated to  $x_3$  and  $a_3$  populations, using the definition of EPM and CRM, respectively.

While  $q$  is rather commonly employed in several similar studies,<sup>16</sup> in order to have a better representation of the oxidation state of the PSUs in illuminated cells, here we also included the fluorescence parameter  $q_L$ , which provides a linear estimation of the saturation level of PSU as discussed in detail in the work by Kramer et al.<sup>33</sup> This means that  $q_L$  is 1 when all active PSUs are open, while it decreases to 0 when all PSUs are closed and photosynthesis is saturated. Parameter  $q_L = [(F'_m - F'_s)/(F'_m - F'_0)](F'_0/F'_s)$  was measured for 21 different light intensities with a PAM fluorometer. Here  $F'_m$  and  $F'_0$  stands for the same minimal and maximal fluorescence, measured in light adapted cells, while  $F'_s$  stands for stationary fluorescence in illuminated cells. Thus,  $q_L$  can be exploited as an estimation of the relative ratio of  $x_1$  and  $x_2$  populations ( $a_1$  and  $a_2$  according to CRM). The complete set of experimental measurements used in this work is reported in Appendix A.

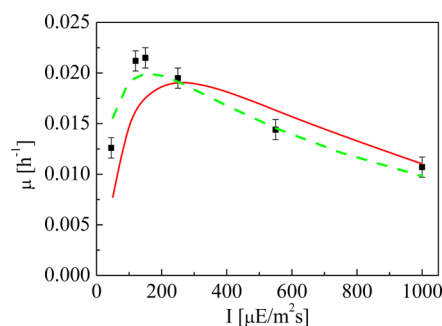
## 5. MODEL DISCRIMINATION AND PRELIMINARY PARAMETERS ESTIMATION

The first objective is to discriminate between the two alternative models so as to select the most suitable one to describe our system. Parameter estimations were performed for both models, based on the entire set of experimental data. The different performance is quantitatively summarized by the  $\chi^2$  test (in the case of EPM, we have  $\chi^2 = 890.1$ , whereas, for CRM,  $\chi^2 = 301.6$ ).

Figure 3 shows the behavior of the two models in representing fluorescence experimental profiles. Although both fittings are quite unsatisfactory, CRM clearly outperforms EPM. In fact,

Figure 3a shows that although EPM provides a slightly better fit for the profile of  $q$  at low light intensities, only CRM is capable of representing the trend at a higher intensity. Figure 3b, however, shows that although EPM performance is still the worst one, both models cannot properly represent the oxidative state of PSU in light adapted cells, consistent with the fact that these kinds of measurements were not considered in building such models. This means that both models are not accurate in estimating the PSU oxidative state.

Also in the predictions of the growth rate constant, reported in Figure 4, CRM outperforms EPM. In fact, EPM does not



**Figure 4.** Growth rate constant predicted by the EP (solid line) and CR models (dashed line) and experimental values of the growth rate constant (black squares).

predict the correct value of light intensity at which the maximum growth rate is reached and underestimates the growth rate constant at low light. CRM correctly predicts the optimal light intensity but it overestimates the growth rate constant at low light.

No additional experiments are needed for discrimination purposes and CRM is then selected, although the experimental data demonstrate that further improvements are required to improve the fitting in the region between 150 and 700  $\mu\text{E}/(\text{m}^2 \text{s})$  with respect to the PSU oxidation state. It is worth underlining that the main difference between EPM and CRM is the presence, in the latter, of a photoacclimation term expressed through the number of PSUs per cell. This suggests that the inclusion of this response is absolutely necessary for an accurate description of the photosynthetic performances and is consistent with its biological relevance, demonstrated by the fact that acclimation responses are conserved in all photosynthetic organisms.

**5.1. Enhancing CRM.** To improve the model, a more detailed description of some fundamental biological phenomena needs to be introduced. According to several works in the



literature, PSII photoinhibition occurs at all light intensities.<sup>29,34,35</sup> Therefore, we assumed that photoinhibition does not depend on the number of active PSUs, but is simply related to light intensity. Above  $I_{cr}$  (the light intensity where photosynthesis is saturated) a second process (photoprotection) is activated: in these conditions a fraction of the energy absorbed does not lead either to photochemistry or to PSU damage but is simply dissipated (e.g., because of NPQ). Accordingly, eq 6 has been modified to represent the different behavior above and below  $I_{cr}$ . Furthermore, in eq 6, the reaction order with respect to the light intensity is assumed to be 0.5: since there is no clear physical reason for setting such a value, here we decided to increase the model flexibility and its capability of incorporating all energy dissipation phenomena, by making the reaction order a parameter ( $\alpha$ ) to be estimated.

As mentioned above, photoprotection mechanisms are activated when photochemical reactions are saturated, and accordingly parameter  $I_{cr}$  plays a key role in representing this behavior. The value of critical light intensity when photosynthesis is saturated has been fixed to  $150 \mu\text{E}/(\text{m}^2 \text{ s})$ , which is the light intensity in response to which *Nannochloropsis* growth is maximal, and the limit over which the growth rate is not linearly dependent on light anymore.<sup>17</sup> This choice is well consistent with the observation of NPQ dependence from light intensity observed in *Nannochloropsis* cells, which shows activation only over this limit (NPQ data are reported in Appendix B). This parameter thus depends on both light intensity and the number of active PSUs.

Also eq 8 representing photoacclimation is modified to allow for a higher flexibility: the hyperbolic form is retained, but the light exponent becomes an additional parameter ( $\alpha_2$ ) to be estimated (see later on, eq 13).

Finally, the representation of the maintenance factor is modified, too. In CRM the maintenance factor is treated as a constant (in fact, this is a typical assumption). However, the maintenance factor should vary with light intensity to account for the metabolic cost of repairing damaged PSUs.<sup>36</sup> In order to do this, the easiest way is to express the maintenance factor as a linear function of the damaged PSUs fraction. Since the fluorescence measurements return the number of active PSUs, the variable term of the maintenance factor was related to the difference between the maximum value of fluorescence ( $q_{\max}$ ) and the current value of fluorescence ( $q$ ). Finally the fluorescence measurements  $q$  and  $q_L$  are related to the oxidation state of the PSUs, as discussed in section 4.

The modified CRM (mCRM) is thus constituted by the following set of equations:

$$\frac{da_2}{dt} = k_a I a_1 - k_d a_2 \quad (10)$$

$$\frac{da_3}{dt} = \begin{cases} k_{i,0} I a_t - k_r a_3 & \text{if } I \leq I_{cr} \\ k_{i,0} I_{cr} a_t + k_{i,1} (I - I_{cr})^\alpha (a_1 + a_2) - k_r a_3 & \text{if } I > I_{cr} \end{cases} \quad (11)$$

$$a_1 + a_2 + a_3 = a_t \quad (12)$$

$$a_t = \frac{1}{k_c + I^{\alpha_2}} \quad (13)$$

$$\mu = k_p k_2 a_2 - M \quad (14)$$

$$M = M_0 + k_M (q_{\max} - q) \quad (15)$$

$$q = q_{\max} \frac{a_1 + a_2}{a_t} \quad (16)$$

$$q_L = \frac{a_1}{a_1 + a_2} \quad (17)$$

where a new vector of the model parameters is  $\hat{\theta} = [k_a, k_d, k_{i,0}, k_{i,1}, k_r, k_p, k_c, k_M, M_0, \alpha, \alpha_2]$ .

## 6. IDENTIFIABILITY ANALYSIS

In order to be reliable and suitable for process simulation and optimization, the model for photosynthetic biomass growth has to be identified against the experimental data. It can be verified that mCRM shows identifiability issues if a parameter estimation is performed on the whole parameters vector  $\hat{\theta}$ . The estimation of the model parameters is characterized by large confidence intervals for some parameters, and more than one set of optimal values can be determined (i.e., the model is not uniquely identifiable). In order to overcome this problem both global (or structural) and practical identifiability of the model have been studied. First of all, the global identifiability has been verified using a differential algebra based method. Afterward, the practical identifiability has been assessed through a sensitivity analysis and a model reparameterization has been performed.

**6.1. Global Identifiability Analysis.** The first step for testing model identifiability is represented by global identifiability analysis. Global identifiability analysis is in fact a necessary condition for practical identifiability and can provide the minimum number of observations required to identify a unique set of optimal parameter values. The two hypotheses, upon which structural identifiability analysis rely, are as follows: (i) complete absence of measurement errors and (ii) a perfectly accurate model structure. Those two assumptions refer to an ideal case, and therefore, once a model is verified to be globally identifiable, practical identifiability, too, needs assessing. Several methods to study the structural identifiability of nonlinear models have been developed in the past two decades and are nicely reviewed in the work by Miao et al.<sup>37</sup>

A rather common approach is given by the series expansion method (e.g., Dochain et al.<sup>38</sup> applied this approach to kinetic models of activated sludge respiration). It requires that functions representing the model are infinitely differentiable, as the method involves the calculation of arbitrary order derivatives. Several examples of application of this method can be found in the literature. However, the series expansion method has a serious drawback: for high dimensional models, high order derivatives are necessary and the resulting equations can easily become too complicated to solve. In fact, it was verified that in this case the series expansion method leads to an intractable problem. Thus, in order to overcome the difficulties related to high order derivatives calculation and the resolution of the resulting equations, a method based on differential algebra was taken into account. In particular, the software package DAISY<sup>39</sup> has been used here. DAISY, as with all differential algebra based methods, requires equations of the models to be written in the form of differential polynomials.<sup>37</sup> For this reason, mCRM has been slightly modified as detailed in Appendix C.

Results indicate that a single experiment is not sufficient to identify a unique value of all of the parameters. However, if (at least) three parallel experiments are carried out at different light intensities and both concentration and fluorescence

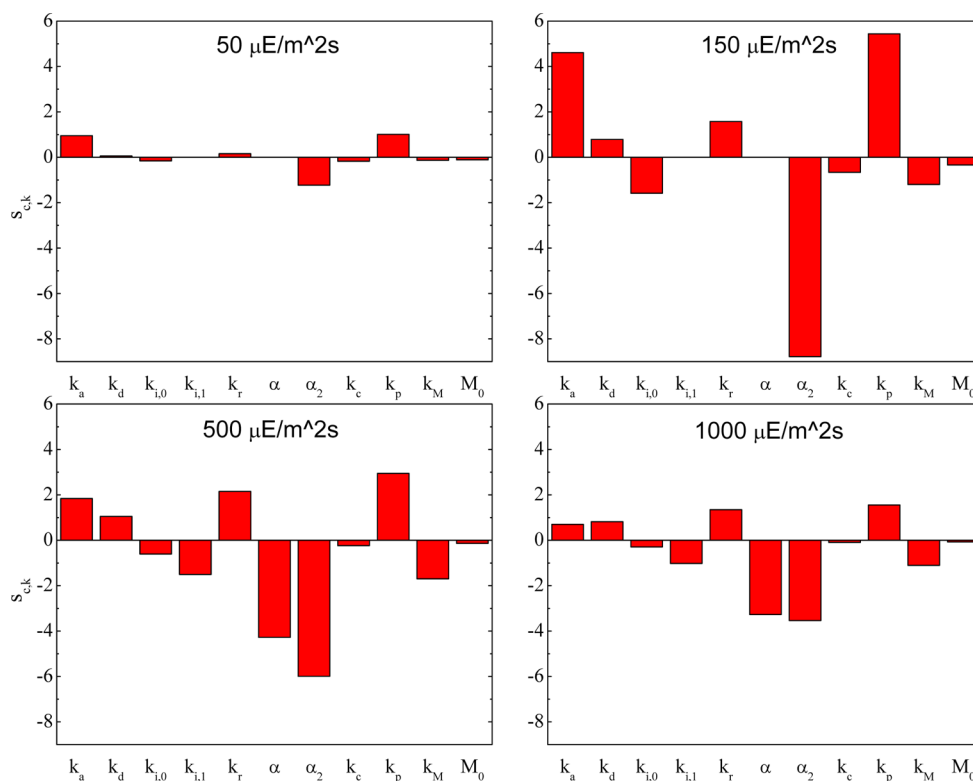


Figure 5. Final values of dynamic sensitivities for mCRM, evaluated at different light intensities.

parameters are measured during the experiments, the model is globally identifiable. This means that the experienced identifiability issues depend on practical identifiability. To tackle the problem, a sensitivity analysis and a reparameterization will be performed. Note that in several cases, where more complex models need considering, the analysis of the model practical identifiability may be the only test that can be carried out, since the verification of global identifiability may not be viable.

**6.2. Sensitivity and Correlation Analysis.** As the model is structurally identifiable, its practical identifiability will now be tested through a sensitivity analysis. The sensitivity  $q_{ik}$  of the  $i$ th response to the  $k$ th model parameter is defined as

$$q_{ik} = \frac{\partial y_i}{\partial \theta_k} \approx \frac{y'_i - y_i}{\Delta \theta_k} \quad i = 1 \dots N_M, \quad k = 1 \dots N_\theta \quad (18)$$

where  $y_i$  is the  $i$ th measured responses predicted by the model,  $y'$  is the same response obtained from a perturbed value of the  $k$ th parameter  $\theta_k$ , and  $\Delta \theta_k$  is the perturbation ( $N_M$  is the number of measured responses and  $N_\theta$  is the number of model parameters). The principal goal of sensitivity analysis is to evaluate the impact of each parameter on the measured responses and to underline the presence of correlation among specific subsets of model parameters. In an overparameterized model, near-zero sensitivity values would be obtained, leading to the nonidentifiability of some subsets of model parameters. In the most desirable case, sensitivity profiles should be clearly distinct and far from being symmetrical (i.e., they should present a low mutual correlation). As the sensitivity analysis requires prespecified parameter values, values by a preliminary

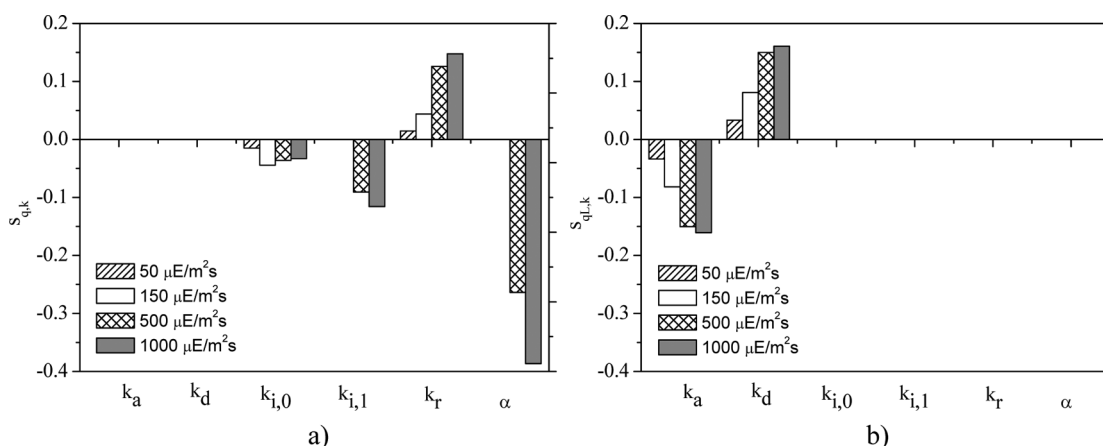
parameter estimation have been considered. Perturbation  $\Delta \theta_k$  was set equal to 1% of the parameter values.

First, the sensitivity profiles of biomass concentration will be taken into account. Although not shown here for the sake of conciseness, it was verified that all sensitivity profiles behave as exponential curves. As a consequence, the final value of sensitivities (value at 120 h) is sufficient to analyze the system behavior in different experimental conditions. The sensitivity profiles evaluated at four different light intensities are reported in Figure 5.

For light intensities under the critical value, the sensitivities related to parameters  $k_{i,0}$  and  $k_r$  show opposite values. Above the critical value of light intensity, the sensitivity of  $k_r$  is the opposite of the sum of the sensitivities of  $k_{i,0}$  and  $k_{i,1}$ . This suggests that it may be difficult to exploit biomass concentration measurements to identify  $k_{i,0}$  and  $k_r$ .

Another critical aspect is related to the fact that parameters  $k_a$  and  $k_p$  exhibit a very similar sensitivity at low light conditions (availability of growth data at high light intensities may be necessary for a robust estimation of these parameters). Also it should be noticed that the parameter related to photoacclimation,  $k_c$ , and the parameter representing the maintenance factor in the dark,  $M_0$ , exhibit a very similar (and low) sensitivity in all light conditions.

Fluorescence profiles ( $q$  and  $q_L$ ) are also considered. Because the dynamics of PSUs are fast with respect to the sampling time, and our measurements are steady state measurements, only the steady state values of sensitivities will be reported. In Figure 6a,b the values of the sensitivities of  $q$  and  $q_L$  are reported for four different light conditions. The sensitivities of parameters  $\alpha_2$ ,  $k_c$ ,  $k_p$ ,  $k_M$ , and  $M_0$  are not reported, since they are zero for both fluorescence measurements (they are related to biomass growth and are not concerned with the oxidation state of the PSUs).



**Figure 6.** Final values for dynamic sensitivities for the Camacho Rubio model evaluated at different light intensities considering literature values for model parameters.

Considering the steady state sensitivity values of the fluorescence measurements, the following critical aspects can be noticed: (1) parameters  $k_a$  and  $k_d$  do not affect the dark fluorescence measurements, while they are completely anticorrelated if light fluorescence measurements are available, showing opposite sensitivities in all light conditions; (2) parameters  $k_{i,0}$ ,  $k_{i,1}$ ,  $k_r$  and  $\alpha$  are not affected by light fluorescence measurements; (3) parameter  $k_r$  is anticorrelated with  $k_{i,0}$ , under the critical value of light intensity, and with  $k_{i,1}$ , above the critical value of light intensity; (4) the sensitivity of  $k_{i,0}$  is always quite small.

The results suggest that a model reparameterization may be necessary to help tackling the issue.<sup>24</sup>

**6.3. Reparameterization of the Camacho Rubio Model.** Because the fluorescence data are measurements of the PSU oxidation state under steady state conditions, an analytical expression for fluorescence measurements can be derived from eqs 10 to 12:

$$\frac{q}{q_{\max}} = \frac{a_t - a_3}{a_t} = \begin{cases} 1 - R_0 I & \text{if } I \leq I_{\text{cr}} \\ \frac{1 - R_0 I_{\text{cr}}}{R_1 (I - I_{\text{cr}})^{\alpha} + 1} & \text{if } I > I_{\text{cr}} \end{cases} \quad (19)$$

$$q_L = \frac{a_1}{a_1 + a_2} = \frac{1}{R_2 I + 1} \quad (20)$$

with  $R_0 = k_{i,0}/k_r$ ,  $R_1 = k_{i,1}/k_r$  and  $R_2 = k_a/k_d$ . The interesting aspect is that only four parameters affect the steady state values of the fluorescence measurements: the three ratios  $R_0$ ,  $R_1$ , and  $R_2$  and parameter  $\alpha$ . This suggests that, in order to have a practically identifiable model, the values of two parameters affecting the PSU dynamics have to be fixed if no dynamics in the fluorescence data can be included in the data set. It was chosen to fix the values of  $k_r$  and  $k_d$ , as those parameters represent the rate constant of the recovery and deexcitation processes, respectively. An approximated estimation can be obtained, considering the time scales of the processes involved: according to literature,<sup>40</sup> values of  $100 \text{ s}^{-1}$  for  $k_a$  and of  $2.22 \times 10^{-4} \text{ s}^{-1}$  for  $k_r$  were assumed. The very low sensitivity to available measurements and the high correlation of  $k_c$  and  $M_0$  make it impossible to identify the two parameters. However, in the case of  $M_0$ , previous experiments suggest it to be between 5% and 10% of the maximum growth rate.<sup>41</sup> Here we set  $M_0 = 1.5 \times 10^{-3} \text{ h}^{-1}$  ( $\sim 8\%$  of the maximum growth rate). For parameter  $k_c$  the preliminary estimation suggests that a “small” value is required for a good description of data. Since in eq 13 we

have that  $I^{\alpha} \gg 1$ , we verified that  $k_c = 1$  is a good approximation for representing all experimental conditions. After reparameterization, the vector of parameters to estimate is  $\hat{\theta}^* = [R_0, R_1, R_2, k_p, k_M, \alpha, \alpha_2]$ .

## 7. RESULTS AND DISCUSSION

In the first part of this section the mCRM parameter estimation results will be presented and discussed. In the second part two additional growth curves and two new measurements of dark fluorescence will be used to validate the model.

**7.1. Model Calibration.** In the estimation procedures, parameters are normalized with respect to the initial values obtained by the preliminary parameter estimation, to increase numerical robustness. The results of parameters estimation are reported in Table 3, along with the confidence intervals and

**Table 3.** Estimated Values of Parameters of the Reparameterized mCRM, Normalised Values (with Respect to the Initial Values) of the Parameters, Confidence Intervals (conf int), and  $t$ -Student Values for Parameters<sup>a</sup>

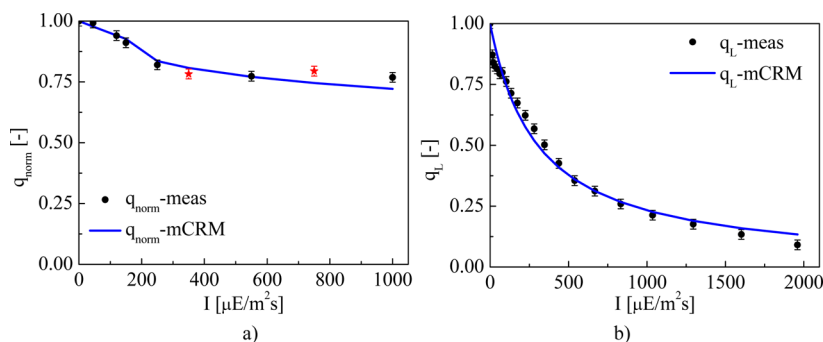
param	estd value	normalized value	95% conf int	$t$ -value	95%
$R_0$	$4.93 \times 10^{-4}$	1.12	0.21	5.20	
$R_1$	$1.34 \times 10^{-2}$	0.65	0.21	3.07	
$R_2$	$3.30 \times 10^{-3}$	2.77	0.41	6.72	
$k_p$	$1.48 \times 10^{-6}$	0.41	0.12	3.35	
$k_M$	$1.12 \times 10^{-1}$	1.96	0.91	2.15	
$\alpha$	0.45	0.45	0.048	9.52	
$\alpha_2$	0.30	0.30	0.078	3.89	

<sup>a</sup>The reference  $t$ -value is equal to 1.67.

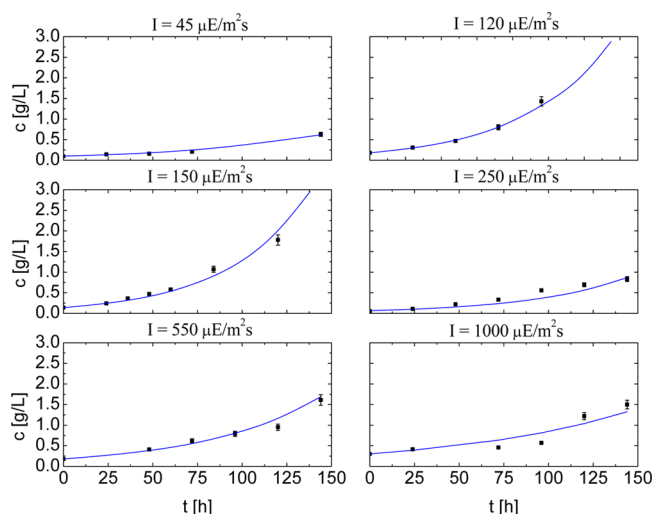
the  $t$ -values (for a statistically precise estimation of a model parameter the  $t$ -value has to be greater than a reference  $t$ -value). The  $t$ -value statistic shows that the parameter values are estimated in a statistically satisfactory way.

The profiles of fluorescence, predicted after the identification of the reparameterized mCRM, are reported in Figure 7. We can observe that the model correctly fits both the dark fluorescence profile (Figure 7a) and the light fluorescence measurements (Figure 7b). Biomass growth profiles are rather well represented by the model as illustrated in Figure 8, where the six different illuminating conditions are represented.

In Figure 9 the growth rate constant predicted by the model is reported along with the experimental value (i.e., the value obtained



**Figure 7.** Measurement (black circles) and predicted values of (a)  $q$  and (b)  $q_L$ . Blue solid lines represent the profiles predicted by the modified Camacho Rubio model. Red stars in panel a represent the experimental data used for model validation.



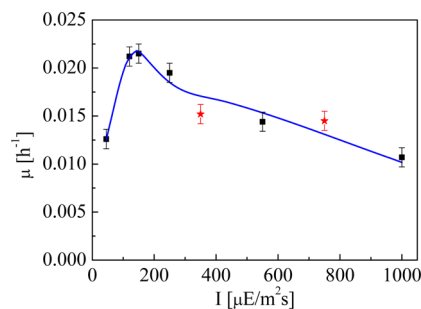
**Figure 8.** Biomass concentration profiles at different light intensities predicted by the modified Camacho Rubio model. Black circles represent the experimental measurements.

fitting the experimental data of growth during the exponential phase with an exponential curve). We can observe that, for all light intensities at which an experiment was carried out, the model well describes the experimental values of the growth rate constant. This suggests that, thanks to the fundamental input of the fluorescence parameters in illuminated cells ( $q_L$ ), the model is capable of reproducing with sufficient accuracy the basic processes of photosynthesis: photochemistry, light damage, and also energy dissipation.

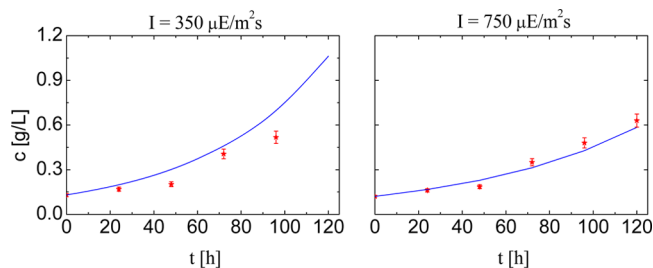
From a statistical point of view, the predicted profiles have a  $\chi^2$  value of 117.7 (whereas  $\chi^2 = 890.1$  in the case of EPM and  $\chi^2 = 301.6$  for CRM).

**7.2. Model Validation.** In order to validate the model, two experiments, not used for model calibration, will be taken into account. In particular two growth curves at 350 and 750  $\mu\text{E}/(\text{m}^2 \text{s})$  have been considered. For both cultures also the value of  $q$  has been measured and was exploited for model validation. In Figures 7a and 9 the validation points are represented by red stars. In Figure 10 the growth curves have been reported.

We can observe that the predictions are accurate for the parameter  $q$ . Also the growth curves are predicted in a sufficiently accurate way, although the growth at 350  $\mu\text{E}/(\text{m}^2 \text{s})$  is slightly overestimated by the model, while the growth at 750  $\mu\text{E}/(\text{m}^2 \text{s})$  is somewhat underestimated in the



**Figure 9.** Growth rate constant predicted by the modified Camacho Rubio model (solid line) and experimental values of the growth rate constant (black circles). Red stars represent the experimental values of the growth rate constant for the experiments used in the model validation.



**Figure 10.** Biomass concentration profiles at different light intensities predicted by the modified Camacho Rubio model. Red stars represent the experimental measurements used for the model validation.

final part. From a statistical point of view, validation leads to a  $\chi^2$  value of 226.23.

## 8. CONCLUSION

A literature model<sup>10</sup> was selected and modified to describe microalgae growth through a rigorous identification procedure. The estimation of the model parameters was performed considering experimental data (growth profiles and fluorescence measurements) on *N. salina*. Results show that the developed model well represents biomass growth over a wide range of light intensities. The modified model was also able to reproduce fluorescence measurements, including the light profile of PSU saturation. This suggests that the proposed model is accurate enough to represent all major processes of photosynthesis, photochemistry, PSU damage, and energy dissipation. While results in reproducing experimental data are fully satisfactory, it

Table 4. Measured Biomass Concentration Profiles at Different Light Intensities<sup>a</sup>

time (h)	c (g/L)							
	I = 50 μE/(m <sup>2</sup> s)	I = 120 μE/(m <sup>2</sup> s)	I = 150 μE/(m <sup>2</sup> s)	I = 250 μE/(m <sup>2</sup> s)	I = 350 μE/(m <sup>2</sup> s)	I = 550 μE/(m <sup>2</sup> s)	I = 750 μE/(m <sup>2</sup> s)	I = 1000 μE/(m <sup>2</sup> s)
0	9.93 × 10 <sup>-2</sup>	1.81 × 10 <sup>-1</sup>	1.42 × 10 <sup>-1</sup>	5.49 × 10 <sup>-2</sup>	1.30 × 10 <sup>-1</sup>	1.80 × 10 <sup>-1</sup>	1.21 × 10 <sup>-1</sup>	3.02 × 10 <sup>-1</sup>
24	1.44 × 10 <sup>-1</sup>	3.07 × 10 <sup>-1</sup>	2.40 × 10 <sup>-1</sup>	1.10 × 10 <sup>-1</sup>	1.69 × 10 <sup>-1</sup>	–	1.61 × 10 <sup>-1</sup>	4.12 × 10 <sup>-1</sup>
36	–	–	3.60 × 10 <sup>-1</sup>	–	–	–	–	–
48	1.57 × 10 <sup>-1</sup>	4.69 × 10 <sup>-1</sup>	4.68 × 10 <sup>-1</sup>	2.21 × 10 <sup>-1</sup>	2.02 × 10 <sup>-1</sup>	4.10 × 10 <sup>-1</sup>	1.86 × 10 <sup>-1</sup>	–
60	–	–	5.76 × 10 <sup>-1</sup>	–	–	–	–	–
72	2.07 × 10 <sup>-1</sup>	8.01 × 10 <sup>-1</sup>	–	3.32 × 10 <sup>-1</sup>	4.06 × 10 <sup>-1</sup>	6.19 × 10 <sup>-1</sup>	3.50 × 10 <sup>-1</sup>	4.56 × 10 <sup>-1</sup>
84	–	–	1.07	–	–	–	–	–
96	–	1.43	–	5.57 × 10 <sup>-1</sup>	5.17 × 10 <sup>-1</sup>	7.87 × 10 <sup>-1</sup>	4.80 × 10 <sup>-1</sup>	5.70 × 10 <sup>-1</sup>
120	–	–	1.78	6.92 × 10 <sup>-1</sup>	–	9.50 × 10 <sup>-1</sup>	6.30 × 10 <sup>-1</sup>	1.22
144	6.31 × 10 <sup>-1</sup>	–	–	8.32 × 10 <sup>-1</sup>	–	1.61	–	1.50

<sup>a</sup>Experiments carried out at 350 and 750 μE/(m<sup>2</sup>s) have been used for mCRM model validation.

should be underlined that algae growing in an industrial scale photobioreactor are exposed to different conditions. In particular light is not homogeneously distributed because of cells shading, and illumination intensity is not constant because of diurnal changes and cells mixing. Finally nutrient availability can also be limiting. Future efforts will be made to expand the model to include these phenomena, also by designing appropriate experiments.

## APPENDIX A

In this appendix the entire set of data used in this work is reported. Table 4 reports the growth data at different light intensities. Each datum is the mean of two experiments conducted in parallel in two identical reactors. Table 5 contains

Table 5. Values of Growth Rate Constant and Coefficient of Determination R<sup>2</sup> Obtained from the Linear Regression of the Growth Curves, Reported in a Semilog Scale<sup>a</sup>

I [μE/(m <sup>2</sup> s)]	μ (h <sup>-1</sup> )	R <sup>2</sup>
45	0.0126	0.976
120	0.0212	0.998
150	0.0215	0.987
250	0.019	0.954
350	0.0152	0.953
550	0.0144	0.984
750	0.0145	0.975
1000	0.0107	0.875

<sup>a</sup>Data points used for the growth rate constant determination are reported in Table 4.

the values of the growth rate constant as obtained from the linear regression of the growth curves reported in a semilog scale; the coefficient of determination R<sup>2</sup> is also included. From every biomass culture, a sample was taken during the exponential growth phase and the value of q was measured: in Table 6 the available measurements are reported. Finally, the second fluorescence parameter q<sub>L</sub> was measured for 21 different light intensities using a PAM fluorometer. The data points obtained by the PAM fluorometer are reported in Table 7.

## APPENDIX B

Chlorophyll fluorescence was determined *in vivo* using Dual PAM 100 from WALZ. Cells were grown as in Sforza et al.,<sup>17</sup> at 50 μmol m<sup>-2</sup> s<sup>-1</sup>. After 20 min of dark adaptation cells were

Table 6. Measured Value of Parameter q and Its Normalized Value q<sub>norm</sub> at Different Light Intensity (Normalization Factor Set to 0.65, the Value of q Corresponding to Fully Active PSUs)<sup>a</sup>

I [μE/(m <sup>2</sup> s)]	q	q <sub>norm</sub>
0	0.650	1.00
46	0.645	0.992
120	0.611	0.940
150	0.592	0.911
250	0.533	0.820
350	0.509	0.783
550	0.503	0.773
750	0.516	0.794
1000	0.499	0.769

<sup>a</sup>The measurements taken at 350 and 750 μE/(m<sup>2</sup>s) have been used for model validation.

exposed to increasing actinic light (60 s for each point). After each light treatment NPQ was determined by a saturating flash and calculated as (F<sub>m</sub> - F<sub>m</sub>')/F<sub>m</sub>'. Data are illustrated in Figure 11.

## APPENDIX C

Global identifiability analysis was performed using the software package DAISY.<sup>39</sup> Given that DAISY is based on differential algebra techniques, it requires equations of the model to be written in the form of differential polynomials. Thus, in eqs 11 and 13 terms (I - I<sub>cr</sub>)<sup>α</sup> and I<sup>α<sub>2</sub></sup> have to be modified. In particular it is possible to define two additional variables, I<sub>α</sub> and I<sub>α<sub>2</sub></sub>, to be the Taylor series expansion of (I - I<sub>cr</sub>)<sup>α</sup> and I<sup>α<sub>2</sub></sup>, respectively. Accordingly

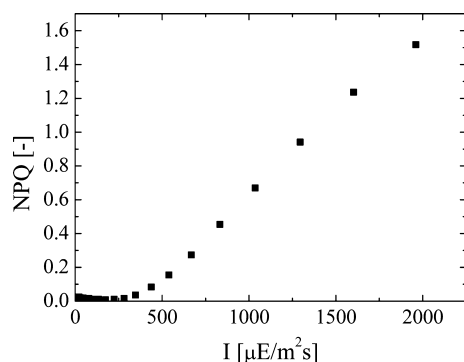
$$I_{\alpha} = (I - I_{cr})^{\bar{\alpha}} \left( \sum_{n=0}^{+\infty} \frac{(\alpha - \bar{\alpha})^n}{n!} \ln^n(I - I_{cr}) \right) \quad (C1)$$

$$I_{\alpha_2} = I^{\bar{\alpha}_2} \left( \sum_{n=0}^{+\infty} \frac{(\alpha_2 - \bar{\alpha}_2)^n}{n!} \ln^n(I) \right) \quad (C2)$$

where  $\bar{\alpha}$  and  $\bar{\alpha}_2$  are parameter values at which the series expansion has been considered. In this work we considered McLaurin series truncated at the second order and a value of 0.5 for  $\bar{\alpha}$  and  $\bar{\alpha}_2$ . The system of eqs 10–17, where eqs C1 and C2 have been used to substitute (I - I<sub>cr</sub>)<sup>α</sup> and I<sup>α<sub>2</sub></sup>, has been verified to be globally identifiable of several values of

**Table 7. Measured Value of Parameter  $q_L$  at Different Light Intensity**

$I$ [ $\mu\text{E}/(\text{m}^2\text{s})$ ]	$q_L$
0	0.650
14	0.567
21	0.545
30	0.536
45	0.529
61	0.516
78	0.519
103	0.495
134	0.464
174	0.438
224	0.405
281	0.369
347	0.326
438	0.277
539	0.231
668	0.203
833	0.168
1036	0.138
1295	0.114
1602	0.0871
1960	0.0592

**Figure 11.** Measured NPQ values for different light intensities.

parameters  $\bar{\alpha}$  and  $\bar{\alpha}_2$ , if at least three experiments are carried out at different light intensities.

## AUTHOR INFORMATION

### Corresponding Author

\*Fax: +39.049.827.5461. E-mail: fabrizio.bezzo@unipd.it.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by ERC Starting Grant BIOLEAP No. 309485 to T.M. and by project PRIN 2012XSAWYM “Improving biofuels and high added value molecules production from unicellular algae” to T.M. and F.B. A.B. and F.B. gratefully acknowledge Fondazione Cariparo for Grant Progetto Dottorati di Ricerca 2012 under whose framework this research has been carried out.

## REFERENCES

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

(2) Mata, T.; Martins, A.; Caetano, N. Microalgae for biodiesel production and other applications: A review. *Renewable Sustainable Energy Rev.* **2010**, *14*, 217–232.

(3) Greenwell, H.; Laurens, L.; Shields, R.; Lovitt, R.; Flynn, K. Placing microalgae on the biofuels priority list: A review of the technological challenges. *J. R. Soc., Interface* **2010**, *7*, 703–726.

(4) Vajda, S.; Godfrey, K.; Rabitz, H. Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Math. Biosci.* **1989**, *93*, 217–248.

(5) Li, Z.; Wakao, S.; Fisher, B.; Niyogi, K. Sensing and responding to excess light. *Annu. Rev. Plant Biol.* **2009**, *60*, 239–260.

(6) Ross, O.; Geider, R. New cell based model of photosynthesis and photoacclimation: Accumulation and mobilisation of energy reserves in phytoplankton. *Mar. Ecol.: Prog. Ser.* **2009**, *383*, 53–71.

(7) Marshall, H.; Geider, R.; Flynn, K. A mechanistic model of photoinhibition. *New Phytol.* **2000**, *145*, 347–359.

(8) Prézelin, B. B. Light reactions in photosynthesis. *Can. J. Fish. Aquat. Sci.* **1981**, *210*, 1–43.

(9) Fasham, M.; Platt, T. Photosynthetic response of phytoplankton to light: A physiological model. *Proc. R. Soc. London, Ser. B* **1983**, *66*, 355–370.

(10) Camacho Rubio, F.; Garcia Camacho, F.; Fernandez Sevilla, J. M.; Christi, Y.; Molina Grima, E. A mechanistic model of photosynthesis in microalgae. *Biotechnol. Bioeng.* **2003**, *81*, 459–473.

(11) Eilers, P.; Peeters, J. A model for the relationship between light intensity and the rate of photosynthesis in phytoplankton. *Ecol. Modell.* **1988**, *42*, 199–215.

(12) Garcia-Camacho, F.; Sanchez-Miron, A.; Molina-Grima, E.; Camacho-Rubio, F.; Merchuck, J. A mechanistic model of photosynthesis in microalgae including photoacclimation dynamics. *J. Theor. Biol.* **2012**, *304*, 1–15.

(13) Han, B. P. Photosynthesis-irradiance response at physiological level: A mechanistic model. *J. Theor. Biol.* **2001**, *148*, 121–127.

(14) Pahlow, M. Linking chlorophyll–nutrient dynamics to the Redfield N:C ratio with a model of optimal phytoplankton growth. *Mar. Ecol.: Prog. Ser.* **2005**, *287*, 33–43.

(15) Papadakis, I.; Kotzabasis, K.; Lika, K. Modeling the dynamic modulation of the light energy in photosynthetic algae. *J. Theor. Biol.* **2012**, *300*, 254–264.

(16) Wu, X.; Merchuck, J. C. A model integrating fluid dynamics in photosynthesis and photoinhibition processes. *Chem. Eng. Sci.* **2001**, *56*, 3527–3538.

(17) Sforza, E.; Simionato, D.; Giacometti, G.; Bertuccio, A.; Morosinotto, T. Adjusted light and dark cycles can optimize photosynthetic efficiency in algae growing in photobioreactors. *PLoS One* **2012**, No. 7, No. e38975.

(18) Box, G.; Hill, W. Discrimination among mechanistic models. *Technometrics* **1967**, *9*, 57–61.

(19) Stewart, W.; Henson, T.; Box, G. Model discrimination and criticism with single-response data. *AIChE J.* **1996**, *42*, 3055–3062.

(20) Akaike, H. A New Look at the Statistical Model Identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723.

(21) Stewart, W. E.; Shon, Y.; Box, G. E. P. Discrimination and Goodness of Fit of Multiresponse Mechanistic Models. *AIChE J.* **1998**, *44*, 1403–1412.

(22) Alberton, A.; Schwaab, M.; Lobao, M.; Pinto, J. Design of experiments for discrimination of rival models based on the expected number of eliminated models. *Chem. Eng. Sci.* **2012**, *75*, 120–131.

(23) Chen, B.; Asprey, S. On the design of optimally informative dynamic experiments for model discrimination in multiresponse nonlinear situations. *Ind. Eng. Chem. Res.* **2003**, *42*, 1379–1390.

(24) Meshkat, N.; Anderson, C.; Distefano, J. Finding identifiable parameter combination in nonlinear ODE models and the rational reparameterization of their input-output equations. *Math. Biosci.* **2011**, *233*, 19–31.

(25) Galvanin, F.; Ballan, C. C.; Barolo, M.; Bezzo, F. A general model-based design of experiments approach to achieve practical identifiability of pharmacokinetic and pharmacodynamic models. *J. Pharmacokinet. Pharmacodyn.* **2013**, *40*, 451–467.

- (26) *gPROMS model validation guide* (v. 3.6); Process Systems Enterprise: London, 2012.
- (27) Maxwell, K.; Johnson, G. Chlorophyll fluorescence—A practical guide. *J. Exp. Bot.* **2000**, *51* (345), 659–668.
- (28) Barber, J.; Anderson, B. Too much of a good thing: Light can be bad for photosynthesis. *Trends Biochem. Sci.* **1992**, *17*, 61–66.
- (29) Nixon, P.; Michoux, F.; Yu, J.; Boehm, M.; Komenda, J. Recent advances in understanding the assembly and repair of photosystem II. *Ann. Bot.* **2010**, *106*, 1–16.
- (30) Raven, J. The cost of photoinhibition. *Physiol. Plant.* **2011**, *142*, 87–104.
- (31) Simionato, D.; Sforza, E.; Corteggiani, C.; Bertucco, A.; Giacometti, G.; Morosinotto, T. Acclimation of *Nannochloropsis gaditana* to different illumination regimes: Effects on lipids accumulation. *Bioresour. Technol.* **2011**, *102*, 6026–6032.
- (32) Kolber, Z.; Falkowski, P. Use of active fluorescence to estimate phytoplankton photosynthesis in situ. *Limnol. Oceanogr.* **1993**, *38*, 1646–1665.
- (33) Kramer, D.; Johnson, G.; Kuirats, O.; Edwards, G. New fluorescence parameters for determination of QA redox state and excitation energy fluxes. *Photosynth. Res.* **2004**, *79*, 209–218.
- (34) Aro, E. M.; McCaffery, S.; Anderson, J. M. Photoinhibition and D1 protein degradation in peas acclimated to different growth irradiances. *Plant Physiol.* **1993**, *103*, 835–843.
- (35) Miyao, M. Involvement of active oxygen species in degradation of the D1 protein under strong illumination in isolated subcomplexes of photosystem II. *Biochemistry* **1994**, *33*, 9722–9730.
- (36) Jansen, M. A. K.; Mattoo, A. K.; Edelman, M. D1-D2 protein degradation in the chloroplast. Complex light saturation kinetics. *Eur. J. Biochem.* **1999**, *260*, 527–532.
- (37) Miao, H.; Xia, X.; Perelson, A. S.; Wu, H. On identifiability of non linear ODE models and applications in viral dynamics. *SIAM Rev. Soc. Ind. Appl. Math.* **2011**, *53*, 3–39.
- (38) Dochain, D.; Vanrolleghem, P. A.; Van Daele, M. Structural identifiability of biokinetic models of activated sludge respiration. *Water Res.* **1995**, *29*, 2571–2578.
- (39) Bellu, G.; Saccomani, M. P.; Audoly, S.; D'Angiò, L. DAISY: A new software tool to test global identifiability of biological and physiological systems. *Comput. Methods Programs Biomed.* **2007**, *88*, 52–61.
- (40) Han, B.; Virtanen, M.; Koponen, J.; Straskraba, M. Effect of photoinhibition on algal photosynthesis: A dynamic model. *J. Plankton Res.* **2000**, *22*, 865–885.
- (41) Ryther, J. H. The ratio of photosynthesis to respiration in marine plankton algae and its effect upon the measurement of productivity. *Deep-Sea Res.* **1954**, *2*, 134–139.

# APPENDIX II

## Response of the Photosynthetic Apparatus to Salt Stress in a Resistant and Susceptible *Oryza sativa* variety

### Author names and affiliations

Giorgio Perin<sup>1</sup>, Elide Formentin<sup>2</sup>, Fiorella Lo Schiavo<sup>2</sup>,  
Tomas Morosinotto<sup>1</sup>

<sup>1</sup>PAR-Lab\_Padua Algae Research Laboratory, Department of Biology,  
University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy

<sup>2</sup> Department of Biology, University of Padova, Via U. Bassi 58/B,  
35121 Padova, Italy

The following results were obtained from a collaboration between our group and Prof. Fiorella Lo Schiavo group (from the biology department of Padova University). These results are being included in a manuscript (Formentin E. et al. – in preparation) revealing the physiological, biochemical and molecular features involved in the salt-stress response in rice (*Oryza sativa*) plants.



**Abstract.**

Rice is commonly cultivated in areas close to the sea and its fields are sometimes exposed to sea floods which induce a strong salt-stress, often heavily compromising productivity. For this reason, it is highly important to investigate the salt response in a species with such a strong economic and social relevance.

In this work we investigated the physiological, biochemical and molecular responses to salt stress of two of the most common Italian rice cultivars (*O. sativa* v. *Baldo* and *O. sativa* v. *Vialone Nano*), which differs significantly for their sensitivity to these conditions. Here we present in particular the responses involving the photosynthetic apparatus, monitoring parameters such as PSII maximum quantum yield, NPQ activation and relaxation kinetic, in response to exposure to higher salt concentrations. In this analysis, we employed an imaging apparatus to monitor photosynthetic parameters, which allowed to identify specific features between different leaves areas. Taken together the data showed that v. *Baldo* salt tolerance is correlated with the ability to activate a stress response also in the photosynthetic apparatus, such as the induction of a stronger NPQ. On the other side the sensitive *Vialone Nano* variety did not show the ability to activate such a response, indicating that NPQ activation ability can be used as a biomarker for salt stress response.

## Materials and Methods.

### *Plant material and growth conditions*

Seeds of rice plants (*Oryza sativa ssp. japonica v. Baldo* and *v. Vialone Nano*) were surface-sterilized by soaking in 70 % (w/v) ethanol solution for 1 min and consequently washed in water. Germination was conducted on moist filter paper for 72 h at 26 °C in the dark and later germinated seeds were moved to seed-holders (containing agar 0.65 % (w/v)) of hydroponic tubs (Araponics), supplied with the hydroponic solution of table 1.

**Table1. Hydroponic solution composition in water.**

<b>Macro-elements</b>	
<b>Element</b>	<b>Concentration</b>
<b>KNO<sub>3</sub></b>	1.5 mM
<b>Ca(NO<sub>3</sub>)<sub>2</sub></b>	1 mM
<b>MgSO<sub>4</sub></b>	0.5 mM
<b>NH<sub>4</sub>H<sub>2</sub>PO<sub>4</sub></b>	250 µM
<b>Micro-elements</b>	
<b>Na<sub>2</sub>O<sub>3</sub>Si</b>	30 µM
<b>FeNaEDTA</b>	25 µM
<b>MnCl<sub>2</sub></b>	9 µM
<b>ZnSO<sub>4</sub></b>	0.8 µM
<b>CuSO<sub>4</sub></b>	0.3 µM
<b>(NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub></b>	250 µM
<b>H<sub>3</sub>BO<sub>3</sub></b>	46 µM

They were grown in a phytotron with 150 µmol photons m<sup>-2</sup> s<sup>-1</sup> at 26±1 °C/21±1 °C, 16/8 h day/night photoperiod. When the second leaf was raising, plants were transferred to fresh hydroponic solution (Control – C group) or to fresh hydroponic solution supplied with 100 mM NaCl, 10 mM CaCl<sub>2</sub>, 20 mM MgSO<sub>4</sub>, 10 mM Na<sub>2</sub>SO<sub>4</sub> (Salt-treated, Salt - T group) and incubated under identical conditions, for 13

days. For the recovery experiment, after 3 or 6 days of salt treatment, plants were placed in fresh hydroponic solution without salt, for 7 days.

#### *Photosynthetic parameters evaluation*

*In vivo* chlorophyll fluorescence measurements were performed with an imaging apparatus (FluorCam FC-800 – Photon Systems Instruments, Brno, Czech Republic). Plants were dark adapted for 30 minutes and later sacrificed. The chlorophyll fluorescence value of dark-adapted samples ( $F_0$ ) was measured applying a non-actinic white light source (intensity  $< 0.05 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ).

Maximum chlorophyll fluorescence value of dark-adapted samples ( $F_m$ ) was instead measured applying a saturating light pulse (intensity =  $3500 \mu\text{mol photons m}^{-2} \text{s}^{-1}$  and 800 ms duration).

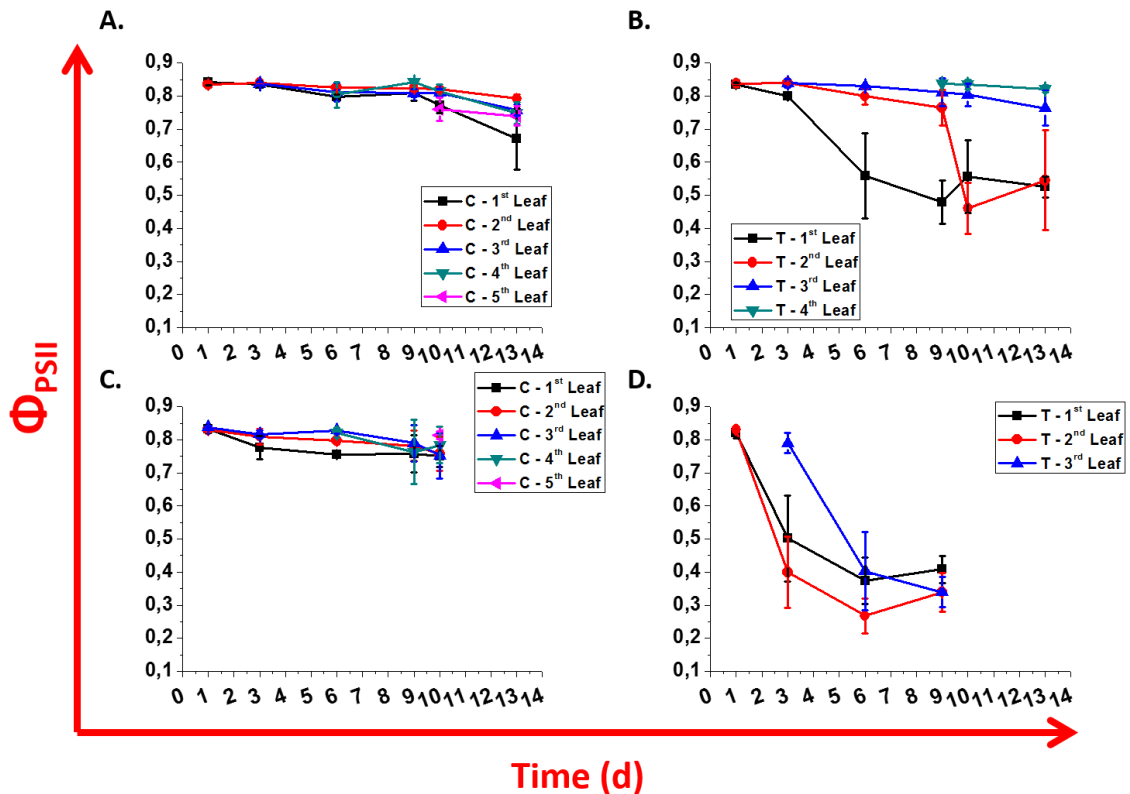
PSII functionality was expressed as PSII maximum quantum yield and it was calculated according to (Maxwell and Johnson, 2000). PSII quantum yield was monitored over time, using control plants not exposed to salt as reference. Leaves were then exposed to  $500 \mu\text{mol photons m}^{-2} \text{s}^{-1}$  for 5 minutes to evaluate non-photochemical quenching (NPQ) activation kinetic. Later light was switched off for 3 minutes to evaluate NPQ relaxation. NPQ was calculated according to (Maxwell and Johnson, 2000).  $qE$  (energy-dependent quenching) values, to which we refer as the major contributor to NPQ (de Bianchi et al., 2010), relaxes within minutes. To calculate its value we subtracted the NPQ value registered after 2 minutes of dark relaxation ( $\text{NPQ}_{D2}$ ) to the NPQ value monitored at the end of the light period ( $\text{NPQ}_L$ );  $qE = \text{NPQ}_L - \text{NPQ}_{D2}$ .

## Results and Discussion.

The scope of the work was to evaluate the effect of the salt stress on the photosynthetic efficiency for two Italian rice (*Oryza sativa ssp. japonica*) varieties, *Baldo* and *Vialone Nano* (VN). These are both largely employed in the Italian agriculture, but they differ for their ability to withstand salt stress, with the former showing a much higher tolerance. The characterization of the photosynthetic apparatus presented here will be then integrated with other physiological, biochemical and molecular characterizations (Formentin E. et al. – in preparation).

### Photosynthesis evaluation in salt treated plants.

Rice plants were cultivated until the second leaf arose (see methods for details) and after were transferred in a fresh growth media (day 0), supplied with 100mM NaCl, 10 mM CaCl<sub>2</sub>, 20 mM MgSO<sub>4</sub>, 10 mM Na<sub>2</sub>SO<sub>4</sub> (Salt-treated, Salt - T group) or without salt (Control – C group). The photosynthetic parameters were thus monitored over time for 13 days. In figure 1 the effect of the salt treatment on photosynthesis was evaluated monitoring PSII functionality from PSII maximum quantum yield ( $\Phi_{PSII}$ ).



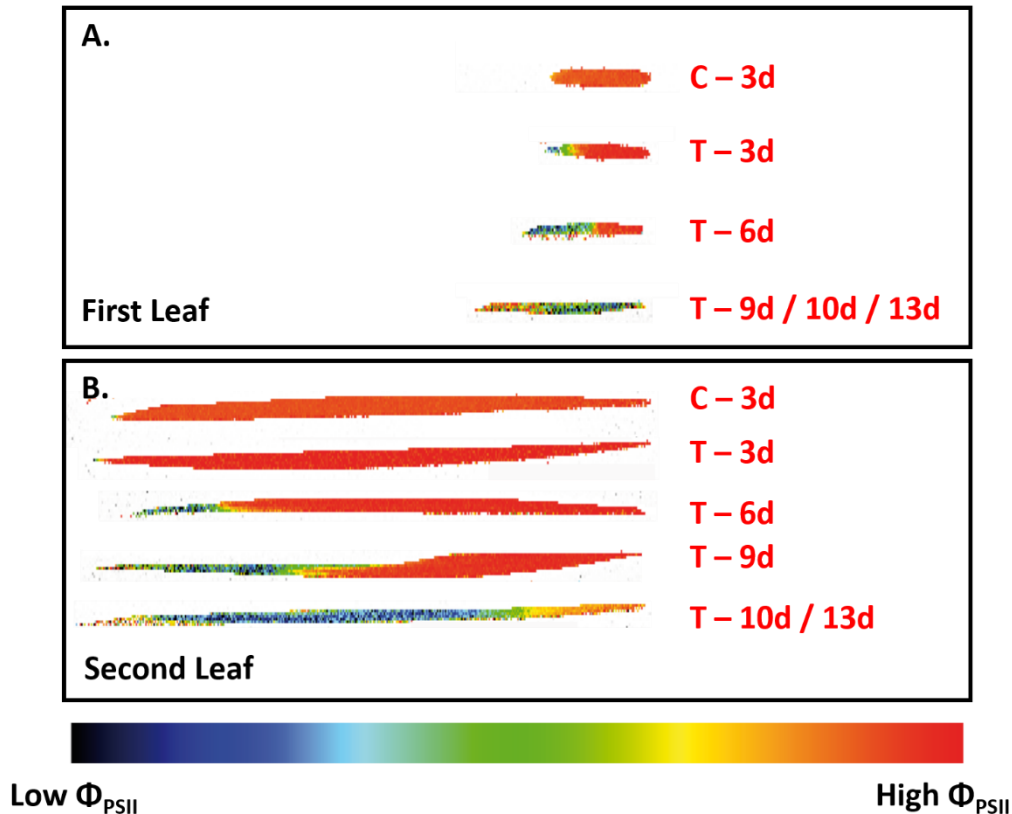
**Figure 1. PSII maximum quantum yield ( $\Phi_{PSII}$ ) trend during 13 days of salt treatment for *O. sativa* plants (v. *Baldo* and v. *Vialone Nano*).** A. and B. PSII maximum quantum yield for rice plants v. *Baldo* in control (A) and in salt treatment (B). C. and D. PSII maximum quantum yield for rice plants v. *Vialone Nano* (VN) in control (C) and in salt treatment (D). and v. *Vialone Nano* (VN) during salt treatment prolonged for 13 days. Plants not exposed to salt treatment - C; Plants exposed to salt treatment - T; black – first leaf, red – second leaf, blu – third leaf, green – fourth leaf, violet – fifth leaf. Data are expressed as average of 6 biological replicates  $\pm$  SD.

For both varieties in control leaves (Figure 1A and C) this parameter is constant over time, with some decrease in the last days. When instead plants were treated with salt (Figure 1B and D), we observed a much diversified response depending from the leaf number. In the case of the *Baldo* variety (more resistant), the first leaf of treated plants shows a weak decrease in PSII maximum quantum yield after 3 days while at day 6 the decrease is major (Figure 1B). The second leaf instead shows the same strong decrement only after the 10<sup>th</sup> day. On the other side, the third leaf didn't show any strong decrease and the fourth, that entirely arises after the salt treatment, has begun to show the same photosynthetic efficiency of the control samples (Figure 1B).

On the other side, the *VN* variety plants, which are more sensitive to the treatment, showed instead a different behavior. Just after 3 days both the first and the second leaves show huge decreases in  $\Phi_{PSII}$  suggesting they are not photosynthetically active anymore. The third leaf is instead photosynthetically active until the 6<sup>th</sup> day, when also in this case PSII maximum quantum yield drastically decreases (Figure 1D). With the salt treatment, *VN* variety plants were not able to support photoautotrophic growth anymore and therefore are not able to produce a fourth leaf in these conditions (Figure 1D). For this reason the experiment was ended prematurely.

PSII functionality thus highlights that *Baldo* plants suffered from salt stress in the first two leaves but have substantially unaffected photosynthesis of the third and fourth leaves. On the contrary, *VN* plants are strongly affected by salt treatment and showed huge impairments of the photosynthetic efficiency. The use of an imaging apparatus (see methods) allowed us to observe the behavior of different part of the leaf. This could represent an additional source of information with respect to the more common photosynthesis characterization of plants exposed to different abiotic stresses, using common pulse-amplitude-modulation (PAM) instruments (Du et al., 2010; Zhu et al., 2011; Zulfugarov et al., 2014). Figure 2 shows the images of the PSII maximum quantum yield ( $\Phi_{PSII}$ ) for the first (Figure 2A) and the second leaves (Figure 2B) of *Baldo* plants under salt stress. These clearly show that PSII maximum quantum yield ( $\Phi_{PSII}$ ) decrease is not homogeneous, in fact while the apical region of the leaf shows a drastic decrease, part of the leaf is photosynthetically active. In the sensitive variety this

response is not observable and all the leaf area shows a strong decrease of  $\Phi_{PSII}$ . These patterns likely match with the activation of specific mechanisms of NaCl allocation inside the leaf, highlighting that *Baldo* plants leaves could have evolved mechanisms to compartmentalize NaCl in contrast to *VN* plants leaves. This strategy could be advantageous also from the photosynthetic point of view since it is better to sacrifice completely a part of the leaf while keeping the rest intact, than losing efficiency in all the cells.

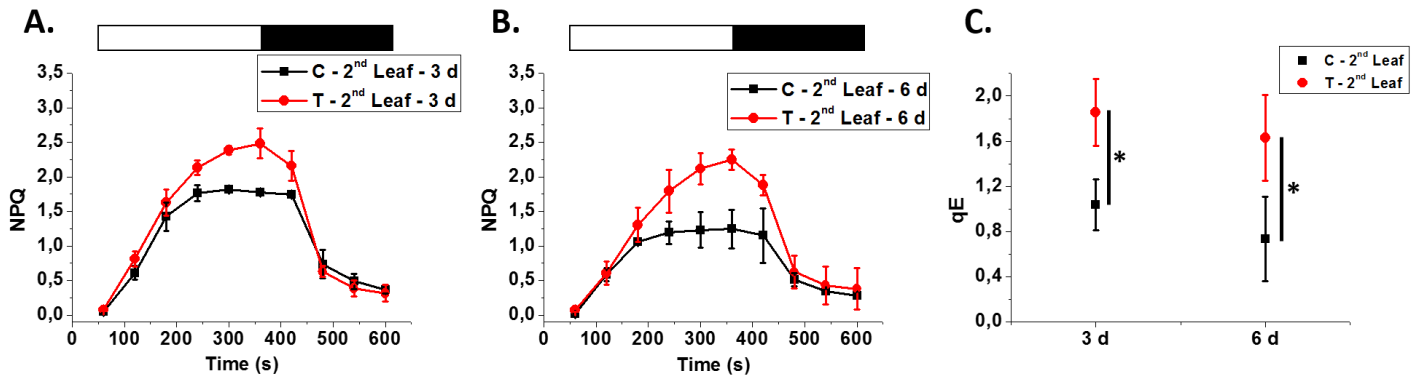


**Figure 2. PSII maximum quantum yield ( $\Phi_{PSII}$ ) images for *v. Baldo* plants leaves.** A. Fluorescence images depicting PSII maximum quantum yield values for the whole first leaf area. B. Fluorescence images depicting PSII maximum quantum yield values for the whole second leaf area. Images were taken along the salt treatment (T), using plants not exposed to salt stress as reference (Control –C). The region on the left represents the apical part of the leaf.

Chl fluorescence measurements can provide other information, such as the ability to respond to intense light, activating a Non-photochemical quenching (NPQ) response (Erickson et al., 2015).

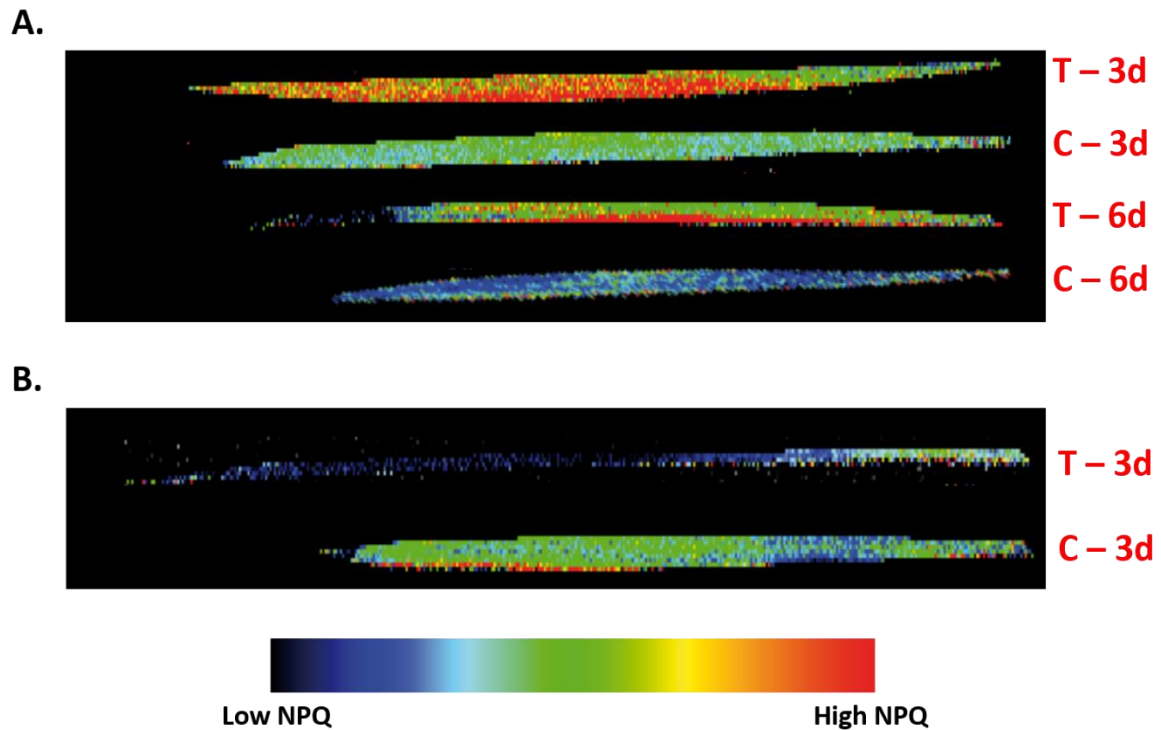
Figure 3 shows the example of the second leaf of variety *Baldo* plants after 3 (Figure 3A) and 6 days (Figure 3B) of salt treatment (T), compared with controls plants (C). As shown, the treated plants

showed a higher NPQ (red curves in Figure 3A and B). This increase is clearly due to the fast relaxed kinetic of the NPQ (qE, also quantified in Figure 3C). This qE component is known to be induced in stress conditions, such as high light (Ballottari et al., 2007; Gerotto et al., 2011) and its increase suggests an acclimation of the photosynthetic apparatus to the stress conditions experienced.

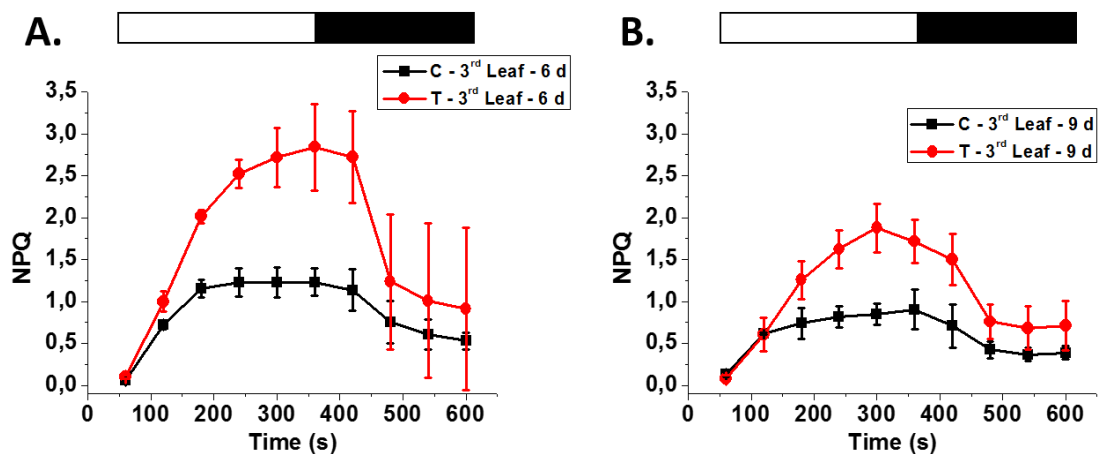


**Figure 3. NPQ kinetics and qE values for *v. Baldo* plants second leaves.** NPQ kinetics for the second leaf of rice plants *v. Baldo*, after 3 (A) and 6 days (d) (B) of salt treatment. White bars - light induction, black bars – dark relaxation (see methods for the detailed protocol). C. qE values for the second leaf of rice plants *v. Baldo*, after 3 and 6 days (d) of salt treatment; it was calculated according to the formula:  $NPQ_L - NPQ_{D2}$  (see methods). In red, salt treated plants – T; in black, control plants – C; asterisks show statistical relevant differences (one-way ANOVA, p-value < 0.05). Data are expressed as average of 6 biological replicates  $\pm$  SD.

If we take a look to the fluorescence images (Figure 4A) we see that the capacity of a higher NPQ activation is particularly present in those leaf areas maintaining their PSII functionality, indicating that they have activated a response to the salt conditions. In fact, the same *Baldo* leaf was analyzed for PSII maximum quantum yield (Figure 2B) and for NPQ (Figure 4A) activation. The same activation trend was monitored for the third leaf of *v. Baldo* plants (Figure 5). For *VN* plants, on the contrary, we did not observe any stronger NPQ activation with respect to control plants (data not shown), not even in those specific leaf regions not affected in the PSII functionality (Figure 4B).



**Figure 4. NPQ images for *v. Baldo* and *v. VN* plants leaves.** Fluorescence images depicting NPQ values for the whole second leaf area for *v. Baldo* (A) and *v. VN* (B) plants. Images were taken after 3 or 6 days (d) of salt treatment (T), using plants not exposed to salt stress as reference (Control – C). The region on the left represents the apical part of the leaf.

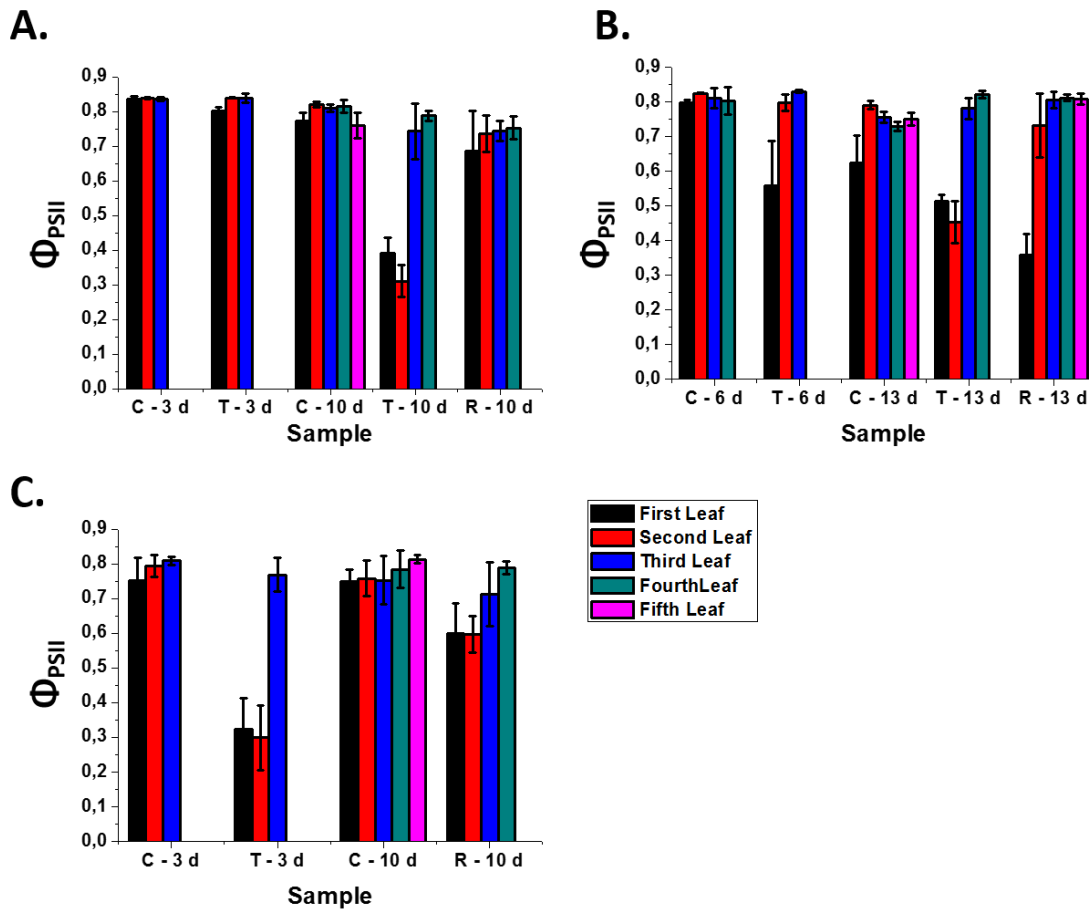


**Figure 5. NPQ kinetics for the third leaf of *O. sativa* plants (*v. Baldo*).** NPQ kinetics were registered after 6 (A) and 9 (B) days (d) of salt treatment. White bars - light induction, black bars – dark relaxation (see methods for the detailed protocol). In red, salt treated plants – T; in black, control plants – C. Data are expressed as average of 6 biological replicates  $\pm$  SD.

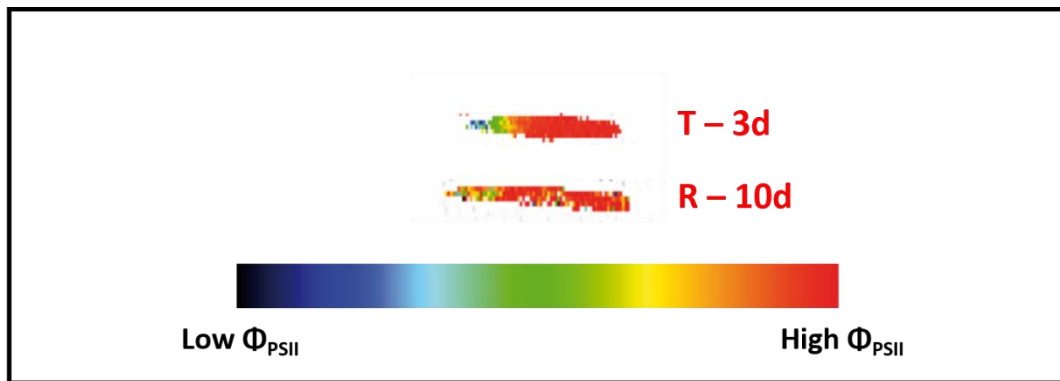


*Photosynthetic efficiency of recovered plants.*

After 3 or 6 days, some of the plants belonging to both varieties were moved in a new fresh medium without salt. PSII maximum quantum efficiency ( $\Phi_{PSII}$ ) was thus monitored after 7 days to evaluate the eventual photosynthesis recovery (Figure 6).



**Figure 6. PSII maximum quantum yield ( $\Phi_{PSII}$ ) during the recovery experiment.** PSII maximum quantum yield ( $\Phi_{PSII}$ ) for rice plants *v. Baldo* (A and B) and *v. Vialone Nano* (VN) (C), during salt treatment prolonged for 3 (A and C) or 6 (B) days (d) and after putting them in a recovery medium (R) without salt, for 7 days. Data for the recovery of VN plants PSII functionality after 6 days of salt treatment are not present since plants, at this time point, were dead. C- plants not exposed to salt treatment; T - plants exposed to salt treatment; R – plants in the recovery medium. black – first leaf, red – second leaf, blu – third leaf, green – fourth leaf, violet – fifth leaf. Data are expressed as average of 6 biological replicates  $\pm$  SD.



**Figure 7. PSII maximum quantum yield ( $\Phi_{PSII}$ ) image for the first leaf of *v. Baldo* plants. B.** Fluorescence images depicting PSII maximum quantum yield values for the whole first leaf area. Images were taken after 3 days (d) of salt treatment and after 7 days of recovery (R) in a control medium, without salt. The region on the left represents the apical part of the leaf.

Even though *Baldo* plants are able to withstand the stress, the latter, also after 3 days, still affects the plants, that are not able to synthesize a fifth leaf after 7 days of recovery, unlike the control (Figure 6A). Taking a look to the fluorescence images, the first leaf in fact shows a recovered PSII functionality (Figure 7), highlighting a PSII maximum quantum yield increase also in the leaf areas originally compromised. 6 days of salt treatment show instead their effect on the PSII maximum quantum yield of the first leaf, which, also after 7 days of recovery, is photosynthetically compromised, much more than for control plants (Figure 6B). However this time the synthesis of a new leaf (fifth leaf) is not affected. For *VN* plants, after 3 days of salt treatment, the survival of just the third leaf (Figure 6C) is sufficient to recover the PSII functionality of the first and the second leaves (Figure 6C) and also the ability to synthesize a new one (the fourth). These data highlights again a different mechanism of salt response for the two varieties, pointing out the capacity for *Baldo* plants to limit the stress response to the first leaf, which can be sacrificed to support photosynthesis activity of the third and fourth ones. While *VN* plants, despite more salt sensible, are still able to recover from the stress if not prolonged over time.

Results presented here are consistent with literature data reporting the induction of a stronger NPQ activation in plants exposed to abiotic stresses (e.g. salt and drought) in several species. As example, plants showing salt tolerance (e.g. *Brassica campestris*) indeed show a faster NPQ activation kinetic which also leads to higher NPQ amplitude after light exposition (Zhu et al., 2011). This event was also

described in *P. patens*, a representative of the early land colonization stage (Azzabi et al., 2012), highlighting how the NPQ activation evolved in plants just after terrestrial habitats colonization.

### **Conclusion.**

These results suggest that increased NPQ activation could be used in combination with  $\Phi_{PSII}$  as a biomarker of the response to salt stress, evidencing that plants are putting in place the appropriate responses. Application of PAM instruments could therefore provide a first evaluation of salt exposure directly on the field, highlighting the severity of the stress.

## References.

- Azzabi, G., Pinnola, A., Betterle, N., Bassi, R., and Alboresi, A.** (2012). Enhancement of non-photochemical quenching in the Bryophyte *Physcomitrella patens* during acclimation to salt and osmotic stress. *Plant Cell Physiol.* **53**: 1815–25.
- Ballottari, M., Dall’Osto, L., Morosinotto, T., and Bassi, R.** (2007). Contrasting behavior of higher plant photosystem I and II antenna systems during acclimation. *J. Biol. Chem.* **282**: 8947–8958.
- de Bianchi, S., Ballottari, M., Dall’osto, L., and Bassi, R.** (2010). Regulation of plant light harvesting by thermal dissipation of excess energy. *Biochem. Soc. Trans.* **38**: 651–60.
- Du, H., Wang, N., Cui, F., Li, X., Xiao, J., and Xiong, L.** (2010). Characterization of the beta-carotene hydroxylase gene *DSM2* conferring drought and oxidative stress resistance by increasing xanthophylls and abscisic acid synthesis in rice. *Plant Physiol.* **154**: 1304–18.
- Erickson, E., Wakao, S., and Niyogi, K.K.** (2015). Light stress and photoprotection in *Chlamydomonas reinhardtii*. *Plant J.* **82**: n/a–n/a.
- Gerotto, C., Alboresi, A., Giacometti, G.M., Bassi, R., and Morosinotto, T.** (2011). Role of PSBS and LHCSR in *Physcomitrella patens* acclimation to high light and low temperature. *Plant. Cell Environ.* **34**: 922–32.
- Maxwell, K. and Johnson, G.N.** (2000). Chlorophyll fluorescence - A practical guide. *J. Exp. Bot.* **51**: 659–668.
- Zhu, S.-Q., Chen, M.-W., Ji, B.-H., Jiao, D.-M., and Liang, J.-S.** (2011). Roles of xanthophylls and exogenous ABA in protection against NaCl-induced photodamage in rice (*Oryza sativa* L) and cabbage (*Brassica campestris*). *J. Exp. Bot.* **62**: 4617–25.
- Zulfugarov, I.S. et al.** (2014). Production of superoxide from Photosystem II in a rice (*Oryza sativa* L.) mutant lacking PsbS. *BMC Plant Biol.* **14**: 242.



# APPENDIX III

## *In-Situ* Solid state NMR spectroscopy on intact photosynthetic thylakoid membranes

### Author names and affiliations

Fatemeh Azadi<sup>1</sup>, Karthick Babu Sai<sup>1</sup>, Sankar Gupta<sup>1</sup>, Diana Simionato<sup>2</sup>, Giorgio Perin<sup>2</sup>, Tomas Morosinotto<sup>2</sup> and Anjali Pandit<sup>1</sup>

<sup>1</sup>Gorlaeus Laboratories, Faculty of Science, Leiden Institute of Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

<sup>2</sup>PAR-Lab\_Padua Algae Research Laboratory, Department of Biology, University of Padova, Via U. Bassi 58/B, 35121 Padova, Italy

The following experimental work was performed during an established collaboration between our group and Prof. Anjali Pandit group (from the Leiden Institute of Chemistry of Leiden University). These results were included in a conference paper (Azadi F. et al., 2015) published during the Euromar conference, that took place in Prague, Czech Republic, in June 2015.

**Abstract.**

In oxygenic photosynthesis, sophisticated regulation mechanisms have evolved to enable the splitting of water via P680, the strongest oxidizer found in nature, while preventing the system from photo-damage. Under excess light, remodeling of the photosynthetic membrane takes place and molecular switching of light-harvesting antenna proteins into a photo-protective, light-quenching state. Photosynthetic thylakoid membranes are densely packed with proteins (~70%) and abundant with photosystems (PSII and PSI) that contain light-harvesting pigment-protein complexes. Here we explore the structure and dynamics of protein and lipid components *in situ* by applying high-resolution Magic-Angle Spinning NMR on uniformly  $^{13}\text{C}$  -  $^{15}\text{N}$  labeled thylakoid membranes of the green alga *Chlamydomonas reinhardtii* in active state (WT) and in photo-protective state (*npq2* mutant). Data are compared with results on isolated *C. reinhardtii* light-harvesting complexes.

## Materials and Methods.

### *Chlamydomonas reinhardtii strains and growth conditions.*

*C. reinhardtii* strains *cw15* and *npq2*, kindly provided by Prof. Roberto Bassi from University of Verona, were employed in this work. The first is a cell wall-less mutant (Clarke et al., 1982) used as WT, while the second is affected in zeaxanthin epoxidase (ZE) activity (Niyogi et al., 1997). Both were cultivated in liquid tris-acetate phosphate (TAP) medium, in Erlenmeyer flasks, at 100 rpm agitation and 21 °C in a growth chamber. Continuous illumination was provided from cool-white fluorescent lamps under low ( $< 25 \mu\text{moles photons m}^{-2} \text{s}^{-1}$ ) photosynthetically active radiation (400-700 nm). The TAP medium (Hooper, 1989) used to grow labeled cells, was prepared using  $^{13}\text{C}$  label sodium acetate (Sigma-Aldrich) and  $^{15}\text{N}$  label ammonium chloride (Sigma-Aldrich). Cultures in labeled medium were set up starting from  $\text{OD}_{750} = 0.1$  and cells were grown until  $\text{OD}_{750} = 1$ . Three steps of culture were performed sequentially, always starting from the former value and reaching the latter, in order to be sure to label  $> 99.9\%$  of all the harvested cells with  $^{13}\text{C}$  and  $^{15}\text{N}$  atoms.

### *Thylakoid isolation.*

Cells were harvested by 10 minutes of centrifugation at 4 °C, at 3500 x g and then washed twice in isolation medium A (IMA, 10mM MES pH 6.5, 2mM KCl, 5mM EDTA pH 8, 1M sorbitol).

After centrifugation, cells were resuspended in cold IMA buffer with 0.5% milk powder and 1 mM PMSF, 1 mM DNP- $\epsilon$ -amino-n-caproic acid and 1 mM benzamidine, and then disrupted at 4 °C using an ultrasonic homogenizer (Sonic Rupter 400 – OMNI International - PBI) for 5 s, with the maximum power.

Immediately after rupture, the samples were centrifuged for 15 minutes at 2500g at 4 °C to collect unbroken cells on the bottom of the tube. The latter were again resuspended in IMA buffer containing inhibitors and milk powder and treated again with the sonicator. This step was repeated 3 times to be sure to break all the harvested cells.

At every step, the collected supernatant contained the thylakoids. The latter was centrifuged for 15 minutes at 2500 g at 4 °C, to eliminate cells debris. The supernatant was then centrifuged for 30 minutes at 40.000 g at 4 °C to collect the thylakoids. The pellet, containing the thylakoids, was washed twice with isolation medium B (IMB, 10mM MES pH 6.5, 2mM KCl, 5mM EDTA pH 8) and later resuspended in T3 buffer (50 mM Hepes-KOH pH 7.5, 5 mM MgCl<sub>2</sub>, 50% glycerol). Immediately it was frozen in liquid nitrogen and stored at -80 °C until use. All steps were performed at 4 °C and in dim light. Thylakoids total pigments were extracted with 80 % acetone,



and the chlorophyll concentration of the samples was determined spectrophotometrically using specific extinction coefficients (Porra et al., 1989) and the acetone spectra fitting previously described in (Croce et al., 2002).

#### *Thylakoid solubilization and sucrose gradients.*

Thylakoid membranes corresponding to 3 mg of total Chl were washed with 50 mM EDTA and then solubilized for 20 minutes on ice in 3 ml of final 1.2 %  $\alpha$ -DM and 10 mM HEPES (pH 7.5), after vortexing for 1 min. The solubilized samples were centrifuged at 15000 x g for 20 min to eliminate any unsolubilized material, and the supernatant with the photosynthetic complexes was then fractionated by ultracentrifugation in a 0.1 – 1 M sucrose gradient containing 0.06 %  $\alpha$ -DM and 10 mM HEPES (pH 7.5) (141000 x g, 40 hours, 4°C). The green fractions of the sucrose gradient were then harvested with a syringe.

#### *Spectroscopic analysis.*

Absorption spectra of single sucrose bands were determined between 350 and 750 nm using a Cary Series 100 UV-VIS spectrophotometer (Agilent Technologies).

#### *Pigment analysis.*

The content of individual carotenoids was determined using HPLC (Beckman System Gold), as described (Färber and Jahns, 1998). The peaks of each sample were identified through the retention time and absorption spectrum (Jeffrey et al., 1997).

## Results and discussion.

Our work focused on the cultivation of *C. reinhardtii*, in a  $^{13}\text{C}$  and  $^{15}\text{N}$  label TAP growth media, to isolate the trimeric light-harvesting complex (LHC) proteins from its thylakoid membranes. The latter were then analyzed with the solid-state nuclear magnetic resonance (NMR) technique by the people in Prof. Anjali Pandit group, to explore the structure and dynamics of these proteins together with membrane lipid components, when photosynthetic membranes are in active state (WT *cw15*) or in photo-protective state (*npq2* mutant).

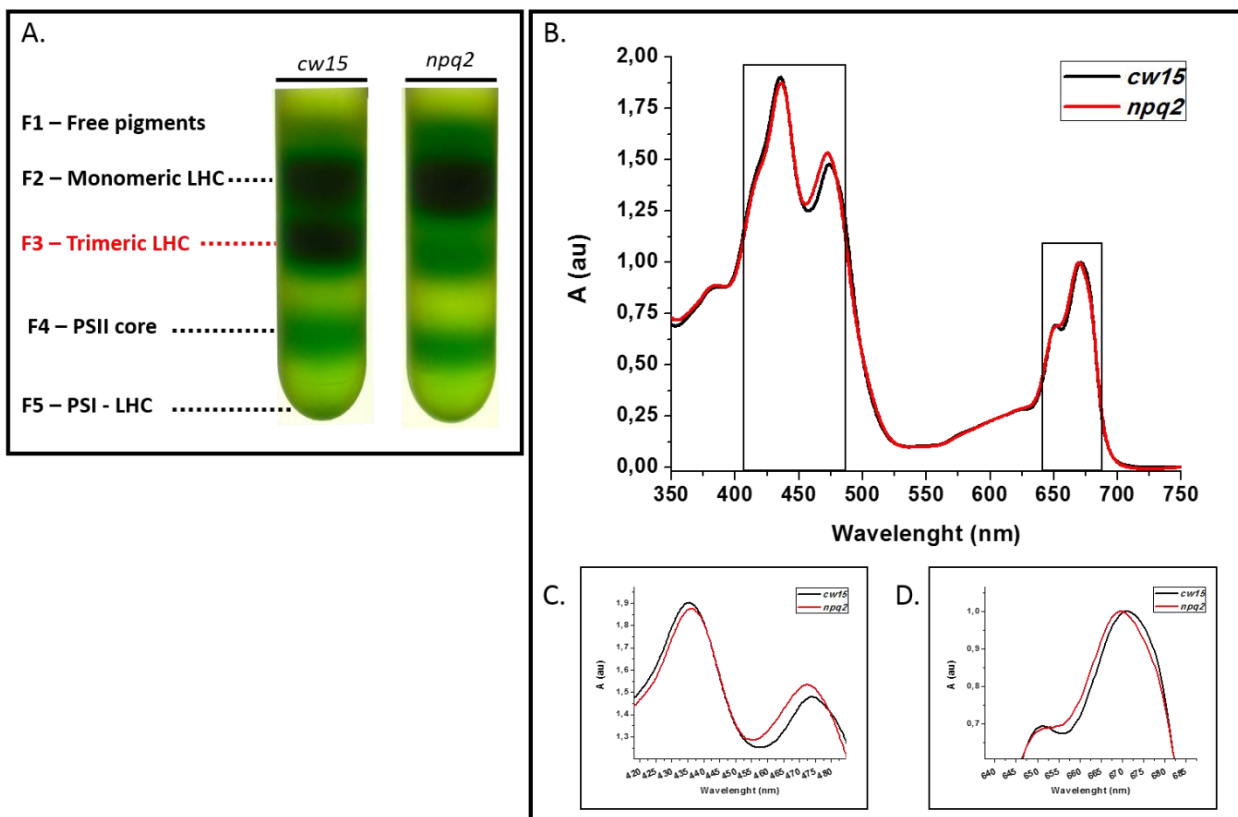
### *C. reinhardtii* cells labeling and thylakoid extraction.

*Chlamydomonas reinhardtii* strains *cw15* and *npq2* were chosen for this work and cultivated according to the following considerations. To let the samples be detectable through the solid-state NMR technique, the cells needed to be labeled with carbon ( $^{13}\text{C}$ ) and nitrogen ( $^{15}\text{N}$ ) isotopes. Since *Chlamydomonas reinhardtii* is a photosynthetic organism, it fixes  $\text{CO}_2$  from the atmosphere to support its photoautotrophic metabolism. However it is also able to grow with a heterotrophic metabolism, in presence of acetate as carbon source (Dent et al., 2015). Therefore, to be sure that the majority of the harvested cells were labeled both in carbon and nitrogen atoms, we decided to exploit the latter metabolic feature, providing the carbon source in form of labeled sodium acetate. Since we wanted to isolate the photosynthetic membranes in an active state, the photoautotrophic metabolism couldn't be completely shut down, therefore we decided to provide the cultures a low light intensity, to be sure to leave the latter active, but not too much to prevent the acetate assimilation from the medium (Heifetz et al., 2000). Moreover, we performed three-sequential-steps of cultivation in these conditions, providing a 1 / 10 dilution at every step. This assured the labeling of at least the 99.9 % of the harvested cells, at the end of the third step of cultivation.

Since the NMR-based techniques generally require a huge amount of material to obtain a detectable signal (between 3 and 5 mg of protein sample), we needed to harvest a lot of cells to isolate the sufficient amount of labeled trimeric LHC to be use form the Prof. Anjali Pandit group. Therefore, this imposed to process a lot of cells at a time for the thylakoids isolation in a time-saving perspective.

We developed a thylakoids isolation protocol, at a higher scale than that generally used in common laboratory procedures (see methods for details). Pigment-binding complexes from the photosynthetic apparatus of both strains were then isolated with sucrose gradients (Figure 1A), in order to separate trimeric LHC to be further analyzed by solid-state NMR.

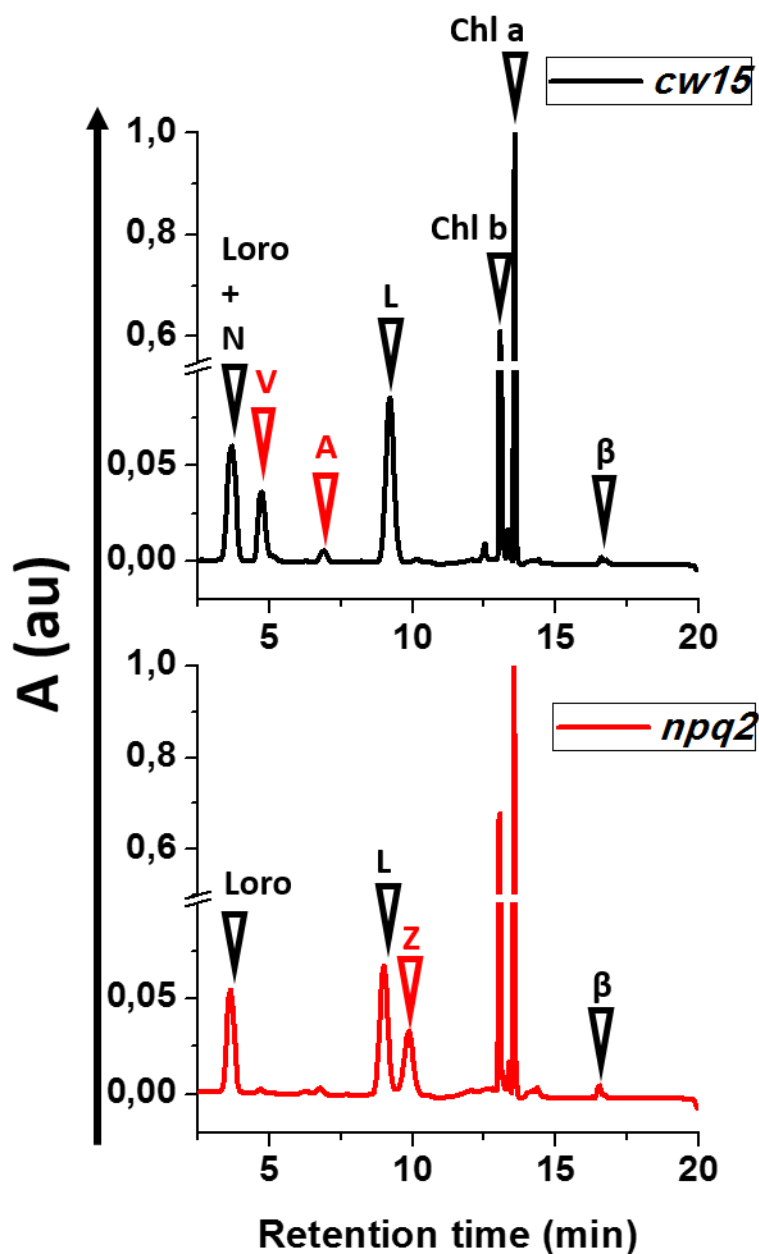
As shown in figure 1A the *C. reinhardtii* *npq2* mutant showed a reduced trimeric LHC content with respect to the *cw15* strain.



**Figure 1. Trimeric light-harvesting complex (LHC) proteins isolation from the *cw15* and the *npq2* strains of *C. reinhardtii*.** A. Isolation of pigment-binding proteins from *C. reinhardtii* labeled thylakoids by sucrose gradient centrifugation. Four distinct bands are visible (F1-F4), while the fifth is present in the bottom of the tube (F5). Bands identification is reported, according to spectroscopic analysis. B. Absorption spectra of trimeric LHC isolated from the sucrose gradients in A. C and D show a magnification of the carotenoids absorption region (C) and of the chlorophyll absorption region (D).

Despite leaving unaltered the nature of the LHC proteins found in its photosynthetic apparatus, the reduction in number of trimeric LHC associated to every PSII was already shown in literature (Polle et al., 2001) for the *npq2* mutant. Its inability to synthesize antheraxanthin, violaxanthin and neoxanthin from zeaxanthin and the accumulation of the latter, also in low light conditions, induces a change in LHC pigment composition (Figure 1B), which switches the carotenoids absorption peak toward the blue region of the electromagnetic spectrum (Figure 1C) as well as the chlorophyll a and b absorption peaks (Figure 1D). The peculiar individual carotenoids composition of the trimeric LHC of the *npq2* mutant was validated in figure 2.

The isolated trimeric LHC proteins were therefore validated for the further analyses of Prof. Anjali Pandit group. The latter will highlight the structural changes that usually undergo in a photosynthetic apparatus when the environmental conditions trigger a quenching state of LHC proteins to meet the photoprotection needs.



**Figure 2. Individual carotenoids composition of trimeric LHC of the *cw15* and the *npq2* strains of *C. reinhardtii*.** Traces were normalized on the peak of Chl a. Identification of lettered peaks is as follows: Loro, Loroanthin; N, neoxanthin; V, violaxanthin; A, antheraxanthin; L, lutein; Z, zeaxanthin;  $\beta$ ,  $\beta$ -carotene.

## References.

- Clarke, A., Coulson, G., and Morris, G.J.** (1982). Relationship between Phospholipid Breakdown and Freezing Injury in a Cell Wall-Less Mutant of *Chlamydomonas reinhardtii*. *Plant Physiol.* **70**: 97–103.
- Croce, R., Canino, G., Ros, F., and Bassi, R.** (2002). Chromophore Organization in the Higher-Plant Photosystem II Antenna Protein CP26. *Biochemistry* **41**: 7334–7343.
- Dent, R.M., Sharifi, M.N., and Malnoe, A.** (2015). Large-scale insertional mutagenesis of *Chlamydomonas* supports phylogenomic functional prediction of photosynthetic genes and analysis of classical acetate-requiring mutants. **82**: 337–351.
- Färber, A. and Jahns, P.** (1998). The xanthophyll cycle of higher plants: influence of antenna size and membrane organization. *Biochim. Biophys. Acta - Bioenerg.* **1363**: 47–58.
- Heifetz, P.B., Förster, B., Osmond, C.B., Giles, L.J., and Boynton, J.E.** (2000). Effects of acetate on facultative autotrophy in *Chlamydomonas reinhardtii* assessed by photosynthetic measurements and stable isotope analyses. *Plant Physiol.* **122**: 1439–45.
- Hooper, J.K.** (1989). *The Chlamydomonas Sourcebook. A Comprehensive Guide to Biology and Laboratory Use.* Elizabeth H. Harris. Academic Press, San Diego, CA, 1989. xiv, 780 pp., illus. \$145. *Science* **246**: 1503–4.
- Jeffrey, S.W., Mantoura, R.F.C., and Wright, S.W.** (1997). *Phytoplankton pigments in oceanography: guidelines to modern methods.* Monogr. Oceanogr. Methodol.
- Niyogi, K.K., Bjorkman, O., and Grossman, A.R.** (1997). *Chlamydomonas Xanthophyll Cycle Mutants Identified by Video Imaging of Chlorophyll Fluorescence Quenching.* *Plant Cell* **9**: 1369–1380.
- Polle, J.E.W., Niyogi, K.K., and Melis, A.** (2001). Absence of lutein, violaxanthin and neoxanthin affects the functional chlorophyll antenna size of photosystem-II but not that of photosystem-I in the green alga *Chlamydomonas reinhardtii*. *Plant Cell Physiol.* **42**: 482–91.
- Porra, R.J., Thompson, W.A., and Kriedemann, P.E.** (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta - Bioenerg.* **975**: 384–394.

## Abbreviations

ACCase, acetyl-CoA carboxylase

AGP, arabinogalactan protein

Alg-Feed, economic scenario in which algae biomass will sustain part of the global economy

AS, functional antenna size

BAU, business as usual

bHLH, basic helix-loop-helix

C, control

Car, carotenoid

CCAP, Culture Collection of Algae and Protozoa

CCS, Carbon Capture and Storage technologies

Chl, chlorophyll

<sup>1</sup>Chl\*, singlet chlorophyll

<sup>3</sup>Chl\*, triplet chlorophyll

CpSRP, chloroplast signal recognition particle

*Crluc*, *Chlamydomonas reinhardtii* luciferase gene

CRM, Camacho Rubio model

DAG, diacylglycerol

DBMIB, dibromothymoquinone

DCMU, 3-(3,4-dichlorophenyl)-1,1-dimethylurea

DGAT, Acyl-CoA : diacylglycerol acyltransferase

DGDG, digalactosyldiacylglycerol

DGTS, diacylglyceryltrimethylhomoserine

DHA, docosahexaenoic acid

DHAP, dihydroxyacetone phosphate

DMF, N,N'-dimethylformamide

EMS, ethyl methane sulfonate

EPA, eicosapentaenoic acid  
EPM, Eilers and Peeters model  
ER, endoplasmic reticulum  
ESI, positive electrospray ionization  
ETR, electron transport rate  
F, fluorescence  
FAS, fatty acids synthase  
FMDV 2A, foot-and-mouth disease-virus 2A  
GAP, glyceraldehyde-3-phosphate  
GMO, genetically modified organism  
GO, gene ontology  
HL, high light  
HPLC, high-pressure (performance) liquid chromatography  
ID, identifier  
IPCC, Intergovernmental Panel on Climate Change  
IPP, isopentenyl-pyrophosphate  
LC-PUFA, long chain poly-unsaturated fatty acids  
LDSP, lipid droplet surface protein  
LHC, light harvesting complex  
LHCX / LHCSR, stress-related light-harvesting complex  
LIL, LHC-like  
LL, low light  
ME, malic enzyme  
MGDG, monogalactosyldiacylglycerol  
ML, medium light  
N, nitrogen  
NMR, nuclear magnetic resonance  
NPQ, non-photochemical quenching  
 $^1\text{O}_2$ , singlet oxygen

OCP, orange carotenoids binding protein  
OD, optical density  
PAM, pulse-amplitude-modulation  
PAR, photosynthetic active radiation  
PBR, photobioreactor  
PC, phosphatidylcholine  
PDC, pyruvate dehydrogenase complex  
PDK, pyruvate dehydrogenase kinase  
PDMS, poly(dimethylsiloxane)  
PDS, phytoene desaturase  
PE, phosphatidylethanolamine  
PG, phosphatidylglycerol  
PS, photosystem  
PSI, photosystem I  
PSII, photosystem II  
PSBS, photosystem II subunit S  
PSU, photosynthetic unit  
PSY, phytoene synthase  
PUFA, poly-unsaturated fatty acids  
qE, fast relaxed kinetic of the NPQ (energy dependent NPQ component)  
RLU, relative luminescence unit  
*Rluc*, *Renilla reniformis* luciferase gene  
RNAi, RNA interference  
ROS, reactive oxygen species  
RT, room temperature  
RT-PCR, reverse transcriptase polymerase chain reaction  
RuBisCO, ribulose 1,5biphosphate carbosilase-oxygenase  
RUBOX, Ring-U-BOX domain containing protein  
SDS, sodium-dodecyl-sulphate



SQDG, sulfoquinovosyldiacylglycerol  
T, treatment  
TAG, triacylglycerol  
TAP, tris-acetate phosphate medium  
TE, thioesterase  
TEF, total electron flow  
TF, transcription factors  
TGF- $\beta$ , transforming growth factor- $\beta$   
TKL, Transketolase  
TLA, truncated light harvesting  
TPT, triose phosphate/phosphate translocator  
UBI, ubiquitin  
UTR, untranslated region  
VDE, violaxanthin de-epoxidase  
*VN, Oryza sativa variety Vialone Nano*  
WT, wild type  
ZE, zeaxanthin epoxidase  
ZEO, zeocin  
 $\alpha$ -DM, n-Dodecyl  $\alpha$ -D-maltoside  
 $\Phi_{\text{PSII}}$ , PSII maximum quantum yield